Differential roles of epigenetic changes and Foxp3 expression in regulatory T cell-specific transcriptional regulation

Hiromasa Morikawa*[1,2], Naganari Ohkura*[1], Alexis Vandenbon[3], Masayohi Itoh[4], Sayaka Nago-Sato[4], Hideya Kawaji[4], Timo Lassmann[4], Piero Carninci[4], Yoshihide Hayashizaki[4], Alistair R.R. Forrest[4], Daron Standley[3], Hiroshi Date[2], Shimon Sakaguchi[1] and the FANTOM consortium

1. Department of Experimental Immunology, World Premier International Immunology Frontier Research Center, Osaka University, Suita 565-0871, Japan
2. Department of Thoracic Surgery, Kyoto University, 54 Shogoin-Kawahara-cho, Sakyo-ku, Kyoto 606-8507, Japan
3. Department of Systems Immunology, World Premier International Immunology Frontier Research Center, Osaka University, Suita 565-0871, Japan
4. RIKEN Omics Center, Yokohama 230-0045, Japan
*These authors contribute equally to this work
Correspondence: Shimon Sakaguchi
Department of Experimental Immunology, World Premier International Immunology Frontier Research Center, Osaka University, Suita 565-0871, Japan
Mail: shimon@ifrec.osaka-u.ac.jp

Keywords: regulatory T cells, DNA methylation, Foxp3, Cap analysis of gene expression

Abstract

Naturally occurring regulatory T (Treg) cells are engaged in the maintenance of immune tolerance and homeostasis. The development of Treg cells requires both the expression of the transcription factor Foxp3 and the establishment of Treg-type DNA hypomethylation pattern. By transcriptional start site (TSS) cluster analysis, we here assessed possible correlation of genome-wide DNA methylation pattern or Foxp3-binding pattern with Treg-specific gene expression. We found that Treg-specific DNA hypomethylated regions were closely correlated with Treg-up-regulated TSS clusters, whereas Foxp3-binding regions had no significant correlation with either up- or down-regulated clusters, in non-activated Treg cells. On the other hand, in activated Treg cells, Foxp3-binding regions showed a strong correlation with down-regulated clusters. In accordance with the properties, these two regions were mostly different in location on the genome. These results collectively indicate that Treg-specific DNA hypomethylation is conducive to up-regulation in steady state Treg cells whereas Foxp3 expression to down-regulation of its target genes in activated Treg cells. Thus, the combination of the two events is required for the establishment of Treg-specific gene expression and function.

Introduction

Naturally occurring $CD25^+CD4^+$ regulatory T (Treg) cells engage in the maintenance of immunological self-tolerance and homeostasis by suppressing aberrant or excessive immune responses harmful to the host (Sakaguchi et al. 2008). The transcription factor Foxp3, which is specifically expressed in $CD25^+CD4^+$ Treg cells, has been shown to play crucial roles in Treg cell development and function (Fontenot et al. 2003a; Hori et al. 2003; Khattri et al. 2003a). The essential role of Foxp3 is best illustrated by *Foxp3* gene mutations. IPEX (Immune dysregulation, polyendocrinopathy, enteropathy, X-linked) syndrome in humans and Scurfy mutant strain in mice, both bearing *Foxp3* mutations, spontaneously develop severe autoimmunity and systemic inflammation because of developmental or functional failure of natural Treg cells (Bennett et al. 2001; Fontenot et al. 2003a). In addition, ectopic expression of Foxp3 confers suppressor function on peripheral $CD4^+CD25^-$ conventional T (Tconv) cells (Fontenot et al. 2003a; Hori et al. 2003). Foxp3 has therefore been considered as a master regulator or lineage-specification factor for Treg cells.

Alteration of the epigenome appears to be another important factor for establishing the Treg cell lineage. Epigenetic gene modifications such as DNA methylation, histone modifications, nucleosome positioning, and microRNAs are essential for controlling gene expression, in particular, for the stabilization and fixation of cell lineages (Ashraf and Ip 1998; Willingham et al. 2005; Waterland 2006; Jirtle and Skinner 2007; Kim et al. 2009; Musri and Parrizas 2012). We have recently demonstrated that the development of Treg cells requires the establishment of Treg-specific DNA hypomethylation pattern (Ohkura et al. 2012). The process is independent of Foxp3 expression and necessary for $Foxp3^+$ T cells to acquire Foxp3-independent gene expression, lineage stability, and suppressive activity. However, it still remains elusive how the two events, Foxp3 expression and epigenetic modification, contribute to the establishment of Treg-type gene expression.

In this report, we have attempted to elucidate the effects of Treg-specific DNA hypomethylation on Treg-type transcriptional regulation, and also analyzed possible differences between epigenome-dependent transcriptional regulation and Foxp3-dependent one. We show that the role of each regulation is different depending on the state of Treg cell activation; i.e., the genes with Treg-specific DNA

hypomethylation tend to be up-regulated in Treg cells in a steady state whereas the genes with Foxp3-binding regions tend to be down-regulated in activated Treg cells. These results, along with the result of the previous study (Ohkura et al. 2012), strongly support the concept that Treg-specific transcriptional regulation requires the combination of Foxp3 induction and the establishment of Treg-type DNA hypomethylation and that each event has a distinct role in the regulation. These results would facilitate our understanding of the molecular mechanisms by which specific transcriptional networks are established in natural Treg cells to determine and maintain their functions.

This work is part of FANTOM5 project. Data downloads, genomic tools and co-published manuscripts are summarized here http://fantom.gsc.riken.jp/5/.

Results

TSS clusters showing Treg-specific expression

We first obtained whole gene expression and transcriptional start site (TSS) profiles of Treg cells by the cap analysis of gene expression (CAGE) with single-molecule sequencer Heliscope as a part of the FANTOM5 project. CAGE tags are short-length nucleotide sequence tags that allow us to determine where transcription starts, and to examine the whole gene expression accurately, since the method is free from the biases, as inherent in DNA microarray analyses, originated from PCR amplification and sequence similarity. By mapping CAGE tags onto mouse genomes, we found 48,374 and 45,705 potential TSSs in $CD25^{+}CD4^{+}$ natural Treg cells and $CD25^{-}CD4^{+}CD44^{low}$ Tconv cells, respectively (Fig. 1A). The TSS expression profiles were mostly common in the Balb/cA and C57/BL6J mouse strains (Fig. S1). Many of them were located within the gene body regions, referenced by NCBI RefSeq, in both Treg and Tconv cells (Fig. 1B). In the *Foxp3* gene, for example, several potential TSSs were present within the gene body regions, and some of them were transcribed from its antisense strand (Fig. 1C). Interestingly, three TSS clusters were located on the *Foxp3* intron 1 region corresponding to the conserved non-coding sequence 2 (CNS2) region (Zheng et al. 2010), suggesting that RNA polymerase II at CNS2 transcribed bi-directionally a novel class of enhancer RNAs, as previously indicated (Orom et al. 2010). The expression of the bidirectionally transcribed TSS clusters around CNS2 was correlated with the high level expression of *Foxp3* (data not shown), consisting with the report that bidirectional TSS clusters are robust predictors for enhancer activity (Andersson et al. 2013). The potential TSSs identified in Treg or Tconv cells consisted of only one third of the TSSs found in all samples analyzed in FANTOM5, indicating that only a small portion of the genes was specifically expressed in those cells (Fig. 1A). We also found that 23,583 TSSs were non-annotated TSSs, based on the NCBI RefSeq; in particular, 336 non-annotated TSSs were specifically expressed in Treg cells. TSSs with opposite direction to their cognate gene TSSs were also frequently present in a variety of genes, in accordance with the previous reports from FANTOM3 (Katayama et al. 2005). In addition, TSS clusters expressed significantly were correlated with histone H3 lysine 4 trimethylation (H3K4me3) modification, a marker for euchromatin associated with transcriptionally permissive state, and non-expressed TSS clusters with

histone H3 lysine 27 trimethyaltion (H3K27me3), a marker for heterochromatin associated with transcriptionally repressive state (Fig. 1D). This indicates that the location of TSS clusters is tightly linked to a transcriptionally permissive state.

Among more than 45,000 TSS clusters, only 697 and 536 clusters were up- or down-regulated in Treg cells compared with Tconv cells, respectively (Fig 2A). To examine possible involvement of transcription factors in the differential expression, we analyzed transcription factor-binding motifs located in the -400 to +100 bp regions (promoter regions) from these differentially regulated TSS clusters. Interestingly, motifs for Rel, Nf-κb, Tbp and Irf4, which have been classified as T cell receptor (TCR) stimulation-dependent transcriptional activators (Li and Verma 2002), were specifically enriched in the promoter regions of the Treg up-regulated TSS clusters (Fig. 2B). Similar enrichment (79% consistency) was also observed in the samples derived from Balb/cA strain. The results indicate that differential recruitment of transcription factors to the specific promoter regions may contribute to the Treg-type gene expression profiles.

We also examined the effect of TCR stimulation on TSS clusters in Tconv and Treg cells. Whereas a large number of TSS clusters was up- or down-regulated in Tconv cells by TCR stimulation, down-regulated TSS clusters were dominant in Treg cells (Fig. 2C). This tendency was also observed in the Balb/cA strain (data not shown). In addition, while the numbers of up-regulated TSS clusters in Tconv cells were larger than those in Treg cells, more than half of the TSS clusters up-regulated in Treg cells were commonly up-regulated in Tconv cells (Fig. 2D). Down-regulated TSS clusters also showed a similar pattern in those T cells.

We next performed clustering analysis of transcription factor binding motifs in promoter regions by utilizing Over-Representation Index (ORI). ORI for the motifs were calculated regarding the genes up- or down-regulated in Treg or Tconv cells after TCR stimulation with anti CD3/CD28 stimulation or PMA/ionomycin treatment (Fig. 3A). Those significantly overrepresented motifs in up- or down-regulated genes after TCR stimulation were mostly shared between Treg and Tconv cells (Fig. S2A-D). This indicates that Treg and Tconv cells are regulated by similar transcription factors after TCR stimulation although they show different gene expression profiles. Moreover, the up-regulated TSS clusters in Treg and Tconv cells after TCR stimulation were prone to possess CpG islands in their promoters (Fig. 3B).

Treg-specific DNA hypomethylation pattern

We next analyzed genome-wide DNA methylation status in Treg cells by methylated DNA immunoprecipitating sequencing (MeDIP-seq) (Ohkura et al. 2012). As an example of MeDIP-seq, a small region of the *Foxp3* intron 1 locus (corresponding to the CNS2 region) showed differential DNA methylation pattern, which was confirmed by bisulfate sequencing (Fig. 4A). By analyzing the details of the whole DNA methylation pattern, we found that only 301 methylation peaks (0.19% of total peaks) showed Treg-dominant hypomethylation compared with Tconv cells (Fig. 4B). Treg-dominant hypomethylated regions were 500 bp to 1 kbp long (Fig. 4C), and a half of the regions were present within gene body regions, especially in intron 1 or intron 2 (Fig. 4B). Intergenic regions also retained a large portion of the differential peaks; however, the frequency of appearance was very low compared with the frequency within the gene body regions. In addition, the differences were rarely detected in CpG islands or 5' upstream regions of TSSs (Fig. 4B, and data not shown). Thus, the differential DNA methylation appears to be predominantly established in gene body regions, and the methylation status of promoters or CpG islands per se may not play a pivotal role in Treg-specific transcriptional regulation. These observations are consistent with the previous reports showing that most differentially methylated regions are located in CpG poor regions distal from annotated promoters (Schmidl et al. 2009).

Correlation between Treg-specific TSS clusters and Treg-specific hypomethylated regions

To examine possible effects of Treg-specific DNA hypomethylation pattern on Treg-type gene expression, we analyzed correlation between TSS clusters and the location of Treg-specific hypomethylated regions. TSS clusters located in the regions upstream of Treg-specific DNA demethylated regions (TSDR) were prone to be up-regulated in Treg cells compared with Tconv cells (Fig. 5A), whereas TSS clusters on the downstream regions of TSDR were rarely detected. Interestingly, significant correlations between TSDR and Treg-specific up-regulation, such as those of *Foxp3, Ikzf2, Ikzf4* and *Il2ra,* were observed independently of TCR stimulation and the distance from TSS clusters to TSDR. Calculate significances based on hyper geometric distribution before and after TCR stimulation were $p=5.29\times10^{-11}$ and $p=3.04\times10^{-20}$,

respectively. These results indicate that TSDR within gene body regions might function as enhancer regions and thereby preferentially contribute to the transcriptional regulation. In addition, DNA methylation status was reversely correlated with the DNaseI hypersensitive regions in 5' flanking regions of TSS clusters (Fig. 5B), indicating that DNA hypomethylated regions prone to possess open chromatin structures that allow transcription factors assemble on the regions. Thus, the DNA methylation status would be important for the Treg-specific transcriptional regulation irrespective of their activation states.

We also examined possible involvement of transcription factors in hypomethylation-mediated Treg-specific gene expression, since differences in DNA methylation and histone modification status could determine the accessibility of transcription factors to the loci (Thurman et al. 2012). Search for enriched DNA sequence motifs within the Treg-specific hypomethylated regions revealed that motifs for Myb, Creb1, Irf5, Ets1, Arnt, Hif1a, Mfi2, Atf1 and Sp100 were enriched within TSDR compared with control genomic regions (Fig. 5C). Given that Ets1 and Creb1 bind to their target sites in a demethylation-dependent manner (Kim et al. 2009; Polansky et al. 2010), the results suggest that some of these transcription factors activate their target genes via direct binding to demethylated TSDR.

Foxp3-binding regions in Treg cells

Assuming that Foxp3 is a lineage determination factor for Treg cells, we next examined how Foxp3 contributed to the Treg-specific gene expression. Foxp3-binding regions (Samstein et al. 2012) were predominantly present in gene body regions, in particular around the TSSs. Transcription factor-binding motifs for Foxo3, Runx1, Irf4, and Ets1 were enriched within 500 bp regions from the Foxp3 binding sites, being consistent with the observations that Foxp3 can associate with some of those transcription factors in Treg cells (Fig. 6A) (Rudra et al. 2012).

The regions around the Foxp3-binding sites tended to be demethylated and highly sensitive to DNaseI in both Treg and Tconv cells (Fig. 6B), suggesting that the accessibility of the Foxp3-binding sites is similar between Treg and Tconv cells. Likewise, the regions bound by Ets1, Foxo1 and Elf1 in Treg cells showed similar profiles in DNaseI hypersensitivity and DNA hypomethylation status between Treg and Tconv cells (Fig. S3). These results are in line with the previous report that enhancer

landscape was similar between Treg and Tconv cells (Samstein et al. 2012).

Respective contributions of Treg-specific DNA hypomethylation and Foxp3 expression to Treg-specific TSS clusters

Foxp3-binding regions and Treg-specific DNA hypomethylated regions were mostly different in Treg cells (Fig. S4A), and were partially correlated with the Treg-specific TSS clusters. To examine each contribution to the Treg-type gene expression, we compared the gene expression profiles of TSSs located on the adjacent regions of TSDR or Foxp3-bidning sites.

Firstly, to examine whether contribution of TSDR to Treg-specific gene regulation is independent of Foxp3 expression, we analyzed expression profiles of genes locating within 10 kb from TSDR in *Foxp3*^{gfpko} mice (Samstein et al. 2012). In these mice, the inserting Green Fluorescent Protein (GFP) to the *Foxp3* locus marked Treg-committed cells, although the *Foxp3* gene itself was disrupted. Clustering of microarray data revealed that a fraction of genes associated with Treg function were similarly regulated in both Foxp3-null Treg cells derived from *Foxp3*^{gfpko} mice and wild-type Treg cells (Fig. 7A). For example, *Foxp3, Ikzf2*, *Ikzf4* and *Il2ra*, all of which were shown to be important for Treg-specific gene regulation and function (Fontenot et al. 2003b; Hori et al. 2003; Khattri et al. 2003b; Sugimoto et al. 2006; Pan et al. 2009), were commonly up-regulated in both. It indicates that a fraction of genes possessing TSDR are properly controlled without Foxp3.

Next we calculated cumulative distribution of up- or down-regulated TSS clusters around the binding sites of Foxp3 in Treg cells. TSS clusters around the Foxp3-binding sites (± 500 bp regions of the Foxp3-binding sites) scarcely correlated with the Treg-specific gene regulation in steady state Treg cells. However, those clusters were highly correlated with down-regulation, but not up-regulation, after TCR stimulation in Treg cells (Fig. 7C). As shown in Fig. 7B, the frequency of the down-regulated TSS clusters locating near the Foxp3-binding sites was higher than those calculated by all TSS clusters as control, while the frequency of up-regulated ones was similar to the control values (Fig. 7B, S5A). In addition, the significant correlation between the location and the transcriptional regulation was not observed as to the other Treg-associated transcription factors including Ets1, Elf1, Foxo1 and Cbfb (Fig. S4B). These results suggest that Foxp3-binding sites are highly correlated with the TCR

stimulation-dependent down-regulation in Treg cells. We also observed that majority of up-regulated TSS clusters in Tconv cells were not up-regulated in Treg cells after TCR stimulation (Fig. S5B), suggesting that TCR stimulation-dependent up-regulation was mostly inhibited in Foxp3-expressing Treg cells. In contrast, TSS clusters with TSDR were dominantly detected in the up-regulated ones, not in the down-regulated ones, of steady state Treg cells (Fig. 7D). This correlation was scarcely retained after TCR stimulation.

Altogether, in Treg cells, TSDR chiefly serves up-regulation of gene expression in a steady state, whereas Foxp3 mainly engages in gene repression following TCR stimulation. And also, TSDR and Foxp3 had respective roles in the transcriptional regulation of Treg cells (Fig. 8).

Discussion

In this study, we demonstrated that Treg-specific DNA hypomethylated regions and Foxp3-binding regions were distinctly correlated with TSS clusters showing Treg-specific expression. We have previously demonstrated that Treg-type DNA hypomethylation pattern is required for Foxp3$^+$ T cells to acquire Treg-specific gene expression, lineage stability, and the suppressive function of Treg cells. In line with the findings, the present results showed that Treg-specific TSS clusters located in the adjacent regions of TSDR were mostly different from those of the Foxp3-binding regions. Moreover, transcription factor-binding motifs found in the hypomethylated regions were also different from those in Foxp3-binding regions. These findings strongly support the notion that Treg cell development is achieved by the combination of the establishment of Treg-specific DNA hypomethylation pattern and Foxp3 induction.

Functional differences between Foxp3-dependent regulation and TSDR-mediated one would be important for the understanding of the Treg-specific gene regulation. We observed that TSS clusters possessing TSDR were prone to be up-regulated in non-activated Treg cells, while TSS clusters possessing Foxp3-biding sites tended to be down-regulated in activated Treg cells. These findings are consistent with the observation that suppressive function of Foxp3 is only evident after TCR stimulation (Takahashi et al. 1998) and that DNA hypomethylation is linked to the transcriptionally permissive state, which enables transcription factors to bind to their target loci (Thurman et al. 2012). Recently, it was revealed that active modification of the chromatin landscape was established during the course of the Treg cell development, and that Foxp3 contributed to the Treg function by exploiting the preformed enhancer network (Samstein et al. 2012). This is also consistent with our observation that a limited number of loci was specifically demethylated in Treg cells, and that the DNA demethylation pattern was completely established independently of Foxp3 expression (Ohkura et al. 2012). In addition, Schmidl et al. show that FOXP3 binds to regions that are enhancers in Treg cells but also at regions that normally enhancers in Tconv cells in human (Schmidl et al. , manuscript in preparation). Thus, these findings suggest that the establishment of Treg-type DNA hypomethylation pattern would be a primary factor for

establishing the preformed Treg-type enhancer network in Treg cells, and that Foxp3 functions upon the established enhancer landscape. In accordance of this idea, several reports showed that Foxp3-dependent transcriptional regulation in Tconv cells is largely diverged from the transcriptional regulation observed in Treg cells, especially in steady state Treg cells. For example, some of the genes specifically expressed in Treg cells showed no correlation with Foxp3 expression (Gavin et al. 2007). Analysis of *Foxp3*^{gfpko} mice, in which the *Foxp3* gene is disrupted by inserting the *GFP* gene, demonstrated that Foxp3^{-}GFP^{+} T cells expressed several Treg signature genes (Gavin et al. 2007; Lin et al. 2007). Thus, while Foxp3 has been referred to as a master regulator for Treg cells, the key factor for establishing the Treg-type whole gene expression profiles would be the preformed enhancer landscape including Treg-specific DNA hypomethylation pattern.

In contrast, Fu et al. recently reported that Treg-type gene expression profiles could be recapitulated by the combination of Foxp3 and several transcription factors (i.e., quintet transcription factors: Eos, Irf4, Satb1, Lef1 and Gata1) in Tconv cells (Fu et al. 2012). It suggests that the preformed Treg-type enhancer landscape is not essential for the recapitulation of Treg-type gene expression. One possible explanation of the discrepancy is that the Treg-type DNA hypomethylation mainly contributes to the stability of gene expression and the expression of several key transcription factors including quintet transcription factors. In fact, TSDR were found in genes for some of the quintet transcription factors and key transcription factors for Treg cell function (Ohkura, et al. 2012, and data not shown). Further study would be required for clarifying the interconnections between DNA hypomethylation-mediated control and transcription factor-mediated one in Treg cells.

Treg-specific RNA transcripts might be involved in specific functions in Treg cells. We have identified 23,583 un-annotated RNA transcripts in Treg cells, and some of them were specifically up-regulated in Treg cells, and expressed independently of TCR stimulation. It is therefore likely that Treg cells express thousands of Treg-specific unknown molecules, and that those molecules might contribute to specific phenotypes observed in steady state Treg cells. In addition, many of the non-annotated TSS clusters are located within intergenic regions, and some of non-annotated TSS clusters locating on gene body regions were antisense transcripts. These unannotated TSS clusters are

also identified in human Treg cells, and some of them are confirmed as splicing variants of Treg signature genes, such as *Foxp3* and *Ctla4*, by rapid amplification of cDNA ends-PCR (Schmidl et al. , manuscript in preparation). Since intergenic regions and antisense strands of gene body regions are prone not to possess long open reading frames, the majority of these non-annotated TSS clusters seem to reflect the expression of non-coding RNAs. As in the case of other cell-types (Guttman et al. 2011), non-coding RNAs up-regulated in Treg cells might play important roles in the development and function of Treg cells. In fact, we found that Treg cells exhibited several antisense RNA transcripts from the *Foxp3* CNS2 region, which has been shown to have a pivotal role in Foxp3 induction and its stability. Given that antisense RNA products function as a novel class of enhancer RNAs for their cognate gene expressions (Orom et al. 2010), the observation supports the notion that novel RNA transcripts in Treg cells, such as antisense transcripts of CNS2, would participate in the establishment of Treg-type gene expression.

We found that several transcription factor-binding motifs, such as those for Ets1 and Creb1, were frequently observed in Treg-specific DNA hypomethyalted regions. It has been postulated that DNA methylation inhibits the recognition of DNA by some proteins (Prokhortchouk and Defossez 2008; Defossez and Stancheva 2011), and is generally associated with gene repression (Weber and Schubeler 2007). In accordance with this notion, Ets1-binding to the *Foxp3* CNS2 region was only observed when CNS2 was demethylated (Polansky et al. 2010). Creb/Atf was also shown to bind to the *Foxp3* CNS2 region in a demethylation-dependent manner (Kim and Leonard 2007). In addition, we found that H3K4me3 modification, an euchromatin marker associated with transcriptionally permissive state, accumulated in the majority of Treg-specific demethylated regions. Thus, Treg-specific demethylation together with accompanied epigenetic modifications would be prerequisite for specific gene expression via facilitating the binding of transcription factors to specific loci. These epigenetic changes would consequently lead to the specific gene expression, and the augmentation of its stability. On the other hand, in the adjacent regions of Foxp3-binding sites, different sets of binding motifs for transcription factors, such as Gabpa, Elk4 and Spi1 were frequently detected. Motifs for Foxo3 and Runx1, both of which were shown to associate with Foxp3 protein in Treg cells (Ono et al. 2007; Kerdiles et al. 2010;

Ouyang et al. 2010), were also enriched in the Foxp3-binding regions. These observations suggest that Foxp3 and its associating transcription factors are assembled on the preformed enhancer regions in an activation-dependent manner. Thus, TSDR and Foxp3 would respectively contribute to the Treg-type gene expression via utilizing different sets of transcription factors.

In conclusion, we have shown that Treg cells possess genome-wide Treg-type DNA hypomethylation pattern in addition to the expression of Foxp3. Importantly, Treg-type DNA hypomethylation highly correlates with transcriptional up-regulation in steady state Treg cells. It suggests that while functional transcription factors are mostly shared in Treg and Tconv cells, chromatin structures determined by specific epigenetic changes would contribute to gene expression by enabling transcription factors to on specific loci. In contrast, TSS clusters with Foxp3-bindng sites correlate with transcriptional down-regulation in activated Treg cells. These Foxp3-binding sites were mostly similar between Treg and Tconv cells in DNA methylation status and DNaseI hypersensitivity. It suggests that specifically expressed Foxp3 controls its target genes under same chromatin condition in an activation-dependent manner (Fig. 8). Altogether, these findings suggest that both Treg-specific hypomethylation and Foxp3 expression have respective roles in the development of Treg cells. This model of distinct contributions of epigenetic changes and transcription factors to natural Treg cell development can be applied for peripheral generation of stable induced Treg cells, and exploited to control a variety of physiological or pathological immune responses via targeting Treg cell generation and its functional stability. The epigenetic information along with transcription factor-dependent transcriptional networks would facilitate our understanding of the developmental process and function of Treg cells.

Methods

Mice, cell sorting and cell culture

C57BL/6JJcl mice and Balb/cAJcl mice were purchased from CLEA Japan (Tokyo, Japan). $CD4^+$ T cells were isolated from splenic and lymph node as previously described (Hori et al. 2003). $CD4^+CD25^+$ T cells (Treg cells) and $CD4^+CD25^-CD44^{low}$ T cells (Tconv cells) were purified by sorting with a cell sorter (MoFlo, Beckman Coulter). For *in vitro* TCR stimulation of Tconv cells, plate coated anti-CD3 (1mg/ml) and anti-CD28 (1mg/ml) for 6 hrs or phorbol 12-myristate 13-acetate (20ng/ml) and ionomycin (1μM) for 2 hrs were used.

Antibodies

PE-conjugated anti-Il2ra (PC61), PE-Cy7-cojugated anti-CD4 (RM4.5), FITC-conjugated anti-CD44 (IM7), APC-conjugated anti-CD8a (53-6.7), anti–B220 (RA3-6B2), anti-CD16/32(2.4G2), anti-NK1.1 (PK136) were obtained from BD PharMingen, Biolegend or eBioscience. Anti-CD3 (2C11) and anti-CD28 (37.51) were used for in vitro T cell stimulation. Mouse recombinant IL-2 was a gift from Shionogi Co. (Osaka, Japan).

RNA preparation

Total RNAs for CAGE were extracted from sorted cells using miRNeasy mini kit (Qiagen, Japan).

CAGE Tag expression profiling

CAGE tag sequencing of each cells and sequence alignment to reference genome (mm9) were performed as a part of FANTOM5 (Table S1) (Forrest et al. 2013). Each TSS cluster expression level were calculated from normalized tag count at each robust TSS clusters defined in FANTOM5.

Processing of CAGE data

Regions with differential levels of transcription initiation were defined based on a method described by Audic and Claverie (Audic and Claverie 1997). In brief, the probability of observing y mapped CAGE reads in a certain TSS region in sample 1,

given x, the number of mapped reads in sample 2 is given by:

$$p(y|x) = \left(\frac{N_2}{N_1}\right)^y \frac{(x+y)!}{x!y!\left(1+\frac{N_2}{N_1}\right)^{(x+y+1)}} \quad [1]$$

where N_1 and N_2 represent the total number of mapped CAGE tags, in sample 1 and sample 2, respectively. A p value for differential transcription initiation in the TSS region was calculated from the cumulative distribution of these probabilities:

$$P(y|x) = \sum_{i=y}^{\infty} p(i|x) = 1 - \sum_{i=0}^{y-1} p(i|x) \quad [2]$$

Finally, a correction for multiple testing was done by multiplying $P(y|x)$ by the number of tests (= 61,949, e.g. the number of TSS clusters containing at least 1 mapped read in the samples of interest). The threshold p value for significant differential transcription initiation was set to 0.01 (after correction for multiple testing).

Regulatory motif over-representation analysis

A set of 543 position weight matrices (PWMs) was prepared from the Jaspar database (Bryne et al. 2008), including binding preferences for human, mouse, and rat transcription factors (TFs), as well as core promoter motifs. The binding motif for Foxp3 was not included in the Jaspar database, so it was inferred from FoxpP3 ChIP data (see section bellow). For each PWM a threshold score was set in a way that resulted in about 1 predicted site per PWM per 5 kb.

For the further analysis, we made a distinction between the enrichment of regulatory motifs in 1) a set of promoters relative to the genomic set of promoters, 2) a set of promoters relative to another set of promoters, and 3) a set of regions relative to the genome-wide tendencies.

1. ORI analysis of a set of promoters relative to the genomic set of promoters

Transcription start sites (TSSs) of CAGE data were decided in FANTOM5 as robust peaks (version mm9 of the mouse genome was used), and promoter regions were

defined as the 500 bps upstream of summits of the TSS clusters. For a given set of promoter sequences S for each PWM motif, the Over-Representation Index (ORI) was calculated as follows (Bajic et al. 2004):

$$\mathrm{ORI}_S = \frac{Density_S}{Density_G} \times \frac{Proportion_S}{Proportion_G} \qquad [3]$$

where $Density_S$ and $Density_G$ represent the average number of predicted TFBSs per sequence for the PWM of interest in the set of promoters S and in the genome-wide set of promoters G, and $Proportion_S$ and $Proportion_G$ represent the proportion of sequences containing at least 1 predicted TFBS for the PWM of interest in the set of promoters S and in the genome-wide set of promoters G, respectively.

The significance of each ORI_S value was evaluated using a random sampling approach, whereby from the genome-wide set of promoters a set of promoters was randomly sampled, and a $\mathrm{ORI}_{Sampled}$ was calculated. In order to limit the influence of GC content biases, this sampling was done in a way that the sampled set contains the same number of sequences that are annotated to be associated with CpG islands as the original set S. Sampling was done 1 million times, and the p values of ORI_S was defined as the ratio of $\mathrm{ORI}_{Sampled}$ values being equal to or higher than ORI_S.

2. ORI analysis of a set of promoters relative to another set of promoters

The approach is basically identical to the one using the genome-wide set of promoters as reference, except that now ORI_S is defined as

$$\mathrm{ORI}_S = \frac{Density_S}{Density_{S'}} \times \frac{Proportion_S}{Proportion_{S'}} \qquad [4]$$

where S' is a second set of promoters, different from the genome-wide set, used as reference. Here too, a random sampling approach is used for estimating the significance of ORI_S. Two sets of sequences are sampled, one corresponding to S and one to S', and from the predicted TFBSs in both sets $\mathrm{ORI}_{Sampled}$ is calculated. Importantly, here too, the randomly sampled sets contain the same number of sequences associated with CpG islands as in S and S', respectively. Sampling was done 10,000 times, and the p value of ORI_S was defined as described above.

3. ORI analysis a set of regions relative to the genome-wide tendencies

In this case, predicted TFBSs in the regions from position -500 to +500 relative to the summit the mapped read peak representing differentially demethylated regions or Foxp3-bound regions were used. We observed that, in practice, these regions had specific GC content profiles that were hard to account for using random sampling approaches. Instead, we used the genome-wide tendencies of each motif to construct a large reference set of artificial sequences with the same GC content profile as the input set of sequences (See Supplementary material for a detailed description). The ORI_S was defined as:

$$\mathrm{ORI}_S = \frac{Density_S}{Density_{G'}} \times \frac{Proportion_S}{Proportion_{G'}} \qquad [3]$$

where G' represents the large set of artificial sequences. The significance of ORI_S was estimated as described above, by randomly sampling from G'.

In the 20 sets of promoter sequences of genes with differential expression between Treg and Tconv cells, and between non-stimulated and stimulated Tconv and Treg cells, ORI_S values and corresponding p values were calculated for all 543 PWMs. A subset of 117 PWMs that had ORI_S>2 and p value <0.001 in at least one of the 20 sets of promoters was picked up for the construction of a heat-map. The heat-map was constructed by hierarchical clustering using the pairwise Pearson correlation coefficient of log(ORI) values as measure of similarity between pairs of PWMs (rows). The 20 sets of promoters were clustered in the same way (columns). Color intensities in the heat-map reflect ORI values.

Prediction of the Foxp3-binding motif

Biding sites sequences were obtained from published Foxp3-binding sites data (Zheng et al. 2007). Motif sequence of Foxp3 was discovered by Weeder (Pavesi and Pesole 2006) and identify the most enriched motif as the motif of Foxp3 (Fig. S6).

DNA methylation analysis

Raw MeDIP-seq data of Treg and Tconv cells were obtained from the previous report (Ohkura et al. 2012). Each tag data were mapped to the UCSC mouse genome mm9

using Bowtie program with default parameters (Langmead et al. 2009). To detect the specifically demethylated regions in Treg cells against Tconv cells, we used MACS1.4 program (Feng et al. 2012) for peak calling of Tconv cells against Treg cells with the Pvalue option=1e-15 as a cutoff for peak detection (Table S2). The threshold is set severely lower to detect definitely TSDR like CNS2 in *Foxp3*.

ChIP-seq data analysis

Raw ChIP-seq data of H3K4me3, H3K27me3 histone modifications in mouse T cells (SRP000706) and raw ChIP-seq data of Ets1, Foxo1, Elf1 and Cbfb in mouse T cells (SRP015626) were obtained from Short Read Archive database. DNaseI HS-seq data were obtained from the ENCODE repository. The tag data were mapped to the UCSC mouse genome mm9 using Bowtie program with default parameters. MACS1.4 program was used to identify the significant peaks of each mapped data. Top 2000 peaks were used as representative binding site of each transcription factor in further analysis. For analysis of Foxp3, we used 2,886 peaks reported by the previous report as its binding sites (Samstein et al. 2012). HOMER program (Heinz et al. 2010) was used to calculate each ChIP-seq tag distribution at selected regions.

Data access

The sequence tags of Methylated DNA Immunoprecipitation can be downloaded from the DDBJ database (accession number is DRA000868).

Acknowledgements

We thank R. Ishii for technical assistance. This study was supported by Grant-in-Aid for Specially Promoted Research 20002007 to S.S.. FANTOM5 was made possible by a Research Grant for RIKEN Omics Science Center from MEXT to Yoshihide Hayashizaki and a Grant of the Innovative Cell Biology by Innovative Technology (Cell Innovation Program) from the MEXT, Japan to YH. We would like to thank all members of the FANTOM5 consortium for contributing to generation of samples and analysis of the data-set and thank GeNAS for data production.

Author contribution

H.M performed experiments, computational analyses and wrote the manuscript. N.O contributed to planning, supervision and manuscript writing. A.V contributed to computational analyses. D.S and H.D contributed to planning and supervision. M.I. and S.N-S. were responsible for CAGE data production, T.L. was responsible for tag mapping, H.K. managed the data handling, P.C., Y.H. and A.R.R.F. were responsible for FANTOM5 management and concept. S.S initiated, planned and supervised the study and contributed to manuscript writing.

Disclosure declaration

The authors declare no competing interests.

Figure legends

Figure 1. TSS clusters identified in Treg and Tconv cells

(A) Distribution of TSS clusters identified in Treg and Tconv cells by CAGE. X- and Y-axes indicate TSS clusters and normalized expression levels, respectively. (B) Annotation of TSS clusters identified in Treg or Tconv cells. (C) TSS clusters of the *Foxp3* locus. Upper and lower peaks showed CAGE tags originated from the sense and antisense strands, respectively. Arrowheads indicate robust TSS clusters defined in the FANTOM5 main paper (Forrest et al. 2013). A right panel show the magnification of the *Foxp3* CNS2 locus. TSS locations determined by FANTOM5 are indicated by vertical lines. (D) The positional relationships between H3K4me3/H3K27me3 modification and TSS clusters. Density distribution of H3K4me3 or H3K27me3 ChIP-Seq tags (Wei et al. 2009) was calculated on the distance from TSS clusters.

Figure 2. Transcriptional regulation is different between Treg and Tconv cells.

(A) Comparison of TSS cluster expression between Tconv (X-axis) and Treg (Y-axis) cells. Red and blue dots indicate significantly up- or down-regulated TSS clusters in Treg cells, respectively. Tpm represents tags per million tags. (B) Shown are enriched transcription factor-binding motifs in the promoter regions of TSS clusters that are up-regulated in Treg cells. Enriched transcription factor-binding motifs, their belonging family names, over-representation index (ORI) scores, and p-values, are shown. (C) Significantly up- (red) or down-regulated (blue) TSS clusters after TCR stimulation in Treg (left panel) or Tconv (right panel) cells. X- and Y-axes indicate TSS cluster expression before and after TCR stimulation, respectively. (D) Venn diagrams illustrating the overlap in TSS clusters. Up- or down-regulated TSS clusters after TCR stimulation in Treg cells were compared with those in Tconv cells.

Figure 3. Transcription factors associated with TCR stimulation-dependent regulation are mostly shared between Treg and Tconv cells.

(A) A heat-map shows ORI values of transcription factor-binding motifs enriched in the promoter regions of TSS clusters significantly up- or down-regulated by TCR stimulation. X-axis indicates sets of up- or down-regulated TSS clusters after TCR

stimulation under different conditions (described in below). Y-axis indicates motifs significantly enriched after TCR stimulation in at least one sample. Abbreviations: CD3/28, anti-CD3 and anti-CD28 antibody stimulation; PMA/iono, stimulation with phorbol 12-myristate 13-acetate (PMA) and ionomycin; up, up-regulated TSS clusters after TCR stimulation; down, down-regulated ones; Treg>, up-regulated TSS clusters in non-stimulated Treg cells compared to Tconv cells; Treg<, down-regulated ones in non-stimulated Treg cells; C57BL6J, samples obtained from mouse C57BL6J strain; Balb/cA, those from Balb/cA. The upper panel shows a heat-map scale with its distribution. (B) The ratio of TSS clusters possessing CpG islands in their promoters. Treg>Tconv and Treg<Tconv represent TSS clusters showing up- and down-regulated in Treg cells compared to Tconv cells, respectively. “Up in stim” and “down in stim” represents TSS clusters up- or down-regulated after TCR stimulation, respectively.

Figure 4. Treg cells exhibit Treg-specific DNA hypomethylated regions.

(A) DNA methylation pattern of the *Foxp3* locus by MeDIP-seq. Confirmation of the differences by bisulfite sequencing are also shown in below. Black and open circles indicate methylated and unmethylated CpG residues. Each column represents each CpG residue in the *Foxp3* CNS2 region. (B) The ratio of Treg-specific DNA demethylated regions (TSDR) within total DNA methylation peaks of Tconv cells determined by MeDIP-seq and MACS1.4 program. Right panels show the annotation of TSDR in location and property (CpG island or not). (C) Accumulated numbers of TSDR sorted by their nucleotide length on the genome.

Figure 5. TSDR are correlated with TSS clusters up-regulated in steady state Treg cells.

(A) Expression profiles of TSS clusters sorted by positional relation to TSDR. Red and blue dots indicate significantly up- or down-regulated TSS clusters in Treg cells, respectively. Upper and lower panels show expression profiles without and with TCR stimulation, respectively. (B) Relationships between DNA methylation status and DNaseI hypersensitivity. Heat-maps show normalized tag counts of MeDIP-seq (X-axis) and DNaseI-HS-seq (Y-axis) within 500 bp from all TSS clusters. Upper and lower panels in right show the distribution of TSS clusters with more than 10 tags per

million tags (tpm) expression and those without significant expression in Treg cells, respectively. (C) Transcription factor-binding motifs significantly enriched within 500 bp from TSDR.

Figure 6. Chromatin status of Foxp3-binding sites is similar between Treg and Tconv cells.

(A) Transcription factor-binding motifs significantly enriched within 500 bp from Foxp3-binding sites. (B) Heat-maps show normalized tag counts of MeDIP-seq (X-axis) and DNaseI-HS-seq (Y-axis) within 500 bp from Foxp3-binding sites of Treg cells, in Treg (left panel) and Tconv cells (right panel).

Figure 7. TSDR and Foxp3 respectively contribute to Treg-specific gene regulation.

(A) Gene expression profiles (Samstein et al. 2012) were compared among Tconv, Treg, and Foxp3-null Treg cells. Shown are the profiles of genes associated with TSS clusters that located within 10 kbp from TSDR and showed up-regulation in Treg cells compared to Tconv cells. A upper panel show a heat-map scale with its distribution. (B) Cumulative distribution of TSS clusters within 1 Gbp regions from Foxp3-binding sites. Blue and red lines indicate TSS clusters significantly down- and up-regulated TSS clusters in Treg cells after TCR stimulation, respectively. Cumulative distribution of all TSS clusters is also shown as a negative control (gray line). (C, D) Numbers of TSS clusters locating within 10 kbp from Foxp3-binding sites (C) or TSDR (D). Left columns showed numbers of up- or down-regulated TSS clusters in steady state Treg cells compared to steady state Tconv cells (Treg vs. Tconv). Right columns showed those in activated Treg cells compared to steady state Treg cells (Stim Treg vs. Treg). Samples of activated Treg cells were obtained from Treg cells stimulated with anti-CD3 and anti-CD28 antibodies for 6 hrs.

Figure 8. Both chromatin structures and transcription factors regulate Treg-specific gene expression coordinately.

Models of Treg-specific gene regulation by chromatin structures and transcription factors. In a steady state, Treg-specific gene regulations are mainly dependent on chromatin structures specifically established in Treg cells. In contrast, under activated

conditions, Foxp3 becomes functional, and contributes to the gene regulation, especially to the repression of its target genes.

Figure S1. TSS cluster expression profiles are mostly conserved in the C57BL/6J and Balb/c A mouse strains.

Comparison of TSS cluster expression profiles between C57/BL6J (X-axis) and Balb/cA (Y-axis) mouse strains. Expression profiles of Treg and Tconv cells were examined under non-stimulated, CD3/CD28-stimulated (CD3/28), and PMA/ionomycin-stimulated (PMA/iono) conditions.

Figure S2. Over representative regulatory motifs in promoters of TSS clusters regulated by TCR stimulation.

Shown are significantly over representative regulatory motifs in promoter regions of TSS clusters that were regulated by TCR stimulation. Enriched transcription factor-binding motifs, their ORI scores, and p-values, are shown. (A, B) Motifs found in promoters of TSS clusters up- (A) or down-regulated (B) after TCR stimulation in Treg cells. (C, D) Motifs found in promoters of TSS clusters up- (C) or down-regulated (D) after TCR stimulation in Tconv cells.

Figure S3. DNA methylation status and DNaseI hypersensitivity of several transcription factor-biding sites.

Heat-maps show normalized tag counts of MeDIP-seq (X-axis) and DNaseI-HS-seq (Y-axis) of the indicated transcription factor-binding sites (500 bp region) in Treg (A) and Tconv (B) cells.

Figure S4. TSDR and Foxp3-binding sites are mostly different in location.

(A) Venn diagram illustrating the commonality between genomic regions having Foxp3-binding sites and those having TSDR. (B) Cumulative distribution of TSS clusters within 1 Gbp regions from the indicated transcription factor-binding sites. Red and blue lines indicate TSS clusters significantly up- and down-regulated TSS clusters in Treg cells after TCR stimulation, respectively. Cumulative distributions of all TSS

clusters are also shown as a negative control (gray line).

Figure S5. Expression profiles of TSS clusters related to Foxp3-binding sites.
(A) Expression profiles of TSS clusters sorted by the distance from Foxp3-binding sites to TSS clusters. Upper and lower panels show TSS cluster expression without and with TCR stimulation, respectively. Red and blue dots indicate significantly up- or down-regulated TSS clusters in Treg cells compared to Tconv cells, respectively. (B) TSS clusters significantly up-regulated in Tconv cells after TCR stimulation were plotted onto the Treg-expression profiles compared between non-stimulated and stimulated conditions (upper panels). TSS clusters up-regulated in Treg cells after TCR stimulation were also plotted onto the Tconv-expression profiles under non-stimulated and stimulated conditions (lower panels).

Figure S6. Predicted Foxp3-binding sequence motif.
(A) Foxp3-binidng sequence motif was predicted by the sequence information of the Foxp3-binding sites. (B) The frequency matrix for Foxp3-binding sequence is also shown.

Table S1. All samples used in this study.
Cell type, background, and uniquely mapped CAGE tag counts are shown.

Table S2. Data production and analysis of MeDIP-seq data.
(A) Data production of MeDIP-seq data derived from Treg and Tconv cells using Hiseq-2000 sequencer. (B) Read alignment of MeDIP-seq data by bowtie. (C) Peak analysis of all MeDIP-seq data by MACS 1.4.

References (reference は legends の前に移動予定)

Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Zhao X, Schmidl C, Valen E, Li K. 2013. Systematic in-vivo characterization of active enhancers across the human body. *Nature*.

Ashraf SI, Ip YT. 1998. Transcriptional control: repression by local chromatin modification. *Curr Biol* **8**(19): R683-686.

Audic S, Claverie JM. 1997. The significance of digital gene expression profiles. *Genome Res* **7**(10): 986-995.

Bajic VB, Choudhary V, Hock CK. 2004. Content analysis of the core promoter region of human genes. *In silico biology* **4**(2): 109-125.

Bennett CL, Christie J, Ramsdell F, Brunkow ME, Ferguson PJ, Whitesell L, Kelly TE, Saulsbury FT, Chance PF, Ochs HD. 2001. The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. *Nat Genet* **27**(1): 20-21.

Bryne JC, Valen E, Tang MH, Marstrand T, Winther O, da Piedade I, Krogh A, Lenhard B, Sandelin A. 2008. JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. *Nucleic Acids Res* **36**(Database issue): D102-106.

Defossez PA, Stancheva I. 2011. Biological functions of methyl-CpG-binding proteins. *Progress in molecular biology and translational science* **101**: 377-398.

Feng J, Liu T, Qin B, Zhang Y, Liu XS. 2012. Identifying ChIP-seq enrichment using MACS. *Nature protocols* **7**(9): 1728-1740.

Fontenot JD, Gavin MA, Rudensky AY. 2003a. Foxp3 programs the development and function of CD4+CD25+ regulatory T cells. *Nat Immunol* **4**(4): 330-336.

-. 2003b. Foxp3 programs the development and function of CD4+CD25+ regulatory T cells. *Nat Immunol* **4**(4): 330-336.

Forrest ARR, Kawaji H, Rehli M, Baillie JK, Hoon MJLd. 2013. A promoter level mammalian expression atlas. *Nature*.

Fu W, Ergun A, Lu T, Hill JA, Haxhinasto S, Fassett MS, Gazit R, Adoro S, Glimcher L, Chan S et al. 2012. A multiply redundant genetic switch 'locks in' the transcriptional signature of regulatory T cells. *Nat Immunol* **13**(10): 972-980.

Gavin MA, Rasmussen JP, Fontenot JD, Vasta V, Manganiello VC, Beavo JA, Rudensky AY. 2007. Foxp3-dependent programme of regulatory T-cell differentiation. *Nature* **445**(7129): 771-775.

Guttman M, Donaghey J, Carey BW, Garber M, Grenier JK, Munson G, Young G, Lucas AB, Ach R, Bruhn L et al. 2011. lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* **477**(7364): 295-300.

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular cell* **38**(4): 576-589.

Hori S, Nomura T, Sakaguchi S. 2003. Control of regulatory T cell development by the transcription factor Foxp3. *Science* **299**(5609): 1057-1061.

Jirtle RL, Skinner MK. 2007. Environmental epigenomics and disease susceptibility. *Nat Rev Genet* **8**(4): 253-262.

Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M, Nishida H, Yap CC, Suzuki M, Kawai J et al. 2005. Antisense transcription in the mammalian transcriptome. *Science* **309**(5740): 1564-1566.

Kerdiles YM, Stone EL, Beisner DR, McGargill MA, Ch'en IL, Stockmann C, Katayama CD, Hedrick SM. 2010. Foxo transcription factors control regulatory T cell development and function. *Immunity* **33**(6): 890-904.

Khattri R, Cox T, Yasayko SA, Ramsdell F. 2003a. An essential role for Scurfin in CD4+CD25+ T regulatory cells. *Nat Immunol* **4**(4): 337-342.

-. 2003b. An essential role for Scurfin in CD4+CD25+ T regulatory cells. *Nat Immunol* **4**(4): 337-342.

Kim HP, Leonard WJ. 2007. CREB/ATF-dependent T cell receptor-induced FoxP3 gene expression: a role for DNA methylation. *J Exp Med* **204**(7): 1543-1551.

Kim JK, Samaranayake M, Pradhan S. 2009. Epigenetic mechanisms in mammals. *Cell Mol Life Sci* **66**(4): 596-612.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**(3): R25.

Li Q, Verma IM. 2002. NF-kappaB regulation in the immune system. *Nat Rev Immunol* **2**(10): 725-734.

Lin W, Haribhai D, Relland LM, Truong N, Carlson MR, Williams CB, Chatila TA. 2007. Regulatory T cell development in the absence of functional Foxp3. *Nat Immunol* **8**(4): 359-368.

Musri MM, Parrizas M. 2012. Epigenetic regulation of adipogenesis. *Curr Opin Clin Nutr Metab Care*.

Ohkura N, Hamaguchi M, Morikawa H, Sugimura K, Tanaka A, Ito Y, Osaki M, Tanaka Y, Yamashita R, Nakano N et al. 2012. T cell receptor stimulation-induced epigenetic changes and foxp3 expression are independent and complementary events required for treg cell development. *Immunity* **37**(5): 785-799.

Ono M, Yaguchi H, Ohkura N, Kitabayashi I, Nagamura Y, Nomura T, Miyachi Y, Tsukada T, Sakaguchi S. 2007. Foxp3 controls regulatory T-cell function by interacting with AML1/Runx1. *Nature* **446**(7136): 685-689.

Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, Lai F, Zytnicki M, Notredame C, Huang Q et al. 2010. Long noncoding RNAs with enhancer-like function in human cells. *Cell* **143**(1): 46-58.

Ouyang W, Beckett O, Ma Q, Paik JH, DePinho RA, Li MO. 2010. Foxo proteins cooperatively control the differentiation of Foxp3+ regulatory T cells. *Nat Immunol* **11**(7): 618-627.

Pan F, Yu H, Dang EV, Barbi J, Pan X, Grosso JF, Jinasena D, Sharma SM, McCadden EM, Getnet D et al. 2009. Eos mediates Foxp3-dependent gene silencing in CD4+ regulatory T cells. *Science* **325**(5944): 1142-1146.

Pavesi G, Pesole G. 2006. Using Weeder for the discovery of conserved transcription factor binding sites. *Current protocols in bioinformatics / editoral board, Andreas D Baxevanis [et al]* **Chapter 2**: Unit 2 11.

Polansky JK, Schreiber L, Thelemann C, Ludwig L, Kruger M, Baumgrass R, Cording S, Floess S, Hamann A, Huehn J. 2010. Methylation matters: binding of Ets-1 to the demethylated Foxp3 gene contributes to the stabilization of Foxp3 expression in regulatory T cells. *J Mol Med*

(Berl) **88**(10): 1029-1040.

Prokhortchouk E, Defossez PA. 2008. The cell biology of DNA methylation in mammals. *Biochim Biophys Acta* **1783**(11): 2167-2173.

Rudra D, deRoos P, Chaudhry A, Niec RE, Arvey A, Samstein RM, Leslie C, Shaffer SA, Goodlett DR, Rudensky AY. 2012. Transcription factor Foxp3 and its protein partners form a complex regulatory network. *Nat Immunol* **13**(10): 1010-1019.

Sakaguchi S, Yamaguchi T, Nomura T, Ono M. 2008. Regulatory T cells and immune tolerance. *Cell* **133**(5): 775-787.

Samstein RM, Arvey A, Josefowicz SZ, Peng X, Reynolds A, Sandstrom R, Neph S, Sabo P, Kim JM, Liao W et al. 2012. Foxp3 exploits a pre-existent enhancer landscape for regulatory T cell lineage specification. *Cell* **151**(1): 153-166.

Schmidl C, Klug M, Boeld TJ, Andreesen R, Hoffmann P, Edinger M, Rehli M. 2009. Lineage-specific DNA methylation in T cells correlates with histone methylation and enhancer activity. *Genome Res* **19**(7): 1165-1174.

Sugimoto N, Oida T, Hirota K, Nakamura K, Nomura T, Uchiyama T, Sakaguchi S. 2006. Foxp3-dependent and -independent molecules specific for CD25+CD4+ natural regulatory T cells revealed by DNA microarray analysis. *International immunology* **18**(8): 1197-1209.

Takahashi T, Kuniyasu Y, Toda M, Sakaguchi N, Itoh M, Iwata M, Shimizu J, Sakaguchi S. 1998. Immunologic self-tolerance maintained by CD25+CD4+ naturally anergic and suppressive T cells: induction of autoimmune disease by breaking their anergic/suppressive state. *Int Immunol* **10**(12): 1969-1980.

Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B et al. 2012. The accessible chromatin landscape of the human genome. *Nature* **489**(7414): 75-82.

Waterland RA. 2006. Epigenetic mechanisms and gastrointestinal development. *J Pediatr* **149**(5 Suppl): S137-142.

Weber M, Schubeler D. 2007. Genomic patterns of DNA methylation: targets

and function of an epigenetic mark. *Curr Opin Cell Biol* **19**(3): 273-280.

Wei G, Wei L, Zhu J, Zang C, Hu-Li J, Yao Z, Cui K, Kanno Y, Roh TY, Watford WT et al. 2009. Global mapping of H3K4me3 and H3K27me3 reveals specificity and plasticity in lineage fate determination of differentiating CD4+ T cells. *Immunity* **30**(1): 155-167.

Willingham AT, Orth AP, Batalov S, Peters EC, Wen BG, Aza-Blanc P, Hogenesch JB, Schultz PG. 2005. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* **309**(5740): 1570-1573.

Zheng Y, Josefowicz S, Chaudhry A, Peng XP, Forbush K, Rudensky AY. 2010. Role of conserved non-coding DNA elements in the Foxp3 gene in regulatory T-cell fate. *Nature* **463**(7282): 808-812.

Zheng Y, Josefowicz SZ, Kas A, Chu TT, Gavin MA, Rudensky AY. 2007. Genome-wide analysis of Foxp3 target genes in developing and mature regulatory T cells. *Nature* **445**(7130): 936-940.