American Society of Hematology
2021 L Street NW, Suite 900,
Washington, DC 20036
Phone: 202-776-0544 | Fax 202-776-0545
editorial@hematology.org

Integrative single-cell chromatin and transcriptome analysis of human plasma cell differentiation

Tracking no: BLD-2023-023237R1

Elina Alaterre (Institute of Human Genetics, France) Sara Ovejero (CHU Montpellier, France) Caroline Bret (CHU Montpellier, France) laure dutrieux (Institute of human genetics, France) Dassou Sika (Institute of Human Genetics, France) Raul Fernandez Perez (Department of Pathology, Hematopathology Section, Hospital Clinic, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), University of Barcelona, Spain, Spain) Marion Espéli (Université de Paris, IRSL, EMiLy, Inserm U1160, France) Thierry Fest (CHU de Rennes, France) Michel Cogné (Centre National de la Recherche Scientifique, France) José Martín-Subero (Department of Pathology, Hematopathology Section, Hospital Clinic, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), University of Barcelona, Spain, Spain) Pierre Milpied (Centre d'Immunologie Marseille-Luminy (CIML), France) Giacomo Cavalli (CNRS,) Jerome Moreaux (CHU Montpellier, France)

Abstract:
Plasma cells (PC) are highly specialized cells representing the end stage of B cell differentiation. We have shown that PC differentiation can be reproduced *in vitro* using elaborate culture systems. The molecular changes occurring during PC differentiationare recapitulated in this *in vitro* differentiation model. However, a major challenge exists to decipher the spatiotemporal epigenetic and transcriptional programs that drives the early stages of PC differentiation. We combined single cell (sc) RNA-seq and single cell ATAC-seq to decipher the trajectories involved in PC differentiation. ScRNA-seq experiments revealed a strong heterogeneity of the preplasmablastic and plasmablastic stages. Among genes that were commonly identified using scATAC-seq and scRNA-seq, we identified several transcription factors with significant stage specific potential importance in PC differentiation. Interestingly, differentially accessible peaks characterizing the preplasmablastic stage were enriched in motifs of *BATF3*, *FOS* and *BATF*,belonging to the AP-1 transcription factor family, that may represent key transcriptional nodes involved in PCD. Integration of transcriptomic and epigenetic data at the single cell level revealed that a population of preplasmablasts already undergone epigenetic remodeling related to PC profile together with UPR activation and are committed to differentiate in PC. These results and the supporting data generated with our *in vitro* PC differentiation model provide a unique resource for the identification of molecular circuits that are crucial for early and mature plasma cell maturation and biological functions. These data thus provide critical insights into epigenetic- and transcriptional-mediated reprogramming events that sustain PC differentiation.

Conflict of interest: No COI declared

COI notes:

Preprint server: No;

Author contributions and disclosures: EA and SO performed research and participated in the writing of the paper. CB, ME, MC, TF, GC and PM participated in the research and in the writing of the paper. JM supervised the research and the writing of the paper.

Non-author contributions and disclosures: No;

Agreement to Share Publication-Related Data and Data Sharing Statement: Single cell ATAC-seq and single cell RNA-seq data have been deposited in the Gene Expression Omnibus repository under the accession numbers GSE242324 and GSE242330.

Clinical trial registration information (if any):

Integrative single-cell chromatin and transcriptome analysis of human plasma cell differentiation

Elina Alaterre[1*], Sara Ovejero[1,2*], Caroline Bret[1,2], Laure Dutrieux[1], Dassou Sika[1], Raul Fernandez Perez[9], Marion Espéli[3], Thierry Fest[4], Michel Cogné[5], José Ignacio Martin-Subero[9,10,11], Pierre Milpied[6], Giacomo Cavalli[1] and Jérôme Moreaux[1,2,7,8]

1. Institute of Human Genetics, UMR CNRS-UM 9002, Montpellier, France
2. Department of Biological Hematology, CHU Montpellier, Montpellier, France
3. INSERM U1160 EMiLy, Institut de Recherche Saint-Louis, Université Paris-Cité, Paris.
4. Université de Rennes 1, INSERM, Établissement Français du Sang de Bretagne, Team B_DEVIL, UMR_S1236, Rennes, France; Laboratoire d'Hématologie, Centre Hospitalier Universitaire, Rennes, France.
5. Institut National de La Santé et de La Recherche Médicale, Unité Mixte de Recherche U1236, Université de Rennes, Etablissement Français Du Sang Bretagne, F-35000, Rennes, France; Centre Hospitalier Universitaire de Rennes, SITI, Pôle Biologie, F-35033, Rennes, France.
6. Aix Marseille Université, CNRS, INSERM, Centre d'Immunologie de Marseille-Luminy, Marseille, France.
7. University of Montpellier, UFR Medicine, Montpellier, France
8. Institut Universitaire de France (IUF), Paris, France
9. Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain
10 Departament de Fonaments Clínics, Facultat de Medicina, Universitat de Barcelona, Barcelona, Spain
11 Institució Catalana de Recerca i Estudis Avançats (ICREA) Barcelona Spain
* These authors contributed equally: Elina Alaterre, Sara Ovejero.

Corresponding Authors:
Pr Jérôme Moreaux
Laboratory for Monitoring Innovative Therapies
Department of Biological Hematology
Hôpital Saint-Eloi - CHRU de Montpellier
80, av. Augustin Fliche
34295 Montpellier Cedex 5
IGH - Institute of Human Genetics
CNRS UMR-UM 9002
http://www.igh.cnrs.fr
phone: +33 (0)4 67 33 79 03
fax: +33 (0)4 67 33 70 36
mail: jerome.moreaux@igh.cnrs.fr

Running Title: scATAC- and RNAseq analysis of PC differentiation

Word count: 4400 words

Data sharing statement:

For original data, please contact jerome.moreaux@igh.cnrs.fr

Single cell ATAC-seq and single cell RNA-seq data are available at Gene Expression Omnibus repository under the accession numbers GSE242324 and GSE242330.

Key Points

1/ Our results illustrate a complex and dynamic pattern of epigenetic and transcriptomic modifications in early PC genesis.

2/ Preplasmablasts already undergo epigenetic remodeling related to mature PC together with UPR priming through mTORC1 pathway activation.

Abstract

Plasma cells (PC) are highly specialized cells representing the end stage of B cell differentiation. We have shown that PC differentiation can be reproduced *in vitro* using elaborate culture systems. The molecular changes occurring during PC differentiation are recapitulated in this *in vitro* differentiation model. However, a major challenge exists to decipher the spatiotemporal epigenetic and transcriptional programs that drives the early stages of PC differentiation. We combined single cell (sc) RNA-seq and single cell ATAC-seq to decipher the trajectories involved in PC differentiation. ScRNA-seq experiments revealed a strong heterogeneity of the preplasmablastic and plasmablastic stages. Among genes that were commonly identified using scATAC-seq and scRNA-seq, we identified several transcription factors with significant stage specific potential importance in PC differentiation. Interestingly, differentially accessible peaks characterizing the preplasmablastic stage were enriched in motifs of *BATF3, FOS* and *BATF,* belonging to the AP-1 transcription factor family, that may represent key transcriptional nodes involved in PCD. Integration of transcriptomic and epigenetic data at the single cell level revealed that a population of preplasmablasts already undergone epigenetic remodeling related to PC profile together with UPR activation and are committed to differentiate in PC. These results and the supporting data generated with our *in vitro* PC differentiation model provide a unique resource for the identification of molecular circuits that are crucial for early and mature plasma cell maturation and biological functions. These data thus provide critical insights into epigenetic- and transcriptional-mediated reprogramming events that sustain PC differentiation.

Introduction

Plasma cells (PC) are highly specialized cells representing the end stage of B cell differentiation. They play an important role in humoral immunity[1]. On the transcriptional level, the differentiation of B cells into PCs is associated with substantial and coordinated changes [2–5].

Several in vitro models of human B to PC differentiation (PCD) were reported[6–11]. These systems could be used for functional interrogation in human cells related to the different stages of B to PC differentiation and are suited to high-throughput molecular characterization and experiments[2,3,12–14]. We have shown that PC generation can be modeled using multi-step culture systems where various combinations of activation molecules and cytokines are subsequently applied in order to reproduce the sequential cell differentiation occurring in the different organs/tissues *in vivo.* In these model, memory B cells (MBCs) differentiate into pre-plasmablasts (prePBs), plasmablasts (PBs), early PCs and, finally, into long-lived PCs (LLPCs), which may survive and produce continuously high amounts of immunoglobulins (Igs) for months *in vitro*[8]. The phenotype of *in vitro*-generated PBs is similar to the phenotype of the few PBs detected in the peripheral blood[6–8]. Moreover, the molecular events occurring during differentiation of B cells into PCs are recapitulated in these *in vitro* differentiation models[2,6,8,15].

Recently, we used next-generation sequencing technology to generate a comprehensive transcriptome database encompassing human in vitro PCD. Our results reveal 8419 differentially expressed genes classified into four temporal gene expression patterns[2]**.** Additionally, our analysis revealed numerous novel transcriptional regulators and helicases (BATF2, BHLHA15/MIST1, EZH2, WHSC1/MMSET, BLM and MYB) with consistent stage-specific overexpression and potential importance in PCD. Furthermore, our analysis revealed the upregulation of epigenetic factors at preplasmablast (PrePB) stage, a critical step during which cells actively proliferate and start secreting immunoglobulins. Finally, we have experimentally validated a role of for the BLM helicase and the histone methytransferase EZH2 in regulating cell survival, proliferation and maturation in PCD[16,17]. However, a major challenge exists to decipher the spatiotemporal epigenetic and transcriptional programs that drive the early stages of PCD[18,19].

In this study, we combined single cell RNA-seq and single cell ATAC-seq to decipher the trajectories involved in PCD. Our analyses reveal considerable transcriptional and epigenetic heterogeneity during the preplasmablastic stage of human PCD. Epigenetic analysis of the different stages suggests that BATF3 target genes may represent a key transcriptional node involved in PCD. Integration of transcriptome and epigenetic data at the single cell level showed that some prePBs already had an epigenetic profile similar to that of PCs in association with ER priming.

Materials and methods

Cell cultures

Peripheral blood cells from healthy donors were purchased from the French Blood Center (Toulouse, France) and $CD19^+CD27^+$ MBCs were purified as described[6]. From purified peripheral blood MBCs, prePB, PB, and PCs were generated using a three-step *in vitro* model as reported[6,7]. Standard culture conditions comprised 21% O_2, 5% CO_2, and 37°C and cells were cultured in Iscove's modified Dulbecco medium (Invitrogen, Waltham, USA) supplemented with 10% fetal bovine serum (Eurobio, Les Ulis, France). MBCs ($1.5x10^5$cells/ml) were doubly activated for 4 days by CpG oligodeoxynucleotide and CD40 ligand using a cocktail comprising 10µg/ml of phosphorothioate CpG oligodeoxynucleotide 2006 (Sigma-Aldrich, Saint-Louis, USA), 50ng/ml of histidine tagged sCD40L (R&D systems, Minneapolis, USA) and 5µg/ml of anti-poly-histidine mAb (R&D systems) with IL-2 (400U/ml) (R&D systems), IL-10 (50ng/ml) (R&D systems) and IL-21 (100ng/ml) (Peprotech, Cranbury, USA) cytokines in six-well culture plates. PBs were generated from prePBs ($2.5x10^5$cells/ml) by removing activating molecules and changing cytokine cocktail composed of IL-2 (400U/ml), IL-6 (50ng/ml) (Peprotech), IL-10 (50ng/ml) and IL-15 (10ng/ml) (Peprotech). Finally, PBs ($5.0x10^5$cells/ml) were differentiated into PCs adding IL-6 (50ng/ml), IL-15 (10ng/ml) and IFN-α (500U/ml) (R&D systems).

Flow cytometry and cell sorting

PrePBs, PBs and PCs were respectively purified at day (D) 4, D7 and D10 using Facs Aria cell sorter (Becton Dickinson, Franklin Lakes, USA) with a purity >95% as well as peripheral blood MBCs. MBCs were sorted using APC-conjugated anti-CD19 mAb and PE-conjugated anti-CD27 (BD Biosciences, #555415 and #555441, respectively). Cells produced in the culture system during differentiation were sorted using FITC-conjugated anti-CD20, PE-conjugated anti-CD38 and APC-conjugated anti-CD138 mAbs (Beckman Coulter, Brea, USA, #6602381, #A07779 and #B49219, respectively) for D4 prePBs ($CD20^-CD38^-$), D7 PBs ($CD20^-CD38^+CD138^-$), D10 PCs ($CD20^-CD38^+CD138^+$).

Supplementary information concerning methodology are included in Supplementary experiment procedures.

Results

Transcriptional features of stages during normal B to PC differentiation

The experimental strategy applied to obtain single-cell RNA-seq profiles on the four populations generated during B to normal PCD is illustrated in Figure 1A. The uniform manifold approximation and projection (UMAP) of the 6,392 cells showed three distinct compartments composed of MBCs, PCs and prePB/PB cells (Figure 1B). This projection revealed a highly specific transcriptomic profile

for MBCs and PCs, and a strong heterogeneity of the prePB and PB without a clear distinction between the two populations. There was no distinction between the replicates of cells generated with MBCs from the two healthy donors (Supplemental Figure 1). Each stage exhibited more than 300 differentially expressed genes that helped clearly distinguish the stages (Figure 1C). When we compared each stage with the rest of the dataset, the prePB stage presented the most differentially expressed genes with almost 2000 DEGs showing that the most important changes take place during this stage. As expecting, B cell transcription factors (TFs) (*PAX5*, *BCL6* and *BACH2*) were expressed in MBCs while the PC TFs (*IRF4*, *PRDM1* and *XBP1*) were strongly expressed in PCs (Figure 1D). The heatmap of the top10 DEGs of the four stages showed dynamic changes during B to normal PCD (Figure 1E). We then focused on genes that are differentially regulated during transitions, from MBCs to prePB, after B cell activation; from prePB to PB, when cells start to secrete antibody; and from PB to PCs. Gene set enrichment analysis of these deregulated genes validated the results previously reported using bulk RNAseq analyses (Figure 1F and Supplemental Table I)[7,20].

Single-cell chromatin accessibility reveals an over-representation of AP-1 TFs in prePB stage

To determine the variations in chromatin opening accompanying B to PCD, we applied single-cell ATAC-Seq, profiling in total 7,721 individual cells. PrePB, PB and PC stages were clearly separated from the MBC stage on the UMAP representation using peaks identified using MACS2[21] (Figure 2A). The number of differentially accessible peaks was higher in prePB (4660) than in other stages (MBC: 641; PB: 44; PC: 105) (Figure 2B). A fraction of cells at the prePB, PB and PC was characterized by very similar ATAC-Seq profiles, highlighting a strong similarity in chromatin structure even if prePB and PB cells are very different from PC at the transcriptomic level (Figure 1B). Interestingly, we observed a clear-cut chromatin decompaction at the prePB stage (Figure 2C), which is associated with a large number of ATAC-Seq peaks (Figure 2B). Moreover, the number of differentially accessible regions identified in each stage using sc-ATACseq strongly correlated with the number of differentially expressed genes identified using sc-RNAseq (R^2 = 0.9949; *p*-value < 0.001) (Supplemental Figure 2). The annotation of these differentially accessible peaks in each stage revealed a higher proportion of peaks localized on genes than on distal elements (Figure 2D). Pairwise comparisons between MBC and prePB, prePB and PB, and PB versus PC showed that the greatest chromatin changes were observed between MBC and prePB (251 open peaks in MBC and 4854 in prePB) after B cell activation (Figures 2D-G). We also observed significant changes between prePB and PB. Among genes that were differentially expressed at the transcriptome level, 29 genes for MBC, 170 genes for prePB and 11 genes for PC presented also chromatin remodeling (Figure 2H). These results revealed that B cell activation led to major epigenetic and transcriptomic remodeling. Among genes that were commonly identified using ATAC-seq and RNA-seq, we identified TFs such as *FOXP1* and *PAX5* for MBC; *ARID3A*, *BATF*, *BATF3*, *E2F4*, *ETS1*, *IKZF1*, *IRF2*, *MYB*, *SOX4*, *SPIB*, *SREBF2*, *STAT3*, *TFDP1* and *ZNF511* for prePB; and finally, *PRDM1* for PC (Table I). Motif

enrichment analysis using JASPAR and CIS-BP databases revealed a significant enrichment of motifs related to all the TF identified in the study and listed in Table I except ZNF511 (Supplemental resource 1). *BATF* and *BATF3* had differentially accessible peaks localized on the core gene and on distal elements (Figure 2I). Interestingly, differentially accessible peaks characterizing the prePB stage were enriched in motifs of BATF3, FOS and BATF belonging to the AP-1 TF family[19] (Figure 2J). In MBC, we found peaks enriched in KLF4, SPIB and KLF9 TF motifs whereas an enrichment in TCF4, ASCL1 and IRF4 motifs was identified in PCs. At the single cell level, the majority of prePB had a medium to high expression of AP-1 TF family (*BATF*, *BATF3*, *FOS*, *FOSB*, *FOSL1*, *FOSL2*, *JUN* and *JUND*)[19] (Figure 2K). Moreover, among the differentially accessible peaks in the prePB stage, 38% presented BATF3 motif (Figure 2L) corresponding to 1479 genes potentially regulated by BATF3 TF (Figure 2M). Among these genes, we identified 140 genes also upregulated in prePB, including *LDHA*, *EZH2*, *CXCR4*, *BIRC3*, *MKI67*, *TRAF1*, *IL21R*, *PAX5*, *CCL5*, *IKZF1*, *IRF5* and *CCR7*. The significant overexpression of BATF3 TF was validated at protein level in prePB (supplemental Figure 3A-B). Since BATF3 TF was previously identified operating in short impulse manner at prePB stage[2], BATF3 target genes may represent a key transcriptional node involved in PCD.

Integrating sc-RNAseq and sc-ATACseq reveals a more mature subpopulation of prePB characterized by an epigenetic profile of PC

To integrate sc-RNAseq and sc-ATACseq datasets, we used the top 50 differentially expressed genes from each stage identified with sc-RNAseq dataset to find anchors and predict cell stage of sc-ATACseq dataset. For the sc-ATACseq dataset, a gene activity matrix was calculated using the number of reads localized within genes. UMAP representation of transferred data showed a good superposition of sc-RNAseq and sc-ATACseq datasets, in particular for MBC and PC stages (Figures 3A-B). Almost a half of prePB from the ATACseq dataset were not predicted as prePB and almost a quarter of PB were not predicted as PB (Figure 3C). Interestingly, the remaining prePB were predicted as PB and PC and the remaining PB were predicted as PC (Figure 3D), revealing that some prePB and PB were characterized by a more mature epigenetic profile. Pairwise comparison between prePB predicted as prePB and prePB predicted as PC revealed key marker genes of PC including *XBP1*, *FAM46C*[20], *MZB1*[22] or *BTG2*[12] (Figure 3E and Supplemental Table II). These data underline that a subpopulation of prePBs already undergone epigenetic remodeling related to PC profile. Using RNA-seq data, the PC cell genes, such as *IFI6*, associated with open chromatin in prePB are still not expressed compared to mature PC (Figure 3F). To validate these results, we performed CHIP-seq of the histone marks H3K4me3, H3K27ac and H3K36me3. H3K36me3 is associated with transcriptional elongation in the gene body, H3K27ac with active regulatory elements including enhancers and promoters, and H3K4me3 with active/promiscuous promoters. The PC genes FAM46C, XBP1, MZB1 and IFI6 already showed active chromatin marks in prePB cells (Supplemental Figure 4). These results suggested that a population of prePB is already committed to generate antibody-secreting cells.

Pseudotemporal analysis of prePB and PB subpopulations by single-cell transcriptomic analysis

Knn-based clustering revealed 7 subpopulations including MBC in cluster 1, prePB in clusters 2 and 3, PB in clusters 4 and 5 and PC in clusters 6 and 7 (Figure 4A and Supplemental Figure 5A). 40% of the analyzed cells displayed a transcriptional profile associated with S-G2-M stages of the cell cycle (Figures 4B-C and Supplemental Figure 5B) represented mainly by prePB and PB[6] (Figures 4D-E). We selected only prePB and PB associated to S-G2-M stages in order to focus on the processes occurring during the transition from prePB to PB, while minimizing biases from cell cycle states (Figure 5A). To unravel potential differentiation trajectories and understand the progression between stages, cells were computationally ordered along pseudotime computed using Monocle3 (Figures 5B-C). We focused on genes differentially regulated along this trajectory, in particular genes coding TFs, epigenetic regulators and proteins involved in ligand/receptor interactions (Supplemental Table III). We identified six groups of deregulated genes: genes downregulated in cluster 1 (early prePB), genes downregulated in cluster 2 (mature prePB), genes upregulated in cluster 1, genes upregulated in cluster 2, genes first downregulated and then upregulated, (impulsed down) and the opposite (impulsed up) (Figure 5D). The majority of differentially expressed genes are downregulated (78.8%), and mostly in C2 (52.5%) (Figure 5E). Among TFs, *BATF3*, *IRF5*, *RUNX3*, *SPIB*, *PAX5*, *STAT5A*, *AHR*, *JUNB*, *STAT6* and *KLF6* were downregulated in the first instance, and *PHB2*, *TFAM*, *ETS1*, *TFDP1*, *YY1*, *E2F4*, *ZNF146*, *TP53*, *MAX* and *CEBPZ* were downregulated in C2 (Figure 5F and Supplemental Table IV). Concerning the epigenetic components, *AICDA*, *EZH2*, *EED*, *PRMT2* and *NCOA4* were downregulated during the transition from early to mature prePB stages while *PCNA*, *PRMT1*, *SET*, *MBD2* and *HDAC1* were downregulated during the transition from prePB to PB. The protein expression of AID in prePB along with a significant induction of 53BP1 and γH2AX, which characterize the presence of DNA strand breaks, was validated at protein level (Supplemental Figure 3). Investigating genes involved in ligand/receptor interactions, we found that *IL2RA*, *IL21R* and *CD40* are downregulated in a first instance after B cell activation (early prePB). The B cell markers *CD19*, *CD22*, *CD83*, *CCR7*, *CCL17* and *CCL22*[2,6,7] are downregulated in C1 (Figure 5F, Supplemental Figure 6 and Supplemental Table IV). *TACI* expression was downregulated in C2. PC surface markers *CD27*, *CD38*, *SLAMF7*, *BCMA* and *ITGA4* were upregulated in C1 together with *IL-6R*, *IL-6ST* and *INSR* (Figure 5F, Supplemental Figure 6 and Supplemental Table IV).

Subclustering of prePB and PB stages by single-cell transcriptomic analysis

The heterogeneity of the prePB stage encouraged us to increase the number of clusters to identity new transitional subpopulations of prePB. We obtained 5 clusters composed of 4 clusters of prePB and a unique cluster of PB (Figure 6A). Genes deregulated along differentiation process and identified using pseudotime analysis (Figure 5F and Supplemental Table IV) were used to order clusters in particular clusters 2, 3 and 4, corresponding to more mature prePB (Figure 6B). Pairwise comparisons revealed a

larger variation of gene expression observed during the first phases (cluster 1: 684 DEG) and at the end (cluster 5: 437 DEG) of the differentiation (Figure 6C and Supplemental Table V). The heatmap of the top 50 differentially expressed genes revealed very specific transcriptomic profiles for cluster 1, the early prePB, and cluster 5, the PB (Figure 6D). In cluster 1 (early prePB), we found several ligands and receptors strongly expressed after B cell activation (Figure 6D). This cluster was also characterized by an overexpression of B cell TFs (*IRF5*, *ZFP36L1*, *SPIB*, *BATF3*, *RUNX3* and *PAX5*) (Figure 6E) and *AICDA*. The cluster 5 (PB) was notably represented by an overexpression of PC receptors, such as *TNFRSF17 (BCMA)*, *SLAMF7 (CS1)*, *CD27*, *CD79A* and *CD38*, the PC TF *XBP1* and the *PSAP* ligand. Mature prePB were divided in 3 clusters (cluster 2: early mature prePB, cluster 3: transitional mature prePB and cluster 4: mature prePB; Figure 6D) and expressed specific markers like the *ETS1* and *ATF5* TFs overexpressed in early mature prePB (cluster 2). ATF5 is a TF involved in the survival pathway CREB3L2-ATF5-MCL1[17]. ETS1 was shown to mediate the transcriptional upregulation of MCL1 antiapoptotic factor and recruit AID to DNA sequence from the Igh locus[23]. *EGR1* and *FOS* are expressed later in transitional mature prePB (cluster 3). EGR1 TF participates in PCD program[24]. *KLF2* is expressed in mature prePB (clusters 4) and PB (cluster 5) (Figure 6E). KLF2 is involved in the control of PC homing in the bone marrow by controlling the expression of β7-integrin[25]. The clusters of mature prePB were also distinguished by the expression of some genes coding ligands and receptors, respectively *CALR*, *HLA-DRB6* and *SLC1A5*, *NCL*, *CANX* for cluster 2, *GPI*, *CD70*, *HLA-DRB6* and *TFRC*, *CXCR4*, *ENO1*, *F2R* for cluster 3, *CCL3* and *ITGA4* for cluster 4. Gene set enrichment analysis revealed that cluster 1 is enriched in genes regulated by NF-kB and STAT5, respectively in response to TNF and IL-2 stimulation (Figure 6F and Supplemental Table VI). This cluster was also enriched in genes involved in inflammatory response and p53 pathway. We also found genes up-regulated by the activation of the PI3K/Akt/mTOR signaling in cluster 1 and genes up-regulated through the activation of mTORC1 complex in all other cluster with a greater enrichment in cluster 2. In parallel, in this cluster, we observed an enrichment in genes involved in unfolded protein response (UPR) and MYC targets. This gene set was also found in cluster 3 in addition to genes involved in oxidative phosphorylation and glycolysis, and targets of E2F TF involved in cell cycle. Cluster 5 was enriched in genes involved in UPR in association with protein secretion.

Dual Activation of UPR during prePB and PB transition

We decided to focus on the dual activation of the endoplasmic reticulum stress observed in the first cluster of more mature prePB and later in PB. Pairwise comparison between cluster 2 and clusters 1, 3 and 4 corresponding to nearest clusters showed that 111 genes were overexpressed in cluster 2, including 19 genes (*ASNS*, *SLC7A5*, *HSPA5*, *HSP90B1*, *CALR*, *HSPA9*, *SERP1*, *PSAT1*, *SSR1*, *EDEM1*, *XPOT*, *TARS1*, *SPCS3*, *DNAJC3*, *PDIA6*, *HYOU1*, *EIF4EBP1* and *HERPUD1)* involved in UPR (Figures 7A-C and Supplemental Table VII). The cluster 5 overexpressed 179 genes compared to the cluster 4 including 10 genes involved in endoplasmic reticulum stress. Among them, 5 genes were

commonly found in clusters 2 and 5, whereas 5 genes were specific to the cluster 5 (Figures 7B-C and Supplemental Table VII). We also compared the genes upregulated in clusters 2 and 5 to identify other genes potentially involved in the UPR (Figure 7D). As reported in mice[26], the first activation of UPR, in prePB, is associated with an overexpression of genes involved in mTORC1 signaling whereas the second activation was associated with a downregulation of mTORC1 signaling genes and an overexpression of genes involved in protein secretion (Figures 7E-F). The heatmap of the genes involved in the UPR showed a clear distinction between the first activation occurring in early mature prePB (cluster 2) and the second activation associated with protein secretion in PB and PCs (Figure 7G). Early UPR activation was associated to a strong expression of *ASNS*, *SLC7A5*, *HSPA5*, *PSAT1*, *XPOT* and *EIF4EBP1* while the second activation was characterized by a strong expression of *TMBIM6*, *HERPUD1*, *VIMP* and *XBP1* in PCs. Interestingly, *HSPA5* coding a member of the heat shock protein (HSP) 70 family named binding immunoglobulin protein (BiP) was only co-expressed with one of the three transmembrane endoplasmic reticulum stress sensors in cluster 2 (Figure 7H). Interestingly, in cluster 2 we observed an imbalance in the ratio of reads corresponding to immunoglobulin light and heavy chains (Figure 7I), with a higher number of reads corresponding to IGH compared to IGL (Supplemental Figures 7A-B) that could explain the release of BiP, firstly described as an immunoglobulin heavy chain-binding protein[27], from its luminal domain at this specific moment. In cluster 2, only the Ire1 pathway was activated, known to splice XBP-1 (sXBP-1) to produce a highly active TF[28]. Moreover, the ligase responsible to the ligation of sXBP1 was also co-expressed in cluster 2 (Supplemental Figure 8A) and pseudotime analysis showed that *HSPA5* was first expressed, followed by a strong expression of *XBP1* (Figure 7J and Supplemental Figure 8B). We also detected some sXBP1 reads confirming that the splicing of XBP-1 occurred after the first UPR activation in early mature prePB (Supplemental Figure 9). The expression of BiP in prePBs together with the induction of XBP1 splicing was validated at the protein level (Supplemental Figure 3). To investigate the role of mTORC1-mediated UPR activation in PCD, we used rapamycin, which is an acute inhibitor of mTORC1. The drug was added from day 2 to day 4 or from day 2 to day 7 to investigate the effect in prePB. When used from day 2 to day 4, rapamycin treatment significantly affected the proliferation after activation of MBCs (Supplemental figure 10A). At day 4, 7 and 10, global cell counts were significantly decreased by 51%, 75% and 56%, respectively (Supplemental Figure 10A)). Rapamycin did not significantly affect cell viability at day 4, 7 and 10 (Supplemental Figure 10B). At the cellular level, the percentage of prePBs at day 4 was not affected by rapamycin (Supplemental Figure 10C). Conversely, at day 7, the percentage of prePBs was significantly increased whereas the percentage of PBs was significantly decreased under mTORC1 inhibition compared to control (Supplemental Figure 10C). At day 10, the percentage of mature PCs was significantly reduced (Supplemental Figure 10C). When used from day 2 to day 7, rapamycin treatment induced the same effects, resulting in inhibition of PCD (Supplemental Figure 10A, B and

D). The PI3K inhibitor idelalisib was used as a control. PI3K inhibition significantly affected proliferation without significantly affecting PCD (Supplemental Figure 10A-D).

The second UPR activation starting in cluster 5 is clearly associated to immunoglobulin gene expression (Figure 7K). Altogether, these data indicate that prePB already prime the UPR through mTORC1 pathway activation in order to prepare for PC function. XBP1 driven UPR activation will then be coordinated in PB in order to cope with the increase in antibody synthesis.

To validate our results, we used the large human tonsil atlas dataset[29] (supplemental Figure 11A). Among the 209,786 cells constituting the human tonsil atlas dataset, we selected the germinal center B cells (GCBC) and PC in S and G2/M phases (Supplemental Figure S11B-D). We identified a subpopulation of prePB characterized by low levels of MS4A1 and CD38 together with high expression of BATF, BATF3, EZH2, MYB, BLM, AICDA, NSD2 and PCNA (supplemental Figure 11E-G). These prePB presented a significant enrichment in MYC target genes, E2F target genes, mTORC1 signaling, oxidative phosphorylation, glycolysis, inflammatory response (Supplemental figure 11I) already identified in the PrePB of our in vitro PCD model (Figure 6F). Taken together, these results demonstrate the identification of transitional prePB cells in the human tonsil as previously reported[6].

Discussion

Herein, using sc-RNA-seq and sc-ATAC-seq of an in vitro PCD model, we provide direct evidence for epigenetic and transcriptional transition during preplasmablastic stage associated with PC genesis. Integration of chromatin accessibility and transcriptomic data revealed a more mature population of preplasmablastic cells characterized by open chromatin in PC genes without significant expression. Among them, we identified MZB1, FAM46C and XBP1. MZB1 is required for differentiation of PB and PC. MZB1 depletion resulted in deregulation of BLIMP1 target genes. Furthermore, MZB1 is required for the trafficking and maintenance of bone marrow PCs in mice[22]. FAM46C plays a role in sustaining ER biogenesis and secretory capacity in PC[20]. XBP1 is essential to support the UPR response and adaptation to Ig secretion[16]. Pseudotemporal analyses identified maturation trajectories in prePB with early prePB characterized by downregulation of B cell markers and B cell TFs together with upregulation of PC markers, adhesion molecules and growth factor receptors. The transition from early prePB to more mature prePB is associated with downregulation of *AICDA* and of PRC2 complex subunits. We previously reported that EZH2 is upregulated in prePB to repress B cell and PC transcriptional programs and sustain a transient prePB immature proliferative state that support their amplification[12]. Furthermore, the observed coregulation of *AICDA* and PRC2 complex genes support the reported role of EZH2 in DNA damage response inhibition in order to stimulate the survival of activated B cells during AID-mediated somatic hypermutation of Ig genes[30]. In mature prePB, we could identify a significant heterogeneity with sequential early activation of UPR followed by EGR1 and FOS activation and PC homing control mediated by KLF2. The first wave of the UPR activation

is associated with the mTORC1 pathway[31,32]. This mTORC1 mediated UPR activation was recently reported in murine activated B cells driving PC priming[26,33] before *XBP1* gene expression. Among the UPR genes commonly identified in murine activated B cells and human PrePB of our model, XPOT, ASNS asparagine synthetase, SLC7A5 amino acid transporter and PSAT1 metabolic enzyme were identified. SLC7A5 and ASNS are involved in protein synthesis. HSPA5, HSP90B1 and HSPA9 genes coding chaperones and facilitator of disulfide bond formation PDIA6 were also induced in prePB early UPR wave. Concomitantly, activation of MCL1-mediated PC-survival pathway was induced in these prePB with ETS1 and ATF5 overexpression[17,23]. This pathway is known to be activated in light-zone GC B cells that differentiate into PCs[18]. Transitional prePB overexpress EGR1, FOS, CXCR4 and TFRC. c-FOS/AP-1 positively regulates BLIMP1 expression and terminal PCD[19,34] and in malignant PCs[35]. TFRC coding CD71 is regulated by BLIMP1 in PCD and is known to modulate mTORC1[22]. In mice, EGR1 depletion in B cells inhibit PCD *in vitro* and in vivo[36]. CXCR4 overexpression promotes PC migration and maintenance in the BM[36]. Mature prePB overexpress KLF2 that participate in BM PC homing through the control of β7-integrin expression together with ITGA4 driving PC motility through VCAM-1 gradient[37] and interaction with stromal cells[38]. In human, MBCs are known to induce PCs faster and with reduced input signals compared to other B cells[39]. Our results revealed that a population of prePBs already undergone epigenetic remodeling related to PC profile together with UPR activation and are committed to differentiate in PC. We could confirm the presence of transitional prePBs in a large human tonsil atlas dataset[29]. A major challenge is to determine the functional contribution of identified epigenetic and transcriptional changes involved in PC generation. The human PCD models developed by several groups[6,8–11,18] and the data generated during their characterization[2,3,12,40,41] may be of particular importance for future functional validation studies using CRISPR-Cas9 mediated deletion[13,42].

No significant association could be defined between gene expression signature in the transitional prePB stages reported and malignant PC counterpart associated with multiple myeloma (MM) cancer. Among the differentially expressed genes identified in the five clusters of prePB and PB, high expression of *ETS1*[43], *NCL*, *SET*, *TFRC*[44] and *ENO1*[45] belonging to early and transitional mature PrePB were associated with significantly poor outcome in MM (Supplemental Figures 12A-B). However, high expression of *CD40*[46], *CD82*[47], *CD22*, *CALR*, *SLAMF7*[48] and *CD27*[49] belonging to early PrePB, early mature PrePB and PB were associated to good prognosis (Supplemental Figures 13A-B).

In sum, our results illustrate a complex and dynamic pattern of epigenetic and transcriptomic modifications in early PC genesis. These results and the supporting data generated provide a resource for the identification of molecular circuits that are crucial for early and mature PC biological function and survival. These data thus provide critical insights into epigenetic- and transcriptional-mediated reprogramming events that sustain PCD.

Acknowledgements

The J. Moreaux research group was supported by grants from INCA PLBIO19 FATidique, PLBIO22 PIC-ASO, ANR-18-CE15-0010-01 PLASMADIFF-3D, SIRIC Montpellier Cancer (INCa-DGOS-INSERM- ITMO Cancer_ 18004), ARC foundation PGA EpiMM3D, ARC foundation PGA RF20180207070 BAR-B cells, Institut Carnot CALYM, Labex EpiGenMed, FFRMG (AAP-FFRMG-2021), the European Union (Project 101097094 — ELMUMY), INSERM PSCI 2020 Smooth-MM, AAP ECOPHYTO – PELYCANO, MUSE LabUM Epigenmed, MSDAVENIR EpiMuM-3D, AAP READYNOV and Institut Universitaire de France. SO is supported by a grant from Fondation de France.

Authorship Contributions

EA and SO performed research and participated in the writing of the paper. CB, ME, MC, TF, JIMS, GC and PM participated in the research and in the writing of the paper. DS and RFP participated in the bioinformatic analyses. LD participated in the research. JM supervised the research and the writing of the paper.

Disclosure of Conflicts of Interest

The authors declare no conflict of interest.

Tables

Table I: Transcription factors and epigenetic enzymes upregulated and showing a more open chromatin state

	MBC	prePB	PB	PC
TFs	FOXP1, PAX5	ARID3A, BATF, BATF3, E2F4, ETS1, IKZF1, IRF2, MYB, SOX4, SPIB, SREBF2, STAT3, TFDP1, ZNF511	-	PRDM1
EEs	KDM2B	GATAD2A	-	-

Figures Legends

Figure 1: Single-cell transcriptomics analysis of memory B cells, pre-plasmablasts, plasmablasts and plasma cells during B to plasma cell differentiation. (A) Schematic representation of the *in vitro* model of B to PC differentiation. Memory B cells from human peripheral blood were purified and cultured with activating molecules, sCD40L and ODN, and cytokines, IL-2, IL-10 and IL-21 to

obtain pre-plasmablasts at day 4. Cells were then cultured with IL-2, IL-6, IL-10, IL-15 and IL-21 cytokines to obtain plasmablasts at day 7. Finally, plasmablasts were cultured with IL-6, IL-15 and IFNα until day 10 to obtain PCs. Flow cytometry gating of CD19+/CD27+ memory B cells at day 0, CD20-/CD38- preplasmablasts at day 4, CD20-/CD38+ plasmablasts at day 7 and CD38+/CD138+ PCs at day 10. Schematic representation of the BD Rhapsody single-cell analysis system used in this study. MBC, prePB, PB and PC were thawed and tagged with 4 different tags to associate, after sequencing, each read to one stage. The four populations (almost 10,000 cells) were pooled and leaded onto a cartridge composed of more than 200,000 wells. Unique barcoded beads were added in excess and after washing, each cell was associated to a unique bead, allowing the association of each read to a unique cell. Then, cells were lysed and mRNA was hybridized on the beads. To finish, beads were recovered to synthetize cDNA and amplify libraries prior to sequence. **(B)** UMAP representation of the four stages identified using tags and demultiplexing. **(C)** Number of positive differentially expressed genes identified for the four stages using pairwise comparisons (One stage *versus* All other cells). **(D)** mRNA expression of B cell transcription factors: *BACH2*, *BCL6* and *PX5*; and PC transcription factors: *IRF4*, *PRDM1* and *XBP1*. **(E)** Heatmap of the top 10 genes up-regulated of each stage. **(F)** Gene ontology enrichment analysis showing both pathways enriched in up-regulated and down-regulated genes during transitions: from MBC to prePB, from prePB to PB and from PB to PC.

Figure 2: Single-cell chromatin accessibility of memory B cells, pre-plasmablasts, plasmablasts and plasma cells during B to plasma cell differentiation. (A) UMAP representation of the four stages analyzed separately and then merged together. Peaks detected with MACS2 peakcalling were used for UMAP representation. **(B)** Number of differentially accessible peaks identified for the four stages using pairwise comparisons (One stage *versus* All other cells). **(C)** Cells were fixed with 4% PFA for 10 minutes at different time points: MBCs (Day 0), PrePBs (Day 4), PBs (Day 7) and PCs (Day 10). Immunofluorescence to detect H3K27me3 levels (green) was performed with an anti-H3K27me3 antibody. DNA was stained with DAPI (red). Scale bar = 10 μm. **(D)** Proportion of peaks localized on genes (in blue) and distal elements (in green) for each stage. **(E, F, G)** Volcano plots of differentially accessible peaks identified during transitions: from MBC to prePB, from prePB to PB and from PB to PC, respectively. Peaks identified as significantly accessible were colored in blue (*P*-value < 0.05 and log2(Fold Change) > 0.25). **(H)** Venn diagrams representing genes that were up-regulated in RNA-seq dataset (in blue) and/or associated with more open chromatin in ATAC-seq dataset (in green). **(I)** Peak tracks of BATF and BATF3 revealing differentially accessible peaks on BATF and BATF3 genes and on distal elements. **(J)** Transcription factor motif enrichment of differentially accessible peaks for each stage. **(K)** mRNA expression of transcription factors belonging to the AP-1 family in the four stages using the RNA-seq dataset. **(L)** Proportion of differentially accessible peaks in the prePB stage enriched in BATF3 motif. **(M)** Venn diagram of the number of genes up-regulated in the prePB stage identified using the RNA-seq dataset and the number of genes

associated with a more open chromatin enriched in BATF3 motif identified using ATAC-seq dataset. Common genes represented potential BATF3 targets.

Figure 3: Integration of sc-RNAseq and sc-ATACseq datasets. (A, B) UMAP representation of sc-RNAseq and sc-ATACseq merged dataset. First, using the top50 differentially expressed genes of each stage from the sc-RNAseq dataset and the number of reads within genes of interest from the sc-ATACseq dataset, a gene activity matrix was calculated for sc-ATACseq dataseq to find and set anchors. Gene expression values of sc-ATACseq dataset were predicted using the global gene expression values of sc-RNAseq dataset and identified anchors. Both sc-RNAseq and sc-ATACseq gene expression matrices were finally merged. **(C)** Number of cells predicted using the gene activity matrix *versus* actual number of observed cells at each stage (MBC, prePB, PB and PC). **(D)** Percentage of prePB and PB predicted as MBC, prePB, PB and PC. **(E)** Volcano plots showing differentially expressed genes (using the gene activity matrix) between the prePB predicted as prePB and the prePB predicted as PC. Genes identified as significantly differentially expressed were colored in blue (*P*-value < 0.05 and log2(Fold Change) > 0.25). **(F)** *IFI6* expression observed in MBC, prePB, PB and PC using respectively sc-RNAseq (above) and sc-ATACseq (below) datasets. High and low expression were represented in dark blue and in yellow, respectively.

Figure 4: Identification of subpopulations within the different stage of B to plasma cell differentiation. (A) Seurat k-nearest neighbors clustering identified 7 clusters: 1 cluster corresponding to MBC and 2 clusters for each stage of prePB, PB and PC. **(B)** Identification of quiescent cells (G1) and proliferative cells (S and G2M) using the Seurat cell-cycle scoring. **(C)** mRNA expression of CDC20, CDK1, MKI67 and PCNA involved in cell cycle. High and low expression were represented in dark blue and in yellow, respectively. **(D)** Cell cycle distribution of each stage. **(E)** Heatmap displaying the average expression of selected genes in clusters identified in A.

Figure 5: Pseudotemporal analysis of prePB and PB subpopulations. (A) UMAP representation of proliferative prePB and PB. **(B)** UMAP projection colored by normalized pseudotime analysis. **(C)** Clusters identified using the monocle package and used to define trajectories. **(D)** Temporal gene expression patterns from prePB to PB. **(E)** Proportion of genes and genes coding transcription factors (TFs), epigenetic enzymes (EEs) and proteins involved in ligand/receptor interactions deregulated along the trajectory according to the expression patterns defined in D. **(F)** Plots of the expression of top differentially expressed genes coding for TFs, EEs, ligands and receptors in function of pseudotime.

Figure 6: Identification of new subpopulations of prePB and PB stages. (A) Seurat k-nearest neighbors clustering identified 5 clusters, including 4 clusters for prePB and 1 cluster for PB. **(B)** Violin plots representing the expression of top marker genes identified using the pseudotime analysis for each cluster. **(C)** Number of positive differentially expressed genes identified in the 5 clusters using pairwise comparisons (One cluster *versus* All other cells). **(D)** The heatmap showed the top 50

genes up-regulated in each cluster. Keys genes coding TFs, EEs, ligands and receptors were indicated and colored in grey, red, green and blue, respectively. **(E)** Expression levels of transcription factors identified in D. High and low expression were represented in dark blue and in yellow, respectively. **(F)** Gene set enrichment analysis of the whole genes up-regulated in each cluster.

Figure 7: Dual Activation of UPR during prePB and PB stages. (A, B) Volcano plots showing differentially expressed genes between the cluster 2 (C2) *versus* clusters 1, 3 et 4 (C1-3-4) and cluster 5 (C5) *versus* cluster 4 (C4), respectively. Genes identified as significantly differentially expressed were colored in blue (*P*-value < 0.05 and log2(Fold Change) > 0.25). **(C)** Venn diagram representing genes involved in UPR and up-regulated in C2 and/or C5. **(D)** Venn diagram of genes up-regulated in C2 and/or C5. Common genes were potentially involved in UPR. **(E, F)** Gene set enrichment analysis showing both pathways enriched in up-regulated and down-regulated genes in C2 and C5 compared to C1-3-4 and C4, respectively. **(G)** Heatmap displaying the expression of genes involved in UPR and up-regulated in C2 and/or C5 for each cluster of proliferating prePB and PB, as well as quiescent MBC, PB and PC. **(H)** Visualization of cells simultaneously co-expressing HSPA5 (in green) and ERN1, EIF2AK3 or ATF6 (in red) genes. Yellow dots correspond to the co-expression of the two genes. **(I)** Boxplots representing the log2 ratio of IGL and IGH read counts per cell in each cluster. **(J)** Plots of HSPA5 and XBP1 mRNA expression in function of pseudotime. **(K)** Violin plots of the main immunoglobulin genes expressed in PCs for each cluster of proliferating prePB and PB, as well as quiescent MBC, PB and PC.

References

1. Shapiro-Shelef M, Calame K. Regulation of plasma-cell development. *Nat Rev Immunol*. 2005;5(3):230–242.
2. Kassambara A, Herviou L, Ovejero S, et al. RNA-sequencing data-driven dissection of human plasma cell differentiation reveals new potential transcription regulators. *Leukemia*. 2021;35(5):1451–1462.
3. Kassambara A, Jourdan M, Bruyer A, et al. Global miRNA expression analysis identifies novel key regulators of plasma cell differentiation and malignant plasma cell. *Nucleic Acids Res*. 2017;45(10):5639–5652.
4. Caron G, Hussein M, Kulis M, et al. Cell-Cycle-Dependent Reconfiguration of the DNA Methylome during Terminal Differentiation of Human B Cells into Plasma Cells. *Cell Rep*. 2015;13(5):1059–1071.
5. Nutt SL, Hodgkin PD, Tarlinton DM, Corcoran LM. The generation of antibody-secreting plasma cells. *Nat Rev Immunol*. 2015;15(3):160–171.
6. Jourdan M, Caraux A, Caron G, et al. Characterization of a Transitional Preplasmablast Population in the Process of Human B Cell to Plasma Cell Differentiation. *The Journal of Immunology*. 2011;187(8):3931–3941.
7. Jourdan M, Caraux A, De Vos J, et al. An in vitro model of differentiation of memory B cells into plasmablasts and plasma cells including detailed phenotypic and molecular characterization. *Blood*. 2009;114(25):5173–5181.
8. Jourdan M, Cren M, Robert N, et al. IL-6 supports the generation of human long-lived plasma cells in combination with either APRIL or stromal cell-soluble factors. *Leukemia*. 2014;28(8):1647–1656.
9. Cocco M, Care MA, Saadi A, et al. A dichotomy of gene regulatory associations during the

activated B-cell to plasmablast transition. *Life Sci Alliance*. 2020;3(10):e202000654.
10. Cocco M, Stephenson S, Care MA, et al. In vitro generation of long-lived human plasma cells. *J Immunol*. 2012;189(12):5773–5785.
11. Haas M, Caron G, Chatonnet F, et al. PIM2 kinase has a pivotal role in plasmablast generation and plasma cell survival, opening up novel treatment options in myeloma. *Blood*. 2022;139(15):2316–2337.
12. Herviou L, Jourdan M, Martinez A-M, Cavalli G, Moreaux J. EZH2 is overexpressed in transitional preplasmablasts and is involved in human plasma cell differentiation. *Leukemia*. 2019;
13. Xiong E, Popp O, Salomon C, et al. A CRISPR/Cas9-mediated screen identifies determinants of early plasma cell differentiation. *Front Immunol*. 2022;13:1083119.
14. Pignarre A, Chatonnet F, Caron G, et al. Plasmablasts derive from CD23- activated B cells after the extinction of IL-4/STAT6 signaling and IRF4 induction. *Blood*. 2021;137(9):1166–1180.
15. Kassambara A, Rème T, Jourdan M, et al. GenomicScape: an easy-to-use web tool for gene expression data analysis. Application to investigate the molecular events in the differentiation of B cells into plasma cells. *PLoS Comput Biol*. 2015;11(1):e1004077.
16. Iwakoshi NN, Lee A-H, Vallabhajosyula P, et al. Plasma cell differentiation and the unfolded protein response intersect at the transcription factor XBP-1. *Nat Immunol*. 2003;4(4):321–329.
17. Sheng Z, Li L, Zhu LJ, et al. A genome-wide RNA interference screen reveals an essential CREB3L2-ATF5-MCL1 survival pathway in malignant glioma with therapeutic implications. *Nat Med*. 2010;16(6):671–677.
18. Santamaria K, Desmots F, Leonard S, et al. Committed Human CD23-Negative Light-Zone Germinal Center B Cells Delineate Transcriptional Program Supporting Plasma Cell Differentiation. *Front Immunol*. 2021;12:744573.
19. Duan M, Nguyen DC, Joyner CJ, et al. Understanding heterogeneity of human bone marrow plasma cell maturation and survival pathways by single-cell analyses. *Cell Rep*. 2023;42(7):112682.
20. Fucci C, Resnati M, Riva E, et al. The Interaction of the Tumor Suppressor FAM46C with p62 and FNDC3 Proteins Integrates Protein and Secretory Homeostasis. *Cell Rep*. 2020;32(12):108162.
21. Zhang Y, Liu T, Meyer CA, et al. Model-based Analysis of ChIP-Seq (MACS). *Genome Biology*. 2008;9(9):R137.
22. Andreani V, Ramamoorthy S, Pandey A, et al. Cochaperone Mzb1 is a key effector of Blimp1 in plasma cell differentiation and β1-integrin function. *Proc Natl Acad Sci U S A*. 2018;115(41):E9630–E9639.
23. Grundström C, Kumar A, Priya A, Negi N, Grundström T. ETS1 and PAX5 transcription factors recruit AID to Igh DNA. *Eur J Immunol*. 2018;48(10):1687–1697.
24. Oh Y-K, Jang E, Paik D-J, Youn J. Early Growth Response-1 Plays a Non-redundant Role in the Differentiation of B Cells into Plasma Cells. *Immune Netw*. 2015;15(3):161–166.
25. Winkelmann R, Sandrock L, Porstner M, et al. B cell homeostasis and plasma cell homing controlled by Krüppel-like factor 2. *Proc Natl Acad Sci U S A*. 2011;108(2):710–715.
26. Gaudette BT, Jones DD, Bortnick A, Argon Y, Allman D. mTORC1 coordinates an immediate unfolded protein response-related transcriptome in activated B cells preceding antibody secretion. *Nat Commun*. 2020;11(1):723.
27. Haas IG, Wabl M. Immunoglobulin heavy chain binding protein. *Nature*. 1983;306(5941):387–389.
28. Yoshida H, Matsui T, Yamamoto A, Okada T, Mori K. XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. *Cell*. 2001;107(7):881–891.
29. Massoni-Badosa R, Aguilar-Fernández S, Nieto JC, et al. An atlas of cells in the human tonsil. *Immunity*. 2024;57(2):379-399.e18.
30. Caganova M, Carrisi C, Varano G, et al. Germinal center dysregulation by histone methyltransferase EZH2 promotes lymphomagenesis. *J Clin Invest*. 2013;123(12):5009–5022.
31. Benhamron S, Pattanayak SP, Berger M, Tirosh B. mTOR activation promotes plasma cell differentiation and bypasses XBP-1 for immunoglobulin secretion. *Mol Cell Biol*. 2015;35(1):153–166.
32. Jones DD, Gaudette BT, Wilmore JR, et al. mTOR has distinct functions in generating versus sustaining humoral immunity. *J Clin Invest*. 2016;126(11):4250–4261.

33. Lemarié M, Chatonnet F, Caron G, Fest T. Early Emergence of Adaptive Mechanisms Sustaining Ig Production: Application to Antibody Therapy. *Front Immunol*. 2021;12:671998.
34. Ohkubo Y, Arima M, Arguni E, et al. A role for c-fos/activator protein 1 in B lymphocyte terminal differentiation. *J Immunol*. 2005;174(12):7703–7710.
35. Fan F, Bashari MH, Morelli E, et al. The AP-1 transcription factor JunB is essential for multiple myeloma cell proliferation and drug resistance in the bone marrow microenvironment. *Leukemia*. 2017;31(7):1570–1581.
36. Biajoux V, Natt J, Freitas C, et al. Efficient Plasma Cell Differentiation and Trafficking Require Cxcr4 Desensitization. *Cell Rep*. 2016;17(1):193–205.
37. Fooksman DR, Schwickert TA, Victora GD, et al. Development and migration of plasma cells in the mouse lymph node. *Immunity*. 2010;33(1):118–127.
38. Roth K, Oehme L, Zehentmeier S, et al. Tracking plasma cell differentiation and survival. *Cytometry A*. 2014;85(1):15–24.
39. Bernasconi NL, Traggiai E, Lanzavecchia A. Maintenance of Serological Memory by Polyclonal Activation of Human Memory B Cells. *Science*. 2002;298(5601):2199–2202.
40. Verstegen NJ, Pollastro S, Unger P-PA, et al. Single-cell analysis reveals dynamics of human B cell differentiation and identifies novel B and antibody-secreting cell intermediates. *eLife*. 12:e83578.
41. van Buijtenen E, Janssen W, Vink P, et al. Integrated Single-Cell (Phospho-)Protein and RNA Detection Uncovers Phenotypic Characteristics and Active Signal Transduction of Human Antibody-Secreting Cells. *Mol Cell Proteomics*. 2023;22(2):100492.
42. Caeser R, Gao J, Di Re M, Gong C, Hodson DJ. Genetic manipulation and immortalized culture of ex vivo primary human germinal center B cells. *Nat Protoc*. 2021;16(5):2499–2519.
43. Fulciniti M, Amin S, Nanjappa P, et al. Significant biological role of sp1 transactivation in multiple myeloma. *Clin Cancer Res*. 2011;17(20):6500–6509.
44. Gu Z, Wang H, Xia J, et al. Decreased ferroportin promotes myeloma cell growth and osteoclast differentiation. *Cancer Res*. 2015;75(11):2211–2221.
45. Ray A, Song Y, Du T, Chauhan D, Anderson KC. Preclinical validation of Alpha-Enolase (ENO1) as a novel immunometabolic target in multiple myeloma. *Oncogene*. 2020;39(13):2786–2796.
46. Tai Y-T, Podar K, Mitsiades N, et al. CD40 induces human multiple myeloma cell migration via phosphatidylinositol 3–kinase/AKT/NF-κB signaling. *Blood*. 2003;101(7):2762–2769.
47. Tohami T, Drucker L, Shapiro H, Radnay J, Lishner M. Overexpression of tetraspanins affects multiple myeloma cell survival and invasive potential. *FASEB J*. 2007;21(3):691–699.
48. Ferguson ID, Patiño-Escobar B, Tuomivaara ST, et al. The surfaceome of multiple myeloma cells suggests potential immunotherapeutic strategies and protein markers of drug resistance. *Nat Commun*. 2022;13(1):4121.
49. Chu B, Bao L, Wang Y, et al. CD27 antigen negative expression indicates poor prognosis in newly diagnosed multiple myeloma. *Clin Immunol*. 2020;213:108363.

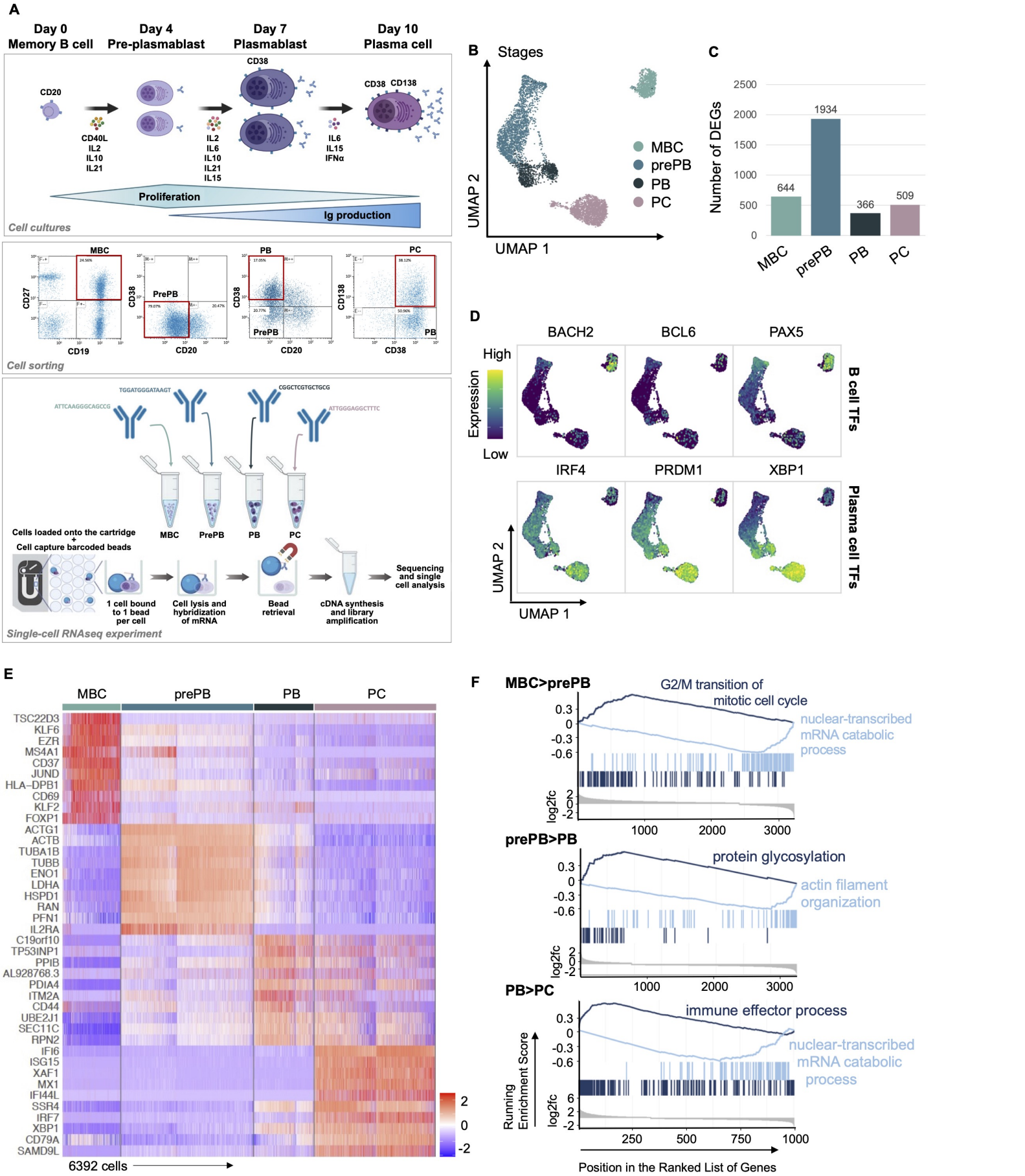

A
Day 0
Memory B cell
Day 4
Pre-plasmablast
Day 7
Plasmablast
Day 10
Plasma cell
CD20
CD40L
IL2
IL10
IL21
CD38
IL2
IL6
IL10
IL21
IL15
CD38 CD138
IL6
IL15
IFNα
Proliferation
Ig production
Cell cultures
MBC
PrePB
PB
PC
CD27
CD19
CD38
CD20
CD138
CD38
Cell sorting
TGGATGGGATAAGT
CGGCTCGTGCTGCG
ATTCAAGGGCAGCCG
ATTGGGAGGCTTTC
MBC
PrePB
PB
PC
Cells loaded onto the cartridge
+
Cell capture barcoded beads
1 cell bound to 1 bead per cell
Cell lysis and hybridization of mRNA
Bead retrieval
cDNA synthesis and library amplification
Sequencing and single cell analysis
Single-cell RNAseq experiment
B
Stages
MBC
prePB
PB
PC
UMAP 2
UMAP 1
C
Number of DEGs
2500
2000
1500
1000
500
0
644
1934
366
509
MBC
prePB
PB
PC
D
BACH2
BCL6
PAX5
High
Expression
Low
B cell TFs
IRF4
PRDM1
XBP1
Plasma cell TFs
UMAP 2
UMAP 1
E
MBC
prePB
PB
PC
TSC22D3
KLF6
EZR
MS4A1
CD37
JUND
HLA–DPB1
CD69
KLF2
FOXP1
ACTG1
ACTB
TUBA1B
TUBB
ENO1
LDHA
HSPD1
RAN
PFN1
IL2RA
C19orf10
TP53INP1
PPIB
AL928768.3
PDIA4
ITM2A
CD44
UBE2J1
SEC11C
RPN2
IFI6
ISG15
XAF1
MX1
IFI44L
SSR4
IRF7
XBP1
CD79A
SAMD9L
2
0
-2
6392 cells
F
MBC>prePB
G2/M transition of mitotic cell cycle
nuclear-transcribed mRNA catabolic process
0.3
0
-0.3
-0.6
log2fc
2
0
-2
1000
2000
3000
prePB>PB
protein glycosylation
actin filament organization
PB>PC
immune effector process
nuclear-transcribed mRNA catabolic process
Running Enrichment Score
6
250
500
750
1000
Position in the Ranked List of Genes

Figure 1.

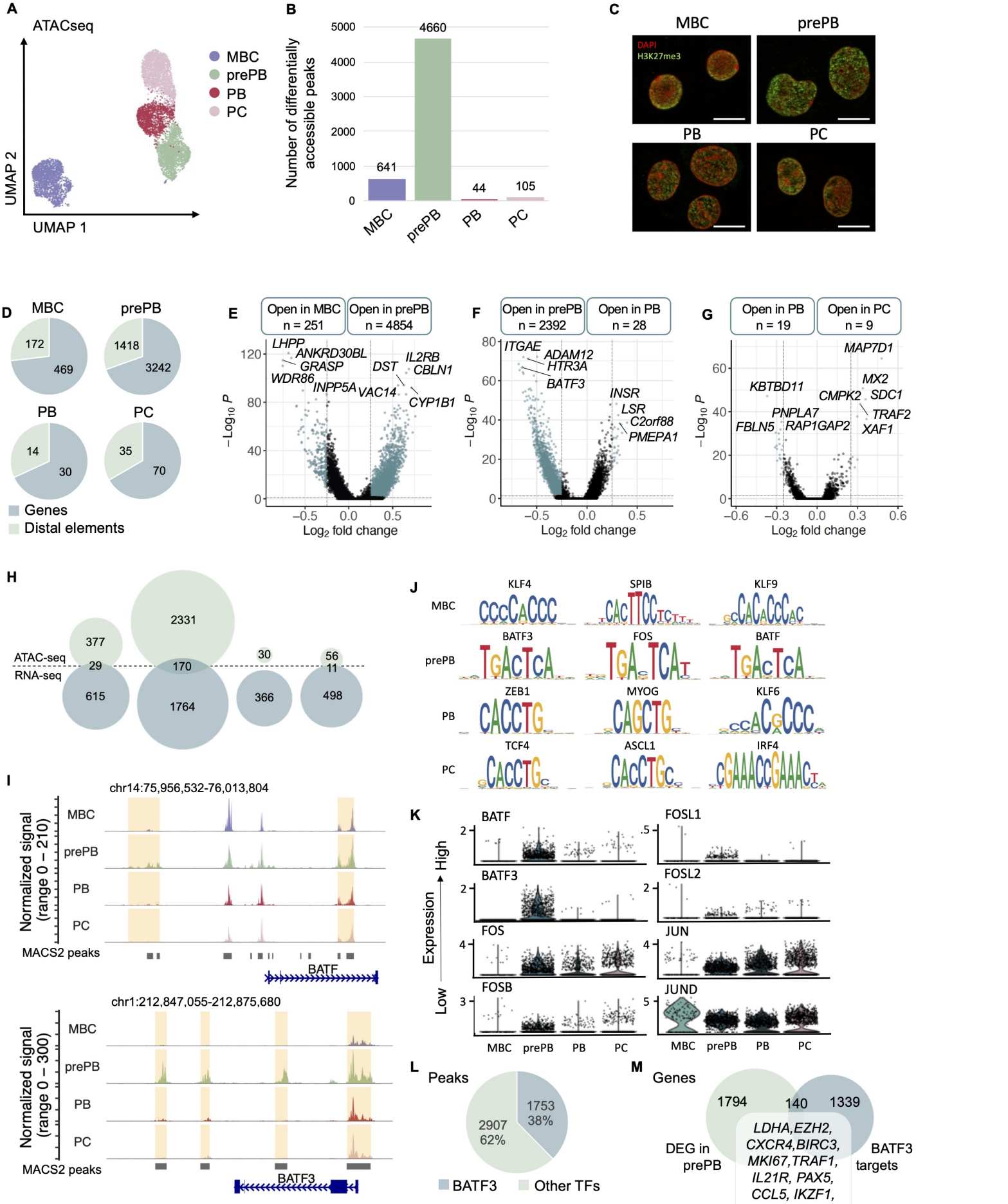

A
ATACseq
MBC
prePB
PB
PC
UMAP 2
UMAP 1
B
Number of differentially accessible peaks
5000
4000
3000
2000
1000
0
641
4660
44
105
MBC
prePB
PB
PC
C
MBC
prePB
DAPI
H3K27me3
PB
PC
D
MBC
prePB
172
469
1418
3242
PB
PC
14
30
35
70
Genes
Distal elements
E
Open in MBC
n = 251
Open in prePB
n = 4854
LHPP
ANKRD30BL
GRASP
WDR86
INPP5A
DST
VAC14
IL2RB
CBLN1
CYP1B1
−Log10 P
Log2 fold change
F
Open in prePB
n = 2392
Open in PB
n = 28
ITGAE
ADAM12
HTR3A
BATF3
INSR
LSR
C2orf88
PMEPA1
−Log10 P
Log2 fold change
G
Open in PB
n = 19
Open in PC
n = 9
MAP7D1
KBTBD11
MX2
CMPK2
SDC1
PNPLA7
TRAF2
FBLN5
RAP1GAP2
XAF1
−Log10 P
Log2 fold change
H
2331
377
ATAC-seq
29
170
30
56
11
RNA-seq
615
1764
366
498
I
chr14:75,956,532-76,013,804
Normalized signal (range 0 – 210)
MBC
prePB
PB
PC
MACS2 peaks
BATF
chr1:212,847,055-212,875,680
Normalized signal (range 0 – 300)
MBC
prePB
PB
PC
MACS2 peaks
BATF3
J
MBC
KLF4
SPIB
KLF9
prePB
BATF3
FOS
BATF
PB
ZEB1
MYOG
KLF6
PC
TCF4
ASCL1
IRF4
K
Expression
High
Low
BATF
FOSL1
BATF3
FOSL2
FOS
JUN
FOSB
JUND
MBC
prePB
PB
PC
L
Peaks
1753
38%
2907
62%
BATF3
Other TFs
M
Genes
1794
140
1339
DEG in prePB
BATF3 targets
LDHA,EZH2, CXCR4,BIRC3, MKI67,TRAF1, IL21R, PAX5, CCL5, IKZF1, IRF5, CCR7, …

Figure 2.

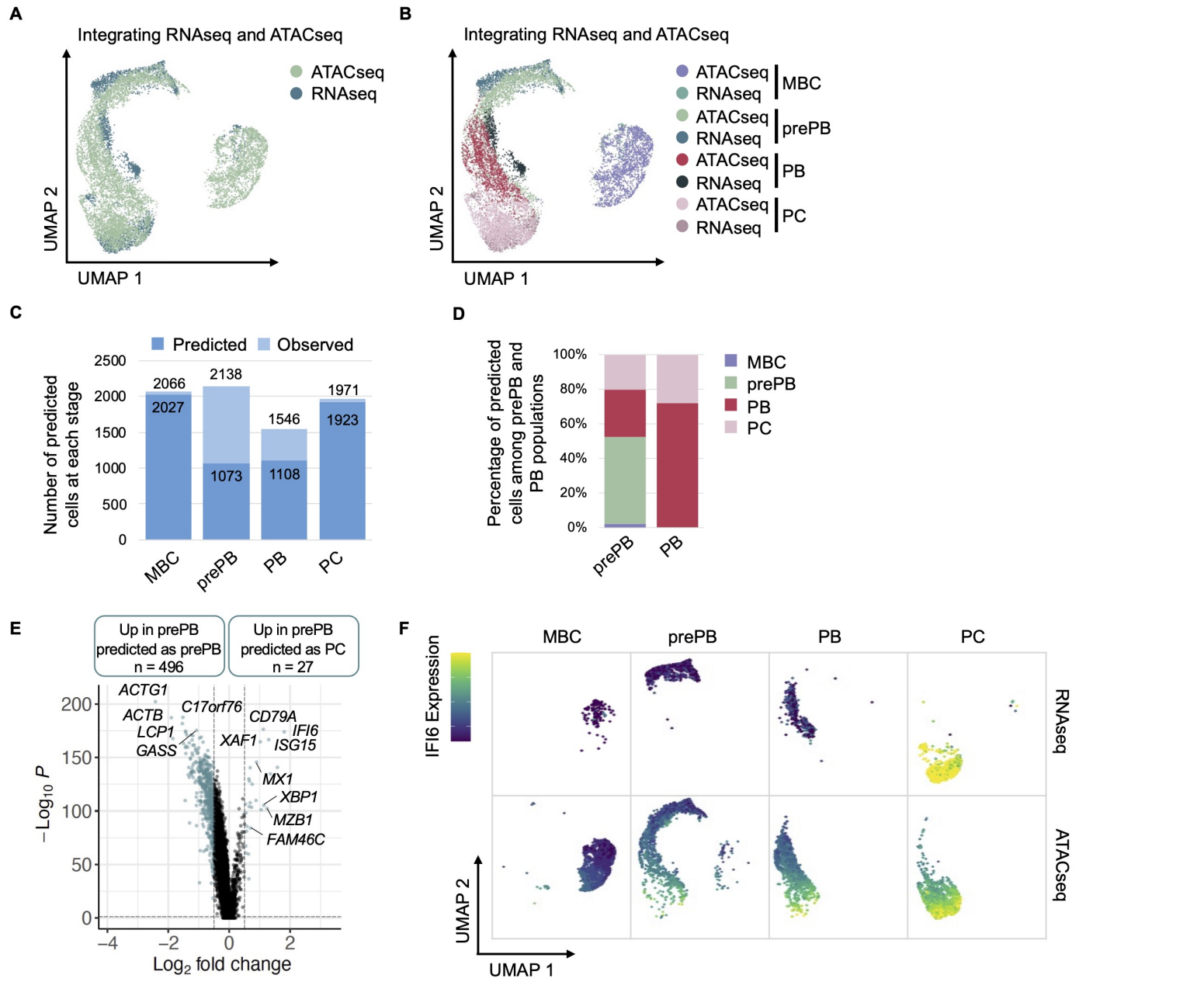

A
Integrating RNAseq and ATACseq
ATACseq
RNAseq
UMAP 2
UMAP 1
B
Integrating RNAseq and ATACseq
ATACseq
RNAseq
MBC
prePB
PB
PC
UMAP 2
UMAP 1
C
Predicted
Observed
Number of predicted cells at each stage
2066
2027
2138
1073
1546
1108
1971
1923
MBC
prePB
PB
PC
D
Percentage of predicted cells among prePB and PB populations
MBC
prePB
PB
PC
prePB
PB
E
Up in prePB predicted as prePB n = 496
Up in prePB predicted as PC n = 27
ACTG1
ACTB
LCP1
GASS
C17orf76
CD79A
XAF1
IFI6
ISG15
MX1
XBP1
MZB1
FAM46C
−Log10 P
Log2 fold change
F
MBC
prePB
PB
PC
IFI6 Expression
RNAseq
ATACseq
UMAP 2
UMAP 1

Figure 3.

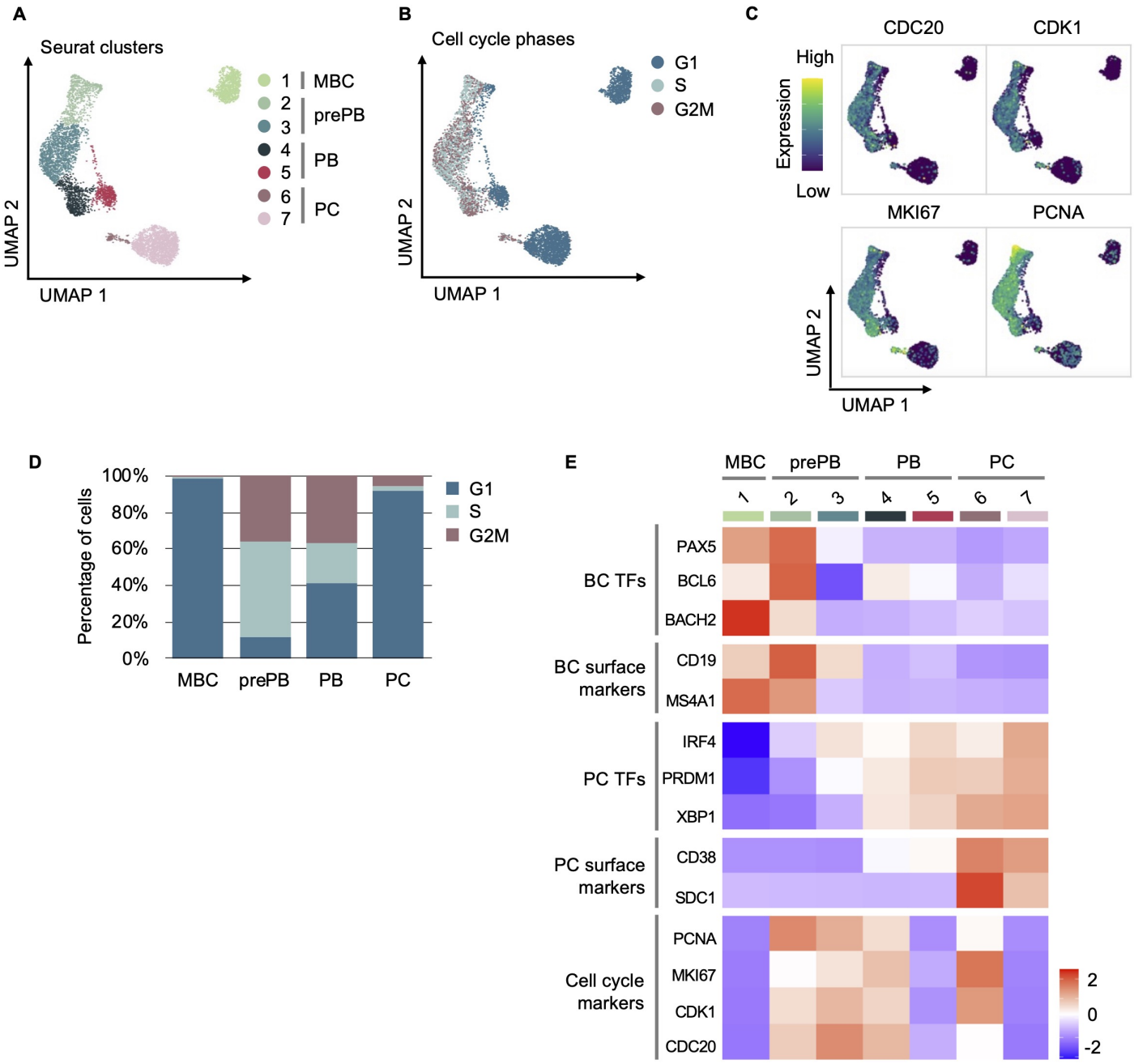

A
Seurat clusters
1
2
3
4
5
6
7
MBC
prePB
PB
PC
UMAP 2
UMAP 1
B
Cell cycle phases
G1
S
G2M
UMAP 2
UMAP 1
C
CDC20
CDK1
High
Expression
Low
MKI67
PCNA
UMAP 2
UMAP 1
D
Percentage of cells
100%
80%
60%
40%
20%
0%
MBC
prePB
PB
PC
G1
S
G2M
E
MBC
prePB
PB
PC
1
2
3
4
5
6
7
BC TFs
PAX5
BCL6
BACH2
BC surface markers
CD19
MS4A1
PC TFs
IRF4
PRDM1
XBP1
PC surface markers
CD38
SDC1
Cell cycle markers
PCNA
MKI67
CDK1
CDC20
2
0
-2

Figure 4.

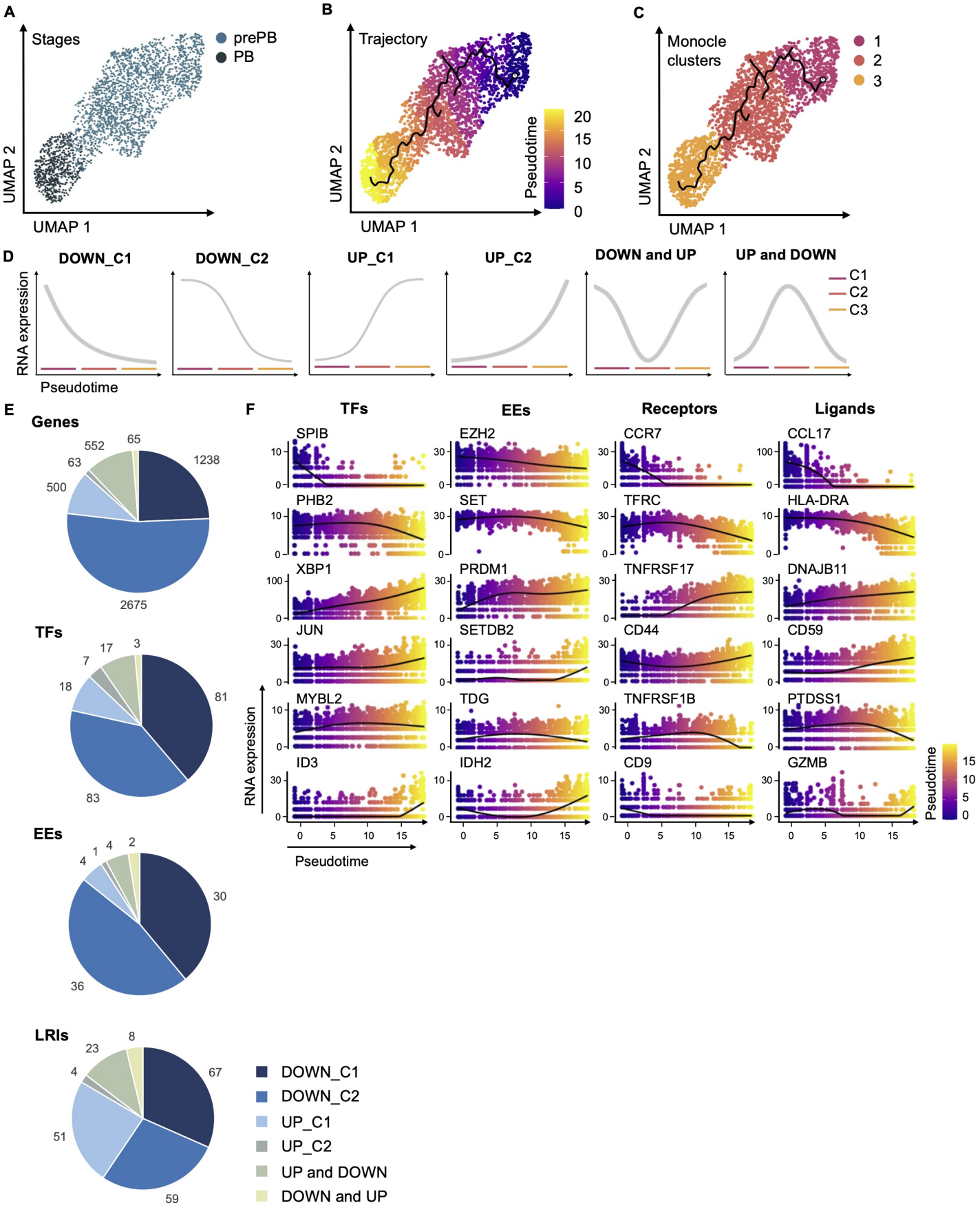

A
Stages
prePB
PB
UMAP 2
UMAP 1
B
Trajectory
Pseudotime
20
15
10
5
0
C
Monocle clusters
1
2
3
D
DOWN_C1
DOWN_C2
UP_C1
UP_C2
DOWN and UP
UP and DOWN
RNA expression
Pseudotime
C1
C2
C3
E
Genes
1238
2675
500
63
552
65
TFs
81
83
18
7
17
3
EEs
30
36
4
1
4
2
LRIs
67
51
4
23
8
DOWN_C1
DOWN_C2
UP_C1
UP_C2
UP and DOWN
DOWN and UP
F
TFs
EEs
Receptors
Ligands
SPIB
EZH2
CCR7
CCL17
PHB2
SET
TFRC
HLA-DRA
XBP1
PRDM1
TNFRSF17
DNAJB11
JUN
SETDB2
CD44
CD59
MYBL2
TDG
TNFRSF1B
PTDSS1
ID3
IDH2
CD9
GZMB
RNA expression
Pseudotime
Pseudotime

Figure 5.

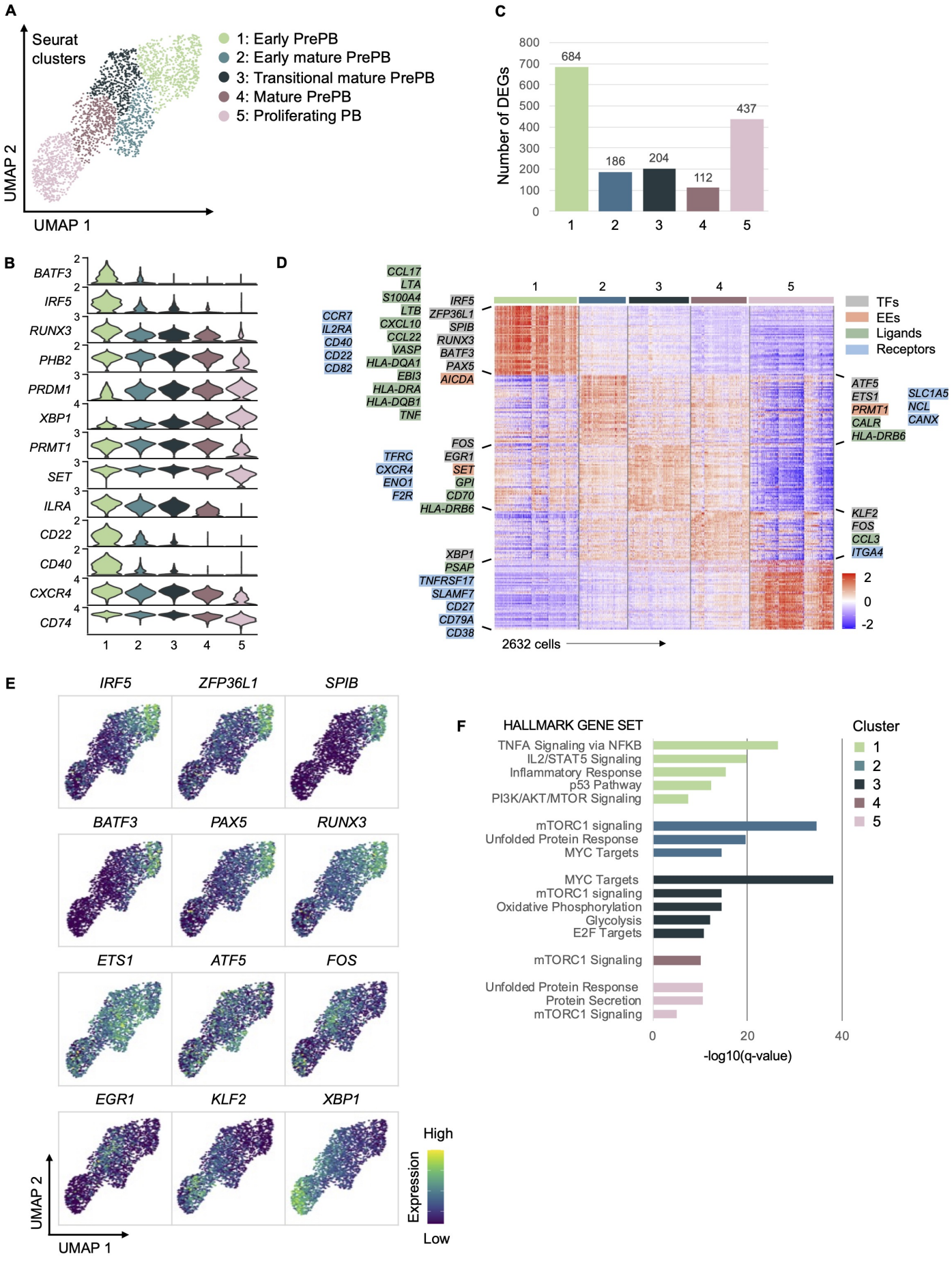

A
Seurat clusters
1: Early PrePB
2: Early mature PrePB
3: Transitional mature PrePB
4: Mature PrePB
5: Proliferating PB
UMAP 2
UMAP 1
C
Number of DEGs
684
186
204
112
437
B
BATF3
IRF5
RUNX3
PHB2
PRDM1
XBP1
PRMT1
SET
ILRA
CD22
CD40
CXCR4
CD74
D
TFs
EEs
Ligands
Receptors
2632 cells
E
IRF5
ZFP36L1
SPIB
BATF3
PAX5
RUNX3
ETS1
ATF5
FOS
EGR1
KLF2
XBP1
High
Expression
Low
F
HALLMARK GENE SET
TNFA Signaling via NFKB
IL2/STAT5 Signaling
Inflammatory Response
p53 Pathway
PI3K/AKT/MTOR Signaling
mTORC1 signaling
Unfolded Protein Response
MYC Targets
MYC Targets
mTORC1 signaling
Oxidative Phosphorylation
Glycolysis
E2F Targets
mTORC1 Signaling
Unfolded Protein Response
Protein Secretion
mTORC1 Signaling
-log10(q-value)
Cluster

Figure 6

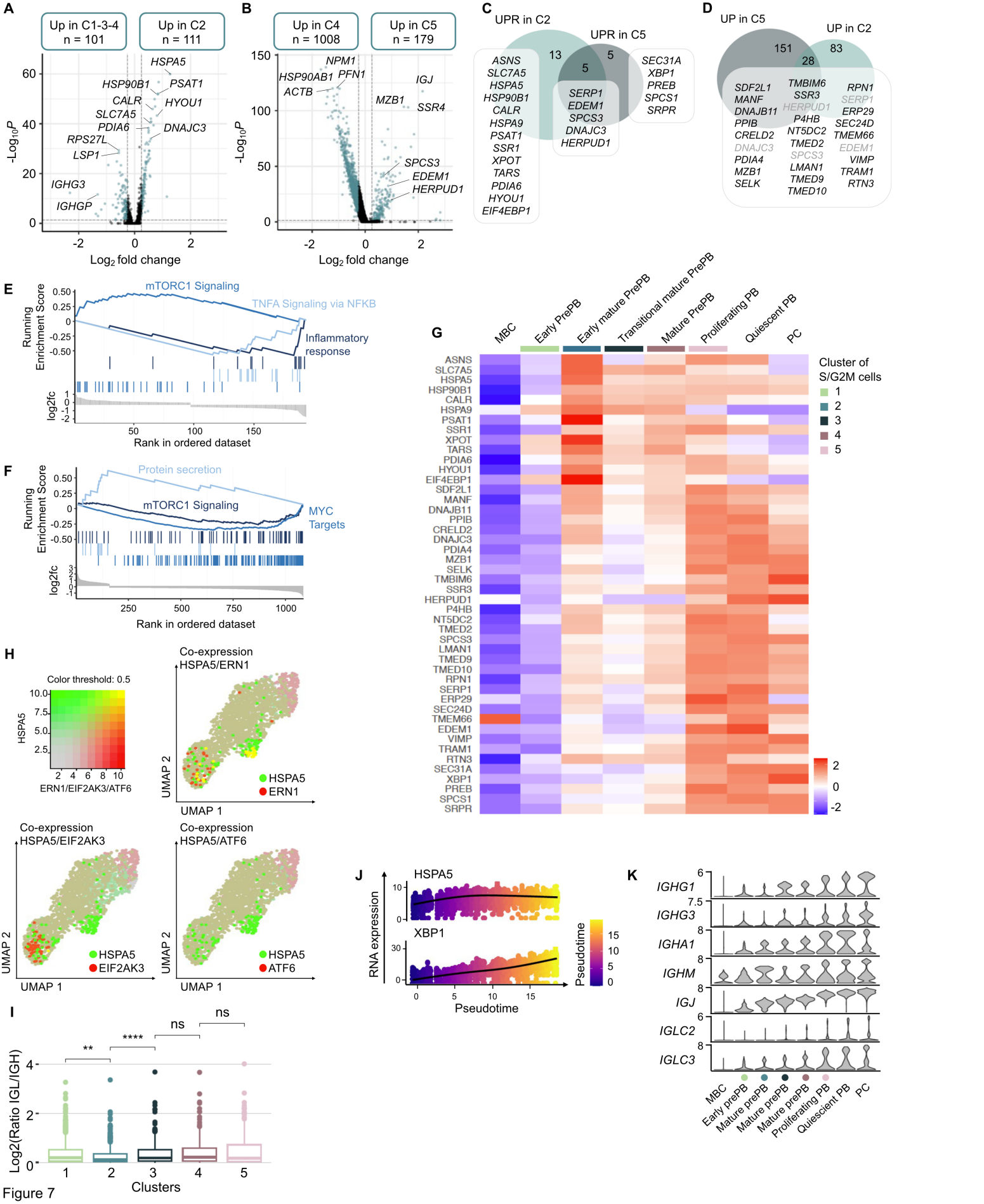
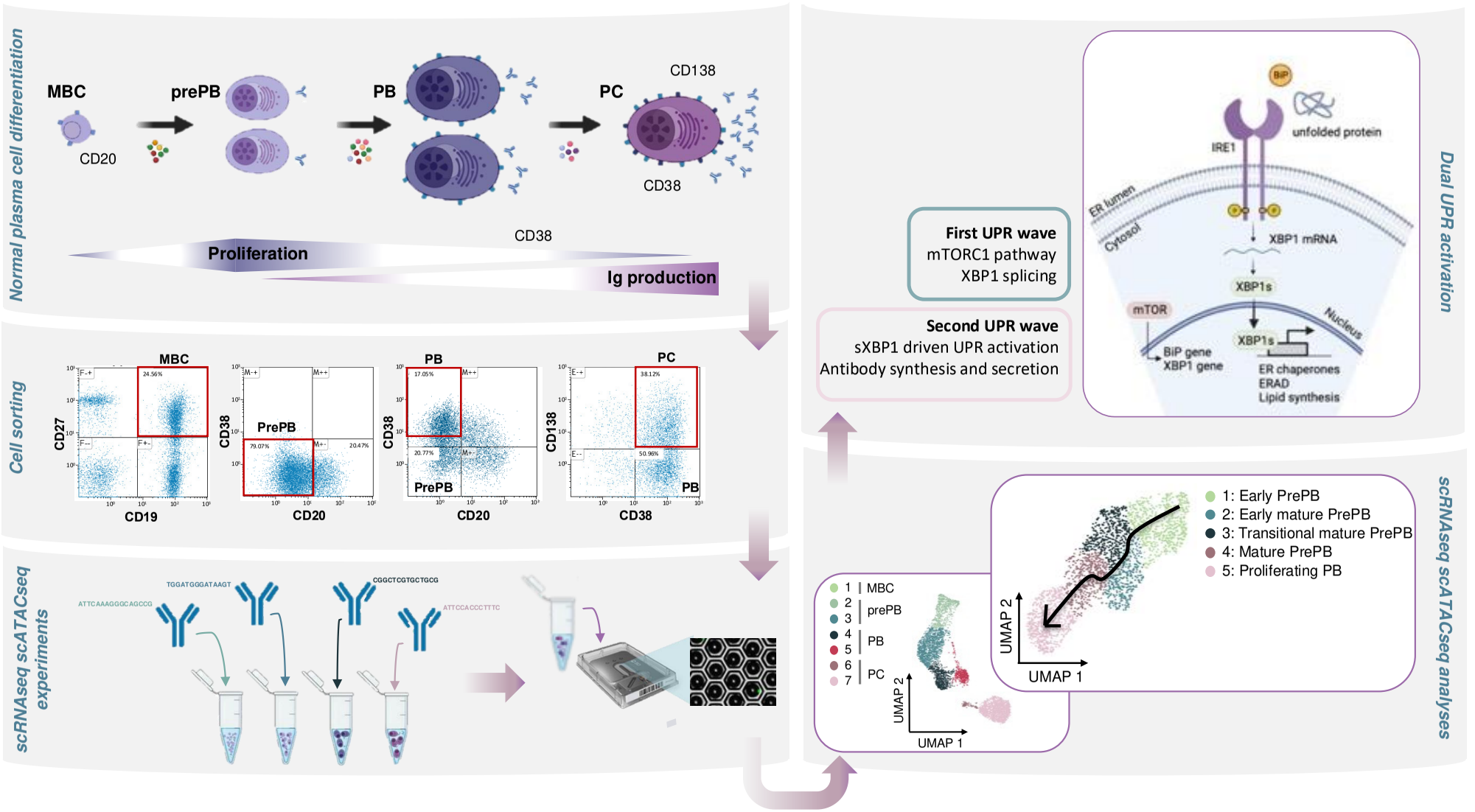
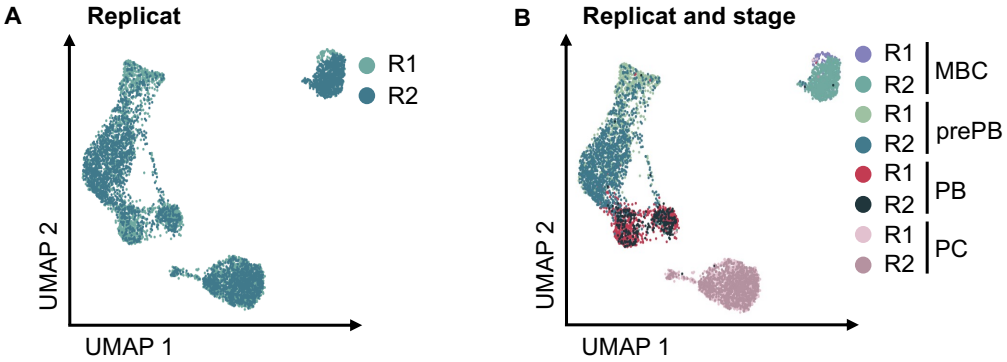

A
Up in C1-3-4 n = 101
Up in C2 n = 111
HSPA5
HSP90B1
PSAT1
CALR
HYOU1
SLC7A5
PDIA6
DNAJC3
RPS27L
LSP1
IGHG3
IGHGP
-Log10P
Log2 fold change
B
Up in C4 n = 1008
Up in C5 n = 179
NPM1
HSP90AB1
PFN1
ACTB
IGJ
MZB1
SSR4
SPCS3
EDEM1
HERPUD1
C
UPR in C2
UPR in C5
ASNS
SLC7A5
HSPA5
HSP90B1
CALR
HSPA9
PSAT1
SSR1
XPOT
TARS
PDIA6
HYOU1
EIF4EBP1
SERP1
EDEM1
SPCS3
DNAJC3
HERPUD1
SEC31A
XBP1
PREB
SPCS1
SRPR
D
UP in C5
UP in C2
SDF2L1
MANF
DNAJB11
PPIB
CRELD2
DNAJC3
PDIA4
MZB1
SELK
TMBIM6
SSR3
HERPUD1
P4HB
NT5DC2
TMED2
SPCS3
LMAN1
TMED9
TMED10
RPN1
SERP1
ERP29
SEC24D
TMEM66
EDEM1
VIMP
TRAM1
RTN3
E
mTORC1 Signaling
TNFA Signaling via NFKB
Inflammatory response
Running Enrichment Score
log2fc
Rank in ordered dataset
F
Protein secretion
mTORC1 Signaling
MYC Targets
Running Enrichment Score
log2fc
Rank in ordered dataset
G
MBC
Early PrePB
Early mature PrePB
Transitional mature PrePB
Mature PrePB
Proliferating PB
Quiescent PB
PC
Cluster of S/G2M cells
H
Color threshold: 0.5
HSPA5
ERN1/EIF2AK3/ATF6
Co-expression HSPA5/ERN1
Co-expression HSPA5/EIF2AK3
Co-expression HSPA5/ATF6
UMAP 1
UMAP 2
HSPA5
ERN1
EIF2AK3
ATF6
I
Log2(Ratio IGL/IGH)
Clusters
ns
****
**
J
HSPA5
XBP1
RNA expression
Pseudotime
K
IGHG1
IGHG3
IGHA1
IGHM
IGJ
IGLC2
IGLC3
MBC
Early prePB
Mature preBP
Mature preBP
Mature preBP
Proliferating PB
Quiescient PB
PC

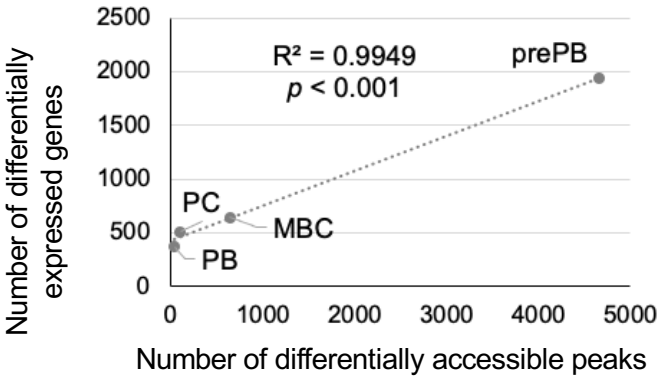
Figure 7


Normal plasma cell differentiation
MBC
CD20
prePB
PB
PC
CD138
CD38
Proliferation
CD38
Ig production
Cell sorting
MBC
CD27
CD19
PrePB
CD38
CD20
PB
PrePB
CD38
CD20
PC
CD138
CD38
PB
scRNAseq scATACseq experiments
TGGATGGGATAAGT
ATTCAAAGGGCAGCCG
CGGCTCGTGCTGCG
ATTCCACCCTTTC
First UPR wave
mTORC1 pathway
XBP1 splicing
Second UPR wave
sXBP1 driven UPR activation
Antibody synthesis and secretion
BiP
IRE1
unfolded protein
ER lumen
Cytosol
XBP1 mRNA
XBP1s
mTOR
Nucleus
XBP1s
BiP gene
XBP1 gene
ER chaperones
ERAD
Lipid synthesis
Dual UPR activation
1 MBC
2
3 prePB
4
5 PB
6
7 PC
UMAP 2
UMAP 1
1: Early PrePB
2: Early mature PrePB
3: Transitional mature PrePB
4: Mature PrePB
5: Proliferating PB
UMAP 2
UMAP 1
scRNAseq scATACseq analyses

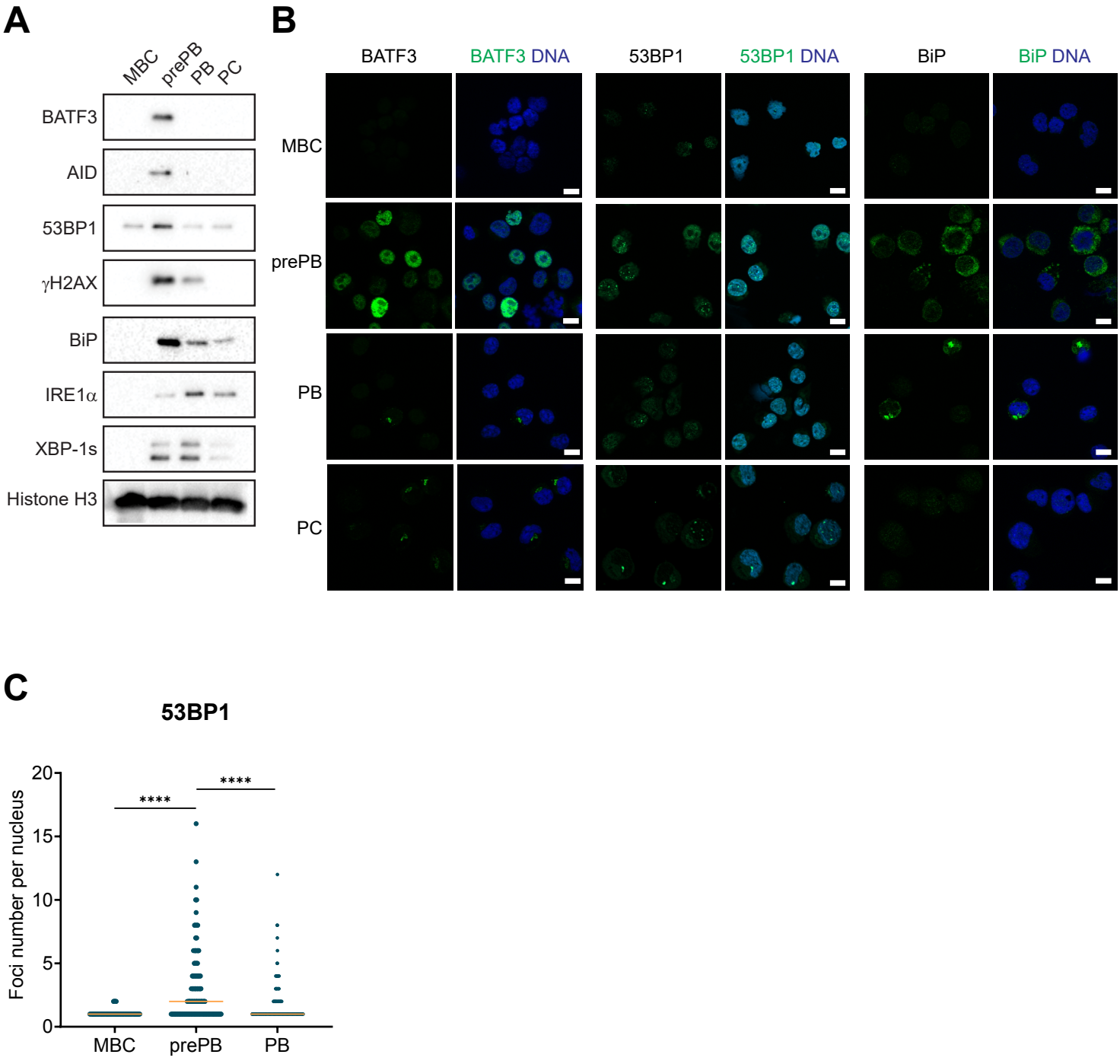
Supplemental Figure 1. UMAP projection of sc-RNAseq of the B to plasma cell differentiation using memory B cells from two healthy donors. This representation was obtained after batch effect correction. (A) UMAP projection was colored according to the replicate. (B) UMAP projection was according to the stage and the replicate.



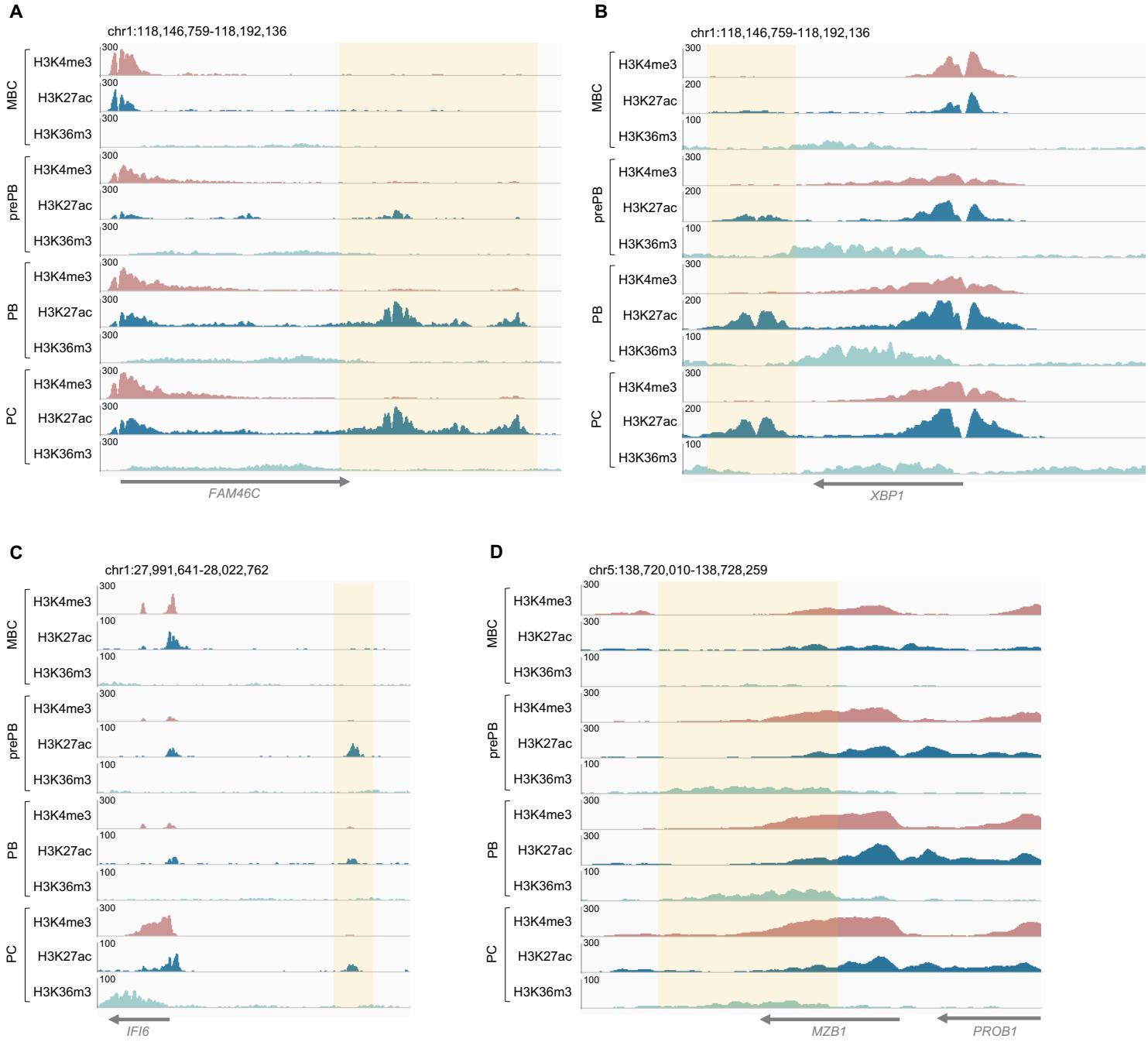
Supplemental Figure 2. Correlation between the number of differentially expressed genes identified using sc-RNAseq dataset and the number of differentially accessible peaks identified using sc-ATACseq dataset. Coefficient of determination R2 represents the square of the Pearson correlation coefficient (r) (Pearson correlation test).



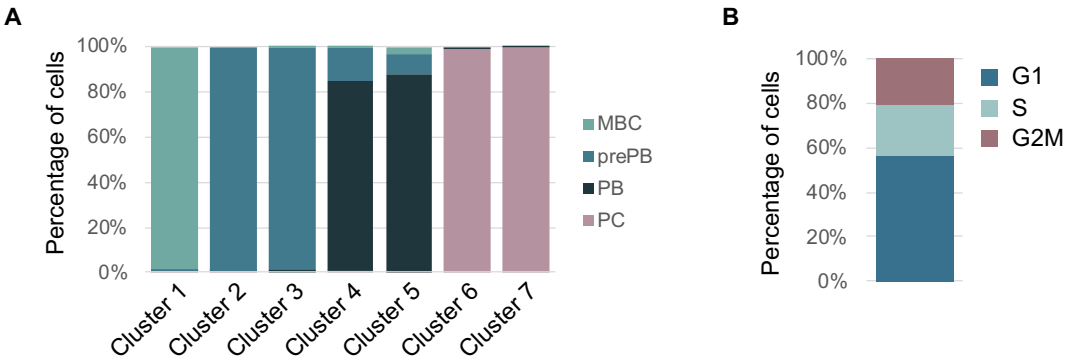
Supplemental Figure 3. Cell populations were FACS sorted at each time point (MBC: Day 0, prePB: Day 4; PB: Day 7; PC: Day 10) and the indicated proteins were analyzed by western blotting with specific antibodies. Histone H3 was used as loading control. Results show one representative out of 3 independent experiments (A). Cells were FACS sorted at each time point as explained before, fixed with 4% PFA-PBS (10 min) and immunofluorescence was performed to detect BATF3 and BiP proteins. DNA was stained with DAPI. Scale bars: 10 µm. Results show one representative out of 3 independent experiments (B). Quantification of 53BP1 foci in MBC, PrePB and PB.**** p value < 0.0001.



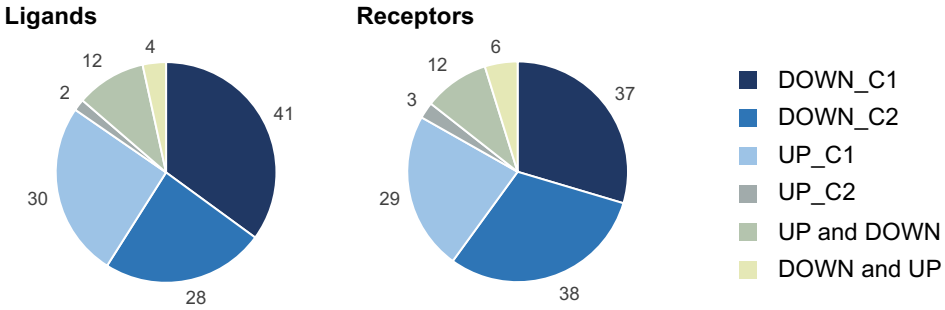
Supplemental Figure 4. Gene tracks of H3K4me3, H3K27ac and H3K36me3 ChIP-seq occupancy near PC genes in MBC, prePB, PB and PC stages. PC genes already demonstrated an enrichment at enhancer region (H3K27ac mark) for *FAM46C* (A), *XBP1* (B) and *IFI6* (C) and in the gene body (H3K36me3 mark) of *MZB1* (D) in prePB stage. The x axis shows the genomic position. The y axis shows signal coverage of ChIP-seq occupancy in units of reads per bin mapped reads.



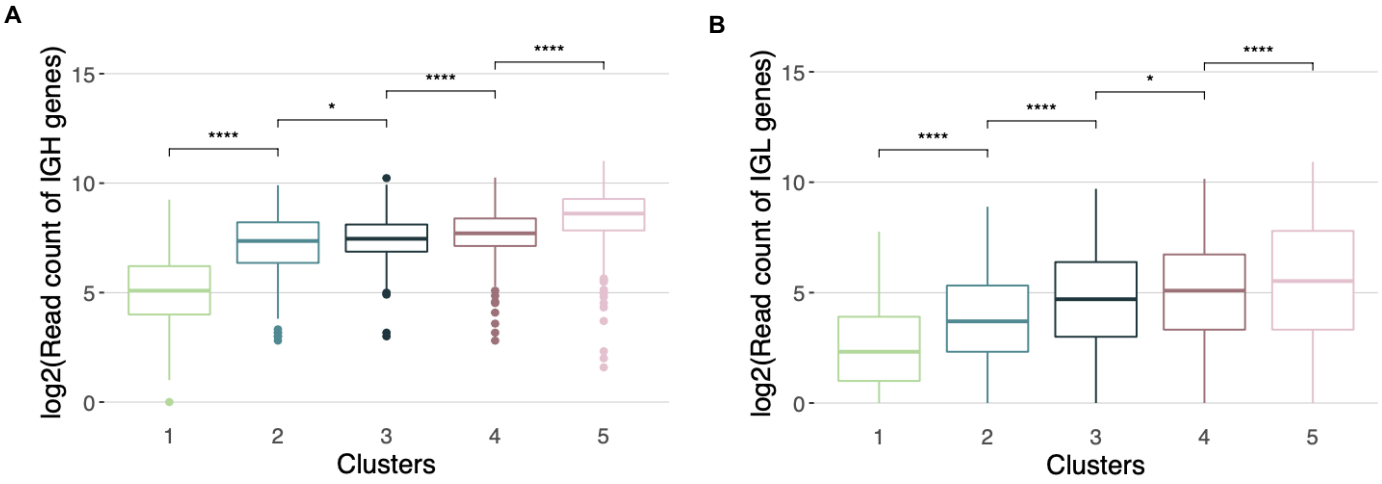
Supplemental Figure 5. Distribution of cells. (A) Distribution of MBC, prePB, PB and PC in each cluster identified in Figure 4A. (B) Percentage of quiescent (G1) and proliferative (S and G2M) cells among all cells.



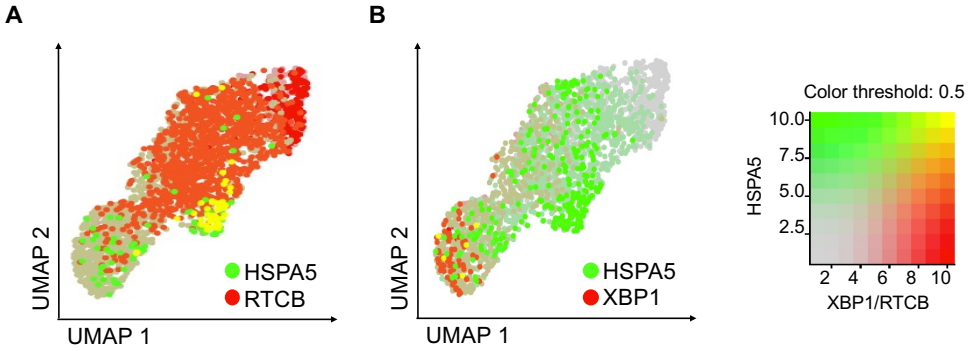
Supplemental Figure 6. Proportion of genes coding ligands and receptors deregulated along the trajectory according to the expression patterns defined in Figure 5D.



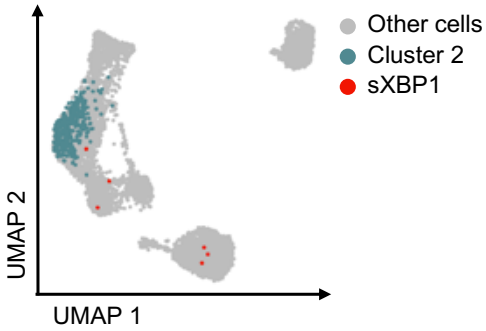
Supplemental Figure 7. Boxplots representing the log2 of read count of IGH (A) and IGL (B) genes per cell in each cluster.



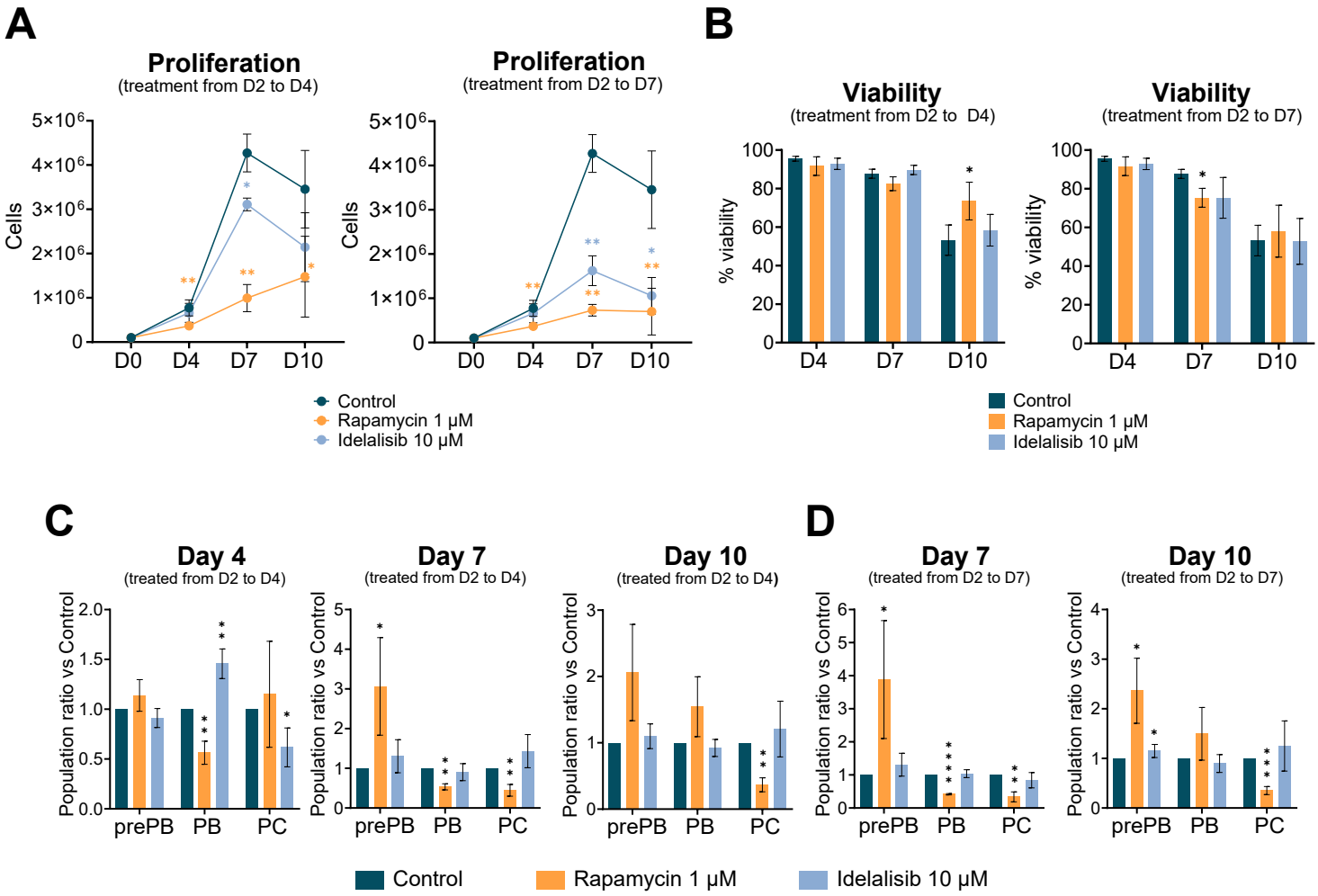
Supplemental Figure 8. Visualization of cells simultaneously co-expressing HSPA5 (in green) and RTCB (A) or XBP1 (B) (in red) genes. Yellow dots correspond to the co-expression of the two genes.



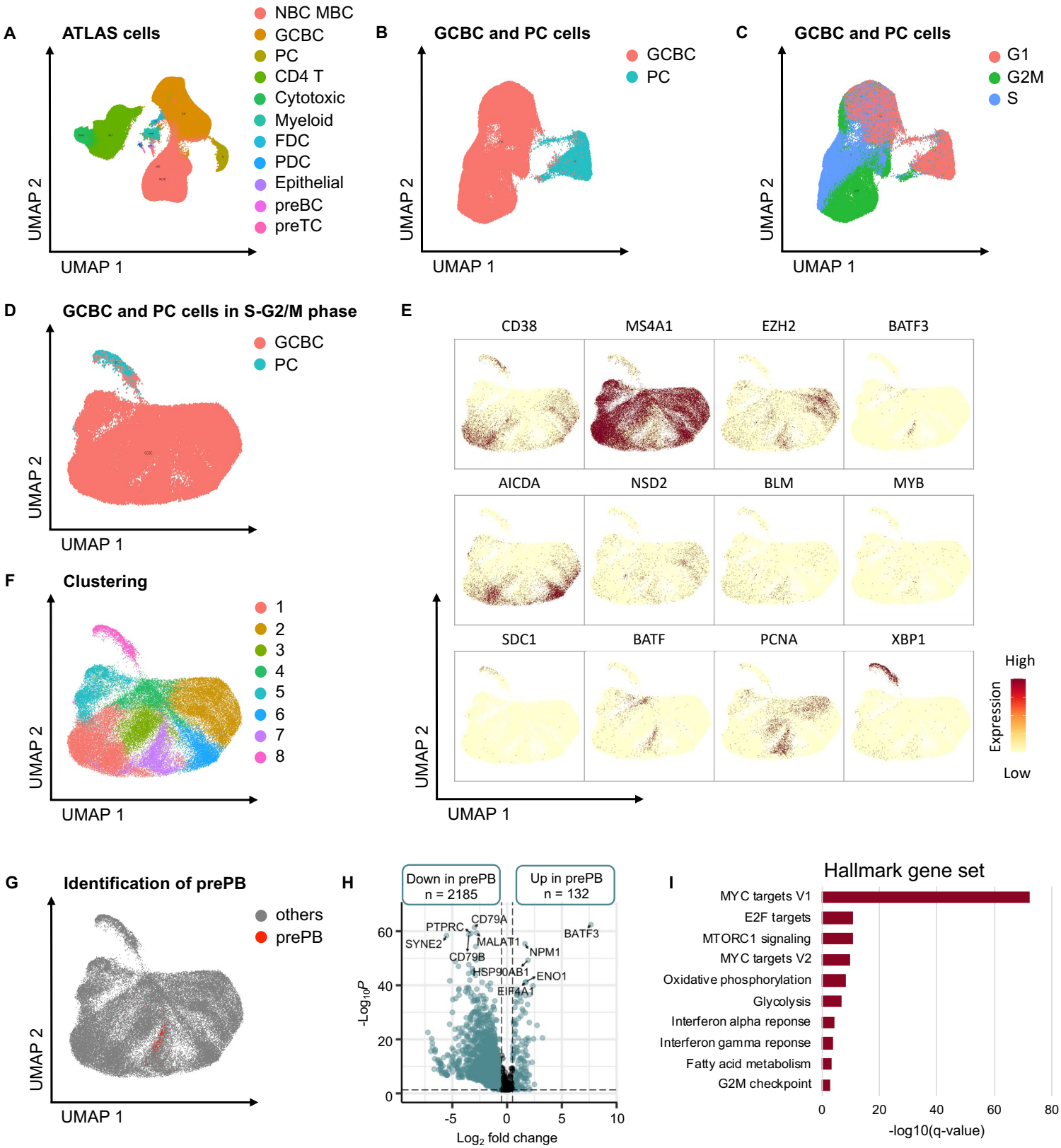
Supplemental Figure 9. UMAP projection of the whole cells (quiescent and proliferative) indicating cells from the cluster 2 (in blue) and cells expressing the spliced form of XBP1 (in red).



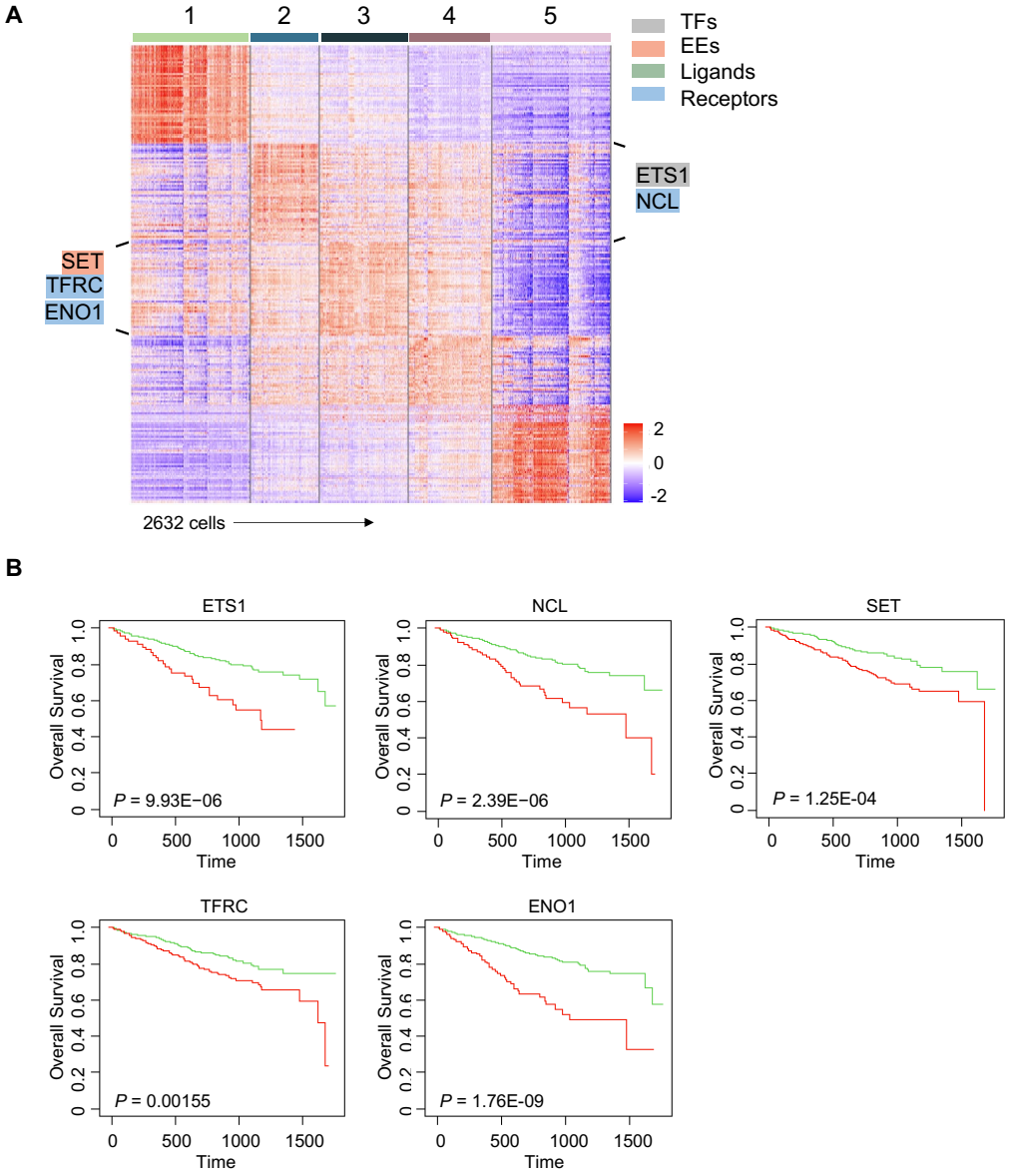
Supplemental Figure 10. Cells were treated with rapamycin (1 µM) or idelalisib (10 µM) from day 2 to day 4 (left charts) or from day 2 to day 7 (right charts). (A) Proliferation rate and the cumulative growth curve were calculated at each time point. (B) Viability was calculated as the percentage of live cells versus total cells in each condition. Results show the mean + SD of 4 independent experiments. Statistical significance was calculated using a Student's t-test for paired samples. *: p-value < 0.05; **: p-value <0.01. At each time point, cell populations were marked with specific fluorescent antibodies, quantified by flow cytometry and represented as percentage of total cells. prePB: CD20- CD38- CD138-; PB: CD20- CD38+ CD138-; PC: CD20- CD38+ CD138+. Results show the mean + SD of 4 independent experiments. Statistical significance was calculated using a Student's t-test for paired samples. *: p-value < 0.05; **: p-value <0.01; ***: p-value <0.001; ****: p-value < 0.0001.



Supplemental Figure 11. Identification of prePB cells in the public Tonsil Atlas dataset. (A) The UMAP representation showed the 209,786 cells composing the Tonsil Atlas dataset. (B) Then, we filtered in 92,218 GCBC and 14,256 PC cells from the dataset and (C) applied the Seurat cell-cycle scoring to identify quiescent cells (G1) and proliferative cells (S and G2M). (D) We filtered out quiescent cells to keep proliferative GCBC and PC cells (n=56,990). (E) mRNA expression of markers with high (*BATF3*, *BATF*, *EZH2*, *MYB*, *BLM*, *AICDA*, *NSD2* and *PCNA*) and low (*CD38*, *MS4A1*, *SDC1* and *XBP1*) expression in prePB. (F) Clustering of proliferative GCBC and PC cells to isolate prePB (SDC1-, CD38-, XBP1-, CD20-, BATF+, BATF3+, PCNA+). (G) Cells with high expression of *BATF3* from the cluster 7 were identified as prePB (n=143). (H) Volcano plots showing differentially expressed genes between the identified prePB *versus* the other cells. (I) Gene set enrichment analysis of the whole genes up-regulated in the identified prePB.


A
ATLAS cells
NBC MBC
GCBC
PC
CD4 T
Cytotoxic
Myeloid
FDC
PDC
Epithelial
preBC
preTC
UMAP 2
UMAP 1
B
GCBC and PC cells
GCBC
PC
UMAP 2
UMAP 1
C
GCBC and PC cells
G1
G2M
S
UMAP 2
UMAP 1
D
GCBC and PC cells in S-G2/M phase
GCBC
PC
UMAP 2
UMAP 1
E
CD38
MS4A1
EZH2
BATF3
AICDA
NSD2
BLM
MYB
SDC1
BATF
PCNA
XBP1
High
Expression
Low
UMAP 2
UMAP 1
F
Clustering
1
2
3
4
5
6
7
8
UMAP 2
UMAP 1
G
Identification of prePB
others
prePB
UMAP 2
UMAP 1
H
Down in prePB
n = 2185
Up in prePB
n = 132
PTPRC
CD79A
SYNE2
MALAT1
BATF3
CD79B
NPM1
HSP90AB1
ENO1
EIF4A1
-Log10P
Log2 fold change
I
Hallmark gene set
MYC targets V1
E2F targets
MTORC1 signaling
MYC targets V2
Oxidative phosphorylation
Glycolysis
Interferon alpha reponse
Interferon gamma reponse
Fatty acid metabolism
G2M checkpoint
-log10(q-value)

Supplemental Figure 12. Key marker genes identified in prePB and PB stages associated with bad prognosis in multiple myeloma patients. (A) The heatmap showed the top 50 genes up-regulated in each cluster. Genes associated to bad prognosis in the CoMMpass cohort (n = 674) were indicated and were colored in grey, red, green and blue if they code for TFs, EEs, ligands and receptors, respectively. (B) Kaplan-Meier curves of genes identified in A in CoMMpass cohort. Red curves represent high-risk group associated with high gene expression and green curves correspond to low-risk group associated with low gene expression.



Supplemental Figure 13. Key marker genes identified in prePB and PB stages associated with good prognosis in multiple myeloma patients. (A) The heatmap showed the top 50 genes up-regulated in each cluster. Genes associated to good prognosis in the CoMMpass cohort (n = 674) were indicated and were colored in grey, red, green and blue if they code for TFs, EEs, ligands and receptors, respectively. (B) Kaplan-Meier curves of genes identified in A in CoMMpass cohort. Red curves represent low-risk group associated with high gene expression and green curves correspond to high-risk group associated with low gene expression.