DOI:
10.1126/sciimmunol.abm9060

Document Version
Peer reviewed version

Two subsets of human marginal zone B cells resolved by global analysis of lymphoid tissues and blood

Authors: Jacqueline H.Y. Siu[1], Michael J. Pitcher[2†], Thomas J. Tull[2†], Rebekah L. Velounias[2], William Guesdon[2], Lucia Montorsi[3, 4], Krishnaa T. Mahbubani[1], Richard Ellis[5], Pawan Dhami[5], Katrina Todd[5], Ulrich D. Kadolsky[5], Michelle Kleeman[5], David P. D'Cruz[2], Kourosh Saeb-Parsy[1], Mats Bemark[6,7‡], Gavin J. Pettigrew[1‡], Jo Spencer[2‡*]

Affiliations:

[1] Department of Surgery, University of Cambridge and NIHR Cambridge Biomedical Research Centre, Cambridge, CB2 0QQ, UK.

[2] School of Immunology and Microbial Sciences, King's College London, Guy's Campus, London, SE1 9RT, UK.

[3] School of Cancer Sciences, King's College London, Guy's Campus, London, UK.

[4] Cancer Systems Biology Laboratory, Francis Crick Institute, London, UK.

[5] NIHR Guy's and St Thomas' Biomedical Research Centre, Guy's and St Thomas NHS Foundation Trust, Guy's Hospital, London SE1 9RT.

[6] Department of Microbiology and Immunology, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, SE 405 30 Gothenburg, Sweden.

[7] Department of Clinical Immunology and Transfusion Medicine, Region Västra Götaland, Sahlgrenska University Hospital, Gothenburg, Sweden.

[†] These authors contributed equally to this work

[‡] Shared senior authorship

[*] Correspondence for this submission

Jo Spencer
Department of Immunobiology
Guy's Hospital, Borough Wing
St Thomas' St

London SE1 9RT, UK

Email: jo.spencer@kcl.ac.uk
Tel: +44 207 848 9609

https://orcid.org/0000-0002-7202-2431

One Sentence Summary:

Human MZB cell subsets differ in cell surface, transcriptome, clone relatives, mutation burden, microanatomy and relevance to disease.

Word count: 7,943

Abstract:

B cells generate antibodies that are essential for immune protection, but their subgroups are poorly defined. Here we perform undirected deep profiling of B cells in matched human lymphoid tissues from deceased transplant organ donors and blood. In addition to identifying unanticipated features of tissue-based B cell differentiation, we resolve two subsets of marginal zone B (MZB) cells differing in cell surface and transcriptomic profiles, clonal relationships to other subsets, enrichment of genes in the NOTCH pathway, distribution bias within splenic marginal zone microenvironment, and immunoglobulin repertoire diversity and hypermutation frequency. Each subset is present in spleen, gut-associated lymphoid tissue, mesenteric lymph nodes, and blood. MZB cells and the lineage from which they are derived are depleted in lupus nephritis. Here we show that this depletion is of only one MZB subset. The other remains unchanged as a proportion of total B cells compared to health. Thus, it is important to factor MZB cell heterogeneity into studies of human B cell responses and pathology.

Main Text:

INTRODUCTION

B cells maintain health by generating affinity-matured and innate-like antibody responses, and as immune regulators that present antigen and secrete cytokines. B cell responses occur in lymphoid tissues that differ in fundamental microarchitecture and mechanisms of antigen acquisition (*1*). For example, gut-associated lymphoid tissue (GALT) in the Peyer's patches and appendix chronically sample predominantly particulate antigens from the gut lumen via a specialised follicle associated epithelium (FAE). This results in sustained germinal centre (GC) responses (*2*). In contrast, lymph nodes (LN) receive soluble or complexed antigens via afferent lymphatics (*3*). Apart from mesenteric lymph nodes (mLN) that are also associated with modulating immunity on mucosal surfaces and that are chronically stimulated by gut-derived antigen, most lymph nodes in healthy adults tend to be quiescent and lack GC (*4, 5*). Similarly, the spleen, which receives antigens delivered from the blood, is also generally immunologically quiescent in a healthy human (*6*).

Despite the intrinsic importance of tissue microanatomy and antigen acquisition for B cell responses, there has been a tendency to depend on studies of blood by flow cytometry, which is highly directed, to establish definitions of human B cell subsets and changes within these subsets in disease (*7*). Blood, which only accommodates a small fraction of all lymphocytes, contains a conglomerate population of cells with different migratory biases as demonstrated by differences in expression of receptors for tissue site-associated endothelial ligands (*8, 9*). Some subsets in blood, such as the circulating marginal zone B (MZB) cells, are named according to splenic microanatomy by shared phenotype, but the extent to which they are truly analogous is unclear (*10-13*). In addition, whilst the connectivity between different tissues that are bridged by migrating cells has been identified by repertoire analysis and deep phenotypic comparisons, the extent to which different B cell subsets share clonal relatives and whether they remain tissue resident or migrate between tissue sites is unknown (*14-17*).

Here we determine the characteristics and relative distribution of human B cell subsets in matched samples of spleen, mLN and appendiceal GALT using deep phenotyping, single cell transcriptomics and clonotype analysis. The strategy was to use experimentally unconnected methods to group cells and then to infer identities of the populations thus identified from published work. We observe several subsets of activated B cells in tissues and uncover evidence

for tissue-based maturation from transitional (TS) to naive B cells. We then focus our attention on two clear subsets of MZB cells in tissues and blood and that we designate MZB-1 and MZB-2, since the nature of such cells is a controversial area needing further clarification (*18*).

Cells with the phenotype $CD27^{+}IgM^{+}IgD^{+}$ have previously been demonstrated to develop from precursors expressing CD45RB via a pathway involving ligation of NOTCH2 (*9, 13, 17-22*). Despite this, these cells are often referred to as 'unswitched memory' cells, especially those in blood (*23-25*). Going forwards, it will be important to distinguish not only the unique derivation of $CD27^{+}IgM^{+}IgD^{+}$ cells but also variants of MZB cells that arise from and differentiate along different cellular pathways and that are associated differently with disease into our understanding of B cell responses and pathology.

RESULTS

B cell phenotypic variants differ in abundance across human tissues

Mass cytometry using a panel of 28 antibodies was used to investigate variation in B cells across human lymphoid tissues (Table S1). We analysed 8 matched samples of appendix, mLN, and spleen from deceased transplant organ donors with no known disease (Table S2). Following quality control (Fig. S1, A and B), data were batch normalised using the internal spleen control (Fig. S1 C) (*26, 27*). We then made a 3-way undirected comparison of the differential abundance of B cells in the tissues according to both position on the UMAP and intensity of expression of all markers except IgG, using the Cydar package (*27*). Rather than clustering, 'hyperspheres' generated by Cydar representing the cells with both similar position on the UMAP and similar median marker intensity were analysed (*28*). Only hyperspheres containing 100 cells or more were used for downstream analysis (Fig. S1 D).

The expression of markers within individual hyperspheres, including CD27, IgD, IgM, and IgA that are helpful for designation of broad B cell classifications, were visualised on the UMAP (Fig. 1 A) (*19, 21, 22, 29*). Cell counts of hyperspheres identified similar distribution between donors within spleen and appendix samples, whilst mLN samples were more variable falling between the other organs or overlapping with GALT (Fig. S1 E). The cell count for individual hyperspheres on the UMAP demonstrated similar cell densities, with slightly higher number of cells within $CD27^-$ and $CD27^+IgM^{high}$ hyperspheres than those representing class switched-cells (Fig. S1 F). A three-way comparison of differential abundance of hyperspheres between tissues was performed. Hyperspheres with significantly different abundances between tissues were predominantly $CD27^+IgM^+$, suggesting that the predominant inter-tissue variability is among unswitched, antigen-experienced cells (Fig. 1, A and B). For technical reasons, we were not able to use IgG in our analysis. However, a region of the UMAP in between IgA^+ and IgM^+ expressing cells that was $IgA^-IgM^-IgD^-$ and mostly $CD27^+$ did not contain hyperspheres representing significant abundance differences between tissues; this likely corresponded to IgG expressing B cells.

We next compared the phenotypic features of cells in the hyperspheres that differed significantly in abundance between tissues, using hierarchical clustering and k-means. This identified 5 major $CD27^+$ clusters of hyperspheres and one $CD27^-$ cluster (Fig. 1 C). The hyperspheres in each

cluster were colour coded according to the dendrogram and located on the UMAP for each tissue (Fig. 1 D).

The single hypersphere in cluster 2, that was $CD27^-$, had significantly higher abundance in spleen, contained cells expressing IgM, IgD, CCR7 and CD10, although it did not express high levels of characteristic TS B cell markers CD38 or CD24. This small cluster was most closely linked in the dendrogram to the two further clusters of hyperspheres (clusters 3 and 4) that were also both relatively more abundant in spleen, and that had the $CD27^+IgM^+IgD^+$ phenotype of MZB cells. While both phenotypically resembled MZB cells, one subset, which was designated MZB-1, had significantly higher expression of CCR7, BAFFR, CD24 and CD27 than the other, which was designated MZB-2 (Fig. 1 D). The remaining clusters of hyperspheres 1, 5 and 6, that were all most abundant in GALT and mLN, comprised IgA-expressing B cells including some CD10 expressing GC variants (cluster 1), IgM expressing GC cells (cluster 5) and IgM-only cells with the phenotype $CD27^+IgM^+IgD^-$ (cluster 6) (Fig. 1, B-D).

We were interested in understanding MZB-1 and MZB-2 groups of cells, that Cydar identified as significant areas of interest, in more detail. To quantify the abundance differences between MZB-1 and MZB-2 more directly, B cell subsets were identified and grouped into bubbles using viSNE and SPADE as shown in the schematic diagram (Fig. 1 E, S1 G). The MZB subset was subdivided based on CCR7 expression (Fig. 1 F). BAFFR, CD24, and CD27 expression agreed with the phenotypes associated with the two MZB subsets found in the unsupervised analysis (Fig. 1 G). MZB-1 cells were significantly more abundant than MZB-2 cells when data from the three tissues were combined ($p < 0.05$, ANOVA). Both MZB subsets were more abundant in the spleen than in the other tissues (Fig. 1 H).

Therefore undirected analysis of relative abundance of B cells defined by phenotype between tissues identified that most differences were in the $CD27^+$ subsets and suggested the existence of MZB heterogeneity.

B cell complexity across tissues dissected by single cell transcriptomics

To understand B cell variability in tissues in more depth, three representative sets of tissues from those analysed by CyTOF were selected for analysis by single cell RNA sequencing (scRNAseq) (Table S2). To facilitate identification of B cell subsets, sorted $CD19^+$ cells were surface labelled with Total-Seq-C antibodies prior to capture on the 10x chromium controller (Fig. S2 A and

Table S1). Gene expression, antibody detection tag (ADT) and V(D)J libraries were then prepared according to the manufacturer's instructions and sequenced. Following quality control and normalisation, data from all tissues and donors were integrated (Fig. S2, B-D). After checking that cells did not separate according to donor on a UMAP (Fig. S2 D), cell clusters were subsequently identified according to the expression of 3000 variable genes. The expression of key genes and cell surface markers were visualised on the UMAP (Fig. 2 A).

Fifteen distinct clusters were identified and named by reference to transcript profile and cell surface phenotype (Table S3 and Table S4) and visualised on the UMAP (Fig. 2 B), and summarised in Fig. 2 C. Clusters includeding TS, naive, activated naive (aNAV), GC, and class switched memory B cells were identifiable according to previous studies (*22, 30*). Markers that identified them are shown in Table S3. The aNAV subset, that is a reported precursor to antibody secreting cells derived from the extrafollicular response, was distinguished from naive and TS B cell subsets by its high expression of *CD19* and *ITGAX (CD11c)*, and low expression of *CXCR5, CD24*, and *CD38* (Fig. S2 E) (*21, 23*).

Four clusters of cells that expressed markers of activation but no classic markers of GC (*BCL6, BCL7a*), were identified and designated activated B cell (AcB) 1 to 4. AcB1 resembled GC B cells most closely, expressing *JUN, JUNB* and *FOS*. AcB2 had strong expression of the proliferation antigen *PCNA* but lacked other markers of activation. AcB3 was designated 'activated' because it expressed the long non-coding RNA *MALAT1* (Table S3). *MALAT1* is linked to class switching through its role in the alternative-non-homozygous end joining pathway (*31*). This cluster also expressed the *ZEB2* transcription factor associated with aNAV and the extrafollicular response (*7*). AcB4 had a strong signature of interferon regulated genes suggesting prior activation via this route.

Two small non-adjacent clusters were termed double negative (DN) because they lacked both CD27 and IgD (Fig. 2 C). However, they did not separate according to the classical features that discriminate between subsets of DN cells in blood (Fig. S2 F) (*23*).

Two clusters of MZB cells defined by cell surface $CD27^{+}IgM^{+}IgD^{+}$ phenotype and presence of *CD27, IGHM, IGHD, CD1C* and *PLD4* transcripts were identified (Fig. 2 A-C) (*22*).

Frequencies of somatic hypermutation of *IGHV* genes (SHM) of cells in clusters were generally consistent with expected properties according to cluster definitions by reference to previously

published work (Fig. 2 D and Fig. S3 A) (*17, 32*). The extent of SHM was highest in memory B cells and lowest in naive and TS B cells. GC B cells had undergone relatively few rounds of SHM. Clusters AcB1-3 each had undergone moderate SHM. Cells in MZB-2 had significantly lower levels of SHM than MZB-1 across all tissues (Fig. S3, B and C). The relative abundance of the B cell subsets between tissues was compared (Fig. 2, E and F). GC cells and AcB1 cells were both more abundant in the appendix than in mLN or spleen. The frequency of GC cells in these frozen samples of GALT, approximately 15%, was very similar to frequencies observed in freshly isolated GALT in a previous study (*17*). IgM-only B cells were more abundant in mLN than appendix as were AcB3 cells. The MZB-1 subset was proportionally underrepresented in appendix compared to spleen with mLN containing similar numbers as the latter. MZB-2 cells were lower in frequency than MZB-1 cells, but not significantly more or less proportionately abundant at any tissue site.

Thus this analysis identified complexity in B cell subsets in tissues and provided further evidence for the existence of MZB subsets.

Developmental relationships between B cell subsets in tissues

We investigated the relationships between cells in subsets using BCR and RNA velocity analysis. Using the Immcantation pipeline for analysis of adaptive immune receptor repertoire in heavy chain genes, clones of two or more cells were identified with the following criteria: identical IGH V-J gene usage, identical CDR3 length, minimum sequence similarity in CDR3 based on the hamming distance between each sequence and its nearest neighbour (Fig. S4, A and B) (*33*). Clonal abundance (cells in clones as a proportion of the entire repertoire) was greater in spleen than other tissues (Fig. S4 C). Of the antibody isotypes IgG tended to have greater clonal abundance (Fig. S4 C). Clonal diversity between isotypes, tissues, and subsets were visualised using a diversity profile curve; spleen and IgG had lower diversity compared to other tissues and isotypes (Fig. S4 D). There was no significant CDR3 length difference across isotypes, tissues, or subsets (Fig. S4 E).

A total of 29614 cells with both transcriptomic and IGH sequences were observed across all subsets, tissues, and donors including 4482 clones (groups of two or more related cells) (Table S5). Of the 4482 clones, 4025 were observed within a single tissue and 457 were observed across two or three tissue sites. We considered the possibility that expectation is shaped by data from

bulk sequencing protocols that pool identical sequences in a sample to a single sequence. We therefore looked at the frequencies of cells in clones with intraclonal identity or intraclonal variation. Of the 3696 clones observed in single tissue sites, 3393 (91.8%) contained identical clone members and would therefore have been pooled as single sequences in bulk sequencing data. Of 457 clones with members in more than one tissue, 116 (25.4%) contained identical clone members. The B cell subset containing cells with greatest tendency to be found in local clones in each tissue was AcB2, that was designated as an activated subset because of its expression proliferation antigen, *PCNA* (Fig. 3 A).

We used a correlation matrix to evaluate the frequency of clone sharing between different B cell subsets, to identify potential developmental relationships between subsets within tissue sites (Fig. 3 B). Significant clone sharing was observed between TS and naive B cells in both appendix and spleen. Clone members were unmutated and only observed locally within the same tissue and not shared between different tissues. When both heavy and light chain rearrangements were captured, these were shared by clone members (Figure S5A and B). Thus, this indicates proliferation once heavy and light chain genes have rearranged, rather than proliferation at the pre-B stage, and is consistent with the hypothesis that TS mature to naive B cells in GALT and spleen involving cell division (Fig. S5, A and B). The lack of shared clones between tissues suggests that clone members are co-clustered spatially and that clone members become diluted once they leave the tissue site. The existence of such local clones involving immature cells in tissues is consistent with a low level of proliferation at these stages as observed previously (*9*, *34*).

In the appendix, MZB-1 and IgM-only B cell subsets were significantly clonally related, as were GC cells and DN-B cells, that locate adjacent to each other on the UMAP plot. In spleen, significant clonal relationships were observed between the DN-B and AcB1 subsets, that were also proximal in the UMAP. No significant clonal relationships between subsets were observed in mLN (Fig. 3 B).

Circos plots were used to visualise clone sharing between subsets across tissues. The MZB clones most likely spanning tissues were those containing MZB-1, though it should be noted that the tendency for a clone to span tissues is related to both the properties of the cells and the sizes of the clones that will both contribute to the probability of identifying them. In two of the three

donors, clones containing MZB-1 spanned all three tissues. Clones containing widely disseminated memory cells and IgM-only B cells were also observed (Fig. 3 C).

Potential developmental trajectories were analysed by RNA velocity. The value and restrictions of this method are debated (*35*). However, it provides a useful tool to enable speculation on developmental continuity between clusters of cells. The analysis largely supported the findings above (Fig. 3 D) (*36*). Transcriptomic progression from TS to naive B cells was observed in GALT and spleen. Developmental connections between GC, AcB1 and DN-B were apparent. AcB2 that has the greatest tendency to contain clonally related cells and has a proliferation signature, potentially reflecting an extrafollicular response, appeared to have connection to aNAV supporting other studies (*23*). There was evidence of memory B cell association with GC in GALT (*32*). MZB-1 and MZB-2 showed no notable connection (Fig. 3 D).

The cluster AcB3 characterised by the expression of *MALAT1* and *ZEB2* appeared to be a common 'destination' for developmental trajectories, despite no evidence on the repertoire analysis of developmental associations with other subsets, possibly representing a transcriptomic state, rather than a stage in progression of specific subsets.

This analysis suggests tissue based homeostatic proliferation during TS to naïve B cells maturation. No evidence of clonal relationship between MZB-1 and MZB-2 was observed.

Human MZB subsets have different derivation and occupy different microanatomical space.

The above analyses indicate the presence in humans of two subsets of MZB cells that so far differ in their: phenotype (Fig. 1 C); SHM (Fig. 2 D, Fig. S3 B); transcriptome (Fig. 2, B and C); their clonal relationships with other clusters (Fig. 3 B); and the tendency for clone members to be identified at multiple sites (Fig. 3, A and C) (though it should be noted that tendency for clone members to be observed at multiple sites is influenced by population and clone size). No suggestion of a developmental link between MZB-1 and MZB-2 was observed by RNA velocity analysis (Fig. 3D).

Differentially expressed genes between MZB-1 and MZB-2 subsets are illustrated in a volcano plot (Fig. 4 A). Genes selectively expressed by MZB-2 cells tended to be associated with cellular activation such as the HLA alleles, *CD83* and *MIF* (*37*, *38*). In addition, MZB-2 cells selectively expressed the RNA helicase *DDX21*, which has been implicated in recognition of viral RNA (*39*). Interestingly, *CD83* is consistently more abundantly expressed in the appendix in both

MZB-1 and MZB-2 subsets. The preferential expression of *CD37* by MZB-1 cells is consistent with its potential higher mobility (Fig. 4 A and B) (*40*).

MZB development in mice and humans involves ligation of NOTCH2 on B cells by delta like-1 (*19, 41, 42*). We therefore sought evidence of enrichment of genes in the NOTCH pathway in the MZB subsets and found it in the MZB-2 subset only (Fig. 4 C).

The term 'marginal zone' originally referred to the zone of B cells on the periphery of the white pulp and in the perifollicular region in human GALT (*12, 43*). We therefore asked if the subset complexity in MZB cells we observe by analysis of cells in suspension was also present in the splenic MZ by imaging mass cytometry (*44*).

The GC, composed of B cells, T cells, and macrophages, had areas of high B cell proliferation (Fig. 5 A). The GC was surrounded by a mantle zone of $IgD^{+}CD27^{-}$ naive B cells (Fig. 5 B). The mantle was in turn surrounded by $IgD^{+/-}CD1c^{+}CD27^{+}$ cells, corresponding phenotypically to MZB cells identified in cell suspension (Fig. 5, B-D and Fig. S6 A). Visualisation of masked cells highlighted that the most peripheral B cell subset in the MZ was $CD1c^{-}$, and likely corresponded to memory B cells, that also had a peripheral distribution in GALT in an earlier study (Fig. 5, B and D) (*17*). DDX21 expression, a feature of MZB-2 cell subset (Fig. 4 B), was significantly higher in $CD1c^{+}$ than $CD1c^{-}$ cells in the MZ (Fig. 5 E).
Consistent with expression of DDX21 by an MZB subset, we observed punctate nuclear and cytoplasmic DDX21 staining in a proportion of MZB cells (Fig. 5 F). We identified a distribution bias in DDX21 expressing $CD1c^{+}$ cells in the MZ, with these being closer to the centre than cells lacking DDX21. Relative to the distance between a cell and the centre point (average X and Y position of all cells per image), the proportion of high and low DDX21 expression in $CD1c^{+}$ cells was compared (Fig. 5, G-K and Fig. S6 B). If there was no spatial bias along the radial axis, the expected slope of the weighted linear regression would be zero. In other words, based on the linear regression slopes, DDX21 had a radial-biased distribution whereas DNA did not (Fig. 5 J). Visualisation of DDX21 high compared to DDX21 low MZB masked cells illustrated this radial-biased distribution (Fig. 5 K).

In summary only MZB-2 were enriched in genes of the NOTCH pathway. Microanatomically, $CD1c^{+}$ MZB and memory B cells occupied different locations within the MZ. In addition, of the

intermingled CD1c expressing MZB subsets, DDX21^{+} MZB-2 cells located significantly closer to the GC.

Analysis of MZB subsets in blood

We and others have previously studied MZB cells in blood (*11, 22, 30, 45*), and identified a differentiation pathway from IgM^{hi} TS cells that is reduced in severe systemic lupus erythematosus (SLE) (*22*). We therefore explored if either or both of the MZB subsets identified here are analogous to previously studied circulating MZB cells.

To address this, we analysed scRNA sequencing data acquired from sorted blood CD19^{+} B cells from 3 healthy control donors and 3 patients with severe SLE (Table S2). Following normalisation and quality control, data from all 6 donors were pooled (Fig. S7, A-C), and clustered according to 2000 variable genes. The 10 clusters identified by gene and surface protein expression (Table S6) included 2 MZB clusters with the surface phenotype CD27^{+}IgM^{+}IgD^{+} and expression of *CD1C* and *PLD4* that are hallmarks of MZB cells (Fig. 6 A-C and Fig S7 D) (*22*). We named these clusters MZB-1 and MZB-2 by comparison with MZB-1 and MZB-2 identified in tissues above.

By analysis of the VDJ libraries (Fig. S8, A and B), we identified that the frequency of SHM in each circulating subset was consistent with the corresponding subset in tissue (Fig. 2 D, Fig. 6 D, Fig. S3 A, and Fig. S8 C). The MZB-1 subset had higher SHM than MZB-2 subset in blood (Fig. S8 C).

In both healthy and severe SLE patients, there was a significant tendency for MZB-1 cells to have clonal relationships with IgM-only B cells as in tissues (Fig. 6 E). In contrast, MZB-2 showed a significant tendency to be clonally related to aNAV and DN B cells in health. A clonal link between aNAV and DN B cells was observed in health and severe SLE, consistent with published work (*22, 23*) (Fig. 6 E).

To cross reference the independent datasets acquired by mass cytometry and single cell transcriptomic analyses of tissues and blood, we compared gene expression of those proteins that distinguished the two subsets of tissue MZB in the original mass cytometry analysis (Fig. 1 C); this gene set also discriminated MZB-1 and MZB-2 in tissues and blood (Fig. 6, G and H).

We recently described that MZB cells can derive from IgM^{hi} TS B cells, a feature that is associated with gut homing and that is depleted in severe lupus (*22*). We therefore asked how this pathway relates to MZB-1 and MZB-2. Whilst both subsets of MZB in blood were IgM^{hi}, the MZB-1 subset was visibly linked to the naive B cell pool by an IgM^{hi} bridge as described previously (Fig. 6 F). In addition, expression of gut-homing ß7 integrin was associated with MZB-1 in both blood and tissues (Fig. 6, I and J).

We observed that MZB-1 cells tend to be more abundant than MZB-2 cells in the blood in health, but that the frequency of the MZB-1 subset was consistently reduced in severe SLE (Fig. 6 K). As in previous studies, DN, aNAV and class switched memory B cells were increased in frequency in severe SLE (*22*, *23*) (Fig. 6 K).

We then used CCR7 to distinguish between MZB-1 and MZB-2 by flow cytometry (Fig. S9), using a panel comprising antibodies to CD19, CD27, IgD, IgM, CD1c, CCR7 and β7 integrin. Although flow cytometry does not have the power to resolve two clearly distinct entities, as sharply as the single cell approaches developed in the paper, it was clear that the frequency of MZB-1 cells was significantly greater than MZB-2 in health, but not in severe SLE (Fig. 6 L). MZB-1 was associated with greater ß7 integrin expression than MZB-2 in health ($p < 0.01$, paired t-test) (Fig. 6 M), consistent with Fig. 6 I.

In summary, MZB-1 and MZB-2 are identifiable in blood and share features of these subsets in tissues. MZB-1 but not MZB-2 cells are depleted from the blood in lupus nephritis.

DISCUSSION

Here we describe subsets of human B cells in matched appendix, mLN and spleen samples from deceased transplant donors using multiparameter unsupervised methods. In addition to expected and novel tissue-associated subsets of activated B cells, we observed two subsets of MZB cells, both of which are found at highest frequency in the spleen, within the MZ. MZB-1 are consistently more abundant than MZB-2 across all tissues independently and blood by CyTOF, scRNAseq and flow cytometry. The MZB-2 subset has a more diverse repertoire and fewer V region mutations.

Our recent analysis of circulating human B cell subsets identified IgM^{hi} intermediary stages in the development of human MZB cells from TS B cells (*22*). Here we refine those findings, by showing that the IgM^{hi} developmental pathway is relevant to the MZB-1 subset. In comparison,

the MZB-2 subset has a signature of NOTCH related genes, suggesting that this subset matures by NOTCH ligation, as described previously in humans and mice (*19, 41*). Our findings suggest that the two MZB subsets are developmentally unrelated because whilst they each showed clonal relationships to cells in other clusters that is similar between tissues and blood, they were not significantly clonally related to each other in any dataset. RNA velocity analysis identified no evidence of developmental links between MZB subsets.

The MZ is traditionally associated with innate like immune responses to bacterial antigens with repeating subunit structures, so-called T-independent type 2 antigens. Children less than 2 years of age (when MZB cells are poorly developed and consist of non-clonal cells with low mutation loads) and also individuals who were splenectomised early in life are particularly at risk of infections with bacterial pathogens (*46-48*). Patients with severe SLE are also susceptible to infection with pneumococci, suggesting that it is principally the MZB-1 population, which we show here to be depleted in severe SLE, that provides protection (*49*). In contrast, the MZB-2 subset, expressed the RNA helicase DDX21. DDX21, that is associated with regulation of RNA (*50*), is also associated with anti-viral immunity (*39, 51*). When members of this subset were detected in blood in health or lupus nephritis, they exhibited clonal relationships with the aNAV and DN subsets. aNAV and DN cells have been linked to the extrafollicular response in lupus and also SARS-CoV-2 infection (*23, 52*).
Mouse MZB cell development is dependent on NOTCH2 (*41, 42*). Mouse MZB cells are generally a naïve subset with low frequencies of somatic hypermutation, suggesting that human MZB-2 may be analogous to the mouse MZB subset (*53*). Heterogeneity within the mouse MZB subset has been described; for instance, NOTCH2 mediated plasticity in mouse MZB lineage (*54-56*). Thus, the differences between human and mouse MZB cells that have been noted for many years may relate to differences in relative proportions of MZB cell subsets derived through different pathways (*10*).

Whereas the importance of IgA in mucosal immunology is well defined, the role of IgM is less clear. There are both memory and clonally related plasma cells that express only IgM in human gut, suggesting that IgM-only cells are not merely recently activated naive B cells (*14*). The high number of mutations in the IgM-only B cell subset supports this concept. We have previously identified that $CD27^+IgM^+IgD^+$ cells in GALT can be clonally related to IgM-only cells (*17*). However, IgM-only cells, class switched memory cells and $CD27^+IgM^+IgD^+$ cells were not

observed in the same clonal families in this previous study (*17*). This suggests that IgM-only cells in tissues are a heterogenous population including both unswitched memory and MZB subtypes (*17*). The separate nature of MZB and classical memory is also supported by their different microanatomical distributions in GALT in a previous study (*17*) and in the spleen here. Thus the B cell subset located most peripherally in the marginal zone appears to be memory cells rather than MZB themselves. Here we identify IgM-only cells as a distinct transcriptome driven cluster, with MZB-1 relatives and clone members that are disseminated between the three tissues. Of note, many IgM-only cells were clustered alongside IgA and IgG class switched memory cells, consistent with the concept that the IgM-only subset contains distinct, albeit yet uncharacterised, functional groups.

The GC cells predominatly used the IgM isotype and had high relative abundance of cells with low mutation frequencies. These features suggests that the role of GALT GCs in maturation of the naive repertoire and in sustaining IgM only responses may have been underestimated (*57*).

We observed small clones of TS and naive B cells in appendix and spleen. These could only be observed within single sites of lymphoid tissue suggesting that related cells occur within the same piece of tissue and that when they join the systemic circulating pool, clone members become separated and no longer detectable due to dilution. Detection of clones at multiple sites reflects not only the tendency of cells to migrate but also the size of the clone and therefore the probability of identifying clone members. When light chains were captured in addition to the heavy chains by which clones were identified, these were identical between clone members supporting the concept of local proliferation of cells that were already fully mature in terms of BCR rearrangement. Activated TS cells have been observed previously in GALT (*9*). Selection occurs as B cells mature from TS to naive follicular and MZ cells, and it is possible that checkpoints in appendix and spleen involving cell division are involved in shaping the naive B cell repertoire (*58*).

We identified 4 subsets of AcB that are likely to represent different activation states or stages of differentiation. Only one, AcB3, was also identifiable in blood. AcB1 is transcriptionally similar to GC B cells; however, it does not express classic GC markers such as *BCL6*. In addition, while clonally related to GC cells, the AcB1 subset has higher mutation frequency, suggesting that it is a more terminally differentiated subset. Thus, the AcB1 subset could reflect cells that have just

exited the GC, possibly on their way to undergo class switch recombination. AcB2 cells, that had the greatest tendency to be observed as clones, also expressed the proliferation antigen *PCNA*. This subset may represent extrafollicular proliferations of B cells. This subset was associated with aNAV cells that are components of extrafollicular responses in other studies (*23*).

Malignancies of marginal zone B cells comprise different histogenetic types including marginal zone B cell lymphoma of mucosa-associated lymphoid tissue (MALT lymphoma) (*59*) and splenic marginal zone lymphoma (SMZL) (*60*). It is possible that the two subsets of MZB cells that we describe could represent their benign analogues. MZB-1 could be benign analogues of MALT lymphoma. This malignancy can be driven by bacterial infections (*59*). These tumours tend to express IgM and can circulate from GALT via blood to the spleen where they tend to occupy the MZ (*61*). In contrast MZB-2 could be benign analogues of SMZL. Cases of SMZL tend to localise to the spleen but not other secondary lymphoid tissues; and they may arise through translocations involving the NOTCH pathway, activation of which is observed here in MZB-2 only (*60*). Understanding benign analogues of lymphoma subtypes may help future identification of drivers and thus potential therapeutic pathway inhibitors.

It is important to consider the caveats of the work we present. It should be borne in mind that this study is a deep observational study. Whilst it is possible to extrapoloate function and developmental origin from such datasets, it will be important to carry out functional validation assays of such inferred data in the future. Most of the tissues analysed in the study were frozen samples of cells from tissue and blood donors. Whilst this allows us to carry out experiments with precious samples, validation using unprocessed cells will be important in the future. The blood studied here was exclusively from female donors because of the strong female gender bias of lupus nephritis. It will be important to validate this work across samples from donors representing the diversity of gender and ethnicity of the human population in the future.

Overall, the deep analysis of B cells in tissues that we present, which combined undirected methods of grouping similar cells with knowledge and reference-based subset alignments to the groups identified, provides a more accurate vision of tissue-based subsets and their interrelatedness within and between tissues than was previously available. There were organ-specific expression patterns within subsets, demonstrating that the local microarchitecture and milieu will determine cellular functions. The human MZB cell subset is considered to have

putative innate immune function; our demonstration of the complexities within this subset highlights the requirement to better understand how its distinct microanatomical features relate to function and disease.

MATERIALS AND METHODS

Study design

This study sought to understand the diverse features of and interrelationships between B cells in human tissues, and their counterparts in blood in health and the severe autoimmune disease lupus nephritis. A series of deep observational methods were used including interrogation of cell surface phenotype by mass cytometry, of transcriptome at the single cell level and of imaging mass cytometry to place cells defined by complex marker sets in histological context.

Experimental subject details

Human tissue was obtained from deceased adult transplant organ donors with research ethics committee (REC) approval and informed consent from the donor family (reference 15/EE/0152, East of England Cambridge South Research Ethics Committee). Studies of human tissues were approved by London, Camberwell St Giles Research Ethics Committee (study 11/LO/1274 Immunology of the intestine; features associated with autoimmunity). (Table S2 A)

Blood was obtained from SLE patients and healthy controls with informed consent and REC approval (REC reference 11/LO/1433: Immune regulation in autoimmune rheumatic disease, London–City Road & Hampstead Research Ethics Committee). For donor details and diagnostic criteria, see Table S2 A, B.

Sample processing

Matched appendix, mLN, and spleen were collected from deceased transplant organ donors and stored in University of Wisconsin solution at 4°C (*62*). All tissue preparation and lymphocyte isolation procedures were performed with RPMI-1640 containing heat inactivated 10% FCS, 2 mM L-glutamine, 100 IU/mL penicillin and 100 μg/mL streptomycin (RPMI-P/S) unless stated otherwise.

Cell suspensions from spleen were taken from approximately $2cm^3$ pieces of tissue. Whole mLN were taken and cell suspensions prepared from the entire node. Cell suspensions from appendix used half of the appendix. Appendix, cut into 1-2mm pieces, was incubated at 37°C for 30 minutes in medium with 1 mg/mL collagenase IV (Sigma-Aldrich) and 1 mg/mL DNAse I (Roche). Lightly digested appendix as well as fresh mLN were then teased through a 70 μm nylon cell strainer, washed with medium, and re-suspended for cryopreservation. Small sections

of spleen were placed in gentleMACS C tubes (Miltenyi) topped up with PBS + 2% FCS where the gentleMACS setting B ran three times in the gentleMACS Dissociator. The solution was then filtered through a 70 μm nylon cell strainer (diluted with PBS as necessary). Spleen solution was layered onto Lymphoprep (Stemcell Technologies) according to manufacturer's instructions and then centrifuged at 800 x g for 20 minutes at 4°C with no brakes. The buffy coat was collected, washed with medium, centrifuged, and re-suspended in red blood cell lysis buffer (0.17M ammonium chloride) for 7 minutes at room temperature to lyse any remaining erythrocytes. Cells from each tissue were cryopreserved in freezing medium of FCS + 10% dimethyl sulfoxide (DMSO) in aliquots of $1x10^7$ cells.

Blood samples were diluted 1:1 in RPMI-P/S, and then layered onto Ficoll before centrifugation. The buffy coat layer was collected, and cells were washed. PBMCs were cryopreserved in FCS + 10% DMSO in aliquots of 1×10^7 cells.

Mass cytometry

Frozen samples were thawed in a 37°C water bath and washed in 2 mL of RPMI-1640 + 10% FBS, 0.1 mg/mL DNase I (Roche), and transferred to a larger volume of 10 mL and stained as described previously (*22*).

Pre-processing and normalisation of mass cytometry data

FCS files were normalised using bead standards and the Normalizer program developed by Nolan's group (v0.3) (*26*). The result of the bead normalization is visualized in Fig. S1 A. The mass cytometry data was initially manually gated in Cytobank (https://mrc.cytobank.org/) using the gating strategy shown in Fig. S1 B to identify live $CD19^+$ B cells for downstream analysis. For batch normalisation purposes, each paired tissue set was run in the same batch with an internal biological sample from the same spleen as reference. Internal controls between batches were normalised by transforming the pooled intensity distribution of all batches towards a reference distribution (*27*). The range normalisation method scaled the marker intensities so that the distribution range was the same for each batch. By comparing and matching the 1st and 99th percentiles of each batch to the reference distribution, a linear function was then defined and applied to all markers of all samples during normalisation. In other words, overall distribution of each batch was adjusted based on the minor distribution shifts of the internal control for that batch. The effects of normalisation on the internal controls are visualised in Figure S1 C.

Differential abundance analysis of mass cytometry data

After the data were normalised, cells were allocated into hyperspheres, and then tested for differential abundance between tissues for each hypersphere while controlling for false discovery rate (*27*). Hyperspheres represent cells in regions of the data with similar marker phenotype. Importantly, unlike clustering, a cell can be allocated into more than one hypersphere. Low-abundance hyperspheres with average counts below 100 were filtered out (Fig. S1 D). Multidimensional scaling plot was used to determine if abundance differences were attributed to different tissue type and not biological variation (Fig. S1 E). As developed previously by Lun *et al*., empirical Bayes method was used to allow hypersphere-specific variation estimates even when replicate numbers were small(*27, 63*).

To compare the cell counts between samples, significant differences in cell abundance between conditions were tested. The null hypothesis of no change in the average counts of cells between conditions for each hyperspheres was tested using negative binomial generalised linear model (NB GLMs) implementation in edgeR package (*63*). To control false discovery rates, spatial false discovery rate of 5% was used which considers hypersphere densities.

Manual gating of mass cytometry data using SPADE

viSNE was run on with 1,500,000 total events using all B cell markers except for CD19 and CD45 with 10000 iterations, perplexity 50, and theta 0.5 on the Cytobank platform. SPADE was then run on the viSNE coordinates, and B cell subsets were identified and grouped into bubbles (Fig. S1 G).

Cell sorting and CITE-seq antibody staining

Cryopreserved samples were thawed in a 37°C water bath and washed in RPMI with penicillin and streptomycin. After cell counting, 3 million cells per sample re-suspended in PBS with 1% BSA were transferred to Eppendorf LoBind Microcentrifuge tubes and washed. Cells were processed and stained as described previously with antibodies in Table S2 (*22*). before sorted $CD19^+$ cells were loaded onto the 10X Chromium controller.

Imaging mass cytometry

A panel containing 10 metal-tagged antibodies (Table S1) was designed to identify and characterize immune populations expected to be present in the splenic white pulp. Formalin Fixed Paraffin Embedded (FFPE) samples of human spleen from three anonymous healthy adult donors, were cut in 4 μm-thick sections. Briefly, sections were deparaffinized, rehydrated and

subjected to antigen retrieval using a pressure cooker. Tissues were then blocked and incubated with the mix of metal-conjugated antibodies contained in the panel overnight at 4°C and subsequently incubated with the DNA intercalator Iridium Cell-ID™ Intercalator-Ir (Fluidigm) before being air-dried. Slides were then inserted into the Hyperion Imaging System and photographed to aid region selection. A total of 10 regions of approximatively 1mm^2 were selected to contain at least one identifiable GC and subsequently laser-ablated at 200 Hz frequency at 1μm/pixel resolution.

Imaging mass cytometry analysis

Single cell segmentation was performed prior to single cell protein expression analysis. Using CellProfiler, DNA staining was used to identify cell nuclei and B cells were detected by masking nuclei which also contained CD20 (*64*). Resulting images showed morphologically appropriate cell boundaries and centres. Pixel-level composite images were created using histoCAT (*65*). Cytomapper was used to overlay the single cell metadata onto the cell segmentation masks (*66*). Marker expressions were scaled by arcsinh with a factor of 0.1 for visualisation purposes. The CD1c threshold of 1.2 was used. Rather than manually designating the centre of the GC, the centre point was determined by calculating the average X and Y position of every CD1c$^+$ cell for each image. Images with multiple follicles were excluded from downstream analysis (SPL_5, SPL_7). The distance between the centre point and each CD1c$^+$ cell was calculated. This distance distribution was scaled such that the overall radius of each image was 1. In the histogram, counts were scaled by image such that the maximum count was 1 and the bins for the spatial distances were 0.05. Marker expression for each cell was classified as high or low if they were above or below the median marker expression per image. Linear regression was weighted by the cell counts in each bin. Paired and one sample t-tests were performed and illustrated using ggpubr (*67*).

Single cell RNA sequencing library preparation

Sorted CD19$^+$ cell populations from three donors were loaded onto a 10X Genomics Chromium Controller and the libraries (5' gene expression, VDJ, ADT) were prepared according to manufacturer's guidelines (Table S2). The Illumina HiSeq 2500 High Output platform was used for sequencing (30-100-100 sequencing configuration). Transcript alignment and generation of feature-barcode matrices for downstream analysis were performed using the 10X Genomics Cell Ranger workflow.

Single cell transcriptome analysis of tissues

Using the Seurat R package (Version 3.2.2), sorted $CD19^+$ cells with high mitochondrial transcripts, low/high number of unique genes per cell, and low/high total RNA transcripts were filtered out (*68*). The threshold used was 3 times the mean absolute deviation of each sample (Fig. S2, B-C). B cells were isolated based on the expression of B cell specific genes (*CD79A, CD79B, CD19, MS4A1*) and absence of T cell specific genes (*CD2, CD3D, CD3E, CD3G, CD4, CD8B, CD7*), and *TCR* genes. Genes that were only expressed in 10 cells or fewer in the entire dataset were filtered out. In addition, *IGHV* and *TCR* genes were removed prior to downstream analysis. The data was transformed and integrated using the SCTransform and Integration workflow (Fig. S2 D). For more efficient clustering and dimension reduction, PCA was performed on the top 3000 transcriptomic variable features, and the top 25 principal components were used for downstream analysis. Cells were clustered using the standard Seurat clustering with resolution of 1.3. Genes were deemed significantly differentially expressed using the Wilcoxon rank sum test with a log fold change threshold of 0.25 and expressed in 10% of either population (Seurat::FindClusters function). Clusters with similar differentially expressed gene lists were combined, and all the clusters were annotated using the gene lists shown in Table S3. Gene set enrichment analysis was done using the Broad Institute's GSEA software (*70, 71*). All relevant NOTCH gene sets were downloaded from the Molecular Signatures Database (MSigDB) and corrected for our dataset's background. Spliced and unspliced RNA count matrices were obtained by scVelo (*72*). RNA velocity was calculated using scVelo (*36*).

Single cell transcriptome analysis of blood

A similar workflow as described above was performed on sorted $CD19^+$ blood from healthy and SLE patients with the following modifications. Genes that were only expressed in 3 cells or fewer in the entire dataset were filtered out. The standard Seurat integration workflow was used to normalise and integrated the samples (Fig. S7, A-C). PCA was performed on the top 2000 variable features, and the top 20 principal components were used for downstream analysis. Cells were clustered using the standard Seurat graph-based approach with resolution of 1.8. Clusters were annotated using differentially expressed gene lists (Table S6) where genes were deemed significantly differentially expressed as described above.

Single cell BCR analysis

The BCR repertoire analysis used the Immcantation framework (Version 4.0) (*33*). V, D, and J genes were assigned using IgBLAST. Non-productive sequences were removed. The following criteria identified clones: identical IGH V-J gene usage, identical CDR3 junctional sequence length, and CDR3 junctional sequences has a minimum of % sequence similarity based on the hamming distance between each sequence and its nearest neighbour. Clonal threshold at 0.15 or 0.1 was determined from the Hamming distance (Fig. S4, B and Fig. S8 B). To normalize the donor effect when looking at the ratio of detected disseminated clones to those found at one site only, the ratio in each subset was divided by the average dissemination ratio in that donor. To determine the clonal correlation within and between tissues (subset compositions on a per cell basis within the clone), the Spearman correlations per donor along with the pairwise p-values were calculated. These p-values were then corrected for multiple interference using Holm's method in the R package RcmdrMisc (*73*). Somatic hypermutation was calculated using only the heavy chain and productive rearrangements where the entire sequence is compared to the germline sequences to identify R and S mutations.

Flow cytometry

Cryopreserved PBMCs from healthy donors (n=10) and lupus nephritis patients (n=10) were thawed, washed and rested in RPMI with DNAse and incubated for 45 minutes at 37°C before staining on ice for 15 min with pre-titrated antibodies. All antibodies used were mouse anti-human, except Beta7 FITC that was rat anti-human. Data was acquired on a BD Fortessa flow cytometer. FCS files were analysed with FlowJo Version 10.8.1 (Treestar) as described in Fig. S9.

Statistics, data analysis, and visualisation

All statistics, data analysis, and visualisation were done in R (3.5.2, 4.0.2) unless stated otherwise (*74*). Boxplots are in the style of Tukey where lower and upper hinges correspond to 25^{th} and 75^{th} percentiles, and whiskers extends from hinge to largest value no further than 1.5 times the inter-quartile range from the hinge. Data beyond the end of the whiskers are considered outlying points, and plotted individually. We used paired t-test and paired Wilcoxon test to compare two groups with paired variables that are normally and not normally distributed respectively. We used two-way ANOVA to analyse experiments with multiple groups. Outliers were assessed by box plot method, normality was assessed using Shapiro-Wilk’s normality test

and homogeneity of variances was assessed by Levene's test. There were no extreme outliers (points 3 times the interquartile range), residuals were normally distributed ($p > 0.05$) and there was homogeneity of variances ($p > 0.05$). Consequently, an analysis of simple main effects for subsets was performed with statistical significance receiving a TukeyHSD adjustment. All pairwise comparisons (estimated marginal means) were analysed between the different sample types organised by subsets, and p-values were adjusted using Bonferroni. Differentially expressed genes p values were calculated using Wilcoxon test with Bonferroni correction for multiple comparisons. Statistically significant correlations were identified using Spearman correlation along with pairwise p-values that were then corrected using Holm's method. Significance is indicated as follows: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; **** $p<0.0001$.

Uniform Manifold Approximation and Projection (UMAP) was implemented using uwot package (*75*). Heatmaps were visualised using ComplexHeatmap package (*76*). Circos plots were built using circlize package (*77*). Volcano plot was drawn using EnhancedVolcano package(*78*). Correlation plots were drawn using corrplot package (*79*). Other single cell sequencing figures were drawn using Seurat (3.2.2), ggplot2, and ggpubr packages (*67, 68, 80*).

Supplementary Materials

Fig. S1: Quality control, preliminary gating, batch normalisation, and preliminary analysis of hyperspheres for mass cytometry

Fig. S2: Sort strategy, 10x genomics workflow & quality control, batch integration for single cell transcriptomics, and key gene expressions

Fig. S3: ANOVA and Tukey post-hoc comparison of mutation frequencies between subsets in tissues

Fig. S4: BCR workflow for single cell analysis, clonal clustering threshold, clonal characteristics

Fig. S5: Examples of clones involving naïve and transitional B cells

Fig. S6: Masking of $CD1c^+$ B cells and centre point calculation for image mass cytometry analysis

Fig. S7: 10x genomics workflow, integration of B cells in blood, ADT surface protein expression, and mutation frequency statistics

Fig. S8: BCR workflow for single cell analysis, clonal clustering threshold, mutation frequency statistics in PBMC.

Fig. S9: Flow cytometry workflow for healthy and SLE PBMC

Table S1: List of reagents

Table S2: Donor details

Table S3: Differentially expressed genes in tissues

Table S4: Cell counts of B cell subsets in each tissue and donor

Table S5: Cell counts with transcriptome and BCR data and number of cells in clones found in one tissue compared to multiple tissues

Table S6: Differentially expressed genes in PBMC

Supplementary Excel S1: Raw data excel file

References:

1. T. Junt, E. Scandella, B. Ludewig, Form follows function: lymphoid tissue microarchitecture in antimicrobial immune defence. *Nat Rev Immunol* **8**, 764-775 (2008).
2. T. Yamanaka, L. Helgeland, I. N. Farstad, H. Fukushima, T. Midtvedt, P. Brandtzaeg, Microbial colonization drives lymphocyte accumulation and differentiation in the follicle-associated epithelium of Peyer's patches. *J Immunol* **170**, 816-822 (2003).
3. M. T. Drayson, W. L. Ford, Afferent lymph and lymph borne cells: their influence on lymph node function. *Immunobiology* **168**, 362-379 (1984).
4. S. A. Elmore, Histopathology of the lymph nodes. *Toxicol Pathol* **34**, 425-454 (2006).
5. P. Luscieti, T. Hubschmid, H. Cottier, M. W. Hess, L. H. Sobin, Human lymph node morphology as a function of age and site. *J Clin Pathol* **33**, 454-461 (1980).
6. B. S. Wilkins, D. H. Wright, *Illustrated Pathology of the Spleen*. (Cambridge: Cambridge University Press, Cambridge, 2000), pp. 204.
7. I. Sanz, C. Wei, S. A. Jenks, K. S. Cashman, C. Tipton, M. C. Woodruff, J. Hom, F. E. Lee, Challenges and Opportunities for Consistent Classification of Human B Cell and Plasma Cell Populations. *Front Immunol* **10**, 2458 (2019).
8. L. S. Rott, M. J. Briskin, E. C. Butcher, Expression of alpha4beta7 and E-selectin ligand by circulating memory B cells: implications for targeted trafficking to mucosal and systemic sites. *J Leukoc Biol* **68**, 807-814 (2000).
9. A. Vossenkämper, P. A. Blair, N. Safinia, L. D. Fraser, L. Das, T. J. Sanders, A. J. Stagg, J. D. Sanderson, K. Taylor, F. Chang, L. M. Choong, D. P. D'Cruz, T. T. Macdonald, G. Lombardi, J. Spencer, A role for gut-associated lymphoid tissue in shaping the human B cell repertoire. *Journal of Experimental Medicine* **210**, 1665-1674 (2013).
10. J. C. Weill, S. Weller, C. A. Reynaud, Human marginal zone B cells. *Annu Rev Immunol* **27**, 267-285 (2009).
11. S. Weller, M. C. Braun, B. K. Tan, A. Rosenwald, C. Cordier, M. E. Conley, A. Plebani, D. S. Kumararatne, D. Bonnet, O. Tournilhac, G. Tchernia, B. Steiniger, L. M. Staudt, J. L. Casanova, C. A. Reynaud, J. C. Weill, Human blood IgM "memory" B cells are circulating splenic marginal zone B cells harboring a prediversified immunoglobulin repertoire. *Blood* **104**, 3647-3654 (2004).
12. R. E. Mebius, M. A. Nolte, G. Kraal, Development and function of the splenic marginal zone. *Crit Rev Immunol* **24**, 449-464 (2004).
13. A. Kibler, B. Budeus, E. Homp, K. Bronischewski, V. Berg, L. Sellmann, F. Murke, A. Heinold, F. M. Heinemann, M. Lindemann, I. Bekeredjian-Ding, P. A. Horn, C. J. Kirschning, R. Küppers, M. Seifert, Systematic memory B cell archiving and random display shape the human splenic marginal zone throughout life. *Journal of Experimental Medicine* **218**, (2021).
14. G. Magri, L. Comerma, M. Pybus, J. Sintes, D. Llige, D. Segura-Garzon, S. Bascones, A. Yeste, E. K. Grasset, C. Gutzeit, M. Uzzan, M. Ramanujam, M. C. van Zelm, R. Albero-Gonzalez, I. Vazquez, M. Iglesias, S. Serrano, L. Marquez, E. Mercade, S. Mehandru, A. Cerutti, Human Secretory IgM Emerges from Plasma Cells Clonally Related to Gut Memory B Cells and Targets Highly Diverse Commensals. *Immunity* **47**, 118-134 e118 (2017).
15. N. M. Weisel, F. J. Weisel, D. L. Farber, L. Borghesi, Y. Shen, W. Ma, E. T. Luning Prak, M. Shlomchik, Comprehensive analyses of B cell compartments across the human body reveal novel subsets and a gut resident memory phenotype. *Blood*, (2020).
16. N. Nair, E. W. Newell, C. Vollmers, S. R. Quake, J. M. Morton, M. M. Davis, X. S. He, H. B. Greenberg, High-dimensional immune profiling of total and rotavirus VP6-specific intestinal and circulating B cells by mass cytometry. *Mucosal Immunol* **9**, 68-82 (2016).
17. Y. Zhao, M. Uduman, J. H. Y. Siu, T. J. Tull, J. D. Sanderson, Y. B. Wu, J. Q. Zhou, N. Petrov, R. Ellis, K. Todd, K. M. Chavele, W. Guesdon, A. Vossenkamper, W. Jassem, D. P. D'Cruz, D. J. Fear, S. John, D. Scheel-Toellner, C. Hopkins, E. Moreno, N. L. Woodman, F. Ciccarelli, S. Heck, S. H. Kleinstein, M. Bemark, J. Spencer, Spatiotemporal segregation of human marginal zone and memory B cell populations in lymphoid tissue. *Nat Commun* **9**, 3857 (2018).
18. D. Nemazee, Natural history of MZ B cells. *Journal of Experimental Medicine* **218**, (2021).
19. M. Descatoire, S. Weller, S. Irtan, S. Sarnacki, J. Feuillard, S. Storck, A. Guiochon-Mantel, J. Bouligand, A. Morali, J. Cohen, E. Jacquemin, M. Iascone, C. Bole-Feysot, N. Cagnard, J. C. Weill, C. A. Reynaud,

Identification of a human splenic marginal zone B cell precursor with NOTCH2-dependent differentiation properties. *J Exp Med* **211**, 987-1000 (2014).
20. J. L. Johnson, R. L. Rosenthal, J. J. Knox, A. Myles, M. S. Naradikian, J. Madej, M. Kostiv, A. M. Rosenfeld, W. Meng, S. R. Christensen, S. E. Hensley, J. Yewdell, D. H. Canaday, J. Zhu, A. B. Mcdermott, Y. Dori, M. Itkin, E. J. Wherry, N. Pardi, D. Weissman, A. Naji, E. T. L. Prak, M. R. Betts, M. P. Cancro, The Transcription Factor T-bet Resolves Memory B Cell Subsets with Distinct Tissue Distributions and Antibody Specificities in Mice and Humans. *Immunity* **52**, 842-855.e846 (2020).
21. C. M. Tipton, C. F. Fucile, J. Darce, A. Chida, T. Ichikawa, I. Gregoretti, S. Schieferl, J. Hom, S. Jenks, R. J. Feldman, R. Mehr, C. Wei, F. E. Lee, W. C. Cheung, A. F. Rosenberg, I. Sanz, Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus. *Nat Immunol* **16**, 755-765 (2015).
22. T. J. Tull, M. J. Pitcher, W. Guesdon, J. H. Y. Siu, C. Lebrero-Fernández, Y. Zhao, N. Petrov, S. Heck, R. Ellis, P. Dhami, U. D. Kadolsky, M. Kleeman, Y. Kamra, D. J. Fear, S. John, W. Jassem, R. W. Groves, J. D. Sanderson, M. D. Robson, D. P. D'Cruz, M. Bemark, J. Spencer, Human marginal zone B cell development from early T2 progenitors. *Journal of Experimental Medicine* **218**, (2021).
23. S. A. Jenks, K. S. Cashman, E. Zumaquero, U. M. Marigorta, A. V. Patel, X. Wang, D. Tomar, M. C. Woodruff, Z. Simon, R. Bugrovsky, E. L. Blalock, C. D. Scharer, C. M. Tipton, C. Wei, S. S. Lim, M. Petri, T. B. Niewold, J. H. Anolik, G. Gibson, F. E. Lee, J. M. Boss, F. E. Lund, I. Sanz, Distinct Effector B Cells Induced by Unregulated Toll-like Receptor 7 Contribute to Pathogenic Responses in Systemic Lupus Erythematosus. *Immunity* **49**, 725-739 e726 (2018).
24. J. B. Moroney, A. Vasudev, A. Pertsemlidis, H. Zan, P. Casali, Integrative transcriptome and chromatin landscape analysis reveals distinct epigenetic regulations in human memory B cells. *Nat Commun* **11**, 5435 (2020).
25. S. De Biasi, D. Lo Tartaro, M. Meschiari, L. Gibellini, C. Bellinazzi, R. Borella, L. Fidanza, M. Mattioli, A. Paolini, L. Gozzi, D. Jaacoub, M. Faltoni, S. Volpi, J. Milic, M. Sita, M. Sarti, C. Pucillo, M. Girardis, G. Guaraldi, C. Mussini, A. Cossarizza, Expansion of plasmablasts and loss of memory B cells in peripheral blood from COVID-19 patients with pneumonia. *Eur J Immunol* **50**, 1283-1294 (2020).
26. R. Finck, E. F. Simonds, A. Jager, S. Krishnaswamy, K. Sachs, W. Fantl, D. Pe'er, G. P. Nolan, S. C. Bendall, Normalization of mass cytometry data with bead standards. *Cytometry A* **83**, 483-494 (2013).
27. A. T. L. Lun, A. C. Richard, J. C. Marioni, Testing for differential abundance in mass cytometry data. *Nat Methods* **14**, 707-709 (2017).
28. L. McInnes, J. Healy, J. Melville, UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. 2018.
29. M. Bemark, J. Holmqvist, J. Abrahamsson, K. Mellgren, Translational Mini-Review Series on B cell subsets in disease. Reconstitution after haematopoietic stem cell transplantation - revelation of B cell developmental pathways and lineage phenotypes. *Clin Exp Immunol* **167**, 15-25 (2012).
30. D. Bagnara, M. Squillario, D. Kipling, T. Mora, A. M. Walczak, L. Da Silva, S. Weller, D. K. Dunn-Walters, J. C. Weill, C. A. Reynaud, A Reassessment of IgM Memory Subsets in Humans. *J Immunol* **195**, 3716-3724 (2015).
31. L. Kato, N. A. Begum, A. M. Burroughs, T. Doi, J. Kawai, C. O. Daub, T. Kawaguchi, F. Matsuda, Y. Hayashizaki, T. Honjo, Nonimmunoglobulin target loci of activation-induced cytidine deaminase (AID) share unique features with immunoglobulin genes. *Proc Natl Acad Sci U S A* **109**, 2479-2484 (2012).
32. C. Lindner, I. Thomsen, B. Wahl, M. Ugur, M. K. Sethi, M. Friedrichsen, A. Smoczek, S. Ott, U. Baumann, S. Suerbaum, S. Schreiber, A. Bleich, V. Gaboriau-Routhiau, N. Cerf-Bensussan, H. Hazanov, R. Mehr, P. Boysen, P. Rosenstiel, O. Pabst, Diversification of memory B cells drives the continuous adaptation of secretory antibodies to gut microbiota. *Nat Immunol* **16**, 880-888 (2015).
33. N. T. Gupta, J. A. Vander Heiden, M. Uduman, D. Gadala-Maria, G. Yaari, S. H. Kleinstein, Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. *Bioinformatics* **31**, 3356-3358 (2015).
34. M. S. Cabatingan, M. R. Schmidt, R. Sen, R. T. Woodland, Naive B lymphocytes undergo homeostatic proliferation in response to B cell deficit. *J Immunol* **169**, 6795-6805 (2002).
35. V. Bergen, R. A. Soldatov, P. V. Kharchenko, F. J. Theis, RNA velocity-current challenges and future perspectives. *Mol Syst Biol* **17**, e10282 (2021).
36. V. Bergen, M. Lange, S. Peidli, F. A. Wolf, F. J. Theis, Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol* **38**, 1408-1414 (2020).

37. L. Krzyzak, C. Seitz, A. Urbat, S. Hutzler, C. Ostalecki, J. Glasner, A. Hiergeist, A. Gessner, T. H. Winkler, A. Steinkasserer, L. Nitschke, CD83 Modulates B Cell Activation and Germinal Center Responses. *J Immunol* **196**, 3581-3594 (2016).
38. D. Starlets, Y. Gore, I. Binsky, M. Haran, N. Harpaz, L. Shvidel, S. Becker-Herman, A. Berrebi, I. Shachar, Cell-surface CD74 initiates a signaling cascade leading to cell proliferation and survival. *Blood* **107**, 4807-4816 (2006).
39. G. Chen, C. H. Liu, L. Zhou, R. M. Krug, Cellular DDX21 RNA helicase inhibits influenza A virus replication but is counteracted by the viral NS1 protein. *Cell Host Microbe* **15**, 484-493 (2014).
40. L. Yeung, M. J. Hickey, M. D. Wright, The Many and Varied Roles of Tetraspanins in Immune Cell Recruitment and Migration. *Front Immunol* **9**, 1644 (2018).
41. T. Saito, S. Chiba, M. Ichikawa, A. Kunisato, T. Asai, K. Shimizu, T. Yamaguchi, G. Yamamoto, S. Seo, K. Kumano, E. Nakagami-Yamaguchi, Y. Hamada, S. Aizawa, H. Hirai, Notch2 is preferentially expressed in mature B cells and indispensable for marginal zone B lineage development. *Immunity* **18**, 675-685 (2003).
42. D. R. Gibb, M. El Shikh, D. J. Kang, W. J. Rowe, R. El Sayed, J. Cichy, H. Yagita, J. G. Tew, P. J. Dempsey, H. C. Crawford, D. H. Conrad, ADAM10 is essential for Notch2-dependent marginal zone B cell development and CD23 cleavage in vivo. *J Exp Med* **207**, 623-635 (2010).
43. J. Spencer, T. Finn, K. A. Pulford, D. Y. Mason, P. G. Isaacson, The human gut contains a novel population of B lymphocytes which resemble marginal zone cells. *Clin Exp Immunol* **62**, 607-612 (1985).
44. C. Giesen, H. A. Wang, D. Schapiro, N. Zivanovic, A. Jacobs, B. Hattendorf, P. J. Schuffler, D. Grolimund, J. M. Buhmann, S. Brandt, Z. Varga, P. J. Wild, D. Gunther, B. Bodenmiller, Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat Methods* **11**, 417-422 (2014).
45. D. Bautista, C. Vásquez, P. Ayala-Ramírez, J. Téllez-Sosa, E. Godoy-Lozano, J. Martínez-Barnetche, M. Franco, J. Angel, Differential Expression of IgM and IgD Discriminates Two Subpopulations of Human Circulating IgM+IgD+CD27+ B Cells That Differ Phenotypically, Functionally, and Genetically. *Frontiers in Immunology* **11**, (2020).
46. P. L. Amlot, A. E. Hayes, Impaired human antibody response to the thymus-independent antigen, DNP-Ficoll, after splenectomy. Implications for post-splenectomy infections. *Lancet* **1**, 1008-1011 (1985).
47. W. Timens, A. Boes, T. Rozeboom-Uiterwijk, S. Poppema, Immaturity of the human splenic marginal zone in infancy. Possible contribution to the deficient infant immune response. *J Immunol* **143**, 3200-3206 (1989).
48. S. Weller, M. Mamani-Matsuda, C. Picard, C. Cordier, D. Lecoeuche, F. Gauthier, J. C. Weill, C. A. Reynaud, Somatic diversification in the absence of antigen-driven responses is the hallmark of the IgM+ IgD+ CD27+ B cell repertoire in infants. *J Exp Med* **205**, 1331-1342 (2008).
49. A. Danza, G. Ruiz-Irastorza, Infection risk in systemic lupus erythematosus patients: susceptibility factors and preventive strategies. *Lupus* **22**, 1286-1294 (2013).
50. E. Calo, R. A. Flynn, L. Martin, R. C. Spitale, H. Y. Chang, J. Wysocka, RNA helicase DDX21 coordinates transcription and ribosomal RNA processing. *Nature* **518**, 249-253 (2015).
51. W. Wu, Y. Qu, S. Yu, S. Wang, Y. Yin, Q. Liu, C. Meng, Y. Liao, Z. Ur Rehman, L. Tan, C. Song, X. Qiu, W. Liu, C. Ding, Y. Sun, Caspase-Dependent Cleavage of DDX21 Suppresses Host Innate Immunity. *mBio* **12**, e0100521 (2021).
52. M. C. Woodruff, R. P. Ramonell, D. C. Nguyen, K. S. Cashman, A. S. Saini, N. S. Haddad, A. M. Ley, S. Kyu, J. C. Howell, T. Ozturk, S. Lee, N. Suryadevara, J. B. Case, R. Bugrovsky, W. Chen, J. Estrada, A. Morrison-Porter, A. Derrico, F. A. Anam, M. Sharma, H. M. Wu, S. N. Le, S. A. Jenks, C. M. Tipton, B. Staitieh, J. L. Daiss, E. Ghosn, M. S. Diamond, R. H. Carnahan, J. E. Crowe, Jr., W. T. Hu, F. E. Lee, I. Sanz, Extrafollicular B cell responses correlate with neutralizing antibodies and morbidity in COVID-19. *Nat Immunol* **21**, 1506-1516 (2020).
53. T. Lopes-Carvalho, J. F. Kearney, Development and selection of marginal zone B cells. *Immunol Rev* **197**, 192-205 (2004).
54. H. Song, J. Cerny, Functional Heterogeneity of Marginal Zone B Cells Revealed by Their Ability to Generate Both Early Antibody-forming Cells and Germinal Centers with Hypermutation and Memory in Response to a T-dependent Antigen. *Journal of Experimental Medicine* **198**, 1923-1935 (2003).
55. P. M. Dammers, A. Visser, E. R. Popa, P. Nieuwenhuis, F. G. Kroese, Most marginal zone B cells in rat express germline encoded Ig VH genes and are ligand selected. *J Immunol* **165**, 6156-6169 (2000).

56. M. Lechner, T. Engleitner, T. Babushku, M. Schmidt-Supprian, R. Rad, L. J. Strobl, U. Zimber-Strobl, Notch2-mediated plasticity between marginal zone and follicular B cells. *Nat Commun* **12**, 1111 (2021).
57. J. A. Roco, L. Mesin, S. C. Binder, C. Nefzger, P. Gonzalez-Figueroa, P. F. Canete, J. Ellyard, Q. Shen, P. A. Robert, J. Cappello, H. Vohra, Y. Zhang, C. R. Nowosad, A. Schiepers, L. M. Corcoran, K. M. Toellner, J. M. Polo, M. Meyer-Hermann, G. D. Victora, C. G. Vinuesa, Class-Switch Recombination Occurs Infrequently in Germinal Centers. *Immunity* **51**, 337-350 e337 (2019).
58. H. Wardemann, S. Yurasov, A. Schaefer, J. W. Young, E. Meffre, M. C. Nussenzweig, Predominant autoantibody production by early human B cell precursors. *Science* **301**, 1374-1377 (2003).
59. A. C. Wotherspoon, C. Doglioni, T. C. Diss, L. Pan, A. Moschini, M. de Boni, P. G. Isaacson, Regression of primary low-grade B-cell gastric lymphoma of mucosa-associated lymphoid tissue type after eradication of Helicobacter pylori. *Lancet* **342**, 575-577 (1993).
60. D. Rossi, V. Trifonov, M. Fangazio, A. Bruscaggin, S. Rasi, V. Spina, S. Monti, T. Vaisitti, F. Arruga, R. Fama, C. Ciardullo, M. Greco, S. Cresta, D. Piranda, A. Holmes, G. Fabbri, M. Messina, A. Rinaldi, J. Wang, C. Agostinelli, P. P. Piccaluga, M. Lucioni, F. Tabbo, R. Serra, S. Franceschetti, C. Deambrogi, G. Daniele, V. Gattei, R. Marasca, F. Facchetti, L. Arcaini, G. Inghirami, F. Bertoni, S. A. Pileri, S. Deaglio, R. Foa, R. Dalla-Favera, L. Pasqualucci, R. Rabadan, G. Gaidano, The coding genome of splenic marginal zone lymphoma: activation of NOTCH2 and other pathways regulating marginal zone development. *J Exp Med* **209**, 1537-1551 (2012).
61. J. Spencer, T. C. Diss, P. G. Isaacson, A study of the properties of a low-grade mucosal B-cell lymphoma using a monoclonal antibody specific for the tumour immunoglobulin. *J Pathol* **160**, 231-238 (1990).
62. J. H. Southard, F. O. Belzer, Organ preservation. *Annu Rev Med* **46**, 235-247 (1995).
63. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2010).
64. C. Mcquin, A. Goodman, V. Chernyshev, L. Kamentsky, B. A. Cimini, K. W. Karhohs, M. Doan, L. Ding, S. M. Rafelski, D. Thirstrup, W. Wiegraebe, S. Singh, T. Becker, J. C. Caicedo, A. E. Carpenter, CellProfiler 3.0: Next-generation image processing for biology. *PLOS Biology* **16**, e2005970 (2018).
65. D. Schapiro, H. W. Jackson, S. Raghuraman, J. R. Fischer, V. R. T. Zanotelli, D. Schulz, C. Giesen, R. Catena, Z. Varga, B. Bodenmiller, histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data. *Nat Methods* **14**, 873-876 (2017).
66. N. Eling, N. Damond, T. Hoch, B. Bodenmiller, cytomapper: an R/Bioconductor package for visualization of highly multiplexed imaging data. *Bioinformatics*, (2020).
67. A. Kassambara. (2020).
68. T. Stuart, A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W. M. Mauck, 3rd, Y. Hao, M. Stoeckius, P. Smibert, R. Satija, Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902 e1821 (2019).
69. A. Kassambara. (2020).
70. V. K. Mootha, C. M. Lindgren, K. F. Eriksson, A. Subramanian, S. Sihag, J. Lehar, P. Puigserver, E. Carlsson, M. Ridderstrale, E. Laurila, N. Houstis, M. J. Daly, N. Patterson, J. P. Mesirov, T. R. Golub, P. Tamayo, B. Spiegelman, E. S. Lander, J. N. Hirschhorn, D. Altshuler, L. C. Groop, PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* **34**, 267-273 (2003).
71. A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, J. P. Mesirov, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550 (2005).
72. G. La Manno, R. Soldatov, A. Zeisel, E. Braun, H. Hochgerner, V. Petukhov, K. Lidschreiber, M. E. Kastriti, P. Lönnerberg, A. Furlan, J. Fan, L. E. Borm, Z. Liu, D. Van Bruggen, J. Guo, X. He, R. Barker, E. Sundström, G. Castelo-Branco, P. Cramer, I. Adameyko, S. Linnarsson, P. V. Kharchenko, RNA velocity of single cells. *Nature* **560**, 494-498 (2018).
73. J. Fox. (2020).
74. R Core Team. (2020).
75. J. Melville. (2020).
76. Z. Gu, R. Eils, M. Schlesner, Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847-2849 (2016).
77. Z. Gu, L. Gu, R. Eils, M. Schlesner, B. Brors, circlize Implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811-2812 (2014).
78. K. Blighe, S. Rana, M. Lewis. (2020).

79. T. Wei, V. Simko. (2017).
80. H. Wickham, ggplot2: Elegant Graphics for Data Analysis. *Use R*, 1-212 (2009).

Acknowledgments:
We thank the deceased organ donors, donor families and the Cambridge Biorepository for Translational Medicine for access to tissue samples. This research was supported by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London and/or the NIHR Clinical Research Facility. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care . We would also like to thank Silvia Cellone Trevelin for help with the histology, Charles Armitage for testing antibodies, and Francesca Ciccarelli and Michele Bortolomeazzi for guidance with image analysis.

Funding:
This research was funded in part, by the Wellcome Trust 220872/Z/20/Z. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.
Medical Research Council of Great Britain grant MR/R000964/1 and MR/P021964/1
The Lupus Trust
Chan Zuckerberg Initiative (KTM, KSP)
Gates Cambridge Trust (JHS)
Canadian Centennial Scholarship Fund (JHS)
Swedish Research Council (MB)
County Council of Västra Götaland (MB)
Cancer Research UK KHP Cancer Centre (JS)

Author contributions:

Conceptualization and design of study: JHS, JS, GJP, TJT, MB

Sample identification and collection: JHS, KTM, KSP, GJP

Data acquisition and methodology: JHS, TJT, RLV, LM, RE, PD, UDK, MK, KT, JS

Data analysis: JHS, MJP, WG, TJT

Supervision and funding: JS, GJP, MB, DDC

Writing manuscript: JHS, JS, GJP, MB, TJT

Competing interests:

The authors declare no competing interests.

Data and materials availability:

All raw and processed next-generation sequencing data have been deposited with GEO under accession numbers GSE193869, GSE193867 and GSE193868. Mass cytometry and image mass cytometry data available upon request. Code is available on Zenodo (doi:10.5281/zenodo.5849834).

Figures:

Fig. 1: Differentially abundant B cell subpopulations between human lymphoid tissues (A). UMAP plot of the median positions of hyperspheres for $CD19^+$ B cells in concatenated human appendix, mLN, and spleen. Each point represents a hypersphere coloured by the median intensity of selected markers (bounded by the 5th and 95th percentiles of the intensities across all cells) for that hypersphere. **(B).** UMAP coloured by scaled cell count within each hypersphere for each tissue. The larger points represent the significantly differentially abundant hyperspheres detected at a spatial FDR of 5%. **(C).** Heatmap of markers expressed in hyperspheres with significantly difference abundances, scaled for each marker individually by row to highlight differences between clusters. Hyperspheres were clustered by hierarchical clustering and k-means with coloured dendrogram to identify clusters. **(D).** UMAP coloured by cluster as in **(C). (E).** SPADE on viSNE was used to manually identify B cell subsets (GC = germinal centre, TS = transitional, MZB = marginal zone B). Nodes represent a cluster of phenotypically similar cells, and the size of a node is proportional to the number of cells represented by it. MZB-1 and MZB-2 subsets are highlighted in red. **(F)**. CCR7 expression, as represented by the colour, was used to distinguish MZB-1 and MZB-2. **(G).** Significantly different median expression of BAFFR, CD24, and CD27 between the two MZB SPADE populations for each sample (n = 24). Statistics were assessed by paired t-test. **(H).** Relative proportion of MZB-1 and MZB-2 SPADE subsets in each lymphoid tissue. Statistics were assessed by ANOVA and TukeyHSD post-hoc. *$p<0.05$, ***$p<0.01$ and ****$p<0.0001$.

Fig. 2: Overview of B cell composition of human lymphoid tissues from single cell transcriptome and BCR profiling (A). UMAP visualisation of select mRNA transcripts (top) and ADT surface protein (bottom) of total $CD19^+$ B cells. **(B).** UMAP visualisation of B cell composition in lymphoid tissues coloured by B cell subsets, annotated based on Seurat

unsupervised clusters (AcB1-4, activated B cells 1-4; aNAV, activated naive; DN, double negative; GC, germinal centre B cells; MZB, marginal zone B cells; PB, plasmablast; TS, transitional B cells). **(C).** Dot plot illustrating marker gene expression for B cell subtypes. The following gene groups were used: GC.RNA.1 (*BCL6, BCL7A*); IgG.RNA.1 (*IGHG1*, *IGHG2, IGHG3, IGHG4*); IgA.RNA.1 (*IGHA1, IGHA2*). **(D).** Frequency of somatic mutations in *IGHV* genes used by B cells in each B cell subset. **(E).** UMAP visualisation of B cell subsets in each lymphoid tissue. **(F).** Relative proportion of B cell subsets in each lymphoid tissue (n = 3). Statistics for differential abundance between matched tissues for each subset were assessed by estimated marginal means with Bonferroni correction for multiple comparisons. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$ and ****$p < 0.0001$.

Fig. 3: Clonal relationship and dissemination of B cell subsets within and between tissues. **(A).** Percentage of cells in each subset and tissue (shape) that are part of a clone with members in one tissue (blue) or a clone with members in two or more tissues (red) compared to the total number of cells in each subset. **(B)**. Clonal relatedness between B cell subsets within each tissue (tendency for some subsets to be found together within the same clone) in the single cell dataset illustrated in a correlation plot. Colour represents the correlation matrix coefficients. Statistically significant tendencies for clonal relatedness were identified using Spearman correlation along with pairwise p-values that were then corrected using Holm's method. **(C).** Circos plots showing the clonal relationships between B cell subsets in the appendix (blue text), mLN (black text), and spleen (red text) for Donor A (left), B (middle), C (right). Clonally related sequences across tissue-subsets are connected by lines, with the top 5% most frequent connections coloured in dark grey and with sequences spanning MZB-1 across tissues coloured in red. All other connections are coloured light grey. **(D)**. RNA velocities of B cells for each tissue shown on the UMAP plot in Fig. 2 B. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

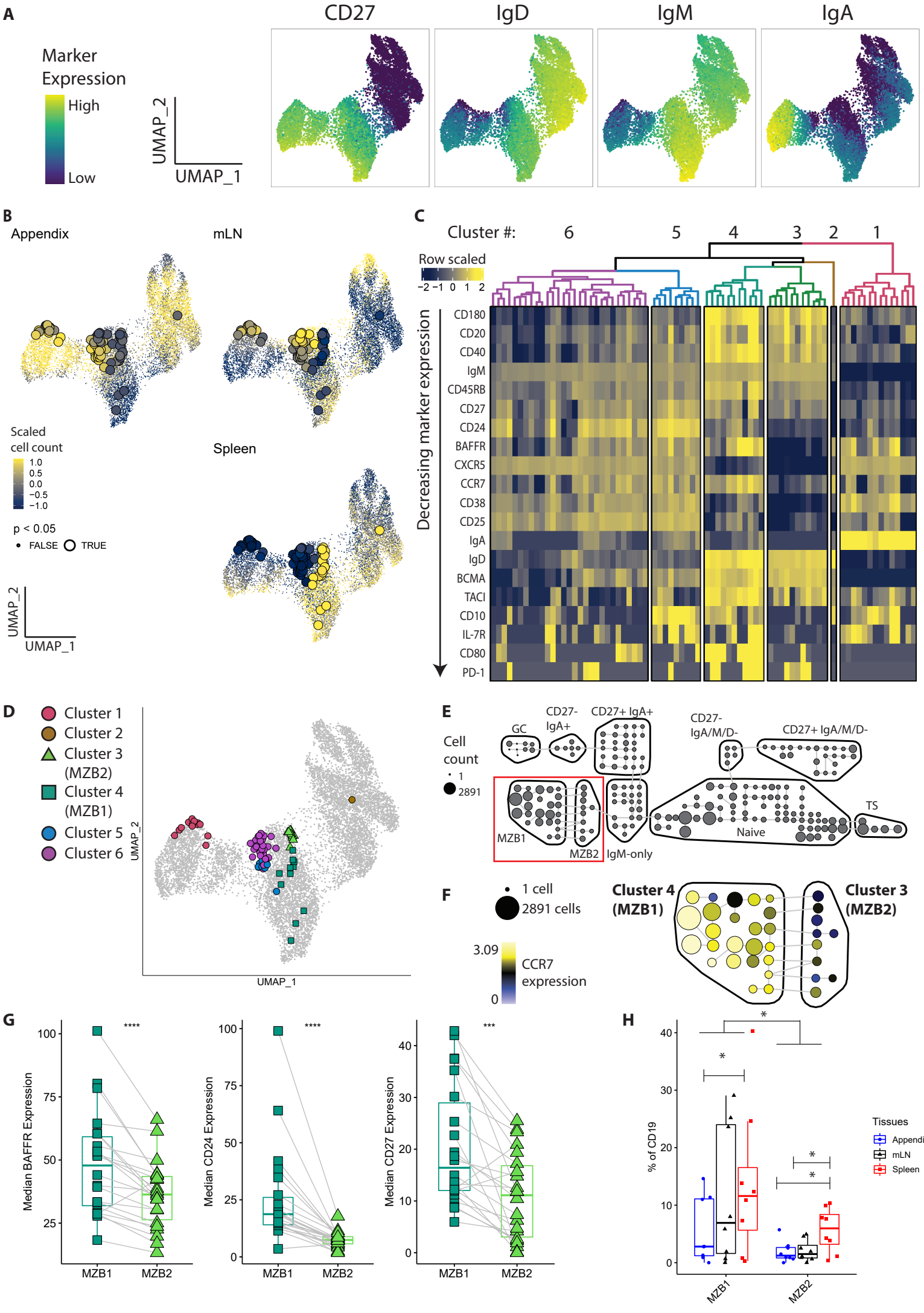
Fig. 4: Transcriptomic differences between MZB-1 and MZB-2 B cell subsets. (A). Volcano plot comparing differentially expressed genes (n = 3157) between MZB-1 (left) and MZB-2 (right). Cut offs for significant p value is 10^{-6}, and $\log_2$FC is 0.25. Red dots highlight the significantly differentially expressed genes. P values were calculated using Wilcoxon test with Bonferroni correction for multiple comparisons. Due to space constraints, only select markers are

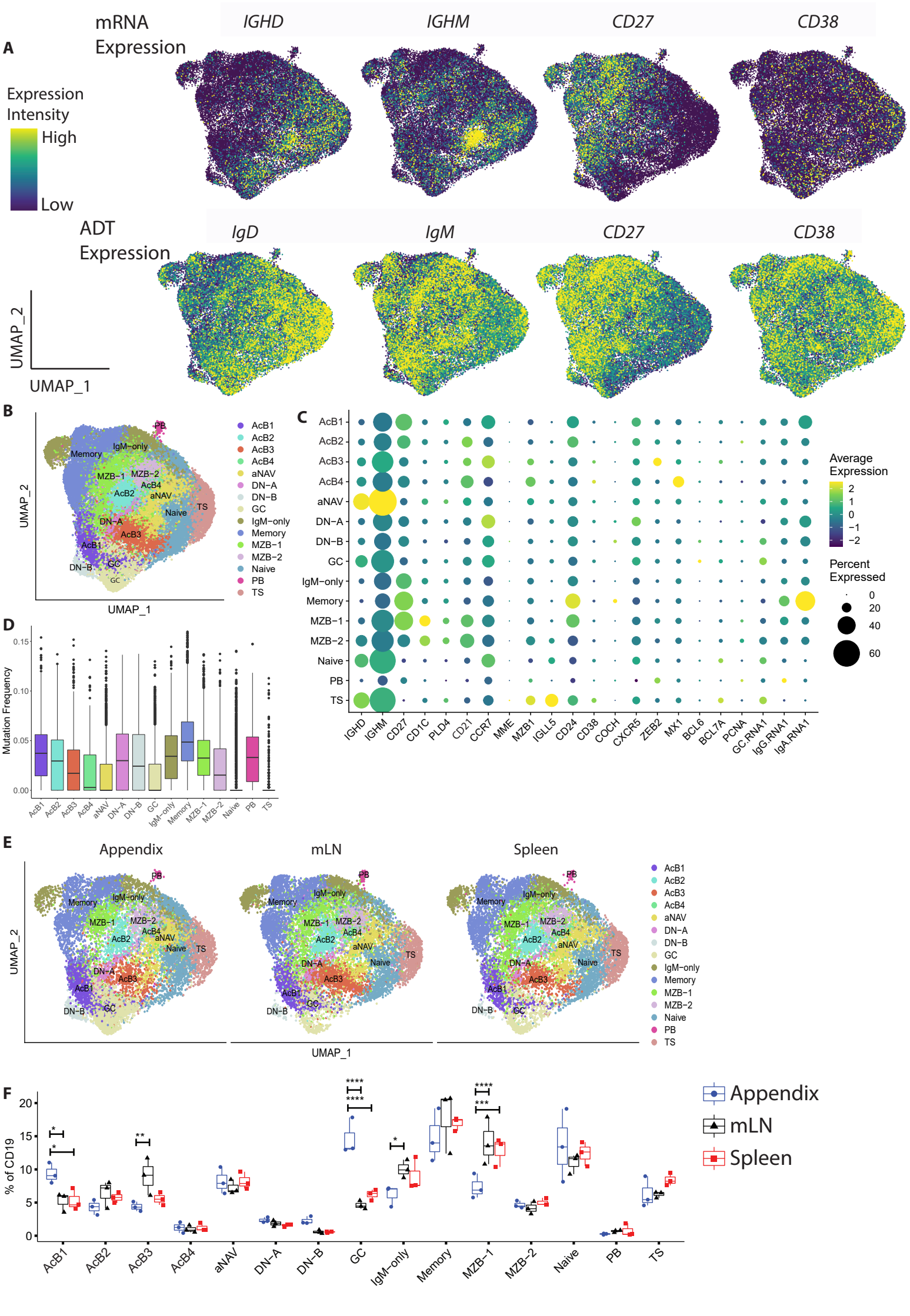
labelled. **(B).** Violin plots demonstrating gene expression of *DDX21, CD83, MIF, TXNIP, LTB,* and *CD37* in MZB-1 and MZB-2 B cells across the three tissues. All comparisons of gene expression within each tissue between MZB-1 and MZB-2 are significant at $p<10^{-12}$ using Kolmogorov-Smirnov test with Bonferroni correction for multiple comparisons. Values for individual cells are shown as dots. **(C).** Enrichment plot of the GSEA analysis using NOTCH gene sets from ImmuneSigDB between MZB-1 and MZB-2 cells, corrected for dataset background. Gene set is significantly enriched in the MZB-2 subset at nominal p value < 1%.

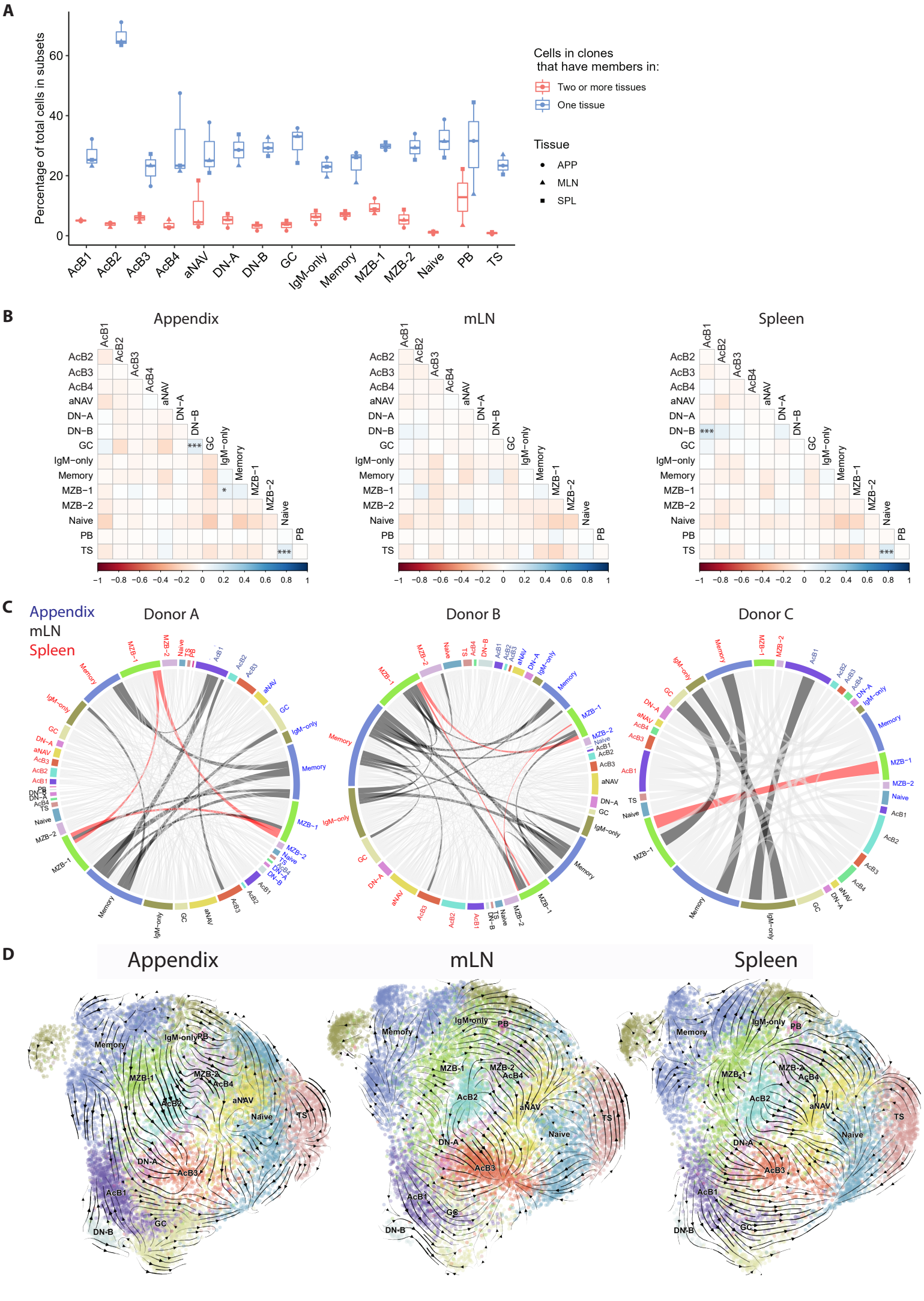
Fig. 5: Spatial distribution of MZB in the spleen. Visualisation of microanatomy of human spleen by imaging mass cytometry. A representative example of spleen from 10 regions of interest. **(A).** Composite pixel-level image to visualise CD3 (T cells; red), CD20 (B cells; green), CD68 (macrophages; blue) and Ki67 (proliferation; white). **(B).** CD27 (red), CD1c (green), and IgD (blue). **(C).** CD1c gating threshold for segmented $CD20^+$ single cells. **(D).** Visualisation of $CD1c^{+/-}$ cell type (coloured by outline) and CD1c marker expression (coloured by fill) on $CD20^+$ segmentation masks. **(E).** DDX21 expression levels between $CD1c^-$ and $CD1c^+$ B cells. Statistics for different expression levels between groups were assessed by paired t-test. **** $p < 0.0001$. **(F).** CD1c in green and DDX21 in red are visualised as a composite pixel-level image. **(G).** Scatterplot indicating the location of every $CD1c^+$ B cell (black), and the centre point of these cells (red). **(H).** Distribution of the number of cells (y-axis; maximum value scaled to 1) at the various binned spatial distances between the centre point and every $CD1c^+$ cell (x-axis; mean scaled to 0.5). Colour represents high (blue) and low (red) DDX21 expression. **(I).** Proportion of high DDX21 expression at each binned spatial distance from the previous histogram. A linear regression, weighted by cell counts, was performed (blue line). Grey indicates 95% confidence intervals. **(J).** Boxplot summary of the weighted linear regression slopes from regions of interest with a single follicle (n = 8) for DDX21 and DNA. One sample t-tests were used to compare the slopes for each marker to zero. **(K).** Visualisation of DDX21 high (red) and DDX21 low (blue) MZB cells.

Fig. 6: Comparison of MZB subsets in blood of healthy and SLE donors. (A). UMAP visualisation of B cell composition in blood coloured by B cell subsets, annotated based on Seurat unsupervised clusters (AcB3, activated B cells 3; aNAV, activated naive; DN, double

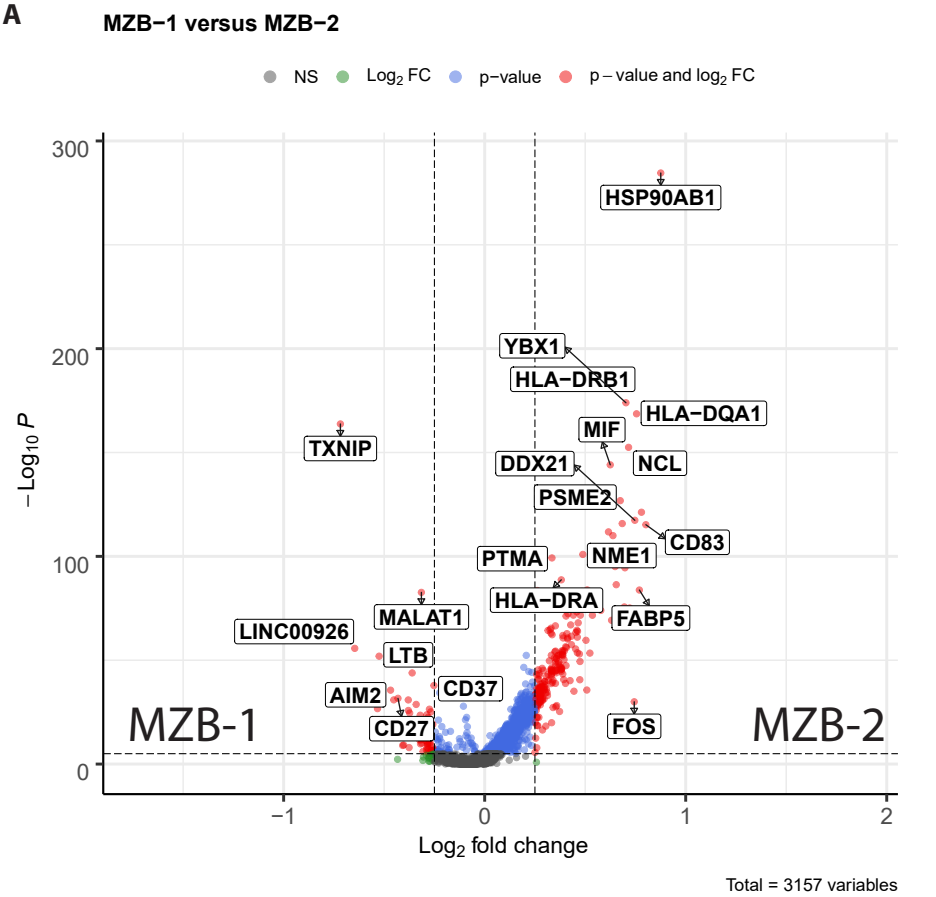
negative; GC, germinal centre B cells; MZB, marginal zone B cells; PB, plasmablast; TS, transitional B cells). **(B).** UMAP visualisation of *PLD4* mRNA transcripts. **(C).** UMAP visualisation of *CD1C* mRNA transcripts. **(D).** Frequency of somatic mutations in *IGHV* associated with each B cell subset. **(E).** Clonal relatedness between B cell subsets within blood of healthy and SLE patients illustrated in a correlation plot. Colour represents the correlation matrix coefficients. Significant relationships were assessed using Spearman correlation along with pairwise p-values that were then corrected using Holm's method. **(F).** UMAP visualisation of IgM ADT surface protein expression in healthy blood B cells. Arrow indicates the IgM^{hi} bridge linking MZB-1 and naive B cells. **(G) and (H):** Dot plots illustrating the gene expressions of markers that discriminated between two subsets of B cells by mass cytometry in Fig. 1. **(G).** for B cell subtypes in healthy blood and **(H).** lymphoid tissues. **(I) and (J)**: Dot plots illustrating ADT surface protein expression of ß7 integrin for B cell subtypes in **(I).** healthy blood and **(J).** tissues. **(K).** Relative proportion of B cell subsets from healthy and SLE blood (n = 3). Statistics for differential abundance between samples for each subset were assessed by estimated marginal means with Bonferroni correction for multiple comparisons. **(L).** Relative proportion of MZB-1 and MZB-2 in health and lupus determined using flow cytometry (n = 10). Statistics for differential abundance between subsets were assessed by paired Wilcoxon test. **(M).** ß7 median fluorescent intensity (MFI) for MZB-1 and MZB-2 in health and lupus. Statistics between health and lupus were assessed by Wilcoxon test. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns = not significant.


A
CD27
IgD
IgM
IgA
Marker Expression
High
Low
UMAP_2
UMAP_1
B
Appendix
mLN
Spleen
Scaled cell count
1.0
0.5
0.0
−0.5
−1.0
p < 0.05
FALSE
TRUE
UMAP_2
UMAP_1
C
Cluster #: 6 5 4 3 2 1
Row scaled
−2 −1 0 1 2
Decreasing marker expression
CD180
CD20
CD40
IgM
CD45RB
CD27
CD24
BAFFR
CXCR5
CCR7
CD38
CD25
IgA
IgD
BCMA
TACI
CD10
IL-7R
CD80
PD-1
D
Cluster 1
Cluster 2
Cluster 3 (MZB2)
Cluster 4 (MZB1)
Cluster 5
Cluster 6
UMAP_2
UMAP_1
E
Cell count
1
2891
GC
CD27- IgA+
CD27+ IgA+
CD27- IgA/M/D-
CD27+ IgA/M/D-
TS
Naive
MZB1
MZB2
IgM-only
F
1 cell
2891 cells
3.09
0
CCR7 expression
Cluster 4 (MZB1)
Cluster 3 (MZB2)
G
Median BAFFR Expression
Median CD24 Expression
Median CD27 Expression
****
****
***
MZB1
MZB2
H
% of CD19
Tissues
Appendix
mLN
Spleen
*
*
*
*
*
MZB1
MZB2

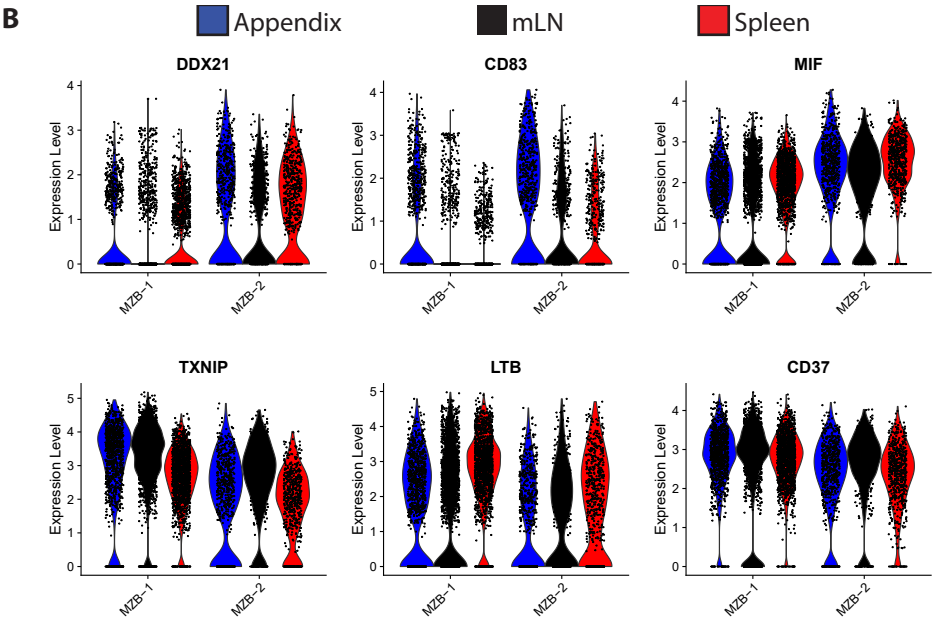

A
mRNA Expression
IGHD
IGHM
CD27
CD38
Expression Intensity
High
Low
ADT Expression
IgD
IgM
CD27
CD38
UMAP_2
UMAP_1
B
PB
IgM−only
Memory
MZB−1
MZB−2
AcB4
AcB2
aNAV
TS
Naive
DN−A
AcB3
AcB1
DN−B
GC
GC
UMAP_2
UMAP_1
AcB1
AcB2
AcB3
AcB4
aNAV
DN−A
DN−B
GC
IgM−only
Memory
MZB−1
MZB−2
Naive
PB
TS
C
AcB1
AcB2
AcB3
AcB4
aNAV
DN−A
DN−B
GC
IgM−only
Memory
MZB−1
MZB−2
Naive
PB
TS
IGHD
IGHM
CD27
CD1C
PLD4
CD21
CCR7
MME
MZB1
IGLL5
CD24
CD38
COCH
CXCR5
ZEB2
MX1
BCL6
BCL7A
PCNA
GC.RNA1
IgG.RNA1
IgA.RNA1
Average Expression
2
1
0
−1
−2
Percent Expressed
0
20
40
60
D
Mutation Frequency
0.15
0.10
0.05
0.00
AcB1
AcB2
AcB3
AcB4
aNAV
DN−A
DN−B
GC
IgM−only
Memory
MZB−1
MZB−2
Naive
PB
TS
E
Appendix
mLN
Spleen
UMAP_2
UMAP_1
PB
IgM−only
Memory
MZB−1
MZB−2
AcB4
AcB2
aNAV
Naive
TS
DN−A
AcB3
AcB1
DN−B
GC
AcB1
AcB2
AcB3
AcB4
aNAV
DN−A
DN−B
GC
IgM−only
Memory
MZB−1
MZB−2
Naive
PB
TS
F
% of CD19
20
15
10
5
0
AcB1
AcB2
AcB3
AcB4
aNAV
DN−A
DN−B
GC
IgM−only
Memory
MZB−1
MZB−2
Naive
PB
TS
Appendix
mLN
Spleen


A
Percentage of total cells in subsets
0
20
40
60
AcB1
AcB2
AcB3
AcB4
aNAV
DN-A
DN-B
GC
IgM-only
Memory
MZB-1
MZB-2
Naive
PB
TS
Cells in clones
that have members in:
Two or more tissues
One tissue
Tissue
APP
MLN
SPL
B
Appendix
mLN
Spleen
C
Appendix
mLN
Spleen
Donor A
Donor B
Donor C
D
Appendix
mLN
Spleen

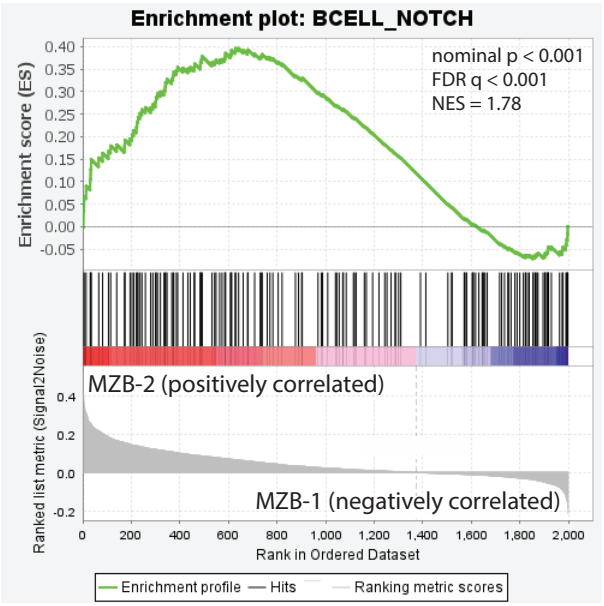
A

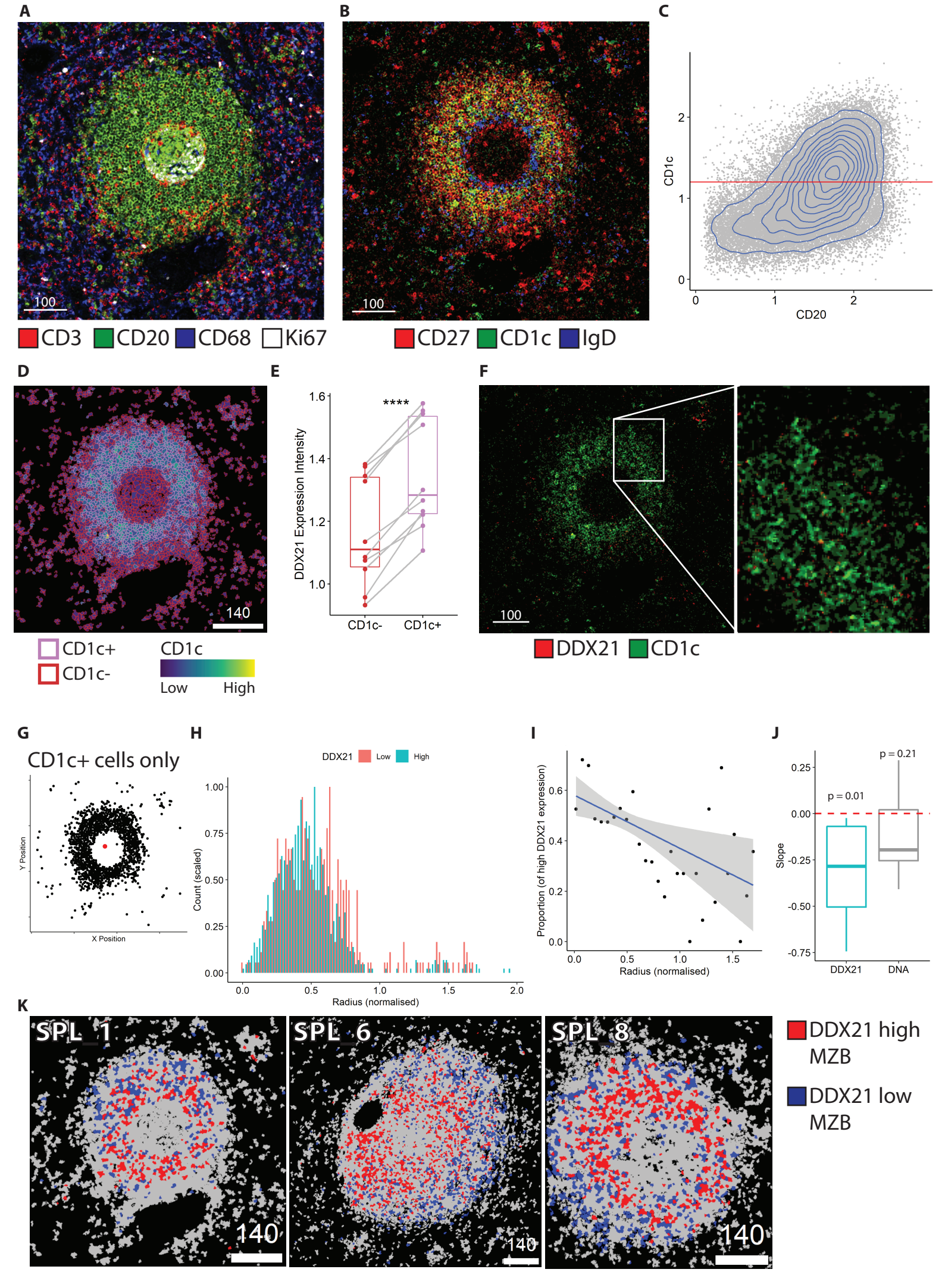

MZB−1 versus MZB−2
NS
Log2 FC
p−value
p − value and log2 FC
HSP90AB1
YBX1
HLA−DRB1
HLA−DQA1
MIF
TXNIP
DDX21
NCL
PSME2
CD83
PTMA
NME1
HLA−DRA
FABP5
MALAT1
LINC00926
LTB
AIM2
CD37
CD27
FOS
MZB-1
MZB-2
−Log10 P
Log2 fold change
Total = 3157 variables

B


Appendix
mLN
Spleen
DDX21
CD83
MIF
TXNIP
LTB
CD37
Expression Level
MZB−1
MZB−2

C


Enrichment plot: BCELL_NOTCH
nominal p < 0.001
FDR q < 0.001
NES = 1.78
Enrichment score (ES)
MZB-2 (positively correlated)
MZB-1 (negatively correlated)
Ranked list metric (Signal2Noise)
Rank in Ordered Dataset
Enrichment profile
Hits
Ranking metric scores


A
B
C
100
100
CD1c
CD20
CD3 CD20 CD68 Ki67
CD27 CD1c IgD
D
E
F
140
****
DDX21 Expression Intensity
CD1c-
CD1c+
100
CD1c+
CD1c-
CD1c
Low
High
DDX21 CD1c
G
H
I
J
CD1c+ cells only
Y Position
X Position
DDX21 Low High
Count (scaled)
Radius (normalised)
Proportion (of high DDX21 expression)
Radius (normalised)
p = 0.01
p = 0.21
Slope
DDX21
DNA
K
SPL_1
SPL_6
SPL_8
140
140
140
DDX21 high MZB
DDX21 low MZB

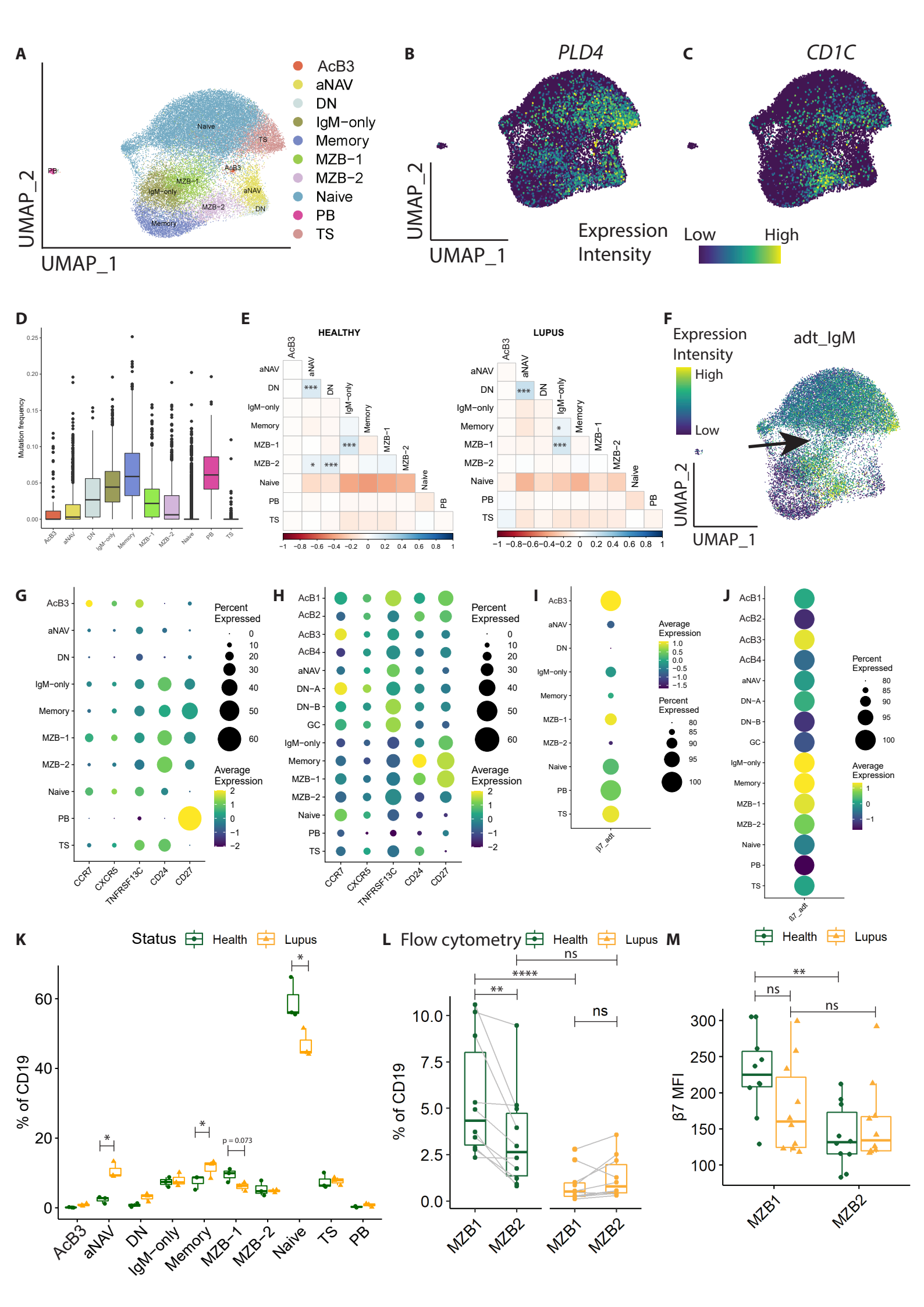

A
AcB3
aNAV
DN
IgM−only
Memory
MZB−1
MZB−2
Naive
PB
TS
UMAP_1
UMAP_2
B
PLD4
C
CD1C
Expression Intensity
Low
High
D
Mutation frequency
E
HEALTHY
LUPUS
F
Expression Intensity
High
Low
adt_IgM
G
Percent Expressed
Average Expression
CCR7
CXCR5
TNFRSF13C
CD24
CD27
H
AcB1
AcB2
AcB4
DN−A
DN−B
GC
I
β7_adt
J
K
Status
Health
Lupus
% of CD19
p = 0.073
L
Flow cytometry
ns
****
**
M
β7 MFI
MZB1
MZB2

Figure S1

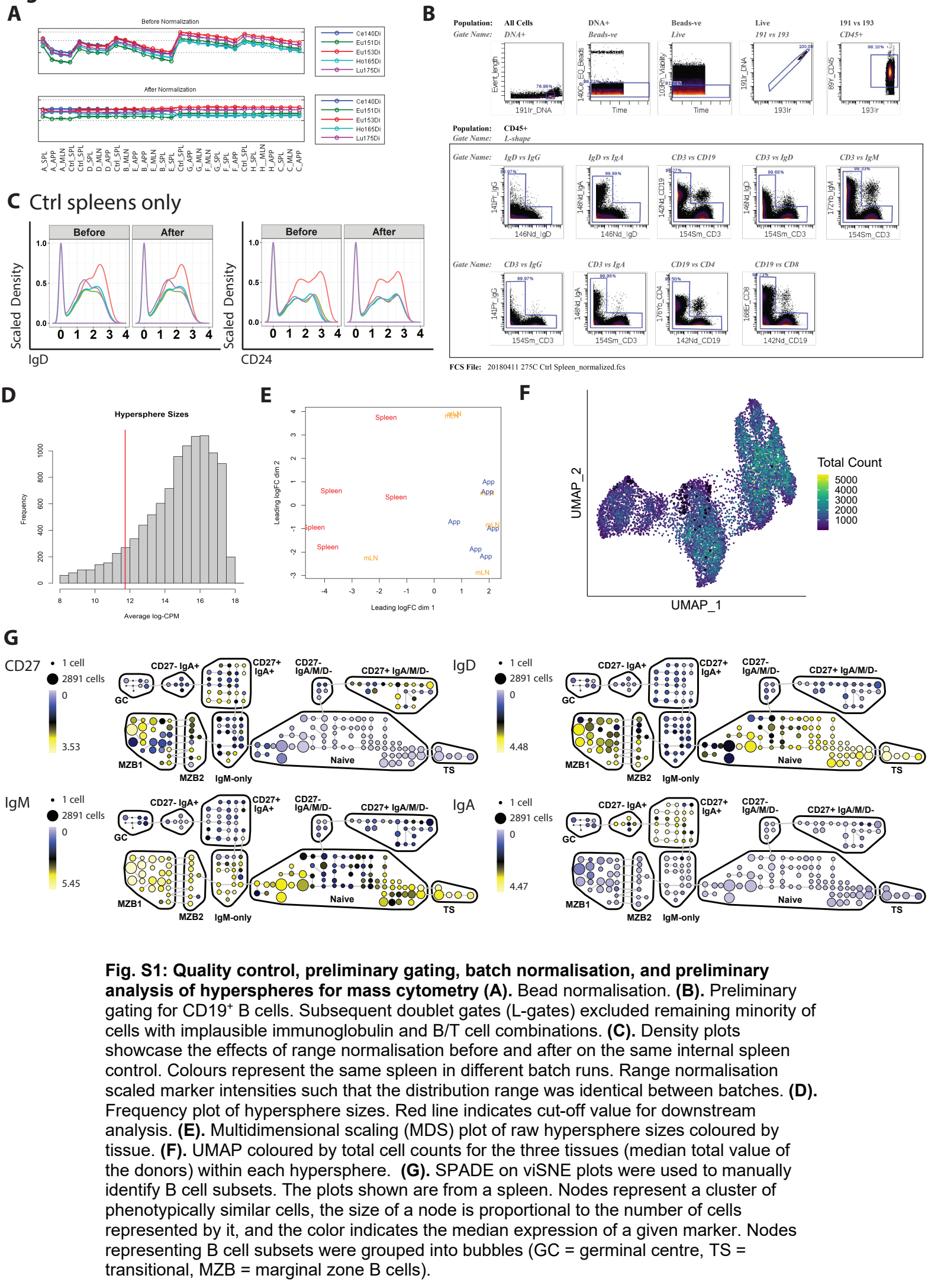


Fig. S1: Quality control, preliminary gating, batch normalisation, and preliminary analysis of hyperspheres for mass cytometry (A). Bead normalisation. **(B).** Preliminary gating for CD19^{+} B cells. Subsequent doublet gates (L-gates) excluded remaining minority of cells with implausible immunoglobulin and B/T cell combinations. **(C).** Density plots showcase the effects of range normalisation before and after on the same internal spleen control. Colours represent the same spleen in different batch runs. Range normalisation scaled marker intensities such that the distribution range was identical between batches. **(D).** Frequency plot of hypersphere sizes. Red line indicates cut-off value for downstream analysis. **(E).** Multidimensional scaling (MDS) plot of raw hypersphere sizes coloured by tissue. **(F).** UMAP coloured by total cell counts for the three tissues (median total value of the donors) within each hypersphere. **(G).** SPADE on viSNE plots were used to manually identify B cell subsets. The plots shown are from a spleen. Nodes represent a cluster of phenotypically similar cells, the size of a node is proportional to the number of cells represented by it, and the color indicates the median expression of a given marker. Nodes representing B cell subsets were grouped into bubbles (GC = germinal centre, TS = transitional, MZB = marginal zone B cells).

Figure S2

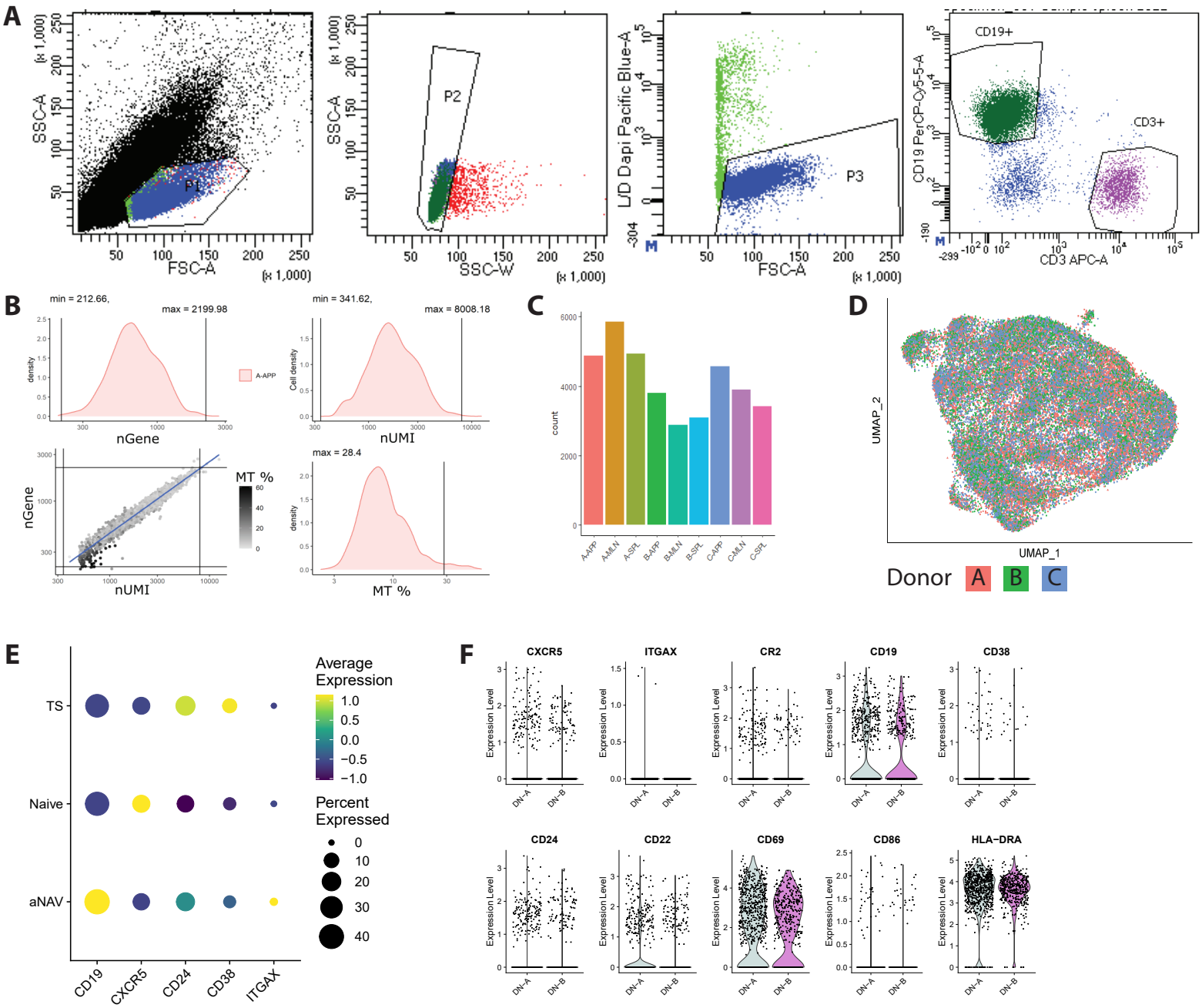


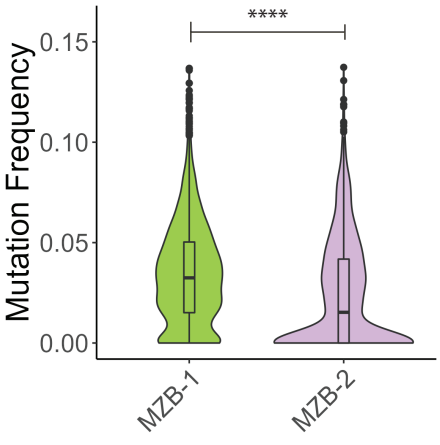
Fig. S2: Sort strategy, 10X genomics workflow & quality control, batch integration for single cell transcriptomics, and key gene expressions (A). Gating strategy to sort live CD19^{+} B cells. **(B).** Cell barcodes processed where barcodes with transcripts counts per cell (nUMI, unique molecular identifiers) and genes detected per cell (nGene) values above/below 3 median absolute deviation (MAD) and mitochondria percentage above MAD were removed. **(C).** Cell counts after quality control processing. **(D).** UMAP visualisation of B cell composition in tissues after integration, coloured by sample. **(E).** Dot plot illustrating *CD19, CXCR5, CD24, CD38, ITGAX (CD11c)* gene expression for TS, naive, and aNAV B cell subsets. **(F).** Violin plots of key DN differentiating genes for DN-A and DN-B subtypes.

Figure S3

A

Subset combination	diff	lwr	upr	p adj
AcB3-AcB1	-0.014242	-0.017108	-0.011376	0
AcB4-AcB1	-0.017732	-0.022963	-0.0125	0
aNAV-AcB1	-0.021997	-0.024625	-0.01937	0
GC-AcB1	-0.023043	-0.025622	-0.020465	0
Memory-AcB1	0.012805	0.010449	0.015161	0
MZB-2-AcB1	-0.014409	-0.017426	-0.011391	0
Naive-AcB1	-0.031569	-0.03395	-0.029189	0
TS-AcB1	-0.036434	-0.039095	-0.033772	0
aNAV-AcB2	-0.01588	-0.018737	-0.013024	0
GC-AcB2	-0.016926	-0.019738	-0.014115	0
Memory-AcB2	0.018922	0.016313	0.021531	0
Naive-AcB2	-0.025452	-0.028083	-0.022821	0
TS-AcB2	-0.030317	-0.033205	-0.027429	0
IgM-only-AcB3	0.012329	0.009538	0.01512	0
Memory-AcB3	0.027047	0.024605	0.029489	0
MZB-1-AcB3	0.010671	0.008094	0.013249	0
Naive-AcB3	-0.017327	-0.019793	-0.014862	0
TS-AcB3	-0.022192	-0.024929	-0.019454	0
Memory-AcB4	0.030536	0.025525	0.035548	0
TS-AcB4	-0.018702	-0.023864	-0.01354	0
DN-A-aNAV	0.018332	0.014195	0.022469	0
IgM-only-aNAV	0.020085	0.01754	0.02263	0
Memory-aNAV	0.034802	0.032646	0.036958	0
MZB-1-aNAV	0.018427	0.016118	0.020736	0
Naive-aNAV	-0.009572	-0.011755	-0.007389	0
TS-aNAV	-0.014436	-0.016923	-0.01195	0
GC-DN-A	-0.019378	-0.023484	-0.015271	0
Memory-DN-A	0.01647	0.0125	0.020441	0
Naive-DN-A	-0.027904	-0.031889	-0.023919	0
TS-DN-A	-0.032768	-0.036927	-0.028609	0
GC-DN-B	-0.016657	-0.021576	-0.011737	0
Memory-DN-B	0.019192	0.014385	0.023998	0
Naive-DN-B	-0.025182	-0.030001	-0.020364	0
TS-DN-B	-0.030047	-0.035011	-0.025083	0
IgM-only-GC	0.021131	0.018636	0.023625	0
Memory-GC	0.035848	0.033752	0.037945	0
MZB-1-GC	0.019473	0.017219	0.021726	0
Naive-GC	-0.008526	-0.01065	-0.006402	0
TS-GC	-0.01339	-0.015825	-0.010955	0
Memory-IgM-only	0.014718	0.012454	0.016981	0
MZB-2-IgM-only	-0.012496	-0.015442	-0.00955	0
Naive-IgM-only	-0.029657	-0.031946	-0.027367	0
TS-IgM-only	-0.034521	-0.037101	-0.03194	0
MZB-1-Memory	-0.016376	-0.01837	-0.014381	0
MZB-2-Memory	-0.027214	-0.029831	-0.024596	0
Naive-Memory	-0.044374	-0.046222	-0.042527	0
TS-Memory	-0.049238	-0.051436	-0.047041	0
MZB-2-MZB-1	-0.010838	-0.013583	-0.008093	0
Naive-MZB-1	-0.027999	-0.030023	-0.025975	0
TS-MZB-1	-0.032863	-0.035211	-0.030515	0
Naive-MZB-2	-0.017161	-0.019801	-0.014521	0
TS-MZB-2	-0.022025	-0.024921	-0.019129	0
GC-AcB3	-0.008801	-0.011459	-0.006144	3.00E-14
MZB-2-aNAV	0.007589	0.004724	0.010453	1.01E-13
AcB3-AcB2	-0.008125	-0.011203	-0.005047	1.03E-13
MZB-2-AcB2	-0.008292	-0.011511	-0.005073	1.22E-13
MZB-2-GC	0.008635	0.005815	0.011455	1.30E-13
MZB-1-AcB4	0.014161	0.009082	0.01924	1.34E-13
DN-B-aNAV	0.015611	0.010665	0.020556	1.37E-13
IgM-only-AcB4	0.015819	0.010628	0.02101	1.44E-13
Naive-AcB4	-0.013838	-0.018861	-0.008814	1.49E-13
TS-PB	-0.035758	-0.047374	-0.024141	1.50E-13
DN-A-AcB3	0.010577	0.006283	0.01487	1.64E-13
aNAV-AcB3	-0.007755	-0.01046	-0.005051	1.73E-13
MZB-2-DN-A	-0.010743	-0.015139	-0.006348	1.77E-13
PB-Naive	0.030893	0.019338	0.042448	1.96E-13
DN-A-AcB4	0.014066	0.007935	0.020197	8.43E-13
TS-Naive	-0.004864	-0.007089	-0.00264	1.27E-11
AcB4-AcB2	-0.011615	-0.016965	-0.006264	1.92E-11
AcB2-AcB1	-0.006117	-0.009127	-0.003107	5.78E-10
PB-GC	0.022367	0.01077	0.033965	6.43E-09
PB-aNAV	0.021321	0.009713	0.03293	4.92E-08
DN-B-AcB4	0.011345	0.004642	0.018048	9.86E-07
MZB-2-DN-B	-0.008022	-0.013185	-0.002858	1.41E-05
DN-B-AcB3	0.007855	0.002779	0.012932	1.58E-05
IgM-only-AcB2	0.004204	0.001266	0.007143	0.00012
MZB-1-AcB1	-0.003571	-0.006067	-0.001074	0.00012
PB-AcB4	0.017056	0.004597	0.029515	0.00034
DN-B-AcB1	-0.006387	-0.011422	-0.001351	0.00161
PB-MZB-2	0.013733	0.00203	0.025436	0.00612
PB-Memory	-0.013481	-0.025031	-0.001931	0.00665
PB-AcB3	0.013566	0.001901	0.025231	0.00703
GC-AcB4	-0.005312	-0.010432	-0.000192	0.03308
MZB-1-AcB2	0.002546	-0.00019	0.005283	0.10108
IgM-only-DN-B	0.004474	-0.000519	0.009467	0.1387
DN-A-AcB1	-0.003665	-0.00791	0.000579	0.1831
aNAV-AcB4	-0.004266	-0.009411	0.000879	0.24011
IgM-only-AcB1	-0.001913	-0.004629	0.000804	0.52387
MZB-1-IgM-only	-0.001658	-0.004068	0.000752	0.5656
AcB4-AcB3	-0.00349	-0.008761	0.001781	0.63169
MZB-2-AcB4	0.003323	-0.002032	0.008678	0.73227
MZB-1-DN-B	0.002816	-0.002061	0.007694	0.82301
DN-A-AcB2	0.002452	-0.001939	0.006842	0.85685
PB-AcB2	0.005441	-0.00626	0.017142	0.96377
PB-DN-B	0.005711	-0.006667	0.018089	0.9661
DN-B-DN-A	-0.002721	-0.008686	0.003243	0.96917
GC-aNAV	-0.001046	-0.003443	0.001351	0.9792
IgM-only-DN-A	0.001753	-0.002442	0.005947	0.98599
PB-MZB-1	0.002895	-0.008685	0.014474	0.99995
PB-DN-A	0.002989	-0.009088	0.015067	0.99996
PB-IgM-only	0.001237	-0.010392	0.012866	1
PB-AcB1	-0.000676	-0.012323	0.010971	1
DN-B-AcB2	-0.00027	-0.005429	0.004889	1
MZB-2-AcB3	-0.000167	-0.003252	0.002918	1
MZB-1-DN-A	9.48E-05	-0.003961	0.00415	1

B



C

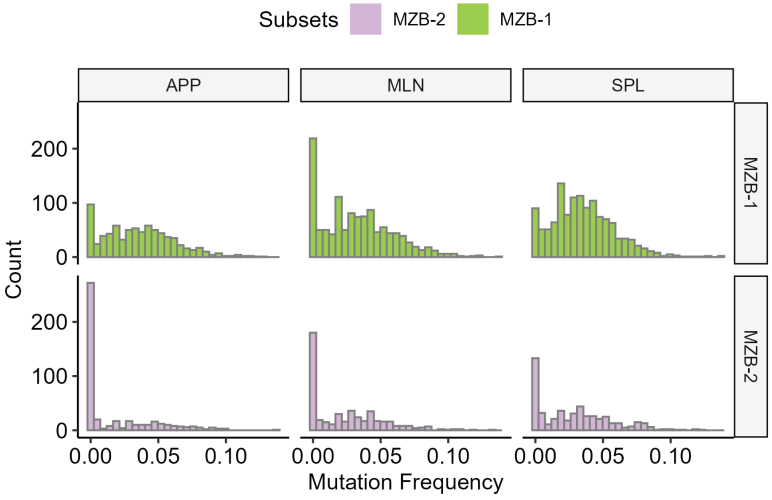


Fig. S3: ANOVA and Tukey post-hoc comparison of mutation frequencies between subsets in tissues After a significant ANOVA test, TukeyHSD post-hoc was performed. **(A)** The multiple comparison results between subsets are shown where 'diff' indicates the difference between observed means, 'lwr' is the lower end point of the interval, 'upr' is the upper end point, and 'p adj' is the p-value after adjustment for multiple comparisons. **(B)** Mutation frequency differences between MZB-1 and MZB-2. **** $p < 0.0001$. **(C)** Distribution of mutation frequency in MZB-1 (top) or MZB-2 (bottom) cells across tissues (columns).

Figure S4

A

Donor	Tissue	Cells captured in BCR library	Mean read pairs/ cell	Number of cells identified in both transcriptome and BCR datasets by unique barcodes
A	Appendix	4937	3435	4279
A	mLN	7602	5130	3642
A	Spleen	7357	5181	4646
B	Appendix	4068	8647	2123
B	mLN	3055	8617	2245
B	Spleen	3942	7354	1582
C	Appendix	4901	5006	3971
C	mLN	4541	5018	3745
C	Spleen	4705	5163	3381

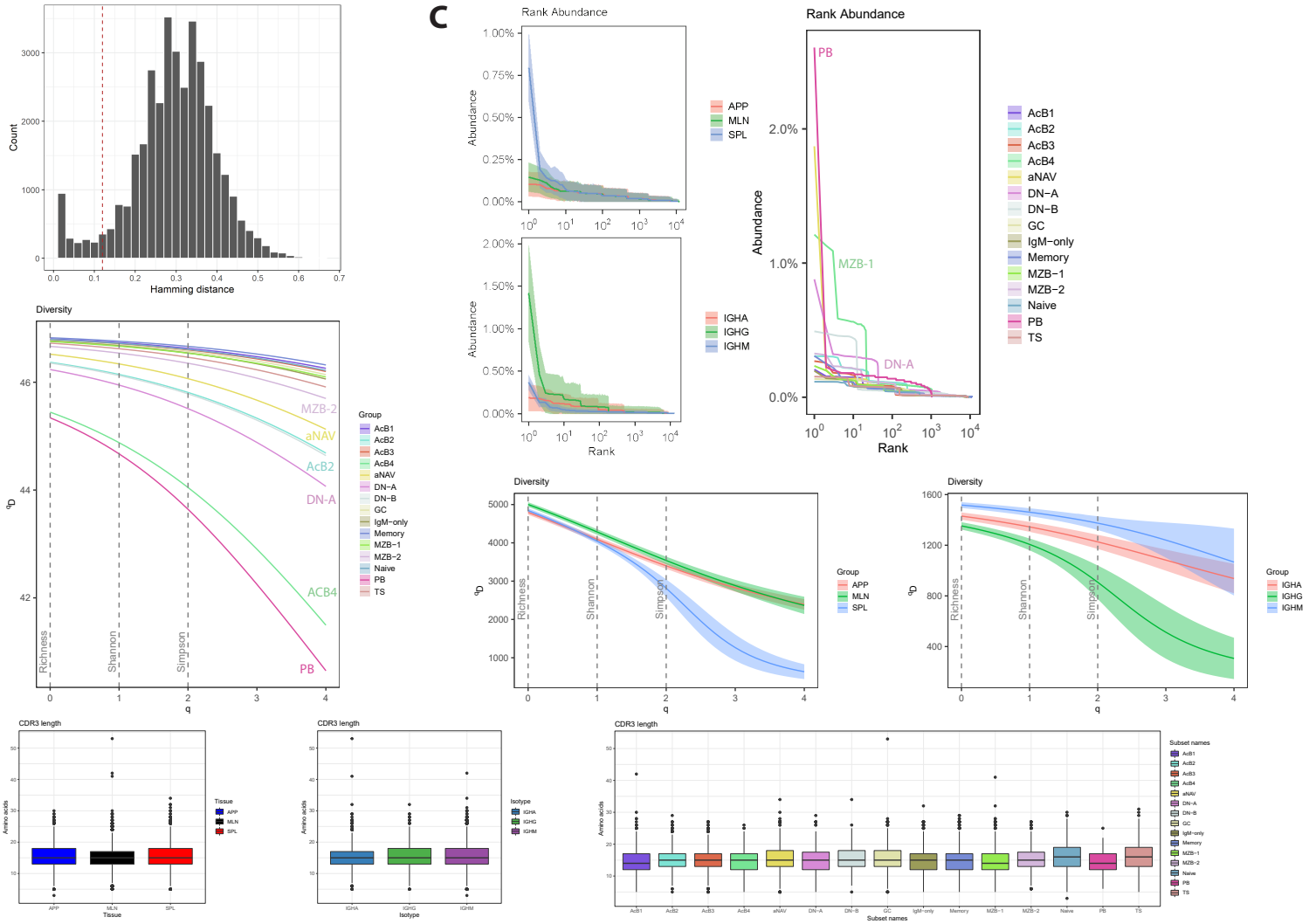


Fig. S4: BCR workflow for single cell analysis, clonal clustering threshold, clonal characteristics (A). Number of cells captured in BCR library, mean read pairs per cell, and number of cells identified in both transcriptomic and BCR datasets for each tissue sample. **(B).** Using the Immcantation pipeline, the distance between sequences and its nearest-neighbour was calculated to determine the threshold for separating clonal groups. The threshold (red dashed line) of 0.15 was determined by inspection of the distance-to-nearest plot for tissues. **(C).** Clone abundance (the size of each as a percent of the repertoire) is plotted against the rank of each clone, where the rank is sorted by size from larger (rank 1, left) to smaller (right). The shaded areas are 95% confidence intervals. For visulisation clarity, confidence intervals are not shown for the subsets (no significance). **(D).** Diversity profile curve that plots Hill diversity scores (^qD) on the y-axis as a function of diversity orders (q). Different values of q represent different measures of diversity: q = 0 (Species Richness), q = 1 (Shannon Entropy), q = 2 (Inverse Simpson Index). The shaded areas represent the 95% confidence interval. For visualisation clarity, confidence intervals are not shown for subsets (no significance). **(E).** Distribution of the CDR3 amino acid length.

Figure S5

A Examples of clones involving naïve and transitional B cells in single human tissues.

Tissue ID and cell type	Cell ID	Clone ID	IGHV (mutations from germline)	IGHD (mutations from germline)	IGHJ (mutations from germline)	IGLV (mutations from germline)
MLN, naive	TCACAAGTCTCGAGTA-1	2875	IGHV4-39*01 (0)	IGHD5-24*01 (0)	IGHJ4*02 (0)	IGLV3-19*01 (0)
MLN, transitional	TCACGAACACACAGAG-1	2875	IGHV4-39*01 (0)	IGHD5-24*01 (0)	IGHJ4*02 (0)	IGLV3-19*01 (0)
MLN, transitional	TCAGGATAGAAGGTGA-1	2875	IGHV4-39*01 (0)	IGHD5-24*01 (0)	IGHJ4*02 (0)	IGLV3-19*01 (0)
MLN, naive	CGTAGGCCACGGCCAT-1	2991	IGHV4-61*01 (0)	IGHD6-19*01 (0)	IGHJ3*02 (0)	IGKV3-11*01 (0)
MLN, transitional	CGTGTAACAGTATAAG-1	2991	IGHV4-61*01 (0)	IGHD6-19*01 (0)	IGHJ3*02 (0)	IGKV3-11*01 (0)
MLN, naive	CTGCTGTAGCCACGTC-1	2991	IGHV4-61*01 (0)	IGHD6-19*01 (0)	IGHJ3*02 (0)	IGKV3-11*01 (0)
APP, transitional	CGTGTCTGTACACCGC-1	5253	IGHV3-21*01 (0)	IGHD3-16*01 (0)	IGHJ6*02 (0)	IGLV3-21*03 (0)
APP, Naive	GGACAGAGTATGCTTG-1	5253	IGHV3-21*01 (0)	IGHD3-16*01 (0)	IGHJ6*02 (0)	IGLV3-21*03 (0)
APP, naive	GTTCTCGTCGGAGGTA-1	5253	IGHV3-21*01 (0)	IGHD3-16*01 (0)	IGHJ6*02 (0)	IGLV3-21*03 (0)
APP, naive	CATTATCTCTTTACAC-1	6828	IGHV7-4-1*02 (0)	IGHD4-17*01 (0)	IGHJ6*02 (0)	IGKV4-1*01 (0)
APP, transitional	CTCATTACATCGATGT-1	6828	IGHV7-4-1*02 (0)	IGHD4-17*01 (0)	IGHJ6*02 (0)	IGKV4-1*01 (0)
APP, transitional	TTGAACGCACCAGGCT-1	6828	IGHV7-4-1*02 (0)	IGHD4-17*01 (0)	IGHJ6*02 (0)	IGKV4-1*01 (0)
MLN, naive	ATGTGTGTCCAAACTG-1	9984	IGHV3-23*01 (0)	IGHD3-22*01 (0)	IGHJ4*02 (0)	(0)
MLN, transitional	CACAGTAAGCCAGGAT-1	9984	IGHV3-23*01 (0)	IGHD3-22*01 (0)	IGHJ4*02 (0)	(0)
MLN, transitional	CAGAGAGGTGCGGTAA-1	9984	IGHV3-23*01 (0)	IGHD3-22*01 (0)	IGHJ4*02 (0)	(0)

B **Result of IMGT junction analysis.** Identical sequence including junctions were observed for all cells sharing a clone ID. Colour code matches table above.

Clone 2875 IGHM

IGHV4-39*01	N1	IGHD5-24*01	N2	IGHJ4*02
tgtgcgagac.	taagtatagtccggga	....agatggctacaat...	gctacacct	..tactttgactactgg

Clone 2875 IGL

IGLV3-19*01	IGLJ2*01
tgtaactcccgggacagcagtggtaaccat..	.gtggtattc

Clone 2991 IGHM

IGHV4-61*01	N1	IGHD6-19*01	N2	P	IGHJ3*02
tgtgcgag...	gcaaat	..gtttagcagtggctggtac	ttcgccgggg	a	tgatgcttttgatatctgg

Clone 2991 IGK

IGKV3-11*01	IGKJ4*01
tgtcagcagcgtagcaactggc....	..tcactttc

Clone 5253 IGHM

IGHV3-21*01	N1	IGHD3-16*01	N2	IGHJ6*02
tgtgcgagag.	ccc	tggctacggggg.....	gggagtat	.....tactactactacggtatggacgtctgg

Clone 5253 IGL

IGLV3-21*03	N	IGLJ3*02
tgtcaggtgtgggatagtagtagtgatc....	c	ttgggtgttc

Clone 6828 IGHM

IGHV7-4-1*02	N	IGHD4-17*01	N	IGHJ6*02
tgtgcgag...	gtg	.gactacggtg.....	tatg	......actactactacggtatggacgtctgg

Clone 6828 IGK

IGKV4-1*01	IGKJ4*01
tgtcagcaatattatagtactcctc.	..tcactttc

Clone 9984 IGHM

IGHV3-23*01	N	IGHD3-22*01	N	IGHJ4*02
tgtgcgaaag.	g	.........tgatagtagtggt.........	cccg	............actgg

Clone 9984 IGK

IGKV3-11*01	N	IGKJ5*01
tgtcagcagcgtagcaactggcct..	agg	....accttc

Fig. S5: Examples of clones involving naïve and transitional B cells (A) Examples of cells in clones that involve naïve and transitional B cells, colour coded by clone. Table includes tissue, subset, 10x cell id (barcode), clone id (determined using Immcantation), the IGH/ IGL alleles, and number of mutations from germline. **(B)** IMGT junction analysis for examples shown above, colour coded by clone.

Figure S6

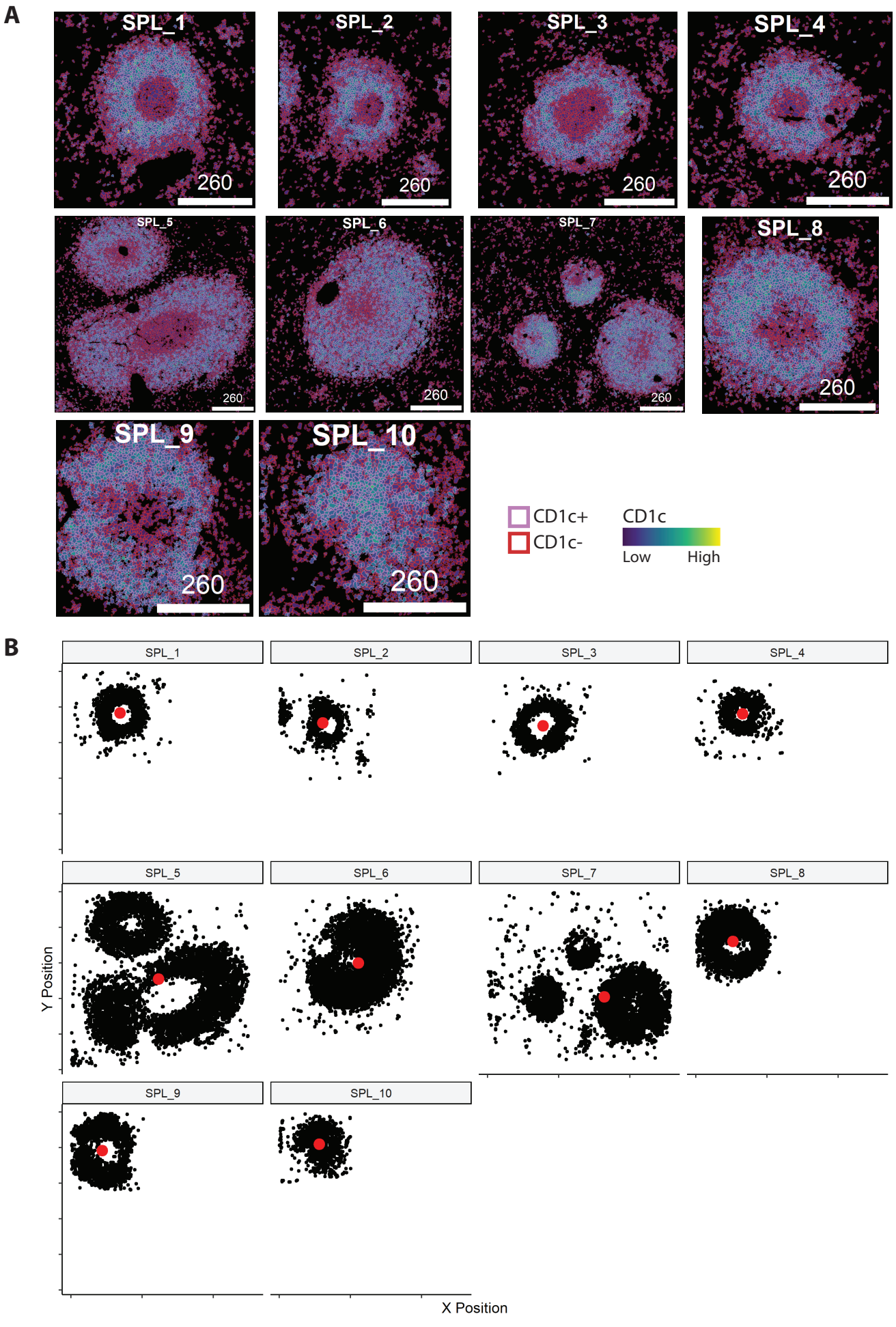


Fig. S6: Masking of CD1c^{+} B cells and centre point calculation for image mass cytometry analysis (A). Visualisation of CD1c$^{+/-}$ cell type (coloured by outline) and CD1c marker expression (coloured by fill) on CD20^{+} segmentation masks. **(B).** Scatterplot indicating the location of every CD1c^{+} B cell (black), and the centre point (average X and Y position) of these cells (red). SPL_5 and SPL_7 were excluded from downstream analysis.

Figure S7

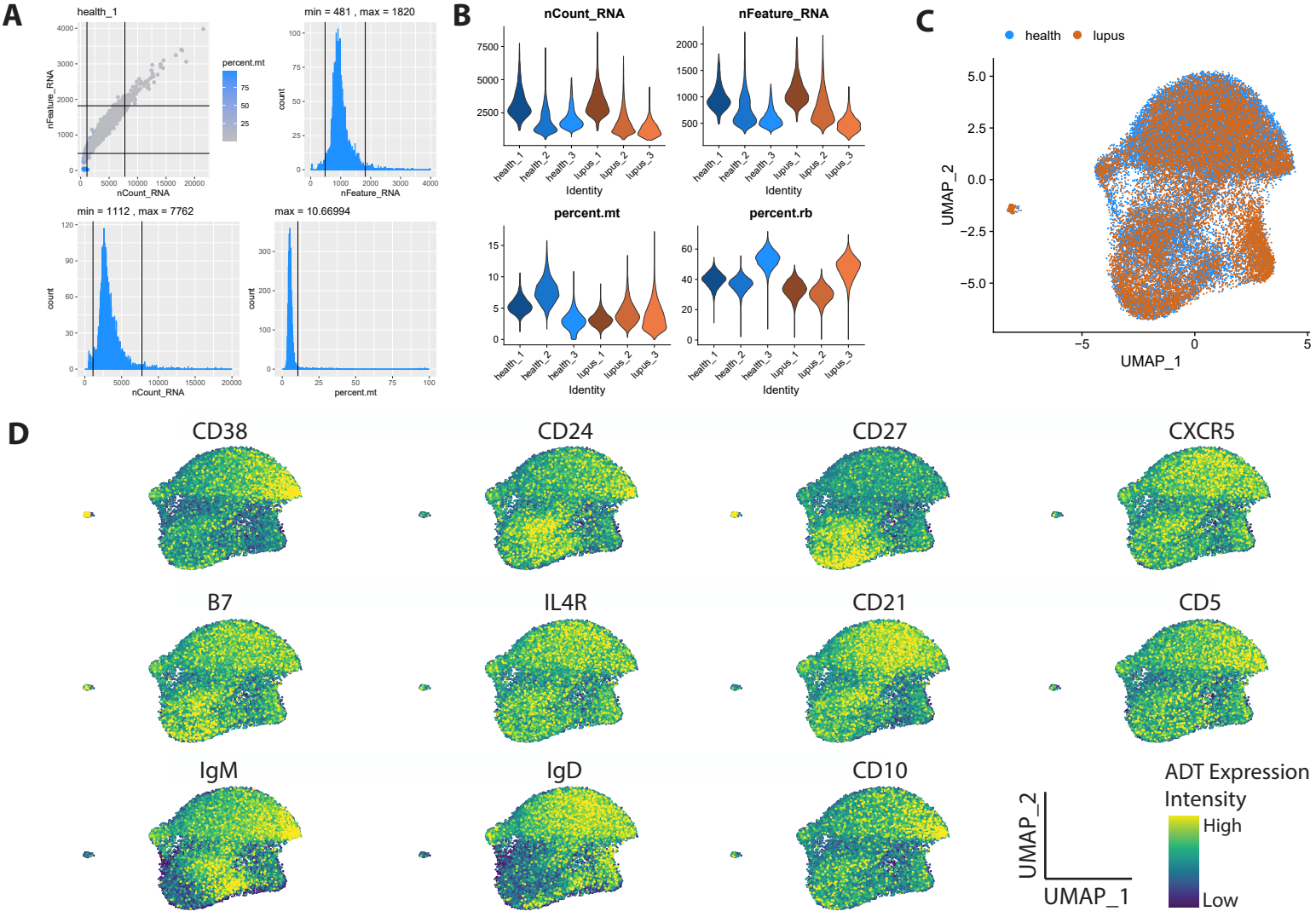


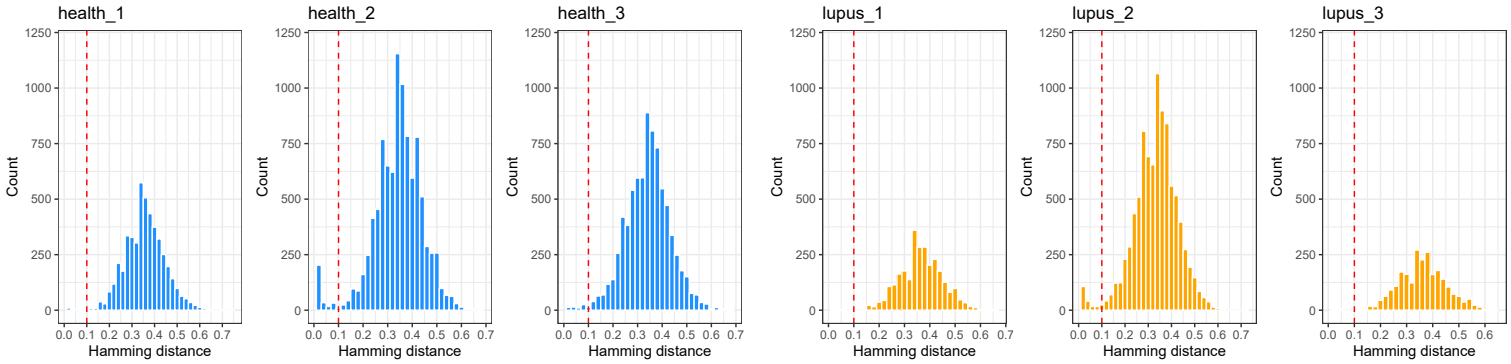
Fig. S7: 10x genomics workflow, integration of B cells in blood, ADT surface protein expression, and mutation frequency statistics (A). Example of quality control processing where thresholds were set automatically based on 3 median absolute deviation. **(B).** Summary of nCount, nFeature, mitochondrial percentage, ribosomal percentage after auto processing. **(C).** UMAP visualisation of B cell composition in blood after integration, coloured by sample. **(D).** UMAP visualisation of ADT surface protein expression.

Figure S8

A

Donor	Cells captured in BCR library	Mean read pairs/ cell	Number of cells identified in both transcriptome and BCR datasets by unique barcodes
HCD 1	5997	7199	5745
HCD 2	12640	5252	8523
HCD 3	14487	5201	13581
SLE 1	4170	11306	3705
SLE 2	12144	5016	9522
SLE 3	6440	4490	5875

B



C

Subset combinations	diff	lwr	up	p adj
DN-AcB3	0.01699	0.00917	0.0248	0
IgM-only-AcB3	0.02886	0.02195	0.03577	0
Memory-AcB3	0.04523	0.03835	0.0521	0
Naive-AcB3	-0.0158	-0.0226	-0.0091	0
PB-AcB3	0.04415	0.03564	0.05265	0
TS-AcB3	-0.0172	-0.024	-0.0104	0
DN-aNAV	0.01913	0.01472	0.02353	0
IgM-only-aNAV	0.031	0.02854	0.03347	0
Memory-aNAV	0.04737	0.04499	0.04975	0
MZB-1-aNAV	0.01113	0.0088	0.01346	0
Naive-aNAV	-0.0137	-0.0156	-0.0118	0
PB-aNAV	0.04629	0.04075	0.05183	0
TS-aNAV	-0.0151	-0.0173	-0.0128	0
IgM-only-DN	0.01187	0.00754	0.01621	0
Memory-DN	0.02824	0.02395	0.03253	0
MZB-1-DN	-0.008	-0.0123	-0.0037	0
MZB-2-DN	-0.0151	-0.0195	-0.0107	0
Naive-DN	-0.0328	-0.0369	-0.0288	0
PB-DN	0.02716	0.02057	0.03375	0
TS-DN	-0.0342	-0.0384	-0.03	0
Memory-IgM-only	0.01637	0.01411	0.01862	0
MZB-1-IgM-only	-0.0199	-0.0221	-0.0177	0
MZB-2-IgM-only	-0.027	-0.0294	-0.0245	0

Subset combinations	diff	lwr	upr	p adj
Naive-IgM-only	-0.0447	-0.0464	-0.043	0
PB-IgM-only	0.01529	0.0098	0.02078	0
TS-IgM-only	-0.0461	-0.0482	-0.0439	0
MZB-1-Memory	-0.0362	-0.0383	-0.0341	0
MZB-2-Memory	-0.0433	-0.0457	-0.041	0
Naive-Memory	-0.0611	-0.0627	-0.0595	0
TS-Memory	-0.0624	-0.0644	-0.0604	0
MZB-2-MZB-1	-0.0071	-0.0094	-0.0048	0
Naive-MZB-1	-0.0248	-0.0264	-0.0233	0
PB-MZB-1	0.03516	0.02973	0.04059	0
TS-MZB-1	-0.0262	-0.0281	-0.0242	0
Naive-MZB-2	-0.0178	-0.0196	-0.0159	0
PB-MZB-2	0.04224	0.03671	0.04777	0
TS-MZB-2	-0.0191	-0.0213	-0.0169	0
PB-Naive	0.06	0.05474	0.06525	0
TS-PB	-0.0613	-0.0667	-0.0559	0
MZB-2-aNAV	0.00405	0.00149	0.00661	2.41E-05
MZB-1-AcB3	0.00899	0.00213	0.01585	0.00141
TS-Naive	-0.0013	-0.0028	6.89E-05	0.07818
aNAV-AcB3	-0.0021	-0.0091	0.00481	0.9936
MZB-2-AcB3	0.00191	-0.005	0.00885	0.99735
PB-Memory	-0.0011	-0.0065	0.00437	0.99981

Fig. S8: BCR workflow for single cell, clonal clustering threshold, mutation frequency statistics in PBMC. (A). Number of cells captured in BCR library, mean read pairs per cell, and number of cells identified in both transcriptomic and BCR datasets for each PBMC sample. **(B).** Using the Immcantation pipeline, the distance between the sequences and its nearest-neighbour was calculated to determine the threshold for separating clonal groups. The threshold of 0.1 was determined by inspection of the distance-to-nearest plot for PBMC. **(C).** After a significant ANOVA test, Tukey HSD post-hoc was performed. The multiple comparison results between subsets is shown where 'diff' indicates the difference between observed means, 'lwr' is the lower end point of the interval, 'upr' is the upper end point, and 'p adj' is the p-value after adjustment for multiple comparisons.

Figure S9

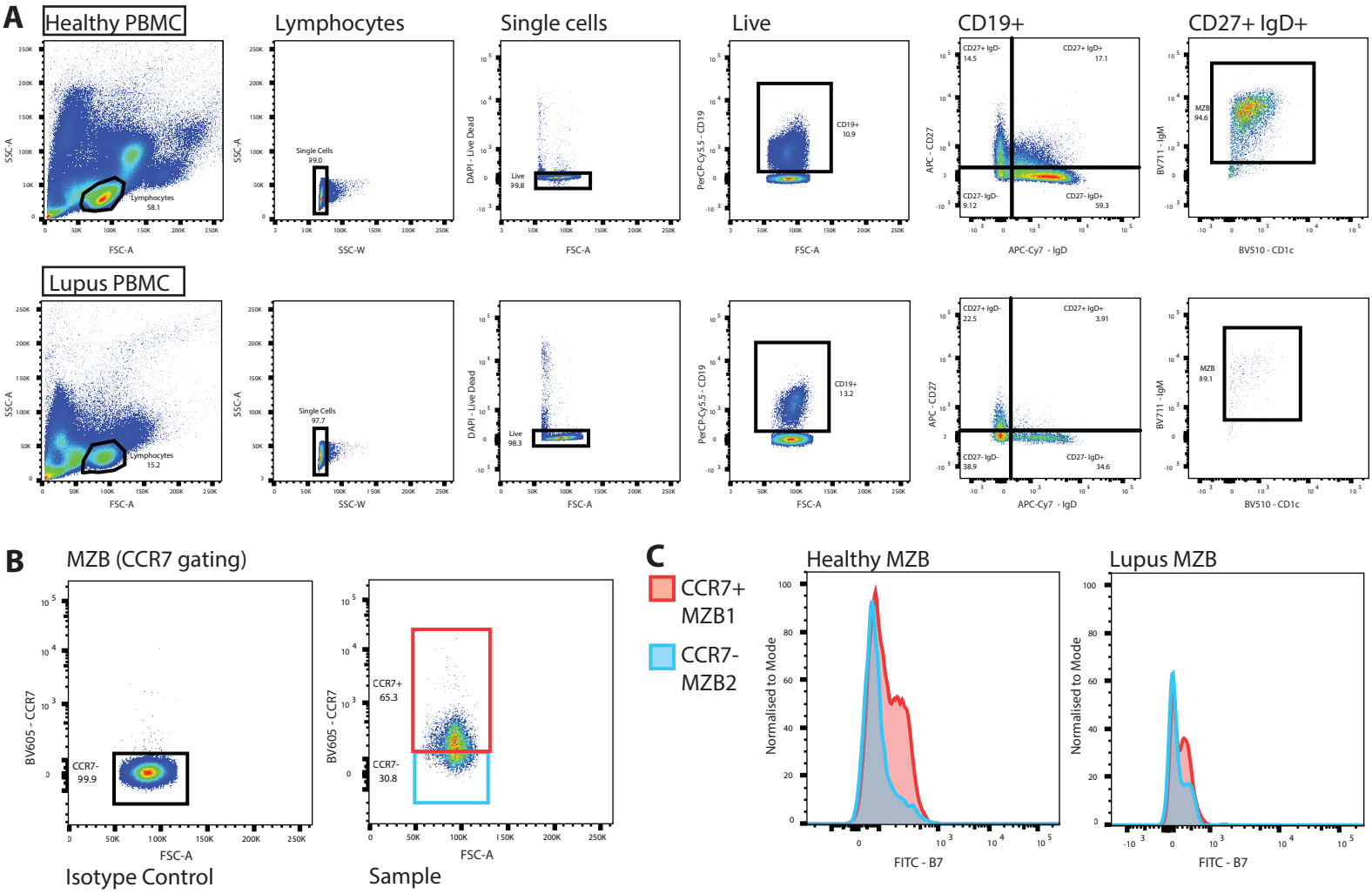


Fig. S9: Flow cytometry workflow for healthy and SLE PBMC (A). Gating strategy for MZB cells in health and lupus PBMC. Lymphocytes were selected from all events, then CD19+ cells were gated from live cells, single cells. Total MZB were gated as CD27+IgD+IgM+CD1c+ **(B).** Within the total MZB gate, MZB1 were gated as CCR7+ (red) and MZB2 were gated as CCR7- (blue), by reference to the isotype control. **(C).** β7 MFI was used for validation. Histogram of ß7 expression in $CCR7^{+}$ MZB-1 (red) and $CCR7^{-}$ MZB-2 subsets (blue).

Table S1

Experiment	Metal/ Antibody	Target	Clone	Working Dilution	Supplier
10x CITE-seq	TotalSeq-C0138	CD5	UCHT2	8 µg/mL	Biolegend
10x CITE-seq	TotalSeq-C0062	CD10	HI10a	8 µg/mL	Biolegend
10x CITE-seq	TotalSeq-C0181	CD21	Bu32	8 µg/mL	Biolegend
10x CITE-seq	TotalSeq-C0180	CD24	ML5	8 µg/mL	Biolegend
10x CITE-seq	TotalSeq-C0154	CD27	O323	8 µg/mL	Biolegend
10x CITE-seq	TotalSeq-C0389	CD38	HIT-2	8 µg/mL	Biolegend
10x CITE-seq	Totalseq-C0214	Integrin ß7	FIB504	8 µg/mL	Biolegend
10x CITE-seq	TotalSeq-C0363	CD124 (IL-4Rα)	G077F6	8 µg/mL	Biolegend
10x CITE-seq	TotalSeq-C0136	IgM	MHM-88	8 µg/mL	Biolegend
10x CITE-seq	TotalSeq-C0384	IgD	IA62	8 µg/mL	Biolegend
10x CITE-seq	TotalSeq-C0144	CXCR5	J252D4	8 µg/mL	Biolegend
CyTOF	158Gb	CD10	HI10a	1:200	Fluidigm
CyTOF	143Nd	CD127 (IL-7R)	A019D5	1:200	Fluidigm
CyTOF	163Dy	CD180	MHR73-11	1:200	Biolegend
CyTOF	153Eu	CD185 (CXCR5)	RF8B2	1:200	Fluidigm
CyTOF	142Nd	CD19	HIB19	1:200	Fluidigm
CyTOF	159Tb	CD197 (CCR7)	G043H7	1:200	Fluidigm
CyTOF	171Yb	CD20	2H7	1:200	Fluidigm
CyTOF	166Er	CD24	ML5	1:100	Fluidigm
CyTOF	169Tm	CD25	2A3	1:100	Fluidigm
CyTOF	164Dy	CD267 (TACI)	1A1	1:200	Biolegend
CyTOF	155Gd	CD268 (BAFFR)	11C1	1:200	Fluidigm
CyTOF	173Yb	CD269 (BCMA)	19F2	1:100	Biolegend
CyTOF	167Er	CD27	O323	1:200	Fluidigm
CyTOF	174Yb	CD279 (PD-1)	EH12.2H7	1:200	Fluidigm
CyTOF	154Sm	CD3	UCHT1	1:200	Fluidigm
CyTOF	170Er	CD307d (FcRL4)	413D12	1:200	Biolegend
CyTOF	165Ho	CD307e (FcRL5)	413D12	1:200	Biolegend
CyTOF	144Nd	CD38	HIT2	1:200	Fluidigm
CyTOF	176Yb	CD4	RPA-T4	1:200	Fluidigm
CyTOF	162Dy	CD40	5C3	1:200	Biolegend
CyTOF	89Y	CD45	HI30	1:200	Fluidigm
CyTOF	145Nd	CD45RB	MEM-55	1:100	Fluidigm
CyTOF	168Er	CD8	SK1	1:200	Fluidigm
CyTOF	161Dy	CD80 (B7-1)	2D10.4	1:200	Fluidigm
CyTOF	148Nd	IgA	Polyclonal	1:200	Fluidigm
CyTOF	146Nd	IgD	IA6-2	1:200	Fluidigm
CyTOF	172Yb	IgM	MHM-88	1:100	Fluidigm
CyTOF	141Pr	IgG	Polyclonal	1:100	Fluidigm
IMC	153Eu	DDX21	EPR14495	1:400	Abcam + in-house conjugation
IMC	154Sm	IgD	EPR6416	1:300	Abcam + in-house conjugation
IMC	159Tb	CD68	KP1	1:400	Fluidigm
IMC	161Dy	CD20	H1	1:250	Fluidigm
IMC	164Dy	CD1c	OTI2F4	1:200	Abcam + in-house conjugation
IMC	167Er	GzB	EPR20129-217	1:300	Fluidigm

IMC	168Er	Ki67	B56	1:400	Fluidigm
IMC	170Er	CD3	Polyclonal, C-Terminal	1:800	Fluidigm
IMC	171Yb	CD27	EPR8569	1:300	Fluidigm
IMC	175Lu	CD25	EPR6452	1:50	Fluidigm
Flow cytometry	DAPI	L/D	na	0.1ug/mL diluted 1:1000	Invitrogen
Flow cytometry	PerCP Cy5.5	CD19	HIB19	1:50	Biolegend
Flow cytometry	APC	CD27	M-T271	1:50	Biolegend
Flow cytometry	APC Cy7	IgD	IA6-2	1:50	Biolegend
Flow cytometry	BV711	IgM	MHM-88	1:50	Biolegend
Flow cytometry	BV510	CD1c	L161	1:50	Biolegend
Flow cytometry	BV605	CCR7	G043H7	1:25	Biolegend
Flow cytometry	FITC	B7	FIB504	1:50	Biolegend
Flow cytometry	APC	isotype	MOPC-21	1:50	Biolegend
Flow cytometry	APC Cy7	isotype	MOPC-173	1:50	Biolegend
Flow cytometry	BV711	isotype	MOPC-21	1:50	Biolegend
Flow cytometry	BV605	isotype	MOPC-173	1:25	Biolegend

Table S2

Experiment	Donor	Disease	DCD/ DBD	Cause of Death	Age	Sex	Ethnicity	Auto-antibodies	Medication	Tissue	Cells captured for 10x	Sequencing depth for 10x	Median genes/ cell for 10x	Number of reads for 10x
(internal control)	Ctrl SPL	Healthy	DCD	Respiratory failure	63	M	--	--	--	Spleen	--	--	--	--
CyTOF, 10x	A	Healthy	DBD	Intracranial haemorrhage	57	M	--	--	--	Appendix	5082	13285	696	81,463,015
CyTOF, 10x	A	Healthy	DBD	Intracranial haemorrhage	57	M	--	--	--	mLN	7088	17438	502	211,360,416
CyTOF, 10x	A	Healthy	DBD	Intracranial haemorrhage	57	M	--	--	--	Spleen	5458	19804	748	212,673,535
CyTOF, 10x	B	Healthy	DBD	Trauma - Accident	77	F	--	--	--	Appendix	3963	16161	589	84,552,949
CyTOF, 10x	B	Healthy	DBD	Trauma - Accident	77	F	--	--	--	mLN	2978	17462	610	82,788,540
CyTOF, 10x	B	Healthy	DBD	Trauma - Accident	77	F	--	--	--	Spleen	3447	20020	716	144,666,189
CyTOF, 10x	C	Healthy	DBD	Intracranial haemorrhage	50	M	--	--	--	Appendix	5156	19675	882	104,752,104
CyTOF, 10x	C	Healthy	DBD	Intracranial haemorrhage	50	M	--	--	--	mLN	4214	21829	851	186,466,848
CyTOF, 10x	C	Healthy	DBD	Intracranial haemorrhage	50	M	--	--	--	Spleen	4110	20969	1223	217,759,551
CyTOF	D	Healthy	DCD	Intracranial haemorrhage	54	M	--	--	--	APP, MLN, SPL	--	--	--	--
CyTOF	E	Healthy	DCD	Hypoxic brain damage	44	F	--	--	--	APP, MLN, SPL	--	--	--	--
CyTOF	F	Healthy	DBD	Intracranial haemorrhage	49	M	--	--	--	APP, MLN, SPL	--	--	--	--
CyTOF	G	Healthy	DCD	Trauma - Accident	60	M	--	--	--	APP, MLN, SPL	--	--	--	--
CyTOF	H	Healthy	DBD	Intracranial haemorrhage	70	F	--	--	--	APP, MLN, SPL	--	--	--	--
10x	HCD 1	Healthy	--	--	34	F	Caucasian	--	--	Blood	4138	19764	934	117,398,259
10x	HCD 2	Healthy	--	--	43	F	Caucasian	--	--	Blood	6849	21428	681	191,544,113
10x	HCD 3	Healthy	--	--	42	F	African Caribbean	--	--	Blood	11445	20030	566	281,804,567
10x	SLE 1	SLE, lupus nephritis	--	--	31	F	African Caribbean	dsDNA, Sm, RNP,	HCQ, MMF	Blood	3825	27382	1051	104,734,793
10x	SLE 2	SLE, inflam arthritis, cutaneous lupus	--	--	48	F	Caucasian	ANA, DNA, RNP	HCQ	Blood	10260	20052	722	205,734,921
10x	SLE 3	SLE, lupus nephritis	--	--	30	F	African Caribbean	ANA, Sm, RNP, La, C1Q & SLE, LN	HCQ, MMF, PRED	Blood	6000	13057	470	78,341,625
Flow cytometry	Health1	Healthy	--	--	33	F		--	--	Blood	--	--	--	--
Flow cytometry	Health2	Healthy	--	--	34	F	Asian Indian	--	--	Blood	--	--	--	--
Flow cytometry	Health3	Healthy	--	--	30	F	Caucasian	--	--	Blood	--	--	--	--
Flow cytometry	Health4	Healthy	--	--	26	F	Caucasian	--	--	Blood	--	--	--	--
Flow cytometry	Health5	Healthy	--	--	29	F	SE Asian	--	--	Blood	--	--	--	--
Flow cytometry	Health6	Healthy	--	--	35	F	Caucasian	--	--	Blood	--	--	--	--
Flow cytometry	Health7	Healthy	--	--	32	F	Caucasian	--	--	Blood	--	--	--	--

Flow cytometry	Health8	Healthy	--	--	36	F	South American	--	--	Blood	--	--	--	--
Flow cytometry	Health9	Healthy	--	--	42	F	Hispanic	--	--	Blood	--	--	--	--
Flow cytometry	Health10	Healthy	--	--	28	F	African Caribbean	--	--	Blood	--	--	--	--
Flow cytometry	Lupus1	SLE, lupus nephritis	--	--	35	F		ANA, Sm, RNP	HCQ, PRED	Blood	--	--	--	--
Flow cytometry	Lupus2	SLE, lupus nephritis	--	--	68	F		ANA, DNA, Sm	HCQ, PRED	Blood	--	--	--	--
Flow cytometry	Lupus3	SLE, lupus nephritis	--	--	46	F		ANA, Sm, RNP	HCQ, MMF	Blood	--	--	--	--
Flow cytometry	Lupus4	SLE, lupus nephritis	--	--	47	F		ANA, DNA	HCQ, PRED	Blood	--	--	--	--
Flow cytometry	Lupus5	SLE, lupus nephritis	--	--	54	F		ANA, DNA, Ro	PRED	Blood	--	--	--	--
Flow cytometry	Lupus6	SLE, lupus nephritis	--	--	36	F		ANA, DNA, Ro	HCQ, MMF, PRED	Blood	--	--	--	--
Flow cytometry	Lupus7	SLE, lupus nephritis	--	--	43	F		ANA, DNA, Sm, RNP	HCQ, MMF, PRED	Blood	--	--	--	--
Flow cytometry	Lupus8	SLE, lupus nephritis	--	--	49	F		ANA, DNA, Ro, C1Q	HCQ, MMF, PRED	Blood	--	--	--	--
Flow cytometry	Lupus9	SLE, lupus nephritis	--	--	65	F		ANA, DNA, Ro, Sm, RNP, La	HCQ, PRED	Blood	--	--	--	--
Flow cytometry	Lupus10	SLE, lupus nephritis	--	--	24	F		ANA, DNA, Sm	MMF, PRED	Blood	--	--	--	--

Table S3

	p_val	avg_log2FC	pct.1	pct.2	p_val_adj	subset
JUN	0	1.918973887	0.795	0.362	0	AcB1
RGS2	0	1.160348318	0.223	0.043	0	AcB1
DUSP1	0	1.553727411	0.619	0.263	0	AcB1
KLF6	0	1.393243606	0.703	0.365	0	AcB1
NR4A1	0	1.553224913	0.321	0.087	0	AcB1
JUNB	0	1.915214657	0.71	0.304	0	AcB1
FOSB	0	1.434296997	0.397	0.107	0	AcB1
NR4A2	3.59E-286	1.170112686	0.265	0.065	6.13E-282	AcB1
RHOB	3.37E-265	1.250168384	0.361	0.12	5.75E-261	AcB1
H3F3B	4.85E-261	1.073737525	0.894	0.714	8.28E-257	AcB1
KLF2	1.77E-258	1.183873897	0.545	0.249	3.02E-254	AcB1
FOS	6.42E-252	1.271395462	0.527	0.236	1.10E-247	AcB1
PPP1R15A	4.08E-241	1.061532291	0.396	0.148	6.96E-237	AcB1
GPR183	6.67E-231	1.232951124	0.524	0.252	1.14E-226	AcB1
RGS1	3.28E-213	1.25736591	0.234	0.064	5.61E-209	AcB1
CD69	4.23E-204	1.011528571	0.691	0.422	7.23E-200	AcB1
AREG	4.51E-179	0.819335765	0.262	0.085	7.71E-175	AcB1
YPEL5	8.43E-176	1.022446455	0.435	0.207	1.44E-171	AcB1
UBC	2.69E-166	0.872452718	0.794	0.588	4.60E-162	AcB1
JUND	5.35E-159	0.827740227	0.321	0.128	9.14E-155	AcB1
MTRNR2L1	0	1.526653519	0.912	0.635	0	AcB2
AL138963.3	0	1.940577951	0.635	0.175	0	AcB2
TPT1	9.11E-223	0.493703506	0.99	0.979	1.56E-218	AcB2
HIST1H1E	1.09E-216	1.509194392	0.579	0.312	1.87E-212	AcB2
MT-ND6	8.18E-152	1.269323934	0.658	0.404	1.40E-147	AcB2
RPL13A	4.66E-113	0.479198633	0.959	0.946	7.96E-109	AcB2
MTRNR2L8	2.28E-111	1.35085275	0.239	0.091	3.89E-107	AcB2
MT-CO1	8.32E-108	0.406976123	0.997	0.993	1.42E-103	AcB2
RPL4	1.28E-98	0.53145297	0.851	0.773	2.18E-94	AcB2
MT-ND5	6.38E-75	0.581695823	0.922	0.834	1.09E-70	AcB2
RPS20	5.81E-73	0.54596612	0.81	0.78	9.92E-69	AcB2
RPL27A	3.63E-71	0.517714846	0.822	0.792	6.21E-67	AcB2
MT-ATP8	5.58E-61	0.703957279	0.751	0.665	9.53E-57	AcB2
EEF2	1.46E-33	0.354288126	0.775	0.742	2.50E-29	AcB2
DDX5	2.11E-23	0.342283696	0.636	0.584	3.60E-19	AcB2
ACTN4	2.53E-17	0.369424351	0.12	0.071	4.32E-13	AcB2
HIST1H1C	4.88E-17	0.497977093	0.279	0.213	8.34E-13	AcB2
Sep-09	4.36E-14	0.394035773	0.388	0.326	7.44E-10	AcB2
CR2	5.00E-14	0.439113052	0.227	0.172	8.55E-10	AcB2
PLCG2	1.12E-13	0.551904008	0.238	0.18	1.91E-09	AcB2
PRRC2A	4.34E-12	0.29622535	0.106	0.067	7.42E-08	AcB2
RPL17	6.10E-12	0.33256581	0.634	0.626	1.04E-07	AcB2
TGOLN2	1.34E-10	0.357648911	0.203	0.155	2.30E-06	AcB2
OGT	1.12E-09	0.306952093	0.114	0.078	1.92E-05	AcB2
BRD4	7.09E-09	0.331364807	0.149	0.11	0.000121	AcB2
MYH9	1.24E-07	0.40281884	0.291	0.25	0.002122	AcB2
MALAT1	0	1.273655795	1	0.997	0	AcB3
DDX17	5.90E-161	1.284650939	0.528	0.323	1.01E-156	AcB3

PTPRC	8.76E-138	0.947561513	0.741	0.611	1.50E-133	AcB3
FAM129C	8.50E-121	1.204668659	0.453	0.274	1.45E-116	AcB3
LINC00926	3.62E-119	1.047221231	0.64	0.495	6.19E-115	AcB3
RSRP1	1.13E-112	1.108763249	0.454	0.286	1.93E-108	AcB3
NKTR	2.91E-109	1.179124074	0.285	0.138	4.97E-105	AcB3
SMCHD1	2.34E-108	1.083917463	0.539	0.38	4.00E-104	AcB3
GABPB1-AS	2.79E-105	1.251422108	0.363	0.204	4.77E-101	AcB3
ANKRD44	3.16E-105	1.114552057	0.416	0.252	5.39E-101	AcB3
MT-CO1	1.52E-100	0.491610018	0.991	0.993	2.60E-96	AcB3
ATM	5.90E-96	1.094538381	0.34	0.189	1.01E-91	AcB3
ARGLU1	2.18E-95	0.927375939	0.518	0.37	3.72E-91	AcB3
NEAT1	5.50E-95	1.29427232	0.37	0.221	9.39E-91	AcB3
TXNIP	4.74E-93	0.606811507	0.9	0.904	8.09E-89	AcB3
AFF3	6.74E-88	0.953515481	0.473	0.329	1.15E-83	AcB3
POLR2J3.1	1.19E-87	0.774951877	0.1	0.028	2.03E-83	AcB3
PARP15	5.38E-85	1.145803281	0.325	0.188	9.19E-81	AcB3
PNISR	1.55E-84	0.979050897	0.558	0.437	2.65E-80	AcB3
CELF2	1.14E-83	1.062000623	0.403	0.261	1.94E-79	AcB3
AC119396.	7.84E-83	0.959305708	0.356	0.213	1.34E-78	AcB3
IFITM1	1.35E-195	2.424273402	0.914	0.399	2.31E-191	AcB4
GBP4	5.52E-192	1.530856123	0.377	0.05	9.43E-188	AcB4
STAT1	3.87E-154	1.634747644	0.517	0.111	6.62E-150	AcB4
GBP1	3.54E-121	0.945361418	0.206	0.023	6.04E-117	AcB4
EPSTI1	1.13E-56	0.825293641	0.267	0.068	1.94E-52	AcB4
IRF1	2.38E-56	1.407360053	0.61	0.308	4.06E-52	AcB4
UBE2L6	8.62E-52	1.045906263	0.48	0.2	1.47E-47	AcB4
ISG15	4.35E-43	1.114207309	0.25	0.074	7.43E-39	AcB4
IFITM2	1.53E-40	0.94325694	0.76	0.534	2.62E-36	AcB4
SAMD9L	1.44E-37	0.669553976	0.174	0.043	2.46E-33	AcB4
PSME2	2.17E-35	0.940249246	0.603	0.366	3.71E-31	AcB4
GBP2	1.84E-33	0.587257079	0.115	0.023	3.15E-29	AcB4
GBP5	3.59E-31	0.464985866	0.127	0.029	6.14E-27	AcB4
XAF1	8.89E-29	0.56824201	0.169	0.049	1.52E-24	AcB4
APOL6	9.32E-27	0.411991349	0.11	0.025	1.59E-22	AcB4
PLSCR1	8.61E-26	0.614295904	0.164	0.05	1.47E-21	AcB4
IFI6	1.71E-24	0.629304326	0.157	0.048	2.92E-20	AcB4
PSMB9	8.00E-24	0.699723358	0.664	0.483	1.37E-19	AcB4
HERC5	1.33E-23	0.599501751	0.135	0.039	2.27E-19	AcB4
ISG20	2.19E-23	0.859813003	0.664	0.477	3.74E-19	AcB4
CD74	0	0.579726299	1	0.996	0	aNAV
HLA-DRB1	0	0.889614733	0.977	0.922	0	aNAV
HLA-DRA	1.57E-290	0.644824107	0.982	0.961	2.69E-286	aNAV
HLA-DPA1	1.35E-242	0.663882603	0.947	0.9	2.30E-238	aNAV
HLA-DRB5	3.56E-228	0.886645894	0.829	0.682	6.09E-224	aNAV
HLA-DQA1	2.41E-182	0.755943084	0.847	0.756	4.12E-178	aNAV
IGHM	2.47E-179	1.176748382	0.639	0.462	4.22E-175	aNAV
HLA-DPB1	8.42E-168	0.547219275	0.938	0.893	1.44E-163	aNAV
IGHD	2.36E-155	0.965159842	0.366	0.18	4.03E-151	aNAV
HLA-DQB1	2.31E-147	0.617014051	0.843	0.76	3.94E-143	aNAV
CD79A	2.66E-138	0.48819568	0.953	0.9	4.54E-134	aNAV

MS4A1	4.01E-129	0.522931505	0.907	0.854	6.85E-125	aNAV
HLA-DMA	1.25E-66	0.535251225	0.612	0.52	2.13E-62	aNAV
CD72	8.71E-59	0.620124523	0.333	0.221	1.49E-54	aNAV
SYNGR2	1.51E-55	0.53181455	0.609	0.522	2.58E-51	aNAV
HLA-DQA2	3.91E-50	0.707946473	0.35	0.251	6.68E-46	aNAV
CYBA	6.50E-48	0.307560248	0.889	0.849	1.11E-43	aNAV
CD79B	9.93E-44	0.388456547	0.698	0.643	1.70E-39	aNAV
FCER2	1.39E-42	0.598042192	0.399	0.301	2.37E-38	aNAV
ACTB	5.65E-37	0.374972069	0.95	0.937	9.65E-33	aNAV
PFN1	2.62E-36	0.277950469	0.858	0.846	4.47E-32	aNAV
LAPTM5	6.52E-34	0.317564226	0.783	0.758	1.11E-29	aNAV
CALR	2.05E-31	0.572525914	0.349	0.273	3.50E-27	aNAV
DUSP2	0	2.949404502	0.661	0.123	0	DN-A
EGR3	0	0.76876104	0.128	0.005	0	DN-A
NFKBID	0	2.198077263	0.436	0.044	0	DN-A
CCL4	1.14E-249	2.650680298	0.236	0.026	1.94E-245	DN-A
EGR1	3.33E-241	1.414084405	0.205	0.02	5.69E-237	DN-A
CD83	5.49E-233	2.242036481	0.744	0.286	9.38E-229	DN-A
NR4A1	6.60E-163	1.927945203	0.395	0.096	1.13E-158	DN-A
NFKBIA	2.42E-122	1.686313569	0.606	0.279	4.14E-118	DN-A
CCL3	4.05E-102	1.206015937	0.119	0.015	6.91E-98	DN-A
BCL2A1	4.42E-101	1.603472727	0.39	0.137	7.55E-97	DN-A
CD69	5.58E-98	1.826727935	0.717	0.434	9.54E-94	DN-A
KDM6B	1.61E-95	0.801685737	0.206	0.044	2.74E-91	DN-A
ZC3H12A	6.60E-88	0.741469767	0.169	0.033	1.13E-83	DN-A
TNFAIP3	2.48E-81	0.652269937	0.155	0.03	4.23E-77	DN-A
SGK1	1.54E-75	0.695820777	0.131	0.023	2.63E-71	DN-A
NR4A2	1.11E-71	1.112214465	0.251	0.075	1.89E-67	DN-A
TNF	4.17E-70	0.740891445	0.147	0.031	7.12E-66	DN-A
TRAF4	4.51E-68	0.721130576	0.198	0.052	7.71E-64	DN-A
IER2	2.69E-64	1.386344433	0.499	0.259	4.60E-60	DN-A
MIR155HG	1.52E-60	1.019153522	0.192	0.054	2.60E-56	DN-A
AREG	0	4.02456895	0.996	0.086	0	DN-B
NR4A1	1.07E-128	1.722590652	0.44	0.098	1.82E-124	DN-B
FOS	1.53E-104	1.650491717	0.667	0.25	2.61E-100	DN-B
FOSB	3.58E-78	1.23610876	0.409	0.122	6.11E-74	DN-B
DUSP1	6.09E-69	1.438417866	0.607	0.283	1.04E-64	DN-B
NR4A2	5.68E-68	0.936030191	0.298	0.075	9.70E-64	DN-B
PPP1R15A	7.25E-68	1.168909894	0.453	0.161	1.24E-63	DN-B
JUNB	1.10E-67	1.373539325	0.656	0.326	1.88E-63	DN-B
BTG1	2.08E-62	0.874676383	0.96	0.908	3.55E-58	DN-B
CD83	3.17E-60	1.272129507	0.598	0.292	5.41E-56	DN-B
CD69	1.69E-52	1.32770539	0.704	0.436	2.89E-48	DN-B
RGS2	1.48E-51	0.908009603	0.216	0.053	2.53E-47	DN-B
KLF6	6.30E-51	1.001883698	0.671	0.384	1.08E-46	DN-B
KLF2	1.00E-49	1.116411608	0.547	0.265	1.71E-45	DN-B
ZNF331	3.77E-48	0.960674492	0.256	0.074	6.45E-44	DN-B
JUN	1.03E-47	1.223356316	0.649	0.387	1.76E-43	DN-B
RGS1	2.33E-46	0.894080083	0.253	0.073	3.99E-42	DN-B
H3F3B	1.25E-42	0.768217348	0.882	0.724	2.13E-38	DN-B

RHOB	2.44E-39	0.830337843	0.344	0.133	4.17E-35	DN-B
LY9	4.72E-39	0.956348142	0.411	0.185	8.07E-35	DN-B
DUSP1	0	1.674798833	0.587	0.258	0	GC
CD83	0	1.428452861	0.583	0.268	0	GC
KLF6	0	1.205545364	0.648	0.363	0	GC
CD69	0	1.942433978	0.788	0.406	0	GC
BTG1	0	0.94897949	0.97	0.903	0	GC
FOS	0	2.452640561	0.736	0.209	0	GC
JUNB	0	1.537680409	0.614	0.303	0	GC
KLF2	0	1.261945264	0.529	0.243	0	GC
FOSB	1.82E-300	1.069872439	0.327	0.107	3.12E-296	GC
JUN	6.33E-298	1.224291461	0.641	0.367	1.08E-293	GC
NR4A2	2.48E-293	0.954440731	0.241	0.062	4.23E-289	GC
NR4A1	3.16E-284	1.188197338	0.283	0.085	5.40E-280	GC
SLC2A3	3.63E-261	1.048457783	0.464	0.218	6.21E-257	GC
PPP1R15A	6.32E-222	0.906750489	0.353	0.146	1.08E-217	GC
YPEL5	9.48E-203	1.034613309	0.411	0.204	1.62E-198	GC
CLEC2B	8.23E-201	1.181871743	0.284	0.109	1.41E-196	GC
CXCR4	4.54E-196	0.916221968	0.765	0.61	7.75E-192	GC
ARL4A	6.87E-195	0.852965453	0.247	0.087	1.17E-190	GC
TSC22D3	1.49E-188	0.869837979	0.736	0.566	2.54E-184	GC
LY9	3.43E-184	0.909752535	0.364	0.17	5.86E-180	GC
LGALS1	0	2.034781856	0.441	0.12	0	IgM-only
B2M	1.72E-103	0.253039453	0.995	0.991	2.95E-99	IgM-only
FAU	1.29E-100	0.287386101	0.967	0.957	2.21E-96	IgM-only
KLK1	1.67E-95	0.742854109	0.232	0.11	2.85E-91	IgM-only
LINC01857	2.35E-50	0.507166424	0.351	0.238	4.01E-46	IgM-only
CD27	8.32E-31	0.306400658	0.343	0.249	1.42E-26	IgM-only
LINC01781	0	1.274435721	0.496	0.183	0	Memory
S100A4	0	1.57226633	0.55	0.234	0	Memory
JCHAIN	0	1.942766217	0.474	0.188	0	Memory
RPS14	0	0.362751528	0.987	0.967	0	Memory
IGHA2	0	1.729737119	0.179	0.03	0	Memory
IGHA1	0	2.131407143	0.361	0.112	0	Memory
SSPN	1.36E-231	0.505398213	0.138	0.036	2.33E-227	Memory
CD27	7.08E-216	0.568968067	0.42	0.225	1.21E-211	Memory
S100A6	4.04E-213	0.708680057	0.448	0.257	6.90E-209	Memory
VIM	5.36E-177	0.576636574	0.547	0.361	9.16E-173	Memory
AIM2	3.56E-172	0.499873073	0.287	0.141	6.09E-168	Memory
TNFRSF13B	3.47E-120	0.415128121	0.315	0.182	5.93E-116	Memory
SCIMP	1.01E-101	0.397470068	0.27	0.156	1.72E-97	Memory
HLA-C	1.50E-101	0.301165491	0.894	0.821	2.56E-97	Memory
SELENOM	2.44E-99	0.352455105	0.232	0.126	4.17E-95	Memory
COTL1	6.62E-98	0.362027584	0.535	0.382	1.13E-93	Memory
CD24	2.94E-85	0.36619351	0.328	0.211	5.03E-81	Memory
ITGB1	1.50E-68	0.372343748	0.14	0.072	2.56E-64	Memory
S100A10	5.92E-63	0.38217089	0.245	0.156	1.01E-58	Memory
LINC01857	2.82E-163	0.644859202	0.411	0.226	4.82E-159	MZB-1
CD27	5.60E-139	0.533958943	0.415	0.236	9.57E-135	MZB-1
CD1C	2.60E-129	0.580367718	0.233	0.106	4.45E-125	MZB-1

ITM2C	3.33E-126	0.541722076	0.295	0.151	5.69E-122	MZB-1
TNFRSF13B	1.12E-111	0.465667888	0.335	0.186	1.91E-107	MZB-1
AIM2	5.81E-94	0.41883943	0.276	0.15	9.93E-90	MZB-1
PTPN18	2.70E-89	0.38990098	0.208	0.103	4.61E-85	MZB-1
COTL1	5.82E-78	0.435384381	0.534	0.39	9.95E-74	MZB-1
GALNTL6	4.90E-75	0.27582424	0.127	0.054	8.37E-71	MZB-1
CLECL1	6.97E-69	0.34735599	0.24	0.137	1.19E-64	MZB-1
CR2	1.66E-63	0.349130714	0.269	0.163	2.84E-59	MZB-1
RHEX	4.35E-57	0.331956494	0.196	0.11	7.44E-53	MZB-1
DNASE1L3	4.88E-55	0.268821248	0.182	0.1	8.33E-51	MZB-1
SMIM14	9.83E-55	0.400552166	0.402	0.287	1.68E-50	MZB-1
MYC	3.23E-54	0.34326193	0.271	0.171	5.52E-50	MZB-1
SMARCB1	1.01E-46	0.278820181	0.46	0.34	1.72E-42	MZB-1
SLA	3.88E-42	0.300960417	0.166	0.097	6.63E-38	MZB-1
MT-ND6	3.09E-13	0.375088188	0.468	0.411	5.28E-09	MZB-1
HSP90AB1	0	1.266489392	0.961	0.781	0	MZB-2
NME1	0	0.7244138	0.239	0.038	0	MZB-2
FABP5	1.75E-231	1.246246489	0.383	0.121	2.99E-227	MZB-2
PSME2	1.46E-216	1.028416709	0.694	0.353	2.49E-212	MZB-2
SRM	8.56E-209	0.777397176	0.29	0.077	1.46E-204	MZB-2
YBX1	8.47E-208	0.942849623	0.899	0.686	1.45E-203	MZB-2
PRDX1	9.08E-206	1.147045977	0.641	0.324	1.55E-201	MZB-2
MIF	1.37E-204	0.923504774	0.876	0.642	2.33E-200	MZB-2
ENO1	1.66E-201	1.101683444	0.598	0.291	2.83E-197	MZB-2
NCL	3.46E-200	1.008852128	0.844	0.576	5.91E-196	MZB-2
NHP2	6.30E-196	0.908455064	0.556	0.246	1.08E-191	MZB-2
PAICS	8.36E-192	0.638335522	0.255	0.064	1.43E-187	MZB-2
LILRA4	2.20E-191	0.758563223	0.213	0.047	3.75E-187	MZB-2
LDHA	2.13E-183	1.057535606	0.453	0.184	3.64E-179	MZB-2
TPI1	9.80E-182	1.008584266	0.722	0.423	1.67E-177	MZB-2
HSPD1	1.61E-176	1.009067737	0.613	0.311	2.75E-172	MZB-2
DDX21	2.07E-176	1.008697678	0.588	0.288	3.54E-172	MZB-2
RPL22L1	6.68E-167	0.957108049	0.762	0.5	1.14E-162	MZB-2
RAN	2.11E-162	0.905023297	0.755	0.49	3.60E-158	MZB-2
NOLC1	2.22E-161	0.653855679	0.302	0.095	3.79E-157	MZB-2
TMSB10	0	0.677710405	0.97	0.931	0	Naive
STAG3	0	1.483766769	0.419	0.19	0	Naive
PLPP5	0	1.073036062	0.39	0.149	0	Naive
TCL1A	0	0.876309909	0.678	0.279	0	Naive
IL4R	0	1.149653651	0.305	0.083	0	Naive
RPL18A	0	0.375754617	0.991	0.981	0	Naive
SESN1	3.70E-305	0.992860751	0.383	0.163	6.33E-301	Naive
FCER2	5.00E-255	0.756952123	0.511	0.28	8.55E-251	Naive
RPS19	8.67E-199	0.311319302	0.99	0.98	1.48E-194	Naive
SELL	9.74E-160	0.748581855	0.567	0.404	1.66E-155	Naive
HLA-DPB1	1.29E-143	0.349769799	0.945	0.889	2.21E-139	Naive
BACH2	1.35E-142	0.665757637	0.301	0.155	2.31E-138	Naive
RPS5	6.91E-135	0.271726052	0.978	0.952	1.18E-130	Naive
ADK	5.55E-127	0.646729111	0.292	0.158	9.49E-123	Naive
LINC00926	5.60E-124	0.521295759	0.632	0.486	9.57E-120	Naive

SKAP1	1.82E-115	0.466790937	0.157	0.063	3.11E-111	Naive
CD200	6.95E-113	0.506057077	0.176	0.076	1.19E-108	Naive
MTSS1	2.56E-106	0.570424178	0.232	0.119	4.37E-102	Naive
CD52	2.68E-94	0.291456243	0.948	0.909	4.59E-90	Naive
CD37	5.06E-90	0.276822557	0.955	0.92	8.64E-86	Naive
YBX3	7.67E-89	0.43030319	0.188	0.092	1.31E-84	Naive
MEF2C	8.59E-89	0.548597483	0.604	0.506	1.47E-84	Naive
IGHD	2.59E-81	0.430471961	0.296	0.18	4.42E-77	Naive
HVCN1	7.81E-74	0.471987649	0.413	0.301	1.33E-69	Naive
BTLA	6.95E-72	0.439642114	0.262	0.16	1.19E-67	Naive
SDC3	0	1.330522606	0.113	0	0	PB
FCGR3A	0	1.721120557	0.175	0.001	0	PB
SLC40A1	0	2.408500954	0.226	0.001	0	PB
AIF1	0	4.307908635	0.503	0.016	0	PB
CTSL	0	1.325309896	0.147	0	0	PB
MS4A6A	0	1.769156556	0.198	0.001	0	PB
IGSF6	0	2.754307395	0.266	0.001	0	PB
CST3	0	5.037577468	0.384	0.002	0	PB
TYROBP	0	4.099981462	0.35	0.004	0	PB
LILRB2	0	1.136008223	0.136	0.001	0	PB
VCAM1	0	2.20255873	0.192	0	0	PB
FCER1G	0	4.251064086	0.362	0.001	0	PB
TMEM176E	0	1.337372322	0.119	0	0	PB
RASSF4	0	1.146765378	0.13	0.001	0	PB
APOC1	0	3.904374217	0.333	0.001	0	PB
APOE	0	5.294829403	0.362	0.002	0	PB
SELENOP	0	3.490784991	0.305	0.001	0	PB
RBP7	0	1.187907451	0.13	0	0	PB
C1QA	0	5.331890608	0.367	0.001	0	PB
C1QC	0	5.132043997	0.362	0.001	0	PB
C1QB	0	5.539760768	0.367	0.001	0	PB
FABP3	0	1.290745701	0.13	0	0	PB
S100A9	0	1.20632887	0.113	0	0	PB
CD5L	0	3.195213445	0.288	0	0	PB
DAB2	0	1.510535816	0.158	0	0	PB
CSF1R	0	1.429405636	0.153	0	0	PB
CPVL	0	2.119510442	0.198	0	0	PB
CXCL12	0	2.085439754	0.243	0	0	PB
CD163	0	1.874573304	0.181	0	0	PB
SPIC	0	1.439310925	0.124	0	0	PB
CMKLR1	0	1.031123528	0.102	0	0	PB
SERPINA1	0	1.060571565	0.102	0	0	PB
FPR1	0	1.10218105	0.119	0	0	PB
TMSB10	0	0.767159437	0.988	0.932	0	TS
TCL1A	0	2.58932188	0.985	0.281	0	TS
AL139020.1	0	1.174534997	0.22	0.026	0	TS
IGLL5	0	1.384234145	0.283	0.064	0	TS
RPL18A	2.38E-265	0.392494848	0.996	0.982	4.06E-261	TS
CD74	4.41E-224	0.395306239	0.998	0.996	7.54E-220	TS
CD37	4.93E-223	0.549422119	0.973	0.921	8.41E-219	TS

CD52	6.73E-175	0.519632995	0.97	0.909	1.15E-170	TS
CD79B	4.00E-135	0.662950015	0.774	0.638	6.84E-131	TS
HLA-DRB1	5.31E-123	0.386543454	0.98	0.923	9.06E-119	TS
IGHD	1.54E-104	0.657779397	0.353	0.183	2.64E-100	TS
RNASE6	8.80E-103	0.836310695	0.482	0.317	1.50E-98	TS
FCER2	8.77E-93	0.5814774	0.48	0.296	1.50E-88	TS
PCDH9	1.44E-88	0.424161953	0.13	0.042	2.47E-84	TS
CD79A	1.70E-85	0.368851414	0.953	0.9	2.90E-81	TS
PLPP5	7.88E-82	0.60756171	0.315	0.169	1.35E-77	TS
NCF1	4.22E-80	0.50882688	0.707	0.598	7.21E-76	TS
HLA-DPB1	2.47E-79	0.330344917	0.95	0.892	4.21E-75	TS
YBX3	2.80E-79	0.453678016	0.216	0.096	4.78E-75	TS
HLA-DRA	2.88E-76	0.279208974	0.989	0.961	4.92E-72	TS
SKAP1	5.77E-73	0.423106633	0.166	0.068	9.86E-69	TS
CDCA7L	1.83E-66	0.526653992	0.237	0.122	3.13E-62	TS

Table S4

Donor	Tissue	Subset Names	Cell Count	Percentage of CD19
A	APP	AcB1	441	9.051724138
B	APP	AcB1	304	7.981097401
C	APP	AcB1	504	11.02844639
A	MLN	AcB1	207	3.542094456
B	MLN	AcB1	177	6.165099269
C	MLN	AcB1	231	5.941358025
A	SPL	AcB1	202	4.105691057
B	SPL	AcB1	223	7.24025974
C	SPL	AcB1	159	4.660023447
A	APP	AcB2	153	3.140394089
B	APP	AcB2	167	4.384352849
C	APP	AcB2	247	5.404814004
A	MLN	AcB2	421	7.203969884
B	MLN	AcB2	230	8.011145942
C	MLN	AcB2	160	4.115226337
A	SPL	AcB2	325	6.605691057
B	SPL	AcB2	153	4.967532468
C	SPL	AcB2	197	5.773739742
A	APP	AcB3	181	3.715106732
B	APP	AcB3	164	4.305592019
C	APP	AcB3	235	5.142231947
A	MLN	AcB3	685	11.72142368
B	MLN	AcB3	175	6.09543713
C	MLN	AcB3	355	9.130658436
A	SPL	AcB3	325	6.605691057
B	SPL	AcB3	144	4.675324675
C	SPL	AcB3	188	5.50996483
A	APP	AcB4	61	1.252052545
B	APP	AcB4	78	2.04778157
C	APP	AcB4	19	0.415754923
A	MLN	AcB4	49	0.838466804
B	MLN	AcB4	46	1.602229188
C	MLN	AcB4	21	0.540123457
A	SPL	AcB4	42	0.853658537
B	SPL	AcB4	61	1.980519481
C	SPL	AcB4	31	0.90855803
A	APP	aNAV	505	10.36535304
B	APP	aNAV	301	7.902336571
C	APP	aNAV	292	6.389496718
A	MLN	aNAV	380	6.502395619
B	MLN	aNAV	194	6.757227447
C	MLN	aNAV	330	8.487654321
A	SPL	aNAV	475	9.654471545
B	SPL	aNAV	242	7.857142857
C	SPL	aNAV	243	7.121922626
A	APP	DN-A	100	2.052545156
B	APP	DN-A	84	2.205303229

C	APP	DN-A	128	2.800875274
A	MLN	DN-A	135	2.310061602
B	MLN	DN-A	40	1.393242773
C	MLN	DN-A	72	1.851851852
A	SPL	DN-A	83	1.68699187
B	SPL	DN-A	41	1.331168831
C	SPL	DN-A	58	1.699882767
A	APP	DN-B	93	1.908866995
B	APP	DN-B	75	1.96902074
C	APP	DN-B	135	2.95404814
A	MLN	DN-B	51	0.872689938
B	MLN	DN-B	10	0.348310693
C	MLN	DN-B	20	0.514403292
A	SPL	DN-B	30	0.609756098
B	SPL	DN-B	15	0.487012987
C	SPL	DN-B	21	0.615474795
A	APP	GC	869	17.83661741
B	APP	GC	502	13.17931216
C	APP	GC	593	12.97592998
A	MLN	GC	310	5.3045859
B	MLN	GC	117	4.07523511
C	MLN	GC	174	4.475308642
A	SPL	GC	264	5.365853659
B	SPL	GC	215	6.980519481
C	SPL	GC	218	6.389214537
A	APP	IgM-only	213	4.371921182
B	APP	IgM-only	273	7.167235495
C	APP	IgM-only	322	7.04595186
A	MLN	IgM-only	511	8.744010951
B	MLN	IgM-only	330	11.49425287
C	MLN	IgM-only	385	9.902263374
A	SPL	IgM-only	375	7.62195122
B	SPL	IgM-only	367	11.91558442
C	SPL	IgM-only	256	7.502930832
A	APP	Memory	553	11.35057471
B	APP	Memory	533	13.99317406
C	APP	Memory	878	19.21225383
A	MLN	Memory	724	12.38877481
B	MLN	Memory	596	20.75931731
C	MLN	Memory	797	20.49897119
A	SPL	Memory	777	15.79268293
B	SPL	Memory	546	17.72727273
C	SPL	Memory	595	17.43845252
A	APP	MZB-1	281	5.767651888
B	APP	MZB-1	262	6.878445786
C	APP	MZB-1	429	9.387308534
A	MLN	MZB-1	1046	17.89869952
B	MLN	MZB-1	390	13.58411703
C	MLN	MZB-1	419	10.77674897
A	SPL	MZB-1	681	13.84146341

B	SPL	MZB-1	321	10.42207792
C	SPL	MZB-1	490	14.36107855
A	APP	MZB-2	218	4.47454844
B	APP	MZB-2	202	5.303229194
C	APP	MZB-2	194	4.245076586
A	MLN	MZB-2	238	4.072553046
B	MLN	MZB-2	94	3.274120515
C	MLN	MZB-2	201	5.169753086
A	SPL	MZB-2	232	4.715447154
B	SPL	MZB-2	146	4.74025974
C	SPL	MZB-2	195	5.715123095
A	APP	Naive	932	19.12972085
B	APP	Naive	510	13.38934103
C	APP	Naive	372	8.140043764
A	MLN	Naive	706	12.0807666
B	MLN	Naive	267	9.299895507
C	MLN	Naive	452	11.6255144
A	SPL	Naive	695	14.12601626
B	SPL	Naive	325	10.55194805
C	SPL	Naive	431	12.63188746
A	APP	PB	7	0.143678161
B	APP	PB	12	0.315043318
C	APP	PB	13	0.284463895
A	MLN	PB	48	0.821355236
B	MLN	PB	17	0.592128178
C	MLN	PB	19	0.488683128
A	SPL	PB	9	0.182926829
B	SPL	PB	46	1.493506494
C	SPL	PB	6	0.175849941
A	APP	TS	265	5.439244663
B	APP	TS	342	8.978734576
C	APP	TS	209	4.573304158
A	MLN	TS	333	5.698151951
B	MLN	TS	188	6.548241031
C	MLN	TS	252	6.481481481
A	SPL	TS	405	8.231707317
B	SPL	TS	235	7.62987013
C	SPL	TS	324	9.495896835

Table S5

Tissue	Subset	Total Cell Count per Subset and Tissue	Number of cells found in clones that have members in 2 or more tissues	Number of cells found in clones that have members in one tissue	Percentage of total cells found in clones that have members in 2 or more tissues	Percentage of total cells found in clones that have members in one tissue
APP	AcB1	928	42	279	4.525862069	30.06465517
APP	AcB2	433	14	277	3.233256351	63.97228637
APP	AcB3	400	24	62	6	15.5
APP	AcB4	120	3	55	2.5	45.83333333
APP	aNAV	853	25	314	2.930832356	36.8112544
APP	DN-A	231	6	64	2.597402597	27.70562771
APP	DN-B	243	4	69	1.646090535	28.39506173
APP	GC	1637	27	558	1.649358583	34.08674404
APP	IgM-only	554	20	139	3.610108303	25.09025271
APP	Memory	1435	82	372	5.714285714	25.92334495
APP	MZB-1	769	93	200	12.09362809	26.00780234
APP	MZB-2	456	11	144	2.412280702	31.57894737
APP	Naive	1597	7	589	0.438321853	36.8816531
APP	PB	19	NA	6	NA	31.57894737
APP	TS	698	3	161	0.429799427	23.06590258
MLN	AcB1	522	28	101	5.363984674	19.348659
MLN	AcB2	467	13	282	2.78372591	60.38543897
MLN	AcB3	849	38	178	4.475853946	20.96584217
MLN	AcB4	93	5	16	5.376344086	17.20430108
MLN	aNAV	819	35	181	4.273504274	22.1001221
MLN	DN-A	150	8	33	5.333333333	22
MLN	DN-B	58	2	15	3.448275862	25.86206897
MLN	GC	499	16	146	3.206412826	29.25851703
MLN	IgM-only	840	49	148	5.833333333	17.61904762
MLN	Memory	1588	114	262	7.17884131	16.49874055
MLN	MZB-1	1183	86	295	7.269653423	24.93660186
MLN	MZB-2	481	25	110	5.197505198	22.86902287
MLN	Naive	1298	16	369	1.232665639	28.42835131
MLN	PB	29	1	4	3.448275862	13.79310345
MLN	TS	756	6	175	0.793650794	23.14814815
SPL	AcB1	475	22	106	4.631578947	22.31578947
SPL	AcB2	556	20	313	3.597122302	56.29496403
SPL	AcB3	492	34	108	6.910569106	21.95121951
SPL	AcB4	107	2	22	1.869158879	20.56074766
SPL	aNAV	831	151	162	18.17087846	19.49458484
SPL	DN-A	151	10	48	6.622516556	31.78807947
SPL	DN-B	49	2	13	4.081632653	26.53061224
SPL	GC	598	28	133	4.682274247	22.24080268
SPL	IgM-only	766	59	154	7.702349869	20.10443864
SPL	Memory	1534	121	380	7.887874837	24.77183833
SPL	MZB-1	1269	112	367	8.825847124	28.92040977
SPL	MZB-2	506	44	118	8.695652174	23.3201581
SPL	Naive	1360	17	300	1.25	22.05882353
SPL	PB	9	2	4	22.22222222	44.44444444
SPL	TS	906	9	167	0.993377483	18.43267108

Table S6

	p_val	avg_log2FC	pct.1	pct.2	p_val_adj	subset
S100A10	0	2.1895349	0.839	0.265	0	Memory
S100A11	0	1.0620257	0.46	0.13	0	Memory
S100A6	0	1.0758518	0.783	0.458	0	Memory
S100A4	0	1.8044352	0.716	0.267	0	Memory
AIM2	0	1.1155493	0.362	0.063	0	Memory
GRAMD1C	0	0.9847	0.245	0.026	0	Memory
HOPX	0	1.0302005	0.264	0.045	0	Memory
PDE4D	0	0.9518577	0.258	0.045	0	Memory
CPNE5	0	1.0203142	0.414	0.111	0	Memory
GSTK1	0	1.1199579	0.615	0.254	0	Memory
CD99	0	1.1497736	0.59	0.22	0	Memory
AK8	0	0.5181849	0.101	0.005	0	Memory
AHNAK	0	1.1653356	0.444	0.126	0	Memory
ITGB1	0	1.6951947	0.508	0.114	0	Memory
CD27	0	1.0648954	0.336	0.059	0	Memory
GAPDH	0	1.236583	0.846	0.535	0	Memory
SSPN	0	0.9409184	0.24	0.03	0	Memory
BTG1	0	-1.2329988	0.761	0.9	0	Memory
TESC	0	0.6875186	0.165	0.018	0	Memory
TRAC	0	0.7620507	0.195	0.029	0	Memory
LINC01781	0	-0.9433463	0.082	0.24	0	Naive
S100A10	0	-1.1269764	0.214	0.441	0	Naive
S100A11	0	-1.0614847	0.061	0.28	0	Naive
S100A6	0	-0.9194879	0.369	0.632	0	Naive
S100A4	0	-1.3774609	0.178	0.465	0	Naive
AIM2	0	-0.8138972	0.016	0.18	0	Naive
PLEK	0	-0.728175	0.125	0.304	0	Naive
RPS23	0	0.2612065	0.999	0.997	0	Naive
HLA-A	0	-0.5088766	0.638	0.811	0	Naive
HLA-B	0	-0.538726	0.749	0.886	0	Naive
PLPP5	0	1.0407018	0.486	0.251	0	Naive
RPL8	0	0.3669396	0.995	0.993	0	Naive
LSP1	0	-0.6154457	0.368	0.562	0	Naive
CD82	0	-0.6664491	0.113	0.293	0	Naive
AHNAK	0	-0.8520152	0.072	0.256	0	Naive
VIM	0	-0.7573234	0.355	0.579	0	Naive
ITGB1	0	-1.1461779	0.058	0.261	0	Naive
CD27	0	-0.800157	0.013	0.17	0	Naive
GAPDH	0	-0.9029774	0.457	0.694	0	Naive
CLECL1	0	-0.7764314	0.018	0.174	0	Naive
SOX4	0	1.3228043	0.313	0.044	0	TS
TCL1A	0	1.1877695	0.906	0.482	0	TS
AL139020.	0	1.3619385	0.4	0.087	0	TS
IGLL5	0	1.5587114	0.517	0.166	0	TS
CD9	1.91E-304	0.7396354	0.12	0.013	3.09E-300	TS
PPP1R14A	1.23E-274	0.8860416	0.367	0.123	1.98E-270	TS
LIMS2	9.84E-242	0.8419964	0.199	0.047	1.59E-237	TS
PLD4	9.75E-194	0.9012326	0.252	0.084	1.58E-189	TS

VPREB3	3.91E-192	0.7854093	0.691	0.434	6.31E-188	TS
TMSB10	6.47E-189	0.3605072	1	0.996	1.04E-184	TS
MZB1	2.55E-182	0.7039988	0.331	0.135	4.12E-178	TS
PCDH9	3.52E-167	0.746652	0.235	0.079	5.68E-163	TS
CD79B	9.40E-155	0.5915956	0.855	0.701	1.52E-150	TS
RPS24	3.54E-152	0.3172725	0.999	0.994	5.72E-148	TS
FAM129C	1.19E-150	0.733743	0.528	0.304	1.93E-146	TS
HRK	7.43E-149	0.5660752	0.101	0.02	1.20E-144	TS
S100A10	1.15E-140	-1.3176631	0.108	0.333	1.86E-136	TS
IGHM	1.83E-116	0.4914843	0.79	0.577	2.96E-112	TS
CD38	2.41E-116	0.5353144	0.138	0.041	3.89E-112	TS
MACROD2	7.10E-115	0.6248978	0.226	0.091	1.15E-110	TS
LINC01781	0	1.8836694	0.573	0.117	0	IgM-only
TMSB10	0	-0.5724788	0.993	0.997	0	IgM-only
TCL1A	0	-2.7875897	0.047	0.554	0	IgM-only
IGHA1	2.13E-298	0.5712987	0.383	0.124	3.44E-294	IgM-only
IGHM	1.12E-290	-1.4099302	0.256	0.622	1.81E-286	IgM-only
AIM2	5.53E-265	0.8875689	0.27	0.074	8.92E-261	IgM-only
COCH	4.86E-226	0.9129057	0.185	0.042	7.85E-222	IgM-only
IGHA2	7.68E-206	0.7740723	0.15	0.031	1.24E-201	IgM-only
PLPP5	1.41E-202	-1.4186196	0.116	0.404	2.28E-198	IgM-only
BTG1	3.57E-200	-0.8201484	0.805	0.894	5.77E-196	IgM-only
HLA-DRB1	2.94E-199	-0.5932083	0.942	0.967	4.74E-195	IgM-only
IGHD	1.35E-185	-1.529556	0.043	0.305	2.18E-181	IgM-only
S100A4	4.81E-185	0.819779	0.543	0.286	7.77E-181	IgM-only
JCHAIN	4.41E-178	0.5091753	0.458	0.23	7.12E-174	IgM-only
SSPN	4.83E-163	0.6539664	0.155	0.039	7.80E-159	IgM-only
RPLP1	1.87E-160	0.2752162	1	0.999	3.03E-156	IgM-only
RPS14	7.76E-150	0.3542498	0.999	0.995	1.25E-145	IgM-only
LGALS1	1.76E-149	0.6013904	0.293	0.115	2.84E-145	IgM-only
CD27	9.26E-136	0.5635799	0.209	0.072	1.49E-131	IgM-only
KLF10	6.66E-129	0.558913	0.131	0.035	1.08E-124	IgM-only
FGR	0	1.5194288	0.366	0.075	0	aNAV
FCRL5	0	1.352711	0.388	0.108	0	aNAV
FCRL3	0	1.4434431	0.42	0.121	0	aNAV
LTB	0	-2.2621972	0.301	0.895	0	aNAV
MPP6	0	1.6849491	0.323	0.033	0	aNAV
HSPB1	0	2.0311085	0.552	0.199	0	aNAV
EMP3	0	1.7405001	0.868	0.574	0	aNAV
RGS2	1.84E-304	1.2187143	0.249	0.048	2.97E-300	aNAV
PPP1R14A	8.89E-303	1.5468159	0.414	0.127	1.44E-298	aNAV
ZEB2	1.21E-283	1.3259506	0.379	0.112	1.96E-279	aNAV
MS4A1	4.46E-280	0.8242638	0.969	0.894	7.20E-276	aNAV
RPS23	3.35E-250	-0.5398541	0.998	0.998	5.42E-246	aNAV
RHOB	1.98E-244	1.4495351	0.376	0.124	3.21E-240	aNAV
VPREB3	2.28E-243	-1.8240748	0.068	0.477	3.69E-239	aNAV
SLC11A1	3.92E-240	0.6228223	0.102	0.009	6.32E-236	aNAV
RPL8	1.46E-228	-0.6071851	0.989	0.994	2.35E-224	aNAV
S100A11	3.56E-225	1.2885082	0.4	0.144	5.75E-221	aNAV
RPS3A	3.46E-214	-0.4972792	0.992	0.996	5.59E-210	aNAV

RPL32	2.77E-213	-0.4103799	0.998	0.999	4.47E-209	aNAV
RPLP0	2.88E-211	-0.6211981	0.916	0.976	4.65E-207	aNAV
TMSB10	0	-0.5862722	0.995	0.997	0	MZB-1
LINC01857	0	1.2016929	0.444	0.118	0	MZB-1
TCL1A	0	-1.9782059	0.171	0.545	0	MZB-1
CD27	1.64E-206	0.667327	0.238	0.07	2.66E-202	MZB-1
TNFRSF13I	9.66E-165	0.6539695	0.263	0.095	1.56E-160	MZB-1
GPR183	2.51E-155	0.6749383	0.294	0.116	4.06E-151	MZB-1
COTL1	8.34E-139	0.6165658	0.523	0.288	1.35E-134	MZB-1
AIM2	1.19E-116	0.4985437	0.21	0.079	1.92E-112	MZB-1
IGHG2	3.30E-114	0.3975311	0.215	0.083	5.32E-110	MZB-1
MARCKS	2.45E-112	0.5290745	0.212	0.082	3.95E-108	MZB-1
MEF2C	2.85E-99	-0.763962	0.507	0.635	4.61E-95	MZB-1
SELENOM	2.90E-96	0.4219037	0.188	0.073	4.69E-92	MZB-1
PLPP5	7.31E-96	-0.9780743	0.226	0.395	1.18E-91	MZB-1
ARHGAP24	7.57E-92	0.4832293	0.38	0.205	1.22E-87	MZB-1
RPL3	1.06E-90	0.3546503	0.948	0.891	1.71E-86	MZB-1
IGLL5	5.03E-90	-1.2689919	0.053	0.206	8.13E-86	MZB-1
HLA-DRB1	1.33E-82	-0.3637149	0.963	0.965	2.16E-78	MZB-1
SMARCB1	1.58E-81	0.4696513	0.358	0.199	2.55E-77	MZB-1
IL4R	2.29E-79	-1.0020703	0.06	0.201	3.70E-75	MZB-1
CD1C	4.07E-79	0.4172579	0.132	0.047	6.57E-75	MZB-1
CD1C	0	1.1629904	0.306	0.042	0	MZB-2
CRIP1	0	1.277135	0.539	0.142	0	MZB-2
PPP1R14A	0	1.2027931	0.469	0.127	0	MZB-2
KLK1	0	0.8290057	0.175	0.015	0	MZB-2
COTL1	1.08E-243	1.1574487	0.648	0.289	1.74E-239	MZB-2
TNFRSF13I	2.55E-218	0.8435977	0.345	0.096	4.12E-214	MZB-2
CIB1	2.69E-217	1.0205066	0.694	0.322	4.34E-213	MZB-2
AHNAK	1.93E-201	0.9045651	0.422	0.141	3.11E-197	MZB-2
IGHG2	1.68E-184	0.817308	0.295	0.083	2.71E-180	MZB-2
ANXA4	1.32E-182	0.7382721	0.319	0.094	2.13E-178	MZB-2
CLECL1	5.64E-179	0.6846872	0.284	0.078	9.10E-175	MZB-2
TCL1A	3.59E-173	-1.7687019	0.202	0.531	5.79E-169	MZB-2
HCK	3.92E-172	0.5997252	0.242	0.06	6.33E-168	MZB-2
CD82	7.46E-166	0.8108356	0.461	0.18	1.20E-161	MZB-2
ACP5	5.02E-162	0.718569	0.3	0.092	8.11E-158	MZB-2
SYK	5.89E-155	0.7950806	0.446	0.176	9.52E-151	MZB-2
MRPL40	1.33E-152	0.6539401	0.367	0.128	2.15E-148	MZB-2
CXCR3	1.00E-149	0.4712986	0.127	0.022	1.62E-145	MZB-2
FGR	3.33E-149	0.6643932	0.275	0.082	5.38E-145	MZB-2
VOPP1	6.42E-149	0.6594901	0.29	0.091	1.04E-144	MZB-2
TNFRSF1B	0	1.566679	0.438	0.048	0	DN
MPP6	0	1.5504217	0.418	0.044	0	DN
SYT1	0	1.0236124	0.214	0.008	0	DN
RBPMS2	0	0.6107815	0.142	0.005	0	DN
ITGAX	0	1.4044317	0.338	0.017	0	DN
CST7	0	1.3786593	0.216	0.001	0	DN
HCST	0	2.2064894	0.61	0.087	0	DN
NKG7	0	2.8101465	0.588	0.006	0	DN

LILRB2	0	0.9723621	0.186	0.005	0	DN
DUSP2	3.11E-297	2.2542476	0.358	0.036	5.02E-293	DN
FGR	4.26E-280	1.6764018	0.544	0.085	6.89E-276	DN
RGS2	5.80E-268	1.7030098	0.418	0.053	9.37E-264	DN
ENC1	2.63E-262	0.7979837	0.188	0.011	4.25E-258	DN
SLC11A1	1.18E-261	0.8269779	0.198	0.012	1.90E-257	DN
AC090152.	2.08E-248	0.6689551	0.158	0.008	3.35E-244	DN
ADGRE5	2.49E-241	1.328812	0.398	0.052	4.03E-237	DN
S100A11	6.64E-240	1.8179671	0.674	0.151	1.07E-235	DN
TIMP1	1.76E-216	1.1521272	0.274	0.028	2.85E-212	DN
SRGN	6.25E-212	1.6014159	0.43	0.07	1.01E-207	DN
ARL4D	7.82E-208	0.8361178	0.196	0.015	1.26E-203	DN
RPS12	4.92E-77	-1.9060749	0.66	1	7.94E-73	AcB3
RPL41	6.77E-77	-1.3956988	0.915	1	1.09E-72	AcB3
RPS4X	6.04E-67	-1.7088248	0.68	0.995	9.75E-63	AcB3
RPL30	9.42E-62	-1.3506456	0.771	1	1.52E-57	AcB3
RPLP1	3.03E-61	-1.3204688	0.804	1	4.90E-57	AcB3
RPL32	2.60E-60	-1.2915134	0.817	0.999	4.20E-56	AcB3
RPL10	5.16E-59	-1.3328007	0.856	0.999	8.33E-55	AcB3
RPS5	9.77E-59	-1.7299048	0.523	0.978	1.58E-54	AcB3
RPL19	8.90E-57	-1.271862	0.791	0.999	1.44E-52	AcB3
RPS28	4.69E-56	-1.3047115	0.752	0.999	7.58E-52	AcB3
RPS13	5.12E-56	-1.4183974	0.699	0.997	8.26E-52	AcB3
RPS8	5.97E-56	-1.5748188	0.732	0.998	9.64E-52	AcB3
RPS7	8.54E-56	-1.3979731	0.667	0.994	1.38E-51	AcB3
RPL13	1.13E-55	-1.2785389	0.83	0.999	1.83E-51	AcB3
RPS14	1.28E-55	-1.4473524	0.66	0.997	2.07E-51	AcB3
RPS18	5.28E-55	-1.4515779	0.66	0.995	8.53E-51	AcB3
RPL8	9.51E-54	-1.4205836	0.745	0.995	1.54E-49	AcB3
RPL18	1.27E-53	-1.333477	0.745	0.995	2.06E-49	AcB3
RPL11	1.76E-53	-1.246592	0.797	0.999	2.84E-49	AcB3
RPL35A	6.28E-53	-1.3312306	0.693	0.995	1.01E-48	AcB3
CNKSR1	0	0.5404841	0.159	0	0	PB
SEMA4A	0	1.0750325	0.317	0.007	0	PB
SLAMF7	0	1.6045248	0.455	0.004	0	PB
MIXL1	0	0.7285988	0.207	0.001	0	PB
CHPF	0	0.9801154	0.276	0.001	0	PB
NT5DC2	0	0.4140284	0.11	0	0	PB
FNDC3B	0	1.0200544	0.297	0.002	0	PB
LINC02362	0	0.9781741	0.207	0.001	0	PB
ELL2	0	1.7098799	0.517	0.018	0	PB
TXNDC5	0	1.8477312	0.476	0.004	0	PB
PRDM1	0	1.8997937	0.455	0.006	0	PB
MAN1A1	0	1.4314806	0.366	0.006	0	PB
CAV1	0	1.0978609	0.297	0.003	0	PB
PDIA4	0	3.009791	0.828	0.058	0	PB
AQP3	0	2.1959605	0.538	0.002	0	PB
SLC44A1	0	1.3077415	0.428	0.014	0	PB
NUCB2	0	1.8730977	0.545	0.018	0	PB
HRASLS2	0	1.4286778	0.331	0.001	0	PB

FKBP11	0	2.7384711	0.779	0.037	0	PB
IGF1	0	0.6492425	0.145	0	0	PB