



Single-Cell Transcriptomic Analyses Define Distinct Peripheral B Cell Subsets and Discrete Development Pathways

OPEN ACCESS

Alexander Stewart^{1†}, Joseph Chi-Fung Ng^{2†}, Gillian Wallis¹, Vasiliki Tsioligka¹, Franca Fraternali^{2*†} and Deborah K. Dunn-Walters^{1*†}

¹ School of Biosciences and Medicine, University of Surrey, Guildford, United Kingdom, ² Randall Centre for Cell and Molecular Biophysics, School of Basic and Medical Biosciences, King's College London, London, United Kingdom

Edited by:

Ignacio Sanz,
Emory University, United States

Reviewed by:

Dagmar Scheel-Toellner,
University of Birmingham,
United Kingdom
Claude-Agnes Reynaud,
Université Paris Descartes, France

*Correspondence:

Deborah Dunn-Walters
d.dunn-walters@surrey.ac.uk
Franca Fraternali
franca.fraternali@kcl.ac.uk

[†]These authors have contributed
equally to this work and share
first authorship

[†]These authors have contributed
equally to this work and share
senior authorship

Specialty section:

This article was submitted to
B Cell Biology,
a section of the journal
Frontiers in Immunology

Received: 03 September 2020

Accepted: 24 February 2021

Published: 18 March 2021

Citation:

Stewart A, Ng JC-F,
Wallis G, Tsioligka V, Fraternali F
and Dunn-Walters DK (2021)
Single-Cell Transcriptomic
Analyses Define Distinct Peripheral
B Cell Subsets and Discrete
Development Pathways.
Front. Immunol. 12:602539.
doi: 10.3389/fimmu.2021.602539

Separation of B cells into different subsets has been useful to understand their different functions in various immune scenarios. In some instances, the subsets defined by phenotypic FACS separation are relatively homogeneous and so establishing the functions associated with them is straightforward. Other subsets, such as the “Double negative” (DN, CD19+CD27-IgD-) population, are more complex with reports of differing functionality which could indicate a heterogeneous population. Recent advances in single-cell techniques enable an alternative route to characterize cells based on their transcriptome. To maximize immunological insight, we need to match prior data from phenotype-based studies with the finer granularity of the single-cell transcriptomic signatures. We also need to be able to define meaningful B cell subsets from single cell analyses performed on PBMCs, where the relative paucity of a B cell signature means that defining B cell subsets within the whole is challenging. Here we provide a reference single-cell dataset based on phenotypically sorted B cells and an unbiased procedure to better classify functional B cell subsets in the peripheral blood, particularly useful in establishing a baseline cellular landscape and in extracting significant changes with respect to this baseline from single-cell datasets. We find 10 different clusters of B cells and applied a novel, geometry-inspired, method to RNA velocity estimates in order to evaluate the dynamic transitions between B cell clusters. This indicated the presence of two main developmental branches of memory B cells. A T-independent branch that involves IgM memory cells and two DN subpopulations, culminating in a population thought to be associated with Age related B cells and the extrafollicular response. The other, T-dependent, branch involves a third DN cluster which appears to be a precursor of classical memory cells. In addition, we identify a novel DN4 population, which is IgE rich and closely linked to the classical/precursor memory branch suggesting an IgE specific T-dependent cell population.

Keywords: B cells, single-cellRNAseq, cell atlas, B cell development, B cell subsets, memory B cells

INTRODUCTION

B cells can differentiate into plasma cells and secrete large amounts of antibody. There are, however, many more B cell functions which contribute to an effective immune response. B cells are highly effective antigen presenting cells, capable of presenting both protein and lipid antigens to T and NKT cells. B cell – T cell interactions trigger a variety of activation signals in both directions resulting in enduring affinity matured B cell memory, that may also be class switched, and activated T helper cells. B cells can also be activated *via* TLRs, producing pro-inflammatory cytokines such as IL6, TNF α and IFN γ in the process and resulting in differentiation into short-lived plasmablasts. The former, T-dependent, response will involve formation of germinal centers over time and, since it is dependent on T cells for maturation which have also been through tolerance checkpoints, it would normally have low risk of producing autoantibodies. The latter, extrafollicular, B cell response has the advantage of being more rapid, but also runs some risk of producing lower specificity antibodies. B cells can also be regulatory, producing IL10 and ensuring that autoreactive responses are not perpetuated.

In studying different functions of human B cells in health and disease most studies rely upon phenotypic differentiation in FACS analyses from peripheral blood using IgD and CD27, or CD24 and CD38, in conjunction with the pan B cell marker CD19. For example, the CD19+CD27+IgD+ IgM memory population (1, 2) is reduced in the elderly as a percentage of total B cells (3, 4) This has important consequences for older people, since the IgM memory population is thought to provide protection against the bacterial polysaccharide T-independent antigens. Higher dimensional phenotyping shows that the IgM memory population in the blood is heterogeneous and further age-related differences are also seen (5), although the likely functional significance of this age related heterogeneity has yet to be determined.

The CD19+IgD-CD27- ‘double negative’ (DN) cells are of particular interest. Many different roles have been ascribed to this population; ‘memory precursors’, ‘exhausted memory cells’, ‘tissue based memory’, ‘extrafollicular ASC precursors’ and ‘atypical memory’ (6–15) or the most recent nomenclature DN1 (memory precursor) and DN2 (extrafollicular ASC precursor B cells) (16). DN cells are increased in older people, and in chronic infections such as HIV (6, 7, 10, 11). DN cells are also expanded in autoimmune disease such as Systemic Lupus Erythematosus (SLE) (14, 17) where they are responsive to IFN γ and thought to be precursors for pathogenic antibody secreting cells (8, 9, 18). Repertoire studies to try and clarify the relationship of DN cells to classical CD19+CD27+IgD-memory cells have been carried out and find both evidence for a close relationship with classical cells, with clones shared in both populations, as well as a difference in overall average repertoire character, with less hypermutation and larger complementarity determining regions (19). It is therefore likely that the DN compartment is functionally heterogeneous and only with high resolution techniques, such as single-cell transcriptomics, will it be possible to tease apart the sub-populations.

Single-cell transcriptomics is rapidly becoming a key methodology in biology thanks to its high resolution in terms of individual cells and high dimensional data. It offers the ability to discriminate between subsets of heterogeneous populations to understand individual contributions which may have previously been confounded by the Simpson’s paradox of studying averaged data. Unsupervised clustering algorithms offer us the chance to define subsets transcriptionally and interrogate the results to find tractable markers for use in phenotypic distinction of the same. Information about the possible functions of cell clusters can be inferred from the transcriptome relative to other clusters. scRNAseq is particularly useful in B cell immunology, where it has made pairing of heavy and light chain sequences possible (20).

Curated reference databases, such as the Human Cell Atlas (21), and baseline transcriptomic profiles of particular immune sites (22) are important for comparison to experimental datasets as well as providing new insights into already well studied biological systems. A frequently occurring problem in single-cell datasets, however, is the relative lack of B cells, given that most human experiments are run on readily available PBMCs; where B cells only make up between 5–10% of the total. Hence B cells are often under-described in single-cell experiments, and often only divided into naive, memory and plasma cells which may miss important distinctions in disease. This is a general problem but has been particularly evident in recent COVID-19 papers (16, 23–29). We have therefore produced a dataset looking exclusively at five commonly defined B cell populations by phenotype to improve the resolution and better understand the heterogeneity present in B cells. This phenotyping followed by transcriptomic characterization was performed to give the clearest possible identity of known functionally relevant populations at the transcriptomic level. By producing this dataset, we reconcile known phenotypic populations with those found in the transcriptome and can identify previously unknown populations in high dimensional space, thus providing a basic map of peripheral blood B cells as a reference dataset for the B cell community. We describe 10 major populations of B cells, including a novel population of IgE-expressing cells in the double negative compartment.

MATERIALS AND METHODS

Samples, Library Preparation and Sequencing

Peripheral blood mononuclear cells were isolated from a male healthy volunteer aged 25 (HB6) using BD-bioscience lithium heparin vacutainersTM and Ficoll-Paque Plus (GE-healthcare). REC authority 17/LO/1877.

Cells were stained in BD Horizon Brilliant stain buffer with (Biologend: CD19 HIB19 BV421, CD27 M-T271 FITC, IgD AI6-2 BV785, CD10 HI10a BV605, Live/Dead Zombie NIR), doublet and dead cells removed and lymphocytes gated on FSC/SSC, and sorted into Transitional (CD19+IgD+CD27-CD10+), Naive (CD19+IgD+CD27-CD10-), IgM Memory (CD19+IgD+CD27+),

Classical Memory (CD19+IgD-CD27+) and Double Negative (CD19+IgD-CD27-) populations on a FACSAria (BD Biosciences) at 4°C (**Supplementary Figure S1**). Cells were washed twice (5 min at 400 g, supernatant removed and replaced) in PBS supplemented with 0.04% non-acetylated BSA with a final spin through a 40 µM cell strainer. Populations were counted and run on a Chromium 10X controller using 5' chemistry (10X Genomics) in individual lanes with an expected recovery rate of 4,000 cells per lane, according to the manufacturer's instructions. Libraries were generated according to 10X Genomics instructions and run on a High Output HiSeq2500 at 1 library per lane in 30-10-100 format. Additional HB34 (male, 36 y.o.) and HB78 (male, 25 y.o.) B cell samples were isolated using StemCell (Kit), FACS sorted in the same manner and each population tagged with Biolegend TotalSeq-C Hashtag antibodies, washed on the Curiox laminar flow system (9 washes at 5 ul/sec) before being recombined in equal ratios to run on the 10X at 4,000 cells per lane; all other methods were identical.

Data Preprocessing, Clustering and Differential Expression

Data was processed through CellRanger (10X Genomics, v3.1.0) and aligned to the GRCh38 genome. The raw transcript count matrix was loaded into R (v3.6.3) using the Seurat (v3.0) package (30, 31). Cells were selected for further analyses according to the following criteria: (i) express zero *CD3E*, *GZMB*, *CD14*, *FCER1A*, *GCCR3A*, *LYZ*, *PPBP* and *CD8A* transcripts, to exclude any non-B cells and; (ii) express at least 200 distinct genes. Additionally, cells with total transcript count in the top 1% percentile were removed, as these cells were manually inspected to express transcripts of multiple V gene families per cell, indicating possible cell clumps tagged with the same barcode. Gene counts were log-normalized and the top 2,000 variably expressed genes were extracted using a variance-stabilizing transformation (vst) as implemented in the FindVariableFeatures function in Seurat. The following genes were removed from the list of variably expressed genes in order to prevent downstream dimensionality reduction and clustering to reflect individual/clonotype specific gene usage: all Ig V, D, J genes (extracted using the regular expression [regex] “*IG[HKL][VD]*”), Ig constant genes (*IGHM*, *IGHD*, *IGHE*, *IGHA[1-2]*, *IGHG[1-4]*, *IGKC*, *IGLC[1-7]*, and *AC233755.1* [which encodes *IGHV4-38-2*]), *IGLL* genes, T-cell receptor genes (regex “*TR[ABGD][CV]*”). This pruned the variably expressed gene list to $n = 1,840$ genes. Principal Component Analysis (PCA) was then performed on this pruned gene list. Surveying the first 50 principal components, the proportion of variance explained plateaued at ~ 1.5% from the 15th PC onwards. Uniform Manifold Approximation and Projection (UMAP) was performed based on the first 14 PCs, using the implementation in the python umap-learn package with correlation as the distance measure in the PC space. UMAP projections were produced on both two-dimensional and three-dimensional spaces. To define cell clusters, a shared nearest neighbor (SNN) graph was constructed using Seurat::FindNeighbors based on the first 14 PCs, and cell clusters were defined on the SNN graph with Seurat::FindClusters (resolution parameter = 1). Clusters were named according to manual

inspection for their composition in terms of the original FACS-defined populations. Differential expression was examined using the Wilcoxon rank-sum test provided in Seurat::FindMarkers. Analysis for the additional samples followed the same procedure, retaining principal components that explained at least 1.5% of the variance. We followed the data integration protocol in Seurat to map all samples to the same UMAP projection. The integrated dataset contains a total of 12,973 cells.

Trajectory and RNA Velocity Analysis

To study the trajectory across the Seurat-defined cell subsets, a spanning tree across the data points was inferred using the monocle3 package (v0.2.0), based on the 3D UMAP embedding produced as detailed above. To estimate RNA velocity, spliced and unspliced transcripts were enumerated using the velocity package (v0.17) (32). RNA velocity stream was mapped onto the UMAP space according to the pipeline provided in the SeuratWrappers package (v0.1.0) which uses functionalities provided in the velocity.R package (v0.6) (32).

Based on the inferred velocity stream we also scored the transition between cell clusters using a geometrical approach. We evaluated, for each arrow depicting the velocity stream, its alignment with the projection towards cluster centroids, using the following procedure (illustrated in **Supplementary Figure S2**): first, for each arrow A_i with starting point X_i and ending point Y_i , we determined the ‘starting cluster’ (i.e. the original cell identity from which transition is considered) by a majority vote of the identities of cells within the grid around X_i as given when the velocity stream was overlaid. We next considered, for every ‘destination’ cell cluster, the distance between X_i and its cluster centroid Z_j . This permitted calculation of the angle q between A_i and X_iZ_j , using the cosine formula:

$$\cos\theta_{i,j} = \frac{(X_iY_i)^2 + (Y_iZ_j)^2 - (X_iZ_j)^2}{2(X_iY_i)(X_iZ_j)} \quad (1)$$

for each arrow I and each ‘destination’ cluster j . q should approach zero (and hence $\cos\theta = 1$) if the arrow is perfectly aligned to the direction pointing from the starting cluster to the destination. This forms the basis to define a ‘transition score’, taking into effect of the strength of the velocity (represented by the size of the arrow projected – i.e. X_iY_i ; it should positively contribute to the score) and the distance X_iZ_j (which should negatively impact the score). This score, here denoted S , is given by:

$$S_{i,j} = \frac{(1 - \cos\theta_{i,j})X_iY_i}{X_iZ_j} \quad (2)$$

For every I and j the $(1 - \cos\theta)$ functional form ensures $S_{i,j}$ increases as the arrow becomes more geometrically align to point to cluster j . This was performed for every arrow I and every cluster j . To summarize scores from individual i 's into a composite transition score (CTS) between every starting cluster j_1 and destination cluster j_2 , all i 's are classified by their starting cluster and $S_{i,j}$ are summed together for each j . Finally the CTS, i.e. $\sum_{i \in j_1, j_2} S_{i,j}$, are adjusted by a weighting factor based on the

distance between cluster centroids d_1 (for cluster j_1) and d_2 (for j_2). We first computed the pairwise distance between d_1 and d_2 . Then, for each d_2 (which corresponds to the destination cluster), we scaled all distances into a weighting factor w of range [0,1]:

$$w = 1 - \frac{z - \min(z)}{\max(z) - \min(z)} \quad (3)$$

for all distances z pointing towards d_2 . Again, this functional form implies heavier weight is given for clusters which are close to one another in the dimensionality-reduced space. The final CTS was scaled by this factor:

$$CTS_{j_1, j_2} = w \times \sum_{i \in j_1, j_2} S_{i, j} \quad (4)$$

For our datasets RNA velocity streams were considered on two perspectives to view the 3D UMAP space: (1) UMAP1 vs UMAP3 (2) UMAP2 vs UMAP3, as visually these two perspectives separate the clusters the best. All arrows visualized in these two perspectives contributed towards the final CTS. These were plotted in **Figure 2F** in \log_2 scale to dampen extreme signals.

Protein-Protein Interaction Network Inference

The R package ACTIONet (v1.0) (33) was used to infer protein-protein interaction networks (PPINs) for (i) each FACS-defined population; (ii) each cell cluster defined by Seurat, and; (iii) cells positive for T-bet (*TBX21*) transcripts. ACTIONet takes as input a reference protein interactome and uses the transcriptomic data to filter and score protein interactions relevant to a given cell subset. Here, the Unified PPIN (UniPPIN), which we previously compiled and recently updated (April 2020; Marzuoli, Ng and Fraternali, manuscript in preparation), was used as the reference interactome. UniPPIN contains in total of 17,997 proteins and 647,508 interactions, after mapping UniProt identifiers to HGNC gene symbols using the biomaRt (v2.42.1) R package. To reduce dimensionality, genes expressed in less than 10 cells were omitted, leaving a total of $n = 12,867$ genes for PPIN inference. Raw gene counts were used to calculate a ‘gene specificity’ score using the `compute.annotations.feature.specifcity` function in ACTIONet, for each cell subset as defined above. This step enhances signals from genes exhibiting expression patterns specific to distinct cell subpopulations and penalizes commonly-expressed genes. These gene specificities were subsequently used to infer a PPIN for each cell subset using ACTIONet:run.SCINET.annotations. Each node (i.e. protein) in the network is scored by a ‘specificity’ metric, which quantifies the centrality of the protein in the network scaled by its gene expression specificity. Since ribosomal and mitochondrial proteins frequently formed cliques (highly interconnected components) in biological networks, in order to mediate the impact of the inclusion of these proteins in analyzing the topological and functional properties of the inferred PPINs, all ribosomal and mitochondrial proteins were removed from all the inferred networks prior to any statistical analyses and functional annotations.

Functional Annotation

Gene sets from Gene Ontology (GO) biological processes (BP) were downloaded from MSigDB (v7.1). Cluster-specific markers were

annotated for GO BPs, by testing for gene set overrepresentation using Fisher exact tests. Gene Set Enrichment Analysis (GSEA) was also carried out (using the `fgsea` package [v1.4.1]) on the ACTIONet-inferred PPINs, with the proteins ranked using the ACTIONet per-node specificity metric.

Data Visualization

Reads coverage was visualized using the Integrated Genomics Viewer (IGV, v2.8.2). All statistics and data visualization were performed under the R statistical computing environment (v3.6.3). Heatmaps of gene expression were produced using `pheatmap` (v1.0.12). Visualizations of PPINs were produced using `visNetwork` (v2.0.9). Volcano plots were produced using `EnhancedVolcano` (v1.4.0). Visualization of three-dimensional data embeddings were performed using the `rgl` package (v0.100.54). All other data visualizations were produced using functionalities provided in Seurat and the `ggplot2` package (v3.3.0).

RESULTS

B Cell Subsets

B cell sorted single-cell transcriptome libraries from five FACS sorted populations based on IgD/CD27/CD10 (Transitional ‘Trans’, Naive, IgM Memory ‘M-mem’, Classical Memory ‘C-mem’ and Double Negative ‘DN’) clustered into 10 unique populations using UMAP (**Figure 1A**). Of the 10 clusters formed in the UMAP analysis in all but 1 case (DN1) over 75% of the cells matched with their original FACS sorted population (**Figure 1B**); the disparity between phenotype and transcriptome was possibly due to impure sorting. Differential gene expression analysis highlights these discrete populations (**Figure 1C**). Even though these cells were CD19-sorted only 34.4% percent of cells had CD19+ in the transcriptome, CD20 proving to be a much more reliable B cell marker in the transcriptome (**Figure 1D**) (30). Similarly, CD27 mRNA was not as abundant as CD27 surface protein, particularly in the C-mem cells. Importantly, we have explicitly filtered any immunoglobulin-related genes prior to clustering to avoid biasing the clustering algorithms by light chain or class of antibody. Even so, we find that *IGHE* and *IGHA2* cells can be differentiated by their transcriptome: DN1 being enriched for *IGHA2* and DN4 for *IGHE*. After finding the unusual DN4 IgE population, further investigation leads us to postulate that it is due to a cat hair allergy, exposure 3 days prior to sampling. M-mem1 cells were distinguished from M-mem2 by higher IgD expression (**Figure 1D**). Transitional and Naive cells are best distinguished from other B cells by expression of *TCL1A*, higher in transitional, present in Naive and absent in all other cells.

Relationships Between B Cell Subsets

We performed dimensionality reduction and pseudotime analyses to decipher the relationships between clusters and RNA velocity measurements were added to superimpose directionality information onto the trajectory (**Figure 2**). Four

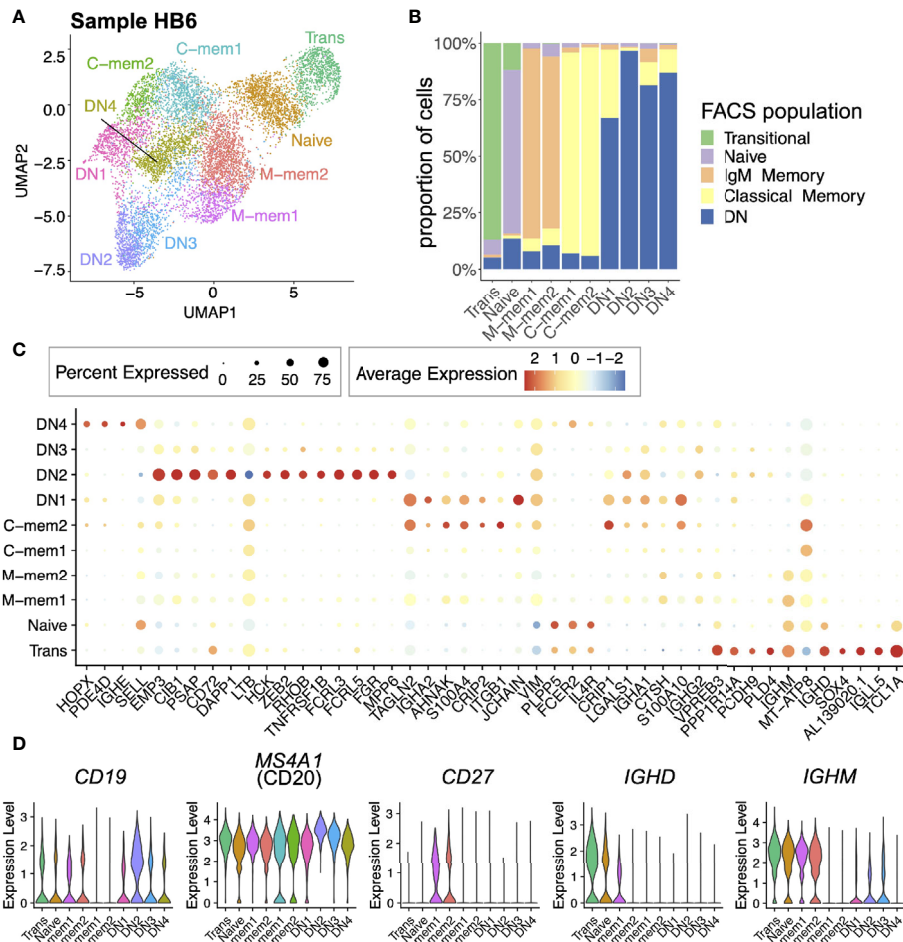


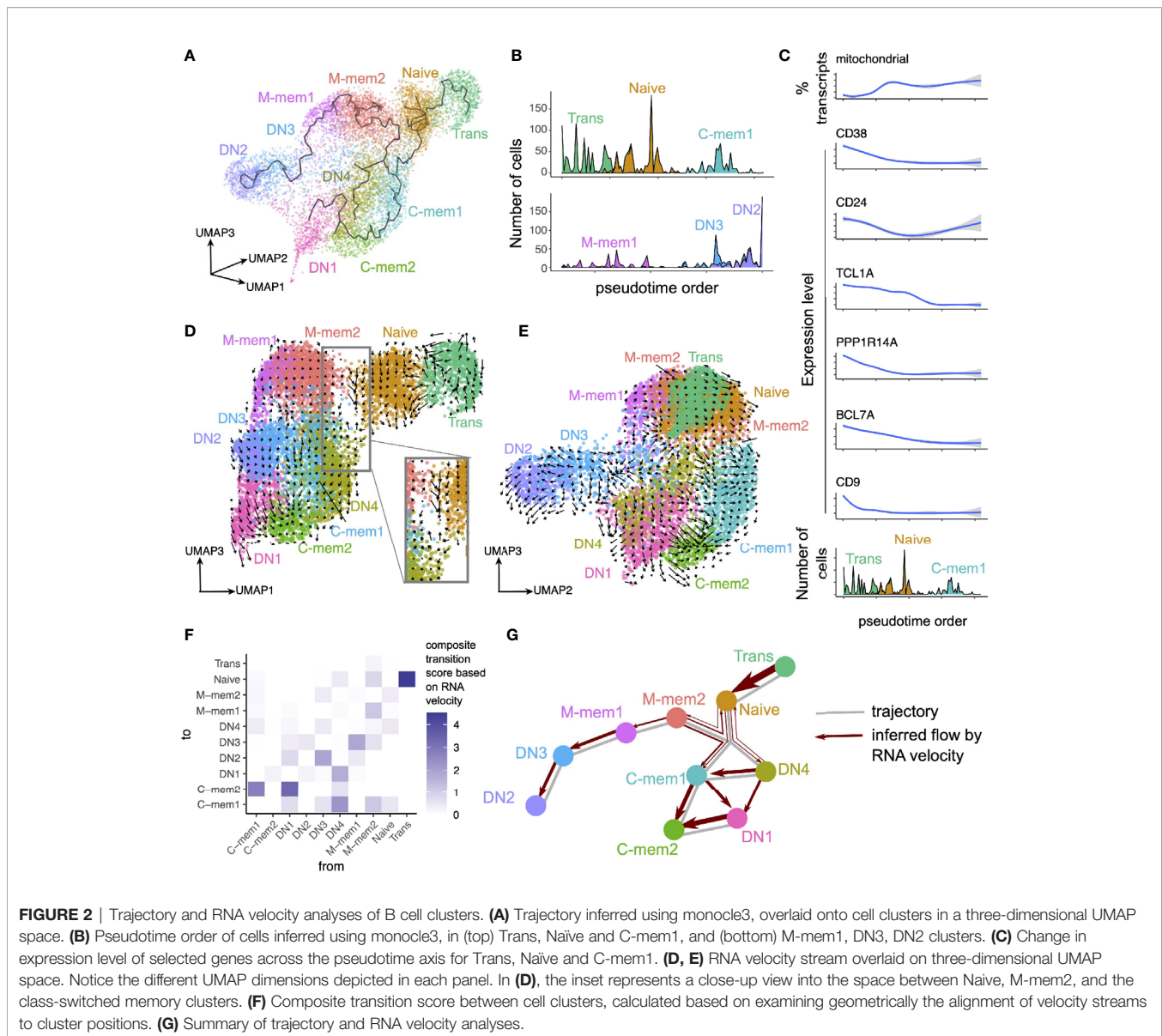
FIGURE 1 | Single-cell atlas of peripheral B cell subsets for one individual (HB6). **(A)** Two dimensional UMAP projection of scRNA-seq data of peripheral B cell subsets from sample HB6. **(B)** Breakdown of each cell cluster defined using scRNA-seq data, in terms of the FACS-defined B cell identities. **(C)** Expression of top markers for each cell cluster. Here, only markers with average log fold difference > 0.75 for at least one cluster are included. **(D)** Expression of *CD19*, *MS4A1* (*CD20*), *CD27*, *IGHD*, and *IGHM* across cell clusters.

distinct branches of B cell clusters were seen in three dimensions (**Figure 2A**); 1) a transitional and naïve branch, 2) a classical memory and DN1 branch, 3) an IgM and DN3/2 branch and 4) a separate branch for the IgE-high DN4s (**Figures 2A, B**). In branch 1 the clear direction of development was transitional to naïve to memory and markers such as *CD38* and *TCL1A* are lost in the progression (**Figure 2C**). The relationships between the different memory populations are more complex (**Figures 2A, D, E**). We therefore developed a new method, by performing geometry-based calculations on the mapped velocity landscape (**Supplementary Figure S2**; also see Materials and Methods), to summarize the directionality, strength and position of individual RNA velocity streams between individual cell clusters. This produced a composite transition score (see Materials and Methods) that quantifies the developmental flow of cell clusters (**Figure 2F**) which can be mapped conceptually onto the inferred trajectory (**Figure 2G**). From this the M-mem2 population almost appears as a separate originating singularity with flow

out of the IgM memory cluster, to DN3/2 and C-mem1, but with relatively little flow from Naïve into the M-mem clusters themselves (**Figures 2F, G**). The unidirectional flow from DN3 to termination at DN2 is strong. We also see a very strong flow from DN1/4 towards C-mem1/2 terminating in C-mem2.

A UMAP projection considering only the DN population highlights the distinct and robust partitioning of these cells into 4 subpopulations (**Figure 3A**). DN3 and DN2 display similar immunoglobulin class distribution, although DN3 expressed IgM which is absent in DN2. DN1 is IgA2-rich, while DN4 is IgE-rich (**Figure 3B**).

As a unifying marker both DN3 and DN2 have higher expression of *RHOB* (Ras Homolog Family Member B, involved in intracellular protein trafficking) compared to DN1 and DN4 (**Figure 3C**). The DN2 population at the terminal point of the M-mem/DN development branch is of interest, expressing T-bet (*TBX21*), *CD11c* (*ITGAX*), and lacking *CD21* (**Figure 3D**, **Supplementary Figure S3**) concomitant with an identity of Age-



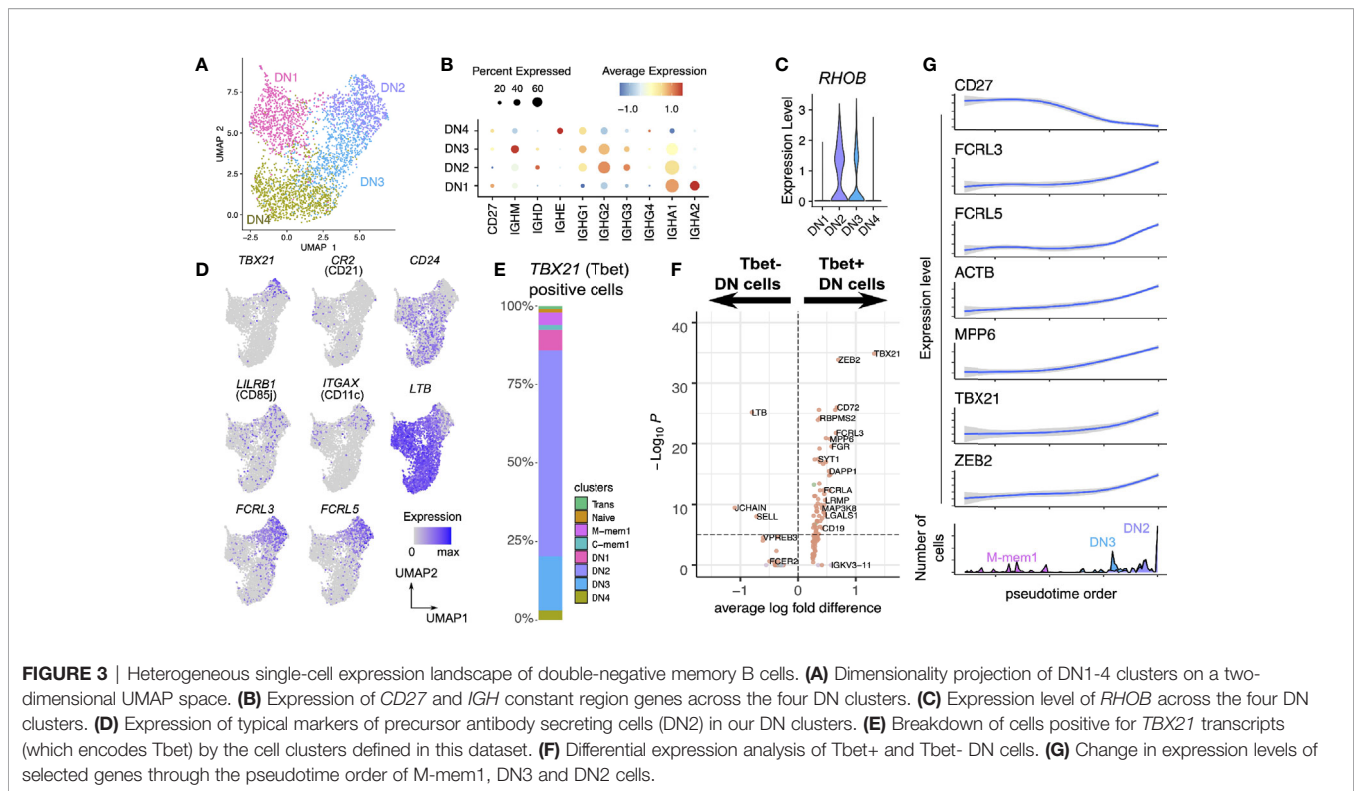
related B cells (34). This population is also enriched in the inhibitory receptors *FCRL3*, *FCRL5* and lacks lymphotoxin-B (*LTB*), *CD24* and *CXCR5* (Figure 3D). These are features in common with a previously identified precursor population for extrafollicular antibody secreting cells (ASCs) (16). T-bet has been shown to mediate IFN γ -dependent (but not IL4 dependent) development of antibody secreting cells from B cells (35). With this in mind we looked at all cells that were T-bet positive and showed that the large majority were in the DN population in the order DN2>DN3>DN1>DN4 (Figure 3E). Other transcriptional markers for the T-bet+ cells include *ACTB*, *MPP6*, (Figures 3F, G) and *ALOX5AP*, *GSTP1*, *LAPTM5* (Figure 4A)

The DN1 population expresses high levels of *IGHA2* and *JCHAIN* (Figure 4). Upregulation of Transgelin-2 (*TAGLN2*) expression suggests an activated state (36). Inferred protein-protein interaction network (PPIN) analysis of the DN1

population shows genes enriched in proteins involved in actin filament formation and organization (Figure 4C). This population falls on a separate differentiation pathway from DN3 and DN2 and more closely resembles classical Memory B cells in differential gene expression (Figures 1C, 2G).

The IgE-expressing DN4 population has very high levels of *FCER2* expression, the low affinity receptor for IgE, as well as *IL4R*, *IL13RA1*, and the co-stimulator *CD40* (Figures 4D, E); all forming part of the same activation network (Figure 4F). This IgE population is therefore in an active state and suggests the donor was undergoing an allergic response.

Traditionally human memory B cells were known as CD27+IgD- and we have designated these “classical memory”. The transcriptomic clustering reveals two different subpopulations in our data: C-mem1 and C-mem2. The main differences between the two are in expression of mitochondrial and ribosomal genes



(Supplementary Figure S5A). This might reflect a low overall gene expression by C-mem1, meaning ribosomal genes are over-represented (Supplementary Figure S5A). Coupled with their relative positioning in the pseudotime/RNA velocity analysis (Figure 2G) this could indicate a more quiescent population awaiting activation for C-mem1, while C-mem2 are more active with higher active mitochondrial gene expression. Comparison of C-mem with DN indicates that DN cells have higher levels of *CD20* (*MS4A1*) and HLA genes (Figure 5A). Analysis of individual genes and inferred PPINs (Figure 5B; Supplementary Figure S5B) shows genes predominantly involved in protein translation are expressed by the C-mem1, while actin networks involved in movement, division and immune synapse formation, form the top networks for C-mem2.

IgM memory cells display a number of genes of interest (Figure 6A) including *CD44* involved in cell to cell interactions and activation (37, 38) along with *MARCKS* involved in actin cross linking (39), *TCF4* (encoding E2-2) which recognized the E-box binding site originally identified in immunoglobulin enhancement (40) and *SMARCB1* which relieves repressive chromatin structure (41). Such a signature of expression would suggest an active group of cells involved in cell-to-cell interaction. We also highlight very strong expression of *CD1C*, a marginal zone B cell marker (42), in a small but distinct number of IgM memory cells.

The two IgM memory subpopulations are predominately defined by the quantity of transcript rather than expression of any one set of genes, M-mem2 containing far fewer transcripts. Taken together with the position of these two subsets relative to one another in the pseudotime and velocity analysis, these suggest an activation pathway from a quiescent population M-

mem2 to the active M-mem1 population. Most of the gene markers shown in Figure 6A are expressed in both the M-mem1 and M-mem2 clusters (Supplementary Figure S5) except *CD1C* and *AP3B1* which are uniquely expressed in M-mem1 cells (Figure 6C). Both genes are involved in presentation of lipid antigens (43). Very few other genes were differentially expressed between M-mem1 and M-mem2 (Figure 6D). The inferred protein interaction networks were more informative, showing translational housekeeping networks dominating the M-mem2 population while M-mem1 networks include both cell adhesion and structural networks suggestive of cell binding (Figure 6E).

In order to confirm these B cell populations PBMCs from two additional people were also processed. We projected the transcriptomes of cells from all three samples to the same UMAP dimensionality-reduced space, and demonstrated that all the same cell populations exist in all individuals (Figure 7A); albeit in some cases at different ratios, as is the case with a lower proportion of DN3 in the two additional individuals (Figure 7B). Additionally, we find a greater overlap between DN1/4 and classical memory highlighting the transcriptomic similarity between these two class switched populations. We also confirmed that cells clustered in the UMAP space show similar expression patterns of marker genes (Supplementary Figure S6) across three individuals.

DISCUSSION

The separation of B cells into functionally different populations is useful in trying to understand immune responses to challenge.

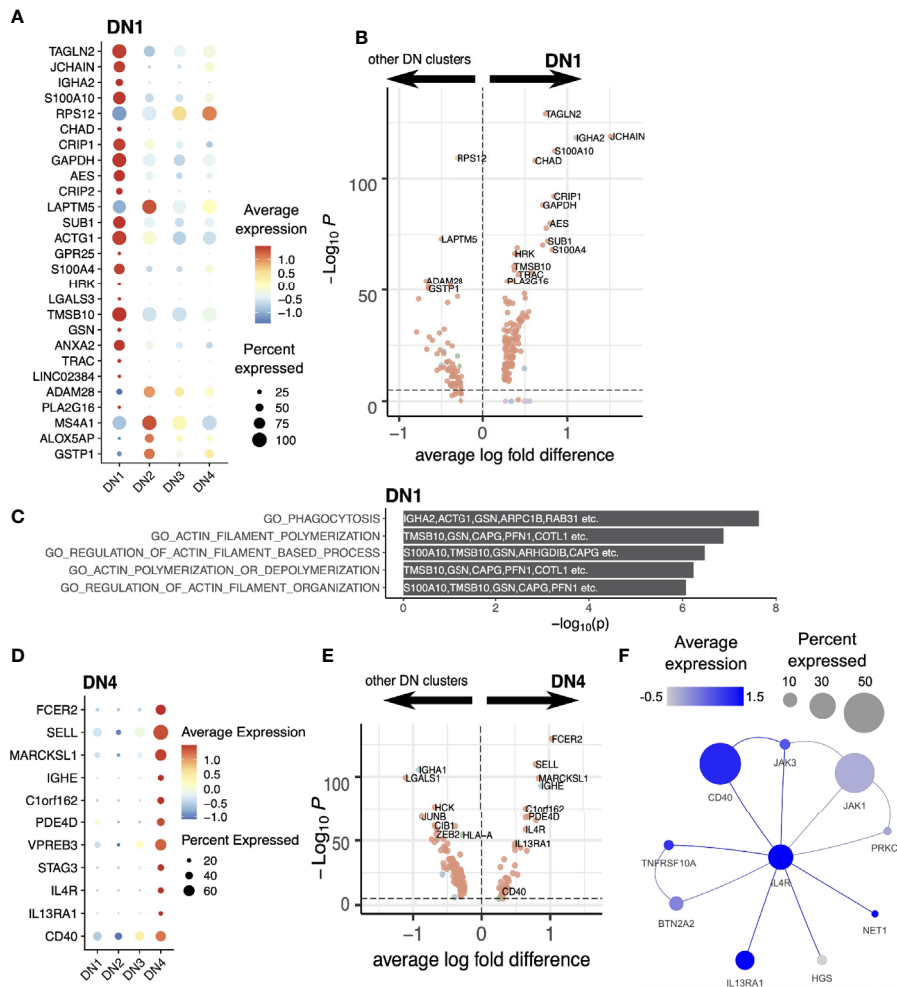


FIGURE 4 | DN1 and DN4 cells. **(A)** Expression level of DN1 markers illustrated with a dot plot. **(B)** Differential expression analysis of DN1 compared with other DN clusters. **(C)** Gene ontologies of DN1 markers. The top 5 pathways in this GO overrepresentation analysis were illustrated. Top markers overlapping these pathways are noted on the plot. **(D)** Expression levels of DN4 markers illustrated with a dot plot. **(E)** Differential expression analysis of DN4 compared with other DN clusters. **(F)** Inferred protein-protein interaction network of *IL4R* and its direct neighbors. The expression levels of these genes in the DN4 cluster were mapped onto the nodes of the network.

Typically, this has been done using surface markers to separate cell types and then ascribing functions to the subsets. However, averaging effects over a heterogeneous population can quite often mask important information. The advent of single-cell technologies enables more precise differentiation between cell types and more understanding of their possible functions based on the genes transcribed. There is a wealth of immunological data based on phenotypic subset separation and maximization of insight to be gained from all available information requires the unification of phenotype and transcriptome techniques. Here we used scRNAseq on the major peripheral B cell populations of an individual, FACS sorted by IgD, CD27 and CD10 (Transitional, Naive, IgM Memory, Classical Memory, Double Negative), to consolidate the phenotype with the transcriptome. Building upon conventional clustering, pseudotime and RNA velocity analyses, we devised a geometry-

inspired method to summarize information from individual RNA velocity streams, and derived quantifications for the transitions between cell clusters (Figures 2F, G). RNA velocity analysis has been widely developed and applied since its conception (32, 44), providing insights into the transcriptional dynamics of single cells. Individual velocity streams could be difficult to interpret, especially in systems where multi-way transitions amongst several cell clusters are possible. It is worth noting that as a result of phenotypic characterization, in order to ensure the clearest transcriptomes of functionally relevant phenotypes, intermediates between states might have been missed. Here we devised a method to build upon these individual velocity streams and provide a summary score, which allowed us to consider the dynamical relationships between all the discovered cell clusters in an unbiased manner (Figure 2G).

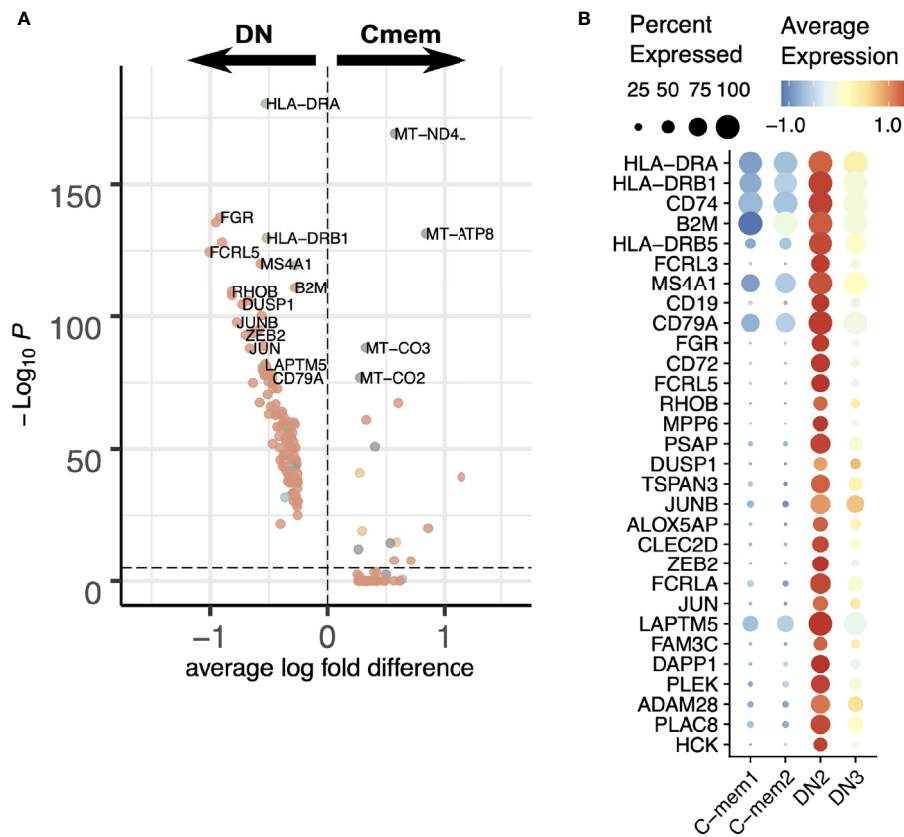


FIGURE 5 | Comparison between classical memory (C-mem) and double-negative (DN) memory B cells. **(A)** Differential expression analysis of C-mem and DN cells. **(B)** Expression of differentially expressed genes between C-mem1/2 and DN2/3 visualized in a dot plot.

We have found that the DN population (IgD-CD27⁻) clusters into four different sub-populations. Our DN1/4 populations both have increased levels of switched Ig genes and *CD24* expression. Sporadic expression of *CD21* and *CD11c* in the DN1/4 makes them hard to identify, however, *CD24* and switched Ig expressions suggests they are ‘precursor memory’ or ‘switched memory’ cells (16, 45). This is highlighted by the RNA velocity data which displays high flow between our precursor memory (DN1/4) and classical memory populations whose development is closely linked (46). The ‘closeness’ of this relationship and flow between populations is further highlighted by *HB34* and *HB78* and the proximity of DN1/4 cells to the classical memory cells; particularly when there is only partial coverage in the transcriptome of the main marker (*CD27*) for differentiating the two populations.

Our DN2 population differentially expressed more *TBX21* and *FCRL5* but no *CD24*, matching the descriptions of Age-related B cells and extrafollicular ASCs (8, 14, 16, 34). The DN3 population shares no previously described gene expression profile, but the RNA velocity data shows DN3 cells flow predominantly into DN2 cells. It is possible the DN3

population are activated naïve (aNAV) cells which develop into precursor ASCs, but the only marker we find for the aNAV population is high IgM expression with the other major markers (e.g. increased *TRL7*, *CD11C* and decreased *CXCR5*, *CD21*) lacking sufficient coverage in the transcriptome to make a definitive conclusion. Additionally, our velocity flow from IgM memory to DN3/2 suggests an alternative activation pathway distinct from the activated naïve (aNAV) cell pathway previously described in SLE patients (8).

Together the transcriptome, UMAP analysis and RNA trajectories, all suggest that DN1/4 and DN2/3 are separate DN memory compartments that contribute to two distinctly different branches of development with hardly any crossover between them (**Figures 2F, G**). T-independent B cells responses are known to expand the IgM memory cell population and plasma cell lineage commitment (47–50) which would appear, on the evidence provided here, to be occurring through an IgM memory/DN3/DN2 pathway. We therefore suggest that these two pathways represent the T-independent (DN2/3) and T-dependent (DN1/4) B cell developmental pathways. The end points of both these pathways, DN2 and C-mem2, have RNA

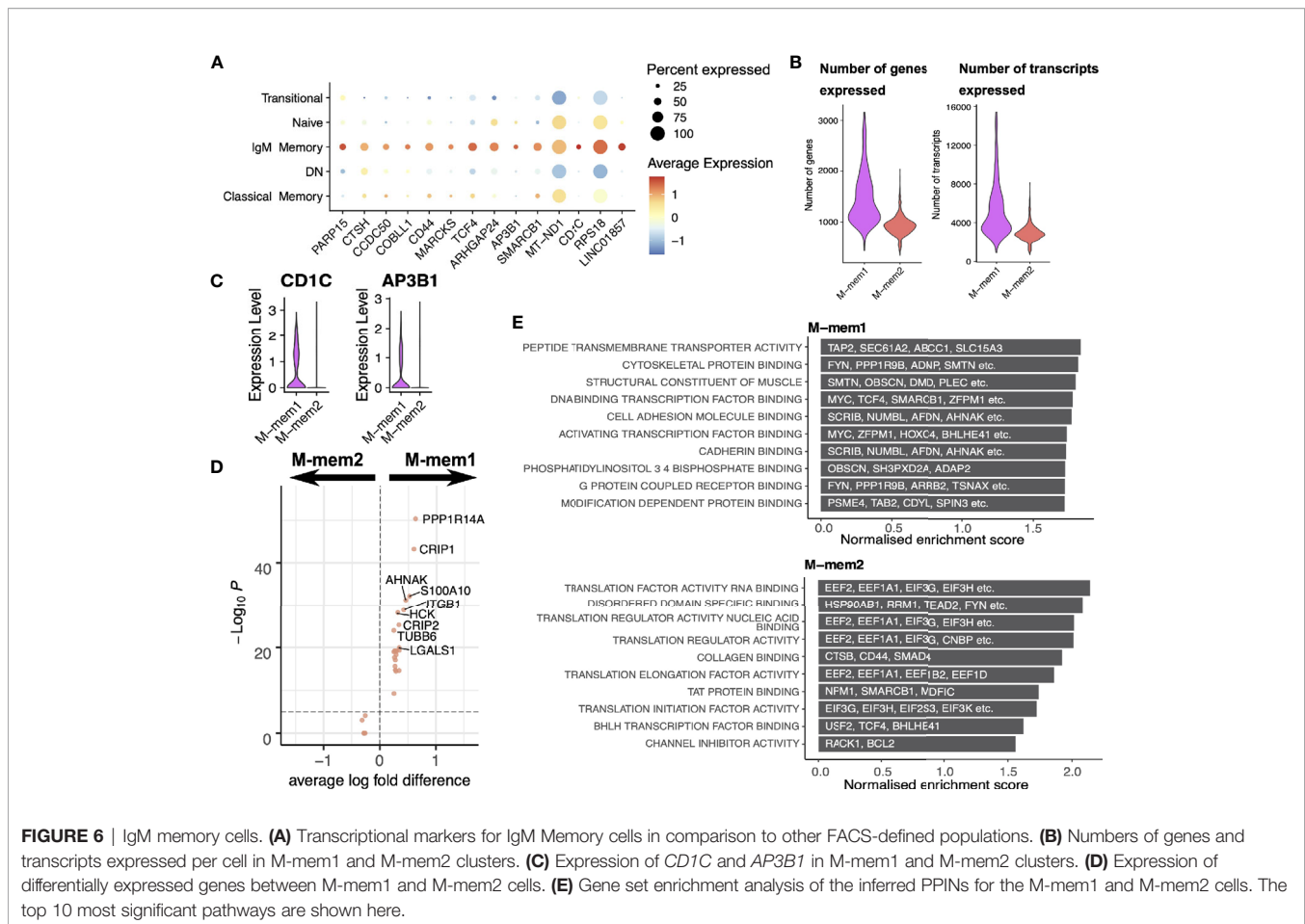


FIGURE 6 | IgM memory cells. **(A)** Transcriptional markers for IgM Memory cells in comparison to other FACS-defined populations. **(B)** Numbers of genes and transcripts expressed per cell in M-mem1 and M-mem2 clusters. **(C)** Expression of *CD1C* and *AP3B1* in M-mem1 and M-mem2 clusters. **(D)** Expression of differentially expressed genes between M-mem1 and M-mem2 cells. **(E)** Gene set enrichment analysis of the inferred PPINs for the M-mem1 and M-mem2 cells. The top 10 most significant pathways are shown here.

velocity arrows continuing to point outwards which may indicate further development, for example into ASCs.

IgM memory cells display several genes involved in actin regulation and cell to cell interactions that might suggest formation of B cell synapses. Expression of the gene *MARCKS* has recently been found to increase lateral motility of the BCR by modulating the cytoskeleton (39), *CD44* is a known lymphocyte activator (37, 38), *TCF4* (E2-2) a developmental transcription factor which regulates the IgG enhancer (24, 40), and *SMARCB1* a chromatin re-modeler that relieves repressive structures (51). The presence of a distinct number of IgM cells strongly expressing *CD1C*, a marginal zone B cell marker (52), is supportive of the idea that these are circulating marginal zone cells following a T-independent pathway of differentiation (42, 53, 54); this separation of IgM memory into two clusters is clearly visible in all three patients. In contrast to other clusters, which have limitations as to their interactions with other groups, the RNA velocity data seems to suggest that cells from the Mmem2 cluster could give rise to cells in many different groups in all branches of the model (Figure 2F).

It is worth noting that the DN1 population is enriched in IgA, particularly genes encoding IgA2 and J chain. This, despite exclusion of immunoglobulin related genes from the clustering,

indicates that IgA2-expressing cells are functionally different from the IgG-expressing cells (Figure 3B). In a similar fashion, we also found an IgE-expressing sub-population of DN cells, DN4. This population is over represented in the data due to our enrichment of the DN subset by FACS sorting and, in reality, represents approximately 0.5% of total B cells in this volunteer; DN making up 1.46% of this patient's B cells (Supplementary Figure S1) and 34.21% of cells in the DN clusters are DN4. Even this level is unusually high and is assumed to be the result of exposure to cat hair 3 days prior to sampling although no symptoms were present at blood draw. CD27-IgE+ cells have been seen previously (55), alongside CD27+IgE+ cells, but in these data DN4 is the only cluster to show IgE expression (Figures 1C, 3B). The study of IgE-expressing cells is usually hampered by their scarcity, the poor polyadenylation of IgE transcripts and the difficulties in distinguishing IgE-expressing cells from those binding to IgE via FcεRII (56, 57). Single cell transcriptomics can circumvent the antibody staining issues and shows here that this population seems to be activated, given the evident RNA velocity stream into and out of DN4 from other clusters (Figure 2G). The presence of *CD24* transcripts and RNA velocity flow into the classical memory population suggests these are memory precursors, although this is difficult to reconcile with

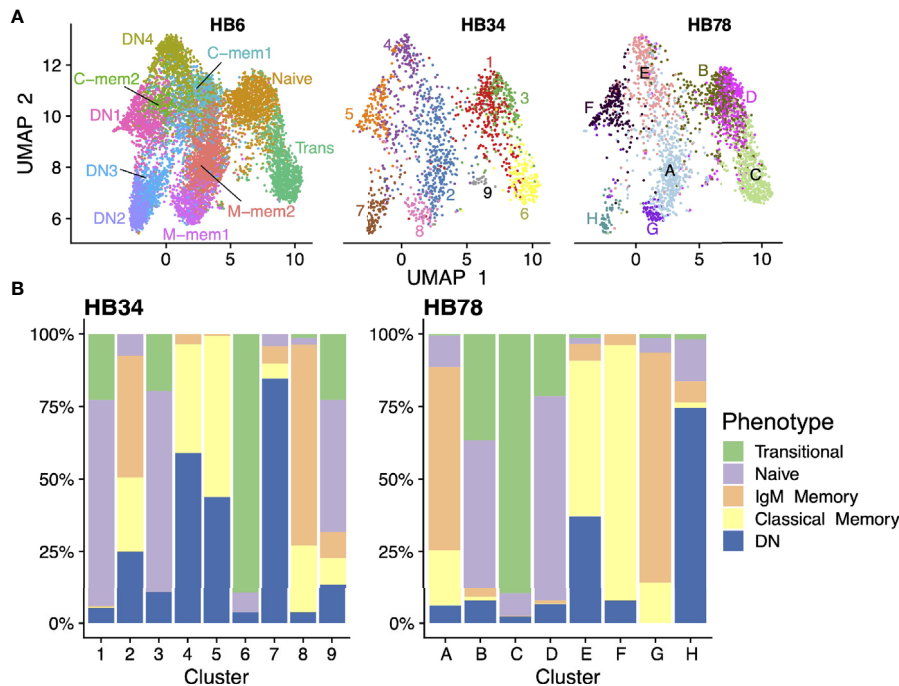


FIGURE 7 | Validation in additional donor samples. **(A)** scRNAseq data of three individuals (HB6 [on whose cells all analyses above were based], HB34, and HB78) were projected onto the same dimensionality-reduced space using UMAP. Cells were clustered separately in each sample and colored accordingly. **(B)** Breakdown of each cell cluster in the HB34 and HB78 data in terms of their phenotypic identities.

the lack of IgE in the classical population. The tendency for some DN4 velocity arrows to point away from the UMAP plot might indicate that these cells are precursor to ASCs, as we have suggested for DN2 and C-mem2.

We also briefly note that transitional and naive cell populations are transcriptionally homogenous. The T1, T2 and T3 transitional cells that have previously been distinguished by gradations of surface markers such as CD10, or CD24+CD38+, are not clustered separately in this data; potentially the result of read depth issues inherent in single-cell approaches and therefore requiring surface marker annotation (e.g. CITE-seq). RNA velocity highlights the highly directional development of transitional cells into naive cells, but only indicates small velocities out of the naive pool into memory. This may look different in an immune challenged volunteer where one would expect active differentiation between naïve and memory. However, the naïve to memory differentiation processes may occur in tissue and not be apparent in the blood.

The occasional disparities between mRNA and protein level gene expression means that markers in the transcriptome do not always translate into the proteome and *vice versa*. Our atlas of peripheral blood B cells can be used as a tool to identify the same cell subsets in other single-cell transcriptome datasets, using referenced-based bioinformatic approaches. It can also be used to identify new markers for tractable methods of B cell sorting (**Supplementary Figure S6**). It is worth noting that transcriptome data can suffer from lack of sequencing depth, so previously well-known phenotypic markers may not always be reliable at a single cell

level. For example, CD27 suffers from incomplete transcript coverage. When clustering data from B cells we strongly suggest removal of variable genes and immunoglobulin constant regions as these tend to skew UMAP results. For example, the kappa lambda light chain distinction between cells is clear and the resulting skew can mask important transcriptomic differences between other, more functional, distinctions. We also note that in this dataset we only find *TBX21*+ cells in the DN population and not in the naive compartment as with SLE (8, 9, 58). Additionally, we did not find a population of *FCRL4*+ DN cells typical of HIV patients (59), highlighting the uniqueness of these populations to particular immune disorders.

In summary, we show branching pathways of B cell development that appear to separate into T dependent and T-independent pathways. The subset of B cells previously known as “Double negative” contains a variety of different cells belonging to both pathways. DN3 and DN2 being in the T-independent pathway closely related to IgM Memory cells, and DN1 and DN4 being more closely related to classical memory cells. In addition, the serendipitous use of an allergic individual after allergen contact has helped to show an IgE-expressing B cell subset within the DN family (DN4), but clustering on its own separate branch of the RNA velocity pathway. That said, it seems to have strong links to the T-dependent development pathway containing classical memory cells and the DN1 cluster rather than the IgM memory/DN3/DN2 development pathway. The data we have collated can be used to map these B cell subsets in other single cell transcriptomic data.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ebi.ac.uk/arrayexpress/>, E-MTAB-9544.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Surrey Ethics Committee, UEC/2017/052/FHMS. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AS, FF, and DD-W devised and designed the experiment. AS carried out the experiment and data acquisition, with FACS sorting performed by GW and VT. JN and AS performed

REFERENCES

- Fischer M, Klein U, Küppers R. Molecular single-cell analysis reveals that CD5-positive peripheral blood B cells in healthy humans are characterized by rearranged V κ genes lacking somatic mutation. *J Clin Invest* (1997) 100:1667–76. doi: 10.1172/JCI119691
- Klein U, Rajewsky K, Küppers R. Human immunoglobulin (Ig)M+IgD+ peripheral blood B cells expressing the CD27 cell surface antigen carry somatically mutated variable region genes: CD27 as a general marker for somatically mutated (memory) B cells. *J Exp Med* (1998) 188:1679–89. doi: 10.1084/jem.188.9.1679
- Shi Y, Yamazaki T, Okubo Y, Uehara Y, Sugane K, Agematsu K. Regulation of Aged Humoral Immune Defense against Pneumococcal Bacteria by IgM Memory B Cell. *J Immunol* (2005) 175:3262–7. doi: 10.4049/jimmunol.175.5.3262
- Griffin DO, Holodick NE, Rothstein TL. Human B1 cells in umbilical cord and adult peripheral blood express the novel phenotype CD20+CD27+CD43+CD70-. *J Exp Med* (2011) 208:67–80. doi: 10.1084/jem.20101499
- Martin V, Wu YC, Kipling D, Dunn-Walters D. Ageing of the B-cell repertoire. *Philos Trans R Soc B Biol Sci* (2015) 370:e20140237. doi: 10.1098/rstb.2014.0237
- Kardava L, Moir S, Shah N, Wang W, Wilson R, Buckner CM, et al. Abnormal B cell memory subsets dominate HIV-specific responses in infected individuals. *J Clin Invest* (2014) 124:3252–62. doi: 10.1172/JCI74351
- Portugal S, Obeng-Adjei N, Moir S, Crompton PD, Pierce SK. Atypical memory B cells in human chronic infectious diseases: An interim report. *Cell Immunol* (2017) 321:18–25. doi: 10.1016/j.cellimm.2017.07.003
- Jenks SA, Cashman KS, Zumaquero E, Marigorta UM, Patel AV, Wang X, et al. Distinct Effector B Cells Induced by Unregulated Toll-like Receptor 7 Contribute to Pathogenic Responses in Systemic Lupus Erythematosus. *Immunity* (2018) 49:725–39.e6. doi: 10.1016/j.immuni.2018.08.015
- Tipton CM, Fucile CF, Darce J, Chida A, Ichikawa T, Gregoretti I, et al. Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus. *Nat Immunol* (2015) 16:755–65. doi: 10.1038/ni.3175
- Weiss GE, Crompton PD, Li S, Walsh LA, Moir S, Traore B, et al. Atypical Memory B Cells Are Greatly Expanded in Individuals Living in a Malaria-Endemic Area. *J Immunol* (2009) 183:2176–82. doi: 10.4049/jimmunol.0901297
- Moir S, Ho J, Malaspina A, Wang W, DiPoto AC, O'Shea MA, et al. Evidence for HIV-associated B cell exhaustion in a dysfunctional memory B cell compartment in HIV-infected viremic individuals. *J Exp Med* (2008) 205:1797–805. doi: 10.1084/jem.20072683

bioinformatics analysis. All authors contributed to the article and approved the submitted version.

FUNDING

The authors are grateful to UKRI, MRC, and BBSRC for their support (MR/L01257X/2, BB/T002212/1). NCSi4P National Core Study Immunity Programme-Cellular assay.

ACKNOWLEDGMENTS

We thank the donor for their blood.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2021.602539/full#supplementary-material>

- Bulati M, Buffa S, Candore G, Caruso C, Dunn-Walters DK, Pellicanò M, et al. B cells and immunosenescence: A focus on IgG+IgD-CD27- (DN) B cells in aged humans. *Ageing Res Rev* (2011) 10:274–84. doi: 10.1016/j.arr.2010.12.002
- Menard LC, Habte S, Gonsiorek W, Lee D, Banas D, Holloway DA, et al. B cells from African American lupus patients exhibit an activated phenotype. *JCI Insight* (2016) 1:e87310. doi: 10.1172/jci.insight.87310
- Wei C, Anolik J, Cappione A, Zheng B, Pugh-Bernard A, Brooks J, et al. A New Population of Cells Lacking Expression of CD27 Represents a Notable Component of the B Cell Memory Compartment in Systemic Lupus Erythematosus. *J Immunol* (2007) 178:6624–33. doi: 10.4049/jimmunol.178.10.6624
- Bagnara D, Squillario M, Kipling D, Mora T, Walczak AM, Da Silva L, et al. A Reassessment of IgM Memory Subsets in Humans. *J Immunol* (2015) 195:3716–24. doi: 10.4049/jimmunol.1500753
- Sanz I, Wei C, Jenks SA, Cashman KS, Tipton C, Woodruff MC, et al. Challenges and Opportunities for Consistent Classification of Human B Cell and Plasma Cell Populations. *Front Immunol* (2019) 10:2458. doi: 10.3389/fimmu.2019.02458
- You X, Zhang R, Shao M, He J, Chen J, Liu J, et al. Double Negative B Cell Is Associated With Renal Impairment in Systemic Lupus Erythematosus and Acts as a Marker for Nephritis Remission. *Front Med* (2020) 7:85. doi: 10.3389/fmed.2020.00085
- Zickert A, Oke V, Parodis I, Svenungsson E, Sundström Y, Gunnarsson I. Interferon (IFN)- λ is a potential mediator in lupus nephritis. *Lupus Sci Med* (2016) 3:e000170. doi: 10.1136/lupus-2016-000170
- Wu Y-CB, Kipling D, Dunn-Walters DK. The Relationship between CD27 Negative and Positive B Cell Populations in Human Peripheral Blood. *Front Immunol* (2011) 2:81. doi: 10.3389/fimmu.2011.00081
- Dunn-Walters D, Townsend C, Sinclair E, Stewart A. Immunoglobulin gene analysis as a tool for investigating human immune responses. *Immunol Rev* (2018) 284:132–47. doi: 10.1111/imr.12659
- Rozenblatt-Rosen O, Stubbington MJT, Regev A, Teichmann SA. The Human Cell Atlas: From vision to reality. *Nature* (2017) 550:451–3. doi: 10.1038/550451a
- King HW, Orban N, Riches JC, Clear AJ, Warnes G, Teichmann SA, et al. Antibody repertoire and gene expression dynamics of diverse human B cell states during affinity maturation. *Sci Immunol* (2021) 6:eabe6291. doi: 10.1101/2020.04.28.054775
- Rubtsov AV, Rubtsova K, Kappler JW, Marrack P. TLR7 drives accumulation of ABCs and autoantibody production in autoimmune-prone mice. *Immunol Res* (2013) 55:210–6. doi: 10.1007/s12026-012-8365-8

24. Wang L-H, Baker NE. Developmental Cell Review E Proteins and ID Proteins: Helix-Loop-Helix Partners in Development and Disease. *Dev Cell* (2015) 35:269–80. doi: 10.1016/j.devcel.2015.10.019
25. Wilk AJ, Rustagi A, Zhao NQ, Roque J, Martínez-Colón GJ, McKechnie JL, et al. A single-cell atlas of the peripheral immune response in patients with severe COVID-19. *Nat Med* (2020) 26:1070–6. doi: 10.1038/s41591-020-0944-y
26. Wen W, Su W, Tang H, Le W, Zhang X, Zheng Y, et al. Immune cell profiling of COVID-19 patients in the recovery stage by single-cell sequencing. *Cell Discovery* (2020) 6:1–18. doi: 10.1038/s41421-020-0168-9
27. Fan X, Chi X, Ma W, Zhong S, Dong Y, Zhou W, et al. Single-cell RNA-seq and V(D)J profiling of immune cells in COVID-19 patients. *medRxiv* (2020). doi: 10.1101/2020.05.24.20101238
28. Lee JS, Park S, Jeong HW, Ahn JY, Choi SJ, Lee H, et al. Immunophenotyping of COVID-19 and influenza highlights the role of type I interferons in development of severe COVID-19. *Sci Immunol* (2020) 5:1554. doi: 10.1126/sciimmunol.abd1554
29. Cao Y, Su B, Guo X, Sun W, Deng Y, Bao L, et al. Potent Neutralizing Antibodies against SARS-CoV-2 Identified by High-Throughput Single-Cell Sequencing of Convalescent Patients' B Cells. *Cell* (2020) 182:73–84.e16. doi: 10.1016/j.cell.2020.05.025
30. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* (2018) 36:411–20. doi: 10.1038/nbt.4096
31. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, et al. Comprehensive Integration of Single-Cell Data. *Cell* (2019) 177:1888–902.e21. doi: 10.1016/j.cell.2019.05.031
32. La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, et al. RNA velocity of single cells. *Nature* (2018) 560:494–8. doi: 10.1038/s41586-018-0414-6
33. Mohammadi S, Davila-Velderrain J, Kellis M. Reconstruction of Cell-type-Specific Interactomes at Single-Cell Resolution. *Cell Syst* (2019) 9:559–568.e4. doi: 10.1016/j.cels.2019.10.007
34. Cancro MP. Age-Associated B Cells. *Annu Rev Immunol* (2020) 38:315–40. doi: 10.1146/annurev-immunol-092419-031130
35. Stone SL, Peel JN, Scharer CD, Risley CA, Chisolm DA, Schultz MD, et al. T-bet Transcription Factor Promotes Antibody-Secreting Cell Differentiation by Limiting the Inflammatory Effects of IFN- γ on B Cells. *Immunity* (2019) 50:1172–87.e7. doi: 10.1016/j.immuni.2019.04.004
36. Kiso K, Yoshifuji H, Oku T, Hikida M, Kitagori K, Hirayama Y, et al. Transglutinin-2 is upregulated on activated B-cells and expressed in hyperplastic follicles in lupus erythematosus patients. *PLoS One* (2017) 12:e0184738. doi: 10.1371/journal.pone.0184738
37. Rafi A, Nagarkatti M, Nagarkatti PS. Hyaluronate-CD44 interactions can induce murine B-cell activation. *Blood* (1997) 89:2901–8. doi: 10.1182/blood.v89.8.2901
38. Föger N, Marhaba R, Zöllner M. CD44 supports T cell proliferation and apoptosis by apposition of protein kinases. *Eur J Immunol* (2000) 30:2888–99. doi: 10.1002/1521-4141(200010)30:10<2888::AID-IMMU2888>3.0.CO;2-4
39. Xu C, Fang Y, Yang Z, Jing Y, Zhang Y, Liu C, et al. MARCKS regulates tonic and chronic active B cell receptor signaling. *Leukemia* (2019) 33:710–29. doi: 10.1038/s41375-018-0244-4
40. Wöhner M, Tagoh H, Bilic I, Jaritz M, Poliakova DK, Fischer M, et al. Molecular functions of the transcription factors E2A and E2-2 in controlling germinal center B cell and plasma cell development. *J Exp Med* (2016) 213:1201–21. doi: 10.1084/jem.20152002
41. Zhao K, Wang W, Rando OJ, Xue Y, Swiderek K, Kuo A, et al. Rapid and phosphoinositide-dependent binding of the SWI/SNF-like BAF complex to chromatin after T lymphocyte receptor signaling. *Cell* (1998) 95:625–36. doi: 10.1016/S0092-8674(00)81633-5
42. Weller S, Braun MC, Tan BK, Rosenwald A, Cordier C, Conley ME, et al. Human blood IgM “memory” B cells are circulating splenic marginal zone B cells harboring a prediversified immunoglobulin repertoire. *Blood* (2004) 104:3647–54. doi: 10.1182/blood-2004-01-0346
43. Barral DC, Brenner MB. CD1 antigen presentation: How it works. *Nat Rev Immunol* (2007) 7:929–41. doi: 10.1038/nri2191
44. Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol* (2020) 38:1–7. doi: 10.1038/s41587-020-0591-3
45. Fecteau JF, Côté G, Néron S. A New Memory CD27 – IgG + B Cell Population in Peripheral Blood Expressing V H Genes with Low Frequency of Somatic Mutation. *J Immunol* (2006) 177:3728–36. doi: 10.4049/jimmunol.177.6.3728
46. Wu YCB, Kipling D, Dunn-Walters DK. The relationship between CD27 negative and positive B cell populations in human peripheral blood. *Front Immunol* (2011) 2:81. doi: 10.3389/fimmu.2011.00081
47. Hua Z, Hou B. TLR signaling in B-cell development and activation. *Cell Mol Immunol* (2013) 10:103–6. doi: 10.1038/cmi.2012.61
48. Alugupalli KR, Akira S, Lien E, Leong JM. MyD88- and Bruton's Tyrosine Kinase-Mediated Signals Are Essential for T Cell-Independent Pathogen-Specific IgM Responses. *J Immunol* (2007) 178:3740–9. doi: 10.4049/jimmunol.178.6.3740
49. Weller S, Bonnet M, Delagrèverie H, Israel L, Chrabieh M, Maròdi L, et al. IgM+IgD+CD27+ B cells are markedly reduced in IRAK-4-, MyD88-, and TIRAP- but not UNC-93B-deficient patients. *Blood* (2012) 120:4992–5001. doi: 10.1182/blood-2012-07-440776
50. Simchoni N, Cunningham-Rundles C. TLR7- and TLR9-Responsive Human B Cells Share Phenotypic and Genetic Characteristics. *J Immunol* (2015) 194:3035–44. doi: 10.4049/jimmunol.1402690
51. Yaniv M. Chromatin remodeling: From transcription to cancer. *Cancer Genet* (2014) 207:352–7. doi: 10.1016/j.cancergen.2014.03.006
52. Fairhurst RM, Wang CX, Sieling PA, Modlin RL, Braun J. CD1-restricted T cells and resistance to polysaccharide-encapsulated bacteria. *Immunol Today* (1998) 19:257–9. doi: 10.1016/S0167-5699(97)01235-8
53. Weller S, Faili A, Garcia C, Braun MC, Le Deist F, De Saint Basile G, et al. CD40-CD40L independent Ig gene hypermutation suggests a second B cell diversification pathway in humans. *Proc Natl Acad Sci U S A* (2001) 98:1166–70. doi: 10.1073/pnas.98.3.1166
54. Krutzmann S, Rosado MM, Weber H, Germing U, Tournilhac O, Peter HH, et al. Human immunoglobulin M memory B cells controlling Streptococcus pneumoniae infections are generated in the spleen. *J Exp Med* (2003) 197:939–45. doi: 10.1084/jem.20022020
55. Berkowska MA, Heeringa JJ, Hajdarbegovic E, Van Der Burg M, Thio HB, Van Hagen PM, et al. Human IgE+ B cells are derived from T cell-dependent and T cell-independent pathways. *J Allergy Clin Immunol* (2014) 134:688–97.e6. doi: 10.1016/j.jaci.2014.03.036
56. Karnowski A, Achatz-Straussberger G, Klockenbusch C, Achatz G, Lamers MC. Inefficient processing of mRNA for the membrane form of IgE is a genetic mechanism to limit recruitment of IgE-secreting cells. *Eur J Immunol* (2006) 36:1917–25. doi: 10.1002/eji.200535495
57. Acharya M, Borland G, Edkins AL, MacLellan LM, Matheson J, Ozanne BW, et al. CD23/Fc ϵ R2: Molecular multi-tasking. *Clin Exp Immunol* (2010) 162:12–23. doi: 10.1111/j.1365-2249.2010.04210.x
58. Myles A, Sanz I, Cancro MP. T-bet + B cells: A common denominator in protective and autoreactive antibody responses? *Curr Opin Immunol* (2019) 57:40–5. doi: 10.1016/j.coi.2019.01.002
59. Knox JJ, Buggert M, Kardava L, Seaton KE, Eller MA, Canaday DH, et al. T-bet+ B cells are induced by human viral infections and dominate the HIV gp140 response. *JCI Insight* (2017) 2:e92943. doi: 10.1172/jci.insight.92943

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Stewart, Ng, Wallis, Tsioligka, Fraternali and Dunn-Walters. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.