



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Cell type--specific DNA methylation at intragenic CpG islands in the immune system

Citation for published version:

Deaton, AM, Webb, S, Kerr, ARW, Illingworth, RS, Guy, J, Andrews, R & Bird, A 2011, 'Cell type--specific DNA methylation at intragenic CpG islands in the immune system' *Genome Research*, vol. 21, no. 7, pp. 1074-1086. DOI: 10.1101/gr.118703.110

Digital Object Identifier (DOI):

[10.1101/gr.118703.110](https://doi.org/10.1101/gr.118703.110)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Genome Research

Publisher Rights Statement:

Freely available online through the Genome Research Open Access option.

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Cell type-specific DNA methylation at intragenic CpG islands in the immune system

Aimée M. Deaton,¹ Shaun Webb,¹ Alastair R.W. Kerr,¹ Robert S. Illingworth,¹ Jacky Guy,¹ Robert Andrews,² and Adrian Bird^{1,3}

¹Wellcome Trust Centre for Cell Biology, University of Edinburgh, Edinburgh EH9 3JR, United Kingdom; ²Wellcome Trust Sanger Centre, Hinxton, Cambridge CB10 1SA, United Kingdom

Human and mouse genomes contain a similar number of CpG islands (CGIs), which are discrete CpG-rich DNA sequences associated with transcription start sites. In both species, ~50% of all CGIs are remote from annotated promoters but, nevertheless, often have promoter-like features. To determine the role of CGI methylation in cell differentiation, we analyzed DNA methylation at a comprehensive CGI set in cells of the mouse hematopoietic lineage. Using a method that potentially detects ~33% of genomic CpGs in the methylated state, we found that large differences in gene expression were accompanied by surprisingly few DNA methylation changes. There were, however, many DNA methylation differences between hematopoietic cells and a distantly related tissue, brain. Altered DNA methylation in the immune system occurred predominantly at CGIs within gene bodies, which have the properties of cell type-restricted promoters, but infrequently at annotated gene promoters or CGI flanking sequences (CGI “shores”). Unexpectedly, elevated intragenic CGI methylation correlated with silencing of the associated gene. Differentially methylated intragenic CGIs tended to lack H3K4me3 and associate with a transcriptionally repressive environment regardless of methylation state. Our results indicate that DNA methylation changes play a relatively minor role in the late stages of differentiation and suggest that intragenic CGIs represent regulatory sites of differential gene expression during the early stages of lineage specification.

[Supplemental material is available for this article. The sequencing and gene expression data from this study have been submitted to the NCBI Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo>) under accession nos. GSE25688 and GSE25578, respectively.]

The dinucleotide sequence CpG is the predominant site of DNA methylation in the vertebrate genome, but not all CpGs are methylated. Two fractions can be discerned based on CpG density and methylation status: The bulk of the genome is CpG-deficient and predominantly methylated (~80%), while discrete regions called CpG islands (CGIs) are on average tenfold more CpG-rich, usually unmethylated, and associated with the histone modification H3K4me3 (Guenther et al. 2007; Mikkelsen et al. 2007; Illingworth and Bird 2009; Thomson et al. 2010). The majority of gene promoters (~60%) are included within CGIs. CGI methylation is invariably associated with promoter silencing, although silenced CGI promoters often remain in a nonmethylated state (Weber et al. 2007). For example, the alpha globin CGI is unmethylated even in nonerythroid tissues (Bird et al. 1987). DNA methylation-associated gene silencing is well documented during X chromosome inactivation, imprinting, and cancer (Edwards and Ferguson-Smith 2007; Jones and Baylin 2007; Payer and Lee 2008), but recent genome-wide studies have described many additional instances in normal somatic cells. A number of these described the acquisition of CGI methylation in somatic cell lineages, compared to the germline, where CGIs are almost invariably hypomethylated (Schilling and Rehli 2007; Weber et al. 2007; Illingworth et al. 2008). CGI methylation has also been

analyzed during differentiation of embryonic stem cells (Mohn et al. 2008). Most DNA methylation studies have focused on CGIs occurring at annotated gene promoters (Weber et al. 2007; Meissner et al. 2008; Mohn et al. 2008), but it has become apparent that CGIs remote from annotated transcription start sites (TSSs), located either between genes or within the body of a transcription unit, exhibit a high degree of tissue-specific methylation (Illingworth et al. 2008; Rauch et al. 2009; Maunakea et al. 2010). These so-called “orphan CGIs” account for about half of all CGIs in human and mouse genomes (Illingworth et al. 2010). Despite the absence of annotated promoters within orphan CGIs, many are marked by H3K4me3 and RNA polymerase II (RNAPII) and give rise to detectable transcripts (Illingworth et al. 2010; Maunakea et al. 2010). Therefore, CpG methylation at these sites may be involved in silencing novel uncharacterized promoters. These findings raise questions about the functional significance of orphan CGI promoters and the dynamics of their methylation during development and differentiation.

To address these issues, we have studied cells of the immune system, which are derived from a common progenitor, the hematopoietic stem cell (Fig. 1A), and therefore offer a convenient system in which to investigate the role of DNA methylation in differentiation. Pure primary cells from the immune lineage can be isolated using fluorescence activated cell sorting (FACS) or magnetic bead purification, avoiding the need to analyze DNA methylation patterns in mixed lineage tissues or in cultured cell lines where DNA methylation is known to be abnormal (Antequera et al. 1990; Jones et al. 1990; Smiraglia et al. 2001; Meissner et al. 2008). Changes in DNA methylation at specific genes have already been

³Corresponding author.
E-mail a.bird@ed.ac.uk.

Article published online before print. Article, supplemental material, and publication date are at <http://www.genome.org/cgi/doi/10.1101/gr.118703.110>. Freely available online through the *Genome Research* Open Access option.

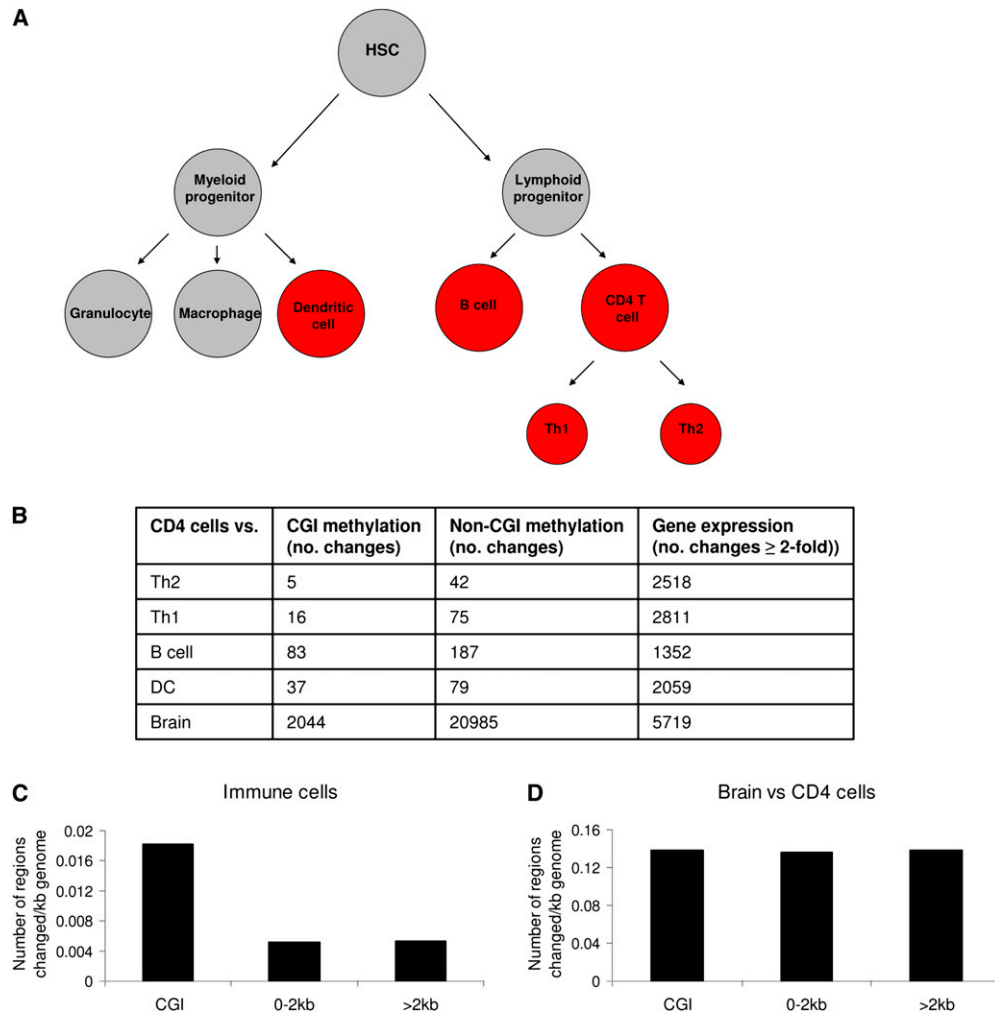


Figure 1. Cell type-specific methylation in the hematopoietic lineage detected by MAP-seq preferentially occurs at CGIs. (A) The immune cell lineage with the cell types investigated shown in red. HSC: hematopoietic stem cell. (B) The number of CGI methylation, non-CGI-associated methylation, and gene expression changes observed when CD4 T-cells were compared to Th2, Th1, B cells, dendritic cells (DC), and brain. (C, D) The location of methylation differences detected by MAP-seq. The fraction of the genome that can be interrogated by MAP-seq was categorized as overlapping a CGI (CGI), not overlapping but within 2 kb of a CGI (0–2 kb), or >2 kb from a CGI (>2 kb). The location of DNA methylation changes was then determined and expressed as the number of methylation changes occurring per kilobase (kb) of genome in each category. (C) Location of all DNA methylation changes occurring between different immune cells. (D) Location of DNA methylation changes occurring between brain and CD4 cells.

documented in this system, most notably during T helper cell differentiation. Upon infection, CD4⁺ T helper cells are activated and then specialize further in response to different kinds of infection (Reiner 2007; Wan 2010). Th1 cells coordinate the immune response to intracellular bacteria and viruses, while Th2 cells respond to infection by extracellular parasites. The key cytokine gene loci, *Ifng* in Th1 cells and *Il4* in Th2 cells, undergo changes in DNA methylation during T helper cell differentiation (Lee et al. 2002; Santangelo et al. 2002; Winders et al. 2004; Schoenborn et al. 2007), a process that can be recapitulated in vitro.

In this study, we have coupled the biochemical isolation of methylated CpG-rich DNA with high-throughput sequencing methods (MAP-seq) (Illingworth et al. 2010) to allow comprehensive, unbiased profiling of DNA methylation in immune cells. We identified cell type-specific CGI methylation occurring during T helper cell differentiation and compared methylation patterns between T helper cell subtypes and more distantly re-

lated B cells and dendritic cells (Fig. 1A). The number of changes in gene expression between these cell types far exceeded the number of DNA methylation changes. Nevertheless, alterations in DNA methylation were enriched at genes with immune system function. CGIs showed more dynamic methylation patterns than elsewhere in the genome and, intriguingly, most CGI methylation changes occurred at intragenic orphan CGIs. There was a general inverse correlation between increased DNA methylation and gene expression, yet differentially methylated intragenic CGIs often lacked the active histone modification H3K4me3, regardless of DNA methylation state. This suggests that although these intragenic CGIs are likely to be sites of transcriptional initiation in other lineages, they do not appear to act as promoters in immune cells. Our findings suggest that CGI methylation plays a limited role in the terminal differentiation of immune cells. However, the predominance of cell type-specific methylation at intragenic CGIs raises the possibility that they

represent major regulatory sites of differential gene expression during lineage specification.

Results

Methylation changes between immune cells are small in number and show an association with developmental relatedness

We used MBD affinity purification coupled to high-throughput sequencing to assess DNA methylation in B cells, dendritic cells, and CD4 T-cells isolated from mouse spleen, as well as in vitro differentiated Th1 and Th2 cells (Fig. 1A). MAP-seq interrogates sites with a CpG density of 3.3 CpGs per 100 bp and above (Supplemental Fig. S1) and therefore potentially detects 7.4 million CpG sites in the genome. This includes all CGIs and many relatively CpG-rich regions. MAP-seq gave ~17-fold coverage over these regions. Comparison of DNA methylation patterns between different immune cell types revealed surprisingly few differences both at CGIs and non-CGI regions (Fig. 1B; Supplemental Table S2). In general, the number of DNA methylation differences mirrored developmental distance, with just 16 CGIs showing differential methylation between CD4 T-cells and Th1 cells, compared to 83 CGIs showing differential methylation between CD4 T-cells and more distantly related B-cells. Nine CGIs showing differential methylation were confirmed by bisulfite genomic sequencing, and all were consistent with the MAP-seq results (Supplemental Fig. S2; Supplemental Table S3). However, bisulfite analysis of methylation on individual DNA strands did not support the occurrence of allele-specific DNA methylation changes (Supplemental Fig. S2). We then compared our immune cell methylation data to a more distantly related tissue. For this, we used a previously determined MAP-seq profile for mouse brain (Skene et al. 2010). The number of DNA methylation differences between brain and CD4 cells was over 2000 at CGIs, about 100-fold higher than between CD4 cells and Th1 cells, and ~20,000 at non-CGI-associated regions (Fig. 1B). Interestingly, brain showed a general loss of CGI methylation compared to immune cells (data not shown). Our findings suggest that the vast majority of DNA methylation changes arise early in lineage commitment, while only a few are acquired during the later stages of terminal differentiation. A similar conclusion has been drawn by analysis of promoter methylation during the differentiation of embryonic stem cells to neurons (Mohn et al. 2008).

We then compared our DNA methylation data with gene expression profiles in the equivalent cell populations using Illumina BeadChip expression arrays. Gene expression differences of twofold or more between immune cell types were numerous and, in contrast to DNA methylation changes, did not correlate closely with developmental relatedness. For example, more genes were differentially expressed between closely related CD4 cells and Th1 cells (2811) than between more diverged CD4 and dendritic cells (2059; χ^2 -test, $P < 2.2 \times 10^{-16}$) (Fig. 1A,B). Gene expression, therefore, reflects functional adaptation of immune cells, rather than their developmental relationship. Comparing CD4 cells with brain revealed that the number of DNA methylation changes (>22,000) is significantly higher than the number of gene expression differences (5719). This is in contrast to comparisons between different immune cell types, where gene expression differences vastly outnumber DNA methylation changes. These findings suggest that differential methylation, at least at the 7.4 million sites we are able to detect using MAP-seq, may not play

a major role in regulation of gene expression during terminal differentiation.

Cell type-specific DNA methylation in the immune system detected by MAP-seq occurs preferentially at CGIs

We analyzed where most of the cell type-specific methylation in the immune system occurs and found that proportionally more changes take place at CGIs than elsewhere in the genome. This is in contrast to recent reports suggesting that tissue-specific methylation preferentially occurs at the edges of CGIs (Irizarry et al. 2009; Ji et al. 2010). We categorized the fraction of the genome that can be interrogated by MAP-seq as overlapping a CGI (CGI), not overlapping but within 2 kb of a CGI (0–2kb), or >2 kb from a CGI (>2 kb) and observed that the flanks of CGIs showed no preference for cell type-specific methylation compared to the rest of the genome at the 7.4 million CpG sites examined in our study. This was the case when methylation differences between cells of the immune system were examined (Fig. 1C) and also when CD4 T-cells were compared to a distantly related tissue, brain (Fig. 1D). We used a comprehensive, biologically defined mouse CGI set (Illingworth et al. 2010) to examine the location of methylation changes between cell types. Using other bioinformatically defined CGI sets for analysis did not result in enrichment for DNA methylation changes at CGI shores (Supplemental Fig. S3). In addition, we specifically examined the average number of MAP-seq hits in the flanks of CGIs differentially methylated in dendritic cells compared to Th1 cells and brain compared to CD4 cells. The greatest differences in read number between the different cell types were seen over the body of the CGI and not in the 2-kb flanking regions (Supplemental Fig. S4).

CGI methylation differences preferentially occur within genes that have immune system function

We asked whether the relatively small number of DNA methylation changes at CGIs are preferentially associated with genes involved in immune cell function. Analysis of gene ontology terms associated with genes showing cell type-specific CGI methylation revealed an enrichment for genes involved in immunity and defense, cytokine/chemokine-mediated immunity, and developmental processes (see Methods for gene ontology analysis) (Supplemental Table S4). These ontology terms were the only ones over-represented. In contrast, genes showing differential CGI methylation between distantly related brain and CD4 cells were involved in a broader range of biological processes, only some of which were associated with neuronal function (see Supplemental Table S4).

Genes that show differential methylation in immune system cells were also examined individually with respect to function, and many were found to have documented roles in the immune system (Table 1). Significantly, the only methylation difference between Th1 and Th2 cells occur in the *Gata3* gene (Fig. 3F,G, see below), which encodes the transcription factor determining Th2 cell fate (Nawijn et al. 2001). The only gene-associated CGIs to show a DNA methylation difference between CD4 cells and both Th1 and Th2 cells are located in *Dtx1*, a gene which suppresses T-cell activation and is down-regulated when T-cells are activated (Liu and Lai 2005), and *Kcnn4* (Fig. 4A, see below), which codes for a calcium-activated potassium channel involved in T-cell activation (Begenisich et al. 2004). We conclude that, despite their low number, DNA methylation changes colocalize with genes that play functionally important roles in that cell type.

Table 1. Cell type-specific CGI methylation affects genes with immune system function

Gene	Cell types showing a change in CGI methylation (hypomethylated cell type in brackets)	Function	References
<i>Gata3</i>	Th1 (vs. Th2)	Th2-specific transcription factor	(Nawijn et al. 2001)
<i>Dtx1</i>	Th1, Th2 (vs. CD4/DC)	Suppression of T-cell activation	(Liu and Lai 2005)
<i>Kcnn4</i>	CD4, DC, B cell (vs. Th1/Th2)	Potassium channel involved in T-cell activation	(Begenisich et al. 2004)
<i>Fgr</i>	CD4 (vs. DC)	Regulation of myeloid cell chemokine signaling	(Zhang et al. 2005)
<i>Sema4a</i>	CD4, Th1, Th2 (vs. DC)	Co-stimulation, T-cell differentiation	(Kumanogoh et al. 2005)
<i>Bcl11b</i>	Th1 (vs. DC)	Thymocyte development	(Wakabayashi et al. 2003)
<i>Runx3</i>	B cell (vs. CD4)	Transcription factor, promotes Th1 cell fate	(Djuretic et al. 2007)
<i>Klf2</i>	DC (vs. Th1)	Transcription factor regulating IL-2 production early in T-cell activation	(Wu and Lingrel 2005)
<i>Il10ra</i>	CD4 (vs. B cell)	Interleukin-10 receptor, regulation of immune response.	(Pils et al. 2010)
<i>Tnfrsf25</i>	DC (vs. Th1/Th2)	Pro-inflammatory gene	(Mookherjee et al. 2006)
<i>Cxcr5</i>	CD4 (vs. B cell)	Chemokine receptor	(Hardtke et al. 2005; Junt et al. 2005)
<i>Notch 2</i>	Th1, Th2, CD4 (vs. DC)	Signaling during immune cell differentiation	(Cheng et al. 2003; Fiorini et al. 2009)
<i>Il16</i>	Th1 (vs. DC)	T-cell chemoattractant, regulation of T-cell growth.	(Cruikshank and Little 2008)

A selection of the immune system genes that display differential CGI methylation is shown along with the relevant references, although this list is not exhaustive. For each example, the cell type(s) showing a decrease in methylation (hypomethylation) is indicated in parentheses.

Cell type-specific methylation preferentially occurs at intragenic CGIs and this negatively correlates with gene expression

CGIs are frequently associated with gene promoters, but we recently identified large numbers of “orphan” CGIs that are not associated with annotated TSSs (Illingworth et al. 2010). Orphan CGIs account for approximately half of CGIs in both human and mouse and consist of two subsets: intragenic and intergenic CGIs (Fig. 2A,B). Strikingly, differential CGI methylation between all immune cell types predominantly occurred at non-TSS-associated CGIs, specifically those located within transcription units (Fig. 2C,D). Altogether, 2% of intragenic CGIs were involved in DNA methylation transitions, compared to 0.28% of annotated promoter CGIs (Fig. 2D). Thus, intragenic CGIs are distinct from TSS-associated and intergenic CGIs with regard to cell type-specific methylation. When brain and CD4 cells are compared, the majority of tissue-specific methylation also occurs at intragenic CGIs (Supplemental Fig. S5).

Given the increased frequency of differential methylation detected at intragenic CGIs, we examined the relationship between gene expression and cell type-specific methylation for gene-associated intragenic CGIs in immune cells. In most cases (74%), alterations in CGI methylation were not associated with discernable differences in expression of the associated gene. Focusing on the 26% of cases where DNA methylation and gene expression both changed, we observed that increased intragenic CGI methylation was most often linked to decreased expression of the gene in which the CGI was located. Similarly, decreased methylation associated with increased gene expression (70% of cases) (Fig. 3A,B). To investigate whether this negative correlation between cell type-specific intragenic CGI methylation and gene expression was a general

phenomenon and not an artefact of the small number of methylation differences detected between different immune system cells, we compared CD4 cells to distantly related brain, which shows hypomethylation of CGIs compared to immune cells. In brain, decreased DNA methylation at intragenic CGIs tended to associate with increased gene expression (82% of cases) (Fig. 3C,D). This supports the negative correlation between cell type-specific intragenic CGI methylation and expression of the associated gene that was observed between immune cell types.

A key immune cell regulator showing a negative correlation between intragenic CGI methylation and expression is the *Gata3* gene, which is expressed specifically in Th2-cells (Fig. 3E) and determines Th2 cell identity (Nawijn et al. 2001). The only methylation difference between Th1 and Th2-cells occurred at an intragenic CGI in the *Gata3* gene (Fig. 3F) and was confirmed by bisulfite sequencing and found to be relatively subtle (54% in Th1 compared to 72% methylation in Th2-cells) when in vitro differentiated Th1 and Th2-cells were compared (Fig. 3G). However, IL4-producing Th2 cells isolated from a Th2-inducing infection with *Schistosoma mansoni* showed complete demethylation of this CGI after an 8-wk infection (8% methylation) (Fig. 3G). This confirmed that DNA methylation changes identified at intragenic CGIs using in vitro T-cell differentiation are relevant in vivo. Complete demethylation of the *Gata3* CGI, as seen in vivo, may require a longer time period than is possible in vitro.

Intragenic CGI methylation at *Kcnn4* and *Gata3* shows an inverse relationship with H3K4me3

Unmethylated CGIs are frequently associated with the transcriptionally permissive histone modification H3K4me3, regardless of gene activity (Guenther et al. 2007; Mikkelsen et al. 2007; Illingworth

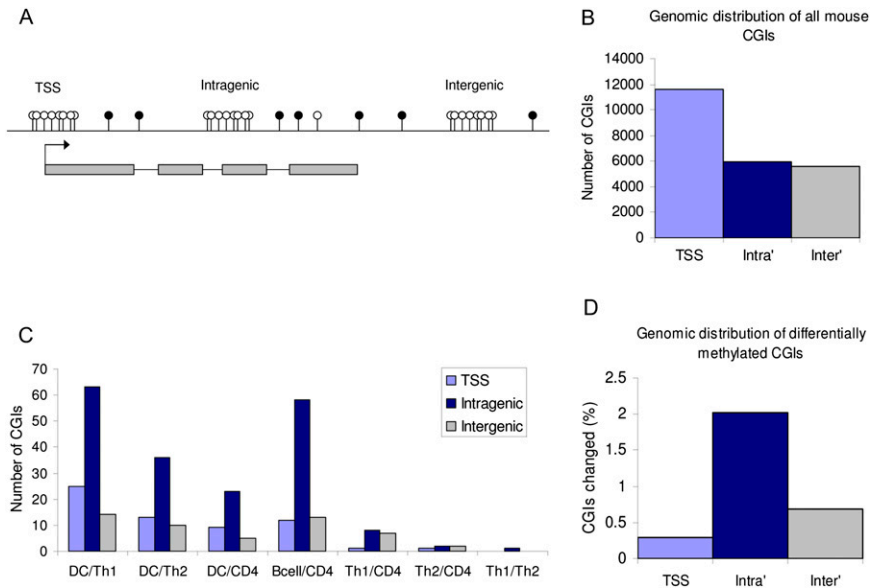


Figure 2. Cell type-specific methylation occurs at intragenic CGIs. (A) CGIs can be transcription start site (TSS)-associated, intragenic, or intergenic. (B) The distribution of CGIs in the mouse genome (Illingworth et al. 2010). (C) Pairwise comparisons of immune cell types [dendritic cells (DC), Th1 cells, Th2 cells, B cells, and CD4 T-cells] showing the location of methylation changes in each comparison. (D) The percentage of CGIs in each class showing a DNA methylation change in any of the immune cell comparisons. Each CGI was only counted once even if it changed in multiple comparisons.

et al. 2010; Thomson et al. 2010). This is true for both TSS and intragenic CGIs, although the proportion of intragenic CGIs marked by H3K4me3 is smaller, with 41% of intragenic CGIs positive for H3K4me3 in embryonic stem (ES) cells compared to 96% of TSS CGIs (Illingworth et al. 2010). Methylated CGIs, on the other hand, do not associate with H3K4me3, reflecting transcriptional silencing at these sites (Meissner et al. 2008; Illingworth et al. 2010; Maunakea et al. 2010). Recent studies have suggested a mechanistic explanation for this phenomenon whereby the protein CFP1 binds to unmethylated CGIs and facilitates recruitment of the SET1 H3K4 methyltransferase complex (Thomson et al. 2010). We reasoned that cell type-specific DNA methylation at intragenic CGIs should lead to a decrease in H3K4me3 at these sites. This was, indeed, the case at a CGI within the body of the *Kcm4* gene, which is marked by H3K4me3 in Th1 cells, where it is hypomethylated, but not in B cells or dendritic cells, where it is heavily methylated (Fig. 4A,B).

Because of its pivotal role in Th2 cell differentiation (Nawijn et al. 2001), the *Gata3* gene is of particular interest and, furthermore, shows the only methylation difference we observed between Th1 and Th2 cells. At the *Gata3* intragenic CGI (Fig. 4C), a small decrease in DNA methylation in Th2 cells (Fig. 3G) was accompanied by a large increase in H3K4me3 (Fig. 4D). On the other hand, H3K4me3 levels at the unmethylated TSS CGI did not vary between Th1 and Th2 cells (Fig. 4C,D) raising the possibility that the intragenic CGI has a regulatory role at this locus.

Silent chromatin configurations at differentially methylated intragenic CGIs

We next asked whether an association between decreased DNA methylation and increased H3K4me3 at intragenic CGIs, such as that observed at *Kcm4* and *Gata3*, was a common feature of cell type-specific methylation. We examined H3K4me3 distribution

genome-wide in Th1 and dendritic cells, as these cell types showed the greatest number of differences in CGI methylation. Of 63 intragenic CGIs showing differential methylation between Th1 and dendritic cells, 55% lacked H3K4me3 in both cell types (Fig. 5A). Of the remaining 45%, almost half showed differential H3K4me3 similar to that detected at *Kcm4* and *Gata3*. In the remainder of cases, changes in DNA methylation occurred at CGIs positive for H3K4me3 in both immune cell types, or differential H3K4me3 was undetectable using our parameters (Fig. 5A).

The absence of the H3K4me3 mark at most differentially methylated intragenic CGIs suggests that they are associated with a transcriptionally inert chromatin environment, regardless of the DNA methylation state. It follows that a DNA methylation-independent mechanism may explain the absence of this "active" chromatin modification. Polycomb-mediated silencing, marked by the presence of H3K27me3, is a potential alternative as 58% of CGIs lacking H3K4me3 in mouse brain have this chromatin mark (Thomson et al. 2010). These histone modifications can also occur together in the context of bivalent chromatin (Bernstein et al. 2006; Mikkelsen et al. 2007). H3K27me3 enrichment at differentially methylated CGIs located in the gene bodies of *Tead3* and *Kctd17* was tested using ChIP-qPCR in Th1 and dendritic cells. Levels were compared to those at the *Sox2* promoter, which has been reported to be silenced by H3K27me3 (Mikkelsen et al. 2007), and the CGI promoter of the housekeeping gene *Actb* (Fig. 5B). We found that the *Tead3* and *Kctd17* intragenic CGIs were marked by H3K27me3 in both Th1 and dendritic cells, although levels varied between the two genes (Fig. 5B). This suggests that silencing by polycomb, as well as, or as an alternative to, DNA methylation may explain the lack of H3K4me3 and hence the transcriptionally inert state at many cell type-specifically methylated intragenic CGIs.

To understand the relationship between H3K4me3 and H3K27me3 at differentially methylated intragenic CGIs, we performed H3K27me3 ChIP-seq in Th1 and dendritic cells. We found that half of the differentially methylated intragenic CGIs negative for H3K4me3 were marked by H3K27me3, indicating that polycomb silencing contributes to a transcriptionally inert state at many of these CGIs but does not account for all sites devoid of this active chromatin modification. H3K27me3 levels at *Tead3*, *Kctd17* (Fig. 5C,D), *Sox2*, and *Actb* (Supplemental Fig. S6A,B) were consistent with ChIP-qPCR results. Examples of other differentially methylated CGIs which lack H3K4me3 but are marked by H3K27me3 (located at the *Tnfrif2* and *2810459M11Rik* loci) are shown in Supplemental Figure S6C,D. However, H3K27me3 itself is not refractory to H3K4me3, as evidenced by the existence of a small number of differentially methylated bivalent intragenic CGIs, marked by H3K4me3 and H3K27me3 in both Th1 and dendritic cells (Supplemental Fig. S7). Altogether, our findings highlight that CGIs showing cell type-specific differences in DNA

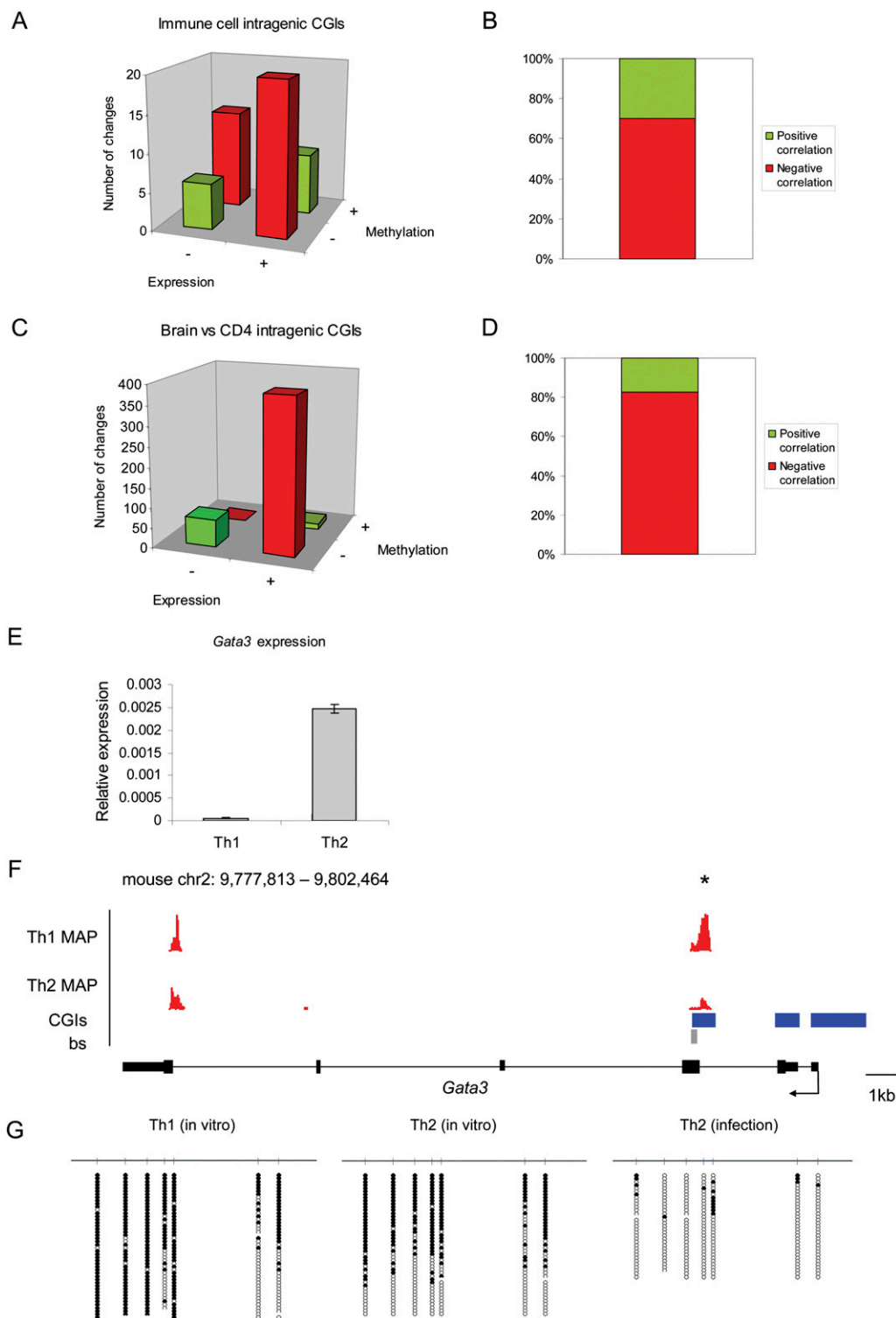


Figure 3. Differential intragenic CGI methylation negatively correlates with gene expression. Genes showing both a change in DNA methylation at an intragenic CGI and a change in gene expression were plotted. (A) All immune cell comparisons where (+) denotes an increase and (-) denotes a decrease in expression or methylation. (B) The percentage of methylation changes showing a positive or negative correlation with gene expression in immune cells. (C) Brain compared with CD4 cells. (+) denotes an increase and (-) a decrease in expression or methylation. (D) The percentage of methylation changes between brain and CD4 cells showing a positive or negative correlation with gene expression. (E) Expression of *Gata3* (relative to *Eef1A1*) in Th1 and Th2 cells verified by qRT-PCR. (F) MAP-seq read density profile (red) showing differential methylation at an intragenic CGI (asterisk) in the *Gata3* gene. CGIs are shown in blue. The arrow indicates the origin and direction of transcription. (G) Confirmation of *Gata3* methylation by bisulfite sequencing of in vitro differentiated Th1 and Th2 cells as well as Th2 cells isolated from a *Schistosoma mansoni* infection. The gray bar on the MAP-seq profile (panel F) indicates the region interrogated. (Filled circles) Methylated CpGs. (Empty circles) Unmethylated CpGs.

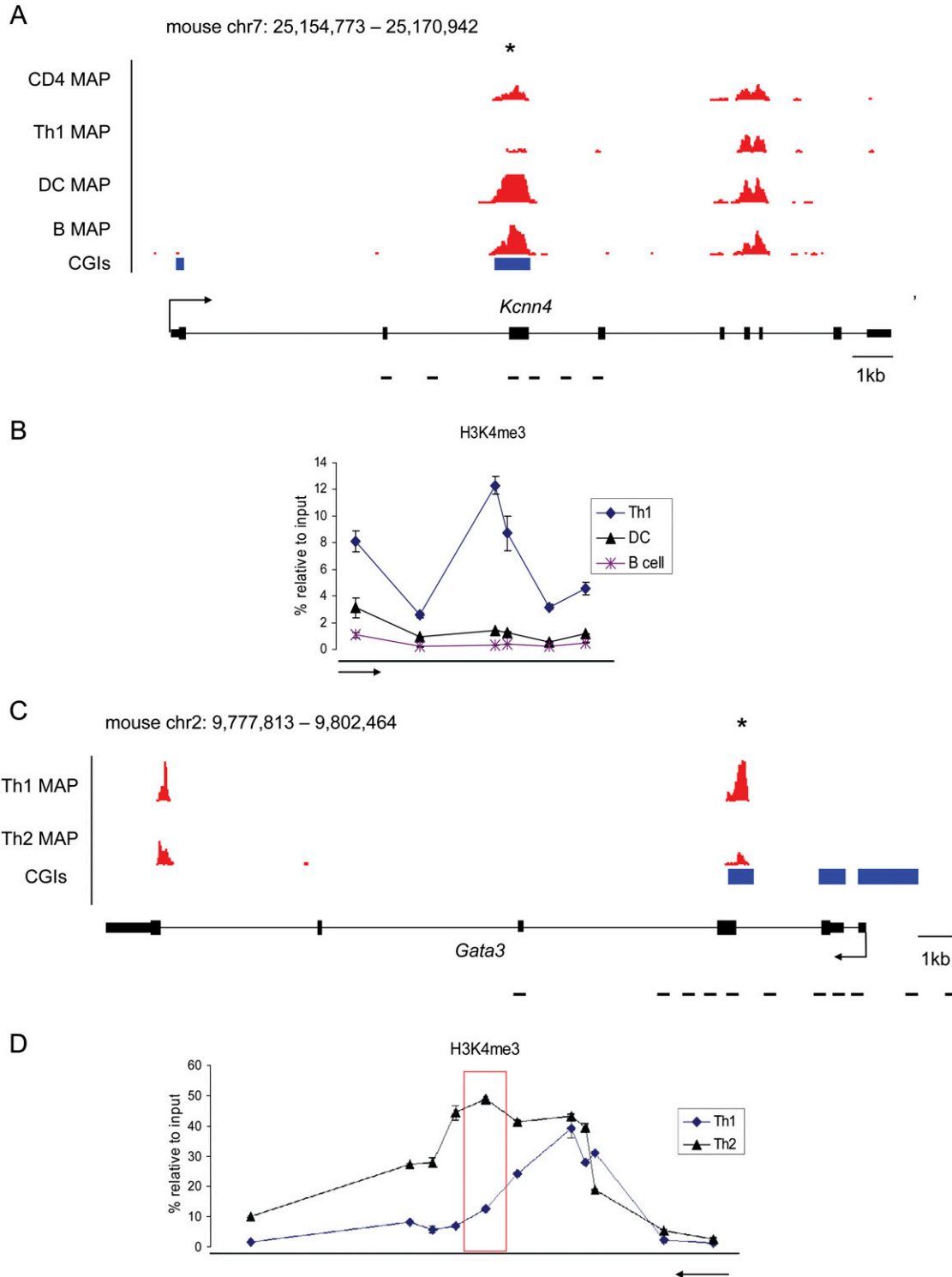


Figure 4. Decreased DNA methylation is associated with increased H3K4me3 at the *Kcnn4* and *Gata3* intragenic CGIs. (A) MAP-seq read density profile (red) showing a methylation difference between Th1 cells, dendritic cells (DC), and B cells at a CGI in *Kcnn4* gene (asterisk). (B) ChIP combined with qPCR across the *Kcnn4* locus reveals specific enrichment for H3K4me3 in Th1 cells where the intragenic CGI is hypomethylated. (C) The methylation difference between Th1 and Th2 cells at the *Gata3* locus (asterisk) is associated with a difference in H3K4me3 specifically over the intragenic CGI (red box) (D). For both genes, the arrow indicates the origin and direction of transcription, blue indicates CGIs, and black lines indicate the position of PCR primer pairs.

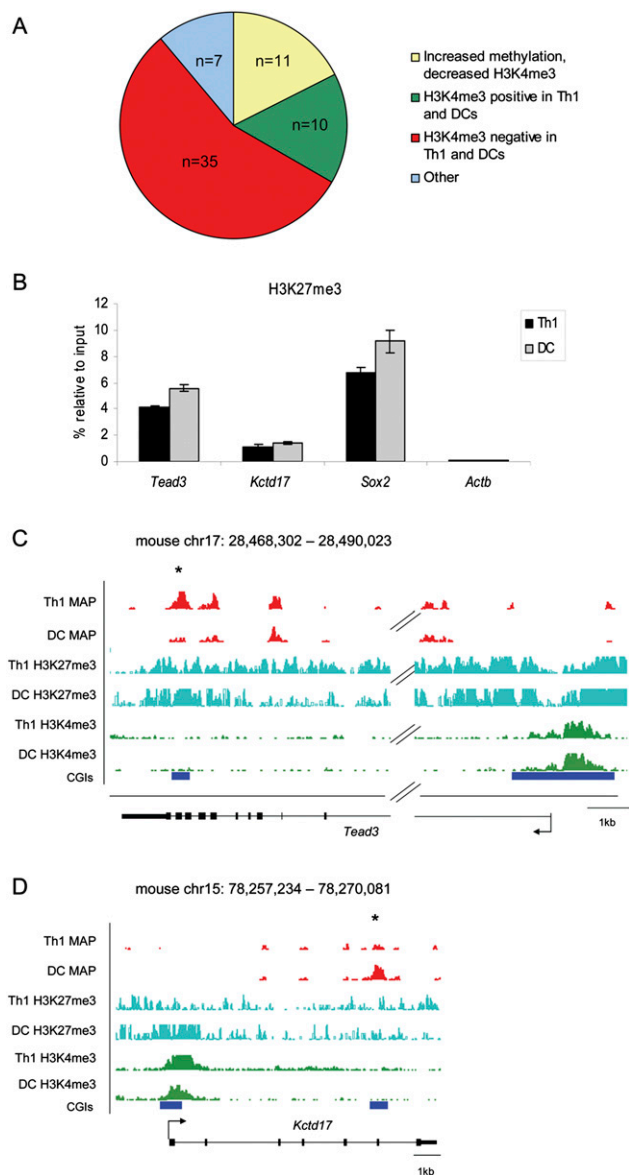


Figure 5. Most differentially methylated intragenic CGIs are depleted for H3K4me3 in immune cells. (A) Association of intragenic CGIs differentially methylated in Th1 and dendritic cells with H3K4me3 as assessed by ChIP-seq. (Yellow) CGIs where increased DNA methylation is associated with decreased H3K4me3 ($n = 11$). (Green) CGIs that are positive for H3K4me3 in both cell types despite a difference in methylation ($n = 10$). (Red) CGIs that lack H3K4me3 in both cell types ($n = 35$). (Blue) CGIs showing a nonsignificant change in H3K4me3 ($n = 5$) or where increased DNA methylation is associated with increased H3K4me3 ($n = 2$). (B) ChIP-PCR reveals enrichment for H3K27me3 at *Tead3* and *Kctd17* in Th1 and dendritic cells (DC) as well as at *Sox2* (positive control). The active *Actb* gene acted as a negative control. These results were confirmed and extended using H3K27me3 ChIP-seq. MAP-seq (red), H3K27me3 ChIP-seq (cyan), and H3K4me3 (green) read density profiles for intragenic CGIs in the (C) *Tead3* and (D) *Kctd17* genes that lack H3K4me3 in Th1 and dendritic cells, despite showing differential methylation between the two cell types (asterisked CGI).

methylation are often transcriptionally inert regardless of methylation state. They also demonstrate the existence of multiple modes of regulation for intragenic CGIs and highlight the complexity of silencing mechanisms at these regions.

Intragenic CGIs showing cell type-specific methylation in the immune system may be promoters in other tissues

Many intragenic CGIs act as promoters for alternative transcripts of unknown functional significance (Illingworth et al. 2010; Maunakea et al. 2010). More than half are associated with one or more signs of promoter activity, including H3K4me3, RNAPII, and the presence of a capped transcript (Illingworth et al. 2010; Maunakea et al. 2010). It is likely that this proportion will increase as further tissues are examined. Methylated intragenic CGIs, on the other hand, lack both H3K4me3 and RNAPII (Illingworth et al. 2010; Maunakea et al. 2010). Surprisingly, we found that the majority of intragenic CGIs showing cell type-specific methylation in the immune system lack H3K4me3 even when hypomethylated. Comparison between immune cells and brain, however, indicates that this reflects tissue-specific marking of intragenic CGIs by H3K4me3. In brain, 13% of intragenic CGIs that were negative for H3K4me3 in dendritic cells were positive for H3K4me3. Conversely, 11.5% of intragenic CGIs that lacked H3K4me3 in brain were H3K4me3-positive in dendritic cells (all intragenic CGIs were examined) (Fig. 6A). A similar trend was observed when dendritic cells were compared to ES cells (Supplemental Fig. S8). Our findings emphasize the highly variable activity of intragenic CGI promoters between tissues. This contrasts

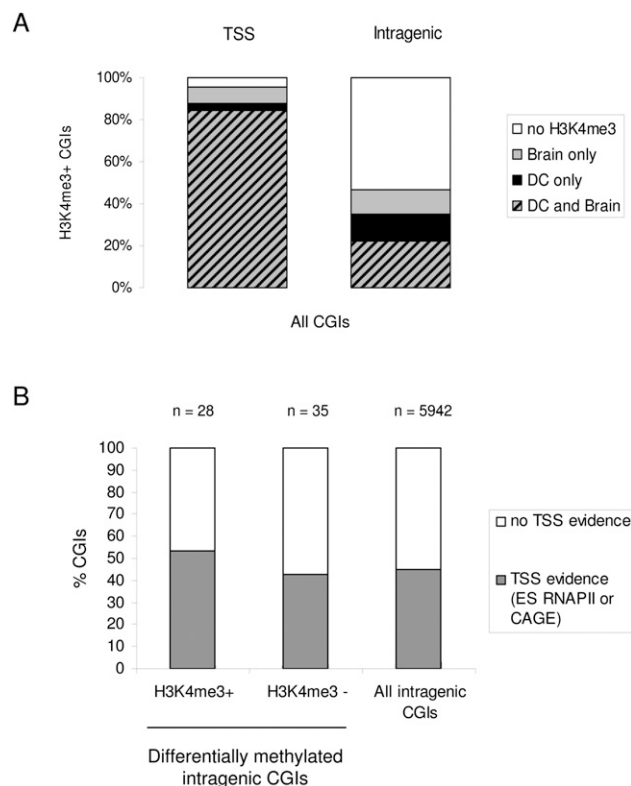


Figure 6. Intragenic CGIs frequently show tissue-specific H3K4me3, and differentially methylated intragenic CGIs have evidence for promoter function. (A) The percentage of TSS and intragenic CGIs that are positive for H3K4me3 in brain (gray), dendritic cells (black), or both (stripes). (B) RNAPII binding data for ES cells and CAGE data for mouse tissues was used to infer whether intragenic CGIs have evidence for association with a TSS (see Methods). H3K4me3 positive and H3K4me3 negative differentially methylated intragenic CGIs are shown, along with all intragenic CGIs in the mouse genome.

with the situation for annotated TSS CGIs, the vast majority of which possess H3K4me3 in both dendritic cells and brain (Fig. 6A).

To test intragenic CGIs differentially methylated in the immune system more directly for promoter activity in other cell types, we looked for direct evidence of transcriptional initiation at these sites by examining RNAPII binding in ES cells (Illingworth et al. 2010) and cap analysis of gene expression (CAGE) data from a panel of mouse tissues that did not include the immune cells investigated here (Faulkner et al. 2009). Approximately 40%–50% of these CGIs show evidence for transcriptional initiation by at least one of these criteria, and this is comparable to the proportion of all intragenic CGIs that are associated with TSS evidence (Fig. 6B). These findings suggest that, although the majority of differentially methylated intragenic CGIs are associated with a repressed chromatin state in the immune system, they may represent active promoters in other cell types. One possibility is that they function during the early stages of commitment to the hematopoietic lineage when the DNA methylation pattern is set.

Discussion

Cell type–specific methylation is rare between cells from the same lineage

We examined DNA methylation in terminally differentiated cells from the hematopoietic lineage and found only a small number of differences between cells both at CGIs and at the non-CGI-associated regions amenable to interrogation by MAP-seq. Induction of Th1 and Th2 cells from CD4 T-cells was accompanied by only a handful of detectable changes genome wide. In contrast to the relative homogeneity of DNA methylomes in this lineage, we observed very different DNA methylation patterns when immune cells were compared with brain. This confirms that major changes in DNA methylation status are detectable using this methodology. Furthermore, it suggests that large changes in DNA methylation patterns may occur early in lineage commitment, with far fewer alterations accompanying terminal differentiation. A similar conclusion was drawn for annotated promoters in an *in vitro* differentiation model of glutamatergic pyramidal neurons (Mohn et al. 2008). Dynamic alterations in DNA methylation appeared to occur prior to the multipotent progenitor state, leaving relatively few additional changes in the later stages of differentiation. Our data extends this conclusion to a major fraction of the genome in a different cell lineage.

In contrast to the relative stability of DNA methylome, gene expression programs showed large numbers of greater than two-fold changes between different cells of the hematopoietic lineage. The discrepancy between large scale transcriptional variations arising during terminal differentiation and far fewer alterations in the DNA methylome leads us to infer that DNA methylation dynamics do not play a major role in determining cell type–specific gene expression programs during late differentiation. A caveat to this conclusion is that MAP-seq does not sample the entire DNA methylome. We estimate that the conditions used for MAP-seq recover DNA with a methyl-CpG density of 3.3 CpGs per 100 bp or above, which potentially includes 6.5% of the mouse genome (excluding repeats). This means that over one-third (7.4 million) of the 21.3 million CpG sites are potentially detectable if densely methylated. The detectable fraction includes all CGIs in the genome. While it remains possible that critical DNA methylation changes in undetected regions of the genome are involved in choreographing gene expression, our data rule out the possibility that modulation of CGI methylation plays a large role in this.

CpG islands are more dynamically methylated than their flanking DNA sequences

Several studies have reported that the regions flanking CGIs (so-called “CpG island shores”) are much more likely to be differentially methylated than CGIs themselves (Doi et al. 2009; Irizarry et al. 2009). This conclusion was recently extended to the hematopoietic lineage (Ji et al. 2010). The methodology relies on cleavage with a DNA methylation-specific type I restriction enzyme, followed by hybridization to microarrays carrying relatively GC-rich regions of the genome. Using a distinct technology (MAP-seq) that also preferentially sees CpG-rich sequences, we did not detect more variable DNA methylation in CGI shores. In fact, between immune system cells, we observed fewer changes in CGI flanks than in CGIs themselves. Moreover, CGI flanks were not more variable in the cells we studied than other detectable regions in the genome at large. A resolution of this discrepancy awaits the widespread application of whole methylome sequencing approaches (Lister et al. 2009).

Intragenic CGIs most frequently show cell type–specific methylation

We recently described a comprehensive annotation of CGIs in human and mouse, identifying three classes: TSS, intragenic, and intergenic CGIs (Illingworth et al. 2010). The latter two classes were grouped as “orphan CGIs” to reflect our ignorance about their functional significance. This study represents the first analysis of cell type–specific methylation using this comprehensive CGI set. Our results show that CGIs located within gene bodies are far more likely to show cell type–specific methylation in the immune system than other CGIs. This demonstrates that intragenic CGIs are distinct from TSS and intergenic CGIs, even though other features such as DNA sequence characteristics, frequent presence of H3K4me3, and RNAPII are shared with the other classes. Our results suggest that intragenic CGIs may share a distinct biological role. Most previous studies examining tissue-specific CGI methylation have focused on annotated promoters (Weber et al. 2007; Meissner et al. 2008; Mohn et al. 2008) and agree with our findings demonstrating that TSS CGIs rarely show differential methylation. It is likely that a fuller understanding of intragenic CGIs will be important for elucidating the extent and significance of cell type–specific methylation.

Differentially methylated intragenic CGIs tend to be silent in immune cells regardless of DNA methylation state

Intragenic CGIs are frequently associated with H3K4me3 when they are unmethylated (Illingworth et al. 2010; Maunakea et al. 2010). We investigated whether cell type–specific CGI methylation was associated with a specific loss of H3K4me3 in the methylated cell type. Although this was the case for some CGIs (e.g., *Kcnn4* and *Gata3*), when all intragenic CGIs differentially methylated in Th1 and dendritic cells were examined, the majority lacked H3K4me3, even in the cell type where they were unmethylated. This suggests that these CGIs are associated with a repressive chromatin environment regardless of methylation state. Previous studies have found that CGIs negative for H3K4me3 in brain are often marked by H3K27me3 (Thomson et al. 2010), and this was the case for half of the differentially methylated intragenic CGIs lacking H3K4me3 in Th1 cells and dendritic cells.

Differentially methylated intragenic CGIs may represent novel promoters

There is evidence for transcriptional initiation at many non-TSS-associated CGIs (Illingworth et al. 2010; Maunakea et al. 2010). Accordingly it has been suggested that CGIs within gene bodies act as alternative promoters for the associated gene and that cell type-specific methylation at these sites reflects silencing of these promoters (Illingworth et al. 2010; Maunakea et al. 2010). This has been shown to be the case for intragenic CGIs in the *Shank3* and *Nfix* genes (Maunakea et al. 2010). However, we found that the majority of intragenic CGIs showing cell type-specific methylation were deficient in H3K4me3 in immune cells even when unmethylated. This implies that these CGIs are not active promoters in immune cells and conflicts with the idea that differential methylation reflects silencing of alternative promoters. Differential DNA methylation in the immune system may reflect stochastic acquisition of methylation at already silent loci or cell type-dependent differences in the use of DNA methylation as a silencing mechanism. There is, however, evidence that these differentially methylated intragenic CGIs may represent promoters active in other cell types. As intragenic CGIs are distinct from other CGIs in terms of their more frequent cell type-specific methylation, one possibility is that they act as promoters for transcripts with particularly restricted expression patterns and/or noncoding transcripts. This may explain why they have escaped annotation by conventional means.

Intragenic CGI methylation may affect gene expression by an indirect mechanism

We observed that increased intragenic CGI methylation, when associated with a gene expression change, negatively correlates with gene expression. The functional significance of this correlation is unclear, as these CGIs are often far from the gene's annotated promoter and do not appear to have novel promoter activity in immune cells. One possibility is that intragenic CGI methylation may impede gene expression by affecting transcriptional elongation. For example, transcriptional elongation through densely methylated CpGs has been reported to be less efficient than through unmethylated CpG-rich sequences (Lorincz et al. 2004). It is evident, however, that intragenic methylation is not refractory to transcription because the *Airm* CGI is present within the body of the expressed *Igf2r* gene (Sleutels et al. 2002), and gene body methylation at lower density is characteristic of actively transcribed genes (Ball et al. 2009; Deng et al. 2009; Rauch et al. 2009). Accordingly, we observed methylated intragenic CGIs in a number of actively transcribed genes in immune cells. Despite these precedents, it cannot be ruled out that dense intragenic methylation can modulate transcription in certain instances.

Another way in which intragenic CGI methylation might affect gene expression is through regulation of noncoding transcripts that affect transcription of the associated gene. Intriguingly, there is evidence that long intergenic noncoding RNAs (ncRNAs) can positively regulate the expression of nearby genes (Orom et al. 2010). It is possible that ncRNAs originating from some intragenic CGIs may also play such a role. Silencing of the regulatory ncRNA promoter by DNA methylation might explain the negative correlation between CGI methylation and gene expression observed for some differentially methylated intragenic CGIs.

A third possible function of intragenic CGI methylation is in modulating alternative splicing. It has been proposed that a closed chromatin environment over an alternatively spliced exon might

slow down the passage of RNAPII and give the splicing machinery more time to recognize that exon, thus facilitating its inclusion in the transcript (Kornblihtt 2006). There is evidence that siRNA-mediated histone modifications (Allo et al. 2009) and the SWI/SNF nucleosome remodelling complex (Batsche et al. 2006) affect splicing in such a manner. It is, therefore, conceivable that differences in intragenic CGI methylation might alter transcript processing in specific cell types. Further studies are now required to decipher how intragenic CGIs can modulate the expression of their associated genes as well as the mechanisms allowing their precise regulation during cell lineage commitment.

Methods

Immune cell isolation

Cells were isolated from the spleens of C57/BL6 mice, and biological samples consisted of pools of individual mice aged 6–12 wk (see Supplemental Table S1 for sample details). Red blood cells were lysed using red blood cell lysing buffer (Sigma) according to the manufacturer's instructions. To isolate conventional myeloid dendritic cells, plasmacytoid dendritic cells (which are lymphoid in origin) were first depleted using PDCA microbeads (Miltenyi Biotec), and then CD11c+ dendritic cells were positively selected using CD11c microbeads (Miltenyi Biotec), both used according to the manufacturer's instructions; cell purity was typically 73%–90%. CD4 T-cells were purified in a similar manner using CD4 microbeads (Miltenyi Biotec), giving purities of 87%–97%. For gene expression analysis, naive CD4 cells (CD44 low, CD62L high) were sorted using a BD FACS Aria machine with purity >90%. B cells (CD19+) were sorted in a similar manner with purity >90%.

Immune cell culture

To generate differentiated Th1 and Th2 cell populations, cells isolated from the spleen were stimulated with plate-bound α -CD3 and α -CD28 (1.6 μ g; BD Pharmingen), and recombinant murine IL-2 (20ng/mL; Peprotech) was added. For Th1 differentiation, IL-12 (5ng/mL; R&D Systems) and anti-IL4 (5 μ g/mL; BD Pharmingen) were added. For Th2 polarization, IL4 (10ng/mL; Peprotech) and anti-IL-12 (5 μ g/mL; BioLegend) were added. To assess cytokine production by intracellular staining, PMA (50ng/mL), ionomycin (500ng/mL), and Golgistop (1:2000; BD Pharmingen) were added to cultures on day 7. After 4 h, cells were harvested, permeabilized, and stained with antibodies against IFNG (Biolegend) or IL4 (BD Pharmingen), to confirm Th1 and Th2 cell identity, respectively. For DNA methylation analysis, cells were sorted based on IFNG production (Th1 cells) or IL4 production (Th2 cells).

Th2 cells were also isolated from mice infected with *Schistosoma mansoni*. At 8 wk post-infection, spleens were harvested, and CD4 cells were purified with CD4 microbeads (Miltenyi Biotec) and then stimulated with PMA (10ng/mL), ionomycin (1 μ g/mL), and Golgistop (1:1000) for 5 h. Cells were permeabilized and stained with IL4 and IFNG antibodies. The IL4+ cell population was sorted as above.

DNA extraction and preparation for MAP

Cells were incubated in buffer containing 10mM Tris-HCl pH 8, 400mM NaCl, 3mM EDTA pH 8, 1% SDS supplemented with proteinase K (40 μ g/mL) at 55°C overnight. DNA was then extracted using phenol:chloroform:isoamylalcohol, followed by isopropanol precipitation. DNA was fragmented to an average size of 500 bp

using a Diagenode Bioruptor, and Illumina sequencing adaptors were added prior to MAP as previously described (Skene et al. 2010). For each cell type, two biological replicates were prepared.

MAP chromatography

MAP was performed using two sequential rounds of chromatography as previously described (Skene et al. 2010) and tested by performing PCR for known methylated regions.

Bisulfite sequencing

DNA was sonicated for 10 s using a Diagenode Bioruptor prior to bisulfite treatment. Bisulfite treatment and sequencing were performed as previously described (Illingworth et al. 2008).

RNA extraction and reverse transcription

RNA was extracted using TRI Reagent (Sigma) according to the manufacturer's instructions and DNase-treated using DNA-free (Ambion). Reverse transcription was carried out using M-MLV reverse transcriptase (Promega) and random hexamer primers (Roche).

Quantitative PCR

Quantitative PCR was carried out using SYBR Green SensiMix (Quantace) according to the manufacturer's instructions on a LightCycler 480 (Roche). Primer sequences are available upon request.

Illumina BeadChip arrays

RNA was labeled using the TotalPrep RNA Amplification kit (Ambion) and hybridized to Illumina MouseWG-6 BeadChip arrays according to the manufacturer's instructions. Three biological replicates were carried out for each cell type. Bead level data were summarized using Illumina GenomeStudio, and data were normalized using the average normalization method in GenomeStudio. Subsequent analysis was carried out using GeneSpring GX10 (Agilent Technologies). In order to identify changes in gene expression between cell types, genes with low expression (Flags absent) across all cell types were removed, and statistical analysis was carried out to find genes that were differentially expressed between any of the cell types (one-way ANOVA, $P \leq 0.5$). Pairwise comparisons were performed on this gene set, and genes changing at least twofold were deemed to be differentially expressed. Brain gene expression studies were carried out on whole male brain from wild-type 5-wk-old C57/BL6 mice, and analysis was performed as for immune cells. Sample information is given in Supplemental Table S1.

Chromatin immunoprecipitation (ChIP)

ChIP was performed on Th1, Th2, and dendritic cells as previously described (Schmiedeberg et al. 2009). For each immunoprecipitation, cross-linked chromatin from ~4 million cells was incubated with either 5 μ g anti-H3K4me3 (07-473; Millipore) or 5 μ g anti-H3K27me3 (07-499; Millipore). Immunoprecipitated DNA was then analyzed using quantitative PCR or high-throughput sequencing. Prior to sequencing, Illumina adaptors were ligated to ChIP DNA samples as previously described (Skene et al. 2010).

Library preparation and Illumina sequencing

To produce Illumina libraries for sequencing, DNA to which Illumina adaptors had been attached was amplified by 10–12 cycles of

PCR with primers complementary to the adaptor sequences, using Phusion polymerase (Finnzymes), and the DNA was purified using QIAquick PCR Purification columns (Qiagen). The purified DNA was attached to an Illumina flow cell for cluster generation. Libraries were sequenced on the Illumina Genome Analyzer following the standard protocol to generate single-end 37-bp reads. Reads were mapped to the mouse genome (mm9-NCBI37; repeat-masked) using MAQ (Li et al. 2008) (<http://maq.sourceforge.net/>), and reads with a mapping score of ≥ 30 were retained. One lane of sequencing was carried out for each sample, and replicate samples were combined to give two lanes of sequencing for each cell type. Detailed sample information is given in Supplemental Table S1.

Analysis of high-throughput sequencing data

Mapped high-throughput sequencing data in the form of WIG files were analyzed as described previously (Illingworth et al. 2010) using a set of tools interfaced with an in-house version of the Galaxy application (Taylor et al. 2007). Briefly, raw sequencing data were normalized to the average read number across all samples of a given purification type in order to account for variable sequence depth between samples. Peak-finding was then carried out to identify regions of prominent enrichment using the following parameters: read height (H), length in bp (L), and gap permitted in the length parameter (G). These were adjusted for each purification type so that regions of known methylation or histone modification status were isolated. The X and Y chromosomes were excluded from analyses in order to avoid spurious results due to fluctuations in the proportion of male and female DNA in samples. All analysis parameters are outlined in Supplemental Table S1.

Peak saturation analysis was carried out using representative samples from each purification type (i.e., MAP-seq, H3K4me3 ChIP-seq, and H3K27me3 ChIP-seq). See Supplemental Methods and Supplemental Figure S9 for further details.

To identify regions showing differential enrichment for a particular modification (i.e., DNA methylation or H3K4me3), peaks for each sample were combined to give a set of regions of interest over which a sliding window analysis was performed. For each region, the average number of hits per base was calculated in 100-bp windows with a 20-bp slide. Values were then compared between samples giving a ratio for each window. If both windows being compared contained less than four reads, this ratio was set to one in order to remove spurious hits attributed to small fluctuations at low read depth. Differentially enriched regions were identified as those containing nine out of 13 contiguous windows with a \log_2 ratio ≥ 1.8 (MAP-seq) or a \log_2 ratio ≥ 2 (H3K4me3 ChIP-seq). These parameters were selected so that known differentially modified regions were identified (e.g., X-linked CGIs which are differentially methylated between males and females). Details of all differentially methylated CGIs identified are given in Supplemental Table S2.

Differentially modified regions were examined for association with CGIs, RefSeq genes, TSSs, and other genomic features using standard tools available on Galaxy. An association was defined as direct overlap with a particular feature by at least 1 bp. The CGIs examined were from the mouse CGI set characterized by Illingworth et al. (2010), and RefSeq-annotated genes were used as the gene set. CGIs were categorized as follows: TSS (overlap with the annotated start site of any RefSeq gene), intragenic (not overlapping the TSS but located anywhere in the transcription unit), and intergenic (no gene overlap). Brain methylation data was obtained from MAP-seq on cerebellum (Skene et al. 2010) and was analyzed in parallel with the other MAP-seq samples. Brain H3K4me3 data was from Thomson et al. (2010), ES cell RNAPII

ChIP-seq data was from Illingworth et al. (2010), and CAGE data was from a study by Faulkner et al. (2009). RNAPII peaks, as defined by Illingworth and colleagues, or regions with more than five CAGE tags, were taken as evidence of a TSS.

Analysis of gene ontology terms was carried out using the PANTHER database (<http://www.pantherdb.org>; Thomas et al. 2003). A version of the binomial test (Cho and Campbell 2000) was used to identify overrepresented ontology terms with $P \leq 0.05$, indicating significant enrichment. All CGI-associated genes were used as a reference.

Acknowledgments

We thank M. Waterfall for assistance with FACS sorting and L. Evenden and A. Condie of the Wellcome Trust Clinical Research Facility, Edinburgh, for carrying out the Illumina BeadChip experiments. *Schistosoma mansoni*-infected mice were a gift from P. Cook. We thank K. Auger and J. Parkhill for coordinating high-throughput sequencing and E. Gibson and J. Burton for processing the sequencing samples. We also thank L. Hing and D. DeSousa for help with bisulfite sequencing and T. Clouaire for helpful comments on the manuscript. This work was funded by a Wellcome Trust studentship to A.M.D. and by grants from the Wellcome Trust and Medical Research Council, UK.

References

- Allo M, Buggiano V, Fededa JP, Petrillo E, Schor I, de la Mata M, Agirre E, Plasm M, Eyras E, Elela SA, et al. 2009. Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nat Struct Mol Biol* **16**: 717–724.
- Antequera F, Boyes J, Bird A. 1990. High levels of de novo methylation and altered chromatin structure at CpG islands in cell lines. *Cell* **62**: 503–514.
- Ball MP, Li JB, Gao Y, Lee JH, LeProust EM, Park IH, Xie B, Daley GQ, Church GM. 2009. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nat Biotechnol* **27**: 361–368.
- Batsche E, Yaniv M, Muchardt C. 2006. The human SWI/SNF subunit Brm is a regulator of alternative splicing. *Nat Struct Mol Biol* **13**: 22–29.
- Begenisich T, Nakamoto T, Ovit CE, Nehrke K, Brugnara C, Alper SL, Melvin JE. 2004. Physiological roles of the intermediate conductance, Ca²⁺-activated potassium channel Kcnn4. *J Biol Chem* **279**: 47681–47687.
- Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, et al. 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**: 315–326.
- Bird AP, Taggart MH, Nicholls RD, Higgs DR. 1987. Non-methylated CpG-rich islands at the human alpha-globin locus: Implications for evolution of the alpha-globin pseudogene. *EMBO J* **6**: 999–1004.
- Cheng P, Nefedova Y, Miele L, Osborne BA, Gabrilovich D. 2003. Notch signaling is necessary but not sufficient for differentiation of dendritic cells. *Blood* **102**: 3980–3988.
- Cho RJ, Campbell MJ. 2000. Transcription, genomes, function. *Trends Genet* **16**: 409–415.
- Cruikshank W, Little F. 2008. Interleukin-16: The ins and outs of regulating T-cell activation. *Crit Rev Immunol* **28**: 467–483.
- Deng J, Shoemaker R, Xie B, Gore A, LeProust EM, Antosiewicz-Bourget J, Egli D, Maherali N, Park IH, Yu J, et al. 2009. Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming. *Nat Biotechnol* **27**: 353–360.
- Djuretic IM, Levanon D, Negreanu V, Groner Y, Rao A, Ansel KM. 2007. Transcription factors T-bet and Runx3 cooperate to activate Ifng and silence Il4 in T helper type 1 cells. *Nature Immunol* **8**: 145–153.
- Doi A, Park IH, Wen B, Murakami P, Aryee MJ, Irizarry R, Herb B, Ladd-Acosta C, Rho J, Loewer S, et al. 2009. Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells, and fibroblasts. *Nat Genet* **41**: 1350–1353.
- Edwards CA, Ferguson-Smith AC. 2007. Mechanisms regulating imprinted genes in clusters. *Curr Opin Cell Biol* **19**: 281–289.
- Faulkner GJ, Kimura Y, Daub CO, Wani S, Plessy C, Irvine KM, Schroder K, Cloonan N, Steptoe AL, Lassmann T, et al. 2009. The regulated retrotransposon transcriptome of mammalian cells. *Nat Genet* **41**: 563–571.
- Fiorini E, Merck E, Wilson A, Ferrero I, Jiang W, Koch U, Auderset F, Laurenti E, Tacchini-Cottier F, Pierres M, et al. 2009. Dynamic regulation of notch 1 and notch 2 surface expression during T cell development and activation revealed by novel monoclonal antibodies. *J Immunol* **183**: 7212–7222.
- Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. 2007. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130**: 77–88.
- Hardtke S, Ohl L, Forster R. 2005. Balanced expression of CXCR5 and CCR7 on follicular T helper cells determines their transient positioning to lymph node follicles and is essential for efficient B-cell help. *Blood* **106**: 1924–1931.
- Illingworth RS, Bird AP. 2009. CpG islands—“a rough guide”. *FEBS Lett* **583**: 1713–1720.
- Illingworth R, Kerr A, Desousa D, Jorgensen H, Ellis P, Stalker J, Jackson D, Clee C, Plumb R, Rogers J, et al. 2008. A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS Biol* **6**: e22. doi: 10.1371/journal.pbio.0060022.
- Illingworth RS, Gruenewald-Schneider U, Webb S, Kerr ARW, James KD, Turner DJ, Smith C, Harrison DJ, Andrews R, Bird AP. 2010. Orphan CpG islands identify numerous conserved promoters in the mammalian genome. *PLoS Genet* **6**: e1001134. doi: 10.1371/journal.pgen.1001134.
- Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, Cui H, Gabo K, Rongione M, Webster M, et al. 2009. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet* **41**: 178–186.
- Ji H, Ehrlich LI, Seita J, Murakami P, Doi A, Lindau P, Lee H, Aryee MJ, Irizarry RA, Kim K, et al. 2010. Comprehensive methylome map of lineage commitment from haematopoietic progenitors. *Nature* **467**: 338–342.
- Jones PA, Baylin SB. 2007. The epigenomics of cancer. *Cell* **128**: 683–692.
- Jones PA, Wolkowicz MJ, Harrington MA, Gonzales F. 1990. Methylation and expression of the Myo D1 determination gene. *Philos Trans R Soc Lond* **326**: 277–284.
- Junt T, Fink K, Forster R, Senn B, Lipp M, Muramatsu M, Zinkernagel RM, Ludewig B, Hengartner H. 2005. CXCR5-dependent seeding of follicular niches by B and Th cells augments antiviral B cell responses. *J Immunol* **175**: 7109–7116.
- Kornblihtt AR. 2006. Chromatin, transcript elongation, and alternative splicing. *Nat Struct Mol Biol* **13**: 5–7.
- Kumanogoh A, Shikina T, Suzuki K, Uematsu S, Yukawa K, Kashiwamura S, Tsutsui H, Yamamoto M, Takamatsu H, Ko-Mitamura EP, et al. 2005. Nonredundant roles of Sema4A in the immune system: defective T cell priming and Th1/Th2 regulation in Sema4A-deficient mice. *Immunity* **22**: 305–316.
- Lee D, Agarwal S, Rao A. 2002. Th2 lineage commitment and efficient IL-4 production involves extended demethylation of the IL-4 gene. *Immunity* **16**: 649–660.
- Li H, Ruan J, Durbin R. 2008. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res* **18**: 1851–1858.
- Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, et al. 2009. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**: 315–322.
- Liu WH, Lai MZ. 2005. Deltex regulates T-cell activation by targeted degradation of active MEK1. *Mol Cell Biol* **25**: 1367–1378.
- Lorincz MC, Dickerson DR, Schmitt M, Groudine M. 2004. Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. *Nat Struct Mol Biol* **11**: 1068–1075.
- Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD, Johnson BE, Hong C, Nielsen C, Zhao Y, et al. 2010. Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* **466**: 253–257.
- Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, et al. 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**: 766–770.
- Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, et al. 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**: 553–560.
- Mohn F, Weber M, Rebhan M, Roloff TC, Richter J, Stadler MB, Bibel M, Schubeler D. 2008. Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol Cell* **30**: 755–766.
- Mookherjee N, Brown KL, Bowdish DM, Doria S, Falsafi R, Hokamp K, Roche FM, Mu R, Doho GH, Pistollic J, et al. 2006. Modulation of the TLR-mediated inflammatory response by the endogenous human host defense peptide LL-37. *J Immunol* **176**: 2455–2464.

- Nawijn MC, Dingjan GM, Ferreira R, Lambrecht BN, Karis A, Grosveld F, Savelkoul H, Hendriks RW. 2001. Enforced expression of GATA-3 in transgenic mice inhibits Th1 differentiation and induces the formation of a T1/ST2-expressing Th2-committed T cell compartment in vivo. *J Immunol* **167**: 724–732.
- 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**: 46–58.
- Payer B, Lee JT. 2008. X chromosome dosage compensation: How mammals keep the balance. *Annu Rev Genet* **42**: 733–772.
- Pils MC, Pisano F, Fasnacht N, Heinrich JM, Groebe L, Schippers A, Rozell B, Jack RS, Muller W. 2010. Monocytes/macrophages and/or neutrophils are the target of IL-10 in the LPS endotoxemia model. *Eur J Immunol* **40**: 443–448.
- Rauch TA, Wu X, Zhong X, Riggs AD, Pfeifer GP. 2009. A human B cell methylome at 100-base pair resolution. *Proc Natl Acad Sci* **106**: 671–678.
- Reiner SL. 2007. Development in motion: Helper T-cells at work. *Cell* **129**: 33–36.
- Santangelo S, Cousins DJ, Winkelmann NE, Staynov DZ. 2002. DNA methylation changes at human Th2 cytokine genes coincide with DNase I hypersensitive site formation during CD4(+) T cell differentiation. *J Immunol* **169**: 1893–1903.
- Schilling E, Rehli M. 2007. Global, comparative analysis of tissue-specific promoter CpG methylation. *Genomics* **90**: 314–323.
- Schmiedebeg L, Skene P, Deaton A, Bird A. 2009. A temporal threshold for formaldehyde crosslinking and fixation. *PLoS ONE* **4**: e4636. doi: 10.1371/journal.pone.0004636.
- Schoenborn JR, Dorschner MO, Sekimata M, Santer DM, Shnyreva M, Fitzpatrick DR, Stamatoyannopoulos JA, Wilson CB. 2007. Comprehensive epigenetic profiling identifies multiple distal regulatory elements directing transcription of the gene encoding interferon-gamma. *Nat Immunol* **8**: 732–742.
- Skene PJ, Illingworth RS, Webb S, Kerr AR, James KD, Turner DJ, Andrews R, Bird AP. 2010. Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. *Mol Cell* **37**: 457–468.
- Sleutels F, Zwart R, Barlow DP. 2002. The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* **415**: 810–813.
- Smiraglia DJ, Rush LJ, Fruhwald MC, Dai Z, Held WA, Costello JF, Lang JC, Eng C, Li B, Wright FA, et al. 2001. Excessive CpG island hypermethylation in cancer cell lines versus primary human malignancies. *Hum Mol Genet* **10**: 1413–1419.
- Taylor J, Schenck I, Blankenberg D, Nekrutenko A. 2007. Using galaxy to perform large-scale interactive data analyses. In *Current protocols in bioinformatics* (ed. AD Baxeavanis et al.), pp. 10.5.1–10.5.25. Wiley, New York.
- Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, Diemer K, Muruganujan A, Narechania A. 2003. PANTHER: A library of protein families and subfamilies indexed by function. *Genome Res* **13**: 2129–2141.
- Thomson JP, Skene PJ, Selfridge J, Clouaire T, Guy J, Webb S, Kerr AR, Deaton A, Andrews R, James KD, et al. 2010. CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature* **464**: 1082–1086.
- Wakabayashi Y, Watanabe H, Inoue J, Takeda N, Sakata J, Mishima Y, Hitomi J, Yamamoto T, Utsuyama M, Niwa O, et al. 2003. Bcl11b is required for differentiation and survival of alphabeta T lymphocytes. *Nature Immunol* **4**: 533–539.
- Wan YY. 2010. Multi-tasking of helper T cells. *Immunology* **130**: 166–171.
- Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schubeler D. 2007. Distribution, silencing potential, and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet* **39**: 457–466.
- Winders BR, Schwartz RH, Bruniquel D. 2004. A distinct region of the murine IFN-gamma promoter is hypomethylated from early T cell development through mature naive and Th1 cell differentiation but is hypermethylated in Th2 cells. *J Immunol* **173**: 7377–7384.
- Wu J, Lingrel JB. 2005. Kruppel-like factor 2, a novel immediate-early transcriptional factor, regulates IL-2 expression in T lymphocyte activation. *J Immunol* **175**: 3060–3066.
- Zhang H, Meng F, Chu CL, Takai T, Lowell CA. 2005. The Src family kinases Hck and Fgr negatively regulate neutrophil and dendritic cell chemokine signaling via PIR-B. *Immunity* **22**: 235–246.

Received December 1, 2010; accepted in revised form April 15, 2011.