

ABSTRACT

Title of Dissertation: DYNAMIC ANALYSIS OF CD4+ T CELL
EPIGENETIC STATUS IN CHICKENS
FOLLOWING MDV INFECTION AND
DURING DIFFERENTIATION IN MICE

Yi Ding,

Doctor of Philosophy, 2018

Dissertation directed by: Professor Dr. Jiuzhou Song,
Department of Animal and Avian Sciences

Epigenetic modifications constitute a complicated regulatory network controlling various biological processes from cell development to immune responses. The mechanisms through which CD4+ T cells react to environmental stimuli, including virus intrusion and differentiation signals, represent the fundamental cell biological question of how the external microenvironment influences intrinsic transcriptional networks. This dissertation investigates the epigenetic status changes in CD4+ T cells induced by Marek's disease virus (MDV) infection in chickens and during differentiation in mice. First, a genome-wide gene expression analysis in the immune organs from resistant line 6₃ and susceptible line 7₂ chickens was performed to explore Marek's disease (MD) resistance mechanisms. MDV infection influences

both cytokine-cytokine receptor interaction and cellular development in resistant and susceptible chickens. Second, we examined the epigenetic status of CD4⁺ T cells induced by MDV infection, including chromatin accessibility and chromosome organization. Our results revealed extensive epigenetic modification changes caused by MDV infection. Only resistant line 6₃ chickens could initiate robust adaptive immune responses at the transcription level, and the increase in chromatin accessibility and chromosome reorganization represented by A/B compartment flipping were related to up-regulated genes induced by MDV infection at 10 days post-infection in line 6₃ chickens. Finally, we investigated CD4⁺ T cells plasticity during Th1 helper cell differentiation. We showed “early” (48 hours) CD4⁺ T cells were plastic for cellular reprogramming while “late” (72 hours) cells lost reprogram plasticity and became committed to Th1 cell fate. T-bet, the Th1 cell master regulator, was not the direct determinant of Th1 cell plasticity. Our integrative analysis of multiple “omics” datasets revealed dynamic and genome-wide changes of chromatin accessibility associated with the process of cellular differentiation and commitment. We predicted that several candidate regulators could contribute to cellular plasticity, including Mxi1, JunB, BATF, IRF4, and Hif-1 α . We observed that substantial alterations of chromatin interactions occurred at the IRF4 locus across differentiation time. Conditional deletion of IRF4 in CD4⁺ T cells impacted the expression of T cell activation and differentiation genes, including T-bet, and extended Th1 cell plasticity during the differentiation process. Our findings provided deeper understanding of CD4⁺ T cell commitment and responses toward viral infection.

DYNAMIC ANALYSIS OF CD4+ T CELL EPIGENETIC STATUS IN CHICKENS
FOLLOWING MDV INFECTION AND DURING DIFFERENTIATION IN MICE

by

Yi Ding

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2018

Advisory Committee:

Professor Dr. Jiuzhou Song, Chair

Senior Investigator Dr. Keji Zhao

Professor Dr. Tom Porter

Associate Professor Dr. Stephen Mount

Research Biologist Dr. George Liu

Assistant Professor Dr. Li Ma

© Copyright by
Yi Ding
2018

Acknowledgements

My advisor Dr. Jiuzhou Song, who has passion and motivation toward science, influenced everybody in the lab especially me deeply. Dr. Song consistently encouraged us with a spirit of enterprise and an open mind regarding research and teaching, and a positive attitude towards life. Without his supervision and continuous encouragement, this dissertation would not have been possible.

Besides my advisor, I would like to thank my committee members, Dr. Keji Zhao, Dr. George Liu, Dr. Tom Porter, Dr. Stephen Mount, and Dr. Li Ma for their consistent help and guidance over the years, immeasurable encouragement, insightful comments, and constructive criticism. Without those, I would not be where I am today.

I would give my thanks to all my fellow students and postdocs in Dr. Jiuzhou Song's lab for their feedback, cooperation, and friendship. With the help from including but not limited to Dr. Juan Luo, Dr. Yanghua He, Dr. Fei Tian, Dr. Apratim Mitra, Dr. Jose Carrillo, Fei Zhan, Dr. Yaokun Li, Dr. Hao Bai, Jianan Liu, Wentao Cai in Dr. Jiuzhou Song's lab, I learnt how to become a eligible PhD candidate, devote myself into research, and develop a critical point of view.

My special thanks go to Dr. Keji Zhao, for offering me the research opportunities in his lab in the National Heart, Lung, and Blood Institute (NHLBI), NIH, and leading me working on exciting projects. I also would like to thank every extraordinary researcher in his lab including but not limited to Dr. Kairong Cui, Dr. WaiLim Ku, Dr. Katarzyna Placek, Dr. Gang Ren, and Qingsong Tang. I will remember the stimulating discussions, the hand-by-hand instruction on experimental procedures, the

enlightening troubleshooting suggestions, and all the laughs we have had together for my lifetime. Before my journey in Dr. Keji Zhao's lab, I could not understand that the impossible is possible with every step of hardwork, and the unreachable is reachable if you are willing to get your feet wet.

I would say thank you to my friends for accepting nothing less than excellence from me. Last but not the least, I am so grateful to my parents, who have supported me in overcoming numerous obstacles from the day of my birth and helping me spiritually. I am also thankful to my husband, Dr. Wentao Luan, and other family members who have supported me along the way.

Table of Contents

Acknowledgements.....	ii
Table of Contents.....	iv
List of Tables	vii
List of Figures	viii
Chapter 1 : Literature Review	1
Marek's Disease Resistance in Chickens.....	1
<u>Pathogenesis of Marek's virus infection</u>	1
<u>Classification of MDV</u>	2
<u>Immune responses to Marek's disease virus infection</u>	4
<u>Resistance/Susceptibility to Marek's disease</u>	7
Helper T Cell Diversity and Plasticity	11
<u>Lineage relationships of helper T cell subsets</u>	11
<u>Cytokine environment and transcription factors in fate determination</u>	13
<u>Flexibility of CD4+ T cell subsets</u>	19
<u>Mechanisms for T helper cell plasticity</u>	22
Active Gene Regulatory Elements and Chromatin Structure	27
<u>Gene sequence and chromosome spatial localizations</u>	27
<u>Chromatin hypersensitive sites</u>	29
<u>3D structure and chromatin interaction</u>	35
Rationale and Significance	41
Chapter 2 : Transcriptomic Analysis of Chicken Immune Organs Affected by Marek's Disease Virus Infection	46
Abstract	46
Introduction.....	47
Materials and Methods.....	48
<u>Animals and experimental design</u>	48
<u>RNA preparation and sequencing</u>	49
<u>Differential expression and pathway analysis</u>	49
<u>Principal component analysis and Venn diagram construction</u>	50
<u>Quantitative Real-time RT-PCR</u>	50
Results.....	51
<u>Gene expression profiles of immune organs in MDV challenge experiments</u> ...	51
<u>Genome-wide profiling of the spleen transcriptome at late cytolytic phase</u>	53
<u>Thymus and bursa transcriptome profiling at late cytolytic phase</u>	55
<u>Pathway analysis to reveal networks and biofunctions involved in MDV infection</u>	58
<u>The potential native dissimilarity between lines makes the resistance different</u>	63
<u>Experimental validation of differentially expressed genes</u>	64
Discussion	66

Chapter 3 : Gene Expression, Chromatin Accessibility and Chromosome Organization in CD4+ T Cells Induced by Marek's Disease Virus Infection in Chickens.....	70
Abstract	70
Introduction.....	71
Materials and Methods.....	72
<u>Experimental animals and chicken CD4+ T cell isolation</u>	72
<u>RNA preparation and sequencing</u>	73
<u>CD4+ T cell ATAC-seq</u>	74
<u>CD4+ T cell 3e Hi-C assay</u>	75
<u>Whole genome gene expression analysis</u>	75
<u>Alignment and peak identification of ATAC-seq data</u>	77
<u>3e Hi-C data mapping and analysis</u>	77
Results.....	78
<u>Transcriptional activity changes caused by MDV infection</u>	78
<u>Genes related to immune response were highly enriched in line 6₃ chickens</u>	80
<u>Correlation between chromatin accessibility and differential gene expression</u> ..	83
<u>Gene expression and chromatin organization</u>	87
Discussion	90
Chapter 4 : Dynamic Epigenetic Status and Plasticity during Mouse CD4+ T Cell Differentiation.....	94
Abstract	94
Introduction.....	95
Materials and Methods.....	100
<u>Mice</u>	100
<u>Cell isolation</u>	101
<u>In vitro cell differentiation and FACS sorting</u>	101
<u>Flow cytometric analysis and intracellular staining</u>	102
<u>DNase-seq</u>	102
<u>RNA-Seq</u>	103
<u>ChIP-seq experiments</u>	103
<u>Hyperactive Tnp purification and transposome complex assembly</u>	104
<u>TrAC-looping (Transposition-mediated Analysis of Chromatin Looping)</u>	104
<u>Data analysis</u>	106
<u>Statistics</u>	112
Results.....	113
<u>Th1 like cells can be reprogrammed to express Foxp3 in a time-dependent manner</u>	113
<u>T-bet is not the direct determinant of Th1 cell lineage commitment and plasticity</u>	116
<u>Dynamic chromatin states for the different stages of differentiated Th1 cells</u> . 120	
<u>Identification of potential regulators governing Th1 lineage commitment</u>	120
<u>Deletion of IRF4 results in enhanced Th1 cell reprogramming capability</u>	127
<u>IRF4-regulated transcriptional profiles in differentiated Th1 cells</u>	129
<u>Three-dimensional architecture is conserved in differentiating cells</u>	138
<u>Local epigenetic environments of the IRF4 locus during differentiation</u>	141

Discussion	143
Chapter 5 : Conclusions and Future Directions	147
Appendices.....	152
Appendix I. Top Pathways of Differentially Expressed Genes Between MDV Infected and Uninfected Chickens	152
Appendix II. Mapping Quality of Chicken CD4+ T Cell RNA-Seq Data.....	154
Appendix III. Genome Browser View of Chicken 10 DPI RNA-Seq Data.....	155
Appendix IV. Sequencing Statistics of Chicken CD4+ T Cell ATAC-Seq and Hi-C	156
Appendix V. Mouse Th1 Helper Cell Time-Course Differentiation Experiments	158
Appendix VI. Sequencing Statistics of Mouse RNA-Seq and DNase-Seq	159
Appendix VII. PCA of Mouse DNase-Seq and RNA-Seq Data.....	161
Bibliography	162

List of Tables

Table 2.1 Enriched networks in organs of chicken lines after MDV infection	60
Table 2.2 Top disease and bio-functions in chicken organs after MDV infection	62

List of Figures

Chapter 1

Figure 1.1 CD4+ T Cell Differentiation	13
Figure 1.2 Plastic CD4+ T Cell Model	21
Figure 1.3 DNA Hypersensitive Sites.....	30
Figure 1.4 Connection Between Promoters and Enhancers.....	36

Chapter 2

Figure 2.1 PCA Plots.	53
Figure 2.2 Differentially Expressed Genes in Chicken Spleen Tissues.....	55
Figure 2.3 Differentially Expressed Genes in Thymus Tissue	56
Figure 2.4 Differentially Expressed Genes in Bursa Tissue	56
Figure 2.5 Venn Diagram Indicating Overlap Genes	58
Figure 2.6 Cytokine-cytokine Receptor Interaction Pathway	63
Figure 2.7 Real-time PCR Validation.....	65

Chapter 3

Figure 3.1 Venn Diagram of DEG Lists	79
Figure 3.2 Hierarchical Clustering and Gene Ontology	83
Figure 3.3 Chromatin Accessibility and Gene Expression	86
Figure 3.4 Reproducibility of Hi-C Data	88
Figure 3.5 A/B Compartment and Correlation with Gene Expression	89

Chapter 4

Figure 4.1 Th1 Cell Reprogramming.....	116
Figure 4.2 T-bet Function in Th1 Cell Fate	119
Figure 4.3 Analysis of the DNase-seq and RNA-seq Data	126
Figure 4.4 IRF4 Controls Th1 Differentiation.....	128
Figure 4.5 Analysis of IRF4-deleted Th1 Cells.	133
Figure 4.6 IRF4 Role During Th1 Differentiation.....	138
Figure 4.7 TrAC-looping Principle.....	140
Figure 4.8 Compartment Analysis of Different Time Points.....	141
Figure 4.9 Epigenetic Changes at the IRF4 Locus	142

Chapter 1 : Literature Review

Marek's Disease Resistance in Chickens

Pathogenesis of Marek's virus infection

In 1907 a Hungarian veterinarian Jozsef Marek first reported a generalized polyneuritis in four chickens [1], and half a century later this disease was named after him. After infection, the sciatic nerve and areas of the spinal cord were infiltrated with mononuclear cells [2, 3]. Marek's disease (MD) is a chicken's contagious and neuropathic disease induced by Marek's disease virus (MDV). MDV is a cell-associated alphaherpesvirus that transforms T lymphocytes and triggers mononuclear tissue infiltration such as peripheral nerves, muscle, visceral organs, and skin [4-6].

MDV infection involves several distinct phases: the early cytolytic stage, latency phase, and late cytolytic phase. MDV is pervasive and can survive for a relatively long time in the environment, and almost every chicken worldwide could face MDV challenge from its first day of life [7]. The first phase involves the virus inhalation from the respiratory tract [8]. The phagocytic cells (including macrophages and dendritic cells) in the respiratory tract could become the very early infected cell population [8].

MDV can be detected in spleen, thymus and bursa of Fabricius within 24 hours of contact [9]. MDV confronts the primary target cells (B lymphocytes and activated CD4⁺ T lymphocytes) and reaches replication peak from three to seven days post-infection [10-12]. With the circulation of the infected CD4⁺ T cells, MDV enters the feather-follicle epithelium and gets ready to spread to uninfected birds that come into contact with the infected chicken [13-15]. The decreased expression of major

histocompatibility complex (MHC) class I molecules on the infected cell surface is caused by an MDV early gene and is retained through the lytic cycle. This mechanism helps the virus-infected cells become 'transparent' to the CD8+ T cells cytotoxic immune response [16] and allows MDV to spread to organs and the skin. MDV can produce viral interleukin-8 (vIL-8) [17-20] and attract B cells, T cells, and monocytes for efficient lytic replication [21]. Another protein, viral lipase (vLIP), is also necessary for efficient lytic replication of MDV in chickens [22].

Seven to eight days after infection, MDV progression switches from the cytolytic phase to latency without producing infectious progeny in the activated CD4+ T cells and infected B cells [23-25]. There is not enough information about the latent stage mechanisms, and it is nearly impossible to separate the latently infected cells from transformed cells because the transitions and boundaries between the two states are not clearly defined. The early cytolytic and latent phases could occur in all MDV-infected chickens. Depending on the genetic resistance of the host [26, 27], the reactivation from latency to late cytolytic infection occurs around 2-3 weeks post-infection in susceptible chickens [7]. The formation of multifocal lymphoma [28] is the most significant outcome of MDV infection. It is assumed that only a small subset of latently infected T cells will proliferate and spread to develop neoplasms [8], and the tumors in the bird might derive from few or even one originally transformed cells [29, 30].

Classification of MDV

MDV was thought to be closely associated with Epstein-Barr virus (EBV) for a long time because of its similar biological properties to EBV. It was the discovery of the

double-stranded DNA structure with inverted repeats nucleotide sequences of MDV [17, 18, 31] that made people believe MDV is genetically closely related to human herpesvirus 1 (herpes simplex virus type 1, HSV-1) and human herpesvirus 3 (varicella-zoster virus, VZV) [7]. MDV is one member of the genus *Mardivirus*, and three species, including MDV, GaHV-3, and turkey herpesvirus 1, have been grouped in the same genus. The protein similarity between the three viruses ranges from 50% to 80% [32], and only MDV leads to disease in chickens. The genomes of three viruses contain a unique long (U_L) and a unique short (U_S) segment bracketed by inverted repeats [7]. Homologous genes are largely located in the U_L and U_S segments, whereas genus- and virus-specific genes are predominantly located in the inverted repeat regions. MDV DNA can be randomly integrated into chicken chromosomes, as no recombination hot spots have been reported. However, telomeric repeats are added to the ends of integrated MDV genomes [31].

During its evolution, MDV has adopted effective strategies to recruit and infect chicken cells. In the inverted repeated regions, MDV has a spliced gene that mimics interleukin-8, referred to as vIL-8. MDV vIL8 is a chemoattractant for mononuclear cells, and deletion of vIL-8-encoding open reading frames led to deficient lytic replication of MDV. Another protein MDV glycoprotein C (gC) is one of the major antigens. Overexpression of gC *in vitro* resulted in a substantial decrease in MDV replication.

Malignant transformation by MDV involves the basic leucine zipper protein Meq, which is similar to the oncoproteins Fos and Jun. Wild-type Meq has 339 amino acids and is expressed during the lytic and latent phases of infection [33]. Meq has a similar

structure to the Jun/Fos family and forms a heterodimer with Jun/Fos family members. It is possible that the Meq-c-Jun interaction activates the cathepsin-like protein JTAP-1, JAC, and the heparin-binding epidermal-growth-factor-like growth factor (HB-EGF). As an oncogene, Meq could repress or activate host and viral gene expression by binding to target promoter regions and thus changing the cell phenotypes [34, 35]. By binding to AP-1 sites, the oncogenic pathway was activated [36, 37]. Consistent with this observation, the overexpression of Meq led to downregulation of apoptotic genes *Fas* and *DAP5* *in vitro* [35].

The existence of a viral homolog of telomerase RNA (TR), termed vTR, promotes MDV-induced T-cell lymphomagenesis. MDV vTR has 88% sequence identity to chicken telomerase RNA and might play a role in MDV genome integration by adding telomeric elongations at the viral genome ends.

Immune responses to Marek's disease virus infection

Over the years, our understanding of immune responses to MDV infection has expanded dramatically, although some challenges remain unsolved. In general, the infection with MDV may initiate the host's innate and acquired immune responses. The innate responses involve the rapid activation of macrophages, natural killer (NK) cells, and cytokines, nitric oxide (NO), and other soluble mediators [38]. Acquired immune responses contain mainly the development of antigen-specific antibodies and cytotoxic T cells (CD8⁺ T cells).

MDV can be detected within 48 hours in the spleen, and at 3 days post-infection the first innate immune response starts with the increasing activity of IFN- γ mRNA

expression and NK cell-mediated lysis. The upregulation of IFN- γ mRNA initiates the expression of iNOS and IL-1 β mRNA in macrophages. IL-1 β mRNA upregulation in macrophages may trigger CTL development by increasing IL-2 expression and the production of NO. The increased level of circulating soluble mediator nitric oxide (NO) produced by activated macrophages and spleen cells might assist restraining viral replication [39] and also suppress T cell proliferation which blocks the potential MDV target cells. However, IFN- γ might have a negative effect on chickens since it allows the establishment of latency by initiating IL-8 receptor expression on activated T cells [8]. NK cells are the first firewall of defense against MDV by targeting virus-infected cells, which have changed or decreased MHC-1 expression levels.

In nonspecific immune responses, the functions of cytokines and soluble mediators have been extensively studied during the early cytolytic stage, while the information is not complete for the latency and tumor stages. The infection of MDV in B and T cells may cause the deregulation of interleukins and IFNs expressions during all phases. The mRNA level of IFN- γ is upregulated in splenocytes, and MHC class II expression is increased. Studies in the 2000s suggested both types of IFN (IFN- α and IFN- γ) could inhibit MDV replication *in vitro* and *in vivo* [40] especially in early pathogenesis. Meanwhile, as early as 3 days post-infection (DPI) MDV infection could up-regulate *IFN- γ* , *IL-1 β* , and *IL-8* genes in spleens from infected chickens [41]. IFN- γ itself would stimulate the expression of IL-1 β , which is not only a chemoattractant for T cells but also a stimulus of generation of IL-2 and CTL [42]. The expression of IL-8 receptors on T cells is stimulated by IFN- γ [43]. Since IL-8 is

a chemoattractant for T cells [44], the IL-8 receptors on T cells may be necessary to attract activated T lymphocytes to virus-infected B lymphocytes through the production of vIL-8 [45]. The differential expression of *IFN- γ* , *IL-6*, and *IL-18* genes in susceptible and resistant chicken lines may indicate that these cytokines play an important role in causing lymphomas in susceptible lines and maintaining latency in resistant chickens [46]. Macrophages may also be involved in defending against MDV infection because some reports show the inhibition of MDV replication *in vitro* is much more effective by macrophages from infected chickens compared to those from uninfected chickens [47]. Decreased or lost macrophage activities in infected chickens could increase the MD incidence and reduce the animal's survival time [48]. Adaptive immune responses to MDV usually arise around 5 to 7 DPI and involve the development of MDV-specific antibodies and cytotoxic T lymphocytes [49]. MDV genome could encode 103 viral proteins including structural glycoproteins gB, gE, and gI. Some antibodies are produced to target MDV based on the structural glycoproteins. For example, monospecific antibodies to two gB monomers could detect MDV antigens on the infected cell membranes and inhibit virus penetration of host cells and cell-to-cell spread [50-52]. Maternal antibodies obtained from vaccinated mothers to children have been proven to reduce MD pathogenesis severity when measuring the morbidity, mortality and tumor formation frequencies [53-55]. Infection in both the genetically susceptible and resistant chicken lines could lead to a similar early cytolytic infection. However, the reactivation and tumor formation stages only occur in the susceptible line. This might indicate that CTLs are critical in preventing virus reactivation from the latent stage and also cell transformation.

Immunosuppression is accompanied by CD4 cell transformation in MD-susceptible chicken lines [56]. The most direct consequence of immunosuppression by MDV is the depletion of B and T cells by induction of apoptosis [57]. Immune tolerance in animal bodies is also associated with Treg cell activity. Some researchers suggest that activated Treg cells may be involved in MDV-infected immunosuppression in chickens. The detailed information of Treg cells will be discussed later in the following sections. In general, chicken CD4⁺CD25⁺ T cells have been classified as Treg cells although CD25 is not the exclusive cell marker of Treg cells. However, no Foxp3-like (Foxp3 is the master regulator of Treg cells in humans and mice) genes have been identified in the chicken genome [58]. With the limitation of knowledge on chicken Treg cells, the studies of immunosuppression mechanisms and their relationship with Treg cells are lagging behind. Some researchers try to separate chicken Treg cells by identifying more cell-specific markers. In 2017, TGF- β ⁺ Treg cells were identified as a subgroup of CD4⁺CD25⁺ chicken T cells and showed a higher percentage in MD-susceptible chicken lines [59]. The MDV-induced lymphoma cells in the susceptible lines adopt Treg phenotype and secrete soluble molecules that can inhibit T cell function, which might lead to immunosuppression in chickens [59].

Resistance/Susceptibility to Marek's disease

Following MDV challenge, different chicken strains show various phenotypic responses. These discoveries facilitate the understanding of MD resistance mechanisms. Many studies have used different populations to estimate the heritability

of MD resistance in chickens. However, the results are not consistent among estimations (range from 0.1 to 0.61) [60, 61]. This may be due to the various populations, chicken lines, ages, virus strains and MDV exposure procedures. The first breakthrough was the observed association between MD resistance and inheritance of the B blood group locus [62], which is a known marker for the chicken MHC. The expression levels of class I MHC molecules on the uninfected cell surface vary depending on the haplotypes [63]. Several subsequent experiments further confirmed the responsibility of major histocompatibility complex (MHC) molecule in controlling MD resistance in chickens, although other non-MHC genes might also be involved. For example, B²¹ gives high resistance regardless of genetic background, and B¹⁹ is correlated with susceptibility. B¹, B⁴/B¹³, B⁵, B¹², and B¹⁵ are associated with moderate susceptibility, and B², B⁶, and B¹⁴ provide moderate resistance relying on other factors. The most famous MHC associated resistant and susceptible chickens are the N and P chicken lines, with B²¹ and B¹⁹ haplotypes respectively [64]. When different haplotypes are crossed to generate MHC heterozygotes, their offspring show various MD resistance. B²/B²¹ chickens have the most significant MD resistance (9% MD incidence); chickens with B²/B¹³ and B²/B¹⁴ haplotypes are moderately susceptible (42-43% MD incidence) [65]. MDV infection is thought to induce low levels of CTLs, and the major viral transcription factor ICP4 is found in N chickens but not in the P line [66].

Different from lines N and P, line 6 (MD resistant) and 7 (MD susceptible) inbred chickens developed by the Avian Disease and Oncology Laboratory (ADOL) [4] both carry the B2 haplotype but show significant differences in MD resistance. B

haplotypes cannot also explain the different responses to MDV infection in several outbred and commercial chicken flocks. Hence other genes should also have strong effects on the overall MD resistance levels. Approximately 99% of line 7₂ chickens develop tumors after infection with vv+MDV, however almost none of the line 6₃ chickens become infected with MDV. The susceptible line 7₂ chickens showed higher rates of viral replication and higher viral loads from the early infection stage through their lifetime. In contrast, the resistant line 6₂ chickens had a gradual increase in viral load with a peak around 10 DPI before falling to low levels [67]. The transplantation of thymus from line 7 to line 6 increased the line 6 recipient chickens' susceptibility, while transplantation of thymus from line 6 to line 7 chickens did not improve recipients' resistance [68]. The innate immune response might help explain the early differences between the two lines, and at later stages of MD the line 6 clearance of infection indicates the impact of adaptive immune response. Substantial damage to the immune system during the early cytolytic infection [11] or the inherent variability in immune responses between the two chicken lines [67] might contribute to the line 7's failure of defense.

To unveil the mechanism of non-MHC associated MD resistance, some studies attempted to compare the divergence between the inbred lines 6 and 7 chickens. Lymphocyte surface markers and alloantigens designated as Ly-4 [69], Bu-1 and Th-1 [70] were shown to be partially associated with MD resistance. Genomic mapping revealed the MDV1 resistant gene in a region on chromosome 1 [71]. Subsequent studies using microarrays and RNA-seq also identified several non-MHC genes related to MD resistance and susceptibility [39, 72-79]. Some cytokines and immune-

related genes showed a significant transcriptional difference between line 6 and line 7 chickens before and after MDV infection. For example, *IL-6* and *IL-18* genes were significantly higher in splenocytes of line 7, and higher *IL-2* transcriptional level in line 7 chickens was only observed at 21 dpi. *IFN- γ* had significant expression differences in the two lines during the early infection stage. Systematic analyses to uncover pathways related to MD [80, 81] have also expanded our knowledge of MD resistance mechanisms.

Non-MHC associated MD resistance may play necessary roles in immune responses and is affected by various genetic and environmental factors. The current QTL and genomic mapping experiments serve as the cornerstone of further investigation. Examination of MDV induced transcriptional effects has also discovered a list of important host cytokines and viral genes. Nevertheless, none of these loci alone could adequately explain the MD-resistance and susceptibility mechanism. Besides, environmental factors could also affect the infection outcome, which indicates the involvement of epigenetic modifications during the MD progression. Hence, line 6 and line 7 inbred chickens are precious animal models to study the progression of MD and the interplay between genes and environment influencing the phenotypic outcomes. Research on microRNAs and lincRNAs [82] has revealed their functionality in regulating gene expression and controlling MDV replication. For example, inhibition of *gga-mir-221* and *gga-mir-222* partially soothed the miRNA-mediated suppression. The *linc-GALMD3* in chickens could *cis*-regulate its downstream neighboring *gga-miR-223* expression and trans-regulate gene expression [82]. With the decrease of *linc-GALMD3* expression, MDV replication was repressed,

indicating that *linc-GALMD3* could be a necessary regulator in chicken MD progression. The regulation of miRNAs by DNA methylation to control cancer metastasis [83, 84] provides a new research angle, and some researchers found DNA methylation heterogeneity between the resistant L6₃ and susceptible L7₂ chickens [85]. A better understanding of the MD progression mechanism requires further studies on the epigenetics and the system biology network.

Helper T Cell Diversity and Plasticity

Lineage relationships of helper T cell subsets

CD4⁺ helper T cells play a significant role in adaptive immunity and immune-mediated diseases like autoimmunity, asthma, allergic responses, and tumor immunity [86]. In general, they assist B cells to produce antibodies, maintain and promote CD8 T cell responses, control macrophage function, coordinate adaptive immune responses against various pathogens, and properly suppress/control autoimmunity regarding the magnitude and persistence. When their numbers and functions are decreased or lost, the individual becomes vulnerable to countless infectious attractions and may develop severe disorders. In this part of the review, we discuss the functions and relationships among T helper cells, the development requirements, the plasticity of T helper cells, and the epigenetic regulation of key cytokines and transcription factors.

Under environmental stimuli, i.e., cytokines, naive CD4 T cells could differentiate into specialized subsets depending on particular TCR activation. The lineages of T helper (Th) cells, including Th1, Th2, Th17, iTreg etc., although coming from the

same precursors, have distinct cytokine production profiles and functions (Figure 1.1 [87]). The first discovery of distinct CD4 populations was the Th1/Th2 scenario, with different cytokine production and patterns of cell surface molecules. In general, IFN- γ is the signature cytokine of Th1 cells; IL-2, TNF- α , and lymphotoxin are also the products of many Th1 cells. On the contrary, Th2 cells do not have the ability to produce IFN- γ or lymphotoxin; their signature cytokines are IL-4, IL-5, and IL-13. TNF- α , IL-2, and IL-9 could also be the products, with modest amounts in some Th2 cells. Although researchers are not sure if *in vitro*-generated IFN- γ -producing and IL-4-producing CD4 T cells (i.e., Th1 and Th2) cells truly reflect the biological responses generated under physiologic conditions [86], *in vitro* differentiation could serve as an important model to understand the CD4 T cell functions. Within several years of the discovery of these two lineages, methods were developed to differentiate naive CD4 T cells into IL-4-producing (Th2) cells [88, 89] and IFN- γ -producing (Th1) cells *in vitro* [90]. In the 2000s, Th17 cells [91-94], with the signature cytokines of IL-17A, IL-17F, and IL-22, were discovered. The fourth major population, regulatory T cells (Tregs) [95-97], which can be induced to differentiate *in vitro* from naive CD4 T cells, were discovered also in the 2000s. Induced Treg cells (iTregs) share the same characteristics of regulatory T cells (nTregs), whose lineages are developed in the thymus and distinct from the naive CD4 T cell progenitors.

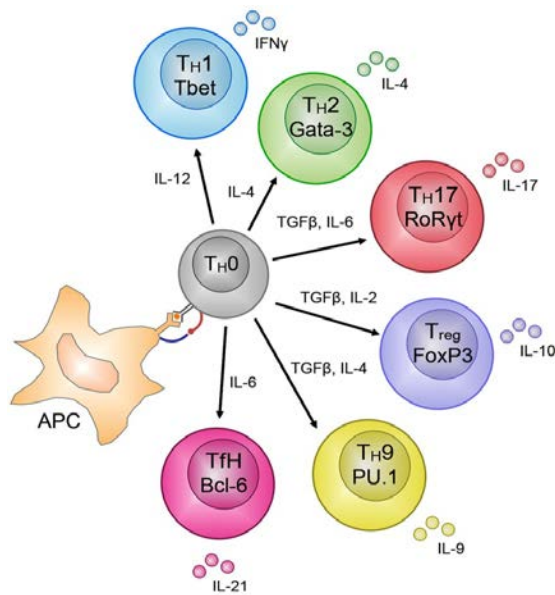


Figure 1.1 CD4⁺ T Cell Differentiation

(Russ BE, Prier JE, Rao S, Turner SJ: **T cell immunity as a tool for studying epigenetic regulation of cellular differentiation**. 2013.)

Th0 (naive) CD4⁺ T cells are able to differentiate into different subsets depending on the soluble molecules (cytokines), e.g., IL-12 for Th1 cells and IL-4 for Th2 cells, during activation by antigen presenting cells (APC). The key transcription factors regulating and maintaining lineage-specific effector functions can be used as the biomarkers to distinguish the corresponding different subsets, e.g., T-bet for Th1 cells, Gata-3 for Th2 cells, and Foxp3 for Treg cells. The corresponding T effector cells can secrete cytokines, e.g., IFN- γ for Th1 cells and IL-10 for Treg cells, to react to pathogen intrusion and mediate autoimmunity.

Cytokine environment and transcription factors in fate determination

The central power of the cytokine environment

The differentiated subsets of CD4 effector/regulatory cells can be determined significantly by their expressed transcription factors *in vivo*. However, *in vitro* the main determinants of the cell state are the set of cytokines present during the T cell receptor (TCR)-mediated activation process [86]. It is an interesting discovery that some key inducing cytokines could also be the major product of the induced differentiated subsets. IL-4-producing CD4 T cells (Th2 cells) could be differentiated

from naive CD4 T cells if IL-4 and IL-2 were present at the time of stimulation by cognate antigens [88, 98-100]. IFN- γ plays an important role in Th1 cells differentiation path [101] and IFN- γ is also the Th1 cell's signature cytokine. Similarly, the iTregs' induction from naive CD4 T cells depends on T cell activation in the company of TGF- β and IL-2 meanwhile Tregs themselves are also good TGF- β producers [86]. *In vitro* Th17 differentiation could be efficiently achieved in the environment of TGF- β and IL-6 or IL-21 and TGF- β [102-106]. Comparably, Th17 cells produce IL-21, and IL-21 can positively promote the Th17 differentiation progression.

Transcription factors and T helper cell fate determination

The master transcription factors for each T helper cell type and the signaling transducer and activator of transcription (STAT) proteins are another two important elements mediating T helper cell fate determination, although through different mechanisms. STATs are regulated by cytokine signaling-mediated posttranslational modifications, including tyrosine and/or serine/threonine phosphorylation [86]. During naive CD4⁺ TCR-mediated activation, the cytokine environment for the cells predominantly determines the CD4 T cell fate by activating the latent signal transducers and activators of transcription (STAT) proteins [107] and triggering the STAT signaling pathway [108-110]. STATs are crucial not only for master regulator's induction but also for cell cytokine's production when cooperating with master regulators. The relationship between master transcription factors and STATs is not simply separated, which we will discuss in detail later.

T-bet and STAT4 for Th1 cells

T-bet, the major transcription factor, and STAT4 are indispensable for inducing Th1-specific gene productions, like IFN- γ , IL-18R1, IL-12R β 2, and Hlx, and Th1 cell differentiation [111]. T-bet can target chromatin remodeling to *IFN- γ* alleles and induce IL-12 receptor β 2 expression [112] by interacting with Hlx [113]. Also, Treg subset cells can express T-bet, and T-bet promotes chemokine receptor CXCR3 on Treg cells. Those T-bet⁺ Treg cells can accumulate at Th1 cell-mediated inflammation and control Th1 responses [114]. T-bet is also found to influence the fate of Th17 cells by directly regulating IL-23R transcription [115] because Th17 cells rely on optimal IL-23 production for survival [116, 117]. T-bet expression is also required for IFN- γ production control in NK cells [118], although not in CD8 cells, which are controlled by another T-box family member Eomesodermin (Eomes) [119-121].

STAT4 is required for Th1 cell development from naive CD4⁺ T cells [122] through induction of Interleukin 12 (IL-12). Upon IL-12 binding, two chains of IL-12 receptors activate JAK2 and TYK2. These two tyrosine kinases phosphorylate STAT4, which translocates into the nucleus to activate its target gene transcription [123]. GATA3 overexpression in developing Th1 cells suppresses Th1 development through downregulation of STAT4, and overexpression of STAT4 could facilitate Th2 $\rightarrow$ Th1 conversion even in cells lacking T-bet [124].

GATA3 and STAT5 for Th2 cells, and ROR γ t and STAT3 for Th17 cells

IL2 and IL4 are necessary for Th2 differentiation [125, 126] *in vitro*, and exogenous or endogenous IL4 promotes GATA3 expression, which is the master transcription factor of Th2 cells [127-129]. In the absence of STAT5 activation, GATA3 alone does not have the capability to induce IL-4 production [125] and *vice versa* [130]. Hence, the orchestra between GATA3 and STAT5 guarantees IL-4 production, and STAT5 activation is crucial for GATA3 expression maintenance in differentiated Th2 cells [131, 132]. Similar experiments were performed on both ROR γ t and STAT3 for Th17 cell differentiation to prove the collaborative effect of these two molecules by discovering that deletion of either led to almost complete loss of IL-17 production [86, 133-135].

Foxp3 for regulatory T cells

Naturally occurring CD4⁺ regulatory T cells, the majority of which express CD25 [interleukin (IL)-2 receptor α -chain] [136-138], play an important role in deleting potentially hazardous self-reactive lymphocytes in the periphery [139]. As a mature T helper cell subpopulation and their functional fluctuation results in spontaneous occurrences of organ-specific autoimmune diseases and other human diseases, including type-1 diabetes, allergies, and cancer [140-144]. In rodents and humans, development of CD4⁺CD25⁺ Treg cells involves the forkhead family transcription factor Foxp3 [145]. *In vitro* retroviral gene transfer of Foxp3 could convert naive T cells to iTreg cells whose phenotype is similar to that of naturally occurring CD4⁺ regulatory T cells [139]. Foxp3 deletion contributes to the lack of functional Treg cells in the periphery and occurrence of systemic autoimmunity [145]. Foxp3

expression is also necessary for maintaining Treg cells' function [146] because decreased expression of Foxp3 in mouse Treg cells could abolish their suppressor function [147]. Although Foxp3 is not necessary for survival of Treg precursors, Foxp3 could amplify and stabilize molecular features of Treg precursor cells (including increased CD25 and CTLA4) and make them permanent [148]. Conditional deletion of *Foxp3* gene in fully differentiated peripheral Treg cells led to loss of 'normal' gene expression pattern and suppressor function. Upon loss of Foxp3, 'ex-Treg' cells were able to produce immune effector cytokines (including IFN- γ , IL-4, and IL-17) [149].

Foxp3 expression level must be constitutively high to induce maximum suppressive capacity [146] regardless of the state of T-cell activation in order to convert human CD4⁺ T cells into potent and stable Treg cells since activation-induced expression of Foxp3 is not sufficient to re-program human Teff into Treg cells [150-152]. And some researchers show that the Foxp3 capacity to reprogram human T cells into Treg cells is dose- and time-dependent [146]. Hence, the simple classification of T cells as Foxp3⁺ or Foxp3⁻ is not adequate to decide their suppressive function, and Foxp3 expression levels should be quantified over the course of T-cell activation before predictions of the T cell functional status can be made, at least in human Treg cells [146].

Multiple master transcription factors are coexpressed in Treg and Th effector cells.

Although as we discussed above that in most cases a transcription factor is dominant in one cell type, the phenomenon of coexpression of these factors has also been

observed. For example, T-bet, GATA3, and ROR γ t are expressed by the Treg cell subsets. T-bet and GATA3 expression level in Treg cells is much lower than that in T effector cells, and Treg cells' CXCR3 expression relies on T-bet existence [153]. Although T-bet deletion, especially in Treg cells, has no correlation with Th1-related autoimmune diseases [153], it remains unclear whether T-bet expression in Treg cells plays a role in suppressing Th1-related immune responses [154, 155].

Previous studies have shown that all T cells including all Treg cells express GATA3 at a basal level; however, GATA3 could be expressed at a relatively high level in particularly activated Treg cells [153, 156, 157] and regulate several “Th2-specific” genes' expression (including Ccr8 and Il1rl1) in Treg cells [158]. Contrary to T-bet expression in Treg cells, deletion of GATA3 specifically in Treg cells could cause mild to severe Th2-like autoimmunity [153, 156, 157, 159]. However, the causality of Th2-like autoimmunity is whether the absence of the “Th2-like” Treg cell subset or the removal of GATA3 basal expression in all Treg cells remains unclear [160].

Similarly, probably due to the shared differentiation program between Th17 and Treg cells, ROR γ t and Foxp3 could be coexpressed, and this stage was initially treated as intermediate before lineage commitment to either Th17 or Treg cell fate [161-163]. However, recent studies might indicate that ROR γ t/Foxp3-coexpressing cells are a distinct subset of Treg cells [164] and Th2-related gut inflammation could be induced by deletion of ROR γ t in Foxp3-expressing cells [165].

Coexpression of multiple master transcription factors is not only observed in Treg cells since T-bet/ROR γ t-coexpressing cells are able to produce both IL-17 and IFN- γ

[166-168] and are called Th1* cells. GATA3/T-bet and GATA3/ROR γ t dual expression is also discovered in mice and humans during type 2 immune responses, including helminth infection and asthmatic inflammation [169-171]. Nevertheless, particular functions of mixed phenotype cells and the synergistic/antagonistic effect of two master regulators on the expression of lineage-specific genes remain unclear and deserve further investigation [160].

Flexibility of CD4⁺ T cell subsets

The first Th1-Th2 paradigm enlightened the community to characterize cytokines, signaling molecules and important effector transcription factors [129], however, this simplistic dichotomy also underestimated the complexity of flexibility in polarization of CD4⁺ cells [172, 173]. The initial idea that effector T helper cell subsets were distinct lineages was based on a series of *in vitro* experiments showing once naive cells had been polarized into Th1 or Th2 cells, these phenotypes would be reserved even when transferred into new polarizing conditions [174] and stable cytokine production had been established relying on a stable transcriptional program [175]. Expression of IL-12R complex β 2 chain was relying on T-bet activation through IFN- γ and STAT1 signaling [112] and conferred IL-12 responsiveness to Th1 cell lineage commitment [176]. IL-4 (Th2 cell cytokine) repressed IL-12 signaling and thus antagonizing Th1 cell differentiation [177]. With more studies on flexibility in polarization for Th1/Th2 cells both in mouse and human cells [178-180], the relative plasticity of early Th1 cells compared to rapidly committed Th2 cells might be caused by early loss of IL-12 receptor in Th2 cells [181]. Continued stimulation with

antigens *in vitro* or chronic diseases *in vivo* might lead to the irreversible CD4⁺ T cells commitment [182].

The interaction between Th1 and Th2 cells may also occur between Th17 and Th1 or Th2 cells [176] since IFN- γ and IL-4 both inhibit IL-17 induction [92]. With the loss of T-bet in Th1 cells or Growth factor independent 1 (Gfi-1) which is induced by IL-4 in Th2 cells, increase of IL-17 expression in CD4⁺ T cells was observed [134, 183, 184]. Similarly, the relationship between Th17 and iTreg cells highly depends on TGF- β level. After *in vitro* ROR γ t and Foxp3 induction by TGF- β , cells did not secrete IL-17 but possessed the dual potential to differentiate into either the Th17 or Treg cell lineage based on the cytokine environment and also the concentration of TGF- β [176]. Thymus-derived CD4⁺CD25⁺Foxp3⁺ regulatory T cells could produce IL-17 after exposure to IL-6 and would differentiate into Th17 cells [185]. The physiological significance of Treg conversion into effector cells might associate with restraining infections with specific viruses, microbial and fungal pathogens *in vivo* [176]. The local microenvironment (the exposure to specific cytokines, cell-cell, and/or cell-microorganism interplay) controls the directions of alternative cell fates that Foxp3⁺ cells could take on [176].

Some researchers proposed a model similar to energy levels of an electron in the atom (Figure 1.2 [186]) to explain the plastic relationship between CD4 subsets. With fewer energy levels, cells are more stable. In most cases, transitions occur from a higher-energy to lower-energy state to make cells more stable. As indicated in Figure 1.2, Th1 phenotype can arise from Th17 cells through the transition from higher to lower energy levels. As suggested by experiments, being polarized in Th17

conditions for 3 weeks long, cells still failed to preserve IL-17 and IL-17F secretion and were converted to Th1 or Th2 cells in the presence of IL-12 or IL-4, respectively [166, 187]. *In vivo* memory CD4⁺CD62L^{low} Th17 cells behaved more resistant to conversion stimulus [187], and this might suggest differences between *in vitro*- and *in vivo*-generated Th17 cells. Naive CD4⁺ T cells are most unstable and could differentiate to one of the known subsets after TCR stimulation and specific cytokine environment.

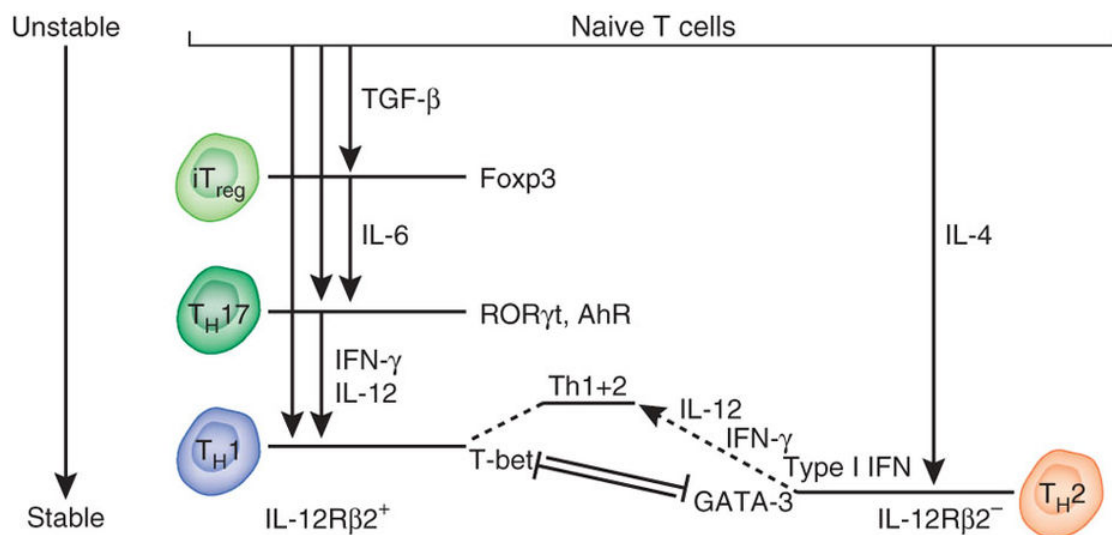


Figure 1.2 Plastic CD4⁺ T Cell Model

(Murphy KM, Stockinger B: **Effector T cell plasticity: flexibility in the face of changing circumstances**. *Nature immunology* 2010, **11**(8):674-680.)

This model describes transitions of plastic CD4⁺ T cells and suggests a potential hierarchy of stability similar to energy levels of an electron in the atom. Unstable naive CD4⁺ T cells can be differentiated into various subsets after TCR stimulation and specific conditions. iTreg, Th17, and Th1 cells maintain IL-12 receptor's expression and thus remaining responsive to IL-12 signals, whereas Th2 cells lose IL-12R expression in early differentiation path. An immediate state of Th1+2 could be induced from Th2 cells under influence of type I interferons, IFN- γ , and IL-12.

Dynamic expression of several transcription factors has been implicated in modulating cell plasticity and development. Both Treg and Th17 cells could produce

IFN- γ through T-bet induction, and losing Foxp3 expression might cause Treg cells to acquire Th2 or Th17 phenotype [154, 161, 166, 188-191]. Interestingly, T-bet and GATA3 may assist Foxp3 in suppressing ROR γ t in a redundant manner to maintain immune tolerance of Treg cells [153]. Th17 cell differentiation requires the existence of IRF4 both *in vitro* and *in vivo* [192] since its absence led to decreased ROR γ t expression, up-regulated Foxp3 expression, and no IL-17 production [176]. Runx1 is another important transcription factor that forms a complex with ROR γ t and participates in Th17 cell differentiation [134]. An interaction between Runx1 and Foxp3 is essential for Treg cells' suppressive ability [193] and Runx-CBF β complexes control Foxp3 expression in regulatory T cells [194]. An iterative system approach has been applied to identify Th17 regulatory networks, and multiple Th17 regulators, including Fosl2, have been associated with cellular plasticity [195]. Conditional deletion of *Fosl2* in CD4⁺ T cells enabled IFN γ production in Th17 and Th2 *in vitro* cultures, especially when subsequently exposing Th17 cells to Th1-skewing conditions [195].

Mechanisms for T helper cell plasticity

One core question to be asked is what elements could decide the relative stability between CD4⁺ subsets. As discussed above, microenvironments act as primary factors influencing cell differentiation and also stability. At the same time, the network of interactions between transcription factors, clonality or heterogeneity of cells (including *in vivo* versus *in vitro* differences), and epigenetic modifications (involve the chromatin modifications, small noncoding RNAs etc.) are also important factors.

Transcriptional regulatory network

Although we often observe the expression of the same transcription factors in various cell types at different developmental stages, the function of transcription factors depends on the cell type and stage. One canonical instance is GATA3, whose expression is detected in T cell development. When GATA3 is expressed at high levels, the Th2 cell fate is determined; when it is expressed at low levels, it influences other cell types besides Th2. Meanwhile, many transcription factors may function as upstream, downstream, or cofactors of master regulators to form a regulatory network. Studies of Th17 cell differentiation network [195, 196] indicate that besides ROR γ t, STAT3, IRF4, BATF, and Hif1 α , which were previously known involving Th17 cell differentiation, other transcription factors including Sp4, ETV6, EGR2, FOSL2, and IRF8 are also important connecting nodes regulating IL-17 expression positively or negatively in the network landscape. Small influences either from the extracellular stimuli or cell-intrinsic fluctuation on core regulatory nodes and pathways could alter the balance among master regulators, thus affecting cell differentiation and stability in a cell type- and stage-specific manner.

Cell clonality, heterogeneity and *in vivo* versus *in vitro* discrepancy

The mystery of CD4⁺ T cell plasticity or stability is further complicated by concerns of clonality or heterogeneity in the cell population. As discussed above, early Th1 cells remain reversible to a Th2 phenotype for a longer time period compared to “resistant” Th2 cells into Th1 phenotype [181], and this could be the outcome of

functional IL-4 signaling in Th1 cells or uncommitted cells present in the heterogeneous Th1 population [181]. *In vitro* polarization would never generate 100% identical cell populations and we should be cautious when we try to describe a population as plastic and assign properties to a ‘subset’ in the mixed population which is not represented by any individual cell [186]. If applicable, discriminating individual cell plasticity from heterogeneity should utilize the current single-cell technology (including scRNA-Seq, single cell imaging etc.) to remove artifacts coming from the not carefully examined mixed population. Also, the difference between differentiation conditions *in vitro* and *in vivo* may contribute substantially to cell polarization and heterogeneity, and *in vitro* cultured cells might not behave identically as *in vivo* generated cells as indicated by Th17 and Treg cells examples. *In vivo* generated Th17 cells stay more stable compared to *in vitro* counterparts [187]. In most cases, Treg cells could maintain their suppressive capacity and phenotype in a Th1-inducing environment while under highly inflammatory conditions leading to host death, Treg cells would lose their conventional phenotype and raise an effector profile with IFN- γ production [154]. Hence, caution should be taken when we interpret our data from the heterogeneous cell population derived specifically *in vitro* and/or *in vivo* conditions.

Chromatin and epigenetic views

Epigenetics describes heritable alterations in phenotype or gene expression without changes in DNA sequence. Histone protein N-terminal tails can be covalently modified through acetylation, methylation, and phosphorylation [197], and those

changes are related with activating or repressing gene transcription by relaxing or condensing the chromatin structure [198]. Differentiation of T helper cells is coupled to epigenetic changes in the chromatin structure, histone and DNA modifications, and small noncoding RNAs [199-201]. With broad usage of chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-Seq), genome-wide examination of histone modifications and DNA methylation in CD4⁺ T cell subsets provides more insights into mechanisms of T helper cell plasticity. Trimethylation of histone H3 lysine 4 (H3K4me3) is highly enriched at active promoters/enhancers near gene transcription start sites and is a permissive mark, whereas trimethylation of histone H3 lysine 27 (H3K27me3) is associated with downregulation or inactivity of nearby genes. Bivalent modifications (both permissive H3K4me3 and repressive H3K27me3 marks) at the same genomic regions may suggest poised genes are ready for switching to other phenotypes in response to cytokine environmental alterations [176]. Bivalent modifications are usually discovered in promoters of many master regulator genes, involving *Tbx21*, *Gata3*, *Rorc*, and *Bcl6* etc. [202, 203]. Bivalent marks at *Tbx21* (encoding T-bet) and *Gata3* genes in Th2 and Th1 cells respectively may help explain plasticity between Th1 and Th2 cells under certain inflammatory conditions [204]. *Tbx21*, *Gata3*, and *Rorc* are bivalently modified in Treg cells, and *in vitro*-generated Th17 cells show bivalent status at *Tbx21* and *Gata3* loci. Such ‘uncertainty’ of master transcription factors might associate with unstable status of Th17 and Treg cells compared to Th1/Th2 cells (Figure 1.2). Such bivalent modifications may contribute to occurrence of mixed phenotype cells co-expressing multiple master regulators and bring lineage switch

when cells encounter transient extracellular signals during and/or after differentiation. Perturbation of the balance between master regulators may result in cell heterogeneity and flexibility, as observed in Treg cells [153].

DNA methylation and microRNAs also provide tools to control stability of differentiated phenotypes [186, 205-207], especially in Treg cells. Thymus-derived nTreg cells and *in vitro* derived iTreg cells show different capabilities for reprogramming and also Foxp3 stability when subjected to other cell-types polarizing conditions. When nTreg cells were cultured in Th1 or Th17 cell-differentiation medium, a substantial fraction of IFN γ +Foxp3+ or IL-17+Foxp3+ cells occurred without decreased Foxp3 expression [162, 185, 202]. However, iTreg cells could rapidly upregulate IL-17 expression and lose Foxp3 expression when cultured in Th17 cell-differentiation medium [162]. A highly conserved CpG-rich element upstream of *Foxp3* exon-1 locus [Treg cell-specific demethylated region (TSDR)] [208-212] between nTreg and iTreg cells might explain phenotypic variations. This region is methylated in iTreg cells and activated human conventional T cells while fully demethylated in nTreg cells. By inhibition of DNA methylation in iTreg cells, Foxp3 expression becomes stable [213]. It seems evident that DNA methylation has a general role in controlling T helper cell differentiation and plasticity rather than only the *Foxp3* locus. Another tool could be the microRNAs. The deficiency of miR-146a in Treg cells led to loss of immunological tolerance and suppressive function of Treg [214], and miR-155 plays an important role in Th1, Th2 and Treg cell development [215-217]. MiR-326 promotes Th17 differentiation by targeting Ets-1, a negative regulator of Th17 differentiation [218]. MicroRNA miR-29 manages T-box

transcription factors (T-bet and Eomes) and IFN γ production in Th1 differentiation [219, 220]. Those few examples demonstrate significant roles of miRNAs in T cell differentiation, and it is reasonable to assume miRNAs must be involved in T cell plasticity and contribute to regulatory networks among crucial transcription factors.

Active Gene Regulatory Elements and Chromatin Structure

Gene sequence and chromosome spatial localizations

Although various cell types of an organism share identical DNA information, their genomes have diverse structural and organizational alterations during differentiation and replication and such fluctuation could influence gene expression and cellular stability [221]. Two-meter long DNA in mammalian cells is packed with histones and folded through different mechanisms into chromosomes, which occupy specific locations in the nucleus. Chromosomes are highly diverse in lengths, and are intricately folded in three dimensions [222]. Multiple levels of folding produce extensive contacts between various genomic regions [223]. During cell cycle replication and differentiation, structural re-organization of chromosomes will happen and affect gene expression and cellular activities, which in turn alter the interaction between chromosomal regions [223].

Half a century ago, a gene locus on a single chromosome was treated as the main determinant of gene activity. Due to the shortage of profiling techniques, regulatory attributes were characterized as a one-dimensional sequence. Advances in high-throughput and computational technologies have resulted in genome sequencing of thousands of organisms across all three domains of life [224]. Cost-effective sequencing techniques have exceeded manual annotation, because input data have

increased by orders of magnitude [225]. Next-Generation Sequencing methods provide researchers the possibility to sequence up to one gigabase in a single run [226]. Techniques including chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) have yielded large datasets on transcription factor binding and post-translational histone modifications in different cell types. The versatile data obtained from NGS are now analyzed to provide insight into gene expression and loci structure [227]. Keeping pace with various high-throughput technologies, the flood of biological data, and increasing information about component interactions, multidimensional genome studies have emerged [228]. One-dimension study focuses on DNA sequence information, and two-dimension study reveals cellular components and interaction network domains. Three-dimension studies emphasize the importance of spatial localization of a gene to genome function [229], and four-dimension investigates genome-scale sequence changes during evolution [228]. Generating two-dimensional reconstructions allows information in one-dimension to be correlated with another biological context [230]. In most cases, two-dimension information is helpful for evaluating one-dimension function. Two-dimension information can help analyze and explain experimental results and sometimes can spark new biological discovery. Also, reconstructions lay the foundation for computation of phenotypic traits and the field of systems biology [231]. Studies of chromosome three-dimensional structures have emerged as researchers agree that both the genomic location and the spatial localization of a gene are functionally critical. With the growing number of fully sequenced genomes and high efficient NGS technologies, more and more researchers realize that genomes are evolutionally altering. Some

researchers have applied high-accuracy mass spectroscopy and NGS methods to re-sequence small and substantial portion of *E. coli* genomes [232, 233]. They discovered that those genomes have adaptive evolution and the sequence alteration information can shed light on mechanisms and functions of genome adaptive evolutionary changes.

Chromatin hypersensitive sites

The structure of chromatin is not static nor uniform, as some chromatin regions have to be “open” for transcriptional regulations by transcription factors while “closed DNA” would stay compact stopping transcription from happening (Figure 1.3 [234]). It is always a key question to identify locations of regulatory elements in the genome to understand the regulatory machinery. Regulatory DNA regions coordinate cell-specific gene expression [235], and in those regions nucleosomes can be modified to mediate accessibility of chromatin-binding proteins to DNA during transcription [236]. Those chromatin accessible regions involve a wide range of regulatory features including promoters, enhancers, silencers, insulators, and boundary elements [234]. DNase I hypersensitive sites (DHSs) give important information about the existence of regulatory elements and the chromatin state in cells [235, 237, 238]. Since the discovery of DHSs 30 years ago, they have been widely used as markers of regulatory DNA regions, and techniques based on DHS data have revealed thousands of regulatory elements in various types of cells. Thanks to the explosive development of Next-Generation Sequencing, DNase sequencing (DNase-seq) and assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq) have

become the universal methodology to identify DNA accessibility on a genome-wide scale [237-240].

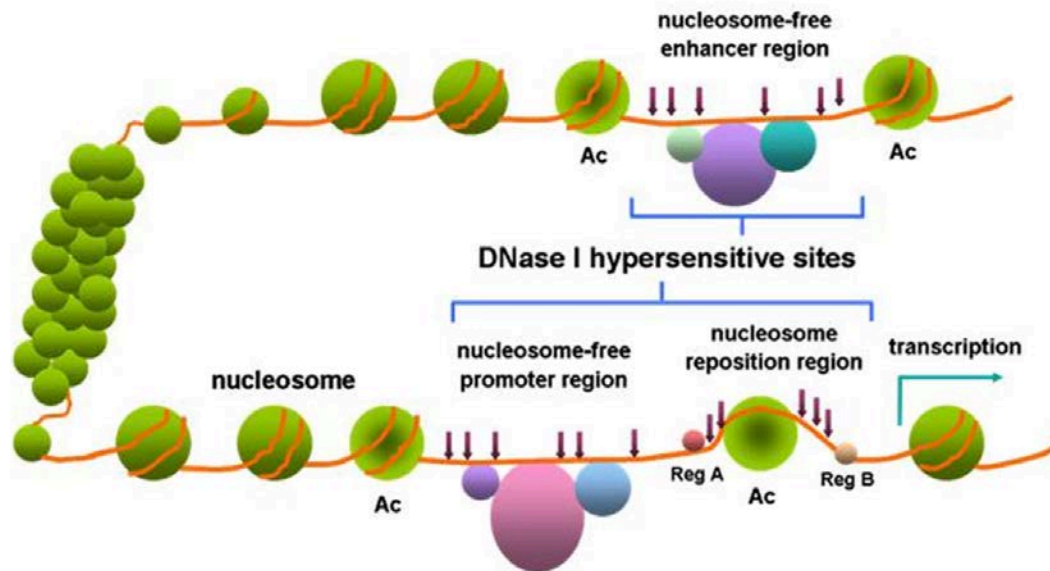


Figure 1.3 DNA Hypersensitive Sites

(Wang Y-M, Zhou P, Wang L-Y, Li Z-H, Zhang Y-N, Zhang Y-X: **Correlation between DNase I hypersensitive site distribution and gene expression in HeLa S3 cells.** *PloS one* 2012, 7(8):e42414.)

This figure demonstrates DNA hypersensitive sites (DHS) in the chromatin.

The single-cell DNase-Seq method has been available through these years for single-cell or limited-cell inputs [241]. Jin et al. utilized the scDNase-Seq to map DHSs at the single-cell level in several different cell types [241], which indicate the cell-to-cell variation within the population. The conventional DNase-seq technique requires millions of cells, which limits applications to cell lines or abundant primary cells [237]. The key idea of scDNase-Seq is to include a bacterial circular carrier DNA to minimize sample DNA loss during experiments [236]. Also, DNase-I digestion optimization and PCR conditions to amplify desired small DNA fragments are other

major changes compared to the bulk population method, though general procedures are similar to the conventional DNase-seq. The first step of scDNase-seq is to permeabilize cellular and nuclear membranes to allow DNase I enzyme entry to the nucleus. Genomic regions unbound by histones, transcription factors and/or other proteins are “open” and more accessible to the DNase I enzyme. This step will generate short DNA fragments from DHSs, and those small fragments are purified and sequenced using next-generation techniques. Data processing is initiated by alignment of short reads to the corresponding genome, and only unique and nonredundant reads are retained for further analysis. In conventional DNase-seq data, DHS peaks are defined as the enrichment of multiple peaks at specific genomic locations, which can be captured by existing peak-calling software like MACS. However, some analysis modifications are needed when applying to fewer than ten cells, since not enough reads would be identified at each DHS.

DNase-seq was applied to nuclei isolated from *C. elegans* embryos and L1 arrest larvae to profile high-resolution transcription factor binding maps [242]. DNase-seq along with RNA-seq assays helped identify bone morphogenetic protein 4 (BMP4) signaling in breast cancer [243]. Applications of DNase-seq are not limited to hypersensitive sites since DNase I could be used to precisely map translational positions of *in vivo* nucleosomes [244]. DNase-seq also has the potential to predict transcription factor occupancy via footprints [245] and reveal the regulatory interaction between specific transcription factors and genes [246]. Conventional DHSs usually span hundreds to thousands of nucleotides and their lengths are much longer than true TF binding motifs, which makes it difficult to infer exact TF binding

loci in real experiments. However, when sequencing depths increase to a very high level, the “footprints” of TF binding could be precisely observed. The ENCODE project [247, 248] described a metric sensitive to sharp drops in the DNase signal within a DHS as the Footprint Occupancy Score (FOS) [249]. This enlightened the pathway to reconstruct TF-TF interaction networks in different cell types and organisms [250]. Footprint detection requires dedicated algorithms [246], and it is necessary to evaluate and modify algorithm parameters based on specific experimental conditions.

Another common method, ATAC-seq, is very useful to identify open chromatin regions and infer DNA-binding proteins using ‘footprinting’ [238]. This technique involves prokaryotic hyperactive Tn5 transposase [251, 252] to tag regulatory regions. In ATAC-seq Tn5 transposase integrates its adaptor into regulatory regions of chromatin, while in less accessible regions such transposition is less favorable. With Tn5 insertions, PCR amplifiable DNA fragments are preferentially generated at open chromatin regions [238]. Similar to DNase-seq, ATAC-seq simultaneously interrogates transcription factor occupancy and chromatin accessibility genome-wide by analyzing insertion positions and insertion lengths generated during transposition reactions. The single-cell assay for transposase-accessible chromatin (scATAC-seq) technique captures and assays individual cells using a programmable microfluidics platform with optimization parameters [253]. Transposition and PCR were processed on the integrated fluidics circuit then libraries were pooled including cell-identifying barcoded primers. DNA accessibility maps from individual lymphoblastoid cells obtained from scATAC-seq data shared similar characteristics of bulk ATAC-seq

data [253]. ATAC-seq was used to identify the repertoire of active promoters and enhancers during normal development and oncogenesis *in vivo* [254]. Some researchers integrated ATAC-seq and RNA-seq to characterize human alpha cells and beta cells and identified endocrine cell type-specific *cis*-regulatory elements [255]. ATAC-seq was also utilized to identify chromatin accessibility changes in the model organism *Caenorhabditis elegans* from embryogenesis to adulthood [256]. EOR-1 was discovered as a unique transcription factor regulating chromatin dynamics during *C. elegans* development [256]. In the immune system, naive CD8⁺ T cells could differentiate into effector cells in response to acute infection and maintain a long-lived pool of memory cells after infection clearance. Genome-wide chromatin accessibility and gene expression of CD8⁺ T cells reacting to acute and chronic viral infection were characterized using ATAC-seq and RNA-seq techniques [257]. Those differentially accessible regions were enriched for transcription factor binding sites including E2A, BATF, IRF4, T-bet, and TCF1, which were known for regulating T cell development and differentiation [257]. Several candidate biomarkers associated with aging immunodeficiency were identified when system immunology approaches and chromatin accessibility assays were applied in human donor cells [258].

The traditional ATAC-seq technique is optimized for exactly 50,000 cells for the sake of proper cellular DNA to Tn5 transposase ratio. ATAC-seq sensitivity and specificity would decrease considerably when starting with 500 or 5,000 cells [238]. Meanwhile, the cost-effect ratio of scDNase-seq [241] per sample is much lower than the commercially available transposase for ATAC-seq, though it is feasible to independently produce the ATAC-seq transposase in research labs [259]. Single-cell

ATAC-seq (scATAC-seq) protocols [260] involve either a programmable microfluidics platform or combinatorial indexing of nuclei. Compared to scDNase-seq, these techniques could efficiently process a larger number of single cells at the same time although they require special knowledge and acquisition with the microfluidics machine. Every single cell can get more than 300,000 unique reads on average with saturation scDNase-seq, while individual cells passing quality filters for scATAC-seq get about 5,000 unique reads. Less than 10% of the bulk ATAC-seq identified promoters are detected in any individual cell scATAC-seq, while about 50% of bulk DHS promoter sites can be detected in scDNase-seq data. Hence, it is important to combine results from hundreds of single cell scATAC-seq data to draw trustworthy conclusions considering limitations of scATAC-seq ability to capture enough DHSs for an individual cell. 50,000 formaldehyde-fixed cells and larger cell numbers have been tested using ATAC-seq [261], while the scDNase-seq protocol is available for both native and formaldehyde-cross-linked cells. In this aspect, the scDNase-seq protocol increases the flexibility of experiment plan and sample preparation. Undoubtedly, the ATAC-seq protocol is also keeping improving to increase signal-to-background ratio and work on difficult cell lines, rare primary cells, and clinically relevant frozen tissues [262]. The new technique is called Omni-ATAC, which is an improved ATAC-seq protocol. Researchers demonstrated that both Omni-ATAC and DNase-seq protocols were capable of identifying substantial numbers of peaks not found using the standard ATAC-seq method [262]. Lastly, formaldehyde-assisted isolation of regulatory elements sequencing (FAIRE-seq) is another assay of DNA accessibility, although not as that popular as DNase-seq and ATAC-seq.

FAIRE-seq is more sensitive for distal regulatory elements detection compared to DNase-seq [263, 264]. However, the FAIRE-seq protocol requires 1 million to 10 million cells, which makes its application to rare primary cells and other difficult-to-get specimens infeasible. In a nutshell, both scDNase-seq and ATAC-seq protocols have been widely used in a range of cells and tissues to detect genome-wide DNA hypersensitive sites and regulatory regions. With each technique's powers and limitations in mind, it is the researchers' responsibility to choose the optimal protocol for special experiment materials in hand.

3D structure and chromatin interaction

The old-fashion concept that locations of genes on the chromosome were the primary determinants of gene activity has been abandoned substantially since more and more scientific evidence showed gene activities were highly influenced by the three-dimensional organization of regulatory elements and the corresponding chromatin interaction. Over the past decade a series of approaches have been developed to study three-dimensional chromosome folding based on sequence profiling technologies, such as high-throughput sequencing to detect the genome conformation and chromatin interaction (Hi-C and ChIA-PET) [265, 266]. These methods substantially rely on the chromosome conformation capture (3C) technique [222]. The underlying hypothesis is that if any pair of loci in the genome is in physically close proximity (possibly in the range of 10-100nm), frequencies of contacts should be high compared to random occurrence [267, 268]. Those techniques enable the study of how various cellular activities are affected by constraints of chromatin and chromosomal structure. DNA looping is a crucial element to bind promoters and distally located enhancers,

and influence the function of transcription initiation complexes (Figure 1.4 [223]). Fluorescence in situ hybridization (FISH) has been used to measure the influence of genomic distance on enhancers [269, 270].

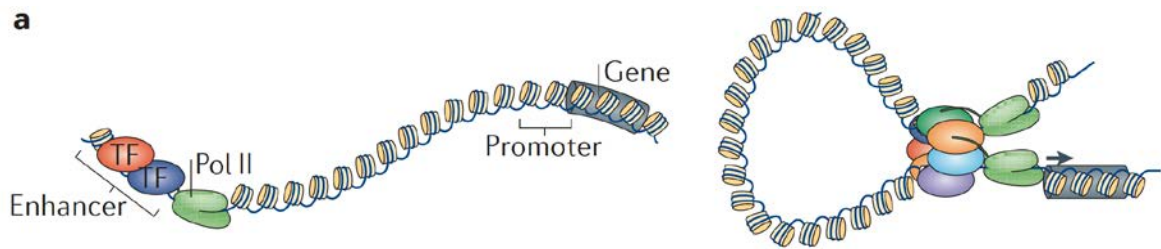


Figure 1.4 Connection Between Promoters and Enhancers

(Pombo A, Dillon N: **Three-dimensional genome architecture: players and mechanisms**. *Nature reviews Molecular cell biology* 2015, **16**(4):245-257.)

This figure indicates the direct connection between promoters and enhancers through looping out of intervening genomic sequences. RNA polymerase II would bind to transcription factor binding enhancers before transcription activation.

The rapid development of 3C-based techniques, including chromosome conformation capture on chip (4C) [271], 3C combined with high-throughput sequencing (4C-seq) [272], multiplexed 3C sequencing (3C-seq) [273], carbon-copy chromosome conformation capture (5C) [274], capture-C [275], genome conformation capture (GCC) [276], Hi-C [265], tethered conformation capture (TCC) [277], and targeted chromatin capture (T2C) [278], has accelerated the study of chromatin organization, making it possible to map interactions across different genomic distances with varying levels of resolution [223]. The core concept of 3C-based methods involves crosslinking and ligation of interacting regions. Those ligated regions are considered as chromatin contacts. However, background noise could come from both basal levels of random ligation [222] and inherent biases in restriction enzyme cutting efficiency.

Also, the fluctuation (temperature, time etc.) initiated in the process of crosslinking would highly contribute to variations among samples. We should be cautious when interpreting results from those ligation-based techniques.

One of the most well studied DNA looping mechanisms is the formation of active chromatin hubs involving the locus control regions (LCR). ACH-type contacts have been reported in erythroid-specific mouse β -globin gene [279], mouse CD8 locus [280], human RUNX1 [281] and CFTR loci [282]. Active chromatin hubs formation might be able to create a specific chromatin folding structure. However, contacts between enhancers, promoters and silencers can be highly flexible and diverse rather than a single pattern like ACHs. The current discovery supports the function of cohesin and CCCTC-binding factor (CTCF) that promote interactions between distally located sequences. The core cohesin complex forms a ring structure and is loaded onto DNA in the cell cycle [283]. Cohesin is also involved in controlling gene expression through the stabilization of long-range interactions between *cis*-regulatory elements [223], and large numbers of cohesin-binding sites colocalize with CTCF binding in the mouse genome [284]. CTCF attracts great interest because it has been associated with insulators, which are sequences that block the action of enhancers on promoters [285], and is involved in promoting long-range interactions that stimulate transcription [223]. Ren et al. demonstrated that CTCF decreases cell-to-cell gene expression variations by stabilizing enhancer-promoter interactions [286].

Some genomic regions contain high levels of chromatin interactions, which are interspersed with genomic regions with fewer interactions [223]. Some researchers define the self-interacting regions as topologically associating domains (TADs). The

occurrence of topological domains is stable across different cell types and partially conserved among species. Housekeeping genes, rather than enhancers or transcription repression associated histone marks, are often enriched in boundaries between TADs. TAD is an individually folded structure with particular chromatin activities, which has the size of a kilo to megabase pairs. The central regions have high numbers of chromatin interactions, along with the high frequency of CTCF and cohesin binding events. Interestingly, CTCF binding is involved in inhibition of inter-interactions between TADs and delineation of boundaries between TADs [223]. There also exists highly dynamic structures of submegabase scale, i.e. smaller-size domains named sub-TADs [287-289], and chromatin loops [288, 290]. As complex entities controlled by different parameters, TADs' chromatin interactions are likely to depend on active chromatin hubs (ACH) formation through enhancer-promoter organizations. The inner envelope of the nucleus is covered by lamins, and lamins help form the nuclear lamina. The nuclear lamina plays a vital role in organizing the spatial chromosome structure through contacts with chromatin and maintaining spatial locations on the inner nucleus. Lamina-associated domains (LADs) are regions of condensed chromatin bound by the nuclear lamina and are enriched for histone 3 Lys9 dimethylation marks, which are related to heterochromatin and silent genes. LADs have a relatively low gene density and most of genes are inactive, indicating a repressive role in gene regulation. During cell differentiation, constitutive LADs (cLADs) maintain attachment with the lamina while facultative LADs (fLADs) become detached to activate gene expression within fLADs. After mitosis, some

LADs can relocate to the nucleolus periphery, although the underlying mechanism remains unknown [291, 292].

As discussed above, high-resolution 3C-based techniques and microscopy technology have revealed some complex chromatin organization and cellular functions [293, 294]. However, it is necessary to note that different experiment protocols may require specific computational algorithms and pipelines for profiling 3D chromatin interactions. A standard sequence-based 3D data processing involves raw reads mapping, fragment filtering, read-pair filtering, normalization, and downstream analysis [221]. Most critical parts are choices of statistical models and quantitative analyses since they are the prerequisite for correctly interpreting the data and discovering the chromosome folding mechanism. The Strings and Binders Switch (SBS) model was generated relying on the hypothesis that chromatin conformations happen during the interaction of specific binding sites with diffusing binding factors [221] and it has been used to decipher dynamic chromatin structures in spatial and temporal dimensions [295, 296]. TargetFinder, developed to reconstruct regulatory landscapes, has been reported to correctly infer interactions up to 2 Mb apart with a concentration on high-resolution intra-TAD interactions rather than inter-TAD interactions [297, 298]. A Mixture Poisson Regression Model and a power-law decay background were proposed to delineate specific sets of interacting genomic loci and regions [280], and integration analysis with DNase-seq and ChIP-seq data proved that Hi-C data contain informative interacting loci [299]. The multiscale correlation evaluation (MCORE) algorithm was introduced to quantify and compare chromatin patterns depending on mapped sequencing reads fluctuation spectrum [300]. Bortle et

al. concentrated on highly conserved insulator and Polycomb multiprotein complexes, which were previously known for mediating long-range interactions and nuclear structures [301]. Network analysis combined with epigenomic landscape was proposed recently to identify proteins and chromatin marks interceding genomic interactions [302]. Hi-C contact maps are the basis for Hi-C data analysis; a linear regression-based chromosome-level adjustment method called caICB [303] was developed to correct the copy number bias in the Hi-C interaction matrix. In order to infer 3D spatial associations within TADs from histone modification, chromatin accessibility, and gene expression profiles, EpiTensor was introduced to computationally characterize hotspots as critical factors maintaining 3D contacts [304].

It is another exciting research focus to apply 3C-based high-throughput profiling techniques to disease biology. Several studies have shown aberrant 3D chromatin structures and interactions in human diseases. Three-dimensional genome architecture rearrangements were directly observed in human diseases especially in cancer, and interaction frequencies were tissue-specific [305]. The complicated relationship between gene expression and chromatin 3D structure has been observed in breast cancer [306] and leukemia [299]. The recent study found that genomic duplications in human patient cells and mice could result in the formation of new chromatin domains (neo-TADs), and neo-TADs, in turn, could influence pathology consequences [307]. A new concept, enhancer hijacking, was proposed to explain the rearrangement of TAD boundaries, and enhancer hijacking was reported to control cancer-related gene overexpression in colorectal cancer [308]. Genetic mutations influencing

chromosome neighborhood boundaries were discovered in many types of cancer, and oncogene activation could happen in those malignant cells [309]. Human IDH mutant glioma had abnormal TADs and initiated oncogene expression because of hypermethylation at cohesin and CTCF binding sites [310]. TADs' disruption has led to human limb malformations due to the rearrangement of chromatin interactions between promoters and enhancers [311].

Within a short period of time, the numbers of 3D genome organization studies explosively increased, and various new techniques are under active development. Molecular imaging techniques, including quantitative 3D telomere fluorescence in situ hybridization (FISH) analysis and 3D super-resolution imaging (3D-SIM) [312, 313], could become potent tools to validate genome-wide profiling results and their functional characteristics. We nevertheless should be cautious with the nature of ligation-based methods and corresponding data analysis. Multi-source information integration from different angles would be expected in studies to accurately interpret transcriptional regulation and functional entities [314].

Rationale and Significance

CD4⁺ helper T cells play a significant role in host defense and immune-mediated diseases. CD4⁺ cells assist B cells in making antibodies, inducing macrophages, and recruiting neutrophils, eosinophils, and basophils to infection and inflammation sites. Antigen-inexperienced naive CD4⁺ T cells can differentiate and develop into specialized subsets of T helper cells in response to many microbial pathogens for various immune-related purposes. A further investigation of signals that control

CD4⁺ T cell lineage commitment versus flexibility/plasticity has clear implications for disease pathogenesis and new therapeutic applications, to fight infections and to manage autoimmune and allergic inflammatory disorders. A variety of autoimmunity diseases are associated with particular T cell types, and new therapeutic approaches for treating those diseases may be pragmatic if T cell responses are plastic. This topic of the mechanisms by which CD4⁺ T cells respond to environmental stimuli including virus infection and cell differentiation also represents a fundamentally essential and intriguing cell biological question of how the microenvironment impacts the organism's intrinsic gene expression control networks.

Marek's disease (MD), which is highly contagious and caused by MDV infection, has a high mortality rate and results in considerable economic loss to the poultry industry. The current over-usage of vaccines and drugs promotes greater MDV virulence. Since CD4⁺ T cells are primary targets of Marek's disease virus (MDV), it is reasonable to infer their unique and vital functions in disease progression. It is also justifiable to hypothesize that the viral infection in chickens would in turn influence CD4⁺ T cell differentiation and plasticity. Therefore, a better understanding of CD4⁺ T cell differentiation and plasticity would benefit substantially towards MD resistance.

Other studies in mammals and yeasts have revealed significant involvement of chromatin accessibility changes and three-dimensional chromosome reorganization during development, disease progression, and viral infection, as discussed in the above literature review. Hence, we hypothesize that CD4⁺ T cell progression is universally navigated by epigenetic factors including DNA hypersensitive sites and chromatin topologically associated domains. The further testing and exploration of

this hypothesis will provide deep insight into the mechanism of MD progression and facilitate disease control and prevention in the poultry industry. Meanwhile, such genome research helps reveal the significance of genetic and environmental interaction networks behind complex diseases, viral infection, and even primary living activities.

One aspect to test our hypothesis is to identify genes and regulators most likely related to MD-resistance and the universal influence of MDV infection on the hosts, including chromatin accessibility and three-dimensional chromosome organization. Two chicken lines with varying Marek's disease resistance induced by Marek's disease virus (MDV) infection provide an excellent model to study how CD4⁺ T cells behave differently in these two chicken lines, from the aspect of transcription activity and epigenetic factors, such as DNA hypersensitive sites and chromatin interactions, and how the fluctuation would explain the biological and disease response differences. We also anticipate studying CD4⁺ T cell differentiation in detail with regards to time, environment, and plasticity. However, the T cell differentiation model is well established in mice rather than chickens, which makes the further mechanistic study in chicken CD4⁺ cells technically tricky. In order to investigate the interaction between the animal's genome and environmental factors, especially finding key regulators that control the animal's immune response when facing changes in environment and viral infection, the mouse CD4⁺ T cell differentiation model is utilized in my study. My research is trying to answer questions if T helper subsets are intrinsically distinct cell populations and if they are capable of changing or adjusting their committed cell fates. Factors determining commitment and plasticity, including

transcription factors and epigenetic modifications, will also be investigated. High-throughput and more sophisticated approaches discussed earlier in this introduction will generate comprehensive data sets and allow genome-wide analysis of CD4⁺ T cells' epigenetic status.

Although at first glance, different studies on chicken CD4⁺ T cell changes induced by Marek's disease virus infection and on mouse CD4⁺ T cells plasticity with the focus on Th1 subsets have no apparent relationships, they are inherently and intricately related. The time-dependent plasticity of Th1 helper cells identified in mouse CD4⁺ T cells highly mimics the biological process that occurred in Marek's disease resistant and susceptible chicken lines. The viral infection in chickens is only a system of environmental changes that could influence cell behavior and function, like the cytokine circumstances during the mouse Th1 cell differentiation process. To the best of our knowledge, our intriguing findings in mouse CD4⁺ T cells would facilitate our understanding of how phenotypic differences in the two inbred chicken lines correlate with CD4⁺ T cell differentiation in MD progression.

Outcomes of this dissertation will further our understanding of epigenetic factors' influence on CD4⁺ T cell differentiation and immune responses, in general, and the effect of MDV infection on host genomes, in particular. This dissertation consists of the following three parts:

1. To investigate tissue-specific gene expression induced by MDV infection in the spleen, thymus, and bursa of chickens.
2. To reveal chromatin accessibility and organization changes induced by MDV infection in chicken CD4⁺ T cells.

3. To decipher the mechanism of dynamic epigenetic status and plasticity during mouse CD4⁺ T cell differentiation.

Chapter 2 : Transcriptomic Analysis of Chicken Immune Organs Affected by Marek's Disease Virus Infection

Abstract

Over the past decades, MDV has dramatically evolved towards more virulent strains and constructs a persistent threat to the world's poultry industry. We performed genome-wide gene expression analysis in the spleen, thymus, and bursa tissues from MD-resistant line 6₃ and susceptible line 7₂ to explore the mechanism of MD resistance and susceptibility. By utilizing the robust RNA sequencing approach, we identified confident genes and pathways that were associated with transcriptional response to MDV infection. The transcriptome analysis revealed a tissue-specific expression pattern among immune organs when confronting MDV. At pathway and network levels, MDV infections influenced both cytokine-cytokine receptor interaction and cellular development in resistant and susceptible chicken lines. Meanwhile, we also observed the different genetic responses between the two chicken lines: some pathways like Herpes simplex infection and Influenza A were found in line 6₃ spleen tissues whereas metabolic related pathways and DNA replication could only be observed in line 7₂ chickens. In summary, our research renders new perception into the MD progression mechanism and beckons further gene function studies into MD-resistance.

Introduction

Marek's disease (MD) is a chicken's highly contagious and neuropathic disease caused by Marek's disease virus (MDV), which is a cell-associated alphaherpesvirus that transforms T lymphocytes and triggers mononuclear tissue infiltration such as peripheral nerves, muscle, visceral organs, and skin [4-6]. MDV infection involves the early cytolytic stage, latency phase, and late cytolytic phase; at each stage of infection, the virus targets immune components of the host system [315]. MDV confronts primary target cells (B lymphocytes and activated CD4⁺ T lymphocytes) and reaches replication peak from three to seven days post-infection [10-12]. 7-8 days after infection, MDV progression switches from cytolytic phase to latency without producing infectious progeny in activated CD4⁺ T cells and infected B cells. Reactivation from latency to late cytolytic infection occurs around 2-3 weeks post-infection in susceptible chickens [7]. The reactivation period (late cytolytic phase) is usually companioned with tumor formation and other acute disease symptoms.

High-throughput gene expression analysis has been widely used to understand host-virus interactions. Before the RNA-Seq method, microarrays were helpful to detect host gene expression in chicken embryo fibroblasts [316] and spleen tissue [75, 77, 79] using MD-resistant and MD-susceptible chicken models [39, 78, 80]. *CD8A*, *IL8*, *USP1*, and *CTLA4* genes were considered as significant genes associated with MD-resistance or –susceptibility by testing temporal transcriptome changes using three representative chicken lines [39]. By taking advantage of next-generation sequencing techniques, researchers have characterized differential genes in the spleen of broiler

and layer chickens, finding that TLR receptor and JAK/STAT signaling pathways were enriched following MDV infection [317].

Well-defined chicken models involve two highly inbred chicken lines 6₃ and 7₂, sub-lines of lines 6 and 7, which have been bred since 1939 and served as unique resources to explore the mechanistic difference towards MD response. Availability of chicken genome sequence [318] and next-generation sequencing techniques have altered our capability to identify more critical factors for MD resistance using these two chicken lines. Here we conducted genome-wide profiling of spleen, thymus, and bursa transcriptomes in MD-resistant line 6₃ and MD-susceptible line 7₂ using mRNA sequencing (RNA-Seq). Age-matched controls aiming at testing innate distinction among inbred chicken lines and those induced by MDV infection were implemented in our experimental design. Also, comparisons among immune organs would further uncover gene candidates related to different MD response.

Materials and Methods

Animals and experimental design

Two inbred lines of White Leghorn (line 6₃ and line 7₂) were hatched, raised and maintained in Avian Disease and Oncology Laboratory (Michigan, USDA). Chickens from each line were separated into infected and non-infected groups respectively. Chickens from the infected group were injected intra-abdominally with a partially attenuated virulent plus strain of MDV (648A passage 40) at 5 days after hatching with a viral dosage of 500 plaque-forming units (PFU). At 21 days post-infection, chickens from each treatment and control group were sacrificed following standard

animal ethics and usage guidelines. Samples from spleen, thymus, and bursa organs were gathered, frozen and stored at -20 C until RNA extraction.

RNA preparation and sequencing

Spleen, bursa, and thymus samples were randomly selected from infected and control groups from line 6₃ and line 7₂ chickens. Approximately 30~50mg of the spleen, thymus, and bursa tissues were homogenized in TRizol Reagent (Qiagen, USA), and total RNA in infected and control groups was extracted by using RNeasy Mini Kit (Qiagen, USA). The mRNA isolation was performed using Oligotex mRNA Mini Kit (Qiagen USA) according to the manufacturer's instruction. Biological replicates were applied for further RNA-Seq library construction and analysis. About 300ng of mRNA was used to synthesize the first and second strands of cDNA by SuperScript™ III Reverse Transcriptase (Invitrogen, USA) and Second Strand cDNA Synthesis Kit (NEB, USA). After sonication, the dscDNA fragment ends were repaired by T4 and Klenow DNA polymerase and underwent library construction procedures following Illumina. Each library was identified by adding a 6-bp adaptor and sequenced at 50-bp read by an Illumina HiSeq 2500 Sequencer.

Differential expression and pathway analysis

Raw sequencing data were checked for quality considering read counts, overall read quality, and read distribution, etc. The alignment to the reference genome (ICGSC Gallus_gallus-4.0/galGal4) downloaded from the UCSC Genome Brower (<http://genome.ucsc.edu/>) was performed by using an ultrafast memory-efficient short read aligner Bowtie [319]. First 15 bps from the original raw 50-bp reads were

trimmed to control the mapping quality. The counting matrix was generated by the `summarizeOverlaps` function in R. R package `edgeR` [320] and corresponding complementary functions GLM were executed to analyze count reads data and perform comparative analysis. The threshold of differentially expressed genes was set as 0.1 FDR. David [321] and Ingenuity Pathways Analysis (IPA) [322] were utilized to analyze the biological process, molecular functions and pathways enriched for those differentially expressed genes.

Principal component analysis and Venn diagram construction

After alignment of sequencing reads and the generation of counting matrix by `summarizeOverlaps` function in R, an integrated, normalized data matrix for the spleen, bursa, and thymus was created using Bioconductor's `DESeq` [323] package. Principal components facilitate dimensionality reduction and noise filter, which identifies directions of maximum variance in the original data. The function *prcomp* in the R statistic package was used to perform PCA on the integrated, normalized data matrix. The top PCs were visualized using the `factoextra` package in R. The Venn diagram was constructed using `Venn diagram` function in R.

Quantitative Real-time RT-PCR

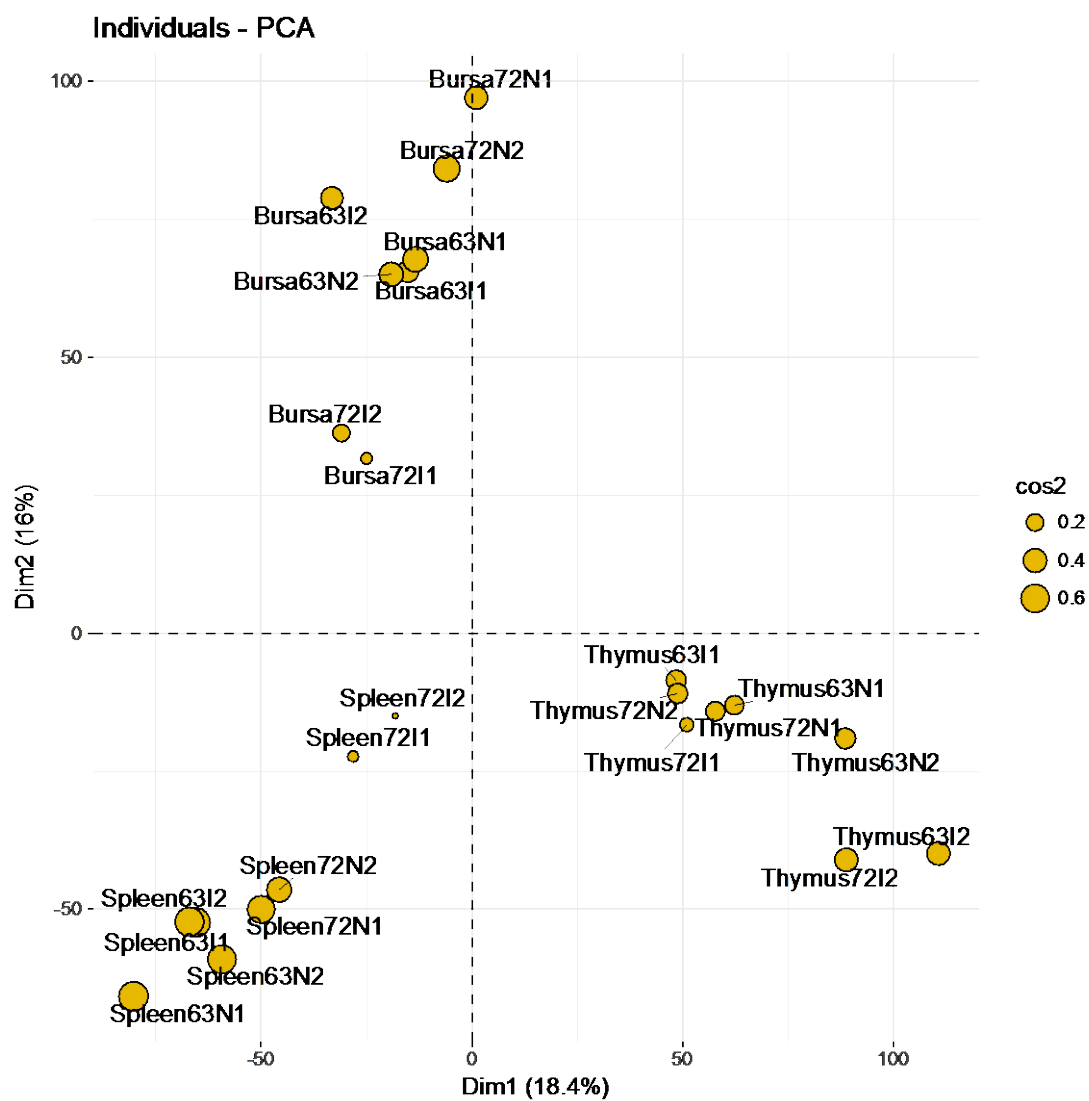
Several significant genes based on RNA-Seq analysis were validated by real-time PCR using the synthesized dsDNA described previously [39]. Real-time reactions were conducted with an iQ SYBR Green PCR Kit (Bio-rad, USA) according to manufacturer's instructions. Primers were designed using Primer3 (<http://fokker.wi.mit.edu/primer3/input.htm>). The melting temperature was 60°C, and

the length of amplicons was between 50-200bp. Ct values were calculated based on normalization of the housekeeping gene GAPDH, and three technical replicates were performed.

Results

Gene expression profiles of immune organs in MDV challenge experiments

To discover genes that might involve in MDV response, we performed transcriptome analysis on the spleen, thymus, and bursa of Fabricius of two chicken lines at 21 days post-MDV infection. An integrated gene expression dataset of three immune organs was created as described in Materials and Methods. The data matrix contained normalized gene expression measurements of 24 samples for 17,108 genes. To explore whether the samples would form distinct groups based on their gene expression profiles, we used principal component analysis (PCA) on the data matrix, and the results were visualized by scatter plots (Figure 2.1). The preliminary PCA plot indicated the chicken samples tended to cluster by organs regardless of MDV infection when we examined the first two principal components, indicating tissue-specific gene expression patterns after MDV infection. Also, the similarity of gene expression patterns between the spleen and thymus was higher compared to other combinations when we plotted PC2 and PC3.



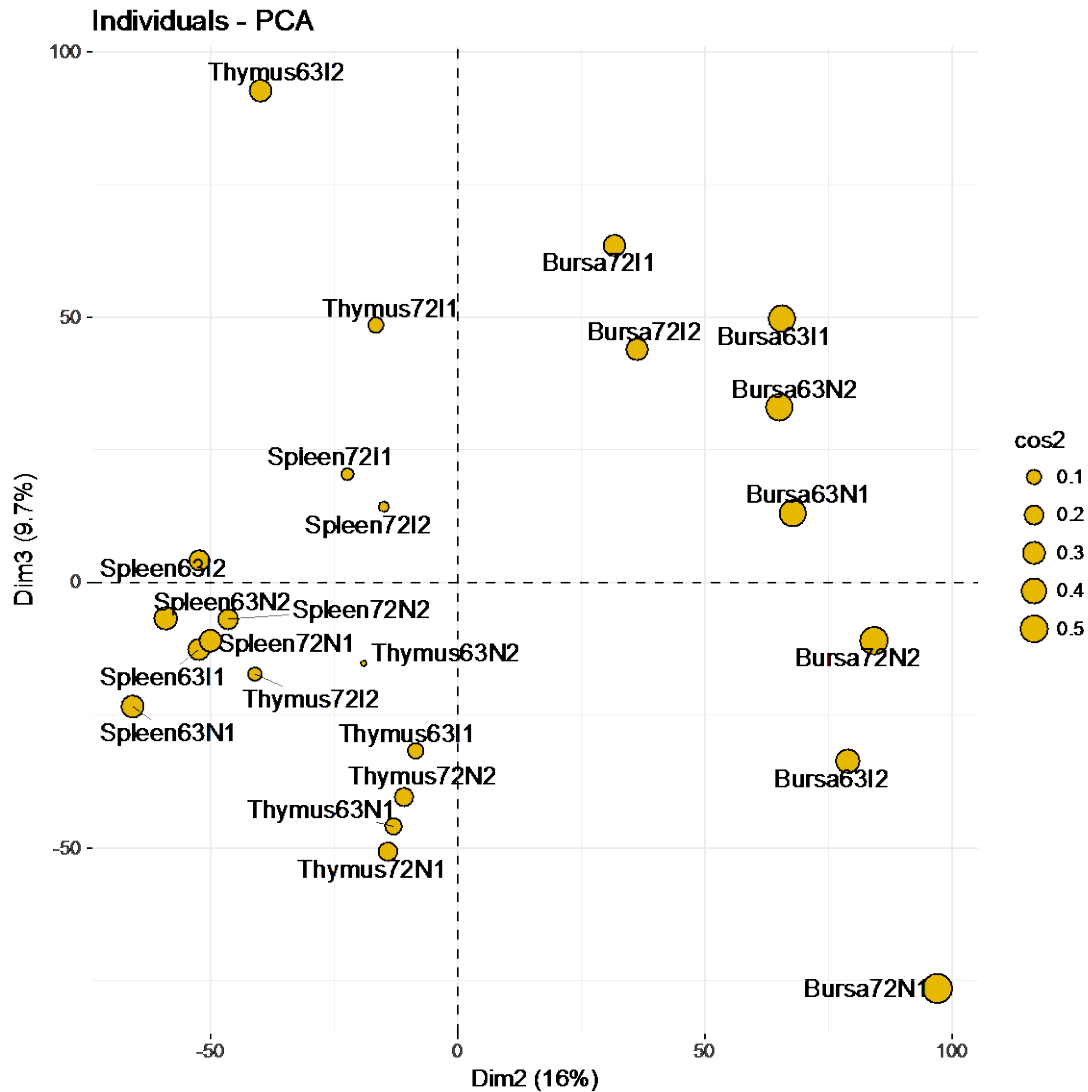


Figure 2.1 PCA Plots.

Two-dimensional PCA plots of 24 individuals indicate tissue-specific gene expression patterns with a slight similarity between chicken spleen and thymus organs. Cos2 values indicate quality of the representation for individuals on the principal components. A higher cos2 represents a better representation of the individual on the principal component.

Genome-wide profiling of the spleen transcriptome at late cytolytic phase

For the spleen tissues, more than 211 million sequence reads were generated and mapped to November 2011 version of the chicken genome (galGal4). Read counts

associated with annotated Ensembl genes [324] were calculated, and differential gene expression was profiled by the statistical package edgeR [320]. Comparisons were carried out between infected and control birds within each line, and birds underwent non-infection status from two lines to indicate disease response and baseline transcription differences.

The behavior discrepancy between the two chicken lines (L6₃ and L7₂) is remarkable when challenged by MDV infection. We witnessed a significantly larger number of differentially expressed (DE) genes in the comparison between MDV-infected and non-infected line 7₂ chickens (susceptible chicken line) than that between MD-resistance chickens. In general, the analysis detected 817 differentially expressed (DE) genes between infected and control line 6₃ chickens, and 4584 DE genes in L7₂ chickens using a threshold of FDR<0.1 (Figure 2.2). Two chicken lines share 522 DE genes regardless of the changing direction. Interestingly, the direction of changes in these two lines is similar: a majority of the DE transcripts in the resistant line 6₃ (563 out of 817, 68.91%) and the susceptible line 7₂ (2959 out of 4584, 64.55%) were downregulated. In those shared 522 DE genes, 136 genes were upregulated both in line 6₃ and 7₂ infected birds, and 365 genes were downregulated while the remaining 21 genes displayed opposite alteration trends in resistant and susceptible chicken lines. When comparing the innate difference between two chicken lines, a majority of DE genes (801 out of 1118, 71.64%) exhibited higher expression levels in line 6₃. Almost half (548 out of 1118) of these genes were also differentially expressed in susceptible line 7₂ in response to MDV infection (84.49% downregulated). Those distinctions

between the two inbred chicken lines might reflect the baseline transcription that could be the possible contribution of varying MD-resistant phenotypes.

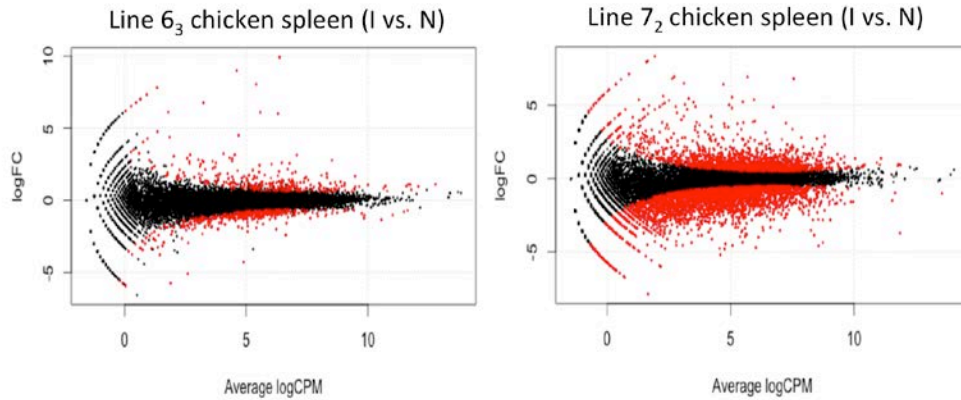


Figure 2.2 Differentially Expressed Genes in Chicken Spleen Tissues

In chicken spleen tissues, the analysis detected 817 differentially expressed genes between MDV infected and non-infected L6₃ chickens (left panel), and 4584 genes in L7₂ chickens (right panel) with FDR<0.1. Each red dot represents the individual differentially expressed gene.

Thymus and bursa transcriptome profiling at late cytolytic phase

Following the similar experimental design and analysis methodology, we examined the genome-wide transcription levels in the thymus and bursa organs at the late cytolytic phase for the above chicken lines. Similarly, considerably larger number of DE genes were observed in line 7₂ birds for both tissues. In thymus tissues, we discovered 734 DE genes in susceptible chicken lines while only 14 DE genes in resistant chicken lines using a threshold of FDR<0.1 (Figure 2.3). In the bursa organ, 259 and 5387 DE genes were detected in resistant and susceptible chickens respectively with FDR<0.1 (Figure 2.4). Interestingly, the majority of DE genes were

upregulated in response to MDV infection in bursa and thymus tissues except for line 6₃ chicken thymus.

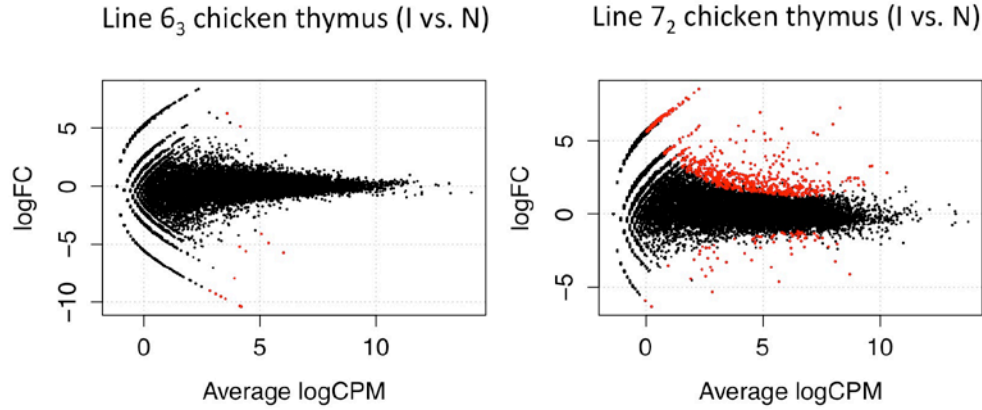


Figure 2.3 Differentially Expressed Genes in Thymus Tissue

The volcano plot is showing the discovered 734 DE genes in susceptible line 7₂ chickens while 14 DE genes in resistant line 6₃ birds using a threshold of FDR<0.1 in the thymus tissue.

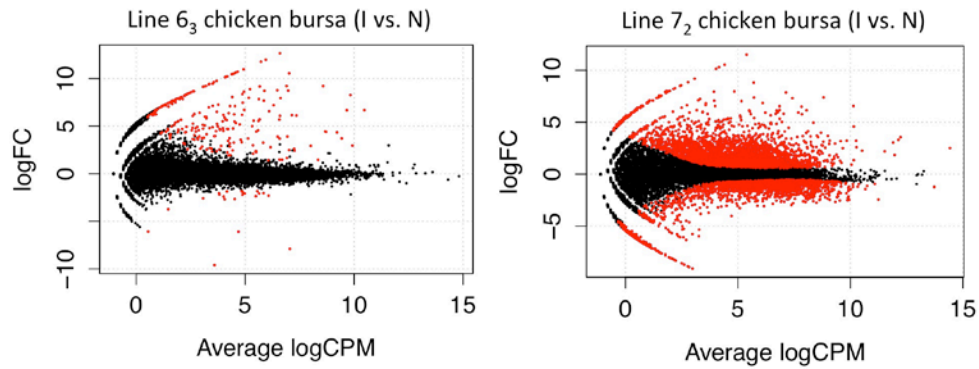


Figure 2.4 Differentially Expressed Genes in Bursa Tissue

In the bursa organ, 259 and 5387 DE genes were detected in line 6₃ and line 7₂ chickens respectively with FDR<0.1.

Comparisons between control birds in thymus and bursa organs revealed the intrinsic difference between chicken lines. Approximately 57% (266 out of 465) DE genes showed higher expression levels in the thymus of resistant line 6₃ chickens and 54.74% (1542 out of 2817) in the bursa of line 6₃ chickens. In summary, by conducting two sets of comparisons in different immune organs, we were able to uncover the innate and MDV-induced variation in two inbred chicken lines. A slightly larger proportion of genes exhibited higher baseline expression levels in the resistant line. However, more significant numbers of DE genes were detected in the susceptible chickens than in the resistant chickens at the late stage of infection, probably due to transcriptional activation during the late cytolytic phase in MDV-susceptible chickens.

Witnessing the distinct features of these chicken lines, we attempted to characterize genes associated with MD-resistance and –susceptibility to pair-wise comparisons. We made the following hypothesis: genes differentially expressed after MDV infection in all three immune organs are likely a reflection of the typical host response and an immune network to viral infection; genes that are differentially expressed merely in one immune tissue could be part of an organ-specific host response and function to viral infection. According to this rationale, we identified common responses and organ-specific functional genes after MDV challenges (Figure 2.5). The majority of DE genes were only present in one tissue, which indicated the tissue-specific expression patterns of MDV infection. Based on KEGG pathway, the significant pathway for those shared DE genes involved the cytokine-cytokine receptor interaction.

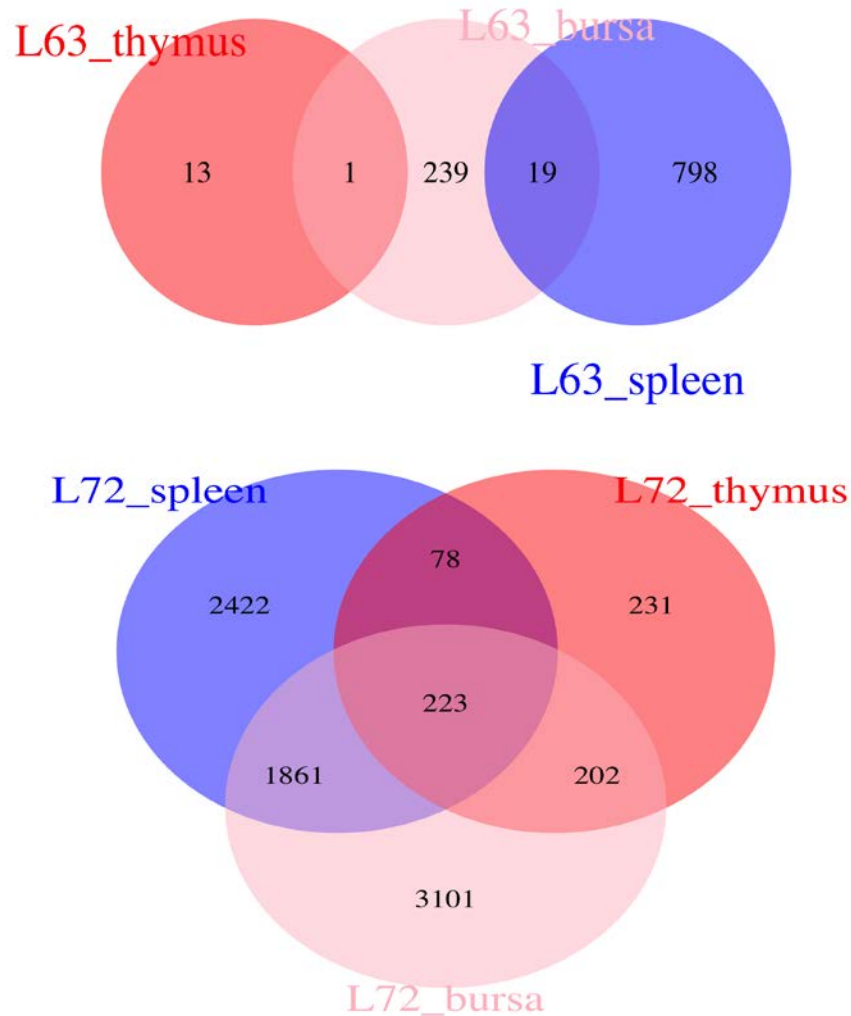


Figure 2.5 Venn Diagram Indicating Overlap Genes

The Venn diagram indicates the overlap differentially expressed genes among different organs after Marek's disease virus infection.

Pathway analysis to reveal networks and biofunctions involved in MDV infection

In the DE gene list, some well-known immune factors have been identified in previous studies. These include chemokine receptor 1 (Xcr1), chemokine receptor 2 (CCR2), interleukin 1 receptor, type II (IL1R2), type I (IFNA) interferon, and type II (IFNG) interferon. Additionally, several other innate immune function genes were influenced, such as FYN oncogene related to SRC, FGR, YES (FYN), and CD247

molecule (cd247), which are involved in natural killer cell-mediated cytotoxicity. Interestingly, some profoundly affected genes have no known function in innate immune responses, including collagen type II alpha 1 (COL2A1) and growth factor receptor-bound protein 2 (GRB2).

To identify enriched biological function and networks associated with the differentially expressed genes after MDV infection, we utilized the Ingenuity Pathway Analysis (IPA) and DAVID Bioinformatics Resources to analyze the gene sets obtained by comparing to age-matched uninfected controls. Top networks (from IPA) influenced by MDV infection in three organs are shown in Table 2.1. From these networks, we noticed that the virus has an evident and broad impact on host gene expression in all immune organs of two chicken lines. Also, if we considered those genes associated with diseases and bio-functions, we discovered that by MDV infection, many genes were influenced, which are involved in metabolism, tissue development, and immune-related disorders. Those genes indicated broad similarities in the host response to MDV infection. Nevertheless, some unique networks, such as digestive system development and function, gastrointestinal disease, and hepatic system development and function, were only shown in the spleen of line 7₂ chickens after MDV infection (Table 2.1). To gain more insight into the bio-function associated with MDV infection, we utilized the IPA program to analyze disease and bio-function of those differentially expressed genes (Table 2.2). As might be expected, cancer and immunological illness were observed.

Complementary to the discoveries from IPA, DAVID displays biological pathways that were altered (compared to a relevant control) during the host response to MDV

infection (Appendix I). Based on KEGG pathways [325], immune-related pathways, such as cytokine-cytokine receptor interaction (Figure 2.6), phagosome, herpes simplex infection, and lysosome, were involved in MDV infection response. Also interestingly we could notice the changes in focal adhesion and tight junction pathways, and this might relate to the viral progression in host bodies since MDV requires cell-to-cell contact for dispersal [80], and the virus might promote such communication by altering the expression of related tight junction formation genes.

Table 2.1 Enriched networks in organs of chicken lines after MDV infection

Lines	Organ	Enriched Network Functions
L6 ₃	Spleen	Cellular Development, Embryonic Development, Nervous System Development and Function Cancer, Respiratory Disease, Developmental Disorder Cell Death and Survival, Cellular Development, Cellular Function and Maintenance Developmental Disorder, Hereditary Disorder, Immunological Disease Cellular Function and Maintenance, Connective Tissue Development and Function
	Bursa	Cell-To-Cell Signaling and Interaction, Hematological System Development and Function, Immune Cell Trafficking Cardiac Hypertrophy, Cardiovascular Disease, Developmental Disorder Cellular Development, Embryonic Development, Organismal Development Cell-mediated Immune Response, Cellular Development, Cellular Function and Maintenance

	Thymus	Cardiac Arrhythmia, Cardiovascular Disease, Organismal Injury and Abnormalities Cell Death and Survival, Cell Morphology, Cellular Compromise
L7 ₂	Spleen	Cell-To-Cell Signaling and Interaction, Cell Signaling, Cell Morphology Post-Translational Modification, Dermatological Diseases and Conditions, Developmental Disorder Digestive System Development and Function, Gastrointestinal Disease, Hepatic System Development and Function Hematological Disease, Metabolic Disease, Protein Synthesis Embryonic Development, Organismal Development, Cellular Movement
	Bursa	Cellular Development, Cellular Growth and Proliferation, Hematological System Development and Function Cell-mediated Immune Response, Cellular Function and Maintenance Carbohydrate Metabolism, Cell Death and Survival, Cellular Movement DNA Replication, Recombination, and Repair, Developmental Disorder
	Thymus	Infectious Disease, Connective Tissue Disorders, Dental Disease Metabolic Disease, Lipid Metabolism, Small Molecule Biochemistry Humoral Immune Response, Protein Synthesis, Inflammatory ResponseCell-To-Cell Signaling and Interaction, Cell Signaling, Molecular Transport Cell Cycle, Cellular Development, Cellular Growth and Proliferation

Table 2.2 Top disease and bio-functions in chicken organs after MDV infection

Lines	Organ	Diseases and Disorders	Physiological System Development and Function
L6 ₃	Spleen	Metabolic Disease, Inflammatory Response, Immunological Disease, Cancer, Organismal Injury and Abnormalities	Cardiovascular System Development and Function, Hepatic System Development and Function, Tissue Development, Connective Tissue Development and Function, Embryonic Development
	Bursa	Cancer, Organismal Injury and Abnormalities, Inflammatory Response	Hematological System Development and Function, Cell-mediated Immune Response, Hematopoiesis, Lymphoid Tissue Structure and Development, Immune Cell Trafficking
	Thymus	Cardiovascular Disease, Organismal Injury and Abnormalities, Connective Tissue Disorders, Developmental Disorder, hereditary Disorder	Cardiovascular System Development and Function, Embryonic Development, Organ Morphology, Organismal Development, Skeletal and Muscular System Development and Function
L7 ₂	Spleen	Cancer, Gastrointestinal Disease, Organismal Injury and Abnormalities, Reproductive System Disease, Hepatic System Disease	Cardiovascular System Development and Function, Organismal Survival, Organismal Development, Tissue Morphology, Embryonic Development
	Bursa	Cancer, Organismal Injury and Abnormalities, Inflammatory Response, Hematological Response, Immunological Disease	Hematological System Development and Function, Tissue Morphology, Hematopoiesis, Cell-mediated Immune Response, Lymphoid Tissue Structure and Development
	Thymus	Cancer, Inflammatory Response, Infectious Disease, Cardiovascular Disease, Gastrointestinal Disease	Tissue Morphology, Organismal Development, Hematological System Development and Function, Immune Cell Trafficking, Humoral Immune Response

CYTOKINE-CYTOKINE RECEPTOR INTERACTION

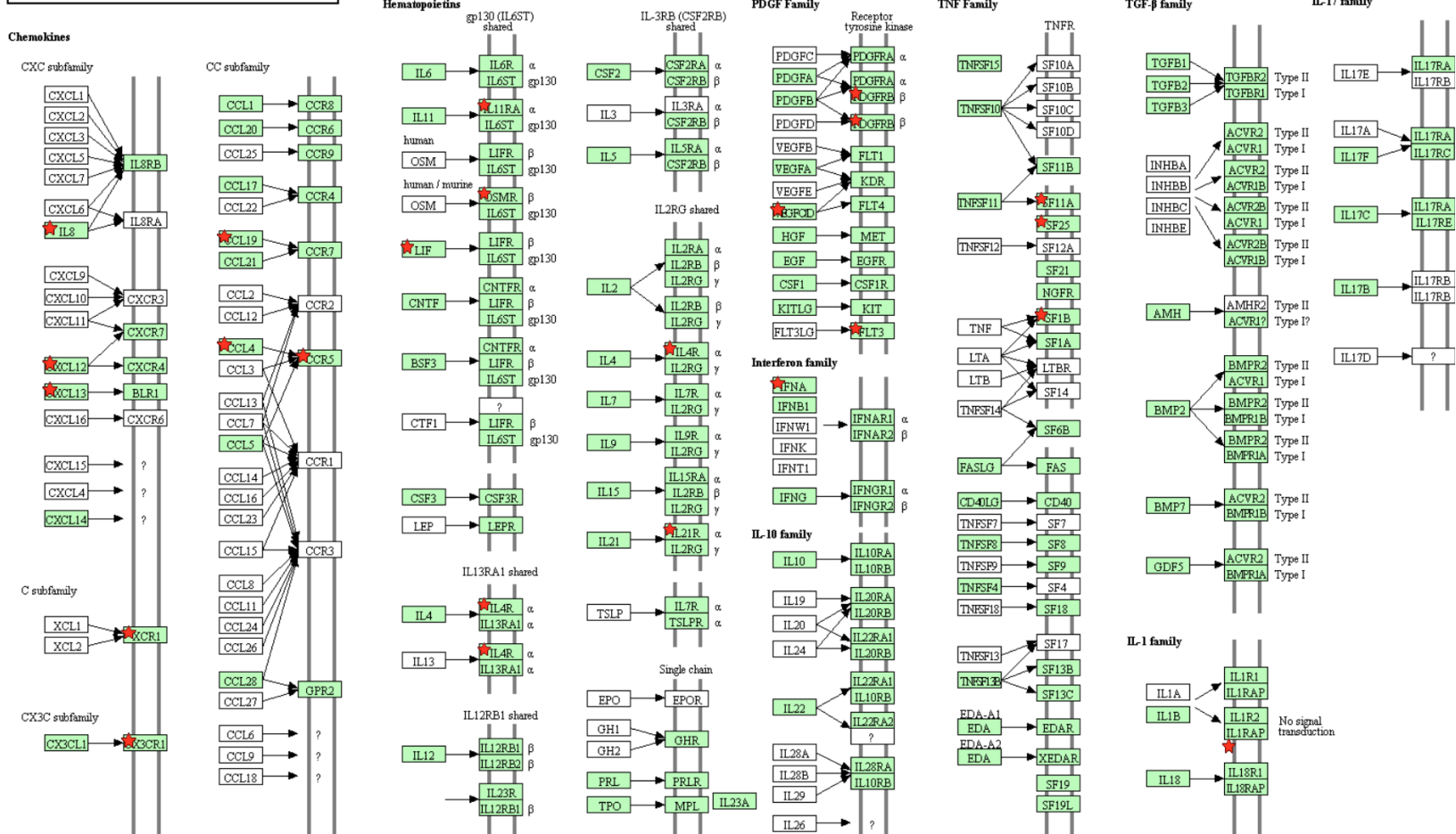


Figure 2.6 Cytokine-cytokine Receptor Interaction Pathway

One example of the enriched pathway induced by MDV infection in line 6₃ chickens. Several genes from this pathway display differential expression patterns, as indicated by red stars in the figure.

The potential native dissimilarity between lines makes the resistance different

The underlying mechanism for the different response to MDV infection between two lines might not be straightforward since MD is a complex disease. One possible hypothesis is that before MDV infection, line 6₃ chickens could promote better immune system development and higher expression of innate immune-related genes

compared to line 7₂ chickens, while after MDV infection, line 6₃ chickens mount stronger induced immune response by regulating adaptive immunity gene activities. As demonstrated previously, baseline transcription differences were discovered between two lines (1118 DE genes in the spleen, 465 DE genes in the thymus, and 2817 DE genes in the bursa) in tissues from the control (uninfected) birds. Among those genes, examples like *Apoa4*, *C3*, *IL1RL1*, *LOC395914*, *Rag2*, *Smad5*, *CATHL2*, *TLR4*, *TNFRSF1A*, *Cxcl14*, and *IL1B* are involved in immune system development and innate immune response.

Gene expression disparity was also witnessed between these two lines after MDV infection (3852 DE genes in the spleen, 222 DE genes in the thymus, and 1439 DE genes in the bursa). We could notice changes of adaptive immunity genes including *IL18*, *il10*, *IRF7*, *TAP2*, *SWAP70*, *Fas*, *CD28*, *TNFSF13B*, *TNFRSF13C*, *B-MA1*, *CD40LG*, *ada*, *HSPD1*, *PRKCD*, and *IFNG*. For example, if we focused on the DE genes that only existed in the spleen tissue of line 6₃ but were not seen in the spleen tissue of line 7₂ chickens after MDV infection, we would detect two very interesting pathways: toll-like receptor signaling pathway and herpes simplex infection pathway. Therefore, regardless of MDV infection the inborn discrepancy between two lines might contribute to MD resistance mechanism.

Experimental validation of differentially expressed genes

To confirm the RNA-Seq reliability and differential expression among contrasting treatment groups, we randomly selected several DE genes and analyzed by qPCR with dscDNA as the template from sequencing samples. A reference gene with stable expression is necessary to avoid distortions in qPCR, so *GAPDH* was used as an

internal reference [78]. Notably, in the spleen organ, eight genes were examined in both L6₃ and L7₂ chickens in response to MDV infection (Figure 2.7). These eight genes manifested the concordant change direction compared to the RNA-Seq analysis-estimation, and the gene expression difference was verified. In general, the validation outcomes from qPCR assay provide us with further confidence in our analysis.

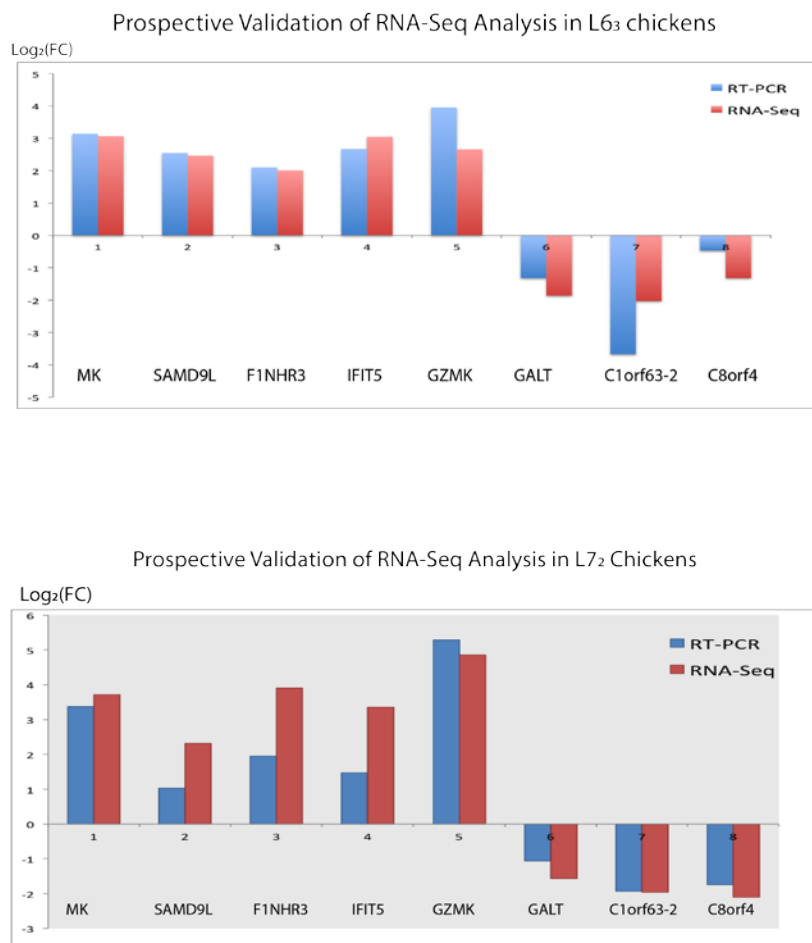


Figure 2.7 Real-time PCR Validation.

Eight genes (five up-regulated genes and three down-regulated genes) identified in spleen tissues of L6₃ and L7₂ chickens were selected for validation. Gene expression difference is confirmed if the direction of change is concordant (as estimated by RNA-Seq analysis).

Discussion

With extensive studies, cytokine production in response to MDV infection and their potential immune functions against MD have been described including IL4 [46], IL6 [326, 327], IL10 [328], IL18 [80, 326, 327], and IFN- γ [41, 329]. For instance, several studies together indicated the lacking association of IFN- γ expression with resistance against MD [327]. Meanwhile, increased IL6 and IL18 expression levels were observed in splenocytes of susceptible chickens. However, those cytokines alone do not adequately manifest interactions among vital immunological mediators and their receptors.

We noticed several interesting DE genes including a set of chemokine receptors and interleukin receptors. For example, chemokine receptors XCR1, CCR2, and CX3CR1 were upregulated in the resistant line's spleen tissue at the late cytolytic phase, and interleukin receptors IL1R2, IL21R, IL4R, and IL11RA underwent dramatic expression changes. The XCL1-XCR1 axis is necessary for efficient cytotoxic immune response mediated by CD8⁺ T cells [330]. Also, CCR2 is indispensable for macrophage-dependent inflammatory reactions and regulates monocyte and macrophage recruitment [331], and has been correlated with delayed AIDS progression in HIV infection [332, 333]. The expression of CX3CR1 appears to induce both adhesion and migration of leukocytes [334]. Considering potential roles of these chemokine receptors in innate immunity and adaptive immunity against HIV infection etc., it's reasonable to infer that upregulation of these chemokine receptors in MDV-infected birds from the resistant line might suggest a stronger anti-viral response compared to the susceptible line. Meanwhile, the over-expressed interleukin

receptor IL1R2 in the resistant chicken spleen indicates regulation of inflammatory response; researchers have revealed IL1R2's function as a proinflammatory factor by activating several inflammatory cytokines' expression [335]. Also in the chicken spleen, the upregulated IL4R (ENSGALG00000006313) promotes differentiation of Th2 cells [336], and IL21R (ENSGALG00000006318) fosters T cells, B cells, and natural killer cells' proliferation and differentiation [337].

Other interesting genes include signal transducer and activator of transcription 1 (STAT1) and leukemia inhibitory factor (LIF). STAT1 was discovered through its involvement in interferon (IFN) signaling and its function in regulating cell growth and differentiation, immune response, antiviral activity, and homeostasis [338]. Challenged with chemical carcinogens, mice lacking STAT1 showed more rapid and frequent tumor development [339]. There was little or no STAT1 expression even after IFN treatment in some tumor cells and tumor-derived cell lines, and STAT1-depleted cells resisted apoptosis in response to TNF or IFN-gamma [339, 340]. ChIP-seq experiments using IFN γ -stimulated HeLa S3 cells indicated the genome-wide binding sites of STAT1 mainly fall into promoters and intronic regions, and suggested the high complexity of STAT1-mediated gene regulatory mechanism [341]. Leukemia inhibitory factor (LIF) was observed to stimulate proliferation of breast, kidney, and prostate cancer cells [342]. LIF inhibited Th17 cells differentiation by exerting an opposite effect on STAT3 phosphorylation, which is required for Th17 cell differentiation, in experimental autoimmune encephalomyelitis mice [343]. When examining STAT1 and LIF expression levels in spleen tissues from the resistant chickens, it is interesting to note that STAT1 was overexpressed (with a logFC of

1.43), and LIF was downregulated (with a logFC -3.33) compared to the non-infected control birds. In resistant chickens, the high level of STAT1 might promote the expression of immune-related genes to fight against MDV while the inhibition of LIF could activate T cell differentiation under the MDV challenge. Since these two genes were only differentially expressed in resistant rather than susceptible chicken lines, their biological functions may contribute to the phenotypic variations induced by MDV infection.

In the three immune organs, we discovered common pathways including cytokine-cytokine receptor interactions and cell adhesion molecules (CAMs) induced by MDV infection. The cytokines and cell adhesion molecules indicate the involvement of innate as well as adaptive inflammatory host defense against MDV infection and immune cell-cell interactions. In fact, most DE genes (as shown in Figure 2.5) and enriched pathways (Appendix I) identified in our study were tissue-specific. For instance, Herpes simplex infection and Influenza A pathways were mainly enriched in line 6₃ chicken spleen tissues, while tight junction and focal adhesion were enriched in line 6₃ chicken bursa tissues. In line 7₂ chickens, pathways associated with metabolism, cell cycle, and DNA replication were highly enriched in three organs induced by MDV infection. This observation is in line with previous findings that networks related to cell-mediated immune response were specifically enriched in line 6₃ chickens [39, 344], broiler and layer chickens [317].

Although in recent years many studies examining gene expression changes related to MD have revealed the interaction between virus and hosts to some extent, the results from similar experiments vary significantly. Admittedly, the complex nature of

Marek's disease introduces confounding sources of variations, such as virus strains, the genetic background of birds and experimental procedures. Hence, it is arbitrary to compare results from experiments performed under different circumstances simply.

In summary, we have carried out a comprehensive analysis of three immune organs' transcriptome in inbred chicken lines showing differential reactions to MD with a focus on the spleen tissue because all stages of the MDV life cycle occur in the spleen [345]. To control the intrinsic transcriptional fluctuation and take full advantage of the similar genetic background of these inbred lines, we designed and performed a pair-wise experiment based on chicken lines and infection time. This methodology not only enabled us to characterize genes highly associated with MD resistance and susceptibility but also reveal a universal impact of MDV infection on the hosts. By using DAVID and IPA analysis, we observed remarkable distinctions between two lines both in the natural state and in response to MDV infection. We discovered enriched networks in metabolism, tissue development, gene expression, and cell signaling. Although our data have provided candidate genes and pathways controlling the different physiological response between two chicken lines, functional studies are necessary to validate impacts of those intriguing genes/pathways and elucidate the underlying mechanism associated with MD resistance.

Chapter 3 : Gene Expression, Chromatin Accessibility and Chromosome Organization in CD4⁺ T Cells Induced by Marek's Disease Virus Infection in Chickens

Abstract

Two inbred line 6₃ and line 7₂ chickens show distinct phenotypes under the infection of Marek's disease virus (MDV). As primary targets of MDV and immune response regulators, CD4⁺ T cells play a critical role in defending against viral infection. It is beneficial to understand epigenetic alterations occurred in CD4⁺ T cells for revealing MD resistance mechanism. To achieve this goal, we performed experiments to measure the differential response induced by MDV infection in CD4⁺ T cells between line 6₃ and line 7₂ chickens in both aspects of transcriptome and epigenome including chromatin accessibility and chromosome organization. We observed that more significant transcriptional activity changes occurred at 10 DPI compared to 5 DPI. Gene ontology analysis showed that at 10 DPI line 6₃ chickens' specific genes were enriched in robust adaptive immune response while there was no enrichment in line 7₂ chickens. The increase of chromatin accessibility levels and chromosome structural changes represented by compartment reorganization were mainly associated with line 6₃ chicken CD4⁺ T cells' up-regulated genes induced by MDV infection at 10 DPI. Our results validate the existence of widespread epigenetic modifications caused by MDV infection. Such changes in resistant chicken lines may contribute to the epigenetic regulatory mechanism for regulation of CD4⁺ T cells' transcriptional activity and control of MD progression outcomes.

Introduction

Chromatin accessibility, determined by “open” or “closed” conformation, is related to regulatory activities of gene expression. A Chromosome’s three-dimensional reorganization has been reported to associate with human and mice diseases. However, studies on the association between those epigenetic changes and poultry health, especially chickens, are limited. Marek’s disease (MD), with a high mortality rate, has resulted in a considerable economic loss to the poultry industry. Since the first discovery of Marek’s disease over 100 years, we have witnessed remarkable progress in vaccine development and understanding of immunity against MD; however, the over usage of vaccines conversely promotes greater virulence and more acute symptoms in susceptible birds. More broad knowledge of MDV susceptibility in chickens requires information from genetic and other environmental factors. The two highly inbred chicken lines 6₃ and 7₂, sub-lines of lines 6 and 7, have been bred since 1939 and served as unique resources to explore the mechanistic difference towards MD response. The availability of chicken genome sequence and next-generation sequencing techniques have altered our capability to identify more critical factors for MD resistance [73, 317].

As target cells of MDV, CD4⁺ T cells play an essential role in viral replication and disease resistance. The study on how CD4⁺ T cells respond to viral infection in the epigenetic aspect would significantly benefit our understanding of the disease resistance mechanism. In the present research, ATAC-seq and Hi-C experiments on chicken splenic CD4⁺ T cells from MD-resistant line 6₃ and MD-susceptible line 7₂ chickens were performed to characterize chromatin accessibility and organization

after MDV infection. We also conducted genome-wide profiling of splenic CD4⁺ T cells using mRNA sequencing (RNA-Seq) with limited cell numbers with the Smart-seq2 protocol [346] to capture the fluctuation in transcriptional activities. Age-matched control chickens without MDV infection were implemented in our experimental design.

Materials and Methods

Experimental animals and chicken CD4⁺ T cell isolation

Two inbred lines of White Leghorn (line 6₃ and line 7₂) were hatched, raised and maintained in Avian Disease and Oncology Laboratory (Michigan, USDA). Chickens from each line were divided into infected (treatment) and non-infected (control) groups respectively. Chickens from the treatment group were injected intra-abdominally with a partially attenuated virulent plus strain of MDV (648A passage 10) at 5 days after hatching with a viral dosage of 500 plaque-forming units (PFU). At 5 and 10 days post-infection, chickens from the treatment and control groups were sacrificed following standard animal ethics and usage guidelines. Samples from the spleen tissue were temporarily stored in the medium (PBS with 2% FBS and 1 mM EDTA) for CD4⁺ T cell isolation preparation. The spleen was meshed and filtered through a 70- μ m mesh nylon strainer to make a cell suspension in medium (PBS with 2% FBS and 1 mM EDTA). Pellets were collected by centrifuge at 300g for 5 minutes. 5 ml of ACK buffer was added to lyse blood red cells for 5 minutes at room temperature. The cell pellets were recollected by centrifuge at 300g for 5 minutes. Then the cells were resuspended at a cell concentration of 2×10^8 cells/ml. 5%

Normal mouse serum (Cat. 0050-01, SouthernBiotech) was added to block non-specific antibody binding. 2 μ g of mouse anti-chicken CD4-PE antibody (Cat. No. 8210-09, SouthernBiotech) was added and mixed well. The binding was complete by incubating at room temperature for 15 minutes. Excess antibody was washed away by one time of wash with 10-fold excess medium (PBS with 2% FBS and 1 mM EDTA). Cells were resuspended to original volume for next selection step. EasySep PE Selection Cocktail from EasySep Positive Selection, PE Selection Kit (Cat. #18557, STEMCELL Technologies) was added at 100 μ l/ml cells and mixed well. Second binding was complete by incubating at room temperature for 15 minutes. EasySep Magnetic Nanoparticles from EasySep Positive Selection, PE Selection Kit (Cat. #18557, STEMCELL Technologies) were added at 50 μ l/ml cells and incubated at room temperature for 10 minutes. The splenic CD4⁺ T cells were enriched after three times of wash using a magnetic stand. After the last wash step, the tubes were removed from the magnet and cells were resuspended in an appropriate amount of medium (PBS with 2% FBS and 1 mM EDTA). The CD4⁺ T cell purity was accessed by flow cytometry, and an average 90%-95% purity was achieved. The positively selected CD4⁺ T cells were then processed differently according to the following protocol's specific requirements.

RNA preparation and sequencing

Approximately 50,000 isolated splenic CD4⁺ T cells were homogenized and stored in DNA/RNA Shield (Cat. # R1100, Zymo Research) for total RNA extraction. Total RNA from splenic CD4⁺ T cells in infected and non-infected groups was extracted using Quick-RNA MicroPrep Kit (Cat. # R1050, Zymo Research, USA) according to

the manufacturer's instruction. Polyadenylated RNA was synthesized following the instructions in the Smart-seq2 protocol [346]. About 100ng of double-stranded cDNA was used for sonication to obtain random size DNA fragments. The dsDNA fragment ends were repaired by T4 and Klenow DNA polymerase and underwent library construction procedures following Illumina protocol instructions. Each library was identified by adding a 6-bp adaptor and sequenced at 50-bp read by an Illumina HiSeq 3000 Sequencer, as previously described [39].

CD4+ T cell ATAC-seq

Approximately 50,000 isolated CD4+ T cells were cross-linked by adding formaldehyde to a final concentration of 1% and rotated at room temperature for 10 minutes. 0.125M final concentration of glycine was added to stop the cross-link reaction. The cell pellets were collected by centrifuge at 300g for 5 minutes at 4 °C. The cell pellets were washed twice with ice-cold PBS and stored in a -80°C freezer for later usage. The cells were thawed on ice and transposition reaction using Tn5 transposase was conducted on the resuspended nuclei at 37 °C for 30 minutes, following the ATAC-seq protocol [238]. EDTA was added to stop transposition reaction, and nuclei were incubated with proteinase K at 65 °C overnight to reverse-crosslink. Cell DNA was purified using the conventional phenol-chloroform precipitation method. The transposed DNA fragments were amplified using PCR with 8-bp adaptors added on the ends. Filtration in a 2% agarose gel was used to select PCR product fragments (DNA plus adaptors) from 200 to 500 bp. After purification, DNA quality was assessed by Qubit assay (Life Technology, USA) and was adjusted

for sequencing on the Illumina HiSeq 3000 Sequencer following manufacturer protocols.

CD4⁺ T cell 3e Hi-C assay

The three enzyme Hi-C (3e Hi-C) was utilized according to in situ Hi-C [288] with some modifications. Isolated CD4⁺ T cells were cross-linked with 1% formaldehyde for 10 minutes at room temperature. 1-2 million cells were lysed in 10 ml lysis buffer with rotation at 4°C for one hour and then incubated with 400 μ L 1X NEB cutsmart buffer with 0.1% SDS at 65°C for 10 minutes. Cell chromatin was afterward digested with 20 Units CviQ I (NEB, USA) and 20 Units CviA II (NEB, USA) at 25°C for 20 minutes, then with 20 Units Bfa I (NEB, USA) at 37 °C for 20 minutes. 600 μ L wash buffer (10 mM NaCl, 1 mM EDTA, 0.1% Triton X-100) was used twice to wash the cells. The DNA ends were marked with biotin containing dCTP, dTTP, biotin-14-dATP, and ligated using T4 DNA ligase. After 65°C overnight reverse crosslink, the samples' DNA biotin labels were removed using T4 DNA polymerase. Sonication (Diagenode, USA) was used to fragment DNA into 300-500 bp. Fragmented DNA was end-repaired and A-added. Dynabeads MyOne Streptavidin C1 beads (Invitrogen, USA) were used to enrich the biotinylated DNA fragments according to manufacturer instructions. Illumina paired-end adaptors were ligated to the DNA on beads, and the enriched DNA was amplified using PCR with appropriate parameters. 300 to 700 bp DNA fragments were selected using 2% E-gel and sequenced on Hi-Seq 3000.

Whole genome gene expression analysis

Mapping and Data Visualization. The raw RNA-Seq reads were mapped onto the chicken reference genome using Tophat [347] (by selecting options `–transcriptome` only and allowing 5 multi-hits). The gene annotation files for the chicken reference galGal4 was downloaded from UCSC. To quantify gene expression for a fair comparison among libraries, we computed RPKM values (Reads Per Kilobase of transcript per Million mapped reads) for each gene using a Python script `rpkmforgenes.py` [348]. Genes with RPKM values <1 were excluded from the following analysis. The function *prcomp* in the R statistic package was used to perform PCA analysis to detect the existence of batch effects and testing reproducibility of the data. For data visualization, we transformed the mapped bed files to bedgraph files using a window size equal to 20, and we uploaded the bedgraph files to the WashU Epigenome Browser [349].

Differential Expression, Clustering, and Ontology Analysis. The transcript numbers of genes were used as an input for the R package edgeR [320] for computing the differential genes using two replicates. In order to get a better fit for the assumed statistical distribution in edgeR, we only considered the expressed genes as mentioned above. The thresholds of FDR and the fold change were set as 0.01 and 1.5, respectively. We applied the K-medoids (K=6) clustering method to the gene expression data for all four conditions. To perform clustering, we computed the variance of gene expression across all four experimental conditions. Genes with dispersion, defined as $sd/mean$, lower than 1 were filtered out. Then log-transformation was used for the gene expression data, and the clustering results were visualized by plotting the z-score values of gene expression across experimental

conditions. Gene ontology analysis using the online Gene consortium platform [350] was used on the list of differentially expressed genes.

Alignment and peak identification of ATAC-seq data

We mapped the raw ATAC-seq data onto the chicken reference genome (galGal4) using Bowtie2 with default parameters (MAPQ>10). Peak enriched regions, which are mainly DNase hypersensitivity sites (DHSs), were identified using SICER with E-value cutoff =0.01. We combined any two of the discovered DHS sites if they overlap > 1bp from different experimental conditions. To determine various DHS sites, we computed the number of reads in each of the combined DHSs for each library. We calculated the quantity CPM (count per million base pair) divided by peak lengths to quantify chromatin accessibility of those regions. Using the raw read counts in the DHSs we applied the same strategy as the RNA-Seq data to identify the differential DHSs.

3e Hi-C data mapping and analysis

Mapping, Visualization, and Normalization. We mapped the raw Hi-C data onto the chicken reference genome (galGal4) using Bowtie2 (the pair-ended mode allowing discordant pairs and MAPQ>30). We applied two ways to visualize the Hi-C interaction data. First, we used a bin with size = 1kb among the reference genome; we merely kept the PETS between any bin pairs if the numbers of PETS between the two bin pairs were more or equal to 2. We transformed the bed files to WashU visualization format and visualized the interaction data using WashU Epigenome Browser [349]. Second, we divided the genome into bins with equal sizes (bin size

=40kb), such that an interaction matrix was obtained. To normalize our Hi-C data for further analysis, we first filtered out the distance factor in the interaction matrix by dividing an expected interaction matrix. Then, we applied the matrix balancing method (KR-matrix balancing) regressing out the other elements.

A/B compartments and Flipping. In order to compute the A/B compartments for the interaction matrices, we considered each row of the normalized matrix as a variable containing different values. We interpreted these different values as the interaction profiles of the genome location corresponding to the bin of the row. Then we computed the Pearson correlation between any two rows such that we obtained a correlation matrix with the same dimension as the interaction matrix. PCA analysis was applied to the correlation matrix, in which the first principal component is the indicator for the A/B compartment. We consider those bins that have the switching of the sign of PC1 vector to be the flipping of A/B compartments. In particular, the A compartment is defined as related to euchromatin regions, and the B compartment is defined as relevant to the heterochromatin regions. By combining with the ATAC-seq data, we recognized the flipping from A to B or B to A.

Results

Transcriptional activity changes caused by MDV infection

We sampled two time-points of MD progression (5 and 10 DPI) by including 4 biological replicates in each treatment condition. The sequencing reads from chicken CD4⁺ T cell RNA-Seq were analyzed and the mapping statistics of libraries is summarized in Appendix II. The average mapping rate was about 65%. Using small

number of cells, we could still capture the transcriptional activity of genes and notice the expression differences induced by MDV infection in CD4⁺ T cells (Appendix III). We computed differentially expressed genes between same lines with and without infection and compared the up-regulated and down-regulated gene lists among different chicken lines for different infection days (i.e., 5 DPI and 10 DPI) (Figure 3.1). In general, our detected gene expression levels for CD4⁺ T cells at 5 DPI were much lower compared to those at 10 DPI, and more transcriptional activity changes were observed at 10 DPI regardless of chicken lines. There were 570 differentially expressed genes (319 up-regulated and 251 down-regulated) in line 6₃ chickens induced by MDV infection at 10 DPI, and there were 303 DEGs (224 up-regulated and 79 down-regulated) in line 7₂ chickens. The genes' overlapping information can be found in Figure 3.1.

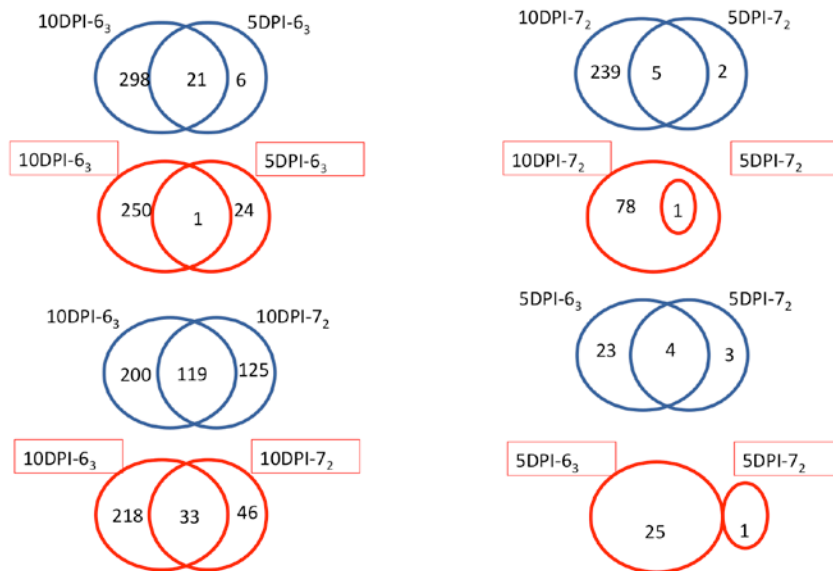


Figure 3.1 Venn Diagram of DEG Lists

(In collaboration with Dr. Wai Lim Ku from Dr. Keji Zhao's lab in NHLBI, NIH)

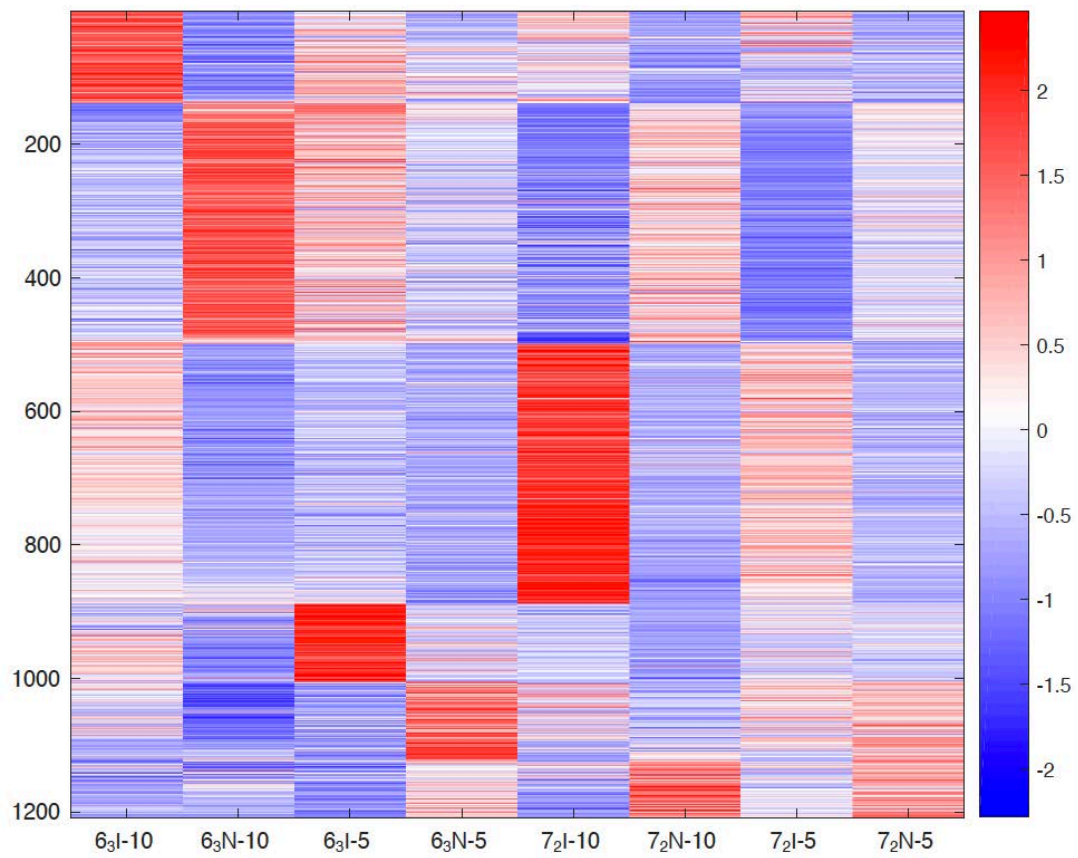
Each set in this figure stands for the number of differentially expressed genes comparing within the same line with and without MDV infection. Blue sets indicate the up-regulated genes in infected groups and red sets stand for the down-regulated genes in infected groups. For instance, there were 319 (298+21) up-regulated genes when comparing line 6₃ infected chickens to line 6₃ non-infected (control) chickens at 10 DPI. The Venn diagram shows the overlapping gene numbers identified between different sets as indicated in the figure.

Genes related to immune response were highly enriched in line 6₃ chickens

In order to visualize the dynamic transcriptional activities initiated by MDV infection, the hierarchical clustering of gene expression changes was plotted. We extracted genes with dispersion larger than 1 over all conditions and performed the K-medoids clustering (K=6) (Figure 3.2). It was obvious that some genes' expression levels increased dynamically over time as shown in the cluster. In line 6₃ chickens, the infection of MDV caused a dramatic change in the gene expression profile from 5 DPI to 10 DPI, while the patterns in line 7₂ chickens remained constant regardless of MDV infection. Line 6₃ chickens are resistant to MDV infection without the development of tumors, and it may be due to the lack of MDV latency establishment around 2 weeks after infection. Hence, we mainly focused on 10 DPI, representing the latent stage of MD progression, and conducted functional annotation of the differentially expressed genes identified in CD4⁺ T cells of line 6₃ and line 7₂ chickens at 10 DPI (Figure 3.2). In line 6₃ chickens, the most significantly enriched pathways were associated with the substantial adaptive immune response including lymphocytes proliferation and activation, interleukin production, and regulations of cytokine production. However, those immune pathways were not observed in line 7₂ chickens after MDV infection, and only metabolic processes were enriched. It is

reasonable to infer that the failure of line 7₂ chickens to induce potent immune response after MDV infection may lead to the expansion of MDV during recovery from latency.

A



B

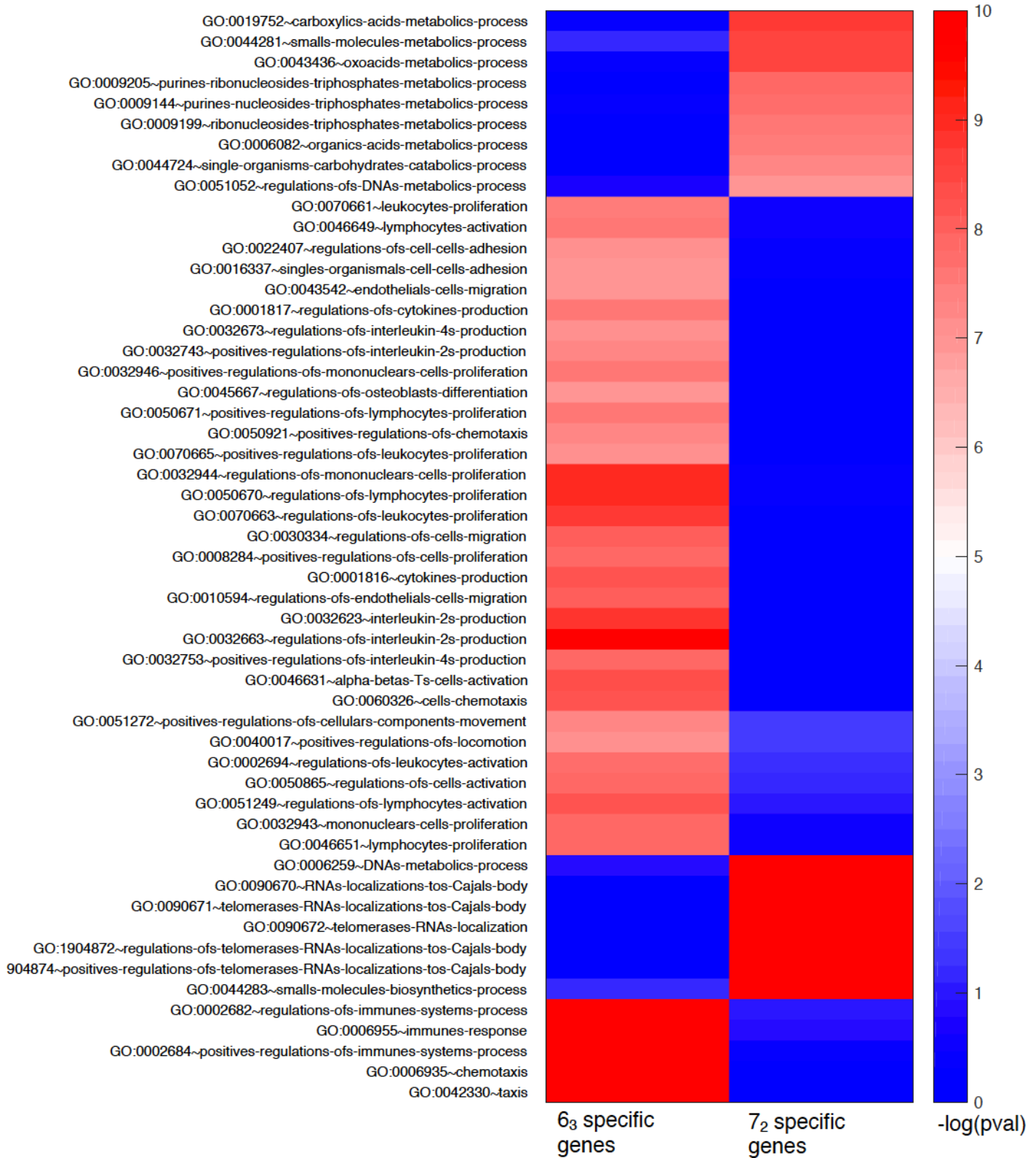


Figure 3.2 Hierarchical Clustering and Gene Ontology

(In collaboration with Dr. Wai Lim Ku from Dr. Keji Zhao's lab in NHLBI, NIH)

A) Unsupervised clustering reveals striking patterns of dynamic gene expression at each time-point. Each row indicates a single gene and the color represents its relative expression level. Genes with dispersion larger than 1 over all conditions are included in K-medoids clustering ($K = 6$). Z-score is computed across conditions and clustered as shown. (B) Gene Ontology analysis shows the enriched pathway of CD4⁺ T cells in lines 6₃ and 7₂ chickens after MDV infection. The color represents the $-\log$ of the p-value. A high value indicates the specific genes are more enriched in the gene ontology terms (redder).

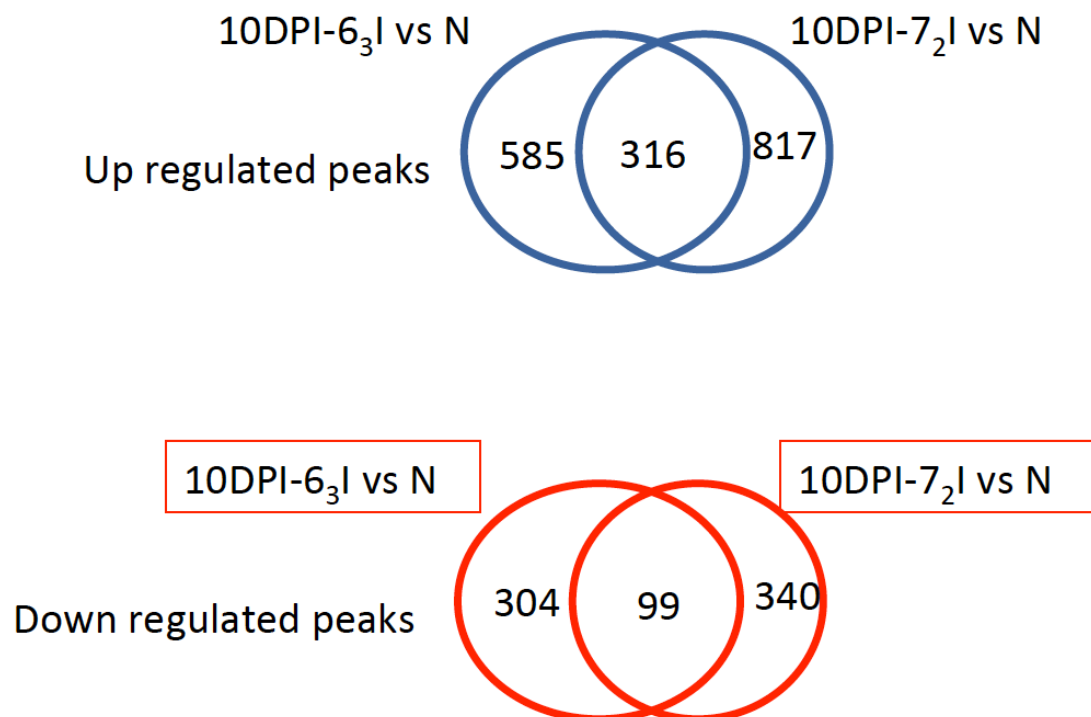
Correlation between chromatin accessibility and differential gene expression

To reveal the potential roles of chromatin accessibility on gene expression associated with MDV infection, we generated ATAC-seq data using the next-generation sequencing in chicken CD4⁺ T cells with similar experimental settings as RNA-Seq. The unique mapped reads were used to identify the peak enrichment regions, which are mainly DNA hypersensitive sites (DHSs). Up-regulated and down-regulated peaks were computed by comparing the same line chickens with and without MDV infection. Our ATAC-seq data generated enough unique reads to characterize the genome-wide DHS changes (Appendix III and Appendix IV). Here we mainly focused on 10 DPI for two reasons: 1) 10 DPI is a critical time-point for establishing latency in susceptible chicken lines, and no significant enrichment of virus load is observed in resistant chickens and 2) Our RNA-Seq data suggested a more interesting gene expression pattern in 10 DPI chicken lines, and we hoped to study the relationship between chromatin accessibility and transcriptional activities. Here we identified 901 up-regulated and 403 down-regulated peaks in line 6₃ chickens (infected vs. control) at 10 DPI and 1133 up-regulated and 439 down-regulated peaks in line 7₂ chickens (infected vs. control) at 10 DPI (Figure 3.3). There were 316

overlapping up-regulated and 99 overlapping down-regulated peaks between line 6₃ and line 7₂ chickens induced by MDV infection (Figure 3.3). We named those genes that were only up-regulated in one treatment condition as that treatment-specific genes. In order to find the relationship between chromatin accessibility and gene expression, we annotated 4 groups of genes (i.e., L6₃ non-infected specific genes, L7₂ non-infected specific genes, L6₃ infected specific genes, and L7₂ infected specific genes) with the DNA hypersensitive sites (DHSs). Then we plotted the chromatin accessibility density in each condition for those treatment-specific genes (Figure 3.3). For line 6₃ non-infected specific genes (specifically down-regulated genes when comparing line 6₃ infected to non-infected chickens), chromatin accessibility of CD4⁺ T cells in both line 6₃ and line 7₂ non-infected (control) chickens was higher ($P < 0.01$). For both line 7₂ non-infected specific genes (i.e., specifically down-regulated genes induced by MDV infection) and line 7₂ infected specific genes (i.e., specifically up-regulated genes induced by MDV infection), chromatin accessibility of those genes in CD4⁺ T cells did not change significantly ($P > 0.05$). Interestingly, line 6₃ infected specific genes (i.e., specifically up-regulated genes induced by MDV infection in line 6₃ chickens) had significantly higher accessibility in line 6₃ infected chickens compared to non-infected chickens ($P < 0.01$); such difference was not observed in line 7₂ chickens ($P > 0.05$). Among those comparisons, we found that only line 6₃ specifically up-regulated genes induced by MDV infection were accompanied by the increase of chromatin accessibility levels. Gene expression regulation involves a wide range of mechanisms and sophisticated programs including DNA methylation, histone modifications, non-coding RNAs, nucleosome

positioning, and chromatin architectures etc. It is reasonable to infer that no correlation between other chicken lines' specific genes (except line 6₃ infected specific genes) and chromatin accessibility is the outcome of the complex gene regulation networks.

A



B

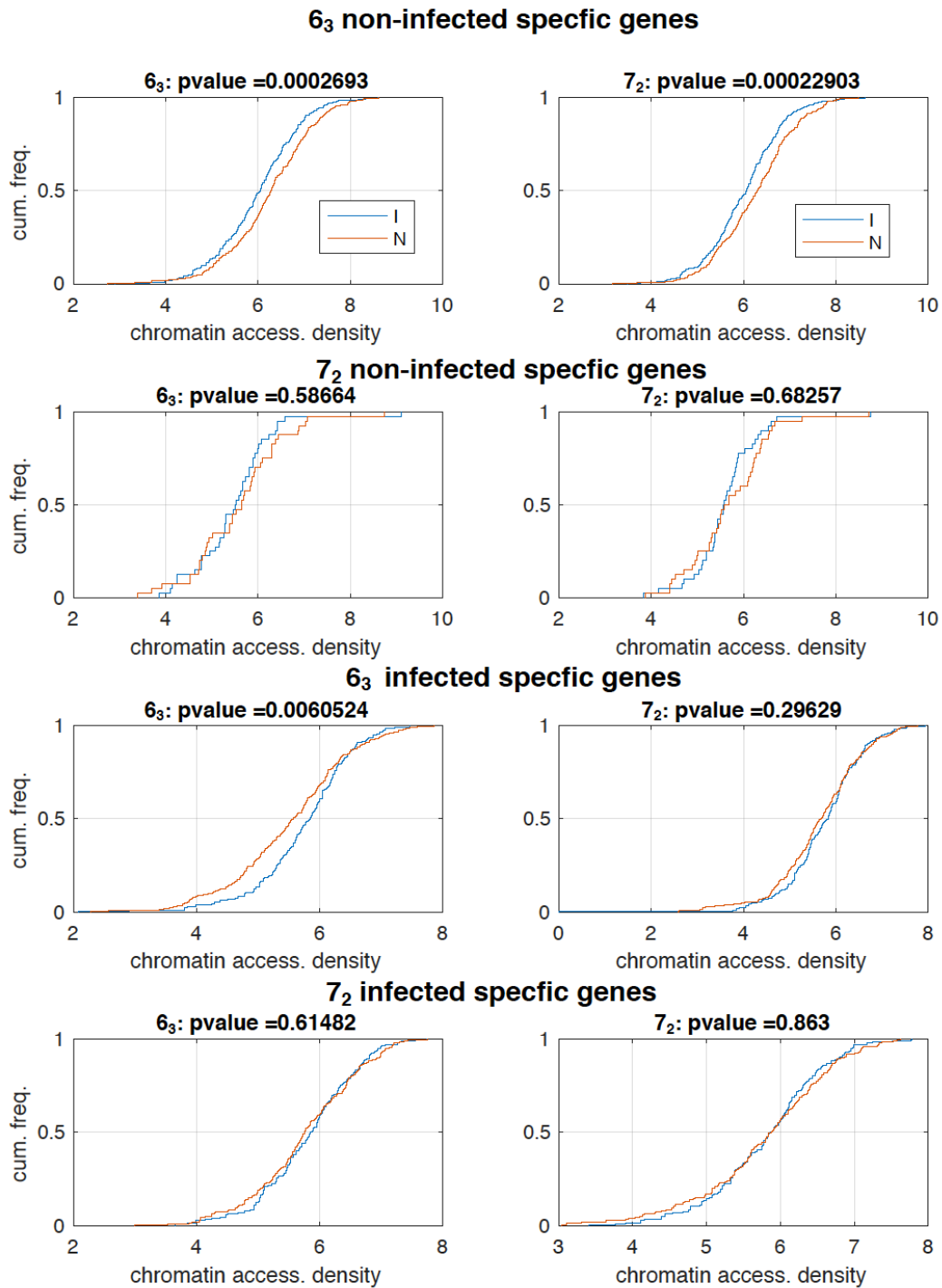


Figure 3.3 Chromatin Accessibility and Gene Expression

(In collaboration with Dr. Wai Lim Ku from Dr. Keji Zhao's lab in NHLBI, NIH)

(A) Each set in this figure stands for the number of up-regulated (blue) or down-regulated (red) peaks comparing within the same line with and without MDV infection (infected vs. control). The Venn diagram shows the overlapping peak numbers identified in each set as indicated in the figure. (B) CDF plots of chromatin accessibility between infected and non-infected chickens for specific gene sets. Specific genes indicate their expression levels are higher in that treatment condition. The left panel stands for line 6₃ groups, and the right panel represents line 7₂ groups. Infected groups are annotated with blue lines, and non-infected groups are colored red.

Gene expression and chromatin organization

To investigate if chromatin reorganization could happen after MDV infection, primarily at 10 DPI, we performed Hi-C experiments on line 6₃ and line 7₂ chickens with and without MDV infection using two biological replicates (Appendix IV). The reproducibility between two biological replicates was high (Figure 3.4). By further analysis of A/B compartments, we observed 2317 numbers of A/B compartment flipping events whole genome-wide from two chicken lines. Some A/B compartment flips were only witnessed in CD4⁺ T cells of line 6₃ chickens (infected vs. control) (Figure 3.5). In general, the A compartment is often associated with open and active chromatin profiles and the B compartment is often related to repressive chromatin accessibility. Here we focused on the A/B compartment flipping caused by MDV infection in line 6₃ chicken CD4⁺ T cells and separated genomic regions based on repressive to active (B to A) flips and active to repressive (A to B) flips in the non-infected (control) group. For repressive to active flipping regions that occurred in non-infected line 6₃ chickens, we did not observe gene expression changes induced by MDV infection (Figure 3.5). For non-infected line 6₃ chickens active to repressive flipping regions, gene expression levels were higher in line 6₃ infected chickens compared to non-infected controls, although the P-value was not that significant ($P =$

0.073). This observation implies that the open chromosome structure in line 6₃ chickens induced by MDV infection, which is indicated by the active to repressive flipping in non-infected controls, was correlated with local higher gene expression levels. This finding is in line with our previous observation that the expression of specifically up-regulated genes in line 6₃ chickens was related to the increase of chromatin accessibility levels.

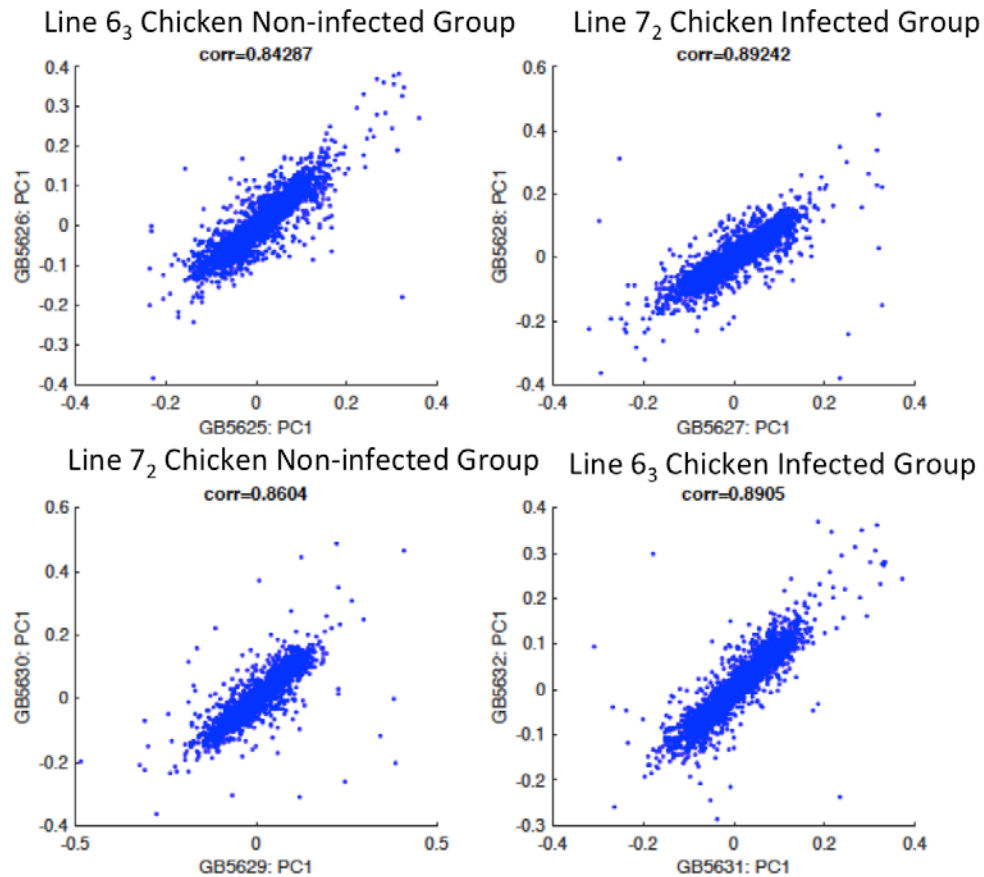
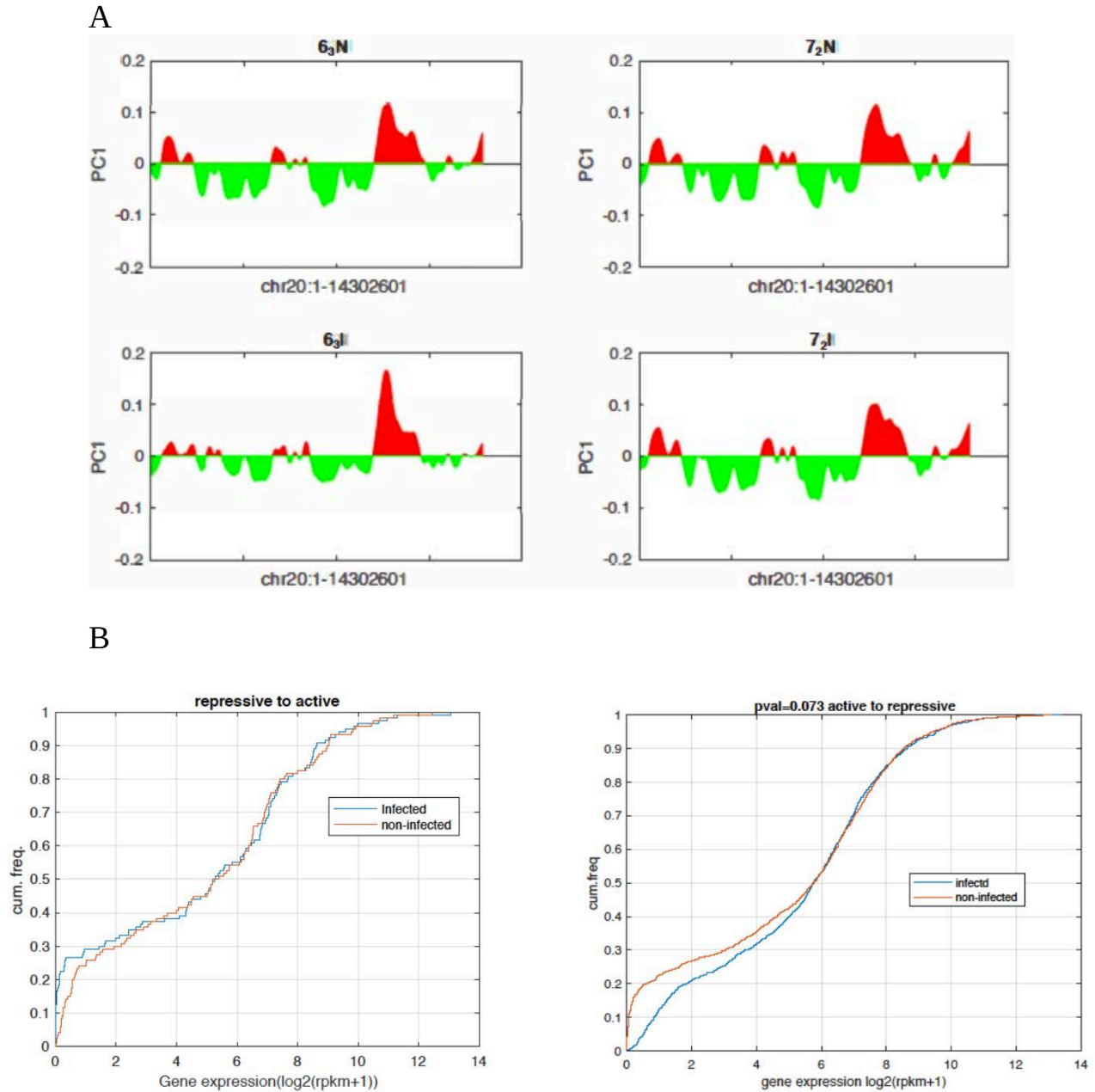


Figure 3.4 Reproducibility of Hi-C Data

(In collaboration with Dr. Wai Lim Ku from Dr. Keji Zhao's lab in NHLBI, NIH)
This figure shows the correlation of the first principal component of Hi-C data for two biological replicates. Compartment analysis is performed for the Hi-C data in both infected and non-infected chickens. Scatter plots for the PC1 values obtained from the compartment analysis are plotted between two replicates.



Discussion

ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) is a technique used in molecular biology to characterize chromatin accessibility levels, based on the genome-wide sequencing of regions sensitive for Tn5 transposase insertion. Identifying enrichment peaks from ATAC-seq is the primary step to find candidate regulatory regions. Hi-C is widely used to identify chromatin interactions in the nucleus comprehensively. Although these techniques have been extensively used in mammals and yeasts, their usage in vertebrates, especially chickens, is limited. We hoped to study the relationship between chromatin structures (including accessibility and three-dimensional organization) and gene expression induced by Marek's disease virus infection. Consequently, we chose CD4⁺ T cells, the primary targets of MDV, as our research focus, and applied these techniques to our unique chicken models (line 6₃ and line 7₂ birds).

In total, thousands of significantly enriched peaks were identified within the genome of CD4⁺ T cells induced by MDV infection. Transcriptional activity changes induced by MDV infection in line 6₃ resistant chickens suggested a boost of robust immune response against the viral infection, which was not observed in line 7₂ susceptible chickens. This phenomenon might help explain the intrinsic difference between these two chicken lines since the strong immune response in line 6₃ chickens play a critical role in restraining viral replication after the latency phase. Line 7₂ chickens remained "blind" to the virus intrusion and became vulnerable to viral replication and tumor formation after the latency phase. The gene expression alterations were correlated with chromatin accessibility changes in line 6₃ chickens, and our data identified 1304

differentially enriched peaks induced by MDV infection in line 6₃ chickens. Although we observed some connections between chromatin accessibility and gene expression, the lack of complete chicken protein binding motif database stopped us from searching for potential transcription factors controlling the difference between line 6₃ and line 7₂ chickens. The differentially expressed gene lists provide us many essential candidate genes responsible for MD resistance; however, factors in the center of the complex regulatory network remain unknown. Hopefully, with the development of more chicken genome databases including protein-binding motifs in the near future, we could explore vital regulatory factors influencing MD resistance beyond the general correlation between chromatin accessibility and gene expression.

Our Hi-C data suggested that the global chromosome reorganization was induced by MDV infection, especially in line 6₃ resistant chickens. Since viral infection could cause changes of chromatin domains in general, it is reasonable to believe some alterations of local interactions (including promoter-enhancer-interactions) would influence the expression of essential immune-related transcription factors. At the most basic level, TADs (topologically associating domains) stand for the genome physical structure. Usually, within the TAD the connections are self-associated, and between the barrier regions in neighboring TADs there is an “insulation layer”. Reports show that there are relatively fewer interactions across the TAD boundaries, and housekeeping genes are typically found at the boundaries. The stability of TADs is critical for maintaining the normal physiological activities of the cell. In our study, the A/B compartment flipping initiated by MDV infection may happen in TADs or even most local levels consisting individual “loops” [288] or “insulation

neighborhoods” [289]. This hypothesis is in line with previous findings that sub-TADs, loops, and insulation neighborhoods seemed to vary between different cell lineages [288, 289] and were involved in cell-type-specific regulatory events. Some researchers believe that TADs are the large and invariant scaffold of genome structure within which cell-type-specific sub-domains (including loops and insulation neighbors) can adjust their dynamic structure to regulate the lineage-specific gene expression [351]. The chromosome structure changes observed in line 6₃ resistant chickens could provide us critical information about potential promoter-enhancer interactions. These interactions may influence the transcription factor binding and the expression of essential immune-related genes. The interplay among these factors may eventually determine MDV infection outcomes.

In summary, our study identified more significant transcriptional activities in chicken CD4⁺ T cells at 10 DPI compared to 5 DPI. At the transcriptional level, pathways related to robust immune response were enriched in line 6₃ chickens rather than line 7₂ chickens at 10 DPI. Increased chromatin accessibility and substantial chromosome structure changes were associated with specifically up-regulated genes in line 6₃ chicken CD4⁺ T cells induced by MDV infection at 10 DPI. Our analysis confirmed the extensive epigenetic alteration caused by MDV infection and indicated the potential epigenetic regulation on CD4⁺ T cells’ transcriptional activities. Future directions of this study may fall into the following aspects: functional validations of both chromosome structure changes and transcriptional activities in chicken CD4⁺ T cells related to MDV infection. Fluorescence in situ hybridization (FISH) would be a useful technique to reveal the spatial organization of specific regions. Also,

conditional gene knockout chickens could be used to validate genes' biological functions under physiological conditions after MDV infection.

Chapter 4 : Dynamic Epigenetic Status and Plasticity during Mouse CD4⁺ T Cell Differentiation

Abstract

As a subset of T lymphocytes, T helper cells can regulate other immune cells by releasing cytokines to mediate the immune response. Current T cell differentiation models invoke separate subsets governed by lineage-specifying transcription factors. However, the knowledge on the plasticity of Th1 cell lineage commitment is not complete. Here we show that in the presence of Treg lineage-defining cytokines, Th1 cells could reprogram to adopt a T-bet⁺Foxp3⁺ “Th1+Treg” phenotype in a time-dependent manner. We discovered a transition stage between 48 hours (plastic) and 72 hours (committed) during the Th1 differentiation path. Interestingly, T-bet was not the direct factor determining the Th1 cell reprogramming efficiency. DNase-Seq, RNA-Seq, and chromosome interaction data at different time points indicated genome-wide chromatin accessibility changes associated with differentiation time. Integration of multiple data sets allowed inference of potential Th1 regulators, including IRF4, as key determinants of cellular plasticity. Further conditional deletion of IRF4 in CD4⁺ T cells enabled Foxp3 expression in Th1 *in vitro* cultures when subsequently exposing Th1 like cells to Treg-skewing conditions. Our findings provide a deeper understanding of Th1 cell plasticity and have obvious implications for disease pathogenesis and new therapeutic applications to fight infection and to control autoimmune and allergic inflammatory disorders.

Introduction

CD4⁺ T cells' function in adaptive immunity and immune-mediated diseases, including autoimmunity, asthma, allergic responses, and tumor immunity [86], has been widely recognized by assisting B cells to produce antibodies, stimulating CD8⁺ T cells responses, mediating macrophage functions, and suppressing autoimmunity related to the magnitude and persistence. Unusual fluctuation in CD4⁺ T cell numbers or functions may result in vulnerability to severe disease disorders. Under environmental changes, naive CD4⁺ T cells have the capability to differentiate to distinct functional T helper cell subsets, including Th1, Th2, Th17, iTreg, depending on particular TCR activation. After the first discovery of Th1/Th2 scenario, different cytokine profiles and master transcription factors of each T helper cell subset have been characterized. In particular, IFN- γ is the signature cytokine of Th1 cells with the expression of T-bet; IL-4, IL-5, and IL-13 are the signature cytokines of Th2 cells with the expression of GATA-3; Th17 cells have the signature cytokines of IL-17A, IL-17F, and IL-22 with the master regulator ROR γ t. The induced Treg cells [95-97], which can be induced from naive CD4⁺ T cells, share similar characteristics of naturally occurring regulatory T cells developed in the thymus and different from the naive CD4 T cell progenitors. The expression of Foxp3 is critical for the normal Treg cell function as an immune-suppressor by suppressing induction and proliferation of effector T cells. The deletion of Foxp3 could result in the lack of functional Treg cells in the periphery and the occurrence of systemic autoimmunity [145].

The simplistic dichotomy of Th1-Th2 paradigm underestimated the complicated flexibility of CD4⁺ T cells [172, 173]. The initial concept that effector T cells were

distinct committed lineages was based on the stable signature cytokine production. For example, the Th2 cell cytokine IL-4 could repress IL-12 signaling and thus antagonize Th1 differentiation [177]. Studies on Th1/Th2 plasticity in mouse and human cells have revealed the relative flexibility of early Th1 cells compared to rapidly committed Th2 cells, which may be the outcome of early loss of IL-12 receptors in Th2 cells [181]. TGF- β profoundly determines the balance between Th17 and iTreg cells since after *in vitro* ROR γ t and Foxp3 induction cells have the dual potential to differentiate into either the Th17 or Treg cell lineage relying on the TGF- β level [176]. Thymus-derived CD4⁺CD25⁺Foxp3⁺ regulatory T cells could differentiate into Th17 cells after exposure to IL-6 [185], and such physiological Treg conversion into effector T cells might fight against specific viruses, microbial and fungal pathogens *in vivo* [176]. Transcription factors' dynamic expression across different subsets, such as IFN- γ production through T-bet induction in Treg and Th17 cells [154, 161, 166, 188, 191], T-bet and GATA3 coexistence in Treg cells to suppress ROR γ t expression and maintain immune tolerance [153], has implied CD4⁺ T cell plasticity. A model similar to the energy levels of an electron in the atom [186] has been proposed to explain the plastic relationships between CD4⁺ T cells, i.e., the transition is more likely to occur from a higher-energy (unstable) to a lower-energy (stable) state between different CD4⁺ T cell subsets. Although this model provides us with a broad view of CD4⁺ T cell plasticity scenario, the underlying mechanism such as the molecular basis remains unclear.

Recent studies have explained the mechanism for T helper cell plasticity to some extent with the focus on the cellular microenvironment, the interaction networks

between transcription factors, and epigenetic modifications. Transcription factors may function as the upstream, downstream, or cofactors of those lineage-specific master regulators to form a regulatory network controlling cell fate and plasticity. Genome-wide studies such as the Th17 cell differentiation network [195, 196] discovered core Th17 regulatory systems associated with cellular plasticity including Fosl2. Conditional deletion of Fosl2 in CD4⁺ T cells enabled IFN γ production in Th17 and Th2 *in vitro* cultures, especially when subsequently exposing Th17 cells to Th1-skewing conditions [195]. The balance of transcription factor networks helps maintain the cell function and stability while extracellular or cell-intrinsic fluctuation on core regulatory nodes and pathways could alter cell fate in a cell type- and stage-specific manner. The differentiation and plasticity of CD4⁺ T cells are also coupled to epigenetic changes in chromatin structure, histone and DNA modifications [199-201]. The bivalent modifications (both permissive H3K4me3 and repressive H3K27me3 marks) exist in the promoters of many regulatory genes, including Tbx21, Gata3, Rorc, and Bcl6, etc. [202, 203]. This phenomenon may indicate the cells' readiness for adopting other phenotypes in response to cellular microenvironment changes [176]. The bivalent marks at Tbx21 (encoding T-bet) in Th2 cells, Gata3 in Th1 cells, Tbx21, Gata3 and Rorc in Treg cells, and Tbx21 and Gata3 loci of *in vitro*-generated Th17 cells, may initiate lineage switch upon transient extracellular signals during and after differentiation. Disturbance of the equilibrium between master regulators may lead to cell heterogeneity and flexibility, as reported in Treg cells [153]. However, to the best of our knowledge, the answer to what factors control the lineage cellular

plasticity is not clear, and it is necessary to understand the lineage commitment status and evaluate cellular plasticity under environmental stimuli.

The chromosome structures like chromatin looping and interactions could influence gene transcriptional activity during cellular differentiation and disease development. In eukaryotes, the genomes are organized into topologically associating domains (TADs) and lamina-associated domains (LADs). The involvement of three-dimensional reorganization is observed in diseases such as breast cancer, colorectal cancer, and leukemia. The 3C-based techniques, including Hi-C (an unbiased genome-wide chromatin interaction mapping technique), have been widely used from yeast to mammals to decipher the chromatin architecture. The 3C-based methods involve crosslinking and ligation of interacting regions, which could contribute to the background noise coming from the ligation of random contacts between chromatin and also the biases in restriction enzyme cutting. Here, we introduce a new high-resolution technique for mapping genome-wide interactions among regulatory regions without chromatin fragmentation and ligation. We named this technique as Transposase-mediated Analysis of Chromatin looping (TrAC-looping). This method utilizes the concept of transposition, a nuclear process that results in genomic rearrangements like insertions and deletion. Transposases could recognize a pair of 19 bp Mosaic End (ME) sequences to initiate transposition. When only one 19 bp ME oligonucleotide sequence is present, transposition results in fragmentation and jointing of target DNA to the ME-containing adaptor, which is the principle of ATAC-seq to label the open chromatin regions. The long DNA spacer between the MEs makes it possible to dimerize the transposase-ME complex. With the power of

transposases, we designed a “bivalent ME linker” containing a pair of MEs using a 30 bp oligonucleotide spacer to detect genomic chromatin interactions. The short spacer does not allow intramolecular dimerization of the transposase-ME complex and thus tend to form a tetramer complex through the bivalent MEs. The two MEs could be inserted into the genome in *cis* or in *trans*. Inserting one ME into each of the two interacting chromatin regions would create a “cross-linking” of the interacting chromatin regions by the oligonucleotide linker. We applied the transposase-bivalent ME linker complexes to formaldehyde-crosslinked cells to allow transposition and covalent linking formation within two interacting chromatin regions. The conjoined DNA products were reverse cross-linked, purified and digested with *Nla*III and *Mlu*CI. The biotinylated bivalent linker-containing DNA fragments were isolated using streptavidin beads; the bound DNA was eluted and circularized; the self-circularized DNA was amplified and sequenced on next-generation sequencing platforms.

In this study, we examined the differentiation process from naive CD4⁺ T cells to Th1 cells and tested Th1 cells’ capability to alter phenotypic characteristics under secondary environmental stimuli *in vitro*. We differentiated naive CD4⁺ T cells in Th1 cell-polarizing cytokines for different time points (24hrs, 48hrs, 72hrs) and then switched to Treg cell-skewing conditions to test Foxp3 induction as a measurement for cellular plasticity. Th1 effectors, when fully committed, could not be converted to Foxp3⁺ Treg cells *in vivo* [352], and thus we can use Foxp3 induction as an indicator of the “pre-commitment” stage of Th1 cells. We discovered that Th1 cells after 48 hours of differentiation were plastic and allowed Foxp3 induction while Th1 cells

after 72 hours of differentiation lost the plasticity and became more committed to their Th1 helper cell fate. To further uncover the molecular mechanism of time-dependent plasticity, we performed different genome-wide measurements including chromatin accessibility and gene expression. We also utilized the new technique TrAC-looping, developed in our lab, on naive CD4⁺ T cells, 48-hour Th1 cells, and 72-hour Th1 cells with the purpose to link chromatin interaction alterations with the different cellular plasticity observed between 48 hours and 72 hours differentiation time. By integrating the information of chromatin accessibility and gene expression for differentiated Th1 cells, we predicted several transcription factors, including IRF4, which might be critical for governing the different status between 48hr Th1 like cells and 72hr Th1 like cells. Further conditional deletion of IRF4 in CD4⁺ T cells experiments confirmed our predictions and provided more insights into cellular plasticity.

Materials and Methods

Mice

The C57BL/6 T-bet-ZsGreen reporter mice were generously provided by Jinfang Zhu's lab and had been previously described [353]. IRF4^{fl/fl} CD4-Cre mice on the C57BL/6 background were provided by Stefan Muljo's lab. In detail, IRF4^{fl/fl} mice on mixed C57BL/6 and 129 backgrounds and CD4-Cre⁺ mice on the C57BL/6 background were purchased from Jackson Laboratory. IRF4^{fl/fl} CD4-Cre mice were generated by crossing IRF4^{fl/fl} mice with CD4-Cre⁺ mice and subsequently backcrossed to C57BL/6 background mice for at least ten generations. C57BL/6

background mice were purchased from Taconic. Male and female mice were bred and maintained in an NHLBI specific-pathogen-free animal facility. All experiments were performed on 6- to 10-week-old mice in accordance with a protocol approved by the NHLBI Animal Care and Use Committee. Mice were sex-matched and age-matched.

Cell isolation

Mouse naive CD4⁺ cells were isolated from the lymph nodes and spleens of C57BL/6, IRF4^{fl/fl}CD4-Cre⁺, and IRF4^{fl/fl}CD4-Cre⁺ mice using magnetic selection following the protocol provided by the manufacturer (Stem Cell, EasySep Mouse CD4⁺CD62L⁺ T cell Isolation Kit) or were sorted by flow cytometry on a FACS Aria II cell analyzer (BD Biosciences) for CD4⁺CD8⁻CD62L⁺CD44^{low}. The purity of naive CD4⁺ T cells was examined by flow cytometry for CD4⁺, CD8⁻ and CD62L⁺ on a FACSCanto II (BD Biosciences) or a BD LSR II (BD Biosciences) and was analyzed using FlowJo. Over 98% purity of CD4⁺CD8⁻CD62L⁺ cells was considered for further experiments. All antibodies used for sorting and flow-cytometry analyses were purchased from eBiosciences: anti-mouse CD4 (RM4-5), anti-mouse CD8 (53-6.7), and anti-mouse CD62L (MEL-14).

In vitro cell differentiation and FACS sorting

Naive CD4⁺ T cells were incubated in Th1 cell-differentiation medium in the presence of the following: plate-bound anti-CD3 (Clone: 145-2C11, 1 μ g/ml, eBioscience), anti-CD28 (Clone: 37.51, 2 μ g/ml, eBioscience), anti-IL-4 (10 μ g/ml, BioXCell), IL-12 (Cat# 210-12, 10 ng/ml, Peprotech), and IL-2 (Cat# 212-12, 50 U/ml, Peprotech). After 1-3 days in Th1 culture conditions, cells were washed by PBS twice and incubated in the Treg differentiation condition for additional 24 hours

in the presence of the following: anti-IL-4 (10 μ g/ml, BioXCell), anti-IL-12 (Clone: R1-5D9, 10 μ g/ml, BioXCell), anti-IFN- γ (Clone: R4-6A2, 10 μ g/ml, BioXCell), IL-2 (Cat# 212-12, 100 U/ml, Peprotech) and TGF- β (Cat# 240-B, 5 ng/ml, R&D Systems). For T-bet-ZsGreen reporter mice, cells were sorted by flow cytometry on a FACS Aria II cell analyzer (BD Biosciences) for ZsGreen⁺ and ZsGreen⁻. DAPI staining was used to remove dead cells during FACS sorting.

Flow cytometric analysis and intracellular staining

Cells were stained for surface markers with anti-mouse CD4 (RM4-5, eBioscience) in PBS with 2% FBS. Cells were washed then fixed and permeabilized with Foxp3 Staining Buffer Set (eBioscience) in accordance with the manufacturer's instructions. Intracellular staining was performed by washing cells and adding the following antibodies: anti-human/mouse T-bet (eBio4B10, eBiosciences), anti-mouse IFN γ (XMG1.2, eBiosciences), and anti-mouse Foxp3 (FJK-16s, eBioscience). Dead cells were removed from analysis using a LIVE/DEAD Fixable Aqua Dead Cell Stain Kit (Life Technologies) or eBioscience Fixable Viability Dye eFluor 780 (eBiosciences). All the antibodies for staining cell surface markers or cytokines were purchased from either BD Biosciences or eBiosciences.

DNase-seq

DNase-seq assays of CD4⁺ T cells at different differentiation stages were performed as described previously [236, 241]. In brief, 50,000 live cells were sorted and collected by flow cytometry on a FACS Aria II cell analyzer (BD Biosciences) by using a LIVE/DEAD Fixable Aqua Dead Cell Stain Kit (Life Technologies) or eBioscience Fixable Viability Dye eFluor 780 (eBiosciences) to remove dead cells.

Live cells were fixed for 10 min in 1% formaldehyde and lysed in 30 μ l cell lysis buffer (10 mM Tris-HCl, pH 7.5, 10 mM NaCl, 3 mM MgCl₂, 0.1% Triton X-100). For DNase I digestion, 8/3 unit of DNase I (Cat# 04716728001, Roche) was added to the cells and incubated at 37°C for 5 min. 80 μ l of stop buffer (10 mM Tris-HCl, pH 7.5, 10 mM NaCl, 0.15% SDS, 10 mM EDTA) containing 2 μ l of 20 mg/ml proteinase K was added to quench the reaction. The mixture was incubated at 65°C overnight, and cellular DNA was extracted using phenol-chloroform followed by ethanol precipitation. The libraries were prepared using Illumina kits as previously described [354] and one-step PCR which preferentially amplifies the small DNA fragments derived from the DNA hypersensitive sites was used. The fragments between 160 and 260 bp were isolated on E-gel and sequenced on Illumina HiSeq 3000.

RNA-Seq

Approximately 50,000 CD4⁺ T cells were homogenized and stored in QIAzol Lysis Reagent (Cat# 79306, Qiagen) for following total RNA extraction. Total RNA from different treatment groups was extracted using miRNeasy Micro Kit (Cat# 217084, Qiagen) according to the manufacturer's instruction. Polyadenylated RNA was synthesized following the instructions in the Smart-seq2 protocol [346]. About 100ng of double-stranded cDNA was sonicated on Bioruptor (Diagenode), and the libraries were constructed using Illumina kits as previously described [354] and sequenced on Illumina HiSeq 3000.

ChIP-seq experiments

We used chromatin from approximately $1-2 \times 10^6$ cells for each ChIP-Seq library and

antibodies to IRF4 (sc-6059, Santa Cruz). ChIP-seq assays were performed as previously described [158, 355]. In brief, cells for ChIP-seq were fixed in 1% formaldehyde for 10 min and sonicated. The sonicated chromatin was incubated with 1-2 μ g of antibodies at 4 °C overnight, and enriched chromatin was washed sequentially twice with low-salt buffer, twice with high-salt buffer, twice with LiCl buffer, and twice with TE (10 mM Tris-HCl). Crosslinks were reversed by incubation overnight at 65 °C, and ChIP DNA was purified by phenol-chloroform extraction and ethanol precipitation. Libraries were prepared using Illumina kits as described [354] and sequenced on Illumina HiSeq 3000.

Hyperactive Tnp purification and transposome complex assembly

The preparation details were described somewhere else in a manuscript under review.

In brief, the hyperactive Tnp (pET15b-His6Tnp) construct was assembled by NEB Gibson builder (NEB cat# E5510S) into pET-15b (Novagen cat# 69661-3), and competent cells BL21 Gold (DE3) (Agilent cat# 230132) were transformed with pET15b-His6Tnp. Single colonies were cultured to induce Tnp expression, and Tnp was purified in the cell lysates. 20 μ M Bio67F and 20 μ M 67bpR were annealed by heating at 98°C for 5 min, and we named this bivalent ME linker as “67bpR-Bio67F”. Similarly, ME3ddCTPtop and 5NH2MEbottom were annealed to form NHddCTP. The transposome complex was formed in a mixture of glycerol, annealed NHddCTP, 67bpR-Bio67F, and HisTnp enzyme at room temperature for 20 min.

TrAC-looping (Transposition-mediated Analysis of Chromatin Looping)

Approximately 10 million FACS-sorted live naive CD4⁺ T cells and Th1 cells were cross-linked using 1% formaldehyde (Thermo Scientific Cat# 28908) for 5 minutes at

room temperature, followed by adding glycine to 0.125M and washing twice with PBS. The cell pellets were resuspended in 100 ml cold lyse buffer sitting on ice for 15 minutes. The cells were spun down with the gentle centrifuge and resuspended in 500 μ l cold lyse buffer containing 1x reaction buffer (50 μ M Tris-acetate pH 7.5, 0.15 M potassium acetate, 10 mM magnesium acetate, 4 mM spermidine) and 8 μ l transposon complex. The reaction mixture was split into 10 Eppendorf tubes to increase the transposition efficiency. The reaction was incubated at 50°C for 1.5 hours and stopped by adding 2.5 μ l 0.5M EDTA, 1.6 μ l 10% SDS and 3 μ l Protease K (Roche Cat# 03115828001). After the reverse crosslink was performed at 65°C overnight, the genomic DNA was extracted by phenol-chloroform and ethanol precipitation. The DNA was dissolved in elution buffer and repaired for the 9bp gap using T4 DNA polymerase (NEB cat# M0203L) in the presence of dNTP. The reaction was abolished by adding EDTA, and larger than 300 bp DNA fragments were purified using Ampure XP beads (Agencourt Cat# A63881) following the manufacturers' protocol. The target DNA was eluted with 1x cutsmart buffer (NEB) and equally split into two restriction enzyme cutting reactions: Nla III and MluCI, respectively, at 37°C for 3 hours. Streptavidin C1 beads (Invitrogen Cat# 65001) were used to enrich the biotinylated bivalent linker-containing DNA fragments following the manufactures' instructions. After the wash steps, the beads containing target DNA fragments were resuspended in elution buffer containing Proteinase K and incubated at 55°C for 4 hours. The DNA was extracted using phenol-chloroform and ethanol precipitation. The DNA was intramolecularly circularized in the presence of T7 DNA ligase (NEB) at room temperature overnight followed by phenol-chloroform

extraction. The DNA pellets were resuspended in sampling buffer from Templify100 Kit (GE Healthcare Cat# 25-6400-10) and denatured. Rolling cycle amplification was performed by adding reactions buffer plus enzyme mix to incubate at 30°C overnight according to the Templify100 Kit protocol. 1 µl of RCA product was used as the template for PCR amplification with Illumina adaptor primers N501 and N701. 250-650 bp DNA fragments were isolated from the PCR products using E-gel and paired-end sequenced on Illumina HiSeq 3000.

Data analysis

Sequencing Library Mapping. Raw sequence reads from the DNase-seq and ChIP-seq libraries were aligned to the reference mouse genome (mm9) using Bowtie2 [319]. Only one read per position matching was allowed, and redundant reads were removed for each aligned library. Raw sequence reads from the RNA-seq libraries were aligned to the reference mouse genome mm9 using Tophat (--max-multihits = 15) [347].

Cluster Analysis. DNA hypersensitive sites (DHS) were identified in DNase-seq data using the software SICER [356]. We generated a combined set of DHSs among all DNase-seq libraries by combining any two DHSs if they overlapped. Finally, we obtained a total of ~53000 combined DHSs. The read density was calculated for each library at the combined DHS regions, where the reads density at a DHS region is the number of reads in the DHS region divided by the total number of reads of the library and the size of the DHS region. Finally, we obtained the read density at the combined DHS regions for 4 time points (naive, 24, 48 and 72 hours). The K-medoids

clustering method ($k=5$) was applied to the read density data, which in turn generated five different clusters.

Gene Ontology Analysis. The DHS regions in the five clusters were annotated with genes using the GREAT tool [357]. Gene ontology analysis was then performed for the associated genes in each cluster using the Gene Ontology Consortium [350].

Footprint Analysis. We applied the method called Protein Interaction Quantification (PIQ) [358] to estimate the TF-gene binding profiles. First, we downloaded the set of motifs in mouse (HOCOMOCOv10_MOUSE) [359]. The predicted motif binding sites were generated by the script `pwmmatch.exact.r`. The footprints were detected with the script `perftf.r`. To create the TF-gene binding profiles between 426 TFs and ~23000 genes, we annotated the binding sites by genes using HOMER (command: `annotatePeaks.pl`) [360]. A cutoff (cutoff = 700) was set for the binding score obtained from PIQ in order to determine whether a TF binds to the binding sites/genes.

Footprint Analysis with Modification. We also proposed a method to reconstruct the functional TF-gene binding profiles. In this study, we selected the potential regulators for the difference between 48hr Th1 cells and 72hr Th1 cells based on both networks generated by the footprint analysis and the following proposed method.

Consider a TF-gene network that has N TFs and M genes. From the footprint analysis, a $N \times M$ matrix P is obtained, where $P_{ij} = 1$ if the TF i binds to regions associated with gene j , otherwise $P_{ij} = 0$.

The expression level of a gene is regulated by different factors and is also associated with its mRNA degradation rate. Given the mRNA expression of gene i at time t to be $X_i(t)$, then $X_i(t)$ generally follows the following differential equation:

$$\frac{dX_i(t)}{dt} = b(t) - a(t)X_i(t), \quad (1)\text{Error!}$$

Bookmark not defined.

where $b(t)$ and $a(t)$ refer to the production and degradation rate of $X_i(t)$. We consider that the production rate of $X_i(t)$ mainly depends on the regulation by transcription factors, which in turn depends on the protein level of TFs and the efficiency of the TF-gene regulations. For simplicity, we assume that the protein levels of TFs are equal to their mRNA expression levels. Therefore, $b(t) = K_{ji}(t)P_{ji}(t)X_j(t)$, where gene j can produce the transcription factor (TF j) protein that regulates gene i , and $K_{ji}(t)$ is the efficiency of the regulation between TF j and gene i . $K_{ji}(t)$ can either have positive and negative values which refer to the activation or repression processes, respectively.

First, we consider the simplest case where only one TF (TF j) regulates the mRNA expression of gene i . By applying the trapezoidal rule and considering that $t_1 - t_0 = 1$, then Eq. (1) becomes:

$$\frac{dX_i(t)}{dt} = K_{ji}(t)P_{ji}(t)X_j(t) - aX_i(t),$$

$$\int_{t_0}^{t_1} \frac{dX_i(t)}{dt} dt = \int_{t_0}^{t_1} (K_{ji}(t)P_{ji}(t)X_j(t) - aX_i(t)) dt,$$

$$X_i(t_1) - X_i(t_0) \approx \frac{[K_{ji}(t_1)P_{ji}(t_1)X_j(t_1) + K_{ji}(t_0)P_{ji}(t_0)X_j(t_0)] - a[X_i(t_1) + X_i(t_0)]}{2},$$

$$\begin{aligned} & \left(1 + \frac{a}{2}\right)X_i(t_1) - \left(1 - \frac{a}{2}\right)X_i(t_0) \\ &= \frac{[K_{ji}(t_1)P_{ji}(t_1)X_j(t_1) + K_{ji}(t_0)P_{ji}(t_0)X_j(t_0)]}{2}. \end{aligned} \quad (2)$$

Eq. (2) is a special case referring that only one TF regulates a gene. General, if many TFs regulate gene i , then Eq. 2 becomes

$$dX'_i = \sum_{\forall j \in TF} \frac{[K_{ji}(t_1)P_{ji}(t_1)X_j(t_1) + K_{ji}(t_0)P_{ji}(t_0)X_j(t_0)]}{2}, \quad (3)$$

where $dX'_i = \left(1 + \frac{a}{2}\right)X_i(t_1) - \left(1 - \frac{a}{2}\right)X_i(t_0)$. Eq (3) can be rewritten in a matrix form

$$dX'_i = \mathbf{Q}\mathbf{Y},$$

where $\mathbf{Q}$ is a $1 \times w$ vector and $\mathbf{Y}$ is a $w \times 1$ vector. w is equal to the total number of non-zero $P_{ji}(t_1)$ and $P_{ji}(t_0)$ for $j = 1 \dots, N$. We can solve $K_{ji}(t)$ by solving $\mathbf{Q}$. As an example for showing how to solve $K_{ji}(t)$, we consider that there is a set of TFs denoted as S , in which contains q elements denoted as $S_j, j = 1, \dots, q$. All TFs in S bind to gene i at both time points. Therefore, $P_{S_j i}(t_1)$ and $P_{S_j i}(t_0)$ equal to 1 for $j = 1, \dots, q$ and $w = 2q$. Then $\mathbf{Q}$ and $\mathbf{Y}$ can be expressed as

$$\mathbf{Q} = \left[K_{S_1 i}(t_0), K_{S_1 i}(t_1), \dots, K_{S_j i}(t_0), K_{S_j i}(t_1), \dots, K_{S_q i}(t_0), K_{S_q i}(t_1) \right],$$

and

$$\mathbf{Y} = \begin{bmatrix} X_{S_1}(t_0) \\ X_{S_1}(t_1) \\ \vdots \\ X_{S_j}(t_0) \\ X_{S_j}(t_1) \\ \vdots \\ X_{S_q}(t_0) \\ X_{S_q}(t_1) \end{bmatrix}.$$

Finally, $\mathbf{Q}$ can be solved by the following equation

$$\mathbf{Q} = d\mathbf{X}'_i \mathbf{Y}^T (\mathbf{Y} \mathbf{Y}^T)^{-1}.$$

As a result, the weighted TF-gene network can be obtained by computing K_{ji} for all gene i and TF j . Note that $K_{ji} = 0$ if $P_{ji} = 0$.

Identification of potential regulators. To identify the potential regulators, we computed the TF-gene networks, P and K , using the footprint analysis and modified footprint analysis described above, respectively. We defined a quantity $D_j(t)$ indicating the total binding of TF j at time t . To distinguish the quantity for the two networks, a superscript was added to $D_j(t)$ such that $D_j^P(t) = \sum_{i=1}^M P_{ji}(t)$ and $D_j^K(t) = \sum_{i=1}^M K_{ji}(t)$. Then we computed a ranking denoted as R for the TFs based on the difference of $D_j(t_1)$ versus $D_j(t_0)$, where t_1 and t_0 refer to 72 hours and 48 hours,

respectively. Then a TF j was considered to be a potential regulator if it satisfied the following three criteria:

1. The TF j has a top ranking, which means R_j is smaller than a cutoff.
2. The mRNA expression for the TF gene j is higher than a cutoff (RPKM > 10).
3. The folder change of mRNA expression for the TF gene j is either larger or smaller than 1 and -1, respectively.

TrAC-looping PETs mapping and significant interaction identification. The paired-end tags (PETs) from TrAC-looping libraries were mapped to the mouse reference genome (mm9) using Bowtie2 [319], and the PETs with low mapping quality were filtered out, and only one uniquely mapped PET was maintained at each coordinate. The identified PETs were separated into three categories depending on the PETs' end locations. Self-PETs, both ends located in the same hypersensitive site, were produced by two insertions of the tetramer-Tn5 complex into the same hypersensitive site and did not involve the interaction among regulatory regions. Inter-PETs, two ends located in two different hypersensitive sites, potentially indicated the existence of the interaction between two regions and provided information about chromatin organization and gene regulatory networks. Background-PETs, at least one end did not locate in hypersensitive sites, gave information about three-dimensional chromatin organization. We mostly focused on the intra-PETs since they helped reveal the possible enhancer-promoter interactions. The number of PETs linked two hypersensitive sites should follow a hyper-geometric distribution, similar to the ChIA-PET data [361, 362]. We assumed the inter-PETs number in a TrAC-looping

library is N , and two hypersensitive sites are denoted as p and q . The numbers of inter-PETs with one end located in hypersensitive site p and hypersensitive site q are n_p and n_q , respectively. The following formula was obtained:

$$P(I_{pq}) = \frac{\binom{n_p}{I_{pq}} \binom{2N-n_p}{n_q-I_{pq}}}{\binom{2N}{n_q}}$$

I_{pq} is the observed number of PETs linked hypersensitive site p and q . The p-value for the hyper-geometric test could be calculated to decide if I_{pq} is over-represented. Multiple testing corrections were performed, and only significant interactions were maintained for further analysis (FDR<0.05 and strength ≥ 2 PET). If we observed 3 or more PETs in one interaction, we believed this was a strong and confident interaction.

Compartment Analysis. Each mouse chromosome was divided into 40Kb bins, and the interaction matrix of each chromosome was generated for TrAC-looping data from mouse naive CD4⁺ T cell, 48-hour Th1 cells, and 72-hour Th1 cells. The interaction matrix was normalized by filtering out the distance factor using the KR-normalization. The correlation matrix was obtained by computing the Pearson correlation between any two interactions. The compartment score was the first principal component obtained by the PCA analysis on the correlation matrix.

Statistics

We applied a two-sided Mann-Whitney test (two-tailed throughout the manuscript) for any two groups of interests. The Mann-Whitney test is a nonparametric method that tests whether two probability distributions differ without the requirement of prior knowledge about the distributions [363]. All the statistical examinations were

performed using R ([http:// www.r-project.org/](http://www.r-project.org/)) with proper parameter adjustments.

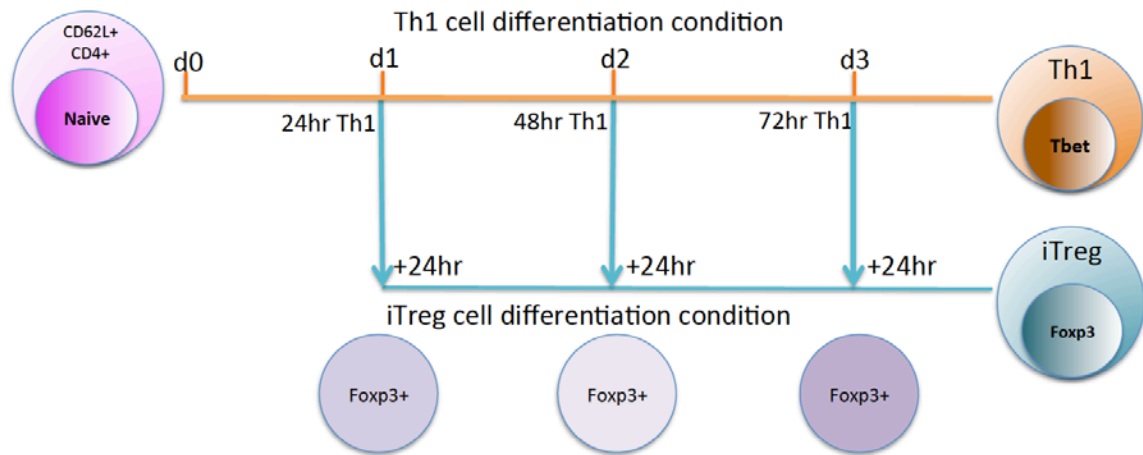
Results

Th1 like cells can be reprogrammed to express Foxp3 in a time-dependent manner

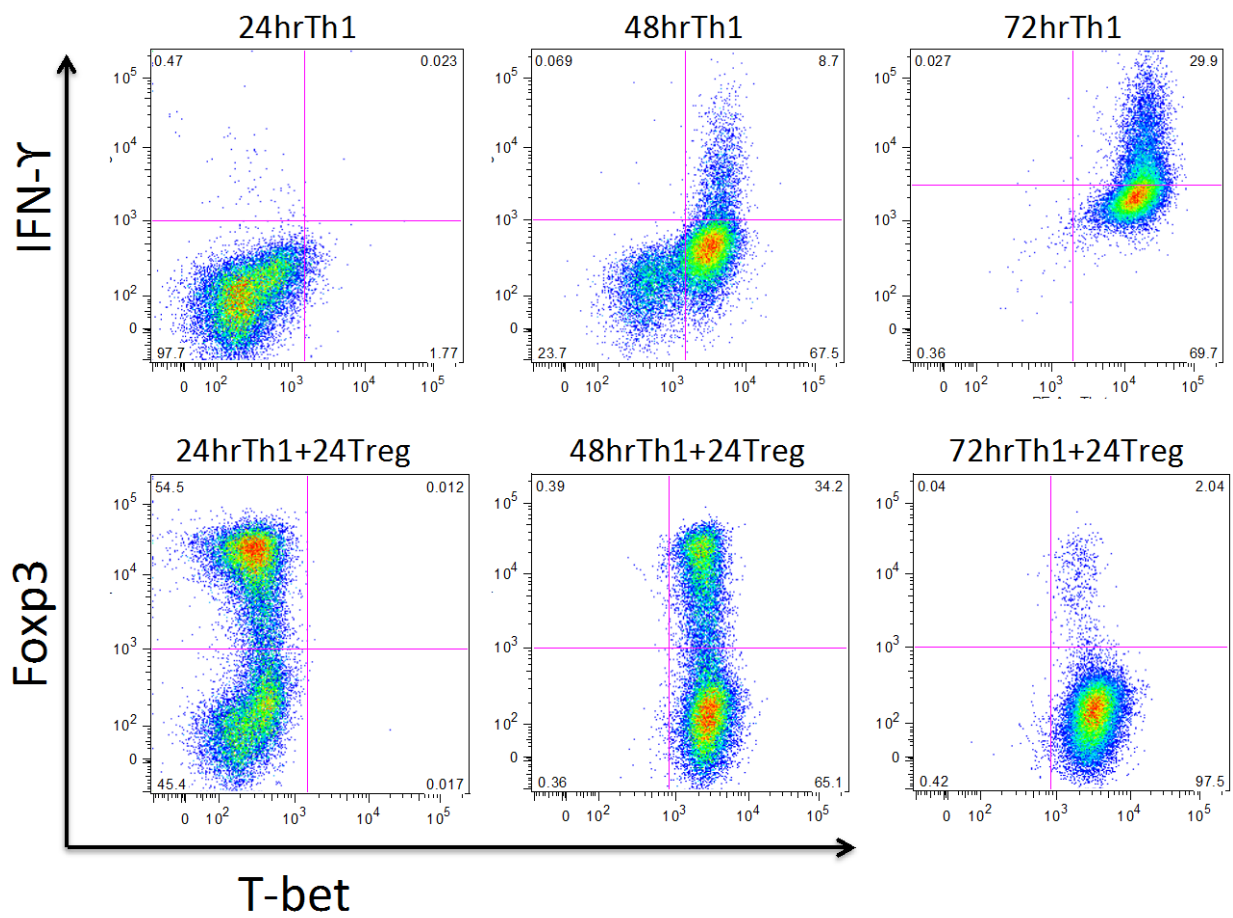
In order to understand the plasticity and commitment of the Th1 cell lineage, naive CD4⁺ T cells, defined as CD62L⁺ and CD44 low, were differentiated into Th1 cells and challenged under iTreg cell-polarizing conditions *in vitro* at different time points (Figure 4.1(A)). After 1 day of Th1 differentiation, although more than 95% cells remained negative for T-bet and IFN- γ , approximately 50% cells induced Foxp3 expression under Treg-inducing conditions for additional 24 hours. After 2 days of Th1 differentiation, the majority of cells became T-bet positive with the up-regulation of IFN- γ . In response to Treg-inducing stimuli for 24 hours, approximately 35% of cells induced Foxp3 expression while almost all cells remained positive for T-bet. However, after 3 days of differentiation, cells that were positive for T-bet were resistant to reprogramming towards Foxp3 expressing Treg-like cells. The secondary stimulation in the presence of TGF β resulted in only minor reduction in the amount of T-bet protein, and Foxp3 expression remained very low when compared with the amount observed in 1 day and 2 day differentiated Th1 cells (Figure 4.1(B)). At the same time, TGF β induced strong iTreg cell differentiation and Foxp3 expression in naive CD4⁺ T cells (Appendix V). Such phenomenon was observed in several independent experiments (Figure 4.1(C)). Thus, the plasticity of Th1 cells was highly

time-dependent, and the Th1 cell commitment stage was established after 3 days of differentiation.

A



B



C

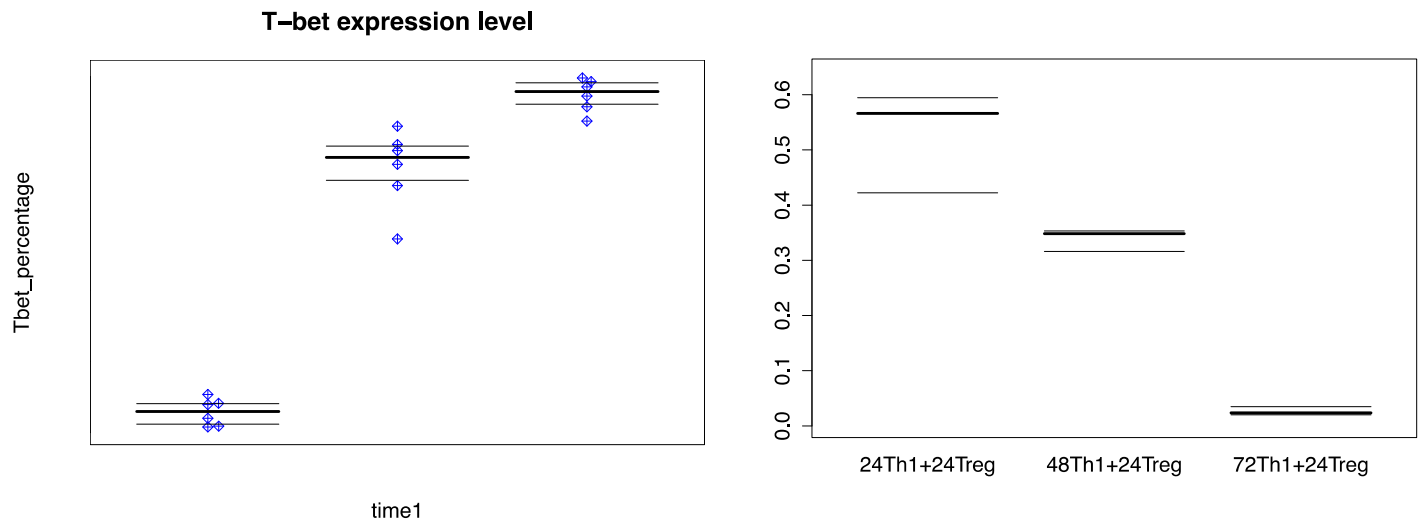


Figure 4.1 Th1 Cell Reprogramming

(A) Naive CD4⁺ T cell populations were polarized in Th1 culture conditions for 24 hours, 48 hours, and 72 hours. At each stage, the cells were collected and then polarized in iTreg culture conditions for additional 24 hours to observe their Foxp3 expression. (B) At the indicated time points (24 hours, 48 hours, 72 hours, and additional 24 hours), T-bet and Foxp3 protein expression were determined in Th1 like cells by FACS. Percentage of protein and cytokine production is indicated. Data are representative of at least two independent experiments. (C) Six independent experiments indicate T-bet and Foxp3 protein expression in Th1 like cells at the indicated time points as in (B).

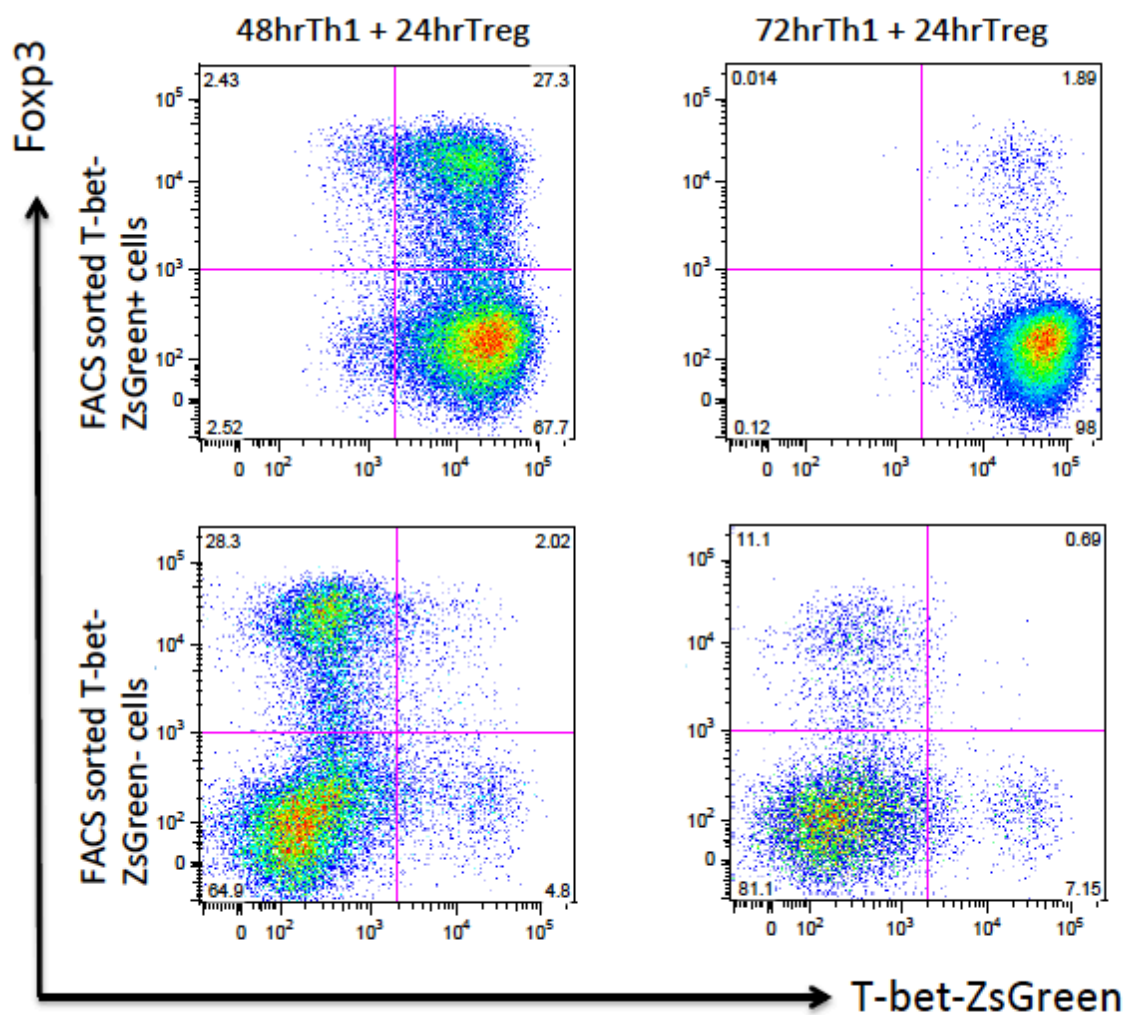
T-bet is not the direct determinant of Th1 cell lineage commitment and plasticity

To understand if T-bet, the master transcription factor of Th1 cells, directly determines how Th1 cells respond to a secondary stimulation, we used T-bet-ZsGreen reporter mice [353] to separate CD4⁺T-bet⁺ and CD4⁺T-bet⁻ cells under Th1 cell-polarizing conditions for 2 days and 3 days, and then challenged under iTreg cell-skewing conditions for additional 1 day *in vitro*. After the secondary stimulation in iTreg culture conditions, the protein expression of Foxp3 in CD4⁺T-bet⁺ cells did not change substantially compared to 2-day differentiated Th1 cells as indicated in Figure

4.1(B) (Figure 4.2(A)). Although the percentages of CD4⁺T-bet⁻ cells under Th1 cell-polarizing conditions for 2 days and 3 days were relatively low compared to CD4⁺T-bet⁺ cells, CD4⁺T-bet⁻ cells could still produce Foxp3 at a similar scale compared to CD4⁺T-bet⁺ cells (Figure 4.2(B)). In order to obtain enough CD4⁺T-bet⁻ cells for FACS analysis as well as understand the dynamic of Th1 cell differentiation programs, 43 hours and 67 hours under Th1-cell polarizing conditions were also included. Notably, the expression of Foxp3 in CD4⁺T-bet⁺ cells after the 24 hours secondary stimulation showed a negative correlation with their culture time under Th1-polarizing conditions. And such substantial changes related to Th1 cell plasticity would mainly occur between 2 days and 3 days of differentiation. The ratio of Foxp3-expressing cells for CD4⁺T-bet⁻ cells after the secondary stimulation fluctuated, which might be caused by incomplete T cell activation or technical variations. We noticed that Foxp3 induction for 2-day CD4⁺T-bet⁺ and CD4⁺T-bet⁻ cells was similar, while 3-day CD4⁺T-bet⁺ and CD4⁺T-bet⁻ cells showed a slight difference in Foxp3 expression after the secondary stimulation (Figure 4.2(B)). Nevertheless, CD4⁺T-bet⁺ cells contributed to the majority or almost all of the cell populations on 2 days and 3 days of differentiation (data not shown). And this might explain why the ratio of Foxp3 expression in CD4⁺T-bet⁺ cells under additional iTreg-polarizing conditions was almost the same compared to that in the unsorted population as indicated in Figure 4.1. And virtually the majority of the CD4⁺T-bet⁺ cells remained positive for T-bet at later time points (Figure 4.2(A)). Thus, it is fair to infer that T-bet, the master regulator of Th1 cells, might not be the direct determinant of Th1-like

cells' behavior toward environmental stimuli alterations. The key factors controlling such cell plasticity and commitment pattern deserve a more in-depth investigation.

A



B

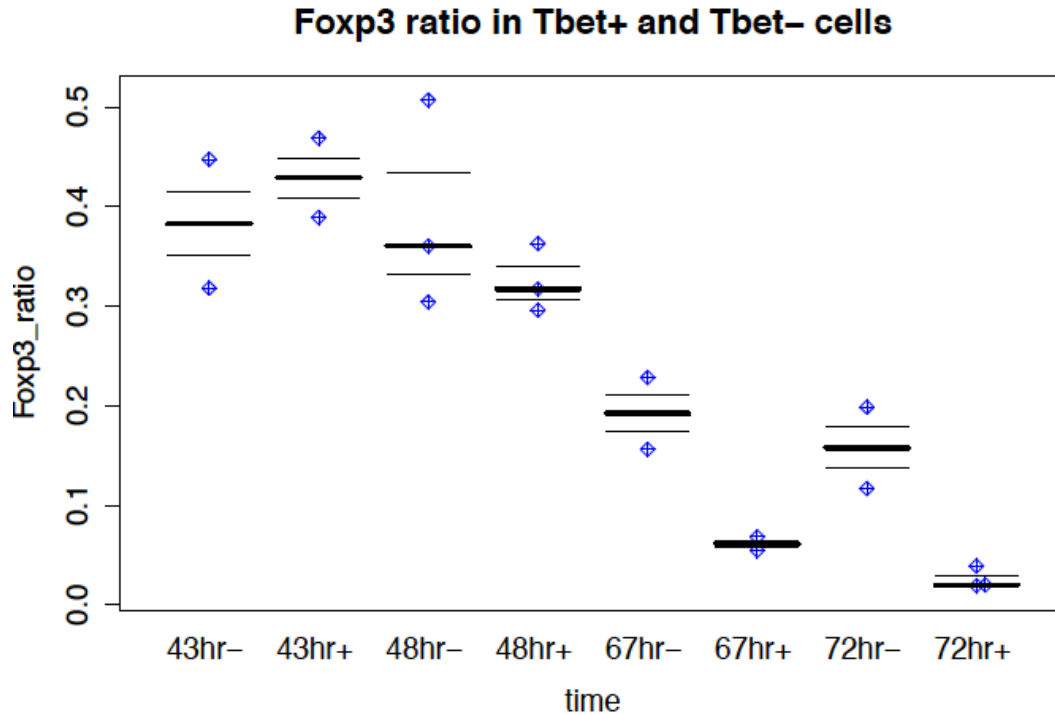


Figure 4.2 T-bet Function in Th1 Cell Fate

T-bet alone is not the only factor determining Th1 cell lineage commitment and plasticity. Naive CD4⁺ T cells isolated from T-bet-ZsGreen reporter mice were differentiated to Th1 cells for 2 days and 3 days before separated into T-bet-ZsGreen⁺ and T-bet-ZsGreen⁻ populations by cell sorting and then challenged under iTreg cell-polarizing conditions *in vitro* culture for additional 24 hours.

(A) Sorted T-bet-ZsGreen⁺ cells and T-bet-ZsGreen⁻ cells from T-bet-ZsGreen reporter mice under Th1 cell-polarizing conditions for 48 hours and 72 hours, respectively, were separately cultured in iTreg cell-polarizing conditions for additional 24 hours. Flow cytometric analysis was used to assess the expression of Foxp3 and T-bet-ZsGreen. 48hrTh1 and 72hrTh1 indicate the culture time under Th1 cell-polarizing conditions, and 24hrTreg refers to the additional 24 hours under iTreg cell-polarizing conditions. Numbers show the percentage of cells in each quadrant or gate. Data are representative of at least two independent experiments. (B) T-bet-ZsGreen⁺ and T-bet-ZsGreen⁻ cells were also separated under Th1 cell-polarizing conditions for 43 hours and 67 hours. The x axle labels (43hr+/-, 48hr+/-, 63hr+/-, and 72hr+/-) refer to the T-bet-ZsGreen⁺ or T-bet-ZsGreen⁻ cells cultured in Th1 cell-polarizing conditions for corresponding 43, 48, 63, and 72 hours, respectively, and then challenged in iTreg skewing conditions for additional 24 hours. Cells were collected for FACS analysis on Foxp3 expression in several independent experiments as indicated.

Dynamic chromatin states for the different stages of differentiated Th1 cells

We next characterized the chromatin states in naive CD4⁺ T cells, cultured Th1 cells over the course of differentiation, as well as the cells under additional iTreg-polarizing conditions. DNase-seq libraries were generated for the 8 experimental conditions (including 4 time points), in which there were two replicates for each library (Appendix VI and Appendix VII). In total, there were 54,000 DNase hypersensitive sites (DHSs) detected after combining the overlapped DHSs from all libraries. To characterize the chromatin state of Th1 differentiation, we applied the K-medoids clustering (k=5) method to the DNase-seq data on differentiated Th1 cells (0, 24, 48, 72 hours). The cluster results (Figure 4.3(A)) showed that the accessibility of the DHS regions decreased in two clusters (C_2 and C_4) and increased in the other three clusters (C_1 , C_3 , and C_5) during differentiation. Also, the transition from low accessible to high accessible regions occurred at different time points (DHSs in C_1 , C_3 , and C_5 became accessible at 72, 24, and 48 hours respectively during Th1 differentiation). Gene ontology analysis for the five clusters of DHS regions showed that they were associated with different biological processes (Figure 4.3(B)), which indicated that the Th1 differentiation process was dynamic, with a distinct emphasis on T cell activation or differentiation depending on time.

Identification of potential regulators governing Th1 lineage commitment

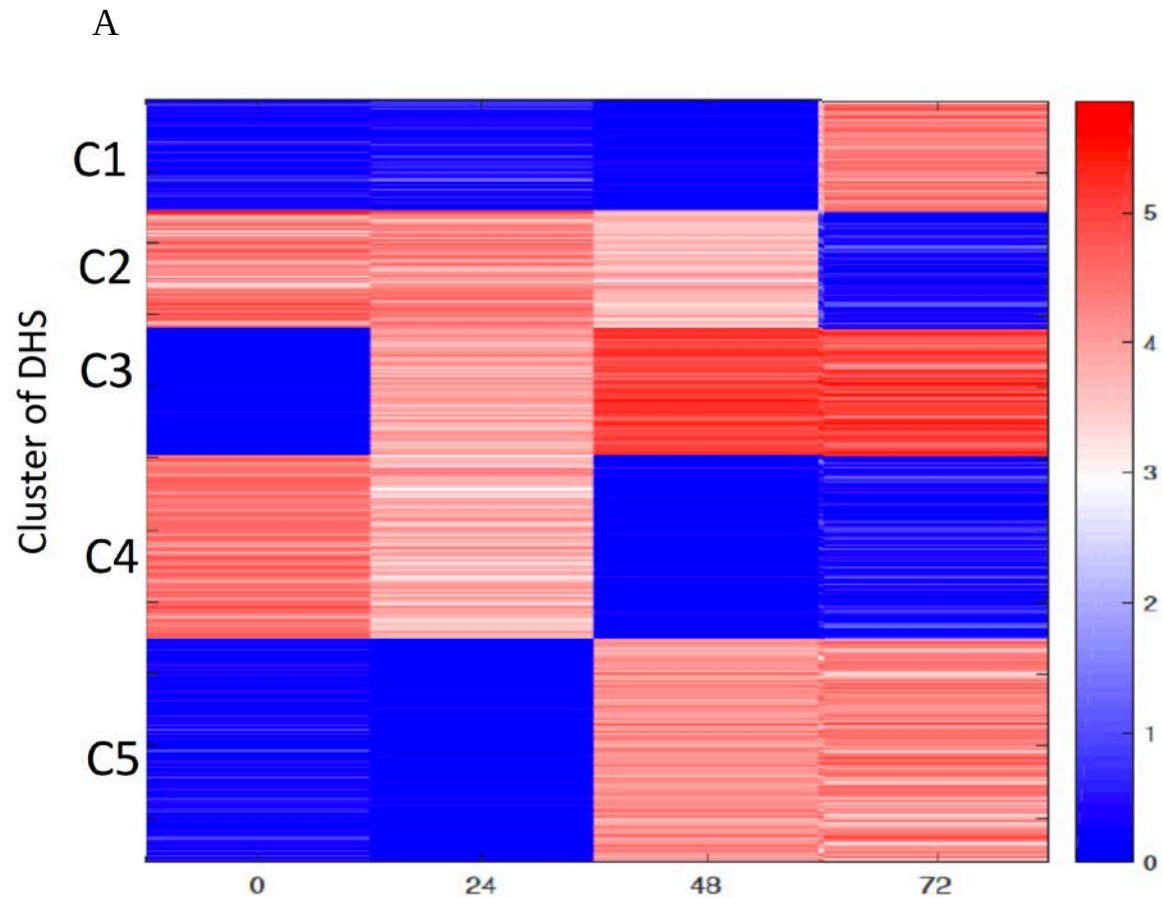
Based on both FACS staining and DNase-seq data, Th1 cells at different differentiation stages exhibited dynamic chromatin states, gene expression, and cellular function. More importantly, there existed a critical transition between 48-hour Th1 cells and 72-hour Th1 cells, where 72-hour Th1 cells were fully committed while

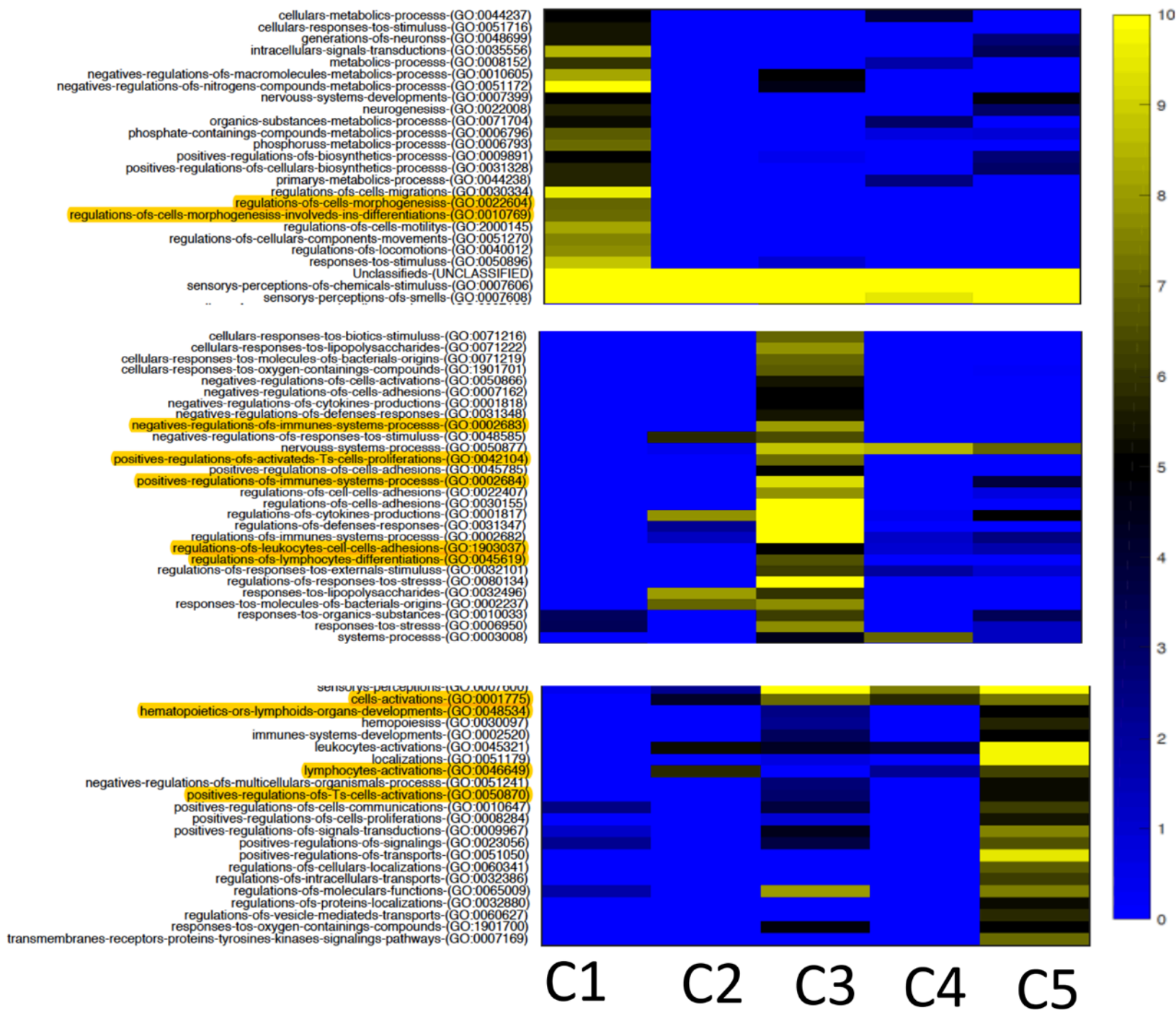
48-hour Th1 cells still exhibited cellular plasticity. Therefore, it would be interesting to elucidate the key regulators governing this critical transition of cell states.

We reconstructed the TF-gene binding networks for the two differentiated stages using DNase-seq data, such that key regulators could be predicted by examining differences between TF-gene networks [195]. Previous studies applied footprint (FP) analysis to DNase-seq or ATAC-seq data to reconstruct the TF-gene binding network, which in turn identified the related regulators [364]. However, the bindings between TFs and target genes predicted by FP analysis do not necessarily imply a functional relationship between the TF and its target genes. Therefore, we developed a modified footprint analysis method (named as MFP) to reconstruct the TF-gene binding networks by integrating the gene expression data with the footprint analysis.

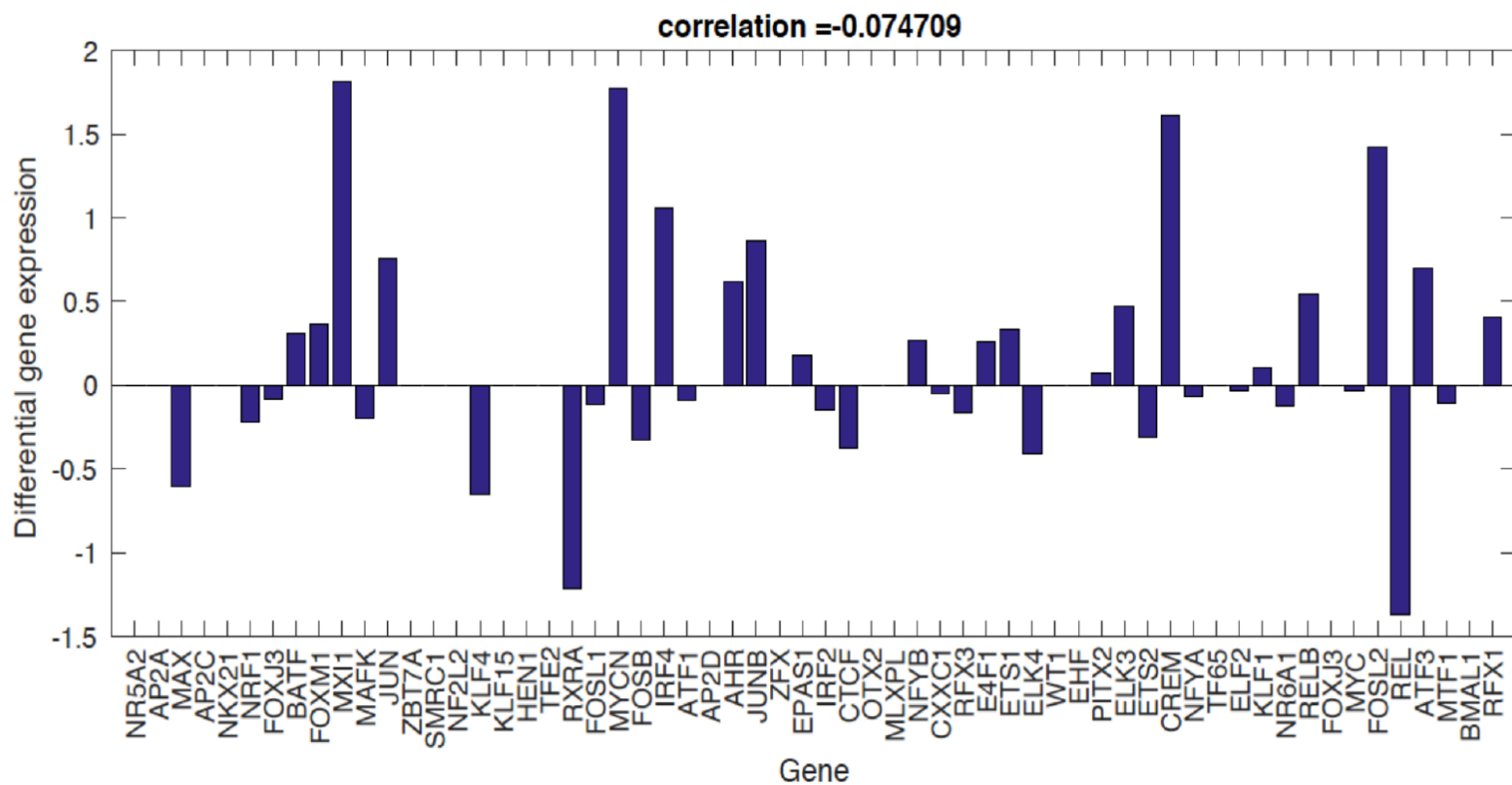
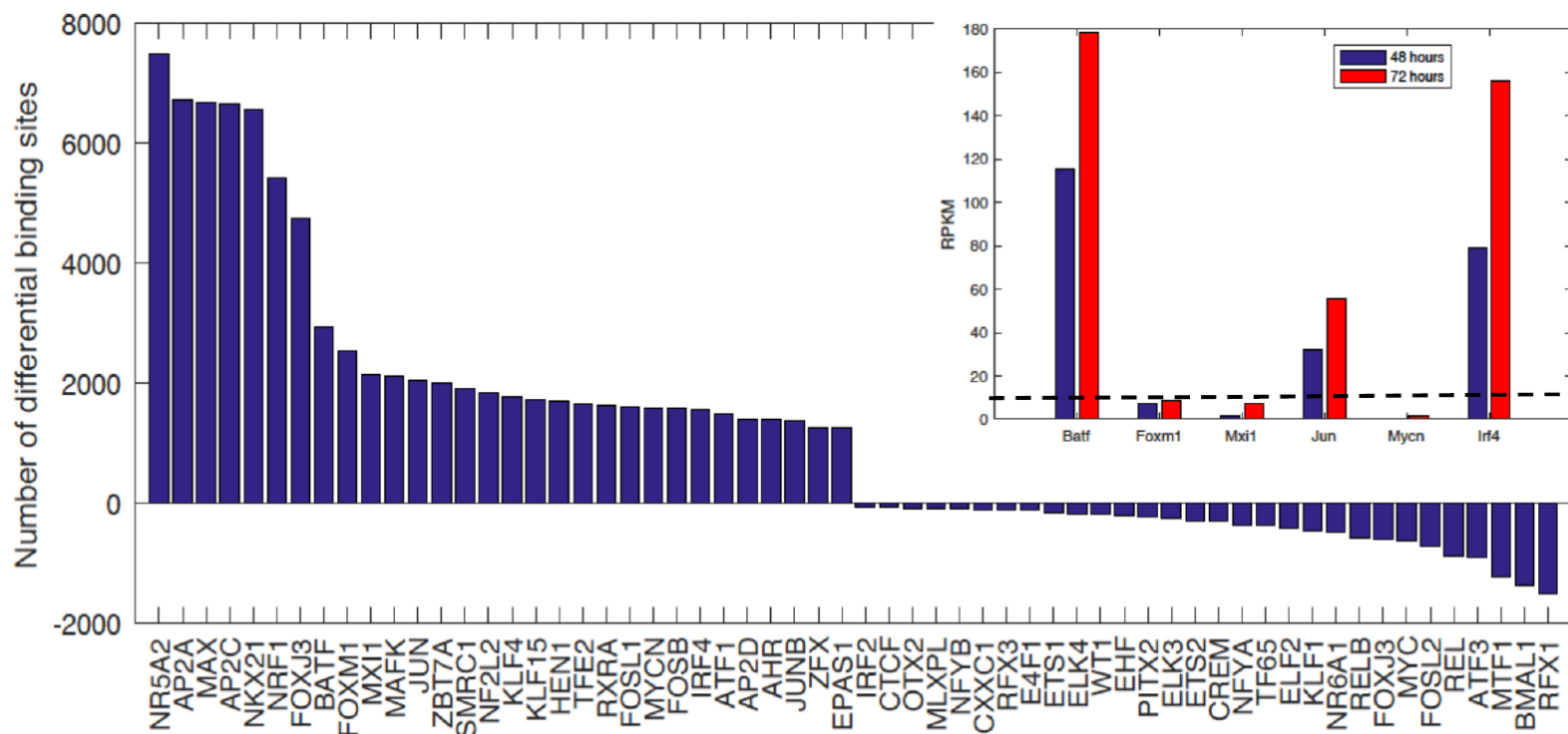
We studied the TF-gene networks obtained by both FP analysis using the PIQ method [358] and our MFP method in parallel. For the FP method, 619,029 and 726,695 interactions were detected between 426 TFs and 23,288 genes for the 48-hour Th1 and 72-hour Th1 cells. For the MFP method, 356,153 and 410,868 interactions were detected between 426 TFs and 23,288 genes for the 48-hour Th1 and 72-hour Th1 cells. Next, we predicted the potential regulators by using the transcription factor binding recovered by the FP and MFP analysis independently. We assigned a value to each TF at each time point, and the number referred to the total number of bindings of the TF. For each TF, we subtracted the value at 72 hours by that at 48 hours. The TFs were ranked in a descending order using the values after subtraction. The top-ranked TFs (top 30 in the FP and MFP analysis) were considered to be potential regulators if they satisfied two criteria: 1) the TF genes were expressed at both time points

(rpkm $\geq$ 10) and 2) their fold changes of gene expression at 72 hours versus 48 hours were greater than a cutoff (>2-fold). Interestingly, we found that IRF4 was the consistent potential regulator predicted using TF-genes' binding profiles from both FP (Figure 4.3(B)) and MFP analyses (Figure 4.3(C)).

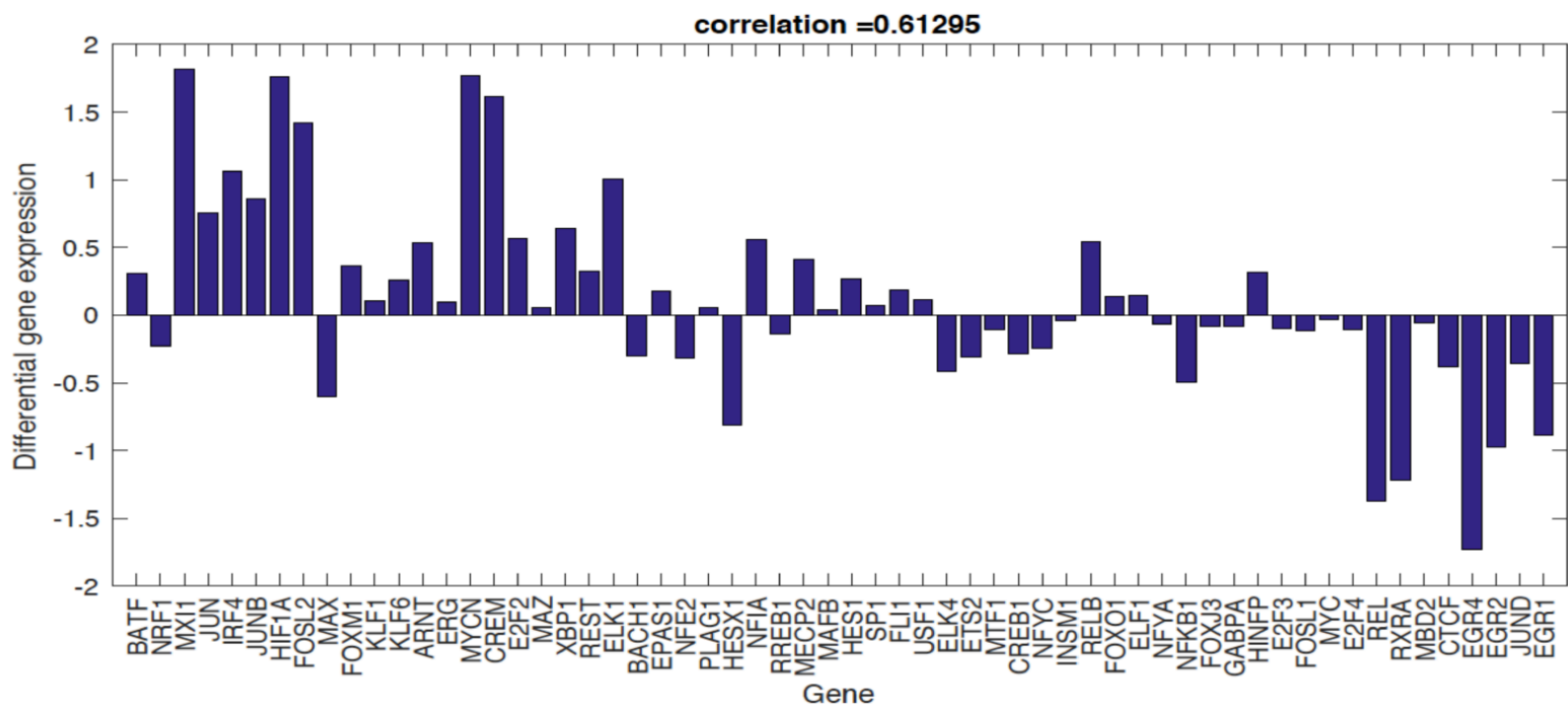
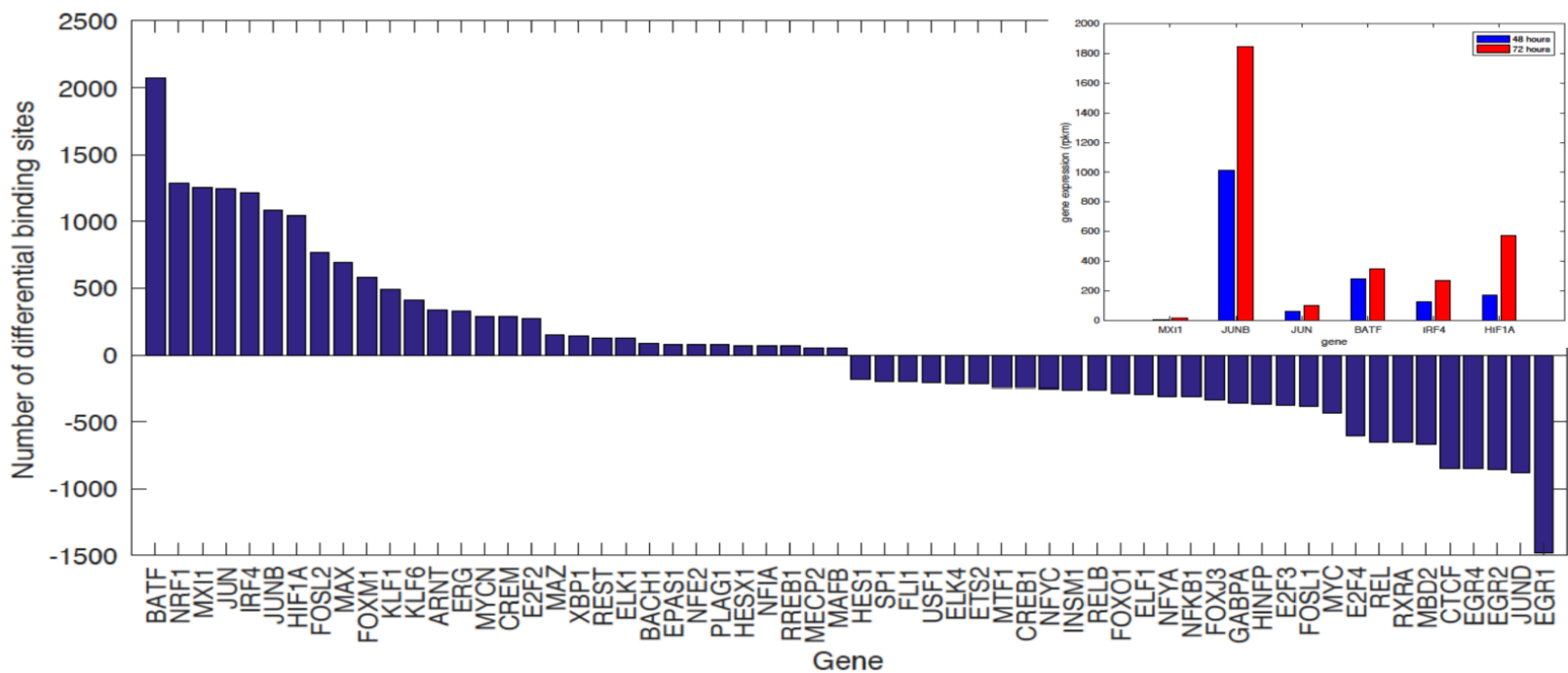




B



C



D

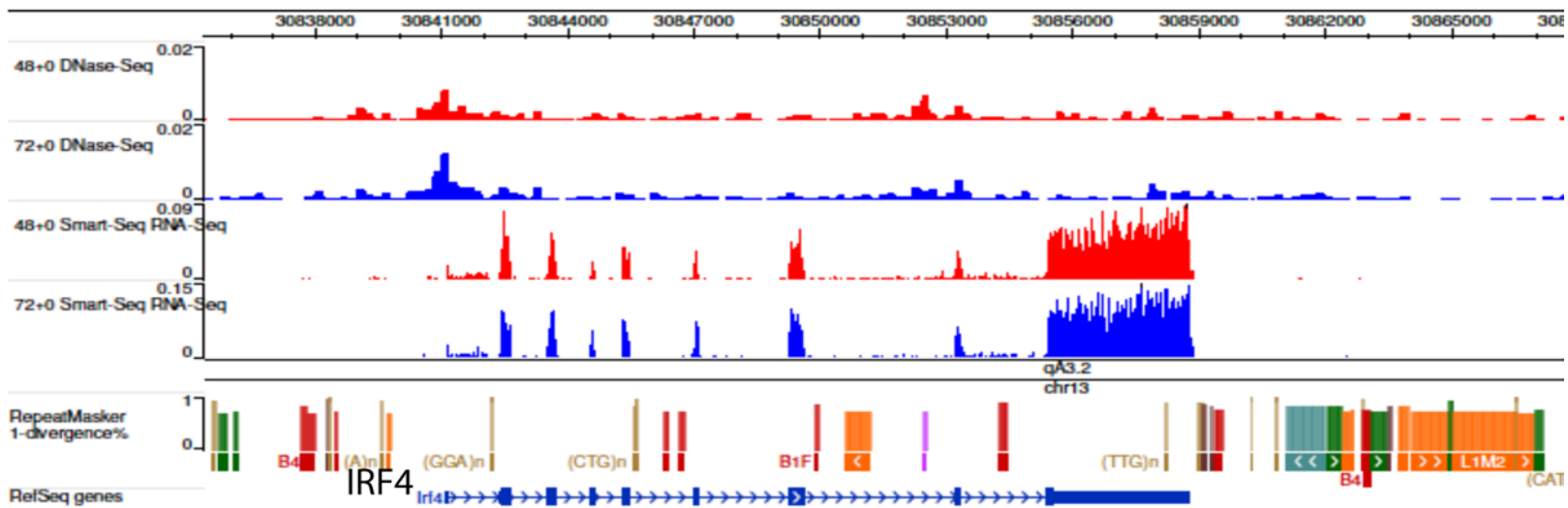


Figure 4.3 Analysis of the DNase-seq and RNA-seq Data

(In collaboration with Dr. Wai Lim Ku from Dr. Keji Zhao's lab in NHLBI, NIH)

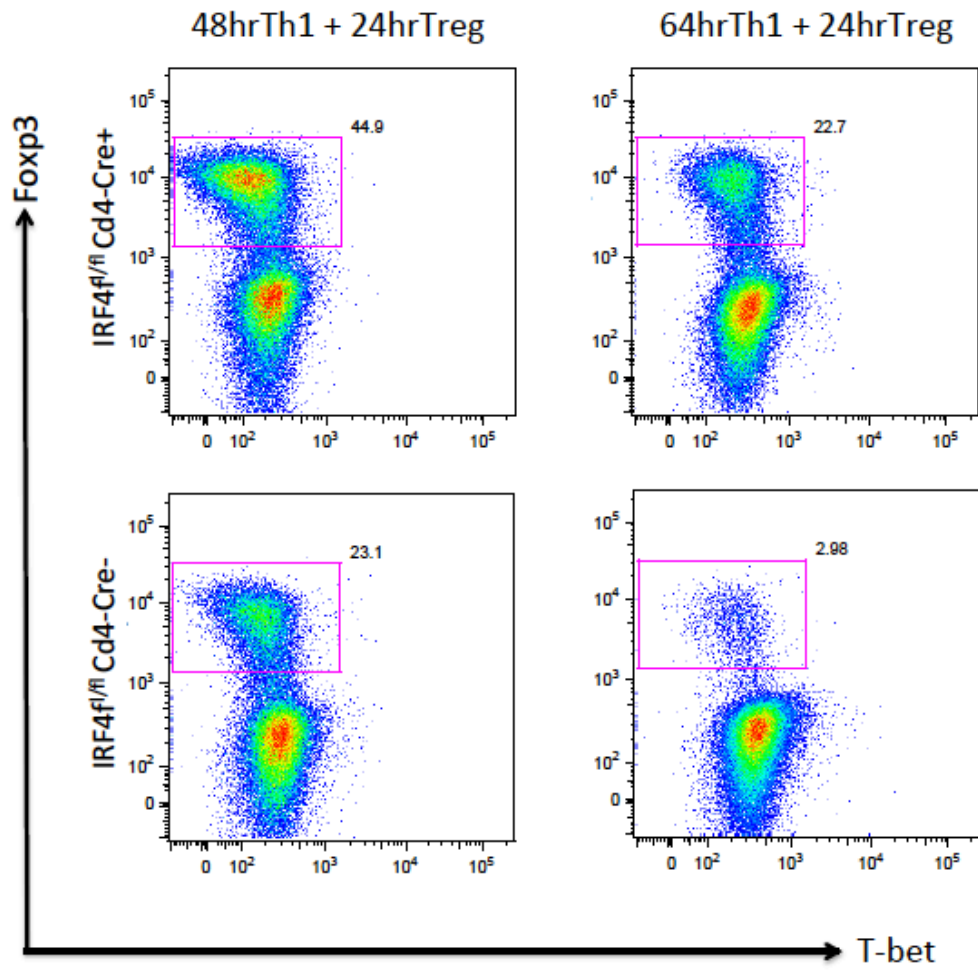
(A) **(The top panel)** Five clusters (c1, c2, c3, c4, and c5) of the DNA hypersensitive sites (DHSs) were detected at 0, 24, 48 and 72 hours after Th1 differentiation. The clusters of DHS were generated by the K-medoids clustering method. The color spectrum refers to the DNA accessibility where the more accessible DHS is in red, and the less accessible DHS is in blue. For the y-axis, one row refers to the average the DNA accessibility of 10 DHSs across the four time points. **(The bottom panel)** The DHSs in the five clusters were annotated with genes using the GREAT tools. Gene ontology analysis was then performed for the associated genes in each cluster. The heatmap in the bottom panel showed the significance of the selected gene ontology terms enriched in genes in each of the five clusters. Each row in the y-axis of the heatmap refers to one of the selected gene ontology terms while each column in the x-axis refers to the cluster (c1, c2, c3, c4, and c5) in the heatmap in the top panel. More yellow in the color spectrum refers to a stronger enrichment while more blue refers to a weaker enrichment. (B) **(The top panel)** The TF-binding profiles were recovered using the DNase footprint analysis. TFs were ranked based on the number of different binding sites between 72 hours and 48 hours (72 hours vs. 48 hours). A positive number refers to more binding sites in 72 hours compared to 48 hours. The numbers of different binding sites for the top and bottom 30 ranked TFs are plotted in the top panel. Six TFs (Batf, Foxm1, Mxi1, Jun, Mycn, and IRF4) were selected. A subplot in the top right corner shows their mRNA expression levels in RPKM (y-axis) at 48 hours (blue) and 72 hours (red). The dashed black horizontal line refers to the cutoff of RPKM values at 10. **(The bottom panel)** The bottom panel shows a bar plot of the fold change of gene expression levels between 72 and 48 hours (72 vs. 48 hours) for the top and bottom 30 ranked TFs. (C) **(The top panel)** The TF-binding profiles were recovered using the MFP method. Another six TFs (Mxi1, JunB, Jun,

Batf, IRF4, and Hif1a) were selected to make the subplot in the top right corner. (**The bottom panel**) The bottom panel shows a bar plot of the fold changes of gene expression levels between 72 and 48 hours (72 vs. 48 hours) for the top and bottom 30 ranked TFs. (D) A genome browser snapshot at the gene locus of IRF4 showing the DNase-seq and RNA-seq data at 48 and 72 hours. The panel order from top to bottom is DNase-seq data at 48 hours (red), DNase-seq data at 72 hours (blue), RNA-seq data at 48 hours (red), RNA-seq data at 72 hours (blue).

Deletion of IRF4 results in enhanced Th1 cell reprogramming capability

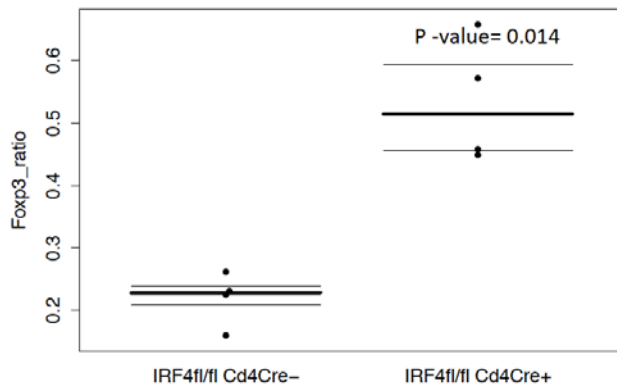
The interferon regulatory factor family TF IRF4 was among the highest-ranking candidates observed to regulate Th1 cell commitment (Figure 4.3(D)). To investigate its capacity in Th1 cell differentiation and reprogram, we used mice with conditional deficiency of IRF4 in CD4⁺ T cells (IRF4^{fl/fl} CD4-Cre). The Th1 cells of IRF4^{fl/fl}CD4-Cre- mice maintained a similar capacity compared to wild-type mice (as indicated in Figure 4.1) to express Foxp3 after the subsequent exposure to Treg-skewing conditions in the presence of TGF β for additional 24 hours (Figure 4.4(A)). However, conditionally IRF4-deleted Th1 like cells could be induced to produce more Foxp3, such that the frequency of Foxp3⁺ cells in IRF4^{-/-} Th1 cells exceeded the frequency in wild-type or IRF4^{fl/fl}CD4-Cre- Th1 cells significantly (P-value < 0.05) (Figure 4.4). This finding indicated a necessary role of IRF4 in restraining Th1 cells reprogramming ability toward Foxp3-expressing Treg like cells.

A



B

48hrTh1 + 24hrTreg cells Foxp3+ ratio



C

64hrTh1 + 24hrTreg cells Foxp3+ ratio

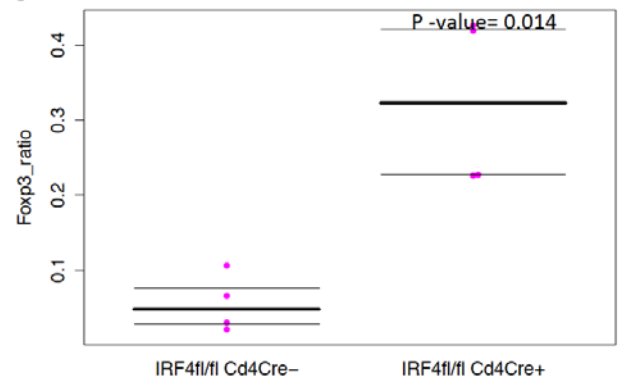


Figure 4.4 IRF4 Controls Th1 Differentiation

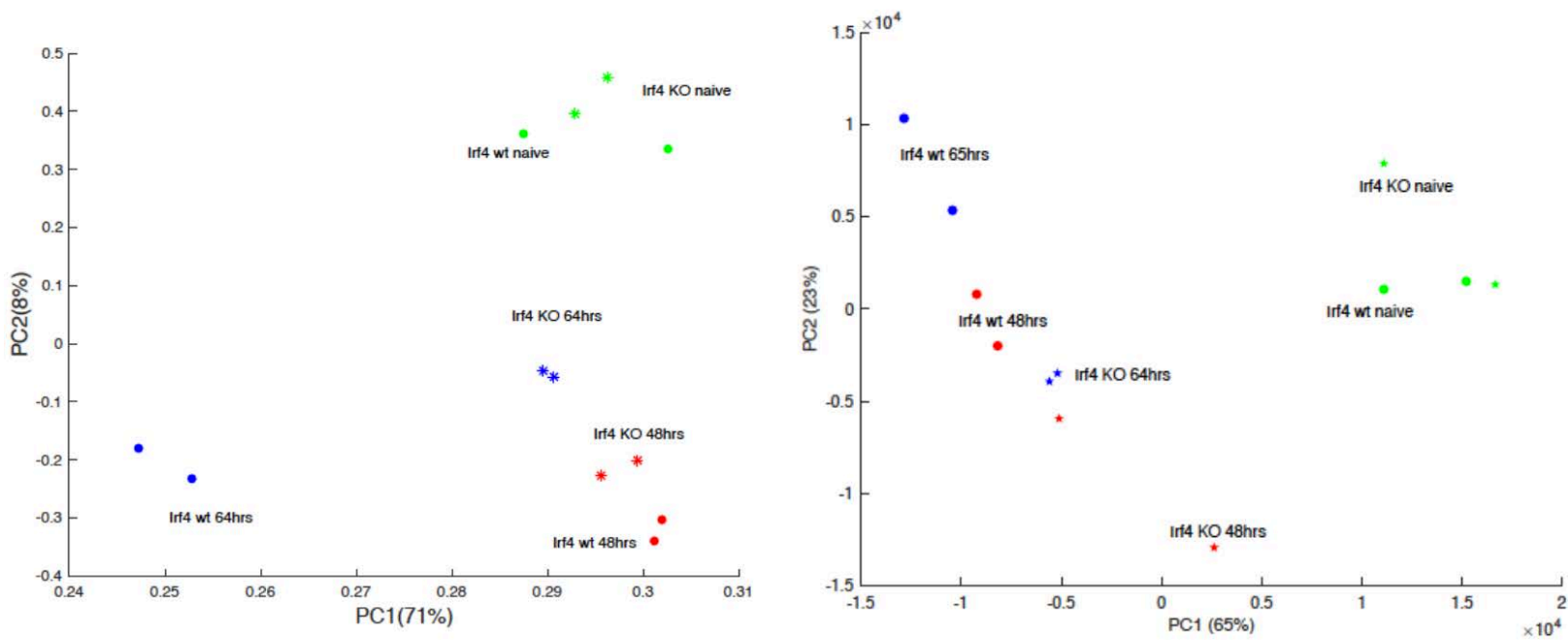
(A) Naive CD4⁺ T cells isolated from IRF4^{fl/fl} CD4-Cre⁺ and IRF4^{fl/fl} CD4-Cre⁻ mice were differentiated to Th1 cells for 48 hours and 64 hours, followed by iTreg culture conditions for additional 24 hours. At the indicated time points (48 + 24 hours, and 64 + 24 hours), T-bet and Foxp3 protein expression levels were determined by FACS. (B) The expression level of Foxp3 for cells cultured in Th1 conditions for 48 hours then re-stimulated in Treg conditions is indicated (n = 4 mice per group). Data are representative of at least three independent experiments. P-values are indicated (Mann-Whitney test). (C) The percentage of Foxp3 expression for cells cultured in Th1 conditions for 64 hours before re-stimulation is indicated (n = 4 mice per group). Data are representative of at least three independent experiments. P-values are indicated (Mann-Whitney test).

IRF4-regulated transcriptional profiles in differentiated Th1 cells

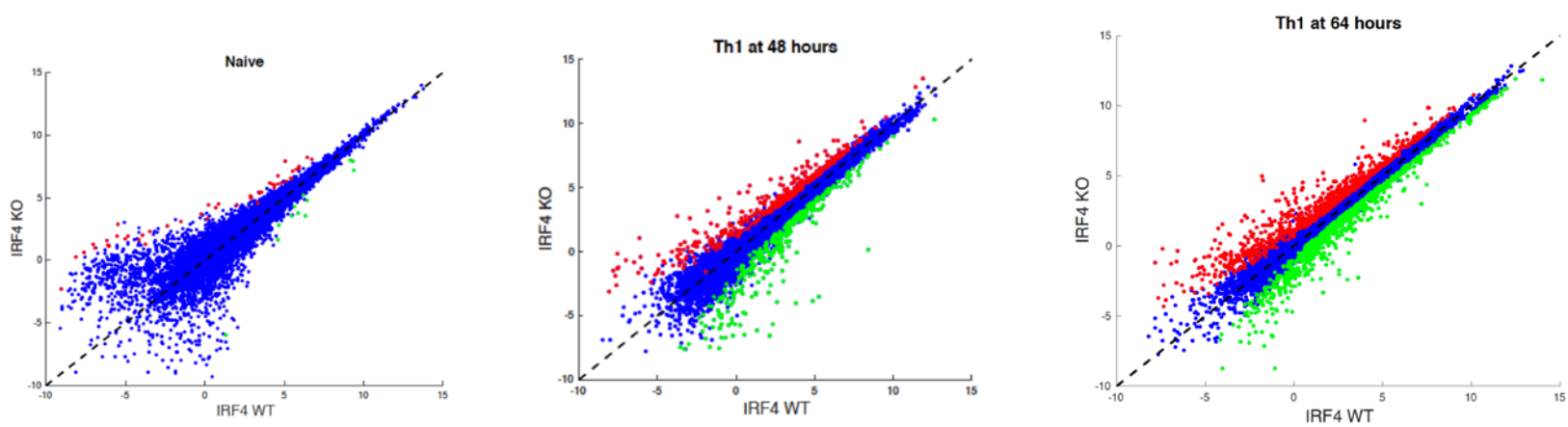
To understand the role of IRF4 in Th1 differentiation, we generated RNA-seq and DNase-seq libraries for naive cells under Th1 conditions for 0, 48, and 64 hours from IRF4^{fl/fl} CD4-Cre⁺ and IRF4^{fl/fl} CD4-Cre⁻ mice. The PCA plots of the RNA-Seq data and DNase-Seq data showed that libraries from IRF4^{fl/fl} CD4-Cre⁻ mice at different time points were separated, which implied the dissimilarity among IRF4^{fl/fl} CD4-Cre⁻ mice libraries at different time points (Figure 4.5(A)). The libraries of IRF4-deleted CD4⁺ T cells at 0 hours were more similar to their corresponding IRF4^{fl/fl} CD4-Cre⁻ control libraries compared to other time points. However, the libraries of IRF4-deleted Th1 cells at 64 hours were located closer to the libraries of IRF4^{fl/fl} CD4-Cre⁻ controls at 48 hours compared to 64 hours IRF4^{fl/fl} CD4-Cre⁻ controls. This observation was in line with our IRF4 deletion experiments *in vitro* (Figure 4.4) and indicated that IRF4 was required for Th1 cell differentiation and lineage commitment. We computed the differential genes between IRF4-deleted Th1 cells and IRF4^{fl/fl} CD4-Cre⁻ controls at 0, 48, and 64 hours. The scatter plots of the read densities in IRF4-deleted Th1 cells and controls including the differential genes were shown in Figure 4.5(B). Overall, there were 50 (15), 373 (321), 872 (847) up-regulated (down-

regulated) genes when comparing IRF4-deleted Th1 cells to controls at 0, 48, and 64 hours, respectively. This result indicated the strong influence of IRF4 to the expression of target genes at 64 hours during differentiation. The heat map of gene ontology analysis for the differential genes at 48 and 64 hours showed that the target genes of IRF4 in 64 hours Th1 cells were more associated with the development of T-helper type 1 cells compared to those in 48 hours Th1 cells (Figure 4.5(C)). Moreover, the conditional deletion of IRF4 in CD4⁺ T cells impacted the expression of T-bet, which is the master transcription factor of Th1 cells, thus delaying Th1 differentiation status (Figure 4.5(D)).

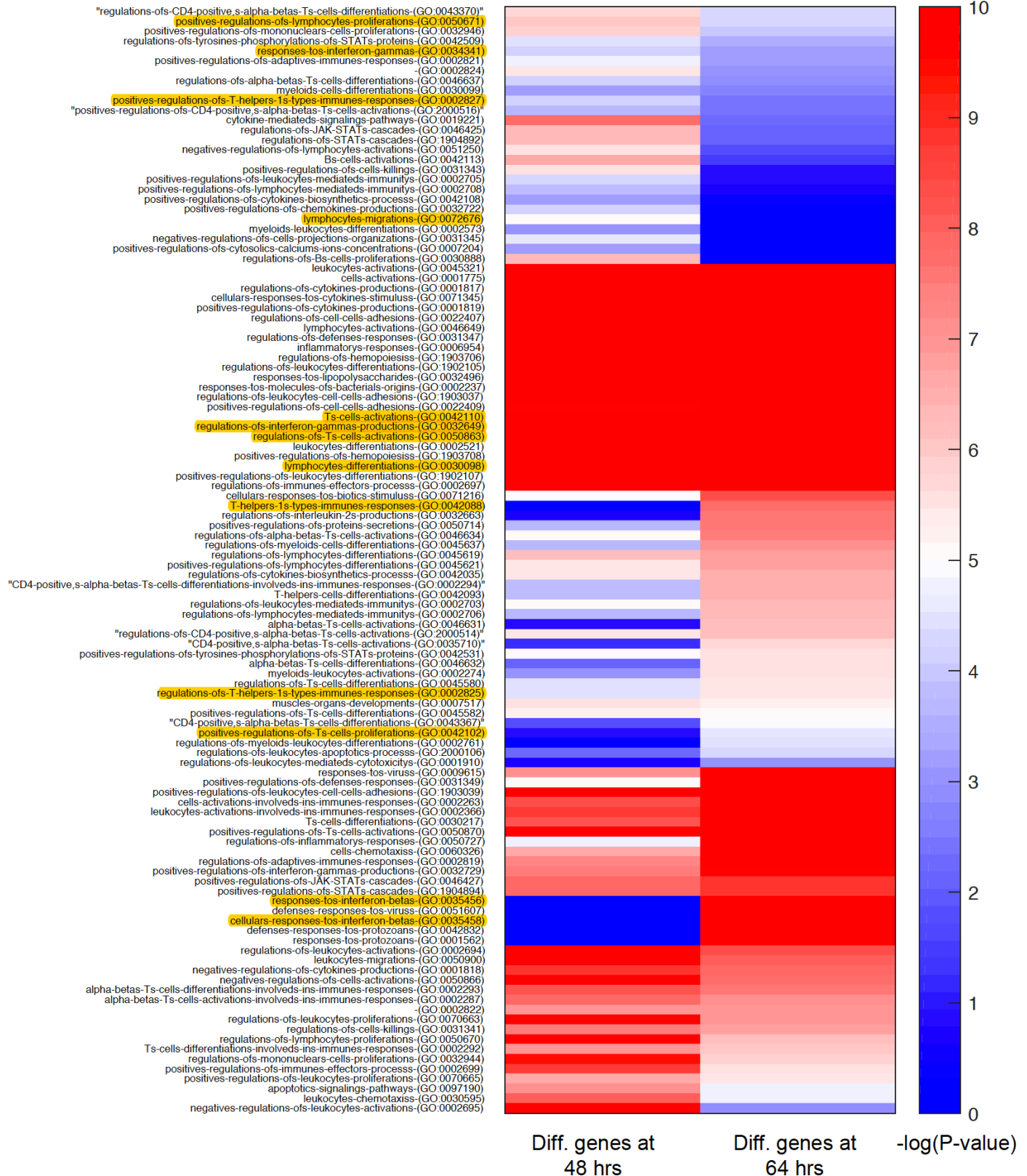
A



B



C



D

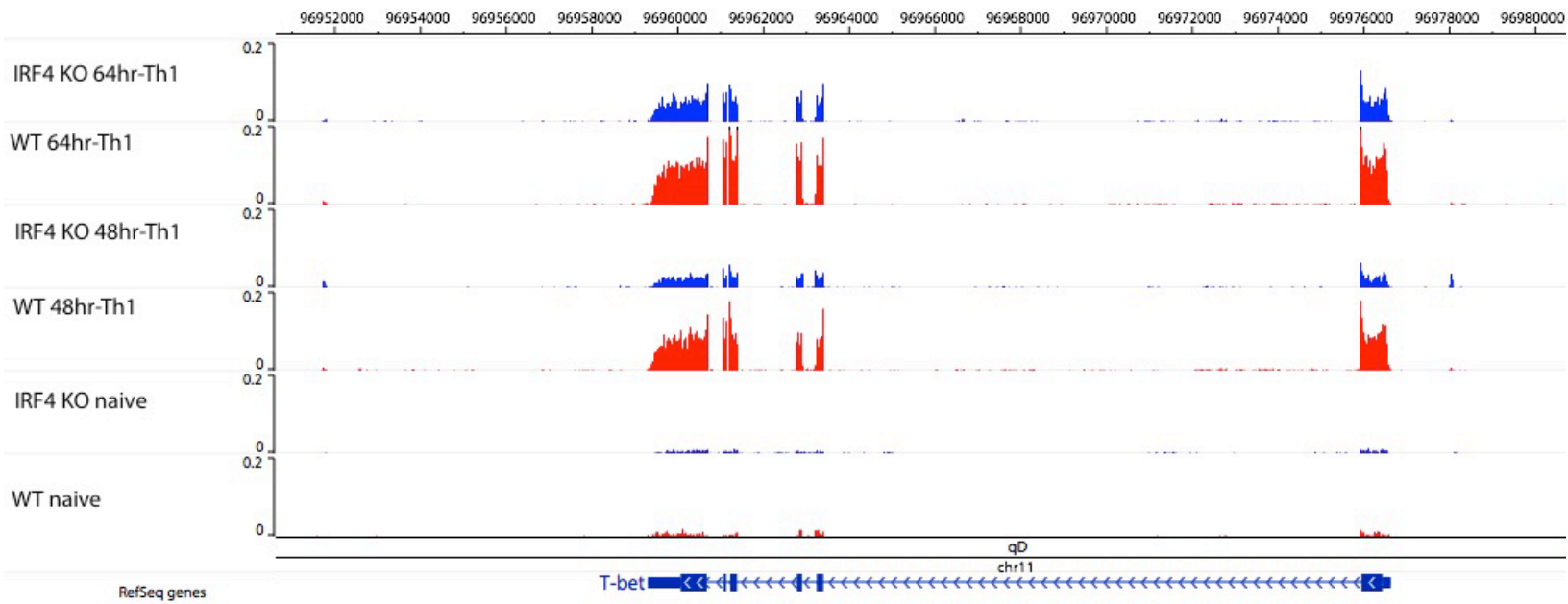


Figure 4.5 Analysis of IRF4-deleted Th1 Cells.

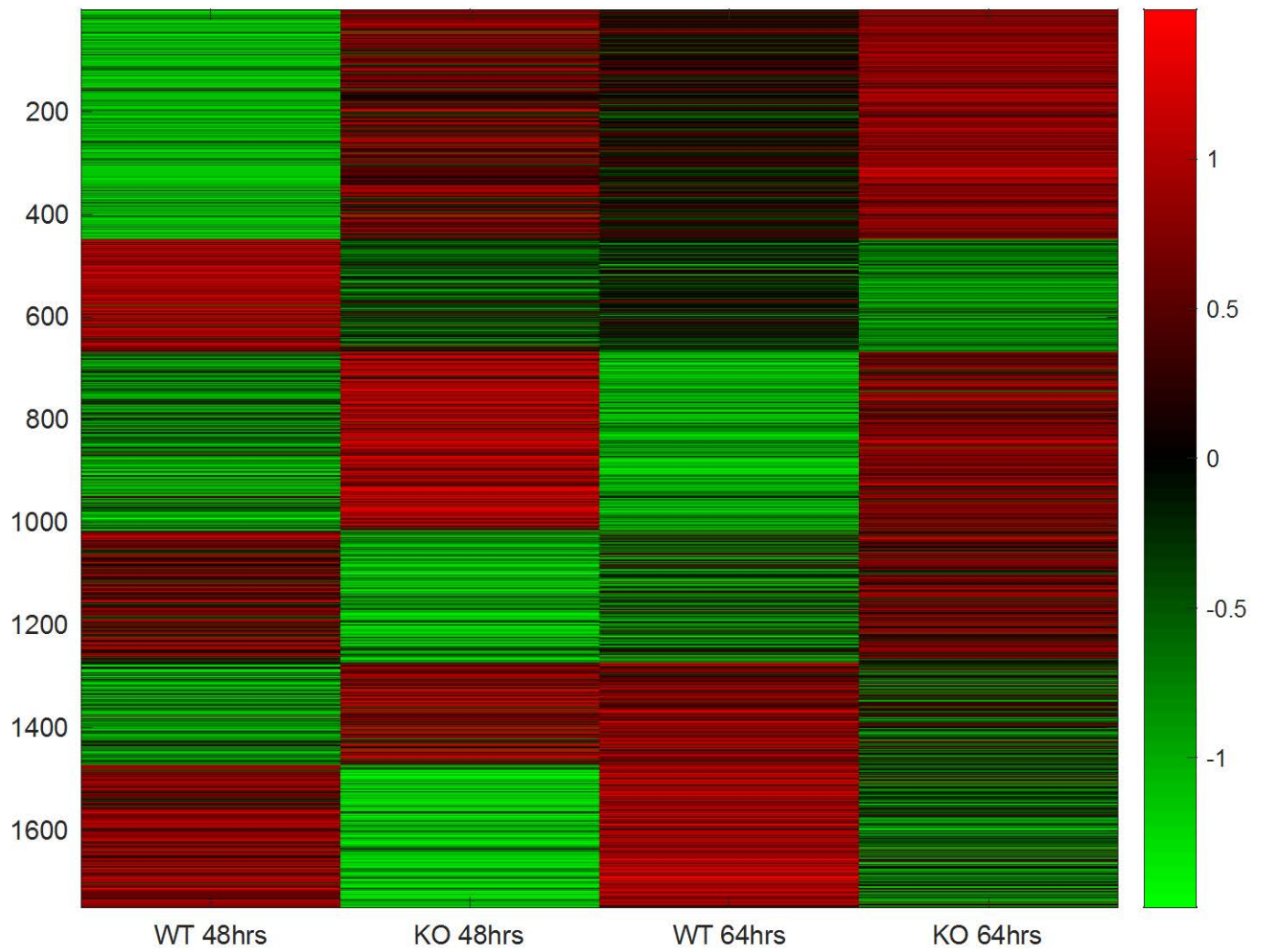
(In collaboration with Dr. Wai Lim Ku from Dr. Keji Zhao's lab in NHLBI, NIH)
 Analysis for IRF4-deleted Th1 cells and IRF4^{fl/fl} CD4-Cre- controls at 0, 48, and 64 hours. (A) PCA plots for control and IRF4-deleted Th1 cells at 0, 48, and 64 hours using DNase-Seq data (**The left panel**) and RNA-Seq data (**The right panel**). (B) The scatter plots of the transcript (in log2 RPKM) in IRF4-deleted cells versus control cells at 0 (left), 48 (middle), and 64 (right) hours. The up-regulated genes are highlighted in red while the down-regulated genes are highlighted in green. The differential genes were selected if FDR<0.05 and abs(fold-change)>1.5. (C) The heat map shows the enriched gene ontology terms for the differential genes (IRF4-deleted Th1 cells vs. IRF4^{fl/fl} CD4-Cre- controls) found at 48 and 64 hours. (D) Genome browser snapshots for the T-bet locus indicate the expression level of T-bet in IRF4-deleted Th1 cells and IRF4^{fl/fl} CD4-Cre- control cells at 0, 48, and 64 hours.

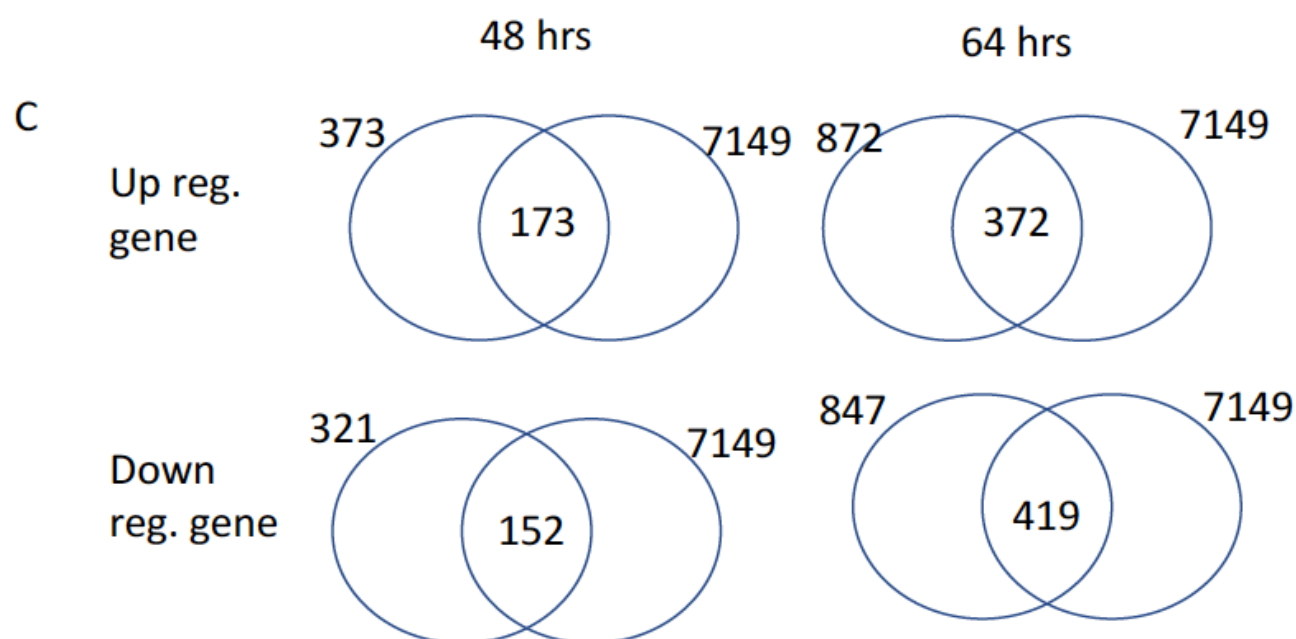
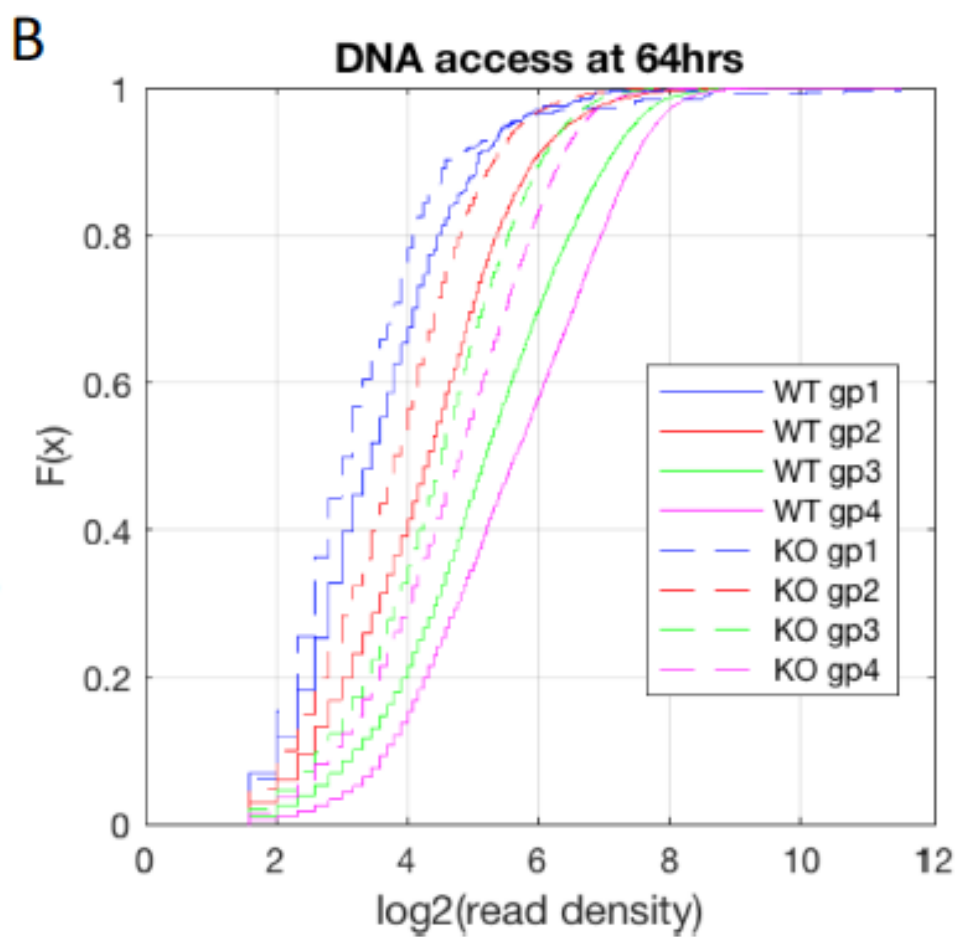
Next, we summarized the differential genes (IRF4-deleted Th1 cells vs. IRF4^{fl/fl} CD4-Cre- control cells) at 48 and 64 hours and performed the K-medoids clustering method to the normalized gene expression matrix (with k=6) for both IRF4-deleted Th1 cells and IRF4^{fl/fl} CD4-Cre- control cells (Figure 4.6(A)). 7149 genes were annotated with the IRF4 binding using the IRF4 ChIP-seq data in CD4 T cells

cultured under Th1 cell-polarizing conditions for 3 days. Figure 4.6(C) showed the overlap between those differential genes and IRF4 binding genes, indicating that IRF4 might influence Th1 differentiation by direct binding to both up-regulated and down-regulated genes. This observation suggested that IRF4 could be either a transcriptional activator or repressor.

To investigate whether the IRF4 binding influence chromatin states, we further separated the DHSs into four groups based on the IRF4 binding levels: group1 ($\text{rpkm} \geq 0$ and $\text{rpkm} \leq 5$), group 2 ($\text{rpkm} \geq 5$ and $\text{rpkm} \leq 10$), group 3 ($\text{rpkm} \geq 10$ and $\text{rpkm} \leq 15$), and group 4 ($\text{rpkm} \geq 15$). We then compared chromatin accessibility measured by the reads density between IRF4^{fl/fl} CD4-Cre- control cells and IRF4-deleted Th1 cells at 64 hours. Figure 4.6(B) showed that IRF4 had a positive role in enhancing accessibility. We tried to detect the “true” differential genes under the influence of IRF4 deletion by only including those genes with both IRF4 bindings and accessibility either in IRF4-deleted Th1 cells or IRF4^{fl/fl} CD4-Cre- controls. Therefore, finally, the up-regulated and down-regulated genes were narrowed down to about 300-400 genes. Gene ontology analysis was applied to these two sets of genes. Figure 4.6(D) showed that the final gene lists were highly related to the T cell activation processes, which suggested the dual function of IRF4 in both T cell activation and differentiation processes.

A





D

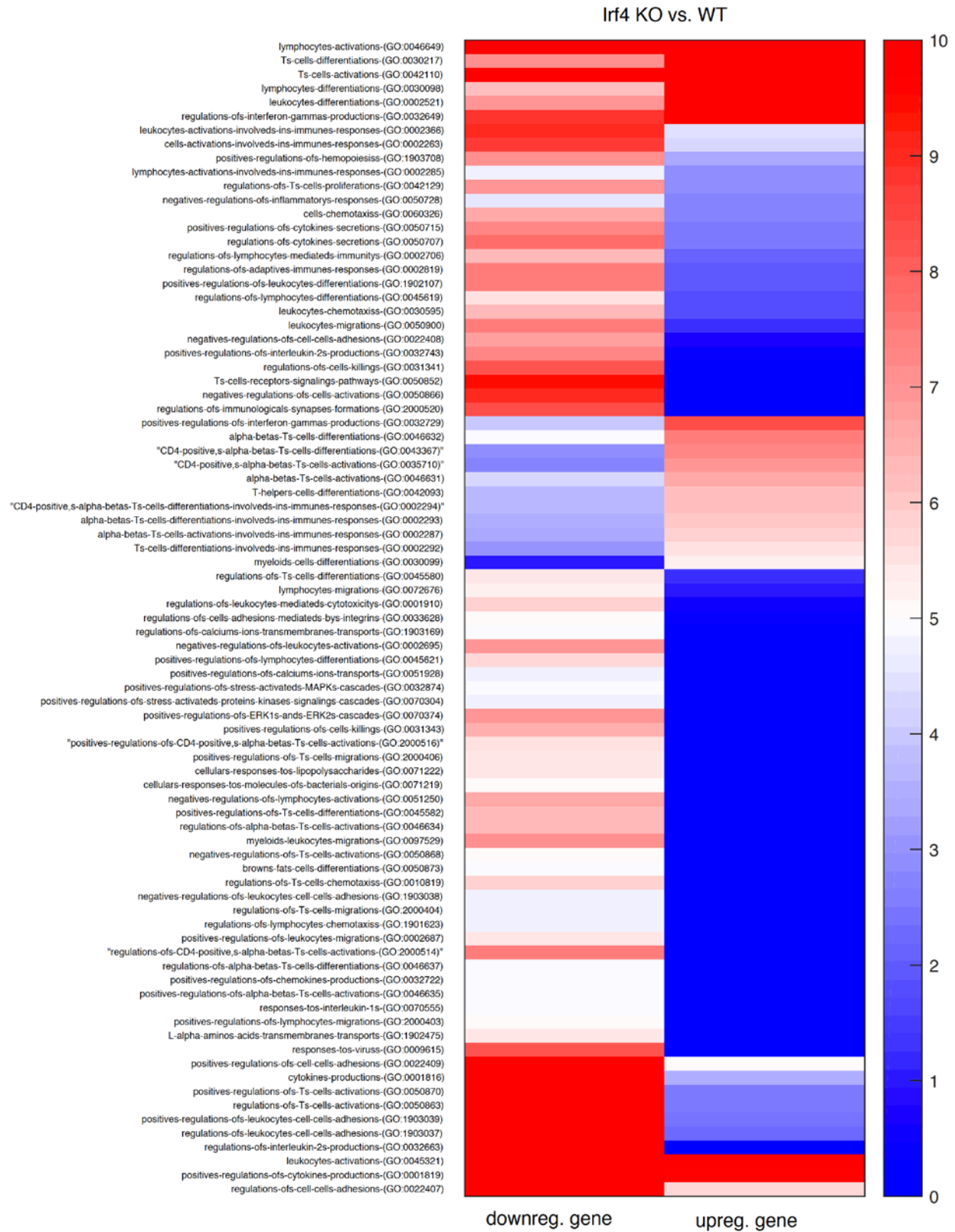


Figure 4.6 IRF4 Role During Th1 Differentiation

(In collaboration with Dr. Wai Lim Ku from Dr. Keji Zhao's lab in NHLBI, NIH)

(A) We pooled the differential genes (IRF4-deleted Th1 cells vs. IRF4^{fl/fl} CD4-Cre-control cells) at 48 and 64 hours and normalized the gene expression levels by the z-score. The heat map using the K-medoids clustering method shows that some genes were showing opposite expression patterns between 48 hours and 64 hours. (B) DNA hypersensitive sites at 64 hours in both control and IRF4-deleted Th1 cells were annotated with the IRF4 binding in CD4 cells. We separated the DHSs into four groups based on the IRF4 binding levels. The solid line in the CDF plot refers to control cells at 64 hours while the dashed line refers to IRF4-deleted Th1 cells at 64 hours. (C) Up and down-regulated genes at 48 hours and 64 hours were annotated with the IRF4 binding. The Venn diagram shows the overlap between those differential genes and IRF4 binding genes. (D) Up and down-regulated genes at 64 hours were further filtered to about 300-400 genes. The heat map shows the p-value of the enrichment of GO terms.

Three-dimensional architecture is conserved in differentiating cells

The principle of TrAC-looping involves the transposition of bivalent Mosaic End sequences at one genomic region or between two interacting chromatin regions, followed by purification of biotinylated bivalent linker-containing DNA fragments (Figure 4.7). The enriched biotinylated bivalent linker-containing DNA fragments contain information of the long-range chromosome interactions. Followed by amplification and library preparations, those fragments could be identified by paired-end sequencing on the next-generation sequencing platform. By performing one assay, we could simultaneously identify both accessible chromatin regions and chromosome interactions since transposition preferentially happens at DNA hypersensitive sites (DHSs).

We found that TrAC-looping possessed the following characteristics: (1) a lower noise level compared with Hi-C and ChIA-PET; (2) detection of accessible chromatin

regions; (3) discovery of both long-range and intermediate-range interactions; and (4) capture of effective promoter-enhancer interactions (supporting data unpublished).

Because the three-dimensional nuclear structure contributes to gene regulation, it is of great interest to understand how chromatin interactions would change during CD4⁺ T cell differentiation. Hence, we applied TrAC-looping technique to Th1 helper cells at different time points of differentiation (i.e., 48 hours and 72 hours). By analyzing TrAC-looping libraries from naive CD4 T cells, 48-hour Th1 cells, and 72-hour Th1 cells, we did not observe substantial compartment flipping changes as indicated by A- and B-compartments [265] genome-wide as examples shown in Figure 4.8. This indicates that the three-dimensional genome architecture was conserved during Th1 helper cell differentiation and domain boundaries remained largely invariant, at least at 48 hours and 72 hours.

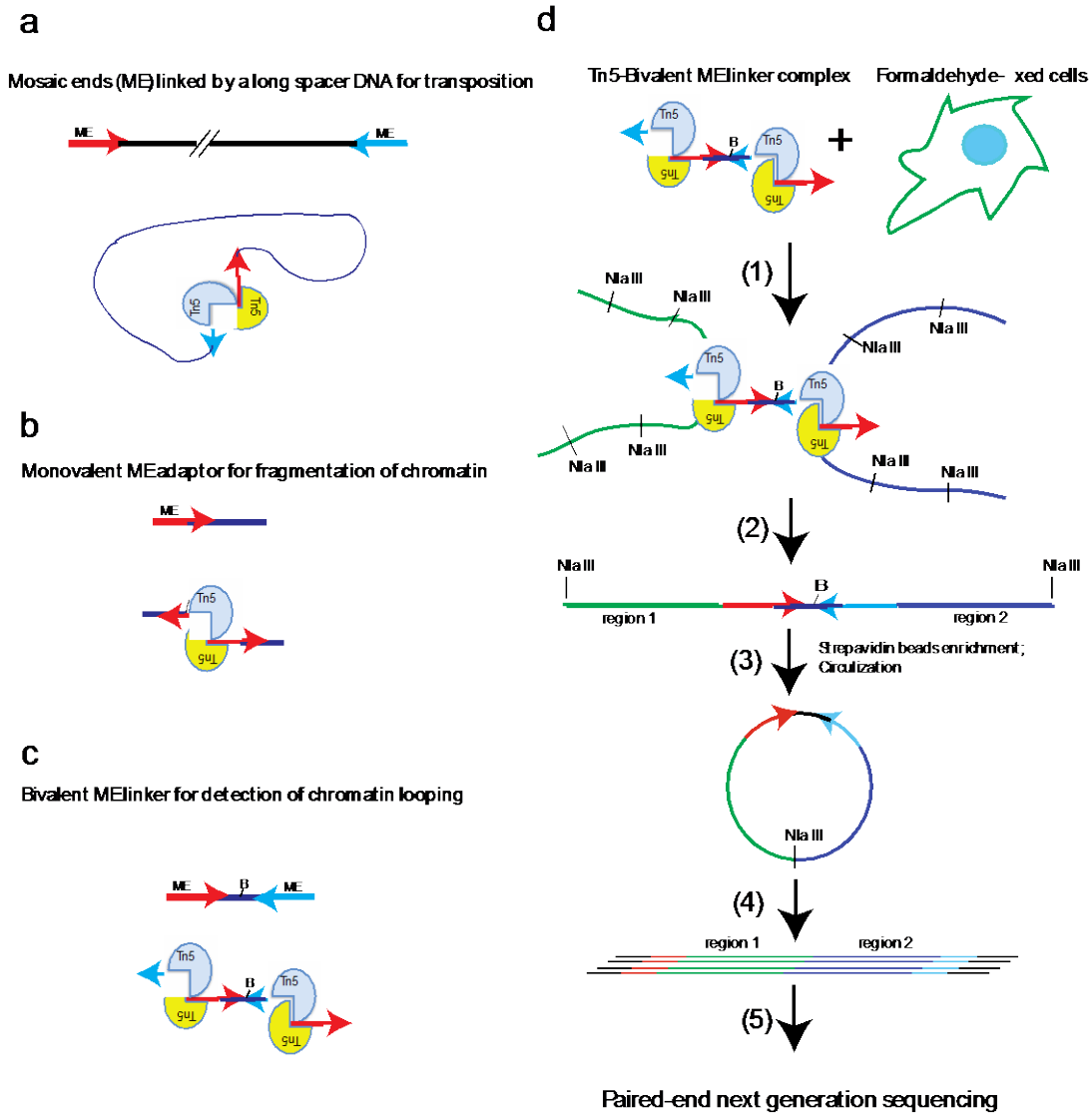


Figure 4.7 TrAC-looping Principle

(In collaboration with Qingsong Tang from Dr. Keji Zhao's lab in NHLBI, NIH)

(A) The simplified structure of Mosaic ends (ME) sequences and Tn5 transposase complexes. (B) The simplified structure of monovalent ME linkers and dimeric Tn5 complexes used in ATAC-seq and other DNA fragmentation applications. (C) The simplified structure of the bivalent ME linker, the monovalent ME adaptor and the hyperactive Tnp enzyme assembling tetrameric Tn5 complex used in TrAC-looping. (D) The pipeline of performing the TrAC-looping method to detect chromatin interactions: (1) in trans transposition of the bivalent linker; (2) DNA fragmentation and enrichment of biotinylated DNA; (3) DNA circularization; (4) rolling cycle amplification of circular DNA; (5) paired-end sequencing.

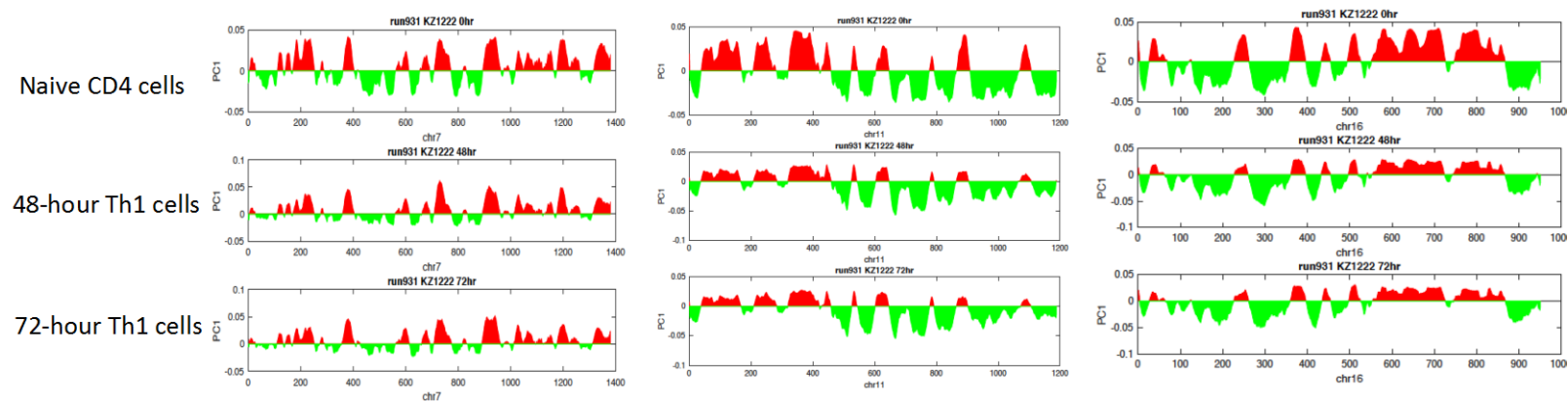


Figure 4.8 Compartment Analysis of Different Time Points

(In collaboration with Dr. Wai Lim Ku from Dr. Keji Zhao's lab in NHLBI, NIH)
 The compartment distributions of naive CD4⁺ T cells, 48-hour Th1 cells, and 72-hour Th1 cells at genomic regions on chromosome 7, 11, and 16. Red peaks stand for the A compartment, and green peaks represent the B compartment.

Local epigenetic environments of the IRF4 locus during differentiation

Although the global domain boundaries stayed stable during CD4⁺ T cell differentiation, we were still interested in investigating whether the “local” chromatin changes occurred at the IRF4 locus. RNA-Seq, DNase-Seq, and TrAc-looping data (Appendix VI and Appendix VII) at the naive CD4⁺ T cell stage, the 48-hour Th1 helper cell stage, and the 72-hour Th1 helper cell stage indicated that substantial epigenetic alterations happened at the IRF4 locus during cell differentiation (Figure 4.9). Especially, the numbers of chromatin looping induced by cell differentiation at the IRF4 locus increased dramatically across time. At the naive stage, IRF4 expression was at the basal level with almost no interactions with other chromatin regions; from 48 hours to 72 hours, the expression level increased accompanied by

more long and immediate chromosome interactions. This observation, in another aspect, confirmed the important role of IRF4 in driving Th1 cell differentiation.

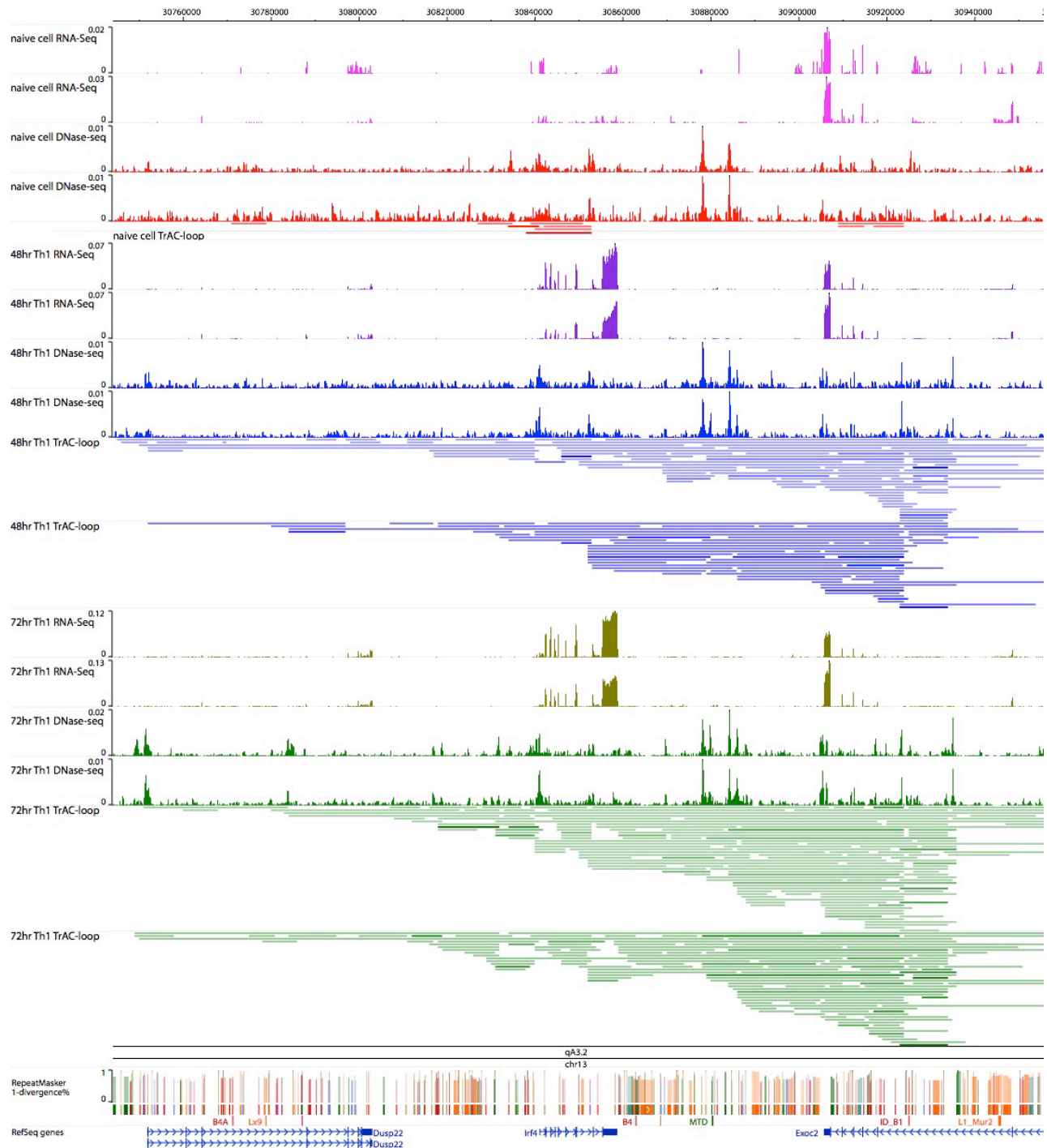


Figure 4.9 Epigenetic Changes at the IRF4 Locus

(In collaboration with Dr. Wai Lim Ku from Dr. Keji Zhao's lab in NHLBI, NIH)
IRF4 gene expression, chromatin accessibility, and spatial interactions in mouse naive CD4⁺ T cells, 48-hour Th1 cells, and 72-hour Th1 cells. The same color panels represent biological replicates. Conditions are indicated in the figure.

Discussion

In this study, we have investigated the plasticity and lineage commitment of Th1 cells by using *in vitro* culture and genome-wide analysis of gene expression, chromatin accessibility, and chromatin interactions. Several interesting discoveries were obtained. First, the early stage T-bet⁺ Th1 cells were plastic and could be reprogrammed to express Foxp3, a marker gene of Treg cells, under the challenge of Treg lineage-defining cytokines. On the contrary, the late stage T-bet⁺ Th1 cells lost the plasticity of expressing Foxp3. Second, the watershed between 48-hour Th1 cells and 72-hour Th1 cells' lineage reprogramming was not entirely dependent on T-bet expression. Third, the chromatin accessibility profiles of Th1 cells at different differentiation stages, especially 48 hours and 72 hours, confirmed the phenotypic difference on the chromatin basis and our modified footprint analysis algorithm implied a functional relationship between the transcription factors and target genes. Among those transcription factors, we predicted the potential regulators, including Mxi1, JunB, Jun, BATF, IRF4, and Hif-1 α , could be essential for the Th1 lineage commitment. And finally, IRF4 was experimentally validated to drive Th1 cell commitment by altering the chromatin accessibility status during differentiation.

The plasticity of CD4⁺ T cells has captured the intense research interest for many years, and Th1 cells were reported to be more susceptible to Th2 cell-favoring manipulations [181], implying that the cells were plastic to change their fate.

However, it remains unclear what factors control such plasticity and lineage commitment. Here, we provide evidence that such plasticity is time-dependent and the chromatin accessibility changes are associated with different differentiation stages. The partially plastic behavior of the Th1 cell lineage is in line with the observations in PDL1-mediated Th1 inhibition [365] and retinoic acid's function as a Th1 fate commitment stabilizer [366]. Hence our results demonstrate the flexibility of the Th1 cell lineage and molecular changes associated with the lineage maintenance.

In addition, to promote Th1 differentiation, IRF4 was found to actively modify chromatin accessibility and control gene expression. It is well known that in Th17 cells, BATF-IRF4 complexes act cooperatively as pioneer factors influencing the expression of crucial Th17 genes and controlling chromatin accessibility to assist STAT3 and ROR γ t binding [195]. BATF-IRF4 complexes are highly expressed in Th2, Th9, and Th17 cells, and some researchers believe that BATF-IRF4 complexes are in the central regulatory node of lineage-specific transcription factor networks. BATF deficiency does not affect Th1 cell differentiation [367]; however, our data suggest BATF and IRF4 are the primary stabilizers of Th1 cell lineage commitment. Induction of BATF and IRF4 may limit plasticity in early Th1 cells and silence other lineage-associated genes. Another transcription factor Hif-1 α was reported to control the lineage choice between Th17 and Treg cells by binding Foxp3 and ROR γ t simultaneously but functioning differently [368]. It is possible that Hif-1 α would influence the committed Th1 cell status by binding to critical Th1 associated genes. It will be of great interest to study transcriptional targets of BATF, IRF4, and Hif-1 α in Th1 cells.

It is reasonable to assume the phenotype we observed in our IRF4 deletion experiment could be caused by the destruction of BATF-IRF4 complexes. We hypothesize that BATF-IRF4 complexes are critical to maintain Th1-cell fates and restrain the conversion to other effector lineages. Although the previous study suggested that differentiated Th1 effectors could not efficiently develop into Foxp3+ Tregs *in vivo* [352], the “controversy” observed in our research might, in turn, indicate the significant influence of time on T helper cell differentiation. Meanwhile, Foxp3+ Tregs were able to convert into Th1 [369] and Th17 [162] cells *in vivo*. All information delineates the considerably various flexibility and inter-convertibility of differentiated CD4+ T cell populations. The question of whether there is a universal mechanism explaining the diverse plasticity among different cell types or such behavior is highly cell-type specific is still unanswered. At least, our research shows that IRF4 (highly probably BATF-IRF4 complexes) plays an important role in modulating Th1 cell plasticity. In the absence of IRF4 during Th1 cell differentiation, cells become unstable and susceptible to acquire Th17 cells phenotypes.

The finding in our study that no dramatic 3D genome architecture changes occurred during Th1 helper cell differentiation is in line with previous studies showing that chromatin structures, especially the topologically associating domains (TADs), remained constant among different cell types in the same species. The question of whether TADs are inherently distinct from sub-TADs, loops, or insulation neighborhoods remains unanswered in the field. Some researchers hypothesize that TADs are fundamentally different from sub-TADs, loops, and insulation neighborhoods because those “subunits” at least partially differ between different cell

lineages and may regulate lineage-specific gene expression while TADs serve as barriers of the physically isolated units of genome organization. Indeed, the alterations of local chromatin interactions observed at the IRF4 locus during Th1 cell differentiation may provide another piece of evidence to the inherent differences between TADs and their “subunits”.

In summary, our study gives insight into the mechanism balancing Th1 cell stability and plasticity in their gene expression programs and chromatin accessibility. A caveat is that our findings are derived from *in vitro* experiments, and the extreme polarizing conditions may exaggerate CD4⁺ T cell plasticity. It will be meaningful to substantiate our key discoveries from *in vitro* studies under physiological conditions. *In vivo* tests of IRF4’s impact on Th1 cell differentiation would be anticipated, and future experiments to evaluate the functionality of Th1 derived “Treg” cells and their implications in the pathogenesis of autoimmune diseases [365] may be informative.

Chapter 5 : Conclusions and Future Directions

The long-term pursuit of our lab is to integrate the knowledge from statistical genetics and computational epigenetic methods to decipher the interaction between animal genomes and environmental factors, and their relationship to disease-prevention and performance. In my graduate research, my general interest falls into several aspects: 1) how CD4⁺ T cells in chickens change their epigenetic status to influence the transcriptional activity and their different phenotypes towards MDV infection; 2) how CD4⁺ T cells, especially Th1 helper cells, undergo differentiation and commitment towards their cell fates; and 3) what are the underlying epigenetic machinery controlling cellular plasticity and responses towards viral infection. The work presented here contributes some novel insights into the above topics.

We first studied the effect of late cytolytic MDV infection on gene expression in three immune organs of two inbred chicken lines showing distinct degrees of resistance to MD. The transcriptome profile suggested the different behavior of spleen, bursa, and thymus when confronting MDV. Integration of differential gene expression and pathway analyses enabled the discovery of immune-related factors involved in MD resistance. For instance, Herpes simplex infection and Influenza A pathways were discovered in line 6₃ chickens, whereas pathways associated with metabolism, cell cycle, and DNA replication were observed in line 7₂ chickens. Our research provided further insight into the mechanism of MD progression and underlined the significance of genetic research behind complex diseases.

The transcriptional changes in chicken tissues may not adequately represent the cellular response against viral infection. Being the primary targets of Marek's disease

virus, CD4⁺ T cells are both the victims and the fighters against viral infection. It is highly reasonable to assume the innate differences between CD4⁺ T cells of the two chicken lines distinguish the symptom of disease progression. Our second study on chicken CD4⁺ T cells, in both transcriptome and chromatin accessibility aspects, observed more significant transcriptional activity changes at 10 DPI. The increase of chromatin accessibility and 3D genome architecture changes represented by A/B compartment flipping were associated with the specifically up-regulated genes in line 6₃ chicken CD4⁺ T cells induced by MDV infection at 10 DPI. The epigenetic alterations in CD4⁺ T cells from resistant chickens may contribute to the disease outcomes.

Cytokines directing Th1 (infection with intracellular pathogens) and Th2 types (infection of extracellular pathogens) of the immune response in the avian immune system have been reported to mimic the similar process in mammals [370]. However, the limitations of knowledge and techniques available for separating chicken CD4⁺ lineages restrain more in-depth study on cell differentiation and commitment. Hence, the next step in our exploration is the extension of our efforts to investigate the dynamic epigenetic changes occurred during Th1 differentiation in mice. Our *in vitro* studies discovered that the differentiation status of Th1 cells was not static and cells differently behaved when confronting a subsequent micro-environmental change during differentiation over time. The difference between 48-hour “plastic” Th1 cells and 72-hour “committed” Th1 cells was not directly caused by T-bet since at both stages T-bet expression levels were relatively high. Our integrative analysis at different time points indicated that the genome-wide transcriptional regulatory

networks were controlling Th1 activation and differentiation processes. Our modified footprinting analysis revealed potential Th1 regulators such as BATF, Jun, IRF4, Hif-1 α , Foxm1, Mxi1, and Mycn. The functional validation of IRF4 confirmed our integrative analysis results. Indeed, our ChIP-Seq assays confirmed that the binding of IRF4 could serve as both transcriptional activation and repression, and IRF4 was involved both in T cell activation and differentiation. Our discoveries gave further insights into CD4⁺ T cell lineage commitment and implications for disease pathogenesis associated with intracellular infection from bacteria and viruses.

Since our Hi-C experiment has revealed the dynamic interacting regions after MDV infection, one future direction of our study would be the functional validation of those changes in chromosome organization and transcriptional activity in chicken CD4⁺ T cells. 5C methods and three-dimensional fluorescence in situ hybridization (FISH) could be used to confirm the spatial organization of targeted regions in the nucleus impacted by MDV infection. At the current stage the transgenic and knockout model systems are not well established in chickens; a possible future exploration of our study could be the establishment of conditional gene deletion chickens to validate biological functions. One common technique applied in chicken models involves RNA interference (RNAi) to knockdown interest genes *in vitro*, although it might not truly represent the physiological response after viral infection. It is always preferable to test candidate targets *in vivo* to truly understand their functions under physiological conditions after MDV infection. Also, the extensive study on mouse CD4⁺ T cell differentiation cannot be performed on the chicken model in our lab due to some funding and resource limitations at this moment. Nevertheless, it is of great interest to

exam the subsets of chicken CD4⁺ T cells under the influence of MDV infection. It will be promising to study gene functions both *in vitro* and *in vivo* associated with resistance to MDV infection.

Another possible direction of our future research is to decipher the regulatory network during CD4⁺ T cell differentiation. We will systematically examine Th1 helper cells and identify novel proteins and genes with respect to T cell differentiation and plasticity. Although we discovered several candidate genes, including BATF, IRF4, and Hif-1 α , we believe there remain undiscovered factors controlling this complicated biological process. Our current integrative study on mouse Th1 cells mainly relies on the available transgenic and knockout mice models. Until now, only the dynamic regulatory network of Th17 subtypes has been validated systematically [195, 196] using transgenic and knockout mice models. Alternatively, the *in vivo* small RNA interference screen on mouse T cells [371] or the currently-developed *in vivo* CRISPR screening could facilitate our future studies. High knockout efficiency might be reached by introducing retroviral CRISPR sgRNAs into T cells from mice constitutively expressing Cas9. Also, it will be of great interest to evaluate the protein function *in vitro* and *in vivo* after identifying their critical roles in Th1 cell differentiation and plasticity. Gene knockouts and knockdowns stop the expression of protein indirectly and may cause delayed protein disruption and initiation of compensatory mechanisms [372]. Considering the possible disadvantages of both knockouts and small RNAi knockdowns, it may be worthy of trying the new technique Trim-Away [372], which does not require prior DNA or mRNA modification, to directly degrade mammalian endogenous proteins. We anticipate

discovering and validating the CD4⁺ cells' dynamic regulatory network in our subsequent future studies.

Appendices

Appendix I. Top Pathways of Differentially Expressed Genes Between MDV Infected and Uninfected Chickens

Lines	Organ	Category	Term	Count	%	P-value
L63	Spleen	KEGG_PATHWAY	Phagosome	18	2.5	4.3E-4
		KEGG_PATHWAY	Cytokine-cytokine receptor interaction	21	2.9	6.3E-4
		KEGG_PATHWAY	Herpes simplex infection	19	2.6	6.43E-4
		KEGG_PATHWAY	Influenza A	16	2.2	4.2E-3
L72	Spleen	KEGG_PATHWAY	Cytokine-cytokine receptor interaction	76	1.8	2.9E-5
		KEGG_PATHWAY	DNA replication	21	0.5	4.7E-5
		KEGG_PATHWAY	Focal adhesion	80	1.9	2.0E-4
		KEGG_PATHWAY	Proteasome	24	0.6	3.5E-5
		KEGG_PATHWAY	Cell cycle	50	1.2	1.2E-3
		KEGG_PATHWAY	Cell adhesion molecules (CAMs)	38	1.0	2.2E-2
		KEGG_PATHWAY	Glycosphingolipid biosynthesis	9	0.2	2.0E-2
		KEGG_PATHWAY	Gap junction	34	0.8	2.2E-2
L72	Thymus	KEGG_PATHWAY	Drug metabolism	6	1.1	2.1E-3
		KEGG_PATHWAY	Linoleic acid metabolism	5	0.9	2.2E-3
		KEGG_PATHWAY	Cytokine-cytokine receptor interaction	13	2.3	3.0E-3
		KEGG_PATHWAY	Arachidonic acid metabolism	6	1.1	3.5E-3
		KEGG_PATHWAY	PPAR signaling pathway	6	1.1	4.2E-2
		KEGG_PATHWAY	Metabolism of xenobiotics by cytochrome P450	4	0.7	6.7E-2
		KEGG_PATHWAY	Retinol metabolism	4	0.7	6.7E-2
L63	Bursa	KEGG_PATHWAY	Focal adhesion	8	4.7	4.3E-4
		KEGG_PATHWAY	Cardiac muscle contraction	4	2.4	8.0E-3

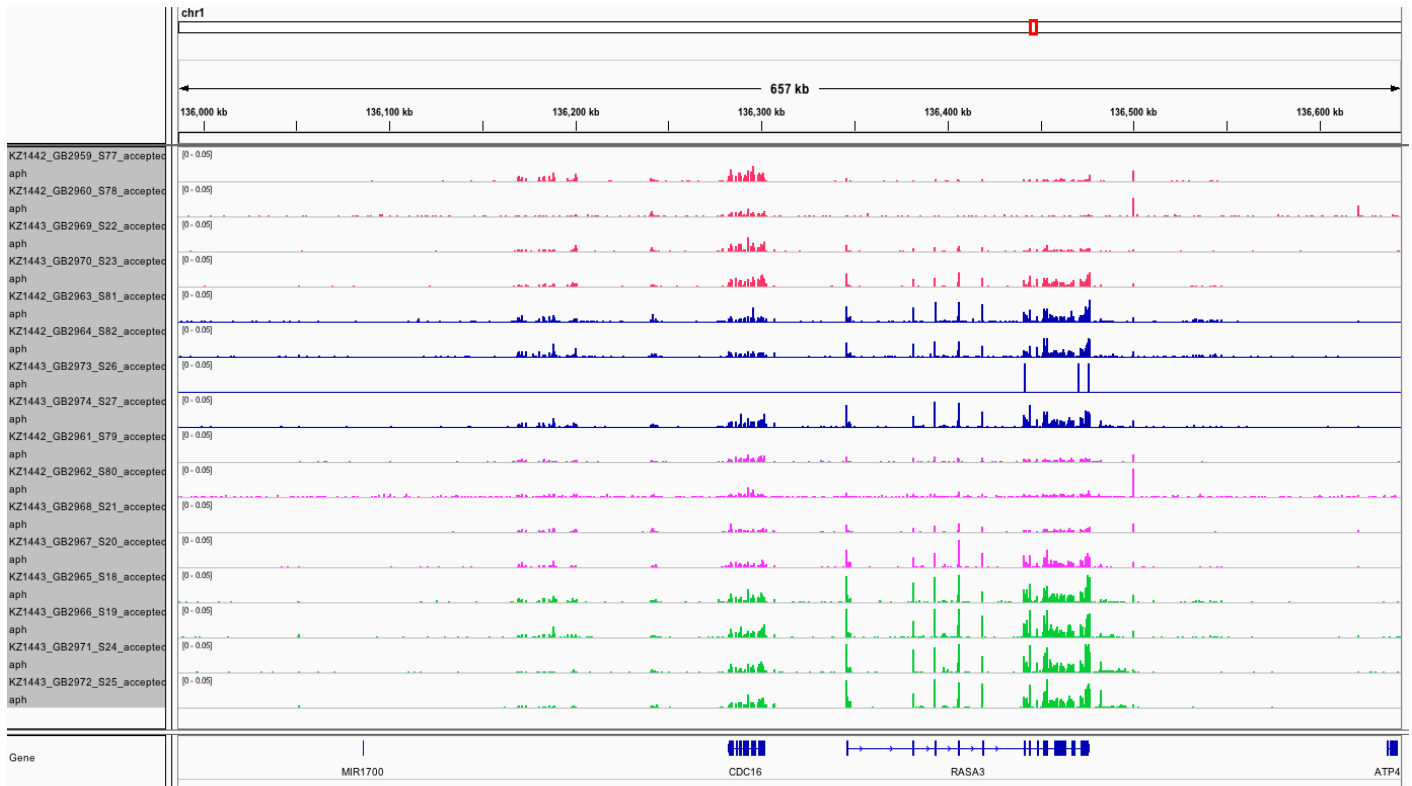
		KEGG_PATHWAY	Tight junction	5	3.0	1.1E-2
		KEGG_PATHWAY	Regulation of actin cytoskeleton	5	3.0	5.3E-2
L72	Bursa	KEGG_PATHWAY	Cell cycle	52	1.2	6.6E-5
		KEGG_PATHWAY	DNA replication	17	0.4	9.6E-4
		KEGG_PATHWAY	Drug metabolism	17	0.4	1.8E-3
		KEGG_PATHWAY	Focal adhesion	69	1.6	4.6E-3
		KEGG_PATHWAY	Cytokine-cytokine receptor interaction	57	1.3	9.3E-3
		KEGG_PATHWAY	Mismatch repair	10	0.2	2.3E-2
		KEGG_PATHWAY	Natural killer cell mediated cytotoxicity	29	0.7	2.5E-2
		KEGG_PATHWAY	Keratan sulfate biosynthesis	9	0.2	2.7E-2
		KEGG_PATHWAY	Lysosome	38	0.9	2.8E-2
		KEGG_PATHWAY	Metabolism of xenobiotics by cytochrome P450	14	0.3	4.2E-2
		KEGG_PATHWAY	P53 signaling pathway	24	0.5	5.2E-2
		KEGG_PATHWAY	Glutathione metabolism	16	0.4	5.5E-2
		KEGG_PATHWAY	ECM-receptor interaction	29	0.7	6.7E-2

Appendix II. Mapping Quality of Chicken CD4+ T Cell RNA-Seq Data

Samples	Total Reads	Mapped Reads	Mappability
Chicken CD4+ 5DPI 72I1	23675171	18890960	0.8
Chicken CD4+ 5DPI 72I2	32455921	23897488	0.74
Chicken CD4+ 5DPI 72I3	24892194	14740125	0.59
Chicken CD4+ 5DPI 72I4	31803542	20305020	0.64
Chicken CD4+ 5DPI 63I1	26762477	21024347	0.79
Chicken CD4+ 5DPI 63I2	28946935	21917892	0.76
Chicken CD4+ 5DPI 63I3	27062574	15779601	0.58
Chicken CD4+ 5DPI 63I4	26199506	13025392	0.5
Chicken CD4+ 5DPI 72N1	28754814	22022737	0.77
Chicken CD4+ 5DPI 72N2	30587769	25968512	0.85
Chicken CD4+ 5DPI 72N3	33784268	21849043	0.65
Chicken CD4+ 5DPI 72N4	34235424	23342131	0.68
Chicken CD4+ 5DPI 63N1	28018733	22241946	0.79
Chicken CD4+ 5DPI 63N2	27965138	19584236	0.7
Chicken CD4+ 5DPI 63N3	33211607	17925904	0.54
Chicken CD4+ 5DPI 63N4	31316580	18674368	0.6
Chicken CD4+ 10DPI 72I1	41813064	30006796	0.72
Chicken CD4+ 10DPI 72I2	31704926	17644144	0.56
Chicken CD4+ 10DPI 72I3	36879805	28092494	0.76
Chicken CD4+ 10DPI 72I4	42102517	33128832	0.79
Chicken CD4+ 10DPI 63I1	28120963	18349904	0.65
Chicken CD4+ 10DPI 63I2	27912984	16584697	0.59
Chicken CD4+ 10DPI 63I3	29461980	22322684	0.76
Chicken CD4+ 10DPI 63I4	30665325	20670219	0.67
Chicken CD4+ 10DPI 72N1	31909727	23495082	0.74
Chicken CD4+ 10DPI 72N2	32321537	23669101	0.73
Chicken CD4+ 10DPI 72N3	58097	4040	0.07
Chicken CD4+ 10DPI 72N4	37065732	29907198	0.81
Chicken CD4+ 10DPI 63N1	36239778	29294207	0.81
Chicken CD4+ 10DPI 63N2	32755078	26547886	0.81
Chicken CD4+ 10DPI 63N3	33257675	26026680	0.78
Chicken CD4+ 10DPI 63N4	33718178	26144844	0.78

Appendix III. Genome Browser View of Chicken 10 DPI RNA-Seq Data

10 days post infection CD4⁺ T cell RNA-Seq data at two gene locations (CDC16 and RASA3). Red tracks stand for line 7₂ infected chicken groups; blue tracks are representative of line 7₂ control (non-infected) chicken groups; pink tracks describe line 6₃ infected chicken groups; green tracks are from line 6₃ control (non-infected) chicken groups.



Appendix IV. Sequencing Statistics of Chicken CD4+ T Cell ATAC-Seq and Hi-C

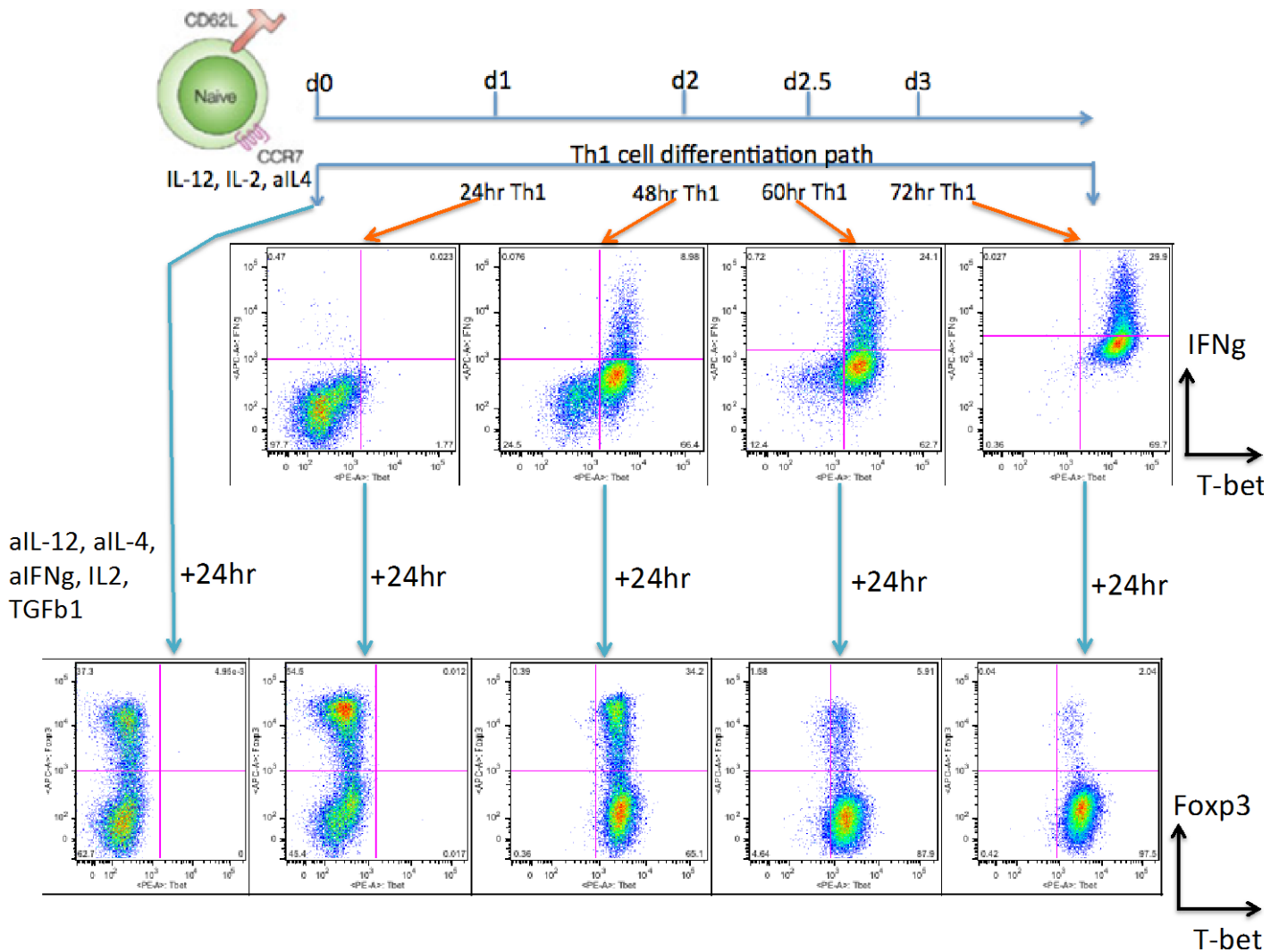
Data Type	Infection Days	Treatment	mapped reads	unique reads	mappability	redundancy
ATAC-Seq	5DPI	72I1	11768969	2482436	0.26	0.79
ATAC-Seq	5DPI	72I2	11694089	2739462	0.25	0.77
ATAC-Seq	5DPI	63I1	13837098	3050443	0.32	0.78
ATAC-Seq	5DPI	63I2	11258562	2436313	0.23	0.78
ATAC-Seq	5DPI	72N1	17496378	4869838	0.34	0.72
ATAC-Seq	5DPI	72N2	13423037	3460935	0.29	0.74
ATAC-Seq	5DPI	63N1	15395060	3231288	0.29	0.79
ATAC-Seq	5DPI	63N2	18969056	4755214	0.39	0.75
ATAC-Seq	5DPI	72I3	3013740	518033	0.07	0.83
ATAC-Seq	5DPI	72I4	2470761	382260	0.05	0.85
ATAC-Seq	5DPI	63I3	2388284	417845	0.05	0.83
ATAC-Seq	5DPI	63I4	3584420	507549	0.07	0.86
ATAC-Seq	5DPI	72N3	9084117	2309942	0.21	0.75
ATAC-Seq	5DPI	72N4	2512559	570865	0.06	0.77
ATAC-Seq	5DPI	63N3	3354328	766876	0.07	0.77
ATAC-Seq	5DPI	63N4	3568311	605177	0.07	0.83
ATAC-Seq	10DPI	72I1	12274856	3009029	0.26	0.75
ATAC-Seq	10DPI	72I2	13310045	3652150	0.29	0.73
ATAC-Seq	10DPI	63I1	16575201	3448491	0.31	0.79
ATAC-Seq	10DPI	63I2	9163130	2278413	0.21	0.75
ATAC-Seq	10DPI	72N1	21771969	5322494	0.50	0.76
ATAC-Seq	10DPI	72N2	24470220	5546747	0.52	0.77
ATAC-Seq	10DPI	63N1	18520872	3303014	0.38	0.82
ATAC-Seq	10DPI	63N2	22611940	5655647	0.50	0.75
ATAC-Seq	10DPI	63I3	16744357	4594846	0.39	0.73
ATAC-Seq	10DPI	63I4	25251268	5781834	0.55	0.77
ATAC-Seq	10DPI	72I3	19351536	4313788	0.44	0.78
ATAC-Seq	10DPI	72I4	20619577	4061553	0.40	0.80
ATAC-Seq	10DPI	63N3	23144451	5364684	0.49	0.77
ATAC-Seq	10DPI	63N4	17025848	3573780	0.29	0.79
ATAC-Seq	10DPI	72N3	17267014	4856067	0.42	0.72
ATAC-Seq	10DPI	72N4	23403197	4640971	0.48	0.80

Hi-C Samples	Total reads	mapped reads	mappability	unique reads	IntraChr	IntraChr/unique
10DPI_63N1	45536461	30113889	0.66	23230219	18706987	0.81
10DPI_63N2	40621486	26789642	0.66	13399616	10549629	0.79
10DPI_72I1	34982031	25701646	0.73	19197578	17339389	0.90
10DPI_72I2	42421931	30199492	0.71	23394345	19991098	0.85
10DPI_72N1	49689947	33074781	0.67	20982400	17627393	0.84
10DPI_72N2	66960608	45130153	0.67	33354751	28531923	0.86
10DPI_63I1	53331385	35951997	0.67	27753266	23580232	0.85
10DPI_63I2	56989261	37845693	0.66	29987296	25597176	0.85

Hi-C Samples	InterChr	<1K	1KTo2M	1KTo2M/intraChr	>2M
10DPI_63N1	4523232	8303134	5304903	0.283578697	5644457
10DPI_63N2	2849987	4020252	3463206	0.328277516	3382986
10DPI_72I1	1858189	12691551	2792973	0.161076783	2087546
10DPI_72I2	3403247	13609519	3501603	0.175158113	3209190
10DPI_72N1	3355007	6938310	6211333	0.352368215	4975174
10DPI_72N2	4822828	14590065	7806644	0.27361086	6794625
10DPI_63I1	4173034	12106898	6483142	0.274939704	5548509
10DPI_63I2	4390120	13907218	6517396	0.254613868	5765336

Appendix V. Mouse Th1 Helper Cell Time-Course Differentiation Experiments

As indicated in the figure, at different time points (0 hours, 24 hours, 48 hours, 60 hours, and 72 hours) T-bet and Foxp3 protein expression were determined in Th1-like cells by FACS. Data are representative of at least two independent experiments.



Appendix VI. Sequencing Statistics of Mouse RNA-Seq and DNase-Seq

Sample ID	Library ID	Library Type	Sample description	raw_reads	mapped reads	mappability
KZ1201	GB0117	RNA-Seq	GB0117-taonic exp5 naïve Smart-Seq RNA-Seq	19905344	16228330	0.82
KZ1201	GB0118	RNA-Seq	GB0118-taonic exp5 0+24 Smart-Seq RNA-Seq	20263283	16437058	0.81
KZ1201	GB0119	RNA-Seq	GB0119-taonic exp5 24+0 Smart-Seq RNA-Seq	17708991	14456224	0.82
KZ1201	GB0120	RNA-Seq	GB0120-taonic exp5 24+24 Smart-Seq RNA-Seq	16830999	13592600	0.81
KZ1201	GB0121	RNA-Seq	GB0121-taonic exp5 48+0 Smart-Seq RNA-Seq	18700782	15174214	0.81
KZ1201	GB0122	RNA-Seq	GB0122-taonic exp5 48+24 Smart-Seq RNA-Seq	17163249	13601256	0.79
KZ1201	GB0123	RNA-Seq	GB0123-taonic exp5 60+0 Smart-Seq RNA-Seq	19294604	15490008	0.80
KZ1201	GB0124	RNA-Seq	GB0124-taonic exp5 60+24 Smart-Seq RNA-Seq	21461138	17554265	0.82
KZ1201	GB0125	RNA-Seq	GB0125-taonic exp5 72+0 Smart-Seq RNA-Seq	17226044	13992081	0.81
KZ1201	GB0126	RNA-Seq	GB0126-taonic exp5 72+24 Smart-Seq RNA-Seq	17921527	14444931	0.81

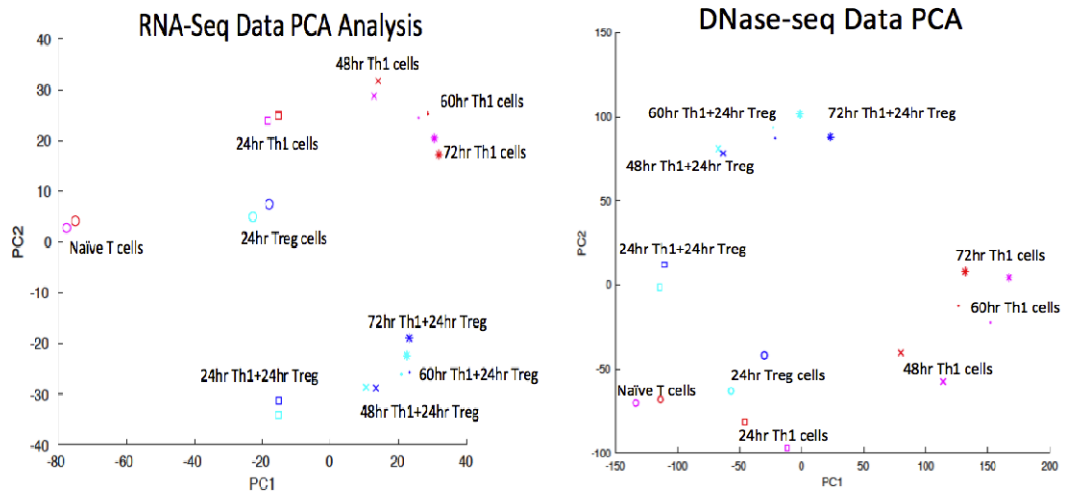
Sample ID	Library ID	Library Type	Sample description	raw_reads	mapped reads	mappability
KZ1165	GA9961	RNA-Seq	GA9961-taonic exp6 naïve CD4 RNA-Seq	18739810	15148423	0.81
KZ1165	GA9962	RNA-Seq	GA9962-taonic exp6 0+24 RNA-Seq	20900330	16817794	0.80
KZ1165	GA9963	RNA-Seq	GA9963-taonic exp6 24+0 RNA-Seq	18540827	14812208	0.80
KZ1165	GA9964	RNA-Seq	GA9964-taonic exp6 24+24 RNA-Seq	21544030	17294954	0.80
KZ1165	GA9965	RNA-Seq	GA9965-taonic exp6 48+0 RNA-Seq	20680830	16411367	0.79
KZ1165	GA9966	RNA-Seq	GA9966-taonic exp6 48+24 RNA-Seq	19941798	15592751	0.78
KZ1165	GA9967	RNA-Seq	GA9967-taonic exp6 60+0 RNA-Seq	21214413	16986336	0.80
KZ1165	GA9968	RNA-Seq	GA9968-taonic exp6 60+24 RNA-Seq	22545857	18451140	0.82
KZ1165	GA9969	RNA-Seq	GA9969-taonic exp6 72+0 RNA-Seq	21913028	17669580	0.81
KZ1165	GA9970	RNA-Seq	GA9970-taonic exp6 72+24 RNA-Seq	21874902	17441012	0.80

Library ID	Sample description	raw_reads	mapped reads	unique reads	mappability	redundancy
GA9731	mouse GA9731-taonic exp5 naive CD4 DNase-Seq	24097870	19401963	18059083	0.81	0.07
GA9732	mouse GA9732-taonic exp5 0+24 DNase-Seq	23641615	19174854	16084635	0.81	0.16
GA9733	mouse GA9733-taonic exp5 24+0 DNase-Seq	17075686	10640597	4642394	0.62	0.56
GA9734	mouse GA9734-taonic exp5 24+24 DNase-Seq	27816978	22746076	20840012	0.82	0.08
GA9735	mouse GA9735-taonic exp5 48+0 DNase-Seq	22590786	18006276	15550197	0.80	0.14
GA9736	mouse GA9736-taonic exp5 48+24 DNase-Seq	18781725	12978737	10096899	0.69	0.22
GA9737	mouse GA9737-taonic exp5 60+0 DNase-Seq	22121017	18529759	17101229	0.84	0.08
GA9738	mouse GA9738-taonic exp5 60+24 DNase-Seq	25705767	21746840	20174476	0.85	0.07
GA9739	mouse GA9739-taonic exp5 72+0 DNase-Seq	17291996	13208117	10549679	0.76	0.20
GA9740	mouse GA9740-taonic exp5 72+24 DNase-Seq	24105642	20117665	18170338	0.83	0.10

Sample description	raw_reads	mapped reads	unique reads	mappability	redundancy
Mouse naive CD4 DNase-Seq repeat	30859406	21242206	16557071	0.69	0.22
Mouse 0+24 DNase-Seq repeat	33946329	25278418	20300059	0.74	0.20
Mouse 24+0 DNase-Seq repeat	31185836	23438691	19115940	0.75	0.18
Mouse 24+24 DNase-Seq repeat	32621457	23627938	18569349	0.72	0.21
Mouse 48+0 DNase-Seq repeat	29981711	22312179	18075464	0.74	0.19
Mouse 48+24 DNase-Seq repeat	42254280	33458477	24593756	0.79	0.26
Mouse 60+0 DNase-Seq repeat	34348762	27027985	22029523	0.79	0.18
Mouse 60+24 DNase-Seq repeat	38725459	29345215	23073540	0.76	0.21
Mouse 72+0 DNase-Seq repeat	29926105	24462372	20092765	0.82	0.18
Mouse 72+24 DNase-Seq repeat	44882148	31444900	23794885	0.70	0.24

Appendix VII. PCA of Mouse DNase-Seq and RNA-Seq Data

The principal component analysis revealed the dynamic changes occurred during Th1 helper cell differentiation and under the secondary cytokine stimulation in Treg culture conditions.



Bibliography

1. Nazerian K: **Marek's disease: a neoplastic disease of chickens caused by a herpesvirus.** *Advances in cancer research* 1973, **17**:279-315.
2. Bacon L, Witter R, Silva R: **Characterization and experimental reproduction of peripheral neuropathy in White Leghorn chickens.** *Avian Pathology* 2001, **30**(5):487-499.
3. Gimeno I, Witter R, Reed W: **Four distinct neurologic syndromes in Marek's disease: effect of viral strain and pathotype.** *Avian diseases* 1999:721-737.
4. Bacon L, Hunt H, Cheng H: **A review of the development of chicken lines to resolve genes determining resistance to diseases.** *Poultry Science* 2000, **79**(8):1082-1093.
5. Biggs P: **The history and biology of Marek's disease virus.** In: *Marek's Disease*. Springer; 2001: 1-24.
6. Witter R, Calnek B, Buscaglia C, Gimeno I, Schat K: **Classification of Marek's disease viruses according to pathotype: philosophy and methodology.** *Avian pathology* 2005, **34**(2):75-90.
7. Osterrieder N, Kamil JP, Schumacher D, Tischler BK, Trapp S: **Marek's disease virus: from miasma to model.** *Nature Reviews Microbiology* 2006, **4**(4):283-294.
8. Calnek B: **Pathogenesis of Marek's disease virus infection.** In: *Marek's Disease*. Springer; 2001: 25-55.
9. Barrow AD, Burgess SC, Baigent SJ, Howes K, Nair VK: **Infection of macrophages by a lymphotropic herpesvirus: a new tropism for Marek's disease virus.** *Journal of General Virology* 2003, **84**(10):2635-2645.
10. Shek W, Calnek B, Schat K, Chen C: **Characterization of Marek's disease virus-infected lymphocytes: discrimination between cytolytically and latently infected cells.** *Journal of the National Cancer Institute* 1983, **70**(3):485-491.
11. Baigent SJ, Davison T: **Development and composition of lymphoid lesions in the spleens of Marek's disease virus-infected chickens: association with virus spread and the pathogenesis of Marek's disease.** *Avian Pathology* 1999, **28**(3):287-300.
12. Baigent SJ, Ross L, Davison T: **Differential susceptibility to Marek's disease is associated with differences in number, but not phenotype or location, of pp38+ lymphocytes.** *Journal of general virology* 1998, **79**(11):2795-2802.
13. Carrozza J, Fredrickson T, Prince R, Luginbuhl R: **Role of desquamated epithelial cells in transmission of Marek's disease.** *Avian diseases* 1973:767-781.
14. Niikura M, Witter R, Jang H, Ono M, Mikami T, Silva R: **MDV glycoprotein D is expressed in the feather follicle epithelium of infected chickens.** *Acta virologica* 1999, **43**(2-3):159-163.

15. Johnson EA, Burke C, Fredrickson T, DiCapua R: **Morphogenesis of Marek's Disease Virus in Feather Follicle Epithelium 2.** *Journal of the National Cancer Institute* 1975, **55**(1):89-99.
16. Hunt H, Lupiani B, Miller M, Gimeno I, Lee L, Parcells M: **Marek's disease virus down-regulates surface expression of MHC (B Complex) Class I (BF) glycoproteins during active but not latent infection of chicken cells.** *Virology* 2001, **282**(1):198-205.
17. Tulman E, Afonso C, Lu Z, Zsak L, Rock D, Kutish G: **The genome of a very virulent Marek's disease virus.** *Journal of Virology* 2000, **74**(17):7980-7988.
18. Lee LF, Wu P, Sui D, Ren D, Kamil J, Kung HJ, Witter RL: **The complete unique long sequence and the overall genomic organization of the GA strain of Marek's disease virus.** *Proceedings of the National Academy of Sciences* 2000, **97**(11):6091-6096.
19. Peng Q, Shirazi Y: **Characterization of the Protein Product Encoded by a Splicing Variant of the Marek's Disease VirusEco-Q Gene (Meq).** *Virology* 1996, **226**(1):77-82.
20. Parcells MS, Lin S-F, Dienglewicz RL, Majerciak V, Robinson DR, Chen H-C, Wu Z, Dubyak GR, Brunovskis P, Hunt HD: **Marek's disease virus (MDV) encodes an interleukin-8 homolog (vIL-8): characterization of the vIL-8 protein and a vIL-8 deletion mutant MDV.** *Journal of virology* 2001, **75**(11):5159-5173.
21. Cui X, Lee LF, Reed WM, Kung H-J, Reddy SM: **Marek's disease virus-encoded vIL-8 gene is involved in early cytolytic infection but dispensable for establishment of latency.** *Journal of virology* 2004, **78**(9):4753-4760.
22. Kamil JP, Tischer BK, Trapp S, Nair VK, Osterrieder N, Kung H-J: **vLIP, a viral lipase homologue, is a virulence factor of Marek's disease virus.** *Journal of virology* 2005, **79**(11):6984-6996.
23. Schat KA, Chen C-LH, Shek WR, Calnek BW: **Surface Antigens on Marek's Disease Lymphoblastoid Tumor Cell Lines 2 3.** *Journal of the National Cancer Institute* 1982, **69**(3):715-720.
24. Schat KA, Chen C, Calnek BW, Char D: **Transformation of T-lymphocyte subsets by Marek's disease herpesvirus.** *Journal of virology* 1991, **65**(3):1408-1413.
25. MORIMURA T, OHASHI K, SUGIMOTO C, ONUMA M: **Pathogenesis of Marek's disease (MD) and possible mechanisms of immunity induced by MD vaccine.** *Journal of veterinary medical science* 1998, **60**(1):1-8.
26. Bacon L, Hunt H, Cheng H: **Genetic resistance to Marek's disease.** In: *Marek's Disease*. Springer; 2001: 121-141.
27. Schat KA, Xing Z: **Specific and nonspecific immune responses to Marek's disease virus.** *Developmental & Comparative Immunology* 2000, **24**(2):201-221.
28. Pappenheimer AM, Dunn LC, Seidlin S: **Studies on fowl paralysis (neurolymphomatosis gallinarum).** *Journal of Experimental Medicine* 1929, **49**(1):87-102.

29. Delecluse H-J, Hammerschmidt W: **Status of Marek's disease virus in established lymphoma cell lines: herpesvirus integration is common.** *Journal of virology* 1993, **67**(1):82-92.
30. Delecluse H-J, Schüller S, Hammerschmidt W: **Latent Marek's disease virus can be activated from its chromosomally integrated state in herpesvirus-transformed lymphoma cells.** *The EMBO journal* 1993, **12**(8):3277.
31. Cebrian J, Kaschka-Dierich C, Berthelot N, Sheldrick P: **Inverted repeat nucleotide sequences in the genomes of Marek disease virus and the herpesvirus of the turkey.** *Proceedings of the National Academy of Sciences* 1982, **79**(2):555-558.
32. Kingham BF, Zelnik V, Kopáček J, Majerčíak V, Ney E, Schmidt CJ: **The genome of herpesvirus of turkeys: comparative analysis with Marek's disease viruses.** *Journal of General Virology* 2001, **82**(5):1123-1135.
33. Liu J-L, Ye Y, Qian Z, Qian Y, Templeton DJ, Lee LF, Kung H-J: **Functional interactions between herpesvirus oncoprotein MEQ and cell cycle regulator CDK2.** *Journal of virology* 1999, **73**(5):4208-4219.
34. Suchodolski PF, Izumiya Y, Lupiani B, Ajithdoss DK, Lee LF, Kung H-J, Reddy SM: **Both homo and heterodimers of Marek's disease virus encoded Meq protein contribute to transformation of lymphocytes in chickens.** *Virology* 2010, **399**(2):312-321.
35. Levy AM, Gilad O, Xia L, Izumiya Y, Choi J, Tsalenko A, Yakhini Z, Witter R, Lee L, Cardona CJ: **Marek's disease virus Meq transforms chicken cells via the v-Jun transcriptional cascade: a converging transforming pathway for avian oncoviruses.** *Proceedings of the National Academy of Sciences of the United States of America* 2005, **102**(41):14831-14836.
36. Huguier S, Baguet J, Perez S, Van Dam H, Castellazzi M: **Transcription factor ATF2 cooperates with v-Jun to promote growth factor-independent proliferation in vitro and tumor formation in vivo.** *Molecular and cellular biology* 1998, **18**(12):7020-7029.
37. Levy AM, Izumiya Y, Brunovskis P, Xia L, Parcells MS, Reddy SM, Lee L, Chen H-W, Kung H-J: **Characterization of the chromosomal binding sites and dimerization partners of the viral oncoprotein Meq in Marek's disease virus-transformed T cells.** *Journal of virology* 2003, **77**(23):12841-12851.
38. Schat K, Markowski-Grimsrud C: **Immune responses to Marek's disease virus infection.** In: *Marek's Disease*. Springer; 2001: 91-120.
39. Yu Y, Luo J, Mitra A, Chang S, Tian F, Zhang H, Yuan P, Zhou H, Song J: **Temporal transcriptome changes induced by MDV in marek's disease-resistant and-susceptible inbred chickens.** *BMC genomics* 2011, **12**(1):501.
40. Xing Z, Schat KA: **Inhibitory effects of nitric oxide and gamma interferon on in vitro and in vivo replication of Marek's disease virus.** *Journal of virology* 2000, **74**(8):3605-3612.
41. Xing Z, Schat K: **Expression of cytokine genes in Marek's disease virus - infected chickens and chicken embryo fibroblast cultures.** *Immunology* 2000, **100**(1):70-76.

42. Hadden J: **Recent advances in the preclinical and clinical immunopharmacology of interleukin-2: emphasis on IL-2 as an immunorestorative agent.** *Cancer detection and prevention* 1987, **12**(1-6):537-552.
43. Tani K, Su SB, Utsunomiya I, Oppenheim JJ, Wang JM: **Interferon - γ maintains the binding and functional capacity of receptors for IL - 8 on cultured human T cells.** *European journal of immunology* 1998, **28**(2):502-507.
44. Gesser B, Lund M, Lohse N, Vestergaard C, Matsushima K, Sindet-Pedersen S, Jensen SL, Thestrup-Pedersen K, Larsen CGnj: **IL-8 induces T cell chemotaxis, suppresses IL-4, and up-regulates IL-8 production by CD4+ T cells.** *Journal of leukocyte biology* 1996, **59**(3):407-411.
45. Liu J, Lin S, Xia L, Brunovskis P, Li D, Davidson I, Lee L, Kung H: **MEQ and V-IL8: cellular genes in disguise?** *Acta virologica* 1998, **43**(2-3):94-101.
46. Heidari M, Zhang HM, Sharif S: **Marek's disease virus induces Th-2 activity during cytolytic infection.** *Viral immunology* 2008, **21**(2):203-214.
47. Kodama H, Mikami T, Inoue M, Izawa H: **Inhibitory Effects of Macrophages Against Marek's Disease Virus Plaque Formation in Chicken Kidney Cell Cultures 2.** *Journal of the National Cancer Institute* 1979, **63**(5):1267-1271.
48. Haffer K, Sevoian M, Wilder M: **The role of the macrophage in Marek's disease: in vitro and in vivo studies.** *International journal of Cancer* 1979, **23**(5):648-656.
49. Abd-Ellatieff HA, Abou-Rawash A-RA, Ellakany HF, Yanai T: **Th1 and Th2 cytokines activity during transformation and lymphoma formation stage in chicken naturally infected with Marek's disease virus.** *IOSR Journal of Agriculture and Veterinary Science* 2017, **10**:72-80.
50. Ikuta K, Ueda S, Kato S, Hirai K: **Identification with monoclonal antibodies of glycoproteins of Marek's disease virus and herpesvirus of turkeys related to virus neutralization.** *Journal of virology* 1984, **49**(3):1014-1017.
51. Davidson I, Becker Y, Malkinson M: **Monospecific antibodies to Marek's disease virus antigen B dimer (200 kDa) and monomer (130 and 60 kDa) glycoproteins neutralize virus infectivity and detect the antigen B proteins in infected cell membranes.** *Archives of virology* 1991, **121**(1-4):125-139.
52. Pereira L: **Function of glycoprotein B homologues of the family herpesviridae.** *Infectious agents and disease* 1994, **3**(1):9-28.
53. Burgoyne GH, Witter R: **Effect of passively transferred immunoglobulins on Marek's disease.** *Avian diseases* 1973:824-837.
54. Chubb R, Churchill A: **Effect of maternal antibody on Marek9s disease.** *Veterinary Record* 1969, **85**(11):303-305.
55. Payne L, Rennie M: **Pathogenesis of Marek's disease in chicks with and without maternal antibody.** *Journal of the National Cancer Institute* 1973, **51**(5):1559-1573.

56. Boodhoo N, Gurung A, Sharif S, Behboudi S: **Marek's disease in chickens: a review with focus on immunology.** *Veterinary research* 2016, **47**(1):119.
57. Morimura T, Ohashi K, Kon Y, Hattori M, Sugimoto C, Onuma M: **Apoptosis and CD8-down-regulation in the thymus of chickens infected with Marek's disease virus.** *Archives of virology* 1996, **141**(11):2243-2249.
58. Andersen KG, Nissen JK, Betz AG: **Comparative genomics reveals key gain-of-function events in Foxp3 during regulatory T cell evolution.** *Frontiers in immunology* 2012, **3**.
59. Gurung A, Kamble N, Kaufer BB, Pathan A, Behboudi S: **Association of Marek's Disease induced immunosuppression with activation of a novel regulatory T cells in chickens.** *PLoS pathogens* 2017, **13**(12):e1006745.
60. Von Krosigk C, McClary C, Vielitz E, Zander D: **Selection for resistance to Marek's disease and its expected effects on other important traits in White Leghorn strain crosses.** *Avian diseases* 1972:11-19.
61. Gavora J, Grunder A, Spencer J, Gowe R, Robertson A, Speckmann G: **An assessment of effects of vaccination on genetic resistance to Marek's disease.** *Poultry science* 1974, **53**(3):889-897.
62. Hansen M: **Differences in susceptibility to Marek's disease in chickens carrying two different B locus blood group alleles.** *Poultry Science* 1967, **46**:1268-1268.
63. KAUFMAN J, VÖLK H, WALLNY HJ: **A "minimal essential Mhc" and an "unrecognized Mhc": two extremes in selection for polymorphism.** *Immunological reviews* 1995, **143**(1):63-88.
64. Pazderka F, Longenecker BM, Law GR, Stone HA, Ruth RF: **Histocompatibility of chicken populations selected for resistance to Marek's disease.** *Immunogenetics* 1975, **2**(1):93-100.
65. Jarosinski KW, Hunt HD, Osterrieder N: **Down-regulation of MHC class I by the Marek's disease virus (MDV) UL49. 5 gene product mildly affects virulence in a haplotype-specific fashion.** *Virology* 2010, **405**(2):457-463.
66. Omar A, Schat K: **Characterization of Marek's disease herpesvirus - specific cytotoxic T lymphocytes in chickens inoculated with a non - oncogenic vaccine strain of MDV.** *Immunology* 1997, **90**(4):579-585.
67. Bumstead N, Sillibourne J, Rennie M, Ross N, Davison F: **Quantification of Marek's disease virus in chicken lymphocytes using the polymerase chain reaction with fluorescence detection.** *Journal of virological methods* 1997, **65**(1):75-81.
68. Powell P, Mustill BM, Rennie M, Lee LF: **The mechanism of genetic resistance to Marek's disease in chickens.** *International journal of cancer* 1982, **29**(2):169-174.
69. Fredericksen TL, Longenecker BM, Pazderka F, Gilmour DG, Ruth RF: **A T-cell antigen system of chickens: Ly-4 and Marek's disease.** *Immunogenetics* 1977, **5**(1):535-552.
70. Gilmour DG, Brand A, Donnelly N, Stone HA: **Bu-1 and Th-1, two loci determining surface antigens of B or T lymphocytes in the chicken.** *Immunogenetics* 1976, **3**(1):549-563.

71. Bumstead N: **Genomic mapping of resistance to Marek's disease.** *Avian Pathology* 1998, **27**(S1):S78-S81.
72. Dong K, Chang S, Xie Q, Black-Pyrkosz A, Zhang H: **Comparative transcriptomics of genetically divergent lines of chickens in response to Marek's disease virus challenge at cytolytic phase.** *PLoS One* 2017, **12**(6):e0178923.
73. Schmid M, Smith J, Burt DW, Aken BL, Antin PB, Archibald AL, Ashwell C, Blackshear PJ, Boschiero C, Cheng HH: **Third Report on Chicken Genes and Chromosomes 2015: Transcriptome variation in response to Marek's disease virus acute infection.** *Cytogenetics* 2015, **145**:154-163.
74. Heidari M, Huebner M, Kireev D, Silva RF: **Transcriptional profiling of Marek's disease virus genes during cytolytic and latent infection.** *Virus Genes* 2008, **36**(2):383-392.
75. Heidari M, Sarson AJ, Huebner M, Sharif S, Kireev D, Zhou H: **Marek's disease virus-induced immunosuppression: array analysis of chicken immune response gene expression profiling.** *Viral immunology* 2010, **23**(3):309-319.
76. Chen C, Li H, Xie Q, Shang H, Ji J, Bai S, Cao Y, Ma Y, Bi Y: **Transcriptional profiling of host gene expression in chicken liver tissues infected with oncogenic Marek's disease virus.** *Journal of general virology* 2011, **92**(12):2724-2733.
77. Kano R, Konnai S, Onuma M, Ohashi K: **Microarray analysis of host immune responses to Marek's disease virus infection in vaccinated chickens.** *Journal of Veterinary Medical Science* 2009, **71**(5):603-610.
78. Sarson A, Parvizi P, Lepp D, Quinton M, Sharif S: **Transcriptional analysis of host responses to Marek's disease virus infection in genetically resistant and susceptible chickens.** *Animal genetics* 2008, **39**(3):232-240.
79. Sarson A, Abdul-Careem M, Zhou H, Sharif S: **Transcriptional analysis of host responses to Marek's disease viral infection.** *Viral immunology* 2006, **19**(4):747-758.
80. Smith J, Sadeyen J-R, Paton IR, Hocking PM, Salmon N, Fife M, Nair V, Burt DW, Kaiser P: **Systems analysis of immune responses in Marek's disease virus-infected chickens identifies a gene involved in susceptibility and highlights a possible novel pathogenicity mechanism.** *Journal of virology* 2011, **85**(21):11146-11158.
81. Kumar S, Kunec D, Buza JJ, Chiang H-I, Zhou H, Subramaniam S, Pendarvis K, Cheng HH, Burgess SC: **Nuclear Factor kappa B is central to Marek's Disease herpesvirus induced neoplastic transformation of CD30 expressing lymphocytes in-vivo.** *BMC systems biology* 2012, **6**(1):123.
82. Han B, He Y, Zhang L, Ding Y, Lian L, Zhao C, Song J, Yang N: **Long intergenic non-coding RNA GALMD3 in chicken Marek's disease.** *Scientific Reports* 2017, **7**.
83. Lujambio A, Calin GA, Villanueva A, Roperio S, Sánchez-Céspedes M, Blanco D, Montuenga LM, Rossi S, Nicoloso MS, Faller WJ: **A microRNA DNA methylation signature for human cancer metastasis.** *Proceedings of the National Academy of Sciences* 2008, **105**(36):13556-13561.

84. Lujambio A, Esteller M: **How epigenetics can explain human metastasis: a new role for microRNAs.** *Cell cycle* 2009, **8**(3):377-382.
85. Luo J, Yu Y, Chang S, Tian F, Zhang H, Song J: **DNA methylation fluctuation induced by virus infection differs between MD-resistant and-susceptible chickens.** *Frontiers in genetics* 2012, **3**.
86. Zhu J, Yamane H, Paul WE: **Differentiation of effector CD4 T cell populations.** *Annual review of immunology* 2009, **28**:445-489.
87. Russ BE, Prier JE, Rao S, Turner SJ: **T cell immunity as a tool for studying epigenetic regulation of cellular differentiation.** 2013.
88. Le Gros G, Ben-Sasson SZ, Seder R, Finkelman F, Paul W: **Generation of interleukin 4 (IL-4)-producing cells in vivo and in vitro: IL-2 and IL-4 are required for in vitro generation of IL-4-producing cells.** *Journal of Experimental Medicine* 1990, **172**(3):921-929.
89. Hsieh C-S, Heimberger AB, Gold JS, O'Garra A, Murphy KM: **Differential regulation of T helper phenotype development by interleukins 4 and 10 in an alpha beta T-cell-receptor transgenic system.** *Proceedings of the National Academy of Sciences* 1992, **89**(13):6065-6069.
90. Hsieh C-S, Macatonia SE, Tripp CS, Wolf SF, O'Garra A, Murphy KM: **Development of TH1 CD4+ T Cells Through IL-12.** *Science* 1993, **260**:547.
91. Park H, Li Z, Yang XO, Chang SH, Nurieva R, Wang Y-H, Wang Y, Hood L, Zhu Z, Tian Q: **A distinct lineage of CD4 T cells regulates tissue inflammation by producing interleukin 17.** *Nature immunology* 2005, **6**(11):1133.
92. Harrington LE, Hatton RD, Mangan PR, Turner H, Murphy TL, Murphy KM, Weaver CT: **Interleukin 17-producing CD4+ effector T cells develop via a lineage distinct from the T helper type 1 and 2 lineages.** *Nature immunology* 2005, **6**(11):1123.
93. Wilson NJ, Boniface K, Chan JR, McKenzie BS, Blumenschein WM, Mattson JD, Basham B, Smith K, Chen T, Morel F: **Development, cytokine profile and function of human interleukin 17-producing helper T cells.** *Nature immunology* 2007, **8**(9):950.
94. Chen D: **Diversification of T-helper-cell lineages: finding the family root of IL-17-producing cells.** *Nature reviews Immunology* 2006, **6**(4):329.
95. Fu S, Zhang N, Yopp AC, Chen D, Mao M, Chen D, Zhang H, Ding Y, Bromberg JS: **TGF - β Induces Foxp3+ T - Regulatory Cells from CD4+ CD25- Precursors.** *American Journal of Transplantation* 2004, **4**(10):1614-1627.
96. Fantini MC, Becker C, Monteleone G, Pallone F, Galle PR, Neurath MF: **Cutting edge: TGF- β induces a regulatory phenotype in CD4+ CD25- T cells through Foxp3 induction and down-regulation of Smad7.** *The Journal of Immunology* 2004, **172**(9):5149-5153.
97. Walker MR, Kasprovicz DJ, Gersuk VH, Bènard A, Van Landeghen M, Buckner JH, Ziegler SF: **Induction of FoxP3 and acquisition of T regulatory activity by stimulated human CD4+ CD25-T cells.** *The Journal of clinical investigation* 2003, **112**(9):1437-1443.

98. King IL, Mohrs M: **IL-4-producing CD4+ T cells in reactive lymph nodes during helminth infection are T follicular helper cells.** *Journal of Experimental Medicine* 2009, **206**(5):1001-1007.
99. Piccinni M-P, Giudizi M-G, Biagiotti R, Beloni L, Giannarini L, Sampognaro S, Parronchi P, Manetti R, Annunziato F, Livi C: **Progesterone favors the development of human T helper cells producing Th2-type cytokines and promotes both IL-4 production and membrane CD30 expression in established Th1 cell clones.** *The Journal of Immunology* 1995, **155**(1):128-133.
100. Takeda K, Tanaka T, Shi W, Matsumoto M: **Essential role of Stat6 in IL-4 signalling.** *Nature* 1996, **380**(6575):627.
101. Lighvani AA, Frucht DM, Jankovic D, Yamane H, Aliberti J, Hissong BD, Nguyen BV, Gadina M, Sher A, Paul WE: **T-bet is rapidly induced by interferon- γ in lymphoid and myeloid cells.** *Proceedings of the National Academy of Sciences* 2001, **98**(26):15137-15142.
102. Yang L, Anderson DE, Baecher-Allan C, Hastings WD, Bettelli E, Oukka M, Kuchroo VK, Hafler DA: **IL-21 and TGF- β are required for differentiation of human TH17 cells.** *Nature* 2008, **454**(7202):350-352.
103. Mangan PR, Harrington LE, O'quinn DB, Helms WS, Bullard DC, Elson CO, Hatton RD, Wahl SM, Schoeb TR, Weaver CT: **Transforming growth factor- β induces development of the TH17 lineage.** *Nature* 2006, **441**(7090):231-234.
104. Bettelli E, Korn T, Oukka M, Kuchroo VK: **Induction and effector functions of TH17 cells.** *Nature* 2008, **453**(7198):1051-1057.
105. Bettelli E, Korn T, Kuchroo VK: **Th17: the third member of the effector T cell trilogy.** *Current opinion in immunology* 2007, **19**(6):652-657.
106. Korn T, Bettelli E, Gao W, Awasthi A, Jäger A, Strom TB, Oukka M, Kuchroo VK: **IL-21 initiates an alternative pathway to induce proinflammatory TH17 cells.** *Nature* 2007, **448**(7152):484-487.
107. Mitchell TJ, John S: **Signal transducer and activator of transcription (STAT) signalling and T - cell lymphomas.** *Immunology* 2005, **114**(3):301-312.
108. Darnell Jr JE, Kerr IM, Stark GR: **Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins.** *Science-AAAS-weekly paper edition-including guide to scientific information* 1994, **264**(5164):1415-1420.
109. Gadina M, Hilton D, Johnston JA, Morinobu A, Lighvani A, Zhou Y-J, Visconti R, O'Shea JJ: **Signaling by type I and II cytokine receptors: ten years after.** *Current opinion in immunology* 2001, **13**(3):363-373.
110. Levy D: **Physiological significance of STAT proteins: investigations through gene disruption in vivo.** *Cellular and molecular life sciences* 1999, **55**(12):1559-1567.
111. Szabo SJ, Kim ST, Costa GL, Zhang X, Fathman CG, Glimcher LH: **A novel transcription factor, T-bet, directs Th1 lineage commitment.** *Cell* 2000, **100**(6):655-669.

112. Mullen AC, High FA, Hutchins AS, Lee HW, Villarino AV, Livingston DM, Kung AL, Cereb N, Yao T-P, Yang SY: **Role of T-bet in commitment of TH1 cells before IL-12-dependent selection.** *Science* 2001, **292**(5523):1907-1910.
113. Mullen AC, Hutchins AS, High FA, Lee HW, Sykes KJ, Chodosh LA, Reiner SL: **Hlx is induced by and genetically interacts with T-bet to promote heritable TH1 gene induction.** *Nature immunology* 2002, **3**(7):652-658.
114. Koch MA, Perdue NR, Killebrew JR, Urdahl KB, Campbell DJ: **The transcription factor T-bet controls regulatory T cell homeostasis and function during type 1 inflammation.** *Nature immunology* 2009, **10**(6):595-602.
115. Gocke AR, Cravens PD, Ben L-H, Hussain RZ, Northrop SC, Racke MK, Lovett-Racke AE: **T-bet regulates the fate of Th1 and Th17 lymphocytes in autoimmunity.** *The Journal of Immunology* 2007, **178**(3):1341-1348.
116. Fitch E, Harper E, Skorcheva I, Kurtz SE, Blauvelt A: **Pathophysiology of psoriasis: recent advances on IL-23 and Th17 cytokines.** *Current rheumatology reports* 2007, **9**(6):461-467.
117. McKenzie BS, Kastelein RA, Cua DJ: **Understanding the IL-23–IL-17 immune pathway.** *Trends in immunology* 2006, **27**(1):17-23.
118. Szabo SJ, Sullivan BM, Stemann C, Satoskar AR, Sleckman BP, Glimcher LH: **Distinct effects of T-bet in TH1 lineage commitment and IFN- γ production in CD4 and CD8 T cells.** *Science* 2002, **295**(5553):338-342.
119. Pearce EL, Mullen AC, Martins GA, Krawczyk CM, Hutchins AS, Zediak VP, Banica M, DiCioccio CB, Gross DA, Mao C-a: **Control of effector CD8+ T cell function by the transcription factor Eomesodermin.** *Science* 2003, **302**(5647):1041-1043.
120. Banerjee A, Gordon SM, Intlekofer AM, Paley MA, Mooney EC, Lindsten T, Wherry EJ, Reiner SL: **Cutting edge: The transcription factor eomesodermin enables CD8+ T cells to compete for the memory cell niche.** *The Journal of Immunology* 2010, **185**(9):4988-4992.
121. Takemoto N, Intlekofer AM, Northrup JT, Wherry EJ, Reiner SL: **Cutting Edge: IL-12 inversely regulates T-bet and eomesodermin expression during pathogen-induced CD8+ T cell differentiation.** *The Journal of Immunology* 2006, **177**(11):7515-7519.
122. Kaplan MH: **A Critical Regulator of Inflammation In Vivo.** *Immunologic Research* 2005, **32**(1):231-241.
123. Wurster AL, Tanaka T, Grusby MJ: **The biology of Stat4 and Stat6.** *Oncogene* 2000, **19**(21):2577.
124. Usui T, Nishikomori R, Kitani A, Strober W: **GATA-3 suppresses Th1 development by downregulation of Stat4 and not through effects on IL-12R β 2 chain or T-bet.** *Immunity* 2003, **18**(3):415-428.
125. Cote-Sierra J, Foucras G, Guo L, Chiodetti L, Young HA, Hu-Li J, Zhu J, Paul WE: **Interleukin 2 plays a central role in Th2 differentiation.** *Proceedings of the National Academy of Sciences of the United States of America* 2004, **101**(11):3880-3885.

126. Shimoda K, Van Deursen J, Sangster MY, Sarawar SR: **Lack of IL-4-induced Th2 response and IgE class switching in mice with disrupted Stat6 gene.** *Nature* 1996, **380**(6575):630.
127. Kurata H, Lee HJ, O'Garra A, Arai N: **Ectopic expression of activated Stat6 induces the expression of Th2-specific cytokines and transcription factors in developing Th1 cells.** *Immunity* 1999, **11**(6):677-688.
128. Lee HJ, Takemoto N, Kurata H, Kamogawa Y, Miyatake S, O'garra A, Arai N: **GATA-3 induces T helper cell type 2 (Th2) cytokine expression and chromatin remodeling in committed Th1 cells.** *Journal of Experimental Medicine* 2000, **192**(1):105-116.
129. Murphy KM, Reiner SL: **The lineage decisions of helper T cells.** *Nature reviews Immunology* 2002, **2**(12):933.
130. Zhu J, Min B, Hu-Li J, Watson CJ, Grinberg A, Wang Q, Killeen N, Urban JF, Guo L, Paul WE: **Conditional deletion of Gata3 shows its essential function in TH1-TH2 responses.** *Nature immunology* 2004, **5**(11):1157.
131. Yagi R, Zhu J, Paul WE: **An updated view on transcription factor GATA3-mediated regulation of Th1 and Th2 cell differentiation.** *International immunology* 2011, **23**(7):415-420.
132. Mukherjee S, Rasky AJ, Lundy PA, Kittan NA, Kunkel SL, Maillard IP, Kowalski PE, Kousis PC, Guidos CJ, Lukacs NW: **STAT5-Induced Lunatic Fringe during Th2 Development Alters Delta-like 4-Mediated Th2 Cytokine Production in Respiratory Syncytial Virus-Exacerbated Airway Allergic Disease.** *The Journal of Immunology* 2014, **192**(3):996-1003.
133. Chen Z, Laurence A, Kanno Y, Pacher-Zavisin M, Zhu B-M, Tato C, Yoshimura A, Hennighausen L, O'Shea JJ: **Selective regulatory function of Socs3 in the formation of IL-17-secreting T cells.** *Proceedings of the National Academy of Sciences* 2006, **103**(21):8137-8142.
134. Zhang F, Meng G, Strober W: **Interactions among the transcription factors Runx1, ROR γ t and Foxp3 regulate the differentiation of interleukin 17-producing T cells.** *Nature immunology* 2008, **9**(11):1297-1306.
135. Zhu J, Paul WE: **Heterogeneity and plasticity of T helper cells.** *Cell research* 2010, **20**(1):4.
136. Maloy KJ, Powrie F: **Regulatory T cells in the control of immune pathology.** *Nature immunology* 2001, **2**(9):816-822.
137. Sakaguchi S: **Regulatory T cells: key controllers of immunologic self-tolerance.** *Cell* 2000, **101**(5):455-458.
138. Shevach EM: **CD4⁺ CD25⁺ suppressor T cells: more questions than answers.** *Nature Reviews Immunology* 2002, **2**(6):389-400.
139. Hori S, Nomura T, Sakaguchi S: **Control of regulatory T cell development by the transcription factor Foxp3.** *Science* 2003, **299**(5609):1057-1061.
140. Itoh M, Takahashi T, Sakaguchi N, Kuniyasu Y, Shimizu J, Otsuka F, Sakaguchi S: **Thymus and autoimmunity: production of CD25⁺ CD4⁺ naturally anergic and suppressive T cells as a key function of the thymus in maintaining immunologic self-tolerance.** *The Journal of Immunology* 1999, **162**(9):5317-5326.

141. Asano M, Toda M, Sakaguchi N, Sakaguchi S: **Autoimmune disease as a consequence of developmental abnormality of a T cell subpopulation.** *Journal of Experimental Medicine* 1996, **184**(2):387-396.
142. Ueda H, Howson JM, Esposito L, Heward J, Chamberlain G, Rainbow DB, Hunter KM, Smith AN, Di Genova G, Herr MH: **Association of the T-cell regulatory gene CTLA4 with susceptibility to autoimmune disease.** *Nature* 2003, **423**(6939):506-511.
143. Sakaguchi S, Ono M, Setoguchi R, Yagi H, Hori S, Fehervari Z, Shimizu J, Takahashi T, Nomura T: **Foxp3⁺ CD25⁺ CD4⁺ natural regulatory T cells in dominant self - tolerance and autoimmune disease.** *Immunological reviews* 2006, **212**(1):8-27.
144. Curiel TJ, Coukos G, Zou L, Alvarez X, Cheng P, Mottram P, Evdemon-Hogan M, Conejo-Garcia JR, Zhang L, Burow M: **Specific recruitment of regulatory T cells in ovarian carcinoma fosters immune privilege and predicts reduced survival.** *Nature medicine* 2004, **10**(9):942-949.
145. Ziegler SF: **FOXP3: of mice and men.** *Annu Rev Immunol* 2006, **24**:209-226.
146. Allan SE, Song - Zhao GX, Abraham T, McMurchy AN, Levings MK: **Inducible reprogramming of human T cells into Treg cells by a conditionally active form of FOXP3.** *European journal of immunology* 2008, **38**(12):3282-3289.
147. Wan YY, Flavell RA: **Regulatory T-cell functions are subverted and converted owing to attenuated Foxp3 expression.** *Nature* 2007, **445**(7129):766-770.
148. Gavin MA, Rasmussen JP, Fontenot JD, Vasta V, Manganiello VC, Beavo JA, Rudensky AY: **Foxp3-dependent programme of regulatory T-cell differentiation.** *Nature* 2007, **445**(7129):771-775.
149. Williams LM, Rudensky AY: **Maintenance of the Foxp3-dependent developmental program in mature regulatory T cells requires continued expression of Foxp3.** *Nature immunology* 2007, **8**(3):277-285.
150. Allan SE, Alstad AN, Merindol N, Crellin NK, Amendola M, Bacchetta R, Naldini L, Roncarolo MG, Soudeyns H, Levings MK: **Generation of potent and stable human CD4⁺ T regulatory cells by activation-independent expression of FOXP3.** *Molecular Therapy* 2008, **16**(1):194-202.
151. Allan SE, Crome SQ, Crellin NK, Passerini L, Steiner TS, Bacchetta R, Roncarolo MG, Levings MK: **Activation-induced FOXP3 in human T effector cells does not suppress proliferation or cytokine production.** *International immunology* 2007, **19**(4):345-354.
152. Tran DQ, Ramsey H, Shevach EM: **Induction of FOXP3 expression in naive human CD4⁺ FOXP3⁻ T cells by T-cell receptor stimulation is transforming growth factor- β -dependent but does not confer a regulatory phenotype.** *Blood* 2007, **110**(8):2983-2990.
153. Yu F, Sharma S, Edwards J, Feigenbaum L, Zhu J: **Dynamic expression of transcription factors T-bet and GATA-3 by regulatory T cells maintains immunotolerance.** *Nature immunology* 2015, **16**(2):197-206.
154. Oldenhove G, Bouladoux N, Wohlfert EA, Hall JA, Chou D, O'Brien S, Blank R, Lamb E, Natarajan S, Kastenmayer R: **Decrease of Foxp3⁺ Treg cell**

- number and acquisition of effector cell phenotype during lethal infection.** *Immunity* 2009, **31**(5):772-786.
155. Yamaguchi T, Wing JB, Sakaguchi S: **Two modes of immune suppression by Foxp3⁺ regulatory T cells under inflammatory or non-inflammatory conditions.** In: *Seminars in immunology: 2011*; Elsevier; 2011: 424-430.
 156. Wang Y, Su MA, Wan YY: **An essential role of the transcription factor GATA-3 for the function of regulatory T cells.** *Immunity* 2011, **35**(3):337-348.
 157. Rudra D, Chaudhry A, Niec RE, Arvey A, Samstein RM, Leslie C, Shaffer SA, Goodlett DR, Rudensky AY: **Transcription factor Foxp3 and its protein partners form a complex regulatory network.** *Nature immunology* 2012, **13**(10):1010-1019.
 158. Wei G, Abraham BJ, Yagi R, Jothi R, Cui K, Sharma S, Narlikar L, Northrup DL, Tang Q, Paul WE: **Genome-wide analyses of transcription factor GATA3-mediated gene regulation in distinct T cell types.** *Immunity* 2011, **35**(2):299-311.
 159. Wohlfert EA, Grainger JR, Bouladoux N, Konkel JE, Oldenhove G, Ribeiro CH, Hall JA, Yagi R, Naik S, Bhairavabhotla R: **GATA3 controls Foxp3⁺ regulatory T cell fate during inflammation in mice.** *The Journal of clinical investigation* 2011, **121**(11):4503.
 160. Fang D, Zhu J: **Dynamic balance between master transcription factors determines the fates and functions of CD4 T cell and innate lymphoid cell subsets.** *Journal of Experimental Medicine* 2017, **214**(7):1861-1876.
 161. Komatsu N, Okamoto K, Sawa S, Nakashima T, Oh-Hora M, Kodama T, Tanaka S, Bluestone JA, Takayanagi H: **Pathogenic conversion of Foxp3⁺ T cells into TH17 cells in autoimmune arthritis.** *Nature medicine* 2014, **20**(1):62-68.
 162. Yang XO, Nurieva R, Martinez GJ, Kang HS, Chung Y, Pappu BP, Shah B, Chang SH, Schluns KS, Watowich SS: **Molecular antagonism and plasticity of regulatory and inflammatory T cell programs.** *Immunity* 2008, **29**(1):44-56.
 163. Zhou L, Lopes JE, Chong MM, Ivanov II, Min R, Victora GD, Shen Y, Du J, Rubtsov YP, Rudensky AY: **TGF- β -induced Foxp3 inhibits TH17 cell differentiation by antagonizing ROR γ t function.** *Nature* 2008, **453**(7192):236-240.
 164. Sefik E, Geva-Zatorsky N, Oh S, Konnikova L, Zemmour D, McGuire AM, Burzyn D, Ortiz-Lopez A, Lobera M, Yang J: **Individual intestinal symbionts induce a distinct population of ROR γ ⁺ regulatory T cells.** *Science* 2015, **349**(6251):993-997.
 165. Ohnmacht C, Park J-H, Cording S, Wing JB, Atarashi K, Obata Y, Gaboriau-Routhiau V, Marques R, Dulauroy S, Fedoseeva M: **The microbiota regulates type 2 immunity through ROR γ t⁺ T cells.** *Science* 2015, **349**(6251):989-993.
 166. Lee YK, Turner H, Maynard CL, Oliver JR, Chen D, Elson CO, Weaver CT: **Late developmental plasticity in the T helper 17 lineage.** *Immunity* 2009, **30**(1):92-107.

167. Lexberg MH, Taubner A, Albrecht I, Lepenies I, Richter A, Kamradt T, Radbruch A, Chang HD: **IFN - γ and IL - 12 synergize to convert in vivo generated Th17 into Th1/Th17 cells.** *European journal of immunology* 2010, **40**(11):3017-3027.
168. Hirota K, Duarte JH, Veldhoen M, Hornsby E, Li Y, Cua DJ, Ahlfors H, Wilhelm C, Tolaini M, Menzel U: **Fate mapping of IL-17-producing T cells in inflammatory responses.** *Nature immunology* 2011, **12**(3):255-263.
169. Wang Y-H, Voo KS, Liu B, Chen C-Y, Uygungil B, Spoede W, Bernstein JA, Huston DP, Liu Y-J: **A novel subset of CD4+ TH2 memory/effector cells that produce inflammatory IL-17 cytokine and promote the exacerbation of chronic allergic asthma.** *Journal of Experimental Medicine* 2010, **207**(11):2479-2491.
170. Cosmi L, Maggi L, Santarlasci V, Capone M, Cardilicchia E, Frosali F, Querci V, Angeli R, Matucci A, Fambrini M: **Identification of a novel subset of human circulating memory CD4+ T cells that produce both IL-17A and IL-4.** *Journal of Allergy and Clinical Immunology* 2010, **125**(1):222-230. e224.
171. Peine M, Rausch S, Helmstetter C, Fröhlich A, Hegazy AN, Köhl AA, Grevelding CG, Höfer T, Hartmann S, Löhning M: **Stable T-bet+ GATA-3+ Th1/Th2 hybrid cells arise in vivo, can develop directly from naive precursors, and limit immunopathologic inflammation.** *PLoS biology* 2013, **11**(8):e1001633.
172. Gor DO, Rose NR, Greenspan NS: **TH1-TH2: a procrustean paradigm.** *Nature immunology* 2003, **4**(6):503-505.
173. Murphy KM: **In search of the CTD.** *Nature immunology* 2003, **4**(7):645-645.
174. Murphy E, Shibuya K, Hosken N, Openshaw P, Maino V, Davis K, Murphy K, O'garra A: **Reversibility of T helper 1 and 2 populations is lost after long-term stimulation.** *Journal of Experimental Medicine* 1996, **183**(3):901-913.
175. Bird JJ, Brown DR, Mullen AC, Moskowitz NH, Mahowald MA, Sider JR, Gajewski TF, Wang C-R, Reiner SL: **Helper T cell differentiation is controlled by the cell cycle.** *Immunity* 1998, **9**(2):229-237.
176. Zhou L, Chong MM, Littman DR: **Plasticity of CD4+ T cell lineage differentiation.** *Immunity* 2009, **30**(5):646-655.
177. Szabo SJ, Dighe AS, Gubler U, Murphy KM: **Regulation of the interleukin (IL)-12R β 2 subunit expression in developing T helper 1 (Th1) and Th2 cells.** *Journal of Experimental Medicine* 1997, **185**(5):817-824.
178. Ahmadzadeh M, Farber DL: **Functional plasticity of an antigen-specific memory CD4 T cell population.** *Proceedings of the national academy of sciences* 2002, **99**(18):11802-11807.
179. Paliard X, de Waal Malefijt R, Yssel H, Blanchard D, Chretien I, Abrams J, De Vries J, Spits H: **Simultaneous production of IL-2, IL-4, and IFN-gamma by activated human CD4+ and CD8+ T cell clones.** *The Journal of Immunology* 1988, **141**(3):849-855.
180. Sundrud MS, Grill SM, Ni D, Nagata K, Alkan SS, Subramaniam A, Unutmaz D: **Genetic reprogramming of primary human T cells reveals functional**

- plasticity in Th cell differentiation.** *The Journal of Immunology* 2003, **171**(7):3542-3549.
181. Szabo SJ, Jacobson NG, Dighe AS, Gubler U, Murphy KM: **Developmental commitment to the Th2 lineage by extinction of IL-12 signaling.** *Immunity* 1995, **2**(6):665-675.
 182. O'Garra A: **Cytokines induce the development of functionally heterogeneous T helper cell subsets.** *Immunity* 1998, **8**(3):275-283.
 183. Zhu J, Guo L, Min B, Watson CJ, Hu-Li J, Young HA, Tsichlis PN, Paul WE: **Growth factor independent-1 induced by IL-4 regulates Th2 cell proliferation.** *Immunity* 2002, **16**(5):733-744.
 184. Zhu J, Davidson TS, Wei G, Jankovic D, Cui K, Schones DE, Guo L, Zhao K, Shevach EM, Paul WE: **Down-regulation of Gfi-1 expression by TGF- β is important for differentiation of Th17 and CD103+ inducible regulatory T cells.** *Journal of Experimental Medicine* 2009, **206**(2):329-341.
 185. Xu L, Kitani A, Fuss I, Strober W: **Cutting edge: regulatory T cells induce CD4+ CD25- Foxp3- T cells or are self-induced to become Th17 cells in the absence of exogenous TGF- β .** *The Journal of Immunology* 2007, **178**(11):6725-6729.
 186. Murphy KM, Stockinger B: **Effector T cell plasticity: flexibility in the face of changing circumstances.** *Nature immunology* 2010, **11**(8):674-680.
 187. Lexberg MH, Taubner A, Förster A, Albrecht I, Richter A, Kamradt T, Radbruch A, Chang HD: **Th memory for interleukin - 17 expression is stable in vivo.** *European journal of immunology* 2008, **38**(10):2654-2664.
 188. Mathur AN, Chang H-C, Zisoulis DG, Kapur R, Belladonna ML, Kansas GS, Kaplan MH: **T-bet is a critical determinant in the instability of the IL-17-secreting T-helper phenotype.** *Blood* 2006, **108**(5):1595-1601.
 189. Bending D, De La Peña H, Veldhoen M, Phillips JM, Uyttenhove C, Stockinger B, Cooke A: **Highly purified Th17 cells from BDC2. 5NOD mice convert into Th1-like cells in NOD/SCID recipient mice.** *The Journal of clinical investigation* 2009, **119**(3):565.
 190. Komatsu N, Mariotti-Ferrandiz ME, Wang Y, Malissen B, Waldmann H, Hori S: **Heterogeneity of natural Foxp3+ T cells: a committed regulatory T-cell lineage and an uncommitted minor population retaining plasticity.** *Proceedings of the National Academy of Sciences* 2009, **106**(6):1903-1908.
 191. Rivas MN, Burton OT, Wise P, Charbonnier L-M, Georgiev P, Oettgen HC, Rachid R, Chatila TA: **Regulatory T cell reprogramming toward a Th2-cell-like lineage impairs oral tolerance and promotes food allergy.** *Immunity* 2015, **42**(3):512-523.
 192. Brüstle A, Heink S, Huber M, Rosenplänter C, Stadelmann C, Yu P, Arpaia E, Mak TW, Kamradt T, Lohoff M: **The development of inflammatory TH-17 cells requires interferon-regulatory factor 4.** *Nature immunology* 2007, **8**(9):958-966.
 193. Ono M, Yaguchi H, Ohkura N, Kitabayashi I, Nagamura Y, Nomura T, Miyachi Y, Tsukada T, Sakaguchi S: **Foxp3 controls regulatory T-cell function by interacting with AML1/Runx1.** *Nature* 2007, **446**(7136):685-689.

194. Rudra D, Egawa T, Chong MM, Treuting P, Littman DR, Rudensky AY: **Runx-CBF β complexes control expression of the transcription factor Foxp3 in regulatory T cells.** *Nature immunology* 2009, **10**(11):1170-1177.
195. Ciofani M, Madar A, Galan C, Sellars M, Mace K, Pauli F, Agarwal A, Huang W, Parkurst CN, Muratet M: **A validated regulatory network for Th17 cell specification.** *Cell* 2012, **151**(2):289-303.
196. Yosef N, Shalek AK, Gaublomme JT, Jin H, Lee Y, Awasthi A, Wu C, Karwacz K, Xiao S, Jorgolli M: **Dynamic regulatory network controlling TH17 cell differentiation.** *Nature* 2013, **496**(7446):461-468.
197. Kouzarides T: **Chromatin modifications and their function.** *Cell* 2007, **128**(4):693-705.
198. Goldberg AD, Allis CD, Bernstein E: **Epigenetics: a landscape takes shape.** *Cell* 2007, **128**(4):635-638.
199. Ansel KM, Lee DU, Rao A: **An epigenetic view of helper T cell differentiation.** *Nature immunology* 2003, **4**(7):616-623.
200. Merkenschlager M, Wilson CB: **RNAi and chromatin in T cell development and function.** *Current opinion in immunology* 2008, **20**(2):131-138.
201. Kanno Y, Vahedi G, Hirahara K, Singleton K, O'Shea JJ: **Transcriptional and epigenetic control of T helper cell specification: molecular mechanisms underlying commitment and plasticity.** *Annual review of immunology* 2012, **30**:707-731.
202. Wei G, Wei L, Zhu J, Zang C, Hu-Li J, Yao Z, Cui K, Kanno Y, Roh T-Y, Watford WT: **Global mapping of H3K4me3 and H3K27me3 reveals specificity and plasticity in lineage fate determination of differentiating CD4⁺ T cells.** *Immunity* 2009, **30**(1):155-167.
203. Nakayamada S, Kanno Y, Takahashi H, Jankovic D, Lu KT, Johnson TA, Sun H-w, Vahedi G, Hakim O, Handon R: **Early Th1 cell differentiation is marked by a Tfh cell-like transition.** *Immunity* 2011, **35**(6):919-931.
204. Krawczyk CM, Shen H, Pearce EJ: **Functional plasticity in memory T helper cell responses.** *The Journal of Immunology* 2007, **178**(7):4080-4088.
205. Wilson CB, Rowell E, Sekimata M: **Epigenetic control of T-helper-cell differentiation.** *Nature Reviews Immunology* 2009, **9**(2):91-105.
206. Zhou X, Jeker LT, Fife BT, Zhu S, Anderson MS, McManus MT, Bluestone JA: **Selective miRNA disruption in T reg cells leads to uncontrolled autoimmunity.** *Journal of Experimental Medicine* 2008, **205**(9):1983-1991.
207. Kuchen S, Resch W, Yamane A, Kuo N, Li Z, Chakraborty T, Wei L, Laurence A, Yasuda T, Peng S: **Regulation of microRNA expression and abundance during lymphopoiesis.** *Immunity* 2010, **32**(6):828-839.
208. Floess S, Freyer J, Siewert C, Baron U, Olek S, Polansky J, Schlawe K, Chang H-D, Bopp T, Schmitt E: **Epigenetic control of the foxp3 locus in regulatory T cells.** *PLoS biology* 2007, **5**(2):e38.
209. Baron U, Floess S, Wiczorek G, Baumann K, Grützkau A, Dong J, Thiel A, Boeld TJ, Hoffmann P, Edinger M: **DNA demethylation in the human FOXP3 locus discriminates regulatory T cells from activated FOXP3⁺ conventional T cells.** *European journal of immunology* 2007, **37**(9):2378-2389.

210. Kim H-P, Leonard WJ: **CREB/ATF-dependent T cell receptor–induced FoxP3 gene expression: a role for DNA methylation.** *Journal of Experimental Medicine* 2007, **204**(7):1543-1551.
211. Nagar M, Vernitsky H, Cohen Y, Dominissini D, Berkun Y, Rechavi G, Amariglio N, Goldstein I: **Epigenetic inheritance of DNA methylation limits activation-induced expression of FOXP3 in conventional human CD25– CD4+ T cells.** *International immunology* 2008, **20**(8):1041-1055.
212. Huehn J, Polansky JK, Hamann A: **Epigenetic control of FOXP3 expression: the key to a stable regulatory T-cell lineage?** *Nature Reviews Immunology* 2009, **9**(2):83-89.
213. Polansky JK, Kretschmer K, Freyer J, Floess S, Garbe A, Baron U, Olek S, Hamann A, von Boehmer H, Huehn J: **DNA methylation controls Foxp3 gene expression.** *European journal of immunology* 2008, **38**(6):1654-1663.
214. Lu L-F, Boldin MP, Chaudhry A, Lin L-L, Taganov KD, Hanada T, Yoshimura A, Baltimore D, Rudensky AY: **Function of miR-146a in controlling Treg cell-mediated regulation of Th1 responses.** *Cell* 2010, **142**(6):914-929.
215. Lu L-F, Thai T-H, Calado DP, Chaudhry A, Kubo M, Tanaka K, Loeb GB, Lee H, Yoshimura A, Rajewsky K: **Foxp3-dependent microRNA155 confers competitive fitness to regulatory T cells by targeting SOCS1 protein.** *Immunity* 2009, **30**(1):80-91.
216. Rodriguez A, Vigorito E, Clare S, Warren MV, Couttet P, Soond DR, van Dongen S, Grocock RJ, Das PP, Miska EA: **Requirement of bic/microRNA-155 for normal immune function.** *Science* 2007, **316**(5824):608-611.
217. Banerjee A, Schambach F, DeJong CS, Hammond SM, Reiner SL: **Micro - RNA - 155 inhibits IFN - γ signaling in CD4+ T cells.** *European journal of immunology* 2010, **40**(1):225-231.
218. Du C, Liu C, Kang J, Zhao G, Ye Z, Huang S, Li Z, Wu Z, Pei G: **MicroRNA miR-326 regulates TH-17 differentiation and is associated with the pathogenesis of multiple sclerosis.** *Nature immunology* 2009, **10**(12):1252-1259.
219. Ma F, Xu S, Liu X, Zhang Q, Xu X, Liu M, Hua M, Li N, Yao H, Cao X: **The microRNA miR-29 controls innate and adaptive immune responses to intracellular bacterial infection by targeting interferon-[gamma].** *Nature immunology* 2011, **12**(9):861-869.
220. Steiner DF, Thomas MF, Hu JK, Yang Z, Babiarz JE, Allen CD, Matloubian M, Blelloch R, Ansel KM: **MicroRNA-29 regulates T-box transcription factors and interferon- γ production in helper T cells.** *Immunity* 2011, **35**(2):169-181.
221. Tang B, Cheng X, Xi Y, Chen Z, Zhou Y, Jin VX: **Advances in Genomic Profiling and Analysis of 3D Chromatin Structure and Interaction.** *Genes* 2017, **8**(9):223.
222. Dekker J, Marti-Renom MA, Mirny LA: **Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data.** *Nature Reviews Genetics* 2013, **14**(6):390-403.

223. Pombo A, Dillon N: **Three-dimensional genome architecture: players and mechanisms.** *Nature reviews Molecular cell biology* 2015, **16**(4):245-257.
224. Janssen P, Goldovsky L, Kunin V, Darzentas N, Ouzounis CA: **Genome coverage, literally speaking.** *EMBO reports* 2005, **6**(5):397-399.
225. Bakke P, Carney N, DeLoache W, Gearing M, Ingvorsen K, Lotz M, McNair J, Penumetcha P, Simpson S, Voss L: **Evaluation of three automated genome annotations for *Halorhabdus utahensis*.** *PloS one* 2009, **4**(7):e6291.
226. Hall N: **Advanced sequencing technologies and their wider impact in microbiology.** *Journal of Experimental Biology* 2007, **210**(9):1518-1525.
227. Morozova O, Marra MA: **Applications of next-generation sequencing technologies in functional genomics.** *Genomics* 2008, **92**(5):255-264.
228. Reed JL, Famili I, Thiele I, Palsson BO: **Towards multidimensional genome annotation.** *Nature Reviews Genetics* 2006, **7**(2):130-141.
229. Thanbichler M, Viollier PH, Shapiro L: **The structure and function of the bacterial chromosome.** *Current opinion in genetics & development* 2005, **15**(2):153-162.
230. Covert MW, Knight EM, Reed JL, Herrgard MJ, Palsson BO: **Integrating high-throughput and computational data elucidates bacterial networks.** *Nature* 2004, **429**(6987):92-96.
231. Palsson B: **Systems biology: properties of reconstructed networks.** 2006. *Cambridge: Cambridge Univ Pr CrossRef Google Scholar*.
232. Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, Wang MD, Zhang K, Mitra RD, Church GM: **Accurate multiplex polony sequencing of an evolved bacterial genome.** *Science* 2005, **309**(5741):1728-1732.
233. Honisch C, Raghunathan A, Cantor CR, Palsson BØ, van den Boom D: **High-throughput mutation detection underlying adaptive evolution of *Escherichia coli*-K12.** *Genome research* 2004, **14**(12):2495-2502.
234. Wang Y-M, Zhou P, Wang L-Y, Li Z-H, Zhang Y-N, Zhang Y-X: **Correlation between DNase I hypersensitive site distribution and gene expression in HeLa S3 cells.** *PloS one* 2012, **7**(8):e42414.
235. Consortium EP: **An integrated encyclopedia of DNA elements in the human genome.** *Nature* 2012, **489**(7414):57-74.
236. Cooper J, Ding Y, Song J, Zhao K: **Genome-wide mapping of DNase I hypersensitive sites in rare cell populations using single-cell DNase sequencing.** *Nature protocols* 2017, **12**(11):2342.
237. Boyle AP, Davis S, Shulha HP, Meltzer P, Margulies EH, Weng Z, Furey TS, Crawford GE: **High-resolution mapping and characterization of open chromatin across the genome.** *Cell* 2008, **132**(2):311-322.
238. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ: **Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position.** *Nature methods* 2013, **10**(12):1213-1218.
239. Crawford GE, Holt IE, Whittle J, Webb BD, Tai D, Davis S, Margulies EH, Chen Y, Bernat JA, Ginsburg D: **Genome-wide mapping of DNase**

- hypersensitive sites using massively parallel signature sequencing (MPSS).** *Genome research* 2006, **16**(1):123-131.
240. Crawford GE, Davis S, Scacheri PC, Renaud G, Halawi MJ, Erdos MR, Green R, Meltzer PS, Wolfsberg TG, Collins FS: **DNase-chip: a high-resolution method to identify DNase I hypersensitive sites using tiled microarrays.** *Nature methods* 2006, **3**(7):503-509.
 241. Jin W, Tang Q, Wan M, Cui K, Zhang Y, Ren G, Ni B, Sklar J, Przytycka TM, Childs R: **Genome-wide detection of DNase I hypersensitive sites in single cells and FFPE tissue samples.** *Nature* 2015, **528**(7580):142-146.
 242. Ho MC, Quintero-Cadena P, Sternberg PW: **Genome-wide discovery of active regulatory elements and transcription factor footprints in *Caenorhabditis elegans* using DNase-seq.** *Genome research* 2017, **27**(12):2108-2119.
 243. Ampuja M, Rantapero T, Rodriguez-Martinez A, Palmroth M, Alarmo E, Nykter M, Kallioniemi A: **Integrated RNA-seq and DNase-seq analyses identify phenotype-specific BMP4 signaling in breast cancer.** *BMC genomics* 2017, **18**(1):68.
 244. Zhong J, Luo K, Winter PS, Crawford GE, Iversen ES, Hartemink AJ: **Mapping nucleosome positions using DNase-seq.** *Genome research* 2016, **26**(3):351-364.
 245. He HH, Meyer CA, Chen M-W, Zang C, Liu Y, Rao PK, Fei T, Xu H, Long H, Liu XS: **Refined DNase-seq protocol and data analysis reveals intrinsic bias in transcription factor footprint identification.** *Nature methods* 2014, **11**(1):73-78.
 246. Barozzi I, Bora P, Morelli MJ: **Comparative evaluation of DNase-seq footprint identification strategies.** *Frontiers in genetics* 2014, **5**.
 247. Consortium EP: **The ENCODE (ENCyclopedia of DNA elements) project.** *Science* 2004, **306**(5696):636-640.
 248. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B: **The accessible chromatin landscape of the human genome.** *Nature* 2012, **489**(7414):75-82.
 249. Neph S, Vierstra J, Stergachis AB, Reynolds AP, Haugen E, Vernot B, Thurman RE, John S, Sandstrom R, Johnson AK: **An expansive human regulatory lexicon encoded in transcription factor footprints.** *Nature* 2012, **489**(7414):83-90.
 250. Neph S, Stergachis AB, Reynolds A, Sandstrom R, Borenstein E, Stamatoyannopoulos JA: **Circuitry and dynamics of human transcription factor regulatory networks.** *Cell* 2012, **150**(6):1274-1286.
 251. Goryshin IY, Reznikoff WS: **Tn5 in vitro transposition.** *Journal of Biological Chemistry* 1998, **273**(13):7367-7374.
 252. Adey A, Morrison HG, Xun X, Kitzman JO, Turner EH, Stackhouse B, MacKenzie AP, Caruccio NC, Zhang X, Shendure J: **Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition.** *Genome biology* 2010, **11**(12):R119.

253. Buenrostro JD, Wu B, Litzenburger UM, Ruff D, Gonzales ML, Snyder MP, Chang HY, Greenleaf WJ: **Single-cell chromatin accessibility reveals principles of regulatory variation.** *Nature* 2015, **523**(7561):486-490.
254. Davie K, Jacobs J, Atkins M, Potier D, Christiaens V, Halder G, Aerts S: **Discovery of transcription factors and regulatory regions driving in vivo tumor development by ATAC-seq and FAIRE-seq open chromatin profiling.** *PLoS genetics* 2015, **11**(2):e1004994.
255. Ackermann AM, Wang Z, Schug J, Naji A, Kaestner KH: **Integration of ATAC-seq and RNA-seq identifies human alpha cell and beta cell signature genes.** *Molecular metabolism* 2016, **5**(3):233-244.
256. Daugherty AC, Yeo RW, Buenrostro JD, Greenleaf WJ, Kundaje A, Brunet A: **Chromatin accessibility dynamics reveal novel functional enhancers in *C. elegans*.** *Genome research* 2017, **27**(12):2096-2107.
257. Scott-Browne JP, López-Moyado IF, Trifari S, Wong V, Chavez L, Rao A, Pereira RM: **Dynamic changes in chromatin accessibility occur in CD8⁺ T cells responding to viral infection.** *Immunity* 2016, **45**(6):1327-1340.
258. Ucar D, Márquez EJ, Chung C-H, Marches R, Rossi RJ, Uyar A, Wu T-C, George J, Stitzel ML, Palucka AK: **The chromatin accessibility signature of human immune aging stems from CD8⁺ T cells.** *Journal of Experimental Medicine* 2017, **214**(10):3123-3144.
259. Picelli S, Björklund ÅK, Reinius B, Sagasser S, Winberg G, Sandberg R: **Tn5 transposase and tagmentation procedures for massively scaled sequencing projects.** *Genome research* 2014, **24**(12):2033-2040.
260. Cusanovich DA, Daza R, Adey A, Pliner HA, Christiansen L, Gunderson KL, Steemers FJ, Trapnell C, Shendure J: **Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing.** *Science* 2015, **348**(6237):910-914.
261. Chen X, Shen Y, Draper W, Buenrostro JD, Litzenburger U, Cho SW, Satpathy AT, Carter AC, Ghosh RP, East-Seletsky A: **ATAC-seq reveals the accessible genome by transposase-mediated imaging and sequencing.** *Nature methods* 2016, **13**(12):1013-1020.
262. Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S, Satpathy AT, Rubin AJ, Montine KS, Wu B: **An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues.** *Nature methods* 2017, **14**(10):959.
263. Song L, Zhang Z, Grasfeder LL, Boyle AP, Girosi PG, Lee B-K, Sheffield NC, Gräf S, Huss M, Keefe D: **Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity.** *Genome research* 2011, **21**(10):1757-1767.
264. Simon JM, Girosi PG, Davis IJ, Lieb JD: **Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA.** *Nature protocols* 2012, **7**(2):256-267.
265. Lieberman-Aiden E, Van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO: **Comprehensive mapping of long-range interactions reveals folding principles of the human genome.** *science* 2009, **326**(5950):289-293.

266. Rusk N: **When ChIA PETs meet Hi-C.** *Nature Methods* 2009, **6**(12):863-863.
267. Van Steensel B, Dekker J: **Genomics tools for unraveling chromosome architecture.** *Nature biotechnology* 2010, **28**(10):1089-1095.
268. Dekker J, Rippe K, Dekker M, Kleckner N: **Capturing chromosome conformation.** *science* 2002, **295**(5558):1306-1311.
269. Wijgerde M, Grosveld F, Fraser P: **Transcription complex stability and chromatin dynamics in vivo.** *Nature* 1995, **377**(6546):209.
270. Dillon N, Trimborn T, Strouboulis J, Fraser P, Grosveld F: **The effect of distance on long-range chromatin interactions.** *Molecular cell* 1997, **1**(1):131-139.
271. Simonis M, Klous P, Splinter E, Moshkin Y, Willemsen R, De Wit E, Van Steensel B, De Laat W: **Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture–on-chip (4C).** *Nature genetics* 2006, **38**(11):1348-1354.
272. Naka K, Hirao A: **Maintenance of genomic integrity in hematopoietic stem cells.** *International journal of hematology* 2011, **93**(4):434-439.
273. Stadhouders R, Kolovos P, Brouwer R, Zuin J, Van Den Heuvel A, Kockx C, Palstra R-J, Wendt KS, Grosveld F, Van Ijcken W: **Multiplexed chromosome conformation capture sequencing for rapid genome-scale high-resolution detection of long-range chromatin interactions.** *Nature protocols* 2013, **8**(3):509-524.
274. Dostie J, Richmond TA, Arnaout RA, Selzer RR, Lee WL, Honan TA, Rubio ED, Krumm A, Lamb J, Nusbaum C: **Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements.** *Genome research* 2006, **16**(10):1299-1309.
275. Hughes JR, Roberts N, McGowan S, Hay D, Giannoulitou E, Lynch M, De Gobbi M, Taylor S, Gibbons R, Higgs DR: **Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment.** *Nature genetics* 2014, **46**(2):205-212.
276. Rodley C, Bertels F, Jones B, O'sullivan J: **Global identification of yeast chromosome interactions using Genome conformation capture.** *Fungal Genetics and Biology* 2009, **46**(11):879-886.
277. Kalhor R, Tjong H, Jayathilaka N, Alber F, Chen L: **Genome architectures revealed by tethered chromosome conformation capture and population-based modeling.** *Nature biotechnology* 2012, **30**(1):90-98.
278. Kolovos P, van de Werken HJ, Kepper N, Zuin J, Brouwer RW, Kockx CE, Wendt KS, van IJcken WF, Grosveld F, Knoch TA: **Targeted Chromatin Capture (T2C): a novel high resolution high throughput method to detect genomic interactions and regulatory elements.** *Epigenetics & chromatin* 2014, **7**(1):10.
279. Tolhuis B, Palstra R-J, Splinter E, Grosveld F, de Laat W: **Looping and interaction between hypersensitive sites in the active β -globin locus.** *Molecular cell* 2002, **10**(6):1453-1465.
280. Ktistaki E, Garefalaki A, Williams A, Andrews SR, Bell DM, Foster KE, Spilianakis CG, Flavell RA, Kosyakova N, Trifonov V: **CD8 locus nuclear**

- dynamics during thymocyte development.** *The Journal of Immunology* 2010, **184**(10):5686-5695.
281. Markova EN, Kantidze OL, Razin SV: **Transcriptional regulation and spatial organisation of the human AML1/RUNX1 gene.** *Journal of cellular biochemistry* 2011, **112**(8):1997-2005.
 282. Blackledge NP, Ott CJ, Gillen AE, Harris A: **An insulator element 3' to the CFTR gene binds CTCF and reveals an active chromatin hub in primary cells.** *Nucleic acids research* 2009, **37**(4):1086-1094.
 283. Nasmyth K, Haering CH: **Cohesin: its roles and mechanisms.** *Annual review of genetics* 2009, **43**:525-558.
 284. Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, Jarmuz A, Canzonetta C, Webster Z, Nesterova T: **Cohesins functionally associate with CTCF on mammalian chromosome arms.** *Cell* 2008, **132**(3):422-433.
 285. Ong C-T, Corces VG: **CTCF: an architectural protein bridging genome topology and function.** *Nature reviews Genetics* 2014, **15**(4):234-246.
 286. Ren G, Jin W, Cui K, Rodriguez J, Hu G, Zhang Z, Larson DR, Zhao K: **CTCF-mediated enhancer-promoter interaction is a critical regulator of cell-to-cell variation of gene expression.** *Molecular cell* 2017, **67**(6):1049-1058. e1046.
 287. Phillips-Cremins JE, Sauria ME, Sanyal A, Gerasimova TI, Lajoie BR, Bell JS, Ong C-T, Hookway TA, Guo C, Sun Y: **Architectural protein subclasses shape 3D organization of genomes during lineage commitment.** *Cell* 2013, **153**(6):1281-1295.
 288. Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES: **A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping.** *Cell* 2014, **159**(7):1665-1680.
 289. Downen JM, Fan ZP, Hnisz D, Ren G, Abraham BJ, Zhang LN, Weintraub AS, Schuijers J, Lee TI, Zhao K: **Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes.** *Cell* 2014, **159**(2):374-387.
 290. Beagan JA, Duong MT, Titus KR, Zhou L, Cao Z, Ma J, Lachanski CV, Gillis DR, Phillips-Cremins JE: **YY1 and CTCF orchestrate a 3D chromatin looping switch during early neural lineage commitment.** *Genome Research* 2017.
 291. Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W: **Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions.** *Nature* 2008, **453**(7197):948-951.
 292. Peric-Hupkes D, Meuleman W, Pagie L, Bruggeman SW, Solovei I, Brugman W, Gräf S, Flicek P, Kerkhoven RM, van Lohuizen M: **Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation.** *Molecular cell* 2010, **38**(4):603-613.
 293. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B: **Topological domains in mammalian genomes identified by analysis of chromatin interactions.** *Nature* 2012, **485**(7398):376-380.

294. Jin F, Li Y, Dixon JR, Selvaraj S, Ye Z, Lee AY, Yen C-A, Schmitt AD, Espinoza CA, Ren B: **A high-resolution map of the three-dimensional chromatin interactome in human cells.** *Nature* 2013, **503**(7475):290-294.
295. Nicodemi M, Pombo A: **Models of chromosome structure.** *Current opinion in cell biology* 2014, **28**:90-95.
296. Barbieri M, Chotalia M, Fraser J, Lavitas L-M, Dostie J, Pombo A, Nicodemi M: **Complexity of chromatin folding is captured by the strings and binders switch model.** *Proceedings of the National Academy of Sciences* 2012, **109**(40):16173-16178.
297. Dai C, Li W, Tjong H, Hao S, Zhou Y, Li Q, Chen L, Zhu B, Alber F, Zhou XJ: **Mining 3D genome structure populations identifies major factors governing the stability of regulatory communities.** *Nature communications* 2016, **7**.
298. Whalen S, Truty RM, Pollard KS: **Enhancer-promoter interactions are encoded by complex genomic signatures on looping chromatin.** *Nature genetics* 2016, **48**(5):488-496.
299. Lan X, Witt H, Katsumura K, Ye Z, Wang Q, Bresnick EH, Farnham PJ, Jin VX: **Integration of Hi-C and ChIP-seq data reveals distinct types of chromatin linkages.** *Nucleic acids research* 2012, **40**(16):7690-7704.
300. Molitor J, Mallm J-P, Rippe K, Erdel F: **Retrieving Chromatin Patterns from Deep Sequencing Data Using Correlation Functions.** *Biophysical Journal* 2017, **112**(3):473-490.
301. Van Bortle K, Corces VG: **Nuclear organization and genome function.** *Annual review of cell and developmental biology* 2012, **28**:163-187.
302. Pancaldi V, Carrillo-de-Santa-Pau E, Javierre BM, Juan D, Fraser P, Spivakov M, Valencia A, Rico D: **Integrating epigenomic data and 3D genomic structure with a new measure of chromatin assortativity.** *Genome biology* 2016, **17**(1):152.
303. Imakaev M, Fudenberg G, McCord RP, Naumova N, Goloborodko A, Lajoie BR, Dekker J, Mirny LA: **Iterative correction of Hi-C data reveals hallmarks of chromosome organization.** *Nature methods* 2012, **9**(10):999-1003.
304. Zhu Y, Chen Z, Zhang K, Wang M, Medovoy D, Whitaker JW, Ding B, Li N, Zheng L, Wang W: **Constructing 3D interaction maps from 1D epigenomes.** *Nature communications* 2016, **7**.
305. Engreitz JM, Agarwala V, Mirny LA: **Three-dimensional genome architecture influences partner selection for chromosomal translocations in human disease.** *PLoS One* 2012, **7**(9):e44196.
306. Barutcu AR, Lajoie BR, McCord RP, Tye CE, Hong D, Messier TL, Browne G, van Wijnen AJ, Lian JB, Stein JL: **Chromatin interaction analysis reveals changes in small chromosome and telomere clustering between epithelial and breast cancer cells.** *Genome biology* 2015, **16**(1):214.
307. Franke M, Ibrahim DM, Andrey G, Schwarzer W, Heinrich V, Schöpflin R, Kraft K, Kempfer R, Jerković I, Chan W-L: **Formation of new chromatin domains determines pathogenicity of genomic duplications.** *Nature* 2016.

308. Weischenfeldt J, Dubash T, Drainas AP, Mardin BR, Chen Y, Stütz AM, Waszak SM, Bosco G, Halvorsen AR, Raeder B: **Pan-cancer analysis of somatic copy-number alterations implicates IRS4 and IGF2 in enhancer hijacking.** *Nature genetics* 2017, **49**(1):65-74.
309. Hnisz D, Weintraub AS, Day DS, Valton A-L, Bak RO, Li CH, Goldmann J, Lajoie BR, Fan ZP, Sigova AA: **Activation of proto-oncogenes by disruption of chromosome neighborhoods.** *Science* 2016:aad9024.
310. Flavahan WA, Drier Y, Liao BB, Gillespie SM, Venteicher AS, Stemmer-Rachamimov AO, Suvà ML, Bernstein BE: **Insulator dysfunction and oncogene activation in IDH mutant gliomas.** *Nature* 2016, **529**(7584):110-114.
311. Lupiáñez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz JM, Laxova R: **Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions.** *Cell* 2015, **161**(5):1012-1025.
312. Chowdhury R, Ganeshan B, Irshad S, Lawler K, Eisenblätter M, Milewicz H, Rodriguez-Justo M, Miles K, Ellis P, Groves A: **The use of molecular imaging combined with genomic techniques to understand the heterogeneity in cancer metastasis.** *The British journal of radiology* 2014, **87**(1037):20140065.
313. Knecht H, Mai S: **The Use of 3D Telomere FISH for the Characterization of the Nuclear Architecture in EBV-Positive Hodgkin's Lymphoma.** *Epstein Barr Virus: Methods and Protocols* 2017:93-104.
314. Tang B, Zhou Y, Wang C-M, Huang TH-M, Jin VX: **Integration of DNA methylation and gene transcription across nineteen cell types reveals cell type-specific and genomic region-dependent regulatory patterns.** *Scientific reports* 2017, **7**(1):3626.
315. Tian F, Luo J, Zhang H, Chang S, Song J: **MiRNA expression signatures induced by Marek's disease virus infection in chickens.** *Genomics* 2012, **99**(3):152-159.
316. Morgan RW, Sofer L, Anderson AS, Bernberg EL, Cui J, Burnside J: **Induction of host gene expression following infection of chicken embryo fibroblasts with oncogenic Marek's disease virus.** *Journal of virology* 2001, **75**(1):533-539.
317. Perumbakkam S, Muir WM, Black-Pyrkosz A, Okimoto R, Cheng HH: **Comparison and contrast of genes and biological pathways responding to Marek's disease virus infection using allele-specific expression and differential expression in broiler and layer chickens.** *BMC genomics* 2013, **14**(1):64.
318. Hillier LW, Miller W, Birney E, Warren W, Hardison RC, Ponting CP, Bork P, Burt DW, Groenen MA, Delany ME: **Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution.** *Nature* 2004, **432**(7018):695-716.
319. Langmead B, Salzberg SL: **Fast gapped-read alignment with Bowtie 2.** *Nature methods* 2012, **9**(4):357-359.

320. Robinson MD, McCarthy DJ, Smyth GK: **edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.** *Bioinformatics* 2010, **26**(1):139-140.
321. Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, Guo Y, Stephens R, Baseler MW, Lane HC: **DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists.** *Nucleic acids research* 2007, **35**(suppl 2):W169-W175.
322. Dessì A, Atzori L, Noto A, Adriaan Visser GH, Gazzolo D, Zanardo V, Barberini L, Puddu M, Ottonello G, Atzei A: **Metabolomics in newborns with intrauterine growth retardation (IUGR): urine reveals markers of metabolic syndrome.** *Journal of Maternal-Fetal and Neonatal Medicine* 2011, **24**(S2):35-39.
323. Anders S, Huber W: **Differential expression analysis for sequence count data.** *Genome biol* 2010, **11**(10):R106.
324. Flicek P, Amode MR, Barrell D, Beal K, Billis K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fitzgerald S: **Ensembl 2014.** *Nucleic acids research* 2013:gkt1196.
325. Kanehisa M, Goto S: **KEGG: kyoto encyclopedia of genes and genomes.** *Nucleic acids research* 2000, **28**(1):27-30.
326. Kaiser P, Underwood G, Davison F: **Differential cytokine responses following Marek's disease virus infection of chickens differing in resistance to Marek's disease.** *Journal of virology* 2003, **77**(1):762-768.
327. Abdul-Careem MF, Hunter BD, Parvizi P, Haghighi HR, Thanthrige-Don N, Sharif S: **Cytokine gene expression patterns associated with immunization against Marek's disease in chickens.** *Vaccine* 2007, **25**(3):424-432.
328. Lian L, Qu L, Sun H, Chen Y, Lamont S, Liu C, Yang N: **Gene expression analysis of host spleen responses to Marek's disease virus infection at late tumor transformation phase.** *Poultry science* 2012, **91**(9):2130-2138.
329. Kano R, Konnai S, Onuma M, Ohashi K: **Cytokine profiles in chickens infected with virulent and avirulent Marek's disease viruses: Interferon - gamma is a key factor in the protection of Marek's disease by vaccination.** *Microbiology and immunology* 2009, **53**(4):224-232.
330. Lei Y, Takahama Y: **XCL1 and XCR1 in the immune system.** *Microbes and Infection* 2012, **14**(3):262-267.
331. Weisberg SP, Hunter D, Huber R, Lemieux J, Slaymaker S, Vaddi K, Charo I, Leibel RL, Ferrante Jr AW: **CCR2 modulates inflammatory and metabolic effects of high-fat feeding.** *Journal of Clinical Investigation* 2006, **116**(1):115.
332. Faure S, Meyer L, Costagliola D, Vaneensberghe C, Genin E, Autran B, French A, Delfraissy J-F, McDermott DH, Murphy PM: **Rapid progression to AIDS in HIV+ individuals with a structural variant of the chemokine receptor CX3CR1.** *Science* 2000, **287**(5461):2274-2277.
333. Arenzana-Seisdedos F, Parmentier M: **Genetics of resistance to HIV infection: Role of co-receptors and co-receptor ligands.** In: *Seminars in immunology: 2006*; Elsevier; 2006: 387-403.

334. Imai T, Hieshima K, Haskell C, Baba M, Nagira M, Nishimura M, Kakizaki M, Takagi S, Nomiyama H, Schall TJ: **Identification and molecular characterization of fractalkine receptor CX 3 CR1, which mediates both leukocyte migration and adhesion.** *Cell* 1997, **91**(4):521-530.
335. Mar A-C, Lee H-J, Lee T-C: **Proinflammatory activity of type II interleukin 1 receptor, a decoy receptor.** *Cancer Research* 2013, **73**(8 Supplement):1423-1423.
336. Fernandez-Botran R, Sanders V, Mosmann T, Vitetta E: **Lymphokine-mediated regulation of the proliferative response of clones of T helper 1 and T helper 2 cells.** *The Journal of experimental medicine* 1988, **168**(2):543-558.
337. Parrish-Novak J, Dillon SR, Nelson A, Hammond A, Sprecher C, Gross JA, Johnston J, Madden K, Xu W, West J: **Interleukin 21 and its receptor are involved in NK cell expansion and regulation of lymphocyte function.** *Nature* 2000, **408**(6808):57-63.
338. Ramana CV, Chatterjee-Kishore M, Nguyen H, Stark GR: **Complex roles of Stat1 in regulating gene expression.** *Oncogene* 2000, **19**(21):2619-2627.
339. Lee C-K, Smith E, Gimeno R, Gertner R, Levy DE: **STAT1 affects lymphocyte survival and proliferation partially independent of its role downstream of IFN- γ .** *The Journal of Immunology* 2000, **164**(3):1286-1292.
340. Sun WH, Pabon C, Alsayed Y, Huang PP, Jandeska S, Uddin S, Platanias LC, Rosen ST: **Interferon- α resistance in a cutaneous T-cell lymphoma cell line is associated with lack of STAT1 expression.** *Blood* 1998, **91**(2):570-576.
341. Satoh J-i, Tabunoki H: **A comprehensive profile of ChIP-Seq-based STAT1 target genes suggests the complexity of STAT1-mediated gene regulatory mechanisms.** *Gene regulation and systems biology* 2013, **7**:41.
342. Kellokumpu - Lehtinen P, Talpaz M, Harris D, Van Q, Kurzrock R, Estrov Z: **Leukemia - inhibitory factor stimulates breast, kidney and prostate cancer cell proliferation by paracrine and autocrine pathways.** *International journal of cancer* 1996, **66**(4):515-519.
343. Cao W, Yang Y, Wang Z, Liu A, Fang L, Wu F, Hong J, Shi Y, Leung S, Dong C: **Leukemia inhibitory factor inhibits T helper 17 cell differentiation and confers treatment effects of neural progenitor cell therapy in autoimmune disease.** *Immunity* 2011, **35**(2):273-284.
344. Mitra A, Luo J, He Y, Gu Y, Zhang H, Zhao K, Cui K, Song J: **Histone modifications induced by MDV infection at early cytolytic and latency phases.** *BMC genomics* 2015, **16**(1):311.
345. Baigent SJ, Davison F: **Marek's disease virus: biology and life cycle.** In: *Marek's disease, an evolving problem.* Elsevier Academic Press Great Britain; 2004: 62-77.
346. Picelli S, Faridani OR, Björklund ÅK, Winberg G, Sagasser S, Sandberg R: **Full-length RNA-seq from single cells using Smart-seq2.** *Nature protocols* 2014, **9**(1):171-181.
347. Trapnell C, Pachter L, Salzberg SL: **TopHat: discovering splice junctions with RNA-Seq.** *Bioinformatics* 2009, **25**(9):1105-1111.

348. Ramsköld D, Wang ET, Burge CB, Sandberg R: **An abundance of ubiquitously expressed genes revealed by tissue transcriptome sequence data.** *PLoS computational biology* 2009, 5(12):e1000598.
349. Zhou X, Maricque B, Xie M, Li D, Sundaram V, Martin EA, Koebbe BC, Nielsen C, Hirst M, Farnham P: **The human epigenome browser at Washington University.** *Nature methods* 2011, 8(12):989.
350. Consortium GO: **The Gene Ontology (GO) database and informatics resource.** *Nucleic acids research* 2004, 32(suppl_1):D258-D261.
351. Dixon JR, Gorkin DU, Ren B: **Chromatin domains: the unit of chromosome organization.** *Molecular cell* 2016, 62(5):668-680.
352. Caretto D, Katzman SD, Villarino AV, Gallo E, Abbas AK: **Cutting edge: the Th1 response inhibits the generation of peripheral regulatory T cells.** *The journal of immunology* 2010, 184(1):30-34.
353. Zhu J, Jankovic D, Oler AJ, Wei G, Sharma S, Hu G, Guo L, Yagi R, Yamane H, Punkosdy G: **The transcription factor T-bet is induced by multiple pathways and prevents an endogenous Th2 cell program during Th1 cell responses.** *Immunity* 2012, 37(4):660-673.
354. Hu G, Tang Q, Sharma S, Yu F, Escobar TM, Muljo SA, Zhu J, Zhao K: **Expression and regulation of intergenic long noncoding RNAs during T cell development and differentiation.** *Nature immunology* 2013.
355. Barski A, Cuddapah S, Cui K, Roh T-Y, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K: **High-resolution profiling of histone methylations in the human genome.** *Cell* 2007, 129(4):823-837.
356. Zang C, Schones DE, Zeng C, Cui K, Zhao K, Peng W: **A clustering approach for identification of enriched domains from histone modification ChIP-Seq data.** *Bioinformatics* 2009, 25(15):1952-1958.
357. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G: **GREAT improves functional interpretation of cis-regulatory regions.** *Nature biotechnology* 2010, 28(5):495.
358. Sherwood RI, Hashimoto T, O'donnell CW, Lewis S, Barkal AA, Van Hoff JP, Karun V, Jaakkola T, Gifford DK: **Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape.** *Nature biotechnology* 2014, 32(2):171.
359. Kulakovskiy IV, Vorontsov IE, Yevshin IS, Soboleva AV, Kasianov AS, Ashoor H, Ba-Alawi W, Bajic VB, Medvedeva YA, Kolpakov FA: **HOCOMOCO: expansion and enhancement of the collection of transcription factor binding sites models.** *Nucleic acids research* 2015, 44(D1):D116-D125.
360. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK: **Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities.** *Molecular cell* 2010, 38(4):576-589.
361. Chepelev I, Wei G, Wangsa D, Tang Q, Zhao K: **Characterization of genome-wide enhancer-promoter interactions reveals co-expression of interacting genes and modes of higher order chromatin organization.** *Cell research* 2012, 22(3):490.

362. Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, Orlov YL, Velkov S, Ho A, Mei PH: **An oestrogen-receptor- α -bound human chromatin interactome.** *Nature* 2009, **462**(7269):58.
363. Hollander M, Wolfe DA: **Nonparametric statistical methods.** 1999.
364. Wang Y, Jiang R, Wong WH: **Modeling the causal regulatory network by integrating chromatin accessibility and transcriptome data.** *National science review* 2016, **3**(2):240-251.
365. Amarnath S, Mangus CW, Wang JC, Wei F, He A, Kapoor V, Foley JE, Massey PR, Felizardo TC, Riley JL: **The PDL1-PD1 axis converts human TH1 cells into regulatory T cells.** *Science translational medicine* 2011, **3**(111):111ra120-111ra120.
366. Brown CC, Esterhazy D, Sarde A, London M, Pullabhatla V, Osma-Garcia I, al-Bader R, Ortiz C, Elgueta R, Arno M: **Retinoic acid is essential for Th1 cell lineage stability and prevents transition to a Th17 cell program.** *Immunity* 2015, **42**(3):499-511.
367. Schraml BU, Hildner K, Ise W, Lee W-L, Smith WA-E, Solomon B, Sahota G, Sim J, Mukasa R, Cemurski S: **The AP-1 transcription factor Batf controls T H 17 differentiation.** *Nature* 2009, **460**(7253):405.
368. Dang EV, Barbi J, Yang H-Y, Jinasena D, Yu H, Zheng Y, Bordman Z, Fu J, Kim Y, Yen H-R: **Control of T H 17/T reg balance by hypoxia-inducible factor 1.** *Cell* 2011, **146**(5):772-784.
369. Zhou X, Bailey-Bucktrout SL, Jeker LT, Penaranda C, Martínez-Llordella M, Ashby M, Nakayama M, Rosenthal W, Bluestone JA: **Instability of the transcription factor Foxp3 leads to the generation of pathogenic memory T cells in vivo.** *Nature immunology* 2009, **10**(9):1000.
370. Kidane FA, Bilic I, Mitra T, Wernsdorf P, Hess M, Liebhart D: **In situ hybridization to detect and localize signature cytokines of T-helper (Th) 1 and Th2 immune responses in chicken tissues.** *Veterinary immunology and immunopathology* 2016, **175**:51-56.
371. Chen R, Bélanger S, Frederick MA, Li B, Johnston RJ, Xiao N, Liu Y-C, Sharma S, Peters B, Rao A: **In vivo RNA interference screens identify regulators of antiviral CD4+ and CD8+ T cell differentiation.** *Immunity* 2014, **41**(2):325-338.
372. Clift D, McEwan WA, Labzin LI, Konieczny V, Mogessie B, James LC, Schuh M: **A method for the acute and rapid degradation of endogenous proteins.** *Cell* 2017, **171**(7):1692-1706. e1618.