Interchromosomal Translocations as a Means to Map Chromosome Territories in Breast Cancer

Cancer Informatics
Volume 18: 1–10

DOI: 10.1177/1176935119842573

Fritz F Parl[1], William D Dupont[2] and Philip S Crooke[3]

[1]Department of Pathology, Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, TN, USA. [2]Department of Biostatistics, Vanderbilt University, Nashville, TN, USA. [3]Department of Mathematics, Vanderbilt University, Nashville, TN, USA.

ABSTRACT: The genome-wide identification of mutated genes is an important advance in our understanding of tumor biology, but several fundamental questions remain open. How do these genes act together to promote cancer development and, a related question, how are they spatially arranged in the nucleus to allow coordinated expression? We examined the nuclear topography of mutated genes in breast cancer and their relation to chromosome territories (CTs). We performed a literature review and analyzed 1 type of mutation, interchromosomal translocations, in 1546 primary breast cancers to infer the spatial arrangement of chromosomes. The cosegregation of all observed fusion genes was used to create a matrix of genome-wide CT contacts and develop a tentative CT map of breast cancer. Regression analysis was performed to determine the association between CTs and all types of mutations. Chromosomes 17, 11, 8, and 1 had the majority of interchromosomal fusions and are presumably clustered in the nuclear center, whereas chromosomes 22, 21, X, and 18 had the lowest number of contacts, likely reflecting a more peripheral position. Regression analysis revealed that there was no significant association between chromosome length indicated by the number of base pairs per chromosome and the number of total (inter- and intrachromosomal) translocations, point mutations, or copy number aberrations (CNAs). The gene density of chromosomes (genes/Mb) was significantly correlated with total translocations ($P = .02$), but not with point mutations $P = .19$ and CNAs $P = .62$. Finally, the association of the 3 genetic alterations with the CT map deduced from the interchromosomal fusions was significant, ie, total translocations $P = 7 \times 10^{-11}$, point mutations $P = .01$, CNAs $P = .002$. In conclusion, we developed a tentative CT map and observed a spatial association with genetic alterations in breast cancer.

KEYWORDS: genetics, mutation

RECEIVED: December 7, 2018. **ACCEPTED:** March 17, 2019.

TYPE: Original Research

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: W.D. Dupont was supported by NCI grants R01 CA050468 and P30 CA068485, and NCATS grant UL1 TR000445.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Fritz F Parl, Department of Pathology, Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, TN 37232, USA.
Email: fritz.parl@vanderbilt.edu

Introduction

Whole-genome sequencing has permitted the comprehensive identification of somatic mutations in malignancies and enabled the definition of mutational signatures characteristic of individual tumor types. For example, a whole-genome analysis of the somatic genetics of 560 breast cancers identified 3 479 652 mutations.[1] A bioinformatics approach was used to extract 20 different mutational signatures and define driver and passenger mutations. The detailed analysis identified 93 protein-coding genes carrying 1628 likely driver mutations. Although the definitive identification of mutated genes is an important advance in our understanding of tumor biology, several fundamental questions remain open. How do these genes act together to promote cancer development and, a related question, how are they spatially arranged in the nucleus to allow coordinated expression? Little is known about the nuclear topography of mutated genes in cancer.

Mapping studies have shown that chromosomes are nonrandomly arranged in the nucleus.[2] Each of the interphase chromosomes is confined to a discrete region of the nucleus, referred to as a chromosome territory (CT). CTs intermingle with neighboring CTs during interphase[3] but are organized into patterns, ie, some chromosomes localize toward the periphery, often touching the nuclear membrane, whereas others are located toward the center of the nucleus. The chromosome arms are mostly kept apart from each other. Even the 2 copies of the same chromosome within the same nucleus often occupy distinct positions and have different immediate neighbors.[4] In this study, we pursued 2 linked goals: first, to define the nuclear topography of breast cancer by developing a tentative CT map, and second, to determine whether we can use the CT map to analyze the spatial distribution of mutated genes. Three-dimensional (3D) information on the juxtaposition of mutated genes would allow a better understanding of functional interactions between cancer genes in the malignant process.

To date, the spatial arrangement of CTs has been studied by elaborate 3-dimensional fluorescence in situ hybridization (3D FISH) protocols and chromosome conformation capture (3 C) techniques, such as Hi-C.[5,6] Because these methods are technically demanding, they have been largely limited to small sample sizes, eg, cell lines. To circumvent this limitation and develop a CT map of a large number of tumors, we take advantage of one characteristic of malignancies, namely, interchromosomal translocations, to map the nuclear topography of CTs in breast cancer. The physical proximity of CTs contributes to the probability of interchromosomal translocations, which occur when unrepaired double-strand breaks from separate chromosomes undergo illegitimate joining.[3] There is a significant correlation between their proximity in health and translocation frequency in malignancy.[7] CTs that are closer to each

Table 1. Primary breast cancers with translocations.

NO. OF TUMORS	NO. OF FUSIONS	TIER 1 + 2	TOTAL TRANSLOCATIONS		REFERENCE
			INTRACHROMOSOMAL, N (%)	INTERCHROMOSOMAL, N (%)	
22	165	71	47 (66.2)	24 (33.8)	Banerji et al[11]
120	70	32	26 (81.2)	6 (18.8)	Kim et al[8]
212	118	105	73 (69.5)	32 (30.5)	Ma et al[9]
173	12	12	7 (58.3)	5 (41.7)	Matissek et al[12]
1019	3767	3567	2572 (72.1)	995 (27.9)	Yoshihara et al[10]
1546	4132	3787	2725 (72.0)	1062 (28.0)	

other in the nuclear space are likely to form fusion junctions more often than distant CTs. A fusion gene is a hybrid formed from 2 distinct genes that undergo chromosomal rearrangement.[8-10] When the genes are located on the same chromosome, their fusion leads to an intrachromosomal rearrangement, whereas interchromosomal fusion transcripts occur between 2 genes located on different chromosomes. In this study, we analyzed interchromosomal translocations in primary breast cancer to infer the spatial clustering of chromosomes. The cosegregation of all observed fusion genes in 1546 tumors was used to create a matrix of genome-wide CT contacts and develop a tentative CT map of breast cancer. Chromosomes 17, 11, 8, and 1 had the majority of interchromosomal fusions suggesting that they are clustered near each other in the nuclear center, whereas chromosomes 22, 21, X, and 18 had the lowest number of contacts, likely reflecting a wider distance between each other and a more peripheral position. The detailed analysis of mutated genes revealed that they were more frequently located on chromosomes near each other than on chromosomes separated by longer distances. Although the cause of this spatial association is uncertain, these findings offer a better understanding of short- and long-range nuclear interactions.

Methods

We performed a literature review of 1 type of genomic rearrangement in primary breast cancer, namely, translocations. The translocation studies were selected on the basis of 3 criteria. (1) Only genome-wide translocation studies were included to ensure an unbiased approach. (2) Only studies distinguishing intra- and interchromosomal translocations were selected because only data from the latter type of gene fusion provide information to construct a CT map. (3) All studies performed transcript sequencing to detect gene fusion events, but only those using a number of filtering criteria to flag false positive fusions were included. Generally, the fusion transcript lists were classified into tiers based on level of evidence, with tier 1 the most accurate. In this study, we accepted tiers 1 and 2 as valid fusion events. The described bioinformatics selection identified 5 studies.[8-12] The interchromosomal fusion genes and their chromosomal locations were tabulated and the number of fusion genes entered into a matrix of all possible chromosome pairs. The R software package gplots (https://cran.r-project.org/web/packages/gplots/gplots.pdf) was used to generate a heatmap from the interchromosomal fusion gene matrix.

To ascertain the genomic distribution of mutated genes, we used the METABRIC trial[13,14] (www.cbioportal.org) to determine the chromosomal distribution of point mutations and CNAs in a large genome-wide database. We regressed total (inter- and intrachromosomal) translocations, point mutations, and CNAs against the chromosome length (number of base pairs per chromosome), against gene densities (genes/Mb per chromosome), and against the number of fusion events observed in each chromosome using simple linear regressions. In these analyses, the unit of observation is the chromosome. Two-sided *P* values from these analyses test the null hypothesis that the expected slope of the relationship between these variables is zero. Sensitivity analyses in which individual chromosomes with influential residuals were eliminated from our analyses did not substantially affect our conclusions. These analyses were conducted in Stata.[15]

Results

We examined 1546 primary breast cancers from 5 studies with a total of 3787 translocations, of which 1062 (28.0%) were interchromosomal (Table 1). The detailed analysis of the interchromosomal fusions is summarized as a matrix in Table 2. Chromosomes 17, 11, 8, and 1 are most frequently involved in interchromosomal translocations, whereas chromosomes 22, 21, X, and 18 are at the opposite, low end of the spectrum. The heatmap in Figure 1 displays the data matrix visually to represent the distribution of chromosomes: The red color identifies chromosomes that have frequent contacts with multiple other chromosomes (eg, 17, 11, 8, and 1), whereas chromosomes with infrequent contacts with few other chromosomes (eg, 22, 21, X, and 18) are depicted in yellow. Because interchromosomal fusions can only occur between adjacent chromosomes, we can tentatively infer their relative locations in the nucleus. We hypothesize that chromosomes that are centrally located are

Table 2. Matrix of interchromosomal fusion events.

	CHR17	CHR11	CHR8	CHR1	CHR6	CHR20	CHR19	CHR12	CHR10	CHR3	CHR2	CHR7	CHR16	CHR9	CHR14	CHR4	CHR5	CHR15	CHR13	CHR22	CHR21	CHRX	CHR18
chr17		38	13	17	12	19	23	6	4	13	3	8	6	7	2	2	3	5	3	3	7	1	2
chr11	38		20	5	10	8	8	13	11	5	6	9	5	5	8	3	2	5	2	4	2	3	3
chr8	13	20		15	14	3	12	5	9	4	15	5	7	7	10	4	7	1	5	0	0	4	1
chr1	17	5	15		19	8	11	9	3	14	7	5	6	5	2	6	7	3	7	3	4	1	2
chr6	12	10	14	19		3	6	6	6	5	9	2	1	3	7	3	4	0	4	5	1	1	2
chr20	19	8	3	8	3		9	8	3	5	2	4	7	2	2	9	1	4	1	9	4	5	3
chr19	23	8	12	11	6	9		2	7	1	3	3	6	5	3	2	1	3	3	1	4	1	1
chr12	6	13	5	9	6	8	2		8	6	6	4	3	4	5	2	2	5	5	2	4	2	2
chr10	4	11	9	3	6	3	7	8		2	5	4	1	7	5	1	7	8	4	5	3	0	0
chr3	13	5	4	14	5	5	1	6	2		2	5	1	1	1	3	5	2	4	4	2	4	1
chr2	3	6	15	7	9	2	3	6	5	2		4	1	3	1	1	2	5	1	1	0	3	1
chr7	8	9	5	5	2	4	3	4	4	5	4		4	5	1	7	1	2	1	1	5	0	0
chr16	6	5	7	6	1	7	6	3	1	1	1	4		2	3	5	3	5	1	1	5	0	0
chr9	7	5	7	5	3	2	5	4	7	1	3	5	2		6	1	2	1	2	1	0	2	1
chr14	2	8	10	2	7	2	3	5	5	1	1	1	3	6		1	2	0	7	1	0	1	0
chr4	2	3	4	6	3	9	2	2	1	3	1	7	5	1	1		5	5	0	0	4	2	1
chr5	3	2	7	7	4	1	1	2	7	5	2	1	3	2	2	5		4	1	1	0	1	1
chr15	5	5	1	3	0	4	3	5	8	2	5	2	5	1	0	5	4		1	0	0	0	2
chr13	3	2	5	7	4	1	3	5	4	4	1	1	1	2	7	0	1	1		2	0	0	3
chr22	3	4	0	3	5	9	1	2	5	4	1	1	1	1	1	0	1	0	2		1	0	2
chr21	7	2	0	4	1	4	4	4	3	2	0	5	5	0	0	4	0	0	0	1		0	0
chrX	1	3	4	1	1	5	1	2	0	4	3	0	0	2	1	2	1	0	0	0	0		0
chr18	2	3	1	2	2	3	1	2	0	1	1	0	0	1	0	1	1	2	3	2	0	0	
Total	197	175	161	159	123	119	115	109	103	90	81	80	73	72	68	67	62	61	57	47	46	31	28

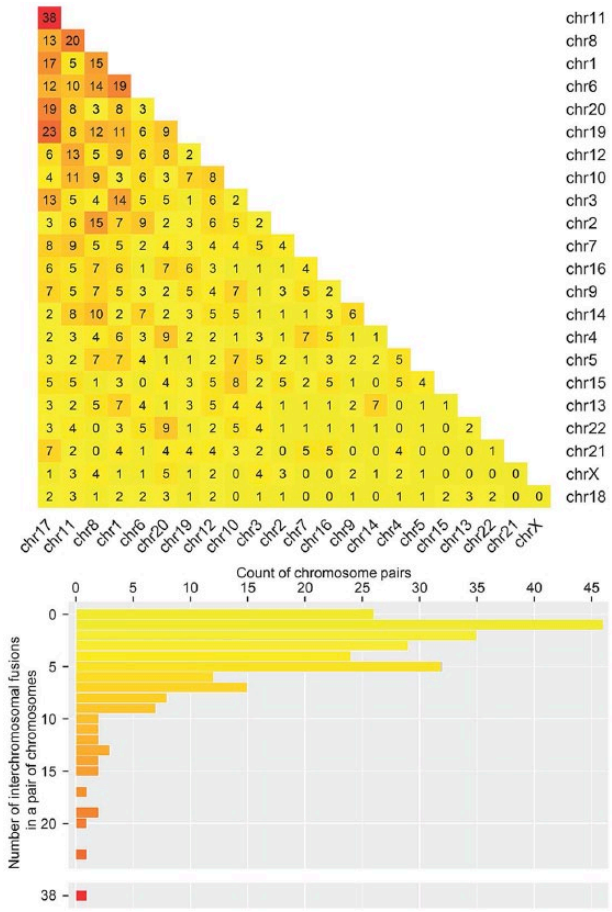

Figure 1. Heatmap and histogram of interchromosomal fusion events. Top: The heatmap shows the observed number of interchromosomal contacts between all pairs of chromosomes. Bottom: The histogram shows how frequently each number of contacts occurred as indicated by a color spectrum ranging from none (yellow) to a maximum of 38 (red).

more likely to be adjacent to multiple chromosomes than are chromosomes near the periphery.

The tentative CT map was derived as an average of 1546 primary breast cancers. To estimate the degree of variability between cancers, we performed a separate analysis of tumors with multiple interchromosomal translocations. The largest numbers of translocations were 9, found in 3 tumors, and even 13 interchromosomal translocations, found in a single tumor (Table 3). The matrix analysis of this subgroup of tumors (Table 4) showed that chromosomes 17, 11, 8, and 1 participated more frequently in interchromosomal fusions (6, 3, 1, and 6 times, respectively) than chromosomes 22, 21, X, and 18 (1, 0, 0, and 0 times, respectively), suggesting a similar pattern for individual tumors and the entire study group.

Although the CT map is tentative, it allows the examination of some basic and clinical questions. For example, is there an association between CTs and chromosome length, indicated by the number of base pairs per chromosome? Chromosomes differ not only in length but also in gene density, indicated as genes/Mb per chromosome. Hence, is there an association between CTs and gene density? Malignancies contain 3 main types of genetic alterations, namely, point mutations, copy number aberrations (CNAs), and rearrangements, which include inter- and intrachromosomal translocations. Is there an association of CTs with any of these alterations? The spatial association of all genetic alterations in breast cancer with the CT map is summarized in Table 5, in which we combined inter- and intrachromosomal fusions in a total translocation group. The studies used to generate the tentative CT map from interchromosomal fusions provided incomplete information about the other 2 main classes of genetic alterations, namely, point mutations and CNAs. Therefore, we chose a different genome-wide database of 2509 breast cancers from the METABRIC trial[13,14] (www.cbioportal.org) to examine the distribution of these genetic alterations in the CT map. Figure 2A to C shows scatterplots of total translocations, point mutations, and CNAs by the number of base pairs for each chromosome. The notation adjacent to each dot gives its chromosome number. Regression analysis revealed that there was no significant association between chromosome length and the number of total translocations, point mutations or CNAs (Figure 2A to C). In other words, there is no association between chromosome number or length and frequency of genetic alteration. Figure 2D to F is analogous to Figure 2A to C but is with respect to the gene density of each chromosome. The gene density was significantly correlated with total translocations ($P = .02$), but not with point mutations $P = .19$ and CNAs $P = .62$. Finally, the association of the three genetic alterations with the CT map inferred from the interchromosomal fusions was significant, ie, total translocations $P = 7 \times 10^{-11}$, point mutations $P = .01$, and CNAs $P = .002$ (Figure 2H to I). These results suggest that genetic alterations are more frequently located on chromosomes near each other than on chromosomes separated by longer distances.

Discussion

Chromosomal translocations resulting in gene fusions are thought to play critical carcinogenic roles via various mechanisms, such as oncogene activation, tumor suppressor deletion/downregulation, and the creation of novel proteins capable of altering cellular pathways.[16] For example, fusion events are frequently characterized by 1 gene that is expressed at relatively high levels in the nonfused state fused to another gene that is expressed at relatively low levels in the nonfused state. According to the oncogenic gene fusion model, the strong promoter of the 5′ gene upregulates the expression of the oncogenic 3′ gene.[10]

Somatic mutations including translocations occur nonrandomly across the genome. Causes of mutations include exposure to DNA damaging agents and deficiencies in DNA repair pathways, but the mechanisms that underpin the nonrandom distribution of mutations across the genome remain to be defined.[17] To better understand the nonrandom distribution, we decided to investigate the extent of the distribution and develop a tentative physical map of mutations and

Table 3. Breast cancers with 9 or more interchromosomal translocations.[10].

FILENAME	GENE_A	GENE_B	A_CHR	B_CHR	FUSIONPAIR
TCGA-AC-A6IW-01A-12R-A33J-07	ALPL	FILIP1L	1	3	ALPL__FILIP1L
TCGA-AC-A6IW-01A-12R-A33J-07	LRP8	TMEM217	1	6	LRP8__TMEM217
TCGA-AC-A6IW-01A-12R-A33J-07	RAP1GAP	NIT2	1	3	RAP1GAP__NIT2
TCGA-AC-A6IW-01A-12R-A33J-07	ZZZ3	NCK1	1	3	ZZZ3__NCK1
TCGA-AC-A6IW-01A-12R-A33J-07	FNDC3B	UCK2	3	1	FNDC3B__UCK2
TCGA-AC-A6IW-01A-12R-A33J-07	SIAH2	KIAA1430	3	4	SIAH2__KIAA1430
TCGA-AC-A6IW-01A-12R-A33J-07	GALNT10	IQCA1	5	2	GALNT10__IQCA1
TCGA-AC-A6IW-01A-12R-A33J-07	TRPS1	PAFAH1B1	8	17	TRPS1__PAFAH1B1
TCGA-AC-A6IW-01A-12R-A33J-07	BRWD1	ABCA10	21	17	BRWD1__ABCA10
TCGA-AN-A0AM-01A-11R-A034-07	KIAA1244	AC093158.1	6	1	KIAA1244__AC093158.1
TCGA-AN-A0AM-01A-11R-A034-07	B4GALNT3	SLCO4A1	12	20	B4GALNT3__SLCO4A1
TCGA-AN-A0AM-01A-11R-A034-07	CAND1	TUSC5	12	17	CAND1__TUSC5
TCGA-AN-A0AM-01A-11R-A034-07	FRS2	NRG3	12	10	FRS2__NRG3
TCGA-AN-A0AM-01A-11R-A034-07	USP15	PKP4	12	2	USP15__PKP4
TCGA-AN-A0AM-01A-11R-A034-07	RPAIN	LRRC27	17	10	RPAIN__LRRC27
TCGA-AN-A0AM-01A-11R-A034-07	TBCD	NFIX	17	19	TBCD__NFIX
TCGA-AN-A0AM-01A-11R-A034-07	MBTPS2	EDEM2	X	20	MBTPS2__EDEM2
TCGA-AN-A0AM-01A-11R-A034-07	POLA1	TIMELESS	X	12	POLA1__TIMELESS
TCGA-AO-A03V-01A-11R-A115-07	GPR35	MELK	2	9	GPR35__MELK
TCGA-AO-A03V-01A-11R-A115-07	SNX16	PCCA	8	13	SNX16__PCCA
TCGA-AO-A03V-01A-11R-A115-07	UNC13B	DNER	9	2	UNC13B__DNER
TCGA-AO-A03V-01A-11R-A115-07	MGMT	LGI4	10	19	MGMT__LGI4
TCGA-AO-A03V-01A-11R-A115-07	PLEKHA1	FAM187B	10	19	PLEKHA1__FAM187B
TCGA-AO-A03V-01A-11R-A115-07	SLTM	ATE1	15	10	SLTM__ATE1
TCGA-AO-A03V-01A-11R-A115-07	GNAL	CSNK1E	18	22	GNAL__CSNK1E
TCGA-AO-A03V-01A-11R-A115-07	ZNF586	SRC	19	20	ZNF586__SRC
TCGA-AO-A03V-01A-11R-A115-07	ZNF587	SRC	19	20	ZNF587__SRC
TCGA-AQ-A04L-01B-21R-A10J-07	MOBKL1A	PPAPDC1A	4	10	MOBKL1A__PPAPDC1A
TCGA-AQ-A04L-01B-21R-A10J-07	C7orf46	MUC4	7	3	C7orf46__MUC4
TCGA-AQ-A04L-01B-21R-A10J-07	OSBPL3	CORIN	7	4	OSBPL3__CORIN
TCGA-AQ-A04L-01B-21R-A10J-07	RAPGEF5	PPAPDC1A	7	10	RAPGEF5__PPAPDC1A
TCGA-AQ-A04L-01B-21R-A10J-07	ROR2	KCNC2	9	12	ROR2__KCNC2
TCGA-AQ-A04L-01B-21R-A10J-07	ANK3	SP2	10	17	ANK3__SP2
TCGA-AQ-A04L-01B-21R-A10J-07	ANK3	CHN2	10	7	ANK3__CHN2
TCGA-AQ-A04L-01B-21R-A10J-07	GFRA1	DGKB	10	7	GFRA1__DGKB
TCGA-AQ-A04L-01B-21R-A10J-07	IGHMBP2	C7orf72	11	7	IGHMBP2__C7orf72
TCGA-AQ-A04L-01B-21R-A10J-07	PAK1	ARHGEF18	11	19	PAK1__ARHGEF18
TCGA-AQ-A04L-01B-21R-A10J-07	CDH1	NPFFR2	16	4	CDH1__NPFFR2
TCGA-AQ-A04L-01B-21R-A10J-07	ARHGEF18	C11orf67	19	11	ARHGEF18__C11orf67
TCGA-AQ-A04L-01B-21R-A10J-07	PGPEP1	KDM4C	19	9	PGPEP1__KDM4C

Table 4. Matrix of Interchromosomal fusion events in tumors with 9 or more interchromosomal translocations.

	CHR17	CHR11	CHR8	CHR1	CHR6	CHR20	CHR19	CHR12	CHR10	CHR3	CHR2	CHR7	CHR16	CHR9	CHR14	CHR4	CHR5	CHR15	CHR13	CHR22	CHR21	CHRX	CHR18
chr17		0	1	0	0	0	1	1	2	0	0	0	0	0	0	0	0	0	0	0	1	0	0
chr11	0		0	0	0	0	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
chr8	1	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
chr1	0	0	0		2	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0
chr6	0	0	0	2		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
chr20	0	0	0	0	0		2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
chr19	1	2	0	0	0	2		0	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0
chr12	1	0	0	0	0	1	0		1	0	1	0	0	1	0	0	0	0	0	0	0	1	0
chr10	2	0	0	0	0	0	2	1		0	0	3	0	0	0	1	0	1	0	0	0	0	0
chr3	0	0	0	4	0	0	0	0	0		0	1	0	0	0	1	0	0	0	0	0	0	0
chr2	0	0	0	0	0	0	0	1	0	0		0	0	2	0	0	1	0	0	0	0	0	0
chr7	0	1	0	0	0	0	0	0	3	1	0		0	0	0	1	0	0	0	0	0	0	0
chr16	0	0	0	0	0	0	0	0	0	0	0	0		0	0	1	0	0	0	0	0	0	0
chr9	0	0	0	0	0	0	1	1	0	0	2	0	0		0	0	0	0	0	0	0	0	0
chr14	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
chr4	0	0	0	0	0	0	0	0	1	1	0	1	1	0	0		0	0	0	0	0	0	0
chr5	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0		0	0	0	0	0	0
chr15	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0		0	0	0	0	0
chr13	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0
chr22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	1
chr21	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0
chrX	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0		0
chr18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	
Total	6	3	1	6	0	4	3	4	5	2	3	1	1	0	0	0	0	0	0	1	0	0	0

Table 5. Association of chromosome territory map with genetic alterations.

CHROMOSOME	BASE PAIRS[a]		DENSITY	TOTAL TRANSLOCATIONS		POINT MUTATIONS		CNAS	
	N	%	GENES/MB	N	%	N	%	N	%
17	83257441	2.7	13.68	524	13.8	1586	12.1	6435	9.3
11	135086622	4.5	9.16	329	8.7	463	3.5	3184	4.6
8	145138636	4.8	4.36	238	6.3	408	3.1	12978	18.8
1	248956422	8.2	7.86	370	9.8	841	6.4	22234	32.2
6	170805979	5.6	5.86	235	6.2	1014	7.7	2181	3.2
20	64444167	2.1	8.22	149	3.9	100	0.8	2147	3.1
19	58617616	1.9	22.53	206	5.4	616	4.7	1784	2.6
12	133275309	4.4	7.37	194	5.1	851	6.5	1900	2.7
10	133797422	4.4	9.16	161	4.3	607	4.6	860	1.2
3	198295559	6.5	5.2	183	4.8	1525	11.6	2265	3.3
2	242193529	8.0	4.87	174	4.6	656	5.0	765	1.1
7	159345973	5.3	5.37	114	3.0	917	7.0	1955	2.8
16	90338345	3.0	8.67	99	2.6	549	4.2	3907	5.7
9	138394717	4.6	5.3	121	3.2	294	2.2	1136	1.6
14	107043718	3.5	5.37	123	3.2	663	5.0	651	0.9
4	190214555	6.3	3.77	89	2.4	138	1.0	373	0.5
5	181538259	6.0	4.62	97	2.6	809	6.2	1189	1.7
15	101991189	3.4	5.33	103	2.7	192	1.5	581	0.8
13	114364328	3.8	2.65	57	1.5	154	1.2	585	0.8
22	50818468	1.7	8.15	65	1.7	216	1.6	324	0.5
21	46709983	1.5	4.43	53	1.4	133	1.0	330	0.5
X	156040895	5.1	5.19	57	1.5	311	2.4	899	1.3
18	80373285	2.7	3.29	46	1.2	105	0.8	460	0.7
	3031042417	100.0		3787	100.0	13148	100.0	69123	100.0

Abbreviation: CNAs = copy number aberrations.
[a]Ensembl Genome Browser Release 87 (December 2016).

their chromosomal locations. As investigative tool we used interchromosomal translocations, which are unique among mutations because they involve 2 chromosomes rather than a single chromosome. This unique characteristic allowed us to determine the number of fusions between chromosomes, made possible only because they are neighbors. We used genomic data on interchromosomal translocations in 1546 primary breast cancers and found that the number of interchromosomal contacts was highest for chromosomes 17, 11, 8, and 1, whereas 22, 21, X, and 18 had infrequent contacts with other chromosomes (Table 2; Figure 1). The observation that the 4 chromosomes 17, 11, 8, and 1 accounted for 32.6% (346/1062) fusion events suggests a shorter distance between them compared with 22, 21, X, and 18, which accounted for only 7.2% (76/1062), implying a longer distance between the latter chromosomes. The deduced physical proximity between chromosomes allowed us to create a tentative CT map with centrally located chromosomes 17, 11, 8, and 1, and chromosomes 22, 21, X, and 18 positioned near the nuclear periphery. We want to emphasize that the constructed CT map does not represent a definitive view of the nuclear interior but a model inferred from the frequency of interchromosomal translocations and the distribution of these events. It is plausible that a group of chromosomes with frequent interchromosomal

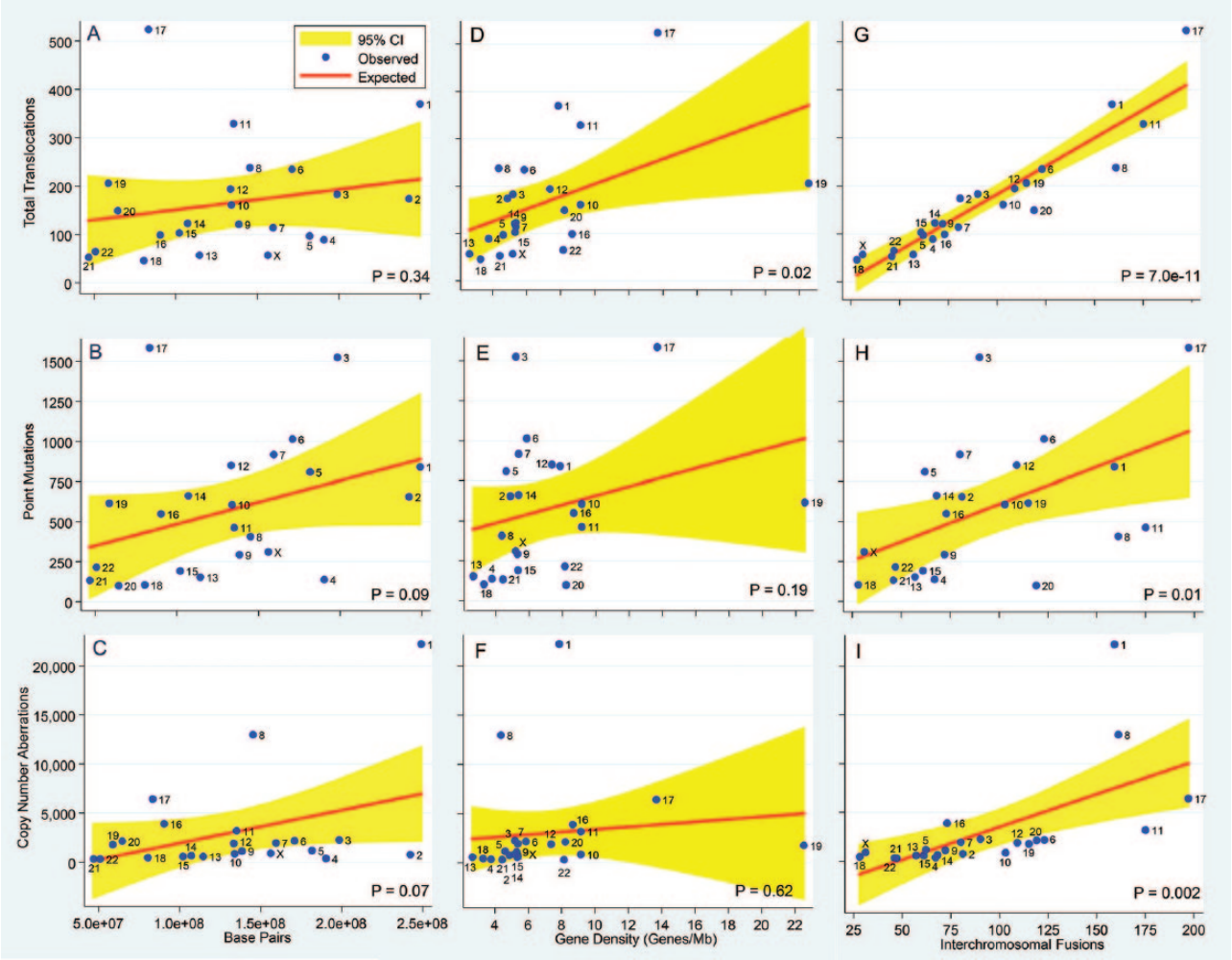

Figure 2. Scatterplots of total translocations, point mutations, and copy number aberrations (CNAs) against number of base pairs (A, B, & C), chromosome densities (D, E, & F), and interchromosomal fusions (G, H, & I) for each chromosome. Each chromosome is labeled adjacent to its dot. The red lines show the expected number of events under simple linear regression. *P* values are with respect to the null hypotheses and the slopes of these regressions are zero.

fusions, such as 17, 11, 8, and 1, are physically in proximity in the nucleus, though not necessarily in the center. We argue that it is easier to pack objects closer together as a cluster in the center of a 3D space than near its boundary, but cannot rule out a peripheral location. This means that we cannot precisely predict, for example, where chromosomes 17 and 18 are located in the nucleus. Both chromosomes have approximately the same number of base pairs and are of equal length (Table 3), but chromosome 17 participates in 7 times more interchromosomal translocations than chromosome 18 (197 and 28, respectively; Table 2). The number of fusions with neighboring chromosomes makes it likely that chromosome 17 is more centrally located than chromosome 18. The CT pattern observed for the entire group of breast cancers was also found in individual tumors containing multiple interchromosomal translocations (Tables 3, 4).

The aggregated database shown in Table 2 and Figure 1 is based on the analysis of 5 genome-wide studies. A recent transcriptomic study performed RNA sequencing of 55 breast cancer samples and identified 370 fusion genes.[18] Interestingly, in spite of the difference in methodology, over half of the fusion genes were found to be situated on chromosomes 17, 8, 1, 20, 6, and 11, similar to the results of our study.

CTs have mostly been studied in monolayer cultures using tumor cell lines, fibroblasts, or lymphocytes.[19] Although these cells offer several technical advantages, even "normal" cells in monolayer cultures have a nuclear structure different from that of the same cells in tissues.[20] This complicates comparisons between normal and neoplastic cell nuclei and might obscure important points of difference in spatial organization. CT maps have been reported of nonepithelial cells, eg, fibroblasts and lymphoblastoid cells.[6,21] There is limited information available on mammary epithelial cells. One study examined the nuclear positioning of CTs in the benign MCF10A cell line and its malignant counterpart MCF10CA1a.[22] The comparison revealed different CT patterns between 10A and CA1a cells. The study was limited to 9 chromosomes (1, 4, 11, 12, 15, 16, 18, 21 and X) and only determined interactions between

individual chromosome pairs, making it difficult to draw comparisons to the present investigation. Edgren et al[23] investigated chromosomal translocations and fusion genes in 4 breast cancer cell lines (BT-474, KPL-4, MCF-7, SK-BR-3). The authors identified and validated 27 fusion genes of which 11 were found in BT-474, while the other cells contained 3, 3, and 10, respectively. Eight of the 27 fusion genes arose from interchromosomal translocations; interestingly, chromosome 17 participated in 6 of the 8 fusion pairs.

In addition to the genetic evidence presented here, there is molecular biological evidence for spatial CT arrangement. The estrogen receptor α (ERα) regulates the transcription of thousands of genes by dynamic long-range chromatin interactions.[24] ERα dimers are recruited to multiple estrogen response elements (EREs), which interact with one another and possibly other factors such as FoxA1 and RNA polymerase II to form chromatin-looping structures around target genes.[25] A genome-wide chromatin interaction analysis using paired-end tag sequencing (ChIA-PET) applied to MCF-7 breast cancer cells identified numerous ERα-bound chromatin interaction regions, in which distal ERα-binding sites interact with proximal sites, forming chromatin loops. Sequencing of large-scale ChIA-PET libraries identified intrachromosomal as well as interchromosomal interactions.[25] This 3D architecture may partition individual genes into subcompartments of nuclear space such as interaction-anchor-associated genes and interaction-loop associated genes for differential transcriptional activation or repression.

The deduced CT map, even as a tentative model, allowed us to examine the spatial distribution of somatic mutations in breast cancer. Regression analysis revealed that there was no significant association between chromosome length and the number of total translocations, point mutations, or CNAs (Figure 2A to C). Interestingly, gene density was significantly correlated with total translocations ($P = .02$), but not with point mutations $P = .19$ and CNAs $P = .62$ (Figure 2D to F). Finally, the association of the 3 genetic alterations with the CT map inferred from the interchromosomal fusions was significant, ie, total translocations $P = 7 \times 10^{-11}$, point mutations $P = .01$, and CNAs $P = .002$ (Figure 2G to I). In other words, the number of genetic alterations on the deduced central chromosomes exceeded expectations, whereas the corresponding number on the deduced peripheral chromosomes was lower than expected.

This study has several limitations. Although CTs are characterized by a certain constancy of nuclear position, we want to caution that there probably is some variation between individual tumors, which our analysis does not ascertain. Similarly, we do not address the CT architecture of normal mammary epithelium and therefore cannot define specific topographic changes associated with malignancy. Interestingly, 4C experiments and 3D FISH analysis indicated that the breast carcinoma amplified sequence (BCAS1-4) genes, which are located on chromosomes 1, 17, and 20, are near each other in normal human mammary epithelial cells (HMEC).[26] The analysis of malignant MCF-7 cells revealed frequent translocations of BCAS genes. These results suggest that the spatial proximity of the BCAS genes in normal breast cells contributes to their frequent oncogenic translocations. Presently, it is poorly understood on what basis CTs assume their position in the nuclear space and which mechanisms maintain this position. In MCF-7 cells, estrogens were shown to play a role in inducing chromosome movements by regulating the actions of nuclear motor proteins, such as myosin-I and dynein light chain 1,[27] but a role in maintaining chromosome positions is unknown. The nonrandom positioning of CTs and somatic mutations observed in this study suggests some functional relevance of the association between genetic alterations and CT map. Nuclear receptors possibly play a role, as the androgen receptor was shown to trigger intra- and interchromosomal interactions in LNCaP prostate cancer cells. Genotoxic stress generated site-selective double-strand breaks and nonrandom translocations.[28] These findings suggest that nuclear receptors may induce chromosomal proximity, thereby increasing the probability of a gene fusion when subjected to DNA damaging agents.[29] A recent study provided genomic-epidemiologic evidence that estrogens promote breast cancer development.[30] The underlying mechanisms remain unknown. Given the many actions of estrogens, future investigations will be required to show whether any effect on CT architecture may play a contributory role.

Conclusions

We used interchromosomal translocations to develop a tentative CT map of breast cancer. This study indicates a spatial association of genetic alterations with CTs in breast cancer.

Author Contributions

FFP, WDD, and PSC conceived and designed the study. FFP, WDD, PSC helped in development of methodology. FFP, WDD, PSC analysed the data (e.g., statistical analysis, biostatistics, computational analysis) and also helped in the interpretation of data. FFP, WDD, PSC contributed to the writing, reviewing, and/or revision of the manuscript. FFP, WDD, PSC helped in administrative, technical, or material support (i.e., reporting or organizing data, constructing databases) of the data. FFP, WDD, PSC study supervised the paper.

ORCID iD

Fritz F Parl https://orcid.org/0000-0001-9244-3447

REFERENCES

1. Nik-Zainal S, Davies H, Staaf J, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature*. 2016;534:47–54.
2. Cremer T, Cremer M. Chromosome territories. *Cold Spring Harb Perspect Biol*. 2010;2:a003889.
3. Branco MR, Pombo A. Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLoS Biol*. 2006;4:e138.
4. Meaburn KJ, Misteli T. Cell biology: chromosome territories. *Nature*. 2007;445:379–781.
5. Beagrie RA, Scialdone A, Schueler M, et al. Complex multi-enhancer contacts captured by genome architecture mapping. *Nature*. 2017;543:519–524.

6. Lieberman-Aiden E, van Berkum NL, Williams L, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. 2009;326:289–293.
7. Misteli T. Beyond the sequence: cellular organization of genome function. *Cell*. 2007;128:787–800.
8. Kim J, Kim S, Ko S, et al. Recurrent fusion transcripts detected by whole-transcriptome sequencing of 120 primary breast cancer samples. *Genes Chromosomes Cancer*. 2015;54:681–691.
9. Ma Y, Ambannavar R, Stephans J, et al. Fusion transcript discovery in formalin-fixed paraffin-embedded human breast cancer tissues reveals a link to tumor progression. *PLoS ONE*. 2014;9:e94202.
10. Yoshihara K, Wang Q, Torres-Garcia W, et al. The landscape and therapeutic relevance of cancer-associated transcript fusions. *Oncogene*. 2015;34:4845–4854.
11. Banerji S, Cibulskis K, Rangel-Escareno C, et al. Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature*. 2012;486:405–409.
12. Matissek KJ, Onozato ML, Sun S, et al. Expressed gene fusions as frequent drivers of poor outcomes in hormone receptor-positive breast cancer. *Cancer Discov*. 2018;8:336–353.
13. Curtis C, Shah SP, Chin SF, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*. 2012;486:346–352.
14. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov*. 2012;2:401–404.
15. StataCorp. *StataStatistical Software: Release 15*. College Station, TX: StataCorp; 2017.
16. Mertens F, Johansson B, Fioretos T, Mitelman F. The emerging complexity of gene fusions in cancer. *Nat Rev Cancer*. 2015;15:371–381.
17. Schuster-Bockler B, Lehner B. Chromatin organization is a major influence on regional mutation rates in human cancer cells. *Nature*. 2012;488:504–507.
18. Fimereli D, Fumagalli D, Brown D, et al. Genomic hotspots but few recurrent fusion genes in breast cancer. *Genes Chromosomes Cancer*. 2018;57:331–338.
19. Rao SS, Huntley MH, Durand NC, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*. 2014;159:1665–1680.
20. Wiech T, Timme S, Riede F, et al. Human archival tissues provide a valuable source for the analysis of spatial genome organization. *Histochem Cell Biol*. 2005;123:229–238.
21. Wang S, Su JH, Beliveau BJ, et al. Spatial organization of chromatin domains and compartments in single chromosomes. *Science*. 2016;353:598–602.
22. Fritz AJ, Stojkovic B, Ding H, et al. Wide-scale alterations in interchromosomal organization in breast cancer cells: defining a network of interacting chromosomes. *Hum Mol Genet*. 2014;23:5133–5146.
23. Edgren H, Murumagi A, Kangaspeska S, et al. Identification of fusion genes in breast cancer by paired-end RNA-sequencing. *Genome Biol*. 2011;12:R6.
24. Liu MH, Cheung E. Estrogen receptor-mediated long-range chromatin interactions and transcription in breast cancer. *Mol Cell Endocrinol*. 2014;382:624–632.
25. Fullwood MJ, Liu MH, Pan YF, et al. An oestrogen receptor-alpha-bound human chromatin interactome. *Nature*. 2009;462:58–64.
26. Zeitz MJ, Ay F, Heidmann JD, et al. Genomic interaction profiles in breast cancer reveal altered chromatin architecture. *PLoS ONE*. 2013;8:e73974.
27. Hu Q, Kwon YS, Nunez E, et al. Enhancing nuclear receptor-induced transcription requires nuclear motor and LSD1-dependent gene networking in interchromatin granules. *Proc Natl Acad Sci U S A*. 2008;105:19199–19204.
28. Lin C, Yang L, Tanasa B, et al. Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. *Cell*. 2009;139:1069–1083.
29. Mani RS, Tomlins SA, Callahan K, et al. Induced chromosomal proximity and gene fusions in prostate cancer. *Science*. 2009;326:1230.
30. Parl FF, Crooke PS, Plummer WD, Dupont WD. Genomic-epidemiologic evidence that estrogens promote breast cancer development. *Cancer Epidemiol Biomarkers Prev*. 2018;27:899–907.