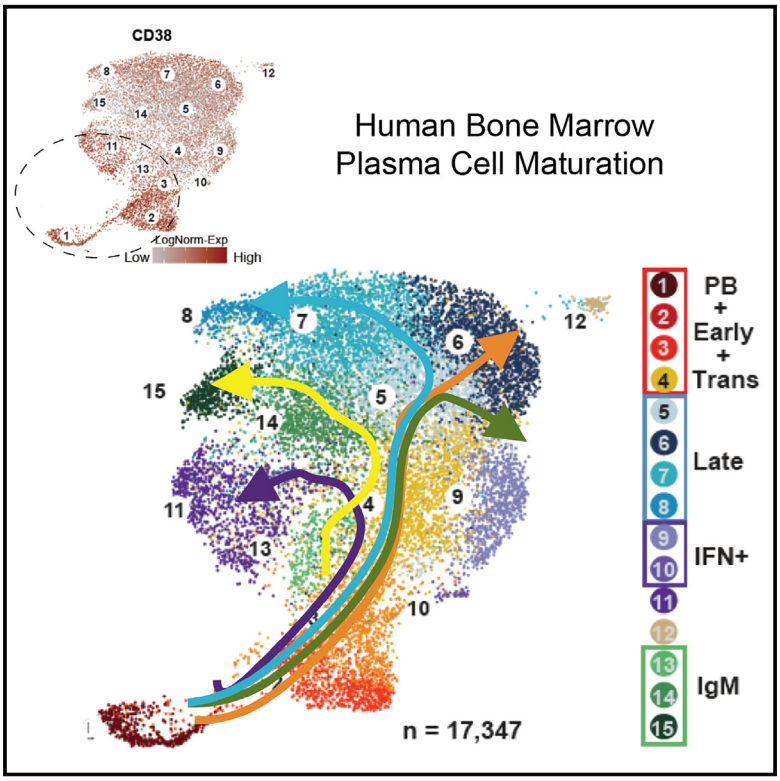
Cell Reports

Article

Understanding heterogeneity of human bone marrow plasma cell maturation and survival pathways by single-cell analyses

Graphical abstract



Authors

Meixue Duan, Doan C. Nguyen, Chester J. Joyner, ..., Ignacio Sanz, Greg Gibson, F. Eun-Hyung Lee

Correspondence

f.e.lee@emory.edu

In brief

Although long-lived plasma cells (LLPCs) have been identified in human bone marrow (BM), the heterogeneity and mechanistic maturation trajectories have not been clearly defined. Duan et al. demonstrate the heterogeneity of the human BM plasma cells by providing a single-cell transcriptional atlas with a roadmap to LLPC maturation.

Highlights

- Early BM ASCs contain clusters with high class I and II, IgM, IFN, and mitochondrial genes
- Late BM ASCs are distinguished by G2M checkpoints, mTOR signaling, and TNFRSF members
- Early ASCs have higher CD38 expression and are selectively depleted with anti-CD38 antibodies
- There is heterogeneity of late BM ASCs by degree of TNF signaling through NF-κB

Duan et al., 2023, Cell Reports *42*, 112682
July 25, 2023
https://doi.org/10.1016/j.celrep.2023.112682

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Understanding heterogeneity of human bone marrow plasma cell maturation and survival pathways by single-cell analyses

Meixue Duan,[1] Doan C. Nguyen,[2,13] Chester J. Joyner,[2,11,13] Celia L. Saney,[2,11] Christopher M. Tipton,[2,3] Joel Andrews,[4] Sagar Lonial,[4] Caroline Kim,[2] Ian Hentenaar,[2] Astrid Kosters,[3] Eliver Ghosn,[3] Annette Jackson,[5,6] Stuart Knechtle,[6] Stalinraja Maruthamuthu,[7] Sindhu Chandran,[8] Tom Martin,[8] Raja Rajalingam,[7] Flavio Vincenti,[9] Cynthia Breeden,[10] Ignacio Sanz,[3] Greg Gibson,[1,12] and F. Eun-Hyung Lee[2,3,12,14,*]

[1]School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA, USA
[2]Department of Medicine, Division of Pulmonary, Allergy, Critical Care, and Sleep Medicine, Emory University, Atlanta, GA, USA
[3]Department of Medicine, Division of Rheumatology, Lowance Center for Human Immunology, Emory University, Atlanta, GA, USA
[4]Department of Hematology and Medical Oncology, Winship Cancer Institute, Emory University, Atlanta, GA, USA
[5]Departments of Immunology, Duke University, Durham, NC, USA
[6]Department of Surgery, Duke University, Durham, NC, USA
[7]Immunogenetics and Transplantation Laboratory, Department of Surgery, University of California San Francisco, San Francisco, CA, USA
[8]Department of Medicine, University of California San Francisco, San Francisco, CA, USA
[9]Division of Nephrology, Department of Medicine, University of California San Francisco, San Francisco, CA, USA
[10]Emory Transplant Center, Department of Surgery, School of Medicine, Emory University, Atlanta, GA, USA
[11]Present address: Center for Vaccines and Immunology, Department of Infectious Diseases, College of Veterinary Medicine, University of Georgia, Athens, GA, USA
[12]Senior author
[13]These authors contributed equally
[14]Lead contact
*Correspondence: f.e.lee@emory.edu
https://doi.org/10.1016/j.celrep.2023.112682

SUMMARY

Human bone marrow (BM) plasma cells are heterogeneous, ranging from newly arrived antibody-secreting cells (ASCs) to long-lived plasma cells (LLPCs). We provide single-cell transcriptional resolution of 17,347 BM ASCs from five healthy adults. Fifteen clusters are identified ranging from newly minted ASCs (cluster 1) expressing MKI67 and high major histocompatibility complex (MHC) class II that progress to late clusters 5–8 through intermediate clusters 2–4. Additional ASC clusters include the following: immunoglobulin (Ig) M predominant (likely of extra-follicular origin), interferon responsive, and high mitochondrial activity. Late ASCs are distinguished by G2M checkpoints, mammalian target of rapamycin (mTOR) signaling, distinct metabolic pathways, CD38 expression, utilization of tumor necrosis factor (TNF)-receptor superfamily members, and two distinct maturation pathways involving TNF signaling through nuclear factor κB (NF-κB). This study provides a single-cell atlas and molecular roadmap of LLPC maturation trajectories essential in the BM microniche. Altogether, understanding BM ASC heterogeneity in health and disease enables development of new strategies to enhance protective ASCs and to deplete pathogenic ones.

INTRODUCTION

The existence of long-lived plasma cells (LLPCs) that provide a lifetime of humoral protection after vaccination and infection is well established in mice and humans. Activated lymph node B cells differentiate into early antibody-secreting cells (ASCs), which home to bone marrow (BM) niches where a fraction may survive as LLPCs, which we previously identified within the $CD19^{-}CD38^{hi}CD138^{+}$ cell population.[1–6] Whether mere migration to protective niches is sufficient for the establishment of an LLPC compartment or, instead, additional maturation in the BM is required was unclear. Thus, we recently showed that early maturation of nascent ASCs takes place in the BM through morphologic, transcriptomic, and epigenomic changes that presumably enable their ultimate differentiation into LLPCs.[6] Accordingly, we postulated that peripheral ASCs arriving in the marrow undergo further maturation locally to generate *bona fide* LLPC.

We have also developed an *in vitro* human BM mimetic system, containing soluble factors from mesenchymal stromal cells (MSCs), a proliferation-inducing ligand (APRIL), and hypoxia, which sustains ASC survival for up to 56 days in culture, thereby overcoming previous experimental limitations in the field imposed by the rarity and *ex vivo* frailty of ASCs. This approach

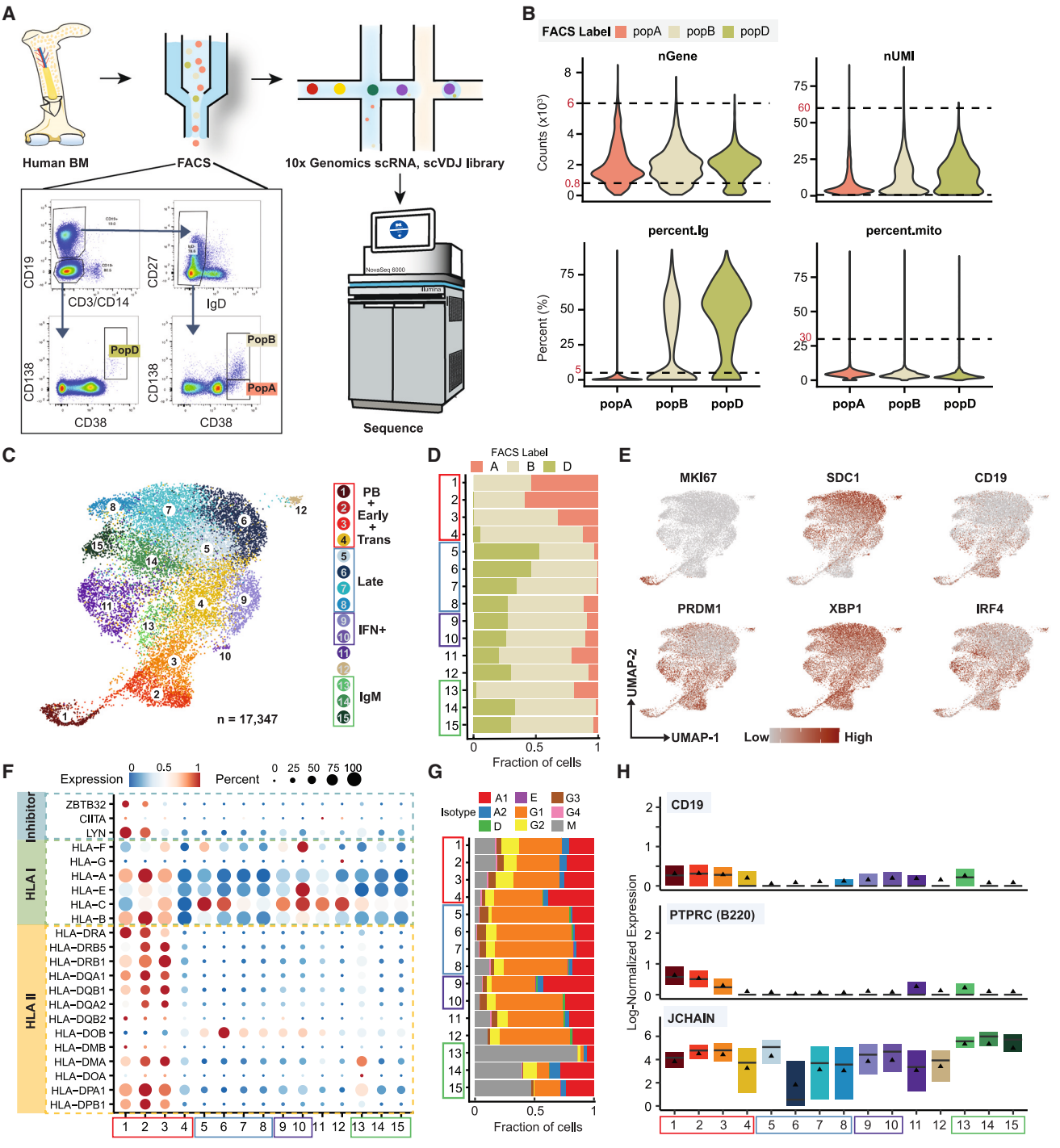


Figure 1. Single-cell transcriptomic profiling of bone marrow PCs

(A) Schematic of single-cell RNA profiling for human BMPCs.

(B) Criterion for removing bad-quality cells. Dashed lines show the cutoffs that are labeled in red.

(C) scRNA-seq cell clusters of combined data from five healthy BMs and visualized in UMAP colored by cell types. Red and blue boxes highlight the early and late stage of BMPC maturation, respectively. The purple box highlights the path toward IFN-response PC subgroups and the green box highlights IgM-dominant cell populations.

(D) The fraction of cells from fluorescence-activated cell sorting (FACS)-sorted cell population in each cell subgroup identified in (C).

(E) Key PC-associated master gene expression. The redder the dot, the higher the log-normalized gene expression.

(legend continued on next page)

enabled a molecular roadmap charting the progression of nascent ASCs into mature ASCs demarcated by upregulation of CD138 expression and downregulation of CD19 as early as day 14.[7] This early work showed that engagement of sequential transcriptomic and epigenetic programs promoting resistance to apoptosis is essential for survival.[6] Although candidate LLPC BM maturation programs were revealed, bulk analyses could not directly interrogate the heterogeneity of the BM LLPC compartment.

Here, we have performed extensive single-cell analysis of BM ASCs. Our studies identify a large degree of cell heterogeneity and maturation trajectories. Starting from nascent KI67+ ASCs with the highest major histocompatibility complex (MHC) class II expression, we could assign early and late ASCs distinguished by differences in G2M check points, E2F targets, mammalian target of rapamycin (mTOR) signaling, and metabolic pathways for fatty acid and oxidative phosphorylation as well as differential expression of members of the tumor necrosis factor (TNF) receptor superfamily and CD38 expression. Finally, terminal differentiation of late immunoglobulin (Ig) G ASCs followed two distinct paths with differential utilization of the TNFα signaling pathway via nuclear factor κB (NF-κB). In all, this study provides a single-cell atlas and molecular roadmap trajectories of LLPC maturation in the human BM.

RESULTS

Single-cell transcriptomic profiling of human plasma cells in the BM

To characterize the heterogeneity of human bone marrow plasma cells (BMPCs), the three major populations previously identified were sorted from five healthy adults without recent immunization or infection[3]: pop A ($CD19^{+}CD38^{hi}CD138^{-}$), pop B ($CD19^{+}CD38^{hi}CD138^{+}$), and pop D ($CD19^{-}CD38^{hi}CD138^{+}$) (Figure 1A). Although pop D constitutes the main reservoir of LLPC, to test the distinct molecular programs reflecting their generation, regulation, and survival, we evaluated the BM ASCs at the single-cell level using 5′-directed cDNA library construction to characterize both the single-cell transcriptomic profiles (single-cell RNA sequencing [scRNA-seq]) and matching single-cell V(D)J repertoire sequencing (scVDJ-seq). After exclusion of non-ASC contaminating cells (including B cells), low-quality cells, dying cells, and doublets, the remaining cells were identified as *bona fide* ASCs with the characteristic expression of *XBP1*, *IRF4*, *PRDM1*, *CD27*, *CD138*, and *CD19* (Figures 1B, 1C, 1E, S1C, and S2A). In total, we retained 17,347 ASCs.

Identification of 15 clusters of BM ASCs

Clustering was performed with the removal of known Ig genes and without knowledge of subsets A, B, and D *a priori*. However, our initial clustering required the elimination of two subsets falsely identified solely on the basis of very high expression of two misannotated Ig VH genes. Ultimately, we identified 15 robust clusters representing different cell states or types based on gene expression profiles (Figures 1C, 1E, and S2A–S2C). After establishing the 15 clusters, we re-incorporated the Ig genes and found no specific VH gene driving a particular cluster. These 15 distinct clusters (Figure 1C) could be grouped as early (clusters c1–3), transitional or intermediate (c4), and late (c5–8) ASC based on human leukocyte antigen (HLA) class I and II expression (Figure 1F). In addition, we identified two interferon (IFN)-response-dominated clusters (c9 and 10); and three IgM-predominant clusters (c13–15, Figure 1G) that contained the majority of BM IgM ASCs. Finally, we identified a mitochondrial-high cluster (c11) with characteristics of early ASCs and a minor cluster (c12), likely representing dying ASCs. The differentially expressed marker genes used to adjudicate each cluster are shown in Figure S2F and Table 1.

Separation of early and late clusters by HLA class II gene expression

c1 distinctly expressed MKI67, CD43 (*SPN*), and CD27 and had low expression of CD20 (*MS4A1*) (Figures S2F and S2H). This pattern is consistent with the definition of proliferative plasmablasts (PB) or early-minted ASCs, representing new BM arrivals from active immune responses ($CD20^{low}CD27^{high}CD43^{high}$).[8] As they mature into resting LLPCs, proliferative ASCs typically shed B cell markers such as surface Ig and MHC class II genes (HLA gene complex).[3,4] c2 and c3 lacked Ki67 expression but retained the highest levels of HLA class II genes, which subsequently decreased in the transitional c4 (Figure 1F). Thus, the initial ASCs seeding the BM could be further divided into early stages, including proliferative and non-proliferative cells (c1, and c2 and 3, respectively), and transitional stages (c4). Interestingly, c1–3 were devoid of pop D, which could be first detected at low frequencies within c4 (Figure 1D).

As expected, *PRDM1*, encoding BLIMP1, which drives plasma cell (PC) commitment and extinguishes class II transactivator *CIITA* and MHC class II gene expression, was expressed from the early ASC stages[9] (Figure 1F). In turn, *CIITA* was universally extinguished on nearly all BM ASC subsets except for a small fraction of early c1. HLA-DOB, which suppresses peptide loading of class II molecules, was upregulated in the late ASC clusters, which is consistent with reports that HLA-DOB is less affected by *CIITA* than other family members.[10] HLA class I expression followed similar trends except for HLA-C. In all, HLA class II expression provides clear separation of early and late BM ASC subsets.

Marker genes of human BM ASC clusters

We next identified cell-type-specific markers across the 15 BM ASC subgroups (Figure S2F; Table S1). In addition to the defining features, Ki-67 and HLA class II, other markers of early ASCs included *LYN* and *ZBTB32*. LYN is a tyrosine protein kinase downstream of BCR signaling that can diminish proliferation while driving terminal PC differentiation[11] (Figure 1F). Interestingly, LYN deficiency leads to a 20-fold increase of BM PCs

(F) Dot plot for expression of human MHC class I, II, and inhibitors of class II genes. Colors represent minimum-maximum normalized mean expression of marker genes in each cell group, and sizes indicate the proportion of cells expressing marker genes.
(G) The fraction of isotypes identified by scVDJ-seq data in each cell subgroup identified in (C).
(H) Boxplots showing the expression levels of the indicated genes. The solid triangle represents the average value.

Table 1. Differentially expressed marker genes define the 15 clusters

Cluster ID	Cell label	Markers	Cluster ID	Cell label	Markers
1	PB	MKI67, UBE2C, BIRC5, SPN, TUBB	9	pc9	ATF5, PSAT1, CEBPB, ERN1
2	early 1	CD52, PEBP1, MHC class II (except HLA-DOB)	10	IFN+	MX1, XAF1, IRF7, STAT1
3	early 2	CD79A, CD74	11	Mito-high	MALAT1, IRF4, PRDM1, ZBTB20, MTRNR2L12, FCRL5
4	trans	IGLV3-25	12	pc12	JUND, CXCR4, SIK1B
5	late 1	SMOC1, MOXD1	13	IgM1	CCDC88A, FOXP1, IGHM, KLHL14, MS4A1
6	late 2	CD9, CST3, CD63, TIMP1	14	IgM2	JCHAIN, RGS2, TNFRSF4
7	late 3	RHOB, CDKN1A, RGCC	15	IgM3	CCL3, CCL4
8	late 4	JUNB, FOS, EGR1, NR4A1	–	–	–

and induces autoimmune disease in mouse models.[12] Thus, progressive extinction of LYN expression in late clusters may contribute to enhanced survival of the more mature PCs.[13] Also informative for early maturation were B cell markers *CD19* and *PTPRC* (B220), which were strongly expressed in early IgG-dominant c1–c3 as well as in c11 and IgM-dominated c13 (Figures 1G and 1H). The progressive loss of *CD19* and *PTPRC* during B cell differentiation into PC and the ability of $CD19^{+}$ $B220^{+}$ cells to secrete low-affinity IgM antibodies[14,15] suggests early populations.

The transitional nature of c4 was documented by the progressive loss of early features including MHC class II and initial expression of late markers. In turn, c5 shared with c4 higher levels of transcripts characteristic of late subsets, including *MDK*, *ITM2B*, *LMNA*, *AREG*, and *TIMP1* (Tables 1 and S1).

Late c6–c7, and to a lesser extent c8, expressed higher levels of *CD9* and *CST3*. The CD9 tetraspanin modifies multiple cellular events of relevance for BM ASCs, including adhesion, migration, proliferation, and survival. Although *CD9* expression has been considered to mark GC-derived human PCs,[16] *CD9* is also considered a marker of mouse PCs derived from marginal zone and B1 B cells in primary T-dependent responses.[17]

Finally, genes involved in regulation of cell-cycle arrest, apoptosis, and survival were preferentially expressed in c7 (*CDKN1A*/p21 and *RGCC*) and c8 (*JUNB*, *FOS*, *EGR1*, *NR4A1*; Table 1). P21 is a potent cyclin-dependent kinase inhibitor whose expression regulates cell-cycle progression and is tightly controlled by the tumor suppressor p53, which mediates cell-cycle arrest and can promote apoptosis in a context-dependent fashion.[18] Early Growth Response (*EGR1*) is a nuclear transcriptional regulator of multiple tumor suppressors, including p53. Notably, *EGR1* is rapidly induced by growth factors, apoptotic signals, and hypoxia, a feature of the BM microenvironment that determines ASC survival.[19,20] Of note, it has been shown to play a non-redundant role in PC differentiation.[21] While the induction of the orphan nuclear receptor *NR4A1* (Nur77) is best recognized as a consequence of antigen receptor engagement in B and T cells,[22] *NR4A1* is also induced by other stimuli, including endoplasmic reticulum (ER) stress, which is present at high levels in PCs secondary to high Ig synthesis.[23] Interestingly, *NR4A1* can modify the pro/anti-apoptotic balance of the Bcl2 family and has enhanced binding to anti-apoptotic Bcl-B, which is prominently expressed in PCs.[24,25]

Notably, c7 and c8 are separated from other late ASC populations (c5 and c6) by the highest levels of the TNFα-signaling NF-κB pathway. These marker genes, especially of the late clusters, provide essential clues to unique mechanisms that support survival.

IFN-responsive genes (IFN+) were the major markers in c9 and c10. C9 highly expressed *ATF5*, *CEBPB*, and *ERN1*, whereas c10 expressed the classical IFN-response signature (*STAT1*, *IRF7*, *ISG15*, *IFITM1*, *IFI6*, *MX1* and *OAS1*).

We identified all five isotypes (IgM, IgG, IgA, and a small fraction of IgD and IgE), as well as the four IgG subclasses (IgG1–IgG4), in BM ASCs (Figure 1G). The majority of ASC clusters were dominated by IgG, predominantly IgG1. Compared with early stages, late stages (c5–8, and c10–12) had expanded proportions of the IgG isotype ($p < 2.2e^{-16}$).

Mature IgM ASC populations in the human BM

C13–15 contained the largest repository of IgM cells (40%–80%) (Figure 1G), which represented a large majority of c13. Interestingly, these IgM-predominant clusters displayed an overall distinct transcriptome and could in turn be split into early (c13) and late clusters (c14 and 15), according to the HLA expression and scarcity of pop D cells in c13 (Figure 1D). However, it is likely the maturation programs of IgM versus IgG trajectories would follow different paths. Thus, the IgM clusters are candidates for the human counterpart to mouse IgM LLPCs, which accumulate in the spleen in a GC-independent fashion and contribute to protective IgM responses.[26–28]

IgM ASCs (c13–15) are further defined by higher expression of *CCR10*, *JCHAIN*, *FHL1*, *PHACTR1*, and *RAMP2* relative to IgG-dominant cell populations (Figures S3A and 1H). Since C-C motif receptor 10 (CCR10) and JCHAIN are widely expressed in mucosal ASCs,[29,30] the IgM-dominant cell populations may have mucosal origins. Additional differentially expressed genes (DEGs) distinguishing IgM- from IgG-dominant clusters included 77 *CCR10* co-expressed genes summarized in Figure S3B. These *CCR10*-related genes included *EBI2* (*GPR183*), *JCHAIN*, *FOXP1*, and several genes involved in regulation of lymphocyte activation, such as *FCRL3*, *TNFSF9*, *CLECL1*, and *FGL2*. *FOXP1* is known to impair the formation of germinal centers (GCs) and to repress human PC differentiation but are highly

expressed in mature primary human B cells (e.g., naive B, memory B cells) as well as in mouse follicular B and B-1 cells.[31,32] This profile is consistent with a GC-independent extra-follicular origin of the late IgM ASCs.[32,33]

Relationship of BM ASC single-cell clusters to previously described ASC populations

Consistent with previous models of BM ASC maturation,[3] *CD19* expression was notable in early (c1–4) and extinguished in late clusters, while SDC1 (*CD138*) expression was highest in the late clusters (Figure 1E). *XBP1*, an essential transcription factor (TF) associated with the ASC unfolded protein response (UPR), was increased in all clusters. Contrary to mouse studies, *PRDM1* expression was higher in early subsets and did not continuously increase in all late subsets (Figures 1E and S4A), a finding noted by bulk transcriptomes and intracellular BLIMP1 staining.[3] Thus, *PRDM1* upregulation was limited to a small fraction of late clusters c7, c8, and c15, indicating that the high expression of this essential PC TF may be required for early ASC differentiation yet less important for terminal differentiation or maturation and long-term survival.[3,34,35]

The majority of pop D was found in the late stages of ASC maturation and, Interestingly, it was distributed across c5–c8, which also included a minority of pop A cells but transcriptionally resembled pop B (Figures 1D and S4D). A smaller fraction of pop D was spread across late c5–c12 and c14–c15. In contrast, pop B contributed nearly half of the cells in each of the early and late clusters, again suggesting that it corresponds to an intermediary BM population.

Both these early (c1–4) and a "mito-high" subgroup (c11) had higher *CD19* and *PRDM1* expression, higher number of total detected genes, and lower percentage of Ig transcripts relative to late mature ASCs (Figures 1E and S4A–S4C). Combined with their higher expression of *IRF4*, *ZBTB20*, and *FCRL5* (Table 1), this cluster appears to belong with the early ASC populations.[36]

Identification of four maturation paths for IgG BM ASCs through trajectory analysis

IgG ASCs comprise the major isotype in the late clusters; thus, we used Slingshot to construct maturation trajectories to focus on IgG BM ASCs and visualize the dynamic alteration of gene sets by pseudotime for each path onto the uniform manifold approximation and projection (UMAP) (Figures 2A and 2B). This approach predicted five maturation trajectories. Path 1 and 2 project to late c6 and c7 and 8 respectively. Path 3a and 3b, which directed to the same terminal c9, were merged into a single path 3. Finally, path 4 led toward c11. There was distinct separation between c1 and c2, but the remaining non-proliferative clusters in each path tend to be shared between adjacent stages/clusters. This result suggests the maturation of BM ASCs is more of a continuum of functional processes rather than an ordered sequence of discrete cell states.

Next, using hallmark pathway enrichment scores,[37] we discovered four patterns describing the dynamic alteration of gene sets during BM ASC maturation (Figures 2C and S5A). From each pattern, a pathway example for visualization is shown in Figure 2D. Projecting scaled enrichment score (ES) from each pattern in Figure 2C onto pseudo-space paths inferred in Figure 2B, patterns 1, 2, and 3 exhibit a linear-like decrease in the pre-late stages but slightly different enrichment patterns in the late phase (Figure 2E). More specifically, pattern 1, which includes pathways of the UPR, reactive oxygen species (ROS), oxidative phosphorylation (Oxphos), and fatty acid metabolism (FA_METAB), decreases but then plateaus in late phase of maturation for paths 1, 2, and 3. Only path 4 decreases continuously into the late phase (Figure 2E). Pattern 2 includes the hallmark pathways hypoxia, UV response up, and p53 pathway, showing the same maturation trends but an increase in late states of path 2 corresponding to c7 and c8. Pattern 3 contains IL6-JAK-STAT3 signaling, inflammatory response, and apoptosis, revealing similar maturation trends in all four paths. Pattern 4, notably including TNFα signaling via NF-κB and KRAS signaling down, is most intriguing since the enrichment diverges between the late ASC in path 1 and 2. In addition to the IgG trajectories, we found high ES in IgM-pre-dominant lineage (c13–15) for TNFα signaling via NF-κB (Figure S5B) in c15 (Figure S5C).

TFs of ASC maturation and survival

To understand distinct enrichment patterns of gene sets regulated by specific TFs, we imputed the potential functions of TFs in regulating BM ASC maturation. We observed 205 differentially expressed TFs that are expressed in at least 10% of assigned cell clusters (Table S2). The most abundant 144 TFs were assigned to the cell cluster or defined stage of maturation (Figures 3A and S2C) with the highest gene expression. These "cluster-associated" TFs were observed in nine of 15 cell groups. Notably absent were any TFs defining the transitional and late-stage clusters c4–c7 or c14 and c15.

Adding protein-protein interaction information from the STRING database,[38] we observed that some TFs are cluster specific (early or late) (Table S2). Thus, 48 out of 144 (33%) of the TFs were c1 associated and typically known to be involved in regulating basic functions such as RNA/DNA binding, transcription, metabolic processes, RNA stabilization (YBX1, SUB1), high-mobility group protein B (HMGB) damage-associated molecular patterns (DAMPs),[39] (HMGB1, HMGB2, and HMGB3), and cell-cycle progression (MYBL2)[40]. Similarly, TFs important for immune activation, signaling, differentiation, and survival (EGR1, JUN, JUNB, FOS, NR4A1, and ZFP36L1) were in c8.

Two other large groups of 25 (∼17%) and 43 (∼30%) TFs were identified in c11 and 12, respectively. High mitochondrial gene expression in c11 suggests high stress and potentially eventual death. c12 was characterized by accumulation of ATF4, the first stable product that can trigger the activation of the PERK pathway, which promotes ASC apoptosis when ER stress is unabated.[41,42] Thus, this small cluster may indeed be a death-prone ASC.

Bifurcation of LLPC differentiation

To characterize the bifurcation of late IgG ASC maturation in path 1 (c5 vs. c6) and path 2 (c5 vs. c7), we used gene set enrichment analysis (GSEA) prerank analysis and identified six and 19 gene sets that were significantly activated or repressed (Table S3). Although IFN-γ response and allograft rejection were downregulated in c6, inflammation and MYC targets v1 were upregulated in c7, and only TNFα signaling via NF-κB was downregulated in

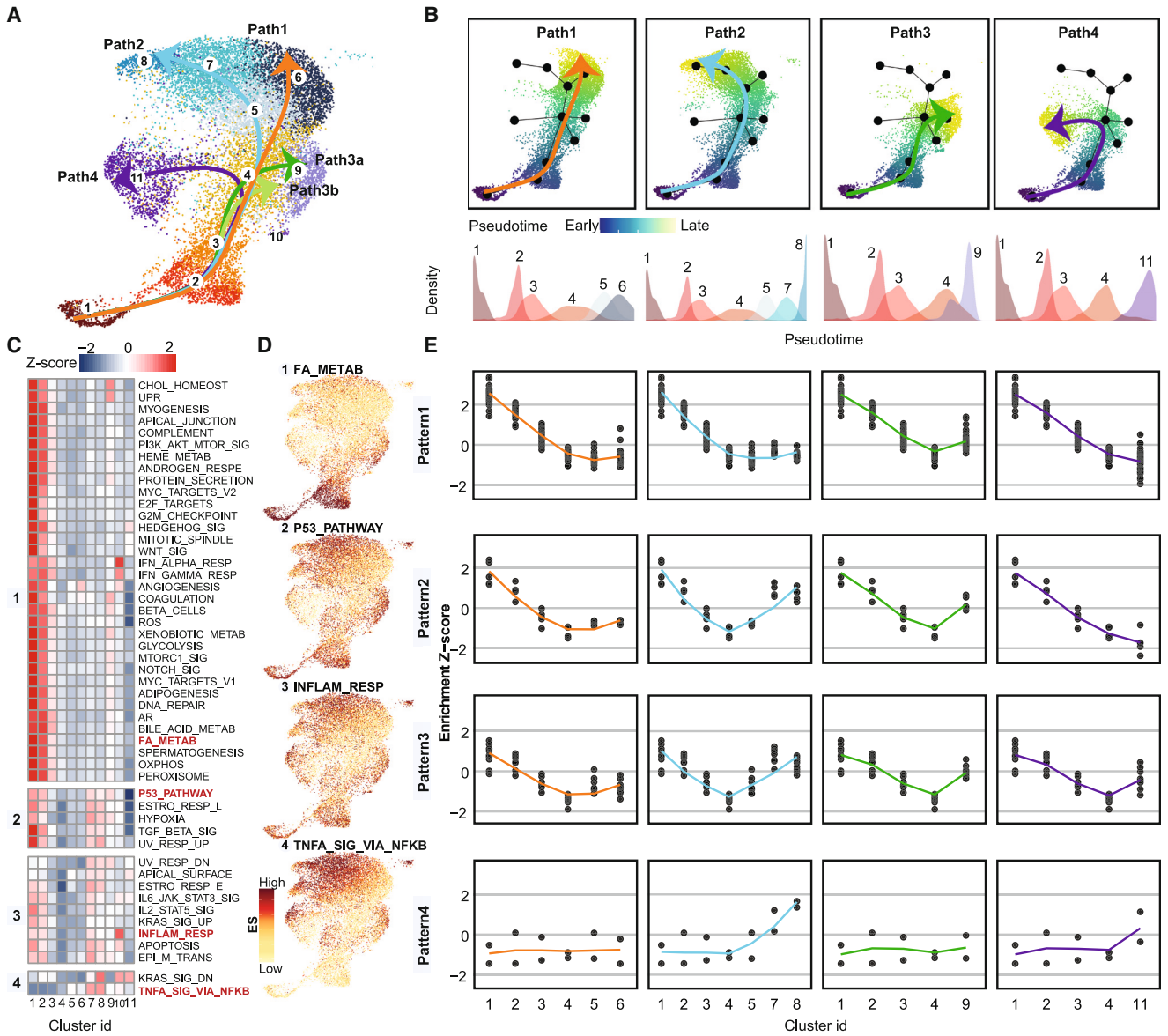


Figure 2. Trajectory and hallmark pathway enrichment analysis to distinguish predicted BMPC maturation paths

(A) UMAP plot shows predicted paths of IgG1-dominant BMPC maturation. Arrows indicate the maturation direction.

(B) UMAP plots show cell cluster located in each maturation path and colored by predicted pseudotime. (Top) Each solid black dot represents a cell population used to predict path in (A). The darker the blue and the lighter the green, the earlier and later stages of maturation, respectively. (Bottom) The density plots show the distribution of scRNA-seq-identified cluster on pseudotime space of each path.

(C) Heatmap shows the row-scaled enrichment scores (ESs) for hallmark pathway enrichment analysis in the IgG1-dominant cell populations; from left to right is c1–c11.

(D) Projected ES of indicated pathway example from each pattern onto UMAP. The darker the red, the higher the enrichment.

(E) Dot plot next to the example UMAP visualization in (D) shows the scaled ES from each pattern by cluster from each path. The x axis is ordered cell populations corresponding to the cell order in each path in (B). Loess method fitted lines of ES alteration trends were colored by predicted paths in (B).

c6 and upregulated in c7 (false discovery rate $\leq$0.001) (Figure 3B). After evaluating the pathway ES for each individual separately, we found that this difference was not driven by any one subject (Figure 3C).

By comparing DEGs between c5 vs. c6 and c5 vs. c7, the Venn diagram shows top DEGs that are up or down in c6 or c7 (Figure 3D; Table S4). The expression of corresponding genes is visualized in Figure 3E. Subsequently, we compared the TNFα signaling via NF-κB directly between c6 and c7 and discovered 78 genes contributing to the significant enrichment (Figures S6A and S6B; Table S5). Out of those 78 leading edge genes, 34 ($\sim$44%) are among the top DEGs, including *TNFAIP3*,

CellPress
OPEN ACCESS

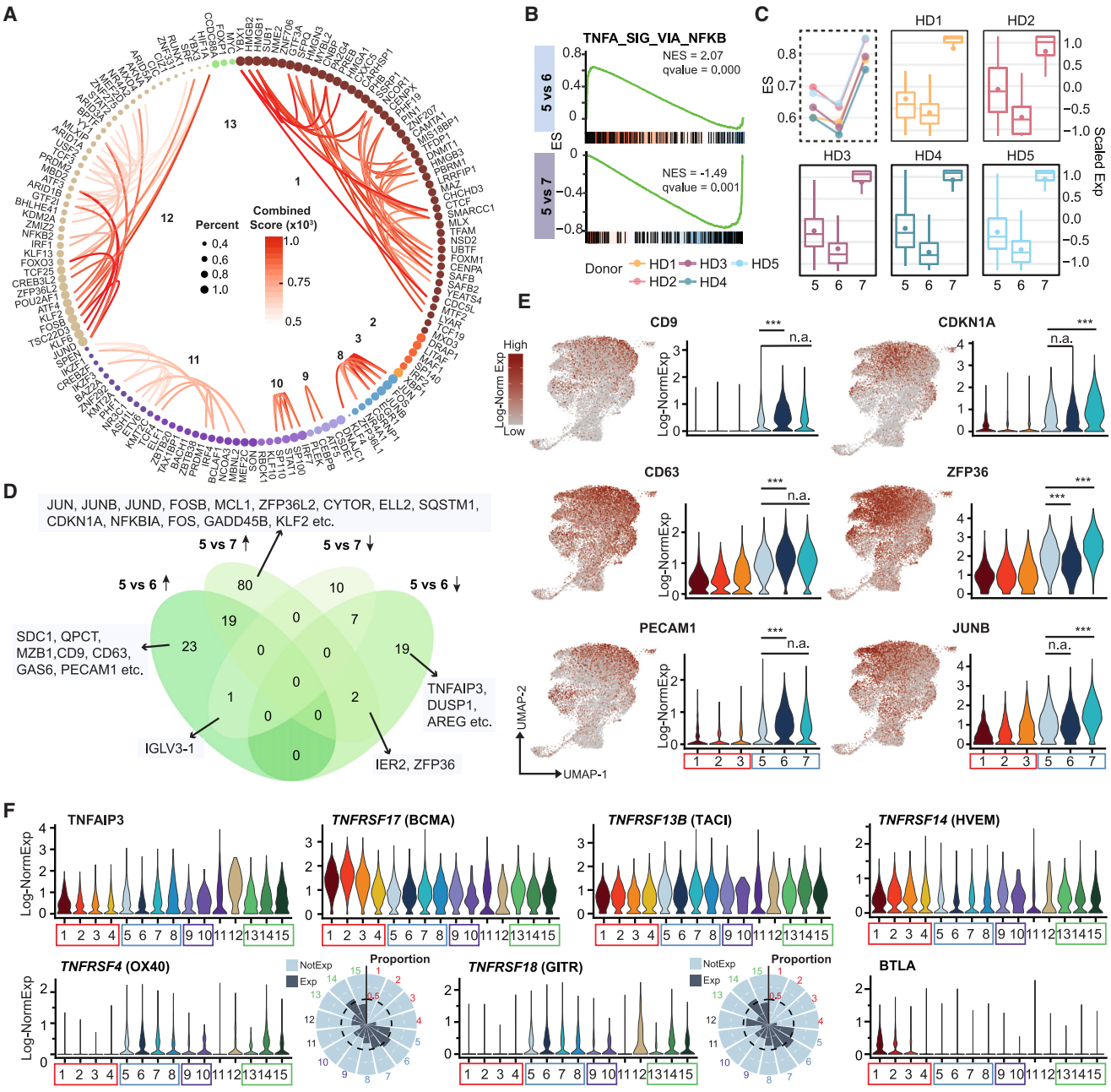


Figure 3. Genes and hallmark pathways distinguish late-phase maturation fate

(A) Cell-cluster-associated TFs; each node represents a TF, colored by associated cell cluster and scaled by the expressed percentage of cells in the cluster. The lines between nodes are inferred protein-protein interactions from the STRING database. The redder the line, the more confident the inferred interactions. Numbers in the circles show the assigned cell cluster ID.

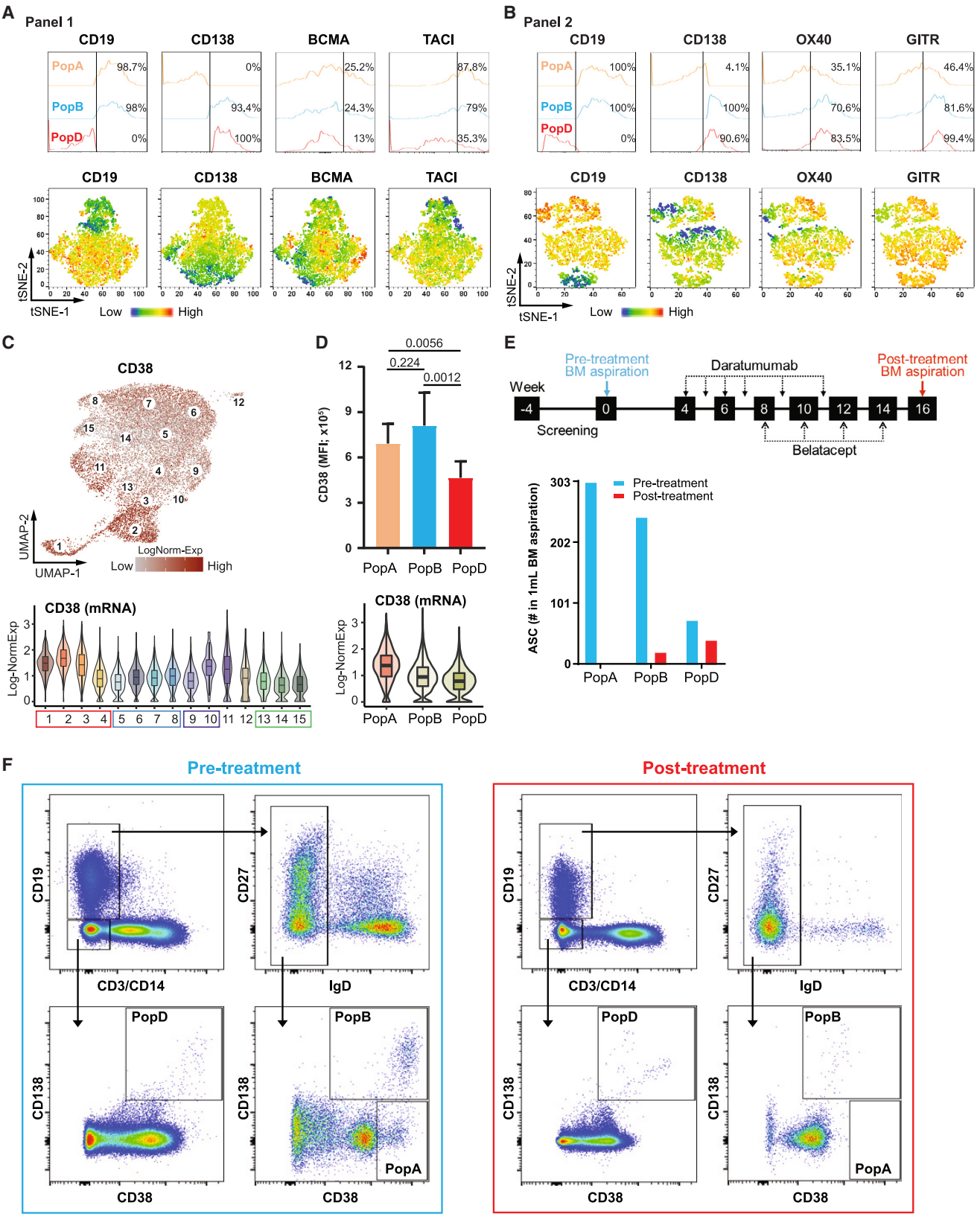
(B) GSEA of the most indicated pathways. TNFα signaling via NF-κB that are differentially enriched between 5 vs. 6 and 5 vs. 7.

(C) TNFα signaling via NF-κB hallmark pathway ES separated for each of the five individual subjects (in dashed box and y axis is on the left) and the distribution of scaled corresponding pathway enriched maturation-associated DEG expression in c5, c6, and c7 (in solid boxes and y axis is on the right).

(D) Comparison of top DEGs between c5 vs. c6 and c5 vs. c7 (see star methods), based on the sign of average log fold-change (avglogFC), the DEGs were divided into up-/downregulated in c6 or c7 groups. Venn diagram shows the results of the DEG comparison.

(E) Gene expression examples from the results of the comparisons in (D).

(F) Gene expression of genes from TNF family. The x axis is cell cluster ID and y axis is the log-normalized gene expression. Circular bar plot shows the proportion of cells in each cell subgroup showing expression of corresponding genes. Black dashed line indicates the proportion of 0.5 and numbers indicate the cell cluster IDs. NotExp, not expressed; Exp, expressed. ***Bonferroni-adjusted p value <0.001; n.a., not available.



(legend on next page)

FOS, *NFKBIA*, *CDKN1A*, and *ZFP36* (Figure S6C; Tables S4 and S5).

Differentially expressed TNFRSF family members between early and late ASC

TNF and TNF superfamily cytokine signaling play important roles in B cells and PC survival and function or cell death. The best known factors BAFF or APRIL (*TNFSF13*) and their recognized receptors BAFF receptor (BAFFR:*TNFRSF13C*), Transmembrane Activator and CAML-interactor (TACI:*TNFRSF13B*) and B cell maturation antigen (BCMA:*TNFRSF13A/17*) are essential for ASC survival.[43,44] APRIL binds strongly to BCMA and moderately to TACI, whereas BAFF binds weakly to BCMA and strongly to TACI.[45] Although BAFF has been suggested to be important in mouse BM ASCs,[46] APRIL is the critical cytokine for ASC survival and maturation.[7,47] Finally, APRIL and not BAFF can bind to heparin sulfate proteoglycans (HSPGs) such as CD138, which concentrates APRIL at the cell surface, thereby increasing its effectiveness.[48]

We found little difference in expression of the members of the TNFRSF family between the two late IgG paths1 and 2; however, there were major differences between the early and late clusters (Figures 3F, S7A, and S7B). BCMA showed high expression in early ASCs and was significantly downregulated in late ASC clusters, while TACI expression was significantly increased in late clusters, albeit to modest levels (Figures 3F and S7A; Table S6). Although TACI has been described in tonsil and BM ASCs,[49] we found high expression of TACI in c14, a late IgM-predominant ASC cluster (Table S6). In mice, TACI expression is highest on mature innate-like B cells such as marginal zone and B-1 B cells, which is critical for T-independent responses[50,51]; however TACI may also have a role in late BM ASCs. Finally, healthy BM ASCs do not express APRIL or BAFF, showing the need for exogenous sources of these survival cytokines.

Other receptors in the TNFRSF considered to play roles in T cell activation were differentially regulated in early vs. late BM ASC clusters (Figure 3F). For example, OX40 (*TNFRSF4*) and GITR (*TNFRSF18*) were both significantly upregulated in the late ASC clusters. In contrast, HVEM (*TNFRSF14*) expression was substantially increased in the early c2, 3, 4, 9, and 13 (Figure S7A; Table S6). In all, the sequential TNFRSF programs in early and late ASCs provide important insights into BM maturation.

Although protein expression may be concordant with gene expression,[52] BCMA and TACI surface expression may be more variable and less concordant with gene expression (Figure 4A). Based on our previous BM pop A, B, and D, both BCMA and TACI decrease surface protein expression with maturation in pop D. In contrast, as BM ASCs mature, flow phenotyping revealed higher surface expression of OX40 and GITR in pop D, validating these expression profiles (Figures 3F and 4B).

Although CD38 expression increases during B cell differentiation to ASC precursors,[53] its downregulation during ASC maturation was unanticipated; however, this feature may actually be consistent with enhanced survival in tumor cells.[54] Pop D had lower mean fluorescence intensity (MFI) of CD38 by flow cytometry, which was concordant with gene expression (Figures 4C and 4D). Concordantly, we observed a selective loss of early BM subsets with anti-CD38 therapy (daratumumab). In a sensitized patient with a broad array of HLA antibodies awaiting the second kidney transplant, only the late mature ASC subsets remained after daratumumab and belatacept for 14 weeks (under an immune tolerance network [ITN] protocol ITN090ST; Figure 4E). We found depletion of pop A and B with 100% and 92.2% reduction respectively, while pop D remained (Figures 4E and 4F). Of the 59 HLA-specific serum antibodies pre-treatment, only 16 remained after treatment. Thus, early BM ASCs are the most susceptible to anti-CD38 therapies as predicted from the single-cell analysis. Combined, these data may help understand the therapeutic targeting achieved by anti-CD38 agents currently used for the treatment of PC malignancies and for highly sensitized patients and autoimmune conditions[55–58] and finally provides an atlas and deep insights into the mechanistic implications of selective depletion of BM ASC subsets.

Mature BM ASCs downregulate pro-apoptotic genes and upregulate pro-survival genes

Previous bulk RNA sequencing (RNA-seq) and assay for transposase-accessible chromatin with sequencing (ATAC-seq) indicated that pop D significantly upregulates pro-survival genes *BCL2* and *MCL1*, despite enhanced chromatin accessibility being present only for *BCL2*[6] (Figure S7C). Here, we examined the single-cell expression of pro-survival, intrinsic and extrinsic pro-apoptotic genes, cell cycle, and cell-cycle arrest across the 15 clusters (Figure 5A). *TSC22D3*, *MCL1*, and *BCL2* are the essential pro-survival genes for BM ASC maturation. *TSC22D3* (glucocorticoid-induced leucine zipper [GILZ]) can inhibit the transcriptional activity of *FOXO3*, which leads to the further suppression of BIM-induced apoptosis, albeit in T cells.[59] Of the late clusters, *TSC22D3* expression is high in c5 and c7 and the highest in c12, which has the corresponding highest expression of BIM (*BCL2L11*). *BCL2* is elevated to a similar degree in all late-stage LLPCs (c5–8), whereas *MCL1* shows even higher expression late in path 2 (c7 and 8; Figures 5A and 5B). Conversely, both intrinsic and extrinsic pro-apoptosis genes are reduced in the LLPC (path 1 and 2). These trends are consistent with a molecular basis for refractoriness to apoptosis in the late ASC.[60]

Figure 4. Experimental validation of early- and late-phase representative markers

(A and B) Histograms (upper) and tSNE heat-maps (bottom) of expression of ASC surface markers (A) CD19, CD138, BCMA, and TACI (panel 1), or (B) CD19, CD138, OX40, and GITR (panel 2).

(C) CD38 log-normalized gene expression visualized by UMAP (top) and violin plot grouped by cluster ID (bottom).

(D) CD38 MFI measured from n = 9 healthy BM aspirates (top) and log-normalized gene expression visualized violin plot grouped by FACS-sorted cell labels (bottom).

(E) Study regimen and timeline of the donor-specific antibodies (DSAs) (ClinicalTrials.gov identifier: NCT04827979, top). Quantitation of BM ASC subsets (# ASC/mL BM) pre and post treatment with daratumumab and belatacept (bottom).

(F) Flow cytometry pre and post treatment.

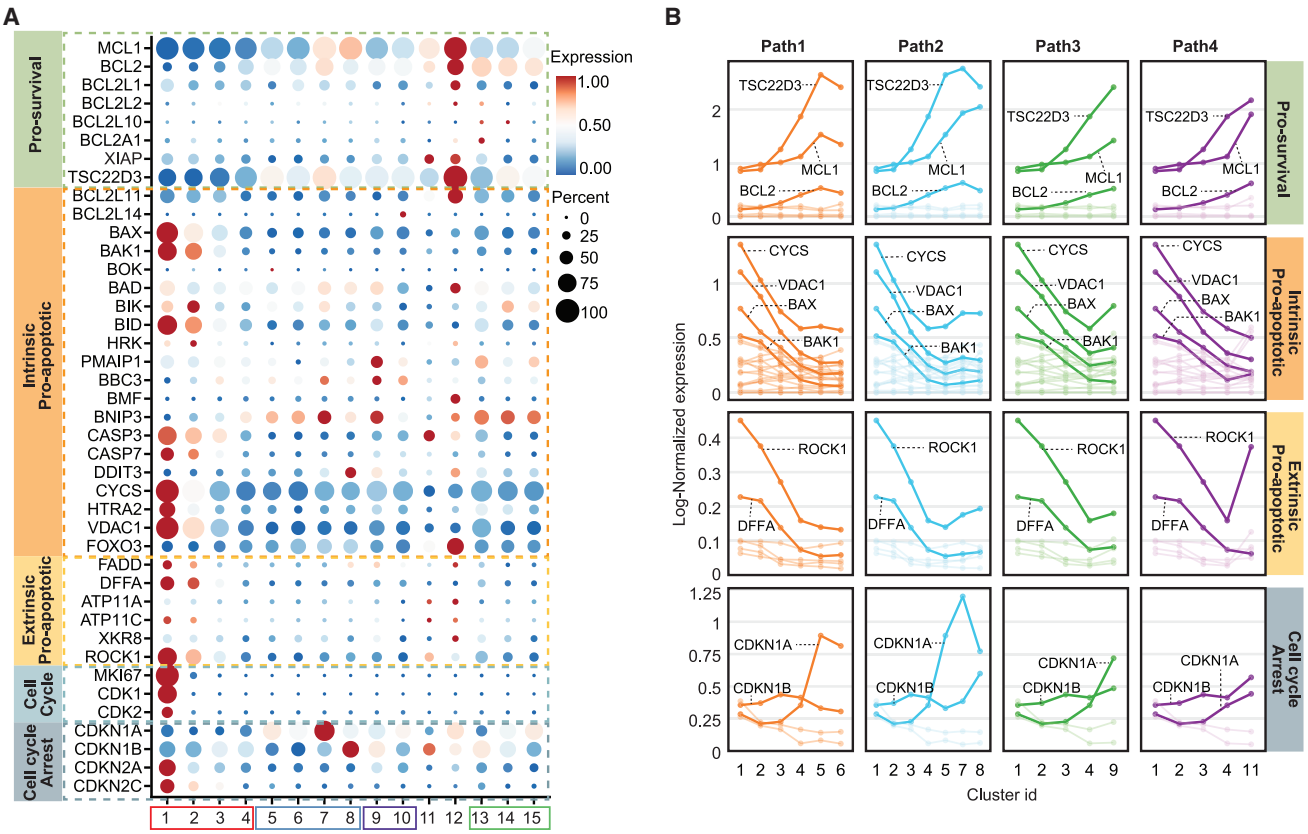


Figure 5. Exploration of apoptotic gene expressions in BMPC clusters

(A) Dot plot showing the expression of genes related to pro-survival, intrinsic pro-apoptotic, extrinsic pro-apoptotic, cell-cycle progression, and cell-cycle arrest functions in BMPC subgroups.

(B) Dynamic gene expression alterations in four maturation paths defined in Figure 2B with the same color coding and ordering. The labeled thick solid lines show the genes with high expressions and variations during the BMPC maturation.

Inability to mature from early to late ASC likely leads to cell death (Figures 5A and 5B). c1 (PB) and c2 stages of BM ASC maturation appeared to be primed for apoptosis with higher expression of pro-apoptotic genes, including the mitochondrial outer membrane permeabilization (MOMP) activators *BAX* and *BAK1* and mediator VDAC1, and released apoptogenic proteins from the intermembrane of mitochondria cytochrome *c* (CYCS) and a serine protease OMI (encoded by *HTRA2*) that neutralizes the caspase-inhibitory proteins to directly and indirectly involve caspase activation. The DNA fragmentation factor (DFFA) and apoptotic membrane blebbing gene *ROCK1* were also upregulated.[61–64] The PB stage also had a high abundance of transcripts important for proliferation, such as *MKI67*, *CDK1*, *CDK2*, as well as cell-cycle arrest-associated genes p16 (*CDKN2A*) and p18 (*CDKN2C*). Two other cell-cycle-inhibiting genes, p21 (*CDKN1A*) and p27 (*CDKN1B*), showed higher expression in c7 and c8, respectively.[65] Cell-cycle arrest is largely regulated by activation of either one or both of the p16/PrB and p21/p53 pathways, which may be differentially regulated in early and late ASCs respectively.

c6 showed high expression of *CD138* (SDC1), a member of the heparan sulfate proteoglycan family which has been reported to promote ASC survival by regulating *BCL2* and *MCL1*.[66] We also observed that *CD63*, *CD9*, and *IL5RA* were upregulated in the late-stage BM ASCs (Figure S7D). *CD9* and *CD63* have been reported to associate with the metastatic ability of tumor cells, such that higher expression promotes decreased cell motility.[67] Interleukin (IL) 6 is known to be important for BM ASC survival,[7,68] but, interestingly, IL6R expression was highest in the PB stage and c2, whereas IL6ST (binding of IL6 and IL6R) showed higher gene expression in c11 and 12 but not in the other late ASC clusters (Figure S7D). Hence, these results suggest that most BM ASCs likely utilize bound IL6 through the soluble IL6R.

Isotype characteristics of BM ASCs

Our single-cell analysis provided a thorough representation of the whole V(D)J repertoire of human BM ASCs. Indeed, of the 17,347 cells with single-cell transcriptomic information, 83% also had matched VDJ:V_H (by scVDJ-seq) sequences including isotype identification. The V_H repertoire included 11,853 clonal lineages, with a majority representing singletons (10,344 cells), while 1,509 lineages contained at least two cells (Figure S8A). Of these, 421 were observed in only one of the 15 clusters, whereas 1,088 lineages were present in at least two cell clusters, mostly in adjacent UMAP spatial clusters (Figure S8B). Each

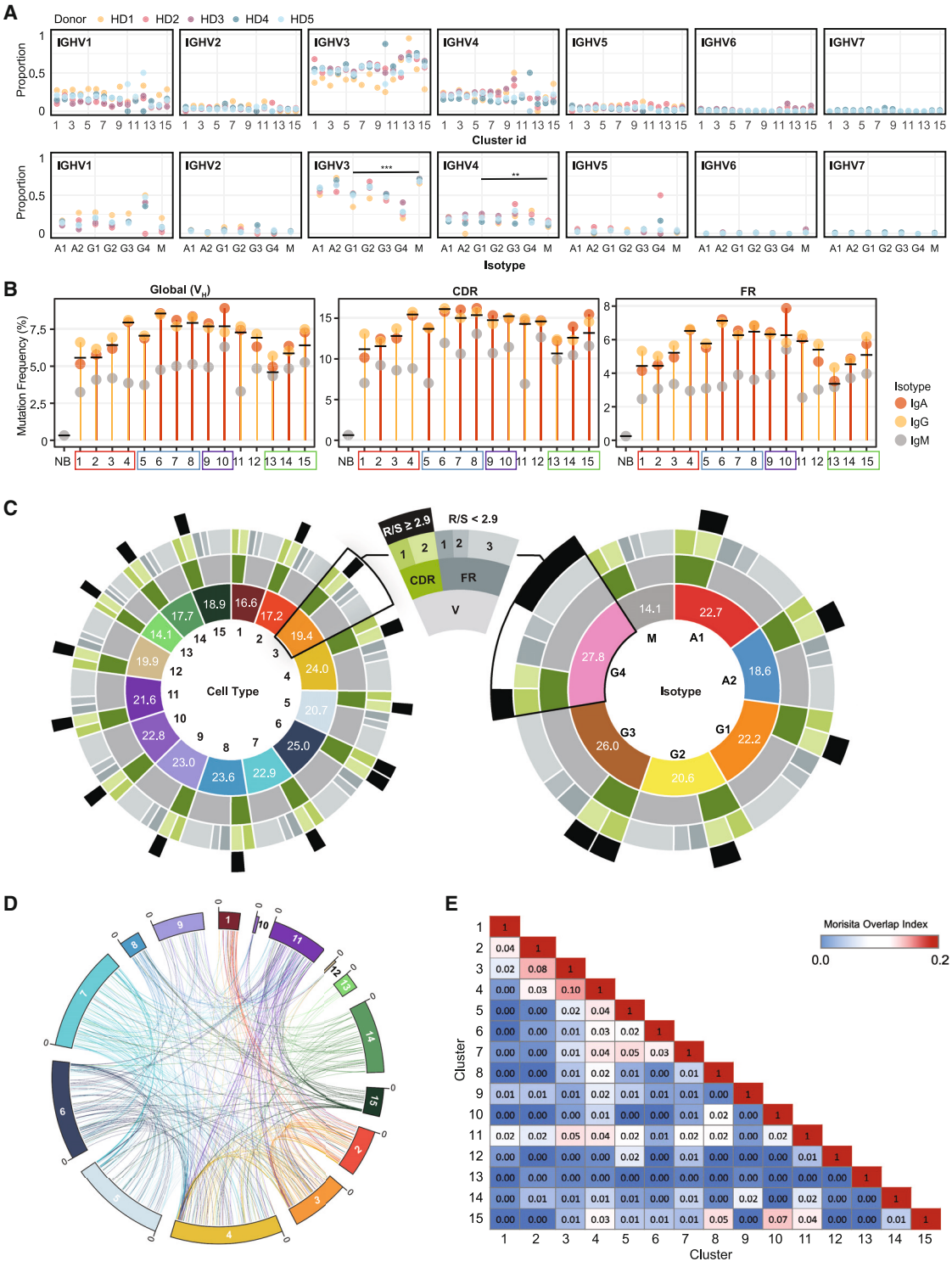


Figure 6. Mutation rate, similarity, and connectivity of clones measured by scVDJ-seq

(A) Summary of Ig heavy-chain family gene usage. The y axis shows the proportion of cells from each cluster (top) and Ig isotype and IgG subclass (bottom). Black bar shows the comparison objects and asterisks indicate the significance of statistical test (***$p < 0.001$, **$p < 0.01$).

(B) The average mutation frequency in the whole region V (global), CDR, and FR of IgA, IgG, and IgM isotypes from each cell population. The solid black bar in each cell cluster indicates the overall average mutation frequency. NB, naive B cell as control.

(legend continued on next page)

cluster contained a similar percentage of singletons except c12, which contained few cells and the highest proportion of non-matched VDJ cells (Figures S8C and S8D). Nearly every cell (98.6% ± 0.8%) had a consensus Ig gene with greater than 100 transcripts (Figure S9), and single ASCs in the transitional and late stage had a higher proportion of cells with large numbers of Ig transcripts. The polyclonal repertoire was not unexpected since BM ASCs are the combined result of a lifetime of antigen exposures reflecting the historical serum antibody record.

V(D)J repertoire characteristics of BM ASC clusters

V_H gene usage was not biased among ASC clusters, Ig isotypes, and individuals. Higher IGHV1, IGHV3, and IGHV4 family genes reflect the predicted distribution of these larger families (Figure 6A). Similarly, there was no consistent bias of individual VH genes within the VH families.

The frequency and distribution of somatic hypermutation (SHM) provides important information regarding the maturation and antigenic selection of ASCs and, by extension, may offer important insight into their differentiation from separate B cell sources. As expected, BM ASCs had higher global mutation frequencies relative to peripheral blood naive B (NB) cells (Figure 6B; NB = 0.33%, BM ASC = 6.89% ± 1.11%, p = $8.512e^{-13}$). The BM ASCs also displayed SHM frequencies comparable with peripheral memory B cells (7.1%).[69]

Overall, SHM frequencies ranged from 4.8% to 8.6%, with the lowest frequency observed in c13 and the highest in c6 (Figure 6B). IgG and IgA generally had significantly higher mutation frequencies than IgM isotypes, but there was not a difference between IgA and IgG isotypes or IgG subclasses within each identified cell cluster (Figure 6B; Table S7). Interestingly, the highest SHM mutation number was found in the small IgG4 fraction across clusters at 27.8 on average (Figure 6C). Early stage ASCs of the IgG lineage, especially those with MHC class II gene expression (c1–c3) as well as all IgM lineage clusters, had a lower mutation frequency than late-stage IgG ASCs. This result was evident across the global V region or within the framework region (FR) and complementarity-determining region (CDR), whether considered in isolation or combined (Figure 6B). There was a trend toward higher average mutation rates as BM ASCs mature, suggesting either a potential survival advantage of cells with higher mutation frequencies or, alternatively, preferential origination from late GC reactions. c6, which had the highest average number of mutations, also had the highest global V_H, CDR, and FR region mutation frequencies (permutation test, $p < 0.05$; Figures S10A–S10D).

A replacement-silent (R/S) mutation ratio greater than 2.9 is suggestive of antigen selection[70,71] and most of the clusters (14 out of 15) had a CDR2 R/S > 2.9. Only c6 showed an R/S > 2.9 in both CDR1 and CDR2, while c1 had no regions with R/S > 2.9 (Figure 6C). Overall, VDJ analysis showed that most BM ASCs are highly mutated and antigen selected.

Repertoire connectivity between early- and late-stage BM ASCs

Among the five individuals, we identified 11,853 V_H lineages, and all shared lineages were within a single individual, demonstrating no public clonotypes. Clonal lineages were defined as the same V, J, CDR3 length, and 85% CDR3 homology. Of the 11,853 V_H lineages identified, 1,088 (9.2%), were present in at least two ASC clusters, thereby indicating a significant level of cluster interrelationship (Figure S11B). However, to track the identical clone, we used 98% homology to follow the same clone in the pseudotrajectories. Since IgG1 ASCs are the dominant isotype in the BM, we show the connectivity of identical clones of IgG1 ASCs in c1–c15 in aggregate and by individual subject (Figures 6D and S11A).

We applied the Morisita overlap index to map lineage connectivity across populations. Cell populations were highly connected within the early (c1–c4) or the late stage (c5–c8) (Figures 6D and 6E). The degree of connectivity was substantially greater after removal of singletons, a strategy that enriches for larger clones and increases the sensitivity of detection (Figures S11B and S11C). Nearly identical clonal connectivity of IgG1 ASC in c1–c8 are shown (Figure S12).

Although some identical clones were detected in multiple cell clusters, one of the biggest, lineage #334, contained 11 cells distributed across early and late phases of IgG ASCs as well as c9, c10, c11, and c15 (Figure S13A). The VDJ sequences in this clone were identical, as shown in the alignment (Figure S13B). There were several clones with eight cells with identical sequences (385 and 280) that progressed through the early to late phases, validating potential paths for ASC maturation (Figures S13A and S13B). While these observations show representative clones, many other inter-cluster clones contained identical sequences as well. These observations are consistent with a dynamic BM ASC compartment composed of clones that likely undergo further maturation in the BM microniche.

DISCUSSION

ASCs represent a highly specialized effector compartment whose main physiological function is the production of protective antibodies, which, in some cases, may persist for the lifespan of the host, owing to the existence of LLPCs. While active infections trigger antibody production through newly made ASCs, either from naive or pre-existing memory B cells, long-standing antibodies represent a person's immune history and contribute serological memory capable of preventing infection by previous agents. In contrast to other immune effector cells that perform their function during acute activation and are short lived, LLPCs are unique in that they constitutively perform their

(C) Summary of the average mutation numbers by scRNA-seq-identified cell populations (left) and scVDJ-seq-identified isotypes (right). The innermost circle shows the cluster or isotype. Numbers in each section show the average number of mutations in region V. The second circle shows CDR and FR of region V. The third circle further breaks CDR and FR into CDR1, CDR2, FR1, FR2, and FR3. The outer-most circle black bar indicates the replacement-to-silence (R/S) ratio greater than 2.9.
(D) Compiled Circos plot connecting individual subjects' clones from cell c1 to c15. Lineages were colored by the latest cell cluster.
(E) Heatmap showing the average Morisita overlap index of five subjects, with 0 indicating no similarity and 1 indicating identical repertoires.

effector function while in a resting state and endowed with prolonged survival. In contrast, short-term ASCs, referred to as plasmablasts, perform the same effector function and may follow two different fates: post-proliferative apoptosis or subsequent maturation into LLPCs. It is well established that LLPCs persist in specialized survival microniches in the BM. However, it has remained uncertain whether they arrive in the BM as fully differentiated cells that only require admission to the protective niche or, instead, begin their BM journey from immature to mature PCs. In the latter scenario, it remains to be determined whether all new ASC arrivals to the BM represent potential LLPC precursors, or LLPCs derive only from specific ASC subsets. In either situation, critical outstanding questions include the specific phenotype of LLPCs among mature BM ASCs and the regulatory and survival trajectories of BM ASCs into LLPCs.

Our single-cell studies provide an in-depth analysis of human BM ASCs, thereby contributing substantial original insight into these questions. We provide a high-resolution atlas of the BM ASC, including the heterogeneity of the LLPC compartment. For example, this study offered mechanisms of known BM PC therapies (anti-CD38) together with discovery of novel therapeutic targets for mature BM plasma, such as GITR, OX40, CD9, CD63, ZBTB20, and IL5RA.

A finding of central significance is the identification of IgM BM ASC clusters with heterogeneous properties. While long-term IgM memory B cells are well established,[72] the ability to generate mature IgM LLPCs has not been explored despite descriptions in the mouse spleen demonstrating persistent immunity.[28] Combined, trajectory studies of clonal connectivity across these clusters show the presence of mature IgM BM PCs that might contribute long-lasting serum IgM antibodies. Future studies will be required to determine the precise contribution of extrafollicular (EF) and GC-driven pathways.[28,73–75]

Once they arrive in the BM, ASCs follow two major differentiation trajectories leading to mature PCs. These two paths offer differences of TNF signaling through NF-κB. AP-1 factors are also of particular importance among upregulated DEGs in path 2 (*JUN*, *JUNB*, *JUND*, *FOS*, *FOSB*, *MCL1*, *ZFP36L2*, *ELL2*, *CDKN1A*, *PRDM1*, *NFKBIA*, and others). Interestingly, JUNB is involved in proliferation and survival[76] in diseased PCs (multiple myeloma [MM]) but is abundantly expressed in healthy ASCs, where it likely plays an anti-apoptotic role. Similarly, in MM, c-JUN is important for caspase-mediated c-ABL cleavage inducing apoptosis but may function quite differently in healthy BM ASCs. Whether the NF-κB/AP-1/STAT3 inflammatory regulatory network described in the transformation of human cancers[77] also plays a role in LLPCs will require further study.

Trajectory analysis also points to an important role for cell-cycle regulation in late progression to terminally differentiated PCs through path 2. For example, p21, a marker capable of inhibiting a range of CDKs, is activated in a p53-dependent manner and was upregulated in c7 to mediate cell-cycle arrest.[78] In turn, the tumor repressor p16 has also been described as a hallmark of cellular senescence.[79–81] Exit from cell cycle irreversibly[82] is a feature of cellular senescence and potential mechanism of LLPC survival. Path 2 was also characterized by higher expression of ZFP36L1, an RNA-binding protein (RBP), facilitates newly formed ASCs homing to the BM microniche.[83] Although a role of RBP in ASC differentiation is unclear, these proteins may actually modulate ASC migration and survival in late ASCs.

Our study also provides insight into the metabolic and survival underpinnings of late-stage BM ASCs. Of particular interest for the latter property is the differential expression of receptors belonging to the TNFRSF family, specifically BCMA, which binds APRIL to enhance survival through MCL-1 in mice and to prolong survival in human LLPCs *in vitro* BM mimetic systems.[44,47] Interestingly, while BCMA was expressed across all BM ASCs, it decreased 2-fold in more mature ASCs, thereby pointing to an important role in securing survival in early maturation. In contrast, TACI transcription experienced a modest increase from early to late BM ASCs, a finding consistent with its heterogeneous expression in MM[84] and indicative of an unrecognized role in physiological LLPC maintenance. Although TACI is important for mouse ASC survival,[85] its actual role and degree of redundancy with BCMA remain to be understood in humans. Finally, cell-cell contact through OX40 and GITR, which are only expressed in late ASCs, intimates a model where LLPCs may have limited motility due to cell-cell contact, as suggested in recent mouse studies.[86]

In our highly sensitized patients treated with daratumumab and belatacept, daratumumab likely resulted in the specific loss of early BM ASC subsets; however, we cannot totally rule out the role of belatacept in inhibiting newly generated ASC immigrants into the BM site. Despite these caveats, belatacept likely inhibited only a small proportion of BM ASCs within a 16-week period. Clearly, additional studies with single agents will be needed.

Our study offers major advances to the field of LLPC biology. First, we optimized isolation and viability of rare human PCs.[7] Second, our experimental and analytical design achieved mRNA sequencing of sufficient depth for quantitative measurement of non-Ig genes, a goal readily compromised by the abundance of Ig transcripts representing up to 90% of some BM ASC total transcripts. Combined, our findings demonstrate substantial heterogeneity of mature BM ASCs and suggest that the LLPC potential may be present in several BM compartments. This possibility needs to be formally tested by localization of antigen-specific ASCs to the different clusters as previously done by our groups with the broader ASC populations. This goal could also be accomplished through the transcriptomes of identifying single ASCs of predetermined specificity.

Limitations of the study

Limitations of this study include the small number of healthy adult subjects and the total number of *bona fide* PCs used in the analysis. However, it is good to have a representation of all the known human BM ASC subsets since we captured the rare CD19-CD138+ cells. Another limitation is the nature of the single BM aspirate that offers a one-time snapshot, while repeat sampling would afford longitudinal temporal trajectories. While providing valuable transcriptional information, this study does not provide epigenetic and functional activity of metabolic pathways.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2023.112682.

ACKNOWLEDGMENTS

We also thank our Emory team of clinical coordinators and donors who made this study possible. We also thank Shuya Kyu, Robert E. Karaffa, and Kametha T. Fife of the Emory Flow Cytometry Core (EFCC) for technical support and Kitza Williams for coordinating ITN samples. This work was supported by NIH grants 1R01AI121252, U01AI141993, 1P01AI125180, U54CA260563, U19AI110483, and UM1AI109565, and the Bill and Melinda Gates Foundation.

AUTHOR CONTRIBUTIONS

Experimental design, F.E.L.; scRNA-seq and scVDJ-seq data analysis, interpretation, M.D., C.J.J., A.K., and E.G.; scVDJ-seq data supervision, C.M.T. and I.S.; human BM aspirates, S.L. and J.A.; experiment performance, D.C.N., C.S., C.J.J., C.K., I.H., S.K., and S.M.; ITN sample supervision, R.R., F.V., A.J., S.J.K., S.C., T.M., and C.B.; supervision, F.E.L., and G.G.; writing – review & editing, M.D., F.E.L., I.S., and G.G.

DECLARATION OF INTERESTS

F.E.L. is the founder of Micro-Bplex, Inc., serves on the scientific board of Be Biopharma, and is a recipient of grants from the BMGF and Genentech, Inc. I.S. has consulted for GSK, Pfizer, Kayverna, Johnson & Johnson, Celgene, Bristol Myer Squibb, and Visterra. F.E.L, D.C.N., and I.S. are inventors of the issued patents: 9/21/21 US 11,124766 B2 PCT/US2016/036650 and 9/21/21 US 11, 125757 B2 for the PC survival media.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: September 2, 2022
Revised: December 28, 2022
Accepted: June 6, 2023
Published: June 24, 2023

REFERENCES

1. Medina, F., Segundo, C., Campos-Caro, A., González-García, I., and Brieva, J.A. (2002). The heterogeneity shown by human plasma cells from tonsil, blood, and bone marrow reveals graded stages of increasing maturity, but local profiles of adhesion molecule expression. Blood *99*, 2154–2161.
2. González-García, I., Ocaña, E., Jiménez-Gómez, G., Campos-Caro, A., and Brieva, J.A. (2006). Immunization-induced perturbation of human blood plasma cell pool: progressive maturation, IL-6 responsiveness, and high PRDI-BF1/BLIMP1 expression are critical distinctions between antigen-specific and nonspecific plasma cells. J. Immunol. *176*, 4042–4050.
3. Halliley, J., Tipton, C., Liesveld, J., Rosenberg, A., Darce, J., Gregoretti, I., Popova, L., Kaminiski, D., Fucile, C., Albizua, I., et al. (2015). Long-lived plasma cells are contained within the CD19(-)CD38(hi)CD138(+) subset in human bone marrow. Immunity *43*, 132–145.
4. Garimalla, S., Nguyen, D.C., Halliley, J.L., Tipton, C., Rosenberg, A.F., Fucile, C.F., Saney, C.L., Kyu, S., Kaminski, D., Qian, Y., et al. (2019). Differential transcriptome and development of human peripheral plasma cell subsets. JCI Insight *4*, e126732.
5. Nguyen, D.C., Duan, M., Ali, M., Ley, A., Sanz, I., and Lee, F.E.H. (2021). Plasma cell survival: the intrinsic drivers, migratory signals, and extrinsic regulators. Immunol. Rev. *303*, 138–153.
6. Joyner, C.J., Ley, A.M., Nguyen, D.C., Ali, M., Corrado, A., Tipton, C., Scharer, C.D., Mi, T., Woodruff, M.C., Hom, J., et al. (2022). Generation of human long-lived plasma cells by developmentally regulated epigenetic imprinting. Life Sci. Alliance *5*, e202101285.
7. Nguyen, D.C., Garimalla, S., Xiao, H., Kyu, S., Albizua, I., Galipeau, J., Chiang, K.Y., Waller, E.K., Wu, R., Gibson, G., et al. (2018). Factors of the bone marrow microniche that support human plasma cell survival and immunoglobulin secretion. Nat. Commun. *9*, 3698.
8. Inui, M., Hirota, S., Hirano, K., Fujii, H., Sugahara-Tobinai, A., Ishii, T., Harigae, H., and Takai, T. (2015). Human CD43+ B cells are closely related not only to memory B cells phenotypically but also to plasmablasts developmentally in healthy individuals. Int. Immunol. *27*, 345–355.
9. Yoon, H.S., Scharer, C.D., Majumder, P., Davis, C.W., Butler, R., Zinzow-Kramer, W., Skountzou, I., Koutsonanos, D.G., Ahmed, R., and Boss, J.M. (2012). ZBTB32 is an early repressor of the CIITA and MHC class II gene expression during B cell differentiation to plasma cells. J. Immunol. *189*, 2393–2403.
10. Ting, J.P.Y., and Trowsdale, J. (2002). Genetic control of MHC class II expression. Cell *109*, S21–S33.
11. Shih, T.A.Y., Roederer, M., and Nussenzweig, M.C. (2002). Role of antigen receptor affinity in T cell-independent antibody responses in vivo. Nat. Immunol. *3*, 399–406.
12. Lamagna, C., Hu, Y., DeFranco, A.L., and Lowell, C.A. (2014). B cell-specific loss of Lyn kinase leads to autoimmunity. J. Immunol. *192*, 919–928.
13. Infantino, S., Jones, S.A., Walker, J.A., Maxwell, M.J., Light, A., O'Donnell, K., Tsantikos, E., Peperzak, V., Phesse, T., Ernst, M., et al. (2014). The tyrosine kinase Lyn limits the cytokine responsiveness of plasma cells to restrict their accumulation in mice. Sci. Signal. *7*, ra77.
14. Chernova, I., Jones, D.D., Wilmore, J.R., Bortnick, A., Yucel, M., Hershberg, U., and Allman, D. (2014). Lasting antibody responses are mediated by a combination of newly formed and established bone marrow plasma cells drawn from clonally distinct precursors. J. Immunol. *193*, 4971–4979.
15. Radbruch, A., Muehlinghaus, G., Luger, E.O., Inamine, A., Smith, K.G.C., Dörner, T., and Hiepe, F. (2006). Competence and competition: the challenge of becoming a long-lived plasma cell. Nat. Rev. Immunol. *6*, 741–750.

16. Yoon, S.-O., Zhang, X., Lee, I.Y., Spencer, N., Vo, P., and Choi, Y.S. (2013). CD9 is a novel marker for plasma cell precursors in human germinal centers. Biochem. Biophys. Res. Commun. *431*, 41–46.

17. Won, W.J., and Kearney, J.F. (2002). CD9 is a unique marker for marginal zone B cells, B1 cells, and plasma cells in mice. J. Immunol. *168*, 5605–5611.

18. Fridman, J.S., and Lowe, S.W. (2003). Control of apoptosis by p53. Oncogene *22*, 9030–9040.

19. Sperandio, S., Fortin, J., Sasik, R., Robitaille, L., Corbeil, J., and de Belle, I. (2009). The transcription factor Egr1 regulates the HIF-1alpha gene during hypoxia. Mol. Carcinog. *48*, 38–44.

20. Sperandio, S., Tardito, S., Surzycki, A., Latterich, M., and de Belle, I. (2009). TOE1 interacts with p53 to modulate its transactivation potential. FEBS Lett. *583*, 2165–2170.

21. Oh, Y.K., Jang, E., Paik, D.J., and Youn, J. (2015). Early growth response-1 plays a non-redundant role in the differentiation of B cells into plasma cells. Immune Netw. *15*, 161–166.

22. Tan, C., Mueller, J.L., Noviski, M., Huizar, J., Lau, D., Dubinin, A., Molofsky, A., Wilson, P.C., and Zikherman, J. (2019). Nur77 links chronic antigen stimulation to B cell tolerance by restricting the survival of self-reactive B cells in the periphery. J. Immunol. *202*, 2907–2923.

23. Pu, Z.Q., Yu, T.F., Liu, D., Jin, C.W., Sadiq, E., Qiao, X., Li, X., Chen, Y., Zhang, J., Tian, M., et al. (2021). NR4A1 enhances MKP7 expression to diminish JNK activation induced by ROS or ER-stress in pancreatic beta cells for surviving. Cell Death Dis. *7*, 133.

24. Luciano, F., Krajewska, M., Ortiz-Rubio, P., Krajewski, S., Zhai, D., Faustin, B., Bruey, J.M., Bailly-Maitre, B., Lichtenstein, A., Kolluri, S.K., et al. (2007). Nur77 converts phenotype of Bcl-B, an antiapoptotic protein expressed in plasma cells and myeloma. Blood *109*, 3849–3855.

25. Godoi, P.H.C., Wilkie-Grantham, R.P., Hishiki, A., Sano, R., Matsuzawa, Y., Yanagi, H., Munte, C.E., Chen, Y., Yao, Y., Marassi, F.M., et al. (2016). Orphan nuclear receptor NR4A1 binds a novel protein interaction site on anti-apoptotic B cell lymphoma gene 2 family proteins. J. Biol. Chem. *291*, 14072–14084.

26. Shapiro-Shelef, M., and Calame, K. (2005). Regulation of plasma-cell development. Nat. Rev. Immunol. *5*, 230–242.

27. Nutt, S.L., Hodgkin, P.D., Tarlinton, D.M., and Corcoran, L.M. (2015). The generation of antibody-secreting plasma cells. Nat. Rev. Immunol. *15*, 160–171.

28. Bohannon, C., Powers, R., Satyabhama, L., Cui, A., Tipton, C., Michaeli, M., Skountzou, I., Mittler, R.S., Kleinstein, S.H., Mehr, R., et al. (2016). Long-lived antigen-induced IgM plasma cells demonstrate somatic mutations and contribute to long-term protection. Nat. Commun. *7*, 11826.

29. Castro, C.D., and Flajnik, M.F. (2014). Putting J chain back on the map: how might its expression define plasma cell development? J. Immunol. *193*, 3248–3255.

30. Shirakawa, A.K., Nagakubo, D., Hieshima, K., Nakayama, T., Jin, Z., and Yoshie, O. (2008). 1,25-dihydroxyvitamin D3 induces CCR10 expression in terminally differentiating human B cells. J. Immunol. *180*, 2786–2795.

31. Patzelt, T., Keppler, S.J., Gorka, O., Thoene, S., Wartewig, T., Reth, M., Förster, I., Lang, R., Buchner, M., and Ruland, J. (2018). Foxp1 controls mature B cell survival and the development of follicular and B-1 B cells. Proc. Natl. Acad. Sci. USA *115*, 3120–3125.

32. Sagardoy, A., Martinez-Ferrandis, J.I., Roa, S., Bunting, K.L., Aznar, M.A., Elemento, O., Shaknovich, R., Fontán, L., Fresquet, V., Perez-Roger, I., et al. (2013). Downregulation of FOXP1 is required during germinal center B-cell function. Blood *121*, 4311–4320.

33. Toellner, K.M. (2015). FOXP1 inhibits plasma cell differentiation. Blood *126*, 2076–2077.

34. Tellier, J., Shi, W., Minnich, M., Liao, Y., Crawford, S., Smyth, G.K., Kallies, A., Busslinger, M., and Nutt, S.L. (2016). Blimp-1 controls plasma cell function through the regulation of immunoglobulin secretion and the unfolded protein response. Nat. Immunol. *17*, 323–330.

35. Kallies, A., Hasbold, J., Tarlinton, D.M., Dietrich, W., Corcoran, L.M., Hodgkin, P.D., and Nutt, S.L. (2004). Plasma cell ontogeny defined by quantitative changes in blimp-1 expression. J. Exp. Med. *200*, 967–977.

36. Chevrier, S., Emslie, D., Shi, W., Kratina, T., Wellard, C., Karnowski, A., Erikci, E., Smyth, G.K., Chowdhury, K., Tarlinton, D., and Corcoran, L.M. (2014). The BTB-ZF transcription factor Zbtb20 is driven by Irf4 to promote plasma cell differentiation and longevity. J. Exp. Med. *211*, 827–840.

37. Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. *1*, 417–425.

38. Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. *47*, D607–D613.

39. Boada-Romero, E., Martinez, J., Heckmann, B.L., and Green, D.R. (2020). The clearance of dead cells by efferocytosis. Nat. Rev. Mol. Cell Biol. *21*, 398–414.

40. Tarasov, K.V., Tarasova, Y.S., Tam, W.L., Riordon, D.R., Elliott, S.T., Kania, G., Li, J., Yamanaka, S., Crider, D.G., Testa, G., et al. (2008). B-MYB is essential for normal cell cycle progression and chromosomal stability of embryonic stem cells. PLoS One *3*, e2478.

41. Sano, R., and Reed, J.C. (2013). ER stress-induced cell death mechanisms. Biochim. Biophys. Acta *1833*, 3460–3470.

42. Zhu, H., Bhatt, B., Sivaprakasam, S., Cai, Y., Liu, S., Kodeboyina, S.K., Patel, N., Savage, N.M., Sharma, A., Kaufman, R.J., et al. (2019). Ufbp1 promotes plasma cell development and ER expansion by modulating distinct branches of UPR. Nat. Commun. *10*, 1084.

43. O'Connor, B.P., Raman, V.S., Erickson, L.D., Cook, W.J., Weaver, L.K., Ahonen, C., Lin, L.L., Mantchev, G.T., Bram, R.J., and Noelle, R.J. (2004). BCMA is essential for the survival of long-lived bone marrow plasma cells. J. Exp. Med. *199*, 91–98.

44. Peperzak, V., Vikström, I., Walker, J., Glaser, S.P., LePage, M., Coquery, C.M., Erickson, L.D., Fairfax, K., Mackay, F., Strasser, A., et al. (2013). Mcl-1 is essential for the survival of plasma cells. Nat. Immunol. *14*, 290–297.

45. Day, E.S., Cachero, T.G., Qian, F., Sun, Y., Wen, D., Pelletier, M., Hsu, Y.M., and Whitty, A. (2005). Selectivity of BAFF/BLyS and APRIL for binding to the TNF family receptors BAFFR/BR3 and BCMA. Biochemistry *44*, 1919–1931.

46. Benson, M.J., Dillon, S.R., Castigli, E., Geha, R.S., Xu, S., Lam, K.P., and Noelle, R.J. (2008). Cutting edge: the dependence of plasma cells and independence of memory B cells on BAFF and APRIL. J. Immunol. *180*, 3655–3659.

47. C. J. Joyner, A.M. Ley, D.C. Nguyen, M. Ali, A. Corrado, C. Tipton, C.D. Scharer, T. Mi, M.C. Woodruff, J. Hom et al., Maturation of human long-lived plasma cells results in resistance to apoptosis by transcriptional and epigenetic regulation.Preprint at bioRxiv, https://doi.org/10.1101/2021.05.22.4452692021.2005.2022.445269 (2021).

48. Ingold, K., Zumsteg, A., Tardivel, A., Huard, B., Steiner, Q.G., Cachero, T.G., Qiang, F., Gorelik, L., Kalled, S.L., Acha-Orbea, H., et al. (2005). Identification of proteoglycans as the APRIL-specific binding partners. J. Exp. Med. *201*, 1375–1383.

49. Darce, J.R., Arendt, B.K., Wu, X., and Jelinek, D.F. (2007). Regulated expression of BAFF-binding receptors during human B cell differentiation. J. Immunol. *179*, 7276–7286.

50. von Bülow, G.U., Russell, H., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., and Bram, R.J. (2000). Molecular cloning and functional characterization of murine transmembrane activator and CAML interactor (TACI) with chromosomal localization in human and mouse. Mamm. Genome *11*, 628–632.

51. Ozcan, E., Garibyan, L., Lee, J.J.Y., Bram, R.J., Lam, K.P., and Geha, R.S. (2009). Transmembrane activator, calcium modulator, and cyclophilin ligand interactor drives plasma cell differentiation in LPS-activated B cells. J. Allergy Clin. Immunol. *123*, 1277–1286.e5.

52. Liu, Y., Beyer, A., and Aebersold, R. (2016). On the dependency of cellular protein levels on mRNA abundance. Cell *165*, 535–550.

53. Arce, S., Luger, E., Muehlinghaus, G., Cassese, G., Hauser, A., Horst, A., Lehnert, K., Odendahl, M., Hönemann, D., Heller, K.D., et al. (2004). CD38 low IgG-secreting cells are precursors of various CD38 high-expressing plasma cell populations. J. Leukoc. Biol. *75*, 1022–1028.

54. Chmielewski, J.P., Bowlby, S.C., Wheeler, F.B., Shi, L., Sui, G., Davis, A.L., Howard, T.D., D'Agostino, R.B., Jr., Miller, L.D., Sirintrapun, S.J., et al. (2018). CD38 inhibits prostate cancer metabolism and proliferation by reducing cellular NAD(+) pools. Mol. Cancer Res. *16*, 1687–1700.

55. Cole, S., Walsh, A., Yin, X., Wechalekar, M.D., Smith, M.D., Proudman, S.M., Veale, D.J., Fearon, U., Pitzalis, C., Humby, F., et al. (2018). Integrative analysis reveals CD38 as a therapeutic target for plasma cell-rich pre-disease and established rheumatoid arthritis and systemic lupus erythematosus. Arthritis Res. Ther. *20*, 85.

56. Humbel, M., Bellanger, F., Fluder, N., Horisberger, A., Suffiotti, M., Fenwick, C., Ribi, C., and Comte, D. (2021). Restoration of NK cell cytotoxic function with elotuzumab and Daratumumab promotes elimination of circulating plasma cells in patients with SLE. Front. Immunol. *12*, 645478.

57. Facon, T., Kumar, S., Plesner, T., Orlowski, R.Z., Moreau, P., Bahlis, N., Basu, S., Nahi, H., Hulin, C., Quach, H., et al. (2019). Daratumumab plus lenalidomide and dexamethasone for untreated myeloma. N. Engl. J. Med. *380*, 2104–2115.

58. Ostendorf, L., Burns, M., Durek, P., Heinz, G.A., Heinrich, F., Garantziotis, P., Enghard, P., Richter, U., Biesen, R., Schneider, U., et al. (2020). Targeting CD38 with Daratumumab in refractory systemic lupus erythematosus. N. Engl. J. Med. *383*, 1149–1155.

59. Asselin-Labat, M.L., David, M., Biola-Vidamment, A., Lecoeuche, D., Zennaro, M.C., Bertoglio, J., and Pallardy, M. (2004). GILZ, a new target for the transcription factor FoxO3, protects T lymphocytes from interleukin-2 withdrawal-induced apoptosis. Blood *104*, 215–223.

60. Sarosiek, K.A., Fraser, C., Muthalagu, N., Bhola, P.D., Chang, W., McBrayer, S.K., Cantlon, A., Fisch, S., Golomb-Mello, G., Ryan, J.A., et al. (2017). Developmental regulation of mitochondrial apoptosis by c-myc governs age- and tissue-specific sensitivity to cancer therapeutics. Cancer Cell *31*, 142–156.

61. Madden, S.D., Donovan, M., and Cotter, T.G. (2007). Key apoptosis regulating proteins are down-regulated during postnatal tissue development. Int. J. Dev. Biol. *51*, 415–423.

62. Baines, C.P., Kaiser, R.A., Sheiko, T., Craigen, W.J., and Molkentin, J.D. (2007). Voltage-dependent anion channels are dispensable for mitochondrial-dependent cell death. Nat. Cell Biol. *9*, 550–555.

63. Sebbagh, M., Renvoizé, C., Hamelin, J., Riché, N., Bertoglio, J., and Bréard, J. (2001). Caspase-3-mediated cleavage of ROCK I induces MLC phosphorylation and apoptotic membrane blebbing. Nat. Cell Biol. *3*, 346–352.

64. Nagata, S. (2005). DNA degradation in development and programmed cell death. Annu. Rev. Immunol. *23*, 853–875.

65. Coller, H.A., Sang, L., and Roberts, J.M. (2006). A new description of cellular quiescence. PLoS Biol. *4*, e83.

66. McCarron, M.J., Park, P.W., and Fooksman, D.R. (2017). CD138 mediates selection of mature plasma cells by regulating their survival. Blood *129*, 2749–2759.

67. Barrena, S., Almeida, J., Yunta, M., López, A., Fernández-Mosteirín, N., Giralt, M., Romero, M., Perdiguer, L., Delgado, M., Orfao, A., and Lazo, P.A. (2005). Aberrant expression of tetraspanin molecules in B-cell chronic lymphoproliferative disorders and its correlation with normal B-cell maturation. Leukemia *19*, 1376–1383.

68. Kimberley, F.C., Hahne, M., and Medema, J.P. (2009). APRIL hath put a spring of youth in everything": relevance of APRIL for survival. J. Cell. Physiol. *218*, 1–8.

69. Tipton, C.M., Fucile, C.F., Darce, J., Chida, A., Ichikawa, T., Gregoretti, I., Schieferl, S., Hom, J., Jenks, S., Feldman, R.J., et al. (2015). Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus. Nat. Immunol. *16*, 755–765.

70. Jukes, T.H., and King, J.L. (1979). Evolutionary nucleotide replacements in DNA. Nature *281*, 605–606.

71. Shlomchik, M.J., Marshak-Rothstein, A., Wolfowicz, C.B., Rothstein, T.L., and Weigert, M.G. (1987). The role of clonal selection and somatic mutation in autoimmunity. Nature *328*, 805–811.

72. Palm, A.K.E., and Henry, C. (2019). Remembrance of things past: long-term B cell memory after infection and vaccination. Front. Immunol. *10*, 1787.

73. Racine, R., McLaughlin, M., Jones, D.D., Wittmer, S.T., MacNamara, K.C., Woodland, D.L., and Winslow, G.M. (2011). IgM production by bone marrow plasmablasts contributes to long-term protection against intracellular bacterial infection. J. Immunol. *186*, 1011–1021.

74. Bortnick, A., and Allman, D. (2013). What is and what should always have been: long-lived plasma cells induced by T cell-independent antigens. J. Immunol. *190*, 5913–5918.

75. Paus, D., Phan, T.G., Chan, T.D., Gardam, S., Basten, A., and Brink, R. (2006). Antigen recognition strength regulates the choice between extrafollicular plasma cell and germinal center B cell differentiation. J. Exp. Med. *203*, 1081–1091.

76. Fan, F., Bashari, M.H., Morelli, E., Tonon, G., Malvestiti, S., Vallet, S., Jarahian, M., Seckinger, A., Hose, D., Bakiri, L., et al. (2017). The AP-1 transcription factor JunB is essential for multiple myeloma cell proliferation and drug resistance in the bone marrow microenvironment. Leukemia *31*, 1570–1581.

77. Ji, Z., He, L., Regev, A., and Struhl, K. (2019). Inflammatory regulatory network mediated by the joint action of NF-kB, STAT3, and AP-1 factors is involved in many human cancers. Proc. Natl. Acad. Sci. USA *116*, 9453–9462.

78. Jung, Y.S., Qian, Y., and Chen, X. (2010). Examination of the expanding pathways for the regulation of p21 expression and activity. Cell. Signal. *22*, 1003–1012.

79. Hernandez-Segura, A., Nehme, J., and Demaria, M. (2018). Hallmarks of cellular senescence. Trends Cell Biol. *28*, 436–453.

80. Baker, D.J., Wijshake, T., Tchkonia, T., LeBrasseur, N.K., Childs, B.G., van de Sluis, B., Kirkland, J.L., and van Deursen, J.M. (2011). Clearance of p16Ink4a-positive senescent cells delays ageing-associated disorders. Nature *479*, 232–236.

81. Burd, C.E., Sorrentino, J.A., Clark, K.S., Darr, D.B., Krishnamurthy, J., Deal, A.M., Bardeesy, N., Castrillon, D.H., Beach, D.H., and Sharpless, N.E. (2013). Monitoring tumorigenesis and senescence in vivo with a p16(INK4a)-luciferase model. Cell *152*, 340–351.

82. Hinze, C., and Boucrot, E. (2018). Endocytosis in proliferating, quiescent and terminally differentiated cells. J. Cell Sci. *131*, jcs216804.

83. Saveliev, A., Bell, S.E., and Turner, M. (2021). Efficient homing of antibody-secreting cells to the bone marrow requires RNA-binding protein ZFP36L1. J. Exp. Med. *218*, e20200504.

84. Novak, A.J., Darce, J.R., Arendt, B.K., Harder, B., Henderson, K., Kindsvogel, W., Gross, J.A., Greipp, P.R., and Jelinek, D.F. (2004). Expression of BCMA, TACI, and BAFF-R in multiple myeloma: a mechanism for growth and survival. Blood *103*, 689–694.

85. Ou, X., Xu, S., and Lam, K.P. (2012). Deficiency in TNFRSF13B (TACI) expands T-follicular helper and germinal center B cells via increased ICOS-ligand expression but impairs plasma cell survival. Proc. Natl. Acad. Sci. USA *109*, 15401–15406.

86. Benet, Z., Jing, Z., and Fooksman, D.R. (2021). Plasma cell dynamics in the bone marrow niche. Cell Rep. *34*, 108733.

87. Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. *36*, 411–420.

88. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. Cell *177*, 1888–1902.e21.

89. Street, K., Risso, D., Fletcher, R.B., Das, D., Ngai, J., Yosef, N., Purdom, E., and Dudoit, S. (2018). Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genom. *19*, 477.

90. Finak, G., McDavid, A., Yajima, M., Deng, J., Gersuk, V., Shalek, A.K., Slichter, C.K., Miller, H.W., McElrath, M.J., Prlic, M., et al. (2015). MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. Genome Biol. *16*, 278.

91. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550.

92. Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell,, E.W. (2019). Dimensionality Reduction for Visualizing Single-Cell Data Using UMAP. Nat Biotechnol.

93. Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. *20*, 296.

94. Xu, C., and Su, Z. (2015). Identification of cell types from single-cell transcriptomes using a novel clustering method. Bioinformatics *31*, 1974–1980.

95. Rand, W.M. (1971). Objective criteria for the evaluation of clustering methods. J. Am. Stat. Assoc. *66*, 846–850.

96. Xu, C., and Su, Z. (2015). Identification of cell types from single-cell transcriptomes using a novel clustering method. Bioinformatics *31*, 1974–1980.

97. Soneson, C., and Robinson, M.D. (2018). Bias, robustness and scalability in single-cell differential expression analysis. Nat. Methods *15*, 255–261.

98. Wittkowski, K.M., Lee, E., Nussbaum, R., Chamian, F.N., and Krueger, J.G. (2004). Combining several ordinal measures in clinical studies. Stat. Med. *23*, 1579–1592.

99. Chaussabel, D., Quinn, C., Shen, J., Patel, P., Glaser, C., Baldwin, N., Stichweh, D., Blankenship, D., Li, L., Munagala, I., et al. (2008). A modular analysis framework for blood genomics studies: application to systemic lupus erythematosus. Immunity *29*, 150–164.

100. Saelens, W., Cannoodt, R., Todorov, H., and Saeys, Y. (2019). A comparison of single-cell trajectory inference methods. Nat. Biotechnol. *37*, 547–554.

101. Zhang, H.M., Chen, H., Liu, W., Liu, H., Gong, J., Wang, H., and Guo, A.Y. (2012). AnimalTFDB: a comprehensive animal transcription factor database. Nucleic Acids Res. *40*, D144–D149.

102. Lambert, S.A., Jolma, A., Campitelli, L.F., Das, P.K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T.R., and Weirauch, M.T. (2018). The human transcription factors. Cell *175*, 598–599.

103. Alamyar, E., Duroux, P., Lefranc, M.P., and Giudicelli, V. (2012). IMGT((R)) tools for the nucleotide analysis of immunoglobulin (IG) and T cell receptor (TR) V-(D)-J repertoires, polymorphisms, and IG mutations: IMGT/V-QUEST and IMGT/HighV-QUEST for NGS. Methods Mol. Biol. *882*, 569–604.

104. Ghraichy, M., Galson, J.D., Kovaltsuk, A., von Niederhäusern, V., Pachlopnik Schmid, J., Recher, M., Jauch, A.J., Miho, E., Kelly, D.F., Deane, C.M., and Trück, J. (2020). Maturation of the human immunoglobulin heavy chain repertoire with age. Front. Immunol. *11*, 1734.

105. Gupta, N.T., Vander Heiden, J.A., Uduman, M., Gadala-Maria, D., Yaari, G., and Kleinstein, S.H. (2015). Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. Bioinformatics *31*, 3356–3358.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
IgD–FITC	BD Biosciences	Cat. No. 555778
CD3-BV711	BioLegend	Cat. No. 317328
CD14-BV711	BioLegend	Cat. No. 301838
CD19-PE-Cy7	BD Biosciences	Cat. No. 560911
CD38-V450	BD Biosciences	Cat. No. 561378
CD138-APC	Miltenyi Biotech	Cat. No. 130-117-395
CD27-APC-e780	eBiosciences	Cat. No. 5016160
LiveDead	Invitrogen	Cat. No. L34966
IgD-Brilliant Violet 480	BD Biosciences	Cat. No. 566138
CD3-BUV737	BD Biosciences	Cat. No. 612750
CD14-BUV737	BD Biosciences	Cat. No. 612763
CD19-Spark NIR 685	BioLegend	Cat. No. 302270
CD38-Brilliant Violet 785	BioLegend	Cat. No. 303530
CD138-APC-R700	BD Biosciences	Cat. No. 566050
CD27-Brilliant Violet 711	BioLegend	Cat. No. 356430
BCMA-Brilliant Violet 421	BioLegend	Cat. No. 357520
TACI-PE-Cy7	BioLegend	Cat. No. 311908
OX40-Brilliant Violet 510	BioLegend	Cat. No. 350026
GITR-Brilliant Violet 605	BD Biosciences	Cat. No. 747664
Zombie NIR Fixable Viability Kit	BioLegend	Cat. No. 423106
IgD-BB700	BD Biosciences	Cat. No. 566538
CD3-PE-Cy5	ThermoFisher	Cat. No. 2363822
CD14-PE-Cy5	ThermoFisher	Cat. No. 2319032
CD19-BUV395	BD Biosciences	Cat. No. 740287
CD38-FITC	Cytognos	Cat. No. CYT-38F2-A
CD138-PE-Cy7	BioLegend	Cat. No. 356514
CD27-BV605	BD Biosciences	Cat. No. 562655
Live/Dead	ThermoFisher	Cat. No. L34962
Biological Samples		
Bone marrow aspirate samples from a total of 11 healthy donors and 1 patient with high donor-specific antibodies awaiting kidney transplant.	Emory University and the ITN protocol (ClinicalTrials.gov Identifier: NCT04827979).	This paper.
Critical Commercial Assays		
EasySep cell isolation kit	StemCell	Custom-designed
Zombie NIR Fixable Viability Kit	BioLegend	Cat. No. 423106
Chromium Single Cell Human BCR Amplification kit	10X Genomics	Cat. No. PN-1000253
Chromium Next GEM Single Cell 5′ Reagent Kit, v1	10X Genomics	Cat. No. PN-1000165
Deposited Data		
scRNA-seq, scVDJ-seq raw and processed data	This paper	GSE230705
Original code	This paper	https://doi.org/10.5281/zenodo.7903579
Software and Algorithms		
Cellranger 3.0.1	10X Genomics	https://support.10xgenomics.com/single-cell-gene-expression/software/downloads/3.0/
Seurat 3.2.2	Butler et al.[87]	http://satijalab.org/seurat

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
SCTransform	Stuart et al.[88]	https://github.com/satijalab/sctransform
Slingshot	Street et al[89]	https://bioconductor.org/packages/release/bioc/html/slingshot.html
MAST	Finak et al.[90]	https://github.com/RGLab/MAST
GSEA v4.0.3	Subramanian et al.[91]	https://www.gsea-msigdb.org/gsea/index.jsp
UMAP	Becht et al.[92]	https://github.com/lmcinnes/umap
Molecular Signatures Database	Liberzon et al.[37]	http://software.broadinstitute.org/gsea/msigdb/index.jsp
STRING protein-protein interaction	Szklarczyk et al.[38]	https://string-db.org/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, F. Eun-Hyung Lee (f.e.lee@emory.edu).

Materials availability

This study did not generate new reagents. Commercially available reagents are listed in the key resources table.

Data and code availability

The raw and processed scRNA-seq and scVDJ-seq data generated during this study are publicly available at the Gene Expression Omnibus (GEO) GSE230705. The GEO accession number for these datasets is listed in the key resources table.

The custom code used to process and visualize the data is deposited at Zenodo, and the accession DOI for the code is listed in the key resources table.

Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

For the single cell BM data, 5 healthy adults were enrolled at Emory University Healthy in 2018 between the ages of 25 ± 4 years. All were female. An additional six healthy adults were enrolled at Emory University between 2019 and 2022 for flow cytometry validation. The mean age was 28.5 ± 6.4 years old and five were female. One male adult age 60 years old was enrolled at University of California San Francisco (UCSF) in the ITN ATTAIN trial who was highly sensitized. All studies were approved by the Institutional Review Board at Emory University and informed consent was provided.

METHOD DETAILS

Cell sorting and library construction

Bone marrow aspirate was obtained under sterile conditions from the iliac crest from each of the 5 healthy adults. Mononuclear cells were isolated by Ficoll density gradient centrifugation and enriched by a custom-designed negative selection EasySep cell isolation kit from StemCell that removes CD66b+/GPA+/CD3+/CD14+ cells to limit sorting time of fragile BM ASC. Enriched mononuclear cells were stained with the following anti-human antibodies: IgD–FITC (Cat. no. 555778; BD Biosciences), CD3-BV711 (Cat. no. 317328; BioLegend), CD14-BV711 (Cat. no. 301838; BioLegend), CD19-PE-Cy7 (Cat. no. 560911; BD Biosciences), CD38-V450 (Cat. no. 561378; BD Biosciences), CD138-APC (Cat. no. 130-117-395; Miltenyi Biotech), CD27-APC-e780 (Cat. no. 5016160; eBiosciences), and LiveDead (L34966; Invitrogen). ASC subsets were FACS-sorted on a BD FACSAria II using a standardized sorting procedure with rainbow calibration particles to ensure consistency of sorts between individuals. ASC subsets were sorted as: popA ($CD19^{+}CD38^{hi}CD138^{-}$), popB ($CD19^{+}CD38^{hi}CD138^{+}$) and popD ($CD19^{-}CD38^{hi}CD138^{-}$). Up to 17,000 cells were FACS-sorted from each population into RPMI with 5% FBS to maintain viability.

FACS-sorted cells were kept on ice until proceeding with 10x Genomics processing. Cells were centrifuged at 500xg for 10 min at 4C to remove media. Cells were then washed twice in 0.04% BSA in PBS. During last wash, media was aspirated and cell volume was measured using a pipette. The whole sample was then taken for processing using the 10x Genomics 5′ v1 Single Cell platform. V(D)J and 5′ Gene Expression (GEX) libraries were constructed for each sample, following 10x Genomics protocol. QC for each library was performed at each step by Bioanalyzer. Final libraries were quantified by kappa qPCR and sequenced on a NovaSeq.

QUANTIFICATION AND STATISTICAL ANALYSIS

Pre-processing of 10x genomics scRNA-seq data and quality control

The 10x Genomics single-cell transcriptomic sequenced raw reads were aligned to GRCh38 reference and quantified per cell barcode using Cellranger v3.0.1 (https://support.10xgenomics.com/single-cell-gene-expression/software/downloads/latest). Genes expressed in at least 0.1%–0.3% (depending on the sample size) of the total cell population were regarded as expressed genes and retained for downstream analysis. To filter low-quality cells, we excluded cells with $\geq$ 30% of their UMIs coming from mitochondrial genes, or $\leq$ 800 total number of detected genes, or total number of UMIs $\leq$ 1000. We first retained cells with a detected gene number between 200 and 800 (labeled as LowGN in Figure S1A), but none of these genes were significantly expressed in the cell group. Moreover, we removed cells with total number of genes or UMIs $\geq$ 6,000 or 60,000, respectively, to control for potential doublets. Additionally, we filtered out cells with immunoglobulin genes corresponding UMI count percentage $\leq$ 5% to avoid contaminated non-antibody secreting cells. Then, we excluded contaminated cells expressing diagnostic cell markers, eg. CD3E (T cell), CD16 (encoded by *FCGR3A*) and CD14 (Monocytes), NKG7, GNLY (Natural killer cells), HBB (Erythrocytes), CD20 (encoded by *MS4A1*), PAX5, IRF8 (B cells) (Figure S1B).

Normalization and cell cluster detection

The scRNA-seq data was next analyzed using version 3.2.2 of the Seurat package.[87,88] The gene expression counts of each cell were normalized using regularized negative binomial regression (SCTransform) to account for sequencing depth.[93] Next, we selected the top 3,000 highly variable features (HVFs) based on their standardized expected variance after variance-stabilizing transformation, removed all the immunoglobulin genes from the HVF list, and then used HVF data from SCTransform residuals to perform principal component analysis (PCA). In the first run of our scRNA-seq analysis pipeline, we falsely included 2 misannotated Ig genes AC233755.1 and AC233755.2 in the HVF list, which resulted in 15 clusters plus 2 additional AC gene-driven cell clusters (Figure S2B). Subsequently, we further excluded these two genes before cell clusters detection. Next, we utilized Canonical Correlation Analysis (CCA) from r package Seurat to anchor each dataset, removing individual effect and generating an integrated dataset.[87] Using this CCA-integrated dataset, a graph-based clustering method was applied to build a shared nearest neighbor (SNN) graph in PCA space, after which the Louvain community identification algorithm was applied to group cells at the set resolutions, with higher values leading to a greater number of clusters.[94]To assess the stability of the clustering based on different combinations of running parameters (dimension numbers = 50, 60 and 70, PC numbers = 30, 50, and 70 and resolution parameter = 0.2, 0.5, 1.0 and 1.5), we computed the Rand Index (RI) between pairs of classifications derived with different parameters. Accordingly, Rand Index (RI),[95] which calculates the concordance of pairwise relationships between all pairs of cells, typically results in 86.5% similarity on average in cluster designation for each individual cell (Figure S2D).RI is a similarity measurement taking values from 0 (low) to 1 (high), which computes the proportion of cell pairs that are in agreement between cell clusters from two different parameters. Finally, we used 70 dimensions to anchor individual datasets and account for subject differences, and 50 PC were used to construct an SNN graph to detect cell clusters.[88,96] A resolution of 1.0 gave the most stable cell clusters (Figure S2D). 15 cell clusters from the first run were retained with the cells from the two AC gene-driven clusters dispersed throughout the remaining 15 clusters. Finally, these clusters were visualized in two dimensions using uniform manifold approximation and projection (UMAP).[92] Notably, the 15 subgroups were detected in all 5 individuals, with no identifiable subject-specific variability or Ig driven clusters (Figures S2C, S2E and S2G).

Differentially expressed and marker gene detection

Differentially expressed genes (DEG) between two cell clusters were identified based on LogNormalize SCTransform-adjusted count matrix, which is from data slot of SCTransform function in the Seurat r package. It's worth noting that the HVF list with excluded Ig genes was used to detect cell clusters, the whole gene expression list which included Ig genes was used to detect cell markers and DEGs. It was performed by using zero-inflated generalized linear models including individual as a random covariate, with the MAST package,[90] which was proven effectiveness on both real measured and simulated single-cell data.[97] Cell cluster marker gene detection used the same settings but compared gene expression between cells from each selected cell cluster versus all remaining cells from the other clusters. Marker genes were defined as significantly up-regulated genes in associated cell clusters with average natural-log fold change (avgLogFC) greater than 0.25 and Bonferroni adjusted p value less than 0.05, while top DEGs were also included down-regulated genes with avgLogFC less than -0.25.

GSEA hallmark enrichment analysis

For gene set enrichment analysis, all expressed genes were ranked in descending order by multiplying the negative log-p value (NLP) derived from DEG analysis by the sign of the avgLogFC between the two clusters. The resultant pre-ranked gene list was used as input into GSEA v4.0.3 Preranked analysis.[91] The enrichment score is derived from a multivariate U score.[98,99] Briefly, scores were calculated by averaging normalized expression levels for all the transcripts that were identified as maturation-associated DEGs, which were differentially expressed in any two adjacent stages of plasma cell maturation in Figure 1C or differentially expressed between early and late cell clusters and annotated in the selected HALLMARK pathways.[37]

Trajectory analysis

For the IgG dominant cell trajectory analysis, we used the slingshot method as it can detect the bifurcation, multifurcation, linear and tree-like differentiation topology.[89] We ran slingshot r package v1.4.0 (*99*) with UMAP embeddings as input data, cell cluster ids as cell labels and cluster 1 (PB) as the starting point, and all other values using default settings. Although numerous methods have been developed for single cell trajectory imputation, strikingly, out of the 45 methods reviewed by this study,[100] only three (PAGA and RaceID/StemID) can be used to detect disconnected graphs, but none of these releases are currently stable. We assumed that each lineage had its own progenitor population, so for the accuracy and stability of inference, we focused on dissecting the maturation paths only for IgG lineage cell populations which are likely to originate with the PB (cluster 1), whereas the progenitor for IgM remains unclear.

Transcription factor analysis

Transcription factors were combined from AnimalTFDB[101] and known human transcription factors.[102] For downstream analysis, we only focused on the TFs that were either identified as HVFs or markers of any one of cell clusters or differentially expressed in the comparison of any adjacent two clusters of PC maturation or between the early and late phases of PC maturation. Only those TFs expressed in at least 10% of any one of the BMPC subgroups were included in detection of cluster-specific TFs; and only those expressed in at least 20% of associated cell clusters and having avgLogFC greater than or equal to 0.25 between the two cell groups with the highest and lowest gene expression were selected for visualization. Averaging the expression of TFs by cluster id and defined maturation stage, each TF was assigned to the cell cluster or stage with the highest expression. For those TFs showing greater than or equal to 0.25 avgLogFC between assigned cell cluster and cell cluster with the second highest expression were labeled as cluster distinct. The same criterion was used to label stage-distinct TFs. Associations between pairs of TFs were evaluated using annotations in the STRING protein-protein interaction database[38]; only interactions that have a combined score greater than 500 and exist within the assigned cell cluster were retained for visualization.

Multicolor flow cytometry for experimental validation

MNC were isolated from 4 (Panel 1) and 2 (Panel 2) BM aspirate samples from heathy donors using a Ficoll density gradient and stained with the following anti-human antibodies: IgD-Brilliant Violet 480 (Cat. No. 566138; BD Biosciences), CD3-BUV737 (Cat. No. 612750; BD Biosciences), CD14-BUV737 (Cat. No. 612763; BD Biosciences), CD19-Spark NIR 685 (Cat. No. 302270; BioLegend), CD38-Brilliant Violet 785 (Cat. No. 303530; BioLegend), CD138-APC-R700 (Cat. No. 566050; BD Biosciences), CD27-Brilliant Violet 711 (Cat. No. 356430; BioLegend), BCMA-Brilliant Violet 421 (Cat. No. 357520; BioLegend), TACI-PE-Cy7 (Cat. No. 311908; BioLegend), OX40-Brilliant Violet 510 (Cat. No. 350026; BioLegend), GITR-Brilliant Violet 605 (Cat. No. 747664; BD Biosciences), and Zombie NIR Fixable Viability Kit (Cat. No. 423106; BioLegend). The anti-human antibodies that were used for staining BM MNC isolated from the patient with high donor-specific antibodies (DSA) awaiting kidney transplant in the ITN protocol (ClinicalTrials.gov Identifier: NCT04827979) include: IgD-BB700 (Cat. No. 566538; BD Biosciences), CD3-PE-Cy5 (Cat. No. 2363822; ThermoFisher), CD14-PE-Cy5 (Cat. No. 2319032; ThermoFisher), CD19-BUV395 (Cat. No. 740287; BD Biosciences), CD38-FITC (Cat. No. CYT-38F2-A; Cytognos), CD138-PE-Cy7 (Cat. No. 356514; BioLegend), CD27-BV605 (Cat. No. 562655; BD Biosciences), and Live/Dead (Cat. No. L34962; ThermoFisher). Cells were acquired on a Cytek Aurora spectral flow cytometer using Cytek SpectroFlo software (Cytek; the HD BM samples) or an LSR Fortessa X20 (special order research product with 5 lasers; BD Biosciences) and analyzed using FlowJo software (v10.8.1; with DownSample (v3.3.1) plugin; BD Biosciences).

Single cell VDJ sequencing (scVDJ-seq) and analysis

Cells were counted using a Bio-Rad TC10 cell counter and verified via hemocytometer. Cell numbers were adjusted to 1,000 cells per μl to allow for 10,000 single cells per sample loaded in the 10x Genomics Chromium device. The manufacturer's standard protocol for Chromium Next GEM Single Cell 5′ Reagent Kit, v2 and Chromium Single Cell Human BCR Amplification kit was used to generate libraries. Libraries were sequenced on an Illumina NovaSeq (paired-end; 2 × 150 bp; read 1:26 cycles; i7 index: 8 cycles, i5 index: 0 cycles; read 2: 98 cycles) such that more than 70% saturation could be achieved with a sequence depth of 5,000 reads per cell for VDJ libraries.

Analysis of single cell VDJ data was conducted using Cellranger v3.1.0 via the 10x Genomics cloud interface and an in-house developed informatics pipeline for clonal clustering and SHM analysis.[69] Fasta files from the Cellranger output were annotated with metadata and aligned to germline sequences using the IMGT/HighV-QUEST web portal.[103] All data from IMGT/HighV-QUEST were retained through the process and were used for mutation calculations and alignment analyses. The definition of lineages/clones is consistent with previous publication.[69] The frequency and distribution of somatic hypermutation were ascertained on the basis of non-gap mismatches of expressed sequences with the closest germline V_H sequence. Mutation frequencies were determined by calculation of the number of mutations in V regions relative to the number of bases in non-gap V regions. The ratio of replacement mutations to silent mutations were calculated for CDR and framework regions separately from the corresponding V_H areas. In sequences with non-zero replacement but zero silent mutations, the number of silent mutations was set to 1.[104] Merging of gene expression and VDJ data sets and subsequent analysis was conducted using in-house developed pipelines in conjunction with the immcantation pipeline[105] and Seurat. Circular visualization plots were created with Circos software v0.69-6.

Permutation test for mutation detection

In order to assess whether mutation numbers differ among clusters, we first randomly shuffled mutation frequency data in all the cells and grouped them using current cell clusters. We then performed ANOVA test and Tukey's HSD (honest significant difference) consecutively. After repeating the previous steps one hundred times, we applied multiple comparison adjusted p values from the TukeyHSD test to calculate the p value for permutation test: P(permutation test) = (Number of permutation tests showing p value from TukeyHSD < p value obtained from running real data)/100.

Cell Reports, Volume *42*

Supplemental information

Understanding heterogeneity of human bone marrow plasma cell maturation and survival pathways by single-cell analyses

Meixue Duan, Doan C. Nguyen, Chester J. Joyner, Celia L. Saney, Christopher M. Tipton, Joel Andrews, Sagar Lonial, Caroline Kim, Ian Hentenaar, Astrid Kosters, Eliver Ghosn, Annette Jackson, Stuart Knechtle, Stalinraja Maruthamuthu, Sindhu Chandran, Tom Martin, Raja Rajalingam, Flavio Vincenti, Cynthia Breeden, Ignacio Sanz, Greg Gibson, and F. Eun-Hyung Lee

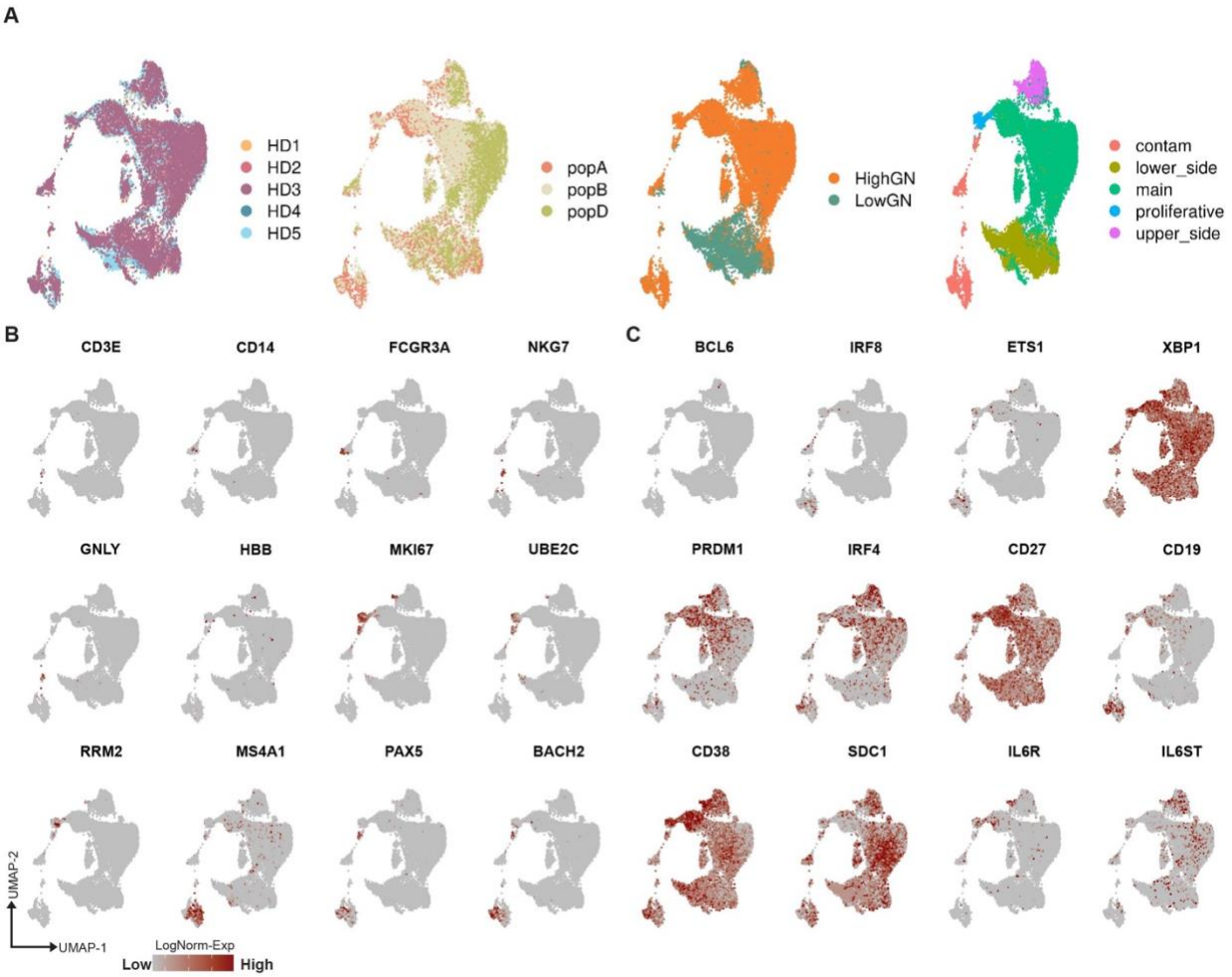


Figure S1. Integrated single-cell transcriptomic datasets. A) UMAP plots of single-cell transcriptomic profiles from 5 subjects before 2-step cell quality control and colored by sorted subject id, FACS-sorted labels, gene number group and cell subsets, which is mainly grouped by the location of cell population on UMAP, proliferating status and whether they are contaminating cells. HighGN: high gene number, indicating cells have greater than 800 detected genes, which is correspond to the gene cutoff in Fig. 1B; LowGN: low gene number, representing cells have greater than 200 detected genes but less or equal than 800 genes; contam: contaminating cell populations. B) and C) Contaminated cell population (B) and plasma cell (C) associated master gene expression. The redder the dot indicates the higher log-normalized gene expression.

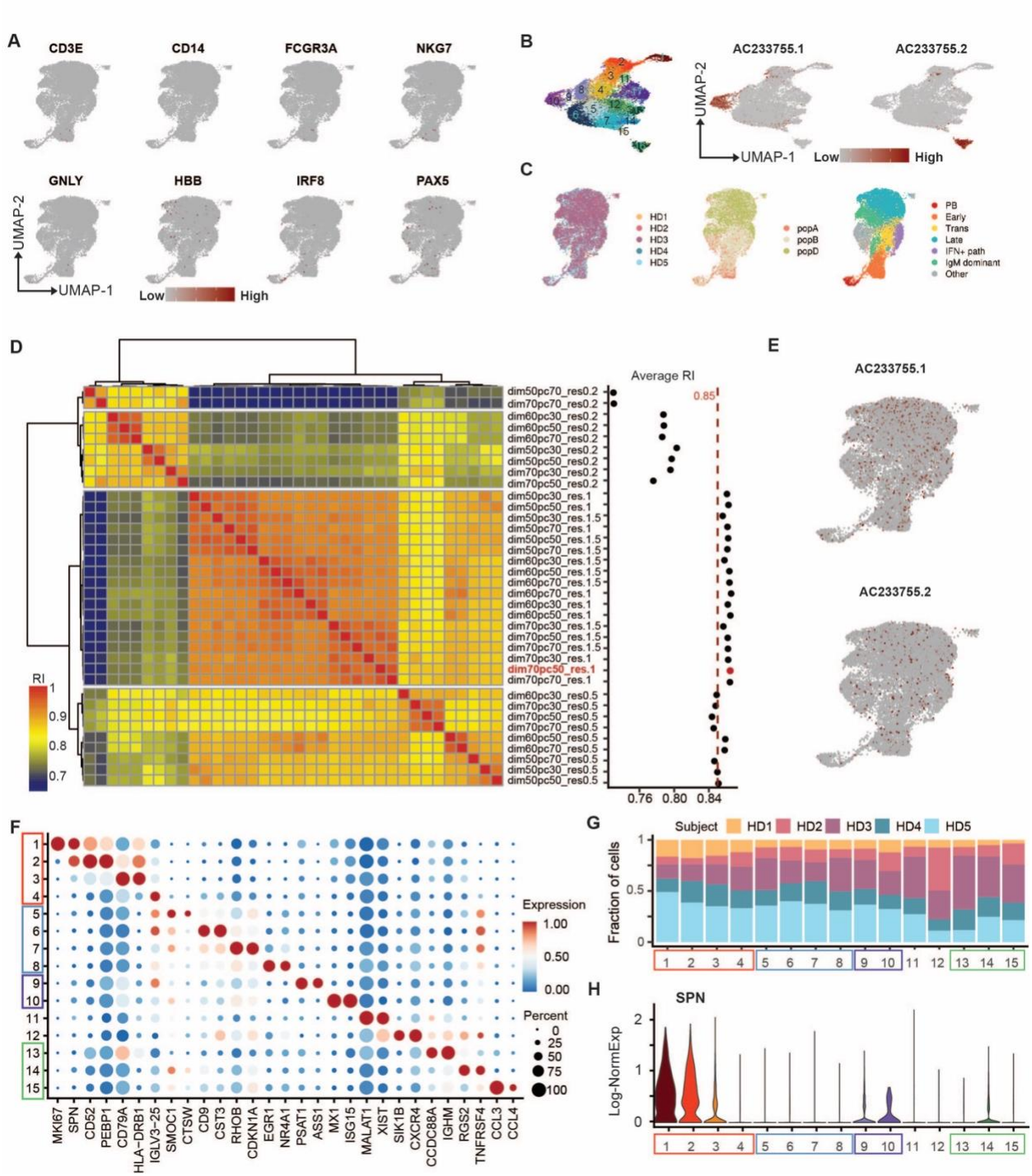


Figure S2. Stability and reliability of cell clusters. A) Expression of contaminated cell markers in current UMAP. B) Cell clusters (left) when not removing AC233755.1 and AC233755.2 genes before detecting cell clusters in the first run. Expression of AC genes in the previous UMAP (right). C) Current UMAP colored by sampled subjects, FACS-sorted cell populations and defined stages

of BMPC maturation. D) Two-way hierarchical clustering of RI by comparing cell clustering results from running different combinations of key cell clustering parameters (left), the parameter information was shown in the middle labels and scatter plot showed the average RI of each row of heatmap (right). Dashed red line labeled the RI value 0.85. (dim: dimension; pc: PC; res: resolution.) E) Expression of AC genes in current UMAP. F) Dot plots for expression of marker genes in each subgroup of bone marrow plasma cell defined in Fig. 1C. Here and in later figures, colors represent Min-Max normalized mean expression of marker genes in each cluster, and dot sizes indicate the proportion of cells expressing each marker gene. G) Fraction of cells in current scRNA-seq data defined cell clusters by sampled subjects. H) Log-normalized gene expression of SPN in cell clusters.

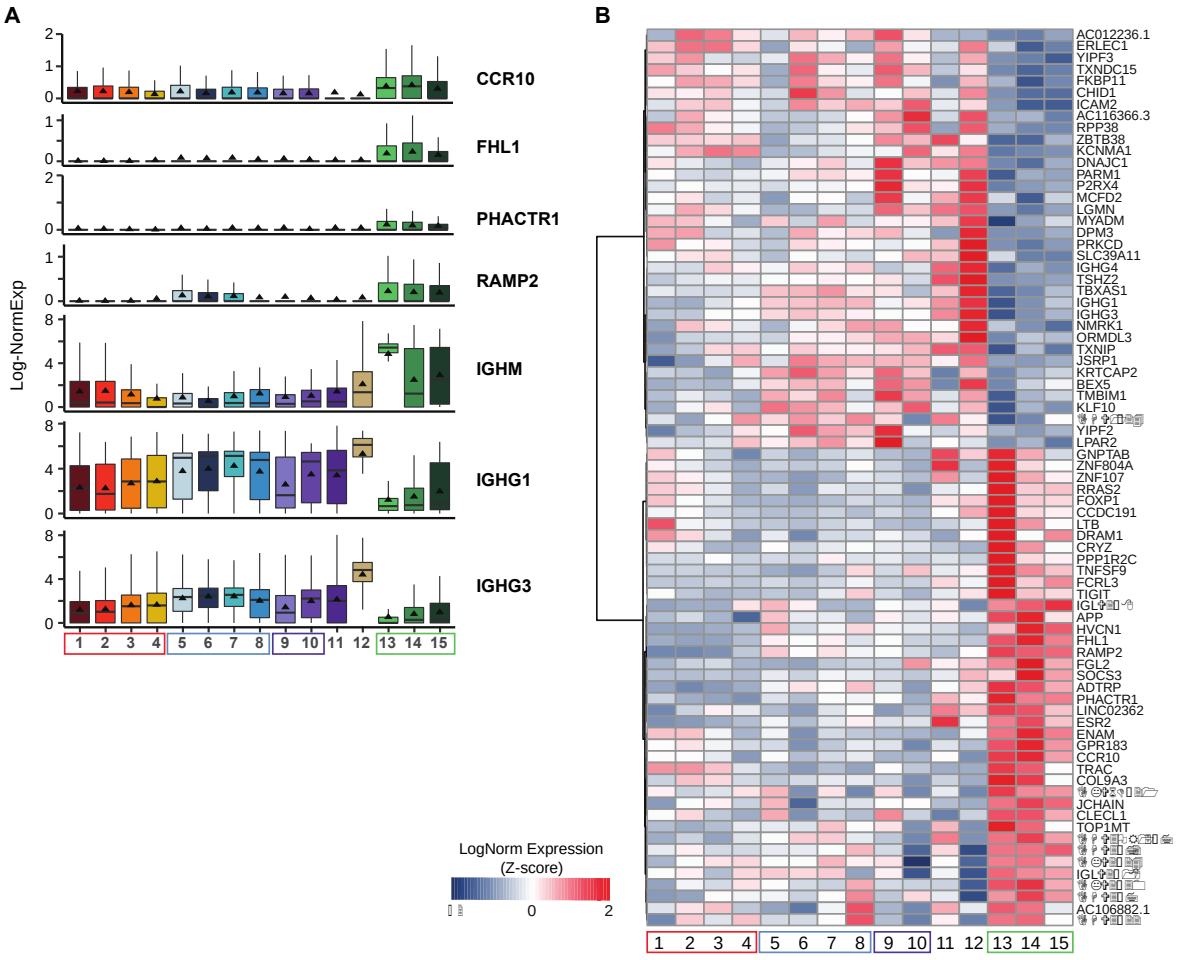


Figure S3. The features and pathways characterize the IgM dominant cell populations. A) Boxplots showing the expression of indicative IgM highly expressed genes as well as three Ig genes associated with somatic recombination of immune receptors in (C). B) Heatmap showed the row-scaled gene expression of CCR10 highly correlated genes (|Pearson correlation coefficient| ≥ 0.6, '||' implies absolute value).

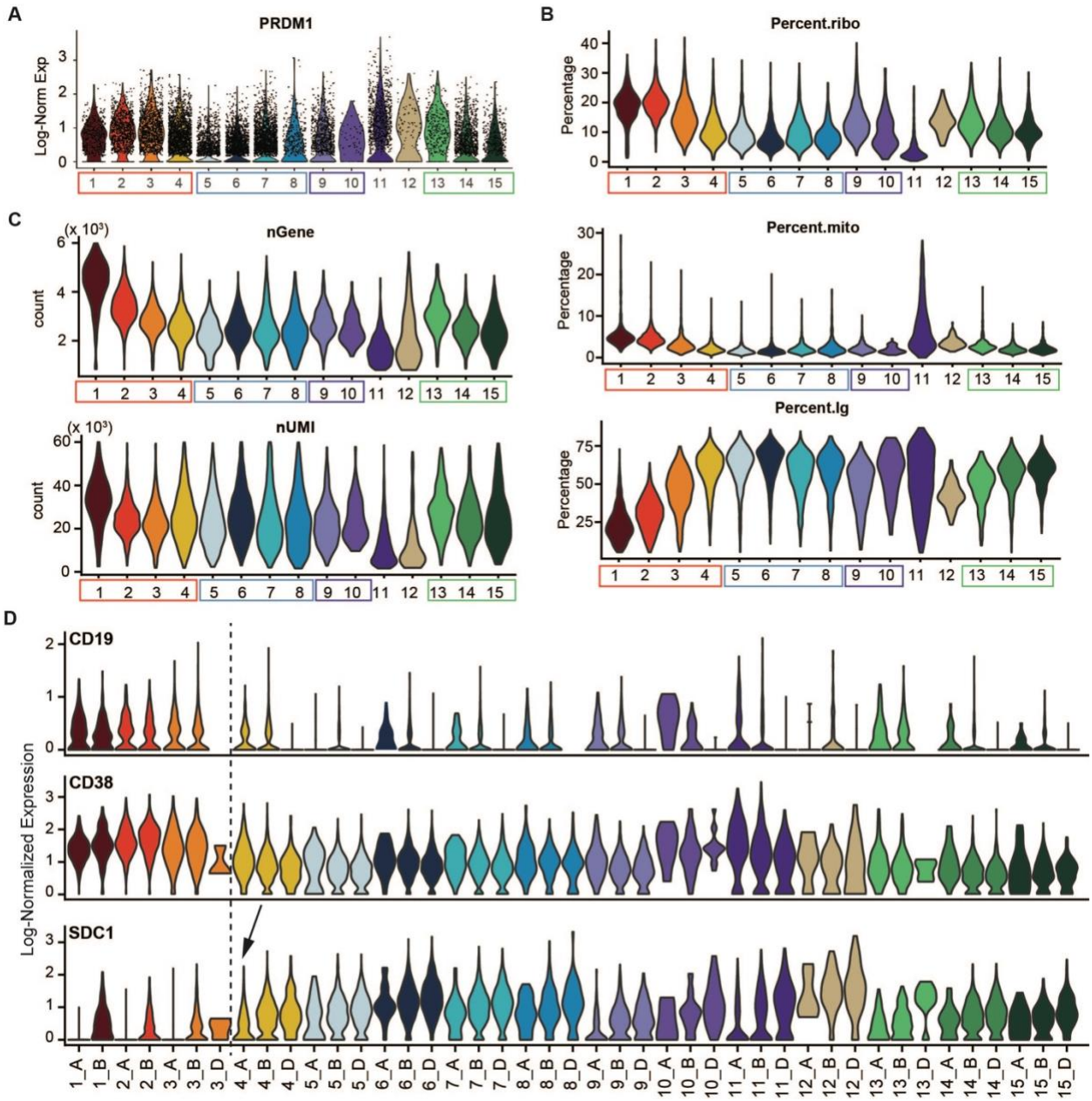


Figure S4. Distribution of selected features in cell clusters. A) Log-normalized gene expression of PRDM1 in cell clusters. B) and C) Violin plots show the percentage of transcripts coming from ribosomal genes (ribo), mitochondrial-encoded mitochondrial genes (mito), and immunoglobulin genes (Ig) (B) and the distribution of total detected gene numbers (nGene), total UMIs (nUMI) (C). D) X-axis is regrouped cell populations when taking both scRNA-seq defined cell subgroups and

FACS-sorted cell labels into account. Y-axis is the log-normalized gene expression. Arrow points out the FACS-sorted popA cells starting CD138 (encoded by SDC1) transcripts.

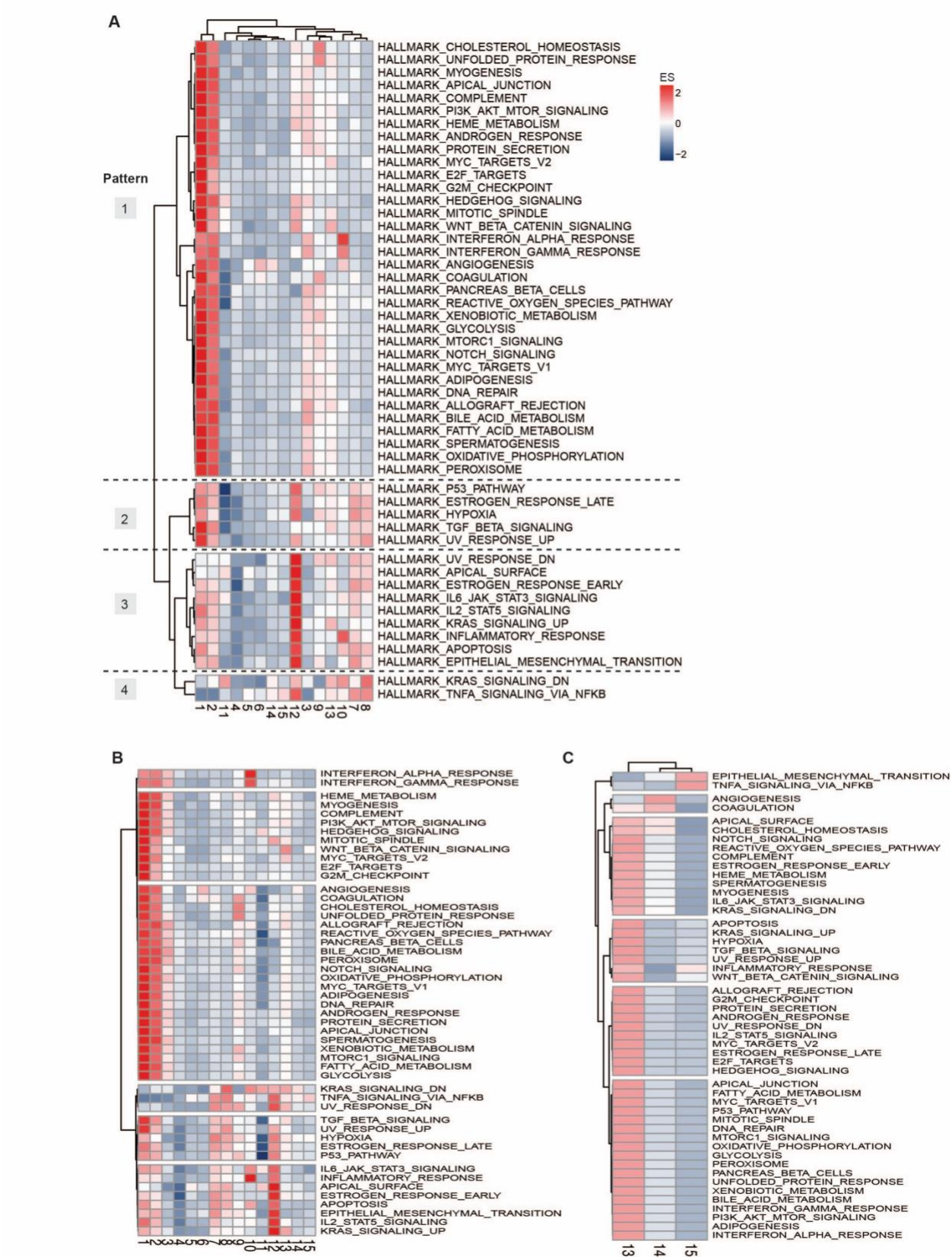


Figure S5. Patterns of Hallmark pathway enrichment analysis. A) Heatmap showed the two-

way hierarchical clustering of enrichment scores in each pathway. X-axis is the cell cluster id. The redder the color, the higher of enrichment scores. Numbers in the left-side of the heatmap showed the defined patterns. B) Heatmap showed the row hierarchical clustering of enrichment scores in each pathway. C) Heatmap showed the two-way hierarchical clustering of enrichment scores in IgM dominant cell lineage using DEGs are specifically differentially expressed in IgM lineage compared with DEGs detected by comparisons made between IgG1 dominant clusters.

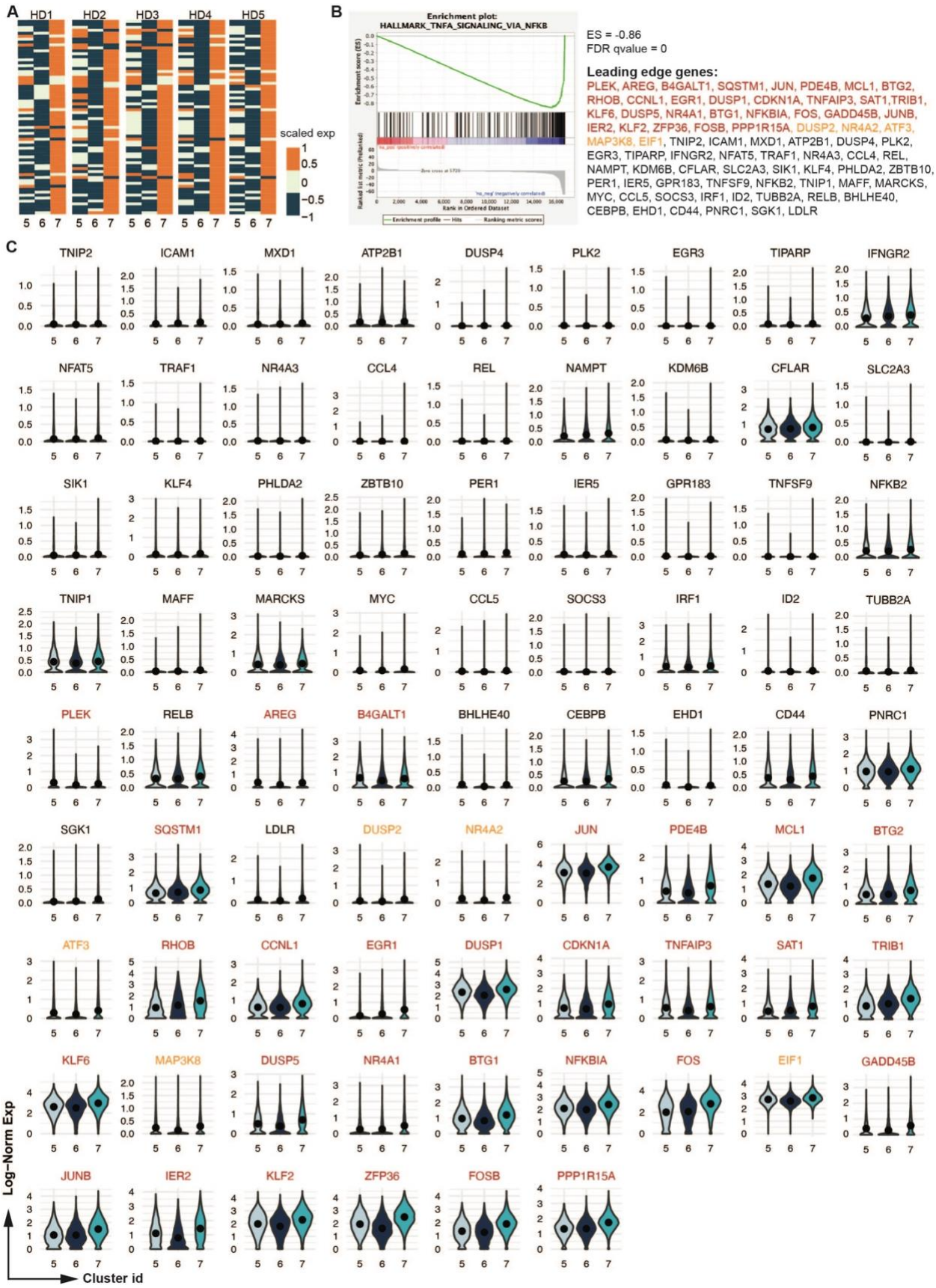

A
HD1
HD2
HD3
HD4
HD5
5 6 7
scaled exp
1
0.5
0
−0.5
−1
B
Enrichment plot:
HALLMARK_TNFA_SIGNALING_VIA_NFKB
Enrichment score (ES)
Ranked list metric (PreRanked)
Rank in Ordered Dataset
Enrichment profile
Hits
Ranking metric scores
ES = -0.86
FDR qvalue = 0
Leading edge genes:
PLEK, AREG, B4GALT1, SQSTM1, JUN, PDE4B, MCL1, BTG2, RHOB, CCNL1, EGR1, DUSP1, CDKN1A, TNFAIP3, SAT1,TRIB1, KLF6, DUSP5, NR4A1, BTG1, NFKBIA, FOS, GADD45B, JUNB, IER2, KLF2, ZFP36, FOSB, PPP1R15A, DUSP2, NR4A2, ATF3, MAP3K8, EIF1, TNIP2, ICAM1, MXD1, ATP2B1, DUSP4, PLK2, EGR3, TIPARP, IFNGR2, NFAT5, TRAF1, NR4A3, CCL4, REL, NAMPT, KDM6B, CFLAR, SLC2A3, SIK1, KLF4, PHLDA2, ZBTB10, PER1, IER5, GPR183, TNFSF9, NFKB2, TNIP1, MAFF, MARCKS, MYC, CCL5, SOCS3, IRF1, ID2, TUBB2A, RELB, BHLHE40, CEBPB, EHD1, CD44, PNRC1, SGK1, LDLR
C
TNIP2
ICAM1
MXD1
ATP2B1
DUSP4
PLK2
EGR3
TIPARP
IFNGR2
NFAT5
TRAF1
NR4A3
CCL4
REL
NAMPT
KDM6B
CFLAR
SLC2A3
SIK1
KLF4
PHLDA2
ZBTB10
PER1
IER5
GPR183
TNFSF9
NFKB2
TNIP1
MAFF
MARCKS
MYC
CCL5
SOCS3
IRF1
ID2
TUBB2A
PLEK
RELB
AREG
B4GALT1
BHLHE40
CEBPB
EHD1
CD44
PNRC1
SGK1
SQSTM1
LDLR
DUSP2
NR4A2
JUN
PDE4B
MCL1
BTG2
ATF3
RHOB
CCNL1
EGR1
DUSP1
CDKN1A
TNFAIP3
SAT1
TRIB1
KLF6
MAP3K8
DUSP5
NR4A1
BTG1
NFKBIA
FOS
EIF1
GADD45B
JUNB
IER2
KLF2
ZFP36
FOSB
PPP1R15A
Log−Norm Exp
Cluster id

Figure S6. Genes driven the difference of TNFα signaling via NFKB pathway in late LLPC. A) Heatmap showed the scaled corresponding TNFα singaling via NFκB pathway enriched maturation-associated DEG expressions in cluster 5, 6 and 7 in each individual. B) GSEA prerank analysis results using preranked genes between cluster 6vs7 (left); enrichment score, FDR qvalue, and leading-edge genes were listed (right). Gene colored in red were DEGs from Table S4 and in orange were cluster 6vs7 specific DEGs included in Table S6. C) Expression of differentially expressed leading-edge genes (|avgLogFC| > 0.25, Bonferroni adjusted p value < 0.05) from (B) in cluster 5, 6 and 7, the order of genes was corresponding to the gene rank in preranked gene list in Table S6.

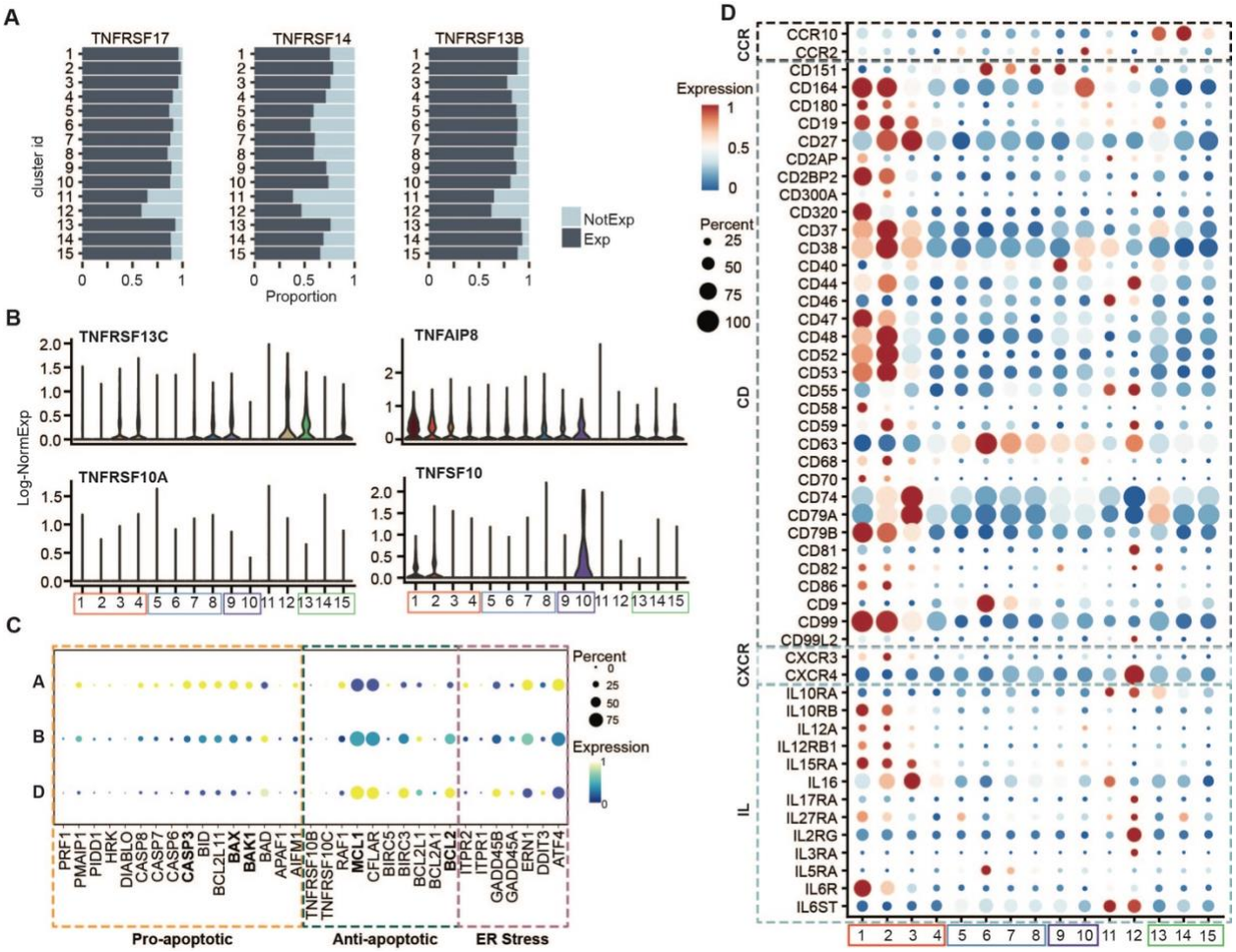


Figure S7. Visualization of indicative genes. A) Light and dark blue bar plots showed the proportion of cells in each cluster not expressed (NotExp) and expressed (Exp) gene TNFRSF17, TNFRSF14 and TNFRSF13B. B) Violin plots showed the expression of remaining expressed TNF family genes. C) Dotplot visualized the gene expression of pro-apoptotic, anti-apoptotic and ER-stress associated genes based on the FACS-sorted cell labels. The bluer the dots, the lower the gene expression. The size of dots represented the percentage of cells from associated cell population expressed the corresponding genes. D) Dotplot showed the gene expression from CCR, CD, CXCR ad interleukin (IL) families.

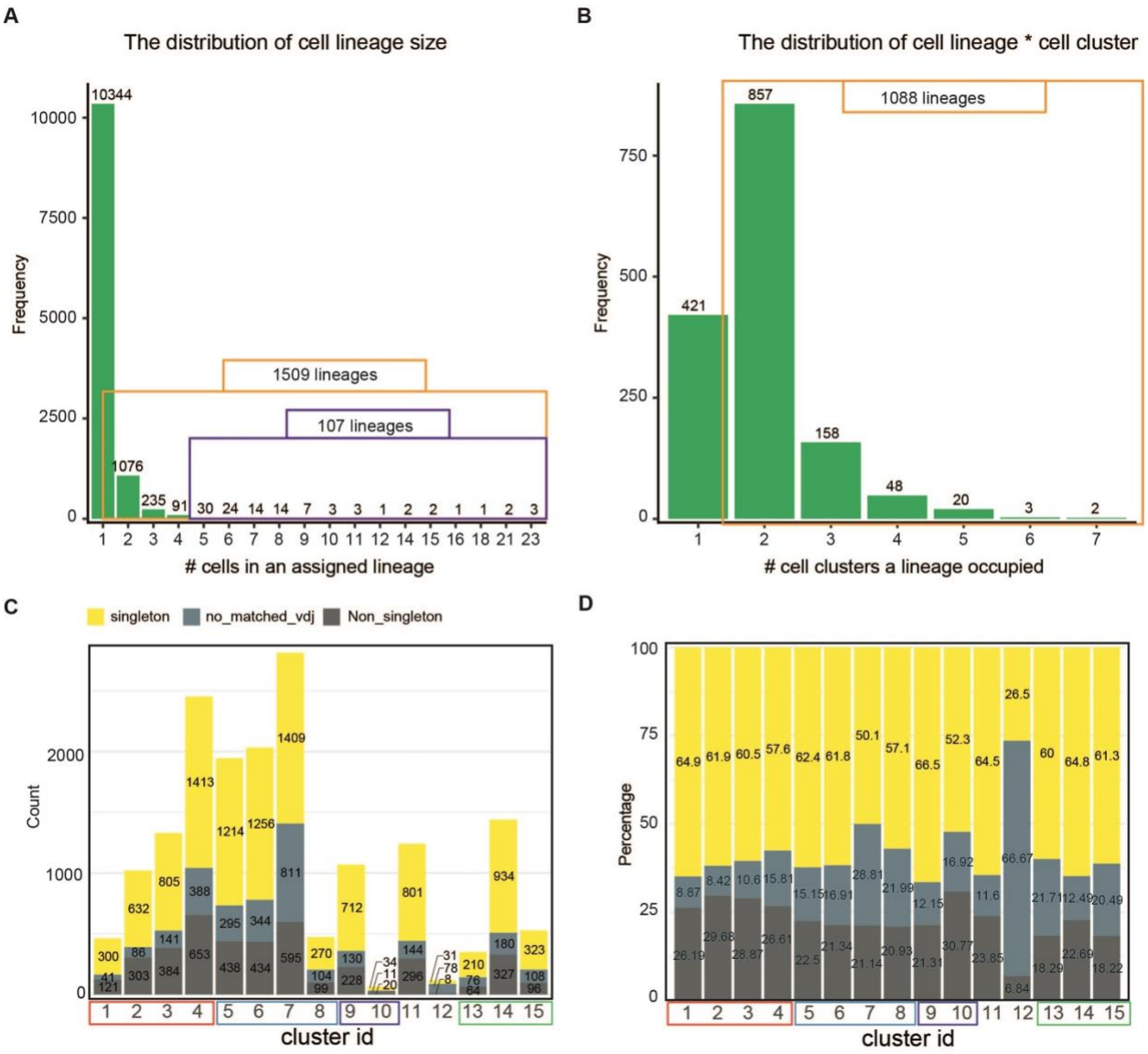
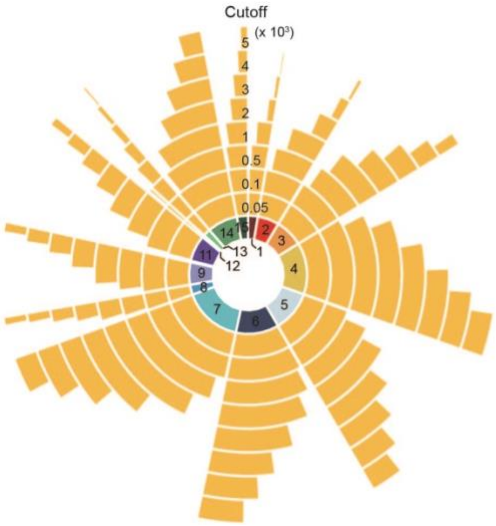


Figure S8. Summary of scVDJ-seq data detected clones/lineages. A) The distribution of lineage size. X-axis is the number of cells contained in detected cell lineages. Y-axis is the number of cell lineages. Orange and purple boxes summarize the total number of lineages that have at least 2 or 5 cells. B) The distribution of overlapped cell clusters of each non-singleton lineages. Orange boxes summarize the number of lineages having at least 2 cells and overlapped with at least 2 cell clusters. C) and D) Count the total number (C) and percentage (D) of cells in each cell cluster by if cells have matched scVDJ-seq information, if they are from singleton lineages (yellow) or not (dark grey). Cells with no matched scVDJ information were colored in ash.



cluster_id	count	50	100	500	1000	2000	3000	4000	5000
1	462	99.6	97.6	86.8	66.9	31.2	16.7	8.9	3
2	1021	99.8	99.3	87.3	66.4	33.3	15.8	6.7	2.5
3	1330	99.2	98	87.1	74.3	49.9	33.6	20.1	11.7
4	2454	99.9	99.5	92	79.6	58.5	41.4	29.5	20.6
5	1947	99.4	98.8	88.5	76.8	55	39	27.7	19.3
6	2034	99.5	99.1	94.4	83.3	63.9	46.8	35.1	26.2
7	2815	99.6	99.3	91.5	79.1	53.9	36.1	23.8	15.1
8	473	98.9	98.1	90.1	79.5	58.4	42.1	27.5	20.9
9	1070	99.6	99.2	87.4	72.2	45.6	27.9	18.2	11.3
10	65	96.9	96.9	89.2	76.9	52.3	38.5	18.5	13.8
11	1241	98	94.8	76.1	57.9	36.8	23.5	14.6	9.2
12	117	100	99.1	72.6	40.2	6.8	0.9	0	0
13	350	99.7	99.7	97.7	83.7	59.1	36.9	24	12.6
14	1441	99.7	99.4	95.2	85.7	64.1	45.2	29.8	18.9
15	527	100	100	93	80.3	56.9	41.7	26.6	18

Figure S9. Summary of Ig heavy chain gene transcript numbers in each cell cluster. Circos plot summarizes the proportion of cells in each cell cluster that pass the criterion by using different cutoffs to gate Ig heavy chain genes. The most inner circle is colored as Fig. 1C and numbers indicate corresponding cell cluster id. From inner to outer circle, the cutoffs are 50, 100, 500, 1000, 2000, 3000, 4000 and 5000. The table summarizes the percentage of cells in each cell cluster that passed cutoffs. The count column shows the total number of cells in each cell cluster.

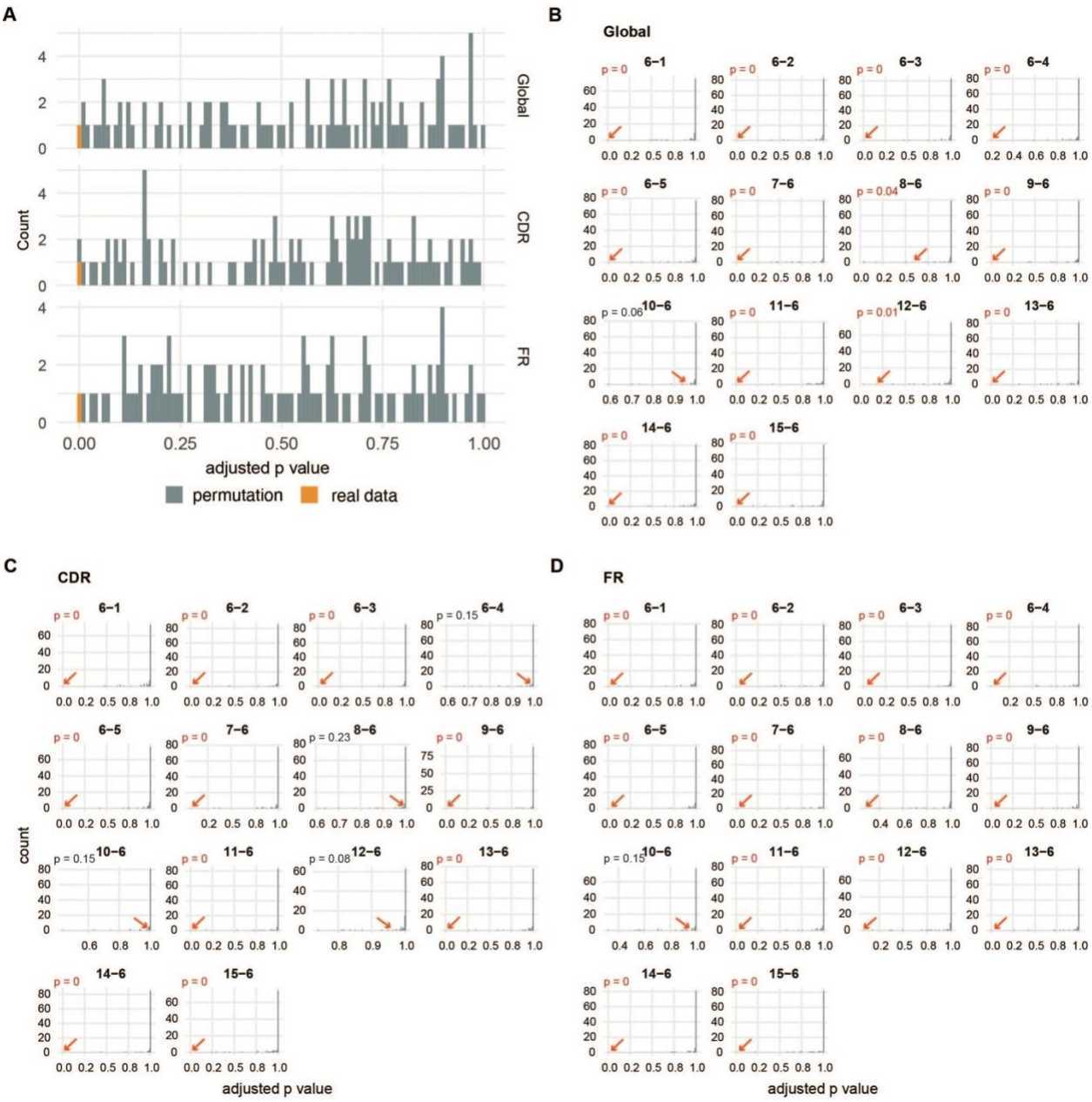


Figure S10. Permutation results for the mutation frequency test. A) Barplot showing the distribution of p values from ANOVA test for the global heavy chain variable region (Global), CDR and FR mutation frequency using real data and shuffled data from permutation test. Orange bars and grey bars indicated the p values from using real and randomly shuffled datasets. B) – D) The distribution of p values from Tukey's HSD (honestly significant difference) tests for the Global (B), CDR (C) and FR (D) mutation frequency between cluster 6 and other 14 clusters by using real

and randomly shuffled datasets. Red arrows pointed out the adjusted p value from Tukey's HSD test using real data, left upper side showed the p value from permutation test and significant results were colored in red. Titles indicated the objects of comparison.

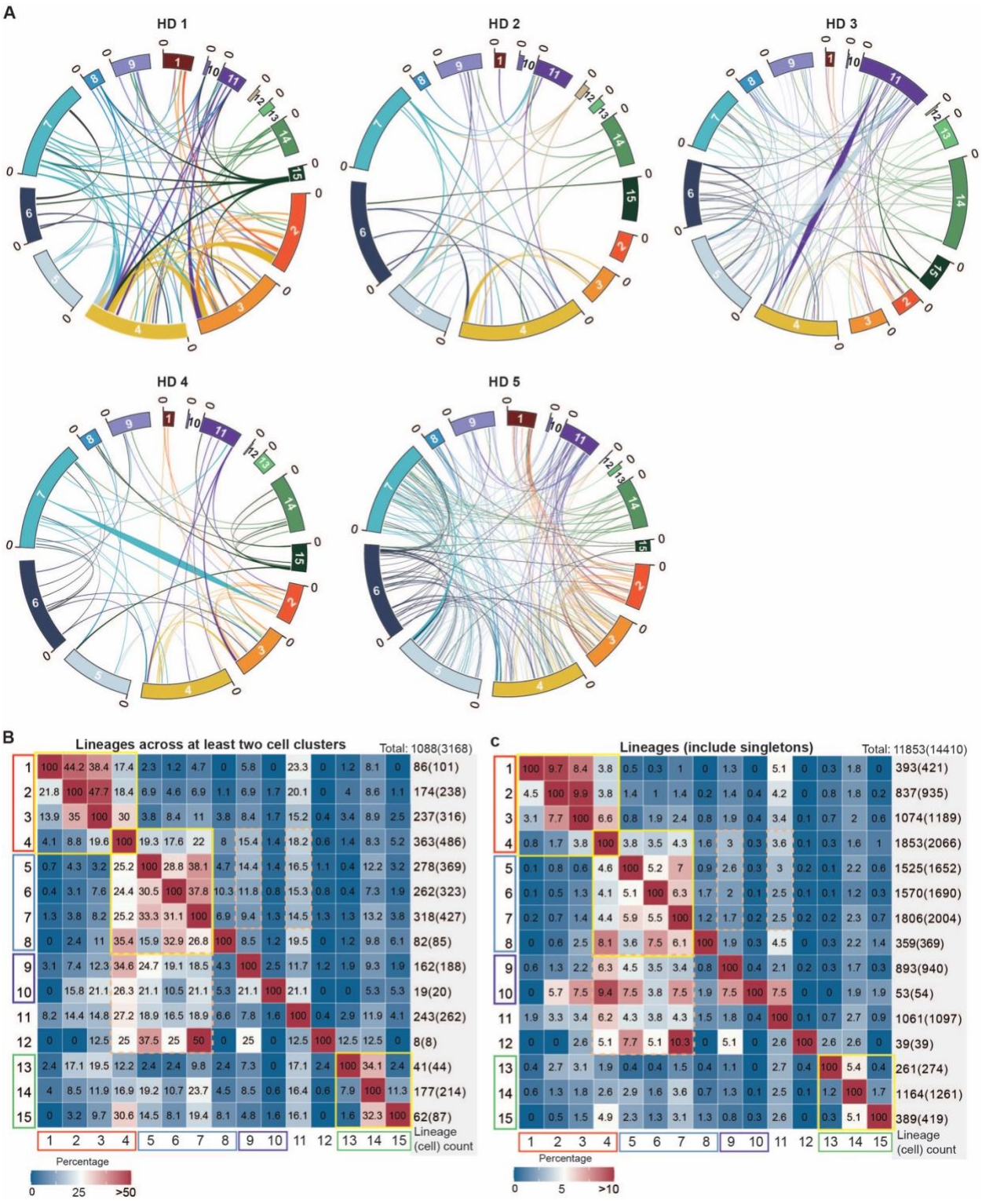


Figure S11. Clonal connectivity visualization in all cell clusters. A) Circos plot from main IgG1 dominant cell clusters 1 to 15 in a connecting detected lineages in each donor. Lineages were

colored by the latest cell cluster. B) – C) Percentage of connected clones between any two cell clusters when only considering lineages across at least two cell clusters (B) or lineages include singletons (C). Numbers in each block showed the percentage of lineages from each row cluster that were connected with other. The numbers showing in the grey boxes recorded the total number of lineages and cells showing in bracket in associated cell cluster. Yellow boxes highlighted the highly connected cell blocks and dashed orange box highlighted those clones showing high connectivity in one direction but not in another.

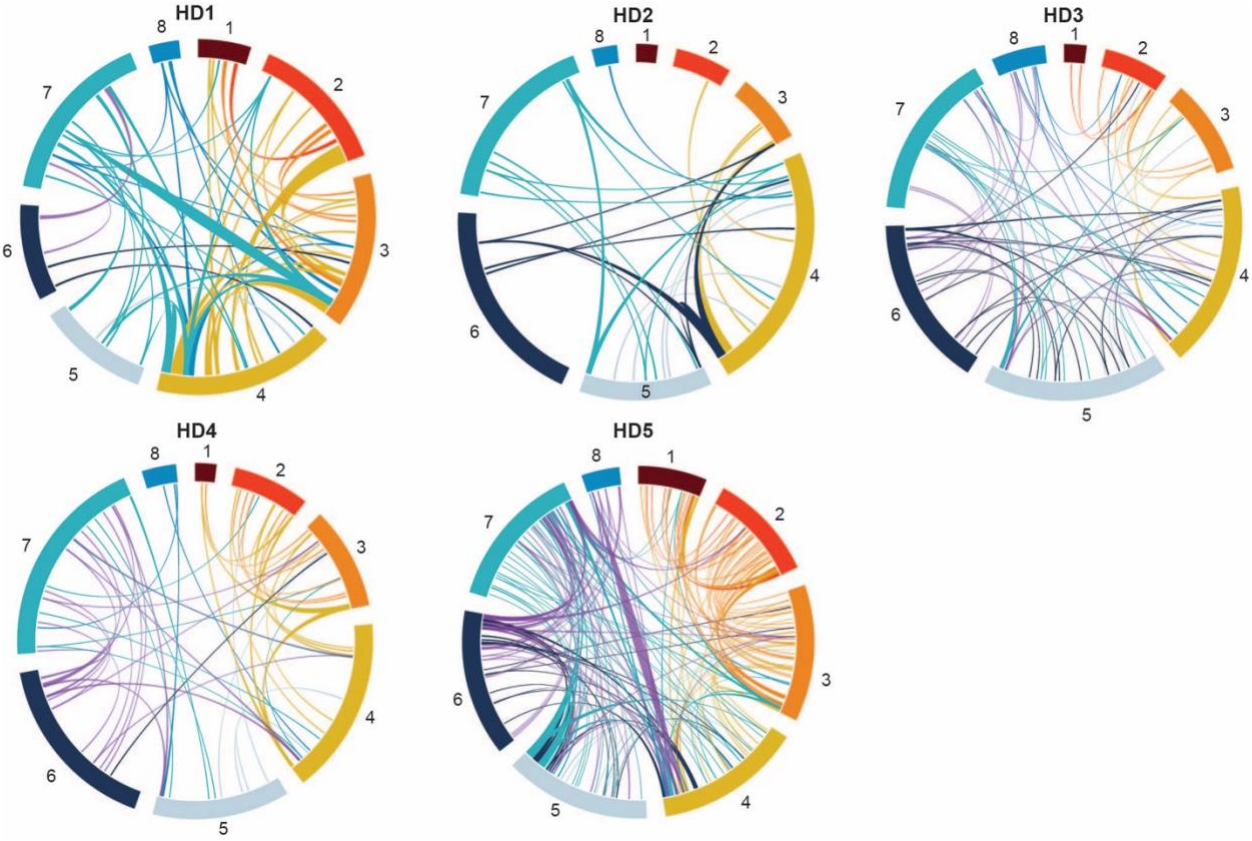


Figure S12. Clonal connectivity of nearly identical clones in IgG1-dominant clusters 1 to 8. A homology of 98% was used within the CDR3 region to cluster clones within each of the 5 subjects. Cell clusters are represented by the outer ring colors and sequences are grouped into clones in size descending order from counterclockwise to clockwise. Lines connecting cell clusters indicate a clone that was found in multiple clusters.

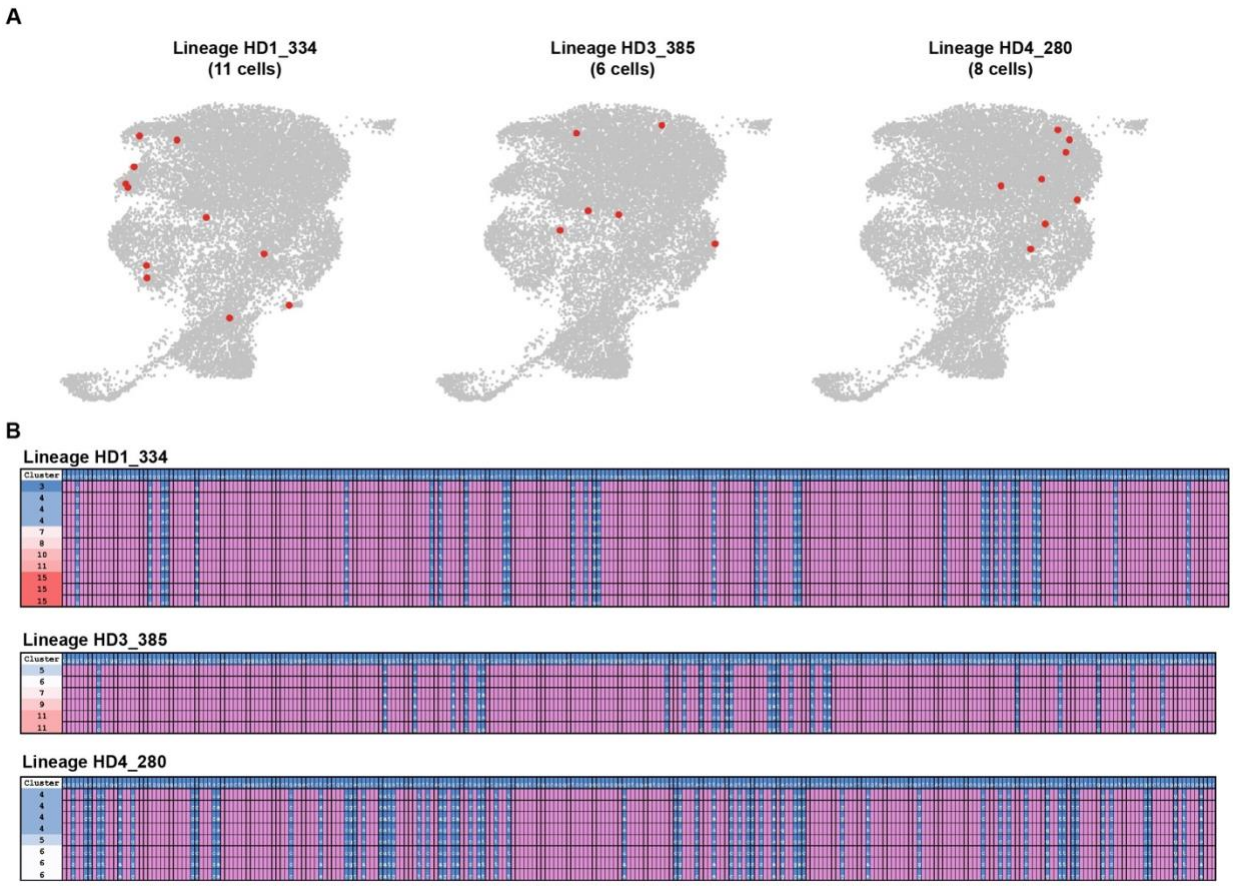


Figure S13. Cellular representation and alignment of clonal VH sequences that span multiple clusters. A) Clones were clustered by identical V, J, CDR3 length and 98% CDR3 homology to identify near identical clones and plotted on UMAP. Cells highlighted in red are members of the same clone. B) V region genes from the same clones were aligned against germline sequences. Blue squares indicate mutations in the individual sequences when compared to germline sequences and pink squares indicate matching nucleotides to the germline sequence. The cluster of each cell where the transcript was identified to be from is shown on the left.