Clusters of gene duplicates:
Identification, evolutionary consequences, and phylogenetic patterns

By

Juan Felipe Ortiz Quiñonez

Dissertation
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY
in
Biological Sciences

May 10, 2019

Nashville, Tennessee

Approved:
Professor Antonis Rokas, Ph.D.
Professor Patrick Abbot, Ph.D.
Professor Tony Capra, Ph.D.
Professor Maulik Patel, Ph.D.
Professor Matthew Hahn, Ph.D.

DEDICATION

To all living things, big and small. To all time things, young and old. To evolution, that beautiful process that transforms chaos into beautiful chaos.

ACKNOWLEDGMENTS

It was Brigitte Gavio, PhD, my undergraduate thesis advisor, who was able to crack open my head and stuff it with the most extreme scientific rigor, at the same time as she made my heart receptive to love biology intensely.

She taught me that I was going to be allowed publish my first paper only after I have learned techniques, methods, and phylosophical arguments to defend my experiments until I ended mentally exhausted, and psychologically inadequate to engage in any meaningful human interaction.

More importantly, she told me that any amount of intellectual training was pointless if I did not love the experiment I was trying to work on. She taught me that pretending to know, or using fancy terms was just embarrasing, if I did not truly believed that my research topic was the most absolutely beautiful piece of reality I could ever have the honor to interact with.

She told me that in order to do science, I have to be an expert in my research topic, but more importantly, that if I wanted to be a good scientist, I have to love my research topic.

I also thank my laboratory, my advisor and the funding sources that made my research possible.

Finally, I want to thank the members of my cohort, who helped me with emotional support many times. Especially Michelle and Stephanie were very supportive, and were islands of sanity and kindness, in the sea of madness that my head was in the course of my studies here.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Traits of great importance for the biology of organisms and their evolution are driven by clusters of tandemly duplicated genes (CTDG). CTDGs like the Hox, protocadherin and β-globin genes are fundamental in the development of animals [1–3], and genes in CTDGs like the major histocompatibility complex (MHC) and the killer-cell immunoglobulin-like receptor (KIR) are important in immunity [4, 5] and even speciation [6]. Moreover, expansions and contractions of CTDGs have been associated with lineage-specific traits such as venom in snakes and platypus [7, 8], and increased olfaction perception in rodents and elephants [9]. Despite their biological importance, CTDGs are often studied in a case by case basis, limiting the scope of comparative and evolutionary analyses in the topic.

I will start with discussing the molecular processes that generate CTDGs. Then, I will cover fundamental aspects of animal biology where CTDGs are involved, and the ways in which CTDGs can generate molecular diversity, from molecular phenotypes, to potentially speciation-inducing variability. I will also review the ways in which CTDG have been studied in the past, some challenges we have when studying CTDGs. By the end of this dissertation I will propose an “evolution research-friendly” definition of CTDG, as well as strategies about how to study CTDGs in evolutionary context.

1.1 Gene duplication and the evolution of gene families

Back in the early 1900’s, it was observed that variations in the number of loci via partial or complete chromosome duplications, had the potential to increase phenotypic diversity (reviewed in [10]).It was well known allelic mutations alone could not account for all the variation seen in vertebrates, and the role of gene duplication in the evolution of vertebrates was proposed back in the 1960’s [11]. With the available tools of that time, various genes were inferred to be duplicated via whole genome duplication or other mechanisms, by comparing the C-value (or the number of chromosomes) of two related taxa to their gene copy number. For example, from the similar C-values estimated for placental mammals and the presence two copies of the enzyme lactate dehydrogenase, it was inferred that whole genome duplication events did not occur after the diversification of mammals [11]. On the other hand, work done on plants of the genus Anemone showed that although some species had the same number of chromosomes, their length tend to be a multiple of a reference, suggesting that segmental duplication, but not whole genome duplication took place in those species [11, 12].

The model of gene family evolution by gene duplication proposed by Ohno tried to explain how gene duplicates were retained via accumulation of beneficial mutations. However, beneficial mutations are quite rare, and there are significantly more duplicated genes retained in genomes than what it would be predicted by Ohno's model [13]. Study of this observation led to the proposal of the Duplication, Degeneration, Complementation model (DDC) [14], and other models (reviewed by Demuth and Hahn [15]). Since the main mechanisms by which gene families evolve involve genome or chromosome level changes, like tandem gene duplication, translocations, or even whole genome duplications [16], knowing how genome evolution is studied is important to further our understanding of how gene families in general, and clusters of tandemly duplicated genes (CTDG) in particular, evolve.

1.2 Genome dynamics and the evolution of gene families

Genomes evolve by insertion, deletion, replacement, and rearrangement of DNA regions [17]. However, evolutionary frameworks able to model how structural variation occur in genomes (e.g. genomic translocations, inversions, segmental duplications) remain limited in scope and resolution, mainly because several molecular mechanisms can result in a specific duplication outcome. For example, an tandem inverted duplication can be obtained in one step if a DNA break is fixed using a homologous region on the same chromatid upstream of the damaged site. Conversely, a similar inverted duplication can be obtained in at least two steps, one duplication via non-homologous recombination, and one inversion (fig. 1.1 [16]). Importantly, given that several molecular mechanisms can generate very similar outcomes, calculating the the rates at which a genome structure evolves (fig. 1.1; e.g. duplication inversion) remains a challenge, which makes the statistical modeling of such processes difficult, hampering the design of the above mentioned evolutionary framework to study genome structure evolution.

Gene families evolve in this dynamic context of genomic rearrangements. A gene family should include all the genes that come from the same ancestor. However, from a practical perspective, gene families are often defined by clustering genes by sequence similarity (reviewed by Demuth and Hahn [15]). In conjunction with gene sequence divergence, gene families evolve by increasing or decreasing the

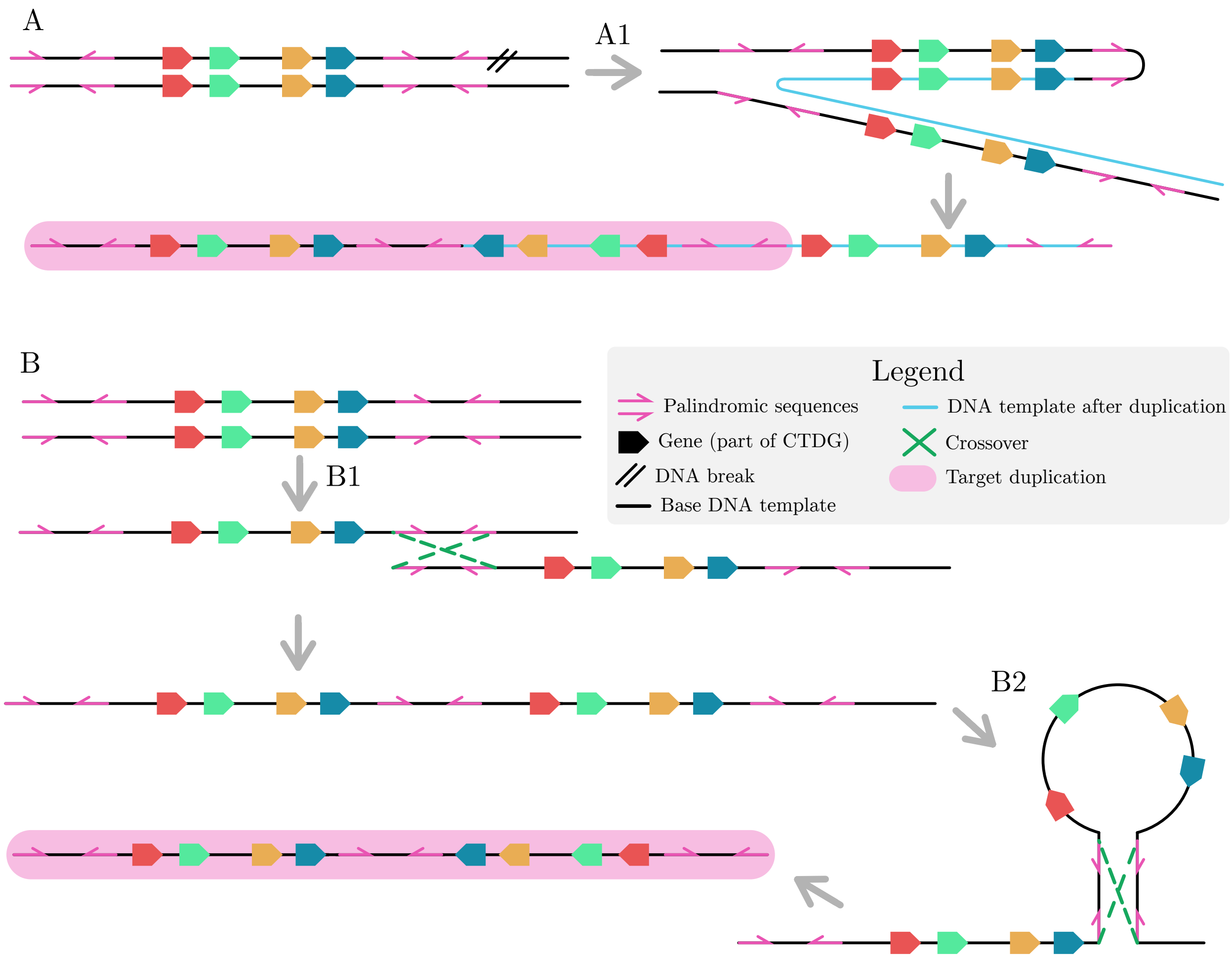

Figure 1.1: CTDG duplication inversion can occur by mechanisms of varying complexity. A. CTDG inversion duplication by DNA break repair and template switching. A.1: After a DNA break, palindromic sequences can force a the damaged DNA to use itself as template to repair the damage, and palindromic sequences at the other end of the damage can provoke template switching, causing the repair mechanisms to continue in the homologous strand (modified from Reams and Roth, 2015). B: CTDG duplication inversion by non-homologous recombination (B1) and inversion (B2) (modified from [18]). Note that B1 and B2 are independent events.

number of genes in them [15].

Interestingly, the processes that give rise to these gene family expansions and contractions influence the probability of gene duplicate retention. It has been well documented that after whole genome duplications (WGD), genes coding for proteins that are strongly regulated in their relative concentration to other proteins will tend to be retained to maintain the stoichiometry of the processes they are involved in. This is known as the gene balance hypothesis [19]. On the other hand, especially in eukaryotes, the genes families that tend to show significant expansions (i.e. increase in gene copies) are often involved with increasing pathogenicity, resistance to toxic substances, or sensing environmental signals in response to stress [15].

1.3 Studying CTDG evolution

1.3.1 Genome-level mechanisms involved in CTDG remodeling

Substantial work has been done documenting cases of gene duplication and its consequences. However, less work has been done deciphering the molecular mechanisms driving tandem gene duplication, which usually give rise to CTDGs, mainly because of practical and theoretical challenges. One of the challenges of modeling CTDG evolution is gene conversion. Since gene conversion removes sequence variation of gene duplicates in a CTDG, using classic phylogenetic methods becomes risky because the sequence variation necessary to perform phylogenetic analyses deteriorates after gene conversion, which in turn can obscure inferences about the timing of duplication events, because instead of observing the diversification history of a CTDG along the CTDG phylogeny, we would see the history of gene conversion events (at best), or a very poorly resolved tree (at worst). However, gene conversion has been hypothesized to be an important biological mechanism for maintaining the integrity of the function of CTDGs. For example, the protocadherin CTDG generates high diversity of receptor types by splicing several variable exons (exons that contribute the sequence diversity of the spliced gene) expressed independently with constant exons (exons that are always expressed in the spliced gene). Interestingly, evidence exists supporting that although gene conversion is not common among the variable exons, it is a common occurrence between the constant exons, which ensures sequence and function stability of the spliced gene [20]. As CTDGs are genomic features (regions of the genome with specific characteristics), they are influenced by many evolutionary processes, both at the scale of the genes they contain, as well as the different molecular mechanisms that generate them. For example, CTDGs can be formed (or amplified) by genomic exchanges between sister chromatids as a way of resolving DNA lesions, and they can produce duplications in the same strand or inversions, depending on the chromatid that was used as template to repair the lesion (fig. 1.1 [16]). Moreover, CTDGs can be amplified one, a few, or several genes at the time, depending on the size of the duplicated genomic region [21]. In adition to genome-level mechanisms that modify, add, or remove gene duplicates, CTDGs can also be remodeled over evolutionary time by inversions, translocations, and other genomic-level rearrangements, shadowing the biological processes that generated them [22].

1.3.2 CTDGs are structural features of the genome

When performing phylogenetic studies on genes, population-genetics and evolutionary theory provide a robust background from which evolutionary hypotheses can be tested. Since point mutations are, as the name implies, restricted to one nucleotide, the spatial relationships between a focal nucleotide and others is not part of the intrinsic characteristics of the mutation. However, since CTDGs are clusters, (i.e. a spatially determined gene organization), when studying a CTDG, not only the homology relationships between their genes is required to define said CTDG, but also the spatial relationships between the genes in a CTDG and its genomic neighborhood are intrinsic characteristics of the CTDG [23] (fig. 1.2).

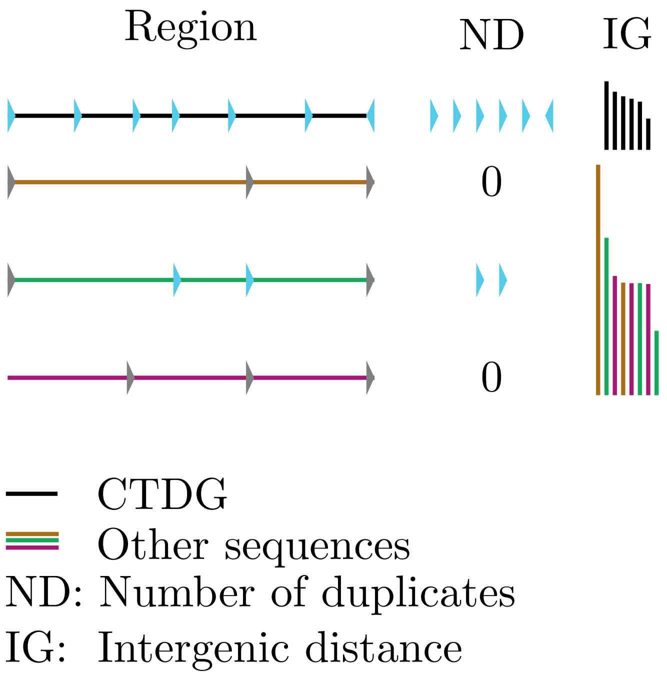

Figure 1.2: Representation of the defining characteristics of a CTDG in four hypothetical genomic regions of equal length. The number of gene duplicates in a CTDG (ND) is not enough to describe a CTDG. Specifically, the intergenic distances (IG) inside (black horizontal lines) and outside (colored horizontal lines) CTDGS can vary substantially, depend on the length and number of duplicates of a CTDG, and should be considered in the context of the genomic background specific to each species.

Looking at the complexity introduced in defining a CTDG, it is worth asking if we need to define what a CTDG is. Are CTDGs really biological entities, spatially and evolutionarily constrained? In other words, are the parameters and characteristics used to define a CTDG (fig. 1.2) of such evolutionary relevance to require delineating yet another biological definition?

1.4 The importance of CTDGs in biological systems and their influence in phenotypic variation

1.4.1 Conserved CTDGs driving polymorphic traits

Arguably some of the most well-known CTDGs are the four Hox CTDGs. Hox genes are transcription factors containing a homeodomain region, and drive the anterior-posterior patterning of body plans in metazoans [24]. The importance of the four sets of Hox genes as a CTDGs relies on the fact that they are expressed along an animal axis during development in the order they exist in the CTDGs (i.e. Hox genes at the 5' end are mostly expressed in the posterior section of the embryo, and Hox genes at the 3' end are mostly expressed in the anterior section of the embryo). The importance of the Hox CTDGs as such is accentuated by the identification of the Hox CTDGs in almost all bilaterian species with their genes mostly in the same order. The conservation of gene order across almost 1,000 million years supports the idea that Hox CTDGs work as a biological entity that must be conserved to achieve their function of body plan patterning [24]. For example, the expression pattern and timing of some 3' Hox genes are important in the development of the little skate (*Leucoraja erinacea*) pectoral fin, but not in the development of the anterior fins of the related bamboo shark [25].

Many other gene families with several duplication events have been shown to have a crucial role in metazoan early development. Zinc finger-associated domain genes (ZAD), which are expressed in the female germline in *D. melanogaster*; NHR genes involved in molting and metamorphosis in *C. elegans*; and T-box genes, expanded in *C. elegans* and involved in cell-fate determination; are some interesting examples [26]. Notably, although many genes from the examples presented above are located in close proximity to each other and are recognized as lineage-specific expansions, they were not explicitly characterized as part of CTDGs.

It is important to note that although the general organization of each paralogous Hox gene in each of the Hox CTDGs is maintained, other defining parameters of CTDGs are not. For example, it has been observed that only one Hox CTDG exists in the ascidian *Ciona intestinalis*, which lacks some Hox genes conserved in other species. Moreover, the Hox CTDG in *C. intestinalis* is not as compact as in other species, even suggesting that it is not one but several CTDGs [27]. Similarly, a significant amount of transposable elements have been observed in the Hox CTDGs of the anole lizard (*Anolis carolinensis*), increasing the length of such CTDGs [28], and non-conserved gene order and length of the only Hox CTDG in the sea urchin (*Strongylocentrotus purpuratus*) has been studied [29]. It is tempting to speculate that differences in the genomic organization of the conserved Hox CTDG can explain part of the body plan architecture in some animals, especially in non-vertebrate chordates, like *C. intestinalis* or *S. purpuratus*. However, we have to keep in mind that although conserved, some level of variation in the Hox CTDGs exists, and those "deviations" from the norm, are part of the spectrum of natural variation inherent to the Hox CTDGs [30].

The β-globin CTDG has been extensively characterized, and its conservation across mammals has been noted [1]. Interestingly, the expression pattern of each globin in the CTDG is also developmentally restricted, and mutations disturbing the structure of the CTDG can cause diseases such as thalassaemia [31] and the "hemoglobin Kenya" [32]. A region 5' of the genes in the β-globin CTDG is considered conserved in mammals, and it has been identified as a locus control region (LCR), responsible of the regulation of genes in the β-globin CTDG [33, 34].

The Hox and β-globin CTDGs are examples of how some level of conservation is needed to ensure the proper function of a CTDG. However, there are CTDGs that are highly polymorphic, in part because generating molecular diversity is an integral part of their function.

1.4.2 Molecular diversity-generating CTDGs

Protocadherins are mainly expressed in synaptic regions on neurons to avoid improper neuronal connections, and are necessary for the neuron expressing them to connect to their correct target neuron [35]. As expected, such a function requires high levels of molecular diversity. There are two sections of the protocadherin CTDG: the variable region, and the constant region [36]. A protocadherin protein is produced by splicing together one exon from the variable region, and three exons from the constant region [35]. Moreover, even higher molecular diversity is achieved by regulating the expression of each protocadherin exon in a monoallelic manner [37].

Another CTDG that is very well known for its polymorphism-generating characteristic is the vertebrate-specific MHC CTDG. The main function of MHC genes is to generate glycoproteins to present peptides to T cells [5]. Similarly to the protocadherin CTDG, the MHC CTDG-encoded proteins are composed by a conserved domain, and two or three extracellular domains (for class I and class II molecules respectively), and a more conserved transmembrane domain. Molecular and phylogenetic evidence suggest that the MHC CTDG maintains high levels of molecular diversity via tandem duplication, unequal crossover, somatic recombination, and gene conversion on different sections of the CTDG [5].

Of interest is another CTDG involved in immune response, the KIR CTDG, which encode for receptors that interact with motifs of the MHC class I molecules to elicit or inhibit immune response in natural killer cells. Similarly to the MHC and protocadherin CTDGs, it is a CTDG where significant polymorphisms have been observed [4]. However, in the KIR CTDG, genomic and phylogenetic analyses suggest that most of the molecular diversity of the KIR CTDG was generated by insertion and movement of transposable elements in and out of the CTDG [4]. As opposed to the the protocadherin and MHC CTDGs, the diversity is not all contained in one organism, but as at least 30 different haplotypes observed across human populations [38].

1.4.3 Lineage-specific CTDG remodeling and their phenotypic products

Some of the more charismatic CTDGs involve lineage-specific expansions or contractions that are correlated with phenotypes characteristic to that lineage. A classic example involves the CTDG expansion of olfaction receptors in rodents and elephants, which rely heavily on olfaction cues [9].

The reduction in ORs has been associated with the evolution of color vision in primates [39, 40] (but see [41]), which caused them to rely more on visual cues, and less on olfaction. Gene loss in OR CTDGs in primates (via pseudogenization) concurrent with the evolution of trichromatic vision in primates has been put forward as evidence for the hypothesis of gene turnover as a response to ecological changes [9, 39, 40, 42]. In a similar fashion, it has been observed that the number of OR has decreased significantly in platypuses and dolphins, concurrent with their evolution to semi-aquatic or fully aquatic lifestyles, as well as the evolution of bill sense and echolocation sensory systems respectively [43]. Similarly, the number of genes in CTDG showing high levels of molecular variation, like the phospholipase-A [7], defensins [8], and kallikreins [44] gene families (not in all cases explicitly characterized as CTDGs) were expanded in venomous crotalous snakes, platypuses, and the king cobra (respectively) compared to non-venomous relatives.

Recently, the expansion of the Apolipoprotein D (ApoD) CTDG has been reported in different teleost fishes [45]. In this case, it is of interest that non-expanded ApoD genes tend to have conserved expression patterns, but expanded ApoD orthologs tended to be expressed in sub-sets of the tissues that their non-expanded orthologs expression pattern, and even got expressed in tissues their non-expanded orthologs did not [45].

Another interesting example of how CTDGs can promote phenotypic variation is the MHC CTDG, this time, not regarding MHC molecules, but speciation rates. It was reported that lineage-specific expansions and contractions of the MHC locus in several branches of the teleosts phylogeny were correlated with speciation rates in those lineages [6]. This discovery is particularly interesting because it support the hypothesis explaining how the number of MHC copies is restricted to the habitat of specific species [46]. Briefly, that hypothesis shows how if the balance of MHC copy number in an organism is broken (in hybrids from parents with different MHC copy number), the organism would have less fitness to the habitat of either of its parents, stabilizing species by reducing the fitness of the hybrids. This example, although different in mechanism, is similar to the well-known case of the chromosomal inversion causing barriers to recombination in heliconia butterflies [47]. In the case of heliconia butterflies, local chromosomal rearrangements can cause maintenance of linkage disequilibrium to the point of promoting speciation [47].

The prolactin CTDG is an interesting but more complicated example. Although it was first identified for its ability to stimulate lactation in rabbits, it has been implicated in functions as diverse as electrolyte balance, metamorphosis, brain development, behavior, and immunoregulation [48] across vertebrates. However, in rodents and bovids, where prolactin is part of a CTDG, one prolactin mainly expressed in the pituitary gland, and the rest of the prolactins in the CTDG are expressed the mammary gland and placenta, and other tissues. Interestingly, independent lineage-specific expansions of the prolactin CTDG have been reported in bovids and rodents [49]. However, the wide expression patterns of its genes make the interpretation of these expansions difficult to interpret.

As seen from these examples, CTDGs can drive high levels of phenotypic diversity by various mechanisms. Genes in CTDGs can be expressed in different tissues and developmental stages, drastically changing the phenotypic output of their molecular function (e.g. Hox CTDGs driving developmental programs in different parts of the embryo, or pancreatic enzymes expressed in a snake venom gland). CTDGs can also achieve high levels of phenotypic diversity by concentrating somatic recombination among their genes (e.g. MHC), or combining their genes via splicing (e.g. protocadherins). Also, high degrees of diversity can be achieved by CTDGs if recurrent CTDG expansions are permissible in lineage-specific manner (e.g. OR, ApoD, prolactin). Importantly, variation in CTDG copy number has been linked to speciation events, which reiterates the importance of CTDG in the biology of genes, organisms, population, and species.

The high conservation and biological importance of CTDGs like the Hox [50], β-globins [51], prolactin [52], or protocadherins [20] make a strong case justifying the need of considering a CTDG as a biological entity. However, we are still in need of a theoretical framework able to explain that the characteristics defining a CTDG are exploited by the biology of the genomes that harbor them, and that CTDGs are subject to evolutionary forces. In order to start proposing such a model, we first need to see how have CTDGs been defined and studied.

1.5 History of CTDG discovery and definitions

1.5.1 Identifying CTDGs before genome sequencing technologies

One of the first CTDGs ever characterized was the histone cluster. There was evidence suggesting that the partially characterized S9 mRNA from cleaving sea urchin embryos contained three RNA regions that coded for the three major histone fractions. Using RNA-DNA hybridization and DNA denaturation experiments, Kedes and Birnstiel showed that the S9 mRNA fraction not only contained at least the three major histone coding sequences, but also that the S9 mRNA contained very similar sequences in close proximity [53].

A few years later, it became evident that five non-α-globins existed and evolved from the duplication of a β-globin gene [54]. Data from pathologies like the "hemoglobin Kenya" [32] and the "hemoglobin Lepore" [55], in which the fusion of β-globin with γ- and δ- globins (respectively) is observed, suggested that the hemoglobin Kenya was a product of non-homologous recombination of tandemly arranged genes, which could also include other members of the globin gene family. More direct evidence supporting the hypothesis of the globin cluster came in the form of a restriction enzyme physical map, confirming that the delta and beta globins were next to each other, and promoting further research, which confirmed the existence of the beta-globin cluster (reviewed in [1]).

1.5.2 Gene-specific sequencing and CTDG identification

Advances in molecular biology techniques and sequencing benefited research in CTDGs. Through gene sequencing, it was known that the human growth hormone has high sequence similarity to the human prolactin gene and to other growth hormone-related genes (reviewed in [56]). Later on, the organization of the growth hormone CTDG was characterized by probing a phage clone library of human genomic sequences with the human growth hormone gene clone, and recognizing that several overlapping clones contained several copies of the growth hormone probe [57]. Similar techniques were used to identify the CTDG harboring the myosin heavy chains (myHC) [58]. However, the case of the myHC CTDG, in-situ hybridization experiments were used to further characterize the chromosomal location of the CTDG to the band q12 of the chromosome 14 [59]. Similarly, several olfaction receptor (OR) CTDGs were identified by in-situ hybridization of PCR products of the OR3B and OR5B genes [60]. Given the wide distribution of the OR CTDGs and their biological and evolutionary importance, the OR CTDG on chromosome 17 was independently sequenced, providing an even more detailed picture of that specific CTDG at that time [42].

1.5.3 CTDGs in the genomic era

As expected, the advent of genome sequencing in the late 1990s brought valuable tools and resources to the field of CTDG genomics. In 1999, by using BLAST searches against sections of available human genomic sequences, Wu and collaborators were able to locate the protocadherin CTDG composed by 52 gene duplicates to the band q31 of the human chromosome 5 [61]. With improved bioinformatics methods, the development of better computational resources, and a significant rise in the number of available vertebrate genomes, comparative studies on the structure and evolution of CTDGs across species was made possible. One such instance of early comparative analyses of CTDGs took place studying the evolution of the Siglec CTDG. By sequencing four Siglec CTDGs in Chimpanzee, Baboon, Mouse, and Rat; comparisons of their gene structure with that of the human Siglec CTDG provided valuable evidence supporting a hypothesis of CTDG expansion, diversification, and contraction in the Siglec CTDG [62]. That study provided the starting material for even larger studies, including human, chimpanzee, macaque, mouse, rat, dog, and opossum, where an interesting pattern of inversion duplication was observed, and it was hypothesized that such inverse duplication occurred over 180 million years ago, before the diversification of mammals [63].

Since at least the early 70s, CTDGs have been a fascinating and fertile field of molecular biology research. However, as can be observed from the studies mentioned above, the techniques, parameters, and types of results (e.g. chromosomal bands from in-situ hybridization, vs. restriction maps) are so different, that integrative approaches able to compare not only genes in CTDGs, but CTDGs in genomes using different CTDG identification techniques have been impractical. For example, how close two in-situ hybridization signals have to be to be as close as two mRNAs being in an arbitrarily short stretch of DNA cut by a pair of restriction enzymes?.

Given how important gene duplication is in the evolution of gene families and how dynamic the genome is in terms of its structure and gene order, being able to analyze CTDGs in a comparative framework is fundamental to understand how gene families and their genomic environment interact in the course of evolution. However, such an endeavor has been difficult to pursue because a historical lack of a unified method for defining and identifying CTDGs. As will be noted, although some effort has been done in the genomics community to develop CTDG identification methods, more work is needed to give the field of CTDG genomics the solid theoretical and methodological foundation it needs to make use of its full potential.

1.6 CTDGs definitions: challenges and proposed solutions

The first computational approaches to identify CTDGs were based only on sequence similarity. The idea was to use BLAST or related algorithms to identify all the genes similar to a query. For example, Martin and collaborators used blastn to find 6 killer cell immunoglobulin-like receptors (KIR) on an available section of chromosome 19 [4]. The authors used the resources available at the time, and leveraged them to produce a map of the putative KIR CTDG. However, although they report the intergenic distance between the KIR duplicates in the KIR CTDG as short (1–3 kbp), it is not clear how do they define the intergenic distances of those genes "short enough" to be considered part of a CTDG. It is important to mention that at that time, the human genome lacked relevant chromosome-level genomic sequences, so accounting for the gene duplicate density in the genome was not possible.

With the availability of more and longer genomic sequences, other methods for identifying CTDGs emerged. For example, a Casein and Enamel matrix protein CTDG was identified using sequence searches, but in this case, the structure of the duplicates in the cluster (i.e. intron – exon organization) in human and mouse was characterized using a dot-plot approach. In this case, the 12 duplicates of the cluster span more than 776 kbp on the human chromosome 4 [64]. Similarly, CTDGs were predicted in the human genome using dotplots from a self-alignment of the human genome [65]. Interestingly, the same authors used inter-species dotplots to predict the beta-globin CTDG in the galago genome [65]. However, discussion of the gene duplicate density of CTDGs in different species was lacking.

More recent methods have tried to include estimations of gene density in their calculations. For example, Cannon and collaborators used the average gene density of 50 Kbp or 30 Kbp in *Arabidopsis thaliana* as their length unit, so that tandem duplications are operationally defined as a set of homologous genes in a genomic region of 50 Kbp [66]. This threshold is based on the observation that there is on average one gene per 5 Kbp in A. thaliana, and that most of the regions between 5 Kbp and 50 Kbp contain densities of duplicated genes above their genome-wide estimate for a region of 5 Kbp.

In order to account for differences in gene family sizes, the authors placed N genes in random places on a 100 Mbp hypothetical genome, and calculated the "expected values" of having tandem duplications on a gene family of N genes "by chance alone". However, it must be pointed out that this null model could be greately improved, mainly because it assumes that genes are not duplicated from an existing gene, but somehow duplicated several times instantly and then put in random places in the genome. Although null models of gene duplication are needed, special attention should be put on proposing models or standardizing expected values that are biologically relevant.

In order to appreciate the importance of considering the gene duplicate density of a CTDG comparatively to the genome it belongs to, let us consider the hypothetical example of the fusion of two homologous chromosomes and a gene translocation event (fig. 1.3). In such a scenario, and considering only the fact that two genes from a gene family are in the same chromosome (as implied in articles describing CTDGs without considering gene duplicate density in their methods), spurious CTDGs could be called. In the case of the gene translocation (fig. 1.3 A.), if the gene copy is retro-transposed in the same chromosome (and exon-intron structure are not compared), a spurious CTDG could be called. In the case of the chromosome fusion, every gene will be clustered because there would be a copy of itself at the other end of the (now fused) chromosome (fig. 1.3 B.). Both of these events are biologically different from CTDGs that evolved by tandem gene duplications like the Hox, globin, or protocadherin

CTDGs, and should be studied independently. The problem of defining a CTDG without taking into account the gene duplicate density of the CTDG candidate with respect of that of comparable regions across the genome is apparent from the above mentioned hypothetical chromosome, because such a "definition" would make a categorical claim (this region is a CTDG) from only the observation of a quantity (the gene duplicates are in a X bp region; X being any number of base pairs). A more practical example, and in depth discussion was crafted by Duboule, in its criticism against the representation of the Hox cluster as a clean and organized CTDG across species, where he explains how it is misleading to classify Hox CTDGs across species as equally organized, without considering the immense diversity of such a CTDG [29, 67] and the phenotypes it can drive [30]. It is important to point out that I am assuming that closely (compared with its genomic background) spaced genes are generated by tandem gene duplication. However, substantial evidence exists supporting that claim [7, 61, 68–75].

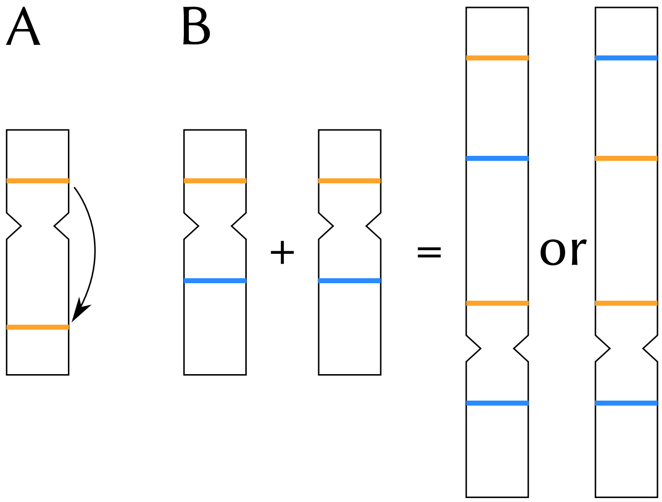

Figure 1.3: Scenarios where processes other than tandem gene duplication could result in spurious CTDGs if gene densities are not taking into account. A. Gene transposition. B. Chromosome fusion. In both cases exemplify how observing two genes of the same gene family (orange or blue) in the same chromosome is not enough evidence to call a CTDG

As shown above, defining what a CTDG is not only challenging, but also of prime importance if we want to study their evolution. In order to test if CTDGs are biological entities subject to selection, first we have to be able to define and identify CTDGs. However, the variability in gene density across chromosomes and more importantly, across species, make the identification of regions of "high" gene duplicate density challenging. Moreover, the methodological incongruence across CTDG identification (e.g. FISH, restriction maps, sequencing) has made comparisons between CTDGs, especially between species, difficult. In the first chapter, I propose a definition of CTDG that takes into account the specificities of gene density and gene duplicates positions in a genome, and implement such definition as an algorithm (CTDGFinder) that uses a spatial clustering algorithm to identify potential CTDGs, and then validates them taking into account the intergenic distance composition in the target genome. I evaluated the performance of CTDGFinder by identifying several well-known CTDGs (β-globin, Hox, protocadherin, galectin, Siglec, luteinizing hormone beta, prolactin, and growth hormone). Interestingly, all those CTDGs were first identified by several methods, but all were recovered by CTDGFinder. However, although previously identified as one CTDG, CTDGFinder identified two CTDGs for each of the galectin and siglec CTDGs, I found that the separation of each of those CTDGs corresponds to important evolutionary (cluster duplication) and regulatory (tissue expression enhancement) events in the Siglec and galectin CTDG respectively.

Are CTDGs just a statistical anomaly coming from random gene duplication?. Are CTDGs temporary errors during the evolution of gene families that are swiftly removed from a species genome?. Or, are CTDGs conserved features of genomes? Using CTDGFinder to map the repertoire of CTDGs on 22 selected mammalian genomes, chapter 2 is devoted to tackle these questions. By studying the distribution of CTDGs across genomes, I show how there are CTDGs that are indeed conserved across genomes, but are not uniformly distributed across chromosomes. Digging deeper into this observation, I discuss in how the human chromosome 19 (the most clustered chromosome in the human genome) evolved from two ancestral chromosomes, one of which has been highly clustered at least since the divergence of the eutherian lineage. Moreover, I show how the syntenic regions of chromosome 19 with other mammals have been under contraction, possibly consolidating CTDGs and increasing their gene duplicates density. Finally, I discuss evolutionary hypotheses derived from these analyses with high potential of improving our understanding of genome evolution and gene family organization.

Chapter 2

A novel homology-based algorithm for identifying closely spaced clusters of tandemly duplicated genes[1]

2.1 Abstract

Closely spaced clusters of tandemly duplicated genes (CTDGs) contribute to the diversity of many phenotypes, including chemosensation, snake venom, and animal body plans. CTDGs have traditionally been identified subjectively as genomic neighborhoods containing several gene duplicates in close proximity; however, CTDGs are often highly variable with respect to gene number, intergenic distance, and synteny. This lack of formal definition hampers the study of CTDG evolutionary dynamics and the discovery of novel CTDGs in the exponentially growing body of genomic data. To address this gap, we developed a novel homology-based algorithm, CTDGFinder, which formalizes and automates the identification of CTDGs by examining the physical distribution of individual members of families of duplicated genes across chromosomes. Application of CTDGFinder accurately identified CTDGs for many well-known gene clusters (e.g. Hox and beta-globin gene clusters) in the human, mouse, and 20 other mammalian genomes. Differences between previously annotated gene clusters and our inferred CTDGs were due to the exclusion of nonhomologs that have historically been considered parts of specific gene clusters, the inclusion or absence of genes between the CTDGs and their corresponding gene clusters, and the splitting of certain gene clusters into distinct CTDGs. Examination of human genes showing tissue-specific enhancement of their expression by CTDGFinder identified members of several well-known gene clusters (e.g. cytochrome P450s and olfactory receptor) and revealed tht they were unequally distributed across tissues. By formalizing and automating CTDG annotation, CTDGFinder will facilitate our understanding of CTDG evolutionary dynamics, their functional implications, and how they are associated with phenotipic diversity.

---

[1] This chapter was published in Molecular Biology and Evolution. 2007 [23]

2.2 Introduction

Gene duplications are among the most frequent types of mutational changes in genomes [16] and arguably the largest source of novel gene functions [76–78]. Gene duplication can occur by many different mechanisms [78], including transposition [79], polyploidization [80, 81], and recombination [82]. Recombination-based gene duplication results in tandem gene duplication, in which the gene duplicates lie adjacent to each other and are closely spaced on the chromosome [42, 61, 64]. Chromosomal regions containing multiple homologs that have arisen through tandem gene duplication are common features of genomes, and are often described as clusters of tandemly duplicated genes (CTDGs) [4, 20, 36, 83–85].

Notable examples of CTDGs include the vertebrate protocadherin gene clusters [20, 86], the vertebrate and invertebrate olfactory receptor gene cluster [9, 40, 87], the vertebrate natural killer cell receptor gene clusters [88], and the Hox gene clusters found in one or more copies across metazoans [4, 20, 75]. Many CTDGs contribute to traits that are highly variable, such as the composition of snake venom [44], the architecture of animal body plans [89], the olfactory repertoire [42] or the immune response [4].

Given the many CTDGs from diverse gene families found in a wide diversity of organisms, the absence of a formal definition of what constitutes a “cluster of tandemly duplicated genes” is surprising. The standard, informal definition that unites the known examples of CTDGs is that they represent groups of duplicated genes that are closely spaced [70], although CTDGs sometimes also contain non-homologous genes [90]. As practical as this definition may be, it is subjective. For example, should two genes located adjecent to each other on a chromosome be considered a cluster? Answering this question is challenging without considering the probability of observing two duplicates next to each other in the chromosome, which in turn requires knowledge of the number and distribution of duplictes in the genome a well as comparison of the intergenic distance between genes in the cluster with those in the rest of the chromosome.

An examination of the organization of the Hox gene cluster across diverse metazoans, which is often protrayed as a conserved, organized, and temporally and spatially clustered set of duplicated genes [24, 89, 91], is a good case in point. Whereas genes in vertebrate Hox gene clusters are typically closely spaced and encoded on the same strand, Hox gene clusters in other animal phyla show striking differences in the number of genes that are members of the gene cluster, in their intergenic spacing, as well as in their general organization [30]. For example, the Hox gene cluster in sea urchin *Strongylocentrotus purpuratus* contains genes exhibiting long intergenic distances and are encoded in both strands [29] (fig. 2.1). In the fruit fly *Drosophila melanogaster*, the Hox "gene cluster" is actually composed by two distinct clusters; the ANT-C cluster, which contains 11 non-Hox genes and five Hox genes encoded in both strands (five non-Hox genes and four Hox genes are in the 3' strand and the others in the 5' strand), and the BX-C cluster, which contains three Hox genes in the 3' strand and two non-Hox genes, one in each strand (fig. 2.1). In contrast, the HoxD gene cluster in the Mouse *Mus musculus* -one of the four Hox clusters in this organism- is composed by nine contiguous homologous genes in thes same strand (fig. 2.1).

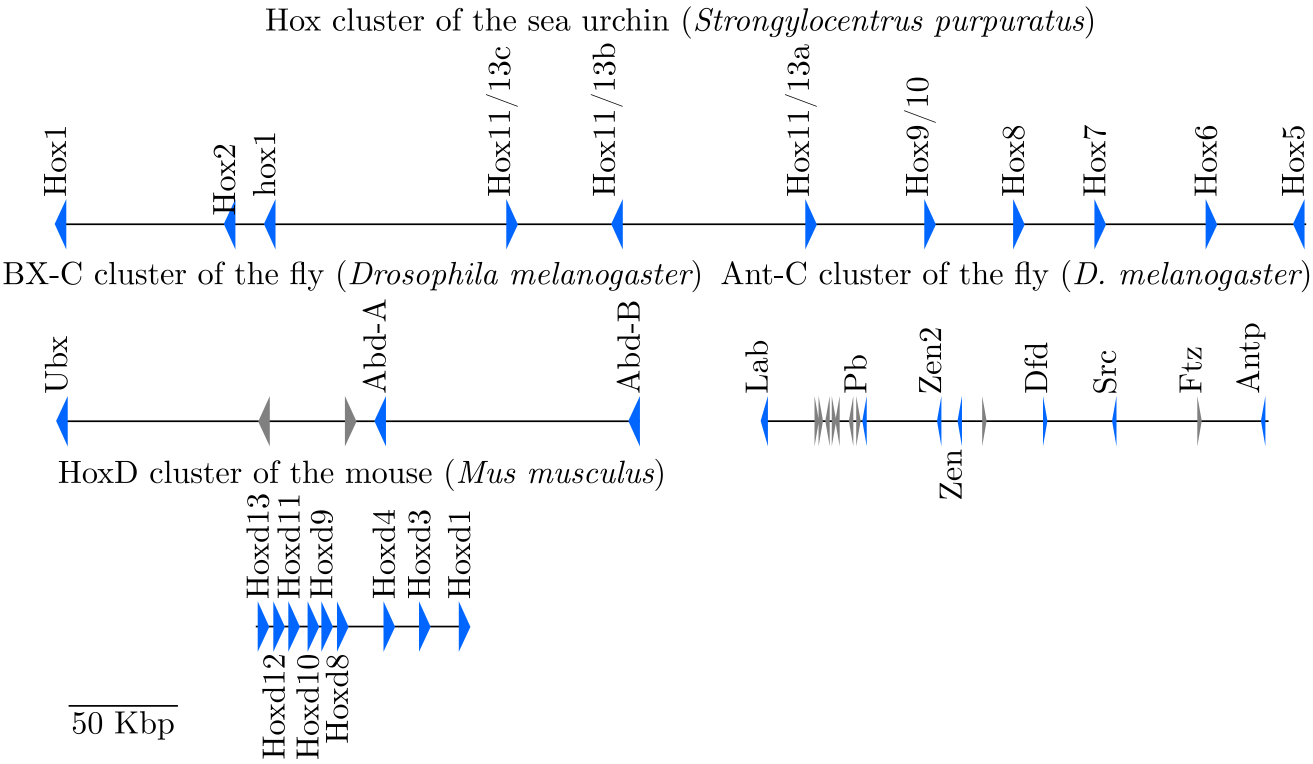

Figure 2.1: An illustration of the variation in the genomic organization of Hox "gene clusters" in three different animal species. The Hox gene cluster of the sea urchin (*Strongylocentrotus purpuratus*), the Bithorax complex (BX-C), and the Antennapedia complex (ANT-C) clusters of the fruit fly (*Drosophila melanogaster*), and the HoxD cluster of the mouse (*Mus musculus*). Genes belonging to the Hox gene family are shown in blue and intervening non-Hox genes in gray.

2.3 Methods

Gene Annotation

Genome annotation for all the mammalian species with assembled chromosomes available from the NCBI was downloaded from ftp://ftp.ncbi.nih.gov/genomes (table B.47). Genes were annotated using CDS (coding sequence) features from their annotation GenBank files. If two CDS features had the same start or end coordinate, the longest gene was selected. If two genes in the same strand had overlapping coordinates but different start coordinates, the gene with the start coordinate downstream from the other gene was removed.

CTDGFinder Validation on Previously Described (Reference) Gene Clusters

Representative protein sequences from the prolactin, growth hormone, Hox, Galectin, luteinizing hormone beta (LHB), β-globin, Siglec, galectin, and protocadherin (PCDH) gene clusters were downloaded from the NCBI (table B.1). Gene family cluster analysis was performed for each downloaded set of sequences using the CTDGFinder algorithm described below.

Phylogenetic Analysis of Orthologous CTDGs

In order to identify orthologous CTDGs, phylogenetic trees were built using the protein sequences from all the CTDGs in a gene family. The best substitution model was selected using ProtTest [92]. The phylogenetic tree was estimated using the rapid bootstrap function implemented in RAxML, version 8.2.4 [93], and 100 bootstrap pseudoreplicates were performed. The tree was rooted using evolutionary information about each family, and the leaves containing genes from known CTDGs in human and mouse were used to identify the clade with CTDG orthologs (selected clade). Given that CTDGs in human and mouse could have experienced gene deletions or accelerated rates of evolution, sister branches of the selected clade were also inspected for the presence of members of the known CTDGs. The set of CTDG orthologs was defined as the CTDGs with all their genes in the selected clade.

CTDGFinder Programing Environment

CTDGfinder was coded in Python 3.43 using the Pandas library v0.16.2 for table manipulation, Numpy v1.9.2 [94] for numerical arrays manipulation, BioPython v1.65 [95] for sequence manipulation and communication with the NCBI servers, scikit-learn v0.16.1 [96] for statistical analysis, and matplotlib v1.4.3 [97] and Bokeh v0.9.2 for graphics generation. CTDGFinder is freely available from https://github.com/biofilos/ctdg_finder.

Assessing the Degree of Clustered Genes with Human Tissue-Enhanced Expression

The "RNA gene data" expression dataset was downloaded from the Protein Atlas (tissue dataset) [98], and genes that were tissue enhanced were selected. Protein sequences from these selected genes were downloaded using the BioMart API from Ensembl. In cases where several protein sequences were mapped to the same gene identifier, only the longest proteins sequence was used. CTDGFinder was run for each of the selected genes. To evaluate whether a given gene was a member of a CTDG, its coordiantes were contrasted with those of all the clustered genes found by CTDGFinder. If the coordinates of a selected gene overlapped with those of a clustred gene, the gene was annotated as clustered.

Gene Family Cluster Analysis (CTDGFinder Algorithm)

Each query protein or protein set was used as a query for BLASTP (default E-value: 0.001 was used; it can be changed by the user) against all the proteome of the species under study. BLASTP hits that were either more than three times the length of the query or less than one third the length of the query were removed (length ratio less than 0.3). The set of BLASTP hits in each chromosome (pre-clusters) was processed using the meanshift algorithm as implemented in scikit-learn [96]. Since the meanshift algorithm is implemented to identify clusters in more than one dimension, the start coordinates of the pre-clusters were considered to be the x coordinates and the y coordinates were set to zero. The mean plus the standard deviation of all intergenic distances in the chromosomes where a given pre-cluster resides were used as the bandwidth parameter for the meanshift implementation. In the meanshift step, the pre-clusters can be ignored, subdivided, or confirmed as proto-clusters. Its output consists of a set of proto-clusters per chromosome.

In order to extract clusters that have more duplciates than expected in the genome, 1,000 random regions of the length of each proto-cluster were taken from the genome for the following analysis: an all vs. BLASTP search of the complete set of proteomes under study was used to extract the gene duplicates of the genes in the region, and the maximum number of hits was considered the maximum number of gene duplicates. Proto-clusters containing more gene duplicates than the 95^{th} percentile from the genome-wide sample set were considered clusters of tandemly duplicated genes (CTDG).

2.4 Results

Defining a Cluster of tandemly duplicated genes

We define a cluster of tandemly duplicated genes (CTDG) in a given genome as a genomic region that contains a statistically significant higher number of tandemly duplicated genes from a specific gene family than the average background genomic region of the same length.

The CTDG Finder algorithm

To formally define CTDGs we developed the CTDGFinder algorithm (fig. 2.2), which uses sequence similarity and the density of the distribution of duplicated genes across the genome to statistically assess and demarcate the presence of CTDGs in a genome. CTDGFinder is written in Python and is freely available from https://github.com/biofilos/ctdg_finder. Briefly, given a query reference protein sequence, or a set of sequences, and a subject genome or set of genomes, CTDGFinder uses the BLAST algorithm [99] to identify sequences that statistically significantly similar (homologs) in the subject genome(s). The sets of closely spaced duplicated genes identified on each of the chromosomes are considered candidate clusters.

Whether the genes in these candidate clusters comprise one or more genuine clusters of tandemly duplicated genes (CTDGs) is evaluated and determined by the meanshift algorithm [100]. Specifically, the meanshift algorithm treats a parameter as an empirical density function, and its objective is to find the region(s) of the parameter space with the highest density (or densities). In the context of the distribution of duplicated genes across a given chromosome, the meanshift algorithm identifies the genomic region(s) with the highest density (or densities) of duplicated genes of the protein reference sequence(s). Statistical assessment is performed by comparing the number of duplicated genes present in the genomic region(s) identified by the meanshift algorithm against an empirical distribution of duplicated genes. Such an empirical distribution is obtained by counting the highest number of duplicated genes from any gene family contained in each of 1,000 randomly sampled genomic regions from the same genome of length equal to that of the cluster candidate from the meanshift step. Genomic regions with more duplicated genes than the genome-wide 95th percentile of this empirical distribution are considered CTDGs (an overview of the algorithm is shown in fig. 2.2).

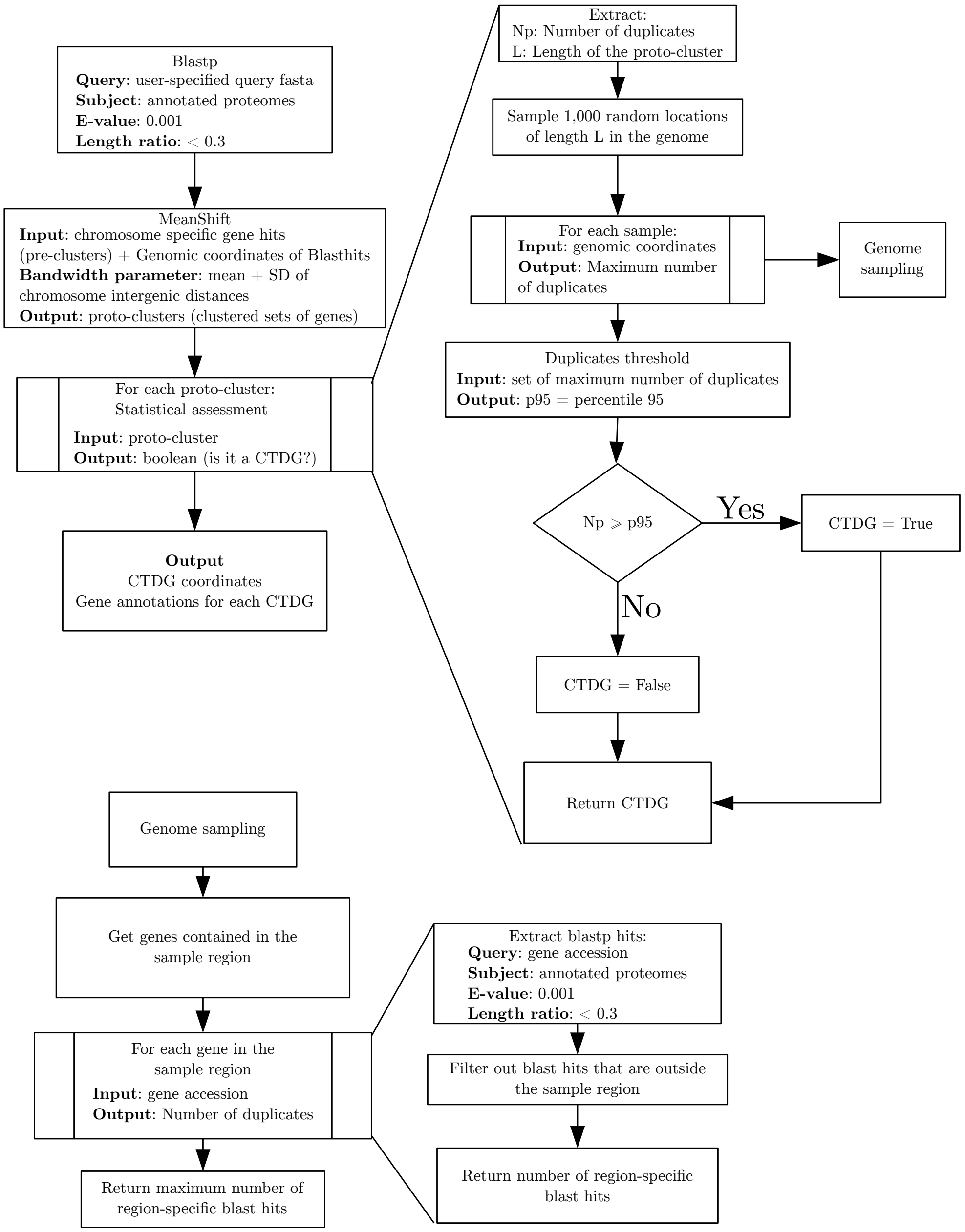

Figure 2.2: Overview of the CTDGFinder algorithm

CTDGFinder Recovers Several Well-Known CTDGs

To evaluate the performance of CTDGFinder, we first examined whether it was able to identify a diverse set of previously characterized, well-known gene clusters in the human and mouse genomes. To capture as much sequence diversity as possible, we retrieved all the genes from each of eight published gene clusters in either mouse (*Mus musculus*) or in the human (*Homo sapiens*) genome and used them as a combined query in CTDGFinder to identify CTDGs in the same or other genome. The list of mouse and human gene clusters used, and of the mouse and human genes used as queries for CTDGFinder are described in (table B.1).

CTDGFinder Performance When Searching the Same Genome

Using previously reported genes from eight gene clusters found in human and/or mouse genomes as queries, CTDGFinder correctly identified the CTDGs in their corresponding genomes (table 2.1). For example, all five genes reported to be part of the growth hormone gene cluster on the human chromosome 17 [101], all 26 protocadherin genes on the mouse chromosome 18 [102], all 11 HoxA genes in the human (chromosome 7 [103]) and mouse (chromosome 6 [103]) genomes, all five beta-globin genes in the human (chromosome 11 [34]) and mouse (chromosome 5 [104]) genomes, and all seven duplicated genes that are part of the luteinizing hormone beta (LHB) gene cluster in the human chromosome 19 [90] were identified as statistically significant CTDGs by CTDGFinder.

CTDGFinder also correctly identified known additional CTDGs in searches of the same genome (table 2.1). For example, using all 11 HoxA genes in the human chromosome 7 as query, CTDGFinder correctly identified CTDGs corresponding to the human HoxA, HoxB, HoxC, and HoxD clusters in the human genome; the same results were obtained for mouse (table 2.1).

The correspondance between the identified CTDGs and the previously described gene clusters was very good (table 2.1) and the few observed differences fell into three categories. The first concerned differences associated with the exclusion of nonhomologous genes that have historically been annotated as part of the cluster. For exaple, the RUVBL2 and NTF5 genes have historically been considered parts of mammalian LHB gene clusters, but their inclusion simply reflects the knowledge that those genes flank the LHB genes in human and chimpanzee.

The second category included cases in which the CTDGs contained one or a few additional genes not originally described as part of the cluster. For example, the HoxB cluster in both the human (chromosome 17) and mouse (chromosome 11) genomes included Hoxb13, which was not originally described as part of the mammalian HoxB gene cluster [103], but was subsequently added to it [105]. Similarly, using the 52 protocadherin genes in human chromosome 5 as a query [61], CTDGFinder identified a CTDG that additionally contained PCDHB16, whereas using the four Siglec genes on mouse chromosome 7 as a query [102], CTDGFinder identified a 7-gene CTDG that additionally contained genes 4931406B18Rik, Iglon5, and Vsig10L (table 2.1).

The third and arguably most interesting category included cases in which CTDGFinder split a previously described gene cluster into two distinct sub-clusters, which may or may not be both CTDGs. For example, CTDGFinder identified two distinct galectin cTDGs, a 4-gene and a 3-gene one, instead of the single 7-gene cluster previously reported to reside on human chromosome 19 [106]. CTDGFinder identified two separate CTDGs because the average intergenic distances of the genes in the 3-gene and 4-gene CTDGs are 13 and 38 kb, respectively, which are significantly smaller than the 790 kb that separates the to CTDGs. Interestingly, examination of tissue expression patterns of the genes in the two galectin CTDGs using data from the ProteinAtlas project [98] showed that expression of the genes in the 4-gene CTDG was enhanced in the placenta and bone marrow, whereas expression of the 3-gene CTDG was enhanced in skin and digestive tract (table B.2). Similarly, genes previously reported as part of the 9-gene Siglec cluster on human chromosome 19 [63] were identified by CTDGFinder as part of two separate CTDGs, containing four and seven genes (with mean intergenic distances of 38 and 29 kb), respectively, separated by 142 kb. In addition to the genes reported in the literature, the 4-gene CTDG additionally contained the SiglecL1 gene, and the 7-gene CTDG additionally contained the loc1053722490 gene. Interestingly, the two CTDGs directly correspond with the A and B sub-clusters that resulted from an inverse duplication of the CD33rSiglec cluster in eutherian mammals [63]. No significant difference in tissue expression patterns was found between the CTDGs (table B.2). Finally, using the 26 genes in the prolactin cluster on mouse chromosome 13 [107] as query, CTDGFinder identified as single 24-gene CTDG (table 2.1). An additional candidate cluster located 13.7 Mbp away and comprised of the genes Prl2c3, Prl2c2, and Prl2c5 was also identified, but it was below the 95th percentile of the empirical distribution of gene duplicates for a genomic region of that size and was not recognized as a CTDG.

Table 2.1: Comparison of eight human and mouse CTDGs inferred by CTDGFinder with the corresponding gene clusters previously described in the literature. Query Species: species* (chromosome). Query Genes: number of genes-gene family (chromosome). Subject Species: species* (chromosome). Subject Gene cluster: subject gene cluster (Chromosome), G. CTDG: Genes according to CTDGFinder. *species: M: Mouse, H: Human

Query		Subject		G. CTDG	Notes	References
Species	Genes	Species	Gene Cluster			
H	11-HoxA (7)	H	HoxA (7)	11	None	[83, 103, 105]
		H	HoxB (17)	10	None	
		H	HoxC (12)	9	None	
		H	HoxD (2)	9	None	
M	11-HoxA (6)	M	HoxA (6)	11	None	
		M	HoxB (11)	10	None	
		M	HoxC (15)	9	None	
		M	HoxD (2)	9	None	
M	7-LHB (19)	H	LHB (19)	7	Genes RUVBL2 and NTF5 were not used in the search because they are not homologous to LHB genes.	[90]
H	5-GH (17)	H	GH (17)	5	None	[101]
M	26-prolactin (13)	M	Prolactin (13)	24	In the 24-gene CTDG, Prl2c1was identified between Prl2a1and Prl4a1, instead of the reported Plf. In addition to the reported gene duplicates, Gm3821 was also identified as part of the CTDG. The remaining three genes (Prl2c4, Prl2c2, and Prl2c5) were identified as an independent cluster candidate but did not pass the statistical threshold to be considered a CTDG.	[107]
H	7-galectin (19)	H	Galectins (19)	4/3	Genes previously reported as part of a single cluster were identified by CTDGFinder as part of two separate CTDGs containing four and three galectin genes, respectively. LGALS17 was not included in the list of query genes and was not identified as part of the CTDG because it is annotated as a pseudogene.	[106]
M	2-galectin (7)	M	Galectins (7)	2	None	[108]
H	52-PCDH (5)	H	PCDH (5)	53	In addition to the genes reported in the literature, PCDHB16 was identified as part of the CTDG	[61]
M	26-PCDH (18)	M	PCDH (18)	26	None	[102]
H	5-β-globin (11)	H	β-Globins (11)	5	None	[34]
M	5-β-globin (7)	M	β-Globins (7)	5	None	[33, 104]
H	9-SIGLEC (19)	H	SIGLECs (19)	4/7	Genes previously reported as aprt of a single gene cluster were identified as part of two CTDGs, containing 4 and 7 genes, respectively. In addition to the genes reported in the literature, the 4-gene CTDG additionally contained SiglecL1, and the 7-gene CTDG contained loc105372490	[63]
M	4-SIGLEC (7)	M	SIGLECs (7)	7	In addition to the genes reported in the literature, the CTDG additionally contained 4931406B18Rik, Iglon, and Vsig10L	[62]

CTDGFinder Performance When Searching Different genomes

Given that CTDGFinder correctly identified several well known gene clusters in searches of the same genome, including known additional homologous CTDGs (table 2.1), we next sought to examine the performance of CTDGFinder on the human genome when using gene queries from the mouse genome (table 2.2) as well as on the mouse genome when using gene queries from the human genome (table 2.2).

In general, CTDGFinder correctly identified CTDGs for several well-known gene clusters in the human and mouse genomes. For example, using the 11 hoxA genes from the human genome as a query, CTDGFinder correctly identified a CTDG corresponding to the mouse HoxA gene cluster (table 2.2); similarly, using the 11 mouse HoxA genes as a query, CTDGFinder identified the human HoxA gene cluster (table 2.2). The same was true for the mouse galectin and β-globin gene clsuters, which were identified using their human homologs (table 2.2), as well as for the human β-globin gene cluster, which was identified by CTDGFinder using mouse homologs (table 2.2). CTDGFinder also correctly identified known additional homologous CTDGs in a given genome. For example, using all 11 HoxA genes from human chromosome 7, CTDGFinder correctly identified CTDGs corresponding to the HoxA, HoxB, HoxC, and HoxD clusters in the mouse genome (table 2.2). Likewise, using the mouse HoxA genes as a query. CTDGFinder correctly identified CTDGs corresponding to the human HoxA, HoxB, HoxC, and HoxD gene clusters (table 2.2).

Similarly to the results performance of CTDGFinder when searching the same genome, the differences between the inferred CTDGs and the previously described gene clusters when searching other genomes fell into three categories. The first category concerned differences associated with the exclusion of nonhomologous genes that have historically been considered to be parts of specific gene clusters, the second included single or a few gene differences between the CTDGs and their corresponding gene clusters, and the third category, cases in which a previously described gene cluster was split by CTDGFinder into two distinct CTDGs (table 2.2).

The most conspicuous differences between the identified CTDGs and the previously described gene clusters were observed in the mouse and human Siglec gene clusters. Specifically, using the Siglec genes on human chromosome 19 as a query, CTDGFinder identified a 6-gene CTDG on mouse chromosome 7 that contained two additional genes (4931406B18Rik and Vsig10L) in addition to those previously described for the mouse Siglec gene cluster [62]. Furthermore, using the four Siglec genes on mouse chromosome 7 as a query, CTDGFinder identified two separate CTDGs 69.0 Kb from each other, containing five and seven genes (with average intergenic distances of 44.5 and 29.4 kb), respectively, that correspond to the A and B sub-clusters previously identified by Cao and collaborators [63] (table 2.2). In addition to the genes previously reported [63], the 5-gene CTDG additionally contained SiglecL1 and VSIG10L, and the 7-gene CTDG additionally contained loc1053722490 (table 2.2). Interestingly, the human gene VSIG10L was identified as part of this CTDG only whem using mouse Siglec genes as queries (table 2.1, and table 2.2)

Table 2.2: Comparison of eight human and mouse CTDGs inferred by CTDGFinder with the corresponding gene clusters previously described in the literature. Query species: pecies* (chromosome). Query genes: number of genes-gene famliy (chromosome). Subject Gene cluster: Subject gene cluster (Chromosome), G. CTDG: Genes according to CTDGFinder. *Species: M: mouse, H: human

Query		Subject		G. CTDG	Notes	References
Species	Genes	Species	Gene Cluster			
H	11-HoxA (7)	M	HoxA (6)	11	None	[83, 103, 105]
		M	HoxB (11)	10	None	
		M	HoxC (15)	9	None	
		M	HoxD (2)	9	None	
H	7-galectin (19)	M	Galectins (7)	2	None	[108]
H	52-PCDH (5)	M	PCDH (18)	24	Pcdh1 and Pcdh12 were not found (these two genes were found as part of the CTDG when mouse genes were used as the query)	[102]
H	5-β-globin. (11)	M	β-globin (7)	5	None	[33, 104]
H	9-SIGLEC (19)	M	SIGLECs (7)	6	In addition to the genes reported in the literature, 4931406B18Rik and Vsig10L were identified as part of the CTDG	[62]
M	11-HoxA (6)	H	HoxA (7)	11	None	[83, 103, 108]
		H	HoxB (17)	10	None	
		H	HoxC (12)	9	None	
		H	HoxD (2)	9	None	
M	2-galectin (7)	H	Galectins (19)	4/3	Genes previously reported as part of a signle cluster were identified by CTDGFinder as parts of two separate CTDGs containing four and three galectin genes, prespectively	[106]
M	26-PCDH (18)	H	PCDH (5)	55	In addition to the genes reported in the literature, PCDHB16, PCDH1, PCDH12 were identified as part of the CTDG	[61]
M	5-β-globin (7)	H	β-Globins (11)	5	None	[34]
M	4-SIGLEC (7)	H	SIGLECs (19)	5/7	Genes previously identified as part of a single cluster were identified as parts of two separate CTDGs, containing five and seven genes, respectively. In addition to the genes reported in the literature, the 5-gene CTDG additionally contained SiglecL1 and VSIG10L, and the 7-gene CTDG additionally contained loc1053722490	[63]

Identifying CTDGs Across Placental Mammals

Given that CTDGFinder performed very well in recovering several well-known CTDGs in the human and mouse genomes, we next used all the paralogs from each of six published gene clusters in the human genome and used them as a combined query in CTDGFinder (table B.1) to identify CTDGs from the same gene families (galectin, hox, β-globin, Siglec, protocadherin, LHB, and growth hormone/prolactin) in 20 other mammalian genomes (fig. 2.3 and tables B.3 to B.14).

Overall, CTDGFinder identified all the CTDGs that are expected to be present and conserved in these 20 mammalian genomes (tables B.3 to B.14). for example, running CTDGFinder with the HoxA genes frm human as a query identified four Hox clusters in all the studied genomes, with the exceptions of HoxA and HoxB in vole (*Microtus ochrogaster*), HoxB in orangutan (*Pongo abelii*), and HoxC in opossum (*Monodelphis domestica*) (tables B.3 to B.14). The reason for these exceptions is that these clusters are present in nonassembled scaffolds that are not part of the standard genome assembly peovided by GenBank.

Using the human β-globins as the query, CTDGFinder also identified both α- and β-globin CTDGs in all species, except in vole, where only the α-globin was found, and in horse (*Equus caballus*) and rabbit (*Oryctolagus cuniculus*) where only the β-globin was found (α- and β-globin CTDGs were differentiated by constructing their phylogeny; see methods). The number of genes contained in the β-globin CTDG varied across species (fig. 2.3). The largest β-globin CTDG was found in goat (*Capra hircus*) and stemmed from a previously reported β-globin cluster duplication [51], followed by the rat (*Rattus norvegicus*) 8-gene CTDG. Most species contained either 5-, 4-, or 3-gene CTDGs, but opossum (*Monodelphis domestica*) contained a 2-gene CTDG. this variation is likely due to both gene duplicates gain and loss, as well as errors in annotation.

Running CTDGFinder with the 52 reported genes from the human protocadherin cluster [61] as a query identified one protocadherin CTDG per species, with the number of duplictes per CTDG ranging from 15 in dog (*Canis lupus familiaris*), to 53 in human (tables B.10 and B.11). the only exception was pig, where a 13-gene CTDG (average intergenic distance of 13.1 kb) and a 20-gene CTDG (average intergenic distance of 16.0 kb) were found in the same chromosome but separated by 213 kb.

The galectin CTDG in all species was identified by CTDGFinder using the 7-gene human galectin cluster reported by Than and collaborators [106] as a query. Orthologous galectin CTDGs were

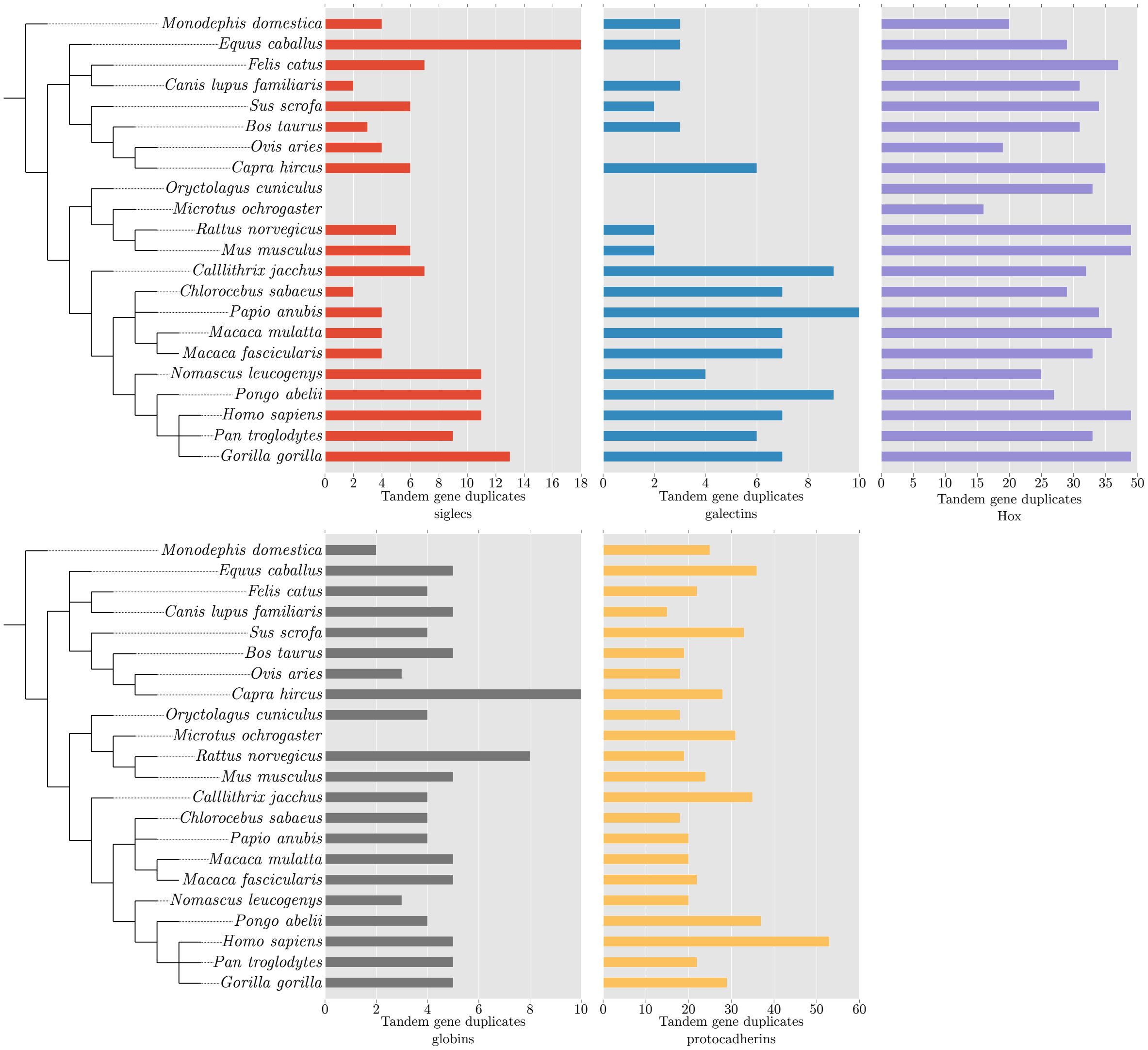

Figure 2.3: Distribution of CTDGs in five sepected gene families across 22 mammalian genomes. For gene families with more than one CTDG in a given organism, numbers of gene duplicates reflect the total number found in all CTDGs associated with that gene family.

extracted from all the galectin CTDGs using phylogenetic trees rooted on the clade containing the galectins LGALS1 and LGALS2, following a previously reported galectin phylogeny [108] (tables B.3 and B.4). No orthologous CTDGs were found in cat (*Felis catus*), vole, rabbit, and sheep, although vole and cat contained other homologous galectin CTDGs. One orthologous 3-gene CTDG was found in cow, dog, and horse, whereas mouse, rat, and pig, contained one orthologous 2-gene CTDG. All other mammals contained two orthologous galectin CTDGs containing 2-3, and 3-7 genes respectively (tables B.3 and B.4).

Using the protein sequences from the previously reported Siglec cluster in human [63], CTDGFinder identified a Siglec CTDG in cow, dog, goat, green monkey (*Chrlorocebus sabaeus*), horse, cat, rhesus macaque (*Macaca mulatta*), mouse, sheep, chimpanzee (*Pan troglodytes*), baboon (*Papio anubis*), orangutan, rat, and pig (tables B.13 and B.14). The number of paralogs in these CTDGs ranged from two in dog and green monkey, to 11 in orangutan and 18 in horse. No Siglec CTDG was found in rabbit and vole. Two Siglec CTDGs were found in marmoset (*Callithrix jacchus*), gorilla (*Gorilla gorilla*), crab-eating macaque (*Macaca fascicularis*), human, and gibbon (*Nomascus leucogenys*). In all these cases, the number of genes in the first (ordered by genomic coordinates) CTDG (ranging between two and six gene duplicates) was smaller or equal to the number of genes in the second CTDG (2-9 gene duplicates) (fig. 2.3 and tables B.13 and B.14).

Finally, CTDGFinder accurately identified CTDGs for clusters that are known to be taxonomically restricted to certain lineages. For example, using genes from the human LHB cluster, CTDGFinder identified an LHB CTDDG in primate lineages (table B.9), consistent with the previously described primate-specific presence of this gene cluster [90]. However, a small CTDG of two LHB duplicates was identified in cow and horse; Interestingly, the duplications that led to the formation of these CTDGs appear to have been independent of the duplications that led to the formation of the primate LHB cluster. Similarly, using the genes from the prolactin cluster in mouse as query, CTDGFinder was able to identify CTDGs for previously described prolactin clusters in certain rodents (rat and vole) and bovids (cow) [107, 109] (table B.12). Among the bovids, a previously unreported 12-gene prolactin CTDG was identified in goat (*Capra hircus*).

What Fraction of Genes Showing Enhanced Expression in Human tissues is Clustered?

To illustrate the potential of CTDGFinder as a tool for furthering our understanding of the function of CTDGs in the human genome, we used the 3,450 genes showing tissue-specific enhancement of their expression in 32 human tissues and organs (genes showing tissue-specific enhancement were those exhibiting "at least five-fold higher mRNA levels in a particular tissue as compared with average levels in all tissues") [98] as individual queries in CTDGFinder to examine whether they were clustered or not. On average, 25% of these enhanced genes were part of CTDGs. However, clustered genes were unequally distributed across tissues, with the percentage of clustered genes per tissue ranging from 56% (49/88) in the appendix, to 12% (74/597) in the cerebral cortex (fig. 2.4).

Examination of the enhanced genes that reside within CTDGs identified members of several well-

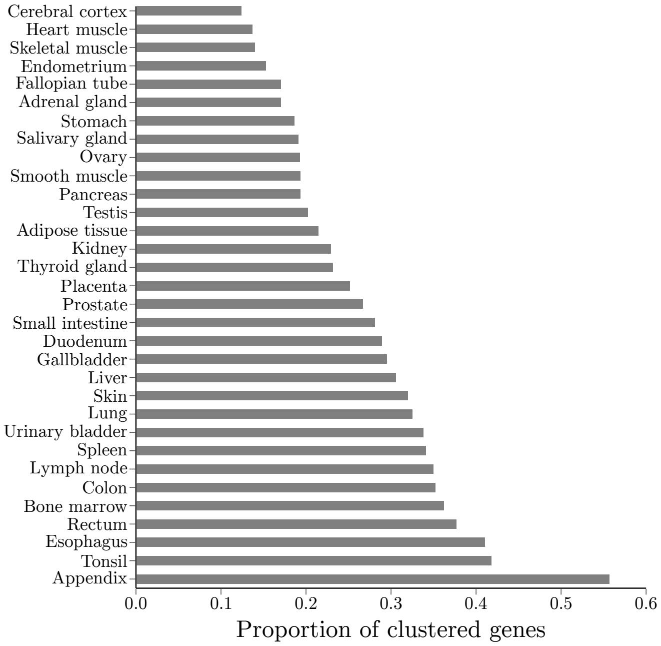

Figure 2.4: Relative abundance of clustered genes (i.e. genes contained in CTDGs) in each tissue showing tissue-specific enhancement of their expression across 32 human tissues and organs.

known gene clusters. In general, gene families that are known to be parts of clusters like olfactory receptors [40], phospholipase A [110], golgins [111], histones [112], cytochromes P450s [75], aquaporins [113], myosin chains [114], Hox genes [83], and protocadherins [61], were enhanced in different tissues (tables B.15 to B.46). For example, nine of the 49 clustered genes with enhanced expression in the appendix are chemokine ligands or receptors, six are eukocyte immunoglobulin-like receptors, and four are Siglecs. Similarly, 20 of the 75 clustered genes in the bone marrow are annotated as part of the histone cluster 1, 32 of 72 clustered genes in the lymph node are involved in the immune system (e.g. chemokines, MHC, T-cell, and B-cell associated), and six of 12 clustered genes in the endometrium belong to the Hox gene family (clusters A and D). Finally, in the palcenta, clustered genes encode for proteins important in tissue structure remodelling like metallopeptidases (2 of 41 clustered genes), collagen (3 genes), and gap junction proteins (3 genes), as well as immunity related genes (one interleukin receptor) (tables B.15 to B.46).

Usage recomendations

CTDGs are features of the spatial rearrengement of genes in chromosomes. For that reason, genomes with high quality assemblies (e.g. genomes with assembled chromosomes) and gene annotation will give optimal results. However, genomes with fragmented assemblies can also be used, as long as they contain the coordinates of genes present in their corresponding scaffolds along with their protein sequence information. Although CTDGFinder works for any genome, its application in genomes with highly fragmented assemblies might be problematic because of the lack of continuity of non-assembled chromosomes. This fragmentation might also influence the statistical sampling process by lowering the number of gene duplicates per sample (because the sampled regions will be shorter), leading to spurious genome-wide 95th percentile thresholds. Similarly, application of CTDGFinder in poorly annotated genomic regions (i.e. regions where many actual genes are unnanotated), especially ones containing CTDGs, will be problematic, due to inaccurate information about the structural arrangement of genes.

To further explore our algorithm's behavior with assemblies of different quality, we examined CTDGFinder's performance in chromosome-level and scaffold-level assemblies and their corresponding gene annotations of the human genome (chromosome-level assembly: 32,010 genes in 260 linkage groups; 23 of these linkage groups are chromosome-level in size, the remaining 237 have fewer than 238 genes, and 173 linkage groups have fewer than 10 genes; scaffold-level assembly: 32,010 genes in 335 linkage groups; none of these linkage groups are chromosome-level in size, and 202 have fewer than less than 10 genes). We used the same queries as previously (see section "Identifying CTDGs Across Placental Mammals") and employed CTDGFinder to identify CTDGs in Hox, protocadherin, globin, galectin, and Siglec families. In all but one case, the same number of TDGs was found for each human assembly. The only exception was in the protocadherin gene family, where PCDHGA2 was found as part of a CTDG using the chromosome-level assembly, but not when using the scaffold-level assembly because the gene was not present in this assembly's annotation. Furthermore, comparison of the mean 95th percentile (MP95) values showed that they were slighty lower when using scaffold-level assemblies (MP95: 2.34) compared to when using the chromosome-level assembly (MP95: 2.47, but such a difference was not statistically significant (T-test Pvalue: 0.700), suggesting that the process used in our statistical sampling is robust to differences in the quality of the assembly.

Data input: CTDGFinder uses a series of annotation files and one pre-computed all vs. all BLASTP results output to run. Specification and helper scripts, as well as instructions on how to use them can be accessed on the "extras" directory of the CTDGFinder repository (https://github.

com/biofilos/ctdg_finder). CTDGFinder requires protein sequences as its input and cannot be used with genomic DNA sequences. If genomic DNA sequences were used (via TBLASTN), multiple BLAST hits could either be due to the presence of duplicated genes or to the presence of duplicated domains in one protein sequence. Furthermore, users should take into account that gene duplicates from different gene families can vary in their evolutionary dynamics. All the sequences (and their BLAST hits) from a query sequence set provided by the user are assumed to belong to a gene family, and all the hits from them will be treated as hits of that gene family. With this in mind, the query sequence set should include as many sequences belonging to the gene family under study as possible. In that way, as much sequence divergence as possible will be accounted for, including highly diverged gene sequences from a gene family, increasing the chances of a comprehensive search for CTDGs. However, CTDGFinder can also be run using only one sequence as query. In this case, CTDGFinder will perform well if the sequence divergence between the genes in the family under study is low, but not so well if the sequence divergence between gene family members is very high (highly divergent genes might be missed in the BLAST search).

To help the user in the delimitation of a query sequence set (especially in highly divergent gene families, CTDGFinder can be run in iterative mode (option –iterative). In this mode, CTDGFinder will perform the BLAST search and length overlap filters using the query sequence set provided by the user, and then all the complete sequences resulting from the first BLAST step will be extracted and used as a new query. A good case in point is the Hox cluster. Using a query containing the genes from the HoxA cluster without EVX1 (a very distant homolog to the Hox genes) in iterative mode, CTDGFinder identified all the CTDs incuding EVX1 and EVX2. Whereas powerful, it is important to note that, since the itertive mode uses all the proteins from a BLAST search as the new queries, it might also result in spurious results because the sequence search space will be so big that sequences with poor sequence similarity can get included in CTDGs.

Finally, CTDGFinder is coded to use several CPU cores, if needed (this is useful when testing various E-value thresholds and sequence inputs). As a reference, the mean running wall time of CTDGFinder when using the queries reported in table B.1 for human and mouse was 59.38s in default mode, and 5.05m in –iterative mode (this time estimate does not include calculation of pre-computed all vs. all BLAST results output).

Sensitivity of the BLAST E-value threshold: CTDGFinder uses the E-value extracted from BLAST (plus the length overlap filter encoded in CTDGFinder) to identify potential gene duplicates. Selecting an appropriate E-value threshold depends on the variability of the gene family in question. To explore the effects of varying the E-value threshold on CTDG inference, CTDGFinder was run using E-values ranging from 1^{-10} to 1 for all the gene families under study in human and mouse. The number of CTDGs and total number of gene duplicates (TDG) identified for the galectin, globin, Hox, and protocadherin gene families did not change in either human or mouse (number of CTDGs and TDG did not change for the primate-specific LHB CTDG) (fig. A.1). The number of CTDGs and TDGs in the GH gene family did not change in human, but changed in mouse. this is because a query containing GH was able to identify the distantly related (rodent and bovid specific) prolactins at high (0.1, 1) E-values. Similarly, using a query containing prolactin sequences yielded variable results on both human and mouse. The number of prolactin CTDGs was observed to change in mouse because there is a small cluster of four genes that was seen at the edge of several percentile 95 thresholds. It might seem counter-intuitive that decreasing the E-value will increase the number of clusters. However, if the genes in question have enough similarity, but the number of gene duplicates is close to the 95^{th} percentile of the empirical distribution of gene duplicates, a low E-value might cause such empirical distribution to filter out more potential duplicates, thus resulting in a lower 95^{th} percentile threshold. For this reason, it might be advisable to increase the number of sampled regions used to build the empirical distribution when using very lo E-values. Finally, the Siglec gene family showed variations in both the number of CTDGs and TDGs. This is a consequence of the high sequence variation between the Siglec CTDGs and the genes in the Siglec family.

2.5 Discussion

From the clssic globin and Hox gene clusters [1, 103] to the more recenty described venom-related gene clusters in snakes [7, 44], the study of protein families whose genes are closely spaced on chromosomes has greatly enhanced our understanding of physiology, development, and the genetic basis of phenotypic diversity [3, 8, 87, 115]. In the post-genomic era, however, the fact that such gene clusters are still defined arbitrarily and in different ways in each gene family or genome, makes attempts for the kind of comparative analyses required to understand the dynamics of gene cluster evolution and function problematic.

Our formal definition and identification of clusters of tandemly duplicated genes (CTDGs; implemented as CTDGFinder) through a statistical approach that takes into account both intergenic distance and homology solves this problem. This approach not only enables the comparison of the genomic arrangement of well-known clusters in different model organisms, but also the discovery of novel gene clusters or novel gene arrangements in genomes. For example, by using previously known sequences of the prolactin gene cluster in mouse, CTDGFinder was able to find novel prolactin CTDGs in goat and sheep (table B.12), providing further evidence supporting the independent evolution of prolactin clusters in ruminants and rodents [56]. Thus, coupled with robust phylogenetic analysis at the gene and species levels, CTDGFinder could be very useful for distinguishing between different types of clusters (e.g. primary, secondary, and independently evolved) comprised of paralogous genes and the evolutionary history of their assemblies [116].

Using a formal definition for CTDGs also has the potential to help identify interesting gene arrangement and orientation features in clusters and inform our understanding of the evolutionary steps that explain their current assembly. For example, according to CTDGFinder, the single Siglec gene cluster on the human chromosome 19 is actually comprised of two distinct CTDGs (tables 2.1 and 2.2). Interestingly, the two inferred CTDGs correspond precisely to the two sub-clusters (A and B) that Cao and co-workers previously identified and inferred to have been generated through a large-scale inverse duplication of the ancestral Siglec locus in a vertebrate ancestor [63]. Such inversions can help stabilize the size of a gene cluster by reducing the effectiveness of recombination to add or remove additional gene duplicates to the existing cluster [63, 117].

Availability of a formal definition and means of characterizing CTDGs will also facilitate efforts to understand the functional implications of clustering. For example, CTDGFinder inferred two galectin CTDGs on the human chromosome 19 (table 2.1), rather than a single cluster [118]. This split into two CTDGs appears to be informative for function; genes on one CTDG showed enhanced expression placenta and bone marrow, whereas genes on the other CTDG are enhanced in skin and digestive tract (table B.2). At a broader level, CTDGFinder can be used to understand whether tissue-specific expression of clustered genes is evenly distributed across human tissues and organs (fig. 2.4), which in turn could be associated with the types of functions of genes expected to function in them (e.g. many secreted proteins are known to be expressed such as the liver and salivary glands) [98]. Additionally, given that CTDGFinder can simultaneously analyze the genomes of multiple species, a formal definition enables the investigation of the types of functional categories of genes (e.g. immunity, metabolism) that tend to be clustered in the genomes of organisms from diverse lineages and which potentially could be implicated in the generation of interesting lineage-specific phenotypes.

Finally, a formal definition for CTDGs has the potential to greatly aid in understanding the relationship between the specific genomic organization of a given gene cluster with its mechanism of regulation. For example, using a diverse and rich body of data on animal Hox gene cluster organization, Duboule [30] has argued that Hox gene cluster organization in different animals (fig. 2.1) has strong implications for how the activity of Hox genes is regulated in these organisms. The first step toward answering this question, not just for the animal Hox gene clusters, but for the wide diversity of gene families forming gene clusters, is the availability of a clear, precise, unambiguous, and easy to implement definition, such the one provided in this study.

2.6 Acknowledgements

We thank members of the Rokas lab for critical feedback on the work described in this study. This work was conducted in part using the resources of the Advanced Computing Center for Research and Education at Vanderbilt University. JFO was supported by the Graduate Program in Biological Sciences at Vanderbilt University. This work was supported in part by the March of Dimes through the March of Dimes Prematurity Research Center Ohio Collaborative and the National Science Foundation (DEB-1442113 to A.R.).

Chapter 3

The Genomic Geography and Evolution of Clusters of Tandemly Duplicated Genes in the Human and Mammal Genomes

3.1 Abstract

Clusters of tandemly duplicated genes (CTDGs) are nearly ubiquitous in life's genomes, and are associated with several well-known gene families, such as olfactory receptors, zinc fingers, and immunity-related genes, as well as with several highly variable traits, including olfaction, body plan architecture, and pathogen resistance. However, these observations are usually anecdotal, restricted to specific cases, and lacking evolutionary context. In this study, we use a robust statistical approach to characterize the CTDG repertoire and analyze the distribution of CTDGs across 18 mammal genomes, including human. We found that, on average, 16% of the genes in each species are parts of CTDGs. Although genes in CTDGs are enriched for several biological processes, these tend to be involved in the interactions between the organism and its environment. We further found that mammalian CTDGs are not uniformly distributed across chromosomes and that orthologs of the human chromosome 19 are among the most clustered chromosomes in nearly all mammalian genomes analyzed. We also found evidence that the human chromosome 19 was formed by a fusion event that occurred before the diversification of the rodent and primate lineages and maintained its high density of CTDGs during its subsequent evolution. Finally, using chromosome-level alignments across mammalian genomes, we show how the syntenic regions of the human chromosome 19 have been shrinking, increasing their gene density and possibly increasing the compactness of its CTDGs. These results suggest that CTDGs are a major feature of mammalian genomes and provide novel insights into the origin and evolution of regions with unusually high densities of CTDGs.

3.2 Introduction

Clusters of tandemly duplicated genes (CTDGs) are a familiar feature of genomes across the tree of life. CTDGs have been associated with phenotypic diversity in a wide variety of traits and organisms, such as in cellulose degradation in bacteria [119], copper resistance in yeasts [120], antimicrobial protection and venom in mammals and reptiles [8], and pathogen resistance in plants [121]. The classic example of how CTDGs are involved in the generation of phenotypic diversity are the Hox gene clusters found in diverse metazoans; variation in the function and regulation of the Hox genes is one of the major contributors to the diversity of metazoan body plans [24, 91]. Similarly, the gene duplications, conversions, and functional divergence of genes in the globin CTDGs in vertebrates are responsible for the functional diversity in the binding and transport of oxygen exhibited by members of this gene family [65, 122].

One of the major dimensions of variation in CTDGs across genomes is the number of genes contained in a cluster. For example, the CTDGs formed by olfactory receptors (ORs) can be highly variable, from a single OR and the absence of a CTDG in elephant sharks, to thousands of ORs and 345 CTDGs in mouse [9, 40]. A similar trend is exhibited by genes in Killer Cell lectin-like receptor (KLRA) CTDGs; whereas only one KLRA pseudogene is found in human chromosome 12, 11 and 25 functional KLRA genes (plus 9 and 4 pseudogenized copies) are found in the syntenic regions of mouse chromosome 6 and rat chromosome 4, respectively [88].

In the human genome, in addition to the above mentioned OR, Hox, and globin examples, CTDGs are also known for gene families whose functions have been implicated in diverse processes, such as brain development (e.g; Protocadherin [20]), immunity (e.g; KIR and Siglec [4, 63]), and pregnancy (e.g; Psg [115]). Interestingly, some of these CTDGs are known to be near each other in the genome. By far the most conspicuous aggregation of CTDGs to a single genomic region is human chromosome 19, which is characterized by an unusually high density of CTDGs [123]. For example, zinc finger genes [124], cytochrome P450 genes [75], Siglec genes [63], and gonadotropin β-subunit genes [125] all exist as CTDGs in chromosome 19.

Although many human CTDGs have been well-characterized [69, 75, 118, 124–127] and the high density of CTDGs in chromosome 19 is well-established [123], a systematic study of the distribution of CTDGs in humans and the degree to which they are conserved in other mammals is lacking. For example, are CTDGs a common feature of the human genome, and more generally of mammalian genomes, or a rare one? Are CTDGs randomly distributed across human and mammalian chromosomes? How conserved are human CTDGs and CTDG neighborhoods in other primates and mammals?

To address these questions, we comprehensively characterized the genomic geography and evolution of CTDGs in the human and select high-quality mammalian genomes. We found that CTDGs are common genomic features, being present in all mammalian genomes examined and significantly associated with biological processes, such as olfaction, signaling, and immunity. Furthermore, mammalian CTDGs are not uniformly distributed; the most spectacular case of this deviation is the human chromosome 19 and its orthologs in mammals, which are 1.75 times more densely populated with CTDGs than the average mammalian chromosome. Interestingly, we found that human chromosome 19 formed from the fusion of two ancestral chromosomes, only one of which was CTDG-rich. By exploring the relationship between gene family dynamics and CTDG membership, we observed that gene number expansion and contraction events are concentrated at the base of the mammalian phylogeny, suggesting that most CTDGs present in mammalian genomes stabilized their gene numbers prior to the diversification of the major mammalian lineages.

Aside from mapping the genomic geography of CTDGs across mammals and tracing their evolution, this work provides a methodological framework for the genome-wide evolutionary study of CTDGs as well as support for the idea that CTDGs are biologically relevant features of genomes. Our data show that CTDGs are not mere statistical artifacts distributed at random; rather, they are concentrated in some regions, with human chromosome 19 being an extreme such example. Furthermore, CTDGs can be both diverse in gene numbers and density at the interchromosomal level, but also stable across (chromosome) orthologs. These observations open interesting avenues of research. For example, understanding how evolutionary processes modulate the diversification and stabilization of CTDGs could yield insights into how the evolution of genes, gene families, and chromosomes are related.

3.3 Methods

Data acquisition

The genomes of select primates (*Callithrix jacchus*, marmoset; *Chlorocebus sabaeus*, green monkey; *Gorilla gorilla*, gorilla; *Homo sapiens*, human; *Macaca mulatta*, macaque; *Nomascus leucogenys*, gibbon; *Pan troglodytes*, chimpanzee; *Papio anubis*, baboon; *Pongo abelii*, orangutan), and select mammals (*Bos taurus*, cow; *Canis familiaris*, dog; *Equus caballus*, horse; *Felis catus*, cat; *Mon-*

odelphis domestica, opossum; *Mus musculus*, mouse; *Ovis aries*, sheep; *Rattus norvegicus*, rat; *Sus scrofa*, pig) were downloaded from the REST Ensembl service (https://rest.ensembl.org, https://rest.ensemblgenomes.org; date of access: Aug 10, 2018). For each species, gene annotations, proteome sequences, and chromosome lengths were downloaded from Ensembl, keeping only genome contigs with more than 10 genes and longer than 10 Kbp (see table B.48 for genome assembly version information).

CTDG identification

To identify clusters of tandemly duplicated genes, we used a modified version of the CTDGFinder software [23]. Instead of using blast results from a given sequence query, CTDGFinder was modified to use the Panther database of protein families [128]. Briefly, all the HMM profiles of protein domains in the Panther (v. 12 [128]) database were first used to annotate the proteomes of each genome using the HMMer sequence search software, version 3.1b2 [129]. HMMer subject-query hits with less than 30% mutual coverage were discarded.

Next, for each protein domain identified in a genome, the chromosomal coordinates of all the proteins in each chromosome (or genomic scaffold) that contained that protein domain were processed by the MeanShift algorithm [100] to obtain CTDG candidates (i.e.the genomic region(s) with the highest densities of that domain). Statistical significance, i.e.whether a CTDG candidate was considered a CTDG, was evaluated by comparing the number of proteins belonging to a Panther family in the genomic region identified by the MeanShift algorithm against an empirical distribution obtained by counting the highest number of genes that belong to a Panther family in each of 1,000 randomly sampled genomic regions from the same genome of length equal to that of the CTDG candidate. A CTDG candidate was considered a CTDG if its number of genes belonging to a Panther family was equal of greater than the 95th percentile value of the distribution generated by the examination of 1,000 randomly sampled genomic regions. In that way, CTDGFinder analysis resulted in clusters containing proteins that shared at least one Panther family.

The next step consisted of collapsing CTDGs with overlapping coordinates, so that potential nested CTDGs were collapsed into a single CTDG. As a result, each CTDG may contain genes that belong to one or more Panther families. Finally, we rerun the MeanShift algorithm to further refine CTDG boundaries.

As CTDGFinder uses a statistical approach to identify clusters of tandemly duplicated genes, some CTDGs also contain genes that do not share homology with the set of tandemly duplicated genes that defines the cluster (e.g; a 5-gene CTDG may be comprised of 4 duplicated genes and 1 unrelated gene). Those unrelated genes were not considered when calculating various gene-based metrics (e.g; total gene length, number of genes), but were considered for the calculation of various cluster-based metrics (e.g; length of clustered regions, percentage of chromosome length spanned by CTDGs), as they contributed to the length of the entire cluster.

GO enrichment

To explore any potential biases in the gene functional categories present in CTDGs, we performed GO enrichment analyses. Per gene GO annotation from human and mouse was downloaded from the BioMart service on Ensembl (last accessed Aug 10, 2018), and the gene ontology files were downloaded from geneontology.org (last accessed April 29, 2017). GO enrichment analysis was performed by comparing the distributions of clustered genes in each chromosome against all protein-coding genes in the genome for human and mouse using the python module Goatools (doi: 10.5281/zenodo.31628). Statistical significance was assessed using the Benjamini/Hochberg false discovery rate method [130] with a significance cutoff of 0.005.

Chromosome orthology calling

To compare the chromosomal distribution of CTDGs across species, we inferred orthologous protein-coding regions for all human chromosomes. To do so, we first downloaded Ensembl's orthology assignments for all the genes in the genomes we examined using Ensembl's REST service (www.rest.ensembl.org; last accessed March 10, 2018). Next, given two chromosomes (A, B), one from the human genome and the other from a different mammalian genome, we calculated the proportion of genes in chromosome A that have orthologs in chromosome B ($O_{A,B}$). Similarly, we calculated the proportion of genes in chromosome B that have orthologs in chromosome A ($O_{B,A}$). Finally, if at least one of $O_{A,B}$ or $O_{B,A}$ was greater than 0.25, we considered chromosomes A and B to be orthologs.

Clusteredness estimation

The distribution of CTDGs is not uniform across chromosomes. To compare the distribution of CTDGs between species, we first quantified how clustered each chromosome is by calculating how different each chromosome's distribution of CTDGs was to all the other chromosomes in a given genome. Examination of the ranking of each chromosome in a given genome based on its proportion of CTDGs provided an assessment of how clustered each chromosome was in that genome. We call this ranking the clusteredness of a chromosome.

The clustering pattern of a chromosome was measured by calculating the distribution of clustered genes in non-overlapping 10Kb windows on each chromosome. That distribution was normalized, so that it could be treated as a probability mass function (PMF). To compare the clustering pattern of two chromosomes, we calculated the Jensen-Shannon distance (JSD) [131] from their PMFs. The all versus all pairwise JSDs were then used to generate a graph, where each chromosome corresponds to each node in the graph, and each JSD between each pair of chromosomes corresponds to each edge in the graph. We then used the farness centrality measure (FCM) [132] to measure how distant a chromosome was to the rest of the chromosomes in a genome. Briefly, the FCM for a given chromosome A was computed by adding all the JSDs connecting to A, and then normalized by dividing the raw FCM value by the number of nodes minus one. The clusteredness rank of a chromosome was computed by sorting the chromosomes according to their normalized FCM values, so that the chromosome with rank 1 was the chromosome with the most different clustering pattern relative to the rest of the chromosomes.

Chromosome length evolution

The observation that certain chromosomes remained highly clustered through evolutionary time proved to be an interesting aspect of our analysis. To gain further insight into the evolutionary processes underlying this conservation, we examined the dynamics of CTDG expansions and contractions of all human chromosomes specifically along the path from the mammalian common ancestor to the human lineage tip in each chromosome. Multiple chromosome alignments for all human chromosomes were downloaded from Ensembl, using their "25 mammals" EPO alignments [133], including inferred syntenic sequences for internal nodes in the chromosome alignment tree accompanying each alignment (alignment tree). The dynamics of contractions and expansions during the mammalian history of each human chromosome were examined by performing linear correlations between the length of the alignment at each node on all alignment trees (for each human chromosome), and the relative age of each node (defined as the distance from the root of the tree, to the most recent common ancestor of the focal node and the human tip, normalized by the length of the tree).

Gene copy dynamics

It has been well established that gene duplication and loss are the main processes by which gene families evolve [134]. Since CTDGs are typically the result of a series of gene duplication and loss events, we investigated the gene number dynamics of clustered genes. The OMA standalone version [135] was used to identify hierarchical orthology groups [136] of the clustered genes found in this study. Gene gain and loss events were estimated using the PyHam python module, using the mammalian species phylogeny obtained from the TimeTree database [137].

3.4 Results

CTDGs are abundant and non-uniformly distributed across mammalian genomes

To get a comprehensive picture of the distribution of CTDGs across species, we used CTDGFinder to identify CTDGs in select high-quality mammal genomes (see methods). A CTDG is described here as a region of the genome that contains a significantly higher number of duplicated genes than the average region of the same length on that genome. CTDGs can be described in terms of their sequence length as well as in terms of the number of clustered genes they contain. On average, 16.0% (standard deviation 4.3%) of the genes in a mammalian genome are clustered. The genome with the smallest fraction of clustered genes was that of pig (6.2%) and that with the largest was mouse (24.3%). Examination of the distribution of CTDGs in the context of the mammalian phylogeny reveals a small decrease in the fraction of clustered genes in the primate lineage (fig. 3.1A; average percentage of genes clustered in primates = 14.4%, average in non-primate mammals = 17.7%). Our GO enrichment analysis shows that the biological processes most commonly enriched in clustered genes in human and mouse are related to olfaction and immunity (fig. A.2 and table B.49).

Importantly, CTDGs are not uniformly distributed across chromosomes (fig. 3.1A, fig. A.3). In all species, there are at most two chromosomes that seem to have a higher fraction of clustered genes compared with the rest of the chromosomes in each genome (fig. A.3). This observation is of particular interest, because it suggests that extreme clustering patterns could have been conserved in orthologous chromosomes during the evolution of mammals, giving rise to the idea that the genomic organization of CTDGs is subject to selection, and serves as a relevant hypothesis to be tested in this study.

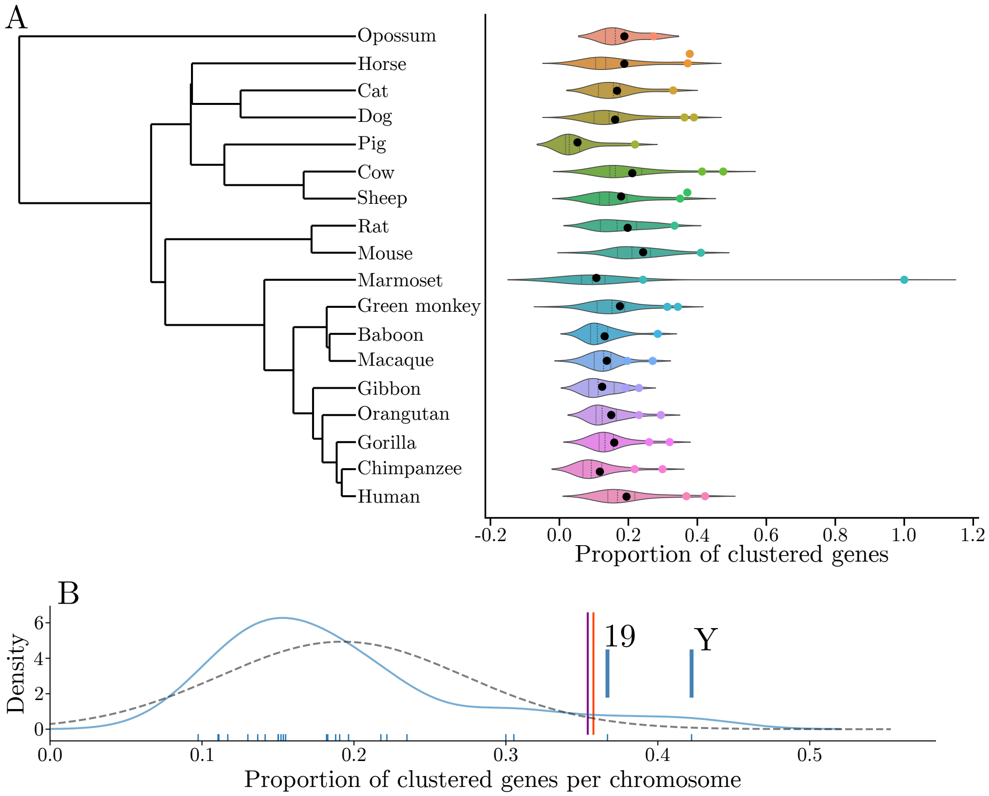

Figure 3.1: A large proportion of genes in mammalian genomes belongs to CTDGs. (A) Distribution of CTDGs across the mammalian phylogeny. Violin plot: distribution of the proportion of clustered genes across chromosomes. Lines inside the plots show the interquartile ranges. Black dots: Mean proportion of clustered genes in each species. Colored dots: chromosomes with higher proportion of clustered genes than the 95th percentile. (B) Detail of the distribution of the proportion of clustered genes in the human genome. Ticks: data points for each chromosome. Blue curve: KDE estimate. Dashed curve: Fitted normal distribution. Red line: 95th percentile. Purple line: mean + standard deviation. Labeled blue bars: Chromosome with higher proportion of clustered genes than the 95th percentile.

One fifth of human genes are part of CTDGs

To infer the distribution of clustered genes in the human genome, we used the CTDGFinder algorithm to examine the frequency of CTDG occurrence across all human chromosomes. We found that 3,821 / 19,665 (19.4%) human genes belong to 961 CTDGs, which collectively occupy 163.5 Mbp

(5.3%) of the human genome. The average length of clustered genes (i.e; belonging to CTDGs) is 18 Kbp, which is significantly lower than an average length of 53 Kbp for non-CTDG genes (Mann-Whitney P-value: 6.6×10^{-256}). The high number of clustered ORs in mammalian genomes as well as their intron-less structure [40] might bias the average length of clustered genes. Removal of all ORs from the mammalian genomes examined did not change the average length of non-clustered genes (53 Kbp) but slightly increased the average length of clustered genes changes to 20.1 Kbp. However, the difference in length between clustered and non-clustered genes remained statistically significant (Mann-Whitney P-value: 5.4×10^{-151}).

Consistent with our previous results [23], the CTDGFinder software accurately identified several well-known CTDGs (fig. 3.2). For example, the Siglec [126], Hox [103], and globin [1] CTDGs were all correctly identified in the human genome. Similarly, the HLA genes [138] were found in 3 CTDGs composed of 12, 4, and 19 genes in chromosome 6; the protocadherin CTDG was found in chromosome 5 with 55 genes [20]; the previously reported ORs in chromosome 11 [40] were recovered as 14 CTDGs that contained a total of 193 genes; and the histone cluster in chromosome 6 [112] was found in three CTDGs of 47, 4, and 15 gene duplicates. Interestingly, the longest CTDG in the human genome is the protocadherin CTDG [61] on chromosome 5 with 55 genes.. The 4-gene beta-globin CTDG was found in embedded in an OR-rich CTDG on chromosome 11 [33]. The next largest CTDG, with 46 duplicates was a histone-containing CTDG on chromosome 6. A comprehensive table with the CTDGs (and their genes) for all mammalian genomes analyzed in this study, including the human genome, can be found in table B.50.

Human chromosomes 19 and Y have the highest densities of CTDGs

To test if CTDGs are uniformly distributed across the human genome, we characterized each of the human chromosomes in terms of their number and total length of clustered genes, as well as the number and length of CTDGs. We found that on average, 5.92% of each chromosome's length is covered by CTDGs and on average, 19.23% of genes in each chromosome belong to CTDGs. The two chromosomes that exhibit the highest percentages of clustered genes in the human genome are chromosome Y (19 clustered genes / 45 genes; 42.2%) and chromosome 19 (525 clustered genes / 1,431 genes; 36.7%) (fig. 3.1B fig. 3.1C). The high percentage of clustered genes in chromosome Y

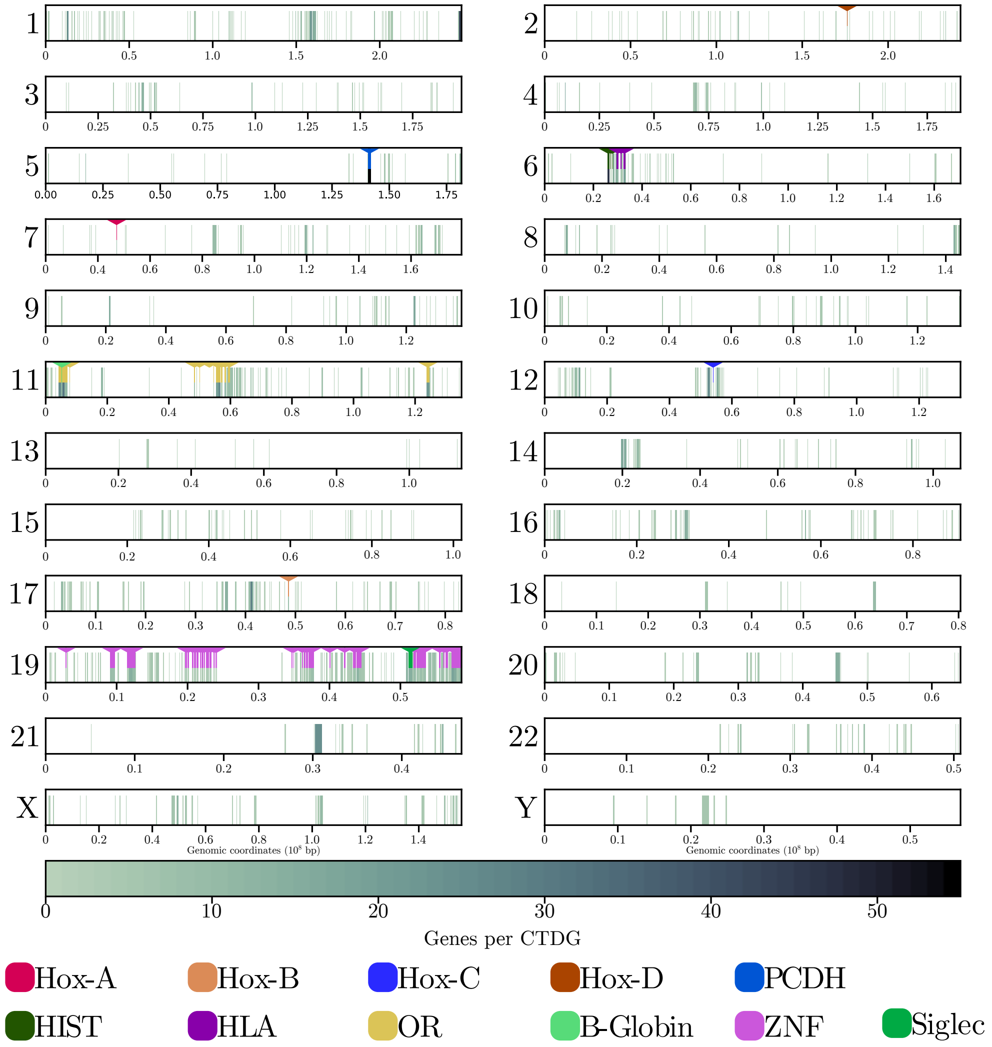

Figure 3.2: Distribution of CTDGs across all human chromosomes. Each plot represents a chromosome. The location and length of all CTDGs are shown as rectangles in the bottom section of each plot. Representative CTDGs are highlighted by a rectangle in the top section of the plots. An extended polygon is included in short CTDGs to improve visibility.

is likely due to its small number of genes. Furthermore, the percentage of chromosome Y length spanned by CTDGs is 2.9%, which is below the average (5.9%) and on par with several other human chromosomes. The longest CTDGs in chromosome Y contain RNA-binding motif Y chromosome (RBMY) genes, which encode for nuclear proteins involved in spermatogenesis [139]. Other CTDGs include deleted in azoospermia (DAZ) genes [140], and testis-specific protein Y-linked (TSPY) genes [141].

In contrast to chromosome Y, whose high percentage of CTDGs appears to be at least in part a consequence of its small gene number, the much larger chromosome 19 appears to be genuinely densely populated with CTDGs. Specifically, not only more than one third of its genes belong to CTDGs (525

/ 1,431), but also 32.0% of its length is composed by CTDGs, by far the highest of human chromosomes. For comparison, the chromosome with the second highest percentage of chromosome length spanned by CTDGs is chromosome 11, with only 12.1%. Furthermore, the percentage of the length of chromosome 19 that is spanned by CTDGs is higher than the 95th percentile of the data, and higher than the mean plus a standard deviation of a normal distribution fitted to the percentage of chromosome length spanned by CTDGs in all the human chromosomes (fig. 3.1B). These data suggest that CTDGs are not uniformly distributed across chromosomes, with chromosome 19 being by far the most notable deviation (fig. 3.1B). The CTDG with most genes in chromosome 19 contains 19 zinc finger duplicated genes. In addition to these zinc finger CTDG, chromosome 19 also contains the KIR [4], kallikrein [127], Siglec [63], OR [40], CEACAM [63], PSG [69], and Galectin [118] CTDGs.

Gene family dynamics of CTDGs

To gain insight into when gene families associated with CTDGs experienced gene gain and loss events, we estimated gene duplication and loss events on genes that are parts of CTDGs using a graph-based approach as implemented by the OMA standalone version [136]. We found that most of the duplication events in the dataset preceded the diversification of eutherians. Interestingly, most gene gains and losses in clustered genes occurred in the rodent lineage (fig. 3.3).

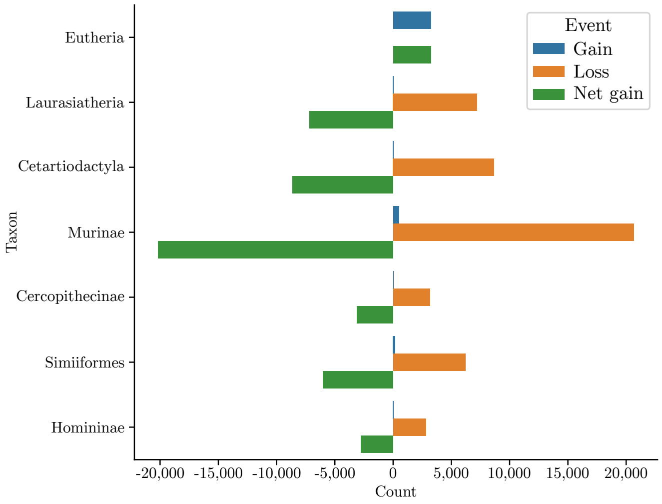

Figure 3.3: Gene family dynamics across the mammalian lineage. Gene gains, losses and net gains (gene gains + gene losses) in different mammalian clades

The extreme clustering of human chromosome 19 is highly conserved across mammals

Our finding that there are at most three chromosomes in each mammalian genome that contained a significantly higher fraction of clustered genes than the rest of the chromosomes raised the possibility that these highly clustered chromosomes are in fact orthologous and have maintained their clusteredness during the evolution of mammals. We found that in all the genomes studied here, except dog and pig, at least one of the five most clustered chromosomes is orthologous to the human chromosome 19 (hc19 ortholog). In the case of the dog genome, the two chromosomes with highest clusteredness rank are orthologs of the human chromosome 11, which is the third most clustered human chromosome (below chromosome 19 and chromosome Y) and is also known to be rich in CTDGs, including an OR CTDG [40] and a fibroblast growth factor CTDG [142]. Most mammalian genomes contain a single hc19 ortholog that is highly clustered; that is the case for all primates, except gibbon, rodents, as well as horse (fig. 3.4). Moreover, in genomes that contain two chromosomes considered by our cutoffs orthologous to hc19 (see methods), except gibbon, only one of these two chromosomes is highly clustered.

The gibbon chromosomes are the exception, presumably because the gibbon genome has experienced a very large number of genomic rearrangements, especially around transposable elements (TEs), compared with other primates [143]. Although CTDGs tend to be depleted in TEs, the significantly higher amount of TEs in the gibbon genome could have increased the probability of TE insertion in CTDGs [123], increasing the likelihood of CTDGs from the hc19 orthologs in gibbon being broken up by TE insertions, shuffled, or transposed to other chromosomes. To test whether the highly rearranged genome influenced the distribution of CTDGs in gibbon, we run phylogenetic independent contrasts [144] on the mean fraction of chromosome length spanned by clusters, and the mean fraction of clustered genes across chromosomes in the primates in our dataset. We found that although gibbon has the lowest mean fraction of chromosome length spanned by CTDGs, it differs from marmoset (the second lowest value) in less than 0.002 PIC units (mean PIC: 0.05; fig. 3.5.A). Similarly, the mean fraction of clustered genes per chromosome in gibbon shows little variation from the rest of the primates (fig. 3.5.B).

To understand the evolution of clustering of human chromosomes, we used the EPO genomic chromosome alignments from Ensembl [133], including inferred ancestral sequences, to estimate the rela-

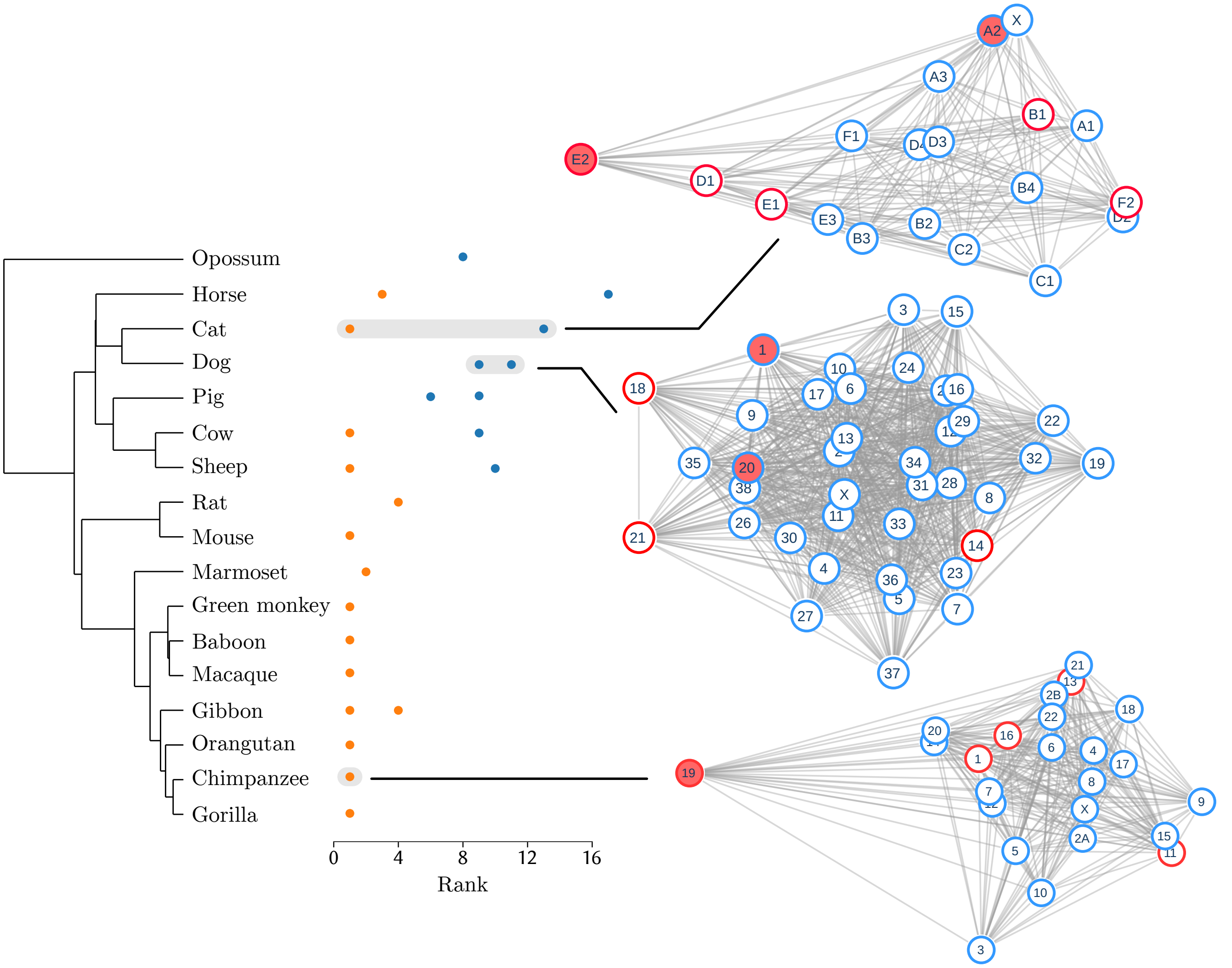

Figure 3.4: Clusteredness rank of the chromosome 19 orthologs on selected species. Orange: Chromosomes on the top 5 rank. Blue: Chromosomes on ranks above 5. To the right, network representations of selected species. Red border: Chromosomes in the top 5 rank. Pink: Human chromosome 19 orthologs. Edges are drawn using 1 - farness as weight.

tive time (in phylogenetic tree units) at which the syntenic regions of each of the human chromosomes across 26 mammals changed in size. By examining the change in the length of syntenic sequences (extant and ancestral) for each human chromosome in the context of the mammalian phylogeny, we inferred the rate at which each the orthologous regions of each chromosome shrunk or expanded (fig. 3.6). Specifically, for all the human multiple chromosome alignments we built linear regressions of the aligned genomic ancestral sequences at each node and the ages of the nodes examined. The slopes of the regressions between changes in alignment length and evolutionary time are positive for all human chromosomes, which suggests that all human chromosomes, or more precisely all aligned regions of human chromosomes, have been shrinking (fig. 3.6.A). Furthermore, along with chromosome Y

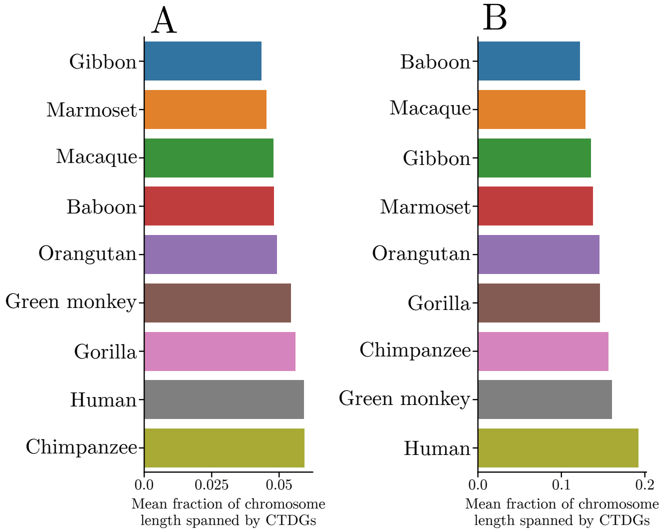

Figure 3.5: Phylogenetically corrected variation in CTDG length and composition across primates. A: Mean fraction of chromosome region spanned by CTDGs. B: Mean fraction of clustered genes per chromosome.

(Pearson R value = 0.17; fig. 3.6.D), chromosome 19 shows the largest slope (Pearson R value = 0.19; fig. 3.6.C), suggesting that human chromosome 19 has been shrinking at a higher rate than the rest of the human chromosomes. To test if the observed behavior was a byproduct of chromosome size, we carried out a correlation of the calculated slope against chromosome size. We found a strong but non-significant correlation between slope and chromosome length (Pearson coefficient: -0.34, P-value: 0.1), suggesting that although chromosome length is a confounding variable in our calculations, it is not significant to explain the variance in the calculated slopes for each chromosome. Both chromosome 19 an Y showed by far the highest slope of all human chromosomes, including when compared to other small human chromosomes, such as chromosomes 20, 21, and 22 (fig. 3.6.B).

3.5 Discussion

Historically, the discovery and description of CTDGs has been piecemeal and subjective [23, 30]. Although studies on subjectively defined CTDGs have been extremely useful and have greatly advanced our knowledge of their evolution and function [20, 85, 91, 118], a comprehensive investigation of CTDG evolution has so far been lacking. In this study, we provide the first systematic characterization of the genomic geography and evolution of CTDGs in mammalian genomes. Although it is hardly surprising that mammalian genomes harbor CTDGs, our findings suggest that CTDGs are highly preva-

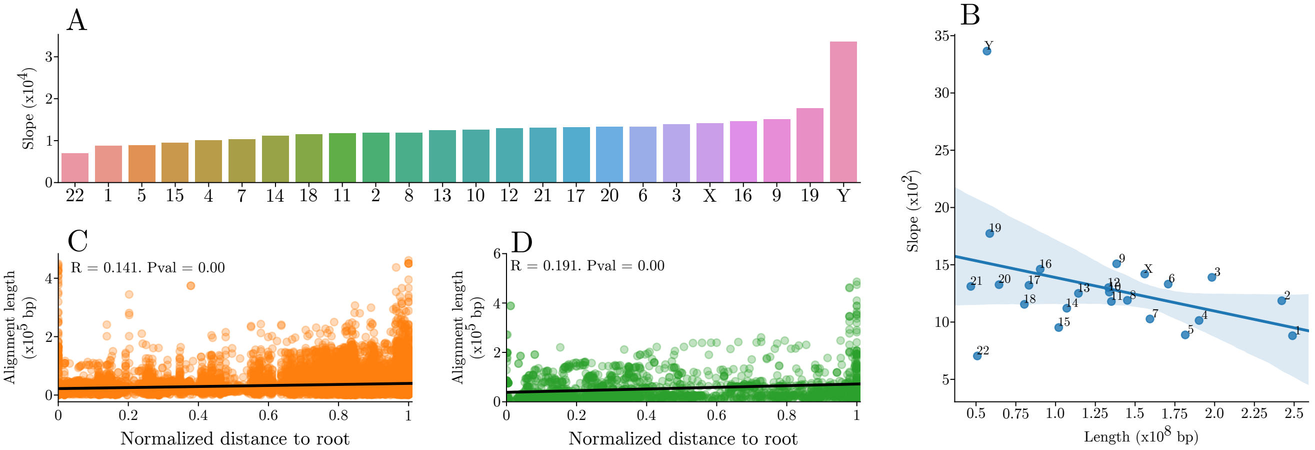

Figure 3.6: Rate of human chromosome shrinking in the mammalian lineage. A: slope calculated from correlations between the total alignment length of all EPO Ensembl alignments for each human chromosome and the normalized distance to the root of the tree of the EPO alignment at each node. B: Influence of chromosome length in the slope calculated in A. C: Example correlation for chromosome 19 (R: Pearson R value; Pval: P-value). D: Example correlation for chromosome Y.

lent – nearly one in five mammalian genes are part of CTDGs – and constitute common and conserved features of mammalian genomes.

CTDGs are not uniformly distributed across mammalian genomes; the same is true for the human genome (fig. 3.2). Our results showcase that CTDGs can vary widely in terms of their gene density (number of genes / Kbp), length, and number of genes. For example, with a cluster length of 147.5 Kbp and 10 gene duplicates, the iconic Hox-A cluster in humans [3] (fig. 3.2, chromosome 7) is more dense (0.07 genes per Kbp) than a zinc finger CTDG in the same genome [124] (fig. 3.2, chromosome 19) with 9 genes over 444.3 Kbp (0.02 genes per Kbp). The diversity of CTDG length and gene density can be further appreciated in the 14 OR-containing CTDG in chromosome 11 on the human genome [40] (fig. 3.2, chromosome 11), which have sizes ranging from 30 Kbp to 917 Kbp. Moreover, while the longest CTDG (1.4 Mbp) has 10 gene duplicates, the second longest CTDG (1.2 Mbp) has 43 gene duplicates.

From the genomic landscape perspective presented here, we can visualize how chromosome 19 is especially CTDG-rich, with more than 13% of the clustered genes in the human genome being located there (fig. 3.2, chromosome 19). Furthermore, our study found that 36.7% of the genes in chromosome 19 are part of CTDGs, a substantial increase over the previously reported estimate of 30% [145].

Importantly, CTDGs can contain more than one cluster; such is the case of the 5-gene globin cluster, which is located inside an 8-gene OR cluster [33] (fig. 3.2 chromosome 11). In terms of the high physical linkage of their genes and their involvement in complex phenotypes, it could be said that CTDGs resemble supergenes [146]. However, we did not see a high difference in recombination rate on sites inside and outside CTDGs (fig. A.4), with the expected exception of chromosome 6, where the MHC complex resides, and has been observed to be highly polymorphic, with high recombination observed inside the MHC cluster [5].

The high density of CTDGs in chromosome 19 puzzled us because the evolutionary maintenance of all these CTDGs "in place" likely required strong selection. By tracing the evolutionary history of chromosome 19 "clusteredness", we found that chromosome 19 was formed by a chromosomal fusion before the diversification of primates and rodents (fig. 3.3). Interestingly, we found that most of the CTDGs in the human chromosome 19 come from only one of the two proposed chromosome 19 orthologs, which points to a model where CTDGs have indeed been maintained in the mammalian lineage. Furthermore, examination of chromosome-level sequence alignments suggests that the syntenic regions of chromosome 19 have been shrinking (fig. 3.6.A). A caveat of our experiments is that we can only examine the evolution of the parts of chromosome 19 that are syntenic and alignable across species; thus, it is possible that the chromosome itself could be expanding if the amount of non-homologous sequence has increased during evolutionary time. That caveat notwithstanding, we hypothesize that the high density of CTDGs might have evolved by contractions in chromosome length.

Why would CTDGs in mammalian genomes appear to be so well conserved? One likely explanation is that mutations that alter the position and / or orientation of genes within CTDGs can be deleterious. In classic examples, like the Hox and beta-globin CTDGs, regulatory elements close to the CTDGs have been identified to be important in the expression of CTDG genes, and importantly, mutations that change the position or orientation of genes in a CTDG can have phenotypic consequences [33, 147].

The gibbon genome showed an interesting pattern supporting the hypothesis of CTDG conservation: if CTDGs were not conserved, the gibbon genome would have significantly more CTDGs of smaller size, because its history of drastic genome rearrangements [143] would have broken existing CTDGs randomly, increasing their number but decreasing their size (in number of genes. However, we found that the gibbon genome has mean fractions of clustered regions and genes across chromosomes similar to those of the rest of the primates (fig. 3.5). This suggests that the higher rate of genomic rearrangements in the gibbon genome did not significantly affect the length and composition of its CTDGs.

A model in which CTDG maintenance is necessary for conservation of gene regulation could explain how, because CTDGs can function as “regulatory domains” of the genes they harbor, CTDGs are maintained across evolutionary time. Similarly, large sets of co-regulated duplicated genes in close proximity would optimize their concomitant expression, because that arrangement would minimize the number of chromatin relaxation events needed for the transcription machinery to access those genes. For example, it has been shown that chromatin interactions via CTCF binding sites are necessary for the co-expression of the gene duplicates in the fibrinogen cluster [148].

Several additional models, not necessarily at odds with one another, could explain how CTDGs could be maintained. Right after gene duplication, it is reasonable to assume that the regulatory programs of any of the duplicates in the CTDG would activate most of the duplicates. If the increased expression of the genes in the CTDG is selected for, then a limited number of regulatory elements would be necessary. From that it follows that the regulatory elements of most duplicates in a CTDG will tend to degrade (or not being duplicated along with the gene in the first place). At that point, the CTDG would be maintained because most of the duplicates would be disregulated without the few remaining regulatory elements left. At the other end of the functional spectrum, we can imagine how CTDGs can remain duplicate-dense to increase the combinatorial potential of the CTDG. For example, the protocadherin CTDG harbors exons that can be independently regulated. In that way, many different protocadherin isoforms can be produced, increasing their functional potential as drivers of neuronal identity specification [37, 149]. Also, different combinations of Hox genes are expressed at different time points in the development of vertebrates, depending on the regulatory programs that are required [3]. Under this model, genes in a CTDG would be maintained in close proximity to facilitate the expression of selected but different genes in the CTDG, increasing their functional repertoire.

Gene family dynamics (gene gains and gene losses) have been of great interest in evolutionary genomics research [15, 150]. Assuming neutral evolution [13], gene families rich in CTDGs should have a higher rate of gene duplication than gene families depleted in CTDGs because the probability of gene duplication for a gene family would be directly correlated with the number of existing gene duplicates. However, our findings suggest that duplication events have generally remained stable across gene families, regardless of the number of clustered genes in each gene family. Overall, our analysis shows that most of the gene gains per family occurred prior to the origin of mammals, which aligns with previous reports supporting that only a minority of mammal-specific homologous gene groups experienced gene duplication events in the mammalian phylogeny [151]). These data, along with our experiment on the conservation of CTDGs in chromosome 19, are consistent with a model by which gene families greatly expanded before mammals diversified, and then stabilized in gene number. If this model is supported by additional data, it would call our attention to include parameters about the "genomic geography" of CTDGs to build a model about the evolutionary mechanisms behind gene family evolution in particular, and genome structure in general.

3.6 Acknowledgements

This work was conducted in part using the resources of the Advanced Computing Center for Research and Education at Vanderbilt University. This work was supported in part by the March of Dimes through the March of Dimes Prematurity Research Center Ohio Collaborative, The Burroughs Wellcome Trust Preterm Birth Initiative, the National Science Foundation (DEB-1442113 to A.R.), and the Guggenheim Foundation.

Chapter 4

Spatial, evolutionary, and functional order conservation in the mammalian Keratin cluster. Implications for the study of complex homology in clustered gene families

4.1 Abstract

After gene duplication, a decrease in selective pressure is often invoked. However, there are more retained gene duplicates in genomes than what we would expect according to the model of relaxation of selection after gene duplication. A particularly dramatic example of this observation are clusters of tandemly duplicated genes (CTDG). CTDGs can be very gene-rich and dynamic, and they have been observed driving important and polymorphic traits, like body plan architecture, olfaction, and venom composition. Such dynamic nature tend to give the idea that CTDGs are continuosly changing. However, important CTDGs as the β-globin or Hox CTDGs are evidence that CTDGs can be conserved in gene order throughout long evolutionary time-scales. Here, we want to explore the extent at which CTDGs can be conserved in gene order across mammalian genomes. In order to acomplish that, we developed LOLCAT, an algorithm that identifies subsets of CTDGs that are ordered both in the genome as well as in the phylogeny, which we call "sequentially duplicated clusters" (SDC). We find that the Keratin type I CTDG is highly ordered. In horse, for example, all the genes in the cluster follow the order of divergence of its phylogeny. We leverage LOLCAT to propose novel Keratin orthologs in the cat and cow genomes. We also show how LOLCAT can be used to identify an SDC in the cystatin CTDG, in which the divergence pattern of such SDC correspond to the expression of its genes in humans. With this work, we present the Keratin CTDG as an example of how comparatively long CTDGs can be maintained in order over evolutionary time, and how that information can be leveraged to help in the resolution of complex orthology calling problems.

4.2 Introduction

After gene duplication, it is commonly reported in the literature that the selection pressure in at least one of the copies will tend to decrease because only one copy is required to maintain its function. Following that model (proposed by Ohno [11]), the gene copy that is not conserved can accept more mutations without interfeering with its function, increasing the probability of positive selection [90, 152–154], resulting most of the time in the accumulation of degenerative mutations ending in the pseudogenization of one copy, the subfunctionalization of both copies via reciprocally deleterious mutations, or less likely, the generation of novel functions via accumulation of beneficial mutations in one copy (neofunctionalization) [78].

Following that model, we would expect that gene duplicates will tend to be removed from the genome, which will result in a low number of gene duplicates surviving in a genome. However, important gene families like the Hox genes [3], olfaction receptors (OR) [155], β-globins [1], and prolactins (PRL) [85] have not only preserved at least part of their gene duplicates, but their high copy number is essential to their biological function; and in general, the high prevalence of gene duplicates in genomes across the tree of life has been reported extensively [23, 77, 78].

The gene families referenced above and many others [23] are not only rich in gene duplicates, but also exist in one or more clusters of tandemly duplicated genes (CTDG) in mammals. However, the relationship between CTDG dynamics (gene duplicate gains and losses) and its impact on the phylogenetic pattern of divergence of genes in CTDGs remains poorly understood.

We reasoned that if a CTDG expands linearly (i.e. adding one gene at the time at the end of the CTDG), that should be reflected in the phylogenetic tree built from genes in that CTDG as a "ladder-like" tree or clade (fig. 4.1A). We call this kind of CTDGs where the genomic order resembles the phylogenetic order of divergence a "sequentially duplicated cluster" (SDC). It is important to note that SDC can be sub-sets of a CTDG, when for example, only one section of the CTDG was expanded linearly.

It is reasonable to expect a very low number of SDCs in a genome, especially because an SDC would require a very coordinated, and even lower than expected evolutionary rate, compared with other rapidly evolving gene families after several rounds of gene duplication [78]. We hypothesize that evidence of SDC suggests that the generation or expansion of CTDGs can lower the evolutionary rate of the newly duplicated genes, securing their place in both the genome and their phylogeny. We call this phenomenon "genomic locking". Although it is not a very common occurrence, a decrease in evolutionary rate after CTDG amplification has been reported in the prolactin/growth hormone (Prl/GH)

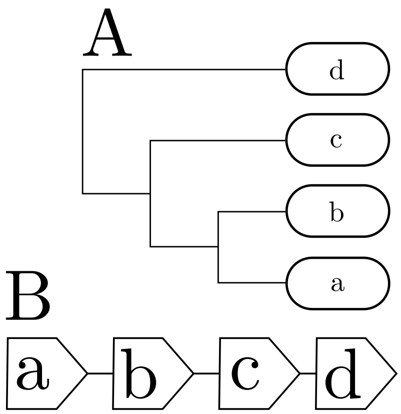

Figure 4.1: Schematic of an SDC. Phylogenetic (A) and genomic organization (B) patterns expected in an ideal SDC

CTDG, where it was hypothesized that the evolutionary rate in the expansion of Prl/GH increased because of changes in the internal or external environments in primates, rodents, and artyodactyls, and then, after the gene duplicated and changed to meet the new environment, its evolutionary rate slowed down [156].

Although this hypothesis remains to be thoroughly tested, the observation of a decrease in evolutionary rate after a CTDG expansion remains an important data point. Other important and well-studied CTDG, like the Hox and β-globin CTDGs have shown how gene order has to be maintained to avoid deleterious phenotypes [34, 147]. All these examples considered, we sought to find cases of SDCs across the mammalian CTDG repertoire, so that interesting hypothesis of how CTDGs evolve and are maintained could be tested.

In order to identify SDCs, we developed the algorithm "Ladder-like Ordered Linear Cluster Analysis for Trees" (LOLCAT). Briefly, LOLCAT flags clades of a CTDG-specific phylogenetic tree as sequentially duplicated clusters (SDC) if the phyletic order of the genes in the CTDG is correlated with the genomic organization of those genes (see methods).

Here, we report the prevalence of SDCs in the Type I Keratin CTDG (Keratin CTDG) across mammalian genomes, and the potential application of SDC analysis in cases of high CTDG dynamics where orthology calling is complicated by extensive paralogy, which is problematic when using syntenic and phylogenetic methods alone.

4.3 Methods

4.3.1 Data acquisition

The genomes of selected mammals (*Bos taurus*; cow, *Chlorocebus sabaeus*; green monkey, *Equus caballus*; horse, *Felis catus*; cat, *Homo sapiens*; human, *Macaca mulatta*; macaque, *Monodelphis domestica*; opossum, *Mus musculus*; mouse, *Nomascus leucogenys*; gibbon, *Rattus norvegicus*; rat) were downloaded from the REST Ensembl service (http://rest.ensembl.org; date of access: August 10, 2018). Chromosomes (or scaffolds) with less than 10 genes or shorter than 10 kbp were not included in the analysis. Assembly versions can be found in table B.47.

4.3.2 CTDG identification

CTDGs were identified using CTDGFinder with 1,000 replicates [23], following the procedure from chapter 3.

4.3.3 Phylogenetic analysis

In order to perform phylogenetic reconstructions using only homologous genes, subclusters were extracted from the collapsed super-clusters, so that all the genes in each subcluster have at least one gene family in common. Protein alignments were performed for the genes in each cluster using MAFFT [157], and Pal2nal [158] was used to convert protein alignments to codon alignments. Phylogenetic trees for the codon alignments of each cluster were estimated using Raxml [93] with 100 bootstrap replicates and using the GTR+G substitution model with the arguments “-# 100 -f a -m GTRGAMMA”.

4.3.4 Identifying sequentially-duplicated clusters (SDC): The LOLCAT algorithm

A strict SDC is defined as a cluster (or sub-cluster) where the gene at one end of a cluster is also the earliest diverged gene in a clade of a phylogenetic tree built from the clustered genes in that cluster (cluster tree), and further genes in the cluster are closer to the tip of the clade, generating a high correlation between the gene order distances (number of genes between two genes) and phyletic distances (number of nodes between two genes in a tree) of all possible pairs of different genes in a cluster.

In order to identify SDC, we took all clusters with four or more clustered genes for each species, estimated their phylogenetic trees, and calculated the Kendall tau [159] correlation on each node for each maximum likelihood cluster tree (focal node) and on the nodes of all bootstrap replicates that contain the same genes as those in the focal node (bootstrap tau distribution) (see below). A focal node was identified as a SDC if they passed all the following tests:

- Filter 1: The p-value of the tau correlation on the maximum likelihood tree is less than 0.05.
- Filter 2: More than half of the bootstrap replicates have significant tau correlation values.
- Filter 3: The tau correlation on the maximum likelihood tree (ML tree) has a value that is different to the mean of the bootstrap replicates on less than the standard deviation of the bootstrap replicates tau.
- Filter 4: The tau correlation of the "perfect" ladder tree has a p-value less than 0.05.
- Filter 5: The tau correlation on the ML tree is closer to the tau value of the perfect tree than to the mean of 100 tau values calculated on random trees.
- Filter 6: The tau value on the ML tree is greater than the percentile 95 of the tau distribution on 100 random trees.

4.3.5 Tau-based phylogenetic-genomic correlation

In order to calculate the correlation between genomic organization and phylogenetic divergence between genes A and B, we counted the number of genes between A and B (genomic distance), and the number of node between nodes A and B in the CTDG-specific tree (tree distance). Because CTDGs usually undergo genomic rearrangements, we calculated all vs all pairwise genomic distances and tree distances for all the genes in the CTDG. In this way, a more biologically relevant environment of the CTDG is taken into account.

4.4 Results and discussion

LOLCAT identified clades of the Keratin CTDG in several mammalian genomes (fig. 4.2 red branches). The most striking example was seen in the horse Keratin CTDG, where the divergence pattern of the genes in the cluster linearly reflect their order in the genome. However, the cow Keratin CTDG called our attention because the SDC identified by LOLCAT was inverted with respect to the general organization of the CTDG. The Keratin genes that are part of the SDC we report here (KRT42, KRT17, KRT16, KRT14) are Keratins related to the basal epidermis and hyperkeratosis (K-SDC) that are also conserved in terrestrial and aquatic mammals [160]. The only exception is KRT16, which is not reported in the study, but it exists in Ensembl as protein-coding gene, and was identified by us as part of the CTDG [160].

Interestingly, the orientation of transcription is the same as the rest of the Keratins in the CTDG, which suggests that the anticorrelation of their divergence and genomic order could be caused by evolutionary mechanisms. In support of that hypothesis, a similar behavior was observed in the organization of those genes (or their orthologs) in cat, rat, green monkey, and macaque (fig. A.5). With these data, we show how LOLCAT was able to identify a clade of Keratins with conserved function (the K-SDC) among a larger set of Keratins (the Keratin CTDG). Furthermore, we show how the SDC contains genes that are genomically and functionally conserved across mammals with significant diverse life histories (e.g. human, cow, dolphins, and whales) [160], supporting the hypothesis of orthology of the genes in that SDC.

Another example to illustrate how we can resolve situations of complex homology in clustered genes is the relationship between KRT31 and KRT34, and KRT32 and KRT35. In many CTDGs, several rounds of gene expansions and contractions make the use of synteny by itself a problematic tool to define orthology relationships, because it is not clear which genes in a CTDG should be aligned, especially with highly dynamic CTDGs. In fig. 4.2 we can see how several gene sub-families are inferred (colored lines) by a sequence similarity method. By looking at the sequential divergence in the same species and the position of neighboring genes, the following pattern is noted: KRT35 and KRT32 are in sister clades, and KRT32 is next to KRT35 in the genome. A similar pattern can be observed for the KRT32 and KRT35 across the Keratin CTDG in several species (fig. A.5). It is important to

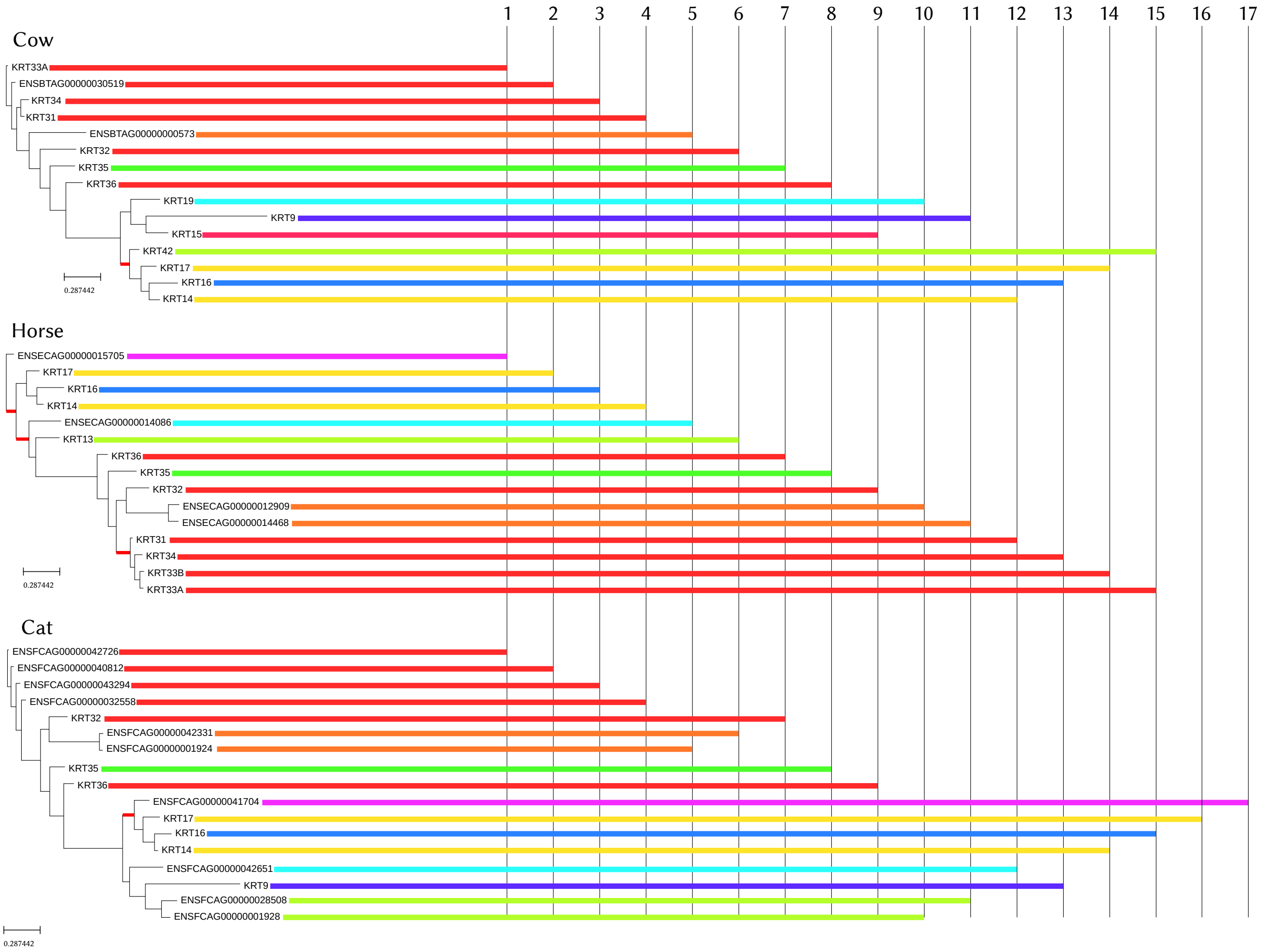

Figure 4.2: High correlation between gene order and phylogenetic divergence of the genes in the Keratin CTDG in the cow, cat, and horse genomes. The phylogenetic trees are rooted on the first gene in the CTDG, so that divergence steps can be appreciated in the vertical axis. Colored lines represent gene sub-families from a one-to-one orthology network based on OMA [135]. Each line connects a gene to its position in the CTDG. Red branches in the phylogenetic tree represent SDCs with high statistical significance

highlight that although these kinds of patterns can be observed manually by looking at the synteny and a gene family-level gene tree, with LOLCAT, this process can be automated.

Using gene family trees to identify orthologs usually involves finding clades in the phylogeny where the genes in that clade recapitulate the overall topology of the species tree where they come from. However, although this method can account for gene duplicates, its reliability suffers when sequence similarity is high, especially within species, and more so if recurrent gene conversion has taken place. In a similar manner, sequence similarity based methods of orthology calling suffer from high sequence similarity because it would call many-to-many orthologous relationships between all the gene duplicates that belong to all the species represented in one cluster of sequences, leaving pair-wise orthology relationships unclear.

Looking at the phylogenetic tree of all the Keratin CTDG in this study (fig. A.6, specifically the clade in fig. 4.3), we can see how LOLCAT can aid orthology calling procedures. As expected with a gene duplicate-rich phylogeny, the branch lengths are short, and support values in many sub-terminal branches are low (fig. 4.3). Moreover, the branching pattern on the tree suggests a scenario of three independent CTDG expansions. However, from knowledge of how the Keratin CTDG is organized, as well as the functions of the genes in it [160], it is unlikely that a subset of a CTDG would expand three times maintaining the relative positions of all the independently duplicated genes and the flanking regions of the duplication across species. Although more data is needed to generalize this hypothesis, the high conservation of gene order in all three species suggests that a CTDG expansion occurred before the divergence of horses, cows and cats, followed by selection acting in maintaining the order of the duplicated genes. If on the other hand, a sequence similarity method was used, all the genes in our focal clade (fig. 4.3) would have been identified as belonging to the same orthology group (red bars in fig. 4.2). Importantly, a gene-family level phylogeny is still useful in this case. For example, in cases where we were the identification of gene sub-families is not possible, potential orthologs could be found in the same relative order (fig. 4.3, color bars).

By using the SDCs identified by LOLCAT, we can see a clearer picture of potential novel orthology calls. For example, observing that the SDC containing KRT31, KRT34, KRT33B, and KRT33A in horse, we can observe that genes in the same orthology group (using sequence similarity) follow the same order in cow (i.e. in horse: KRT31 follows KRT34, KRT33B, and KRT33A, and in cow: KRT31 is followed by KRT34, ENSBTAG00000030519, and KRT33A), a pattern that extend all the way to KRT36 in both cow and horse. In this scenario, we suggest that ENSBTAG00000030519 is the ortholog of KRT33B. Similarly, we can follow the orthologs of the cat Keratin CTDG, and identify ENSFCAG00000042726, ENSFCAG00000040812, ENSFCAG00000043294, and ENSFCAG00000032558, which we suggest, are orthologs of KRT31, KRT34, KRT33B, and KRT33A respectively. It is very important to keep in mind that we are assuming that the orientation of transcription of these genes has not changed during the evolution of the CTDG, as we see its genes in the same orientation in the species studied here.

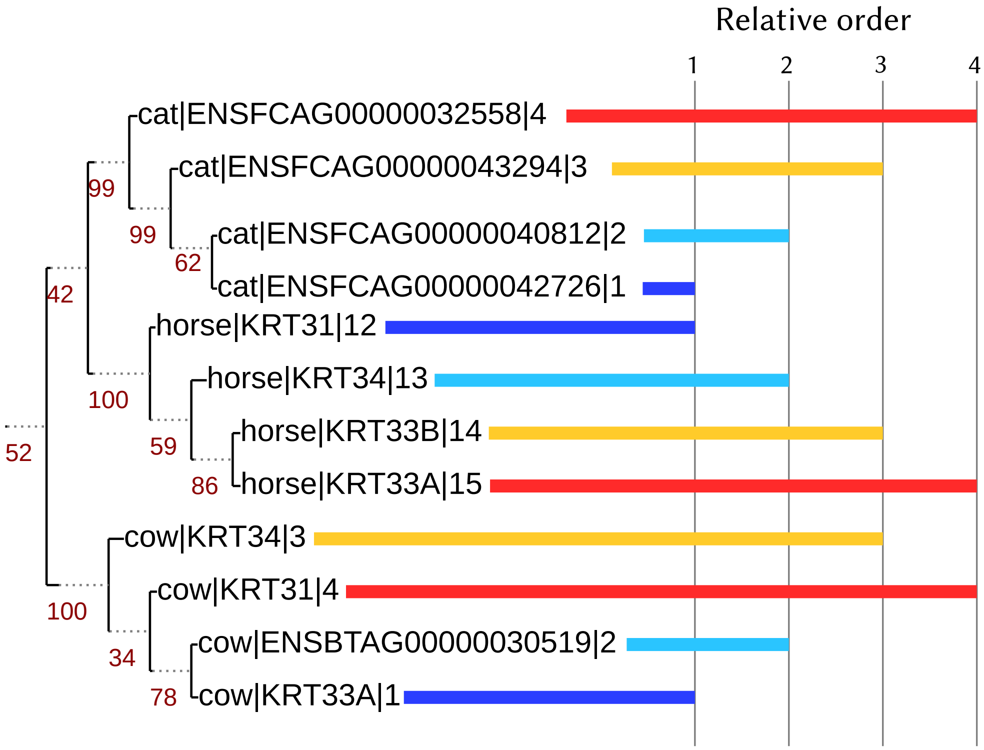

Figure 4.3: Clade from gene family-level phylogenetic tree containing all the Keratin CTDGs that contain orthologs of a horse SDC in cow and cat. Vertical lines: relative order of the genes with respect to one another. Color bars: equivalent relative position of a gene in their corresponding CTDG

The technique we described above can be summarized as follow:

1. Use a sequence-based orthology-calling algorithm to find gene families and subfamilies.

2. Identify CTDGs using CTDGFinder.

3. Build phylogenetic trees of each species-specific CTDG.

4. Run LOLCAT to dentify SDCs in all the CTDG-specific trees.

5. Analyze the relative positions of sequence-based orthologs, paying special attention to relative positions (i.e. in a species, genes A, B, and C in positions 2, 3, and 4 should be analyzed if sequence-based orthologs are in positions 10, 11, and 12) and look for potential orthologs.

6. If the sequence-based orthology calling algorithm was not able to resolve more than one gene family, build a phylogenetic tree with all the genes in the focal CTDG from all species under study, and select the clade containing the SDC identified by LOLCAT.

7. Call orthologs that share homology at the sequence, phylogenetic, and now, shared gene order level.

Another interesting example is the cystatin type II CTDG. Although it is a very ancient gene family, cystatins have expanded several times across multiple animal lineages [161]. Moreover, its origin can be traced back to ancestral lineages of eukaryotes [162]. In vertebrates, cystatins have evolved from their inhibitory activity to be involved in innate immunity [162]. Out of the eight members of the cystatin CTDG, an SDC of four genes were identified by LOLCAT (fig. 4.4). Interestingly, the cystatins in the SDC cluster according to their expression patterns (according to the GTEx database), when compared with all the cystatins in the CTDG (fig. A.7).

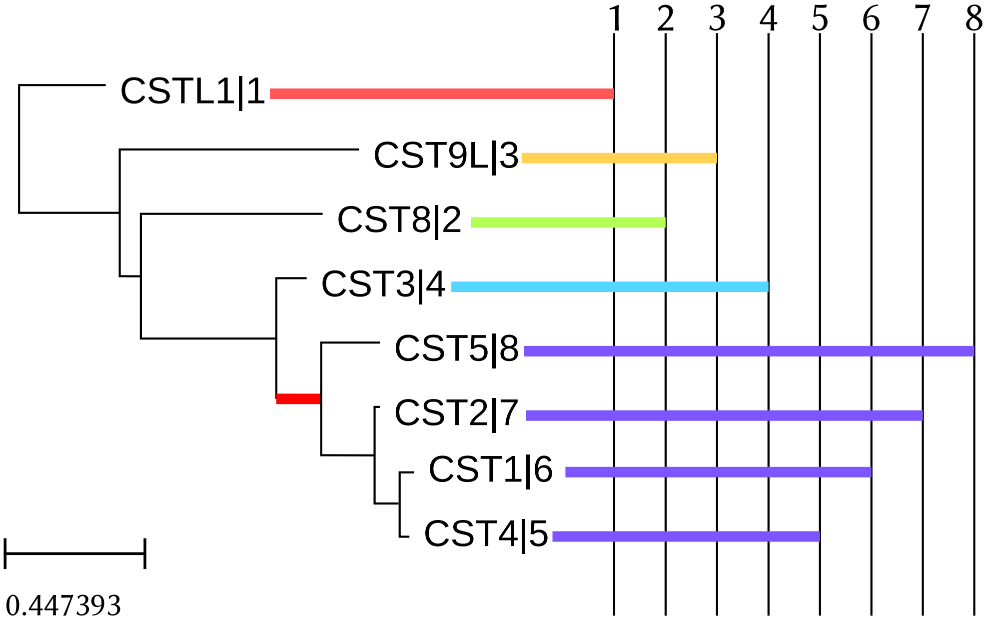

Figure 4.4: Human cystatin CTDG. Legend follows fig. 4.2

Defining homology relationships is challenging in CTDGs because of the high similarity between sets of genes of the same gene family in the same species. Although SDCs are very rare, here we report the striking case of the Keratin CTDG SDC across mammals. We also show how by combining the analysis of gene order in time (phylogenetically) and space (genomically) we can strengthen orthology hypotheses between genes. Given that gene order conservation has been observed to have functional consequences in important phylogenetically conserved CTDGs like the β-globin [34] and Hox CTDGs [147], with this report we show how by using genomic and phylogenetic information simultaneously, the spatial conservation of important orthologous genes in the Keratin CTDG were identified as a proof of concept of a tool that can aid in the identification of orthologous relationships in complex gene families.

Finally, we present LOLCAT, a valuable tool in aiding the often difficult task of orthology calling. Importantly, LOLCAT is best suited to be used in CTDGs with relatively ordered genes. Also, it has to be noted that this method is insufficient to identify orthologs by itself, and it should be used in conjunction with other methods to leverage its potential.

Chapter 5

Conclusions

CTDGs can be defined, and more importantly, computationally identified (chapter 2). With such a definition, evolutionarily and functionally interesting features of CTDGs can be uncovered by systematically applying it to the study of CTDGs in mammalian genomes. CTDGs are integral parts of mammalian genomes (chapter 3[23]), and evidence exist strongly suggesting that the same is true for other groups [66, 80, 163, 164].

CTDGs are not only common among mammals, but also maintained as such. Genomic contraction is likely to be at play in the evolutionary processes that have preserved the mammalian repertoire of CTDGs, and future research in that direction is bound to be exciting. Given that CTDGs are indeed biological entities likely subject to evolutionary forces, we can use characteristics of CTDGs as part of our toolkit for describing and discovering the intricate processes of the evolution of genomes and gene families (chapter 4).

There is still a lot that we do not know, and many new avenues of research have opened. For example, now that we can describe linear CTDGs (in 1D along the length of the genome), what about three dimensional CTDGs? We know that chromatin conformation is an active source of genic regulation, and with the tools now at our disposal, we could start describing how CTDGs organize in three dimensional space, and maybe more interestingly, how gene duplicates that are not close along the genome, can come together to form transient but functionally important chromatin-level CTDGs. We now can also start to think again the level at which our assumptions of orthology have to go (or can go). By systematically adding a new level of biological organization (the CTDG), we can now incorporate it to our protocols for modeling the evolution of gene families.

Finally, with tools like CTDGFinder and LOLCAT, we can start developing more sophisticated algorithms to model how the genome as a whole is organized, potentially informing robust statistical models that incorporate not only gene family descriptions, but also the parameters that describe their organization in the genome.

Bibliography

[1] Argiris Efstratiadis, James W. Posakony, Tom Maniatis, Richard M. Lawn, Catherine O'Connell, Richard A. Spritz, Jon K. Deriel, Bernard G. Forget, Sherman M. Weissman, Jerry L. Slightom, Ann E. Blechl, Oliver Smithies, Francisco E. Baralle, Carol C. Shoulders, and Nicholas J. Proudfoot. The structure and evolution of the human β-globin gene family. *Cell*, 21(3):653–668, October 1980.

[2] Sonoko Hasegawa, Shun Hamada, You Kumode, Shigeyuki Esumi, Shota Katori, Emi Fukuda, Yasuo Uchiyama, Takahiro Hirabayashi, Peter Mombaerts, and Takeshi Yagi. The protocadherin-α family is involved in axonal coalescence of olfactory sensory neurons into glomeruli of the olfactory bulb in mouse. *Molecular and Cellular Neuroscience*, 38(1):66–79, May 2008.

[3] Peter W.H. Holland and Jordi Garcia-Fernández. HoxGenes and Chordate Evolution. *Developmental Biology*, 173(2):382–395, February 1996.

[4] A. M. Martin, E. M. Freitas, C. S. Witt, and F. T. Christiansen. The genomic organization and evolution of the natural killer immunoglobulin-like receptor (KIR) gene cluster. *Immunogenetics*, 51(4-5):268–280, April 2000.

[5] Meredith Yeager and Austin L. Hughes. Evolution of the mammalian MHC: Natural selection, recombination, and convergent evolution. *Immunological Reviews*, 167(1):45–58, April 2006.

[6] Martin Malmstrøm, Michael Matschiner, Ole K. Tørresen, Bastiaan Star, Lars G. Snipen, Thomas F. Hansen, Helle T. Baalsrud, Alexander J. Nederbragt, Reinhold Hanel, Walter Salzburger, Nils C. Stenseth, Kjetill S. Jakobsen, and Sissel Jentoft. Evolution of the immune system influences speciation rates in teleost fishes. *Nature Genetics*, 48(10):1204–1210, October 2016.

[7] Naoki Ikeda, Takahito Chijiwa, Kazumi Matsubara, Naoko Oda-Ueda, Shosaku Hattori, Yoichi Matsuda, and Motonori Ohno. Unique structural characteristics and evolution of a cluster of

venom phospholipase A2 isozyme genes of Protobothrops flavoviridis snake. *Gene*, 461(1-2):15–25, August 2010.

[8] C. M. Whittington, A. T. Papenfuss, P. Bansal, A. M. Torres, E. S.W. Wong, J. E. Deakin, T. Graves, A. Alsop, K. Schatzkamer, C. Kremitzki, C. P. Ponting, P. Temple-Smith, W. C. Warren, P. W. Kuchel, and K. Belov. Defensins and the convergent evolution of platypus and reptile venom genes. *Genome Research*, 18(6):986–994, May 2008.

[9] Yoshihito Niimura, Atsushi Matsui, and Kazushige Touhara. Extreme expansion of the olfactory receptor gene repertoire in African elephants and evolutionary dynamics of orthologous gene groups in 13 placental mammals. *Genome Research*, 24(9):1485–1496, September 2014.

[10] John S. Taylor and Jeroen Raes. Duplication and Divergence: The Evolution of New Genes and Old Ideas. *Annual Review of Genetics*, 38(1):615–643, November 2004.

[11] Susumu Ohno, Ulrich Wolf, and Niels B. Atkin. Evolution from fish to mammals by gene duplication. *Hereditas*, 59(6):169–187, 1967.

[12] Klaus Rothfels, Elizabeth Sexsmith, Margaret Heimburger, and Margarida O. Krause. Chromosome size and DNA content of species of anemone L. and related genera (Ranunculaceae). *Chromosoma*, 20(1):54–74, March 1966.

[13] M. Lynch. The Evolutionary Fate and Consequences of Duplicate Genes. *Science*, 290(5494):1151–1155, November 2000.

[14] Allan Force, Michael Lynch, F Bryan Pickett, Angel Amores, Yi-lin Yan, and John Postlethwait. Preservation of Duplicate Genes by Complementary, Degenerative Mutations. *Genetics*, 151:1531–1545, April 1999.

[15] Jeffery P. Demuth and Matthew W. Hahn. The life and death of gene families. *BioEssays*, 31(1):29–39, January 2009.

[16] Andrew B. Reams and John R. Roth. Mechanisms of Gene Duplication and Amplification. *Cold Spring Harbor Perspectives in Biology*, 7(2):a016592, February 2015.

[17] Evan E. Eichler and David Sankoff. Structural Dynamics of Eukaryotic Chromosome Evolution. *Science*, 301(5634):793–797, August 2003.

[18] A.J.F. Griffiths, W.M. Gelbart, J.H. Miller, and Richard C. Lewontin. *Modern Genetic Analysis*. Freeman, New York, 1999.

[19] J. A. Birchler and R. A. Veitia. Gene balance hypothesis: Connecting issues of dosage sensitivity across biological disciplines. *Proceedings of the National Academy of Sciences*, 109(37):14746–14753, September 2012.

[20] J. P. Noonan. Gene Conversion and the Evolution of Protocadherin Gene Cluster Diversity. *Genome Research*, 14(3):354–366, March 2004.

[21] Yu Zhang, Giltae Song, Tomáš Vinař, Eric D. Green, Adam Siepel, and Webb Miller. Reconstructing the Evolutionary History of Complex Human Gene Clusters. In Martin Vingron and Limsoon Wong, editors, *Research in Computational Molecular Biology*, Lecture Notes in Computer Science, pages 29–49. Springer Berlin Heidelberg, 2008.

[22] Guillaume Bourque and Pavel A Pevzner. Genome-Scale Evolution: Reconstructing Gene Orders in the Ancestral Species. *Genome Research*, 12:26–36, 2002.

[23] Juan F. Ortiz and Antonis Rokas. CTDGFinder: A Novel Homology-Based Algorithm for Identifying Closely Spaced Clusters of Tandemly Duplicated Genes. *Molecular Biology and Evolution*, 34(1):215–229, January 2017.

[24] Jordi Garcia-Fernàndez. The genesis and evolution of homeobox gene clusters. *Nature Reviews Genetics*, 6(12):881–892, December 2005.

[25] Tetsuya Nakamura, Jeff Klomp, Joyce Pieretti, Igor Schneider, Andrew R. Gehrke, and Neil H. Shubin. Molecular mechanisms underlying the exceptional adaptations of batoid fins. *Proceedings of the National Academy of Sciences*, 112(52):15940–15945, December 2015.

[26] Richard R Copley. The animal in the genome: Comparative genomics and evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1496):1453–1461, April 2008.

[27] Antonietta Spagnuolo, Filomena Ristoratore, Anna Di Gregorio, Francesco Aniello, Margherita Branno, and Roberto Di Lauro. Unusual number and genomic organization of Hox genes in the tunicate Ciona intestinalis. *Gene*, 309(2):71–79, May 2003.

[28] N. Di-Poi, J. I. Montoya-Burgos, and D. Duboule. Atypical relaxation of structural constraints in Hox gene clusters of the green anole lizard. *Genome Research*, 19(4):602–610, April 2009.

[29] R. Andrew Cameron, Lee Rowen, Ryan Nesbitt, Scott Bloom, Jonathan P. Rast, Kevin Berney, Cesar Arenas-Mena, Pedro Martinez, Susan Lucas, Paul M. Richardson, Eric H. Davidson, Kevin J. Peterson, and Leroy Hood. Unusual gene order and organization of the sea urchin hox cluster. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 306B(1):45–58, January 2006.

[30] D. Duboule. The rise and fall of Hox gene clusters. *Development*, 134(14):2549–2560, June 2007.

[31] J. R. Mann, A. S. MacNeish, D. Bannister, J. B. Clegg, W. G. Wood, and D. J. Weatherall. $\Delta\beta$-Thalassaemia in a Chinese Family. *British Journal of Haematology*, 23(4):393–402, 1972.

[32] A G Kendall, P J Ojwang, W A Schroeder, and T H Huisman. Hemoglobin Kenya, the product of a gamma-beta fusion gene: Studies of the family. *American Journal of Human Genetics*, 25(5):548–563, September 1973.

[33] M. Bulger, J. H. von Doorninck, N. Saitoh, A. Telling, C. Farrell, M. A. Bender, G. Felsenfeld, R. Axel, and M. Groudine. Conservation of sequence and structure flanking the mouse and human -globin loci: The -globin genes are embedded within an array of odorant receptor genes. *Proceedings of the National Academy of Sciences*, 96(9):5129–5134, April 1999.

[34] Padraic P. Levings and Jörg Bungert. The human β-globin locus control region: A center of attraction. *European Journal of Biochemistry*, 269(6):1589–1599, March 2002.

[35] Kerry M. Goodman, Rotem Rubinstein, Hanbin Dan, Fabiana Bahna, Seetha Mannepalli, Göran Ahlsén, Chan Aye Thu, Rosemary V. Sampogna, Tom Maniatis, Barry Honig, and Lawrence

Shapiro. Protocadherin cis-dimer architecture and recognition unit diversity. *Proceedings of the National Academy of Sciences*, 114(46):E9829–E9837, November 2017.

[36] Takeshi Yagi. Clustered protocadherin family: Clustered protocadherin family. *Development, Growth & Differentiation*, 50:S131–S140, April 2008.

[37] Shigeyuki Esumi, Naoki Kakazu, Yusuke Taguchi, Teruyoshi Hirayama, Ayako Sasaki, Takahiro Hirabayashi, Tsuyoshi Koide, Takashi Kitsukawa, Shun Hamada, and Takeshi Yagi. Monoallelic yet combinatorial expression of variable exons of the protocadherin-α gene cluster in single neurons. *Nature Genetics*, 37(2):171–176, February 2005.

[38] Raja Rajalingam. Diversity of Killer Cell Immunoglobulin-Like Receptors and Disease. *Clinics in Laboratory Medicine*, 38(4):637–653, December 2018.

[39] Yoav Gilad, Victor Wiebe, Molly Przeworski, Doron Lancet, and Svante Pääbo. Loss of Olfactory Receptor Genes Coincides with the Acquisition of Full Trichromatic Vision in Primates. *PLOS Biology*, 2(1):e5, January 2004.

[40] Yoshihito Niimura. Evolutionary dynamics of olfactory receptor genes in chordates: Interaction between environments and genomic contents. *Human Genomics*, 4(2):107, 2009.

[41] Atsushi Matsui, Yasuhiro Go, and Yoshihito Niimura. Degeneration of Olfactory Receptor Gene Repertories in Primates: No Direct Link to Full Trichromatic Vision. *Molecular Biology and Evolution*, 27(5):1192–1200, May 2010.

[42] Gustavo Glusman, Alona Sosinsky, Edna Ben-Asher, Nili Avidan, Dina Sonkin, Anita Bahar, André Rosenthal, Sandra Clifton, Bruce Roe, Concepción Ferraz, Jacques Demaille, and Doron Lancet. Sequence, Structure, and Evolution of a Complete Human Olfactory Receptor Gene Cluster. *Genomics*, 63(2):227–245, January 2000.

[43] Harumi Saito, Qiuyi Chi, Hanyi Zhuang, Hiroaki Matsunami, and Joel D. Mainland. Odor Coding by a Mammalian Receptor Repertoire. *Sci. Signal.*, 2(60):ra9–ra9, March 2009.

[44] F. J. Vonk, N. R. Casewell, C. V. Henkel, A. M. Heimberg, H. J. Jansen, R. J. R. McCleary, H. M. E. Kerkkamp, R. A. Vos, I. Guerreiro, J. J. Calvete, W. Wuster, A. E. Woods, J. M.

Logan, R. A. Harrison, T. A. Castoe, A. P. J. de Koning, D. D. Pollock, M. Yandell, D. Calderon, C. Renjifo, R. B. Currier, D. Salgado, D. Pla, L. Sanz, A. S. Hyder, J. M. C. Ribeiro, J. W. Arntzen, G. E. E. J. M. van den Thillart, M. Boetzer, W. Pirovano, R. P. Dirks, H. P. Spaink, D. Duboule, E. McGlinn, R. M. Kini, and M. K. Richardson. The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. *Proceedings of the National Academy of Sciences*, 110(51):20651–20656, December 2013.

[45] Langyu Gu and Canwei Xia. Cluster expansion of apolipoprotein D (ApoD) genes in teleost fishes. *BMC Evolutionary Biology*, 19(1), December 2019.

[46] Christophe Eizaguirre, Tobias L. Lenz, Martin Kalbe, and Manfred Milinski. Divergent selection on locally adapted major histocompatibility complex immune genes experimentally proven in the field. *Ecology Letters*, 15(7):723–731, 2012.

[47] Paul Jay, Annabel Whibley, Lise Frézal, María Ángeles Rodríguez de Cara, Reuben W. Nowell, James Mallet, Kanchon K. Dasmahapatra, and Mathieu Joron. Supergene Evolution Triggered by the Introgression of a Chromosomal Inversion. *Current Biology*, 28(11):1839–1845.e3, June 2018.

[48] Christine Bole-Feysot, Vincent Goffin, Marc Edery, Nadine Binart, and Paul A Kelly. Prolactin (PRL) and Its Receptor: Actions, Signal Transduction Pathways and Phenotypes Observed in PRL Receptor Knockout Mice. 19(3):44, 1998.

[49] Dustin O. Wiemers, Long-jiang Shao, Rupasri Ain, Guoli Dai, and Michael J. Soares. The Mouse Prolactin Gene Family Locus. *Endocrinology*, 144(1):313–325, January 2003.

[50] David E. K. Ferrier and Peter W. H. Holland. Ancient origin of the Hox gene cluster. *Nature Reviews Genetics*, 2(1):33–38, January 2001.

[51] S. C. Hardies, M. H. Edgell, and C. A. Hutchison. Evolution of the mammalian beta-globin gene cluster. *Journal of Biological Chemistry*, 259(6):3748–3756, March 1984.

[52] Isabel A Forsyth and Michael Wallis. Growth Hormone and Prolactin—Molecular and Functional Evolution. *Journal of Mammary Gland Biology and Neoplasia*, page 22, 2003.

[53] Laurence Kedes and Max Birnstiel. Reiteration and Clustering of DNA Sequences Complementary to Histone Messenger RNA. *Nature New Biology*, 230(14):165–169, April 1971.

[54] R. A. Flavell, J. M. Kooter, E. De Boer, P. F. R. Little, and R. Williamson. Analysis of the β-δ-globin gene loci in normal and hb lepore DNA: Direct determination of gene linkage and intergene distance. *Cell*, 15(1):25–41, September 1978.

[55] Corrado Baglioni. The fusion of two peptide chains in hemoglobin lepor and its interpretation as a genetic deletion. *Proceedings of the National Academy of Sciences of the United States of America*, 48(11):1880–1886, November 1962.

[56] Walter L. Miller and Norman L. Eberhardt. Structure and Evolution of the Growth Hormone Gene Family. *Endocrine Reviews*, 4(2):97–130, April 1983.

[57] V. J. Kidd and G. F. Saunders. Linkage arrangement of human placental lactogen and growth hormone genes. *Journal of Biological Chemistry*, 257(18):10673–10680, September 1982.

[58] L. A. Leinwand, L. Saez, E. McNally, and B. Nadal-Ginard. Isolation and characterization of human myosin heavy chain genes. *Proceedings of the National Academy of Sciences*, 80(12):3716–3720, June 1983.

[59] H. Qin, J. Kemp, M.-Y. Yip, P. R. L. Lam-Po-Tang, J. F. Y. Hoh, and B. J. Morris. Localization of human cardiac β-myosin heavy chain gene (MYH7) to chromosome 14q12 by in situ hybridization. *Cytogenetic and Genome Research*, 54(1-2):74–76, 1990.

[60] Sylvie Rouquier, Sylvie Taviaux, Barbara J. Trask, Véronique Brand-Arpon, Ger van den Engh, Jacques Demaille, and Dominique Giorgi. Distribution of olfactory receptor genes in the human genome. *Nature Genetics*, 18(3):243, March 1998.

[61] Qiang Wu and Tom Maniatis. A Striking Organization of a Large Family of Human Neural Cadherin-like Cell Adhesion Genes. *Cell*, 97(6):779–790, June 1999.

[62] T. Angata, E. H. Margulies, E. D. Green, and A. Varki. Large-scale sequencing of the CD33-related Siglec gene cluster in five mammalian species reveals rapid evolution by multiple mech-

anisms. *Proceedings of the National Academy of Sciences*, 101(36):13251–13256, September 2004.

[63] Huan Cao, Bernard de Bono, Katherine Belov, Emily S. Wong, John Trowsdale, and Alexander David Barrow. Comparative genomics indicates the mammalian CD33rSiglec locus evolved by an ancient large-scale inverse duplication and suggests all Siglecs share a common ancestral region. *Immunogenetics*, 61(5):401–417, May 2009.

[64] K. Kawasaki and K. M. Weiss. Mineralized tissue and vertebrate evolution: The secretory calcium-binding phosphoprotein gene cluster. *Proceedings of the National Academy of Sciences*, 100(7):4060–4065, April 2003.

[65] Giltae Song, Cathy Riemer, Benjamin Dickins, Hie Lim Kim, Louxin Zhang, Yu Zhang, Chih-Hao Hsu, Ross C. Hardison, NISC Comparative Sequencing Program, Eric D. Green, and Webb Miller. Revealing Mammalian Evolutionary Relationships by Comparative Analysis of Gene Clusters. *Genome Biology and Evolution*, 4(4):586–601, 2012.

[66] Steven B Cannon, Arvind Mitra, Andrew Baumgarten, Nevin D Young, and Georgiana May. The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana. *BMC Plant Biology*, page 21, 2004.

[67] T. Ikuta, N. Yoshida, N. Satoh, and H. Saiga. Ciona intestinalis Hox gene cluster: Its dispersed structure and residual colinear expression in development. *Proceedings of the National Academy of Sciences*, 101(42):15118–15123, October 2004.

[68] Thomas J. Y. Kono, Alex B. Brohammer, Suzanne E. McGaugh, and Candice N. Hirsch. Tandem Duplicate Genes in Maize Are Abundant and Date to Two Distinct Periods of Time. *G3: Genes, Genomes, Genetics*, 8(9):3049–3058, September 2018.

[69] Stephan Teglund, Anne Olsen, Wasif Noor Khan, Lars Frångsmyr, and Sten Hammarström. The Pregnancy-Specific Glycoprotein (PSG) Gene Cluster on Human Chromosome 19: Fine Structure of the 11 PSG Genes and Identification of 6 New Genes Forming a Third Subgroup within the Carcinoembryonic Antigen (CEA) Family. *Genomics*, 23(3):669–684, October 1994.

[70] G J Graham. Tandem Genes and Clustered Genes. *Journal of Theoretical Biology*, 175:71–87, 1995.

[71] Evolution and Organization of the Fibrinogen Locus on Chromosome 4: Gene Duplication Accompanied by Transposition and Inversion Author(s): Jeffrey A. Kant, Albert J. Fornace, Debra Saxe, Melvin I. Simon, O. Wesley McBride and Gerald R. Crabtree.

[72] Simon Schiml, Friedrich Fauser, and Holger Puchta. Repair of adjacent single-strand breaks is often accompanied by the formation of tandem sequence duplications in plant genomes. *Proceedings of the National Academy of Sciences*, 113(26):7266–7271, June 2016.

[73] C. Y. Huang, P.-Y. Chen, M.-D. Huang, C.-H. Tsou, W.-N. Jane, and A. H. C. Huang. Tandem oleosin genes in a cluster acquired in Brassicaceae created tapetosomes and conferred additive benefit of pollen vigor. *Proceedings of the National Academy of Sciences*, 110(35):14480–14485, August 2013.

[74] Seongsik Yun, Michael Furlong, Mikang Sim, Minah Cho, Sumi Park, Eun Bee Cho, Arfaxad Reyes-Alcaraz, Jong-Ik Hwang, Jaebum Kim, and Jae Young Seong. Prevertebrate Local Gene Duplication Facilitated Expansion of the Neuropeptide GPCR Superfamily. *Molecular Biology and Evolution*, 32(11):2803–2817, November 2015.

[75] S.M.G. Hoffman, P. Fernandez-Salguero, F.J. Gonzalez, and H.W. Mohrenweiser. Organization and evolution of the cytochrome P450 CYP2A-2B-2F subfamily gene cluster on human chromosome 19. *Journal of Molecular Evolution*, 41(6), December 1995.

[76] Dan I. Andersson and Diarmaid Hughes. Gene Amplification and Adaptive Evolution in Bacteria. *Annual Review of Genetics*, 43(1):167–195, December 2009.

[77] Michael Lynch and John S. Conery. The evolutionary demography of duplicate genes. *Genome Evolution*, 3:35–44, 2003.

[78] Jianzhi Zhang. Evolution by gene duplication: An update. *Trends in Ecology & Evolution*, 18(6):292–298, June 2003.

[79] M. Freeling, E. Lyons, B. Pedersen, M. Alam, R. Ming, and D. Lisch. Many or most genes in Arabidopsis transposed after the origin of the order Brassicales. *Genome Research*, 18(12):1924–1937, October 2008.

[80] David Grant, Perry Cregan, and Randy C. Shoemaker. Genome Organization in Dicots: Genome Duplication in Arabidopsis and Synteny between Soybean and Arabidopsis. *Proceedings of the National Academy of Sciences*, 97(8):4168–4173, 2000.

[81] Lorenzo Carretero-Paulet and Mario A. Fares. Evolutionary Dynamics and Functional Specialization of Plant Paralogs Formed by Whole and Small-Scale Genome Duplications. *Molecular Biology and Evolution*, 29(11):3541–3551, November 2012.

[82] Christopher D. Krause and Sidney Pestka. Cut, copy, move, delete: The study of human interferon genes reveal multiple mechanisms underlying their evolution in amniotes. *Cytokine*, 76(2):480–495, December 2015.

[83] Robb Krumlauf. Evolution of the vertebrateHox homeobox genes. *BioEssays*, 14(4):245–252, April 1992.

[84] James A. MacLean, Diego Lorenzetti, Zhiying Hu, Will J. Salerno, Jonathan Miller, and Miles F. Wilkinson. Rhox homeobox gene cluster: Recent duplication of three family members. *genesis*, 44(3):122–129, March 2006.

[85] S.M. Khorshed Alam, Rupasri Ain, Toshihiro Konno, Jennifer K. Ho-Chen, and Michael J. Soares. The rat prolactin gene family locus: Species-specific gene family expansion. *Mammalian Genome*, 17(8):858–877, August 2006.

[86] Q. Wu. Comparative DNA Sequence Analysis of Mouse and Human Protocadherin Gene Clusters. *Genome Research*, 11(3):389–404, March 2001.

[87] Elissa A. Hallem, Anupama Dahanukar, and John R. Carlson. Insect odor and taste receptors. *Annual Review of Entomology*, 51(1):113–135, January 2006.

[88] James Kelley, Lutz Walter, and John Trowsdale. Comparative Genomics of Natural Killer Cell Receptor Gene Clusters. *PLoS Genetics*, 1(2):e27, 2005.

[89] J. W. Pendleton, B. K. Nagai, M. T. Murtha, and F. H. Ruddle. Expansion of the Hox gene family and the evolution of chordates. *Proceedings of the National Academy of Sciences*, 90(13):6300–6304, July 1993.

[90] Pille Hallast, Janna Saarela, Aarno Palotie, and Maris Laan. High divergence in primate-specific duplicated regions: Human and chimpanzee Chorionic Gonadotropin Beta genes. *BMC Evolutionary Biology*, 8(1):195, 2008.

[91] D. Lemons. Genomic Evolution of Hox Gene Clusters. *Science*, 313(5795):1918–1922, September 2006.

[92] Diego Darriba, Guillermo L. Taboada, Ramón Doallo, and David Posada. ProtTest 3: Fast selection of best-fit models of protein evolution. *Bioinformatics*, 27(8):1164–1165, April 2011.

[93] Alexandros Stamatakis. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 22(21):2688–2690, November 2006.

[94] Travis E. Oliphant. Python for Scientific Computing. *Computing in Science & Engineering*, 9(3):10–20, 2007.

[95] P. J. A. Cock, T. Antao, J. T. Chang, B. A. Chapman, C. J. Cox, A. Dalke, I. Friedberg, T. Hamelryck, F. Kauff, B. Wilczynski, and M. J. L. de Hoon. Biopython: Freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, 25(11):1422–1423, June 2009.

[96] Fabian Pedregosa, Gael Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, and David Cournapeau. Scikit-learn: Machine Learning in Python. *MACHINE LEARNING IN PYTHON*, page 6.

[97] John D. Hunter. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, 9(3):90–95, May 2007.

[98] M. Uhlen, L. Fagerberg, B. M. Hallstrom, C. Lindskog, P. Oksvold, A. Mardinoglu, A. Sivertsson, C. Kampf, E. Sjostedt, A. Asplund, I. Olsson, K. Edlund, E. Lundberg, S. Navani, C. A.-

K. Szigyarto, J. Odeberg, D. Djureinovic, J. O. Takanen, S. Hober, T. Alm, P.-H. Edqvist, H. Berling, H. Tegel, J. Mulder, J. Rockberg, P. Nilsson, J. M. Schwenk, M. Hamsten, K. von Feilitzen, M. Forsberg, L. Persson, F. Johansson, M. Zwahlen, G. von Heijne, J. Nielsen, and F. Ponten. Tissue-based map of the human proteome. *Science*, 347(6220):1260419–1260419, January 2015.

[99] Stephen F AltschuP, Warren Gish, Webb Miller, and Eugene W Myers. Basic Local Alignment Search Tool. *Journal of Molecular Biology*, 215:403–410, 1990.

[100] Dorin Comaniciu. Mean Shift: A Robust Approach Toward Feature Space Analysis. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 24(5):17, 2002.

[101] Yuhua Su, Stephen A. Liebhaber, and Nancy E. Cooke. The Human Growth Hormone Gene Cluster Locus Control Region Supports Position-independent Pituitary- and Placenta-specific Expression in the Transgenic Mouse. *Journal of Biological Chemistry*, 275(11):7902–7909, March 2000.

[102] Naohiro Kohmura, Kouji Senzaki, Shun Hamada, Nobuyuki Kai, Rie Yasuda, Masahiko Watanabe, Hiroki Ishii, Masahiro Yasuda, Masayoshi Mishina, and Takeshi Yagi. Diversity Revealed by a Novel Family of Cadherins Expressed in Neurons at a Synaptic Complex. *Neuron*, 20(6):1137–1151, June 1998.

[103] Robb Krumlauf. Hox genes in vertebrate development. *Cell*, 78(2):191–201, July 1994.

[104] Steven Weaver, Mary B. Comer, Carolyn L. Jahn, Clyde A. Hutchison, and Marshall Hall Edgell. The adult β-globin genes of the “single” type mouse C57BL. *Cell*, 24(2):403–411, May 1981.

[105] L. Zeltser, C. Desplan, and N. Heintz. Hoxb-13: A new Hox gene in a distant region of the HOXB cluster maintains colinearity. *Development*, 122(8):2475–2484, August 1996.

[106] N. G. Than, R. Romero, M. Goodman, A. Weckle, J. Xing, Z. Dong, Y. Xu, F. Tarquini, A. Szilagyi, P. Gal, Z. Hou, A. L. Tarca, C. J. Kim, J.-S. Kim, S. Haidarian, M. Uddin, H. Bohn, K. Benirschke, J. Santolaya-Forgas, L. I. Grossman, O. Erez, S. S. Hassan, P. Zavodszky, Z. Papp, and D. E. Wildman. A primate subfamily of galectins expressed at the maternal-fetal

interface that promote immune cell death. *Proceedings of the National Academy of Sciences*, 106(24):9731–9736, June 2009.

[107] David G Simmons, Saara M Rawn, Alastair Davies, Martha Hughes, and James C Cross. Spatial and temporal expression of the 23 murine Prolactin/Placental Lactogen-related genes is not associated with their position in the locus. *BMC Genomics*, 9(1):352, 2008.

[108] Denis Houzelstein, Isabelle R. Gonçalves, Andrew J. Fadden, Sukhvinder S. Sidhu, Douglas N. W. Cooper, Kurt Drickamer, Hakon Leffler, and Françoise Poirier. Phylogenetic Analysis of the Vertebrate Galectin Family. *Molecular Biology and Evolution*, 21(7):1177–1187, July 2004.

[109] M. Wallis. The expanding growth hormone/prolactin family. *Journal of Molecular Endocrinology*, 9(3):185–188, December 1992.

[110] Jay A. Tischfield, Yu-Rong Xia, Diana M. Shih, Ivana Klisak, Ju Chen, Sandra J. Engle, Aristotle N. Siakotos, Michelle V. Winstead, Jeffrey J. Seilhamer, Valérie Allamand, Gabor Gyapay, and Aldons J. Lusis. Low-Molecular-Weight, Calcium-Dependent Phospholipase A2Genes Are Linked and Map to Homologous Chromosome Regions in Mouse and Human. *Genomics*, 32(3):328–333, March 1996.

[111] Devin P Locke, Nicoletta Archidiacono, Doriana Misceo, Maria Francesca Cardone, Stephane Deschamps, Bruce Roe, Mariano Rocchi, and Evan E Eichler. Refinement of a chimpanzee pericentric inversion breakpoint to a segmental duplication cluster. *Genome Biology*, page 9, 2003.

[112] Werner Albig, Petra Kioschis, Annemarie Poustka, Konstanze Meergans, and Detlef Doenecke. Human Histone Gene Organization: Nonregular Arrangement within a Large Cluster. *Genomics*, 40(2):314–322, March 1997.

[113] Roderick Nigel Finn, François Chauvigné, Jón Baldur Hlidberg, Christopher P. Cutler, and Joan Cerdà. The Lineage-Specific Evolution of Aquaporin Gene Clusters Facilitated Tetrapod Terrestrial Adaptation. *PLoS ONE*, 9(11):e113686, November 2014.

[114] A. Weiss, D. McDonough, B. Wertman, L. Acakpo-Satchivi, K. Montgomery, R. Kucherlapati, L. Leinwand, and K. Krauter. Organization of human and mouse skeletal myosin heavy chain gene clusters is highly conserved. *Proceedings of the National Academy of Sciences*, 96(6):2958–2963, March 1999.

[115] Saara M. Rawn and James C. Cross. The Evolution, Regulation, and Function of Placenta-Specific Genes. *Annual Review of Cell and Developmental Biology*, 24(1):159–181, November 2008.

[116] David E. K. Ferrier. Evolution of Homeobox Gene Clusters in Animals: The Giga-Cluster and Primary vs. Secondary Clustering. *Frontiers in Ecology and Evolution*, 4, April 2016.

[117] Claudio Passanantil, Brendan Davies, Miriam Ford, and Mike Fried. Structure of an inverted duplication formed as a first step in a gene amplification event: Implications for a model of gene. page 7.

[118] N.G. Than, R. Romero, Y. Xu, O. Erez, Z. Xu, G. Bhatti, R. Leavitt, T.H. Chung, H. El-Azzamy, C. LaJeunesse, B. Wang, A. Balogh, G. Szalai, S. Land, Z. Dong, S.S. Hassan, T. Chaiworapongsa, M. Krispin, C.J. Kim, A.L. Tarca, Z. Papp, and H. Bohn. Evolutionary origins of the placental expression of chromosome 19 cluster galectins and their complex dysregulation in preeclampsia. *Placenta*, 35(11):855–865, November 2014.

[119] Wolfgang Liebl, Peter Ruile, Karin Bronnenmeier, Katrin Riedel, Friedrich Lottspeich, and Ingrid Greif. Analysis of a Thermotoga maritima DNA fragment encoding two similar thermostable cellulases, CelA and CelB, and characterization of the recombinant enzymes. *Microbiology*, 142(9):2533–2542, 1996.

[120] Michael Karin, Rick Najarian, Alois Haslinger, Pablo Valenzuela, Juliet Welch, and Seymour Fogel. Primary structure and transcription of an amplified genetic locus: The CUPI locus of yeast. *Proceedings of the National Academy of Sciences*, 81:337–341, January 1984.

[121] Richard W. Michelmore and Blake C. Meyers. Clusters of Resistance Genes in Plants Evolve by Divergent Selection and a Birth-and-Death Process. *Genome Research*, 8(11):1113–1130, November 1998.

[122] A. M. Runck, H. Moriyama, and J. F. Storz. Evolution of Duplicated -Globin Genes and the Structural Basis of Hemoglobin Isoform Differentiation in Mus. *Molecular Biology and Evolution*, 26(11):2521–2532, November 2009.

[123] Jane Grimwood, Laurie A. Gordon, Anne Olsen, Astrid Terry, Jeremy Schmutz, Jane Lamerdin, Uffe Hellsten, David Goodstein, Olivier Couronne, Mary Tran-Gyamfi, Andrea Aerts, Michael Altherr, Linda Ashworth, Eva Bajorek, Stacey Black, Elbert Branscomb, Sean Caenepeel, Anthony Carrano, Chenier Caoile, Yee Man Chan, Mari Christensen, Catherine A. Cleland, Alex Copeland, Eileen Dalin, Paramvir Dehal, Mirian Denys, John C. Detter, Julio Escobar, Dave Flowers, Dea Fotopulos, Carmen Garcia, Anca M. Georgescu, Tijana Glavina, Maria Gomez, Eidelyn Gonzales, Matthew Groza, Nancy Hammon, Trevor Hawkins, Lauren Haydu, Isaac Ho, Wayne Huang, Sanjay Israni, Jamie Jett, Kristen Kadner, Heather Kimball, Arthur Kobayashi, Vladimer Larionov, Sun-Hee Leem, Frederick Lopez, Yunian Lou, Steve Lowry, Stephanie Malfatti, Diego Martinez, Paula McCready, Catherine Medina, Jenna Morgan, Kathryn Nelson, Matt Nolan, Ivan Ovcharenko, Sam Pitluck, Martin Pollard, Anthony P. Popkie, Paul Predki, Glenda Quan, Lucia Ramirez, Sam Rash, James Retterer, Alex Rodriguez, Stephanine Rogers, Asaf Salamov, Angelica Salazar, Xinwei She, Doug Smith, Tom Slezak, Victor Solovyev, Nina Thayer, Hope Tice, Ming Tsai, Anna Ustaszewska, Nu Vo, Mark Wagner, Jeremy Wheeler, Kevin Wu, Gary Xie, Joan Yang, Inna Dubchak, Terrence S. Furey, Pieter DeJong, Mark Dickson, David Gordon, Evan E. Eichler, Len A. Pennacchio, Paul Richardson, Lisa Stubbs, Daniel S. Rokhsar, Richard M. Myers, Edward M. Rubin, and Susan M. Lucas. The DNA sequence and biology of human chromosome 19. *Nature*, 428(6982):529–535, April 2004.

[124] S Lukic, J-C Nicolas, and A J Levine. The diversity of zinc-finger genes on human chromosome 19 provides an evolutionary mechanism for defense against inherited endogenous retroviruses. *Cell Death & Differentiation*, 21(3):381–387, March 2014.

[125] Liina Nagirnaja, Kristiina Rull, Liis Uusküla, Pille Hallast, Marina Grigorova, and Maris Laan. Genomics and genetics of gonadotropin beta-subunit genes: Unique FSHB and duplicated LH-B/CGB loci. *Molecular and Cellular Endocrinology*, 329(1-2):4–16, November 2010.

[126] Huan Cao and Paul R. Crocker. Evolution of CD33-related siglecs: Regulating host immune

functions and escaping pathogen exploitation?: Evolution of CD33-related siglecs. *Immunology*, 132(1):18–26, January 2011.

[127] Pernilla Lindahl, Torbjörn Säll, Anders Bjartell, Anna M. Johansson, Hans Lilja, and Christer Halldén. Copy Number Variants in the Kallikrein Gene Cluster. *PLoS ONE*, 8(7):e69097, July 2013.

[128] Huaiyu Mi, Xiaosong Huang, Anushya Muruganujan, Haiming Tang, Caitlin Mills, Diane Kang, and Paul D. Thomas. PANTHER version 11: Expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Research*, 45(D1):D183–D189, January 2017.

[129] Sean R. Eddy. Accelerated Profile HMM Searches. *PLoS Computational Biology*, 7(10):e1002195, October 2011.

[130] Yoav Benjamini and Yosef Hochberg. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1):289–300, 1995.

[131] Jianhua Lin. Divergence Measures Based on the Shannon Entropy. *IEEE transactions on information theory*, 37(145-151):7, January 1991.

[132] Peter V Marsden. Egocentric and sociocentric measures of network centrality. *Social Networks*, 24(4):407–422, October 2002.

[133] A. J. Vilella, J. Severin, A. Ureta-Vidal, L. Heng, R. Durbin, and E. Birney. EnsemblCompara GeneTrees: Complete, duplication-aware phylogenetic trees in vertebrates. *Genome Research*, 19(2):327–335, December 2008.

[134] Michael Lynch, Martin O'Hely, Bruce Walsh, and Allan Force. The Probability of Preservation of a Newly Arisen Gene Duplicate. *Genetics*, 159:1789–1804, December 2001.

[135] Adrian M. Altenhoff, Nives Škunca, Natasha Glover, Clément-Marie Train, Anna Sueki, Ivana Piližota, Kevin Gori, Bartlomiej Tomiczek, Steven Müller, Henning Redestig, Gaston H. Gonnet,

and Christophe Dessimoz. The OMA orthology database in 2015: Function predictions, better plant support, synteny view and other improvements. *Nucleic Acids Research*, 43(D1):D240–D249, January 2015.

[136] Adrian M. Altenhoff, Manuel Gil, Gaston H. Gonnet, and Christophe Dessimoz. Inferring Hierarchical Orthologous Groups from Orthologous Gene Pairs. *PLOS ONE*, 8(1):e53786, January 2013.

[137] Sudhir Kumar, Glen Stecher, Michael Suleski, and S. Blair Hedges. TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Molecular Biology and Evolution*, 34(7):1812–1819, July 2017.

[138] Jingjing Tao, Kunkai Su, Chengbo Yu, Xiaoli Liu, Wei Wu, Wei Xu, Bingxun Jiang, Rui Luo, Jian Yao, Jiawei Zhou, Yan Zhan, Chao Ye, Wenji Yuan, Xianzhong Jiang, Wenyan Cui, Ming D. Li, and Lianjuan Li. Fine mapping analysis of HLA-DP/DQ gene clusters on chromosome 6 reveals multiple susceptibility loci for HBV infection. *Amino Acids*, 47(12):2623–2634, December 2015.

[139] Ning-Ning Chai, Haiyan Zhou, Joseph Hernandez, Hossein Najmabadi, Shalender Bhasin, and Pauline H. Yen. Structure and Organization of theRBMYGenes on the Human Y Chromosome: Transposition and Amplification of an Ancestral AutosomalhnRNPGGene. *Genomics*, 49(2):283–289, April 1998.

[140] Richa Saxena, Laura G. Brown, Trevor Hawkins, Raaji K. Alagappan, Helen Skaletsky, Mary Pat Reeve, Renee Reijo, Steve Rozen, Mary Beth Dinulos, Christine M. Disteche, and David C. Page. The DAZ gene cluster on the human Y chromosome arose from an autosomal gene that was transposed, repeatedly amplified and pruned. *Nature Genetics*, 14(3):292–299, November 1996.

[141] T. Vogel and J. Schmidtke. Structure and function of TSPY, the Y-chromosome gene coding for the “testis-specific protein”. *Cytogenetic and Genome Research*, 80(1-4):209–213, 1998.

[142] Debasish Roy and Gloria M. Calaf. Allelic loss at chromosome 11q13 alters FGF3 gene ex-

pression in a human breast cancer progression model. *Oncology Reports*, 32(6):2445–2452, December 2014.

[143] Lucia Carbone, R. Alan Harris, Sante Gnerre, Krishna R. Veeramah, Belen Lorente-Galdos, John Huddleston, Thomas J. Meyer, Javier Herrero, Christian Roos, Bronwen Aken, Fabio Anaclerio, Nicoletta Archidiacono, Carl Baker, Daniel Barrell, Mark A. Batzer, Kathryn Beal, Antoine Blancher, Craig L. Bohrson, Markus Brameier, Michael S. Campbell, Oronzo Capozzi, Claudio Casola, Giorgia Chiatante, Andrew Cree, Annette Damert, Pieter J. de Jong, Laura Dumas, Marcos Fernandez-Callejo, Paul Flicek, Nina V. Fuchs, Ivo Gut, Marta Gut, Matthew W. Hahn, Jessica Hernandez-Rodriguez, LaDeana W. Hillier, Robert Hubley, Bianca Ianc, Zsuzsanna Izsvák, Nina G. Jablonski, Laurel M. Johnstone, Anis Karimpour-Fard, Miriam K. Konkel, Dennis Kostka, Nathan H. Lazar, Sandra L. Lee, Lora R. Lewis, Yue Liu, Devin P. Locke, Swapan Mallick, Fernando L. Mendez, Matthieu Muffato, Lynne V. Nazareth, Kimberly A. Nevonen, Majesta O'Bleness, Cornelia Ochis, Duncan T. Odom, Katherine S. Pollard, Javier Quilez, David Reich, Mariano Rocchi, Gerald G. Schumann, Stephen Searle, James M. Sikela, Gabriella Skollar, Arian Smit, Kemal Sonmez, Boudewijn ten Hallers, Elizabeth Terhune, Gregg W. C. Thomas, Brygg Ullmer, Mario Ventura, Jerilyn A. Walker, Jeffrey D. Wall, Lutz Walter, Michelle C. Ward, Sarah J. Wheelan, Christopher W. Whelan, Simon White, Larry J. Wilhelm, August E. Woerner, Mark Yandell, Baoli Zhu, Michael F. Hammer, Tomas Marques-Bonet, Evan E. Eichler, Lucinda Fulton, Catrina Fronick, Donna M. Muzny, Wesley C. Warren, Kim C. Worley, Jeffrey Rogers, Richard K. Wilson, and Richard A. Gibbs. Gibbon genome and the fast karyotype evolution of small apes. *Nature*, 513(7517):195–201, September 2014.

[144] Joseph Felsenstein. Phylogenies and the Comparative Method. *The American Naturalist*, 125(1):1–15, 1985.

[145] Paramvir Dehal, Paul Predki, Anne S. Olsen, Art Kobayashi, Peg Folta, Susan Lucas, Miriam Land, Astrid Terry, Carol L. Ecale Zhou, Sam Rash, Qing Zhang, Laurie Gordon, Joomyeong Kim, Christopher Elkin, Martin J. Pollard, Paul Richardson, Dan Rokhsar, Ed Uberbacher, Trevor Hawkins, Elbert Branscomb, and Lisa Stubbs. Human Chromosome 19 and Related

Regions in Mouse: Conservative and Lineage-Specific Evolution. *Science*, 293(5527):104–111, July 2001.

[146] Tanja Schwander, Romain Libbrecht, and Laurent Keller. Supergenes and Complex Phenotypes. *Current Biology*, 24(7):R288–R294, March 2014.

[147] François Spitz, Federico Gonzalez, and Denis Duboule. A Global Control Region Defines a Chromosomal Regulatory Landscape Containing the HoxD Cluster. *Cell*, 113(3):405–417, May 2003.

[148] C. Espitia Jaimes, R. J. Fish, and M. Neerman-Arbez. Local chromatin interactions contribute to expression of the fibrinogen gene cluster. *Journal of Thrombosis and Haemostasis*, 16(10):2070–2082, October 2018.

[149] Julie L. Lefebvre, Dimitar Kostadinov, Weisheng V. Chen, Tom Maniatis, and Joshua R. Sanes. Protocadherins mediate dendritic self-avoidance in the mammalian nervous system. *Nature*, 488(7412):517–521, August 2012.

[150] Ricard Albalat and Cristian Cañestro. Evolution by gene loss. *Nature Reviews Genetics*, 17(7):379–391, July 2016.

[151] Thomas L. Dunwell, Jordi Paps, and Peter W. H. Holland. Novel and divergent genes in the evolution of placental mammals. *Proc. R. Soc. B*, 284(1864):20171357, October 2017.

[152] Roberto Feuda, Ferdinand Marlétaz, Michael A. Bentley, and Peter W.H. Holland. Conservation, Duplication, and Divergence of Five Opsin Genes in Insect Evolution. *Genome Biology and Evolution*, 8(3):579–587, February 2016.

[153] Findley R. Finseth, Yuzhu Dong, Arpiar Saunders, and Lila Fishman. Duplication and Adaptive Evolution of a Key Centromeric Protein in Mimulus , a Genus with Female Meiotic Drive. *Molecular Biology and Evolution*, 32(10):2694–2706, October 2015.

[154] Mia T. Levine, Helen M. Vander Wende, Emily Hsieh, EmilyClare P. Baker, and Harmit S. Malik. Recurrent Gene Duplication Diversifies Genome Defense Repertoire in Drosophila. *Molecular Biology and Evolution*, 33(7):1641–1653, July 2016.

[155] Joerg Fleischer. Mammalian olfactory receptors. *Frontiers in Cellular Neuroscience*, 3, 2009.

[156] Michael Wallis. Episodic Evolution of Protein Hormones: Molecular Evolution of Pituitary Prolactin. *Journal of Molecular Evolution*, 50(5):465–473, May 2000.

[157] K. Katoh and H. Toh. Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics*, 9(4):286–298, March 2008.

[158] M. Suyama, D. Torrents, and P. Bork. PAL2NAL: Robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research*, 34(Web Server):W609–W612, July 2006.

[159] Roger Newson. Parameters behibd ”nonparametric” statistics: Kendall’s tau, Somers’ D and median differences. *The Stata Journal*, 2(1):45–64, 2002.

[160] Florian Ehrlich, Heinz Fischer, Lutz Langbein, Silke Praetzel-Wunder, Bettina Ebner, Katarzyna Figlak, Anton Weissenbacher, Wolfgang Sipos, Erwin Tschachler, and Leopold Eckhart. Differential Evolution of the Epidermal Keratin Cytoskeleton in Terrestrial and Aquatic Mammals. *Molecular Biology and Evolution*, 36(2):328–340, February 2019.

[161] Patrícia de Sousa-Pereira, Francisco Amado, Joana Abrantes, Rita Ferreira, Pedro J. Esteves, and Rui Vitorino. An evolutionary perspective of mammal salivary peptide families: Cystatins, histatins, statherin and PRPs. *Archives of Oral Biology*, 58(5):451–458, May 2013.

[162] Dušan Kordiš and Vito Turk. Phylogenomic analysis of the cystatin superfamily in eukaryotes and prokaryotes. *BMC Evolutionary Biology*, 9(1):266, December 2009.

[163] Chih-Yao Hsu, Chung-Shien Wu, and Shu-Miaw Chaw. Birth of Four Chimeric Plastid Gene Clusters in Japanese Umbrella Pine. *Genome Biology and Evolution*, 8(6):1776–1784, June 2016.

[164] Vladimir Kanazin, Laura Fredrick Marek, and Randy C. Shoemaker. Resistance Gene Analogs are Conserved and Clustered in Soybean.

Appendix A

Figures

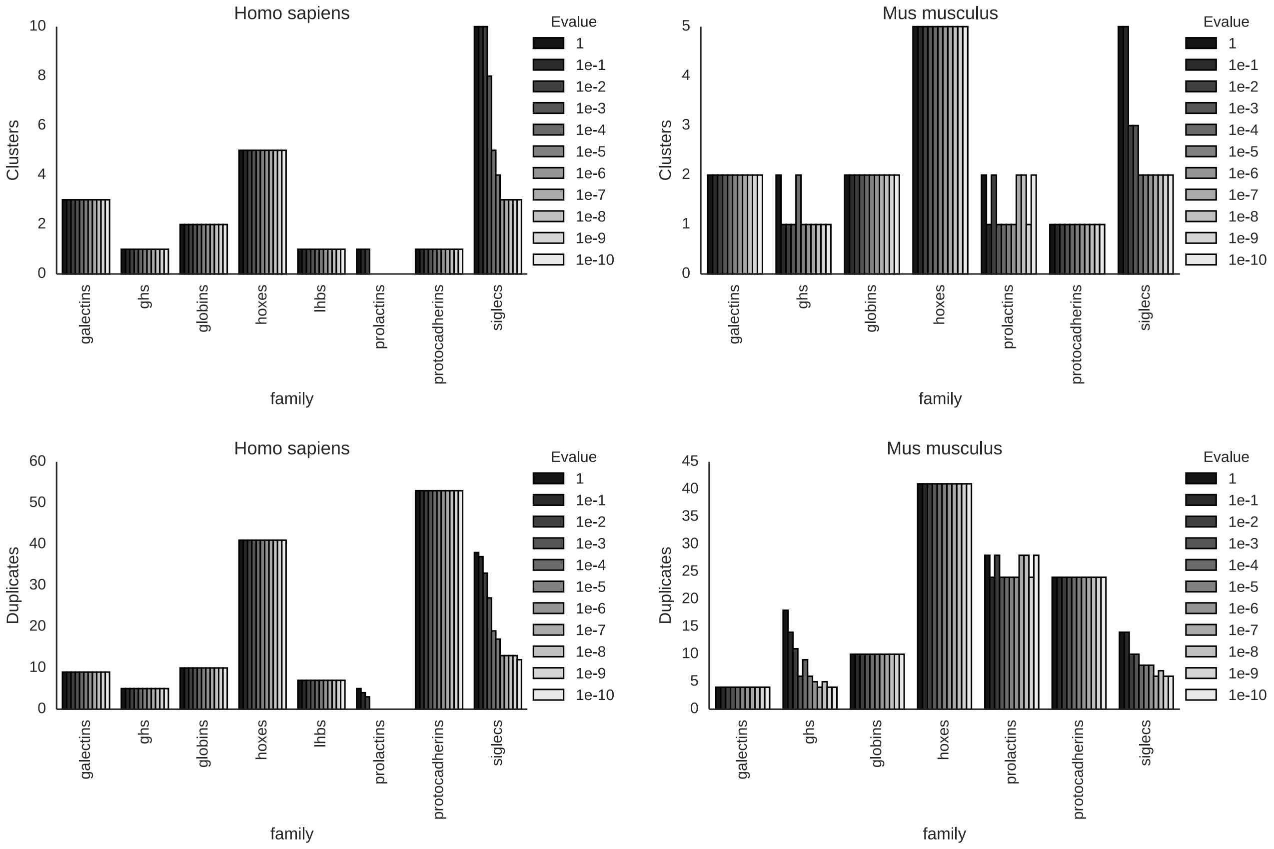

Figure A.1: Effect of BLAST E-value in the number of CTDGs (top) and numbr of gene duplicates (bottom) in human (left) and mouse (right) identified by CTDGFinder.

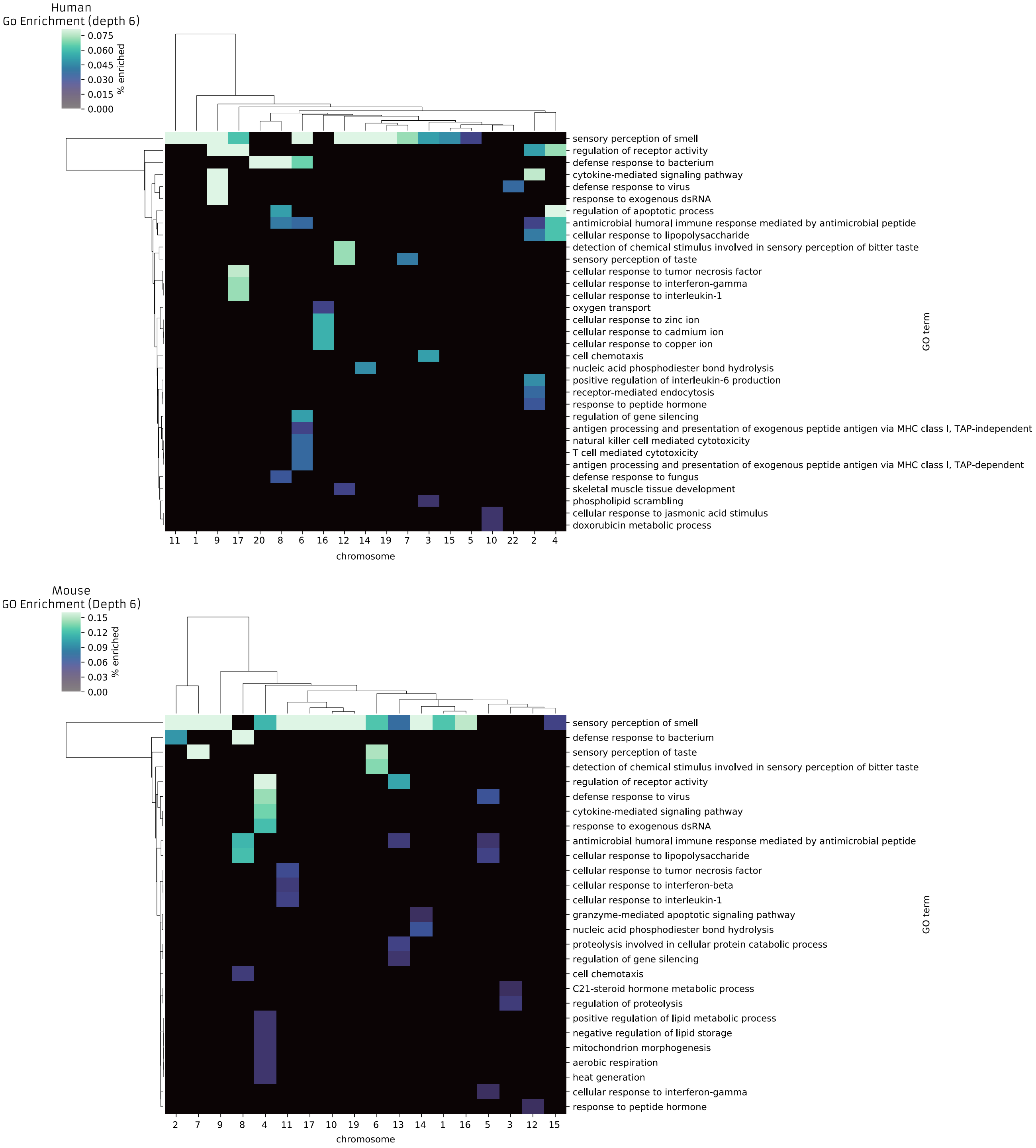

Figure A.2: GO enrichment analysis of clustered genes compared to non-clustered genes in human and mouse. The percentage of clustered genes in each chromosome enriched for biological processes is used to color the heatmap. The GO terms represented in the heatmap belong to the depth 6 because it is the depth in the ontology graph with the highest diversity of terms.

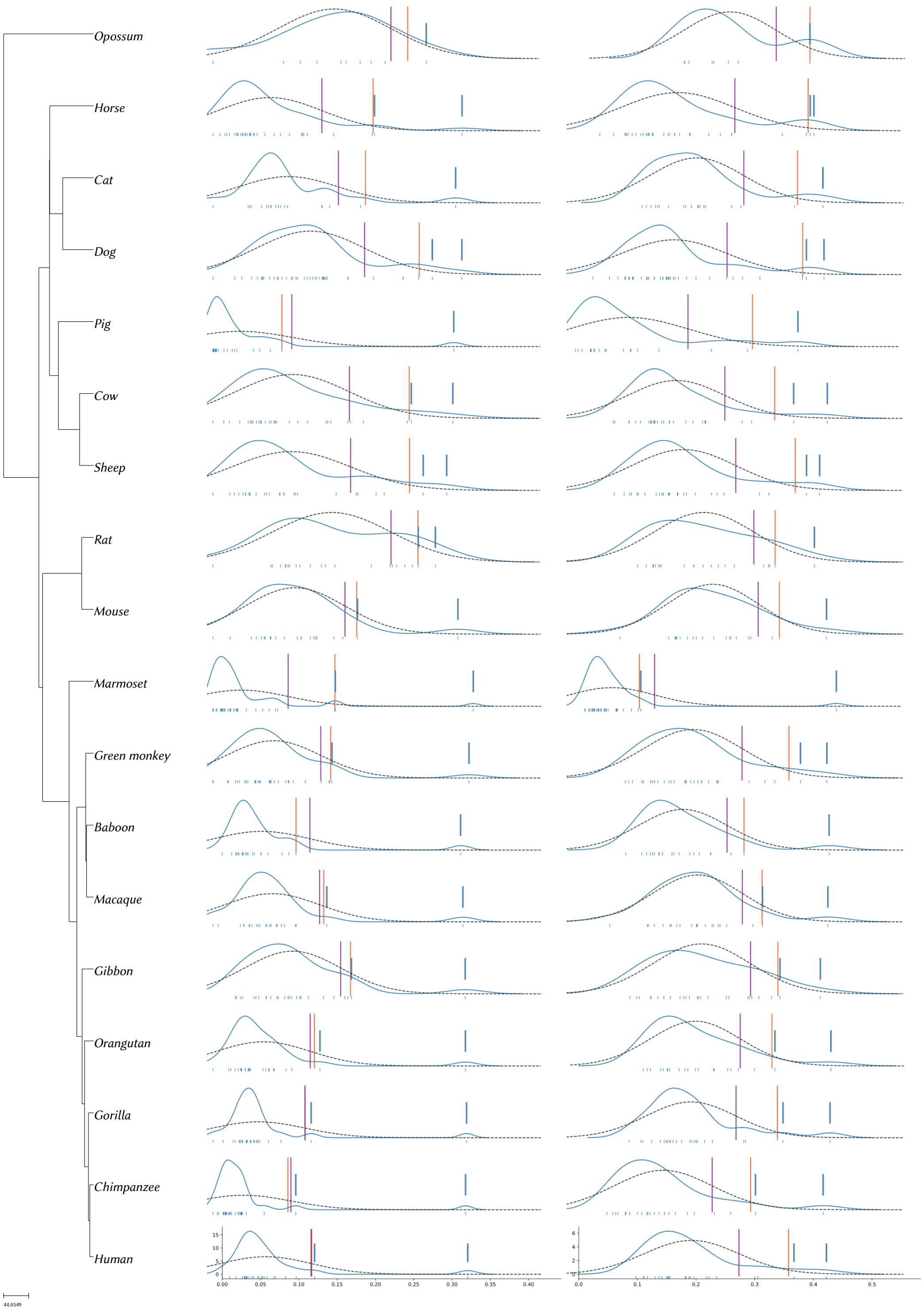

Figure A.3: Distribution of CTDGs across mammalian species. X axis: fraction of chromosome length spanned by CTDGs (A), and fraction of clustered genes (i.e., genes in CTDGs) per chromosome (B). Y axis: density. Axes are plotted only on the human distribution to increase readability. Dotted line: fitted normal distribution, blue line: kernel density estimate, orange line: percentile 95, purple line: standard deviation plus the mean of the standard distribution fitted to the data, bottom blue ticks: data for every chromosome or scaffold, middle blue lines: data for chromosomes beyond the percentile 95. All plots are plotted on the same scale on both axes.

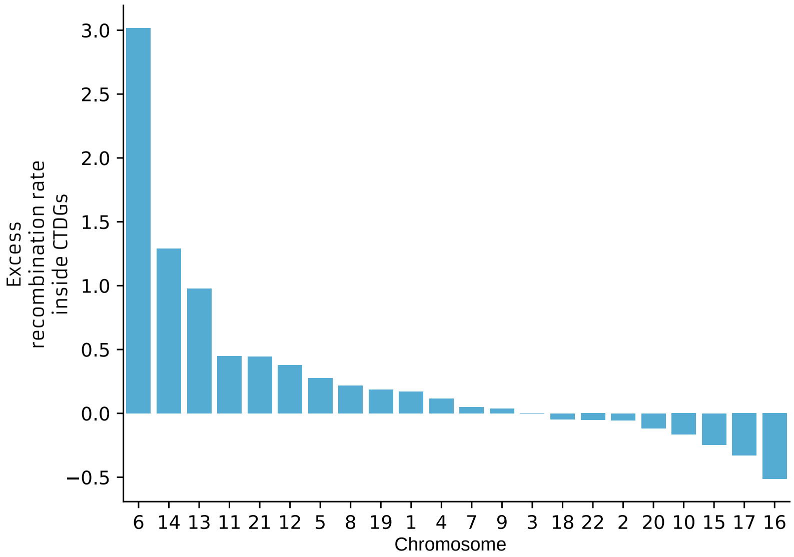

Figure A.4: Mean recombination rate difference between regions inside and outside of CTDGs. The excess recombination rate inside CTDGs was calculated by subtracting the mean recombination rate outside CTDGs from the mean recombination rate inside CTDGs.

Figure A.5: High correlation between gene order and phylogenetic divergence of the genes in the keratin CTDG in mammalian genomes. The figure follows the legend in fig. 4.2 The image can be accessed at https://figshare.com/account/home#/projects/61319

Figure A.6: Gene-family level phylogenetic tree of all the keratins under study. Blue: focus clade in fig. 4.3. Image can be accessed at https://figshare.com/account/home#/projects/61319

Figure A.7: Expression pattern of the human genes in the CST CTDG (From the GTEx database).Image can be accessed at https://figshare.com/account/home#/projects/61319

Appendix B

Tables

Table B.1: Query sequences used for each gene family in human and mouse

Species	Family	Accession	Symbol
human	galectins	NP_001819.2	CLC
human	galectins	XP_011525176.1	LGALS13
human	galectins	NP_064514.1	LGALS14
human	galectins	NP_001177370.2	LGALS16
human	galectins	XP_011525276.1	LGALS4
human	galectins	NP_002298.1	LGALS7
human	galectins	NP_001035972.1	LGALS7B
human	ghs	NP_001308.1	CSH1
human	ghs	NP_072171.1	CSH2
human	ghs	NP_001309.3	CSHL1
human	ghs	XP_011522914.1	GH1
human	ghs	NP_072052.1	GH2
human	globins	NP_000509.1	HBB
human	globins	NP_000510.1	HBD
human	globins	NP_005321.1	HBE1
human	globins	NP_000550.2	HBG1
human	globins	NP_000175.1	HBG2
human	hoxes	NP_005513.1	HOXA1
human	hoxes	NP_061824.3	HOXA10
human	hoxes	NP_005514.1	HOXA11
human	hoxes	NP_000513.2	HOXA13
human	hoxes	NP_006726.1	HOXA2
human	hoxes	NP_109377.1	HOXA3
human	hoxes	NP_002132.3	HOXA4
human	hoxes	NP_061975.2	HOXA5
human	hoxes	NP_076919.1	HOXA6
human	hoxes	NP_008827.2	HOXA7
human	hoxes	NP_689952.1	HOXA9
human	lhbs	NP_000728.1	CGB
human	lhbs	NP_203695.2	CGB1
human	lhbs	NP_203696.2	CGB2
human	lhbs	NP_149032.1	CGB5
human	lhbs	NP_149133.1	CGB7
human	lhbs	NP_149439.1	CGB8
human	lhbs	XP_011525277.1	LHB
human	protocadherins	NP_113598.1	PCDHA1
human	protocadherins	NP_114065.1	PCDHA10
human	protocadherins	NP_114067.1	PCDHA11
human	protocadherins	NP_114070.1	PCDHA12
human	protocadherins	NP_114071.1	PCDHA13

Table B.1: Query sequences used for each gene family in human and mouse

Species	Family	Accession	Symbol
human	protocadherins	NP_113683.1	PCDHA2
human	protocadherins	NP_113685.1	PCDHA3
human	protocadherins	NP_113688.1	PCDHA4
human	protocadherins	NP_113689.1	PCDHA5
human	protocadherins	NP_114036.1	PCDHA6
human	protocadherins	NP_114040.1	PCDHA7
human	protocadherins	NP_114062.1	PCDHA8
human	protocadherins	NP_054724.1	PCDHA9
human	protocadherins	NP_114088.2	PCDHAC1
human	protocadherins	NP_114089.1	PCDHAC2
human	protocadherins	NP_037472.2	PCDHB1
human	protocadherins	NP_061753.1	PCDHB10
human	protocadherins	NP_061754.1	PCDHB11
human	protocadherins	NP_061755.1	PCDHB12
human	protocadherins	NP_061756.1	PCDHB13
human	protocadherins	NP_061757.1	PCDHB14
human	protocadherins	NP_061758.1	PCDHB15
human	protocadherins	NP_061759.1	PCDHB2
human	protocadherins	NP_061760.1	PCDHB3
human	protocadherins	NP_061761.1	PCDHB4
human	protocadherins	NP_056484.2	PCDHB5
human	protocadherins	NP_061762.2	PCDHB6
human	protocadherins	NP_061763.1	PCDHB7
human	protocadherins	NP_061993.3	PCDHB8
human	protocadherins	NP_061992.3	PCDHB9
human	protocadherins	NP_114382.1	PCDHGA1
human	protocadherins	NP_114479.1	PCDHGA10
human	protocadherins	NP_114480.1	PCDHGA11
human	protocadherins	NP_115265.1	PCDHGA12
human	protocadherins	NP_114398.1	PCDHGA2
human	protocadherins	NP_114400.1	PCDHGA3
human	protocadherins	NP_114442.2	PCDHGA4
human	protocadherins	NP_114443.1	PCDHGA5
human	protocadherins	NP_114475.1	PCDHGA6
human	protocadherins	NP_114476.1	PCDHGA7
human	protocadherins	NP_054723.1	PCDHGA8
human	protocadherins	NP_114478.1	PCDHGA9
human	protocadherins	NP_115266.1	PCDHGB1
human	protocadherins	NP_115267.1	PCDHGB2
human	protocadherins	NP_115268.2	PCDHGB3
human	protocadherins	NP_115269.1	PCDHGB4
human	protocadherins	NP_115270.1	PCDHGB5
human	protocadherins	NP_115271.1	PCDHGB6
human	protocadherins	NP_115272.1	PCDHGB7
human	protocadherins	NP_115778.1	PCDHGC3
human	protocadherins	NP_115782.1	PCDHGC4
human	protocadherins	NP_115783.1	PCDHGC5
human	siglecs	XP_011525833.1	CD33

Table B.1: Query sequences used for each gene family in human and mouse

Species	Family	Accession	Symbol
human	siglecs	NP_001164632.1	SIGLEC10
human	siglecs	NP_201586.1	SIGLEC12
human	siglecs	NP_001092082.1	SIGLEC14
human	siglecs	NP_003821.1	SIGLEC5
human	siglecs	NP_001171018.1	SIGLEC6
human	siglecs	XP_006723203.1	SIGLEC7
human	siglecs	XP_011525037.1	SIGLEC8
human	siglecs	XP_011525035.1	SIGLEC9
mouse	galectins	NP_034836.1	Lgals4
mouse	galectins	NP_032522.2	Lgals7
mouse	globins	NP_032245.1	Hbb-bh1
mouse	globins	XP_006508067.1	Hbb-bh2
mouse	globins	NP_001188320.1	Hbb-bs
mouse	globins	NP_032246.2	Hbb-bt
mouse	globins	NP_032247.1	Hbb-y
mouse	hoxes	NP_034579.3	Hoxa1
mouse	hoxes	NP_032289.2	Hoxa10
mouse	hoxes	NP_034580.1	Hoxa11
mouse	hoxes	NP_032290.1	Hoxa13
mouse	hoxes	NP_034581.1	Hoxa2
mouse	hoxes	NP_034582.1	Hoxa3
mouse	hoxes	NP_032291.1	Hoxa4
mouse	hoxes	NP_034583.1	Hoxa5
mouse	hoxes	NP_034584.1	Hoxa6
mouse	hoxes	XP_006505675.1	Hoxa7
mouse	hoxes	NP_034586.1	Hoxa9
mouse	prolactins	NP_035294.2	Prl
mouse	prolactins	NP_064375.1	Prl2a1
mouse	prolactins	NP_079808.1	Prl2b1
mouse	prolactins	XP_006516640.1	Prl2c2
mouse	prolactins	XP_006516641.1	Prl2c3
mouse	prolactins	NP_036084.2	Prl2c4
mouse	prolactins	XP_011242563.1	Prl2c5
mouse	prolactins	NP_080172.1	Prl3a1
mouse	prolactins	NP_032891.1	Prl3b1
mouse	prolactins	NP_038794.1	Prl3c1
mouse	prolactins	NP_001192251.1	Prl3d1
mouse	prolactins	XP_006516702.1	Prl3d2
mouse	prolactins	XP_006516703.1	Prl3d3
mouse	prolactins	NP_035295.1	Prl4a1
mouse	prolactins	NP_076235.1	Prl5a1
mouse	prolactins	NP_035296.1	Prl6a1
mouse	prolactins	NP_032956.1	Prl7a1
mouse	prolactins	XP_006516644.1	Prl7a2
mouse	prolactins	XP_011242659.1	Prl7b1
mouse	prolactins	XP_006516796.1	Prl7c1
mouse	prolactins	XP_006516643.1	Prl7d1
mouse	prolactins	NP_082753.1	Prl8a1

Table B.1: Query sequences used for each gene family in human and mouse

Species	Family	Accession	Symbol
mouse	prolactins	NP_034218.1	Prl8a2
mouse	prolactins	NP_035297.1	Prl8a6
mouse	prolactins	NP_076230.1	Prl8a8
mouse	prolactins	XP_006516792.1	Prl8a9
mouse	protocadherins	XP_006526384.1	Pcdh1
mouse	protocadherins	NP_059074.2	Pcdh12
mouse	protocadherins	NP_473413.1	Pcdha1
mouse	protocadherins	NP_444356.1	Pcdhb1
mouse	protocadherins	NP_444365.1	Pcdhb10
mouse	protocadherins	NP_444366.3	Pcdhb11
mouse	protocadherins	NP_444367.1	Pcdhb12
mouse	protocadherins	NP_444368.1	Pcdhb13
mouse	protocadherins	NP_444369.3	Pcdhb14
mouse	protocadherins	NP_444370.1	Pcdhb15
mouse	protocadherins	NP_444371.3	Pcdhb16
mouse	protocadherins	NP_444372.2	Pcdhb17
mouse	protocadherins	NP_444373.1	Pcdhb18
mouse	protocadherins	NP_444374.2	Pcdhb19
mouse	protocadherins	NP_444357.1	Pcdhb2
mouse	protocadherins	NP_444375.2	Pcdhb20
mouse	protocadherins	NP_444376.1	Pcdhb21
mouse	protocadherins	NP_444377.3	Pcdhb22
mouse	protocadherins	NP_444358.2	Pcdhb3
mouse	protocadherins	NP_444359.3	Pcdhb4
mouse	protocadherins	NP_444360.1	Pcdhb5
mouse	protocadherins	NP_444361.1	Pcdhb6
mouse	protocadherins	NP_444362.3	Pcdhb7
mouse	protocadherins	NP_444363.1	Pcdhb8
mouse	protocadherins	NP_444364.3	Pcdhb9
mouse	protocadherins	NP_291062.1	Pcdhga1
mouse	siglecs	NP_001104528.1	Cd33
mouse	siglecs	XP_006541383.1	Siglece
mouse	siglecs	NP_663556.1	Siglecf
mouse	siglecs	XP_006540925.1	Siglecg

Table B.2: Expression enhancement of genes in the CTDs from galectins and Siglecs in chromosome 19

Ensembl ID	NCBI ID	symbol	cluster	order	enhanced
ENSG00000105695	NP_542167.1	MAG	siglecs-19_4	1	cerebral cortex, ovary
ENSG00000012124	NP_001172029.1	CD22	siglecs-19_4	2	lymph node, tonsil
ENSG00000105388	NP_004354.3	CEACAM5	siglecs-19_5	1	colon, rectum
ENSG00000086548	NP_002474.4	CEACAM6	siglecs-19_5	2	bone marrow, colon, esophagus, gallbladder, lung, rectum
ENSG00000243130	NP_976032.2	PSG11	siglecs-19_3	1	placenta
ENSG00000242221	NP_112536.2	PSG2	siglecs-19_3	2	placenta
ENSG00000204941	NP_002772.3	PSG5	siglecs-19_3	3	placenta
ENSG00000243137	NP_001263424.1	PSG4	siglecs-19_3	4	placenta
ENSG00000129450	NP_055256.1	SIGLEC9	siglecs-19_2	1	appendix
ENSG00000168995	NP_001264130.1	SIGLEC7	siglecs-19_2	2	spleen
ENSG00000105383	NP_001171079.1	CD33	siglecs-19_2	3	non_enhanced
ENSG00000179213	NP_775906.1	SIGLECL1	siglecs-19_2	4	testis
ENSG00000142512	NP_001164632.1	SIGLEC10	siglecs-19_1	1	appendix
ENSG00000105366	NP_055257.2	SIGLEC8	siglecs-19_1	2	non_enhanced
ENSG00000254521	NP_201586.1	SIGLEC12	siglecs-19_1	3	spleen
ENSG00000105492	NP_001171018.1	SIGLEC6	siglecs-19_1	4	placenta
ENSG00000105501	NP_003821.1	SIGLEC5	siglecs-19_1	6	appendix, bone marrow
ENSG00000254415	NP_001092082.1	SIGLEC14	siglecs-19_1	7	appendix, spleen
ENSG00000205076	NP_002298.1	LGALS7	galectins-19_2	1	skin
ENSG00000178934	NP_001035972.1	LGALS7B	galectins-19_2	2	skin
ENSG00000171747	NP_006140.1	LGALS4	galectins-19_2	3	colon, duodenum, rectum, small intestine
ENSG00000105198	NP_037400.1	LGALS13	galectins-19_1	1	placenta
ENSG00000249861	NP_001177370.2	LGALS16	galectins-19_1	2	placenta
ENSG00000006659	NP_064514.1	LGALS14	galectins-19_1	3	placenta
ENSG00000105205	NP_001819.2	CLC	galectins-19_1	4	bone marrow

Table B.3: Mammalian CTDGs based on the 7 galectins on human chromosome 19

Species	Chromosome	Cluster	Gene duplicates
Bos taurus	18	galectins-18_1	3
	19	galectins-19_1	2
	5	galectins-5_1	2
Callithrix jacchus	22	galectins-22_1	6
	22	galectins-22_2	3
Canis lupus familiaris	1	galectins-1_1	3
	10	galectins-10_1	2
Capra hircus	18	galectins-18_1	3
	18	galectins-18_2	3
Chlorocebus sabaeus	6	galectins-6_1	4
	6	galectins-6_2	3
Equus caballus	10	galectins-10_1	3
	11	galectins-11_1	4
Felis catus	B4	galectins-B4_1	2
Gorilla gorilla	19	galectins-19_1	4
	19	galectins-19_2	3
	22	galectins-22_1	2
Homo sapiens	19	galectins-19_1	4
	19	galectins-19_2	3
	22	galectins-22_1	2
Macaca fascicularis	16	galectins-16_1	2
	19	galectins-19_1	5
	19	galectins-19_2	2
Macaca mulatta	10	galectins-10_1	2
	19	galectins-19_1	5
	19	galectins-19_2	2
Microtus ochrogaster	15	galectins-15_1	2
Monodelphis domestica	4	galectins-4_1	3
	8	galectins-8_1	4
Mus musculus	15	galectins-15_1	2
	7	galectins-7_1	2
Nomascus leucogenys	11	galectins-11_1	2
	17	galectins-17_1	2
	7b	galectins-7b_1	2
Pan troglodytes	17	galectins-17_1	2
	19	galectins-19_1	3
	19	galectins-19_2	3
	22	galectins-22_1	2
Papio anubis	10	galectins-10_1	2
	19	galectins-19_1	7
	19	galectins-19_2	3
Pongo abelii	11	galectins-11_1	2
	17	galectins-17_1	2
	19	galectins-19_1	6
	19	galectins-19_2	3
	22	galectins-22_1	2
Rattus norvegicus	1	galectins-1_1	2
	7	galectins-7_1	2
Sus scrofa	5	galectins-5_1	2
	6	galectins-6_1	2

Table B.4: Mammalian CTDGs based on the 2 galectins on mouse chromosome 7

Species	Chromosome	Cluster	Gene duplicates
Bos taurus	18	galectins_mm-18_1	3
	19	galectins_mm-19_1	2
	5	galectins_mm-5_1	2
Callithrix jacchus	22	galectins_mm-22_1	6
	22	galectins_mm-22_2	3
Canis lupus familiaris	1	galectins_mm-1_1	3
	10	galectins_mm-10_1	2
Capra hircus	18	galectins_mm-18_1	3
	18	galectins_mm-18_2	3
Chlorocebus sabaeus	6	galectins_mm-6_1	4
	6	galectins_mm-6_2	3
Equus caballus	10	galectins_mm-10_1	3
	11	galectins_mm-11_1	4
Felis catus	B4	galectins_mm-B4_1	2
Gorilla gorilla	19	galectins_mm-19_1	4
	19	galectins_mm-19_2	3
	22	galectins_mm-22_1	2
Homo sapiens	19	galectins_mm-19_1	4
	19	galectins_mm-19_2	3
	22	galectins_mm-22_1	2
Macaca fascicularis	16	galectins_mm-16_1	2
	19	galectins_mm-19_1	4
	19	galectins_mm-19_2	2
Macaca mulatta	10	galectins_mm-10_1	2
	19	galectins_mm-19_1	5
	19	galectins_mm-19_2	2
Microtus ochrogaster	15	galectins_mm-15_1	2
Monodelphis domestica	4	galectins_mm-4_1	3
	8	galectins_mm-8_1	4
Mus musculus	15	galectins_mm-15_1	2
	7	galectins_mm-7_1	2
Nomascus leucogenys	11	galectins_mm-11_1	2
	17	galectins_mm-17_1	2
	7b	galectins_mm-7b_1	2
Pan troglodytes	17	galectins_mm-17_1	2
	19	galectins_mm-19_1	3
	19	galectins_mm-19_2	3
	22	galectins_mm-22_1	2
Papio anubis	10	galectins_mm-10_1	2
	19	galectins_mm-19_1	6
	19	galectins_mm-19_2	3
Pongo abelii	11	galectins_mm-11_1	2
	17	galectins_mm-17_1	2
	19	galectins_mm-19_1	5
	19	galectins_mm-19_2	3
	22	galectins_mm-22_1	2
Rattus norvegicus	1	galectins_mm-1_1	2
	7	galectins_mm-7_1	2
Sus scrofa	5	galectins_mm-5_1	2
	6	galectins_mm-6_1	2

Table B.5: Mammalian β-globin CTDG using the 5 β-globin genes in the human chromosome 11

species	chromosome	cluster	gene duplicates
Bos taurus	15	globins-15_1	5
	25	globins-25_1	6
Callithrix jacchus	11	globins-11_1	4
	12	globins-12_1	3
Canis lupus familiaris	21	globins-21_1	5
	6	globins-6_1	4
Capra hircus	15	globins-15_1	10
	25	globins-25_1	4
Chlorocebus sabaeus	1	globins-1_1	4
	5	globins-5_1	5
Equus caballus	7	globins-7_1	5
Felis catus	D1	globins-D1_1	4
	E3	globins-E3_1	3
Gorilla gorilla gorilla	11	globins-11_1	5
	16	globins-16_1	6
Homo sapiens	11	globins-11_1	5
	16	globins-16_1	5
Macaca fascicularis	14	globins-14_1	5
	20	globins-20_1	4
Macaca mulatta	14	globins-14_1	5
	20	globins-20_1	6
Microtus ochrogaster	7	globins-7_1	5
Monodelphis domestica	4	globins-4_1	2
	6	globins-6_1	5
Mus musculus	11	globins-11_1	5
	7	globins-7_1	5
Nomascus leucogenys	15	globins-15_1	3
	18	globins-18_1	2
Oryctolagus cuniculus	1	globins-1_1	4
Ovis aries	15	globins-15_1	3
	24	globins-24_1	3
Pan troglodytes	11	globins-11_1	5
	16	globins-16_1	2
Papio anubis	14	globins-14_1	4
	20	globins-20_1	5
Pongo abelii	11	globins-11_1	4
Rattus norvegicus	1	globins-1_1	8
	10	globins-10_1	4
Sus scrofa	3	globins-3_1	4
	9	globins-9_1	4

Table B.6: Mammalian β-globin CTDG using the 5 β-globin genes in the mouse chromosome 7

species	chromosome	cluster	gene duplicates
Bos taurus	15	globins_mm-15_1	5
	25	globins_mm-25_1	6
Callithrix jacchus	11	globins_mm-11_1	4
	12	globins_mm-12_1	3
Canis lupus familiaris	21	globins_mm-21_1	5
	6	globins_mm-6_1	4
Capra hircus	15	globins_mm-15_1	10
	25	globins_mm-25_1	4
Chlorocebus sabaeus	1	globins_mm-1_1	4
	5	globins_mm-5_1	5
Equus caballus	7	globins_mm-7_1	5
Felis catus	D1	globins_mm-D1_1	4
	E3	globins_mm-E3_1	3
Gorilla gorilla gorilla	11	globins_mm-11_1	5
	16	globins_mm-16_1	6
Homo sapiens	11	globins_mm-11_1	5
	16	globins_mm-16_1	5
Macaca fascicularis	14	globins_mm-14_1	5
	20	globins_mm-20_1	4
Macaca mulatta	14	globins_mm-14_1	5
	20	globins_mm-20_1	6
Microtus ochrogaster	7	globins_mm-7_1	5
Monodelphis domestica	4	globins_mm-4_1	2
	6	globins_mm-6_1	5
Mus musculus	11	globins_mm-11_1	5
	7	globins_mm-7_1	5
Nomascus leucogenys	15	globins_mm-15_1	3
	18	globins_mm-18_1	2
Oryctolagus cuniculus	1	globins_mm-1_1	4
Ovis aries	15	globins_mm-15_1	3
	24	globins_mm-24_1	3
Pan troglodytes	11	globins_mm-11_1	5
	16	globins_mm-16_1	2
Papio anubis	14	globins_mm-14_1	4
	20	globins_mm-20_1	5
Pongo abelii	11	globins_mm-11_1	4
Rattus norvegicus	1	globins_mm-1_1	8
	10	globins_mm-10_1	4
Sus scrofa	3	globins_mm-3_1	4
	9	globins_mm-9_1	4

Table B.7: Mammalian Hox CTDG using the 11 HoxA genes in the human chromosome 7

species	chromosome	cluster	gene duplicates
Bos taurus	19	hoxes-19_1	7
	2	hoxes-2_1	7
	4	hoxes-4_1	9
	5	hoxes-5_1	8
Callithrix jacchus	5	hoxes-5_1	9
	5	hoxes-5_2	2
	6	hoxes-6_1	7
	8	hoxes-8_1	9
	9	hoxes-9_1	7
Canis lupus familiaris	14	hoxes-14_1	9
	25	hoxes-25_1	2
	27	hoxes-27_1	6
	36	hoxes-36_1	7
	9	hoxes-9_1	9
Capra hircus	19	hoxes-19_1	9
	2	hoxes-2_1	8
	4	hoxes-4_1	9
	5	hoxes-5_1	9
Chlorocebus sabaeus	10	hoxes-10_1	7
	11	hoxes-11_1	6
	16	hoxes-16_1	6
	21	hoxes-21_1	10
	3	hoxes-3_1	2
Equus caballus	11	hoxes-11_1	8
	17	hoxes-17_1	3
	18	hoxes-18_1	6
	4	hoxes-4_1	8
	6	hoxes-6_1	7
Felis catus	A1	hoxes-A1_1	3
	A2	hoxes-A2_1	10
	B4	hoxes-B4_1	9
	C1	hoxes-C1_1	9
	E1	hoxes-E1_1	9
Gorilla gorilla gorilla	12	hoxes-12_1	8
	13	hoxes-13_1	2
	2B	hoxes-2B_1	11
	5	hoxes-5_1	9
	7	hoxes-7_1	11
Homo sapiens	12	hoxes-12_1	9
	13	hoxes-13_1	2
	17	hoxes-17_1	10
	2	hoxes-2_1	9
	7	hoxes-7_1	12
Macaca fascicularis	11	hoxes-11_1	8
	12	hoxes-12_1	8
	16	hoxes-16_1	7
	17	hoxes-17_1	2
	3	hoxes-3_1	10
Macaca mulatta	11	hoxes-11_1	8
	12	hoxes-12_1	8
	16	hoxes-16_1	10
	17	hoxes-17_1	2
	3	hoxes-3_1	10
Microtus ochrogaster	15	hoxes-15_1	8

Table B.7: Mammalian Hox CTDG using the 11 HoxA genes in the human chromosome 7

species	chromosome	cluster	gene duplicates
	2	hoxes-2_1	2
	4	hoxes-4_1	8
Monodelphis domestica	2	hoxes-2_1	7
	4	hoxes-4_1	7
	4	hoxes-4_2	2
	8	hoxes-8_1	6
Mus musculus	11	hoxes-11_1	10
	15	hoxes-15_1	9
	2	hoxes-2_1	9
	5	hoxes-5_1	2
	6	hoxes-6_1	11
Nomascus leucogenys	11	hoxes-11_1	7
	17	hoxes-17_1	9
	22a	hoxes-22a_1	9
	9	hoxes-9_1	2
Oryctolagus cuniculus	10	hoxes-10_1	10
	19	hoxes-19_1	6
	4	hoxes-4_1	8
	7	hoxes-7_1	9
Ovis aries	11	hoxes-11_1	5
	2	hoxes-2_1	3
	3	hoxes-3_1	7
	4	hoxes-4_1	4
Pan troglodytes	12	hoxes-12_1	7
	13	hoxes-13_1	3
	17	hoxes-17_1	8
	2B	hoxes-2B_1	9
	7	hoxes-7_1	9
Papio anubis	11	hoxes-11_1	7
	12	hoxes-12_1	8
	16	hoxes-16_1	8
	17	hoxes-17_1	2
	3	hoxes-3_1	11
Pongo abelii	12	hoxes-12_1	7
	13	hoxes-13_1	2
	2B	hoxes-2B_1	9
	7	hoxes-7_1	11
Rattus norvegicus	10	hoxes-10_1	10
	12	hoxes-12_1	2
	3	hoxes-3_1	7
	4	hoxes-4_1	13
	7	hoxes-7_1	9
Sus scrofa	11	hoxes-11_1	3
	12	hoxes-12_1	8
	15	hoxes-15_1	8
	18	hoxes-18_1	12
	5	hoxes-5_1	6
	9	hoxes-9_1	2

Table B.8: Mammalian Hox CTDG using the 11 HoxA genes in the mouse chromosome 6

species	chromosome	cluster	gene duplicates
Bos taurus	19	hoxes_mm-19_1	7
	2	hoxes_mm-2_1	7
	4	hoxes_mm-4_1	9
	5	hoxes_mm-5_1	8
Callithrix jacchus	5	hoxes_mm-5_1	9
	5	hoxes_mm-5_2	2
	6	hoxes_mm-6_1	7
	8	hoxes_mm-8_1	9
	9	hoxes_mm-9_1	7
Canis lupus familiaris	14	hoxes_mm-14_1	9
	25	hoxes_mm-25_1	2
	27	hoxes_mm-27_1	6
	36	hoxes_mm-36_1	7
	9	hoxes_mm-9_1	9
Capra hircus	19	hoxes_mm-19_1	9
	2	hoxes_mm-2_1	8
	4	hoxes_mm-4_1	9
	5	hoxes_mm-5_1	9
Chlorocebus sabaeus	10	hoxes_mm-10_1	7
	11	hoxes_mm-11_1	6
	16	hoxes_mm-16_1	6
	21	hoxes_mm-21_1	10
	3	hoxes_mm-3_1	3
Equus caballus	11	hoxes_mm-11_1	8
	17	hoxes_mm-17_1	3
	18	hoxes_mm-18_1	6
	4	hoxes_mm-4_1	8
	6	hoxes_mm-6_1	7
Felis catus	A1	hoxes_mm-A1_1	3
	A2	hoxes_mm-A2_1	10
	B4	hoxes_mm-B4_1	9
	C1	hoxes_mm-C1_1	9
	E1	hoxes_mm-E1_1	9
Gorilla gorilla gorilla	12	hoxes_mm-12_1	8
	13	hoxes_mm-13_1	3
	2B	hoxes_mm-2B_1	11
	5	hoxes_mm-5_1	9
	7	hoxes_mm-7_1	11
Homo sapiens	12	hoxes_mm-12_1	9
	13	hoxes_mm-13_1	3
	17	hoxes_mm-17_1	10
	2	hoxes_mm-2_1	9
	7	hoxes_mm-7_1	11
Macaca fascicularis	11	hoxes_mm-11_1	8
	12	hoxes_mm-12_1	8
	16	hoxes_mm-16_1	7
	17	hoxes_mm-17_1	3
	3	hoxes_mm-3_1	10
Macaca mulatta	11	hoxes_mm-11_1	8
	12	hoxes_mm-12_1	8
	16	hoxes_mm-16_1	10
	17	hoxes_mm-17_1	2
	3	hoxes_mm-3_1	10
Microtus ochrogaster	15	hoxes_mm-15_1	8

Table B.8: Mammalian Hox CTDG using the 11 HoxA genes in the mouse chromosome 6

species	chromosome	cluster	gene duplicates
	2	hoxes_mm-2_1	2
	4	hoxes_mm-4_1	8
Monodelphis domestica	2	hoxes_mm-2_1	7
	4	hoxes_mm-4_1	6
	4	hoxes_mm-4_2	2
	8	hoxes_mm-8_1	6
Mus musculus	11	hoxes_mm-11_1	10
	15	hoxes_mm-15_1	9
	2	hoxes_mm-2_1	9
	5	hoxes_mm-5_1	3
	6	hoxes_mm-6_1	11
Nomascus leucogenys	11	hoxes_mm-11_1	7
	17	hoxes_mm-17_1	8
	22a	hoxes_mm-22a_1	9
	9	hoxes_mm-9_1	2
Oryctolagus cuniculus	10	hoxes_mm-10_1	10
	19	hoxes_mm-19_1	6
	4	hoxes_mm-4_1	8
	7	hoxes_mm-7_1	9
Ovis aries	11	hoxes_mm-11_1	5
	2	hoxes_mm-2_1	3
	3	hoxes_mm-3_1	7
	4	hoxes_mm-4_1	4
Pan troglodytes	12	hoxes_mm-12_1	7
	13	hoxes_mm-13_1	3
	17	hoxes_mm-17_1	8
	2B	hoxes_mm-2B_1	9
	7	hoxes_mm-7_1	9
Papio anubis	11	hoxes_mm-11_1	7
	12	hoxes_mm-12_1	8
	16	hoxes_mm-16_1	8
	17	hoxes_mm-17_1	2
	3	hoxes_mm-3_1	11
Pongo abelii	12	hoxes_mm-12_1	7
	13	hoxes_mm-13_1	3
	2B	hoxes_mm-2B_1	9
	7	hoxes_mm-7_1	11
Rattus norvegicus	10	hoxes_mm-10_1	10
	12	hoxes_mm-12_1	2
	3	hoxes_mm-3_1	7
	4	hoxes_mm-4_1	13
	7	hoxes_mm-7_1	9
Sus scrofa	11	hoxes_mm-11_1	3
	12	hoxes_mm-12_1	8
	15	hoxes_mm-15_1	8
	18	hoxes_mm-18_1	12
	5	hoxes_mm-5_1	6
	9	hoxes_mm-9_1	2

Table B.9: Mammalian LHB CTDG using the 7 LHB genes in the human chromosome 19

species	chromosome	cluster	gene duplicates
Bos taurus	15	lhbs-15_1	2
Chlorocebus sabaeus	6	lhbs-6_1	2
Equus caballus	7	lhbs-7_1	2
Gorilla gorilla gorilla	19	lhbs-19_1	3
Homo sapiens	19	lhbs-19_1	7
Macaca fascicularis	19	lhbs-19_1	3
Macaca mulatta	19	lhbs-19_1	2
Papio anubis	19	lhbs-19_1	2
Pongo abelii	19	lhbs-19_1	6

Table B.10: Mammalian protocadherin CTDG using the 52 protocadherin genes in the human chromosome 5

species	chromosome	cluster	gene duplicates
Bos taurus	7	protocadherins-7_1	19
Callithrix jacchus	2	protocadherins-2_1	35
Canis lupus familiaris	2	protocadherins-2_1	15
Capra hircus	7	protocadherins-7_1	28
Chlorocebus sabaeus	23	protocadherins-23_1	18
Equus caballus	14	protocadherins-14_1	36
Felis catus	A1	protocadherins-A1_1	22
Gorilla gorilla gorilla	5	protocadherins-5_1	29
Homo sapiens	5	protocadherins-5_1	53
Macaca fascicularis	6	protocadherins-6_1	22
Macaca mulatta	6	protocadherins-6_1	20
Microtus ochrogaster	18	protocadherins-18_1	31
Monodelphis domestica	1	protocadherins-1_1	25
Mus musculus	18	protocadherins-18_1	24
Nomascus leucogenys	2	protocadherins-2_1	20
Oryctolagus cuniculus	3	protocadherins-3_1	18
Ovis aries	5	protocadherins-5_1	18
Pan troglodytes	5	protocadherins-5_1	22
Papio anubis	6	protocadherins-6_1	20
Pongo abelii	5	protocadherins-5_1	37
Rattus norvegicus	18	protocadherins-18_1	19
Sus scrofa	2	protocadherins-2_1	20
	2	protocadherins-2_2	13

Table B.11: Mammalian protocadherin CTDG using the 26 protocadherin genes in the mouse chromosome 18

species	chromosome	cluster	gene duplicates
Bos taurus	7	protocadherins_mm-7_1	19
	7	protocadherins_mm-7_2	2
Callithrix jacchus	2	protocadherins_mm-2_1	37
Canis lupus familiaris	2	protocadherins_mm-2_1	15
Capra hircus	7	protocadherins_mm-7_1	27
	7	protocadherins_mm-7_2	2
Chlorocebus sabaeus	23	protocadherins_mm-23_1	19
Equus caballus	14	protocadherins_mm-14_1	38
Felis catus	A1	protocadherins_mm-A1_1	24
Gorilla gorilla gorilla	5	protocadherins_mm-5_1	29
	5	protocadherins_mm-5_2	2
Homo sapiens	5	protocadherins_mm-5_1	55
Macaca fascicularis	6	protocadherins_mm-6_1	23
Macaca mulatta	6	protocadherins_mm-6_1	21
Microtus ochrogaster	18	protocadherins_mm-18_1	31
Monodelphis domestica	1	protocadherins_mm-1_1	26
	7	protocadherins_mm-7_1	2
Mus musculus	18	protocadherins_mm-18_1	26
Nomascus leucogenys	2	protocadherins_mm-2_1	22
Oryctolagus cuniculus	3	protocadherins_mm-3_1	22
Ovis aries	5	protocadherins_mm-5_1	17
Pan troglodytes	5	protocadherins_mm-5_1	22
Papio anubis	6	protocadherins_mm-6_1	21
Pongo abelii	5	protocadherins_mm-5_1	38
Rattus norvegicus	18	protocadherins_mm-18_1	20
Sus scrofa	11	protocadherins_mm-11_1	2
	2	protocadherins_mm-2_1	19
	2	protocadherins_mm-2_2	12

Table B.12: Mammalian Prolactin using the 26 Prolactin genes in the mouse chromosome 13

species	chromosome	cluster	gene duplicates
Bos taurus	23	prolactins-23_1	11
Callithrix jacchus	5	prolactins-5_1	3
Capra hircus	23	prolactins-23_1	10
Chlorocebus sabaeus	16	prolactins-16_1	2
Gorilla gorilla gorilla	5	prolactins-5_1	4
Macaca fascicularis	16	prolactins-16_1	5
Macaca mulatta	16	prolactins-16_1	4
Microtus ochrogaster	16	prolactins-16_1	20
Mus musculus	13	prolactins-13_1	24
Ovis aries	20	prolactins-20_1	7
Pan troglodytes	17	prolactins-17_1	4
Papio anubis	16	prolactins-16_1	4
Rattus norvegicus	17	prolactins-17_1	22
	17	prolactins-17_2	5

Table B.13: Mammalian Siglec using the 9 Siglec genes in the human chromosome 19

species	chromosome	cluster	gene duplicates
Bos taurus	18	siglecs-18_1	3
	3	siglecs-3_1	2
Callithrix jacchus	22	siglecs-22_1	5
	22	siglecs-22_2	2
	22	siglecs-22_3	2
Canis lupus familiaris	1	siglecs-1_1	2
Capra hircus	18	siglecs-18_1	6
	18	siglecs-18_2	2
Chlorocebus sabaeus	6	siglecs-6_1	2
Equus caballus	10	siglecs-10_1	18
Felis catus	E2	siglecs-E2_1	7
Gorilla gorilla gorilla	19	siglecs-19_1	6
	19	siglecs-19_2	7
	19	siglecs-19_3	2
	19	siglecs-19_4	2
Homo sapiens	19	siglecs-19_1	7
	19	siglecs-19_2	4
	19	siglecs-19_3	4
	19	siglecs-19_4	2
	19	siglecs-19_5	2
Macaca fascicularis	19	siglecs-19_1	3
	19	siglecs-19_2	2
	19	siglecs-19_3	2
Macaca mulatta	19	siglecs-19_1	4
Microtus ochrogaster	21	siglecs-21_1	2
Monodelphis domestica	4	siglecs-4_1	4
	4	siglecs-4_3	2
Mus musculus	7	siglecs-7_1	6
	7	siglecs-7_2	2
Nomascus leucogenys	10	siglecs-10_1	9
	10	siglecs-10_2	2
Ovis aries	14	siglecs-14_1	4
Pan troglodytes	19	siglecs-19_1	9
	19	siglecs-19_2	2
	2A	siglecs-2A_1	2
Papio anubis	19	siglecs-19_1	4
	19	siglecs-19_2	2
Pongo abelii	19	siglecs-19_1	11
Rattus norvegicus	1	siglecs-1_1	5
	1	siglecs-1_2	3
	13	siglecs-13_1	2
Sus scrofa	6	siglecs-6_1	6
	6	siglecs-6_2	3
	9	siglecs-9_1	2
	9	siglecs-9_2	2

Table B.14: Mammalian Siglec using the 4 Siglec genes in the mouse chromosome 7

species	chromosome	cluster	gene duplicates
Bos taurus	18	siglecs_mm-18_1	5
	3	siglecs_mm-3_1	2
Callithrix jacchus	22	siglecs_mm-22_1	5
	22	siglecs_mm-22_2	2
	22	siglecs_mm-22_3	2
Canis lupus familiaris	1	siglecs_mm-1_1	2
Capra hircus	18	siglecs_mm-18_1	7
	18	siglecs_mm-18_2	2
Chlorocebus sabaeus	6	siglecs_mm-6_1	2
Equus caballus	10	siglecs_mm-10_1	17
	10	siglecs_mm-10_2	2
Felis catus	E2	siglecs_mm-E2_1	7
	F1	siglecs_mm-F1_1	3
Gorilla gorilla gorilla	19	siglecs_mm-19_1	14
	19	siglecs_mm-19_2	2
Homo sapiens	19	siglecs_mm-19_1	7
	19	siglecs_mm-19_2	5
	19	siglecs_mm-19_3	3
	19	siglecs_mm-19_4	2
Macaca fascicularis	19	siglecs_mm-19_1	2
	19	siglecs_mm-19_2	2
Macaca mulatta	19	siglecs_mm-19_1	4
Monodelphis domestica	4	siglecs_mm-4_1	4
Mus musculus	7	siglecs_mm-7_1	7
	7	siglecs_mm-7_2	2
	7	siglecs_mm-7_3	2
Nomascus leucogenys	10	siglecs_mm-10_1	10
	10	siglecs_mm-10_2	2
Ovis aries	14	siglecs_mm-14_1	3
Pan troglodytes	19	siglecs_mm-19_1	9
	19	siglecs_mm-19_2	2
	2A	siglecs_mm-2A_1	2
Papio anubis	19	siglecs_mm-19_1	4
	19	siglecs_mm-19_2	2
Pongo abelii	19	siglecs_mm-19_1	11
Rattus norvegicus	1	siglecs_mm-1_1	5
	1	siglecs_mm-1_2	4
Sus scrofa	10	siglecs_mm-10_1	2
	6	siglecs_mm-6_1	6
	6	siglecs_mm-6_2	3
	9	siglecs_mm-9_1	3

Table B.15: Clustered genes with enhanced expression in adipose tissue

Ensembl	Gene	Gene description
ENSG00000166106	ADAMTS15	ADAM metallopeptidase with thrombospondin type 1 motif, 15
ENSG00000165810	BTNL9	Butyrophilin-like 9
ENSG00000205502	C2CD4B	C2 calcium-dependent domain containing 4B
ENSG00000174807	CD248	CD248 molecule, endosialin
ENSG00000275385	CCL18	Chemokine (C-C motif) ligand 18 (pulmonary and activation-regulated)
ENSG00000163734	CXCL3	Chemokine (C-X-C motif) ligand 3
ENSG00000122756	CNTFR	Ciliary neurotrophic factor receptor
ENSG00000106823	ECM2	Extracellular matrix protein 2, female organ and adipocyte specific
ENSG00000094963	FMO2	Flavin containing monooxygenase 2 (non-functional)
ENSG00000164849	GPR146	G protein-coupled receptor 146
ENSG00000102287	GABRE	Gamma-aminobutyric acid (GABA) A receptor, epsilon
ENSG00000137878	GCOM1	GRINL1A complex locus 1
ENSG00000165478	HEPACAM	Hepatic and glial cell adhesion molecule
ENSG00000196917	HCAR1	Hydroxycarboxylic acid receptor 1
ENSG00000159387	IRX6	Iroquois homeobox 6
ENSG00000205362	MT1A	Metallothionein 1A
ENSG00000205364	MT1M	Metallothionein 1M
ENSG00000157654	PALM2-AKAP2	PALM2-AKAP2 readthrough
ENSG00000135917	SLC19A3	Solute carrier family 19 (thiamine transporter), member 3
ENSG00000120332	TNN	Tenascin N
ENSG00000122691	TWIST1	Twist family bHLH transcription factor 1

Table B.16: Clustered genes with enhanced expression in adrenal gland

Ensembl	Gene	Gene description
ENSG00000085563	ABCB1	ATP-binding cassette, sub-family B (MDR/TAP), member 1
ENSG00000174156	GSTA3	Glutathione S-transferase alpha 3
ENSG00000070886	EPHA8	EPH receptor A8
ENSG00000106004	HOXA5	Homeobox A5
ENSG00000214510	SPINK13	Serine peptidase inhibitor, Kazal type 13 (putative)
ENSG00000112936	C7	Complement component 7
ENSG00000260869	AC002310.13	Uncharacterized protein (ECO:0000313 Ensembl:ENSP00000455659)
ENSG00000123999	INHA	Inhibin, alpha
ENSG00000188488	SERPINA5	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 5
ENSG00000027644	INSRR	Insulin receptor-related receptor
ENSG00000135914	HTR2B	5-hydroxytryptamine (serotonin) receptor 2B, G protein-coupled
ENSG00000197859	ADAMTSL2	ADAMTS-like 2
ENSG00000137878	GCOM1	GRINL1A complex locus 1
ENSG00000112038	OPRM1	Opioid receptor, mu 1
ENSG00000162594	IL23R	Interleukin 23 receptor
ENSG00000142494	SLC47A1	Solute carrier family 47 (multidrug and toxin extrusion), member 1
ENSG00000135111	TBX3	T-box 3
ENSG00000144355	DLX1	Distal-less homeobox 1
ENSG00000106038	EVX1	Even-skipped homeobox 1
ENSG00000203722	RAET1G	Retinoic acid early transcript 1G
ENSG00000115297	TLX2	T-cell leukemia homeobox 2
ENSG00000135409	AMHR2	Anti-Mullerian hormone receptor, type II

Table B.17: Clustered genes with enhanced expression in appendix

Ensembl	Gene	Gene description
ENSG00000160593	AMICA1	Adhesion molecule, interacts with CXADR antigen 1
ENSG00000111729	CLEC4A	C-type lectin domain family 4, member A
ENSG00000166523	CLEC4E	C-type lectin domain family 4, member E
ENSG00000205846	CLEC6A	C-type lectin domain family 6, member A
ENSG00000172243	CLEC7A	C-type lectin domain family 7, member A
ENSG00000178789	CD300LB	Free fatty acid receptor 2
ENSG00000186407	CD300E	CD300e molecule
ENSG00000102970	CCL17	Chemokine (C-C motif) ligand 17
ENSG00000102962	CCL22	Chemokine (C-C motif) ligand 22
ENSG00000163823	CCR1	Chemokine (C-C motif) receptor 1
ENSG00000121807	CCR2	Chemokine (C-C motif) receptor 2
ENSG00000169245	CXCL10	Chemokine (C-X-C motif) ligand 10
ENSG00000169248	CXCL11	Chemokine (C-X-C motif) ligand 11
ENSG00000163735	CXCL5	Chemokine (C-X-C motif) ligand 5
ENSG00000163464	CXCR1	Chemokine (C-X-C motif) receptor 1
ENSG00000180871	CXCR2	Chemokine (C-X-C motif) receptor 2
ENSG00000064886	CHI3L2	Chitinase 3-like 2
ENSG00000203710	CR1	Complement component (3b/4b) receptor 1 (Knops blood group)
ENSG00000197405	C5AR1	Complement component 5a receptor 1
ENSG00000163599	CTLA4	Cytotoxic T-lymphocyte-associated protein 4
ENSG00000127507	EMR2	Egf-like module containing, mucin-like, hormone receptor-like 2
ENSG00000131355	EMR3	Egf-like module containing, mucin-like, hormone receptor-like 3
ENSG00000150337	FCGR1A	Fc fragment of IgG, high affinity Ia, receptor (CD64)
ENSG00000162747	FCGR3B	Fc fragment of IgG, low affinity IIIb, receptor (CD16b)
ENSG00000171051	FPR1	Formyl peptide receptor 1
ENSG00000171049	FPR2	Formyl peptide receptor 2
ENSG00000126262	FFAR2	Free fatty acid receptor 2
ENSG00000185897	FFAR3	Free fatty acid receptor 3
ENSG00000154451	GBP5	Guanylate binding protein 5
ENSG00000140678	ITGAX	Integrin, alpha X (complement component 3 receptor 4 subunit)
ENSG00000104974	LILRA1	Leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 1
ENSG00000187116	LILRA5	Leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 5
ENSG00000244482	LILRA6	Leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 6
ENSG00000104972	LILRB1	Leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 1
ENSG00000131042	LILRB2	Leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 2
ENSG00000204577	LILRB3	Leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 3
ENSG00000019169	MARCO	Macrophage receptor with collagenous structure
ENSG00000262406	MMP12	Matrix metallopeptidase 12 (macrophage elastase)
ENSG00000103313	MEFV	Mediterranean fever
ENSG00000137558	PI15	Peptidase inhibitor 15
ENSG00000027869	SH2D2A	SH2 domain containing 2A
ENSG00000142512	SIGLEC10	Sialic acid binding Ig-like lectin 10
ENSG00000254415	SIGLEC14	Sialic acid binding Ig-like lectin 14

Table B.17: Clustered genes with enhanced expression in appendix

Ensembl	Gene	Gene description
ENSG00000105501	SIGLEC5	Sialic acid binding Ig-like lectin 5
ENSG00000129450	SIGLEC9	Sialic acid binding Ig-like lectin 9
ENSG00000196209	SIRPB2	Signal-regulatory protein beta 2
ENSG00000101916	TLR8	Toll-like receptor 8
ENSG00000124731	TREM1	Triggering receptor expressed on myeloid cells 1
ENSG00000173535	TNFRSF10C	Tumor necrosis factor receptor superfamily, member 10c, decoy without an intracellular domain

Table B.18: Clustered genes with enhanced expression in bone marrow

Ensembl	Gene	Gene description
ENSG00000128383	APOBEC3A	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3A
ENSG00000179750	APOBEC3B	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3B
ENSG00000140379	BCL2A1	BCL2-related protein A1
ENSG00000172322	CLEC12A	C-type lectin domain family 12, member A
ENSG00000166523	CLEC4E	C-type lectin domain family 4, member E
ENSG00000172243	CLEC7A	C-type lectin domain family 7, member A
ENSG00000007129	CEACAM21	Carcinoembryonic antigen-related cell adhesion molecule 21
ENSG00000186407	CD300E	CD300e molecule
ENSG00000110848	CD69	CD69 molecule
ENSG00000277632	CCL3	Chemokine (C-C motif) ligand 3
ENSG00000275302	CCL4	Chemokine (C-C motif) ligand 4
ENSG00000276070	CCL4L1	Chemokine (C-C motif) ligand 4-like 1
ENSG00000263020	CSNK2B-LY6G5B-1181	Chimera CSNK2B-LY6G5B splicing isoform 1181 (ECO:0000313 EMBL:CCI79697.1); Uncharacterized protein (ECO:0000313 Ensembl:ENSP00000365040)
ENSG00000133048	CHI3L1	Chitinase 3-like 1 (cartilage glycoprotein-39)
ENSG00000213937	CLDN9	Claudin 9
ENSG00000197721	CR1L	Complement component (3b/4b) receptor 1-like
ENSG00000186529	CYP4F3	Cytochrome P450, family 4, subfamily F, polypeptide 3
ENSG00000137441	FGFBP2	Fibroblast growth factor binding protein 2
ENSG00000171051	FPR1	Formyl peptide receptor 1
ENSG00000126266	FFAR1	Free fatty acid receptor 1
ENSG00000126262	FFAR2	Free fatty acid receptor 2
ENSG00000121742	GJB6	Gap junction protein, beta 6, 30kDa
ENSG00000197465	GYPE	Glycophorin E (MNS blood group)
ENSG00000060558	GNA15	Guanine nucleotide binding protein (G protein), alpha 15 (Gq class)
ENSG00000130656	HBZ	Hemoglobin, zeta
ENSG00000113749	HRH2	Histamine receptor H2
ENSG00000278463	HIST1H2AB	Histone cluster 1, H2ab
ENSG00000196866	HIST1H2AD	Histone cluster 1, H2ad
ENSG00000196747	HIST1H2AI	Histone cluster 1, H2ai
ENSG00000276368	HIST1H2AJ	Histone cluster 1, H2aj
ENSG00000278677	HIST1H2AM	Histone cluster 1, H2am
ENSG00000274290	HIST1H2BE	Histone cluster 1, H2be
ENSG00000278588	HIST1H2BI	Histone cluster 1, H2bi
ENSG00000273703	HIST1H2BM	Histone cluster 1, H2bm
ENSG00000274641	HIST1H2BO	Histone cluster 1, H2bo
ENSG00000278272	HIST1H3C	Histone cluster 1, H3c
ENSG00000197409	HIST1H3D	Histone cluster 1, H3d
ENSG00000277775	HIST1H3F	Histone cluster 1, H3f
ENSG00000275379	HIST1H3I	Histone cluster 1, H3i
ENSG00000197153	HIST1H3J	Histone cluster 1, H3j
ENSG00000278637	HIST1H4A	Histone cluster 1, H4a
ENSG00000278705	HIST1H4B	Histone cluster 1, H4b
ENSG00000274618	HIST1H4F	Histone cluster 1, H4f
ENSG00000197238	HIST1H4J	Histone cluster 1, H4j
ENSG00000184270	HIST2H2AB	Histone cluster 2, H2ab
ENSG00000184260	HIST2H2AC	Histone cluster 2, H2ac
ENSG00000255398	HCAR3	Hydroxycarboxylic acid receptor 3

Table B.18: Clustered genes with enhanced expression in bone marrow

Ensembl	Gene	Gene description
ENSG00000169896	ITGAM	Integrin, alpha M (complement component 3 receptor 3 subunit)
ENSG00000105371	ICAM4	Intercellular adhesion molecule 4 (Landsteiner-Wiener blood group)
ENSG00000125538	IL1B	Interleukin 1, beta
ENSG00000115607	IL18RAP	Interleukin 18 receptor accessory protein
ENSG00000170442	KRT86	Keratin 86
ENSG00000167633	KIR3DL1	Killer cell immunoglobulin-like receptor, three domains, long cytoplasmic tail, 1
ENSG00000125498	KIR2DL1	Killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 1
ENSG00000243772	KIR2DL3	Killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 3
ENSG00000116852	KIF21B	Kinesin family member 21B
ENSG00000239998	LILRA2	Leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 2
ENSG00000187116	LILRA5	Leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 5
ENSG00000131042	LILRB2	Leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 2
ENSG00000204424	LY6G6F	Lymphocyte antigen 6 complex, locus G6F
ENSG00000238243	OR2W3	Olfactory receptor, family 2, subfamily W, member 3
ENSG00000170909	OSCAR	Osteoclast associated, immunoglobulin-like receptor
ENSG00000141506	PIK3R5	Phosphoinositide-3-kinase, regulatory subunit 5
ENSG00000109272	PF4V1	Platelet factor 4 variant 1
ENSG00000105501	SIGLEC5	Sialic acid binding Ig-like lectin 5
ENSG00000268500	SIGLEC5	Sialic acid binding Ig-like lectin 5
ENSG00000059804	SLC2A3	Solute carrier family 2 (facilitated glucose transporter), member 3
ENSG00000124731	TREM1	Triggering receptor expressed on myeloid cells 1
ENSG00000112195	TREML2	Triggering receptor expressed on myeloid cells-like 2
ENSG00000204613	TRIM10	Tripartite motif containing 10
ENSG00000232810	TNF	Tumor necrosis factor
ENSG00000276043	UHRF1	Ubiquitin-like with PHD and ring finger domains 1
ENSG00000254979	RP11-872D17.8	Uncharacterized protein (ECO:0000313 Ensembl:ENSP00000431536)
ENSG00000093134	VNN3	Vanin 3
ENSG00000229676	ZNF492	Zinc finger protein 492

Table B.19: Clustered genes with enhanced expression in cerebral cortex

Ensembl	Gene	Gene description
ENSG00000149305	HTR3B	5-hydroxytryptamine (serotonin) receptor 3B, ionotropic
ENSG00000008300	CELSR3	Cadherin, EGF LAG seven-pass G-type receptor 3
ENSG00000075429	CACNG5	Calcium channel, voltage-dependent, gamma subunit 5
ENSG00000185933	CALHM1	Calcium homeostasis modulator 1
ENSG00000105270	CLIP3	CAP-GLY domain containing linker protein 3
ENSG00000183196	CHST6	Carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 6
ENSG00000168329	CX3CR1	Chemokine (C-X3-C motif) receptor 1
ENSG00000133048	CHI3L1	Chitinase 3-like 1 (cartilage glycoprotein-39)
ENSG00000144355	DLX1	Distal-less homeobox 1
ENSG00000269502	DMRTC1	DMRT-like family C1
ENSG00000184911	DMRTC1B	DMRT-like family C1B
ENSG00000184845	DRD1	Dopamine receptor D1
ENSG00000116661	FBXO2	F-box protein 2
ENSG00000166206	GABRB3	Gamma-aminobutyric acid (GABA) A receptor, beta 3
ENSG00000165478	HEPACAM	Hepatic and glial cell adhesion molecule
ENSG00000167178	ISLR2	Immunoglobulin superfamily containing leucine-rich repeat 2
ENSG00000167755	KLK6	Kallikrein-related peptidase 6
ENSG00000213424	KRT222	Keratin 222
ENSG00000205809	KLRC2	Killer cell lectin-like receptor subfamily C, member 2
ENSG00000205810	KLRC3	Killer cell lectin-like receptor subfamily C, member 3
ENSG00000183908	LRRC55	Leucine rich repeat containing 55
ENSG00000033122	LRRC7	Leucine rich repeat containing 7
ENSG00000144893	MED12L	Mediator complex subunit 12-like
ENSG00000187243	MAGED4B	Melanoma antigen family D, 4B
ENSG00000108852	MPP2	Membrane protein, palmitoylated 2 (MAGUK p55 subfamily member 2)
ENSG00000104722	NEFM	Neurofilament, medium polypeptide
ENSG00000125820	NKX2-2	NK2 homeobox 2
ENSG00000177432	NAP1L5	Nucleosome assembly protein 1-like 5
ENSG00000196368	NUDT11	Nudix (nucleoside diphosphate linked moiety X)-type motif 11
ENSG00000105088	OLFM2	Olfactomedin 2
ENSG00000196071	OR2L13	Olfactory receptor, family 2, subfamily L, member 13
ENSG00000125510	OPRL1	Opiate receptor-like 1
ENSG00000082556	OPRK1	Opioid receptor, kappa 1
ENSG00000243444	PALM2	Paralemmin 2
ENSG00000235961	PNMA6A	Paraneoplastic Ma antigen family member 6A
ENSG00000182013	PNMAL1	Paraneoplastic Ma antigen family-like 1
ENSG00000059915	PSD	Pleckstrin and Sec7 domain containing
ENSG00000184388	PABPC1L2B	Poly(A) binding protein, cytoplasmic 1-like 2B
ENSG00000182870	GALNT9	Polypeptide N-acetylgalactosaminyltransferase 9
ENSG00000153233	PTPRR	Protein tyrosine phosphatase, receptor type, R
ENSG00000250120	PCDHA10	Protocadherin alpha 10
ENSG00000249158	PCDHA11	Protocadherin alpha 11
ENSG00000204967	PCDHA4	Protocadherin alpha 4
ENSG00000081842	PCDHA6	Protocadherin alpha 6
ENSG00000204963	PCDHA7	Protocadherin alpha 7
ENSG00000243232	PCDHAC2	Protocadherin alpha subfamily C, 2
ENSG00000177839	PCDHB9	Protocadherin beta 9
ENSG00000204956	PCDHGA1	Protocadherin gamma subfamily A, 1
ENSG00000253846	PCDHGA10	Protocadherin gamma subfamily A, 10
ENSG00000254245	PCDHGA3	Protocadherin gamma subfamily A, 3
ENSG00000253537	PCDHGA7	Protocadherin gamma subfamily A, 7

Table B.19: Clustered genes with enhanced expression in cerebral cortex

Ensembl	Gene	Gene description
ENSG00000253305	PCDHGB6	Protocadherin gamma subfamily B, 6
ENSG00000240184	PCDHGC3	Protocadherin gamma subfamily C, 3
ENSG00000169313	P2RY12	Purinergic receptor P2Y, G-protein coupled, 12
ENSG00000175077	RTP1	Receptor (chemosensory) transporter protein 1
ENSG00000117152	RGS4	Regulator of G-protein signaling 4
ENSG00000242173	ARHGDIG	Rho GDP dissociation inhibitor (GDI) gamma
ENSG00000212864	RNF208	Ring finger protein 208
ENSG00000249884	RNF103-CHMP3	RNF103-CHMP3 readthrough
ENSG00000198794	SCAMP5	Secretory carrier membrane protein 5
ENSG00000153253	SCN3A	Sodium channel, voltage-gated, type III, alpha subunit
ENSG00000132164	SLC6A11	Solute carrier family 6 (neurotransmitter transporter), member 11
ENSG00000173898	SPTBN2	Spectrin, beta, non-erythrocytic 2
ENSG00000184905	TCEAL2	Transcription elongation factor A (SII)-like 2
ENSG00000182916	TCEAL7	Transcription elongation factor A (SII)-like 7
ENSG00000095970	TREM2	Triggering receptor expressed on myeloid cells 2
ENSG00000092445	TYRO3	TYRO3 protein tyrosine kinase
ENSG00000248751	RP1-130H16.18	Uncharacterized protein (ECO:0000313 Ensembl:ENSP00000400319)
ENSG00000255641	NKG2-E	Uncharacterized protein (ECO:0000313 Ensembl:ENSP00000437563)
ENSG00000169884	WNT10B	Wingless-type MMTV integration site family, member 10B
ENSG00000152977	ZIC1	Zic family member 1
ENSG00000174963	ZIC4	Zic family member 4
ENSG00000152467	ZSCAN1	Zinc finger and SCAN domain containing 1
ENSG00000276302	RP5-874C20.8	Predicted intracellular proteins

Table B.20: Clustered genes with enhanced expression in colon

Ensembl	Gene	Gene description
ENSG00000104267	CA2	Carbonic anhydrase II
ENSG00000172828	CES3	Carboxylesterase 3
ENSG00000079385	CEACAM1	Carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein)
ENSG00000105388	CEACAM5	Carcinoembryonic antigen-related cell adhesion molecule 5
ENSG00000275718	CCL15	Chemokine (C-C motif) ligand 15
ENSG00000166920	C15orf48	Chromosome 15 open reading frame 48
ENSG00000165215	CLDN3	Claudin 3
ENSG00000237289	CKMT1B	Creatine kinase, mitochondrial 1B
ENSG00000188833	ENTPD8	Ectonucleoside triphosphate diphosphohydrolase 8
ENSG00000186564	FOXD2	Forkhead box D2
ENSG00000171124	FUT3	Fucosyltransferase 3 (galactoside 3(4)-L-fucosyltransferase, Lewis blood group)
ENSG00000089356	FXYD3	FXYD domain containing ion transport regulator 3
ENSG00000044012	GUCA2B	Guanylate cyclase activator 2B (uroguanylin)
ENSG00000113249	HAVCR1	Hepatitis A virus cellular receptor 1
ENSG00000170689	HOXB9	Homeobox B9
ENSG00000128713	HOXD11	Homeobox D11
ENSG00000170178	HOXD12	Homeobox D12
ENSG00000179914	ITLN1	Intelectin 1 (galactofuranose binding)
ENSG00000174562	KLK15	Kallikrein-related peptidase 15
ENSG00000171431	KRT20	Keratin 20
ENSG00000114248	LRRC31	Leucine rich repeat containing 31
ENSG00000177984	LCN15	Lipocalin 15
ENSG00000183034	OTOP2	Otopetrin 2
ENSG00000188257	PLA2G2A	Phospholipase A2, group IIA (platelets, synovial fluid)
ENSG00000162896	PIGR	Polymeric immunoglobulin receptor
ENSG00000178172	SPINK6	Serine peptidase inhibitor, Kazal type 6
ENSG00000115616	SLC9A2	Solute carrier family 9, subfamily A (NHE2, cation proton antiporter 2), member 2
ENSG00000160180	TFF3	Trefoil factor 3 (intestinal)
ENSG00000204616	TRIM31	Tripartite motif containing 31
ENSG00000242366	UGT1A1	UDP glucuronosyltransferase 1 family, polypeptide A1
ENSG00000267881	CEA	Uncharacterized protein (ECO:0000313 Ensembl:ENSP00000469926)

Table B.21: Clustered genes with enhanced expression in duodenum

Ensembl	Gene	Gene description
ENSG00000186038	HTR3E	5-hydroxytryptamine (serotonin) receptor 3E, ionotropic
ENSG00000164270	HTR4	5-hydroxytryptamine (serotonin) receptor 4, G protein-coupled
ENSG00000198074	AKR1B10	Aldo-keto reductase family 1, member B10 (aldose reductase)
ENSG00000227471	AKR1B15	Aldo-keto reductase family 1, member B15
ENSG00000162482	AKR7A3	Aldo-keto reductase family 7, member A3 (aflatoxin aldehyde reductase)
ENSG00000114771	AADAC	Arylacetamide deacetylase
ENSG00000275718	CCL15	Chemokine (C-C motif) ligand 15
ENSG00000258941	RP11-407N17.3	CTAGE family member 5 (ECO:0000313 Ensembl:ENSP00000452252)
ENSG00000108242	CYP2C18	Cytochrome P450, family 2, subfamily C, polypeptide 18
ENSG00000165841	CYP2C19	Cytochrome P450, family 2, subfamily C, polypeptide 19
ENSG00000167600	CYP2S1	Cytochrome P450, family 2, subfamily S, polypeptide 1
ENSG00000106258	CYP3A5	Cytochrome P450, family 3, subfamily A, polypeptide 5
ENSG00000186204	CYP4F12	Cytochrome P450, family 4, subfamily F, polypeptide 12
ENSG00000188833	ENTPD8	Ectonucleoside triphosphate diphosphohydrolase 8
ENSG00000205436	EXOC3L4	Exocyst complex component 3-like 4
ENSG00000165140	FBP1	Fructose-1,6-bisphosphatase 1
ENSG00000176920	FUT2	Fucosyltransferase 2 (secretor status included)
ENSG00000073605	GSDMB	Gasdermin B
ENSG00000156510	HKDC1	Hexokinase domain containing 1
ENSG00000257743	RP11-1220K2.2	Homo sapiens maltase-glucoamylase (alpha-glucosidase) (LOC93432), mRNA.
ENSG00000133328	HRASLS2	HRAS-like suppressor 2
ENSG00000171431	KRT20	Keratin 20
ENSG00000179528	LBX2	Ladybird homeobox 2
ENSG00000177984	LCN15	Lipocalin 15
ENSG00000150526	MIA2	Melanoma inhibitory activity 2
ENSG00000169688	MT1B	Metallothionein 1B
ENSG00000169894	MUC3A	Mucin 3A, cell surface associated
ENSG00000162896	PIGR	Polymeric immunoglobulin receptor
ENSG00000165828	PRAP1	Proline-rich acidic protein 1
ENSG00000172023	REG1B	Regenerating islet-derived 1 beta
ENSG00000242173	ARHGDIG	Rho GDP dissociation inhibitor (GDI) gamma
ENSG00000146039	SLC17A4	Solute carrier family 17, member 4
ENSG00000142583	SLC2A5	Solute carrier family 2 (facilitated glucose/fructose transporter), member 5
ENSG00000197165	SULT1A2	Sulfotransferase family, cytosolic, 1A, phenol-preferring, member 2
ENSG00000196228	SULT1C3	Sulfotransferase family, cytosolic, 1C, member 3
ENSG00000204610	TRIM15	Tripartite motif containing 15
ENSG00000204614	TRIM40	Tripartite motif containing 40
ENSG00000135220	UGT2A3	UDP glucuronosyltransferase 2 family, polypeptide A3
ENSG00000142539	CTD-2545M3.6	UDP glucuronosyltransferase 2 family, polypeptide A3

Table B.22: Clustered genes with enhanced expression in endometrium

Ensembl	Gene	Gene description
ENSG00000128710	HOXD10	Homeobox D10
ENSG00000143768	LEFTY2	Left-right determination factor 2
ENSG00000253293	HOXA10	Homeobox A10
ENSG00000257184	HOXA10-HOXA9	Uncharacterized protein (ECO:0000313 Ensembl:ENSP00000421799)
ENSG00000170166	HOXD4	Homeobox D4
ENSG00000128709	HOXD9	Homeobox D9
ENSG00000122691	TWIST1	Twist family bHLH transcription factor 1
ENSG00000078098	FAP	Fibroblast activation protein, alpha
ENSG00000124939	SCGB2A1	Secretoglobin, family 2A, member 1
ENSG00000135914	HTR2B	5-hydroxytryptamine (serotonin) receptor 2B, G protein-coupled
ENSG00000128713	HOXD11	Homeobox D11

Table B.23: Clustered genes with enhanced expression in esophagus

Ensembl	Gene	Gene description
ENSG00000108602	ALDH3A1	Aldehyde dehydrogenase 3 family, member A1
ENSG00000132746	ALDH3B2	Aldehyde dehydrogenase 3 family, member B2
ENSG00000179593	ALOX15B	Arachidonate 15-lipoxygenase, type B
ENSG00000178363	CALML3	Calmodulin-like 3
ENSG00000158477	CD1A	CD1a molecule
ENSG00000180871	CXCR2	Chemokine (C-X-C motif) receptor 2
ENSG00000137975	CLCA2	Chloride channel accessory 2
ENSG00000134755	DSC2	Desmocollin 2
ENSG00000137857	DUOX1	Dual oxidase 1
ENSG00000140254	DUOXA1	Dual oxidase maturation factor 1
ENSG00000135373	EHF	Ets homologous factor
ENSG00000094963	FMO2	Flavin containing monooxygenase 2 (non-functional)
ENSG00000171124	FUT3	Fucosyltransferase 3 (galactoside 3(4)-L-fucosyltransferase, Lewis blood group)
ENSG00000156413	FUT6	Fucosyltransferase 6 (alpha (1,3) fucosyltransferase)
ENSG00000153294	GPR115	G protein-coupled receptor 115
ENSG00000138271	GPR87	G protein-coupled receptor 87
ENSG00000188910	GJB3	Gap junction protein, beta 3, 31kDa
ENSG00000121742	GJB6	Gap junction protein, beta 6, 30kDa
ENSG00000060558	GNA15	Guanine nucleotide binding protein (G protein), alpha 15 (Gq class)
ENSG00000069812	HES2	Hes family bHLH transcription factor 2
ENSG00000277224	HIST1H2BF	Histone cluster 1, H2bf
ENSG00000277157	HIST1H4D	Histone cluster 1, H4d
ENSG00000136689	IL1RN	Interleukin 1 receptor antagonist
ENSG00000129451	KLK10	Kallikrein-related peptidase 10
ENSG00000167757	KLK11	Kallikrein-related peptidase 11
ENSG00000186474	KLK12	Kallikrein-related peptidase 12
ENSG00000167755	KLK6	Kallikrein-related peptidase 6
ENSG00000167916	KRT24	Keratin 24
ENSG00000186442	KRT3	Keratin 3
ENSG00000180155	LYNX1	Ly6/neurotoxin 1
ENSG00000197353	LYPD2	LY6/PLAUR domain containing 2
ENSG00000205364	MT1M	Metallothionein 1M
ENSG00000102891	MT4	Metallothionein 4
ENSG00000149573	MPZL2	Myelin protein zero-like 2
ENSG00000133020	MYH8	Myosin, heavy chain 8, skeletal muscle, perinatal
ENSG00000182938	OTOP3	Otopetrin 3
ENSG00000163218	PGLYRP4	Peptidoglycan recognition protein 4
ENSG00000243708	PLA2G4B	Phospholipase A2, group IVB (cytosolic)
ENSG00000143217	PVRL4	Poliovirus receptor-related 4
ENSG00000005001	PRSS22	Protease, serine, 22
ENSG00000101213	PTK6	Protein tyrosine kinase 6
ENSG00000175591	P2RY2	Purinergic receptor P2Y, G-protein coupled, 2
ENSG00000164520	RAET1E	Retinoic acid early transcript 1E
ENSG00000203722	RAET1G	Retinoic acid early transcript 1G
ENSG00000165799	RNASE7	Ribonuclease, RNase A family, 7
ENSG00000189334	S100A14	S100 calcium binding protein A14
ENSG00000188643	S100A16	S100 calcium binding protein A16
ENSG00000196754	S100A2	S100 calcium binding protein A2
ENSG00000197632	SERPINB2	Serpin peptidase inhibitor, clade B (ovalbumin), member 2
ENSG00000168447	SCNN1B	Sodium channel, non-voltage-gated 1, beta subunit

Table B.23: Clustered genes with enhanced expression in esophagus

Ensembl	Gene	Gene description
ENSG00000132164	SLC6A11	Solute carrier family 6 (neurotransmitter transporter), member 11
ENSG00000205413	SAMD9	Sterile alpha motif domain containing 9
ENSG00000104055	TGM5	Transglutaminase 5
ENSG00000131015	ULBP2	UL16 binding protein 2
ENSG00000248751	RP1-130H16.18	Uncharacterized protein (ECO:0000313 Ensembl:ENSP00000400319)
ENSG00000267881	CEA	Uncharacterized protein (ECO:0000313 Ensembl:ENSP00000469926)
ENSG00000186806	VSIG10L	V-set and immunoglobulin domain containing 10 like

Table B.24: Clustered genes with enhanced expression in fallopian tube

Ensembl	Gene	Gene description
ENSG00000147117	ZNF157	Zinc finger protein 157
ENSG00000173467	AGR3	Anterior gradient 3
ENSG00000139330	KERA	Keratocan
ENSG00000235098	ANKRD65	Ankyrin repeat domain 65
ENSG00000058085	LAMC2	Laminin, gamma 2
ENSG00000175920	DOK7	Docking protein 7
ENSG00000163286	ALPPL2	Alkaline phosphatase, placental-like 2
ENSG00000120075	HOXB5	Homeobox B5
ENSG00000174156	GSTA3	Glutathione S-transferase alpha 3
ENSG00000180509	KCNE1	Potassium voltage-gated channel, Isk-related family, member 1
ENSG00000110195	FOLR1	Folate receptor 1 (adult)
ENSG00000180881	CAPS2	Calcyphosine 2
ENSG00000217236	SP9	Sp9 transcription factor
ENSG00000172000	ZNF556	Zinc finger protein 556
ENSG00000145002	FAM86B2	Family with sequence similarity 86, member B2
ENSG00000101443	WFDC2	WAP four-disulfide core domain 2
ENSG00000128652	HOXD3	Homeobox D3
ENSG00000168077	SCARA3	Scavenger receptor class A, member 3
ENSG00000168515	SCGB1D1	Secretoglobin, family 1D, member 1
ENSG00000124939	SCGB2A1	Secretoglobin, family 2A, member 1
ENSG00000182742	HOXB4	Homeobox B4
ENSG00000070731	ST6GALNAC2	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2
ENSG00000198889	DCAF12L1	DDB1 and CUL4 associated factor 12-like 1
ENSG00000198354	DCAF12L2	DDB1 and CUL4 associated factor 12-like 2
ENSG00000107014	RLN2	Relaxin 2
ENSG00000172789	HOXC5	Homeobox C5
ENSG00000128645	HOXD1	Homeobox D1
ENSG00000108511	HOXB6	Homeobox B6
ENSG00000197208	SLC22A4	Solute carrier family 22 (organic cation/zwitterion transporter), member 4
ENSG00000167755	KLK6	Kallikrein-related peptidase 6
ENSG00000198353	HOXC4	Homeobox C4
ENSG00000174586	ZNF497	Zinc finger protein 497
ENSG00000164393	GPR111	G protein-coupled receptor 111
ENSG00000259120	SMIM6	Small integral membrane protein 6
ENSG00000174326	SLC16A11	Solute carrier family 16, member 11
ENSG00000184925	LCN12	Lipocalin 12
ENSG00000197576	HOXA4	Homeobox A4

Table B.25: Clustered genes with enhanced expression in gallbladder

Ensembl	Gene	Gene description
ENSG00000138316	ADAMTS14	ADAM metallopeptidase with thrombospondin type 1 motif, 14
ENSG00000227471	AKR1B15	Aldo-keto reductase family 1, member B15
ENSG00000100739	BDKRB1	Bradykinin receptor B1
ENSG00000163751	CPA3	Carboxypeptidase A3 (mast cell)
ENSG00000138615	CILP	Cartilage intermediate layer protein, nucleotide pyrophosphohydrolase
ENSG00000124875	CXCL6	Chemokine (C-X-C motif) ligand 6
ENSG00000164220	F2RL2	Coagulation factor II (thrombin) receptor-like 2
ENSG00000168542	COL3A1	Collagen, type III, alpha 1
ENSG00000171903	CYP4F11	Cytochrome P450, family 4, subfamily F, polypeptide 11
ENSG00000140279	DUOX2	Dual oxidase 2
ENSG00000176678	FOXL1	Forkhead box L1
ENSG00000213424	KRT222	Keratin 222
ENSG00000139330	KERA	Keratocan
ENSG00000100079	LGALS2	Lectin, galactoside-binding, soluble, 2
ENSG00000139329	LUM	Lumican
ENSG00000187123	LYPD6	LY6/PLAUR domain containing 6
ENSG00000196611	MMP1	Matrix metallopeptidase 1 (interstitial collagenase)
ENSG00000137673	MMP7	Matrix metallopeptidase 7 (matrilysin, uterine)
ENSG00000106809	OGN	Osteoglycin
ENSG00000204967	PCDHA4	Protocadherin alpha 4
ENSG00000139988	RDH12	Retinol dehydrogenase 12 (all-trans/9-cis/11-cis)
ENSG00000124233	SEMG1	Semenogelin I
ENSG00000100665	SERPINA4	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 4
ENSG00000198075	SULT1C4	Sulfotransferase family, cytosolic, 1C, member 4
ENSG00000172236	TPSAB1	Tryptase alpha/beta 1
ENSG00000197253	TPSB2	Tryptase beta 2 (gene/pseudogene)
ENSG00000196620	UGT2B15	UDP glucuronosyltransferase 2 family, polypeptide B15
ENSG00000248235	AC037459.4	Uncharacterized protein (ECO:0000313 Ensembl:ENSP00000428700)

Table B.26: Clustered genes with enhanced expression in heart muscle

Ensembl	Gene	Gene description
ENSG00000102287	GABRE	Gamma-aminobutyric acid (GABA) A receptor, epsilon
ENSG00000112276	BVES	Blood vessel epicardial substance
ENSG00000175920	DOK7	Docking protein 7
ENSG00000157654	PALM2-AKAP2	PALM2-AKAP2 readthrough
ENSG00000160678	S100A1	S100 calcium binding protein A1
ENSG00000249884	RNF103-CHMP3	RNF103-CHMP3 readthrough
ENSG00000159387	IRX6	Iroquois homeobox 6
ENSG00000109846	CRYAB	Crystallin, alpha B
ENSG00000170276	HSPB2	Heat shock 27kDa protein 2
ENSG00000127472	PLA2G5	Phospholipase A2, group V
ENSG00000126890	CTAG2	Cancer/testis antigen 2
ENSG00000121743	GJA3	Gap junction protein, alpha 3, 46kDa
ENSG00000137821	LRRC49	Leucine rich repeat containing 49

Table B.27: Clustered genes with enhanced expression in kidney

Ensembl	Gene	Gene description
ENSG00000177465	ACOT4	Acyl-CoA thioesterase 4
ENSG00000205669	ACOT6	Acyl-CoA thioesterase 6
ENSG00000197859	ADAMTSL2	ADAMTS-like 2
ENSG00000157399	ARSE	Arylsulfatase E (chondrodysplasia punctata 1)
ENSG00000185758	CLDN24	Claudin 24
ENSG00000156284	CLDN8	Claudin 8
ENSG00000169031	COL4A3	Collagen, type IV, alpha 3 (Goodpasture antigen)
ENSG00000258941	RP11-407N17.3	CTAGE family member 5 ECO:0000313—Ensembl:ENSP00000452252
ENSG00000173253	DMRT2	Doublesex and mab-3 related transcription factor 2
ENSG00000162897	FCAMR	Fc receptor, IgA, IgM, high affinity
ENSG00000076258	FMO4	Flavin containing monooxygenase 4
ENSG00000156413	FUT6	Fucosyltransferase 6 (alpha (1,3) fucosyltransferase)
ENSG00000243955	GSTA1	Glutathione S-transferase alpha 1
ENSG00000244067	GSTA2	Glutathione S-transferase alpha 2
ENSG00000113249	HAVCR1	Hepatitis A virus cellular receptor 1
ENSG00000078399	HOXA9	Homeobox A9
ENSG00000108511	HOXB6	Homeobox B6
ENSG00000120068	HOXB8	Homeobox B8
ENSG00000180818	HOXC10	Homeobox C10
ENSG00000128713	HOXD11	Homeobox D11
ENSG00000175879	HOXD8	Homeobox D8
ENSG00000027644	INSRR	Insulin receptor-related receptor
ENSG00000188883	KLRG2	Killer cell lectin-like receptor subfamily G, member 2
ENSG00000144214	LYG1	Lysozyme G-like 1
ENSG00000125144	MT1G	Metallothionein 1G
ENSG00000161270	NPHS1	Nephrosis 1, congenital, Finnish type (nephrin)
ENSG00000100344	PNPLA3	Patatin-like phospholipase domain containing 3
ENSG00000124157	SEMG2	Semenogelin II
ENSG00000166828	SCNN1G	Sodium channel, non-voltage-gated 1, gamma subunit
ENSG00000132874	SLC14A2	Solute carrier family 14 (urea transporter), member 2
ENSG00000145217	SLC26A1	Solute carrier family 26 (anion exchanger), member 1
ENSG00000142494	SLC47A1	Solute carrier family 47 (multidrug and toxin extrusion), member 1
ENSG00000173930	SLCO4C1	Solute carrier organic anion transporter family, member 4C1
ENSG00000183018	SPNS2	Spinster homolog 2 (Drosophila)
ENSG00000198203	SULT1C2	Sulfotransferase family, cytosolic, 1C, member 2
ENSG00000127412	TRPV5	Transient receptor potential cation channel, subfamily V, member 5
ENSG00000121236	TRIM6	Tripartite motif containing 6
ENSG00000167165	UGT1A6	UDP glucuronosyltransferase 1 family, polypeptide A6
ENSG00000145626	UGT3A1	UDP glycosyltransferase 3 family, polypeptide A1
ENSG00000158955	WNT9B	Wingless-type MMTV integration site family, member 9B
ENSG00000203995	ZYG11A	Zyg-11 family member A, cell cycle regulator

Table B.28: Clustered genes with enhanced expression in liver

Ensembl	Gene	Gene description
ENSG00000262660	SLC25A10	39S ribosomal protein L12, mitochondrial ECO:0000313—Ensembl:ENSP00000439565; Mitochondrial dicarboxylate carrier ECO:0000313—Ensembl:ENSP00000461324; cDNA FLJ60124, highly similar to Mitochondrial dicarboxylate carrier ECO:0000313—EMBL:BAG59593.1
ENSG00000166743	ACSM1	Acyl-CoA synthetase medium-chain family member 1
ENSG00000005187	ACSM3	Acyl-CoA synthetase medium-chain family member 3
ENSG00000183549	ACSM5	Acyl-CoA synthetase medium-chain family member 5
ENSG00000205669	ACOT6	Acyl-CoA thioesterase 6
ENSG00000160323	ADAMTS13	ADAM metallopeptidase with thrombospondin type 1 motif, 13
ENSG00000118137	APOA1	Apolipoprotein A-I
ENSG00000114771	AADAC	Arylacetamide deacetylase
ENSG00000157399	ARSE	Arylsulfatase E (chondrodysplasia punctata 1)
ENSG00000091262	ABCC6	ATP-binding cassette, sub-family C (CFTR/MRP), member 6
ENSG00000081041	CXCL2	Chemokine (C-X-C motif) ligand 2
ENSG00000163347	CLDN1	Claudin 1
ENSG00000126218	F10	Coagulation factor X
ENSG00000088926	F11	Coagulation factor XI
ENSG00000166278	C2	Complement component 2
ENSG00000244731	C4A	Complement component 4A (Rodgers blood group)
ENSG00000224389	C4B	Complement component 4B (Chido blood group)
ENSG00000258941	RP11-407N17.3	CTAGE family member 5 ECO:0000313—Ensembl:ENSP00000452252
ENSG00000108242	CYP2C18	Cytochrome P450, family 2, subfamily C, polypeptide 18
ENSG00000165841	CYP2C19	Cytochrome P450, family 2, subfamily C, polypeptide 19
ENSG00000160870	CYP3A7	Cytochrome P450, family 3, subfamily A, polypeptide 7
ENSG00000171903	CYP4F11	Cytochrome P450, family 4, subfamily F, polypeptide 11
ENSG00000186529	CYP4F3	Cytochrome P450, family 4, subfamily F, polypeptide 3
ENSG00000165140	FBP1	Fructose-1,6-bisphosphatase 1
ENSG00000243955	GSTA1	Glutathione S-transferase alpha 1
ENSG00000244067	GSTA2	Glutathione S-transferase alpha 2
ENSG00000213931	HBE1	Hemoglobin, epsilon 1
ENSG00000114378	HYAL1	Hyaluronoglucosaminidase 1
ENSG00000169688	MT1B	Metallothionein 1B
ENSG00000100344	PNPLA3	Patatin-like phospholipase domain containing 3
ENSG00000165828	PRAP1	Proline-rich acidic protein 1
ENSG00000100003	SEC14L2	SEC14-like 2 (S. cerevisiae)
ENSG00000100012	SEC14L3	SEC14-like 3 (S. cerevisiae)
ENSG00000100665	SERPINA4	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 4
ENSG00000188488	SERPINA5	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 5
ENSG00000174327	SLC16A13	Solute carrier family 16, member 13
ENSG00000167165	UGT1A6	UDP glucuronosyltransferase 1 family, polypeptide A6
ENSG00000196620	UGT2B15	UDP glucuronosyltransferase 2 family, polypeptide B15
ENSG00000145626	UGT3A1	UDP glycosyltransferase 3 family, polypeptide A1
ENSG00000093134	VNN3	Vanin 3

Table B.29: Clustered genes with enhanced expression in lung

Ensembl	Gene	Gene description
ENSG00000178084	HTR3C	5-hydroxytryptamine (serotonin) receptor 3C, ionotropic
ENSG00000134917	ADAMTS8	ADAM metallopeptidase with thrombospondin type 1 motif, 8
ENSG00000127249	ATP13A4	ATPase type 13A4
ENSG00000198183	BPIFA1	BPI fold containing family A, member 1
ENSG00000105974	CAV1	Caveolin 1, caveolae protein, 22kDa
ENSG00000105971	CAV2	Caveolin 2
ENSG00000275385	CCL18	Chemokine (C-C motif) ligand 18 (pulmonary and activation-regulated)
ENSG00000169031	COL4A3	Collagen, type IV, alpha 3 (Goodpasture antigen)
ENSG00000172752	COL6A5	Collagen, type VI, alpha 5
ENSG00000206384	COL6A6	Collagen, type VI, alpha 6
ENSG00000197446	CYP2F1	Cytochrome P450, family 2, subfamily F, polypeptide 1
ENSG00000142973	CYP4B1	Cytochrome P450, family 4, subfamily B, polypeptide 1
ENSG00000186160	CYP4Z1	Cytochrome P450, family 4, subfamily Z, polypeptide 1
ENSG00000094963	FMO2	Flavin containing monooxygenase 2 (non-functional)
ENSG00000110195	FOLR1	Folate receptor 1 (adult)
ENSG00000069122	GPR116	G protein-coupled receptor 116
ENSG00000013588	GPRC5A	G protein-coupled receptor, class C, group 5, member A
ENSG00000266524	GDF10	Growth differentiation factor 10
ENSG00000278637	HIST1H4A	Histone cluster 1, H4a
ENSG00000105371	ICAM4	Intercellular adhesion molecule 4 (Landsteiner-Wiener blood group)
ENSG00000176842	IRX5	Iroquois homeobox 5
ENSG00000256980	KHDC1L	KH homology domain containing 1-like
ENSG00000125498	KIR2DL1	Killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 1
ENSG00000188883	KLRG2	Killer cell lectin-like receptor subfamily G, member 2
ENSG00000019169	MARCO	Macrophage receptor with collagenous structure
ENSG00000198502	HLA-DRB5	Major histocompatibility complex, class II, DR beta 5
ENSG00000204301	NOTCH4	Notch 4
ENSG00000170909	OSCAR	Osteoclast associated, immunoglobulin-like receptor
ENSG00000173391	OLR1	Oxidized low density lipoprotein (lectin-like) receptor 1
ENSG00000251664	PCDHA12	Protocadherin alpha 12
ENSG00000188015	S100A3	S100 calcium binding protein A3
ENSG00000100012	SEC14L3	SEC14-like 3 (S. cerevisiae)
ENSG00000214491	SEC14L6	SEC14-like 6 (S. cerevisiae)
ENSG00000121068	TBX2	T-box 2
ENSG00000124731	TREM1	Triggering receptor expressed on myeloid cells 1
ENSG00000095917	TPSD1	Tryptase delta 1
ENSG00000111981	ULBP1	UL16 binding protein 1
ENSG00000267261	CTD-2132N18.3	Uncharacterized protein (ECO:0000313 Ensembl:ENSP00000468275)

Table B.30: Clustered genes with enhanced expression in lymph node

Ensembl	Gene	Gene description
ENSG00000163568	AIM2	Absent in melanoma 2
ENSG00000156127	BATF	Basic leucine zipper transcription factor, ATF-like
ENSG00000069493	CLEC2D	C-type lectin domain family 2, member D
ENSG00000198178	CLEC4C	C-type lectin domain family 4, member C
ENSG00000197992	CLEC9A	C-type lectin domain family 9, member A
ENSG00000184293	CLECL1	C-type lectin-like 1
ENSG00000012124	CD22	CD22 molecule
ENSG00000178562	CD28	CD28 molecule
ENSG00000167286	CD3D	CD3d molecule, delta (CD3-TCR complex)
ENSG00000198851	CD3E	CD3e molecule, epsilon (CD3-TCR complex)
ENSG00000160654	CD3G	CD3g molecule, gamma (CD3-TCR complex)
ENSG00000117091	CD48	CD48 molecule
ENSG00000110448	CD5	CD5 molecule
ENSG00000013725	CD6	CD6 molecule
ENSG00000110848	CD69	CD69 molecule
ENSG00000066294	CD84	CD84 molecule
ENSG00000173578	XCR1	Chemokine (C motif) receptor 1
ENSG00000102970	CCL17	Chemokine (C-C motif) ligand 17
ENSG00000137077	CCL21	Chemokine (C-C motif) ligand 21
ENSG00000112486	CCR6	Chemokine (C-C motif) receptor 6
ENSG00000163599	CTLA4	Cytotoxic T-lymphocyte-associated protein 4
ENSG00000189057	FAM111B	Family with sequence similarity 111, member B
ENSG00000162894	FAIM3	Fas apoptotic inhibitory molecule 3
ENSG00000163534	FCRL1	Fc receptor-like 1
ENSG00000143297	FCRL5	Fc receptor-like 5
ENSG00000162897	FCAMR	Fc receptor, IgA, IgM, high affinity
ENSG00000122025	FLT3	Fms-related tyrosine kinase 3
ENSG00000159618	GPR114	G protein-coupled receptor 114
ENSG00000147138	GPR174	G protein-coupled receptor 174
ENSG00000125245	GPR18	G protein-coupled receptor 18
ENSG00000120436	GPR31	G protein-coupled receptor 31
ENSG00000178732	GP5	Glycoprotein V (platelet)
ENSG00000145649	GZMA	Granzyme A (granzyme 1, cytotoxic T-lymphocyte-associated serine esterase 3)
ENSG00000113088	GZMK	Granzyme K (granzyme 3; tryptase II)
ENSG00000154016	GRAP	GRB2-related adaptor protein
ENSG00000189152	GRAPL	GRB2-related adaptor protein-like
ENSG00000154451	GBP5	Guanylate binding protein 5
ENSG00000197153	HIST1H3J	Histone cluster 1, H3j
ENSG00000140511	HAPLN3	Hyaluronan and proteoglycan link protein 3
ENSG00000131203	IDO1	Indoleamine 2,3-dioxygenase 1
ENSG00000163600	ICOS	Inducible T-cell co-stimulator
ENSG00000172183	ISG20	Interferon stimulated exonuclease gene 20kDa
ENSG00000103522	IL21R	Interleukin 21 receptor
ENSG00000172578	KLHL6	Kelch-like family member 6
ENSG00000139187	KLRG1	Killer cell lectin-like receptor subfamily G, member 1
ENSG00000213809	KLRK1	Killer cell lectin-like receptor subfamily K, member 1
ENSG00000239961	LILRA4	Leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 4
ENSG00000186818	LILRB4	Leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 4
ENSG00000167618	LAIR2	Leukocyte-associated immunoglobulin-like receptor 2
ENSG00000122224	LY9	Lymphocyte antigen 9

Table B.30: Clustered genes with enhanced expression in lymph node

Ensembl	Gene	Gene description
ENSG00000226979	LTA	Lymphotoxin alpha
ENSG00000237541	HLA-DQA2	Major histocompatibility complex, class II, DQ alpha 2
ENSG00000179344	HLA-DQB1	Major histocompatibility complex, class II, DQ beta 1
ENSG00000167984	NLRC3	NLR family, CARD domain containing 3
ENSG00000117215	PLA2G2D	Phospholipase A2, group IID
ENSG00000173200	PARP15	Poly (ADP-ribose) polymerase family, member 15
ENSG00000078589	P2RY10	Purinergic receptor P2Y, G-protein coupled, 10
ENSG00000163564	PYHIN1	Pyrin and HIN domain family, member 1
ENSG00000127074	RGS13	Regulator of G-protein signaling 13
ENSG00000188404	SELL	Selectin L
ENSG00000180096	SEPT1	Septin 1
ENSG00000170054	SERPINA9	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 9
ENSG00000089012	SIRPG	Signal-regulatory protein gamma
ENSG00000117090	SLAMF1	Signaling lymphocytic activation molecule family member 1
ENSG00000162739	SLAMF6	SLAM family member 6
ENSG00000269404	SPIB	Spi-B transcription factor (Spi-1/PU.1 related)
ENSG00000145850	TIMD4	T-cell immunoglobulin and mucin domain containing 4
ENSG00000100721	TCL1A	T-cell leukemia/lymphoma 1A
ENSG00000213231	TCL1B	T-cell leukemia/lymphoma 1B
ENSG00000112195	TREML2	Triggering receptor expressed on myeloid cells-like 2
ENSG00000074966	TXK	TXK tyrosine kinase
ENSG00000181896	ZNF101	Zinc finger protein 101

Table B.31: Clustered genes with enhanced expression in ovary

Ensembl	Gene	Gene description
ENSG00000154736	ADAMTS5	ADAM metallopeptidase with thrombospondin type 1 motif, 5
ENSG00000167311	ART5	ADP-ribosyltransferase 5
ENSG00000135409	AMHR2	Anti-Mullerian hormone receptor, type II
ENSG00000154263	ABCA10	ATP-binding cassette, sub-family A (ABC1), member 10
ENSG00000154258	ABCA9	ATP-binding cassette, sub-family A (ABC1), member 9
ENSG00000006606	CCL26	Chemokine (C-C motif) ligand 26
ENSG00000169515	CCDC8	Coiled-coil domain containing 8
ENSG00000269502	DMRTC1	DMRT-like family C1
ENSG00000134201	GSTM5	Glutathione S-transferase mu 5
ENSG00000197757	HOXC6	Homeobox C6
ENSG00000103742	IGDCC4	Immunoglobulin superfamily, DCC subclass, member 4
ENSG00000123999	INHA	Inhibin, alpha
ENSG00000115461	IGFBP5	Insulin-like growth factor binding protein 5
ENSG00000163884	KLF15	Kruppel-like factor 15
ENSG00000147145	LPAR4	Lysophosphatidic acid receptor 4
ENSG00000205978	NYNRIN	NYN domain and retroviral integrase containing
ENSG00000214456	PLIN5	Perilipin 5
ENSG00000136546	SCN7A	Sodium channel, voltage-gated, type VII, alpha subunit
ENSG00000126549	STATH	Statherin
ENSG00000092445	TYRO3	TYRO3 protein tyrosine kinase
ENSG00000172000	ZNF556	Zinc finger protein 556
ENSG00000197360	ZNF98	Zinc finger protein 98

Table B.32: Clustered genes with enhanced expression in pancreas

Ensembl	Gene	Gene description
ENSG00000197360	ZNF98	Zinc finger protein 98
ENSG00000161270	NPHS1	Nephrosis 1, congenital, Finnish type (nephrin)
ENSG00000126266	FFAR1	Free fatty acid receptor 1
ENSG00000126549	STATH	Statherin
ENSG00000269533	AC003002.6	Uncharacterized protein (ECO:0000313 Ensembl:ENSP00000472277)
ENSG00000104826	LHB	Luteinizing hormone beta polypeptide

Table B.33: Clustered genes with enhanced expression in placenta

Ensembl	Gene	Gene description
ENSG00000139567	ACVRL1	Activin A receptor type II-like 1
ENSG00000138316	ADAMTS14	ADAM metallopeptidase with thrombospondin type 1 motif, 14
ENSG00000166106	ADAMTS15	ADAM metallopeptidase with thrombospondin type 1 motif, 15
ENSG00000163286	ALPPL2	Alkaline phosphatase, placental-like 2
ENSG00000109321	AREG	Amphiregulin
ENSG00000150048	CLEC1A	C-type lectin domain family 1, member A
ENSG00000126890	CTAG2	Cancer/testis antigen 2
ENSG00000196337	CGB7	Chorionic gonadotropin, beta polypeptide 7
ENSG00000042304	C2orf83	Chromosome 2 open reading frame 83
ENSG00000184697	CLDN6	Claudin 6
ENSG00000187498	COL4A1	Collagen, type IV, alpha 1
ENSG00000134871	COL4A2	Collagen, type IV, alpha 2
ENSG00000080573	COL5A3	Collagen, type V, alpha 3
ENSG00000198223	CSF2RA	Colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage)
ENSG00000100122	CRYBB1	Crystallin, beta B1
ENSG00000152207	CYSLTR2	Cysteinyl leukotriene receptor 2
ENSG00000064195	DLX3	Distal-less homeobox 3
ENSG00000108813	DLX4	Distal-less homeobox 4
ENSG00000197594	ENPP1	Ectonucleotide pyrophosphatase/phosphodiesterase 1
ENSG00000121743	GJA3	Gap junction protein, alpha 3, 46kDa
ENSG00000187513	GJA4	Gap junction protein, alpha 4, 37kDa
ENSG00000265107	GJA5	Gap junction protein, alpha 5, 40kDa
ENSG00000069812	HES2	Hes family bHLH transcription factor 2
ENSG00000106302	HYAL4	Hyaluronoglucosaminidase 4
ENSG00000131203	IDO1	Indoleamine 2,3-dioxygenase 1
ENSG00000123243	ITIH5	Inter-alpha-trypsin inhibitor heavy chain family, member 5
ENSG00000081985	IL12RB2	Interleukin 12 receptor, beta 2
ENSG00000167618	LAIR2	Leukocyte-associated immunoglobulin-like receptor 2
ENSG00000187123	LYPD6	LY6/PLAUR domain containing 6
ENSG00000204632	HLA-G	Major histocompatibility complex, class I, G
ENSG00000173391	OLR1	Oxidized low density lipoprotein (lectin-like) receptor 1
ENSG00000178226	PRSS36	Protease, serine, 36
ENSG00000113555	PCDH12	Protocadherin 12
ENSG00000171631	P2RY6	Pyrimidinergic receptor P2Y, G-protein coupled, 6
ENSG00000135917	SLC19A3	Solute carrier family 19 (thiamine transporter), member 3
ENSG00000126549	STATH	Statherin
ENSG00000104953	TLE6	Transducin-like enhancer of split 6
ENSG00000159450	TCHH	Trichohyalin
ENSG00000254979	RP11-872D17.8	Uncharacterized protein (ECO:0000313 Ensembl:ENSP00000431536)
ENSG00000115596	WNT6	Wingless-type MMTV integration site family, member 6
ENSG00000180532	ZSCAN4	Zinc finger and SCAN domain containing 4

Table B.34: Clustered genes with enhanced expression in prostate

Ensembl	Gene	Gene description
ENSG00000166743	ACSM1	Acyl-CoA synthetase medium-chain family member 1
ENSG00000179593	ALOX15B	Arachidonate 15-lipoxygenase, type B
ENSG00000121270	ABCC11	ATP-binding cassette, sub-family C (CFTR/MRP), member 11
ENSG00000121005	CRISPLD1	Cysteine-rich secretory protein LCCL domain containing 1
ENSG00000186526	CYP4F8	Cytochrome P450, family 4, subfamily F, polypeptide 8
ENSG00000106038	EVX1	Even-skipped homeobox 1
ENSG00000166105	GLB1L3	Galactosidase, beta 1-like 3
ENSG00000169224	GCSAML	Germinal center-associated, signaling and motility-like
ENSG00000196866	HIST1H2AD	Histone cluster 1, H2ad
ENSG00000277075	HIST1H2AE	Histone cluster 1, H2ae
ENSG00000273802	HIST1H2BG	Histone cluster 1, H2bg
ENSG00000203814	HIST2H2BF	Histone cluster 2, H2bf
ENSG00000128713	HOXD11	Homeobox D11
ENSG00000244355	LY6G6D	Lymphocyte antigen 6 complex, locus G6D
ENSG00000111049	MYF5	Myogenic factor 5
ENSG00000109063	MYH3	Myosin, heavy chain 3, skeletal muscle, embryonic
ENSG00000185792	NLRP9	NLR family, pyrin domain containing 9
ENSG00000180785	OR51E1	Olfactory receptor, family 51, subfamily E, member 1
ENSG00000082556	OPRK1	Opioid receptor, kappa 1
ENSG00000184388	PABPC1L2B	Poly(A) binding protein, cytoplasmic 1-like 2B
ENSG00000100003	SEC14L2	SEC14-like 2 (S. cerevisiae)
ENSG00000007908	SELE	Selectin E
ENSG00000124233	SEMG1	Semenogelin I
ENSG00000164647	STEAP1	Six transmembrane epithelial antigen of the prostate 1
ENSG00000141469	SLC14A1	Solute carrier family 14 (urea transporter), member 1 (Kidd blood group)
ENSG00000125508	SRMS	Src-related kinase lacking C-terminal regulatory tyrosine and N-terminal myristylation sites
ENSG00000165125	TRPV6	Transient receptor potential cation channel, subfamily V, member 6
ENSG00000279219	AL008723.1	Transient receptor potential cation channel, subfamily V, member 6

Table B.35: Clustered genes with enhanced expression in rectum

Ensembl	Gene	Gene description
ENSG00000134028	ADAMDEC1	ADAM-like, decysin 1
ENSG00000106541	AGR2	Anterior gradient 2
ENSG00000104267	CA2	Carbonic anhydrase II
ENSG00000172828	CES3	Carboxylesterase 3
ENSG00000079385	CEACAM1	Carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein)
ENSG00000105388	CEACAM5	Carcinoembryonic antigen-related cell adhesion molecule 5
ENSG00000275718	CCL15	Chemokine (C-C motif) ligand 15
ENSG00000106178	CCL24	Chemokine (C-C motif) ligand 24
ENSG00000006606	CCL26	Chemokine (C-C motif) ligand 26
ENSG00000166920	C15orf48	Chromosome 15 open reading frame 48
ENSG00000156284	CLDN8	Claudin 8
ENSG00000186564	FOXD2	Forkhead box D2
ENSG00000089356	FXYD3	FXYD domain containing ion transport regulator 3
ENSG00000113249	HAVCR1	Hepatitis A virus cellular receptor 1
ENSG00000170689	HOXB9	Homeobox B9
ENSG00000179914	ITLN1	Intelectin 1 (galactofuranose binding)
ENSG00000171431	KRT20	Keratin 20
ENSG00000114248	LRRC31	Leucine rich repeat containing 31
ENSG00000204033	LRIT2	Leucine-rich repeat, immunoglobulin-like and transmembrane domains 2
ENSG00000153898	MCOLN2	Mucolipin 2
ENSG00000183034	OTOP2	Otopetrin 2
ENSG00000130035	GALNT8	Polypeptide N-acetylgalactosaminyltransferase 8
ENSG00000168229	PTGDR	Prostaglandin D2 receptor (DP)
ENSG00000146039	SLC17A4	Solute carrier family 17, member 4
ENSG00000115616	SLC9A2	Solute carrier family 9, subfamily A (NHE2, cation proton antiporter 2), member 2
ENSG00000160180	TFF3	Trefoil factor 3 (intestinal)
ENSG00000204616	TRIM31	Tripartite motif containing 31
ENSG00000267881	CEA	Uncharacterized protein (ECO:0000313 Ensembl:ENSP00000469926)
ENSG00000221874	ZNF321P	Zinc finger protein 321, pseudogene

Table B.36: Clustered genes with enhanced expression in salivary gland

Ensembl	Gene	Gene description
ENSG00000161798	AQP5	Aquaporin 5
ENSG00000187527	ATP13A5	ATPase type 13A5
ENSG00000198183	BPIFA1	BPI fold containing family A, member 1
ENSG00000125999	BPIFB1	BPI fold containing family B, member 1
ENSG00000142224	IL19	Interleukin 19
ENSG00000176842	IRX5	Iroquois homeobox 5
ENSG00000174562	KLK15	Kallikrein-related peptidase 15
ENSG00000214456	PLIN5	Perilipin 5
ENSG00000111344	RASAL1	RAS protein activator like 1 (GAP1 like)
ENSG00000165125	TRPV6	Transient receptor potential cation channel, subfamily V, member 6
ENSG00000186806	VSIG10L	V-set and immunoglobulin domain containing 10 like
ENSG00000101443	WFDC2	WAP four-disulfide core domain 2
ENSG00000267748	CTB-102L5.4	Predicted intracellular protein

Table B.37: Clustered genes with enhanced expression in skeletal muscle

Ensembl	Gene	Gene description
ENSG00000170667	RASA4B	RAS p21 protein activator 4B
ENSG00000006788	MYH13	Myosin, heavy chain 13, skeletal muscle
ENSG00000133020	MYH8	Myosin, heavy chain 8, skeletal muscle, perinatal
ENSG00000260458	KCNJ18	Potassium inwardly-rectifying channel, subfamily J, member 18
ENSG00000160678	S100A1	S100 calcium binding protein A1
ENSG00000255526	NEDD8-MDP1	NEDD8-MDP1 readthrough
ENSG00000111049	MYF5	Myogenic factor 5
ENSG00000170276	HSPB2	Heat shock 27kDa protein 2
ENSG00000180818	HOXC10	Homeobox C10
ENSG00000146054	TRIM7	Tripartite motif containing 7
ENSG00000180806	HOXC9	Homeobox C9
ENSG00000155367	PPM1J	Protein phosphatase, Mg2+/Mn2+ dependent, 1J
ENSG00000131152	RP11-178L8.4	Protein phosphatase, Mg2+/Mn2+ dependent, 1J
ENSG00000266953	RP11-618P17.4	Uncharacterized protein (ECO:0000313 Ensembl:ENSP00000475682)

Table B.38: Clustered genes with enhanced expression in skin

Ensembl	Gene	Gene description
ENSG00000147160	AWAT2	Acyl-CoA wax alcohol acyltransferase 2
ENSG00000132746	ALDH3B2	Aldehyde dehydrogenase 3 family, member B2
ENSG00000198483	ANKRD35	Ankyrin repeat domain 35
ENSG00000187527	ATP13A5	ATPase type 13A5
ENSG00000256660	CLEC12B	C-type lectin domain family 12, member B
ENSG00000205863	C1QTNF9B	C1q and tumor necrosis factor related protein 9B
ENSG00000062038	CDH3	Cadherin 3, type 1, P-cadherin (placental)
ENSG00000255221	CARD17	Caspase recruitment domain family, member 17
ENSG00000158477	CD1A	CD1a molecule
ENSG00000137975	CLCA2	Chloride channel accessory 2
ENSG00000196337	CGB7	Chorionic gonadotropin, beta polypeptide 7
ENSG00000163347	CLDN1	Claudin 1
ENSG00000172752	COL6A5	Collagen, type VI, alpha 5
ENSG00000163354	DCST2	DC-STAMP domain containing 2
ENSG00000176797	DEFB103A	Defensin, beta 103A
ENSG00000177243	DEFB103B	Defensin, beta 103B
ENSG00000064195	DLX3	Distal-less homeobox 3
ENSG00000137857	DUOX1	Dual oxidase 1
ENSG00000140254	DUOXA1	Dual oxidase maturation factor 1
ENSG00000143590	EFNA3	Ephrin-A3
ENSG00000261150	EPPK1	Epiplakin 1
ENSG00000161243	FBXO27	F-box protein 27
ENSG00000164393	GPR111	G protein-coupled receptor 111
ENSG00000153294	GPR115	G protein-coupled receptor 115
ENSG00000138271	GPR87	G protein-coupled receptor 87
ENSG00000188910	GJB3	Gap junction protein, beta 3, 31kDa
ENSG00000069812	HES2	Hes family bHLH transcription factor 2
ENSG00000123388	HOXC11	Homeobox C11
ENSG00000204866	IGFL2	IGF-like family member 2
ENSG00000115598	IL1RL2	Interleukin 1 receptor-like 2
ENSG00000016402	IL20RA	Interleukin 20 receptor, alpha
ENSG00000176842	IRX5	Iroquois homeobox 5
ENSG00000168970	JMJD7-PLA2G4B	JMJD7-PLA2G4B readthrough
ENSG00000167757	KLK11	Kallikrein-related peptidase 11
ENSG00000167767	KRT80	Keratin 80
ENSG00000171403	KRT9	Keratin 9
ENSG00000204033	LRIT2	Leucine-rich repeat, immunoglobulin-like and transmembrane domains 2
ENSG00000213903	LTB4R	Leukotriene B4 receptor
ENSG00000213906	LTB4R2	Leukotriene B4 receptor 2
ENSG00000232629	HLA-DQB2	Major histocompatibility complex, class II, DQ beta 2
ENSG00000183695	MRGPRX2	MAS-related GPR, member X2
ENSG00000137675	MMP27	Matrix metallopeptidase 27
ENSG00000102891	MT4	Metallothionein 4
ENSG00000105088	OLFM2	Olfactomedin 2
ENSG00000163218	PGLYRP4	Peptidoglycan recognition protein 4
ENSG00000143217	PVRL4	Poliovirus receptor-related 4
ENSG00000260458	KCNJ18	Potassium inwardly-rectifying channel, subfamily J, member 18
ENSG00000151006	PRSS53	Protease, serine, 53
ENSG00000101213	PTK6	Protein tyrosine kinase 6
ENSG00000164520	RAET1E	Retinoic acid early transcript 1E
ENSG00000139988	RDH12	Retinol dehydrogenase 12 (all-trans/9-cis/11-cis)

Table B.38: Clustered genes with enhanced expression in skin

Ensembl	Gene	Gene description
ENSG00000165799	RNASE7	Ribonuclease, RNase A family, 7
ENSG00000204618	RNF39	Ring finger protein 39
ENSG00000197632	SERPINB2	Serpin peptidase inhibitor, clade B (ovalbumin), member 2
ENSG00000132164	SLC6A11	Solute carrier family 6 (neurotransmitter transporter), member 11
ENSG00000217236	SP9	Sp9 transcription factor
ENSG00000173898	SPTBN2	Spectrin, beta, non-erythrocytic 2
ENSG00000104055	TGM5	Transglutaminase 5
ENSG00000167723	TRPV3	Transient receptor potential cation channel, subfamily V, member 3
ENSG00000137747	TMPRSS13	Transmembrane protease, serine 13
ENSG00000168671	UGT3A2	UDP glycosyltransferase 3 family, polypeptide A2
ENSG00000135925	WNT10A	Wingless-type MMTV integration site family, member 10A

Table B.39: Clustered genes with enhanced expression in small intestine

Ensembl	Gene	Gene description
ENSG00000186038	HTR3E	5-hydroxytryptamine (serotonin) receptor 3E, ionotropic
ENSG00000164270	HTR4	5-hydroxytryptamine (serotonin) receptor 4, G protein-coupled
ENSG00000134028	ADAMDEC1	ADAM-like, decysin 1
ENSG00000198074	AKR1B10	Aldo-keto reductase family 1, member B10 (aldose reductase)
ENSG00000162482	AKR7A3	Aldo-keto reductase family 7, member A3 (aflatoxin aldehyde reductase)
ENSG00000173467	AGR3	Anterior gradient 3
ENSG00000118137	APOA1	Apolipoprotein A-I
ENSG00000114771	AADAC	Arylacetamide deacetylase
ENSG00000085563	ABCB1	ATP-binding cassette, sub-family B (MDR/TAP), member 1
ENSG00000152672	CLEC4F	C-type lectin domain family 4, member F
ENSG00000185015	CA13	Carbonic anhydrase XIII
ENSG00000275718	CCL15	Chemokine (C-C motif) ligand 15
ENSG00000106178	CCL24	Chemokine (C-C motif) ligand 24
ENSG00000165215	CLDN3	Claudin 3
ENSG00000258941	RP11-407N17.3	CTAGE family member 5 ECO:0000313—Ensembl:ENSP00000452252
ENSG00000108242	CYP2C18	Cytochrome P450, family 2, subfamily C, polypeptide 18
ENSG00000106258	CYP3A5	Cytochrome P450, family 3, subfamily A, polypeptide 5
ENSG00000186204	CYP4F12	Cytochrome P450, family 4, subfamily F, polypeptide 12
ENSG00000197635	DPP4	Dipeptidyl-peptidase 4
ENSG00000188833	ENTPD8	Ectonucleoside triphosphate diphosphohydrolase 8
ENSG00000154269	ENPP3	Ectonucleotide pyrophosphatase/phosphodiesterase 3
ENSG00000073605	GSDMB	Gasdermin B
ENSG00000044012	GUCA2B	Guanylate cyclase activator 2B (uroguanylin)
ENSG00000156510	HKDC1	Hexokinase domain containing 1
ENSG00000257743	RP11-1220K2.2	Homo sapiens maltase-glucoamylase (alpha-glucosidase) (LOC93432), mRNA.
ENSG00000133328	HRASLS2	HRAS-like suppressor 2
ENSG00000187242	KRT12	Keratin 12
ENSG00000171431	KRT20	Keratin 20
ENSG00000170421	KRT8	Keratin 8
ENSG00000177984	LCN15	Lipocalin 15
ENSG00000150526	MIA2	Melanoma inhibitory activity 2
ENSG00000169894	MUC3A	Mucin 3A, cell surface associated
ENSG00000188257	PLA2G2A	Phospholipase A2, group IIA (platelets, synovial fluid)
ENSG00000165828	PRAP1	Proline-rich acidic protein 1
ENSG00000172023	REG1B	Regenerating islet-derived 1 beta
ENSG00000146039	SLC17A4	Solute carrier family 17, member 4
ENSG00000142583	SLC2A5	Solute carrier family 2 (facilitated glucose/fructose transporter), member 5
ENSG00000197165	SULT1A2	Sulfotransferase family, cytosolic, 1A, phenol-preferring, member 2
ENSG00000196228	SULT1C3	Sulfotransferase family, cytosolic, 1C, member 3
ENSG00000167723	TRPV3	Transient receptor potential cation channel, subfamily V, member 3
ENSG00000204610	TRIM15	Tripartite motif containing 15
ENSG00000204614	TRIM40	Tripartite motif containing 40
ENSG00000135220	UGT2A3	UDP glucuronosyltransferase 2 family, polypeptide A3

Table B.40: Clustered genes with enhanced expression in smooth muscle

Ensembl	Gene	Gene description
ENSG00000128710	HOXD10	Homeobox D10
ENSG00000143768	LEFTY2	Left-right determination factor 2
ENSG00000253293	HOXA10	Homeobox A10
ENSG00000138615	CILP	Cartilage intermediate layer protein, nucleotide pyrophosphohydrolase
ENSG00000188828	GLRA4	Glycine receptor, alpha 4
ENSG00000137558	PI15	Peptidase inhibitor 15
ENSG00000162891	IL20	Interleukin 20
ENSG00000145936	KCNMB1	Potassium large conductance calcium-activated channel, subfamily M, beta member 1
ENSG00000106819	ASPN	Asporin
ENSG00000135914	HTR2B	5-hydroxytryptamine (serotonin) receptor 2B, G protein-coupled
ENSG00000037965	HOXC8	Homeobox C8
ENSG00000168542	COL3A1	Collagen, type III, alpha 1

Table B.41: Clustered genes with enhanced expression in spleen

Ensembl	Gene	Gene description
ENSG00000166523	CLEC4E	C-type lectin domain family 4, member E
ENSG00000175544	CABP4	Calcium binding protein 4
ENSG00000177675	CD163L1	CD163 molecule-like 1
ENSG00000186074	CD300LF	CD300 molecule-like family member f
ENSG00000276409	CCL14	Chemokine (C-C motif) ligand 14
ENSG00000163739	CXCL1	Chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)
ENSG00000124875	CXCL6	Chemokine (C-X-C motif) ligand 6
ENSG00000163464	CXCR1	Chemokine (C-X-C motif) receptor 1
ENSG00000203710	CR1	Complement component (3b/4b) receptor 1 (Knops blood group)
ENSG00000134830	C5AR2	Complement component 5a receptor 2
ENSG00000070886	EPHA8	EPH receptor A8
ENSG00000188820	FAM26F	Family with sequence similarity 26, member F
ENSG00000162894	FAIM3	Fas apoptotic inhibitory molecule 3
ENSG00000162747	FCGR3B	Fc fragment of IgG, low affinity IIIb, receptor (CD16b)
ENSG00000163534	FCRL1	Fc receptor-like 1
ENSG00000122025	FLT3	Fms-related tyrosine kinase 3
ENSG00000154016	GRAP	GRB2-related adaptor protein
ENSG00000189152	GRAPL	GRB2-related adaptor protein-like
ENSG00000213203	GIMAP1	GTPase, IMAP family member 1
ENSG00000114378	HYAL1	Hyaluronoglucosaminidase 1
ENSG00000255398	HCAR3	Hydroxycarboxylic acid receptor 3
ENSG00000140678	ITGAX	Integrin, alpha X (complement component 3 receptor 4 subunit)
ENSG00000167633	KIR3DL1	Killer cell immunoglobulin-like receptor, three domains, long cytoplasmic tail, 1
ENSG00000240403	KIR3DL2	Killer cell immunoglobulin-like receptor, three domains, long cytoplasmic tail, 2
ENSG00000125498	KIR2DL1	Killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 1
ENSG00000243772	KIR2DL3	Killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 3
ENSG00000189013	KIR2DL4	Killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 4
ENSG00000205809	KLRC2	Killer cell lectin-like receptor subfamily C, member 2
ENSG00000205810	KLRC3	Killer cell lectin-like receptor subfamily C, member 3
ENSG00000183542	KLRC4	Killer cell lectin-like receptor subfamily C, member 4
ENSG00000134539	KLRD1	Killer cell lectin-like receptor subfamily D, member 1
ENSG00000213809	KLRK1	Killer cell lectin-like receptor subfamily K, member 1
ENSG00000104974	LILRA1	Leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 1
ENSG00000239998	LILRA2	Leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 2
ENSG00000239961	LILRA4	Leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 4
ENSG00000104972	LILRB1	Leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 1
ENSG00000105609	LILRB5	Leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 5
ENSG00000103313	MEFV	Mediterranean fever
ENSG00000166927	MS4A7	Membrane-spanning 4-domains, subfamily A, member 7
ENSG00000189430	NCR1	Natural cytotoxicity triggering receptor 1

Table B.41: Clustered genes with enhanced expression in spleen

Ensembl	Gene	Gene description
ENSG00000164128	NPY1R	Neuropeptide Y receptor Y1
ENSG00000164129	NPY5R	Neuropeptide Y receptor Y5
ENSG00000180644	PRF1	Perforin 1 (pore forming protein)
ENSG00000109272	PF4V1	Platelet factor 4 variant 1
ENSG00000173200	PARP15	Poly (ADP-ribose) polymerase family, member 15
ENSG00000181631	P2RY13	Purinergic receptor P2Y, G-protein coupled, 13
ENSG00000171631	P2RY6	Pyrimidinergic receptor P2Y, G-protein coupled, 6
ENSG00000150681	RGS18	Regulator of G-protein signaling 18
ENSG00000154133	ROBO4	Roundabout, axon guidance receptor, homolog 4 (Drosophila)
ENSG00000180096	SEPT1	Septin 1
ENSG00000254521	SIGLEC12	Sialic acid binding Ig-like lectin 12 (gene/pseudogene)
ENSG00000254415	SIGLEC14	Sialic acid binding Ig-like lectin 14
ENSG00000268500	SIGLEC5	Sialic acid binding Ig-like lectin 5
ENSG00000168995	SIGLEC7	Sialic acid binding Ig-like lectin 7
ENSG00000168356	SCN11A	Sodium channel, voltage-gated, type XI, alpha subunit
ENSG00000159450	TCHH	Trichohyalin
ENSG00000249240	AC069368.3	Uncharacterized protein (ECO:0000313 Ensembl:ENSP00000397942
ENSG00000255641	NKG2-E	Uncharacterized protein (ECO:0000313 Ensembl:ENSP00000437563

Table B.42: Clustered genes with enhanced expression in stomach

Ensembl	Gene	Gene description
ENSG00000104267	CA2	Carbonic anhydrase II
ENSG00000106541	AGR2	Anterior gradient 2
ENSG00000197353	LYPD2	LY6/PLAUR domain containing 2
ENSG00000146755	TRIM50	Tripartite motif containing 50
ENSG00000111780	AL021546.6	Glutamyl-tRNA(Gln) amidotransferase subunit C, mitochondrial (ECO:0000313 Ensembl:ENSP00000229384)
ENSG00000019102	VSIG2	V-set and immunoglobulin domain containing 2
ENSG00000125850	OVOL2	Ovo-like zinc finger 2
ENSG00000198203	SULT1C2	Sulfotransferase family, cytosolic, 1C, member 2
ENSG00000198074	AKR1B10	Aldo-keto reductase family 1, member B10 (aldose reductase)
ENSG00000115616	SLC9A2	Solute carrier family 9, subfamily A (NHE2, cation proton antiporter 2), member 2
ENSG00000125508	SRMS	Src-related kinase lacking C-terminal regulatory tyrosine and N-terminal myristylation sites
ENSG00000125820	NKX2-2	NK2 homeobox 2
ENSG00000125999	BPIFB1	BPI fold containing family B, member 1
ENSG00000167600	CYP2S1	Cytochrome P450, family 2, subfamily S, polypeptide 1
ENSG00000227471	AKR1B15	Aldo-keto reductase family 1, member B15
ENSG00000108602	ALDH3A1	Aldehyde dehydrogenase 3 family, member A1
ENSG00000259120	SMIM6	Small integral membrane protein 6
ENSG00000155428	TRIM74	Tripartite motif containing 74
ENSG00000042980	ADAM28	ADAM metallopeptidase domain 28

Table B.43: Clustered genes with enhanced expression in testis

Ensembl	Gene	Gene description
ENSG00000167311	ART5	ADP-ribosyltransferase 5
ENSG00000258529	ALG9	ALG9, alpha-1,2-mannosyltransferase
ENSG00000102048	ASB9	Ankyrin repeat and SOCS box containing 9
ENSG00000161798	AQP5	Aquaporin 5
ENSG00000121270	ABCC11	ATP-binding cassette, sub-family C (CFTR/MRP), member 11
ENSG00000140798	ABCC12	ATP-binding cassette, sub-family C (CFTR/MRP), member 12
ENSG00000256660	CLEC12B	C-type lectin domain family 12, member B
ENSG00000008300	CELSR3	Cadherin, EGF LAG seven-pass G-type receptor 3
ENSG00000126890	CTAG2	Cancer/testis antigen 2
ENSG00000198930	CSAG1	Chondrosarcoma associated gene 1
ENSG00000173894	CBX2	Chromobox homolog 2
ENSG00000000460	C1orf112	Chromosome 1 open reading frame 112
ENSG00000184697	CLDN6	Claudin 6
ENSG00000160318	CLDND2	Claudin domain containing 2
ENSG00000269897	COMMD3-BMI1	COMMD3-BMI1 readthrough
ENSG00000271079	CTAGE15	CTAGE family, member 15
ENSG00000271321	CTAGE6	CTAGE family, member 6
ENSG00000236761	CTAGE9	CTAGE family, member 9
ENSG00000197446	CYP2F1	Cytochrome P450, family 2, subfamily F, polypeptide 1
ENSG00000163354	DCST2	DC-STAMP domain containing 2
ENSG00000198889	DCAF12L1	DDB1 and CUL4 associated factor 12-like 1
ENSG00000198354	DCAF12L2	DDB1 and CUL4 associated factor 12-like 2
ENSG00000064218	DMRT3	Doublesex and mab-3 related transcription factor 3
ENSG00000205085	FAM71F2	Family with sequence similarity 71, member F2
ENSG00000146276	GABRR1	Gamma-aminobutyric acid (GABA) A receptor, rho 1
ENSG00000088053	GP6	Glycoprotein VI (platelet)
ENSG00000278662	GOLGA6L10	Golgin A6 family-like 10
ENSG00000261794	GOLGA8H	Golgin A8 family, member H
ENSG00000179938	GOLGA8J	Golgin A8 family, member J
ENSG00000249931	GOLGA8K	Golgin A8 family, member K
ENSG00000188626	GOLGA8M	Golgin A8 family, member M
ENSG00000268738	HSFX2	Heat shock transcription factor family, X linked 2
ENSG00000138646	HERC5	HECT and RLD domain containing E3 ubiquitin protein ligase 5
ENSG00000130656	HBZ	Hemoglobin, zeta
ENSG00000196787	HIST1H2AG	Histone cluster 1, H2ag
ENSG00000275221	HIST1H2AK	Histone cluster 1, H2ak
ENSG00000276410	HIST1H2BB	Histone cluster 1, H2bb
ENSG00000124635	HIST1H2BJ	Histone cluster 1, H2bj
ENSG00000106302	HYAL4	Hyaluronoglucosaminidase 4
ENSG00000204866	IGFL2	IGF-like family member 2
ENSG00000167178	ISLR2	Immunoglobulin superfamily containing leucine-rich repeat 2
ENSG00000123999	INHA	Inhibin, alpha
ENSG00000027644	INSRR	Insulin receptor-related receptor
ENSG00000164509	IL31RA	Interleukin 31 receptor A
ENSG00000167216	KATNAL2	Katanin p60 subunit A-like 2
ENSG00000178502	KLHL11	Kelch-like family member 11
ENSG00000170486	KRT72	Keratin 72
ENSG00000256980	KHDC1L	KH homology domain containing 1-like
ENSG00000176681	LRRC37A	Leucine rich repeat containing 37A
ENSG00000160886	LY6K	Lymphocyte antigen 6 complex, locus K

Table B.43: Clustered genes with enhanced expression in testis

Ensembl	Gene	Gene description
ENSG00000185674	LYG2	Lysozyme G-like 2
ENSG00000144893	MED12L	Mediator complex subunit 12-like
ENSG00000108852	MPP2	Membrane protein, palmitoylated 2 (MAGUK p55 subfamily member 2)
ENSG00000255526	NEDD8-MDP1	NEDD8-MDP1 readthrough
ENSG00000186086	NBPF6	Neuroblastoma breakpoint family, member 6
ENSG00000136327	NKX2-8	NK2 homeobox 8
ENSG00000179873	NLRP11	NLR family, pyrin domain containing 11
ENSG00000122824	NUDT10	Nudix (nucleoside diphosphate linked moiety X)-type motif 10
ENSG00000196368	NUDT11	Nudix (nucleoside diphosphate linked moiety X)-type motif 11
ENSG00000204657	OR2H2	Olfactory receptor, family 2, subfamily H, member 2
ENSG00000159961	OR3A3	Olfactory receptor, family 3, subfamily A, member 3
ENSG00000127530	OR7C1	Olfactory receptor, family 7, subfamily C, member 1
ENSG00000188000	OR7D2	Olfactory receptor, family 7, subfamily D, member 2
ENSG00000112038	OPRM1	Opioid receptor, mu 1
ENSG00000188086	PRSS45	Protease, serine, 45
ENSG00000155367	PPM1J	Protein phosphatase, Mg2+/Mn2+ dependent, 1J
ENSG00000204963	PCDHA7	Protocadherin alpha 7
ENSG00000204961	PCDHA9	Protocadherin alpha 9
ENSG00000223638	RFPL4A	Ret finger protein-like 4A
ENSG00000161939	RNASEK-C17orf49	RNASEK-C17orf49 readthrough
ENSG00000214510	SPINK13	Serine peptidase inhibitor, Kazal type 13 (putative)
ENSG00000188488	SERPINA5	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 5
ENSG00000138378	STAT4	Signal transducer and activator of transcription 4
ENSG00000142583	SLC2A5	Solute carrier family 2 (facilitated glucose/fructose transporter), member 5
ENSG00000181240	SLC25A41	Solute carrier family 25, member 41
ENSG00000170374	SP7	Sp7 transcription factor
ENSG00000173898	SPTBN2	Spectrin, beta, non-erythrocytic 2
ENSG00000205238	SPDYE2	Speedy/RINGO cell cycle regulator family member E2
ENSG00000173678	SPDYE2B	Speedy/RINGO cell cycle regulator family member E2B
ENSG00000260097	SPDYE6	Speedy/RINGO cell cycle regulator family member E6
ENSG00000145850	TIMD4	T-cell immunoglobulin and mucin domain containing 4
ENSG00000213231	TCL1B	T-cell leukemia/lymphoma 1B
ENSG00000166984	TCP10L2	T-complex 10-like 2
ENSG00000215029	TCP11X2	T-complex 11 family, X-linked 2
ENSG00000274808	AC003976.2	TBC1 domain family member 3B
ENSG00000260287	TBC1D3G	TBC1 domain family, member 3G
ENSG00000273513	TBC1D3K	TBC1 domain family, member 3K
ENSG00000183632	TP53TG3	TP53 target 3
ENSG00000261509	TP53TG3B	TP53 target 3B
ENSG00000205457	TP53TG3C	TP53 target 3C
ENSG00000205456	TP53TG3D	TP53 target 3D
ENSG00000162931	TRIM17	Tripartite motif containing 17
ENSG00000146755	TRIM50	Tripartite motif containing 50
ENSG00000183439	TRIM61	Tripartite motif containing 61
ENSG00000170236	USP50	Ubiquitin specific peptidase 50
ENSG00000111981	ULBP1	UL16 binding protein 1
ENSG00000131019	ULBP3	UL16 binding protein 3

Table B.43: Clustered genes with enhanced expression in testis

Ensembl	Gene	Gene description
ENSG00000251569	RP11-724O16.1	Uncharacterized protein (ECO:0000313 Ensembl:ENSP00000424363)
ENSG00000182931	WFDC10B	WAP four-disulfide core domain 10B
ENSG00000152467	ZSCAN1	Zinc finger and SCAN domain containing 1
ENSG00000197279	ZNF165	Zinc finger protein 165
ENSG00000157429	ZNF19	Zinc finger protein 19
ENSG00000159915	ZNF233	Zinc finger protein 233
ENSG00000229676	ZNF492	Zinc finger protein 492
ENSG00000198633	ZNF534	Zinc finger protein 534
ENSG00000187187	ZNF546	Zinc finger protein 546
ENSG00000188868	ZNF563	Zinc finger protein 563
ENSG00000196109	ZNF676	Zinc finger protein 676
ENSG00000196946	ZNF705A	Zinc finger protein 705A
ENSG00000242852	ZNF709	Zinc finger protein 709
ENSG00000269067	ZNF728	Zinc finger protein 728
ENSG00000183850	ZNF730	Zinc finger protein 730
ENSG00000186777	ZNF732	Zinc finger protein 732
ENSG00000185869	ZNF829	Zinc finger protein 829
ENSG00000197360	ZNF98	Zinc finger protein 98
ENSG00000203995	ZYG11A	Zyg-11 family member A, cell cycle regulator

Table B.44: Clustered genes with enhanced expression in thyroid gland

Ensembl	Gene	Gene description
ENSG00000127249	ATP13A4	ATPase type 13A4
ENSG00000116985	BMP8B	Bone morphogenetic protein 8b
ENSG00000064218	DMRT3	Doublesex and mab-3 related transcription factor 3
ENSG00000137857	DUOX1	Dual oxidase 1
ENSG00000140254	DUOXA1	Dual oxidase maturation factor 1
ENSG00000112796	ENPP5	Ectonucleotide pyrophosphatase/phosphodiesterase 5 (putative)
ENSG00000179938	GOLGA8J	Golgin A8 family, member J
ENSG00000188626	GOLGA8M	Golgin A8 family, member M
ENSG00000206127	GOLGA8O	Golgin A8 family, member O
ENSG00000186399	GOLGA8R	Golgin A8 family, member R
ENSG00000074706	IPCEF1	Interaction protein for cytohesin exchange factors 1
ENSG00000135480	KRT7	Keratin 7
ENSG00000188883	KLRG2	Killer cell lectin-like receptor subfamily G, member 2
ENSG00000184925	LCN12	Lipocalin 12
ENSG00000152939	MARVELD2	MARVEL domain containing 2
ENSG00000125144	MT1G	Metallothionein 1G
ENSG00000197822	OCLN	Occludin
ENSG00000221970	OR2A1	Olfactory receptor, family 2, subfamily A, member 1
ENSG00000212807	OR2A42	Olfactory receptor, family 2, subfamily A, member 42
ENSG00000177212	OR2T33	Olfactory receptor, family 2, subfamily T, member 33
ENSG00000177462	OR2T8	Olfactory receptor, family 2, subfamily T, member 8
ENSG00000238243	OR2W3	Olfactory receptor, family 2, subfamily W, member 3
ENSG00000153246	PLA2R1	Phospholipase A2 receptor 1, 180kDa
ENSG00000172780	RAB43	RAB43, member RAS oncogene family
ENSG00000196420	S100A5	S100 calcium binding protein A5
ENSG00000142319	SLC6A3	Solute carrier family 6 (neurotransmitter transporter), member 3
ENSG00000104953	TLE6	Transducin-like enhancer of split 6
ENSG00000160180	TFF3	Trefoil factor 3 (intestinal)
ENSG00000256229	ZNF486	Zinc finger protein 486
ENSG00000160352	ZNF714	Zinc finger protein 714
ENSG00000203995	ZYG11A	Zyg-11 family member A, cell cycle regulator

Table B.45: Clustered genes with enhanced expression in tonsil

Ensembl	Gene	Gene description
ENSG00000163568	AIM2	Absent in melanoma 2
ENSG00000124557	BTN1A1	Butyrophilin, subfamily 1, member A1
ENSG00000069493	CLEC2D	C-type lectin domain family 2, member D
ENSG00000198178	CLEC4C	C-type lectin domain family 4, member C
ENSG00000184293	CLECL1	C-type lectin-like 1
ENSG00000158481	CD1C	CD1c molecule
ENSG00000012124	CD22	CD22 molecule
ENSG00000178562	CD28	CD28 molecule
ENSG00000167286	CD3D	CD3d molecule, delta (CD3-TCR complex)
ENSG00000198851	CD3E	CD3e molecule, epsilon (CD3-TCR complex)
ENSG00000160654	CD3G	CD3g molecule, gamma (CD3-TCR complex)
ENSG00000117091	CD48	CD48 molecule
ENSG00000013725	CD6	CD6 molecule
ENSG00000137077	CCL21	Chemokine (C-C motif) ligand 21
ENSG00000102962	CCL22	Chemokine (C-C motif) ligand 22
ENSG00000272980	CCR6	Chemokine (C-C motif) receptor 6
ENSG00000112486	CCR6	Chemokine (C-C motif) receptor 6
ENSG00000064886	CHI3L2	Chitinase 3-like 2
ENSG00000082293	COL19A1	Collagen, type XIX, alpha 1
ENSG00000163599	CTLA4	Cytotoxic T-lymphocyte-associated protein 4
ENSG00000176797	DEFB103A	Defensin, beta 103A
ENSG00000177243	DEFB103B	Defensin, beta 103B
ENSG00000189057	FAM111B	Family with sequence similarity 111, member B
ENSG00000162894	FAIM3	Fas apoptotic inhibitory molecule 3
ENSG00000163534	FCRL1	Fc receptor-like 1
ENSG00000143297	FCRL5	Fc receptor-like 5
ENSG00000162746	FCRLB	Fc receptor-like B
ENSG00000162897	FCAMR	Fc receptor, IgA, IgM, high affinity
ENSG00000159618	GPR114	G protein-coupled receptor 114
ENSG00000147138	GPR174	G protein-coupled receptor 174
ENSG00000125245	GPR18	G protein-coupled receptor 18
ENSG00000171657	GPR82	G protein-coupled receptor 82
ENSG00000154016	GRAP	GRB2-related adaptor protein
ENSG00000189152	GRAPL	GRB2-related adaptor protein-like
ENSG00000163600	ICOS	Inducible T-cell co-stimulator
ENSG00000136689	IL1RN	Interleukin 1 receptor antagonist
ENSG00000112116	IL17F	Interleukin 17F
ENSG00000103522	IL21R	Interleukin 21 receptor
ENSG00000164485	IL22RA2	Interleukin 22 receptor, alpha 2
ENSG00000129451	KLK10	Kallikrein-related peptidase 10
ENSG00000167755	KLK6	Kallikrein-related peptidase 6
ENSG00000172578	KLHL6	Kelch-like family member 6
ENSG00000167916	KRT24	Keratin 24
ENSG00000126337	KRT36	Keratin 36
ENSG00000170454	KRT75	Keratin 75
ENSG00000167767	KRT80	Keratin 80
ENSG00000239961	LILRA4	Leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 4
ENSG00000122224	LY9	Lymphocyte antigen 9
ENSG00000226979	LTA	Lymphotoxin alpha
ENSG00000237541	HLA-DQA2	Major histocompatibility complex, class II, DQ alpha 2
ENSG00000167984	NLRC3	NLR family, CARD domain containing 3
ENSG00000117215	PLA2G2D	Phospholipase A2, group IID

Table B.45: Clustered genes with enhanced expression in tonsil

Ensembl	Gene	Gene description
ENSG00000173200	PARP15	Poly (ADP-ribose) polymerase family, member 15
ENSG00000078589	P2RY10	Purinergic receptor P2Y, G-protein coupled, 10
ENSG00000127074	RGS13	Regulator of G-protein signaling 13
ENSG00000165799	RNASE7	Ribonuclease, RNase A family, 7
ENSG00000196754	S100A2	S100 calcium binding protein A2
ENSG00000188404	SELL	Selectin L
ENSG00000170054	SERPINA9	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 9
ENSG00000197632	SERPINB2	Serpin peptidase inhibitor, clade B (ovalbumin), member 2
ENSG00000089012	SIRPG	Signal-regulatory protein gamma
ENSG00000117090	SLAMF1	Signaling lymphocytic activation molecule family member 1
ENSG00000162739	SLAMF6	SLAM family member 6
ENSG00000269404	SPIB	Spi-B transcription factor (Spi-1/PU.1 related)
ENSG00000100721	TCL1A	T-cell leukemia/lymphoma 1A
ENSG00000213231	TCL1B	T-cell leukemia/lymphoma 1B

Table B.46: Clustered genes with enhanced expression in urinary bladder

Ensembl	Gene	Gene description
ENSG00000112116	IL17F	Interleukin 17F
ENSG00000204616	TRIM31	Tripartite motif containing 31
ENSG00000108405	P2RX1	Purinergic receptor P2X, ligand-gated ion channel, 1
ENSG00000100739	BDKRB1	Bradykinin receptor B1
ENSG00000258691	RP11-404P21.8	Uncharacterized protein ECO:0000313 Ensembl:ENSP00000450984
ENSG00000128422	KRT17	Keratin 17
ENSG00000163735	CXCL5	Chemokine (C-X-C motif) ligand 5
ENSG00000120068	HOXB8	Homeobox B8
ENSG00000184995	IFNE	Interferon, epsilon
ENSG00000163739	CXCL1	Chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)
ENSG00000242366	UGT1A1	UDP glucuronosyltransferase 1 family, polypeptide A1
ENSG00000198183	BPIFA1	BPI fold containing family A, member 1
ENSG00000124875	CXCL6	Chemokine (C-X-C motif) ligand 6
ENSG00000262406	MMP12	Matrix metallopeptidase 12 (macrophage elastase)
ENSG00000137648	TMPRSS4	Transmembrane protease, serine 4
ENSG00000007908	SELE	Selectin E
ENSG00000141469	SLC14A1	Solute carrier family 14 (urea transporter), member 1 (Kidd blood group)
ENSG00000186723	OR10H1	Olfactory receptor, family 10, subfamily H, member 1
ENSG00000133048	CHI3L1	Chitinase 3-like 1 (cartilage glycoprotein-39)
ENSG00000105991	HOXA1	Homeobox A1
ENSG00000179934	CCR8	Chemokine (C-C motif) receptor 8
ENSG00000135480	KRT7	Keratin 7
ENSG00000167165	UGT1A6	UDP glucuronosyltransferase 1 family, polypeptide A6

Table B.47: Genome assemblies used in the study

Scientific name	Assembly ID	Chromosome	chromosome accession
Bos taurus	313678	01	AC_000158.1
	313678	02	AC_000159.1
	313678	03	AC_000160.1
	313678	04	AC_000161.1
	313678	05	AC_000162.1
	313678	06	AC_000163.1
	313678	07	AC_000164.1
	313678	08	AC_000165.1
	313678	09	AC_000166.1
	313678	10	AC_000167.1
	313678	11	AC_000168.1
	313678	12	AC_000169.1
	313678	13	AC_000170.1
	313678	14	AC_000171.1
	313678	15	AC_000172.1
	313678	16	AC_000173.1
	313678	17	AC_000174.1
	313678	18	AC_000175.1
	313678	19	AC_000176.1
	313678	20	AC_000177.1
	313678	21	AC_000178.1
	313678	22	AC_000179.1
	313678	23	AC_000180.1
	313678	24	AC_000181.1
	313678	25	AC_000182.1
	313678	26	AC_000183.1
	313678	27	AC_000184.1
	313678	28	AC_000185.1
	313678	29	AC_000186.1
	313678	X	AC_000187.1
Callithrix jacchus	6608	01	NC_013896.1
	6608	02	NC_013897.1
	6608	03	NC_013898.1
	6608	04	NC_013899.1
	6608	05	NC_013900.1
	6608	06	NC_013901.1
	6608	07	NC_013902.1
	6608	08	NC_013903.1
	6608	09	NC_013904.1
	6608	0Y	NC_013919.1
	6608	10	NC_013905.1
	6608	11	NC_013906.1
	6608	12	NC_013907.1
	6608	13	NC_013908.1
	6608	14	NC_013909.1
	6608	15	NC_013910.1
	6608	16	NC_013911.1
	6608	17	NC_013912.1
	6608	18	NC_013913.1
	6608	19	NC_013914.1
	6608	20	NC_013915.1
	6608	21	NC_013916.1
	6608	22	NC_013917.1
	6608	X	NC_013918.1
	317138	01	NC_006583.3

Table B.47: Genome assemblies used in the study

Scientific name	Assembly ID	Chromosome	chromosome accession
	317138	02	NC_006584.3
	317138	03	NC_006585.3
	317138	04	NC_006586.3
	317138	05	NC_006587.3
	317138	06	NC_006588.3
	317138	07	NC_006589.3
	317138	08	NC_006590.3
	317138	09	NC_006591.3
	317138	10	NC_006592.3
	317138	11	NC_006593.3
	317138	12	NC_006594.3
	317138	13	NC_006595.3
	317138	14	NC_006596.3
	317138	15	NC_006597.3
	317138	16	NC_006598.3
	317138	17	NC_006599.3
	317138	18	NC_006600.3
	317138	19	NC_006601.3
	317138	20	NC_006602.3
	317138	21	NC_006603.3
	317138	22	NC_006604.3
	317138	23	NC_006605.3
	317138	24	NC_006606.3
	317138	25	NC_006607.3
	317138	26	NC_006608.3
	317138	27	NC_006609.3
	317138	28	NC_006610.3
	317138	29	NC_006611.3
	317138	30	NC_006612.3
	317138	31	NC_006613.3
	317138	32	NC_006614.3
	317138	33	NC_006615.3
	317138	34	NC_006616.3
	317138	35	NC_006617.3
	317138	36	NC_006618.3
	317138	37	NC_006619.3
	317138	38	NC_006620.3
	317138	X	NC_006621.3
Capra hircus	524258	01	NC_022293.1
	524258	02	NC_022294.1
	524258	03	NC_022295.1
	524258	04	NC_022296.1
	524258	05	NC_022297.1
	524258	06	NC_022298.1
	524258	07	NC_022299.1
	524258	08	NC_022300.1
	524258	09	NC_022301.1
	524258	10	NC_022302.1
	524258	11	NC_022303.1
	524258	12	NC_022304.1
	524258	13	NC_022305.1
	524258	14	NC_022306.1
	524258	15	NC_022307.1
	524258	16	NC_022308.1
	524258	17	NC_022309.1

Table B.47: Genome assemblies used in the study

Scientific name	Assembly ID	Chromosome	chromosome accession
	524258	18	NC_022310.1
	524258	19	NC_022311.1
	524258	20	NC_022312.1
	524258	21	NC_022313.1
	524258	22	NC_022314.1
	524258	23	NC_022315.1
	524258	24	NC_022316.1
	524258	25	NC_022317.1
	524258	26	NC_022318.1
	524258	27	NC_022319.1
	524258	28	NC_022320.1
	524258	29	NC_022321.1
	524258	X	NC_022322.1
Chlorocebus sabaeus	132581	01	NC_023642.1
	132581	02	NC_023643.1
	132581	03	NC_023644.1
	132581	04	NC_023645.1
	132581	05	NC_023646.1
	132581	06	NC_023647.1
	132581	07	NC_023648.1
	132581	08	NC_023649.1
	132581	09	NC_023650.1
	132581	0Y	NC_023672.1
	132581	10	NC_023651.1
	132581	11	NC_023652.1
	132581	12	NC_023653.1
	132581	13	NC_023654.1
	132581	14	NC_023655.1
	132581	15	NC_023656.1
	132581	16	NC_023657.1
	132581	17	NC_023658.1
	132581	18	NC_023659.1
	132581	19	NC_023660.1
	132581	20	NC_023661.1
	132581	21	NC_023662.1
	132581	22	NC_023663.1
	132581	23	NC_023664.1
	132581	24	NC_023665.1
	132581	25	NC_023666.1
	132581	26	NC_023667.1
	132581	27	NC_023668.1
	132581	28	NC_023669.1
	132581	29	NC_023670.1
	132581	X	NC_023671.1
	286598	01	NC_009144.2
	286598	02	NC_009145.2
	286598	03	NC_009146.2
	286598	04	NC_009147.2
	286598	05	NC_009148.2
	286598	06	NC_009149.2
	286598	07	NC_009150.2
	286598	08	NC_009151.2
	286598	09	NC_009152.2
	286598	10	NC_009153.2
	286598	11	NC_009154.2

Table B.47: Genome assemblies used in the study

Scientific name	Assembly ID	Chromosome	chromosome accession
	286598	12	NC_009155.2
	286598	13	NC_009156.2
	286598	14	NC_009157.2
	286598	15	NC_009158.2
	286598	16	NC_009159.2
	286598	17	NC_009160.2
	286598	18	NC_009161.2
	286598	19	NC_009162.2
	286598	20	NC_009163.2
	286598	21	NC_009164.2
	286598	22	NC_009165.2
	286598	23	NC_009166.2
	286598	24	NC_009167.2
	286598	25	NC_009168.2
	286598	26	NC_009169.2
	286598	27	NC_009170.2
	286598	28	NC_009171.2
	286598	29	NC_009172.2
	286598	30	NC_009173.2
	286598	31	NC_009174.2
	286598	X	NC_009175.2
Felis catus	250841	A1	NC_018723.2
	250841	A2	NC_018724.2
	250841	A3	NC_018725.2
	250841	B1	NC_018726.2
	250841	B2	NC_018727.2
	250841	B3	NC_018728.2
	250841	B4	NC_018729.2
	250841	C1	NC_018730.2
	250841	C2	NC_018731.2
	250841	D1	NC_018732.2
	250841	D2	NC_018733.2
	250841	D3	NC_018734.2
	250841	D4	NC_018735.2
	250841	E1	NC_018736.2
	250841	E2	NC_018737.2
	250841	E3	NC_018738.2
	250841	F1	NC_018739.2
	250841	F2	NC_018740.2
	250841	X	NC_018741.2
Gorilla gorilla gorilla	415668	01	NC_018424.1
	415668	03	NC_018427.1
	415668	04	NC_018428.1
	415668	05	NC_018429.1
	415668	06	NC_018430.1
	415668	07	NC_018431.1
	415668	08	NC_018432.1
	415668	09	NC_018433.1
	415668	10	NC_018434.1
	415668	11	NC_018435.1
	415668	12	NC_018436.1
	415668	13	NC_018437.1
	415668	14	NC_018438.1
	415668	15	NC_018439.1
	415668	16	NC_018440.1

Table B.47: Genome assemblies used in the study

Scientific name	Assembly ID	Chromosome	chromosome accession
	415668	17	NC_018441.1
	415668	18	NC_018442.1
	415668	19	NC_018443.1
	415668	20	NC_018444.1
	415668	21	NC_018445.1
	415668	22	NC_018446.1
	415668	2A	NC_018425.1
	415668	2B	NC_018426.1
	415668	X	NC_018447.1
Homo sapiens	599081	01	NC_000001.11
	599081	02	NC_000002.12
	599081	03	NC_000003.12
	599081	04	NC_000004.12
	599081	05	NC_000005.10
	599081	06	NC_000006.12
	599081	07	NC_000007.14
	599081	08	NC_000008.11
	599081	09	NC_000009.12
	599081	0Y	NC_000024.10
	599081	10	NC_000010.11
	599081	11	NC_000011.10
	599081	12	NC_000012.12
	599081	13	NC_000013.11
	599081	14	NC_000014.9
	599081	15	NC_000015.10
	599081	16	NC_000016.10
	599081	17	NC_000017.11
	599081	18	NC_000018.10
	599081	19	NC_000019.10
	599081	20	NC_000020.11
	599081	21	NC_000021.9
	599081	22	NC_000022.11
	599081	X	NC_000023.11
Macaca fascicularis	704988	01	NC_022272.1
	704988	02	NC_022273.1
	704988	03	NC_022274.1
	704988	04	NC_022275.1
	704988	05	NC_022276.1
	704988	06	NC_022277.1
	704988	07	NC_022278.1
	704988	08	NC_022279.1
	704988	09	NC_022280.1
	704988	10	NC_022281.1
	704988	11	NC_022282.1
	704988	12	NC_022283.1
	704988	13	NC_022284.1
	704988	14	NC_022285.1
	704988	15	NC_022286.1
	704988	16	NC_022287.1
	704988	17	NC_022288.1
	704988	18	NC_022289.1
	704988	19	NC_022290.1
	704988	20	NC_022291.1
	704988	X	NC_022292.1
	588981	01	NC_027893.1

Table B.47: Genome assemblies used in the study

Scientific name	Assembly ID	Chromosome	chromosome accession
	588981	02	NC_027894.1
	588981	03	NC_027895.1
	588981	04	NC_027896.1
	588981	05	NC_027897.1
	588981	06	NC_027898.1
	588981	07	NC_027899.1
	588981	08	NC_027900.1
	588981	09	NC_027901.1
	588981	0Y	NC_027914.1
	588981	10	NC_027902.1
	588981	11	NC_027903.1
	588981	12	NC_027904.1
	588981	13	NC_027905.1
	588981	14	NC_027906.1
	588981	15	NC_027907.1
	588981	16	NC_027908.1
	588981	17	NC_027909.1
	588981	18	NC_027910.1
	588981	19	NC_027911.1
	588981	20	NC_027912.1
	588981	X	NC_027913.1
Microtus ochrogaster	504458	01	NC_022009.1
	504458	02	NC_022010.1
	504458	04	NC_022011.1
	504458	05	NC_022012.1
	504458	06	NC_022013.1
	504458	07	NC_022014.1
	504458	08	NC_022015.1
	504458	10	NC_022016.1
	504458	15	NC_022017.1
	504458	16	NC_022018.1
	504458	17	NC_022019.1
	504458	18	NC_022020.1
	504458	19	NC_022021.1
	504458	21	NC_022022.1
	504458	22	NC_022023.1
	504458	24	NC_022024.1
	504458	26	NC_022025.1
	504458	X	NC_022026.1
Monodelphis domestica	237648	01	NC_008801.1
	237648	02	NC_008802.1
	237648	03	NC_008803.1
	237648	04	NC_008804.1
	237648	05	NC_008805.1
	237648	06	NC_008806.1
	237648	07	NC_008807.1
	237648	08	NC_008808.1
	237648	X	NC_008809.1
Mus musculus	558528	01	NC_000067.6
	558528	02	NC_000068.7
	558528	03	NC_000069.6
	558528	04	NC_000070.6
	558528	05	NC_000071.6
	558528	06	NC_000072.6
	558528	07	NC_000073.6

Table B.47: Genome assemblies used in the study

Scientific name	Assembly ID	Chromosome	chromosome accession
	558528	08	NC_000074.6
	558528	09	NC_000075.6
	558528	0Y	NC_000087.7
	558528	10	NC_000076.6
	558528	11	NC_000077.6
	558528	12	NC_000078.6
	558528	13	NC_000079.6
	558528	14	NC_000080.6
	558528	15	NC_000081.6
	558528	16	NC_000082.6
	558528	17	NC_000083.6
	558528	18	NC_000084.6
	558528	19	NC_000085.6
	558528	X	NC_000086.7
Nomascus leucogenys	506498	02	NC_019817.1
	506498	03	NC_019818.1
	506498	04	NC_019819.1
	506498	05	NC_019820.1
	506498	06	NC_019821.1
	506498	08	NC_019823.1
	506498	09	NC_019824.1
	506498	10	NC_019825.1
	506498	11	NC_019826.1
	506498	12	NC_019827.1
	506498	13	NC_019828.1
	506498	14	NC_019829.1
	506498	15	NC_019830.1
	506498	16	NC_019831.1
	506498	17	NC_019832.1
	506498	18	NC_019833.1
	506498	19	NC_019834.1
	506498	1a	NC_019816.1
	506498	20	NC_019835.1
	506498	21	NC_019836.1
	506498	22a	NC_019837.1
	506498	23	NC_019838.1
	506498	24	NC_019839.1
	506498	25	NC_019840.1
	506498	7b	NC_019822.1
	506498	X	NC_019841.1
Oryctolagus cuniculus	182491	01	NC_013669.1
	182491	02	NC_013670.1
	182491	03	NC_013671.1
	182491	04	NC_013672.1
	182491	05	NC_013673.1
	182491	06	NC_013674.1
	182491	07	NC_013675.1
	182491	08	NC_013676.1
	182491	09	NC_013677.1
	182491	10	NC_013678.1
	182491	11	NC_013679.1
	182491	12	NC_013680.1
	182491	13	NC_013681.1
	182491	14	NC_013682.1
	182491	15	NC_013683.1

Table B.47: Genome assemblies used in the study

Scientific name	Assembly ID	Chromosome	chromosome accession
	182491	16	NC_013684.1
	182491	17	NC_013685.1
	182491	18	NC_013686.1
	182491	19	NC_013687.1
	182491	20	NC_013688.1
	182491	21	NC_013689.1
	182491	X	NC_013690.1
Ovis aries	585171	01	NC_019458.2
	585171	02	NC_019459.2
	585171	03	NC_019460.2
	585171	04	NC_019461.2
	585171	05	NC_019462.2
	585171	06	NC_019463.2
	585171	07	NC_019464.2
	585171	08	NC_019465.2
	585171	09	NC_019466.2
	585171	10	NC_019467.2
	585171	11	NC_019468.2
	585171	12	NC_019469.2
	585171	13	NC_019470.2
	585171	14	NC_019471.2
	585171	15	NC_019472.2
	585171	16	NC_019473.2
	585171	17	NC_019474.2
	585171	18	NC_019475.2
	585171	19	NC_019476.2
	585171	20	NC_019477.2
	585171	21	NC_019478.2
	585171	22	NC_019479.2
	585171	23	NC_019480.2
	585171	24	NC_019481.2
	585171	25	NC_019482.2
	585171	26	NC_019483.2
	585171	X	NC_019484.2
Pan troglodytes	255628	01	NC_006468.3
	255628	03	NC_006490.3
	255628	04	NC_006471.3
	255628	05	NC_006472.3
	255628	06	NC_006473.3
	255628	07	NC_006474.3
	255628	08	NC_006475.3
	255628	09	NC_006476.3
	255628	0Y	NC_006492.3
	255628	10	NC_006477.3
	255628	11	NC_006478.3
	255628	12	NC_006479.3
	255628	13	NC_006480.3
	255628	14	NC_006481.3
	255628	15	NC_006482.3
	255628	16	NC_006483.3
	255628	17	NC_006484.3
	255628	18	NC_006485.3
	255628	19	NC_006486.3
	255628	20	NC_006487.3
	216221	21	NC_006488.3

Table B.47: Genome assemblies used in the study

Scientific name	Assembly ID	Chromosome	chromosome accession
	255628	22	NC_006489.3
	255628	2A	NC_006469.3
	255628	2B	NC_006470.3
	255628	X	NC_006491.3
Papio anubis	399268	01	NC_018152.1
	399268	02	NC_018153.1
	399268	03	NC_018154.1
	399268	04	NC_018155.1
	399268	05	NC_018156.1
	399268	06	NC_018157.1
	399268	07	NC_018158.1
	399268	08	NC_018159.1
	399268	09	NC_018160.1
	399268	10	NC_018161.1
	399268	11	NC_018162.1
	399268	12	NC_018163.1
	399268	13	NC_018164.1
	399268	14	NC_018165.1
	399268	15	NC_018166.1
	399268	16	NC_018167.1
	399268	17	NC_018168.1
	399268	18	NC_018169.1
	399268	19	NC_018170.1
	399268	20	NC_018171.1
	399268	X	NC_018172.1
Pongo abelii	395158	01	NC_012591.1
	395158	03	NC_012594.1
	395158	04	NC_012595.1
	395158	05	NC_012596.1
	395158	06	NC_012597.1
	395158	07	NC_012598.1
	395158	08	NC_012599.1
	395158	09	NC_012600.1
	395158	10	NC_012601.1
	395158	11	NC_012602.1
	395158	12	NC_012603.1
	395158	13	NC_012604.1
	395158	14	NC_012605.1
	395158	15	NC_012606.1
	395158	16	NC_012607.1
	395158	17	NC_012608.1
	395158	18	NC_012609.1
	395158	19	NC_012610.1
	395158	20	NC_012611.1
	395158	21	NC_012612.1
	395158	22	NC_012613.1
	395158	2A	NC_012592.1
	395158	2B	NC_012593.1
	395158	X	NC_012614.1
	191871	01	NC_005100.4
	191871	02	NC_005101.4
	191871	03	NC_005102.4
	191871	04	NC_005103.4
	191871	05	NC_005104.4
	191871	06	NC_005105.4

Table B.47: Genome assemblies used in the study

Scientific name	Assembly ID	Chromosome	chromosome accession
	191871	07	NC_005106.4
	191871	08	NC_005107.4
	191871	09	NC_005108.4
	191871	0Y	NC_024475.1
	191871	10	NC_005109.4
	191871	11	NC_005110.4
	191871	12	NC_005111.4
	191871	13	NC_005112.4
	191871	14	NC_005113.4
	191871	15	NC_005114.4
	191871	16	NC_005115.4
	191871	17	NC_005116.4
	191871	18	NC_005117.4
	191871	19	NC_005118.4
	191871	20	NC_005119.4
	191871	X	NC_005120.4
Sus scrofa	304498	01	NC_010443.4
	304498	02	NC_010444.3
	304498	03	NC_010445.3
	304498	04	NC_010446.4
	304498	05	NC_010447.4
	304498	06	NC_010448.3
	304498	07	NC_010449.4
	304498	08	NC_010450.3
	304498	09	NC_010451.3
	304498	0Y	NC_010462.2
	304498	10	NC_010452.3
	304498	11	NC_010453.4
	304498	12	NC_010454.3
	304498	13	NC_010455.4
	304498	14	NC_010456.4
	304498	15	NC_010457.4
	304498	16	NC_010458.3
	304498	17	NC_010459.4
	304498	18	NC_010460.3
	304498	X	NC_010461.4

Table B.48: Genome assemblies used in the study

Species	Assembly
Bos taurus	UMD3.1
Callithrix jacchus	ASM275486v1
Canis familiaris	CanFam3.1
Chlorocebus sabaeus	ChlSab1.1
Equus caballus	EquCab2
Felis catus	Felis_catus_9.0
Gorilla gorilla	.gorGor4
Homo sapiens	GRCh38
Macaca mulatta	Mmul_8.0.1
Monodelphis domestica	monDom5
Mus musculus	GRCm38
Nomascus leucogenys	Nleu_3.0
Ovis aries	Oar_v3.1
Pan troglodytes	Pan_tro_3.0
Papio anubis	Panu_3.0
Pongo abelii	PPYG2
Rattu norvegicus	Rnor_6.0
Sus scrofa	Sscrofa11.1

Table B.49: GO analysis. Enriched categories at selected depth in mouse and human for biological processes

species	GO	GO term	depth	p_fdr_bh	chromosome
Homo_sapiens	GO:0007608	sensory perception of smell	6	4.21708527557463e-88	11
	GO:0007608	sensory perception of smell	6	1.51889289806345e-42	1
	GO:0042742	defense response to bacterium	6	2.49246965790805e-18	8
	GO:0007608	sensory perception of smell	6	6.83740406785496e-16	9
	GO:0010469	regulation of receptor activity	6	2.1816439221646e-10	9
	GO:0007608	sensory perception of smell	6	2.57584435500666e-10	14
	GO:0019221	cytokine-mediated signaling pathway	6	3.42331211596682e-10	9
	GO:0007608	sensory perception of smell	6	3.94614768135192e-10	12
	GO:0042981	regulation of apoptotic process	6	4.14541896792115e-10	4
	GO:0043330	response to exogenous dsRNA	6	3.54659245786316e-09	9
	GO:0042742	defense response to bacterium	6	6.33326383821042e-09	20
	GO:0051607	defense response to virus	6	1.40463821165743e-08	9
	GO:0001580	detection of chemical stimulus involved in sensory perception of bitter taste	6	2.0936345831244e-08	12
	GO:0007608	sensory perception of smell	6	1.6237690680113e-07	6
	GO:0007608	sensory perception of smell	6	8.95473280489903e-07	3
	GO:0042742	defense response to bacterium	6	9.4731027594783e-07	6
	GO:0061844	antimicrobial humoral immune response mediated by antimicrobial peptide	6	1.46604910393086e-06	4
	GO:0007608	sensory perception of smell	6	2.27218634231578e-06	7
	GO:0019221	cytokine-mediated signaling pathway	6	2.69346697512603e-06	2
	GO:0071280	cellular response to copper ion	6	5.39497154508523e-06	16
	GO:0071294	cellular response to zinc ion	6	5.39497154508523e-06	16
	GO:0050909	sensory perception of taste	6	6.28302854672165e-06	12
	GO:0032755	positive regulation of interleukin-6 production	6	3.48635937297813e-05	2
	GO:0007608	sensory perception of smell	6	3.90730946548007e-05	17
	GO:0010469	regulation of receptor activity	6	3.98679438606913e-05	17
	GO:0071276	cellular response to cadmium ion	6	4.36238559344426e-05	16
	GO:0071347	cellular response to interleukin-1	6	4.43720029985695e-05	17
	GO:0060968	regulation of gene silencing	6	5.57063119243323e-05	6
	GO:0007608	sensory perception of smell	6	0.000152978213098681	15
	GO:0071222	cellular response to lipopolysaccharide	6	0.000224423556929832	2
	GO:0071222	cellular response to lipopolysaccharide	6	0.000437148143630299	4
	GO:0010469	regulation of receptor activity	6	0.00111760327551263	4
	GO:0060326	cell chemotaxis	6	0.00119344671030012	3
	GO:0061844	antimicrobial humoral immune response mediated by antimicrobial peptide	6	0.00127791985224567	8
	GO:0071356	cellular response to tumor necrosis factor	6	0.00147701290796805	17
	GO:0050832	defense response to fungus	6	0.00178046332380012	8
	GO:0061844	antimicrobial humoral immune response mediated by antimicrobial peptide	6	0.00315270349244047	6

Table B.49: GO analysis. Enriched categories at selected depth in mouse and human for biological processes

species	GO	GO term	depth	p_fdr_bh	chromosome
	GO:0061844	antimicrobial humoral immune response mediated by antimicrobial peptide	6	0.00343849558507611	2
	GO:0042981	regulation of apoptotic process	6	0.0040983603675449	8
	GO:0051607	defense response to virus	6	0.00568638407712805	22
	GO:0007608	sensory perception of smell	6	0.0105336552513843	19
	GO:0071346	cellular response to interferon-gamma	6	0.0141552031584823	17
	GO:0001913	T cell mediated cytotoxicity	6	0.0168449576082598	6
	GO:0002479	antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent	6	0.0168449576082598	6
	GO:0007608	sensory perception of smell	6	0.0206794194344771	5
	GO:0071395	cellular response to jasmonic acid stimulus	6	0.026070920673981	10
	GO:0044598	doxorubicin metabolic process	6	0.026070920673981	10
	GO:0006898	receptor-mediated endocytosis	6	0.026078800805003	2
	GO:0090305	nucleic acid phosphodiester bond hydrolysis	6	0.0289255278891878	14
	GO:0015671	oxygen transport	6	0.0335946376154424	16
	GO:0050909	sensory perception of taste	6	0.0337748240101389	7
	GO:0043434	response to peptide hormone	6	0.0356518949965671	2
	GO:0007519	skeletal muscle tissue development	6	0.0399517058031403	12
	GO:0002480	antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-independent	6	0.0419652787460946	6
	GO:0010469	regulation of receptor activity	6	0.04336368910453	2
	GO:0017121	phospholipid scrambling	6	0.0473752717688974	3
	GO:0042267	natural killer cell mediated cytotoxicity	6	0.0477821534159313	6
	GO:0007608	sensory perception of smell	6	1.5573814555991e-193	2
	GO:0007608	sensory perception of smell	6	1.05333535679868e-96	7
	GO:0007608	sensory perception of smell	6	5.47829066916301e-81	9
	GO:0007608	sensory perception of smell	6	2.11915150107754e-51	11
	GO:0007608	sensory perception of smell	6	1.1862110678921e-42	10
	GO:0007608	sensory perception of smell	6	1.91126023091457e-39	19
	GO:0042742	defense response to bacterium	6	2.04774212931842e-31	8
	GO:0007608	sensory perception of smell	6	1.24348631245485e-28	17
	GO:0007608	sensory perception of smell	6	4.72625166733471e-19	1
	GO:0061844	antimicrobial humoral immune response mediated by antimicrobial peptide	6	5.10783721114977e-18	8
	GO:0007608	sensory perception of smell	6	4.02085660117788e-17	16
	GO:0050909	sensory perception of taste	6	9.91742881468636e-16	6
	GO:0007608	sensory perception of smell	6	1.13417335328823e-15	14
	GO:0051607	defense response to virus	6	3.82318541680495e-15	4
	GO:0007608	sensory perception of smell	6	1.36400573777545e-14	4
	GO:0001580	detection of chemical stimulus involved in sensory perception of bitter taste	6	1.42510573269166e-14	6

Mus_musculus

Table B.49: GO analysis. Enriched categories at selected depth in mouse and human for biological processes

species	GO	GO term	depth	p_fdr_bh	chromosome
	GO:0043330	response to exogenous dsRNA	6	1.40925817992277e-13	4
	GO:0071222	cellular response to lipopolysaccharide	6	1.49083611705873e-13	8
	GO:0019221	cytokine-mediated signaling pathway	6	8.65099933621852e-13	4
	GO:0007608	sensory perception of smell	6	6.22341209611904e-12	6
	GO:0050909	sensory perception of taste	6	4.18612975716179e-09	7
	GO:0010469	regulation of receptor activity	6	1.51436227727289e-08	4
	GO:0007608	sensory perception of smell	6	1.72905965910573e-05	13
	GO:0007608	sensory perception of smell	6	2.81001463829565e-05	15
	GO:0070584	mitochondrion morphogenesis	6	4.79467675752365e-05	4
	GO:0031649	heat generation	6	4.79467675752365e-05	4
	GO:0010888	negative regulation of lipid storage	6	4.79467675752365e-05	4
	GO:0042742	defense response to bacterium	6	7.16881290803292e-05	2
	GO:0030162	regulation of proteolysis	6	7.62735091574265e-05	3
	GO:0051607	defense response to virus	6	0.000151009567621272	5
	GO:0008207	C21-steroid hormone metabolic process	6	0.000244732043703717	3
	GO:0071346	cellular response to interferon-gamma	6	0.000260236343985398	5
	GO:0060326	cell chemotaxis	6	0.000427454875925172	8
	GO:0043434	response to peptide hormone	6	0.000612157928884465	12
	GO:0061844	antimicrobial humoral immune response mediated by antimicrobial peptide	6	0.000921899395375275	13
	GO:0071347	cellular response to interleukin-1	6	0.000986021486345496	11
	GO:0061844	antimicrobial humoral immune response mediated by antimicrobial peptide	6	0.00124632968879284	5
	GO:0045834	positive regulation of lipid metabolic process	6	0.00148056717221677	4
	GO:0060968	regulation of gene silencing	6	0.00267686322425951	13
	GO:0035458	cellular response to interferon-beta	6	0.00310198318263796	11
	GO:0009060	aerobic respiration	6	0.00456861398687385	4
	GO:0051603	proteolysis involved in cellular protein catabolic process	6	0.00881974252133851	13
	GO:0010469	regulation of receptor activity	6	0.0126573443398326	13
	GO:0008626	granzyme-mediated apoptotic signaling pathway	6	0.0138175087093708	14
	GO:0090305	nucleic acid phosphodiester bond hydrolysis	6	0.026305875252421	14
	GO:0071222	cellular response to lipopolysaccharide	6	0.0267891457713527	5
	GO:0071356	cellular response to tumor necrosis factor	6	0.0270830571741903	11

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_10X0	65265657	65462923	2	197266
	1	1_11X0	66968007	67046927	2	78920
	1	1_12X0	67559268	67659266	2	99998
	1	1_13X0	69935163	70106415	2	171252
	1	1_14X0	73740503	73802039	3	61536
	1	1_15X0	74545919	74769754	2	223835
	1	1_16X0	77469357	77611958	2	142601
	1	1_17X0	81097636	81209114	4	111478
	1	1_18X0	83643182	83721672	3	78490
	1	1_19X0	84062791	84211374	2	148583
	1	1_1X0	401206	492432	2	91226
	1	1_20X0	95125675	95166199	2	40524
	1	1_21X0	98303615	98369907	3	66292
	1	1_22X0	107879883	108058677	2	178794
	1	1_23X0	109661962	109729694	2	67732
	1	1_24X0	116974391	117168355	2	193964
	1	1_25X0	117636823	117835976	5	199153
	1	1_26X0	119605842	119823457	3	217615
	1	1_27X0	120236686	120315813	2	79127
	1	1_28X0	121863151	121886919	2	23768
	1	1_29X0	122821128	123273104	4	451976
	1	1_2X0	1376827	1627137	4	250310
	1	1_30X0	130588843	130710118	2	121275
	1	1_31X0	131089308	131170538	3	81230
	1	1_32X0	136573249	136676796	2	103547
	1	1_33X0	136683311	136790649	2	107338
	1	1_34X0	143119272	143204865	2	85593
	1	1_35X0	143569076	143755849	2	186773
	1	1_36X0	144067291	144196576	3	129285
	1	1_37X0	145801880	145887681	2	85801
	1	1_38X0	145901591	146077070	4	175479
	1	1_39X0	146437114	146583242	4	146128
	1	1_3X0	4775170	4956408	4	181238
	1	1_40X0	147404396	147570736	2	166340
	1	1_41X0	150054221	150153934	3	99713
	1	1_4X0	5035870	5096443	2	60573
	1	1_5X0	8811792	8963815	2	152023
	1	1_6X0	41700079	41824958	2	124879
	1	1_7X0	41955547	42469403	9	513856
	1	1_8X0	54481399	54516574	2	35175
	1	1_9X0	58015564	58125875	2	110311
	10	10_10X0	21920977	22009402	6	88425
	10	10_11X0	23636371	23934477	6	298106
	10	10_11X1	23330788	23632510	10	301722
	10	10_11X2	23978360	24283878	13	305518
	10	10_11X3	22552579	22900718	10	348139
	10	10_11X4	23099397	23221949	4	122552
	10	10_12X0	24646142	25548213	26	902071
	10	10_13X0	25589275	25654241	5	64966
	10	10_14X0	26113954	26632354	14	518400
	10	10_15X0	26667693	26686302	2	18609
	10	10_16X0	26880575	27524781	33	644206
	10	10_16X1	27641548	28256440	22	614892
	10	10_16X2	28281890	28578629	2	296739
	10	10_17X0	35903021	36013346	2	110325

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	10	10_18X0	37346275	37664678	4	318403
	10	10_19X0	42749352	42808107	2	58755
	10	10_1X0	5409681	5655684	2	246003
	10	10_20X0	43278649	43484579	2	205930
	10	10_21X0	52648527	52771273	2	122746
	10	10_22X0	58569062	58711813	3	142751
	10	10_23X0	59961836	60037841	2	76005
	10	10_24X0	60202862	60207710	2	4848
	10	10_25X0	61600903	61751865	2	150962
	10	10_26X0	65492593	65558771	4	66178
	10	10_27X0	76943981	76979552	2	35571
	10	10_28X0	80135428	80167935	2	32507
	10	10_29X0	82467738	82532119	3	64381
	10	10_2X0	7777600	7964929	3	187329
	10	10_30X0	85358158	85449397	4	91239
	10	10_31X0	86557298	86762452	2	205154
	10	10_32X0	86883739	87132202	3	248463
	10	10_3X0	10072118	10151271	2	79153
	10	10_4X0	12218639	12291146	2	72507
	10	10_5X0	20207483	20354153	2	146670
	10	10_6X0	20569774	20600789	2	31015
	10	10_7X0	20665546	20934454	8	268908
	10	10_8X0	21325414	21377051	2	51637
	10	10_9X0	21652617	21663680	2	11063
	11	11_10X0	19578982	19780728	2	201746
	11	11_11X0	21526906	21761348	2	234442
	11	11_12X0	26124239	26174650	2	50411
	11	11_13X0	27200797	27263699	2	62902
	11	11_14X0	44707028	44785035	4	78007
	11	11_15X0	46349422	46706152	10	356730
	11	11_16X0	47008475	47174623	4	166148
	11	11_17X0	47915426	47995400	2	79974
	11	11_18X0	49060227	49173648	4	113421
	11	11_19X0	49249763	49264015	2	14252
	11	11_1X0	65466	181573	2	116107
	11	11_20X0	56528811	56642712	2	113901
	11	11_21X0	67246030	67289445	3	43415
	11	11_22X0	72067123	72312020	4	244897
	11	11_23X0	74984268	75023692	2	39424
	11	11_24X0	92215374	92374682	3	159308
	11	11_25X0	93121981	93158576	2	36595
	11	11_26X0	93268704	94190302	35	921598
	11	11_27X0	95514365	95586792	2	72427
	11	11_28X0	97582042	97652610	2	70568
	11	11_29X0	98509474	98541363	2	31889
	11	11_2X0	1747909	1891332	3	143423
	11	11_30X0	98568364	98594176	2	25812
	11	11_31X0	100210927	100227150	2	16223
	11	11_32X0	103181661	103328557	6	146896
	11	11_33X0	104373528	104505740	2	132212
	11	11_34X0	106244291	106440173	11	195882
	11	11_35X0	107191850	107261718	3	69868
	11	11_3X0	1935216	1957806	2	22590
	11	11_4X0	2690660	2762524	2	71864
	11	11_5X0	2768909	2782703	2	13794

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	11	11_6X0	6721485	7221257	6	499772
	11	11_7X0	7229706	7414636	2	184930
	11	11_8X0	10105101	10122607	2	17506
	11	11_9X0	13572118	13596639	2	24521
	12	12_10X0	90508105	90541989	3	33884
	12	12_11X0	91148973	91156716	2	7743
	12	12_1X0	11363648	11443817	2	80169
	12	12_2X0	21514646	21593686	2	79040
	12	12_3X0	29991971	30039901	2	47930
	12	12_4X0	32317772	32357227	2	39455
	12	12_5X0	36310779	36374313	3	63534
	12	12_6X0	80095128	80244063	3	148935
	12	12_7X0	80714189	80724166	2	9977
	12	12_8X0	82907167	83115415	2	208248
	12	12_9X0	87654333	87684129	2	29796
	13	13_10X0	43654794	44214277	12	559483
	13	13_11X0	47400413	47449390	2	48977
	13	13_12X0	50354055	50400983	13	46928
	13	13_13X0	52563657	52576188	2	12531
	13	13_14X0	53169426	53304213	2	134787
	13	13_15X0	53668954	53941056	4	272102
	13	13_16X0	54247557	54255011	2	7454
	13	13_17X0	54591299	54606206	2	14907
	13	13_18X0	60208924	60232034	2	23110
	13	13_19X0	61082324	61187210	2	104886
	13	13_1X0	216569	339068	2	122499
	13	13_20X0	61314568	61615456	7	300888
	13	13_21X0	61900820	62140119	2	239299
	13	13_22X0	63196226	63329709	3	133483
	13	13_22X1	62850752	62950669	3	99917
	13	13_23X0	65532018	65592912	2	60894
	13	13_24X0	67835164	67974925	3	139761
	13	13_25X0	74202932	74339570	7	136638
	13	13_26X0	74526199	75284423	13	758224
	13	13_27X0	75354606	75362820	2	8214
	13	13_28X0	75885062	75942663	2	57601
	13	13_2X0	17540994	17639994	2	99000
	13	13_3X0	17788337	17883719	3	95382
	13	13_4X0	28204600	28275573	2	70973
	13	13_5X0	41031399	41107079	2	75680
	13	13_6X0	42217371	42247239	2	29868
	13	13_7X0	42463338	42566091	5	102753
	13	13_8X0	43253789	43422283	2	168494
	13	13_9X0	43493944	43508178	2	14234
	14	14_10X0	79372712	79733755	5	361043
	14	14_11X0	79862685	79896055	2	33370
	14	14_1X0	1136304	1322214	4	185910
	14	14_2X0	1445334	1788087	4	342753
	14	14_3X0	1825982	2240765	5	414783
	14	14_4X0	2266062	2540475	3	274413
	14	14_5X0	2563160	2856167	11	293007
	14	14_6X0	23263815	23395443	2	131628
	14	14_7X0	24847257	24976948	2	129691
	14	14_8X0	25105062	25179651	2	74589
	14	14_9X0	46644609	46891306	5	246697

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	15	15_10X0	29001571	29233495	2	231924
	15	15_11X0	29257448	29389165	5	131717
	15	15_12X0	29441565	29469747	3	28182
	15	15_13X0	38041580	38340230	3	298650
	15	15_14X0	43413545	43585976	2	172431
	15	15_15X0	43703911	43770016	2	66105
	15	15_16X0	45175108	45575215	15	400107
	15	15_17X0	46368597	46412894	2	44297
	15	15_18X0	46417938	46732271	17	314333
	15	15_19X0	46902643	47005492	2	102849
	15	15_1X0	771399	1296403	9	525004
	15	15_1X1	182645	578564	6	395919
	15	15_20X0	48405343	49085845	33	680502
	15	15_20X1	47384234	48332456	35	948222
	15	15_20X10	55103366	55122512	2	19146
	15	15_20X2	50412563	51073006	29	660443
	15	15_20X3	49098192	49535828	18	437636
	15	15_20X4	49562175	49852276	13	290101
	15	15_20X5	49922808	50404606	17	481798
	15	15_20X6	51316034	51797639	16	481605
	15	15_20X7	52184590	52605536	10	420946
	15	15_20X8	53492850	53564982	2	72132
	15	15_20X9	54196170	54224051	2	27881
	15	15_21X0	55676067	55973229	6	297162
	15	15_22X0	56827092	56974713	3	147621
	15	15_23X0	61698326	61767248	2	68922
	15	15_24X0	75032941	75068534	2	35593
	15	15_25X0	77719053	77966408	5	247355
	15	15_26X0	79563651	79981032	22	417381
	15	15_26X1	80244344	80829405	23	585061
	15	15_26X2	79135786	79508721	18	372935
	15	15_26X3	80842358	81738412	22	896054
	15	15_27X0	81869684	81956829	3	87145
	15	15_28X0	82511487	83157432	21	645945
	15	15_29X0	83365200	83515102	4	149902
	15	15_2X0	5912037	6400322	8	488285
	15	15_30X0	83580546	83605727	2	25181
	15	15_31X0	83767120	84279377	15	512257
	15	15_32X0	84361364	84404937	2	43573
	15	15_33X0	84485248	84691574	7	206326
	15	15_34X0	84899737	85011313	3	111576
	15	15_3X0	6585516	6654015	2	68499
	15	15_4X0	22547742	22557344	2	9602
	15	15_5X0	22566929	22572575	2	5646
	15	15_6X0	24779413	24851695	2	72282
	15	15_7X0	25419219	25524857	2	105638
	15	15_8X0	27871672	27934085	3	62413
	15	15_9X0	28909315	28963808	2	54493
	16	16_10X0	24929465	24998201	2	68736
	16	16_11X0	27702967	27840009	2	137042
	16	16_12X0	29663583	29721104	2	57521
	16	16_13X0	36963321	36979962	2	16641
	16	16_14X0	38049259	38192433	3	143174
	16	16_15X0	39506590	39759444	5	252854
	16	16_16X0	42376258	42493773	2	117515

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	16	16_17X0	42700409	42713502	2	13093
	16	16_18X0	42881664	42912457	3	30793
	16	16_19X0	47798009	47952892	2	154883
	16	16_1X0	297300	759737	5	462437
	16	16_20X0	52264771	52379152	2	114381
	16	16_21X0	53272046	53302109	2	30063
	16	16_22X0	55718640	56118319	12	399679
	16	16_23X0	56552275	56706787	2	154512
	16	16_24X0	64965975	65015946	2	49971
	16	16_25X0	65028277	65210675	3	182398
	16	16_26X0	65545497	65771215	2	225718
	16	16_27X0	70886830	71016271	2	129441
	16	16_28X0	72740451	72832974	2	92523
	16	16_29X0	77490629	77911587	4	420958
	16	16_2X0	925083	1069419	3	144336
	16	16_3X0	2117057	2121226	2	4169
	16	16_4X0	2557488	2683555	2	126067
	16	16_5X0	2763988	3108506	3	344518
	16	16_6X0	3168635	3250256	2	81621
	16	16_7X0	4113295	4365926	3	252631
	16	16_8X0	4402474	4612613	7	210139
	16	16_9X0	4765914	5123974	3	358060
	17	17_10X0	63278104	63311098	2	32994
	17	17_11X0	63611774	63678163	3	66389
	17	17_12X0	64899652	64913110	2	13458
	17	17_13X0	64913867	65014401	2	100534
	17	17_14X0	71551242	71619967	3	68725
	17	17_15X0	72690558	72754666	2	64108
	17	17_16X0	72808678	73202019	11	393341
	17	17_17X0	73273379	73319008	5	45629
	17	17_18X0	73425746	73452741	2	26995
	17	17_19X0	73784639	73828749	2	44110
	17	17_1X0	2823044	2884552	2	61508
	17	17_20X0	74319809	74519770	4	199961
	17	17_21X0	74607593	74613377	2	5784
	17	17_22X0	74663544	74880531	2	216987
	17	17_2X0	14230809	14333746	2	102937
	17	17_3X0	25033671	25057633	2	23962
	17	17_4X0	35768716	35873721	2	105005
	17	17_5X0	44506802	44665571	2	158769
	17	17_6X0	44966897	45223845	6	256948
	17	17_7X0	51260767	51648371	3	387604
	17	17_8X0	55035080	55054749	2	19669
	17	17_9X0	56173049	56275072	2	102023
	18	18_10X0	25446367	25504738	3	58371
	18	18_11X0	25608275	25753920	3	145645
	18	18_12X0	34394579	34556068	4	161489
	18	18_13X0	34757953	34810221	3	52268
	18	18_14X0	34918530	35117221	3	198691
	18	18_15X0	35576307	35587974	2	11667
	18	18_16X0	36095734	36243479	2	147745
	18	18_17X0	39204698	39218132	2	13434
	18	18_18X0	45337460	45440643	2	103183
	18	18_19X0	45562957	45689249	2	126292
	18	18_1X0	1772754	1821405	2	48651

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	18	18_20X0	45705247	45891786	3	186539
	18	18_21X0	46033469	46078935	4	45466
	18	18_22X0	46184320	46239044	2	54724
	18	18_23X0	46243828	46314341	4	70513
	18	18_24X0	46398986	46410960	2	11974
	18	18_25X0	46721282	46890070	7	168788
	18	18_26X0	47193774	47422682	5	228908
	18	18_26X1	47431025	47883111	9	452086
	18	18_26X2	46960041	47180829	6	220788
	18	18_27X0	47928221	47971681	2	43460
	18	18_28X0	48769012	48810095	3	41083
	18	18_29X0	49214936	49281315	2	66379
	18	18_2X0	2590736	2611931	3	21195
	18	18_30X0	49552815	49580639	2	27824
	18	18_31X0	49872449	50013518	3	141069
	18	18_32X0	50040253	50053543	2	13290
	18	18_33X0	50345655	50666554	5	320899
	18	18_34X0	51373917	51507155	3	133238
	18	18_35X0	51707795	52334826	15	627031
	18	18_36X0	52380158	52675994	8	295836
	18	18_36X1	52732800	53007832	5	275032
	18	18_37X0	53634883	53777767	2	142884
	18	18_38X0	54753334	54781917	2	28583
	18	18_39X0	55173175	55212513	2	39338
	18	18_3X0	2662463	2838848	3	176385
	18	18_40X0	55235067	55287983	2	52916
	18	18_41X0	55369206	55444508	4	75302
	18	18_42X0	55753309	55832924	3	79615
	18	18_43X0	56175613	56250611	3	74998
	18	18_44X0	57338709	57499297	13	160588
	18	18_45X0	57537670	57910816	17	373146
	18	18_45X1	57938695	58220103	11	281408
	18	18_46X0	58494237	58682409	7	188172
	18	18_46X1	58688854	59004703	3	315849
	18	18_48X0	60352473	60466851	2	114378
	18	18_48X1	60117464	60344819	4	227355
	18	18_48X2	60545413	61032896	6	487483
	18	18_48X3	59692586	60077406	4	384820
	18	18_49X0	61040309	61417516	7	377207
	18	18_4X0	12379750	12434808	3	55058
	18	18_50X0	61608937	61951470	7	342533
	18	18_51X0	62052428	62121184	3	68756
	18	18_52X0	62194105	62260863	2	66758
	18	18_53X0	62347609	62638443	8	290834
	18	18_54X0	62717575	62735263	2	17688
	18	18_55X0	62827495	63486978	26	659483
	18	18_56X0	63520186	63604895	4	84709
	18	18_57X0	63610150	63683298	3	73148
	18	18_58X0	63687553	63923147	14	235594
	18	18_58X1	63968610	64282112	7	313502
	18	18_59X0	64607953	64891741	9	283788
	18	18_59X1	65538898	65980686	11	441788
	18	18_59X2	64908196	65078903	5	170707
	18	18_59X3	65132596	65506427	7	373831
	18	18_5X0	13811387	13964622	2	153235

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	18	18_6X0	16551061	16732300	2	181239
	18	18_7X0	24031967	24098654	2	66687
	18	18_8X0	24106722	24165538	6	58816
	18	18_9X0	24765330	24854379	2	89049
	19	19_10X0	19641109	19757359	2	116250
	19	19_11X0	19833725	19952293	5	118568
	19	19_12X0	20615339	20796711	4	181372
	19	19_13X0	21725905	21844511	2	118606
	19	19_14X0	23403657	23430667	2	27010
	19	19_15X0	24470226	24802537	22	332311
	19	19_16X0	24854533	24913053	2	58520
	19	19_17X0	25010757	25135095	2	124338
	19	19_18X0	25530011	25634931	2	104920
	19	19_19X0	26949113	26975114	2	26001
	19	19_1X0	118123	150708	2	32585
	19	19_20X0	27331075	27434258	3	103183
	19	19_21X0	27455597	27461346	2	5749
	19	19_22X0	27467161	27532634	3	65473
	19	19_23X0	27893175	27904808	2	11633
	19	19_24X0	28310097	28371440	3	61343
	19	19_25X0	28913867	29001212	2	87345
	19	19_26X0	29956798	30251091	6	294293
	19	19_27X0	33673396	33739141	2	65745
	19	19_28X0	34483124	34619450	4	136326
	19	19_29X0	36610213	36675085	2	64872
	19	19_2X0	6313771	6464109	3	150338
	19	19_30X0	36811806	36845371	2	33565
	19	19_31X0	36877017	36927313	2	50296
	19	19_32X0	37296591	37320363	2	23772
	19	19_33X0	38428299	38596261	8	167962
	19	19_34X0	39194570	39296505	2	101935
	19	19_35X0	40894257	40955437	2	61180
	19	19_36X0	41017271	41048228	2	30957
	19	19_37X0	41515535	41841840	13	326305
	19	19_38X0	41846176	41867171	3	20995
	19	19_39X0	41903870	42113449	14	209579
	19	19_3X0	9209310	9352851	8	143541
	19	19_40X0	42185570	42484047	15	298477
	19	19_41X0	42666504	42689873	2	23369
	19	19_42X0	42749145	42809906	2	60761
	19	19_43X0	42812569	42903081	2	90512
	19	19_44X0	42960226	43132624	3	172398
	19	19_45X0	43316471	43365350	2	48879
	19	19_46X0	43508614	43571700	5	63086
	19	19_47X0	44310879	44412456	2	101577
	19	19_48X0	44437199	44458077	3	20878
	19	19_49X0	45444316	45459216	2	14900
	19	19_4X0	9441828	9484176	2	42348
	19	19_50X0	45973041	46071153	2	98112
	19	19_52X0	48447902	48487616	2	39714
	19	19_53X0	48570157	48656396	2	86239
	19	19_54X0	48750715	48867609	3	116894
	19	19_55X0	49175892	49292988	2	117096
	19	19_56X0	50537804	50562978	2	25174
	19	19_57X0	50827280	50849924	2	22644

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	19	19_58X0	50867137	50984331	4	117194
	19	19_59X0	51461750	51469340	2	7590
	19	19_5X0	11865527	11951411	2	85884
	19	19_60X0	51499066	51528822	2	29756
	19	19_61X0	53318969	53377423	3	58454
	19	19_62X0	54633144	54666987	2	33843
	19	19_63X0	55749013	55822126	2	73113
	19	19_64X0	55904213	55918416	2	14203
	19	19_65X0	56359579	56377735	2	18156
	19	19_66X0	57097872	57121003	2	23131
	19	19_67X0	57328131	57621461	7	293330
	19	19_68X0	61185603	61229571	2	43968
	19	19_69X0	61873253	62151579	4	278326
	19	19_6X0	12536191	12573533	2	37342
	19	19_70X0	63621226	63745798	3	124572
	19	19_7X0	14568542	14831157	11	262615
	19	19_8X0	15098771	15184294	3	85523
	19	19_9X0	16110980	16234839	4	123859
	2	2_10X0	72049607	72250165	2	200558
	2	2_11X0	79939485	80043984	2	104499
	2	2_12X0	89517811	89858016	3	340205
	2	2_13X0	90139395	90301820	2	162425
	2	2_14X0	92325180	92587607	3	262427
	2	2_15X0	105338358	105397646	2	59288
	2	2_16X0	106906773	107014997	3	108224
	2	2_17X0	107544683	107576619	2	31936
	2	2_18X0	107921714	107942525	2	20811
	2	2_19X0	120845481	120913882	4	68401
	2	2_1X0	916140	994208	2	78068
	2	2_20X0	120975092	120993968	2	18876
	2	2_21X0	121847012	121975687	2	128675
	2	2_22X0	122030511	122146489	3	115978
	2	2_23X0	122503653	122649721	2	146068
	2	2_24X0	122814786	122835401	2	20615
	2	2_25X0	126276335	126518906	2	242571
	2	2_26X0	129315264	129378160	2	62896
	2	2_27X0	130769172	130795255	3	26083
	2	2_28X0	132470750	132529468	2	58718
	2	2_29X0	133088083	133376607	8	288524
	2	2_2X0	7139738	7344001	2	204263
	2	2_30X0	133953059	133983362	2	30303
	2	2_31X0	135812788	136103406	5	290618
	2	2_3X0	19257062	19326224	2	69162
	2	2_4X0	20759890	20856159	8	96269
	2	2_5X0	24489074	24506672	2	17598
	2	2_6X0	26619585	26798046	2	178461
	2	2_7X0	34297726	34571758	3	274032
	2	2_8X0	44930623	44973211	2	42588
	2	2_9X0	71099131	71285350	2	186219
	20	20_1X0	13234	133868	4	120634
	20	20_2X0	6721327	6794112	2	72785
	20	20_3X0	9179368	9315274	2	135906
	20	20_4X0	10154909	10265817	2	110908
	20	20_5X0	10300553	10314075	2	13522
	20	20_6X0	23212633	23355111	2	142478

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	20	20_7X0	24033747	24107687	3	73940
	20	20_8X0	58563065	58664249	2	101184
	20	20_9X0	71182950	71318130	3	135180
	21	21_10X0	34231528	34734990	6	503462
	21	21_11X0	34903758	34952418	2	48660
	21	21_12X0	35112322	35267979	10	155657
	21	21_13X0	47130288	47198449	2	68161
	21	21_14X0	49637067	49738270	2	101203
	21	21_15X0	55515505	55579813	2	64308
	21	21_16X0	59496451	59830426	9	333975
	21	21_17X0	60927758	61032824	5	105066
	21	21_18X0	62100003	62136163	2	36160
	21	21_19X0	62701850	62772148	2	70298
	21	21_1X0	736183	784819	2	48636
	21	21_20X0	66852437	66882899	2	30462
	21	21_21X0	68150604	68151509	2	905
	21	21_22X0	69418979	69466051	2	47072
	21	21_23X0	71453627	71596399	6	142772
	21	21_2X0	20002742	20030816	2	28074
	21	21_3X0	20082490	20685028	14	602538
	21	21_4X0	20800158	20886141	2	85983
	21	21_5X0	22661292	22871700	3	210408
	21	21_6X0	26543823	26600056	2	56233
	21	21_7X0	26811825	26962870	2	151045
	21	21_8X0	31483268	31537227	3	53959
	21	21_9X0	34088423	34230892	2	142469
	22	22_10X0	49557992	49630194	6	72202
	22	22_11X0	50593627	50601799	2	8172
	22	22_12X0	50655648	50756457	4	100809
	22	22_13X0	50789615	50892474	2	102859
	22	22_14X0	50928327	50969031	2	40704
	22	22_15X0	51335299	51482027	2	146728
	22	22_16X0	52143306	52221924	9	78618
	22	22_17X0	53236358	53343856	5	107498
	22	22_18X0	53567797	54057181	9	489384
	22	22_19X0	55497704	55714644	2	216940
	22	22_1X0	5650	34397	3	28747
	22	22_20X0	57122412	57204951	2	82539
	22	22_21X0	58768219	58783528	2	15309
	22	22_22X0	59609594	59642750	2	33156
	22	22_2X0	6789385	7024065	3	234680
	22	22_3X0	11674127	11894086	4	219959
	22	22_4X0	12015691	12356183	3	340492
	22	22_5X0	12613418	12679638	2	66220
	22	22_6X0	16367996	16523810	5	155814
	22	22_7X0	16877938	16896124	2	18186
	22	22_8X0	48615397	48669271	3	53874
	22	22_9X0	49282960	49552848	4	269888
	23	23_10X0	19369192	19395134	2	25942
	23	23_11X0	20708929	20758331	2	49402
	23	23_12X0	22130935	22261583	3	130648
	23	23_13X0	22997916	23122686	2	124770
	23	23_14X0	24346988	24397956	2	50968
	23	23_15X0	24823652	24977317	6	153665
	23	23_16X0	25351507	25656497	11	304990

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	23	23_17X0	25855146	25980368	2	125222
	23	23_18X0	26916511	27011946	3	95435
	23	23_19X0	27083669	27196793	4	113124
	23	23_1X0	6975320	7220390	12	245070
	23	23_20X0	27225270	27243367	2	18097
	23	23_21X0	27331773	27338328	2	6555
	23	23_22X0	27407611	27446642	5	39031
	23	23_23X0	27451817	27560628	5	108811
	23	23_24X0	27600055	27867532	7	267477
	23	23_25X0	28177338	28224731	2	47393
	23	23_26X0	28299437	28506312	5	206875
	23	23_27X0	28530978	28676470	6	145492
	23	23_28X0	28828555	29207889	25	379334
	23	23_28X1	29223821	29463723	11	239902
	23	23_28X2	29480530	29785578	15	305048
	23	23_29X0	30079857	30343084	9	263227
	23	23_2X0	7426487	7571700	2	145213
	23	23_30X0	30349199	30696900	14	347701
	23	23_31X0	30713382	30768030	16	54648
	23	23_32X0	30978065	30996482	2	18417
	23	23_33X0	31111555	31127004	4	15449
	23	23_34X0	31235365	31252079	2	16714
	23	23_35X0	31343551	31373121	2	29570
	23	23_36X0	31456886	31674743	32	217857
	23	23_37X0	31721014	31861069	4	140055
	23	23_38X0	31914984	31918508	2	3524
	23	23_39X0	34357710	34818877	10	461167
	23	23_3X0	9944807	10045045	2	100238
	23	23_40X0	34861310	35113759	4	252449
	23	23_41X0	45312387	45372316	2	59929
	23	23_42X0	47328131	47352621	2	24490
	23	23_43X0	50354843	50394279	2	39436
	23	23_44X0	50394684	50417050	2	22366
	23	23_45X0	50515222	50678845	5	163623
	23	23_46X0	52388044	52445489	3	57445
	23	23_4X0	11093689	11132861	2	39172
	23	23_5X0	13057053	13181324	3	124271
	23	23_6X0	15081867	15163188	4	81321
	23	23_7X0	15901126	15921069	2	19943
	23	23_8X0	16640766	16794939	2	154173
	23	23_9X0	17257357	17269998	2	12641
	24	24_10X0	47246084	47297664	2	51580
	24	24_11X0	62248909	62638833	12	389924
	24	24_1X0	314607	511067	3	196460
	24	24_2X0	4231234	4278276	2	47042
	24	24_3X0	12694119	12765314	2	71195
	24	24_4X0	22080132	22100704	2	20572
	24	24_5X0	25388382	25496186	2	107804
	24	24_6X0	25913549	26448151	7	534602
	24	24_7X0	37820258	37910130	2	89872
	24	24_8X0	43983880	44024553	2	40673
	24	24_9X0	45787843	45924550	2	136707
	25	25_10X0	2536214	2855519	12	319305
	25	25_11X0	14469282	14643306	2	174024
	25	25_12X0	18102959	18143260	2	40301

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	25	25_13X0	18207129	18413889	3	206760
	25	25_14X0	18539756	18656582	2	116826
	25	25_15X0	21587069	21620834	2	33765
	25	25_16X0	25139417	25268874	2	129457
	25	25_17X0	26361862	26381153	2	19291
	25	25_18X0	26437044	26547977	2	110933
	25	25_19X0	26831623	27161868	6	330245
	25	25_1X0	192905	217650	6	24745
	25	25_20X0	27222530	27230804	2	8274
	25	25_21X0	27382282	27418925	2	36643
	25	25_22X0	27431639	27447066	2	15427
	25	25_23X0	27450356	27501201	3	50845
	25	25_24X0	27602222	27736719	2	134497
	25	25_25X0	27811921	27837967	2	26046
	25	25_26X0	33212217	33493476	3	281259
	25	25_27X0	33983208	34083970	4	100762
	25	25_28X0	35018574	35039183	2	20609
	25	25_29X0	35930035	36088161	2	158126
	25	25_2X0	292324	347325	2	55001
	25	25_30X0	36120546	36276673	3	156127
	25	25_31X0	36673965	36855872	2	181907
	25	25_32X0	36905467	36947947	3	42480
	25	25_33X0	37025335	37242576	3	217241
	25	25_34X0	37343014	37469998	4	126984
	25	25_35X0	37546485	37587556	2	41071
	25	25_36X0	39006993	39089918	4	82925
	25	25_37X0	42259657	42282988	2	23331
	25	25_3X0	517559	552281	2	34722
	25	25_4X0	649253	660655	2	11402
	25	25_5X0	987390	1022334	4	34944
	25	25_6X0	1524318	1742107	4	217789
	25	25_7X0	1796660	1945919	2	149259
	25	25_8X0	2126944	2275410	8	148466
	25	25_9X0	2433850	2438483	2	4633
	26	26_10X0	21823912	21896298	2	72386
	26	26_11X0	24275663	24309613	3	33950
	26	26_12X0	25060238	25119647	3	59409
	26	26_13X0	37192556	37366138	3	173582
	26	26_14X0	42824587	42877424	3	52837
	26	26_15X0	43034075	43086628	2	52553
	26	26_16X0	43257908	43272421	2	14513
	26	26_17X0	49504091	49620555	2	116464
	26	26_18X0	50573881	50681525	4	107644
	26	26_1X0	34600	106867	2	72267
	26	26_2X0	337785	402566	2	64781
	26	26_3X0	10211198	10540469	5	329271
	26	26_4X0	11069934	11127464	4	57530
	26	26_5X0	14447529	14463796	2	16267
	26	26_6X0	15931985	16538258	7	606273
	26	26_7X0	18430347	18443469	2	13122
	26	26_8X0	19201799	19222118	2	20319
	26	26_9X0	21711025	21730353	2	19328
	27	27_1X0	4788447	6196146	18	1407699
	27	27_2X0	13208048	13210908	2	2860
	27	27_3X0	15326026	15370080	2	44054

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	27	27_4X0	16630109	16850282	4	220173
	27	27_5X0	32774355	32915668	2	141313
	27	27_6X0	33816038	34625508	6	809470
	27	27_7X0	34686577	34774452	2	87875
	27	27_8X0	37164686	37209034	2	44348
	27	27_9X0	41695665	41883521	2	187856
	28	28_10X0	35592783	35854765	5	261982
	28	28_11X0	35923906	35958219	2	34313
	28	28_12X0	39505113	39527943	2	22830
	28	28_13X0	41912322	41924801	2	12479
	28	28_14X0	42612012	42640763	2	28751
	28	28_15X0	44819689	45003218	5	183529
	28	28_16X0	46026267	46053236	3	26969
	28	28_1X0	2556212	2855792	7	299580
	28	28_2X0	12923499	13246045	4	322546
	28	28_3X0	25207673	25249032	2	41359
	28	28_4X0	25335074	25399768	2	64694
	28	28_5X0	25752421	25924368	2	171947
	28	28_6X0	27613178	27685834	2	72656
	28	28_7X0	29074151	29105086	2	30935
	28	28_8X0	30008049	30009656	2	1607
	28	28_9X0	30993339	31048603	2	55264
	29	29_10X0	29933714	29979585	2	45871
	29	29_11X0	32757309	32829939	2	72630
	29	29_12X0	33546279	33703975	2	157696
	29	29_13X0	36684172	36809433	2	125261
	29	29_14X0	37098498	37165575	2	67077
	29	29_15X0	37639043	37725179	4	86136
	29	29_16X0	37892577	38026747	2	134170
	29	29_17X0	38135609	38668543	13	532934
	29	29_18X0	38727762	39270024	13	542262
	29	29_19X0	39326930	39635792	7	308862
	29	29_1X0	5175376	5474084	5	298708
	29	29_20X0	39662643	40007570	7	344927
	29	29_21X0	40034226	40466766	8	432540
	29	29_22X0	40940932	41102449	3	161517
	29	29_23X0	41378327	41504719	3	126392
	29	29_24X0	41645576	41846745	2	201169
	29	29_25X0	41944850	42372760	8	427910
	29	29_26X0	42426069	42451969	2	25900
	29	29_27X0	42478123	42608367	2	130244
	29	29_28X0	42735180	42791900	2	56720
	29	29_29X0	43126404	43154681	2	28277
	29	29_2X0	5682483	5892565	4	210082
	29	29_30X0	43416170	43471503	2	55333
	29	29_31X0	43648614	43696730	2	48116
	29	29_32X0	44370408	44460617	2	90209
	29	29_33X0	44766936	44771529	2	4593
	29	29_34X0	45299519	45347624	3	48105
	29	29_35X0	45844928	45971982	2	127054
	29	29_36X0	45988625	46046958	2	58333
	29	29_37X0	46143363	46203015	2	59652
	29	29_38X0	46969455	47011851	3	42396
	29	29_39X0	47591667	47699877	3	108210
	29	29_3X0	9962193	9997407	2	35214

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	29	29_40X0	48988149	49026889	2	38740
	29	29_41X0	49842969	49926003	2	83034
	29	29_42X0	50036101	50062631	2	26530
	29	29_43X0	50218484	50287648	2	69164
	29	29_44X0	50560713	50699835	2	139122
	29	29_4X0	17974605	18121808	2	147203
	29	29_5X0	26390466	26553359	4	162893
	29	29_6X0	26693329	26734926	3	41597
	29	29_7X0	27820502	28423743	27	603241
	29	29_7X1	26802126	27265974	21	463848
	29	29_7X2	27301299	27787786	11	486487
	29	29_8X0	28595576	28767534	3	171958
	29	29_9X0	29429085	29459846	2	30761
	3	3_10X0	10977984	11355939	15	377955
	3	3_10X1	11412357	12189914	20	777557
	3	3_11X0	12308303	12907593	6	599290
	3	3_12X0	13743911	13784728	2	40817
	3	3_13X0	14019444	14055462	2	36018
	3	3_14X0	14755582	14915294	2	159712
	3	3_15X0	15521494	15591685	3	70191
	3	3_16X0	15603805	15636992	2	33187
	3	3_17X0	18881347	18970566	2	89219
	3	3_18X0	19361111	19557450	2	196339
	3	3_19X0	19779460	19806629	2	27169
	3	3_1X0	20751	96237	2	75486
	3	3_20X0	19994998	20047228	2	52230
	3	3_21X0	20770059	20825680	11	55621
	3	3_22X0	20871006	20975445	2	104439
	3	3_23X0	21519979	21605094	2	85115
	3	3_24X0	21859235	21999397	2	140162
	3	3_25X0	22310554	22349850	2	39296
	3	3_26X0	22768703	22899870	2	131167
	3	3_27X0	23510977	23558293	3	47316
	3	3_28X0	26158672	26414580	3	255908
	3	3_29X0	26735793	26889208	2	153415
	3	3_2X0	7881817	8040506	4	158689
	3	3_30X0	32047399	32284496	4	237097
	3	3_31X0	32851233	33035163	3	183930
	3	3_32X0	33768050	33880598	6	112548
	3	3_33X0	33952094	34013930	2	61836
	3	3_34X0	39885182	39948945	2	63763
	3	3_35X0	42703661	42766060	2	62399
	3	3_36X0	53475696	53862657	3	386961
	3	3_37X0	54031912	54525762	8	493850
	3	3_38X0	54587049	55075636	7	488587
	3	3_39X0	57499184	57792193	5	293009
	3	3_3X0	8229236	8508122	3	278886
	3	3_40X0	59252316	59351124	2	98808
	3	3_41X0	66282837	66318248	2	35411
	3	3_42X0	86404311	86640839	4	236528
	3	3_43X0	93978825	94046956	2	68131
	3	3_44X0	94049248	94197896	2	148648
	3	3_45X0	98660627	98722981	2	62354
	3	3_46X0	99666161	99957408	5	291247
	3	3_47X0	100285863	100338715	2	52852

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	3	3_48X0	100401226	100419716	2	18490
	3	3_49X0	103238486	103422206	2	183720
	3	3_4X0	8519572	8742728	4	223156
	3	3_50X0	103480291	103689535	8	209244
	3	3_51X0	104603593	104612391	2	8798
	3	3_52X0	106216832	106275253	2	58421
	3	3_53X0	106917951	107097489	2	179538
	3	3_54X0	108508805	108555472	2	46667
	3	3_55X0	110338488	110558237	3	219749
	3	3_56X0	111481702	111525013	4	43311
	3	3_57X0	113943011	114031371	4	88360
	3	3_58X0	118191441	118223089	2	31648
	3	3_59X0	119687614	120111016	16	423402
	3	3_5X0	8794793	9293445	8	498652
	3	3_6X0	9468431	9576282	2	107851
	3	3_7X0	9605197	9670496	2	65299
	3	3_8X0	9695129	10000540	8	305411
	3	3_9X0	10236168	10683411	10	447243
	4	4_10X0	49263419	49472239	2	208820
	4	4_11X0	52173110	52235915	2	62805
	4	4_12X0	65671028	66090181	3	419153
	4	4_13X0	69252420	69402503	11	150083
	4	4_14X0	74836426	74911776	2	75350
	4	4_15X0	76705105	76725301	2	20196
	4	4_16X0	83280521	83478290	5	197769
	4	4_17X0	88965021	89161368	4	196347
	4	4_18X0	93475459	93524304	2	48845
	4	4_19X0	94861073	94949979	3	88906
	4	4_1X0	11060376	11079756	2	19380
	4	4_20X0	99016785	99140956	2	124171
	4	4_21X0	99787244	99948547	2	161303
	4	4_22X0	103799172	103862938	2	63766
	4	4_23X0	105874500	106937743	56	1063243
	4	4_24X0	106996940	107062131	2	65191
	4	4_25X0	107184504	107216578	3	32074
	4	4_26X0	107295010	107347462	3	52452
	4	4_27X0	107720864	108378588	16	657724
	4	4_28X0	113021969	113630153	10	608184
	4	4_29X0	113675490	114151331	11	475841
	4	4_2X0	12389567	12520270	3	130703
	4	4_30X0	114199462	114212312	2	12850
	4	4_31X0	115246102	115360689	2	114587
	4	4_3X0	14276357	14296237	2	19880
	4	4_4X0	25258619	25333477	2	74858
	4	4_5X0	27854574	27890570	2	35996
	4	4_6X0	32077891	32269539	2	191648
	4	4_7X0	33026387	33193999	2	167612
	4	4_8X0	40432906	40643369	3	210463
	4	4_9X0	48956493	49092515	2	136022
	5	5_10X0	30070404	30103142	3	32738
	5	5_11X0	30155763	30298109	2	142346
	5	5_12X0	30713144	30868119	4	154975
	5	5_13X0	30918838	31097555	4	178717
	5	5_14X0	31483512	31776650	7	293138
	5	5_14X1	31859948	32264561	5	404613

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	5	5_15X0	34028553	34198996	2	170443
	5	5_16X0	44351615	44720976	11	369361
	5	5_17X0	56334244	56350443	2	16199
	5	5_18X0	56863815	57026537	3	162722
	5	5_19X0	57303355	57318372	2	15017
	5	5_1X0	4824769	5021025	3	196256
	5	5_20X0	57486017	57699615	5	213598
	5	5_21X0	59178071	59717840	19	539769
	5	5_21X1	59793588	60158008	12	364420
	5	5_21X2	58751279	59170684	10	419405
	5	5_21X3	58260957	58619384	8	358427
	5	5_21X4	57934910	58208574	9	273664
	5	5_22X0	72049628	72082472	3	32844
	5	5_23X0	74631082	75084280	7	453198
	5	5_24X0	75810521	75827375	2	16854
	5	5_25X0	89258893	89509795	3	250902
	5	5_26X0	95516924	95531308	2	14384
	5	5_27X0	97402843	97453094	2	50251
	5	5_28X0	99035752	99243044	8	207292
	5	5_29X0	99974365	100401104	13	426739
	5	5_29X1	100479659	100961494	9	481835
	5	5_29X2	99463834	99870803	6	406969
	5	5_2X0	10330469	10342660	2	12191
	5	5_30X0	101133130	101545688	4	412558
	5	5_31X0	101686492	101730873	2	44381
	5	5_32X0	102076231	102168602	4	92371
	5	5_33X0	102940224	103207769	4	267545
	5	5_34X0	103634595	103779250	4	144655
	5	5_35X0	103987752	104173568	4	185816
	5	5_36X0	104302583	104415588	3	113005
	5	5_37X0	105538144	105764041	3	225897
	5	5_38X0	106016835	106126587	2	109752
	5	5_39X0	106157909	106216757	2	58848
	5	5_3X0	20909663	21119087	4	209424
	5	5_40X0	106975989	107052441	2	76452
	5	5_41X0	107649324	107709872	2	60548
	5	5_42X0	109933973	110017878	2	83905
	5	5_43X0	111045317	111088530	2	43213
	5	5_44X0	113596984	113625976	4	28992
	5	5_45X0	114494264	114627819	2	133555
	5	5_46X0	114966366	115075690	3	109324
	5	5_47X0	115517820	115608935	2	91115
	5	5_48X0	119821435	119837101	2	15666
	5	5_4X0	26136686	26252315	9	115629
	5	5_5X0	26695587	26759487	2	63900
	5	5_6X0	26769555	26985686	3	216131
	5	5_7X0	27171925	27910785	30	738860
	5	5_8X0	28038532	28106733	2	68201
	5	5_9X0	28887967	29172757	5	284790
	6	6_10X0	85177921	85716159	6	538238
	6	6_11X0	85926576	86865690	12	939114
	6	6_12X0	86917981	87094952	3	176971
	6	6_13X0	87189736	87239005	2	49269
	6	6_14X0	90232762	90312046	3	79284
	6	6_15X0	90559882	90824270	7	264388

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	6	6_16X0	90992364	91480129	4	487765
	6	6_17X0	92600177	92646326	3	46149
	6	6_18X0	93633306	93758039	2	124733
	6	6_19X0	98916665	98990087	2	73422
	6	6_1X0	2385668	2423445	2	37777
	6	6_20X0	99655644	99754969	2	99325
	6	6_21X0	103971537	104056137	2	84600
	6	6_22X0	104341538	104453232	2	111694
	6	6_23X0	115687877	115858942	2	171065
	6	6_24X0	117798819	117923220	2	124401
	6	6_2X0	5344468	5499008	2	154540
	6	6_3X0	23096104	23254634	2	158530
	6	6_4X0	26547390	26916216	6	368826
	6	6_5X0	37479804	37793279	3	313475
	6	6_6X0	59670233	59706713	2	36480
	6	6_7X0	68405247	68530673	2	125426
	6	6_8X0	69144752	69187082	2	42330
	6	6_9X0	73838854	73997338	2	158484
	7	7_10X0	11459646	11831715	3	372069
	7	7_11X0	13922169	13939754	2	17585
	7	7_12X0	14527360	15288535	21	761175
	7	7_13X0	15302198	15485011	5	182813
	7	7_14X0	15647684	15709872	5	62188
	7	7_15X0	15894542	16252264	7	357722
	7	7_16X0	17051782	17254000	5	202218
	7	7_17X0	17757622	17814125	3	56503
	7	7_18X0	18306841	18582713	3	275872
	7	7_19X0	18709398	18834685	2	125287
	7	7_1X0	2184137	2394402	4	210265
	7	7_20X0	18960639	18970399	2	9760
	7	7_21X0	19076010	19190363	3	114353
	7	7_22X0	19235826	19258805	2	22979
	7	7_23X0	19846208	19936959	3	90751
	7	7_24X0	20816181	20841683	2	25502
	7	7_25X0	21132389	21429075	3	296686
	7	7_26X0	21440493	21478598	2	38105
	7	7_27X0	21907306	22195989	8	288683
	7	7_28X0	22368335	22385179	2	16844
	7	7_29X0	23329260	23425468	2	96208
	7	7_2X0	2453803	2555050	3	101247
	7	7_30X0	40445995	40479104	2	33109
	7	7_31X0	41154755	41780369	16	625614
	7	7_32X0	42084486	42914770	25	830284
	7	7_32X1	43044539	43567568	17	523029
	7	7_32X2	43593814	43978409	9	384595
	7	7_32X3	44096134	44522295	6	426161
	7	7_33X0	44797942	45034833	7	236891
	7	7_34X0	45354486	45462547	2	108061
	7	7_35X0	53487090	53507322	2	20232
	7	7_36X0	53585216	54281944	25	696728
	7	7_37X0	54624665	54702923	2	78258
	7	7_38X0	54857213	55041418	4	184205
	7	7_39X0	61446179	61734503	3	288324
	7	7_3X0	2948734	3074600	2	125866
	7	7_40X0	63379480	63641028	3	261548

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	7	7_41X0	64524873	64684500	2	159627
	7	7_42X0	70213686	70348692	2	135006
	7	7_43X0	98582890	98878164	3	295274
	7	7_4X0	5655109	5671232	2	16123
	7	7_5X0	7972554	8330618	15	358064
	7	7_6X0	8389985	8651746	7	261761
	7	7_7X0	8754284	8805446	2	51162
	7	7_8X0	8999056	9715604	16	716548
	7	7_9X0	9911860	10557186	12	645326
	7	7_9X1	10687060	11340452	13	653392
	8	8_10X0	65834278	66114341	4	280063
	8	8_11X0	70804358	70896918	2	92560
	8	8_12X0	71663195	71702901	2	39706
	8	8_13X0	72403760	72596739	3	192979
	8	8_14X0	73147853	73184369	2	36516
	8	8_15X0	76468680	76531334	2	62654
	8	8_16X0	77260361	77368154	2	107793
	8	8_17X0	77394838	77422532	2	27694
	8	8_18X0	82396095	82491694	2	95599
	8	8_19X0	83524639	83581560	2	56921
	8	8_1X0	7099561	7162249	4	62688
	8	8_20X0	85453132	85579683	4	126551
	8	8_21X0	95761537	96153861	10	392324
	8	8_22X0	100138374	100144156	2	5782
	8	8_23X0	101484812	101552954	2	68142
	8	8_24X0	103906601	103943464	2	36863
	8	8_25X0	104101561	104206275	2	104714
	8	8_26X0	104218993	104296480	2	77487
	8	8_2X0	8573278	8633235	3	59957
	8	8_3X0	22114512	22151549	2	37037
	8	8_4X0	22563337	23310956	38	747619
	8	8_5X0	39437808	39514616	2	76808
	8	8_6X0	43780986	43972570	3	191584
	8	8_7X0	60207472	60386110	3	178638
	8	8_8X0	60483506	60682800	12	199294
	8	8_9X0	62935793	63029685	2	93892
	9	9_10X0	71468912	71804207	18	335295
	9	9_11X0	71833315	71913826	2	80511
	9	9_12X0	75979324	76116593	3	137269
	9	9_13X0	85813503	85961158	2	147655
	9	9_14X0	88232044	88639753	7	407709
	9	9_15X0	91215612	91292179	2	76567
	9	9_16X0	97750746	98014739	3	263993
	9	9_17X0	103460841	103477813	2	16972
	9	9_1X0	274516	320732	2	46216
	9	9_2X0	12969175	12982550	2	13375
	9	9_3X0	25681086	25766095	2	85009
	9	9_4X0	34605304	34689575	3	84271
	9	9_5X0	34872096	34893768	2	21672
	9	9_6X0	39618967	39759908	2	140941
	9	9_7X0	45452891	45522649	2	69758
	9	9_8X0	61667230	61836839	2	169609
	9	9_9X0	64859552	64929935	2	70383
	X	X_10X0	39187430	39363370	7	175940
	X	X_11X0	39805525	40247095	7	441570

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	X	X_12X0	50449123	50531696	3	82573
	X	X_13X0	55171775	55451996	7	280221
	X	X_14X0	57142324	57477875	6	335551
	X	X_15X0	57664213	57715859	2	51646
	X	X_16X0	58242198	58505529	2	263331
	X	X_17X0	63896594	64080463	2	183869
	X	X_18X0	78187378	78639361	3	451983
	X	X_19X0	82642268	82779631	2	137363
	X	X_1X0	3646867	3701262	2	54395
	X	X_20X0	84029852	84080527	2	50675
	X	X_21X0	84662118	84745441	2	83323
	X	X_22X0	85499886	85702893	3	203007
	X	X_23X0	90997109	91275259	3	278150
	X	X_24X0	92287878	92320415	4	32537
	X	X_25X0	95546166	95714844	2	168678
	X	X_26X0	99440294	99529359	3	89065
	X	X_27X0	102429175	102885917	3	456742
	X	X_28X0	107357735	107383173	2	25438
	X	X_29X0	111387160	111633453	3	246293
	X	X_2X0	4355493	4518975	3	163482
	X	X_30X0	121951935	121979509	2	27574
	X	X_31X0	123603172	124337612	9	734440
	X	X_32X0	126010550	126053438	2	42888
	X	X_33X0	135436138	135503524	2	67386
	X	X_34X0	138379924	138472330	2	92406
	X	X_35X0	138986761	139033616	2	46855
	X	X_36X0	140154706	140215657	2	60951
	X	X_37X0	141075572	141104299	2	28727
	X	X_38X0	148319908	148780901	6	460993
	X	X_3X0	15127648	15454970	3	327322
	X	X_4X0	18812423	19117275	4	304852
	X	X_5X0	23534850	23651841	2	116991
	X	X_6X0	24998138	25009992	2	11854
	X	X_7X0	32630887	32979646	5	348759
	X	X_8X0	33460693	33634304	2	173611
	X	X_9X0	37593403	37598760	2	5357
	NTIC01000001.1	NTIC01000001.1_10X0	27736090	27757925	2	21835
	NTIC01000001.1	NTIC01000001.1_11X0	29476204	29528012	2	51808
	NTIC01000001.1	NTIC01000001.1_12X0	30426664	30489181	3	62517
	NTIC01000001.1	NTIC01000001.1_13X0	33052806	33154785	3	101979
	NTIC01000001.1	NTIC01000001.1_14X0	33725699	33777264	3	51565
	NTIC01000001.1	NTIC01000001.1_15X0	35170811	35190226	2	19415
	NTIC01000001.1	NTIC01000001.1_16X0	35301117	35318773	2	17656
	NTIC01000001.1	NTIC01000001.1_17X0	38688573	38708887	2	20314
	NTIC01000001.1	NTIC01000001.1_18X0	40550940	40582060	2	31120
	NTIC01000001.1	NTIC01000001.1_19X0	41611939	41662007	2	50068
	NTIC01000001.1	NTIC01000001.1_1X0	2304640	2319373	2	14733
	NTIC01000001.1	NTIC01000001.1_20X0	45523252	45531801	2	8549
	NTIC01000001.1	NTIC01000001.1_21X0	45694938	46041683	12	346745
	NTIC01000001.1	NTIC01000001.1_22X0	55275461	55332760	2	57299
	NTIC01000001.1	NTIC01000001.1_23X0	58191586	58272983	2	81397
	NTIC01000001.1	NTIC01000001.1_24X0	63851807	64042297	5	190490
	NTIC01000001.1	NTIC01000001.1_25X0	73800536	73873017	5	72481
	NTIC01000001.1	NTIC01000001.1_26X0	75147638	75167773	2	20135
	NTIC01000001.1	NTIC01000001.1_27X0	76074673	76141527	2	66854

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	NTIC01000001.1	NTIC01000001.1_28X0	87448829	87487790	2	38961
	NTIC01000001.1	NTIC01000001.1_29X0	88069765	88325077	6	255312
	NTIC01000001.1	NTIC01000001.1_2X0	3086113	3135242	4	49129
	NTIC01000001.1	NTIC01000001.1_30X0	90169078	90244201	2	75123
	NTIC01000001.1	NTIC01000001.1_31X0	103951536	104054204	2	102668
	NTIC01000001.1	NTIC01000001.1_32X0	131975407	132110665	4	135258
	NTIC01000001.1	NTIC01000001.1_33X0	132807368	132881480	2	74112
	NTIC01000001.1	NTIC01000001.1_34X0	134848223	134932164	2	83941
	NTIC01000001.1	NTIC01000001.1_35X0	196527611	196550257	2	22646
	NTIC01000001.1	NTIC01000001.1_36X0	209300568	209363670	3	63102
	NTIC01000001.1	NTIC01000001.1_3X0	4946745	4961532	2	14787
	NTIC01000001.1	NTIC01000001.1_4X0	6954853	6992621	2	37768
	NTIC01000001.1	NTIC01000001.1_5X0	16770724	16837142	3	66418
	NTIC01000001.1	NTIC01000001.1_6X0	22416800	22462273	2	45473
	NTIC01000001.1	NTIC01000001.1_7X0	22953971	23019145	2	65174
	NTIC01000001.1	NTIC01000001.1_8X0	23217562	23574158	11	356596
	NTIC01000001.1	NTIC01000001.1_9X0	25269848	25329253	2	59405
	NTIC01000002.1	NTIC01000002.1_10X0	68132297	68199479	2	67182
	NTIC01000002.1	NTIC01000002.1_11X0	68642698	68759075	9	116377
	NTIC01000002.1	NTIC01000002.1_12X0	69343250	69358745	2	15495
	NTIC01000002.1	NTIC01000002.1_13X0	69763776	70162056	13	398280
	NTIC01000002.1	NTIC01000002.1_14X0	70700064	70707318	2	7254
	NTIC01000002.1	NTIC01000002.1_15X0	71946722	71967126	2	20404
	NTIC01000002.1	NTIC01000002.1_16X0	83032567	83080362	2	47795
	NTIC01000002.1	NTIC01000002.1_17X0	108507433	108666945	4	159512
	NTIC01000002.1	NTIC01000002.1_18X0	108852405	108923808	2	71403
	NTIC01000002.1	NTIC01000002.1_19X0	110212497	110257684	3	45187
	NTIC01000002.1	NTIC01000002.1_1X0	18571902	18632932	2	61030
	NTIC01000002.1	NTIC01000002.1_20X0	110496022	110546198	2	50176
	NTIC01000002.1	NTIC01000002.1_21X0	111377366	111514156	3	136790
	NTIC01000002.1	NTIC01000002.1_22X0	113412579	113490985	2	78406
	NTIC01000002.1	NTIC01000002.1_23X0	115212679	115276670	4	63991
	NTIC01000002.1	NTIC01000002.1_24X0	117812819	117848530	2	35711
	NTIC01000002.1	NTIC01000002.1_25X0	118111282	118621184	10	509902
	NTIC01000002.1	NTIC01000002.1_26X0	118896305	119213195	11	316890
	NTIC01000002.1	NTIC01000002.1_27X0	120783004	120822272	2	39268
	NTIC01000002.1	NTIC01000002.1_2X0	18981618	19088618	3	107000
	NTIC01000002.1	NTIC01000002.1_3X0	44481864	44624695	3	142831
	NTIC01000002.1	NTIC01000002.1_4X0	51114046	51125563	2	11517
	NTIC01000002.1	NTIC01000002.1_5X0	63772390	63788528	2	16138
	NTIC01000002.1	NTIC01000002.1_6X0	64727951	64751459	3	23508
	NTIC01000002.1	NTIC01000002.1_7X0	66699045	66779775	2	80730
	NTIC01000002.1	NTIC01000002.1_8X0	66928496	67660162	22	731666
	NTIC01000002.1	NTIC01000002.1_9X0	68034328	68116471	2	82143
	NTIC01001113.1	NTIC01001113.1_10X0	26404491	26488635	2	84144
	NTIC01001113.1	NTIC01001113.1_11X0	28036735	28103657	2	66922
	NTIC01001113.1	NTIC01001113.1_12X0	46726210	46772377	3	46167
	NTIC01001113.1	NTIC01001113.1_13X0	50777043	50806624	2	29581
	NTIC01001113.1	NTIC01001113.1_14X0	61100231	61177424	2	77193
	NTIC01001113.1	NTIC01001113.1_15X0	62852050	62890928	2	38878
	NTIC01001113.1	NTIC01001113.1_16X0	66910265	67652934	25	742669
	NTIC01001113.1	NTIC01001113.1_16X1	65473707	66607667	31	1133960
	NTIC01001113.1	NTIC01001113.1_16X2	68136963	68445447	9	308484
	NTIC01001113.1	NTIC01001113.1_16X3	69232761	69433952	5	201191
	NTIC01001113.1	NTIC01001113.1_16X4	64620231	64963682	4	343451

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	NTIC01001113.1	NTIC01001113.1_16X5	63546125	63608888	2	62763
	NTIC01001113.1	NTIC01001113.1_16X6	80222689	80278620	2	55931
	NTIC01001113.1	NTIC01001113.1_17X0	104228815	104265619	2	36804
	NTIC01001113.1	NTIC01001113.1_18X0	108409832	109959505	41	1549673
	NTIC01001113.1	NTIC01001113.1_19X0	110096043	110109329	2	13286
	NTIC01001113.1	NTIC01001113.1_1X0	306174	746375	9	440201
	NTIC01001113.1	NTIC01001113.1_20X0	110727219	111125364	13	398145
	NTIC01001113.1	NTIC01001113.1_21X0	111789512	112156578	7	367066
	NTIC01001113.1	NTIC01001113.1_22X0	112268567	112314955	2	46388
	NTIC01001113.1	NTIC01001113.1_23X0	112469898	113003467	8	533569
	NTIC01001113.1	NTIC01001113.1_24X0	113971838	114151013	3	179175
	NTIC01001113.1	NTIC01001113.1_25X0	114426770	114475369	2	48599
	NTIC01001113.1	NTIC01001113.1_26X0	114730995	114805060	2	74065
	NTIC01001113.1	NTIC01001113.1_27X0	115398958	115439291	2	40333
	NTIC01001113.1	NTIC01001113.1_28X0	115609176	115744360	4	135184
	NTIC01001113.1	NTIC01001113.1_29X0	115913038	115972664	2	59626
	NTIC01001113.1	NTIC01001113.1_2X0	9416736	9423039	2	6303
	NTIC01001113.1	NTIC01001113.1_30X0	116741252	116772586	2	31334
	NTIC01001113.1	NTIC01001113.1_31X0	116957778	117016582	2	58804
	NTIC01001113.1	NTIC01001113.1_32X0	117715289	117808681	2	93392
	NTIC01001113.1	NTIC01001113.1_33X0	118796118	118860842	2	64724
	NTIC01001113.1	NTIC01001113.1_34X0	119642931	119713123	2	70192
	NTIC01001113.1	NTIC01001113.1_35X0	121575249	121707585	3	132336
	NTIC01001113.1	NTIC01001113.1_36X0	124296734	124327923	2	31189
	NTIC01001113.1	NTIC01001113.1_37X0	124642386	124737173	2	94787
	NTIC01001113.1	NTIC01001113.1_3X0	11423811	11504772	2	80961
	NTIC01001113.1	NTIC01001113.1_4X0	14286240	14326356	3	40116
	NTIC01001113.1	NTIC01001113.1_5X0	15329836	15352833	2	22997
	NTIC01001113.1	NTIC01001113.1_6X0	15665064	15798761	5	133697
	NTIC01001113.1	NTIC01001113.1_7X0	15846535	15870403	2	23868
	NTIC01001113.1	NTIC01001113.1_8X0	21269701	22039987	19	770286
	NTIC01001113.1	NTIC01001113.1_9X0	22219995	22407617	5	187622
	NTIC01002224.1	NTIC01002224.1_10X0	93316509	93548200	8	231691
	NTIC01002224.1	NTIC01002224.1_11X0	93607922	94416689	12	808767
	NTIC01002224.1	NTIC01002224.1_12X0	99150163	99735422	9	585259
	NTIC01002224.1	NTIC01002224.1_13X0	100570555	100748669	3	178114
	NTIC01002224.1	NTIC01002224.1_14X0	100844883	100886436	2	41553
	NTIC01002224.1	NTIC01002224.1_15X0	101169988	101300225	4	130237
	NTIC01002224.1	NTIC01002224.1_16X0	103466764	103533167	2	66403
	NTIC01002224.1	NTIC01002224.1_17X0	108344359	108434980	2	90621
	NTIC01002224.1	NTIC01002224.1_18X0	112489650	112614733	4	125083
	NTIC01002224.1	NTIC01002224.1_1X0	12942900	12966555	2	23655
	NTIC01002224.1	NTIC01002224.1_2X0	15479702	15485587	2	5885
	NTIC01002224.1	NTIC01002224.1_3X0	27980947	28127208	10	146261
	NTIC01002224.1	NTIC01002224.1_4X0	38576647	38663058	2	86411
	NTIC01002224.1	NTIC01002224.1_5X0	49428823	49448092	2	19269
	NTIC01002224.1	NTIC01002224.1_6X0	51034214	51166178	3	131964
	NTIC01002224.1	NTIC01002224.1_7X0	52471706	52486965	2	15259
	NTIC01002224.1	NTIC01002224.1_8X0	56051330	56136445	2	85115
	NTIC01002224.1	NTIC01002224.1_9X0	93247200	93260115	2	12915
	NTIC01003335.1	NTIC01003335.1_10X0	99722097	100261811	9	539714
	NTIC01003335.1	NTIC01003335.1_1X0	6430665	6903495	18	472830
	NTIC01003335.1	NTIC01003335.1_2X0	8262588	8294580	2	31992
	NTIC01003335.1	NTIC01003335.1_3X0	14508193	14537041	2	28848
	NTIC01003335.1	NTIC01003335.1_4X0	36434382	37092542	5	658160

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	NTIC01003335.1	NTIC01003335.1_5X0	40526577	40588321	2	61744
	NTIC01003335.1	NTIC01003335.1_6X0	42554125	42614419	2	60294
	NTIC01003335.1	NTIC01003335.1_7X0	54040687	54068597	2	27910
	NTIC01003335.1	NTIC01003335.1_8X0	67691072	68215798	7	524726
	NTIC01003335.1	NTIC01003335.1_9X0	71808594	71900644	3	92050
	NTIC01004446.1	NTIC01004446.1_10X0	31701328	31728125	2	26797
	NTIC01004446.1	NTIC01004446.1_11X0	32634088	32660880	2	26792
	NTIC01004446.1	NTIC01004446.1_12X0	33427231	33446847	2	19616
	NTIC01004446.1	NTIC01004446.1_13X0	36157905	36198365	3	40460
	NTIC01004446.1	NTIC01004446.1_14X0	40116970	40152669	2	35699
	NTIC01004446.1	NTIC01004446.1_15X0	40550803	40590402	3	39599
	NTIC01004446.1	NTIC01004446.1_16X0	44944770	45012827	2	68057
	NTIC01004446.1	NTIC01004446.1_17X0	70749995	70966612	5	216617
	NTIC01004446.1	NTIC01004446.1_18X0	71438693	71619774	4	181081
	NTIC01004446.1	NTIC01004446.1_19X0	75243428	75260060	2	16632
	NTIC01004446.1	NTIC01004446.1_1X0	14765831	14805726	2	39895
	NTIC01004446.1	NTIC01004446.1_20X0	79497644	79512572	2	14928
	NTIC01004446.1	NTIC01004446.1_21X0	85711027	85745178	3	34151
	NTIC01004446.1	NTIC01004446.1_22X0	86523196	86564900	2	41704
	NTIC01004446.1	NTIC01004446.1_23X0	98608731	98816764	4	208033
	NTIC01004446.1	NTIC01004446.1_24X0	105298255	105330665	2	32410
	NTIC01004446.1	NTIC01004446.1_25X0	105553437	105567435	2	13998
	NTIC01004446.1	NTIC01004446.1_2X0	14867658	15034687	3	167029
	NTIC01004446.1	NTIC01004446.1_3X0	17347084	17383561	2	36477
	NTIC01004446.1	NTIC01004446.1_4X0	22356848	22382037	2	25189
	NTIC01004446.1	NTIC01004446.1_5X0	22741583	23638951	16	897368
	NTIC01004446.1	NTIC01004446.1_6X0	23639653	23694213	3	54560
	NTIC01004446.1	NTIC01004446.1_7X0	23824775	23951912	3	127137
	NTIC01004446.1	NTIC01004446.1_8X0	24628812	24765034	3	136222
	NTIC01004446.1	NTIC01004446.1_9X0	27046123	27119986	3	73863
	NTIC01005557.1	NTIC01005557.1_10X0	20483272	20561776	2	78504
	NTIC01005557.1	NTIC01005557.1_11X0	28234921	28326155	4	91234
	NTIC01005557.1	NTIC01005557.1_12X0	32795252	32813472	2	18220
	NTIC01005557.1	NTIC01005557.1_13X0	36902147	36926216	2	24069
	NTIC01005557.1	NTIC01005557.1_14X0	38899217	38935453	2	36236
	NTIC01005557.1	NTIC01005557.1_15X0	63278189	63361349	2	83160
	NTIC01005557.1	NTIC01005557.1_16X0	64461671	64512515	2	50844
	NTIC01005557.1	NTIC01005557.1_17X0	69128215	69200178	2	71963
	NTIC01005557.1	NTIC01005557.1_18X0	84215052	84281828	2	66776
	NTIC01005557.1	NTIC01005557.1_1X0	816967	885596	3	68629
	NTIC01005557.1	NTIC01005557.1_2X0	2047421	2082703	2	35282
	NTIC01005557.1	NTIC01005557.1_3X0	2089064	2106697	2	17633
	NTIC01005557.1	NTIC01005557.1_4X0	4037872	4085902	2	48030
	NTIC01005557.1	NTIC01005557.1_5X0	7407631	7545825	3	138194
	NTIC01005557.1	NTIC01005557.1_6X0	15658761	15723476	2	64715
	NTIC01005557.1	NTIC01005557.1_7X0	16141534	16295033	3	153499
	NTIC01005557.1	NTIC01005557.1_8X0	18762369	18912621	5	150252
	NTIC01005557.1	NTIC01005557.1_9X0	19218016	19602096	8	384080
	NTIC01006668.1	NTIC01006668.1_10X0	26954844	27068878	5	114034
	NTIC01006668.1	NTIC01006668.1_11X0	27142949	27153083	2	10134
	NTIC01006668.1	NTIC01006668.1_12X0	27953749	28005857	3	52108
	NTIC01006668.1	NTIC01006668.1_13X0	59566258	59593710	2	27452
	NTIC01006668.1	NTIC01006668.1_14X0	81567925	81651814	2	83889
	NTIC01006668.1	NTIC01006668.1_15X0	92511094	92580271	2	69177
	NTIC01006668.1	NTIC01006668.1_1X0	5685846	5751787	3	65941

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	NTIC01006668.1	NTIC01006668.1_2X0	8260997	8380045	4	119048
	NTIC01006668.1	NTIC01006668.1_3X0	15960265	16013258	3	52993
	NTIC01006668.1	NTIC01006668.1_4X0	19986292	20227255	7	240963
	NTIC01006668.1	NTIC01006668.1_5X0	21353473	21832500	7	479027
	NTIC01006668.1	NTIC01006668.1_6X0	22112337	22144010	3	31673
	NTIC01006668.1	NTIC01006668.1_7X0	25129084	25171232	2	42148
	NTIC01006668.1	NTIC01006668.1_8X0	25438393	25501477	2	63084
	NTIC01006668.1	NTIC01006668.1_9X0	25534751	25552536	3	17785
	NTIC01007779.1	NTIC01007779.1_1X0	44214071	44672475	12	458404
	NTIC01007779.1	NTIC01007779.1_2X0	44696773	44753147	3	56374
	NTIC01007779.1	NTIC01007779.1_3X0	45072839	45108370	2	35531
	NTIC01007779.1	NTIC01007779.1_4X0	46087718	46323021	5	235303
	NTIC01007779.1	NTIC01007779.1_5X0	53392267	53603152	4	210885
	NTIC01008890.1	NTIC01008890.1_10X0	61722034	61875786	3	153752
	NTIC01008890.1	NTIC01008890.1_11X0	69343173	69380894	2	37721
	NTIC01008890.1	NTIC01008890.1_1X0	9024319	9099126	3	74807
	NTIC01008890.1	NTIC01008890.1_2X0	18550262	18614718	2	64456
	NTIC01008890.1	NTIC01008890.1_3X0	25518602	25554827	2	36225
	NTIC01008890.1	NTIC01008890.1_4X0	27940363	28009285	2	68922
	NTIC01008890.1	NTIC01008890.1_5X0	29320852	29347128	2	26276
	NTIC01008890.1	NTIC01008890.1_6X0	36383972	36459706	2	75734
	NTIC01008890.1	NTIC01008890.1_7X0	36836328	36863275	2	26947
	NTIC01008890.1	NTIC01008890.1_8X0	41720183	41802325	2	82142
	NTIC01008890.1	NTIC01008890.1_9X0	44087065	44104965	2	17900
	NTIC01010001.1	NTIC01010001.1_10X0	32498301	32974384	12	476083
	NTIC01010001.1	NTIC01010001.1_11X0	33057543	33286363	4	228820
	NTIC01010001.1	NTIC01010001.1_12X0	33948527	34016456	5	67929
	NTIC01010001.1	NTIC01010001.1_1X0	1348922	1415528	2	66606
	NTIC01010001.1	NTIC01010001.1_2X0	11668683	11669852	2	1169
	NTIC01010001.1	NTIC01010001.1_3X0	23164849	23223215	2	58366
	NTIC01010001.1	NTIC01010001.1_4X0	25007313	25058246	2	50933
	NTIC01010001.1	NTIC01010001.1_5X0	25173609	25332995	3	159386
	NTIC01010001.1	NTIC01010001.1_6X0	28917523	29020712	2	103189
	NTIC01010001.1	NTIC01010001.1_7X0	29068819	29205434	4	136615
	NTIC01010001.1	NTIC01010001.1_8X0	29207349	29278002	3	70653
	NTIC01010001.1	NTIC01010001.1_9X0	32351413	32366055	2	14642
	NTIC01011112.1	NTIC01011112.1_10X0	187324918	187471099	4	146181
	NTIC01011112.1	NTIC01011112.1_11X0	187848393	187901677	3	53284
	NTIC01011112.1	NTIC01011112.1_12X0	190038263	190071280	2	33017
	NTIC01011112.1	NTIC01011112.1_13X0	191320662	191375835	2	55173
	NTIC01011112.1	NTIC01011112.1_14X0	198534198	199209989	5	675791
	NTIC01011112.1	NTIC01011112.1_1X0	50293202	50364137	2	70935
	NTIC01011112.1	NTIC01011112.1_2X0	130097976	130128509	2	30533
	NTIC01011112.1	NTIC01011112.1_3X0	130621079	130827245	4	206166
	NTIC01011112.1	NTIC01011112.1_4X0	151647818	151725766	2	77948
	NTIC01011112.1	NTIC01011112.1_5X0	161323662	161415168	2	91506
	NTIC01011112.1	NTIC01011112.1_6X0	161779020	162414636	13	635616
	NTIC01011112.1	NTIC01011112.1_7X0	162638691	162664541	2	25850
	NTIC01011112.1	NTIC01011112.1_8X0	184972204	185097862	4	125658
	NTIC01011112.1	NTIC01011112.1_9X0	187184000	187245414	2	61414
	NTIC01011113.1	NTIC01011113.1_10X0	6878457	6936874	3	58417
	NTIC01011113.1	NTIC01011113.1_11X0	7394611	7471547	4	76936
	NTIC01011113.1	NTIC01011113.1_12X0	26981677	27174245	14	192568
	NTIC01011113.1	NTIC01011113.1_12X1	26641147	26724333	6	83186
	NTIC01011113.1	NTIC01011113.1_12X10	9085076	9240355	5	155279

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	NTIC01011113.1	NTIC01011113.1_12X11	35256629	35361700	5	105071
	NTIC01011113.1	NTIC01011113.1_12X12	36253781	36398586	5	144805
	NTIC01011113.1	NTIC01011113.1_12X13	8667151	8844589	5	177438
	NTIC01011113.1	NTIC01011113.1_12X14	9290836	9452554	5	161718
	NTIC01011113.1	NTIC01011113.1_12X15	21221720	21325026	5	103306
	NTIC01011113.1	NTIC01011113.1_12X16	8871138	8982357	4	111219
	NTIC01011113.1	NTIC01011113.1_12X17	23269032	23332658	5	63626
	NTIC01011113.1	NTIC01011113.1_12X18	23033660	23167054	4	133394
	NTIC01011113.1	NTIC01011113.1_12X19	24198831	24282310	4	83479
	NTIC01011113.1	NTIC01011113.1_12X2	23832022	23970516	7	138494
	NTIC01011113.1	NTIC01011113.1_12X20	22618946	22695381	3	76435
	NTIC01011113.1	NTIC01011113.1_12X21	36923876	37109271	5	185395
	NTIC01011113.1	NTIC01011113.1_12X22	22771657	22844013	3	72356
	NTIC01011113.1	NTIC01011113.1_12X23	35140231	35190844	4	50613
	NTIC01011113.1	NTIC01011113.1_12X24	26317797	26413640	4	95843
	NTIC01011113.1	NTIC01011113.1_12X25	35693508	35775249	4	81741
	NTIC01011113.1	NTIC01011113.1_12X26	24392752	24554233	5	161481
	NTIC01011113.1	NTIC01011113.1_12X27	36635578	36744042	4	108464
	NTIC01011113.1	NTIC01011113.1_12X28	34926985	35025814	4	98829
	NTIC01011113.1	NTIC01011113.1_12X29	36133962	36232313	3	98351
	NTIC01011113.1	NTIC01011113.1_12X3	21766829	21888335	6	121506
	NTIC01011113.1	NTIC01011113.1_12X30	36778358	36892451	3	114093
	NTIC01011113.1	NTIC01011113.1_12X31	20901503	21013505	3	112002
	NTIC01011113.1	NTIC01011113.1_12X32	21059709	21173799	3	114090
	NTIC01011113.1	NTIC01011113.1_12X33	16579721	16685306	3	105585
	NTIC01011113.1	NTIC01011113.1_12X34	36469599	36568627	3	99028
	NTIC01011113.1	NTIC01011113.1_12X35	26156471	26198271	3	41800
	NTIC01011113.1	NTIC01011113.1_12X36	22462194	22573959	3	111765
	NTIC01011113.1	NTIC01011113.1_12X37	16449828	16509597	2	59769
	NTIC01011113.1	NTIC01011113.1_12X38	22900742	23016017	3	115275
	NTIC01011113.1	NTIC01011113.1_12X39	25136304	25212615	3	76311
	NTIC01011113.1	NTIC01011113.1_12X4	21399085	21541517	6	142432
	NTIC01011113.1	NTIC01011113.1_12X40	26516222	26610618	3	94396
	NTIC01011113.1	NTIC01011113.1_12X41	17248627	17338222	3	89595
	NTIC01011113.1	NTIC01011113.1_12X42	19641069	19700987	2	59918
	NTIC01011113.1	NTIC01011113.1_12X43	16160857	16205971	2	45114
	NTIC01011113.1	NTIC01011113.1_12X44	11432182	11459354	2	27172
	NTIC01011113.1	NTIC01011113.1_12X45	17808126	17857865	2	49739
	NTIC01011113.1	NTIC01011113.1_12X46	21952426	22016077	2	63651
	NTIC01011113.1	NTIC01011113.1_12X47	11901184	11953112	2	51928
	NTIC01011113.1	NTIC01011113.1_12X48	25620631	25689870	2	69239
	NTIC01011113.1	NTIC01011113.1_12X5	21580390	21654485	4	74095
	NTIC01011113.1	NTIC01011113.1_12X6	12311229	12483991	7	172762
	NTIC01011113.1	NTIC01011113.1_12X7	25874899	25996784	6	121885
	NTIC01011113.1	NTIC01011113.1_12X8	21690616	21751699	3	61083
	NTIC01011113.1	NTIC01011113.1_12X9	20703314	20815185	6	111871
	NTIC01011113.1	NTIC01011113.1_13X0	38004019	38039615	3	35596
	NTIC01011113.1	NTIC01011113.1_14X0	38155212	38236979	2	81767
	NTIC01011113.1	NTIC01011113.1_15X0	38708413	38713281	2	4868
	NTIC01011113.1	NTIC01011113.1_16X0	38842743	38913227	3	70484
	NTIC01011113.1	NTIC01011113.1_17X0	39144629	39156879	2	12250
	NTIC01011113.1	NTIC01011113.1_18X0	39207597	39278382	3	70785
	NTIC01011113.1	NTIC01011113.1_19X0	39352022	39419222	2	67200
	NTIC01011113.1	NTIC01011113.1_1X0	28735	107255	2	78520
	NTIC01011113.1	NTIC01011113.1_20X0	39573391	39588954	2	15563

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	NTIC01011113.1	NTIC01011113.1_21X0	39950050	40105596	5	155546
	NTIC01011113.1	NTIC01011113.1_22X0	40359945	40438440	3	78495
	NTIC01011113.1	NTIC01011113.1_23X0	41264929	41406596	4	141667
	NTIC01011113.1	NTIC01011113.1_23X1	41156246	41219880	2	63634
	NTIC01011113.1	NTIC01011113.1_24X0	41469678	41540609	3	70931
	NTIC01011113.1	NTIC01011113.1_25X0	41609076	41747435	4	138359
	NTIC01011113.1	NTIC01011113.1_26X0	41939779	42122524	8	182745
	NTIC01011113.1	NTIC01011113.1_26X1	42130495	42197249	3	66754
	NTIC01011113.1	NTIC01011113.1_26X2	42231656	42363257	4	131601
	NTIC01011113.1	NTIC01011113.1_27X0	42371994	42505908	4	133914
	NTIC01011113.1	NTIC01011113.1_28X0	43731102	43750175	2	19073
	NTIC01011113.1	NTIC01011113.1_29X0	44764449	44799215	2	34766
	NTIC01011113.1	NTIC01011113.1_2X0	747240	802825	3	55585
	NTIC01011113.1	NTIC01011113.1_30X0	46086994	46133707	2	46713
	NTIC01011113.1	NTIC01011113.1_3X0	836442	873310	2	36868
	NTIC01011113.1	NTIC01011113.1_4X0	932383	980589	2	48206
	NTIC01011113.1	NTIC01011113.1_5X0	3254122	3327431	2	73309
	NTIC01011113.1	NTIC01011113.1_6X0	4003854	4036305	2	32451
	NTIC01011113.1	NTIC01011113.1_7X0	4524744	4537189	2	12445
	NTIC01011113.1	NTIC01011113.1_8X0	5177170	5274502	5	97332
	NTIC01011113.1	NTIC01011113.1_9X0	6455723	6547399	4	91676
	NTIC01011113.1	NTIC01011113.1_9X1	6566638	6598101	2	31463
	NTIC01011113.1	NTIC01011113.1_9X2	6665127	6713144	2	48017
	NTIC01012224.1	NTIC01012224.1_1X0	1599594	1643770	3	44176
	NTIC01012224.1	NTIC01012224.1_2X0	2460801	2538360	2	77559
	NTIC01012224.1	NTIC01012224.1_3X0	7527880	7603129	2	75249
	NTIC01012224.1	NTIC01012224.1_4X0	9239044	9308581	2	69537
	NTIC01012224.1	NTIC01012224.1_5X0	10277840	10461897	4	184057
	NTIC01012224.1	NTIC01012224.1_6X0	10637377	10682459	2	45082
	NTIC01012224.1	NTIC01012224.1_7X0	12757624	13433006	11	675382
	NTIC01012224.1	NTIC01012224.1_8X0	13496463	13555099	2	58636
	NTIC01013335.1	NTIC01013335.1_10X0	34162814	34263296	3	100482
	NTIC01013335.1	NTIC01013335.1_11X0	34305386	34636287	14	330901
	NTIC01013335.1	NTIC01013335.1_12X0	34682391	34853468	5	171077
	NTIC01013335.1	NTIC01013335.1_13X0	35201429	35429559	5	228130
	NTIC01013335.1	NTIC01013335.1_14X0	35689548	35731659	2	42111
	NTIC01013335.1	NTIC01013335.1_15X0	35935045	35975851	2	40806
	NTIC01013335.1	NTIC01013335.1_16X0	36154798	36195155	2	40357
	NTIC01013335.1	NTIC01013335.1_17X0	37694674	37770376	3	75702
	NTIC01013335.1	NTIC01013335.1_18X0	37787430	37818843	2	31413
	NTIC01013335.1	NTIC01013335.1_19X0	39411140	39459447	2	48307
	NTIC01013335.1	NTIC01013335.1_1X0	18371	93452	2	75081
	NTIC01013335.1	NTIC01013335.1_20X0	40413595	40571322	3	157727
	NTIC01013335.1	NTIC01013335.1_21X0	42860802	43003145	5	142343
	NTIC01013335.1	NTIC01013335.1_22X0	43827559	43892586	2	65027
	NTIC01013335.1	NTIC01013335.1_23X0	44107319	44130725	2	23406
	NTIC01013335.1	NTIC01013335.1_24X0	44946367	45009131	2	62764
	NTIC01013335.1	NTIC01013335.1_2X0	14618013	14658413	2	40400
	NTIC01013335.1	NTIC01013335.1_3X0	20487955	20600210	3	112255
	NTIC01013335.1	NTIC01013335.1_4X0	21831419	22068611	4	237192
	NTIC01013335.1	NTIC01013335.1_5X0	29792373	29859255	2	66882
	NTIC01013335.1	NTIC01013335.1_6X0	31655353	31764253	3	108900
	NTIC01013335.1	NTIC01013335.1_7X0	32269087	32776597	11	507510
	NTIC01013335.1	NTIC01013335.1_8X0	33175338	33269675	2	94337
	NTIC01013335.1	NTIC01013335.1_9X0	33710110	33772376	2	62266

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	NTIC01014446.1	NTIC01014446.1_10X0	31015012	31067677	3	52665
	NTIC01014446.1	NTIC01014446.1_11X0	31735493	31865255	6	129762
	NTIC01014446.1	NTIC01014446.1_12X0	32475873	32668207	3	192334
	NTIC01014446.1	NTIC01014446.1_1X0	2121591	2188382	3	66791
	NTIC01014446.1	NTIC01014446.1_2X0	13471171	13541342	2	70171
	NTIC01014446.1	NTIC01014446.1_3X0	13787572	13807986	2	20414
	NTIC01014446.1	NTIC01014446.1_4X0	15929478	15961708	2	32230
	NTIC01014446.1	NTIC01014446.1_5X0	20844557	20861077	2	16520
	NTIC01014446.1	NTIC01014446.1_6X0	21650445	21700978	2	50533
	NTIC01014446.1	NTIC01014446.1_7X0	21860015	21932630	3	72615
	NTIC01014446.1	NTIC01014446.1_8X0	22168643	22288705	4	120062
	NTIC01014446.1	NTIC01014446.1_9X0	30742086	30867600	3	125514
	NTIC01015557.1	NTIC01015557.1_1X0	17552144	17555312	2	3168
	NTIC01015557.1	NTIC01015557.1_2X0	20358910	20407523	2	48613
	NTIC01015557.1	NTIC01015557.1_3X0	22054900	22093540	2	38640
	NTIC01016668.1	NTIC01016668.1_1X0	87255	112962	3	25707
	NTIC01016668.1	NTIC01016668.1_3X0	1422060	1469264	4	47204
	NTIC01016668.1	NTIC01016668.1_4X0	1755638	1788768	2	33130
	NTIC01016668.1	NTIC01016668.1_5X0	1935510	1971631	2	36121
	NTIC01018890.1	NTIC01018890.1_1X0	655309	776510	3	121201
	NTIC01021112.1	NTIC01021112.1_1X0	581355	621442	2	40087
	NTIC01022223.1	NTIC01022223.1_10X0	117079125	117219029	3	139904
	NTIC01022223.1	NTIC01022223.1_11X0	120581206	120617683	2	36477
	NTIC01022223.1	NTIC01022223.1_12X0	124456862	124485553	2	28691
	NTIC01022223.1	NTIC01022223.1_13X0	125337276	125513082	6	175806
	NTIC01022223.1	NTIC01022223.1_14X0	127407393	127553123	3	145730
	NTIC01022223.1	NTIC01022223.1_15X0	130301677	131006094	4	704417
	NTIC01022223.1	NTIC01022223.1_16X0	131045445	131057149	2	11704
	NTIC01022223.1	NTIC01022223.1_17X0	133646300	134045855	13	399555
	NTIC01022223.1	NTIC01022223.1_18X0	134230962	134247067	2	16105
	NTIC01022223.1	NTIC01022223.1_19X0	135226106	135244771	2	18665
	NTIC01022223.1	NTIC01022223.1_1X0	37564071	37643285	4	79214
	NTIC01022223.1	NTIC01022223.1_20X0	135451597	135505947	4	54350
	NTIC01022223.1	NTIC01022223.1_21X0	135591600	135601498	3	9898
	NTIC01022223.1	NTIC01022223.1_22X0	135674905	136344355	6	669450
	NTIC01022223.1	NTIC01022223.1_23X0	136813316	137134911	5	321595
	NTIC01022223.1	NTIC01022223.1_24X0	138118054	138305416	5	187362
	NTIC01022223.1	NTIC01022223.1_25X0	143186110	143314295	3	128185
	NTIC01022223.1	NTIC01022223.1_26X0	144369451	144433613	2	64162
	NTIC01022223.1	NTIC01022223.1_27X0	166712835	167363760	32	650925
	NTIC01022223.1	NTIC01022223.1_28X0	167745609	167763085	4	17476
	NTIC01022223.1	NTIC01022223.1_29X0	167945267	168003922	2	58655
	NTIC01022223.1	NTIC01022223.1_2X0	53739445	53827957	3	88512
	NTIC01022223.1	NTIC01022223.1_30X0	168450092	168530330	13	80238
	NTIC01022223.1	NTIC01022223.1_31X0	168544133	168635830	5	91697
	NTIC01022223.1	NTIC01022223.1_32X0	168647419	169088351	14	440932
	NTIC01022223.1	NTIC01022223.1_33X0	169657466	169693314	2	35848
	NTIC01022223.1	NTIC01022223.1_3X0	54008888	54026283	2	17395
	NTIC01022223.1	NTIC01022223.1_4X0	63215799	63216621	2	822
	NTIC01022223.1	NTIC01022223.1_5X0	65161454	65227854	2	66400
	NTIC01022223.1	NTIC01022223.1_6X0	97106123	97196966	5	90843
	NTIC01022223.1	NTIC01022223.1_7X0	114030199	114153265	4	123066
	NTIC01022223.1	NTIC01022223.1_8X0	114642016	114696274	2	54258
	NTIC01022223.1	NTIC01022223.1_9X0	116932859	117003983	4	71124
	NTIC01024446.1	NTIC01024446.1_1X0	178454	197865	5	19411

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	NTIC01025557.1	NTIC01025557.1_1X0	75802	139815	5	64013
	NTIC01025557.1	NTIC01025557.1_2X0	312363	376084	4	63721
	NTIC01025557.1	NTIC01025557.1_3X0	449930	511763	2	61833
	NTIC01025557.1	NTIC01025557.1_4X0	600981	618844	2	17863
	NTIC01031111.1	NTIC01031111.1_1X0	186079	213353	2	27274
	NTIC01033278.1	NTIC01033278.1_10X0	97758891	97786330	2	27439
	NTIC01033278.1	NTIC01033278.1_11X0	97818972	97861875	2	42903
	NTIC01033278.1	NTIC01033278.1_12X0	98652825	99278359	5	625534
	NTIC01033278.1	NTIC01033278.1_13X0	114424229	114485416	2	61187
	NTIC01033278.1	NTIC01033278.1_14X0	124377643	124409019	2	31376
	NTIC01033278.1	NTIC01033278.1_15X0	147268139	147296396	2	28257
	NTIC01033278.1	NTIC01033278.1_16X0	156271390	156307305	2	35915
	NTIC01033278.1	NTIC01033278.1_1X0	1597430	1720203	3	122773
	NTIC01033278.1	NTIC01033278.1_2X0	3980253	4006610	2	26357
	NTIC01033278.1	NTIC01033278.1_3X0	13326504	13381335	2	54831
	NTIC01033278.1	NTIC01033278.1_4X0	80450674	80545491	2	94817
	NTIC01033278.1	NTIC01033278.1_5X0	85411732	85488737	2	77005
	NTIC01033278.1	NTIC01033278.1_6X0	91909253	91943009	3	33756
	NTIC01033278.1	NTIC01033278.1_7X0	93773656	94068899	9	295243
	NTIC01033278.1	NTIC01033278.1_8X0	94307698	94416376	3	108678
	NTIC01033278.1	NTIC01033278.1_9X0	97403840	97469500	2	65660
	NTIC01033390.1	NTIC01033390.1_1X0	116912	136254	2	19342
	NTIC01034389.1	NTIC01034389.1_10X0	38547640	38569328	2	21688
	NTIC01034389.1	NTIC01034389.1_11X0	56116496	56133655	2	17159
	NTIC01034389.1	NTIC01034389.1_12X0	56651998	56664292	2	12294
	NTIC01034389.1	NTIC01034389.1_13X0	61586975	61620280	2	33305
	NTIC01034389.1	NTIC01034389.1_14X0	62179071	62252494	2	73423
	NTIC01034389.1	NTIC01034389.1_15X0	65532798	65545836	2	13038
	NTIC01034389.1	NTIC01034389.1_16X0	66302750	66330066	2	27316
	NTIC01034389.1	NTIC01034389.1_17X0	67076596	67161732	3	85136
	NTIC01034389.1	NTIC01034389.1_18X0	67714459	67721434	2	6975
	NTIC01034389.1	NTIC01034389.1_19X0	67755513	67867817	3	112304
	NTIC01034389.1	NTIC01034389.1_1X0	3831903	4086894	5	254991
	NTIC01034389.1	NTIC01034389.1_20X0	68094208	68099864	2	5656
	NTIC01034389.1	NTIC01034389.1_21X0	68238743	68250767	2	12024
	NTIC01034389.1	NTIC01034389.1_22X0	68750342	68796618	2	46276
	NTIC01034389.1	NTIC01034389.1_23X0	70930405	71265746	6	335341
	NTIC01034389.1	NTIC01034389.1_24X0	79422120	79444732	2	22612
	NTIC01034389.1	NTIC01034389.1_25X0	82658504	82755716	2	97212
	NTIC01034389.1	NTIC01034389.1_26X0	82901536	82931957	5	30421
	NTIC01034389.1	NTIC01034389.1_27X0	84909369	84988806	2	79437
	NTIC01034389.1	NTIC01034389.1_28X0	85058617	85085785	3	27168
	NTIC01034389.1	NTIC01034389.1_29X0	85142208	85226393	2	84185
	NTIC01034389.1	NTIC01034389.1_2X0	5563054	5836001	11	272947
	NTIC01034389.1	NTIC01034389.1_30X0	85376459	85644375	13	267916
	NTIC01034389.1	NTIC01034389.1_31X0	85665475	85907250	14	241775
	NTIC01034389.1	NTIC01034389.1_32X0	85924466	85939421	3	14955
	NTIC01034389.1	NTIC01034389.1_33X0	85949066	86279608	12	330542
	NTIC01034389.1	NTIC01034389.1_34X0	88942184	89029444	2	87260
	NTIC01034389.1	NTIC01034389.1_35X0	91756449	91963256	6	206807
	NTIC01034389.1	NTIC01034389.1_36X0	92412574	92509699	3	97125
	NTIC01034389.1	NTIC01034389.1_37X0	93379460	93477768	5	98308
	NTIC01034389.1	NTIC01034389.1_38X0	100198721	100217464	2	18743
	NTIC01034389.1	NTIC01034389.1_39X0	109282761	109317706	2	34945
	NTIC01034389.1	NTIC01034389.1_3X0	7047340	7149107	3	101767

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	NTIC01034389.1	NTIC01034389.1_40X0	109353927	109422393	2	68466
	NTIC01034389.1	NTIC01034389.1_41X0	109659125	109696362	2	37237
	NTIC01034389.1	NTIC01034389.1_42X0	109801447	109824217	2	22770
	NTIC01034389.1	NTIC01034389.1_43X0	111033053	111222260	10	189207
	NTIC01034389.1	NTIC01034389.1_44X0	111850778	111926636	2	75858
	NTIC01034389.1	NTIC01034389.1_45X0	114452581	114481588	3	29007
	NTIC01034389.1	NTIC01034389.1_46X0	119479145	119521606	2	42461
	NTIC01034389.1	NTIC01034389.1_47X0	123790246	123955064	5	164818
	NTIC01034389.1	NTIC01034389.1_48X0	129206454	129250307	2	43853
	NTIC01034389.1	NTIC01034389.1_49X0	139295130	139364106	2	68976
	NTIC01034389.1	NTIC01034389.1_4X0	8431041	8446065	2	15024
	NTIC01034389.1	NTIC01034389.1_50X0	142331382	142494111	3	162729
	NTIC01034389.1	NTIC01034389.1_51X0	149502076	149578831	3	76755
	NTIC01034389.1	NTIC01034389.1_52X0	150952568	150988376	2	35808
	NTIC01034389.1	NTIC01034389.1_53X0	153730769	153790397	3	59628
	NTIC01034389.1	NTIC01034389.1_54X0	156764654	156827832	2	63178
	NTIC01034389.1	NTIC01034389.1_55X0	158391881	158416569	2	24688
	NTIC01034389.1	NTIC01034389.1_5X0	9986919	10013959	2	27040
	NTIC01034389.1	NTIC01034389.1_6X0	28166603	28445138	7	278535
	NTIC01034389.1	NTIC01034389.1_7X0	31245967	31322888	2	76921
	NTIC01034389.1	NTIC01034389.1_8X0	37921226	38072107	6	150881
	NTIC01034389.1	NTIC01034389.1_9X0	38312682	38439211	3	126529
	NTIC01035500.1	NTIC01035500.1_10X0	133953084	133972815	2	19731
	NTIC01035500.1	NTIC01035500.1_11X0	134197106	134272435	2	75329
	NTIC01035500.1	NTIC01035500.1_12X0	146779831	146801085	2	21254
	NTIC01035500.1	NTIC01035500.1_13X0	146869506	146889925	2	20419
	NTIC01035500.1	NTIC01035500.1_1X0	241776	288181	2	46405
	NTIC01035500.1	NTIC01035500.1_2X0	16254859	16308832	2	53973
	NTIC01035500.1	NTIC01035500.1_3X0	17113950	17142892	2	28942
	NTIC01035500.1	NTIC01035500.1_4X0	47736296	47843920	10	107624
	NTIC01035500.1	NTIC01035500.1_5X0	72899847	72933157	2	33310
	NTIC01035500.1	NTIC01035500.1_6X0	115454719	115606775	3	152056
	NTIC01035500.1	NTIC01035500.1_7X0	131445538	131478924	2	33386
	NTIC01035500.1	NTIC01035500.1_8X0	132846589	132969406	3	122817
	NTIC01035500.1	NTIC01035500.1_9X0	133559249	133594273	2	35024
	NTIC01036611.1	NTIC01036611.1_10X0	80933111	81037658	3	104547
	NTIC01036611.1	NTIC01036611.1_11X0	86651523	86688740	4	37217
	NTIC01036611.1	NTIC01036611.1_12X0	89034090	89131202	2	97112
	NTIC01036611.1	NTIC01036611.1_13X0	97107761	97173775	2	66014
	NTIC01036611.1	NTIC01036611.1_14X0	98549116	98576690	3	27574
	NTIC01036611.1	NTIC01036611.1_15X0	99192362	99202996	2	10634
	NTIC01036611.1	NTIC01036611.1_16X0	100334247	100393063	2	58816
	NTIC01036611.1	NTIC01036611.1_17X0	101737144	101773808	2	36664
	NTIC01036611.1	NTIC01036611.1_18X0	104582338	104605643	2	23305
	NTIC01036611.1	NTIC01036611.1_19X0	105149838	105183144	2	33306
	NTIC01036611.1	NTIC01036611.1_1X0	9443161	9635750	4	192589
	NTIC01036611.1	NTIC01036611.1_20X0	107301335	107358156	2	56821
	NTIC01036611.1	NTIC01036611.1_21X0	108163497	108176253	2	12756
	NTIC01036611.1	NTIC01036611.1_22X0	108356399	108380689	3	24290
	NTIC01036611.1	NTIC01036611.1_23X0	111011736	111076840	2	65104
	NTIC01036611.1	NTIC01036611.1_24X0	113424028	113451262	2	27234
	NTIC01036611.1	NTIC01036611.1_25X0	129197133	129248595	2	51462
	NTIC01036611.1	NTIC01036611.1_26X0	147502386	147736192	6	233806
	NTIC01036611.1	NTIC01036611.1_27X0	147927320	147953724	2	26404
	NTIC01036611.1	NTIC01036611.1_2X0	11151852	11177358	2	25506

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	NTIC01036611.1	NTIC01036611.1_3X0	11219549	11281089	2	61540
	NTIC01036611.1	NTIC01036611.1_4X0	12422742	12437358	2	14616
	NTIC01036611.1	NTIC01036611.1_5X0	31808480	32195950	7	387470
	NTIC01036611.1	NTIC01036611.1_6X0	52540872	52548700	2	7828
	NTIC01036611.1	NTIC01036611.1_7X0	73970936	73991820	2	20884
	NTIC01036611.1	NTIC01036611.1_8X0	77267639	77301238	2	33599
	NTIC01036611.1	NTIC01036611.1_9X0	79310352	79321980	2	11628
	NTIC01038832.1	NTIC01038832.1_1X0	3149931	3200645	2	50714
	NTIC01038832.1	NTIC01038832.1_2X0	3203944	3294767	2	90823
	NTIC01038832.1	NTIC01038832.1_3X0	10191452	10202527	2	11075
	NTIC01038832.1	NTIC01038832.1_4X0	16331525	16395815	2	64290
	NTIC01038833.1	NTIC01038833.1_10X0	71654567	71682210	2	27643
	NTIC01038833.1	NTIC01038833.1_11X0	89841538	89872638	4	31100
	NTIC01038833.1	NTIC01038833.1_12X0	93806508	93833614	2	27106
	NTIC01038833.1	NTIC01038833.1_13X0	104789506	104861130	2	71624
	NTIC01038833.1	NTIC01038833.1_14X0	107214593	107271429	2	56836
	NTIC01038833.1	NTIC01038833.1_1X0	11604883	11673653	2	68770
	NTIC01038833.1	NTIC01038833.1_2X0	26389686	26669751	5	280065
	NTIC01038833.1	NTIC01038833.1_3X0	26882640	26967927	3	85287
	NTIC01038833.1	NTIC01038833.1_4X0	27693782	27771006	4	77224
	NTIC01038833.1	NTIC01038833.1_5X0	55102749	55140550	2	37801
	NTIC01038833.1	NTIC01038833.1_6X0	66625776	66833908	4	208132
	NTIC01038833.1	NTIC01038833.1_7X0	68895709	69033399	3	137690
	NTIC01038833.1	NTIC01038833.1_8X0	69596866	69615901	2	19035
	NTIC01038833.1	NTIC01038833.1_9X0	71261805	71548173	6	286368
	NTIC01038834.1	NTIC01038834.1_10X0	54275966	54353057	2	77091
	NTIC01038834.1	NTIC01038834.1_11X0	67049861	67138787	2	88926
	NTIC01038834.1	NTIC01038834.1_12X0	79095161	79173049	4	77888
	NTIC01038834.1	NTIC01038834.1_13X0	79492072	79502717	2	10645
	NTIC01038834.1	NTIC01038834.1_14X0	80483393	80499039	2	15646
	NTIC01038834.1	NTIC01038834.1_15X0	81718764	81781014	3	62250
	NTIC01038834.1	NTIC01038834.1_16X0	82126989	82182966	2	55977
	NTIC01038834.1	NTIC01038834.1_17X0	82393398	82609838	8	216440
	NTIC01038834.1	NTIC01038834.1_18X0	82859439	83551662	14	692223
	NTIC01038834.1	NTIC01038834.1_19X0	92517207	92589236	2	72029
	NTIC01038834.1	NTIC01038834.1_1X0	1076393	1111724	2	35331
	NTIC01038834.1	NTIC01038834.1_20X0	104714616	104787812	4	73196
	NTIC01038834.1	NTIC01038834.1_21X0	106342100	106444308	3	102208
	NTIC01038834.1	NTIC01038834.1_22X0	110652497	110730700	3	78203
	NTIC01038834.1	NTIC01038834.1_23X0	130568579	130602493	2	33914
	NTIC01038834.1	NTIC01038834.1_24X0	130684933	130864490	3	179557
	NTIC01038834.1	NTIC01038834.1_2X0	3816853	3855635	2	38782
	NTIC01038834.1	NTIC01038834.1_3X0	7929534	7965243	2	35709
	NTIC01038834.1	NTIC01038834.1_4X0	8491758	8517202	2	25444
	NTIC01038834.1	NTIC01038834.1_5X0	9534634	9892253	11	357619
	NTIC01038834.1	NTIC01038834.1_6X0	30367850	30388980	2	21130
	NTIC01038834.1	NTIC01038834.1_7X0	36134002	36201008	2	67006
	NTIC01038834.1	NTIC01038834.1_8X0	39593056	39636475	2	43419
	NTIC01038834.1	NTIC01038834.1_9X0	51739878	51811592	2	71714
	1	1_10X0	70049818	70176078	2	126260
	1	1_11X0	72099571	72168267	2	68696
	1	1_12X0	78438091	78657310	2	219219
	1	1_13X0	89677533	89881004	3	203471
	1	1_14X0	93471149	93688867	4	217718
	1	1_15X0	98831898	98959173	4	127275

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_16X0	100345157	100866365	18	521208
	1	1_16X1	101108258	101563617	15	455359
	1	1_16X10	115702047	115971886	6	269839
	1	1_16X11	111486389	111834729	8	348340
	1	1_16X12	99893640	100284657	6	391017
	1	1_16X13	117436239	117734060	7	297821
	1	1_16X14	110542524	110890184	7	347660
	1	1_16X15	115992633	116197901	5	205268
	1	1_16X16	101582593	101877099	6	294506
	1	1_16X17	116658081	116814413	7	156332
	1	1_16X18	103467177	103724523	7	257346
	1	1_16X19	112734986	112992652	7	257666
	1	1_16X2	105567193	106028331	16	461138
	1	1_16X20	115323206	115657149	5	333943
	1	1_16X21	113905414	114324703	8	419289
	1	1_16X22	115007134	115306304	5	299170
	1	1_16X23	104520326	104914837	6	394511
	1	1_16X24	116302327	116589575	4	287248
	1	1_16X25	102362733	102960845	7	598112
	1	1_16X26	103910451	104206379	4	295928
	1	1_16X27	113262460	113652076	5	389616
	1	1_16X28	108644945	108719964	4	75019
	1	1_16X29	108161285	108221577	3	60292
	1	1_16X3	102972294	103403633	12	431339
	1	1_16X30	107162476	107256782	3	94306
	1	1_16X31	111902537	112271208	3	368671
	1	1_16X32	106520300	106561318	2	41018
	1	1_16X33	107614264	107644255	2	29991
	1	1_16X34	99130539	99454847	5	324308
	1	1_16X4	111181194	111443670	6	262476
	1	1_16X5	101896556	102228796	10	332240
	1	1_16X6	110916106	111170054	6	253948
	1	1_16X7	105119927	105555559	12	435632
	1	1_16X8	99495014	99875655	10	380641
	1	1_16X9	117054817	117423691	8	368874
	1	1_17X0	118778224	118849659	2	71435
	1	1_1X0	4876977	4931752	2	54775
	1	1_2X0	13213440	13638925	9	425485
	1	1_3X0	24388280	24431456	2	43176
	1	1_4X0	29506545	29669676	2	163131
	1	1_5X0	40602291	40636942	2	34651
	1	1_6X0	49269241	49495146	3	225905
	1	1_7X0	54932810	54977116	2	44306
	1	1_8X0	56894631	56978049	3	83418
	1	1_9X0	66270560	66331231	2	60671
	10	10_10X0	25910928	25918970	2	8042
	10	10_11X0	26932351	27015748	2	83397
	10	10_12X0	27370669	27567023	4	196354
	10	10_13X0	28132195	28149721	2	17526
	10	10_14X0	28503873	28539012	2	35139
	10	10_15X0	35494461	35644453	3	149992
	10	10_16X0	40579698	41003665	6	423967
	10	10_17X0	43399980	43469225	2	69245
	10	10_18X0	45138538	45197417	2	58879
	10	10_19X0	46272962	46319695	2	46733

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	10	10_1X0	301089	423604	2	122515
	10	10_20X0	47340619	47403507	2	62888
	10	10_21X0	67766127	67793587	2	27460
	10	10_22X0	69166611	69190289	2	23678
	10	10_2X0	1005900	1134325	3	128425
	10	10_3X0	1546387	1564073	2	17686
	10	10_4X0	1576236	1585087	2	8851
	10	10_5X0	15901193	15976729	3	75536
	10	10_6X0	16037166	16159020	2	121854
	10	10_7X0	21709300	21775013	2	65713
	10	10_8X0	22367367	22489371	2	122004
	10	10_9X0	23956823	23998240	2	41417
	11	11_10X0	50365924	50523747	2	157823
	11	11_11X0	51227595	51334380	2	106785
	11	11_12X0	51334891	51359207	2	24316
	11	11_13X0	52167237	52342202	2	174965
	11	11_14X0	52399132	52504571	7	105439
	11	11_15X0	58085495	58109339	2	23844
	11	11_16X0	60467150	60692527	11	225377
	11	11_17X0	64028093	64033788	2	5695
	11	11_18X0	65692887	66033099	2	340212
	11	11_19X0	67416463	67456917	2	40454
	11	11_1X0	239003	444959	6	205956
	11	11_20X0	67528480	67625798	2	97318
	11	11_21X0	67641411	67691754	3	50343
	11	11_22X0	68338935	68560565	2	221630
	11	11_2X0	509648	957886	6	448238
	11	11_3X0	2516079	2919685	8	403606
	11	11_4X0	3508845	3527501	2	18656
	11	11_5X0	9511237	9599151	2	87914
	11	11_6X0	20598888	20683073	2	84185
	11	11_7X0	40668156	40880272	2	212116
	11	11_8X0	41223261	41264379	2	41118
	11	11_9X0	45179230	45376504	2	197274
	12	12_10X0	8115793	8232093	2	116300
	12	12_11X0	9914891	10055604	6	140713
	12	12_12X0	10788656	10806651	2	17995
	12	12_13X0	11559843	11774631	4	214788
	12	12_14X0	14353487	14385171	2	31684
	12	12_15X0	15017055	15189507	2	172452
	12	12_16X0	15721491	15767598	2	46107
	12	12_17X0	17538987	17685906	2	146919
	12	12_18X0	17791239	17836200	2	44961
	12	12_19X0	18447343	18570222	2	122879
	12	12_1X0	892388	1021690	4	129302
	12	12_20X0	19854129	19908417	2	54288
	12	12_21X0	20320070	20443200	4	123130
	12	12_22X0	35415803	35450007	3	34204
	12	12_23X0	48437341	48537879	2	100538
	12	12_24X0	62542269	62601339	2	59070
	12	12_25X0	67674646	67749620	2	74974
	12	12_26X0	71917210	71937285	2	20075
	12	12_2X0	1049069	1148566	5	99497
	12	12_3X0	1155288	1201102	5	45814
	12	12_4X0	1281472	1300710	3	19238

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	12	12_5X0	1574724	1641321	2	66597
	12	12_6X0	2055181	2607784	16	552603
	12	12_7X0	4897221	4913701	2	16480
	12	12_8X0	4942447	5250404	3	307957
	12	12_9X0	6243549	6258366	2	14817
	13	13_10X0	58276083	58752217	7	476134
	13	13_11X0	58911538	59177515	5	265977
	13	13_12X0	59229514	59343734	3	114220
	13	13_13X0	62152698	62232992	3	80294
	13	13_14X0	62434792	62531005	3	96213
	13	13_15X0	62804225	62934656	3	130431
	13	13_1X0	2068832	2126209	2	57377
	13	13_2X0	21377316	21634830	2	257514
	13	13_3X0	36773699	37009111	9	235412
	13	13_4X0	37019086	37289442	4	270356
	13	13_5X0	37317105	37945942	4	628837
	13	13_6X0	38061904	38205547	6	143643
	13	13_7X0	38312372	38355436	3	43064
	13	13_8X0	43910590	44015356	2	104766
	13	13_9X0	44608274	44653460	2	45186
	14	14_10X0	19419136	19431295	2	12159
	14	14_11X0	20565886	20681700	3	115814
	14	14_12X0	22066067	22085230	2	19163
	14	14_13X0	36971420	37183274	2	211854
	14	14_14X0	40267026	40422046	12	155020
	14	14_15X0	43365233	43756571	3	391338
	14	14_16X0	46657635	46881015	2	223380
	14	14_17X0	55437622	55492691	2	55069
	14	14_1X0	467363	649378	6	182015
	14	14_2X0	862354	959328	2	96974
	14	14_3X0	1993715	2849348	19	855633
	14	14_3X1	1184208	1626812	8	442604
	14	14_4X0	6562815	6665649	3	102834
	14	14_5X0	7516959	7711366	2	194407
	14	14_6X0	7907075	7958861	2	51786
	14	14_7X0	11703906	11770158	3	66252
	14	14_8X0	13542668	13743053	2	200385
	14	14_9X0	16074785	16139843	2	65058
	15	15_10X0	16798320	16977405	2	179085
	15	15_11X0	17225363	17625150	12	399787
	15	15_12X0	17842734	18123719	9	280985
	15	15_13X0	18641193	18646850	2	5657
	15	15_14X0	23039070	23051112	2	12042
	15	15_15X0	31670414	31877417	4	207003
	15	15_16X0	49196161	49229170	2	33009
	15	15_17X0	52220730	52270169	3	49439
	15	15_18X0	53174157	53313711	2	139554
	15	15_19X0	55781571	56019541	2	237970
	15	15_1X0	235912	300786	4	64874
	15	15_20X0	59580561	59608218	2	27657
	15	15_2X0	1086549	1096519	2	9970
	15	15_3X0	3156394	3302289	2	145895
	15	15_4X0	4547920	4590989	2	43069
	15	15_5X0	6145898	6356719	3	210821
	15	15_6X0	6703533	7006185	3	302652

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	15	15_7X0	7168197	7197784	4	29587
	15	15_8X0	13510165	13694493	6	184328
	15	15_9X0	13994580	14040870	2	46290
	16	16_10X0	14919165	14932939	2	13774
	16	16_11X0	15082147	15266417	2	184270
	16	16_12X0	15841169	15951359	2	110190
	16	16_13X0	23005420	23040012	2	34592
	16	16_14X0	25316588	25407289	2	90701
	16	16_15X0	25500368	26551122	7	1050754
	16	16_16X0	27445601	27583484	2	137883
	16	16_17X0	42953884	43000036	2	46152
	16	16_18X0	44466300	44516482	2	50182
	16	16_19X0	47010183	47013037	2	2854
	16	16_1X0	1021225	1040923	2	19698
	16	16_20X0	53978452	54028773	2	50321
	16	16_21X0	54043532	54090367	4	46835
	16	16_22X0	58871018	59033519	6	162501
	16	16_2X0	5692102	6568564	21	876462
	16	16_3X0	6650574	6705704	2	55130
	16	16_4X0	6754820	7436097	28	681277
	16	16_5X0	9263325	9325940	2	62615
	16	16_6X0	13598123	13642331	2	44208
	16	16_7X0	13743404	13753963	2	10559
	16	16_8X0	14141642	14712580	10	570938
	16	16_9X0	14752828	14901510	8	148682
	17	17_10X0	36970573	37250891	8	280318
	17	17_11X0	37499225	37829005	12	329780
	17	17_12X0	38345758	38400360	2	54602
	17	17_13X0	39326012	39426229	2	100217
	17	17_14X0	39489119	39499535	2	10416
	17	17_15X0	44826667	44861903	2	35236
	17	17_16X0	48626020	48643917	2	17897
	17	17_17X0	53973029	54185647	2	212618
	17	17_18X0	54341484	54558020	3	216536
	17	17_19X0	58289327	58425391	2	136064
	17	17_1X0	773790	875117	2	101327
	17	17_20X0	58497412	58536691	2	39279
	17	17_21X0	59119612	59176095	6	56483
	17	17_22X0	59893335	59936715	2	43380
	17	17_23X0	60104958	60134605	2	29647
	17	17_24X0	60353864	60495279	2	141415
	17	17_25X0	60853493	60924770	2	71277
	17	17_26X0	61094689	61136824	2	42135
	17	17_27X0	61544781	61805810	2	261029
	17	17_2X0	21332621	21537241	2	204620
	17	17_3X0	21567211	21687062	2	119851
	17	17_4X0	29496128	29688217	2	192089
	17	17_5X0	30645843	30816120	2	170277
	17	17_6X0	31202968	31439118	2	236150
	17	17_7X0	34853502	34919586	4	66084
	17	17_8X0	34977298	35065259	2	87961
	17	17_9X0	36597936	36687197	2	89261
	18	18_10X0	38654390	38744500	2	90110
	18	18_11X0	39168192	39607517	30	439325
	18	18_11X1	39711186	40258633	29	547447

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	18	18_11X2	41249258	41571773	10	322515
	18	18_11X3	40499425	40885975	12	386550
	18	18_11X4	41020077	41242043	10	221966
	18	18_11X5	38878947	39140355	2	261408
	18	18_12X0	45197226	45244801	2	47575
	18	18_13X0	46077958	46148399	2	70441
	18	18_14X0	46296960	46324933	2	27973
	18	18_15X0	46422530	46473960	2	51430
	18	18_16X0	47135617	47213754	2	78137
	18	18_17X0	48383274	48468782	3	85508
	18	18_18X0	49802351	49859927	2	57576
	18	18_19X0	49905161	49953881	2	48720
	18	18_1X0	1464409	1618925	2	154516
	18	18_20X0	49960671	50017854	2	57183
	18	18_21X0	50034554	50167485	2	132931
	18	18_22X0	50638097	50800845	5	162748
	18	18_23X0	51594274	51681152	2	86878
	18	18_24X0	51925010	51981842	2	56832
	18	18_25X0	52542498	52584443	2	41945
	18	18_26X0	53180772	53236479	2	55707
	18	18_27X0	53382682	53463130	4	80448
	18	18_28X0	53503572	53744235	4	240663
	18	18_29X0	53855282	54037163	2	181881
	18	18_2X0	11353904	11730184	7	376280
	18	18_30X0	54188081	54338105	4	150024
	18	18_31X0	54533388	54670815	3	137427
	18	18_32X0	55131441	55166311	2	34870
	18	18_33X0	55217817	55309983	2	92166
	18	18_34X0	55489013	55814753	4	325740
	18	18_3X0	12517157	12700157	2	183000
	18	18_4X0	12818100	12941647	2	123547
	18	18_5X0	17020338	17225582	2	205244
	18	18_6X0	25434478	25662393	5	227915
	18	18_7X0	37037184	37514501	18	477317
	18	18_8X0	37584024	37620583	2	36559
	18	18_9X0	37867183	38285155	17	417972
	19	19_1X0	29990064	30241064	2	251000
	19	19_2X0	30862296	31034421	2	172125
	19	19_3X0	50846379	51069137	2	222758
	2	2_10X0	42375787	42425609	3	49822
	2	2_11X0	42611905	42762902	2	150997
	2	2_12X0	43144071	43260246	2	116175
	2	2_13X0	54441812	54545430	2	103618
	2	2_14X0	55264868	55383113	2	118245
	2	2_15X0	58781680	58887526	3	105846
	2	2_16X0	59001473	59045465	3	43992
	2	2_17X0	59602961	59643370	4	40409
	2	2_18X0	60236180	60306083	2	69903
	2	2_19X0	66826189	66967274	2	141085
	2	2_1X0	23186095	23253489	2	67394
	2	2_20X0	67940747	68163060	2	222313
	2	2_21X0	68694870	68824437	2	129567
	2	2_22X0	68874543	68996711	3	122168
	2	2_23X0	69353359	69492506	2	139147
	2	2_24X0	72678384	72920438	2	242054

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	2	2_25X0	74553802	74580211	2	26409
	2	2_26X0	75428213	75478812	2	50599
	2	2_27X0	76649363	76670143	3	20780
	2	2_28X0	78166449	78224596	2	58147
	2	2_29X0	78687534	78837719	5	150185
	2	2_2X0	28368004	28506240	2	138236
	2	2_30X0	80888958	81149288	5	260330
	2	2_31X0	83737052	83829486	2	92434
	2	2_32X0	84396073	84408369	2	12296
	2	2_33X0	84545930	84570127	2	24197
	2	2_3X0	29932127	30034106	2	101979
	2	2_4X0	30095983	30314174	2	218191
	2	2_5X0	30411502	30431684	2	20182
	2	2_6X0	30674594	30703389	2	28795
	2	2_7X0	35828592	35850523	2	21931
	2	2_8X0	35944005	36480128	9	536123
	2	2_9X0	36813586	36900729	2	87143
	20	20_10X0	39029018	39047356	3	18338
	20	20_11X0	39059245	39161325	4	102080
	20	20_12X0	39199236	39334107	2	134871
	20	20_13X0	41970311	42051815	4	81504
	20	20_14X0	42270477	42384217	4	113740
	20	20_14X1	42511431	42616823	3	105392
	20	20_15X0	43541276	43568517	2	27241
	20	20_16X0	46391445	46588842	8	197397
	20	20_17X0	46611481	46694052	3	82571
	20	20_18X0	46773530	46809732	2	36202
	20	20_19X0	47145651	47356182	10	210531
	20	20_19X1	47424053	47616143	6	192090
	20	20_1X0	27659	49035	2	21376
	20	20_20X0	47844170	47968581	3	124411
	20	20_20X1	48021369	48220048	3	198679
	20	20_21X0	49425534	49443595	2	18061
	20	20_22X0	49458696	49539566	2	80870
	20	20_23X0	50808989	50827914	3	18925
	20	20_24X0	51239732	51459125	8	219393
	20	20_25X0	51579751	51720118	5	140367
	20	20_26X0	52449871	52520931	3	71060
	20	20_27X0	53238798	53465258	3	226460
	20	20_28X0	53722942	53745885	2	22943
	20	20_29X0	53916844	53981643	2	64799
	20	20_2X0	3098937	3171166	2	72229
	20	20_30X0	54208840	54225841	2	17001
	20	20_31X0	54345210	54414119	2	68909
	20	20_32X0	55465078	55704949	3	239871
	20	20_33X0	55713800	55747466	2	33666
	20	20_34X0	56138659	56161071	2	22412
	20	20_35X0	56209815	56390521	5	180706
	20	20_36X0	56966668	57024344	2	57676
	20	20_37X0	57416935	57525318	2	108383
	20	20_38X0	57783153	57880169	3	97016
	20	20_3X0	4167354	4260128	2	92774
	20	20_5X0	8342660	8367658	2	24998
	20	20_6X0	37458129	37465768	2	7639
	20	20_8X0	37738670	37746979	2	8309

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	20	20_9X0	37797267	37889736	7	92469
	21	21_10X0	26203333	26296027	2	92694
	21	21_11X0	28047095	28654490	33	607395
	21	21_11X1	27225856	27598763	20	372907
	21	21_11X2	28687790	29605725	36	917935
	21	21_11X3	27613634	28035192	14	421558
	21	21_11X4	26635476	27198390	17	562914
	21	21_12X0	30082563	30640747	15	558184
	21	21_13X0	30650626	30707418	2	56792
	21	21_14X0	31359682	31493248	8	133566
	21	21_15X0	37658340	37918226	4	259886
	21	21_16X0	40598093	40704315	4	106222
	21	21_17X0	40834559	41006448	4	171889
	21	21_18X0	50334769	50369183	2	34414
	21	21_19X0	50448936	50733704	9	284768
	21	21_1X0	10273971	10388785	4	114814
	21	21_2X0	13787148	13820314	2	33166
	21	21_3X0	15943521	15998477	2	54956
	21	21_4X0	20660163	20800622	2	140459
	21	21_5X0	21915956	22056918	2	140962
	21	21_6X0	22834671	23001371	2	166700
	21	21_7X0	24929780	24996055	2	66275
	21	21_8X0	25917318	25947503	2	30185
	21	21_9X0	25976207	25998114	2	21907
	22	22_1X0	2511379	3112499	3	601120
	22	22_2X0	9388832	9469152	2	80320
	22	22_3X0	49288343	49334270	2	45927
	22	22_4X0	49936039	49951896	2	15857
	22	22_5X0	52182115	52395759	2	213644
	22	22_6X0	60572511	60596983	2	24472
	23	23_10X0	30603285	30749184	3	145899
	23	23_11X0	35205836	35249236	2	43400
	23	23_12X0	35619227	35698170	2	78943
	23	23_13X0	41411292	41739546	4	328254
	23	23_14X0	42509756	42538734	2	28978
	23	23_15X0	43645532	43748625	2	103093
	23	23_16X0	43979890	44054097	2	74207
	23	23_17X0	44082290	44262234	3	179944
	23	23_18X0	45782630	45910513	5	127883
	23	23_19X0	46289398	46363365	2	73967
	23	23_1X0	1821414	1940465	3	119051
	23	23_20X0	51963107	52019047	2	55940
	23	23_2X0	7932254	8218841	4	286587
	23	23_3X0	8280793	8613345	3	332552
	23	23_4X0	8936791	9019127	2	82336
	23	23_5X0	9923607	9979936	3	56329
	23	23_6X0	11790314	11986393	2	196079
	23	23_7X0	12295777	12508186	3	212409
	23	23_8X0	19547145	19787455	2	240310
	23	23_9X0	27439868	27899273	3	459405
	24	24_10X0	20614030	20989827	12	375797
	24	24_11X0	21284562	21513489	2	228927
	24	24_12X0	22275899	22555981	6	280082
	24	24_13X0	24578230	24647433	2	69203
	24	24_14X0	26774177	26906360	3	132183

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	24	24_15X0	32487264	32568464	5	81200
	24	24_16X0	32720435	33102815	6	382380
	24	24_17X0	33163491	33177526	2	14035
	24	24_18X0	33684698	33755706	2	71008
	24	24_19X0	46548720	46765965	2	217245
	24	24_1X0	73973	202313	5	128340
	24	24_2X0	559973	601649	2	41676
	24	24_3X0	1931076	2036434	2	105358
	24	24_5X0	16825655	16847632	2	21977
	24	24_6X0	18182856	18194002	2	11146
	24	24_7X0	18700293	18782729	2	82436
	24	24_8X0	19097392	19573560	7	476168
	24	24_9X0	20361797	20457489	2	95692
	25	25_10X0	40364426	40420187	2	55761
	25	25_11X0	42868638	43042947	2	174309
	25	25_12X0	44166152	44183489	2	17337
	25	25_13X0	45033066	45161632	2	128566
	25	25_14X0	49982598	50260238	8	277640
	25	25_1X0	9145011	9200046	2	55035
	25	25_2X0	11701226	11855169	3	153943
	25	25_3X0	17920820	18004160	3	83340
	25	25_4X0	25904179	25964005	2	59826
	25	25_5X0	27419275	27502327	2	83052
	25	25_6X0	32463497	32498570	2	35073
	25	25_7X0	32946078	33147802	3	201724
	25	25_8X0	33781879	33816626	2	34747
	25	25_9X0	37245728	37296628	2	50900
	26	26_10X0	27364199	27632004	9	267805
	26	26_10X1	27265388	27354548	3	89160
	26	26_10X2	25886955	26080218	5	193263
	26	26_10X3	25571077	25753391	4	182314
	26	26_10X4	26151663	26451731	5	300068
	26	26_10X5	26760401	27165446	5	405045
	26	26_10X6	26570289	26704730	3	134441
	26	26_11X0	27782117	27844572	2	62455
	26	26_12X0	28304432	28347613	2	43181
	26	26_13X0	28526356	28580865	5	54509
	26	26_14X0	28750742	28834679	3	83937
	26	26_15X0	30170326	30215776	2	45450
	26	26_16X0	30428085	30497645	2	69560
	26	26_17X0	31013607	31177352	2	163745
	26	26_18X0	38419869	38622343	4	202474
	26	26_1X0	46172	257312	5	211140
	26	26_2X0	7896837	7983848	2	87011
	26	26_3X0	9423486	9470336	3	46850
	26	26_4X0	10385074	10451332	3	66258
	26	26_5X0	16269219	16369615	2	100396
	26	26_6X0	16836790	16869102	2	32312
	26	26_7X0	23542659	23612282	3	69623
	26	26_8X0	24884853	25026600	2	141747
	26	26_9X0	25050205	25455562	6	405357
	27	27_10X0	5333820	5463452	3	129632
	27	27_11X0	5509630	5686692	3	177062
	27	27_12X0	5861071	5901362	2	40291
	27	27_13X0	5928245	6075130	6	146885

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	27	27_13X1	6107219	6347985	4	240766
	27	27_14X0	8056308	8174464	2	118156
	27	27_15X0	25780103	25974768	3	194665
	27	27_16X0	26035792	26198320	2	162528
	27	27_17X0	26348490	26549041	3	200551
	27	27_18X0	31792144	31803070	2	10926
	27	27_19X0	33177639	33302370	2	124731
	27	27_1X0	40637	306017	11	265380
	27	27_20X0	33417456	33460377	2	42921
	27	27_21X0	34739537	34899902	6	160365
	27	27_22X0	35144651	35215448	2	70797
	27	27_22X1	35309329	35386012	2	76683
	27	27_23X0	35864121	36145641	7	281520
	27	27_23X1	36152112	36376701	4	224589
	27	27_23X2	35565795	35838821	4	273026
	27	27_24X0	36412031	36696052	3	284021
	27	27_25X0	37064014	37113663	2	49649
	27	27_26X0	37519570	37568357	2	48787
	27	27_27X0	37935664	37985995	3	50331
	27	27_28X0	39792638	40045812	3	253174
	27	27_29X0	40440576	40509766	2	69190
	27	27_2X0	848200	877029	2	28829
	27	27_30X0	41478516	41648367	2	169851
	27	27_31X0	42269697	42336459	2	66762
	27	27_3X0	1210345	1323370	8	113025
	27	27_4X0	1729206	1795509	2	66303
	27	27_5X0	1802588	2004180	3	201592
	27	27_6X0	2609661	2883787	10	274126
	27	27_6X1	2189327	2545103	11	355776
	27	27_7X0	3012761	3090670	2	77909
	27	27_8X0	4727561	4751001	3	23440
	27	27_9X0	4801623	4944783	2	143160
	28	28_10X0	32666938	32687981	2	21043
	28	28_11X0	40950955	41035293	3	84338
	28	28_12X0	41078615	41122819	2	44204
	28	28_1X0	2260719	2385710	3	124991
	28	28_2X0	3484422	3541294	3	56872
	28	28_3X0	13807720	13877118	2	69398
	28	28_4X0	13988517	14301924	2	313407
	28	28_5X0	15814429	15842455	3	28026
	28	28_6X0	16568909	16591387	2	22478
	28	28_7X0	27139260	27212128	3	72868
	28	28_8X0	32251189	32465175	3	213986
	28	28_9X0	32479841	32527475	2	47634
	29	29_1X0	4773409	5052699	2	279290
	29	29_2X0	7367481	7512253	2	144772
	29	29_3X0	23231812	23399408	2	167596
	29	29_4X0	28489947	28701204	4	211257
	29	29_5X0	31784610	31975411	4	190801
	29	29_6X0	41767804	41802495	3	34691
	3	3_10X0	51814598	51833514	2	18916
	3	3_11X0	52015706	52042064	2	26358
	3	3_12X0	52738294	52759981	2	21687
	3	3_13X0	53815273	54056768	4	241495
	3	3_14X0	58403292	58611898	2	208606

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	3	3_15X0	64529721	64641059	2	111338
	3	3_16X0	70809592	71002819	2	193227
	3	3_17X0	73520706	73570744	3	50038
	3	3_1X0	8632397	8825682	2	193285
	3	3_2X0	12681688	12919295	3	237607
	3	3_3X0	23999219	24226507	2	227288
	3	3_4X0	27731288	27790126	2	58838
	3	3_5X0	29799435	29975589	3	176154
	3	3_6X0	30776403	31186599	3	410196
	3	3_7X0	31306813	31571035	9	264222
	3	3_8X0	32003457	32097303	2	93846
	3	3_9X0	36314104	36365870	2	51766
	30	30_10X0	29702776	29777531	2	74755
	30	30_11X0	36845157	36990110	2	144953
	30	30_12X0	37291848	37329458	2	37610
	30	30_13X0	37706678	37903702	5	197024
	30	30_14X0	37906058	38029576	2	123518
	30	30_1X0	410672	698401	7	287729
	30	30_1X1	41360	272687	5	231327
	30	30_2X0	7361964	7462665	2	100701
	30	30_3X0	8868294	9164608	4	296314
	30	30_4X0	10065515	10148830	3	83315
	30	30_5X0	11521841	11588202	4	66361
	30	30_6X0	16358958	16453807	2	94849
	30	30_7X0	17455927	17595323	2	139396
	30	30_8X0	17889709	18116499	2	226790
	30	30_9X0	22760781	22876424	2	115643
	31	31_10X0	35860596	35929863	2	69267
	31	31_11X0	36638762	36684660	3	45898
	31	31_12X0	37653072	37746203	2	93131
	31	31_13X0	38183532	38261972	2	78440
	31	31_14X0	38306559	38396459	3	89900
	31	31_15X0	38834608	38956271	2	121663
	31	31_16X0	39301588	39448160	2	146572
	31	31_1X0	22169406	22289456	2	120050
	31	31_2X0	25041093	25090720	2	49627
	31	31_3X0	25144046	25315783	7	171737
	31	31_4X0	25345975	25583403	2	237428
	31	31_5X0	27674126	27721393	2	47267
	31	31_6X0	27869402	28049722	4	180320
	31	31_7X0	28759607	28948914	4	189307
	31	31_8X0	29777154	29859118	2	81964
	31	31_9X0	31316431	31423731	4	107300
	32	32_1X0	557199	600003	2	42804
	32	32_2X0	1613746	1728227	2	114481
	32	32_3X0	7001312	7089031	2	87719
	32	32_4X0	11695052	11978818	3	283766
	32	32_5X0	21249740	21437450	4	187710
	32	32_6X0	24292979	24405528	2	112549
	33	33_1X0	5093779	5397165	10	303386
	33	33_2X0	14457713	14493209	2	35496
	33	33_3X0	17272560	17629170	3	356610
	33	33_4X0	24713002	24836184	2	123182
	33	33_5X0	25640915	25801465	3	160550
	33	33_6X0	27928744	28044355	2	115611

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	33	33_7X0	31340428	31360330	2	19902
	34	34_10X0	34407701	34470243	3	62542
	34	34_1X0	122610	210758	2	88148
	34	34_2X0	11210939	11344518	3	133579
	34	34_3X0	16446944	16620899	2	173955
	34	34_4X0	16851376	17089640	5	238264
	34	34_5X0	19221060	19318160	4	97100
	34	34_6X0	19675614	19831943	2	156329
	34	34_7X0	22302592	22412102	2	109510
	34	34_8X0	26332805	26498821	2	166016
	34	34_9X0	32331034	32624469	2	293435
	35	35_10X0	25263532	25719965	12	456433
	35	35_11X0	26016794	26179533	4	162739
	35	35_12X0	26256379	26394245	4	137866
	35	35_1X0	3147428	3263311	3	115883
	35	35_2X0	23071318	23169723	2	98405
	35	35_3X0	23677318	23691758	4	14440
	35	35_4X0	23725328	23852228	4	126900
	35	35_5X0	23974030	24289167	29	315137
	35	35_6X0	24543339	24558879	4	15540
	35	35_7X0	24742416	24796822	2	54406
	35	35_8X0	24999530	25061935	12	62405
	35	35_9X0	25078788	25194856	2	116068
	36	36_1X0	3541248	3818650	2	277402
	36	36_2X0	5569724	5793654	2	223930
	36	36_3X0	7438641	7647681	2	209040
	36	36_4X0	11108050	11561455	3	453405
	36	36_5X0	14206403	14434299	2	227896
	36	36_6X0	16485227	16502192	2	16965
	36	36_7X0	19887524	19996296	8	108772
	36	36_8X0	21271603	21330692	2	59089
	36	36_9X0	30488591	30677658	2	189067
	37	37_1X0	9838063	9992935	2	154872
	37	37_2X0	10283661	10401822	3	118161
	37	37_3X0	12503905	12737748	3	233843
	37	37_4X0	23545856	23596146	2	50290
	37	37_5X0	24839130	24943267	3	104137
	37	37_6X0	25453926	25475491	2	21565
	37	37_7X0	25794367	26007702	4	213335
	37	37_8X0	26008461	26079488	2	71027
	38	38_10X0	21893001	21999598	2	106597
	38	38_11X0	22023393	22080856	2	57463
	38	38_12X0	22096531	22320700	6	224169
	38	38_13X0	22472210	23284259	29	812049
	38	38_14X0	23294565	23493501	6	198936
	38	38_1X0	103799	221830	3	118031
	38	38_2X0	1481531	1600682	2	119151
	38	38_3X0	1670091	2001965	3	331874
	38	38_4X0	2037450	2127510	2	90060
	38	38_5X0	2196391	2351385	2	154994
	38	38_6X0	15426910	15485200	2	58290
	38	38_7X0	19683202	19778927	2	95725
	38	38_8X0	20009666	20278374	2	268708
	38	38_9X0	21378144	21768148	9	390004
	38	38_9X1	20943916	21211572	5	267656

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	4	4_10X0	35928497	36251858	2	323361
	4	4_11X0	37383255	37552910	2	169655
	4	4_12X0	49006946	49578074	3	571128
	4	4_13X0	53018990	53131340	3	112350
	4	4_14X0	57837526	58013981	3	176455
	4	4_15X0	58800921	59009920	3	208999
	4	4_16X0	60024886	60327235	2	302349
	4	4_17X0	60454573	60473800	2	19227
	4	4_18X0	72461733	72577186	2	115453
	4	4_1X0	63038	141487	4	78449
	4	4_2X0	153460	436743	5	283283
	4	4_3X0	19054718	19159847	2	105129
	4	4_4X0	20073139	20340882	3	267743
	4	4_5X0	20354231	20516136	2	161905
	4	4_6X0	29394768	29433622	2	38854
	4	4_7X0	29493062	29514409	2	21347
	4	4_8X0	32464825	32477693	2	12868
	4	4_9X0	34996577	35018160	2	21583
	5	5_10X0	15576742	15752710	2	175968
	5	5_11X0	15801888	15823057	2	21169
	5	5_12X0	16740359	16776768	3	36409
	5	5_13X0	19022616	19064649	2	42033
	5	5_14X0	19324444	19396997	2	72553
	5	5_15X0	21180081	21186071	2	5990
	5	5_16X0	27133771	27179203	2	45432
	5	5_17X0	28853642	29194593	8	340951
	5	5_18X0	29293036	29337240	2	44204
	5	5_19X0	31883544	32037775	2	154231
	5	5_1X0	753035	931452	3	178417
	5	5_20X0	32059463	32244022	3	184559
	5	5_21X0	32844494	33018422	5	173928
	5	5_22X0	33471196	33605342	2	134146
	5	5_23X0	34618393	34936610	5	318217
	5	5_24X0	39203759	39401249	4	197490
	5	5_25X0	40493711	40664010	3	170299
	5	5_26X0	40832203	41040385	2	208182
	5	5_27X0	43197365	43368441	2	171076
	5	5_28X0	56566388	56693429	2	127041
	5	5_29X0	60051282	60279124	2	227842
	5	5_2X0	1082047	1316834	2	234787
	5	5_30X0	63043496	63121684	2	78188
	5	5_31X0	64661108	64853343	2	192235
	5	5_32X0	66283528	66347925	3	64397
	5	5_33X0	69927984	70038748	2	110764
	5	5_34X0	75500419	75518686	3	18267
	5	5_35X0	76332872	76369090	2	36218
	5	5_36X0	77246058	77272685	2	26627
	5	5_37X0	80759112	80913018	2	153906
	5	5_38X0	82159592	82204373	2	44781
	5	5_39X0	82640259	82726778	5	86519
	5	5_3X0	4457959	4517483	2	59524
	5	5_4X0	5736511	5812794	2	76283
	5	5_5X0	8752716	8813176	2	60460
	5	5_6X0	9520760	9673068	4	152308
	5	5_7X0	10297144	10698707	17	401563

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	5	5_7X1	9826884	10260671	16	433787
	5	5_8X0	15377812	15415551	3	37739
	5	5_9X0	15446739	15560441	5	113702
	6	6_10X0	11952306	12129347	6	177041
	6	6_11X0	15792897	15825118	3	32221
	6	6_12X0	16808971	16996535	4	187564
	6	6_13X0	17096996	17148633	3	51637
	6	6_14X0	17149403	17870087	13	720684
	6	6_15X0	18014717	18115188	2	100471
	6	6_16X0	19179574	19272157	2	92583
	6	6_17X0	22150057	22176450	2	26393
	6	6_18X0	22440182	22593478	2	153296
	6	6_19X0	24558190	24639226	2	81036
	6	6_19X1	24862966	24940992	2	78026
	6	6_1X0	5657889	5963867	3	305978
	6	6_20X0	24993089	25027339	2	34250
	6	6_21X0	25765909	25902478	2	136569
	6	6_22X0	27800109	27989560	2	189451
	6	6_23X0	37793337	38115352	7	322015
	6	6_24X0	38170608	38173203	2	2595
	6	6_25X0	38200359	38471005	5	270646
	6	6_26X0	38595117	38712112	2	116995
	6	6_27X0	38762103	38964641	4	202538
	6	6_28X0	39477693	39506955	2	29262
	6	6_29X0	39820852	39832346	2	11494
	6	6_2X0	6488348	6595281	4	106933
	6	6_30X0	40158250	40233303	2	75053
	6	6_31X0	40324263	40343504	5	19241
	6	6_32X0	40646351	40718462	2	72111
	6	6_33X0	41343610	41497254	3	153644
	6	6_34X0	42188441	42207944	2	19503
	6	6_35X0	42259702	42315412	2	55710
	6	6_36X0	43296811	43475039	3	178228
	6	6_37X0	58692792	59048718	3	355926
	6	6_38X0	59150614	59411940	3	261326
	6	6_39X0	61510741	61704722	4	193981
	6	6_3X0	7150741	7179455	2	28714
	6	6_40X0	68364900	68400785	2	35885
	6	6_41X0	75515640	75710605	2	194965
	6	6_4X0	7645981	7666484	2	20503
	6	6_5X0	9272713	9462751	2	190038
	6	6_6X0	9524958	9587011	3	62053
	6	6_7X0	9688452	9909036	5	220584
	6	6_8X0	9931501	10102598	6	171097
	6	6_9X0	10192601	10245017	2	52416
	7	7_10X0	12555151	12617869	2	62718
	7	7_11X0	13411792	13501541	2	89749
	7	7_12X0	15747616	15792956	3	45340
	7	7_13X0	15813247	16080403	3	267156
	7	7_14X0	16406353	16624085	2	217732
	7	7_15X0	24045302	24299082	2	253780
	7	7_16X0	25349786	25484264	2	134478
	7	7_17X0	26525328	26561320	2	35992
	7	7_18X0	27159022	27253654	2	94632
	7	7_19X0	27568036	27845626	5	277590

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	7	7_1X0	205382	327607	3	122225
	7	7_20X0	28888977	28995600	3	106623
	7	7_21X0	39755785	40022337	2	266552
	7	7_22X0	40334172	40451710	2	117538
	7	7_23X0	40481121	40857017	9	375896
	7	7_24X0	40956024	40991742	2	35718
	7	7_25X0	41140931	41176758	2	35827
	7	7_26X0	41752921	41900917	2	147996
	7	7_27X0	42233311	42249405	2	16094
	7	7_28X0	42373920	42431709	3	57789
	7	7_29X0	42443820	42466673	2	22853
	7	7_2X0	731601	888325	2	156724
	7	7_30X0	45378307	45489588	2	111281
	7	7_31X0	54665927	54758964	5	93037
	7	7_32X0	57523710	57607975	2	84265
	7	7_33X0	57983729	58457593	7	473864
	7	7_34X0	69866914	69894878	2	27964
	7	7_3X0	1598553	1658556	2	60003
	7	7_4X0	5674454	5752812	2	78358
	7	7_5X0	5818734	5903839	2	85105
	7	7_6X0	5932705	6047097	4	114392
	7	7_7X0	6048957	6107032	3	58075
	7	7_8X0	6484113	6643653	3	159540
	7	7_9X0	10789098	10853332	2	64234
	8	8_10X0	28480478	28538838	2	58360
	8	8_11X0	35565835	35748709	3	182874
	8	8_12X0	38913832	38952619	2	38787
	8	8_13X0	41653831	41695447	2	41616
	8	8_14X0	44109780	44266287	4	156507
	8	8_15X0	46853060	46911577	3	58517
	8	8_16X0	48027599	48218449	2	190850
	8	8_17X0	48330836	48561208	3	230372
	8	8_18X0	50271098	50285876	2	14778
	8	8_19X0	63324277	63596747	9	272470
	8	8_1X0	332660	796834	7	464174
	8	8_20X0	64974782	65026775	2	51993
	8	8_21X0	68578262	68606872	2	28610
	8	8_22X0	72847361	73267073	5	419712
	8	8_23X0	73639115	74224248	8	585133
	8	8_2X0	2210577	2952167	19	741590
	8	8_3X0	3342703	3356136	2	13433
	8	8_4X0	3838617	4309140	8	470523
	8	8_5X0	4368782	4410598	2	41816
	8	8_6X0	4438211	4511843	4	73632
	8	8_7X0	22550532	22669527	2	118995
	8	8_8X0	26305253	26364891	2	59638
	8	8_9X0	26780137	26958893	2	178756
	9	9_10X0	12181059	12266468	2	85409
	9	9_11X0	16436630	16475722	2	39092
	9	9_12X0	18373992	18387336	2	13344
	9	9_13X0	19385788	19395605	2	9817
	9	9_14X0	19418182	19504401	2	86219
	9	9_15X0	20173910	20185103	2	11193
	9	9_16X0	20539480	20719680	3	180200
	9	9_17X0	20754213	20853837	2	99624

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	9	9_18X0	20873946	20922435	2	48489
	9	9_19X0	20975032	20989021	2	13989
	9	9_1X0	1839640	1855777	2	16137
	9	9_20X0	21147248	21418028	14	270780
	9	9_21X0	21677999	21688332	2	10333
	9	9_22X0	21710269	21971964	10	261695
	9	9_23X0	22462014	22493495	2	31481
	9	9_24X0	24178438	24245280	2	66842
	9	9_25X0	24792889	24961243	10	168354
	9	9_26X0	26001253	26023071	2	21818
	9	9_27X0	26352034	26402205	2	50171
	9	9_28X0	26599959	26661574	2	61615
	9	9_29X0	32899293	32937785	2	38492
	9	9_2X0	2991130	3014639	2	23509
	9	9_30X0	35166353	35246249	2	79896
	9	9_31X0	37681034	37823880	8	142846
	9	9_32X0	38934058	39010042	5	75984
	9	9_33X0	45844004	45848761	2	4757
	9	9_34X0	45876114	45902636	2	26522
	9	9_35X0	47061492	47317920	8	256428
	9	9_36X0	47368564	47431893	2	63329
	9	9_38X0	48287202	48299478	2	12276
	9	9_39X0	48528903	48567038	4	38135
	9	9_3X0	4101050	4165054	2	64004
	9	9_40X0	48644113	48808734	8	164621
	9	9_41X0	49602669	49758479	9	155810
	9	9_42X0	49786682	49799337	2	12655
	9	9_43X0	49839252	49959981	2	120729
	9	9_44X0	54017916	54036050	2	18134
	9	9_45X0	55503369	55536086	2	32717
	9	9_46X0	56397995	56465016	2	67021
	9	9_47X0	58385846	58488207	2	102361
	9	9_48X0	59969865	60118956	4	149091
	9	9_4X0	4714086	4731882	2	17796
	9	9_5X0	5708132	5976435	9	268303
	9	9_6X0	10081637	10172753	2	91116
	9	9_7X0	10287952	10328751	2	40799
	9	9_8X0	11497182	11535099	2	37917
	9	9_9X0	11613175	11703416	2	90241
	JH373243.1	JH373243.1_1X0	242582	259598	2	17016
	X	X_10X0	21525932	21654736	2	128804
	X	X_11X0	22982302	23029413	3	47111
	X	X_12X0	35883708	36108217	3	224509
	X	X_13X0	37677493	37884192	2	206699
	X	X_14X0	41353292	41652600	3	299308
	X	X_15X0	42273806	42299487	2	25681
	X	X_16X0	43916754	44069130	2	152376
	X	X_17X0	44451636	44682895	2	231259
	X	X_18X0	46501987	46615697	2	113710
	X	X_19X0	48017384	48090398	2	73014
	X	X_1X0	780894	911428	4	130534
	X	X_20X0	48420527	48429586	2	9059
	X	X_21X0	54516776	54684042	3	167266
	X	X_22X0	56770402	56860380	2	89978
	X	X_23X0	61174658	61420672	2	246014

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	X	X_24X0	75402944	75519218	5	116274
	X	X_25X0	76017820	76047740	2	29920
	X	X_26X0	76272606	76317833	2	45227
	X	X_27X0	76373430	76755594	7	382164
	X	X_28X0	77023034	77055971	2	32937
	X	X_29X0	79603965	79859594	2	255629
	X	X_2X0	1012952	1120489	2	107537
	X	X_30X0	82335592	82602970	2	267378
	X	X_31X0	91471708	91527183	2	55475
	X	X_32X0	92057968	92129748	3	71780
	X	X_33X0	99375977	99413876	2	37899
	X	X_34X0	105397067	105463162	2	66095
	X	X_35X0	105634893	105934389	5	299496
	X	X_36X0	117556011	117810235	7	254224
	X	X_37X0	118088226	118339465	5	251239
	X	X_38X0	118820517	118999359	2	178842
	X	X_39X0	119870193	120032312	3	162119
	X	X_3X0	1412931	1461118	2	48187
	X	X_40X0	120946336	121095076	4	148740
	X	X_41X0	121489009	121987418	5	498409
	X	X_42X0	122185170	122567736	3	382566
	X	X_4X0	1516352	1635990	4	119638
	X	X_5X0	2893379	2925206	3	31827
	X	X_6X0	9334062	9399887	2	65825
	X	X_7X0	11511669	11567726	2	56057
	X	X_8X0	11913239	12040221	2	126982
	X	X_9X0	19878167	19906268	2	28101
	1	1_10X0	5563977	5594234	2	30257
	1	1_11X0	6507755	6607145	2	99390
	1	1_12X0	6740211	6808456	2	68245
	1	1_13X0	6828682	6995259	2	166577
	1	1_14X0	7594677	7665576	2	70899
	1	1_15X0	8385020	8714763	4	329743
	1	1_16X0	9390825	9451818	2	60993
	1	1_17X0	9632645	9679813	2	47168
	1	1_18X0	10428432	10483159	2	54727
	1	1_19X0	10632566	10804423	4	171857
	1	1_1X0	93334	607893	7	514559
	1	1_20X0	10886690	10925983	2	39293
	1	1_21X0	11063265	11302472	2	239207
	1	1_22X0	11470195	11669890	6	199695
	1	1_23X0	11910318	12055159	3	144841
	1	1_24X0	12594379	12653780	2	59401
	1	1_25X0	12719333	12877704	2	158371
	1	1_26X0	13036412	13685000	16	648588
	1	1_27X0	13694537	13785652	2	91115
	1	1_28X0	13813835	13849516	2	35681
	1	1_29X0	14087468	14206125	6	118657
	1	1_2X0	639345	673763	2	34418
	1	1_30X0	14441116	14682393	6	241277
	1	1_31X0	14846517	15205273	10	358756
	1	1_32X0	16359212	17106975	29	747763
	1	1_32X1	15738207	16192481	8	454274
	1	1_33X0	21293958	21330477	2	36519
	1	1_34X0	30591273	30785774	2	194501

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_35X0	45930332	46050947	2	120615
	1	1_36X0	46334562	46540441	5	205879
	1	1_37X0	46672577	46894914	7	222337
	1	1_38X0	49840286	49954092	2	113806
	1	1_39X0	56856497	57089676	5	233179
	1	1_3X0	824343	987610	3	163267
	1	1_40X0	57822351	57880699	2	58348
	1	1_41X0	57891612	58062732	3	171120
	1	1_42X0	58555321	58773887	4	218566
	1	1_43X0	61912772	62460659	15	547887
	1	1_43X1	60877898	61869312	19	991414
	1	1_44X0	63067083	63085138	2	18055
	1	1_45X0	63293676	63313293	2	19617
	1	1_46X0	63346807	63441743	4	94936
	1	1_47X0	64479642	64542207	2	62565
	1	1_48X0	65235775	65278807	2	43032
	1	1_49X0	66935584	67021182	2	85598
	1	1_4X0	1659470	1738477	2	79007
	1	1_50X0	67864609	68002490	2	137881
	1	1_51X0	69219985	69381551	2	161566
	1	1_52X0	76869431	76897841	2	28410
	1	1_53X0	80881379	81190817	7	309438
	1	1_54X0	86212761	86271843	2	59082
	1	1_55X0	93889198	94334747	9	445549
	1	1_56X0	96283235	96480359	5	197124
	1	1_57X0	103257347	103263952	2	6605
	1	1_58X0	105288865	105374664	2	85799
	1	1_59X0	105923408	106071176	3	147768
	1	1_5X0	1936830	1966613	2	29783
	1	1_60X0	108155641	108203609	3	47968
	1	1_61X0	109186886	109209424	2	22538
	1	1_62X0	109270750	109495224	2	224474
	1	1_63X0	109509979	109645145	5	135166
	1	1_64X0	109681065	109732014	3	50949
	1	1_65X0	110684458	110791569	2	107111
	1	1_66X0	115162700	115717525	12	554825
	1	1_67X0	115896167	116080125	5	183958
	1	1_68X0	116796896	117001796	4	204900
	1	1_69X0	119969370	120057023	2	87653
	1	1_6X0	2110467	2211389	2	100922
	1	1_70X0	121465883	121535609	2	69726
	1	1_71X0	124974027	125211322	2	237295
	1	1_72X0	125360454	125454965	2	94511
	1	1_7X0	3035455	3046673	2	11218
	1	1_8X0	3086903	3182381	2	95478
	1	1_9X0	4746696	4859102	3	112406
	10	10_10X0	76492533	76674928	2	182395
	10	10_11X0	86782240	86943829	3	161589
	10	10_12X0	87778167	87908589	2	130422
	10	10_13X0	89471730	89732944	3	261214
	10	10_14X0	102469669	102531724	2	62055
	10	10_15X0	104022811	104145413	3	122602
	10	10_16X0	104760045	104794854	2	34809
	10	10_17X0	105150421	105387810	2	237389
	10	10_18X0	113028552	113335312	2	306760

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	10	10_19X0	116947905	117552943	3	605038
	10	10_1X0	11163604	11388999	2	225395
	10	10_20X0	118388992	118479499	2	90507
	10	10_21X0	118546464	118565610	2	19146
	10	10_22X0	119635523	119797366	2	161843
	10	10_23X0	126087495	126098496	2	11001
	10	10_2X0	36679830	36715809	2	35979
	10	10_3X0	42942284	43281177	2	338893
	10	10_4X0	47394093	47652689	2	258596
	10	10_5X0	57590838	57608079	2	17241
	10	10_6X0	58589300	58717615	2	128315
	10	10_7X0	61654829	61766081	10	111252
	10	10_8X0	63860341	64054358	2	194017
	10	10_9X0	74455068	74659366	2	204298
	11	11_10X0	10621391	11143994	12	522603
	11	11_11X0	12816729	12857124	2	40395
	11	11_12X0	12990594	13119541	2	128947
	11	11_13X0	14706660	14723889	2	17229
	11	11_14X0	20880399	21283796	3	403397
	11	11_14X1	20555846	20761161	2	205315
	11	11_15X0	42421910	42595484	2	173574
	11	11_16X0	44686450	44811615	3	125165
	11	11_17X0	45097076	45525212	10	428136
	11	11_18X0	45976700	46131155	2	154455
	11	11_19X0	46180387	46206161	3	25774
	11	11_1X0	192398	264719	2	72321
	11	11_20X0	48082135	48166813	2	84678
	11	11_21X0	48315946	49081642	23	765696
	11	11_22X0	49288374	49512541	3	224167
	11	11_23X0	49525641	49600335	2	74694
	11	11_24X0	50049509	50166267	9	116758
	11	11_25X0	50746518	50764346	2	17828
	11	11_26X0	51210918	51585370	10	374452
	11	11_27X0	51871563	52098440	4	226877
	11	11_28X0	52710758	52854167	3	143409
	11	11_29X0	53348742	53370322	2	21580
	11	11_2X0	4398675	4473916	2	75241
	11	11_30X0	63831673	63886761	2	55088
	11	11_31X0	70937927	71104778	3	166851
	11	11_32X0	76182151	76193549	2	11398
	11	11_33X0	86374285	86586891	4	212606
	11	11_34X0	94087620	94099997	2	12377
	11	11_35X0	108163299	108252836	3	89537
	11	11_36X0	108642694	108693815	2	51121
	11	11_37X0	115701408	115822700	2	121292
	11	11_38X0	116529346	116609511	2	80165
	11	11_39X0	118042809	118057643	2	14834
	11	11_3X0	4854570	5087335	3	232765
	11	11_40X0	128240319	128497779	7	257460
	11	11_4X0	6364330	6499738	3	135408
	11	11_5X0	6653445	6860104	4	206659
	11	11_6X0	7125367	7177324	3	51957
	11	11_7X0	7848703	8017559	2	168856
	11	11_8X0	8190370	8321274	4	130904
	11	11_9X0	9481275	10398481	20	917206

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	12	12_10X0	11252192	11302325	2	50133
	12	12_11X0	13592088	13666863	2	74775
	12	12_12X0	15214777	15224237	2	9460
	12	12_13X0	15334587	15517890	4	183303
	12	12_14X0	26242384	26293083	2	50699
	12	12_15X0	26311334	26419634	2	108300
	12	12_16X0	26485828	26544285	2	58457
	12	12_17X0	29219827	29301530	2	81703
	12	12_18X0	30083128	30382212	2	299084
	12	12_19X0	30691018	30700266	2	9248
	12	12_1X0	1211487	1456008	7	244521
	12	12_20X0	34865596	35019206	5	153610
	12	12_21X0	44605030	44735442	4	130412
	12	12_22X0	44793253	44990183	2	196930
	12	12_23X0	45768187	45930904	4	162717
	12	12_24X0	45955863	46069944	2	114081
	12	12_25X0	46818779	46888890	2	70111
	12	12_26X0	58362756	58706182	12	343426
	12	12_27X0	73938042	74056806	2	118764
	12	12_28X0	74186116	74372729	3	186613
	12	12_29X0	78461987	78678284	3	216297
	12	12_2X0	2601380	2648073	3	46693
	12	12_30X0	80307380	80577500	2	270120
	12	12_31X0	102908479	103053678	4	145199
	12	12_32X0	104617683	104939414	2	321731
	12	12_33X0	108453156	108527219	2	74063
	12	12_3X0	3231325	3283549	2	52224
	12	12_4X0	4562864	4714386	2	151522
	12	12_5X0	4759388	4785385	2	25997
	12	12_6X0	4863165	4988978	4	125813
	12	12_7X0	8312650	8333950	2	21300
	12	12_8X0	9634483	9873767	2	239284
	12	12_9X0	10205705	10238350	2	32645
	13	13_10X0	57536749	57560893	2	24144
	13	13_11X0	62606556	62786048	2	179492
	13	13_12X0	68648685	68707885	2	59200
	13	13_13X0	69831005	69833031	2	2026
	13	13_14X0	87456260	87523098	2	66838
	13	13_15X0	87745023	88074946	3	329923
	13	13_16X0	88083200	88322644	2	239444
	13	13_17X0	94865936	94914030	2	48094
	13	13_1X0	3909430	4046875	2	137445
	13	13_2X0	13936016	14073757	2	137741
	13	13_3X0	21498772	21740910	6	242138
	13	13_4X0	36203004	36397302	3	194298
	13	13_5X0	40762296	40855171	2	92875
	13	13_6X0	40889682	40973691	5	84009
	13	13_7X0	41681213	41918259	2	237046
	13	13_8X0	46170872	46248983	2	78111
	13	13_9X0	57252676	57354318	3	101642
	14	14_10X0	15789013	16151326	9	362313
	14	14_11X0	16609594	17917273	51	1307679
	14	14_12X0	18807427	18883885	2	76458
	14	14_13X0	21458986	21499850	2	40864
	14	14_14X0	21562847	21579820	2	16973

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	14	14_15X0	28066956	28163996	4	97040
	14	14_16X0	32762310	32779931	2	17621
	14	14_17X0	38145559	38179457	2	33898
	14	14_18X0	60027801	60287200	2	259399
	14	14_19X0	62213115	62281127	2	68012
	14	14_1X0	214100	254059	2	39959
	14	14_20X0	63390433	63438917	2	48484
	14	14_21X0	67931940	68196066	2	264126
	14	14_22X0	70072330	70296463	2	224133
	14	14_23X0	80011536	80234046	2	222510
	14	14_2X0	437662	464578	2	26916
	14	14_3X0	1095719	1159085	2	63366
	14	14_4X0	1165361	1187796	2	22435
	14	14_5X0	2948502	3005561	2	57059
	14	14_6X0	5673163	6114028	6	440865
	14	14_7X0	6137368	6364617	2	227249
	14	14_8X0	11622043	11871767	4	249724
	14	14_9X0	15303456	15403012	2	99556
	15	15_10X0	38855348	38954333	2	98985
	15	15_11X0	39359612	39496030	5	136418
	15	15_12X0	41206957	41394054	3	187097
	15	15_13X0	41807727	41877377	2	69650
	15	15_14X0	43281433	43309187	2	27754
	15	15_15X0	44106757	44457509	4	350752
	15	15_16X0	51011771	51102249	2	90478
	15	15_17X0	51487831	51535955	3	48124
	15	15_18X0	56472634	56560045	2	87411
	15	15_19X0	59519231	59985569	3	466338
	15	15_1X0	80909	182744	4	101835
	15	15_20X0	68632197	68902292	2	270095
	15	15_21X0	76748220	77005095	3	256875
	15	15_22X0	81468965	81978207	3	509242
	15	15_23X0	84516478	84619660	2	103182
	15	15_24X0	87357409	87646708	2	289299
	15	15_25X0	88437957	88494018	3	56061
	15	15_2X0	467210	867457	3	400247
	15	15_3X0	933070	1255925	4	322855
	15	15_4X0	2753810	2817514	3	63704
	15	15_5X0	5315240	5393468	3	78228
	15	15_6X0	5732928	5913734	2	180806
	15	15_7X0	19529565	19605777	3	76212
	15	15_8X0	30400222	30577484	2	177262
	15	15_9X0	32112205	32180384	2	68179
	16	16_10X0	6796842	6803655	2	6813
	16	16_11X0	7425990	7500428	3	74438
	16	16_12X0	8159400	8322320	2	162920
	16	16_13X0	9647777	9994402	6	346625
	16	16_14X0	14922584	14952958	2	30374
	16	16_15X0	15047426	15179169	2	131743
	16	16_16X0	17421229	17532828	2	111599
	16	16_17X0	18125252	18312756	4	187504
	16	16_18X0	22379129	22621676	4	242547
	16	16_19X0	23762444	23918783	2	156339
	16	16_1X0	1392571	1424077	2	31506
	16	16_20X0	24042254	24061827	2	19573

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	16	16_21X0	27763555	27872329	5	108774
	16	16_22X0	28739499	29033380	6	293881
	16	16_23X0	29350391	29567906	6	217515
	16	16_24X0	29620618	29708614	2	87996
	16	16_25X0	29875024	29897544	2	22520
	16	16_26X0	32627819	32711674	2	83855
	16	16_27X0	35120581	35207874	3	87293
	16	16_28X0	35232927	35247997	2	15070
	16	16_29X0	42176524	42195616	2	19092
	16	16_2X0	2668273	2947421	8	279148
	16	16_30X0	42573240	42642046	2	68806
	16	16_31X0	42795782	42829785	2	34003
	16	16_32X0	42866741	42947655	2	80914
	16	16_33X0	43188022	43226329	2	38307
	16	16_34X0	43340576	43366474	2	25898
	16	16_35X0	43573867	43605220	2	31353
	16	16_36X0	45038943	45116378	2	77435
	16	16_37X0	45791865	45811469	2	19604
	16	16_38X0	46730972	46756452	2	25480
	16	16_39X0	46944921	47170038	6	225117
	16	16_3X0	3019602	3090330	2	70728
	16	16_40X0	51452567	51494596	2	42029
	16	16_41X0	52278573	52697992	5	419419
	16	16_42X0	54383204	54569097	3	185893
	16	16_43X0	56809159	56950999	2	141840
	16	16_44X0	57348824	57410337	5	61513
	16	16_45X0	57550703	57669720	2	119017
	16	16_46X0	60473029	60586596	2	113567
	16	16_47X0	60735274	60876861	2	141587
	16	16_48X0	61236566	61261845	2	25279
	16	16_49X0	62439813	62453871	2	14058
	16	16_4X0	3226582	3366741	2	140159
	16	16_50X0	62487453	62600459	2	113006
	16	16_51X0	63418091	63449079	3	30988
	16	16_52X0	63619346	63680571	2	61225
	16	16_53X0	63922862	64070445	3	147583
	16	16_54X0	64115059	64242521	2	127462
	16	16_55X0	64267371	64333138	2	65767
	16	16_56X0	64394926	64422674	2	27748
	16	16_57X0	64617219	64887488	16	270269
	16	16_58X0	64913512	65085587	4	172075
	16	16_59X0	65100839	65156850	8	56011
	16	16_5X0	3914406	4022678	2	108272
	16	16_60X0	65174629	65189986	3	15357
	16	16_61X0	65199256	65532372	12	333116
	16	16_62X0	66079089	66118079	2	38990
	16	16_63X0	66194174	66266381	2	72207
	16	16_64X0	68247695	68327753	2	80058
	16	16_65X0	68941746	69139382	10	197636
	16	16_66X0	71741500	71802051	3	60551
	16	16_67X0	74611703	74645562	2	33859
	16	16_6X0	4556690	4575643	2	18953
	16	16_7X0	6243242	6369738	2	126496
	16	16_8X0	6395927	6402606	2	6679
	16	16_9X0	6433573	6551456	3	117883

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	17	17_10X0	26082451	26117080	2	34629
	17	17_11X0	28868627	29123565	4	254938
	17	17_12X0	29957508	29977487	2	19979
	17	17_13X0	30873666	31008036	5	134370
	17	17_14X0	32816553	32947190	3	130637
	17	17_15X0	34919373	34925980	2	6607
	17	17_16X0	35983430	36284900	3	301470
	17	17_17X0	36320541	36339416	3	18875
	17	17_18X0	38609944	38762277	2	152333
	17	17_19X0	38763206	39851127	14	1087921
	17	17_1X0	1609752	1630150	2	20398
	17	17_20X0	40203936	40211676	2	7740
	17	17_21X0	40301880	40353659	4	51779
	17	17_22X0	40359961	40439433	2	79472
	17	17_23X0	40557156	41188248	5	631092
	17	17_24X0	41941869	41964872	2	23003
	17	17_25X0	42142073	42428787	7	286714
	17	17_26X0	42688085	42903326	3	215241
	17	17_27X0	43029422	43534390	12	504968
	17	17_28X0	43975284	44512637	10	537353
	17	17_29X0	44639024	44723809	6	84785
	17	17_2X0	1775955	1922579	5	146624
	17	17_30X0	45061170	45118612	2	57442
	17	17_31X0	45363109	45378192	2	15083
	17	17_32X0	45705703	45884925	5	179222
	17	17_33X0	46006665	46527748	19	521083
	17	17_34X0	54541267	54866058	2	324791
	17	17_35X0	68956102	69029801	2	73699
	17	17_36X0	69207684	69328865	3	121181
	17	17_37X0	70534821	70827478	3	292657
	17	17_3X0	4674162	4854128	2	179966
	17	17_4X0	19575876	19771592	5	195716
	17	17_5X0	20264481	20318785	2	54304
	17	17_6X0	21604067	21727772	2	123705
	17	17_7X0	22390429	22452699	3	62270
	17	17_8X0	22551309	22721751	3	170442
	17	17_9X0	25237852	25430061	2	192209
	18	18_1X0	5317542	5397100	2	79558
	18	18_2X0	15765671	16283764	11	518093
	18	18_3X0	45418568	45526501	3	107933
	18	18_4X0	48603965	48711604	2	107639
	18	18_5X0	49174409	49709529	7	535120
	18	18_6X0	62501879	62529131	2	27252
	18	18_7X0	65996234	66123266	2	127032
	18	18_8X0	72264815	72289105	2	24290
	19	19_10X0	14928160	15079766	2	151606
	19	19_11X0	18406678	18492020	2	85342
	19	19_12X0	19716180	19735194	2	19014
	19	19_13X0	20275294	20383464	2	108170
	19	19_14X0	21625056	21749288	5	124232
	19	19_15X0	23778433	23829631	2	51198
	19	19_16X0	24667282	24688691	2	21409
	19	19_17X0	25576379	25714342	2	137963
	19	19_18X0	26397163	26462210	2	65047
	19	19_19X0	26606974	26711541	2	104567

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	19	19_1X0	2718883	2788076	2	69193
	19	19_20X0	32743124	32760798	2	17674
	19	19_21X0	33223380	33235016	2	11636
	19	19_2X0	2981757	3373070	20	391313
	19	19_3X0	3461735	3485863	2	24128
	19	19_4X0	3587711	3768500	15	180789
	19	19_4X1	3877187	4092479	6	215292
	19	19_5X0	7000793	7077848	2	77055
	19	19_6X0	7087145	7104316	2	17171
	19	19_7X0	7325057	7399221	2	74164
	19	19_8X0	11567528	11614168	2	46640
	19	19_9X0	13276913	13387453	3	110540
	2	2_10X0	33750108	33763172	2	13064
	2	2_11X0	35287063	35885391	13	598328
	2	2_12X0	36207901	36490190	2	282289
	2	2_13X0	37379651	37633307	5	253656
	2	2_14X0	50380003	50502438	2	122435
	2	2_15X0	52043746	52082029	2	38283
	2	2_16X0	52832181	53019336	5	187155
	2	2_17X0	54101063	54221973	2	120910
	2	2_18X0	54589437	55097223	7	507786
	2	2_19X0	58743922	58954790	4	210868
	2	2_1X0	143029	155292	2	12263
	2	2_20X0	59109287	59153458	2	44171
	2	2_21X0	61548881	61953481	16	404600
	2	2_22X0	62016228	62069376	2	53148
	2	2_23X0	65254285	65385918	2	131633
	2	2_24X0	80096955	80164248	2	67293
	2	2_25X0	85289689	85368507	2	78818
	2	2_26X0	85685932	85879005	2	193073
	2	2_27X0	86190399	86243795	3	53396
	2	2_28X0	88254996	88397620	5	142624
	2	2_29X0	89608625	89758425	2	149800
	2	2_2X0	657623	676109	2	18486
	2	2_3X0	1255203	1574689	2	319486
	2	2_4X0	7576483	7595602	2	19119
	2	2_5X0	17359487	17428657	2	69170
	2	2_6X0	18000473	18026506	2	26033
	2	2_7X0	18100770	18707188	12	606418
	2	2_8X0	25374755	25449958	2	75203
	2	2_9X0	33344808	33370366	2	25558
	20	20_10X0	6094610	6370125	5	275515
	20	20_11X0	6736327	6781445	2	45118
	20	20_12X0	6987905	7028557	2	40652
	20	20_13X0	7203378	7240183	2	36805
	20	20_14X0	7794884	7976819	2	181935
	20	20_15X0	8025074	8200960	2	175886
	20	20_16X0	8696241	8766830	3	70589
	20	20_17X0	10405138	10459478	2	54340
	20	20_18X0	10696549	11189545	8	492996
	20	20_19X0	12236145	12399497	2	163352
	20	20_1X0	2344502	2512372	5	167870
	20	20_20X0	12570343	12877977	3	307634
	20	20_22X0	14264136	14359290	2	95154
	20	20_23X0	16511609	16815651	3	304042

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	20	20_24X0	16953811	17208671	3	254860
	20	20_25X0	22184849	22356924	3	172075
	20	20_26X0	22870880	23027509	3	156629
	20	20_27X0	23799177	23890926	3	91749
	20	20_28X0	23933737	23999072	2	65335
	20	20_29X0	25153065	25460131	4	307066
	20	20_2X0	3126040	3463901	7	337861
	20	20_2X1	2665737	2984269	3	318532
	20	20_30X0	29803801	29845754	2	41953
	20	20_31X0	43752370	43905691	2	153321
	20	20_32X0	44056783	44386713	6	329930
	20	20_33X0	46786102	47006646	4	220544
	20	20_34X0	48306631	48426171	2	119540
	20	20_35X0	54539392	54586030	2	46638
	20	20_36X0	65714618	65938024	2	223406
	20	20_37X0	80047073	80214422	2	167349
	20	20_38X0	84647795	84835691	2	187896
	20	20_39X0	85481622	85505076	2	23454
	20	20_3X0	3606293	3729479	2	123186
	20	20_40X0	85802179	86018055	4	215876
	20	20_41X0	86370868	86412982	2	42114
	20	20_42X0	89444165	89625474	2	181309
	20	20_43X0	90758455	90770198	2	11743
	20	20_44X0	92374930	92474941	3	100011
	20	20_45X0	93161877	93360398	2	198521
	20	20_46X0	94823783	94873616	2	49833
	20	20_47X0	96801900	97056051	3	254151
	20	20_48X0	98092785	98130083	4	37298
	20	20_49X0	99381677	99601860	2	220183
	20	20_4X0	3759597	3828692	2	69095
	20	20_50X0	100275390	100405616	2	130226
	20	20_51X0	100499350	100618015	2	118665
	20	20_52X0	101081197	101295467	2	214270
	20	20_53X0	108502050	108568371	2	66321
	20	20_54X0	110002864	110020854	2	17990
	20	20_55X0	111772973	111837377	2	64404
	20	20_56X0	112375528	112635992	6	260464
	20	20_57X0	113234210	113299970	2	65760
	20	20_58X0	115176446	115520828	5	344382
	20	20_59X0	116241569	116276253	2	34684
	20	20_5X0	3848946	4148730	8	299784
	20	20_60X0	116801843	116836844	2	35001
	20	20_61X0	118833949	119018681	7	184732
	20	20_62X0	119070235	119152854	2	82619
	20	20_63X0	119932021	119945917	2	13896
	20	20_64X0	120116169	120144258	3	28089
	20	20_65X0	121823328	121921621	2	98293
	20	20_66X0	122683373	122748899	2	65526
	20	20_67X0	125143114	125167014	2	23900
	20	20_68X0	125325167	125496335	2	171168
	20	20_69X0	129877142	130038604	2	161462
	20	20_6X0	4635328	4777298	2	141970
	20	20_7X0	4873008	5095868	4	222860
	20	20_8X0	5142874	5510659	14	367785
	20	20_9X0	5572822	5730061	6	157239

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	21	21_10X0	55098261	55118113	2	19852
	21	21_11X0	55878526	55912015	2	33489
	21	21_12X0	58701578	58783351	2	81773
	21	21_13X0	76578665	76721139	2	142474
	21	21_14X0	76829443	77040725	2	211282
	21	21_15X0	85197537	85258627	2	61090
	21	21_16X0	91490582	91495125	2	4543
	21	21_17X0	92654174	92744715	2	90541
	21	21_18X0	97270993	97329596	2	58603
	21	21_19X0	98898373	99021350	4	122977
	21	21_1X0	2243635	2868696	4	625061
	21	21_1X1	1564412	2090936	3	526524
	21	21_20X0	103059991	103266363	4	206372
	21	21_21X0	108198147	108294078	2	95931
	21	21_22X0	110516303	110673519	4	157216
	21	21_23X0	110697544	110873336	2	175792
	21	21_24X0	110988631	111024576	2	35945
	21	21_25X0	111044530	111693873	48	649343
	21	21_26X0	111750945	111811847	2	60902
	21	21_27X0	111901400	111932041	2	30641
	21	21_28X0	112058368	112375680	4	317312
	21	21_29X0	112404776	112957981	13	553205
	21	21_2X0	12905280	12937891	2	32611
	21	21_30X0	117637075	118058029	7	420954
	21	21_30X1	118330789	118542147	3	211358
	21	21_31X0	118596947	118927211	6	330264
	21	21_32X0	118982603	118996471	2	13868
	21	21_33X0	119209675	119412054	4	202379
	21	21_3X0	20153875	20268556	6	114681
	21	21_4X0	27278788	27739116	3	460328
	21	21_5X0	31123397	31275551	12	152154
	21	21_6X0	39011258	39041146	2	29888
	21	21_7X0	41240863	41326663	2	85800
	21	21_8X0	52036401	52054985	2	18584
	21	21_9X0	53607541	53739875	3	132334
	22	22_10X0	11679585	11709454	3	29869
	22	22_11X0	12003994	12050903	2	46909
	22	22_12X0	13133783	13270856	6	137073
	22	22_13X0	13300799	13520345	4	219546
	22	22_14X0	14160680	14213603	3	52923
	22	22_15X0	45933785	45960164	2	26379
	22	22_16X0	46857569	47078844	2	221275
	22	22_17X0	51497238	51750586	2	253348
	22	22_18X0	55809901	55930454	2	120553
	22	22_19X0	58134973	58337124	3	202151
	22	22_1X0	1967936	2030755	2	62819
	22	22_20X0	67862783	68009761	2	146978
	22	22_21X0	71435925	71474328	2	38403
	22	22_22X0	82314491	82620915	5	306424
	22	22_2X0	4110354	4367759	3	257405
	22	22_3X0	5919544	6209700	8	290156
	22	22_4X0	7354489	7847353	8	492864
	22	22_5X0	8135046	8242958	4	107912
	22	22_6X0	10510423	10732104	2	221681
	22	22_7X0	11112953	11309980	3	197027

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	22	22_8X0	11343214	11507096	2	163882
	22	22_9X0	11545455	11665418	4	119963
	23	23_10X0	53834653	54047935	3	213282
	23	23_11X0	59423334	59500836	2	77502
	23	23_12X0	77589060	77841467	2	252407
	23	23_13X0	79097429	79453113	2	355684
	23	23_14X0	80407085	80828920	6	421835
	23	23_15X0	82617506	82825039	4	207533
	23	23_1X0	5704129	5965068	2	260939
	23	23_2X0	35124433	35223137	2	98704
	23	23_3X0	43307415	43331401	2	23986
	23	23_4X0	43422340	44137770	16	715430
	23	23_5X0	44139360	44277678	2	138318
	23	23_6X0	44481056	44590074	2	109018
	23	23_7X0	50365269	50977838	6	612569
	23	23_8X0	51087285	51447401	2	360116
	23	23_9X0	52664076	52899375	3	235299
	24	24_10X0	40780596	40796107	2	15511
	24	24_11X0	41690705	41740654	2	49949
	24	24_12X0	44909001	44971836	2	62835
	24	24_13X0	50788229	50834407	3	46178
	24	24_14X0	52124793	52342724	2	217931
	24	24_15X0	52497521	52785978	3	288457
	24	24_16X0	72058780	72415507	10	356727
	24	24_17X0	73548165	73570348	2	22183
	24	24_18X0	74118210	74146592	2	28382
	24	24_19X0	78250796	78292773	2	41977
	24	24_1X0	1122631	1132596	2	9965
	24	24_20X0	81073662	81113509	2	39847
	24	24_21X0	83490807	84916725	54	1425918
	24	24_2X0	1189063	1322462	4	133399
	24	24_3X0	1381483	1419852	2	38369
	24	24_4X0	1471725	1577798	4	106073
	24	24_5X0	13310846	13381185	2	70339
	24	24_6X0	26742644	26833870	2	91226
	24	24_7X0	27396911	27604295	2	207384
	24	24_8X0	29350749	29409575	2	58826
	24	24_9X0	37700513	37911612	3	211099
	25	25_10X0	22454382	22594836	4	140454
	25	25_11X0	22629729	22891920	3	262191
	25	25_12X0	23341490	23577695	2	236205
	25	25_13X0	23682834	23774451	2	91617
	25	25_14X0	23809352	24195085	3	385733
	25	25_15X0	24261651	24474635	2	212984
	25	25_16X0	25835906	26220776	6	384870
	25	25_17X0	27092717	27283350	2	190633
	25	25_18X0	27997709	28054120	2	56411
	25	25_19X0	32357993	32507705	4	149712
	25	25_1X0	1272427	1279200	2	6773
	25	25_20X0	46131827	46353557	2	221730
	25	25_21X0	46683783	46915475	3	231692
	25	25_22X0	46964056	46981195	2	17139
	25	25_23X0	49452289	49530723	2	78434
	25	25_24X0	55306282	55473175	2	166893
	25	25_25X0	57728614	57995613	4	266999

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	25	25_26X0	59322913	59462598	3	139685
	25	25_27X0	60464445	60506661	2	42216
	25	25_28X0	65608303	65735381	2	127078
	25	25_29X0	78990362	79083762	2	93400
	25	25_2X0	1315164	1335504	2	20340
	25	25_30X0	84031794	84346550	3	314756
	25	25_31X0	84465283	84955353	15	490070
	25	25_31X1	85017722	85525423	10	507701
	25	25_32X0	85688714	85699760	2	11046
	25	25_3X0	1666053	1788663	2	122610
	25	25_4X0	3830059	3884151	2	54092
	25	25_5X0	5846298	6021530	2	175232
	25	25_6X0	8717436	8782379	2	64943
	25	25_7X0	16834957	16922213	2	87256
	25	25_8X0	21655969	22245727	4	589758
	25	25_9X0	22389314	22447732	3	58418
	26	26_10X0	34329322	34546842	2	217520
	26	26_11X0	38038771	38109636	4	70865
	26	26_12X0	39701479	39794173	3	92694
	26	26_13X0	40836266	41157965	4	321699
	26	26_14X0	41429556	41501956	2	72400
	26	26_15X0	41666966	41703106	2	36140
	26	26_16X0	42737613	42855306	2	117693
	26	26_17X0	53676125	53746399	2	70274
	26	26_18X0	57939330	57980041	2	40711
	26	26_1X0	4701321	4776822	3	75501
	26	26_2X0	8454857	8628081	2	173224
	26	26_3X0	8628454	8861425	3	232971
	26	26_4X0	9314572	9358304	2	43732
	26	26_5X0	9772914	9926109	2	153195
	26	26_6X0	18158859	18252581	2	93722
	26	26_7X0	25791219	25916205	2	124986
	26	26_8X0	31471362	31624905	2	153543
	26	26_9X0	32786271	32866073	2	79802
	27	27_1X0	1361261	1424270	2	63009
	27	27_2X0	2024719	2190398	2	165679
	27	27_3X0	11459181	11514671	3	55490
	27	27_4X0	34248751	34277315	2	28564
	27	27_5X0	34376271	34525196	2	148925
	27	27_6X0	42763302	42829791	2	66489
	27	27_7X0	48419685	48518469	2	98784
	28	28_10X0	13538395	13689447	5	151052
	28	28_11X0	15826717	15910594	4	83877
	28	28_12X0	16195607	16206819	2	11212
	28	28_13X0	20492188	20523900	2	31712
	28	28_1X0	1620359	1698311	2	77952
	28	28_2X0	1730062	1864827	2	134765
	28	28_3X0	2189520	2384341	2	194821
	28	28_4X0	8363057	8601921	2	238864
	28	28_5X0	9242909	9303108	2	60199
	28	28_6X0	10036337	10069695	2	33358
	28	28_7X0	10646228	10680182	2	33954
	28	28_8X0	13000838	13056332	3	55494
	28	28_9X0	13216548	13519507	5	302959
	29	29_10X0	23753309	23799771	2	46462

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	29	29_11X0	24016071	24185634	2	169563
	29	29_1X0	3553618	3769890	2	216272
	29	29_2X0	7165964	7196387	2	30423
	29	29_3X0	7394265	7439666	2	45401
	29	29_4X0	8286828	8315678	2	28850
	29	29_5X0	20230361	20774009	14	543648
	29	29_6X0	20969559	20984931	4	15372
	29	29_7X0	21033443	21532139	13	498696
	29	29_8X0	22027035	22922519	46	895484
	29	29_9X0	23390401	23412459	2	22058
	3	3_10X0	81241412	81456247	2	214835
	3	3_11X0	91110380	91171042	3	60662
	3	3_1X0	782251	848286	3	66035
	3	3_2X0	6923832	7087285	3	163453
	3	3_3X0	9779925	9847431	2	67506
	3	3_4X0	11159342	11285607	2	126265
	3	3_5X0	19389455	19453679	2	64224
	3	3_6X0	26333023	26624133	2	291110
	3	3_7X0	29951153	30045216	2	94063
	3	3_8X0	77911915	77954297	2	42382
	3	3_9X0	78615177	78636548	2	21371
	4	4_10X0	70886113	71110005	3	223892
	4	4_11X0	73289652	73352434	2	62782
	4	4_12X0	90600690	90863268	3	262578
	4	4_1X0	1033125	1258866	3	225741
	4	4_2X0	14300065	14389987	2	89922
	4	4_3X0	17127782	17154071	2	26289
	4	4_4X0	34795751	34912475	2	116724
	4	4_5X0	49082787	49380051	2	297264
	4	4_6X0	51270612	51348249	2	77637
	4	4_7X0	52101410	52214584	2	113174
	4	4_8X0	65506474	65644787	2	138313
	4	4_9X0	66557778	66699405	2	141627
	5	5_10X0	18596144	18638719	2	42575
	5	5_11X0	18698330	18929729	4	231399
	5	5_12X0	20891416	21107824	2	216408
	5	5_13X0	21396204	21431962	2	35758
	5	5_14X0	25046701	25153703	2	107002
	5	5_15X0	26536532	26688389	2	151857
	5	5_16X0	26842746	27298904	8	456158
	5	5_16X1	27466162	27752954	4	286792
	5	5_17X0	27771384	27839907	3	68523
	5	5_18X0	27957324	28086838	2	129514
	5	5_19X0	28517720	28845973	2	328253
	5	5_1X0	108804	134681	5	25877
	5	5_20X0	33883290	34032184	2	148894
	5	5_21X0	41514205	41788465	4	274260
	5	5_22X0	42478706	42638611	15	159905
	5	5_23X0	43297897	43348599	3	50702
	5	5_24X0	43495843	43620471	3	124628
	5	5_25X0	52244909	52393532	5	148623
	5	5_26X0	52634602	52670788	2	36186
	5	5_27X0	53443037	53521566	2	78529
	5	5_28X0	54578928	54609465	2	30537
	5	5_29X0	55205338	55228612	2	23274

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	5	5_2X0	214964	290575	2	75611
	5	5_30X0	58578146	58775339	2	197193
	5	5_31X0	59195753	59345108	2	149355
	5	5_32X0	59440951	59468890	3	27939
	5	5_33X0	60229389	60295452	3	66063
	5	5_34X0	60685372	60705571	2	20199
	5	5_35X0	71898786	71965002	3	66216
	5	5_36X0	73864461	74065999	2	201538
	5	5_3X0	511660	610359	4	98699
	5	5_4X0	1106239	1154126	4	47887
	5	5_5X0	1864680	2105739	4	241059
	5	5_6X0	2494390	2532107	3	37717
	5	5_7X0	2756747	2759364	2	2617
	5	5_8X0	2813268	3144023	11	330755
	5	5_9X0	14972646	15242811	2	270165
	6	6_1X0	576840	613498	3	36658
	6	6_2X0	43867596	44139733	16	272137
	6	6_2X1	49917351	50033632	7	116281
	6	6_2X10	50565540	50691326	5	125786
	6	6_2X100	43396029	43428297	2	32268
	6	6_2X101	36708311	36798288	2	89977
	6	6_2X102	41235094	41272077	2	36983
	6	6_2X104	10527281	10630923	2	103642
	6	6_2X105	45539904	45652774	2	112870
	6	6_2X11	14126086	14263494	4	137408
	6	6_2X12	46043328	46298909	7	255581
	6	6_2X13	46605245	46777298	6	172053
	6	6_2X14	11199304	11242024	2	42720
	6	6_2X15	50706724	50846241	5	139517
	6	6_2X16	48935957	49164925	8	228968
	6	6_2X17	37500752	37665791	5	165039
	6	6_2X18	44783533	44971487	7	187954
	6	6_2X19	5463660	5560141	6	96481
	6	6_2X2	49757454	49910500	7	153046
	6	6_2X20	48725956	48833798	5	107842
	6	6_2X21	34132832	34286645	6	153813
	6	6_2X22	47206172	47453199	8	247027
	6	6_2X23	2607592	2715272	5	107680
	6	6_2X24	48298005	48415552	5	117547
	6	6_2X25	45804981	45910505	4	105524
	6	6_2X26	11060810	11179237	4	118427
	6	6_2X27	37710952	37808975	4	98023
	6	6_2X28	31770639	32035860	5	265221
	6	6_2X29	35314764	35520745	6	205981
	6	6_2X3	46887065	47076855	8	189790
	6	6_2X30	30736966	30862658	5	125692
	6	6_2X31	44690163	44762648	3	72485
	6	6_2X32	8527129	8641252	4	114123
	6	6_2X33	37369280	37528939	5	159659
	6	6_2X34	31599643	31750252	4	150609
	6	6_2X35	13454313	13533933	5	79620
	6	6_2X36	50303866	50437236	5	133370
	6	6_2X37	32173099	32425668	7	252569
	6	6_2X38	8341934	8494361	5	152427
	6	6_2X39	33842635	33903894	5	61259

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	6	6_2X4	48116111	48235938	8	119827
	6	6_2X40	7175032	7258421	5	83389
	6	6_2X41	30227442	30378419	5	150977
	6	6_2X42	45936423	46026825	3	90402
	6	6_2X43	45442191	45514681	3	72490
	6	6_2X44	50512152	50554791	2	42639
	6	6_2X45	10677518	10812080	4	134562
	6	6_2X46	31369241	31412073	2	42832
	6	6_2X47	11283323	11384361	4	101038
	6	6_2X48	30052066	30104530	4	52464
	6	6_2X49	32064611	32119775	2	55164
	6	6_2X5	14267401	14387104	6	119703
	6	6_2X50	37837799	37893327	2	55528
	6	6_2X51	35907670	36040856	4	133186
	6	6_2X52	36900795	36980995	3	80200
	6	6_2X53	45057672	45197170	4	139498
	6	6_2X54	37050665	37191754	4	141089
	6	6_2X55	43091818	43206758	4	114940
	6	6_2X56	38127434	38187342	3	59908
	6	6_2X57	48441029	48577623	3	136594
	6	6_2X58	44568501	44642207	3	73706
	6	6_2X59	41930398	42001841	3	71443
	6	6_2X6	44156925	44336295	7	179370
	6	6_2X60	31455723	31584336	3	128613
	6	6_2X61	34521297	34585127	3	63830
	6	6_2X62	9322550	9348284	3	25734
	6	6_2X63	2751525	2932138	4	180613
	6	6_2X64	37933500	38059378	3	125878
	6	6_2X65	48600935	48708934	2	107999
	6	6_2X66	18699200	18773499	2	74299
	6	6_2X67	33403952	33436904	3	32952
	6	6_2X68	21193762	21301974	3	108212
	6	6_2X69	8695074	8856679	4	161605
	6	6_2X7	47457094	47624160	7	167066
	6	6_2X70	18417812	18661700	3	243888
	6	6_2X71	45279250	45382338	3	103088
	6	6_2X72	39493609	39639623	3	146014
	6	6_2X73	44367624	44479651	3	112027
	6	6_2X74	31207357	31332205	3	124848
	6	6_2X75	49559862	49733627	4	173765
	6	6_2X76	31017409	31138693	3	121284
	6	6_2X77	10949715	10964498	2	14783
	6	6_2X78	6031366	6063842	2	32476
	6	6_2X79	18241912	18347718	2	105806
	6	6_2X8	50163707	50271946	5	108239
	6	6_2X80	5244000	5328511	2	84511
	6	6_2X81	40633448	40672069	2	38621
	6	6_2X82	42247343	42273330	2	25987
	6	6_2X83	18135054	18164749	2	29695
	6	6_2X84	29652797	29690489	2	37692
	6	6_2X85	17686410	17727715	2	41305
	6	6_2X86	33555240	33629223	2	73983
	6	6_2X87	20482362	20580589	2	98227
	6	6_2X88	13966702	14023148	2	56446
	6	6_2X89	34879118	34894949	2	15831

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	6	6_2X9	50040160	50155300	4	115140
	6	6_2X90	19569082	19629281	2	60199
	6	6_2X91	38288537	38361742	2	73205
	6	6_2X92	34639506	34720781	2	81275
	6	6_2X93	29860068	29887779	2	27711
	6	6_2X94	4229480	4261126	2	31646
	6	6_2X95	28343907	28421452	2	77545
	6	6_2X96	38828375	38853804	2	25429
	6	6_2X97	49249280	49313455	2	64175
	6	6_2X98	47726002	47752204	2	26202
	6	6_2X99	42979674	43017879	2	38205
	7	7_10X0	24507130	24542832	3	35702
	7	7_11X0	30731175	30810720	2	79545
	7	7_12X0	35682130	35777229	2	95099
	7	7_13X0	36739566	37066994	3	327428
	7	7_14X0	47190294	47590593	5	400299
	7	7_15X0	51011592	51183015	2	171423
	7	7_16X0	75097492	75162750	2	65258
	7	7_17X0	101106483	101157111	3	50628
	7	7_18X0	102238477	102381251	2	142774
	7	7_19X0	109774918	109801536	2	26618
	7	7_1X0	88633	428292	3	339659
	7	7_20X0	111371014	111398159	2	27145
	7	7_21X0	129294170	129297035	2	2865
	7	7_22X0	132144420	132208155	2	63735
	7	7_23X0	133960679	134009241	2	48562
	7	7_2X0	486135	1145673	5	659538
	7	7_3X0	17971037	18123659	2	152622
	7	7_4X0	18203827	18330741	2	126914
	7	7_5X0	18415786	18445142	2	29356
	7	7_6X0	18849839	18925760	3	75921
	7	7_7X0	21876132	21972739	3	96607
	7	7_8X0	22202527	22633914	12	431387
	7	7_9X0	22844202	23308287	4	464085
	8	8_10X0	49049238	49362947	2	313709
	8	8_11X0	51919808	52151274	2	231466
	8	8_12X0	76697007	76949291	5	252284
	8	8_13X0	80617898	80848707	4	230809
	8	8_14X0	136926456	137440909	13	514453
	8	8_15X0	137471524	137905923	4	434399
	8	8_16X0	137932364	138058553	2	126189
	8	8_17X0	138319103	138503661	2	184558
	8	8_18X0	138547817	139173282	12	625465
	8	8_1X0	6643556	7166128	19	522572
	8	8_2X0	16314446	16494621	2	180175
	8	8_3X0	21102267	21351673	3	249406
	8	8_4X0	21791422	21825801	2	34379
	8	8_5X0	22430579	22633039	3	202460
	8	8_6X0	23039282	23083222	2	43940
	8	8_7X0	37017545	37809910	6	792365
	8	8_8X0	37930562	38027632	2	97070
	8	8_9X0	40706666	40766236	2	59570
	9	9_10X0	47588923	47608805	2	19882
	9	9_11X0	51811052	51914021	3	102969
	9	9_12X0	56262099	56335764	3	73665

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	9	9_13X0	61917440	62099221	2	181781
	9	9_14X0	62111648	62424006	3	312358
	9	9_15X0	82035967	82242574	5	206607
	9	9_16X0	82711612	82805599	5	93987
	9	9_17X0	86311950	86328775	2	16825
	9	9_18X0	87938712	88183458	2	244746
	9	9_19X0	90622613	90745716	2	123103
	9	9_1X0	981926	1003963	2	22037
	9	9_20X0	94149126	94250545	2	101419
	9	9_21X0	96493488	96527841	3	34353
	9	9_22X0	97305010	97349422	2	44412
	9	9_23X0	109238862	109453689	4	214827
	9	9_24X0	115311395	115553178	2	241783
	9	9_25X0	115556589	115595766	2	39177
	9	9_26X0	115807486	115821508	2	14022
	9	9_2X0	5013032	5423972	8	410940
	9	9_3X0	5609912	5635905	2	25993
	9	9_4X0	5751299	5997657	2	246358
	9	9_5X0	7617999	7794933	2	176934
	9	9_6X0	37806893	38113520	5	306627
	9	9_7X0	39135578	39302669	4	167091
	9	9_8X0	40971653	41015776	2	44123
	9	9_9X0	44067469	44092484	2	25015
	KE148049.1	KE148049.1_1X0	364420	389662	2	25242
	KE148049.1	KE148049.1_2X0	389997	390939	2	942
	KE148049.1	KE148049.1_3X0	1509879	1550652	2	40773
	KE148108.1	KE148108.1_1X0	880125	941359	2	61234
	KE148108.1	KE148108.1_2X0	1512691	1560669	2	47978
	X	X_10X0	45281336	45571891	9	290555
	X	X_11X0	46267358	46299773	2	32415
	X	X_12X0	46415031	46678837	4	263806
	X	X_13X0	48082447	48269666	2	187219
	X	X_14X0	48365285	48422046	2	56761
	X	X_15X0	48621533	48832931	2	211398
	X	X_16X0	48994883	49096937	3	102054
	X	X_17X0	50898324	51013942	2	115618
	X	X_18X0	51328223	51346037	2	17814
	X	X_19X0	52780705	52929005	2	148300
	X	X_1X0	837104	1018013	2	180909
	X	X_20X0	59863866	60082546	3	218680
	X	X_21X0	61474680	61555480	2	80800
	X	X_22X0	67764914	68180602	3	415688
	X	X_23X0	89870253	89974273	4	104020
	X	X_24X0	90299327	90313878	2	14551
	X	X_25X0	90584653	90734092	3	149439
	X	X_26X0	90984686	91525594	10	540908
	X	X_27X0	91894198	91966688	3	72490
	X	X_28X0	105660614	105693641	2	33027
	X	X_29X0	110422761	110488319	2	65558
	X	X_2X0	1108267	1314997	4	206730
	X	X_30X0	110522126	110583996	2	61870
	X	X_31X0	116797877	117124879	4	327002
	X	X_32X0	118400889	118589347	2	188458
	X	X_33X0	124264066	124470984	3	206918
	X	X_34X0	125174289	125354377	2	180088

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	X	X_35X0	127243534	127662410	5	418876
	X	X_36X0	127769179	127859858	2	90679
	X	X_37X0	127917536	127941164	2	23628
	X	X_38X0	128099682	128583432	5	483750
	X	X_39X0	128725454	128827287	2	101833
	X	X_3X0	11367110	11404029	2	36919
	X	X_40X0	129121944	129130892	2	8948
	X	X_4X0	13708259	13783846	2	75587
	X	X_5X0	24612338	24661850	2	49512
	X	X_6X0	28622915	28654000	4	31085
	X	X_7X0	38664042	38931854	3	267812
	X	X_8X0	40834159	41054641	2	220482
	X	X_9X0	45068146	45229653	3	161507
	X	X_9X1	44586388	44632856	2	46468
	1	1_10X0	33814996	34174512	4	359516
	1	1_11X0	34382342	34648231	4	265889
	1	1_12X0	35909927	35926783	2	16856
	1	1_13X0	38900183	38997995	3	97812
	1	1_14X0	39498462	39960976	7	462514
	1	1_15X0	48188588	48293987	2	105399
	1	1_16X0	57167856	57439533	3	271677
	1	1_17X0	62535401	62606059	2	70658
	1	1_18X0	69201657	69225318	2	23661
	1	1_19X0	70495340	70656636	3	161296
	1	1_1X0	171786	367589	8	195803
	1	1_20X0	71391094	71642584	3	251490
	1	1_21X0	80229112	80324827	5	95715
	1	1_22X0	82865162	82890774	2	25612
	1	1_23X0	85736310	85756295	2	19985
	1	1_24X0	88922677	89000284	3	77607
	1	1_25X0	92397477	92609573	4	212096
	1	1_26X0	94539643	94563167	2	23524
	1	1_27X0	110153086	110262807	2	109721
	1	1_28X0	113934089	114014970	2	80881
	1	1_29X0	116200062	116252509	3	52447
	1	1_2X0	9681191	9695057	2	13866
	1	1_30X0	118791350	118908967	2	117617
	1	1_31X0	118911236	119119328	5	208092
	1	1_32X0	119506525	119544940	2	38415
	1	1_33X0	119849077	119973481	2	124404
	1	1_34X0	123884172	123913284	2	29112
	1	1_35X0	138610187	138731050	2	120863
	1	1_36X0	139717547	139786052	2	68505
	1	1_37X0	144263880	144328545	4	64665
	1	1_38X0	145647731	145733123	3	85392
	1	1_39X0	146626190	146941954	4	315764
	1	1_3X0	9888824	9959038	2	70214
	1	1_40X0	147182214	147231499	2	49285
	1	1_41X0	148242346	148343175	2	100829
	1	1_42X0	155719515	156205474	19	485959
	1	1_42X1	156736323	157063408	15	327085
	1	1_42X2	155245477	155640245	17	394768
	1	1_42X3	154744595	155214568	17	469973
	1	1_42X4	156280102	156725179	17	445077
	1	1_43X0	157264136	157412002	4	147866

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_44X0	157443268	157908962	9	465694
	1	1_45X0	159891517	160771968	32	880451
	1	1_45X1	158361968	158797075	22	435107
	1	1_45X2	158825183	159126310	11	301127
	1	1_45X3	159155661	159402461	13	246800
	1	1_45X4	159472712	159882028	14	409316
	1	1_46X0	160882311	160938397	4	56086
	1	1_47X0	161212225	161223006	2	10781
	1	1_48X0	161476365	161527244	2	50879
	1	1_49X0	161674601	161975043	3	300442
	1	1_4X0	9990756	10217825	3	227069
	1	1_50X0	162045426	162053202	2	7776
	1	1_51X0	162098055	162228522	2	130467
	1	1_52X0	162300084	162331242	2	31158
	1	1_53X0	162711436	163203258	9	491822
	1	1_53X1	162420224	162699085	5	278861
	1	1_54X0	172980142	173037957	2	57815
	1	1_55X0	180152005	180200369	2	48364
	1	1_56X0	183724796	183777520	2	52724
	1	1_57X0	184185846	184314578	2	128732
	1	1_58X0	185704994	185755011	2	50017
	1	1_5X0	10228660	10310340	4	81680
	1	1_6X0	15482866	15636794	4	153928
	1	1_7X0	26079852	26116965	2	37113
	1	1_8X0	26857483	26886829	3	29346
	1	1_9X0	28713095	28785092	2	71997
	10	10_10X0	9751450	9801460	3	50010
	10	10_11X0	9901269	10089412	3	188143
	10	10_12X0	10121912	10173436	3	51524
	10	10_13X0	10561597	10607424	2	45827
	10	10_14X0	10718535	10825623	2	107088
	10	10_15X0	11119008	11133587	2	14579
	10	10_16X0	11324800	11463585	4	138785
	10	10_16X1	11581227	11802423	4	221196
	10	10_17X0	12689854	13008250	4	318396
	10	10_17X1	13044461	13297099	3	252638
	10	10_18X0	13546960	13668047	2	121087
	10	10_19X0	14122675	14240249	4	117574
	10	10_1X0	5687725	5844139	3	156414
	10	10_20X0	14324494	14397253	3	72759
	10	10_21X0	14484913	14746342	8	261429
	10	10_22X0	14807576	15232553	11	424977
	10	10_23X0	15325897	15476758	5	150861
	10	10_24X0	16138750	16254828	2	116078
	10	10_25X0	16445800	16583422	4	137622
	10	10_26X0	17408742	17423267	2	14525
	10	10_27X0	17779191	18054135	4	274944
	10	10_28X0	18071866	18110291	2	38425
	10	10_29X0	18269934	18306134	3	36200
	10	10_2X0	5973976	6004849	2	30873
	10	10_30X0	18702940	18765702	4	62762
	10	10_31X0	19126132	19162578	2	36446
	10	10_32X0	19737712	19756461	2	18749
	10	10_33X0	19992502	20056083	2	63581
	10	10_34X0	20256588	20520701	13	264113

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	10	10_34X1	20484423	20920489	12	436066
	10	10_34X2	20959230	21337161	8	377931
	10	10_35X0	21413871	21914919	16	501048
	10	10_35X1	22018035	22712666	17	694631
	10	10_36X0	22957780	23035974	2	78194
	10	10_37X0	23271422	23500381	5	228959
	10	10_38X0	23629243	23894462	5	265219
	10	10_39X0	23933632	24265756	9	332124
	10	10_3X0	6118948	6143853	2	24905
	10	10_40X0	24355535	24373121	2	17586
	10	10_41X0	24471617	24670422	5	198805
	10	10_42X0	24700515	24746232	4	45717
	10	10_43X0	24807141	24949082	4	141941
	10	10_44X0	24962529	25065391	3	102862
	10	10_45X0	25185543	25558442	9	372899
	10	10_46X0	25586551	25710319	9	123768
	10	10_47X0	25730254	25793033	2	62779
	10	10_48X0	27135450	27333069	4	197619
	10	10_48X1	27926863	28177463	6	250600
	10	10_48X2	26078963	26492472	8	413509
	10	10_48X3	27698089	27892824	5	194735
	10	10_48X4	27396763	27615342	5	218579
	10	10_48X5	26855824	27109226	5	253402
	10	10_48X6	26548788	26832268	3	283480
	10	10_49X0	28827298	28884741	2	57443
	10	10_4X0	6289312	6329854	3	40542
	10	10_50X0	41629332	41729551	2	100219
	10	10_51X0	59240929	59364975	3	124046
	10	10_52X0	60557764	60639255	2	81491
	10	10_53X0	64808972	64821348	2	12376
	10	10_54X0	64980076	65058211	3	78135
	10	10_55X0	65103492	65242920	3	139428
	10	10_56X0	74307724	74384294	2	76570
	10	10_57X0	78661274	78846196	11	184922
	10	10_58X0	78874830	78932322	2	57492
	10	10_59X0	82434211	82566499	3	132288
	10	10_5X0	6445506	6494352	2	48846
	10	10_6X0	6503587	6562124	2	58537
	10	10_7X0	7020275	7171296	5	151021
	10	10_8X0	7776219	8006217	5	229998
	10	10_8X1	8302182	8631441	7	329259
	10	10_8X2	7538476	7753460	4	214984
	10	10_8X3	7232727	7516933	6	284206
	10	10_8X4	8011911	8264185	5	252274
	10	10_8X5	8690022	9020254	5	330232
	10	10_9X0	9043021	9096109	2	53088
	11	11_10X0	15078834	15181535	2	102701
	11	11_11X0	15646282	15727932	2	81650
	11	11_12X0	15812147	15849932	2	37785
	11	11_13X0	18057259	18146410	2	89151
	11	11_14X0	18561823	18575592	2	13769
	11	11_15X0	19501957	19511565	2	9608
	11	11_16X0	19529053	19618185	2	89132
	11	11_17X0	20226994	20254973	3	27979
	11	11_18X0	20385008	20422020	2	37012

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	11	11_19X0	20610215	20726070	3	115855
	11	11_1X0	486060	509047	2	22987
	11	11_20X0	20782867	20922462	4	139595
	11	11_21X0	20974011	20989849	2	15838
	11	11_22X0	21153831	21388420	15	234589
	11	11_23X0	21442194	21495779	2	53585
	11	11_24X0	21594713	21619271	6	24558
	11	11_25X0	21665928	21978588	13	312660
	11	11_26X0	22402568	22421804	2	19236
	11	11_27X0	22489533	22546131	2	56598
	11	11_28X0	24021983	24088074	2	66091
	11	11_29X0	24581790	24743227	9	161437
	11	11_2X0	3100454	3117805	2	17351
	11	11_30X0	25846533	25880472	2	33939
	11	11_31X0	26077679	26135367	2	57688
	11	11_32X0	26164813	26193215	2	28402
	11	11_33X0	26331318	26384543	2	53225
	11	11_34X0	31308080	31368384	2	60304
	11	11_35X0	32411678	32446134	3	34456
	11	11_36X0	32458346	32543784	3	85438
	11	11_37X0	34635880	34710859	2	74979
	11	11_38X0	37197929	37345300	6	147371
	11	11_39X0	37593800	37760847	4	167047
	11	11_3X0	4384048	4409682	2	25634
	11	11_40X0	38505372	38594301	3	88929
	11	11_41X0	41572047	41665036	3	92989
	11	11_42X0	42554741	42750103	4	195362
	11	11_43X0	45398478	45421641	2	23163
	11	11_44X0	46844902	47054375	11	209473
	11	11_44X1	46515370	46829915	12	314545
	11	11_45X0	47108438	47166952	2	58514
	11	11_46X0	47869036	47954974	2	85938
	11	11_47X0	49427292	49457246	2	29954
	11	11_48X0	50061536	50067419	2	5883
	11	11_49X0	50104227	50203849	4	99622
	11	11_4X0	5558246	5619093	2	60847
	11	11_50X0	50403867	50409273	2	5406
	11	11_51X0	50941750	51004840	3	63090
	11	11_52X0	51609145	51741179	2	132034
	11	11_53X0	52794922	53118147	5	323225
	11	11_54X0	57517336	57734761	4	217425
	11	11_55X0	58772116	58850188	3	78072
	11	11_55X1	58885760	59012179	2	126419
	11	11_5X0	6171628	6186523	2	14895
	11	11_6X0	6916663	6935551	2	18888
	11	11_7X0	7133017	7362725	6	229708
	11	11_8X0	10866361	10903989	2	37628
	11	11_9X0	11536750	11767408	3	230658
	11	11_9X1	11789275	11942751	2	153476
	12	12_10X0	22106465	22252101	5	145636
	12	12_11X0	22584866	22738323	4	153457
	12	12_12X0	23172031	23298145	3	126114
	12	12_13X0	23489077	23597667	2	108590
	12	12_14X0	23720139	24009444	5	289305
	12	12_15X0	24963310	25007976	2	44666

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	12	12_16X0	25676638	25761547	2	84909
	12	12_17X0	26588658	26622089	2	33431
	12	12_18X0	27060312	27184398	2	124086
	12	12_19X0	27201205	27254212	2	53007
	12	12_1X0	8924682	8954869	2	30187
	12	12_20X0	27260357	27294513	2	34156
	12	12_21X0	27344032	27412265	2	68233
	12	12_22X0	28690211	28849553	3	159342
	12	12_23X0	30633633	30661665	2	28032
	12	12_24X0	30763486	30826955	2	63469
	12	12_25X0	31933166	31965740	2	32574
	12	12_26X0	32956829	32985269	4	28440
	12	12_2X0	15713281	15951780	12	238499
	12	12_2X1	13251410	13524613	10	273203
	12	12_2X10	18206495	18350547	4	144052
	12	12_2X11	12963245	13206778	7	243533
	12	12_2X12	14647818	14871912	8	224094
	12	12_2X13	17808211	18008300	6	200089
	12	12_2X14	19013823	19365715	10	351892
	12	12_2X15	15454415	15673769	6	219354
	12	12_2X16	13801478	14009095	5	207617
	12	12_2X17	16072601	16375498	6	302897
	12	12_2X18	16524741	16845441	5	320700
	12	12_2X19	17321481	17557129	4	235648
	12	12_2X2	12367499	12756614	11	389115
	12	12_2X20	15216596	15397865	4	181269
	12	12_2X3	12762955	12936767	8	173812
	12	12_2X4	18025703	18186013	7	160310
	12	12_2X5	14168110	14540233	11	372123
	12	12_2X6	18381455	18509953	7	128498
	12	12_2X7	13566468	13718578	7	152110
	12	12_2X8	19485206	19714077	9	228871
	12	12_2X9	18614128	18816340	9	202212
	12	12_3X0	19790776	19821168	2	30392
	12	12_4X0	19932778	20106576	5	173798
	12	12_5X0	20134061	20377023	6	242962
	12	12_6X0	20526491	20583744	3	57253
	12	12_7X0	20748362	20843908	2	95546
	12	12_8X0	20892575	21266254	9	373679
	12	12_9X0	21722550	21872504	3	149954
	13	13_10X0	10728292	10749290	2	20998
	13	13_11X0	11305602	11402983	2	97381
	13	13_12X0	11664189	11749229	2	85040
	13	13_13X0	18476085	18591972	3	115887
	13	13_14X0	18678936	18726900	3	47964
	13	13_15X0	18727672	18745158	2	17486
	13	13_16X0	18758124	18791366	2	33242
	13	13_17X0	18869481	18877190	2	7709
	13	13_18X0	18968572	19170856	5	202284
	13	13_18X1	19255114	19419183	2	164069
	13	13_19X0	19550152	19697585	2	147433
	13	13_1X0	220152	260756	2	40604
	13	13_20X0	20725748	20811832	2	86084
	13	13_21X0	23124260	23154654	2	30394
	13	13_22X0	24883561	25195609	3	312048

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	13	13_23X0	27009483	27042804	2	33321
	13	13_24X0	27685267	27823960	2	138693
	13	13_25X0	29616043	29765350	2	149307
	13	13_26X0	39354533	39692102	13	337569
	13	13_27X0	39836937	39965259	3	128322
	13	13_28X0	40133354	40168055	2	34701
	13	13_29X0	40380844	40544772	2	163928
	13	13_2X0	2086922	2160263	4	73341
	13	13_30X0	40600871	40641344	2	40473
	13	13_31X0	40817360	40826731	2	9371
	13	13_32X0	41461557	41476068	2	14511
	13	13_33X0	41883769	41894384	2	10615
	13	13_34X0	41969344	42023029	2	53685
	13	13_35X0	42263771	42333694	4	69923
	13	13_36X0	42432800	42456189	2	23389
	13	13_3X0	5960510	5996766	3	36256
	13	13_4X0	6790713	6882342	4	91629
	13	13_5X0	6996565	7509191	7	512626
	13	13_6X0	7617868	7672541	2	54673
	13	13_7X0	7917367	8012374	3	95007
	13	13_8X0	8924938	9020418	2	95480
	13	13_9X0	10175447	10200055	2	24608
	14	14_10X0	35492138	36117354	18	625216
	14	14_11X0	36214470	36231271	2	16801
	14	14_12X0	43157130	43254600	2	97470
	14	14_13X0	51450957	51479862	2	28905
	14	14_14X0	51724785	51793182	2	68397
	14	14_15X0	58197614	58319155	2	121541
	14	14_16X0	71808888	72037550	3	228662
	14	14_17X0	86340792	86389187	2	48395
	14	14_18X0	88195790	88362393	3	166603
	14	14_19X0	91769416	91887290	2	117874
	14	14_1X0	130219	153790	2	23571
	14	14_20X0	92566005	92669710	2	103705
	14	14_21X0	92739545	92969632	8	230087
	14	14_22X0	93194580	93327388	4	132808
	14	14_23X0	93627806	93714239	2	86433
	14	14_2X0	1339186	1758352	5	419166
	14	14_3X0	3327906	3852292	8	524386
	14	14_4X0	4539203	4569558	2	30355
	14	14_5X0	22058284	22087546	2	29262
	14	14_6X0	26678565	26849125	3	170560
	14	14_7X0	27119141	27164836	2	45695
	14	14_8X0	27721132	27925728	3	204596
	14	14_9X0	35083166	35164668	2	81502
	15	15_10X0	17759330	17817376	2	58046
	15	15_11X0	17933462	18062586	2	129124
	15	15_12X0	18796493	18835902	2	39409
	15	15_13X0	18894401	18906282	2	11881
	15	15_14X0	24573402	24616068	3	42666
	15	15_15X0	28885798	28910821	2	25023
	15	15_16X0	29348808	29361671	2	12863
	15	15_17X0	31757501	31780383	2	22882
	15	15_18X0	33372243	33414151	3	41908
	15	15_19X0	54785442	54826184	2	40742

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	15	15_1X0	1518938	1714008	3	195070
	15	15_20X0	58633592	58806751	2	173159
	15	15_21X0	60376732	60562250	2	185518
	15	15_22X0	68826720	68946749	2	120029
	15	15_2X0	7202599	7642168	6	439569
	15	15_3X0	10024295	10136919	2	112624
	15	15_4X0	11758708	11769143	2	10435
	15	15_5X0	12775412	12839785	3	64373
	15	15_6X0	13230708	13256724	2	26016
	15	15_7X0	13353758	13662976	4	309218
	15	15_8X0	15701256	16054212	9	352956
	15	15_9X0	16359629	17119239	24	759610
	16	16_10X0	37279526	37319365	2	39839
	16	16_11X0	37355649	37429465	2	73816
	16	16_12X0	38709075	38787184	3	78109
	16	16_13X0	40021149	40083053	5	61904
	16	16_14X0	40348736	40751509	8	402773
	16	16_15X0	41737842	41979909	6	242067
	16	16_16X0	45088489	45132230	2	43741
	16	16_17X0	45996473	46077255	2	80782
	16	16_18X0	46916752	46970874	2	54122
	16	16_19X0	51689223	51797209	2	107986
	16	16_1X0	1727451	1839275	2	111824
	16	16_20X0	66256406	66706204	5	449798
	16	16_21X0	69298174	69415432	2	117258
	16	16_22X0	73947604	74029664	2	82060
	16	16_23X0	79601661	79770199	3	168538
	16	16_24X0	80484607	80509005	2	24398
	16	16_25X0	82062506	82218415	3	155909
	16	16_26X0	83671446	83795225	5	123779
	16	16_27X0	84099707	84201621	2	101914
	16	16_2X0	6940051	6962502	2	22451
	16	16_3X0	10473388	10615256	2	141868
	16	16_4X0	34940926	34991796	3	50870
	16	16_5X0	35683399	35693020	2	9621
	16	16_6X0	35753038	35832035	4	78997
	16	16_7X0	36950327	36966058	3	15731
	16	16_8X0	36977158	37083266	4	106108
	16	16_9X0	37115629	37234990	2	119361
	17	17_10X0	62922476	62948641	2	26165
	17	17_11X0	68395707	68536595	3	140888
	17	17_12X0	79888991	79936368	3	47377
	17	17_1X0	1038093	1122812	3	84719
	17	17_2X0	1830522	1873245	2	42723
	17	17_3X0	7516029	7672956	3	156927
	17	17_4X0	14967009	15031146	2	64137
	17	17_5X0	19133042	19199710	2	66668
	17	17_6X0	22407468	22419872	2	12404
	17	17_7X0	28846488	28913106	2	66618
	17	17_8X0	36348467	36442071	2	93604
	17	17_9X0	55653571	55750331	3	96760
	18	18_1X0	33400202	33440434	2	40232
	18	18_2X0	51258772	51275856	2	17084
	18	18_3X0	54540670	54647154	8	106484
	18	18_4X0	76188969	76329380	3	140411

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	18	18_5X0	78369777	78603605	3	233828
	19	19_10X0	35624674	35669124	2	44450
	19	19_11X0	36754148	36931354	3	177206
	19	19_12X0	44838120	44922356	2	84236
	19	19_13X0	47707821	47736371	2	28550
	19	19_14X0	56380318	56819084	13	438766
	19	19_15X0	58168261	58274747	2	106486
	19	19_1X0	2120160	2210127	2	89967
	19	19_2X0	10244433	10310009	3	65576
	19	19_3X0	21923494	22082190	2	158696
	19	19_4X0	22433760	22460296	2	26536
	19	19_5X0	22765571	22840016	2	74445
	19	19_6X0	24672174	24752928	4	80754
	19	19_7X0	24784339	24929063	3	144724
	19	19_8X0	27781751	27885521	2	103770
	19	19_9X0	31151181	31205878	3	54697
	2	2_10X0	24467444	24579065	3	111621
	2	2_11X0	30009867	30027396	2	17529
	2	2_12X0	30879376	30925749	2	46373
	2	2_13X0	32050237	32070218	3	19981
	2	2_14X0	33515526	33562790	2	47264
	2	2_15X0	34048368	34215662	6	167294
	2	2_16X0	34700133	34728290	2	28157
	2	2_17X0	36238925	36503259	5	264334
	2	2_18X0	37393127	37422963	2	29836
	2	2_19X0	37461298	37581374	2	120076
	2	2_1X0	11343274	11745360	9	402086
	2	2_20X0	39522328	39622081	2	99753
	2	2_21X0	39795584	39806895	2	11311
	2	2_22X0	39938409	39956907	2	18498
	2	2_23X0	41335752	41406626	2	70874
	2	2_24X0	44041698	44062299	2	20601
	2	2_25X0	47305618	47330572	2	24954
	2	2_26X0	52365001	52394973	2	29972
	2	2_27X0	53011408	53222398	3	210990
	2	2_28X0	53666192	53698815	2	32623
	2	2_29X0	54305193	54406754	3	101561
	2	2_2X0	15104801	15336181	7	231380
	2	2_30X0	58559677	58592798	2	33121
	2	2_31X0	60609118	60694575	2	85457
	2	2_32X0	69687661	69783995	3	96334
	2	2_33X0	71161734	71188687	2	26953
	2	2_34X0	78275329	78376812	2	101483
	2	2_35X0	79340078	79384295	3	44217
	2	2_3X0	16174377	16216990	2	42613
	2	2_4X0	17616117	17698404	3	82287
	2	2_5X0	18282527	18472757	3	190230
	2	2_6X0	21435171	21523022	2	87851
	2	2_7X0	22113495	22202820	2	89325
	2	2_8X0	22384701	22422483	4	37782
	2	2_9X0	24305643	24417976	2	112333
	20	20_10X0	25624045	25709907	2	85862
	20	20_11X0	26069023	26133406	14	64383
	20	20_12X0	26161356	26489190	14	327834
	20	20_13X0	26538750	26913171	13	374421

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	20	20_14X0	26981915	27024133	2	42218
	20	20_15X0	27836373	28241955	19	405582
	20	20_15X1	27461487	27826604	16	365117
	20	20_15X2	28303442	28730112	16	426670
	20	20_16X0	28779273	28951488	6	172215
	20	20_17X0	28994106	29462636	9	468530
	20	20_18X0	30176091	30433336	5	257245
	20	20_19X0	30522205	30835411	7	313206
	20	20_1X0	3298143	3443474	3	145331
	20	20_20X0	30893697	31330405	9	436708
	20	20_21X0	31337014	31429707	5	92693
	20	20_22X0	31436721	31450853	2	14132
	20	20_23X0	31476634	31481111	2	4477
	20	20_24X0	31550871	31566581	2	15710
	20	20_25X0	31652406	31671369	2	18963
	20	20_26X0	31828901	31895951	2	67050
	20	20_27X0	33339018	33557879	10	218861
	20	20_27X1	32596235	32791723	6	195488
	20	20_27X2	32804256	33198487	5	394231
	20	20_28X0	33737871	33849668	2	111797
	20	20_29X0	35799670	35820026	2	20356
	20	20_2X0	3635183	3694885	2	59702
	20	20_30X0	36991854	37002510	2	10656
	20	20_31X0	38732413	38853886	3	121473
	20	20_32X0	40551166	40718695	6	167529
	20	20_33X0	41398410	41415550	2	17140
	20	20_34X0	42089837	42269299	4	179462
	20	20_35X0	44738570	44760346	2	21776
	20	20_36X0	45409022	45569354	2	160332
	20	20_37X0	46087067	46133920	2	46853
	20	20_38X0	47664031	47871684	4	207653
	20	20_39X0	48613949	48732717	2	118768
	20	20_3X0	23673573	23849242	4	175669
	20	20_40X0	49863357	49909879	2	46522
	20	20_41X0	50347190	50613800	7	266610
	20	20_4X0	23928324	24171536	33	243212
	20	20_5X0	24552641	24631896	3	79255
	20	20_6X0	24738927	24842914	5	103987
	20	20_7X0	25014455	25148611	2	134156
	20	20_8X0	25223154	25237952	4	14798
	20	20_9X0	25518584	25521664	2	3080
	21	21_1X0	24483	389195	6	364712
	21	21_2X0	622122	660094	2	37972
	21	21_3X0	735918	1403836	12	667918
	21	21_4X0	4003592	4101057	2	97465
	21	21_5X0	16399502	16540847	2	141345
	21	21_6X0	17208490	17303723	2	95233
	21	21_7X0	29204050	29315110	3	111060
	21	21_8X0	45485999	45557611	2	71612
	22	22_10X0	23759109	23850908	4	91799
	22	22_11X0	25183607	25257354	2	73747
	22	22_12X0	26363890	26427974	2	64084
	22	22_13X0	28509821	28569166	2	59345
	22	22_14X0	34230815	34335813	4	104998
	22	22_15X0	34540836	34760986	7	220150

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	22	22_16X0	34828382	34846432	2	18050
	22	22_17X0	35366312	35375355	2	9043
	22	22_18X0	49278323	49304560	2	26237
	22	22_19X0	49781910	49792733	2	10823
	22	22_1X0	1877201	2078128	6	200927
	22	22_2X0	5040971	5095032	2	54061
	22	22_3X0	18380455	18401535	2	21080
	22	22_4X0	19691112	19701748	2	10636
	22	22_5X0	20231208	20347901	2	116693
	22	22_6X0	20676164	21097254	5	421090
	22	22_7X0	21854932	21975477	2	120545
	22	22_8X0	22136128	22270987	4	134859
	22	22_9X0	22440848	22562585	5	121737
	23	23_10X0	50484877	50582294	2	97417
	23	23_11X0	50605846	50769988	5	164142
	23	23_12X0	55235202	55352684	4	117482
	23	23_1X0	623659	736197	3	112538
	23	23_2X0	2765176	2846437	2	81261
	23	23_3X0	2880954	2972946	2	91992
	23	23_4X0	10468559	10504292	2	35733
	23	23_5X0	12858704	12928003	2	69299
	23	23_6X0	26780635	26873890	2	93255
	23	23_7X0	40099096	40389005	17	289909
	23	23_8X0	40990896	41029286	2	38390
	23	23_9X0	49792490	49842485	2	49995
	24	24_10X0	44677905	44703864	2	25959
	24	24_1X0	3921230	3984063	2	62833
	24	24_2X0	7906025	8085887	3	179862
	24	24_3X0	11297091	11337005	2	39914
	24	24_4X0	13946226	13993686	2	47460
	24	24_5X0	19292112	19351227	3	59115
	24	24_6X0	20677112	20902469	3	225357
	24	24_7X0	36942535	37291564	11	349029
	24	24_8X0	38694878	38719121	2	24243
	24	24_9X0	42214122	42242442	2	28320
	25	25_10X0	15434364	15512774	2	78410
	25	25_11X0	17646984	17699028	2	52044
	25	25_12X0	17847085	17910426	2	63341
	25	25_13X0	18097099	18193473	2	96374
	25	25_14X0	18206728	18253221	2	46493
	25	25_15X0	19128597	19144604	2	16007
	25	25_16X0	25109204	25195850	3	86646
	25	25_17X0	26698220	27034620	20	336400
	25	25_17X1	26482907	26661712	9	178805
	25	25_17X2	26070331	26462611	16	392280
	25	25_18X0	27136728	27153522	2	16794
	25	25_19X0	30393182	30451168	2	57986
	25	25_20X0	31271108	31299217	2	28109
	25	25_21X0	31315171	31335137	2	19966
	25	25_22X0	32679227	32695330	2	16103
	25	25_23X0	35406766	35517432	3	110666
	25	25_24X0	35629136	35753320	2	124184
	25	25_25X0	36801475	36825451	2	23976
	25	25_26X0	38177356	38523613	10	346257
	25	25_27X0	38629189	38678986	2	49797

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	25	25_2X0	1102823	1255388	3	152565
	25	25_3X0	3403237	3550703	2	147466
	25	25_4X0	3987144	4129819	2	142675
	25	25_5X0	4132818	4269257	3	136439
	25	25_6X0	7672053	7704874	2	32821
	25	25_7X0	10704281	10959714	9	255433
	25	25_7X1	10431381	10646422	7	215041
	25	25_8X0	10979092	11018161	2	39069
	25	25_9X0	14266532	14274149	2	7617
	26	26_10X0	38418197	38534299	2	116102
	26	26_11X0	39054925	39070073	2	15148
	26	26_12X0	39576125	39664636	2	88511
	26	26_13X0	39723149	39832504	6	109355
	26	26_14X0	41057463	41218017	2	160554
	26	26_1X0	24313408	24439348	2	125940
	26	26_2X0	27166569	27207225	2	40656
	26	26_3X0	27262301	27430211	4	167910
	26	26_4X0	27885646	27897027	2	11381
	26	26_5X0	29897051	30082749	4	185698
	26	26_6X0	30892092	30967695	2	75603
	26	26_7X0	32397606	32438938	2	41332
	26	26_8X0	37021987	37090380	2	68393
	26	26_9X0	37835367	37878249	3	42882
	27	27_1X0	2801126	2852221	2	51095
	27	27_2X0	5111652	5188915	2	77263
	27	27_3X0	6718445	6803569	2	85124
	27	27_4X0	13590423	13636512	2	46089
	27	27_5X0	22586694	22625927	2	39233
	27	27_6X0	24009616	24065887	2	56271
	27	27_7X0	26480660	26484458	2	3798
	27	27_8X0	33042424	33096410	2	53986
	27	27_9X0	33317878	33514252	7	196374
	28	28_10X0	39440090	39559781	2	119691
	28	28_11X0	40153584	40204842	2	51258
	28	28_12X0	40322519	40413530	2	91011
	28	28_13X0	45680975	45707939	2	26964
	28	28_14X0	45716635	45734492	2	17857
	28	28_1X0	350029	400314	2	50285
	28	28_2X0	3517946	3671706	3	153760
	28	28_3X0	8070376	8081927	2	11551
	28	28_4X0	16786570	16996770	4	210200
	28	28_5X0	34275830	34312069	2	36239
	28	28_6X0	34369221	34384493	2	15272
	28	28_7X0	35928335	35996231	4	67896
	28	28_8X0	37765149	37807188	2	42039
	28	28_9X0	38520641	38707111	7	186470
	29	29_1X0	79031	251974	3	172943
	29	29_2X0	501638	810135	3	308497
	29	29_3X0	2264381	2280938	2	16557
	29	29_4X0	21458270	21503315	2	45045
	29	29_5X0	26322180	26475885	2	153705
	29	29_6X0	28371632	28396858	2	25226
	29	29_7X0	28641073	29303208	13	662135
	3	3_10X0	21964965	22041392	2	76427
	3	3_11X0	22365168	22394822	2	29654

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	3	3_12X0	23269213	23322405	2	53192
	3	3_13X0	24110610	24133714	2	23104
	3	3_14X0	36259276	36555709	5	296433
	3	3_15X0	36933469	37059062	2	125593
	3	3_16X0	39569308	39720636	2	151328
	3	3_17X0	44959511	44984355	2	24844
	3	3_18X0	50358221	50432146	2	73925
	3	3_19X0	53962816	54024955	2	62139
	3	3_1X0	1393445	1541223	2	147778
	3	3_20X0	54619894	54696527	2	76633
	3	3_21X0	59105837	59216533	2	110696
	3	3_22X0	60077014	60132447	3	55433
	3	3_23X0	61688780	62104983	10	416203
	3	3_24X0	62121985	62235961	4	113976
	3	3_25X0	66605784	67117210	8	511426
	3	3_26X0	76415086	76436616	2	21530
	3	3_27X0	80157317	80202758	2	45441
	3	3_28X0	80719113	80863124	2	144011
	3	3_29X0	88384273	88442929	3	58656
	3	3_2X0	7911072	8383333	6	472261
	3	3_30X0	107579827	107690114	2	110287
	3	3_31X0	114517954	114589746	3	71792
	3	3_32X0	118098093	118172615	2	74522
	3	3_3X0	8978911	9045064	7	66153
	3	3_4X0	9610450	9652893	3	42443
	3	3_5X0	9773162	9878005	3	104843
	3	3_6X0	17158083	17241080	4	82997
	3	3_7X0	17469089	17495216	2	26127
	3	3_8X0	18297852	18311869	2	14017
	3	3_9X0	18864350	19011211	2	146861
	30	30_1X0	12689	72169	2	59480
	30	30_2X0	9294065	9448119	2	154054
	30	30_3X0	22072642	22092901	2	20259
	30	30_4X0	28282116	28310221	2	28105
	30	30_5X0	29622406	29751076	3	128670
	31	31_1X0	2380989	2629905	3	248916
	31	31_2X0	7669293	7685494	2	16201
	31	31_3X0	10313741	10460366	7	146625
	31	31_4X0	16745031	16957903	6	212872
	31	31_5X0	19577657	19711636	2	133979
	4	4_10X0	38674493	38794452	3	119959
	4	4_11X0	40177116	40194251	2	17135
	4	4_12X0	43420453	43511825	2	91372
	4	4_13X0	49266235	49334113	2	67878
	4	4_14X0	53209070	53213287	2	4217
	4	4_15X0	58124458	58273515	11	149057
	4	4_16X0	61445609	61576149	2	130540
	4	4_17X0	63049608	63153923	2	104315
	4	4_18X0	73625826	73679217	2	53391
	4	4_19X0	78715737	78720542	2	4805
	4	4_1X0	7617053	7734447	2	117394
	4	4_20X0	79608248	79912973	4	304725
	4	4_21X0	83600830	83639489	2	38659
	4	4_22X0	84868298	84944623	2	76325
	4	4_23X0	92756790	92821386	2	64596

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	4	4_24X0	94566327	94760210	6	193883
	4	4_25X0	94788854	94937186	2	148332
	4	4_26X0	94980395	95391597	27	411202
	4	4_26X1	95406020	95907408	25	501388
	4	4_27X0	95957150	96009680	2	52530
	4	4_28X0	96685820	97416466	26	730646
	4	4_28X1	96116872	96658223	10	541351
	4	4_29X0	101160025	101454771	7	294746
	4	4_2X0	8766008	8997179	5	231171
	4	4_30X0	101740352	101931465	2	191113
	4	4_31X0	101957689	102353280	9	395591
	4	4_32X0	102446502	102554371	2	107869
	4	4_33X0	103426974	103546182	2	119208
	4	4_34X0	104766812	104802401	2	35589
	4	4_3X0	9211896	9759479	6	547583
	4	4_4X0	16280516	16300689	2	20173
	4	4_5X0	23666301	23733586	2	67285
	4	4_6X0	29638016	29662239	2	24223
	4	4_7X0	34202958	34271084	2	68126
	4	4_8X0	36749161	36792765	2	43604
	4	4_9X0	37417410	37529711	2	112301
	5	5_10X0	7290194	7481006	4	190812
	5	5_11X0	9407486	9555547	2	148061
	5	5_12X0	14141220	14172564	2	31344
	5	5_13X0	16305639	16427719	2	122080
	5	5_14X0	17332683	17397208	2	64525
	5	5_15X0	19521400	19567481	3	46081
	5	5_16X0	19577742	19756201	3	178459
	5	5_17X0	26005544	26077751	2	72207
	5	5_18X0	30661442	30832144	3	170702
	5	5_19X0	35676703	35961144	5	284441
	5	5_1X0	99128	244214	3	145086
	5	5_20X0	36152427	36296376	4	143949
	5	5_21X0	36816101	37108811	7	292710
	5	5_22X0	37235331	37337374	2	102043
	5	5_23X0	37365143	37424718	2	59575
	5	5_24X0	37577853	37718801	5	140948
	5	5_25X0	37777677	37905781	2	128104
	5	5_26X0	37953158	38236743	10	283585
	5	5_27X0	38333107	38510401	4	177294
	5	5_28X0	38589608	39136702	27	547094
	5	5_29X0	39172736	39366239	5	193503
	5	5_2X0	1472507	1588210	2	115703
	5	5_30X0	39642856	39720453	2	77597
	5	5_31X0	40053383	40398658	4	345275
	5	5_32X0	41215248	41254498	2	39250
	5	5_33X0	41436475	41472932	2	36457
	5	5_34X0	42095065	42241744	2	146679
	5	5_35X0	42744773	42801833	3	57060
	5	5_36X0	42813626	42838583	2	24957
	5	5_37X0	44263530	44294640	2	31110
	5	5_38X0	45434842	45481225	2	46383
	5	5_39X0	46152185	46179899	2	27714
	5	5_3X0	1662114	1973994	3	311880
	5	5_40X0	46341891	46390153	2	48262

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	5	5_41X0	47063193	47116826	7	53633
	5	5_42X0	48228678	48349159	2	120481
	5	5_43X0	52067023	52148091	2	81068
	5	5_44X0	55580924	55726678	2	145754
	5	5_45X0	56934870	57072201	3	137331
	5	5_46X0	57666620	57752244	2	85624
	5	5_47X0	58419854	58459646	4	39792
	5	5_48X0	58523950	58577764	2	53814
	5	5_49X0	74638952	74765769	2	126817
	5	5_4X0	2000657	2083747	2	83090
	5	5_50X0	74897622	75267540	4	369918
	5	5_51X0	77308114	77552085	5	243971
	5	5_52X0	78753553	78838921	2	85368
	5	5_53X0	84251017	84287586	2	36569
	5	5_5X0	2738080	2956869	3	218789
	5	5_6X0	2987183	3114942	3	127759
	5	5_7X0	3117106	3173462	3	56356
	5	5_8X0	3377848	3460033	2	82185
	5	5_9X0	6026628	6140251	3	113623
	6	6_10X0	32541188	32768813	3	227625
	6	6_11X0	33940005	34012044	2	72039
	6	6_12X0	34261123	34301295	2	40172
	6	6_13X0	34515391	34567899	3	52508
	6	6_14X0	35113757	35289380	5	175623
	6	6_15X0	35706105	35743004	2	36899
	6	6_16X0	36554193	36867535	6	313342
	6	6_17X0	37104343	37449068	7	344725
	6	6_18X0	37755158	37867260	4	112102
	6	6_19X0	38227853	38358104	2	130251
	6	6_1X0	6594703	6625833	2	31130
	6	6_20X0	38981213	39117394	3	136181
	6	6_20X1	38705588	38826110	4	120522
	6	6_21X0	40662156	40692395	2	30239
	6	6_22X0	40816676	40933662	2	116986
	6	6_23X0	42299000	42311564	2	12564
	6	6_24X0	47333344	47546471	3	213127
	6	6_25X0	64255509	64372882	2	117373
	6	6_26X0	66223613	66525494	14	301881
	6	6_26X1	65860190	66149961	6	289771
	6	6_27X0	66694302	66871233	5	176931
	6	6_28X0	66914624	67033054	3	118430
	6	6_29X0	67407035	67534652	2	127617
	6	6_2X0	7857487	7956835	3	99348
	6	6_30X0	67559953	67610811	2	50858
	6	6_31X0	69092922	69159160	2	66238
	6	6_32X0	69280851	69593753	17	312902
	6	6_32X1	69613165	69974648	13	361483
	6	6_33X0	70186370	70365718	3	179348
	6	6_34X0	70376121	70431193	2	55072
	6	6_35X0	70802998	70919290	9	116292
	6	6_36X0	71251019	71273173	2	22154
	6	6_37X0	72048941	72345930	14	296989
	6	6_37X1	72673423	73072029	17	398606
	6	6_37X2	72376664	72652110	13	275446
	6	6_37X3	71668228	72033893	15	365665

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	6	6_37X4	73145302	73432228	10	286926
	6	6_38X0	73651511	73842892	5	191381
	6	6_39X0	74322075	74408059	2	85984
	6	6_3X0	8790796	8809653	2	18857
	6	6_40X0	74602745	74650356	4	47611
	6	6_41X0	75028276	75042308	2	14032
	6	6_4X0	18326672	18508778	2	182106
	6	6_5X0	19572895	19591143	2	18248
	6	6_6X0	25628453	25729705	7	101252
	6	6_7X0	29614146	29695221	2	81075
	6	6_8X0	31491824	31559085	2	67261
	6	6_9X0	32136373	32203036	2	66663
	7	7_10X0	5103751	5145681	3	41930
	7	7_11X0	5553148	5729534	2	176386
	7	7_12X0	12297642	12815411	10	517769
	7	7_13X0	14574017	14619023	3	45006
	7	7_14X0	20466837	20472354	2	5517
	7	7_15X0	22305735	22367036	2	61301
	7	7_16X0	22839080	22956190	3	117110
	7	7_17X0	24834459	24883126	3	48667
	7	7_18X0	25762756	25783939	2	21183
	7	7_19X0	25832965	26025047	2	192082
	7	7_1X0	824546	889918	2	65372
	7	7_20X0	26040708	26146595	5	105887
	7	7_21X0	26182418	26210911	3	28493
	7	7_22X0	30077950	30282327	2	204377
	7	7_23X0	74109976	74564185	22	454209
	7	7_23X1	72492906	72741849	16	248943
	7	7_23X10	76148604	76435111	14	286507
	7	7_23X11	51337075	51681352	14	344277
	7	7_23X12	31794364	32059106	10	264742
	7	7_23X13	32387514	32564876	6	177362
	7	7_23X14	50693540	51235233	16	541693
	7	7_23X15	32095313	32343701	8	248388
	7	7_23X16	30826281	31108228	10	281947
	7	7_23X17	31409764	31742769	10	333005
	7	7_23X18	31142920	31387099	7	244179
	7	7_23X19	75793223	76067305	8	274082
	7	7_23X2	73625529	73859244	12	233715
	7	7_23X20	32865297	33091755	7	226458
	7	7_23X21	51804657	52127114	11	322457
	7	7_23X22	32608482	32790425	7	181943
	7	7_23X23	71392364	71674767	7	282403
	7	7_23X24	50091153	50618913	9	527760
	7	7_23X25	33134497	33458721	7	324224
	7	7_23X26	49644897	49796362	6	151465
	7	7_23X27	45895273	46098084	5	202811
	7	7_23X28	45641390	45861923	4	220533
	7	7_23X29	47096680	47447956	5	351276
	7	7_23X3	73360904	73611938	14	251034
	7	7_23X30	44107613	44376746	4	269133
	7	7_23X31	48202603	48476881	4	274278
	7	7_23X32	68984877	69021981	4	37104
	7	7_23X33	46440497	46606643	3	166146
	7	7_23X34	42448236	42663053	3	214817

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	7	7_23X35	48605491	48854805	3	249314
	7	7_23X36	56530526	56664799	3	134273
	7	7_23X37	37470386	37545283	2	74897
	7	7_23X38	46818732	46842487	2	23755
	7	7_23X39	59964239	59990418	2	26179
	7	7_23X4	73056176	73342230	14	286054
	7	7_23X40	52666867	52736938	2	70071
	7	7_23X41	66311060	66419063	2	108003
	7	7_23X42	34587598	34646519	2	58921
	7	7_23X43	36702926	36844166	2	141240
	7	7_23X44	38847607	38907555	2	59948
	7	7_23X45	68413959	68490292	2	76333
	7	7_23X46	67506431	67645493	2	139062
	7	7_23X47	69851732	69886120	2	34388
	7	7_23X5	72758804	73040745	14	281941
	7	7_23X6	73872631	74099559	12	226928
	7	7_23X7	77070400	77502095	21	431695
	7	7_23X8	72059592	72461941	17	402349
	7	7_23X9	74603241	75289118	20	685877
	7	7_24X0	83443346	83697430	4	254084
	7	7_25X0	86624979	86804684	4	179705
	7	7_26X0	86988866	86999025	2	10159
	7	7_27X0	92233214	92328729	2	95515
	7	7_28X0	96116000	96165734	2	49734
	7	7_29X0	98226271	98343539	2	117268
	7	7_2X0	1467857	1706771	5	238914
	7	7_3X0	1811263	1833888	2	22625
	7	7_4X0	2272120	2564781	4	292661
	7	7_5X0	3388372	3449475	2	61103
	7	7_6X0	3975900	3995720	2	19820
	7	7_7X0	4105595	4184430	2	78835
	7	7_8X0	4196690	4289454	2	92764
	7	7_9X0	4554655	4580669	2	26014
	8	8_10X0	6502073	6581188	2	79115
	8	8_11X0	12970733	13129839	4	159106
	8	8_12X0	13204590	13244872	2	40282
	8	8_13X0	13681720	13770313	2	88593
	8	8_14X0	19159655	19189973	2	30318
	8	8_15X0	19465590	19516482	2	50892
	8	8_16X0	21708977	21792180	2	83203
	8	8_17X0	22831535	22842138	2	10603
	8	8_18X0	29820452	30034900	6	214448
	8	8_19X0	34226372	34270844	2	44472
	8	8_1X0	380778	393066	2	12288
	8	8_20X0	51959710	52472129	7	512419
	8	8_21X0	52855303	52952120	2	96817
	8	8_22X0	55612300	55708503	5	96203
	8	8_23X0	64711737	64834380	2	122643
	8	8_24X0	79778980	80329754	14	550774
	8	8_25X0	88906186	88964604	2	58418
	8	8_2X0	494062	575759	4	81697
	8	8_3X0	730710	793425	2	62715
	8	8_4X0	1259667	1293683	2	34016
	8	8_5X0	1477085	1557499	5	80414
	8	8_6X0	2160539	2900415	7	739876

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	8	8_7X0	4648067	4994903	13	346836
	8	8_7X1	3730193	4144803	12	414610
	8	8_7X2	4210606	4590612	5	380006
	8	8_8X0	5039099	5144387	2	105288
	8	8_9X0	5399186	5454425	2	55239
	9	9_10X0	82888818	82974531	2	85713
	9	9_11X0	83277475	83413846	5	136371
	9	9_1X0	3139003	3330058	4	191055
	9	9_2X0	6377131	6574536	5	197405
	9	9_3X0	27236703	27307849	2	71146
	9	9_4X0	40092531	40232047	2	139516
	9	9_5X0	48814812	48937340	2	122528
	9	9_6X0	70970701	71101356	3	130655
	9	9_7X0	81449605	81890492	9	440887
	9	9_8X0	82267556	82288331	2	20775
	9	9_9X0	82679925	82823979	3	144054
	Un0002	Un0002_1X0	212558	218731	2	6173
	Un0002	Un0002_2X0	947007	990041	2	43034
	Un0011	Un0011_1X0	674945	768934	5	93989
	Un0011	Un0011_1X1	576506	652695	4	76189
	X	X_10X0	33782100	33805328	2	23228
	X	X_11X0	38414539	38504509	2	89970
	X	X_12X0	38862249	39046638	3	184389
	X	X_13X0	39612602	39637956	2	25354
	X	X_14X0	41459952	41495756	2	35804
	X	X_15X0	41832762	41982785	3	150023
	X	X_16X0	43536477	43548409	2	11932
	X	X_17X0	46027201	46249978	3	222777
	X	X_18X0	51957728	52212146	3	254418
	X	X_19X0	58087764	58135854	2	48090
	X	X_1X0	401485	573171	5	171686
	X	X_20X0	58554971	58970146	4	415175
	X	X_21X0	62114201	62242381	2	128180
	X	X_22X0	73993741	74047798	2	54057
	X	X_23X0	79858878	79973513	6	114635
	X	X_24X0	80084932	80186168	2	101236
	X	X_25X0	80734275	80886290	3	152015
	X	X_26X0	81087979	81145079	3	57100
	X	X_27X0	81170086	81376212	5	206126
	X	X_28X0	81658764	81711148	2	52384
	X	X_29X0	85349355	85364067	2	14712
	X	X_2X0	2165874	2281485	2	115611
	X	X_30X0	94846177	94854097	2	7920
	X	X_31X0	94973941	95017631	2	43690
	X	X_32X0	104330844	104606922	5	276078
	X	X_33X0	107155373	107212085	2	56712
	X	X_34X0	107667301	107945324	3	278023
	X	X_35X0	109926517	110055896	2	129379
	X	X_36X0	118567949	118821107	3	253158
	X	X_37X0	121806322	121866030	2	59708
	X	X_38X0	121921614	121951428	2	29814
	X	X_3X0	6123963	6148502	2	24539
	X	X_4X0	8675768	8739205	2	63437
	X	X_5X0	10764705	10822729	2	58024
	X	X_6X0	19230850	19394876	2	164026

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	X	X_7X0	20017304	20097605	3	80301
	X	X_8X0	21376937	21436625	3	59688
	X	X_9X0	30115574	30131941	2	16367
	A1	A1_10X0	83397995	84379011	13	981016
	A1	A1_11X0	84425205	84501213	3	76008
	A1	A1_12X0	84609646	84625358	2	15712
	A1	A1_13X0	85029833	86173628	13	1143795
	A1	A1_14X0	86694732	87760829	19	1066097
	A1	A1_15X0	87925132	88019680	5	94548
	A1	A1_16X0	88309953	88403634	2	93681
	A1	A1_17X0	91415868	91757293	8	341425
	A1	A1_18X0	93365920	93807007	8	441087
	A1	A1_19X0	94440092	94455744	2	15652
	A1	A1_1X0	1246106	1280216	2	34110
	A1	A1_20X0	111420516	111501695	2	81179
	A1	A1_21X0	111650599	111844072	2	193473
	A1	A1_22X0	118930367	118950349	2	19982
	A1	A1_23X0	119048617	120032422	18	983805
	A1	A1_24X0	125374200	125708445	4	334245
	A1	A1_25X0	126487223	126597315	2	110092
	A1	A1_26X0	138297019	138408794	2	111775
	A1	A1_27X0	139238076	139357624	2	119548
	A1	A1_28X0	143147870	143349887	3	202017
	A1	A1_29X0	145392911	145455902	2	62991
	A1	A1_2X0	7445463	7603603	3	158140
	A1	A1_30X0	176823765	176999476	2	175711
	A1	A1_31X0	193007116	193125356	3	118240
	A1	A1_32X0	197966892	198104280	3	137388
	A1	A1_33X0	198375096	198425140	2	50044
	A1	A1_34X0	198916441	199603991	4	687550
	A1	A1_35X0	200182786	200448240	2	265454
	A1	A1_36X0	200605285	200646804	3	41519
	A1	A1_37X0	213067192	213218889	3	151697
	A1	A1_38X0	230109432	230207940	2	98508
	A1	A1_39X0	241229899	241384936	3	155037
	A1	A1_3X0	7629927	7915911	3	285984
	A1	A1_4X0	10118138	10173442	2	55304
	A1	A1_5X0	19434187	19505810	2	71623
	A1	A1_6X0	29371829	29448873	2	77044
	A1	A1_7X0	71317891	71347691	2	29800
	A1	A1_8X0	73605683	73834027	2	228344
	A1	A1_9X0	82236813	82261461	2	24648
	A2	A2_10X0	4810690	4833282	2	22592
	A2	A2_11X0	5225821	5312596	2	86775
	A2	A2_12X0	6076142	6109521	2	33379
	A2	A2_13X0	6854639	7387829	24	533190
	A2	A2_14X0	7673742	7839662	8	165920
	A2	A2_15X0	8428854	8660492	2	231638
	A2	A2_16X0	8680385	8985041	7	304656
	A2	A2_17X0	9023502	9041511	2	18009
	A2	A2_18X0	10381921	10568458	4	186537
	A2	A2_19X0	10670027	11203245	23	533218
	A2	A2_1X0	326164	442221	3	116057
	A2	A2_20X0	11527303	11572351	2	45048
	A2	A2_21X0	11641844	11703566	2	61722

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	A2	A2_22X0	11729703	11821334	4	91631
	A2	A2_23X0	13181021	13393146	2	212125
	A2	A2_24X0	14625706	14649658	2	23952
	A2	A2_25X0	15607533	15962529	6	354996
	A2	A2_26X0	16186344	16266990	4	80646
	A2	A2_27X0	18238312	18397503	2	159191
	A2	A2_28X0	18711803	18737288	2	25485
	A2	A2_29X0	18854550	18893969	2	39419
	A2	A2_2X0	735084	874024	2	138940
	A2	A2_30X0	18917798	19060720	2	142922
	A2	A2_31X0	19090509	19196521	4	106012
	A2	A2_32X0	19207642	19225899	3	18257
	A2	A2_33X0	20356777	20362138	2	5361
	A2	A2_34X0	20482773	20564886	5	82113
	A2	A2_35X0	20576017	20872282	3	296265
	A2	A2_36X0	21325790	21376288	3	50498
	A2	A2_37X0	51150755	51173939	2	23184
	A2	A2_38X0	56620059	56690160	2	70101
	A2	A2_39X0	63310755	63333416	2	22661
	A2	A2_3X0	1260252	1343841	2	83589
	A2	A2_40X0	77339699	77517232	4	177533
	A2	A2_41X0	78413186	78574503	2	161317
	A2	A2_42X0	78730276	78860478	2	130202
	A2	A2_43X0	93005934	93194177	2	188243
	A2	A2_44X0	95577393	95641466	2	64073
	A2	A2_45X0	95744848	95809353	2	64505
	A2	A2_46X0	97274863	97298027	2	23164
	A2	A2_47X0	98962864	98979229	2	16365
	A2	A2_48X0	100169466	100286924	3	117458
	A2	A2_49X0	101759625	101777991	2	18366
	A2	A2_4X0	1931450	1999208	3	67758
	A2	A2_50X0	111293539	111355503	2	61964
	A2	A2_51X0	113454011	113486795	2	32784
	A2	A2_52X0	117155307	117366167	2	210860
	A2	A2_53X0	120508432	120657611	10	149179
	A2	A2_54X0	124016903	124146188	2	129285
	A2	A2_55X0	125515854	125736705	2	220851
	A2	A2_56X0	142804811	142859691	2	54880
	A2	A2_57X0	146852657	146882846	2	30189
	A2	A2_58X0	148084247	148185034	4	100787
	A2	A2_59X0	155999270	156046935	2	47665
	A2	A2_5X0	2376671	2406601	2	29930
	A2	A2_60X0	157878202	158054043	5	175841
	A2	A2_61X0	158082937	158252558	2	169621
	A2	A2_62X0	158278509	158525965	7	247456
	A2	A2_63X0	158608283	158666908	2	58625
	A2	A2_64X0	159128624	159656934	19	528310
	A2	A2_64X1	158756726	158773712	2	16986
	A2	A2_65X0	163688275	163970622	8	282347
	A2	A2_65X1	164076593	164266702	4	190109
	A2	A2_66X0	164779293	164942939	3	163646
	A2	A2_67X0	165904496	166005434	2	100938
	A2	A2_6X0	2459508	2513786	2	54278
	A2	A2_7X0	2945630	3176588	3	230958
	A2	A2_8X0	4178642	4248767	2	70125

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	A2	A2_9X0	4574798	4633370	2	58572
	A3	A3_10X0	28770772	29163529	7	392757
	A3	A3_11X0	29480924	29588428	2	107504
	A3	A3_12X0	30091772	30106541	2	14769
	A3	A3_13X0	31442239	31462792	2	20553
	A3	A3_14X0	43557616	43793384	2	235768
	A3	A3_15X0	46682892	46785228	2	102336
	A3	A3_16X0	48125271	48165225	2	39954
	A3	A3_17X0	48539090	48790136	5	251046
	A3	A3_18X0	54362825	54386202	2	23377
	A3	A3_19X0	54789638	55001604	3	211966
	A3	A3_1X0	473823	487572	2	13749
	A3	A3_20X0	60173311	60610674	6	437363
	A3	A3_21X0	63112338	63165556	2	53218
	A3	A3_22X0	64888976	64909159	2	20183
	A3	A3_23X0	64916019	64976197	2	60178
	A3	A3_24X0	66060418	66109103	2	48685
	A3	A3_25X0	67165083	67237449	2	72366
	A3	A3_26X0	68937539	69173692	2	236153
	A3	A3_27X0	88560563	88591379	2	30816
	A3	A3_28X0	90150994	90185689	2	34695
	A3	A3_29X0	91896516	92155214	2	258698
	A3	A3_2X0	935155	956092	2	20937
	A3	A3_30X0	93186391	93205201	2	18810
	A3	A3_31X0	97118492	97158190	2	39698
	A3	A3_32X0	102707457	102719424	2	11967
	A3	A3_33X0	103823473	103883903	2	60430
	A3	A3_34X0	104466811	104595833	6	129022
	A3	A3_35X0	104847813	105121118	6	273305
	A3	A3_36X0	105443069	105539931	2	96862
	A3	A3_37X0	107248901	107318041	4	69140
	A3	A3_38X0	110804107	111068177	2	264070
	A3	A3_39X0	112646816	112848696	2	201880
	A3	A3_3X0	10220290	10582692	3	362402
	A3	A3_40X0	120945032	121135169	2	190137
	A3	A3_41X0	123890980	123930523	2	39543
	A3	A3_42X0	142245983	142360114	2	114131
	A3	A3_4X0	15418513	15886480	9	467967
	A3	A3_5X0	21621896	21682800	2	60904
	A3	A3_6X0	25947066	26218074	6	271008
	A3	A3_7X0	26950261	27180522	2	230261
	A3	A3_8X0	27464840	27750096	10	285256
	A3	A3_9X0	27907492	27998401	2	90909
	B1	B1_10X0	46802626	46947588	2	144962
	B1	B1_11X0	55169166	55245150	4	75984
	B1	B1_12X0	65332371	65477749	4	145378
	B1	B1_13X0	67049409	67077339	2	27930
	B1	B1_14X0	74054061	74201925	2	147864
	B1	B1_15X0	75180157	75229105	3	48948
	B1	B1_16X0	121425166	121572701	2	147535
	B1	B1_17X0	124654503	124850748	4	196245
	B1	B1_18X0	134459201	134747591	3	288390
	B1	B1_19X0	135608920	135677274	2	68354
	B1	B1_1X0	6862001	7285871	15	423870
	B1	B1_20X0	139569351	139652229	2	82878

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	B1	B1_21X0	140234423	140332491	3	98068
	B1	B1_22X0	145026793	145137202	2	110409
	B1	B1_23X0	146137267	146202668	3	65401
	B1	B1_24X0	147886171	148130339	6	244168
	B1	B1_25X0	148280851	148423061	4	142210
	B1	B1_26X0	151391237	151523240	3	132003
	B1	B1_27X0	151594328	151735888	2	141560
	B1	B1_28X0	151873775	152380631	7	506856
	B1	B1_29X0	166609454	166656055	2	46601
	B1	B1_2X0	14494111	14496995	2	2884
	B1	B1_30X0	167270017	167382020	2	112003
	B1	B1_31X0	196182899	196318277	2	135378
	B1	B1_32X0	202637266	202814159	2	176893
	B1	B1_33X0	203057188	203129886	2	72698
	B1	B1_34X0	203553160	203595586	3	42426
	B1	B1_3X0	17132773	17177817	2	45044
	B1	B1_4X0	18747295	18796631	2	49336
	B1	B1_5X0	33772754	33993822	4	221068
	B1	B1_6X0	34677425	34713964	2	36539
	B1	B1_7X0	42285640	42353856	2	68216
	B1	B1_8X0	44675737	44782301	2	106564
	B1	B1_9X0	44820973	45915545	6	1094572
	B2	B2_10X0	25709866	25764996	2	55130
	B2	B2_11X0	25959399	26086326	3	126927
	B2	B2_12X0	31998620	32106374	3	107754
	B2	B2_13X0	32459068	32708658	14	249590
	B2	B2_14X0	32786138	32802651	3	16513
	B2	B2_15X0	32899623	32917912	2	18289
	B2	B2_16X0	33601519	34161918	17	560399
	B2	B2_17X0	34400254	34927411	13	527157
	B2	B2_18X0	34928370	35056143	2	127773
	B2	B2_19X0	37191471	37484539	3	293068
	B2	B2_1X0	182990	287315	4	104325
	B2	B2_20X0	40384282	40496742	3	112460
	B2	B2_21X0	42223048	42364092	5	141044
	B2	B2_22X0	43102230	43121551	2	19321
	B2	B2_23X0	43865094	44057276	4	192182
	B2	B2_24X0	46738287	46772529	2	34242
	B2	B2_25X0	47435700	47597590	2	161890
	B2	B2_26X0	48173171	48225825	2	52654
	B2	B2_27X0	50128657	50330747	2	202090
	B2	B2_28X0	50394304	50505509	5	111205
	B2	B2_29X0	51137746	51269579	2	131833
	B2	B2_2X0	456506	933531	16	477025
	B2	B2_2X1	1008325	1608861	12	600536
	B2	B2_30X0	52598447	52666783	2	68336
	B2	B2_31X0	53099907	53212149	5	112242
	B2	B2_32X0	69029091	69054367	3	25276
	B2	B2_33X0	82686749	82780564	2	93815
	B2	B2_34X0	97479898	97537395	2	57497
	B2	B2_35X0	101525020	101610730	2	85710
	B2	B2_36X0	102829064	102977924	2	148860
	B2	B2_37X0	107535705	107556822	2	21117
	B2	B2_38X0	107744865	107825341	3	80476
	B2	B2_39X0	107903272	108090683	3	187411

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	B2	B2_3X0	1670159	2377396	15	707237
	B2	B2_40X0	117775944	117834212	2	58268
	B2	B2_41X0	122436580	122566981	10	130401
	B2	B2_42X0	126417765	126556639	2	138874
	B2	B2_43X0	147038122	147294934	3	256812
	B2	B2_44X0	152771416	152813899	2	42483
	B2	B2_4X0	2864965	2974633	4	109668
	B2	B2_5X0	3017273	3083965	13	66692
	B2	B2_6X0	3303811	3376343	2	72532
	B2	B2_7X0	3650291	3683666	3	33375
	B2	B2_8X0	3848276	4375363	34	527087
	B2	B2_9X0	18342087	18397816	2	55729
	B3	B3_10X0	34040725	34081039	2	40314
	B3	B3_11X0	34384613	34529863	2	145250
	B3	B3_12X0	41871081	41956171	2	85090
	B3	B3_13X0	54501445	54631742	2	130297
	B3	B3_14X0	55687102	55784055	2	96953
	B3	B3_15X0	60712194	60778466	4	66272
	B3	B3_16X0	62195163	62289726	3	94563
	B3	B3_17X0	63200835	63499117	4	298282
	B3	B3_18X0	64934040	65032739	2	98699
	B3	B3_19X0	72925540	73384251	18	458711
	B3	B3_19X1	71684579	72487790	19	803211
	B3	B3_19X2	72516547	72907883	9	391336
	B3	B3_1X0	169537	257051	3	87514
	B3	B3_20X0	73680749	74055854	9	375105
	B3	B3_21X0	74538141	74613941	5	75800
	B3	B3_22X0	74707687	75046337	9	338650
	B3	B3_23X0	75861602	75873932	2	12330
	B3	B3_24X0	76134102	76187655	2	53553
	B3	B3_25X0	76333775	76775834	5	442059
	B3	B3_26X0	76786889	76792801	2	5912
	B3	B3_27X0	76871619	76915707	2	44088
	B3	B3_28X0	76948977	77039271	6	90294
	B3	B3_29X0	87948476	88015727	2	67251
	B3	B3_2X0	2295287	2366668	2	71381
	B3	B3_30X0	99746093	99815525	2	69432
	B3	B3_31X0	100263725	100455887	2	192162
	B3	B3_32X0	102100682	102162931	2	62249
	B3	B3_33X0	109467600	109658305	2	190705
	B3	B3_34X0	112999193	113049485	2	50292
	B3	B3_35X0	115835505	115875025	2	39520
	B3	B3_36X0	118262090	118319532	4	57442
	B3	B3_37X0	121030010	121152135	4	122125
	B3	B3_38X0	122255515	122450772	2	195257
	B3	B3_39X0	122573871	122803344	3	229473
	B3	B3_3X0	4789473	5028926	3	239453
	B3	B3_40X0	139523969	139807122	8	283153
	B3	B3_41X0	140746202	140782610	2	36408
	B3	B3_42X0	147353622	147385738	2	32116
	B3	B3_43X0	149331443	149567911	4	236468
	B3	B3_4X0	6097999	6119446	2	21447
	B3	B3_5X0	6864744	6914412	2	49668
	B3	B3_6X0	7082475	7107932	2	25457
	B3	B3_7X0	23202606	23251363	2	48757

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	B3	B3_8X0	28201062	28331424	2	130362
	B3	B3_9X0	33313520	33638726	7	325206
	B4	B4_10X0	42111645	42150955	2	39310
	B4	B4_11X0	42625173	42765863	4	140690
	B4	B4_12X0	43178711	43219672	2	40961
	B4	B4_13X0	43423309	43777656	4	354347
	B4	B4_14X0	45950652	46515462	15	564810
	B4	B4_14X1	46682489	47198348	7	515859
	B4	B4_15X0	47404518	47554803	4	150285
	B4	B4_16X0	48894949	48944542	2	49593
	B4	B4_17X0	49055164	49178786	2	123622
	B4	B4_18X0	55969638	56181363	3	211725
	B4	B4_19X0	56297917	56739299	4	441382
	B4	B4_1X0	4647400	4756740	2	109340
	B4	B4_20X0	75476989	75602856	2	125867
	B4	B4_21X0	77091252	77115370	2	24118
	B4	B4_22X0	77421474	77688337	13	266863
	B4	B4_23X0	77873336	78058983	5	185647
	B4	B4_24X0	78103371	78243644	3	140273
	B4	B4_25X0	78638813	78783470	2	144657
	B4	B4_26X0	78833170	78856432	3	23262
	B4	B4_27X0	80485189	80563067	2	77878
	B4	B4_28X0	80691531	81404817	27	713286
	B4	B4_29X0	81586089	81790853	3	204764
	B4	B4_2X0	4880297	4901318	2	21021
	B4	B4_30X0	81798199	81857158	2	58959
	B4	B4_31X0	82272248	82391286	9	119038
	B4	B4_32X0	83243128	83913559	12	670431
	B4	B4_33X0	83920081	84344041	16	423960
	B4	B4_34X0	84788004	84795610	2	7606
	B4	B4_35X0	85336438	85489412	4	152974
	B4	B4_36X0	95382387	95421346	2	38959
	B4	B4_37X0	101300560	101442619	3	142059
	B4	B4_38X0	106072285	106085024	2	12739
	B4	B4_39X0	115077290	115215798	3	138508
	B4	B4_3X0	6932921	7093599	2	160678
	B4	B4_40X0	133206301	133379325	4	173024
	B4	B4_41X0	133721516	133808110	2	86594
	B4	B4_42X0	136806527	136855281	2	48754
	B4	B4_43X0	138316070	138440135	2	124065
	B4	B4_44X0	139202110	139278814	2	76704
	B4	B4_45X0	144023385	144040369	2	16984
	B4	B4_4X0	32281833	32284806	2	2973
	B4	B4_5X0	37702640	37766339	2	63699
	B4	B4_6X0	38470960	38638010	2	167050
	B4	B4_7X0	39605489	39678939	2	73450
	B4	B4_8X0	40048663	40284635	3	235972
	B4	B4_9X0	41824162	41885635	2	61473
	C1	C1_10X0	18933269	18953915	2	20646
	C1	C1_11X0	20588589	20831703	2	243114
	C1	C1_12X0	24000748	24134760	2	134012
	C1	C1_13X0	24501323	24615669	3	114346
	C1	C1_14X0	24661942	24800857	2	138915
	C1	C1_15X0	26731350	26768600	3	37250
	C1	C1_16X0	26936721	27273552	3	336831

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	C1	C1_17X0	27612394	27836249	3	223855
	C1	C1_18X0	29443825	29489297	2	45472
	C1	C1_19X0	30762788	30913196	2	150408
	C1	C1_1X0	4313808	4335007	2	21199
	C1	C1_20X0	31276759	31621451	4	344692
	C1	C1_21X0	33086699	33093959	2	7260
	C1	C1_22X0	33907567	33985339	5	77772
	C1	C1_23X0	34041184	34214588	2	173404
	C1	C1_24X0	37427626	37591443	5	163817
	C1	C1_25X0	37823167	37845802	2	22635
	C1	C1_26X0	42818160	42997559	2	179399
	C1	C1_27X0	55937699	56113955	2	176256
	C1	C1_28X0	66279091	66315202	2	36111
	C1	C1_29X0	71893129	71981885	2	88756
	C1	C1_2X0	6954541	7037106	2	82565
	C1	C1_30X0	73278691	73502144	4	223453
	C1	C1_31X0	75718031	76035010	5	316979
	C1	C1_32X0	76168097	76518449	3	350352
	C1	C1_33X0	92694783	92822796	2	128013
	C1	C1_34X0	93833849	93891971	2	58122
	C1	C1_35X0	93948234	93973751	3	25517
	C1	C1_36X0	95397040	95488248	4	91208
	C1	C1_37X0	96697920	96839725	2	141805
	C1	C1_38X0	100404983	100538684	2	133701
	C1	C1_39X0	103867202	103895794	2	28592
	C1	C1_3X0	8503267	8522773	3	19506
	C1	C1_40X0	105085104	105182459	9	97355
	C1	C1_41X0	105900812	105952068	2	51256
	C1	C1_42X0	106375493	106519337	2	143844
	C1	C1_43X0	106870612	106926677	2	56065
	C1	C1_44X0	107600004	107860481	3	260477
	C1	C1_45X0	121575202	121771473	2	196271
	C1	C1_46X0	142544889	142782343	2	237454
	C1	C1_47X0	144346196	144374394	2	28198
	C1	C1_48X0	151897263	152153839	2	256576
	C1	C1_49X0	160935839	161132383	2	196544
	C1	C1_4X0	12208353	12478761	5	270408
	C1	C1_50X0	163234640	163252940	2	18300
	C1	C1_51X0	166849207	166958083	10	108876
	C1	C1_52X0	168261159	168319785	2	58626
	C1	C1_53X0	178525279	178639310	2	114031
	C1	C1_54X0	180440415	180583939	2	143524
	C1	C1_55X0	188088286	188248547	3	160261
	C1	C1_56X0	190521003	190746651	3	225648
	C1	C1_57X0	202003835	202038612	2	34777
	C1	C1_58X0	203278105	203384065	3	105960
	C1	C1_59X0	203887110	203919611	2	32501
	C1	C1_5X0	14049275	14065651	2	16376
	C1	C1_60X0	204244176	204261684	2	17508
	C1	C1_61X0	204452399	204524509	2	72110
	C1	C1_62X0	211573760	211624943	2	51183
	C1	C1_63X0	214061075	214239050	2	177975
	C1	C1_64X0	215349937	215367733	2	17796
	C1	C1_6X0	14579870	14735256	4	155386
	C1	C1_7X0	15205603	15263563	2	57960

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	C1	C1_8X0	16779945	16801352	3	21407
	C1	C1_9X0	17998315	18050096	2	51781
	C2	C2_10X0	12286779	12331238	2	44459
	C2	C2_11X0	14519367	14744401	7	225034
	C2	C2_12X0	14775112	14939039	10	163927
	C2	C2_13X0	14999473	15049786	2	50313
	C2	C2_14X0	17941310	18079078	2	137768
	C2	C2_15X0	46303024	46636890	11	333866
	C2	C2_16X0	56329297	56364686	2	35389
	C2	C2_17X0	58994055	59852915	5	858860
	C2	C2_18X0	68301636	68466781	3	165145
	C2	C2_19X0	74410021	74467454	3	57433
	C2	C2_1X0	690866	855467	2	164601
	C2	C2_20X0	79887867	80006213	2	118346
	C2	C2_21X0	83090317	83185609	4	95292
	C2	C2_22X0	85343349	85569128	5	225779
	C2	C2_23X0	85789854	85970585	2	180731
	C2	C2_24X0	98343587	98411985	3	68398
	C2	C2_25X0	106810136	106995947	2	185811
	C2	C2_26X0	108429994	108521147	2	91153
	C2	C2_27X0	114798519	114871313	2	72794
	C2	C2_28X0	115312106	115447000	5	134894
	C2	C2_29X0	117030564	117232899	3	202335
	C2	C2_2X0	1931045	1970247	2	39202
	C2	C2_30X0	117251309	117355124	2	103815
	C2	C2_31X0	117609049	117675809	2	66760
	C2	C2_32X0	118890691	118918857	2	28166
	C2	C2_33X0	119697806	120019973	4	322167
	C2	C2_34X0	126081574	126174828	2	93254
	C2	C2_35X0	126599621	126611506	2	11885
	C2	C2_36X0	131213849	131393253	4	179404
	C2	C2_37X0	145600476	145811284	2	210808
	C2	C2_38X0	150292461	150504268	3	211807
	C2	C2_39X0	150823095	151019727	2	196632
	C2	C2_3X0	2097954	2179394	2	81440
	C2	C2_40X0	152888809	152945150	3	56341
	C2	C2_41X0	153837644	153931766	2	94122
	C2	C2_42X0	154204418	154575132	3	370714
	C2	C2_43X0	154761369	154807988	2	46619
	C2	C2_44X0	160890371	161112948	6	222577
	C2	C2_4X0	2622907	2736711	2	113804
	C2	C2_5X0	3740947	3787817	3	46870
	C2	C2_6X0	4037956	4211024	2	173068
	C2	C2_7X0	9362511	9489134	4	126623
	C2	C2_8X0	11009600	11090113	2	80513
	C2	C2_9X0	11922324	12131244	4	208920
	D1	D1_10X0	15775945	15886646	5	110701
	D1	D1_11X0	15914466	15937041	2	22575
	D1	D1_12X0	20825081	21650253	37	825172
	D1	D1_13X0	21808137	21962848	5	154711
	D1	D1_14X0	22981851	23062198	2	80347
	D1	D1_15X0	25802481	25882749	2	80268
	D1	D1_16X0	27178226	27245514	2	67288
	D1	D1_17X0	31020204	31282015	3	261811
	D1	D1_18X0	40718073	40735298	2	17225

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	D1	D1_19X0	45256823	45365204	3	108381
	D1	D1_1X0	1041682	1095582	2	53900
	D1	D1_20X0	48416645	48462235	2	45590
	D1	D1_21X0	49021426	49055297	2	33871
	D1	D1_22X0	56153339	56305639	2	152300
	D1	D1_23X0	57520355	57663148	2	142793
	D1	D1_24X0	58436795	58524393	2	87598
	D1	D1_25X0	59939149	59974257	2	35108
	D1	D1_26X0	102983531	103319580	22	336049
	D1	D1_26X1	62868573	63277661	28	409088
	D1	D1_26X10	105555897	106242153	17	686256
	D1	D1_26X11	104893800	105394049	14	500249
	D1	D1_26X12	102049984	102413344	9	363360
	D1	D1_26X13	61564512	61869785	7	305273
	D1	D1_26X14	104204842	104611749	7	406907
	D1	D1_26X15	66602402	66737395	7	134993
	D1	D1_26X16	76106891	76359899	6	253008
	D1	D1_26X17	60594157	60662532	2	68375
	D1	D1_26X18	98404583	98435996	2	31413
	D1	D1_26X2	63331417	63832472	32	501055
	D1	D1_26X3	103334132	103957530	29	623398
	D1	D1_26X4	102472759	102947949	26	475190
	D1	D1_26X5	62278137	62844017	21	565880
	D1	D1_26X6	64283782	64580790	17	297008
	D1	D1_26X7	64596461	65077672	17	481211
	D1	D1_26X8	63852886	64255104	15	402218
	D1	D1_26X9	65628517	65916145	15	287628
	D1	D1_27X0	106313305	106348934	2	35629
	D1	D1_28X0	106445060	107005953	10	560893
	D1	D1_29X0	107142348	107268885	2	126537
	D1	D1_2X0	1185300	1523789	9	338489
	D1	D1_30X0	107322817	107357926	2	35109
	D1	D1_31X0	107813442	107946791	3	133349
	D1	D1_32X0	108144180	108229203	3	85023
	D1	D1_33X0	108383310	108560008	2	176698
	D1	D1_34X0	108672301	108860598	2	188297
	D1	D1_35X0	108895298	109025191	3	129893
	D1	D1_36X0	109146517	109220722	2	74205
	D1	D1_37X0	109799072	109838100	2	39028
	D1	D1_38X0	110016229	110060090	2	43861
	D1	D1_39X0	110662157	110754620	2	92463
	D1	D1_3X0	3159883	3205621	2	45738
	D1	D1_40X0	111611415	111663409	3	51994
	D1	D1_41X0	112170307	112302513	2	132206
	D1	D1_42X0	112319760	112379624	2	59864
	D1	D1_43X0	112471565	112547844	2	76279
	D1	D1_44X0	113852454	113932711	3	80257
	D1	D1_45X0	115941577	115972658	2	31081
	D1	D1_46X0	116112828	116184884	2	72056
	D1	D1_47X0	117010020	117052986	4	42966
	D1	D1_48X0	117475373	117503775	2	28402
	D1	D1_4X0	9944204	9950758	2	6554
	D1	D1_5X0	11794370	11863214	2	68844
	D1	D1_6X0	12129089	12170690	2	41601
	D1	D1_7X0	14536192	14572621	3	36429

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	D1	D1_8X0	15511698	15553097	2	41399
	D1	D1_9X0	15580212	15760878	2	180666
	D2	D2_10X0	47212575	47226341	2	13766
	D2	D2_11X0	50243443	50366472	6	123029
	D2	D2_12X0	51448884	51471007	2	22123
	D2	D2_13X0	55346636	55364381	2	17745
	D2	D2_14X0	56668363	56745366	2	77003
	D2	D2_15X0	61786458	61856093	2	69635
	D2	D2_16X0	63741699	63770204	3	28505
	D2	D2_17X0	64539353	64577413	2	38060
	D2	D2_18X0	75270043	75438934	4	168891
	D2	D2_19X0	79667235	79773246	2	106011
	D2	D2_1X0	8237286	8472843	5	235557
	D2	D2_20X0	80787481	80863562	2	76081
	D2	D2_21X0	80884144	80939726	2	55582
	D2	D2_22X0	81104713	81119575	2	14862
	D2	D2_23X0	89735098	89804906	2	69808
	D2	D2_24X0	89945036	90021125	3	76089
	D2	D2_2X0	8945464	9024849	4	79385
	D2	D2_3X0	9032416	9387491	3	355075
	D2	D2_4X0	32030104	32312154	3	282050
	D2	D2_5X0	32326288	32521678	2	195390
	D2	D2_6X0	37493054	37563286	2	70232
	D2	D2_7X0	41696937	41740619	3	43682
	D2	D2_8X0	41801532	41826613	2	25081
	D2	D2_9X0	44873327	44896457	2	23130
	D3	D3_10X0	19129510	19167619	4	38109
	D3	D3_11X0	19351902	19401542	2	49640
	D3	D3_12X0	19873044	19960991	2	87947
	D3	D3_13X0	20087626	20978521	23	890895
	D3	D3_14X0	21010839	21152233	2	141394
	D3	D3_15X0	22467406	22544707	3	77301
	D3	D3_16X0	22658201	22685257	2	27056
	D3	D3_17X0	29133106	29140016	2	6910
	D3	D3_18X0	29545630	29724898	2	179268
	D3	D3_19X0	30226320	30386350	2	160030
	D3	D3_1X0	188601	393985	6	205384
	D3	D3_20X0	30510341	30581872	2	71531
	D3	D3_21X0	33119732	33204203	2	84471
	D3	D3_22X0	42846050	42874695	2	28645
	D3	D3_23X0	54938177	55477895	7	539718
	D3	D3_24X0	55838221	55930718	2	92497
	D3	D3_25X0	58788924	58886151	5	97227
	D3	D3_26X0	68165709	68280217	2	114508
	D3	D3_27X0	83423901	83788475	9	364574
	D3	D3_28X0	92286837	92344642	2	57805
	D3	D3_2X0	7105249	7121198	2	15949
	D3	D3_3X0	8241568	8331200	2	89632
	D3	D3_4X0	9768437	9788669	2	20232
	D3	D3_5X0	10679741	10742441	3	62700
	D3	D3_6X0	11018073	11049300	2	31227
	D3	D3_7X0	16708259	16805677	2	97418
	D3	D3_8X0	17268767	17306978	2	38211
	D3	D3_9X0	18872313	19070236	3	197923
	D4	D4_10X0	49515759	49734355	2	218596

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	D4	D4_11X0	57605532	57659204	2	53672
	D4	D4_12X0	58369591	58473684	2	104093
	D4	D4_13X0	58501171	58521712	2	20541
	D4	D4_14X0	59359088	59530864	2	171776
	D4	D4_15X0	59580881	59707239	8	126358
	D4	D4_16X0	61810495	61943413	2	132918
	D4	D4_17X0	65328673	65357579	2	28906
	D4	D4_18X0	67858145	68164198	12	306053
	D4	D4_19X0	68192803	68244376	2	51573
	D4	D4_1X0	557757	682408	4	124651
	D4	D4_20X0	71657868	71665658	2	7790
	D4	D4_21X0	71999733	72196573	2	196840
	D4	D4_22X0	72864469	72928065	2	63596
	D4	D4_23X0	75170784	75236902	2	66118
	D4	D4_24X0	75296646	75406496	2	109850
	D4	D4_25X0	75416057	75479258	2	63201
	D4	D4_26X0	82562382	82635052	3	72670
	D4	D4_27X0	83503671	84068216	22	564545
	D4	D4_28X0	84157548	84175270	2	17722
	D4	D4_29X0	87618275	87665954	2	47679
	D4	D4_2X0	6200087	6342302	2	142215
	D4	D4_30X0	88511717	88539257	2	27540
	D4	D4_31X0	88709017	88854208	7	145191
	D4	D4_32X0	90101901	90118801	2	16900
	D4	D4_33X0	94262068	94282577	2	20509
	D4	D4_34X0	95210292	95582295	11	372003
	D4	D4_35X0	95667650	95705432	2	37782
	D4	D4_3X0	6375807	6434998	2	59191
	D4	D4_4X0	10150871	10286208	2	135337
	D4	D4_5X0	28236740	28631178	5	394438
	D4	D4_6X0	28634900	28700386	2	65486
	D4	D4_7X0	30880303	31013347	2	133044
	D4	D4_8X0	47624644	47829647	10	205003
	D4	D4_9X0	48177372	48224606	2	47234
	E1	E1_10X0	9312036	9434936	3	122900
	E1	E1_11X0	10515512	10588147	2	72635
	E1	E1_12X0	10627083	10727218	2	100135
	E1	E1_13X0	13277248	13410278	2	133030
	E1	E1_14X0	13517254	13581729	2	64475
	E1	E1_15X0	13629763	13944794	16	315031
	E1	E1_16X0	15083678	15107018	2	23340
	E1	E1_18X0	17972834	18183659	4	210825
	E1	E1_19X0	22086119	22151908	4	65789
	E1	E1_1X0	548845	620493	2	71648
	E1	E1_20X0	23276062	23419041	6	142979
	E1	E1_21X0	28163604	28240694	2	77090
	E1	E1_22X0	30565120	30645904	3	80784
	E1	E1_23X0	30656244	30683302	3	27058
	E1	E1_24X0	37153363	37181285	2	27922
	E1	E1_25X0	37216123	37268946	2	52823
	E1	E1_26X0	37468269	37503401	2	35132
	E1	E1_27X0	37595724	37622235	2	26511
	E1	E1_28X0	38619061	38794088	10	175027
	E1	E1_29X0	39342093	39401107	2	59014
	E1	E1_2X0	1678893	1816711	2	137818

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	E1	E1_30X0	41069694	41100928	2	31234
	E1	E1_31X0	41584837	41836820	11	251983
	E1	E1_32X0	41848907	41866794	3	17887
	E1	E1_33X0	41885825	42116875	10	231050
	E1	E1_34X0	42150814	42409233	17	258419
	E1	E1_35X0	42566804	42585367	2	18563
	E1	E1_36X0	42640520	42695860	2	55340
	E1	E1_37X0	42702218	42801439	2	99221
	E1	E1_38X0	42838902	42968081	3	129179
	E1	E1_39X0	43165089	43207086	2	41997
	E1	E1_3X0	2062837	2068652	2	5815
	E1	E1_40X0	43345086	43373384	3	28298
	E1	E1_41X0	43982243	44067681	2	85438
	E1	E1_42X0	44083943	44093512	2	9569
	E1	E1_43X0	45505655	45599347	2	93692
	E1	E1_44X0	47943443	48020476	2	77033
	E1	E1_45X0	49718426	49881003	3	162577
	E1	E1_46X0	51980946	52216642	3	235696
	E1	E1_46X1	51790884	51952174	2	161290
	E1	E1_47X0	52888259	52932658	2	44399
	E1	E1_48X0	56566483	56603927	3	37444
	E1	E1_48X1	56722631	56820071	2	97440
	E1	E1_49X0	57008187	57029478	2	21291
	E1	E1_4X0	2088360	2146538	3	58178
	E1	E1_50X0	57782398	57798066	2	15668
	E1	E1_51X0	58347244	58401897	2	54653
	E1	E1_52X0	60691595	60749382	3	57787
	E1	E1_53X0	62384799	62395796	2	10997
	E1	E1_54X0	63189125	63213958	2	24833
	E1	E1_5X0	2344817	2349721	2	4904
	E1	E1_6X0	2454425	2465860	2	11435
	E1	E1_7X0	2855987	2913361	3	57374
	E1	E1_8X0	3465847	3603925	2	138078
	E1	E1_9X0	4628000	4939733	6	311733
	E2	E2_10X0	4965132	5189799	3	224667
	E2	E2_11X0	5232809	5272941	2	40132
	E2	E2_12X0	5444071	5618276	5	174205
	E2	E2_13X0	5795305	5963117	4	167812
	E2	E2_14X0	6051602	6075799	2	24197
	E2	E2_15X0	6117907	6510592	10	392685
	E2	E2_16X0	6568745	6750065	7	181320
	E2	E2_17X0	6783124	6836072	2	52948
	E2	E2_18X0	6935311	7280593	12	345282
	E2	E2_19X0	7307799	7492431	14	184632
	E2	E2_1X0	1169875	1446159	12	276284
	E2	E2_1X1	1447681	1812964	10	365283
	E2	E2_1X2	257716	651741	10	394025
	E2	E2_1X3	696310	1054460	7	358150
	E2	E2_20X0	8184301	8220235	2	35934
	E2	E2_21X0	8640978	8699570	2	58592
	E2	E2_22X0	8993440	9089126	2	95686
	E2	E2_23X0	10817737	10838735	2	20998
	E2	E2_24X0	11051895	11201608	4	149713
	E2	E2_25X0	11205584	11278816	2	73232
	E2	E2_26X0	11745946	12008840	6	262894

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	E2	E2_27X0	12773342	13029029	10	255687
	E2	E2_27X1	12514105	12756252	5	242147
	E2	E2_27X2	12029609	12472641	10	443032
	E2	E2_28X0	13183349	13311077	3	127728
	E2	E2_29X0	13787762	13991096	6	203334
	E2	E2_2X0	2074369	2714117	17	639748
	E2	E2_30X0	14267383	14279562	2	12179
	E2	E2_31X0	14302325	14624492	4	322167
	E2	E2_32X0	14874679	14919922	2	45243
	E2	E2_33X0	14955516	15151579	3	196063
	E2	E2_34X0	15262296	15279177	2	16881
	E2	E2_35X0	16737191	17112020	8	374829
	E2	E2_35X1	16038564	16378728	7	340164
	E2	E2_35X2	16399727	16713826	5	314099
	E2	E2_35X3	17199884	17571529	7	371645
	E2	E2_36X0	17634122	17749102	4	114980
	E2	E2_37X0	17764987	17793466	2	28479
	E2	E2_38X0	18151242	18219558	3	68316
	E2	E2_39X0	18224496	18263907	2	39411
	E2	E2_3X0	2726528	3016931	4	290403
	E2	E2_40X0	18365988	18413911	4	47923
	E2	E2_41X0	18542918	18635688	2	92770
	E2	E2_42X0	19771098	19847674	2	76576
	E2	E2_43X0	28414304	28577467	2	163163
	E2	E2_44X0	35202882	35308423	2	105541
	E2	E2_45X0	35869710	35924873	5	55163
	E2	E2_46X0	36482889	36526085	3	43196
	E2	E2_47X0	36642159	36758477	3	116318
	E2	E2_48X0	44402372	44542502	4	140130
	E2	E2_49X0	44740339	44786401	2	46062
	E2	E2_4X0	3073465	3132626	7	59161
	E2	E2_50X0	44908547	44953849	2	45302
	E2	E2_51X0	46189993	46356337	2	166344
	E2	E2_52X0	49982690	49999067	2	16377
	E2	E2_53X0	51009633	51065302	2	55669
	E2	E2_54X0	51835574	51862176	3	26602
	E2	E2_55X0	57575137	57695914	2	120777
	E2	E2_56X0	61213504	61266958	2	53454
	E2	E2_57X0	62674976	62895882	4	220906
	E2	E2_5X0	3154464	3358078	3	203614
	E2	E2_6X0	3446856	3465224	2	18368
	E2	E2_7X0	3560074	4126363	12	566289
	E2	E2_8X0	4149307	4222630	3	73323
	E2	E2_9X0	4420370	4779372	6	359002
	E3	E3_10X0	10280270	10388428	4	108158
	E3	E3_11X0	10879014	11196826	3	317812
	E3	E3_12X0	20096448	20315178	3	218730
	E3	E3_13X0	20444250	20512064	5	67814
	E3	E3_14X0	20527419	20561593	2	34174
	E3	E3_15X0	20641598	20648360	2	6762
	E3	E3_16X0	20734079	21060549	6	326470
	E3	E3_17X0	21181426	21210187	2	28761
	E3	E3_18X0	22832028	22918895	2	86867
	E3	E3_19X0	24657862	24707258	2	49396
	E3	E3_1X0	1040545	1071783	2	31238

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	E3	E3_20X0	25831393	25860463	2	29070
	E3	E3_21X0	26126915	26289528	2	162613
	E3	E3_22X0	27441040	27526578	2	85538
	E3	E3_23X0	28080805	28166691	2	85886
	E3	E3_24X0	28395579	28502345	2	106766
	E3	E3_25X0	28556936	28591493	2	34557
	E3	E3_26X0	29368696	29514954	2	146258
	E3	E3_27X0	31475930	31604466	2	128536
	E3	E3_28X0	41742143	42051116	12	308973
	E3	E3_29X0	42124804	42127476	2	2672
	E3	E3_2X0	4693461	4835396	4	141935
	E3	E3_30X0	42277351	42376378	6	99027
	E3	E3_31X0	42623725	42712784	2	89059
	E3	E3_32X0	42764289	42970804	3	206515
	E3	E3_33X0	43532762	43553618	3	20856
	E3	E3_34X0	44005981	44089788	4	83807
	E3	E3_35X0	44302051	44381359	2	79308
	E3	E3_36X0	44461815	44481309	6	19494
	E3	E3_3X0	6585372	6633422	2	48050
	E3	E3_4X0	6749879	6897168	5	147289
	E3	E3_5X0	6966460	7072376	2	105916
	E3	E3_6X0	7212875	7320663	3	107788
	E3	E3_7X0	7737167	7915493	2	178326
	E3	E3_8X0	9192637	9214262	2	21625
	E3	E3_9X0	9691595	9708715	2	17120
	F1	F1_10X0	37255005	37591991	6	336986
	F1	F1_11X0	37706049	37769177	3	63128
	F1	F1_12X0	37771136	37883379	4	112243
	F1	F1_13X0	37911708	37997598	2	85890
	F1	F1_14X0	38135159	38215948	2	80789
	F1	F1_15X0	42194974	42252965	2	57991
	F1	F1_16X0	42865612	43041114	2	175502
	F1	F1_17X0	43499264	43632861	2	133597
	F1	F1_18X0	43826053	43874113	2	48060
	F1	F1_19X0	43959161	44086165	3	127004
	F1	F1_1X0	2553327	2660844	2	107517
	F1	F1_20X0	45322841	45448610	2	125769
	F1	F1_21X0	45514028	45841739	3	327711
	F1	F1_22X0	45881175	45961396	2	80221
	F1	F1_23X0	46030209	46193419	2	163210
	F1	F1_24X0	59917693	59986440	2	68747
	F1	F1_25X0	61413457	61468306	2	54849
	F1	F1_26X0	65634589	65749549	4	114960
	F1	F1_27X0	65946660	66520586	9	573926
	F1	F1_28X0	66789784	66849014	2	59230
	F1	F1_29X0	66862244	67116521	8	254277
	F1	F1_2X0	4014243	4047355	2	33112
	F1	F1_30X0	67423576	67638683	8	215107
	F1	F1_31X0	67845999	68146645	13	300646
	F1	F1_32X0	68178530	68462798	8	284268
	F1	F1_33X0	68511308	68879024	6	367716
	F1	F1_34X0	68999281	69035663	2	36382
	F1	F1_35X0	69186516	69222290	2	35774
	F1	F1_36X0	69784275	69927054	3	142779
	F1	F1_37X0	70274303	70291839	2	17536

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	F1	F1_38X0	70416244	70478633	3	62389
	F1	F1_39X0	70490133	70518030	2	27897
	F1	F1_3X0	9713426	9807373	2	93947
	F1	F1_40X0	71470402	71542514	2	72112
	F1	F1_4X0	11748449	11933209	2	184760
	F1	F1_5X0	14188813	14294921	3	106108
	F1	F1_6X0	15454766	15651818	5	197052
	F1	F1_7X0	17907992	18046650	2	138658
	F1	F1_8X0	27899420	27946399	3	46979
	F1	F1_9X0	30172552	30256352	2	83800
	F2	F2_1X0	7731961	7963322	2	231361
	F2	F2_2X0	10302927	10451477	2	148550
	F2	F2_3X0	31876352	32176957	5	300605
	F2	F2_4X0	35390315	35599672	3	209357
	F2	F2_5X0	68381929	68644070	2	262141
	F2	F2_6X0	84118508	84403691	11	285183
	F2	F2_7X0	84413816	84719417	4	305601
	F2	F2_8X0	85076631	85186299	2	109668
	F2	F2_9X0	85228033	85732719	7	504686
	X	X_10X0	38676872	38874471	2	197599
	X	X_11X0	41810145	42488341	5	678196
	X	X_12X0	43168903	43194195	2	25292
	X	X_13X0	44776919	44942813	2	165894
	X	X_14X0	45393142	45712838	4	319696
	X	X_15X0	47800650	47940356	2	139706
	X	X_16X0	50197949	50318152	2	120203
	X	X_17X0	59522397	59698023	3	175626
	X	X_18X0	61815792	61848260	2	32468
	X	X_19X0	66528771	66993283	4	464512
	X	X_1X0	85872	131925	2	46053
	X	X_20X0	83743876	83822462	4	78586
	X	X_21X0	84261380	84284887	2	23507
	X	X_22X0	85146674	85485861	5	339187
	X	X_23X0	86272290	86313403	2	41113
	X	X_24X0	90344972	90579781	2	234809
	X	X_25X0	100039630	100086380	2	46750
	X	X_26X0	110675889	110820860	4	144971
	X	X_27X0	113469683	113536376	2	66693
	X	X_28X0	113897424	113983190	2	85766
	X	X_29X0	125894744	126048243	2	153499
	X	X_2X0	912887	993943	2	81056
	X	X_30X0	128653120	129125892	5	472772
	X	X_31X0	130003983	130014191	2	10208
	X	X_3X0	1064259	1225581	2	161322
	X	X_4X0	1755811	1872577	4	116766
	X	X_5X0	9893077	9951025	2	57948
	X	X_6X0	12064094	12124435	2	60341
	X	X_7X0	20642129	20663484	2	21355
	X	X_8X0	23702073	23751176	2	49103
	X	X_9X0	32701182	33120440	5	419258
	1	1_100X0	184789150	184894737	2	105587
	1	1_101X0	184982989	185376746	3	393757
	1	1_102X0	186247233	186509643	3	262410
	1	1_103X0	186544250	186678319	4	134069
	1	1_104X0	186680402	186746983	3	66581

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_105X0	186866591	187118298	4	251707
	1	1_106X0	187250586	187593042	3	342456
	1	1_107X0	200749767	200814708	2	64941
	1	1_108X0	203726153	203879412	2	153259
	1	1_109X0	205598534	205653469	2	54935
	1	1_10X0	15398872	15455989	3	57117
	1	1_110X0	207664221	207822171	2	157950
	1	1_111X0	208155016	208275631	5	120615
	1	1_112X0	227044723	227398287	5	353564
	1	1_113X0	228160044	228634145	14	474101
	1	1_113X1	227527719	228138332	14	610613
	1	1_11X0	16899654	17251281	5	351627
	1	1_12X0	19153418	19186776	2	33358
	1	1_13X0	19810757	20067199	5	256442
	1	1_14X0	20623092	20691803	2	68711
	1	1_15X0	21947991	21988452	2	40461
	1	1_16X0	22630688	22655871	3	25183
	1	1_17X0	24151694	24223040	2	71346
	1	1_18X0	25473284	25475893	2	2609
	1	1_19X0	27480226	27751002	2	270776
	1	1_1X0	1068880	1304565	5	235685
	1	1_20X0	31786077	31923700	2	137623
	1	1_21X0	32384581	32538776	3	154195
	1	1_22X0	32639255	32773713	2	134458
	1	1_23X0	32794971	32825324	2	30353
	1	1_24X0	33515808	33738180	2	222372
	1	1_25X0	35051948	35056675	2	4727
	1	1_26X0	35281435	35414157	2	132722
	1	1_27X0	36121949	36380381	3	258432
	1	1_28X0	39895865	40051327	2	155462
	1	1_29X0	40927036	41031218	3	104182
	1	1_2X0	6222143	6245067	2	22924
	1	1_30X0	42649687	42661031	2	11344
	1	1_31X0	43674800	43866888	2	192088
	1	1_32X0	44638714	44661793	2	23079
	1	1_33X0	45545326	45592616	2	47290
	1	1_34X0	46181537	46212275	2	30738
	1	1_35X0	47414880	47736260	5	321380
	1	1_36X0	53382932	53495653	2	112721
	1	1_37X0	54647355	54673243	2	25888
	1	1_38X0	68045324	68238772	2	193448
	1	1_39X0	79603445	79647470	2	44025
	1	1_3X0	8679346	8744665	2	65319
	1	1_40X0	85887876	86009378	2	121502
	1	1_41X0	87409483	87589876	3	180393
	1	1_42X0	90096914	90506317	8	409403
	1	1_43X0	90638809	91018357	3	379548
	1	1_44X0	104984571	105071528	2	86957
	1	1_45X0	109658623	109802317	2	143694
	1	1_46X0	110923482	111000192	2	76710
	1	1_47X0	111059109	111132146	3	73037
	1	1_48X0	111956853	112114771	3	157918
	1	1_49X0	112666480	112890435	5	223955
	1	1_4X0	9551843	9648400	2	96557
	1	1_50X0	118020906	118276880	3	255974

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_51X0	118458217	118694057	3	235840
	1	1_52X0	120942748	121142972	3	200224
	1	1_53X0	125511476	125790731	4	279255
	1	1_54X0	127005347	127158804	2	153457
	1	1_55X0	128101934	128133148	2	31214
	1	1_56X0	128764455	128868553	12	104098
	1	1_57X0	129499385	129526989	2	27604
	1	1_58X0	129749179	129820120	2	70941
	1	1_59X0	130032742	130054506	2	21764
	1	1_5X0	11385916	11413306	3	27390
	1	1_60X0	130286534	130441134	4	154600
	1	1_61X0	130724941	130744584	2	19643
	1	1_62X0	130891203	130961446	2	70243
	1	1_63X0	131568703	132063187	9	494484
	1	1_64X0	132360850	132413077	2	52227
	1	1_65X0	134098485	134130996	2	32511
	1	1_66X0	134143409	134211552	2	68143
	1	1_67X0	134685362	134839501	2	154139
	1	1_68X0	134983547	135165615	2	182068
	1	1_69X0	135929098	135970630	2	41532
	1	1_6X0	11586596	11601084	2	14488
	1	1_70X0	136185432	136231871	2	46439
	1	1_71X0	136597590	136886922	5	289332
	1	1_72X0	137146959	137443827	6	296868
	1	1_73X0	137484059	137904965	11	420906
	1	1_74X0	137929439	138177826	4	248387
	1	1_75X0	138295799	138653037	8	357238
	1	1_76X0	138924802	139083823	5	159021
	1	1_77X0	139106097	139171691	2	65594
	1	1_78X0	139496271	140114931	9	618660
	1	1_79X0	140434029	140659572	4	225543
	1	1_7X0	12408503	12495786	2	87283
	1	1_80X0	141419779	141718128	2	298349
	1	1_81X0	142007102	142151266	2	144164
	1	1_82X0	145901533	146052124	2	150591
	1	1_83X0	147570938	147609073	2	38135
	1	1_84X0	148608442	148753811	3	145369
	1	1_85X0	150093657	150347246	5	253589
	1	1_86X0	150794732	150830015	2	35283
	1	1_87X0	152798189	152958465	2	160276
	1	1_88X0	159040687	159116711	2	76024
	1	1_89X0	161681673	161902939	3	221266
	1	1_8X0	12541679	12863937	10	322258
	1	1_90X0	162262832	162494289	2	231457
	1	1_91X0	165729199	165748380	2	19181
	1	1_92X0	171997674	172212864	2	215190
	1	1_93X0	176400507	176566962	3	166455
	1	1_94X0	180507279	180524702	2	17423
	1	1_95X0	181027665	181083232	2	55567
	1	1_96X0	181863476	182106833	4	243357
	1	1_97X0	182972854	183020187	2	47333
	1	1_98X0	183140900	183308555	3	167655
	1	1_99X0	183483343	183557705	2	74362
	1	1_9X0	12935077	13369978	9	434901
	10	10_10X0	64338078	64341526	2	3448

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	10	10_11X0	66447978	66673071	4	225093
	10	10_12X0	82561574	82859284	4	297710
	10	10_13X0	93677748	94185454	3	507706
	10	10_14X0	94556477	94588656	2	32179
	10	10_15X0	98521994	98543103	2	21109
	10	10_16X0	101235470	101297399	2	61929
	10	10_17X0	101506073	101649652	2	143579
	10	10_18X0	102919024	103780336	11	861312
	10	10_19X0	107452837	107469102	2	16265
	10	10_1X0	1013608	1044673	2	31065
	10	10_20X0	109129279	109557130	4	427851
	10	10_21X0	111960146	111980726	2	20580
	10	10_22X0	115778622	115871133	2	92511
	10	10_23X0	116165158	116342041	2	176883
	10	10_24X0	117912593	117919169	2	6576
	10	10_25X0	118051300	118060125	2	8825
	10	10_26X0	118110574	118143692	3	33118
	10	10_27X0	118920863	118969125	2	48262
	10	10_28X0	120386795	120392219	2	5424
	10	10_29X0	131419015	131618779	3	199764
	10	10_2X0	4884680	5335138	5	450458
	10	10_30X0	137905786	138020609	4	114823
	10	10_31X0	138206846	138221375	2	14529
	10	10_32X0	139870881	139911994	2	41113
	10	10_3X0	5686105	5937579	2	251474
	10	10_4X0	7646890	7827007	2	180117
	10	10_5X0	34645510	34707560	2	62050
	10	10_6X0	40479578	40658257	2	178679
	10	10_7X0	40672753	40978561	3	305808
	10	10_8X0	58481224	58491931	2	10707
	10	10_9X0	62811730	62838793	2	27063
	11	11_10X0	3961882	3973686	2	11804
	11	11_11X0	5294550	5992462	21	697912
	11	11_11X1	4840189	5259191	19	419002
	11	11_11X2	4073739	4766186	20	692447
	11	11_11X3	6491593	6659503	8	167910
	11	11_12X0	6669799	6744555	2	74756
	11	11_13X0	54927146	55733930	24	806784
	11	11_13X1	57943378	58857398	15	914020
	11	11_13X10	70849497	71262277	6	412780
	11	11_13X11	18105458	18463133	6	357675
	11	11_13X12	55745271	56092791	5	347520
	11	11_13X13	63734250	64025141	4	290891
	11	11_13X14	62116803	62371704	4	254901
	11	11_13X15	65749605	65926005	4	176400
	11	11_13X16	18829277	18922969	3	93692
	11	11_13X17	50850132	50907674	3	57542
	11	11_13X18	68987303	69111404	3	124101
	11	11_13X19	7581542	7715346	3	133804
	11	11_13X2	57160839	57835753	12	674914
	11	11_13X20	34618539	34798605	2	180066
	11	11_13X21	14811960	14928930	2	116970
	11	11_13X22	68269947	68296253	2	26306
	11	11_13X23	44350334	44386623	2	36289
	11	11_13X24	49264263	49292436	2	28173

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	11	11_13X25	72315077	72379026	2	63949
	11	11_13X26	64730635	64822534	2	91899
	11	11_13X27	66484809	66669025	2	184216
	11	11_13X28	54446466	54806456	3	359990
	11	11_13X3	59490572	59909462	7	418890
	11	11_13X4	59919452	60387159	7	467707
	11	11_13X5	60932229	61560674	9	628445
	11	11_13X6	62521941	62998336	7	476395
	11	11_13X7	56309356	56641596	7	332240
	11	11_13X8	48515452	48802334	6	286882
	11	11_13X9	58970563	59462027	5	491464
	11	11_14X0	73059358	73093952	2	34594
	11	11_15X0	75732021	75870916	2	138895
	11	11_16X0	77042968	77203802	2	160834
	11	11_17X0	85178723	85206857	2	28134
	11	11_18X0	88612419	88727073	2	114654
	11	11_19X0	89410943	89500180	3	89237
	11	11_1X0	104252	117753	3	13501
	11	11_20X0	102077856	102134545	2	56689
	11	11_21X0	102321124	102705794	8	384670
	11	11_22X0	104671621	104849449	4	177828
	11	11_23X0	113869863	113962605	2	92742
	11	11_24X0	114505708	114659534	3	153826
	11	11_25X0	116796453	116843583	3	47130
	11	11_26X0	117834035	117856772	2	22737
	11	11_27X0	117922384	118118758	2	196374
	11	11_28X0	118138730	118270866	5	132136
	11	11_29X0	118313335	118348321	2	34986
	11	11_2X0	293254	608911	4	315657
	11	11_30X0	119172696	119195327	2	22631
	11	11_31X0	119305692	119424017	2	118325
	11	11_32X0	121343737	121367956	2	24219
	11	11_33X0	122935460	123197470	2	262010
	11	11_34X0	123867475	124696550	21	829075
	11	11_35X0	124888241	125078044	5	189803
	11	11_36X0	125832001	125960592	2	128591
	11	11_37X0	126114846	126231002	2	116156
	11	11_38X0	129028993	129114306	2	85313
	11	11_39X0	130605495	130676683	2	71188
	11	11_3X0	641912	673069	2	31157
	11	11_40X0	134111063	134349511	2	238448
	11	11_41X0	134474130	134573653	2	99523
	11	11_4X0	771748	844009	2	72261
	11	11_5X0	846844	928711	2	81867
	11	11_6X0	1642689	1742193	2	99504
	11	11_7X0	1931891	1963618	2	31727
	11	11_8X0	2996471	3007658	2	11187
	11	11_9X0	3390565	3415436	2	24871
	12	12_10X0	10214946	10439392	5	224446
	12	12_11X0	10648697	11032126	9	383429
	12	12_12X0	11114412	11245471	3	131059
	12	12_13X0	12721571	12782507	2	60936
	12	12_14X0	12904369	13072928	3	168559
	12	12_15X0	14712115	14729798	2	17683
	12	12_16X0	20599676	20983682	3	384006

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	12	12_17X0	28664920	28688525	2	23605
	12	12_18X0	29174236	29387729	3	213493
	12	12_19X0	29790359	29851921	2	61562
	12	12_1X0	156132	230141	2	74009
	12	12_20X0	30005565	30191878	5	186313
	12	12_21X0	30496240	30989281	12	493041
	12	12_22X0	31501800	31519306	2	17506
	12	12_23X0	32113648	32231054	9	117406
	12	12_24X0	32685912	32752425	2	66513
	12	12_25X0	32765037	32992434	3	227397
	12	12_26X0	33223150	33961620	24	738470
	12	12_27X0	34142951	34224549	2	81598
	12	12_28X0	36195038	36218301	3	23263
	12	12_29X0	36265031	36419112	2	154081
	12	12_2X0	4391686	4467138	2	75452
	12	12_30X0	36901946	37037087	3	135141
	12	12_31X0	37090218	37302110	4	211892
	12	12_32X0	39850602	40023510	2	172908
	12	12_33X0	63791173	63933563	2	142390
	12	12_34X0	67309869	67361634	2	51765
	12	12_35X0	68880022	68881683	2	1661
	12	12_36X0	74626436	74788815	3	162379
	12	12_37X0	80013808	80025869	2	12061
	12	12_38X0	90499866	90648291	3	148425
	12	12_39X0	112845314	112963646	3	118332
	12	12_3X0	4829030	5071922	3	242892
	12	12_40X0	113344408	113390270	2	45862
	12	12_41X0	121211171	121302078	2	90907
	12	12_42X0	133170794	133414818	7	244024
	12	12_4X0	6698208	6905916	4	207708
	12	12_5X0	7133698	7178912	3	45214
	12	12_6X0	7874547	8025123	4	150576
	12	12_7X0	8564856	8651044	3	86188
	12	12_8X0	8938095	9303095	3	365000
	12	12_9X0	9445082	9944950	12	499868
	13	13_10X0	82827315	82847694	2	20379
	13	13_11X0	85627500	85856149	2	228649
	13	13_12X0	96015577	96078688	3	63111
	13	13_1X0	1098648	1103032	2	4384
	13	13_2X0	1690101	1771967	3	81866
	13	13_3X0	9431907	9588738	3	156831
	13	13_4X0	12534381	12603314	2	68933
	13	13_5X0	22924052	22989119	2	65067
	13	13_6X0	23116760	23184975	2	68215
	13	13_7X0	30318819	30578295	2	259476
	13	13_8X0	34037294	34126374	2	89080
	13	13_9X0	82099225	82148632	2	49407
	14	14_10X0	32994964	33086441	2	91477
	14	14_11X0	33651649	33852863	2	201214
	14	14_12X0	35465624	35526487	2	60863
	14	14_13X0	37168208	37184774	2	16566
	14	14_14X0	47836549	47878131	2	41582
	14	14_15X0	51133365	51179834	2	46469
	14	14_16X0	53809711	54126929	4	317218
	14	14_17X0	57180341	57243517	3	63176

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	14	14_18X0	58550434	58766011	2	215577
	14	14_19X0	59081650	59190424	2	108774
	14	14_1X0	8709925	8768409	2	58484
	14	14_20X0	65082851	65109532	2	26681
	14	14_21X0	78334438	78678854	10	344416
	14	14_22X0	79781333	79811754	2	30421
	14	14_23X0	80301447	80366268	2	64821
	14	14_24X0	84306856	84345125	2	38269
	14	14_25X0	90247840	90828992	8	581152
	14	14_2X0	20760987	20884413	4	123426
	14	14_3X0	20953232	21001452	2	48220
	14	14_4X0	21058769	21727642	10	668873
	14	14_5X0	21940867	21992670	2	51803
	14	14_6X0	22329131	22342924	2	13793
	14	14_7X0	23200140	23291456	2	91316
	14	14_8X0	23702761	23803924	4	101163
	14	14_9X0	25152864	25611185	15	458321
	14	14_9X1	24349163	24842740	12	493577
	15	15_10X0	26970830	27106646	2	135816
	15	15_11X0	28637637	28744230	2	106593
	15	15_12X0	29988285	30137637	2	149352
	15	15_13X0	30428628	30683343	2	254715
	15	15_14X0	33459647	33508627	2	48980
	15	15_15X0	37786078	37794190	2	8112
	15	15_16X0	43804986	43900855	2	95869
	15	15_17X0	46734775	46786185	2	51410
	15	15_18X0	52268758	52421674	2	152916
	15	15_19X0	52862374	52885931	2	23557
	15	15_1X0	1871450	1929036	2	57586
	15	15_20X0	53339023	53563070	5	224047
	15	15_21X0	53564393	53718677	2	154284
	15	15_22X0	57338452	57412171	3	73719
	15	15_23X0	67220771	67255120	2	34349
	15	15_24X0	68353778	68386143	2	32365
	15	15_2X0	18163623	18277915	2	114292
	15	15_3X0	19490460	19566372	2	75912
	15	15_4X0	19763021	19862283	2	99262
	15	15_5X0	20174459	20349731	2	175272
	15	15_6X0	21118747	21297032	2	178285
	15	15_7X0	21355870	21460879	3	105009
	15	15_8X0	21756204	21859889	2	103685
	15	15_9X0	23251686	23324655	4	72969
	16	16_10X0	3221471	3581880	9	360409
	16	16_11X0	4628179	4678109	2	49930
	16	16_12X0	16447601	16704651	2	257050
	16	16_13X0	18733669	18778559	2	44890
	16	16_14X0	20652272	20691809	2	39537
	16	16_15X0	20752280	21132713	5	380433
	16	16_16X0	23411279	23621817	2	210538
	16	16_17X0	23922439	23956989	2	34550
	16	16_18X0	27710282	27808553	2	98271
	16	16_19X0	29431814	29639302	2	207488
	16	16_1X0	29960	66564	4	36604
	16	16_20X0	30124829	30268793	2	143964
	16	16_21X0	30501330	30903024	9	401694

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	16	16_22X0	31114497	31161936	2	47439
	16	16_23X0	31183534	31206774	2	23240
	16	16_24X0	31207059	31273488	3	66429
	16	16_25X0	31388600	31560024	3	171424
	16	16_26X0	31679261	31859620	2	180359
	16	16_27X0	33816901	33828339	2	11438
	16	16_28X0	37998031	38016274	2	18243
	16	16_29X0	38527476	38676978	2	149502
	16	16_2X0	171328	207705	4	36377
	16	16_30X0	46348813	46466620	3	117807
	16	16_31X0	47190140	47313665	10	123525
	16	16_32X0	48003244	48059749	3	56505
	16	16_33X0	48192497	48336322	3	143825
	16	16_34X0	48834512	48872427	2	37915
	16	16_35X0	55176590	55203205	2	26615
	16	16_36X0	57537624	57611930	3	74306
	16	16_37X0	57915245	58081900	4	166655
	16	16_38X0	58250606	58302003	2	51397
	16	16_39X0	59070939	59088770	2	17831
	16	16_3X0	298609	369436	2	70827
	16	16_40X0	59674589	59849412	2	174823
	16	16_41X0	60981151	61213173	2	232022
	16	16_42X0	61323656	61401466	2	77810
	16	16_43X0	62503526	62544821	2	41295
	16	16_44X0	63141258	63178775	3	37517
	16	16_45X0	66332062	66353619	2	21557
	16	16_46X0	66604288	66665033	2	60745
	16	16_47X0	66798294	66866208	3	67914
	16	16_48X0	75370942	75425930	2	54988
	16	16_49X0	75962194	76037210	2	75016
	16	16_4X0	614415	711658	4	97243
	16	16_50X0	77834025	77846416	2	12391
	16	16_5X0	786044	809672	2	23628
	16	16_6X0	1260878	1278733	2	17855
	16	16_7X0	2021616	2280012	3	258396
	16	16_8X0	2689205	2750978	2	61773
	16	16_9X0	2822185	2976354	4	154169
	17	17_10X0	7041738	7300669	5	258931
	17	17_11X0	7367864	7377147	4	9283
	17	17_12X0	7440970	7448388	2	7418
	17	17_13X0	8069702	8150205	3	80503
	17	17_14X0	8854086	8963039	2	108953
	17	17_15X0	10370490	10734448	6	363958
	17	17_16X0	17339857	17396680	2	56823
	17	17_17X0	19681513	19902313	3	220800
	17	17_18X0	26118104	26247207	2	129103
	17	17_19X0	39962901	40097568	2	134667
	17	17_1X0	1202481	1268064	2	65583
	17	17_20X0	40539852	40709054	2	169202
	17	17_21X0	40859261	40949127	2	89866
	17	17_22X0	42941878	43166091	2	224213
	17	17_23X0	59280414	59387335	2	106921
	17	17_24X0	77941703	77997579	2	55876
	17	17_25X0	80877855	81016950	2	139095
	17	17_26X0	92182629	92241936	2	59307

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	17	17_27X0	94396753	94643076	3	246323
	17	17_2X0	1615736	1650875	2	35139
	17	17_3X0	2989718	3308152	10	318434
	17	17_4X0	3400070	3479629	2	79559
	17	17_5X0	3556082	3795619	4	239537
	17	17_6X0	4366321	4473396	2	107075
	17	17_7X0	4610466	4963294	3	352828
	17	17_8X0	5033443	5129968	3	96525
	17	17_9X0	6895662	7016451	2	120789
	18	18_10X0	72364631	72448861	2	84230
	18	18_1X0	906961	967760	2	60799
	18	18_2X0	2556958	2561358	2	4400
	18	18_3X0	11753887	11779930	2	26043
	18	18_4X0	14332804	14510222	2	177418
	18	18_5X0	27752079	28322305	7	570226
	18	18_6X0	32080359	32214624	4	134265
	18	18_7X0	42621794	42780100	2	158306
	18	18_8X0	44078511	44160465	4	81954
	18	18_9X0	60983214	61525298	11	542084
	19	19_10X0	5599182	5636089	3	36907
	19	19_11X0	6215680	6245735	2	30055
	19	19_12X0	6322404	6469650	2	147246
	19	19_13X0	6823061	6836162	2	13101
	19	19_14X0	7506826	7732364	5	225538
	19	19_15X0	8915448	9034650	5	119202
	19	19_16X0	56577417	57094534	21	517117
	19	19_16X1	50104673	50393453	15	288780
	19	19_16X10	57526734	57916477	10	389743
	19	19_16X11	57184541	57468645	9	284104
	19	19_16X12	11343250	11673323	8	330073
	19	19_16X13	43497634	43784334	8	286700
	19	19_16X14	55326722	55566973	7	240251
	19	19_16X15	11916737	12139250	5	222513
	19	19_16X16	14398535	14729336	9	330801
	19	19_16X17	55726913	55974656	8	247743
	19	19_16X18	53836711	54104352	7	267641
	19	19_16X19	42586776	42895933	8	309157
	19	19_16X2	50516144	50834699	12	318555
	19	19_16X20	9150600	9542363	8	391763
	19	19_16X21	43904899	44182885	7	277986
	19	19_16X22	36887789	37276086	8	388297
	19	19_16X23	36178349	36410290	5	231941
	19	19_16X24	15440496	15806510	8	366014
	19	19_16X25	35333513	35538939	7	205426
	19	19_16X26	40383499	40680301	7	296802
	19	19_16X27	50921465	51212729	6	291264
	19	19_16X28	12145814	12397162	5	251348
	19	19_16X29	40975494	41214265	6	238771
	19	19_16X3	52412718	52918044	13	505326
	19	19_16X30	9889273	10107424	6	218151
	19	19_16X31	53548788	53664553	4	115765
	19	19_16X32	52065684	52322793	6	257109
	19	19_16X33	34803641	34962319	6	158678
	19	19_16X34	42944009	43109819	4	165810
	19	19_16X35	39045471	39243311	5	197840

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	19	19_16X36	34019992	34260121	5	240129
	19	19_16X37	38299745	38557073	5	257328
	19	19_16X38	39533040	39744026	4	210986
	19	19_16X39	35887101	36133665	5	246564
	19	19_16X4	51214024	51509777	9	295753
	19	19_16X40	39762666	39972899	4	210233
	19	19_16X41	53256500	53418096	5	161596
	19	19_16X42	55126397	55315710	4	189313
	19	19_16X43	36418847	36878358	5	459511
	19	19_16X44	19732378	20083138	5	350760
	19	19_16X45	38765937	38822300	4	56363
	19	19_16X46	49209270	49281358	4	72088
	19	19_16X47	35622476	35762953	3	140477
	19	19_16X48	45300192	45491357	4	191165
	19	19_16X49	20763144	21104086	5	340942
	19	19_16X5	11686551	11910496	7	223945
	19	19_16X50	47071186	47268067	4	196881
	19	19_16X51	48313219	48351514	4	38295
	19	19_16X52	34614066	34652446	3	38380
	19	19_16X53	22366580	22482909	3	116329
	19	19_16X54	21896293	22144914	3	248621
	19	19_16X55	15161397	15364746	3	203349
	19	19_16X56	42138654	42298621	2	159967
	19	19_16X57	22624063	22900743	3	276680
	19	19_16X58	46573265	46595210	2	21945
	19	19_16X59	21254205	21384575	2	130370
	19	19_16X6	43175359	43475037	10	299678
	19	19_16X60	19075110	19125928	2	50818
	19	19_16X61	49647565	49679509	2	31944
	19	19_16X62	41929022	42019668	2	90646
	19	19_16X63	14157588	14248170	2	90582
	19	19_16X64	20437343	20543242	2	105899
	19	19_16X65	44848332	44944340	2	96008
	19	19_16X66	19532957	19595137	2	62180
	19	19_16X67	23297164	23418589	2	121425
	19	19_16X68	32739685	32816927	2	77242
	19	19_16X69	34424188	34456983	2	32795
	19	19_16X7	54851290	55070073	9	218783
	19	19_16X70	47958710	48008565	2	49855
	19	19_16X71	41481597	41535610	2	54013
	19	19_16X72	56148228	56549610	3	401382
	19	19_16X8	51615655	51934629	9	318974
	19	19_16X9	54128448	54397102	8	268654
	19	19_1X0	324937	473205	2	148268
	19	19_2X0	611511	648094	4	36583
	19	19_4X0	2597633	2707906	4	110273
	19	19_5X0	2748397	2806890	2	58493
	19	19_6X0	2870710	2925816	2	55106
	19	19_7X0	3492383	3525588	2	33205
	19	19_8X0	3724465	3893199	2	168734
	19	19_9X0	5349596	5436347	2	86751
	20	20_10X0	30478343	30771273	2	292930
	20	20_11X0	31664618	31990833	5	326215
	20	20_12X0	34314298	34367076	2	52778
	20	20_13X0	37091765	37147881	2	56116

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	20	20_14X0	43938945	44086939	6	147994
	20	20_15X0	44306416	44655149	7	348733
	20	20_16X0	44723508	44749957	2	26449
	20	20_17X0	58922868	59010805	2	87937
	20	20_18X0	62479254	62498821	2	19567
	20	20_19X0	63039346	63056643	2	17297
	20	20_1X0	22737	223986	5	201249
	20	20_2X0	349686	474121	2	124435
	20	20_3X0	1448710	1908780	6	460070
	20	20_4X0	2950513	2964188	2	13675
	20	20_5X0	4570797	4589365	2	18568
	20	20_6X0	18226994	18518507	2	291513
	20	20_7X0	21671535	21792242	2	120707
	20	20_8X0	23740828	24182793	7	441965
	20	20_9X0	29897322	30094875	6	197553
	21	21_10X0	33548967	33603568	3	54601
	21	21_11X0	34773996	34921691	2	147695
	21	21_12X0	35458245	35587372	2	129127
	21	21_13X0	35692929	35849616	8	156687
	21	21_14X0	37166968	37318533	2	151565
	21	21_1X0	4696573	4701403	2	4830
	21	21_2X0	17896808	18026374	2	129566
	21	21_3X0	19933142	20117847	2	184705
	21	21_4X0	21252774	21310318	2	57544
	21	21_5X0	21356108	21511388	5	155280
	21	21_6X0	21564882	21707489	9	142607
	21	21_7X0	24317513	24505843	4	188330
	21	21_8X0	27203858	27277297	2	73439
	21	21_9X0	32613626	32705029	2	91403
	22	22_10X0	20569110	20747125	2	178015
	22	22_11X0	21175155	21284217	2	109062
	22	22_12X0	22508573	22619164	4	110591
	22	22_13X0	24759777	24810348	2	50571
	22	22_14X0	26539658	26684498	2	144840
	22	22_15X0	27464679	27532617	2	67938
	22	22_16X0	27662829	27772399	2	109570
	22	22_17X0	34153713	34167851	2	14138
	22	22_1X0	3848322	3896558	2	48236
	22	22_2X0	4633306	4678981	2	45675
	22	22_3X0	5818657	5846432	2	27775
	22	22_4X0	5869466	5889900	2	20434
	22	22_5X0	6417219	6504658	2	87439
	22	22_6X0	7268141	7436161	7	168020
	22	22_7X0	13932506	14078833	4	146327
	22	22_8X0	19271590	19353852	2	82262
	22	22_9X0	19782211	19917759	4	135548
	2A	2A_10X0	17794619	18171912	4	377293
	2A	2A_11X0	18933515	18968364	2	34849
	2A	2A_12X0	19885566	19958718	2	73152
	2A	2A_13X0	21064904	21106165	2	41261
	2A	2A_14X0	27669671	27807990	4	138319
	2A	2A_15X0	36036133	36062069	2	25936
	2A	2A_16X0	37916809	37952406	2	35597
	2A	2A_17X0	38494235	38547784	2	53549
	2A	2A_18X0	62175613	62242689	2	67076

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	2A	2A_19X0	63332441	63389165	2	56724
	2A	2A_1X0	463901	532155	3	68254
	2A	2A_20X0	67896747	68081011	2	184264
	2A	2A_21X0	69914323	70125026	2	210703
	2A	2A_22X0	79774031	80017008	2	242977
	2A	2A_23X0	83264809	83335049	2	70240
	2A	2A_2X0	1370785	1409653	2	38868
	2A	2A_3X0	1900636	1943009	2	42373
	2A	2A_4X0	1949933	1968308	2	18375
	2A	2A_5X0	4132607	4193937	2	61330
	2A	2A_6X0	6951648	7443838	6	492190
	2A	2A_7X0	7466704	7702681	2	235977
	2A	2A_8X0	13218999	13393424	3	174425
	2A	2A_9X0	17244477	17306503	3	62026
	2B	2B_10X0	64485103	64502256	2	17153
	2B	2B_11X0	68486581	68630445	11	143864
	2B	2B_12X0	77619087	77812981	2	193894
	2B	2B_13X0	81687547	81799010	2	111463
	2B	2B_14X0	83723909	83901034	2	177125
	2B	2B_15X0	93909049	94043674	3	134625
	2B	2B_16X0	94204846	94693634	3	488788
	2B	2B_17X0	95021371	95028305	2	6934
	2B	2B_18X0	96552669	96808970	3	256301
	2B	2B_19X0	109228395	109269313	2	40918
	2B	2B_1X0	733132	881411	2	148279
	2B	2B_20X0	109861293	109899143	2	37850
	2B	2B_21X0	111269406	111472663	4	203257
	2B	2B_22X0	111884086	111922421	2	38335
	2B	2B_23X0	112080931	112115383	2	34452
	2B	2B_24X0	112473812	112711184	4	237372
	2B	2B_25X0	112714757	112798568	2	83811
	2B	2B_26X0	123557011	123734899	2	177888
	2B	2B_27X0	124262426	124898974	3	636548
	2B	2B_28X0	125758711	125809459	2	50748
	2B	2B_29X0	125874343	125893714	2	19371
	2B	2B_2X0	1269164	1644200	10	375036
	2B	2B_30X0	134043267	134062445	2	19178
	2B	2B_3X0	7734515	7956784	2	222269
	2B	2B_4X0	41266974	41533536	2	266562
	2B	2B_5X0	43356696	43397074	2	40378
	2B	2B_6X0	49405563	50002307	3	596744
	2B	2B_7X0	54239107	54489535	2	250428
	2B	2B_8X0	61868052	62116657	2	248605
	2B	2B_9X0	62172607	62195168	2	22561
	3	3_10X0	45265131	45552342	6	287211
	3	3_11X0	46783729	47300304	8	516575
	3	3_12X0	47655072	47686294	3	31222
	3	3_13X0	50404879	50425387	2	20508
	3	3_14X0	51057382	51129835	2	72453
	3	3_15X0	51169411	51286928	4	117517
	3	3_16X0	51309481	51328866	2	19385
	3	3_17X0	52806362	52929585	6	123223
	3	3_18X0	52953219	53302082	3	348863
	3	3_19X0	53828570	53881760	3	53190
	3	3_1X0	10176034	10206711	2	30677

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	3	3_20X0	54272189	54349355	3	77166
	3	3_21X0	99302580	99765083	10	462503
	3	3_22X0	102976142	102981064	2	4922
	3	3_23X0	109767928	109979554	2	211626
	3	3_24X0	110156864	110179944	2	23080
	3	3_25X0	110737626	110781448	2	43822
	3	3_26X0	114320254	114491307	2	171053
	3	3_27X0	115771263	115919857	2	148594
	3	3_28X0	116855042	116892859	2	37817
	3	3_29X0	120874778	120918550	2	43772
	3	3_2X0	10267205	10323785	2	56580
	3	3_30X0	124235974	124449426	2	213452
	3	3_31X0	126642606	126791469	2	148863
	3	3_32X0	129495007	129822391	2	327384
	3	3_33X0	131897372	131935028	2	37656
	3	3_34X0	133002207	133059653	2	57446
	3	3_35X0	135868335	135948154	2	79819
	3	3_36X0	141087328	141131829	3	44501
	3	3_37X0	141542345	141630184	2	87839
	3	3_38X0	148376148	148789426	4	413278
	3	3_39X0	149600290	149623572	2	23282
	3	3_3X0	24479882	24730334	2	250452
	3	3_40X0	151095592	151162258	2	66666
	3	3_41X0	151546788	151730439	3	183651
	3	3_42X0	153510788	153650595	3	139807
	3	3_43X0	154066703	154137017	2	70314
	3	3_44X0	161202355	161268240	2	65885
	3	3_45X0	162835921	163018471	2	182550
	3	3_46X0	163938953	163984031	2	45078
	3	3_47X0	172589686	172666022	3	76336
	3	3_48X0	186584006	186785566	2	201560
	3	3_49X0	187157330	187229838	3	72508
	3	3_4X0	32771026	33042427	3	271401
	3	3_50X0	189788560	189920037	3	131477
	3	3_51X0	197710995	197733088	2	22093
	3	3_5X0	38952428	39004937	2	52509
	3	3_6X0	39244887	39719821	3	474934
	3	3_7X0	40037037	40099571	2	62534
	3	3_8X0	41255581	41295781	2	40200
	3	3_9X0	43406551	43678793	3	272242
	4	4_10X0	59310267	60100129	5	789862
	4	4_11X0	60139978	60804669	5	664691
	4	4_12X0	80141943	80430162	4	288219
	4	4_13X0	80506409	80638953	3	132544
	4	4_14X0	80711724	80740131	2	28407
	4	4_15X0	80775719	80837981	3	62262
	4	4_16X0	81135011	81183148	3	48137
	4	4_17X0	84265400	84361954	3	96554
	4	4_18X0	84614199	84966288	10	352089
	4	4_19X0	86946222	86980828	3	34606
	4	4_1X0	41753	449548	5	407795
	4	4_20X0	87999506	88127997	2	128491
	4	4_21X0	93131712	93437674	3	305962
	4	4_22X0	98419615	98502624	2	83009
	4	4_23X0	99517515	99882561	3	365046

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	4	4_24X0	110355298	110711962	7	356664
	4	4_25X0	114270967	114446476	2	175509
	4	4_26X0	124135386	124167377	2	31991
	4	4_27X0	167084179	167156666	2	72487
	4	4_28X0	168207983	168347753	2	139770
	4	4_29X0	176031238	176058862	2	27624
	4	4_2X0	3498766	3556560	2	57794
	4	4_30X0	177664335	177753784	3	89449
	4	4_31X0	180955013	181221947	2	266934
	4	4_32X0	182758003	182793780	2	35777
	4	4_33X0	196296827	196299695	2	2868
	4	4_34X0	199100365	199157181	2	56816
	4	4_35X0	200972406	201029196	2	56790
	4	4_3X0	5697158	5962554	2	265396
	4	4_4X0	6798756	6868255	3	69499
	4	4_5X0	8752131	9072334	2	320203
	4	4_6X0	15225883	15364913	2	139030
	4	4_7X0	15455919	15483189	2	27270
	4	4_8X0	47940115	48104391	2	164276
	4	4_9X0	48716292	48789785	2	73493
	5	5_10X0	16094158	16315505	3	221347
	5	5_11X0	19365720	19385424	2	19704
	5	5_12X0	19557069	19637599	2	80530
	5	5_13X0	19822696	19841156	2	18460
	5	5_14X0	21809156	21892636	2	83480
	5	5_15X0	25109029	25178955	2	69926
	5	5_16X0	32854738	32891106	2	36368
	5	5_17X0	32929201	33018945	2	89744
	5	5_18X0	33317004	33355051	2	38047
	5	5_19X0	33472824	33498882	2	26058
	5	5_1X0	355053	393037	2	37984
	5	5_20X0	34749927	34940859	9	190932
	5	5_21X0	35577121	35650684	2	73563
	5	5_22X0	36082748	36318618	2	235870
	5	5_23X0	36598335	36706419	2	108084
	5	5_24X0	38266686	38287857	2	21171
	5	5_25X0	39508613	39521363	2	12750
	5	5_26X0	40513669	40544313	3	30644
	5	5_27X0	41043525	41206805	3	163280
	5	5_28X0	41395573	41485508	2	89935
	5	5_29X0	41551535	41574963	2	23428
	5	5_2X0	3210270	3267061	2	56791
	5	5_30X0	41795234	42075817	15	280583
	5	5_31X0	42130313	42433304	23	302991
	5	5_32X0	42458423	42472143	3	13720
	5	5_33X0	42483294	42773107	11	289813
	5	5_34X0	43125713	43420546	3	294833
	5	5_35X0	43505734	43576607	2	70873
	5	5_36X0	43755585	44022342	2	266757
	5	5_37X0	45323954	45399413	3	75459
	5	5_38X0	47114206	47320641	8	206435
	5	5_39X0	47749907	48081196	6	331289
	5	5_3X0	4588802	4636714	2	47912
	5	5_40X0	48967086	49097076	6	129990
	5	5_41X0	54758487	54885938	2	127451

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	5	5_42X0	59059079	59277037	4	217958
	5	5_43X0	62091286	62196852	2	105566
	5	5_44X0	66087335	66324338	2	237003
	5	5_45X0	79752432	80006372	3	253940
	5	5_46X0	115744250	115845068	2	100818
	5	5_47X0	123013395	123018470	2	5075
	5	5_48X0	124200087	124245320	6	45233
	5	5_49X0	124372133	125054391	14	682258
	5	5_4X0	6368344	6463710	2	95366
	5	5_50X0	125413421	125509518	2	96097
	5	5_51X0	131820853	131958940	4	138087
	5	5_52X0	133765167	133984124	3	218957
	5	5_53X0	135018365	135208801	3	190436
	5	5_54X0	140862450	140943092	2	80642
	5	5_55X0	145582065	146073516	3	491451
	5	5_56X0	159492637	159735535	2	242898
	5	5_57X0	160957720	161194673	2	236953
	5	5_58X0	161428450	161457669	2	29219
	5	5_59X0	162785897	163142323	6	356426
	5	5_5X0	7138169	7160029	2	21860
	5	5_60X0	164992511	165289116	7	296605
	5	5_6X0	8092248	8117401	2	25153
	5	5_7X0	8336247	8522950	6	186703
	5	5_8X0	12879179	12957623	2	78444
	5	5_9X0	13740421	14248587	5	508166
	6	6_10X0	29656834	30222698	15	565864
	6	6_11X0	30414193	30681531	4	267338
	6	6_12X0	30763651	31046002	7	282351
	6	6_13X0	32035010	32102426	2	67416
	6	6_14X0	32153080	32270435	5	117355
	6	6_15X0	32275064	32326767	4	51703
	6	6_16X0	32596543	32634089	2	37546
	6	6_17X0	32781010	33980075	20	1199065
	6	6_18X0	33981046	34107417	2	126371
	6	6_19X0	36421135	36441338	3	20203
	6	6_1X0	2738698	2869418	3	130720
	6	6_20X0	36475377	37183764	4	708387
	6	6_21X0	39878272	40008321	3	130049
	6	6_22X0	41927162	42129621	6	202459
	6	6_23X0	42954617	42976297	2	21680
	6	6_24X0	43838639	44077615	4	238976
	6	6_25X0	47074773	47125565	2	50792
	6	6_26X0	47858849	48025587	2	166738
	6	6_27X0	48744574	48877803	2	133229
	6	6_28X0	49200597	49276668	2	76071
	6	6_29X0	50776328	50941067	3	164739
	6	6_2X0	7414160	7417578	2	3418
	6	6_30X0	51032051	51128362	4	96311
	6	6_31X0	51814880	51945179	2	130299
	6	6_32X0	53169266	53225271	2	56005
	6	6_33X0	53809339	53974183	4	164844
	6	6_34X0	75180493	75378624	6	198131
	6	6_35X0	85332971	85351031	2	18060
	6	6_36X0	91466811	91610344	2	143533
	6	6_37X0	107321993	107404983	2	82990

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	6	6_38X0	110315578	110361637	2	46059
	6	6_39X0	113283134	113476918	2	193784
	6	6_3X0	26344064	27114101	40	770037
	6	6_40X0	114869560	114870743	2	1183
	6	6_41X0	118502825	118529953	2	27128
	6	6_42X0	118719424	118827409	3	107985
	6	6_43X0	134121360	134355592	2	234232
	6	6_44X0	135024530	135118642	5	94112
	6	6_45X0	139522158	139743676	3	221518
	6	6_46X0	146169709	146244050	2	74341
	6	6_47X0	146893249	146955179	2	61930
	6	6_48X0	152636557	152810360	6	173803
	6	6_49X0	162978171	163299301	3	321130
	6	6_4X0	27204111	27226426	2	22315
	6	6_50X0	163381893	163633836	2	251943
	6	6_51X0	170084222	170108944	2	24722
	6	6_5X0	27701259	27716568	4	15309
	6	6_6X0	27927178	28040290	3	113112
	6	6_7X0	28393014	28472977	13	79963
	6	6_8X0	28490374	28619755	2	129381
	6	6_9X0	28683066	29206797	14	523731
	7	7_10X0	65333642	65704661	5	371019
	7	7_11X0	72276742	72435959	4	159217
	7	7_12X0	74048332	74064635	2	16303
	7	7_13X0	75561397	75579370	2	17973
	7	7_14X0	77342690	77368923	3	26233
	7	7_15X0	77995443	78065131	3	69688
	7	7_16X0	78095873	78116696	2	20823
	7	7_17X0	78299399	78454303	3	154904
	7	7_18X0	78472493	78618750	4	146257
	7	7_19X0	78711675	78775925	3	64250
	7	7_1X0	985104	1023748	2	38644
	7	7_20X0	81089447	81106956	2	17509
	7	7_21X0	82719864	82854150	4	134286
	7	7_22X0	84232875	84252224	2	19349
	7	7_23X0	90601100	90817944	2	216844
	7	7_24X0	101363030	101474155	2	111125
	7	7_25X0	106542683	106677835	2	135152
	7	7_26X0	106805177	107013032	2	207855
	7	7_27X0	123155037	123278069	2	123032
	7	7_28X0	129675499	129773769	3	98270
	7	7_29X0	133795871	133941798	3	145927
	7	7_2X0	6433223	6897939	4	464716
	7	7_30X0	141377320	141582915	5	205595
	7	7_31X0	141620802	141845143	2	224341
	7	7_32X0	141878894	141918945	2	40051
	7	7_33X0	142168646	142201072	2	32426
	7	7_34X0	142412703	142461926	6	49223
	7	7_35X0	142559741	142621505	2	61764
	7	7_36X0	142713754	142764558	3	50804
	7	7_37X0	142899431	143226621	6	327190
	7	7_38X0	143385698	143479420	2	93722
	7	7_39X0	143536410	143752050	8	215640
	7	7_3X0	12620510	12622604	2	2094
	7	7_40X0	143827752	143919352	2	91600

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	7	7_41X0	148665935	149371238	9	705303
	7	7_42X0	150050249	150339391	6	289142
	7	7_43X0	150391086	150404706	2	13620
	7	7_44X0	150758963	150807767	2	48804
	7	7_45X0	151599747	151759403	2	159656
	7	7_4X0	16741333	16830866	2	89533
	7	7_5X0	27155167	27302028	10	146861
	7	7_6X0	30792587	30843357	2	50770
	7	7_7X0	33145536	33348118	2	202582
	7	7_8X0	46568612	46602179	2	33567
	7	7_9X0	63124051	63732738	5	608687
	7	7_9X1	63843607	64378374	4	534767
	8	8_10X0	29721196	29964054	4	242858
	8	8_11X0	33400396	33476627	2	76231
	8	8_12X0	33543219	33832065	5	288846
	8	8_13X0	59674805	59857230	2	182425
	8	8_14X0	62584630	62895085	2	310455
	8	8_15X0	73544069	73605193	2	61124
	8	8_16X0	76268699	76365350	2	96651
	8	8_17X0	78244285	78409904	2	165619
	8	8_18X0	81952181	82000298	2	48117
	8	8_19X0	137085773	137088313	2	2540
	8	8_1X0	6656656	7335185	15	678529
	8	8_20X0	144351156	144835224	11	484068
	8	8_21X0	144868538	145321033	4	452495
	8	8_22X0	145690601	145697746	2	7145
	8	8_23X0	145726050	145955999	2	229949
	8	8_24X0	145999798	146644694	9	644896
	8	8_2X0	9644718	9703050	2	58332
	8	8_3X0	11189874	11211277	3	21403
	8	8_4X0	11258075	11311337	2	53262
	8	8_5X0	17306689	17510495	2	203806
	8	8_6X0	22109547	22346844	4	237297
	8	8_7X0	22810363	22836010	2	25647
	8	8_8X0	23438950	23683655	3	244705
	8	8_9X0	29316295	29555506	5	239211
	9	9_10X0	44368970	44484107	2	115137
	9	9_11X0	44599850	44806249	4	206399
	9	9_12X0	49061344	49211981	2	150637
	9	9_13X0	63929602	64052466	3	122864
	9	9_14X0	74350561	74489641	4	139080
	9	9_15X0	76592780	76675496	2	82716
	9	9_16X0	78452095	78911947	3	459852
	9	9_17X0	86693362	86916769	7	223407
	9	9_18X0	86941240	86967136	2	25896
	9	9_19X0	91184462	91193375	2	8913
	9	9_1X0	13474356	13618854	4	144498
	9	9_20X0	91622066	91898871	2	276805
	9	9_21X0	92678047	92781002	2	102955
	9	9_22X0	95506488	95584899	2	78411
	9	9_23X0	95692617	95821468	2	128851
	9	9_24X0	95832619	95896836	2	64217
	9	9_25X0	96641172	96919880	4	278708
	9	9_26X0	100381572	100508787	2	127215
	9	9_27X0	105340029	105414338	2	74309

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	9	9_28X0	105808919	106121581	9	312662
	9	9_29X0	106243128	106252284	2	9156
	9	9_2X0	13678811	13826449	2	147638
	9	9_30X0	107879216	107953953	2	74737
	9	9_31X0	110243184	110291815	2	48631
	9	9_32X0	113306177	113327846	2	21669
	9	9_33X0	116645272	116783084	4	137812
	9	9_34X0	116853579	116877652	2	24073
	9	9_35X0	116926008	117078249	2	152241
	9	9_36X0	118365570	118402621	2	37051
	9	9_37X0	119020523	119161159	4	140636
	9	9_38X0	120181286	120434076	7	252790
	9	9_3X0	14621734	14738270	2	116536
	9	9_4X0	14753690	14774383	2	20693
	9	9_5X0	14801512	14827692	2	26180
	9	9_6X0	16014383	16081508	2	67125
	9	9_7X0	27662284	27701775	2	39491
	9	9_8X0	28192633	28588855	6	396222
	9	9_9X0	34840985	34852482	2	11497
	X	X_10X0	42168517	42177610	2	9093
	X	X_11X0	44176626	44330615	2	153989
	X	X_12X0	47030800	47115622	2	84822
	X	X_13X0	47974003	48059288	2	85285
	X	X_13X1	48475973	48622721	2	146748
	X	X_14X0	48776540	49008331	3	231791
	X	X_15X0	49289383	49316161	2	26778
	X	X_16X0	49760380	49788920	2	28540
	X	X_17X0	49900762	50268392	3	367630
	X	X_18X0	52099262	52190335	3	91073
	X	X_19X0	52379022	52661160	3	282138
	X	X_1X0	1230130	1464877	2	234747
	X	X_20X0	52826035	53438116	7	612081
	X	X_21X0	53991913	54019778	4	27865
	X	X_22X0	55462766	55554485	2	91719
	X	X_23X0	55698072	55874347	2	176275
	X	X_24X0	56081217	56087532	2	6315
	X	X_25X0	58198130	58518938	2	320808
	X	X_26X0	69785495	69990849	3	205354
	X	X_27X0	71062090	71199546	2	137456
	X	X_28X0	71451888	71552504	3	100616
	X	X_29X0	72680561	72783190	2	102629
	X	X_2X0	2282488	2386644	2	104156
	X	X_30X0	80559196	80571816	2	12620
	X	X_31X0	100792633	100828376	2	35743
	X	X_32X0	102116186	102332620	2	216434
	X	X_33X0	102466765	102593045	2	126280
	X	X_34X0	102912745	103508110	9	595365
	X	X_35X0	107877886	108108314	2	230428
	X	X_36X0	116338777	116532630	2	193853
	X	X_37X0	119251860	119318711	2	66851
	X	X_38X0	119943203	119992452	2	49249
	X	X_39X0	120775738	120806453	2	30715
	X	X_3X0	2477619	2695348	4	217729
	X	X_40X0	134834066	134918351	2	84285
	X	X_41X0	135065713	135180348	2	114635

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	X	X_42X0	135367511	135445058	2	77547
	X	X_43X0	142065274	142374528	3	309254
	X	X_44X0	149770605	149883424	2	112819
	X	X_45X0	150799095	150967036	2	167941
	X	X_46X0	153632332	153717052	2	84720
	X	X_47X0	153991090	154537787	5	546697
	X	X_48X0	154781080	154934809	3	153729
	X	X_4X0	12625682	12682582	2	56900
	X	X_5X0	15058485	15142414	2	83929
	X	X_6X0	26320795	26417175	3	96380
	X	X_7X0	28001221	28239974	2	238753
	X	X_8X0	30499692	30530831	4	31139
	X	X_9X0	39115529	39164701	2	49172
	1	1_10X0	16022052	16057285	2	35233
	1	1_11X0	17079031	17401699	5	322668
	1	1_12X0	19267130	19311907	3	44777
	1	1_13X0	19920009	20175194	5	255185
	1	1_14X0	22637580	22661536	3	23956
	1	1_15X0	24119771	24187248	2	67477
	1	1_16X0	25272393	25420806	2	148413
	1	1_17X0	27355391	27626240	2	270849
	1	1_18X0	31617700	31764047	2	146347
	1	1_19X0	32209094	32364312	3	155218
	1	1_1X0	69055	686673	3	617618
	1	1_20X0	32465069	32600990	2	135921
	1	1_21X0	34755047	34795090	4	40043
	1	1_22X0	34986980	35422058	3	435078
	1	1_23X0	35808172	35930528	2	122356
	1	1_24X0	37996770	38046794	2	50024
	1	1_25X0	39572393	39763914	2	191521
	1	1_26X0	40161373	40548169	4	386796
	1	1_27X0	42153421	42164718	2	11297
	1	1_28X0	43172330	43363203	2	190873
	1	1_29X0	44118850	44141631	2	22781
	1	1_2X0	1418422	1630589	5	212167
	1	1_30X0	46040142	46176488	2	136346
	1	1_31X0	46798998	47149735	5	350737
	1	1_32X0	47416072	47440691	2	24619
	1	1_33X0	52726467	52894991	2	168524
	1	1_34X0	84926685	85048470	2	121785
	1	1_35X0	86424086	86580754	3	156668
	1	1_36X0	89052319	89386461	6	334142
	1	1_37X0	89568241	89936001	3	367760
	1	1_38X0	99970018	100083377	3	113359
	1	1_39X0	103571100	103758690	5	187590
	1	1_3X0	1635738	2185190	5	549452
	1	1_40X0	109656390	109740304	5	83914
	1	1_41X0	110211500	110675033	4	463533
	1	1_42X0	111227659	111427777	3	200118
	1	1_43X0	111722064	111755797	2	33733
	1	1_44X0	112701127	112711355	2	10228
	1	1_45X0	116910057	117034250	2	124193
	1	1_46X0	119414931	119514800	2	99869
	1	1_47X0	145866560	145964582	2	98022
	1	1_48X0	147756200	147909257	2	153057

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_49X0	149782689	149887916	11	105227
	1	1_4X0	9850659	9942768	2	92109
	1	1_50X0	150732567	150807434	2	74867
	1	1_51X0	151024742	151046845	2	22103
	1	1_52X0	151287484	151459179	2	171695
	1	1_53X0	152510844	153032900	12	522056
	1	1_54X0	153297862	153348840	2	50978
	1	1_55X0	155018533	155049264	2	30731
	1	1_56X0	155063748	155134857	4	71109
	1	1_57X0	155659443	155857295	2	197852
	1	1_58X0	155900594	156070514	2	169920
	1	1_59X0	156213007	156243321	2	30314
	1	1_5X0	11654857	11674030	2	19173
	1	1_60X0	156815656	156881794	2	66138
	1	1_61X0	157092044	157138367	2	46323
	1	1_62X0	157531964	158100261	6	568297
	1	1_63X0	158179947	158357553	5	177606
	1	1_64X0	158398522	159536007	23	1137485
	1	1_64X1	159802424	160647295	15	844871
	1	1_64X2	160683519	161727664	15	1044145
	1	1_65X0	162497251	162780421	2	283170
	1	1_66X0	166918866	167090631	2	171765
	1	1_67X0	168540765	168582077	2	41312
	1	1_68X0	169603026	169733660	3	130634
	1	1_69X0	171090919	171342084	4	251165
	1	1_6X0	11845712	11858931	2	13219
	1	1_70X0	173714984	173886636	2	171652
	1	1_71X0	179843931	179918171	2	74240
	1	1_72X0	182450146	182671853	3	221707
	1	1_73X0	183023460	183244900	2	221440
	1	1_74X0	196652056	197067267	7	415211
	1	1_75X0	201359015	201421746	2	62731
	1	1_76X0	202122858	202319207	2	196349
	1	1_77X0	203178981	203229671	2	50690
	1	1_78X0	203340628	203491352	2	150724
	1	1_79X0	203795706	203854124	2	58418
	1	1_7X0	12644547	12728400	2	83853
	1	1_80X0	205142505	205527825	3	385320
	1	1_81X0	205579566	205680459	2	100893
	1	1_82X0	205770300	205994439	2	224139
	1	1_83X0	206470495	206734283	3	263788
	1	1_84X0	206768695	206869217	3	100522
	1	1_85X0	206913801	206970625	3	56824
	1	1_86X0	207089283	207793646	6	704363
	1	1_87X0	212608628	212699985	2	91357
	1	1_88X0	220754972	220808976	2	54004
	1	1_89X0	223541609	223776016	2	234407
	1	1_8X0	12774841	13421328	26	646487
	1	1_90X0	225886282	225941383	2	55101
	1	1_91X0	227918656	228061260	2	142604
	1	1_92X0	228395534	228493885	3	98351
	1	1_93X0	235362954	235448932	2	85978
	1	1_94X0	246985323	247329245	5	343922
	1	1_95X0	247450953	248682304	43	1231351
	1	1_96X0	248838280	248858860	2	20580

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_9X0	15438439	15482357	3	43918
	10	10_10X0	68901289	68985067	2	83778
	10	10_11X0	69220303	69401879	2	181576
	10	10_12X0	75037836	75109099	2	71263
	10	10_13X0	79557187	79982614	5	425427
	10	10_14X0	80336108	80367281	2	31173
	10	10_15X0	84220495	84241461	2	20966
	10	10_16X0	87225500	87370695	2	145195
	10	10_17X0	88664458	89252039	5	587581
	10	10_18X0	89302046	89421001	5	118955
	10	10_19X0	93061264	93077890	2	16626
	10	10_1X0	1018907	1048397	2	29490
	10	10_20X0	94683775	95067430	4	383655
	10	10_21X0	97319267	97334709	2	15442
	10	10_22X0	97584374	97676434	2	92060
	10	10_23X0	103446789	103479240	3	32451
	10	10_24X0	104254710	104299417	2	44707
	10	10_25X0	116427867	116645143	4	217276
	10	10_26X0	122849096	122913111	2	64015
	10	10_27X0	123135962	123150672	2	14710
	10	10_28X0	133308941	133352548	3	43607
	10	10_29X0	133420666	133569835	2	149169
	10	10_2X0	4826296	5218947	5	392651
	10	10_3X0	5498697	5526771	2	28074
	10	10_4X0	5638868	5889906	2	251038
	10	10_5X0	7571405	7749339	2	177934
	10	10_6X0	13586968	13655929	2	68961
	10	10_7X0	37828765	38123348	4	294583
	10	10_8X0	43437126	43648704	4	211578
	10	10_9X0	47300386	47326270	2	25884
	11	11_10X0	76657525	76798153	2	140628
	11	11_11X0	78016593	78176302	4	159709
	11	11_12X0	85628573	85656543	2	27970
	11	11_13X0	89710650	90042025	7	331375
	11	11_14X0	93741664	93812378	2	70714
	11	11_15X0	94997173	95051338	3	54165
	11	11_16X0	102317450	102377913	2	60463
	11	11_17X0	102520508	102955705	9	435197
	11	11_18X0	104885718	105023168	2	137450
	11	11_19X0	111786284	111879425	2	93141
	11	11_1X0	298200	327537	4	29337
	11	11_20X0	111908622	111914093	2	5471
	11	11_21X0	113931385	113990310	2	58925
	11	11_22X0	114521716	114706933	3	185217
	11	11_23X0	116789373	116837950	3	48577
	11	11_24X0	117820140	117877031	3	56891
	11	11_25X0	117907454	118119839	2	212385
	11	11_26X0	118136961	118264536	5	127575
	11	11_27X0	118304899	118342631	2	37732
	11	11_28X0	119308529	119728553	3	420024
	11	11_29X0	123753488	124571141	22	817653
	11	11_2X0	494513	805245	3	310732
	11	11_30X0	124748106	124936412	5	188306
	11	11_31X0	125672493	125840069	3	167576
	11	11_32X0	125946056	126021522	2	75466

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	11	11_33X0	128838015	128916731	2	78716
	11	11_34X0	130404925	130476641	2	71716
	11	11_35X0	134276741	134376321	2	99580
	11	11_3X0	832952	867116	2	34164
	11	11_4X0	1607565	1697755	2	90190
	11	11_5X0	1732410	1757452	2	25042
	11	11_6X0	1839673	1938492	2	98819
	11	11_7X0	2132716	2161209	3	28493
	11	11_8X0	5253190	6170408	32	917218
	11	11_8X1	4998983	5324297	13	325314
	11	11_8X10	60280586	60801305	9	520719
	11	11_8X11	63079940	63598715	7	518775
	11	11_8X12	66509079	66721370	8	212291
	11	11_8X13	61846591	62423195	8	576604
	11	11_8X14	72080859	72221949	6	141090
	11	11_8X15	48216810	48489780	7	272970
	11	11_8X16	62649694	63015839	4	366145
	11	11_8X17	58708758	59155034	6	446276
	11	11_8X18	61210754	61481842	5	271088
	11	11_8X19	64555956	64844569	4	288613
	11	11_8X2	56093277	56743811	26	650534
	11	11_8X20	71538025	71884561	4	346536
	11	11_8X21	56988872	57390636	4	401764
	11	11_8X22	67266481	67524517	4	258036
	11	11_8X23	7795905	7939594	4	143689
	11	11_8X24	69698232	69819024	3	120792
	11	11_8X25	49952391	49982535	2	30144
	11	11_8X26	68980021	69013382	2	33361
	11	11_8X27	3217944	3232386	2	14442
	11	11_8X28	73218457	73298599	2	80142
	11	11_8X29	18933813	19060681	2	126868
	11	11_8X3	55343151	56031393	20	688242
	11	11_8X30	63974607	63988346	2	13739
	11	11_8X31	44047981	44084214	2	36233
	11	11_8X32	14966669	15082342	2	115673
	11	11_8X33	54603069	54707902	2	104833
	11	11_8X34	33698261	33736422	2	38161
	11	11_8X35	49028857	49059112	2	30255
	11	11_8X4	4233288	4955713	26	722425
	11	11_8X5	59364412	59866575	10	502163
	11	11_8X6	58030930	58508105	13	477175
	11	11_8X7	59878809	60184666	8	305857
	11	11_8X8	18089438	18588681	12	499243
	11	11_8X9	6199146	7020235	12	821089
	11	11_9X0	75719374	75801535	2	82161
	12	12_10X0	10801532	11186937	18	385405
	12	12_11X0	11265924	11395566	4	129642
	12	12_12X0	12891974	12952147	2	60173
	12	12_13X0	13088743	13216773	2	128030
	12	12_14X0	14803572	14824415	2	20843
	12	12_15X0	20695597	21334900	6	639303
	12	12_16X0	48202083	48562956	3	360873
	12	12_17X0	48857145	49069837	3	212692
	12	12_18X0	49128454	49273330	3	144876
	12	12_19X0	49950754	49974007	3	23253

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	12	12_1X0	190077	262873	2	72796
	12	12_20X0	51915448	51994925	2	79477
	12	12_21X0	52168996	52952900	25	783904
	12	12_22X0	53180700	53416446	3	235746
	12	12_23X0	53425116	53493866	2	68750
	12	12_24X0	53938765	54056030	9	117265
	12	12_25X0	54630812	54648493	2	17681
	12	12_26X0	55129681	55637854	15	508173
	12	12_27X0	55716531	55721794	2	5263
	12	12_28X0	55757275	55829636	3	72361
	12	12_29X0	55932028	56160945	3	228917
	12	12_2X0	4368227	4445614	2	77387
	12	12_30X0	56300177	56316059	2	15882
	12	12_31X0	56763324	56958319	3	194995
	12	12_32X0	57434760	57458008	2	23248
	12	12_33X0	68201351	68253504	2	52153
	12	12_34X0	75334683	75422932	2	88249
	12	12_35X0	80707498	80719673	2	12175
	12	12_36X0	89519408	89526198	2	6790
	12	12_37X0	90971800	91179582	4	207782
	12	12_38X0	111753890	111817529	2	63639
	12	12_39X0	112938352	113011216	2	72864
	12	12_3X0	4809176	5046788	4	237612
	12	12_40X0	113398515	113438241	2	39726
	12	12_41X0	120346222	120469571	5	123349
	12	12_42X0	121132828	121234104	2	101276
	12	12_43X0	122207663	122227534	2	19871
	12	12_44X0	122701293	122730843	3	29550
	12	12_45X0	132921462	133205023	6	283561
	12	12_4X0	6452084	6820808	3	368724
	12	12_5X0	7060676	7109210	3	48534
	12	12_6X0	7812513	7936275	2	123762
	12	12_7X0	8455926	8539255	3	83329
	12	12_8X0	8845004	9208370	3	363366
	12	12_9X0	10030677	10454616	8	423939
	12	12_9X1	9594551	10018619	10	424068
	13	13_10X0	99962964	99986773	2	23809
	13	13_11X0	102799049	102876001	3	76952
	13	13_12X0	113105788	113172380	3	66592
	13	13_1X0	20138255	20193010	2	54755
	13	13_2X0	27792643	27971139	3	178496
	13	13_3X0	28003274	28322633	3	319359
	13	13_4X0	36168794	36297842	2	129048
	13	13_5X0	41127569	41194566	2	66997
	13	13_6X0	52065407	52159861	2	94454
	13	13_7X0	57140918	57170218	5	29300
	13	13_8X0	61409685	61427849	2	18164
	13	13_9X0	99254714	99307405	2	52691
	14	14_10X0	24398786	24438408	2	39622
	14	14_11X0	24505506	24634160	4	128654
	14	14_12X0	36517217	36582607	2	65390
	14	14_13X0	49693129	49783122	2	89993
	14	14_14X0	50331927	50532609	2	200682
	14	14_15X0	52267812	52328606	2	60794
	14	14_16X0	60508951	60724348	3	215397

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	14	14_17X0	61485535	61657964	2	172429
	14	14_18X0	64486839	64533690	2	46851
	14	14_19X0	64914493	65062598	2	148105
	14	14_1X0	19712904	20243820	17	530916
	14	14_1X1	20510453	21058455	16	548002
	14	14_20X0	67685153	67734450	2	49297
	14	14_21X0	70325081	70417045	3	91964
	14	14_22X0	70452157	70535015	2	82858
	14	14_23X0	73537394	73619888	4	82494
	14	14_24X0	74903954	75127080	2	223126
	14	14_25X0	93184993	93212127	2	27134
	14	14_26X0	94283315	94624052	8	340737
	14	14_27X0	95686417	95714196	2	27779
	14	14_28X0	96204839	96264754	2	59915
	14	14_29X0	103100144	103137439	2	37295
	14	14_2X0	21569803	21670920	4	101117
	14	14_3X0	22946228	22982258	2	36030
	14	14_4X0	23025902	23044276	2	18374
	14	14_5X0	23306967	23325334	2	18367
	14	14_6X0	23381990	23435718	2	53728
	14	14_7X0	23638816	24299460	10	660644
	14	14_8X0	24310151	24316946	2	6795
	14	14_9X0	24332310	24340045	2	7735
	15	15_10X0	32392782	32607310	3	214528
	15	15_11X0	34379071	34533260	2	154189
	15	15_12X0	39953975	40277487	4	323512
	15	15_13X0	40695129	40732158	2	37029
	15	15_14X0	41503967	41583586	2	79619
	15	15_15X0	41828107	42156636	6	328529
	15	15_16X0	43197235	43302255	3	105020
	15	15_17X0	43594042	43697217	2	103175
	15	15_18X0	45092655	45165573	2	72918
	15	15_19X0	48740506	48878050	2	137544
	15	15_1X0	21579912	21653276	3	73364
	15	15_20X0	50424387	50546699	2	122312
	15	15_21X0	51751667	51915725	2	164058
	15	15_22X0	57592033	57714745	2	122712
	15	15_23X0	64841883	64930920	2	89037
	15	15_24X0	65329579	65423072	2	93493
	15	15_25X0	73592008	73714509	2	122501
	15	15_26X0	74131240	74176872	2	45632
	15	15_27X0	74719542	74756202	2	36660
	15	15_28X0	74788544	74842676	2	54132
	15	15_29X0	74850512	75019061	2	168549
	15	15_2X0	22015233	22095857	3	80624
	15	15_30X0	75258599	75295807	2	37208
	15	15_31X0	78580890	78621179	2	40289
	15	15_32X0	79844976	79923754	2	78778
	15	15_33X0	82339998	82436762	2	96764
	15	15_34X0	82536788	82628152	2	91364
	15	15_35X0	84603901	84806430	2	202529
	15	15_36X0	89748661	89778692	2	30031
	15	15_37X0	90265705	90314499	2	48794
	15	15_38X0	101803509	101923095	3	119586
	15	15_3X0	22786632	22868384	2	81752

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	15	15_4X0	23128268	23447234	3	318966
	15	15_5X0	23643544	23687330	2	43786
	15	15_6X0	28378664	28848675	4	470011
	15	15_7X0	30083053	30145567	2	62514
	15	15_8X0	30403740	30614561	3	210821
	15	15_9X0	30626200	30685606	2	59406
	16	16_10X0	4340251	4416627	2	76376
	16	16_11X0	14711689	14765413	2	53724
	16	16_12X0	15395754	15515348	2	119594
	16	16_13X0	16350098	16393954	2	43856
	16	16_14X0	18317947	18372785	2	54838
	16	16_15X0	20409534	20697744	4	288210
	16	16_16X0	23182715	23380803	2	198088
	16	16_17X0	23678822	24220611	3	541789
	16	16_18X0	27341289	27450485	2	109196
	16	16_19X0	28342555	28772807	7	430252
	16	16_1X0	152687	181181	5	28494
	16	16_20X0	29443230	29454964	2	11734
	16	16_21X0	29663279	29697524	2	34245
	16	16_22X0	30053090	30070420	2	17330
	16	16_23X0	30395112	31150061	14	754949
	16	16_24X0	31259994	31426505	3	166511
	16	16_25X0	48082973	48235177	2	152204
	16	16_26X0	55803049	55944464	2	141415
	16	16_27X0	56565049	56683603	11	118554
	16	16_28X0	57358772	57416062	3	57290
	16	16_29X0	57542689	57688695	3	146006
	16	16_2X0	269269	352449	2	83180
	16	16_30X0	66552582	66696707	5	144125
	16	16_31X0	66934460	67009314	3	74854
	16	16_32X0	67161689	67226998	2	65309
	16	16_33X0	67975663	67993863	2	18200
	16	16_34X0	68644993	68835541	2	190548
	16	16_35X0	69299682	69339667	4	39985
	16	16_36X0	70299194	70372582	3	73388
	16	16_37X0	71448093	71884190	3	436097
	16	16_38X0	72054609	72077245	2	22636
	16	16_39X0	75204106	75223607	2	19501
	16	16_3X0	649417	690444	2	41027
	16	16_40X0	75478428	75535170	2	56742
	16	16_41X0	75660227	75726442	3	66215
	16	16_42X0	81053665	81096296	2	42631
	16	16_43X0	86510527	86582160	3	71633
	16	16_44X0	87302801	87334273	2	31472
	16	16_45X0	88453317	88686493	2	233176
	16	16_46X0	89917879	89936092	3	18213
	16	16_4X0	1221651	1258998	4	37347
	16	16_5X0	1967273	2209352	4	242079
	16	16_6X0	2493597	2527955	4	34358
	16	16_7X0	2712418	2858154	5	145736
	16	16_8X0	3012456	3018187	2	5731
	16	16_9X0	3088894	3443542	10	354648
	17	17_10X0	7074796	7179564	2	104768
	17	17_11X0	7203474	7234061	2	30587
	17	17_12X0	7420360	7427568	2	7208

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	17	17_13X0	7548891	7561608	3	12717
	17	17_14X0	8039040	8119047	3	80007
	17	17_15X0	8150816	8162975	2	12159
	17	17_16X0	8802737	8912517	2	109780
	17	17_17X0	10300866	10657309	6	356443
	17	17_18X0	15436021	15563561	2	127540
	17	17_19X0	15571491	15682879	2	111388
	17	17_1X0	1742836	1777574	2	34738
	17	17_20X0	16551387	16653853	2	102466
	17	17_21X0	19021863	19159000	2	137137
	17	17_22X0	19533880	19745612	4	211732
	17	17_23X0	27798806	27893365	2	94559
	17	17_24X0	28714281	28859113	2	144832
	17	17_25X0	31305356	31321884	2	16528
	17	17_26X0	34255285	34363231	5	107946
	17	17_27X0	35243089	35558098	6	315009
	17	17_28X0	35872002	36439566	14	567564
	17	17_29X0	37924415	38192541	6	268126
	17	17_2X0	3062669	3433841	11	371172
	17	17_30X0	39904597	39970565	2	65968
	17	17_31X0	40068903	40100725	2	31822
	17	17_32X0	40655620	40984308	11	328688
	17	17_33X0	40993430	41009114	3	15684
	17	17_34X0	41026026	41309253	30	283227
	17	17_35X0	41346092	41624541	15	278449
	17	17_36X0	41837780	41865431	2	27651
	17	17_37X0	42020123	42154916	3	134793
	17	17_38X0	42199168	42388414	3	189246
	17	17_39X0	42609686	42667006	2	57320
	17	17_3X0	3510503	3592350	2	81847
	17	17_40X0	42844601	42858128	2	13527
	17	17_41X0	42950550	42980528	3	29978
	17	17_42X0	43801443	43907942	2	106499
	17	17_43X0	43940815	43953940	2	13125
	17	17_44X0	44899712	44905390	2	5678
	17	17_45X0	45148502	45170035	2	21533
	17	17_46X0	46762506	46886730	2	124224
	17	17_47X0	47844914	47917856	2	72942
	17	17_48X0	48542655	48729178	9	186523
	17	17_49X0	49972939	49993703	2	20764
	17	17_4X0	3679611	3916500	4	236889
	17	17_50X0	50112671	50150630	2	37959
	17	17_51X0	50381238	50468966	2	87728
	17	17_52X0	50508384	50542496	2	34112
	17	17_53X0	51153636	51171744	2	18108
	17	17_54X0	58155133	58171411	2	16278
	17	17_55X0	58192737	58280935	3	88198
	17	17_56X0	61399896	61483997	2	84101
	17	17_57X0	63622710	63714057	2	91347
	17	17_58X0	63872015	63918838	3	46823
	17	17_59X0	64319415	64468642	2	149227
	17	17_5X0	4433688	4537798	2	104110
	17	17_60X0	66877273	67056797	3	179524
	17	17_61X0	68867292	69327182	5	459890
	17	17_62X0	70075225	70180048	2	104823

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	17	17_63X0	74523153	74705255	5	182102
	17	17_64X0	74924275	74949416	2	25141
	17	17_65X0	75874161	75896911	2	22750
	17	17_66X0	76566133	76643838	2	77705
	17	17_67X0	79795204	79839429	2	44225
	17	17_68X0	81703357	81721016	3	17659
	17	17_69X0	82716787	82751197	2	34410
	17	17_6X0	4671385	5028401	3	357016
	17	17_7X0	5078248	5191883	3	113635
	17	17_8X0	7013733	7017520	2	3787
	17	17_9X0	7036075	7043923	2	7848
	18	18_1X0	3252275	3278284	2	26009
	18	18_2X0	13825544	13915707	2	90163
	18	18_3X0	30990086	31549008	6	558922
	18	18_4X0	35242473	35264223	2	21750
	18	18_5X0	45614801	45752520	2	137719
	18	18_6X0	46962768	47035264	5	72496
	18	18_7X0	49482152	49492479	2	10327
	18	18_8X0	63477127	63989371	10	512244
	19	19_10X0	5830626	5903756	5	73130
	19	19_11X0	6426039	6459735	2	33696
	19	19_12X0	6530999	6670588	2	139589
	19	19_13X0	7018971	7058640	5	39669
	19	19_14X0	7382834	7472477	2	89643
	19	19_15X0	7677088	7682854	2	5766
	19	19_16X0	7688758	7769604	3	80846
	19	19_17X0	9093179	9343843	10	250664
	19	19_17X1	9324341	9768617	9	444276
	19	19_18X0	10106395	10231331	3	124936
	19	19_19X0	10287004	10296778	2	9774
	19	19_1X0	544035	695461	2	151426
	19	19_20X0	11767092	12189881	15	422789
	19	19_20X1	12272329	12631921	11	359592
	19	19_20X2	11483427	11738945	3	255518
	19	19_21X0	12670248	12696581	2	26333
	19	19_22X0	14381444	14475362	2	93918
	19	19_23X0	14619117	14776842	2	157725
	19	19_24X0	14799146	15088221	6	289075
	19	19_25X0	15353385	15419026	2	65641
	19	19_26X0	15525333	15697145	4	171812
	19	19_27X0	15728020	15949958	7	221938
	19	19_28X0	16577651	16660130	2	82479
	19	19_29X0	17435509	17460804	2	25295
	19	19_2X0	827836	863238	4	35402
	19	19_30X0	18207961	18255419	2	47458
	19	19_31X0	18868545	18896144	2	27599
	19	19_32X0	19145584	19192138	2	46554
	19	19_33X0	19211973	19262796	2	50823
	19	19_34X0	19516227	19536076	2	19849
	19	19_35X0	19668901	19934626	5	265725
	19	19_35X1	20004418	20200490	3	196072
	19	19_36X0	20936606	21183158	4	246552
	19	19_36X1	21291160	21538660	3	247500
	19	19_36X2	20621367	20807322	2	185955
	19	19_37X0	21726529	21852068	2	125539

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	19	19_38X0	21970683	22196951	3	226268
	19	19_38X1	22286408	22422346	2	135938
	19	19_39X0	23116975	23395446	4	278471
	19	19_39X1	22634324	23003176	4	368852
	19	19_3X0	1269268	1432690	2	163422
	19	19_40X0	23652801	23758891	2	106090
	19	19_42X0	33299934	33382686	2	82752
	19	19_43X0	34677639	34773215	3	95576
	19	19_44X0	34926903	34960703	2	33800
	19	19_45X0	35116262	35169881	4	53619
	19	19_46X0	35248656	35347355	3	98699
	19	19_47X0	35351552	35451765	4	100213
	19	19_48X0	35826380	36032671	7	206291
	19	19_49X0	37412555	37779590	9	367035
	19	19_49X1	36445119	36639537	5	194418
	19	19_49X2	36711240	36952741	6	241501
	19	19_49X3	36334453	36418656	2	84203
	19	19_49X4	37078480	37390770	6	312290
	19	19_49X5	36111203	36238771	3	127568
	19	19_4X0	2841444	2944933	4	103489
	19	19_50X0	38770971	38808805	3	37834
	19	19_51X0	38915404	39032525	4	117121
	19	19_52X0	39243553	39298673	3	55120
	19	19_53X0	39530990	39540333	2	9343
	19	19_54X0	39602501	39738028	4	135527
	19	19_55X0	39997076	40090938	3	93862
	19	19_56X0	40191744	40390187	3	198443
	19	19_57X0	40421592	40444275	2	22683
	19	19_58X0	40843538	41128254	5	284716
	19	19_59X0	41576249	41880980	7	304731
	19	19_5X0	2977580	3045788	2	68208
	19	19_60X0	42076166	42228201	2	152035
	19	19_61X0	42230190	42255132	2	24942
	19	19_62X0	43015120	43269523	5	254403
	19	19_62X1	42754233	42937166	3	182933
	19	19_62X2	42509207	42740384	3	231177
	19	19_63X0	43353659	43719272	8	365613
	19	19_64X0	43833505	44177683	13	344178
	19	19_64X1	44212538	44500523	7	287985
	19	19_65X0	44511010	44878938	6	367928
	19	19_66X0	44942256	44949563	3	7307
	19	19_67X0	45589764	45681988	2	92224
	19	19_68X0	46039748	46231243	4	191495
	19	19_69X0	47309874	47342015	2	32141
	19	19_6X0	3113463	3163769	2	50306
	19	19_70X0	47801243	47843323	2	42080
	19	19_71X0	47968046	48025142	2	57096
	19	19_72X0	48696039	48755390	2	59351
	19	19_73X0	49015980	49055779	8	39799
	19	19_74X0	50819153	51084245	14	265092
	19	19_75X0	51142370	51646889	15	504519
	19	19_76X0	51746420	51825808	3	79388
	19	19_77X0	53044738	53193374	5	148636
	19	19_77X1	52764195	53016122	9	251927
	19	19_77X2	51872550	52160648	10	288098

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	19	19_77X3	52274462	52511142	7	236680
	19	19_77X4	53235385	53580269	9	344884
	19	19_77X5	52527653	52735044	4	207391
	19	19_78X0	53912817	54012664	3	99847
	19	19_79X0	54048366	54339150	9	290784
	19	19_7X0	3958454	4124117	2	165663
	19	19_80X0	54769825	55032865	9	263040
	19	19_80X1	54502907	54753052	8	250145
	19	19_81X0	55132832	55156725	2	23893
	19	19_82X0	55481197	55645621	8	164424
	19	19_83X0	55711433	56061810	8	350377
	19	19_84X0	56087366	56209200	4	121834
	19	19_85X0	57191775	57592390	19	400615
	19	19_85X1	57599885	57933905	16	334020
	19	19_85X2	58278951	58572775	11	293824
	19	19_85X3	57940833	58264092	7	323259
	19	19_85X4	56382752	56671760	7	289008
	19	19_85X5	56774562	57145202	3	370640
	19	19_8X0	4502180	4535224	2	33044
	19	19_9X0	5586999	5668326	2	81327
	2	2_10X0	68945297	68980974	2	35677
	2	2_11X0	70808813	70835822	2	27009
	2	2_12X0	70939318	70995336	2	56018
	2	2_13X0	79025727	79159753	4	134026
	2	2_14X0	85657318	85698747	2	41429
	2	2_15X0	86505444	86623028	2	117584
	2	2_16X0	86785920	86861924	2	76004
	2	2_17X0	89266494	89330421	2	63927
	2	2_18X0	89852186	89960754	4	108568
	2	2_19X0	95148303	95215603	3	67300
	2	2_1X0	18554723	18589564	2	34841
	2	2_20X0	95476967	95599778	2	122811
	2	2_21X0	96760902	96833911	2	73009
	2	2_22X0	96837915	96858095	2	20180
	2	2_23X0	99243456	99304742	2	61286
	2	2_24X0	102008515	102452384	6	443869
	2	2_25X0	108247195	108387640	3	140445
	2	2_26X0	112215865	112340063	2	124198
	2	2_27X0	112773915	113133121	9	359206
	2	2_28X0	118942166	119158889	2	216723
	2	2_29X0	130539095	130582406	2	43311
	2	2_2X0	27377235	27623215	2	245980
	2	2_30X0	151247940	151290057	2	42117
	2	2_31X0	159771851	160062610	3	290759
	2	2_32X0	161992241	162242998	2	250757
	2	2_33X0	169479480	169773385	4	293905
	2	2_34X0	176077472	176190907	9	113435
	2	2_35X0	177548998	177618966	2	69968
	2	2_36X0	188974386	189179879	2	205493
	2	2_37X0	190975642	191151260	2	175618
	2	2_38X0	197500413	197551279	2	50866
	2	2_39X0	201133036	201278161	3	145125
	2	2_3X0	37105143	37317079	2	211936
	2	2_40X0	203706475	203961577	3	255102
	2	2_41X0	218129287	218263848	3	134561

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	2	2_42X0	219245896	219493632	4	247736
	2	2_43X0	230225730	230403730	2	178000
	2	2_44X0	232378534	232460032	3	81498
	2	2_45X0	232526160	232545716	2	19556
	2	2_46X0	233617645	233773310	9	155665
	2	2_47X0	240028965	240046072	2	17107
	2	2_48X0	240676422	240698479	2	22057
	2	2_4X0	39178547	39437301	2	258754
	2	2_5X0	43812842	43878466	2	65624
	2	2_6X0	44941898	45009430	2	67532
	2	2_7X0	48530186	48679605	3	149419
	2	2_8X0	53670293	53860160	2	189867
	2	2_9X0	68643589	68658247	2	14658
	20	20_10X0	23439685	23879750	8	440065
	20	20_11X0	31257664	31473013	7	215349
	20	20_12X0	31819373	32101519	2	282146
	20	20_13X0	32632183	32743523	2	111340
	20	20_14X0	33007600	33309878	5	302278
	20	20_15X0	36574553	36612384	2	37831
	20	20_16X0	38304150	38377023	2	72873
	20	20_17X0	45109452	45790072	15	680620
	20	20_18X0	45857617	45885265	2	27648
	20	20_19X0	49936407	49988886	2	52479
	20	20_1X0	87710	261096	6	173386
	20	20_20X0	50081124	50153637	3	72513
	20	20_21X0	56524486	56536520	2	12034
	20	20_22X0	63528001	63547504	2	19503
	20	20_23X0	64085016	64107171	2	22155
	20	20_2X0	1476140	1939894	6	463754
	20	20_3X0	2296001	2432753	2	136752
	20	20_4X0	2814968	2818066	2	3098
	20	20_5X0	3071620	3084724	2	13104
	20	20_6X0	4686543	4728460	2	41917
	20	20_7X0	18567453	18744216	2	176763
	20	20_8X0	21395367	21514026	2	118659
	20	20_9X0	23045633	23086340	2	40707
	21	21_10X0	41361943	41459213	2	97270
	21	21_11X0	41798227	42010387	2	212160
	21	21_12X0	42312074	42366594	3	54520
	21	21_13X0	43659548	43805293	2	145745
	21	21_14X0	44329891	44462196	2	132305
	21	21_15X0	44538981	44697807	13	158826
	21	21_16X0	45981769	46129805	2	148036
	21	21_1X0	5116343	5133805	2	17462
	21	21_2X0	26835747	26966513	2	130766
	21	21_3X0	30165564	30216073	2	50509
	21	21_4X0	30281309	31038476	25	757167
	21	21_5X0	32578823	32612555	2	33732
	21	21_6X0	33025935	33072420	2	46485
	21	21_7X0	33246774	33437502	4	190728
	21	21_8X0	34364024	34511325	2	147301
	21	21_9X0	36070035	36146566	2	76531
	22	22_10X0	37010859	37025081	2	14222
	22	22_11X0	37570246	37679806	4	109560
	22	22_12X0	38290692	38398522	2	107830

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	22	22_13X0	38957522	39102558	8	145036
	22	22_14X0	42126501	42144577	2	18076
	22	22_15X0	43039776	43187133	2	147357
	22	22_16X0	43879679	43947061	2	67382
	22	22_17X0	44026304	44198717	2	172413
	22	22_18X0	44725261	44862784	2	137523
	22	22_19X0	50252903	50270393	2	17490
	22	22_1X0	21383374	21551461	2	168087
	22	22_20X0	50523568	50526145	2	2577
	22	22_2X0	22484873	22520270	2	35397
	22	22_3X0	23573369	23754468	2	181099
	22	22_4X0	23857000	24005409	8	148409
	22	22_5X0	30285146	30305631	3	20485
	22	22_6X0	30459230	30546682	3	87452
	22	22_7X0	32059170	32361161	4	301991
	22	22_8X0	35648395	35729483	2	81088
	22	22_9X0	36141027	36267530	4	126503
	3	3_10X0	45886509	46903755	13	1017246
	3	3_11X0	49416794	49429310	3	12516
	3	3_12X0	49859255	49903609	2	44354
	3	3_13X0	50159019	50277139	4	118120
	3	3_14X0	50295124	50322777	3	27653
	3	3_15X0	51778563	51903335	5	124772
	3	3_16X0	51933430	52278643	9	345213
	3	3_17X0	52779921	52826917	3	46996
	3	3_18X0	52833150	52897596	2	64446
	3	3_19X0	63864560	64003462	2	138902
	3	3_1X0	9902848	9933620	2	30772
	3	3_20X0	98087173	98533150	11	445977
	3	3_21X0	109293788	109337572	2	43784
	3	3_22X0	112819768	112974912	2	155144
	3	3_23X0	122531773	122730840	3	199067
	3	3_24X0	129087575	129161037	2	73462
	3	3_25X0	139005468	139044892	3	39424
	3	3_26X0	139454772	139539829	2	85057
	3	3_27X0	146192339	146606216	4	413877
	3	3_28X0	147386048	147416719	2	30671
	3	3_29X0	149322368	149487687	3	165319
	3	3_2X0	10816200	11036917	2	220717
	3	3_30X0	151198427	151384812	5	186385
	3	3_31X0	151733927	151824860	2	90933
	3	3_32X0	158666414	158732442	2	66028
	3	3_33X0	159764002	159897349	2	133347
	3	3_34X0	160258180	160448937	2	190757
	3	3_35X0	169793473	169869833	3	76360
	3	3_36X0	183487531	183673008	2	185477
	3	3_37X0	184035104	184106693	3	71589
	3	3_38X0	184249695	184293031	2	43336
	3	3_39X0	186570678	186597103	2	26425
	3	3_3X0	32238679	32503408	3	264729
	3	3_40X0	186613066	186743948	4	130882
	3	3_41X0	194340772	194398354	3	57582
	3	3_4X0	38265812	38318575	2	52763
	3	3_5X0	39263495	39333446	2	69951
	3	3_6X0	40477113	40539794	3	62681

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	3	3_7X0	42658814	42916995	2	258181
	3	3_8X0	42936859	43032292	2	95433
	3	3_9X0	44439770	44737083	8	297313
	4	4_10X0	67820876	68497604	5	676728
	4	4_11X0	68537184	69653247	10	1116063
	4	4_12X0	69995931	70058841	3	62910
	4	4_13X0	70360761	70410194	3	49433
	4	4_14X0	73436247	73504001	2	67754
	4	4_15X0	73740541	74099293	9	358752
	4	4_16X0	74309097	74770156	4	461059
	4	4_17X0	76001275	76041415	3	40140
	4	4_18X0	82359516	82430225	2	70709
	4	4_19X0	86886472	86934732	2	48260
	4	4_1X0	53295	499156	5	445861
	4	4_20X0	87303789	87391149	2	87360
	4	4_21X0	99076292	99435510	7	359218
	4	4_22X0	102911434	103077282	2	165848
	4	4_23X0	109713916	109802179	2	88263
	4	4_24X0	143875313	144127908	3	252595
	4	4_25X0	154563026	154612635	2	49609
	4	4_26X0	155667663	155806824	2	139161
	4	4_27X0	164954446	165060554	3	106108
	4	4_28X0	183318194	183322426	2	4232
	4	4_29X0	186227518	186288780	2	61262
	4	4_2X0	5562439	5758311	2	195872
	4	4_30X0	188091337	188147622	2	56285
	4	4_3X0	6640091	6709880	3	69789
	4	4_4X0	9210657	9364721	20	154064
	4	4_5X0	15715328	15840490	2	125162
	4	4_6X0	15935569	15963236	2	27667
	4	4_7X0	25160663	25279092	2	118429
	4	4_8X0	38772239	38829473	3	57234
	4	4_9X0	48831227	48906828	2	75601
	5	5_10X0	79069767	79131484	2	61717
	5	5_11X0	132294443	132394272	2	99829
	5	5_12X0	132449652	132646079	2	196427
	5	5_13X0	134160996	134226071	2	65075
	5	5_14X0	140676654	140698938	2	22284
	5	5_15X0	140786291	141959062	55	1172771
	5	5_16X0	146203550	146339251	2	135701
	5	5_17X0	147827086	148339849	6	512763
	5	5_18X0	149495896	150289629	4	793733
	5	5_19X0	151277351	151480143	3	202792
	5	5_1X0	1201595	1445430	3	243835
	5	5_20X0	156919389	157109122	3	189733
	5	5_21X0	175440039	175684313	2	244274
	5	5_22X0	178711592	179083537	6	371945
	5	5_23X0	180739042	181368262	10	629220
	5	5_2X0	14581775	14699711	2	117936
	5	5_3X0	17486231	17635053	14	148822
	5	5_4X0	35962874	36066882	2	104008
	5	5_5X0	55024253	55110252	2	85999
	5	5_6X0	55859509	55976314	2	116805
	5	5_7X0	69355492	69370013	2	14521
	5	5_8X0	69419386	69558104	2	138718

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	5	5_9X0	76617373	76835315	3	217942
	6	6_10X0	28080975	28587335	14	506360
	6	6_11X0	29044213	29588968	15	544755
	6	6_12X0	29657210	30346854	12	689644
	6	6_13X0	31269339	31511123	4	241784
	6	6_14X0	31572283	31582425	3	10142
	6	6_15X0	31669118	31721845	8	52727
	6	6_16X0	31809619	31830255	3	20636
	6	6_17X0	31927724	31952053	3	24329
	6	6_18X0	31982068	32035417	2	53349
	6	6_19X0	32394736	33087201	19	692465
	6	6_1X0	1312440	1613897	3	301457
	6	6_20X0	33314406	33457548	2	143142
	6	6_21X0	34417467	34426069	2	8602
	6	6_22X0	35776619	35796850	2	20231
	6	6_23X0	35833034	36547470	4	714436
	6	6_24X0	39188973	39322588	3	133615
	6	6_25X0	41158507	41350864	5	192357
	6	6_26X0	42155406	42194916	2	39510
	6	6_27X0	43037620	43272626	4	235006
	6	6_28X0	46129993	46170971	2	40978
	6	6_29X0	46852512	47018499	2	165987
	6	6_2X0	2832332	2903280	2	70948
	6	6_30X0	47673371	47722021	2	48650
	6	6_31X0	49692358	49866499	3	174141
	6	6_32X0	49946101	50048651	5	102550
	6	6_33X0	50713828	50847613	2	133785
	6	6_34X0	52186387	52244537	2	58150
	6	6_35X0	52750089	52994261	5	244172
	6	6_36X0	73223544	73395133	5	171589
	6	6_37X0	89177508	89315299	2	137791
	6	6_38X0	116249961	116279903	2	29942
	6	6_39X0	116462467	116558575	3	96108
	6	6_3X0	3153669	3227735	2	74066
	6	6_40X0	132538290	132646003	6	107713
	6	6_41X0	149918707	150069066	5	150359
	6	6_42X0	160121936	160452581	3	330645
	6	6_43X0	160531483	160753306	2	221823
	6	6_44X0	166858094	166956589	2	98495
	6	6_45X0	167111807	167158329	2	46522
	6	6_4X0	10722915	10756971	2	34056
	6	6_5X0	25726132	26509935	47	783803
	6	6_6X0	27132370	27147515	4	15145
	6	6_7X0	27388992	27473118	2	84126
	6	6_8X0	27807444	27893843	15	86399
	6	6_9X0	27911185	27958182	2	46997
	7	7_10X0	55887277	55955239	2	67962
	7	7_11X0	63998874	65401135	10	1402261
	7	7_12X0	66075884	66153289	3	77405
	7	7_13X0	73768997	73832693	2	63696
	7	7_14X0	74505992	75146931	4	640939
	7	7_15X0	75769533	75823356	2	53823
	7	7_16X0	90154375	90237678	2	83303
	7	7_17X0	93099513	93148345	2	48832
	7	7_18X0	93906594	93928610	2	22016

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	7	7_19X0	95297676	95435072	3	137396
	7	7_1X0	1055285	1093809	2	38524
	7	7_20X0	97007425	97024950	2	17525
	7	7_21X0	99325879	99394814	3	68935
	7	7_22X0	99416739	99466164	2	49425
	7	7_23X0	99472841	100079167	13	606326
	7	7_24X0	100358015	100400096	2	42081
	7	7_25X0	102347206	102671601	9	324395
	7	7_26X0	107660828	107803207	2	142379
	7	7_27X0	107959506	108130356	2	170850
	7	7_28X0	116499780	116561184	2	61404
	7	7_29X0	122697712	122702967	2	5255
	7	7_2X0	6615617	6706923	3	91306
	7	7_30X0	123848070	123960039	2	111969
	7	7_31X0	130266827	130388114	4	121287
	7	7_32X0	134442350	134579549	3	137199
	7	7_33X0	139340359	139422599	2	82240
	7	7_34X0	141764097	141973773	5	209676
	7	7_35X0	141995879	142222324	2	226445
	7	7_36X0	142749468	142774564	2	25096
	7	7_37X0	142871707	142933723	2	62016
	7	7_38X0	143026158	143053347	2	27189
	7	7_39X0	143183419	143478796	4	295377
	7	7_3X0	16797631	16881987	2	84356
	7	7_40X0	143571801	143902155	5	330354
	7	7_41X0	143935166	144319057	11	383891
	7	7_42X0	149069963	149773479	8	703516
	7	7_43X0	150368790	150408274	4	39484
	7	7_44X0	150450630	150744063	7	293433
	7	7_45X0	150791292	150803808	2	12516
	7	7_46X0	151028422	151227166	3	198744
	7	7_47X0	151956447	152122160	2	165713
	7	7_4X0	17298622	17467234	2	168612
	7	7_5X0	19115468	19145421	2	29953
	7	7_6X0	27100354	27247825	10	147471
	7	7_7X0	30496623	30550761	2	54138
	7	7_8X0	30771417	30925516	3	154099
	7	7_9X0	45888497	45921140	2	32643
	8	8_10X0	24384470	24526970	2	142500
	8	8_11X0	39913284	40016391	2	103107
	8	8_12X0	42697419	42768610	2	71191
	8	8_13X0	55879835	56113982	2	234147
	8	8_14X0	81280867	81531378	4	250511
	8	8_15X0	85245527	85481493	3	235966
	8	8_16X0	94371960	94475109	2	103149
	8	8_17X0	123219957	123274284	2	54327
	8	8_18X0	132024220	132085655	2	61435
	8	8_19X0	142680457	143217170	15	536713
	8	8_1X0	6924693	7056734	6	132041
	8	8_20X0	143246928	143695380	6	448452
	8	8_21X0	143716378	143840630	2	124252
	8	8_22X0	144330565	144950888	6	620323
	8	8_2X0	7332387	7338707	2	6320
	8	8_3X0	7414860	7896711	16	481851
	8	8_4X0	7971661	7977985	2	6324

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	8	8_5X0	10723768	10839847	3	116079
	8	8_6X0	11973937	12318316	7	344379
	8	8_7X0	18216894	18400993	2	184099
	8	8_8X0	23022311	23225061	4	202750
	8	8_9X0	23680988	23706598	2	25610
	9	9_10X0	92389854	92536092	4	146238
	9	9_11X0	94558720	94640249	2	81529
	9	9_12X0	96385941	96853072	3	467131
	9	9_13X0	100442271	100577636	2	135365
	9	9_14X0	104504263	104695485	8	191222
	9	9_15X0	104747688	104774010	2	26322
	9	9_16X0	108854589	108863759	2	9170
	9	9_17X0	109239731	109498313	2	258582
	9	9_18X0	109780456	110169909	3	389453
	9	9_19X0	110243811	110337884	2	94073
	9	9_1X0	976964	1057550	2	80586
	9	9_20X0	111631386	111670246	2	38860
	9	9_21X0	112997120	113056688	2	59568
	9	9_22X0	113150942	113264492	2	113550
	9	9_23X0	113275642	113340272	2	64630
	9	9_24X0	114060992	114333252	3	272260
	9	9_25X0	122476958	123035696	17	558738
	9	9_26X0	124481236	124595808	2	114572
	9	9_27X0	126805006	126885875	2	80869
	9	9_28X0	127885321	127916996	2	31675
	9	9_29X0	129803268	129824134	2	20866
	9	9_2X0	5163863	5339873	3	176010
	9	9_30X0	133351805	133376166	2	24361
	9	9_31X0	135521455	135665913	5	144458
	9	9_32X0	136739525	136985761	6	246236
	9	9_33X0	137217452	137221323	2	3871
	9	9_3X0	5450503	5571254	2	120751
	9	9_4X0	21077105	21482313	16	405208
	9	9_5X0	34551434	34660819	2	109385
	9	9_6X0	34689857	34710069	2	20212
	9	9_7X0	35869263	35958154	2	88891
	9	9_8X0	69035750	69255208	2	219458
	9	9_9X0	81928514	81995256	3	66742
	X	X_10X0	41447460	41730135	3	282675
	X	X_11X0	47370583	48132613	7	762030
	X	X_12X0	48186220	48411910	5	225690
	X	X_13X0	49171926	49200199	2	28273
	X	X_14X0	49303669	49833973	14	530304
	X	X_15X0	51332830	51496596	2	163766
	X	X_16X0	51617024	51712131	3	95107
	X	X_17X0	52369030	52868068	7	499038
	X	X_18X0	52899383	52995472	2	96089
	X	X_19X0	54807599	54931429	2	123830
	X	X_1X0	1190897	1382689	3	191792
	X	X_20X0	55075071	55264730	4	189659
	X	X_21X0	57119683	57137523	2	17840
	X	X_22X0	70040542	70240627	3	200085
	X	X_23X0	71667804	71767199	3	99395
	X	X_24X0	72845042	72943675	2	98633
	X	X_25X0	73003517	73079512	2	75995

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	X	X_26X0	78271480	79172229	4	900749
	X	X_27X0	101390890	101412165	2	21275
	X	X_28X0	101550531	101659891	4	109360
	X	X_29X0	102125688	102142970	2	17282
	X	X_2X0	1403139	1643081	2	239942
	X	X_30X0	102215839	102471812	4	255973
	X	X_31X0	102601137	102752438	4	151301
	X	X_32X0	103062651	103832228	11	769577
	X	X_33X0	104012712	104040864	2	28152
	X	X_34X0	119399060	119471319	2	72259
	X	X_35X0	120070672	120165630	3	94958
	X	X_36X0	120872603	120985788	13	113185
	X	X_37X0	134769567	134854610	2	85043
	X	X_38X0	134990938	135428074	5	437136
	X	X_39X0	141002591	141698739	5	696148
	X	X_3X0	2691187	2816500	2	125313
	X	X_40X0	141880562	142205290	3	324728
	X	X_41X0	146809784	146814726	2	4942
	X	X_42X0	149548210	149931287	7	383077
	X	X_43X0	150568619	150735499	2	166880
	X	X_44X0	151912889	152138578	2	225689
	X	X_45X0	152698767	153183880	6	485113
	X	X_46X0	153344277	153426481	2	82204
	X	X_47X0	153781001	154271805	7	490804
	X	X_48X0	154477769	154653568	4	175799
	X	X_4X0	2903970	3112706	4	208736
	X	X_5X0	12867083	12922202	2	55119
	X	X_6X0	15243989	15315656	2	71667
	X	X_7X0	26138343	26218270	4	79927
	X	X_8X0	27746809	27981449	2	234640
	X	X_9X0	30215560	30252038	4	36478
	Y	Y_1X0	9337510	9530675	5	193165
	Y	Y_2X0	13985772	14056958	2	71186
	Y	Y_3X0	17877410	18028597	2	151187
	Y	Y_4X0	21511425	22417881	6	906456
	Y	Y_5X0	23130710	23291356	2	160646
	Y	Y_6X0	24763069	24905683	2	142614
	1	1_10X0	21417868	21439751	2	21883
	1	1_11X0	30671807	30809523	2	137716
	1	1_12X0	33899882	33941890	4	42008
	1	1_13X0	34137390	34278728	2	141338
	1	1_14X0	34996853	35240875	3	244022
	1	1_15X0	39758944	39854362	3	95418
	1	1_16X0	41498189	41509857	2	11668
	1	1_17X0	42673041	42851810	2	178769
	1	1_18X0	43677899	43699181	2	21282
	1	1_19X0	46386958	46609330	5	222372
	1	1_1X0	5721388	5745344	2	23956
	1	1_20X0	52391262	52517160	2	125898
	1	1_21X0	66974565	67164240	2	189675
	1	1_22X0	78667141	78707531	2	40390
	1	1_23X0	85164567	85286801	2	122234
	1	1_24X0	86609493	86831674	4	222181
	1	1_25X0	89295544	89585755	5	290211
	1	1_26X0	89757538	90116384	3	358846

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_27X0	104406691	104545286	2	138595
	1	1_28X0	109112281	109430375	4	318094
	1	1_29X0	110560523	110633021	2	72498
	1	1_2X0	8172105	8229020	2	56915
	1	1_30X0	110683374	110765376	4	82002
	1	1_31X0	111536587	111698738	3	162151
	1	1_32X0	112237421	112392245	4	154824
	1	1_33X0	117584924	117846493	3	261569
	1	1_34X0	117987422	118253289	3	265867
	1	1_35X0	120566187	120789947	4	223760
	1	1_36X0	122615510	122767466	2	151956
	1	1_37X0	123383712	123478850	2	95138
	1	1_38X0	124142938	124243290	5	100352
	1	1_39X0	125192083	125270204	2	78121
	1	1_3X0	10908209	10936004	3	27795
	1	1_40X0	125788737	125930397	2	141660
	1	1_41X0	126375222	126441613	2	66391
	1	1_42X0	127057124	127506935	7	449811
	1	1_43X0	129706852	129743132	2	36280
	1	1_44X0	129774499	129831298	2	56799
	1	1_45X0	130478140	130672368	2	194228
	1	1_46X0	132162950	132469737	4	306787
	1	1_47X0	132828475	133030850	3	202375
	1	1_48X0	133087364	133500833	8	413469
	1	1_49X0	133560456	133780493	4	220037
	1	1_4X0	11111921	11126510	2	14589
	1	1_50X0	133877030	134018568	2	141538
	1	1_51X0	134023754	134272076	4	248322
	1	1_52X0	134529817	134833889	6	304072
	1	1_53X0	135224316	135612488	6	388172
	1	1_54X0	135626765	135698601	2	71836
	1	1_55X0	135795118	135947997	3	152879
	1	1_56X0	136330228	136442018	3	111790
	1	1_58X0	139399220	139509958	5	110738
	1	1_59X0	142502258	142649131	2	146873
	1	1_5X0	11951831	12035774	2	83943
	1	1_60X0	158788494	158931493	2	142999
	1	1_61X0	159121555	159502939	4	381384
	1	1_62X0	159635112	159697208	3	62096
	1	1_63X0	159737125	159843965	3	106840
	1	1_64X0	159883541	160140399	3	256858
	1	1_65X0	161047274	161435603	3	388329
	1	1_66X0	161507945	161642356	2	134411
	1	1_67X0	163121273	163289331	3	168058
	1	1_68X0	163408391	163460211	2	51820
	1	1_69X0	163709094	163883751	2	174657
	1	1_6X0	12101823	12173980	3	72157
	1	1_70X0	164425659	164617715	2	192056
	1	1_71X0	165353036	165408869	2	55833
	1	1_72X0	169798574	170196614	4	398040
	1	1_73X0	174094300	174335200	2	240900
	1	1_74X0	184080614	184270055	2	189441
	1	1_75X0	184700328	184850513	3	150185
	1	1_76X0	184963571	184981354	2	17783
	1	1_77X0	187504092	187579951	2	75859

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_78X0	193681272	193840438	2	159166
	1	1_79X0	196268245	196541900	5	273655
	1	1_7X0	15764261	16110899	3	346638
	1	1_80X0	197890148	198034618	3	144470
	1	1_81X0	204460302	204586161	2	125859
	1	1_82X0	223619189	223956009	6	336820
	1	1_83X0	224536243	224758735	7	222492
	1	1_83X1	224922015	225308826	7	386811
	1	1_83X2	224081578	224485815	5	404237
	1	1_84X0	225328849	225409362	2	80513
	1	1_85X0	225491690	225502813	2	11123
	1	1_8X0	18059924	18094249	2	34325
	1	1_9X0	18774231	18967019	3	192788
	10	10_10X0	35900674	35922951	2	22277
	10	10_11X0	37574938	37602947	2	28009
	10	10_12X0	56190839	56229789	2	38950
	10	10_13X0	56591369	56870687	7	279318
	10	10_14X0	58931960	58943610	2	11650
	10	10_15X0	62244512	62308495	2	63983
	10	10_16X0	63903192	64351138	20	447946
	10	10_16X1	64377833	64838781	14	460948
	10	10_16X2	63540975	63781594	8	240619
	10	10_17X0	65015402	65112721	2	97319
	10	10_18X0	65847954	65974462	7	126508
	10	10_19X0	66185997	66229225	2	43228
	10	10_1X0	176770	192989	2	16219
	10	10_20X0	72302255	72384150	3	81895
	10	10_21X0	73941860	74118507	2	176647
	10	10_22X0	77569269	77647273	2	78004
	10	10_23X0	78061977	78159250	3	97273
	10	10_24X0	78991676	79006376	2	14700
	10	10_25X0	79556239	79670576	2	114337
	10	10_26X0	80941100	81076521	6	135421
	10	10_27X0	84233025	84254021	2	20996
	10	10_28X0	84655900	84725269	2	69369
	10	10_29X0	85196574	85332659	2	136085
	10	10_2X0	7897622	7918860	2	21238
	10	10_30X0	86040452	86106072	2	65620
	10	10_31X0	86227526	86359327	2	131801
	10	10_3X0	18731938	18991541	7	259603
	10	10_4X0	19245372	19269310	2	23938
	10	10_5X0	25679078	25755599	2	76521
	10	10_6X0	31011020	31086183	3	75163
	10	10_7X0	32605302	32881400	7	276098
	10	10_8X0	33134644	33266007	2	131363
	10	10_9X0	34252759	34469542	5	216783
	11	11_10X0	11854882	12046483	4	191601
	11	11_11X0	13575092	13622808	2	47716
	11	11_12X0	13758713	13885645	2	126932
	11	11_13X0	15478562	15498729	2	20167
	11	11_14X0	21664920	22284236	5	619316
	11	11_15X0	44354831	44593441	2	238610
	11	11_16X0	46735386	46790919	2	55533
	11	11_17X0	47137727	47359948	3	222221
	11	11_18X0	47415827	47560796	3	144969

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	11	11_19X0	48028736	48179162	2	150426
	11	11_1X0	4982966	5219997	3	237031
	11	11_20X0	48235211	48258525	3	23314
	11	11_21X0	50215007	50302516	2	87509
	11	11_22X0	50825168	51598241	26	773073
	11	11_23X0	51815034	52050229	3	235195
	11	11_24X0	52064984	52135549	2	70565
	11	11_25X0	52624099	52703596	2	79497
	11	11_26X0	53324699	53342856	2	18157
	11	11_27X0	53826135	54359577	13	533442
	11	11_28X0	54663973	54912271	3	248298
	11	11_29X0	55690181	55740862	2	50681
	11	11_2X0	6551789	6692259	2	140470
	11	11_30X0	56226272	56250674	2	24402
	11	11_31X0	56737533	56739995	2	2462
	11	11_32X0	66942738	66998300	2	55562
	11	11_33X0	74198542	74384295	3	185753
	11	11_34X0	79595962	79608309	2	12347
	11	11_35X0	90638572	90793975	2	155403
	11	11_36X0	112949466	113044703	3	95237
	11	11_37X0	113422309	113476239	2	53930
	11	11_38X0	120556862	120698210	2	141348
	11	11_39X0	121419351	121523373	2	104022
	11	11_3X0	7395069	7471314	4	76245
	11	11_40X0	133134466	133146363	2	11897
	11	11_41X0	133271667	133496812	6	225145
	11	11_4X0	8143677	8274437	2	130760
	11	11_5X0	8691640	8789430	3	97790
	11	11_6X0	9138466	9589453	4	450987
	11	11_7X0	9955495	10415249	11	459754
	11	11_7X1	10421057	10899261	9	478204
	11	11_8X0	11243168	11384425	3	141257
	11	11_9X0	11589406	11755886	5	166480
	12	12_10X0	81915682	82278974	3	363292
	12	12_11X0	82634669	82705985	2	71316
	12	12_12X0	82914892	82983757	2	68865
	12	12_13X0	85399569	85656296	3	256727
	12	12_14X0	98707196	98764511	2	57315
	12	12_15X0	100155672	100356252	4	200580
	12	12_16X0	101380020	101402492	2	22472
	12	12_17X0	101567158	101695641	2	128483
	12	12_18X0	113345323	113562990	2	217667
	12	12_19X0	114916292	114969850	2	53558
	12	12_1X0	37805208	37840831	2	35623
	12	12_20X0	115037071	115056369	2	19298
	12	12_21X0	116262623	116281343	2	18720
	12	12_2X0	43967236	44627543	3	660307
	12	12_3X0	48722453	48975989	2	253536
	12	12_4X0	52908075	53422419	3	514344
	12	12_5X0	63122233	63180019	6	57786
	12	12_6X0	64596862	64663679	2	66817
	12	12_7X0	65608246	65791334	2	183088
	12	12_8X0	76381451	76493465	2	112014
	12	12_9X0	78594581	78771452	2	176871
	13	13_10X0	87415726	87453786	2	38060

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	13	13_11X0	89337846	89411830	2	73984
	13	13_12X0	90024279	90266094	2	241815
	13	13_13X0	90316572	91339881	36	1023309
	13	13_14X0	91539713	91577592	2	37879
	13	13_15X0	91793453	91819016	2	25563
	13	13_16X0	97426523	97852787	6	426264
	13	13_17X0	97873962	98100030	2	226068
	13	13_18X0	103701910	103963251	2	261341
	13	13_19X0	107272612	107371690	2	99078
	13	13_1X0	24373607	24419914	2	46307
	13	13_20X0	107765464	108130599	7	365135
	13	13_21X0	108616722	108965385	8	348663
	13	13_2X0	37575434	37796911	2	221477
	13	13_3X0	39721582	39922942	2	201360
	13	13_4X0	44534928	44591487	2	56559
	13	13_5X0	45687701	45756031	2	68330
	13	13_6X0	69105645	69133662	2	28017
	13	13_7X0	69980059	70015671	2	35612
	13	13_8X0	71915229	71941635	2	26406
	13	13_9X0	80394553	80503678	3	109125
	14	14_10X0	6856605	6928474	2	71869
	14	14_11X0	6947703	7123340	2	175637
	14	14_12X0	7699288	7873539	4	174251
	14	14_13X0	9615109	9664565	2	49456
	14	14_14X0	10751171	10820551	2	69380
	14	14_15X0	11028686	11219720	4	191034
	14	14_16X0	11285512	11325364	2	39852
	14	14_17X0	11458753	11710380	2	251627
	14	14_18X0	11893175	12061808	2	168633
	14	14_19X0	12345611	12488523	3	142912
	14	14_1X0	110278	144220	3	33942
	14	14_20X0	12815468	13071746	2	256278
	14	14_21X0	13158985	13312542	2	153557
	14	14_22X0	13465075	14146625	10	681550
	14	14_23X0	14605499	14662210	3	56711
	14	14_24X0	14741750	14818719	2	76969
	14	14_25X0	14879188	14914947	2	35759
	14	14_26X0	14928883	15090658	4	161775
	14	14_27X0	15261356	15620764	5	359408
	14	14_28X0	16259429	16274221	2	14792
	14	14_29X0	16783899	17753466	23	969567
	14	14_29X1	16411995	16756746	2	344751
	14	14_2X0	336292	650195	3	313903
	14	14_30X0	21966911	22003896	2	36985
	14	14_31X0	31246834	31428560	2	181726
	14	14_32X0	47726366	47949074	4	222708
	14	14_33X0	48115451	48334096	2	218645
	14	14_34X0	58633171	58798675	2	165504
	14	14_35X0	59602193	59665234	2	63041
	14	14_36X0	59703823	59837922	3	134099
	14	14_37X0	60513540	60830475	6	316935
	14	14_38X0	61378615	61895283	14	516668
	14	14_38X1	62508758	63161921	16	653163
	14	14_38X2	61961031	62463051	11	502020
	14	14_39X0	63769657	63789190	2	19533

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	14	14_3X0	682980	716720	2	33740
	14	14_40X0	64062640	64161768	4	99128
	14	14_41X0	65209328	65289724	2	80396
	14	14_42X0	66048497	66094250	2	45753
	14	14_43X0	67199373	67220017	2	20644
	14	14_44X0	67893897	67974454	2	80557
	14	14_45X0	68836519	68974231	2	137712
	14	14_46X0	82247149	82518375	8	271226
	14	14_47X0	87592998	87697790	4	104792
	14	14_48X0	95415721	95855252	9	439531
	14	14_49X0	97784838	97885629	3	100791
	14	14_4X0	934546	1103789	2	169243
	14	14_50X0	106779891	106863900	2	84009
	14	14_51X0	107399418	107529648	3	130230
	14	14_52X0	109647545	109663626	2	16081
	14	14_53X0	110757331	110963359	2	206028
	14	14_54X0	110999344	111088277	3	88933
	14	14_55X0	111140735	111177842	2	37107
	14	14_56X0	112176792	112274741	2	97949
	14	14_57X0	116717171	117306748	9	589577
	14	14_58X0	117487437	117629067	3	141630
	14	14_59X0	123109919	123181479	2	71560
	14	14_5X0	1675881	1750381	2	74500
	14	14_60X0	127294660	127395240	2	100580
	14	14_6X0	1954752	1984024	2	29272
	14	14_7X0	3147195	3208893	3	61698
	14	14_8X0	5601869	5632976	2	31107
	14	14_9X0	6620971	6719952	2	98981
	15	15_10X0	26072666	26129560	2	56894
	15	15_11X0	26142497	26246049	2	103552
	15	15_12X0	27820787	27900955	2	80168
	15	15_13X0	29208881	29304495	2	95614
	15	15_14X0	30709218	30718601	2	9383
	15	15_15X0	34783519	34815878	2	32359
	15	15_16X0	34837880	35020261	3	182381
	15	15_17X0	42578574	42720291	2	141717
	15	15_18X0	44411959	44786662	2	374703
	15	15_19X0	48156902	48265425	5	108523
	15	15_1X0	1306544	1565235	7	258691
	15	15_20X0	48450739	48653648	2	202909
	15	15_21X0	49435211	49541488	2	106277
	15	15_22X0	49550518	49577801	2	27283
	15	15_23X0	50503855	50573440	2	69585
	15	15_24X0	50762686	50903686	2	141000
	15	15_25X0	61917015	61955305	2	38290
	15	15_26X0	62417156	62863531	7	446375
	15	15_27X0	78886369	78989472	2	103103
	15	15_28X0	79122520	79213084	2	90564
	15	15_29X0	83488227	83704828	2	216601
	15	15_2X0	2692610	2825013	4	132403
	15	15_30X0	98301363	98353550	2	52187
	15	15_31X0	103804795	104116150	2	311355
	15	15_32X0	108366403	108447747	2	81344
	15	15_33X0	110561089	110635107	2	74018
	15	15_3X0	3445397	3471486	2	26089

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	15	15_4X0	4785860	4942041	2	156181
	15	15_5X0	4996385	5021737	2	25352
	15	15_6X0	10973816	11006589	2	32773
	15	15_7X0	12048737	12134726	2	85989
	15	15_8X0	14493560	14567628	2	74068
	15	15_9X0	16053612	16669649	7	616037
	16	16_10X0	8818641	8985690	2	167049
	16	16_11X0	10349193	10701495	6	352302
	16	16_12X0	19354954	19509709	2	154755
	16	16_13X0	22619968	22761218	2	141250
	16	16_14X0	24968066	25092207	4	124141
	16	16_15X0	28231966	28334332	5	102366
	16	16_16X0	29225120	29319417	2	94297
	16	16_16X1	29429254	29508216	2	78962
	16	16_17X0	29829331	30057421	7	228090
	16	16_18X0	30138021	30174707	2	36686
	16	16_19X0	31375631	31463242	3	87611
	16	16_1X0	1634499	1671348	2	36849
	16	16_20X0	39008489	39042382	2	33893
	16	16_21X0	39403637	39441589	2	37952
	16	16_22X0	39551548	39574183	2	22635
	16	16_23X0	41273004	41357911	3	84907
	16	16_24X0	41995071	42073489	2	78418
	16	16_25X0	45875502	45906580	2	31078
	16	16_26X0	46524894	46792662	10	267768
	16	16_27X0	46852802	47080449	16	227647
	16	16_28X0	47139340	47415353	15	276013
	16	16_29X0	47675293	47702397	2	27104
	16	16_2X0	3031685	3190129	7	158444
	16	16_30X0	47770519	47865253	2	94734
	16	16_31X0	48047928	48202574	3	154646
	16	16_32X0	48701782	48730811	3	29029
	16	16_33X0	49594672	49745188	2	150516
	16	16_34X0	49790718	49804289	2	13571
	16	16_35X0	51068854	51092359	2	23505
	16	16_36X0	51761726	51882363	2	120637
	16	16_37X0	60485460	60529405	2	43945
	16	16_38X0	62299473	62551636	3	252163
	16	16_38X1	62557782	62732651	2	174869
	16	16_39X0	63458900	63506092	2	47192
	16	16_3X0	3377605	3464020	2	86415
	16	16_40X0	68101151	68348449	6	247298
	16	16_41X0	69590636	69613179	2	22543
	16	16_42X0	70325363	70399526	2	74163
	16	16_43X0	73631207	73651080	2	19873
	16	16_44X0	76644815	76685342	2	40527
	16	16_4X0	4351408	4445345	2	93937
	16	16_5X0	5027050	5072227	3	45177
	16	16_6X0	6841450	6971369	2	129919
	16	16_7X0	7033496	7152558	3	119062
	16	16_8X0	7396781	7403674	2	6893
	16	16_9X0	8055212	8140969	3	85757
	17	17_1X0	734276	818188	3	83912
	17	17_2X0	20399455	20471439	2	71984
	17	17_3X0	31438598	31532624	2	94026

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	17	17_4X0	80961828	81016207	2	54379
	17	17_5X0	81653500	81671223	2	17723
	17	17_6X0	94584340	94644592	3	60252
	18	18_1X0	6212823	6241579	2	28756
	18	18_2X0	19546578	20081801	6	535223
	18	18_3X0	23942441	24073945	3	131504
	18	18_4X0	34568699	34687515	2	118816
	18	18_5X0	35895455	35902363	2	6908
	18	18_6X0	53050700	53634388	11	583688
	18	18_7X0	64660112	64743444	2	83332
	19	19_1X0	595162	630937	4	35775
	19	19_2X0	1814850	1910340	2	95490
	19	19_3X0	51644839	51745989	6	101150
	19	19_3X1	51747281	51856787	6	109506
	19	19_3X10	39322416	39402995	4	80579
	19	19_3X11	11965244	12194508	8	229264
	19	19_3X12	47905690	48187977	6	282287
	19	19_3X13	33197909	33434800	5	236891
	19	19_3X14	48400566	48568201	5	167635
	19	19_3X15	46698731	46792555	4	93824
	19	19_3X16	2727783	2869190	5	141407
	19	19_3X17	47099448	47176579	3	77131
	19	19_3X18	39409307	39514388	3	105081
	19	19_3X19	9270507	9410095	5	139588
	19	19_3X2	46187527	46384135	10	196608
	19	19_3X20	32954132	33168665	6	214533
	19	19_3X21	37412579	37584473	5	171894
	19	19_3X22	31221016	31372354	5	151338
	19	19_3X23	47489211	47639123	5	149912
	19	19_3X24	12480922	12638240	5	157318
	19	19_3X25	44420002	44460439	5	40437
	19	19_3X26	48677653	48818730	5	141077
	19	19_3X27	49255774	49408271	5	152497
	19	19_3X28	50986772	51144842	5	158070
	19	19_3X29	36626697	36730591	4	103894
	19	19_3X3	50007777	50155152	7	147375
	19	19_3X30	39129163	39287997	4	158834
	19	19_3X31	52416799	52611685	4	194886
	19	19_3X32	51194450	51300560	3	106110
	19	19_3X33	46413820	46561206	4	147386
	19	19_3X34	15631871	15741939	4	110068
	19	19_3X35	32253423	32439668	4	186245
	19	19_3X36	51881321	52015945	4	134624
	19	19_3X37	33451062	33554310	3	103248
	19	19_3X38	52649417	52774200	4	124783
	19	19_3X39	31743026	31830265	4	87239
	19	19_3X4	51366821	51497538	6	130717
	19	19_3X40	39529520	39725048	4	195528
	19	19_3X41	35780972	35870785	3	89813
	19	19_3X42	36817657	36952390	4	134733
	19	19_3X43	46597503	46682668	2	85165
	19	19_3X44	38533242	38682671	4	149429
	19	19_3X45	32801118	32910904	3	109786
	19	19_3X46	49561390	49670029	4	108639
	19	19_3X47	10294749	10368088	3	73339

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	19	19_3X48	2923802	3102333	4	178531
	19	19_3X49	46975396	47052497	3	77101
	19	19_3X5	49807031	49982745	7	175714
	19	19_3X50	21296371	21443278	3	146907
	19	19_3X51	47737172	47816245	3	79073
	19	19_3X52	31012815	31096180	3	83365
	19	19_3X53	9482612	9561008	3	78396
	19	19_3X54	12320409	12404968	3	84559
	19	19_3X55	52088255	52150025	2	61770
	19	19_3X56	32456064	32582448	3	126384
	19	19_3X57	7926620	7953275	3	26655
	19	19_3X58	50696004	50777367	2	81363
	19	19_3X59	44062839	44123927	2	61088
	19	19_3X6	51514503	51631574	4	117071
	19	19_3X60	20482827	20563458	2	80631
	19	19_3X61	29258114	29335724	2	77610
	19	19_3X62	5744366	5825402	2	81036
	19	19_3X63	20170091	20242138	2	72047
	19	19_3X64	42681348	42718947	2	37599
	19	19_3X65	50191689	50352893	3	161204
	19	19_3X66	38329220	38409472	2	80252
	19	19_3X67	9728240	9769238	2	40998
	19	19_3X68	15874298	15916141	2	41843
	19	19_3X69	20909896	21017074	2	107178
	19	19_3X7	35283788	35505012	7	221224
	19	19_3X70	34681722	34768738	2	87016
	19	19_3X71	30830017	30866790	2	36773
	19	19_3X72	30637191	30670840	2	33649
	19	19_3X73	45465141	45482718	2	17577
	19	19_3X74	19798781	19849355	2	50574
	19	19_3X75	14643370	14725276	2	81906
	19	19_3X76	34526070	34552659	2	26589
	19	19_3X77	32115069	32193378	2	78309
	19	19_3X78	43319676	43361143	2	41467
	19	19_3X79	5990167	6017957	2	27790
	19	19_3X8	50832982	50933203	5	100221
	19	19_3X80	37873921	37927101	2	53180
	19	19_3X81	36169571	36192637	2	23066
	19	19_3X82	35932736	35995503	2	62767
	19	19_3X83	40115383	40196623	2	81240
	19	19_3X84	11768583	11938217	3	169634
	19	19_3X85	32665735	32732728	2	66993
	19	19_3X86	6623090	6660690	2	37600
	19	19_3X87	21722029	21805776	2	83747
	19	19_3X88	34987245	35043416	2	56171
	19	19_3X9	47199458	47383444	7	183986
	19	19_4X0	52908802	53047287	5	138485
	19	19_5X0	53207696	53374001	3	166305
	19	19_5X1	53071963	53189634	2	117671
	19	19_6X0	53533912	53606820	6	72908
	2	2_10X0	90334811	90494805	5	159994
	2	2_11X0	91991409	92026709	3	35300
	2	2_12X0	92042286	92162830	2	120544
	2	2_13X0	92212181	92350349	2	138168
	2	2_14X0	93048408	93263106	2	214698

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	2	2_15X0	95728254	95835992	5	107738
	2	2_16X0	96139855	96658703	8	518848
	2	2_17X0	97848380	98152066	8	303686
	2	2_18X0	99722997	99975244	3	252247
	2	2_19X0	102153717	102212720	3	59003
	2	2_1X0	15410396	15416673	2	6277
	2	2_20X0	103770188	104178110	3	407922
	2	2_21X0	104245450	104599468	4	354018
	2	2_22X0	106210063	106322605	3	112542
	2	2_23X0	108943909	109022100	3	78191
	2	2_24X0	119320343	119323090	2	2747
	2	2_25X0	123974830	124049991	3	75161
	2	2_26X0	134271483	134447873	2	176390
	2	2_27X0	135999502	136170781	2	171279
	2	2_28X0	137986114	138165594	5	179480
	2	2_29X0	138580634	138682505	2	101871
	2	2_2X0	19099854	19339426	3	239572
	2	2_30X0	145817751	145929891	2	112140
	2	2_31X0	154342015	154373763	2	31748
	2	2_32X0	155170731	155541118	4	370387
	2	2_33X0	168109712	168186442	2	76730
	2	2_34X0	170567101	170637603	2	70502
	2	2_35X0	171339526	171795059	3	455533
	2	2_36X0	180767013	180966731	2	199718
	2	2_37X0	188603470	188866928	3	263458
	2	2_38X0	192269205	192500801	2	231596
	2	2_39X0	200034272	200324876	2	290604
	2	2_3X0	30401140	30435751	2	34611
	2	2_40X0	201129736	201152966	2	23230
	2	2_4X0	33918821	34029267	2	110446
	2	2_5X0	43754524	43952182	2	197658
	2	2_6X0	55272897	55493263	2	220366
	2	2_7X0	56440216	56468740	2	28524
	2	2_8X0	89348243	89394948	3	46705
	2	2_9X0	89721651	90304310	7	582659
	20	20_10X0	15836815	16014631	2	177816
	20	20_11X0	19489582	19528943	2	39361
	20	20_12X0	19586197	19850699	4	264502
	20	20_13X0	21859871	22086016	2	226145
	20	20_14X0	22382264	22418347	2	36083
	20	20_15X0	26133577	26241718	2	108141
	20	20_16X0	27557415	27602833	3	45418
	20	20_17X0	27997461	28417041	5	419580
	20	20_18X0	28708601	28717484	2	8883
	20	20_19X0	28935604	28954086	2	18482
	20	20_1X0	133491	140054	2	6563
	20	20_20X0	29006611	29026849	2	20238
	20	20_21X0	29103625	29228660	2	125035
	20	20_22X0	29599485	29735559	2	136074
	20	20_23X0	31856092	32006397	2	150305
	20	20_24X0	39787442	39996826	4	209384
	20	20_25X0	40701830	40845274	9	143444
	20	20_26X0	41570914	41623502	2	52588
	20	20_27X0	41771511	41902093	3	130582
	20	20_28X0	50832600	50960940	2	128340

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	20	20_29X0	51201330	51271908	2	70578
	20	20_2X0	250465	335687	2	85222
	20	20_30X0	52338273	52354943	2	16670
	20	20_31X0	53029844	53219285	2	189441
	20	20_32X0	55730650	55765790	2	35140
	20	20_33X0	56689428	56709201	2	19773
	20	20_34X0	59491501	59513675	2	22174
	20	20_35X0	59781529	59838972	2	57443
	20	20_36X0	73329914	73506974	2	177060
	20	20_3X0	573622	676754	4	103132
	20	20_4X0	751950	770654	2	18704
	20	20_5X0	1185640	1248571	5	62931
	20	20_6X0	1956805	2230149	3	273344
	20	20_7X0	3511451	3853413	7	341962
	20	20_8X0	3909696	3914811	2	5115
	20	20_9X0	4138177	4175282	2	37105
	3	3_10X0	16315080	16323601	3	8521
	3	3_11X0	16475981	16579751	4	103770
	3	3_12X0	16653628	16707227	2	53599
	3	3_13X0	19965199	20100745	2	135546
	3	3_14X0	34972290	35011126	2	38836
	3	3_15X0	39973880	40067033	2	93153
	3	3_16X0	41512194	41556625	2	44431
	3	3_17X0	42431267	42767956	3	336689
	3	3_18X0	43618309	43684649	2	66340
	3	3_19X0	50154362	50228737	2	74375
	3	3_1X0	641695	788789	2	147094
	3	3_20X0	51038132	52054194	7	1016062
	3	3_21X0	55152381	55197564	2	45183
	3	3_22X0	63012114	63046778	2	34664
	3	3_23X0	70456384	70498016	2	41632
	3	3_24X0	80969612	81040584	3	70972
	3	3_25X0	89064659	89094946	2	30287
	3	3_26X0	92732291	92814479	2	82188
	3	3_27X0	103740502	103827182	2	86680
	3	3_28X0	118405398	118454888	2	49490
	3	3_29X0	119274266	119293936	2	19670
	3	3_2X0	2036035	2193666	10	157631
	3	3_30X0	120653710	120786929	3	133219
	3	3_31X0	124912903	125068388	4	155485
	3	3_32X0	125092064	125377444	4	285380
	3	3_33X0	125527350	125595276	3	67926
	3	3_34X0	127040567	127065979	2	25412
	3	3_35X0	133330638	133547278	2	216640
	3	3_36X0	149851609	150040663	3	189054
	3	3_37X0	156403331	156498340	3	95009
	3	3_38X0	160580724	160824804	4	244080
	3	3_39X0	168162158	168612380	7	450222
	3	3_3X0	2269945	2389398	2	119453
	3	3_40X0	168648749	169351393	31	702644
	3	3_41X0	169423894	169517469	2	93575
	3	3_42X0	174934256	175473485	12	539229
	3	3_43X0	175622411	176055915	7	433504
	3	3_44X0	176533834	176576033	3	42199
	3	3_45X0	176618955	176963224	7	344269

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	3	3_46X0	177026174	177039996	2	13822
	3	3_47X0	178175099	178329857	2	154758
	3	3_4X0	2916524	3077330	2	160806
	3	3_5X0	4522386	4702123	2	179737
	3	3_6X0	5222289	5303169	2	80880
	3	3_7X0	13249961	13458156	3	208195
	3	3_8X0	13626899	13670768	2	43869
	3	3_9X0	16123962	16163368	3	39406
	4	4_10X0	28331150	28383361	2	52211
	4	4_11X0	28406707	29017996	14	611289
	4	4_12X0	29497128	30048864	10	551736
	4	4_13X0	30176724	30618888	5	442164
	4	4_14X0	30772889	31065690	7	292801
	4	4_15X0	31982538	32364224	5	381686
	4	4_16X0	32430770	32441942	3	11172
	4	4_17X0	32680549	32700878	2	20329
	4	4_18X0	33284317	33930487	14	646170
	4	4_19X0	34158412	34313174	2	154762
	4	4_1X0	2757695	2880681	3	122986
	4	4_20X0	36703464	36722913	3	19449
	4	4_21X0	36763736	37068127	3	304391
	4	4_22X0	40180925	40264572	3	83647
	4	4_23X0	42110733	42322096	5	211363
	4	4_24X0	43167687	43189902	2	22215
	4	4_25X0	44053247	44291719	4	238472
	4	4_26X0	47940313	48108481	2	168168
	4	4_27X0	48754835	48911411	3	156576
	4	4_28X0	50750835	50877447	2	126612
	4	4_29X0	50971154	51060530	4	89376
	4	4_2X0	3063263	3131387	2	68124
	4	4_30X0	51720991	51845261	2	124270
	4	4_31X0	53157019	53203695	2	46676
	4	4_32X0	53753292	53915049	5	161757
	4	4_33X0	71261227	71409832	3	148605
	4	4_34X0	71536395	71559367	2	22972
	4	4_35X0	87242754	87379803	2	137049
	4	4_36X0	105086923	105163826	2	76903
	4	4_37X0	116248266	116588544	8	340278
	4	4_38X0	134701575	134782968	2	81393
	4	4_39X0	134819417	134929788	8	110371
	4	4_3X0	10688854	10829054	2	140200
	4	4_40X0	151262036	151357342	3	95306
	4	4_41X0	156685249	156863781	3	178532
	4	4_42X0	161258066	161577888	3	319822
	4	4_4X0	10860294	10943065	2	82771
	4	4_5X0	26313543	26864906	33	551363
	4	4_6X0	26979981	27138030	7	158049
	4	4_7X0	27529455	27551288	4	21833
	4	4_8X0	27763472	27866289	3	102817
	4	4_9X0	28220633	28305384	12	84751
	5	5_10X0	51970151	52044180	2	74029
	5	5_11X0	57372842	57501507	2	128665
	5	5_12X0	58528129	58550730	2	22601
	5	5_13X0	59776006	60272317	4	496311
	5	5_14X0	60487858	60840023	8	352165

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	5	5_15X0	61188565	61254256	2	65691
	5	5_16X0	64240491	64324102	3	83611
	5	5_17X0	64629700	64660237	2	30537
	5	5_18X0	64794349	64917765	2	123416
	5	5_19X0	64989796	65758594	7	768798
	5	5_1X0	33212	436731	3	403519
	5	5_20X0	65821479	66406396	3	584917
	5	5_21X0	85650955	85751647	2	100692
	5	5_22X0	86783877	87112844	3	328967
	5	5_23X0	97723224	98170094	8	446870
	5	5_24X0	101765303	101921844	2	156541
	5	5_25X0	154466961	154518292	3	51331
	5	5_26X0	163629967	163660425	2	30458
	5	5_27X0	165209379	165300671	3	91292
	5	5_28X0	168437914	168684091	2	246177
	5	5_29X0	183728375	183731240	2	2865
	5	5_2X0	5790055	6031691	2	241636
	5	5_30X0	186737041	186796452	2	59411
	5	5_31X0	188585296	188640881	2	55585
	5	5_3X0	6866918	6933158	3	66240
	5	5_4X0	8846137	9100224	2	254087
	5	5_5X0	15014622	15146609	2	131987
	5	5_6X0	15254249	15283603	2	29354
	5	5_7X0	38420183	38477080	3	56897
	5	5_8X0	48073565	48248417	2	174852
	5	5_9X0	48873247	48942282	2	69035
	6	6_10X0	75787720	75850059	2	62339
	6	6_11X0	80290540	80491819	2	201279
	6	6_12X0	93751393	94006490	3	255097
	6	6_13X0	99499289	99756270	2	256981
	6	6_14X0	119842917	119895982	2	53065
	6	6_15X0	129984049	130098581	2	114532
	6	6_16X0	138612471	139742391	14	1129920
	6	6_17X0	146451826	146471177	2	19351
	6	6_18X0	146617729	146945758	2	328029
	6	6_19X0	148187885	148423916	3	236031
	6	6_1X0	1104408	1335115	3	230707
	6	6_20X0	149372640	149585661	3	213021
	6	6_21X0	155076323	155261188	2	184865
	6	6_22X0	173806005	174053650	2	247645
	6	6_23X0	177222773	177655476	6	432703
	6	6_24X0	179688434	180023263	6	334829
	6	6_2X0	17969646	17985425	3	15779
	6	6_3X0	24808770	24847673	2	38903
	6	6_4X0	36304442	36408819	2	104377
	6	6_5X0	43735264	43805927	4	70663
	6	6_6X0	55939390	56018484	2	79094
	6	6_7X0	56777664	56895184	2	117520
	6	6_8X0	67592590	67778619	2	186029
	6	6_9X0	73276060	73398552	2	122492
	7	7_10X0	19618245	19721133	3	102888
	7	7_11X0	21396124	21467918	2	71794
	7	7_12X0	25066892	25188113	2	121221
	7	7_13X0	26781485	26891788	2	110303
	7	7_14X0	28148250	28283761	2	135511

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	7	7_15X0	42009619	42094773	2	85154
	7	7_16X0	50158778	50302128	2	143350
	7	7_17X0	51327715	51370808	2	43093
	7	7_18X0	51830633	52066722	5	236089
	7	7_19X0	52067355	52247095	2	179740
	7	7_1X0	265580	344466	2	78886
	7	7_20X0	55970667	56039722	3	69055
	7	7_21X0	62742032	62956031	2	213999
	7	7_22X0	67772073	67805595	2	33522
	7	7_23X0	80689454	81451890	17	762436
	7	7_24X0	81723149	82168928	10	445779
	7	7_25X0	82856589	83790535	31	933946
	7	7_26X0	84373252	84390256	2	17004
	7	7_27X0	84734495	84786086	2	51591
	7	7_28X0	85419680	85752185	9	332505
	7	7_29X0	85813921	85856958	2	43037
	7	7_2X0	448331	501285	2	52954
	7	7_30X0	85907612	86030974	3	123362
	7	7_31X0	112344677	112540053	2	195376
	7	7_32X0	114199260	114250936	2	51676
	7	7_33X0	126425395	126469132	2	43737
	7	7_34X0	129842097	129902213	2	60116
	7	7_35X0	132448186	132818685	5	370499
	7	7_36X0	135927485	135972339	2	44854
	7	7_37X0	137296954	137513685	2	216731
	7	7_38X0	137822988	137944642	2	121654
	7	7_39X0	156540600	156899480	8	358880
	7	7_3X0	702302	744749	2	42447
	7	7_40X0	157963093	157985672	2	22579
	7	7_41X0	158479212	158543624	2	64412
	7	7_42X0	165398131	165436292	2	38161
	7	7_43X0	166060309	166073396	2	13087
	7	7_44X0	168800930	169544212	24	743282
	7	7_44X1	168040903	168737827	12	696924
	7	7_4X0	6695644	6758478	2	62834
	7	7_5X0	8225284	8439101	2	213817
	7	7_6X0	16481657	16599416	2	117759
	7	7_7X0	17876463	17953536	2	77073
	7	7_8X0	18241495	18553429	4	311934
	7	7_9X0	18824778	19000378	2	175600
	8	8_10X0	53899344	54073708	2	174364
	8	8_11X0	79548509	79808496	4	259987
	8	8_12X0	79893071	79955690	2	62619
	8	8_13X0	83645619	83845623	2	200004
	8	8_14X0	141939978	142462583	10	522605
	8	8_15X0	142487349	142942358	3	455009
	8	8_16X0	143345549	143533854	2	188305
	8	8_17X0	143579433	144287666	8	708233
	8	8_1X0	7160820	7249528	4	88708
	8	8_2X0	7311414	7318063	2	6649
	8	8_3X0	7325493	7783871	10	458378
	8	8_4X0	17961794	18147148	2	185354
	8	8_5X0	22995672	23245314	3	249642
	8	8_6X0	23682824	23714068	2	31244
	8	8_7X0	24322606	24549908	3	227302

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	8	8_8X0	39447589	40388354	5	940765
	8	8_9X0	40457840	40557197	2	99357
	9	9_10X0	41803182	41830563	2	27381
	9	9_11X0	45111246	45207772	2	96526
	9	9_12X0	48223310	48243202	2	19892
	9	9_13X0	52253184	52287636	2	34452
	9	9_14X0	52636135	52729141	4	93006
	9	9_15X0	63056637	63237027	2	180390
	9	9_16X0	63476671	63560983	2	84312
	9	9_17X0	83778833	83998147	5	219314
	9	9_18X0	84486216	84582174	4	95958
	9	9_19X0	88273278	88290574	2	17296
	9	9_1X0	913528	941976	2	28448
	9	9_20X0	89898366	90425459	5	527093
	9	9_21X0	92820319	92837894	2	17575
	9	9_22X0	99260846	99273693	2	12847
	9	9_23X0	100080199	100127969	2	47770
	9	9_24X0	112459157	112559723	3	100566
	9	9_25X0	118622983	118846691	2	223708
	9	9_2X0	4806984	5119927	4	312943
	9	9_3X0	5363761	5390259	2	26498
	9	9_4X0	5511469	5763490	2	252021
	9	9_5X0	7540123	7702331	2	162208
	9	9_6X0	15096648	15188304	4	91656
	9	9_7X0	36429212	36526730	2	97518
	9	9_8X0	37355142	37596586	3	241444
	9	9_9X0	38976991	39066107	3	89116
	X	X_10X0	47549556	47656198	2	106642
	X	X_11X0	48018946	48248612	3	229666
	X	X_12X0	48291285	48655837	9	364552
	X	X_13X0	49404362	49432108	2	27746
	X	X_14X0	49548220	49838189	4	289969
	X	X_15X0	51318443	51522830	2	204387
	X	X_16X0	51965278	52215023	2	249745
	X	X_17X0	52235844	52276468	2	40624
	X	X_18X0	54235419	54361592	2	126173
	X	X_19X0	54567967	54831547	6	263580
	X	X_1X0	1948730	2161963	4	213233
	X	X_20X0	66023397	66155234	3	131837
	X	X_21X0	66415731	66485087	2	69356
	X	X_22X0	67509636	67696324	3	186688
	X	X_23X0	72618093	73163639	3	545546
	X	X_24X0	95735627	95842409	4	106782
	X	X_25X0	96185710	96201475	2	15765
	X	X_26X0	96235167	96260533	2	25366
	X	X_27X0	96410566	96575499	4	164933
	X	X_28X0	96811438	97378239	10	566801
	X	X_29X0	97681180	97907035	5	225855
	X	X_2X0	12062000	12116636	2	54636
	X	X_30X0	110186598	110389354	2	202756
	X	X_31X0	113830518	113881081	2	50563
	X	X_32X0	114580677	114587031	2	6354
	X	X_33X0	124292250	124325807	2	33557
	X	X_34X0	128747401	128831352	2	83951
	X	X_35X0	128967392	129411763	5	444371

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	X	X_36X0	135575959	135725349	4	149390
	X	X_37X0	137152598	137345108	2	192510
	X	X_38X0	143330159	143568867	4	238708
	X	X_39X0	144200523	144369786	2	169263
	X	X_3X0	14535447	14615904	2	80457
	X	X_40X0	146370818	146436197	5	65379
	X	X_41X0	146814184	146894902	2	80718
	X	X_42X0	147136484	147645981	5	509497
	X	X_43X0	147954961	147992780	2	37819
	X	X_4X0	23749843	23799068	2	49225
	X	X_5X0	25596541	25674062	4	77521
	X	X_6X0	27277210	27548236	2	271026
	X	X_7X0	29844572	29878773	4	34201
	X	X_8X0	41646373	41683956	2	37583
	X	X_9X0	43706588	43977723	2	271135
	Y	Y_1X0	8812707	8904928	2	92221
	Y	Y_2X0	9880133	10069059	5	188926
	1	1_100X0	422367952	422421200	2	53248
	1	1_101X0	423445557	423780386	4	334829
	1	1_102X0	423857986	423907912	2	49926
	1	1_103X0	424058012	424549205	8	491193
	1	1_103X1	424728685	425261645	4	532960
	1	1_104X0	427039549	427168140	2	128591
	1	1_105X0	436577238	436727707	2	150469
	1	1_106X0	446202775	446427046	3	224271
	1	1_107X0	447482720	447554919	3	72199
	1	1_108X0	448585857	448676099	3	90242
	1	1_109X0	449446423	449477493	2	31070
	1	1_10X0	25374083	25403160	2	29077
	1	1_110X0	450245258	450382836	2	137578
	1	1_111X0	450823878	451103034	2	279156
	1	1_112X0	452742989	452799100	2	56111
	1	1_113X0	452830272	452969001	2	138729
	1	1_114X0	453298184	453332933	2	34749
	1	1_115X0	456067741	456086845	2	19104
	1	1_116X0	460277763	460443080	2	165317
	1	1_117X0	460621050	460654230	2	33180
	1	1_118X0	467276506	467690244	6	413738
	1	1_119X0	467941998	468483534	8	541536
	1	1_11X0	38583197	39122130	7	538933
	1	1_120X0	468793176	468834847	2	41671
	1	1_121X0	469163420	469247486	5	84066
	1	1_122X0	472322030	472408735	2	86705
	1	1_123X0	473616008	473724141	2	108133
	1	1_124X0	480065039	480068691	2	3652
	1	1_125X0	482389499	482575582	2	186083
	1	1_126X0	500942573	501387594	3	445021
	1	1_127X0	503076118	503161801	2	85683
	1	1_128X0	508690287	508964768	2	274481
	1	1_129X0	524002961	524195706	2	192745
	1	1_12X0	51850227	51947342	3	97115
	1	1_130X0	527094269	527251009	6	156740
	1	1_131X0	561992900	562031356	3	38456
	1	1_132X0	562960190	563037000	2	76810
	1	1_133X0	564232992	564509493	2	276501

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_134X0	566232477	566242491	2	10014
	1	1_135X0	574040213	574642321	4	602108
	1	1_136X0	574712956	574878074	2	165118
	1	1_137X0	578293248	578354141	2	60893
	1	1_138X0	591006982	591338507	7	331525
	1	1_139X0	591437139	591624446	2	187307
	1	1_13X0	58150944	58271251	2	120307
	1	1_140X0	598108311	598167088	2	58777
	1	1_141X0	600244740	600369388	2	124648
	1	1_142X0	650470940	650782998	2	312058
	1	1_143X0	653513967	653658790	2	144823
	1	1_144X0	655027464	655082154	2	54690
	1	1_145X0	657765974	657886203	2	120229
	1	1_146X0	661345080	661626283	2	281203
	1	1_147X0	668880033	669114393	4	234360
	1	1_148X0	669147301	669732031	13	584730
	1	1_149X0	669738928	669844579	2	105651
	1	1_14X0	59479151	59698335	2	219184
	1	1_150X0	677557999	677698383	2	140384
	1	1_151X0	683713015	684050057	6	337042
	1	1_152X0	684347676	684445964	2	98288
	1	1_153X0	684760119	685029173	2	269054
	1	1_154X0	691481736	691611129	3	129393
	1	1_155X0	695000105	695280858	2	280753
	1	1_156X0	697587972	697660377	2	72405
	1	1_157X0	698572510	698713705	2	141195
	1	1_158X0	703523979	703839215	5	315236
	1	1_159X0	705293348	705693318	6	399970
	1	1_15X0	59758224	60179533	6	421309
	1	1_160X0	706989483	707022055	2	32572
	1	1_161X0	710519288	710689056	3	169768
	1	1_162X0	711315114	711821171	8	506057
	1	1_163X0	712257005	712689751	3	432746
	1	1_164X0	712733801	713263462	20	529661
	1	1_165X0	714128911	714776346	11	647435
	1	1_165X1	713287528	713951450	7	663922
	1	1_165X2	714937039	715379022	4	441983
	1	1_166X0	715429485	715902454	13	472969
	1	1_167X0	716608305	716677051	2	68746
	1	1_168X0	718189042	718314318	2	125276
	1	1_169X0	729096059	729178031	3	81972
	1	1_16X0	60283380	60332966	2	49586
	1	1_170X0	737606142	737640540	2	34398
	1	1_171X0	738252380	738514097	3	261717
	1	1_172X0	739261608	739307126	2	45518
	1	1_173X0	739815551	739886863	2	71312
	1	1_174X0	739960170	740450991	11	490821
	1	1_175X0	745213361	745682828	7	469467
	1	1_175X1	746047916	746211807	3	163891
	1	1_17X0	65447060	65473884	2	26824
	1	1_18X0	67004409	67042549	2	38140
	1	1_19X0	76489213	76680079	5	190866
	1	1_1X0	8847731	8941370	3	93639
	1	1_20X0	82181556	82294903	4	113347
	1	1_21X0	83133874	83325346	2	191472

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_22X0	84077064	84129743	2	52679
	1	1_23X0	84485594	84508595	2	23001
	1	1_24X0	84800076	85329695	10	529619
	1	1_25X0	90834595	91206009	2	371414
	1	1_26X0	93430656	93565140	3	134484
	1	1_27X0	107484259	107745054	2	260795
	1	1_28X0	108570469	108613768	3	43299
	1	1_29X0	111096843	111214706	2	117863
	1	1_2X0	10060250	10235314	3	175064
	1	1_30X0	112114927	112364297	3	249370
	1	1_31X0	117149674	117281888	2	132214
	1	1_32X0	121112605	121277728	2	165123
	1	1_33X0	121533211	121665507	2	132296
	1	1_34X0	122228637	122270079	2	41442
	1	1_35X0	123190487	123296907	2	106420
	1	1_36X0	123498003	123524259	2	26256
	1	1_37X0	151235723	151364330	2	128607
	1	1_38X0	157043441	157052180	2	8739
	1	1_39X0	163476084	163667264	2	191180
	1	1_3X0	13828644	14325613	8	496969
	1	1_40X0	166329081	168208656	47	1879575
	1	1_40X1	168244191	169189695	12	945504
	1	1_40X2	169498857	169667068	3	168211
	1	1_41X0	170212782	170385738	6	172956
	1	1_42X0	170467178	171634192	37	1167014
	1	1_43X0	171746301	171909355	5	163054
	1	1_44X0	173161124	173693430	8	532306
	1	1_45X0	173712491	173722825	2	10334
	1	1_46X0	181105906	181182202	2	76296
	1	1_47X0	182637711	182778666	2	140955
	1	1_48X0	196673848	196906130	2	232282
	1	1_49X0	198476511	198562520	2	86009
	1	1_4X0	16487669	17311834	26	824165
	1	1_50X0	199334643	199551397	2	216754
	1	1_51X0	200908630	201086588	2	177958
	1	1_52X0	203042567	203165940	4	123373
	1	1_53X0	203268960	203445166	2	176206
	1	1_54X0	222107840	222178308	2	70468
	1	1_55X0	247670483	248168550	3	498067
	1	1_56X0	248338992	248641642	2	302650
	1	1_57X0	250398982	250518691	3	119709
	1	1_58X0	262010808	262243076	3	232268
	1	1_59X0	267145687	267197939	2	52252
	1	1_5X0	20593118	20769860	2	176742
	1	1_60X0	272039248	272113070	2	73822
	1	1_61X0	275824649	275928604	2	103955
	1	1_62X0	287967759	288165057	2	197298
	1	1_63X0	294929032	295177494	2	248462
	1	1_64X0	297352689	297661583	4	308894
	1	1_65X0	297946596	298771689	13	825093
	1	1_66X0	306615747	306732397	2	116650
	1	1_67X0	306966870	307152760	2	185890
	1	1_68X0	312973116	313158920	2	185804
	1	1_69X0	318470218	318605884	2	135666
	1	1_6X0	21765170	21839986	2	74816

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_70X0	325600334	325685871	2	85537
	1	1_71X0	328157916	328638211	7	480295
	1	1_72X0	328647384	328741482	2	94098
	1	1_73X0	332292316	332319924	2	27608
	1	1_74X0	332637457	333339521	16	702064
	1	1_75X0	333698839	333808820	2	109981
	1	1_76X0	336860842	336905187	2	44345
	1	1_77X0	342879191	343235495	3	356304
	1	1_78X0	345209460	345539051	3	329591
	1	1_79X0	346385770	346403060	2	17290
	1	1_7X0	22458804	22600542	2	141738
	1	1_80X0	351857109	352013882	3	156773
	1	1_81X0	359857941	360156883	2	298942
	1	1_82X0	364256294	364259552	2	3258
	1	1_83X0	366797232	367043272	2	246040
	1	1_84X0	369413345	369469695	2	56350
	1	1_85X0	377965826	378185003	2	219177
	1	1_86X0	378518591	378736565	3	217974
	1	1_87X0	379038476	379267402	3	228926
	1	1_88X0	380434192	380612602	2	178410
	1	1_89X0	384854750	385116307	2	261557
	1	1_8X0	23042533	23350852	5	308319
	1	1_90X0	385459498	385561772	2	102274
	1	1_91X0	389063021	389138321	2	75300
	1	1_92X0	399394932	399700021	3	305089
	1	1_93X0	400111951	400283451	4	171500
	1	1_94X0	401380029	401674143	2	294114
	1	1_95X0	402840476	403034863	2	194387
	1	1_96X0	404401017	404546855	3	145838
	1	1_97X0	414417241	414625167	2	207926
	1	1_98X0	415906570	416138255	3	231685
	1	1_99X0	416168933	416683704	14	514771
	1	1_9X0	23352212	23592797	2	240585
	2	2_100X0	284679064	284777134	3	98070
	2	2_101X0	285842341	286083848	3	241507
	2	2_102X0	287170439	287375465	2	205026
	2	2_103X0	291786877	291815400	2	28523
	2	2_104X0	293990380	294231222	3	240842
	2	2_105X0	296539218	296969034	4	429816
	2	2_106X0	298460628	298632399	2	171771
	2	2_107X0	300545662	300629878	2	84216
	2	2_108X0	301326777	301548920	4	222143
	2	2_109X0	312522955	312846275	2	323320
	2	2_10X0	36491789	36837155	4	345366
	2	2_110X0	314010207	314295966	3	285759
	2	2_111X0	332354647	332712145	13	357498
	2	2_112X0	341398693	341443175	2	44482
	2	2_113X0	346430324	346623930	2	193606
	2	2_114X0	369135575	369265423	2	129848
	2	2_115X0	377020421	377189116	2	168695
	2	2_116X0	384574059	384751672	4	177613
	2	2_117X0	406713166	407116293	13	403127
	2	2_118X0	407136967	407176398	2	39431
	2	2_119X0	412753809	412997574	3	243765
	2	2_11X0	37999152	38027262	2	28110

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	2	2_120X0	431242916	431380856	2	137940
	2	2_121X0	442243855	442321336	3	77481
	2	2_122X0	442563452	442978064	3	414612
	2	2_123X0	451262038	451314739	2	52701
	2	2_124X0	451494662	451812402	2	317740
	2	2_125X0	456829654	457259107	3	429453
	2	2_126X0	458981153	459182710	2	201557
	2	2_127X0	466347750	466635297	2	287547
	2	2_128X0	471998786	472244729	2	245943
	2	2_129X0	477235853	477439237	2	203384
	2	2_12X0	38107615	38639215	6	531600
	2	2_130X0	478664373	478735358	2	70985
	2	2_131X0	478806289	478824655	2	18366
	2	2_132X0	480171926	480482082	3	310156
	2	2_133X0	481121540	481332128	7	210588
	2	2_134X0	488739000	489094300	3	355300
	2	2_135X0	491969194	492270894	3	301700
	2	2_136X0	493197216	493374604	6	177388
	2	2_137X0	493421635	493426268	2	4633
	2	2_138X0	493797770	493802342	2	4572
	2	2_139X0	493813155	494062422	2	249267
	2	2_13X0	41328377	41433462	2	105085
	2	2_140X0	494306094	494405211	2	99117
	2	2_141X0	494856488	495000068	5	143580
	2	2_142X0	495916736	495979495	2	62759
	2	2_143X0	496461344	496657330	2	195986
	2	2_144X0	497807659	498013410	3	205751
	2	2_144X1	497395409	497488086	2	92677
	2	2_145X0	498864656	499020992	9	156336
	2	2_146X0	500917520	500993124	3	75604
	2	2_147X0	505448037	505499653	2	51616
	2	2_148X0	506638529	506776918	2	138389
	2	2_149X0	514690760	514908861	2	218101
	2	2_14X0	43169438	43346414	2	176976
	2	2_150X0	515113265	515144838	2	31573
	2	2_151X0	518892146	519270090	6	377944
	2	2_152X0	521045982	521115093	3	69111
	2	2_153X0	521805003	522419218	10	614215
	2	2_154X0	523542878	523648624	2	105746
	2	2_155X0	523789760	523856003	2	66243
	2	2_156X0	524325425	524414180	2	88755
	2	2_157X0	525850338	525998491	2	148153
	2	2_158X0	528722843	528770451	2	47608
	2	2_159X0	530274795	530432458	3	157663
	2	2_15X0	52245331	52400503	4	155172
	2	2_160X0	531405874	531483278	2	77404
	2	2_161X0	532076152	532290326	4	214174
	2	2_162X0	532316023	532508373	2	192350
	2	2_163X0	533095001	533235701	2	140700
	2	2_164X0	535431126	535723943	7	292817
	2	2_165X0	537366524	537828154	6	461630
	2	2_166X0	538198109	538754445	6	556336
	2	2_167X0	539372242	539737288	4	365046
	2	2_168X0	539877272	540161442	2	284170
	2	2_16X0	54158121	54314937	2	156816

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	2	2_17X0	58303631	58441656	3	138025
	2	2_18X0	60627963	60654638	2	26675
	2	2_19X0	61819596	62046079	3	226483
	2	2_1X0	2003769	2075466	2	71697
	2	2_20X0	63345257	63657410	7	312153
	2	2_21X0	66154429	66334103	2	179674
	2	2_22X0	73835500	73891114	2	55614
	2	2_23X0	77175047	77451641	3	276594
	2	2_24X0	83156028	83161477	2	5449
	2	2_25X0	102285475	102707170	3	421695
	2	2_26X0	103068472	103292567	2	224095
	2	2_27X0	104958857	105062964	2	104107
	2	2_28X0	106008049	106291783	2	283734
	2	2_29X0	107145909	107168443	2	22534
	2	2_2X0	2397225	2513380	2	116155
	2	2_30X0	107458261	107576625	2	118364
	2	2_31X0	107744697	107963541	3	218844
	2	2_32X0	108789476	109465444	6	675968
	2	2_33X0	110585428	110879302	2	293874
	2	2_34X0	112662411	113079663	3	417252
	2	2_35X0	113139394	113377464	4	238070
	2	2_36X0	113380406	113450717	2	70311
	2	2_37X0	113719432	114177140	5	457708
	2	2_38X0	120835088	121011773	2	176685
	2	2_39X0	134976177	135271725	2	295548
	2	2_3X0	3004396	3190181	2	185785
	2	2_40X0	164317798	164546244	4	228446
	2	2_41X0	164625005	165189395	5	564390
	2	2_42X0	165770117	165839552	2	69435
	2	2_43X0	165939502	166222196	5	282694
	2	2_44X0	167396193	168330642	21	934449
	2	2_44X1	166715007	167338184	7	623177
	2	2_45X0	168478824	169276585	9	797761
	2	2_46X0	171005611	171129192	2	123581
	2	2_47X0	174202812	174373939	3	171127
	2	2_48X0	189629649	189662906	2	33257
	2	2_49X0	189676882	189791667	3	114785
	2	2_4X0	7154315	7187582	2	33267
	2	2_50X0	190440527	190702092	2	261565
	2	2_51X0	190755421	190927995	2	172574
	2	2_52X0	191924997	192022257	3	97260
	2	2_53X0	192999537	193218033	2	218496
	2	2_54X0	194024737	194214386	3	189649
	2	2_55X0	194268149	194564286	4	296137
	2	2_56X0	194983030	195458971	13	475941
	2	2_57X0	195972287	196076470	3	104183
	2	2_58X0	196193408	196759169	15	565761
	2	2_59X0	197211273	197474011	3	262738
	2	2_5X0	7458787	7854592	4	395805
	2	2_60X0	197597196	197858415	4	261219
	2	2_61X0	198066662	198346871	2	280209
	2	2_62X0	200127027	200235342	2	108315
	2	2_63X0	201101199	201365724	9	264525
	2	2_64X0	201484853	201507000	2	22147
	2	2_65X0	203256670	203302947	2	46277

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	2	2_66X0	203484901	203544192	2	59291
	2	2_67X0	203900532	203920647	2	20115
	2	2_68X0	210828866	210999570	2	170704
	2	2_69X0	213075910	213121840	2	45930
	2	2_6X0	15049257	15216689	2	167432
	2	2_70X0	214568792	214592962	2	24170
	2	2_71X0	222052174	222108593	2	56419
	2	2_72X0	228016137	228172310	2	156173
	2	2_73X0	231306172	231347245	2	41073
	2	2_74X0	232216828	232314723	2	97895
	2	2_75X0	236350698	236462603	2	111905
	2	2_76X0	239793780	239873508	2	79728
	2	2_77X0	250273125	250661065	6	387940
	2	2_77X1	250801548	251172080	6	370532
	2	2_77X2	249761007	250178947	4	417940
	2	2_78X0	251239203	251508429	8	269226
	2	2_79X0	251694342	251756752	2	62410
	2	2_7X0	15380776	15706478	7	325702
	2	2_80X0	253091912	253260555	3	168643
	2	2_81X0	254153238	254453115	4	299877
	2	2_82X0	255494945	255677085	3	182140
	2	2_83X0	255809118	255932246	2	123128
	2	2_84X0	257748746	257856937	3	108191
	2	2_85X0	257888326	258375340	12	487014
	2	2_86X0	258383457	258450957	2	67500
	2	2_87X0	258650739	259693314	14	1042575
	2	2_88X0	261422579	261899343	6	476764
	2	2_89X0	265479489	266210149	19	730660
	2	2_89X1	264904049	265445303	10	541254
	2	2_8X0	24066480	24241707	2	175227
	2	2_90X0	266311039	266779070	7	468031
	2	2_91X0	267432261	267558003	3	125742
	2	2_92X0	267628535	267804655	5	176120
	2	2_93X0	267981732	267987078	2	5346
	2	2_94X0	268413240	268885486	10	472246
	2	2_95X0	269111369	269728881	8	617512
	2	2_96X0	270106252	270430980	6	324728
	2	2_97X0	275178486	275348369	2	169883
	2	2_98X0	280146866	280382258	2	235392
	2	2_99X0	283574492	283887939	6	313447
	2	2_9X0	31438650	32070785	9	632135
	3	3_10X0	94393556	94396085	2	2529
	3	3_11X0	112604110	112756611	2	152501
	3	3_12X0	143603320	144480445	14	877125
	3	3_13X0	144551559	144617440	11	65881
	3	3_14X0	144768547	144959856	5	191309
	3	3_15X0	145416780	145786884	4	370104
	3	3_16X0	151767568	152220582	6	453014
	3	3_17X0	162436586	162736824	2	300238
	3	3_18X0	183699430	183789084	2	89654
	3	3_19X0	184001514	184254331	2	252817
	3	3_1X0	2308886	2449026	2	140140
	3	3_20X0	186009121	186035378	2	26257
	3	3_21X0	199990816	200114529	2	123713
	3	3_22X0	200748948	201490659	6	741711

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	3	3_23X0	230470651	230474964	2	4313
	3	3_24X0	231579225	231582996	2	3771
	3	3_25X0	282199347	282308632	2	109285
	3	3_26X0	300038674	300195990	2	157316
	3	3_27X0	302431077	302572690	2	141613
	3	3_28X0	331630980	331736656	2	105676
	3	3_29X0	333633802	333702607	2	68805
	3	3_2X0	11643161	11766800	2	123639
	3	3_30X0	343367858	343430583	2	62725
	3	3_31X0	345018485	345347842	3	329357
	3	3_32X0	346785909	347060180	2	274271
	3	3_33X0	362626632	362668300	2	41668
	3	3_34X0	378016852	378238887	2	222035
	3	3_35X0	392526698	392532183	2	5485
	3	3_36X0	400316809	400388346	2	71537
	3	3_37X0	406668723	406988658	5	319935
	3	3_38X0	424231521	424346826	2	115305
	3	3_39X0	428776794	428995158	5	218364
	3	3_3X0	18337333	18464806	2	127473
	3	3_40X0	429176551	429305895	2	129344
	3	3_41X0	430472869	430671698	3	198829
	3	3_42X0	430865325	431665649	4	800324
	3	3_43X0	432002681	432485550	6	482869
	3	3_44X0	432538411	433229661	7	691250
	3	3_45X0	433577723	434190900	7	613177
	3	3_46X0	434255385	434543581	2	288196
	3	3_47X0	435903760	435927400	2	23640
	3	3_48X0	436078629	436176443	4	97814
	3	3_49X0	436592067	436791849	3	199782
	3	3_4X0	19050009	19229098	2	179089
	3	3_50X0	437379176	437616126	3	236950
	3	3_51X0	438815320	439254032	4	438712
	3	3_52X0	439278754	439560157	4	281403
	3	3_53X0	440020486	440589636	3	569150
	3	3_54X0	443004712	443082565	2	77853
	3	3_55X0	444636654	444828383	2	191729
	3	3_56X0	445056945	446177368	8	1120423
	3	3_57X0	447534879	447923502	3	388623
	3	3_58X0	447952069	448179722	3	227653
	3	3_59X0	449391797	449507635	2	115838
	3	3_5X0	36920809	36981977	2	61168
	3	3_60X0	450226609	450602243	11	375634
	3	3_61X0	454082892	454332559	2	249667
	3	3_62X0	459013037	459057113	2	44076
	3	3_63X0	459648559	459821167	2	172608
	3	3_64X0	460598500	461446520	6	848020
	3	3_65X0	462905191	463096613	4	191422
	3	3_66X0	464110005	464506719	6	396714
	3	3_67X0	475944106	476065035	2	120929
	3	3_68X0	478123178	479166479	20	1043301
	3	3_68X1	476917001	477920208	13	1003207
	3	3_68X2	480388927	480434562	2	45635
	3	3_69X0	480972827	481008425	2	35598
	3	3_6X0	40520640	40898974	3	378334
	3	3_70X0	482320600	482381148	2	60548

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	3	3_71X0	484475100	484654826	3	179726
	3	3_72X0	486224264	486245634	2	21370
	3	3_73X0	490227773	490347754	2	119981
	3	3_74X0	493575202	494930278	18	1355076
	3	3_75X0	495034161	495295824	2	261663
	3	3_76X0	495734205	495788049	2	53844
	3	3_77X0	496195451	496439413	5	243962
	3	3_78X0	500497690	501268948	14	771258
	3	3_78X1	499519009	500400453	14	881444
	3	3_78X2	502931849	503528859	11	597010
	3	3_78X3	503602861	504459541	12	856680
	3	3_78X4	501950690	502617610	10	666920
	3	3_78X5	501342542	501783390	4	440848
	3	3_79X0	509527030	509815917	7	288887
	3	3_7X0	46065496	46261096	2	195600
	3	3_80X0	513321608	513705389	3	383781
	3	3_81X0	514161395	514168371	2	6976
	3	3_82X0	518847228	518943415	2	96187
	3	3_83X0	519483705	519836352	5	352647
	3	3_84X0	523886211	524525811	20	639600
	3	3_84X1	523356518	523871375	13	514857
	3	3_84X2	524512502	525216976	14	704474
	3	3_84X3	521714968	522659263	16	944295
	3	3_84X4	522733398	523267984	8	534586
	3	3_84X5	525413116	525903891	13	490775
	3	3_84X6	520738553	521645184	10	906631
	3	3_85X0	525942863	527012867	14	1070004
	3	3_86X0	527161035	527413146	7	252111
	3	3_87X0	527480568	527577123	7	96555
	3	3_8X0	56605138	56853376	4	248238
	3	3_9X0	91644557	91994580	3	350023
	4	4_10X0	68636528	68658557	2	22029
	4	4_11X0	73575796	73738152	2	162356
	4	4_12X0	86687828	86727367	2	39539
	4	4_13X0	87086018	87377997	3	291979
	4	4_14X0	89340816	89390419	2	49603
	4	4_15X0	90872329	91059405	2	187076
	4	4_16X0	94030410	94140074	2	109664
	4	4_17X0	94218280	94446626	2	228346
	4	4_18X0	102054418	102216894	2	162476
	4	4_19X0	124430853	124535513	2	104660
	4	4_1X0	6229411	6374624	2	145213
	4	4_20X0	153668926	153724306	2	55380
	4	4_21X0	176610994	176733784	2	122790
	4	4_22X0	182292747	182314378	2	21631
	4	4_23X0	187417681	187540762	9	123081
	4	4_24X0	205641166	205883275	2	242109
	4	4_25X0	208199290	208402301	2	203011
	4	4_26X0	222801260	224060274	32	1259014
	4	4_28X0	231719143	231765523	3	46380
	4	4_29X0	231822466	231969994	4	147528
	4	4_2X0	8347357	8463955	3	116598
	4	4_30X0	231987926	232324779	2	336853
	4	4_31X0	232433162	232649563	2	216401
	4	4_32X0	233867325	233920354	3	53029

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	4	4_33X0	237047544	237257217	2	209673
	4	4_34X0	238276723	238359109	2	82386
	4	4_35X0	240988423	241109240	2	120817
	4	4_36X0	241316336	241322341	2	6005
	4	4_37X0	250103436	250365549	6	262113
	4	4_38X0	258736645	258809483	2	72838
	4	4_39X0	258976469	259562430	8	585961
	4	4_3X0	9882867	10149082	2	266215
	4	4_40X0	270659572	271072246	18	412674
	4	4_40X1	271113994	271576590	14	462596
	4	4_41X0	271830394	272005241	3	174847
	4	4_42X0	278471712	278629247	2	157535
	4	4_43X0	280670556	280771209	2	100653
	4	4_44X0	287664560	287724816	3	60256
	4	4_45X0	287921712	288126774	3	205062
	4	4_46X0	289097245	289225803	3	128558
	4	4_47X0	298302797	298500675	3	197878
	4	4_48X0	304203517	304331747	2	128230
	4	4_49X0	304985727	304988615	2	2888
	4	4_4X0	11131637	11191037	2	59400
	4	4_50X0	322975254	323216419	4	241165
	4	4_51X0	335259135	335450689	2	191554
	4	4_52X0	337561676	337705141	2	143465
	4	4_53X0	340395774	340575315	2	179541
	4	4_54X0	342405284	342580170	2	174886
	4	4_55X0	343427383	343864728	6	437345
	4	4_56X0	345193475	345244428	2	50953
	4	4_57X0	351009274	351882484	36	873210
	4	4_57X1	351999319	352607242	30	607923
	4	4_57X2	353420178	354546870	33	1126692
	4	4_57X3	350193753	350945204	26	751451
	4	4_57X4	352678696	353401311	23	722615
	4	4_57X5	349118336	350096776	21	978440
	4	4_57X6	348008774	348528870	7	520096
	4	4_57X7	347434995	347742740	4	307745
	4	4_57X8	346152482	346332816	2	180334
	4	4_58X0	355261794	355281831	2	20037
	4	4_59X0	357572247	357645875	2	73628
	4	4_5X0	13069991	13141092	2	71101
	4	4_60X0	358492245	358781882	6	289637
	4	4_61X0	360641183	360693446	2	52263
	4	4_62X0	365688946	365739736	3	50790
	4	4_63X0	366942174	367051537	2	109363
	4	4_64X0	367759495	367779811	2	20316
	4	4_65X0	370005149	370640062	9	634913
	4	4_65X1	369209273	369971090	7	761817
	4	4_65X2	368507977	368919007	3	411030
	4	4_66X0	371714865	371842104	2	127239
	4	4_67X0	373501541	373599977	2	98436
	4	4_68X0	374136185	374600565	5	464380
	4	4_69X0	376124876	376161462	2	36586
	4	4_6X0	14326433	14586665	4	260232
	4	4_70X0	376944383	377015555	2	71172
	4	4_71X0	377465373	377482862	2	17489
	4	4_72X0	377531232	377926814	9	395582

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	4	4_73X0	378038333	378076923	2	38590
	4	4_74X0	378176928	378462337	3	285409
	4	4_75X0	382386534	382497747	2	111213
	4	4_76X0	382669438	383228509	9	559071
	4	4_77X0	406792831	407475420	23	682589
	4	4_77X1	409272089	409867131	16	595042
	4	4_77X10	399663716	400410861	13	747145
	4	4_77X11	398408151	398855948	8	447797
	4	4_77X12	410275623	410786043	9	510420
	4	4_77X13	413286874	413983096	12	696222
	4	4_77X14	383877534	384615962	12	738428
	4	4_77X15	408451675	409050182	10	598507
	4	4_77X16	403851237	404270004	7	418767
	4	4_77X17	404320525	404686290	6	365765
	4	4_77X18	405121575	405910024	8	788449
	4	4_77X19	411000893	411550756	8	549863
	4	4_77X2	397480620	398335239	21	854619
	4	4_77X20	384937019	385291431	7	354412
	4	4_77X21	411588506	412038499	4	449993
	4	4_77X22	407489390	408270507	7	781117
	4	4_77X23	401365828	401776270	5	410442
	4	4_77X24	386388036	386922518	5	534482
	4	4_77X25	412100372	412694707	5	594335
	4	4_77X26	401798564	402141026	4	342462
	4	4_77X27	402182296	402770085	4	587789
	4	4_77X28	400664122	401197881	5	533759
	4	4_77X29	386962221	387501409	4	539188
	4	4_77X3	403360710	403779535	10	418825
	4	4_77X30	385560170	385847670	4	287500
	4	4_77X31	396676788	396877423	3	200635
	4	4_77X32	388059305	388069598	3	10293
	4	4_77X4	409873153	410220847	7	347694
	4	4_77X5	405808093	406663702	14	855609
	4	4_77X6	402822588	403301797	9	479209
	4	4_77X7	398911282	399522369	12	611087
	4	4_77X8	404738644	405178928	8	440284
	4	4_77X9	383243949	383610461	12	366512
	4	4_78X0	415111149	415188923	2	77774
	4	4_79X0	417584933	417637638	4	52705
	4	4_7X0	14833079	14927168	2	94089
	4	4_80X0	418687168	418970094	3	282926
	4	4_81X0	421657940	421719997	2	62057
	4	4_82X0	423521429	423564234	2	42805
	4	4_83X0	425628394	425729407	2	101013
	4	4_84X0	425804702	425939098	2	134396
	4	4_85X0	428513858	428541602	2	27744
	4	4_86X0	429213527	429232973	2	19446
	4	4_87X0	429771965	429986356	5	214391
	4	4_88X0	430031677	430376142	3	344465
	4	4_89X0	433688125	433797584	2	109459
	4	4_8X0	19328224	19625495	2	297271
	4	4_9X0	21891920	21947212	2	55292
	5	5_10X0	23086162	23173853	2	87691
	5	5_11X0	27052674	27118629	2	65955
	5	5_12X0	31659288	31772837	2	113549

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	5	5_13X0	32516441	32593137	2	76696
	5	5_14X0	33543316	33687399	2	144083
	5	5_15X0	34846674	34905637	2	58963
	5	5_16X0	46800244	46982442	2	182198
	5	5_17X0	51618231	52086700	6	468469
	5	5_18X0	78560404	78642441	2	82037
	5	5_19X0	80667489	80731862	2	64373
	5	5_1X0	1693959	1804034	2	110075
	5	5_20X0	82376246	82379300	2	3054
	5	5_21X0	87146109	87429419	2	283310
	5	5_22X0	100114340	100741124	8	626784
	5	5_23X0	101293523	101322689	2	29166
	5	5_24X0	110254677	110352567	2	97890
	5	5_25X0	120564380	120638049	3	73669
	5	5_26X0	120741861	120902096	2	160235
	5	5_27X0	122235155	122301677	3	66522
	5	5_28X0	154263000	154505661	3	242661
	5	5_29X0	154663611	154896226	2	232615
	5	5_2X0	3164868	3218980	3	54112
	5	5_30X0	155075565	155795041	7	719476
	5	5_31X0	174176246	174273834	2	97588
	5	5_32X0	174856735	175079115	2	222380
	5	5_33X0	185150455	185434095	2	283640
	5	5_34X0	202601638	202709118	2	107480
	5	5_35X0	211848944	212087999	2	239055
	5	5_36X0	213532001	213566921	2	34920
	5	5_37X0	213659888	213822297	2	162409
	5	5_38X0	228309165	228328481	2	19316
	5	5_39X0	237426247	237471296	2	45049
	5	5_3X0	5583831	5788811	2	204980
	5	5_40X0	237689607	237717198	2	27591
	5	5_41X0	238518489	238551629	3	33140
	5	5_42X0	241635147	242043907	11	408760
	5	5_43X0	242367763	242677426	4	309663
	5	5_44X0	289822123	291178020	48	1355897
	5	5_44X1	286678651	287557470	30	878819
	5	5_44X2	288639511	289660032	31	1020521
	5	5_44X3	287568941	288618981	26	1050040
	5	5_44X4	285022991	285765759	20	742768
	5	5_44X5	285817177	286664232	16	847055
	5	5_44X6	254349270	254391787	3	42517
	5	5_44X7	280499930	280549700	2	49770
	5	5_45X0	291533138	291568902	3	35764
	5	5_46X0	291591808	291739292	2	147484
	5	5_47X0	292011868	292850291	32	838423
	5	5_48X0	294192183	295134901	20	942718
	5	5_48X1	293297713	294068881	15	771168
	5	5_49X0	295273525	295320917	2	47392
	5	5_4X0	9949596	10056218	2	106622
	5	5_50X0	295512649	295805976	5	293327
	5	5_51X0	295865194	296480733	15	615539
	5	5_52X0	296506798	297232527	8	725729
	5	5_53X0	297275455	298292590	12	1017135
	5	5_54X0	298457368	298909490	8	452122
	5	5_55X0	299110179	299249931	2	139752

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	5	5_56X0	299671414	299736399	2	64985
	5	5_57X0	300549994	300654862	3	104868
	5	5_58X0	300844005	301175304	3	331299
	5	5_59X0	301923441	302733969	11	810528
	5	5_5X0	12669359	12899776	3	230417
	5	5_60X0	303671561	303719422	2	47861
	5	5_61X0	304488292	304612433	3	124141
	5	5_6X0	18169650	18212074	2	42424
	5	5_7X0	20924861	21038281	3	113420
	5	5_8X0	21251775	21396257	3	144482
	5	5_9X0	21642603	21874879	3	232276
	6	6_10X0	58317074	58525276	2	208202
	6	6_11X0	59600676	59637734	2	37058
	6	6_12X0	59722360	60018139	4	295779
	6	6_13X0	62738212	62834571	2	96359
	6	6_14X0	69543952	69641724	2	97772
	6	6_15X0	80548698	80814644	4	265946
	6	6_16X0	83481810	83588390	2	106580
	6	6_17X0	85651351	85960433	7	309082
	6	6_17X1	85131160	85387323	4	256163
	6	6_18X0	85963223	86023203	3	59980
	6	6_19X0	99635121	99907995	3	272874
	6	6_1X0	4334958	4403471	2	68513
	6	6_20X0	104993020	105036280	2	43260
	6	6_21X0	119141989	119205147	2	63158
	6	6_22X0	119530104	119607450	2	77346
	6	6_23X0	120297473	120352066	2	54593
	6	6_24X0	122554899	122793581	2	238682
	6	6_25X0	124030508	124052681	2	22173
	6	6_26X0	134894273	135150303	12	256030
	6	6_27X0	141236610	141968313	18	731703
	6	6_28X0	146919080	147665261	19	746181
	6	6_29X0	147753199	147756704	2	3505
	6	6_2X0	16959323	17078882	2	119559
	6	6_30X0	148026051	148144806	6	118755
	6	6_31X0	148552314	148686548	3	134234
	6	6_32X0	149396422	149447414	4	50992
	6	6_33X0	150666230	150987011	4	320781
	6	6_34X0	151342446	151461413	2	118967
	6	6_35X0	155062071	155223256	3	161185
	6	6_36X0	157464085	157611189	2	147104
	6	6_37X0	158202942	158367682	2	164740
	6	6_38X0	158640325	158774045	2	133720
	6	6_39X0	160472912	160613438	2	140526
	6	6_3X0	26402604	26496653	2	94049
	6	6_40X0	163267476	163343326	2	75850
	6	6_41X0	167082640	167221527	2	138887
	6	6_42X0	190455693	190830647	2	374954
	6	6_43X0	193130111	193225867	2	95756
	6	6_44X0	195029250	195083400	2	54150
	6	6_45X0	196346131	196507306	3	161175
	6	6_46X0	197659597	197742237	3	82640
	6	6_47X0	198153832	198233621	2	79789
	6	6_48X0	199685576	200174636	6	489060
	6	6_49X0	200206528	200233888	2	27360

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	6	6_4X0	26632576	26769021	3	136445
	6	6_50X0	201195825	201264440	2	68615
	6	6_51X0	203334509	203679825	3	345316
	6	6_52X0	203745136	203873467	2	128331
	6	6_53X0	204069448	204324641	2	255193
	6	6_54X0	204348106	204401543	3	53437
	6	6_55X0	242711106	242744017	2	32911
	6	6_56X0	273394526	273436480	2	41954
	6	6_57X0	274904415	274999119	2	94704
	6	6_58X0	275076223	275094461	2	18238
	6	6_59X0	275998784	276213137	3	214353
	6	6_5X0	35511085	35662828	5	151743
	6	6_60X0	277014357	277381782	2	367425
	6	6_61X0	277754122	277771393	2	17271
	6	6_62X0	278888725	278986598	2	97873
	6	6_63X0	283847974	283888151	3	40177
	6	6_64X0	288151353	288623510	3	472157
	6	6_65X0	291638600	291751962	2	113362
	6	6_6X0	36297238	36336868	2	39630
	6	6_7X0	43732288	43985014	2	252726
	6	6_8X0	51562449	51690526	3	128077
	6	6_9X0	52952306	53186354	2	234048
	7	7_10X0	60708523	60796118	2	87595
	7	7_11X0	63957305	64750886	10	793581
	7	7_12X0	65591021	66136167	6	545146
	7	7_13X0	73293162	73535599	5	242437
	7	7_14X0	77793419	77934447	3	141028
	7	7_15X0	84044936	84284186	4	239250
	7	7_16X0	87629205	87903736	2	274531
	7	7_17X0	93903085	93936628	3	33543
	7	7_18X0	97598758	97626309	2	27551
	7	7_19X0	98340941	98759965	3	419024
	7	7_1X0	16810375	16816649	2	6274
	7	7_20X0	111212521	111345617	2	133096
	7	7_21X0	149143148	149149972	2	6824
	7	7_22X0	171396300	171687760	4	291460
	7	7_23X0	173073828	173134880	2	61052
	7	7_24X0	173143095	173393173	4	250078
	7	7_25X0	173825605	173887036	2	61431
	7	7_26X0	174479235	174598998	3	119763
	7	7_27X0	176420021	176496540	2	76519
	7	7_28X0	177185449	177347015	2	161566
	7	7_29X0	192090339	192416520	3	326181
	7	7_2X0	18941428	19235470	10	294042
	7	7_30X0	206841185	207034826	4	193641
	7	7_31X0	207349310	207516907	5	167597
	7	7_32X0	209516965	209640400	3	123435
	7	7_33X0	210052534	210138083	2	85549
	7	7_34X0	211738093	211765872	2	27779
	7	7_35X0	212616950	213003482	4	386532
	7	7_36X0	232396755	232460313	3	63558
	7	7_37X0	242746147	243021928	4	275781
	7	7_38X0	250265880	250450048	3	184168
	7	7_39X0	250524568	250642450	3	117882
	7	7_3X0	24889815	25022744	2	132929

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	7	7_40X0	256914766	257075247	2	160481
	7	7_41X0	260534534	260854726	4	320192
	7	7_4X0	32610321	32632765	2	22444
	7	7_5X0	46805620	47077298	4	271678
	7	7_6X0	47275679	47402050	2	126371
	7	7_7X0	49447390	49828321	3	380931
	7	7_8X0	52278591	52390517	2	111926
	7	7_9X0	53430724	53539857	2	109133
	8	8_10X0	26288900	26441708	2	152808
	8	8_11X0	28481007	28634995	2	153988
	8	8_12X0	29487504	29572812	2	85308
	8	8_13X0	29709452	29831292	2	121840
	8	8_14X0	31073147	31925666	15	852519
	8	8_15X0	33709917	34003258	8	293341
	8	8_16X0	34418021	34510963	4	92942
	8	8_17X0	35700724	36472960	14	772236
	8	8_18X0	36555381	36626420	2	71039
	8	8_19X0	44181571	44329217	3	147646
	8	8_1X0	10224051	10262213	2	38162
	8	8_20X0	54961258	55039642	2	78384
	8	8_21X0	67200831	67474419	4	273588
	8	8_22X0	79499931	79742603	2	242672
	8	8_23X0	80506375	80835191	4	328816
	8	8_24X0	91637386	91910982	4	273596
	8	8_25X0	92516758	92655215	2	138457
	8	8_26X0	97655840	97899851	2	244011
	8	8_27X0	105169817	105222916	2	53099
	8	8_28X0	105296785	105617290	3	320505
	8	8_29X0	107654952	107763404	2	108452
	8	8_2X0	11270239	11492668	2	222429
	8	8_30X0	108172191	108258922	2	86731
	8	8_31X0	108535844	108624191	3	88347
	8	8_32X0	110968941	111227953	2	259012
	8	8_33X0	111348962	111432847	2	83885
	8	8_34X0	113474625	113849610	6	374985
	8	8_35X0	114146633	114284130	2	137497
	8	8_36X0	116894546	116995280	2	100734
	8	8_37X0	121178199	121358533	2	180334
	8	8_38X0	121761906	122038747	2	276841
	8	8_39X0	122488085	122905056	3	416971
	8	8_3X0	12683416	12702804	2	19388
	8	8_40X0	129230897	129431385	2	200488
	8	8_41X0	159517389	159678975	2	161586
	8	8_42X0	159833412	160076643	2	243231
	8	8_43X0	174330894	174691536	5	360642
	8	8_44X0	181920805	182120898	3	200093
	8	8_45X0	188491416	188563233	2	71817
	8	8_46X0	190295767	190395231	3	99464
	8	8_47X0	192874802	193000456	2	125654
	8	8_48X0	201569232	201696401	2	127169
	8	8_49X0	204320975	204583618	4	262643
	8	8_4X0	13000178	13112991	2	112813
	8	8_50X0	204627118	204836985	3	209867
	8	8_51X0	204871162	204905255	2	34093
	8	8_52X0	204930250	205433162	21	502912

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	8	8_53X0	205494486	205585367	2	90881
	8	8_54X0	205775088	205908711	5	133623
	8	8_55X0	206189807	207357678	14	1167871
	8	8_56X0	213308335	213589653	7	281318
	8	8_56X1	213888317	214326619	5	438302
	8	8_57X0	214354999	214658454	3	303455
	8	8_58X0	214976940	215215880	6	238940
	8	8_59X0	215594194	215853903	2	259709
	8	8_5X0	13308611	13540077	2	231466
	8	8_60X0	228958362	229069066	3	110704
	8	8_61X0	245487976	245723537	2	235561
	8	8_62X0	249941132	250069984	2	128852
	8	8_63X0	259421285	259941484	3	520199
	8	8_64X0	262500029	262628597	2	128568
	8	8_65X0	266020785	266190850	4	170065
	8	8_66X0	266632871	266790293	2	157422
	8	8_67X0	271058486	271574611	7	516125
	8	8_68X0	271589088	271724539	2	135451
	8	8_69X0	274636898	275189449	3	552551
	8	8_6X0	13776874	14101565	3	324691
	8	8_70X0	283745768	284064963	2	319195
	8	8_71X0	284115845	284420003	2	304158
	8	8_72X0	292251651	292414847	3	163196
	8	8_73X0	297227600	297429569	11	201969
	8	8_74X0	300227620	300391539	2	163919
	8	8_75X0	304069648	304139676	2	70028
	8	8_76X0	306062196	306117729	2	55533
	8	8_77X0	311255240	311441313	2	186073
	8	8_7X0	15092722	15310459	2	217737
	8	8_8X0	17179240	17291465	2	112225
	8	8_9X0	23142176	23196571	2	54395
	Un	Un_10X0	14679317	14793232	2	113915
	Un	Un_11X0	15043299	15129643	2	86344
	Un	Un_12X0	16593547	16923637	5	330090
	Un	Un_13X0	17052906	17296070	9	243164
	Un	Un_14X0	17540639	17732336	2	191697
	Un	Un_15X0	17933807	17982549	2	48742
	Un	Un_16X0	22449061	22534880	3	85819
	Un	Un_17X0	23346908	23387569	2	40661
	Un	Un_18X0	24748100	24986437	3	238337
	Un	Un_19X0	25590584	25719698	7	129114
	Un	Un_1X0	231287	260450	2	29163
	Un	Un_20X0	26274688	26365435	3	90747
	Un	Un_21X0	26618413	26748834	3	130421
	Un	Un_22X0	26900104	26979843	3	79739
	Un	Un_23X0	27764598	27788499	2	23901
	Un	Un_24X0	30212587	30310363	5	97776
	Un	Un_25X0	30671467	30772511	2	101044
	Un	Un_26X0	30925688	31126119	5	200431
	Un	Un_27X0	37130722	37308380	2	177658
	Un	Un_28X0	37553322	37651826	2	98504
	Un	Un_29X0	41439657	41494299	2	54642
	Un	Un_2X0	6886953	6963741	2	76788
	Un	Un_30X0	43057180	43135434	2	78254
	Un	Un_31X0	43298894	43404698	2	105804

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	Un	Un_32X0	43479837	43610158	2	130321
	Un	Un_33X0	44138522	44183291	3	44769
	Un	Un_34X0	45911330	45971631	2	60301
	Un	Un_35X0	49216362	49250299	2	33937
	Un	Un_36X0	50396467	50621203	4	224736
	Un	Un_37X0	50656506	50690903	2	34397
	Un	Un_38X0	52259754	52343110	3	83356
	Un	Un_39X0	52399928	52428514	4	28586
	Un	Un_3X0	7888584	7970629	2	82045
	Un	Un_40X0	52631243	52707703	2	76460
	Un	Un_41X0	52756516	52786781	2	30265
	Un	Un_42X0	53293228	53317321	3	24093
	Un	Un_43X0	53732483	53775676	4	43193
	Un	Un_44X0	54509060	54535762	2	26702
	Un	Un_45X0	55186950	55193196	2	6246
	Un	Un_46X0	55789330	55999940	4	210610
	Un	Un_46X1	56132324	56272546	2	140222
	Un	Un_46X2	56412985	56531497	2	118512
	Un	Un_47X0	58893235	59038404	2	145169
	Un	Un_48X0	67146240	67282436	2	136196
	Un	Un_49X0	76814471	76976746	2	162275
	Un	Un_4X0	8188660	8331363	4	142703
	Un	Un_50X0	83278669	83382463	2	103794
	Un	Un_51X0	87745337	87913958	2	168621
	Un	Un_52X0	92134438	92256029	2	121591
	Un	Un_5X0	8398623	8470407	2	71784
	Un	Un_6X0	9502265	9791492	6	289227
	Un	Un_6X1	9246779	9500052	4	253273
	Un	Un_7X0	9931202	10086927	2	155725
	Un	Un_8X0	10131621	10202905	2	71284
	Un	Un_9X0	13719846	13896203	2	176357
	X	X_10X0	32568123	32718207	5	150084
	X	X_11X0	35982925	36053789	2	70864
	X	X_12X0	39010656	39124199	2	113543
	X	X_13X0	39846700	40018106	2	171406
	X	X_14X0	42203468	42635756	4	432288
	X	X_15X0	42739747	42862902	2	123155
	X	X_16X0	43782299	43808557	2	26258
	X	X_17X0	56516128	57108936	5	592808
	X	X_18X0	69461976	69581479	2	119503
	X	X_19X0	69642559	69710977	3	68418
	X	X_1X0	3167955	3197265	2	29310
	X	X_20X0	70671827	70774347	3	102520
	X	X_21X0	71527369	71580736	2	53367
	X	X_22X0	71639844	72092433	6	452589
	X	X_23X0	72310089	72350285	2	40196
	X	X_2X0	3518646	3538121	2	19475
	X	X_3X0	3627247	3699969	3	72722
	X	X_4X0	5272421	5315105	2	42684
	X	X_5X0	7156418	7337017	2	180599
	X	X_6X0	10816683	10919647	2	102964
	X	X_7X0	21673456	22044378	3	370922
	X	X_8X0	25259452	25376289	3	116837
	X	X_9X0	26319142	26394198	2	75056
	1	1_10X0	40580081	40727681	2	147600

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_11X0	53297095	53318035	2	20940
	1	1_12X0	60746358	61000320	3	253962
	1	1_13X0	65100619	65123193	2	22574
	1	1_14X0	72825322	72874884	2	49562
	1	1_15X0	74153989	74279624	3	125635
	1	1_16X0	74771892	74804179	2	32287
	1	1_17X0	87086695	87127912	3	41217
	1	1_18X0	87190844	87208295	2	17451
	1	1_19X0	88055481	88220002	9	164521
	1	1_1X0	18236473	18265080	2	28607
	1	1_20X0	92479681	92986391	9	506710
	1	1_21X0	93478993	93509733	2	30740
	1	1_22X0	93861346	94017147	3	155801
	1	1_23X0	106870234	107610484	12	740250
	1	1_24X0	117697131	118202307	6	505176
	1	1_25X0	130576716	130661610	2	84894
	1	1_26X0	130800902	130880791	3	79889
	1	1_27X0	130882074	131024974	4	142900
	1	1_28X0	131056889	131276951	3	220062
	1	1_29X0	133323966	133329722	2	5756
	1	1_2X0	19103022	19234624	2	131602
	1	1_30X0	133619862	133661380	4	41518
	1	1_31X0	133900834	134048277	3	147443
	1	1_32X0	134138646	134190029	2	51383
	1	1_33X0	135805057	135852260	2	47203
	1	1_34X0	153665709	153844142	3	178433
	1	1_35X0	153849542	153900208	3	50666
	1	1_36X0	156004599	156068861	2	64262
	1	1_37X0	156524012	156642716	2	118704
	1	1_38X0	161017756	161131511	2	113755
	1	1_39X0	162793883	162984527	5	190644
	1	1_3X0	20730905	20790617	2	59712
	1	1_40X0	164048234	164150026	3	101792
	1	1_41X0	166483613	166681835	3	198222
	1	1_42X0	170232749	170286769	2	54020
	1	1_43X0	170907273	171948073	14	1040800
	1	1_44X0	172082832	172196389	2	113557
	1	1_45X0	172223513	172298064	2	74551
	1	1_46X0	172322087	172374085	2	51998
	1	1_47X0	172475358	172602551	5	127193
	1	1_48X0	172956768	173271137	6	314369
	1	1_49X0	173455354	174031755	13	576401
	1	1_4X0	21349641	21386373	3	36732
	1	1_50X0	174041118	174449602	13	408484
	1	1_51X0	180893108	180938400	2	45292
	1	1_52X0	184786776	184846451	2	59675
	1	1_53X0	191098325	191218039	2	119714
	1	1_54X0	195036826	195171168	3	134342
	1	1_5X0	34498410	34560920	2	62510
	1	1_6X0	36471620	36528237	2	56617
	1	1_7X0	36530193	36547231	3	17038
	1	1_8X0	37905923	37957759	2	51836
	1	1_9X0	40084698	40551705	6	467007
	10	10_10X0	51480640	51496613	3	15973
	10	10_11X0	58222972	58236701	3	13729

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	10	10_12X0	62429209	62651218	3	222009
	10	10_13X0	63024512	63077962	2	53450
	10	10_14X0	74990841	75054061	2	63220
	10	10_15X0	75771233	75860240	7	89007
	10	10_16X0	75954514	76189265	3	234751
	10	10_17X0	77689957	77862882	23	172925
	10	10_18X0	77897575	77986905	2	89330
	10	10_19X0	78517428	79035802	10	518374
	10	10_1X0	19591949	19760053	3	168104
	10	10_20X0	79148797	79492154	4	343357
	10	10_21X0	79689020	79892655	5	203635
	10	10_22X0	80167778	80288408	2	120630
	10	10_23X0	80602896	80615783	2	12887
	10	10_24X0	80629656	80671976	2	42320
	10	10_25X0	81119340	81262975	3	143635
	10	10_26X0	81367509	81600900	5	233391
	10	10_27X0	81607025	82379636	12	772611
	10	10_28X0	86763972	86838389	2	74417
	10	10_29X0	97482367	97682454	4	200087
	10	10_2X0	22158586	22374139	3	215553
	10	10_30X0	100335676	100382560	6	46884
	10	10_31X0	104144317	104196203	7	51886
	10	10_32X0	107482908	107494737	2	11829
	10	10_33X0	111985448	112160327	4	174879
	10	10_34X0	117233761	117292868	3	59107
	10	10_35X0	118204942	118354105	10	149163
	10	10_36X0	127348857	127370431	2	21574
	10	10_37X0	127759721	128000678	9	240957
	10	10_38X0	128254150	128270153	2	16003
	10	10_39X0	128490860	128744075	5	253215
	10	10_3X0	23851462	23905343	2	53881
	10	10_40X0	129540539	130211149	34	670610
	10	10_40X1	128989431	129527167	28	537736
	10	10_41X0	130385316	130497379	4	112063
	10	10_4X0	23920356	24109534	15	189178
	10	10_5X0	33766424	33879475	2	113051
	10	10_6X0	34041223	34127984	3	86761
	10	10_7X0	34282190	34301320	2	19130
	10	10_8X0	39890018	40142255	6	252237
	10	10_9X0	45289402	45372479	2	83077
	11	11_10X0	52144593	52213987	2	69394
	11	11_11X0	53864542	54028090	3	163548
	11	11_12X0	55124815	55233837	3	109022
	11	11_13X0	58199556	58222771	2	23215
	11	11_14X0	58307069	58323346	2	16277
	11	11_15X0	58379043	58842042	26	462999
	11	11_16X0	58847158	58904225	3	57067
	11	11_17X0	58917908	58991457	4	73549
	11	11_18X0	59248033	59333552	2	85519
	11	11_19X0	59461197	59506644	2	45447
	11	11_1X0	4031783	4118831	3	87048
	11	11_20X0	59570699	59614225	3	43526
	11	11_21X0	61208745	61267185	2	58440
	11	11_22X0	61301631	61378345	2	76714
	11	11_23X0	62711265	62789462	2	78197

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	11	11_24X0	67078300	67371502	6	293202
	11	11_25X0	68432121	68552698	2	120577
	11	11_26X0	69125896	69197809	3	71913
	11	11_27X0	69682932	69695809	2	12877
	11	11_28X0	69816566	69822180	2	5614
	11	11_29X0	70055881	70170811	4	114930
	11	11_2X0	7197782	7213923	2	16141
	11	11_30X0	70213910	70220994	2	7084
	11	11_31X0	70248963	70352031	3	103068
	11	11_32X0	71092236	71201298	2	109062
	11	11_33X0	72451638	72550301	2	98663
	11	11_34X0	73026585	73138294	2	111709
	11	11_35X0	73240310	73300363	2	60053
	11	11_36X0	73774849	74045272	12	270423
	11	11_36X1	74053776	74365609	10	311833
	11	11_36X2	73329741	73603520	11	273779
	11	11_37X0	77044292	77151925	2	107633
	11	11_38X0	78073376	78293745	4	220369
	11	11_39X0	79513385	79530589	2	17204
	11	11_3X0	22834744	22982184	2	147440
	11	11_40X0	82035571	82179812	6	144241
	11	11_41X0	82954661	83286726	8	332065
	11	11_42X0	83525778	83664683	5	138905
	11	11_43X0	83703991	83711348	2	7357
	11	11_44X0	85832551	85916097	2	83546
	11	11_45X0	87794209	87875536	3	81327
	11	11_46X0	87888959	87915003	3	26044
	11	11_47X0	93949817	93968293	2	18476
	11	11_48X0	94284652	94339509	2	54857
	11	11_49X0	94489599	94507355	2	17756
	11	11_4X0	32276893	32300873	5	23980
	11	11_50X0	94565047	94645216	2	80169
	11	11_51X0	94996064	95026087	2	30023
	11	11_52X0	95120119	95146263	2	26144
	11	11_53X0	96194316	96368256	9	173940
	11	11_54X0	96954219	97024731	2	70512
	11	11_55X0	98626360	98677708	3	51348
	11	11_56X0	98753586	98974942	3	221356
	11	11_57X0	99279959	99543158	11	263199
	11	11_58X0	99550130	99566630	4	16500
	11	11_59X0	99579976	99986597	30	406621
	11	11_5X0	46454969	46844332	6	389363
	11	11_60X0	100011195	100269871	15	258676
	11	11_61X0	100441927	100472741	2	30814
	11	11_62X0	100545694	100613342	2	67648
	11	11_63X0	100623011	100738215	2	115204
	11	11_64X0	100780731	100939479	3	158748
	11	11_65X0	101119938	101161787	2	41849
	11	11_66X0	101325087	101341938	2	16851
	11	11_67X0	101406839	101424796	3	17957
	11	11_68X0	102017336	102088515	2	71179
	11	11_69X0	102099930	102107416	2	7486
	11	11_6X0	48806404	48850555	2	44151
	11	11_70X0	103116231	103139876	2	23645
	11	11_71X0	103727364	103817957	2	90593

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	11	11_72X0	105975208	106005443	2	30235
	11	11_73X0	107703218	107911362	3	208144
	11	11_74X0	109934273	110337622	5	403349
	11	11_75X0	111023860	111076821	2	52961
	11	11_76X0	114890041	115133992	11	243951
	11	11_77X0	115322925	115346926	2	24001
	11	11_78X0	116105752	116131126	2	25374
	11	11_79X0	116675303	116775507	2	100204
	11	11_7X0	48865249	49294314	14	429065
	11	11_7X1	49317548	49564860	13	247312
	11	11_80X0	119022962	119086221	3	63259
	11	11_81X0	120725399	120731944	2	6545
	11	11_82X0	121036747	121081130	3	44383
	11	11_83X0	121116215	121148319	2	32104
	11	11_84X0	121421401	121443505	2	22104
	11	11_8X0	50814564	51070685	11	256121
	11	11_9X0	51234125	51256264	2	22139
	12	12_10X0	72939892	73113456	3	173564
	12	12_11X0	76348900	76396950	2	48050
	12	12_12X0	79175551	79212755	2	37204
	12	12_13X0	81376991	81568474	6	191483
	12	12_14X0	83987861	84104453	6	116592
	12	12_15X0	85473890	85709087	3	235197
	12	12_16X0	87494721	88321481	16	826760
	12	12_17X0	102719975	102758672	3	38697
	12	12_18X0	103616675	104414329	22	797654
	12	12_19X0	105147033	105222793	6	75760
	12	12_1X0	4715549	4769267	2	53718
	12	12_20X0	105563226	105605428	2	42202
	12	12_21X0	108825873	108856815	2	30942
	12	12_22X0	111421561	111454678	2	33117
	12	12_23X0	113489511	113546465	2	56954
	12	12_24X0	114406477	114416895	2	10418
	12	12_2X0	31438219	31559969	2	121750
	12	12_3X0	33928425	33959829	2	31404
	12	12_4X0	35925620	36004087	2	78467
	12	12_5X0	55089202	55263020	5	173818
	12	12_6X0	56531958	56613284	2	81326
	12	12_7X0	69241832	69310687	2	68855
	12	12_8X0	69746848	69893163	2	146315
	12	12_9X0	72618954	72664909	2	45955
	13	13_10X0	27057570	28151611	22	1094041
	13	13_11X0	32845673	33335370	10	489697
	13	13_12X0	33395761	33936083	6	540322
	13	13_13X0	34074274	34130354	2	56080
	13	13_14X0	34948084	34994144	2	46060
	13	13_15X0	49504810	49624501	4	119691
	13	13_16X0	55530410	55681256	2	150846
	13	13_17X0	59700083	59744917	4	44834
	13	13_18X0	60781887	61570127	11	788240
	13	13_19X0	62453472	62607499	3	154027
	13	13_1X0	3803485	3854730	3	51245
	13	13_20X0	62700108	62777089	2	76981
	13	13_21X0	65138089	65153152	2	15063
	13	13_22X0	66119547	66227573	2	108026

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	13	13_23X0	67038596	67843113	27	804517
	13	13_24X0	73536747	73700780	3	164033
	13	13_25X0	74371426	74482943	2	111517
	13	13_26X0	93616675	93674302	2	57627
	13	13_27X0	95511732	95702739	3	191007
	13	13_28X0	99319647	99411328	2	91681
	13	13_29X0	100144063	100446511	4	302448
	13	13_2X0	4059591	4609174	9	549583
	13	13_30X0	100498114	100616971	2	118857
	13	13_31X0	100651343	100656059	2	4716
	13	13_32X0	111425932	111549947	2	124015
	13	13_33X0	112464081	112562074	2	97993
	13	13_34X0	113093827	113180897	2	87070
	13	13_35X0	119335528	119408997	2	73469
	13	13_36X0	119828089	120318178	14	490089
	13	13_3X0	8885994	8960945	3	74951
	13	13_4X0	21072349	21703965	19	631616
	13	13_5X0	21715763	22043676	22	327913
	13	13_6X0	22087878	22621135	18	533257
	13	13_7X0	22772399	23250323	16	477924
	13	13_8X0	23463748	23488847	2	25099
	13	13_9X0	23531050	23934851	35	403801
	14	14_10X0	30886521	30943289	3	56768
	14	14_11X0	33923587	33947198	2	23611
	14	14_12X0	33988417	34076405	2	87988
	14	14_13X0	37054830	37073743	2	18913
	14	14_14X0	41049117	41092197	2	43080
	14	14_15X0	41131788	41185149	3	53361
	14	14_16X0	41278805	41772441	13	493636
	14	14_16X1	41901813	42524282	15	622469
	14	14_17X0	42594502	43115190	10	520688
	14	14_18X0	43179867	43784745	14	604878
	14	14_19X0	43818913	44103534	5	284621
	14	14_1X0	3030007	4087136	20	1057129
	14	14_20X0	44123851	44750340	14	626489
	14	14_21X0	44851235	45003820	2	152585
	14	14_22X0	49894258	51256112	44	1361854
	14	14_23X0	51361756	51922773	14	561017
	14	14_24X0	52035082	52495325	8	460243
	14	14_25X0	54883441	54898137	2	14696
	14	14_26X0	54941926	54994626	2	52700
	14	14_27X0	55578575	55591101	2	12526
	14	14_28X0	55650207	55788857	4	138650
	14	14_29X0	55854036	55896781	2	42745
	14	14_2X0	4983998	6125742	27	1141744
	14	14_30X0	55941453	56262260	15	320807
	14	14_31X0	57034460	57133281	3	98821
	14	14_32X0	59244441	59395381	5	150940
	14	14_33X0	62977524	63037846	4	60322
	14	14_34X0	64075445	64113755	3	38310
	14	14_35X0	68083863	68124846	2	40983
	14	14_36X0	68497336	68655842	3	158506
	14	14_37X0	69171802	69194662	2	22860
	14	14_38X0	79426511	79454816	2	28305
	14	14_39X0	121911253	121965195	2	53942

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	14	14_3X0	6216224	6720251	10	504027
	14	14_40X0	122456795	122479852	2	23057
	14	14_4X0	6736742	7643972	20	907230
	14	14_5X0	12283976	12324454	2	40478
	14	14_6X0	14175095	14408198	3	233103
	14	14_7X0	19532896	19602581	3	69685
	14	14_8X0	21684946	21748626	2	63680
	14	14_9X0	25942539	26269436	6	326897
	15	15_10X0	79229391	79256674	2	27283
	15	15_11X0	82371615	82794294	7	422679
	15	15_12X0	83491874	83595157	2	103283
	15	15_13X0	84112621	84186206	2	73585
	15	15_14X0	84232043	84342978	2	110935
	15	15_15X0	88862186	89149628	3	287442
	15	15_16X0	98186013	98498406	11	312393
	15	15_17X0	98672217	98894123	5	221906
	15	15_18X0	98931425	99038110	4	106685
	15	15_19X0	99579055	99605477	3	26422
	15	15_1X0	9279829	9370955	2	91126
	15	15_20X0	100304789	100384436	5	79647
	15	15_21X0	101128522	101213684	2	85162
	15	15_22X0	101348323	102032027	27	683704
	15	15_23X0	102356781	102436404	2	79623
	15	15_24X0	102921103	103036852	9	115749
	15	15_25X0	103419968	103448459	2	28491
	15	15_2X0	63775971	63902936	4	126965
	15	15_3X0	74714839	75599481	23	884642
	15	15_4X0	75940952	75975734	2	34782
	15	15_5X0	76535721	76639958	2	104237
	15	15_6X0	76853010	76918010	3	65000
	15	15_7X0	77045278	77757003	14	711725
	15	15_8X0	78283446	78493756	5	210310
	15	15_9X0	78899667	78919032	2	19365
	16	16_10X0	30256378	30310779	3	54401
	16	16_11X0	35832874	35972605	2	139731
	16	16_12X0	36122169	36455392	10	333223
	16	16_13X0	36603869	36726741	2	122872
	16	16_14X0	38347022	38362297	2	15275
	16	16_15X0	44765736	44960867	4	195131
	16	16_16X0	45382135	45532050	3	149915
	16	16_17X0	48283787	48319723	2	35936
	16	16_18X0	48816856	48874496	3	57640
	16	16_19X0	58613675	58638739	2	25064
	16	16_1X0	3747887	3872396	3	124509
	16	16_20X0	58717433	59362216	26	644783
	16	16_21X0	85795476	85901828	2	106352
	16	16_22X0	88505807	88563183	2	57376
	16	16_23X0	88610705	88829484	13	218779
	16	16_24X0	88868917	89206394	19	337477
	16	16_25X0	89403218	89420111	2	16893
	16	16_26X0	91225457	91271933	2	46476
	16	16_27X0	91372899	91565623	5	192724
	16	16_28X0	92295616	92358874	2	63258
	16	16_29X0	93607889	93690990	3	83101
	16	16_2X0	3884619	3923912	3	39293

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	16	16_30X0	97447037	97560899	2	113862
	16	16_3X0	16860671	16869255	2	8584
	16	16_4X0	17894223	17900024	2	5801
	16	16_5X0	18620466	18629939	2	9473
	16	16_6X0	19285743	19625101	8	339358
	16	16_7X0	20611625	20646485	2	34860
	16	16_8X0	22921745	23080170	4	158425
	16	16_9X0	26356642	26482763	2	126121
	17	17_10X0	19361949	19812536	5	450587
	17	17_11X0	20540517	20914363	9	373846
	17	17_11X1	20029425	20503195	9	473770
	17	17_11X2	20993727	21314912	5	321185
	17	17_12X0	21383748	21845759	12	462011
	17	17_12X1	21879570	22007301	4	127731
	17	17_12X2	22144359	22455838	5	311479
	17	17_13X0	22547941	22619133	2	71192
	17	17_14X0	22943184	23077425	2	134241
	17	17_15X0	23290934	23479597	4	188663
	17	17_16X0	23556769	23609698	3	52929
	17	17_17X0	23679365	23684018	2	4653
	17	17_18X0	23833360	24045954	12	212594
	17	17_19X0	24264279	24410201	2	145922
	17	17_1X0	7352064	7385464	2	33400
	17	17_20X0	24473550	24718057	4	244507
	17	17_21X0	25298394	25374385	6	75991
	17	17_22X0	25736040	25754327	2	18287
	17	17_23X0	25823627	25919642	4	96015
	17	17_24X0	26138828	26211324	2	72496
	17	17_25X0	28692568	28778698	2	86130
	17	17_26X0	29519281	29549593	2	30312
	17	17_27X0	31125306	31165053	3	39747
	17	17_28X0	32303676	32350581	2	46905
	17	17_29X0	32468462	33394641	24	926179
	17	17_2X0	12419972	12675838	3	255866
	17	17_30X0	33919332	34600896	20	681564
	17	17_31X0	34856384	34872507	3	16123
	17	17_32X0	34956436	34979285	3	22849
	17	17_33X0	35071443	35115428	4	43985
	17	17_34X0	35194626	35205351	3	10725
	17	17_35X0	35262730	35474563	7	211833
	17	17_36X0	36006269	36286490	9	280221
	17	17_37X0	36772910	37023199	11	250289
	17	17_37X1	36323201	36672219	6	349018
	17	17_38X0	37085191	37467152	20	381961
	17	17_38X1	37492529	37851553	17	359024
	17	17_38X2	37903522	38336097	13	432575
	17	17_39X0	40221777	40319207	2	97430
	17	17_3X0	12841115	13052280	5	211165
	17	17_40X0	42657091	42742179	2	85088
	17	17_41X0	44078845	44105764	2	26919
	17	17_42X0	47385404	47400584	2	15180
	17	17_43X0	48125605	48367021	9	241416
	17	17_44X0	55864892	55949500	3	84608
	17	17_45X0	56100591	56290511	3	189920
	17	17_46X0	56562939	56606294	2	43355

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	17	17_47X0	56843103	56888904	2	45801
	17	17_48X0	57032925	57053582	2	20657
	17	17_49X0	78852564	79020816	2	168252
	17	17_4X0	13221188	13236454	2	15266
	17	17_50X0	84658488	84698963	2	40475
	17	17_51X0	85621031	85688254	2	67223
	17	17_5X0	13948373	14001765	2	53392
	17	17_6X0	14995787	15041537	2	45750
	17	17_7X0	17064113	17146831	2	82718
	17	17_8X0	17876471	18022704	4	146233
	17	17_9X0	18049424	18598157	7	548733
	18	18_10X0	62566463	62741387	3	174924
	18	18_11X0	68337603	68429251	2	91648
	18	18_12X0	78101649	78206440	2	104791
	18	18_13X0	84610509	84681996	2	71487
	18	18_1X0	19962970	20604521	9	641551
	18	18_2X0	23954705	24020320	3	65615
	18	18_3X0	34579846	34624601	2	44755
	18	18_4X0	36766530	36791694	2	25164
	18	18_5X0	36930285	38284357	61	1354072
	18	18_6X0	43728069	44206018	8	477949
	18	18_7X0	52615982	52663733	2	47751
	18	18_8X0	60212080	60392627	4	180547
	18	18_9X0	61045200	61132149	2	86949
	19	19_10X0	8062209	8218673	2	156464
	19	19_11X0	8591254	8628299	2	37045
	19	19_12X0	8735839	8741223	2	5384
	19	19_13X0	8883732	8895064	4	11332
	19	19_14X0	10041548	10196870	3	155322
	19	19_15X0	10718143	10829856	2	111713
	19	19_16X0	10997025	11392790	10	395765
	19	19_17X0	11414841	11640834	8	225993
	19	19_18X0	11647280	11691150	2	43870
	19	19_19X0	13445014	13920044	22	475030
	19	19_19X1	11828077	12395212	21	567135
	19	19_19X2	12861742	13103284	12	241542
	19	19_19X3	13145642	13411500	8	265858
	19	19_19X4	12422265	12719117	8	296852
	19	19_1X0	3913493	3981645	3	68152
	19	19_20X0	20601961	20727562	2	125601
	19	19_21X0	25505706	25679010	3	173304
	19	19_22X0	29321344	29334670	2	13326
	19	19_23X0	29367455	29471157	2	103702
	19	19_24X0	33459609	33517776	2	58167
	19	19_24X1	33555167	33769142	2	213975
	19	19_25X0	33776546	34122687	5	346141
	19	19_26X0	34550710	34650009	6	99299
	19	19_27X0	37685581	37701528	2	15947
	19	19_28X0	39007019	39186756	3	179737
	19	19_29X0	39287074	39568529	3	281455
	19	19_2X0	4035407	4059295	3	23888
	19	19_30X0	39608842	39886769	4	277927
	19	19_31X0	39992424	40187286	4	194862
	19	19_32X0	41829970	41848132	2	18162
	19	19_33X0	44203288	44407709	4	204421

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	19	19_34X0	45150680	45235812	2	85132
	19	19_35X0	47105353	47157764	3	52411
	19	19_36X0	47854970	47886324	2	31354
	19	19_37X0	58670365	58777533	3	107168
	19	19_3X0	4083519	4139609	2	56090
	19	19_4X0	4756571	4811585	4	55014
	19	19_5X0	5689131	5771401	2	82270
	19	19_6X0	6067842	6080324	2	12482
	19	19_7X0	6306456	6345742	2	39286
	19	19_8X0	7557477	7639642	2	82165
	19	19_9X0	7673061	7802667	2	129606
	2	2_10X0	71528113	71546754	2	18641
	2	2_11X0	74655616	74765142	11	109526
	2	2_12X0	75935849	75981967	3	46118
	2	2_13X0	83955693	83993045	7	37352
	2	2_14X0	84600781	84678041	2	77260
	2	2_15X0	84838850	84944586	2	105736
	2	2_16X0	84980461	84993886	2	13425
	2	2_17X0	85136225	86456269	61	1320044
	2	2_17X1	87387656	88006776	35	619120
	2	2_17X10	104071066	104115349	2	44283
	2	2_17X11	98662237	98667301	2	5064
	2	2_17X12	93835438	93869157	2	33719
	2	2_17X2	88755741	89125729	21	369988
	2	2_17X3	88387062	88746886	18	359824
	2	2_17X4	88024702	88344727	14	320025
	2	2_17X5	86482799	87372478	37	889679
	2	2_17X6	89143508	89618245	21	474737
	2	2_17X7	89629940	90401674	28	771734
	2	2_17X8	111230818	111665814	22	434996
	2	2_17X9	111754251	112157007	17	402756
	2	2_18X0	119751328	119818104	2	66776
	2	2_19X0	120027507	120314165	5	286658
	2	2_1X0	24153161	24350936	6	197775
	2	2_20X0	121023650	121109795	3	86145
	2	2_21X0	121460235	121506665	2	46430
	2	2_22X0	122279247	122347972	4	68725
	2	2_23X0	122439747	122528040	2	88293
	2	2_24X0	125563088	125673502	2	110414
	2	2_25X0	126707360	126783444	2	76084
	2	2_26X0	127634874	127729932	2	95058
	2	2_27X0	128926268	129018563	2	92295
	2	2_28X0	129299610	129371139	2	71529
	2	2_29X0	130405239	130408256	2	3017
	2	2_2X0	25234983	25244261	2	9278
	2	2_30X0	130576173	130582554	2	6381
	2	2_31X0	131909928	131956130	3	46202
	2	2_32X0	147083316	147186404	2	103088
	2	2_33X0	148404466	148443563	2	39097
	2	2_34X0	148750396	148875692	5	125296
	2	2_35X0	150117156	150485091	10	367935
	2	2_36X0	151086786	151103041	2	16255
	2	2_37X0	152459055	152623053	12	163998
	2	2_38X0	153649454	153729907	2	80453
	2	2_39X0	153875045	154240902	6	365857

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	2	2_3X0	25466709	25825537	13	358828
	2	2_40X0	154257854	154320646	2	62792
	2	2_41X0	156091958	156180238	2	88280
	2	2_42X0	158258094	158332852	2	74758
	2	2_43X0	164177467	164718068	21	540601
	2	2_44X0	164805098	164826871	2	21773
	2	2_45X0	165351298	165388259	2	36961
	2	2_46X0	167607850	167661556	3	53706
	2	2_47X0	175010319	175432182	7	421863
	2	2_47X1	175651005	176237609	8	586604
	2	2_48X0	176317638	176811223	8	493585
	2	2_48X1	176916769	177202829	4	286060
	2	2_48X2	177264654	177578199	5	313545
	2	2_49X0	177768044	178078425	7	310381
	2	2_4X0	30952959	30967933	2	14974
	2	2_50X0	181193721	181213185	2	19464
	2	2_5X0	32587625	32620804	2	33179
	2	2_6X0	33406108	33462353	2	56245
	2	2_7X0	35354218	35423327	3	69109
	2	2_8X0	36376778	37520603	31	1143825
	2	2_9X0	51924451	51973004	2	48553
	3	3_10X0	83007861	83049863	3	42002
	3	3_11X0	86986586	86999441	2	12855
	3	3_12X0	87258939	87457836	3	198897
	3	3_13X0	87778244	87816052	2	37808
	3	3_14X0	88242874	88287210	2	44336
	3	3_15X0	88368622	88384433	3	15811
	3	3_16X0	89081939	89101099	2	19160
	3	3_17X0	89271740	89338028	3	66288
	3	3_18X0	89350219	89377549	2	27330
	3	3_19X0	92436809	92993426	23	556617
	3	3_1X0	10012548	10301174	5	288626
	3	3_20X0	93651542	94266784	13	615242
	3	3_21X0	94310089	94347350	2	37261
	3	3_22X0	95243060	95282076	2	39016
	3	3_23X0	95499256	95556403	2	57147
	3	3_24X0	96219865	96270132	9	50267
	3	3_25X0	98446608	98859794	8	413186
	3	3_26X0	100993532	101110278	2	116746
	3	3_27X0	105973711	106406182	6	432471
	3	3_28X0	107896247	108044859	7	148612
	3	3_29X0	108094448	108146146	2	51698
	3	3_2X0	14641727	14900770	3	259043
	3	3_30X0	109080469	109168457	2	87988
	3	3_31X0	113247178	113532428	5	285250
	3	3_32X0	116631164	116681260	2	50096
	3	3_33X0	135307745	135397827	2	90082
	3	3_34X0	138217760	138455499	6	237739
	3	3_35X0	142496900	142638008	5	141108
	3	3_36X0	144729916	145099106	7	369190
	3	3_37X0	146121672	146195512	2	73840
	3	3_38X0	151730922	151762892	2	31970
	3	3_3X0	20215620	20275733	2	60113
	3	3_4X0	30624267	30699843	3	75576
	3	3_5X0	59096448	59220259	5	123811

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	3	3_6X0	59729790	60040160	5	310370
	3	3_7X0	64081642	64844765	8	763123
	3	3_8X0	67457998	67515523	2	57525
	3	3_9X0	82032045	82145787	2	113742
	4	4_10X0	49381778	49503120	3	121342
	4	4_11X0	52825399	52971567	5	146168
	4	4_12X0	52989284	53021576	2	32292
	4	4_13X0	56740005	56744925	2	4920
	4	4_14X0	57943373	58009124	2	65751
	4	4_15X0	59003210	59041903	2	38693
	4	4_16X0	60418977	60821463	6	402486
	4	4_16X1	59956804	60222599	5	265795
	4	4_17X0	61224307	62150863	11	926556
	4	4_18X0	62291547	62391769	2	100222
	4	4_19X0	62408797	62470887	2	62090
	4	4_1X0	3172083	3185358	2	13275
	4	4_20X0	63143275	63365878	4	222603
	4	4_21X0	73401032	73920760	10	519728
	4	4_22X0	88522025	88880201	30	358176
	4	4_23X0	96042660	96664154	8	621494
	4	4_24X0	101835134	101990014	6	154880
	4	4_25X0	108181934	108301096	2	119162
	4	4_26X0	111919555	112300468	5	380913
	4	4_27X0	114906282	114926013	2	19731
	4	4_28X0	115108682	115641899	10	533217
	4	4_29X0	116173384	116299841	2	126457
	4	4_2X0	3995936	4076931	2	80995
	4	4_30X0	117724475	117765648	2	41173
	4	4_31X0	118665519	118934840	13	269321
	4	4_32X0	119637704	119658954	2	21250
	4	4_33X0	124656807	124708611	2	51804
	4	4_34X0	126332732	126533472	3	200740
	4	4_35X0	127311421	127358181	4	46760
	4	4_36X0	129353628	129440818	2	87190
	4	4_37X0	129489014	129558372	2	69358
	4	4_38X0	135686287	135752140	2	65853
	4	4_39X0	136880129	136898803	3	18674
	4	4_3X0	12089441	12146731	2	57290
	4	4_40X0	137401554	137430540	2	28986
	4	4_41X0	138216627	138261332	2	44705
	4	4_42X0	138725325	138882817	6	157492
	4	4_43X0	140727355	140952586	5	225231
	4	4_44X0	141384610	141416014	2	31404
	4	4_45X0	141683563	141723419	2	39856
	4	4_46X0	141814962	141846359	2	31397
	4	4_47X0	143394428	143853235	14	458807
	4	4_47X1	143894237	144408464	13	514227
	4	4_48X0	144453771	144786583	7	332812
	4	4_49X0	145514884	145899975	4	385091
	4	4_4X0	33063112	33163588	2	100476
	4	4_50X0	146097345	146198749	2	101404
	4	4_51X0	146449023	146539395	2	90372
	4	4_52X0	146971976	147848366	11	876390
	4	4_53X0	148145716	148166424	3	20708
	4	4_54X0	149485294	149564508	2	79214

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	4	4_55X0	150119283	150168482	2	49199
	4	4_56X0	152008803	152027661	2	18858
	4	4_57X0	155656943	155793173	2	136230
	4	4_5X0	41033074	41098183	2	65109
	4	4_6X0	41657498	41769474	2	111976
	4	4_7X0	42612195	42773993	5	161798
	4	4_8X0	43694784	43855463	6	160679
	4	4_9X0	45848947	45950774	2	101827
	5	5_10X0	31151142	31180117	2	28975
	5	5_11X0	37313269	37425055	2	111786
	5	5_12X0	43818891	43912375	2	93484
	5	5_13X0	64924680	64960047	2	35367
	5	5_14X0	72695980	72868446	2	172466
	5	5_15X0	73292895	73341028	2	48133
	5	5_16X0	86231752	86745815	7	514063
	5	5_17X0	86889767	87490871	11	601104
	5	5_18X0	87514689	87591594	3	76905
	5	5_19X0	87666233	87824321	3	158088
	5	5_1X0	5664831	5749326	2	84495
	5	5_20X0	90460897	90578605	4	117708
	5	5_21X0	90759378	90905938	7	146560
	5	5_22X0	92321347	92363277	3	41930
	5	5_23X0	93925078	94876892	13	951814
	5	5_24X0	94988626	95804708	11	816082
	5	5_25X0	99966027	100038869	2	72842
	5	5_26X0	103963461	104021838	2	58377
	5	5_27X0	105014153	105343929	7	329776
	5	5_28X0	108797193	109453387	10	656194
	5	5_29X0	109674545	110129102	7	454557
	5	5_2X0	11417626	11922788	7	505162
	5	5_2X1	10947805	11345021	6	397216
	5	5_30X0	110786094	110874145	2	88051
	5	5_31X0	112353801	112451738	2	97937
	5	5_32X0	114809036	114823468	3	14432
	5	5_33X0	114901353	114937915	2	36562
	5	5_34X0	120460688	120483932	2	23244
	5	5_35X0	120730333	120921652	9	191319
	5	5_36X0	121518613	121545482	2	26869
	5	5_37X0	121604572	121660514	2	55942
	5	5_38X0	123863570	123880020	2	16450
	5	5_39X0	134184027	134456716	3	272689
	5	5_3X0	14910122	15032998	4	122876
	5	5_40X0	134914249	134987472	3	73223
	5	5_41X0	135560448	135573049	2	12601
	5	5_42X0	136038496	136064326	2	25830
	5	5_43X0	137821952	137871758	3	49806
	5	5_44X0	138021276	138065767	3	44491
	5	5_45X0	138085084	138648836	6	563752
	5	5_46X0	139406387	139427800	2	21413
	5	5_47X0	139807978	139826843	2	18865
	5	5_48X0	143172186	143292356	6	120170
	5	5_49X0	145204668	145291278	3	86610
	5	5_4X0	15471703	15714271	6	242568
	5	5_50X0	145345287	146113285	8	767998
	5	5_51X0	147188696	147307270	3	118574

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	5	5_5X0	17373180	17480936	2	107756
	5	5_6X0	21519179	21645557	2	126378
	5	5_7X0	24479753	24540592	2	60839
	5	5_8X0	25181502	25265918	2	84416
	5	5_9X0	25935791	26264308	12	328517
	6	6_10X0	40491233	40613310	5	122077
	6	6_11X0	40895270	41525079	14	629809
	6	6_12X0	41621068	41680655	2	59587
	6	6_13X0	41847585	42042851	5	195266
	6	6_14X0	42140936	42435464	5	294528
	6	6_15X0	42460076	43236636	23	776560
	6	6_16X0	47819266	48086593	7	267327
	6	6_17X0	48436613	48633689	4	197076
	6	6_18X0	48647384	48766958	6	119574
	6	6_19X0	48835514	48844775	2	9261
	6	6_1X0	3993792	4008445	2	14653
	6	6_20X0	48904986	48991722	4	86736
	6	6_21X0	52162417	52260802	8	98385
	6	6_22X0	56956466	57432602	17	476136
	6	6_23X0	57842830	58472878	11	630048
	6	6_24X0	66552180	66805332	8	253152
	6	6_25X0	71322788	71379173	2	56385
	6	6_26X0	78370898	78468872	6	97974
	6	6_27X0	83068591	83088243	2	19652
	6	6_28X0	83644542	83677857	2	33315
	6	6_29X0	85707173	85915604	8	208431
	6	6_2X0	5171783	5298455	3	126672
	6	6_30X0	87345650	87388935	3	43285
	6	6_31X0	89714203	90270053	16	555850
	6	6_32X0	113458895	113483140	2	24245
	6	6_33X0	116493611	116584195	5	90584
	6	6_34X0	117841242	117872765	2	31523
	6	6_35X0	118428687	118479320	2	50633
	6	6_36X0	121300227	121365372	2	65145
	6	6_37X0	122553801	122600127	2	46326
	6	6_38X0	122921848	123289870	8	368022
	6	6_39X0	124493113	124636085	5	142972
	6	6_3X0	6863334	6882085	2	18751
	6	6_40X0	124864692	124911705	2	47013
	6	6_41X0	126532551	126740674	3	208123
	6	6_42X0	127015586	127081408	2	65822
	6	6_43X0	128375503	128408932	2	33429
	6	6_44X0	128483567	128581608	2	98041
	6	6_45X0	129268259	129775849	18	507590
	6	6_45X1	129831164	130386874	10	555710
	6	6_45X2	128706442	129186534	7	480092
	6	6_46X0	131486400	131553757	2	67357
	6	6_47X0	131629823	131690063	6	60240
	6	6_48X0	132206795	132600702	5	393907
	6	6_49X0	132656953	133055816	18	398863
	6	6_4X0	23648874	23655136	2	6262
	6	6_50X0	135065651	135118283	2	52632
	6	6_51X0	136830421	136840507	2	10086
	6	6_5X0	24733245	24891805	4	158560
	6	6_6X0	29281141	29335854	2	54713

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	6	6_7X0	30541582	30645363	4	103781
	6	6_8X0	34310605	34423142	4	112537
	6	6_9X0	38627261	38680864	2	53603
	7	7_100X0	104552884	104734362	11	181478
	7	7_101X0	104768160	104862667	3	94507
	7	7_102X0	104880709	105470898	26	590189
	7	7_103X0	106677534	107148329	26	470795
	7	7_104X0	107814580	108212527	18	397947
	7	7_104X1	108255192	108565793	12	310601
	7	7_104X2	108612280	108894420	12	282140
	7	7_105X0	112349394	112413102	2	63708
	7	7_106X0	114634372	114723063	2	88691
	7	7_107X0	119442542	119479255	2	36713
	7	7_108X0	119526341	119784889	4	258548
	7	7_109X0	122159439	122186204	2	26765
	7	7_10X0	6676923	6730421	2	53498
	7	7_110X0	125552120	125633570	2	81450
	7	7_111X0	126760749	126884703	2	123954
	7	7_112X0	127015051	127017352	2	2301
	7	7_113X0	127233255	127615272	9	382017
	7	7_114X0	127712944	127732166	2	19222
	7	7_115X0	127807372	127849000	2	41628
	7	7_116X0	127865993	127885992	2	19999
	7	7_117X0	127886093	127946725	3	60632
	7	7_118X0	128062683	128205501	4	142818
	7	7_119X0	131326099	131336716	2	10617
	7	7_11X0	7114690	7209998	4	95308
	7	7_120X0	131542867	131558014	2	15147
	7	7_121X0	140159999	140705267	19	545268
	7	7_122X0	140818087	140857577	2	39490
	7	7_123X0	140948958	141016908	6	67950
	7	7_124X0	141468822	141493421	2	24599
	7	7_125X0	141633972	141873084	3	239112
	7	7_126X0	142211859	142304594	5	92735
	7	7_127X0	142371148	142387857	2	16709
	7	7_128X0	142443171	142516003	2	72832
	7	7_129X0	142650769	142679533	2	28764
	7	7_12X0	7231327	7566786	5	335459
	7	7_130X0	143764906	143784500	2	19594
	7	7_131X0	144838595	144899881	3	61286
	7	7_132X0	145300828	145324086	2	23258
	7	7_13X0	7670887	7902462	3	231575
	7	7_14X0	8136950	8489075	5	352125
	7	7_15X0	8906772	9223653	4	316881
	7	7_16X0	9754636	10176286	5	421650
	7	7_17X0	10273828	10754364	6	480536
	7	7_18X0	11306378	11614112	4	307734
	7	7_18X1	10900740	11166159	4	265419
	7	7_19X0	12041299	12369273	8	327974
	7	7_19X1	11669599	11931285	5	261686
	7	7_1X0	3332955	3435163	3	102208
	7	7_20X0	12415153	12494338	2	79185
	7	7_21X0	12581470	12733105	4	151635
	7	7_22X0	12768092	13053684	9	285592
	7	7_23X0	13084289	13220289	5	136000

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	7	7_24X0	13450856	13571067	2	120211
	7	7_25X0	15750381	15904020	4	153639
	7	7_26X0	16234726	16259540	2	24814
	7	7_27X0	16671331	16725214	2	53883
	7	7_28X0	17087934	17215760	3	127826
	7	7_29X0	17713238	18070427	5	357189
	7	7_2X0	3489071	4063204	8	574133
	7	7_30X0	18348210	18821607	11	473397
	7	7_31X0	19076258	19206479	2	130221
	7	7_32X0	19671579	19681410	2	9831
	7	7_33X0	19724161	19991104	6	266943
	7	7_34X0	20509617	20912347	15	402730
	7	7_34X1	20101158	20474925	12	373767
	7	7_34X2	21011532	21361291	11	349759
	7	7_34X3	21511780	21844267	7	332487
	7	7_35X0	22370122	22714910	12	344788
	7	7_35X1	21909807	22329717	12	419910
	7	7_35X2	22761697	22984673	5	222976
	7	7_35X3	23019006	23275243	7	256237
	7	7_36X0	23301192	23441922	2	140730
	7	7_37X0	23504642	24055691	10	551049
	7	7_38X0	24081924	24316666	10	234742
	7	7_39X0	24349196	24760311	10	411115
	7	7_3X0	4237754	4397728	3	159974
	7	7_40X0	25077284	25227507	2	150223
	7	7_41X0	25464460	25539819	2	75359
	7	7_42X0	26061497	26315088	3	253591
	7	7_42X1	25802475	25926559	2	124084
	7	7_43X0	26364691	26574144	3	209453
	7	7_44X0	26665227	26939380	5	274153
	7	7_44X1	27029090	27158564	3	129474
	7	7_45X0	27527059	27674598	3	147539
	7	7_46X0	27692502	28008416	8	315914
	7	7_47X0	28071236	28164811	2	93575
	7	7_48X0	28267881	28279489	2	11608
	7	7_49X0	28508836	28524322	2	15486
	7	7_4X0	4845162	4896973	2	51811
	7	7_50X0	28543596	28737482	3	193886
	7	7_51X0	28833854	28866284	2	32430
	7	7_52X0	29768552	30235725	16	467173
	7	7_53X0	30291753	30487223	7	195470
	7	7_54X0	30670722	30700380	2	29658
	7	7_55X0	30818357	30861564	3	43207
	7	7_56X0	30866584	30914327	2	47743
	7	7_57X0	30922389	30944032	2	21643
	7	7_58X0	31042513	31074952	3	32439
	7	7_59X0	31290519	31376916	4	86397
	7	7_5X0	4915217	5053335	10	138118
	7	7_60X0	32209173	32327245	2	118072
	7	7_61X0	32334182	32442722	2	108540
	7	7_62X0	33063449	33374559	4	311110
	7	7_63X0	34021567	34100837	2	79270
	7	7_64X0	35050315	35121928	2	71613
	7	7_65X0	40987555	41045302	2	57747
	7	7_66X0	41373790	41650596	5	276806

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	7	7_67X0	42473386	42867234	6	393848
	7	7_68X0	43187170	43313255	2	126085
	7	7_69X0	43351548	43659353	12	307805
	7	7_6X0	5146218	5454838	5	308620
	7	7_70X0	43690418	44229543	26	539125
	7	7_71X0	45125922	45136231	3	10309
	7	7_72X0	45259031	45288993	2	29962
	7	7_73X0	45465984	45488805	2	22821
	7	7_74X0	45617606	45694402	3	76796
	7	7_75X0	46847082	46958527	4	111445
	7	7_76X0	47185717	47601372	6	415655
	7	7_77X0	47980837	48199288	4	218451
	7	7_78X0	48388526	48643811	4	255285
	7	7_79X0	55935856	56019954	2	84098
	7	7_7X0	5544172	6011010	6	466838
	7	7_80X0	62346569	62381640	2	35071
	7	7_81X0	78902579	78920187	2	17608
	7	7_82X0	79792241	79813439	2	21198
	7	7_83X0	84853571	84920861	3	67290
	7	7_84X0	84940169	85237704	4	297535
	7	7_85X0	85558703	85754981	3	196278
	7	7_86X0	86268730	86697507	15	428777
	7	7_87X0	86795141	87037968	3	242827
	7	7_88X0	90450700	90476001	2	25301
	7	7_89X0	98350668	98501831	2	151163
	7	7_8X0	6131491	6396915	7	265424
	7	7_90X0	99153658	99238619	2	84961
	7	7_91X0	99735145	99768130	2	32985
	7	7_92X0	100472993	100498256	2	25263
	7	7_93X0	100937630	101011974	2	74344
	7	7_94X0	101410886	101512819	2	101933
	7	7_95X0	101899803	101920774	2	20971
	7	7_96X0	102065778	102096396	2	30618
	7	7_97X0	102097114	102111156	2	14042
	7	7_98X0	103491915	103791821	22	299906
	7	7_98X1	103812524	104190205	26	377681
	7	7_98X2	102476773	102857271	24	380498
	7	7_98X3	102877122	103135242	17	258120
	7	7_98X4	103143590	103471358	17	327768
	7	7_99X0	104225507	104483426	9	257919
	7	7_9X0	6439836	6570919	8	131083
	8	8_10X0	40675077	40827037	4	151960
	8	8_11X0	43129807	43193881	2	64074
	8	8_12X0	43519762	43699487	4	179725
	8	8_13X0	45241174	45294859	2	53685
	8	8_14X0	47822143	47825476	2	3333
	8	8_15X0	64980884	65018584	3	37700
	8	8_16X0	66680721	66706792	2	26071
	8	8_17X0	67490758	67548627	3	57869
	8	8_18X0	69427980	69522976	2	94996
	8	8_19X0	69611414	69774911	6	163497
	8	8_1X0	3682824	4137707	8	454883
	8	8_20X0	70083457	70120873	2	37416
	8	8_21X0	71882019	71971716	4	89697
	8	8_22X0	72057660	72110538	3	52878

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	8	8_23X0	83541093	83741173	3	200080
	8	8_24X0	85060055	85080746	2	20691
	8	8_25X0	92855395	92918823	2	63428
	8	8_26X0	93020214	93535707	9	515493
	8	8_27X0	94137197	94180327	3	43130
	8	8_28X0	94745590	94812036	3	66446
	8	8_29X0	94923694	95045250	2	121556
	8	8_2X0	13026034	13076026	3	49992
	8	8_30X0	104250944	104395807	4	144863
	8	8_31X0	104734018	105150109	10	416091
	8	8_32X0	105264648	105326259	2	61611
	8	8_33X0	105973513	105996394	3	22881
	8	8_34X0	110974996	111031751	2	56755
	8	8_35X0	117663333	117759027	2	95694
	8	8_36X0	120608786	120668079	3	59293
	8	8_37X0	121084386	121130644	3	46258
	8	8_3X0	18974940	19497775	17	522835
	8	8_4X0	20892669	21948850	39	1056181
	8	8_5X0	22073637	22162494	2	88857
	8	8_6X0	24532454	24597009	2	64555
	8	8_7X0	24602246	25016922	5	414676
	8	8_8X0	26119534	26160819	2	41285
	8	8_9X0	27368782	27413944	2	45162
	9	9_10X0	30899155	30962856	2	63701
	9	9_11X0	32315920	32399197	2	83277
	9	9_12X0	35452076	35518269	2	66193
	9	9_13X0	35607093	35790112	5	183019
	9	9_14X0	36431884	36541963	3	110079
	9	9_15X0	37402046	37433246	2	31200
	9	9_16X0	37711348	38607417	37	896069
	9	9_16X1	39613544	40188094	26	574550
	9	9_16X2	38657458	39047549	16	390091
	9	9_16X3	39059729	39573314	16	513585
	9	9_17X0	44045672	44142724	4	97052
	9	9_18X0	44973620	45009613	3	35993
	9	9_19X0	45040177	45154152	5	113975
	9	9_1X0	3000922	3038316	9	37394
	9	9_20X0	45370185	45410278	2	40093
	9	9_21X0	46228721	46271919	3	43198
	9	9_22X0	48899214	48964990	2	65776
	9	9_23X0	54981533	55048779	3	67246
	9	9_24X0	57443646	57580912	2	137266
	9	9_25X0	57589452	57653631	2	64179
	9	9_26X0	57676937	57703824	2	26887
	9	9_27X0	58157773	58204306	2	46533
	9	9_28X0	58524300	58651318	2	127018
	9	9_29X0	65101486	65185872	2	84386
	9	9_2X0	5298517	5358823	3	60306
	9	9_30X0	66779193	66834648	2	55455
	9	9_31X0	71479325	71592265	2	112940
	9	9_32X0	73102470	73122382	2	19912
	9	9_33X0	78191947	78347168	5	155221
	9	9_34X0	78367052	78378725	2	11673
	9	9_35X0	89199209	89240240	3	41031
	9	9_36X0	91360500	91389348	2	28848

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	9	9_37X0	92192936	92492456	5	299520
	9	9_38X0	98422961	98509771	2	86810
	9	9_39X0	98842503	98904622	4	62119
	9	9_3X0	7272514	7699585	10	427071
	9	9_40X0	103111807	103266225	5	154418
	9	9_41X0	105130207	105145076	2	14869
	9	9_42X0	106440051	106452920	2	12869
	9	9_43X0	106499992	106627675	4	127683
	9	9_44X0	107569117	107586661	3	17544
	9	9_45X0	107740500	107872829	2	132329
	9	9_46X0	107906873	107949683	2	42810
	9	9_47X0	109142363	109790670	12	648307
	9	9_48X0	110798166	110864628	6	66462
	9	9_49X0	114731173	114844152	2	112979
	9	9_4X0	18443456	18485239	2	41783
	9	9_50X0	121760033	121783818	2	23785
	9	9_51X0	122851914	122931028	3	79114
	9	9_52X0	123590800	123678782	2	87982
	9	9_53X0	123774540	124124731	8	350191
	9	9_5X0	19248357	19497688	8	249331
	9	9_5X1	19883309	20336094	12	452785
	9	9_5X2	19537011	19849747	7	312736
	9	9_5X3	18754553	19177312	10	422759
	9	9_6X0	20345136	20594471	5	249335
	9	9_7X0	21015985	21039036	3	23051
	9	9_8X0	22055700	22307638	5	251938
	9	9_9X0	26765203	26923081	3	157878
	JH584299.1	JH584299.1_1X0	4975	176134	3	171159
	JH584299.1	JH584299.1_2X0	477246	625594	3	148348
	JH584299.1	JH584299.1_3X0	716296	924675	5	208379
	X	X_10X0	21115820	21175653	2	59833
	X	X_11X0	23892784	23958934	2	66150
	X	X_12X0	31117674	31383918	3	266244
	X	X_13X0	36743659	36798807	2	55148
	X	X_14X0	37213804	38129967	33	916163
	X	X_15X0	53243415	53331185	2	87770
	X	X_16X0	53742901	53797696	2	54795
	X	X_17X0	56136572	56212881	2	76309
	X	X_18X0	67682900	67706350	2	23450
	X	X_19X0	73123068	73149450	2	26382
	X	X_1X0	4800496	4954077	2	153581
	X	X_20X0	73342621	73428385	2	85764
	X	X_21X0	73657276	73778925	2	121649
	X	X_22X0	74569153	74595275	2	26122
	X	X_23X0	77837902	78245921	7	408019
	X	X_24X0	101261376	101269023	2	7647
	X	X_25X0	101791225	101798642	2	7417
	X	X_26X0	107089234	107296769	3	207535
	X	X_27X0	114905012	114908510	2	3498
	X	X_28X0	134717963	134809221	4	91258
	X	X_29X0	135167624	135615805	6	448181
	X	X_2X0	6047453	6173015	2	125562
	X	X_30X0	135742733	135891078	4	148345
	X	X_31X0	135993824	136034946	2	41122
	X	X_32X0	136066564	136271978	6	205414

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	X	X_33X0	136448107	136534061	2	85954
	X	X_34X0	136666611	136711478	2	44867
	X	X_35X0	136954988	136976869	2	21881
	X	X_36X0	154001590	154484682	6	483092
	X	X_37X0	154924012	155089793	6	165781
	X	X_38X0	157565947	157598650	2	32703
	X	X_39X0	164438058	164539752	2	101694
	X	X_3X0	7722245	7731064	2	8819
	X	X_40X0	167242732	167330497	2	87765
	X	X_4X0	8327424	8880750	12	553326
	X	X_5X0	8962134	9063861	5	101727
	X	X_6X0	11299257	11328151	10	28894
	X	X_7X0	13632769	13667433	2	34664
	X	X_8X0	20688418	20860208	2	171790
	X	X_9X0	21029589	21089229	2	59640
	Y	Y_1X0	2599099	2720674	2	121575
	Y	Y_2X0	2785165	3782910	10	997745
	Y	Y_3X0	21166083	21243821	2	77738
	10	10_1X0	913475	1498277	23	584802
	10	10_1X1	62169852	62839822	18	669970
	10	10_1X10	72118250	72547182	10	428932
	10	10_1X11	61751958	62118580	8	366622
	10	10_1X12	71460398	71897158	11	436760
	10	10_1X13	4474143	5124522	10	650379
	10	10_1X14	41231916	41628538	9	396622
	10	10_1X15	2727789	3135086	9	407297
	10	10_1X16	72893026	73196547	7	303521
	10	10_1X17	42464410	42661830	5	197420
	10	10_1X18	2385911	2713256	7	327345
	10	10_1X19	72555375	72863437	6	308062
	10	10_1X2	8400	336651	11	328251
	10	10_1X20	58681234	59147976	10	466742
	10	10_1X21	40781847	41199338	8	417491
	10	10_1X22	53625664	53850465	3	224801
	10	10_1X23	54486054	54909797	6	423743
	10	10_1X24	53885721	54068930	3	183209
	10	10_1X25	53241029	53602525	5	361496
	10	10_1X26	54160702	54410068	3	249366
	10	10_1X27	73362679	73464050	4	101371
	10	10_1X28	42004370	42307703	5	303333
	10	10_1X29	69406813	69437776	4	30963
	10	10_1X3	3880060	4454617	13	574557
	10	10_1X30	66453236	66516237	3	63001
	10	10_1X31	59421459	59594848	3	173389
	10	10_1X32	30054469	30274143	3	219674
	10	10_1X33	50880865	51321181	4	440316
	10	10_1X34	51608730	51826800	3	218070
	10	10_1X35	14575697	14724980	3	149283
	10	10_1X36	52457774	52884684	4	426910
	10	10_1X37	55352589	55397447	2	44858
	10	10_1X38	63972390	64051649	2	79259
	10	10_1X39	37341643	37396049	2	54406
	10	10_1X4	640887	901251	8	260364
	10	10_1X40	24740798	24752983	2	12185
	10	10_1X41	70638921	70674184	2	35263

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	10	10_1X5	354561	574220	6	219659
	10	10_1X6	71121991	71355244	11	233253
	10	10_1X7	3581887	3852197	9	270310
	10	10_1X8	64553764	65103687	13	549923
	10	10_1X9	1797912	2237158	10	439246
	10	10_2X0	83118205	83134480	2	16275
	10	10_3X0	86783165	86881675	3	98510
	10	10_4X0	87254907	87299067	2	44160
	10	10_5X0	95110285	95208038	2	97753
	10	10_6X0	106644953	106907524	6	262571
	11	11_10X0	56723117	56852641	2	129524
	11	11_11X0	57218030	57383213	4	165183
	11	11_12X0	57830032	57847177	2	17145
	11	11_13X0	58412602	58534244	8	121642
	11	11_14X0	59014693	59076935	2	62242
	11	11_15X0	59089019	59167372	2	78353
	11	11_16X0	59569546	60345621	27	776075
	11	11_17X0	70887284	70978891	2	91607
	11	11_18X0	71719877	71745892	2	26015
	11	11_19X0	73086553	73193262	2	106709
	11	11_1X0	668329	690727	2	22398
	11	11_20X0	73582593	73762317	3	179724
	11	11_21X0	75594264	75735268	4	141004
	11	11_22X0	76163960	76249899	2	85939
	11	11_23X0	78784485	78866979	2	82494
	11	11_24X0	83308502	83362475	2	53973
	11	11_25X0	84884040	85079547	2	195507
	11	11_26X0	94248249	94332766	3	84517
	11	11_27X0	108355804	108544068	2	188264
	11	11_28X0	108891357	108940191	3	48834
	11	11_29X0	111459385	111608624	4	149239
	11	11_2X0	21288610	21411947	2	123337
	11	11_30X0	121135030	121228908	2	93878
	11	11_3X0	25107488	25127829	2	20341
	11	11_4X0	26518520	26649198	3	130678
	11	11_5X0	28281407	28302775	2	21368
	11	11_6X0	39280658	39371584	2	90926
	11	11_7X0	46961781	47991816	21	1030035
	11	11_8X0	55322323	55347077	2	24754
	11	11_9X0	55826525	56009308	3	182783
	12	12_10X0	28631672	28787051	2	155379
	12	12_11X0	31223079	31497781	5	274702
	12	12_12X0	32868500	33001212	3	132712
	12	12_13X0	35576837	35740856	2	164019
	12	12_14X0	39342836	39444643	2	101807
	12	12_15X0	40872301	41017075	4	144774
	12	12_15X1	41183756	41491150	3	307394
	12	12_16X0	41555378	41605934	2	50556
	12	12_17X0	41619726	41918304	6	298578
	12	12_18X0	42079498	42208974	2	129476
	12	12_19X0	42247753	42313672	2	65919
	12	12_1X0	1602673	1697899	2	95226
	12	12_20X0	42336928	42653243	8	316315
	12	12_21X0	42897957	43255636	7	357679
	12	12_22X0	43355598	43547319	4	191721

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	12	12_23X0	43569271	43906063	8	336792
	12	12_24X0	43967075	44107969	5	140894
	12	12_25X0	44493127	44788024	5	294897
	12	12_26X0	44995585	45042857	2	47272
	12	12_27X0	46458751	46561356	2	102605
	12	12_28X0	46932054	46967803	4	35749
	12	12_29X0	47041594	47105350	3	63756
	12	12_2X0	4079298	4083993	2	4695
	12	12_30X0	47123545	47153049	2	29504
	12	12_31X0	48823081	48877584	2	54503
	12	12_32X0	49188812	49528969	4	340157
	12	12_33X0	50241598	50312078	2	70480
	12	12_34X0	50759532	50895948	2	136416
	12	12_35X0	51152878	51187571	2	34693
	12	12_36X0	51411537	51495610	2	84073
	12	12_37X0	52312389	52400320	6	87931
	12	12_38X0	53182101	53329767	2	147666
	12	12_39X0	56557666	56656413	2	98747
	12	12_3X0	4307597	4437607	2	130010
	12	12_40X0	58026633	58254491	5	227858
	12	12_41X0	60654157	60752697	2	98540
	12	12_42X0	60916017	61175034	3	259017
	12	12_43X0	66342639	66491983	4	149344
	12	12_44X0	67046996	67205271	3	158275
	12	12_45X0	68004773	68089045	3	84272
	12	12_46X0	68135874	68203433	2	67559
	12	12_47X0	73383588	73433392	2	49804
	12	12_48X0	87167567	87316807	2	149240
	12	12_49X0	87415109	87570509	4	155400
	12	12_4X0	5126887	5386589	3	259702
	12	12_50X0	89975572	90132785	3	157213
	12	12_51X0	91516011	91643725	2	127714
	12	12_5X0	9717158	9819716	3	102558
	12	12_6X0	11396437	11407913	2	11476
	12	12_7X0	12492988	12689401	2	196413
	12	12_8X0	13457985	13477202	2	19217
	12	12_9X0	22366988	22448287	2	81299
	13	13_10X0	35019448	35454466	5	435018
	13	13_11X0	35830898	35968422	2	137524
	13	13_12X0	36638415	36653605	2	15190
	13	13_13X0	43053300	43356758	6	303458
	13	13_14X0	43704709	43745161	2	40452
	13	13_15X0	45303807	45425507	2	121700
	13	13_16X0	67612992	67671215	2	58223
	13	13_17X0	75134934	75260886	2	125952
	13	13_18X0	81406459	81530205	4	123746
	13	13_19X0	85646138	85799232	3	153094
	13	13_1X0	174434	197684	2	23250
	13	13_20X0	93633742	93673578	2	39836
	13	13_21X0	94154792	94178239	3	23447
	13	13_22X0	94278798	94342719	2	63921
	13	13_23X0	94593417	94636301	2	42884
	13	13_24X0	95195007	95287240	3	92233
	13	13_25X0	100195934	100632908	8	436974
	13	13_26X0	101017618	101060119	3	42501

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	13	13_27X0	101103411	101301700	5	198289
	13	13_28X0	101347629	101366091	2	18462
	13	13_29X0	101589476	101791461	4	201985
	13	13_2X0	727548	744891	2	17343
	13	13_30X0	102537981	102674887	2	136906
	13	13_3X0	18864574	18896579	2	32005
	13	13_4X0	18966169	19274995	8	308826
	13	13_5X0	19533321	19636023	4	102702
	13	13_6X0	26454472	26513629	2	59157
	13	13_7X0	31472897	31788439	5	315542
	13	13_8X0	33353491	33761260	10	407769
	13	13_9X0	33897556	34030776	2	133220
	14	14_10X0	48365277	48384031	2	18754
	14	14_11X0	50072633	50126503	2	53870
	14	14_12X0	54850610	54930406	2	79796
	14	14_13X0	57817266	58047153	3	229887
	14	14_14X0	59852860	59874101	2	21241
	14	14_15X0	64215308	64252706	2	37398
	14	14_16X0	70599613	70641650	2	42037
	14	14_17X0	72743195	72827210	2	84015
	14	14_18X0	74691688	74998074	7	306386
	14	14_19X0	75343199	75463355	2	120156
	14	14_1X0	617382	656827	2	39445
	14	14_20X0	77482367	77601682	2	119315
	14	14_21X0	83135123	83592567	5	457444
	14	14_22X0	88697221	88716120	2	18899
	14	14_23X0	92002237	92080898	2	78661
	14	14_24X0	92816849	92846004	2	29155
	14	14_25X0	93808406	93835510	2	27104
	14	14_26X0	94458024	94496515	2	38491
	14	14_2X0	1107468	1139354	2	31886
	14	14_3X0	8341456	8413092	2	71636
	14	14_4X0	8799307	8886636	3	87329
	14	14_5X0	11463892	11531710	2	67818
	14	14_6X0	37623326	37659182	2	35856
	14	14_7X0	42313573	42534789	4	221216
	14	14_8X0	47027228	47092810	3	65582
	14	14_9X0	48286181	48357571	2	71390
	15	15_10X0	18388807	18443635	3	54828
	15	15_11X0	20620251	20780752	3	160501
	15	15_12X0	21355988	21439807	2	83819
	15	15_13X0	23548092	23553576	2	5484
	15	15_14X0	30625470	30692762	3	67292
	15	15_15X0	32673606	33102341	9	428735
	15	15_16X0	33259244	33314780	2	55536
	15	15_17X0	40683741	40733154	3	49413
	15	15_18X0	46111233	46217269	3	106036
	15	15_19X0	47303384	47351372	2	47988
	15	15_1X0	578967	686313	2	107346
	15	15_20X0	51543061	51573971	2	30910
	15	15_21X0	60765232	60908325	2	143093
	15	15_22X0	61796442	61886952	2	90510
	15	15_23X0	63570441	63597885	2	27444
	15	15_24X0	64393810	64450620	2	56810
	15	15_25X0	65537044	65606889	3	69845

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	15	15_26X0	65848113	65867749	2	19636
	15	15_27X0	66376899	67958751	42	1581852
	15	15_27X1	68005973	68737495	7	731522
	15	15_28X0	68789464	68874974	2	85510
	15	15_29X0	69648546	69714407	2	65861
	15	15_2X0	4660641	4732811	2	72170
	15	15_30X0	78384085	78387169	2	3084
	15	15_31X0	80038427	80129703	3	91276
	15	15_32X0	80269995	80373461	3	103466
	15	15_33X0	80373517	80487099	2	113582
	15	15_34X0	80790725	80852108	2	61383
	15	15_35X0	105797123	105835423	2	38300
	15	15_36X0	110213286	110353332	4	140046
	15	15_3X0	6243981	6323868	2	79887
	15	15_4X0	9261425	9353872	2	92447
	15	15_5X0	9511223	9658571	2	147348
	15	15_6X0	10414247	10589848	5	175601
	15	15_7X0	10817960	11393372	7	575412
	15	15_8X0	16903407	17035624	5	132217
	15	15_9X0	17301128	17361726	2	60598
	16	16_1X0	22861219	22897108	6	35889
	16	16_2X0	23011524	23218873	3	207349
	16	16_3X0	26902775	26993588	2	90813
	16	16_4X0	27064351	27319802	4	255451
	16	16_5X0	46912729	47079527	2	166798
	16	16_6X0	97871511	98332464	12	460953
	16	16_7X0	98366562	98789722	4	423160
	16	16_8X0	98809585	98935572	2	125987
	16	16_9X0	99795134	100024601	5	229467
	17	17_10X0	35139936	35207872	3	67936
	17	17_11X0	38260280	38291855	2	31575
	17	17_12X0	40005686	40067527	2	61841
	17	17_13X0	40777986	40825779	2	47793
	17	17_14X0	41437259	41485758	2	48499
	17	17_15X0	50489628	50868577	5	378949
	17	17_15X1	49706267	50108982	4	402715
	17	17_15X2	50142909	50347737	2	204828
	17	17_16X0	62949776	62981437	2	31661
	17	17_17X0	81425255	81574570	8	149315
	17	17_18X0	86146981	86202556	3	55575
	17	17_19X0	90207634	90878346	19	670712
	17	17_19X1	89594452	90061797	8	467345
	17	17_19X2	86841341	87224288	7	382947
	17	17_19X3	88157475	88359805	5	202330
	17	17_19X4	88983961	89493409	6	509448
	17	17_19X5	87611793	87852061	5	240268
	17	17_19X6	88549063	88706044	2	156981
	17	17_1X0	2819223	2954681	3	135458
	17	17_20X0	90889170	91204475	6	315305
	17	17_21X0	92349907	92483898	4	133991
	17	17_22X0	94353883	94439719	2	85836
	17	17_23X0	95918068	96252625	3	334557
	17	17_24X0	96854341	96903850	2	49509
	17	17_25X0	96991385	97026082	2	34697
	17	17_26X0	97059853	97198918	5	139065

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	17	17_27X0	99228928	99405561	5	176633
	17	17_2X0	4747066	4881637	6	134571
	17	17_3X0	5713721	5757218	2	43497
	17	17_4X0	6712475	6851826	2	139351
	17	17_5X0	27022354	27063633	2	41279
	17	17_6X0	32282252	32374511	2	92259
	17	17_7X0	34389339	34451586	2	62247
	17	17_8X0	34534105	34703481	5	169376
	17	17_9X0	34720531	34941805	4	221274
	18	18_10X0	67331643	67452737	2	121094
	18	18_11X0	81155343	81305612	2	150269
	18	18_12X0	89181463	89452640	3	271177
	18	18_13X0	101528667	101864925	9	336258
	18	18_14X0	102135494	102260365	4	124871
	18	18_15X0	102683983	102938119	4	254136
	18	18_16X0	104101956	104122084	2	20128
	18	18_17X0	104195505	104301319	4	105814
	18	18_18X0	104557978	104644119	2	86141
	18	18_19X0	104740654	104756301	3	15647
	18	18_1X0	18258188	18588698	3	330510
	18	18_2X0	20319329	20345055	2	25726
	18	18_3X0	24293072	24489453	4	196381
	18	18_4X0	32222119	32285747	2	63628
	18	18_5X0	37143836	37232224	3	88388
	18	18_6X0	37358185	37390914	2	32729
	18	18_7X0	41339313	41360106	2	20793
	18	18_8X0	45539978	45957170	5	417192
	18	18_9X0	66426570	66488845	2	62275
	19	19_10X0	33543256	33890732	12	347476
	19	19_11X0	34207479	34506393	3	298914
	19	19_12X0	34600329	34674934	2	74605
	19	19_13X0	36732919	36861999	5	129080
	19	19_14X0	39685546	39949190	4	263644
	19	19_15X0	45977936	46067434	2	89498
	19	19_16X0	50687879	50876228	2	188349
	19	19_17X0	52851529	53035853	2	184324
	19	19_18X0	65914256	65988127	2	73871
	19	19_19X0	70962683	71323944	6	361261
	19	19_1X0	24023909	24095337	2	71428
	19	19_20X0	72718406	72827196	2	108790
	19	19_21X0	73529198	73574809	2	45611
	19	19_22X0	74188430	74194620	2	6190
	19	19_23X0	74485764	74520538	2	34774
	19	19_24X0	74553486	74560926	2	7440
	19	19_25X0	74591165	74716614	2	125449
	19	19_26X0	76513580	76535447	2	21867
	19	19_27X0	76593192	76685146	2	91954
	19	19_28X0	77193817	77293869	2	100052
	19	19_29X0	77835071	77988253	2	153182
	19	19_2X0	27965449	28050539	2	85090
	19	19_30X0	78137129	78222113	2	84984
	19	19_31X0	79758264	79790114	2	31850
	19	19_3X0	29692561	29877464	2	184903
	19	19_4X0	30961477	31079726	2	118249
	19	19_5X0	31589080	31613219	2	24139

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	19	19_6X0	32810097	32823852	2	13755
	19	19_7X0	32848613	32955049	2	106436
	19	19_8X0	33436114	33494091	5	57977
	19	19_9X0	33527338	33533355	2	6017
	1a	1a_10X0	38471751	38550401	2	78650
	1a	1a_11X0	42737838	42890413	4	152575
	1a	1a_12X0	52799316	52851663	2	52347
	1a	1a_13X0	72383056	72397607	4	14551
	1a	1a_14X0	73054623	73143858	14	89235
	1a	1a_15X0	73171079	73239996	2	68917
	1a	1a_16X0	73272940	73809678	14	536738
	1a	1a_17X0	74633335	75284482	13	651147
	1a	1a_18X0	76139269	76301569	3	162300
	1a	1a_19X0	76371353	76382216	3	10863
	1a	1a_1X0	1568073	1794853	3	226780
	1a	1a_20X0	76485326	76541144	4	55818
	1a	1a_21X0	76880234	77017824	2	137590
	1a	1a_22X0	77230188	78146331	18	916143
	1a	1a_23X0	78147296	78297454	2	150158
	1a	1a_24X0	80669116	80686861	3	17745
	1a	1a_25X0	80726759	81050752	3	323993
	1a	1a_26X0	81545052	81639466	2	94414
	1a	1a_27X0	82210399	82416066	2	205667
	1a	1a_28X0	84197701	84254461	2	56760
	1a	1a_29X0	96710458	96750811	2	40353
	1a	1a_2X0	6041735	6237396	4	195661
	1a	1a_30X0	100086686	100144693	2	58007
	1a	1a_31X0	111105016	111175052	2	70036
	1a	1a_32X0	118477483	118490061	2	12578
	1a	1a_3X0	6332424	6453966	2	121542
	1a	1a_4X0	22455830	22793993	8	338163
	1a	1a_5X0	23235567	23276021	2	40454
	1a	1a_6X0	24941793	25041634	2	99841
	1a	1a_7X0	26433992	26443246	2	9254
	1a	1a_8X0	30591814	30617887	2	26073
	1a	1a_9X0	30673736	30868131	4	194395
	2	2_10X0	40632184	40656937	2	24753
	2	2_11X0	48921407	48999659	2	78252
	2	2_12X0	50096909	50139751	2	42842
	2	2_13X0	55124338	55143797	2	19459
	2	2_14X0	55932489	55988551	2	56062
	2	2_15X0	56159656	56285775	4	126119
	2	2_16X0	65762839	65898773	3	135934
	2	2_17X0	66023683	66081179	4	57496
	2	2_18X0	66774044	66892068	11	118024
	2	2_19X0	67607183	67760600	3	153417
	2	2_1X0	37341	197621	4	160280
	2	2_20X0	79747218	79887776	3	140558
	2	2_21X0	80000368	80067075	3	66707
	2	2_22X0	80070513	80988446	12	917933
	2	2_23X0	81429166	81503857	7	74691
	2	2_24X0	82838991	82954165	2	115174
	2	2_25X0	86651075	86686349	2	35274
	2	2_26X0	89173509	89485149	3	311640
	2	2_27X0	89544869	89587653	2	42784

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	2	2_28X0	90666936	90879387	3	212451
	2	2_29X0	93149574	93207420	2	57846
	2	2_2X0	1823257	2038049	5	214792
	2	2_30X0	111044825	111315344	3	270519
	2	2_31X0	146647485	146748809	2	101324
	2	2_32X0	148075501	148112983	2	37482
	2	2_33X0	148289692	148338836	3	49144
	2	2_3X0	4810121	5049452	2	239331
	2	2_4X0	23954470	24134240	3	179770
	2	2_5X0	29648422	29718196	2	69774
	2	2_6X0	30737414	30980027	3	242613
	2	2_7X0	32732288	33238406	5	506118
	2	2_8X0	39298668	39401495	2	102827
	2	2_9X0	39761233	40499166	16	737933
	20	20_1X0	997959	1027544	2	29585
	20	20_2X0	38715413	38788064	2	72651
	20	20_3X0	39516962	39687927	2	170965
	20	20_4X0	48739873	48793016	3	53143
	20	20_5X0	70858036	70987142	2	129106
	20	20_6X0	79350009	79410511	3	60502
	20	20_7X0	84875674	85078302	3	202628
	20	20_8X0	85128108	85343360	3	215252
	21	21_10X0	86054766	86080345	3	25579
	21	21_1X0	12184545	12225569	2	41024
	21	21_2X0	23150566	23503737	7	353171
	21	21_3X0	45758816	45893526	2	134710
	21	21_4X0	48111815	48264073	2	152258
	21	21_5X0	51801385	51872138	2	70753
	21	21_6X0	58041737	58070644	2	28907
	21	21_7X0	85215789	85316352	2	100563
	21	21_8X0	85490189	85519626	2	29437
	21	21_9X0	85832980	85881469	2	48489
	22a	22a_10X0	40194446	40204741	2	10295
	22a	22a_11X0	40883645	40934123	2	50478
	22a	22a_12X0	42107955	42185784	2	77829
	22a	22a_13X0	42597694	42603713	2	6019
	22a	22a_14X0	43374554	43555178	4	180624
	22a	22a_15X0	43865632	43959520	5	93888
	22a	22a_16X0	44738857	44834282	2	95425
	22a	22a_17X0	51620046	51755329	4	135283
	22a	22a_18X0	52266388	52324118	2	57730
	22a	22a_19X0	53779709	53920090	2	140381
	22a	22a_1X0	52792	306499	5	253707
	22a	22a_20X0	54561773	54578554	2	16781
	22a	22a_21X0	54633310	54805272	3	171962
	22a	22a_22X0	56325628	56385360	2	59732
	22a	22a_23X0	61947596	61964444	2	16848
	22a	22a_24X0	66088249	66205267	9	117018
	22a	22a_25X0	67651362	67719971	2	68609
	22a	22a_26X0	68634545	68834665	2	200120
	22a	22a_27X0	79049884	79161814	2	111930
	22a	22a_28X0	91361645	91508508	3	146863
	22a	22a_29X0	91653743	91818906	2	165163
	22a	22a_2X0	2865380	2904254	2	38874
	22a	22a_30X0	94055963	94314798	3	258835

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	22a	22a_31X0	107431809	107517720	2	85911
	22a	22a_32X0	109803591	109831619	2	28028
	22a	22a_33X0	110181682	110202939	2	21257
	22a	22a_34X0	110352307	110487329	3	135022
	22a	22a_35X0	122129575	122336551	2	206976
	22a	22a_36X0	123587713	123609766	2	22053
	22a	22a_37X0	123675651	123696671	2	21020
	22a	22a_3X0	9865308	9895998	2	30690
	22a	22a_4X0	10411795	10440920	2	29125
	22a	22a_5X0	11502328	11881739	9	379411
	22a	22a_6X0	30798536	30913946	2	115410
	22a	22a_7X0	32770121	32818054	3	47933
	22a	22a_8X0	39759342	39861685	4	102343
	22a	22a_9X0	39902910	39942525	2	39615
	23	23_10X0	27797002	27981570	4	184568
	23	23_11X0	30093245	30174503	2	81258
	23	23_12X0	34303566	34374476	2	70910
	23	23_1X0	14101343	14280551	2	179208
	23	23_2X0	20638465	20651732	2	13267
	23	23_3X0	22252564	22386957	2	134393
	23	23_4X0	22527017	22593901	2	66884
	23	23_5X0	24015473	24456926	11	441453
	23	23_6X0	25026436	25489301	10	462865
	23	23_6X1	24655771	24974590	5	318819
	23	23_7X0	26474377	26528193	3	53816
	23	23_8X0	26715563	26933559	3	217996
	23	23_9X0	27500100	27545935	3	45835
	24	24_10X0	19106850	19289319	5	182469
	24	24_11X0	19845895	19909959	2	64064
	24	24_12X0	21788731	21819182	3	30451
	24	24_13X0	23112273	23176515	2	64242
	24	24_1X0	6588075	6611196	2	23121
	24	24_2X0	8953988	9018849	2	64861
	24	24_3X0	9796058	9895120	2	99062
	24	24_4X0	11702746	11730902	3	28156
	24	24_5X0	12687706	12731972	2	44266
	24	24_6X0	12798334	12963736	10	165402
	24	24_7X0	14988097	15045865	3	57768
	24	24_8X0	15582041	15632976	4	50935
	24	24_9X0	15861281	16053603	4	192322
	25	25_10X0	27772725	27857327	2	84602
	25	25_11X0	28133385	28336757	2	203372
	25	25_12X0	28617942	28651459	3	33517
	25	25_13X0	29774697	29919720	2	145023
	25	25_14X0	30631095	30764661	9	133566
	25	25_15X0	32017409	32142266	2	124857
	25	25_1X0	8432930	8576817	2	143887
	25	25_2X0	13417344	13552654	2	135310
	25	25_3X0	16781472	16834935	2	53463
	25	25_4X0	16935644	17037134	2	101490
	25	25_5X0	17167508	17182751	2	15243
	25	25_6X0	17185284	17206491	2	21207
	25	25_7X0	19783680	19964045	3	180365
	25	25_8X0	20942666	21022594	2	79928
	25	25_9X0	22576748	22643416	2	66668

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	3	3_10X0	45645347	45890422	5	245075
	3	3_11X0	57894850	58045386	6	150536
	3	3_12X0	75882055	76019251	2	137196
	3	3_13X0	92344866	92431184	2	86318
	3	3_14X0	98298019	98482378	2	184359
	3	3_15X0	103479930	103507219	2	27289
	3	3_16X0	103699016	103791555	3	92539
	3	3_17X0	119904274	119911008	2	6734
	3	3_18X0	124398647	124621526	3	222879
	3	3_19X0	137471696	137598499	5	126803
	3	3_1X0	10499807	10513801	2	13994
	3	3_2X0	10700750	10813089	4	112339
	3	3_3X0	17197213	17422530	4	225317
	3	3_4X0	29973085	30018454	2	45369
	3	3_5X0	30805813	30838173	3	32360
	3	3_6X0	33109486	33211435	2	101949
	3	3_7X0	39338244	39543612	3	205368
	3	3_8X0	41289773	41308256	2	18483
	3	3_9X0	45034111	45118703	4	84592
	4	4_10X0	77778429	77890405	2	111976
	4	4_11X0	78604483	78935368	5	330885
	4	4_12X0	79216041	79413637	4	197596
	4	4_13X0	79531606	79603292	2	71686
	4	4_14X0	79669223	80029248	5	360025
	4	4_15X0	80153446	80193664	2	40218
	4	4_16X0	80364655	81194121	11	829466
	4	4_17X0	81376936	81489800	2	112864
	4	4_18X0	82140586	82149009	2	8423
	4	4_19X0	82158871	82323828	3	164957
	4	4_1X0	5702044	5781953	2	79909
	4	4_20X0	82624404	82842766	5	218362
	4	4_21X0	83441918	83699168	4	257250
	4	4_22X0	83865266	84005622	4	140356
	4	4_23X0	84184256	84248641	2	64385
	4	4_24X0	84994458	85036140	2	41682
	4	4_25X0	85232142	85285187	2	53045
	4	4_26X0	85965862	86056690	2	90828
	4	4_27X0	87016250	87076540	3	60290
	4	4_28X0	87686874	87861118	2	174244
	4	4_29X0	87883620	87953330	2	69710
	4	4_2X0	16144769	16590154	9	445385
	4	4_30X0	88080661	88158763	2	78102
	4	4_31X0	89930171	90065981	2	135810
	4	4_32X0	91831048	91842637	2	11589
	4	4_33X0	92603393	92696707	2	93314
	4	4_34X0	92841869	92872237	2	30368
	4	4_35X0	93073414	93206299	2	132885
	4	4_36X0	100039867	100101097	3	61230
	4	4_37X0	100611757	100959783	3	348026
	4	4_38X0	100986279	101104130	5	117851
	4	4_39X0	102405160	102429334	3	24174
	4	4_3X0	33201220	33222778	5	21558
	4	4_40X0	102443525	102564273	4	120748
	4	4_41X0	102606387	102769417	2	163030
	4	4_42X0	105911319	105941853	3	30534

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	4	4_43X0	106271171	106540501	4	269330
	4	4_44X0	107738844	108009958	8	271114
	4	4_45X0	111909946	111965941	3	55995
	4	4_46X0	114139266	114189506	2	50240
	4	4_47X0	120095753	120368280	3	272527
	4	4_48X0	140883874	141308686	10	424812
	4	4_49X0	142709161	142751009	2	41848
	4	4_4X0	34486205	34610174	2	123969
	4	4_50X0	145222657	145334107	5	111450
	4	4_5X0	45202037	45324269	5	122232
	4	4_6X0	49078307	49621688	7	543381
	4	4_7X0	63365375	63391793	2	26418
	4	4_8X0	76252897	77597727	23	1344830
	4	4_9X0	77746515	77760688	2	14173
	5	5_10X0	54799940	54867415	3	67475
	5	5_11X0	54902594	55000332	3	97738
	5	5_12X0	55038292	55296282	3	257990
	5	5_13X0	56043886	56149072	2	105186
	5	5_14X0	56195111	56580915	3	385804
	5	5_15X0	56664067	56794217	2	130150
	5	5_16X0	64804711	64900354	3	95643
	5	5_17X0	68737791	68803441	2	65650
	5	5_18X0	79760393	79856195	2	95802
	5	5_19X0	84666289	84678162	2	11873
	5	5_1X0	209961	734940	10	524979
	5	5_20X0	127910337	127931035	2	20698
	5	5_21X0	141255655	141322389	3	66734
	5	5_22X0	142460090	142498165	3	38075
	5	5_2X0	1029486	1335684	5	306198
	5	5_3X0	20043631	20050414	2	6783
	5	5_4X0	20090883	20112977	2	22094
	5	5_5X0	20443057	20560667	2	117610
	5	5_6X0	22468483	22525514	2	57031
	5	5_7X0	24142955	24294199	2	151244
	5	5_8X0	27037999	27103542	2	65543
	5	5_9X0	54191192	54689868	5	498676
	6	6_10X0	76273326	76358520	2	85194
	6	6_11X0	84727436	84810632	2	83196
	6	6_12X0	93136148	93281514	2	145366
	6	6_13X0	93667316	93715427	2	48111
	6	6_14X0	93887417	94416323	6	528906
	6	6_15X0	94897473	94911094	2	13621
	6	6_16X0	98127521	98212779	3	85258
	6	6_17X0	104266446	104473227	2	206781
	6	6_18X0	108885163	108913745	2	28582
	6	6_1X0	977417	1209717	3	232300
	6	6_2X0	13783815	13891389	2	107574
	6	6_3X0	48635918	48796720	2	160802
	6	6_4X0	49985404	50085164	2	99760
	6	6_5X0	51682970	51817839	2	134869
	6	6_6X0	55513727	55584903	4	71176
	6	6_7X0	57354589	57452803	3	98214
	6	6_8X0	58493665	58828042	4	334377
	6	6_9X0	59092045	59154823	2	62778
	7b	7b_10X0	53039764	53132813	2	93049

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	7b	7b_11X0	53274358	53281767	2	7409
	7b	7b_12X0	53939741	54011577	2	71836
	7b	7b_13X0	54245822	54288982	2	43160
	7b	7b_14X0	54441295	54671627	5	230332
	7b	7b_15X0	60518255	60581411	2	63156
	7b	7b_16X0	61381538	61400167	2	18629
	7b	7b_17X0	78500990	78553226	3	52236
	7b	7b_18X0	79580333	79720713	2	140380
	7b	7b_19X0	88793563	88881259	2	87696
	7b	7b_1X0	550104	566192	2	16088
	7b	7b_20X0	107446737	107449610	2	2873
	7b	7b_21X0	110474963	110556447	2	81484
	7b	7b_22X0	112368946	112422633	2	53687
	7b	7b_2X0	6692594	6760290	2	67696
	7b	7b_3X0	8510029	8539767	3	29738
	7b	7b_4X0	9402434	9453165	2	50731
	7b	7b_5X0	11543180	11642999	4	99819
	7b	7b_6X0	13542795	13561651	2	18856
	7b	7b_7X0	14758293	14803787	2	45494
	7b	7b_8X0	49776326	49830712	2	54386
	7b	7b_9X0	52529778	52862811	7	333033
	8	8_10X0	29778724	30007281	3	228557
	8	8_11X0	45117643	45460449	3	342806
	8	8_12X0	45534922	45649391	2	114469
	8	8_13X0	48345342	48409743	2	64401
	8	8_14X0	53225035	53338611	3	113576
	8	8_15X0	53546082	53613432	2	67350
	8	8_16X0	76754212	76760518	2	6306
	8	8_17X0	76795559	76958121	4	162562
	8	8_18X0	77049438	77308496	29	259058
	8	8_19X0	77392671	77515156	6	122485
	8	8_1X0	66536	142548	2	76012
	8	8_20X0	81862859	81874110	2	11251
	8	8_21X0	82060332	82271528	2	211196
	8	8_22X0	83159947	83188549	2	28602
	8	8_23X0	84111715	84190132	2	78417
	8	8_24X0	91902534	91992039	2	89505
	8	8_25X0	92035708	92154014	2	118306
	8	8_26X0	92165140	92233756	2	68616
	8	8_27X0	101614109	101808584	6	194475
	8	8_28X0	102054556	102063366	2	8810
	8	8_29X0	107071410	107104238	2	32828
	8	8_2X0	2109375	2131979	2	22604
	8	8_30X0	109006734	109027654	2	20920
	8	8_31X0	112604626	112628839	2	24213
	8	8_32X0	114075548	114112592	2	37044
	8	8_33X0	114723269	114847896	5	124627
	8	8_34X0	115918753	116163057	7	244304
	8	8_3X0	2183140	2316613	2	133473
	8	8_4X0	2776488	2915993	3	139505
	8	8_5X0	2976102	3190519	5	214417
	8	8_6X0	11464455	11551438	2	86983
	8	8_7X0	17141465	17266722	3	125257
	8	8_8X0	28566995	28704370	2	137375
	8	8_9X0	29142233	29169907	2	27674

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	9	9_10X0	53363946	54136822	8	772876
	9	9_11X0	54211012	54350023	2	139011
	9	9_12X0	54490243	54558437	3	68194
	9	9_13X0	54873922	54953618	3	79696
	9	9_14X0	58012813	58117208	3	104395
	9	9_15X0	58361069	58719350	9	358281
	9	9_16X0	60708074	60866184	3	158110
	9	9_17X0	61894786	62017897	2	123111
	9	9_18X0	67143722	67436043	3	292321
	9	9_19X0	72334577	72419501	2	84924
	9	9_1X0	8356390	8426130	2	69740
	9	9_20X0	73326217	73656670	3	330453
	9	9_21X0	83850896	84216355	6	365459
	9	9_22X0	87750990	87918230	2	167240
	9	9_23X0	113094208	113401036	5	306828
	9	9_24X0	117207265	117238444	2	31179
	9	9_2X0	11414573	11549789	2	135216
	9	9_3X0	12689199	12868079	3	178880
	9	9_4X0	12978730	13029383	2	50653
	9	9_5X0	14816173	14871759	2	55586
	9	9_6X0	19188783	19527704	5	338921
	9	9_7X0	35573509	35726455	2	152946
	9	9_8X0	37395825	37853124	4	457299
	9	9_9X0	47481132	47483275	2	2143
	GL397432.1	GL397432.1_1X0	230835	325894	6	95059
	GL397444.1	GL397444.1_1X0	966959	1026499	4	59540
	GL397444.1	GL397444.1_2X0	2490463	2515192	2	24729
	GL397460.1	GL397460.1_1X0	5199	10288	2	5089
	GL397460.1	GL397460.1_2X0	48099	119216	6	71117
	GL397460.1	GL397460.1_3X0	237369	286514	4	49145
	GL397460.1	GL397460.1_3X1	187595	225748	3	38153
	GL397460.1	GL397460.1_4X0	302934	371039	6	68105
	GL397460.1	GL397460.1_5X0	411673	431233	2	19560
	GL397460.1	GL397460.1_6X0	516041	530550	2	14509
	GL397460.1	GL397460.1_7X0	741299	768691	2	27392
	GL397460.1	GL397460.1_8X0	1770188	1783512	2	13324
	GL397538.1	GL397538.1_1X0	674909	715116	4	40207
	GL397538.1	GL397538.1_2X0	760914	804442	3	43528
	X	X_10X0	42125285	42392538	5	267253
	X	X_11X0	43180516	43428751	4	248235
	X	X_12X0	43491959	43736718	3	244759
	X	X_13X0	44090738	44174636	2	83898
	X	X_14X0	48053627	48231041	2	177414
	X	X_15X0	63154481	63342539	3	188058
	X	X_16X0	64435260	64549242	2	113982
	X	X_17X0	64800665	64839943	2	39278
	X	X_18X0	70283302	70649570	3	366268
	X	X_19X0	77319541	77492680	2	173139
	X	X_1X0	953972	1037912	2	83940
	X	X_20X0	89054490	89090870	2	36380
	X	X_21X0	90232061	90424986	2	192925
	X	X_22X0	90512728	90556012	2	43284
	X	X_23X0	90861416	91360244	9	498828
	X	X_24X0	103524788	103573265	2	48477
	X	X_25X0	106384387	106447646	2	63259

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	X	X_26X0	121446859	121528682	2	81823
	X	X_27X0	121910079	121996941	2	86862
	X	X_28X0	135716480	135779436	2	62956
	X	X_29X0	136445982	136631598	2	185616
	X	X_2X0	1139206	1434295	4	295089
	X	X_30X0	139056376	139138496	2	82120
	X	X_31X0	139394520	139498122	2	103602
	X	X_3X0	10855360	10910675	2	55315
	X	X_4X0	13247223	13321269	2	74046
	X	X_5X0	24275504	24303011	2	27507
	X	X_6X0	28322282	28358865	2	36583
	X	X_7X0	37353520	37512532	2	159012
	X	X_8X0	37613447	37712351	2	98904
	X	X_9X0	40466054	40635330	2	169276
	1	1_10X0	20645335	20659452	2	14117
	1	1_11X0	20722662	20776025	2	53363
	1	1_12X0	21119180	21467954	6	348774
	1	1_13X0	22406675	22502105	3	95430
	1	1_14X0	26944008	27073109	2	129101
	1	1_15X0	27075758	27139474	2	63716
	1	1_16X0	34452841	34712820	4	259979
	1	1_17X0	54450661	54495155	2	44494
	1	1_18X0	61332244	61434315	2	102071
	1	1_19X0	62895304	63196907	5	301603
	1	1_1X0	1290546	1479046	6	188500
	1	1_20X0	65625202	66040377	5	415175
	1	1_21X0	66179794	66551541	3	371747
	1	1_22X0	77192131	77255046	2	62915
	1	1_23X0	85995810	86030014	2	34204
	1	1_24X0	86103887	86237871	5	133984
	1	1_25X0	86673229	87058005	3	384776
	1	1_26X0	87710082	87982027	8	271945
	1	1_27X0	89330380	89460021	2	129641
	1	1_28X0	93028623	93139453	2	110830
	1	1_29X0	93478527	93573646	2	95119
	1	1_2X0	7058779	7154292	4	95513
	1	1_30X0	96211023	96404843	3	193820
	1	1_31X0	97838727	97978779	2	140052
	1	1_32X0	98228230	98310688	2	82458
	1	1_33X0	98693968	98784214	2	90246
	1	1_34X0	98848183	98906550	10	58367
	1	1_35X0	99623159	99677407	2	54248
	1	1_36X0	99863474	99891563	2	28089
	1	1_37X0	100109335	100283344	2	174009
	1	1_38X0	100688677	100770288	2	81611
	1	1_39X0	101579898	101930299	4	350401
	1	1_3X0	9558414	9563266	2	4852
	1	1_40X0	102253641	102291777	2	38136
	1	1_41X0	102566103	102639148	2	73045
	1	1_42X0	103807551	103839259	2	31708
	1	1_43X0	103853881	103919989	3	66108
	1	1_44X0	104509459	104671086	2	161627
	1	1_45X0	105365778	105451105	3	85327
	1	1_46X0	106317637	107003419	8	685782
	1	1_46X1	105698315	106268329	5	570014

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_47X0	107108011	107649490	12	541479
	1	1_47X1	107710982	108103688	11	392706
	1	1_48X0	108164514	108277577	2	113063
	1	1_49X0	108452661	108823372	8	370711
	1	1_4X0	10521948	10741067	3	219119
	1	1_50X0	109141097	109448988	9	307891
	1	1_51X0	109476936	109541160	2	64224
	1	1_52X0	109569606	109681985	2	112379
	1	1_53X0	109863404	110649461	11	786057
	1	1_53X1	110829042	111175156	6	346114
	1	1_54X0	112790132	112936012	2	145880
	1	1_55X0	119257114	119343098	2	85984
	1	1_56X0	120327318	120571823	4	244505
	1	1_57X0	123359533	123378489	2	18956
	1	1_58X0	123422302	123578504	3	156202
	1	1_59X0	123650073	123832843	7	182770
	1	1_5X0	13972375	14164083	2	191708
	1	1_60X0	123915092	123973280	2	58188
	1	1_61X0	127592697	127756508	2	163811
	1	1_62X0	160568914	160692910	2	123996
	1	1_63X0	160946471	161288907	9	342436
	1	1_64X0	172525876	172562743	2	36867
	1	1_65X0	175438052	175498685	2	60633
	1	1_66X0	175935175	176381603	3	446428
	1	1_67X0	184627807	184734131	2	106324
	1	1_68X0	185222516	185325662	2	103146
	1	1_69X0	187947243	188056039	2	108796
	1	1_6X0	14510336	14883985	3	373649
	1	1_70X0	191400063	191437795	2	37732
	1	1_71X0	195151317	195295348	2	144031
	1	1_72X0	198138027	198315225	2	177198
	1	1_73X0	198708986	198813239	4	104253
	1	1_74X0	201245864	201326718	3	80854
	1	1_75X0	201668544	201821749	2	153205
	1	1_76X0	215545763	215608486	3	62723
	1	1_77X0	225078853	225247042	2	168189
	1	1_78X0	226893623	226959880	2	66257
	1	1_79X0	234199476	234380044	3	180568
	1	1_7X0	16465544	16474553	2	9009
	1	1_80X0	234815179	234964434	5	149255
	1	1_81X0	236691318	236913148	3	221830
	1	1_82X0	236943080	237045220	2	102140
	1	1_83X0	237343563	237431369	2	87806
	1	1_84X0	239011056	239040481	2	29425
	1	1_85X0	239967093	240351917	4	384824
	1	1_86X0	247591841	247687648	2	95807
	1	1_87X0	248099281	248193390	4	94109
	1	1_88X0	253571546	253626541	2	54995
	1	1_89X0	253633735	253733748	2	100013
	1	1_8X0	17386096	17587221	6	201125
	1	1_90X0	259688071	259762804	2	74733
	1	1_91X0	260634425	260705731	3	71306
	1	1_92X0	261810911	261953055	4	142144
	1	1_93X0	262453451	262536528	2	83077
	1	1_94X0	262593750	262710074	4	116324

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_95X0	263809116	263975267	2	166151
	1	1_96X0	266404853	266492080	4	87227
	1	1_9X0	17634305	17818632	2	184327
	10	10_10X0	75465173	75607852	3	142679
	10	10_11X0	78274732	78472545	2	197813
	10	10_12X0	85829334	85861100	3	31766
	10	10_1X0	11689685	11772210	2	82525
	10	10_2X0	18533288	18769088	2	235800
	10	10_3X0	21746108	21831130	2	85022
	10	10_4X0	30217152	30265933	2	48781
	10	10_5X0	32053343	32276062	2	222719
	10	10_6X0	32304264	32342084	2	37820
	10	10_7X0	36253000	36305769	3	52769
	10	10_8X0	64690523	64759576	2	69053
	10	10_9X0	70848921	71223980	3	375059
	11	11_10X0	22453045	22478665	2	25620
	11	11_11X0	23509437	23829312	15	319875
	11	11_12X0	23853979	23909161	2	55182
	11	11_13X0	24008302	24122789	2	114487
	11	11_14X0	24527709	24633665	2	105956
	11	11_15X0	26004562	26125739	2	121177
	11	11_16X0	26321704	26425109	3	103405
	11	11_17X0	26447251	26453354	2	6103
	11	11_18X0	26457911	26529770	3	71859
	11	11_19X0	26731465	26737001	2	5536
	11	11_1X0	5800378	5970414	3	170036
	11	11_20X0	26842600	26853534	2	10934
	11	11_21X0	27256343	27321627	3	65284
	11	11_22X0	27434978	27455340	2	20362
	11	11_23X0	27862277	27948643	2	86366
	11	11_24X0	28880149	29169247	6	289098
	11	11_25X0	32631477	32695198	3	63721
	11	11_26X0	33426696	33548855	4	122159
	11	11_27X0	33629196	33756798	2	127602
	11	11_28X0	35517434	35575612	2	58178
	11	11_29X0	35715931	35750603	2	34672
	11	11_2X0	8522316	8674890	4	152574
	11	11_30X0	35781196	35831177	2	49981
	11	11_31X0	36043787	36079960	2	36173
	11	11_32X0	36195487	36220371	2	24884
	11	11_33X0	37337231	37509285	10	172054
	11	11_34X0	39819278	39884850	2	65572
	11	11_35X0	39957731	39977351	2	19620
	11	11_36X0	40438505	40762811	12	324306
	11	11_37X0	40766480	40786669	4	20189
	11	11_38X0	40811713	41009391	10	197678
	11	11_39X0	41042320	41338864	15	296544
	11	11_3X0	8770204	8807810	2	37606
	11	11_40X0	41511929	41535388	2	23459
	11	11_41X0	41589669	41649382	2	59713
	11	11_42X0	41653598	41748783	2	95185
	11	11_43X0	41807519	41934839	3	127320
	11	11_44X0	42166004	42215648	2	49644
	11	11_45X0	42353843	42402155	2	48312
	11	11_46X0	43114080	43211534	2	97454

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	11	11_47X0	43246612	43259202	2	12590
	11	11_48X0	44759833	44855421	2	95588
	11	11_49X0	47222569	47266860	2	44291
	11	11_4X0	13779812	14117595	14	337783
	11	11_50X0	47346954	47425001	2	78047
	11	11_51X0	47525517	47645553	3	120036
	11	11_52X0	47957629	48061882	2	104253
	11	11_53X0	49293868	49318361	2	24493
	11	11_54X0	49557349	49581823	2	24474
	11	11_55X0	49678876	49711157	2	32281
	11	11_56X0	50014563	50023438	2	8875
	11	11_57X0	51704269	51762847	3	58578
	11	11_58X0	53004485	53036753	2	32268
	11	11_59X0	54124802	54195420	2	70618
	11	11_5X0	14377648	14465104	3	87456
	11	11_60X0	54280969	54295283	2	14314
	11	11_61X0	54739292	54753637	2	14345
	11	11_62X0	55479278	55501215	2	21937
	11	11_63X0	55693429	55946573	9	253144
	11	11_64X0	59478483	59519150	2	40667
	11	11_65X0	60194810	60435471	4	240661
	11	11_66X0	61878125	62006804	3	128679
	11	11_6X0	15377590	15507249	4	129659
	11	11_7X0	18877132	18988041	2	110909
	11	11_8X0	19681858	19865755	4	183897
	11	11_9X0	20719581	20905075	4	185494
	12	12_10X0	34300387	34319227	2	18840
	12	12_11X0	35391480	35536270	3	144790
	12	12_12X0	36801185	37060513	5	259328
	12	12_13X0	39630985	39748949	2	117964
	12	12_14X0	39958532	39971270	2	12738
	12	12_15X0	40138218	40172744	3	34526
	12	12_16X0	44908843	44931429	2	22586
	12	12_17X0	45422916	45570611	2	147695
	12	12_18X0	49071662	49196241	2	124579
	12	12_19X0	49581741	49705466	2	123725
	12	12_1X0	522962	667966	3	145004
	12	12_20X0	50089624	50116738	2	27114
	12	12_21X0	52249182	52667825	15	418643
	12	12_22X0	53140561	53280142	2	139581
	12	12_23X0	59115276	59200099	2	84823
	12	12_24X0	61591772	61639956	2	48184
	12	12_25X0	61651860	61835896	3	184036
	12	12_26X0	68984008	69069280	2	85272
	12	12_27X0	73705977	74114558	6	408581
	12	12_28X0	77948098	78037269	2	89171
	12	12_2X0	2239744	2586767	3	347023
	12	12_3X0	2651299	2738876	2	87577
	12	12_4X0	2818397	3031185	2	212788
	12	12_5X0	3528897	3783235	3	254338
	12	12_6X0	3820935	3956683	4	135748
	12	12_7X0	3958538	4026874	3	68336
	12	12_8X0	22355069	22415876	2	60807
	12	12_9X0	25081358	25241603	2	160245
	13	13_10X0	52521878	52854357	4	332479

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	13	13_11X0	53097296	53110212	2	12916
	13	13_12X0	53446523	53461499	2	14976
	13	13_13X0	58960755	58985079	2	24324
	13	13_14X0	59833973	59935282	2	101309
	13	13_15X0	60050291	60373600	15	323309
	13	13_16X0	60667341	60921921	2	254580
	13	13_17X0	61611933	61711550	4	99617
	13	13_18X0	62197295	62243153	2	45858
	13	13_19X0	64318446	64329675	2	11229
	13	13_1X0	11718823	11865867	2	147044
	13	13_20X0	64408134	64499282	2	91148
	13	13_21X0	66754362	66906063	3	151701
	13	13_22X0	73036214	73118692	4	82478
	13	13_23X0	73265630	74097806	19	832176
	13	13_24X0	79098073	79418351	3	320278
	13	13_2X0	37399961	37599154	2	199193
	13	13_3X0	40956482	40983359	2	26877
	13	13_4X0	41190185	41299663	6	109478
	13	13_5X0	41989254	42167257	2	178003
	13	13_6X0	42254474	42268534	2	14060
	13	13_7X0	42404771	43042300	19	637529
	13	13_8X0	46225254	46252021	2	26767
	13	13_9X0	52030761	52173973	2	143212
	14	14_10X0	34149168	34340963	3	191795
	14	14_11X0	34789200	34803228	2	14028
	14	14_12X0	35337828	35459641	2	121813
	14	14_13X0	38887068	38924638	2	37570
	14	14_14X0	43017755	43097426	2	79671
	14	14_15X0	44257056	44578735	10	321679
	14	14_16X0	44616468	44797352	3	180884
	14	14_17X0	44951568	44993838	4	42270
	14	14_18X0	45067277	45150045	3	82768
	14	14_19X0	45155223	45223904	4	68681
	14	14_1X0	1934718	1955712	3	20994
	14	14_20X0	45356047	45449119	3	93072
	14	14_21X0	45656452	45687144	2	30692
	14	14_22X0	45702859	45820851	2	117992
	14	14_23X0	45892869	46257029	9	364160
	14	14_23X1	46565263	46857852	8	292589
	14	14_23X2	46321079	46536568	6	215489
	14	14_24X0	46905204	46951654	2	46450
	14	14_25X0	47687488	47725503	3	38015
	14	14_26X0	47825668	48092858	6	267190
	14	14_27X0	48113820	48176517	3	62697
	14	14_28X0	48739975	48884833	3	144858
	14	14_29X0	48910893	48921463	2	10570
	14	14_2X0	2006995	2177411	2	170416
	14	14_30X0	49202173	49393112	5	190939
	14	14_30X1	49532621	49582843	2	50222
	14	14_30X2	49728829	49880496	3	151667
	14	14_31X0	50093951	50224092	3	130141
	14	14_32X0	50807110	51077252	9	270142
	14	14_32X1	50403061	50716754	8	313693
	14	14_33X0	51115840	51375030	8	259190
	14	14_34X0	51380571	51687767	6	307196

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	14	14_35X0	52187624	52253835	2	66211
	14	14_36X0	52276414	52405854	2	129440
	14	14_37X0	53371557	53397914	2	26357
	14	14_38X0	53854491	53911095	3	56604
	14	14_39X0	53989046	54057456	5	68410
	14	14_3X0	16007907	16189218	2	181311
	14	14_40X0	54353631	54416588	2	62957
	14	14_41X0	54757605	54832746	3	75141
	14	14_42X0	55362680	55408095	2	45415
	14	14_43X0	55932085	56086295	13	154210
	14	14_44X0	56120052	56543412	14	423360
	14	14_45X0	56646818	56816870	6	170052
	14	14_46X0	56819363	57200168	14	380805
	14	14_47X0	57597302	57892255	9	294953
	14	14_48X0	58191275	58435433	7	244158
	14	14_48X1	58512819	58762845	6	250026
	14	14_49X0	58865825	58932185	2	66360
	14	14_4X0	23434607	23498927	2	64320
	14	14_50X0	59077405	59105747	2	28342
	14	14_51X0	59188170	59297484	6	109314
	14	14_52X0	59539421	59556856	2	17435
	14	14_53X0	59643790	60011146	14	367356
	14	14_54X0	60075634	60367533	5	291899
	14	14_55X0	60399063	60433703	4	34640
	14	14_56X0	60434087	60504466	3	70379
	14	14_57X0	60569935	60578732	2	8797
	14	14_58X0	61826624	62099144	10	272520
	14	14_58X1	61422255	61785173	14	362918
	14	14_58X2	60580844	61065832	12	484988
	14	14_58X3	62459468	62699804	7	240336
	14	14_58X4	62169608	62418633	5	249025
	14	14_5X0	23505870	23562844	4	56974
	14	14_6X0	24794346	24854796	3	60450
	14	14_7X0	24960829	25104433	3	143604
	14	14_8X0	33634983	33731040	4	96057
	14	14_9X0	33986663	34037380	3	50717
	15	15_10X0	27968008	27990954	2	22946
	15	15_11X0	28065641	28300796	2	235155
	15	15_12X0	28322151	28446802	5	124651
	15	15_13X0	28501024	28529013	3	27989
	15	15_14X0	29353626	29452173	2	98547
	15	15_15X0	37098593	37372978	3	274385
	15	15_16X0	44116583	44464961	11	348378
	15	15_17X0	45212904	45257501	2	44597
	15	15_18X0	45291920	45571845	14	279925
	15	15_19X0	48075006	48399333	17	324327
	15	15_19X1	48411559	49266685	31	855126
	15	15_19X2	47700418	48056945	15	356527
	15	15_19X3	47018027	47679562	26	661535
	15	15_19X4	46168959	47008547	29	839588
	15	15_1X0	50615	133079	2	82464
	15	15_20X0	49652293	49717302	4	65009
	15	15_21X0	49927737	49948591	2	20854
	15	15_22X0	49993284	50034889	3	41605
	15	15_23X0	51596036	51628720	2	32684

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	15	15_24X0	52504597	52546345	3	41748
	15	15_25X0	53068022	53356145	5	288123
	15	15_26X0	54189957	54346348	2	156391
	15	15_27X0	59124632	59190048	2	65416
	15	15_28X0	72348534	72389751	2	41217
	15	15_29X0	76826527	77679529	26	853002
	15	15_29X1	76139437	76786772	16	647335
	15	15_2X0	222804	639922	4	417118
	15	15_30X0	77816389	77834803	2	18414
	15	15_31X0	77878362	77965571	2	87209
	15	15_32X0	78393959	78994303	10	600344
	15	15_33X0	79179627	79288198	3	108571
	15	15_34X0	79496430	79908754	10	412324
	15	15_35X0	80000452	80046922	2	46470
	15	15_36X0	80068770	80095965	2	27195
	15	15_37X0	80103848	80412010	8	308162
	15	15_38X0	80547423	80674032	3	126609
	15	15_3X0	5258462	5717067	8	458605
	15	15_4X0	5861839	5967347	2	105508
	15	15_5X0	14891501	14913650	2	22149
	15	15_6X0	21674873	21679815	2	4942
	15	15_7X0	23910103	23985964	2	75861
	15	15_8X0	24567161	24678323	2	111162
	15	15_9X0	26930074	26994743	3	64669
	16	16_1X0	5792030	5861262	2	69232
	16	16_2X0	6839129	6908793	2	69664
	16	16_3X0	9242499	9374702	2	132203
	16	16_4X0	10211260	10319662	2	108402
	16	16_5X0	23267860	23392995	2	125135
	16	16_6X0	24075358	24153334	3	77976
	16	16_7X0	38069831	38131379	2	61548
	16	16_8X0	58378017	58454930	2	76913
	16	16_9X0	71554894	71695091	2	140197
	17	17_10X0	62296759	62396924	2	100165
	17	17_11X0	68936119	69029382	3	93263
	17	17_12X0	70018319	70109962	2	91643
	17	17_13X0	70136803	70485028	11	348225
	17	17_14X0	70578342	70598980	3	20638
	17	17_15X0	70712275	70733791	2	21516
	17	17_1X0	2657181	2715366	3	58185
	17	17_2X0	43174199	43319709	2	145510
	17	17_3X0	43624823	43918149	6	293326
	17	17_4X0	52373620	52391453	2	17833
	17	17_5X0	52931436	53002859	2	71423
	17	17_6X0	53511024	53616466	2	105442
	17	17_7X0	54923732	55017200	2	93468
	17	17_8X0	60655935	60693139	2	37204
	17	17_9X0	61848477	61889731	3	41254
	18	18_10X0	57825384	58178141	10	352757
	18	18_11X0	59208858	59250673	2	41815
	18	18_12X0	63869774	63904907	2	35133
	18	18_13X0	65194253	65195158	2	905
	18	18_14X0	66480650	66495905	2	15255
	18	18_15X0	68448003	68551257	3	103254
	18	18_1X0	18773141	18805714	2	32573

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	18	18_2X0	18812462	19382132	18	569670
	18	18_3X0	19520823	19571940	2	51117
	18	18_4X0	29994614	30061504	3	66890
	18	18_5X0	32747886	32957201	5	209315
	18	18_6X0	33398299	33444372	2	46073
	18	18_7X0	33604837	33736798	7	131961
	18	18_8X0	53953811	54001653	2	47842
	18	18_9X0	57586531	57630796	2	44265
	19	19_10X0	48598004	48967002	5	368998
	19	19_11X0	48970460	49039744	6	69284
	19	19_12X0	49974113	49988891	3	14778
	19	19_13X0	50037255	50136691	4	99436
	19	19_14X0	50170137	50280996	2	110859
	19	19_15X0	50302387	50341594	2	39207
	19	19_16X0	50708003	50856490	2	148487
	19	19_17X0	51520643	51589754	7	69111
	19	19_18X0	52591644	52701058	7	109414
	19	19_19X0	52912035	53340806	7	428771
	19	19_1X0	2538555	2540860	2	2305
	19	19_20X0	54752592	54964101	2	211509
	19	19_21X0	57941567	57958832	2	17265
	19	19_2X0	6716368	6845557	3	129189
	19	19_3X0	11578992	11793283	2	214291
	19	19_4X0	12488293	12553961	2	65668
	19	19_5X0	13236727	13310243	2	73516
	19	19_6X0	14718770	14738919	2	20149
	19	19_7X0	16224437	16373650	7	149213
	19	19_8X0	16730706	16753272	2	22566
	19	19_9X0	48037537	48091789	3	54252
	2	2_10X0	36684455	36795204	2	110749
	2	2_11X0	37506644	37564903	2	58259
	2	2_12X0	40522108	40558659	2	36551
	2	2_13X0	41012453	41220272	3	207819
	2	2_14X0	41867521	41907137	2	39616
	2	2_15X0	47144650	47204625	2	59975
	2	2_16X0	52245976	52275573	2	29597
	2	2_17X0	52423842	52607846	2	184004
	2	2_18X0	68630812	68813039	3	182227
	2	2_19X0	73044621	73143766	2	99145
	2	2_1X0	709822	786165	2	76343
	2	2_20X0	88871549	89257528	13	385979
	2	2_21X0	103251573	103307651	2	56078
	2	2_22X0	104544028	104559224	2	15196
	2	2_23X0	104709999	104749490	3	39491
	2	2_24X0	112778077	112837285	2	59208
	2	2_25X0	119063656	119270385	2	206729
	2	2_26X0	130517833	130693174	2	175341
	2	2_27X0	131379920	131446354	2	66434
	2	2_28X0	132820016	132926561	10	106545
	2	2_29X0	136584803	136602848	2	18045
	2	2_2X0	10332512	10398584	2	66072
	2	2_30X0	138681642	138856360	2	174718
	2	2_31X0	150789011	151059867	2	270856
	2	2_32X0	157033323	157073134	2	39811
	2	2_33X0	183217546	183402771	2	185225

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	2	2_34X0	201796182	202104732	3	308550
	2	2_35X0	202371262	202557903	2	186641
	2	2_36X0	204637945	204873693	3	235748
	2	2_37X0	217607709	217669182	2	61473
	2	2_38X0	219176197	219204184	2	27987
	2	2_39X0	219811922	219844940	2	33018
	2	2_3X0	10416301	10517310	2	101009
	2	2_40X0	220171144	220475937	8	304793
	2	2_41X0	231786976	231978154	2	191178
	2	2_42X0	233074053	233158712	4	84659
	2	2_43X0	233224523	233251567	2	27044
	2	2_44X0	233255802	233275235	2	19433
	2	2_45X0	234111750	234242941	2	131191
	2	2_46X0	234303477	234419987	3	116510
	2	2_47X0	234812335	234959661	2	147326
	2	2_48X0	238570808	238803401	2	232593
	2	2_49X0	240595063	240617850	2	22787
	2	2_4X0	10653132	10679453	2	26321
	2	2_50X0	241618700	241679859	2	61159
	2	2_51X0	243051704	243078666	3	26962
	2	2_52X0	244750617	244806309	2	55692
	2	2_53X0	245350941	245618188	7	267247
	2	2_54X0	246181823	246216518	2	34695
	2	2_55X0	248054811	248351464	5	296653
	2	2_5X0	12957435	13028295	2	70860
	2	2_6X0	18273095	18495035	6	221940
	2	2_7X0	28621797	28737039	4	115242
	2	2_8X0	31653188	31765595	2	112407
	2	2_9X0	36644216	36666948	2	22732
	20	20_10X0	19430738	19457651	2	26913
	20	20_11X0	20765718	20828101	2	62383
	20	20_12X0	22188728	22326873	3	138145
	20	20_13X0	22382209	22443011	4	60802
	20	20_14X0	23054320	23178184	2	123864
	20	20_15X0	24394173	24428779	2	34606
	20	20_16X0	24871703	25013919	4	142216
	20	20_17X0	25666699	26029599	8	362900
	20	20_17X1	25353186	25608591	6	255405
	20	20_18X0	26237128	26339178	3	102050
	20	20_19X0	26463053	26515796	3	52743
	20	20_1X0	7164298	7425930	11	261632
	20	20_20X0	26649266	26666275	3	17009
	20	20_21X0	26727242	26768301	5	41059
	20	20_22X0	26850361	26859917	3	9556
	20	20_23X0	26924091	27141475	6	217384
	20	20_24X0	27446468	27499475	2	53007
	20	20_25X0	27566931	27769637	4	202706
	20	20_26X0	27786720	28008640	8	221920
	20	20_27X0	28111661	28446611	13	334950
	20	20_27X1	28472967	28651401	7	178434
	20	20_27X2	28703395	28964002	9	260607
	20	20_28X0	29190452	29570737	11	380285
	20	20_29X0	29583479	29903426	10	319947
	20	20_2X0	7582445	7728984	2	146539
	20	20_30X0	29930396	29991703	16	61307

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	20	20_31X0	30152364	30184206	2	31842
	20	20_32X0	30305149	30316927	4	11778
	20	20_33X0	30524486	30582945	3	58459
	20	20_34X0	30634402	30842673	24	208271
	20	20_35X0	30890662	31023504	4	132842
	20	20_36X0	31077040	31080597	2	3557
	20	20_37X0	33537858	33930211	8	392353
	20	20_38X0	34080419	34266900	4	186481
	20	20_39X0	49563909	49648102	3	84193
	20	20_3X0	9537073	9545402	2	8329
	20	20_4X0	9880500	9894286	2	13786
	20	20_5X0	9920838	10182291	3	261453
	20	20_6X0	13101448	13224512	3	123064
	20	20_7X0	15158837	15237353	4	78516
	20	20_8X0	15994971	16014486	2	19515
	20	20_9X0	16727427	16964261	4	236834
	21	21_10X0	31445460	31528698	3	83238
	21	21_11X0	32204737	32362290	2	157553
	21	21_12X0	35685217	35750447	2	65230
	21	21_13X0	36240528	36337457	4	96929
	21	21_14X0	36499871	36653161	2	153290
	21	21_15X0	36714148	37214599	14	500451
	21	21_16X0	37359746	37839504	7	479758
	21	21_17X0	37988416	38185249	4	196833
	21	21_18X0	38225332	38547817	6	322485
	21	21_19X0	38621009	39182575	10	561566
	21	21_1X0	4740779	5298388	11	557609
	21	21_20X0	39652537	39813053	3	160516
	21	21_21X0	40081840	40188084	3	106244
	21	21_22X0	40333115	40538442	2	205327
	21	21_23X0	40623582	41131798	10	508216
	21	21_24X0	41152231	41278440	3	126209
	21	21_25X0	41421748	41490013	2	68265
	21	21_26X0	42338172	42383929	2	45757
	21	21_27X0	43015535	43105245	2	89710
	21	21_28X0	43951242	43999246	3	48004
	21	21_29X0	44484585	44610751	2	126166
	21	21_2X0	9341467	9480340	2	138873
	21	21_30X0	44627876	44685101	2	57225
	21	21_31X0	44692716	44722852	2	30136
	21	21_32X0	44780244	44839141	2	58897
	21	21_33X0	45593239	45642961	3	49722
	21	21_34X0	46224390	46329754	3	105364
	21	21_35X0	49743059	49792831	4	49772
	21	21_3X0	17274681	17411025	2	136344
	21	21_4X0	18163745	18649670	3	485925
	21	21_5X0	25476296	25641018	4	164722
	21	21_6X0	25773003	25829605	4	56602
	21	21_7X0	25864695	26184035	13	319340
	21	21_7X1	26242848	26566383	9	323535
	21	21_7X2	26622891	27215656	15	592765
	21	21_8X0	27444759	27619070	5	174311
	21	21_9X0	28322389	28384658	2	62269
	22	22_10X0	41405011	41457647	3	52636
	22	22_11X0	41612410	41662622	2	50212

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	22	22_12X0	50569811	50590967	2	21156
	22	22_1X0	64240	138525	3	74285
	22	22_2X0	10012134	10240383	4	228249
	22	22_3X0	10770400	10835980	4	65580
	22	22_4X0	14113175	14129084	2	15909
	22	22_5X0	15607416	15937730	5	330314
	22	22_6X0	21103942	21179307	2	75365
	22	22_7X0	23110257	23145246	3	34989
	22	22_8X0	23877943	23940751	3	62808
	22	22_9X0	35829080	35990153	3	161073
	23	23_10X0	49154028	49159743	2	5715
	23	23_11X0	61974470	62330628	10	356158
	23	23_1X0	105182	210793	4	105611
	23	23_2X0	3926172	3967571	2	41399
	23	23_3X0	12582187	12641984	2	59797
	23	23_4X0	22019242	22093233	4	73991
	23	23_5X0	25364704	25456840	2	92136
	23	23_6X0	25896411	26427620	7	531209
	23	23_7X0	37747768	37776271	2	28503
	23	23_8X0	43867867	43894748	2	26881
	23	23_9X0	45567409	45696614	2	129205
	24	24_10X0	2535825	2864490	9	328665
	24	24_11X0	14341016	14518770	2	177754
	24	24_12X0	16885341	17040297	2	154956
	24	24_13X0	17926803	17963018	2	36215
	24	24_14X0	18027363	18245264	3	217901
	24	24_15X0	18411601	18526247	2	114646
	24	24_16X0	21338359	21370787	2	32428
	24	24_17X0	24897923	25002160	2	104237
	24	24_18X0	26188152	26294593	2	106441
	24	24_19X0	26444852	26921216	7	476364
	24	24_1X0	147873	158052	3	10179
	24	24_20X0	27012122	27032517	2	20395
	24	24_21X0	27112495	27154401	2	41906
	24	24_22X0	27170160	27185904	2	15744
	24	24_23X0	27188022	27231806	3	43784
	24	24_24X0	27341076	27473467	3	132391
	24	24_25X0	32635145	32902784	3	267639
	24	24_26X0	33395757	33499104	4	103347
	24	24_27X0	33953102	33982087	2	28985
	24	24_28X0	34407750	34426790	2	19040
	24	24_29X0	35409415	35438155	2	28740
	24	24_2X0	220472	272301	2	51829
	24	24_30X0	35554263	35624755	3	70492
	24	24_31X0	36287233	36328608	3	41375
	24	24_32X0	36450665	36645659	3	194994
	24	24_33X0	36692057	36820519	7	128462
	24	24_34X0	36885477	36926200	2	40723
	24	24_35X0	38382711	38460219	3	77508
	24	24_36X0	41446559	41465034	2	18475
	24	24_3X0	489843	535781	2	45938
	24	24_4X0	921705	942423	2	20718
	24	24_5X0	953067	993397	2	40330
	24	24_6X0	1510460	1721314	4	210854
	24	24_7X0	1787914	1928169	2	140255

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	24	24_8X0	2109485	2233876	6	124391
	24	24_9X0	2409319	2411880	2	2561
	25	25_10X0	41554041	41693923	2	139882
	25	25_11X0	41835784	41861362	2	25578
	25	25_12X0	45176441	45216521	2	40080
	25	25_1X0	2035527	2734430	10	698903
	25	25_2X0	12976510	13055784	2	79274
	25	25_3X0	25001146	25371374	3	370228
	25	25_4X0	25385759	25560202	2	174443
	25	25_5X0	30693143	30760364	3	67221
	25	25_6X0	35214535	35281088	4	66553
	25	25_7X0	35349401	35386268	2	36867
	25	25_8X0	38818831	38840028	2	21197
	25	25_9X0	41183169	41193792	2	10623
	26	26_1X0	5666629	6047793	16	381164
	26	26_2X0	12908705	12911506	2	2801
	26	26_3X0	14999909	15047055	2	47146
	26	26_4X0	16362576	16402747	2	40171
	26	26_5X0	32750491	33475095	5	724604
	26	26_6X0	33532844	33625109	2	92265
	26	26_7X0	35955868	36016236	2	60368
	3	3_10X0	14290904	14400494	3	109590
	3	3_11X0	31618393	31675782	2	57389
	3	3_12X0	34278335	34513217	4	234882
	3	3_13X0	39264293	39319994	3	55701
	3	3_14X0	49907517	50018149	3	110632
	3	3_15X0	57272769	57284889	2	12120
	3	3_16X0	57360889	57395975	2	35086
	3	3_17X0	58476359	58552174	2	75815
	3	3_18X0	59304260	59417902	8	113642
	3	3_19X0	59709089	60088697	10	379608
	3	3_1X0	544747	585846	2	41099
	3	3_20X0	61703391	61769403	3	66012
	3	3_21X0	79242090	79306342	2	64252
	3	3_22X0	80333311	80384030	2	50719
	3	3_23X0	84680681	84920666	2	239985
	3	3_24X0	86686321	86871162	2	184841
	3	3_25X0	92818953	92842334	2	23381
	3	3_26X0	96286823	96304041	2	17218
	3	3_27X0	99015266	99191678	2	176412
	3	3_28X0	99200154	99683489	6	483335
	3	3_29X0	103584795	103598419	2	13624
	3	3_2X0	909517	966613	6	57096
	3	3_30X0	104397723	104432413	3	34690
	3	3_31X0	104476131	104619322	3	143191
	3	3_32X0	106215630	106328687	2	113057
	3	3_33X0	111181346	111357456	3	176110
	3	3_34X0	116604900	116622399	2	17499
	3	3_35X0	127257338	127458485	4	201147
	3	3_36X0	132314501	132429375	9	114874
	3	3_37X0	132874525	132937701	2	63176
	3	3_38X0	132949220	133155342	2	206122
	3	3_39X0	133329883	134042991	29	713108
	3	3_3X0	2371399	2501464	2	130065
	3	3_40X0	134169941	134236131	2	66190

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	3	3_41X0	134994204	135193662	3	199458
	3	3_42X0	136144860	136178344	3	33484
	3	3_43X0	136228882	136366740	2	137858
	3	3_44X0	136794891	136929972	2	135081
	3	3_45X0	136976492	137151088	4	174596
	3	3_46X0	137462191	137810591	5	348400
	3	3_47X0	139936867	140063468	2	126601
	3	3_48X0	144527676	144604144	2	76468
	3	3_49X0	150225121	150439510	6	214389
	3	3_4X0	3542728	3634070	4	91342
	3	3_50X0	151427327	151470372	2	43045
	3	3_51X0	161898847	161916030	2	17183
	3	3_52X0	162414465	162604684	4	190219
	3	3_53X0	162903225	162921999	2	18774
	3	3_54X0	163152226	163276634	2	124408
	3	3_55X0	163532333	164035525	15	503192
	3	3_55X1	164104035	164853127	16	749092
	3	3_56X0	176752381	176784127	3	31746
	3	3_57X0	179324267	179532260	4	207993
	3	3_58X0	180166727	180373738	4	207011
	3	3_59X0	193610779	193867360	3	256581
	3	3_5X0	6663184	6678962	2	15778
	3	3_60X0	199925153	199940386	2	15233
	3	3_61X0	201508686	201664371	2	155685
	3	3_62X0	201807778	201842949	2	35171
	3	3_63X0	203291302	203324978	2	33676
	3	3_64X0	203469549	203667549	5	198000
	3	3_65X0	203872448	205197724	32	1325276
	3	3_66X0	205321986	205535076	3	213090
	3	3_67X0	205622220	205709408	2	87188
	3	3_68X0	205985865	206040784	2	54919
	3	3_69X0	206376134	206459013	4	82879
	3	3_6X0	8222957	8250043	2	27086
	3	3_70X0	206582108	207098973	4	516865
	3	3_71X0	207275702	207366557	3	90855
	3	3_72X0	207577228	207878690	5	301462
	3	3_73X0	207881562	207984844	3	103282
	3	3_74X0	209636333	209698301	2	61968
	3	3_75X0	210506053	210559513	2	53460
	3	3_76X0	211137409	211199991	2	62582
	3	3_77X0	213422852	213503025	2	80173
	3	3_78X0	214514393	214557335	2	42942
	3	3_79X0	217850300	217992766	2	142466
	3	3_7X0	11238597	11309195	2	70598
	3	3_80X0	218607386	218672240	2	64854
	3	3_81X0	218742630	218892320	2	149690
	3	3_8X0	12748373	12757704	2	9331
	3	3_9X0	12880614	13322141	12	441527
	4	4_10X0	50075590	50118654	2	43064
	4	4_11X0	51786039	51848168	2	62129
	4	4_12X0	65212844	65646350	5	433506
	4	4_13X0	68787397	68937869	10	150472
	4	4_14X0	74385704	74454728	2	69024
	4	4_15X0	76176920	76192761	2	15841
	4	4_16X0	82590902	82755780	6	164878

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	4	4_17X0	87135683	87141028	2	5345
	4	4_18X0	88181214	88369080	4	187866
	4	4_19X0	92718819	92767515	2	48696
	4	4_1X0	2255574	2448893	2	193319
	4	4_20X0	94073222	94163486	3	90264
	4	4_21X0	98109027	98217241	2	108214
	4	4_22X0	102742335	102801541	2	59206
	4	4_23X0	104791789	105566665	49	774876
	4	4_24X0	105627309	105690505	2	63196
	4	4_25X0	105907079	105976545	3	69466
	4	4_26X0	106006371	106312585	5	306214
	4	4_27X0	106334954	106961405	12	626451
	4	4_28X0	111592405	112197087	8	604682
	4	4_29X0	112226366	112744989	14	518623
	4	4_2X0	11110712	11128973	2	18261
	4	4_30X0	112797026	112817293	2	20267
	4	4_31X0	113001972	113194740	2	192768
	4	4_32X0	113862258	113975698	2	113440
	4	4_3X0	12410438	12539560	3	129122
	4	4_4X0	14251019	14268503	2	17484
	4	4_5X0	14338474	14426196	2	87722
	4	4_6X0	25372592	25441508	2	68916
	4	4_7X0	27866988	27902556	2	35568
	4	4_8X0	40240542	40458528	3	217986
	4	4_9X0	48690960	48821011	2	130051
	5	5_10X0	9069829	9270698	2	200869
	5	5_11X0	9450503	9660161	2	209658
	5	5_12X0	10886065	10903121	2	17056
	5	5_13X0	10960312	11105734	4	145422
	5	5_14X0	11326459	11724281	6	397822
	5	5_14X1	11836336	12084118	3	247782
	5	5_15X0	12500037	12554248	4	54211
	5	5_16X0	13494355	13628213	2	133858
	5	5_17X0	14206784	14264659	3	57875
	5	5_18X0	14743613	14985321	2	241708
	5	5_19X0	15000428	15289527	4	289099
	5	5_1X0	2001405	2165961	3	164556
	5	5_20X0	15629500	15650634	2	21134
	5	5_21X0	16230123	16320581	3	90458
	5	5_22X0	16881793	17411466	4	529673
	5	5_23X0	17475008	17764228	3	289220
	5	5_24X0	17777750	17798765	2	21015
	5	5_25X0	18220154	18515015	8	294861
	5	5_26X0	19584507	19682216	2	97709
	5	5_27X0	36555486	36589395	2	33909
	5	5_28X0	39016773	39576733	11	559960
	5	5_28X1	39593734	40331781	19	738047
	5	5_28X2	37374982	38242276	13	867294
	5	5_28X3	38269294	38974936	13	705642
	5	5_29X0	40576104	40809220	6	233116
	5	5_2X0	2248075	2447777	2	199702
	5	5_30X0	41052615	41329922	4	277307
	5	5_31X0	46711565	46739388	2	27823
	5	5_32X0	49189098	49210025	2	20927
	5	5_33X0	49268610	49972087	15	703477

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	5	5_34X0	50316483	50401204	2	84721
	5	5_35X0	56965953	57257104	6	291151
	5	5_36X0	58896827	59146308	3	249481
	5	5_37X0	60054082	60211077	3	156995
	5	5_38X0	65700254	66100753	4	400499
	5	5_39X0	93487749	93788740	3	300991
	5	5_3X0	2734761	2867743	2	132982
	5	5_4X0	5467620	5488372	2	20752
	5	5_5X0	6693182	7052066	9	358884
	5	5_6X0	7130482	7414174	7	283692
	5	5_7X0	7513910	7561201	2	47291
	5	5_8X0	7753295	8381209	8	627914
	5	5_9X0	8688503	8953241	4	264738
	6	6_10X0	84116959	84839654	11	722695
	6	6_11X0	84886613	85009405	2	122792
	6	6_12X0	85089487	85195356	2	105869
	6	6_13X0	88136611	88256794	4	120183
	6	6_14X0	88474889	88746948	7	272059
	6	6_15X0	88912478	89414148	4	501670
	6	6_16X0	90526788	90571258	3	44470
	6	6_17X0	91541615	91663290	2	121675
	6	6_18X0	96762530	96833647	2	71117
	6	6_19X0	97495356	97570005	2	74649
	6	6_1X0	2564603	2596986	2	32383
	6	6_20X0	101704087	101893341	2	189254
	6	6_21X0	102188189	102311822	2	123633
	6	6_22X0	112285626	112521495	4	235869
	6	6_2X0	21899540	22040542	2	141002
	6	6_3X0	25339921	25489934	3	150013
	6	6_4X0	36005466	36306670	3	301204
	6	6_5X0	57991535	58037166	3	45631
	6	6_6X0	66735049	66924657	2	189608
	6	6_7X0	67463055	67504063	2	41008
	6	6_8X0	72195008	72347656	2	152648
	6	6_9X0	83395525	83926260	7	530735
	7	7_10X0	23285953	23767364	15	481411
	7	7_11X0	23994539	24657202	27	662663
	7	7_11X1	24692680	25196145	19	503465
	7	7_12X0	32810619	32910333	2	99714
	7	7_13X0	33947472	33999790	2	52318
	7	7_14X0	34254112	34572784	4	318672
	7	7_15X0	39795308	39851596	2	56288
	7	7_16X0	40308578	40504372	2	195794
	7	7_17X0	55457452	55598875	2	141423
	7	7_18X0	56789155	56845482	2	56327
	7	7_19X0	62233421	62300717	4	67296
	7	7_1X0	5243233	5490813	2	247580
	7	7_20X0	68194988	68260633	2	65645
	7	7_21X0	69617164	69805671	3	188507
	7	7_22X0	73551879	73584536	2	32657
	7	7_23X0	76438390	76480520	2	42130
	7	7_24X0	79039070	79109447	3	70377
	7	7_25X0	81886695	81984430	3	97735
	7	7_26X0	83066587	83275670	2	209083
	7	7_27X0	83397872	83654245	3	256373

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	7	7_28X0	83681822	83877586	4	195764
	7	7_2X0	7644832	7840879	3	196047
	7	7_3X0	9894376	9969339	2	74963
	7	7_4X0	12049147	12123534	2	74387
	7	7_5X0	18953504	18984672	2	31168
	7	7_6X0	19988579	20139147	2	150568
	7	7_7X0	20353214	20393188	2	39974
	7	7_8X0	20452478	20718944	6	266466
	7	7_9X0	21870715	22843871	62	973156
	8	8_10X0	73866471	73892672	2	26201
	8	8_11X0	82967383	83231704	3	264321
	8	8_12X0	88547398	88563818	2	16420
	8	8_1X0	21064326	21114184	2	49858
	8	8_2X0	21215915	21263198	2	47283
	8	8_3X0	26261261	26398457	2	137196
	8	8_4X0	47992889	48094878	2	101989
	8	8_5X0	52785999	52879447	2	93448
	8	8_6X0	56800675	57018342	2	217667
	8	8_7X0	57670167	57862663	13	192496
	8	8_8X0	57886671	57962047	2	75376
	8	8_9X0	62006022	62144778	3	138756
	9	9_10X0	50990373	51183564	2	193191
	9	9_11X0	57354087	57593537	5	239450
	9	9_12X0	89631198	89894106	4	262908
	9	9_13X0	90028854	90069351	2	40497
	9	9_1X0	542456	553613	2	11157
	9	9_2X0	3960965	4151747	2	190782
	9	9_3X0	13264571	13559172	3	294601
	9	9_4X0	13595454	13986027	5	390573
	9	9_5X0	14292401	14656183	11	363782
	9	9_6X0	33682501	33877561	2	195060
	9	9_7X0	34449585	34587699	2	138114
	9	9_8X0	36031532	36133062	2	101530
	9	9_9X0	36264636	36324915	2	60279
	X	X_10X0	27392853	27486505	2	93652
	X	X_11X0	31759639	31841708	2	82069
	X	X_12X0	33096792	33282762	3	185970
	X	X_13X0	34755314	34776472	2	21158
	X	X_14X0	38780731	38803527	2	22796
	X	X_15X0	40680225	40812225	2	132000
	X	X_16X0	43187759	43223228	2	35469
	X	X_17X0	43337196	43520388	2	183192
	X	X_18X0	46457640	46569706	3	112066
	X	X_19X0	50712450	50878327	2	165877
	X	X_1X0	481198	608881	4	127683
	X	X_20X0	52680779	52700821	2	20042
	X	X_21X0	53363457	53468846	3	105389
	X	X_22X0	53626874	54173336	6	546462
	X	X_23X0	54609055	54686142	2	77087
	X	X_24X0	59512247	59705026	3	192779
	X	X_25X0	66363570	66814725	3	451155
	X	X_26X0	68992686	69105773	2	113087
	X	X_27X0	76387463	76554434	2	166971
	X	X_28X0	76735588	77147108	4	411520
	X	X_29X0	77461651	77732189	3	270538

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	X	X_2X0	621541	773215	2	151674
	X	X_30X0	77957662	78020687	2	63025
	X	X_31X0	78098141	78139835	2	41694
	X	X_32X0	78941075	79083946	2	142871
	X	X_33X0	80062469	80194856	2	132387
	X	X_34X0	80775148	81272063	13	496915
	X	X_35X0	85134097	85307009	2	172912
	X	X_36X0	90178637	90265695	2	87058
	X	X_37X0	94749082	94792273	2	43191
	X	X_38X0	109342823	109391313	2	48490
	X	X_39X0	112732368	112738904	2	6536
	X	X_3X0	1210289	1257152	2	46863
	X	X_40X0	121827573	121855537	2	27964
	X	X_41X0	121919777	122074995	4	155218
	X	X_42X0	122175023	122221841	3	46818
	X	X_43X0	122610814	122617371	2	6557
	X	X_44X0	125570541	125820620	8	250079
	X	X_45X0	126210243	126226543	2	16300
	X	X_46X0	126315607	126412968	2	97361
	X	X_47X0	126480122	126517417	2	37295
	X	X_4X0	1295130	1362876	3	67746
	X	X_5X0	2375775	2956273	10	580498
	X	X_6X0	10344925	10400506	2	55581
	X	X_7X0	12709059	12782477	2	73418
	X	X_8X0	21580141	21621570	2	41429
	X	X_9X0	24886477	24913745	3	27268
	1	1_10X0	15998618	16339673	5	341055
	1	1_11X0	18273546	18307509	2	33963
	1	1_12X0	18925890	19192839	5	266949
	1	1_13X0	19739282	19812299	2	73017
	1	1_14X0	21083636	21119312	2	35676
	1	1_15X0	21764852	21793705	3	28853
	1	1_16X0	23302259	23369700	2	67441
	1	1_17X0	31587441	31677879	2	90438
	1	1_18X0	34074526	34079254	2	4728
	1	1_19X0	39861002	39961487	3	100485
	1	1_1X0	647192	728465	3	81273
	1	1_20X0	41602904	41614377	2	11473
	1	1_21X0	43630433	43652334	2	21901
	1	1_22X0	46635116	46742538	3	107422
	1	1_23X0	47012656	47035956	2	23300
	1	1_24X0	78742962	78786954	2	43992
	1	1_25X0	86845557	87070510	5	224953
	1	1_26X0	89493552	89871149	8	377597
	1	1_27X0	104337711	104433312	3	95601
	1	1_28X0	110051287	110122361	2	71074
	1	1_29X0	116148944	116380700	6	231756
	1	1_2X0	5847950	5870860	2	22910
	1	1_30X0	118767560	118868237	2	100677
	1	1_31X0	124685605	124897634	4	212029
	1	1_32X0	126422634	126481817	3	59183
	1	1_33X0	127723571	127800906	2	77335
	1	1_34X0	127801241	127802180	2	939
	1	1_35X0	128685401	128756413	2	71012
	1	1_36X0	128980418	129010294	2	29876

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_37X0	131363006	131414330	2	51324
	1	1_38X0	133106218	133138450	2	32232
	1	1_39X0	133150474	133222622	3	72148
	1	1_3X0	8334190	8418743	2	84553
	1	1_40X0	133302235	133331478	2	29243
	1	1_41X0	134968297	135009014	2	40717
	1	1_42X0	135219212	135265866	2	46654
	1	1_43X0	135639501	135725135	2	85634
	1	1_44X0	136316244	136491239	5	174995
	1	1_45X0	136537901	136951422	13	413521
	1	1_46X0	136984022	137209798	4	225776
	1	1_47X0	137937943	138095369	5	157426
	1	1_48X0	138264158	138330249	2	66091
	1	1_49X0	139020663	139088329	2	67666
	1	1_4X0	11122124	11148803	3	26679
	1	1_50X0	139155332	139255746	2	100414
	1	1_51X0	139668135	139812421	3	144286
	1	1_52X0	146803296	146845429	2	42133
	1	1_53X0	147965143	148000239	2	35096
	1	1_54X0	149396317	149648346	5	252029
	1	1_55X0	158361604	158435250	2	73646
	1	1_56X0	180205607	180267345	2	61738
	1	1_57X0	182051362	182101503	2	50141
	1	1_58X0	182221151	182391840	3	170689
	1	1_59X0	183855217	183960763	2	105546
	1	1_5X0	11320533	11333860	2	13327
	1	1_60X0	185522631	185619338	2	96707
	1	1_61X0	185654525	185756061	3	101536
	1	1_62X0	185792625	185859112	3	66487
	1	1_63X0	204889857	204944933	2	55076
	1	1_64X0	207443047	207547406	7	104359
	1	1_65X0	226525542	226904111	6	378569
	1	1_66X0	227026663	227811623	23	784960
	1	1_66X1	227845394	228300564	11	455170
	1	1_67X0	228329902	228369681	2	39779
	1	1_68X0	228465323	228476376	2	11053
	1	1_6X0	12135046	12218882	2	83836
	1	1_7X0	12263786	12522402	8	258616
	1	1_8X0	14586224	14638962	3	52738
	1	1_9X0	15187245	15242400	2	55155
	10	10_10X0	81557395	81588054	2	30659
	10	10_11X0	85485513	85506450	2	20937
	10	10_12X0	88294391	88370980	2	76589
	10	10_13X0	89820181	90055837	5	235656
	10	10_14X0	90540909	90652915	5	112006
	10	10_15X0	94438830	94455470	2	16640
	10	10_16X0	98672890	98687993	2	15103
	10	10_17X0	104929774	104963042	3	33268
	10	10_18X0	105747602	105791880	2	44278
	10	10_19X0	124693964	124794448	4	100484
	10	10_1X0	4986183	5482896	6	496713
	10	10_20X0	124977654	124992288	2	14634
	10	10_2X0	5686091	5712583	2	26492
	10	10_3X0	38989431	39285979	4	296548
	10	10_4X0	39558827	39590759	2	31932

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	10	10_5X0	45299796	45479410	4	179614
	10	10_6X0	48032874	48061307	2	28433
	10	10_7X0	69967100	70047879	2	80779
	10	10_8X0	76253600	76315561	2	61961
	10	10_9X0	81375204	81439293	2	64089
	11	11_10X0	95072533	95143368	5	70835
	11	11_11X0	102574902	102622290	2	47388
	11	11_12X0	102765624	103205734	9	440110
	11	11_13X0	105159922	105343870	4	183948
	11	11_14X0	112297993	112303440	2	5447
	11	11_15X0	114327300	114393235	2	65935
	11	11_16X0	117212078	117259693	3	47615
	11	11_17X0	118262975	118285540	2	22565
	11	11_18X0	118582068	118713023	5	130955
	11	11_19X0	118753397	118788381	2	34984
	11	11_1X0	168132	191180	4	23048
	11	11_20X0	124265210	125084931	21	819721
	11	11_21X0	125262472	125456473	4	194001
	11	11_22X0	130969294	131041179	2	71885
	11	11_2X0	421677	478356	2	56679
	11	11_3X0	786256	820168	2	33912
	11	11_4X0	1834136	1925352	2	91216
	11	11_5X0	2143445	2175544	2	32099
	11	11_6X0	5492134	6070052	21	577918
	11	11_6X1	4173353	5037048	28	863695
	11	11_6X10	56725503	57276625	7	551122
	11	11_6X11	54529655	55327181	9	797526
	11	11_6X12	63476449	63850550	6	374101
	11	11_6X13	17915966	18096673	6	180707
	11	11_6X14	49195798	49400399	6	204601
	11	11_6X15	62149022	62389865	6	240843
	11	11_6X16	71373631	71775228	4	401597
	11	11_6X17	7582010	7744600	4	162590
	11	11_6X18	64589047	64878362	4	289315
	11	11_6X19	67553373	67841237	4	287864
	11	11_6X2	55978340	56517719	18	539379
	11	11_6X20	3228208	3577838	4	349630
	11	11_6X21	49700038	49801578	3	101540
	11	11_6X22	69593355	69714319	3	120964
	11	11_6X23	75091019	75157393	3	66374
	11	11_6X24	73953405	74000904	2	47499
	11	11_6X25	73172332	73235985	2	63653
	11	11_6X26	65587821	65676993	2	89172
	11	11_6X27	62979047	63018431	2	39384
	11	11_6X28	66726933	66764354	2	37421
	11	11_6X29	68854102	68887588	2	33486
	11	11_6X3	5118255	5476417	8	358162
	11	11_6X4	55455009	55902195	16	447186
	11	11_6X5	58885988	59673577	15	787589
	11	11_6X6	57788776	58289218	11	500442
	11	11_6X7	6576393	6788215	8	211822
	11	11_6X8	48516076	48781211	8	265135
	11	11_6X9	71790737	72139412	5	348675
	11	11_7X0	75740701	75818134	2	77433
	11	11_8X0	85798484	85827270	2	28786

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	11	11_9X0	89931973	90039694	5	107721
	12	12_10X0	13481966	13530521	2	48555
	12	12_11X0	15408144	15425816	2	17672
	12	12_12X0	32358079	32385005	2	26926
	12	12_13X0	32858939	32895295	2	36356
	12	12_14X0	33693453	33699866	2	6413
	12	12_15X0	34290985	34598440	8	307455
	12	12_16X0	35094431	35111908	2	17477
	12	12_17X0	35721427	35838433	9	117006
	12	12_18X0	36288487	36363510	2	75023
	12	12_19X0	36859927	37669638	26	809711
	12	12_1X0	163961	243334	2	79373
	12	12_20X0	37844275	37929489	2	85214
	12	12_21X0	39923337	39947350	3	24013
	12	12_22X0	40636848	40798077	3	161229
	12	12_23X0	40948297	40964590	2	16293
	12	12_24X0	70308121	70359848	2	51727
	12	12_25X0	83126849	83138886	2	12037
	12	12_26X0	93607644	93755632	3	147988
	12	12_27X0	116040520	116145073	3	104553
	12	12_28X0	116518453	116564059	2	45606
	12	12_29X0	126114923	126167011	3	52088
	12	12_2X0	4569798	4646834	2	77036
	12	12_30X0	136809954	137095616	7	285662
	12	12_3X0	7091490	7134817	2	43327
	12	12_4X0	7388619	7434661	3	46042
	12	12_5X0	8553860	8559774	2	5914
	12	12_6X0	8820861	8919794	3	98933
	12	12_7X0	10021589	10905355	17	883766
	12	12_8X0	11284174	11718019	19	433845
	12	12_9X0	11835986	11901420	2	65434
	13	13_1X0	4699951	4787541	3	87590
	13	13_2X0	15654304	15723804	2	69500
	13	13_3X0	25971685	26047924	2	76239
	13	13_4X0	37125177	37216381	2	91204
	13	13_5X0	84949733	84989522	2	39789
	13	13_6X0	85677841	85701572	2	23731
	13	13_7X0	99080715	99143386	3	62671
	14	14_10X0	8193549	8323295	3	129746
	14	14_11X0	20491526	20557149	2	65623
	14	14_12X0	36401473	36462778	2	61305
	14	14_13X0	48838489	48878515	2	40026
	14	14_14X0	52064955	52121051	2	56096
	14	14_15X0	54710124	55003002	4	292878
	14	14_16X0	58084858	58137037	3	52179
	14	14_17X0	79448239	79780003	9	331764
	14	14_18X0	80850345	80878087	2	27742
	14	14_19X0	81400669	81464803	2	64134
	14	14_1X0	3611710	3978392	10	366682
	14	14_20X0	85513215	85563970	2	50755
	14	14_21X0	88409553	88447764	2	38211
	14	14_22X0	91735078	91800579	3	65501
	14	14_2X0	4286800	4759827	12	473027
	14	14_3X0	5285366	5399835	4	114469
	14	14_4X0	5808963	5902876	2	93913

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	14	14_5X0	7130714	7183198	2	52484
	14	14_6X0	7815346	7826292	2	10946
	14	14_7X0	7900390	7904248	2	3858
	14	14_8X0	8002714	8008842	2	6128
	14	14_9X0	8091015	8133357	2	42342
	15	15_10X0	55781258	55844533	2	63275
	15	15_11X0	59587028	59648901	3	61873
	15	15_12X0	69329966	69380823	2	50857
	15	15_13X0	70524258	70553789	2	29531
	15	15_14X0	82898995	83063960	4	164965
	15	15_1X0	4115808	4201632	2	85824
	15	15_2X0	9298169	9368059	2	69890
	15	15_3X0	10968361	11037213	2	68852
	15	15_4X0	22045402	22119261	2	73859
	15	15_5X0	23783770	23888568	3	104798
	15	15_6X0	24182941	24285880	2	102939
	15	15_7X0	25642355	25714338	4	71983
	15	15_8X0	46185988	46282493	2	96505
	15	15_9X0	55716573	55749924	2	33351
	16	16_10X0	23107892	23142150	2	34258
	16	16_11X0	28309149	28326596	2	17447
	16	16_12X0	29298766	29686209	9	387443
	16	16_13X0	29794377	29809188	2	14811
	16	16_14X0	29891261	29939218	2	47957
	16	16_15X0	29965704	29988358	2	22654
	16	16_16X0	29988408	30053462	3	65054
	16	16_17X0	46458178	46530955	2	72777
	16	16_18X0	47236670	47349758	11	113088
	16	16_19X0	48036737	48095968	3	59231
	16	16_1X0	171289	197819	6	26530
	16	16_20X0	48224884	48371607	3	146723
	16	16_21X0	57386318	57515118	4	128800
	16	16_22X0	57751012	57804820	3	53808
	16	16_23X0	58037995	58105113	2	67118
	16	16_24X0	58865162	58882982	2	17820
	16	16_25X0	61188160	61262451	2	74291
	16	16_26X0	62378457	62413615	2	35158
	16	16_27X0	63007394	63057815	2	50421
	16	16_28X0	66189717	66210257	2	20540
	16	16_29X0	66463903	66524221	2	60318
	16	16_2X0	296914	387668	2	90754
	16	16_30X0	66655636	66731535	3	75899
	16	16_31X0	77770712	77837289	3	66577
	16	16_3X0	713983	770884	4	56901
	16	16_4X0	858698	882371	2	23673
	16	16_5X0	1379887	1393542	2	13655
	16	16_6X0	2362813	2404845	2	42032
	16	16_7X0	2930657	3111271	4	180614
	16	16_8X0	3359730	3704246	9	344516
	16	16_9X0	19895491	19934954	2	39463
	17	17_10X0	9925189	9995771	2	70582
	17	17_11X0	12370920	12392065	2	21145
	17	17_12X0	13656197	13668930	2	12733
	17	17_13X0	13709148	13791331	2	82183
	17	17_14X0	14691428	14721909	3	30481

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	17	17_15X0	14897360	14954361	2	57001
	17	17_16X0	15741216	15764783	2	23567
	17	17_17X0	16002415	16280727	17	278312
	17	17_18X0	16320038	16564642	21	244604
	17	17_19X0	16582223	16595984	3	13761
	17	17_1X0	1744688	1777588	2	32900
	17	17_20X0	16605504	16944661	12	339157
	17	17_21X0	17523085	17561069	2	37984
	17	17_22X0	17645381	17722277	2	76896
	17	17_23X0	19489490	19506467	2	16977
	17	17_24X0	21090900	21317816	7	226916
	17	17_25X0	21629528	21934663	5	305135
	17	17_26X0	22867721	22970353	5	102632
	17	17_27X0	37546087	37635898	2	89811
	17	17_28X0	37895961	38002062	2	106101
	17	17_29X0	38051223	38067665	2	16442
	17	17_2X0	3168889	3581483	11	412594
	17	17_30X0	38871111	38973682	2	102571
	17	17_31X0	40798135	40813916	2	15781
	17	17_32X0	45987866	46346181	6	358315
	17	17_33X0	48559980	48639415	3	79435
	17	17_34X0	49083057	49109495	2	26438
	17	17_35X0	49231926	49269898	2	37972
	17	17_36X0	49500861	49587835	2	86974
	17	17_37X0	49627326	49660406	2	33080
	17	17_38X0	57476153	57565089	3	88936
	17	17_39X0	60753505	60839015	2	85510
	17	17_3X0	3658611	3736494	2	77883
	17	17_40X0	63357857	63375951	2	18094
	17	17_41X0	69550221	69602696	2	52475
	17	17_42X0	73952029	74190740	7	238711
	17	17_43X0	74409068	74435528	2	26460
	17	17_44X0	75432886	75455760	2	22874
	17	17_45X0	79419375	79466780	2	47405
	17	17_4X0	4631162	4730780	2	99618
	17	17_5X0	5323263	5415695	3	92432
	17	17_6X0	7289782	7298506	2	8724
	17	17_7X0	7329083	7431709	3	102626
	17	17_8X0	7682194	7694476	3	12282
	17	17_9X0	9096253	9302482	10	206229
	18	18_1X0	5625999	5674110	2	48111
	18	18_2X0	14749757	14776153	2	26396
	18	18_3X0	28144506	28628723	6	484217
	18	18_4X0	32377006	32508596	4	131590
	18	18_5X0	44227211	44234552	2	7341
	18	18_6X0	60956517	61478675	10	522158
	18	18_7X0	72232581	72317671	2	85090
	19	19_10X0	8077793	8160683	5	82890
	19	19_11X0	59677825	60354721	25	676896
	19	19_11X1	57028992	57546150	17	517158
	19	19_11X10	45875868	46201142	10	325274
	19	19_11X11	57893843	58085694	9	191851
	19	19_11X12	46478674	46754721	6	276047
	19	19_11X13	60740041	60966671	5	226630
	19	19_11X14	12214893	12614094	8	399201

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	19	19_11X15	46207927	46451347	6	243420
	19	19_11X16	60365835	60717594	9	351759
	19	19_11X17	56871279	56999537	5	128258
	19	19_11X18	37238598	37397396	6	158798
	19	19_11X19	42870316	43209147	8	338831
	19	19_11X2	53047628	53373226	14	325598
	19	19_11X20	39238432	39681639	10	443207
	19	19_11X21	13090188	13361180	6	270992
	19	19_11X22	12852952	13064379	4	211427
	19	19_11X23	58893045	59354095	8	461050
	19	19_11X24	9928748	10296260	8	367512
	19	19_11X25	38876035	39149601	7	273566
	19	19_11X26	54547593	54900638	8	353045
	19	19_11X27	38558623	38811603	6	252980
	19	19_11X28	58611474	58888247	5	276773
	19	19_11X29	44816295	45054131	6	237836
	19	19_11X3	56574515	56740992	9	166477
	19	19_11X30	37766443	38068262	7	301819
	19	19_11X31	46764826	47025119	6	260293
	19	19_11X32	16314362	16863061	9	548699
	19	19_11X33	15753774	15890695	6	136921
	19	19_11X34	53828067	54122503	5	294436
	19	19_11X35	55557808	55899249	6	341441
	19	19_11X36	21061685	21348085	4	286400
	19	19_11X37	43543165	43728190	5	185025
	19	19_11X38	20792666	21021265	4	228599
	19	19_11X39	58149301	58449221	5	299920
	19	19_11X4	12640533	12846574	8	206041
	19	19_11X40	45100264	45306551	4	206287
	19	19_11X41	56259462	56459761	5	200299
	19	19_11X42	38136485	38408362	5	271877
	19	19_11X43	36878466	37100695	5	222229
	19	19_11X44	42002201	42251000	4	248799
	19	19_11X45	36483840	36720356	4	236516
	19	19_11X46	41547150	41630905	4	83755
	19	19_11X47	44464849	44722348	4	257499
	19	19_11X48	21831711	22272766	5	441055
	19	19_11X49	10801903	10875064	4	73161
	19	19_11X5	53433769	53748116	10	314347
	19	19_11X50	48174438	48325479	4	151041
	19	19_11X51	42269680	42496534	3	226854
	19	19_11X52	23008285	23367420	4	359135
	19	19_11X53	40762205	40815204	3	52999
	19	19_11X54	18164130	18188625	3	24495
	19	19_11X55	51237009	51276515	3	39506
	19	19_11X56	24642139	24747228	2	105089
	19	19_11X57	50883456	50931652	2	48196
	19	19_11X58	20315410	20366466	2	51056
	19	19_11X59	49443674	49465737	2	22063
	19	19_11X6	9627145	9898312	10	271167
	19	19_11X60	49973644	50001090	2	27446
	19	19_11X61	52185892	52209297	2	23405
	19	19_11X62	47703972	47795468	2	91496
	19	19_11X63	52508529	52540733	2	32204
	19	19_11X64	19482782	19507641	2	24859

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	19	19_11X65	45551250	45620127	2	68877
	19	19_11X7	54132736	54436282	9	303546
	19	19_11X8	60999109	61295049	9	295940
	19	19_11X9	55019245	55506557	13	487312
	19	19_1X0	775851	810568	4	34717
	19	19_2X0	1989840	2084396	2	94556
	19	19_3X0	2931483	3048208	5	116725
	19	19_4X0	3112641	3186957	2	74316
	19	19_5X0	3236001	3306582	2	70581
	19	19_6X0	3886121	3925161	2	39040
	19	19_7X0	6085501	6100902	2	15401
	19	19_8X0	6705462	6741097	2	35635
	19	19_9X0	7369925	7414209	3	44284
	20	20_10X0	39862656	39923715	2	61059
	20	20_11X0	46778887	46935097	6	156210
	20	20_12X0	47189974	47420840	4	230866
	20	20_13X0	47556996	47583295	2	26299
	20	20_14X0	58496947	58509102	2	12155
	20	20_15X0	65702795	65726706	2	23911
	20	20_16X0	66289827	66344805	2	54978
	20	20_1X0	10455	154450	4	143995
	20	20_2X0	1515652	1569818	2	54166
	20	20_3X0	2927405	2941286	2	13881
	20	20_4X0	4583159	4615855	2	32696
	20	20_5X0	23181020	23221913	2	40893
	20	20_6X0	23564438	24046168	8	481730
	20	20_7X0	32625998	32838623	7	212625
	20	20_8X0	34379389	34697581	5	318192
	20	20_9X0	37093070	37180201	2	87131
	21	21_10X0	32826703	32848099	2	21396
	21	21_1X0	16559613	16610666	2	51053
	21	21_2X0	16677126	16800046	6	122920
	21	21_3X0	16857088	16994454	10	137366
	21	21_4X0	19466829	19510994	2	44165
	21	21_5X0	19687325	19871935	4	184610
	21	21_6X0	20816225	20904744	2	88519
	21	21_7X0	22532920	22608991	2	76071
	21	21_8X0	28849754	28899647	3	49893
	21	21_9X0	31104293	31258412	11	154119
	22	22_10X0	25582156	25678103	5	95947
	22	22_11X0	27895601	27951843	2	56242
	22	22_12X0	28813257	28845490	2	32233
	22	22_13X0	30647092	30716537	2	69445
	22	22_14X0	37237484	37254474	2	16990
	22	22_1X0	7019640	7059743	2	40103
	22	22_2X0	8644820	8957022	5	312202
	22	22_3X0	9145534	9235441	2	89907
	22	22_4X0	9871254	9898027	2	26773
	22	22_5X0	10302170	10388521	4	86351
	22	22_6X0	16982817	17126896	4	144079
	22	22_7X0	22250443	22324222	2	73779
	22	22_8X0	22733817	22795723	2	61906
	22	22_9X0	23549608	23563874	2	14266
	2A	2A_10X0	89988142	90053703	2	65561
	2A	2A_11X0	94202730	94276040	3	73310

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	2A	2A_12X0	95917622	95960686	2	43064
	2A	2A_13X0	95967627	95986216	2	18589
	2A	2A_14X0	98147248	98200608	2	53360
	2A	2A_15X0	101022400	101473575	6	451175
	2A	2A_16X0	110883072	111175353	7	292281
	2A	2A_1X0	24586523	24653865	2	67342
	2A	2A_2X0	44802583	44876205	2	73622
	2A	2A_3X0	45960270	46027314	2	67044
	2A	2A_4X0	70349743	70385529	2	35786
	2A	2A_5X0	72276645	72303245	2	26600
	2A	2A_6X0	75988845	76008098	2	19253
	2A	2A_7X0	80656020	80784045	5	128025
	2A	2A_8X0	87377070	87413406	2	36336
	2A	2A_9X0	88499914	88575771	2	75857
	2B	2B_10X0	110415458	110490944	2	75486
	2B	2B_11X0	123450272	123527743	3	77471
	2B	2B_12X0	123593209	123612616	2	19407
	2B	2B_13X0	124812030	124823764	2	11734
	2B	2B_14X0	131847427	131871776	2	24349
	2B	2B_1X0	41051264	41093855	2	42591
	2B	2B_2X0	62359354	62376540	2	17186
	2B	2B_3X0	66384681	66501144	10	116463
	2B	2B_4X0	67861807	67931761	2	69954
	2B	2B_5X0	91746115	91889104	3	142989
	2B	2B_6X0	107562064	107597929	2	35865
	2B	2B_7X0	109040604	109166292	3	125688
	2B	2B_8X0	109774406	109808872	2	34466
	2B	2B_9X0	110169547	110190875	2	21328
	3	3_10X0	52883950	53010273	5	126323
	3	3_11X0	53076811	53088026	2	11215
	3	3_12X0	53906769	53960022	3	53253
	3	3_13X0	100774166	101195375	6	421209
	3	3_14X0	104386635	104464552	4	77917
	3	3_15X0	112314696	112379963	2	65267
	3	3_16X0	134657347	134679410	2	22063
	3	3_17X0	137145070	137224865	2	79795
	3	3_18X0	142431666	142470848	3	39182
	3	3_19X0	142883553	142970170	2	86617
	3	3_1X0	10156457	10187860	2	31403
	3	3_20X0	150865762	150889576	2	23814
	3	3_21X0	154723401	154867974	4	144573
	3	3_22X0	155284591	155354017	2	69426
	3	3_23X0	162264024	162330290	2	66266
	3	3_24X0	173481183	173557744	3	76561
	3	3_25X0	188008679	188081995	3	73316
	3	3_26X0	190726075	190873007	3	146932
	3	3_27X0	198693190	198750549	3	57359
	3	3_2X0	39018355	39071179	2	52824
	3	3_3X0	41331418	41389166	3	57748
	3	3_4X0	43558572	43598286	2	39714
	3	3_5X0	45338865	45695207	7	356342
	3	3_6X0	46889296	47416676	8	527380
	3	3_7X0	47713915	47746400	3	32485
	3	3_8X0	51277982	51343728	2	65746
	3	3_9X0	51378883	51407326	3	28443

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	4	4_10X0	60683417	60716902	2	33485
	4	4_11X0	60792026	60929373	3	137347
	4	4_12X0	61027189	61905306	8	878117
	4	4_13X0	61946152	62601803	5	655651
	4	4_14X0	83167699	83246862	2	79163
	4	4_15X0	89661064	89759442	2	98378
	4	4_16X0	101733174	102135053	6	401879
	4	4_17X0	158447509	158497330	2	49821
	4	4_18X0	167296607	167324240	2	27633
	4	4_19X0	169050920	169071526	2	20606
	4	4_1X0	7108159	7176836	3	68677
	4	4_20X0	187645434	187648304	2	2870
	4	4_21X0	190578681	190635370	2	56689
	4	4_22X0	192526984	192587020	2	60036
	4	4_2X0	9567311	9618691	3	51380
	4	4_3X0	16238435	16263982	2	25547
	4	4_4X0	48010006	48102967	2	92961
	4	4_5X0	54467186	54507339	3	40153
	4	4_6X0	56404297	56772593	9	368296
	4	4_7X0	57007439	57108155	3	100716
	4	4_8X0	60225927	60274055	3	48128
	4	4_9X0	60578689	60650615	3	71926
	5	5_1X0	44942130	44955814	2	13684
	5	5_2X0	59534710	59620936	2	86226
	5	5_3X0	132241291	132342877	2	101586
	5	5_4X0	140690501	140728954	2	38453
	5	5_5X0	140996918	141999312	20	1002394
	5	5_6X0	150269907	150335419	2	65512
	5	5_7X0	157315877	157395026	2	79149
	5	5_8X0	179245586	179632423	6	386837
	5	5_9X0	181508648	181861570	8	352922
	6	6_10X0	28304570	28827144	14	522574
	6	6_11X0	29284994	29816955	17	531961
	6	6_12X0	29948688	30224359	7	275671
	6	6_13X0	30318602	30571698	6	253096
	6	6_14X0	31603516	31663286	2	59770
	6	6_15X0	31718744	31740453	3	21709
	6	6_16X0	31829129	31880936	3	51807
	6	6_17X0	31968676	31985234	2	16558
	6	6_18X0	32152402	32244267	7	91865
	6	6_19X0	32644829	33639492	20	994663
	6	6_1X0	2730688	2863394	3	132706
	6	6_20X0	36127860	36148060	3	20200
	6	6_21X0	41617310	41814642	6	197332
	6	6_22X0	42635039	42656388	2	21349
	6	6_23X0	46655347	46696502	2	41155
	6	6_24X0	48700133	48766329	2	66196
	6	6_25X0	50488743	50590985	5	102242
	6	6_26X0	52785928	52846348	2	60420
	6	6_27X0	53368179	53641341	6	273162
	6	6_28X0	75899442	76071208	6	171766
	6	6_29X0	108535850	108614656	2	78806
	6	6_2X0	3102121	3211762	3	109641
	6	6_30X0	119880337	119904528	2	24191
	6	6_31X0	120085801	120184016	3	98215

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	6	6_32X0	136444426	136552358	6	107932
	6	6_33X0	153987318	154151873	6	164555
	6	6_34X0	171671777	171718240	2	46463
	6	6_3X0	26133234	26294829	4	161595
	6	6_4X0	26381168	26657336	20	276168
	6	6_5X0	26737399	26878507	7	141108
	6	6_6X0	27349694	27364857	4	15163
	6	6_7X0	27580090	27681918	3	101828
	6	6_8X0	28035724	28123561	14	87837
	6	6_9X0	28189655	28293643	2	103988
	7	7_10X0	64146248	64480039	4	333791
	7	7_11X0	64970423	65037350	2	66927
	7	7_12X0	75439427	75499975	2	60548
	7	7_13X0	77711028	77742666	2	31638
	7	7_14X0	78010484	78054625	2	44141
	7	7_15X0	93023518	93107702	2	84184
	7	7_16X0	96780844	96800640	2	19796
	7	7_17X0	98179250	98309312	3	130062
	7	7_18X0	99876508	99895531	2	19023
	7	7_19X0	102262887	102324222	2	61335
	7	7_1X0	5166096	5218644	2	52548
	7	7_20X0	102430804	102601315	4	170511
	7	7_21X0	102619499	102915633	4	296134
	7	7_22X0	103072302	103134524	3	62222
	7	7_23X0	128395714	128497358	2	101644
	7	7_24X0	134838518	134933199	3	94681
	7	7_25X0	139000122	139135693	3	135571
	7	7_26X0	146377229	146577811	4	200582
	7	7_27X0	147405157	147460996	6	55839
	7	7_28X0	147557627	147619473	2	61846
	7	7_29X0	147711872	147738857	2	26985
	7	7_2X0	15861483	15953521	2	92038
	7	7_30X0	148358014	148427504	2	69490
	7	7_31X0	148486100	148779803	8	293703
	7	7_32X0	153579445	154256920	9	677475
	7	7_33X0	154869898	155143304	6	273406
	7	7_34X0	155191326	155204890	2	13564
	7	7_3X0	18182145	18210764	2	28619
	7	7_4X0	26275053	26422541	10	147488
	7	7_5X0	43452372	43551678	4	99306
	7	7_6X0	45327920	45361496	2	33576
	7	7_7X0	48266038	48345833	2	79795
	7	7_8X0	56203457	56211846	2	8389
	7	7_9X0	58161514	58169882	2	8368
	8	8_10X0	147132561	147812293	7	679732
	8	8_1X0	7289685	7376933	3	87248
	8	8_2X0	9865500	9923135	2	57635
	8	8_3X0	11362623	11376556	2	13933
	8	8_4X0	23079993	23105388	2	25395
	8	8_5X0	42424293	42494350	2	70057
	8	8_6X0	63926666	63933462	2	6796
	8	8_7X0	87161506	87287927	13	126421
	8	8_8X0	145396778	145959785	12	563007
	8	8_9X0	145992574	146040816	3	48242
	9	9_10X0	82331848	82550428	7	218580

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	9	9_11X0	82575365	82600895	2	25530
	9	9_12X0	86722533	86731519	2	8986
	9	9_13X0	91115007	91221361	2	106354
	9	9_14X0	91232546	91300673	2	68127
	9	9_15X0	92278136	92308702	2	30566
	9	9_16X0	100579646	100878317	12	298671
	9	9_17X0	100980377	100989516	2	9139
	9	9_18X0	102555765	102630165	2	74400
	9	9_19X0	104936518	104984970	2	48452
	9	9_1X0	5215929	5399968	4	184039
	9	9_20X0	106024532	106055389	2	30857
	9	9_21X0	108016154	108039733	2	23579
	9	9_22X0	111644473	111668682	2	24209
	9	9_23X0	113889294	113965152	4	75858
	9	9_2X0	21586404	22007623	12	421219
	9	9_3X0	35446965	35467570	2	20605
	9	9_4X0	36709890	36867021	4	157131
	9	9_5X0	63596391	63667767	2	71376
	9	9_6X0	65528404	65589846	2	61442
	9	9_7X0	70103210	70255255	4	152045
	9	9_8X0	72223709	72316910	2	93201
	9	9_9X0	74473752	74573380	2	99628
	KV420888.1	KV420888.1_1X0	2448516	2532639	2	84123
	KV420889.1	KV420889.1_1X0	1261024	1321591	2	60567
	KV420904.1	KV420904.1_1X0	91127	126604	2	35477
	KV420904.1	KV420904.1_2X0	347096	386674	2	39578
	KV420904.1	KV420904.1_3X0	530714	557699	2	26985
	KV420905.1	KV420905.1_1X0	458071	490829	2	32758
	KV420905.1	KV420905.1_2X0	584818	610676	3	25858
	KV420913.1	KV420913.1_1X0	161201	316070	11	154869
	KV420955.1	KV420955.1_1X0	1070546	1112729	3	42183
	KV420955.1	KV420955.1_2X0	1153601	1269772	3	116171
	KV420955.1	KV420955.1_3X0	1385393	1438047	2	52654
	KV420955.1	KV420955.1_4X0	1486828	1500640	2	13812
	KV420955.1	KV420955.1_5X0	1848647	1893588	2	44941
	KV420959.1	KV420959.1_1X0	836296	904711	2	68415
	KV421231.1	KV421231.1_1X0	118855	173387	3	54532
	KV421314.1	KV421314.1_1X0	186241	250201	3	63960
	KV421314.1	KV421314.1_2X0	349400	389033	3	39633
	KV421314.1	KV421314.1_3X0	414142	439063	2	24921
	KV421314.1	KV421314.1_4X0	452097	507249	3	55152
	KV421332.1	KV421332.1_1X0	420326	464499	2	44173
	KV421332.1	KV421332.1_2X0	640946	821347	4	180401
	KV421455.1	KV421455.1_1X0	119741	160964	4	41223
	X	X_10X0	51463487	51557865	3	94378
	X	X_11X0	52213279	52449838	4	236559
	X	X_12X0	52501842	52530258	2	28416
	X	X_13X0	54793212	55034824	4	241612
	X	X_14X0	56971245	56988752	2	17507
	X	X_15X0	71118202	71242271	4	124069
	X	X_16X0	100456616	100492514	2	35898
	X	X_17X0	101187883	101295376	3	107493
	X	X_18X0	101823075	101880942	2	57867
	X	X_19X0	101982655	102042765	2	60110
	X	X_1X0	12730019	12785462	2	55443

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	X	X_20X0	102197641	102352943	4	155302
	X	X_21X0	102674505	103015822	6	341317
	X	X_22X0	103216282	103273122	2	56840
	X	X_23X0	103509433	103583468	3	74035
	X	X_24X0	119006733	119073641	2	66908
	X	X_25X0	119737964	119786672	2	48708
	X	X_26X0	120518792	120528927	2	10135
	X	X_27X0	135005248	135066015	2	60767
	X	X_28X0	135131337	135205792	2	74455
	X	X_29X0	146515911	146547069	2	31158
	X	X_2X0	15134747	15206208	2	71461
	X	X_30X0	149209646	149262390	2	52744
	X	X_31X0	152437666	152468677	3	31011
	X	X_32X0	153014008	153092899	2	78891
	X	X_33X0	154001597	154177166	6	175569
	X	X_3X0	24497424	24551772	2	54348
	X	X_4X0	26333535	26356350	2	22815
	X	X_5X0	30404415	30437362	3	32947
	X	X_6X0	47577053	47673389	2	96336
	X	X_7X0	48183034	48269614	2	86580
	X	X_8X0	48315108	48487506	5	172398
	X	X_9X0	49336822	49365044	2	28222
	Y	Y_1X0	2070514	2125870	4	55356
	Y	Y_2X0	10876744	10922271	3	45527
	Y	Y_3X0	12771812	12845471	2	73659
	Y	Y_4X0	15064647	15098735	2	34088
	Y	Y_5X0	21594468	21668154	2	73686
	1	1_10X0	14534859	14566243	2	31384
	1	1_11X0	15375567	15726587	5	351020
	1	1_12X0	18263844	18522458	5	258614
	1	1_13X0	19078188	19142823	2	64635
	1	1_14X0	20910175	20932415	2	22240
	1	1_15X0	22416308	22481177	2	64869
	1	1_16X0	29905913	30043183	2	137270
	1	1_17X0	30498405	30650521	3	152116
	1	1_18X0	30737298	30889934	2	152636
	1	1_19X0	31587954	31808400	2	220446
	1	1_1X0	422417	480580	3	58163
	1	1_20X0	33085615	33090321	2	4706
	1	1_21X0	33313278	33447702	2	134424
	1	1_22X0	34164992	34409997	3	245005
	1	1_23X0	38714270	38817263	3	102993
	1	1_24X0	40434733	40446405	2	11672
	1	1_25X0	41593286	41774183	2	180897
	1	1_26X0	42552723	42575008	2	22285
	1	1_27X0	45152794	45395892	5	243098
	1	1_28X0	45697939	45721630	2	23691
	1	1_29X0	46375887	46571104	2	195217
	1	1_2X0	5423894	5447187	2	23293
	1	1_30X0	51110649	51261916	2	151267
	1	1_31X0	65485101	65674805	2	189704
	1	1_32X0	76981883	77029143	2	47260
	1	1_33X0	82082016	82202240	2	120224
	1	1_34X0	83524754	83744472	4	219718
	1	1_35X0	86079668	86399431	7	319763

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_36X0	86547346	86728064	2	180718
	1	1_37X0	101024434	101083725	2	59291
	1	1_38X0	105572644	105891999	4	319355
	1	1_39X0	106908314	106976123	2	67809
	1	1_3X0	7866190	7935066	2	68876
	1	1_40X0	107040531	107087958	2	47427
	1	1_41X0	107859925	108016214	3	156289
	1	1_42X0	108565312	108728426	4	163114
	1	1_43X0	113777760	114034839	3	257079
	1	1_44X0	114171379	114441380	3	270001
	1	1_45X0	116696577	116798305	2	101728
	1	1_46X0	118020778	118110957	2	90179
	1	1_47X0	119110723	119260275	2	149552
	1	1_48X0	120076684	120147731	7	71047
	1	1_49X0	121016562	121090476	2	73914
	1	1_4X0	8721922	8811978	2	90056
	1	1_50X0	121599971	121733794	2	133823
	1	1_51X0	122183446	122256225	2	72779
	1	1_52X0	122872117	123188605	4	316488
	1	1_53X0	123639899	123699334	2	59435
	1	1_54X0	125372376	125407148	2	34772
	1	1_55X0	125420730	125495762	3	75032
	1	1_56X0	125576895	125608432	2	31537
	1	1_57X0	126143807	126154694	2	10887
	1	1_58X0	126304462	126498025	2	193563
	1	1_59X0	127059655	127096595	2	36940
	1	1_5X0	10523313	10549747	3	26434
	1	1_60X0	127281575	127322516	2	40941
	1	1_61X0	127530515	127575814	2	45299
	1	1_62X0	127906944	128221654	5	314710
	1	1_63X0	128482822	128750271	7	267449
	1	1_64X0	128876949	129258069	4	381120
	1	1_65X0	129288889	129506146	4	217257
	1	1_66X0	129621761	129738618	2	116857
	1	1_67X0	129743693	129990273	3	246580
	1	1_68X0	130254835	130557185	7	302350
	1	1_69X0	130579579	130644617	2	65038
	1	1_6X0	10724465	10738278	2	13813
	1	1_70X0	130944033	131747405	12	803372
	1	1_71X0	131958578	132088265	4	129687
	1	1_72X0	133418748	133529038	5	110290
	1	1_73X0	133872069	133995295	2	123226
	1	1_74X0	135985688	136041712	2	56024
	1	1_75X0	138007310	138156792	2	149482
	1	1_76X0	140878444	140943815	2	65371
	1	1_77X0	152872475	153185435	3	312960
	1	1_78X0	153294509	153515289	3	220780
	1	1_79X0	153657589	153720565	3	62976
	1	1_7X0	11542311	11622563	2	80252
	1	1_80X0	153761929	153869371	3	107442
	1	1_81X0	153907503	154163640	3	256137
	1	1_82X0	154627809	154855058	2	227249
	1	1_83X0	155092077	155316534	2	224457
	1	1_84X0	157131807	157299851	3	168044
	1	1_85X0	157416824	157469077	2	52253

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_86X0	158391303	158583283	2	191980
	1	1_87X0	159289937	159345363	2	55426
	1	1_88X0	163659549	164069769	4	410220
	1	1_89X0	178468027	178687512	3	219485
	1	1_8X0	11676459	11885148	11	208689
	1	1_90X0	181044630	181117706	2	73076
	1	1_91X0	186971694	187119468	2	147774
	1	1_92X0	190925042	191069388	3	144346
	1	1_93X0	192058189	192097760	2	39571
	1	1_94X0	193554532	193719088	2	164556
	1	1_95X0	197279455	197374813	2	95358
	1	1_96X0	215889435	216205726	5	316291
	1	1_97X0	216655287	217167433	11	512146
	1	1_98X0	217186817	217259961	3	73144
	1	1_99X0	217347042	217358096	2	11054
	1	1_9X0	13905008	13979938	2	74930
	10	10_10X0	34452437	34631787	3	179350
	10	10_11X0	35260644	35398105	2	137461
	10	10_12X0	36029063	36042337	2	13274
	10	10_13X0	52785595	52787631	2	2036
	10	10_14X0	54356089	54479306	2	123217
	10	10_15X0	56410831	56775169	8	364338
	10	10_16X0	62028627	62070381	2	41754
	10	10_17X0	63246681	63456176	2	209495
	10	10_18X0	63703565	64125784	11	422219
	10	10_19X0	64498303	64590436	2	92133
	10	10_1X0	715596	733346	2	17750
	10	10_20X0	64757663	64782044	2	24381
	10	10_21X0	65348001	65604543	4	256542
	10	10_22X0	71298929	71442806	4	143877
	10	10_23X0	72990799	73168630	2	177831
	10	10_24X0	76488958	76580204	2	91246
	10	10_25X0	76997044	77089534	3	92490
	10	10_26X0	77578067	77592670	2	14603
	10	10_27X0	78167101	78284477	2	117376
	10	10_28X0	79613421	79675041	3	61620
	10	10_29X0	81852823	81908830	2	56007
	10	10_2X0	7764375	7783713	2	19338
	10	10_30X0	82571292	82578465	2	7173
	10	10_31X0	82769278	82790063	2	20785
	10	10_32X0	83627124	83775849	2	148725
	10	10_33X0	84490638	84556075	2	65437
	10	10_34X0	84700632	84805659	2	105027
	10	10_35X0	90791734	90809513	2	17779
	10	10_3X0	18474277	18500993	2	26716
	10	10_4X0	18576029	18873957	8	297928
	10	10_5X0	19089507	19226345	6	136838
	10	10_6X0	25929533	26005238	2	75705
	10	10_7X0	31052366	31315784	5	263418
	10	10_8X0	32831367	33212703	13	381336
	10	10_9X0	33367812	33495845	2	128033
	11	11_10X0	10851817	10961454	6	109637
	11	11_11X0	11075234	11114800	2	39566
	11	11_12X0	12637677	12696629	2	58952
	11	11_13X0	12833652	12961186	2	127534

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	11	11_14X0	20398293	20936976	5	538683
	11	11_15X0	42193481	42357583	2	164102
	11	11_16X0	44533567	44592762	2	59195
	11	11_17X0	44939161	45159103	5	219942
	11	11_18X0	45275311	45366283	2	90972
	11	11_19X0	45850263	46003318	2	153055
	11	11_1X0	208826	279669	2	70843
	11	11_20X0	46052381	46075420	3	23039
	11	11_21X0	47990253	48071073	2	80820
	11	11_22X0	48222639	48982397	23	759758
	11	11_23X0	49218297	49453490	3	235193
	11	11_24X0	49468369	49536768	2	68399
	11	11_25X0	50006420	50122923	8	116503
	11	11_26X0	50705737	50723631	2	17894
	11	11_27X0	51238595	51527802	5	289207
	11	11_28X0	51858162	52000265	2	142103
	11	11_29X0	52053212	52061865	2	8653
	11	11_2X0	4451091	4529065	2	77974
	11	11_30X0	52691849	52842841	3	150992
	11	11_31X0	53326016	53347653	2	21637
	11	11_32X0	63957324	64012059	2	54735
	11	11_33X0	71161764	71330432	3	168668
	11	11_34X0	76480208	76491772	2	11564
	11	11_35X0	90885304	91030770	3	145466
	11	11_36X0	112733091	112826475	3	93384
	11	11_37X0	113205563	113266184	2	60621
	11	11_38X0	120237253	120369470	2	132217
	11	11_39X0	121071162	121159221	2	88059
	11	11_3X0	4913052	5148019	3	234967
	11	11_40X0	132656019	132920627	6	264608
	11	11_4X0	6448822	6588073	2	139251
	11	11_5X0	6754014	6969798	4	215784
	11	11_6X0	7231770	7305774	5	74004
	11	11_7X0	8251980	8383790	4	131810
	11	11_8X0	9538991	10329050	15	790059
	11	11_9X0	10620131	10689515	2	69384
	12	12_10X0	86912274	87067070	3	154796
	12	12_11X0	89656578	89897903	3	241325
	12	12_12X0	93704924	93772075	2	67151
	12	12_13X0	102585088	102617506	2	32418
	12	12_14X0	104021181	104220174	2	198993
	12	12_15X0	104855709	104879963	2	24254
	12	12_16X0	105241517	105262969	2	21452
	12	12_17X0	105483000	105558611	2	75611
	12	12_18X0	116194271	116372872	2	178601
	12	12_19X0	116868052	117088099	2	220047
	12	12_1X0	23221175	23650884	10	429709
	12	12_20X0	118307224	118399734	3	92510
	12	12_21X0	118467897	118487328	2	19431
	12	12_22X0	119714757	119733571	2	18814
	12	12_2X0	37263013	37295361	2	32348
	12	12_3X0	43205711	43447340	2	241629
	12	12_4X0	57827669	57844432	2	16763
	12	12_5X0	61841647	61951842	10	110195
	12	12_6X0	63288416	63354874	2	66458

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	12	12_7X0	74754671	74866893	2	112222
	12	12_8X0	76627626	76802124	2	174498
	12	12_9X0	86208046	86568670	3	360624
	13	13_10X0	78143073	78290980	6	147907
	13	13_11X0	84811306	84838566	2	27260
	13	13_12X0	86251264	86329218	2	77954
	13	13_13X0	87715457	87984319	4	268862
	13	13_13X1	87266315	87522065	2	255750
	13	13_14X0	88202473	88276934	3	74461
	13	13_15X0	89142698	89188457	2	45759
	13	13_16X0	89195664	89217670	2	22006
	13	13_17X0	91036511	91090803	2	54292
	13	13_18X0	93792437	94235700	6	443263
	13	13_19X0	103493137	103591957	2	98820
	13	13_1X0	23253668	23328117	2	74449
	13	13_20X0	103965729	104337608	9	371879
	13	13_21X0	104802144	104825085	2	22941
	13	13_2X0	26536916	26762877	2	225961
	13	13_3X0	38276351	38538239	2	261888
	13	13_4X0	42991259	43040659	2	49400
	13	13_5X0	44143068	44219645	2	76577
	13	13_6X0	66710468	66715497	2	5029
	13	13_7X0	68140642	68173720	2	33078
	13	13_8X0	70079406	70105979	2	26573
	13	13_9X0	73699259	73717167	2	17908
	14	14_10X0	6673804	6842218	2	168414
	14	14_11X0	7415685	7585226	4	169541
	14	14_12X0	8482860	8576403	2	93543
	14	14_13X0	9264123	9323811	2	59688
	14	14_14X0	9516037	9559792	2	43755
	14	14_15X0	10296992	10357246	2	60254
	14	14_16X0	10519165	10664606	4	145441
	14	14_17X0	10744899	10784273	2	39374
	14	14_18X0	11347605	11585361	7	237756
	14	14_19X0	11836874	11974010	3	137136
	14	14_1X0	125980	153907	3	27927
	14	14_20X0	12646628	12794232	2	147604
	14	14_21X0	12944875	13598936	10	654061
	14	14_22X0	13728373	13764227	2	35854
	14	14_23X0	14051609	14115428	3	63819
	14	14_24X0	14327123	14369590	2	42467
	14	14_25X0	14384324	14601650	3	217326
	14	14_26X0	15627358	15734018	2	106660
	14	14_27X0	15751112	15765432	2	14320
	14	14_28X0	16644620	17030322	9	385702
	14	14_28X1	16255930	16629710	9	373780
	14	14_29X0	21173548	21209582	2	36034
	14	14_2X0	332488	391764	2	59276
	14	14_30X0	30324206	30505020	2	180814
	14	14_31X0	46510094	46725162	5	215068
	14	14_32X0	46865584	47143633	6	278049
	14	14_33X0	50125657	50249721	2	124064
	14	14_34X0	58014385	58075050	2	60665
	14	14_35X0	58148640	58408912	4	260272
	14	14_36X0	60106497	60526536	11	420039

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	14	14_36X1	59329159	60095061	13	765902
	14	14_36X2	60574459	61032673	8	458214
	14	14_37X0	61107481	61300357	2	192876
	14	14_38X0	61922428	61940466	2	18038
	14	14_39X0	62213451	62280130	2	66679
	14	14_3X0	676193	708462	2	32269
	14	14_40X0	63369680	63432016	2	62336
	14	14_41X0	64144558	64187474	2	42916
	14	14_42X0	65245229	65265378	2	20149
	14	14_43X0	65891735	65973492	2	81757
	14	14_44X0	66834923	66970245	2	135322
	14	14_45X0	68215705	68384191	2	168486
	14	14_46X0	76029391	76057640	2	28249
	14	14_47X0	80033664	80191936	6	158272
	14	14_48X0	85283153	85313985	2	30832
	14	14_49X0	92470496	92518561	2	48065
	14	14_4X0	864402	953844	2	89442
	14	14_50X0	92661428	93086861	9	425433
	14	14_51X0	95024726	95178542	4	153816
	14	14_52X0	101968100	101973207	2	5107
	14	14_53X0	103965826	104044617	2	78791
	14	14_54X0	104606896	104712724	2	105828
	14	14_55X0	106757835	106804825	3	46990
	14	14_56X0	107789428	107811385	2	21957
	14	14_57X0	107877535	108078434	2	200899
	14	14_58X0	108092791	108219577	5	126786
	14	14_59X0	108258857	108293932	2	35075
	14	14_5X0	1523612	1600473	2	76861
	14	14_60X0	113689923	114243182	8	553259
	14	14_61X0	114419594	114596887	5	177293
	14	14_62X0	115327752	115465100	2	137348
	14	14_63X0	115605462	115681198	2	75736
	14	14_64X0	119949097	120017547	2	68450
	14	14_65X0	123452295	123676892	2	224597
	14	14_66X0	123822607	123916257	2	93650
	14	14_6X0	1800268	1828838	2	28570
	14	14_7X0	4583826	4698350	3	114524
	14	14_8X0	6348120	6448945	2	100825
	14	14_9X0	6583473	6653075	2	69602
	15	15_10X0	21546830	21610022	2	63192
	15	15_11X0	21622877	21725243	2	102366
	15	15_12X0	21786334	21852521	2	66187
	15	15_13X0	24608660	24701762	2	93102
	15	15_14X0	26082669	26091963	2	9294
	15	15_15X0	30174037	30205450	2	31413
	15	15_16X0	30229081	30418293	4	189212
	15	15_17X0	37581802	37701838	2	120036
	15	15_18X0	40145211	40253532	5	108321
	15	15_19X0	40407986	40607684	2	199698
	15	15_1X0	1241952	1508412	8	266460
	15	15_20X0	41517847	41544713	2	26866
	15	15_21X0	42490249	42556994	2	66745
	15	15_22X0	42751947	42899330	2	147383
	15	15_23X0	53911941	54290627	6	378686
	15	15_24X0	55740803	55743452	2	2649

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	15	15_25X0	69974493	70086292	2	111799
	15	15_26X0	70222606	70419537	3	196931
	15	15_27X0	74650299	74800259	2	149960
	15	15_28X0	89130117	89185988	2	55871
	15	15_29X0	97223955	97255890	2	31935
	15	15_2X0	2615711	2754242	5	138531
	15	15_30X0	98268903	98319093	2	50190
	15	15_31X0	102361139	102491892	4	130753
	15	15_32X0	104313315	104393126	2	79811
	15	15_33X0	106461567	106552290	2	90723
	15	15_3X0	3335010	3390635	3	55625
	15	15_4X0	4665637	4821026	2	155389
	15	15_5X0	4868632	4893385	2	24753
	15	15_6X0	8185795	8208118	2	22323
	15	15_7X0	10194105	10269013	2	74908
	15	15_8X0	11824975	11846106	2	21131
	15	15_9X0	11951235	12164314	2	213079
	16	16_10X0	7116938	7129762	3	12824
	16	16_11X0	7764713	7844157	3	79444
	16	16_12X0	8496517	8662897	2	166380
	16	16_13X0	9991298	10337369	6	346071
	16	16_14X0	15572462	15652645	2	80183
	16	16_15X0	18813805	18962575	2	148770
	16	16_16X0	22043968	22179735	2	135767
	16	16_17X0	27453790	27562130	5	108340
	16	16_18X0	28596949	28741813	3	144864
	16	16_19X0	29059430	29301441	7	242011
	16	16_1X0	1507675	1544476	2	36801
	16	16_20X0	29455967	29529072	2	73105
	16	16_21X0	30162941	30359727	10	196786
	16	16_22X0	31608250	31634044	2	25794
	16	16_23X0	31749291	31787889	2	38598
	16	16_24X0	32032630	32114305	2	81675
	16	16_25X0	32153095	32186650	2	33555
	16	16_26X0	39298001	39385777	3	87776
	16	16_27X0	41879167	41966212	2	87045
	16	16_28X0	46789863	46862704	2	72841
	16	16_29X0	46965915	46989880	2	23965
	16	16_2X0	2820976	3097727	8	276751
	16	16_30X0	47607457	47874881	10	267424
	16	16_31X0	47935948	48168294	18	232346
	16	16_32X0	48281687	48502158	14	220471
	16	16_33X0	48714166	48738496	2	24330
	16	16_34X0	48808995	48900162	2	91167
	16	16_35X0	49081790	49231004	3	149214
	16	16_36X0	49494425	49554584	2	60159
	16	16_37X0	49729538	49743437	2	13899
	16	16_38X0	50601228	50690356	2	89128
	16	16_39X0	50743524	50757300	2	13776
	16	16_3X0	3174764	3256740	2	81976
	16	16_40X0	51960026	51982615	2	22589
	16	16_41X0	52657518	52767709	2	110191
	16	16_42X0	55986315	56021694	4	35379
	16	16_43X0	58783886	59009519	3	225633
	16	16_44X0	60608952	60859068	3	250116

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	16	16_44X1	60866955	61036970	2	170015
	16	16_45X0	61805897	61883058	2	77161
	16	16_46X0	66194447	66373215	4	178768
	16	16_47X0	66575833	66602190	2	26357
	16	16_48X0	67559150	67581015	2	21865
	16	16_49X0	68280899	68356246	2	75347
	16	16_4X0	4121144	4226968	2	105824
	16	16_50X0	69800436	69823839	2	23403
	16	16_51X0	71451813	71498599	2	46786
	16	16_52X0	74396031	74435266	2	39235
	16	16_5X0	4575234	4725166	2	149932
	16	16_6X0	4793923	4832549	3	38626
	16	16_7X0	6547845	6679181	2	131336
	16	16_8X0	6705115	6711719	2	6604
	16	16_9X0	6741766	6862469	3	120703
	17	17_10X0	90308923	90368901	3	59978
	17	17_1X0	848543	868575	2	20032
	17	17_2X0	7085873	7252464	3	166591
	17	17_3X0	9975810	10043109	2	67299
	17	17_4X0	19803851	19877137	2	73286
	17	17_5X0	30488205	30583011	2	94806
	17	17_6X0	35523482	35579639	2	56157
	17	17_7X0	77034360	77076886	2	42526
	17	17_8X0	77740334	77761471	2	21137
	17	17_9X0	80410588	80632432	2	221844
	18	18_1X0	10041129	10070202	2	29073
	18	18_2X0	23171208	23641665	6	470457
	18	18_3X0	27353132	27456805	3	103673
	18	18_4X0	37834608	37952824	2	118216
	18	18_5X0	39199667	39207789	2	8122
	18	18_6X0	56018493	56541796	12	523303
	18	18_7X0	67003380	67088978	2	85598
	19	19_1X0	544567	575425	4	30858
	19	19_2X0	1737846	1812646	2	74800
	19	19_3X0	44066552	44360147	16	293595
	19	19_3X1	49996051	50498908	18	502857
	19	19_3X10	11162736	11471582	10	308846
	19	19_3X11	10904396	11089732	7	185336
	19	19_3X12	46379420	46637310	7	257890
	19	19_3X13	51162797	51338009	7	175212
	19	19_3X14	46139808	46364790	6	224982
	19	19_3X15	10728033	10873292	4	145259
	19	19_3X16	46976465	47153864	6	177399
	19	19_3X17	49009800	49225508	6	215708
	19	19_3X18	50700902	50829376	4	128474
	19	19_3X19	31570770	31832495	7	261725
	19	19_3X2	37560086	37829181	11	269095
	19	19_3X20	42360842	42397542	6	36700
	19	19_3X21	2577066	2764088	5	187022
	19	19_3X22	31858338	32110984	7	252646
	19	19_3X23	47267077	47458676	6	191599
	19	19_3X24	51016849	51094802	3	77953
	19	19_3X25	37984861	38115464	4	130603
	19	19_3X26	49350717	49574286	7	223569
	19	19_3X27	31160512	31301588	4	141076

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	19	19_3X28	8487796	8677565	5	189769
	19	19_3X29	30937723	31115997	4	178274
	19	19_3X3	44385880	44763554	12	377674
	19	19_3X30	36791566	37095476	6	303910
	19	19_3X31	35393294	35582411	5	189117
	19	19_3X32	31345752	31483982	4	138230
	19	19_3X33	8255804	8427424	5	171620
	19	19_3X34	37185492	37384371	5	198879
	19	19_3X35	29918084	30064978	5	146894
	19	19_3X36	50858689	50990589	3	131900
	19	19_3X37	14266484	14444096	4	177612
	19	19_3X38	47598516	47784993	4	186477
	19	19_3X39	34413338	34612643	5	199305
	19	19_3X4	48471476	48638453	9	166977
	19	19_3X40	18275069	18614099	6	339030
	19	19_3X41	2766724	2929803	3	163079
	19	19_3X42	38159235	38335944	4	176709
	19	19_3X43	30743937	30911143	3	167206
	19	19_3X44	30411202	30495260	4	84058
	19	19_3X45	48786300	48990613	4	204313
	19	19_3X46	9201925	9272507	3	70582
	19	19_3X47	45738576	45919730	4	181154
	19	19_3X48	14452506	14623255	3	170749
	19	19_3X49	45379984	45520578	4	140594
	19	19_3X5	47798211	47983958	6	185747
	19	19_3X50	35201273	35305947	3	104674
	19	19_3X51	35922986	36094061	4	171075
	19	19_3X52	33931747	34140340	4	208593
	19	19_3X53	41172115	41235693	3	63578
	19	19_3X54	38435903	38513330	2	77427
	19	19_3X55	19554589	19767682	3	213093
	19	19_3X56	5578380	5615990	3	37610
	19	19_3X57	19231603	19413462	3	181859
	19	19_3X58	10424999	10677240	3	252241
	19	19_3X59	33598170	33654679	3	56509
	19	19_3X6	44982929	45285302	10	302373
	19	19_3X60	29753650	29785874	3	32224
	19	19_3X61	33128897	33176847	3	47950
	19	19_3X62	43272847	43292657	2	19810
	19	19_3X63	18753278	18833991	2	80713
	19	19_3X64	17804331	17856882	2	52551
	19	19_3X65	39128704	39223003	2	94299
	19	19_3X66	13175886	13271453	2	95567
	19	19_3X67	20400693	20472594	2	71901
	19	19_3X68	6168538	6200434	2	31896
	19	19_3X69	43586176	43615127	2	28951
	19	19_3X7	37841518	37971522	4	130004
	19	19_3X70	36342170	36406224	2	64054
	19	19_3X71	7375026	7403211	2	28185
	19	19_3X72	14098648	14148483	2	49835
	19	19_3X73	5329460	5413373	2	83913
	19	19_3X74	29270699	29360633	2	89934
	19	19_3X75	47994453	48142671	3	148218
	19	19_3X76	27982713	28064442	2	81729
	19	19_3X77	13603431	13633384	2	29953

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	19	19_3X78	45570204	45665203	2	94999
	19	19_3X79	6693022	6710164	2	17142
	19	19_3X8	48703058	48782163	4	79105
	19	19_3X80	9030510	9151494	2	120984
	19	19_3X9	50508839	50663630	6	154791
	2	2_10X0	83744233	83898733	5	154500
	2	2_11X0	85335890	85365672	3	29782
	2	2_12X0	85381892	85503909	4	122017
	2	2_13X0	85547039	85671912	2	124873
	2	2_14X0	85757896	85793890	2	35994
	2	2_15X0	86354079	86570904	2	216825
	2	2_16X0	88944791	89053721	4	108930
	2	2_17X0	89352679	89876574	7	523895
	2	2_18X0	91066433	91364965	8	298532
	2	2_19X0	92971315	93218770	2	247455
	2	2_1X0	18251137	18543776	6	292639
	2	2_20X0	95344438	95402875	3	58437
	2	2_21X0	97535216	97591498	2	56282
	2	2_22X0	99241783	99312519	3	70736
	2	2_23X0	101902425	101956283	3	53858
	2	2_24X0	102305949	102501457	2	195508
	2	2_25X0	116380015	116428251	2	48236
	2	2_26X0	126323963	126507241	2	183278
	2	2_27X0	128079009	128144110	2	65101
	2	2_28X0	134768128	134875022	2	106894
	2	2_29X0	135312266	135449916	5	137650
	2	2_2X0	29277206	29314724	2	37518
	2	2_30X0	137191462	137379733	3	188271
	2	2_31X0	137801551	137876485	2	74934
	2	2_32X0	139306586	139335160	2	28574
	2	2_33X0	140140300	140504438	4	364138
	2	2_34X0	147203511	147293234	2	89723
	2	2_35X0	152731724	152809023	2	77299
	2	2_36X0	155160161	155183977	2	23816
	2	2_37X0	173323402	173587471	3	264069
	2	2_38X0	181231972	181335549	2	103577
	2	2_39X0	187930759	188108991	2	178232
	2	2_3X0	32577094	32705841	2	128747
	2	2_4X0	33936772	34075550	2	138778
	2	2_5X0	42164266	42371138	2	206872
	2	2_6X0	44598313	44741102	2	142789
	2	2_7X0	50463855	50495252	2	31397
	2	2_8X0	82769003	82822140	3	53137
	2	2_9X0	83364214	83710055	3	345841
	20	20_10X0	15013856	15194201	2	180345
	20	20_11X0	18550368	18589456	2	39088
	20	20_12X0	18649303	18900837	4	251534
	20	20_13X0	20886535	21111932	2	225397
	20	20_14X0	21402405	21438672	2	36267
	20	20_15X0	25056573	25167298	2	110725
	20	20_16X0	26576900	26623856	5	46956
	20	20_17X0	26745974	26976505	3	230531
	20	20_18X0	27096575	27108899	2	12324
	20	20_19X0	27195343	27245509	2	50166
	20	20_1X0	20709	47849	6	27140

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	20	20_20X0	27268471	27289695	2	21224
	20	20_21X0	27290415	27367872	3	77457
	20	20_22X0	27483830	27651633	3	167803
	20	20_23X0	30600476	30747485	2	147009
	20	20_24X0	32564124	32585727	2	21603
	20	20_25X0	38328721	38602917	5	274196
	20	20_26X0	39223906	39388462	16	164556
	20	20_27X0	40059444	40111259	3	51815
	20	20_28X0	40320617	40380872	2	60255
	20	20_29X0	48897557	49049380	3	151823
	20	20_2X0	145915	228111	2	82196
	20	20_30X0	49290236	49347183	3	56947
	20	20_31X0	49509505	49558940	2	49435
	20	20_32X0	50351887	50381217	3	29330
	20	20_33X0	51051415	51249440	2	198025
	20	20_34X0	54624390	54646089	2	21699
	20	20_35X0	55228338	55260892	2	32554
	20	20_36X0	56328503	56387992	2	59489
	20	20_37X0	57099152	57122796	2	23644
	20	20_38X0	57598717	57625003	2	26286
	20	20_39X0	68460790	68528515	3	67725
	20	20_3X0	467487	569240	4	101753
	20	20_4X0	644140	661288	2	17148
	20	20_5X0	1068589	1146283	6	77694
	20	20_6X0	1853581	2093710	3	240129
	20	20_7X0	2520799	2596183	4	75384
	20	20_8X0	2882525	3233498	9	350973
	20	20_9X0	14307675	14354047	2	46372
	3	3_10X0	33754519	33788846	2	34327
	3	3_11X0	37641652	37880734	4	239082
	3	3_12X0	39235455	39326621	3	91166
	3	3_13X0	40846145	40907282	2	61137
	3	3_14X0	40986470	41134152	4	147682
	3	3_15X0	41157031	41418440	5	261409
	3	3_16X0	41577727	41631254	3	53527
	3	3_17X0	41803324	41845784	2	42460
	3	3_18X0	43600774	43633153	2	32379
	3	3_19X0	44234908	44267834	2	32926
	3	3_1X0	647973	793911	2	145938
	3	3_20X0	45043281	45105780	2	62499
	3	3_21X0	52555633	52720710	2	165077
	3	3_22X0	52934647	53756563	6	821916
	3	3_22X1	53897191	54277123	4	379932
	3	3_23X0	65749110	65784955	2	35845
	3	3_24X0	76297716	76520297	2	222581
	3	3_25X0	78348132	78552036	2	203904
	3	3_26X0	84093937	84242849	9	148912
	3	3_27X0	87528757	87778242	2	249485
	3	3_28X0	92175190	92202657	2	27467
	3	3_29X0	94423086	94503864	2	80778
	3	3_2X0	2062549	2197151	9	134602
	3	3_30X0	105509198	105526809	2	17611
	3	3_31X0	107111226	107235625	3	124399
	3	3_32X0	108594557	108613552	2	18995
	3	3_33X0	112247402	112337664	2	90262

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	3	3_34X0	130484612	130626532	2	141920
	3	3_35X0	146749684	146848933	2	99249
	3	3_36X0	152845607	152942516	3	96909
	3	3_37X0	157038046	157155174	2	117128
	3	3_38X0	162047403	162127818	2	80415
	3	3_39X0	164494764	164707556	2	212792
	3	3_3X0	2896519	3053341	2	156822
	3	3_40X0	165348257	165419243	6	70986
	3	3_41X0	165526400	165586212	2	59812
	3	3_42X0	166305654	166371075	2	65421
	3	3_43X0	166463041	166738986	5	275945
	3	3_44X0	171573338	172010150	7	436812
	3	3_45X0	172299369	172517794	3	218425
	3	3_46X0	172595014	172899008	8	303994
	3	3_47X0	172963605	172977441	2	13836
	3	3_48X0	173300674	173349089	2	48415
	3	3_49X0	174120570	174264738	2	144168
	3	3_4X0	4274295	4326710	3	52415
	3	3_5X0	5149535	5221562	2	72027
	3	3_6X0	13132827	13346289	4	213462
	3	3_7X0	15918349	15955691	2	37342
	3	3_8X0	16031540	16082292	4	50752
	3	3_9X0	19707595	19840073	2	132478
	4	4_10X0	28758004	29115245	5	357241
	4	4_11X0	29439045	29532154	3	93109
	4	4_12X0	29621844	29753362	6	131518
	4	4_13X0	30837796	30927674	2	89878
	4	4_14X0	30974023	31101483	5	127460
	4	4_15X0	31107912	31161652	3	53740
	4	4_16X0	31254869	31282382	2	27513
	4	4_17X0	31838387	32672901	14	834514
	4	4_18X0	35164582	35191334	2	26752
	4	4_19X0	35233679	35539718	3	306039
	4	4_1X0	2641019	2758475	3	117456
	4	4_20X0	38674054	38802092	3	128038
	4	4_21X0	40597148	40818454	5	221306
	4	4_22X0	42554573	42787044	4	232471
	4	4_23X0	45612361	45637673	2	25312
	4	4_24X0	49043793	49208655	3	164862
	4	4_25X0	49292791	49369736	4	76945
	4	4_26X0	50051519	50174874	2	123355
	4	4_27X0	51483908	51537196	2	53288
	4	4_28X0	52151627	52233740	2	82113
	4	4_29X0	69205963	69364234	6	158271
	4	4_2X0	2935965	3001978	2	66013
	4	4_30X0	69519910	69533627	2	13717
	4	4_31X0	84889645	85027203	2	137558
	4	4_32X0	101008965	101085061	2	76096
	4	4_33X0	112908282	113275599	11	367317
	4	4_34X0	125857221	126046023	2	188802
	4	4_35X0	130399793	130477739	2	77946
	4	4_36X0	130516697	130625122	6	108425
	4	4_37X0	146611244	146714624	3	103380
	4	4_38X0	146901740	146928510	2	26770
	4	4_39X0	151707539	151886443	2	178904

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	4	4_3X0	7767015	7812810	2	45795
	4	4_40X0	156167605	156484963	3	317358
	4	4_41X0	156542040	156693439	3	151399
	4	4_4X0	25737310	26295061	32	557751
	4	4_5X0	26382728	26559936	6	177208
	4	4_6X0	26898177	26913839	4	15662
	4	4_7X0	27160040	27221517	2	61477
	4	4_8X0	27589640	27674614	14	84974
	4	4_9X0	27775861	28333511	14	557650
	5	5_10X0	67850804	67948127	3	97323
	5	5_11X0	70899475	70989534	3	90059
	5	5_12X0	71241769	71294461	3	52692
	5	5_13X0	71329254	71358222	2	28968
	5	5_14X0	71425904	71551971	2	126067
	5	5_15X0	71629844	72311045	6	681201
	5	5_16X0	72366640	72937993	4	571353
	5	5_17X0	89405686	89472167	2	66481
	5	5_18X0	90100126	90270842	2	170716
	5	5_19X0	94901734	95326625	7	424891
	5	5_1X0	8113	400300	3	392187
	5	5_20X0	98783678	98951319	2	167641
	5	5_21X0	148841308	148892300	3	50992
	5	5_22X0	149977826	150089529	2	111703
	5	5_23X0	157597259	157626035	2	28776
	5	5_24X0	159177060	159241855	2	64795
	5	5_25X0	177394775	177397658	2	2883
	5	5_26X0	180274249	180335245	2	60996
	5	5_27X0	181966280	182021551	2	55271
	5	5_2X0	6537426	6568091	2	30665
	5	5_3X0	14514433	14662114	2	147681
	5	5_4X0	37654767	37680773	2	26006
	5	5_5X0	54183409	54277703	2	94294
	5	5_6X0	58818576	58899009	2	80433
	5	5_7X0	64063316	64184358	2	121042
	5	5_8X0	65191176	65226912	3	35736
	5	5_9X0	67165387	67623530	12	458143
	6	6_10X0	77957541	78147834	2	190293
	6	6_11X0	91354325	91607551	3	253226
	6	6_12X0	121629387	121784793	2	155406
	6	6_13X0	126638078	126741970	2	103892
	6	6_14X0	134967178	134991320	2	24142
	6	6_15X0	135239075	136245330	15	1006255
	6	6_16X0	142136912	142751487	6	614575
	6	6_17X0	144448311	144683171	3	234860
	6	6_18X0	145624081	145826008	3	201927
	6	6_19X0	151246575	151429844	3	183269
	6	6_1X0	907326	1138947	3	231621
	6	6_20X0	169730837	169975418	2	244581
	6	6_21X0	172873369	173236698	6	363329
	6	6_22X0	175075373	175309007	4	233634
	6	6_2X0	14678279	14791639	2	113360
	6	6_3X0	42813667	42876301	2	62634
	6	6_4X0	51412146	51494879	2	82733
	6	6_5X0	51641609	51818822	2	177213
	6	6_6X0	52263923	52384577	2	120654

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	6	6_7X0	65593674	65732336	2	138662
	6	6_8X0	66675417	66810150	2	134733
	6	6_9X0	71124408	71347684	3	223276
	7	7_10X0	26117325	26255112	2	137787
	7	7_11X0	39786996	39882761	2	95765
	7	7_12X0	48209551	48359033	2	149482
	7	7_13X0	48803766	48850018	2	46252
	7	7_14X0	49314903	49553189	5	238286
	7	7_15X0	53195539	53237045	2	41506
	7	7_16X0	59603415	59812121	2	208706
	7	7_17X0	63534794	63601796	2	67002
	7	7_18X0	64461259	64490890	2	29631
	7	7_19X0	64799140	65053802	2	254662
	7	7_1X0	366233	409136	2	42903
	7	7_20X0	76498472	77180872	7	682400
	7	7_21X0	77564717	77748322	3	183605
	7	7_22X0	78527482	78619110	4	91628
	7	7_23X0	80012975	80030987	2	18012
	7	7_24X0	80384518	80436134	2	51616
	7	7_25X0	81035329	81409344	8	374015
	7	7_26X0	81470090	81509629	2	39539
	7	7_27X0	81564193	81695883	3	131690
	7	7_28X0	93351277	93416605	2	65328
	7	7_29X0	106690389	106785973	2	95584
	7	7_2X0	4080107	4157962	2	77855
	7	7_30X0	107347195	107551070	2	203875
	7	7_31X0	109217453	109265841	2	48388
	7	7_32X0	121424977	121469153	2	44176
	7	7_33X0	124406109	124467484	2	61375
	7	7_34X0	127000502	127195616	4	195114
	7	7_35X0	130267731	130290686	2	22955
	7	7_36X0	131621524	131836877	2	215353
	7	7_37X0	132149858	132265845	2	115987
	7	7_38X0	151133861	151492229	11	358368
	7	7_39X0	152552477	152575069	2	22592
	7	7_3X0	14737519	14855810	2	118291
	7	7_40X0	153077390	153145343	2	67953
	7	7_41X0	160033089	160070533	2	37444
	7	7_4X0	16133459	16208066	2	74607
	7	7_5X0	16496179	16810626	4	314447
	7	7_6X0	17860630	17961154	3	100524
	7	7_7X0	19497368	19568422	4	71054
	7	7_8X0	23107028	23236510	2	129482
	7	7_9X0	24804409	24915421	2	111012
	8	8_10X0	81382699	81584624	2	201925
	8	8_11X0	138235555	138759632	14	524077
	8	8_12X0	139263215	139393641	2	130426
	8	8_13X0	139604467	140292691	8	688224
	8	8_1X0	6559708	7025847	17	466139
	8	8_2X0	21070445	21300406	3	229961
	8	8_3X0	21777415	21808370	2	30955
	8	8_4X0	22418567	22651092	3	232525
	8	8_5X0	37343909	38244755	6	900846
	8	8_6X0	38316262	38415812	2	99550
	8	8_7X0	41164338	41244382	2	80044

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	8	8_8X0	52202157	52371778	2	169621
	8	8_9X0	77362862	77618950	4	256088
	9	9_10X0	60658307	60836362	2	178055
	9	9_11X0	60849047	61159541	3	310494
	9	9_12X0	81092862	81302895	5	210033
	9	9_13X0	81795951	81894517	4	98566
	9	9_14X0	85439029	85455530	2	16501
	9	9_15X0	87092433	87365961	4	273528
	9	9_16X0	89516965	89533271	2	16306
	9	9_17X0	93322481	93422862	2	100381
	9	9_18X0	93661061	93902501	2	241440
	9	9_19X0	95727845	95761302	3	33457
	9	9_1X0	1109569	1129687	2	20118
	9	9_20X0	96547555	96591204	2	43649
	9	9_21X0	108551782	108774480	4	222698
	9	9_22X0	114942343	114982945	2	40602
	9	9_23X0	115197578	115212384	2	14806
	9	9_24X0	125629677	125761422	2	131745
	9	9_2X0	5048952	5422386	4	373434
	9	9_3X0	36250432	36687524	5	437092
	9	9_4X0	38036244	38118086	3	81842
	9	9_5X0	42417345	42442592	2	25247
	9	9_6X0	46029491	46050616	2	21125
	9	9_7X0	49989411	50021816	2	32405
	9	9_8X0	50364075	50471109	4	107034
	9	9_9X0	54897817	54963745	2	65928
	KZ097459.1	KZ097459.1_1X0	31131	41356	2	10225
	KZ097459.1	KZ097459.1_2X0	151059	174197	2	23138
	X	X_10X0	44529894	44625815	2	95921
	X	X_11X0	44990749	45230850	3	240101
	X	X_12X0	45284236	45600711	8	316475
	X	X_13X0	46331592	46359959	2	28367
	X	X_14X0	46475658	46776899	7	301241
	X	X_15X0	48290609	48483176	2	192567
	X	X_16X0	49354215	49627448	3	273233
	X	X_17X0	51847506	52060240	5	212734
	X	X_18X0	61832879	62020937	3	188058
	X	X_19X0	62909437	63033434	3	123997
	X	X_1X0	59676	166011	2	106335
	X	X_20X0	63288792	63295896	2	7104
	X	X_21X0	91777460	91853973	2	76513
	X	X_22X0	92132355	92169593	2	37238
	X	X_23X0	92369254	92458996	3	89742
	X	X_24X0	92693186	93026849	6	333663
	X	X_25X0	108341702	108406884	2	65182
	X	X_26X0	108979748	109029389	2	49641
	X	X_27X0	109745047	109755447	2	10400
	X	X_28X0	119210312	119243386	2	33074
	X	X_29X0	123550732	123638318	2	87586
	X	X_2X0	261604	469527	4	207923
	X	X_30X0	124140346	124200136	2	59790
	X	X_31X0	133491056	133512499	2	21443
	X	X_32X0	137844083	138143598	3	299515
	X	X_33X0	138778684	138980351	2	201667
	X	X_34X0	141267185	141340192	2	73007

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	X	X_35X0	141687299	142195409	5	508110
	X	X_36X0	142388354	142541992	3	153638
	X	X_3X0	6061983	6298432	2	236449
	X	X_4X0	10199151	10254992	2	55841
	X	X_5X0	10657467	10677887	2	20420
	X	X_6X0	12613017	12685330	2	72313
	X	X_7X0	21690383	21739275	2	48892
	X	X_8X0	23526790	23581127	3	54337
	X	X_9X0	43654346	43725125	2	70779
	1	1_10X0	43175139	43272143	3	97004
	1	1_11X0	43272188	43410438	4	138250
	1	1_12X0	43447370	43719625	2	272255
	1	1_13X0	44665793	45056544	3	390751
	1	1_14X0	45132702	45380342	2	247640
	1	1_15X0	46721044	47131282	6	410238
	1	1_16X0	47985171	48183067	2	197896
	1	1_17X0	53373965	53561177	3	187212
	1	1_18X0	57543660	57793833	2	250173
	1	1_19X0	67430398	67580707	2	150309
	1	1_1X0	116903	126574	2	9671
	1	1_20X0	68020989	68230770	3	209781
	1	1_21X0	70944570	71019657	2	75087
	1	1_22X0	77102798	77278249	2	175451
	1	1_23X0	79767826	80037149	5	269323
	1	1_24X0	81420296	81573194	3	152898
	1	1_25X0	82577879	82619106	2	41227
	1	1_26X0	88135662	88280825	2	145163
	1	1_27X0	88573144	88877414	2	304270
	1	1_28X0	89671275	89811417	5	140142
	1	1_29X0	89997938	90870537	12	872599
	1	1_2X0	351153	952872	14	601719
	1	1_2X1	1041415	1497566	10	456151
	1	1_30X0	91024775	91153683	2	128908
	1	1_31X0	91182170	91257501	2	75331
	1	1_32X0	91275075	91585661	8	310586
	1	1_33X0	91872347	92241041	5	368694
	1	1_34X0	92346538	92587360	4	240822
	1	1_35X0	92626329	93054181	10	427852
	1	1_36X0	93106559	93285130	5	178571
	1	1_37X0	93707142	93976694	4	269552
	1	1_38X0	94360560	94407090	2	46530
	1	1_39X0	94619650	94663481	2	43831
	1	1_3X0	1665206	2049511	6	384305
	1	1_40X0	95434803	95620766	2	185963
	1	1_41X0	95658387	95854927	2	196540
	1	1_42X0	96355959	96429082	3	73123
	1	1_43X0	96443599	96474142	2	30543
	1	1_44X0	98181414	98241288	2	59874
	1	1_45X0	98543765	99058625	8	514860
	1	1_46X0	99663376	99733997	2	70621
	1	1_47X0	100138513	100301469	2	162956
	1	1_48X0	100556225	100595512	2	39287
	1	1_49X0	100801197	100874600	2	73403
	1	1_4X0	21083865	21202603	3	118738
	1	1_50X0	101749768	101795644	4	45876

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_51X0	107057302	107211915	2	154613
	1	1_52X0	107261057	107382854	2	121797
	1	1_53X0	111117980	111652426	5	534446
	1	1_54X0	116826618	117047471	4	220853
	1	1_55X0	117613675	117769650	3	155975
	1	1_56X0	118533552	118611257	5	77705
	1	1_57X0	118653858	118720081	2	66223
	1	1_58X0	138931708	139289917	3	358209
	1	1_59X0	139439471	139839660	6	400189
	1	1_5X0	21545454	21666555	2	121101
	1	1_60X0	142352108	142527119	3	175011
	1	1_61X0	143928424	144062112	2	133688
	1	1_62X0	150417592	150455325	2	37733
	1	1_63X0	162007310	162239983	2	232673
	1	1_64X0	176830087	177011428	2	181341
	1	1_65X0	182770434	183063470	4	293036
	1	1_66X0	185810965	185842449	2	31484
	1	1_67X0	186556405	186763265	2	206860
	1	1_68X0	187918047	187929516	2	11469
	1	1_69X0	189550763	189643947	3	93184
	1	1_6X0	23727082	23781855	2	54773
	1	1_70X0	190321929	190627498	4	305569
	1	1_71X0	194143237	194390189	3	246952
	1	1_72X0	195116886	195255349	2	138463
	1	1_73X0	195444720	195482951	3	38231
	1	1_74X0	196702929	196933208	2	230279
	1	1_75X0	197623743	197768344	2	144601
	1	1_76X0	197867574	197963568	2	95994
	1	1_77X0	198492878	198638703	2	145825
	1	1_78X0	202705652	202994441	2	288789
	1	1_79X0	206263802	206314027	2	50225
	1	1_7X0	25828459	26039407	2	210948
	1	1_80X0	207533782	207841626	5	307844
	1	1_81X0	209573051	209635431	2	62380
	1	1_82X0	210147756	210419959	6	272203
	1	1_83X0	211035702	211080531	3	44829
	1	1_84X0	213024724	213415498	5	390774
	1	1_85X0	214391123	214422117	2	30994
	1	1_86X0	214980515	215071434	2	90919
	1	1_87X0	217050049	217445219	6	395170
	1	1_88X0	217492855	217610888	2	118033
	1	1_89X0	218432663	218448229	2	15566
	1	1_8X0	37245559	37352531	2	106972
	1	1_90X0	218629059	218659792	3	30733
	1	1_91X0	221359700	221441633	2	81933
	1	1_92X0	222554669	222558618	2	3949
	1	1_93X0	223906888	223930448	2	23560
	1	1_94X0	229056621	229135603	3	78982
	1	1_9X0	42333912	42655641	3	321729
	10	10_10X0	54853408	54885363	2	31955
	10	10_11X0	55352502	55545226	3	192724
	10	10_12X0	60049166	60122229	2	73063
	10	10_13X0	66022422	66213559	2	191137
	10	10_14X0	66225472	66668785	3	443313
	10	10_15X0	85606251	85722068	3	115817

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	10	10_16X0	87986244	88111388	5	125144
	10	10_17X0	91797842	91814821	2	16979
	10	10_18X0	93459060	93823829	3	364769
	10	10_19X0	100162743	100268778	2	106035
	10	10_1X0	5029068	5462087	5	433019
	10	10_20X0	102534846	102565633	3	30787
	10	10_21X0	103359381	103403709	2	44328
	10	10_22X0	115870236	116096957	4	226721
	10	10_23X0	122426918	122456071	2	29153
	10	10_24X0	122757434	122772408	2	14974
	10	10_2X0	5772699	5800353	2	27654
	10	10_3X0	5936737	6222184	2	285447
	10	10_4X0	7866402	8066629	2	200227
	10	10_5X0	38167185	38485198	4	318013
	10	10_6X0	42041267	42228444	4	187177
	10	10_7X0	45217721	45250150	2	32429
	10	10_8X0	45668575	45800638	2	132063
	10	10_9X0	50685124	50704711	2	19587
	11	11_10X0	8394932	8449094	2	54162
	11	11_11X0	8467513	8643338	2	175825
	11	11_12X0	9258387	9320462	3	62075
	11	11_13X0	9374152	9390756	2	16604
	11	11_15X0	11106266	11162530	2	56264
	11	11_16X0	11353273	11404029	2	50756
	11	11_17X0	11598526	12088887	3	490361
	11	11_18X0	12151237	12224170	2	72933
	11	11_19X0	12437711	12560060	3	122349
	11	11_1X0	21316	597056	7	575740
	11	11_20X0	12634659	13176571	5	541912
	11	11_21X0	13323493	13572912	2	249419
	11	11_22X0	13760335	14054661	5	294326
	11	11_23X0	14383531	14539510	3	155979
	11	11_24X0	15225720	15421185	2	195465
	11	11_25X0	15634197	16450092	13	815895
	11	11_26X0	16467230	16594273	2	127043
	11	11_27X0	16663469	16700357	2	36888
	11	11_28X0	16823493	17219598	4	396105
	11	11_29X0	17702714	18059848	4	357134
	11	11_2X0	630882	668407	2	37525
	11	11_30X0	18315986	18762732	10	446746
	11	11_31X0	19371148	19485055	2	113907
	11	11_32X0	19502344	19515933	2	13589
	11	11_33X0	20173620	20273513	3	99893
	11	11_33X1	19659817	19922755	2	262938
	11	11_34X0	24662152	24697985	2	35833
	11	11_35X0	34647700	34835982	2	188282
	11	11_36X0	51059519	51184685	2	125166
	11	11_37X0	51537968	51748224	5	210256
	11	11_38X0	55052528	55169107	2	116579
	11	11_39X0	62375855	62683363	3	307508
	11	11_3X0	800995	1044817	3	243822
	11	11_40X0	63450805	63516929	2	66124
	11	11_41X0	63551427	63692772	6	141345
	11	11_42X0	65757023	66380788	20	623765
	11	11_42X1	65347028	65701548	6	354520

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	11	11_42X2	64943901	65328323	12	384422
	11	11_42X3	64194162	64880358	14	686196
	11	11_42X4	67084509	67456303	6	371794
	11	11_42X5	68600479	68661419	2	60940
	11	11_43X0	69347408	69379296	2	31888
	11	11_44X0	71168993	71268004	2	99011
	11	11_45X0	72132182	72271802	2	139620
	11	11_46X0	73507658	73681574	2	173916
	11	11_47X0	81406972	81434724	2	27752
	11	11_48X0	85592186	85717233	4	125047
	11	11_49X0	98814418	99267694	6	453276
	11	11_4X0	1689493	1804416	2	114923
	11	11_50X0	101294526	101476464	4	181938
	11	11_51X0	108653719	108659252	2	5533
	11	11_52X0	110715030	110809868	2	94838
	11	11_53X0	111358066	111538021	3	179955
	11	11_54X0	113745243	113763109	2	17866
	11	11_55X0	114746357	114804015	2	57658
	11	11_56X0	114828568	115048940	2	220372
	11	11_57X0	115066172	115201095	5	134923
	11	11_58X0	115249965	115288792	3	38827
	11	11_59X0	116258880	116690577	3	431697
	11	11_5X0	3409602	3439676	2	30074
	11	11_60X0	120016451	120027230	2	10779
	11	11_61X0	120627008	121514157	20	887149
	11	11_62X0	121711222	121913836	5	202614
	11	11_63X0	125932437	126015982	2	83545
	11	11_64X0	131373945	131477875	3	103930
	11	11_6X0	3585683	3702076	2	116393
	11	11_7X0	4471632	4482863	2	11231
	11	11_8X0	6211139	6336738	3	125599
	11	11_9X0	7051083	7084819	2	33736
	11_random	11_random_1X0	455395	481474	2	26079
	11_random	11_random_2X0	592816	664596	2	71780
	12	12_10X0	15281797	15300052	2	18255
	12	12_11X0	21972399	22193204	2	220805
	12	12_12X0	27963807	27998944	2	35137
	12	12_13X0	45720931	45930204	2	209273
	12	12_14X0	48167456	48216845	2	49389
	12	12_15X0	48510749	48730514	5	219765
	12	12_16X0	48848233	48939471	2	91238
	12	12_17X0	49424033	49588613	2	164580
	12	12_18X0	49628791	49653950	3	25159
	12	12_19X0	51614730	51711017	2	96287
	12	12_1X0	4941464	5181880	3	240416
	12	12_20X0	51870777	52717101	22	846324
	12	12_21X0	52953770	53182901	2	229131
	12	12_22X0	53190348	53260085	2	69737
	12	12_23X0	53713386	53810922	8	97536
	12	12_24X0	54415793	54433241	2	17448
	12	12_25X0	54900091	55298240	8	398149
	12	12_26X0	55686111	55922001	5	235890
	12	12_27X0	56552994	56770443	3	217449
	12	12_28X0	65898084	65997581	2	99497
	12	12_29X0	68416021	68475701	2	59680

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	12	12_2X0	6999036	7048556	2	49520
	12	12_30X0	75771715	75958630	3	186915
	12	12_31X0	81210466	81222156	2	11690
	12	12_32X0	91877793	92123453	4	245660
	12	12_33X0	111579472	111780074	2	200602
	12	12_34X0	114850421	114954397	3	103976
	12	12_35X0	115358643	115406397	2	47754
	12	12_36X0	122541666	122683572	2	141906
	12	12_37X0	123387093	123513648	2	126555
	12	12_38X0	136027268	136330758	8	303490
	12	12_3X0	7295050	7376200	3	81150
	12	12_4X0	8173029	8248705	2	75676
	12	12_5X0	8451171	8577945	3	126774
	12	12_6X0	8861851	9294445	3	432594
	12	12_7X0	9777694	10904763	18	1127069
	12	12_8X0	11122618	11604206	15	481588
	12	12_9X0	13532299	13665587	2	133288
	12_random	12_random_1X0	5466880	5750310	2	283430
	12_random	12_random_2X0	10687425	10729543	2	42118
	13	13_10X0	102301876	102325928	2	24052
	13	13_11X0	105053903	105269650	2	215747
	13	13_12X0	115735651	115801705	3	66054
	13	13_1X0	19231933	19315376	3	83443
	13	13_2X0	27232754	27411264	3	178510
	13	13_3X0	27444198	27940323	3	496125
	13	13_4X0	30382011	30451456	2	69445
	13	13_5X0	31942634	32080630	2	137996
	13	13_6X0	41096136	41161469	2	65333
	13	13_7X0	48594999	48895731	2	300732
	13	13_8X0	52370236	52502815	2	132579
	13	13_9X0	101592947	101635074	2	42127
	14	14_10X0	50765233	50971443	2	206210
	14	14_11X0	52696363	52755430	2	59067
	14	14_12X0	61238642	61453018	3	214376
	14	14_13X0	65392881	65434694	2	41813
	14	14_14X0	68660348	68717634	2	57286
	14	14_15X0	71288349	71608036	3	319687
	14	14_16X0	74702328	74780772	3	78444
	14	14_17X0	76066639	76285852	2	219213
	14	14_18X0	76441996	76715412	3	273416
	14	14_19X0	95638032	96011119	11	373087
	14	14_1X0	18932072	19560027	8	627955
	14	14_20X0	97092466	97118433	2	25967
	14	14_21X0	97628615	97678713	2	50098
	14	14_22X0	101810162	101851261	2	41099
	14	14_23X0	104761442	104798731	2	37289
	14	14_24X0	108297670	108769425	12	471755
	14	14_24X1	107387381	108206835	10	819454
	14	14_2X0	19816078	20402743	13	586665
	14	14_3X0	20961698	22029927	35	1068229
	14	14_4X0	22496684	22517937	2	21253
	14	14_5X0	23008940	23022259	2	13319
	14	14_6X0	23092121	23817133	7	725012
	14	14_7X0	23878130	23916776	2	38646
	14	14_8X0	23987463	24120122	4	132659

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	14	14_9X0	45339838	45371470	3	31632
	14_random	14_random_1X0	395447	596169	3	200722
	14_random	14_random_2X0	2296615	2406988	2	110373
	14_random	14_random_3X0	5406193	5988396	5	582203
	14_random	14_random_4X0	7097176	7305495	2	208319
	15	15_10X0	41515683	41588028	4	72345
	15	15_11X0	45278252	45505261	2	227009
	15	15_12X0	47031261	47154576	2	123315
	15	15_13X0	48378343	48537610	2	159267
	15	15_14X0	62438231	62520350	2	82119
	15	15_15X0	70958354	71117830	2	159476
	15	15_16X0	71566981	71614540	2	47559
	15	15_17X0	72024756	72253705	5	228949
	15	15_18X0	72255014	72434586	2	179572
	15	15_19X0	76190877	76264482	3	73605
	15	15_1X0	18725488	18755075	3	29587
	15	15_20X0	79973383	80073268	2	99885
	15	15_21X0	85246245	85296682	2	50437
	15	15_22X0	85483654	85545030	2	61376
	15	15_23X0	99078403	99138272	2	59869
	15	15_2X0	19047271	19129955	2	82684
	15	15_3X0	36519055	36643680	2	124625
	15	15_4X0	37945138	38021010	2	75872
	15	15_5X0	38227314	38240420	2	13106
	15	15_6X0	38292898	38642068	4	349170
	15	15_7X0	39650773	39755075	3	104302
	15	15_8X0	39987205	40144086	4	156881
	15	15_9X0	41311248	41496420	2	185172
	15_random	15_random_1X0	4607639	4811350	2	203711
	15_random	15_random_2X0	9435649	9639154	2	203505
	16	16_10X0	19833475	20219011	5	385536
	16	16_11X0	22281779	22497658	2	215879
	16	16_12X0	22824449	22859895	2	35446
	16	16_13X0	26589072	26698293	2	109221
	16	16_14X0	28015878	28075909	2	60031
	16	16_15X0	28150371	28302541	2	152170
	16	16_16X0	29428158	29686837	9	258679
	16	16_16X1	28909706	29303131	7	393425
	16	16_17X0	29805497	29945066	2	139569
	16	16_18X0	35018150	35171367	2	153217
	16	16_19X0	43009650	43106133	3	96483
	16	16_1X0	152493	172577	3	20084
	16	16_20X0	43812381	43932098	5	119717
	16	16_21X0	44619433	44652631	2	33198
	16	16_22X0	44797321	44885075	3	87754
	16	16_23X0	53838250	53984158	3	145908
	16	16_24X0	54286490	54366825	3	80335
	16	16_25X0	54567522	54751360	2	183838
	16	16_26X0	55361933	55381336	2	19403
	16	16_27X0	60351592	60388983	2	37391
	16	16_28X0	62528313	62568641	2	40328
	16	16_29X0	62846238	62906900	2	60662
	16	16_2X0	609482	715297	4	105815
	16	16_30X0	74219508	74291161	2	71653
	16	16_31X0	76080675	76289385	2	208710

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	16	16_3X0	1240055	1248684	2	8629
	16	16_4X0	2027108	2286345	4	259237
	16	16_5X0	2809504	3033720	4	224216
	16	16_6X0	3153351	3156011	2	2660
	16	16_7X0	3236561	3524232	9	287671
	16	16_8X0	16106107	16308465	2	202358
	16	16_9X0	19729630	19771699	2	42069
	16_random	16_random_1X0	11943690	11950650	2	6960
	16_random	16_random_2X0	12024409	12036695	3	12286
	17	17_10X0	8058330	8138427	2	80097
	17	17_11X0	8850411	9009384	2	158973
	17	17_12X0	10370079	10653164	4	283085
	17	17_13X0	16891020	17013936	2	122916
	17	17_14X0	23345091	23605235	4	260144
	17	17_15X0	28988984	29102765	5	113781
	17	17_16X0	29992788	30254399	4	261611
	17	17_17X0	30648601	30886777	7	238176
	17	17_18X0	31139885	31213662	2	73777
	17	17_19X0	33869318	33938196	3	68878
	17	17_1X0	1604475	1639220	2	34745
	17	17_20X0	41999374	42026240	2	26866
	17	17_21X0	42071423	42171598	2	100175
	17	17_22X0	42877526	42997167	2	119641
	17	17_23X0	45556173	45665207	2	109034
	17	17_24X0	46544551	46580801	3	36250
	17	17_25X0	47059554	47248009	3	188455
	17	17_26X0	47291903	47429029	2	137126
	17	17_27X0	47432523	47524617	2	92094
	17	17_28X0	47590367	47619694	2	29327
	17	17_29X0	47845693	48063136	12	217443
	17	17_2X0	3132555	3328404	5	195849
	17	17_2X1	2979361	3064452	4	85091
	17	17_30X0	48101154	48375069	13	273915
	17	17_31X0	48409478	48415520	2	6042
	17	17_32X0	48425538	48758690	10	333152
	17	17_33X0	53321899	53451726	2	129827
	17	17_34X0	56570669	56761442	3	190773
	17	17_35X0	58596129	58766760	2	170631
	17	17_36X0	59862456	59974991	2	112535
	17	17_37X0	64503367	64702711	5	199344
	17	17_38X0	64924863	64951032	2	26169
	17	17_39X0	65903288	65926771	2	23483
	17	17_3X0	3403890	3486031	2	82141
	17	17_40X0	66628431	66721779	2	93348
	17	17_41X0	68152394	68181954	2	29560
	17	17_42X0	72046745	72058938	2	12193
	17	17_43X0	72825979	72866237	2	40258
	17	17_4X0	3614719	3773174	2	158455
	17	17_5X0	4347129	4454471	2	107342
	17	17_6X0	5042773	5148690	3	105917
	17	17_7X0	7041627	7048790	2	7163
	17	17_8X0	7086413	7190723	3	104310
	17	17_9X0	7438079	7445052	2	6973
	17_random	17_random_10X0	16338439	16341851	2	3412
	17_random	17_random_11X0	16591194	16610706	2	19512

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	17_random	17_random_12X0	17278814	17297836	2	19022
	17_random	17_random_1X0	9426	54502	2	45076
	17_random	17_random_2X0	798551	882014	2	83463
	17_random	17_random_3X0	3532793	3642991	2	110198
	17_random	17_random_4X0	7267820	7487295	4	219475
	17_random	17_random_5X0	11710134	11792062	2	81928
	17_random	17_random_6X0	12412159	12633644	9	221485
	17_random	17_random_7X0	13929441	13955285	2	25844
	17_random	17_random_8X0	14094141	14130013	2	35872
	18	18_1X0	29266653	29293304	2	26651
	18	18_2X0	42955477	43533940	7	578463
	18	18_3X0	44062098	44178135	2	116037
	18	18_4X0	47338997	47477316	5	138319
	18	18_5X0	58101965	58228261	2	126296
	18	18_6X0	76307217	76866635	11	559418
	18	18_7X0	88090650	88182515	2	91865
	19	19_10X0	7011340	7124412	2	113072
	19	19_11X0	7887245	7937726	3	50481
	19	19_12X0	9276014	9419561	5	143547
	19	19_13X0	9515168	9724054	5	208886
	19	19_13X1	9753814	9975439	4	221625
	19	19_14X0	10293309	10413146	2	119837
	19	19_15X0	10461417	10485293	2	23876
	19	19_16X0	12193031	12475643	6	282612
	19	19_16X1	11989101	12126200	4	137099
	19	19_16X2	12638727	12834172	4	195445
	19	19_16X3	11662693	11905627	3	242934
	19	19_17X0	12878745	12908044	2	29299
	19	19_18X0	14520945	14619580	2	98635
	19	19_19X0	14774802	14954977	2	180175
	19	19_1X0	516900	718588	5	201688
	19	19_20X0	15004902	15243315	6	238413
	19	19_21X0	15698733	15761088	2	62355
	19	19_22X0	15999134	16062077	2	62943
	19	19_23X0	19636847	19686847	2	50000
	19	19_24X0	20102737	20248034	3	145297
	19	19_25X0	20435906	20560863	2	124957
	19	19_26X0	20993308	21153556	2	160248
	19	19_27X0	21799339	21910459	2	111120
	19	19_28X0	22430141	22647945	3	217804
	19	19_29X0	23761508	23849336	2	87828
	19	19_30X0	35426008	35465010	2	39002
	19	19_31X0	35666839	35703509	2	36670
	19	19_32X0	35868613	35921004	3	52391
	19	19_33X0	36020037	36120600	3	100563
	19	19_34X0	36125401	36232721	4	107320
	19	19_35X0	36645038	36864736	7	219698
	19	19_36X0	52550995	52794552	12	243557
	19	19_36X1	52861744	53179365	12	317621
	19	19_36X10	60091898	60335134	6	243236
	19	19_36X11	37266243	37592786	7	326543
	19	19_36X12	44967849	45267199	7	299350
	19	19_36X13	55089375	55333818	6	244443
	19	19_36X14	50559816	50610629	7	50813
	19	19_36X15	45481155	45725951	6	244796

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	19	19_36X16	38199723	38412041	4	212318
	19	19_36X17	42795567	43144207	7	348640
	19	19_36X18	51365775	51612710	6	246935
	19	19_36X19	52324526	52466827	6	142301
	19	19_36X2	60363299	60652479	10	289180
	19	19_36X20	37636013	37823292	4	187279
	19	19_36X21	42043572	42352179	5	308607
	19	19_36X22	38419587	38604534	4	184947
	19	19_36X23	57384137	57458269	5	74132
	19	19_36X24	57620300	57870377	5	250077
	19	19_36X25	55683046	55886672	5	203626
	19	19_36X26	59754374	60047372	6	292998
	19	19_36X27	37891034	38188832	4	297798
	19	19_36X28	41375735	41585923	4	210188
	19	19_36X29	57932688	58193349	5	260661
	19	19_36X3	59236059	59724302	13	488243
	19	19_36X30	45736837	46128001	6	391164
	19	19_36X31	56315970	56460125	5	144155
	19	19_36X32	53757679	53922241	3	164562
	19	19_36X33	41145988	41362519	3	216531
	19	19_36X34	39646383	39709872	3	63489
	19	19_36X35	36943163	37226587	4	283424
	19	19_36X36	54000908	54286697	4	285789
	19	19_36X37	47343715	47547060	4	203345
	19	19_36X38	56027739	56176192	4	148453
	19	19_36X39	43353367	43507607	3	154240
	19	19_36X4	56575326	56890693	9	315367
	19	19_36X40	40170429	40234894	3	64465
	19	19_36X41	39850800	39957520	2	106720
	19	19_36X42	54374962	54553305	2	178343
	19	19_36X43	43788425	43878016	2	89591
	19	19_36X44	58864284	59189888	2	325604
	19	19_36X5	45243650	45436445	5	192795
	19	19_36X6	53318484	53658444	9	339960
	19	19_36X7	54757016	54997145	8	240129
	19	19_36X8	58369057	58675812	9	306755
	19	19_36X9	44580297	44912314	9	332017
	19	19_3X0	1941574	2026382	2	84808
	19	19_4X0	2843767	2968363	2	124596
	19	19_5X0	3004483	3057424	2	52941
	19	19_6X0	3127290	3199535	2	72245
	19	19_7X0	4482895	4518368	2	35473
	19	19_8X0	5602142	5683866	2	81724
	19	19_9X0	5840771	5856314	2	15543
	19_random	19_random_1X0	2937949	3049464	2	111515
	19_random	19_random_2X0	3239639	3482269	3	242630
	19_random	19_random_3X0	3613511	3753610	2	140099
	19_random	19_random_4X0	4652628	4819844	2	167216
	19_random	19_random_5X0	5600629	5650548	2	49919
	19_random	19_random_6X0	5984497	6027201	2	42704
	19_random	19_random_7X0	6398700	6518661	2	119961
	19_random	19_random_8X0	7106418	7168780	2	62362
	19_random	19_random_9X0	7427867	7553615	2	125748
	1_random	1_random_1X0	26099087	27283100	4	1184013
	1_random	1_random_2X0	27856178	29736317	7	1880139

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	20	20_10X0	42607397	42780952	4	173555
	20	20_11X0	43048426	43364709	6	316283
	20	20_12X0	43434327	43463006	2	28679
	20	20_13X0	44080914	44101578	2	20664
	20	20_14X0	47662447	47712554	2	50107
	20	20_15X0	54588055	54590534	2	2479
	20	20_16X0	61940714	61958280	2	17566
	20	20_17X0	62539647	62554164	2	14517
	20	20_1X0	16480992	16611464	2	130472
	20	20_2X0	18228117	18267019	2	38902
	20	20_3X0	18631172	18971871	8	340699
	20	20_4X0	25218228	25232173	2	13945
	20	20_5X0	25896216	26031406	2	135190
	20	20_6X0	26692686	26909947	5	217261
	20	20_7X0	27886459	28059846	2	173387
	20	20_8X0	28157973	28591882	13	433909
	20	20_9X0	32882308	32974566	2	92258
	20_random	20_random_1X0	3245792	3474136	2	228344
	20_random	20_random_2X0	5081233	5147944	3	66711
	21	21_10X0	45115703	45273169	2	157466
	21	21_11X0	46076718	46233699	9	156981
	21	21_1X0	27329279	27456498	2	127219
	21	21_2X0	31025396	31173617	7	148221
	21	21_3X0	31223832	31391014	8	167182
	21	21_4X0	34211607	34426995	3	215388
	21	21_5X0	35399246	35475696	2	76450
	21	21_6X0	37097577	37176997	2	79420
	21	21_7X0	42531367	42639527	2	108160
	21	21_8X0	43006463	43216723	2	210260
	21	21_9X0	43603927	43681525	3	77598
	21_random	21_random_1X0	2641418	2656484	2	15066
	22	22_10X0	36492778	36548750	2	55972
	22	22_11X0	37402620	37433056	2	30436
	22	22_12X0	37790371	37869090	2	78719
	22	22_13X0	38353490	38508752	2	155262
	22	22_14X0	39243804	39314753	2	70949
	22	22_15X0	45921781	45941925	2	20144
	22	22_1X0	17266867	17429864	7	162997
	22	22_2X0	17494726	17522254	2	27528
	22	22_4X0	18156239	18221823	2	65584
	22	22_5X0	19120411	19158203	2	37792
	22	22_6X0	25339858	25453732	3	113874
	22	22_7X0	31273450	31335417	2	61967
	22	22_8X0	32137266	32156504	2	19238
	22	22_9X0	34184902	34297113	4	112211
	2a	2a_10X0	24462674	24500202	2	37528
	2a	2a_11X0	31318048	31469923	4	151875
	2a	2a_12X0	39957483	39983828	2	26345
	2a	2a_13X0	41711365	41752036	2	40671
	2a	2a_14X0	64172547	64467828	2	295281
	2a	2a_15X0	66360786	66428025	2	67239
	2a	2a_16X0	67560083	67631990	2	71907
	2a	2a_17X0	72260990	72530580	2	269590
	2a	2a_18X0	74468046	74686263	2	218217
	2a	2a_19X0	84518186	84750934	2	232748

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	2a	2a_1X0	3396303	3465868	3	69565
	2a	2a_20X0	109247456	109262765	2	15309
	2a	2a_2X0	6826444	6890999	2	64555
	2a	2a_3X0	9619254	10128259	6	509005
	2a	2a_4X0	10151152	10394110	2	242958
	2a	2a_5X0	16304549	16723487	4	418938
	2a	2a_6X0	20760225	21154139	9	393914
	2a	2a_7X0	21233354	21270562	2	37208
	2a	2a_8X0	21450437	22297377	23	846940
	2a	2a_9X0	23221472	23298616	2	77144
	2a_random	2a_random_1X0	39391	74755	2	35364
	2a_random	2a_random_2X0	80692	150400	2	69708
	2a_random	2a_random_3X0	11133355	11300445	2	167090
	2a_random	2a_random_4X0	12318454	12333986	2	15532
	2b	2b_10X0	66234263	66348235	9	113972
	2b	2b_11X0	67752940	67824744	2	71804
	2b	2b_12X0	79725835	79929327	2	203492
	2b	2b_13X0	81753792	81926730	2	172938
	2b	2b_14X0	92379117	92547147	3	168030
	2b	2b_15X0	93420367	93551318	2	130951
	2b	2b_16X0	95145817	95403282	3	257465
	2b	2b_17X0	108412227	108470342	2	58115
	2b	2b_18X0	109978878	110107996	3	129118
	2b	2b_19X0	110709578	110745105	2	35527
	2b	2b_1X0	15379699	15601347	2	221648
	2b	2b_20X0	124766111	124861685	2	95574
	2b	2b_21X0	124933737	124953746	2	20009
	2b	2b_2X0	21143074	21169816	3	26742
	2b	2b_3X0	38616628	38876133	2	259505
	2b	2b_4X0	40646455	40690503	2	44048
	2b	2b_5X0	47189490	47483077	2	293587
	2b	2b_6X0	51786827	52039877	2	253050
	2b	2b_7X0	55839086	56356760	3	517674
	2b	2b_8X0	59548316	59812380	2	264064
	2b	2b_9X0	62210746	62228558	2	17812
	3	3_10X0	60006083	60037508	2	31425
	3	3_11X0	93501666	93558742	3	57076
	3	3_12X0	94075144	94426972	5	351828
	3	3_13X0	94485152	94616493	4	131341
	3	3_14X0	96170017	96190403	2	20386
	3	3_15X0	96213142	96335426	2	122284
	3	3_16X0	97177218	97381880	2	204662
	3	3_17X0	100193200	100672889	5	479689
	3	3_18X0	101870988	102157613	7	286625
	3	3_19X0	103733983	104001270	2	267287
	3	3_1X0	3099021	3324283	2	225262
	3	3_20X0	106184861	106242388	3	57527
	3	3_21X0	107350567	107419255	2	68688
	3	3_22X0	107735264	108155670	3	420406
	3	3_23X0	108220396	108527802	4	307406
	3	3_24X0	114465907	114734168	3	268261
	3	3_25X0	133681755	133704325	2	22570
	3	3_26X0	136209625	136301425	2	91800
	3	3_27X0	141513176	141552962	3	39786
	3	3_28X0	141971361	142063204	2	91843

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	3	3_29X0	148904195	149236701	3	332506
	3	3_2X0	7572783	7704482	2	131699
	3	3_30X0	151609353	151675314	2	65961
	3	3_31X0	152101707	152305786	3	204079
	3	3_32X0	154199244	154243872	3	44628
	3	3_33X0	161864020	161932465	2	68445
	3	3_34X0	173354576	173432074	3	77498
	3	3_35X0	187287741	187514032	2	226291
	3	3_36X0	187884223	187966126	3	81903
	3	3_37X0	190425158	190529553	3	104395
	3	3_38X0	194093947	194221255	2	127308
	3	3_39X0	197209548	197511474	2	301926
	3	3_3X0	9952479	10169756	3	217277
	3	3_40X0	198346539	198403641	3	57102
	3	3_4X0	19557166	20007210	3	450044
	3	3_5X0	20337435	20515530	2	178095
	3	3_6X0	23507511	23548048	2	40537
	3	3_7X0	24306699	24519954	2	213255
	3	3_8X0	34768076	35119236	6	351160
	3	3_9X0	58857067	59082826	2	225759
	3_random	3_random_1X0	18639529	18656003	2	16474
	4	4_10X0	71480573	72695560	8	1214987
	4	4_11X0	72766571	72905269	2	138698
	4	4_12X0	73060535	73112999	3	52464
	4	4_13X0	73427152	73470045	3	42893
	4	4_14X0	76650883	76759521	3	108638
	4	4_15X0	77011891	77378866	10	366975
	4	4_16X0	77583380	78082773	4	499393
	4	4_17X0	79329282	79361879	3	32597
	4	4_18X0	80374925	80507003	2	132078
	4	4_19X0	85857444	85935867	2	78423
	4	4_1X0	39763	457093	4	417330
	4	4_20X0	91189192	91285909	2	96717
	4	4_21X0	103343984	103742325	5	398341
	4	4_22X0	107357809	107532144	2	174335
	4	4_23X0	159706987	159710062	2	3075
	4	4_24X0	160420038	160476696	3	56658
	4	4_25X0	161558664	161679109	2	120445
	4	4_26X0	169667186	169694491	2	27305
	4	4_27X0	171410027	172213408	4	803381
	4	4_28X0	174830583	175100547	2	269964
	4	4_29X0	191487273	191490170	2	2897
	4	4_2X0	6362695	6431076	3	68381
	4	4_30X0	194467196	194562760	2	95564
	4	4_31X0	196415863	196480303	2	64440
	4	4_3X0	7382834	7699398	2	316564
	4	4_4X0	15434275	15563433	2	129158
	4	4_5X0	39822110	39883241	3	61131
	4	4_6X0	49568397	49676444	2	108047
	4	4_7X0	50345417	50425636	2	80219
	4	4_8X0	50811419	50956431	2	145012
	4	4_9X0	50999792	51729583	5	729791
	4_random	4_random_1X0	3464664	3715729	2	251065
	5	5_10X0	76738265	76967691	3	229426
	5	5_11X0	79207953	79273758	2	65805

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	5	5_12X0	83695837	83907869	2	212032
	5	5_13X0	97461804	97645577	2	183773
	5	5_14X0	133752322	133851866	2	99544
	5	5_15X0	142164982	142192020	2	27038
	5	5_16X0	142299496	143493885	22	1194389
	5	5_17X0	149568782	150232065	6	663283
	5	5_18X0	150352662	150737447	2	384785
	5	5_19X0	152000386	152250123	3	249737
	5	5_1X0	1208501	1471215	3	262714
	5	5_20X0	153321641	153558195	3	236554
	5	5_21X0	159137805	159369167	3	231362
	5	5_22X0	179763132	180139214	2	376082
	5	5_23X0	181299033	181677914	5	378881
	5	5_24X0	183278599	183769099	4	490500
	5	5_2X0	15428269	15550725	2	122456
	5	5_3X0	56282528	56369891	2	87363
	5	5_4X0	56518781	57085892	3	567111
	5	5_5X0	57124242	57269709	2	145467
	5	5_6X0	71146833	71293106	2	146273
	5	5_7X0	71489762	71611767	3	122005
	5	5_8X0	72257670	72398089	2	140419
	5	5_9X0	73899735	74032026	2	132291
	6	6_10X0	31837499	32003335	3	165836
	6	6_11X0	32044398	32163969	5	119571
	6	6_12X0	32174774	32218001	2	43227
	6	6_13X0	32310281	32318293	2	8012
	6	6_14X0	32427810	32455767	2	27957
	6	6_15X0	32672457	33220331	5	547874
	6	6_16X0	35785343	35806269	3	20926
	6	6_17X0	35840871	36574746	4	733875
	6	6_18X0	39225240	39356220	3	130980
	6	6_19X0	41234657	41432095	5	197438
	6	6_1X0	2825816	2969339	3	143523
	6	6_20X0	42289932	42311433	2	21501
	6	6_21X0	43114270	43365834	4	251564
	6	6_22X0	46122030	46163874	2	41844
	6	6_23X0	46851646	47052531	2	200885
	6	6_24X0	47697289	47744110	2	46821
	6	6_25X0	49833779	50001332	3	167553
	6	6_26X0	50110159	50229419	5	119260
	6	6_27X0	50991192	51129663	2	138471
	6	6_28X0	52537100	52594591	2	57491
	6	6_29X0	53158971	53393601	4	234630
	6	6_2X0	26798935	27630964	36	832029
	6	6_30X0	64833357	65320951	4	487594
	6	6_31X0	73721083	73875712	4	154629
	6	6_32X0	96894605	96926400	2	31795
	6	6_33X0	106988947	107054013	2	65066
	6	6_34X0	118528749	118555835	2	27086
	6	6_35X0	118740734	118837453	3	96719
	6	6_36X0	135410500	135474766	3	64266
	6	6_37X0	135513837	135589140	3	75303
	6	6_38X0	139958374	140128612	3	170238
	6	6_39X0	153033345	153196309	5	162964
	6	6_3X0	28025661	28040029	2	14368

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	6	6_40X0	163643173	163980617	3	337444
	6	6_41X0	164017530	164295362	2	277832
	6	6_42X0	170753016	170774761	2	21745
	6	6_43X0	174164505	174181924	2	17419
	6	6_4X0	28296933	28369006	2	72073
	6	6_5X0	28708529	28791355	14	82826
	6	6_6X0	28825973	28946893	3	120920
	6	6_7X0	28959838	29923668	12	963830
	6	6_8X0	30033706	30553016	11	519310
	6	6_9X0	30686952	31203118	4	516166
	6_cox_hap1	6_cox_hap1_1X0	161620	312005	6	150385
	6_cox_hap1	6_cox_hap1_2X0	778220	796237	2	18017
	6_cox_hap1	6_cox_hap1_3X0	1082119	1178965	2	96846
	6_cox_hap1	6_cox_hap1_4X0	1275513	1335637	2	60124
	6_cox_hap1	6_cox_hap1_5X0	1353385	1385259	4	31874
	6_cox_hap1	6_cox_hap1_6X0	1463679	1483453	2	19774
	6_cox_hap1	6_cox_hap1_7X0	1609129	1632902	2	23773
	6_qbl_hap2	6_qbl_hap2_1X0	123320	264780	2	141460
	6_random	6_random_1X0	23087	66227	2	43140
	7	7_10X0	13004250	13571714	3	567464
	7	7_11X0	14221812	14390628	4	168816
	7	7_12X0	22270735	23421138	7	1150403
	7	7_13X0	27500811	28091072	3	590261
	7	7_14X0	38231369	38265471	2	34102
	7	7_15X0	39271875	39289540	2	17665
	7	7_16X0	45770205	45864918	3	94713
	7	7_17X0	48914432	49180840	2	266408
	7	7_18X0	51101798	51322307	2	220509
	7	7_19X0	57200101	57359549	12	159448
	7	7_1X0	6645535	6711435	3	65900
	7	7_20X0	60580983	60847060	2	266077
	7	7_21X0	65449222	65484620	2	35398
	7	7_22X0	67802393	67895359	2	92966
	7	7_23X0	79006433	79025282	2	18849
	7	7_24X0	80638810	80767916	3	129106
	7	7_25X0	82179616	82199278	2	19662
	7	7_26X0	82994395	83038819	2	44424
	7	7_27X0	85960743	86050941	2	90198
	7	7_28X0	88818398	89024481	2	206083
	7	7_29X0	103837588	103982033	2	144445
	7	7_2X0	8228532	8300415	2	71883
	7	7_30X0	104103904	104307099	2	203195
	7	7_31X0	113027889	113089251	2	61362
	7	7_32X0	120628075	120752068	2	123993
	7	7_33X0	125553651	125612643	2	58992
	7	7_34X0	127209144	127345287	3	136143
	7	7_35X0	131544168	131647055	2	102887
	7	7_36X0	136682747	136778225	2	95478
	7	7_37X0	139062234	139262379	2	200145
	7	7_38X0	139291117	139442037	2	150920
	7	7_39X0	139591793	139635199	2	43406
	7	7_3X0	8376263	8542771	4	166508
	7	7_40X0	139699927	140302779	25	602852
	7	7_41X0	140706635	141673530	15	966895
	7	7_42X0	146817369	147594846	7	777477

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	7	7_43X0	147977427	148007975	2	30548
	7	7_44X0	148064855	148401880	7	337025
	7	7_45X0	148712011	148914946	2	202935
	7	7_46X0	149624907	149790231	2	165324
	7	7_4X0	8558677	8846993	4	288316
	7	7_5X0	9004108	9065109	3	61001
	7	7_6X0	9407780	9451815	2	44035
	7	7_7X0	10265873	10273679	2	7806
	7	7_8X0	10722886	10755668	2	32782
	7	7_9X0	12436355	12476157	2	39802
	7_random	7_random_1X0	10285383	10449747	2	164364
	7_random	7_random_2X0	17218282	17224106	2	5824
	7_random	7_random_3X0	18237122	18245692	2	8570
	8	8_10X0	40347981	40460346	2	112365
	8	8_11X0	43225407	43298173	2	72766
	8	8_12X0	54576530	54902868	2	326338
	8	8_13X0	57719857	57892692	2	172835
	8	8_14X0	77378823	77618423	2	239600
	8	8_15X0	84191798	84446568	5	254770
	8	8_16X0	88240691	88617610	4	376919
	8	8_17X0	151027869	151585278	10	557409
	8	8_18X0	151636958	152056637	3	419679
	8	8_19X0	152465530	152626837	2	161307
	8	8_1X0	7112116	7856830	14	744714
	8	8_20X0	153084612	153343637	7	259025
	8	8_2X0	9373797	9416503	2	42706
	8	8_3X0	17609597	17813172	2	203575
	8	8_4X0	22603729	22860209	3	256480
	8	8_5X0	23329852	23358051	2	28199
	8	8_6X0	23974113	24219265	3	245152
	8	8_7X0	24679313	24723879	2	44566
	8	8_8X0	38513536	38557477	2	43941
	8	8_9X0	39297999	39607187	2	309188
	8_random	8_random_1X0	9057416	9076836	2	19420
	9	9_10X0	88141275	88271546	4	130271
	9	9_11X0	90486942	90572114	2	85172
	9	9_12X0	100698259	100898942	3	200683
	9	9_13X0	100928986	100950767	2	21781
	9	9_14X0	105219517	105228504	2	8987
	9	9_15X0	106621931	106718024	2	96093
	9	9_16X0	109624117	109739684	2	115567
	9	9_17X0	109751351	109819111	2	67760
	9	9_18X0	119300112	119610713	5	310601
	9	9_19X0	119760579	119769982	2	9403
	9	9_1X0	25542296	25569549	2	27253
	9	9_20X0	121372152	121668149	2	295997
	9	9_21X0	123726708	123774888	2	48180
	9	9_22X0	124806440	124839675	2	33235
	9	9_23X0	125192722	125480877	2	288155
	9	9_24X0	126711146	126732771	2	21625
	9	9_25X0	129303450	129414404	2	110954
	9	9_26X0	130136556	130255350	3	118794
	9	9_27X0	130290478	130307710	2	17232
	9	9_28X0	130371388	130525940	2	154552
	9	9_29X0	132504094	132524781	2	20687

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	9	9_2X0	25750625	25954532	2	203907
	9	9_30X0	133733366	133984495	7	251129
	9	9_3X0	26909610	26927831	2	18221
	9	9_4X0	26956024	27068359	2	112335
	9	9_5X0	28226064	28277540	2	51476
	9	9_6X0	40651810	41099976	11	448166
	9	9_7X0	57250597	57372268	2	121671
	9	9_8X0	57637563	57803411	3	165848
	9	9_9X0	62018480	62238066	2	219586
	9_random	9_random_1X0	6830212	6879549	2	49337
	Un	Un_1X0	2095221	2106268	2	11047
	Un	Un_2X0	6844754	6912379	3	67625
	Un	Un_3X0	7053964	7100713	2	46749
	Un	Un_4X0	7517253	8180344	5	663091
	Un	Un_5X0	8770117	8799925	2	29808
	Un	Un_6X0	11978513	11996003	2	17490
	Un	Un_7X0	12113288	12141809	2	28521
	X	X_10X0	42031161	42308029	3	276868
	X	X_11X0	44279402	44502945	2	223543
	X	X_12X0	47214021	47311358	2	97337
	X	X_13X0	48011409	48868879	7	857470
	X	X_14X0	49087801	49154954	2	67153
	X	X_15X0	49960496	49992060	2	31564
	X	X_16X0	50109756	50362221	4	252465
	X	X_17X0	51900379	52092065	2	191686
	X	X_18X0	52192395	52281664	2	89269
	X	X_19X0	52474591	52653048	2	178457
	X	X_1X0	1104597	1246696	2	142099
	X	X_20X0	52940396	53247407	5	307011
	X	X_21X0	55707608	55753336	2	45728
	X	X_22X0	57618950	57728483	3	109533
	X	X_23X0	62905092	63088970	2	183878
	X	X_24X0	67699766	67886781	3	187015
	X	X_25X0	76426524	76886867	3	460343
	X	X_26X0	100495115	100661183	5	166068
	X	X_27X0	101512467	101675964	4	163497
	X	X_28X0	102010354	102785595	10	775241
	X	X_29X0	106789624	107076544	2	286920
	X	X_2X0	1353004	1594993	2	241989
	X	X_30X0	119134167	119176135	2	41968
	X	X_31X0	119916232	119921777	2	5545
	X	X_32X0	134436691	134919800	5	483109
	X	X_33X0	149725731	149908775	2	183044
	X	X_34X0	150640181	150813560	2	173379
	X	X_35X0	152876766	152893187	2	16421
	X	X_36X0	153553022	153644120	2	91098
	X	X_37X0	153898123	154394134	6	496011
	X	X_38X0	154597550	154793515	3	195965
	X	X_3X0	2299755	2423837	2	124082
	X	X_4X0	2530390	2747522	4	217132
	X	X_5X0	7542310	7876823	2	334513
	X	X_6X0	8525980	8775763	2	249783
	X	X_7X0	15200737	15274777	2	74040
	X	X_8X0	26372396	26399761	2	27365
	X	X_9X0	30477878	30505952	3	28074

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	X_random	X_random_1X0	8377115	8399351	2	22236
	1	1_100X0	214446659	214473741	2	27082
	1	1_101X0	215830983	215858034	2	27051
	1	1_102X0	216191886	216253311	2	61425
	1	1_103X0	216941913	216971183	3	29270
	1	1_104X0	218003018	218061693	3	58675
	1	1_105X0	218466289	218488783	2	22494
	1	1_106X0	219145680	219207182	2	61502
	1	1_107X0	219291679	219312240	2	20561
	1	1_108X0	219329574	219388009	2	58435
	1	1_109X0	220009397	220067124	3	57727
	1	1_10X0	60117804	60634546	7	516742
	1	1_10X1	59765835	60088406	5	322571
	1	1_110X0	220353356	220448989	3	95633
	1	1_111X0	221043119	221129654	2	86535
	1	1_112X0	221673590	221725656	2	52066
	1	1_113X0	222619523	222677372	2	57849
	1	1_114X0	222861777	222937633	2	75856
	1	1_115X0	224319918	224833259	5	513341
	1	1_116X0	225591969	225902676	6	310707
	1	1_117X0	227016803	227506822	13	490019
	1	1_118X0	227592635	227992623	8	399988
	1	1_119X0	227998591	228053645	2	55054
	1	1_11X0	61268248	61872975	11	604727
	1	1_11X1	62112797	62558033	5	445236
	1	1_120X0	228225398	228685348	11	459950
	1	1_121X0	229039889	229080410	2	40521
	1	1_122X0	229722768	230269929	19	547161
	1	1_122X1	230297117	230812750	17	515633
	1	1_122X2	229416489	229481119	4	64630
	1	1_122X3	229100752	229108849	2	8097
	1	1_123X0	247473292	247486202	2	12910
	1	1_124X0	248895738	248956446	2	60708
	1	1_125X0	251325923	251365385	2	39462
	1	1_126X0	252102219	252428774	6	326555
	1	1_127X0	252894663	252946170	4	51507
	1	1_128X0	256370850	256386729	2	15879
	1	1_129X0	257970345	258252448	3	282103
	1	1_12X0	62665651	62979941	4	314290
	1	1_130X0	266908369	266934303	3	25934
	1	1_131X0	267607416	267640402	2	32986
	1	1_13X0	63184734	63334176	5	149442
	1	1_14X0	63679769	64254564	7	574795
	1	1_15X0	64330106	64405149	2	75043
	1	1_16X0	64480929	64492526	3	11597
	1	1_17X0	65522118	65857872	10	335754
	1	1_18X0	65937952	66021670	2	83718
	1	1_19X0	66637868	66645214	2	7346
	1	1_1X0	699413	813517	3	114104
	1	1_20X0	66663965	66688422	2	24457
	1	1_21X0	67274428	67351707	2	77279
	1	1_22X0	67660337	67918469	3	258132
	1	1_23X0	68176904	68240174	3	63270
	1	1_24X0	69615603	69879804	6	264201
	1	1_25X0	70222964	70266809	2	43845

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_26X0	70437776	70496347	5	58571
	1	1_27X0	71167156	71447195	7	280039
	1	1_28X0	71765944	71787211	2	21267
	1	1_29X0	71844368	72230961	6	386593
	1	1_2X0	1101965	1619634	5	517669
	1	1_30X0	72338877	72468738	9	129861
	1	1_31X0	72491780	72545329	2	53549
	1	1_32X0	72882806	72899629	2	16823
	1	1_33X0	73040901	73423448	6	382547
	1	1_34X0	73610149	73868695	3	258546
	1	1_35X0	75131348	75195678	2	64330
	1	1_36X0	77745288	77789397	2	44109
	1	1_37X0	78166810	78195328	2	28518
	1	1_38X0	78622590	78663970	2	41380
	1	1_39X0	79065532	79708639	9	643107
	1	1_3X0	22319353	22596475	16	277122
	1	1_40X0	80639785	80912082	6	272297
	1	1_41X0	80920747	81163689	6	242942
	1	1_42X0	81208212	81733957	8	525745
	1	1_43X0	81952067	82003691	2	51624
	1	1_44X0	83003841	83049924	2	46083
	1	1_45X0	83103925	83523310	5	419385
	1	1_45X1	83573032	83969465	5	396433
	1	1_46X0	84328114	84491466	3	163352
	1	1_47X0	84530669	84950210	11	419541
	1	1_48X0	85112834	85201757	2	88923
	1	1_49X0	85236243	85253635	2	17392
	1	1_4X0	22614783	22661377	2	46594
	1	1_50X0	85509662	85517360	2	7698
	1	1_51X0	86073435	86534023	5	460588
	1	1_52X0	86873253	86919236	2	45983
	1	1_53X0	87019975	87047221	2	27246
	1	1_54X0	88206813	88690534	13	483721
	1	1_54X1	87563975	87915794	6	351819
	1	1_55X0	88875375	88881460	2	6085
	1	1_56X0	88910415	88950559	2	40144
	1	1_57X0	89132531	89150862	2	18331
	1	1_58X0	89268938	89309379	3	40441
	1	1_59X0	89314632	89560469	7	245837
	1	1_5X0	37802592	37932084	3	129492
	1	1_60X0	97712177	97784139	2	71962
	1	1_61X0	98398660	98455920	2	57260
	1	1_62X0	98487631	98501249	3	13618
	1	1_63X0	98513116	98570949	2	57833
	1	1_64X0	99299011	100203260	32	904249
	1	1_65X0	100859349	100887853	2	28504
	1	1_66X0	101120325	101131193	3	10868
	1	1_67X0	101249522	101277204	2	27682
	1	1_68X0	101599018	101664436	3	65418
	1	1_69X0	102941231	103014388	2	73157
	1	1_6X0	38528185	38586563	2	58378
	1	1_70X0	103298174	104024682	11	726508
	1	1_71X0	114348038	114422733	2	74695
	1	1_72X0	122981755	123019522	2	37767
	1	1_73X0	140584380	140608946	2	24566

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	1	1_74X0	141532390	141553230	2	20840
	1	1_75X0	146976975	147077568	4	100593
	1	1_76X0	148846679	148863446	2	16767
	1	1_77X0	149578357	150311257	20	732900
	1	1_78X0	156262031	156291052	2	29021
	1	1_79X0	164706276	164756039	2	49763
	1	1_7X0	47972425	48025579	2	53154
	1	1_80X0	165482155	165512744	2	30589
	1	1_81X0	165972440	166037424	2	64984
	1	1_82X0	166915045	166943592	2	28547
	1	1_83X0	166986013	167005839	2	19826
	1	1_84X0	167192810	167206614	2	13804
	1	1_85X0	167587641	168177139	32	589498
	1	1_85X1	168196862	168547009	20	350147
	1	1_85X2	168575090	168905659	20	330569
	1	1_85X3	168929814	169270329	21	340515
	1	1_85X4	169616178	170206460	31	590282
	1	1_85X5	169305232	169598835	15	293603
	1	1_86X0	170693928	171275249	23	581321
	1	1_87X0	172599610	173048604	17	448994
	1	1_87X1	172051215	172567846	18	516631
	1	1_87X2	173128094	173469699	6	341605
	1	1_88X0	189166747	189199376	2	32629
	1	1_89X0	189241593	189541224	4	299631
	1	1_8X0	48069194	48108216	2	39022
	1	1_90X0	192115966	192142854	2	26888
	1	1_91X0	199055947	199330997	7	275050
	1	1_91X1	198456699	198869009	7	412310
	1	1_92X0	199331787	199395363	3	63576
	1	1_93X0	199555722	199623670	2	67948
	1	1_94X0	201905126	201915352	2	10226
	1	1_95X0	203509731	203523358	2	13627
	1	1_96X0	212632366	213452653	24	820287
	1	1_97X0	213676954	213691943	2	14989
	1	1_98X0	213744796	213818222	5	73426
	1	1_99X0	213981845	214423881	6	442036
	1	1_9X0	48273639	48521256	3	247617
	10	10_10X0	15577249	15610826	5	33577
	10	10_11X0	31561895	31915247	6	353352
	10	10_12X0	34155515	34199279	2	43764
	10	10_13X0	34213936	34285494	3	71558
	10	10_14X0	34553284	34805813	13	252529
	10	10_14X1	34300250	34415792	6	115542
	10	10_15X0	36311343	36600469	13	289126
	10	10_16X0	36690871	36716601	2	25730
	10	10_17X0	40462701	40574444	3	111743
	10	10_18X0	43744882	43760930	2	16048
	10	10_19X0	43817947	43892604	4	74657
	10	10_1X0	12311730	12818137	16	506407
	10	10_1X1	11912543	12236107	7	323564
	10	10_20X0	44146097	44408395	10	262298
	10	10_20X1	44714212	45160268	13	446056
	10	10_20X2	44458092	44680749	8	222657
	10	10_20X3	43894429	44122978	9	228549
	10	10_21X0	45169409	45243672	3	74263

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	10	10_22X0	45258887	45335030	6	76143
	10	10_23X0	45598912	45681393	2	82481
	10	10_24X0	45810406	45851984	2	41578
	10	10_25X0	45922873	45968307	3	45434
	10	10_26X0	47490153	47546345	2	56192
	10	10_27X0	47588497	47666921	2	78424
	10	10_28X0	49135277	49196177	2	60900
	10	10_29X0	53781239	53918524	2	137285
	10	10_2X0	12885533	12939435	3	53902
	10	10_30X0	55712043	55783489	3	71446
	10	10_31X0	56286973	56300077	2	13104
	10	10_32X0	56424168	56429748	2	5580
	10	10_33X0	56824505	56831243	2	6738
	10	10_34X0	56854366	56962161	3	107795
	10	10_35X0	59799123	59860987	2	61864
	10	10_36X0	59923213	60573129	17	649916
	10	10_36X1	60728160	61161060	14	432900
	10	10_37X0	62241756	62273119	2	31363
	10	10_38X0	62361019	62428488	4	67469
	10	10_39X0	69412017	69537882	5	125865
	10	10_3X0	13000090	13004441	2	4351
	10	10_40X0	70739800	70886355	5	146555
	10	10_41X0	70926440	71014423	4	87983
	10	10_42X0	75089928	75171774	3	81846
	10	10_43X0	75186572	75202030	2	15458
	10	10_44X0	76274563	76281283	2	6720
	10	10_45X0	81648216	81666523	2	18307
	10	10_46X0	81984560	82045935	2	61375
	10	10_47X0	82220762	82252720	2	31958
	10	10_48X0	82292110	82367047	2	74937
	10	10_49X0	82800704	82833867	2	33163
	10	10_4X0	13173529	13418154	7	244625
	10	10_50X0	82937971	82963919	2	25948
	10	10_51X0	84031955	84215687	10	183732
	10	10_52X0	84891821	84960045	2	68224
	10	10_53X0	86669233	86688730	2	19497
	10	10_54X0	87142116	87459652	12	317536
	10	10_55X0	87467991	87484935	3	16944
	10	10_56X0	87500299	87753668	11	253369
	10	10_57X0	87896771	87920581	2	23810
	10	10_58X0	87950364	88172910	14	222546
	10	10_59X0	88356615	88389347	2	32732
	10	10_5X0	13839250	13892997	2	53747
	10	10_60X0	88459605	88533171	2	73566
	10	10_61X0	89030865	89075585	2	44720
	10	10_62X0	90041591	90049112	2	7521
	10	10_63X0	91200874	91222840	2	21966
	10	10_64X0	94170766	94206535	2	35769
	10	10_65X0	94280703	94365049	2	84346
	10	10_66X0	98237141	98644938	5	407797
	10	10_67X0	99388130	99439114	2	50984
	10	10_68X0	103441549	103685844	7	244295
	10	10_69X0	103874383	103898331	2	23948
	10	10_6X0	14066948	14085855	2	18907
	10	10_70X0	104654717	104681377	2	26660

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	10	10_71X0	104917862	105051965	4	134103
	10	10_72X0	105650862	105723700	2	72838
	10	10_73X0	106767161	106795299	2	28138
	10	10_74X0	108132105	108150511	2	18406
	10	10_75X0	109907759	109953066	3	45307
	10	10_76X0	110251363	110274768	2	23405
	10	10_77X0	110305928	110348652	2	42724
	10	10_78X0	110614713	110642260	2	27547
	10	10_79X0	111547196	111779063	3	231867
	10	10_7X0	14656812	14731086	7	74274
	10	10_8X0	15108072	15125408	2	17336
	10	10_9X0	15205534	15282951	2	77417
	11	11_10X0	44072043	44092877	2	20834
	11	11_11X0	45615509	45688560	3	73051
	11	11_12X0	54513102	54567357	3	54255
	11	11_13X0	55474820	55498163	2	23343
	11	11_14X0	60253608	60882379	5	628771
	11	11_15X0	61462994	61530837	2	67843
	11	11_16X0	67308943	67566972	8	258029
	11	11_17X0	71747565	71757801	2	10236
	11	11_18X0	74014983	74067413	3	52430
	11	11_19X0	81424252	81717521	4	293269
	11	11_1X0	23057512	23063414	2	5902
	11	11_20X0	85014910	85633756	11	618846
	11	11_20X1	85992696	86096397	3	103701
	11	11_21X0	87221613	87240833	2	19220
	11	11_22X0	88376256	88391370	2	15114
	11	11_2X0	28477184	28527890	2	50706
	11	11_3X0	28582897	28781270	11	198373
	11	11_4X0	28842087	29087882	5	245795
	11	11_5X0	31389514	31430517	2	41003
	11	11_6X0	32440237	32508420	2	68183
	11	11_7X0	33812662	33917674	4	105012
	11	11_8X0	42936028	43005097	2	69069
	11	11_9X0	43143882	44003664	25	859782
	12	12_10X0	18106629	18167517	2	60888
	12	12_11X0	18220104	18303041	4	82937
	12	12_12X0	19231092	19294888	3	63796
	12	12_13X0	21661512	21950085	5	288573
	12	12_14X0	23654996	23682381	2	27385
	12	12_15X0	38144855	38161546	2	16691
	12	12_16X0	39308025	39388021	2	79996
	12	12_17X0	41620591	41647205	2	26614
	12	12_18X0	42343123	42376426	2	33303
	12	12_19X0	46989951	47039542	2	49591
	12	12_1X0	2233778	2413324	5	179546
	12	12_20X0	47444104	47482961	2	38857
	12	12_21X0	51846097	51934702	2	88605
	12	12_22X0	52637000	52674737	2	37737
	12	12_2X0	2658547	2954562	7	296015
	12	12_3X0	6671940	6740714	2	68774
	12	12_4X0	9464026	9501213	2	37187
	12	12_5X0	11127754	11187380	3	59626
	12	12_6X0	11252296	11294058	2	41762
	12	12_7X0	11775090	11808977	2	33887

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	12	12_8X0	13323547	13442276	6	118729
	12	12_9X0	17309834	17358617	3	48783
	13	13_10X0	50977431	51030802	2	53371
	13	13_11X0	52625441	52680794	2	55353
	13	13_12X0	56598979	56958561	5	359582
	13	13_13X0	71280652	71324016	3	43364
	13	13_14X0	73671412	73735339	2	63927
	13	13_15X0	84682409	84795081	3	112672
	13	13_16X0	88557860	88622788	2	64928
	13	13_17X0	89285998	91228935	25	1942937
	13	13_18X0	91481461	92505351	20	1023890
	13	13_19X0	99173548	99217786	2	44238
	13	13_1X0	9464906	9480933	2	16027
	13	13_20X0	99271390	99346543	2	75153
	13	13_21X0	99369126	99395842	2	26716
	13	13_22X0	100189236	100214895	2	25659
	13	13_23X0	102698546	102756174	2	57628
	13	13_2X0	26903052	27912716	10	1009664
	13	13_3X0	47126741	47397890	4	271149
	13	13_4X0	47539231	47617064	3	77833
	13	13_5X0	47618452	47743392	4	124940
	13	13_6X0	47785161	48056394	3	271233
	13	13_7X0	48367307	48400632	2	33325
	13	13_8X0	48893412	49243316	3	349904
	13	13_9X0	50735599	50755046	2	19447
	14	14_10X0	18731378	18862407	6	131029
	14	14_11X0	21901441	21979510	3	78069
	14	14_12X0	22072024	22155231	3	83207
	14	14_13X0	22417206	22952822	5	535616
	14	14_14X0	23089699	23461384	5	371685
	14	14_15X0	37435654	37475418	2	39764
	14	14_16X0	45033309	45081991	2	48682
	14	14_17X0	78939903	78973883	2	33980
	14	14_18X0	82818663	82869792	2	51129
	14	14_19X0	84306466	84355528	2	49062
	14	14_1X0	759398	1785698	10	1026300
	14	14_1X1	182678	338157	3	155479
	14	14_20X0	87448692	87465374	2	16682
	14	14_2X0	5086054	5428757	3	342703
	14	14_3X0	6140190	6223394	3	83204
	14	14_4X0	6281336	6369666	3	88330
	14	14_5X0	7073445	7128218	2	54773
	14	14_6X0	11199404	11274578	2	75174
	14	14_7X0	15424044	15952083	7	528039
	14	14_8X0	17195014	17234712	3	39698
	14	14_9X0	18231860	18634309	4	402449
	15	15_10X0	34446878	34453171	2	6293
	15	15_11X0	34520142	34561727	2	41585
	15	15_12X0	34601037	35417273	18	816236
	15	15_13X0	35504845	36472346	24	967501
	15	15_14X0	36878811	36947107	2	68296
	15	15_15X0	44799334	44860604	2	61270
	15	15_16X0	46159511	46237472	5	77961
	15	15_17X0	47433566	47474968	3	41402
	15	15_18X0	51034620	51068421	2	33801

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	15	15_19X0	61692116	61733742	2	41626
	15	15_1X0	2526368	2806536	5	280168
	15	15_20X0	108323179	108376221	2	53042
	15	15_21X0	108891992	108912737	2	20745
	15	15_2X0	5319916	5907758	6	587842
	15	15_3X0	12128252	12364324	3	236072
	15	15_4X0	25502019	25521393	2	19374
	15	15_5X0	26780039	27653317	23	873278
	15	15_5X1	27933211	28323507	19	390296
	15	15_6X0	28445850	28935269	8	489419
	15	15_7X0	30621374	31282563	27	661189
	15	15_7X1	29064750	29922522	30	857772
	15	15_7X2	31395772	32856423	33	1460651
	15	15_7X3	29976103	30612842	20	636739
	15	15_8X0	33605654	33656089	2	50435
	15	15_9X0	34256071	34269851	2	13780
	16	16_10X0	21061237	21089508	2	28271
	16	16_11X0	21399553	21524286	5	124733
	16	16_12X0	24779481	24805079	2	25598
	16	16_13X0	47534349	47537476	2	3127
	16	16_14X0	50152008	50201698	2	49690
	16	16_15X0	51730452	51752936	2	22484
	16	16_16X0	52116975	52177586	2	60611
	16	16_17X0	68586235	68635522	2	49287
	16	16_18X0	68860018	68913628	2	53610
	16	16_19X0	71810377	72210900	5	400523
	16	16_1X0	43814	70786	2	26972
	16	16_20X0	72388880	72403179	2	14299
	16	16_21X0	74785281	74848113	2	62832
	16	16_22X0	75027131	75705489	20	678358
	16	16_23X0	81784348	81834945	3	50597
	16	16_2X0	3754968	3790928	2	35960
	16	16_3X0	4997461	5087844	2	90383
	16	16_4X0	6970382	7026540	3	56158
	16	16_5X0	10250461	10273239	2	22778
	16	16_6X0	14306804	14324208	2	17404
	16	16_7X0	18635445	18667968	2	32523
	16	16_8X0	18716019	18757918	3	41899
	16	16_9X0	19223087	19643236	4	420149
	17	17_10X0	32516679	32911495	5	394816
	17	17_11X0	38849932	39278695	9	428763
	17	17_12X0	39411732	39824299	9	412567
	17	17_13X0	42695322	42927704	7	232382
	17	17_14X0	43423111	43798383	12	375272
	17	17_15X0	43868191	43909627	2	41436
	17	17_16X0	44519860	44841382	19	321522
	17	17_17X0	45078556	45285060	5	206504
	17	17_18X0	45670284	45802478	4	132194
	17	17_19X0	57676569	57984036	7	307467
	17	17_1X0	372551	417480	2	44929
	17	17_20X0	69365453	69862110	11	496657
	17	17_21X0	70040704	70073530	2	32826
	17	17_22X0	89880080	89923423	2	43343
	17	17_2X0	2690062	2705123	2	15061
	17	17_3X0	3710572	3729810	2	19238

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	17	17_4X0	3820606	4273170	10	452564
	17	17_5X0	5320987	5348813	3	27826
	17	17_6X0	14607442	14717420	4	109978
	17	17_7X0	15448986	15555919	4	106933
	17	17_8X0	31441640	31498651	2	57011
	17	17_9X0	31706245	31947751	5	241506
	18	18_10X0	64114933	64177729	2	62796
	18	18_11X0	74141481	74198369	2	56888
	18	18_12X0	81466716	81539065	2	72349
	18	18_1X0	11755446	12092858	5	337412
	18	18_2X0	15856801	15921734	4	64933
	18	18_3X0	27375919	27419281	2	43362
	18	18_4X0	29611547	29638721	2	27174
	18	18_5X0	29951094	31430973	44	1479879
	18	18_6X0	35249137	35294967	2	45830
	18	18_7X0	38177441	38242119	2	64678
	18	18_8X0	55391388	55704904	7	313516
	18	18_9X0	56194019	56458300	3	264281
	19	19_10X0	27835890	28296640	6	460750
	19	19_11X0	28593401	28666993	2	73592
	19	19_12X0	37025375	37107872	3	82497
	19	19_13X0	37221132	37281933	2	60801
	19	19_14X0	37946722	37970537	2	23815
	19	19_15X0	53012332	53056770	3	44438
	19	19_1X0	426	100424	4	99998
	19	19_2X0	118150	462559	10	344409
	19	19_3X0	10392004	10513349	3	121345
	19	19_4X0	10619220	10681145	3	61925
	19	19_5X0	11302125	11341863	4	39738
	19	19_6X0	14459450	14495208	2	35758
	19	19_7X0	24800072	24875137	2	75065
	19	19_8X0	26173771	26194198	2	20427
	19	19_9X0	26944489	26986935	2	42446
	2	2_10X0	102415503	102455609	2	40106
	2	2_11X0	105016626	105089659	2	73033
	2	2_12X0	115074344	115307494	12	233150
	2	2_13X0	116313591	116391959	3	78368
	2	2_14X0	149319445	149444548	5	125103
	2	2_15X0	149715198	150365033	8	649835
	2	2_16X0	157031192	157066781	2	35589
	2	2_17X0	164628566	164684985	2	56419
	2	2_18X0	168228099	168260536	2	32437
	2	2_19X0	181987217	182035032	3	47815
	2	2_1X0	4800361	4859854	3	59493
	2	2_20X0	186425223	187181395	8	756172
	2	2_21X0	188588796	188660179	3	71383
	2	2_22X0	188672405	188704462	2	32057
	2	2_23X0	192623244	193294599	13	671355
	2	2_24X0	193429176	193485313	2	56137
	2	2_25X0	194084081	194718294	5	634213
	2	2_26X0	195553971	195588000	2	34029
	2	2_27X0	196415738	196456481	2	40743
	2	2_28X0	198359754	198424443	9	64689
	2	2_29X0	199528902	199563386	2	34484
	2	2_2X0	23236575	23289266	2	52691

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	2	2_30X0	200572502	200607618	2	35116
	2	2_31X0	203200427	203494391	3	293964
	2	2_32X0	208750356	209001895	3	251539
	2	2_33X0	209766512	210241455	4	474943
	2	2_34X0	210685197	210839295	8	154098
	2	2_35X0	210880777	210934749	2	53972
	2	2_36X0	213173403	213421811	2	248408
	2	2_37X0	216373706	216443518	2	69812
	2	2_38X0	243502073	243740899	4	238826
	2	2_39X0	248178389	248484100	6	305711
	2	2_3X0	35178507	35550204	15	371697
	2	2_40X0	251983339	252068271	2	84932
	2	2_4X0	36011045	36096583	3	85538
	2	2_5X0	46066780	46223959	6	157179
	2	2_6X0	52591159	52871706	3	280547
	2	2_7X0	55055297	55109232	2	53935
	2	2_8X0	86470156	87040452	5	570296
	2	2_9X0	88453812	88763733	3	309921
	20	20_10X0	5374985	5418019	2	43034
	20	20_11X0	5791521	5806097	2	14576
	20	20_12X0	5831573	6251634	4	420061
	20	20_13X0	8484311	8486624	2	2313
	20	20_14X0	9851448	9895984	3	44536
	20	20_15X0	10889573	10952391	2	62818
	20	20_16X0	11530134	11656884	14	126750
	20	20_17X0	13415039	13652909	3	237870
	20	20_18X0	13760822	13825475	4	64653
	20	20_19X0	14557740	14620019	2	62279
	20	20_1X0	437638	1323419	26	885781
	20	20_1X1	34404	422464	5	388060
	20	20_20X0	27181195	27247309	2	66114
	20	20_21X0	41083317	41102056	2	18739
	20	20_22X0	44995755	45259928	6	264173
	20	20_2X0	1647441	2079467	18	432026
	20	20_3X0	2098375	2224463	6	126088
	20	20_4X0	2503401	3282523	15	779122
	20	20_5X0	3791407	4279911	19	488504
	20	20_6X0	4536212	4561152	3	24940
	20	20_7X0	4638663	4964963	12	326300
	20	20_8X0	5057701	5107899	5	50198
	20	20_9X0	5184515	5192000	2	7485
	3	3_10X0	20852767	21364268	19	511501
	3	3_11X0	21680727	21697945	2	17218
	3	3_12X0	22999616	23066658	2	67042
	3	3_13X0	37434620	37480984	2	46364
	3	3_14X0	58164931	58181971	2	17040
	3	3_15X0	61578229	61687758	10	109529
	3	3_16X0	62758758	62803373	2	44615
	3	3_17X0	72385666	72401318	2	15652
	3	3_18X0	75809445	76250840	22	441395
	3	3_18X1	72540538	74014795	50	1474257
	3	3_18X2	77695359	78216338	22	520979
	3	3_18X3	74058255	75089064	29	1030809
	3	3_18X4	78620115	79187984	23	567869
	3	3_18X5	75167456	75772521	24	605065

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	3	3_18X6	77126261	77677349	21	551088
	3	3_18X7	76707492	77110831	15	403339
	3	3_18X8	78234647	78595778	13	361131
	3	3_18X9	76255179	76696107	11	440928
	3	3_19X0	82743155	82777814	2	34659
	3	3_1X0	2686123	3274547	14	588424
	3	3_20X0	102431541	102926409	24	494868
	3	3_20X1	102947730	103598156	24	650426
	3	3_20X2	103637469	104018861	3	381392
	3	3_21X0	111553680	111616142	2	62462
	3	3_22X0	111826297	112084144	4	257847
	3	3_23X0	112365938	112396032	4	30094
	3	3_24X0	114218193	114286802	2	68609
	3	3_25X0	119173818	119247449	2	73631
	3	3_26X0	121825835	121882726	2	56891
	3	3_27X0	123106694	123119460	2	12766
	3	3_28X0	124515978	124545900	2	29922
	3	3_29X0	143084563	143223615	5	139052
	3	3_2X0	4374602	4405194	2	30592
	3	3_30X0	144316595	144574112	4	257517
	3	3_31X0	147928201	148104483	12	176282
	3	3_32X0	149358881	149724352	5	365471
	3	3_33X0	149741312	149791497	2	50185
	3	3_34X0	151641095	151724654	2	83559
	3	3_35X0	158808793	158985814	9	177021
	3	3_36X0	159082346	159115450	2	33104
	3	3_37X0	160567805	160802433	14	234628
	3	3_38X0	161018538	161212188	12	193650
	3	3_39X0	161272385	161291758	2	19373
	3	3_3X0	9792899	9807328	2	14429
	3	3_40X0	162028384	162059524	2	31140
	3	3_41X0	165679307	165741967	2	62660
	3	3_42X0	167184953	167789663	9	604710
	3	3_43X0	176698305	176716146	2	17841
	3	3_4X0	11607225	11641466	2	34241
	3	3_5X0	12448281	12502859	2	54578
	3	3_6X0	14556033	14643897	3	87864
	3	3_7X0	15688992	16319367	15	630375
	3	3_8X0	16610086	17346058	10	735972
	3	3_8X1	16413080	16754189	8	341109
	3	3_9X0	18706988	19559816	16	852828
	3	3_9X1	20303979	20695952	8	391973
	3	3_9X2	19690016	20163642	5	473626
	4	4_10X0	69138525	69457765	11	319240
	4	4_11X0	70614524	70779249	7	164725
	4	4_12X0	70918632	70973842	2	55210
	4	4_13X0	71320859	71383953	3	63094
	4	4_14X0	71512695	71837603	5	324908
	4	4_15X0	71869332	72671505	21	802173
	4	4_16X0	77495663	77747070	7	251407
	4	4_17X0	78233332	78258545	2	25213
	4	4_18X0	78283026	78385577	7	102551
	4	4_19X0	78450724	78462423	2	11699
	4	4_1X0	1355820	1843031	32	487211
	4	4_20X0	79557854	79581208	2	23354

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	4	4_21X0	82131329	82375374	16	244045
	4	4_22X0	87312766	88424305	23	1111539
	4	4_23X0	93888502	93959891	2	71389
	4	4_24X0	99185885	99243351	2	57466
	4	4_25X0	100231561	100252755	2	21194
	4	4_26X0	103049232	103761881	13	712649
	4	4_26X1	102262007	102515012	6	253005
	4	4_26X2	101531378	102147775	10	616397
	4	4_26X3	102592965	103024841	6	431876
	4	4_27X0	106289816	106362103	3	72287
	4	4_28X0	109467272	109532234	5	64962
	4	4_29X0	113900556	113920400	2	19844
	4	4_2X0	6044792	6062641	2	17849
	4	4_30X0	114766714	114786293	2	19579
	4	4_31X0	115322291	115354795	2	32504
	4	4_32X0	116912351	116936066	2	23715
	4	4_33X0	117490035	117790350	10	300315
	4	4_34X0	119143070	119188251	3	45181
	4	4_35X0	119398198	119466491	2	68293
	4	4_36X0	120407855	120414118	2	6263
	4	4_37X0	121675261	122178757	11	503496
	4	4_38X0	145413230	145439843	2	26613
	4	4_39X0	148476943	148711830	10	234887
	4	4_3X0	13564879	13645351	2	80472
	4	4_40X0	149877889	149908980	2	31091
	4	4_41X0	150484030	150520774	2	36744
	4	4_42X0	153876149	153940963	2	64814
	4	4_43X0	155009479	155051429	2	41950
	4	4_44X0	155359921	155401480	2	41559
	4	4_45X0	155792761	156276304	8	483543
	4	4_46X0	156324922	156629949	4	305027
	4	4_47X0	157107469	157154120	3	46651
	4	4_48X0	157383052	157433467	2	50415
	4	4_49X0	157822840	157877633	2	54793
	4	4_4X0	28972434	28993621	2	21187
	4	4_50X0	159563798	159630082	2	66284
	4	4_51X0	163084689	163570864	15	486175
	4	4_51X1	162009407	163049084	16	1039677
	4	4_51X2	163597635	164691405	13	1093770
	4	4_52X0	164876253	165460075	7	583822
	4	4_53X0	165732643	166224565	12	491922
	4	4_54X0	166833320	167202106	17	368786
	4	4_55X0	168832910	168884886	2	51976
	4	4_56X0	170079779	170195107	7	115328
	4	4_57X0	176158640	176528110	4	369470
	4	4_57X1	175729726	176026133	3	296407
	4	4_5X0	32373641	32392007	2	18366
	4	4_6X0	56556507	56615158	2	58651
	4	4_7X0	57855416	57959109	4	103693
	4	4_8X0	61706864	62030874	4	324010
	4	4_9X0	66371511	66421654	2	50143
	5	5_10X0	73492644	73497609	2	4965
	5	5_11X0	75049747	75116490	2	66743
	5	5_12X0	76674886	76756147	2	81261
	5	5_13X0	77245143	78042556	10	797413

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	5	5_14X0	78267248	78324278	2	57030
	5	5_15X0	106865192	107505454	11	640262
	5	5_16X0	133708228	133725656	2	17428
	5	5_17X0	133929532	134526089	6	596557
	5	5_18X0	137497363	137928107	19	430744
	5	5_19X0	138685624	138697641	2	12017
	5	5_1X0	187312	231742	2	44430
	5	5_20X0	144878212	145186357	3	308145
	5	5_21X0	145374690	145423172	4	48482
	5	5_22X0	147554918	147635789	2	80871
	5	5_23X0	147690239	147761983	2	71744
	5	5_24X0	154037202	154106433	2	69231
	5	5_25X0	155246447	155264143	3	17696
	5	5_26X0	156587093	156626520	2	39427
	5	5_27X0	157165888	157285328	5	119440
	5	5_28X0	157759448	157813756	2	54308
	5	5_29X0	159241413	159471144	5	229731
	5	5_2X0	16904939	16995304	2	90365
	5	5_30X0	159931511	159962218	2	30707
	5	5_31X0	160374031	160403373	2	29342
	5	5_32X0	162127810	162751128	12	623318
	5	5_33X0	163867961	164648328	5	780367
	5	5_34X0	164796185	164809705	2	13520
	5	5_35X0	164956280	164977916	3	21636
	5	5_36X0	167142182	167211391	2	69209
	5	5_37X0	169181418	169200109	2	18691
	5	5_3X0	56751657	56771522	2	19865
	5	5_4X0	57358327	57429245	2	70918
	5	5_5X0	58181026	58198782	2	17756
	5	5_6X0	58917208	59327762	8	410554
	5	5_7X0	64614755	64622360	2	7605
	5	5_8X0	69170358	69785744	11	615386
	5	5_9X0	69809495	69849112	2	39617
	6	6_10X0	107460668	107590373	6	129705
	6	6_11X0	109300433	109584545	3	284112
	6	6_12X0	111534548	111572949	2	38401
	6	6_13X0	127500016	128080878	15	580862
	6	6_14X0	129004240	129059050	2	54810
	6	6_15X0	129399505	129440574	2	41069
	6	6_16X0	132735883	132768379	2	32496
	6	6_17X0	135045363	135112775	2	67412
	6	6_18X0	135866739	135903647	2	36908
	6	6_19X0	138092131	139719323	42	1627192
	6	6_19X1	139747307	140715174	14	967867
	6	6_1X0	7935771	7978008	2	42237
	6	6_20X0	140805046	141831632	16	1026586
	6	6_20X1	141866364	142635763	14	769399
	6	6_20X2	142879949	143768985	13	889036
	6	6_2X0	8886730	8956276	2	69546
	6	6_3X0	30038777	30097790	2	59013
	6	6_4X0	53371706	53403214	2	31508
	6	6_5X0	77418096	77508585	2	90489
	6	6_6X0	91532469	91610799	2	78330
	6	6_7X0	95480042	95541393	2	61351
	6	6_8X0	99381362	99424019	2	42657

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	6	6_9X0	102356867	102405750	2	48883
	7	7_10X0	10962330	11213613	5	251283
	7	7_11X0	11325254	11400805	2	75551
	7	7_12X0	11908107	12457513	4	549406
	7	7_13X0	12634216	12899004	5	264788
	7	7_14X0	13378338	13997907	12	619569
	7	7_15X0	14324104	14364178	2	40074
	7	7_16X0	14486638	14828517	5	341879
	7	7_17X0	14865941	15425802	11	559861
	7	7_18X0	15785447	15852930	3	67483
	7	7_19X0	16009639	16462843	9	453204
	7	7_1X0	2906147	3170146	3	263999
	7	7_20X0	20068974	20118466	2	49492
	7	7_20X1	20262680	20480453	2	217773
	7	7_21X0	38742051	38934985	4	192934
	7	7_22X0	49729533	49741540	2	12007
	7	7_23X0	54767827	54855557	3	87730
	7	7_24X0	60087429	60100517	2	13088
	7	7_25X0	60305006	60341264	3	36258
	7	7_26X0	70640075	70661891	2	21816
	7	7_27X0	71065197	71170511	3	105314
	7	7_28X0	104474418	104511234	2	36816
	7	7_29X0	115961731	116106368	8	144637
	7	7_29X1	116275817	116634293	6	358476
	7	7_2X0	3425912	3901163	8	475251
	7	7_30X0	116655145	116677901	2	22756
	7	7_31X0	116980076	117032863	3	52787
	7	7_32X0	117055623	117440342	3	384719
	7	7_33X0	117909961	117978787	3	68826
	7	7_34X0	118473070	118983984	10	510914
	7	7_35X0	119616324	119631425	2	15101
	7	7_36X0	122922022	122963299	2	41277
	7	7_37X0	123599266	123655896	4	56630
	7	7_38X0	129860327	130128589	3	268262
	7	7_39X0	139781784	140108420	9	326636
	7	7_3X0	4182563	4293352	3	110789
	7	7_40X0	140132469	140213045	2	80576
	7	7_41X0	140448465	140469046	2	20581
	7	7_42X0	140614752	140723714	3	108962
	7	7_43X0	141237768	141261735	3	23967
	7	7_44X0	142776580	142860572	2	83992
	7	7_45X0	142994119	143633131	24	639012
	7	7_46X0	143810892	144044629	2	233737
	7	7_47X0	144052202	144109116	2	56914
	7	7_48X0	144531814	144650129	9	118315
	7	7_4X0	4620033	4714534	2	94501
	7	7_5X0	5105770	5451262	7	345492
	7	7_5X1	5593735	6216681	6	622946
	7	7_6X0	6625341	6971557	7	346216
	7	7_6X1	7068703	7311272	3	242569
	7	7_6X2	7461109	7753978	3	292869
	7	7_7X0	8809714	9630611	15	820897
	7	7_8X0	9897983	9976367	2	78384
	7	7_9X0	10418602	10896447	5	477845
	8	8_10X0	31977001	32037531	2	60530

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	8	8_11X0	36815007	37349694	7	534687
	8	8_12X0	39878955	39907478	2	28523
	8	8_13X0	40001030	40014734	2	13704
	8	8_14X0	40191902	40765853	16	573951
	8	8_15X0	41235156	41395825	10	160669
	8	8_16X0	43469442	44043983	26	574541
	8	8_16X1	42260375	42614536	17	354161
	8	8_16X2	43043094	43443531	14	400437
	8	8_16X3	42650438	43024557	13	374119
	8	8_16X4	41898624	42047334	10	148710
	8	8_17X0	49274960	49301125	3	26165
	8	8_18X0	49342067	49456279	5	114212
	8	8_19X0	49676616	49716955	2	40339
	8	8_1X0	2604962	2686160	3	81198
	8	8_20X0	50525091	50561736	3	36645
	8	8_21X0	52729003	52763400	2	34397
	8	8_22X0	53211438	53277859	2	66421
	8	8_23X0	55176847	55182545	2	5698
	8	8_24X0	59561817	59629133	3	67316
	8	8_25X0	62368998	62424303	2	55305
	8	8_26X0	62451329	62478147	2	26818
	8	8_27X0	62946732	62987182	2	40450
	8	8_28X0	70603249	70684115	2	80866
	8	8_29X0	85711795	85720790	2	8995
	8	8_2X0	4320230	4419384	3	99154
	8	8_30X0	98734295	98761840	2	27545
	8	8_31X0	99568958	99918825	4	349867
	8	8_32X0	107106982	107155930	2	48948
	8	8_33X0	113672758	113689681	2	16923
	8	8_34X0	114867062	115188588	5	321526
	8	8_35X0	116325286	116342146	3	16860
	8	8_36X0	116686601	116730061	2	43460
	8	8_37X0	119064351	119141150	5	76799
	8	8_38X0	122398276	122421384	2	23108
	8	8_39X0	127858610	127912860	3	54250
	8	8_3X0	5522739	5901049	10	378310
	8	8_40X0	128171379	128521109	3	349730
	8	8_41X0	128740756	128782456	2	41700
	8	8_42X0	130401470	130421871	2	20401
	8	8_43X0	131888047	132009981	4	121934
	8	8_44X0	132828091	133214257	7	386166
	8	8_4X0	6013207	6061837	2	48630
	8	8_5X0	17775423	18115456	8	340033
	8	8_6X0	18719444	19082478	11	363034
	8	8_6X1	19153903	19549426	9	395523
	8	8_6X2	20054422	20536473	9	482051
	8	8_6X3	19585840	19995597	7	409757
	8	8_6X4	20728817	21341431	10	612614
	8	8_7X0	21380504	21617736	5	237232
	8	8_8X0	22035256	22057209	3	21953
	8	8_9X0	23193181	23292682	3	99501
	9	9_10X0	24446264	24509761	3	63497
	9	9_11X0	26841299	26897219	2	55920
	9	9_12X0	27333956	27351295	2	17339
	9	9_13X0	27408333	27511176	3	102843

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	9	9_14X0	27546528	27556861	2	10333
	9	9_15X0	37727942	37760569	2	32627
	9	9_16X0	41006932	41069929	2	62997
	9	9_17X0	43049638	43108486	2	58848
	9	9_18X0	43112141	43127887	2	15746
	9	9_19X0	44499800	44617210	4	117410
	9	9_1X0	4094995	4327679	5	232684
	9	9_20X0	46840992	47258503	6	417511
	9	9_21X0	64929721	65265149	4	335428
	9	9_22X0	67546408	67785021	3	238613
	9	9_23X0	80118029	80167033	2	49004
	9	9_24X0	82033543	82066047	2	32504
	9	9_25X0	88747957	88749664	2	1707
	9	9_26X0	92962738	93404883	3	442145
	9	9_27X0	94178221	94231966	4	53745
	9	9_28X0	99658487	100138801	11	480314
	9	9_29X0	101222846	101269843	2	46997
	9	9_2X0	4654294	4978892	4	324598
	9	9_30X0	114004095	114327767	4	323672
	9	9_31X0	117439148	117467198	2	28050
	9	9_32X0	119307176	119332967	2	25791
	9	9_3X0	9988057	10012611	2	24554
	9	9_4X0	10158937	10203460	2	44523
	9	9_5X0	10450426	10490416	2	39990
	9	9_6X0	14611561	14864680	8	253119
	9	9_7X0	15610438	15628881	2	18443
	9	9_8X0	19451630	19476646	2	25016
	9	9_9X0	20917673	20982487	2	64814
	KL567939.1	KL567939.1_1X0	72607	131036	7	58429
	KL567939.1	KL567939.1_1X1	33092	61645	3	28553
	X	X_10X0	42951237	43044975	3	93738
	X	X_11X0	45178446	45284341	4	105895
	X	X_12X0	60849014	61232085	6	383071
	X	X_13X0	63897664	63999622	3	101958
	X	X_14X0	71677455	71684773	2	7318
	X	X_15X0	77674150	78399944	6	725794
	X	X_16X0	95344825	95353344	2	8519
	X	X_17X0	105537602	106393459	11	855857
	X	X_18X0	106487870	106523820	2	35950
	X	X_19X0	106556838	106825014	5	268176
	X	X_1X0	1543244	1601395	2	58151
	X	X_20X0	107323215	107372338	2	49123
	X	X_21X0	123350346	123407645	2	57299
	X	X_22X0	123851845	124576133	14	724288
	X	X_23X0	131340045	131382011	2	41966
	X	X_24X0	136715598	137015266	4	299668
	X	X_25X0	157389496	157474219	2	84723
	X	X_2X0	9998415	10031167	2	32752
	X	X_3X0	13989401	13992429	2	3028
	X	X_4X0	14733145	14783792	2	50647
	X	X_5X0	15679254	15707436	2	28182
	X	X_6X0	17962909	18341467	4	378558
	X	X_7X0	23414354	23469041	2	54687
	X	X_8X0	24059516	24629998	4	570482

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	X	X_9X0	28486869	28562803	2	75934
	1	1_10X0	262510765	263734877	26	1224112
	1	1_11X0	263827304	263838299	2	10995
	1	1_12X0	269967008	269985478	2	18470
	1	1_1X0	2003172	2019821	2	16649
	1	1_2X0	31048535	31162453	11	113918
	1	1_3X0	52986462	53007877	3	21415
	1	1_4X0	81891717	81909384	2	17667
	1	1_5X0	201223599	201852294	40	628695
	1	1_6X0	236586271	236634721	6	48450
	1	1_7X0	236709676	236742192	3	32516
	1	1_8X0	245962270	245986294	3	24024
	1	1_9X0	249876927	249884826	2	7899
	10	10_1X0	32093459	32112737	2	19278
	10	10_2X0	40565664	40579521	2	13857
	12	12_10X0	21305930	21440121	12	134191
	12	12_11X0	21456592	21477119	4	20527
	12	12_12X0	24784953	24880134	8	95181
	12	12_13X0	34407732	34470986	6	63254
	12	12_14X0	40779159	40800026	2	20867
	12	12_15X0	49366557	49418344	5	51787
	12	12_16X0	52436012	52442920	2	6908
	12	12_17X0	52729025	52735415	2	6390
	12	12_18X0	52843013	52857023	2	14010
	12	12_1X0	973786	982259	2	8473
	12	12_2X0	2562335	2580284	2	17949
	12	12_3X0	5037172	5055485	2	18313
	12	12_4X0	5523233	5539715	2	16482
	12	12_5X0	6310230	6333338	2	23108
	12	12_6X0	19207143	19218352	2	11209
	12	12_7X0	19996343	20027248	3	30905
	12	12_8X0	20842393	20861461	2	19068
	12	12_9X0	21209793	21255560	4	45767
	13	13_1X0	27940430	27955035	2	14605
	13	13_2X0	34007074	34059910	5	52836
	13	13_3X0	34090023	34112723	2	22700
	13	13_4X0	131325123	131345685	2	20562
	13	13_5X0	138193262	138214816	2	21554
	13	13_6X0	161015933	162092362	30	1076429
	13	13_7X0	194424740	194439042	2	14302
	14	14_10X0	132685064	132699485	2	14421
	14	14_1X0	30105257	30119292	2	14035
	14	14_2X0	48865742	48884625	2	18883
	14	14_3X0	49795092	49822296	4	27204
	14	14_4X0	49951437	49972768	2	21331
	14	14_5X0	50937885	50944093	2	6208
	14	14_6X0	85460113	85481911	2	21798
	14	14_7X0	104551333	104567722	2	16389
	14	14_8X0	108739541	108754749	2	15208
	14	14_9X0	114381027	114413170	3	32143
	15	15_1X0	44804228	44807100	2	2872
	15	15_2X0	78186174	78204157	2	17983
	15	15_3X0	81881362	81990807	10	109445
	15	15_4X0	84406876	84415607	2	8731
	15	15_5X0	133040370	133058400	2	18030

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	15	15_6X0	139040741	139050341	2	9600
	15	15_7X0	139135513	139148523	2	13010
	16	16_1X0	47478023	47499700	2	21677
	17	17_1X0	30220162	30239897	2	19735
	17	17_2X0	30403841	30474471	6	70630
	17	17_3X0	61692968	61698900	2	5932
	17	17_4X0	62537448	62555659	2	18211
	17	17_5X0	62906142	62927297	2	21155
	18	18_1X0	6338614	6348958	2	10344
	18	18_2X0	6516474	6537647	2	21173
	18	18_3X0	7172636	7184900	2	12264
	18	18_4X0	50002921	50022799	2	19878
	2	2_1X0	106967	159276	5	52309
	2	2_2X0	13750130	14531021	36	780891
	2	2_2X1	67104951	67480534	18	375583
	2	2_2X10	52289418	52587060	3	297642
	2	2_2X2	67970426	68292371	12	321945
	2	2_2X3	52699316	53251277	13	551961
	2	2_2X4	12843402	12878983	4	35581
	2	2_2X5	11964449	11991314	3	26865
	2	2_2X6	66291172	66312586	2	21414
	2	2_2X7	64190713	64213028	2	22315
	2	2_2X8	51726854	51743253	2	16399
	2	2_2X9	63348429	63364518	2	16089
	2	2_3X0	69097009	69120461	2	23452
	2	2_4X0	79643342	79650829	2	7487
	2	2_5X0	80364524	80381719	2	17195
	2	2_6X0	142806018	143156550	24	350532
	3	3_10X0	41479477	41491918	4	12441
	3	3_11X0	56762418	56781143	2	18725
	3	3_12X0	59176153	59192898	2	16745
	3	3_13X0	64838491	64850659	2	12168
	3	3_1X0	691931	715980	3	24049
	3	3_2X0	9866016	9886647	2	20631
	3	3_3X0	17346240	17361585	2	15345
	3	3_4X0	17535514	17545171	2	9657
	3	3_5X0	39171390	39173954	2	2564
	3	3_6X0	39302486	39314076	2	11590
	3	3_7X0	39334181	39347202	2	13021
	3	3_8X0	40616643	40635947	2	19304
	3	3_9X0	41016457	41029080	2	12623
	4	4_1X0	88834253	88855793	2	21540
	4	4_2X0	91459849	91981079	30	521230
	4	4_3X0	99168255	99198053	4	29798
	4	4_4X0	100143464	100164252	2	20788
	5	5_10X0	61242636	61263269	3	20633
	5	5_11X0	61441125	61809691	10	368566
	5	5_11X1	61833027	62259578	12	426551
	5	5_12X0	78977079	79209505	15	232426
	5	5_13X0	80512208	80531021	2	18813
	5	5_14X0	100753501	100764767	2	11266
	5	5_1X0	423294	439617	2	16323
	5	5_2X0	14997381	15014019	2	16638
	5	5_3X0	15863303	15892148	3	28845
	5	5_4X0	17482097	17868244	15	386147

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	5	5_4X1	17896310	18222833	14	326523
	5	5_5X0	19167174	19270411	8	103237
	5	5_6X0	20541804	20817846	17	276042
	5	5_6X1	20829059	21099179	21	270120
	5	5_6X2	20112632	20535575	20	422943
	5	5_7X0	21555072	21562417	2	7345
	5	5_8X0	21752204	21760815	2	8611
	5	5_9X0	22718510	22734083	2	15573
	6	6_10X0	47643560	47657784	2	14224
	6	6_11X0	47939897	47953282	3	13385
	6	6_12X0	48698849	48713086	2	14237
	6	6_13X0	54289166	54308064	2	18898
	6	6_14X0	55570635	55725259	13	154624
	6	6_15X0	58482749	58501139	3	18390
	6	6_16X0	58520967	58541867	2	20900
	6	6_17X0	59351448	59370483	2	19035
	6	6_18X0	59671471	59684286	2	12815
	6	6_19X0	59701608	59767006	7	65398
	6	6_1X0	2542675	2554367	2	11692
	6	6_20X0	60511110	60521141	2	10031
	6	6_21X0	60963992	61046846	8	82854
	6	6_22X0	71740302	71774262	3	33960
	6	6_23X0	71920605	71932254	2	11649
	6	6_24X0	80587177	80608011	3	20834
	6	6_25X0	91002163	91048162	4	45999
	6	6_26X0	119076677	119096127	2	19450
	6	6_27X0	164112645	164133008	2	20363
	6	6_28X0	168101472	168146473	5	45001
	6	6_29X0	169146334	169154551	2	8217
	6	6_2X0	12405616	12426270	3	20654
	6	6_3X0	18610593	18676112	6	65519
	6	6_4X0	27618602	27637508	2	18906
	6	6_5X0	28583320	28599368	2	16048
	6	6_6X0	44650787	44696766	4	45979
	6	6_7X0	45005358	45018656	2	13298
	6	6_8X0	45307536	45312539	2	5003
	6	6_9X0	46172561	46842870	13	670309
	7	7_10X0	32817917	32825846	2	7929
	7	7_11X0	37288636	37310170	2	21534
	7	7_12X0	43971250	43993156	2	21906
	7	7_13X0	78624873	79119982	28	495109
	7	7_13X1	76832222	77548071	32	715849
	7	7_13X2	74974308	75164810	4	190502
	7	7_13X3	76321059	76335796	2	14737
	7	7_14X0	79214271	79531982	15	317711
	7	7_1X0	21037717	21118895	7	81178
	7	7_2X0	21220799	21226971	2	6172
	7	7_3X0	22515697	22541285	3	25588
	7	7_4X0	23696398	23706738	3	10340
	7	7_5X0	23791579	23833540	4	41961
	7	7_6X0	23909960	23929754	2	19794
	7	7_7X0	24020323	24040523	2	20200
	7	7_8X0	25058463	25077233	2	18770
	7	7_9X0	31605421	31621132	2	15711
	8	8_1X0	4107834	4123084	2	15250

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	9	9_10X0	61554442	61559946	2	5504
	9	9_11X0	73705304	73726240	2	20936
	9	9_12X0	76624845	76644690	2	19845
	9	9_1X0	1324366	1344990	2	20624
	9	9_2X0	2630813	2833065	14	202252
	9	9_3X0	3408745	3935067	39	526322
	9	9_3X1	4363666	4737880	27	374214
	9	9_3X2	3947756	4353256	30	405500
	9	9_3X3	4750073	5039118	19	289045
	9	9_3X4	5065146	5470853	26	405707
	9	9_3X5	6433177	6685320	4	252143
	9	9_3X6	9378417	9409139	3	30722
	9	9_3X7	5910085	5929626	2	19541
	9	9_4X0	23009082	23031338	2	22256
	9	9_5X0	27072760	27104985	4	32225
	9	9_6X0	39638645	39644103	2	5458
	9	9_7X0	45165815	45188656	2	22841
	9	9_8X0	50613023	51232044	34	619021
	9	9_9X0	51977458	51992249	2	14791
	AEMK02000159.1	AEMK02000159.1_1X0	91359	108542	2	17183
	AEMK02000159.1	AEMK02000159.1_1X1	125549	145791	2	20242
	AEMK02000159.1	AEMK02000159.1_2X0	201389	216885	3	15496
	AEMK02000159.1	AEMK02000159.1_2X1	167227	185058	2	17831
	AEMK02000159.1	AEMK02000159.1_3X0	238458	293941	5	55483
	AEMK02000253.1	AEMK02000253.1_1X0	49131	82484	4	33353
	AEMK02000253.1	AEMK02000253.1_1X1	417	35186	3	34769
	AEMK02000253.1	AEMK02000253.1_2X0	186984	221293	3	34309
	AEMK02000261.1	AEMK02000261.1_1X0	16325	35870	2	19545
	AEMK02000261.1	AEMK02000261.1_2X0	73478	104919	4	31441
	AEMK02000261.1	AEMK02000261.1_3X0	170553	213363	5	42810
	AEMK02000261.1	AEMK02000261.1_3X1	126663	153873	3	27210
	AEMK02000261.1	AEMK02000261.1_4X0	259990	286540	3	26550
	AEMK02000261.1	AEMK02000261.1_5X0	309011	338467	3	29456
	AEMK02000261.1	AEMK02000261.1_6X0	359126	396614	4	37488
	AEMK02000261.1	AEMK02000261.1_7X0	580666	604943	3	24277
	AEMK02000261.1	AEMK02000261.1_7X1	563007	573167	2	10160
	AEMK02000261.1	AEMK02000261.1_7X2	537778	551622	2	13844
	AEMK02000261.1	AEMK02000261.1_7X3	498703	525503	3	26800
	AEMK02000261.1	AEMK02000261.1_7X4	463431	486557	3	23126
	AEMK02000361.1	AEMK02000361.1_1X0	92389	138099	4	45710
	AEMK02000368.1	AEMK02000368.1_1X0	16791	30874	2	14083
	AEMK02000368.1	AEMK02000368.1_2X0	63235	119551	6	56316
	AEMK02000368.1	AEMK02000368.1_3X0	165964	185311	2	19347
	AEMK02000393.1	AEMK02000393.1_1X0	346129	367575	2	21446
	AEMK02000393.1	AEMK02000393.1_2X0	829949	833269	2	3320
	AEMK02000410.1	AEMK02000410.1_1X0	81219	113429	3	32210
	AEMK02000410.1	AEMK02000410.1_2X0	197905	222292	3	24387
	AEMK02000410.1	AEMK02000410.1_2X1	234113	270209	4	36096
	AEMK02000410.1	AEMK02000410.1_2X2	145188	187570	4	42382
	AEMK02000410.1	AEMK02000410.1_2X3	285352	317515	3	32163
	AEMK02000435.1	AEMK02000435.1_1X0	64160	80293	3	16133
	AEMK02000435.1	AEMK02000435.1_1X1	43272	56236	2	12964
	AEMK02000435.1	AEMK02000435.1_1X2	11066	33052	3	21986
	AEMK02000449.1	AEMK02000449.1_1X0	246242	276984	3	30742
	AEMK02000449.1	AEMK02000449.1_2X0	330907	408221	6	77314

Table B.50: CTDG repertoire across selected mammalian genomes

species	chromosome	cluster	start	end	duplicates	length
	AEMK02000449.1	AEMK02000449.1_3X0	622437	685745	6	63308
	AEMK02000449.1	AEMK02000449.1_4X0	776417	787645	2	11228
	AEMK02000449.1	AEMK02000449.1_5X0	865676	951850	7	86174
	AEMK02000449.1	AEMK02000449.1_6X0	1028234	1038248	2	10014
	AEMK02000449.1	AEMK02000449.1_7X0	1270942	1280946	2	10004
	AEMK02000453.1	AEMK02000453.1_1X0	34683	62237	4	27554
	AEMK02000453.1	AEMK02000453.1_1X1	76338	96056	3	19718
	AEMK02000453.1	AEMK02000453.1_1X2	107643	121865	2	14222
	AEMK02000522.1	AEMK02000522.1_1X0	124343	181345	5	57002
	AEMK02000522.1	AEMK02000522.1_2X0	235935	252054	2	16119
	AEMK02000522.1	AEMK02000522.1_3X0	295157	308461	2	13304
	AEMK02000522.1	AEMK02000522.1_4X0	350248	377625	3	27377
	AEMK02000522.1	AEMK02000522.1_5X0	410253	436487	3	26234
	AEMK02000537.1	AEMK02000537.1_1X0	144365	162763	3	18398
	AEMK02000537.1	AEMK02000537.1_2X0	262873	314061	5	51188
	AEMK02000566.1	AEMK02000566.1_1X0	21132	57301	4	36169
	AEMK02000566.1	AEMK02000566.1_2X0	74004	116159	6	42155
	AEMK02000566.1	AEMK02000566.1_3X0	121684	143366	4	21682
	AEMK02000566.1	AEMK02000566.1_3X1	154153	173411	3	19258
	AEMK02000577.1	AEMK02000577.1_1X0	15662	48727	4	33065
	AEMK02000577.1	AEMK02000577.1_1X1	69193	81415	2	12222
	AEMK02000577.1	AEMK02000577.1_1X2	89219	120396	3	31177
	AEMK02000577.1	AEMK02000577.1_2X0	164562	178767	2	14205
	AEMK02000577.1	AEMK02000577.1_3X0	202663	215444	2	12781
	AEMK02000598.1	AEMK02000598.1_1X0	1409321	1438635	3	29314
	AEMK02000598.1	AEMK02000598.1_2X0	1951141	1970891	2	19750
	AEMK02000630.1	AEMK02000630.1_1X0	79981	95535	2	15554
	AEMK02000630.1	AEMK02000630.1_2X0	193917	204170	2	10253
	AEMK02000658.1	AEMK02000658.1_1X0	38093	52426	3	14333
	AEMK02000658.1	AEMK02000658.1_1X1	9275	23843	2	14568
	AEMK02000658.1	AEMK02000658.1_2X0	89395	118604	3	29209
	AEMK02000658.1	AEMK02000658.1_2X1	72196	84629	2	12433
	AEMK02000658.1	AEMK02000658.1_3X0	145796	157462	2	11666
	AEMK02000661.1	AEMK02000661.1_1X0	148594	184890	3	36296
	AEMK02000661.1	AEMK02000661.1_2X0	274376	326050	4	51674
	AEMK02000661.1	AEMK02000661.1_3X0	374416	384751	2	10335
	AEMK02000661.1	AEMK02000661.1_4X0	522588	545022	3	22434
	AEMK02000661.1	AEMK02000661.1_4X1	572032	590348	2	18316
	AEMK02000661.1	AEMK02000661.1_5X0	619445	637697	2	18252
	AEMK02000677.1	AEMK02000677.1_1X0	44774	55466	2	10692
	AEMK02000682.1	AEMK02000682.1_1X0	885009	916136	4	31127
	AEMK02000682.1	AEMK02000682.1_2X0	1027607	1053198	5	25591
	X	X_1X0	33299217	33308230	2	9013
	X	X_2X0	49892308	49898438	2	6130
	X	X_3X0	58696808	58719987	3	23179
	X	X_4X0	58830477	58852879	2	22402
	X	X_5X0	83580964	83599213	2	18249
	X	X_6X0	107894810	107909548	2	14738
	X	X_7X0	125124224	125140510	2	16286
	Y	Y_1X0	19637456	23945894	16	4308438