TMEVPG1, a Long Noncoding RNA within the Immune System

By

Sarah Parker Collier

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

Microbiology and Immunology

May, 2014

Nashville, Tennessee

Approved:

Eric Sebzda, Ph.D.

Andrew J. Link, Ph.D.

Amy S. Major, Ph.D.

Anthony P. Weil, Ph.D.

James G. Patton, Ph.D.

Thomas M. Aune, Ph.D.

ACKNOWLEDGEMENT

To all of my collaborators: Dr. Mark Boothby and Dr. Christopher Williams, Dr. Louise Rollins-Smith and Dr. J. Scott Fites, Dr. Rachel Henry, and Dr. Danyvid Olivares-Villagomez, I appreciate your willingness to help me succeed. Also, I am indebted to my colleagues and friends in the Department of Pathology, Microbiology and Immunology; Infectious Disease; and Cell and Developmental Biology for countless laughs and positive inspiration.

To Dr. Roger Chalkley, Dr. Kim Petrie, and Kate Stuart thank you for the amazing outreach through the Biomedical Research and Education Training office.

I am forever indebted to Dr. Derek Bruff, Dr. Nancy Chick, Dr. Cynthia Brame and Dr. Joe Bandy and the entire Vanderbilt Center for Teaching, my "Land of Oz." Your programs and resources were constant inspiration to me. Thank you for all you do and taking an interest in my passion for teaching.

To Dr. Phillip E. Funk, you have been a fabulous mentor, counsel and friend over these years since the first day of Biology 101 in my first semester at DePaul University. Thank you for realizing my potential and guiding my career in science. You were 100% correct, the lab was where I needed to be. And I'll never stop believing!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

H4Ac: Histone four acetylation

H3K27me3: Histone three lysine 27 tri-methylation

H3K4me1, me2 or me3: Histone 3 lysine four mono-, di-, tri-methylation

H3K9Ac: Histone three lysine 9 acetylation

HATs: Histone acetyltransferases

HDACs: Histone deacetylases

HMT: Histone methyltransferase

HDMT: Histone demethyltransferases

CNSs: Conserved noncoding sequences

lncRNA: long noncoding RNA

IFNG: human interferon gamma mRNA transcript or gene where noted

Ifng: mouse interferon gamma gene mRNA transcript or gene where noted

IFN-γ; interferon gamma protein

CTCF: CCCTC-binding factor

Stat4: Signal transducer and activator of transcription 4

T-bet: T-box protein 21

Eomes: Eomesodermin

SHH: Sonic Hedge Hog

HS: Hypersensitivity site

Th cell: T helper cell

Tc1: Cytotoxic T cell type 1

CHAPTER I

Introduction

Overview

The human genome consists of approximately 20,000 to 25,000 protein-coding genes, accounting for less than 5% of the total genes in the genome (1-3). Predominantly noncoding, it is now appreciated that a significant portion of the transcriptional regulation of protein-coding genes resides within these noncoding segments (3-7). From an evolutionary perspective, it is tempting to speculate why these regions were conserved if they are devoid of purpose (6, 7). It is known that terminal differentiation of cells is dependent upon the coordinate expression of gene networks forming distinct cell lineages and compartmentalized systems; however, the role of noncoding regulatory elements as well as their transcriptional products governing this process remains elusive particularly as it pertains to discovery of new classes of regulatory molecules (8). To uncover the mechanisms by which noncoding elements participate in gene regulation, we utilize the process of CD4$^+$ T cell polarization as a model system where effector fate is established by noncoding regulation of hallmark cytokine genes.

Noncoding Regulation and the Histone Code

Noncoding regulation of genes includes epigenetic architecture, distal conserved elements in the genome, and regulatory RNAs. Arguably, the apical layer of noncoding mechanisms governing gene regulation is access control (9). Within the nuclear

compartment, genomic DNA is efficiently wrapped around histones creating the compact nucleosome promoting further higher order structure (9-12). The octameric core of heterodimers of histone H2A, H2B, H3 and H4 proteins have flexible N-terminal amino acid strands commonly referred to as "tails" that sustain covalent, yet reversible, post-translational modifications resulting in the formation of heterochromatin (a compact, transcriptionally silent genomic status) or euchromatin (a structured, yet more relaxed state permissive of transcription) (Figure 1-1) (9, 10, 12-15). Further, patterns of histone modifications correlate with functional regions of gene loci. Histone four acetylation (H4Ac) is found across actively transcribed genes while histone three, lysine 27 trimethylation (H3K27me3) or H3K9 di-or tri-methylation is often associated with repression (9, 16-18). H3K4me2 marks correlate with poised or actively transcribed genes while H3K4me1 or H3K4me3 marks denote enhancer elements as well as active promoters (Figure 1-1*B*) (9, 13, 17-20). In many cases, hallmark genes of linage-specific networks possess bivalent histone modifications particularly in precursor or founders cells; however, upon commitment histone marks are restored (9)
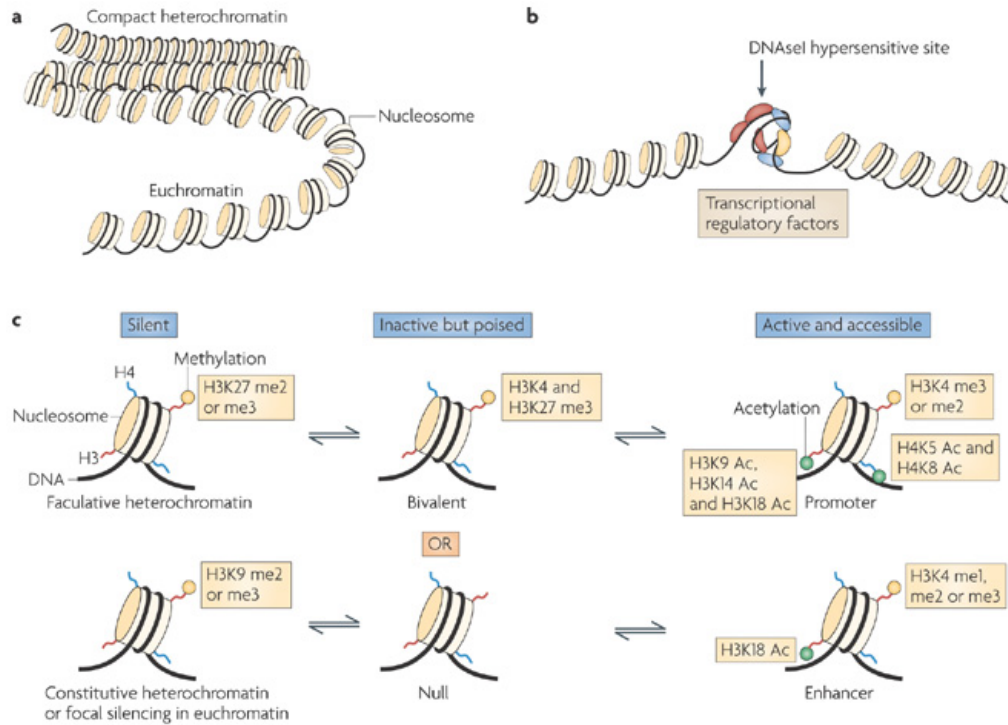
FIGURE 1-1. Post-translational modification of histone proteins influences gene expression (13).

Of the histone modifications described, methylation and acetylation remain the most understood and often correlate with repression or activation of transcription, respectively (9, 21). Not only are these modifications on key lysine residues of the histone tails effective with respect to strengthening or weakening the interaction between histone proteins and DNA but these "marks" are sustained and recognized by a global class of protein "reader" components of histone-modifying complexes, the "writers" and the "erasers" (9). The dynamic interactions between histone acetyltransferases (HATs) and deacetylases (HDACs) as well as histone methyltransferase (HMTs) and demethyltransferases (HDMTs) establish formation of these essential marks (9, 22). The H3K27me3 mark is orchestrated by EZH2 methyltransferase complexes commonly associated with polycomb group proteins (9, 17). In contrast, H3K4me3 marks are established by MLL HMT complexes often guided by members of the Set protein family (23). The formation and removal of acetyl groups on histone tails is conducted by an entirely separate cohort of proteins conferring specificity by recruitment of various histone-modifying complexes (9). At one time, it was predicted that histone marks were irreversible but it is now appreciated that histone marks are highly dynamic and dependent upon cell-specific lineages and signals governing enhancement or repression of gene regulation fortifying the histone code hypothesis (24, 25).

Similar to epigenetic modifications of histone proteins, methylation of CpG dinucleotides exhibits a second example of epigenetic regulation. DNA Methyltransferase (DMNT) enzymes catalyze the methylation mark formation and have been shown to be required for widespread silencing of gene loci (26, 27). During cell division, DMNT1

maintains the heritable patterns of DNA methylation in daughter cells while DMNT3a and 3b are considered "de novo" methyltransferase enzymes regulating gene repression (13, 28). Pertaining to cellular differentiation, precursor cells may exhibit heavy CpG methylation on lineage-specific genes until driven by external stimuli (13, 26-28). These lineage-specific signals promote removal of these marks on necessary genes pertaining to that program while promoting widespread methylation across opposing program-related genes.

Distal Regulatory Elements

Distal regulatory elements contribute significantly to productive gene expression supporting discrete stages of development. Frequently, distal conserved elements are found in close proximity (defined by 50 kilobases up-and downstream from the transcription start site of the respective gene); however, this is not a requirement as enhancer elements are found upwards of a full megabase pair away as evidenced by the gene encoding Sonic the Hedgehog (SHH) (29-31). Moreover, the SHH model in humans establishes an important point regarding the magnitude of impact that distal noncoding sequences have on gene regulation (29-31). For example, one mutation within this long-range distal enhancer results in preaxial polydactyly a developmental phenotype where individuals will develop an extra thumb due to overexpressed SHH (29-31). Certainly, SHH-induced preaxial polydactyly is not the only case of mutations within distal regulatory elements as it is now appreciated that when considering the diverse polymorphisms of the human population,

small alterations at given nucleotides across the genome often residing within the noncoding regions, have significant implications to human health that will continue to be resolved with deeper probing of the genome on an individualized level, namely in the clinic (30, 31).

By definition, distal regulatory elements demonstrate conservation across species, are sensitive to DNase I treatment, bear histone modifications indicative of permissive chromatin and are void of CpG dinucleotide methylation (32, 33). Transactivating factors recognizing canonical binding sequences within these conserved elements allow for regulation of protein-coding genes. Binding of transcription factors to distal sites induces recruitment of histone-modifying complexes promoting remodeling of the locus as well as recruitment of RNA polymerase favoring transcription. Whereas distance might appear to be an obstacle, even within the 50 kilobases up-and downstream of a promoter region, studies of the *beta-globin* and *IL-4* loci demonstrate that three-dimensional (3D) conformation via intrachromosomal looping brings distal enhancers into close proximity to locus control regions and promoters (34-36). In mammals, intrachromosomal looping is largely orchestrated by transcription factors such as CCCTC binding factor (CTCF) as well as cohesion proteins (36, 37). These studies demonstrate a role for the spatial and temporal regulation of genes, however; in general this process is elusive.

Noncoding RNA Regulation

The ENCyclopedia Of DNA Elements (ENCODE) Consortium demonstrates that while proteins may not be encoded within noncoding regions, active transcription occurs unanimously genome-wide (1, 2, 4). While significant in number it remains to be determined whether the majority of transcripts are biologically active. Regulatory RNA species provide a significant proportion of these numerous transcripts (38). Several well-defined RNA species include those involved in the translation of messenger RNA (mRNA) directly via enzymatic mechanisms or indirectly through maintaining levels of various RNA species (38). For example, pre-mRNA transcripts are spliced into mature mRNA by way of the spliceosome comprised of small nuclear ribonuclear proteins (snRNPs). Ribosomal RNA (rRNA) and transfer RNA (tRNA) are directly responsible for mRNA translation while small nucleolar RNAs (snoRNA) regulate rRNA and tRNA processing in the nucleus (38). While these aforementioned categories contribute to maturation of mRNA and subsequent translation, microRNA and small interfering RNA (siRNA) species modulate gene expression by directly interacting with mRNA transcripts (39-41). miRNA are highly conserved RNA molecules spanning approximately 22 nucleotides in length (41). Through complementary base pairing with the 3' untranslated regions of mRNAs, miRNA's promote mRNA transcript instability inhibiting translation (39-41). siRNAs utilize complementary base pairing leading to the destruction of RNA transcripts (39, 40, 42, 43). Although these brief points highlight the fundamental aspects to small noncoding RNA regulation of genes, new species are continuously reported in the literature (Figure 1-2).
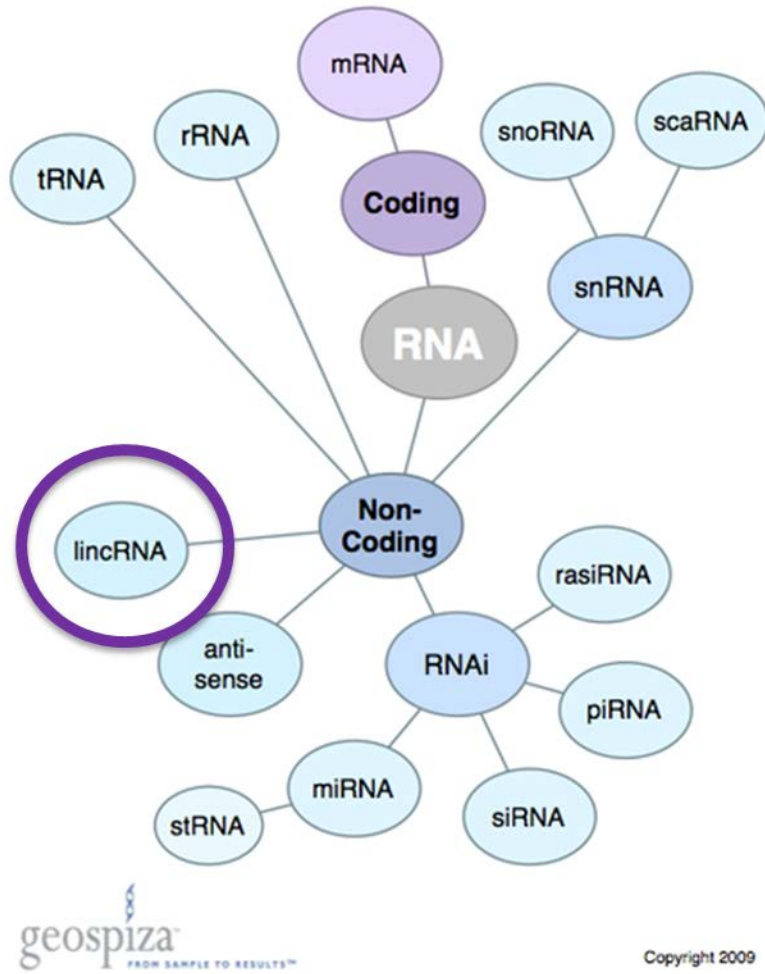
FIGURE 1-2. Noncoding RNA species.

Long Noncoding RNAs

Within the regulatory RNA network, a new species was recently described (5, 44). Identified through next generation sequencing of the transcriptome based upon H3K4me3 promoter marks and H3K36me3 marks through the gene body, thousands of long noncoding RNAs (lncRNAs) were found to map across the human and the mouse genome (44-45). At present, a universal naming system is lacking therefore lncRNAs may also be referred to as large noncoding or long inter-or intragenic noncoding RNAs. To capture all species I will refer to the greater population as lncRNAs. LncRNA expression is not exclusive to mammals, as transcripts are found in plants as well as fungi (46). Despite the infancy of the field, lncRNAs are implicated in nearly all compartmentalized phases of cellular life particularly evident in cell lineage commitment.

Found predominantly within the nuclear compartment, lncRNAs are transcribed generally in the antisense direction relative to surrounding protein-coding genes by RNA polymerase II and exist as an mRNA-like transcript with multiple exons; however, lack protein-coding potential (5, 47). Transcript length is greater than 200 nucleotides, substantially longer than other regulatory RNAs (47). Most lncRNAs undergo post-transcriptional modifications such as splicing, addition of the 5' cap and 3' polyadenylation; however, the purpose of these modifications are poorly understood for these noncoding transcripts (47). LncRNA transcript stability is variable despite exhibiting lower transcript expression relative to comparable mRNAs (48). Sequence conservation is variable and highly divergent (5). Despite this, most genes encoding lncRNAs are located within close proximity to their putative target genes (46, 49, 50). Similar to protein-coding

gene loci, these regions are enriched for DNase I hypersensitivity as well as histone modifications permissive for transcription (32, 33, 46, 51). While each of the aforementioned characteristics do not apply to all known lncRNAs, unequivocally lncRNAs participate in the noncoding regulation of genes.

At present, less than 1% of lncRNAs are functionally described in the literature. The majority of known lncRNAs are assigned repressive function of target gene transcription while a smaller proportion of lncRNAs are defined as enhancers (44-45, 66, 135). LncRNAs mediate transcriptional control via RNA-protein interactions with components of histone-modifying complexes and/or transcription factors, but may also influence target gene transcription via complementary nucleotide base pairing as well as serving as precursors for smaller regulatory RNAs (137, 139). As lncRNAs are expressed in nearly all developmental systems, effort is needed towards evaluating and assigning biological function to these molecules (49, 52).

From a historical perspective, *XIST* is the prototypical lncRNA and the first to be described within the context of embryogenesis and development. Essential for gene dosage in female embryos, *XIST* is transcribed from and mediates a repressive effect *in cis* across the X chromosome to be inactivated (Xi) through a process known as X chromosome inactivation. Interacting with the polycomb repressive complex 2 (PRC2) and the transcription factor YYI, *XIST* "covers" the Xi promoting H3K27me3 covalent histone modifications resulting in transcriptional silencing (53-55). Further, this process is facilitated by five additional lncRNAs. For example, *XIST* transcription is enhanced by

Jpx, a trans-activating lncRNA but is also repressed by the anti-sense *TSIX* lncRNA providing a model for regulation of lncRNA genes by fellow lncRNAs (56, 57). Further, *TSIX* is implicated in methylation of CpG dinucleotides during X chromosome inactivation to repress the *XIST* promoter (5).

Similar to X chromosome inactivation, repressive and enhancer lncRNAs influence regulation of *HOX* genes involved in developmental patterning. *HOTAIR* is transcribed adjacent to the *HOXC* gene and mediates repression of the *HOXD* gene by serving as a molecular scaffold for PRC2-containing histone-modifying complexes as well as LSD1-containing H3K4-demethylase complexes but is unique in that it modulates the *HOXD* locus *in trans* (58-60). The *HOTAIR* mechanism provides the first evidence that lncRNAs are not restricted to neighboring genes for regulation but may mediate effects over great distances. Together, *XIST* and *HOTAIR* provide a framework by which thousands of lncRNAs were shown to associate with PRC2 in embryonic stem cells (61). *HOTTIP,* a *cis*-acting lncRNA involved in *HOX* gene regulation, activates the *HOXA* locus via interaction with theMLL1/Set1 HMT complex promoting H3K4me3 epigenetic modifications (62). Taken together, X chromosome inactivation and *HOX* gene regulation establishes the foundation upon which epigenetic regulation of genes by lncRNAs is modeled.

In addition to epigenetic modifications, lncRNAs facilitate intrachromosomal looping as a mechanism of gene regulation. The *H19* transcript promotes spatial intrachromosomal rearrangement bringing the maternal enhancer in close proximity to the

promoter of the *IGF2/H19* gene in cooperation with an *SRA* lncRNA and p68 (50, 63, 64). Additionally, a subclass of *cis*-acting lncRNA was shown to physically interact with "Mediator" to exert long-range enhancement of target *AURKA* and *SNAI1* genes through intramolecular interactions fortified by chromatin looping (65). A similar effect has been demonstrated in MCF-7 human breast cancer cells where lncRNA-dependent chromatin loops were required for enhancement of transcription (66). These studies offer evidence for additional complexity to the accepted understanding of spatial arrangement as it pertains to gene regulation.

LncRNAs may also serve as molecular responders within cell signaling networks particularly in response to environmental stress promoting apoptosis. *LncRNA-p21* and *PANDA* are located within the *CDKN1A* locus and are regulated by p53, a potent tumor suppressor. *LncRNA-p21* negatively regulates pro-apoptotic genes conferring cell cycle regulation, whereas *PANDA* responds to DNA damage by sequestration of the NF-YA transcription factor resulting in cell survival (67, 68). *NRON* functions as a negative regulator of NFAT downstream of calcium signaling in activated lymphocytes through RNA-protein complex formation, resulting in the sequestration of NFAT in the cytosol (69). The *LncRNA-LET* transcript is suppressed in hepatocellular carcinoma in the presence of a hypoxic microenvironment, an obstacle for invasive and metastatic solid tumors (70).

As expected, lncRNAs are implicated in a host of human diseases affecting millions of people worldwide including various cancers, Crohn's disease, cardiovascular disease and neurological conditions to highlight a few. *HOTAIR* and *MALAT-1* are known to be

predictive markers as well as oncogenic in cancers of the breasts, lungs, colon, prostate, and liver (71, 72). In the case of coronary artery disease leading to heart attack and sudden death, genetic risk through single nucleotide polymorphisms identified the myocardial infarction associated lncRNA transcript (*MIAT*) correlating with incidence of disease (73). From a neurological disease perspective *BACE1-AS,* the antisense transcript relative to the enzyme responsible for cleavage of the amyloid precursor protein, is elevated in post mortem samples from patients with Alzheimer's disease having implications in the Aβ plaque formation (74). Another lncRNA, *MNSP1AS* correlates strongly in patients diagnosed with autism spectrum disorder. Surely as genomic analysis of disease-related tissue proceeds, more cases of lncRNA-associated disease relatedness will emerge.

While significant effort has focused on lncRNA function in embryonic stem cells and other developmental processes, the role of lncRNAs in the immune system remains largely undefined. A foundational array found over 100 lncRNAs expressed in the discrete stages of CD8$^+$ T cell development as well as activation in both humans and mice providing strong evidence for lncRNA presence in the immune system (75). Further, a separate study reported the induction of 1,500 lncRNAs [500 are annotated] in mouse lung tissue upon viral infection with SARS (76). Comparison of influenza virus versus SARS virus infection identified 37 differentially expressed lncRNAs indicating a role for lncRNAs in immune cells particularly in response to microbial pathogenesis. LncRNA involvement within the immune system as it pertains to assignment of biological function is largely unknown and establishes an open area of investigation within this thriving field.

Effector CD4[+] T Helper Cell Polarization

Originating from the thymus, CD4[+] and CD8[+] T cells are functionally naïve yet are carefully selected to identify antigen-derived peptides in the context of major histocompatibility complex (MHC) molecules presented on the surface of antigen presenting cells (APC) (77, 78). Circulating through the blood into secondary lymph tissues, a naïve CD4[+] T cell, for example, is presented its cognate antigen by an activated APC and initiates the process of activation and effector lineage commitment (79, 80). T helper cell polarization depends on three key signals resulting in the expression of essential transcription factors driving hallmark cytokine production: ligation of the T cell receptor (TCR) followed by costimulatory surface molecules upon antigen presentation *in vivo* or immobilized anti-CD3 under experimental conditions *in vitro*, are signals one and two, respectively, mediating downstream NFAT signal transduction (13, 81). This initial cascade results in upregulation of surface molecules including the high affinity IL2 receptor (IL2R) as well as cytokine receptors (82). In addition to T cell receptor stimulation, cytokine stimulation provides the ultimate signal culminating in execution of lineage-specific fates (79, 80). Production of the inflammatory cytokines by effector T helper cells promotes the clearance of pathogens and tumors as well as support humoral immunity through antibody production by B lymphocytes (B cells) (80). At present, there are four well-defined polarization programs based upon the master transcription factors governing the production of hallmark cytokines: Th1, Th2, Th17 and T regulatory (Treg)

while origins of more recently described populations of CD4$^+$ helper cells such as the Th9

and Th22 phenotypes are less defined especially *in vivo* (Figure 1-3) (83-85).
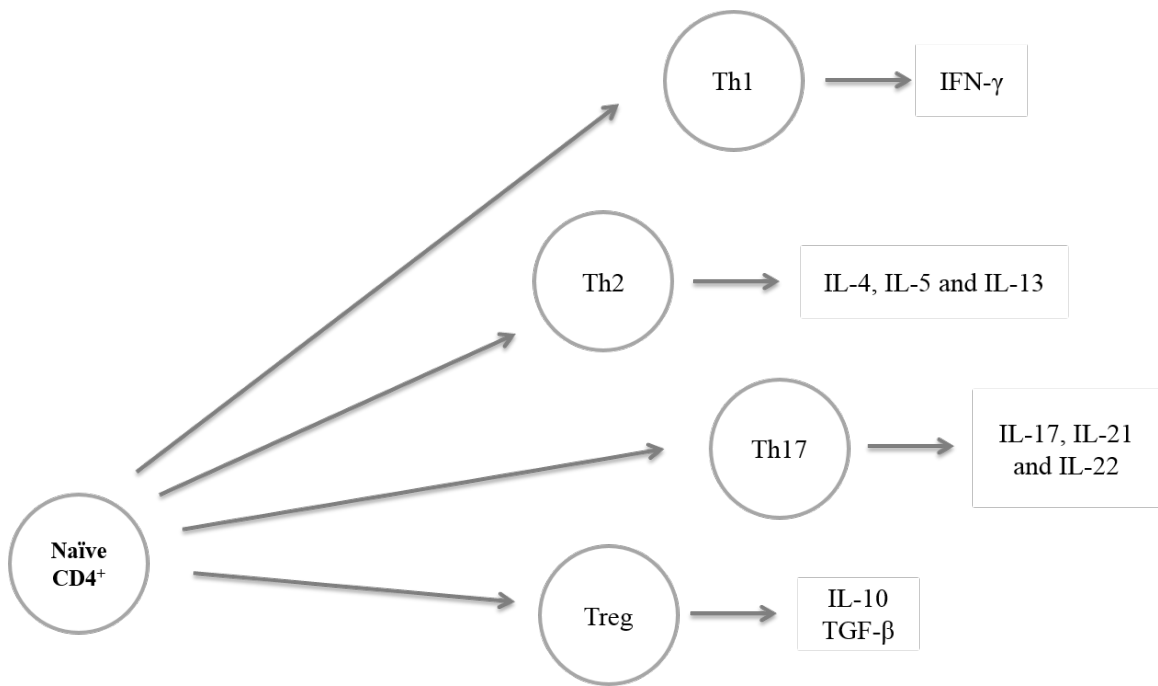
FIGURE 1-3. CD4$^+$ T cell effector subsets and hallmark cytokines.

The Th1 program is defined by interferon gamma (IFN-γ) production contributing to macrophage activation and cell-mediated adaptive immunity (27). Execution of the Th1 program is driven by the presence of IFN-γ, likely generated by natural killer (NK) or natural killer T (NKT) cells as the first responders to pathogen infection in addition to the cytokine IL-12 (26, 86, 87). Signal transduction through Stat1 culminates in T-bet upregulation and together with Stat4, downstream of the IL-12 receptor, governs the stochastic remodeling of the *IFNG* locus while concurrently repressing the hallmark genes of the other lineages (88-95). IFN-γ-producing Th1 cells are integral to the clearance of bacterial infections with pathogens such as *Listeria monocytogenes* responsible for food-borne illness leading to endocarditis, meningitis and spontaneous abortion (96).

Alternatively, the Th2 lineage is established by the cytokine IL-4 under *in vitro* culture conditions as the endogenous source of IL-4 *in vivo* is unknown. Together, Stat6 and the master Th2 transcription factor GATA-3 culminates in cellular expression of IL-4 similar to the process described for the Th1 locus (97, 98). *In vitro* and complementary *in vivo* genetic approaches resulting in the deficiency in either *Stat6* or *Gata-3* demonstrate that both are necessary while GATA-3 alone is sufficient to promote IL-4, IL-5 and IL-13 expression indicative of Th2 polarization (94, 98). Accumulation of H3K9Ac and H3K4me histone marks at these loci are dependent upon GATA-3 (16, 99). Whereas Th2 cells contribute to clearance of extracellular pathogen infections, exacerbated Th2 activities contribute to allergic responses.

While the Th1 and Th2 polarization programs establish one axis of regulation, a second axis exists between Th17 and Treg cells, the remaining effector polarization types discussed herein. In contrast to the Th1 and Th2 polarization paradigm that is seemingly fixed, the Th17 lineage is heterogeneous and dependent upon the concentration of polarizing cytokines IL-6 and TGF-β (83, 100). Molecular mechanisms leading to the production of hallmark cytokines IL-17A, IL-17F, IL-21 and IL-22 are independent of T-bet and GATA-3 driven instead by Stat3, RORγT and RORα (85, 94, 100, 101). Arguably the least understood with respect to signal transduction and gene expression as compared to the other effector subsets, it remains clear that Th17 cells have a definitive role in parasitic worm infection, are highly proinflammatory and thus are implicated in autoimmunity.

T regulatory cells (Tregs) are the remaining subset of classically defined effector cells although their status as effectors is often debated (102). This population exhibits repressive influence on the greater immune response in efforts to balance the often proinflammatory nature of the immune response by preventing exhaustive damage. Induced by TGF-β, Tregs are governed by the master transcriptional regulator Foxp3. Surface molecule expression such as CTLA-4 as well as secretion of IL-10 family cytokines mediates attenuation of immune responses. Similar to Th17 cells, Treg populations are thought to be heterogeneous particularly highlighted by a growing body of literature uncovering the molecular mechanisms of immunosuppression (77, 102).

Interferon Gamma (*IFNG*)

The interferon gamma receptor (IFN-γR) is expressed on all nucleated cells reinforcing the potency of this cytokine (103, 104). Mice lacking IFN-γR expression demonstrate increased susceptibility to *Listeria monocytogenes* as well as *Mycobacterium* species and vaccinia viral infection in addition to decreased IgG2a titers supporting an essential role for IFN-γ in several key areas of adaptive immunity (35, 77, 96, 104). IFN-γ is also implicated in a host of human diseases for example, atherosclerosis. Experimental evidence demonstrates a pro-atherogenic influence of IFN-γ on the early stages of foam cell formation and plaque development resulting in manifestations of chronic vascular inflammation promoting disease (105). Taken together these observations necessitate understanding the mechanism of *IFNG* regulation.

In addition to Th1 cells, effector CD8+ T cells, NK and NKT cells also express significant levels of IFN-γ; however, there are several distinctions in the cell intrinsic mechanisms culminating in IFN-γ expression despite sharing a common hematopoietic precursor (106, 107). NK and NKT cells are phenotypically mature upon entry into the periphery as defined by rapid expression of IFN-γ within hours of stimulation; meanwhile naïve CD8+ and CD4+ T cells require three days of polarization to produce equivalent levels (87, 106). Th1 and NK cells both require T-bet while CD8+ T cells utilize an alternative T-box transcription factor known as eomesodermin (Eomes) for *Ifng* expression (87, 89, 108). Aside from transcription factors, required distal regulatory elements are also varied among these cellular lineages.

Coordinate Expression of the *IFNG* Locus

Productive expression of *IFNG* within immune cells provides an incredibly useful experimental system to investigate mechanisms of noncoding gene regulation in lineage commitment and terminal differentiation as all currently appreciated methods are utilized to regulate the *IFNG* gene.

The *IFNG* locus, which includes 100 kilobases surrounding the gene, sustains progressive epigenetic modifications attributed to the development of primary, effector and memory Th1 or Tc1 cells. An activated, yet unpolarized Th0 cell is void of H4Ac marks across the *Ifng* locus (26, 27, 79). The balance of HDACs and HATs recruited to locus sustains this effect (27, 109). In the presence of Th1 polarizing cytokines, transcription factors such as Stat4, Hlx, Runx3 and T-bet bind across the locus at conserved elements as well as at the promoter recruiting HATs and HMTs to sustain marks permissive for transcription as this phenomenon of seeding and spreading is compromised in Stat4 or T-bet deficient animals (95, 110, 111). As Th1 or Tc1 cells transition from primary to effector-memory cells, these marks are sustained. Further, NK cells exhibit H4Ac across the *Ifng* locus directly *ex vivo* without stimulation supporting the connection between H4Ac marks and rapid IFN-γ production by effector cells (112).

In the presence of IL-4, initiation of the Th2 program also induces histone modifications across the *Ifng* gene(26). Whereas permissive marks are sustained in primary Th1 cells, lineage commitment of Th2 cells results in H3K27me2/me3 marks inducing widespread silencing in a parallel seeding and spreading method governed by Stat6,

GATA-3 and HMT complexes containing EZH2 (16, 98). Additionally, IFN-γ expression is regulated by CpG dinucleotide methylation(26). Distal sequences across the *Ifng* locus as well as the first intron of *IFNG* sustain CpG methylation in undifferentiated cells (13, 26, 27). Moreover, commitment to the Th2 or Th17 lineage induces a hypermethylation at these marks fortifying the repression of the gene while under Th1 conditions these marks are largely lost as indicated by the hypomethylation status in effector Th1 as well as NK cells (87).

Understanding how the *IFNG* locus is regulated by distal noncoding elements of the genome, particularly in the context of Th1 cell differentiation, has been an area of active investigation. The contribution of distal noncoding sequences, commonly referred to as conserved noncoding sequences (CNSs), to *IFNG* regulation has been thoroughly investigated in *in vitro* reporter assays as well as transgenic mouse model systems whereby mice express the human *IFNG* gene in the context of its endogenous locus (Figure 1-4*A*) (26, 27). For example, mice expressing a bacterial artificial chromosome (BAC) transgene containing human *IFNG* in addition to 8.6 kilobases of surrounding genome lacks Th1-specific expression of *IFNG* as Th2 cells also express IFN-γ, however; by increasing the breadth of the locus to 190 or even 210 total kilobases, the Th1-specific expression is restored (112-114). These observations support an essential role for CNSs in proper coordination of lineage-specific gene expression.

Five defined CNSs: CNS-30, CNS-22, CNS-6, CNS-2 and CNS+18/20 demonstrate differential requirements in each stage of Th1 cell lineage commitment as well

as among IFN-γ-producing cells (Figure 1-4*B*) (26-28, 86, 113-117). By way of example, the CNS-30 site located -30 kilobases upstream of *IFNG*, is required by primary, effector and memory Th1 cells as well as NKT cells (114). Recruitment of Runx3 to the CNS-30 site during early stages of T cell polarization promotes RNA polymerase II recruitment to the *IFNG* promoter while deletion of the CNS-30 site has no effect on histone marks across the *IFNG* locus (113). Similarly, this site is required by NKT cell production of *IFNG* but is not necessary for NK cells. The CNS-6 site confers epigenetic remodeling in the transition between naïve and primary Th1 cells (118). Early downstream of TCR signaling, transcription factors such as AP-1 and NFAT are recruited to the CNS-6 site promoting T-bet binding across the *IFNG* locus (79). In contrast, the CNS+18/20 site is required by memory Th1 responses and NKT interferon production while deletion only slightly impairs NK cell *IFNG* expression but has no effect on developing Th1 primary and effector cell populations (114). CNS+18/20 enhances the activity of CNS-6 in effector Th1 cells. T-bet has been shown to associate with CNS-22 which is necessary for primary CD4[+] T cells as well as CD8[+] T cells and NK subsets but not effector stimulated CD4+ T cells (116). These studies illuminate several key observations. First, NK cell production of IFN-γ does not appear to depend on any one CNS whereas Th1 cells ultimately require CNS-30, CNS-6 and CNS+18/20 depending on the polarization status (26, 27). Second, CNSs recruit essential transcriptional regulatory molecules such as Runx3 and T-bet (26). Lastly, differences between cells capable of expressing *IFNG* arise from the need for various CNSs over others demonstrating, at least for *IFNG*, that these sites are instrumental for lineage-specific expression of genes.
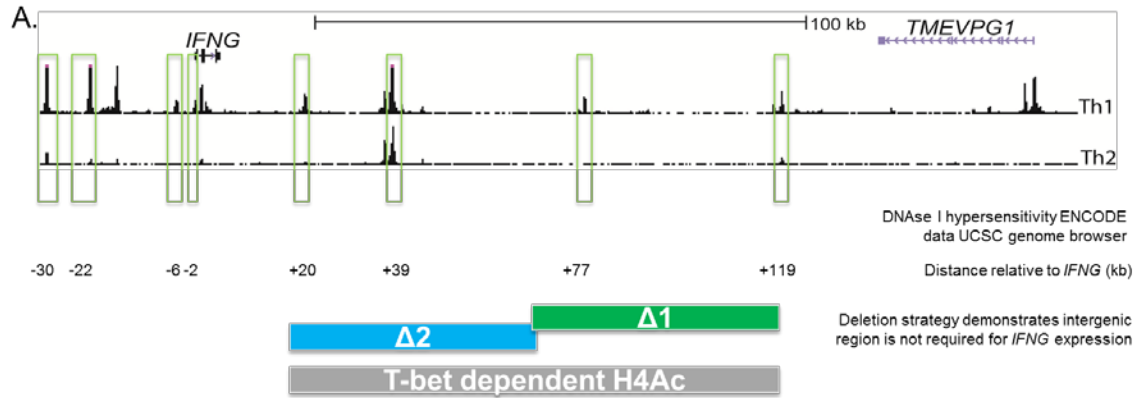
FIGURE 1-4. Analysis of the human *IFNG* locus and distal conserved noncoding sequences. *A,* CNSs across the *IFNG* locus in distance relative to the transcription start site, mapped locations of two large deletions generated to identify distal CNSs and highlighted region of T-bet-dependent H4Ac marks. *B*, Requirement of four CNSs experimentally assessed in among CD4[+], NK and NKT cells. Modified from (26).

RNA Regulation of *IFNG*

Lastly, regulatory RNAs are implicated in attenuation of immune responses particularly as it relates to *IFNG* expression. One miRNA, named miR-125b was recently demonstrated to be highly expressed in naïve T cells and was predicted to regulate 72 mRNA targets including the *IFNG* transcript (119). Further, miR-125b was demonstrated to mediate repression of *IFNG* reinforcing a naïve CD4$^+$ T cell status as overexpression significantly impaired the *IFNG* levels; however, it was determined that the endogenous downregulation of miR-125b was indicative of the transition to a effector-memory T cell state. A second miRNA shown to participate in *IFNG* regulation is miR-29, which specifically targets both T-bet and Eomes preventing expression in CD4$^+$ T cells (120).

TMEVPG1

TMEVPG1 ([Theiler's Murine Encephalitis Virus Possible Gene 1] also referred to as *NeST* [NEttoie Theiler's Pas Salmonella] or *Ifng-As1*) is the first described lncRNA of the immune system involved in regulating a master cytokine such as IFN-γ. Most inbred mouse strains are acutely infected with the murine pathogen Theiler's virus; however, it was observed that the SJL strain relative to the B10.S strain fails to clear the virus after intracerebral inoculation, sustains a chronic infection, and succumbs to demyelinating disease similar to multiple sclerosis (121, 122). In efforts to determine the cause for susceptibility to viral persistence, SJL and B10.S congenic mouse strains were systematically analyzed at various loci including but not limited to MHC haplotypes as

well as at *Ifng* (121, 122). These studies identified that the *Tmevp3* loci located just adjacent to *Ifng* is a region of susceptibility in the context of the demyelinating disease.

At the start of the body of work described herein, *Tmevpg1* the transcribed product of the *Tmevp3* locus was implicated to be a repressor of *Ifng* (121, 122). Predominately expressed in the spleen and thymus of C57BL/6 and B10.S mice, detectable *Tmevpg1* transcript levels were elevated in both unstimulated mouse and human immune cells and upon stimulation inversely correlated with *Ifng* expression (121, 122).

Our work examines the noncoding regulation of *Ifng* particularly in the context of Th1 differentiation, thus I aimed to uncover the functional role of *Tmevpg1* in IFN-γ-producing cells of the adaptive immune system. Chapter II describes *Tmevpg1* expression as it relates to Th1-selective expression of IFN-γ. Further, as described in Chapter III, I establish that together the *Tmevpg1* and *Ifng* genes comprise the Th1 locus maintained by T-bet. Lastly, encouraged by my studies of *Tmevpg1*, in Chapter IV I examined results from an RNA-sequencing experiment to assess the possibility that other lncRNAs are involved in the process of T helper cell lineage commitment. Together, these findings emphasize a crucial role for noncoding regulation of compartmentalized gene networks.

CHAPTER II

Influence of *Tmevpg1* on the expression of *Ifng* by Th1 cells

Overview

   While the majority of the genome is noncoding, the transcriptional control of protein-coding genes resides within these noncoding regions. Long noncoding RNAs, a transcriptional product of noncoding genomic elements, are found in abundance throughout the genome appearing to function as either positive or negative regulators of neighboring genes. For example, both *TMEVPG1* and its mouse orthologue encode lncRNAs and are positioned near the interferon gamma gene. Here we show that transcription of both mouse and human *TMEVPG1* genes is Th1-selective and dependent upon Stat4 and T-bet, transcription factors that drive the Th1 differentiation program. *Ifng* expression is partially restored in $Stat4^{-/-}Tbx21^{-/-}$ cells through co-expression of T-bet and *Tmevpg1* and *Tmevpg1* expression contributes to but is not sufficient to drive Th1-dependent *Ifng* expression. Our results suggest that *TMEVPG1* belongs to the general class of lncRNAs that positively regulate gene transcription.

Materials and Methods

Mice

BALB/cJ.*Stat4*[-/-]*Tbx21*[-/-], DO11.10.*Stat4*[-/-], DO11.10.*Tbx21*[-/-] and wildtype mice were obtained from Christopher L. Williams in the Boothby laboratory. BALB/cJ mice were obtained from the Jackson Laboratory (Bar Harbor, ME). All mice were bred in the Vanderbilt University animal facilities. Research using mice complied with all relevant institutional and federal guidelines and policies.

Human and mouse lymphocyte culture conditions

Healthy human PBMC were isolated by Ficoll–Hypaque density centrifugation (GE Healthcare). $CD4^+$ or $CD8^+$ T cells were sorted by positive selection and stimulated with immobilized anti-CD3 (OKT3; ATCC, Manassas, VA) and soluble anti-CD28 under Th1 (IL-12 10 ng/ml) or Th2 (IL-4 10 ng/ml) polarizing conditions for three days followed by two days of culture with IL-2. Murine BALB/cJ, BALB/cJ.*Stat4*[-/-]*Tbx21*[-/-], DO11.10.*Stat4*[-/-], DO11.10.*Tbx21*[-/-] splenocytes ($1x10^6$ cells/ml) were stimulated with immobilized anti-CD3 (2C11; ATCC, Manassas, VA) or $OVA_{323-339}$ peptide antigen (10 μg/ml) under Th1 (10 ng/ml IL-12 and 10 ug/ml anti–IL-4 11B11; ATCC), Th2 (10 ng/ml IL-4 and 10 μg/ml anti–IFN-γ, R4-642; ATCC), Th17 (10 ng/ml IL-6, 10 ng/ml IL-23, 1 ng/ml TGF$_{β and}$ 10 ug/ml anti-IFN-γ) or Th22 (5 ng/ml IL-1$_β$, 10 ng/ml IL-6 and 5 ng/ml TNF-α) polarizing conditions for three days. $CD4^+$ T cells were purified by positive selection (Miltenyi Biotec). Human or mouse cells were restimulated to generate effector cells by the addition

of 50 ng/ml PMA and 1 μM ionomycin for 6 hours or peptide antigen as described in the *Results* section.

RNA isolation and quantitative RT-PCR

Total RNA was isolated with TriReagent (Ambion, Inc.). cDNA was synthesized with the SSRIII kit (Invitrogen). Transcript levels were determined by SybrGreen quantitative RT-PCR (RT-PCR) using the following primer pairs: human *TMEVPG1* forward 5'aaacgctggaggagaagtca 3' and reverse 5'ttctcctccagcgttttacg 3' and mouse *Tmevpg1* forward 5'cctgaaaatcaccatgcaca 3' and reverse 5'gttttcgggatgtcgtcaaa 3'. Human and murine message levels are expressed as the ratio to *GAPDH* (or *Gapdh*) transcript levels calculated directly from the C_t.

siRNA knockdown

Silencer® Select siRNA duplexes (Ambion, Inc.) were designed against the 5' *Tmevpg1* sequence:

	Sense Strand 5'→ 3'	Antisense Strand 3' → 5'
siRNA 1	GAGAAGAGCCUGAGAGAAA TT	TTCUCUUCUCGGACUCUCU UU
siRNA 2	GCAGACUAAACUAGAUAGU TT	TTCGUCUGAUUUGAUCUAU CA
Ifng siRNA	siRNA ID: 158238	████████████████████
scramble siRNA	CAACUGGGACACAUGUGUU TT	TTGUUGACCCUGUGUACAC AA

siRNA duplexes (30 pmoles) were transfected into cells by Amaxa Nucleofection according to the manufacturer's instructions (Lonza). After four hours of recovery cells were cultured (5 x 10^5 cells/ml) under Th1 conditions for three days before restimulation. IFN-γ was measured in culture supernatants by ELISA (BD OptEIA).

Tmevpg1 sequencing and overexpression

Tmevpg1 cDNA clone AA162222 (Open Biosystems) was sequenced at the Vanderbilt Sequencing Facility. Full length *Tmevpg1* was cloned into the pcDNA3.1/myc-His A overexpression vector (Invitrogen). CMV-*Tmevpg1*, CMV-*Tbx21* or CMV-empty vectors were transfected (1 μg/10^6 cells) into polarized splenocytes using Amaxa Nucleofection (Lonzabio).

Statistical analysis

Statistical significance was determined by Student's T test.

Results

Selective expression of TMEVPG1 and its murine orthologue under Th1 polarizing conditions

Utilizing the UCSC genome browser configuration we identified a gene, AK124066, also named *TMEVPG1*, which is predicted to transcribe a spliced, noncoding mRNA transcript and is positioned approximately 170 kilobases from the *IFNG* coding region (Figure 2-1). The 33 kilobase long *TMEVPG1* gene is comprised of four exons and

encodes an mRNA of 1791 base pairs in length. *TMEVPG1*, similar to *IFNG*, possesses

multiple Th1-specific DNase I hypersensitivity sites at its promoter as well as epigenetic

histone marks, H3K9 acetylation and H3K4 mono- and tri-methylation, that are known to

be associated with active transcription (52, 99). Both *TMEVPG1* and its mouse orthologue

are located on the opposing strand to *IFNG*. The transcriptional start site of mouse *Tmevpg1*

is positioned 117 kilobases from the *Ifng* transcriptional start site and is spliced into a

mature transcript 918 base pairs in length. The promoter region and first intron of *Tmevpg1*

also exhibit considerable sequence conservation with human *TMEVPG1*. These data are

consistent with the possibility that *TMEVPG1* encodes a lncRNA transcript selectively

expressed in Th1 cells relative to Th2 cells.
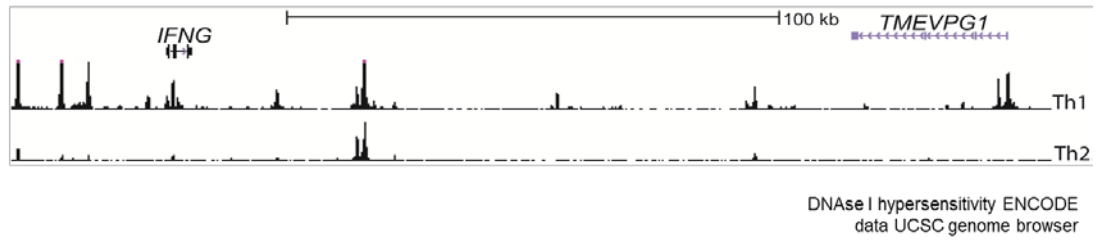
FIGURE 2-1. Relative genomic position of *TMEVPG1* and *IFNG* on human chromosome 12. *TMEVPG1* is located 170 kilobases away from the *IFNG* transcription start site. Below are the Th1, Th2 and Th17 DNase I hypersensitivity sites. Arrowheads indicate the orientation of transcription.

To address this possibility, human CD4$^+$ T cells were stimulated *in vitro* under Th1 or Th2 polarizing conditions for three days, followed by two days of additional culture with IL-2 before restimulation with PMA and ionomycin. Transcript levels of *IFNG*, *TMEVPG1*, and *GAPDH* were determined by RT-PCR. Transcript levels of *IFNG* and *TMEVPG1* were substantially greater in Th1 cultures compared to Th2 cultures (Figure 2-2*A* and 2-2*B*). We also determined transcript levels of *IFNG* and *TMEVPG1* in PBMCs from healthy human control subjects and compared transcript levels by linear regression. We found a positive correlation between *IFNG* and *TMEVPG1* transcript levels relative to *GAPDH* (Figure 2-2*C*). Taken together these results demonstrate that, like *IFNG*, *TMEVPG1* transcript levels are selectively expressed in Th1 cultures relative to Th2 cultures. The linear regression analysis further indicates a strong association between *IFNG* and *TMEVPG1* transcript levels in PBMCs.
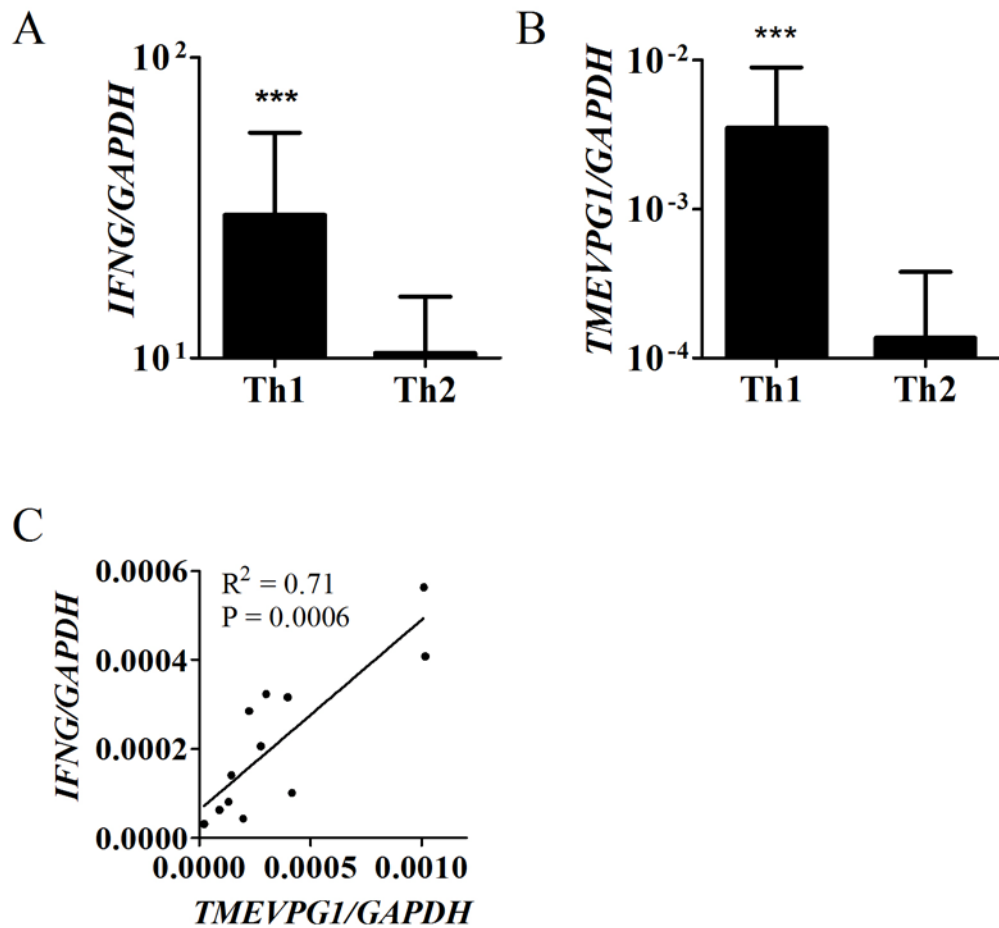
FIGURE 2-2. ***TMEVPG1* is selectively expressed in Th1 cells and positively correlates with *IFNG* expression in human PBMCs.** *A* and *B*, Human CD4[+] T cells (n=12) were cultured under Th1 or Th2 polarizing conditions for three days. *A, IFNG* and *B, TMEVPG1* transcript levels relative to *GAPDH* were determined by SybrGreen RT-PCR. *C*, Relationship between *IFNG* and *TMEVPG1* transcript levels from whole blood samples (PAXgene collection tubes) were determined by linear regression analysis (n=12). Results are expressed as the mean ± the standard deviation of three independent experiments. **p < 0.01 and ***p < 0.001

To assess whether murine *Tmevpg1* was also preferentially expressed in polarized Th1 cells, we measured *Tmevpg1* transcript levels in total splenocytes cultured under Th1 or Th2 polarizing conditions. RNA was isolated after three days in culture at the peak of *Ifng* expression during primary stimulation. *Tmevpg1* and *Ifng* message levels were measured by RT-PCR. Consistent with our results in human lymphocytes, *Tmevpg1* transcript levels were significantly greater in Th1 cells than in Th2 cells (Figure 2-3*A*). *Tmevpg1* transcript levels were also analyzed in Th17 and Th22 polarized cells to confirm Th1 specificity. We also restimulated Th1 effector cells with PMA and ionomycin and followed *Tmevpg1* transcript levels over time. IL-2 was added to these cultures to sustain viability. Restimulation resulted in a marked increase in *Tmevpg1* transcript levels that was sustained over several days (Figure 2-3*B*). In contrast to what we observed in Th1 effector cultures, *Tmevpg1* transcript levels were undetectable in effector CD8$^+$ T cells polarized under Th1 culture conditions, which produced a significant amount of IFN-γ (Figure 2-3*C* and 2-3*D*). We conclude from these experiments that Th1-selective *Tmevpg1* expression is conserved between murine and human lymphocytes but is not expressed by CD8$^+$ T cells under these culture conditions. Further, restimulation of Th1 cells results in greater *Tmevpg1* expression levels than observed in primary Th1 cultures implicating a role for *Tmevpg1* in effector Th1 cells.
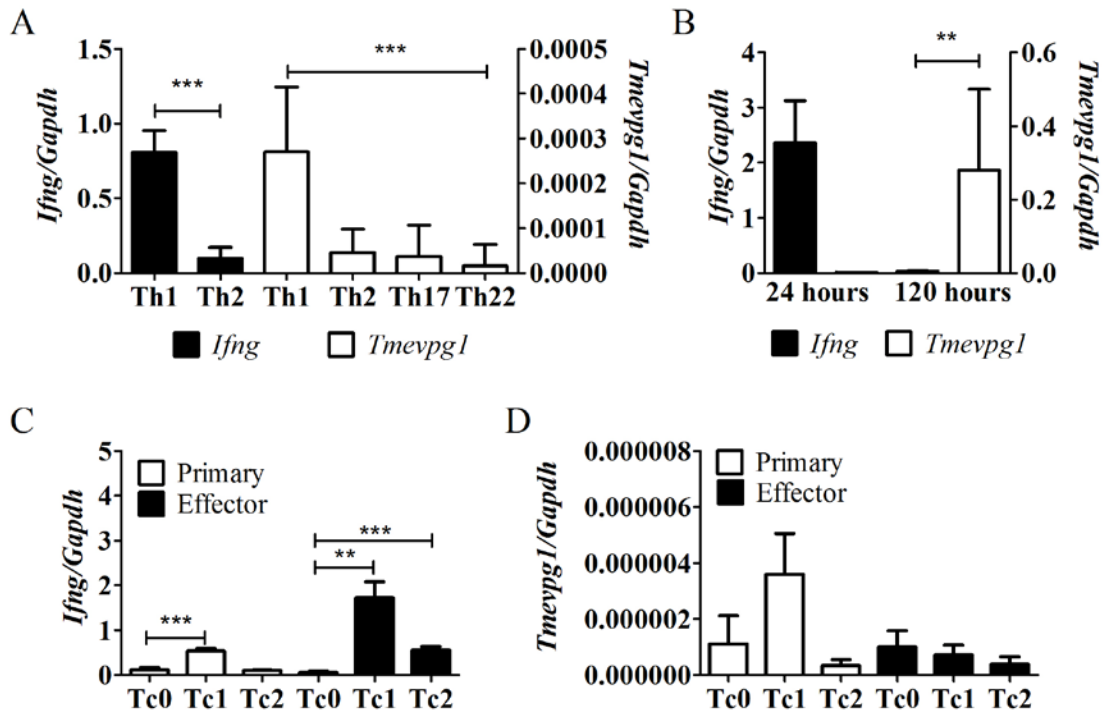
FIGURE 2-3. *Tmevpg1* is selectively expressed in and positively correlates with *IFNG* expression by mouse Th1 cells. *A*, *Ifng* and *Tmevpg1* transcript levels relative to *Gapdh* in Th1, Th2, Th17 or Th22 polarized cultures. *B*, *Ifng* and *Tmevpg1* transcript levels relative to *Gapdh* in CD4[+] effector Th1 cells after restimulation with PMA and ionomycin. Results are expressed as the mean ± standard deviation of three independent experiments. *C, Ifng* and *D, Tmevpg1* transcript levels were measured relative to *Gapdh* in primary and effector CD8[+] Tc0, Tc1 and Tc2 polarized cells by RT-PCR. Results are expressed as the mean and standard error of the mean of at least three independent experiments. **$p < 0.01$ and ***$p < 0.001$

Tmevpg1 is regulated by Th1 transcription factors

Th1-selective expression of IFN-γ in our model system is dependent upon the transcription factors Stat4 and T-bet. Because *Tmevpg1* also displays selective Th1 expression we determined if *Tmevpg1* expression was also dependent upon Stat4 and T-bet. To do so, we isolated splenocytes from DO11.10.*Stat4*[-/-] and DO11.10.*Tbx21*[-/-] (T-bet knockout) transgenic mice and stimulated the cells *in vitro* with OVA$_{323-339}$ peptide and IL-12. After three days, CD4$^+$ T cells were purified and restimulated with OVA$_{323-339}$ peptide for 48 hrs. As expected, *Ifng* transcript levels were substantially diminished in Th1 cells in the absence of Stat4 or T-bet compared to the DO11.10 wildtype control mice (Figure 2-4*A*). *Tmevpg1* transcript levels were also markedly reduced in T cells deficient in either Stat4 or T-bet. Additionally, we examined the expression of *Ifng* and *Tmevpg1* in primary as well as effector polarized *Stat1*[-/-] cells (Figure 2-4*B*). Effector T cells exhibited significant inhibition in the absence of *Stat1*[-/-] whereas little effect was observed for *Ifng* or *Tmevpg1*. These observations support the dependence of *Ifng* and *Tmevpg1* on Th1-specific transcription factors.
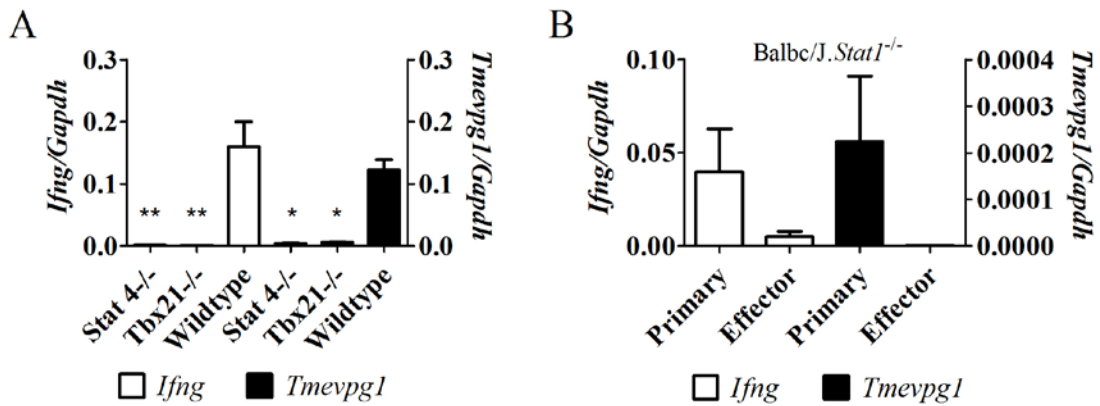
FIGURE 2-4. *Tmevpg1* **induction is dependent upon Stat1, Stat4 and T-bet transcription factors.** *A*, *Ifng* and *Tmevpg1* expression in Th1 cells from DO11.10.*Stat4*⁻/⁻, DO11.10.*Tbx21*⁻/⁻ (T-bet knockout) and wildtype mice relative to *Gapdh*. Results are expressed as the mean ± standard error of the mean. *B*, *Ifng* and *Tmevpg1* expression in Th1 cells from BALB/cJ.*Stat1*⁻/⁻ in primary and effector cultures relative to *Gapdh*. Results are expressed as the mean ± standard deviation of replicates. *p > 0.05 and **p > 0.01

Influence of Tmevpg1 on Ifng transcription

LncRNAs cluster into two functionally distinct categories: repressors and enhancers of transcription of protein-coding genes. We aimed to determine the function of *Tmevpg1* in Th1 cells via siRNA-mediated knockdown of *Tmevpg1*. Nucleofection of primary cells with *Tmevpg1*-specific siRNA duplexes 1 and 2 resulted in a reduction of *Tmevpg1* (Figure 2-5*A*) and *Ifng* (Figure 2-5*B*) transcript levels relative to the scramble siRNA transfected polarized splenocytes. Knockdown of *Tmevpg1* by siRNA duplex 1 or siRNA duplex 2 resulted in a two-fold or four-fold reduction in detectable IFN-γ protein in the culture supernatant, respectively, relative to transfection with a non-specific scrambled siRNA (Figure 2-5*C*). Knockdown with *Ifng* siRNA resulted in a similar decrease in IFN-γ protein concentrations. Nucleofection with both *Tmevpg1* siRNA duplexes 1 and 2 caused a comparable decrease in *Tmevpg1* transcript levels while siRNA knockdown of *Ifng* did not affect transcript levels of *Tmevpg1*. Our conclusion is that *Tmevpg1* plays a role in *Ifng* expression by Th1 cells.
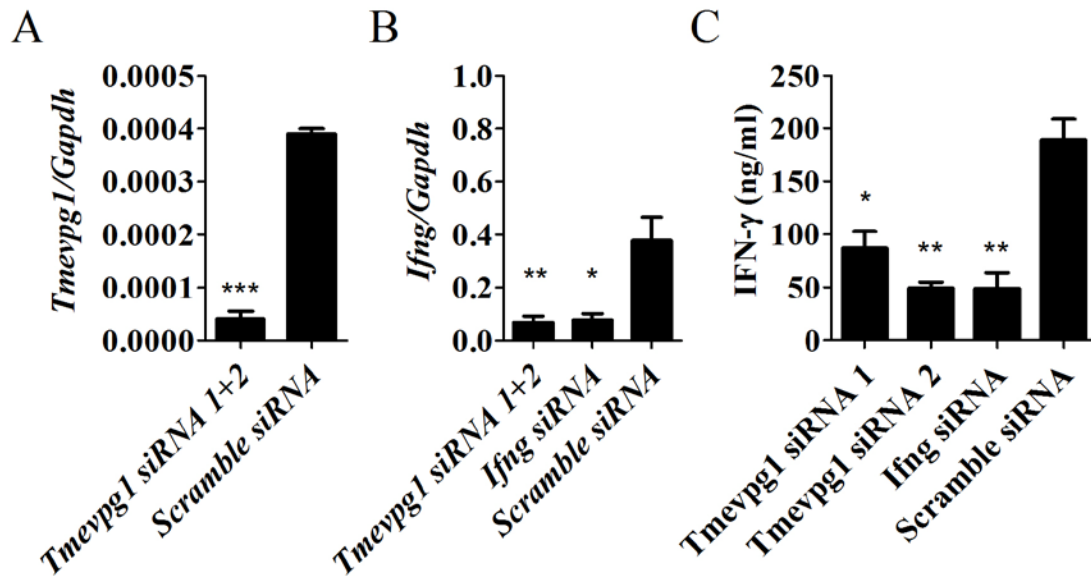
FIGURE 2-5. *Tmevpg1* knockdown impairs *Ifng* expression in CD4[+] T cells. *Tmevpg1*-specific, *Ifng*-specific or scrambled siRNA duplexes were introduced into Th1 polarized CD4[+] T cells by nucleofection. *A, Tmevpg1*, *B*, *Ifng* and *C*, IFN-γ levels (measured by ELISA) were determined after restimulation with PMA and ionomycin. Results are expressed as the mean ± standard deviation of independent experiments. *p > 0.05 and **p > 0.01

Based upon the results from the siRNA knockdown experiments, we determined if overexpression of *Tmevpg1* was sufficient to induce *Ifng* transcription. Full length *Tmevpg1* was cloned into a CMV expression plasmid. Total BALB/cJ splenocytes were stimulated with anti-CD3 under neutral conditions (Th0) or under Th1 or Th2 polarizing conditions for three days. CD4$^+$ T cells were isolated and CMV-*Tmevpg1* or CMV-empty vectors were then transfected (1 μg of plasmid per 10^6 cells). After a period of rest, cells were restimulated with PMA and ionomycin. Nucleofection of CMV-*Tmevpg1* into primary Th1 cells resulted in an increased expression of *Tmevpg1* compared to the empty vector control (Figure 2-6*A*). Ectopic expression of *Tmevpg1* in primary CD4$^+$ T cells resulted in no significant increase in IFN-γ protein in Th0, Th1, or Th2 cells relative to transfection with an empty vector control (Figure 2-6*B*).
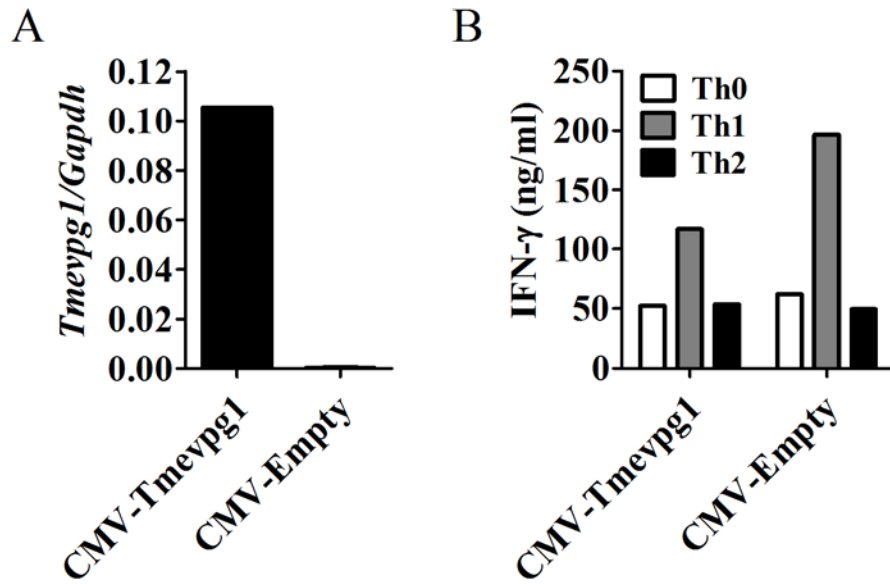
FIGURE 2-6. *Tmevpg1* **is necessary but not sufficient for IFN-γ expression by Th1 cells.** CMV-*Tmevpg1* or CMV-empty expression plasmids were introduced into polarized cultures by nucleofection and cultures were restimulated with PMA and ionomycin. *A, Tmevpg1* transcript levels are expressed relative to *Gapdh. B,* IFN-γ levels in Th0, Th1 and Th2 polarized cultures were determined by ELISA. Results represent the mean of at least three independent experiments.

We further examined whether *Tmevpg1* was sufficient to restore *Ifng* transcript expression in the absence of Stat4 and T-bet. CMV-*Tmevpg1*, CMV-*Tbx21* or CMV-empty vector plasmids were transfected into Th1 polarized BALB/cJ.*Stat4*[-/-]*Tbx21* [-/-] splenocytes. After a period of rest, cultures were restimulated with PMA and ionomycin. *Tmevpg1*, *Tbx21* and *Ifng* transcript levels were determined by RT-PCR. We observed *Tmevpg1* expression to be restored by ectopic expression of T-bet alone in the BALB/cJ.*Stat4*[-/-] *Tbx21* [-/-] cells (Figure 2-7*A*). CMV-*Tmevpg1* and CMV-*Tbx21* co-transfection resulted in a substantial increase in *Ifng* transcript levels relative to transfection of CMV-*Tmevpg1* or CMV-*Tbx21* alone (Figure 2-7*C*). We conclude from these experiments that overexpression of *Tmevpg1 in trans* alone is not sufficient to induce increased *Ifng* transcription in Th0, Th1, or Th2 cells or in *Stat4*[-/-]*Tbx21* [-/-] cells except in the presence of T-bet.
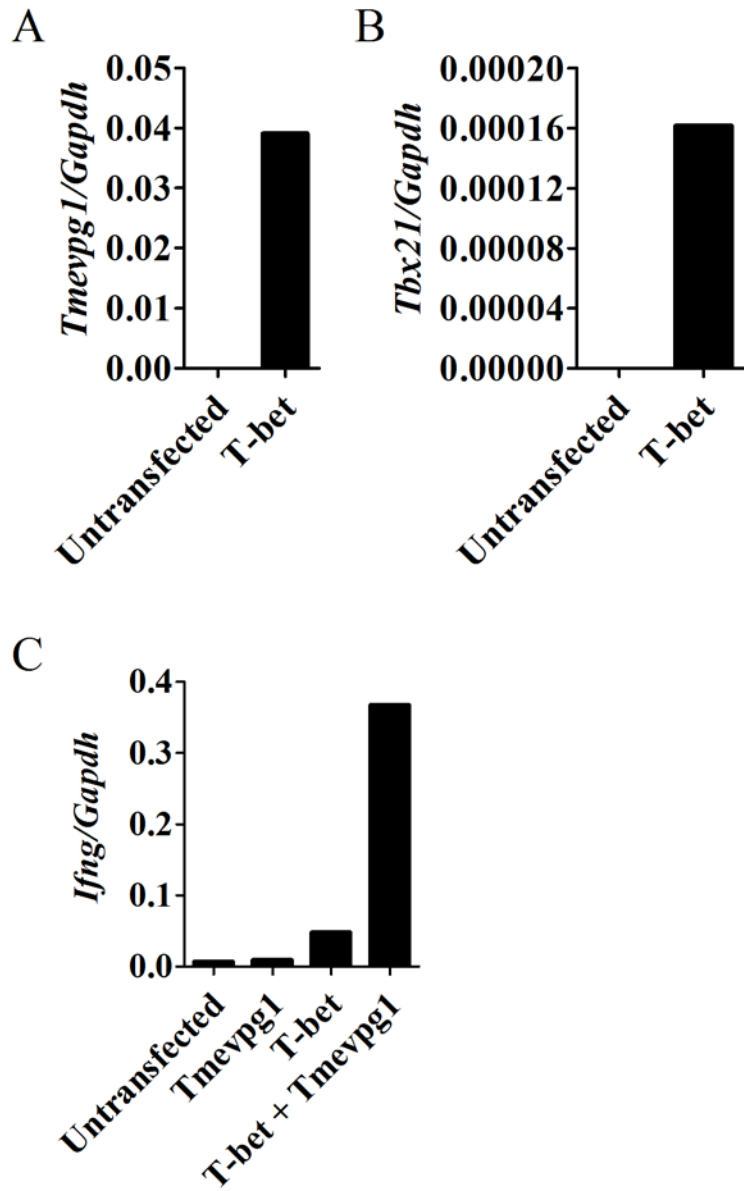
FIGURE 2-7. Cooperative action of T-bet and *Tmevpg1* restores *Ifng* expression. CMV-*Tmevpg1*, CMV-*Tbx21 and/*or CMV-empty vector plasmids were transfected into Th1 polarized BALB/cJ.*Stat4*$^{-/-}$*Tbx21*$^{-/-}$ cells. Cells were restimulated with PMA and ionomycin. *A, Tmevpg1* B, *Tbx21* and *C, Ifng* transcript levels are expressed relative to *Gapdh*. Results represent the mean of at least three independent experiments.

Discussion

Presently, lncRNAs segregate into two functional categories that either repress or enhance the transcription of protein-coding genes. To summarize our results, *Tmevpg1* and its human orthologue, are expressed selectively in Th1 cells relative to Th2, Th17 and Th22 cells and expression is dependent upon the Th1 specific transcription factors, Stat4 and T-bet. Our results also demonstrate that *Tmevpg1* influences *Ifng* transcription in response to the Th1 differentiation program. In contrast, ectopic expression of *Tmevpg1* does not increase *Ifng* transcript levels in Th0, Th1, or Th2 cells; however, *Tmevpg1* is able to partially restore *Ifng* expression when T-bet is also overexpressed. One possible interpretation is that *Tmevpg1* must be expressed from its endogenous locus, or in *cis*, to stimulate *Ifng* transcription. A second possible interpretation, which our data favor, is that *Tmevpg1* must act in concert with T-bet, or other critical trans-activating factors, to influence *Ifng* transcription (Figure 2-8). Other studies of enhancer lncRNAs are consistent with our results as these lncRNAs also fail to stimulate transcription of protein-coding genes *in trans* or require additional transactivation factors to drive their transcription.

A general emerging model is that cell-type specific transcription factors bind to lncRNA promoters to drive their transcription. The lncRNAs bind to ubiquitous proteins required to establish the epigenetic code and by mechanisms that are incompletely understood direct these proteins to their target protein-coding genes. This model does not rule out the possibility that these cell-type specific transcription factors also target protein-coding genes. Our results demonstrate that one lncRNA, *Tmevpg1*, contributes to *Ifng*

expression as part of the Th1 differentiation program. We predict that additional lncRNAs play critical roles in developmental programs required to establish the different functions of the immune system.
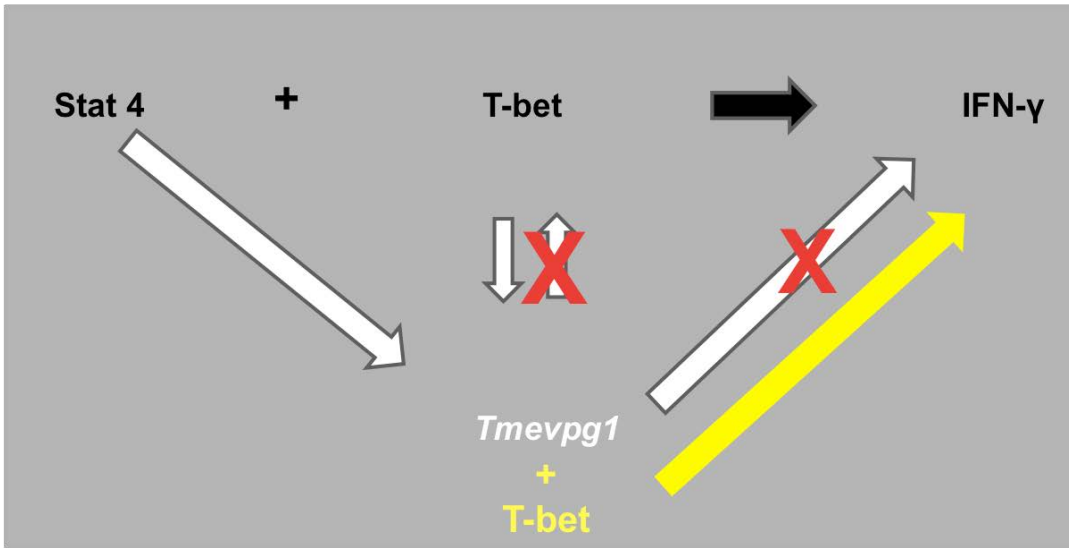
FIGURE 2-8. Summary of *Tmevpg1* mediated effects on *Ifng* expression. Stat 4 and T-bet are known transcriptional regulators involved in production of IFN-γ (black arrows). Here I demonstrate that both Stat4 and T-bet are necessary for *Tmevpg1* expression (white arrows). While *Tmevpg1* alone is not sufficient to induce T-bet or IFN-γ (arrows with red X's), co-expression of *Tmevpg1* and T-bet together is capable to partially restore IFN-γ (shown in yellow).

CHAPTER III

T-bet regulates the Th1 locus from *Ifng* through *Tmevpg1*

Overview

At the transcriptional level *IFNG* gene regulation, specifically in CD4$^+$ Th1 cells, is dependent upon the transcription factors Stat4 and T-bet for expression. Previously, the limits of the mouse *Ifng* gene mapped to +66 downstream marked by an insulator element (**CCCTC-binding factor sites or CTCF site**) (37). This insulator element is located near the mapped 3' end of the *Tmevpg1* gene encoding a Th1-specific lncRNA transcript involved in the regulation of *Ifng*. Whereas substantial interest in lineage-specific gene regulation and consequently expression has focused on protein-coding genes, little is known about how lncRNA genes are regulated. Here I show that T-bet-dependent regulation of the *Tmevpg1* gene is due to epigenetic remodeling of the locus in developing and effector-like Th1 cells, favoring expression. Further, assessment of surrounding distal noncoding sequences identifies four sites capable of enhancer activity which actively recruit inducible transcriptional regulators, such as NFκB and Ets-1. These transcription factors exhibit T-bet dependent recruitment and are partially involved in *Tmevpg1* expression. These findings expand our current understanding of the governing power of T-bet across what we now define as the Th1 locus.

Materials and Methods

Mice

BALB/cJ and DO11.10 mice were obtained from The Jackson Laboratory (Bar Harbor, ME). All mice were bred in the Vanderbilt University animal facilities. Research using mice complied with all relevant institutional and federal guidelines and policies. Human *IFNG* BAC transgenic mice were created on the C57BL/6 genetic background previously (114).

Cultures

Primary BALB/cJ or DO11.10 splenocytes (1×10^6 cells per ml) were stimulated with immobilized anti-CD3 (2C11; American Type Culture Collection) or OVA$_{323-339}$ peptide antigen (10 μg/ml; InvivoGen) respectively, under neutral Th0 (10 ug/ml anti– IFN-γ and 10 μg/ml anti–IL-4), Th1 (10 ug/ml anti–IL-4, 11B11; American Type Culture Collection) or Th2 (10 ng/ml IL-4 and 10 ug/ml anti– IFN-γ, R4-642; American Type Culture Collection) polarizing conditions for 3 days generating primary cultures. Effector cultures were generated by restimulation for an additional 48 hours on immobilized anti-CD3 coated plates. Jurkat T lymphocytes (American Type Culture Collection) were maintained in complete RPMI at a density of 5 x10^5 cells/ml. CD4$^+$ T cells and NK cells were purified by negative magnetic selection (Miltenyi Biotec).

RNA isolation and RT-PCR

Total RNA was isolated with TRI Reagent (Ambion). cDNA was synthesized with the SSRIII kit (Invitrogen). *Tmevpg1* (Life Technologies TaqMan assay ID: Mm01161206_m1) and *Ifng* (TaqMan, Life Technologies) transcript levels were measured by RT-PCR and calculated relative to *Gapdh* (TaqMan, Life Technologies) by the delta-delta Ct method.

Chromatin immunoprecipitation (ChIP)

ChIP procedures were followed as described (114) using the following antibodies: anti-H4Ac (Millipore), Ets-1 (Santa Cruz), NF-κB p65 (Abcam), and anti-rabbit IgG (Sigma). Primers used in amplifying chromatin across the *Tmevpg1* locus are listed in Supplementary Table 3-1. Conserved sequences for HS sites were amplified by using the primers listed in Supplemental Table 3-2.

Cell Transfection and Luciferase Assays

Human HS elements were amplified with the primers listed in Supplemental Table 3-3 from Jurkat T cell genomic DNA and cloned into the minimal promoter luciferase (pGL4.24; Promega) or the promoter-less luciferase construct (pGL4.10; Promega). Jurkat T cells were transfected with the luciferase constructs by the DEAE transfection method (123) at 1 ug of plasmid per 10^6 cells. After overnight recovery, cells were stimulated with 50 ng/ml PMA and 1 μM ionomycin per for 6 hours before luciferase activity was measured with the luciferase activity system (Promega) as described by the manufacturer. Cells were

treated with BAY11-7085 NF-κB inhibitor (Sigma) for one hour at 5 uM concentration per milliliter of culture. Promoter truncations were generated with the primers listed in Supplemental Table 3-4 and cloned into the pGL4.10 construct.

Statistical analysis

Statistical significance was determined by a Student's *t* test.

Results

IFN-γ producing cells share Tmevpg1 expression

Like effector Th1 cells, NK and NKT cells also rapidly express *Ifng* in response to extracellular stimuli (Figure 3-1*A*). Thus, we asked if NK cells also endogenously express *Tmevpg1*. We found that freshly isolated mouse NK cells without stimulation also express high levels of *Tmevpg1* that were similar in magnitude to effector Th1 cells (Figure 3-1*B*). *In vivo*, effector-memory T cells (also referred to as polyclonal memory CD4$^+$ cells) that express *Ifng* accumulate with age. In 4-5 week old mice, only ~5% of T cells are of the effector/memory phenotype while in 14-18 week old mice, about 90% of T cells are of the effector/memory phenotype (124). We verified that CD4$^+$ T cells from 14-18 week old mice express significantly higher endogenous levels of *Ifng* than CD4$^+$ T cells from 4-5 week old mice (Figure 3-1*C*). We also found that CD4$^+$ T cells from 14-18 week old mice express significantly higher endogenous levels of *Tmevpg1* than CD4$^+$ T cells 4-5 week old mice (Figure 3-1*D*). Taken together, from both the NK cell observations and survey of *Tmevpg1* expression in polyclonal memory cells, we speculate that terminally

differentiated populations of primed yet unstimulated effector-like NK and polyclonal memory cells are capable of rapid IFN-γ production because of the elevated levels of endogenous *Tmevpg1* transcripts.
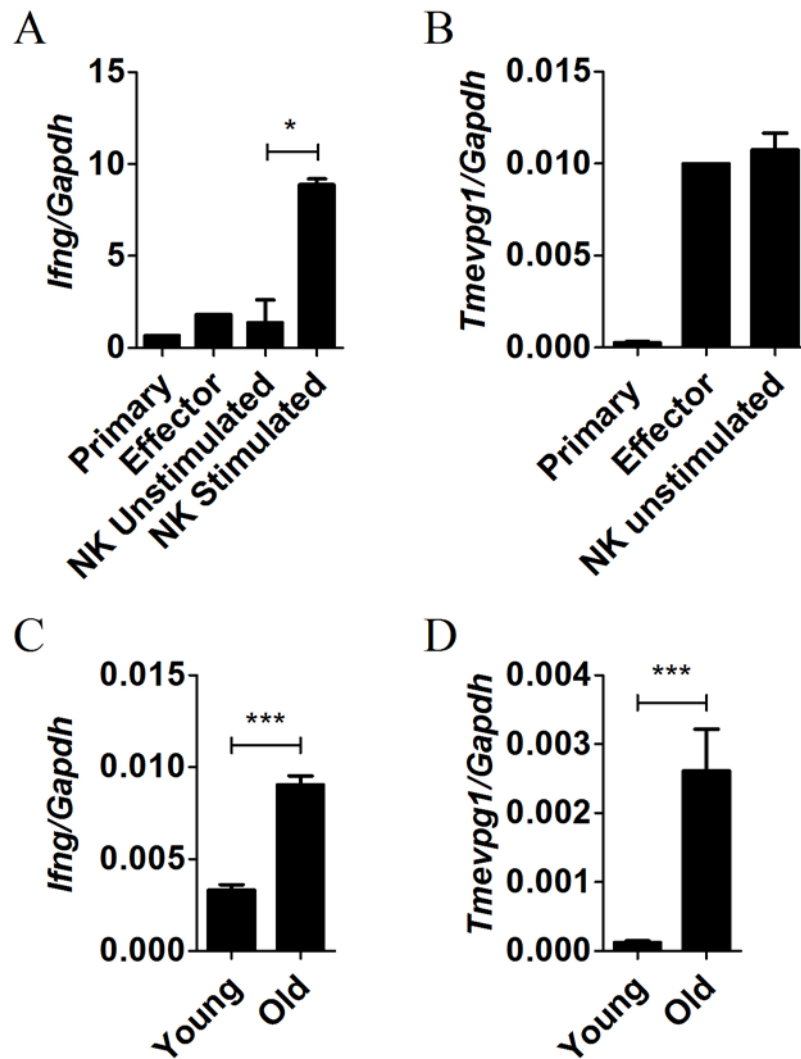
FIGURE 3-1. Differential expression of *Tmevpg1* in Th1 cells, CD4⁺ T cells and NK cells. *A, Ifng* and *B, Tmevpg1* transcript levels were measured relative to *Gapdh* in cultured Th1 cells (levels shown for reference) and purified NK cells. *C, Ifng* and *D, Tmevpg1* transcript levels were measured in purified resting CD4⁺ T cells from young (4-6 weeks old) and old (14- 18 weeks old) animals. Results are expressed as the mean ± standard error of the mean of at least three independent experiments. * p > 0.05 and *** p > 0.001

T-bet physically associates and regulates epigenetic modifications at the Tmevpg1 locus

Th1 polarization promotes the expression of IFN-γ while suppressing Th2-specific genes, a process largely orchestrated by noncoding regulatory elements and transcription factor recruitment such as with T-bet (112). Through ChIP analysis, T-bet is found to associate with the *Ifng* promoter as well as numerous sites along the 140 kilobase intergenic distance between *Ifng* and *Tmevpg1*. Further, T-bet is required for the formation, spreading, and maintenance of H4Ac histone modifications in this region, promoting an open chromatin conformation favoring transcription (112, 116). As *Tmevpg1* expression is also dependent upon and induced by T-bet, we investigated T-bet association with the *Tmevpg1* locus.

Splenic cultures were polarized under Th1 conditions for three days to generate primary cultures and for five days with an additional 48 hours of stimulation with immobilized anti-CD3 to generate effector Th1 cultures. T-bet binding was assessed within +3.0 kilobases upstream relative to the *Tmevpg1* transcription start site through -160 base pairs downstream by ChIP followed by RT-PCR. T-bet was found to associate with chromatin mapping to the *Tmevpg1* locus in primary Th1 cells as well as in effector Th1 cells relative to isotype control immunoprecipitations (Figure 3-2*A*). T-bet was similarly found to associate with the *Ifng* promoter (Figure 3-2*B*) under both stimulation conditions; however, was markedly enriched in effector Th1 cultures. Moreover, the majority of T-bet association in the primary Th1 cultures focused within ± 200 base pairs of the *Tmevpg1* transcriptional start site whereas T-bet association was found to spread for upwards of +3.0 kilobases across the locus in effector cultures. Thus, T-bet binds to the

Tmevpg1 promoter in primary Th1 cells and spreads across up to +3 kilobases of the

Tmevpg1 region as cells further differentiate into effector Th1 cells.
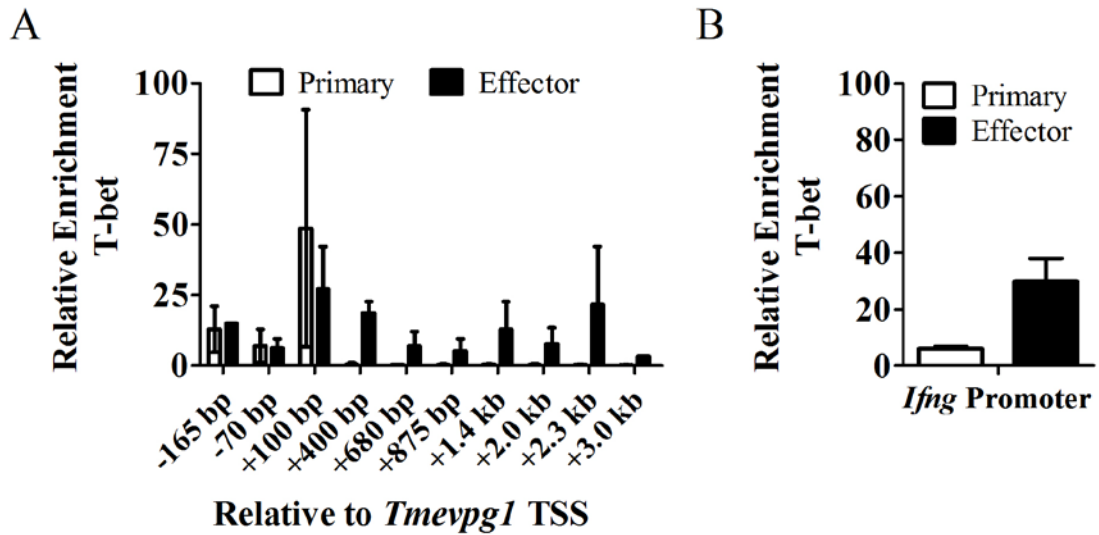
FIGURE 3-2. T-bet associates with the *Tmevpg1* locus in Th1 primary and effector cell cultures. T-bet enrichment across the *A, Tmevpg1* locus and at the *B, Ifng* promoter was measure in primary and effector Th1 cultures by ChIP assays followed by RT-PCR. Results are expressed as the mean ± standard error of the mean from three independent experiments.

Given the above results, we next examined T-bet-dependent H4Ac patterns at the *Tmevpg1* gene, a histone modification associated with permissive transcription and formation of H4Ac marks at *Ifng* and various distal *Ifng* enhancers is known to be dependent, in part upon T-bet. Spleen cell cultures from DO11.10 wildtype or DO11.10.*Tbx21*[-/-] mice were polarized under Th1 conditions for three days to generate primary Th1 cultures and five days with 48 hours additional stimulation with immobilized anti-CD3 to generate effector Th1 cultures. H4Ac histone modifications were assessed by ChIP spanning +3 kilobases upstream to -165 base pairs downstream of the *Tmevpg1* transcriptional start site. Chromatin was enriched and assayed for abundance of H4Ac modifications by RT-PCR relative to immunoprecipitations with an isotype control antibody. In primary Th1 cultures, H4Ac modifications were enriched at genomic regions surrounding the *Tmevpg1* promoter and these modifications required the presence of T-bet (Figure 3-3*A*). H4Ac modifications were also enriched in effector Th1 cells and these were also dependent upon T-bet (Figure 3-3*B*). Further, the level of H4Ac modifications was higher in effector Th1 cultures than primary Th1 cultures and H4Ac modifications appeared to spread across the ~3 kilobase genomic region. These findings suggest that T-bet associates with sites along the *Tmevpg1* locus and is necessary for the seeding and spreading of the H4Ac marks across the *Tmevpg1* genomic region.
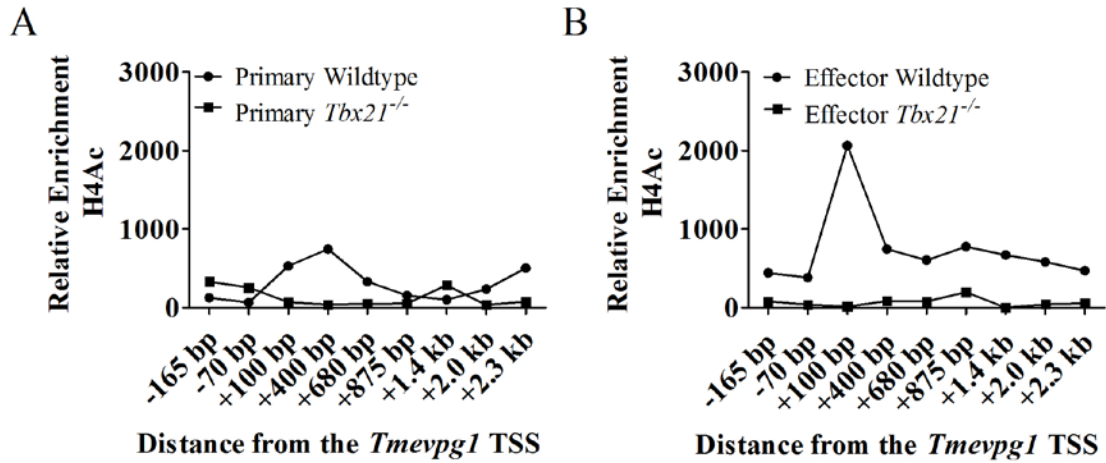
FIGURE 3-3. Seeding and spreading of H4Ac marks across the *Tmevpg1* locus is dependent upon T-bet in primary and effector Th1 cells. H4Ac measurements were assayed by ChIP in *A,* primary and *B*, effector Th1 cultures. Results are expressed as the mean fold enrichment over the isotype control immunoprecipitations in three independent experiments.

TMEVPG1 is surrounded by functional enhancer elements dispensable for IFNG

The *IFNG* locus, including all known *cis* regulatory elements, was previously defined by two insulator sequences located at -63 kilobases and +119 kilobases in humans and -70 kilobases to +66 kilobases in mice (37, 125). Studies utilizing a human BAC transgenic mouse model where mice express the human *IFNG* gene in the context of a 190-210 kilobases human locus demonstrate that 190 kilobases of surrounding noncoding elements are required for Th1-selective expression of *IFNG* whereas a smaller transgene derived from a plasmid containing the intact *IFNG* gene and approximately 3 kilobases of upstream and downstream of genomic sequence was unable to repress *IFNG* expression by Th2 cells (114, 126). Such observations indicated that proper regulation of *IFNG* by developing Th1/Th2 cells was dependent on noncoding regions within the locus.

We identified five, Th1-specific DNase I hypersensitivity sites (HS) between *IFNG* and *TMEVPG1* (Figure 3-4*A*, adapted from the UCSC Genome Browser) indicating that these areas represent accessible, open chromatin. Using an unbiased deletion strategy to identify noncoding regions conferring Th1-specific *IFNG* expression, we examined the requirement of two large deletions of the 190 kilobase BAC transgene mapping to the region between *IFNG* and *TMEVPG1*. CD4$^+$ T cells were purified and cultured under neutral (Th0), Th1 and Th2 conditions for five days to generate primary Th1 cells or subsequently restimulated for 48 hours with immobilized anti-CD3 to generate Th1 effector cells. Human *IFNG* mRNA was measured relative to an endogenous control by RT-PCR under both culture conditions. Compared to the full BAC transgene, deletion 1

and deletion 2 cultures exhibited equivalent abundance of *IFNG* message relative to *Gapdh* in Th0, Th1 and Th2 primary and effector cultures (Figure 3-4*B* and 3-4*C*) coinciding with levels of IFN-γ in the culture supernatants (114). These results support the notion that this intergenic region is dispensable for *IFNG* expression. Therefore, if these HS sites possess transcriptional enhancer activity, it is likely that they target *Tmevpg1* rather than *Ifng*.
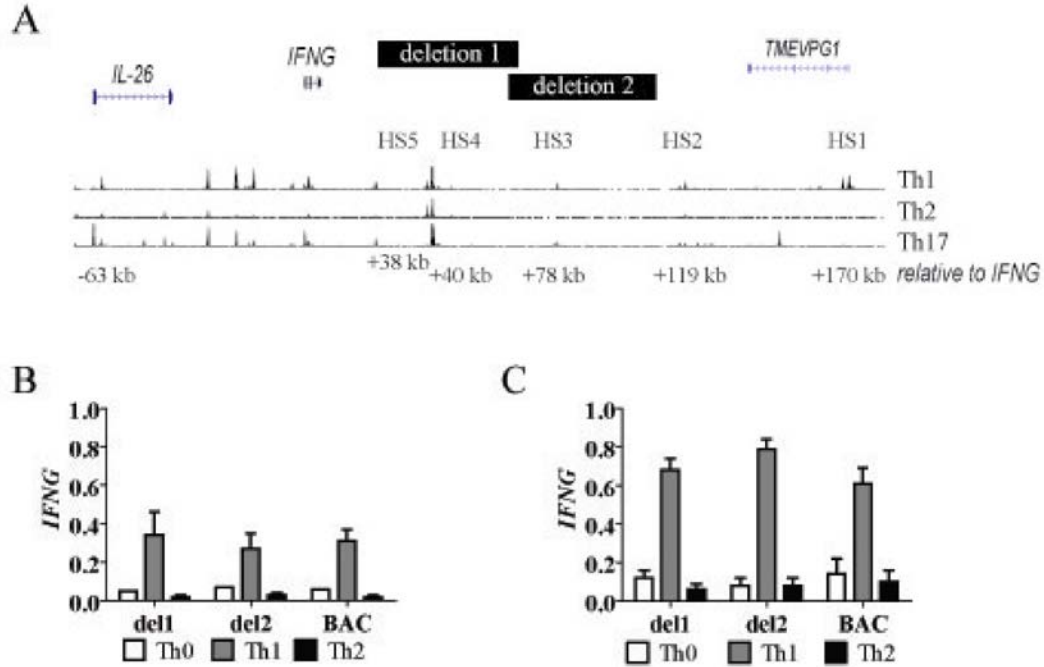
FIGURE 3-4. Enhancers in the intergenic region between *IFNG* and *TMEVPG1* are not required for *IFNG* expression. *A,* Human Th1 locus on chromosome 12 from the UCSC genome browser build. Peaks represent DNase I hypersensitivity in polarized human Th1, Th2 and Th17 cultures. Locations of unbiased BAC deletions (deletion 1 and deletion 2) used to generate human *IFNG* transgenic mice are shown in the intergenic region. Insulator regions at -66 kilobases and +119 kilobases are shown. Distances for each HS site are shown relative to *IFNG*. *B,* Human *IFNG* expression by primary and *C,* effector unpolarized (Th0) Th1 and Th2 murine cells from BAC deletion 1, deletion 2 or the full human 190 kilobases BAC transgene. Results are expressed as the mean ± the standard deviation of *IFNG* message expression relative to *Gapdh*.

60

To explore this hypothesis, we assessed enhancer activity of each HS in the context of *TMEVPG1*. Each HS, numbered one through five, was cloned into the pGL4.24 vector construct, containing a minimal promoter followed by the luciferase gene. The constructs were transfected into Jurkat T lymphocytes at 1 µg per million cells. After overnight recovery, the cells were stimulated with 50 ng/ml PMA and 1 µM ionomycin for six hours and were subsequently assayed for luciferase activity. Transfection with HS1-, HS2-, HS3- or HS4-pGL4.24 constructs resulted in a significant increase in luciferase activity relative to vector alone indicating that these DNA elements possess enhancer activity (Figure 3-5A). Similar levels of enhancer activity were observed in HS1 and HS3 while HS4 had a markedly higher level of enhancer activity relative to vector alone. In contrast, HS5 exhibited a 10-fold decrease in luciferase activity relative to vector alone (Figure 3-5B). These findings indicated that HS1, HS2, HS3 and HS4 were functional enhancer elements while HS5 appeared to have suppressive function.

In addition to having enhancer activity, the HS1 element aligns with the *TMEVPG1* transcriptional start site, thus we also evaluated HS1 for promoter activity. To do so, the HS1 sequence was cloned into the pGL4.10 promoter-less vector construct followed by the luciferase gene and was subsequently transfected into Jurkat T lymphocytes at 1 ug per million cells. After overnight recovery, cells were stimulated with 50 ng/ml PMA and 1 uM ionomycin for 6 hours and assayed for luciferase activity. A three-fold increase in luciferase expression was observed relative the vector control supporting the fact that in addition to enhancer activity, HS1 also contains the *Tmevpg1* promoter (Figure 3-5C).

The HS1 element contains two distinct peaks of Th1-specific DNase I hypersensitivity, either of which could contribute to enhancer as well as promoter activity. Therefore, we employed a traditional promoter analysis truncation approach to identify regions of HS1 that contributed to the observed promoter activity. The HS1 region was divided into four segments, cloned into the pGL4.10 vector construct and assayed for promoter activity relative to the full HS1 construct alone (Figure 3-5*D*). Compared to transfection of vector alone, the larger region spanning either 1-515 base pairs or 263-774 base pairs retained similar promoter activity as HS1; however, deletion of the first 250 base pairs in 263-774 base pairs resulted in a marked decrease in luciferase activity. Additionally, removal of the middle segment leaving only the 1-270 base pair segment resulted in an increase in luciferase activity supporting that this is a region for suppressive control. Further, sequence analysis of the HS1 element identified a TATA sequence motif as well as canonical Ets-1 and NF-κB (RelA) transcription factor binding sites (Figure 3-6*E*). Ets-1 is a transcription factor predicted to function alongside T-bet in contributing to the polarization of CD4$^+$ T cells as mice lacking Ets-1 are impaired in proper Th1 polarization demonstrating the importance for Ets-1 in this process (127-129). These results indicate that *TMEVPG1* is surrounded by functional enhancer sequences while HS1 also possesses promoter activity.
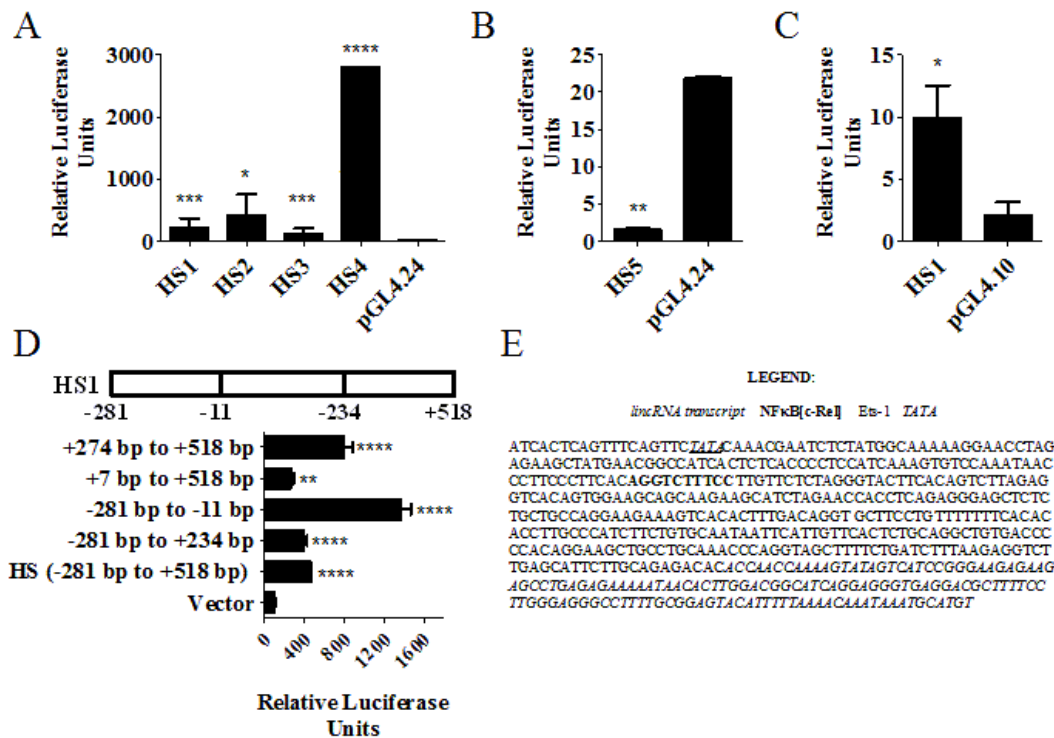
FIGURE 3-5. Survey of the human Th1 locus from *IFNG* to *TMEVPG1*. *A and B,* Human HS elements were cloned into the pGL4.24-luciferase construct under the control of a generic minimal promoter. Constructs were transfected into Jurkat T cells via DEAE transfection. Results are sown as the mean ± the standard error of the mean of two independent transfections. *C,* HS1 promoter activity was evaluated by cloning into the promoter-less pGL4.10 construct in Jurkat T cells. Results are expressed as the mean of three independent transfections ± the standard error of the mean. *D,* Truncations of the HS1 promoter to identify regions contributing to promoter activity was evaluated by cloning fragments into the pGL4.10 construct in Jurkat T cells. Results are expressed as the mean of three independent transfections ± the standard error of the mean. *E,* HS1 sequence analysis identifies binding sites for NF-κB and Ets-1. * p > 0.05, ** p > 0.01, *** p > 0.001 and **** p > 0.0001

Whereas IFN-γ production by primary and effector Th1 cells is dependent upon T-bet, it is known that T-bet is dispensable for *IFNG* expression by memory cells (124). Inducible transcription factors such as NF-☐B, in particular, contribute to the rapid IFN-γ response in effector-memory Th1 cells (86). As such, we examined NF-κB and Ets-1 recruitment to the *Tmevpg1* locus. Spleen cell cultures from DO11.10 wildtype mice were polarized under Th1 conditions for three days for primary cultures and five days with 48 hours additional stimulation with immobilized anti-CD3 to generate effector Th1 cultures. NF-☐B and Ets-1 recruitment to *Tmevpg1* was assessed by ChIP assay spanning +2.3 kilobases upstream relative to the *Tmevpg1* transcriptional start site to -165 base pairs downstream of the *Tmevpg1* transcriptional start site. NF-κB recruitment to *Tmvepg1* was relatively low across the locus particularly in primary Th1 cultures (Figure 3-6*A*). However, a distinctive peak of enrichment was detected at -70 base pairs downstream of the *Tmevpg1* transcription start site in effector Th1 cultures. Conversely, in these same samples, Ets-1 was recruited across the locus in both primary and effector Th1 cultures (Figure 3-6*B*). Thus, both NF-κB and Ets-1, along with T-bet, are recruited to the *Tmevpg1* locus in Th1 cells.
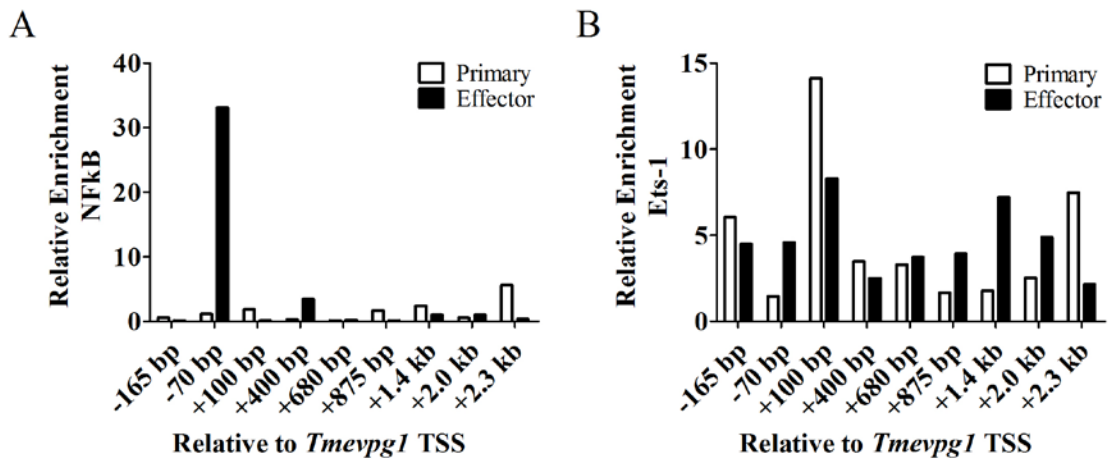
FIGURE 3-6. NF-κB and Ets-1 associate with the *Tmevpg1* locus in primary and effector Th1 cells. *A,* NF-κB and *B*, Ets-1 measurements were assayed by ChIP in primary and effector Th1 cultures. Results are expressed as the mean fold enrichment over the isotype control immunoprecipitations ± the standard error of the mean from three independent experiments.

As both NF-κB and Ets-1 associated with *Tmevpg1* genomic promoter/enhancers, we examined recruitment of these transcription factors to murine HS enhancer elements that were identified by sequence conservation to the above human HS elements (the HS2 element is not conserved between humans and mice). To address this question, spleen cell cultures from DO11.10 wildtype or DO11.10.*Tbx21*$^{-/-}$ mice were polarized under Th1 conditions for three days for primary cultures and five days with 48 hours addition stimulation with immobilized anti-CD3 to generate effector cultures. We found that NF-κB and Ets-1 were enriched at these conserved HS regions and recruitment was largely T-bet dependent (Figure 3-7*A* and 3-7*B*). The magnitude of recruitment to mHS3 and mHS4 was greater than the recruitment to the *Ifng* promoter (compare Figure 3-7*A* & 3-7*B* to Figure 3-7*C*). Ets-1 was also recruited to each of the mHS enhancer sequences in effector cultures and was found to be largely, but not absolutely dependent upon the presence of T-bet (Figure 3-7*D* & Figure 3-7*E*), similar to the pattern observed for NF-κB enrichment. Lastly, Ets-1 associated with the *Ifng* promoter in a T-bet dependent manner in effector cultures (Figure 3-7*F*). These results support that T-bet is required, in part, for inducible transcription factor Ets-1 and NF-κB recruitment to this intergenic enhancer sequences.
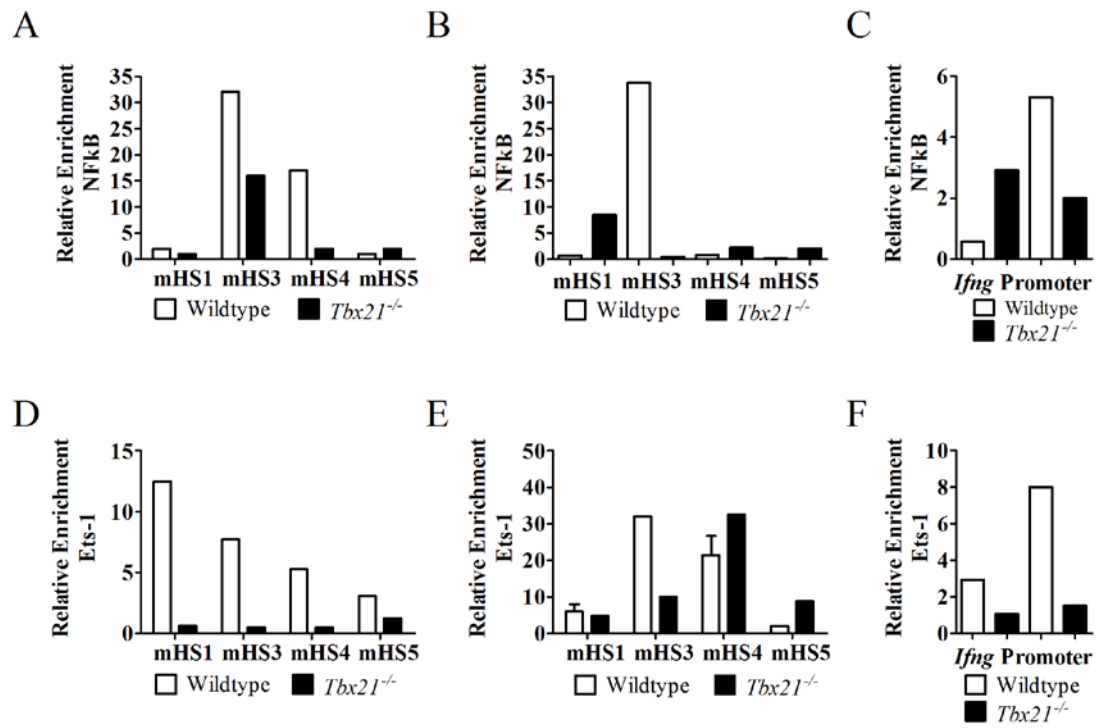
FIGURE 3-7. NF-κB and Ets-1 are enriched at *Tmevpg1* enhancer sequences in a T-bet dependent manner. *A* and *B*, NF-κB and *D* and *E*, Ets-1 transcription factors bind to homologous mouse *Tmevpg1* enhancer sequences in primary (*A* and *D*), effector (*B* and *E*) cultures and at the *Ifng* promoter under both culture conditions (*C* and *E*) in DO11.10 and DO11.10.*Tbx21*⁻/⁻ Th1 cells. Cells were assayed by ChIP for transcription factor binding after 48 hours of restimulation with immobilized anti-CD3. Enhancer primers are designated on the x-axis. Results are expressed as the representative mean fold enrichment and error bars represent the standard error of the mean over the isotype control for three independent cultures.

Additionally, we assessed pharmacological inhibition of NF-κB activation by treatment with the extensively utilized I-kappa alpha inhibitor BAY11-7085. The HS1-, HS2-, HS3-, HS4- and HS5-pGL4.24 constructs were transfected into Jurkat T lymphocytes at 1 μg per million cells (130). After overnight recovery, cells were incubated in the presence of the inhibitor for one hour followed by stimulation with 50 ng/ml PMA and 1 μM ionomycin for six hours before determining luciferase activity. Relative to control cultures, cultures treated with NF-κB inhibitors exhibited a marked decrease in activity of each enhancer: HS1-HS4, to varying degrees (Figure 3-8*A*). Conversely, inhibition of NF-κB activity appeared to reverse the repressor activity of mHS5 (Figure 3-8*B*). Further, treatment of effector cultures with BAY11-7085 before restimulation resulted in impaired *Ifng* as well as *Tmevpg1* expression relative to control cultures. These results support an important role for NF-κB activation and recruitment to HS1-HS5 elements to achieve stimulus-dependent transcriptional enhancer activity.
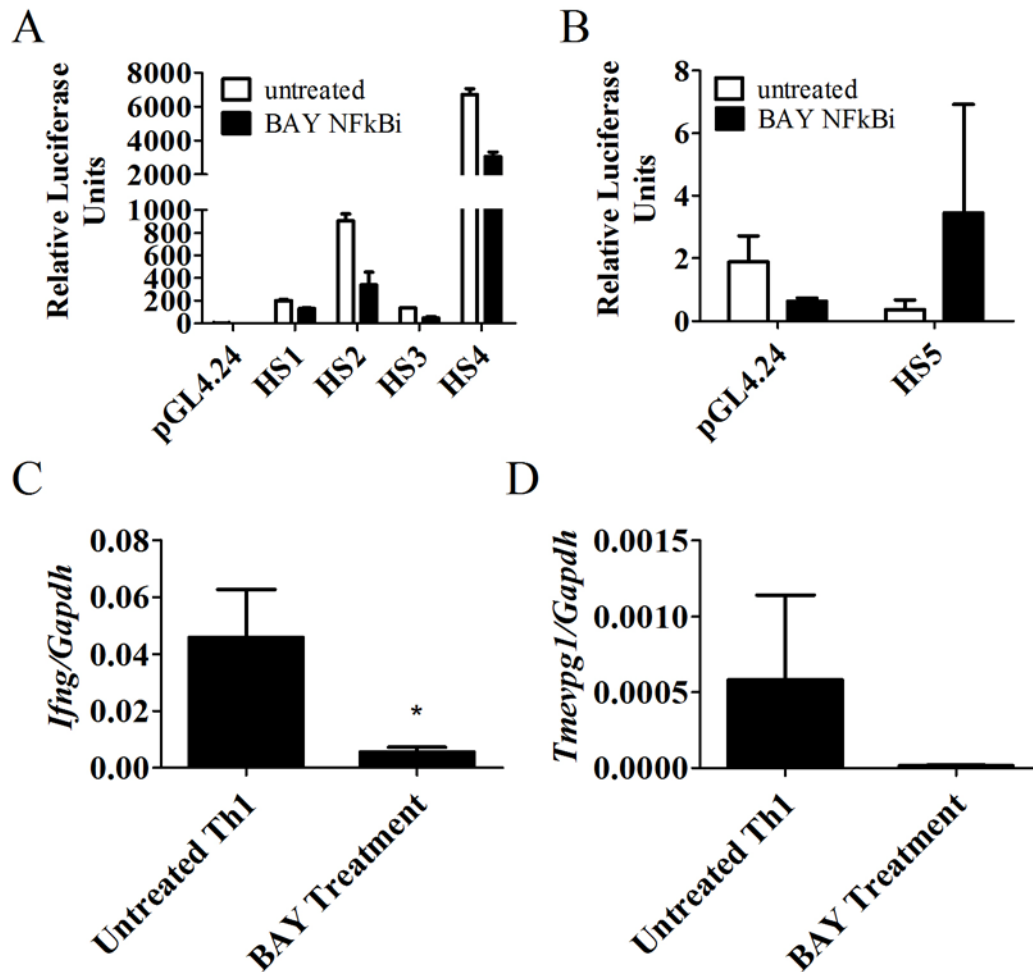
FIGURE 3-8. NF-κB inhibition reduces enhancer activity *in vitro* and *Tmevpg1* expression in effector T cell cultures. NF-κB inhibition by BAY11-7085 in Jurkat T cells and mouse effector T cell cultures. *A,* HS1-HS4-pGL4.24 or *B,* HS5-pGL4.24 enhancer constructs. Constructs were transfected into Jurkat T cells via DEAE transfection. Cells were incubated one hour in the presence of the inhibitor before stimulation with PMA and ionomycin for six hours. Relative luciferase activity was assessed. *C, Ifng* and *D, Tmevpg1* expression in effector T cells cultures. Cells were incubated one hour in the presence of the inhibitor before restimulation with immobilized anti-CD3 for 48 hours. All results in this figure are shown as the mean ± the standard error of the mean of two independent transfections. *$p > 0.05$

Discussion

Here, we provide evidence to extend the breadth of T-bet positive regulation of the Th1 locus from *Ifng* over 170 kilobases upstream to the gene encoding the *Tmevpg1* lncRNA. In previous studies, we found that *Tmevpg1* expression is dependent upon the master Th1 transcription factor T-bet, but this effect was poorly understood. To summarize our results, we find substantial *Tmevpg1* transcript expression by terminally differentiated effector-like cells primed to rapidly produce IFN-γ such as NK cells and polyclonal memory T cells. T-bet associates with the *Tmevpg1* locus promoting H4Ac marks in primary and effector Th1 cells. T-bet is required for the formation and maintenance of H4Ac marks downstream of *Ifng*, however, this region is not required for *IFNG* expression (112). This genomic region between *Tmevpg1* and *Ifng* contains a number of Th1-specific DNase 1 HS sites to which the inducible transcription factors, NF-κB and Ets-1 are recruited in a T-bet dependent fashion. Four of these sites possess strong transcriptional enhancer activity while the fifth possesses transcriptional repressor activity. Further, pharmacological inhibition of NF-κB reduces enhancer activity of the HS1-HS4 elements supporting that enhancement from these sites is mediated, in part, by NF-κB. Taken together, our data support a mechanism by which *Tmevpg1* is transcriptionally regulated by T-bet via epigenetic mechanisms enabling recruitment of inducible transcription factors to both the *Tmevpg1* promoter and gene body as well as distal transcriptional enhancers and repressors (Figure 3-9). Our findings not only provide a detailed description of lncRNA gene regulation as part of a developmental program but also contribute to the accepted mechanism of regulation of *Ifng* during Th1 lineage commitment.
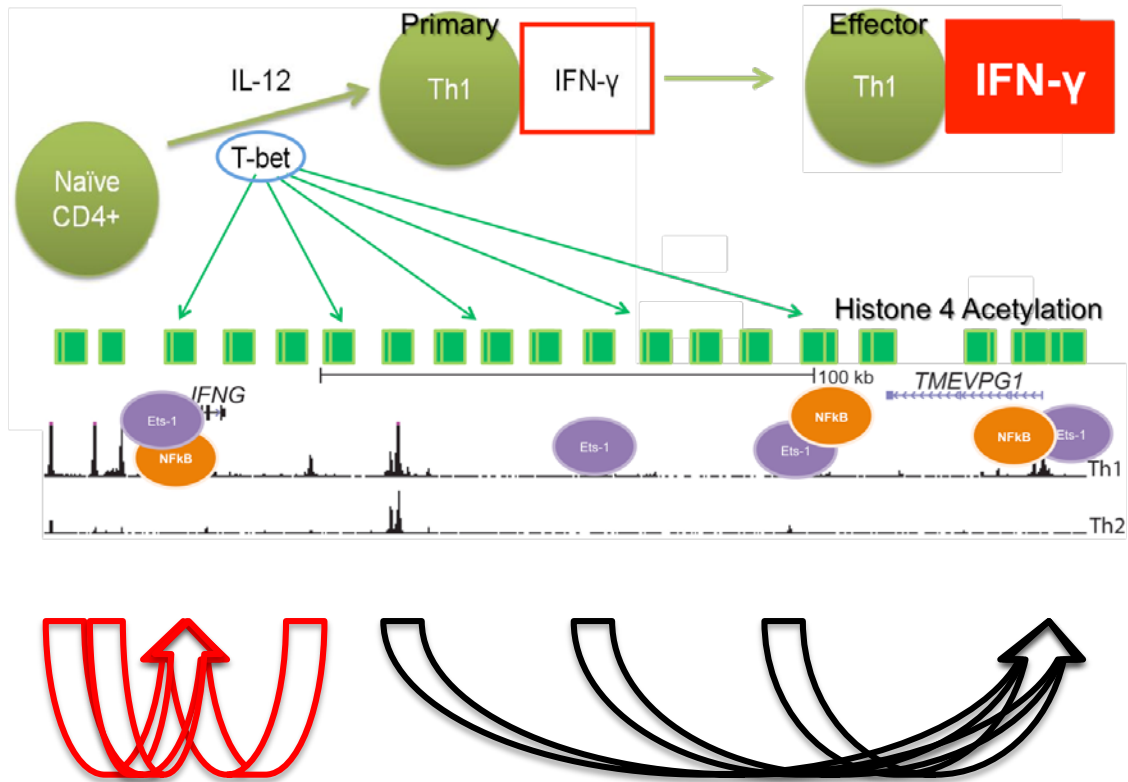
FIGURE 3-9. T-bet-dependent regulation of the Th1 locus. T-bet is required for the initial formation and maintenance of global H4Ac (green boxes) across the Th1 locus extending from *IFNG* through *TMEVPG1* during the primary as well as effector polarization stages of a naïve CD4$^+$ T cell. The open epigenetic confirmation of the locus allows for NF-κB (orange ovals) as well as Ets-1 (purple ovals) to bind to the *IFNG* and *TMEVPG1* promoters inducing expression. These factors also bind to surrounding HS elements within the intergenic region. *IFNG* (red arrows) and *TMEVPG1* (black arrows) demonstrate transcriptional regulation by surrounding genomic elements.

Our model is generally consistent with recent studies examining the requirement of *Tmevpg1* for expression of *Ifng* in response to infection with *Salmonella* (131). Here, *Tmevpg1* expression in CD8$^+$ T cells promotes rapid IFN-γ expression in response to *Salmonella* infection through formation of covalent H3K4me3 marks at the *Ifng* locus correlating with active transcription and rapid expression of IFN-γ by CD8$^+$ T cells (131). Moreover, in an *in vitro* assay, *Tmevpg1* physically associates with the WDR5 component of the MLL/Set1 histone-modifying complex responsible for producing the H3K4me3 mark (131). These findings are generally supported by studies examining function of other lncRNAs. Whereas repressive lncRNAs mediate epigenetic regulation of genes through association with PRC2-containing histone complexes, lncRNAs imposing positive regulation on target genes are commonly found to associate with the MLL/Set1 complex and associated their proteins.

Our results described herein, support these findings and establish a role for *Tmevpg1* in the priming of terminally differentiated cells such as effector Th1, NK and polyclonal memory cells to rapidly express IFN-γ in response to external stimuli. While initial expression of *Tmevpg1* is mediated through T-bet as part of the initial stages of the Th1 polarization program, the inducible burst of *Tmevpg1* in effector cells is mediated, in part, through inducible transcription factors, such as NF-κB and Ets-1.

Supplemental

Primer Name	5' → 3' sequence
-165 base pairs	Fwd ACGCTACATTTTAAGGACATGC
	Rev GCAGAGACACACCAACCAAA
+100 base pairs	Fwd TGGTGTGTCTCTGCAAGAATG
	Rev TCTGCTGCCAGGAAGAAAGT
+200 base pairs	Fwd AGAAGATGGGCAAGGTGT
	Rev TTCCCTTCACAGGTCTTTC
+400 base pairs	Fwd ACTTTGATGGAGGGGTGA
	Rev GGAAGCATGTCCCTATGA
+680 base pairs	Fwd CTTATAAGGCACCGTGTTTC
	Rev GATAATCTGGGTTTGACTCC
+875 base pairs	Fwd GCAGTAAACACCATCACTCA
	Rev CAGCTACCTCCAACAACTTAC
+1.4 kilobases	Fwd ATCCTACCCATCCTTTCAAC
	Rev ATCATGCCTTACCCTCAAC
+2.0 kilobases	Fwd CCTGGGGTACCAGAGAATAA
	Rev CCTAGTGCTTTAACCCTTGA
+2.3 kilobases	Fwd AGTACTGGGGCTAAAGGTGT
	Rev TTCTTTCAGTGAAGCCTACC
+3.0 kilobases	Fwd GAGGTGTTCTCCTGGTTACA
	Rev CAGGTAATAGCTGGATTCGT

SUPPLEMENTAL TABLE 3-1. Primer sequences utilized in ChIP assays.

Primer Name	5'→ 3' sequence
mHS1 (1)	FWD CTAGAATAATTCCACCAAAGCTTAC
	REV GAAGGCTAGAATGAAATGGG
mHS1 (2)	FWD GCCTTCCCTTCTCCTGTC
	REV GAGCCTGAGAGAAAAATAACACTT
mHS4 (1)	FWD AATTAATGCTGTGCTCATCTG
	REV TATCACCCAGCTTCCACTTC
mHS6 (1)	FWD GTATTGTGCAAAATTTTAGGAAC
	REV TATTTATTTCATGCCCCCAA
mHS6 (2)	FWD GGGGCATGAAATAAATAATCTT
	REV CATGGTTCCAATTTAATCCC
mHS8 (1)	FWD TAGAATAATTCCACCAAAGCTTA
	REV GGAAGGCTAGAATGAAATGG
mHS8 (4)	FWD ACCTGTCAAAGTGTGACTTTC
	REV CTCTAGGGTACTTCACAGTCTTAG

SUPPLEMENTAL TABLE 3-2. Primer sequences utilized in ChIP assays in murine T cell cultures.

Primer Name	5' → 3' sequence
Human TMEVPG1 HS 1	FWD GGTACCAAAACAGAATTCCCACCGTG
	REV CTCGAGTAGAATAGTTCAATCAAAGCTT
Human TMEVPG1 HS 2	FWD GGTACCTATAAATGAGCATCAGTCCC
	REV CTCGAGTGAATTCACCAGCTG
Human TMEVPG1 HS 3	FWD GGTACCGGGGGAAAAATACTAGTTAT
	REV CTCGAGTACTTCTGTCATTTGCCCAT
Human TMEVPG1 HS 4	FWD GGTACCATAAATGTCTCTATTCTGTA
	REV CTCGAGGGCAATATTTTAGAAATATT
Human TMEVPG1 HS 5	FWD GGTACCTTTCCATGCACAAATTATAG
	REV CTCGAGATTTATCCCTTTAAAACAGC

SUPPLEMENTAL TABLE 3-3. Primer sequences that used to amplify human enhancer sequences. Enhancer sequences were cloned in the pGL4.24 construct. Forward primers marked with a KPN1 and reverse primers were marked with Xho1 restriction enzyme sites on the 5' end, respectively.

Primer Name	5' → 3' sequence
1-509	FWD [Kpn1] AACCTGGTACCATTAAAAAAACAGAATTCCC
	REV [Xho1] TACGCTCGAGGGTTCTCCTAATTCC
1-270	FWD [Kpn1] AACCTGGTACCATTAAAAAAACAGAAT
	REV [EcoRV] ATCTCGAGGCAGCTTCCTGTT
507-777	FWD [Kpn1] AACCTGGTACCCCAACTGTAAGTG
	REV [EcoRV] CGACAAGATATCATTCCCTAGAATAG
270-509	FWD [Kpn1] ACATAGGTACCCTGCAATTTCAGGTAG C
	REV [Xho1] AATCTCGAGTTCTCCTAATTCCACCCA

SUPPLEMENTAL TABLE 3-4. Primer sequences to generate minimal promoter truncations of the human HS1 element. Enhancer sequences were cloned in the pGL4.10 construct.

CHAPTER IV

RNA-sequencing based discovery of lncRNAs in polarizing human T cell cultures

Overview

Current estimates enumerate 20,000-25,000 protein-coding genes and approximately 10,000 lncRNA-coding genes in the human genome (1, 44, 45). While significant in number, the lncRNA studies lack assignment of biologic meaning to most of the expressed transcripts. Moreover, these genome-wide surveys rely heavily on transformed cell lines possibly excluding entire branches of gene networks involved only in distinct processes of lineage commitment, for example, in T helper cell polarization. Therefore, we took a more biologically relevant approach to lncRNA discovery within the process of human T cell polarization from peripheral blood of a healthy human patient. We hypothesize that significant lncRNA populations will be expressed in these differentiating cells and variation in transcript expression will be largely dependent upon the program initiated: Th1, Th2, or Th17. Previously, lncRNA involvement in the process of lineage commitment, particularly within the immune system, was largely unknown. Utilizing the stochastic fate decisions within $CD4^+$ T cell polarization, we identified 2,788 lncRNA transcripts through a RNA-sequencing analysis of Th1, Th2 and Th17 *in vitro* polarized human PBMCs. Our work with *TMEVPG1* discussed in Chapters II and III establishes an experimental approach to subsequently identify additional lncRNA transcripts involved in execution of polarization programs by naïve helper T cells. This work demonstrates the

identification of large lncRNA networks of expression involved in key developmental fate decisions of cells such as T helper cell commitment.

Materials and Methods

Human lymphocyte culture conditions

Healthy human PBMC were isolated from one volunteers and were stimulated *in vitro* with anti-CD3, 1×10^6 cells/ml, under Th1: IL-12 (5 ng/ml), anti-IL-4 (5 ug/ml) and anti-IL17 (5 ug/ml), Th2: IL-4 (5 ng/ml) and anti-IFN-γ (5 ug/ml), or Th17: IL-1β (10 ng/ml), IL-6 (20 ng/ml), IL-23 (20 ng/ml), TGF-β1 (5 ng/ml) and anti-IFN-γ (5 ug/ml), polarizing conditions. Cultures were harvested after three days to obtain RNA from primary cultures. After five days, cultures were restimulated with anti-CD3 and harvested two days later to obtain RNA from effector cultures.

RNA-sequencing sample preparation

RNA was isolated from cultures using Tri-Reagent (Sigma). Library preparation was performed using the Illumina Tru-Seq RNA kit to generate a polyA-enriched cDNA library. Whole genome RNA sequencing was performed by the Vanderbilt Technologies for Advanced Genomics (VANTAGE) core. 50 base pairs paired-end reads were generated with an Illumina HiSeq 2500. A quality control step was initially performed on the raw data to identify potential outliers before any advanced analysis using tools such as Fastx Toolkit and FastQC. The RNA data were aligned with TopHat and gene expression levels

were quantified using Cufflinks. RPKM (reads per kilobase per million reads) based approaches (Cuffdiff) were used to detect differentially expressed genes. False discovery rate (FDR < 0.05) was used to correct for multiple testing. The UCSC genome browser (GRCh37/hg19 build) were used to determine chromosomal locations and for other computational purposes as described in the text.

Results

T helper cell lineage-specific lncRNA expression

Our understanding of lncRNAs within the immune system is limited thus in efforts to identify lncRNA expression in primary and effector T helper cell polarization, we employed a whole genome RNA-sequencing approach. Chase Spurlock, a trainee in the Aune laboratory, initiated these studies under Dr. Aune's direction. Peripheral blood was collected from a healthy patient and the mononuclear cells were purified followed by *in vitro* culture under Th1, Th2 or Th17 polarizing conditions. We assessed expression of 36 protein-coding genes within the RNA-sequencing analysis as well as through RT-PCR. Quantitative transcript determinations were highly correlated across this set of genes that exhibit a wide range of expression (quantitative range $>10^4$) (Figure 4-1*A*). Thus, quantitation of transcript levels by RNA-sequencing yielded similar results to RT-PCR. Approximately 54% of the total estimated protein-coding genes and about 28% of the total lncRNA encoding genes were expressed at detectable levels in our PBMC cultures relative to the total estimated number of protein and lncRNA encoding genes known (Figure 4-1*A*). Additionally, mean expression levels of all mRNAs identified by the RNA-sequencing

analysis were found to be substantially higher than the mean expression levels of all lncRNAs (about 3-fold) as were the median levels of mRNAs relative to lncRNAs (Figure 4-1*B*). Consistent with findings in other experimental systems, the average expression level of lncRNA transcripts is reduced in comparison to levels of mRNAs in PBMC cultures.
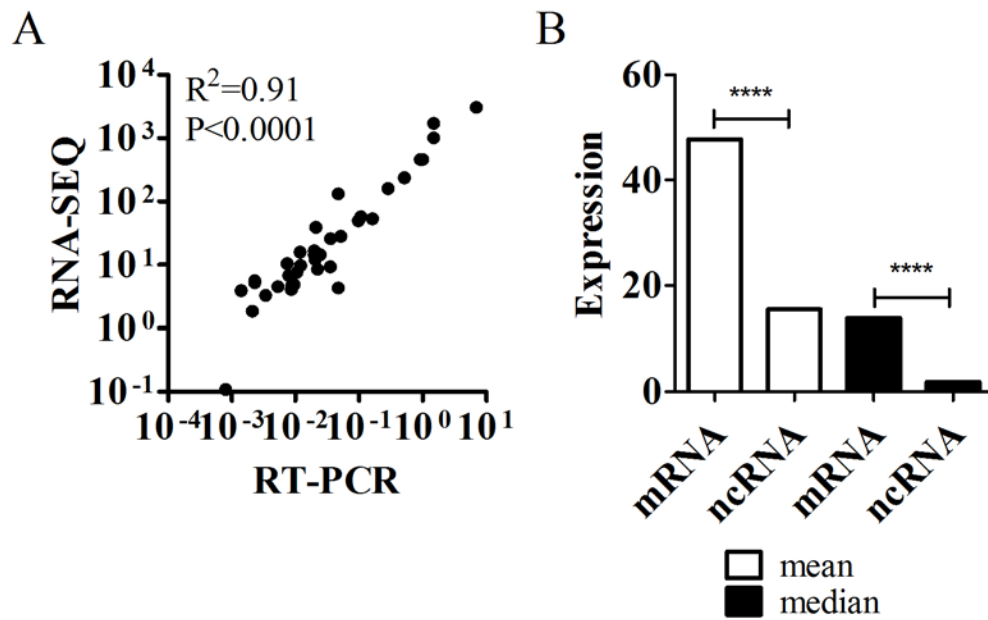
FIGURE 4-1. Expression levels of mRNAs and lncRNAs. *A,* Expression correlation of 36 genes determined by RT-PCR and by RNA-sequencing. R^2 was determined by Pearson's linear regression. *B,* Mean and median transcript levels of all expressed mRNAs and all expressed lncRNAs in PBMC cultures were determined by RNA-sequencing.

We subsequently evaluated the quality of our cultures through detection of genes encoding lineage-specific cytokines and transcriptional regulators. As predicted, *IFNG* was expressed selectively under Th1 polarizing conditions. Similarly, *IL4*, *IL13*, and *IL5* were expressed selectively under Th2 polarizing conditions and *IL17A, IL17F* and *IL10* were expressed predominantly under Th17 polarizing conditions (Figure 4-2*A*). Further, we examined profiles of transcriptional regulators of these well-established effector lineages. Skewing of effector programs is highly dependent upon JAK/STAT signaling mediators downstream of cytokine receptor stimulation for example, *STAT1* and *STAT4* for the Th1 program; *STAT6* for the Th2 program and STAT3 for the Th17 program (Figure 4-3*A*). Whereas *STAT1* and *STAT3* expression correlated with Th1 and Th17 programs, respectively, *STAT4* was expressed in both Th1 and Th2 lineages. Whereas *STAT6* is associated with the Th2 polarization program, we observed equivalent expression levels between all primary and effector cultures. Further, we observed that key transcriptional regulators of Th1: *TBX21*, Th2: *GATA3*, and Th17: *RORC* and *BATF* were substantially expressed in the corresponding T helper cell cultures (Figure 4-3*B*). Taken together, these results demonstrate that our culture conditions exhibited successful T helper lineage-specific expression of transcriptional regulators as well as cytokine genes validating the usefulness of this system in the investigation of lncRNAs involved in T helper polarization.
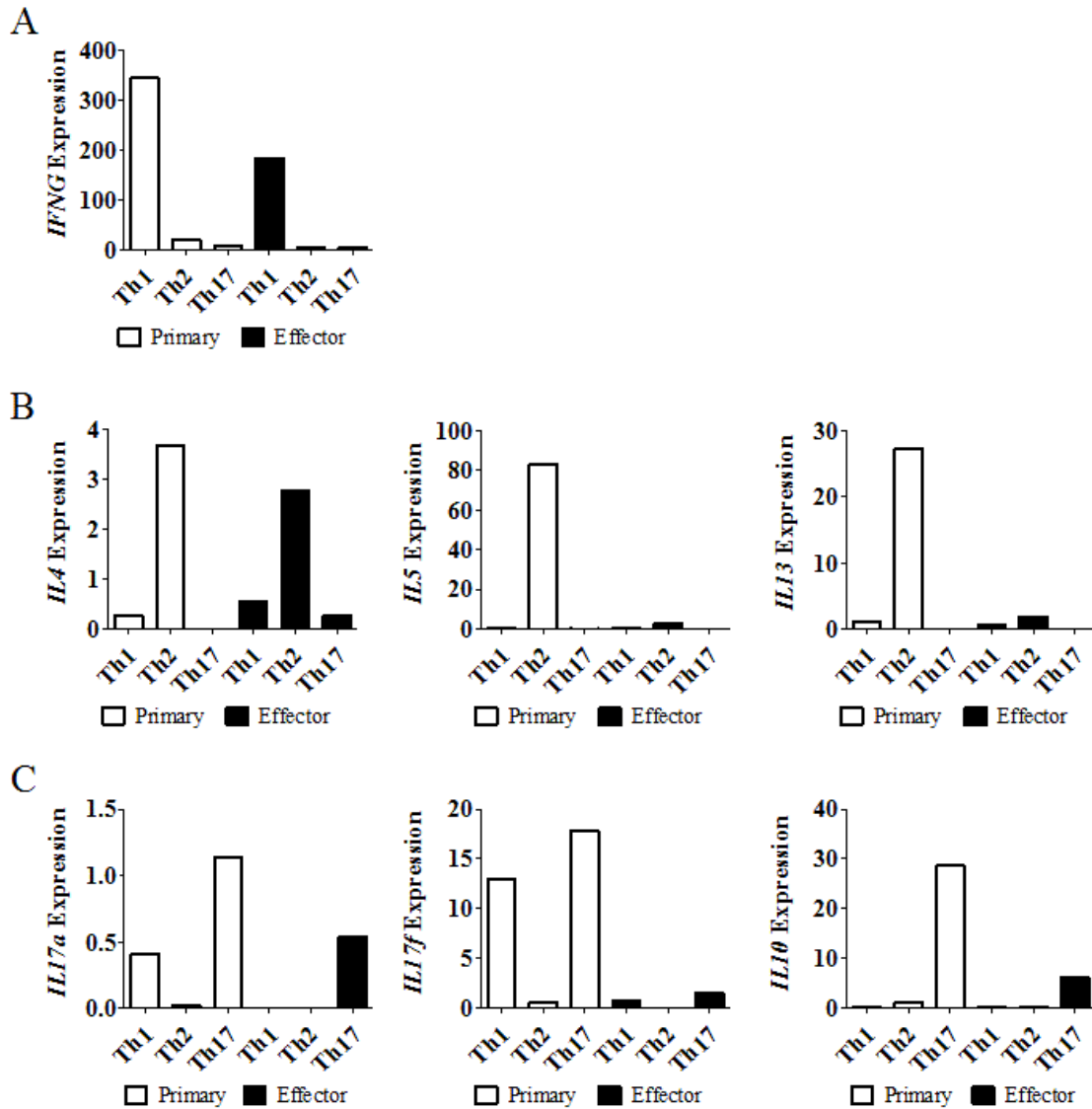
FIGURE 4-2. Lineage-specific cytokine gene expression in T helper cells. Lineage-specific cytokine gene expression in *A*, Th1 *B*, Th2 and *C,* Th17 primary and effector cultures. Results are expressed as normalized sequence reads.

Approximately 2,788 lncRNA transcripts were identified through our RNA-sequencing analysis. To begin to assign meaning to the quantified transcripts, I examined ratios of lncRNAs within each subset relative to the other two subsets. In general, we observed equivalent numbers of lncRNAs expressed at least two-fold greater relative to the other subsets in primary polarized Th1, Th2 and Th17 cells ranging between 277 to 369 individual transcripts (Supplemental Table 4-1). In contrast, transcript numbers derived from Th1 effector cultures exhibited 882 and 868 individual transcripts relative to Th2 and to Th17 cultures, respectively. Similar numbers were observed when Th2 or Th17 derived lncRNAs were compared to those in Th1 cultures; however, Th2 transcripts were fewer in number at 257 relative to Th17 numbers and to a similar extent vice versa between the two cultures (Supplemental Table 4-1). We take these observations to mean that lncRNA genes are differentially induced in Th1, Th2 and Th17 cells in response to polarization under primary or effector stimulation.
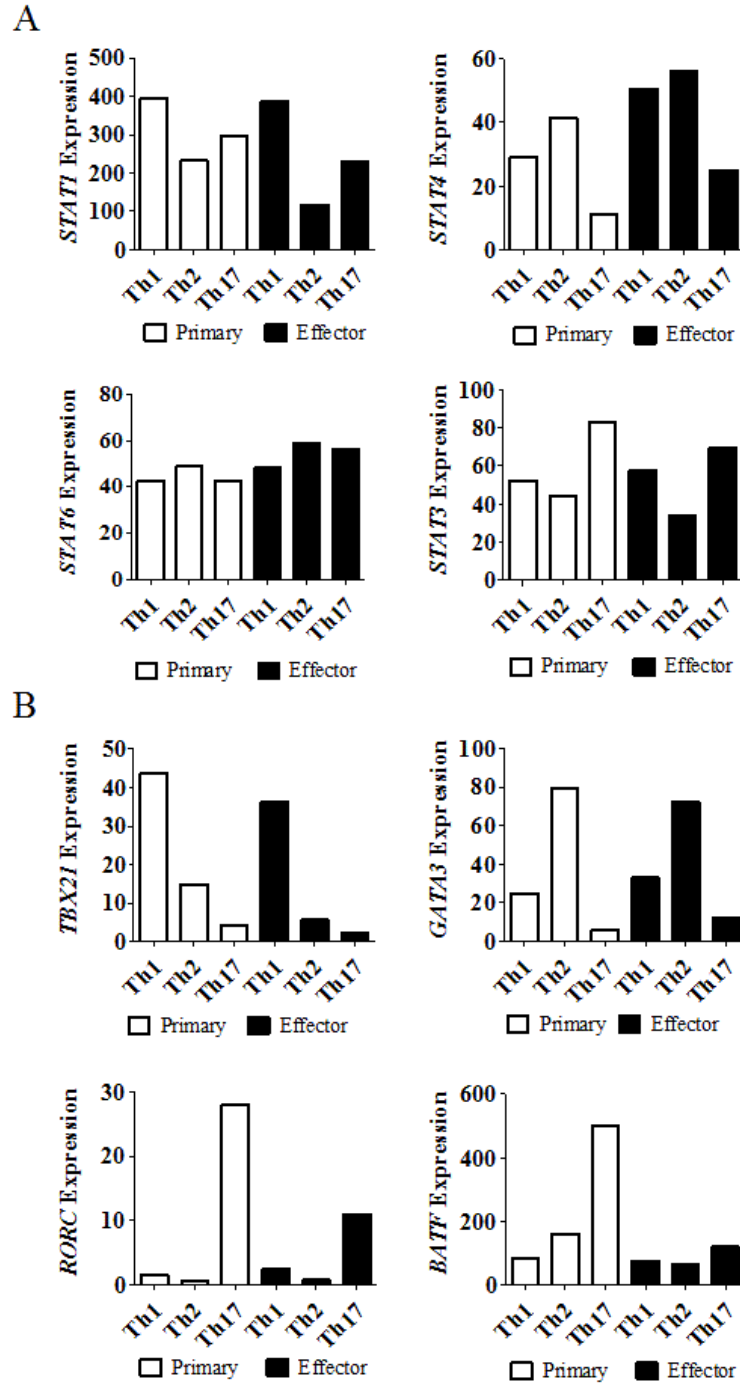
FIGURE 4-3. Master transcriptional regulator gene expression determined by RNA-sequencing. *A*, Expression of *STAT* genes and *B*, genes encoding lineage-specific "master" transcription factors in primary and effector polarized Th1, Th2 and Th17 cell cultures. Results are expressed as normalized sequence reads.

As demonstrated in Chapters II and III, I described a strategy to assign biologic function to a Th1-specific lncRNA that fails to be expressed in Th2 as well as in Th17 cells. Utilizing *Tmevpg1* expression and regulation as a guide, I examined the expression of lncRNA transcripts expressed in each T cell subset. Consistent with our previous findings, *TMEVPG1* is selectively expressed in Th1 cells (Figure 4-4*A*). Similarly, we identified two novel lncRNA transcripts AC007278.2 and AC007278.3 that were also expressed in a Th1-specific manner. AC007278.3 is expressed in both primary and effector Th1 cultures while AC007278.2 is substantially expressed by effector Th1 cultures. Both lncRNAs are transcribed from two separate intron regions of the IL18 receptor accessory protein (IL18RAP), a gene known to be expressed in Th1 cells and regulated by T-bet (132). We further examined 35 additional lncRNAs that demonstrated preferential expression relative to Th2 and Th17 primary or effector cells.

Similarly we discovered *GATA3-AS1*, found adjacent to the gene encoding the Th2 master transcription factor GATA-3. My analysis of the *GATA3-AS1* transcript indicates that it is a 3,035 base pair gene located 1.2 kilobases away from the transcription start site of *GATA-3*. *GATA3-AS1* appears in two forms, a 2.2 kilobase long transcript comprised of two exons and a second 1.369 kilobase long further spliced transcript consisting of four exons (Supplemental Figure 4-1*A*). Transcription of *GATA3-AS1* occurs on the negative strand relative to *GATA-3* and proceeds in the opposing direction; however, whether *GATA-3* and *GATA3-AS1* share a bidirectional promoter is unknown. Sequence conservation is highest among placental mammals. Further, epigenetic marks indicate that the region upstream of the putative promoter for *GATA3-AS-1* is marked with permissive

marks favorable for transcription. DNase I hypersensitivity analysis, through the UCSC genome browser build, suggests that this area is also accessible in polarized human CD4$^+$ Th1, Th2 and Th17 cells. One novel lncRNA, *AC004041.2* is also selectively expressed in primary and effector Th2 cultures relative to primary and effector Th1 and Th17 cultures. We found 20 additional lncRNAs that were preferentially expressed in primary or effector Th2 cells. Lastly, 29 lncRNAs were found to be selectively expressed in primary Th17 cells including *AC004041.2* and *RP11-98D18.3*.
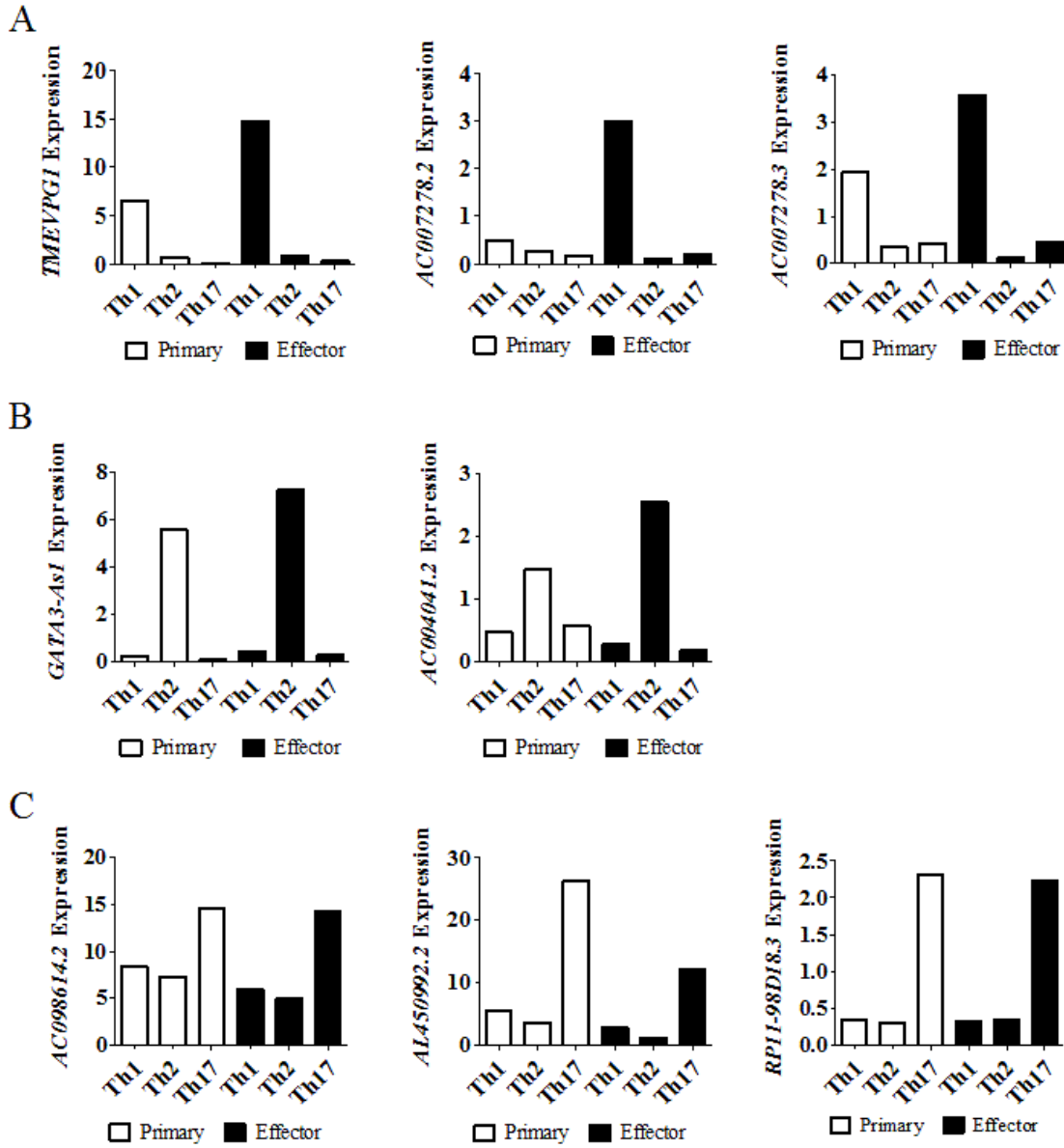
FIGURE 4-4. T helper program-specific lncRNA expression. LncRNA gene expression in *A*, Th1 *B*, Th2 and *C*, Th17 primary and effector stimulated cultures. Results are expressed as normalized sequence reads.

Identification of co-expressed lncRNAs and mRNAs

TMEVPG1 and *IFNG* are co-expressed under Th1 culture conditions and *TMEVPG1* is required for *IFNG* transcription in both tissue culture and infectious models (131, 133). *GATA3* and *GATA3-AS1* are also co-expressed under Th2 culture conditions (Supplemental Figure 4-1*B* and 4-1*C*). For these reasons, we determined the extent to which T helper lineage-specific lncRNAs were co-expressed with adjacent protein-coding genes in the genome. We identified adjacent protein-coding genes using the UCSC genome browser and used Pearson's linear regression to quantify co-expression across primary and effector Th1, Th2, and Th17 cultures. This analysis confirmed that *TMEVPG1* and *IFNG* were co-expressed and that *GAT3-AS1* and *GATA3* were also co-expressed (Supplemental Table 4-2 and Table 4-3). In addition, we found that the novel lncRNA, *AC004041.2*, was co-expressed with *IL4*, *IL13*, and *IL5* genes. The lncRNA *AC004041.2* is located within the *RAD50* gene but is not co-expressed with *RAD50*. Interestingly, the genomic position of *AC004041.2* aligns precisely with the previously described Th2 locus control region required for coordinated expression of *IL4*, *IL13*, and *IL5* (35). We also found that the lncRNAs, *RP11-98D18.3* and *AC007182.6*, were co-expressed with adjacent or nearby *RORC* and *BATF*, respectively. In fact, 85% of lncRNAs expressed in a Th1, Th2, or Th17 lineage-specific manner were adjacent in the genome to co-expressed protein-coding genes (Supplementary Tables 4-2, 4-3 and 4-4). A total of 41% of these lineage-specific lncRNAs were intragenic or within their co-expressed protein-coding gene. A second feature of this genomic organization was that 38% of lineage-specific lncRNAs were localized within clusters of 2 or more co-expressed protein-coding genes

(Supplementary Tables 4-2, 4-3 and 4-4). These results are consistent with the general notion that lncRNAs tend to regulate adjacent protein-coding genes and identify co-expressed protein-coding genes as potential lncRNA targets.

We also identified lncRNAs with selectively reduced levels of expression in primary and effector Th1, Th2, or Th17 culture conditions relative to the opposing culture conditions, respectively. For example, we identified lncRNAs expressed at low levels in Th2 cultures and high levels in Th1 and Th17 cultures (Supplementary table 4-5). As above, we examined expression levels of neighboring protein-coding genes and found that co-expressed protein-coding genes were more likely to be intragenic or next to the lncRNA in question as opposed to far away from the lncRNA in question. Like the lncRNAs, the co-expressed protein-coding genes were uniformly under-expressed in Th2 cultures compared to Th1 and Th17 cultures. Thus, these data support the possibility that these lncRNAs, similar to *TMEVPG1,* may act as transcriptional enhancers of neighboring protein-coding genes.

To further evaluate these points, we compared distances between lineage-specific lncRNAs and co-expressed protein-coding genes as a function of distances in kilobases between lineage-specific lncRNAs and adjacent protein-coding genes and as a function of adjacent protein-coding genes in the genome. Both types of comparisons demonstrated that the majority of protein-coding genes (85%) that contained a lineage-specific lncRNA were co-expressed with the lncRNA (intragenic) (Figure 4-5*A* & 4-5*B*). Approximately 50% of protein-coding genes within 50 kilobases of a lineage-specific lncRNA or were

adjacent to the lineage-specific lncRNA were also co-expressed. This frequency decreased further by moving further away from the lineage-specific lncRNA. Thus, proximity in the genome is an important contributor to whether or not protein-coding genes and lncRNAs are co-expressed. We also asked if lncRNA genes localized on the sense or antisense strands of co-expressed protein-coding genes. We found that there was about a 50:50 distribution between lncRNA genes and co-expressed protein-coding genes transcribed from the same DNA strand or transcribed from the opposite DNA strand (Figure 4-5*C*). Thus, transcription from the same DNA strand or from opposite DNA strands was unbiased for lncRNA genes and co-expressed protein-coding genes.
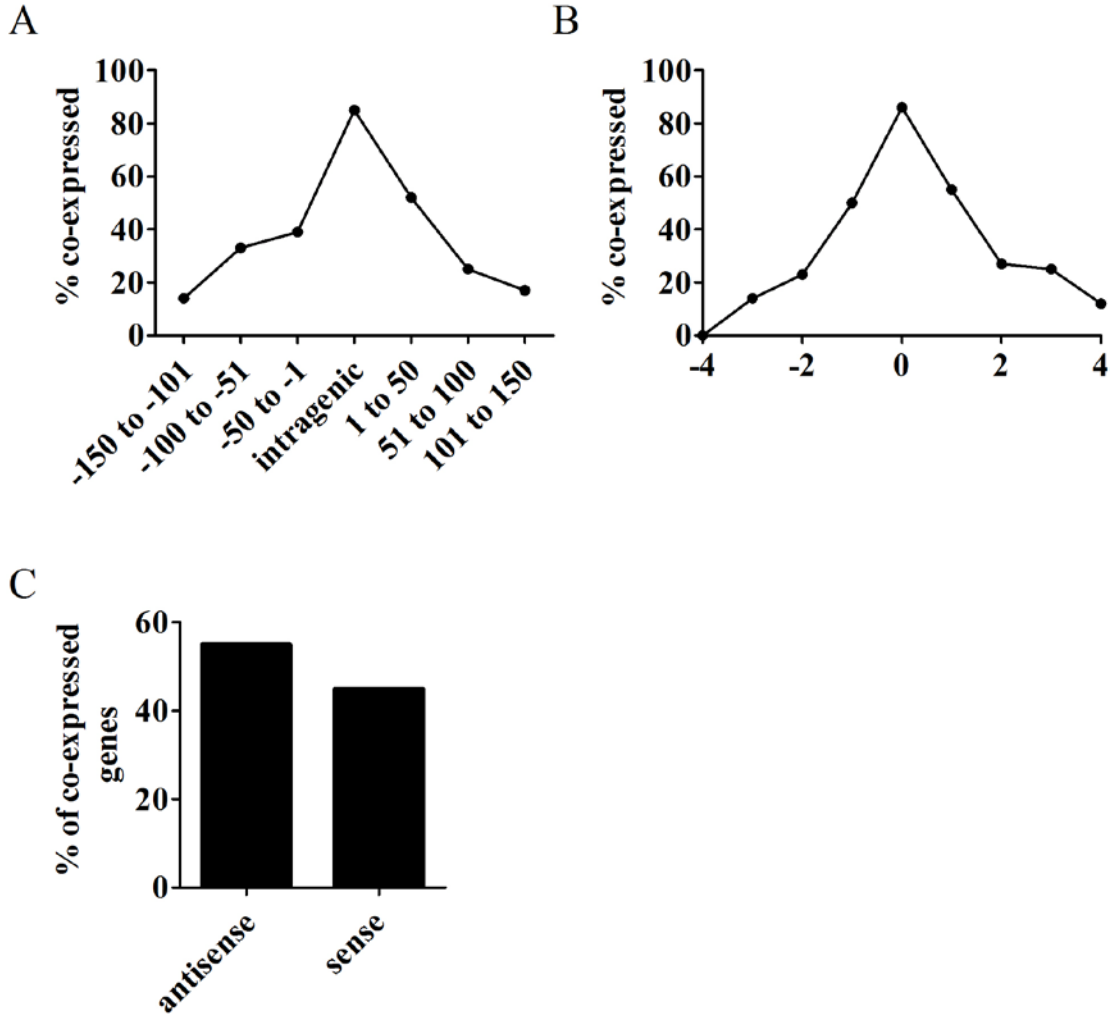
FIGURE 4-5. Frequency of co-expression between lncRNA genes and neighboring protein-coding genes. *A*, Co-expression of lncRNA and protein-coding genes as a function of genomic distance in kilobases. *B,* Co-expression of lncRNA and protein-coding genes as a function of nearby protein-coding genes. *C,* Percentage of co-expressed lncRNA and protein-coding genes as a function of directionality.

Discussion

Previously, lncRNA involvement in the process of lineage commitment, particularly within the immune system, was largely unknown. Utilizing the stochastic fate decisions within CD4$^+$ T cell polarization, we identified 2,788 lncRNA transcripts through a RNA-sequencing analysis of Th1, Th2 and Th17 *in vitro* polarized human PBMCs. Analysis of these transcripts led to several observations. First, each T helper cell program induces specific expression of numerous lncRNA transcripts. Whereas our previous work established evidence for *TMEVPG1*, we now identified *GATA3-AS1* and *RP11-98D18.3* as Th2 and Th17-specific lncRNAs, respectively. Second, through linear regression, we determined that these lncRNAs correlated strongly with neighboring protein-coding genes. Lastly, our analysis of global expression demonstrates a strong correlation for co-expression and genomic distance as it relates to lncRNA expression.

In addition to a screen examining lncRNA expression within various stages of development as well as activation states of CD8$^+$ T cells, our RNA-sequencing results are accompanied by a recent report of greater than 1,500 lncRNA transcripts identified in 42 subsets of murine CD4$^+$ T cells (75, 134). Consistent with our results, global patterns of lncRNA expression were distinct within each polarized T helper cell subset. Additionally, this screen identified significant numbers of unique lncRNA transcripts within the Th1 population relative to Th2 and Th17. Moreover, T-bet was found to associate with the *Tmevpg1* locus, as our data suggest, as well as 209 addition lncRNA genes supporting that master transcriptional regulatory molecules are responsible for not only activation of

hallmark cytokines and suppression of alternative lineages but also initiating patterns of lncRNA expression. The question remains as to the functional requirement of each individual lncRNA and the cooperation of gene as well lncRNA regulatory networks culminating in appropriate cell fate decisions. As the lncRNA field is in its infancy, involvement of lncRNAs as key regulators in essential cellular decisions will become apparent as indicated by our observations.

Supplemental Tables and Figures

Tables

	PRIMARY			EFFECTOR		
Relative cultures	*Th1*	*Th2*	*Th17*	*Th1*	*Th2*	*Th17*
Th1		277	285		882	868
Th2	323		369	988		257
Th17	289	277		968	312	

SUPPLEMENTAL TABLE 4-1. Differentially expressed lncRNA transcripts among T helper cell subsets. The number of transcripts in each row represent the number of relative reads from the subset designated in each row (**Th1, Th2 or Th17**) divided by the relative reads from the culture in each vertical column (*Th1, Th2 or Th17*). Cutoff was 2-fold higher expression and ratios where comparative column had zero reads were excluded.

LncRNA	Genomic position	Protein-coding	R^2	P	Distance away in kilobases	Orientation
RP11-430C7.4	1:204,572,162-204,585,693	*MDM4*	0.27	NS	intragenic	sense
		LRRN2	**0.91**	**0.003**	**0.6**	**antisense**
RP3-340N1.5	1:20,510,734-20,522,541	*PLA2G5*	0.16	NS	99	antisense
		PLA2G2D	0.1	NS	71	sense
		PLA2G2F	0.09	NS	45	antisense
		PLA2G2C	**0.81**	**0.01**	**7**	**sense**
		UBXN10	0	NS	intragenic	antisense
		VWA5B1	0.47	NS	107	antisense
CHRM3-AS2	1:239,870,426-239,882,419	***CHRM3***	**0.88**	**0.005**	**intragenic**	**antisense**
RP11-343J24.1	1:239,867,192-239,882,366	***CHRM3***	**0.92**	**0.003**	**intragenic**	**antisense**
RP11-288L9.4	1:27,992,571-27,998,729	***IFI6***	**0.92**	**0.003**	**intragenic**	**antisense**
		AHDC1	0.31	NS	132	antisense
		FGR	0.12	NS	54	antisense
		FAM76A	0.03	NS	60	sense
		STX12	0	NS	107	sense
RP11-525A16.4	10:112,257,757-112,258,300	***DUSP5***	**0.91**	**0.003**	**intragenic**	**antisense**
		SMC3	0.06	ns	70	sense
RP11-9E13.2	10:70,237,755-70,240,521	*HNRNPH3*	0.02	NS	146	antisense
		RUFY2	0	NS	100	sense
		DNA2	0	NS	63	sense
		SLC25A1	0.045	NS	5	sense
		TET1	0.19	NS	83	sense
RP11-886D15.2	11:104,934,071-104,942,268	***CASP1***	**0.77**	**0.02**	**intragenic**	**sense**
		CARD16	**0.79**	**0.02**	**22**	**sense**
		CASP4	**0.63**	**0.04**	**21**	**sense**

		CASP5	**0.9**	**0.004**	**70**	**sense**
		CARD17	**0.93**	**0.002**	**29**	**sense**
RP11-428C19.5	11:19,321,430-19,329,905	*CSRP3*	0.21	NS	118	antisense
		E2F8	0.04	NS	76	antisense
		NAV2	**0.75**	**0.03**	**413**	**sense**
RP11-326C3.2	11:287,305-288,298	*PSMD13*	0.23	NS	51	sense
		NLRP6	**0.94**	**0.0013**	**9**	**sense**
		ATHL1	**0.99**	**<0.0001**	**2**	**sense**
		IFITM2	0.28	NS	25	sense
		IFITM3	0.34	NS	33	antisense
RP11-113K21.3	11:82,817,622-82,818,003	*C11orf82*	0	NS	205	antisense
		RAB30	**0.65**	**0.04**	**125**	**sense**
		PCF11	0.47	NS	51	sense
		ANKRD42	0.13	NS	88	antisense
		CCDC90B	**0.69**	**0.04**	**155**	**sense**
RP11-23J18.1	12:47,529,545-47,532,153	*AMIGO2*	0.32	NS	58	sense
		PCED1B	**0.8**	**0.01**	**81**	**antisense**
IFNG-AS1	12:68,383,225-68,415,107	*IFNG*	**0.95**	**<0.0001**	**140**	**antisense**
RP11-275I4.2	15:38,964,056-38,970,209	*C15orf53*	0.06	NS	24	sense
		RASGRP1	**0.92**	**0.003**	**184**	**antisense**
AC005838.2	17:15,468,796-15,587,613	*FAM18B2*	**0.09**	**0.004**	**142**	**sense**
		CDRT4	0.08	NS	129	sense
		TRIM16	0.01	NS	intragenic	antisense
		CDRT1	**0.78**	**0.02**	**0**	**antisense**
TOB1-AS1	17:48,939,583-48,987,593	*LUC7L3*	**0.83**	**0.01**	**143**	**sense**
		TOB1	**0.88**	**0.006**	**intragenic**	**antisense**
		SPAG9	0.33	SN	169	antisense

RNF157-AS1	17:74,136,637-74,150,364	*ZACN*	0.02	NS	61	sense
		EXOC7	0	NS	59	antisense
		UBALD2	0.18	NS	125	sense
		RNF157	**0.93**	**0.002**	**intragenic**	**antisense**
CTC-378H22.2	19:42,656,721-42,661,675	*GRIK5*	0.21	NS	131	antisense
		ZNF574	0.07	NS	76	sense
		POU2F2	**0.78**	**0.02**	**intragenic**	**antisense**
		DEDD2	0.6	NS	46	antisense
		ZNF526	0.04	NS	68	sense
		GSK3A	0.01	NS	78	antisense
		ERF	0.09	NS	31	antisense
AC007278.3	2:103,055,173-103,056,935	*IL1RL1*	0.38	NS	128	sense
AC007278.2	2:103,050,416-103,051,800	*IL18R1*	0.6	NS	40	sense
		IL18RAP	**0.95**	**0.001**	**intragenic**	**sense**
		SLC9A4	0.04	NS	34	**sense**
AC009299.2	2:162,079,296-162,111,179	***TANK***	**0.74**	**0.03**	**intragenic**	**antisense**
		PSMD14	0.1	NS	85	antisense
AC008063.2	2:162,836,116-162,974,655	***SLC4A10***	**0.99**	**<0.0001**	**87**	**antisense**
		DPP4	**0.97**	**0.0003**	**intragenic**	**antisense**
		GCG	0.43	NS	163	antisense
		FAP	0.19	NS	191	sense
AF131217.1	21:29,816,870-30,047,170	*N6AMT1*	0.48	NS	201	sense
		LTN1	0	NS	253	sense
		RWDD2B	0.52	NS	350	sense
		USP16	0.05	NS	354	antisense
RP11-377G16.2	4:81,104,434-81,111,323	***PRDM8***	**0.92**	**0.003**	**intragenic**	**antisense**
		FGF5	0.47	ns	83	antisense

CTD-2313F11.1	5:54,317,127-54,319,997	*ESM1*	**0.85**	**0.008**	**intragenic**	**sense**
		GZMK	**0.99**	**<0.0001**	**intragenic**	**antisense**
		GZMA	0.59	NS	81	sense
		CDC20B	0.05	NS	103	antisense
RP11-325F22.2	7:104,581,653-104,602,781	*KMT2E*	0.45	ns	73	sense
RP5-1051J4.4	7:150,446,824-150,447,182	*GIMAP2*	0.39	ns	64	sense
		GIMAP1	0.39	NS	11	sense
		GIMAP5	**0.8**	**0.02**	**6**	**sense**
		TMEM176B	0.25	NS	42	antisense
		TMEM176A	0.16	NS	51	sense

SUPPLEMENTAL TABLE 4-2. Th1 related lncRNA genes and neighboring protein-coding genes.

LncRNA	Genomic position	Protein-coding	R^2	P	Distance away in kilobases	Orientation
RP11-261C10.3	1:243,219,616 -243,265,046	*CEP170*	0.02	NS	44	sense
RP3-395M20.8	1:2,481,359- 2,488,450	*PLCH2*	0.02	NS	74	antisense
		PANK4	0	NS	42	antisense
		HES5	0	NS	21	antisense
		TNFRSF14	**0.85**	**0.008**	**6**	**antisense**
		TTC34	**0.66**	**0.04**	**91**	**sense**
RP11-38L15.3	10:46,951,472 -46,966,835	**SYT15**	**0.93**	**0.002**	**2**	**antisense**
RP11-144G6.12	10:47,213,346 -47,243,502	**ANXA8L1**	**0.94**	**0.002**	**56**	**sense**
		ANXA8	**0.97**	**0.0005**	**56, sense**	**sense**
		FAM25C	**0.97**	**0.0004**	**36, sense**	**sense**
		FAM25B	**0.99**	**<0.0001**	**36, sense**	**sense**
		AGAP9	**0.92**	**0.002**	**intragenic**	**sense**
		AGAP10	**0.96**	**0.0007**	**intragenic**	**sense**
RP11-508M1.3	10:48,927,373 -48,950,972	**PTPN20B**	**0.0004**	**0.97**	**100**	**sense**
		AGAP8	**0.99**	**<0.0001**	**26**	**sense**
		BMS1P1	**0.92**	**0.002**	**intragenic**	**sense**
GATA3-AS1	10:8,092,413- 8,095,447	*TAF3*	0.03	NS	35	sense
		GATA3	**0.87**	**0.007**	**1**	**antisense**
RP11-234B24.4	12:4,809,583- 4,829,268	*NDUFA9*	0.09	NS	51	sense
		GALNT8	0.21	NS	20	sense

		KCNA6	**0.91**	**0.003**	**109**	**sense**
		KCNA1	0.1	NS	110	sense
RP11-977G19.12	12:56,702,652-56,703,233	*SMARCC2*	0.01	NS	119	
		RNF41	**0.87**	**0.006**	**87**	**antisense**
		SLC39A5	0.11	NS	71	sense
		ANKRD52	0.17	NS	50	antisense
		COQ10A	0.05	NS	38	sense
		CS	0.04	NS	23	antisense
		CNPY2	0.3	NS	intragenic	antisense
		PAN2	0.15	NS	8	antisense
		IL23A	0.14	NS	30	sense
		STAT2	0.1	NS	40	antisense
RP11-731F5.1	14:106,067,809-106,071,694	*MTA1*	0.48	NS	81	antisense
		CRIP2	0.11	NS	128	sense
		TMEM121	0.2	NS	75	sense
		IGHE	**0.98**	**0.0001**	**intragenic**	**sense**
CTD-3092A11.2	15:30,780,166-30,782,516	*CHRFAM7A*	0.45	NS	127	
		GOLGA8R	0.23	NS	88	
		GOLGA8Q	0.12	NS	64	
CTD-2616J11.3	19:51,917,552-51,918,219	***IGLON5***	**0.95**	**0.001**	**102**	**sense**
		VSIG10L	**0.95**	**0.0008**	**72**	**antisense**
		ETFB	**0.97**	**0.0004**	**59**	**antisense**
		CLDND2	**0.67**	**0.04**	**45**	**antisense**
		NKG7	0.11	NS	42	sense
		LIM2	0.01	NS	34	antisense
		SIGLEC10	**0.98**	**0.0002**	**intragenic**	**antisense**
		SIGLEC12	**0.68**	**0.04**	**77**	**antisense**
AC017074.2	2:114,435,279-114,461,655	*DDX11L2*	0.35	NS	79	sense

		RPL23AP7	**0.85**	**0.009**	**51**	**sense**
		RABL2A	0.18	NS	35	antisense
		SLC35F5	0.24	NS	10	sense
AC105344.2	2:231,849,083-231,860,746	*CAB39*	0.11	NS	164	sense
		ITM2C	**0.73**	**0.03**	**120**	**sense**
		GPR55	**0.96**	**0.0005**	**74**	**antisense**
		SPATA3	0.31	NS	11	antisense
		PSMD1	0.26	NS	72	antisense
		HTR2B	**0.8**	**0.02**	**123**	**sense**
DGUOK-AS1	2:74,174,769-74,208,568	*STAMBP*	0.38	NS	49	sense
		DGUOK	**0.91**	**0.003**	**intragenic**	**antisense**
		TET3	0.22	NS	99	antisense
AC004041.2	5:131,966,281-131,977,465	*IRF1*	0.13	NS	149	sense
		IL5	**0.89**	**0.004**	**89**	**sense**
		RAD50	0.54	NS	intragenic	antisense
		IL13	**0.9**	**0.004**	**27**	**antisense**
		IL4	**0.83**	**0.008**	**13**	**antisense**
		KIF3A	0.25	NS	32	sense
		CCNI2	0.24	NS	117	antisense
RP11-305L7.1	9:93,867,239-93,869,586	***AUH***	**0.85**	**0.009**	**9**	**sense**
RP11-305L7.3	9:93,881,420-93,925,369	***AUH***	**0.87**	**0.006**	**200**	**sense**

SUPPLEMENTAL TABLE 4-3. Th2 related lncRNA genes and neighboring protein-coding genes.

LncRNA	Genomic position	Protein-coding	R^2	P	Distance away in kilobases	Orientation
RP11-126K1.6	1:151,319,443-151,320,503	VPS72	0.53	NS	171	sense
		PIP5K1A	0.07	NS	148	antisense
		PSMD4	0	NS	80	antisense
		ZNF687	0.01	NS	55	antisense
		PI4KB	0.39	NS	19	sense
		RFX5	**0.7**	**0.03**	**intragenic**	**antisense**
		PSMB4	0.01	NS	53	antisense
		POGZ	0.25	NS	56	sense
RP11-98D18.3	1:151,735,860-151,741,977	**SNX27**	**0.97**	**0.0003**	**64**	**antisense**
		OAZ3	0.26	NS	intragenic	sense
		TDRKH	**0.95**	**0.001**	**11**	**sense**
		LINGO4	**0.73**	**0.03**	**37**	**sense**
		RORC	**0.8**	**0.02**	**41**	**sense**
		C2CD4D	**0.84**	**0.001**	**75**	**sense**
		THEM4	0.41	NS	108	sense
AL450992.2	1:151,814,353-151,822,861	**SNX27**	**0.9**	**0.004**	**143**	**sense**
		OAZ3	0.12	NS	71	sense
		TDRKH	**0.84**	**0.01**	**51**	**antisense**
		LINGO4	**0.97**	**0.0004**	**37**	**antisense**
		RORC	**0.98**	**0.0002**	**10**	**antisense**
		C2CD4D	**0.99**	**<0.00001**	**1**	**antisense**
		THEM4	0.1	NS	intragenic	antisense
RP5-997D24.3	1:54,751,078-54,753,044	CDCP2	0.12	NS	147	antisense
		CYB5RL	0.33	NS	113	antisense
		MRPL37	0.4	NS	86	sense
		SSBP3	0.36	NS	intragenic	antisense
RP11-783K16.5	11:64,013,436-64,015,689	VEGFB	0.49	NS	11	sense
		FKBP2	0.09	NS	2	sense
		PPP1R14B	0.55	NS	intragenic	antisense
		PLCB3	0.6	NS	3	sense

RP11-290L1.3	12:76,424,274-76,425,158	*PHLDA1*	**0.97**	**0.0002**	**intragenic**	**antisense**
		NAP1L1	0.68	0.04	14	antisense
AL928768.3	14:106,170,301-106,170,939	*IGHE*	0.17	NS	intragenic	antisense
		IGHG4	**0.83**	**0.01**	**3**	**sense**
		IGHG2	**0.89**	**0.005**	**59**	**sense**
		IGHA1	**0.85**	**0.009**	**3**	**sense**
		IGHG1	**0.88**	**0.006**	**32**	**sense**
		IGHG3	**0.82**	**0.01**	**65**	**sense**
AC007182.6	14:76,041,231-76,045,931	*BATF*	**0.78**	**0.02**	**28**	**antisense**
		C14orf1	0.04	NS	76	sense
		TTLL5	0.48	NS	81	sense
RP11-488C13.5	14:77,248,083-77,253,067	*VASH1*	**0.99**	**<0.0001**	**intragenic**	**antisense**
		ANGEL1	0.45	NS	1	sense
CTD-2008A1.1	15:45,118,738-45,119,292	*PATL2*	0.43	NS	161	antisense
		B2M	0.18	NS	108	sense
		TRIM69	0.43	NS	58	sense
GAPB1-AS1	15:50,647,664-50,650,501	*SLC27A2*	0.06	NS	119	sense
		HDC	0.2	NS	89	antisense
		GABPB1	0.13	NS	1	antisense
		USP8	0.04	NS	69	sense
RP11-876N24.5	16:11,032,743-11,033,901	*NUBP1*	0.38	NS	169	sense
		TVP23A	**0.76**	**0.02**	**120**	**antisense**
		CIITA	**0.72**	**0.03**	**14**	**sense**
		DEXI	**0.67**	**0.04**	**intragenic**	**antisense**
		CLEC16A	0.05	NS	6	sense
RP11-304L19.3	16:2,144,831-2,147,027	*GFER*	0.12	NS	110	antisense
		SYNGR3	0.57	NS	105	antisense
		ZNF598	0.19	NS	97	sense

		NPW	0.54	NS	75	antisense
		SLC9A3R2	**0.83**	**0.01**	**55**	**antisense**
		TSC2	0.11	NS	6	antisense
		PKD1	0.28	NS	intragenic	sense
		RAB26	**0.67**	**0.04**	**51**	**antisense**
		TRAF7	0.01	NS	58	antisense
		CASKIN1	0.04	NS	83	sense
		MLST8	0.17	NS	111	antisense
AC002331.1	16:26,596,075-26,606,134	*none*				
IL21R-AS1	16:27,458,991-27,464,714	*IL4R*	0.25	NS	82	antisense
		IL21R	**0.99**	**<0.0001**	**intragenic**	**antisense**
		GTF3C1	0.39	NS	13	sense
AC074212.5	19:46,268,042-46,272,311	*GIPR*	0.04	NS	83	sense
		QPCTL	0.14	NS	73	sense
		FBXO46	0.03	NS	34	antisense
		SIX5	**0.72**	**0.03**	**intragenic**	**antisense**
		DMPK	**0.62**	**0.04**	**1**	**antisense**
		DMWD	**0.63**	**0.04**	**14**	**antisense**
		SYMPK	0	NS	46	antisense
AC096579.7	2:89,156,709-89,165,653	***IGKC***	**0.99**	**<0.0001**	**intragenic**	**antisense**
		IGKV4-1	**0.99**	**<0.0001**	**19**	**antisense**
		IGKV5-2	**0.99**	**<0.0001**	**31**	**antisense**
		IGKV6-21	**0.89**	**0.005**	**294**	**sense**
		IGKV2-40	**0.85**	**0.009**	**473**	**sense**
		IGKV2D-26	**0.99**	**<0.0001**	**859**	**antisense**
		IGKV3D-20	**0.99**	**<0.0001**	**912**	**antisense**
		IGKV1D-13	**0.98**	**0.0002**	**1027**	**antisense**
		IGKV3D-11	**0.98**	**0.0001**	**1046**	**antisense**
LINC00176	20:62,665,697-62,671,315	*DNAJC5*	0.23	NS	98	sense

|		UCKL1	0	NS	78	antisense
		ZNF512B	0.56	NS	intragenic	antisense
		SAMD10	**0.69**	**0.04**	**55**	**antisense**
		PRPF6	0.14	NS	53	sense
		SOX18	0.4	NS	1	antisense
		TCEA2	0.13	NS	23	sense
		OPRL1	**0.85**	**0.01**	**45**	**antisense**
RP11-398A8.3	3:72,084,451-72,149,578	**EIF4E3**	**0.69**	**0.04**	**310**	**sense**
		GPR27	0.49	NS	280	antisense
		PROK2	0.09	NS	250	sense
		RYBP	**0.7**	**0.04**	**274**	**sense**
RP11-213H15.3	5:90,606,838-90,610,219	GPR98	0.26	NS	146	antisense
		ARRDC3	0.16	NS	58	sense

SUPPLMENTAL TABLE 4-4. Th17 related lncRNA genes and neighboring protein-coding genes.

LncRNA	Genomic position	Protein-coding	R^2	P	Distance away in kilobases	Orientation
RP11-87G24.6	17:74,953,105-74,954,304	*MGAT5B*	0.91	0.003	10	antisense
RP11-445P17.8	10:5,307,996-5,313,199	*AKR1CL1*	0.49	NS	80	sense
CTB-58E17.9	17:36,871,581-36,876,525	*CWC25*	0.35	NS	102	sense
		PIP4K2B	0.15	NS	61	sense
		PSMB3	0.24	NS	38	antisense
		C17orf96	0.02	NS	44	sense
		MLLT6	**0.76**	**0.02**	**intragenic**	**antisense**
		CISD3	0.17	NS	15	antisense
		PCGF2	0.02	NS	19	sense
RP11-15A1.3	19:44,395,956-44,405,955	*ZNF283*	0.57	NS	42	sense
		ZNF221	0.31	NS	50	antisense
		ZNF45	**0.89**	**0.004**	**11**	**sense**
		ZNF404	0.16	NS	14	sense
RP11-714G18.1	4:186,291,879-186,312,082	***SNX25***	**0.86**	**0.008**	**intragenic**	**sense**
		LRP2BP	**0.81**	**0.009**	**intragenic**	**antisense**
		ANKRD37	0.06	NS	5	antisense
		UFSP2	0	NS	8	antisense
		C4orf47	0	NS	38	sense
		CCDC110	0.03	NS	54	antisense
RP5-1028K7.2	17:38,673,278-38,683,254	*RARA*	0.07	NS	150	sense
		TOP2A	0.2	NS	99	antisense
		IGFBP4	**0.77**	**0.02**	**60**	**sense**
		TNS4	0.2	NS	16	antisense
		CCR7	0.18	NS	27	antisense
		SMARCE1	0.43	NS	100	antisense

CTD-3094K11.1	15:91,382,954-91,384,608	*BLM*	0.03	NS	122	sense
		FURIN	**0.72**	**0.03**	**27**	**sense**
		MAN2A2	**0.92**	**0.003**	**63**	**antisense**
		FES	**0.7**	**0.04**	**43**	**sense**
		UNC45A	0.24	NS	94	sense
		HDDC3	0.01	NS	92	antisense
		RCCD1	0.42	NS	114	sense
RP11-689B22.2	12:109,022,463-109,035,094	*SART3*	0.09	NS	67	antisense
		ISCU	0.29	NS	59	sense
		TMEM119	**0.69**	**0.03**	**31**	**antisense**
		SELPLG	**0.66**	**0.04**	**intragenic**	**antisense**
		CORO1C	0.09	NS	3	antisense
RP11-254F7.2	2:10,179,219-10,180,790	*TAF1B*	0.55	NS	5	antisense
		GRHL1	0.1	NS	37	antisense
		KLF11	**0.69**	**0.04**	**intragenic**	**antisense**
		CYS1	0.02	NS	16	sense
		RRM2	0.51	NS	82	antisense
		C2orf48	0.01	NS	101	antisense
RP11-126O1.5	18:56,337,712-56,339,109	***ALPK2***	**0.97**	**0.0003**	**41**	**sense**
		MALT1	**0.85**	**0.009**	**intragenic**	**antisense**
AC069363.1	17:34,400,226-34,417,203	***CCL3***	**0.85**	**0.009**	**intragenic**	**antisense**
		CCL4	**0.96**	**0.0006**	**14**	**sense**
		TBC1D3B	0.02	NS	77	antisense
		CCL3L1	**0.8**	**0.01**	**105**	**antisense**
		CCL3L3	**0.73**	**0.03**	**105**	**antisense**
		CCL4L2	**0.98**	**0.0001**	**121**	**sense**
RP11-252E2.1	16:75,142,499-75,144,610	***ZNRF1***	**0.98**	**<0.0001**	**intragenic**	**antisense**
		LDHD	0.55	NS	2	sense

		ZFP1	0.55	NS	38	antisense
		CTRB2	0.16	NS	93	sense

SUPPLEMENTAL TABLE 4-5. Th2 down regulated genes lncRNA genes.
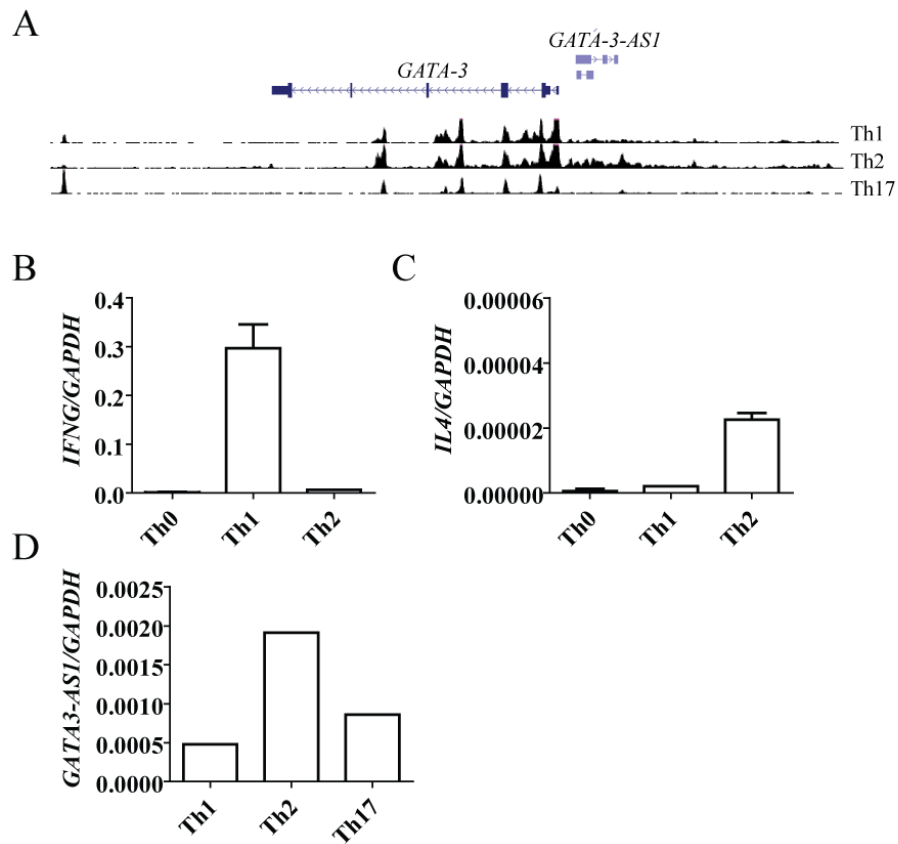
SUPPLEMENTAL FIGURE 4-1. *GATA3-AS1* **demonstrates Th2-specific expression.**
A, GATA3 locus from the UCSC genome browser configuration. *GATA3* gene and *GATA3-AS1* mapped transcripts are shown. DNase I hypersensitivity peaks are shown for human Th1, Th2 and Th17 polarized subsets. *B*, *IFNG* and *C,* IL-4 measurements were normalized relative to *GAPDH* in primary Th1, Th2 or Th17 conditions.

CHAPTER V

DISCUSSION

Herein, I describe my work evaluating the function and biology of the lncRNA, *Tmevpg1* within the context of the Th1 polarization program namely as an enhancer of *Ifng* transcription. In chapter II, I evaluated *Tmevpg1* expression by CD4$^+$ and CD8$^+$ T cells, the major subsets of IFN-γ-producing immune cells, upon various stimulation conditions finding that expression was exclusive to polarized CD4$^+$ Th1 cells. *Tmevpg1* levels correlate strongly with *Ifng* transcript levels and induction of both *Tmevpg1* and *Ifng* transcription in response to Th1 polarization display similar dependence upon Stat4 and T-bet transcription factors. The most significant finding in this work is the discovery that *Tmevpg1* is necessary for efficient *Ifng* expression, an observation that has not been made within the immune system particularly in the context of genes that encode cytokines. A second significant conclusion is that both T-bet and *Tmevpg1* cooperate to induce the high levels of *Ifng* transcripts that define the effector Th1 cell lineage. These studies also demonstrate that *Tmevpg1* is capable of functioning as an RNA molecule *in trans*. As lncRNAs may function either *in cis* or *in trans* to regulate gene expression as diffusible molecules, our evidence suggests that *Tmevpg1* is a *cis*-acting lncRNA regulating the *Ifng* locus just adjacent to the *Tmevpg1* locus.

Shortly after publishing Chapter II in manuscript form, a follow-up report confirming the role of *Tmevpg1* as an enhancer lncRNA capable of regulating expression

of the *Ifng* gene was published by Gomez *et. al.* out of Stanford University (131). In this report, transgenic overexpression of *Tmevpg1* in CD8$^+$ T cells results in rapid IFN-γ responses elicited in response to *Salmonella* pathogenesis highlighting a critical role for this lncRNA in protection against bacterial pathogenesis (131). Moreover, transgenic overexpression of *Tmevpg1* resulted in increased expression of TNF-α, RANTES, and IL-2 in addition to IFN-γ (131). Within this study, a mechanism was proposed for *Tmevpg1* enhancement of *Ifng* transcription. Similar to other described enhancer lncRNAs, it is hypothesized that *Tmevpg1* interacts with WDR5 to recruit the histone methylase complex MLL/Set1 to the *Ifng* locus leading to accumulation of H3K4me3 epigenetic marks across the locus (135). Taken together, my results as well as those presented in this subsequent publication establishes *Tmevpg1* as an enhancer lncRNA contributing to epigenetic remodeling of the *Ifng* locus *in cis* to promote expression of IFN-γ. Second, *Tmevpg1* plays a critical role controlling bacterial pathogenesis, in vivo, by inducing expression of high levels of IFN-γ.

One outstanding question of significant interest to the lncRNA field involves the nature of lncRNA gene regulation, as the lncRNA community has focused effort on identification of new lncRNAs and their functions. As our expertise, historically, includes analysis of regulation of *Ifng* transcription by distal noncoding elements and analysis of epigenetic remodeling, I began to pose similar questions by extending these methods to

111

explore functions of distal noncoding elements and epigenetic remodeling to better understand regulation of the adjacent *Tmevpg1* gene.

As *Tmevpg1* expression is dependent upon Th1 transcriptional regulators, is located adjacent to *Ifng* and is highly expressed in effector Th1 cells, I hypothesized that *Ifng* and *Tmevpg1* share mechanisms of gene regulation. In Chapter III, I examined the epigenetic influence of T-bet demonstrating that establishment and spreading of the positive histone mark H4Ac at the *Tmevpg1* promoter is dependent upon T-bet. This key observation places T-bet in control of epigenetic remodeling between the *Ifng* and *Tmevpg1* genes but also at the *Tmevpg1* transcriptional start site as well as +3.0 kilobases upstream of *Tmevpg1*. The absence of *Tmevpg1* expression in *Tbx21*[-/-] cultures therefore is due, at least in part, to epigenetic restriction, a feature that had not been previously demonstrated for lineage-specific lncRNA expression. As *Tmevpg1* is expressed at high levels in effector Th1 cells, I also examined NK and polyclonal memory T cell populations, which are also capable of rapid production of IFN-γ. Equivalent levels of *Ifng* message and the *Tmevpg1* transcript are observed in unstimulated NK cells as well as stimulated polyclonal T cell memory populations as compared to those levels expressed by cultured effector Th1 cells supporting the view that *Tmevpg1* function may be important for rapid IFN-γ responses by effector-like and memory immune cells.

Within the intergenic distance between the *IFNG and TMEVPG1* genes resides five distal noncoding elements previously shown be nonessential for *IFNG* expression in Th1 cells (Figure 3-4) (114, 115, 136). As these experiments utilized BAC transgenic mice that did not include the *TMEVPG1* locus, I began to assess these genomic elements for enhancer

activity. While four elements demonstrate clear enhancer activity, one was observed to be a repressive element supporting that the *Tmevpg1* gene is surrounded by elements capable of producing dynamic transcriptional regulation. The aforementioned enhancer elements are also sites for transcription factor recruitment. NF-κB and Ets-1 associate at varying levels in effector Th1 cells to these enhancer elements demonstrating that transcriptional regulation of *Tmevpg1* may be induced by these factors within effector Th1 cells.

The degree to which T-bet is capable of regulating *Tmevpg1* is significant as I've shown that T-bet is required for expression, epigenetic remodeling and physically associates with the *Tmevpg1* locus. Data from Chapter II suggest that T-bet and *Tmevpg1* cooperate to induce *Ifng* expression but the role of T-bet in facilitating *Tmevpg1* enhancement of *Ifng* is not well understood. For instance, does *Tmevpg1* localize to the *Ifng* locus?

Previously, it was observed that intrachromosomal looping events across the *Ifng* locus occurred in Th1 cells to promote *Ifng* expression in a T-bet-dependent manner (36, 37). Within this work, the transcription factor CTCF was demonstrated to bind to a conserved site +66 kilobases 3' of *Ifng* and was proposed to establish a boundary for the locus (37). This binding allows for intrachromosomal rearrangement bringing distal CNS elements into close proximity of the *Ifng* promoter. These observations in light of my work are important for two reasons. First, the spatial rearrangement described within the aforementioned body of work is dependent upon T-bet at and spans the *Ifng* locus (137). These findings are consistent with my observation that T-bet regulates the entirety of the Th1 locus. Secondly, the CTCF binding insulator element aligns with the 3' end of the

Tmevpg1 gene. Moving forward with defining the biology of *Tmevpg1*, one area lacking experimental evidence is how *Tmevpg1* localizes or tethers to the *Ifng* locus. As the 3D conformation of the *Ifng* locus data suggests, the *Tmevpg1* gene is brought into close proximity of the *Ifng* promoter allowing for simultaneous co-expression of both *Ifng* and *Tmevpg1* (Figure 5-1). In light of the work establishing a role for *Tmevpg1* physical association with WDR5, perhaps similar to HOTTIP, *Tmevpg1* may function as a molecular scaffold. To further assess this possibility it will be useful to employ a novel technique called ChIRP-seq, chromatin isolation by RNA purification followed by sequencing, to identify not only protein-RNA but also RNA-DNA interactions (137). With this approach, *Tmevpg1*-specific locations across the genome can be determined.
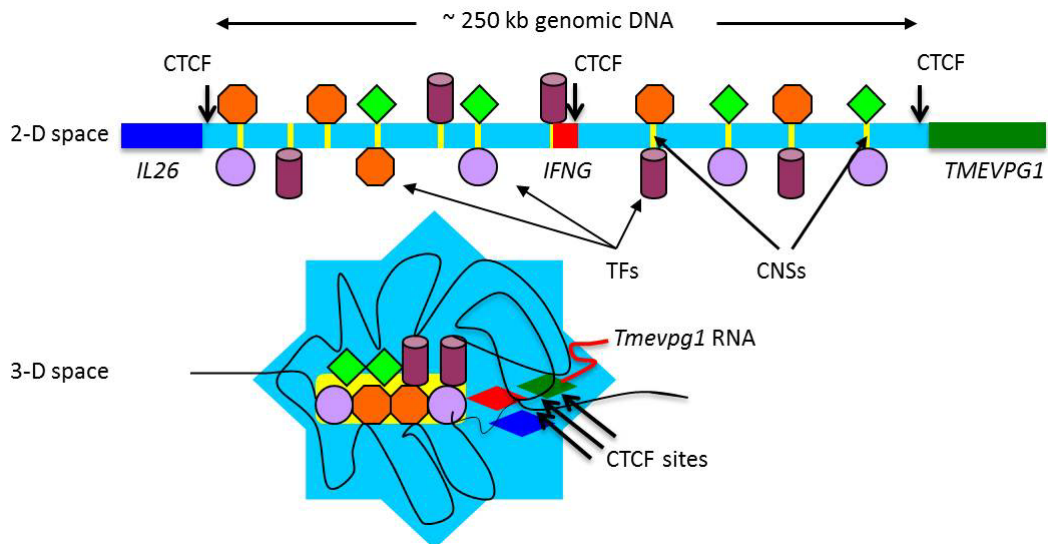
FIGURE 5-1. 3D conformation of the *Ifng* locus brings regulatory elements in close proximity to the promoter. CTCF binding sites are indicated by black arrows. Relative transcription factor binding is representative of interactions with CNSs across the locus. The *TMEVPG1* gene is shown in green and *IFNG* in red (26). Whereas the looping event previously excluded the *TMEVPG1* lncRNA transcript, my work suggests that *TMEVPG1* could co-localize with the *IFNG* locus due to intrachromosomal looping.

In the final data chapter, I describe a pilot RNA-sequencing experiment that our lab has undertaken to identify additional lncRNAs that are involved in the process of Th1, Th2 and Th17 development in human polarized PBMCs. This study identifies nearly 2,800 individual transcripts that are expressed among the three polarization types assessed. We could identify nearly unique lncRNAs elevated in a Th1, Th2 or Th17-specific manner. Genes encoding lineage specific lncRNAs were enriched in the genome adjacent to lineage specific protein-coding genes with immunologic functions. Expression of transcripts detected from RNA-sequencing correlated strongly with RT-PCR results of assessed genes. However, enumerated transcripts and genomic placement among protein-coding genes does not implicate function. Assigning priority for future investigation will be difficult with this significant number of newly discovered lncRNAs; however, beginning with known protein regulators of transcriptional networks such as the master transcription factors will be an optimal place to start. Similar to *TMEVPG1*, we found a lncRNA expressed specifically in Th2 cells named *GATA3-AS1* that is expressed by primary cultured Th2 cells offering feasibility to future experimental investigations. Utilizing a similar strategy that we employed to study *TMEVPG1* including transcription factor dependence, siRNA-mediated suppression and target gene analysis will be essential to assign meaning to the lncRNA networks that may play critical roles in T helper cell differentiation. Our observations are consistent with a recent publication that identified greater than 1,500 individual lncRNA transcripts expressed by 42 examined subsets of murine T cells particular during the process of CD4$^+$ T cell polarization (134). T-bet was confirmed in this study to associate with the *Tmevpg1* gene supporting our findings

described in Chapter III (61). Moreover, in addition to *Tmevpg1,* T-bet was also found to associate with numerous lncRNA-encoding genes within differentiating T cells supporting a greater lncRNA regulatory network driven by master transcriptional regulators.

To date, there is one paradigm that has been established for lncRNA regulation of genes by epigenetic modifications: repressive lncRNAs associate with EZH2 containing methylation complexes that establish and sustain H3K27me3 marks and enhancers associate with MLL/Set1 complexes containing WDR5 promoting H3K4me3 mark formation (38, 99). This paradigm seems to be quite narrow due to the extensive modifications of histones that are formed (e.g. acetylation, phosphorylation, sumolyation) and influence gene expression (9). Additionally, association between lncRNAs and removal of histone marks has not been demonstrated. The question of whether these observations are, in fact, the paradigm or whether it is a consequence of our limited experience remains to be determined. At present the odds are stacked disproportionately as fewer than 1% of lncRNA have been assigned a biological function and a smaller proportion of that number regulate genes at the epigenetic level. I predict that other associations will arise as the breadth of lncRNA function becomes more defined.

If the paradigm for lncRNA association with histone modifying complexes outlined above is correct, it supports the existence of common structural motifs of lncRNA transcripts contributing to lncRNA functional interactions with proteins. While genomic sequence conservation among lncRNAs is certainly not the case, then perhaps RNA structural motifs dictate the limited choices for protein-lncRNA interactions (138). At present, structures of lncRNA are poorly defined reflecting the consensus in the RNA

biology community regarding our understanding of RNA structure (47, 138-140). As RNA is quite dynamic as a molecule, significant strides in identifying domains of protein-lncRNA and further lncRNA-protein-DNA interactions through biochemical and structural methodologies will define the future of this growing field.

The fundamental purpose of my work translates to a broader question of the process of cellular lineage commitment and terminal differentiation. Cell types within compartmentalized biological systems arise from a common precursor; for example, cells of the immune system arise from a common hematopoietic progenitor (141, 142). How a variety of effector cell types develop and maintain terminal differentiation is an incredibly complex question in biology. Cellular differentiation through stochastic niche signals elicits cell type-specific gene expression for example: the hematopoietic progenitors in the thymus begin to develop linearly into immature thymocytes expressing key surface molecules regulated by transcription factor networks or a naïve $CD4^+$ T cells becomes activated and "commits" to an effector polarization program (141, 142). How lineage-specific terminal differentiation is maintained is not known. In light of my observations as well as the work of the greater lncRNAs community, lncRNAs may be instrumental to this process. LncRNAs are expressed in nearly all developmental systems located nearby target protein-coding genes allowing for diffusible accessibility (8). Could perhaps lncRNAs serve to maintain gene expression patterns exhibited by lineage-committed cells, which may, in fact, be regulated by complex lncRNA expression networks?

My work supports the notion that this is at least is a feasible hypothesis: inducible lncRNA gene regulatory networks that are co-expressed with and regulate expression of

protein-coding genes to establish lineage specification. *Tmevpg1* and *Ifng* represent one example. Induction of specific lncRNA networks from progenitor cells to effector Th1 cells that are fully committed to their respective effector phenotype, it becomes likely that lncRNAs contribute to maintenance of lineage-specific gene expression, as is the case for *Tmevpg1*.

In summary, lncRNAs function to regulate target gene transcription through a variety of mechanisms including recruitment or sequestration of histone-modifying complexes and transcription factors, nucleotide homology with mRNA or DNA and as precursors of smaller noncoding RNA. Moreover, these studies pose the question of how many other biological systems are too regulated by lncRNAs. Herein, I describe evidence for the lncRNA *Tmevpg1* contributing to Th1 development. With certainty, future studies will identify additional lncRNAs contributing to all terminally differentiated cell types establishing as well as reinforcing all cell lineage-specific gene programs.

PUBLISHED MANUSCRIPTS

Collier, S.P., Henderson, M.H., and Aune T.M. T-bet dependent regulation of the Th1 locus extends from *Ifng* through the long noncoding RNA *Tmevpg1*. 2014. *Manuscript in preparation.*

Aune TM, Collins PL, **Collier SP**, Henderson MA, Chang S (2013) Epigenetic activation and silencing of the gene that encodes IFN-γ. *Frontiers in Immunology* 4:112 doi:10.3389/fimmu.2013.00112.
http://journal.frontiersin.org/Journal/10.3389/fimmu.2013.00112/full

Collier, S.P., Collins P.L., Williams C.L., Boothby M.R., Aune T.M. 2012. Cutting Edge: Influence of *Tmevpg1*, a Long Intergenic Noncoding RNA, on the Expression of *Ifng* by Th-1 Cells. *The Journal of Immunology* 189: 2084-8.
http://www.jimmunol.org/content/189/5/2084.short

CITATIONS

1.   Consortium, E. P., E. Birney, J. A. Stamatoyannopoulos, A. Dutta, R. Guigo, T. R. Gingeras, E. H. Margulies, Z. Weng, M. Snyder, E. T. Dermitzakis, R. E. Thurman, M. S. Kuehn, C. M. Taylor, S. Neph, C. M. Koch, S. Asthana, A. Malhotra, I. Adzhubei, J. A. Greenbaum, R. M. Andrews, P. Flicek, P. J. Boyle, H. Cao, N. P. Carter, G. K. Clelland, S. Davis, N. Day, P. Dhami, S. C. Dillon, M. O. Dorschner, H. Fiegler, P. G. Giresi, J. Goldy, M. Hawrylycz, A. Haydock, R. Humbert, K. D. James, B. E. Johnson, E. M. Johnson, T. T. Frum, E. R. Rosenzweig, N. Karnani, K. Lee, G. C. Lefebvre, P. A. Navas, F. Neri, S. C. Parker, P. J. Sabo, R. Sandstrom, A. Shafer, D. Vetrie, M. Weaver, S. Wilcox, M. Yu, F. S. Collins, J. Dekker, J. D. Lieb, T. D. Tullius, G. E. Crawford, S. Sunyaev, W. S. Noble, I. Dunham, F. Denoeud, A. Reymond, P. Kapranov, J. Rozowsky, D. Zheng, R. Castelo, A. Frankish, J. Harrow, S. Ghosh, A. Sandelin, I. L. Hofacker, R. Baertsch, D. Keefe, S. Dike, J. Cheng, H. A. Hirsch, E. A. Sekinger, J. Lagarde, J. F. Abril, A. Shahab, C. Flamm, C. Fried, J. Hackermuller, J. Hertel, M. Lindemeyer, K. Missal, A. Tanzer, S. Washietl, J. Korbel, O. Emanuelsson, J. S. Pedersen, N. Holroyd, R. Taylor, D. Swarbreck, N. Matthews, M. C. Dickson, D. J. Thomas, M. T. Weirauch, J. Gilbert, J. Drenkow, I. Bell, X. Zhao, K. G. Srinivasan, W. K. Sung, H. S. Ooi, K. P. Chiu, S. Foissac, T. Alioto, M. Brent, L. Pachter, M. L. Tress, A. Valencia, S. W. Choo, C. Y. Choo, C. Ucla, C. Manzano, C. Wyss, E. Cheung, T. G. Clark, J. B. Brown, M. Ganesh, S. Patel, H. Tammana, J. Chrast, C. N. Henrichsen, C. Kai, J. Kawai, U. Nagalakshmi, J. Wu, Z. Lian, J. Lian, P. Newburger, X. Zhang, P. Bickel, J. S. Mattick, P. Carninci, Y. Hayashizaki, S. Weissman, T. Hubbard, R. M. Myers, J. Rogers, P. F. Stadler, T. M. Lowe, C. L. Wei, Y. Ruan, K. Struhl, M. Gerstein, S. E. Antonarakis, Y. Fu, E. D. Green, U. Karaoz, A. Siepel, J. Taylor, L. A. Liefer, K. A. Wetterstrand, P. J. Good, E. A. Feingold, M. S. Guyer, G. M. Cooper, G. Asimenos, C. N. Dewey, M. Hou, S. Nikolaev, J. I. Montoya-Burgos, A. Loytynoja, S. Whelan, F. Pardi, T. Massingham, H. Huang, N. R. Zhang, I. Holmes, J. C. Mullikin, A. Ureta-Vidal, B. Paten, M. Seringhaus, D. Church, K. Rosenbloom, W. J. Kent, E. A. Stone, N. C. S. Program, C. Baylor College of Medicine Human Genome Sequencing, C. Washington University Genome Sequencing, I. Broad, I. Children's Hospital Oakland Research, S. Batzoglou, N. Goldman, R. C. Hardison, D. Haussler, W. Miller, A. Sidow, N. D. Trinklein, Z. D. Zhang, L. Barrera, R. Stuart, D. C. King, A. Ameur, S. Enroth, M. C. Bieda, J. Kim, A. A. Bhinge, N. Jiang, J. Liu, F. Yao, V. B. Vega, C. W. Lee, P. Ng, A. Shahab, A. Yang, Z. Moqtaderi, Z. Zhu, X. Xu, S. Squazzo, M. J. Oberley, D. Inman, M. A. Singer, T. A. Richmond, K. J. Munn, A. Rada-Iglesias, O. Wallerman, J. Komorowski, J. C. Fowler, P. Couttet, A. W. Bruce, O. M. Dovey, P. D. Ellis, C. F. Langford, D. A. Nix, G. Euskirchen, S. Hartman, A. E. Urban, P. Kraus, S. Van Calcar, N. Heintzman, T. H. Kim, K. Wang, C. Qu, G. Hon, R. Luna, C. K. Glass, M. G. Rosenfeld, S. F. Aldred, S. J. Cooper, A. Halees, J. M. Lin, H. P. Shulha, X. Zhang, M. Xu, J. N. Haidar, Y. Yu, Y. Ruan, V. R. Iyer, R. D. Green, C. Wadelius, P. J. Farnham, B. Ren, R. A. Harte, A. S. Hinrichs, H. Trumbower, H. Clawson, J. Hillman-Jackson, A. S. Zweig, K. Smith, A. Thakkapallayil, G. Barber, R. M. Kuhn, D. Karolchik, L. Armengol, C. P. Bird, P. I. de Bakker, A. D. Kern, N. Lopez-Bigas, J. D. Martin, B. E. Stranger, A. Woodroffe, E. Davydov, A. Dimas, E. Eyras, I. B. Hallgrimsdottir, J. Huppert, M. C. Zody, G. R. Abecasis, X. Estivill, G. G. Bouffard, X. Guan, N. F. Hansen, J. R. Idol, V. V. Maduro, B. Maskeri, J. C. McDowell, M. Park, P. J. Thomas, A. C. Young, R. W. Blakesley, D. M. Muzny, E. Sodergren, D. A. Wheeler, K. C. Worley, H. Jiang, G. M. Weinstock, R. A. Gibbs, T. Graves, R. Fulton, E.

R. Mardis, R. K. Wilson, M. Clamp, J. Cuff, S. Gnerre, D. B. Jaffe, J. L. Chang, K. Lindblad-Toh, E. S. Lander, M. Koriabine, M. Nefedov, K. Osoegawa, Y. Yoshinaga, B. Zhu, and P. J. de Jong. 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 447: 799-816.

2. Harrow, J., F. Denoeud, A. Frankish, A. Reymond, C. K. Chen, J. Chrast, J. Lagarde, J. G. Gilbert, R. Storey, D. Swarbreck, C. Rossier, C. Ucla, T. Hubbard, S. E. Antonarakis, and R. Guigo. 2006. GENCODE: producing a reference annotation for ENCODE. *Genome biology* 7 Suppl 1: S4 1-9.

3. International Human Genome Sequencing, C. 2004. Finishing the euchromatic sequence of the human genome. *Nature* 431: 931-945.

4. Carninci, P., T. Kasukawa, S. Katayama, J. Gough, M. C. Frith, N. Maeda, R. Oyama, T. Ravasi, B. Lenhard, C. Wells, R. Kodzius, K. Shimokawa, V. B. Bajic, S. E. Brenner, S. Batalov, A. R. Forrest, M. Zavolan, M. J. Davis, L. G. Wilming, V. Aidinis, J. E. Allen, A. Ambesi-Impiombato, R. Apweiler, R. N. Aturaliya, T. L. Bailey, M. Bansal, L. Baxter, K. W. Beisel, T. Bersano, H. Bono, A. M. Chalk, K. P. Chiu, V. Choudhary, A. Christoffels, D. R. Clutterbuck, M. L. Crowe, E. Dalla, B. P. Dalrymple, B. de Bono, G. Della Gatta, D. di Bernardo, T. Down, P. Engstrom, M. Fagiolini, G. Faulkner, C. F. Fletcher, T. Fukushima, M. Furuno, S. Futaki, M. Gariboldi, P. Georgii-Hemming, T. R. Gingeras, T. Gojobori, R. E. Green, S. Gustincich, M. Harbers, Y. Hayashi, T. K. Hensch, N. Hirokawa, D. Hill, L. Huminiecki, M. Iacono, K. Ikeo, A. Iwama, T. Ishikawa, M. Jakt, A. Kanapin, M. Katoh, Y. Kawasawa, J. Kelso, H. Kitamura, H. Kitano, G. Kollias, S. P. Krishnan, A. Kruger, S. K. Kummerfeld, I. V. Kurochkin, L. F. Lareau, D. Lazarevic, L. Lipovich, J. Liu, S. Liuni, S. McWilliam, M. Madan Babu, M. Madera, L. Marchionni, H. Matsuda, S. Matsuzawa, H. Miki, F. Mignone, S. Miyake, K. Morris, S. Mottagui-Tabar, N. Mulder, N. Nakano, H. Nakauchi, P. Ng, R. Nilsson, S. Nishiguchi, S. Nishikawa, F. Nori, O. Ohara, Y. Okazaki, V. Orlando, K. C. Pang, W. J. Pavan, G. Pavesi, G. Pesole, N. Petrovsky, S. Piazza, J. Reed, J. F. Reid, B. Z. Ring, M. Ringwald, B. Rost, Y. Ruan, S. L. Salzberg, A. Sandelin, C. Schneider, C. Schonbach, K. Sekiguchi, C. A. Semple, S. Seno, L. Sessa, Y. Sheng, Y. Shibata, H. Shimada, K. Shimada, D. Silva, B. Sinclair, S. Sperling, E. Stupka, K. Sugiura, R. Sultana, Y. Takenaka, K. Taki, K. Tammoja, S. L. Tan, S. Tang, M. S. Taylor, J. Tegner, S. A. Teichmann, H. R. Ueda, E. van Nimwegen, R. Verardo, C. L. Wei, K. Yagi, H. Yamanishi, E. Zabarovsky, S. Zhu, A. Zimmer, W. Hide, C. Bult, S. M. Grimmond, R. D. Teasdale, E. T. Liu, V. Brusic, J. Quackenbush, C. Wahlestedt, J. S. Mattick, D. A. Hume, C. Kai, D. Sasaki, Y. Tomaru, S. Fukuda, M. Kanamori-Katayama, M. Suzuki, J. Aoki, T. Arakawa, J. Iida, K. Imamura, M. Itoh, T. Kato, H. Kawaji, N. Kawagashira, T. Kawashima, M. Kojima, S. Kondo, H. Konno, K. Nakano, N. Ninomiya, T. Nishio, M. Okada, C. Plessy, K. Shibata, T. Shiraki, S. Suzuki, M. Tagami, K. Waki, A. Watahiki, Y. Okamura-Oho, H. Suzuki, J. Kawai, Y. Hayashizaki, F. Consortium, R. G. E. R. Group, and G. Genome Science. 2005. The transcriptional landscape of the mammalian genome. *Science* 309: 1559-1563.

5. Sabin, L. R., M. J. Delas, and G. J. Hannon. 2013. Dogma derailed: the many influences of RNA on the genome. *Molecular cell* 49: 783-794.

6. Mondal, T., M. Rasmussen, G. K. Pandey, A. Isaksson, and C. Kanduri. 2010. Characterization of the RNA content of chromatin. *Genome research* 20: 899-907.

7.   Kapranov, P., J. Cheng, S. Dike, D. A. Nix, R. Duttagupta, A. T. Willingham, P. F. Stadler, J. Hertel, J. Hackermuller, I. L. Hofacker, I. Bell, E. Cheung, J. Drenkow, E. Dumais, S. Patel, G. Helt, M. Ganesh, S. Ghosh, A. Piccolboni, V. Sementchenko, H. Tammana, and T. R. Gingeras. 2007. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* 316: 1484-1488.

8.   Guttman, M., J. Donaghey, B. W. Carey, M. Garber, J. K. Grenier, G. Munson, G. Young, A. B. Lucas, R. Ach, L. Bruhn, X. Yang, I. Amit, A. Meissner, A. Regev, J. L. Rinn, D. E. Root, and E. S. Lander. 2011. lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* 477: 295-300.

9.   Kouzarides, T. 2007. Chromatin modifications and their function. *Cell* 128: 693-705.

10.   Eberharter, A., and P. B. Becker. 2002. Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. *EMBO reports* 3: 224-229.

11.   Kingston, R. E., C. A. Bunker, and A. N. Imbalzano. 1996. Repression and activation by multiprotein complexes that alter chromatin structure. *Genes & development* 10: 905-920.

12.   Luger, K., A. W. Mader, R. K. Richmond, D. F. Sargent, and T. J. Richmond. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* 389: 251-260.

13.   Wilson, C. B., E. Rowell, and M. Sekimata. 2009. Epigenetic control of T-helper-cell differentiation. *Nature reviews. Immunology* 9: 91-105.

14.   Shapiro, I. M., and M. V. Risbud. 2010. Transcriptional profiling of the nucleus pulposus: say yes to notochord. *Arthritis research & therapy* 12: 117.

15.   Shapiro, M. J., and V. S. Shapiro. 2011. Transcriptional repressors, corepressors and chromatin modifying enzymes in T cell development. *Cytokine* 53: 271-281.

16.   Wei, G., L. Wei, J. Zhu, C. Zang, J. Hu-Li, Z. Yao, K. Cui, Y. Kanno, T. Y. Roh, W. T. Watford, D. E. Schones, W. Peng, H. W. Sun, W. E. Paul, J. J. O'Shea, and K. Zhao. 2009. Global mapping of H3K4me3 and H3K27me3 reveals specificity and plasticity in lineage fate determination of differentiating CD4+ T cells. *Immunity* 30: 155-167.

17.   Bernstein, E., E. M. Duncan, O. Masui, J. Gil, E. Heard, and C. D. Allis. 2006. Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. *Molecular and cellular biology* 26: 2560-2569.

18.   Black, J. C., C. Van Rechem, and J. R. Whetstine. 2012. Histone lysine methylation dynamics: establishment, regulation, and biological impact. *Molecular cell* 48: 491-507.

19.   Heintzman, N. D., R. K. Stuart, G. Hon, Y. Fu, C. W. Ching, R. D. Hawkins, L. O. Barrera, S. Van Calcar, C. Qu, K. A. Ching, W. Wang, Z. Weng, R. D. Green, G. E. Crawford, and B. Ren. 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature genetics* 39: 311-318.

20. Bernstein, B. E., E. L. Humphrey, R. L. Erlich, R. Schneider, P. Bouman, J. S. Liu, T. Kouzarides, and S. L. Schreiber. 2002. Methylation of histone H3 Lys 4 in coding regions of active genes. *Proceedings of the National Academy of Sciences of the United States of America* 99: 8695-8700.

21. Sterner, D. E., and S. L. Berger. 2000. Acetylation of histones and transcription-related factors. *Microbiology and molecular biology reviews : MMBR* 64: 435-459.

22. Justin, N., V. De Marco, R. Aasland, and S. J. Gamblin. 2010. Reading, writing and editing methylated lysines on histone tails: new insights from recent structural studies. *Current opinion in structural biology* 20: 730-738.

23. Wysocka, J., T. Swigut, T. A. Milne, Y. Dou, X. Zhang, A. L. Burlingame, R. G. Roeder, A. H. Brivanlou, and C. D. Allis. 2005. WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* 121: 859-872.

24. Wang, Z., C. Zang, J. A. Rosenfeld, D. E. Schones, A. Barski, S. Cuddapah, K. Cui, T. Y. Roh, W. Peng, M. Q. Zhang, and K. Zhao. 2008. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nature genetics* 40: 897-903.

25. Jenuwein, T., and C. D. Allis. 2001. Translating the histone code. *Science* 293: 1074-1080.

26. Aune, T. M., P. L. Collins, S. P. Collier, M. A. Henderson, and S. Chang. 2013. Epigenetic Activation and Silencing of the Gene that Encodes IFN-gamma. *Frontiers in immunology* 4: 112.

27. Aune, T. M., P. L. Collins, and S. Chang. 2009. Epigenetics and T helper 1 differentiation. *Immunology* 126: 299-305.

28. Balasubramani, A., R. Mukasa, R. D. Hatton, and C. T. Weaver. 2010. Regulation of the Ifng locus in the context of T-lineage specification and plasticity. *Immunological reviews* 238: 216-232.

29. Sagai, T., M. Hosoya, Y. Mizushina, M. Tamura, and T. Shiroishi. 2005. Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. *Development* 132: 797-803.

30. Lettice, L. A., S. J. Heaney, L. A. Purdie, L. Li, P. de Beer, B. A. Oostra, D. Goode, G. Elgar, R. E. Hill, and E. de Graaff. 2003. A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Human molecular genetics* 12: 1725-1735.

31. Lettice, L. A., T. Horikoshi, S. J. Heaney, M. J. van Baren, H. C. van der Linde, G. J. Breedveld, M. Joosse, N. Akarsu, B. A. Oostra, N. Endo, M. Shibata, M. Suzuki, E. Takahashi, T. Shinka, Y. Nakahori, D. Ayusawa, K. Nakabayashi, S. W. Scherer, P. Heutink, R. E. Hill, and S. Noji. 2002. Disruption of a long-range cis-acting regulator for Shh causes preaxial polydactyly. *Proceedings of the National Academy of Sciences of the United States of America* 99: 7548-7553.

32. Thurman, R. E., E. Rynes, R. Humbert, J. Vierstra, M. T. Maurano, E. Haugen, N. C. Sheffield, A. B. Stergachis, H. Wang, B. Vernot, K. Garg, S. John, R. Sandstrom, D. Bates, L. Boatman, T. K. Canfield, M. Diegel, D. Dunn, A. K. Ebersol, T. Frum, E. Giste, A. K. Johnson, E. M. Johnson, T. Kutyavin, B. Lajoie, B. K. Lee, K. Lee, D. London, D. Lotakis, S. Neph, F. Neri, E. D. Nguyen, H. Qu, A. P. Reynolds, V. Roach, A. Safi, M. E. Sanchez, A. Sanyal, A. Shafer, J. M. Simon, L. Song, S. Vong, M. Weaver, Y. Yan, Z. Zhang, Z. Zhang, B. Lenhard, M. Tewari, M. O. Dorschner, R. S. Hansen, P. A. Navas, G. Stamatoyannopoulos, V. R. Iyer, J. D. Lieb, S. R. Sunyaev, J. M. Akey, P. J. Sabo, R. Kaul, T. S. Furey, J. Dekker, G. E. Crawford, and J. A. Stamatoyannopoulos. 2012. The accessible chromatin landscape of the human genome. *Nature* 489: 75-82.

33. Boyle, A. P., S. Davis, H. P. Shulha, P. Meltzer, E. H. Margulies, Z. Weng, T. S. Furey, and G. E. Crawford. 2008. High-resolution mapping and characterization of open chromatin across the genome. *Cell* 132: 311-322.

34. Drissen, R., R. J. Palstra, N. Gillemans, E. Splinter, F. Grosveld, S. Philipsen, and W. de Laat. 2004. The active spatial organization of the beta-globin locus requires the transcription factor EKLF. *Genes & development* 18: 2485-2490.

35. Lee, G. R., C. G. Spilianakis, and R. A. Flavell. 2005. Hypersensitive site 7 of the TH2 locus control region is essential for expressing TH2 cytokine genes and for long-range intrachromosomal interactions. *Nature immunology* 6: 42-48.

36. Eivazova, E. R., and T. M. Aune. 2004. Dynamic alterations in the conformation of the Ifng gene region during T helper cell differentiation. *Proceedings of the National Academy of Sciences of the United States of America* 101: 251-256.

37. Sekimata, M., M. Perez-Melgosa, S. A. Miller, A. S. Weinmann, P. J. Sabo, R. Sandstrom, M. O. Dorschner, J. A. Stamatoyannopoulos, and C. B. Wilson. 2009. CCCTC-binding factor and the transcription factor T-bet orchestrate T helper 1 cell-specific structure and function at the interferon-gamma locus. *Immunity* 31: 551-564.

38. Eddy, S. R. 2001. Non-coding RNA genes and the modern RNA world. *Nature reviews. Genetics* 2: 919-929.

39. Miyoshi, K., T. Miyoshi, and H. Siomi. 2010. Many ways to generate microRNA-like small RNAs: non-canonical pathways for microRNA production. *Molecular genetics and genomics : MGG* 284: 95-103.

40. Novina, C. D., and P. A. Sharp. 2004. The RNAi revolution. *Nature* 430: 161-164.

41. Krol, J., I. Loedige, and W. Filipowicz. 2010. The widespread regulation of microRNA biogenesis, function and decay. *Nature reviews. Genetics* 11: 597-610.

42. Shyu, A. B., M. F. Wilkinson, and A. van Hoof. 2008. Messenger RNA regulation: to translate or to degrade. *The EMBO journal* 27: 471-481.

43. McKee, A. E., and P. A. Silver. 2007. Systems perspectives on mRNA processing. *Cell research* 17: 581-590.

44.   Guttman, M., I. Amit, M. Garber, C. French, M. F. Lin, D. Feldser, M. Huarte, O. Zuk, B. W. Carey, J. P. Cassady, M. N. Cabili, R. Jaenisch, T. S. Mikkelsen, T. Jacks, N. Hacohen, B. E. Bernstein, M. Kellis, A. Regev, J. L. Rinn, and E. S. Lander. 2009. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* 458: 223-227.

45.   Guttman, M., M. Garber, J. Z. Levin, J. Donaghey, J. Robinson, X. Adiconis, L. Fan, M. J. Koziol, A. Gnirke, C. Nusbaum, J. L. Rinn, E. S. Lander, and A. Regev. 2010. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nature biotechnology* 28: 503-510.

46.   Mattick, J. S. 2009. The genetic signatures of noncoding RNAs. *PLoS genetics* 5: e1000459.

47.   Mercer, T. R., and J. S. Mattick. 2013. Structure and function of long noncoding RNAs in epigenetic regulation. *Nature structural & molecular biology* 20: 300-307.

48.   Clark, M. B., R. L. Johnston, M. Inostroza-Ponta, A. H. Fox, E. Fortini, P. Moscato, M. E. Dinger, and J. S. Mattick. 2012. Genome-wide analysis of long noncoding RNA stability. *Genome research* 22: 885-898.

49.   Ulitsky, I., and D. P. Bartel. 2013. lincRNAs: genomics, evolution, and mechanisms. *Cell* 154: 26-46.

50.   Nagano, T., and P. Fraser. 2011. No-nonsense functions for long noncoding RNAs. *Cell* 145: 178-181.

51.   Felsenfeld, G., and M. Groudine. 2003. Controlling the double helix. *Nature* 421: 448-453.

52.   Fang, T. C., U. Schaefer, I. Mecklenbrauker, A. Stienen, S. Dewell, M. S. Chen, I. Rioja, V. Parravicini, R. K. Prinjha, R. Chandwani, M. R. MacDonald, K. Lee, C. M. Rice, and A. Tarakhovsky. 2012. Histone H3 lysine 9 di-methylation as an epigenetic signature of the interferon response. *The Journal of experimental medicine* 209: 661-669.

53.   Brown, C. J., A. Ballabio, J. L. Rupert, R. G. Lafreniere, M. Grompe, R. Tonlorenzi, and H. F. Willard. 1991. A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* 349: 38-44.

54.   Brockdorff, N., A. Ashworth, G. F. Kay, P. Cooper, S. Smith, V. M. McCabe, D. P. Norris, G. D. Penny, D. Patel, and S. Rastan. 1991. Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* 351: 329-331.

55.   Jeon, Y., and J. T. Lee. 2011. YY1 tethers Xist RNA to the inactive X nucleation center. *Cell* 146: 119-133.

56.   Lee, J. T. 2009. Lessons from X-chromosome inactivation: long ncRNA as guides and tethers to the epigenome. *Genes & development* 23: 1831-1842.

57.   Tian, D., S. Sun, and J. T. Lee. 2010. The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. *Cell* 143: 390-403.

58.   Rinn, J. L., M. Kertesz, J. K. Wang, S. L. Squazzo, X. Xu, S. A. Brugmann, L. H. Goodnough, J. A. Helms, P. J. Farnham, E. Segal, and H. Y. Chang. 2007. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 129: 1311-1323.

59.   Schorderet, P., and D. Duboule. 2011. Structural and functional differences in the long non-coding RNA hotair in mouse and human. *PLoS genetics* 7: e1002071.

60.   Tsai, M. C., O. Manor, Y. Wan, N. Mosammaparast, J. K. Wang, F. Lan, Y. Shi, E. Segal, and H. Y. Chang. 2010. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* 329: 689-693.

61.   Zhao, J., T. K. Ohsumi, J. T. Kung, Y. Ogawa, D. J. Grau, K. Sarma, J. J. Song, R. E. Kingston, M. Borowsky, and J. T. Lee. 2010. Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Molecular cell* 40: 939-953.

62.   Wang, K. C., Y. W. Yang, B. Liu, A. Sanyal, R. Corces-Zimmerman, Y. Chen, B. R. Lajoie, A. Protacio, R. A. Flynn, R. A. Gupta, J. Wysocka, M. Lei, J. Dekker, J. A. Helms, and H. Y. Chang. 2011. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* 472: 120-124.

63.   Yao, H., K. Brick, Y. Evrard, T. Xiao, R. D. Camerini-Otero, and G. Felsenfeld. 2010. Mediation of CTCF transcriptional insulation by DEAD-box RNA-binding protein p68 and steroid receptor RNA activator SRA. *Genes & development* 24: 2543-2555.

64.   Yang, P. K., and M. I. Kuroda. 2007. Noncoding RNAs and intranuclear positioning in monoallelic gene expression. *Cell* 128: 777-786.

65.   Lai, F., U. A. Orom, M. Cesaroni, M. Beringer, D. J. Taatjes, G. A. Blobel, and R. Shiekhattar. 2013. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature* 494: 497-501.

66.   Li, W., D. Notani, Q. Ma, B. Tanasa, E. Nunez, A. Y. Chen, D. Merkurjev, J. Zhang, K. Ohgi, X. Song, S. Oh, H. S. Kim, C. K. Glass, and M. G. Rosenfeld. 2013. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* 498: 516-520.

67.   Huarte, M., M. Guttman, D. Feldser, M. Garber, M. J. Koziol, D. Kenzelmann-Broz, A. M. Khalil, O. Zuk, I. Amit, M. Rabani, L. D. Attardi, A. Regev, E. S. Lander, T. Jacks, and J. L. Rinn. 2010. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* 142: 409-419.

68.   Hung, T., Y. Wang, M. F. Lin, A. K. Koegel, Y. Kotake, G. D. Grant, H. M. Horlings, N. Shah, C. Umbricht, P. Wang, Y. Wang, B. Kong, A. Langerod, A. L. Borresen-Dale, S. K. Kim, M. van de Vijver, S. Sukumar, M. L. Whitfield, M. Kellis, Y. Xiong, D. J. Wong, and H. Y. Chang. 2011. Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nature genetics* 43: 621-629.

69.   Willingham, A. T., A. P. Orth, S. Batalov, E. C. Peters, B. G. Wen, P. Aza-Blanc, J. B. Hogenesch, and P. G. Schultz. 2005. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* 309: 1570-1573.

70.   Yang, F., X. S. Huo, S. X. Yuan, L. Zhang, W. P. Zhou, F. Wang, and S. H. Sun. 2013. Repression of the long noncoding RNA-LET by histone deacetylase 3 contributes to hypoxia-mediated metastasis. *Molecular cell* 49: 1083-1096.

71.   Bhan, A., I. Hussain, K. I. Ansari, S. Kasiri, A. Bashyal, and S. S. Mandal. 2013. Antisense transcript long noncoding RNA (lncRNA) HOTAIR is transcriptionally induced by estradiol. *Journal of molecular biology* 425: 3707-3722.

72.   Piao, H. L., and L. Ma. 2012. Non-coding RNAs as regulators of mammary development and breast cancer. *Journal of mammary gland biology and neoplasia* 17: 33-42.

73.   Ishii, N., K. Ozaki, H. Sato, H. Mizuno, S. Saito, A. Takahashi, Y. Miyamoto, S. Ikegawa, N. Kamatani, M. Hori, S. Saito, Y. Nakamura, and T. Tanaka. 2006. Identification of a novel non-coding RNA, MIAT, that confers risk of myocardial infarction. *Journal of human genetics* 51: 1087-1099.

74.   Faghihi, M. A., F. Modarresi, A. M. Khalil, D. E. Wood, B. G. Sahagan, T. E. Morgan, C. E. Finch, G. St Laurent, 3rd, P. J. Kenny, and C. Wahlestedt. 2008. Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nature medicine* 14: 723-730.

75.   Pang, K. C., M. E. Dinger, T. R. Mercer, L. Malquori, S. M. Grimmond, W. Chen, and J. S. Mattick. 2009. Genome-wide identification of long noncoding RNAs in CD8+ T cells. *Journal of immunology* 182: 7738-7748.

76.   Peng, X., L. Gralinski, C. D. Armour, M. T. Ferris, M. J. Thomas, S. Proll, B. G. Bradel-Tretheway, M. J. Korth, J. C. Castle, M. C. Biery, H. K. Bouzek, D. R. Haynor, M. B. Frieman, M. Heise, C. K. Raymond, R. S. Baric, and M. G. Katze. 2010. Unique signatures of long noncoding RNA expression in response to virus infection and altered innate immune signaling. *mBio* 1.

77.   Murphy, K. M., and S. L. Reiner. 2002. The lineage decisions of helper T cells. *Nature reviews. Immunology* 2: 933-944.

78.   Olson, M. R., B. E. Russ, P. C. Doherty, and S. J. Turner. 2010. The role of epigenetics in the acquisition and maintenance of effector function in virus-specific CD8 T cells. *IUBMB life* 62: 519-526.

79.   Avni, O., D. Lee, F. Macian, S. J. Szabo, L. H. Glimcher, and A. Rao. 2002. T(H) cell differentiation is accompanied by dynamic changes in histone acetylation of cytokine genes. *Nature immunology* 3: 643-651.

80.   Zhu, J., and W. E. Paul. 2010. Peripheral CD4+ T-cell differentiation regulated by networks of cytokines and transcription factors. *Immunological reviews* 238: 247-262.

81. Hermann-Kleiter, N., and G. Baier. 2010. NFAT pulls the strings during CD4+ T helper cell effector functions. *Blood* 115: 2989-2997.

82. Lupino, E., B. Buccinna, C. Ramondetti, A. Lomartire, G. De Marco, E. Ricotti, P. A. Tovo, M. T. Rinaudo, and M. Piccinini. 2010. In CD28-costimulated human naive CD4+ T cells, I-kappaB kinase controls the expression of cell cycle regulatory proteins via interleukin-2-independent mechanisms. *Immunology* 131: 231-241.

83. Wan, Y. Y., and R. A. Flavell. 2009. How diverse--CD4 effector T cells and their functions. *Journal of molecular cell biology* 1: 20-36.

84. Horiuchi, S., A. Onodera, H. Hosokawa, Y. Watanabe, T. Tanaka, S. Sugano, Y. Suzuki, and T. Nakayama. 2011. Genome-wide analysis reveals unique regulation of transcription of Th2-specific genes by GATA3. *Journal of immunology* 186: 6378-6389.

85. Harrington, L. E., R. D. Hatton, P. R. Mangan, H. Turner, T. L. Murphy, K. M. Murphy, and C. T. Weaver. 2005. Interleukin 17-producing CD4+ effector T cells develop via a lineage distinct from the T helper type 1 and 2 lineages. *Nature immunology* 6: 1123-1132.

86. Balasubramani, A., Y. Shibata, G. E. Crawford, A. S. Baldwin, R. D. Hatton, and C. T. Weaver. 2010. Modular utilization of distal cis-regulatory elements controls Ifng gene expression in T cells activated by distinct stimuli. *Immunity* 33: 35-47.

87. Tato, C. M., G. A. Martins, F. A. High, C. B. DiCioccio, S. L. Reiner, and C. A. Hunter. 2004. Cutting Edge: Innate production of IFN-gamma by NK cells is independent of epigenetic modification of the IFN-gamma promoter. *Journal of immunology* 173: 1514-1517.

88. Zhu, J., D. Jankovic, A. J. Oler, G. Wei, S. Sharma, G. Hu, L. Guo, R. Yagi, H. Yamane, G. Punkosdy, L. Feigenbaum, K. Zhao, and W. E. Paul. 2012. The transcription factor T-bet is induced by multiple pathways and prevents an endogenous Th2 cell program during Th1 cell responses. *Immunity* 37: 660-673.

89. Szabo, S. J., S. T. Kim, G. L. Costa, X. Zhang, C. G. Fathman, and L. H. Glimcher. 2000. A novel transcription factor, T-bet, directs Th1 lineage commitment. *Cell* 100: 655-669.

90. Miller, S. A., and A. S. Weinmann. 2010. Molecular mechanisms by which T-bet regulates T-helper cell commitment. *Immunological reviews* 238: 233-246.

91. Schulz, E. G., L. Mariani, A. Radbruch, and T. Hofer. 2009. Sequential polarization and imprinting of type 1 T helper lymphocytes by interferon-gamma and interleukin-12. *Immunity* 30: 673-683.

92. Afkarian, M., J. R. Sedy, J. Yang, N. G. Jacobson, N. Cereb, S. Y. Yang, T. L. Murphy, and K. M. Murphy. 2002. T-bet is a STAT1-induced regulator of IL-12R expression in naive CD4+ T cells. *Nature immunology* 3: 549-557.

93. Mullen, A. C., F. A. High, A. S. Hutchins, H. W. Lee, A. V. Villarino, D. M. Livingston, A. L. Kung, N. Cereb, T. P. Yao, S. Y. Yang, and S. L. Reiner. 2001. Role of T-bet in commitment of TH1 cells before IL-12-dependent selection. *Science* 292: 1907-1910.

94. O'Shea, J. J., R. Lahesmaa, G. Vahedi, A. Laurence, and Y. Kanno. 2011. Genomic views of STAT function in CD4+ T helper cell differentiation. *Nature reviews. Immunology* 11: 239-250.

95. Thieu, V. T., Q. Yu, H. C. Chang, N. Yeh, E. T. Nguyen, S. Sehra, and M. H. Kaplan. 2008. Signal transducer and activator of transcription 4 is required for the transcription factor T-bet to promote T helper 1 cell-fate determination. *Immunity* 29: 679-690.

96. Hsieh, C. S., S. E. Macatonia, C. S. Tripp, S. F. Wolf, A. O'Garra, and K. M. Murphy. 2008. Pillars article: development of TH1 CD4+ T cells through IL-12 produced by Listeria-induced macrophages. 1993. Science 260(5107): 547-549. *Journal of immunology* 181: 4437-4439.

97. Agarwal, S., and A. Rao. 1998. Modulation of chromatin structure regulates cytokine gene expression during T cell differentiation. *Immunity* 9: 765-775.

98. Zheng, W., and R. A. Flavell. 1997. The transcription factor GATA-3 is necessary and sufficient for Th2 cytokine gene expression in CD4 T cells. *Cell* 89: 587-596.

99. Eissenberg, J. C., and A. Shilatifard. 2010. Histone H3 lysine 4 (H3K4) methylation in development and differentiation. *Developmental biology* 339: 240-249.

100. Ivanov, II, B. S. McKenzie, L. Zhou, C. E. Tadokoro, A. Lepelley, J. J. Lafaille, D. J. Cua, and D. R. Littman. 2006. The orphan nuclear receptor RORgammat directs the differentiation program of proinflammatory IL-17+ T helper cells. *Cell* 126: 1121-1133.

101. Lazarevic, V., X. Chen, J. H. Shim, E. S. Hwang, E. Jang, A. N. Bolm, M. Oukka, V. K. Kuchroo, and L. H. Glimcher. 2011. T-bet represses T(H)17 differentiation by preventing Runx1-mediated activation of the gene encoding RORgammat. *Nature immunology* 12: 96-104.

102. Josefowicz, S. Z., L. F. Lu, and A. Y. Rudensky. 2012. Regulatory T cells: mechanisms of differentiation and function. *Annual review of immunology* 30: 531-564.

103. Chen, J., and X. Liu. 2009. The role of interferon gamma in regulation of CD4+ T-cells and its clinical implications. *Cellular immunology* 254: 85-90.

104. Huang, S., W. Hendriks, A. Althage, S. Hemmi, H. Bluethmann, R. Kamijo, J. Vilcek, R. M. Zinkernagel, and M. Aguet. 1993. Immune response in mice that lack the interferon-gamma receptor. *Science* 259: 1742-1745.

105. McLaren, J. E., and D. P. Ramji. 2009. Interferon gamma: a master regulator of atherosclerosis. *Cytokine & growth factor reviews* 20: 125-135.

106. Godfrey, D. I., S. Stankovic, and A. G. Baxter. 2010. Raising the NKT cell family. *Nature immunology* 11: 197-206.

107. Ansel, K. M., D. U. Lee, and A. Rao. 2003. An epigenetic view of helper T cell differentiation. *Nature immunology* 4: 616-623.

108. Pearce, E. L., A. C. Mullen, G. A. Martins, C. M. Krawczyk, A. S. Hutchins, V. P. Zediak, M. Banica, C. B. DiCioccio, D. A. Gross, C. A. Mao, H. Shen, N. Cereb, S. Y. Yang, T. Lindsten, J. Rossant, C. A. Hunter, and S. L. Reiner. 2003. Control of effector CD8+ T cell function by the transcription factor Eomesodermin. *Science* 302: 1041-1043.

109. Chang, S., P. L. Collins, and T. M. Aune. 2008. T-bet dependent removal of Sin3A-histone deacetylase complexes at the Ifng locus drives Th1 differentiation. *Journal of immunology* 181: 8372-8381.

110. Mullen, A. C., A. S. Hutchins, F. A. High, H. W. Lee, K. J. Sykes, L. A. Chodosh, and S. L. Reiner. 2002. Hlx is induced by and genetically interacts with T-bet to promote heritable T(H)1 gene induction. *Nature immunology* 3: 652-658.

111. Kohu, K., H. Ohmori, W. F. Wong, D. Onda, T. Wakoh, S. Kon, M. Yamashita, T. Nakayama, M. Kubo, and M. Satake. 2009. The Runx3 transcription factor augments Th1 and down-modulates Th2 phenotypes by interacting with and attenuating GATA3. *Journal of immunology* 183: 7817-7824.

112. Chang, S., and T. M. Aune. 2005. Histone hyperacetylated domains across the Ifng gene region in natural killer cells and T cells. *Proceedings of the National Academy of Sciences of the United States of America* 102: 17095-17100.

113. Collins, P. L., M. A. Henderson, and T. M. Aune. 2012. Diverse functions of distal regulatory elements at the IFNG locus. *Journal of immunology* 188: 1726-1733.

114. Collins, P. L., S. Chang, M. Henderson, M. Soutto, G. M. Davis, A. G. McLoed, M. J. Townsend, L. H. Glimcher, D. P. Mortlock, and T. M. Aune. 2010. Distal regions of the human IFNG locus direct cell type-specific expression. *Journal of immunology* 185: 1492-1501.

115. Collins, P. L., M. A. Henderson, and T. M. Aune. 2012. Lineage-specific adjacent IFNG and IL26 genes share a common distal enhancer element. *Genes and immunity* 13: 481-488.

116. Hatton, R. D., L. E. Harrington, R. J. Luther, T. Wakefield, K. M. Janowski, J. R. Oliver, R. L. Lallone, K. M. Murphy, and C. T. Weaver. 2006. A distal conserved sequence element controls Ifng gene expression by T cells and NK cells. *Immunity* 25: 717-729.

117. Balasubramani, A., C. J. Winstead, H. Turner, K. M. Janowski, S. N. Harbour, Y. Shibata, G. E. Crawford, R. D. Hatton, and C. T. Weaver. 2014. Deletion of a conserved cis-element in the Ifng locus highlights the role of acute histone acetylation in modulating inducible gene transcription. *PLoS genetics* 10: e1003969.

118. Shi, M., T. H. Lin, K. C. Appell, and L. J. Berg. 2008. Janus-kinase-3-dependent signals induce chromatin remodeling at the Ifng locus during T helper 1 cell differentiation. *Immunity* 28: 763-773.

119. Rossi, R. L., G. Rossetti, L. Wenandy, S. Curti, A. Ripamonti, R. J. Bonnal, R. S. Birolo, M. Moro, M. C. Crosti, P. Gruarin, S. Maglie, F. Marabita, D. Mascheroni, V. Parente, M. Comelli, E. Trabucchi, R. De Francesco, J. Geginat, S. Abrignani, and M. Pagani. 2011.

Distinct microRNA signatures in human lymphocyte subsets and enforcement of the naive state in CD4+ T cells by the microRNA miR-125b. *Nature immunology* 12: 796-803.

120. Steiner, D. F., M. F. Thomas, J. K. Hu, Z. Yang, J. E. Babiarz, C. D. Allen, M. Matloubian, R. Blelloch, and K. M. Ansel. 2011. MicroRNA-29 regulates T-box transcription factors and interferon-gamma production in helper T cells. *Immunity* 35: 169-181.

121. Vigneau, S., F. Levillayer, H. Crespeau, L. Cattolico, B. Caudron, F. Bihl, C. Robert, M. Brahic, J. Weissenbach, and J. F. Bureau. 2001. Homology between a 173-kb region from mouse chromosome 10, telomeric to the Ifng locus, and human chromosome 12q15. *Genomics* 78: 206-213.

122. Vigneau, S., P. S. Rohrlich, M. Brahic, and J. F. Bureau. 2003. Tmevpg1, a candidate gene for the control of Theiler's virus persistence, could be implicated in the regulation of gamma interferon. *Journal of virology* 77: 5632-5638.

123. Cebrat, M., A. Cebula, A. Laszkiewicz, M. Kasztura, A. Miazek, and P. Kisielow. 2008. Mechanism of lymphocyte-specific inactivation of RAG-2 intragenic promoter of NWC: implications for epigenetic control of RAG locus. *Molecular immunology* 45: 2297-2306.

124. Lai, W., M. Yu, M. N. Huang, F. Okoye, A. D. Keegan, and D. L. Farber. 2011. Transcriptional control of rapid recall by memory CD4 T cells. *Journal of immunology* 187: 133-140.

125. Hadjur, S., L. M. Williams, N. K. Ryan, B. S. Cobb, T. Sexton, P. Fraser, A. G. Fisher, and M. Merkenschlager. 2009. Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature* 460: 410-413.

126. Soutto, M., W. Zhou, and T. M. Aune. 2002. Cutting edge: distal regulatory elements are required to achieve selective expression of IFN-gamma in Th1/Tc1 effector cells. *Journal of immunology* 169: 6664-6667.

127. Grenningloh, R., B. Y. Kang, and I. C. Ho. 2005. Ets-1, a functional cofactor of T-bet, is essential for Th1 inflammatory responses. *The Journal of experimental medicine* 201: 615-626.

128. Nguyen, H. V., E. Mouly, K. Chemin, R. Luinaud, R. Despres, J. P. Fermand, B. Arnulf, and J. C. Bories. 2012. The Ets-1 transcription factor is required for Stat1-mediated T-bet expression and IgG2a class switching in mouse B cells. *Blood* 119: 4174-4181.

129. Lee, C. G., H. K. Kwon, A. Sahoo, W. Hwang, J. S. So, J. S. Hwang, C. S. Chae, G. C. Kim, J. E. Kim, H. S. So, E. S. Hwang, R. Grenningloh, I. C. Ho, and S. H. Im. 2012. Interaction of Ets-1 with HDAC1 represses IL-10 expression in Th1 cells. *Journal of immunology* 188: 2244-2253.

130. Malarkannan, S., J. Regunathan, H. Chu, S. Kutlesa, Y. Chen, H. Zeng, R. Wen, and D. Wang. 2007. Bcl10 plays a divergent role in NK cell-mediated cytotoxicity and cytokine generation. *Journal of immunology* 179: 3752-3762.

131. Gomez, J. A., O. L. Wapinski, Y. W. Yang, J. F. Bureau, S. Gopinath, D. M. Monack, H. Y. Chang, M. Brahic, and K. Kirkegaard. 2013. The NeST long ncRNA controls microbial susceptibility and epigenetic activation of the interferon-gamma locus. *Cell* 152: 743-754.

132. Jenner, R. G., M. J. Townsend, I. Jackson, K. Sun, R. D. Bouwman, R. A. Young, L. H. Glimcher, and G. M. Lord. 2009. The transcription factors T-bet and GATA-3 control alternative pathways of T-cell differentiation through a shared set of target genes. *Proceedings of the National Academy of Sciences of the United States of America* 106: 17876-17881.

133. Collier, S. P., P. L. Collins, C. L. Williams, M. R. Boothby, and T. M. Aune. 2012. Cutting edge: influence of Tmevpg1, a long intergenic noncoding RNA, on the expression of Ifng by Th1 cells. *Journal of immunology* 189: 2084-2088.

134. Hu, G., Q. Tang, S. Sharma, F. Yu, T. M. Escobar, S. A. Muljo, J. Zhu, and K. Zhao. 2013. Expression and regulation of intergenic long noncoding RNAs during T cell development and differentiation. *Nature immunology* 14: 1190-1198.

135. Orom, U. A., T. Derrien, M. Beringer, K. Gumireddy, A. Gardini, G. Bussotti, F. Lai, M. Zytnicki, C. Notredame, Q. Huang, R. Guigo, and R. Shiekhattar. 2010. Long noncoding RNAs with enhancer-like function in human cells. *Cell* 143: 46-58.

136. Collins, P. L., and T. M. Aune. 2010. Keeping one's option open. *Immunity* 32: 581-583.

137. Chu, C., K. Qu, F. L. Zhong, S. E. Artandi, and H. Y. Chang. 2011. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Molecular cell* 44: 667-678.

138. Schroeder, R., A. Barta, and K. Semrad. 2004. Strategies for RNA folding and assembly. *Nature reviews. Molecular cell biology* 5: 908-919.

139. Novikova, I. V., S. P. Hennelly, and K. Y. Sanbonmatsu. 2013. Tackling structures of long noncoding RNAs. *International journal of molecular sciences* 14: 23672-23684.

140. Mercer, T. R., and J. S. Mattick. 2013. Understanding the regulatory and transcriptional complexity of the genome through structure. *Genome research* 23: 1081-1088.

141. Germain, R. N. 2002. T-cell development and the CD4-CD8 lineage decision. *Nature reviews. Immunology* 2: 309-322.

142. Freeman, M. 2000. Feedback control of intercellular signalling in development. *Nature* 408: 313-319.