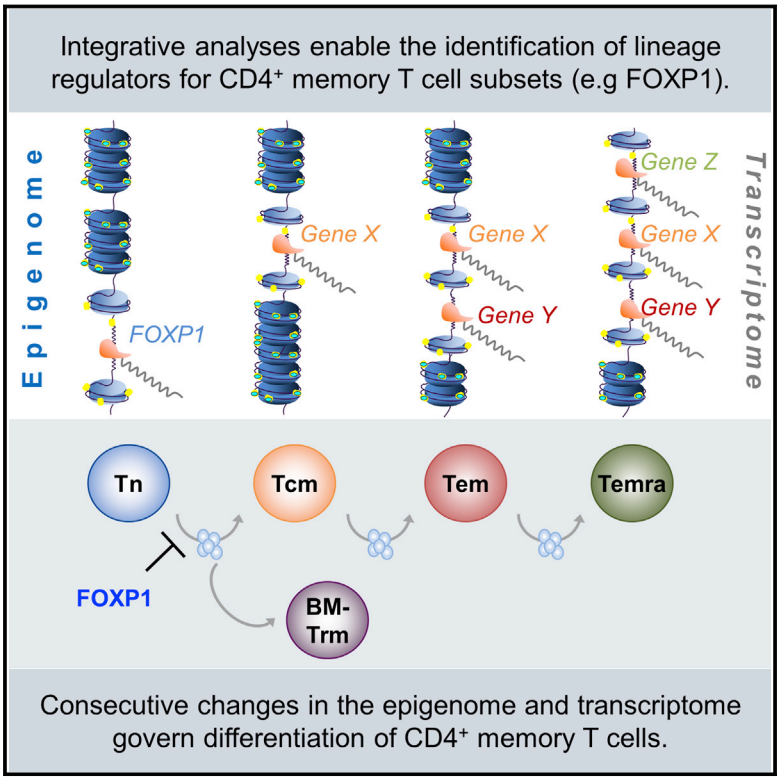
Resource

Immunity

Epigenomic Profiling of Human CD4+ T Cells Supports a Linear Differentiation Model and Highlights Molecular Regulators of Memory Development

Graphical Abstract



Highlights

- Comprehensive epigenomes for human CD4+ T memory subsets generated and analyzed
- Integrative analyses support a linear model of memory T cell differentiation
- Epigenetic control of transcriptional regulators of memory differentiation revealed
- Chromatin changes highlight novel regulators for T memory cell differentiation

Authors

Pawel Durek, Karl Nordström, Gilles Gasparoni, ..., Jörn Walter, Alf Hamann, Julia K. Polansky

Correspondence

julia.polansky@drfz.de

In Brief

As part of the IHEC consortium, Durek et al. (2016) generated deep epigenomes and transcriptomes of CD4+ memory T cell subsets to infer their lineage relationships and to demonstrate the impact of epigenetic regulation on known and novel molecular regulators involved in memory generation. Explore the Cell Press IHEC webportal at www.cell.com/consortium/IHEC.

Durek et al., 2016, Immunity *45*, 1148–1161
November 15, 2016
http://dx.doi.org/10.1016/j.immuni.2016.10.022

CellPress

CellPress

Immunity
Resource

Epigenomic Profiling of Human CD4$^+$ T Cells Supports a Linear Differentiation Model and Highlights Molecular Regulators of Memory Development

Pawel Durek,[1,23] Karl Nordström,[2,23] Gilles Gasparoni,[2,23] Abdulrahman Salhab,[2,23] Christopher Kressler,[1,23] Melanie de Almeida,[1,23] Kevin Bassler,[3] Thomas Ulas,[3] Florian Schmidt,[4,5] Jieyi Xiong,[6] Petar Glažar,[7] Filippos Klironomos,[7] Anupam Sinha,[8] Sarah Kinkley,[9] Xinyi Yang,[9] Laura Arrigoni,[10] Azim Dehghani Amirabad,[4,5] Fatemeh Behjati Ardakani,[4,5] Lars Feuerbach,[11] Oliver Gorka,[12] Peter Ebert,[4] Fabian Müller,[4] Na Li,[9] Stefan Frischbutter,[1] Stephan Schlickeiser,[13] Carla Cendon,[14] Sebastian Fröhler,[6] Bärbel Felder,[15] Nina Gasparoni,[2] Charles D. Imbusch,[11] Barbara Hutter,[11] Gideon Zipprich,[15] Yvonne Tauchmann,[16] Simon Reinke,[17] Georgi Wassilew,[18] Ute Hoffmann,[1] Andreas S. Richter,[10] Lina Sieverling,[11] DEEP Consortium,[19] Hyun-Dong Chang,[14] Uta Syrbe,[20] Ulrich Kalus,[16] Jürgen Eils,[15] Benedikt Brors,[11] Thomas Manke,[10] Jürgen Ruland,[12,21,22] Thomas Lengauer,[4] Nikolaus Rajewsky,[7] Wei Chen,[6] Jun Dong,[14] Birgit Sawitzki,[13] Ho-Ryun Chung,[9] Philip Rosenstiel,[8] Marcel H. Schulz,[4,5] Joachim L. Schultze,[3] Andreas Radbruch,[14] Jörn Walter,[2] Alf Hamann,[1] and Julia K. Polansky[1,24,*]

[1]Experimental Rheumatology, German Rheumatism Research Centre, 10117 Berlin, Germany
[2]Department of Genetics, University of Saarland, 66123 Saarbrücken, Germany
[3]Life and Medical Sciences Institute, Genomics and Immunoregulation, University of Bonn, 53115 Bonn, Germany
[4]Department of Computational Biology and Applied Algorithmics, Max Planck Institute for Informatics, 66123 Saarbrücken, Germany
[5]Excellence Cluster on Multimodal Computing and Interaction, Saarland University, 66123 Saarbrücken, Germany
[6]Berlin Institute for Medical Systems Biology, Max-Delbrück Center for Molecular Medicine, 13125 Berlin, Germany
[7]Systems Biology of Gene Regulatory Elements, Max-Delbrück Center for Molecular Medicine, 13125 Berlin, Germany
[8]Institute of Clinical Molecular Biology, Christian-Albrechts-University, 24105 Kiel, Germany
[9]Otto Warburg Laboratories: Epigenomics at Max Plank Institute for Molecular Genetics, 14195 Berlin, Germany
[10]Max Planck Institute of Immunobiology and Epigenetics, 78108 Freiburg, Germany
[11]Applied Bioinformatics, Deutsches Krebsforschungszentrum, 59120 Heidelberg, Germany
[12]Institute for Clinical Chemistry and Pathobiochemistry, Klinikum rechts der Isar, Technical University 81675 Munich, Germany
[13]Institute of Medical Immunology, Charité University Medicine, 13353 Berlin, Germany
[14]Cell Biology, German Rheumatism Research Centre, 10117 Berlin, Germany
[15]Data Management and Genomics IT, Deutsches Krebsforschungszentrum, 69120 Heidelberg, Germany
[16]Institut für Transfusionsmedizin, Charité University Medicine, 12203 Berlin, Germany
[17]Berlin-Brandenburg Center for Regenerative Therapies, 13353 Berlin, Germany
[18]Center for Musculoskeletal Surgery, Charité University Medicine, 10117 Berlin, Germany
[19]http://www.deutsches-epigenom-programm.de/
[20]Medizinische Klinik für Gastroenterologie, Infektiologie und Rheumatologie, Charité University Medicine, 12000 Berlin, Germany
[21]German Cancer Consortium (DKTK), 59120 Heidelberg, Germany
[22]German Center for Infection Research (DZIF), partner site 81675 Munich, Germany
[23]Co-first authors
[24]Lead Contact
*Correspondence: julia.polansky@drfz.de
http://dx.doi.org/10.1016/j.immuni.2016.10.022

SUMMARY

The impact of epigenetics on the differentiation of memory T (Tmem) cells is poorly defined. We generated deep epigenomes comprising genome-wide profiles of DNA methylation, histone modifications, DNA accessibility, and coding and non-coding RNA expression in naive, central-, effector-, and terminally differentiated CD45RA$^+$ CD4$^+$ Tmem cells from blood and CD69$^+$ Tmem cells from bone marrow (BM-Tmem). We observed a progressive and proliferation-associated global loss of DNA methylation in heterochromatic parts of the genome during Tmem cell differentiation. Furthermore, distinct gradually changing signatures in the epigenome and the transcriptome supported a linear model of memory development in circulating T cells, while tissue-resident BM-Tmem branched off with a unique epigenetic profile. Integrative analyses identified candidate master regulators of Tmem cell differentiation, including the transcription factor FOXP1. This study highlights the importance of epigenomic changes for Tmem cell biology and demonstrates the value of epigenetic data for the identification of lineage regulators.

INTRODUCTION

CD4$^+$ T helper (Th) cells orchestrate the quality and quantity of an adaptive immune reaction and contribute to immunity by

generating a pool of long-lived memory (Tmem) cells, which arise from naive T (Tn) cells after activation by primary antigen encounter. Tmem cells are per se resting, almost non-dividing cells, which can be subdivided into subpopulations based on marker expression, tissue localization and functional properties. Central memory (Tcm) cells appear most similar to Tn cells with respect to their ability to recirculate through blood and lymphoid tissues, the limited effector cytokine commitment, and their high proliferative capacity (Sallusto et al., 1999). In contrast, T effector memory (Tem) cells preferentially home to peripheral tissues and show commitment for the selective production of effector cytokine panels (e.g., IFN-γ, IL-4, and IL-17) characteristic of their functional subtype (Th1, Th2, and Th17, respectively). Their capacity to expand and differentiate is more limited than that of Tcm cells—a feature also found for the so far poorly characterized $CD4^+$ terminally differentiated $CD45RA^+$ memory (Temra) cells (Henson et al., 2012), which feature expression of selected markers of Tn cells (e.g., CD45RA). In addition to these populations circulating through the blood, recent studies have highlighted the importance of tissue-resident memory cells (Carbone et al., 2013; Schenkel and Masopust, 2014). $CD4^+$ Tmem cells from the bone marrow (BM-Tmem) have been shown to constitute a major part of long-term memory in mouse and man (Okhrimenko et al., 2014; Tokoyoda et al., 2009).

The developmental relationship of Tmem cell subsets is not well defined. The question whether different Tmem subtypes represent stages in a sequential linear differentiation process, or whether they branch into different sublineages from early activation stages is still a subject of controversy (Ahmed et al., 2009; Flossdorf et al., 2015; Harrington et al., 2008; Kaech and Cui, 2012). Similarly, master regulators controlling the transit from naive to memory stages, particularly in the human system, are largely unknown, partially due to the lack of suitable experimental systems.

Epigenetic mechanisms play a key role in cell differentiation by controlling expression programs that are stable over time and through cellular generations and hence are prime candidates for the imprinting of stable, heritable expression profiles. Because Tmem cells do not revert to the naive stage, their cellular program seems to be permanently switched, pointing toward epigenetic regulation. Main players in epigenetic regulation are DNA methylation (DNA-meth), histone modifications, and non-coding RNAs, which together direct the rearrangement of the chromatin to promote or to prevent expression of the affected genes. Genome-wide analysis of such epigenetic marks therefore allows for conclusions not only on the current gene expression status but also facilitates insights into the history and the future potential of cells. To date only a few studies on mouse Tmem cells have been published, reporting limited datasets (Crompton et al., 2016; Hashimoto et al., 2013; Komori et al., 2015; Russ et al., 2014). A deep and systematic genome-wide analysis of the epigenetic landscape during human $CD4^+$ Tmem cell differentiation is currently lacking.

As part of the International Human Epigenome Consortium (IHEC) and the German Epigenome Programme (DEEP), we generated comprehensive epigenomic maps of ex vivo isolated isogenic human $CD4^+$ Tn cells and several Tmem cell subsets from the blood and the bone marrow to address the question of whether and how the epigenome contributes to the formation, maintenance, and function of Tmem cell populations in humans. Our data support a model of linear differentiation for circulating human Tmem cells—a topic so far studied only in the murine system. In addition, we find that many known molecular regulators of Tmem cells are under epigenetic control and that epigenetic changes point to novel regulator candidates, which are likely to be involved in Tmem cell differentiation.

RESULTS

Generation of Genome-Wide Epigenetic Datasets of Human $CD4^+$ Tn Cells and Tmem Cell Subsets

To generate comprehensive epigenomic datasets (i.e., class I epigenomes according to the IHEC standards) for the key differentiation stages of human $CD4^+$ Th cells, we sorted $CD4^+$ Tn, Tcm, Tem, and Temra cells from the peripheral blood of healthy human donors by flow cytometry (Figure S1A). To obtain sufficient cell numbers for the subsequent analyses and to mitigate potential inter-donor variations, we used pooled samples of 3–10 female donors (Table S1). For Tn, Tcm, and Tem, analyses of all epigenetic parameters within one replicate were carried out in parallel, i.e., were derived from the same genetic donor pool and therefore represent isogenic samples. For each sample we determined (1) genome-wide DNA-meth profiles, by whole-genome bisulfite sequencing (WBGS) or by reduced representation bisulfite sequencing (RRBS), (2) DNA accessibility maps by nucleosome occupancy and methylome sequencing (NOMe-seq), (3) high-resolution histone modification maps (by Chromatin Immune-Precipitation sequencing, ChIP-seq) for H3K4me1 (= mono-methylation of lysine 4 on histone 3), H3K4me3, H3K9me3, H3K27ac, H3K27me3, and H3K36me3 and, (4) transcriptomes for total RNA (depleted from ribosomal RNAs), messenger RNAs (mRNAs), long non-coding RNAs (lncRNAs), micro RNAs (miRNA), and circular RNAs (circRNA) by deep sequencing of three different RNA libraries (polyadenylated RNAs, small RNAs, and total RNAs depleted from ribosomal RNAs). A selection of these datasets (Figure S1B) was generated for $CD4^+$ BM-Tmem cells, which were separated into the $CD69^+$ tissue-resident and the circulating $CD69^-$ subsets (Figure S1A).

Progressive Segmented Loss of DNA-Meth Correlates with Tmem Cell Differentiation

We profiled the DNA-meth landscape in Tn, Tcm, Tem, and Temra cells using WGBS and observed a strong progressive loss of DNA-meth in the order Tn-Tcm-Tem-Temra with mean methylation levels for the entire genome dropping from 84% in Tn to 67% in Temra (Figure 1A). Loss of methylation predominantly occurred in large domains of up to several hundreds of kilobases (kb), which were decorated with the repressive histone marks H3K27me3 and H3K9me3 (Figure 1B and Figure S2A). Such regions are referred to as "partially methylated domains, PMDs" and can be identified using established software packages ("MethylSeekR," Burger et al., 2013). PMDs contrasted with broad regions that were uniformly fully methylated ("fully methylated regions," FMRs, by MethylSeekR) and to peaks of strong consistent de-methylation typically found in CpG islands

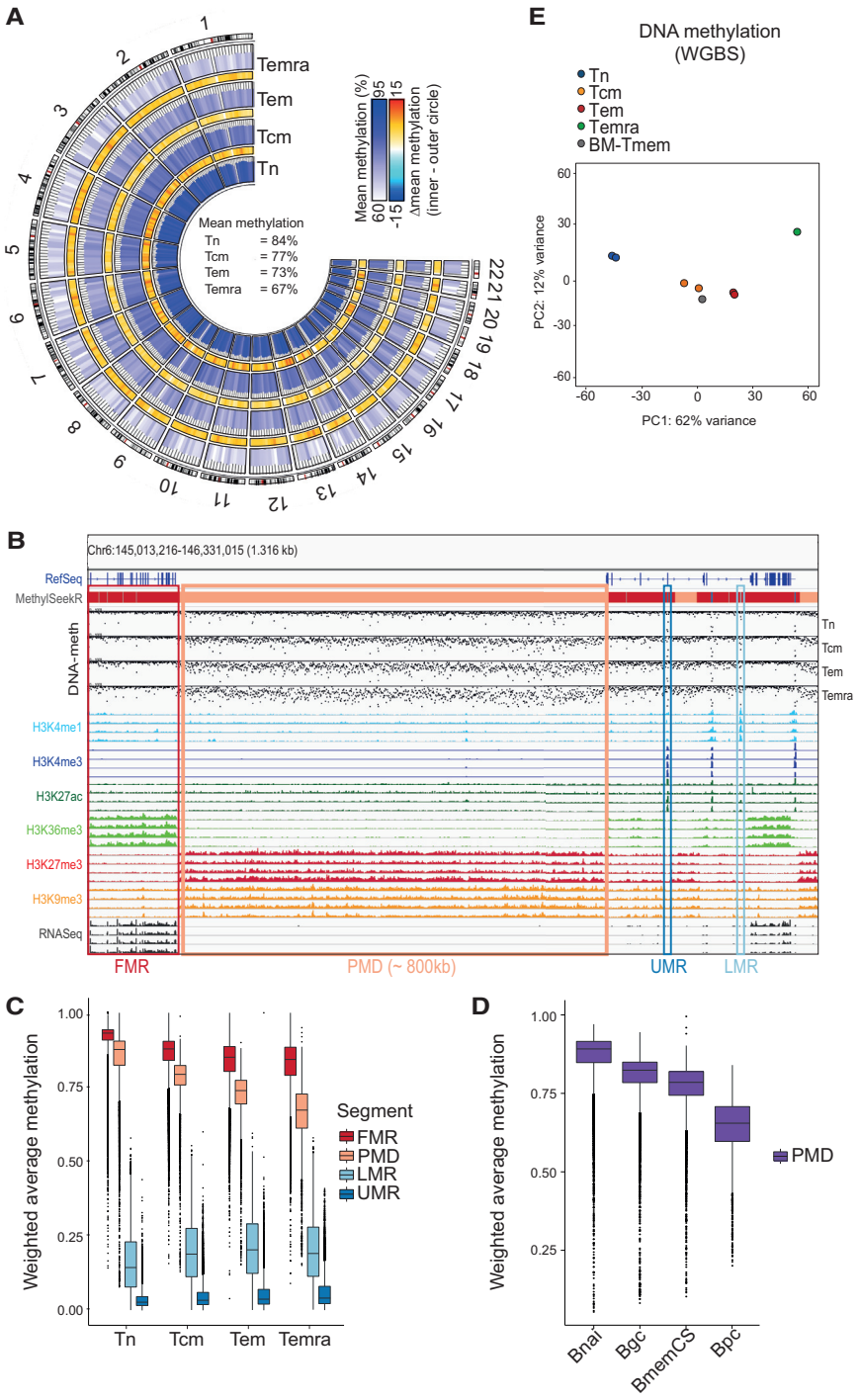


Figure 1. Global Loss of DNA-Meth in Tmem Cells Occurs in Large Heterochromatic Regions

(A) Circle plots of WGBS data (Tn, Tcm, Tem of donor pool Hf03, and Temra, Hf05) are shown. Mean methylation levels in 10 Mb blocks are depicted as color-coded (white-blue) bars. Heatmaps (blue-red) indicate the methylation difference between the adjacent subsets. Total mean methylation for each cell type is given in the center.

(B) Exemplary genomic view of Tn, Tcm, Tem (Hf03 samples), and Temra (Hf05), displaying examples of the genomic segments called by the MethylSeekR software from WGBS data (indicated with boxes): PMD, partially methylated domain; FMR, fully methylated region; LMR, low methylated region; UMR, unmethylated region. The following tracks are shown (top to bottom, each for Tn, Tcm, Tem, Temra): Genes annotated in RefSeq; MethylSeekR-segments; DNA-meth (WGBS); 6 indicated histone modifications; total RNA.

(C) Weighted average DNA-meth across the MethylSeekR segments.

(D) Weighted average methylation across PMDs in B cells (data from the BLUEPRINT project, see Accession codes in Experimental Procedures). Bnai, naive B cells (ERX625136); Bgc, geminal center B cells (ERX715129); BmemCS, class-switched memory B cells (ERX625127); Bpc, plasma cells (ERX301127).

(E) PCA of DNA-meth data (based on WGBS). CpGs with min. coverage = 5 were considered; only CpGs with calls in all indicated samples were used.

("unmethylated regions," UMRs) and transcriptional control regions (e.g., CpG-low promoters and enhancers, "low methylated regions," LMRs). PMDs showed the strongest loss of methylation of all MethylSeekR segments (Figure 1C) and covered up to 67% of the genome (in Tem cells; Figure S2B). Hence, PMDs were responsible for the bulk of the observed global DNA de-methylation in Tmem cell populations. PMD-associated genes generally showed low expression levels compared to FMR-associated genes and were fewer in number (Figure S2C).

We observed a similar segmented loss of global DNA-meth when re-examining DNA-meth profiles from B cells published by the BLUEPRINT consortium (Kulis et al., 2015). Here too, PMDs were the genomic segments that displayed progressive loss of DNA-meth with differentiation into memory B cells and antibody-secreting plasma cells (Figure 1D), indicating that this phenomenon is shared during lymphocyte development.

In a principal-component analysis (PCA), the blood-derived T cell subsets were placed along the main principal component 1 (PC1), in the order Tn-Tcm-Tem-Temra (Figure 1E), which mirrors the DNA de-methylation in PMDs. Temra cells fell at the extreme position along PC1 in relation to Tn cells, suggesting that they are the most differentiated population. However, their inter-donor pool variation was larger compared to other cell types (Figure S2D). In contrast to Temra cells, BM-Tmem cells took an "intermediate" position on PC1 close to circulating Tcm and Tem cells, indicating that their epigenetic imprint toward terminal differentiation is less pronounced (Figure 1E).

These data show that DNA de-methylation in heterochromatic parts of the genome accompanies Tmem cell differentiation in the order of Tn-Tcm-Tem-Temra with BM-Tmem cells clustering with the Tcm and Tem cell populations.

Comprehensive Transcriptome Analyses Reveal a Progressive Change with Tmem Cell Differentiation in the Order of Tn-Tcm-Tem-Temra

We generated full transcriptomes by RNA-seq and determined expression profiles for total RNA, mRNAs, miRNAs, lncRNA, and circRNA for Tn cells and Tmem subsets. Our analysis identified previously described RNAs, as well as previously unknown RNAs (including 981 novel miRNAs, 173 lncRNAs, and 4,826 candidate circRNAs) and many differentially expressed RNAs between the T cell subtypes (Tables S2–S4).

We performed PCA on each of these functionally independent RNA species. Our analysis revealed a consistent pattern with respect to the main component PC1: for all RNA species, the cell types fell along this axis in the strict order of Tn-Tcm-Tem-Temra (Figure 2A). As observed for DNA-meth, BM-Tmem cells took an intermediate position close to Tcm and Tem cells from the blood rather than resembling the most terminally differentiated Temra cell population. For total RNA and lncRNAs, PC2 indicated properties of Tn that were recapitulated in Temra cells and distinguished them from the other memory subsets. Inter-donor pool differences were generally small, except in the miRNA datasets, in which one donor pool became separated by PC2 from all others. Thus, the consistent arrangement of the T cell subtypes on PC1 for all RNA species indicated a progressive change of the transcriptome during Tmem cell differentiation (Tn-Tcm-Tem-Temra).

To validate this further, we performed a co-expression network analysis using the Tn, Tcm, and Tem samples and focused on the 700 most variable genes or on transcriptional regulators (TRs). The topology of both networks showed similar features, with two major gene clusters and a smaller number of genes connecting these two clusters (Figure 2B). Overlaying the expression differences of the included genes revealed that one major cluster was defined by Tn-, the other by Tem cell-associated genes. Tcm cell-associated genes mainly connected the two main clusters, indicating that this population indeed represents an intermediate stage of T cell differentiation.

An additional bioinformatic approach was used to evaluate the mode of differentiation. We used the degree of similarity of the entire transcriptomes between Tn, Tcm, and Tem cells and calculated the likelihood of possible differentiation models: two linear models in the order of Tn-Tcm-Tem or Tn-Tem-Tcm and one bifurcated model in which Tcm and Tem cells arise independently from Tn. As shown in Figure 2C, the linear Tn-Tcm-Tem model had the highest cosine similarity score of 0.98 (max = 1) and was significantly different from the other two models ($p < 10^{-16}$).

These data show that the transcriptome changes progressively during Tmem cell differentiation in the order of Tn-Tcm-Tem-Temra.

Chromatin Accessibility and DNA-Meth Analyses Support a Linear Model of Differentiation for Circulating Tmem Cells

We wanted to clarify whether the linear relationship between the Tmem cell subsets (Tn-Tcm-Tem-Temra) apparent from the DNA-meth and transcriptome data (Figures 1E and 2A), could also be deduced from epigenetic imprints in the chromatin structure. For this, we first analyzed genome-wide DNA accessibility maps, which were generated by NOMe-seq. In a PCA, again a linear arrangement of the blood-derived T cell subsets in the order of Tn-Tcm-Tem-Temra was visible on the 2nd most important component PC2 (Figure 2D). However, the different populations were generally less stringently separated. The main component PC1 separated the replicates Hf03 and Hf04 from Hf06, which reflected a slight change in the NOMe protocol between these samples (Figure S2E and Supplemental Experimental Procedures). In addition, when we called accessible (= nucleosome-depleted) regions (NOMe-peaks) from Tn, Tcm, and Tem cells and compared their degree of accessibility between the cellular subtypes, the vast majority of sites gained or lost accessibility in the order Tn-Tcm-Tem (Figure 2E).

Next, we analyzed global DNA-meth profiles (by RRBS) of blood- and bone-marrow-derived Tmem cell subsets, with the latter population subdivided into a tissue-resident CD69$^+$ and a circulating CD69$^-$ fraction (CD69 being a regulator of tissue egress and marker for tissue-resident cells; Sathaliyawala et al., 2013). While the CD69$^-$ fraction clustered closely to Tcm and Tem cells from the blood, the CD69$^+$ tissue-resident BM-Tmem subfraction deviated from its CD69$^-$ counterpart, as well as from blood-derived circulating populations in PC2 (Figure 2F), indicating a major epigenetic imprint for their tissue residency and specialized function.

Taken together, our results from epigenomic and transcriptomic analyses support a linear model of differentiation for circulating Tmem cells from the blood with the bone-marrow-resident (CD69$^+$) T cell population deviating early and displaying a specific epigenetic imprint (Figure S2F).

Changes in DNA-Meth of Transcriptional Control Elements Are Associated with Tmem Cell Differentiation

DNA-meth can control the expression of genes, which are required for the maintenance of lineage identity, as found for *Foxp3* in regulatory T cells (Huehn et al., 2009). This holds true also for CD4$^+$ Tmem cells, as we found a correlation between DNA-meth and gene-expression changes, when we used an integrative sparse linear regression model measuring DNA-meth in promoters and gene bodies (Figure S2G). While the highest impact on gene expression was computed for predicted TF binding in accessible chromatin sites (NOMe peaks around genes), DNA-meth had a higher regulation potential than miRNAs.

With our genome-wide epigenetic datasets we therefore strived to (1) elucidate to what extent DNA-meth is involved in the regulation of known key Tmem cell checkpoint regulators, and (2) investigate whether epigenomic data could identify novel transcriptional regulators of Tmem differentiation.

For this, we called differentially methylated regions (DMRs) from the WGBS datasets, using the Metilene software (Jühling et al., 2016) applying strict selection criteria (min. #CpGs = 5; min. coverage = 5 reads) and a context-sensitive filtering step to reduce the contribution of the global de-methylation effect observed in PMDs ("adaptive filtering," Supplemental Experimental Procedures). This approach resulted in 1670 DMRs between Tn, Tcm and Tem cells (Table S5) associated with

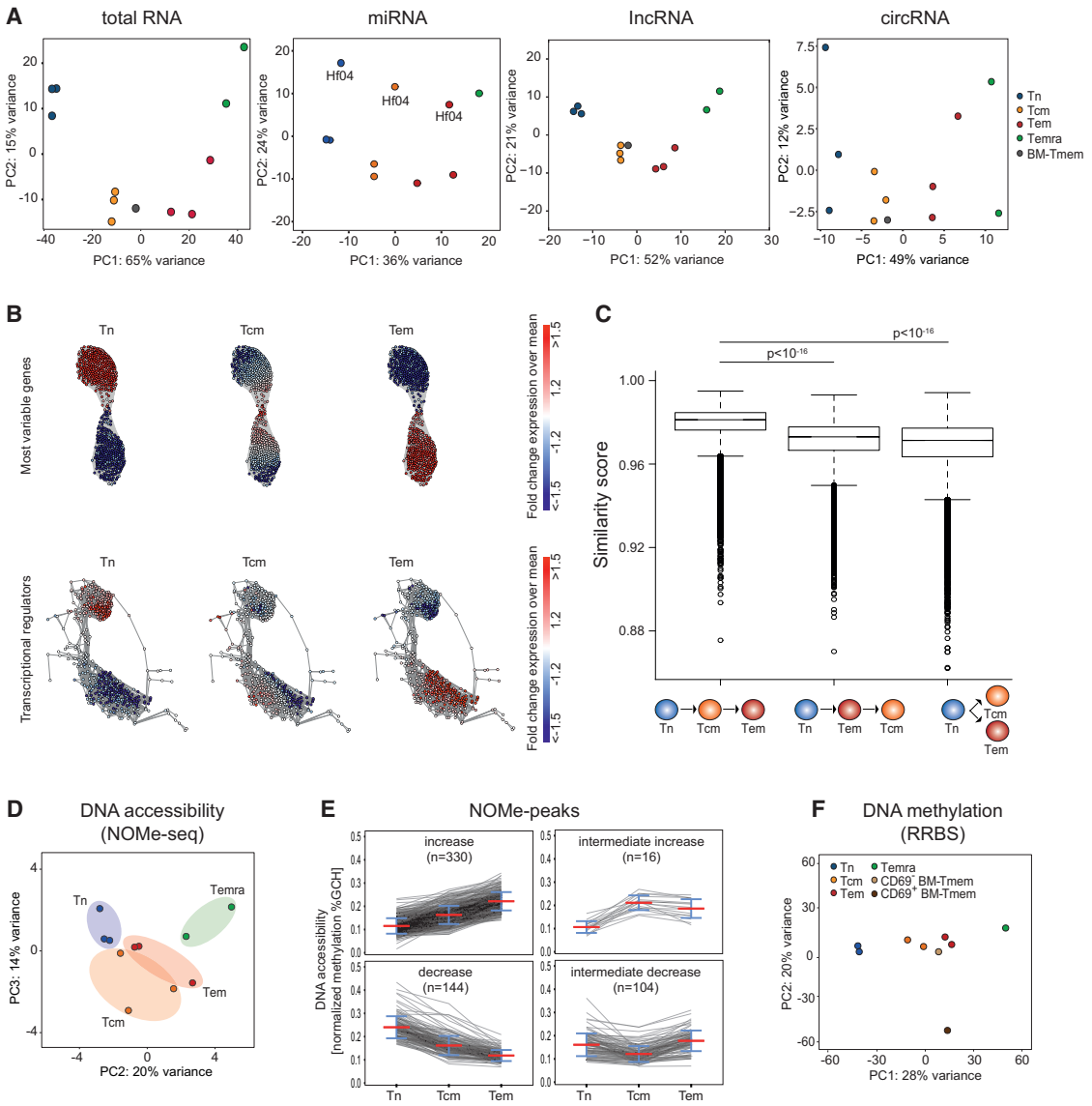


Figure 2. Progressive Changes in the Transcriptomes and in the DNA Accessibility Profiles Support a Linear Differentiation Model for Circulating Human CD4$^+$ Tmem Subsets

(A) PCAs of different RNA species for Tn, Tcm, Tem, Temra cells from blood, and Tmem cells from the bone-marrow (BM-Tmem).

(B) Co-regulation network based on the top 700 most variable genes in the dataset (top) or based on transcriptional regulators (TRs, bottom). Nodes represent genes colored according to the corresponding fold-change to mean expression. Links are unweighted and represent significant correlations.

(C) Three possible differentiation models (x axis) were compared using a designed similarity score (y axis), based on the hypothesis, that T cells that are closer to each other in the differentiation order should show more similar gene-expression profiles. The plot shows the distribution of similarity scores obtained (error bars denote SD estimated from 100.000 bootstrap samples). A significantly higher score was obtained for Tn-Tcm-Tem compared to the other models (bootstrapped t test p value).

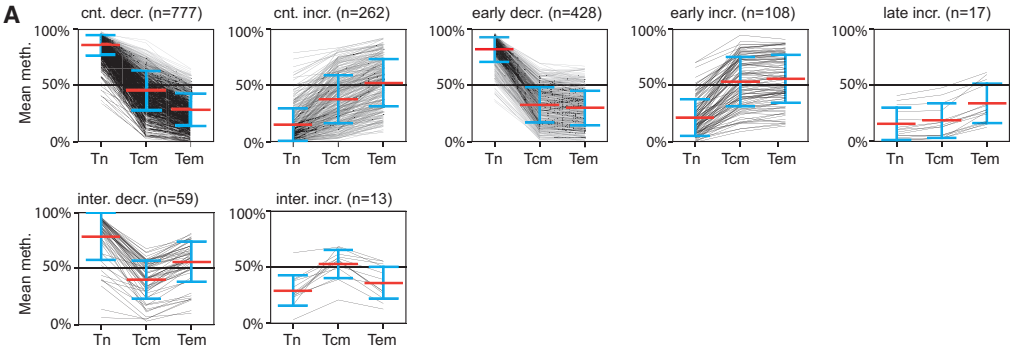
(D) PCA of DNA accessibility data (based on NOMe-seq data).

(E) Visualization of the degree of DNA accessibility (quantile-normalized GCH methylation levels) in consistent nucleosome depleted regions (NOMe-peaks) with a statistical difference between at least two cell types. Bars denote mean and SD.

(F) PCA of DNA-meth data (based on RRBS data). CpGs with min. coverage of 5 were considered; only CpGs with calls in all indicated samples were used.

970 protein-coding genes. These DMRs seemed functionally relevant for the regulation of gene expression as most of them were located within or proximal to genes and were classified as promoters or enhancers according to their histone modification profile (Figure S3A). The majority of these DMRs showed a continuous (Tn > Tcm > Tem, 47%) or early (Tn > Tcm and Tem, 26%) decrease in DNA-meth with Tmem cell generation (Figure 3A).

Next, we analyzed the correlation between DNA-meth changes and gene expression and found that 516 of the DMRs (36%) displayed an inverse correlation to gene expression (Figure S3B). Such DMRs showed the paradigm mode of gene

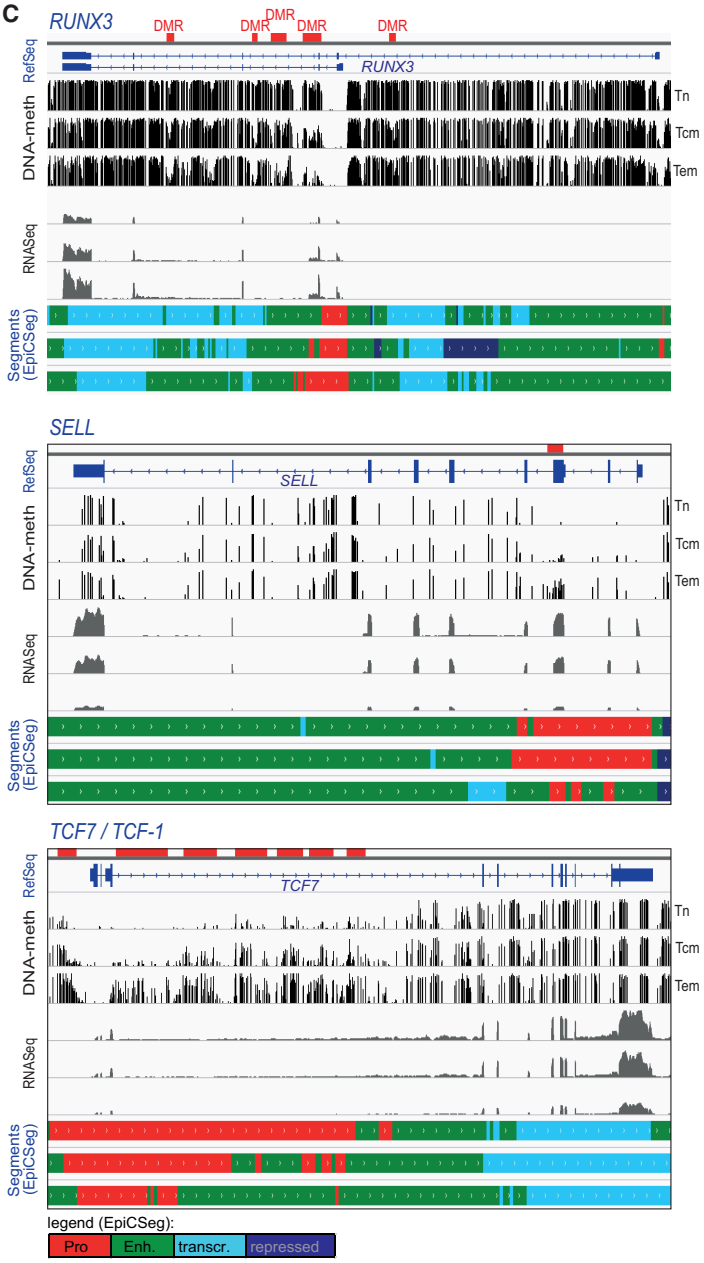


B

Gene Symbol	# of DMRs	methylation change (Tn-Tcm-Tem)	expression change (Tn-Tcm-Tem)
ARNT2	1		
BACH2	1		
	1		
BATF	3		
BCL2	1		
	1		
BCOR	7		
CCR7	1		
CDKN2A	1		
DUSP4	1		
DUSP5	1		
ELF4	1		
FASN	1		
FOXO1	2		
HNRNPLL	3		
IL10RA	1		
IL2RA	1		
IL2RB	3		
LAG3	2		
LEF1	4		
MAF	2		
NCOR2	1		
NFATC2	7		
NOD2	1		
NOTCH1	1		
PDCD1	3		
PPP3CB	1		
PRDM1	2		
RBPJ	1		
RPTOR	4		
	3		
RUNX3	5		
SELL	1		
SLAMF1	1		
SLFN12L	1		
	3		
STAM	1		
TBX21	2		
TCF4	1		
TCF7/TCF-1	7		
TNFRSF1B	3		
TOX	2		
ZBTB32	1		
ZEB2	2		

legend: incr. | decr. | no expr. | no change

C



(legend on next page)

regulation by DNA-meth, in which transcriptional control elements, such as promoter and enhancers, are repressed by increased DNA-meth, whereas a loss of DNA-meth at these elements leads to gene activation. Other DMRs were associated with genes by location, which (1) were not expressed in any cell type, (2) the expression of which did not change, or (3) the expression change correlated to the methylation change (Figure S3B). These classes of DMRs might serve different functions such as (1) preparing a locus for gene expression upon additional environmental signals or locking a locus to prevent alternative cellular fates, (2) stabilizing otherwise transient gene expression, or (3) affecting sites acting as silencers. Furthermore, it cannot be excluded that some DMRs might also act as long-range regulators for distant genes.

These data show that in addition to the large-scale DNA demethylation in PMDs, transcriptional control elements such as promoters and enhancers are targets of epigenetic regulation during Tmem cell differentiation.

Known Regulators of Tmem Differentiation and Function Display DNA-Meth Changes in Transcriptional Control Regions

To test the assumption that key factors regulating Tmem differentiation and function are under epigenetic control, we extracted a list of 144 known memory-related genes according to recent reviews (Figure S3C) and checked for the occurrence of DMRs in their loci. One quarter of these Tmem cell-related genes displayed one or several DMRs (Figure 3B), 95% of which were associated with a promoter or enhancer histone signature (Table S5). The largest group lost DNA-meth with progressive differentiation, which correlated with an increase in expression (Figure 3B). Among them were genes upregulated upon differentiation from naive to memory states, including surface or intracellular receptors such as *PDCD1* (PD-1), *IL2RA*, and *IL2RB*, *NOD2*, *SLAMF1*, and *TNFRSF1B*, but also many transcriptional regulators such as *RUNX3* (Figure 3C, top), *NFATC2*, *BATF*, *MAF*, *TBX21* (T-BET), the CD45-splicing regulator *HNRNPLL*, *PRDM1* (BLIMP-1), *DUSP4*, *DUSP5*, *STAM*, *TOX*, and *ZEB2*. In a smaller group, increased methylation was linked with decreased expression. This group included the signature markers of Tn and Tcm *SELL* (L-SELECTIN; Figure 3C, middle) and *CCR7*, but also several key transcriptional regulators, namely *TCF7* (encodes TCF-1; Figure 3C, bottom), *LEF1*, and *BACH2*, which are known to control the development or maintenance of Tmem cells. In a few genes (*FOXO1* and *BACH2*), loss of DNA-meth was associated with a decrease in expression; accordingly, these DMRs might control transcriptional silencers. In other cases (*NOTCH1*, *SLFN12L*, *RPTOR*, *ZBTB32*, *RBPJ*), a change in methylation was not correlated to changes in expression. It is of high interest to investigate whether loss of DNA-meth in genes of this group is not a requirement for expression but might act by stabilizing transcription, as was found previously for the TSDR (CNS2) enhancer region in the *FOXP3* locus (Huehn et al., 2009; Polansky et al., 2008).

While a causal role of DNA-meth in the regulation of these genes remains to be experimentally demonstrated, these findings provide evidence that epigenetic mechanisms contribute to the developmental regulation of Tmem cells by controlling the expression of key genes.

Integrative Analyses of Epigenomic and Transcriptomic Datasets Facilitate the Identification of Functional Regulator Candidates for Tmem Differentiation

In addition to screening for known developmental regulators, a reciprocal approach can be applied, in which the occurrence of DMRs is used to identify novel candidate genes that might be involved in the control of Tmem differentiation and which undergo direct epigenetic expression control. To this end, 171 DMRs associated with 104 transcriptional regulator genes were identified (Table S5), indicating that these genes might contribute to memory development and maintenance. This list included the gene *FOXP1*.

While these candidates seem to undergo direct expression control by DNA-meth, a different class of regulators might gain or lose functional importance for Tmem cell development and/or function because their binding sites in target genes are being exposed (or blocked) by chromatin remodeling. To identify such "functional epigenetic" regulators, we used an alternative approach combining DNA accessibility data and transcriptomic data. In this, the impact of a given TF on the transcriptional profile is determined by the accessibility of its binding sites within promoters and enhancers in its target genes. We used our DNA accessibility dataset (NOMe-seq data) and computed TF binding affinities to open chromatin regions (NOMe-peaks). Using a machine learning approach (Schmidt et al., 2016), we modeled differential gene expression between T cell subsets based on these TF binding predictions. In this, the impact of each TF to the differential transcriptional profile is calculated and TFs with a strong influence can be extracted. Indeed, comparison of modeled to observed gene-expression changes, as measured by RNA-seq, displayed a high accuracy (Figure S4A) supporting the validity of the approach. A number of regulatory TF candidates for Tn cells and Tmem subsets could be extracted (Figure 4A and 4B). The lists comprised TFs known to control Tmem cells, such as BCL6, E2F2, and RUNX3, as well as new candidates, including AHR, CREB1, ETS1, FLI1, FOXP1, FOXJ3, NFEL2, NRF1, RFX3, and ZFP161. For a number of these (e.g., AHR, FLI1, FOXP1, and RUNX3), we also found an associated DMR in their genes and differential expression during Tmem cell differentiation (Table S5), indicating that these factors not only drive transcriptional profiles during Tmem differentiation but are under epigenetic regulation themselves.

Figure 3. Epigenetic Changes in Known Tmem Cell-Related Genes

(A) Patterns of DNA-meth changes in differentially methylated regions (DMRs). DMRs changing in the order Tn-Tcm-Tem are shown in the upper row. cnt., continuous; inter., intermediate; decr., decrease; incr., increase. Bars denote mean and SD.

(B) List of known Tmem cell-related genes (based on Figure S3C) which display at least one DMR in their loci (color legend shown on the bottom).

(C) Examples of DMR-containing Tmem regulator genes (*RUNX3*, top; *SELL*, middle; *TCF7*/*TCF-1*, bottom) showing the location of the identified DMRs (red, top track). The following tracks are shown in each panel (top to bottom, each for Tn, Tcm, Tem): Gene annotation from RefSeq; DNA-meth (WGBS); polyA-RNA; genome segmentation by EpicCSeg (color legend shown on the bottom).

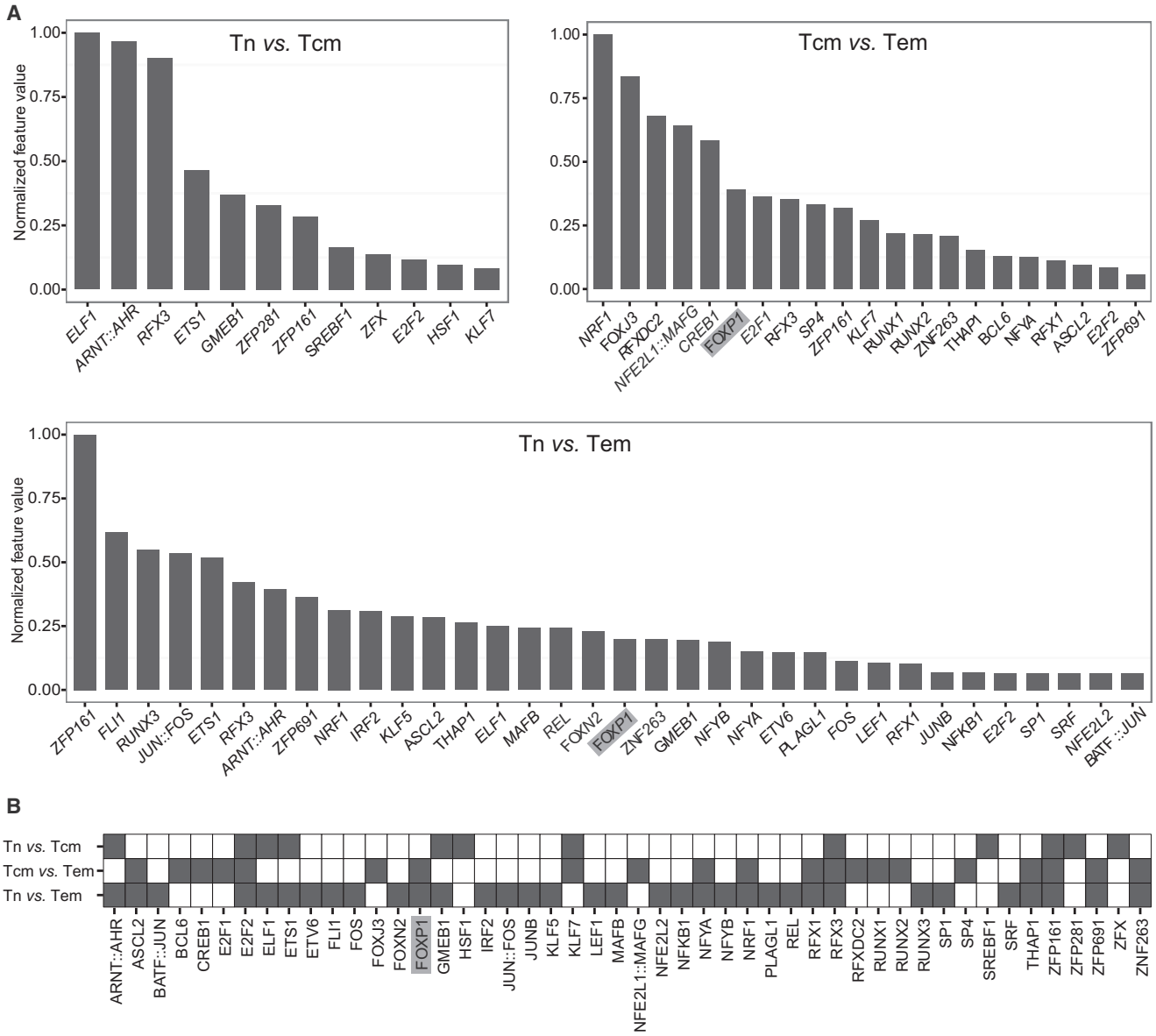


Figure 4. Selection of Tmem Cell Regulator Candidates Based on Their Predicted Binding Affinities in Open Chromatin Regions of Differentially Expressed Genes

(A) Bar plots showing normalized feature values (y axis) for each TF (x axis) computed using a machine learning approach (based on logistic regression classifiers) to predict differentially expressed genes in pairwise comparisons of two cellular subtypes. Differences in predicted TF affinities, calculated from open chromatin regions (NOMe-peaks) in the vicinity of a gene, were used as features in the classification. Large feature values denote a higher impact of the TF on differential gene expression.

(B) Summarized representation of all selected TFs shown in (A). Filled boxes reflect that a TF (column) has been selected as a feature in the respective comparison (row). TFs joint by double colons indicate that both TFs are predicted to bind as a complex. The TF FOXP1 is highlighted in gray.

With these analyses we identified several promising new TFs from epigenomic data which are likely to be involved in Tmem cell generation and function.

FOXP1 Is an Epigenetically Controlled "Naive-Keeping" Checkpoint Regulator

We found the TF FOXP1 to be a particularly interesting candidate for Tmem cell regulation due to several reasons: First, a DMR in the *FOXP1* locus displayed increasing DNA-meth, concomitant with decreased mRNA levels from Tn to Tem (Table S5); second, FOXP1 was predicted to bind to accessible chromatin regions and thus contribute to differential gene expression in Tn versus Tem (Figure 4); and third, *FOXP1* was among the top predicted Tn cell-specific regulators according to an iRegulon (Janky et al., 2014) analysis (Figure S4B), which is based on the enrichment of TF binding sites in genes contributing to the cell-type-specific clusters shown in Figure 2B. Therefore, we selected the TF FOXP1 for a more detailed investigation.

Our data suggested that FOXP1 might act as an important regulator for the Tn-to-Tmem transition. Confirming this, we found that T cell-specific depletion of Foxp1 protein expression in Foxp1 conditional-deficient mice resulted in loss of the naive $CD44^{low}$ phenotype in T cells (Figure 5A). These findings together with published data (Feng et al., 2011; Wei et al., 2016), support the view that Foxp1 acts as a "naive-keeping" factor for T cells. Analyses of DNA-meth in our datasets revealed a DMR in the *FOXP1* locus, which displayed a strong progressive gain of methylation with differentiation (Tn < Tcm < Tem), which was classified as a selective active promoter in Tn cells (Figure 5B) based on the displayed histone modification patterns (by EpiCSeg, Mammana and Chung, 2015). Indeed, a methylation-sensitive promoter activity of the *FOXP1-DMR* was confirmed in luciferase reporter gene assays in primary human $CD4^+$ T cells, as the *FOXP1-DMR* was able to drive luciferase expression when cloned upstream of the reporter gene in the sense orientation, but not when the orientation of the *FOXP1-DMR* was inverted or when the *FOXP1-DMR* was methylated (Figure 5C).

Consistent with the occurrence of a Tn cell-specific promoter, we found the FOXP1 protein expression in human $CD4^+$ T cells to be highest in Tn cells and to be decreased in Tcm, Tem, and Temra cells (Figure 5D). In addition, we found indications in our RNA-seq datasets for three alternative shorter RNA isoforms, which started within or directly downstream of the *FOXP1-DMR* promoter (Figure S4C). All three isoforms showed preferential expression in Tn compared to Tcm and Tem cells as measured by qPCR (Figure 5E). In addition, two of them contain the complete protein coding sequence (Figure 5E), which we verified by single molecule real-time sequencing (data not shown).

Taken together, these results validate FOXP1 as an important gate-keeper for the naive-to-memory transition, which was identified by integrative analyses of epigenomic data. In addition, these analyses also enabled the identification of the epigenetic control mechanisms regulating differential *FOXP1* expression during Tmem cell generation.

DISCUSSION

This study reveals that the differentiation of Tn cells into distinct types of memory cells and their long-term maintenance is connected to major epigenetic and transcriptional reprogramming. This is manifested on a global scale with a genome-wide segmented loss of DNA-meth during differentiation and in gene-specific epigenetic changes, which control the stage-specific expression and/or function of transcriptional regulators.

As our first major finding, we documented a progressive genome-wide loss of DNA-meth upon transition from the naive to the memory stages. This de-methylation was most prominent in "partially methylated domains" (PMDs, Hon et al., 2012; Lister et al., 2009), a feature shared in memory differentiation of B cells, but absent during the differentiation of monocytes into macrophages (Wallner et al., 2016). PMDs have been associated with heterochromatic histone signatures and correlate to regions, which are replicated late during S phase and progressively lose methylation during strong proliferation (Aran et al., 2011). Consistent with this, T and B cells, but not monocytes, undergo extensive proliferation during differentiation as a result of TCR-(BCR-) mediated activation. It is therefore feasible that the observed PMD-associated loss of methylation is a consequence and a signature of highly proliferative episodes in the history of these cells.

This interpretation is supported for $CD4^+$ Tmem cells by two additional observations in our study: (1) the progressive shortening of telomere length in the order of Tn-Tcm-Tem (Figure S5A), and (2) the progressive loss of methylation in PMDs observed in short-term culture of Tn cells proliferating in vitro after TCR-mediated activation (Figures S5B and S5C). It remains to be investigated whether the global de-methylation is just a tolerated bystander effect of proliferation or whether it constitutes a telomere-independent senescence signal for the cell as suggested by studies on hematopoietic stem cells under proliferative stress (Beerman et al., 2013).

These findings are relevant for the interpretation and functional assignment of DNA-meth changes found by gene-specific DNA-meth assays. Using the epigenomic maps of this study as a reference, gene-specific differentially methylated regions (DMRs), which might have (direct) functional relevance for gene expression, can be discriminated from DMRs in regions that are likely to represent a mere imprint from the proliferation history. This discrimination could be of particular relevance when studying epigenetic changes in T cells isolated from chronic stimulatory conditions, e.g., inflammatory diseases such as rheumatoid arthritis or lupus erythematosus where disease-associated methylation differences have been reported (de Andres et al., 2015; Javierre et al., 2010).

The second major conclusion from our global epigenomic analyses sheds light onto the still-controversial subject of the mode of memory differentiation of human $CD4^+$ T cells: Do memory cells originate (1) early after antigen encounter independently of (but in parallel to) the massive expansion of short-lived effector cells (parallel or bifurcative differentiation model, Arsenio et al., 2015) or (2) do they develop from effector cells, which adopt Tmem stages toward the end of the primary effector phase (linear model)? While we could not directly address the positioning of effector cells in relation to memory development with the present dataset, all our findings are more consistent with the linear progression model for circulating Tmem cells: We observed a strong loss of DNA-meth in PMDs upon transition from Tn to Tcm cells, which was further reduced in the more differentiated Tem and Temra phenotypes. These data indicate that Tcm cells would have already passed through a phase of intense proliferation during initial activation and prior to converting into resting memory cells. While the parallel model cannot be formally excluded by this, additional proliferation-independent datasets were similarly more consistent with the linear model, including: (1) patterns of DNA-meth at single-DMR resolution, as well as patterns of DNA accessibility (NOMe-peaks), showed almost exclusively changes in the order of Tn-Tcm-Tem, (2) changes in the transcriptomes grouped the samples along a progressive Tn-Tcm-Tem-Temra cell differentiation axis, (3) network analysis of co-expressed genes placed the Tcm phenotype as intermediate to the Tn- and Tem-associated clusters, and (4) calculation of the similarity of the transcriptomic profiles revealed the linear Tn-Tcm-Tem model as the most likely one.

These conclusions are in part in contrast to conclusions from restricted expression analyses of murine $CD8^+$ single cells

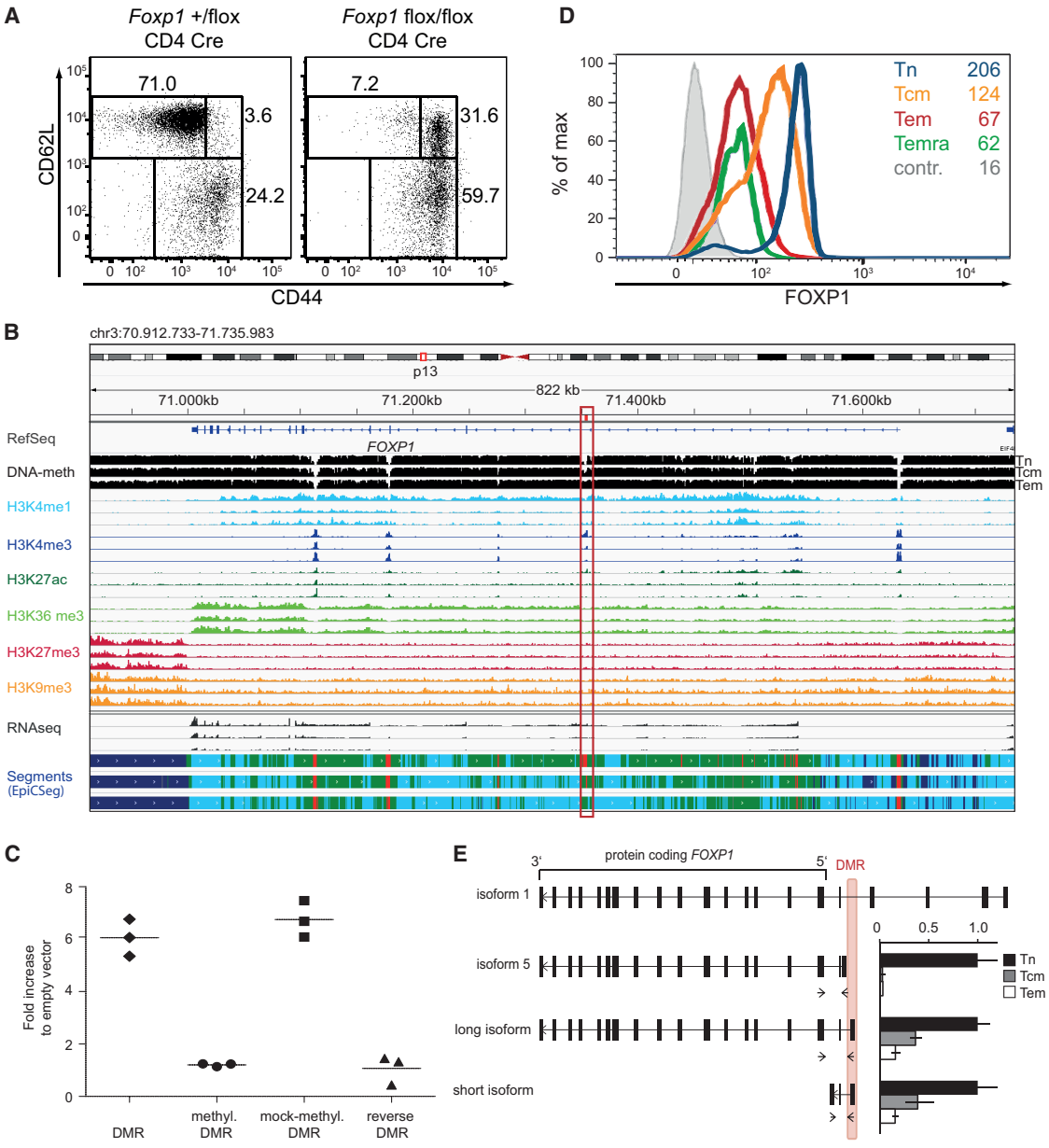


Figure 5. A Newly Identified Methylation-Sensitive Promoter Drives Alternative Coding mRNA Isoforms of the "Naive-Keeping" Transcription Factor FOXP1 in Tn

(A) $CD4^+$ T cells isolated from spleens of *Foxp1*$^{+/flox}$ CD4 Cre and *Foxp1*$^{flox/flox}$ CD4 Cre mice at the age of 6–10 weeks were analyzed by flow cytometry for a $CD44^{high}$ memory phenotype. Dot plots show CD62L and CD44 surface expression after gating on $CD4^+$ living cells.

(B) Genomic view of the human *FOXP1* locus indicating a distinct *FOXP1-DMR* (red box) gaining DNA-meth from Tn to Tcm to Tem cells. The following tracks are shown (each for Tn, Tcm, and Tem cells): RefSeq annotation, DNA-meth (WGBS), 6 histone modifications, total RNA coverage and segmentation by EpiCSeg (red, promoter; green, enhancer; light blue, transcribed; dark blue, repressed).

(C) Luciferase reporter assay testing a methylation-sensitive and orientation-dependent promoter activity of the *FOXP1-DMR* in primary human $CD4^+$ T cells. The Firefly luciferase signal was normalized to the signal of the Renilla transfection control and is shown relative to the empty vector control. One representative experiment performed in triplicates out of two independent experiments is shown.

(D) FOXP1 protein expression in gated human Tn, Tcm, Tem, and Temra cells as assessed by intracellular staining and flow cytometry. Numbers: geometric mean of the FOXP1 signal. Control staining (contr.) was done using the fluorescently labeled secondary antibody only.

(E) Schematic depiction of different human *FOXP1* RNA isoforms. The position of the *FOXP1-DMR* is shown and the protein coding exons are indicated as boxes. Three isoforms start within or shortly downstream of the *FOXP1-DMR*. Their relative expression values are shown for the different cell types (normalized to Tn cells) measured by qPCR using the indicated primers (arrowheads). One representative experiment performed in technical triplicates (mean and SD) out of two independent experiments is shown.

(Arsenio et al., 2015). However, elegant in vivo approaches using adoptive transfer systems of single murine $CD8^+$ memory cells (Gerlach et al., 2013a; Graef et al., 2014) also argue in favor of a linear differentiation model. As similar experiments have not been conducted for $CD4^+$ cells yet and are impossible in the human system, our data represent new insights into this topic.

The distinct positioning of Temra cells in the analysis of genome-wide DNA-meth and transcriptomic data suggests that they represent a late stage of Tmem differentiation and have undergone extensive proliferation. Alternatively, circulating Temra cells might represent survivors of prolonged effector proliferation due to chronic re-activation. This could also explain the enhanced heterogeneity of the Temra samples, since their epigenetic imprint might have been specialized over time. Indeed, increased frequencies of Temra cells have e.g., been reported in response to persistent CMV infection (Derhovanessian et al., 2011) and in liver disease, where high Temra cell numbers represent a significant risk of organ rejection after organ transplantation (Gerlach et al., 2013b). In contrast, Tmem cells isolated from the bone marrow did not display a Temra-like epigenetic phenotype but were positioned between Tcm and Tem cells, indicating that they have preserved significant expansion and differentiation capacity. In addition, the $CD69^+$ tissue-resident subset of BM-Tmem cells displayed a distinct DNA-meth profile, indicating that acquisition of a resident phenotype, too, is linked to significant epigenetic reprogramming.

The third pillar of our study is dedicated to the identification of factors, which drive and/or maintain Tmem cells. In this endeavor, we identified many non-coding RNAs (ncRNAs), which are highly and/or differentially expressed in Tn and Tmem cells. Among them, we identified numerous circRNAs for which the normal linear host transcript is barely detectable (Rybak-Wolf et al., 2015), thus, activity of these genes would have been missed in standard polyA- selected RNA-seq and/or normal linear splice analysis. These and other ncRNA molecules (lncRNAs, miRNAs) may not only reflect but also induce functional consequences during Tmem differentiation. The expression of the ncRNAs appears to be coordinated and finely tuned during Tmem differentiation with a remarkable link to other epigenomic changes. Therefore, our dataset provides a deep basis to further investigate the direct contribution of ncRNAs to Tmem differentiation and to clarify the mutual regulatory impact between ncRNAs and chromatin structure.

In addition to RNAs, we report two classes of protein regulators, which include known and potentially new factors controlling Tmem cell generation and function: (1) TFs, which undergo epigenetic expression control during Tmem cell formation, and (2) TFs, which gain or lose functional importance as their binding sites in target genes are being exposed or closed, respectively, independently of their own expression change. For the first class, we found several widely discussed regulators of Tmem differentiation, which displayed differential DNA-meth in promoter or enhancer regions that anti-correlated with differences in gene expression levels, following the classical paradigm of methylation-controlled gene repression. For several of them (e.g., *IL2RA*, *RUNX3*, *NFATC2*, *MAF*, *BACH2*, *FOXO1*), epigenetic control in Tmem differentiation has not been reported so far and awaits experimental confirmation. Interestingly, among differentially methylated genes was also *HNRNPLL*, involved in alternative splicing of CD45 to the Tmem signature isoform CD45RO, and the two homing-related receptors, *SELL* (L-selectin) and *CCR7*, suggesting that the permanent change in the recirculation pattern with transition from Tcm to Tem is epigenetically fixed. For others, concordant epigenetic changes have already been described in murine $CD8^+$ T cells (e.g., *BATF*, *LEF1*, *PDCD1*, *TBX21*, *TCF7*, *ZBTB32*; Scharer et al., 2013; Youngblood et al., 2011; Hashimoto et al., 2013).

As for the second class of TFs, we identified FOXP1 as one of the top candidates, a less well known but functionally confirmed Tmem regulator in mice (Feng et al., 2011; Wei et al., 2016). Our present analyses support a similar function in human $CD4^+$ Tmem and additionally unravel the epigenetic control of the *FOXP1* gene. Other prime TF candidates include TFs previously implicated in Tmem regulation such as RUNX3, E2F2, LEF1, BCL6, or members of the ELF-, KLF-, or FOXJ- families, as well as CREB1, ETS-1, and JUN-FOS, which are known to be involved in multiple cellular processes of differentiation and activation. Additional interesting candidates include (1) the aryl hydrocarbon receptor, AHR, which has been implicated in differentiation of $CD4^+$ T cells into pro- or anti-inflammatory subsets and, hence, to modulate autoimmune diseases in various animal models (reviewed in Esser et al., 2009; Hanieh, 2014) and (2) the ets-family member FLI-1, which has been reported to affect thymic T cell development, TCR signaling, glycosphingolipid metabolism, and cytokine expression and has been implicated in autoimmune diseases, too (Richard et al., 2013; Sato et al., 2014). Others of the top-predicted TFs have not been directly associated to regulation of Tmem differentiation yet, but are players in potentially relevant cellular processes, such as intracellular signaling (RFX3, via the RAS-MAPK pathway), metabolic processes (NRF1 and SREBF1, associated with mTORC1 signaling), and chromatin remodeling (ZFP161, targeting of the repressive Polycomp complex). Thus, important Tmem cell properties might be under the control of yet neglected transcriptional regulators that could be revealed by the integrated analysis of transcriptomic and epigenomic features.

In conclusion, the comprehensive epigenomic analysis of several human $CD4^+$ Tmem subsets in this study revealed insights into the Tmem differentiation pathway and allowed the identification of relevant epigenetically controlled transcriptional regulators. In addition, these data constitute a resource of normal T cell differentiation, which can serve as a reference for the identification of altered epigenetic signatures in T cells from pathological situations such as chronic inflammatory disease. The challenging task for the future will be the application of ''epigenetic engineering'' to achieve therapeutic re-programming of pathogenic T cells or to optimize T cells for their use in cellular therapy.

EXPERIMENTAL PROCEDURES

T Cell Isolation

PBMCs from blood of healthy female donors or from bone marrow samples of female donors undergoing hip replacements were isolated and enriched for $CD4^+$ T cells using the MACS-technology (Miltenyi Biotech). Tn cells and Tmem subsets were purified by flow cytometry using markers shown in Figure S1A. Donors gave their written and informed consent prior to participating in the study (Ethics committee of the Charite Universitaetsmedizin Berlin, application numbers EA1/116/13 and EA1/105/09).

Epigenomic Data Generation

WGBS was carried out by the combined analysis of two bisulfite-converted libraries using the pre-bisulfite library protocol (Urich et al., 2015) and the TruSeq DNA Methylation kit (Illumina, San Diego, USA). RRBS libraries were prepared as previously published (Boyle et al., 2012). For NOMe-seq, nuclei of fixed cells were extracted and DNA-meth on GpC motifs in accessible chromatin regions was introduced using the M. *CviP*I methyltransferase, followed by WGBS analysis. ChIP-seq for histone modifications was carried out as previously described (Arrigoni et al., 2016; Kinkley et al., 2016). RNA was extracted using the miRNeasy Micro Kit (QIAGEN) and three Illumina sequencing libraries were prepared (small RNA sequencing library, one stranded total RNA, and one stranded mRNA library). Sequencing was carried out on HiSeq 2000 and HiSeq2500 machines (Illumina). Bioinformatical processing of the sequencing reads including mapping to the hg19 reference genome is outlined in the Supplemental Experimental Procedures section.

DNA Methylation Analyses

Genome segmentation based on WGBS data was performed using MethylSeekR (Burger et al., 2013). The methylation levels from both strands were aggregated and weighted average methylation levels were plotted. WGBS data from B cells (Blueprint consortium) was converted to hg19 coordinates using the liftOver tool (Rosenbloom et al., 2015) and segmentation was carried out. Differentially methylated regions (DMRs) were predicted with Metilene (Jühling et al., 2016) in de-novo mode among sites with at least 5× coverage.

Calling of Accessible Chromatin Region

Nucleosome-depleted regions (NOMe-peaks) were identified by segmenting the GCH-methylation signal with a binomial hidden Markov model with two states (1 open/NDR, 0 background) in each sample separately and consistent NOMe-peaks confirmed in all three replicates were selected.

Identification of mRNAs, miRNAs, lncRNAs, and circRNAs

Expression values for total RNA were quantified using TopHat, Htseq-count, and DESeq2 (Anders and Huber, 2010). Cufflinks (Trapnell et al., 2010) was used for the identification of novel lncRNAs. To remove possible coding genes, we estimated the coding potential of novel transcripts using PhyloCSF (Lin et al., 2011) and CPAT (Wang et al., 2013). Mature miRNA read counts were estimated for each sample using miRDeep2 (Friedländer et al., 2012) and miRBase (version 21) annotations. CircRNAs were detected, filtered, and annotated as described before (Memczak et al., 2013).

Co-Expression Network Construction

Expression data of Tn, Tcm, and Tem cells (3 replicates each) was filtered using either a list of human transcriptional regulators (TRs) or the 700 most variable genes (i.e., most significant p values in an ANOVA-based analysis) to get a reduced expression table of present genes. The group of TRs contained transcription factors (TFs), co-factors, RNA-binding proteins and chromatin remodelers originating from the TFCat data base (Fulton et al., 2009). The expression matrices were loaded into BioLayout Express3D (Theocharidis et al., 2009) and co-regulation networks were generated with a Pearson correlation cutoff of 0.9. The predicted gene-gene pairs were visualized by Cytoscape (Shannon et al., 2003) and fold change expression values calculated against the group mean were mapped to the network.

Prediction of Transcriptional Regulators Using a Machine-Learning Approach to Model Differential Gene Expression

We used a machine-learning approach based on a logistic regression classifier with the elastic net penalty (Zou and Hastie, 2005) to model differential gene expression between the Tn, Tcm, and Tem subsets. Because the TF features for the logistic regression classifier, we used the ratio of TF gene scores, which were computed using TEPIC (Schmidt et al., 2016). TFs predicted to contribute to differential gene expression were selected.

Functional Analyses on the TF FOXP1 and the FOXP1-DMR

For the generation of a T cell-specific Foxp1 deletion, a conditional Foxp1 knock-out allele was generated using standard gene targeting techniques in murine ESCs by introducing *loxP* sites into intronic regions flanking exons 10–12 (T. Patzelt, O. Gorka, and J. Ruland, manuscript in preparation). The generated Foxp1-floxed mice were crossed to CD4-Cre animals (Lee et al., 2001).

Intracellular staining of FOXP1 protein in human $CD4^+$ T cells was performed using the Fixation/Permeabilization Buffer set for intracellular Foxp3 staining (eBioscience) in a two-step staining procedure (primary FOXP1 antibody polyclonal #2005, Cell Signaling Technology, DyeLight-649-labeled donkey anti-rabbit secondary antibody #406406, BioLegend). Samples were acquired on a BD LSRFortessa instrument (BDBioscience).

The *FOXP1-DMR* was cloned into the CpG-free Firefly luciferase vector pCpGL (Klug and Rehli, 2006). Treatment with the M.*Sss*I CpG methyltransferase (NEB) allowed selective methylation of the *FOXP1-DMR*. Ex vivo isolated $CD4^+$ T cells were TCR-stimulated for 48 hr and transfected with the *FOXP1-DMR* Firefly vector and a pRL-TK Renilla control vector (Promega) using the Neon™ Transfection System (Life Technologies). Firefly and Renilla luciferase activity were assessed using the Dual Luciferase Assay Kit (Promega) after 24 hr. The Firefly luciferase signal was normalized to the Renilla reporter signal.

Expression levels of the *FOXP1* RNA isoforms were quantified using the platinum SYBR green qPCR superMix-UDG (Thermo Fisher Scientific) on a Step One instrument (Thermo Fisher Scientific). Relative transcript levels were normalized to h*RPS18*. Primer sequences are given in the Supplemental Experimental Procedures section.

ACCESSION NUMBERS

All sequencing data have been deposited at the European Genome-Phenome Archive under the accession number EGAS00001001624. WGBS Blueprint data of B cells are available from the EGA under the accessions EGAD00001001590, EGAD00001001587, EGAD00001001548, and EGAD00001001160.

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures, five tables, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.immuni.2016.10.022.

AUTHOR CONTRIBUTIONS

Sample preparation: S. Frischbutter, S.S., U.H., C.C., Y.T., S.R., G.W., U.K., U.S., J.D., H.-D.C., B.S., A.H., J.K.P.; WGBS and NOMe-seq: G.G., N.G., J.W.; ChIP-seq: S.K., N.L., H.-R.C., L.A., A.S.R., T.M.; RNA-seq: S. Fröhler, W.C.; mapping & management of sequencing data: B.F., G.Z., K.N., P.E., C.D.I., B.H., B.B., J.E.; Data analysis: P.D., K.N., A. Salhab, F.M., J.W., T.L. (DNA-meth); P.D., J.P.K. (integrative DMR analysis), K.N., G.G. (NOMe-seq); X.Y., H.-R.C., P.E., T.L. (ChIP-seq); J.X., W.C. (lncRNAs); P.G., N.R. (circRNAs); F.K., N.R. (miRNAs); A. Sinha, P.R. (total RNAs); L.F., L.S., B.B. (telomere length); T.U., K.B., J.L.S. (co-expression network); F.S., A.D.A., F.B.A., M.H.S. (gene expression prediction and modeling of TF binding, similarity score); O.G., J.R. (Foxp1-KO mice); C.K., MdA, A.H., J.K.P. (functional FOXP1 analyses); C.K., MdA, A.R., J.D., B.S., A.H., J.K.P. (functional data interpretation). J.K.P., J.W., A.H. designed and coordinated the study supported by N.G.; J.K.P., J.W., A.H. wrote the manuscript with contributions from other authors.

ACKNOWLEDGMENTS

We thank René Maier for expert technical assistance, the FCCF of the DRFZ for expert cell sorting, staff from EURICE for project management, and Ulrike Biskup for help with figure layout. Selected data generated by the Blueprint Consortium (www.blueprint-epigenome.eu; EU FP7/2007-2013, grant agreement no. 282510 BLUEPRINT) were used. Funding is as follows: German Epigenome Programme (DEEP) of the Federal Ministry of Education and Research in Germany (BMBF), DFG HA 1505/10-1 to A.H., DFG-SFB650-TP1 to A.H., ERC grant 322865 to J.R., EU FP7 "ONE Study" to B.S., DFG SFB650 to B.S., DFG SFB704 to J.L.S., J.L.S. is a member of the Excellence Cluster Immunosensation, ERC-2010-AdG_20100317 Grant 268987 to A.R. and Priority Programme 1468 Immunobone to A.R.

Received: December 30, 2015
Revised: June 22, 2016
Accepted: July 22, 2016
Published: November 15, 2016

REFERENCES

Ahmed, R., Bevan, M.J., Reiner, S.L., and Fearon, D.T. (2009). The precursors of memory: models and controversies. Nat. Rev. Immunol. *9*, 662–668.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Aran, D., Toperoff, G., Rosenberg, M., and Hellman, A. (2011). Replication timing-related and gene body-specific methylation of active human genes. Hum. Mol. Genet. *20*, 670–680.

Arrigoni, L., Richter, A.S., Betancourt, E., Bruder, K., Diehl, S., Manke, T., and Bönisch, U. (2016). Standardizing chromatin research: a simple and universal method for ChIP-seq. Nucleic Acids Res. *44*, e67.

Arsenio, J., Metz, P.J., and Chang, J.T. (2015). Asymmetric Cell Division in T Lymphocyte Fate Diversification. Trends Immunol. *36*, 670–683.

Beerman, I., Bock, C., Garrison, B.S., Smith, Z.D., Gu, H., Meissner, A., and Rossi, D.J. (2013). Proliferation-dependent alterations of the DNA methylation landscape underlie hematopoietic stem cell aging. Cell Stem Cell *12*, 413–425.

Boyle, P., Clement, K., Gu, H., Smith, Z.D., Ziller, M., Fostel, J.L., Holmes, L., Meldrim, J., Kelley, F., Gnirke, A., and Meissner, A. (2012). Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. Genome Biol. *13*, R92.

Burger, L., Gaidatzis, D., Schübeler, D., and Stadler, M.B. (2013). Identification of active regulatory regions from DNA methylation data. Nucleic Acids Res. *41*, e155.

Carbone, F.R., Mackay, L.K., Heath, W.R., and Gebhardt, T. (2013). Distinct resident and recirculating memory T cell subsets in non-lymphoid tissues. Curr. Opin. Immunol. *25*, 329–333.

Crompton, J.G., Narayanan, M., Cuddapah, S., Roychoudhuri, R., Ji, Y., Yang, W., Patel, S.J., Sukumar, M., Palmer, D.C., Peng, W., et al. (2016). Lineage relationship of CD8(+) T cell subsets is revealed by progressive changes in the epigenetic landscape. Cell. Mol. Immunol. *13*, 502–513.

de Andres, M.C., Perez-Pampin, E., Calaza, M., Santaclara, F.J., Ortea, I., Gomez-Reino, J.J., and Gonzalez, A. (2015). Assessment of global DNA methylation in peripheral blood cell subpopulations of early rheumatoid arthritis before and after methotrexate. Arthritis Res. Ther. *17*, 233.

Derhovanessian, E., Maier, A.B., Hähnel, K., Beck, R., de Craen, A.J., Slagboom, E.P., Westendorp, R.G., and Pawelec, G. (2011). Infection with cytomegalovirus but not herpes simplex virus induces the accumulation of late-differentiated CD4+ and CD8+ T-cells in humans. J. Gen. Virol. *92*, 2746–2756.

Esser, C., Rannug, A., and Stockinger, B. (2009). The aryl hydrocarbon receptor in immunity. Trends Immunol. *30*, 447–454.

Feng, X., Wang, H., Takata, H., Day, T.J., Willen, J., and Hu, H. (2011). Transcription factor Foxp1 exerts essential cell-intrinsic regulation of the quiescence of naive T cells. Nat. Immunol. *12*, 544–550.

Flossdorf, M., Rössler, J., Buchholz, V.R., Busch, D.H., and Höfer, T. (2015). CD8(+) T cell diversification by asymmetric cell division. Nat. Immunol. *16*, 891–893.

Friedländer, M.R., Mackowiak, S.D., Li, N., Chen, W., and Rajewsky, N. (2012). miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. Nucleic Acids Res. *40*, 37–52.

Fulton, D.L., Sundararajan, S., Badis, G., Hughes, T.R., Wasserman, W.W., Roach, J.C., and Sladek, R. (2009). TFCat: the curated catalog of mouse and human transcription factors. Genome Biol. *10*, R29.

Gerlach, C., Rohr, J.C., Perié, L., van Rooij, N., van Heijst, J.W., Velds, A., Urbanus, J., Naik, S.H., Jacobs, H., Beltman, J.B., et al. (2013a). Heterogeneous differentiation patterns of individual CD8+ T cells. Science *340*, 635–639.

Gerlach, U.A., Vogt, K., Schlickeiser, S., Meisel, C., Streitz, M., Kunkel, D., Appelt, C., Ahrlich, S., Lachmann, N., Neuhaus, P., et al. (2013b). Elevation of CD4+ differentiated memory T cells is associated with acute cellular and antibody-mediated rejection after liver transplantation. Transplantation *95*, 1512–1520.

Graef, P., Buchholz, V.R., Stemberger, C., Flossdorf, M., Henkel, L., Schiemann, M., Drexler, I., Höfer, T., Riddell, S.R., and Busch, D.H. (2014). Serial transfer of single-cell-derived immunocompetence reveals stemness of CD8(+) central memory T cells. Immunity *41*, 116–126.

Hanieh, H. (2014). Toward understanding the role of aryl hydrocarbon receptor in the immune system: current progress and future trends. BioMed Res. Int. *2014*, 520763.

Harrington, L.E., Janowski, K.M., Oliver, J.R., Zajac, A.J., and Weaver, C.T. (2008). Memory CD4 T cells emerge from effector T-cell progenitors. Nature *452*, 356–360.

Hashimoto, S., Ogoshi, K., Sasaki, A., Abe, J., Qu, W., Nakatani, Y., Ahsan, B., Oshima, K., Shand, F.H., Ametani, A., et al. (2013). Coordinated changes in DNA methylation in antigen-specific memory CD4 T cells. J. Immunol. *190*, 4076–4091.

Henson, S.M., Riddell, N.E., and Akbar, A.N. (2012). Properties of end-stage human T cells defined by CD45RA re-expression. Curr. Opin. Immunol. *24*, 476–481.

Hon, G.C., Hawkins, R.D., Caballero, O.L., Lo, C., Lister, R., Pelizzola, M., Valsesia, A., Ye, Z., Kuan, S., Edsall, L.E., et al. (2012). Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. Genome Res. *22*, 246–258.

Huehn, J., Polansky, J.K., and Hamann, A. (2009). Epigenetic control of FOXP3 expression: the key to a stable regulatory T-cell lineage? Nat. Rev. Immunol. *9*, 83–89.

Janky, R., Verfaillie, A., Imrichová, H., Van de Sande, B., Standaert, L., Christiaens, V., Hulselmans, G., Herten, K., Naval Sanchez, M., Potier, D., et al. (2014). iRegulon: from a gene list to a gene regulatory network using large motif and track collections. PLoS Comput. Biol. *10*, e1003731.

Javierre, B.M., Fernandez, A.F., Richter, J., Al-Shahrour, F., Martin-Subero, J.I., Rodriguez-Ubreva, J., Berdasco, M., Fraga, M.F., O'Hanlon, T.P., Rider, L.G., et al. (2010). Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. Genome Res. *20*, 170–179.

Jühling, F., Kretzmer, H., Bernhart, S.H., Otto, C., Stadler, P.F., and Hoffmann, S. (2016). metilene: fast and sensitive calling of differentially methylated regions from bisulfite sequencing data. Genome Res. *26*, 256–262.

Kaech, S.M., and Cui, W. (2012). Transcriptional control of effector and memory CD8+ T cell differentiation. Nat. Rev. Immunol. *12*, 749–761.

Kinkley, S., Helmuth, J., Polansky, J.K., Dunkel, I., Gasparoni, G., Fröhler, S., Chen, W., Walter, J., Hamann, A., and Chung, H.R. (2016). reChIP-seq reveals widespread bivalency of H3K4me3 and H3K27me3 in CD4(+) memory T cells. Nat. Commun. *7*, 12514.

Klug, M., and Rehli, M. (2006). Functional analysis of promoter CpG methylation using a CpG-free luciferase reporter vector. Epigenetics *1*, 127–130.

Komori, H.K., Hart, T., LaMere, S.A., Chew, P.V., and Salomon, D.R. (2015). Defining CD4 T cell memory by the epigenetic landscape of CpG DNA methylation. J. Immunol. *194*, 1565–1579.

Kulis, M., Merkel, A., Heath, S., Queirós, A.C., Schuyler, R.P., Castellano, G., Beekman, R., Raineri, E., Esteve, A., Clot, G., et al. (2015). Whole-genome fingerprint of the DNA methylome during human B cell differentiation. Nat. Genet. *47*, 746–756.

Lee, P.P., Fitzpatrick, D.R., Beard, C., Jessup, H.K., Lehar, S., Makar, K.W., Pérez-Melgosa, M., Sweetser, M.T., Schlissel, M.S., Nguyen, S., et al. (2001). A critical role for Dnmt1 and DNA methylation in T cell development, function, and survival. Immunity *15*, 763–774.

Lin, M.F., Jungreis, I., and Kellis, M. (2011). PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. Bioinformatics *27*, i275–i282.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.M., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature *462*, 315–322.

Mammana, A., and Chung, H.R. (2015). Chromatin segmentation based on a probabilistic model for read counts explains a large portion of the epigenome. Genome Biol. *16*, 151.

Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., Maier, L., Mackowiak, S.D., Gregersen, L.H., Munschauer, M., et al. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. Nature *495*, 333–338.

Okhrimenko, A., Grün, J.R., Westendorf, K., Fang, Z., Reinke, S., von Roth, P., Wassilew, G., Kühl, A.A., Kudernatsch, R., Demski, S., et al. (2014). Human memory T cells from the bone marrow are resting and maintain long-lasting systemic memory. Proc. Natl. Acad. Sci. USA *111*, 9229–9234.

Polansky, J.K., Kretschmer, K., Freyer, J., Floess, S., Garbe, A., Baron, U., Olek, S., Hamann, A., von Boehmer, H., and Huehn, J. (2008). DNA methylation controls Foxp3 gene expression. Eur. J. Immunol. *38*, 1654–1663.

Richard, E.M., Thiyagarajan, T., Bunni, M.A., Basher, F., Roddy, P.O., Siskind, L.J., Nietert, P.J., and Nowling, T.K. (2013). Reducing FLI1 levels in the MRL/lpr lupus mouse model impacts T cell function by modulating glycosphingolipid metabolism. PLoS ONE *8*, e75175.

Rosenbloom, K.R., Armstrong, J., Barber, G.P., Casper, J., Clawson, H., Diekhans, M., Dreszer, T.R., Fujita, P.A., Guruvadoo, L., Haeussler, M., et al. (2015). The UCSC Genome Browser database: 2015 update. Nucleic Acids Res. *43*, D670–D681.

Russ, B.E., Olshanksy, M., Smallwood, H.S., Li, J., Denton, A.E., Prier, J.E., Stock, A.T., Croom, H.A., Cullen, J.G., Nguyen, M.L., et al. (2014). Distinct epigenetic signatures delineate transcriptional programs during virus-specific CD8(+) T cell differentiation. Immunity *41*, 853–865.

Rybak-Wolf, A., Stottmeister, C., Glažar, P., Jens, M., Pino, N., Giusti, S., Hanan, M., Behm, M., Bartok, O., Ashwal-Fluss, R., et al. (2015). Circular RNAs in the Mammalian Brain Are Highly Abundant, Conserved, and Dynamically Expressed. Mol. Cell *58*, 870–885.

Sallusto, F., Lenig, D., Förster, R., Lipp, M., and Lanzavecchia, A. (1999). Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. Nature *401*, 708–712.

Sathaliyawala, T., Kubota, M., Yudanin, N., Turner, D., Camp, P., Thome, J.J., Bickham, K.L., Lerner, H., Goldstein, M., Sykes, M., et al. (2013). Distribution and compartmentalization of human circulating and tissue-resident memory T cell subsets. Immunity *38*, 187–197.

Sato, S., Lennard Richard, M., Brandon, D., Jones Buie, J.N., Oates, J.C., Gilkeson, G.S., and Zhang, X.K. (2014). A critical role of the transcription factor fli-1 in murine lupus development by regulation of interleukin-6 expression. Arthritis Rheumatol. *66*, 3436–3444.

Scharer, C.D., Barwick, B.G., Youngblood, B.A., Ahmed, R., and Boss, J.M. (2013). Global DNA methylation remodeling accompanies CD8 T cell effector function. J. Immunol. *191*, 3419–3429.

Schenkel, J.M., and Masopust, D. (2014). Tissue-resident memory T cells. Immunity *41*, 886–897.

Schmidt, F., Gasparoni, N., Ebert, P., Gianmoena, K., Cadenas, C., Polansky, J.K., Barann, M., Sinha, A., Froehler, S., Gasparoni, G., et al. (2016). Combining transcription factor affinities with open-chromatin data for accurate gene expression prediction. bioRxiv. Published online October 19, 2016. http://dx.doi.org/10.1101/081935.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. *13*, 2498–2504.

Theocharidis, A., van Dongen, S., Enright, A.J., and Freeman, T.C. (2009). Network visualization and analysis of gene expression data using BioLayout Express(3D). Nat. Protoc. *4*, 1535–1550.

Tokoyoda, K., Zehentmeier, S., Hegazy, A.N., Albrecht, I., Grün, J.R., Löhning, M., and Radbruch, A. (2009). Professional memory CD4+ T lymphocytes preferentially reside and rest in the bone marrow. Immunity *30*, 721–730.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. *28*, 511–515.

Urich, M.A., Nery, J.R., Lister, R., Schmitz, R.J., and Ecker, J.R. (2015). MethylC-seq library preparation for base-resolution whole-genome bisulfite sequencing. Nat. Protoc. *10*, 475–483.

Wallner, S., Schröder, C., Leitão, E., Berulava, T., Haak, C., Beißer, D., Rahmann, S., Richter, A.S., Manke, T., Bönisch, U., et al. (2016). Epigenetic dynamics of monocyte-to-macrophage differentiation. Epigenetics Chromatin *9*, 33.

Wang, L., Park, H.J., Dasari, S., Wang, S., Kocher, J.P., and Li, W. (2013). CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. Nucleic Acids Res. *41*, e74.

Wei, H., Geng, J., Shi, B., Liu, Z., Wang, Y.H., Stevens, A.C., Sprout, S.L., Yao, M., Wang, H., and Hu, H. (2016). Cutting Edge: Foxp1 Controls Naive CD8+ T Cell Quiescence by Simultaneously Repressing Key Pathways in Cellular Metabolism and Cell Cycle Progression. J. Immunol. *196*, 3537–3541.

Youngblood, B., Oestreich, K.J., Ha, S.J., Duraiswamy, J., Akondy, R.S., West, E.E., Wei, Z., Lu, P., Austin, J.W., Riley, J.L., et al. (2011). Chronic virus infection enforces demethylation of the locus that encodes PD-1 in antigen-specific CD8(+) T cells. Immunity *35*, 400–412.

Zou, H., and Hastie, T. (2005). Regularization and variable selection via the elastic net. J. R. Stat. Soc. Series B Stat. Methodol. *67*, 301–320.