Title
Integrative transcriptome and chromatin landscape analysis reveals distinct epigenetic regulations in human memory B cells

Permalink
https://escholarship.org/uc/item/0rd5p831

Journal
Nature Communications, 11(1)

ISSN
2041-1723

Authors
Moroney, Justin B
Vasudev, Anusha
Pertsemlidis, Alexander
et al.

Publication Date
2020-10-01

DOI
10.1038/s41467-020-19242-6

Peer reviewed
Integrative transcriptome and chromatin landscape analysis reveals distinct epigenetic regulations in human memory B cells

Justin B. Moroney\textsuperscript{1}, Anusha Vasudev\textsuperscript{1}, Alexander Pertsemidis\textsuperscript{2}, Hong Zan\textsuperscript{1} & Paolo Casali\textsuperscript{1}\textsuperscript{✉}

Memory B cells (MBCs) are long-lived and produce high-affinity, generally, class-switched antibodies. Here, we use a multiparameter approach involving CD27 to segregate naïve B cells (NBC), IgD\textsuperscript{+} unswitched (unsw)MBCs and IgG\textsuperscript{+} or IgA\textsuperscript{+} class-switched (sw)MBCs from humans of different age, sex and race. Conserved antibody variable gene expression indicates that MBCs emerge through unbiased selection from NBCs. Integrative analyses of mRNAs, miRNAs, IncRNAs, chromatin accessibility and cis-regulatory elements uncover a core mRNA-ncRNA transcriptional signature shared by IgG\textsuperscript{+} and IgA\textsuperscript{+} swMBCs and distinct from NBCs, while unswMBCs display a transitional transcriptome. Some swMBC transcriptional signature loci are accessible but not expressed in NBCs. Profiling miRNAs reveals downregulated MIR181, and concomitantly upregulated MIR181 target genes such as RASSF6, TOX, TRERF1, TRPV3 and ROR\textalpha, in swMBCs. Finally, IncRNAs differentially expressed in swMBCs cluster proximal to the IgH chain locus on chromosome 14. Our findings thus provide new insights into MBC transcriptional programs and epigenetic regulation, opening new investigative avenues on these critical cell elements in human health and disease.
cell maturation in the bone marrow, peripheral differentiation, and survival is dictated by genetic programs that drive critical pathways, giving rise to plasma cells and memory B cells (MBCs). Upon activation by antigen, mature naïve B cells (NBs) enter the germinal center (GC) reaction to undergo activation-induced deaminase (AID)-mediated somatic hypermutation (SHM) and class switch DNA recombination (CSR), leading to plasma cell or MBC differentiation. MBCs are central to the adaptive immune response, establishing long-term antigen-specific immunity. Their generation and function are critical for the development of anamnestic antibody responses induced by vaccines or natural infections. Effective anamnestic immune response to foreign antigens, such as those on bacteria and viruses, is required for the establishment of “herd immunity” and eradication of an infectious agent in a population as a whole. MBCs generally express somatically mutated, affinity mature B cell surface receptors (BCRs) for antigen and persist in a quiescent state until re-encountering their cognate antigen. While many MBCs express class-switched BCRs, such as immunoglobulin (IgG or IgA, "unswitched" IgD+ IgM- (unswMBCs) also exist, with human switched (swMBCs) and unswMBCs bearing somatically mutated BCRs and expressing surface CD27, a marker of "antigen-experience." Upon reactivation, MBCs rapidly proliferate and differentiate into plasma cells to secrete large amounts of high-affinity antibodies. Alternatively, MBCs re-enter the GC reaction to undergo additional rounds of SHM/CSR. Thus, enhanced responses against previously experienced antigens are contingent upon functional MBCs, which possess intrinsic advantages over NBs, including longer lifespan and accelerated antibody response.

Research to date, primarily in the mouse, has attempted to attribute intrinsic differences between MBCs and NBs to select gene expression programs in MBC differentiation, survival, and reactivation. IL-9 signaling has been suggested to drive MBC development within GCs as well as mediate antibody recall responses. High Bch2 expression biases GC B cells towards MBC differentiation, while reduced Bch2 expression skews MBCs towards plasma cell differentiation upon reactivation. Divergent transcriptional programming has been tentatively identified in MBCs with mouse IgG2a- and IgA+MBCs utilizing T-bet and Rora, respectively, to maintain their distinct identities and functions. As recently reported, the transcription factor (TF) Hhex interacts with the corepressor Tle3 to promote MBC differentiation, but not maintenance of mouse MBCs. Finally, MBCs accelerated and heightened response has been attributed to reduced expression of quiescence factors, such as ZBTB16, KLF4, and KLF9, as well as enhanced intrinsic signaling of class-switched BCR IgY chain. This carries tail tyrosine ITT motifs that reduce IgG+MBCs activation threshold.

The human MBC transcriptome and its regulation, particularly by epigenetic mechanisms, need better understanding. As we and others have shown, coordinated regulation of gene networks is critical in human and mouse B cell differentiation and antibody responses. Such coordinated regulation is mediated by epigenetic modifications and factors, including DNA methylation, histone post-translational modifications, and non-coding RNAs (ncRNAs), such as microRNAs (miRNAs) and long non-coding RNAs (lncRNAs). These layers of epigenetic regulation synergize with TFs and chromatin accessibility to outline distinct gene expression programs, thereby dictating cell functions. Although different miRNA profiles have been reported in different B cell subsets, including MBCs, and lncRNAs have been associated with different stages of B cell development, no comprehensive analysis of the protein-coding and non-coding transcriptome and chromatin accessibility has been reported in human MBCs.

Integrative analysis of differentially expressed mRNAs, miRNAs, lncRNAs, chromatin landscape, and cis-regulatory elements reveals how gene expression intersects with epigenetic regulation to segregate distinct MBCs. Analysis of CD27 IgD+ NBs, CD27 IgD+ unswMBCs, CD27 IgG+ swMBCs and CD27 IgA+ swMBCs as well as total CD27+ NBs and CD27+ MBCs from healthy human subjects defines a differential transcriptional signature that distinguishes CD27+ IgG+ swMBCs and CD27+ IgA+ swMBCs from CD27+ IgD+ NBs, stratifying CD27+ IgD+ unswMBCs between the two. The integration of RNA-Seq with ATAC-Seq data uncovers distinct profiles of cis-regulatory elements in human MBCs. MIR181 downregulation with concomitant upregulation of this microRNA’s target genes indicates an important role for MIR181 in MBC function(s). Further, the inverse correlation of upregulated MIR181 sponge, lncRNA MIAT, with downregulated MIR181a/MIR181b in swMBCs points to MIAT as a regulator that, along with chromatin accessibility, reduces MIR181 expression and promotes swMBC-specific gene expression. Overall, our findings provide evidence for overlapping layers of regulation, including chromatin remodeling, cis-regulatory elements, and distinct sets of miRNAs and lncRNAs, which integrate to dictate gene expression profiles and activation pathways characteristic of human MBCs.

**Results**

Identification and purification of human swMBCs, unswMBCs, and NBs. To elucidate the transcriptional landscape of human MBCs, we set up to identify and isolate MBCs and their NB counterparts. We purified B cells (>99% CD19+) from PBMCs of 7 healthy human subjects (age 20–39, 4 males, 3 females of different race and ethnic backgrounds) by immunomagnetic negative selection (Fig. 1a). CD19+ cells were analyzed for expression of CD27, a marker for antigen-experienced B cells, as well as IgD, IgM, IgG, and IgA. This allowed for the identification of four distinct subsets: CD27+IgD- NBs (63.9 ± 14.3%), CD27-IgD- unswMBCs (97. ± 8.3%), CD27-IgG+ swMBCs (64. ± 3.1%) and CD27-IgA+ swMBCs (2.0 ± 2.0%). A fifth subset identified "double negative" CD27-IgD- accounted for the remaining 15.0% of B cells. CD27- B cells were IgD+IgM+, while CD27- B cells comprised comparable proportions of IgD+ and IgD- B cells. CD27+IgD- B cells were mostly IgM- with only 3.8 ± 0.0% IgD-IgM- B cells. CD27+IgD+ B cells were primarily IgG+ or IgA+ (Fig. 1b). CD19+ cells from subjects B, C, and G were sorted into 4 fractions: CD27+IgD- NBs (70.2 ± 15.2%), CD27-IgD+ unswMBCs (10.1 ± 5.1%), CD27-IgG+ swMBCs (6.6 ± 5.2%) and CD27-IgA+ swMBCs (4.3 ± 2.7%). Reanalysis of the sorted fractions confirmed the identity and purity of the 4 discrete B cell subsets (Fig. 1c). Finally, CD19+ cells from the remaining four subjects (A, D, E, and F) were analyzed and isolated as CD27- (NBs) and CD27+ total (MBCs) (Supplementary Table 1).

IgVH1, JH1, and Jv1 gene expression in human NBs and MBCs. We analyzed the expression of Ig Vg1, IgD, JH1, and Vk, Jk, as well as Vα, Jα genes in CD27-IgD+ NBs, CD27-IgD+ unswMBCs, CD27-IgG+ swMBCs, and CD27-IgA+ swMBCs. The human IgH locus comprises 36–49 functional VH genes segregated into 7 families. In all four subsets within each subject, VH1 family expression was stochastic (r = 0.98), reflecting the genomic representation of the 7 VH1 families, with VH13 genes being the most represented (37.8–59.9%), followed by VH14 (9.7–26.4%) and VH11 (11.3–23.7%) (Fig. 2a). Expression of individual VH1 genes was conserved in each VH1 family, across the 4 subsets in each subject (r = 0.88) and across the three subjects (r = 0.84) (Fig. 2b). As expected, D gene expression was diverse across the four subsets in each subject (r = 0.49) and across the three subjects (r = 0.17) (Fig. 2c, d). J11 gene usage was conserved across the four subsets within each subject (r = 0.90), with a preponderant J15.
immunomagnetic negative selection. The indicated gating strategy and two distinct staining panels. Three independent experiments. The purity of the sorted four B cell subsets as analyzed by (Fig. 3a). A salient feature of MBCs is the load of somatic point-mutations in expressed Ig genes. CD27+ B cells, be they IgG+ , IgA+ or IgD+ exhibited a higher frequency of somatic point-mutations (0.0314, 0.0566 and 0.0417 change/base, respectively), mostly replacement mutations, than NBCs (Fig. 3b, c), with IgG+ swMBCs and IgA+ swMBCs carrying greater mutational burdens than unswMBCs.

19.2–40.5%) and Ig6 (27.8–67.7%) utilization in all three subjects, albeit to differing degrees (Fig. 2c, f).

The human Igk locus comprises 39 functional Vk and 5 Jk genes. Vx and Jk gene expression were comparable across the four subsets within each subject (r = 0.88; 0.86, respectively) and across the three subjects (r = 0.86; 0.82, respectively). Vx gene usage was biased towards Vx3 and Vx4 and away from Vx1 and Vx2. Jk gene usage was biased towards Jk1, Jk3, and Jk5 (Supplementary Fig. 1a–d). The human Igk locus comprises 30 functional Vλ genes segregated into 10 subgroups and 5 functional Jλ-CA clusters. Vx gene expression was comparable across the four subsets within each subject (r = 0.82) and across the three subjects (r = 0.83), with over-representation of Vλ3, Vλ2, and Vλ1 subgroups (Supplementary Fig. 1e, f). Jλ-CA usage was biased towards Jκ3-CA3 (35.2–57.5%) and Jκ2-CA2 (17.9–54.6%), with conserved distributions in all four subsets (r = 0.87) (Supplementary Fig. 1g, h).

IgH CDR3 length and somatic point-mutations in human MBCs. Somatic IgH VhDJh rearrangement determines the sequence and length of the complementary determining region 3 (CDR3), which is critical for BCR-antigen contact. To address IgH CDR3 length and nature as well as Vh multiallo mutation, swMBCs, unswMBCs, and NBCs, the recombined VhDJh CDR3 transcripts were amplified using forward (degenerate) primers for Vh1, Vh2, Vh3, Vh4, Vh5, Vh6, Vh7 genes leader sequences in conjunction with reverse Cμ, Cδ, Cγ, or Cκ isotype-specific primer and analyzed by IMGT/HighV-QUEST. The distribution of CDR3 lengths was largely conserved among NBCs (16.05 ± 0.10 AA’s), unswMBCs (14.86 ± 0.11 AA’s), IgG+ swMBCs (15.53 ± 0.12 AA’s), and IgA+ swMBCs (15.08 ± 0.11 AA’s) (Fig. 3a). A salient feature of MBCs is the load of somatic point-mutations in expressed Ig V(D)J genes, a result of precursor B cells undergoing antigen-driven SHM and selection. Consistent with their non-antigen-experienced status (CD27−), NBCs exhibited a negligible frequency of somatic point-mutations (0.0020 change/base). By contrast, consistent with their antigen-experienced status, CD27+ B cells, be they IgG+ , IgA+ or IgD+ exhibited a higher frequency of somatic point-mutations (0.0314, 0.0566 and 0.0417 change/base, respectively), mostly replacement mutations, than NBCs (Fig. 3b, c), with IgG+ swMBCs and IgA+ swMBCs carrying greater mutational burdens than unswMBCs.

A core transcriptional signature distinguishes IgG+ swMBCs and IgA+ swMBCs from NBCs. Using next-generation sequencing, we identified differentially expressed (DE) mRNA transcripts in CD27+ IgG+ vs. CD27− IgD+ B cells and CD27+ IgA+ vs. CD27− IgD+ B cells by pairwise comparisons. IgCy and IgCa transcripts were significantly increased and IgCu and IgCg correspondingly decreased in CD27+ IgG+ and CD27+ IgA+ swMBCs (Fig. 4a, b). The higher IgCy/I and IgCa/I expression reflected the peripheral blood predominance of IgG1 and IgA2 subclasses and respective B cells. Analysis at the highest level of significance of padj < 10 × 10−30 identified the same 24 (17 upregulated and 7 downregulated) DE mRNAs (not including IgCδ) in both IgG+ and IgA+ swMBCs as compared to their NBC counterparts (Fig. 4a, b and Supplementary Fig. 2a, b, Supplementary Table 2). RSPI7, a ribosomal subunit protein, and RUNX2, a TF essential for CSR to IgA, were highly expressed in IgA+ swMBCs as compared to IgG+ swMBCs (Fig. 4c). IgCe expression in IgG+ swMBCs, but not IgA+ swMBCs, likely reflected the sequential CSR poise from IgG1 to IgE of such swMBCs. Importantly, the “cancellation” of the 24 DE genes when comparing CD27− IgG+ vs. CD27− IgA+ B cells outlined a gene expression profile characteristic and distinctive of swMBCs, regardless of BCR isotype (Fig. 4c and Supplementary Table 2). This was strengthened by an equivalent level of normalized expression of the 24 transcripts in IgG+ swMBCs and IgA+ swMBCs (Fig. 4d).
Fig. 2 Expressed Ig V̄, D, and JH repertoire in human CD27⁺ IgD⁺, CD27⁺ IgD⁻, CD27⁻ IgG⁺, and CD27⁻ IgA⁺ B cells. a, c, e: Ig V₁, D (a), (c), and JH (e) gene expression in CD27⁻ IgD⁺, CD27⁺ IgD⁺, CD27⁺ IgG⁺, and CD27⁻ IgA⁺ B cells, depicted by stacked column for each subject (subjects B, C, G). Ig gene family members grouped within each Ig V₁, D, JH family or subgroup are depicted in shades of the respective family/group color. V₁ family: V₁1 (blue), V₁2 (green), V₁3 (red), V₁4 (yellow), V₁5 (teal), V₁6 (orange), V₁7 (purple). D family: D1 (blue), D2 (green), D3 (red), D4 (yellow), D5 (teal), D6 (orange), D7 (purple). JH family: JH1 (blue), JH2 (green), JH3 (red), JH4 (yellow), JH5 (teal), JH6 (orange). b, d, f: Correlation matrix showing the relationship of individual Ig V₁ (b), D (d), and JH (f) gene expression between B cell subsets in the three subjects analyzed. Color bars depict the correlation coefficients on a scale from 0.5 (white) to 1 (red) (b, f), or from –1 (blue) to 0 (white) to 1 (red) (d), with color, circle size, and value all indicating the overall strength of the correlation. B cell subsets are repeated at the bottom, left to right in CD27⁻ IgD⁺ (gray), CD27⁺ IgD⁺ (lavender), CD27⁺ IgG⁺ (purple), and CD27⁺ IgA⁺ (dark purple).

The transcriptional distance between the four B cell subsets, as measured by principal component analysis (PCA), yielded three discrete clusters, identifying NBCs, unswMBCs, and swMBCs (Fig. 4e and Supplementary Fig. 2c). The 24 DE genes discriminated IgG⁺ swMBCs and IgA⁺ swMBCs from the other two B cell subsets across all three subjects (Fig. 4f). Further, the 24 DE genes discriminated total MBCs (CD27⁺ B cells) from NBCs (CD27⁻ B cells) in the 4 additional healthy subjects (A, D, E, F) (Supplementary Table 1 and Supplementary Fig. 2c, d). Thus, at the highest level of significance, 24 DE mRNAs characterizedly distinguish class-switched IgG⁺ swMBCs and IgA⁺ swMBCs from their NBC counterparts (swMBC core transcriptional signature).

Gene expression outlines distinct signaling and TF activities in swMBCs. Among the 24 swMBC core transcriptional signature genes, six have a role in cell cycle and apoptosis, and six are implicated in gene expression (Fig. 5a). The aggregate mRNA profile (DE transcripts, Padj < 0.05) in MBCs was explored by IPA and revealed significant over-representation in leukocyte extravasation as well as cytokine and MAP kinase signaling pathways. The activation status of such pathways differed significantly between swMBCs and NBCs (Z-score) (Fig. 5b), suggesting a contribution of such pathways to MBC function. As mRNA expression fails to capture post-transcriptional and post-translational events, we inferred the differential activity of TFs in human MBCs by applying a master regulator inference algorithm (MARINA) using a human B cell-specific Bayesian inter-actome31. In swMBCs, MEF2B and BATF were inferred to have increased activity, while GTF2I, EGR3, and multiple zinc-finger TFs were inferred to have decreased activity (Fig. 5c). Two DE genes, RORα, and ZBTB16, were central to a transcriptional network that incorporated MARINA-inferred TFs, nucleus-localized swMBC core transcriptional signature genes, and IPA-identified coactivators and corepressors (Fig. 5d). Probing deeper
into MARINa TF analysis, ZBTB16 was found to be significant (p = 0.034), ranking 15th out of the 621 TFs in the regulon. RORα was not part of the transcriptional “regulon” object31 and, consequently, its activity was not addressed here. Analysis of RORα protein, however, showed it to be significantly expressed in swMBCs and unswMBCs across the seven human subjects (Fig. 5e). Further, RORα protein was analyzed in five discrete tonsil B cell subsets from three additional subjects and was highly expressed in such tonsil swMBCs (Fig. 5f). Thus, our integrative analysis identified signaling pathways and TF networks that are important to MBC identity and/or functions, suggesting a central role for TF RORα in the phenotype/function of human swMBCs.

unswMBCs display a transcriptional profile of transition between NBCs and swMBCs. Although the memory phenotype of unswMBC is an accepted notion, the nature of these CD27+ IgD+ B cells remains a matter of debate. Virtually all CD27+ IgD+ B cells retained expression of surface IgM (Fig. 1b), yet they were reported to express mutated Ig V<sub>H</sub> genes3. These unswMBCs carried significantly more somatic point-mutations than NBCs (Fig. 3c, b), and displayed (pairwise comparison) two DE upregulated transcripts, TFEC and ZBTB32, at $p_{adj} < 10 \times 10^{-30}$, as well as four swMBC core signature transcripts (TRPV3, TACI, COL4A4, and SAMSN1), upregulated at a significance of $p_{adj} < 10 \times 10^{-15}$ (Supplementary Fig. 3a, b and Fig. 4d). PCA showed unswMBCs clustering between NBCs and swMBCs (Fig. 4e) and displayed at lower levels signaling pathways characteristic swMBCs (Fig. 5b). PCA and hierarchical clustering analyses were expanded across the DE gene profile at differing levels of significance and consistently showed CD27+ IgD+ unswMBCs as an independent cluster, slightly closer to swMBCs (Supplementary Fig. 4). Overall, CD27+ IgD+ unswMBCs display a transcriptome of transition from NBCs to swMBCs.

Chromatin accessibility and active gene landscape in MBCs and NBCs. As RNA expression is dependent upon DNA accessibility, we sought to define the chromatin landscape of human MBCs by ATAC-Seq and characterize its intersection with the transcriptome in these B cells (subjects A, D, E, F). Of the 77,388 accessible loci identified, NBCs (CD27− B cells) and total MBCs (CD27+ B cells) displayed distinct patterns of chromatin accessibility, as indicated by component 1 of PCA (74.6% variance) (Fig. 6a). Overall, total MBCs trended towards higher chromatin accessibility, with 43.6% of identified peaks unique to them, 11.9% of peaks unique to NBCs, and 44.5% of peaks accessible in both total MBCs and NBCs (Fig. 6b). Differentially accessible regions (DARs) at $p < 0.05$ were identified in total MBCs vs. NBCs. Overall, 4198 loci were differentially accessible in total MBCs vs. NBCs, with 2169 (51.7%) DARs with increased accessibility and 2029 (48.3%) with decreased accessibility (Fig. 6c).

To determine whether cell-specific open chromatin regions identified by ATAC-Seq correlated with cell-specific gene expression, we integrated our NBC and MBC ATAC-Seq and mRNA-Seq data sets. Overall, 462 and 612 DE mRNAs were at significantly higher or lower levels, respectively, in both IgG+ swMBCs and IgA+ swMBCs compared to NBCs at $p_{adj} < 0.05$. Of the 462 upregulated genes, 161 displayed increased chromatin

---

Fig. 3 CDR3 lengths and somatic point-mutations in recombined Ig V<sub>H</sub>D<sub>JH</sub> gene segments of human MBCs. a Average percentage of total sequences at any given CDR3 length in recombined V<sub>H</sub>D<sub>JH</sub> transcripts expressed by CD27+ IgD+, CD27+ IgG+, and CD27+ IgA+ B cells. b Somatic point-mutations in recombined Ig V<sub>H</sub>D<sub>JH</sub> transcripts expressed by CD27+ IgD+ (4302 transcripts), CD27+ IgG+ (4705 transcripts), CD27+ IgA+ (1564 transcripts), and CD27+ IgA+ (13,549 transcripts) B cells with boxplots depicting the frequencies of point-mutations (change/base). Data are represented as boxplots where the middle line is the median, the lower and upper hinges correspond to the first and third quartiles, the upper whisker extends from the hinge to the largest value no further than 1.5 × IQR from the hinge (where IQR is the inter-quartile range) and the lower whisker extends from the hinge to the smallest value at most 1.5 × IQR of the hinge. *p < 0.05, **p < 0.01, ***p < 0.001, ns not significant (paired two-sided t-test). c Frequency of silent and replacement point-mutations in framework regions (FR) and complementarity-determining regions (CDRs) in all four sorted B cell subsets for each of the three healthy human subjects. Data are mean ± SEM of three independent experiments. B cell subsets are repeated at the bottom, left to right in CD27+ IgD+ (gray), CD27+ IgG (lavender), CD27+ IgA+ (purple), and CD27+ IgA (dark purple).
accessibility nearby or within their loci. Of the 612 downregulated genes, 219 displayed decreased chromatin accessibility nearby and/or within their loci. Thus, of 1074 DE mRNAs, 35.4% displayed corresponding increases or decreases in chromatin accessibility near their genetic loci (Fig. 6d). Chromatin accessibility was altered in the exon and promoter regions of 5 of 7 downregulated and 9 of 17 upregulated genes in the swMBC core transcriptional signature (Supplementary Fig. 5a–c). Gains in chromatin accessibility were associated with genes encoding proteins involved in MAPK, BCR, and Ca2+ signaling and leukocyte migration, validated by pathway analysis of the mRNA transcriptome (Fig. 5b). Losses in chromatin accessibility were associated with genes involved in Sphingolipid, MAPK, and stem-cell signaling, focal adhesion, and autophagy (Fig. 6e).

As the transcriptional regulation of human MBCs is poorly understood, we sought to determine whether MBCs DARs were associated with given cis-regulatory elements. Significant enrichment in DNA-binding factors that poise B cells for proliferation and plasma cell differentiation were found in DARs with increased accessibility in total MBCs, such as OCT2, IRF8, and bZIP motifs43. Conversely, enrichment in DNA-binding factors involved in B cell development and cellular quiescence, such as ERG, RHD, and KLF motifs21,44, were found in decreased accessible DARs of total MBCs (Fig. 6f). To further assess changes in TF activity, we probed for the enrichment of cis-regulatory elements in all accessible loci found in total MBCs over NBCs (ΔMD score). NR1H4- and RXRA-binding motifs were enriched in the open chromatin regions over NBCs in swMBCs based on patterns of mRNA expression (ΔMD score). These differences in motif accessibility were supported by MARINA-inferred changes in related TF activity. For example, BATF activity was inferred to be increased in swMBCs based on patterns of mRNA expression, and BATF-binding motifs were enriched in increased DARs in total MBCs (Figs. 5c and 6f, g).

Taiji was used to develop an integrative transcriptional regulatory network and predict TF activity based off a personalized PageRank algorithm. Consistent with our previous
analyses, Taiji identified TFEC, RORA, and BATF as the TFs with the highest activity in MBCs, while KLF4, HEY2, and LEP1 were identified as the TFs with the highest activity in NBCs (Fig. 6h, i). Consistent with the notion that MBCs are poised to differentiate into plasma cells, PageRank inferred increased activity of PRDM1 and decreased activity of BCL6 in MBCs (Fig. 6h). Thus, the profile of chromatin accessibility contrasted human MBCs with NBCs and outlined signaling pathways and cis-regulatory elements relevant to MBCs.

**MicroRNAs in MBCs and identification of MIR181 as a major gene expression regulator.** While chromatin accessibility is necessary for transcription, gene expression is subject to additional layers of regulation, mostly epigenetic, both at the DNA as well as post-transcriptional level. miRNAs are small ncRNAs, which target mRNA 3′ UTRs to post-transcriptionally repress gene expression\(^\text{45}\). As we have shown, miRNAs play a critical role in modulating B cell SHM, CSR, and plasma cell differentiation\(^\text{26,29,45}\). miRNAs role in MBCs is less understood. Our analysis of the miRNAome identified 19 DE miRNAs, 6 upregulated and 13 downregulated, in both IgA\(^+\) swMBCs and IgA\(^+\) swMBCs vs. NBCs at a significance of \(p < 0.05\) (Fig. 7a, b). These 19 DE miRNAs “canceled out” when IgG\(^+\) swMBCs were compared with IgA\(^+\) swMBCs (Fig. 7c), being expressed at similar levels in IgG\(^+\) swMBCs and IgA\(^+\) swMBCs (Fig. 7d and Supplementary Fig. 3c). Although also expressed in some cases in unswMBCs, they overall discriminated MBCs from their NBC counterparts (Fig. 7e, f). Among the 13 downregulated miRNAs in swMBCs, MIR181a and MIR181b, both MIR181 family members, were decreased by more than 60,000 and 3000 transcripts (Fig. 7d). MIR181a and MIR181b transcript down-regulation were concomitant with decreased chromatin accessibility at these miRNA’s host–gene (HG) promoters (Supplementary Fig. 6a).
**Fig. 6** The chromatin landscape of human MBCs. a Clustering of CD27⁻ NBCs and CD27⁺ total MBCs based on top differentially accessible loci as displayed by PCA (4198). Prediction ellipses define 95% confidence intervals. Each symbol represents an individually sorted subset (n = 4). b Overlap of accessible loci between NBCs and total MBCs depicted by Venn diagram. c Differential chromatin accessibility (p < 0.05) between total MBCs and NBCs as determined by DEseq2. Differentially accessible regions (DARs) with increased (red) or decreased (blue) accessibility depicted by histograms. d Proportion of DE genes that overlap with DARs, as depicted by Venn diagrams. Percentage of genes exhibiting corresponding increases/decreases in chromatin accessibility shown below the number of genes. e Statistically over-represented KEGG pathways associated with increased (red) or decreased (blue) DARs depicted by the histogram. f Significant TF-binding motifs enriched in NBC- or all MBC-specific open chromatin regions as determined by HOMER motif analysis. g For all motifs (dots), the changes in MD score between NBCs and total MBCs (y-axis) are (MA) plotted against the number of motifs within 1.5 kb of any ATAC-Seq peak center (x-axis)—computed by DASh89. Red points depict statistically significant increased MD-scores (p < 0.05). h, i RNA-Seq and ATAC-Seq data sets were integrated using the Taiji PageRank algorithm90 to generate PageRank activity scores and p-values for human Tfs. h All 421 Tfs (p < 0.05) plotted in rank order according to their log2 pagerank score ratio between MBCs vs. NBCs. Each TF PageRank —log2(p) is denoted by color-coding according to legend: more significance, red; less significance, blue. i The top 10 TFs in NBCs and top 10 TFs in MBCs was plotted according to their log2 pagerank score ratio and log2 fold change in transcript expression in MBCs vs. NBCs. The bubble size of each TF is determined by the —log2(p) of Taiji analysis.

Downregulation of MIR181a and MIR181b was expected to relieve the silencing of MIR181 targeted mRNAs. Indeed, 5 of the 17 significantly upregulated mRNAs in swMBC core transcriptional signature, RASSF6, TOX, TREFR1, RORa, and TRPV3, were predicted targets of MIR181a and MIR181b, exhibiting 3’UTR sequences highly complementary with MIR181 seed regions and favorable thermodynamics (Fig. 8a)—overall mRNA profile at repair < 0.05 showing 50 of the 462 mRNAs (10.8%) upregulated in both IgG⁺swMBCs and IgA⁺swMBCs to be predicted targets of MIR181 (Supplementary Fig. 6b). A significant inverse correlation between MIR181a/MIR181b and their 5 mRNA targets indicated the epigenetic release of these select mRNA transcripts in swMBCs (Fig. 8b). Indeed, enforced expression of MIR181a in retroviral-transduced mouse B cells significantly reduced the expression of these five genes by 32.3–72.4% (Fig. 8c, d and Supplementary Fig. 6c).

Further, luciferase reporter assays involving (CD154 and IL-4) induced human CL-01 B cells expressing MIR181 and transfected with wildtype (WT) or mutant (Mut) reporter constructs confirmed TOX, TREFR1, and TRPV3 to be direct and specific targets of MIR181 (Fig. 8e, f)—RASSF6 was shown to be a direct target of MIR181 family members in the context of gastric cancer86; RORa was not tested because of the length and complexity of its 3’UTR, which, however, contains three MIR181 target sites. Thus, MIR181 releases select swMBC gene expression from epigenetic regulation.

Potential regulatory capacities of lnRNAs in human swMBCs. lncRNAs are greater than 200 nucleotides in length and can interact with DNA, RNA, and proteins, thereby mediating a layer of epigenetic regulation to positively or negatively modulate gene expression87. lncRNAs have been implicated in key processes of
T cells and innate lymphocytes. While profiled in B cells, their functional significance in human MBCs has not been explored. A comprehensive analysis by pairwise comparisons at \( p < 0.001 \) of MBC IncRNA landscape identified the same 23 upregulated and 17 downregulated DE IncRNAs in IgG+ swMBCs and IgA+ swMBCs as compared to NBCs—among these 40 DE IncRNAs, 21 had a decreased expression of such lncRNAs and mRNA regulatory interactions (Fig. 7a, b).

Intronic analyses paired with functional and spatial enrichment. By integrated co-expression analysis, we identified significant positive and negative correlations between DE IncRNAs and DE miRNAs in swMBCs as compared to NBCs (Fig. 7a, b). One cluster of positively correlated transcripts found at the telomeric end of chromosome 14, upstream of the \( I g H \) locus contained a downregulated IncRNA, \( T C L 6 \), \( R P 1 1 - 1 6 4 H 1 3 . 1 \), and \( T U N A R \), together with 1 downregulated mRNA, \( T C L 1 A \) (Fig. 10c). The decreased expression of such IncRNAs and mRNA reflected the decreased chromatin accessibility at their respective loci (Fig. 10d).

Interestingly, IncRNA \( A L 9 2 8 7 6 8 . 3 \) found to be expressed at 10.8-fold higher levels in IgG+ MBCs than IgA+ MBCs is located 2.5 kb downstream of the \( I G H A 2 \) locus. IncRNAs \( C O P D A 1 \) and \( R P 1 1 - 7 3 I F 5 . 1 \) were found to be expressed at 6.0 and 4.6-fold higher levels in IgG+ MBCs than IgA+ MBCs and are located in putative regions of \( S T / M I A T \) and \( G I H 2 / G I H E \) loci, respectively (Fig. 10c, d). MIAT, a molecular sponge for \( M I R 1 8 1 b \), was consistently upregulated in swMBCs and unswMBCs of all three subjects and displayed significant negative correlations with \( M I R 1 8 1 a \) and \( M I R 1 8 1 b \) expression (Fig. 10e).

To outline potential IncRNA regulatory functions, \( t r a n s \) and \( c i s \) co-expression correlations of IncRNAs with miRNAs or mRNAs were analyzed in swMBC, unswMBCs, and NBCs (Supplementary Fig. 7a–d). \( T r a n s \) IncRNA-miRNA correlations trended positive \(( r > 0.7, 2.60%) \) and negative \(( r < -0.7, 2.60%) \), while \( c i s \) IncRNA-mRNA correlations showed a wide range of correlation coefficients \(( r > 0.7, 3.97%; r < -0.7, 3.05%) \) (Supplementary Table 3). IncRNAs with strong positive or negative \( c i s \) correlations \(( r > 0.65 \) or \( r < -0.65 \)) with mRNAs genes revealed enrichment in lipid metabolic genes and GPCR signaling (Supplementary Fig. 7e, i). Thus, a lncRNA expression...
profile is characteristic of all CD27+ MBCs as compared to NBCs, with select lncRNAs proximal to the IgH locus displaying potential regulatory activity or acting as miRNA sponges and overall co-expression patterns of different RNA species indicating enrichment in distinct signaling pathways.

**Discussion**

mRNA, miRNA, and lncRNA transcriptomes revealed a common and unique transcriptional profile of human IgG+swMBCs and IgA+swMBCs, distinguishing them from NBCs. At the highest level of significance, swMBCs differed from NBCs in the expression of 24 genes, 17 increased and 7 decreased, with unswMBCs displaying a transcriptional signature of transition. The swMBC transcriptome was enriched in distinct MAPK, migratory and cytokine signaling pathways, likely influencing survival, homing and cell-to-cell interactions. In all (CD27+)

MBCs, the chromatin landscape reflected changes in gene transcription and outlined differentially accessible cis-regulatory elements, potentially influencing cell gene expression and functions. As compared to NBCs, swMBCs non-coding transcriptome revealed a characteristic downregulation of MIR181a/MIR181b, pointing to this miRNA as a key modulator of swMBC gene expression. Also, lncRNA MIAT, a molecular sponge of MIR181, was significantly increased in such swMBCs. MIAT likely released target genes from MIR181-mediated silencing, in concert with decreased accessibility of MIR181 HG promoters. Thus, when compared to NBCs, swMBCs exhibit a distinct mRNA, miRNA, and lncRNA transcriptome, dynamically interacting with a changing chromatin landscape to shape the identity and functions of these B cells.

CD27+IgD+, CD27+IgG+, CD27+IgA+ B cells displayed a pattern of IgVH gene expression conserved across
Fig. 9 lncRNA expression profile in human MBCs and NBCs. a–c Global transcriptional differences of lncRNAs in CD27⁺IgG⁺ vs. CD27⁺IgD⁺ B cells (a), CD27⁺IgA⁺ vs. CD27⁺IgD⁺ B cells (b), as well as CD27⁺IgG⁺ vs. CD27⁺IgA⁺ B cells (c), as depicted by volcano plots. All IncRNAs annotated in human GENCODE v24 GRCh38 are shown, each circle representing 1 IncRNA. –Log₁₀-transformed p is shown on the y axis (p < 0.001 indicated above dashed line). DE IncRNAs at p < 0.05 are highlighted in red (upregulated) or in blue (downregulated). DE IncRNAs at p < 0.001 (with Δ abundance > 100) common to both CD27⁺IgG⁺ and CD27⁺IgA⁺ are annotated. d Normalized (log₂RPKM) expression of the top 40 DE IncRNAs at p < 0.001 in swMBCs as compared to NBCs depicted by histogram for each B cell subset. Data are mean ± SEM of all subjects. e Transcriptome clustering of the sorted subsets performed using the 40 DE IncRNAs depicted by the PCA plot. Prediction ellipses define 95% confidence intervals. Each symbol represents an individually sorted subset (n = 3). f Relative expression profiles of swMBC core transcriptional signature IncRNAs at p < 0.001 compared by heatmap, depicting relative transcriptional changes across CD27⁺IgD⁺, CD27⁺IgD⁺, CD27⁺IgG⁺, and CD27⁺IgA⁺ B cell subsets in each subject (order: B, C, G). Data in a–f depict DE IncRNAs as determined by edgeR. B cell subsets; CD27⁺IgD⁺ (gray), CD27⁺IgG⁺ (lavender), CD27⁺IgA⁺ (purple), and CD27⁺IgA⁺ (dark purple).

NBCs, unswMBCs, and swMBCs within each subject and reflecting the genomic representation of the seven VH gene families. This was associated with a conserved distribution of CDR3 lengths in recombinant VH DJH-Cγ transcripts, indicating an unbiased clonal differentiation of MBCs from their NBC progenitors. This was further supported by the conservation of Ig JH, Vk, Jk, Vκ, and Jκ gene expression across NBCs, unswMBCs, and swMBCs, vastly overlapping across the subjects studied. Reflecting on a different degree of antigen experience, the mutational load in recombinant transcripts was the greatest in swMBCs, intermediate in unswMBCs, and negligible in NBCs.

Our findings outlined a distinct gene expression profile characteristic of human swMBCs, regardless of their effector Ig class, be it IgG⁺ or IgA⁺ and possibly IgE⁺. They extend previous studies, which addressed select human NBCs, conditionally by different antigens, microenvironments, and/or biased genetic programs⁵¹, and identified transcriptional programs not associated with MBCs of specific effector Ig classes⁵²,⁵³. In our seven subjects, NBCs accounted for two-thirds of the B cell repertoire, with unswMBCs and (IgG⁺ and IgA⁺) swMBCs making up for about one-tenth each. The 17 upregulated and seven downregulated mRNA transcripts distinguishing swMBCs from NBCs “canceled out” when comparing IgG⁺ MBCs with IgA⁺ MBCs, leaving the reciprocal IgH expression of the respective IgCy and IgCa transcripts as the distinguishing difference. Expression of IgCe transcripts in IgG⁺ MBCs, but not in IgA⁺ MBCs likely reflected the transition to IgE through sequential CSR from IgM to IgG1 and to IgE⁴².

The phenotype, transcriptome, functional cellular pathways, and gene loci accessibility yielded unifying information on MBCs of different isotypes, as different from NBCs in the same subjects. This was unexpected and per se remarkable, as we found a consistent homogeneity and concordance in the mRNAome, miRNAome, IncRNAome, chromatin accessibility, and inferred pathways of cell functions within MBCs—whether considered as CD27⁺IgG⁺ swMBCs and CD27⁺IgA⁺ swMBCs or total CD27⁺ B cells vs. NBCs, using different B cell isolation approaches (isotype-specific MBCs and total MBCs) and despite the
heterogeneous origin of the B cell samples: seven subjects of different age, sex, and race—a cohort quite different from any cohort of inbred mouse strains. Single-cell RNA-Seq may resolve further DE genes beyond this core transcriptional signature.

The identity and role of unswMBCs (CD27⁺IgD⁺) are a matter of debate. These B cells have been suggested to act as a reservoir of emergent memory precursors, thereby emphasizing the importance of such genes critical to ion channel Ca²⁺ flux, Ca²⁺ signal transduction, or signal transduction, or critical for MBC reactivation59. TOX, TRERF1, and RORA, encode for TOX, a high-mobility group (HMG) box protein, which binds DNA architectural motifs and is critical in immune cell development60. TRERF1, which interacts with p300 and SF-1 to regulate P450cc gene expression, thereby impacting the cell cycle61, and RORA.

RORA was highly expressed in IgG⁺ MBCs and IgA⁺ MBCs, but not in NBCs, suggesting a central role for this TF in the identity of human swMBCs. RORA is a sequence-specific ligand-dependent TF implicated in mouse ILC2 and Th17 cell differentiation as well as mouse IgA⁺ MBC maintenance20,62. RORA was found to be central to the transcriptional network of

![Image](https://example.com/image.png)

**Fig. 10 Characterization of the human MBC IncRNA profile.** a, b IncRNA-to-mRNA co-expression correlations depicted as Circos plots with human chromosomal ideograms (various colors). IncRNA-to-mRNA correlations showing 21 DE IncRNAs (Δ abundance>100; depicted as squares) in swMBCs and the top three positively (a, circles, r > 0) correlated mRNAs and top three negatively (b, triangles, r < 0) correlated mRNAs for each IncRNA with p < 0.05 (Pearson correlation). c Schematic of the terminal end of chromosome 14 (q32.2–q32.3), depicted with DE mRNAs, IncRNAs, and critical B cell gene loci annotated. DE mRNAs and DE IncRNAs in swMBCs, as well as differences in genetic loci between CD27⁺IgG⁺ swMBCs, CD27⁺IgA⁺ swMBCs, and CD27⁺IgD⁺ NBCs are depicted. Downregulated IncRNAs (dark blue), upregulated IncRNAs (dark red), DE IncRNAs between swMBCs (green). d Chromatin accessibility upstream and in the human IgH loci is displayed by IGV gene track. Coverage includes gene and IncRNA introns, exons, promoter regions, and potential enhancer regions. ATAC-Seq signal is normalized for the window of interest, with NBCs depicted in gray and total MBCs in purple. e Correlations between expressed mRNAs and MIAT as calculated across all sorted subsets (n = 12) by Spearman’s rank correlation.
MARINA-inferred TFs, nucleus-localized swMBC signature genes, and IPA-identified coactivators and corepressors. This indicates RORα's importance in MBC identity and/or functions, as further supported by the integrative Tajiri analysis. RORα's high expression in peripheral blood and tonsil swMBCs further suggested an important role of this TF in MBCs maintenance. MUC16, which encodes a 22,000 amino acid membrane-associated adhesion molecule, was significantly upregulated in swMBCs. Through its binding partners galectin-1, galectin-3, and L-selectin, abundantly expressed by CD169+ macrophages, Mucin-16 would stabilize cell-to-cell contact between MBCs and CD169+ subcapsular sinus macrophages to promote antigen subsequent B cell reactivation63. Mesothelin, another Mucin-16 binding partner, potentially home MBCs to lungs, consistent with reports of influenza-specific lung-resident MBCs in the mouse65,66. Thus, alongside changes in leukocyte migratory signaling and positional cues, the upregulation of MUC16 could suggest unique trafficking patterns that promote antigen surveillance by MBCs in the respiratory tract.

The downregulation at the highest level of significance of seven miRNAs in swMBCs suggested these genes are non-essential for differentiation or maintenance of these B cells. Such genes included ZBTB16 (ZBTB16's activity was inferred by MARINA to be significantly decreased in swMBCs as compared to NBCs) as well as six other genes. Of these, IKZF2 codes for a key TF in early B cell development67 and double negative self-reactive T cells68, while TCLI1, an AKT signal transduction molecule, is downregulated in the transition from NBCs to GC B cells69. PCDH9, is a Ca2+ dependent cell-adhesion protein70 and SPRY1 a negative regulator of cell proliferation, migration, and promoter of apoptosis71.

At a lower level of significance (p adj < 0.05), swMBCs’ gene expression profile comprised transcripts that were previously reported, including upregulated CD86, CXCR3, EB12, FAS, and ZBTB32, and downregulated KLF4 and KLF921,50,63,72. IgG+ swMBCs and IgA+ swMBCs displayed decreased expression of BACH2, a gene that skews mouse MBCs toward plasma cell differentiation, upon reactivation18. In unswMBC and swMBCs, the expression of IL-9 and IL-9R was virtually absent, suggesting that IL-9 signaling, which is putatively important for mouse MBC maintenance17, is dispensable for human MBC maintenance and, perhaps, generation. While IL-2RG was abundantly expressed in all four B cell subsets, cytokine receptor subunits that pair with the common-γ chain were differentially expressed in NBCs and swMBCs. Accordingly, human swMBCs downregulate the expression of IL-4R, IL-21R, and IFNγR1, whose signaling contributes to T-bet expression73. In the healthy humans we analyzed, TBX21 expression was absent in swMBCs and unswMBCs, consistent with the segregation of T-bet+ “atypical” MBCs among double negative (CD27−IgD−) B cells74.

As assessed by ATAC sequencing, the genes whose transcription was significantly upregulated in swMBCs were highly accessible, emphasizing the importance of such genes to the identity of swMBCs. While chromatin accessibility is necessary for transcription, it is not sufficient per se for gene expression. Gene transcription is regulated by DNA methylation, histone post-translational modifications and activators, and repressors34,35,75. Additional modulation of gene expression is mediated by epigenetic factors, mainly, non-coding RNAs, such as miRNAs and IncRNAs, at the transcriptional and post-transcriptional level. 13 of 17 gene loci, including RASSF6, TRPV3, TOX, and RORα, were significantly accessible in NBCs, despite negligible levels of the respective transcripts, suggesting mechanisms of post-transcriptional regulation, possibly mediated by ncRNAs.

In this context, we showed MIR181a and MIR181b to be key regulatory IncRNAs, likely acting as epigenetic repressors of MBC-specific gene expression. To date, there has been no investigation addressing the interplay of miRNAs and mRNAs in human MBCs. Our findings showed a great reduction of MIR181a and MIR181b expression in the transition from NBCs to unswMBCs and swMBCs, a transition that would occur after the GC stage37,76. Chromatin accessibility in promoter regions of MIR181 HGs on chromosomes 1 and 9 was reduced in total MBCs, suggesting that decreased MIR181a/MIR181b expression was mediated in part, by chromatin remodeling. Integrated analysis of the miRNA and mRNA data sets predicted MIR181 to further target 11% of the greater upregulated MBC genes (50/462 genes) and 5 of the 17 DE mRNAs upregulated in swMBC. MIR181’s role in regulating such genes was supported by MIR181 complementarity to target gene 3’UTRs and strengthened through functional evidence. Enforced expression of MIR181 in B cells significantly reduced Rassf6, Tox, Trerf1, Rora, and Trvp3 expression, while luciferase reporter assays confirmed Rassf6, Tox, Trerf1, and Trvp3 as direct and specific targets of MIR181, thereby identifying a novel epigenetic mechanism centered on MIR181 and impacting key gene expression in swMBCs.

In human tonsils, IncRNAs display a dynamic relationship between enhancer IncRNAs and genetic elements of GC B cells48. Our findings showed that human circulating IgG+ swMBCs and IgA+ swMBCs as well as unswMBCs displayed nearly identical IncRNA profiles, that discriminate them from their NBC progenitors. In swMBCs, distinct DE IncRNAs disproportionately clustered proximal to or within the IgH locus, a key element of BCR and antibody expression. Accordingly, DE IncRNAs in IgG+ MBCs vs. IgA+ MBCs were associated with either IgCy2, IgCe, or IgGa1 gene exons, thereby indicating a role for such IncRNAs in the regulation of switched BCR expression. Integration of DE mRNAs with DE miRNAs, as complemented by functional assays, identified MIR181 downregulation as a central mechanism of swMBC-specific gene expression. MIAT, a sponge of MIR18149, was upregulated in swMBCs and inversely correlated with MIR181a and MIR181b expression, suggesting that MIAT provides a regulatory mechanism that, along with chromatin accessibility, contributes to reduced MIR181 expression and promotes swMBC-specific gene expression.

Our integrative analysis of transcriptional profiles and chromatin accessibility revealed a dynamic and synergetic epigenetic landscape distinguishing MBCs from NBCs, and suggested that swMBCs emerge stochastically from the NBC pool, not from select elements or subsets of NBCs. Overall, our findings point to a core transcriptional signature, which is characteristic of and shared by swMBCs, regardless of BCR isotype, be it IgG or IgA, and distinct from that of NBCs. Beyond this core transcriptional signature, single-cell RNA-Seq may further tease apart the heterogeneity of human swMBCs, perhaps identifying discrete elements within this memory compartment77. UnswMBCs may act as intermediate elements in the transition of NBCs to swMBCs, which trend closer to swMBCs, consistent with their CD27 expression, light mutational load, Euclidean clustering, and non-coding transcriptional profile. Further, our findings, which stem from the integration of chromatin remodeling, cis-regulatory elements, and distinct expression profiles of mRNAs, miRNAs, and IncRNAs, provide evidence of a key regulatory role for the non-coding transcriptome, particularly MIR181 and MIAT, in the identity, maintenance and, likely, generation of human MBCs. Finally, by providing new insights into the transcriptional and chromatin landscape of MBCs, they open new avenues of investigation on the generation, key drivers, such as RORα, TRERF1, and TOX, identity, and role of MBCs as critical cellular elements in human health and disease.
imported into FlowJo software version 10.6.2 and biexponentially transformed prior to t-SNE analysis. t-SNE was then run on concatenated down-sampled files (100,000 cells). To analyze t-SNE maps for CD138+ and CD27+ expression in dimension 2 and the gated subsets overlaid manually based on markers of interest on the dot-plot, to show the expression of these markers on different tissular subtypes.

**mRNA, miRNA, and IncRNA sequencing.** RNA was extracted from cells using the DirectZol RNA Microprep Kit (Zymogen Research) based on cell number, according to the manufacturer’s instructions and as previously described. RNA integrity was verified using an Agilent Bioanalyzer 2100 (Agilent). Next-generation RNA-Seq for mRNA and non-coding RNA was performed by the Genome Sequencing Facility at the University of Texas Health Science Center San Antonio (UHSCSA) Greehey Children’s Cancer Research Institute. High-quality RNA was processed using an Illumina TruSeq RNA sample prep kit v2 or TruSeq small RNA Sample Prep Kit following the manufacturer’s instructions (Illumina). Clusters were generated using TruSeq Single-Read Cluster Gen. Kit v3-cBot-HS on an Illumina cBot Cluster Generation System. After quality control procedures, individual mRNA-Seq or small RNA-Seq libraries were then pooled based on their respective 6-bp index portion of the TruSeq adapters and sequenced at 50bp paired-end using an Illumina HiSeq 2500. Reads were checked by assurance (QA) pipeline and initial genome alignment (Alignment).

**Flow cytometry.** PBMCs or B cells were stained with fluorescent-conjugated mAbs in Hank’s Buffered Salt Solution plus 0.1% BSA (BSA-HBSS) for 20 min. Total PBMCs were stained with FITC-anti-human-IgG mAb (clone G18-145; BD Pharmingen), APC-anti-human-IgM mAb (clone IS1-8E10; Miltenyi Biotec), PE-anti-human-CD27 mAb (clone M-T271; Biolegend; 1:100), BV421-anti-human-IgD mAb (clone IA6-2; Biolegend; 1:100), BV421-anti-human-CD27 mAb (clone M-T271; Biolegend; 1:100), BV421-anti-human-CD138 mAb (clone M-T271; Biolegend; 1:100), BV421-anti-human-CD19 mAb (clone H1B19; Biolegend; 1:100), PE-anti-human-CD27 mAb (clone M-T271; Biolegend; 1:100), BV421-anti-human-CD27 mAb (clone M-T271; Biolegend; 1:100), BV421-anti-human-CD123 mAb (clone OX41; Biolegend; 1:100), APC-anti-human-IgM mAb (clone IS1-8E10; Miltenyi Biotec), APC-anti-human-IgA mAb (clone IS11-8E10; Miltenyi Biotec), and glycophorin A mAbs conjugated to anti-dextran Abs, which in conjunction with dextran-coated magnetic beads allow for depletion of all cell types other than B-cells. Approximately 5 × 10⁶ B cells were obtained from each subject, validated for purity (>99%) by flow cytometry and CD19+ B cells and used for cell sorting using a BD LSRII flow cytometry (BD Biosciences) or BD FACSCelesta flow cytometer (BD Biosciences) with FACSDiva software (BD Biosciences). Seventy-five percent of these B cells were stained with FITC-anti-human-IgM mAb (clone G18-145; BD Pharmingen), APC-anti-human-IgA mAb (clone IS1-8E10; Miltenyi Biotec), and CD27+ B cells were sorted into separate 1.2 ml Eppendorf tubes containing 0.5 ml of HBSS (Hank’s balanced salt solution) buffer containing 0.1% bovine serum albumin (BSA). The remaining 25% of B cells were stained with anti-IgD BV421 mAb (clone IA6-2; Biolegend) and anti-CD27 PE mAb (clone M-T271; Biolegend), and CD27+ B cells were sorted in 0.5 ml of HBSS + 0.1% BSA buffer. All sorted B cells were subjected to phenotypic analysis for confirmation of their identity and purity. CD27+ and CD27− B cells were directly isolated from PBMCs (Subjects A, D, E, F) by immunomagnetic selection using the EasySep Human Memory B Cell Isolation Kit. This is based on a proprietary two-step method using EasySep Releasable R&Dishes to positively select CD27+ cells, remove magnetic particles, and then deplete non-B cells using EasySep Dextran R&Dishes with antibody complexes. This procedure also allows for the depletion of non-B cells in the CD27− fraction in order to isolate naïve B cells in parallel.

In all seven cases, B cells were subjected to pre-post isolation phenotypic analysis using the following surface markers and fluorophores: PECy7-anti-human-CD19 mAb (clone H1B19; Biolegend; 1:100), PE-anti-human-CD27 mAb (clone M-T271; Biolegend; 1:100), BV421-anti-human-IgG mAb (clone IA6-2; Biolegend; 1:100), BV421-anti-human-IgD mAb (clone IA6-2; Biolegend; 1:100), PE-anti-human-CD27 mAb (clone M-T271; Biolegend; 1:100), BV421-anti-human-IgM mAb (clone IS1-8E10; Miltenyi Biotec; 1:100), and APC-anti-human-IgA mAb (clone IS1-8E10; Miltenyi Biotec; 1:100), and APC-Cy7-anti-IgM mAb (clone MHM-88; Biolegend; 1:100). Data were analyzed using FlowJo software version 10.6.2 (FlowJo LLC).
2100, pooled at an equimolar ratio, and sequenced on a HiSeq 3000 using 50 bp single-end chemistry. Peak calling, gene annotation, and enrichment of TF motifs were performed using DESeq2. Significant changes in TF activity were detected using DAStk and HOMER. Peak coverage and annotation were performed using HOMER24,68.

Retrovirial Mir181 construct and enforced expression. A mouse Mir181a retroviral expression vector was used for overexpression experiments. Briefly, a 270-bp mRNA gene segment containing the Mir181a mRNA hairpin was cloned from mouse chromosome 1 into MDH1-PGK-GFP expression vector. To generate the retrovirus, MDH1-PGK-GFP expression vector, encoding GFP, or MDH1-mir181a-1-1 PGK-GFP expression vector, encoding GFP and mirRNA-181a, together with the pCL-Eco retrovirus-packaging vector (Ligeneric) were used to transfect HEK293 cells by lipofection (Transfection procedures). Supernatants were harvested and used to transduce (by spinning at 400 g for 45 minutes followed by culture with LPS plus IL-4 for 72 hrs) LPS-preactivated (12 hr) C57BL/6 mouse spleen B cells, as we reported52,54. CD19+ GFP+7AAD- transduced B cells were sorted by FAC57. Expression of Rona, Tona, Trefyl, Trp53, and Rauno was induced by transduction with Mir181a expression retroviral construct was analyzed by quantitative RT-PCR78,85.

Quantitative RT-PCR (qRT-PCR) of mRNAs and miRNAs. For the quantification of mRNA transcripts, RNA was extracted. cDNA was synthesized from total RNA with the SuperScript III First-Strand Synthesis System (Invitrogen) using an oligo-dT primer. Transcript expression was measured by qRT-PCR with the appropriate primers using a QuantStudio 3 Real-Time PCR System (Thermo-Fisher) to measure SYBR Green (IQ SYBR Green Supermix, Bio-Rad Laboratories) incorporation with the following protocol: 95 °C for 15 s, 40 cycles of 94 °C for 10 s, 60 °C for 30 s, 72 °C for 30 s. Data acquisition was performed during the 72 °C extension step. Melting curve analysis was performed from 72 to 95 °C. For the quantification of mature miRNA transcripts, RNA was extracted from 0.2 to 5.0 x 10⁶ cells using mirNeasy® Mini Kit (Qiagen) and then reverse-transcribed with miScript II RT Kit (Qiagen) and the miScript Hsp6 buffer. A Bio-Rad MyQ Real-Time PCR Detection System was used to measure SYBR Green (miScript SYBR Green Kit; Qiagen) incorporation, according to manufacturer’s instructions. Mature miRNA forward primers were at 250 nM in conjunction with the Qiangen miScript Universal Primer and normalized to the expression of small nuclear/nucleolar RNAs Rnu6/RNU61/2, Snord61/SNORD61, Snord68/SNORD68, and Snord70/SNORD70. The 2−ΔΔCt method was used for data analysis of qRT-PCR experiments.

Luciferase 3′ UTR reporter assays. The 3′UTR of TOX, TREFR1, and TRPV3 mRNAs were cloned into the psi-CHECK2 miRNA Expression Reporter Vector System (Promega) downstream of the Firefly luciferase85,68. TOX, TREFR1, and TRPV3 3′UTR mutant psiCHECK-2 constructs were generated through site-directed mutagenesis of Mir181 target site seed sequences. The constructs were verified by sequencing. The psiCHECK-2 reporter also contained a Firefly luciferase reporter, which allows for normalization of Renilla luciferase activity between samples. Transfections were used to transfect human CEU Q31-B cells electroporated (500 V, 950 μF, 70 Ω) with a Gene Pulser II (Bio-Rad). Transfected CL-01 B cells were then stimulated with CD154 (1 U/ml) and IL-4 (5 mg/ml) for 24 h. The ability of endogenous Mir181 to repress reporter activity was determined by Firefly luciferase activity and normalized to Renilla luciferase activity, according to the manufacturer’s instructions, using the Luc-Pair® Duo-Luciferase HS Assay Kit (GeneCopaiea).

Bioinformatics. After sequencing, demultiplexing with CASAVA was employed to generate a Fastq file for each sample. Initial data processing was performed by the CRBl (Computational Biology and Bioinformatics Initiative) at UTSCSA. All sequencing reads were trimmed with Trim Galore, aligned with their reference genome (human Hg19 assembly). Bam files from trimming and alignment were processed using HTSeq-count to obtain counts gene in all samples87. RNA expression levels were determined using GENCODE annotation (GENCODE human v24). mRNA and IncRNA sequencing generated 12–21 million reads per sample, while smRNA sequencing generated 0.6–2.5 million reads per sample. Differential expression analysis was performed using the EdgeR package in R post-normalization88. mRNA/IncRNA was removed from downstream analysis if it did not break the threshold of at least 1 RPKM mapped reads across all sample libraries. For V(D)J analysis, the threshold was set >5 mapped reads across all sample libraries to exclude genes that are not productively rearranged. DE mRNA-between two groups was determined based on a Benjamini–Hochberg false discovery rate (FDR)-corrected threshold for statistical significance of FDR < 0.05. DE mRNA and DE IncRNA between two groups was defined based on the criterion of p < 0.05. The transcript read counts were transferred to log2(RPKM) (reads per kilobase per million reads) and used to generate a Manhattan plot, using Clstag0. Volcano plots depicting log2-fold change and raw or adjusted p-values were generated in RCircos plot ideograms using the Circos package in R.

Ingenuity pathway analysis (IPA) and gene functional classifications. Network analysis was performed using IPA (Qiagen) to investigate biological pathways associated with DE genes in the swMBC core transcriptional signature. The IPA database is the largest curated database of published findings on human biology from the public literature. Canonical pathway analysis categorizes function-specific genes present within networks and determines the significance of those genes through statistical over-representation analysis. Further, for pathways with enough informative overlap, the directionality of upregulated and downregulated genes can be used to infer the action of a gene and the directionality of inactivated or a pathway between two states (i.e., NBCs and MBCs), which is quantified independently in IPA as Z-score. Thus, in different cells or different cell states, pathways can be statistically significant as well as be significantly activated or statistically significant and be inactivated. The aggregate mRNA profile was analyzed to identify statistically over-represented pathways based on FDR-adjusted p values < 0.05 and to infer pathway activation status based on Z-score. For the biological function classifications shown in Fig. 5e, DE genes were manually curated using multiple public databases that include gene ontology annotation, as well as published literature.

MARINA. Transcriptional regulators that may underlie the differential expression profiles of human NBCs and NBCs were inferred using the MARINA algorithm90,91 based on the available human B cell interactome. The regulatory direction (positive or negative) for each transcriptional regulator-target gene pair was defined previously90 and developed into a region object for use in the VIPER package. Gene expression was ranked from most downregulated to most upregulated genes between NBCs and MBCs. MARINA quantifies differential activities of transcriptional regulators by measuring enrichment of predicted targets with DE genes at the ends of differentially regulated genes. For any given regulator, enrichment of its positively regulated targets among upregulated genes as well as enrichment of its negatively regulated targets among downregulated genes suggested higher regulator activity in NBCs. The converse was true if positively regulated targets were enriched among downregulated genes and negatively regulated targets were enriched among upregulated genes, indicating lower regulator activity in NBCs. 1000 sample permutations were used to calculate a TFs enrichment score and statistical significance (p < 0.05), which was taken to suggest that a given TF was a potential driver of differences between NBCs and MBCs.

MARINA is the only cell regulatory network model that is solely based on data sets from human B cells (performing better than standard Bayesian networks), therefore making it a good fit for our studies. The MARINA B cell interactome has limitations. It was constructed using HG-U95Av2 Affymetrix microarrays probing relative expression of 12,625 genes. Additionally, the B cell interactome was constructed using an overwhelming proportion (92.6%) of a variety of neoplastic B cell types, including “34 samples of B cell chronic lymphocytic leukemia; 68 samples of diffuse large B cell lymphomas, including cases further classified as immunoblastic or centroblastio; 27 samples of Burkitt lymphoma; 6 samples of follicular lymphoma; 9 samples of primary effusion lymphoma; 8 samples of mantle cell lymphoma; 16 samples of hairy cell leukemia; 4 cell lines derived from HCL; 9 cell lines derived from B cell lines; 5 B cell lines derived from T cell lines; 5 B cell lines derived from BCL6 expression vector”1. The data set also included a Burkitt lymphoma cell line (Ramos) treated in vitro to activate CD40 or BCR signaling and cell lines engineered to stably express BCL6 and BCL6(�PEST) mutant or to conditionally express BCL6 or MYC90,91.

Taiji. The Taiji software is an integrative multi-omics data analysis framework that integrates diverse data sets to construct regulatory networks and identify candidate driver genes. This package contains dependencies on the MAC2S, BWA, and bedGraphToBigWig packages. Briefly, Taiji identifies active regulatory elements, including promoter activities and active enhancers, defined by ATAC-seq peaks. Enhancers are then assigned to their promoters using EpiTensor predicted chromatin interactions. Transcriptional regulatory networks are constructed, by scanning regulatory elements for putative TF binding motifs supplied by the CIS-BP database, ultimately linking TFs to their target genes. Finally, the PageRank algorithm was used to assess the genome-wide influence of individual TFs, and node weights are determined by the z scores of gene expression, ranking TFs higher if they regulate more differentially expressed genes. Edge weights are set to be proportional to TFs expression levels, which filters out minimally expressed TFs. Thus, Taiji and the previous MARINA PageRank analysis was performed using raw ATAC-seq data and RNA-seq gene expression tables comparing MBCs to NBCs. Only transcription factors with a PageRank score > 0.5 in both sample groups were included to filter out less-important factors92,93.

Statistical analysis and reproducibility. All statistical analyses were performed using Excel (Microsoft), GraphPad Prism, or R software environment. Differences in RNA transcript expression were determined by EdgeR, which implements statistical methods on empirical Bayes generalized linear models (GLMs) and determines differential expression through a pairwise F-test. Spearman correlation coefficient analysis was used to measure the strength and direction of mRNA-to-mRNA, mRNA-to-miRNA, mRNA-to-lncRNA, and miRNA-to-lncRNA expression correlations. Data are derived from three or more independent experiments.
References

1. Shen, Y. et al. Distinct gene expression profiles in different B-cell compartments in human peripheral lymphoid organs. *BMC Immunol.* 5, 20 (2004).
2. Chandra, V., Bortnick, A. & Murre, C. AID targeting: old mysteries and new challenges. *Trends Immunol.* 36, 527–535 (2015).
3. Xu, Z. et al. DNA lesions and repair in immunoglobulin class switch recombination and somatic hypermutation. *Ann. N. Y. Acad. Sci.* 1050, 146–162 (2005).
4. Xu, Z. et al. 14-3-3 adaptor proteins recruit AID to 5′-AGCT-3′-rich switch regions for class switch recombination. *Nat. Struct. Mol. Biol.* 17, 1124–1135 (2010).
5. Lam, J. H. & Baumgarth, N. The multifaceted B cell response to influenza viruses. *J. Immunol.* 202, 351–359 (2019).
6. McHeyzer-Williams, L. J. & McHeyzer-Williams, M. G. Antigen-specific memory B cell development. *Annu. Rev. Immunol.* 23, 487–513 (2005).
7. Agematsu, K. et al. B cell subpopulations separated by CD27 and crucial inflammatory signals of the cytidine deaminase AID gene to induce AID expression, class-switch DNA recombination and somatic hypermutation. *Nat. Immunol.* 10, 540–550 (2009).
8. Thale, T. H. et al. Deletion of microRNA-155 reduces autoantibody responses and alleviates lupus-like disease in the Fas(−/−) mouse. *Proc. Natl Acad. Sci. USA* 110, 20194–20199 (2013).
9. Gou, M. et al. EZH2 represses the B cell transcriptional program and regulates antibody-secreting cell metabolism and antibody production. *J. Immunol.* 200, 1039–1052 (2015).
10. Schaller, C. D., Barwick, B. G., Guo, M., Bally, A. P. R. & Boss, J. M. Plasma cell differentiation is controlled by multiple cell division-coupled epigenetic programs. *Nat. Commun.* 9, 1698 (2018).
11. Zan, H. & Casali, P. Epigenetics of peripheral B-cell differentiation and the antibody response. *Front. Immunol.* 6, 631 (2015).
12. Malumbres, R. et al. Differentiation stage-specific expression of microRNAs in B lymphocytes and diffuse large B-cell lymphomas. *Blood* 113, 3754–3764 (2009).
13. Pettit, A. et al. Long noncoding RNA expression during human B-cell development. *Plos ONE* 10, e0138236 (2015).
14. Lefranc, M. P. Antibody informatics: IMGT, the International ImMunoGeneTics Information System. *Microbiol. Spectr.* 2, https://doi.org/10.1128/microbiolspec.IMM-0001-2012 (2014).
15. Brochet, X., Lefranc, M. P. & Giudicelli, V. IMGT/V-QUEST: the highly customized and integrated system for Ig and TR standardized V-J-V-D-J sequence analysis. *Nucleic Acids Res.* 36, W503–W508 (2008).
16. Seifert, M. & Kuppers, R. Human memory B cells. *Leukemia* 30, 2283–2292 (2016).
17. He, J. S. et al. IgG1 memory B cells keep the memory of IgE responses. *Nat. Commun.* 8, 641 (2017).
18. Emilie, D. et al. Oct2 enhances antibody-secreting cell differentiation through regulation of IL-5 receptor alpha chain expression on activated B cells. *J. Exp. Med.* 205, 409–421 (2008).
19. Sondergaard, E. et al. ERG controls B cell development by promoting IgH V-D-J recombination. *Cell Rep.* 29, 2756–2769.e2756 (2019).
20. Zan, H., Tat, C. & Casali, P. MicroRNAs in lupus. *Autoimmunity* 47, 272–285 (2014).
21. Mi, Y. et al. miR-181a-5p promotes the progression of gastric cancer via RASSF6-mediated MAPK signalling activation. *Cancer Lett.* 389, 11–22 (2016).
22. Chen, Y. G., Satpathy, A. T. & Chang, H. Y. Gene regulation in the immune system by long noncoding RNAs. *Nat. Immunol.* 18, 962–972 (2017).
23. Agirre, X. et al. Long non-coding RNAs discriminate the stages and gene regulatory states of human humoral immune response. *Nat. Commun.* 10, 821 (2019).
24. Zang, X. et al. MIAT promotes proliferation and hinders apoptosis by modulating miR-181b/STAT3 axis in ox-LDL-induced atherosclerosis cell models. *BioMed. Pharmacother.* 97, 1078–1085 (2018).
mononuclear cells: cell recovery and viability, population composition, and cell functionality. *Biopreserv. Biobank* **14**, 410–415 (2016).

59. Lefranc, M. P. Immunoinformatics of the V, C, and D domains: IMGT(R) definitive system for IG, TR and IgSF, MH, and MiSF. *Methods Mol. Biol.* **1184**, 59–107 (2014).

60. Pfeiffer, F. et al. Systematic evaluation of error rates and causes in short samples in next-generation sequencing. *Sci. Rep.* **8**, 10950 (2018).

61. Cortes, M. R. et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* **14**, 959–962 (2017).

62. Scharer, C. D. et al. ATAC-seq on biobanked specimens defines a unique chromatin accessibility structure in naive SLE B cells. *Sci. Rep.* **6**, 27030 (2016).

63. Zan, H. et al. Rev1 recruits wing to switch regions and enhances DNA glycosylation for immunoglobulin class switch DNA recombination. *Cell Rep.* **2**, 1220–1232 (2016).

64. Casali, P. et al. Estrogen reverses HDAC inhibitor-mediated repression of Aicda and class-switching in antibody and autoantibody responses by downregulation of miR-26a. *Front. Immunol.* **11**, 491 (2020).

65. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnol.* **37**, 907–915 (2019).

66. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).

67. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

68. Metsalu, T. & Vilo, J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res.* **43**, W566–W570 (2015).

69. Lefebvre, C. et al. A human B-cell interactome identifies MYB and FOXM1 as master regulators of proliferation in germinal centers. *Mol. Syst. Biol.* **6**, 377 (2010).

70. Basset, K. et al. Reverse engineering of regulatory networks in human B cells. *Nat. Genet.* **37**, 382–390 (2005).

71. Zhang, K., Wang, M., Zhao, Y. & Wang, W. Taji: System-level identification of key transcription factors reveals transcriptional waves in mouse embryonic development. *Sci. Adv.* **5**, eaav3262 (2019).

72. Scharer, C. D. et al. Epigenetic programming underpins B cell dysfunction in human SLE. *Nat. Immunol.* **20**, 1071–1082 (2019).

73. Tripodi I. J., Allen M. A., & Dowell R. D. Detecting differential transcription factor activity from ATAC-seq data. *Molecules* **23**, 1136 (2018).

Acknowledgements
We thank Dr. Zhao Lai (UTHSCSA Genome Sequencing Facility) for high-throughput sequencing, and Dr. Yi dong Chen and Dr.Hung I Harry Chen for help with bioinformatics. This work was supported by NIH grants AI138944, AI105813, AI079705, and the Alliance for Lupus Research Target Identification in Lupus Grant ALR 641363 to P.C.

Author contributions
J.M. designed and executed the series of experiments, analyzed results, and wrote the manuscript. A.V. and H.Z. designed and executed some experiments, analyzed some results. A.P. advised with bioinformatic analyses. P.C. conceived the project, designed experiments, and wrote the manuscript.

Competing interests
The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-019-19242-6.

Correspondence and requests for materials should be addressed to P.C.

Peer review information Nature Communications thanks Christopher Scharer 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.
