blood

Prepublished online October 20, 2014;
doi:10.1182/blood-2014-06-582700

Global transcriptome analysis and enhancer landscape of human primary T follicular helper and T effector lymphocytes

Jason S. Weinstein, Kimberly Lezon-Geyda, Yelena Maksimova, Samuel Craft, Yaoping Zhang, Mack Su, Vincent P. Schulz, Joseph Craft and Patrick G. Gallagher

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036.

Blood First Edition Paper, prepublished online October 20, 2014; DOI 10.1182/blood-2014-06-582700

BLOOD/2014/582700 Revision

Global Transcriptome Analysis and Enhancer Landscape of Human Primary T Follicular Helper and T Effector Lymphocytes

Jason S. Weinstein,[1] Kimberly Lezon-Geyda,[2] Yelena Maksimova,[2] Samuel Craft,[2] Yaoping Zhang,[2] Mack Su,[2]

Vincent P. Schulz,[2] Joseph Craft,[1,3*] Patrick G. Gallagher[2,4*]

[1]Department of Internal Medicine (Rheumatology), Yale School of Medicine, New Haven, CT 06520

[2]Department of Pediatrics, Yale University School of Medicine, New Haven, CT, 06520

[3]Department of Immunobiology, Yale University School of Medicine, New Haven, CT 06520

[4]Departments of Pathology and Genetics, University School of Medicine, New Haven, CT 06520

* Co-corresponding and senior authors

Running Title: Human Tfh cell enhancers

Word Count: Abstract: 200; Text 4299; Figures: 7; Tables 3; References 125. This manuscript has a supplement.

Contacts: Joseph Craft, joseph.craft@yale.edu, Tel: (203) 785-7063, Fax: (203) 785-5415 and

Patrick G. Gallagher, patrick.gallagher@yale.edu, Tel: 203 688 2896, Fax: (203) 785 6974.

Key Points

- Transcriptomes and enhancers of human $CD4^+$ T follicular helper and non-Tfh T effector cells reveal cell type-specific differences.
- These data are a significant resource for understanding mechanisms of normal and perturbed T follicular helper cell function.

Abstract

T follicular helper (Tfh) cells are a subset of $CD4^+$ T helper (Th) cells that migrate into germinal centers and promote B cell maturation into memory B and plasma cells. Tfh cells are necessary for promotion of protective humoral immunity following pathogen challenge, but when aberrantly regulated, drive pathogenic antibody formation in autoimmunity and undergo neoplastic transformation in angioimmunoblastic T-cell lymphoma and other primary cutaneous T-cell lymphomas. Limited information is available on the expression and regulation of genes in human Tfh cells. Using a fluorescence activated cell sorting-based strategy, we obtained primary Tfh and non-Tfh T effector (Teff) cells from tonsils and prepared genome-wide maps of active, intermediate, and poised enhancers determined by ChIP-seq, with parallel transcriptome analyses determined by RNA-seq. Tfh cell enhancers were enriched near genes highly expressed in lymphoid cells or involved in lymphoid cell function, with many mapping to sites previously associated with autoimmune disease in genome-wide association studies. A group of active enhancers unique to Tfh cells associated with differentially expressed genes was identified. Fragments from these regions directed expression in reporter gene assays. These data provide a significant resource for studies of T lymphocyte development and differentiation and normal and perturbed Tfh cell function.

Abbreviations: Tfh: $CD4^+$ T follicular helper; Teff: $CD4^+$ T effector; Th: $CD4^+$ T helper; PD-1: programmed cell death receptor 1; interleukin: IL; SLE, systemic lupus erythematosus; RNA-seq: RNA sequencing; FACS: fluorescence activated cell sorting; ChIP: chromatin immunoprecipitation; ChIP-seq: chromatin immunoprecipitation-sequencing; SNP: single nucleotide polymorphism; GWAS: genome-wide association studies; TSS: transcriptional start site; TES: transcriptional end site; GSEA: Gene Set Enrichment Analysis; AITL: angioimmunoblastic T-cell lymphoma; PTCL: peripheral T-cell lymphoma (PTCL)

Introduction

T follicular helper (Tfh) cells are a subset of $CD4^+$ T helper (Th) lymphocytes that migrate into the B cell follicle and provide germinal center (GC) B cells with survival and differentiation signals essential for B cell selection with maturation into memory B cells and long-lived antibody-secreting plasma cells.[1-8] Tfh cells also secrete cytokines that enable B cell isotype class switching appropriate to invading pathogens.[8-10] Tfh cells can be distinguished from other Th cells by down regulation of P-selectin glycoprotein ligand 1 (PSGL-1) required for their emigration from T cell zones of secondary lymphoid organs towards the B cell follicle, and by their sustained expression of the transcriptional repressor B-cell lymphoma 6 (BCL6), the C-X-C chemokine receptor type 5 (CXCR5) needed for their migration into the follicle, and the programmed cell death receptor (PD-1) necessary for proper B cell maturation therein in GCs.[11,12] Although Tfh cells are essential for the GC response, much less is known about their origin, development and function compared to other CD4 Th cell subsets.[13]

Tfh cells are abnormally regulated in several inherited and acquired diseases.[14,15] Expansion of dysfunctional Tfh cells is a major contributor to systemic autoimmunity, including systemic lupus erythematosus (SLE, lupus), Sjogren's syndrome, and rheumatoid arthritis.[16,17] Their malignant transformation results in the phenotype of angioimmunoblastic T-cell lymphoma (AITL), a subset of peripheral T-cell lymphoma (PTCL).[18-22] Tfh cells are thought to be the origin of subtypes of primary cutaneous T-cell lymphoma.[23,24] A possible contributory role for Tfh cells in graft versus host disease also has been suggested.[25]

Recent advances in genomic technologies have revolutionized our understanding of gene expression and gene regulation, and their relationship to mechanisms of human disease.[26] Detailed information on cellular transcriptomes obtained by RNA sequencing (RNA-seq) provides unbiased information on transcript composition and abundance, including detection of novel transcripts, novel isoforms, alternative splicing, and allele-specific expression.[27-29] Similarly, genomic strategies have allowed understanding of programs controlling cellular development and differentiation by providing insight into the regulatory DNA sequences that control or regulate these programs.

Enhancers are DNA regulatory sequences with numerous, complex roles in the control of gene expression,[30-33] participating in cellular development, differentiation and cell fate determination.[34-37] They assist in determining nuclear organization,[33] transcription initiation and the release of RNA polymerase II from promoter pausing,[38] transcriptional competence,[36] and insulator element activity.[39,40] Noncoding RNAs have also been linked to enhancer function[41-47] and intergenic enhancers may act as alternate, tissue-specific promoters generating abundant, spliced, multiexonic $poly(A)^+$ RNAs.[48] Secondary enhancers synergize with primary enhancers to fine tune gene expression.[49,50] Recent studies in three dimensional transcriptional space reveal that turning on and off enhancers during development correlates with promoter activity and that promoter-enhancer interactions are highly cell-type specific varying widely across the genome.[51-54]

Numerous studies characterizing enhancers in human lymphoid cells on a genome wide scale have been performed.[55-65] Despite their biologic relevance, data are not available for human primary Tfh cell enhancers, perhaps because of the difficulty in obtaining adequate samples for analysis. Obtaining sufficient numbers from mice is also

challenging, in light of the challenge differentiating these cells *in vitro*, in comparison to other Th cell subsets.[66] Using a FACS-based strategy, we obtained Tfh cells, and for comparison, non-Tfh T effector cells (hereafter, T effector or Teff cells) from tonsils. Using these purified samples, we constructed and analyzed genome-wide maps of active, intermediate, and poised enhancers, with integration of global transcriptome analyses determined by RNA-seq. Consistent with their predicted function, these important regulatory elements were enriched near genes highly expressed in lymphoid cells or involved in lymphoid cell structure and function. Many Tfh cell enhancers mapped to sites previously associated with autoimmune disease in genome-wide association studies (GWAS). A group of differentially marked active enhancers unique to Tfh cells associated with differentially expressed genes was identified. This group contained genes expressed at high levels, including *PDCD1* and *BCL6*, which are critical for Tfh-cell function. Fragments from several enhancer regions were also associated with directed statistically significant expression in reporter gene assays.

Together, these data provide a significant resource for studies of programs of gene expression in Tfh and non-Tfh Teff cells and their regulation. This will allow a deeper understanding of $CD4^+$ lymphocyte development, differentiation, structure, and function, mechanisms of the GC response, and provide insights into normal and perturbed Tfh cell function including that associated with immune disorders and lymphomas.

Methods

Flow cytometry and cell sorting

To obtain primary human Tfh and Teff cells, human tonsils were cut into small sections and homogenized by crushing followed by straining through a 40μM nylon filter. $CD4^+$ T lymphocytes were enriched using a negative selection biotin-based magnetic separation kit (EasySep; StemCell Technologies) prior to cell surface staining with the antibodies to: CD4 (clone RPA-T4), CD45RA (clone HI-100), TCRβ (clone IP26), PD-1 (clone EH12.1), CXCR5 (clone RF8B2), and PSGL-1 (clone KPL-1)(all from BD Biosciences). Staining of CXCR5 was performed at room temperature (25°C) with 1 hour incubation. Intracellular staining for Foxp3 was performed using Foxp3/Transcription Factor Staining Buffer Set™ (eBiosciences) following the manufacturer's protocol. Stained and rinsed cells were analyzed using an LSRII Multilaser Cytometer (BD Biosciences) or specific populations sorted using a FACSAria (BD Bioscience).

Cell selection and RNA analyses

RNA was isolated from lymphocytes and prepared for RNA-seq analyses as described.[67] Samples were sequenced on an Illumina HiSeq 2000 using 75bp-paired end reads. FASTQ format sequencing reads were aligned to the hg19 genome, NCBI Build 37, using TopHat Version 2.0.4 software with default parameters except minimum anchor length of 12. The EdgeR program was used to identify differences in expression of RefSeq transcripts. Filtering included transcripts with >1 tag/million reads in 3 or more samples. Quantitative real-time PCR was performed to confirm expression levels of RNA transcripts (Supplemental Table S1). Gene Set Enrichment Analysis (GSEA) was performed as described using default parameters except 10,000 permutations were performed.[68]

Chromatin immunoprecipitation (ChIP)

ChIP assays were performed as previously described.[69] Samples were immunoprecipitated with antibodies against monomethyl histone 3 lysine 4 (H3k4me1, Abcam ab8895), trimethyl histone 3 lysine 27 (Abcam ab6002), acetyl histone 3 lysine 27 (H3K27Ac, Abcam ab4729), and nonspecific rabbit IgG (sc-2091, Santa Cruz).

High throughput sequencing and data analyses

DNA processing and high throughput sequencing were performed as described.[67] Sequenced reads were mapped to the human genome (hg19[70] NCBI Build 37) using the BWA alignment program. The Model-based Alignment of ChIP-Seq (MACS) program version 1.4.0rc2 was used to identify H3K4me1 and H3K27Ac peaks with a p-value<10e-5.[71] The MACS2 program version 2.0.10.20131216 was used to identify broad regions bound by H3K27me3 that had an enrichment of four fold or more. Localization of histone modifications relative to known genes was done using customized BedTools scripts. Motif finding was performed using the Homer algorithm (http://homer.salk.edu/homer/motif/). Conservation analyses were performed using PhastCons.[72,73] The Genomic Regions Enrichment Annotations Tool (GREAT) was used to analyze functional significance of *cis*-regulatory regions identified by ChIP-seq.[74]

Validation of ChIP-seq results

Primers were designed for representative binding regions for all 3 antibodies in candidate enhancers identified by the MACS program (Supplemental Table S2). Immunoprecipitated DNA was analyzed by quantitative real-time PCR (iCycler, Bio-Rad) as described.[69]

Identification and analysis of biologically relevant SNPs

The locations of SNPs shown to demonstrate highly significant linkage to immune disorder-related traits (Supplemental Table S3) were obtained from the UCSC genome browser database and the catalog of published Genome-Wide Association Studies (GWAS) compiled by the National Human Genome Research Institute (www.genome.gov/gwastudies).[75] Using BedTools software, enhancers were intersected with immune disorder-related SNPs and overlap identified.

Reporter gene assays

Candidate enhancer regions were amplified using primers flanking the boundaries of called peaks (Supplemental Table S4) and cloned upstream of an SV40 promoter-firefly luciferase reporter cassette in the pGL2Promoter plasmid. Transfections were performed as described.[76] 10^7 Jurkat cells (ATCC, TIB-152) were transfected by electroporation with a single pulse of 300V at 950μF with 15μg of test plasmid and 0.3μg of pRL-TK, a reporter plasmid expressing *Renilla* luciferase driven by the *Herpes simplex* virus thymidine kinase (HSV-TK) promoter (Promega) as described.[77] Two days after transfection, cell extracts were analyzed using the Dual-Luciferase assay according to manufacturer's instructions (Promega).

Data access

The raw data files generated by the RNA seq and ChIP-seq analyses have been submitted to Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/, Reference series GSE58597).

Results

Transcriptome analyses of primary human naïve, Tfh and Teff cells by RNA-seq. Human primary naïve ($CD4^+CD45RA^-TCR\beta^+$), Tfh ($CD4^+CD45RA^-TCR\beta^+PD\text{-}1^{hi}CXCR5^{hi}PSGL\text{-}1^{lo}$) and Teff cells ($CD4^+CD45RA^-TCR\beta^+PD\text{-}1^{lo}CXCR5^{lo}PSGL\text{-}1^{hi}$) cells were isolated from human tonsils using the sorting strategy described (Figure 1A). A population of T follicular regulatory (Tfr) cells which exhibit a phenotype similar to Tfh cells, *i.e.*, $CXCR5^{hi}$ and $PD\text{-}1^{hi}$ expression, may be present within our sorted Tfh cell fraction. To quantify the amount of Tfr cells in our sorted Tfh cell population, we performed intracellular staining for Foxp3 in Tfh cells from 5 different tonsils and found that Tfr cells ($CD4^+CD45RA^-PD\text{-}1^{hi}CXCR5^{hi}PSGL\text{-}1^{lo}Foxp3^+$) comprised < 0.5% of the tonsillar Tfh cell population ($p<0.0079$; Figure 1B). Thus, the small numbers of Tfr cells found in the Tfh cell sorting gate should not significantly bias our transcriptome and genomic analyses.

RNA was isolated and RNA-seq performed to obtain the transcriptome of Tfh, Teff, and naïve T cells. Multidimensional scaling was performed on expressed genes (>1cpm in 3 or more samples) to assess sample relatedness. Samples from each cell type clustered together (Figure 2A), indicating that samples from each cell type were closely related and distinct from the other cell types. Quantitative real-time PCR validated expression levels of representative mRNA transcripts detected by RNA-seq (Supplemental Table S5). Overall, 11,839, 12,007, and 12,202 transcripts were expressed in Tfh, Teff, and naïve T cell RNA, respectively, with 11,307 expressed in all 3 cell types (Supplemental Figure S1).

EdgeR, a Bioconductor software package for examining differential expression of replicated count data, identified the differentially expressed transcripts among the three cell types (Figure 2B). These genes were sorted by absolute differences (fold change) in their expression (Supplemental Table S6) and heat maps displaying patterns of differentially expressed genes prepared (Figure 2C and Supplemental Figure S2). Several of the most differentially expressed genes with increased expression in Tfh cells encoded proteins critical for Tfh cell function including *PDCD1*, *CXCR5*, and *BCL6* (Figure 2D), validating our sorting strategy. In parallel, several differentially expressed genes, including *CCR7* and *IL7R*, exhibited increased expression in Teff cells, consistent with their phenotype (Table 1). The top 4 functional networks, as assessed by Ingenuity Pathway Analysis (IPA) of differentially expressed genes between Tfh and Teff cells, all with scores $\geq$30, included cell movement, hematological system development and function, and immune cell trafficking (network 1); cellular development, hematological system development and function, and hematopoiesis (network 2); cell cycle, dermatological diseases, and gastrointestinal disease (network 3); and hematological system development and function, immune cell trafficking, and inflammatory response (network 4). The top categories associated with the category *Diseases and Disorders* included inflammatory response, connective tissue disorders, skeletal and muscular disorders, inflammatory disease, and immunological disease.

Gene Set Enrichment Analysis (GSEA). Angioimmunoblastic T-cell lymphoma (AITL) is a subtype of peripheral T-cell lymphoma (PTCL) with a molecular signature indicating origin from Tfh cells (see below).[20,78-83] To better understand Tfh cell gene expression in AITL, we performed GSEA using a gene set containing genes up regulated in Tfh

cells compared to Teff cells. We compared 37 AITL samples from the dataset GSE19069,[78] to the 60 non-AITL samples, primarily containing other malignancies of T-cell origin, in the data set (Supplemental Figure S3). The normalized enrichment score, which reflects the degree to which a gene set is overrepresented at the top or bottom of a ranked list of genes, was 2.02 (*p*-value <0.00001) indicating highly significant enrichment of Tfh cell-expressed genes in the AITL samples.

H3K4me1, H3K27me3, and H3K27Ac occupancy in Tfh and Teff cell chromatin. ChIP-seq was performed with antibodies for H3K4me1, H3K27me3, and H3K27Ac using primary Tfh and Teff cell chromatin to generate genome-wide maps of histone architecture. The MACS program was used to identify peaks with a cutoff of *p*-value<10e-5. Validation of histone modification enrichment at selected peaks identified by ChIP-seq was performed by quantitative ChIP PCR for all 4 antibodies (Supplemental Table S7).

Identification of multiple classes of Tfh and Teff cell enhancers. Enhancers are frequently marked by monomethylation of lysine 4 of histone H3 (H3K4me1), with further classification as active and poised. Active enhancers are dynamic participants in gene expression whereas poised enhancers participate in gene expression in response to various cellular stimuli, *e.g.*, differentiation cues. Zentner and colleagues refined this classification, further classifying enhancers as active, poised, and intermediate based on chromatin architecture, conservation, genomic location, levels of gene expression of associated genes, and predicted function of associated genes.[84,85]

Using these definitions, we identified and classified Tfh and Teff cell enhancers as active, intermediate, and poised. There were 17229 active, 28098 intermediate, and 2055 poised enhancers in Tfh cells and 16707 active, 14,555 intermediate, and 940 poised enhancers in Teff cells. Aggregate plots of each enhancer class centered on peaks of H3K4me1 demonstrated high amounts of H3K27Ac in active enhancers and enrichment of H3K27me3 in poised enhancers (Figure 3A-C). Heat maps generated by ranked plots averaged across rows of HeK4me1, H3K27Ac, and/or H3K27me3 peaks revealed patterns of distinct chromatin architecture characteristic of the enhancer classes (Figure 3D).

The human genome was portioned into seven bins relative to RefSeq genes corresponding to exons, introns, promoters, distal (-1 to -50kb), downstream (+1 to +50kb), and intergenic regions. Sites of specific histone modifications and enhancer class were assigned to these bins and percentages calculated (Figure 4A). In Tfh cells, poised enhancers were less frequently found in introns, but were found more frequently in the 5' and 3' distal regions of associated genes. Similar localization of poised enhancers was observed in Teff cells (Figure 4B).

Expression and function of genes associated with enhancer classes in Tfh and Teff cells. Typically, cell and tissue-type specific enhancers act over distances of tens to hundreds of kilobases.[86] Thus, *bona fide* Tfh and Teff cell enhancers are expected to be enriched in the genomic vicinity of genes that are expressed and functional in their respective cells.[87-89] To determine whether Tfh and Teff cell enhancers are localized in this manner, gene expression in Tfh and Teff cells was correlated with each enhancer class. To exclude gene promoters, localization of H3K4me1 within 1kb of annotated transcriptional start sites (TSS) was excluded from the analyses. There was a statistically significant higher expression of genes with active enhancers 1-50 kb from a TSS compared to expression of genes associated with active enhancers >50kb of a TSS in Tfh and Teff cells (both *p*-values < 2.2e-16, Supplemental Figures S4A and S4B).

Levels of gene expression of associated genes were highest for active enhancers and lowest for poised enhancers, with levels of expression between these values for intermediate enhancers in Tfh cells 1-50 kb and >50kb of a TSS.

We also examined whether enhancers were enriched near genes with known Tfh and Teff cell function. We performed a statistical enrichment analysis of functional gene annotations associated with each type of enhancer class.[74] Tfh cell active and intermediate enhancers were associated with genes linked to immune cell-related biological functions whereas poised enhancers were associated with genes involved in cellular development and differentiation (Supplemental Figure S5A). The top enriched phenotypes linked to genes associated with Tfh cell active enhancers were immune and lymphocyte-related phenotypes (Table 2). Teff cell active enhancers were associated with genes linked to immune signaling, Teff intermediate enhancers were associated with genes linked to immune responses, and poised Teff enhancers were associated with genes linked to transcription (Supplemental Figure S5B).

Conservation analyses by enhancer class. Conservation plots using PhastCons conservation scores with the 46-way vertebrate hg19 multiple alignment PhastCons track were constructed for each enhancer class of Tfh and Teff cells. Similar to results previously observed,[85] only poised enhancers showed strong conservation in both cell types (Supplemental Figure S6).

Motif enrichment. The Homer program was utilized to identify over-represented DNA/transcription factor motifs at sites of candidate enhancers. The top over-represented motifs at sites of active enhancers in Tfh cells were ETS, ZIC2, TLX, and MEF2C with over-representation of ETS, RUNX1, and DCE in Teff cells (Supplemental Figure S7).

Enhancer classes and biologically relevant single nucleotide polymorphisms. We explored whether SNPs associated with biologically relevant immune cell traits and with immune cell traits were enriched in Tfh cell enhancers. We utilized the set of non-coding SNPs from the GWAS catalog of the NHGRI (www.genome.gov/gwastudies)[75] and we collected a data set of immune-associated non-coding SNPs. (See Methods; there are 1514 SNPs associated with these terms from the GWAS catalog.) Currently, the functional significance of the overwhelming majority of these SNPs is unknown. SNP locations were compared to the locations of Tfh cell enhancers. For active Tfh enhancers, there was association with 321 SNPs in the GWAS catalog with 86 (5.7%) related to immune cell traits (Supplemental Table S8). For intermediate Tfh enhancers, there was association of 197 SNPs in the total GWAS catalogue with 38 (2.5%) related to immune cell traits. For poised Tfh enhancers, there were no SNPs linked to immune cell traits (Supplemental Table S8). Examples of immune disease-associated SNPs with Tfh cell active enhancers are shown in Supplemental Figure S8.

A subset of enhancers, called super enhancers or stretch enhancers, important for regulating genes critical for cell-type specific identify, have been described.[90,91] Super enhancers span large regions of chromatin, have domains of transcription factor binding sites and are marked by significant amounts of H3K4me1 and H3K27Ac modification. We identified super enhancers in Tfh cell chromatin as described by finding regions with the highest levels of clustered, K27 acetylated chromatin (Figure 5).[90,92] In some cell types, disease-associated SNPs are enriched in super enhancers of relevant cell types, suggesting that altered expression of key cell identity genes may contribute to disease phenotype.[91,92] For Tfh cell super enhancers, there was association with 893 SNPs in the total GWAS catalog with 119 SNPs (13%) linked to immune cell traits (Supplemental Table S9). Relevant SNPs were found in super enhancers near the *MAF*, *IRF4*,

and *BATF* gene loci (Figure 5). For super enhancer-associated SNPs, the majority were associated with rheumatoid arthritis, type 1 diabetes, multiple sclerosis, celiac disease, and ulcerative colitis. Thus SNPs associated with biologically relevant, disease-associated immune-cell traits were significantly enriched in super enhancers compared to active or intermediate Tfh cell enhancers.

Tfh cell type specific enhancers. To assess the cell-type specificity of the enhancers identified, we compared Tfh enhancers marked by non-promoter associated peaks of H3K4 monomethylation to enhancers in 61 different cell types. This identified 1660 enhancers that were present in Tfh cell chromatin but were not present in the other cell types. To further refine cell type specificity, we next compared enhancers marked by non-promoter associated peaks of H3K4 monomethylation in Tfh cells to enhancers in 13 hematopoietic cell types, primarily lymphoid cells. We identified 7166 nonpromoter-associated H3K4me1 peaks in Tfh cell chromatin not present in the 13 other cell types, with 9475 not identified in the 10 lymphoid cell types. Jaccard coefficient clustering revealed that hematopoietic cell enhancers clustered distinctly from human embryonic stem cell enhancers (for comparison) and lymphoid cell enhancers clustered distinctly from other hematopoietic cell enhancers (not shown).

To identify enhancers potentially contributing to differential gene regulation in Tfh and Teff cells, we compared levels of H3K27 acetylation at non-promoter peaks of H3K4me1 in Tfh and Teff cell chromatin, as relative amounts of acetylation have been correlated with enhancer strength or levels of gene expression.[93] There were 1281 differentially acetylated enhancers between the 2 cell types (Figures 6A and 6B). We then determined if the genes nearest these differentially acetylated enhancers were also differentially expressed in the associated cell type. There were 43 differentially marked enhancers associated with differentially expressed genes (Table 3) expressed at higher levels in Tfh cells, including *PDCD1* and *BCL6* (Figure 7), genes associated with Tfh cell function.

Reporter gene assay of Tfh-cell type specific enhancers. Individual reporter gene plasmids were prepared with representative, differentially acetylated enhancer elements associated with differential gene expression in Tfh cells cloned upstream of an SV40 gene promoter-luciferase reporter gene cassette. These included enhancers associated with the B cell lymphoma 6 (*BCL6)*, histidine-ammonia lyase (*HAL)*, suppressor of tumorigenicity 14 (*ST14*), CXXC finger protein 5 (*CXXC5*), secretoglobin, family 3A, member 1 *(SCGB3A1*), Dab reelin signal transducer homolog 1 (*DAB1),* WNK lysine deficient protein kinase 2 (*WNK2*), and guanine nucleotide binding protein (G protein), gamma 4 (*GNG4)* genes. Plasmids were transfected into human T lymphoid Jurkat cells, which express all but one of the associated genes (Supplemental Table S10), cell lysates collected after 2 days, and luciferase activity analyzed. All 7 Tfh cell-specific enhancers directed statistically significant ($p<0.05$) reporter gene activity between 2 and 3.5 times over activity of control (Supplemental Figure S9).

Discussion

Tfh cells are necessary for B cell maturation into memory and long-lived plasma cells in GCs of B cell follicles.[94] They provide signals to cognate B cells via CD40 ligand, programmed cell death 1 (PD-1), and cytokines, including IL-21 and IL-4, promoting B cell proliferation and affinity maturation within the GC.[95-99] Tfh cell development requires the transcription factors Achaete-scute complex homolog 2 (ASCL2) and BCL6, leading to expression of transcripts important for their function, while repressing activation of genes, among them Blimp1 (*PRDM1*), critical for development of other T helper subsets.[5,8,10] ASCL2 and BCL6 up regulation also promotes expression of the chemokine receptor CXCR5 necessary for Tfh cell entry into the B cell follicle following a gradient of its ligand CXCL13.[11,12] Genes encoding these proteins were all highly expressed in the Tfh cells we isolated from tonsils.

Our integrative analyses also revealed that several of these genes were associated with Tfh-cell type specific enhancers, including *PDCD1*, *CXCR5*, and *BCL6*, suggesting that these *cis*-elements are critical for Tfh cell identify and function. However, not all key factors exhibited either differential expression at the transcriptome level and/or Tfh cell-specific enhancers. Although a remote enhancer cannot be excluded, no Tfh cell-type enhancers were identified within 100kb of the *ASCL2* gene locus, indicating additional regulatory factors direct expression of *ASCL2* in Tfh cells.

The master regulator BCL6 regulates a unique program of gene expression essential for Tfh cell differentiation and function.[5,8,100] BCL6 promotes expression of genes important for Tfh cell migration and function, while repressing critical regulators and microRNAs of other T-helper subsets.[8,100] Similar to their relationship in GC B cells, BLC6 and Blimp1 (*PRDM1*) are reciprocal and antagonistic regulators of the Tfh cell phenotype.[5,100-103] The precise mechanism(s) whereby BCL6 controls these processes in Tfh cells is unclear. Numerous functional roles outside of the GC have been described for BCL6, including a late check point function in pre-B cells, generation and maintenance of effector and memory $CD8^+$ and memory $CD4^+$ cells, regulation of effector functions in peripheral Tregs, regulation of Th17 T helper cell differentiation, and acting to constrain immune and inflammatory responses in macrophages (reviewed in [104]). We found BCL6 expressed in Tfh, Teff, and naïve T cells, with markedly increased expression in Tfh cells. In parallel, our studies also identified a Tfh cell-specific enhancer in the *BCL6* gene locus. It will be important to identify the *cis*-regulatory elements that control BCL6 expression in Tfh cells and other related lineages.

Recent reports indicate that there is heterogeneity in Tfh cells,[105,106] including variable populations of Tfh cells circulating in human peripheral blood.[107-111] Although Tfr cells were excluded as a major population of the Tfh cells we obtained from tonsil, variable types of Tfh cells were likely included in the bulk population of Tfh cells. In addition, despite phenotypic and transcriptional differences with other $CD4^+$ Th subsets, Tfh cells secrete many common Teff cytokines such as IL-4, IFN-γ and IL-17, implying a complex relationship between Tfh cells and other CD4 T effector lineages and highlighting the plasticity of Tfh cells.[112] Further refining these varying cellular populations, e.g. by using single cell transcriptome analyses, will provide insight into T effector cell development, differentiation, and function.[113]

Angioimmunoblastic T-cell lymphoma is an uncommon subtype of peripheral T-cell lymphoma with a poor prognosis.[114] Initial studies revealed a pan T-cell phenotype with expression of Tfh markers including CD10, CXCL13,

and PD-1. Gene expression profiling studies identified a specific molecular signature in AITL, indicating origin from Tfh cells,[20,78-83] as well as overlap with other, less well defined PTCL.[78,115] Recent genomics-based studies have identified a group of common mutations AITL, including mutations in *IDH2*, *RHOA*, *TET2*, and *DNMT3*.[20,21,116-120] These studies have led to reclassification of subtypes of PTCL and have allowed assignment of diagnostic and prognostic significance to these subtypes.[20] Combining detailed transcriptome information provided by RNA-seq with genome wide mutation analyses will allow further refinement of these lymphoma subtypes, allowing better assignment of disease diagnosis and prognosis, as well as revealing novel targets for therapeutic strategies.

Genome wide association studies (GWAS) have identified many single nucleotide polymorphisms (SNPs) enriched in patients with autoimmune diseases, leading to the identification of many disease-associated loci.[121,122] Frequently, specific SNPs are found in more than one autoimmune disorder, in line with the observation that some patients suffer from more than one such disorder, with certain polymorphisms likely contributing to common causality.[123] Most variants identified in GWAS studies are outside coding regions,[124] and are enriched for regulatory and transcriptionally functional SNPs.[125] Because expansion of dysfunctional Tfh cells is a major contributor to systemic autoimmunity, we examined the relationship between Tfh cell enhancers and super enhancers and SNPs identified in GWAS. We found many SNPs linked to autoimmune diseases by GWAS in Tfh cell enhancers, particularly multiple sclerosis, rheumatoid arthritis, and type I diabetes, with several linked to multiple such illnesses, with even more significant enrichment in Tfh cell super enhancers. The challenge now is to translate the linkage of autoimmune SNPs and Tfh cell enhancers and super enhancers to a better understanding of Tfh cell development and differentiation and to determine the functional significance of variants associated with quantitative traits linked to autoimmune disease.

Acknowledgements

Supported in part by grants T32HD007094, R01HL106184, RO1HL65448, T32AR07107, RO1AR040072, R21AR062842, and P30AR053495 from the National Institutes of Health and the Alliance for Lupus Research.

Authorship Contributions

Contributions: JSW designed and performed experiments and analyzed data; KLG designed and performed experiments, YM designed and performed experiments, SC performed experiments, YZ performed experiments, MS designed and performed experiments, VPS analyzed data, JC designed experiments, analyzed data and wrote the manuscript, and PGG designed experiments, analyzed data and wrote the manuscript.

Conflict-of-Interest Disclosure

The authors declare no competing financial interests.

Correspondence

Joseph Craft, Departments of Medicine and Immunobiology, Yale University School of Medicine, 300 Cedar Street, P. O. Box 208031, New Haven, CT 06520-8031. Tel: 203 785 2453. Fax: 203. Email: joseph.craft@yale.edu and Patrick G. Gallagher, Departments of Pediatrics, Pathology, and Genetics, Yale University School of Medicine, 333 Cedar Street, P. O. Box 208064, New Haven, CT 06520-8064. Tel: 203 688 2896. Fax: 203 785 6874. Email: patrick.gallagher@yale.edu.

References

1. Choi YS, Yang JA, Crotty S. Dynamic regulation of Bcl6 in follicular helper CD4 T (Tfh) cells. *Curr Opin Immunol*. 2013;25(3):366-372.

2. Crotty S. Follicular helper CD4 T cells (TFH). *Annu Rev Immunol*. 2011;29:621-663.

3. Weinmann AS. Regulatory mechanisms that control T-follicular helper and T-helper 1 cell flexibility. *Immunol Cell Biol*. 2014;92(1):34-39.

4. Crotty S. The 1-1-1 fallacy. *Immunol Rev*. 2012;247(1):133-142.

5. Johnston RJ, Poholek AC, DiToro D, et al. Bcl6 and Blimp-1 are reciprocal and antagonistic regulators of T follicular helper cell differentiation. *Science*. 2009;325(5943):1006-1010.

6. Ramiscal RR, Vinuesa CG. T-cell subsets in the germinal center. *Immunol Rev*. 2013;252(1):146-155.

7. Vinuesa CG, Sanz I, Cook MC. Dysregulation of germinal centres in autoimmune disease. *Nat Rev Immunol*. 2009;9(12):845-857.

8. Yu D, Rao S, Tsai LM, et al. The transcriptional repressor Bcl-6 directs T follicular helper cell lineage commitment. *Immunity*. 2009;31(3):457-468.

9. Johnston RJ, Poholek AC, DiToro D, et al. Bcl6 and Blimp-1 are reciprocal and antagonistic regulators of T follicular helper cell differentiation. *Science*. 2009:325(5943):1006-10.

10. Nurieva RI, Chung Y, Martinez GJ, et al. Bcl6 mediates the development of T follicular helper cells. *Science*. 2009;325(5943):1001-1005.

11. Ansel KM, McHeyzer-Williams LJ, Ngo VN, McHeyzer-Williams MG, Cyster JG. In vivo-activated CD4 T cells upregulate CXC chemokine receptor 5 and reprogram their response to lymphoid chemokines. *J Exp Med*. 1999;190(8):1123-1134.

12. Haynes NM, Allen CD, Lesley R, Ansel KM, Killeen N, Cyster JG. Role of CXCR5 and CCR7 in follicular Th cell positioning and appearance of a programmed cell death gene-1high germinal center-associated subpopulation. *J Immunol*. 2007;179(8):5099-5108.

13. Bauquet AT, Jin H, Paterson AM, et al. The costimulatory molecule ICOS regulates the expression of c-Maf and IL-21 in the development of follicular T helper cells and TH-17 cells. *Nat Immunol*. 2009;10(2):167-175.

14. Ma CS, Deenick EK. Human T follicular helper (Tfh) cells and disease. *Immunol Cell Biol*. 2014;92(1):64-71.

15. Tangye SG, Ma CS, Brink R, Deenick EK. The good, the bad and the ugly - TFH cells in human health and disease. *Nat Rev Immunol*. 2013;13(6):412-426.

16. Yu D, Vinuesa CG. Multiple checkpoints keep follicular helper T cells under control to prevent autoimmunity. *Cell Mol Immunol*;7(3):198-203.

17. Linterman MA, Rigby RJ, Wong RK, et al. Follicular helper T cells are required for systemic autoimmunity. *J Exp Med*. 2009;206(3):561-576.

18. Gaulard P, de Leval L. Follicular helper T cells: implications in neoplastic hematopathology. *Semin Diagn Pathol*. 2011;28(3):202-213.

19. Huang Y, Moreau A, Dupuis J, et al. Peripheral T-cell lymphomas with a follicular growth pattern are derived from follicular helper T cells (TFH) and may show overlapping features with angioimmunoblastic T-cell lymphomas. *Am J Surg Pathol*. 2009;33(5):682-690.

20. Iqbal J, Wright G, Wang C, et al. Gene expression signatures delineate biologic and prognostic subgroups in peripheral T-cell lymphoma. *Blood*. 2014; 123(19):2915-23.

21. Lemonnier F, Couronne L, Parrens M, et al. Recurrent TET2 mutations in peripheral T-cell lymphomas correlate with TFH-like features and adverse clinical parameters. *Blood*. 2012;120(7):1466-1469.

22. Moroch J, Copie-Bergman C, de Leval L, et al. Follicular peripheral T-cell lymphoma expands the spectrum of classical Hodgkin lymphoma mimics. *Am J Surg Pathol*. 2012;36(11):1636-1646.

23. Buder K, Poppe LM, Brocker EB, et al. Primary cutaneous follicular helper T-cell lymphoma: diagnostic pitfalls of this new lymphoma subtype. *J Cutan Pathol*. 2013;40(10):903-908.

24. Battistella M, Beylot-Barry M, Bachelez H, Rivet J, Vergier B, Bagot M. Primary cutaneous follicular helper T-cell lymphoma: a new subtype of cutaneous T-cell lymphoma reported in a series of 5 cases. *Arch Dermatol*. 2012;148(7):832-839.

25. Fu J, Heinrichs J, Yu XZ. Helper T-cell differentiation in graft-versus-host disease after allogeneic hematopoietic stem cell transplantation. *Arch Immunol Ther Exp (Warsz)*. 2014;62(4):277-301

26. Sankaran VG, Gallagher PG. Applications of high-throughput DNA sequencing to benign hematology. *Blood*. 2013.

27. Garber M, Grabherr MG, Guttman M, Trapnell C. Computational methods for transcriptome annotation and quantification using RNA-seq. *Nat Methods*. 2011;8(6):469-477.

28. McGettigan PA. Transcriptomics in the RNA-seq era. *Curr Opin Chem Biol*. 2013;17(1):4-11.

29. Trapnell C, Williams BA, Pertea G, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. 2010;28(5):511-515.

30. Fromm G, Bulger M. A spectrum of gene regulatory phenomena at mammalian beta-globin gene loci. *Biochem Cell Biol*. 2009;87(5):781-790.

31. Higgs DR, Vernimmen D, Wood B. Long-range regulation of alpha-globin gene expression. *Adv Genet*. 2008;61:143-173.

32. Bulger M, Groudine M. Enhancers: the abundance and function of regulatory sequences beyond promoters. *Dev Biol*. 2010;339(2):250-257.

33. Bulger M, Groudine M. Functional and mechanistic diversity of distal transcription enhancers. *Cell*. 2011;144(3):327-339.

34. Ong CT, Corces VG. Enhancers: emerging roles in cell fate specification. *EMBO Rep*. 2012;13(5):423-430.

35. Borok MJ, Tran DA, Ho MCW, Drewell RA. Dissecting the regulatory switches of development: lessons from enhancer evolution in Drosophila. *Development*. 2010;137(1):5-13.

36. Xu J, Watts JA, Pope SD, et al. Transcriptional competence and the active marking of tissue-specific enhancers by defined transcription factors in embryonic and induced pluripotent stem cells. *Genes Dev*. 2009;23(24):2824-2838.

37. Nord AS, Blow MJ, Attanasio C, et al. Rapid and pervasive changes in genome-wide enhancer usage during mammalian development. *Cell*. 2013;155(7):1521-1531.

38. Ong CT, Corces VG. Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nat Rev Genet*. 2011;12(4):283-293.

39. Core LJ, Lis JT. Paused Pol II captures enhancer activity and acts as a potent insulator. *Genes Dev*. 2009;23(14):1606-1612.

40. Maksimenko O, Golovnin A, Georgiev P. Enhancer-promoter communication is regulated by insulator pairing in a *Drosophila* model bigenic locus. *Mol Cell Biol*. 2008;28(17):5469-5477.

41. Mattick JS. Linc-ing Long noncoding RNAs and enhancer function. *Dev Cell*. 2010;19(4):485-486.

42. Orom UA, Derrien T, Beringer M, et al. Long noncoding RNAs with enhancer-like function in human cells. *Cell*. 2010;143(1):46-58.

43. Orom UA, Shiekhattar R. Long non-coding RNAs and enhancers. *Curr Opin Genet Dev*. 2011;21(2):194-198.

44. Orom UA, Shiekhattar R. Noncoding RNAs and enhancers: complications of a long-distance relationship. *Trends Genet*. 2011;27(10):433-439.

45. Yoo EJ, Cooke NE, Liebhaber SA. An RNA-independent linkage of noncoding transcription to long-range enhancer function. *Mol Cell Biol*. 2012;32(10):2020-2029.

46. Kim T-K, Hemberg M, Gray JM, et al. Widespread transcription at neuronal activity-regulated enhancers. *Nature*. 2010;465(7295):182-7.

47. Wang D, Garcia-Bassets I, Benner C, et al. Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. *Nature*. 2011;474(7351):390-394.

48. Kowalczyk MS, Hughes JR, Garrick D, et al. Intragenic enhancers act as alternative promoters. *Mol Cell*. 2012;45(4):447-458.

49. Guerrero L, Marco-Ferreres R, Serrano AL, Arredondo JJ, Cervera M. Secondary enhancers synergise with primary enhancers to guarantee fine-tuned muscle gene expression. *Dev Biol*. 2010;337(1):16-28.

50. Hong JW, Hendrix DA, Levine MS. Shadow enhancers as a source of evolutionary novelty. *Science*. 2008;321(5894):1314.

51. Krivega I, Dean A. Enhancer and promoter interactions-long distance calls. *Curr Opin Genet Dev*. 2012;22(2):79-85.

52. Zhang Y, Wong CH, Birnbaum RY, et al. Chromatin connectivity maps reveal dynamic promoter-enhancer long-range associations. *Nature*. 2013;504(7479):306-310.

53. Kieffer-Kwon KR, Tang Z, Mathe E, et al. Interactome maps of mouse gene regulatory domains reveal basic principles of transcriptional regulation. *Cell*. 2013;155(7):1507-1520.

54. Gibcus JH, Dekker J. The hierarchy of the 3D genome. *Mol Cell*. 2013;49(5):773-782.

55. Wang Z, Zang C, Rosenfeld JA, et al. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet*. 2008;40(7):897-903.

56. Vahedi G, A CP, Hand TW, et al. Helper T-cell identity and evolution of differential transcriptomes and epigenomes. *Immunol Rev*. 2013;252(1):24-40.

57. Schmidl C, Hansmann L, Lassmann T, et al. The enhancer and promoter landscape of human regulatory and conventional T-cell subpopulations. *Blood*. 2014;123(17):e68-78.

58. Andersson R, Gebhard C, Miguel-Escalada I, et al. An atlas of active enhancers across human cell types and tissues. *Nature*. 2014;507(7493):455-461.

59. Samstein RM, Arvey A, Josefowicz SZ, et al. Foxp3 exploits a pre-existent enhancer landscape for regulatory T cell lineage specification. *Cell*. 2012;151(1):153-166.

60. Vahedi G, Takahashi H, Nakayamada S, et al. STATs shape the active enhancer landscape of T cell populations. *Cell*. 2012;151(5):981-993.

61. Hawkins RD, Hon GC, Lee LK, et al. Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell*. 2010;6(5):479-491.

62. Hawkins RD, Larjo A, Tripathi SK, et al. Global chromatin state analysis reveals lineage-specific enhancers during the initiation of human T helper 1 and T helper 2 cell polarization. *Immunity*. 2013;38(6):1271-1284.

63. Wei G, Wei L, Zhu J, et al. 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.

64. Abraham BJ, Cui K, Tang Q, Zhao K. Dynamic regulation of epigenomic landscapes during hematopoiesis. *BMC Genomics*. 2013;14:193.

65. Roh TY, Wei G, Farrell CM, Zhao K. Genome-wide prediction of conserved and nonconserved enhancers by histone acetylation patterns. *Genome Res*. 2007;17(1):74-81.

66. Lu KT, Kanno Y, Cannons JL, et al. Functional and epigenetic studies reveal multistep differentiation and plasticity of in vitro-generated and in vivo-derived follicular T helper cells. *Immunity*. 2011;35(4):622-632.

67. Su MY, Steiner LA, Bogardus H, et al. Identification of biologically relevant enhancers in human erythroid cells. *J Biol Chem*. 2013;288(12):8433-8444.

68. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545-15550.

69. Steiner LA, Maksimova Y, Schulz V, et al. Chromatin Architecture and Transcription Factor Binding Regulate Expression of Erythrocyte Membrane Protein Genes. *Mol Cell Biol*. 2009; 29(20):5399-412.

70. Karolchik D, Kuhn RM, Baertsch R, et al. The UCSC Genome Browser Database: 2008 update. *Nucleic Acids Res*. 2008;36(Database issue):D773-779.

71. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008;9(9):R137.

72. Siepel A, Haussler D. Combining phylogenetic and hidden Markov models in biosequence analysis. *J Comput Biol*. 2004;11(2-3):413-428.

73. Siepel A, Bejerano G, Pedersen JS, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res*. 2005;15(8):1034-1050.

74. McLean CY, Bristor D, Hiller M, et al. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol*. 2010;28(5):495-501.

75. Hindorff LA, Sethupathy P, Junkins HA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A*. 2009;106(23):9362-9367.

76. Jane SM, Ney PA, Vanin EF, Gumucio DL, Nienhuis AW. Identification of a stage selector element in the human gamma-globin gene promoter that fosters preferential interaction with the 5' HS2 enhancer when in competition with the beta-promoter. *EMBO J*. 1992;11(8):2961-2969.

77. Gallagher PG, Liem RI, Wong E, Weiss MJ, Bodine DM. GATA-1 and Oct-1 are required for expression of the human alpha-hemoglobin-stabilizing protein gene. *J Biol Chem*. 2005;280(47):39016-39023.

78. Iqbal J, Weisenburger DD, Greiner TC, et al. Molecular signatures to improve diagnosis in peripheral T-cell lymphoma and prognostication in angioimmunoblastic T-cell lymphoma. *Blood*. 2010;115(5):1026-1036.

79. Piccaluga PP, Agostinelli C, Califano A, et al. Gene expression analysis of angioimmunoblastic lymphoma indicates derivation from T follicular helper cells and vascular endothelial growth factor deregulation. *Cancer Res*. 2007;67(22):10703-10710.

80. Cuadros M, Dave SS, Jaffe ES, et al. Identification of a proliferation signature related to survival in nodal peripheral T-cell lymphomas. *J Clin Oncol*. 2007;25(22):3321-3329.

81. Martinez-Delgado B, Cuadros M, Honrado E, et al. Differential expression of NF-kappaB pathway genes among peripheral T-cell lymphomas. *Leukemia*. 2005;19(12):2254-2263.

82. Ballester B, Ramuz O, Gisselbrecht C, et al. Gene expression profiling identifies molecular subgroups among nodal peripheral T-cell lymphomas. *Oncogene*. 2006;25(10):1560-1570.

83. de Leval L, Rickman DS, Thielen C, et al. The gene expression profile of nodal peripheral T-cell lymphoma demonstrates a molecular link between angioimmunoblastic T-cell lymphoma (AITL) and follicular helper T (TFH) cells. *Blood*. 2007;109(11):4952-4963.

84. Zentner GE, Scacheri PC. The chromatin fingerprint of gene enhancer elements. *J Biol Chem*. 2012;287(37):30888-30896.

85. Zentner GE, Tesar PJ, Scacheri PC. Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Res*. 2011;21(8):1273-1283.

86. May D, Blow MJ, Kaplan T, et al. Large-scale discovery of enhancers from human heart tissue. *Nat Genet*. 2012;44(1):89-93.

87. Bernstein BE, Kamal M, Lindblad-Toh K, et al. Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell*. 2005;120(2):169-181.

88. Heintzman ND, Hon GC, Hawkins RD, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*. 2009;459(7243):108-12.

89. Ernst J, Kheradpour P, Mikkelsen TS, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature*. 2011;473(7345):43-49.

90. Whyte WA, Orlando DA, Hnisz D, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*. 2013;153(2):307-319.

91. Parker SC, Stitzel ML, Taylor DL, et al. Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. *Proc Natl Acad Sci U S A*. 2013;110(44):17921-17926.

92. Hnisz D, Abraham BJ, Lee TI, et al. Super-enhancers in the control of cell identity and disease. *Cell*. 2013;155(4):934-947.

93. Shlyueva D, Stampfel G, Stark A. Transcriptional enhancers: from properties to genome-wide predictions. *Nat Rev Genet*. 2014;15(4):272-286.

94. King C, Tangye SG, Mackay CR. T follicular helper (TFH) cells in normal and dysregulated immune responses. *Annu Rev Immunol*. 2008;26:741-766.

95. Xu J, Foy TM, Laman JD, et al. Mice deficient for the CD40 ligand. *Immunity*. 1994;1(5):423-431.

96. Glatman Zaretsky A, Taylor JJ, King IL, Marshall FA, Mohrs M, Pearce EJ. T follicular helper cells differentiate from Th2 cells in response to helminth antigens. *J Exp Med*. 2009;206(5):991-999.

97. Zotos D, Coquet JM, Zhang Y, et al. IL-21 regulates germinal center B cell differentiation and proliferation through a B cell-intrinsic mechanism. *J Exp Med*. 2010;207(2):365-378.

98. Linterman MA, Beaton L, Yu D, et al. IL-21 acts directly on B cells to regulate Bcl-6 expression and germinal center responses. *J Exp Med*. 2010;207(2):353-363.

99. Good-Jacobson KL, Szumilas CG, Chen L, Sharpe AH, Tomayko MM, Shlomchik MJ. PD-1 regulates germinal center B cell survival and the formation and affinity of long-lived plasma cells. *Nat Immunol*. 2010;11(6):535-542.

100. Nurieva RI, Chung Y, Martinez GJ, et al. Bcl6 mediates the development of T follicular helper cells. *Science*. 2009;325(5943):1001-1005.

101. Awasthi A, Kuchroo VK. Immunology. The yin and yang of follicular helper T cells. *Science*. 2009;325(5943):953-955.

102. Malissen B. Revisiting the follicular helper T cell paradigm. *Nat Immunol*. 2009;10(4):371-372.

103. Oestreich KJ, Mohn SE, Weinmann AS. Molecular mechanisms that control the expression and activity of Bcl-6 in TH1 cells to regulate flexibility with a TFH-like gene profile. *Nat Immunol*. 2012;13(4):405-411.

104. Bunting KL, Melnick AM. New effector functions and regulatory mechanisms of BCL6 in normal and malignant lymphocytes. *Curr Opin Immunol*. 2013;25(3):339-346.

105. Breitfeld D, Ohl L, Kremmer E, et al. Follicular B helper T cells express CXC chemokine receptor 5, localize to B cell follicles, and support immunoglobulin production. *J Exp Med*. 2000;192(11):1545-1552.

106. Tellier J, Nutt SL. The unique features of follicular T cell subsets. *Cell Mol Life Sci*. 2013;70(24):4771-4784.

107. Chung Y, Tanaka S, Chu F, et al. Follicular regulatory T cells expressing Foxp3 and Bcl-6 suppress germinal center reactions. *Nat Med*. 2011;17(8):983-988.

108. Reinhardt RL, Liang HE, Locksley RM. Cytokine-secreting follicular T cells shape the antibody repertoire. *Nat Immunol*. 2009;10(4):385-393.

109. Chevalier N, Jarrossay D, Ho E, et al. CXCR5 expressing human central memory CD4 T cells and their relevance for humoral immune responses. *J Immunol*. 2011;186(10):5556-5568.

110. Morita R, Schmitt N, Bentebibel SE, et al. Human blood CXCR5(+)CD4(+) T cells are counterparts of T follicular cells and contain specific subsets that differentially support antibody secretion. *Immunity*. 2011;34(1):108-121.

111. Linterman MA, Pierson W, Lee SK, et al. Foxp3+ follicular regulatory T cells control the germinal center response. *Nat Med*. 2011;17(8):975-982.

112. Laurent C, Fazilleau N, Brousset P. A novel subset of T-helper cells: follicular T-helper cells and their markers. *Haematologica*. 2010;95(3):356-358.

113. Linterman MA, Liston A, Vinuesa CG. T-follicular helper cell differentiation and the co-option of this pathway by non-helper cells. *Immunol Rev*. 2012;247(1):143-159.

114. Moskowitz AJ, Lunning MA, Horwitz SM. How I treat the peripheral T-cell lymphomas. *Blood*. 2014;123(17):2636-2644.

115. Piccaluga PP, Agostinelli C, Califano A, et al. Gene expression analysis of peripheral T cell lymphoma, unspecified, reveals distinct profiles and new potential therapeutic targets. *J Clin Invest*. 2007;117(3):823-834.

116. Cairns RA, Iqbal J, Lemonnier F, et al. IDH2 mutations are frequent in angioimmunoblastic T-cell lymphoma. *Blood*. 2012;119(8):1901-1903.

117. Odejide O, Weigert O, Lane AA, et al. A targeted mutational landscape of angioimmunoblastic T-cell lymphoma. *Blood*. 2014;123(9):1293-1296.

118. Palomero T, Couronne L, Khiabanian H, et al. Recurrent mutations in epigenetic regulators, RHOA and FYN kinase in peripheral T cell lymphomas. *Nat Genet*. 2014;46(2):166-170.

119. Sakata-Yanagimoto M, Enami T, Yoshida K, et al. Somatic RHOA mutation in angioimmunoblastic T cell lymphoma. *Nat Genet*. 2014;46(2):171-175.

120. Yoo HY, Sung MK, Lee SH, et al. A recurrent inactivating mutation in RHOA GTPase in angioimmunoblastic T cell lymphoma. *Nat Genet*. 2014;46(4):371-375.

121. Ricano-Ponce I, Wijmenga C. Mapping of immune-mediated disease genes. *Annu Rev Genomics Hum Genet*. 2013;14:325-353.

122. Orru V, Steri M, Sole G, et al. Genetic variants regulating immune cell levels in health and disease. *Cell*. 2013;155(1):242-256.

123. Zhernakova A, van Diemen CC, Wijmenga C. Detecting shared pathogenesis from the shared genetics of immune-related diseases. *Nat Rev Genet.* 2009;10(1):43-55.

124. Maurano MT, Humbert R, Rynes E, et al. Systematic localization of common disease-associated variation in regulatory DNA. *Science*. 2012;337(6099):1190-1195.

125. Schaub MA, Boyle AP, Kundaje A, Batzoglou S, Snyder M. Linking disease associations with regulatory information in the human genome. *Genome Res*. 2012;22(9):1748-1759.

Table 1. Top differentially expressed genes in human Tfh and Teff cells

Gene	Mean Tfh	Mean Teff	Absolute Difference	Log Fold Change	False Discovery Rate
PDCD1	1596.6	276.8	1319.8	2.2	0.0035
CXCR5	1395.1	410.6	984.6	1.6	0.0240
TOX2	1018.7	297.1	721.6	1.6	0.0397
TRIM8	504.6	162.8	341.8	1.6	0.0138
GNG4	335.6	15.6	320.0	3.2	0.0009
NBEAL2	118.8	424.2	305.4	-1.6	0.0217
FAM43A	359.6	74.7	284.9	2.0	0.0066
BCL6	392.7	125.5	267.3	1.5	0.0454
KCNK5	325.5	68.7	256.8	2.1	0.0004
MYO7A	273.0	37.5	235.6	2.4	0.0061
KIAA1324	327.3	123.4	203.9	1.3	0.0468
CCR7	41.2	244.9	203.7	-2.4	0.0007
ATP2B4	93.6	289.9	196.4	-1.6	0.0052
VIM	21.3	205.6	184.3	-2.8	0.0004
CORO1B	274.0	91.4	182.6	1.5	0.0148
AAK1	44.0	223.3	179.3	-2.3	0.0000
IL7R	24.0	195.9	171.9	-2.8	0.0001
MIAT	58.5	219.9	161.4	-1.9	0.0331
PREX1	78.7	232.4	153.7	-1.5	0.0446
GIMAP4	88.3	236.8	148.4	-1.4	0.0494

Table. 2. Top enriched annotations of target genes near human Tfh active enhancers*

Top Enriched Phenotypes	Binomial False Discovery RateQ-Value
abnormal bone marrow cell morphology/development	0
abnormal lymphocyte morphology	0
decreased hematopoietic cell number	1.46E-317
abnormal lymphocyte cell number	5.19E-303
abnormal lymphocyte physiology	3.03E-294
abnormal leukopoiesis	7.28E-289
abnormal mononuclear cell differentiation	1.66E-286
abnormal myeloblast morphology/development	9.13E-286
abnormal lymphopoiesis	2.51E-283
decreased leukocyte cell number	8.16E-282
abnormal T cell morphology	2.85E-281
increased hematopoietic cell number	3.07E-258
decreased lymphocyte cell number	9.68E-258
abnormal T cell differentiation	3.37E-250
abnormal immune serum protein physiology	5.79E-250
increased leukocyte cell number	1.30E-240
abnormal T cell number	1.17E-235
abnormal T cell physiology	5.65E-232
increased lymphocyte cell number	1.07E-212
abnormal B cell morphology	1.37E-201

* Unsupervised enrichment analysis of annotated genes in the proximity of candidate enhancer regions identified by active Tfh cell enhancers. The top enriched Mouse Genome Informatics phenotype ontology terms showing highly significant enrichment of genes implicated in immune cell-related phenotypes are shown. Only terms that showed significant enrichment and had a binomial-fold enrichment of ≥2 were considered.

Table 3. Differentially acetylated active enhancers in Tfh cells associated with differentially expressed genes

Gene	Log Fold Change	False Discovery Rate	Tfh Counts Per Million	Teff Counts Per Million
HAL	4.90	2.18E-06	141.69	3.32
GNG4	3.24	9.13E-04	335.59	15.57
MMP17	3.09	3.11E-04	20.03	2.68
B3GAT1	2.98	2.26E-04	70.71	7.37
CXXC5	2.80	1.08E-03	96.13	8.16
SCGB3A1	2.61	2.44E-02	51.56	9.21
DAB1	2.48	1.09E-02	4.56	0.63
WNK2	2.39	2.82E-02	106.74	17.90
MYO7A	2.36	6.14E-03	273.01	37.46
LYN	2.21	3.42E-02	26.90	3.41
PDCD1	2.16	3.49E-03	1596.55	276.78
XXYLT1	2.14	2.49E-03	149.51	25.76
KCNK5	2.12	4.21E-04	325.55	68.70
KIAA0125	2.11	4.46E-02	6.77	1.34
CLCN4	2.09	2.40E-02	14.70	2.54
CNIH3	2.07	1.03E-03	46.11	10.46
TRPV3	2.04	3.63E-02	13.99	2.44
LOC100132078	2.04	4.48E-02	1.22	0.16
CLTCL1	2.00	4.67E-03	48.44	9.80
TMCC2	1.95	4.43E-03	22.93	5.00
MYBL2	1.94	6.94E-03	39.72	10.34
ZNF703	1.93	5.55E-03	69.98	16.59
FAM167A	1.89	4.93E-02	22.13	5.13
CTTN	1.84	1.90E-02	125.32	27.47
TOX2	1.63	3.97E-02	1018.73	297.15
POU3F1	1.60	4.78E-02	48.69	14.29
SMPD3	1.58	2.53E-02	113.37	36.41
TRIM8	1.57	1.38E-02	504.63	162.83
HIP1	1.55	2.33E-02	31.59	10.13
C11orf75	1.54	4.81E-02	110.97	33.14
CASP9	1.49	1.93E-02	67.74	22.34
BCL6	1.49	4.54E-02	392.74	125.48
KIF2C	1.40	3.81E-02	8.99	3.55
GFOD1	1.36	4.46E-02	65.37	25.20

Figure Legends

Figure 1. Fluorescence-activated cell sorting (FACS). A. Primary human Tfh, Teff, and naïve T cell populations were isolated from tonsil via FACS as shown. **B**. Intracellular staining for Foxp3 was used to identify the percentage of Tfr cells within the Tfh cell population. Representative flow cytometry plots (left) show the percentages of $Foxp3^+$ Tfh or Teff cells in the tonsils using the gating strategy as described in (A), with data quantified from 5 different patients (right).

Figure 2. Transcriptome analyses: Human primary Tfh, Teff, and naïve T cells have distinct expression profiles. **A**. Tfh, Teff, and naïve T cell transcriptomes were obtained by RNA-seq and subjected to multidimensional scaling analysis of expressed genes. Symbols representing 3 or 4 biologic replicates of Tfh (red squares), Teff (blue circles), and naïve T cells (green triangles) clustered together, indicating that samples from each cell type are closely related and distinct from the other cell types. **B.** Venn diagram display of differentially expressed genes. **C.** Heat map display of gene expression patterns of differentially expressed genes. Red represents elevated expression while blue represents decreased expression, compared with the row mean. Each column represents a biologic replicate. Genes displayed in **B** and **C** were selected based on fold changes of 2 or more and FDR adjusted *p* value < .05 between cell types. **D**. RNA coverage profiles of representative differentially expressed genes.

Figure 3. Histone modification density and enhancer class in human primary Tfh cell chromatin. The signal density of H3K4me1, H3K27Ac, H3K27me3, and background total input (TI) chromatin is plotted relative to the H3K4me1 peak. A-C. The average signal over all enhancers in the active, intermediate, and poised enhancer classes, respectively. D. Signal for each enhancer in the active (A), intermediate (I), and poised (P) enhancer classes.

Figure 4. Distribution of histone modifications, and active, intermediate and poised enhancers in human primary Tfh and Teff cell chromatin. The human genome was portioned into seven bins relative to RefSeq genes. The percentage of the human genome represented by each bin was color coded, and the distribution of peaks of each histone modification and enhancer class, active, intermediate and poised, in each bin graphed on the color-coded bar. A. Tfh cells. B. Teff cells. C. K562 cells are included as a non-lymphocyte, hematopoietic cell type for comparison. Abbreviations: TSS: transcriptional start site; TES: transcriptional end site.

Figure 5. Super enhancers in Tfh cells. A. Distribution of H3K27Ac normalized ChIP-seq signal across Tfh cell enhancers. Super enhancers are shown in red. Select super enhancer-associated genes are labeled. B. Representative Tfh cell super enhancers associated with immune-related SNPs at three gene loci, *MAF*, *BATF*, and *IRF4*. The called super enhancer is denoted by the thick blue line at the top of the figure. The associated SNP is shown below the super enhancer line. The track of H3K27 acetylated chromatin is shown above the associated gene locus.

Figure 6. Differential histone 3 lysine 27 acetylation in Tfh and Teff cells. Differentially H3K27 acetylated enhancers in Tfh and Teff cell chromatin were identified. A. The signal density of H3K27Ac was plotted relative to the H3K4me1 peak. A. The average signal over all differentially acetylated enhancers in Tfh and Teff cells is shown. B. The signal for each differentially acetylated enhancer in both cell types.

Figure 7. Tfh cell-type specific active enhancers. A. A differentially acetylated enhancer 5' of the *PDCD1* gene locus in Tfh cells is shown (red bar). Normalized RNA sequencing read density from each cell type at the *PDCD1* gene locus is shown below. **B.** A differentially acetylated enhancer in intron 1 of the *BCL6* gene locus in Tfh cells is shown (red bar). Normalized RNA sequencing read density from each cell type at the *BCL6* gene locus is shown below.

Figure 1A

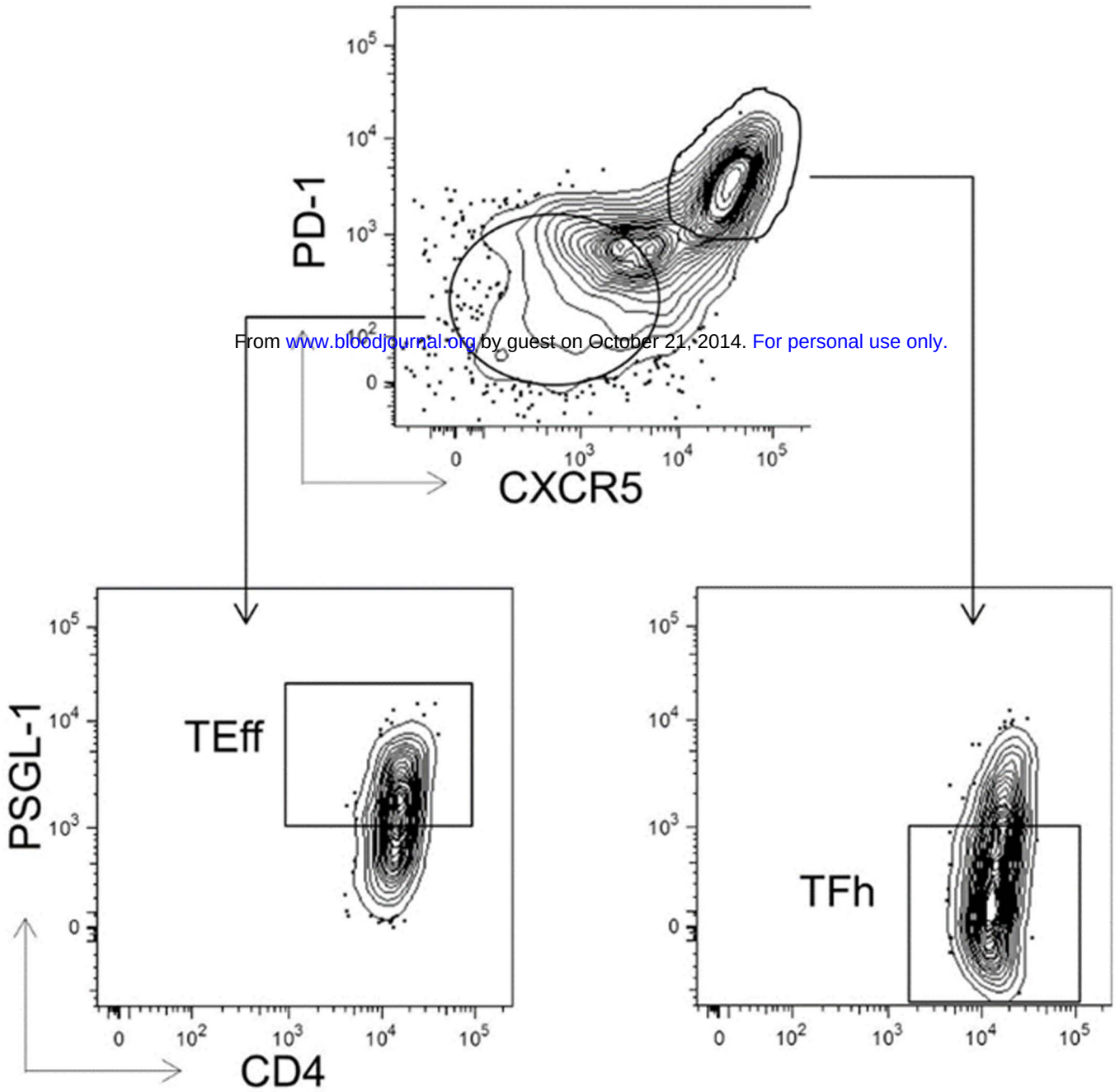
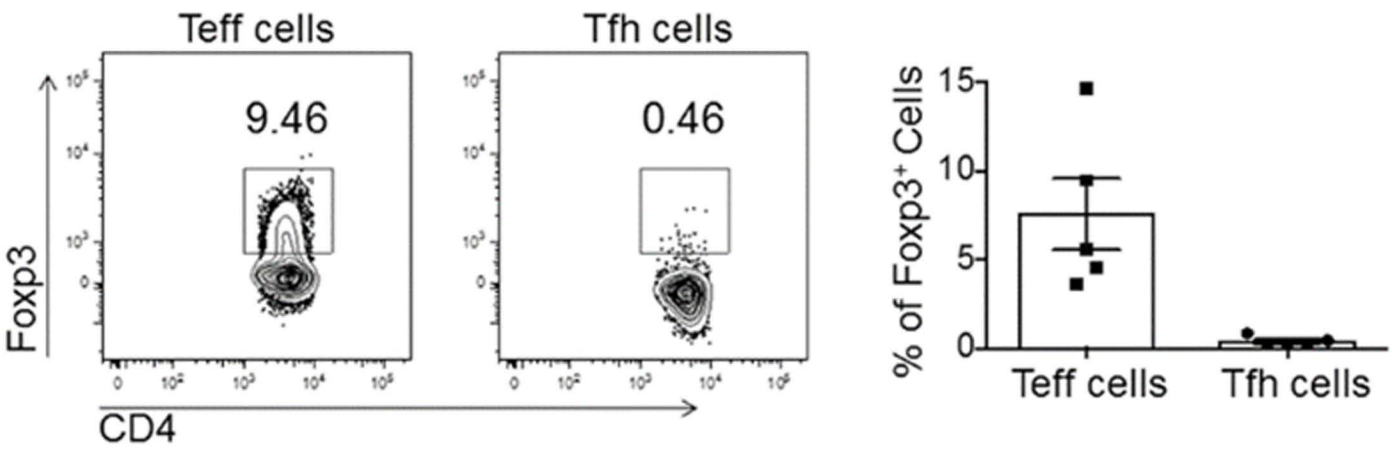
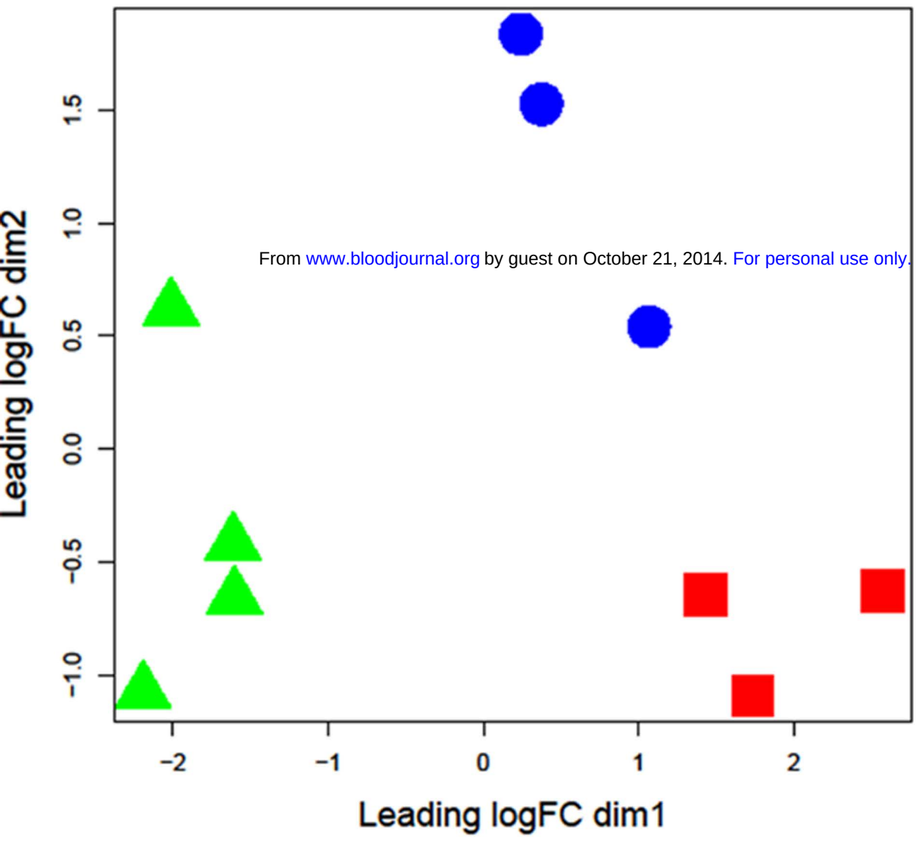


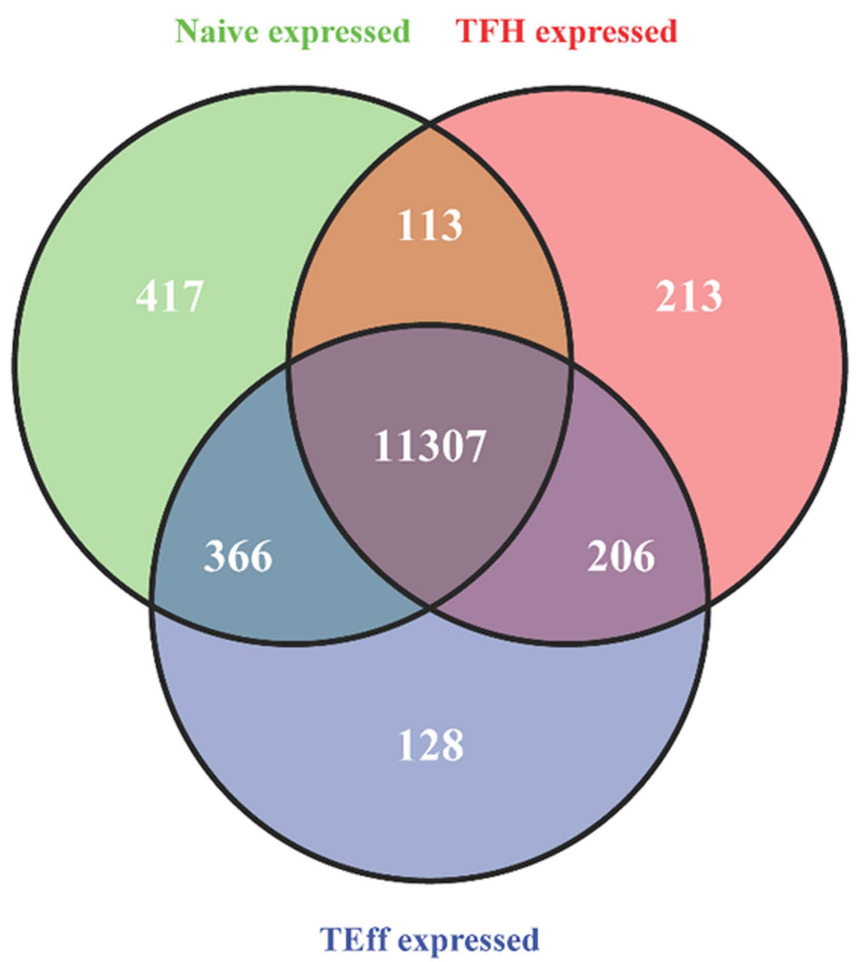
Figure 1B



A



B



C

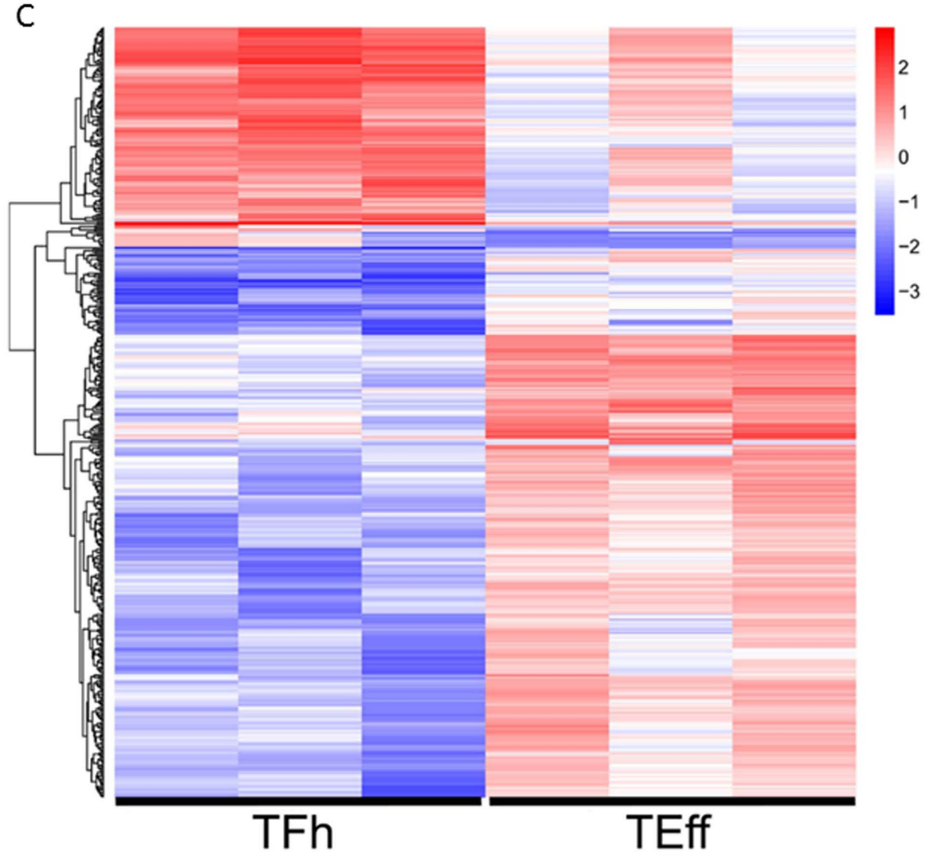


Figure 2

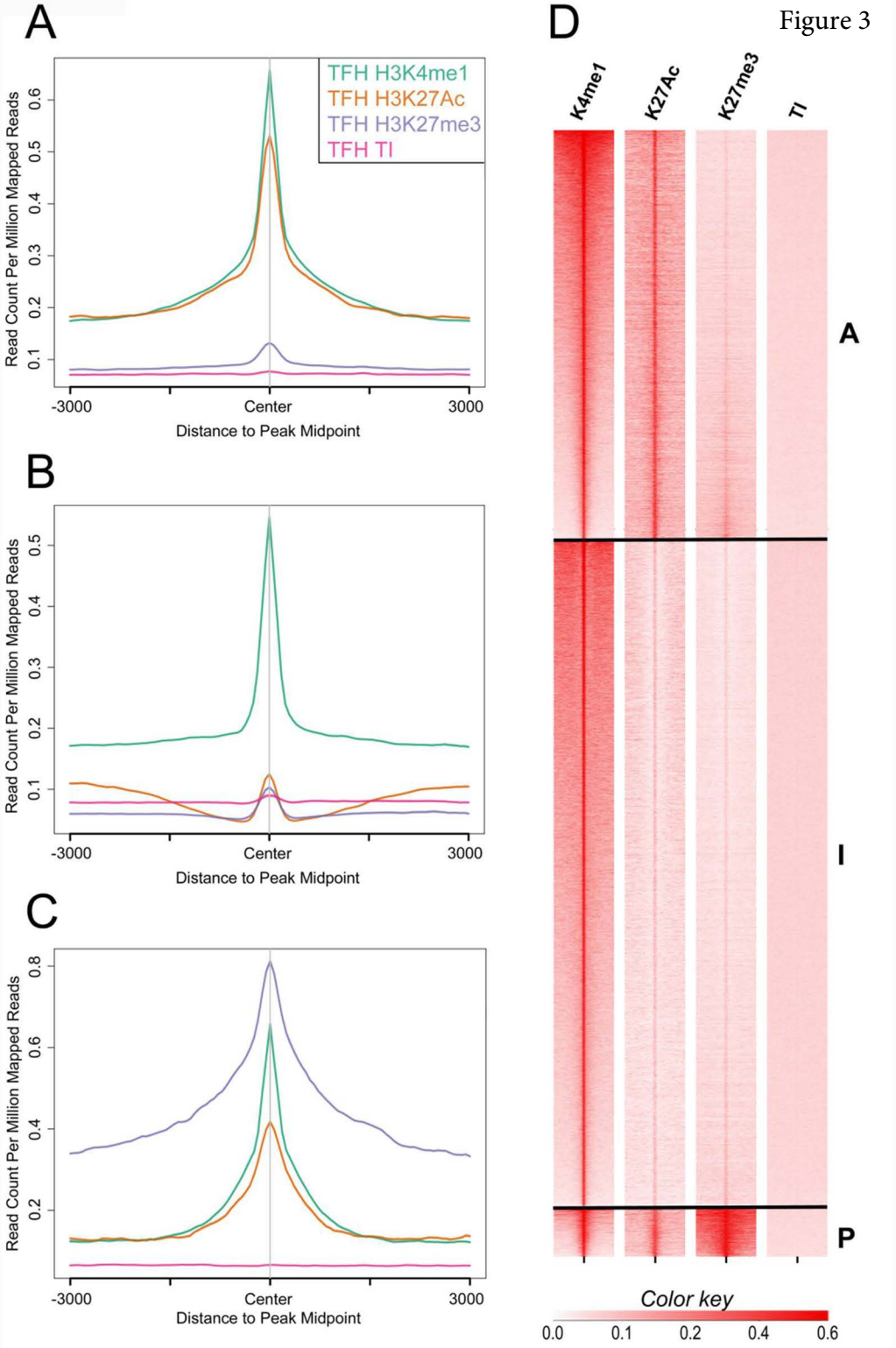
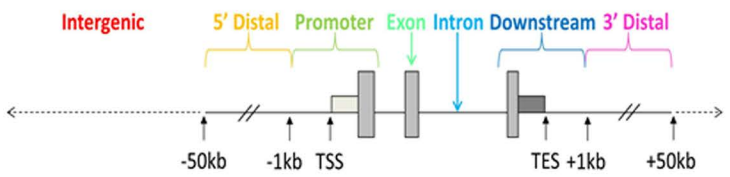
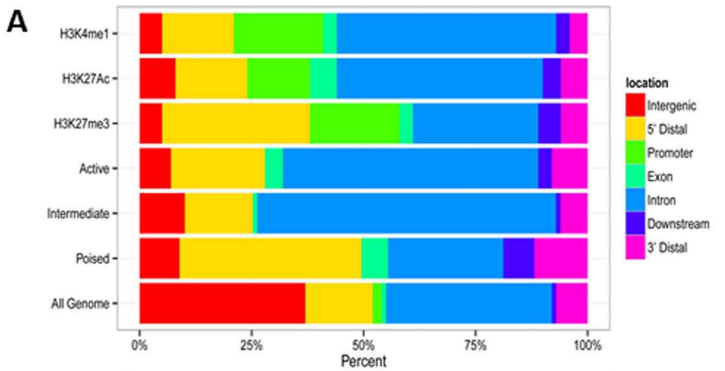
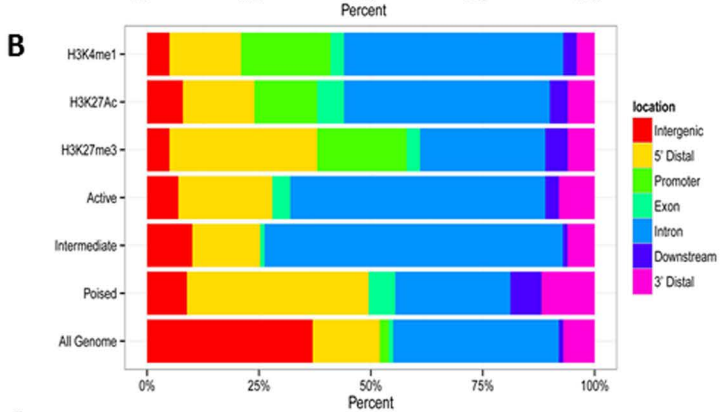

A
TFH H3K4me1
TFH H3K27Ac
TFH H3K27me3
TFH TI
Read Count Per Million Mapped Reads
0.1
0.2
0.3
0.4
0.5
0.6
-3000
Center
3000
Distance to Peak Midpoint
B
Read Count Per Million Mapped Reads
0.1
0.2
0.3
0.4
0.5
-3000
Center
3000
Distance to Peak Midpoint
C
Read Count Per Million Mapped Reads
0.2
0.4
0.6
0.8
-3000
Center
3000
Distance to Peak Midpoint
D
K4me1
K27Ac
K27me3
TI
A
I
P
Color key
0.0
0.1
0.2
0.4
0.6

Figure 4


Intergenic
5' Distal
Promoter
Exon
Intron
Downstream
3' Distal
-50kb
-1kb
TSS
TES
+1kb
+50kb


A
H3K4me1
H3K27Ac
H3K27me3
Active
Intermediate
Poised
All Genome
0%
25%
50%
75%
100%
Percent
location
Intergenic
5' Distal
Promoter
Exon
Intron
Downstream
3' Distal


B
H3K4me1
H3K27Ac
H3K27me3
Active
Intermediate
Poised
All Genome
0%
25%
50%
75%
100%
Percent
location
Intergenic
5' Distal
Promoter
Exon
Intron
Downstream
3' Distal

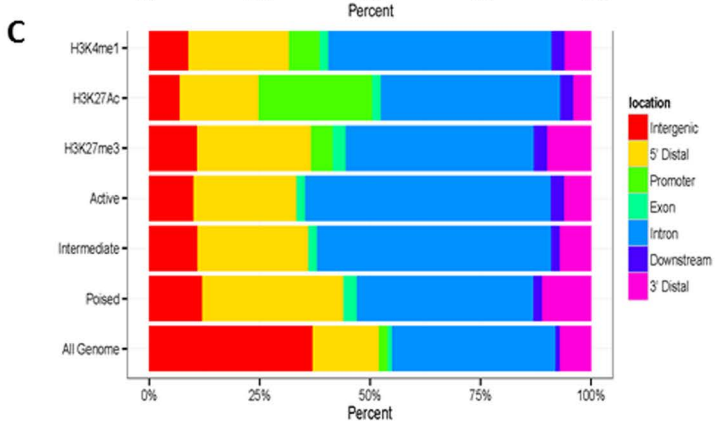

C
H3K4me1
H3K27Ac
H3K27me3
Active
Intermediate
Poised
All Genome
0%
25%
50%
75%
100%
Percent
location
Intergenic
5' Distal
Promoter
Exon
Intron
Downstream
3' Distal

Figure 5A

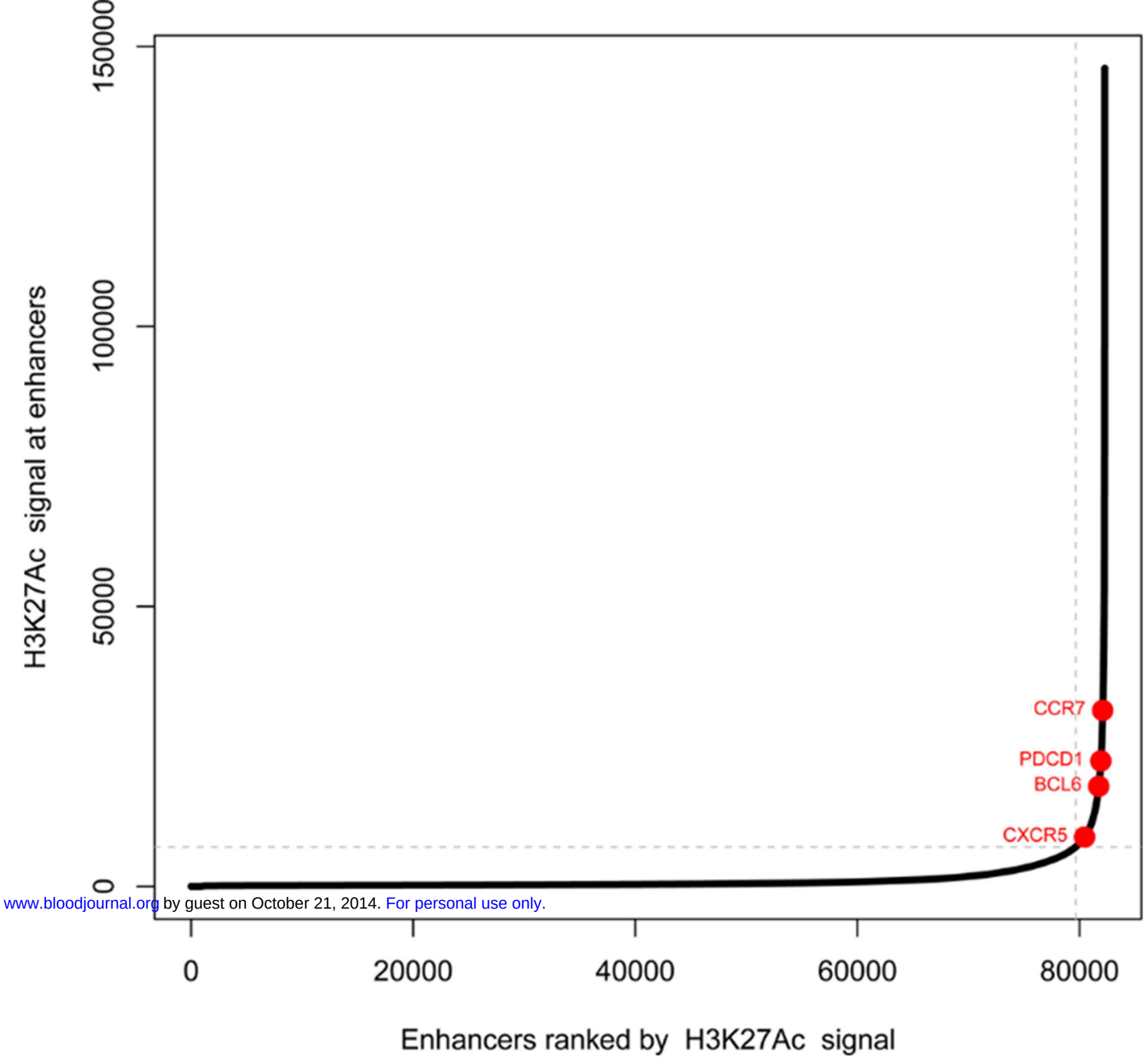


Figure 5B

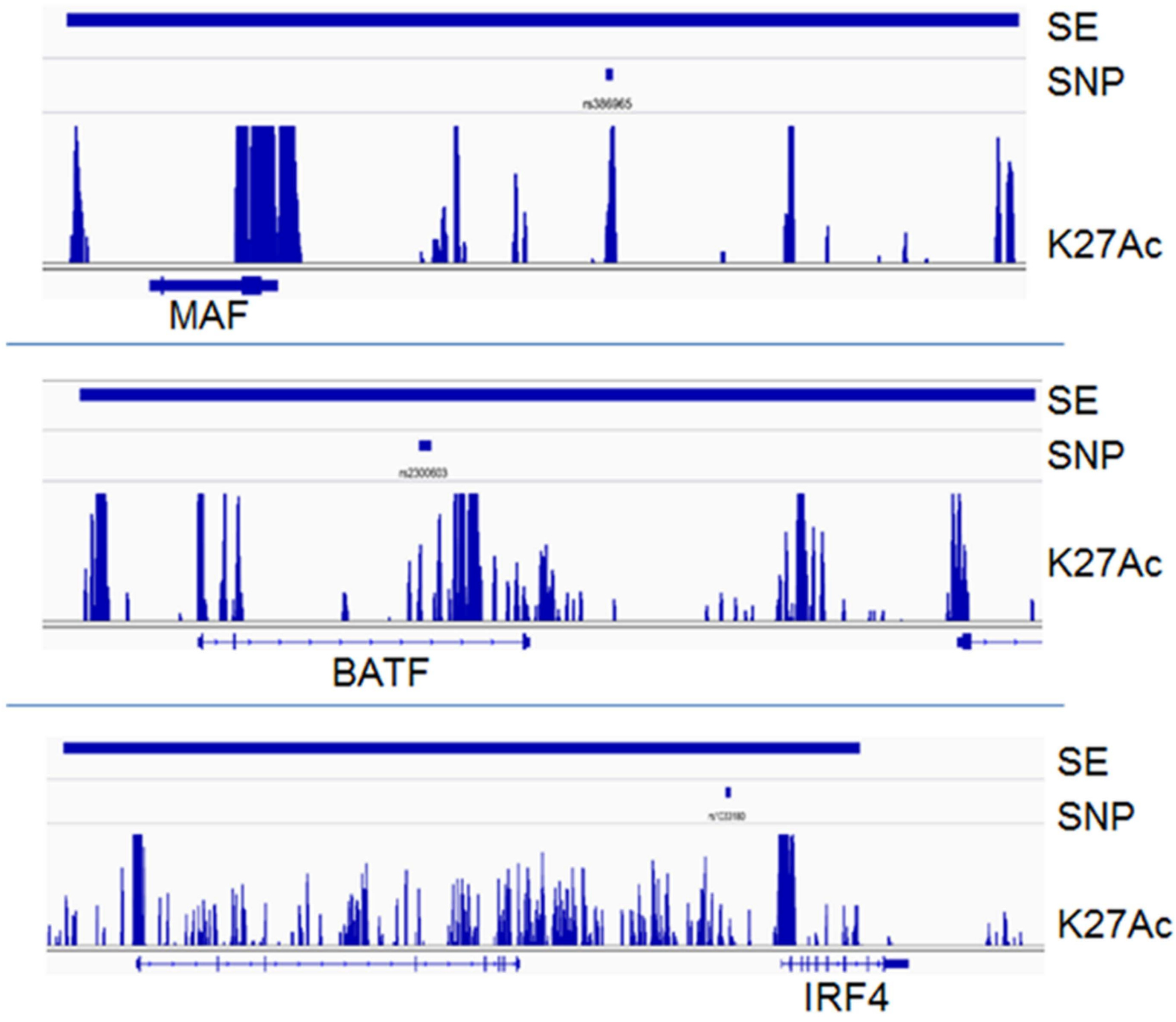


Figure 6

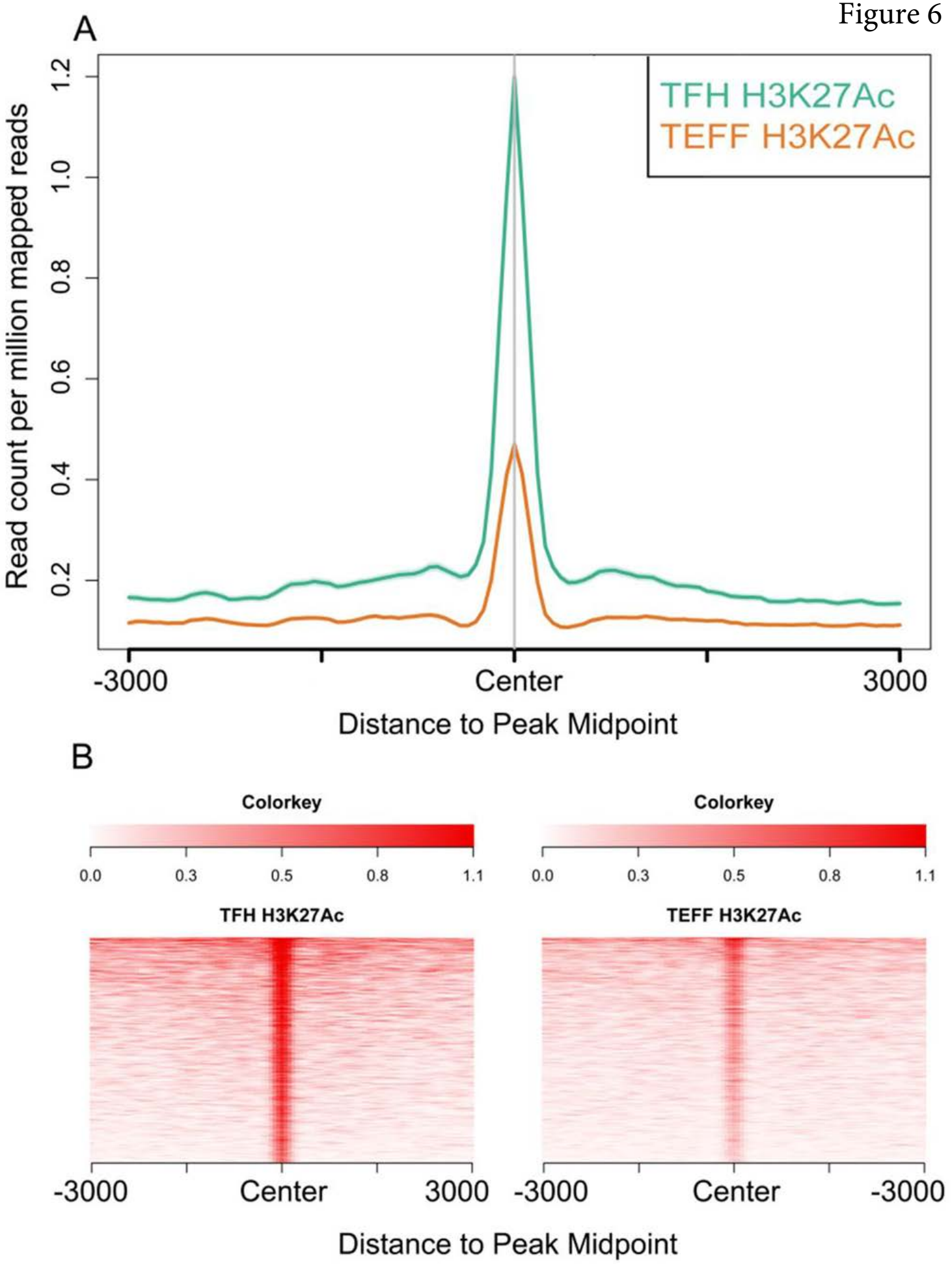
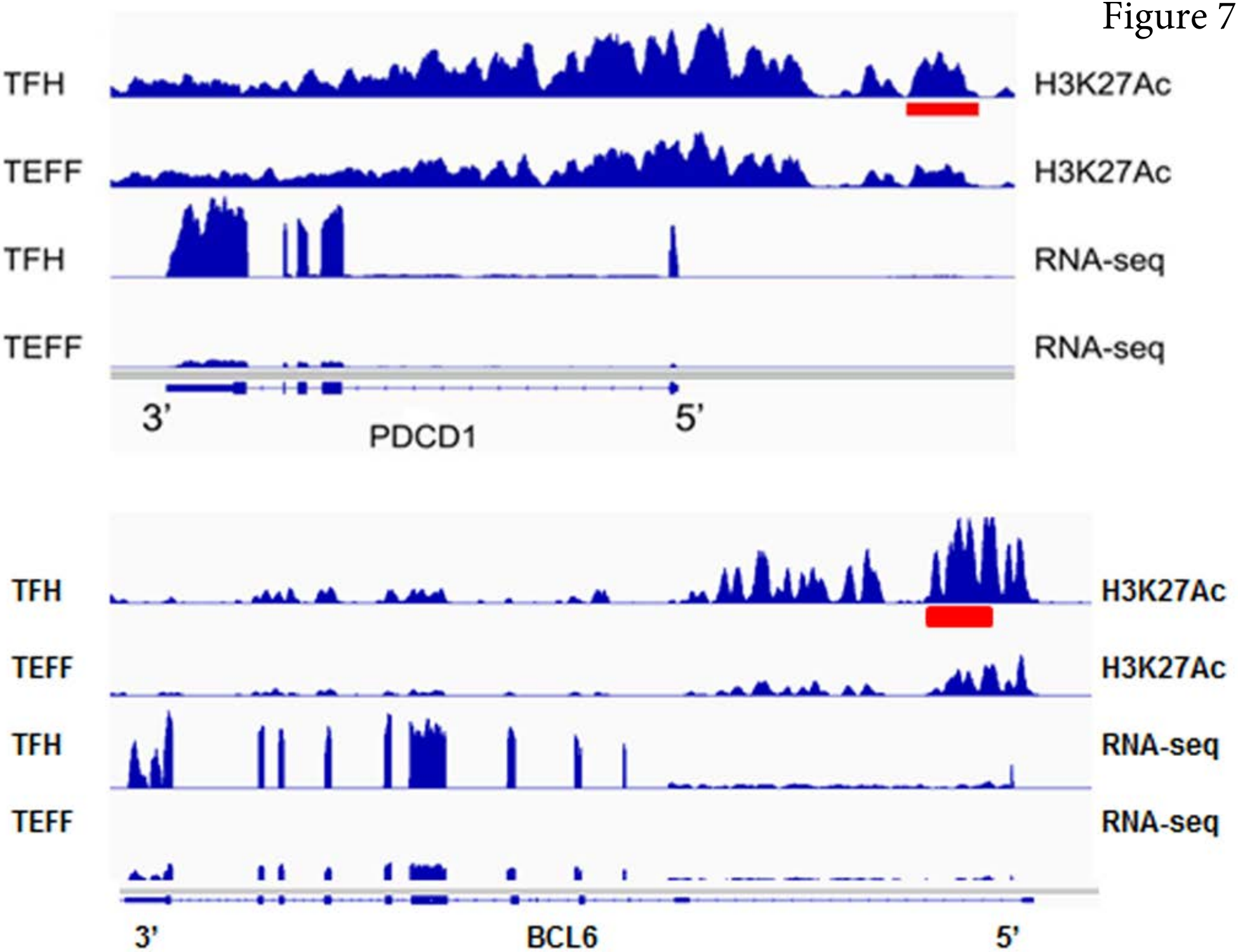

A
TFH H3K27Ac
TEFF H3K27Ac
Read count per million mapped reads
0.2
0.4
0.6
0.8
1.0
1.2
-3000
Center
3000
Distance to Peak Midpoint
B
Colorkey
0.0
0.3
0.5
0.8
1.1
TFH H3K27Ac
TEFF H3K27Ac
-3000
Center
3000
-3000
Center
-3000
Distance to Peak Midpoint

Figure 7


TFH
H3K27Ac
TEFF
H3K27Ac
TFH
RNA-seq
TEFF
RNA-seq
3'
PDCD1
5'
TFH
H3K27Ac
TEFF
H3K27Ac
TFH
RNA-seq
TEFF
RNA-seq
3'
BCL6
5'