SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1. (**A**) Heatmap showing the Pearson correlation coefficient matrices of pairwise comparisons using informative windows as described in Oakes et al. 2016 (10) of T-WGBS data for two biological replicates per cell subpopulation. The genome was tiled into 500-bp windows and only the methylation averages of those with > 4 CpGs were used. (**B**) t-SNE projection of T-WGBS data obtained from naïve and memory B cells isolated from two HIGM2 patients and two healthy controls. (**C**) Heatmap showing the Pearson correlation coefficient matrices using informative windows for our control data and previously published data. Oakes refers to the Oakes et al. 2016 data (10) and ICGC refers to International Cancer Genomic Consortium methylation data. (**D**) Unsupervised t-SNE projection comparing our T-WGBS control data with published datasets. NBC: naïve B cells; ncsMBC: non-class switched memory B cells.

Supplementary Figure 2. Indirect involvement of AID in DNA demethylation dynamics during B cell activation. (**A**) Box and dot plots showing the distribution of expression values of the *AICDA* gene in naïve, memory and germinal center B cells obtained from the BLUEPRINT database (* p-value < 0.05). (**B**) Box and violin plots depicting the distribution of DNA methylation levels of AID off-target DMRs (DMRs associated with genes described as AID off-target). (**C**) Location proportions of P-AID, AID off-target DMRs in the context of CpG islands (CGI) and gene-related regions. (**D**) Bubble plot depicting the enrichment (red) and depletion (blue) of significant ChIP-seq peaks of the annotated histone marks obtained from germinal center B cells (BLUEPRINT database). Color represents the logarithmic-fold change, dot size indicates the percentage of DMRs in the histone mark peaks, and the edge indicates the statistical significance of the enrichment (black: significant, none: not significant; q < 0.01). (**E**) DNA methylation and H3K27ac profiles in the vicinity of the gene *IL21R*. DMR color indicates the type of DMR, either P-AID (green) or AID off-target DMRs (red). H3K27ac ChIP-seq signal in germinal center B cells, super-enhancer and enhancer are depicted. (**F**) Box and dot plots showing the distribution of expression of genes associated with P-AID and AID off-target DMRs. Gene expression data of germinal center B cells was obtained from the BLUEPRINT database. FPKM refers to fragments per kilobase of transcript per million mapped reads. One-sided unpaired Wilcoxon's test was used to test for differences in expression (**** p-value < 0.0001). (**G**) Box and violin plots depicting the distribution of DNA methylation of CpG sites containing WRCY hotspots at super-enhancers of AID off-target genes. Statistical tests: two-tailed Wilcoxon's (A, F, G) and Fisher's exact (D) tests.

Supplementary Figure 3. (**A**) Scheme depicting the potential dynamics of demethylation. (**B**) Scatter plot showing a pairwise comparison of DNA methylation differences between naïve B cells (NB) and germinal center B cells (GBC) of wild type (WT) and *Aicda*$^{-/-}$ mice. Blue dots indicate CpGs potentially demethylated by AID (mouse P-AID). Orange dots indicate those CpGs demethylated in both wild type and *Aicda*$^{-/-}$ mice in the transition from naïve B cells to germinal center B cells (positive control). Gray dots show those CpGs without methylation changes in this transition in either set of mice (negative control.). (**C**) Box and violin plots depicting the distribution of hydroxymethylation values of the three sets of CpGs at different times after LPS/IL4 B cell activation. (**D**) Venn diagram showing the overlap between mouse P-AID CpG-associated genes and AID off-target genes (8). Statistical test: Two-tailed unpaired Wilcoxon's test (**C**) (* p-value < 0.05, **** p-value < 0.0001, ns is not significant).

Supplementary Figure 4. Aberrant DNA methylation in naïve B cells of HIGM2 patients (**A**) Box and violin representations of the distribution of DNA methylation levels of differentially hypermethylated regions in HIGM2 naïve B cells compared to controls. (**B**) Genomic location annotation of DMRs obtained comparing naïve B cells of patients with controls in the context of CpG islands (CGIs) and gene-related regions. (**C**) Genomic location analysis in relation to annotated genes of DMRs situated within CGIs (left) and outside of CGIs (right). **(D)** Bubble chart depicting the enrichment (red) or depletion (blue) of chromatin states obtained from ChromHMM database in germinal center B cells. Color represents logarithmic-fold change, dot size indicates percentage of DMRs in each chromatin state, and the edge indicates the statistical significance of the enrichment (black: significant, none: not significant; q-value < 0.01). (**E**) Smoothed DNA methylation data of altered DMRs associated with *MEF2A*. Tracks that represent H3K27ac ChIP-seq signal and the presence of super-enhancers and enhancers in germinal center B cells are indicated. (**F**) Bubble scatter plot of transcription factor motif enrichment using HOMER software in DMRs aberrantly methylated in HIGM2 naïve B cells. Transcription factors downstream of BCR signaling, as identified in NetPath database (19), are colored according to the transcription factor family. Bubble size corresponds to the logarithm of adjusted p-values.

Supplementary Figure 5. (**A**) Hypermethylated DMRs were identified comparing patient naïve B cells to controls. Box and violin representations of DNA methylation of regions corresponding to hypermethylated DMRs in naïve and memory B cells from control and HIGM2 from this study, as well as data from resting B cells (unstimulated), B cells activated with CD40L/IL4 and B cells infected with Epstein-Barr virus (EBV), obtained from Hansen et al (35). (**B**) Dot plot showing the

DNA methylation values determined by pyrosequencing of naïve B cells (naïve) and EBV-infected naïve B cells (EBV) from healthy donors. Composite CpG methylation levels surrounding BATF and IRF4 (**C**), and JUND, RAD21, ZNF273 and MYC. (**D**) ChIP-seq peaks (2.5 kb) in naïve and memory B cells from HIGM2 patients and controls. (**E**) Volcano plot depicting mRNA expression data comparing wild type with IRF4 (left) or BATF (right) GM12878 knockouts (54). Colored dots represent the genes that are associated with a DMR from the comparison between HIGM2 naïve B cells and control naïve B cells.

Supplementary Figure 6. RNAseq analysis of memory B cells transcriptomes from HIGM2 patients and healthy donors. (**A**) Volcano plot representation of the transcriptomic comparison between memory B cells of HIGM2 patients and controls. In blue, genes that are differentially expressed (absolute of fold change ≥2 and adjusted p-value < 0.05). (**B**) Gene Ontology analysis representation of differentially expressed genes. All ontologies represented are significantly enriched (adjusted p-value < 0.05). Gene ratio refers to the ratio between the unique genes present in the DMRs that are associated with a Gene Ontology term and all the genes potentially included in that ontology geneset. Dot size represents the number of times each term is represented in the DMRs. Dot color represents adjusted p-value. (**C**) Heatmap representation showing TF activity predicted using DoRothEA from mRNA expression of differentially expressed genes in memory B cells of HIGM2 compared to controls. Color represents normalized enrichment scores (NES) where blue and red represent a decreased and increased activity, respectively, in comparison to background.