Specific pre-plasma cell states and local proliferation at the dark zone – medulla interface characterize germinal center-derived plasma cell differentiation in lymph node

Laurine Binet[1], Chuang Dong[1], Noudjoud Attaf[1], Laurine Gil[1], Matthieu Fallet[1], Thomas Boudier[2], Bertrand Escalière[1], Lionel Chasson[1], Carole Siret[1], Serge A. van de Pavert[1], Jean-Marc Navarro[1], Pierre Milpied[1,‡]

[1] Aix Marseille Université, CNRS, INSERM, Centre d'Immunologie de Marseille-Luminy, Marseille, France

[2] Turing Center for Living Systems, CENTURI, Aix Marseille Université, Marseille, France.

[‡] Corresponding author: Pierre Milpied (milpied@ciml.univ-mrs.fr)

ABSTRACT

High affinity antibody-producing plasma cell (PC) production in germinal centers (GC) is crucial for antibody-mediated immune protection after vaccination or infection. The selection of high affinity B cells in the GC light zone instructs PC differentiation in a subset of cells, but the phenotype, differentiation trajectory and spatial localization of those prePC intermediates remain to be characterized. Here, we have used a mouse model to track GC-derived B cells with integrative single-cell and spatial analyses in draining lymph node after immunization. We first identified putative prePC in scRNA-seq datasets, then enriched those cells through their specific surface phenotype for further analysis of their gene expression trajectories and BCR repertoire. We found a continuum of actively proliferating transitional states bridging selected LZ GC B cells and recently exported PCs, with gradually increasing levels of endoplasmic reticulum stress-associated genes and immunoglobulin transcripts. Spatial analyses revealed that recently differentiated PC continued their maturation and proliferation at the interface between the DZ and extensions of the lymph node medulla. Our results provide insights into the intermediate stages and microenvironmental factors involved in the differentiation of GC B cells into PC, with implications for vaccine development and understanding antibody responses.

KEYWORDS

Germinal center, plasma cell, lymph node, differentiation trajectory, localization, transitional states

SHORT RUNNING TITLE

Mapping GC to PC differentiation

INTRODUCTION

B cell immune responses are important for long term protection against pathogens, after natural infection or vaccination. Germinal Centers (GCs), micro-anatomical structures that form within B cell follicles in secondary lymphoid organs after T cell-dependent B cell activation, play a crucial part for long term antibody-based immunity. In GCs, the affinity maturation cyclic process enables the diversification and overall gain in affinity of antigen-specific antibodies expressed as surface B cell receptors (BCR) by GC B cells. In a maturation cycle, B cells first undergo cell division and somatic hypermutation (SHM) in the dark zone (DZ) in order to diversify their BCR before migrating to the light zone (LZ) and being selected based on the affinity of their B-cell receptor (BCR). Most studies have shown that selection broadly follows the affinity-dependent selection model[1]. In this model, B cells which cannot bind antigen die by apoptosis, the low affinity ones differentiate in memory B cells (MBC), the intermediate affinity ones recycle in the DZ and the high affinity ones differentiate in PC[1–4].

The differentiation of GC-derived MBC and plasma cells (PC) results in diversified and affinity-enhanced antibody specificities being expressed in long-lived cell types for long-term immune protection. In that regard, GC-derived PC have been shown to be the primary source of long-lived high-affinity antibody producing cells that home to the bone marrow[5]. The differentiation of GC B cells into long-lived PC results from the induction of gene expression modifications through different signals. First, GC B cells have to internalize the antigen contained in immune complexes at the follicular dendritic cell (FDC) membrane before they present it to T follicular helper cells (T_{FH}) on MHC class II molecules and receive co-stimulatory signals. The BCR signal also plays an important role by repressing *Bcl6* transcription[6,7] and inducing *Irf4* expression[8,9]. BCR affinity allows to be more competitive and to extract more antigen at the FDC surface, which has been correlated with a better help from T_{FH} cells [10–13]. T_{FH} cells allow the selection of GC B cells through a competitive access to co-stimulatory signals such as CD40L and IL21. CD40 signaling reinforces BCR signaling through the blockade of *Bcl6* expression and the induction of *Irf4* in a dose-dependent manner[14–16]. Finally, IL21 was described to have a dual role in the GC. It induces the expression of Bcl6, but when it synergizes with the BCR signal and the CD40 signal, it allows the expression of Irf4 and Blimp1 (encoded by *Prdm1*), two markers important for terminal PC differentiation[17–20].

The precise mapping of stepwise processes for PC differentiation has been mostly studied in *ex vivo* differentiation assays starting from naïve B cells of MBC. Although the situation may differ for the GC to PC differentiation *in vivo*, those assays have generated important findings at the genetic level. Notably, because some GC markers such as Bcl6 or Pax5 repress the PC phenotype[6,21], it is then mandatory for B cells to downregulate the expression of the GC profile in order to differentiate in GC-derived PC[9,22]. Once B cells have switched their gene expression profile towards a PC profile, they undergo several rounds of cell division that allow the hypomethylation of the chromatin parts encoding the proteins responsible for the PC profile in an irreversible way[23]. Thus, PC differentiation goes through a remodeling at the epigenetic, transcriptomic and phenotypic profile. These changes allow the functional remodeling of activated and proliferating GC B cells to antibody-secreting PC.

Several studies have tried to identify intermediate states bridging GC B cells to PC *in vivo*. Indeed, cells intermediate in their transcriptomic profile[24], or LZ GC B cells with PC features such as high expression of *Irf4* in mice[16,25], have been described in recent years. In humans, transcriptomic studies also highlighted intermediate states expressing PC and cell cycle markers[26–28]. Unfortunately, the scarcity of these states in lymphoid organs have most often precluded their in-depth study and the differentiation trajectory between selected GC B cells and differentiated antibody-producing cells remains uncharted. *In silico* modeling experiments have suggested that PC exit the GC from the DZ, but *in vivo* evidence is lacking[29]. Several signals appear to be implicated in PC egress from GC (GPR183 signaling, S1P1 and CXCL12 gradients), but no final mechanism has been identified[30,31]. During early T-dependent responses, plasmablasts accumulate at the GC-T zone interface (GTI), but whether this also occurs for GC-derived PC is not known. PC that are found in the LN medulla, further away from the GC DZ, express less proliferation markers ($Ki67^+$) than those that are closer to the DZ, indicating a possible spatio-temporal axis of GC to PC differentiation[32]. In addition, it was recently shown that post-selection PC expand clonally without further SHM[33], but whether this expansion occurs inside or outside of the GC, or where in the dLN, is still unknown.

Here, we identified GC-to-PC intermediate differentiation states in mouse draining LN after model antigen immunization, that we characterized through integrative analyses of phenotypes, transcriptomes, BCR repertoires, and microanatomic localization. We show that GC-derived PC differentiation implies local proliferation at the interface between the GC DZ and the LN medulla.

RESULTS

Tracking GC B cells and recent GC emigrants after model immunization

We used a GC B cell fate mapping mouse model, *Aicda-Cre-ERT2 x Rosa26-lox-STOP-lox-eYFP*[34], to track GC B cells and their recent progeny in draining lymph nodes after subcutaneaous immunization. After immunization with the T-dependent model antigen 4-hydroxy-3-nitrophenylacetyl-conjugated keyhole lympet hemocyanin (NP-KLH) in Sigma Adjuvant System adjuvant, we gavaged the mice once with tamoxifen to induce eYFP expression in *Aicda* expressing cells, and analyzed cells in draining lymph nodes (dLN) by flow cytometry at day 13 after immunization, 2 to 6 days after tamoxifen gavage (**Figure 1a**). In that model, gavaging with tamoxifen after day 6 ensures that eYFP expression is triggered almost exclusively in GC B cells. Thus, among activated IgD^{neg} B cells, we detected eYFP expression in $CD138^{+}$ PC (**Figure 1b**), $GL7^{+}CD38^{-}$ GC B cells, and $CD38^{+}GL7^{-}$ memory B cells (MBC) (**Figure 1c**). While tamoxifen-induced labeling was already maximal 2 days after gavage in GC B cells, GC-emigrant $eYFP^{+}$ MBC and PC were barely detectable at that time, but increased gradually after 4 to 5 days (**Figure 1d**). We then assessed the proportion of NP-binding cells among IgD^{neg} B cell subsets in dLN at day 14 after NP-KLH immunization and 4 days after tamoxifen gavage (**Figure 1e-f**). Total GC B cells and PC included on average 20% and 30% $NP\text{-}PE^{+}$ cells, respectively, while MBC included less than 5% NP-binding cells (**Figure 1e**). Restricting the analysis to $eYFP^{+}$ cells resulted in higher proportions of NP-binding cells in all subsets, including MBC where 15% on average were $NP\text{-}PE^{+}$ (**Figure 1f**), indicating that PC and MBC exported from the GC from day 10 to day 14 were enriched in cells that could detectably bind antigen. Thus, the *Aicda-Cre-ERT2 x Rosa26-lox-STOP-lox-eYFP* mouse model enabled us to track GC output of PC and MBC in dLN after subcutaneous model antigen immunization.

Single-cell RNA-seq analysis of GC B cells and recent GC emigrants identifies putative prePC

Using the GC B cell fate mapping mouse model in different immunization schemes with either the chicken ovalbumin (OVA) or NP-KLH model antigens (**Figure 2a**), we generated 3 distinct single-cell RNA-seq (scRNA-seq) datasets of IgD^{neg} $eYFP^{+}$ B cells from dLN. In dataset #1, mice were primed with OVA in Alum, and we investigated cells 4 days after tamoxifen gavage, at day 10 and day 20 of either the primary response, or the secondary response after boosting, using the FACS-based 5-prime-end sequencing (FB5P-seq) method for integrative scRNA-seq analysis[35]. In dataset #2, mice were primed with NP-KLH in Sigma Adjuvant System, and we investigated cells 5 days after tamoxifen gavage, at day 16 of the primary response, using droplet-based 5'-end scRNA-seq. In dataset #3, mice were primed with NP-KLH in Sigma Adjuvant System, and we investigated cells 3 days after tamoxifen gavage, at day 10, day 16 and day 22 of the primary response, using droplet-based 5'-end scRNA-seq. After standard quality controls of the resulting scRNA-seq datasets, we integrated all experiments in a single low dimensional UMAP embedding (**Supplementary Figure 1a-b**) and defined gene expression-based subsets by non-supervised clustering (**Figure 2b**). Based on marker genes expression (**Figure 2c**), cluster 0 included mostly quiescent cells and corresponded to LZ GC B cells (*Cd83*, *Il4i1*); clusters 1, 2, 3 and 7 expressed distinct combinations of cell cycle associated genes and corresponded to DZ GC B cells; cluster 5 expressed the typical Myc-induced signature of positively selected LZ GC B cells (*Npm1*, *C1qbp*); cluster 4 included mostly quiescent cells and expressed both LZ and MBC markers (*Klf2*, *Serpinb1a*), suggesting those cells corresponded to preMBC or early MBC; and cluster 6 expressed high levels of immunoglobulin (Ig) coding genes and corresponded to PC. Although there was a clear continuum between quiescent LZ GC B cells and preMBC in the low dimensional embedding, the PC cluster was completely separated and we failed to identify transitional states from GC to PC through non-supervised analyses in that dataset.

We reasoned that a supervised approach would more accurately identify putative prePC transitional states that bridge GC B cells and PC. We thus used a gene signature-based scoring approach to hierarchically "gate" cells in the integrated scRNA-seq dataset (**Supplementary Table 1**, **Figure 2d** and **Supplementary Figure 1c**). In particular, we identified rare cells expressing intermediate levels of a PC-specific gene signature that were embedded in the GC "continent" of the UMAP embedding, which we named "prePC" (**Figure 2e**). Other supervised annotations, DZ, LZ, LZtoDZ, preMBC and PC were consistent with non-supervised clusters (**Figure 2e** and **Supplementary Figure 1d**). Putative PrePC represented very low frequencies (average 0.2%) of sequenced IgD^{neg} eYFP cells in all time points of all datasets (**Figure 2f**). We computed marker genes of the supervised annotated cell subsets in the 3 datasets (**Supplementary Figure 1e**). PreMBC, LZ, DZ, LZtoDZ and PC expressed the expected gene expression programs that have already been described in other mouse GC B cell scRNA-seq datasets[36,37]. Putative prePC expressed *Bst2*, *H2-Q7*, *Glipr1*, *Selplg*, *Itgb7*, *Plac8*, *Cd22*, *Usp8* and *Serpina3g* among other marker genes (**Supplementary Figure**

143 **1e**). In particular, high expression of *Psgl1* and *Bst2* (**Figure 2g**), encoding the surface markers PSGL1 and
144 BST2, suggested that prePC may be identified and enriched by flow cytometry.

145 *Enrichment of putative prePC by flow cytometry*

146 We thus designed a 14-color flow cytometry panel targeting surface markers and transcription factors
147 characteristic of GC, MBC and PC, and including the PSGL1 and BST2 markers, which we applied to analyze
148 IgD^{neg} B cells in dLN of *Aicda-Cre-ERT2 x Rosa26-lox-STOP-lox-eYFP* mice previously immunized with NP-
149 KLH in Sigma Adjuvant System adjuvant and gavaged with tamoxifen (**Figure 3a**). UMAP embedding, based
150 on 10 surface and intracellular markers identified clusters of phenotypically defined GC, MBC and PC, as
151 well as cells located in intermediate areas of the low dimensional embedding (**Figure 3b**). We gated cells
152 situated within GC and PC clusters as "GC-to-PC" and inspected their surface phenotype in comparison with
153 the well-defined GC, MBC and PC clusters. GC-to-PC cells expressed high levels of the GC markers GL7,
154 CD19 and B220 markers, high levels of the PC-specific transcription factor IRF4, intermediate levels of the
155 PC marker CD138, and intermediate levels of PSGL1 (**Figure 3c**). Based on the specific surface phenotype
156 of GC-to-PC cells, we thus reverse engineered a gating strategy (**Figure 3d**) that allowed for the identification
157 and quantification of prePC (**Figure 3e**). At days 14-16 after primary immunization with NP-KLH, prePC
158 accounted for approximately 0.08% of IgD^{neg} B cells in dLN (**Figure 3f**), included significantly lower
159 proportions of NP-binding cells (18% on average) compared to GC B cells (23%) and PC (24%), which was
160 still higher than the proportion of MBC detectably binding NP (2%) (**Figure 3g**).

161 *Characterization of prePC in the GC-to-PC differentiation continuum*

162 In order to gain more insight into the molecular features of prePC, we next used our gating strategy to sort
163 phenotypically defined prePC and compare them directly to DZ, LZ and PC in FB5P-seq (**Figure 4a**). After
164 quality control and cell cycle regression of the resulting dataset, low-dimensional UMAP embedding displayed
165 a clear separation between DZ and LZ GC B cells on one side, and PC on the other side, with a subset of
166 phenotypically defined prePC bridging the 2 cell continents (**Figure 4b**). Consistent with our previous analysis
167 on non-enriched cells (**Supplementary Figure 1e**), single prePC combined the expression of GC B cell
168 signature genes *Ms4a1*, *Cd19*, *Irf8*, with the expression of positive-selection induced gene *Myc*, and the
169 expression of PC differentiation surface markers and transcription factors *Sdc1*, *Prdm1*, *Irf4*, *Selplg* and *Bst2*
170 (**Figure 4c**). Those cells were also actively proliferating, either in S or G2/M phase of the cell cycle (**Figure**
171 **4d**).

172 PC differentiation is characterized by the production of high amounts of antibodies, requiring high level of Ig
173 genes transcription and the induction of a specific endoplasmic reticulum (ER) stress response[31]. PrePC
174 expressed intermediate levels of genes involved in the ER stress response (**Supplementary Figure 2a**), and
175 intermediate levels of Ig transcript counts (**Supplementary Figure 2b**), when compared with GC B cells and
176 PC. We thus defined a continuum of GC-to-PC differentiation based on single-cell gene expression of the
177 ER stress module and Ig transcripts counts (**Figure 4e-g**), in which most phenotypically defined prePC
178 bridged the gap between GC B cells and PC. Another feature of PC differentiation, the gradual loss of antigen-
179 presentation capacity on MHC-II[38], was also intermediate in phenotypically defined prePC (**Figure 4h** and
180 **Supplementary Figure 2c-d**). Functionally, FACS-sorted prePC spontaneously secreted detectable
181 amounts of soluble IgG in *ex vivo* cultures (**Figure 4i**). Thus, prePC with the $GL7^{+}$ $B220^{+}$ $CD138^{int}$ $PSGL1^{+}$
182 phenotype were engaged in the early stages of GC-to-PC differentiation and had already initiated the
183 antibody production program.

184 The distribution of phenotypically defined prePC along the GC-to-PC differentiation continuum showed that
185 those cells were also heterogeneous, and could be further divided based on their degree of differentiation
186 (**Figure 4j-k**). PrePC are rare intermediates in non-enriched datasets (*e.g.* **Figure 2**), but our enriched
187 dataset provided a unique opportunity to finely map the transcriptional changes that pave the GC-to-PC
188 differentiation. We thus clustered genes according to the dynamic evolution of their expression as cells
189 progress through the GC-to-PC continuum (**Figure 4l**), and computed the gene ontology enrichment of gene
190 modules that are progressively lost (clusters 1-3) or induced (clusters 4-6) through differentiation (**Figure**
191 **4m**). Those analyses revealed that very early metabolic reprogramming (nucleotide biosynthesis process,
192 cellular respiration, oxidative phosphorylation, proton motive force−driven mitochondrial ATP synthesis)
193 preceded the gradual increase in antibody production-associated physiological responses (response to ER
194 stress, response to unfolded protein).

BCR analysis of prePC

The affinity maturation trajectory of GC B cells and their progeny is in-part traceable through the analysis of their Ig gene transcripts[39]. Our prePC-enriched FB5P-seq analysis retrieved IgH and IgL sequences for the vast majority of single cells (n=965/1367) For better resolution of the different gene expression programs and their associated BCR sequence features, we re-annotated the cell types and states using the supervised annotation method described in **Supplementary Figure 1c**, resulting in the identification of DZ, LZ, LZtoDZ, PC, preMBC and prePC with specific gene expression programs and phenotypes (**Supplementary Figure 2e**). While close to 90% of DZ, LZ and LZtoDZ subsets carried mutations in their IGH and IgL variable genes, only 75% of preMBC and 40% of prePC and PC were mutated (**Figure 5a-b**). Unmutated and mutated prePC and PC did not differ in their gene expression programs as testified by low-dimensional UMAP embedding (**Figure 5c**), but were slightly different in isotype usage, with unmutated cells being more frequently IgM-positive (**Figure 5d**). IgM-expressing cells represented approximately 25% of both DZ and LZ cells, less than 5% of positively selected LZtoDZ cells, and about 60% of preMBC; in unmutated PC and prePC, 30% were IgM-positive, as compared to 5% and 10%, respectively, for their mutated counterparts.

The supervised annotation showed that prePC were split into two subsets: one expressing intermediate levels of ER stress response genes and the signature of recent positive selection; and the other, mostly quiescent, expressing low levels of ER stress response genes and no sign of activation (**Figure 5e**). We re-annotated those prePC subsets as prePC.Act and prePC.ERstress^{lo} (**Figure 5f-g**) and re-analyzed their surface phenotype and BCR mutations. PrePC.Act cells had the typical prePC phenotype identified in this study (GL7^{+} CD138^{int} B220^{+} CD19^{+}), while prePC.ERstress^{lo} were phenotypically heterogeneous and analogous to PC, suggesting that the latter were more advanced in their differentiation (**Figure 5h**). BCR mutations did not differ between the two prePC subsets (**Figure 5i**). Finally, we identified BCR clonotypes in the dataset and analyzed the compositions in distinct cell types and states of clonal families (groups of 2 or more cells sharing the same IgH and IgL clonotype). We first annotated clonal families depending on whether they contained cells from one or several of the cell states defined by our supervised annotation; then we analyzed for each cell state the proportion of cells being allocated to distinct clonal families (**Figure 5j**). About half of DZ and LZ cells were in clonal families comprising either only GC B cells (either DZ, LZ and/or LZtoDZ), GC and PC, or GC, PC and prePC. By contrast, LZtoDZ cells were more frequently clonally associated to GC, prePC and PC, and preMBC were most often not clonally related to any other cell in the dataset. As expected for transitional states in the GC-to-PC differentiation, prePC.Act cells were often clonally related to both GC and PC.

Overall, the integration of gene expression, surface phenotype and BCR sequence measurements in single cells enriched in putative prePC enabled the fine characterization of the rare intermediate transitional states in the GC-to-PC differentiation.

Recent GC-derived PC proliferate at the DZ-medulla interface

Some of the genes specifically associated with the prePC state encoded chemotaxis (*Cxcr3*, *Gpr183*) or adhesion molecules (*Itga4*, *Itgb1*) (**Figure 6a**), suggesting that prePC may be localized in a specific compartment in the dLN. To assess the location of GC-derived PC differentiation in our model, we first performed whole transcriptome spot-based spatial transcriptomics (10x Genomics Visium) analysis of 4 sections of the same 2 dLN, separated by approximately 50µm in depth (**Figure 6b**). After quality control, we annotated spots with a semi-supervised approach: first using non-supervised spot clustering and marker genes to identify spots corresponding to B follicles, T zones, medulla, GC LZ and GC DZ; then defining three zones at the interface of the GC (Methods), the GC-T zone interface (GTI), the GC-B follicle interface (GBI), and the GC-medulla interface (GMI) (**Figure 6b-c**). We then scored a PC gene expression signature on every spot as a proxy for detecting the presence of PC in the distinct dLN microenvironments and interfaces. The PC score was high in medulla, as expected since PC are known to reside in medullary chords before dLN exit[40]; among GC-proximal interfaces, GMI spots were the main areas with high PC scores (**Figure 6d**). This result suggested that stromal components of the medulla that support PC survival extend to the direct proximity of the GC, at the DZ side. Accordingly, several stromal cell subsets gene expression signatures[41] also scored highly in the medulla and GMI areas (**Supplementary Figure 3a-b**).

We thus turned to high-resolution microscopy to precisely map the location and interactions of eYFP^{+} IRF4^{+} cells, which we defined as recently differentiated PC. Consistent with our spatial transcriptomics results, we found clusters of eYFP^{+} IRF4^{+} cells accumulating in collagen-IV-rich areas in close proximity to the GC DZ (**Supplementary Figure 3c**). At the GMI, eYFP^{+} IRF4^{+} cells were surrounded by Lyve1^{+} lymphatic endothelial cells and PDGFR1^{+} stromal cells (**Figure 6e**). The density of eYFP+ IRF4+ cells in the GMI and

medulla was similar (**Figure 6f**) Although on some sections, the eYFP$^+$ IRF4$^+$ cells in close proximity to the GC DZ seemed to express lower levels of CD138 than those more distal in the medullary areas (**Figure 6e**, panel *iv*), suggesting less advanced differentiation, that observation was not statistically significant when considering all GCs from multiple dLN sections (**Figure 6g**). Our scRNA-seq data and recently published observations[33] have indicated that early differentiating prePC, or recently differentiated PC proliferate extensively. In situ, we showed that a majority of eYFP$^+$ IRF4$^+$ at the GMI expressed Ki67, a hallmark of active proliferation (**Figure 6h**). Since we did not find a GMI area surrounding every GC in every tissue section, we reasoned that the medulla may connect with the GC DZ only at certain contact areas that are not always captured in a 5-10 µm tissue section. We thus run whole-LN 3D-imaging using light sheet microscopy, and could indeed show that every GC in a dLN included a GMI contact area, but that the area was not found at all depths (**Figure 6i** and **Supplementary Movie**).

Overall, we have characterized the GMI, a specific dLN microenvironment defined as an extension of the medulla contacting the GC DZ at some level, as the main environment for exit and proliferation of GC-derived PC.

DISCUSSION

Our study provides a detailed characterization of rare prePC intermediate states in the GC-to-PC differentiation trajectory in draining LN after immunization. We showed that prePC expressed a mixed phenotype expressing both GC and early PC markers in phenotypic and transcriptomic analyses; they were also actively cycling, expressed intermediate levels of MHC-II genes, had initiated metabolic reprogramming and started upregulating UPR programs, and produced antibodies *ex vivo* without any additional signal. Those data indicate that we have identified late intermediate states in the GC-to-PC differentiation trajectory. As prePC spanned different unsupervised clusters in gene expression space, we analysed whether multiple prePC states existed. PrePC could be subdivided based on either their BCR mutational profile, or their gene expression profile. In the latter, we separated ERstress^{low} PrePC from activated PrePC which expressed a gene expression profile (Myc-induced genes) and phenotype (GL7^{hi} CD138^{lo} B220^{int} CD19^{hi}) related to selected LZ GC B cells. PrePC also expressed a distinctive repertoire of homing markers, suggesting they had already homed or were homing to extra-GC areas in draining LN. We found recently GC-exported PC proliferating just outside of the GC DZ, at the interface with medullary tissue microenvironments that connected every GC to the deep medulla. The GC-medulla interface contained Lyve1^{+} and PDGFR1^{+} stromal cells that may serve as a maturation niche for cells in the GC-to-PC differentiation trajectory.

Several studies have identified prePC states as LZ GC B cells expressing high levels of *Irf4* in mice[42,43], and cells expressing both PC and cell cycle markers in humans[5,10,44]. Those studies lacked other information such as their surface phenotype, function, and *in situ* localization, making it difficult to compare them directly with our definition of prePC. Recent studies in human tonsils has identified pre-plasmablasts that appeared integrated in the differentiation trajectory from GC to PC[26,27,45]. Indeed, in that study the authors described cells that were intermediate in their transcriptomic profile, and shared clones with both GC B cells and PC. Even though the authors did not focus their analyses on these states, we could identify some common features with the prePC we have identified in our study in mice (*e.g.* expression of PC lineage genes like *IRF4*, *PRDM1*, *XBP1*, *FKBP11*, and cell cycle progression), highlighting some evolutionary conserved modules activated at the GC-to-PC transition. In another high resolution human tonsil atlas[46], it was described that GC-derived PC branched both from LZ GC B cells and activated LZ-to-DZ B cells, the latter likely representing the counterparts of the activated prePC we describe in our study in mouse LN. Another study characterized human prePC in an *ex vivo* MBC differentiation model using a combination of scRNAseq and ATACseq[47]. Although they originated from blood MBC, and not LN GC B cells, prePC in that study were highly heterogenous and integrated in a continuum of differentiation with stepwise induction of the unfolded protein response, and with AP-1 family transcription factors like BATF playing a likely role at the onset of differentiation. In our study, prePC branched from the LZ-to-DZ GC B cell state, where the transcription factor BATF has been shown to play a key role for metabolic refuelling induced by T cell help[48]. Altogether, our current results on prePC are consistent with previous studies, refine the molecular characterization of those rare intermediate states, and introduce a gating strategy for enriching those cells from mouse GC by flow cytometry, which will be important to specifically analyze that population in future studies.

In silico modeling experiments have suggested that PC exit the GC from the DZ[29]. Several homing receptors appear to be implicated in PC egress from GC (Gpr183, S1PR1, CXCR4), but the precise route that prePC or PC use to exit GCs remained unidentified[30,31]. Based on gene expression, we found that prePC were likely in the process of egress from the GC, but their scarcity precluded their robust identification by confocal microscopy in LN tissue sections. Instead, we analyzed recent GC emigrant PC (RGCE-PC) that were eYFP^{+} IRF4^{+} 3-4 days after tamoxifen induction of eYFP expression in GC B cells. RGCE-PC accumulated at the border of the GC. In previous studies, PC differentiated after early GC-independent B cell activation accumulated at the interface between the nascent GC and the T zone[41,42]. In our study, we showed that RGCE-PC were in fact accumulating and proliferating in DZ-proximal microenvironments characterized by a gene expression and cellular composition similar to the LN medulla. Notably, the GC-medulla interface was surrounded by Lyve1^{+} cells resembling medullary sinus LEC[49], and contained numerous PDGFR1^{+} stromal cells with high gene expression scores for previously described medullary stromal cell subsets (Nr4a1^{+} and Inmt^{+} stromal cells)[41]. Medullary fibroblastic reticular cells have been shown to provide essential survival and maturation cues to plasma cells in lymph nodes[50], but the fact that those cells may be in such a close proximity to the GC DZ has remained unnoticed in imaging studies focusing on B follicle stromal remodeling[51,52]. We attribute that to the fact that most 2D images of thin tissue sections failed to capture the GMI, which bridged medulla and GC DZ only at certain z-depth for a given GC structure. We did not detect GMI areas in non-immunized lymph nodes (unpublished observation), and other studies have shown a tight separation between medulla and B cell follicles in naïve LN[50,53]. However, adjuvanted immunization or infection induces

320 lymphangiogenesis and stromal remodeling[50,54] close to activated B cell follicles, which may initiate the
321 spontaneous organization of a GMI area that supports early PC proliferation, maturation and migration.

322 In our experiments, close to 25% of LZ-to-DZ cells were clonally related to PC or prePC, compared to less
323 than 10% for LZ or DZ GC B cells, suggesting that prePC derive more closely from GC B cells clones selected
324 at the LZ-to-DZ stage. Thus, our study implies that, after being selected in the LZ, a fraction of LZ-to-DZ state
325 GC B cells commit to the prePC state and proliferate at the external border of the DZ, mature locally in the
326 GC-medulla interface, before migrating along the medullary continuum towards efferent lymphatics.

327 Overall, our study provides novel insights into the intermediate stages and microenvironmental factors
328 involved in the differentiation of GC B cells into PC. The identification of the GMI as an important lymph node
329 microenvironment for post-GC PC maturation and clonal expansion opens new perspective for studying
330 factors contributing to PC differentiation. We anticipate that in-depth studies of those intermediate cellular
331 states in different models of infection and vaccination will enable a finer understanding of the mechanisms
332 that are key for generating long-lived high affinity PC, with implications for vaccine development.

333

334

335 **EXPERIMENTAL METHODS**

336 *Mouse models*

337 *Aicda-Cre-ERT2 x Rosa26-lox-STOP-lox-eYFP* mice[34] were bred at the Centre d'Immuno-Phenomique,
338 (Marseille, France), and transferred to the animal care facility of Centre d'Immunologie de Marseille-Luminy
339 for experiments. C57BL/6 mice were purchased from Janvier Labs (Le Genest-Saint-Isle, France). All mice
340 were maintained in the CIML mouse facility under specific pathogen-free conditions. Experimental
341 procedures were conducted in agreement with French and European guidelines for animal care under the
342 authorization number APAFIS #30945-2021040807508680, following review and approval by the local animal
343 ethics committee in Marseille. Mice were used regardless of sex, at ages greater than 7 weeks and less than
344 3 months.

345 Mice were immunized with either 100µg of NP-KLH at 1µg/µL emulsified with Sigma Adjuvant System (SAS)
346 at a 1:1 (v:v) ratio, or 100µg chicken ovalbumin (OVA) at 1µg/µl emulsified with Alum at a 1:1 (v:v) ratio,
347 subcutaneously at the base of the tail, 50µl on each side. For induction of the Cre-ERT2-mediated labelling,
348 we gavaged the mice once with 5mg of tamoxifen (TS648-1G, Sigma) in 200µL of peanut oil (P2144-250 ML,
349 Sigma), at least 6 days after immunization. Mice were euthanized between 10 days and 21 days post-
350 immunization (prime or boost) according to the experiments.

351 *Flow cytometry*

352 Single-cell suspensions from draining lymph nodes were washed and resuspended in FACS buffer (5% fetal
353 calf serum, 2mM EDTA, 5% Brilliant Stain Buffer Plus in PBS 1X) at a concentration of 100 million of cells
354 per ml. Cells were first incubated with FcBlock (Biolegend) for 10 min on ice. Then, cells were incubated with
355 a mix of antibodies (see table1 below) conjugated with fluorochromes 30 min on ice. Cells were washed in
356 PBS, and incubated with the Live/Dead Fixable Aqua Dead Cell Stain (Thermofisher) for 10 min on ice. Cells
357 were then washed again in FACS buffer and resuspended in 2% paraformaldehyde during 50min to preserve
358 the eYFP contained in the cytoplasm. Cells were washed and permeabilized using the FoxP3
359 permeabilization kit (eBioscience) during 30min, then washed again in the permeabilization buffer and
360 incubated with intracellular antibodies for 45min at RT. Cells were finally washed in permeabilization buffer
361 and resuspended in FACS buffer. Cell suspensions were analyzed on the LSRII UV cytometer (Becton
362 Dickinson).

363

mAb	Clone	Manufacturer	Reference
Ki67 - BUV395	B56	BD	564071
PSGL1 (CD162) - BUV737	2PH1	BD	741796
GL7 - Pacific Blue	GL7	Biolegend	144614
CCR6 (CD196) - BV605	29-2L17	Biolegend	129819
CD86 - BV650	GL-1	Biolegend	105035
CD138 - BV711	281-2	BD	563193
CD19 - BV785	6D5	Biolegend	115543
anti-GFP (cross-react eYFP) – AF488		Thermofisher	A-21311
NP - PE		Home made	
B220 - PE-Cy5	RA3-6B2	Biolegend	103209
CXCR4 - PE-Cy7	2B11	Thermofisher	25-9991-80
IRF4 - AF647	IRF4.3E4	Biolegend	646409
CD38 - AF700	90	Thermofisher	56-0381-82
CD3 - APC-Cy7	17A2	BD	560590
Gr1 (Ly6G) - APC-Cy7	1A8	BD	560600
IgD - APC-Cy7	11-26c.2a	Biolegend	405716
Bst2 (CD317) - AF647	129c1	Biolegend	127106

364

365 For cell sorting, cells were pre-enriched using the "Pan B Cell Isolation Kit II mouse" enrichment kit from
366 Miltenyi Biotec (ref. 130-104-443) in which we added a biotinylated anti-IgD antibody for further enrichment
367 of IgD^{neg} B cells (Biolegend ref. 405734). Cells were prepared as mentioned in the protocol provided and
368 passed through LS columns according to the manufacturer's instructions. We collected the negative fraction
369 and processed the cells according to the classical extracellular staining protocol, using antibodies described
370 in the table below. Cells were sorted on the BD Influx™ Cell Sorter, in 96-well plates, with index-sorting mode
371 for recording the fluorescence parameters associated to each sorted cell.

mAb	Clone	Manufacturer	Reference
B220 - FITC	RA3-6B2	Biolegend	103205
CXCR4 - PerCP-eF710	2B11	Thermofisher	46-9991-82
GL7 - Pacific Blue	GL7	Biolegend	144614
PSGL1 (CD162) - AF647	2PH1	BD	562806
CD3 - APC-Cy7	17A2	BD	560590
Gr1 (Ly6G) - APC-Cy7	1A8	BD	560600
IgD - APC-Cy7	11-26c.2a	Biolegend	405716
CD38 - PE	90	Invitrogen	12-0381-82
CD19 - PE-Dazzle594	6D5	Biolegend	115554
CD86 - PE-Cy7	GL-1	Thermofisher	12-0862-82

372

373 Data were analysed using FlowJo (v10.8.1).

374 *FB5P-seq*

375 The protocol was performed as previously described by Attaf et al.[35]. Individual cells were sorted into a 96-
376 well PCR plate, with each well containing 2 µL of lysis buffer. Index sort mode was activated to record the
377 fluorescence intensities of all markers for each individual cell. Flow cytometry standard (FCS) files from the
378 index sort were analyzed using FlowJo software, and compensation parameters were exported as CSV
379 tables for subsequent bioinformatic analysis. Immediately after sorting, plates containing individual cells were
380 stored at -80°C until further processing. Following thawing, reverse transcription was performed, and the
381 resulting cDNA was preamplified for 22 cycles. Libraries were then prepared according to the FB5P-seq
382 protocol. The FB5P-seq data were processed to generate both a single-cell gene count matrix and single-
383 cell B cell receptor (BCR) repertoire sequences for B cell analysis. Two separate bioinformatic pipelines were
384 employed for gene expression and repertoire analysis, as detailed in Attaf et al.[35].

385 *10x 5' scRNA-seq*

386 For experiments described in Figure 2a, cells from draining lymph nodes were washed and resuspended in
387 FACS buffer (PBS containing 5% FCS and 2 mM EDTA) at a concentration of 10^8 cells/ml. Samples from
388 different mice and different time points were processed separately using cell hashing as described[55]. Cells
389 were individually stained with distinct barcoded anti-mouse CD45 antibodies (in-house conjugated) in FACS
390 buffer for 30 minutes on ice. Subsequently, cells were washed and stained with a mix of primary antibodies,
391 then Live/Dead Fixable Aqua Dead Cell Stain (Thermofisher). Live cells of interest (IgD^{neg} $eYFP^{+}$ PC and B
392 cells) from each sample were bulk-sorted using a BD Influx cell sorter. PC and non-PC were captured in
393 distinct wells for droplet-based scRNA-seq to maximize the recovery of BCR sequence information from non-
394 PC. Within each fraction, cells from different samples were pooled and loaded for the 10x Genomics Single
395 Cell 5' v2 workflow. Libraries were prepared according to the manufacturer's instructions with modifications
396 for generating B cell receptor (BCR) sequencing libraries. Following cDNA amplification, SPRI select beads
397 were used to separate the large cDNA fraction derived from cellular mRNAs (retained on beads) from the
398 hashtag oligonucleotide (HTO)-containing fraction (in supernatant). For the mRNA-derived cDNA fraction, 50
399 ng were used to generate the transcriptome library, and 10-20 ng were used for BCR library construction.
400 Gene expression libraries were prepared according to the manufacturer's instructions. For BCR libraries,
401 heavy and light chain cDNAs were amplified by two rounds of PCR (6 cycles + 8 cycles) using external
402 primers recommended by 10x Genomics. Approximately 800 pg of purified amplified cDNA was tagmented

using the Nextera XT DNA Sample Preparation kit (Illumina) and amplified for 12 cycles using the SI-PCR forward primer (10x Genomics) and a Nextera i7 reverse primer (Illumina). For the HTO-containing fraction, 5 ng were used to generate the HTO library. The resulting libraries were pooled and sequenced on an Illumina NextSeq2000 platform with single-indexed paired-end kits following the manufacturer's guidelines.

Ex vivo culture and ELISA

Single-cell suspensions FACS-sorted from draining lymph nodes were washed and resuspended in culture medium (10% FCS, 0.1% 2-Mercapto-ethanol, 1% non-essential amino acids, 1% Sodium Pyruvate, 1% HEPES buffer; 1% L-Glutamine, 1% Peniciline-Streptomycin in RPMI 1640) and placed in culture for different times at 37°C, 5% CO_2. Cell culture supernatants were collected and stored at -80°C until analysis. Mouse IgG concentrations were determined using the LSBio Mouse IgG ELISA Kit (Catalog No. LS-F10451) according to the manufacturer's instructions.

Spatial transcriptomics

Freshly dissected draining lymph nodes were dryed on absorbent paper, embedded in OCT, snap frozen in isopentane over dry ice, and stored at -80°C until processing. On the day of the experiment, 10 µm-thick cryosections were prepared from the region of interest using a cryostat. Four sections separated by 50µm in depth were placed on a Visium slide within the designated capture areas. Hematoxylin and eosin (H&E) staining was performed according to the manufacturer's guidelines with the following modifications: hematoxylin incubation for 30 seconds, bluing buffer for 5 seconds, and eosin incubation for 1 minute 30 seconds, all at room temperature. Following washing, brightfield images of the stained sections were acquired before proceeding to subsequent steps of the 10x Genomics Visium Spatial Gene Expression protocol. The tissue was embedded in OCT medium and a 10 µm section was cut on a cryostat and deposited on the capture area of the Visium slide following the guidelines of the 10x Genomics Visium Spatial Gene Expression protocol. Briefly, we performed permeabilization (18 min), reverse transcription, second strand synthesis, denaturation, cDNA amplification (16 cycles of PCR), and library construction according to the manufacturer's instructions. The resulting libraries were sequenced on an Illumina NextSeq2000, generating an average of 120 million reads per library (Read 1: 28 cycles, Read i7: 10 cycles, Read i5: 10 cycles, Read 2: 79 cycles).

Confocal microscopy

Draining lymph nodes were harvested from immunized mice and fixed in antigen fix solution (DiaPath, ref. P0014) for 2 hours at 4°C. Samples were subsequently washed in phosphate-buffered saline (PBS) and cryoprotected in 30% sucrose overnight. The tissue was then embedded in optimal cutting temperature (OCT) compound and snap-frozen in isopentane. Cryosections (20 µm thickness) were prepared using a cryostat and stored at -20°C until staining. For immunofluorescence staining, sections were rehydrated in 1X PBS for 10 minutes. Non-specific binding was blocked by incubating sections for 30 minutes at room temperature in a blocking solution containing 0.1% Triton X-100, 1% fetal calf serum (FCS), 1% bovine serum albumin (BSA), and 1% serum from the host species of the secondary antibody in 1X PBS. Primary antibodies (see table below) were diluted in blocking solution, and sections were incubated overnight at 4°C. Following washing, slides were mounted using ProLong Gold antifade reagent (Invitrogen, ref. P36930). Confocal and spectral images were acquired using a Zeiss LSM 980 confocal microscope. Image processing for conventional analysis was performed using Zen software, while quantitative analysis was conducted using *QuPath* software (see "*Quantitative analysis of confocal microscopy images*" section for details).

mAb	Clone	Manufacturer	Reference
Lyve1 – eF570	ALY7	Thermofisher	41-0443-82
Ki67 – eF660	SolA15	Thermofisher	50-5698-82
IRF4 – AF594	IRF4.3E4	Biolegend	646409
CD3 – eF450	17A2	Thermofisher	48-0032-82
CD21/35 – AF700	7E9	Biolegend	123431
IgD – SNIR 685	11-26c.2a	Biolegend	405749
PDGFR1 - PE	18A2	Biolegend	323605
anti-GFP (cross-react eYFP) – AF488		Thermofisher	A-21311
Bcl6 – AF647	IG191E/A8	Biolegend	648306

CD138 – PE	281-2	Biolegend	142504
Rabbit anti-mouse Col IV		Abcam	Ab19808
Donkey anti-rabbit IgG – Cy3		Jackson ImmunoResearch	711-165-152

Lightsheet microscopy

Draining lymph nodes were harvested and fixed overnight in 0.4% paraformaldehyde (Electron Microscopy Science, ref. 15714) diluted in phosphate-buffered saline (PBS). Immunostaining was performed using a mild permeabilization protocol without methanol[56]. In short, after fixation, the organs were washed in PBS and incubated overnight at 4°C in blocking buffer (PBS containing 5% donkey serum, 1% rehydrated milk, and 0.4% Triton X-100). Samples were then incubated for 5 days with primary antibodies (see table below) diluted in PBS containing 1% rehydrated milk, 0.4% Triton X-100, 3% donkey serum, and 3% mouse serum. Following incubation, samples were washed in PBS with 0.4% Triton X-100 for 1 day. Secondary antibody incubation was performed similarly for 5 days, followed by another 1-day wash in PBS. Samples were embedded in 1% low-melting agarose and subsequently dehydrated through a graded methanol series (20%, 40%, 60%, 80%, and 2× 100% in PBS) for 1 hour per concentration at room temperature. Lymph nodes were then incubated overnight in 100% dehydrated methanol at room temperature. Clearing was initiated by incubating samples in a 1:1 mixture of methanol and BABB (benzyl alcohol and benzyl benzoate at a 1:2 ratio, Sigma ref. 305197 and Fisher Scientific ref. 10654752), followed by overnight incubation in pure BABB. Finally, samples were transferred to ethyl cinnamate for storage until imaging. Imaging was performed using a LaVision Ultramicroscope II (Miltenyi Biotec). Image stacks were acquired with a step size of 6 µm at x2 magnification using an optic zoom with a numerical aperture of 0.5. Three-dimensional reconstruction and analysis of image stacks were conducted using IMARIS software (Version 9.1.0, Bitplane).

mAb	Clone	Manufacturer	Reference
Chicken anti-GFP	-	AVES	GFP-1020
Rabbit anti-mouse Lyve1	-	Reliatech	103-PA50
Donkey anti-chicken IgY – Cy3	-	Jackson Immunoresearch	703-166-155
Donkey anti rabbit IgG – AF790	-	Jackson Immunoresearch	711-655-152
CD4 – AF488	GK1.5	eBioscience	53-0041-82
IRF4 – AF647	IRF4.3E4	Biolegend	646409

COMPUTATIONAL METHODS

Flow cytometry data analysis

The FlowJo UMAP extension was downloaded from FlowJo Exchange Website and used according to the default parameters (euclidean distance, nearest neighbors = 15, minimum distance = 0.5, number of components = 2). The UMAP presented in Figure 3b was computed based on the following compensated parameters: IRF4, Bst2, CD38, GL7, CD138, CD19, PSGL1, Bcl6, B220, CCR6.

Pre-processing of scRNA-seq datasets

<u>10x Genomics 5'-end sequencing</u>

Raw fastq files from gene expression libraries were processed using *Cell Ranger* and *Cell Ranger VDJ* (v3.0.1 for dataset #2, v6.1.2 for dataset #3), with alignment on the mm10 and vdj_GRCm38_alts_ensembl-7.0.0 reference genomes, respectively. Quality control was performed on each dataset independently to remove poor quality cells based on UMI counts, number of genes detected, percentage of transcripts from mitochondrial genes, and percentage of transcripts from ribosomal protein coding genes. For each cell, gene expression UMI count values were log-normalized with Seurat *NormalizeData* with a scale factor of 10,000 to generate normalized UMI count matrices. HTO barcodes for sample demultiplexing after hashing were counted using CITE-seq-count and were normalized for each cell using a centered log ratio (CLR) transformation across cells implemented in the Seurat function *NormalizeData*. Hashtags were demultiplexed using *MULTIseqDemux* function and barcodes assigned as doublets or negative were excluded from further analysis.

484 FB5P-seq

485 We used a custom bioinformatics pipeline to process fastq files and generate single-cell gene expression
486 matrices and BCR sequence files as previously described[35]. Detailed instructions for running the FB5P-seq
487 bioinformatics pipeline can be found at https://github.com/MilpiedLab/FB5P-seq. Quality control was
488 performed on each dataset independently to remove poor quality cells based on UMI counts, number of
489 genes detected, ERCC spike-in quantification accuracy, and percentage of transcripts from mitochondrial
490 genes. For each cell, gene expression UMI count values were log-normalized with Seurat *NormalizeData*
491 with a scale factor of 10,000 to generate normalized UMI count matrices.

492 Index-sorting FCS files were visualized in FlowJo software and compensated parameters values were
493 exported in CSV tables for further processing. For visualization on linear scales in the R programming
494 software, we applied the hyperbolic arcsine transformation on fluorescence parameters[57].

495 For BCR sequence reconstruction, the outputs of the FB5P-seq pipeline were further processed and filtered
496 with custom R scripts. For each cell, reconstructed contigs corresponding to the same V(D)J rearrangement
497 were merged, keeping the largest sequence for further analysis. We discarded contigs with no constant
498 region identified in-frame with the V(D)J rearrangement. In cases where several contigs corresponding to the
499 same BCR chain had passed the above filters, we retained the contig with the highest expression level. BCR
500 metadata from the *MigMap* and *Blastn* annotations were appended to the gene expression and index sorting
501 metadata for each cell.

502 *Non-supervised analysis*

503 For the analysis described in Figure 2, the three datasets after quality control and normalization were
504 integrated in *Seurat* using *SelectIntegrationFeatures* (2000 features, excluding BCR coding genes),
505 *FindIntegrationAchors*, and *IntegrateData* (30 components). After integration, we used *RunPCA* (30
506 components), *FindNeighbors* (30 components) and *FindClusters* (resolution 0.3) for non-supervised
507 clustering, and *RunUMAP* (30 components) for visualization. Marker genes were computed with
508 *FindAllMarkers* (Wilcoxon assay) and top markers visualized as a *Seurat* dot plot. For the analysis described
509 in Figure 4, we used a standard *Seurat* analysis workflow, including cell cycle regression.

510 *Supervised annotation*

511 Supervised annotation of scRNA-seq datasets were performed as described in Figure 2d and Supplementary
512 Figure 2c. We used the *AddModuleScore* function to compute gene expression scores for every cell in the
513 datasets for the indicated signatures (**Supplementary Table 1**). For DZ and LZ signatures, genes associated
514 to cell cycle ontologies (based on GO terms), were removed from the gene lists prior to scoring, as described
515 in Milpied et al.[58]. Thresholds for "gating" (**Supplementary Figure 1c**) were defined empirically to optimize
516 concordance between supervised annotation and non-supervised clustering. The continuum score in Figure
517 4f was computed by fitting a linear regression on the distribution of B cells in the scatter plot of ER stress
518 score (x-axis) and Normalized Ig transcripts UMI counts (y-axis), and projecting data points on the regression
519 line.

Gene ontology

The evolution profile of genes along the differentiation continuum was computed by applying a kernel smoother on the distribution of data, for each gene expressed in more than 25% cells in the dataset (n=1834 genes after excluding mitochondrial and ribosomal genes). The derivatives of the resulting profiles were hierarchically clustered (pearson's correlation distance) to define groups of co-evolving genes in the continuum. Lists of genes grouped in distinct evolution clusters were then submitted to gene ontology analysis using *gprofiler2*, with default settings, and computing only the results for the "BP: biological processes" categories. The results of multiple gene clusters were appended in a single plot using *ggplot2*.

BCR-seq analysis

Single-cell BCR-seq data were further analyzed with Change-O[59] to compute clonotypes and annotate somatic hypermutation load. The intersection between cell clonotype identities and supervised annotations based on gene expression was used to compute the clonal compositions displayed in Figure 5j.

Spatial Transcriptomics analysis

Spatial transcriptomics FASTQ files were aligned to the mm10 reference genome using *SpaceRanger* v1.3.1. Downstream analysis was performed using the *Seurat* R package v4.2.1, employing log-normalization and the Louvain clustering algorithm. Annotation of the four merged slices was conducted iteratively, beginning with a clustering resolution of 0.1 (50 principal components) to identify main areas (B cell zone, T cell zone, and medulla), followed by a resolution of 0.2 (50 principal components) to reveal germinal centers. GC spots were further subclustered for light zone (LZ) and dark zone (DZ) segregation (resolution 0.2, 30 principal components). Border regions were annotated based on neighboring spots, defining GC-T zone interface (GTI), GC-medulla interface (GMI), and GC-B zone interface (GBI) based on neighbor spots annotation. Annotations were projected onto a UMAP calculated using 50 principal components from the complete dataset. Cell type-specific genes were identified using *FindAllMarkers* (default parameters) and filtered for adjusted p-value < 0.05. For stromal cell signature analysis, gene lists were derived from Rodda et al.[41] (log2FoldChange > 0.5, FDR < 0.05, pct_in > 0.10).

Quantitative analysis of confocal microscopy images

For cell detection and quantification in confocal microscopy images (Figure 6), we used the *QuPath* analysis platform. We annotated manually the GC and medulla areas based on the eYFP staining for the GC and on the IRF4 staining for the medulla. We defined the GC-medulla interface (GMI) as a 50µm-wide border surrounding the GC and intersecting the medulla annotation. We imported a custom script (*Cellpose2coloc*) in *QuPath*[60] to segment and identify cell subsets: we first generated cell segmentation masks based on the different channels selected; then we determined the positivity thresholds for defining cell subsets, and used them for colocalization analysis. We exported the measurements and performed quantification analyses in GraphPad Prism. All scripts are publicly available here: https://github.com/Imagimm-CIML/Detection-of-B-cell-subtypes-in-a-draining-lymph-node-using-a-mask-colocalization-approach?tab=readme-ov-file#readme.

Statistical analyses

Statistical analyses were performed using Graphpad Prism or R softwares with tests and p-value significance criteria detailed in the figure legends.

Data and code availability

Single-cell RNA-seq data have been deposited on NCBI GEO under accession numbers XXXXX. Spatial Transcriptomics data have been deposited on NCBI GEO under accession number XXXXX. Annotated scRNA-seq and spatial transcriptomics objects, as well as all code used to perform analyses will be available upon request to the corresponding author.

ACKNOWLEDGEMENTS

We are grateful to all past and present members of the “Integrative B cell Immunology” lab at Centre d’Immunologie de Marseille-Luminy (CIML) for useful discussions, to Marc Bajénoff for stimulating discussions on lymph node stromal cells, and to Hugues Lelouard for his expertise on spectral confocal microscopy. We acknowledge the Computational Biology, Biostatistics and Modeling (CB2M) hub at CIML for helpful discussions and comments. We thank the Flow Cytometry Core Facility, and the Imagimm facility of CIML. We thank the animal care facility of CIML. We acknowledge Centre de Calcul Intensif d’Aix-Marseille for granting access to its high-performance computing resources. This work was supported by the French National Research Agency through the « Investments for the Future » program (France-BioImaging, ANR-10-INBS-04). This work was supported by grants from ANR (ANR-17-CE15-0009-01 “MoDEx-GC” and ANR-23-CE15-0025-01 “GCselection”) to P.M. This work was supported by institutional grants from INSERM, CNRS and Aix-Marseille University to the CIML. L.B. and N.A. were supported by fellowships from the French Ministry of Research and Higher Education and Fondation pour la Recherche Médicale.

AUTHORSHIP CONTRIBUTIONS

L.B. designed experiments, performed experiments, analyzed the data, performed some of the bioinformatics analyses, prepared the figures and wrote the manuscript. C.D. performed most bioinformatics analyses and prepared the figures. N.A. designed and performed experiments at the initiation of the project. L.G. performed single-cell and spatial transcriptomics experiments. M.F. and T.B. designed and performed the quantitative analysis of confocal microscopy images. B.E. performed the bioinformatics analysis of the spatial transcriptomics data. L.C. performed the spatial transcriptomics experiment. C.S. and S.vdP. supervised the light-sheet microscopy staining, data acquisition, and data analyses. J.M.N. produced hashtag antibodies. P.M. designed experiments, performed experiments, analyzed the data, prepared the figures, wrote the manuscript, acquired funding and supervised the study. All authors revised and approved the manuscript.

DISCLOSURE OF CONFLICTS OF INTEREST

The authors declare no competing financial interests.

591 **REFERENCES**

592 1. Victora GD, Nussenzweig MC. Germinal Centers. *Annu. Rev. Immunol.* 2022;40(1):413–442.
593 2. Smith KGC. The extent of affinity maturation differs between the memory and antibody-forming cell
594 compartments in the primary immune response. *The EMBO Journal*. 1997;16(11):2996–3006.
595 3. Phan TG, Paus D, Chan TD, et al. High affinity germinal center B cells are actively selected into the
596 plasma cell compartment. *The Journal of Experimental Medicine*. 2006;203(11):2419–2424.
597 4. Victora GD, Schwickert TA, Fooksman DR, et al. Germinal Center Dynamics Revealed by Multiphoton
598 Microscopy with a Photoactivatable Fluorescent Reporter. *Cell*. 2010;143(4):592–605.
599 5. Robinson MJ, Ding Z, Dowling MR, et al. Intrinsically determined turnover underlies broad heterogeneity
600 in plasma-cell lifespan. *Immunity*. 2023;S1074761323001838.
601 6. Alinikula J, Nera K-P, Junttila S, Lassila O. Alternate pathways for Bcl6-mediated regulation of B cell to
602 plasma cell differentiation. *Eur. J. Immunol.* 2011;41(8):2404–2413.
603 7. Niu H, Ye BH, Dalla-Favera R. Antigen receptor signaling induces MAP kinase-mediated phosphorylation
604 and degradation of the BCL-6 transcription factor. *Genes & Development*. 1998;12(13):1953–1961.
605 8. Heise N, De Silva NS, Silva K, et al. Germinal center B cell maintenance and differentiation are controlled
606 by distinct NF-κB transcription factor subunits. *Journal of Experimental Medicine*. 2014;211(10):2103–
607 2118.
608 9. Klein U, Casola S, Cattoretti G, et al. Transcription factor IRF4 controls plasma cell differentiation and
609 class-switch recombination. *Nat Immunol*. 2006;7(7):773–782.
610 10. O'Connor BP, Cascalho M, Noelle RJ. Short-lived and Long-lived Bone Marrow Plasma Cells Are Derived
611 from a Novel Precursor Population. *The Journal of Experimental Medicine*. 2002;195(6):737–745.
612 11. Paus D, Phan TG, Chan TD, et al. Antigen recognition strength regulates the choice between
613 extrafollicular plasma cell and germinal center B cell differentiation. *Journal of Experimental Medicine*.
614 2006;203(4):1081–1091.
615 12. Taylor JJ, Pape KA, Steach HR, Jenkins MK. Apoptosis and antigen affinity limit effector cell
616 differentiation of a single naïve B cell. *Science*. 2015;347(6223):784–787.
617 13. Zhang Y, Garcia-Ibanez L, Toellner K-M. Regulation of germinal center B-cell differentiation. *Immunol*
618 *Rev*. 2016;270(1):8–19.
619 14. Basso K, Klein U, Niu H, et al. Tracking CD40 signaling during germinal center development. *Blood*.
620 2004;104(13):4088–4096.
621 15. Saito M, Gao J, Basso K, et al. A Signaling Pathway Mediating Downregulation of BCL6 in Germinal
622 Center B Cells Is Blocked by BCL6 Gene Alterations in B Cell Lymphoma. *Cancer Cell*. 2007;12(3):280–
623 292.
624 16. Ise W, Fujii K, Shiroguchi K, et al. T Follicular Helper Cell-Germinal Center B Cell Interaction Strength
625 Regulates Entry into Plasma Cell or Recycling Germinal Center Cell Fate. *Immunity*. 2018;48(4):702-
626 715.e4.
627 17. Diehl SA, Schmidlin H, Nagasawa M, et al. STAT3-Mediated Up-Regulation of BLIMP1 Is Coordinated
628 with BCL6 Down-Regulation to Control Human Plasma Cell Differentiation. *J Immunol*.
629 2008;180(7):4805–4815.
630 18. Kwon H, Thierry-Mieg D, Thierry-Mieg J, et al. Analysis of Interleukin-21-Induced Prdm1 Gene Regulation
631 Reveals Functional Cooperation of STAT3 and IRF4 Transcription Factors. *Immunity*. 2009;31(6):941–
632 952.
633 19. Kallies A, Hasbold J, Fairfax K, et al. Initiation of Plasma-Cell Differentiation Is Independent of the
634 Transcription Factor Blimp-1. *Immunity*. 2007;26(5):555–566.
635 20. Ding BB, Bi E, Chen H, Yu JJ, Ye BH. IL-21 and CD40L Synergistically Promote Plasma Cell
636 Differentiation through Upregulation of Blimp-1 in Human B Cells. *J.I.* 2013;190(4):1827–1836.
637 21. Nera K-P, Kohonen P, Narvi E, et al. Loss of Pax5 Promotes Plasma Cell Differentiation. *Immunity*.
638 2006;24(3):283–293.
639 22. Sciammas R, Li Y, Warmflash A, et al. An incoherent regulatory network architecture that orchestrates B
640 cell diversification in response to antigen signaling. *Mol Syst Biol*. 2011;7:495.
641 23. Barwick BG, Scharer CD, Bally APR, Boss JM. Plasma cell differentiation is coupled to division-
642 dependent DNA hypomethylation and gene regulation. *Nat Immunol*. 2016;17(10):1216–1225.
643 24. Kräutler NJ, Suan D, Butt D, et al. Differentiation of germinal center B cells into plasma cells is initiated
644 by high-affinity antigen and completed by Tfh cells. *The Journal of Experimental Medicine*.
645 2017;214(5):1259–1267.
646 25. Nakagawa R, Toboso-Navasa A, Schips M, et al. Permissive selection followed by affinity-based
647 proliferation of GC light zone B cells dictates cell fate and ensures clonal breadth. *PNAS*. 2021;118(2):.

26. Holmes AB, Corinaldesi C, Shen Q, et al. Single-cell analysis of germinal-center B cells informs on lymphoma cell of origin and outcome. *Journal of Experimental Medicine*. 2020;217(10):e20200483.
27. King HW, Orban N, Riches JC, et al. Single-cell analysis of human B cell maturation predicts how antibody class switching shapes selection dynamics. *Sci. Immunol.* 2021;6(56):eabe6291.
28. 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.
29. Meyer-Hermann M, Mohr E, Pelletier N, et al. A Theory of Germinal Center B Cell Selection, Division, and Exit. *Cell Reports*. 2012;2(1):162–174.
30. Schwab SR, Cyster JG. Finding a way out: lymphocyte egress from lymphoid organs. *Nature Immunology*. 2007;8(12):1295–1301.
31. Nutt SL, Hodgkin PD, Tarlinton DM, Corcoran LM. The generation of antibody-secreting plasma cells. *Nat Rev Immunol*. 2015;15(3):160–171.
32. Mohr E, Serre K, Manz RA, et al. Dendritic Cells and Monocyte/Macrophages That Create the IL-6/APRIL-Rich Lymph Node Microenvironments Where Plasmablasts Mature. *J Immunol*. 2009;182(4):2113–2123.
33. Sprumont A, Rodrigues A, McGowan SJ, Bannard C, Bannard O. Germinal centers output clonally diverse plasma cell populations expressing high- and low-affinity antibodies. *Cell*. 2023;186(25):5486-5499.e13.
34. Le Gallou S, Nojima T, Kitamura D, Weill J-C, Reynaud C-A. The AID-Cre-ERT2 Model: A Tool for Monitoring B Cell Immune Responses and Generating Selective Hybridomas. *Methods Mol. Biol.* 2017;1623:243–251.
35. Attaf N, Cervera-Marzal I, Dong C, et al. FB5P-seq: FACS-Based 5-Prime End Single-Cell RNA-seq for Integrative Analysis of Transcriptome and Antigen Receptor Repertoire in B and T Cells. *Front. Immunol.* 2020;11:.
36. Laidlaw BJ, Duan L, Xu Y, Vazquez SE, Cyster JG. The transcription factor Hhex cooperates with the corepressor Tle3 to promote memory B cell development. *Nature Immunology*. 2020;21(9):1082–1093.
37. Kennedy DE, Okoreeh MK, Maienschein-Cline M, et al. Novel specialized cell state and spatial compartments within the germinal center. *Nature Immunology*. 2020;1–11.
38. Piskurich JF, Lin K-I, Lin Y, et al. BLIMP-1 mediates extinction of major histocompatibility class II transactivator expression in plasma cells. *Nat Immunol*. 2000;1(6):526–532.
39. Attaf N, Baaklini S, Binet L, Milpied P. Heterogeneity of germinal center B cells: New insights from single-cell studies. *European Journal of Immunology*. 2021;51(11):2555–2567.
40. Sze DM-Y, Toellner K-M, de Vinuesa CG, Taylor DR, MacLennan ICM. Intrinsic Constraint on Plasmablast Growth and Extrinsic Limits of Plasma Cell Survival. *The Journal of Experimental Medicine*. 2000;192(6):813–822.
41. Rodda LB, Lu E, Bennett ML, et al. Single-Cell RNA Sequencing of Lymph Node Stromal Cells Reveals Niche-Associated Heterogeneity. *Immunity*. 2018;48(5):1014-1028.e6.
42. Zhang Y, Tech L, George LA, et al. Plasma cell output from germinal centers is regulated by signals from Tfh and stromal cells. *The Journal of Experimental Medicine*. 2018;215(4):1227–1243.
43. Suan D, Kräutler NJ, Maag JLV, et al. CCR6 Defines Memory B Cell Precursors in Mouse and Human Germinal Centers, Revealing Light-Zone Location and Predominant Low Antigen Affinity. *Immunity*. 2017;47(6):1142-1153.e4.
44. Jellusova J. Metabolic control of B cell immune responses. *Current Opinion in Immunology*. 2020;63:21–28.
45. Corinaldesi C, Holmes AB, Shen Q, et al. Tracking Immunoglobulin Repertoire and Transcriptomic Changes in Germinal Center B Cells by Single-Cell Analysis. *Frontiers in Immunology*. 2022;12:.
46. Massoni-Badosa R, Aguilar-Fernández S, Nieto JC, et al. An atlas of cells in the human tonsil. *Immunity*. 2024;0(0):
47. Alaterre E, Ovejero S, Bret C, et al. Integrative single-cell chromatin and transcriptome analysis of human plasma cell differentiation. *Blood Journal*. 2024;blood.2023023237.
48. Long Z, Phillips B, Radtke D, Meyer-Hermann M, Bannard O. Competition for refueling rather than cyclic reentry initiation evident in germinal centers. *Science Immunology*. 7(69):eabm0775.
49. Takeda A, Salmi M, Jalkanen S. Lymph node lymphatic endothelial cells as multifaceted gatekeepers in the immune system. *Trends in Immunology*. 2022;S1471490622002356.
50. Huang H-Y, Rivas-Caicedo A, Renevey F, et al. Identification of a new subset of lymph node stromal cells involved in regulating plasma cell homeostasis. *Proceedings of the National Academy of Sciences*. 2018;115(29):E6826–E6835.

51. Pikor NB, Mörbe U, Lütge M, et al. Remodeling of light and dark zone follicular dendritic cells governs germinal center responses. *Nat Immunol*. 2020;21(6):649–659.
52. Lütge M, De Martin A, Gil-Cruz C, et al. Conserved stromal–immune cell circuits secure B cell homeostasis and function. *Nat Immunol*. 2023;24(7):1149–1160.
53. Takeuchi A, Ozawa M, Kanda Y, et al. A Distinct Subset of Fibroblastic Stromal Cells Constitutes the Cortex-Medulla Boundary Subcompartment of the Lymph Node. *Front. Immunol.* 2018;9:.
54. Dubey LK, Karempudi P, Luther SA, Ludewig B, Harris NL. Interactions between fibroblastic reticular cells and B cells promote mesenteric lymph node lymphangiogenesis. *Nat Commun*. 2017;8(1):367.
55. Mimitou EP, Cheng A, Montalbano A, et al. Expanding the CITE-seq tool-kit: Detection of proteins, transcriptomes, clonotypes and CRISPR perturbations with multiplexing, in a single assay. *Nat Methods*. 2019;16(5):409–412.
56. Siret C, van de Pavert SA. Three-Dimensional Imaging of Macrophages in Complete Organs. *Tissue-Resident Macrophages*. 2024;2713:297–306.
57. Finak G, Perez J-M, Weng A, Gottardo R. Optimizing transformations for automated, high throughput analysis of flow cytometry data. *BMC Bioinformatics*. 2010;11(1):546.
58. Milpied P, Cervera-Marzal I, Mollichella M-L, et al. Human germinal center transcriptional programs are de-synchronized in B cell lymphoma. *Nat Immunol*. 2018;19(9):1013–1024.
59. Gupta NT, Vander Heiden JA, Uduman M, et al. Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. *Bioinformatics*. 2015;31(20):3356–3358.
60. Bankhead P, Loughrey MB, Fernández JA, et al. QuPath: Open source software for digital pathology image analysis. *Sci Rep*. 2017;7(1):16878.

FIGURES AND LEGENDS

Figure 1

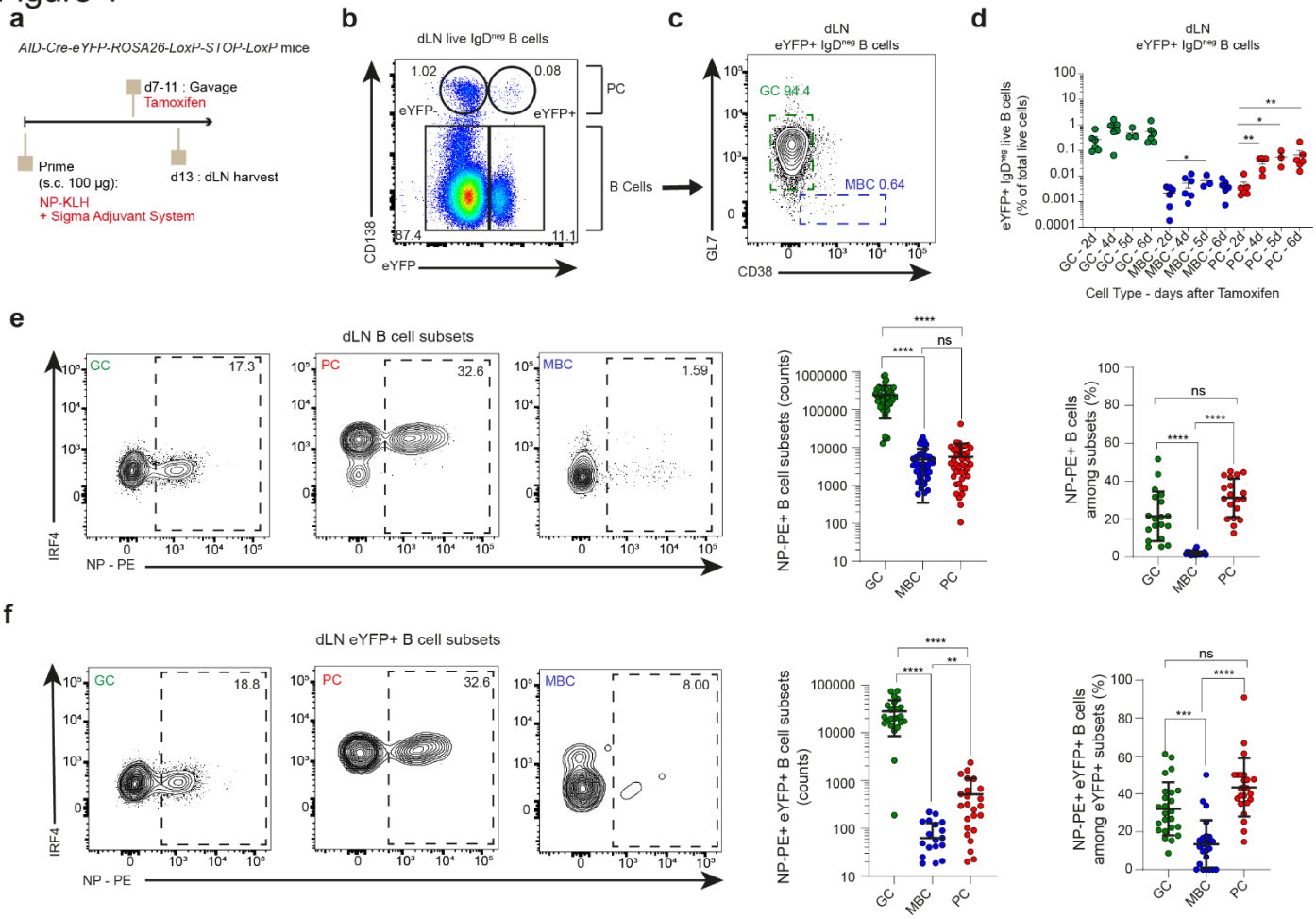
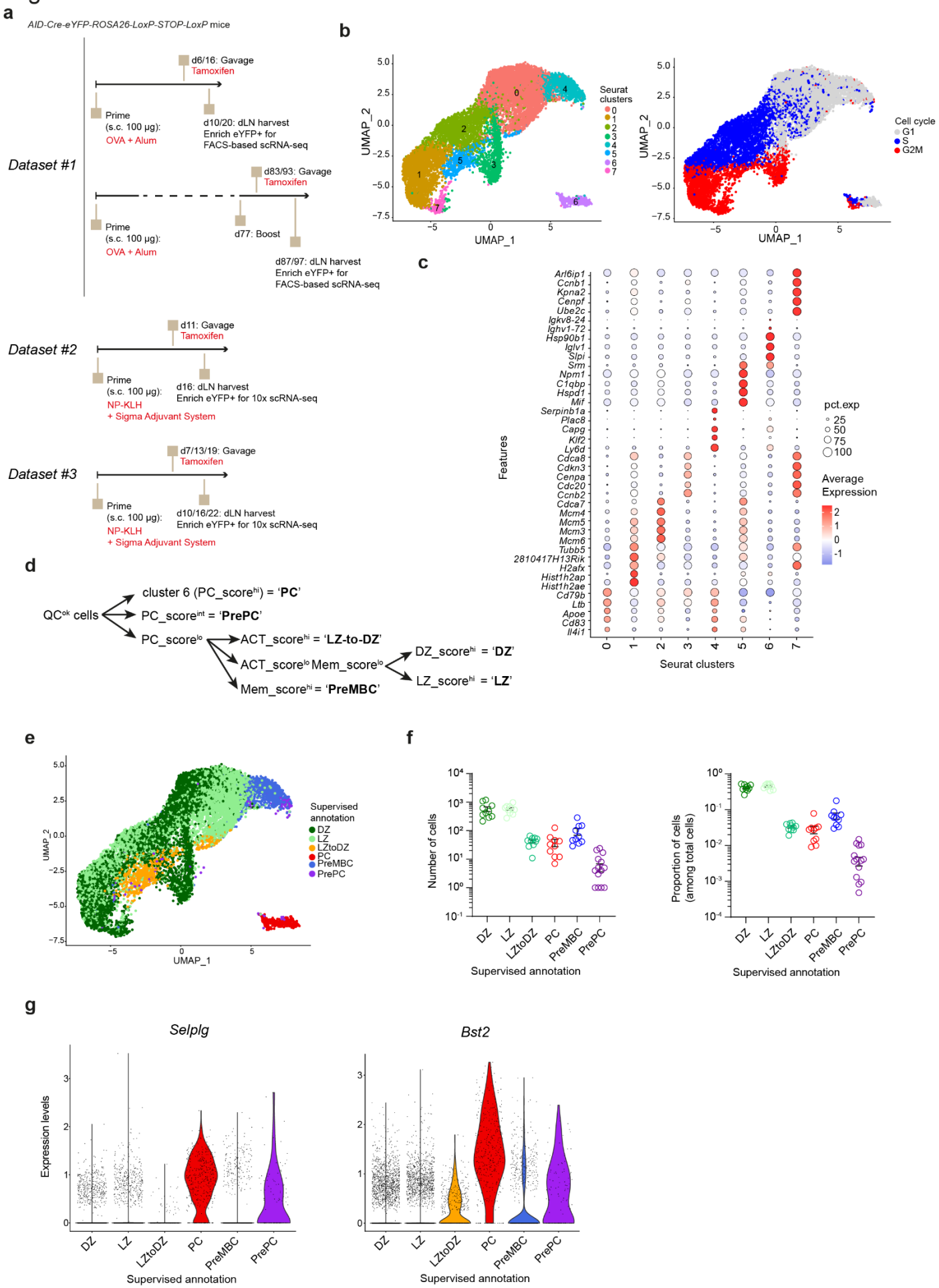


Figure 1: Tracking GC B cells and recent GC emigrants after model immunization.

a, Experimental design. **b**, Flow cytometry gating of eYFP$^+$ and eYFP$^-$ PC (CD138^{hi}) and B cells among IgD^{neg} B cells from dLN in mice gavaged 4 days before sacrifice. **c**, Flow cytometry gating of GC B cells and MBC among eYFP$^+$ B cells. Numbers indicate % of parent gate. **d**, Percentage of the indicated eYFP$^+$ IgD^{neg} B cell types (GC, MBC or PC) among total live cells in dLN at the indicated times after tamoxifen gavage (2, 4, 5, 6 days). **e**, Flow cytometry gating of NP-specific cells among GC, PC and MBC (numbers indicate % of parent gate) (left). Cell counts of NP-specific GC, MBC and PC in dLN (middle). Percentage of NP-specific cells among the indicated cell types (right). **f**, Flow cytometry gating of NP-specific cells among eYFP$^+$ GC, PC and MBC (numbers indicate % of parent gate) (left). Cell counts of eYFP$^+$ NP-specific GC, MBC and PC types in dLN (middle). Percentage of NP-specific cells among the indicated eYFP$^+$ cell types (right). Symbols in panels **d**, **e** and **f** represent individual mice. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$ in Wilcoxon matched pairs signed rank test.

Figure 2


a
AID-Cre-eYFP-ROSA26-LoxP-STOP-LoxP mice
Dataset #1
d6/16: Gavage
Tamoxifen
Prime
(s.c. 100 µg):
OVA + Alum
d10/20: dLN harvest
Enrich eYFP+ for
FACS-based scRNA-seq
d83/93: Gavage
Tamoxifen
Prime
(s.c. 100 µg):
OVA + Alum
d77: Boost
d87/97: dLN harvest
Enrich eYFP+ for
FACS-based scRNA-seq
Dataset #2
d11: Gavage
Tamoxifen
Prime
(s.c. 100 µg):
NP-KLH
+ Sigma Adjuvant System
d16: dLN harvest
Enrich eYFP+ for 10x scRNA-seq
Dataset #3
d7/13/19: Gavage
Tamoxifen
Prime
(s.c. 100 µg):
NP-KLH
+ Sigma Adjuvant System
d10/16/22: dLN harvest
Enrich eYFP+ for 10x scRNA-seq
b
UMAP_2
UMAP_1
Seurat clusters
Cell cycle
G1
S
G2M
c
Features
Seurat clusters
pct.exp
Average Expression
d
QC^ok cells
cluster 6 (PC_score^hi) = 'PC'
PC_score^int = 'PrePC'
PC_score^lo
ACT_score^hi = 'LZ-to-DZ'
ACT_score^lo Mem_score^lo
Mem_score^hi = 'PreMBC'
DZ_score^hi = 'DZ'
LZ_score^hi = 'LZ'
e
Supervised annotation
DZ
LZ
LZtoDZ
PC
PreMBC
PrePC
f
Number of cells
Proportion of cells (among total cells)
Supervised annotation
g
Selplg
Bst2
Expression levels
Supervised annotation

743

744 **Figure 2**: Single-cell RNA-seq analysis of GC B cells and recent GC emigrants identifies putative prePC.

745 **a**, Experimental designs for the generation of three distinct scRNA-seq datasets of GC and GC-derived cells
746 in dLN. **b**, UMAP representation of single-cell gene expression profiles of GC and GC-derived cells integrated
747 from the three datasets, colored by unsupervised Louvain clusters (left) or by cell cycle status (right). **c**, dot
748 plot of the expression of top 5 markers of each unsupervised cluster, as indicated. The percentage of cells
749 from a cluster expressing a marker is coded by the size of the circles, the average expression level by the
750 color. **d**, Hierarchical gating strategy for gene expression score-based supervised annotation of GC and GC-
751 derived cells. **e**, UMAP representation colored by supervised annotations. **f**, Total counts (left) and proportion
752 among total cells (right) of cells from the indicated supervised annotation. Each symbol represents a distinct
753 sample from the three datasets. **g**, Violin plots of gene expression levels of *Selplg* (left) and *Bst2* (right)
754 among cells from the indicated supervised annotations.

755

Figure 3

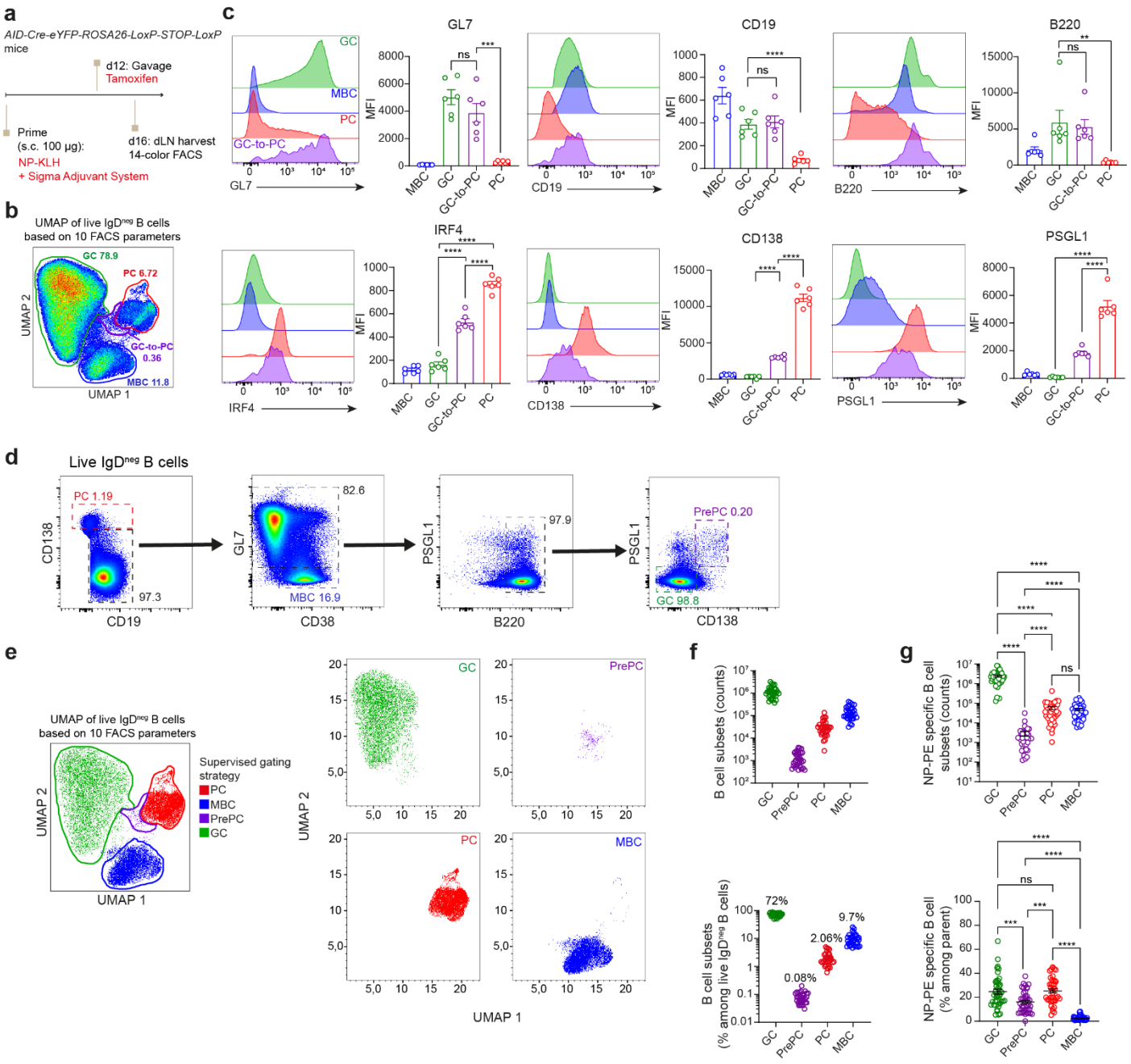
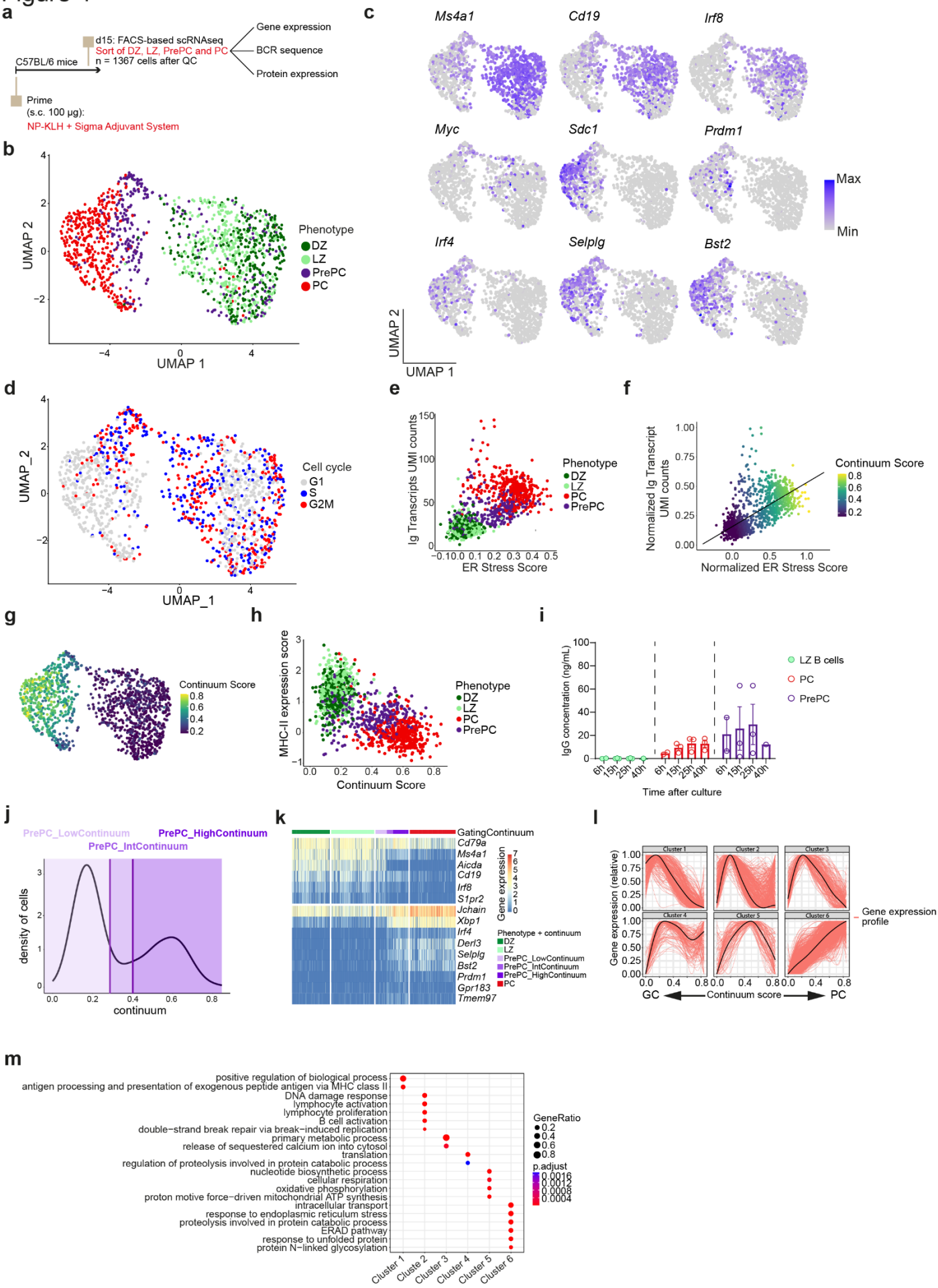


Figure 3: Enrichment of putative prePC by flow cytometry.

a, Experimental design. **b**, UMAP representation of flow cytometry-based protein expression profiles of dLN IgD^{neg} B cells, pseudocolored by embedded cell density. Clusters of GC, PC, MBC and intermediate GC-to-PC are circled and annotated. **c**, Mean Fluorescence Intensity (MFI) histogram and quantification for the indicated GC (GL7, CD19, B220 ; top panels) and PC (IRF4, CD138, PSGL1 ; bottom panels) protein markers in the indicated cell subsets. **d**, Gating strategy to enrich for putative PrePC by flow cytometry. **e**, UMAP projection as in b, representing only the cells from the subsets gated in d (left) and then split by subset (right). **f**, Total counts (top) and proportion among dLN IgD^{neg} B cells (bottom) of cells from the indicated supervised gates defined in d and e. **g**, Total counts (top) and proportion among dLN IgD^{neg} B cells (bottom) of NP-specific cells from the indicated supervised gates defined in d and e. Each symbol in panels c, f and g represent a distinct mouse. *** $p < 0.001$, **** $p < 0.0001$ in Wilcoxon matched pairs signed rank test.

Figure 4


a
C57BL/6 mice
Prime
(s.c. 100 µg):
NP-KLH + Sigma Adjuvant System
d15: FACS-based scRNAseq
Sort of DZ, LZ, PrePC and PC
n = 1367 cells after QC
Gene expression
BCR sequence
Protein expression
b
UMAP 2
UMAP 1
Phenotype
DZ
LZ
PrePC
PC
c
Ms4a1
Cd19
Irf8
Myc
Sdc1
Prdm1
Irf4
Selplg
Bst2
Max
Min
UMAP 2
UMAP 1
d
UMAP_2
UMAP_1
Cell cycle
G1
S
G2M
e
Ig Transcripts UMI counts
ER Stress Score
Phenotype
DZ
LZ
PC
PrePC
f
Normalized Ig Transcript UMI counts
Normalized ER Stress Score
Continuum Score
g
Continuum Score
h
MHC-II expression score
Continuum Score
Phenotype
DZ
LZ
PC
PrePC
i
IgG concentration (ng/mL)
Time after culture
LZ B cells
PC
PrePC
j
PrePC_LowContinuum
PrePC_IntContinuum
PrePC_HighContinuum
density of cells
continuum
k
GatingContinuum
Cd79a
Ms4a1
Aicda
Cd19
Irf8
S1pr2
Jchain
Xbp1
Irf4
Derl3
Selplg
Bst2
Prdm1
Gpr183
Tmem97
Gene expression
Phenotype + continuum
DZ
LZ
PrePC_LowContinuum
PrePC_IntContinuum
PrePC_HighContinuum
PC
l
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Gene expression (relative)
Gene expression profile
GC
Continuum score
PC
m
positive regulation of biological process
antigen processing and presentation of exogenous peptide antigen via MHC class II
DNA damage response
lymphocyte activation
lymphocyte proliferation
B cell activation
double-strand break repair via break-induced replication
primary metabolic process
release of sequestered calcium ion into cytosol
translation
regulation of proteolysis involved in protein catabolic process
nucleotide biosynthetic process
cellular respiration
oxidative phosphorylation
proton motive force-driven mitochondrial ATP synthesis
intracellular transport
response to endoplasmic reticulum stress
proteolysis involved in protein catabolic process
ERAD pathway
response to unfolded protein
protein N-linked glycosylation
Cluster 1
Cluste 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
GeneRatio
p.adjust

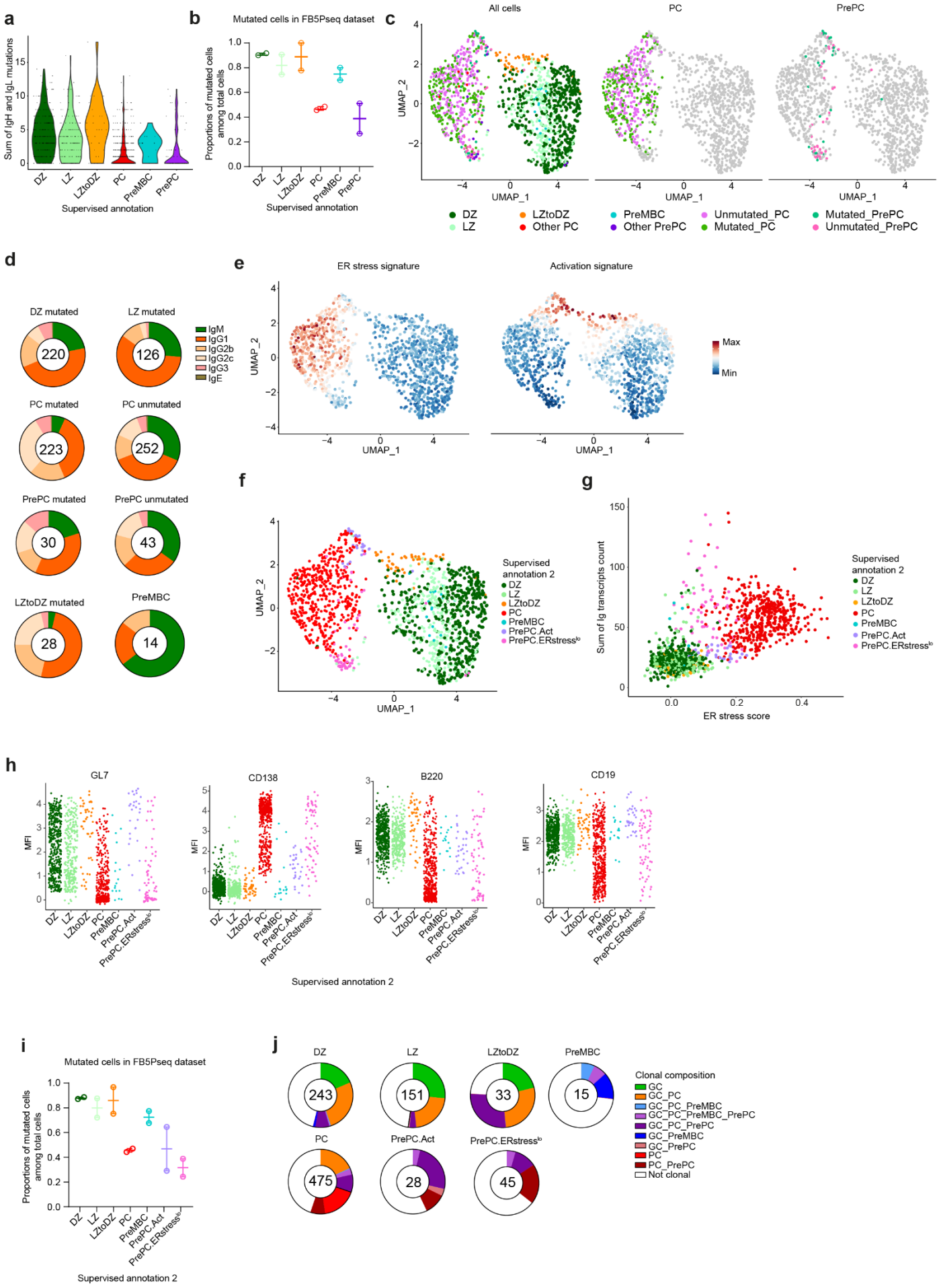
769

770 **Figure 4**: Characterization of prePC in the GC-to-PC differentiation continuum.

771 **a**, Experimental design. **b**, UMAP projection of gene expression profiles of FACS-sorted DZ, LZ, PrePC and
772 PC, after cell cycle regression. Each dot is a cell, colored by the FACS sorting phenotype. **c**, Feature plots
773 showing the expression of the indicated GC (*Ms4a1*, *Cd19*, *Irf8*, *Myc*) and PC (*Sdc1*, *Prdm1*, *Irf4*, *Selplg*,
774 *Bst2*) marker genes in cells laid out as in b. **d**, UMAP projection as in b, with cells colored by cell cycle status.
775 **e**, Scatter plot representation of the ER stress score (x-axis) and the Ig transcripts UMI counts (y-axis) in
776 cells colored by the FACS sorting phenotype. **f**, Scatter plot presented in d, overlayed with the regression
777 line and colored by the GC-to-PC continuum score computed after projecting cells on the regression line and
778 ranking. **g**, UMAP projection as in b, colored by the continuum score. **h**, Scatter plot representation of the
779 continuum score (x-axis) and the MHC-II gene expression score (y-axis) in cells colored by the FACS sorting
780 phenotype. **i**, IgG concentration in culture supernatant of LZ B cells, PC and PrePC at the indicated times
781 after *ex vivo* cell culture. **j**, Cell density histogram distribution of continuum scores for cells in the dataset,
782 indicating the continuum thresholds used for defining low, intermediate and high values for the continuum
783 score in PrePC. **k**, Single-cell gene expression heatmap of the indicated GC (top 6 genes), PC (next 7 genes)
784 and PrePC (bottom 2 genes) markers expression in the different subsets defined by their phenotype and
785 continuum score. **l**, Relative gene expression profile of cells ranked by continuum score, for genes grouped
786 according to their evolution profile (each gene a red line, the average cluster profile a black line). **m**, Results
787 of gene ontology (GO) enrichment analysis for genes in the different evolution clusters defined in l.

788

Figure 5


a
Sum of IgH and IgL mutations
DZ
LZ
LZtoDZ
PC
PreMBC
PrePC
Supervised annotation
b
Mutated cells in FB5Pseq dataset
Proportions of mutated cells among total cells
Supervised annotation
c
All cells
PC
PrePC
UMAP_2
UMAP_1
DZ
LZ
LZtoDZ
Other PC
PreMBC
Other PrePC
Unmutated_PC
Mutated_PC
Mutated_PrePC
Unmutated_PrePC
d
DZ mutated
LZ mutated
PC mutated
PC unmutated
PrePC mutated
PrePC unmutated
LZtoDZ mutated
PreMBC
220
126
223
252
30
43
28
14
IgM
IgG1
IgG2b
IgG2c
IgG3
IgE
e
ER stress signature
Activation signature
Max
Min
f
Supervised annotation 2
DZ
LZ
LZtoDZ
PC
PreMBC
PrePC.Act
PrePC.ERstress^lo
g
Sum of Ig transcripts count
ER stress score
h
GL7
CD138
B220
CD19
MFI
Supervised annotation 2
i
Mutated cells in FB5Pseq dataset
Proportions of mutated cells among total cells
Supervised annotation 2
j
DZ
LZ
LZtoDZ
PreMBC
PC
PrePC.Act
PrePC.ERstress^lo
243
151
33
15
475
28
45
Clonal composition
GC
GC_PC
GC_PC_PreMBC
GC_PC_PreMBC_PrePC
GC_PC_PrePC
GC_PreMBC
GC_PrePC
PC
PC_PrePC
Not clonal

789

Figure 5: BCR analysis of prePC.

a, Violin plot of BCR mutations (sum of IgH and IgL) in the indicated cell subsets from the dataset described in Figure 4. **b**, Proportion of mutated cells among total cells of the indicated subsets. Each dot represents a distinct mouse, bar at median. **c**, UMAP projection as in Figure 4, colored by cell subsets, with PC and PrePC divided into BCR-unmutated and BCR-mutated (left), and split to display only PC (middle) and PrePC (right). **d**, Pie charts of the isotype repartition in the indicated cell subsets, with PC and PrePC divided into BCR-unmutated and BCR-mutated. Numbers indicate the number of cells analyzed for each subset. **e**, UMAP projection as in Figure 4, colored by the expression levels of the ER stress gene expression signature (left) or the activation gene expression signature (right). **f**, UMAP projection as in Figure 4, colored by supervised annotation 2 segragating prePC into prePC.Act and prePC.ERstress^{lo}. **g**, Scatter plot representation of the ER stress score (x-axis) and the Ig transcripts UMI counts (y-axis) in cells colored by supervised annotation 2. **h**, Flow cytometry signal intensity (index sorting) of the indicated surface markers in cells from the indicated subsets (supervised annotation 2). **i**, Proportion of mutated cells among total cells grouped by supervised annotation 2. Each dot represents a distinct mouse, bar at median. **j**, Pie charts of shared clonotype composition within each subset of the supervised annotation 2. Numbers indicate the number of cells analyzed for each subset.

Figure 6

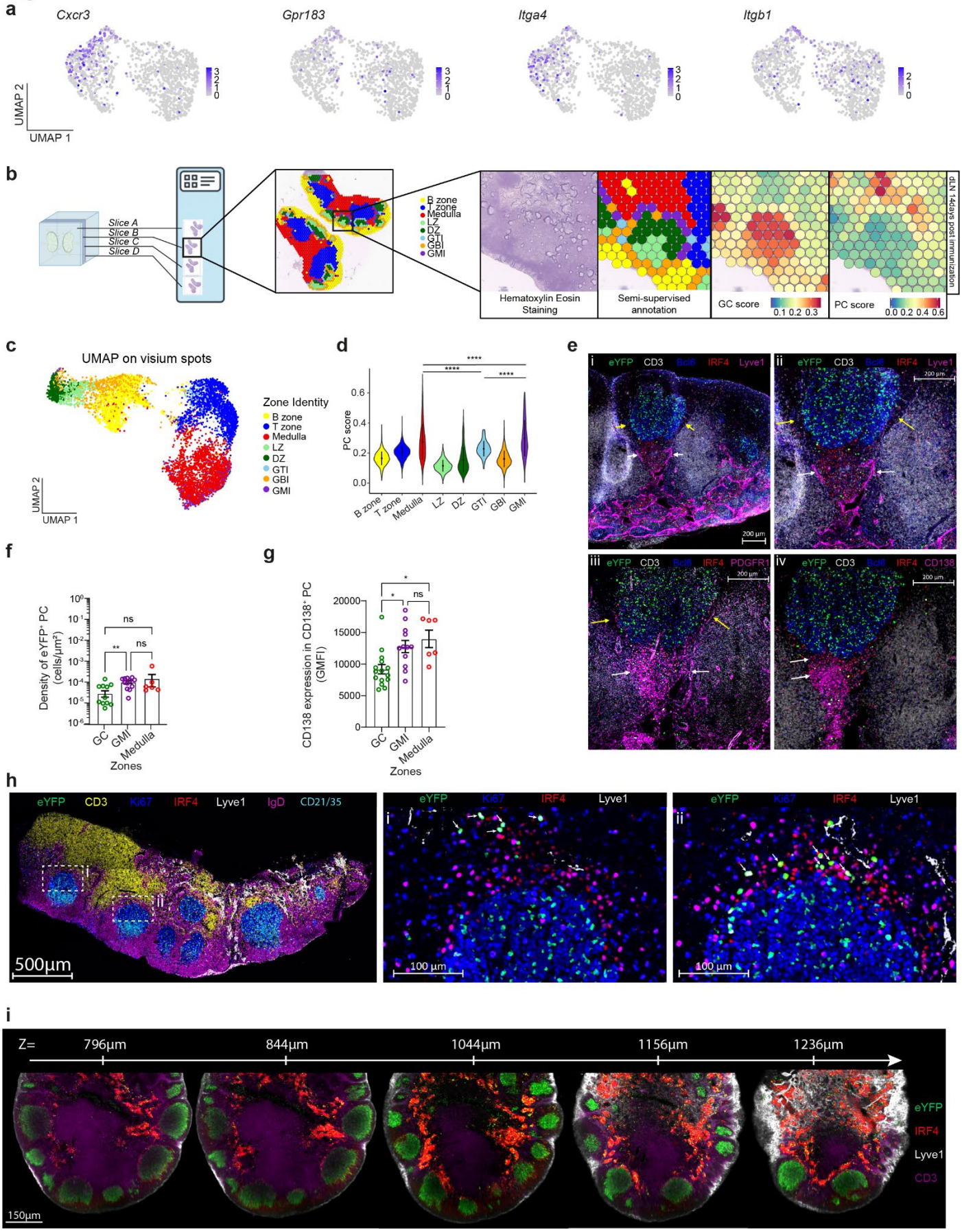
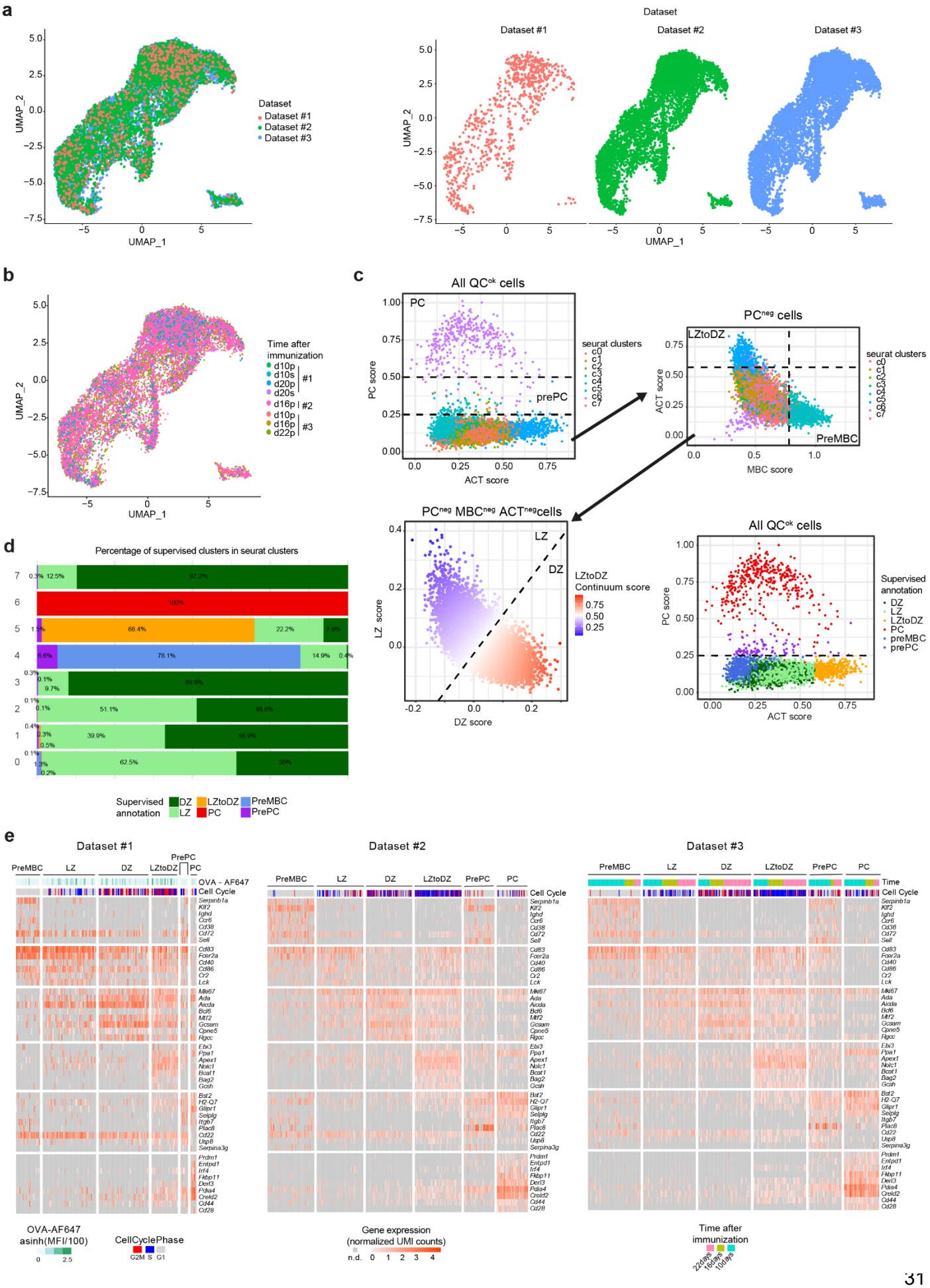


Figure 6: Recent GC-derived PC proliferate at the DZ-medulla interface.

a, Feature plots showing the expression of the indicated genes (*Cxcr3*, *Gpr183*, *Itga4*, *Itgb1*) in cells laid out as in Figure 4b. **b**, Experimental design of the spot-based spatial transcriptomics analysis of dLN, and semi-supervised annotation of spots based on unsupervised spot clustering, gene expression scores, and neighboring spot identity (GTI: GC-T zone interface; GMI: GC-medulla interface; GBI: GC-B zone interface). **c**, Gene expression-based UMAP projection of spatial transcriptomics spots, colored by zonal identity. **d**, Violin plot of PC gene expression score in spots from the indicated zones. **e**, Spectral confocal microscopy images of serial dLN sections (14 days after primary NP-KLH immunization, 3 days after tamoxifen) stained for the indicated markers (scale bar 200 µm). The images show a broad view around on GC (i). We zoomed in this view to show the GTI (ii and iii, yellow arrows) and the GMI (ii and iii, white arrows). The white arrows in iv represent the distinct populations of $CD138^{lo}$ and $CD138^{+}$ cells. **f**, Density of GC-derived PC ($IRF4^{+}$ cells expressing the GC-derived eYFP reporter) quantified from spectral confocal microscopy images after segmentation in the indicated zones. Each dot represents a zone, combining analyses from 2 dLN in 2 distinct experiments. **g**, Expression level of CD138 (geometric mean fluorescence intensity) in mature PC ($IRF4^{+}$ $CD138^{+}$ cells) in the indicated zones. Each dot represents a zone, combining analyses from 2 dLN in 2 distinct experiments. **h**, Spectral confocal microscopy image of the indicated markers in a whole dLN section (left) and zooming in on DZ proximal areas of two GCs (GC i: middle; GC ii: right). Arrows in the zoomed-in panels indicate triple positive $eYFP^{+}$ $IRF4^{+}$ $Ki67^{+}$ GC-derived proliferating PC. **i**, Lightsheet microscopy images of a whole dLN stained with the indicated markers, extracted from different Z-positions as indicated in the top panel. * $p < 0.05$, ** $p < 0.01$, **** $p < 0.0001$ in Wilcoxon matched pairs signed rank test.

829 Supplementary Figures and Legends

Supplementary Figure 1



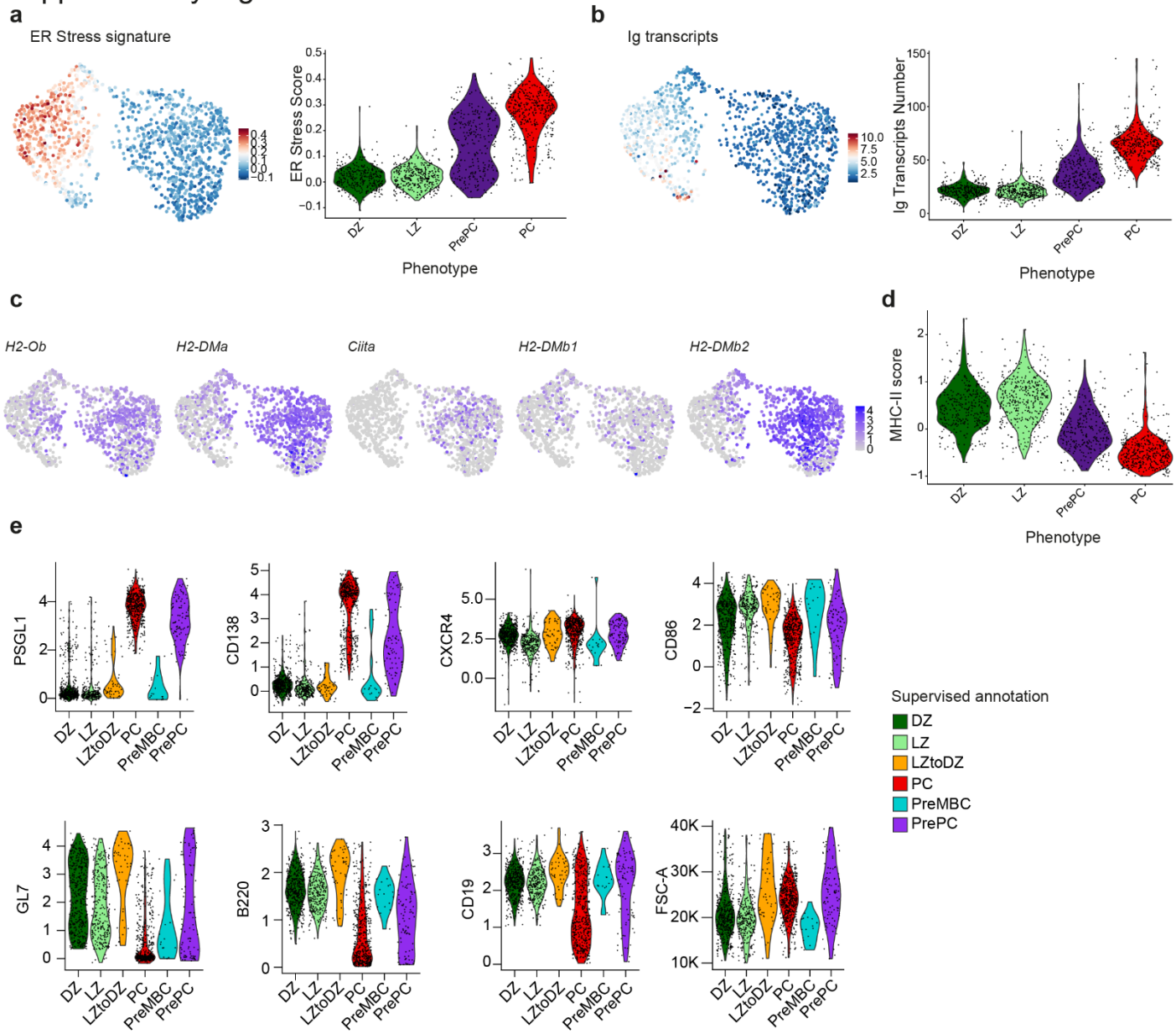
830

831 **Supplementary Figure 1**: supervised annotation of integrated scRNA-seq datasets of GC B cells and recent
832 GC emigrants

833 **a**, UMAP representation of single-cell gene expression profiles of GC and GC-derived cells integrated from
834 the three datasets, colored by dataset identity (left) and split by dataset (right). **b**, UMAP representation of
835 single-cell gene expression profiles of GC and GC-derived cells integrated from the three datasets, colored
836 by time after immunization (d10p: day 10 of primary response, d10s: day 10 of secondary response). **c**,
837 Supervised gating strategy for the integrated scRNA-seq dataset based on gene expression scores for PC,
838 activated (ACT), memory (MBC), DZ and LZ scores. Dashed lines represent the thresholds used to gate the
839 cells in different annotations; the final supervised annotation is represented in the bottom right panel. **d**, Bar
840 graphs of the composition of the unsupervised Louvain clusters in the cell types or states identified by
841 supervised annotation. **e**, Single-cell gene expression heatmaps of the indicated marker genes in the different
842 subsets defined by supervised annotation, also indicating the cell cycle phase (G1, S, G2/M) (all datasets),
843 surface OVA binding (dataset #1) or time after primary immunization (dataset #3).

844

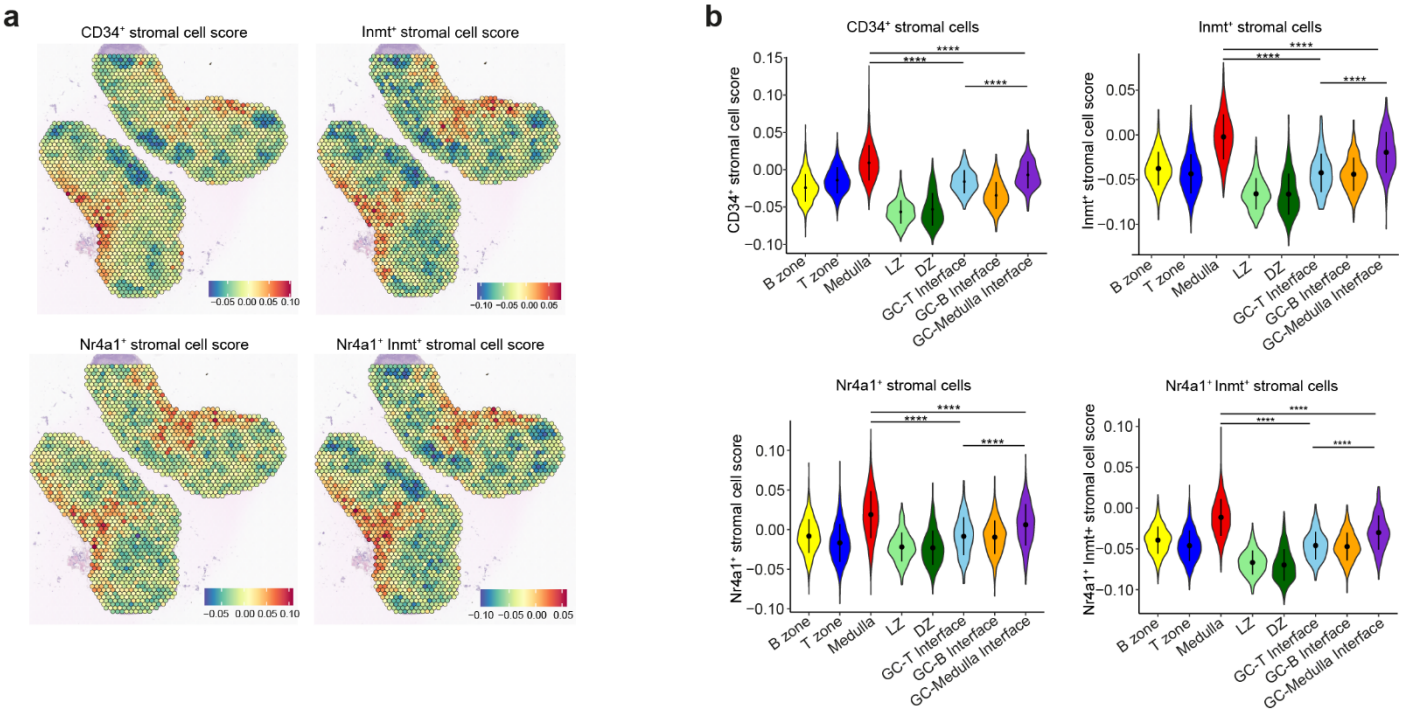
Supplementary Figure 2

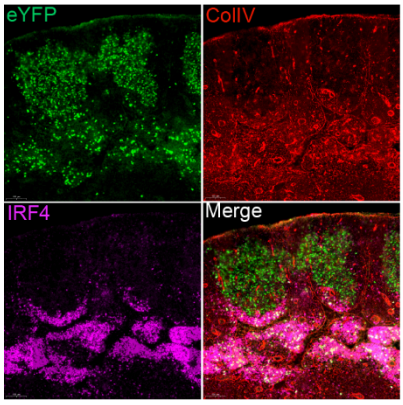


Supplementary Figure 2: Characterization of prePC in the GC-to-PC differentiation continuum.

a, UMAP projection as in Figure 4, colored by the expression levels of the ER stress gene expression signature (left) and violin plot of the expression of the ER stress signature in the cells grouped by phenotype (right). **b**, UMAP projection as in Figure 4, colored by the expression levels of Ig transcripts (left) and violin plot of the expression of Ig transcripts in the cells grouped by phenotype (right). **c**, UMAP projection as in Figure 4, colored by the expression levels of different MHC class II genes (*H2-Ob, H2-DMa, Ciita, H2-DMb1, H2-DMb2*). **d**, Violin plot of the expression of the MHC class II score in the cells grouped by phenotype. The MHC-II gene expression signature was built using the markers depicted in c. **e,** Flow cytometry signal intensity (index sorting) of the indicated surface markers in cells from the indicated subsets (supervised annotation).

Supplementary Figure 3





Supplementary Figure 3: Recent GC-derived PC localize at the DZ-medulla interface.

a, Spatial feature maps showing the expression levels of the stromal cell subset specific gene expression signatures, as indicated. **b**, Violin plots showing the expression of the stromal cell subset specific gene expression signatures indicated in the different zones identified. **c**, Confocal microscopy image of the indicated markers in a dLN section, focusing on two GC and the DZ-proximal area. **** $p < 0.0001$ in Wilcoxon matched pairs signed rank test.