





# B1a and B2 cells are characterized by distinct CpG modification states at DNMT3A-maintained enhancers

Vinay S. Mahajan <sup>1,2,6</sup>, Hamid Mattoo<sup>1,3,6</sup>, Na Sun<sup>1,4,6</sup>, Vinayak Viswanadham <sup>1,5,6</sup>, Grace J. Yuen<sup>1</sup>, Hugues Allard-Chamard<sup>1</sup>, Maimuna Ahmad<sup>1</sup>, Samuel J. H. Murphy <sup>1</sup>, Annaiah Cariappa<sup>1</sup>, Yesim Tuncay<sup>1</sup> & Shiv Pillai <sup>1</sup>✉

The B1 and B2 lineages of B cells contribute to protection from pathogens in distinct ways. The role of the DNA CpG methylome in specifying these two B-cell fates is still unclear. Here we profile the CpG modifications and transcriptomes of peritoneal B1a and follicular B2 cells, as well as their respective proB cell precursors in the fetal liver and adult bone marrow from wild-type and *CD19-Cre Dnmt3a* floxed mice lacking DNMT3A in the B lineage. We show that an underlying foundational CpG methylome is stably established during B lineage commitment and is overlaid with a DNMT3A-maintained dynamic methylome that is sculpted in distinct ways in B1a and B2 cells. This dynamic DNMT3A-maintained methylome is composed of novel enhancers that are closely linked to lineage-specific genes. While DNMT3A maintains the methylation state of these enhancers in both B1a and B2 cells, the dynamic methylome undergoes a prominent programmed demethylation event during B1a but not B2 cell development. We propose that the methylation pattern of DNMT3A-maintained enhancers is determined by the coincident recruitment of DNMT3A and TET enzymes, which regulate the developmental expression of B1a and B2 lineage-specific genes.

<sup>1</sup>Ragon Institute of MGH, MIT, and Harvard, Cambridge, MA, USA. <sup>2</sup>Brigham and Women's Hospital, Boston, MA, USA. <sup>3</sup>Immunology and Inflammation Therapeutic Area, Sanofi, Cambridge, MA, USA. <sup>4</sup>Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA, USA. <sup>5</sup>Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA. <sup>6</sup>These authors contributed equally: Vinay S. Mahajan, Hamid Mattoo, Na Sun, Vinayak Viswanadham. ✉email: [pillai@helix.mgh.harvard.edu](mailto:pillai@helix.mgh.harvard.edu)

B cells are divided into two major sub-lineages. B2 cells, which constitute the majority of all B cells, are short-lived recirculating cells that are metabolically quiescent and can be activated in secondary lymphoid organs where they collaborate with helper T cells to generate specific antibodies. B1a cells, found predominantly in the peritoneal and pleural cavities, are long-lived self-renewing cells that produce natural antibodies as part of innate-like immunity<sup>1,2</sup>. Given the crucial role of DNA methylation in controlling gene expression, lineage specification, and the maintenance of cellular states, we explored the B1 and B2 sub-lineages through the lens of whole-genome cytosine modification dynamics<sup>3,4</sup>. Intergenic CpG island methylation has been linked to lineage specification in the immune system<sup>5</sup>, and it has been suggested that DNA methylation might protect specific cell types from the aberrant activation of transcription factors in a lineage-specific manner<sup>6</sup>. Genome-wide studies of cytosine methylation have been undertaken in human B cells, and it has been shown that early human B-cell development is characterized by the loss of CpG methylation at enhancer sites; additionally, the differentiation of B cells into memory and plasma cells is characterized by the loss of CpG methylation in regions of heterochromatin and a gain of methylation in polycomb-repressed regions<sup>7</sup>. However, the identification of human B1 cells is still controversial and prior studies on human B-cell methylomes did not address the B1 vs. B2 lineage decision.

The enzymes that add methyl marks to cytosines include DNA methyltransferase 1 (DNMT1) and the DNA methyltransferases 3a (DNMT3A) and 3b (DNMT3B). DNMT1 is termed a maintenance methyltransferase because it recognizes hemi-methylated sites in newly replicated DNA and restores symmetric CpG methylation. DNMT3A and DNMT3B are de novo methyltransferases that are active during early embryogenesis and modify unmethylated CpGs symmetrically<sup>8</sup>. Although 5-methylcytosine (5mC) is a stable modification, it can be converted to 5-hydroxymethylcytosine (5hmC) by methylcytosine dioxygenases of the Ten Eleven Translocation Family (TET1, TET2, and TET3)<sup>9–12</sup>. 5hmC itself is also a substrate for TET enzymes and can be progressively oxidized to 5-formyl (5fC) and 5-carboxy (5caC) cytosine, which are recognized and excised by thymine-DNA glycosylase, such that an unmethylated cytosine is restored by the base-excision repair pathway. Unlike 5mC, the 5hmC mark is not maintained by DNMT1 during cell division and appears to undergo DNA replication dependent loss, thereby providing an additional mechanism for DNA demethylation<sup>13</sup>. Thus, the TET enzymes play a critical role in active demethylation of mCpGs.

Hypomethylated CpGs mark gene regulatory regions such as promoters and some enhancers. The 5hmC mark may serve as an indicator of TET activity, but given that this mark is especially enriched in enhancer elements, it has been speculated that it may play an active, functional role beyond being just an intermediate in demethylation<sup>14–16</sup>. Several readers that specifically recognize methylated or unmethylated CpGs have been identified, but a specific reader for 5hmC has not yet been identified. Unlike DNMT1, which is associated with the DNA replication machinery, DNMT3A and the TET enzymes are recruited to specific genomic sites such as gene regulatory regions through their association with transcription factors. Engineered loss of *Dnmt3a* in mouse hematopoietic stem cells has been reported to result in malignant transformation in many hematopoietic lineages including the B lineage, resembling human B-CLL<sup>17–21</sup>. Leukemic transformation in the *Eμ-Tcl1*-transgenic model of CLL is also thought to be driven by the ability of TCL1 to inhibit DNMT3A<sup>22</sup>. Once a CpG methylation pattern is established across the genome by DNMT3A and TET enzymes during B-lineage commitment, it is subsequently maintained across cell

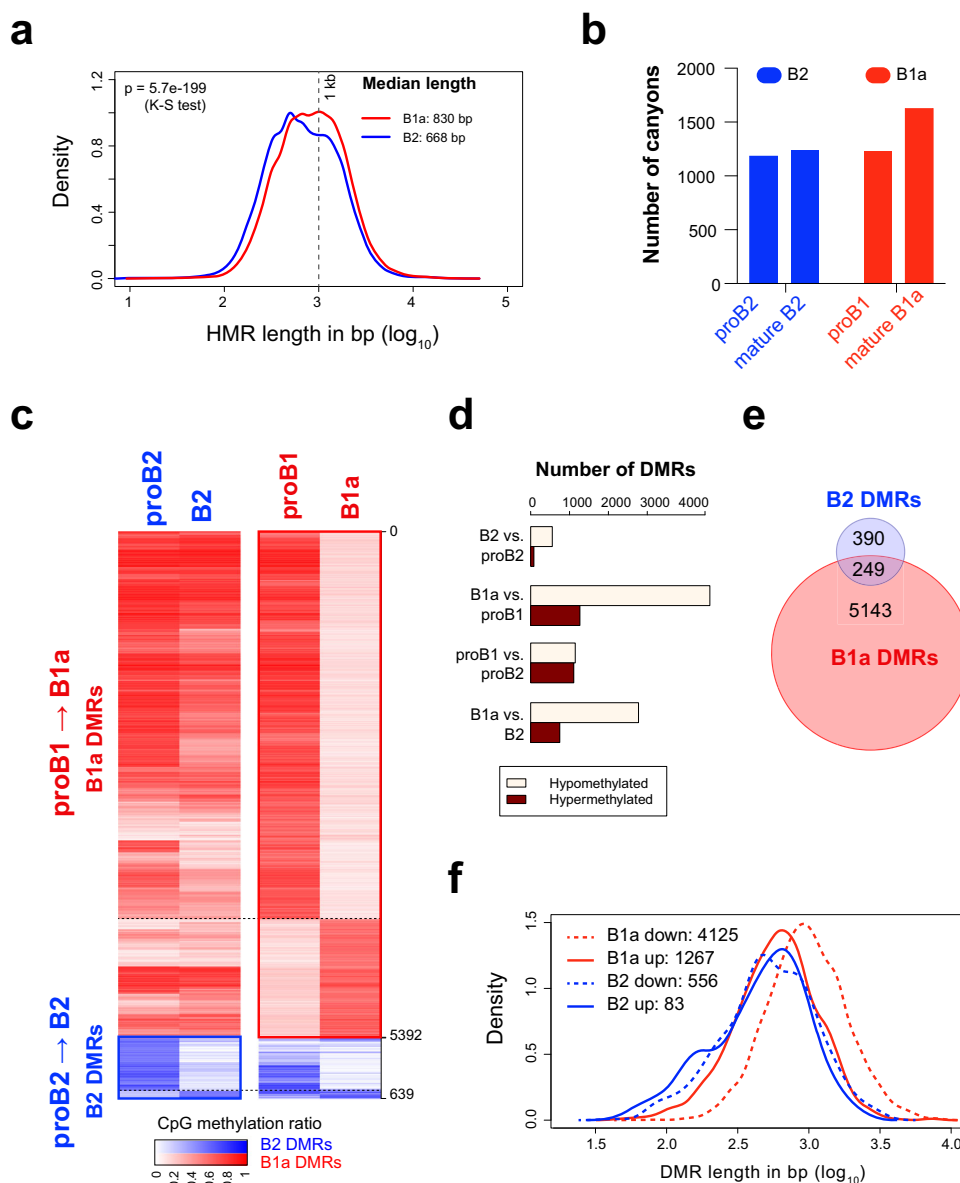
divisions by DNMT1. Mice lacking *Tet2* and *Tet3* in the B lineage retain normal numbers of follicular B2 cells but experience a partial block in early B-cell development and completely lack B1a and marginal zone B cells<sup>23</sup>, suggesting a continued role for DNA demethylation in establishing the B1a cell program after B lineage commitment. In this study, we elucidate the role of DNMT3A during and after B lineage commitment by studying *Dnmt3a*-floxed mice in the *CD19-Cre<sup>+/+</sup>* background at 4 weeks of age, well before the establishment of leukemia and at a time when the B1a cells are still polyclonal<sup>20,21,24</sup>, in order to examine the role of DNMT3A in the absence of the confounding effects of clonal somatic aberrations and widespread changes in CpG methylation that are known to accompany malignant transformation into CLL.

We profile cytosine modifications in B1a and B2 cells and their precursors and show that both B1a and B2 cells exhibit a dynamic layer of lineage-specific and DNMT3A-maintained CpG methylation that is superimposed on a shared foundational methylome. We find that continued action of DNMT3A is required to counteract TET activity in this dynamic methylome layer, and the sites whose CpG methylation is maintained by DNMT3A after B lineage commitment are highly enriched for enhancers, which we have termed DNMT3A-maintained enhancers (DMEs). DME methylation is linked to the regulated expression of lineage-specific genes and reveals a novel role for DNMT3A in both marking and maintaining DNA methylation at specific enhancers that modulate the fate of developing lymphocytes. Our results also provide an epigenetic framework for understanding the reported BCR-dependent plasticity in B-cell sub-lineages<sup>25</sup>.

## Results

**B1a cell development is characterized by a pronounced loss of CpG modification.** We surveyed the genome-wide CpG modification profiles in murine B1a and B2 cell lineages, by performing whole-genome bisulfite sequencing (WGBS) of proB2 cells (Hardy fractions B and C from bone marrow), splenic follicular B2 cells, proB1 cells (Hardy fractions B and C from fetal liver) and peritoneal B1a cells; these populations will hereon be referred to as proB2, B2, proB1 and B1a cells, respectively. We identified about 40,000–50,000 hypomethylated regions (HMRs) that were dispersed across 1.7% to 1.9% of the genome, largely in intergenic and intronic regions in all four cell types<sup>26</sup>. Interestingly, the HMRs in wild-type B1a cells were found to be significantly longer than those in the B2 cells (median lengths of 830 bp and 668 bp, respectively; K–S test  $p < 1e-10$ ), indicating that the B1a lineage may favor CpG demethylation at HMRs across the genome (Fig. 1a). Large HMRs over 3.5 kb have been previously referred to as canyons in studies on murine hematopoietic stem cells<sup>27</sup>. Using a similar cutoff of 3.5 kb, we identified methylation canyons in all four B-cell populations. Mature B1a cells exhibited 30–35% more canyons than all other B-cell stages examined (Fig. 1b). This developmental increase in the number of canyons in B1 cells was reminiscent of the canyon expansion observed in *Dnmt3a*-deficient murine HSCs<sup>27</sup> and suggested that changes in DNMT3A activity may contribute to the programmed demethylation observed during B1a cell differentiation.

We next identified differentially methylated regions (DMRs) that arise during the development of the B1a and B2 cell lineages. DMRs are regions with a high probability of differential methylation where an HMR exists in one methylome but not at the same location in the other. DMRs were particularly abundant between the proB1 and B1a stages. They encompassed ~6 million bp across 5,392 genomic intervals in the B1a lineage (0.42% of the genome), but only about 0.5 million bp across 639 intervals in the B2 lineage (0.04% of the genome), suggesting that the regulation

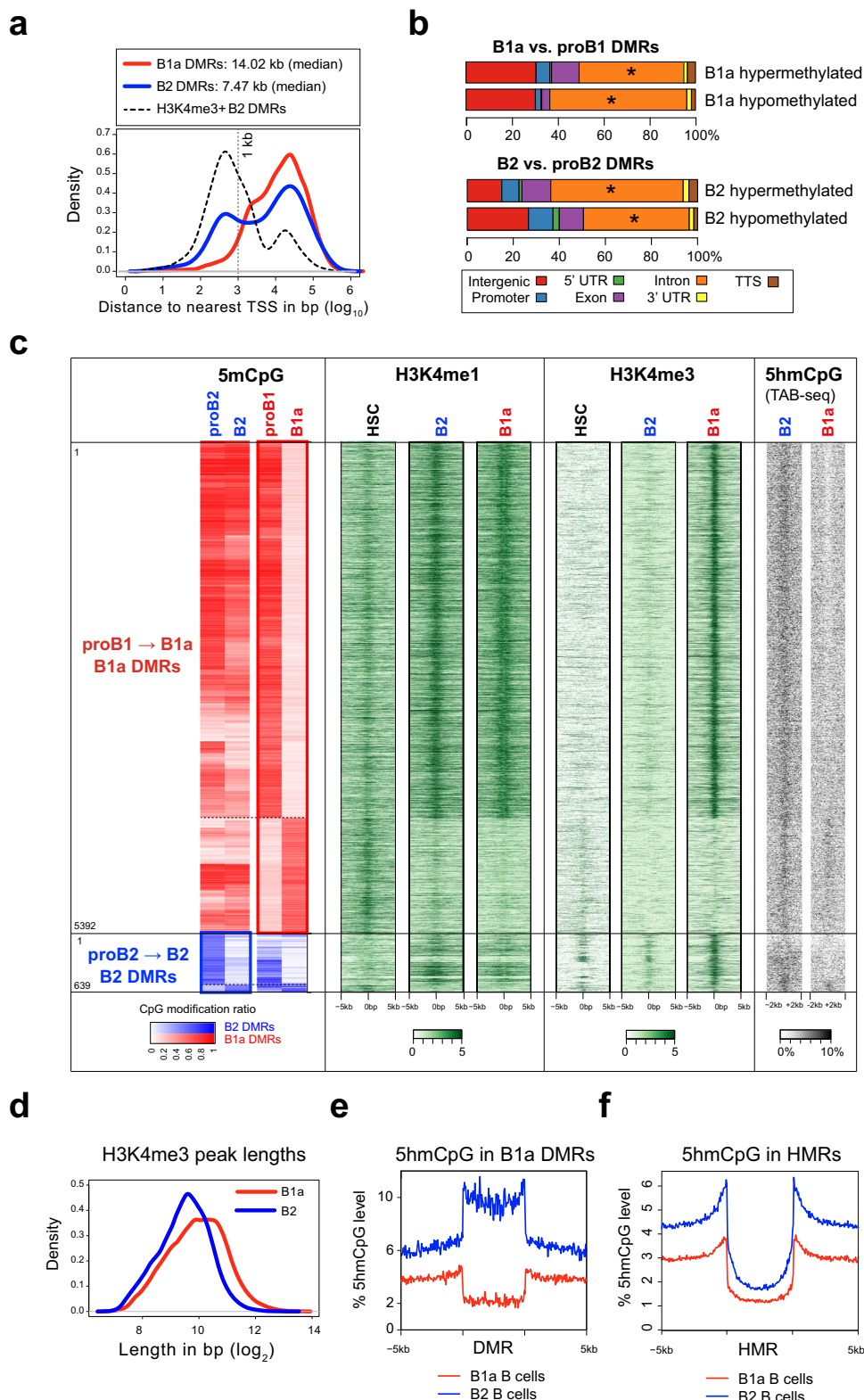


**Fig. 1 B1a cell development is marked by a pronounced loss of CpG modification.** **a** Distribution of HMR lengths in proB and B cells in the B1a and B2 lineage. B1a, B2 and proB2 cells were pooled from three mice each, and proB1 cells were pooled from 14 fetal livers. **b** Number of canyons (HMR length >3.5 kb) identified in proB2, B2 (follicular B), proB1 and B1a cells ( $n = 1$  methylome each; DNA was pooled from multiple mice as indicated in Fig. 1a). **c** CpG modification profiles in developmental DMRs in the B1a and B2 lineages. Average CpG modification ratio in the B1a and B2 DMRs in proB1, B1a, proB2 and B2 cells is depicted in the heatmap. **d** The number of DMRs identified by pairwise comparisons between proB and mature B cells and between corresponding stages of the B1a and B2 cell lineages. **e** Overlap between B2 and B1a DMRs; DMRs with over 50% overlap were considered overlapping. **f** Distribution of development DMR lengths in proB and B cells in the B1a and B2 lineage. Source data are provided as a Source Data file.

of CpG methylation is a far more prominent feature of B1a cell development, compared to B2 cell development (Fig. 1c–e). The majority of the B1a DMRs (~75%) exhibited a loss of CpG modification during B1a cell development (Fig. 1c, d). The median size of the observed B1a DMRs was 857 bp, and ~12% of the hypomethylated B1a DMRs fell within methylation canyons in B1a cells (Fig. 1f)

**Programmed demethylation in B1a cells occurs at enhancers that are demethylated and re-methylated in B2 cells.** The majority of B1a and B2 DMRs were present at a distance from the transcription start site (TSS) in intergenic or intragenic spaces, compatible with a possible enhancer function (Fig. 2a, b). Indeed, many prior studies of CpG methylomes in adult somatic tissues

have found that variably methylated regions are strongly enriched at enhancer sites. Analysis of publicly available ChIP-Seq data for H3K4me1 and H3K4me3 modifications in mature B2 and B1a cells<sup>28</sup> showed that the DMRs that underwent developmental demethylation in both B1 and B2 cells co-localized with H3K4me1 peaks in mature B1a and B2 cells. This suggested that these DMRs might represent developmentally demethylated enhancers that modulate lineage-specific gene expression (Fig. 2c). To assess the significance of these enhancers in the context of hematopoietic development, we also examined H3K4me1 and H3K4me3 marks in HSCs using published ChIP-Seq data. The DMRs that were developmentally hypermethylated in mature B2 or B1a cells exhibited little H3K4me1 signal in B cells, but did bear an H3K4me1 signature in HSCs, suggesting



that these hypermethylated DMRs may represent enhancers that were developmentally silenced by CpG methylation (Fig. 2c). The B1a DMRs exhibited low levels of H3K4me1 in HSCs, suggesting that they may represent primed enhancers in HSCs that subsequently become fully active in the B lineage.

In B2 cells, only a small subset of developmentally demethylated DMRs that were present in the close proximity of annotated transcription start sites (TSS) exhibited an H3K4me3 signal,

consistent with a promoter function (Fig. 2a). However, in B1a cells, most DMRs that were developmentally demethylated acquired some H3K4me3 trimethylation marks (Fig. 2c). Although H3K4me3 is generally associated with promoters, it can be found in enhancers in some contexts<sup>29,30</sup>. The recognition of unmethylated CpGs by the SET1 H3K4 trimethylase complex through CFP1 has been reported to restrict this chromatin modifying enzyme complex to unmethylated promoter regions

**Fig. 2 Extensive loss of enhancer methylation during B1a cell development.** **a** Distribution of distance to nearest TSS for B1a (red) and B2 (blue) DMRs. B2 DMRs that overlap with H3K4me3 ChIP-Seq peaks in B2 cells (black) are separately depicted. **b** DMR genomic features. Asterisks depict genomic features that are significantly enriched with respect to HMRs. **c** The heatmap depicting the CpG modification ratio in B1a and B2 DMRs is reproduced from Fig. 1c to provide a frame of reference on the far left. H3K4me1 and H3K4me3 histone marks are correspondingly plotted +/- 5 kb upstream or downstream of the B1a and B2 DMRs in HSCs 27, B1a and B2 cells 28. An aggregate signal derived from the 3 published replicates per condition is plotted. 5hmCpG levels +/- 2 kb of the DMRs are shown on the far right as measured by TAB-Seq in B1a and B2 cells (one replicate per condition, using cells pooled from three mice each). **d** Distribution of H3K4me3 peak lengths in B1a (red) and B2 (blue) cells. **e** Average profile of 5hmCpG modification (% of total CpGs) in the B1a DMRs in B1a (red) and B2 (blue) cells. The DMRs are scaled to the same length and 5 kb flanking regions are depicted. **f** Average profile of 5hmCpG modification (% of total CpGs) in the HMRs in B1 and B2 cells and 5 kb flanking regions. Source data are provided as a Source Data file.

and prevent CpG methylated enhancers from acquiring ectopic H3K4me3 marks<sup>31</sup>. This CpG methylation barrier to H3K4me3 acquisition appears to break down at many enhancer sites in B1a cells, and the developmental acquisition of ectopic H3K4me3 marks at demethylated enhancers is unique to B1a cells. Interestingly the H3K4me3 peaks were wider in B1a cells than B2 cells across the genome, consistent with a global increase in HMR lengths in B1a cells (Fig. 2d).

WGBS revealed that 90% of the CpGs in DMRs that were developmentally demethylated in B1a cells were modified in B2 cells (Fig. 2c). As WGBS does not distinguish between cytosine methylation and hydroxymethylation, we also analyzed 5-hydroxymethylation (5hmC) levels in mature B1a and B2 cells using TET-assisted bisulfite sequencing (TAB-Seq)<sup>32</sup> (Fig. 2c, far right). Given that the bulk of the bisulfite protection arose from 5mCpG rather than 5hmCpG, we focused on 5mCpG in our analyses. The majority of CpGs in B1 DMRs were modified in B2 cells. About 10 percent of these CpG were hydroxymethylated, indicating that TET enzymes are active at these sites even in B2 cells (Fig. 2e). In contrast, our analyses showed no appreciable level of 5hmC at the B1 DMRs in B1 cells. As previously described in HSCs, 5hmCpGs, a mark of TET activity, were found to be enriched at HMR boundaries in both B1a and B2 cells (Fig. 2f). TET activity at HMR boundaries plays a critical role in maintaining the HMRs by opposing the action of DNMT3A<sup>27,33</sup>.

### B1 lineage-specific gene expression is linked to enhancer demethylation.

Examination of gene expression data in B-cell subsets from the ImmGen Consortium indicated that the genes in the proximity of B1a cell-specific DMRs that lose methylation during B1a cell development are upregulated during differentiation from proB1 to mature B1a cells (either fetal liver Fraction E or peritoneal B1a cells) (Fig. 3a)<sup>34</sup>. These genes are enriched for pathways involved in B-cell receptor signaling and activation (Fig. 3b). In contrast, B1a cell-specific DMRs that gain methylation during development were not associated with statistically significant changes in the expression of neighboring genes (Fig. 3a). Unlike the hypomethylated DMR-proximal genes, the genes in the vicinity of hypermethylated or decommissioned enhancers are not linked to typical lymphocyte functions (Supplementary Table 1). B2 DMRs were less numerous and exhibited little correlation with gene expression changes during development (Supplementary Fig. 1). Although some genes exhibited multiple DMRs in their vicinity, there was no correlation between the number of DMRs and gene expression (Supplementary Fig. 2). Genes whose expression was increased during B1a cell development exhibited a significant overlap with genes that were closest to DMRs that lost CpG methylation during B1a cell development ( $p < 0.01$ ; Fisher's exact test) (Fig. 3c). Of these genes, 40% encode B1a lineage-specific genes, that include transcription factors such as *Arid3a*, *Bhlhe41* and *Zbtb32*, and signaling regulators such as *Siglecg* and *Rasgrp1*, that are known to be required for B1a cell development and function (Fig. 3d)<sup>35–38</sup>.

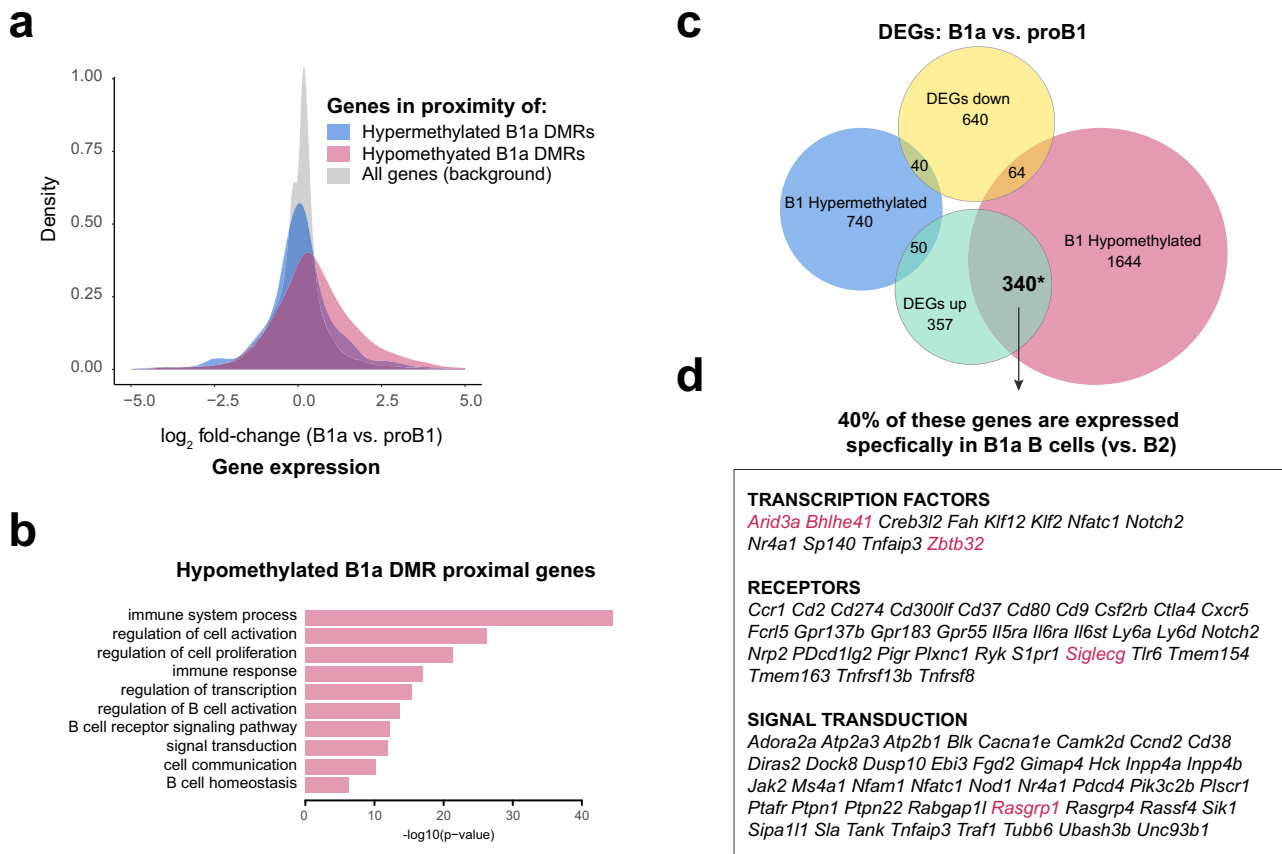
### Deficiency of *Dnmt3a* in the murine B lineage results in the selective expansion of B1a cells.

In order to evaluate the possible contribution of DNMT3A to the B1 vs. B2 cell fate decision, we used *Cd19-Cre* to conditionally delete a floxed *Dnmt3a* allele (*Dnmt3a*<sup>2lox</sup>) in the B lineage, specifically from the proB1 and proB2 stages onwards<sup>39,40</sup>. The direct or indirect loss of DNMT3A in the B lineage has been shown to result in the leukemic transformation of B1a cells<sup>20–22</sup>. Similarly, *Cd19-Cre*<sup>+/+</sup> *Dnmt3a*<sup>2lox/2lox</sup> (*Dnmt3a*<sup>-/-</sup>) mice also exhibited a marked expansion of B1a B cells that progressed to monoclonal leukemia by 9 months of age with a 100% penetrance (Fig. 4a). The increase in B1a cells was initially detectable in the peritoneal cavity as early as 4 weeks of age (Fig. 4a) and became evident in the spleen and bone marrow after 10 weeks of age. The peritoneal B1b B-cell population was progressively displaced by the expanding B1a population with age. We observed no increase in the proportion of CD19<sup>+</sup> B1a lineage progenitors or proB1 cells in the fetal liver<sup>41</sup>, suggesting that the expansion of B1a cells was driven by increased self-renewal or proliferation in the mature B1a cell compartment (Supplementary Fig. 3a). There was a slight decrease in the number of follicular B cells in the spleen and no change in the number of developing B2 cells in the bone marrow of *Dnmt3a*<sup>-/-</sup> mice (Fig. 4b and Supplementary Fig. 3b). The number of marginal zone (MZ) B cells was also reduced at 8 weeks of age (Fig. 4b and Supplementary Fig. 3c). Changes in the MZ B-cell compartment could not be assessed beyond 12 weeks of age, as the splenic marginal zone niche was obliterated by infiltrating B1a cells and MZ B cells were no longer detectable.

At 4 weeks of age, B1a cells lacking DNMT3A remained largely polyclonal, similar to wild-type B1a cells (Fig. 4c and Supplementary Fig. 4). Staining of Igk and Igl light chains showed that the expanded B1a cell pool in the peritoneal cavity underwent a gradual restriction in repertoire from a polyclonal to monoclonal state. Of the eight mice analyzed beyond 30 weeks of age, seven showed only Igk staining while one showed only Igl staining. *Dnmt3a*<sup>-/-</sup> B1a cells exhibited higher Ki-67 staining, consistent with increased self-renewal in this compartment (Fig. 4d). IgH rearrangements from the clonal leukemias were amplified and sequenced, and like unmutated human CLL, were found to be free of somatic hypermutation. To avoid potential confounding effects arising from oncogenic somatic mutations that may have been selected in a leukemic *Dnmt3a*<sup>-/-</sup> clone at a later time point, we used peritoneal *Dnmt3a*<sup>-/-</sup> B1a cells from 4-week old mice, when the cells were still in a polyclonal state, for all subsequent analyses.

### DNMT3A loss reveals a B lineage-specific foundational methylome shared in B1a and B2 cells.

We identified a total of 9,401 and 9,765 genomic intervals in the B2 and B1a cell lineages, respectively, that exhibited differential CpG modification across proB, mature B (wild-type) and *Dnmt3a*-deficient mature B cells, and henceforth refer to them as B2 and B1a DNMT3A-dependent DMRs (DDMRs), respectively (Fig. 5a). Remarkably, the



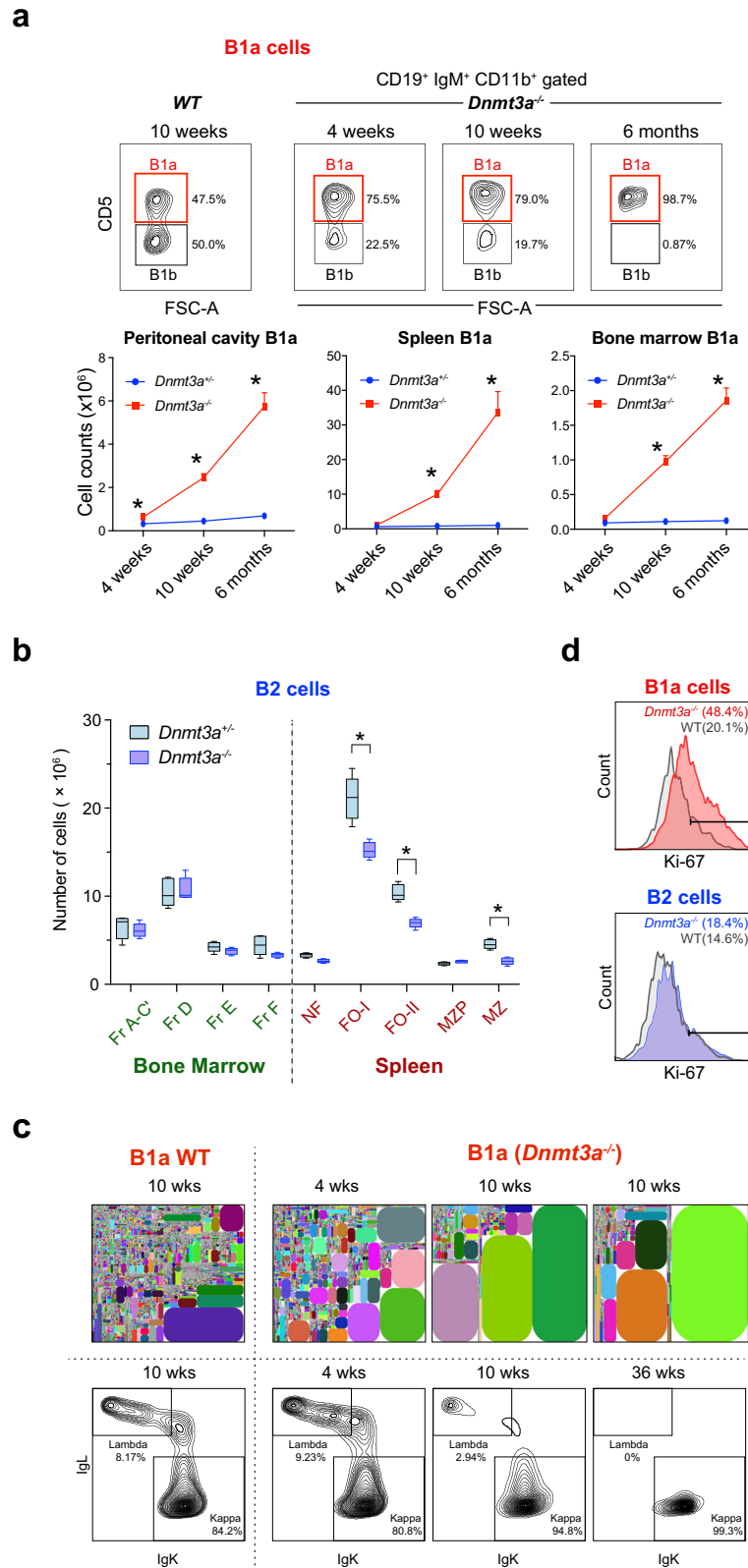
**Fig. 3 B1 lineage-specific gene expression is linked to enhancer demethylation.** **a** Distribution of fold-changes in gene expression between mature B1a (peritoneal B1a or fetal liver FrE) and proB1 (fetal liver FrBC) cells. Genes in the proximity of developmentally hypomethylated or hypermethylated B1a DMRs are depicted alongside the set of all genes. Microarray data from the ImmGen Consortium was used for this analysis. **b** Enriched Gene Ontology (Biological Process) terms for genes in the proximity of hypomethylated B1a DMRs. **c** Overlap between differentially expressed genes and genes in the proximity of DMRs between B1a (peritoneal B1a or fetal liver FrE) and proB1 (fetal liver FrBC) cells. Gene sets with a significant overlap are marked with an asterisk ( $p < 0.001$ , Fisher's exact test). **d** B1a lineage-specific genes in the significantly overlapping set shown in **c**. Source data are provided as a Source Data file.

lineage-specific methylation patterns in mature B2 and B1a cells were almost completely erased in the absence of DNMT3A, unmasking a foundational methylome that was virtually identical in B1a and B2 cells (Fig. 5b). Indeed, there were only 34 differentially methylated intervals between *Dnmt3a*<sup>-/-</sup> B1a and B2 cells. Interestingly, the CpG methylation profile of wild-type B1a cells resembled *Dnmt3a*<sup>-/-</sup> B1a and B2 cells, suggesting that DNMT3A activity may be physiologically downregulated during B1a cell development (Fig. 5b). We found that the number of canyons increased in *Dnmt3a*-deficient B cells compared to their wild-type counterparts (Fig. 5c) and appeared to arise from the expansion and amalgamation of smaller HMRs present at the respective sites in wild-type B cells. The canyons in *Dnmt3a*-deficient B2 and B1a cells exhibited a very high degree of overlap (98%) and were strongly enriched for genes encoding transcription factors, such that for nearly one-third of these canyons, the closest gene encoded a transcription factor.

Comparison with publicly available WGBS data from murine wild-type and *Dnmt3a*-deficient bone marrow hematopoietic stem cells (HSCs) suggested that genomic loci exhibiting *Dnmt3a*-dependent CpG methylation showed very little overlap between HSCs and B cells (Fig. 5d)<sup>27</sup>. Of note, given the use of *Cd19-Cre* for deleting the *Dnmt3a*-floxed allele, the loss of DNMT3A occurs after B lineage commitment in the mice studied. Interestingly, the B lineage DDMRs were predominantly

methylated in HSCs even in the absence of *Dnmt3a* (Fig. 5e). This suggests that the foundational and dynamic methylomes are recast following differentiation into the B lineage whereupon the CpG methylation at the B lineage DDMRs becomes dependent upon DNMT3A, perhaps to counteract the activity of TET2 as suggested by the enrichment of 5hmC at these sites (Fig. 5f, g, h).

The DDMRs were stratified into 12 groups based on their patterns of CpG modification; six patterns accounted for >97% of all DDMRs in B2 and B1a cells (Supplementary Fig. 5a, f). As previously observed, the majority (~75%) of the B1a DDMRs exhibit a loss of CpG modification during the differentiation of proB1 cells into B1a cells (Fig. 5f). Notably, the DDMRs that undergo partial demethylation from the proB1 to the mature B1a stage exhibited almost complete demethylation in *Dnmt3a*<sup>-/-</sup> B cells (Pattern 7 in Fig. 5f), suggesting that even the reduced level of CpG methylation observed in the DDMRs in wild-type B1a cells is maintained by DNMT3A. In B2 cells, which exhibited >65% CpG modification in the DDMRs, about 10% of the CpGs in the B2 DDMRs were 5-hydroxymethylated indicative of ongoing TET activity at these sites (Fig. 5g). In contrast, our analysis showed no appreciable level of 5hmCpG at the B1a DDMRs in B1a cells (Fig. 5h). This suggests that while TET can convert a portion of mCpG generated by DNMT3A at B2 DDMRs to 5hmCpG in B2 cells, the developmentally dampened activity of DNMT3A at B1a DDMRs in B1a cells may lead to a



paucity of the mCpG substrate for TET in B1a cells. In a majority of the B lineage DDMRs, DNMT3A appears to function as a maintenance methyltransferase at the DDMRs even in non-dividing B2 cells, in a manner that is distinct from the replication-coupled maintenance methylation function performed by DNMT1. This set of DDMRs may, therefore, be viewed as the DNMT3A-maintained dynamic methylome in B1a and B2 cells.

**DNMT3A-maintained enhancers (DMEs) represent a novel methylation-sensitive subset of enhancers.** The DDMRs were predominantly in intergenic and intronic regions at a distance from the TSS, consistent with an enhancer function (Fig. 5i, j). To assess this, we examined chromatin features at the DDMRs using CUT&RUN with antibodies against H3K4me1, H3K4me3 and H3K27ac<sup>42</sup> and ATAC-Seq<sup>43</sup>. In both B1a and B2 cells, the

**Fig. 4 *Dnmt3a* deficiency in the B lineage results in the selective expansion of B1a cells.** **a** Expansion of B1a cells (CD19<sup>+</sup> IgM<sup>+</sup> CD11b<sup>+</sup> CD5<sup>+</sup>) in the peritoneal cavity, spleen and bone marrow of *Dnmt3a*<sup>-/-</sup> mice (mean + SD). The top panel shows the gating strategy used to distinguish B1a and B1b cells. Five wild-type and *Dnmt3a*<sup>-/-</sup> mice were analyzed in each group. Statistically significant differences are indicated with asterisks. **b** Number of bone marrow B2 cell precursors (Fractions A-C', D, E, F) and mature splenic B2 cells (newly formed (NF), follicular (FO-I and FO-II), marginal zone precursor (MZP), and marginal zone (MZ) B cells) from 8 week old *Dnmt3a*<sup>-/-</sup> mice and their littermate controls. The experiment was performed two times and a total of eight mice in each group were analyzed. The box and whiskers mark the 25th to 75th percentiles and the minimum or maximum values. Statistically significant differences (two-sided unpaired t-tests, FDR > 1%) are indicated with asterisks ( $p = 0.001104$  (FO-I),  $p = 0.000087$  (FO-II),  $p = 0.000175$  (MZ)). **c** IgH repertoire of peritoneal B1a cells in wild-type and *Dnmt3a*<sup>-/-</sup> mice at varying ages (top row). Each clonotype is represented by a distinctly colored rectangle, with an area proportional to the fraction of total reads. Kappa and lambda Ig light chain staining in B1a cells from wild-type and *Dnmt3a*<sup>-/-</sup> mice at varying ages is also shown (bottom row). **d** Ki-67 levels in B1a cells from the peritoneal cavity and B2 cells from the spleens of 4-week-old wild-type and *Dnmt3a*<sup>-/-</sup> mice. Source data are provided as a Source Data file.

DDMRs were enriched for the H3K4me1 histone modification, a widely studied mark of enhancers (Fig. 6a, b). Indeed, the DDMRs accounted for approximately 11% of all enhancers defined by non-TSS H3K4me1 peaks (7,990 of 74,344) in B cells. Developmentally hypomethylated DDMRs (B1a DDMR patterns 1 and 7, and B2 DDMR pattern 7), acquired H3K27ac marks, which is indicative of an active enhancer state, during development (Fig. 6a, b). B2 and B1a DDMRs overlapped with sites of enhanced chromatin accessibility. In both sets of DDMRs, chromatin accessibility was inversely correlated with CpG methylation (Spearman's correlation coefficient = -0.26 and -0.57, for B2 and B1a DDMRs, respectively). Given that the majority of the DDMRs were partially hypomethylated in wild-type B1a cells, the loss of DNMT3A had a more profound impact on the CpG methylation of DDMRs in B2 cells, concomitant with a gain in chromatin accessibility (Fig. 6b, c). This suggests that CpG methylation may constrain chromatin accessibility at these DNMT3A-maintained enhancers that we refer to as DMEs.

In B2 cells, as is typical of active enhancers, DMEs exhibited both 5mCpG and 5hmCpG, and were marked by H3K4me1 but not H3K4me3 (Fig. 6a). The few B2 DMEs (pattern 1) that were enriched in H3K4me3 were also enriched in TSS, suggesting that they are promoters (Supplementary Fig. 6). In contrast, in B1a cells, the developmentally hypomethylated DMEs (especially pattern 1) exhibited both H3K4me1 and H3K4me3 marks, although they exhibited a minimal overlap with promoters of annotated transcripts (Fig. 6b, d, and Supplementary Fig. 6). As illustrated in Fig. 2c using publicly available CHIP-seq data, and in Fig. 6b with a different antibody in a CUT&RUN assay, we show that acquisition of H3K4me3 marks at DMEs is a unique feature of B1a cell development. H3K4me3 is prominent at promoters, but detectable levels of this modification are also observed at active enhancers bound by RNA Pol II<sup>44</sup>. Although CpG methylation in the DMEs was largely erased in the *Dnmt3a*-deficient B1a and B2 cells, the ablation of *Dnmt3a* had little impact on H3K4me1 or H3K4me3 levels in the DMEs in either B1a or B2 cells and did not result in the acquisition of H3K4me3 marks at DMEs in *Dnmt3a*-deficient B2 cells (Fig. 6a, b). Complete loss of CpG methylation in the DMEs in *Dnmt3a*-deficient B1a cells was also associated with an increase in H3K27ac marks in the B1a DMEs bearing ectopic H3K4me3 enhancer marks (Pattern 1), suggesting that CpG methylation may constrain the activity of DMEs with H3K4me3 marks (Fig. 6b, d). In contrast, in B2 cells, the levels of H3K27ac, H3K4me1 or H3K4me3 in the DMEs were largely unchanged upon DNMT3A loss.

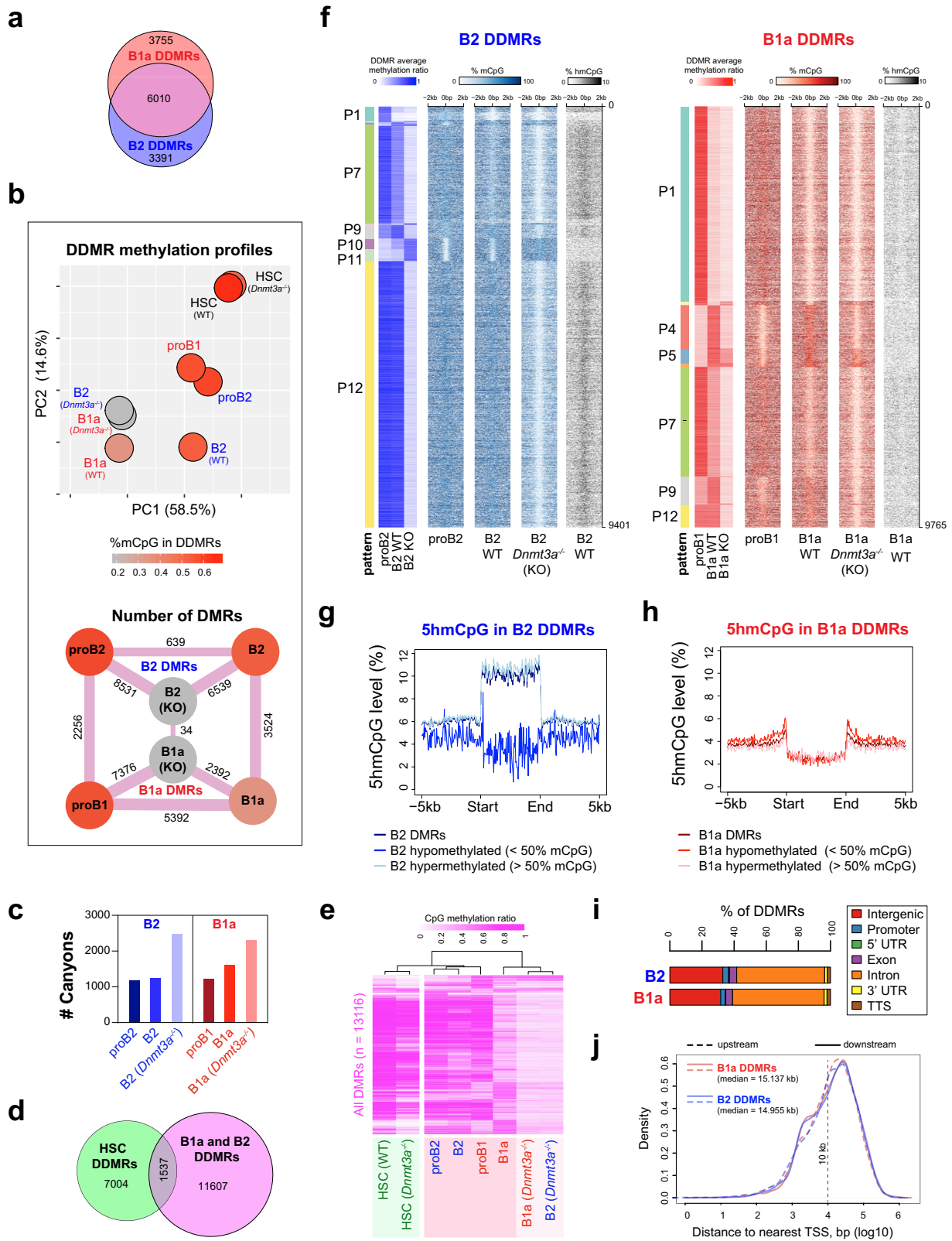
A small number of DDMRs (patterns 10 and 11) exhibited a paradoxical gain in methylation in *Dnmt3a*-knockout B2 cells compared to wild-type. This may be analogous to the aberrant hypermethylation driven by DNMT3B in the absence of DNMT3A in hematopoietic stem cells<sup>45</sup>. Interestingly, these sites were devoid of enhancer marks in our data (Fig. 6a), indicating

that they are not functional enhancers in B cells. This is corroborated by publicly available ChromHMM data on immune cells types (such as CH12 or mouse spleen) where these intervals are largely unmarked, making it harder to ascertain a specific function for these intervals (Supplementary Fig. 7)<sup>46</sup>. We also observed that developmentally hypermethylated B1 DDMRs (Patterns 4 and 9) bore the H3K4me1 mark at the proB1 stage but lost it in mature B1a cells, suggesting that these DDMRs represent enhancers that were developmentally decommissioned by DNMT3A-mediated CpG methylation. Failure to decommission this set of proB1 enhancers in *Dnmt3a*-deficient B1a cells may have contributed to the preleukemic phenotype observed in *Dnmt3a*-deficient B1a cells. The developmentally methylated B1a DDMRs also do not exhibit 5hmCpG marks, suggesting that DNMT3A may not be required for their continued methylation. Failure to methylate patterns 4 and 9 DDMRs is not accompanied by H3K4me3 acquisition. The impact of DNMT3A loss on the pattern of histone marks and cytosine modifications in the two largest categories of B1a and B2 DDMRs is depicted in Fig. 5D.

#### ***Dnmt3a*-dependent CpG methylation of DMEs controls lineage-specific gene expression in B cells.**

The global loss of DME methylation in *Dnmt3a*-deficient B cells resulted in widespread perturbation in gene expression in both B1a (2,872 genes) and B2 (2,018 genes) B cells. The expression of 1,002 of these genes was concordantly altered in B1a and B2 cells, and the expression of 1,710 and 946 genes was altered in a lineage-specific manner in B1a and B2 cells, respectively (Fig. 7a, b). In both *Dnmt3a*-deficient B1a and B2 cells, there was little correlation between the differential expression of genes and the presence of a DME in their vicinity (Fig. 7c). We surmised that this was because DME-proximal genes that are differentially expressed in *Dnmt3a*-deficient B cells secondarily influence genes lacking a DME in their vicinity. In order to explore this notion, we repeated this analysis after stratifying the differentially expressed genes into sets, which may be enriched for master regulatory genes such as lineage-defining genes or those encoding transcription factors. We found that B1a and B2 lineage-specific genes as well as genes that are upregulated during B1a and B2 cell development were significantly enriched in the vicinity of DMEs whose methylation was maintained by DNMT3A (Fig. 7c). Interestingly, genes in the vicinity of DMEs that preserved their CpG modifications despite the loss of DNMT3A (B1a pattern 5, and B2 patterns 10 and 11) were not associated with lineage-specific, developmental or *Dnmt3a*-dependent changes in gene expression. We had previously noticed that the gene closest to nearly one-third of the hypomethylation canyons in *Dnmt3a*<sup>-/-</sup> B cells encoded a transcription factor (TF). This also suggested to us that although the presence of DMEs was not directly correlated with global changes in gene expression upon DNMT3A loss, enhancer demethylation at TF genes may result in widespread secondary changes in gene expression. Indeed, we found that the TF genes that were



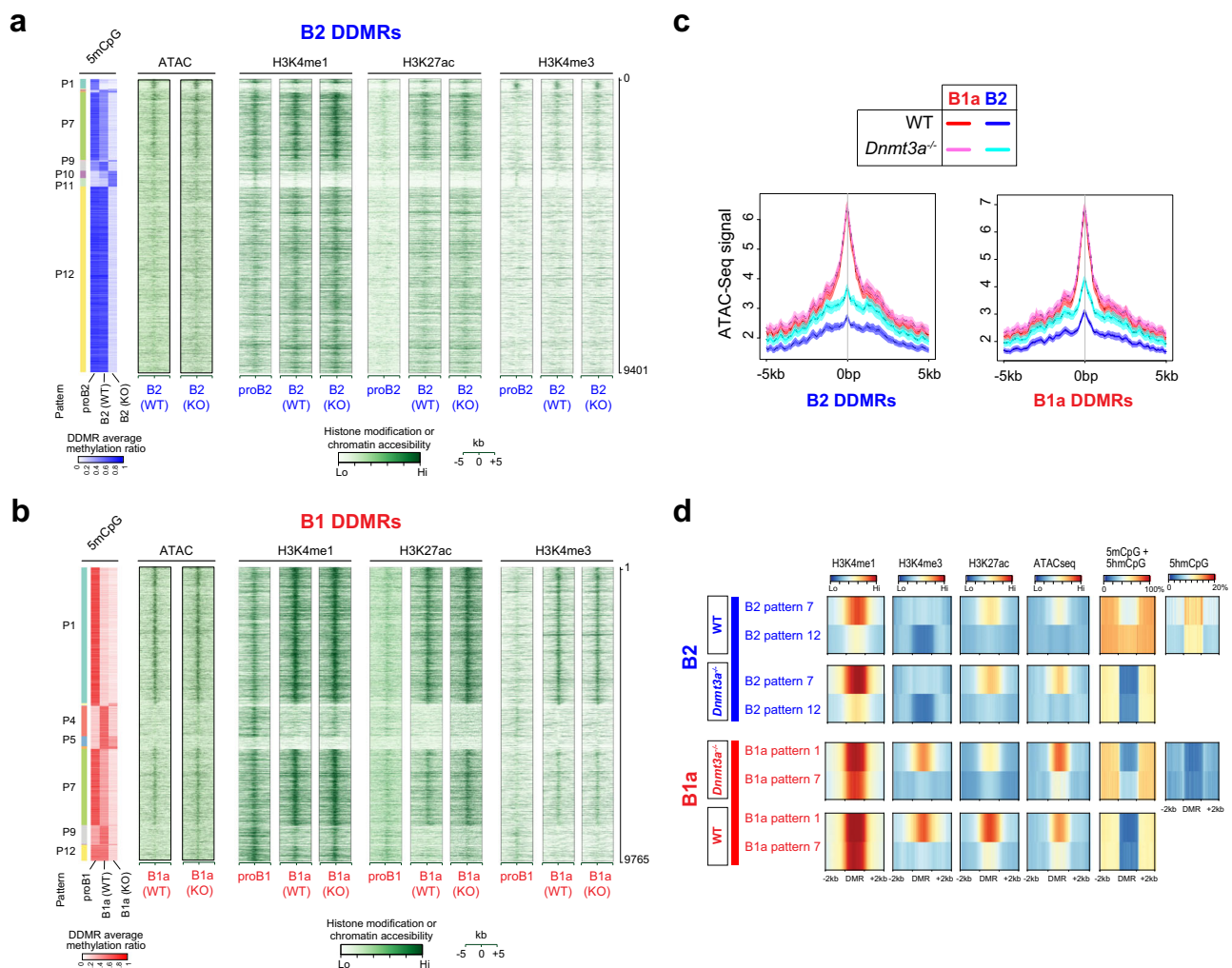


upregulated in *Dnmt3a*-deficient B cells were selectively enriched in the vicinity of DMEs in both B1a and B2 cells (Fig. 7c).

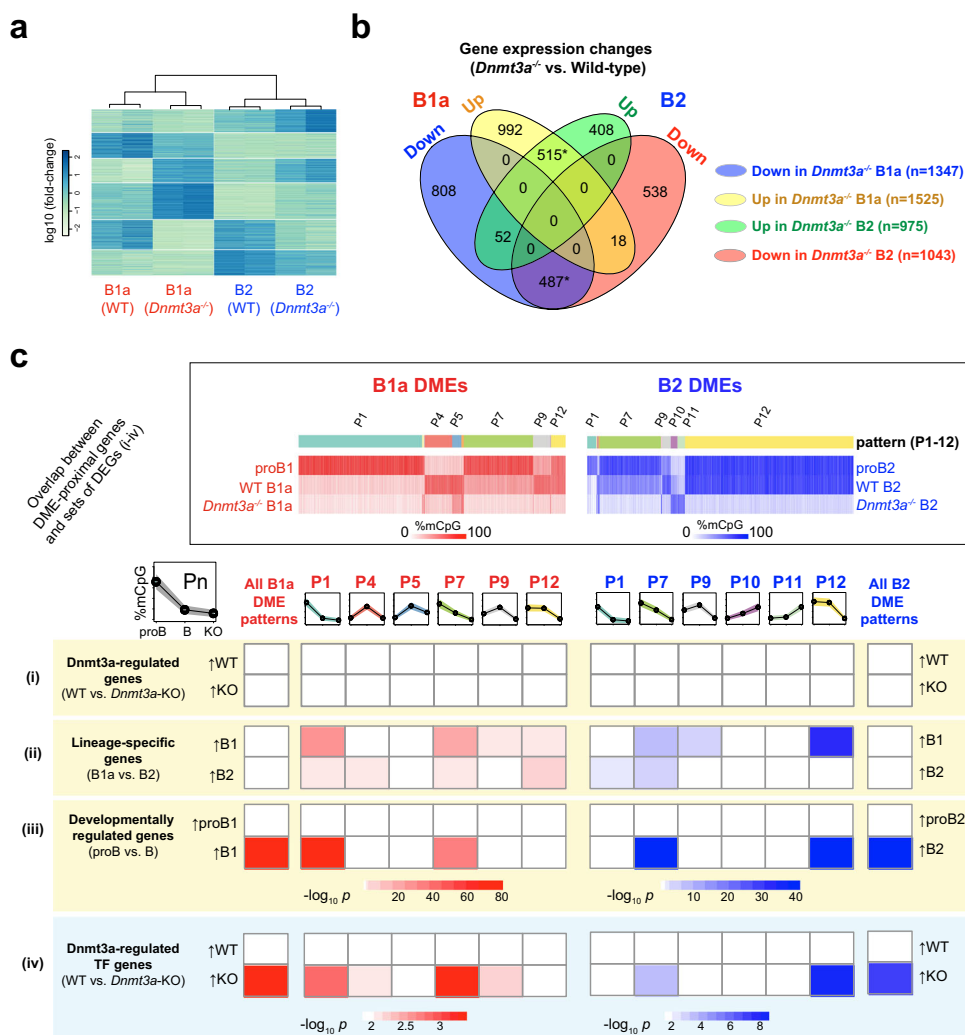
During development, lineage-specific enhancers arise at sites of silent chromatin progressing through “primed”, or “poised” states to the “active” state, each state being marked by specific combinations of chromatin and cytosine modifications<sup>44</sup>. These transitions require site-specific recruitment of chromatin and

cytosine-modifying enzymes through the action of TFs. We analyzed the enrichment of TF-binding motifs in the DMEs and interpreted the data in the context of cell-stage-specific methylation states. Indeed, DMEs with similar patterns of methylation exhibited similar TF motif enrichment, suggesting that particular sets of TFs may be responsible for specific methylation patterns (Fig. 8a). Interestingly, DDMR patterns that are associated with

**Fig. 5 DNMT3A maintains lineage-specific CpG methylation patterns in B1a and B2 cells.** **a** Overlap between B1a and B2 DDMRs. B1a and B2 cells from three wild-type or *Dnmt3a*<sup>-/-</sup> mice were pooled for the analysis. **b** Differences in the methylation profiles in the B lineage DDMRs (DNMT3A-dependent DMRs) observed across proB, mature B and *Dnmt3a*<sup>-/-</sup> B cells in B2 and B1a cell lineages, as well as wild-type and *Dnmt3a*<sup>-/-</sup> HSCs are depicted as a PCA plot (top) and as a graph of the number of the pairwise DMRs (bottom). The %CpG methylation in the DDMRs for each condition is depicted using a color scale. **c** Number of canyons (HMR length >3.5 kb) observed in proB, wild-type mature B and *Dnmt3a*<sup>-/-</sup> B cells in the B2 and B1a lineages (*n* = 1). **d** Degree of overlap between the B lineage DDMRs and HSC DDMRs (wild-type vs. *Dnmt3a*<sup>-/-</sup> HSCs). **e** Differences between the CpG methylation profiles in the B lineage DDMRs across the eight conditions depicted in **b** are illustrated as a hierarchically clustered heatmap. **f** DNMT3A-dependent differentially methylated regions (DDMRs) in the B2 (*n* = 9,401) and B1a (*n* = 9,765) B-cell lineage. Heatmaps depict the average methylation ratio within the B1a and B2 DDMRs as well as the %CpG modification in 100 bp bins depicted ±2 kb from the center of the DDMRs are shown for wild-type proB2, B2, proB1, B1a, as well as *Dnmt3a*<sup>-/-</sup> B1a and B2 cells. The CpG methylation state of B2 DDMRs is shown in shades of blue and B1a DDMRs is shown in shades of red. The six main patterns of methylation observed for B1a and B2 cells are labeled (P1-P12). The CpG methylation profiles in the prominent patterns in B1 and B2 DDMRs are plotted in Fig Supplementary Fig. 5c. **g** and **h** Average 5hmCpG modification level (% of total CpG) in B2 DDMRs in B2 cells (**g**) and B1a DDMRs in B1a cells (**h**). The DDMRs are scaled to the same length and depicted alongside 5 kb flanking regions. **i** Genomic features of DDMRs in B2 and B1a cells. **j** Distribution of distances to the nearest TSS for B1a and B2 DDMRs. Source data are provided as a Source Data file.



**Fig. 6 DDMRs exhibit enhancer marks that are modulated by CpG modification patterns.** **a** and **b** 5mCpG, H3K4me1, H3K27Ac, H3K4me3 and ATAC-Seq signal in B2 and B1a DDMRs in proB, WT B and *Dnmt3a*<sup>-/-</sup> B cells in the B1a and B2 lineages, respectively. An aggregate signal derived from 3 to 4 ATAC-Seq or CUT&RUN replicates per condition is plotted. The average CpG methylation ratio within the B1a and B2 DDMRs is reproduced from Fig. 5E to provide a frame of reference alongside a heatmap of histone modifications, as measured by CUT&RUN or chromatin accessibility signals in 100 bp bins in ±5 kb regions from the center of the DMRs. **c** Profiles of chromatin accessibility in B1a and B2 DDMRs as measured by ATAC-seq. **d** The average profile of modified CpGs, H3K4me1, H3K4me3, H3K27ac, and chromatin accessibility in the two largest patterns of B1a and B2 DDMRs in WT and *Dnmt3a*<sup>-/-</sup> B1a and B2 cells, respectively.

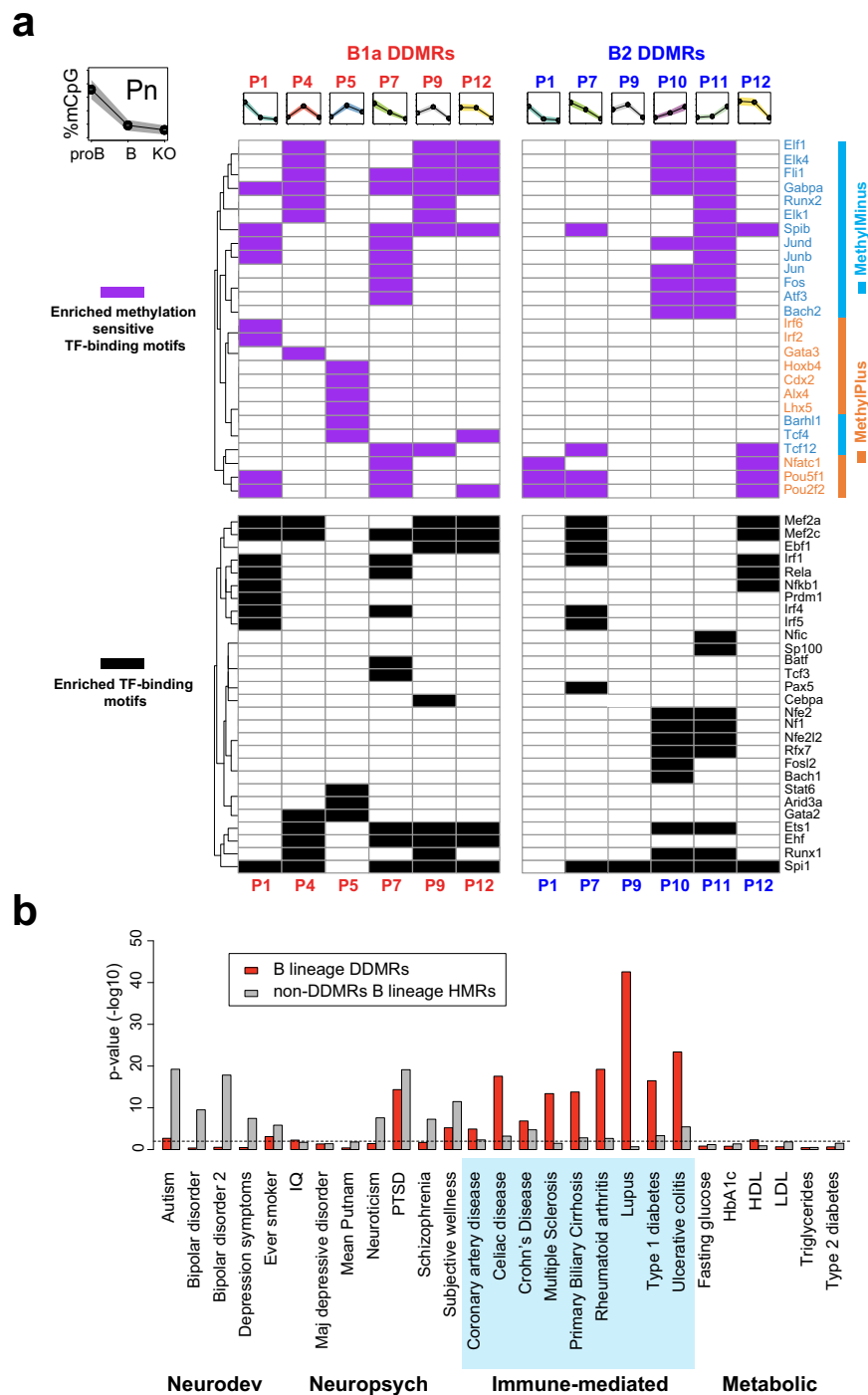


**Fig. 7 B-lineage DMEs are maintained in distinct chromatin states in B2 and B1a cells.** **a** Heatmap of differentially expressed genes between wild-type and *Dnmt3a*<sup>-/-</sup> B1a and B2 cells (two replicates per condition). **b** Genes differentially expressed between wild-type and *Dnmt3a*<sup>-/-</sup> B1a and B2 cells are shown as a Venn diagram. Statistically significant overlaps ( $p < 2.2e-16$ , Fisher’s exact test) are marked with an asterisk. **c** Significance of overlap between genes in the proximity of B1a and B2 DMEs exhibiting distinct patterns of methylation and the following differentially expressed sets of genes (i) all genes whose expression is altered by the loss of *Dnmt3a* in B1a and B2 cells, or (ii) B2 and B1a lineage-specific genes, or (iii) genes whose expression is altered during development (from proB to mature B), or (iv) genes encoding transcription factors whose expression is altered by the loss of *Dnmt3a*. The significance of the overlap was assessed using two-sided Fisher’s exact test and significantly overlapping pairs of gene sets ( $p < 0.001$ ) are colored in pink (B1) or blue (B2) based on the  $p$ -value. Source data are provided as a Source Data file.

changes in gene expression (Fig. 7c) do not share any enriched TF motifs with DDMRs that are not associated with changes in gene expression. In any given B-cell developmental stage, the CpG methylation state of DMEs is likely determined by TF-mediated site-specific recruitment of DNMT3A or TET enzymes<sup>47–49</sup>. Binding motifs for TFs that have been shown to recruit to DNMT3A such as SPI1 (PU.1), FOS, JUN, and ETS1<sup>47</sup>, as well as motifs for TET2-interacting TFs such as EBF1, BATF, TCF3 (E2A), SPI1 and EBF1, were both strongly enriched in the DMEs<sup>48,49</sup>. The transcriptional activator domains of the listed TFs likely determine the enhancer function of the DMEs, and it is noteworthy that the enriched TF-binding motifs include both well-known master regulators of the B-lineage, as well as transcriptional activators in response to BCR signaling.

Interestingly, TFs not only recruit writers and erasers of cytosine modifications, but some can also function as readers of CpG methylation as their binding can be influenced by the degree of CpG methylation at their recognition sites. The impact of cytosine methylation on TF binding has been experimentally

determined in vitro for over 500 TF’s<sup>50</sup>. About half of the TF binding motifs enriched in DMEs had been reported to be sensitive to CpG methylation in the Yin et al. study<sup>50</sup>. Interestingly, the enriched TF-binding motifs for TFs with vital roles in B lineage specification, such as EBF1, TCF3 (E2A), MEF2A, MEF2C, SP1 (PU.1), PAX5, and NFKB1 were either not sensitive to methylation or lacked a CpG site<sup>51</sup>. They were enriched in both methylated and unmethylated intervals (Fig. 8a). ARID3A, which lacks CpGs in its binding motif and can selectively drive B1a lineage development, is selectively enriched in B1a DMEs in the pattern 5 intervals that are methylated during B1a development<sup>35</sup>. In contrast, TFs that are activated in response to BCR signaling such as FOS, JUN, and NFATC1 are sensitive to CpG methylation in their binding sites. The binding of FOS and JUN to their targets is decreased by CpG methylation and the binding of NFATC1 is increased by CpG methylation. FOS and JUN motifs are enriched in hypomethylated intervals in both B1a and B2 cells. Indeed, the developmental demethylation of DMEs in the B1a lineage may render a large number of DMEs



**Fig. 8 DMEs are enriched for TF-binding motifs and SNPs associated with immune-mediated diseases.** **a** Significantly enriched TF-binding motifs in B1a and B2 DDMRs with varying patterns of methylation. Enriched motifs were filtered by  $p < 0.01$  (binomial test run with HOMER) and a difference in motif instances between test and background sets of intervals greater than 5%; the entire set of DDMRs was used as the background. Motifs that are insensitive to CpG methylation or lack a CpG are depicted in black. Methylation-sensitive transcription factor binding motifs are marked in magenta and colored based on whether the colored by TF binding is enhanced (orange) or inhibited (cyan) by CpG methylation. **b** Enrichment of SNPs linked to human disease in GWAS in B lineage DMRs (red) and all HMRs excluding the DMRs (gray) calculated using RiVIERA. Source data are provided as a Source Data file.

responsive to FOS and JUN, which are downstream of BCR signaling.

These B lineage DMEs exhibit elevated sequence conservation across 60 vertebrate species relative to their flanking regions (Supplementary Fig. 5b). SNPs within the corresponding intervals in the human genome have been linked to human disease in genome-wide association studies (GWAS) studies<sup>52</sup>. In particular,

the DME syntenic regions were strongly enriched for SNPs associated with multiple human immune-mediated disorders (Fig. 8b)<sup>53</sup>, but a similar enrichment was not seen for most neurodevelopmental, neuropsychiatric or metabolic disorders. Common variants implicated by GWAS are known to be enriched for regulatory variants and it is indeed possible that the disease-associated SNPs in these DME syntenic regions

directly contribute to disease susceptibility by impacting their enhancer function<sup>54</sup>. From a functional standpoint, the DMEs identified in this study can be considered to be a conserved set of methylation-sensitive enhancers in the B lineage, whose enhancer function is influenced and maintained by DNMT3A-mediated methylation.

## Discussion

Our study has uncovered the underlying foundational methylome that is shared among B cells and maintained, likely by DNMT1, in both B1a and B2 sub-lineages independent of DNMT3A. We also describe a dynamic methylome largely composed of enhancers that are maintained by DNMT3A in a B1a and B2 lineage-specific manner. DNMT3A is required to counteract ongoing TET activity at these enhancers. We therefore refer to these enhancers as DNMT3A-maintained enhancers or DMEs. DMEs undergo widespread programmed CpG demethylation during B1a cell development, which coincides with the induction of the expression of lineage-specific genes in the vicinity of these enhancers. These very same enhancers remain methylated by DNMT3A in B2 cells. However, DMEs are a site of ongoing TET activity in B2 cells, as evidenced by 5hmCpG marks, and are completely demethylated in the absence of DNMT3A. Our results suggest that DNMT3A can function in a maintenance role at DME sites exhibiting ongoing cycles of TET-dependent demethylation and DNMT3A-dependent maintenance re-methylation (Supplementary Fig. 8). We also show that the developmentally demethylated DMEs acquire H3K4me3 modifications in B1a cells. Recent studies analyzing the kinetics cytosine modifications in early passage ES cells following the loss of DNMT3A showed that somatic enhancers in ES cells exhibit signs of coincident DNMT3A and TET activities<sup>55,56</sup>. Given the unique biology and chromatin features at these enhancers, and their role in and sub-lineage determination, we propose that DMEs may represent a specialized category of enhancers and their role in other cell types remains to be explored.

Although the underlying foundational CpG methylomes that are unmasked in *Dnmt3a*-deficient B1a and B2 cells are virtually identical in the two sub-lineages, the cells do not lose their identity. Indeed, differences in other epigenetic marks including the ectopic H3K4me3 marks in DMEs are also preserved despite the loss of DNMT3A. This suggests that other epigenetic features including the H3K4me3 marks in DMEs are likely to be important in maintaining B1a or B2 sub-lineage identity. H3K4me3 marks are overwhelmingly associated with transcription, and their presence at DMEs may indicate the presence of enhancer RNAs. Different levels of H3K4me1 and H3K4me3 may determine whether unidirectional or bidirectional transcription occurs at enhancer sites<sup>57,58</sup>. In B cells, transcription of novel eRNAs has been observed in the context of changes in cytosine modification at CTCF anchors<sup>59</sup>. However, the B1 and B2 DMEs that we identified were not enriched in the CTCF motif, and only 10% overlapped with the set of 4,516 polyadenylated long non-coding RNAs identified in a recent study of eight B-cell types in mice<sup>28</sup>. We did not find evidence of transcripts specifically mapping to DMEs bearing H3K4me3 marks in our transcriptomic data. Given that we enriched polyadenylated RNAs for sequencing, non-polyadenylated enhancer RNAs may have been missed and further studies will be required to assess transcription initiation at DMEs. The role of eRNAs in regulating gene expression is not yet fully understood and the relationship between DME methylation, eRNAs transcription and gene expression will be explored in future studies. However, the loss of DME methylation in *Dnmt3a*-deficient cells is not without consequence and results in

a perturbed transcriptional program manifesting as a preleukemic state in B1a cells.

In B1a cells, DNMT3A activity at DMEs may be opposed, at least in part, by the acquisition of H3K4me3 marks. While some studies have described the acquisition of H3K4me3 marks at enhancers in CD8<sup>+</sup> T cells and in a tumor context, these previous studies did not evaluate them in the context of cytosine modifications<sup>29,30</sup>. The well-characterized mutually inhibitory relationship between CpG methylation and H3K4me3 in mammalian cells may underlie the biochemical basis of the switch in CpG methylation and chromatin modification states of DMEs between B1a and B2 cells. In mouse embryonic stem cells, the CFP1 subunit of the SET1 H3K4 trimethylase complex, which binds unmodified CpGs, constrains H3K4 trimethylase activity to hypomethylated promoters. In the absence of CFP1, ectopic H3K4me3 can accumulate in otherwise methylated enhancers that interact with promoters, resulting in the increased expression of their target genes<sup>31</sup>. Conversely, H3K4me3, unlike H3K4me0/me1, is unable to free the catalytic site of DNMT3A from its autoinhibitory ATRX-DNMT3-DNMT3L (ADD) domain, causing local inhibition of DNMT activity in its vicinity<sup>60,61</sup>. Therefore, we hypothesize that the established recruitment of the CFP1 subunit of the SET1 complex, the primary H3K4 trimethylase, to unmethylated DNA at promoters, also applies in part to unmethylated DMEs<sup>31,62</sup>. Once the H3K4me3 mark is established in the DMEs perhaps in response to a transient opportunity presented by TET-mediated demethylation, H3K4me3 marks may also subsequently prevent further CpG methylation by DNMT3A. Indeed, 5hmCpG is enriched at the boundaries of the DMEs in B1a cells (Fig. 5h). As 5hmCpG is not maintained during replication, the baseline self-renewal of B1 cells in the context of reduced DNMT3A may also contribute to the complete demethylation at the DMEs. Furthermore, mice lacking TET2/3 in the B lineage completely lack B1a and marginal zone B cells<sup>23</sup>. It is likely that TET enzyme mediated demethylation is necessary for the H3K4 trimethylase activity at B1a DMEs.

The importance of DMEs in the B lineage is further highlighted by the fact that they are highly enriched for binding sites of B lineage-determining TFs. The striking enrichment of SNPs linked to human immune-mediated disorders within the DME syntenic regions further highlights the importance of these DMEs in maintaining cellular function, and we speculate that DMEs are likely to be of broader biological significance. Future studies using tissue-specific deletion of DNMT3A in various tissues are likely to reveal key elements of DME-mediated gene regulation in various cell types. In studies of organismal development and cell differentiation, enhancers have been identified as poised, primed, or active, depending on their histone modifications and the functional states of genes in their vicinity<sup>44</sup>. We propose that a subset of enhancers are DMEs and likely also undergo a similar progression from a poised or primed to an active state. However, unlike conventional enhancers, DMEs continue to exhibit localized TET activity even in the active state, which is balanced by DNMT3A. 5hmC is known to be enriched in enhancers in ES cells and is considered to mark enhancer priming in models of ES cell differentiation<sup>63,64</sup>. TET2 has previously been shown to drive the generation of 5hmC at enhancers in proB cells<sup>49</sup>.

Maintenance of the appropriate level of each CpG modification at DMEs may require persistent localization of DNMT3A and TET2 at these sites through interactions with chromatin and transcription factors, as observed in epidermal stem cells<sup>65</sup>. CpG methylation of the majority of B-lineage DMEs is maintained by DNMT3A in B cells, but not in HSCs, suggesting that the dynamic methylome is specific to a particular cell type or lineage. Changes in the transcription factor activities that recruit

DNMT3A and TET enzymes to the DMEs may underlie the loss of CpG methylation seen in B1a cells. Given that PU.1 binds DNMT3A<sup>47</sup> and its binding motif is enriched in DMEs, it is likely to be one of the transcription factors that can recruit DNMT3A to DMEs. Interestingly, the loss of PU.1 (SPI1) in B cells has been shown to result in a switch from a B2 to a B1a cell fate, and B1a cells exhibit a lower level of PU.1<sup>66</sup>. The acquisition of H3K4me3 marks at DMEs in B1a cells may also be a consequence of constitutive BCR signaling, which was recently demonstrated to be sufficient to reprogram B2 cells into a B1a cell fate<sup>25</sup>. Indeed, the H3K4me3 bearing B1a DMEs are enriched in FOS, JUN and NFATC1 motifs, supporting the notion that BCR signaling may play a critical role in the acquisition of ectopic H3K4me3 in the DMEs. On the contrary, DNMT3A loss alone is not sufficient for H3K4me3 acquisition in the DMEs in B2 cells or in the developmentally methylated and decommissioned enhancers in B1a cells, suggesting that the acquisition of H3K4me3 requires a B1-specific factor, likely constitutive BCR signaling or a B1a cell-specific combination of lineage-determining factors such as LIN28, ARID3A, BHLHE41 or ZBTB32.

The local balance between DNMT3A and TET2 has also been suggested to maintain hypomethylation canyon boundaries in murine HSCs. The loss of DNMT3A in HSCs results in the expansion of canyons, as well as increased self-renewal<sup>67</sup>. B1a cell development is also accompanied by a genome-wide increase in the size of HMRs and in the number of canyons observed during B1a cell development, suggesting that the reduction in DNMT3A activity during B1a cell differentiation may not be restricted to DMEs alone but may involve a global reduction in DNMT3A activity. During development of the B1 lineage, B1a cells spontaneously acquire a CpG methylation profile that closely resembles that seen in *Dnmt3a*-deficient B1a and B2 cells. It has been proposed that HSCs modulate their self-renewal capacity by suppressing the level of DNMT3A via miR-29, and the increase in DNMT3A caused by the loss of miR-29a in HSCs is associated with a reduction in HSC self-renewal<sup>68</sup>. We speculate that analogous mechanisms may control the levels of DNMT3A in B1a cells, and consequently their self-renewal capacity. Indeed, the direct or indirect loss of DNMT3A in the B lineage has been shown to result in the leukemic transformation of B1a cells<sup>20–22</sup>. The reduced level of CpG methylation at DMEs in B1a cells may be sufficient to facilitate self-renewal; complete demethylation in the absence of DNMT3A in the knockout context may lead to uncontrolled self-renewal.

Interestingly, two recent studies describe evidence for coordinated action of DNMT3A and TET enzymes at somatic enhancers in ES cells that are highly enriched for transcription factor binding sites<sup>55,56</sup>. Our study suggests that a similar pattern of regulation may also operate at enhancers in fully differentiated somatic cells such as B cells. Having comprehensively analyzed the CpG methylome in an untransformed differentiated hematopoietic cell lineage, our findings may be applicable to a broader range of somatic cells. This study also illustrates how DNMT3A functions as a maintenance methyltransferase across a broad swath of B lineage-specific enhancers in a manner that is distinct from the DNA replication-coupled maintenance methyltransferase function of DNMT1. Our results also provide additional insights into how perturbation of the dynamic methylome promotes the preleukemic phase of IgHV-unmutated chronic lymphocytic leukemia, a leukemia suspected to be of B1a cell origin.

As shown here in B cells, we surmise that large subsets of enhancers, modulated by DNMT3A are likely present in many differentiated cell types and the contribution of DMEs in the dynamic methylome to the differentiation and biology of other lineages will undoubtedly be very revealing. DMEs may provide an additional mode of developmental control of enhancer

function, wherein programmed CpG demethylation and acquisition of H3K4me3 marks facilitate additional modulation of enhancer activity, particularly in the context of lineage-specific gene expression. Our work also suggests how the cellular phenotypic consequences of DNMT3A-mediated epigenetic modifications can be constrained to a specific developmental lineage and how the loss of such epigenetic regulatory mechanisms may contribute to a cell type-specific malignancy.

## Methods

**Mice.** Generation of knockout mice bearing the conditional *Dnmt3a* allele, in which the exon encoding the catalytic domain of DNMT3A (exon 19) is floxed (*Dnmt3a<sup>tm3.1Enl</sup>*, originally referred to as *Dnmt3a<sup>2lox</sup>*) has been described<sup>39</sup>. The mice were a gift from Drs. Taiping Chen and En Li and are presently available from the RIKEN BioResource Center (BRC strain: RBRC03731). *Dnmt3a<sup>2lox/2lox</sup>* mice were backcrossed into the C57BL/6J background (Jax strain:000664) for over ten generations. They were crossed into *Cd19-Cre* mice (Jax strain:006785) to obtain *Cd19-Cre<sup>+/+</sup> Dnmt3a<sup>2lox/2lox</sup>* mice with a conditional deletion of *Dnmt3a* in B cells<sup>40</sup>. These mice are referred to as *Dnmt3a<sup>-/-</sup>* mice in the manuscript for simplicity. C57BL/6J mice were used as controls. All animals were housed at the Massachusetts General Hospital under specific pathogen-free conditions, with the ambient temperature set-point at 70 °F, relative humidity between 30–70% and a 12 h light–dark cycle. All rodent experimental procedures were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC).

**Flow cytometry and isolation of B-cell populations.** Single-cell suspensions from spleens, bone marrow, and peritoneal washes were analyzed using flow cytometry. The clones, suppliers and catalog numbers of the antibodies used in this study are listed in the Reporting Summary. All antibodies were used at a dilution of 1:100, except for anti-mouse CD19, B220 and CD3, which were used at a dilution of 1:200. Cells were sorted based on the gating strategy employed by the ImmGen consortium (Supplementary Fig. 9). Peritoneal B1a B cells were identified by the following surface phenotype: CD19<sup>+</sup> B220<sup>lo</sup> IgM<sup>++</sup> CD23<sup>-</sup> CD43<sup>+</sup> CD5<sup>+</sup>. Splenic B2 (follicular) B cells were identified by the following surface phenotype: CD19<sup>+</sup> B220<sup>+</sup> IgM<sup>+</sup> IgD<sup>+</sup> AA4.1<sup>-</sup> CD23<sup>+</sup> CD43<sup>-</sup> CD5<sup>-</sup>. For purification of proB1 cells (Fraction BC from E16 fetal liver), fetal liver cells from a total of 14 male and female E16 mouse fetuses were pooled and stained with a cocktail of PE-Cy7-conjugated antibodies against Ly6c, Ter119, CD11b and CD3. Lineage-negative cells were enriched using anti-Cy7 magnetic beads (Miltenyi Biotec). The enriched lineage-negative cell fraction was then stained with a cocktail of antibodies following the gating strategy used by the ImmGen Consortium (Supplementary Fig. 9). Similarly, proB2 cells (Fr-BC) were FACS sorted from the bone marrow of 4-week-old mice using the gating strategy used by the ImmGen Consortium (Ly6c<sup>-</sup> Ter119<sup>-</sup> Cd11b<sup>-</sup> CD3<sup>-</sup> CD19<sup>+</sup> IgM<sup>lo</sup> CD93<sup>+</sup> B220<sup>lo</sup> CD43<sup>+</sup> CD24<sup>lo</sup>).

Single-cell suspensions were made from spleens, bone marrow, peritoneal washes or fetal liver. RBCs were lysed with 2 mL ACK lysis buffer (Lonza). Prior to surface staining,  $1 \times 10^6$  cells were labeled with the LIVE/DEAD Fixable Blue Dead Cell Stain kit (Thermo Scientific) at 1:1000 dilution in 1 mL of HBSS for 30 min at room temperature. Fc-receptors were blocked with 2.5  $\mu$ g of clone 2.4G2 (Biolegend) for 20 min on ice. Surface staining was performed using appropriate dilutions of antibodies in 100  $\mu$ L of cell staining buffer (Biolegend, CA) for 30 min in the dark at 4 °C or on ice. Flow cytometry and FACS sorting was performed on a 4-laser BD LSR-II, a 5-laser BD Fortessa, or a BD FACS Aria (BD Biosciences), as appropriate. For RNA-Seq studies, 100,000 live cells from each subset were directly sorted into RLT plus buffer (Qiagen) containing 1% 2-mercaptoethanol (Sigma) and stored at –80 °C until the RNA extraction. For ATAC-Seq, WGBS, and TAB-Seq, 200,000 live cells from each subset were sorted into media containing 1% BSA, spun at 500  $\times$ g at 4 °C for 10 min. Cells were then processed immediately for ATAC-Seq library preparation or resuspended in DNA/RNA shield (Zymo Research) and stored at –80 °C prior to DNA extraction using the QiaAmp DNA Mini kit for WGBS or TAB-Seq.

**RNA-Seq library preparation.** CD19<sup>+</sup> CD23<sup>-</sup> IgM<sup>+</sup> CD5<sup>+</sup> B1a cells (35,000–100,000 cells) were FACS sorted from the peritoneal washes of two *Dnmt3a<sup>-/-</sup>* mice and two littermate controls at 4 weeks of age; CD19<sup>+</sup> IgD<sup>hi</sup> IgM<sup>lo</sup> CD21<sup>+</sup> follicular B cells (100,000 cells per sample) were sorted from spleens of two *Dnmt3a<sup>-/-</sup>* mice and two littermate controls at 8 weeks of age. RNA was isolated from the sorted cells using an RNeasy-plus Micro kit (Qiagen). RNA-Seq libraries were prepared using the KAPA Stranded RNA-Seq Library Preparation kit (KAPA Biosystems). Libraries were assessed for quality using High Sensitivity DNA chips on the Agilent Bioanalyzer, quantified using Qubit fluorometer (Thermo Fisher Scientific), as well as KAPA Library Quantification kit (KAPA Biosystems), and sequenced on the Illumina NextSeq 550 platform.

**WGBS library preparation.** DNA was isolated from peritoneal cavity B1a cells, splenic follicular B cells (*Dnmt3a<sup>-/-</sup>* vs. WT), fetal liver proB1 cells, and bone marrow proB2 cells. WGBS libraries were prepared using 100 ng of DNA using the

Pico Methyl-Seq™ Library Prep Kit (Zymo Research) according to the manufacturer's protocol. The quality of the libraries was assessed using High Sensitivity D5000 Screen Tapes (Agilent 4200 TapeStation) and quantified with KAPA Library Quantification kit. Libraries were sequenced on an Illumina NextSeq 550. Fewer CpGs were modified in proB1 and B1a cells (74.1% and 73.1%, respectively) than in the proB2 and B2 cells (75.9% and 76%, respectively) (Supplementary Table 2). Since CHH and CHG methylation (where H represents A, T, or C) represented <2% of all cytosine methylation in proB and B cells (Supplementary Table 2), we focused our analyses on CpG modifications.

**TAB-Seq library preparation.** TAB-Seq libraries were constructed using the TAB-Seq library preparation kit from Wisegene (catalog #K001)<sup>69</sup>. Briefly, 5-hydroxymethyl cytosines in 500 ng genomic DNA with 0.1% spike-in control DNA (synthesized with 5-hydroxymethyl cytosines or 5-methyl cytosines exclusively) were protected using T4 β-glucosyltransferase, and 5-methyl cytosines were oxidized using recombinant TET1. Following TET1 oxidation, the DNA was subjected to bisulfite conversion and PCR-based sequencing adapter tagging as described for WGBS above (Pico Methyl-Seq Library Prep Kit, Zymo Research). The library was quantified and sequenced on an Illumina Nextseq 550. Genome-wide, 2.6% of CpGs were hydroxymethylated in B1a cells and 3.8% of CpGs were hydroxymethylated in B2 cells (Supplementary Table 2).

**ATAC-Seq library preparation.** ATAC-Seq libraries were prepared as originally described<sup>43</sup>. Fifty thousand freshly sorted cells were pelleted and washed with 50 μL chilled 1× PBS, and with 50 μL lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.1% IGEPAL CA-630). Three replicates were analyzed per condition. The nuclei were pelleted in a cold micro-centrifuge at 550 × g for 10 min and resuspended in a 50 μL transposition reaction with 2.5 μL Tn5 transposase (FC-121-1030; Illumina) to tag open chromatin. The reaction was incubated at 37 °C in a Thermomixer (Eppendorf) at 300 rpm for 30 min. Tagmented DNA was purified using a QIAGEN MinElute kit and amplified with 7 or 11 cycles of PCR, based on the results of a test qPCR. ATAC-Seq libraries were then purified using a QIAGEN PCR cleanup kit and quantified using KAPA library quantification kit (KAPA Biosystems, Roche) and sequenced on the Nextseq 550 platform.

**CUT&RUN library preparation.** CUT&RUN libraries were prepared using rabbit monoclonal antibodies against H3K4me1, H3K4me3 and H3K27Ac (Cell Signaling Technology catalog numbers 5326, 9751, and 8173, respectively) as previously described<sup>42</sup>. The protein-A Mnase fusion protein from a gift from Steven Henikoff.

**Preparation and analysis of Ig Heavy Chain-Repertoire libraries.** RNA from B1a cells was reverse transcribed and amplified using a Qiagen OneStep RT-PCR kit (Qiagen Inc., Valencia, CA, USA) and iRepertoire® mouse BCR heavy chain (MBH) primers (iRepertoire Inc., Huntsville, AL, USA), following the iRepertoire user manual for the Illumina sequencing platform. Following manufacturer's instructions in the iRepertoire manual, Illumina adapters were added using a Qiagen Multiplex PCR kit, and a 350–500 bp bands corresponding to the Ig heavy chain amplicons were gel purified using a Qiagen Gel Extraction kit. The libraries were quantified using Qubit dsDNA quant (Life Technologies) and sequenced on the Illumina MiSeq platform (paired end, 250 bp). For samples from aged mice bearing clonal leukemias, the libraries were Sanger sequenced using the Illumina universal primers as the sequencing primer. Repertoire data was analyzed using the *iRweb* online analysis platform provided by iRepertoire Inc.

**Analysis of WGBS and TAB-Seq data.** The first 5 and last 2 bases were trimmed using *cutadapt v1.11* and the trimmed reads were aligned to the *mm10* mouse reference genome using *Bismark v0.14.370* with the parameters: `--bowtie2 --non_directional --un --ambiguous --phred33-quals`. PCR duplicates were removed using *Picard (v2.5.0)*. Approximately 5× coverage was obtained after removal of duplicates (Supplementary Table 2). The *Bismark* package was used to extract the coverage and methylation status of individual cytosines. We then used the *MethPipe v3.4.326* to identify differentially methylated regions. We implemented *MethPipe* in the following steps. First, we used the program *to-mr* to convert the output BAM file from *Bismark* to *mr* format. Second, we removed duplicate reads using *duplicate-remover* to get rid of the reads resulting from PCR over-amplification before calculating the methylation level. Third, the bisulfite conversion rate, defined as the rate at which unmethylated cytosines in the sample appear as thymidines in the sequenced reads, was measured by the program *bsrate* from *MethPipe* based on the non-CpG cytosines, which are believed to be almost completely unmethylated in most mammalian cells. In all, 98–99% bisulfite conversion efficiency was observed. Finally, the *methcounts* program was used to estimate the methylation level for individual cytosines, calculated as a probability based on the ratio of methylated to total mapped reads at the respective cytosines. Symmetric CpGs were merged using *symmetric-cpgs*. The *bedgraph* output from *Bismark* was converted into the *bigwig* format for visualization on *IGV* and for plotting heatmaps using the *SeqPlots* package<sup>71,72</sup>. WGBS data for HSC (GSE49714) and CD8<sup>+</sup> T cells (GSE107150) were additionally downloaded and

reprocessed using the same methods and parameters as used in this study<sup>27,73</sup>. A similar approach was used for the analysis of TAB-Seq data. Analysis of TAB-Seq spike-in controls showed an efficiency of ~90% for TET oxidation and ~98% for bisulfite conversion.

**Detection of HMRs (hypomethylated regions) and canyons.** HMRs were called using the program *hmr* in the *MethPipe* package, which utilizes a hidden Markov model (HMM)-based approach. A length cutoff of 3.5 kb or longer was used to identify canyons, as previously described<sup>27</sup>. Similarity in the HMRs among different cell types was assessed based on the Jaccard statistic using the *BEDtools* suite<sup>74,75</sup>. The Jaccard statistic between the two sets of genomic intervals ranges from 0 to 1 (no overlap to complete overlap, respectively).

**Identification and annotation of DMRs.** To compare two methylomes, we identified differentially methylated regions (DMRs) using the programs *methdiff* and *dmr* in the *MethPipe* package. The first step was to calculate the differential methylation score for each CpG site using *methdiff*. Then, the *methdiff* and *hmr* output files were used to compute DMRs using the program *dmr*. Finally, we obtained a filtered set of DMRs spanning at least 10 CpGs and having at least 5 significantly differentially methylated CpG as recommended by the *MethPipe* user manual. In addition, the *roimethstat* program was used to estimate the average methylation level in each DMR across all samples of interest.

To annotate the genomic features of DMRs (promoter, intergenic, intron, exon, 3' UTR, 5' UTR, TTS) and identify DMR-proximal genes, we used the *annotatePeaks.pl* script from *HOMER (v4.9)*<sup>76</sup>. We also used the *HOMER* scripts *findGO.pl* to perform functional enrichment tests. The *findMotifsGenome.pl* script was used to identify enriched vertebrate transcription factor binding motifs from JASPAR and *HOMER* databases in the DMR sets of interest (filtered using a *p*-value cutoff <0.01 and a difference between test sets and background intervals >5%; the full set of DMRs was used as the background).

The DMRs were grouped into patterns based on the methylation dynamics across proB, mature B and *Dnmt3a*<sup>-/-</sup> B cells. We created a list of 12 pattern vectors of interest representing CpG methylation ratios (range [0,1]) across the three conditions (low, intermediate and high) (Supplementary Fig. 5a). In each pattern vector, “low,” “intermediate,” and “high” values were set at CpG methylation ratios of 0.2, 0.5, and 0.8, respectively. Each vector represents the CpG methylation ratios across proB, wild-type B, and *Dnmt3a*<sup>-/-</sup> B conditions for a particular B-cell subset. We assigned each B1a and B2 DMR to the pattern vector that yielded the maximum Pearson correlation coefficient (PCC). No DMR was assigned to more than one pattern.

**ChIP-Seq and ATAC-Seq analysis.** We obtained H3K4me1, H3K4me3 ChIP-Seq data in B1a and B2 cells (GSE72017) and H3K4me1 ChIP-Seq data in mouse HSCs (GSE63276) from the NCBI GEO database<sup>28,77</sup>. The ATAC-Seq libraries were generated and sequenced in our laboratory as described above<sup>43</sup>. ChIP-Seq and ATAC-Seq reads were aligned to the *mm10* mouse reference genome. We removed duplicated and mitochondrial reads, and generated *bigwig* files for plotting heatmaps using the *SeqPlots* package<sup>72</sup>. ATAC-Seq or ChIP-Seq peaks were called using *MACS2* with a fold change >5 as a cutoff<sup>78</sup>.

**Transcriptomic analysis.** We aligned the RNA-Seq reads to the *mm10* reference genome and performed transcript quantification using *RSEM v1.25.0*. Refseq gene annotations were used. We used *EBSeq*<sup>79</sup> to identify differentially expressed genes (DEGs) between two biological conditions; we used a cutoff of 0.95 for the posterior probability of differential expression reported by *EBSeq* to determine differentially expressed genes with a false discovery rate controlled at 5%. We used the *HOMER findGO.pl* script<sup>76</sup> to perform functional enrichment tests. Fisher's exact test was used to evaluate the significance of overlap between two gene sets, using all expressed genes as the background.

We also downloaded microarray data for developing and mature B1a and B2 cells (proB cells from bone marrow or fetal liver, follicular B cells from spleen, FrE from fetal liver or bone marrow, and B1a cells from peritoneal cavity) from the ImmGen consortium (GSE15907) and identified differentially expressed genes (*p* < 0.01) using the *RankProd* R package<sup>34,80,81</sup>.

**SNPs enrichment test in DMRs.** We used the probabilistic framework implemented by *RiVIERA* to infer the association of DMRs and SNPs of 27 human traits from four major disease categories (NeuroDegen, NeuroPsych, Immune, and Metabolic)<sup>52</sup>.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

All sequencing datasets generated in the current study have been deposited in the NCBI GEO repository (GSE150497). Source data are provided with this paper.

**Code availability**

All of the code used in the paper was from published studies and is cited in the manuscript. Version numbers are provided in the Reporting Summary.

Received: 13 April 2020; Accepted: 7 March 2021;

Published online: 13 April 2021

**References**

- Hardy, R. R. & Hayakawa, K. Perspectives on fetal derived CD5+ B1 B cells. *Eur. J. Immunol.* **45**, 2978–2984 (2015).
- Montecino-Rodriguez, E. et al. Distinct genetic networks orchestrate the emergence of specific waves of fetal and adult B-1 and B-2 development. *Immunity* **45**, 527–539 (2016).
- Hodges, E. et al. Directional DNA methylation changes and complex intermediate states accompany lineage specificity in the adult hematopoietic compartment. *Mol. Cell* **44**, 17–28 (2011).
- Ji, H. et al. Comprehensive methylome map of lineage commitment from haematopoietic progenitors. *Nature* **467**, 338–342 (2010).
- Deaton, A. M. & Bird, A. CpG islands and the regulation of transcription. *Genes Dev.* **25**, 1010–1022 (2011).
- Bock, C. et al. DNA methylation dynamics during in vivo differentiation of blood and skin stem cells. *Mol. Cell* **47**, 633–647 (2012).
- Kulis, M. et al. Whole-genome fingerprint of the DNA methylome during human B cell differentiation. *Nat. Genet.* **47**, 746–756 (2015).
- Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247–257 (1999).
- He, Y.-F. et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333**, 1303–1307 (2011).
- Ito, S. et al. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129–1133 (2010).
- Ito, S. et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**, 1300–1303 (2011).
- Tahiliani, M. et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–935 (2009).
- Inoue, A. & Zhang, Y. Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science* **334**, 194 (2011).
- Serandour, A. A. et al. Dynamic hydroxymethylation of deoxyribonucleic acid marks differentiation-associated enhancers. *Nucleic Acids Res.* **40**, 8255–8265 (2012).
- Johnson, K. C. et al. 5-Hydroxymethylcytosine localizes to enhancer elements and is associated with survival in glioblastoma patients. *Nat. Commun.* **7**, 13177 (2016).
- Hon, G. C. et al. 5mC oxidation by Tet2 modulates enhancer activity and timing of transcriptome reprogramming during differentiation. *Mol. Cell* **56**, 286–297 (2014).
- Challen, G. A. et al. Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat. Genet.* **44**, 23–31 (2011).
- Gao, X. N. et al. AML1/ETO cooperates with HIF1 $\alpha$  to promote leukemogenesis through DNMT3a transactivation. *Leukemia* **29**, 1730–1740 (2015).
- Mayle, A. et al. Dnmt3a loss predisposes murine hematopoietic stem cells to malignant transformation. *Blood* **125**, 629–638 (2015).
- Haney, S. L. et al. Loss of Dnmt3a induces CLL and PTCL with distinct methylomes and transcriptomes in mice. *Sci. Rep.* **6**, 34222 (2016).
- Haney, S. L. et al. Promoter hypomethylation and expression is conserved in mouse chronic lymphocytic leukemia induced by decreased or inactivated Dnmt3a. *Cell Rep.* **15**, 1190–1201 (2016).
- Palamarchuk, A. et al. Tc1 protein functions as an inhibitor of de novo DNA methylation in B-cell chronic lymphocytic leukemia (CLL). *Proc. Natl Acad. Sci. USA.* **109**, 2555–2560 (2012).
- Orlanski, S. et al. Tissue-specific DNA demethylation is required for proper B-cell differentiation and function. *Proc. Natl Acad. Sci. USA.* **113**, 5018–5023 (2016).
- Zaborsky, N. et al. Exome sequencing of the TCL1 mouse model for CLL reveals genetic heterogeneity and dynamics during disease development. *Leukemia* **33**, 957–968 (2019).
- Graf, R. et al. BCR-dependent lineage plasticity in mature B cells. *Science* **363**, 748–753 (2019).
- Song, Q. et al. A reference methylome database and analysis pipeline to facilitate integrative and comparative epigenomics. *PLoS ONE* **8**, e81148 (2013).
- Jeong, M. et al. Large conserved domains of low DNA methylation maintained by Dnmt3a. *Nat. Genet.* <https://doi.org/10.1038/ng.2836> (2013).
- Brazão, T. F. et al. Long noncoding RNAs in B-cell development and activation. *Blood* **128**, e10–e19 (2016).
- Russ, B. E. et al. Regulation of H3K4me3 at transcriptional enhancers characterizes acquisition of virus-specific CD8+ T cell-lineage-specific function. *Cell Rep.* **21**, 3624–3636 (2017).
- Cao, F. et al. Super-enhancers and broad H3K4me3 domains form complex gene regulatory circuits involving chromatin interactions. *Sci. Rep.* **7**, 2186 (2017).
- Clouaire, T. et al. Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. *Genes Dev.* **26**, 1714–1728 (2012).
- Yu, M. et al. Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell* **149**, 1368–1380 (2012).
- Gu, T. et al. DNMT3A and TET1 cooperate to regulate promoter epigenetic landscapes in mouse embryonic stem cells. *Genome Biol.* **19**, 88 (2018).
- Heng, T. S. P. & Painter, M. W., Immunological Genome Project Consortium. The Immunological Genome Project: networks of gene expression in immune cells. *Nat. Immunol.* **9**, 1091–1094 (2008).
- Zhou, Y. et al. Lin28b promotes fetal B lymphopoiesis through the transcription factor Arid3a. *J. Exp. Med.* **212**, 569–580 (2015).
- Kreslavsky, T. et al. Essential role for the transcription factor Bhlhe41 in regulating the development, self-renewal and BCR repertoire of B-1a cells. *Nat. Immunol.* **18**, 442–455 (2017).
- Jellusova, J. et al. Siglec-G regulates B1 cell survival and selection. *J. Immunol.* **185**, 3277–3284 (2010).
- Guo, B. & Rothstein, T. L. RasGRP1 is an essential signaling molecule for development of B1a cells with autoantigen receptors. *J. Immunol.* **196**, 2583–2590 (2016).
- Kaneda, M. et al. Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* **429**, 900–903 (2004).
- Rickert, R. C., Rajewsky, K. & Roes, J. Impairment of T-cell-dependent B-cell responses and B-1 cell development in CD19-deficient mice. *Nature* **376**, 352–355 (1995).
- Montecino-Rodriguez, E., Leathers, H. & Dorshkind, K. Identification of a B-1 B cell-specified progenitor. *Nat. Immunol.* **7**, 293–301 (2006).
- Skene, P. J., Henikoff, J. G. & Henikoff, S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat. Protoc.* **13**, 1006–1019 (2018).
- Buenostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).
- Calo, E. & Wysocka, J. Modification of enhancer chromatin: what, how, and why? *Mol. Cell* **49**, 825–837 (2013).
- Challen, G. A. et al. Dnmt3a and Dnmt3b have overlapping and distinct functions in hematopoietic stem cells. *Cell Stem Cell* **15**, 350–364 (2014).
- Ernst, J. & Kellis, M. ChromHMM: automating chromatin-state discovery and characterization. *Nat. Methods* **9**, 215–216 (2012).
- Suzuki, M. et al. Site-specific DNA methylation by a complex of PU.1 and Dnmt3a/b. *Oncogene* **25**, 2477–2488 (2006).
- Guilhamon, P. et al. Meta-analysis of IDH-mutant cancers identifies EBF1 as an interaction partner for TET2. *Nat. Commun.* **4**, 2166 (2013).
- Lio, C.-W. et al. Tet2 and Tet3 cooperate with B-lineage transcription factors to regulate DNA modification and chromatin accessibility. *Elife* **5**, e18290 (2016).
- Yin, Y. et al. Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science* **356**, eaaj2239 (2017).
- Lin, Y. C. et al. A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. *Nat. Immunol.* **11**, 635–643 (2010).
- Li, Y. & Kellis, M. Joint Bayesian inference of risk variants and tissue-specific epigenomic enrichments across multiple complex human diseases. *Nucleic Acids Res.* **44**, e144 (2016).
- Chen, L. & Qin, Z. S. traseR: an R package for performing trait-associated SNP enrichment analysis in genomic intervals. *Bioinformatics* **32**, 1214–1216 (2016).
- Gusev, A. et al. Partitioning heritability of regulatory and cell-type-specific variants across 11 common diseases. *Am. J. Hum. Genet.* **95**, 535–552 (2014).
- Ginno, P. A. et al. A genome-scale map of DNA methylation turnover identifies site-specific dependencies of DNMT and TET activity. *Nat. Commun.* **11**, 2680 (2020).
- Charlton, J. et al. TETs compete with DNMT3 activity in pluripotent cells at thousands of methylated somatic enhancers. *Nat. Genet.* **52**, 819–827 (2020).
- Koch, F. et al. Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. *Nat. Struct. Mol. Biol.* **18**, 956–963 (2011).
- Wang, D. et al. Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. *Nature* **474**, 390–394 (2011).
- Benner, C., Isoda, T. & Murre, C. New roles for DNA cytosine modification, eRNA, anchors, and superanchors in developing B cell progenitors. *Proc. Natl Acad. Sci. USA.* **112**, 12776–12781 (2015).



60. Zhang, Y. et al. Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Res.* **38**, 4246–4253 (2010).
61. Guo, X. et al. Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. *Nature* **517**, 640–644 (2015).
62. Brown, D. A. et al. The SET1 complex selects actively transcribed target genes via multivalent interaction with CpG island chromatin. *Cell Rep.* **20**, 2313–2327 (2017).
63. Stroud, H., Feng, S., Morey Kinney, S., Pradhan, S. & Jacobsen, S. E. 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biol.* **12**, R54 (2011).
64. Mahé, E. A. et al. Cytosine modifications modulate the chromatin architecture of transcriptional enhancers. *Genome Res.* **27**, 947–958 (2017).
65. Rinaldi, L. et al. Dnmt3a and Dnmt3b associate with enhancers to regulate human epidermal stem cell homeostasis. *Cell Stem Cell* **19**, 491–501 (2016).
66. Ye, M., Ermakova, O. & Graf, T. PU.1 is not strictly required for B cell development and its absence induces a B-2 to B-1 cell switch. *J. Exp. Med.* **202**, 1411–1422 (2005).
67. Zhang, X. et al. DNMT3A and TET2 compete and cooperate to repress lineage-specific transcription factors in hematopoietic stem cells. *Nat. Genet.* **48**, 1014–1023 (2016).
68. Hu, W. et al. miR-29a maintains mouse hematopoietic stem cell self-renewal by regulating Dnmt3a. *Blood* **125**, 2206–2216 (2015).
69. Yu, M. et al. Tet-assisted bisulfite sequencing of 5-hydroxymethylcytosine. *Nat. Protoc.* **7**, 2159–2170 (2012).
70. Krueger, F. & Andrews, S. R. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–1572 (2011).
71. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24 (2011).
72. Stempor, P. & Ahringer, J. SeqPlots - Interactive software for exploratory data analyses, pattern discovery and visualization in genomics. *Wellcome Open Res.* **1**, 14 (2016).
73. Youngblood, B. et al. Effector CD8 T cells dedifferentiate into long-lived memory cells. *Nature* **552**, 404–409 (2017).
74. Favorov, A. et al. Exploring massive, genome scale datasets with the GenometriCorr package. *PLoS Comput. Biol.* **8**, e1002529 (2012).
75. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
76. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
77. Luo, M. et al. Long non-coding RNAs control hematopoietic stem cell function. *Cell Stem Cell* **16**, 426–438 (2015).
78. Feng, J., Liu, T., Qin, B., Zhang, Y. & Liu, X. S. Identifying ChIP-seq enrichment using MACS. *Nat. Protoc.* **7**, 1728–1740 (2012).
79. Leng, N. et al. EBSeq: An empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics* **29**, 1035–1043 (2013).
80. Del Carratore, F. et al. RankProd 2.0: a refactored bioconductor package for detecting differentially expressed features in molecular profiling datasets. *Bioinformatics* **33**, 2774–2775 (2017).
81. Hong, F. et al. RankProd: a bioconductor package for detecting differentially expressed genes in meta-analysis. *Bioinformatics* **22**, 2825–2827 (2006).

## Acknowledgements

We thank Eliezer Calo for helpful discussions and Deepta Bhattacharya and Ananda Roy for comments on the manuscript. This work was supported by NIH AI110495 and Ragon Institute Strategic Funding to S.P. and support from NIH AI113163 to V.S.M.

## Author contributions

S.P., V.M., and H.M. conceptualized and supervised the study. V.M., H.M., G.Y., H.A., M.A., S.M., and Y.T. performed the experiments. V.M., N.S., and V.V. performed the computational analysis. A.C. helped establish the *Dnmt3a*-deficient mouse colony. V.M. and S.P. wrote the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-22458-9>.

**Correspondence** and requests for materials should be addressed to S.P.

**Peer review information** *Nature Communications* thanks Paolo Casali and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021