nature immunology

Article

https://doi.org/10.1038/s41590-023-01721-9

Memory B cell subsets have divergent developmental origins that are coupled to distinct imprinted epigenetic states

Received: 14 February 2023

Accepted: 28 November 2023

Published online: 10 January 2024

Check for updates

Derrick Callahan[1], Shuchi Smita[1], Stephen Joachim[1], Kenneth Hoehn[2], Steven Kleinstein[2,3,4], Florian Weisel[1], Maria Chikina[1,5] & Mark Shlomchik[1]

Memory B cells (MBCs) are phenotypically and functionally diverse, but their developmental origins remain undefined. Murine MBCs can be divided into subsets by expression of CD80 and PD-L2. Upon re-immunization, CD80/PD-L2 double-negative (DN) MBCs spawn germinal center B cells (GCBCs), whereas CD80/PD-L2 double-positive (DP) MBCs generate plasmablasts but not GCBCs. Using multiple approaches, including generation of an inducible GCBC-lineage reporter mouse, we demonstrate in a T cell-dependent response that DN cells formed independently of the germinal center (GC), whereas DP cells exhibited either extrafollicular (DP_{EX}) or GCBC (DP_{GC}) origins. Chromatin and transcriptional profiling revealed similarity of DN cells with an early memory precursor. Reciprocally, GCBC-derived DP cells shared distinct genomic features with GCBCs, while DP_{EX} cells had hybrid features. Upon restimulation, DP_{EX} cells were more prone to divide, while DP_{GC} cells differentiated toward $IgG1^+$ plasmablasts. Thus, MBC functional diversity is generated through distinct developmental histories, which imprint characteristic epigenetic patterns onto their progeny, thereby programming them for divergent functional responses.

MBCs are central components of humoral immunity. Arising after vaccination or infection, MBCs provide enhanced and more rapid effector functions upon recall. They also enable the secondary response to adapt to genetically altered pathogen variants if they undergo secondary rounds of proliferation and selection. The origins of MBCs, how they are intrinsically different from their naive precursors, and how they provide both enhanced effector function and adapt to evolution of pathogen epitopes are all areas of active research that require further elucidation.

To this end, in mice, we previously defined three major subsets of MBCs based on expression of CD80 and PD-L2: DN for both markers, single-positive for PD-L2 (PD-L2SP) and DP for both markers[1,2]. A CD80 single-positive (CD80SP) subset also exists, although its rarity has precluded further study. Controlling for isotype expression, we found that IgM^+ DN cells could form GCs, while IgM^+ DP cells made a robust and rapid IgG^+ plasmablast response but lacked GC potential; PD-L2SP cells had an intermediate phenotype[3]. Thus, the dual properties of enhanced effector function and flexibility to adapt upon recall are represented in separate subsets of MBCs defined by CD80 and PD-L2 expression.

The origins of these MBC subsets and how their different functions are encoded remain uncertain, despite active work in this area[4–8]. One common thread is that IgM^+, DN-like MBCs arise early in

[1]Department of Immunology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. [2]Department of Pathology, Yale School of Medicine, New Haven, CT, USA. [3]Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT, USA. [4]Department of Immunobiology, Yale School of Medicine, New Haven, CT, USA. [5]Department of Computational and Systems Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. e-mail: mshlomch@pitt.edu

the response[4,9,10], likely from a CD38$^+$GL7$^+$ proliferating precursor[5,6,8]. This generative site may remain active even after the onset of the GC response, albeit at a relatively low level[4,6]. Similarly, a subset of MBCs, variably termed 'GC-memB' and in some ways resembling DP cells, was inferred to derive from GCBCs based on timing of appearance and evidence of extensive proliferation[4,6]. How proliferative precursors and their MBC progeny are linked is not fully understood.

Here, we developed several approaches to connect proliferative precursors to their respective MBC progeny. Through these studies we linked early memory precursors (EMPs) to DN cells. We further used lineage tracing, exploiting a new GCBC-specific inducible Cre, to demonstrate that GCBCs almost exclusively generate DP cells. This analysis, coupled with single-cell RNA sequencing (scRNA-seq), revealed further heterogeneity among DP cells, with the unexpected finding that DP cells contain at least two subsets that in turn differed in developmental origin and subsequent reactivation potential. We then developed additional experimental systems to better dissect out the roles of proliferation, CD40 signals and GCBC differentiation state in the development of particular MBC types.

By carrying out transcriptomic and epigenetic profiling on MBC subsets derived under different conditions, as well as on their biologically linked precursor populations, we were able to define the intrinsic differences between these MBC subsets. Moreover, among the genes and open chromatin regions (OCRs) that are different among MBC subsets, many are first remodeled in respective precursors and remain remodeled in the resting MBC product. Hence, differential characteristics of memory are imprinted at the precursor state. Taken together, these data substantially advance our understanding of the MBC compartments, including how and when they are created, providing fundamental insight into the nature of cellular memory.

Results

Transcriptional and epigenetic analysis of NBCs and MBCs

To determine the transcriptional and epigenetic relationships between naive B cells (NBCs) and MBCs, MBC subsets and follicular (FO) NBCs were subjected to fluorescence-activated cell sorting (FACS) and subjected to RNA-seq and the assay for transposase-accessible chromatin using sequencing (ATAC-seq; Fig. 1a,b and Supplementary Fig. 1a,b). Principal component analysis (PCA) of all expressed genes among FO NBCs and MBCs segregated the cells mainly by memory status along PC1; DN cells were closest to, and DP cells were the farthest from, FO NBCs (Extended Data Fig. 1a), in agreement with previous microarray data[3].

There were 1,304 unique differentially expressed genes (DEGs) comparing FO NBCs and any MBC subset, with more genes expressed highly in MBCs than FO NBCs (Fig. 1c and Supplementary Table 1). DN cells had the fewest DEGs compared to FO NBCs, and DP cells had the most (Fig. 1d). DP cells had the most uniquely expressed genes (Fig. 1e). Many DEGs were shared among at least two MBC subsets, with a substantial number shared among all three subsets, thereby defining an MBC gene signature (Fig. 1e). Notably, there was a gradual progression in expression levels of shared DEGs from DN to PD-L2SP to DP MBCs (Fig. 1f).

In PCA analysis of all OCRs among FO NBCs and MBCs, MBCs clustered together, yet DN cells showed more similarity to FO NBCs than DPs (Extended Data Fig. 1b). There was a total of 14,926 differentially accessible regions (DARs), comparing all MBC types to FO cells, with more unique OCRs up in MBCs compared to FO NBCs (Fig. 1g and Supplementary Table 2). DPs had the most DARs compared to FO cells, and DNs had the fewest (Fig. 1h). DPs also had the most unique DARs, yet DNs and PD-L2SPs had unique DARs, albeit relatively fewer of them (Fig. 1i). Also, similarly to DEGs, there were shared DARs among all three MBC subsets (Fig. 1i), defining a global epigenetic MBC signature. Among the shared DARs, DPs on average showed the greatest differences from FO cells, and DNs the least (Fig. 1j).

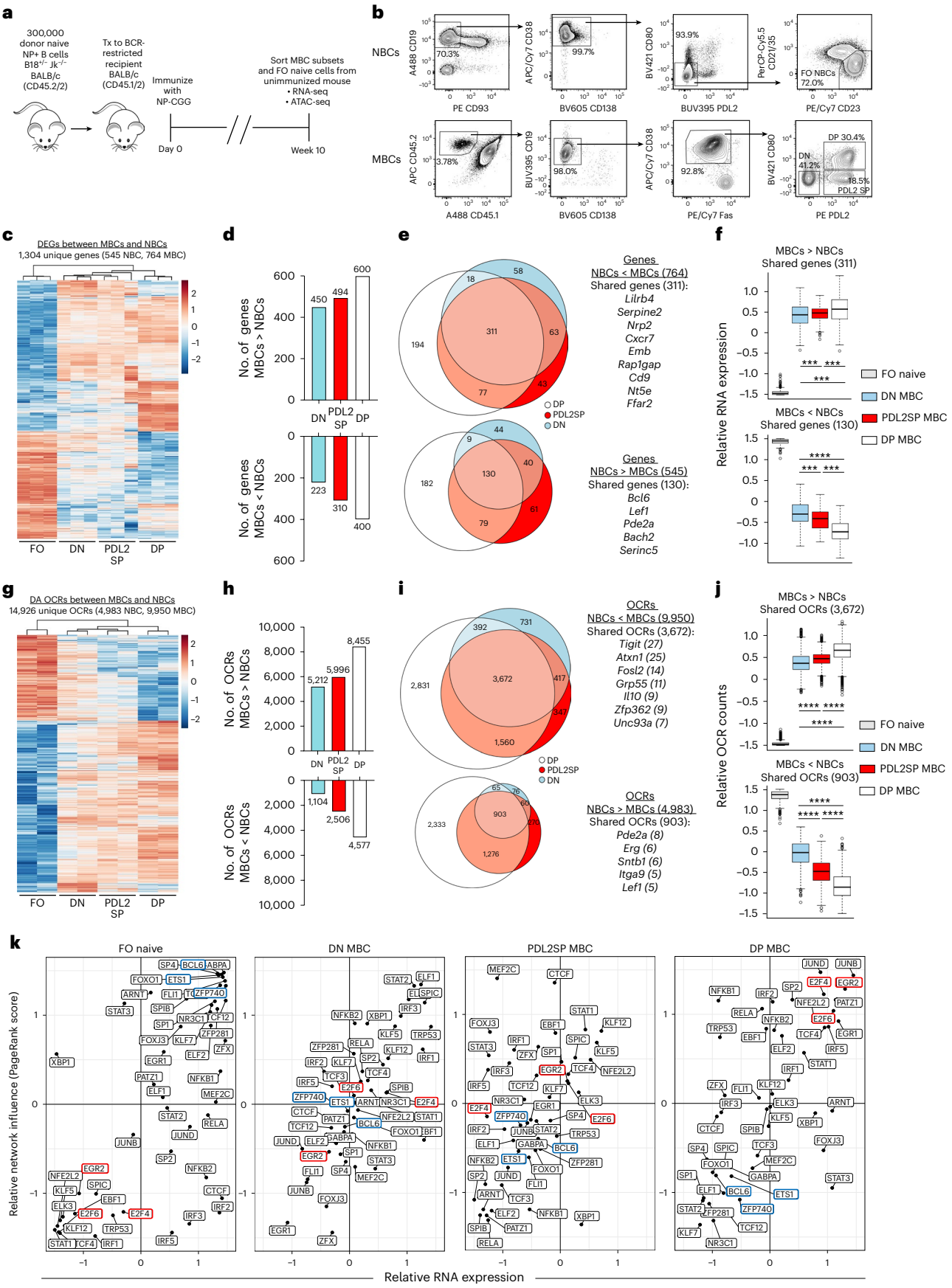
We used the PageRank algorithm[11,12], which integrates RNA-seq and ATAC-seq to rank influence of individual transcription factors (TFs) on the global gene regulatory network (GRN), to infer the top 40 TFs from each cell type. This analysis resulted in only 48 unique TFs (Fig. 1k and Supplementary Table 3), indicative of similar B cell-specific GRNs overall, although relative influence differs. For example, ETS1 showed a progressive decrease in network influence in order of FO, DN, PD-L2SP to DP (Fig. 1k). A similar pattern was observed for BCL6. Given the roles of ETS1 and BCL6 in preventing plasmablast differentiation and the requirement of BCL6 for GCBC formation[13–16], the pattern of relative influence of these TFs matches the in vivo differentiation potential of FO NBCs and MBC subsets. Generally, TFs with a higher rank score in FO NBCs than all MBCs showed a progressive decline from DNs to DPs (Extended Data Fig. 1c). TFs with a higher rank in all MBCs compared to FO NBCs showed an opposite pattern; they descended in ranking from DP, to PD-L2SP, to DN, to FO (Extended Data Fig. 1d). These TFs include nuclear factor-κB, E2f and AP-1 family members, raising the possibility that their relatively high network influence may explain the propensity of DPs to differentiate into plasmablasts upon reactivation[17–19]. Overall, the data highlight a progression of differentiation of MBC subsets from FO cells, with DNs as the most 'naive-like' and DPs as the most terminally differentiated subset.

DN cells and EMPs share similar proliferation profiles

DN cells form before DP cells, with many DP cells forming after the onset of a GC[4]. Therefore, we hypothesized that transcriptional and epigenetic differences between DN cells and DP cells may be imprinted in proliferating upstream precursors—namely, GCBCs for DP cells and early, activated B cells for DN cells. The identity of the precursors for DN cells is unclear, but others have proposed a multipotent, CD38$^+$GL7$^+$, activated, yet undifferentiated B cell at day 2.5 after immunization as the precursor to GC-independent MBCs[5,6,8]. To investigate this, we labeled donor NBCs with violet proliferation dye (VPD) before transfer and immunization and assessed cell phenotype and VPD dilution at days 2.5 and 11.5 to identify both proliferating precursors and newly formed MBCs (Extended Data Fig. 2a). At day 2.5 after immunization, VPD-labeled, CD38$^+$GL7$^+$ donor B cells had divided 3–9 times (Extended Data Fig. 2b). At day 11.5, DNs had a bimodal distribution of VPD peaks: half of the cells divided 3–8 times and the other half 9+ times. The other MBC subsets had diluted all VPD (Extended Data Fig. 2c). The match in division profiles between CD38$^+$GL7$^+$ cells at day 2.5 and DN cells at day 11.5, along with the fact that almost half of the DN cells form in the first

Fig. 1 | Distinct transcriptomic and epigenetic profiles of FO NBCs and MBC subsets. a, Schematic of BALB/c transfer system to generate antigen-specific MBCs. **b**, Gating strategy for FO NBCs and MBCs, first gated on live singlet lymphocytes. For **c**–**f**, $n = 3$ per sample type. **c**, RNA-seq heat map (row z-scores) of DEGs from FO and MBC subsets. Differential expression cutoff for each MBC subset compared to FO NBCs was $\log_2$fold change (FC) ≥ 1, false discovery rate (FDR) ≤ 0.05, and expression level of $\log_2$ expression ≥ 0. **d**, Number of genes in **c** showing either higher or lower expression compared to FO NBCs for each MBC subset. **e**, Venn diagrams showing overlap of DEGs from **c**. **f**, Box plots of scaled expression in each MBC subset of shared DEGs compared to FO NBCs among all three MBC subsets, as depicted in **e**. Box plots display median values and lower and upper quartiles, and the ranges display minimum and maximum values with outliers. **g**–**j**, ATAC-seq analysis between MBC subsets and FO NBCs using the same approach as in **c**–**f**, except the differential accessibility cutoff was $\log_2$FC ≥ 1, FDR ≤ 0.05, $\log_2$ expression ≥ 1 ($n = 2$ for each cell type). **k**, x–y plot showing relative (z-scored) PageRank score of FO NBCs and MBC subsets (y axis, scaled) versus relative (z-scored) RNA expression (x axis, scaled). Top 40 ranked TFs for each cell type were plotted, totaling 48 TFs. Red boxes indicate selected TFs with low relative PageRank score in FO NBCs and high relative PageRank score in DP cells. Blue boxes indicate TFs with high relative PageRank score in FO NBCs and low relative PageRank score in DP cells. P values were calculated using two-tailed paired t-tests (***$P < 0.001$, ****$P < 0.0001$).

https://doi.org/10.1038/s41590-023-01721-9



2 d of the response[4], support the selection of the day 2.5 $CD38^+GL7^+$ cell as a candidate EMP.

RNA-seq and ATAC-seq link DN cells to EMPs and DP cells to GCBCs

We purified day 2.5 EMPs and day 14 GCBCs (further separated by light zone (LZ) and dark zone (DZ) phenotypes) and performed RNA-seq and ATAC-seq (Fig. 2a,b and Supplementary Fig. 2a,b). PCA analysis of all samples clustered MBCs with FO NBCs with respect to transcriptome, but clustered MBCs separately from FO NBCs with respect to OCRs; EMPs and GCBCs clustered separately in both cases (Extended Data Fig. 2d,e). To connect MBC subsets and their putative precursors, we identified 316 DEGs between DN cells and DP cells (Fig. 2c). *k*-means clustering of these DEGs resulted in seven gene groups, distinguished by patterns of expression among MBCs and their precursors (Fig. 2c, Supplementary Fig. 2c and Supplementary Table 4). DN-associated DEG groups (Fig. 2c) separated into expression profiles shared by DN cells and GC (DN/GC), DN only (DN high) or DNs and EMPs (DN/EMP), the largest group. Of DP-associated DEG groups (Fig. 2c), the largest had similar expression in DPs and GCBCs (DP/GC high). There was a highly skewed ($P < 2.2 \times 10^{-16}$) concentration of DEGs into the groups that linked DN and EMP (DN/EMP) groups and DP and GC (DP/GC) groups. Notably, as the $\log_2$FC threshold was increased, DN cells more closely resembled EMPs, while DP cells more closely resembled GCBCs. Hence, the genes with the largest transcriptional differences between DN cells and DP cells have the most similar expression in the proliferating precursor of these MBC subsets, EMPs and GCBCs, respectively (Extended Data Fig. 2f,g).

There were 3,449 unique DARs between DN cells and DP cells (Fig. 2d). *k*-means clustering of the DARs resulted in six overall DAR groups based on accessibility patterns among all cell types (Fig. 2d, Supplementary Fig. 2d and Supplementary Table 5). DN cells clustered closely with EMPs, and DP cells clustered closely with GCBCs (Fig. 2d). As with the transcriptome, the DN/EMP high group was the most prominent of the DN-associated DARs, and the most prominent DP-associated DAR group shared similar expression between DPs and GCBCs (DP/GC high). As with DEGs, there was a highly skewed ($P < 2.2 \times 10^{-16}$) concentration of DARs into the groups that linked DNs with EMPs (DN/EMP) and DPs with GCs (DP/GC). Examples of DN/EMPs DARs were found near the *Foxp1* gene; expression of *Foxp1* also matched accessibility (Fig. 2e). Examples of DP/GC DARs were found near the genes *Basp1* and *Tox* (Fig. 2e). These genes were also some of the most differentially expressed between DN cells and DP cells. Thus, the largest groups of DARs and DEGs link DNs to EMPs and DPs to GCBCs.

We used HOMER[20] to identify TF motifs that were enriched for each DAR group (Fig. 2f, Extended Data Fig. 2h and Supplementary Table 6). Focusing on motifs from expressed TFs (Extended Data Fig. 2i), members of the ETS TF family were enriched among DN/EMP DARs. Oct TF family member motifs, particularly a Pou2f2 (Oct2) motif were enriched in the DP/GC high DAR group; additionally, an EBF motif was also enriched in this group (Extended Data Fig. 2j). Some members of the bHLH TF family were enriched in DN/EMP/GC DARs (Extended Data Fig. 2j). The smaller DAR groups—DP high and DP/EMP high—showed enrichment of nuclear factor-κB, Rel and AP-1 family motifs (Extended Data Fig. 2j). Notably, motifs for binding of these TFs, especially AP-1 family members, were enriched among all DP-associated DAR groups compared to background and DN-associated groups (Extended Data Fig. 2j). The only TF motif specifically enriched in DARs of the DP/EMP/GC group was for Egr2. Together, these data support a role for ETS family members in influencing activity of DN/EMP DARs, while Oct and EBF members may influence activity of DP/GC DARs; AP-1 members may play a role in influencing activity of all DP-associated DARs.

CD40L blockade disproportionately reduces SP and DP MBCs

One prediction of the putative EMP→DN and GCBC→DP linkages is that suppression of GCBC formation should lead to selective reduction of DPs. To test this, we administered anti-CD40L starting at day 3 after immunization (Fig. 3a and Supplementary Fig. 3a). As expected, this blockade resulted in a severe reduction of GCBCs at 4 weeks after immunization (Fig. 3b,c). Supporting the hypothesis, DPs (and PD-L2SPs) were markedly reduced in both frequency and cell number, with a smaller, yet statistically significant effect on DN and CD80SP cell numbers (Fig. 3b,c). However, as CD40L blockade also shuts down the extrafollicular response, we could not exclude the role of this pathway in reducing MBC numbers.

Generation of a GCBC-lineage reporter mouse

Lineage tracing is another approach to test the types of MBC that arise from GCBCs. To perform lineage tracing of GCBCs into MBC progeny with higher fidelity than prior systems[6,7,21], we generated a novel GCBC-specific, tamoxifen-inducible Cre mouse, named GCET-TamCre (Supplementary Fig. 3b). Germinal center-expressed transcript 2 (*Gcet*, *Gcsam*), is a highly expressed, GCBC-specific transcript, yet its deletion resulted in no phenotypic difference compared to wild-type controls[22]. By crossing the *GCET-TamCre* allele with a Rosa26-LSL-YFP mouse, we could administer tamoxifen and use YFP expression to identify cells that developed from GCBCs, allowing us to validate the specificity of the GCET-TamCre and perform lineage tracing of the GCBC progeny. Tamoxifen administration on days 9, 10 and 11 after immunization resulted in ~71% YFP^+ GCBCs at day 13, 3% of non-GCBCs of the B lineage (which include naive cells, activated cells and MBCs), and 3% of plasmablasts (Extended Data Fig. 3a–c). It is likely that the few YFP^+ non-GCBCs and plasmablast cells were derived from the GCBCs labeled during the period of tamoxifen administration. There were no YFP^+ T cells, as expected, and there were virtually no YFP^+ myeloid cells (~0.2% of total splenic cells; Extended Data Fig. 3d).

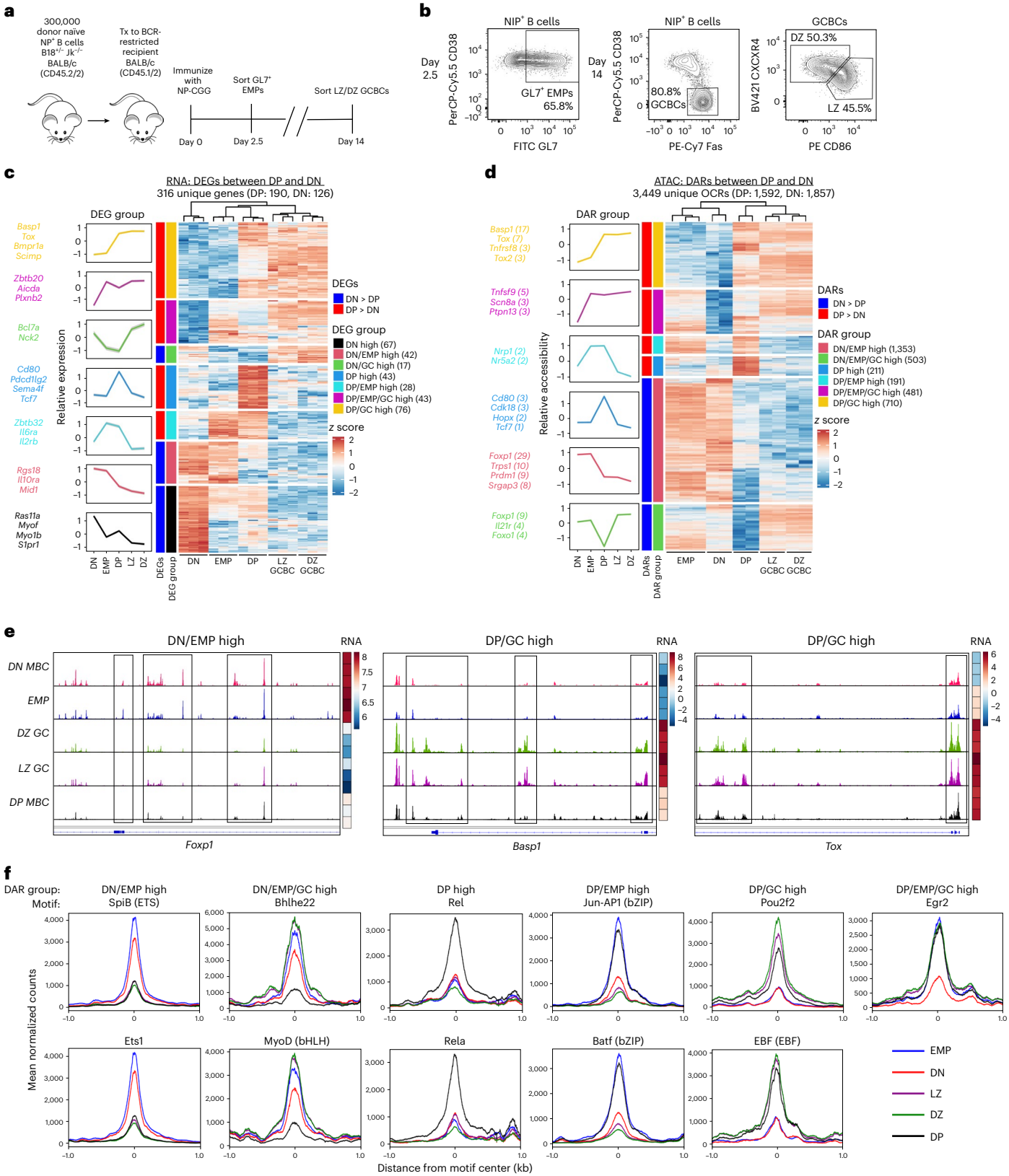
To confirm that the *GCET-TamCre* allele did not display any activity in naive or activated B cells, we gave three doses of tamoxifen 1 d apart in naive mice and checked YFP expression the following day or 12 weeks later; essentially no YFP expression was detected (Extended Data Fig. 3e–g). Additionally, no YFP expression was seen after administering tamoxifen half a day before immunization, followed by two doses 1 d apart, and assessing YFP expression within activated B cells at day 2.5 and week 12 (Extended Data Fig. 3e–g). These data demonstrate that the GCET-TamCre is an effective tool for lineage tracing of cells that are derived from GCBCs, but not from activated B cells.

Fig. 2 | Transcriptomic and epigenetic comparison of proliferating precursors and MBCs. a, Protocol for generating day 2.5 EMPs and day 14 LZ/DZ GCBCs. **b**, Representative flow cytometry gating examples of NIP^+ EMPs and LZ/DZ GCs, first gated on live, $CD45.1^-CD138^-NIP^+$ B cells (Supplementary Fig. 2a,b). **c**, Heat map (row *z*-scores, $n = 3$ per sample type) of expression of DEGs between DNs and DPs across all the indicated cell types, with $\log_2$ FC ≥ 1, FDR ≤ 0.05 and expression ≥ 0 $\log_2$-normalized reads in either cell type. **d**, Heat map (row *z*-scores, $n = 2$ per sample type except $n = 3$ for EMPs) of DA OCRs (DARs) between DNs or DPs, defined as OCRs with log2 FC ≥ 1, FDR ≤ 0.05, and accessibility ≥ 1 $\log_2$-normalized reads in either cell type. For **c** and **d**, *k*-means clustering was used to generate DEG and DAR groups (Supplementary Fig. 2c,d). The resultant heat map was then split by DEG and DAR group annotations and rows were clustered within each group. The average relative expression (**c**) and accessibility (**d**) of each cell type for each group was plotted to the left of the heat maps. Examples of genes (**c**) or OCRs assigned to genes (**d**) corresponding to each group are listed. Parentheses next to listed genes in **d** indicate the number of OCRs associated with the given gene. **e**, Genome track examples of ATAC-seq data from regions characterized in **d** as DN/EMP high (left) or DP/GC high (center and right). To the right of each track is the $\log_2$-normalized expression of the listed gene for each cell type, $n = 3$. **f**, TF motif fold-change enrichment for each DAR group over background (all called peaks) was generated using HOMER. Examples of top TFs for each DAR group were chosen (FC > 1.2, FDR < = 0.01) and the mean normalized counts within 2 kb of the motif center are plotted. Columns represent each DAR group and associated OCRs, and lines represent the listed cell types.

https://doi.org/10.1038/s41590-023-01721-9

GCBCs almost exclusively produce DP MBCs

For GCBC-lineage tracing, we administered tamoxifen on days 9, 10 and 11 after immunization and analyzed cells on day 42 (Fig. 3d and Supplementary Fig. 3c). Residual GCBCs were ~80% YFP$^+$ (Fig. 3e,f), similarly to labeling frequencies immediately after tamoxifen administration (Extended Data Fig. 3b). Among antigen-specific cells at day 42, YFP$^-$ cells were mostly IgM$^+$, while YFP$^+$ cells were split evenly between IgM$^+$ and IgG1$^+$ (Supplementary Fig. 3d). YFP$^+$ MBCs were almost exclusively DP cells, yet only about 50% of DP cells were YFP$^+$ (Fig. 3e,f). Labeling from days 3 to 13 after immunization barely increased the proportion of YFP$^+$ DP cells compared to the day 9–11 labeling window, although earlier labeling slightly increased the proportion of PD-L2SPs



https://doi.org/10.1038/s41590-023-01721-9

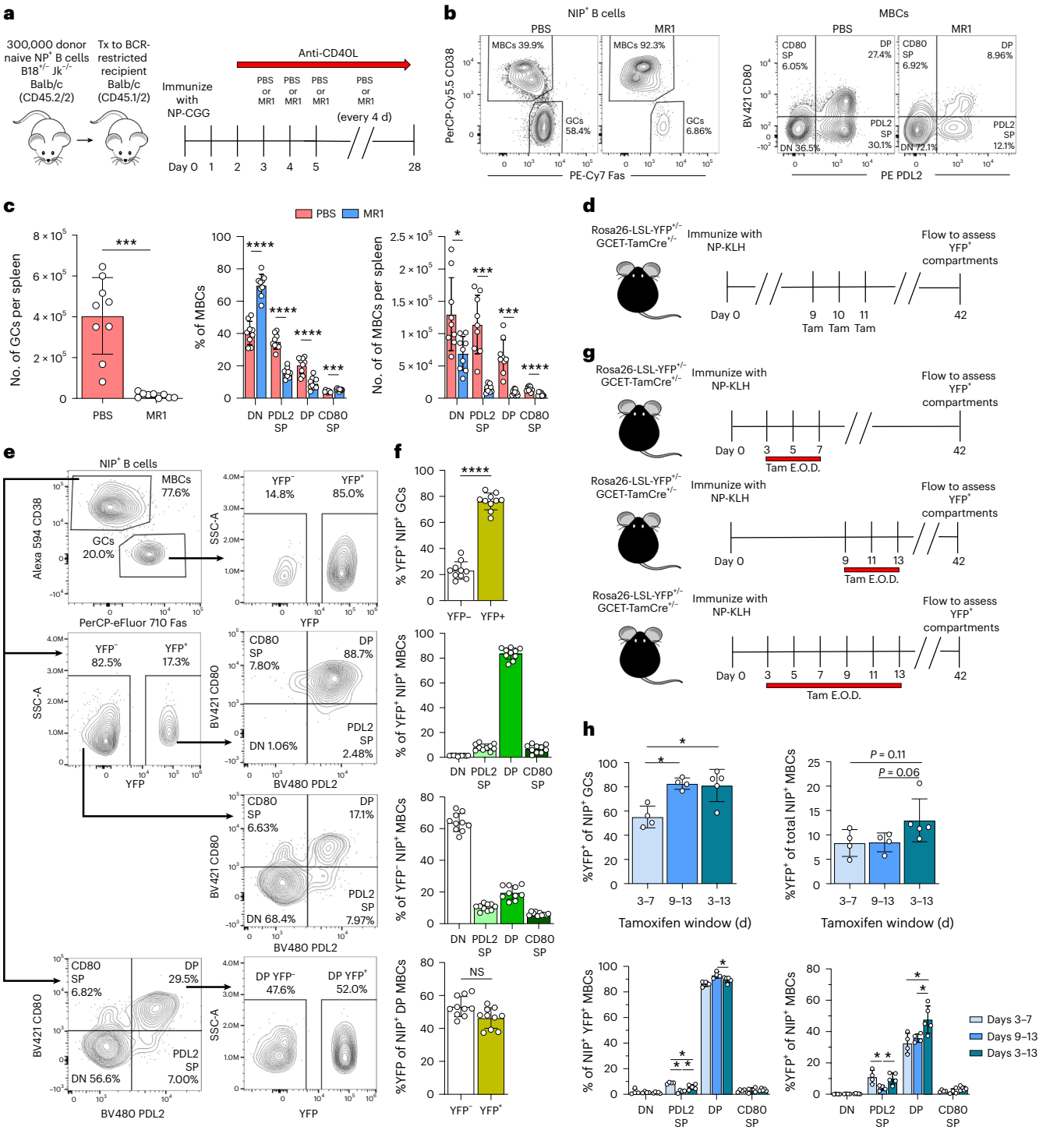


Fig. 3 | GCs disproportionately produce DP MBCs. a, Schematic for MBC generation with anti-CD40L blockade. **b**, Left, representative flow cytometry plots of NIP+B cells (as gated in Supplementary Fig. 3a) produced in **a**, showing CD38^{hi}Fas^{lo} MBCs and CD38^{lo}Fas^{hi} GCBCs; Right, representative flow cytometry plots of MBCs gated from the left plot. **c**, From left to right, number of GCs per spleen, percentage of MBC subsets in total MBCs, and number of MBC subsets per spleen, based on flow cytometry gating in **b**. *P* values were calculated using two-tailed Welch's *t*-test. Bars display the mean ± s.d. Results were compiled from two independent experiments ($n = 9$ for each group). **d**, Mouse model for lineage-tracing GC-derived memory cells. Mice were immunized with 100 μg NP-KLH in alum and given 1 mg tamoxifen by oral gavage at days 9, 10 and 11 and euthanized at day 42. **e**, Representative spectral cytometry plots from splenocytes obtained from mice generated as in **d**, gated on NIP+B cells (Supplementary Fig. 3c); arrows show cascading gating scheme. **f**, From top to bottom, the percentage of YFP+ cells among NIP+ GCBCs, percentage of each MBC subset among YFP+ total MBCs, percentage of each MBC subset among YFP− total MBCs and percentage YFP+ cells among NIP+ DP MBCs. Bars display the mean ± s.d. *P* values were calculated using two-tailed paired *t*-tests. Results are compiled from two independent experiments ($n = 10$ per group). **g**, Schematic of tamoxifen dosing window for three groups of mice given 1 mg of tamoxifen every other day (E.O.D.) from days 3–7, days 9–13 or days 3–13 and euthanized at day 42. **h**, Top left, percentage of YFP+ cells among NIP+ GCBCs; top right, percentage of YFP+ cells among total NIP+ MBCs; bottom left, percentage of each MBC subset among NIP+ YFP+ total MBCs; bottom right, percentage of YFP+ cells within each NIP+MBC subset. Bars display the mean ± s.d. *P* values were calculated using two-tailed Mann–Whitney tests. Results are from one experiment for each time point ($n = 4$ for days 3–7, $n = 4$ for days 9–13, $n = 5$ for days 3–13). *$P <= 0.05$, ***$P < 0.001$, ****$P < 0.0001$ for all plots. NS, not significant.

(Fig. 3g,h), in line with our prior finding that PD-L2SPs form earlier than DP cells[4]. The 80% labeling efficiency of GCBCs compared to 50% labeling of DPs implied a substantial contribution to the DP compartment from a non-GCBC, unlabeled precursor. V-region sequencing of antigen-specific sorted cells from YFP$^+$ and YFP$^-$ DP cells showed that YFP$^+$ DP cells had more mutations than their YFP$^-$ counterparts (Supplementary Fig. 3e). This result suggests that at least some YFP$^-$ DP cells derived from less mutated, potentially non-GCBC precursors; otherwise, the YFP$^-$ DP mutational content would have matched that of YFP$^+$ DPs.

scRNA-seq identifies two subpopulations of DP cells

To further investigate the hypothesis that DP cells comprise distinct cell types, we performed cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq) on purified MBC subsets, including the rare CD80SP population. Cells could be grouped into five clusters (Fig. 4a). A plurality of DP cells was found in cluster 1 (Fig. 4b,c). Because the cell population analyzed was reconstructed from FACS-purified cells, the clusters do not represent natural distributions of MBC subsets; the typical distribution of MBC subsets in this system[4] indicates that cluster 1 would be made up of predominantly DP cells among total splenic MBCs. CD80SPs were overrepresented in cluster 2, to the exclusion of other MBC subsets (Fig. 4b). Overlaying DN and DP scores generated from bulk RNA-seq DEGs on the uniform manifold approximation and projection (UMAP) plot highlighted that cluster 1, which was greatly enriched with DPs, lacked a DN signature, but displayed a strong DP signature; conversely, the DPs that were not in cluster 1 showed a relatively strong DN score (Fig. 4b,d). We hypothesized that DP cells with stronger DN scores may have derived from EMPs (perhaps at a later time point than day 2.5) rather than GCBCs. Supporting the hypothesized EMP→DN pathway, there was a strong association between the DN and EMP scores; conversely, DP and LZ GCBC scores were also correlated (Fig. 4d).

Re-clustering of just the DPs revealed two similarly sized clusters, with DP cluster 1 having high DN/EMP scores and DP cluster 0 having high DP/GC scores, further supporting two major ontogenies of DPs, distinguished by their relationship to GCBCs (Fig. 4e). Cells in cluster 0 (putative GCBC origin) mostly expressed *Ighg1*, while cluster 1 cells (putative EMP origin) mostly expressed *Ighm* (Fig. 4e,f). Among some of the top DEGs in cluster 1 were *Foxp1* (associated with DN/EMP DARs from Fig. 2), as well as *S1pr3* and *Ly6d* (Fig. 4f). *Basp1* and *Tox* were uniquely expressed in cluster 0 (Fig. 4f), matching the increased expression and chromatin accessibility we had previously demonstrated in bulk DP cells and GCBCs.

scRNA-seq data suggested Ly6D as a surface marker to distinguish DP_{GC} versus DP_{EX} in our BALB/c transfer system (Fig. 4f). GCET-TamCre lineage tracing revealed that all GCBC-derived MBCs expressed CD73 and PlexinB2 (Extended Data Fig. 4a), two markers that were reported as more highly expressed on MBCs compared to NBCs[2,5,23] (Figs. 1a and 2c). Applying these markers to our MBC subsets revealed a unique population of PlexinB2^{hi}Ly6D^{lo}CD73$^+$ cells that was present in half of DPs; additionally, these cells overwhelmingly had a DP phenotype (Fig. 4g–i). The same pattern was found after direct immunization of BALB/c mice among polyclonal NIP$^+$MBCs (Extended Data Fig. 4b–d). We thus conclude that DPs comprise at least two major cell types that differ by origin as well as by gene expression profiles and surface markers: GCBC-derived DP (henceforth referred to as DP_{GC}) cells and extrafollicular-derived DP (DP_{EX}) cells.

DP_{GC} and DP_{EX} cells have distinct responses to restimulation

To test for functional differences between DP_{EX} and DP_{GC} cells, we first used NB-21 feeder cells (3T3 cells expressing constitutive CD40L, BAFF and interleukin (IL)-21)[24] to test in vitro function (Fig. 5a). We sort-purified the two (IgG1$^-$) DP populations from our BALB/c transfer system using the newly defined surface markers PlexinB2 and Ly6D (Fig. 5b and Supplementary Fig. 4a,b). In vitro, we found that DP_{GC} cells proliferated less than DP_{EX} cells, yet both DP subsets proliferated less than DN cells (Fig. 5c). IgG1$^-$ DP_{GC} generated more IgG1$^+$ antibody-forming cells (AFCs)/plasmablasts on a per-cell basis than IgG1$^-$ DP_{EX}, although DP_{EX} cells still generated more IgG1$^+$ AFCs than DNs. DP_{EX} cells also generated more IgM$^+$ AFCs on a per-cell basis. We also used the GCET-TamCre lineage-tracing system to seed YFP$^+$ and YFP$^-$ DPs into this in vitro culture system. YFP$^+$ DP cells proliferated less than YFP$^-$ DP cells but produced more IgG1$^+$ AFCs on a per-cell basis, whereas YFP$^-$ DP cells developed more IgM$^+$ AFCs in vitro (Fig. 5d,e and Supplementary Fig. 4c).

We then used the BALB/c transfer system to test in vivo reactivation of DP_{EX} and DP_{GC} cells (Extended Data Fig. 5a). At day 4 after immunization, there were fewer splenic antigen-specific cells emanating from DP_{GC} versus DP_{EX} cells; reciprocally, DP_{GC} cells were more likely to differentiate into IgG1$^+$ AFCs than DP_{EX} cell (Extended Data Fig. 5b). At day 14, neither DP_{EX} nor DP_{GC} cells were capable of forming GCBCs, and both produced similar numbers of total MBCs (Extended Data Fig. 5c). These in vitro and in vivo results reveal baseline differences between DP_{GC} and DP_{EX} cells, illustrating that DP_{GC} cells are less proliferative but more prone to IgG1$^+$ AFC differentiation, yet neither DP subset can produce GCBCs.

T cell-independent immunization produces GC-independent DP cells

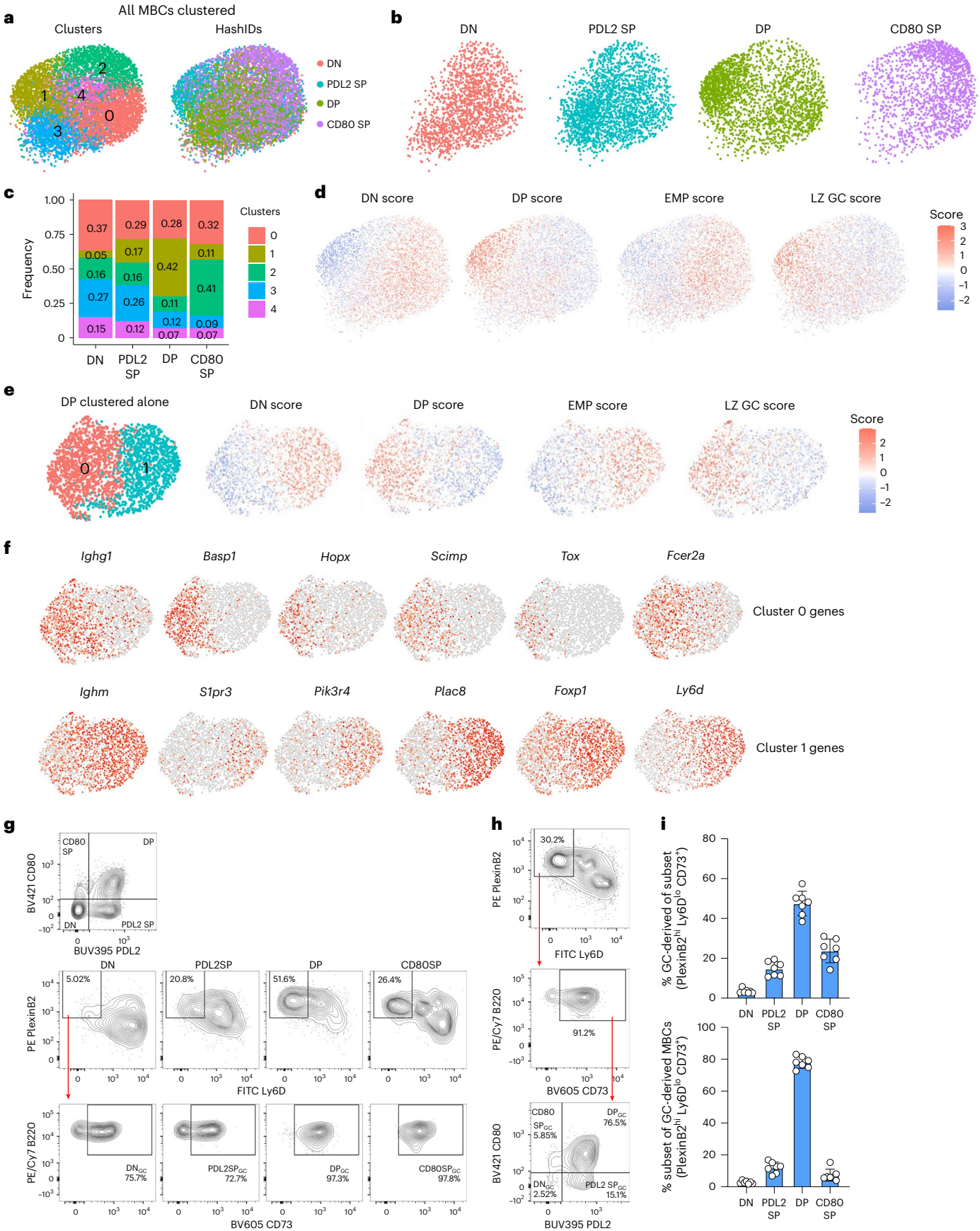
While GCBCs generate almost exclusively DPs, whether, conversely, a GCBC origin is required to establish the DP_{GC} state is unclear. DP_{GC} cells have likely undergone substantially more proliferation and received more T cell help over time than DP_{EX} cells, so these factors alone could be sufficient to establish the transcriptomic and epigenetic state of DP_{GC} cells, independent of GCBC differentiation. To further explore this question, we developed systems to generate TI-DP_{EX} cells that experienced large amounts of proliferation and/or CD40 signals but in a GC-independent fashion.

4-Hydroxy-3-nitrophenylacetyl (NP)-Ficoll is a TI-2 antigen thought to elicit a B cell response independently of T cells. Nonetheless, under some circumstances, NP-Ficoll elicits rapid, transient GC responses[25–27]. Recently, it was suggested that NP-Ficoll elicits TI-MBCs that are GC-dependent, in that they fail to form when B

Fig. 4 | scRNA-seq identifies heterogeneity among DP MBCs, reflective of origin. scRNA-seq with CITE-seq analysis of week 10 MBC subsets from BALB/c transfer system. **a**, Left, UMAP clustering of all MBC subsets identifying five clusters; right, all MBC subsets overlaid on UMAP clusters based on hash-tagged oligonucleotide (HTO) assignments, **b**, Overlay of individual MBC subsets on UMAP clusters ($n = 1$ for each cell type). **c**, Frequency of each MBC subset (determined by HTO assignments) in each cluster. **d**, Left two plots, gene-set enrichment analysis (GSEA) scores were calculated for each cell using DEGs (defined as $\log_2$FC ≥ 1, FDR ≤ 0.05, $\log_2$ expression ≥ 0) between DN and DP MBCs from bulk RNA-seq, and overlaid on the UMAP plot; right two plots, GSEA scores generated from DEGs between EMPs and LZ GCBCs ($\log_2$FC ≥ 4, FDR ≤ 0.05, $\log_2$ expression ≥ 0). Scores are plotted as scaled $-\log_{10}$(*P* value), calculated by two-tailed Wilcoxon's test. **e**, UMAP clustering of DP MBCs only, determined by HTO assignment. GSEA scores were calculated and shown for each cell as in **d**. **f**, Expression of selected genes associated with **e**; cluster 0 or 1 ($\log_2$FC > 0.3, FDR ≤ 0.01). The darker red color indicates higher expression, while gray indicates no expression. **g**–**i**, NIP$^+$ MBCs from the BALB/c transfer system 10 weeks after immunization. **g**, Flow cytometry gating strategy to determine GC-derived MBCs (PlexinB2^{hi}Ly6D^{lo}CD73$^+$) frequencies within CD80/PD-L2 MBC subsets. **h**, Flow cytometry gating strategy to determine frequencies of CD80/PD-L2 MBC subsets within total GC-derived MBCs (PlexinB2^{hi}Ly6D^{lo}CD73$^+$). **i**, Top, percentage of GC-derived cells within each CD80/PD-L2 MBC subset; bottom, percentage of each CD80/PD-L2 subset within total GC-derived MBCs ($n = 7$). Bars display the mean ± s.d.

https://doi.org/10.1038/s41590-023-01721-9

cells lack BCL6 (ref. 27). To generate MBCs that were fully T cell independent, we used our BALB/c transfer system, with the modifications that recipients were depleted of CD4$^+$ T cells (or not) and then immunized with either NP-Ficoll or NP-CGG (Fig. 6a and Supplementary Fig. 5a–d). At day 28, GCBCs were only present in NP-CGG-immunized mice (Fig. 6b–d), as expected. At day 5, NP-Ficoll-immunized mice had



https://doi.org/10.1038/s41590-023-01721-9

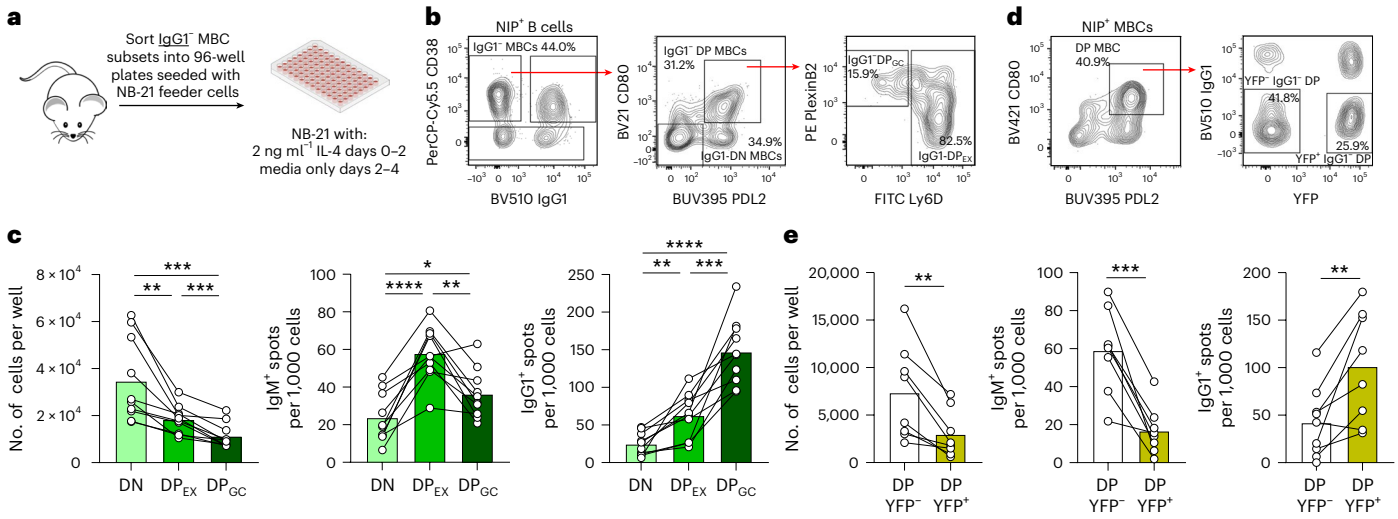


Fig. 5 | GCBC-derived DP cells are less proliferative and more prone to AFC differentiation than extra-GCBC-derived DP cells in vitro. a, Setup for sorting 400 IgG1⁻ MBCs directly into 96-well plates coated with 1,000 NB-21 feeder cells. **b**, Gating strategy for sorting 10 weeks after immunization (NP-CGG) for NIP⁺ IgG1⁻ DP_{EX} and DP_{GC} MBCs from the BALB/c transfer system. **c**, Four days of in vitro culture of indicated subsets from **b** ($n = 10$, 3 independent experiments, one dot connected by lines indicates cells sorted from same mouse); left: cell counts; middle, NP-specific IgM⁺ ELISPOTs per 1,000 counted cells; right, NP-specific IgG1⁺ spots per 1,000 counted cells. **d**, Gating strategy for sorting 8 weeks after immunization (NP-KLH) for NIP⁺ IgG1⁻ YFP⁻ and YFP⁺ DP MBCs from directly immunized GCET-TamCre$^{+/-}$R26-LSL-YFP$^{+/-}$ mice. **e**, Four days of in vitro culture of indicated subsets from **d** ($n = 8$, 2 independent experiments, one dot connected by lines indicates cells sorted from same mouse); left, cell counts; middle, NP-specific IgM⁺ ELISPOTs per 1,000 counted cells; right, NP-specific IgG1⁺ spots per 1,000 counted cells. *P* values were calculated using two-tailed paired *t*-tests (*$P \le 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$). Bars display mean values.

higher frequencies of GCBCs, and much higher total GCBC numbers (Fig. 6b–d) than did NP-CGG-immunized mice. However, CD4-depleted NP-Ficoll-immunized mice had sixfold fewer GCBC-phenotype cells at day 5 than non-depleted NP-Ficoll-immunized mice. Hence, the great majority of the rapid GCBC response to NP-Ficoll in fact depends on CD4⁺ cells.

The quality of TI-MBCs differed substantially from those elicited by NP-CGG. CD80SPs were the most prevalent MBC subset produced, with commensurately very few PD-L2SPs formed (Fig. 6e,f). Despite the large difference in GCBC numbers at day 5 in NP-Ficoll-immunized mice compared to CD4-depleted NP-Ficoll-immunized mice (Extended Data Fig. 6a), there were similar numbers of DPs at day 28 in all groups (Fig. 6f). This finding indicates that DPs produced by NP-Ficoll were largely, if not completely, GCBC and T cell independent; thus, DP differentiation in this case was driven by extensive proliferation alone. Consistent with this conclusion, in the GCET-TamCre system, even though NP-Ficoll-derived GCBCs expressed YFP, the resulting NP-Ficoll-derived DP MBCs did not express YFP (Extended Data Fig. 6b–d). The NP-Ficoll-derived DPs from our transfer system therefore offer proliferation-enhanced TI-DP_{EX} cells as comparators to TD-DP_{EX} and DP_{GC}cells.

CD40 stimulation in the absence of GCBCs produces DP MBCs

To explore the role of CD40 signals and proliferation, independent of GCBC differentiation, we used the same transfer system but upon immunization with NP-CGG in alum also treated T cell-depleted recipient mice with FGK45 (a CD40 agonist) to enhance CD40 signaling (Fig. 7a and Supplementary Fig. 6a–e). Controls received standard NP-CGG immunization with no other treatments. As expected from effective T cell depletion, there were virtually no GCBCs at days 5, 14 and 28 (Fig. 7b–e and Supplementary Fig. 6a–e). However, compared to controls, anti-CD40 administration caused a sharp increase in the numbers of GL7⁺, EMP-like B cells at day 5 (Fig. 7b,f). At both days 5 and 28, anti-CD40 treatment resulted in a marked increase in the frequency and numbers of CD80SP non-GCBCs, with an almost complete loss of PD-L2SPs. DP cells were equivalent in frequency at both timepoints (Fig. 7g), although much higher in number in anti-CD40-treated mice at day 5, compatible with more precursor expansion (Fig. 7h). The numbers of total MBCs, DPs and CD80SPs in anti-CD40-treated mice declined about tenfold from day 5 to day 28, although CD80SP cell numbers remained higher than controls (Fig. 6h and Extended Data Fig. 7a). NP-CGG immunization with T cell depletion but without added anti-CD40 severely reduced numbers of CD80SP cells to levels observed after NP-CGG-only immunization (Extended Data Fig. 7b,c). Given the CD40 signal-driven massive burst of cells at day 5 in the absence of detectable GCBC formation, the resulting DP cells at day 28 can be considered TI-DP_{EX} cells that resulted from a higher degree of proliferation and CD40 signaling. Notably, MBCs derived from these TI systems lacked any cells with a GCBC-derived phenotype, as indicated by expression of PlexinB2/Ly6D/CD73 markers, further validating the GCBC independence of MBCs from these systems (Extended Data Fig. 7d,e).

CD80SP genomic analysis

To investigate the CD80SP population that was abundant in the TI systems, we collected ATAC-seq and RNA-seq data for the MBC subsets in the TI systems and the CD80SPs from the NP-CGG system. Within each system, as anticipated, there were OCRs at the *Cd80* locus that shared higher accessibility between CD80SP and DP MBCs, but lacked accessibility in DN cells; conversely, at the *Pdcd1lg2* locus, CD80SP and DN MBCs lacked accessibility at certain OCRs compared to DPs within each system (Extended Data Fig. 8a). We searched for a shared CD80SP signature across all the systems using GSEA. DARs between CD80SPs and either DNs or DPs were defined in each system as a reference gene set, and then enrichment was assessed between CD80SP and either DN (Extended Data 8b) or DP (Extended Data Fig. 8c) cells. In nearly all cases, DARs that were up or down in CD80SP in one system were enriched in the same direction in the other systems. Hence, there are CD80SP gene expression patterns common to all systems.

We focused on the CD80SPs derived from NP-CGG for deeper analysis, as these would naturally arise in an adjuvant-based vaccination response. While CD80SPs had many OCRs with differential accessibility with DN cells, DP cells had more; on the other hand, CD80SPs had very

https://doi.org/10.1038/s41590-023-01721-9

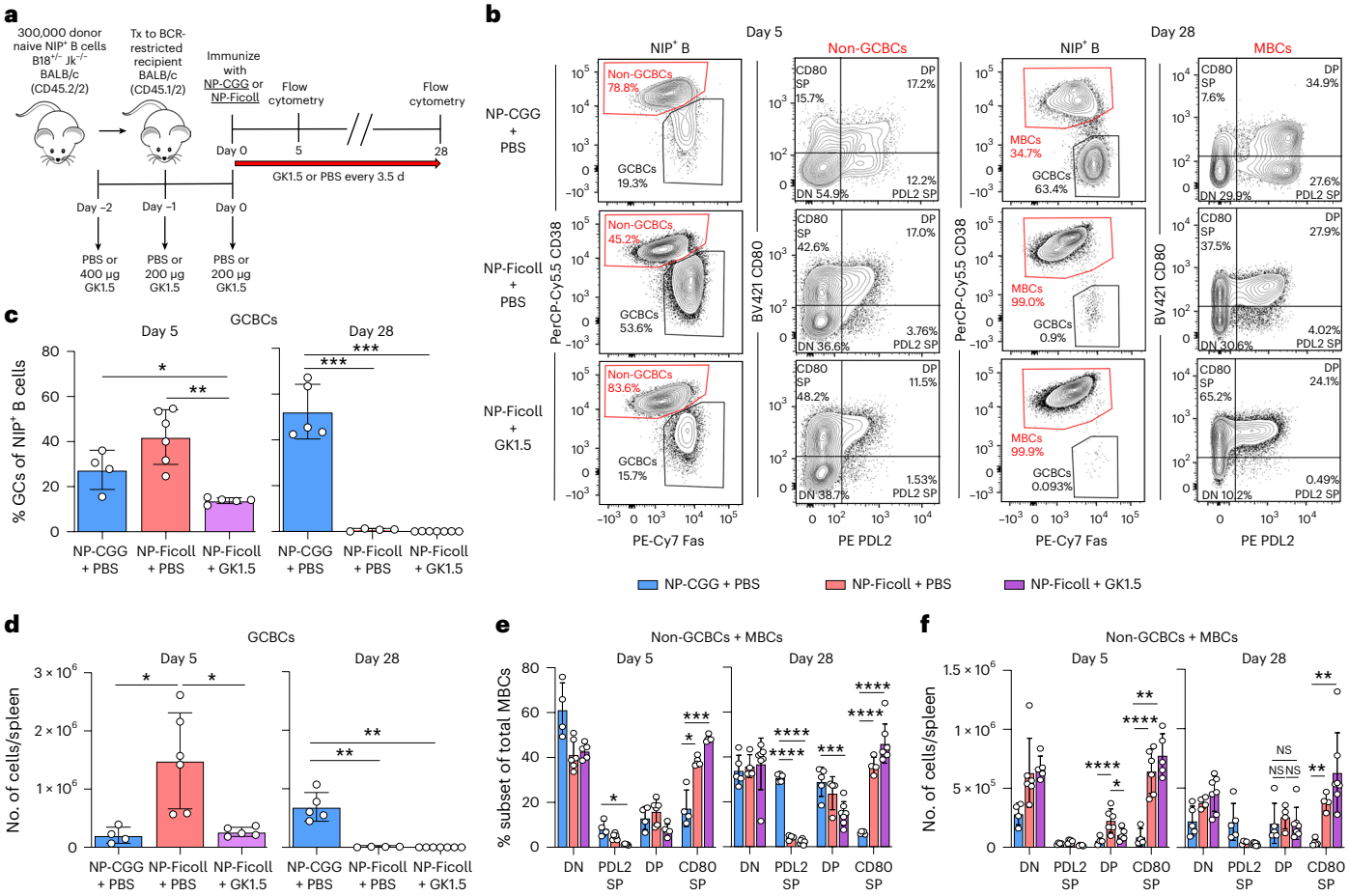


Fig. 6 | NP-Ficoll immunization produces DP MBCs independently of a GC reaction. a, Schematic for the generation of MBCs by treatment with NP-CGG + PBS, NP-Ficoll + PBS or NP-Ficoll + anti-CD4 (GK1.5) in the BALB/c transfer system. Recipient mice were pretreated with PBS or 400/200 µg of GK1.5 at days −2, −1 and 0. Mice were administered donor cells via tail vein injection on day −1, immunized with NP-CGG or NP-Ficoll on day 0 and euthanized on day 5 or day 28. **b**, Representative flow cytometry plots from day 5 (left) and day 28 (right) after immunization. Left column for each day is pre-gated on NIP$^+$B cells (Supplementary Fig. 5a,c). Right column for each day is gated on the non-GC (day 5) or MBC (day 28) populations as drawn in the left column. **c**, Frequency of GCBCs among total NIP$^+$ B cells at days 5 and 28 for each treatment group. **d**, Number of NIP$^+$ GCBCs per spleen at days 5 and 28. **e**, Frequency of each subset among total NIP$^+$ non-GCs (day 5) or MBCs (day 28) for each treatment group. **f**, Number of NIP$^+$ non-GC subsets (day 5) or MBC subsets (day 28) per spleen. Data represent two independent experiments for each day (NP-CGG day 5; $n = 4$, NP-Ficoll + PBS day 5; $n = 6$, NP-Ficoll + GK1.5 day 5; $n = 5$, NP-CGG day 28; $n = 5$, NP-Ficoll + PBS day 28; $n = 4$, NP-Ficoll + GK1.5 day 28; $n = 7$). Bars display the mean ± s.d. P values were calculated using two-tailed Welch's t-test (*$P \leq 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$).

few OCRs that were differentially accessible in DP cells (Extended Data Fig. 8a–d). Most notably, CD80SPs have a large group of OCRs that are not shared with other subsets (Extended Data Fig. 8b,d). Closer examination revealed DARs at the *S1pr5* and *Tbx21* loci that were highly accessible only in NP-CGG-derived CD80SPs (Extended Data Fig. 8e); these CD80SP-specific DARs implied a T-bet/inflammatory signature often associated with age-associated B cells[28]. This notion was supported by the upregulation of select genes associated with T-bet/age-associated B cells[28–33] in NP-CGG-derived CD80SPs compared to other cell types (Extended Data Fig. 8f). Further, T-bet also had the highest relative PageRank score among NP-CGG-derived MBCs (Extended Data Fig. 8g). These findings suggest a relationship to inflammation and the development of CD80SPs in T cell-dependent systems. In fact, CD80SPs are prevalent in systems with chronic inflammation, such as the MRL/lpr mouse model[34]. The relatively low inflammatory environment accompanying NP-CGG immunization with alum may underlie the rarity of CD80SPs in this setting.

Proliferation and CD40 do not solely establish DP_{GC} identity

NP-Ficoll and in vivo anti-CD40-stimulated DP MBCs are proliferation-enhanced TI-MBCs. These TI-DP_{EX} cells, along with TD-DP_{EX} and TD-DP_{GC} cells, allow us to investigate the contributions of GCBC differentiation, proliferation, CD40 signals and other T cell signals in establishing the DP_{GC} epigenetic state. The key comparisons among different types of DP cells used for this analysis and their origins are outlined in Supplementary Fig. 7. If the GC-independent signals associated with DP_{EX} (namely proliferation, CD40, and/or other T cell signals) are able to establish an epigenetic state similar to that of DP_{GC} cells, then we would find no accessibility differences in the DAR groups previously defined in Fig. 2d. Conversely, differences in accessibility would imply a potential requirement for GCBC differentiation in establishing the DP_{GC} state. As hypothesized, for the DARs comprising the two DN-associated DAR groups from Fig. 2d, DP_{GC} cells had lower accessibility than did the comparator DP_{EX} cells in every case (each comparison shown on the x axis; Fig. 8a). The largest DN-associated DAR group, DN/EMP DARs, had the most pronounced difference. Thus, in every system tested, DP_{EX} cells are more 'DN-like' than DP_{GC}. Chromatin regions that were closed specifically in GCBCs and remain closed in DP_{GC} cells may therefore require GCBC differentiation for this closure, as opposed to simply having proliferation or CD40 signals accomplish this. Conversely, the DARs that comprise the four DP-associated DAR groups from Fig. 2d were more enriched

https://doi.org/10.1038/s41590-023-01721-9

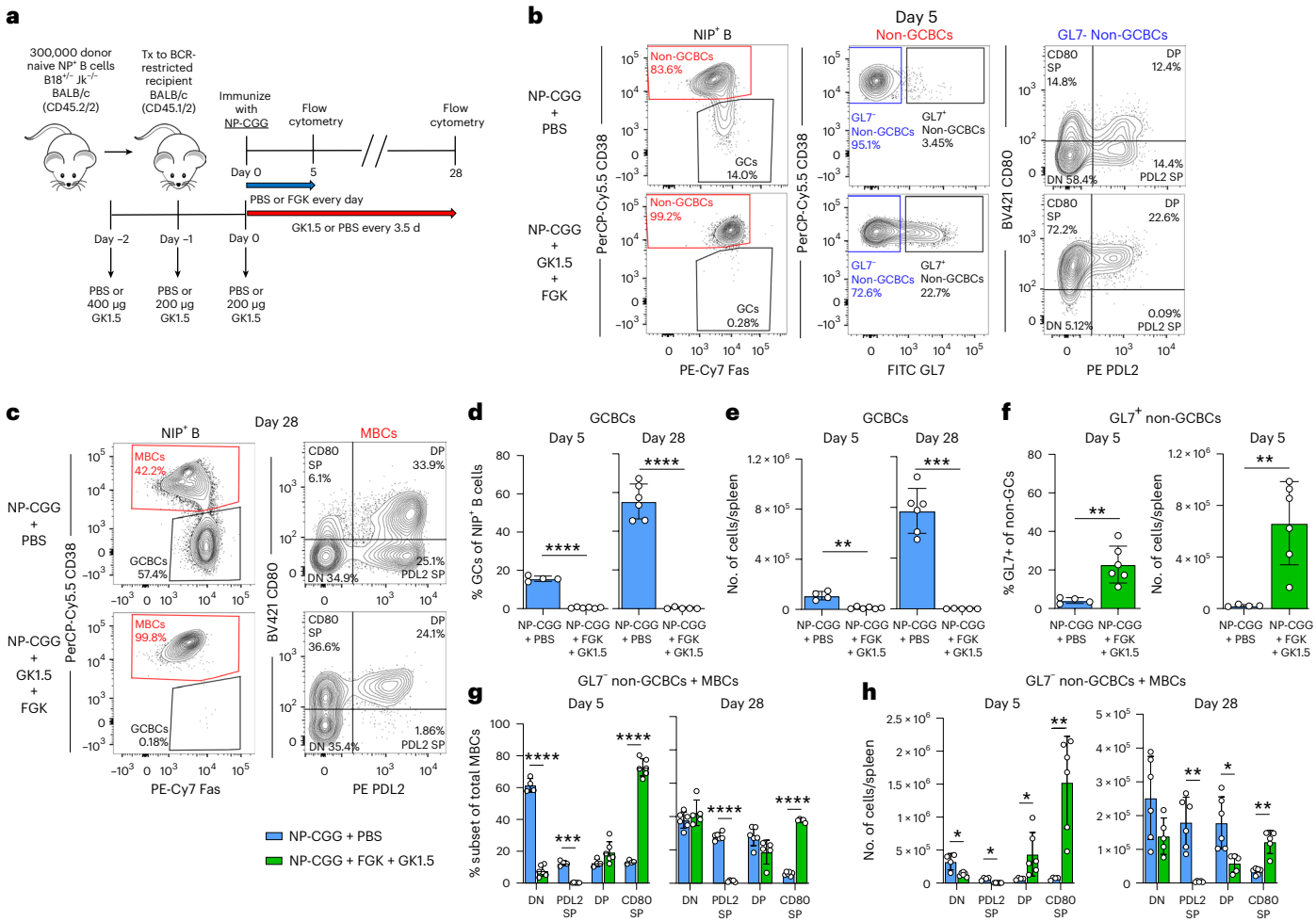


Fig. 7 | CD40 stimulation in the absence of GCBCs produces DP MBCs. **a**, Schematic for the generation of MBCs by NP-CGG + PBS or NP-CGG + GK1.5 + FGK45 (anti-CD40 agonist) in the BALB/c transfer system. Recipient mice were pretreated with PBS or 400/200 µg of GK1.5 (anti-CD4) at days −2, −1 and 0 intraperitoneally (i.p.). Donor cells were administered via tail vein injection on day −1, immunized with NP-CGG on day 0. Some mice were given 25 µg FGK i.p. at days 0.5, 1.5, 2.5, 3.5 and 4.5, and euthanized on day 5 or day 28. **b**, Representative flow cytometry plots from day 5 after immunization. Left column is pre-gated on NIP$^+$ B cells (Supplementary Fig. 6a), the middle column is gated on non-GCs, and right column is gated on GL7$^-$ non-GCs. **c**, Representative flow cytometry plots from day 28 after immunization. Left column is pre-gated on NIP$^+$ B cells (Supplementary Fig. 6d); right column is gated on MBCs. **d**, Frequency of GCBCs among total NIP$^+$ B cells at days 5 and 28. **e**, Number of NIP$^+$ GCBCs at days 5 and 28. **f**, Left, frequency of GL7$^+$ cells among total non-GCs at day 5; right, number of NIP$^+$ GL7$^+$ non-GCs per spleen at day 5. **g**, Frequency of each MBC subset among total NIP$^+$ non-GC (day 5) or MBC (day 28) subsets. **h**, Number of NIP$^+$ non-GC subsets (day 5) or MBC subsets (day 28) per spleen. Data represent two independent experiments for each day (NP-CGG day 5; $n = 4$, NP-CGG + GK1.5 + FGK45 day 5; $n = 6$, NP-CGG + PBS day 28; $n = 6$, NP-CGG + GK1.5 + FGK45 day 28; $n = 5$). Bars display the mean ± s.d. P values were calculated using two-tailed Welch's t-test (*$P \leq 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$).

in DP_{GC} compared to DP_{EX} cells, establishing that, epigenetically, DP_{GC} cells are more 'DP-like' (Fig. 8b). The largest DP-associated DAR group, DP/GC DARs, showed the highest enrichment in DP_{GC} cells compared to the other DAR groups.

While the above analysis clearly indicates a major role for precursor differentiation status—GC or not—in determining the epigenetic fingerprint of specific MBC products, it could still also be the case that proliferation and/or CD40 signals also play a role. To assess this, we re-clustered the DN/EMP and DP/GC DAR groups to identify GC-dependent and GC-independent OCRs more precisely. Overall, these extended MBC groups clustered similarly to the more limited groups presented in Fig. 2 (Fig. 8c,d and Supplementary Table 8). However, DN/EMP DARs could be segregated into two subgroups: a large group of regions that remained open in TI-DP_{EX} cells, reflecting the EMP origin of TI-DP_{EX} cells; and a smaller group that closed in TI-DP_{EX} cells, matching the pattern found in GCBCs and DP_{GC} cells, even though TI-DP_{EX} cells are not GCBC derived (Fig. 8c). TD-DP_{EX} cells were intermediate, but more closely resembled EMPs, TI-DP_{EX} cells and DNs. Reciprocally, within DP/GC DARs, TI-DP_{EX} cells were unable to open most OCRs but were able to open a smaller group, which therefore could be considered proliferation-dependent but not GC-dependent OCRs (Fig. 8d). With respect to DP/GC DARs, TD-DP_{EX} cells clustered with their DP_{GC} counterparts and with GCBCs (Fig. 8d). We conclude that, while extensive proliferation and/or CD40 signals can impact the DP epigenome, proliferation and CD40 signals alone are not sufficient to establish the overall DP_{GC} identity, which in large part is established by precursor GCBC differentiation.

Discussion

Here, we defined the complex relationships between proliferating precursors and their MBC products, demonstrating that MBC lineage-defining and sublineage-defining chromatin remodeling and gene transcription can be traced to features in their respective proliferating precursors. Using systems that dissect the influences of

https://doi.org/10.1038/s41590-023-01721-9

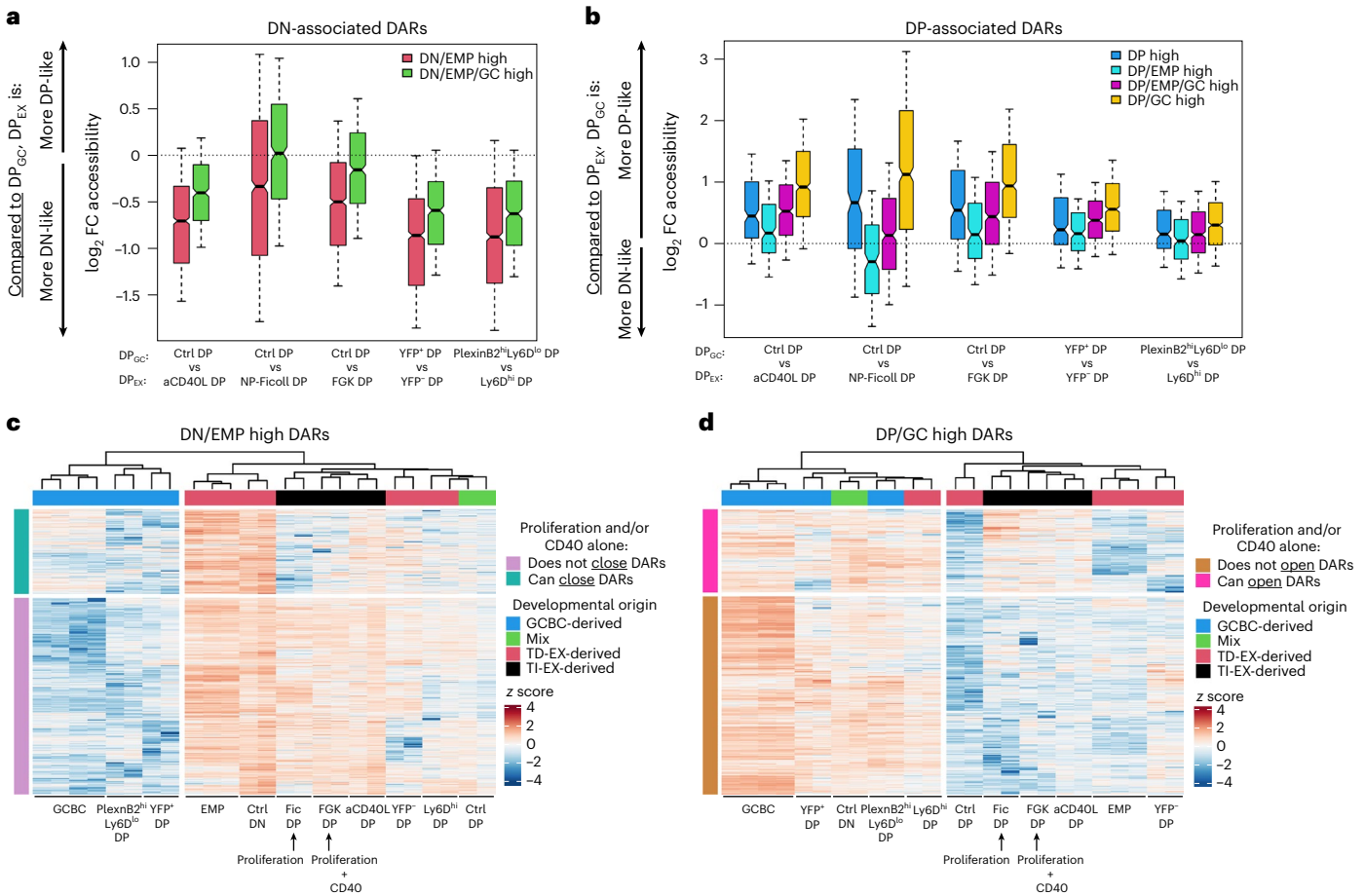


Fig. 8 | Proliferation and CD40 signals are not sufficient to establish the epigenetic identity of DP_{GC}. **a**, Box plots depicting log2FC in accessibility at DN-associated DARs between all DP cells from a T cell-dependent NP-CGG response (containing DP_{GC}) and DP_{EX} generated in different systems, as indicated on the *x* axis. Each bar depicts the distribution across all of the DARs from each DN-associated DAR groups (color coded in the boxed legend), as defined in Fig. 2d. Box plots display median values and lower and upper quartiles, and the ranges display minimum and maximum values with outliers. **b**, Box plots depicting log2FC in accessibility at DP-associated DARs between all DP cells from a T cell-dependent NP-CGG response (containing DP_{GC}) and DP_{EX} generated in different systems, as indicated on the *x* axis. Each bar depicts the distribution across all of the DARs from each DP-associated DAR group (color coded in the boxed legend), as defined in Fig. 2d. Box plots display median values and lower and upper quartiles, and the ranges display minimum and maximum values. For **c** and **d**, heat maps were first generated using *k*-means clustering via the pheatmap package in R. Clusters were then manually assigned as either insufficient or sufficient and re-plotted using the ComplexHeatmap package. EMP: NP-CGG-derived $GL7^+$ EMP (day 2.5), LZ/DZ: NP-CGG-derived LZ and DZ GCBCs (day 14). Ctrl.DP: NP-CGG-derived DP MBC from PBS-treated control mice (week 4), Fic. DP: NP-Ficoll-derived DP MBC from GK1.5-treated mice (week 4), aCD40L.DP: NP-CGG-derived DP MBC from MR1-treated mice (week 4), FGK.DP: NP-CGG-derived DP MBC from GK1.5 and FGK45-treated mice (week 4), DP.YFP.neg: NP-KLH-derived YFP^- DP MBC from GCET-TamCre lineage tracing (week 6), DP.YFP.pos: NP-KLH-derived YFP^+ DP MBC from GCET-TamCre lineage tracing (week 6). $Ly6D^{hi}$ DP: NP-CGG-derived $PlexinB2^{hi/lo}Ly6D^{hi}$ DP MBCs (week 10). $PlexinB2^{hi}Ly6D^{lo}$ DP: NP-CGG-derived $PlexinB2^{hi}Ly6D^{lo}$ DP MBCs (week 10). These cell types are also summarized in Supplementary Fig. 7. **c**,**d**, Heat maps of DN/EMP high DARs (**c**) and DP/GC high DARs (**d**), with each showing whether proliferation and/or CD40 signals are sufficient to close (**c**) or open (**d**) the DARs in the absence of GCBC differentiation. The developmental origin of each cell is indicated as GCBC-derived, mix (GCBC or EX-derived), TI-EX-derived, or TD-EX-derived. Fic and FGK DPs are indicated as cell types that undergo extensive proliferation or extensive proliferation and CD40 signals, absent of GCBC differentiation, respectively.

GCBC state, proliferation and T cell help, we discovered not only that the differentiation state of precursors is an important determinant of chromatin in resultant MBCs, but also that extent of proliferation and possibly T cell-derived signals in precursors can influence the nature of chromatin remodeling and hence the types of MBCs that result. In the process, we discovered that DP MBCs actually comprise two subsets—DP_{EX} and DP_{GC}—which are distinguished by their origins. We further showed that they have distinct functional capacity both in vitro and in vivo. We further defined conditions that favor the development of CD80SP MBCs, which previously were barely explored, and defined their epigenetic and gene expression signatures. Together these findings provide new insights into how and where B cell memory forms and into the qualitatively distinct types of MBCs that can result from various types of stimulation.

Several groups have shown that DN-type MBCs can arise early in the response[4,5,35]. Most recently, Glaros et al. showed that EMPs can generate what they termed 'early MBCs' (eMBCs), and that, when examined very shortly after formation, EMP and eMBC transcriptomes bore similarities; they did not follow out lineage-marked progeny to a later time point, leaving open the question of whether such eMBCs evolve into a more mature transcriptome, nor did they study the epigenome[8]. We advanced upon this work in several ways. First, we provided a more definite, causal and specific linkage of EMPs to their MBC progeny; identifying some of these as DN MBCs (as opposed to 'early' MBCs) using a combination of VPD dilution analysis, targeted interruption of the response at very early time points, lineage tracing and transcriptomic and epigenetic analysis. Unexpectedly, these studies revealed that DP cells comprise two subpopulations—one of GCBC origin (DP_{GC}),

and the other formed outside a GC (DP_{EX}). This conclusion was also supported by scRNA-seq, which identified two distinct subsets of DP cells that aligned with either EMP or GCBC transcriptional programs. Viant et al. focused on actB that formed at later times[6]; consistent with this, but extending their findings, we found that extrafollicular responses accompanied by substantial proliferation and/or CD40 signals can generate DP cells, but also CD80SP cells, neither of which, to our knowledge, had previously been recognized.

Price et al. recently characterized the epigenome and transcriptome of both IgM and IgG murine MBCs, suggesting that they were both primed to undergo plasma cell differentiation[36]. We added a critical new dimension to the analysis by comparing epigenomes and transcriptomes across defined precursors and MBC subsets, which showed distinctive behavior. The linkage of EMPs to DN cells and DP_{EX} cells, and GCBCs to DP_{GC} cells in precursor–product relationships, along with the large swath of DEGs and OCRs that defined differences between MBC subsets and that are shared between these precursor/product pairs, argues mechanistically for a model in which some loci that are remodeled in a developing proliferating precursor retain the remodeled pattern (and in many cases the gene expression) in the resting MBC product. 'Memory' is thus a memory in part of the state of the cell before it became a resting MBC. A perfect match of OCRs would not be expected, as in fact the precursors and products are fundamentally different cell types.

Our data also suggest that the extent of proliferation can influence MBC quality. For example, OCR modules that are shared among both TI-DP_{EX} and DP_{GC} cells but absent in DN cells reflect a distinct vector of influence, which is likely linked to proliferation. Nonetheless, GCBC differentiation state per se is a key determinant of the quality of DPs. In agreement with this notion, the extent of proliferation and CD40 signals received by TI-DP_{EX} was mostly unable to remodel chromatin to match that of DP_{GC}cells. However, while TD-DP_{EX} cells did not completely close OCRs that were closed in DP_{GC} cells, they were able to open many OCRs that were opened in DP_{GC} cells, albeit to a lesser extent. It is therefore possible that T cell signals received by TD-DP_{EX}, but not TI-DP_{EX} (for example, IL-21), are sufficient to open these OCRs independently of a GC; we also cannot rule out other epigenetic differences at these regions. These different vectors—proliferation, CD40 signals, GCBC differentiation and other T cell signals—may even act together, such that a DP_{EX} or DP_{GC} cell that forms later in the response may differ from one that forms earlier, as a result of undergoing more divisions and perhaps getting different amounts and qualities of T cell signals.

Studies here represent salient demonstrations of precursor genetic states imprinting stable changes in resting progeny, of which there are relatively few other documented examples. In human B cells, DNA demethylation occurred in GCBCs and was retained in both MBCs and plasma cells[37]. During mouse development, enhancers that are open in developing precursors are demethylated and can remain in a demethylated yet transcriptionally silent state[38], bearing some similarity to what we have seen, although in MBCs some of the genes first turned on in precursors continue to be transcribed in MBC progeny. Observations comparable to our findings have also been made in memory $CD8^+$ T cells and memory NK cells[39,40]. Interestingly, some skin cell OCRs that are induced in inflammation stay relatively open after resolution of inflammation; these post-inflammation sites healed wounds more quickly[41]. Thus, skin stores epigenetic memory of inflammation that yields functional differences upon challenge, much like we have observed for MBCs.

An important contribution of our work is identifying the different functional capacity of DP_{EX} versus DP_{GC} cells, thus extending the functional heterogeneity of the MBC compartment even beyond the previously identified DN versus DP differences. A major remaining task is to determine how these functional differences are encoded. To this end, we identified TF motifs in DARs that may differentially control DN versus DP development, regardless of origin. The TFs that have more network influence and expression in FO NBCs, typically show a progressive decrease from FO NBCs to DNs, to PD-L2SPs, and then to DPs. This hierarchical influence is commensurate with known functions of these cells upon immunization. The influence of these TFs, which include BCL6 and ETS1, may promote GC differentiation in DN MBCs. Conversely, some TFs with higher expression and network influence in DP cells show progressively lower influence in PD-L2SPs, to DN cells, to FO NBCs. These TFs may enable DP cells to become quickly activated and differentiate into plasmablasts. Some of these TFs, such as AP-1 factors, have been implicated in the development of CD8 and NK cell memory[42]. Additionally, E2F4 and E2F6, which have been associated with cell cycle arrest[18,43], have high influence in DP cells, which may serve to block proliferation in favor of differentiation, as is seen in DP MBCs.

While here we have defined epigenetic and transcriptomic differences among known and newly revealed MBC subsets, a future challenge will be to determine how these epigenetic changes are imprinted in different biological settings. Finally, it will be important to determine how epigenetic states influence the differences in DP_{EX} and DP_{GC} cells upon restimulation, thus more directly explaining how specific MBC subset functions are encoded.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41590-023-01721-9.

References

1. Anderson, S. M., Tomayko, M. M., Ahuja, A., Haberman, A. M. & Shlomchik, M. J. New markers for murine memory B cells that define mutated and unmutated subsets. *J. Exp. Med.* **204**, 2103–2114 (2007).
2. Tomayko, M. M., Steinel, N. C., Anderson, S. M. & Shlomchik, M. J. Cutting edge: hierarchy of maturity of murine memory B cell subsets. *J. Immunol.* **185**, 7146–7150 (2010).
3. Zuccarino-Catania, G. V. et al. CD80 and PD-L2 define functionally distinct memory B cell subsets that are independent of antibody isotype. *Nat. Immunol.* **15**, 631–637 (2014).
4. Weisel, F. J., Zuccarino-Catania, G. V., Chikina, M. & Shlomchik, M. J. A temporal switch in the germinal center determines differential output of memory B and plasma cells. *Immunity* **44**, 116–130 (2016).
5. Taylor, J. J., Pape, K. A. & Jenkins, M. K. A germinal center-independent pathway generates unswitched memory B cells early in the primary response. *J. Exp. Med.* **209**, 597–606 (2012).
6. Viant, C. et al. Germinal center-dependent and -independent memory B cells produced throughout the immune response. *J. Exp. Med.* **218**, e20202489 (2021).
7. Shinnakasu, R. et al. Regulated selection of germinal-center cells into the memory B cell compartment. *Nat. Immunol.* **17**, 861–869 (2016).
8. Glaros, V. et al. Limited access to antigen drives generation of early B cell memory while restraining the plasmablast response. *Immunity* **54**, 2005–2023 (2021).
9. Inamine, A. et al. Two waves of memory B-cell generation in the primary immune response. *Int Immunol.* **17**, 581–589 (2005).
10. Takemori, T., Kaji, T., Takahashi, Y., Shimoda, M. & Rajewsky, K. Generation of memory B cells inside and outside germinal centers. *Eur. J. Immunol.* **44**, 1258–1264 (2014).
11. Yu, B. et al. Epigenetic landscapes reveal transcription factors that regulate $CD8^+$ T cell differentiation. *Nat. Immunol.* **18**, 573–582 (2017).

12. Zhang, K., Wang, M., Zhao, Y. & Wang, W. Taiji: system-level identification of key transcription factors reveals transcriptional waves in mouse embryonic development. *Sci. Adv.* **5**, eaav3262 (2019).
13. Bories, J. C. et al. Increased T-cell apoptosis and terminal B-cell differentiation induced by inactivation of the Ets-1 proto-oncogene. *Nature* **377**, 635–638 (1995).
14. John, S. A., Clements, J. L., Russell, L. M. & Garrett-Sinha, L. A. Ets-1 regulates plasma cell differentiation by interfering with the activity of the transcription factor Blimp-1. *J. Biol. Chem.* **283**, 951–962 (2008).
15. Angelin-Duclos, C., Cattoretti, G., Lin, K. I. & Calame, K. Commitment of B lymphocytes to a plasma cell fate is associated with Blimp-1 expression in vivo. *J. Immunol.* **165**, 5462–5471 (2000).
16. Fukuda, T. et al. Disruption of the *Bcl6* gene results in an impaired germinal center formation. *J. Exp. Med.* **186**, 439–448 (1997).
17. Vasanwala, F. H., Kusam, S., Toney, L. M. & Dent, A. L. Repression of AP-1 function: a mechanism for the regulation of Blimp-1 expression and B lymphocyte differentiation by the B cell lymphoma-6 protooncogene. *J. Immunol.* **169**, 1922–1929 (2002).
18. Tooze, R. M. A replicative self-renewal model for long-lived plasma cells: questioning irreversible cell cycle exit. *Front. Immunol.* **4**, 460 (2013).
19. Scharer, C. D., Barwick, B. G., Guo, M., Bally, A. P. R. & Boss, J. M. Plasma cell differentiation is controlled by multiple cell division-coupled epigenetic programs. *Nat. Commun.* **9**, 1698 (2018).
20. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
21. Dogan, I. et al. Multiple layers of B cell memory with different effector functions. *Nat. Immunol.* **10**, 1292–1299 (2009).
22. Schenten, D., Egert, A., Pasparakis, M. & Rajewsky, K. M17, a gene specific for germinal center (GC) B cells and a prognostic marker for GC B-cell lymphomas, is dispensable for the GC reaction in mice. *Blood* **107**, 4849–4856 (2006).
23. Weisel, N. M. et al. Surface phenotypes of naive and memory B cells in mouse and human tissues. *Nat. Immunol.* **23**, 135–145 (2022).
24. Kuraoka, M. et al. Complex antigens drive permissive clonal selection in germinal centers. *Immunity* **44**, 542–552 (2016).
25. de Vinuesa, C. G. et al. Germinal centers without T cells. *J. Exp. Med.* **191**, 485–494 (2000).
26. Obukhanych, T. V. & Nussenzweig, M. C. T-independent type II immune responses generate memory B cells. *J. Exp. Med.* **203**, 305–310 (2006).
27. Liu, X., Zhao, Y. & Qi, H. T-independent antigen induces humoral memory through germinal centers. *J. Exp. Med.* **219**, e20210527 (2022).
28. Cancro, M. P. Age-associated B cells. *Annu. Rev. Immunol.* **38**, 315–340 (2020).
29. Johnson, J. L. et al. The transcription factor T-bet resolves memory B cell subsets with distinct tissue distributions and antibody specificities in mice and humans. *Immunity* **52**, 842–855 (2020).
30. Kim, C. C., Baccarella, A. M., Bayat, A., Pepper, M. & Fontana, M. F. FCRL5$^+$ memory B cells exhibit robust recall responses. *Cell Rep.* **27**, 1446–1460 (2019).
31. Li, J. et al. CCR5$^+$T-bet$^+$FoxP3$^+$ effector CD4 T cells drive atherosclerosis. *Circ. Res.* **118**, 1540–1552 (2016).
32. Evrard, M. et al. Sphingosine 1-phosphate receptor 5 (S1PR5) regulates the peripheral retention of tissue-resident lymphocytes. *J. Exp. Med.* **219**, e20210116 (2022).
33. Reis, B. S., Hoytema van Konijnenburg, D. P., Grivennikov, S. I. & Mucida, D. Transcription factor T-bet regulates intraepithelial lymphocyte functional maturation. *Immunity* **41**, 244–256 (2014).
34. Nickerson, K. M. et al. Age-associated B cells are heterogeneous and dynamic drivers of autoimmunity in mice. *J. Exp. Med.* **220**, e20221346 (2023).
35. Kaji, T. et al. Distinct cellular pathways select germline-encoded and somatically mutated antibodies into immunological memory. *J. Exp. Med.* **209**, 2079–2097 (2012).
36. Price, M. J., Scharer, C. D., Kania, A. K., Randall, T. D. & Boss, J. M. Conserved epigenetic programming and enhanced heme metabolism drive memory B cell reactivation. *J. Immunol.* **206**, 1493–1504 (2021).
37. Lai, A. Y. et al. DNA methylation profiling in human B cells reveals immune regulatory elements and epigenetic plasticity at Alu elements during B-cell activation. *Genome Res.* **23**, 2030–2041 (2013).
38. Jadhav, U. et al. Extensive recovery of embryonic enhancer and gene memory stored in hypomethylated enhancer DNA. *Mol. Cell* **74**, 542–554 (2019).
39. Schauder, D. M. et al. E2A-regulated epigenetic landscape promotes memory CD8 T cell differentiation. *Proc. Natl Acad. Sci. USA* **118**, e2013452118 (2021).
40. Ruckert, T., Lareau, C. A., Mashreghi, M. F., Ludwig, L. S. & Romagnani, C. Clonal expansion and epigenetic inheritance of long-lasting NK cell memory. *Nat. Immunol.* **23**, 1551–1563 (2022).
41. Naik, S. et al. Inflammatory memory sensitizes skin epithelial stem cells to tissue damage. *Nature* **550**, 475–480 (2017).
42. Lau, C. M. et al. Epigenetic control of innate and adaptive immune memory. *Nat. Immunol.* **19**, 963–972 (2018).
43. Hsu, J. & Sage, J. Novel functions for the transcription factor E2F4 in development and disease. *Cell Cycle* **15**, 3183–3190 (2016).

Methods

Mice

All experiments were conducted under protocols approved by the University of Pittsburgh Institutional Animal Care and Use Committee and adhered to ethical regulations. All mouse strains were kept under specific-pathogen-free conditions in a room at 20–26 °C with 30–70% relative humidity. Mice were kept on a 14-h light/10-h dark schedule. Mice used for donor and recipient transfers and immunizations were between 6 and 12 weeks at the time of primary immunization. Females were used in all cases except for YFP^+ and YFP^- DP MBCs from the GCET-TamCre lineage tracing, in which the ATAC-seq data was 80% female and 20% male. WT BALB/c mice and Rosa26-LSL-YFP mice were purchased from JAX (strains 000651 and 006148, respectively).

GCET mouse generation

The region surrounding the ATG of *Gcet* was analyzed to identify Cas9 target sequences that overlap the start codon. The region chr16:45,610,511–45,610,563 was targeted (Mouse Dec. 2011 (GRCm38/mm10) Assembly). The CRISPR target sites were identified with predicted specificity (off-target effects) and predicted efficiency (on-target cleavage) by various algorithms through CRISPOR (http://crispor.tefor.net/). All target sequences were unique in the genome and had a limited number of potential off-targets. Two target sequences were selected based on a combination of considerations: proximity to the target codons, acceptable profile of potential off-targets and Moreno-Mateos score.

iCre-ERT2-T2A was knocked-in at the start codon of *Gcet* on a C57BL/6J background using the pBlu2KS-iCre-ERT2-T2A vector. This vector was amplified, linearized and assembled with amplified 5′ and 3′ homology arms of the *Gcet* locus using the pBluescript II KS by Gibson assembly, generating the pBlu2KS-Gcet-iCre-ERT2-T2A targeting vector. This targeting vector was successfully microinjected in the pronucleus of fertilized C57BL/6J zygotes, which led to the production of potential founders. Fertilized embryos (C57BL/6J, The Jackson Laboratory) produced by natural mating, were microinjected in the pronuclei with a mixture of 0.33 µM EnGen Cas9 protein (New England Biolabs, M0646T), single-guide RNAs (sgRNAs; sgRNA1: 5′-TAATACGACTCACTATAGGTCTGCAAACAGTTCCCCATG-GTTTTAGAGCTAGAAATAGCA-3′, sgRNA2: 5′-TAATACGACTCACTATAGGTGCTGAGGGGC-CTGCCTCATGTTTTAGAGCTAGAAATAGCA; 21.23 ng $\mu l^{-1} \approx 0.66$ µM) and the pBlu2KS-Gcet-iCre-ERT2-T2A plasmid (10 ng μl^{-1}). The injected zygotes were cultured overnight, and the next day the embryos that developed to the two-cell stage were transferred to the oviducts of pseudopregnant CD1 female recipients.

BALB/c transfer system to generate memory cells

Memory mice were generated as previously described[3,4]. Briefly, splenocytes from naive B1-8i$^{+/-}$ Jk$^{-/-}$ CD45.2/2 mice, in which ~50% of NBCs are NP^+, were isolated using anti-CD4 IgM, anti-CD8 IgM followed by complement depletion. An equivalent of 3×10^5 NP^+ cells were transferred via tail vein injection into naive AM14-Tg mice crossed with Vk8R$^{+/-}$ CD45.1/2 recipients and immunized the following day. The resulting MBCs are >90% NP^+.

Tissue processing

Splenocyte single-cell suspensions were generated by crushing spleens between glass microscope slides in 5 ml R10 (Gibco RPMI 1640, Thermo Fisher Scientific, 21870076; 10% Fetalplex (Gemini Bio Products, 100602/500); 1% penicillin–streptomycin (Thermo Fisher Scientific,10378016); 1% L-glutamine (Thermo Fisher Scientific, A2916801); 1% HEPES pH 7.0–7.5 (Thermo Fisher Scientific, 15630080); 0.00035% 2-mercaptoethanol). Cells were filtered through 80-µm nylon mesh and washed at 300*g* at 4 °C for 5 min. Pellets were resuspended in 3 ml of 25 °C ACK lysis buffer (Thermo Fisher Scientific, A1049201) per spleen and incubated at 25 °C for 2 min, followed by addition of 12 ml of R10 and washing at 300*g* at 4 °C for 5 min. For flow cytometry, cells were resuspended in staining buffer (PBS, 3% Fetalplex, 1 mM EDTA, 0.02% sodium azide). For bead enrichment followed by cell sorting, cells were resuspended in staining buffer without sodium azide.

NP-CGG, NP-KLH, NP-Ficoll prep and immunization

In a 50 ml conical tube, 100 µl of 10 mg m^{-1} NP_{33}CGG (BioSearch Technologies, N-1010-100, Rockland Immunochemicals, D602-0100, conjugated in-house) was added to 1.9 ml PBS, followed by 4 ml of 10% alum. 1 M potassium hydroxide was added dropwise while taking pH until a pH of 6.5 was reached. The tube was spun at 500*g* at 25 °C for 10 min, followed by two subsequent washes in PBS at 500*g* at 25 °C for 10 min. The pellet was resuspended to a final volume of 4 ml in PBS. Then, 50 µl (200 µg) was injected i.p. into each mouse. For NP_{32}-KLH (Bioresearch Technologies, N-5060-25), 600 µl of 1 mg ml^{-1} NP_{32}-KLH was added directly to 1.2 ml of 10% alum, and the mixture was brought to a pH of 6.5 and washed as with NP-CGG. The pellet was resuspended to 1.2 ml PBS and 200 µl (100 µg) was injected i.p. into each mouse. Then, 1 ml of 2 mg ml NP_{55}Ficoll (BioSearch Technologies, F-1420-100) was added to 7 ml of PBS, and 200 µl (50 µg) was injected i.p. into each mouse.

VPD labeling protocol

Cells were prepared as described in 'GCET mouse generation', counted and then washed in staining buffer. Supernatant was aspirated, cells were resuspended to 5×10^6 cells per ml in pre-warmed PBS with 2 µM BD Horizon VPD 450 (BD Biosciences), and then incubated at 37 °C for 10 min. Cells were then washed with 25 °C R10 and suspended in staining buffer.

Tamoxifen administration

A total of 10 mg tamoxifen powder (MP Biomedicals, 156738) was solubilized in 1 ml corn oil by mixing at 37 °C overnight with shaking. The 10 mg ml^{-1} solution was then frozen at −20 °C. Freeze/thaw cycles were avoided. Mice were given 100 µl (1 mg) of tamoxifen via oral gavage at the times indicated.

In vivo antibody administration

Anti-CD40L (clone MR1, BioXcell BE0017-1 and homemade) was injected i.p. at 350 µg per 200 µl. Anti-CD4 (clone GK1.5, homemade), was diluted to either 2 mg ml^{-1} or 1 mg ml^{-1} and 200 µl (400 µg) and 200 µl (200 µg) were injected i.p., respectively. Anti-CD40 agonist (clone FGK45, BioXcell, BE0016-2) was injected i.p. at 25 µg per 200 µl.

RNA collection and sequencing

A total of 25,000–100,000 cells were collected via FACS, washed twice in PBS + 1% Fetalplex, resuspended in 350 µl RLT Plus Buffer (Qiagen) with 1% 2-Me, and frozen at −80 °C. RNA purification was performed using the RNeasy Micro Plus Kit (Qiagen, 70434). Libraries were generated using the SMART-Seq v4 Ultra Low Input RNA kit (Takara Bio) followed by Nextera FLEX using UDI set B indexes (Illumina). Samples were sequenced using NovaSeq 6000 or NextSeq 2000 flowcells to obtain 20 M 2 × 75-bp reads. Some samples were sequenced as 2 × 101-bp reads.

ATAC-seq

The ATAC-seq protocol was adapted from ref. 44. Around 10,000–50,000 viable cells were pelleted in lo-bind 1.5 ml microcentrifuge tubes (Eppendorf, 022431021), and washed once with PBS supplemented with 1% Fetalplex and protease inhibitor (Sigma, P8340, stock at 100×). Cells were resuspended in 50 µl of cold lysis buffer containing 0.1% IGEPAL CA-360 (Sigma, I8896, 0.01% Digitonin (Promega, G9441) and 0.1% Tween-20 (Sigma, 11332465001) in 10 mM Tris-HCl (pH 7.5), 3 mM $MgCl_2$, 2 mM NaCl and mixed by gently pipetting. Samples were incubated on ice for 3 min to isolate nuclei, then washed with 950 µl of 0.1% Tween-20, 10 mM Tris-HCl (pH 7.5), 3 mM $MgCl_2$ and 2 mM NaCl and pelleted at 500*g* for 10 min at 4 °C. Supernatant was aspirated

using P1000 and P200 pipettes, followed by addition of 50 µl transposition mix containing 25 µl TD buffer (Illumina, 20034197), 2.5 µl TDE1 Tagment DNA Enzyme (Illumina, 20034197), 16.5 µl PBS, 0.01% digitonin and 1% Tween-20, brought to 50 µl with molecular-grade H_20. Samples were incubated at 37 °C for 30 min. Transposed DNA was then purified using the Qiagen MinElute PCR purification kit (Qiagen, 28004). To generate libraries, adaptors (IDT for Illumina Nextera DNA Unique Dual Indexes, Illumina, 20027215) were added during amplification using 2× Phusion High Fidelity Mastermix with HF Buffer (Thermo Fisher, F531L). Libraries were cleaned up with Zymo DNA Clean & Concentrator-5 columns (Zymo Research, D4004) and assessed with a Qubit Flex Fluorometer (Thermo Fisher, Q33327) and a TapeStation 4200 (Agilent, G2991BA). Samples were sequenced to obtain 20 M 2 × 75-bp paired-end reads using an Illumina NextSeq 550 sequencer.

Bulk RNA-seq processing and analysis

Samples were aligned to the mm10 genome using the STAR aligner[45]. Gene-level counts were determined using featureCounts[46] (v.2.0.1), and raw counts were quantile normalized to each other for differential expression using the voom method[47] in the limma R package[48]. For normalization of the datasets, the quantile method was used. Differential accessibility analysis was performed using 'limma' R package (Supplementary Tables 1 and 4).

Bulk ATAC-seq processing and analysis

All sequencing reads were trimmed with Trim Galore (v.0.6.5) and aligned with their reference genome (mm10) using Bowtie2 (v.2.3.3)[49] aligner (using default parameters) to produce raw bam files. Further, mitochondrial reads, unmapped reads and low mappability (<30) reads were filtered out using the SAMtools view[50]. PCR duplicates were removed using picard (v.2.18.12) to obtain final filtered bam files. Final bam files were converted to bigwig files using deeptools (v.3.3.0) bamCoverage tool (parameter --bs = 1 --normalizeUsing RPKM --p = max) and bigwig files were used to view genome tracks in IGV (Broad Institute v.2.3.97). Peaks and peaks summit were called using MACS2 (with parameter --q 0.05 -B --nomodel --shift -75 --extsize 150) after merging filtered reads from replicate 1 and replicate 2 from all samples into two files. The irreproducible discovery rate (IDR v.2.0.3) analysis was used to evaluate the reproducibility between the two biological replicates from all samples. Peaks passing the soft IDR threshold (IDR ≤ 0.1) and rank based on *P* value were retained for further analysis. Blacklisted peaks were removed. Raw counts overlapping each peak's summit extended 100 bp up and down (in such a way so that peaks are in the boundary of a called narrow peak region) were quantified using bwtool summary (v.1.0.0) and normalized using the voom method[47] in the limma (v.3.52.2) R package[48]. Differential accessibility analysis was performed using 'limma' R package (Supplementary Tables 2, 5 and 8). TF motif enrichment analysis for selected peaks was performed using HOMER[20] (v.4.9.1; Supplementary Table 6). All gene-set enrichments were performed using the rankSumTestWithCorrelation function in limma, which explicitly corrects for correlation among genes (or OCRs) in the gene set being interrogated.

CITE-seq protocol

MBC subsets were sorted as described above and in Fig. 1a,b. After sorting, cells were washed once in PBS with 0.04% BSA. TotalSeq anti-mouse Hashtag Antibodies recognizing CD45 and MHCI (BioLegend) were added at a 1:50 dilution to each of the sorted samples (DN MBCs: TotalSeq-A0310, PD-L2SP MBCs: TotalSeq-A0311, DP MBCs: TotalSeq-A0312, CD80SP MBCs: TotalSeq-A0312). Cells were stained on ice in PBS with 0.04% BSA for 20 min and then washed once in PBS with 0.04% BSA. Cells were counted and 5,600 of each cell type were loaded into the 10x Genomics Chromium system per the manufacturer's instructions, with a ~57% input recovery. Gene expression and antibody hashtag/feature barcode libraries were generated, their quality was assessed through the Agilent TapeStation High Sensitivity D5000 Screentape, and their amounts were quantified with the KAPA Library Quantification Kit for Illumina Platforms. We used the 3' V2 library; the feature barcode library was generated according to the New York Genome Center protocol[51]. For hash-tagging, we followed the 'feature barcode' instructions from the manufacturer. Libraries were pooled and sequenced by Medgenome on a NovaSeq 6000 (Illumina Biosciences) to get a total of 660 M 2×150-bp reads.

Single-cell RNA-seq data processing and analysis

FASTQ files were generated and aligned to the mouse reference genome mm10 with Cell Ranger 5.0.0 to produce the gene–cell count matrix and cell–antibody count matrix. The 10x raw data from each sample were demultiplexed and FASTQ files were generated using the 'mkfastq' Cell Ranger pipeline (v5.0.0, 10x Genomics). Cell Ranger 'count' was used to align reads to the mm10 reference genome, and mRNA transcript, and HTO unique molecular identifier (UMI) quantification tables were generated. The raw barcode matrix files generated from the Cell Ranger pipeline were further utilized for downstream analysis using the Seurat package (v.3.0.2)[52] in R (v.3.4.3). Cells expressing less than 200 genes, or with greater than 10% of UMIs that mapped to mitochondrial DNA, were filtered out. The HTO tables were added to the dataset and normalized by a centered-log ratio method using the 'NormalizeData' function. The normalized HTO count was used to determine if each gel bead-in-emulsion contained a single cell using the Seurat 'MULTIseqDemux' function and manual inspection of cells, where a cell was considered 'singlet' if expression of a single HTO accounts more than 70% of the total HTO expression in that cell; otherwise, the cell was considered a 'doublet' and removed. UMI, mitochondrial content, hemoglobulin gene and ribosomal gene content scores were 'regressed out' using Seurat's 'ScaleData' function. Variable genes were detected using the 'mean.var.plot' method in 'FindVariableFeatures' function with default cutoff. These variable genes were used for dimensionality reduction based on PCA using the 'RunPCA' function. 'ElbowPlot' was used to assess the first 50 principal components, and the principal components that account for the largest variability in the data were selected for further UMAP dimensional reduction and clustering analysis. To identify distinct groups of cells, unsupervised clustering was performed using the 'FindClusters' function, which calculates the *k*-nearest neighbors according to variable gene expression in all cells, thereby constructing a shared nearest neighbor graph using the Louvain algorithm. To avoid overclustering, we tested different resolution ('res') parameters, ranging from 0.1 to 2 in increments of 0.1, and the clustering progression was assessed and visualized using 'Clustree' (v.0.4.3). Optimal resolution was determined based on continued separation before 'overclustering' as observed by the increasing crossover between clusters. Based on these observations, we chose the resolutions '0.4' and '0.1' for MBC and DP clusters, respectively. Cell clusters were visualized using UMAP dimensional reduction plots. A small cluster expressing high levels of the *Igkc* transcript was removed; cells of interest were $J\kappa^{-/-}$ and therefore the *Igkc*-expressing cells were presumed to be recipient contamination. The 'FindAllMarkers' function with default settings was utilized to find DEGs in each cluster, in comparison to all other clusters, using the Wilcoxon rank-sum test with genes detected in a minimum of 10% of cells, a minimum of 0.25 average log-fold change and a minimum of 0.01 Bonferroni-adjusted *P* value (Supplementary Table 7). Scores for UMAP plots were generated using the AddModuleScore function in Seurat.

PageRank analysis

Taiji PageRank pipeline (v.1.2.0)[12] with default parameters were used to integrate bulk RNA-seq and ATAC-seq data for key TF identification (Supplementary Table 3). Cis-BP database was used for mouse motifs.

Flow and spectral cytometry

Fc receptors on cells were blocked to prevent nonspecific antibody binding at a concentration of 1×10^8 cells per milliliter with anti-CD16/CD32 (clone 2.4G2, made in our laboratory), rat serum (Equitech, SRT30-0100) and mouse serum (Equitech, SM30-0500, when mouse-derived antibodies were used) on ice for 10 min. If mouse-derived antibodies or mouse serum was required, IgM or IgG antibodies were stained first. For flow cytometry, cells were washed and stained with antibody cocktails for 20 min on ice, followed by washing and staining in a live/dead discriminator for 15 min on ice. For samples that required a streptavidin-conjugated fluorophore, it was also added with the live/dead stain. Cells were washed and fixed in 1% paraformaldehyde for 30 min, washed again, and resuspended in staining buffer for analysis on BD LSR II and BD LSR Fortessa flow cytometers using BD Diva software v.8.01, or Cytek Aurora spectral cytometers using SpectroFlo software (v.3.0.1). Analysis of cytometry data was performed using FlowJo v.10. The antibodies used for flow and spectral cytometry were: CD45.2-APC (clone 104, BioLegend, 109814; 1:200 dilution), CD19-A488 (clone 1D3, produced and conjugated in our laboratory; 1:400 dilution), CD93-PE (clone AA4.1, BioLegend, 136504; 1:200 dilution), CD38-APC/Cy7 (clone 90, BioLegend, 102728; 1:200 dilution), CD138-BV605 (clone 281-2, BD, 563147; 1:400 dilution), CD80-BV421 (clone 16-10A1, BD, BDB566285; 1:400 dilution), PD-L2-Biotin (clone TY25, BioLegend, 107203; 1:100 dilution), CD21/35-PerCPCy5.5 (clone 7E9, BioLegend, 123416; 1:100 dilution), CD38-PerCPCy5.5 (clone 90, BioLegend, 102722; 1:400 dilution), CD23-PE/Cy7 (clone B3B4, BioLegend, 101614; 1:200 dilution), CD45.1-A488 (clone A20, produced and conjugated in our laboratory; 1:100 dilution), CD19-BUV395 (clone 1D3, BD, 563557; 1:100 dilution), Fas(CD95)-PE/Cy7 (clone Jo2, BD, 557653; 1:200 dilution), PD-L2-PE (clone TY25, Invitrogen, 12-5986-82; 1:200 dilution), CD45.1-PE (clone A20, Tonbo, 50-0453; 1:200 dilution), B220-BUV395 (clone RA3-6B2, BD, 563793; 1:100 dilution), GL7-FITC (clone GL7, BioLegend, 144604; 1:800 dilution), CXCR4-BV421 (clone L276F12, BioLegend, 146511; 1:100 dilution), CD86-PE (clone PO3, BioLegend, 105008; 1:100 dilution), CD19-APC/Cy7 (clone 1D3, BD, 557655; 1:200 dilution), CD19-BV786 (clone 1D3, BD, 563333; 1:400 dilution), CD45.1-APCeFluor780 (clone A20, Invitrogen, 47-0453-82; 1:50 dilution), TCRb-PerCPCy5.5 (clone H57-597, BD, 560657; 1:100 dilution), CD44-A647 (clone Pgp-1, produced and conjugated in our laboratory; 1:400 dilution), CD19-A647 (clone 1D3, produced and conjugated in our laboratory; 1:800 dilution), MHCII-APC/Cy7 (clone M5/114, BioLegend, 107628; 1:200 dilution), CD11c-PE/Cy7 (clone N418, Tonbo, 60-0114; 1:1,600 dilution), CD11b-PE (clone M1/70, BioLegend, 101208; 1:1,600 dilution), GL7-Pacblue (clone GL7, BioLegend, 144614; 1:800 dilution), CD3e-PE (clone 145-2C11, BioLegend, 100308; 1:100 dilution), CD4-BV605 (clone RM4-5, BD, 563151; 1:1,000 dilution), CD38-A594 (clone 90, produced and conjugated in our laboratory; 1:100 dilution), Fas(CD95)-PerCPeFluor710 (clone 15A7, Invitrogen, 46-0951-82; 1:200 dilution), PD-L2-BV480 (clone TY25, BD, 746756; 1:50 dilution), NIP-APC (conjugated in our laboratory; 1:2,000 dilution), NIP-PE (conjugated in our laboratory; 1:4,000 dilution), SA-BUV395 (BD, 564176; 1:100 dilution), Zombie NIR (BioLegend, 423106; 1:500 dilution), Ghost BV510 (Tonbo, 13-0870; 1:250 dilution), Ly6D-FITC (clone 49-H4, BioLegend, 138605; 1:400 dilution), PlexinB2-PE (clone 3E7, BioLegend, 145903; 1:200 dilution), IgG1-BV510 (clone A85-1, BD optibuild, 740121; 1:200 dilution), CD73-BV605 (clone TY/11.8, BioLegend, 127215; 1:50 dilution), B220-PE/Cy7 (clone RA3-6B2, Tonbo, 60-0452; 1:800 dilution).

Bead purification and cell sorting

Cells were blocked as in the flow cytometry step. For cell sorting, bead purification was first performed by staining total splenocytes in a cocktail of anti-CD4-biotin (clone GK1.5, produced and conjugated in our laboratory), anti-CD8-biotin (clone 53-6.7, produced and conjugated in our laboratory), anti-GR1-biotin (clone RB6.8C5, BioLegend, 108404), anti-CD49b-bio (clone DX5, BioLegend, 108908) and anti-TER119-biotin (clone TER119, produced and conjugated in our laboratory). For naive mouse spleens, anti-CD43-biotin (clone S7, produced and conjugated in our laboratory), anti-CD11b-biotin (clone M1/70, produced and conjugated in our laboratory) and anti-CD11c-biotin (clone N418, produced and conjugated in our laboratory) were also added. Cells were stained on ice for 15 min, washed and resuspended in magnetic streptavidin-coated beads (BD iMag Streptavidin beads, 90 µl beads per ml of buffer) and incubated on ice for 5 min. Cells were brought to 6 ml of staining buffer without azide (sorting buffer). The tubes were then placed in a cold magnet and incubated for 3 min. To increase yield, the supernatant was poured into a new tube, and the original tube was resuspended in sorting buffer. The tube was placed on the magnet again and the process was repeated. The second supernatant was combined with the first, and the cells were washed and resuspended in sorting buffer for counting. Around 1×10^8 cells were then stained for sorting with 7AAD (Tonbo, 13-6993) used as a live/dead discriminator. Sorting was performed by a 5-laser BD Aria II with the 70-µm nozzle into 200 µl of R10.

In vitro NB-21 feeder cultures

NB-21 feeder cells were obtained from H. Singh. Cells were maintained in 175 mm^3 cell culture flasks in DMEM with glucose (Thermo Fisher Scientific, 11965092), supplemented with penicillin–streptomycin and 10% FCS (HyClone, SH30396.03). One day before sorting B cells, feeders were detached from the plate using pre-warmed PBS with 5 mM EDTA, and 1,000 cells per well were plated in a flat-bottom 96-well plate in 200 µl R10 (10% Fetalplex replaced with 10% Hyclone FCS for in vitro cultures). On the day of sorting, 2 µl of 200 ng ml^{-1} (100×) IL-4 (PeproTech, 214-14) was added to each well immediately before sorting. MBC populations were prepped for sorting as stated above, and 400 cells of each population from individual mice were sorted directly into each well. Cultures were incubated at 37 °C, 5% CO_2 for 2 d. At day 2, the medium was completely removed and replaced with fresh R10 without IL-4. On day 4, the medium was completely removed, and 100 µl pre-warmed PBS with 5 mM EDTA was added to detach cells. A total of 150 µl of R10 was added after detachment, cells were spun at 300*g*, for 5 min at 25 °C, and resuspended in R10 for counting. Cells were counted using Trypan blue exclusion and size discrimination; B cells appeared much smaller and rounder than feeder cells and were not present in feeder-only controls.

ELISPOTs

Immulon 4 HBX ELISPOT plates (VWR, 62402-959) were coated with 50 µl of 5 µg ml^{-1} NIP_{17}-BSA overnight in PBS at 4 °C. Liquid from plates was discarded, 50 µl of blocking buffer (PBS, 1% BSA, 0.03% sodium azide) was added, and plates were incubated at 25 °C for at least 1 h. Plates were washed 3× with ELISA wash buffer (PBS with 10% Tween-20, diluted 10× in H_20) then 2× with H_20, liquid was discarded, and 200 µl of 25 °C R10 was placed in each well. Appropriate numbers of cells were added to the top well, and volume was brought to 300 µl in R10 if needed. Next, 3× dilutions were performed down the plate by removing 100 µl from the top well and adding this to the well below. Plates were spun at 300*g*, for 3 min at 25 °C, and then incubated at 37 °C and 5% CO_2 for 5 h. After incubation, plates were washed again with the 3× ELISA wash buffer, 2× H_20 protocol and then 50 µl of goat-ant-mouse IgM-AP or IgG1-AP (Southern Biotech, 1021-04 and 1071-04) in PBS with 2% Fetalplex was added to appropriate wells. Plates were incubated for 1 h at 25 °C or 4 °C overnight. Plates were then washed with 3× ELISA wash buffer, 2× H_20 protocol, and 50 µl of warmed (≥37 °C) BCIP-agarose solution was added; 5 ml BCIP buffer (12.1 g Tris base, 5.8 g NaCl, 1 g $MgCl_2 \times 6H_20$, pH to 9.5), 1 ml 3% liquid low-melt agarose (Fisher Bioreagents, BP-165-25), 33 µl of 0.1 g ml^{-1} BCIP (Thermo Fisher, 34040) in dimethylformamide. Plates were allowed to solidify and develop at 25 °C for 1 h, then placed at 4 °C and large, dark spots were counted the following day under a microscope.

Statistics and reproducibility

Statistics outside next-generation sequencing data were performed using GraphPad Prism (v.9.0.0) or R base statistical package (v.4.1.0). Data distribution was assumed to be normal. The same software was also used for plotting figures. The VennEuler R package (v.1.1) was used for plotting Venn diagrams, base R functions and Prism were used for box plots and bar plots, the pheatmap and ComplexHeatmap R packages were used for plotting heatmaps, and the ggplot2 R package (v.3.4.1) was used to plot x–y plots. Binomial tests were used to determine P values for group sizes in Fig. 2c,d. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment. No statistical methods were used to predetermine sample sizes, but our sample sizes are standard for the field. Data distribution was assumed to be normal, but this was not formally tested.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All bulk RNA-seq data have been deposited in the NCBI's Gene Expression Omnibus (GEO) database and are publicly available under accession number GSE225748.
All bulk ATAC-seq data have been deposited in the NCBI GEO database and are publicly available under accession number GSE225672.
All scRNA-seq data have been deposited in the NCBI GEO database and are publicly available under accession number GSE225673.
The mm10 genome database (https://www.ncbi.nlm.nih.gov/assembly/GCF_000001635.20/) was used to align sequences for the RNA-seq analysis and to align sequencing reads for the ATAC-seq analysis.
All raw data and materials will be made available to investigators upon request. Source data are provided with this paper.

References

44. Corces, M. R. et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* **14**, 959–962 (2017).
45. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
46. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
47. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
48. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
49. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
50. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
51. Stoeckius, M. et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).
52. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902 (2019).

Acknowledgements

We thank H. Singh for extensive discussions and critical reading of the manuscript; S. Naik (New York University), J. Sun (Sloan Kettering Institute) and C. Lau (Cornell University) for critical reading of the manuscript. We thank T. Marinov, M. Berkey, L. Conter and M. Price for technical support; University of Pittsburgh DLAR for excellent animal husbandry; members of the Shlomchik Lab Memory Group for suggestions and discussion throughout the development of this work. We thank the University of Pittsburgh Innovative Technology Development Core, led by S. Gingras, and the Mouse Embryo Services Core, led by C. Bi, for generating the GCET-TamCre mice. This research was supported in part by the University of Pittsburgh Center for Research Computing, RRID:SCR_022735, through the resources provided. Specifically, this work used the HTC cluster, which is supported by National Institutes of Health (NIH) award number S10OD028483. This research was also supported by NIH grants R01-AI43603 and R01-AI105018 to M.S. D.C. was supported by NIH T32 AI060525 (J. Flynn, principal investigator).

Author contributions

D.C. performed most experiments and analyzed all experiments. S.J. performed some GCET-TamCre experiments, and F.W. performed the scRNA-seq experiment. F.W. drove the development of the GCET-TamCre mouse. K.H., under the supervision of S.K., performed Vh186.2 immunoglobulin somatic hypermutation analysis. S.S. completed quality checks and processed bioinformatic data and was advised by M.C. Bioinformatic analyses were performed by D.C. and S.S. The project was conceived by M.S. and D.C., who both designed experiments and interpreted data. M.S. and D.C. wrote the manuscript with input from all the authors.

Competing interests

The authors declare no competing interests.

Additional information

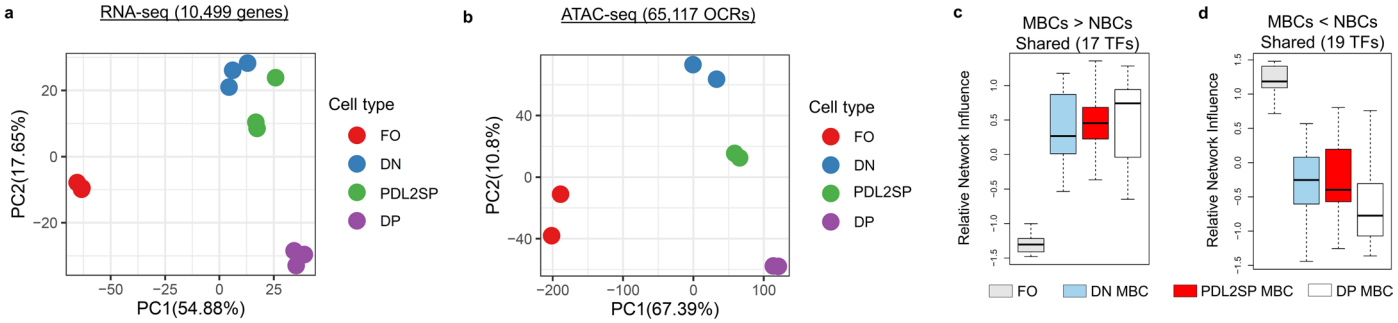
Extended data is available for this paper at https://doi.org/10.1038/s41590-023-01721-9.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41590-023-01721-9.

Correspondence and requests for materials should be addressed to Mark Shlomchik.

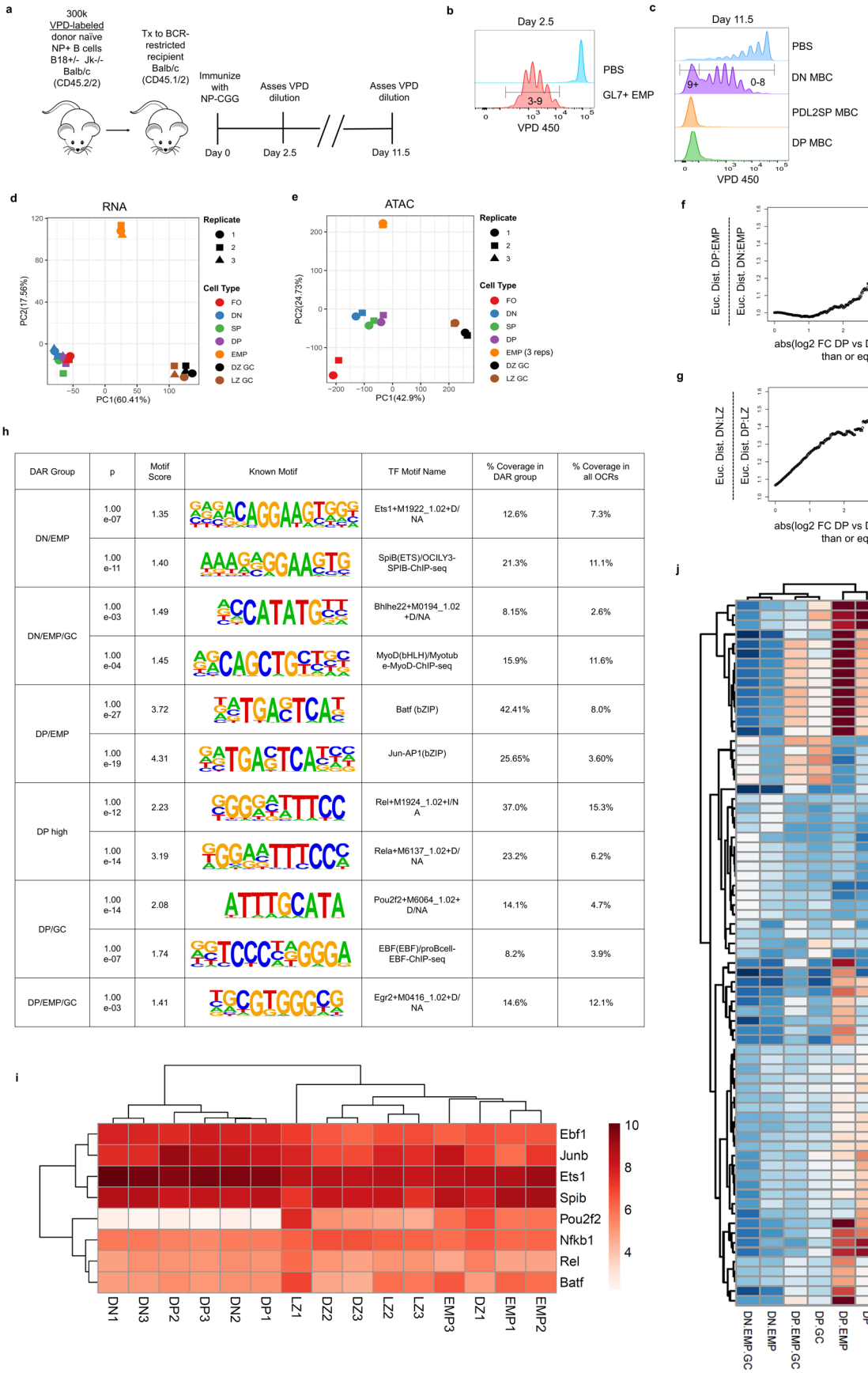
Peer review information *Nature Immunology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editor: L. A. Dempsey, in collaboration with the *Nature Immunology* team. Peer reviewer reports are available.

https://doi.org/10.1038/s41590-023-01721-9



Extended Data Fig. 1 | Distinct transcriptomic and epigenetic profiles of FO NBCs and MBC subsets. a, PCA plot of genes with expression of >= 0 log2 normalized reads in at least one cell type (n=3 per cell type). **b**, PCA plot of OCRs with accessibility of >= 1 log2 normalized reads in at least one cell type (n=2 per cell type). **c**,**d**, Boxplots showing relative PageRank score for TFs with higher (**c**) or lower (**d**) scores in all MBC subsets compared to FO naïve. Boxplots display median values and lower and upper quartiles, the ranges display min and maximum values. n= 2 per cell type.

https://doi.org/10.1038/s41590-023-01721-9



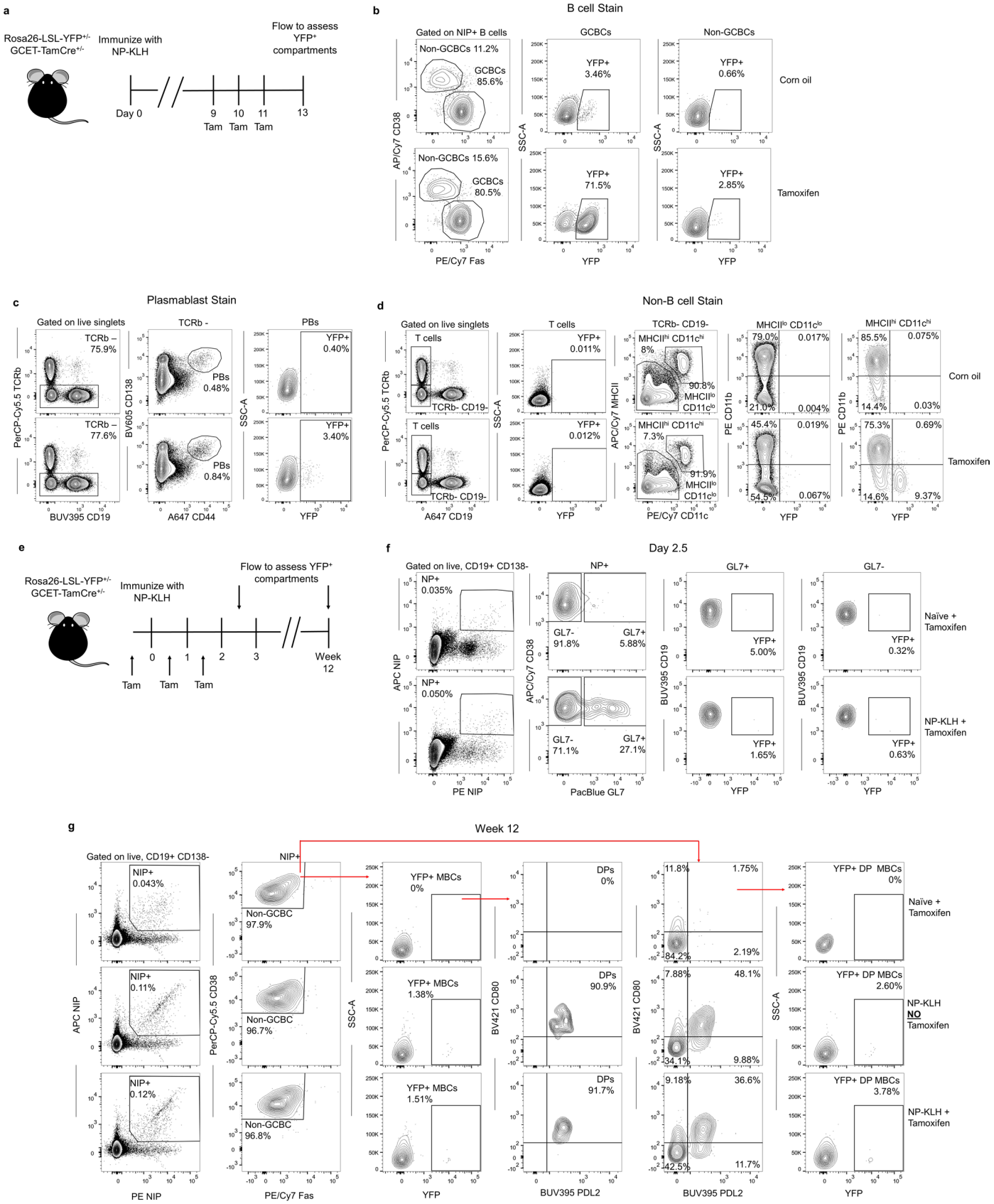
DAR Group	p	Motif Score	Known Motif	TF Motif Name	% Coverage in DAR group	% Coverage in all OCRs
DN/EMP	1.00 e-07	1.35		Ets1+M1922_1.02+D/NA	12.6%	7.3%
	1.00 e-11	1.40		SpiB(ETS)/OCILY3-SPIB-ChIP-seq	21.3%	11.1%
DN/EMP/GC	1.00 e-03	1.49		Bhlhe22+M0194_1.02+D/NA	8.15%	2.6%
	1.00 e-04	1.45		MyoD(bHLH)/Myotube-MyoD-ChIP-seq	15.9%	11.6%
DP/EMP	1.00 e-27	3.72		Batf (bZIP)	42.41%	8.0%
	1.00 e-19	4.31		Jun-AP1(bZIP)	25.65%	3.60%
DP high	1.00 e-12	2.23		Rel+M1924_1.02+I/NA	37.0%	15.3%
	1.00 e-14	3.19		Rela+M6137_1.02+D/NA	23.2%	6.2%
DP/GC	1.00 e-14	2.08		Pou2f2+M6064_1.02+D/NA	14.1%	4.7%
	1.00 e-07	1.74		EBF(EBF)/proBcell-EBF-ChIP-seq	8.2%	3.9%
DP/EMP/GC	1.00 e-03	1.41		Egr2+M0416_1.02+D/NA	14.6%	12.1%

Extended Data Fig. 2 | See next page for caption.

https://doi.org/10.1038/s41590-023-01721-9

Extended Data Fig. 2 | DAR groups are enriched for distinct TF motifs. a, Schematic of BALB/c transfer system using VPD-labeled donor cells. Mice were euthanized at days 2.5 and 11.5**. b,c** Representative flow plots depicting VPD-dilution of listed cell types compared to total transferred B cells in PBS control (unimmunized) at days 2.5 (**b**) and 11.5 (**c**). **d**, PCA plot of genes with expression of >= 0 log2 normalized reads in at least one cell type. **e**, PCA plot of OCRs with accessibility of >=1 log2 normalized reads in at least one cell type. **f**, Ratio of Euclidean distance between DPs and EMPs to Euclidean distance between DNs and EMPs for given log2 FC RNA-seq differential expression thresholds, where Fig. 2c is the reference at abs(log2 FC DP vs DN) >=1. For Euclidean distance, a higher number indicates less similarity; if Euclidean distance ratio is > 1, numerator is less similar than denominator. **g**, Ratio of Euclidean distance between DNs and LZ GCs to Euclidean distance between DPs and LZ GCs for given log2 FC RNA-seq differential expression thresholds, where Fig. 2c is the reference at abs(log2 FC DP vs DN) >=1. **h**, Table of top TF motifs enriched in various DAR groups from Fig. 2d compared to background (all called peaks), computed via HOMER. Examples were manually chosen. p = p.value (calculated using cumulative binomial distribution), Motif Score = motif enrichment score above background, Known Motif = motif taken from HOMER database, TF Motif Name = specific TF and/or TF family, along with source if provided, % Coverage in DAR group = % of OCRs in a DAR group that have at least one motif sequence, % Coverage in all OCRs = % of all called OCRs that have at least one motif sequence. **i**, Heatmap depicting log2 normalized RNA expression of select TFs from Fig. 2f. **j**, Heatmap depicting all TF motifs with enrichment scores of log2 FC > = 1.4 above background, p. value <= 0.001 for any DAR group. Values shown are motif enrichment scores above background.

https://doi.org/10.1038/s41590-023-01721-9

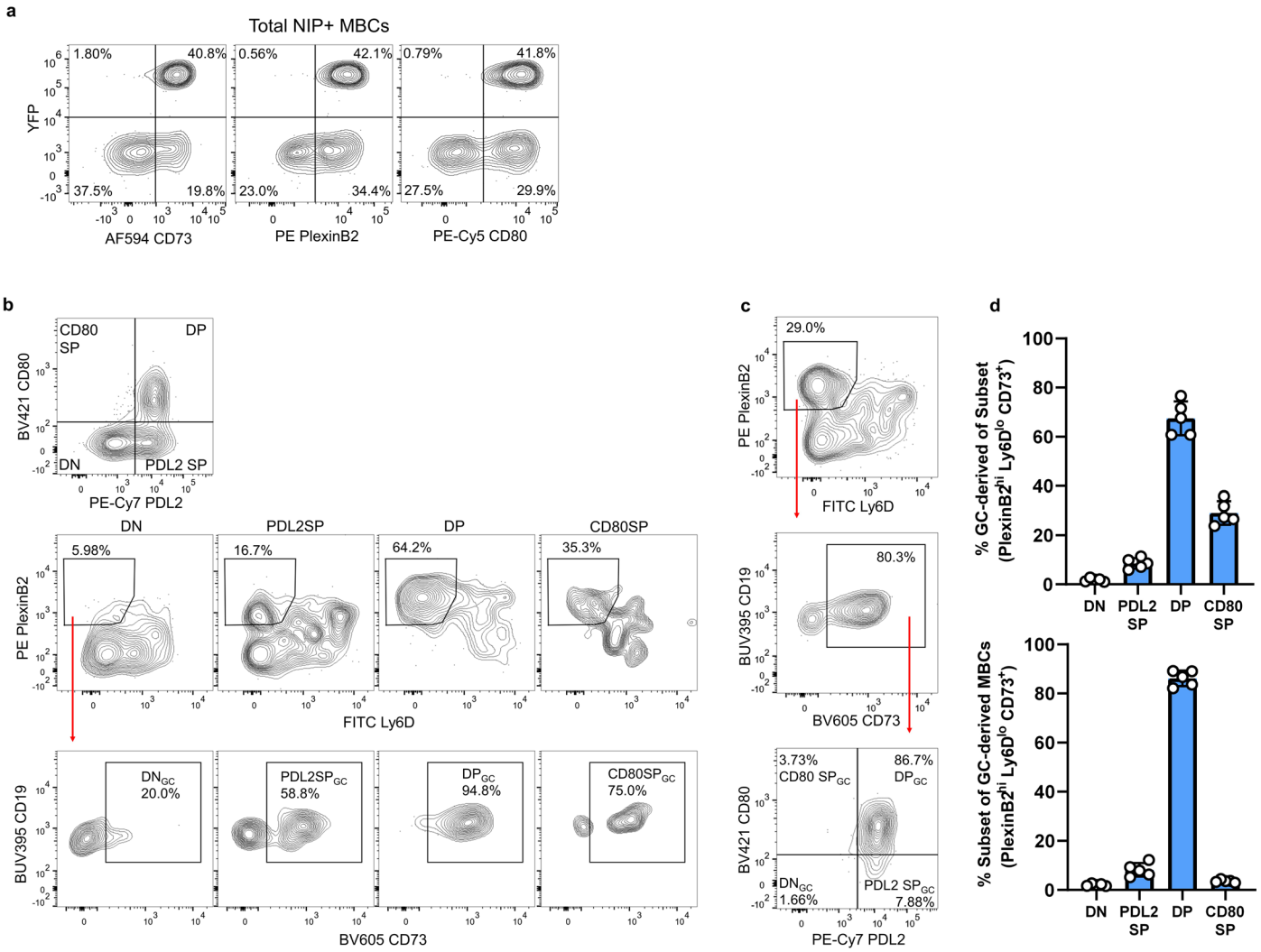


Extended Data Fig. 3 | See next page for caption.

https://doi.org/10.1038/s41590-023-01721-9

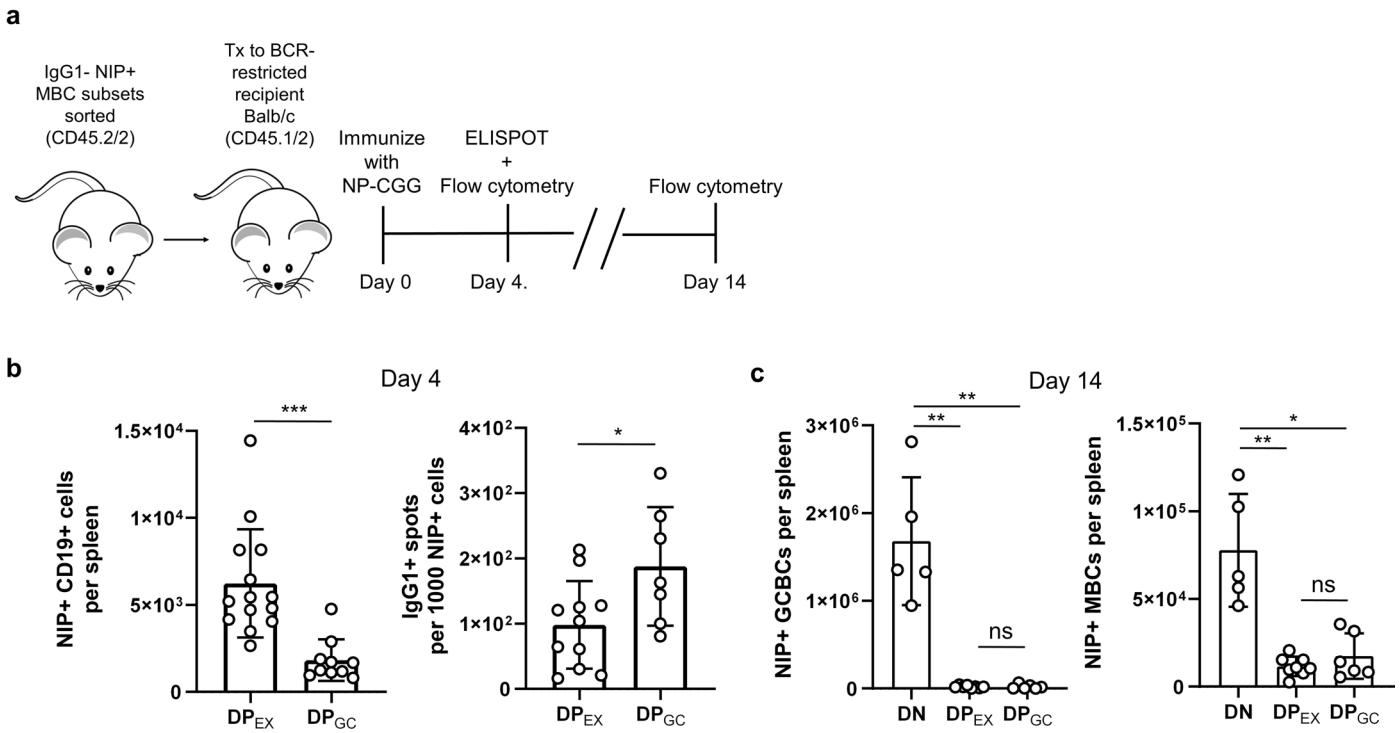
Extended Data Fig. 3 | The GCET-TamCre is GCBC-specific. a, Schematic to assess efficiency of GCBC-labeling and non-specific labeling in GCET-TamCre lineage tracing mice. Mice were given 1 mg tamoxifen on days 9, 10, and 11, post-immunization with NP-KLH and euthanized on day 13. **b**, Representative flow plots depicting YFP expression in NIP^+ GCBCs and NIP^+ non-GCBCs of corn oil control and tamoxifen-treated mice at day 13. **c**, Representative flow plots depicting YFP expression in plasmablasts of corn oil control and tamoxifen-treated mice at day 13. **d**, Representative flow plots depicting YFP expression in non-B cells of corn oil control and tamoxifen-treated mice at day 13. Left two columns depict a total T cell stain, while right 3 columns depict a myeloid lineage stain. **e**, Schematic to assess efficiency of B cell labeling GCET-TamCre lineage tracing mice. Naïve mice were given 1 mg tamoxifen every day for 3 consecutive days, whereas NP-KLH immunized mice were given 1 mg tamoxifen on days −0.5, 0.5, and 1.5 post-immunization. Mice were euthanized 1 day after final tamoxifen dose or 12 weeks later. **f**, Representative flow plots depicting YFP expression in naïve mice after 3 tamoxifen doses (top row) and YFP expression in early activated B cells at day 2.5 after 3 tamoxifen doses (bottom row). **g**, Representative flow plots depicting YFP expression in B cells of naïve mice after 3 tamoxifen doses and chasing out to 12 weeks (top row), YFP expression at week 12 post-immunization in mice not given tamoxifen (middle row), and YFP expression at 12 weeks post-immunization in mice given tamoxifen at days −0.5, 0.5, and 1.5 post-immunization.

https://doi.org/10.1038/s41590-023-01721-9



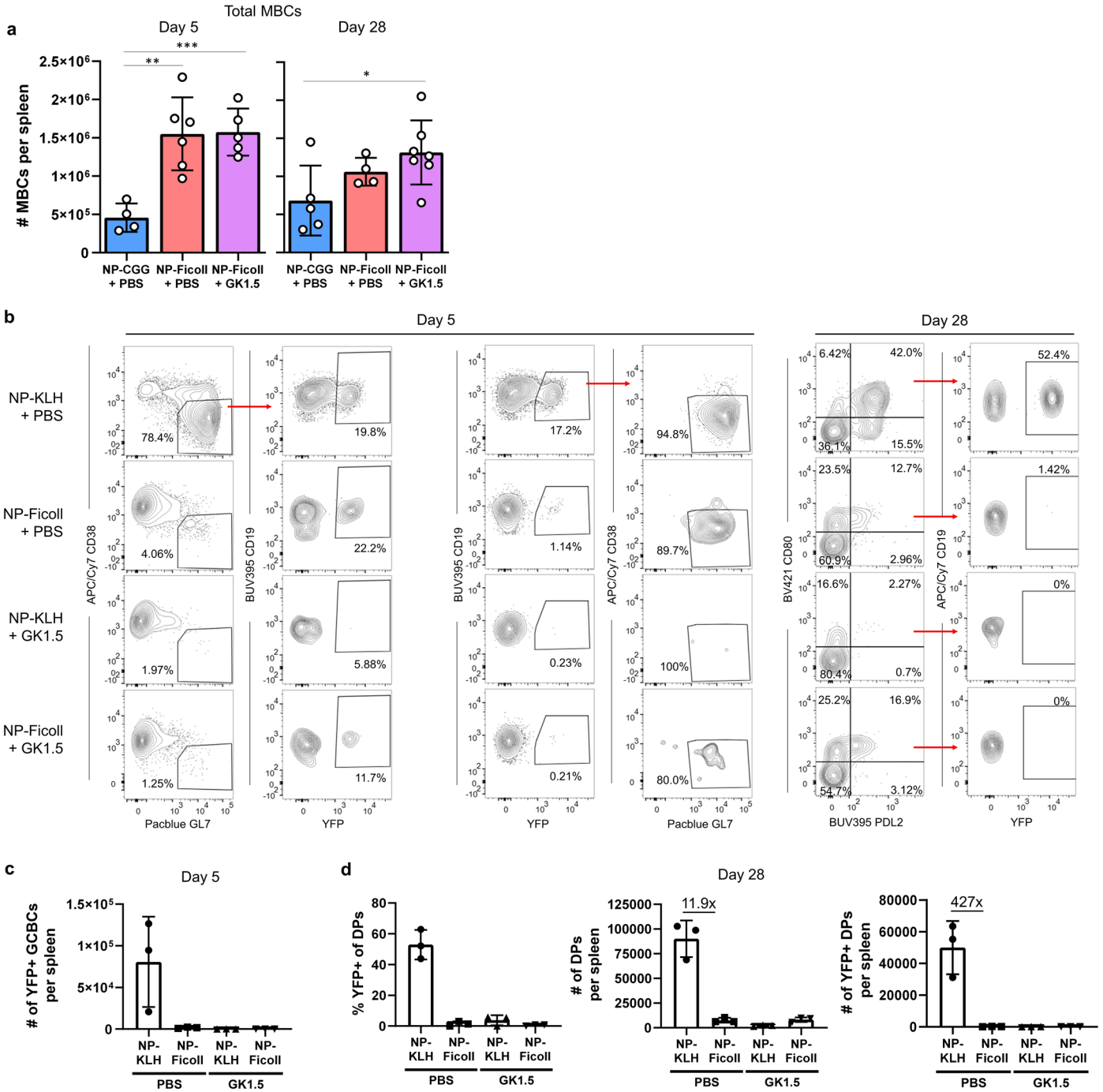
Extended Data Fig. 4 | Ly6D, PlexinB2, and CD73 can be used to identify GCBC-derived MBCs in WT BALB/c mice. a, Flow cytometry plots depicting markers expressed by YFP$^+$ MBCs in week 6 KP-KLH-immunized GCET-TamCre$^{+/-}$R26-LSL-YFP$^{+/-}$ mice (from Fig. 3g; tamoxifen given on days 3–13). **b-d**, NIP$^+$ Memory cells from directly immunized WT BALB/c mice 8 weeks post-NP-CGG immunization. **b**, Flow cytometry gating strategy to determine GC-derived MBCs (PlexinB2^{hi}Ly6D^{lo}CD73$^+$) frequencies within CD80/PD-L2 MBC subsets. **c**, Flow cytometry gating strategy to determine frequencies of CD80/PD-L2 MBC subsets within total GC-derived MBCs (PlexinB2^{hi}Ly6D^{lo}CD73$^+$). **d**, top: percent of GC-derived cells within each CD80/PD-L2 MBC subset; bottom: percent of each CD80/PD-L2 subset within total GC-derived MBCs (n = 5). Bars display mean ± s.d.

https://doi.org/10.1038/s41590-023-01721-9



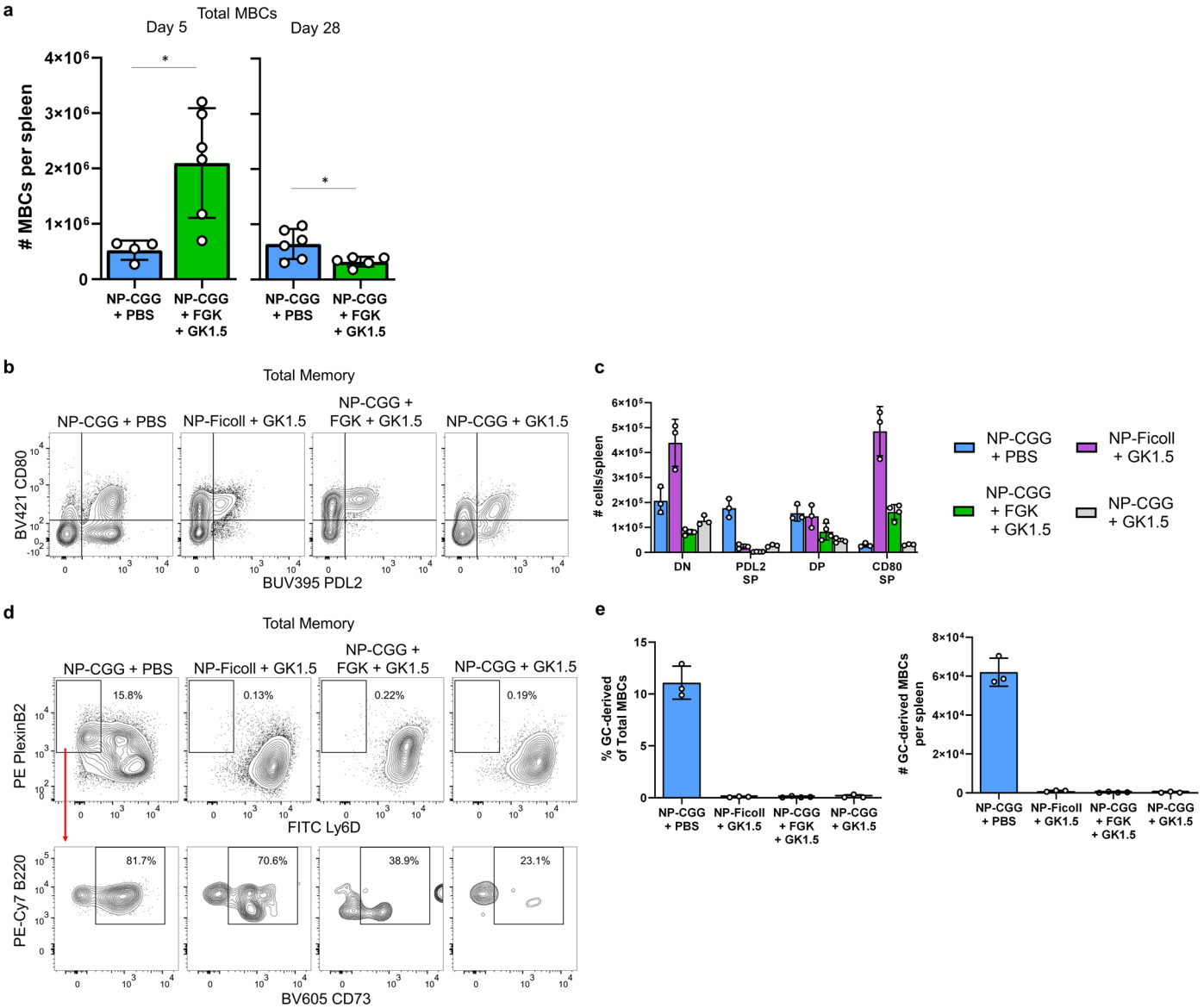
Extended Data Fig. 5 | DP_{EX} and DP_{GC} have distinct functions *in vivo*. **a**, Schematic for *in vivo* transfer and re-activation of 5,000 sorted $IgG1^-$ NIP^+ MBC subsets. **b**, Day 4 *in vivo* data; left: $NIP^+CD19^+CD45.1^-$ cell counts per spleen, counted by flow cytometry (n = 14 for DP_{EX}, n = 10 for DP_{GC}, 4 independent experiments); right: Number of $IgG1^+$ NIP^+ spots per 1000 NIP^+ cells (derived from left panel, (n = 11 for DP_{EX}, n = 7 for DP_{GC}, 3 independent experiments). **c**, Day 14 in vivo data; left: number of $NIP^+CD19^+CD138^-CD45.1^-CD45.2^+CD38^-CD95^+$ GCBCs per spleen; right: number of $NIP^+CD19^+CD138^-CD45.1^-CD45.2^+CD38^+$ GCBCs per spleen. Both panels counted via flow cytometry (n = 5 for DN, n = 8 for DP_{EX}, n = 6 for DP_{GC}, 3 independent experiments). Bars display mean ± s.d. P-values were calculated using two-tailed Welch's t-test (*P < = 0.05, **P < 0.01, ***P < 0.001). Actual p-values are listed in source data.

https://doi.org/10.1038/s41590-023-01721-9



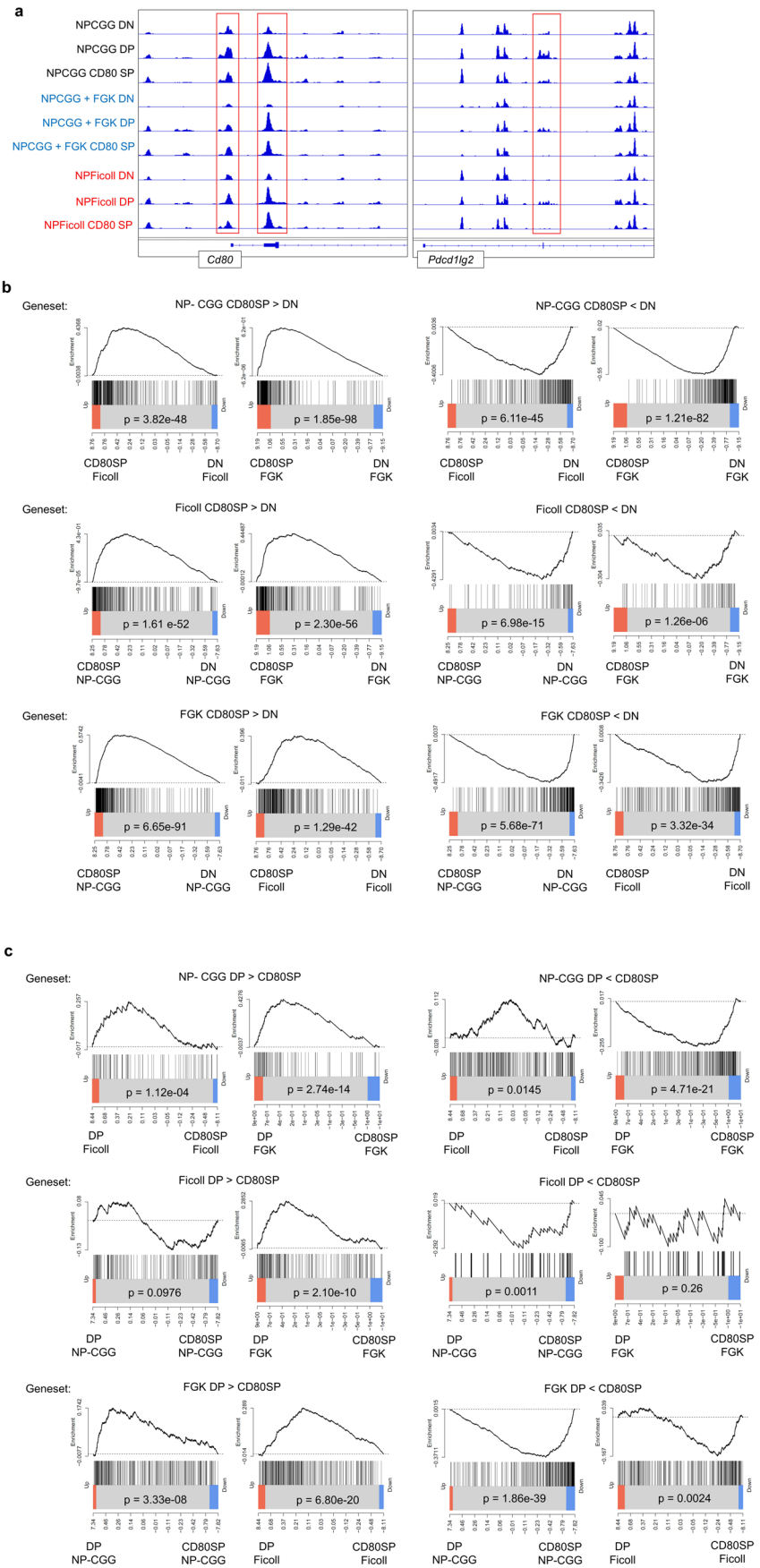
Extended Data Fig. 6 | NP-Ficoll does not produce GCBC-derived MBCs. **a**, Number of total MBCs at Day 5 (left) and Day 28 (right) between NP-CGG + PBS (day 5; n = 4, day 28, n = 5), NP-Ficoll + PBS (day 5; n = 6, day 28, n = 4), and NP-Ficoll + GK1.5 (day 5; n = 5, day 28, n = 7) treated mice from Fig. 5. **b**, Left/middle: Representative flow plots of pre-gated NIP$^+$ B cells from Day 5 post-immunization for NP-KLH + PBS, NP-Ficoll + PBS, NP-KLH + GK1.5, and NP-Ficoll + GK1.5 treated GCET-TamCre$^{+/-}$R26-LSL-YFP$^{+/-}$ mice. Mice were given one dose of tamoxifen on Day 3; right: Representative flow plots of pre-gated NIP + B cells from Day 28 post-immunization for NP-KLH + PBS, NP-Ficoll + PBS, NP-KLH + GK1.5, and NP-Ficoll + GK1.5 treated GCET-TamCre$^{+/-}$R26-LSL-YFP$^{+/-}$ mice. Mice were given one dose of tamoxifen every other day from day 3–13. **c**, Number of YFP$^+$ GCBCs (NIP$^+$CD 19$^+$CD138$^-$CD38$^-$GL7$^+$) per spleen at day 5 (n = 3 per group). **d**, left: Percent YFP$^+$ of DP MBCs (NIP$^+$CD19$^+$CD138$^-$CD38$^+$GL7$^-$), middle: Number of total DP MBCs per spleen, and right: Number of YFP$^+$ DPs per spleen at day 28 (n = 3 per group). Bars display mean ± s.d. P-values were calculated using two-tailed Welch's t-test (*P < = 0.05, **P < 0.01, ***P < 0.001). Actual p-values are listed in source data.

https://doi.org/10.1038/s41590-023-01721-9



Extended Data Fig. 7 | TI-derived MBCs are not GCBC-derived. a, Number of total MBCs at Day 5 (left) and Day 28 (right) between NP-CGG + PBS control (day 5; n = 4, day 28, n = 6) and NP-CGG + GK1.5 + FGK (day 5; n = 6, day 28, n = 5) treated mice. **b**, Representative flow plots of NIP$^+$ MBCs from BALBc transfer system under indicated immunizations at week 4. **c**, Number of MBC subsets per spleen for each indicated immunization. **d**, Representative flow plots of NIP$^+$ MBCs from BALB/c transfer system under indicated immunizations at week 4, showing gating to determine GC-derived MBCs. **e**, left: percent of; right: number of; total MBCs displaying a GC-derived phenotype in the 4 different immunizations. For **b**-**e**, n = 3, except NP-CGG + FGK + GK1.5, where n = 4. Bars display mean ± s.d. P-values were calculated using two-tailed Welch's t-test (*P < = 0.05). Actual p-values are listed in source data.

https://doi.org/10.1038/s41590-023-01721-9

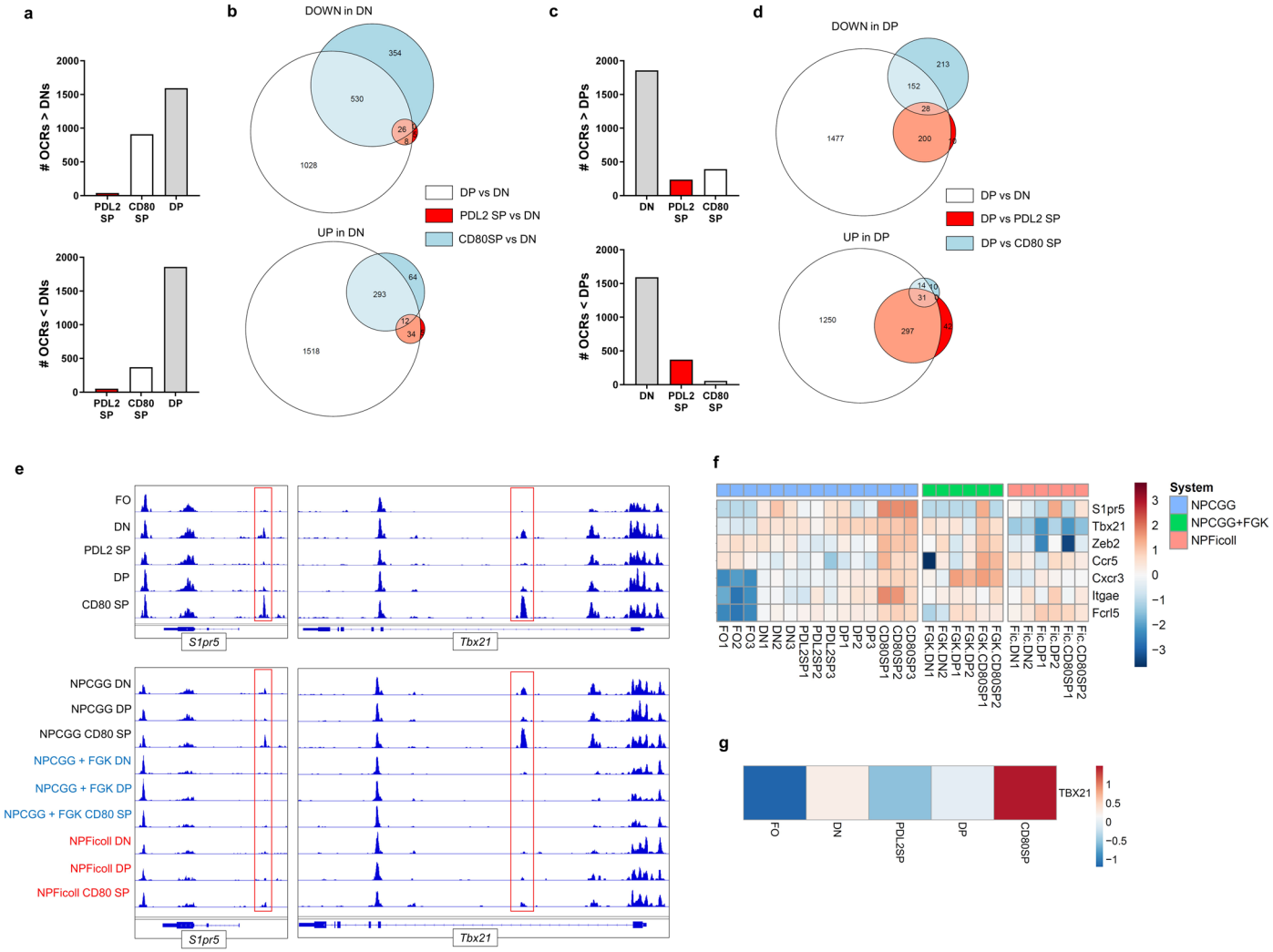


Extended Data Fig. 8 | See next page for caption.

https://doi.org/10.1038/s41590-023-01721-9

Extended Data Fig. 8 | A shared CD80SP epigenetic signature can be found in TD and TI systems. a, Examples of OCRs at *Cd80* and *Pdcd1lg2* that display a consistent pattern of accessibility in CD80SP MBCs within each system (NP-CGG + PBS, NP-Ficoll + GK1.5, NP-CGG + FGK + GK1.5, MBCs from day 28. n = 2 per cell types). **b**, GSEA plots using DARs between CD80SP MBCs and DN MBCs as genesets, showing all possible comparisons across all different systems. **c**, GSEA plots using DARs between CD80SP MBCs and DP MBCs as genesets, showing all possible comparisons across all different systems. Cutoff of log2FC >= 1, p-value <= 0.05 (calculated via two-tailed permutations), accessibility >= 2 log2 normalized counts was used for all genesets; the top 300 DARs, ordered by log2 FC, were used for the geneset unless total DARs were fewer than 300. Number of total DARs (in parentheses): NP-CGG; CD80SP > DN (884), CD80SP < DN (383), DP > CD80SP (93), DP < CD80SP (652). NP-Ficoll + GK1.5; CD80SP > DN (395), CD80SP < DN (83), DP > CD80SP (184), DP < CD80SP (37). NP-CGG + FGK + GK1.5; CD80SP > DN (1,332), CD80SP < DN (503), DP > CD80SP (499), DP < CD80SP (665).

https://doi.org/10.1038/s41590-023-01721-9



Extended Data Fig. 9 | TD-derived CD80SPs exhibit an inflammatory signature. a, Bar graphs depicting OCRs that differ between DN MBCs and each other MBC subset. Top plot depicts OCRs lower in DNs than each MBC subset (x-axis); bottom plot depicts OCRs higher in DNs. Differential accessibility cutoff for each MBC subset compared to DNs is log2FC >= 1, FDR < = 0.05, and log2 accessibility >=1 normalized counts. **b**, Venn diagrams plotting the intersection of DARs that were more accessible in any MBC subset compared to DNs on the top and less accessible compared to DNs on the bottom, as enumerated in **a. c**, Bar graphs depicting OCRs that differ between DP MBCs and each other MBC subset. Top plot depicts OCRs lower in DPs than each MBC subset (x-axis); bottom plot depicts OCRs higher in DPs. Differential accessibility cutoff for each MBC subset compared to DPs is log2FC >=1, FDR < = 0.05, and log2 accessibility >=1 normalized counts. **d**, Venn diagrams plotting the intersection of DARs that were more accessible in any MBC subset compared to DPs on the top and less accessible compared to DPs on the bottom, as enumerated in **c. e**, Select DARs near *S1pr5* and *Tbx21* showing specificity to CD80SP MBCs from the NP-CGG system. **f**, RNA expression heatmap (row z-scored) of select Tbet/ABC-associated genes among different systems and cell types (n = 3 for NP-CGG, n = 2 for other systems). **g**, PageRank score (z-scored) of Tbx21 from NP-CGG derived cell types.

nature portfolio

Corresponding author(s): Mark J. Shlomchik

Last updated by author(s): 11-24-2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

Flow cytometry data were collected with either BD LSRII or BD LSRFortessa using BD DIva software version 8.01 or with Cytek Aurora using SpectroFlo software (v3.0.1). Cell sorting was performed using a BD FACSAria II.

RNA-seq: Libraries were generated using the SMART-Seq v4 Ultra Low Input RNA kit followed by Nextera FLEX using UDI set B indexes. Samples were sequenced using NovaSeq 6000 or NextSeq 2000 sequencers.

ATAC-seq: Libraries were generated using adapters (IDT for Illumina Nextera DNA Unique Dual Indexes, Illumina, #20027215) were added during amplification using 2X Phusion High Fidelity Mastermix with HF Buffer (ThermoFisher #F531L). Libraries were cleaned up with Zymo DNA Clean & Concentrator-5 columns (Zymo Research #D4004) and assessed with a Qubit Flex Fluorometer (ThermoFisher #Q33327) and a TapeStation 4200 (Agilent #G2991BA). Samples were sequenced to obtain 20M 2x75 bp paired-end reads using an Illumina NextSeq 550 sequencer.

scRNAseq: TotalSeq anti-mouse Hashtag Antibodies (Biolegend) were added at a 1:50 dilution to each of the sorted samples (DN MBC: TotalSeq-A0310, PD-L2 SP MBC: TotalSeq-A0311, DP MBC: TotalSeq-A0312) . Cells were stained on ice in PBS with 0.04% BSA for 20 minutes and then washed once in PBS with 0.04% BSA. Cells were counted and 5,600 of each cell type were loaded into the 10x Genomics Chromium system per the manufacturer's instructions, with a ~57% input recovery. Gene expression and antibody hashtag/feature barcode libraries were generated, their quality was assessed through the Agilent TapeStation High Sensitivity D5000 Screentape, and their amounts were quantified with the KAPA Library Quantification Kit for Illumina Platforms. We used the 3' V2 library; the feature barcode library was generated according to the New York Genome Center protocol. For hash-tagging, we followed the "feature barcode" instructions from the manufacturer. Libraries were pooled and sequenced by Medgenome on a NovaSeq 6000.

Data analysis

Flow and spectral cytometry data analysis were performed in FlowJo v10. Flow data was compiled and analyzed in GraphPad Prism (v. 9.0.0).

RNA-seq: Samples were aligned to the mm10 genome using the STAR aligner. Gene-level counts were determined using featureCounts (v.2.0.1), and raw counts were quantile normalized to each other for differential expression using the voom method in the Limma R package. For normalization of the datasets, the Quantile method was used. All gene-set enrichments were performed using the rankSumTestWithCorrelation function in limma, which explicitly corrects for correlation among genes in the gene set being interrogated.

ATAC-seq: All sequencing reads were trimmed with Trim Galore (v.0.6.5) and aligned with their reference genome (mm10) using Bowtie2 (v.2.3.3) aligner (using default parameters) to produce raw bam files. Further, mitochondrial reads, unmapped reads and low mappability (<30) reads were filtered out using the samtools view. PCR duplicates were removed using picard (v. 2.18.12) to obtain final filtered bam files. Final bam files were converted to bigwig files using deeptools (v.3.3.0) bamCoverage tool (parameter -bs=1 --normalizeUsing RPKM -p=max) and bigwig files were used to view genome tracks in IGV (Broad Institute v. 2.3.97). Peaks and peaks summit were called using MACS2 (with parameter -q 0.05 -B --nomodel --shift -75 --extsize 150) after merging filtered reads from replicate 1 and replicate 2 from all samples into two files. The Irreproducible Discovery Rate (IDR v.2.0.3) analysis was used to evaluate the reproducibility between the two biological replicates from all samples. Peaks passing the soft IDR threshold (IDR <=0.1) and rank based on p-value were retained for further analysis. Blacklisted peaks were removed. Raw counts overlapping each peaks summit extended 100bp up and down (in such a way so that peaks are in the boundary of a called narrow peak region) were quantified using bwtool summary (v.1.0.0) and normalized using the voom method in the limma (v.3.52.2) R package. Differential accessibility analysis was performed using "limma" R package. Transcription factor motif enrichment analysis for selected peaks were performed using HOMER24 (v.4.9.1).

scRNAseq: FASTQ files were generated and aligned to the mouse reference genome mm10 with Cell Ranger 5.0.0 to produce the gene-cell count matrix and cell-antibody count matrix. The 10x raw data from each sample were demultiplexed and FASTQ files were generated using the "mkfastq" Cell Ranger pipeline (v5.0.0, 10X Genomics). Cell Ranger "count" was used to align reads to the mm10 reference genome, and mRNA transcript, and HTO unique molecular identifier (UMI) quantification tables were generated. The raw barcode matrix files generated from the Cell Ranger pipeline were further utilized for downstream analysis using the Seurat package (v.3.0.2) in R (v.3.4.3). Cells expressing less than 200 genes, or with greater than 10% of UMIs that mapped to mitochondrial DNA, were filtered out. The HTO tables were added to the dataset and normalized by a centered-log ratio (CLR) method using the "NormalizeData" function. The normalized HTO count was used to determine if each gel bead-in-emulsion contained a single cell using the Seurat "MULTIseqDemux" function and manual inspection of cells, where a cell was considered "singlet" if expression of a single HTO accounts more than 70% of the total HTO expression in that cell; otherwise, the cell was considered a "doublet" and removed. UMI, mitochondrial content, hemoglobulin gene and ribosomal gene content scores were "regressed-out" using Seurat's "ScaleData" function. Variable genes were detected using the "mean.var.plot" method in "FindVariableFeatures" function with default cutoff. These variable genes were used for dimensionality reduction based on principal component analysis (PCA) using the "RunPCA" function. "ElbowPlot" was used to assess the first 50 principal components and the principal components that account for the largest variability in the data were selected for further Uniform Manifold Approximation and Projection (UMAP) dimensional reduction and clustering analysis. In order to identify distinct groups of cells, unsupervised clustering was performed using the "FindClusters" function, which calculates the k-nearest neighbors according to variable gene expression in all cells, thereby constructing a shared nearest neighbor graph using the Louvain algorithm. To avoid overclustering, we tested different resolution ("res") parameters, ranging from 0.1 to 2 in increments of 0.1 and the clustering progression was assessed and visualized using "Clustree" (v 0.4.3). Optimal resolution was determined based on continued separation prior to "overclustering" as observed by the increasing crossover between clusters. Based on these observations we chose the resolutions "0.4" and "0.1", respectively for MBC and DP cluster. Cell clusters were visualized using UMAP dimensional reduction plots. The "FindAllMarkers" function with default settings was utilized to find DEGs in each cluster, in comparison to all other clusters, using the Wilcoxon Rank Sum test with genes detected in a minimum of 10% of cells, a minimum of 0.25 average log-fold change, and a minimum of 0.01 Bonferroni-adjusted P-value.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All bulk RNA-seq data have been deposited in the NCBI's Gene Expression Omnibus (GEO) database and are publicly available under accession number GSE225748
All bulk ATAC-seq data have been deposited in the NCBI GEO database and are publicly available under accession number GSE225672.
All scRNA-seq data have been deposited in the NCBI GEO database and are publicly available under accession number GSE225673.
The mm10 genome database (https://www.ncbi.nlm.nih.gov/assembly/GCF_000001635.20/) was used to align sequences for the RNA-seq analysis and to align sequencing reads for the ATAC-seq analysis. Source data are provided with this paper.
All raw data and materials will be made available to investigators upon request.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	NA
Population characteristics	NA

Recruitment	NA
Ethics oversight	NA

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculations were performed; for RNA-seq and ATAC-seq, sample sizes were based on previous knowledge/publications as well as a review of our sample quality. We especially used the insight of computational biologists (Shuchi Smita and Maria Chikina, who are also authors on the paper). For other data, sample size was also determined based on previous knowledge and common practices in immunology.
Data exclusions	For scRNA-seq, a cluster of Igkc-highly expressing cells was assumed to be contamination and removed (genetically, our cells of interest are Jk-/- and therefore cannot express Igkc).
Replication	At least two independent biological replicates were gathered when possible and all replicates were successful.
Randomization	No randomization occurred. Mice were grouped by age and sex and conditions.
Blinding	Data collection was not blinded. Data analysis was done in an unbiased way using specialized software. Biological groups had to be identified so that the appropriate cells were grouped and analyzed together.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Unconjugated CD16/32 (Clone 2.4G2, for Fc block, produced in our lab, 1:100)
Unconjugated CD4 (Clone GK1.5, for depletion, produced in our lab and BioXcell #BE0017-1, 400 ug or 200 ug in 200 ul injected)
Unconjugated CD40L (Clone MR1, for blockade, produced in our lab, 350 ug in 200 ul injected)
Unconjugated CD40 (Clone FGK-45, for agonism, BioXcell #BE0016-2, 25 ug in 200 ul injected)
CD45.2-APC (Clone 104, Biolegend, #109814, 1:200)
CD19-A488 (Clone 1D3, produced and conjugated in our lab, 1:400)
CD93-PE (Clone AA4.1, Biolegend, #136504, 1:200)
CD38-APC/Cy7 (Clone 90, Biolegend, #102728, 1:200)
CD138-BV605 (Clone 281-2, BD, #563147, 1:400)
CD80-BV421 (Clone 16-10A1, BD, #BDB566285), 1:400
PDL2-Biotin (Clone TY25, Biolegend, #107203, 1:100)
CD21/35-PerCPCy5.5 (Clone 7E9, Biolegend, #123416, 1:100)
CD38-PerCPCy5.5 (Clone 90, Biolegend, #102722, 1:400)
CD23-PE/Cy7 (Clone B3B4, Biolegend, #101614, 1:200)
CD45.1-A488 (Clone A20, produced and conjugated in our lab, 1:100)

CD19-BUV395 (Clone 1D3, BD, #563557, 1:100)
Fas(CD95)-PE/Cy7 (Clone Jo2, BD, #557653, 1:200)
PDL2-PE (Clone TY25, Invitrogen, #12-5986-82, 1:200)
CD45.1-PE (Clone A20, Tonbo, #50-0453, 1:200)
B220-BUV395 (Clone RA3-6B2, BD, #563793, 1:100)
GL7-FITC (Clone GL7, Biolegend, #144604, 1:800)
CXCR4-BV421 (Clone L276F12, Biolegend, #146511, 1:100)
CD86-PE (Clone PO3, Biolegend, #105008, 1:100)
CD19-APC/Cy7 (Clone 1D3, BD, #557655, 1:200)
CD19-BV786 (Clone 1D3, BD, #563333, 1:400)
CD45.1-APCeFluor780 (Clone A20, Invitrogen, #47-0453-82, 1:50)
TCRb-PerCPCy5.5 (Clone H57-597, BD, #560657, 1:100)
CD44-A647 (Clone Pgp-1, produced and conjugated in our lab, 1:400)
CD19-A647 (Clone 1D3, produced and conjugated in our lab, 1:800)
MHCII-APC/Cy7 (Clone M5/114, Biolegend, #107628, 1:200)
CD11c-PE/Cy7 (Clone N418, Tonbo, #60-0114, 1:1600)
CD11b-PE (Clone M1/70, Biolegend, #101208, 1:1600)
GL7-Pacblue (Clone GL7, Biolegend, #144614, 1:800)
CD3e-PE (Clone 145-2C11, Biolegend, #100308, 1:100)
CD4-BV605 (Clone RM4-5, BD, #563151, 1:1000)
CD38-A594 (Clone 90, produced and conjugated in our lab, 1:100)
Fas(CD95)-PerCPeFluor710 (Clone 15A7, Invitrogen, #46-0951-82, 1:200)
PDL2-BV480 (Clone TY25, BD, #746756, 1:50)
NIP-APC (Conjuagted in our lab, 1:2000)
NIP-PE (Conjugated in our lab, 1:4000)
SA-BUV395 (BD, #564176, 1:100)
7AAD (Tonbo, #13-6993, 1:50)
Zombie NIR (Biolegend, #423106, 1:500)
Ghost BV510 (Tonbo, #13-0870, 1:250)
CD4-biotin (Clone GK1.5, produced and conjugated in our lab, 1:800)
CD8-biotin (Clone T1B105, produced and conjugated in our lab, 1:400)
CD43-biotin (Clone S7, produced and conjugated in our lab, 1:800)
GR1-biotin (Clone RB6-8C5, Biolegend, #108404, 1:1600)
TER119-biotin (TER119, produced and conjugated in our lab, 1:400)
CD49b-biotin (Clone DX5, Biolegend, #108908, 1:200)
CD11c-biotin (Clone N418, produced and conjugated in our lab, 1:400)
CD11b-biotin (Clone M1/70, produced and conjugated in our lab, 1:400)
Ly6D-FITC (Clone 49-H4, Biolegend #138605, 1:400)
PlexinB2-PE (Clone 3E7, Biolegend #145903, 1:200)
IgG1-BV510 (Clone A85-1, BD optibuild #740121, 1:200)
CD73-BV605 (Clone TY/11.8, Biolegend #127215, 1:50)
B220-PE/Cy7 (Clone RA3-6B2, Tonbo #60-0452, 1:800)

Validation

Commercial antibodies were validated by vendors with data available online. All have been published. Antibodies made by our lab were titered and tested against appropriate controls or vendor equivalents to validate.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	All mice were bred and housed in specific pathogen-free conditions in a room at 20–26°C with 30–70% relative humidity. Mice were kept on a 14 hour light/10 hour dark schedule. Donor mice and recipient mice for the Balb/c transfer system were 6-12 weeks of age at time of transfer and immunization. GCET-Cre experimental mice were also between ages 6-12 weeks at time of immunization.
Wild animals	No wild animals were used.
Reporting on sex	All donors and recipients in the Balb/c transfer system were females. For GCET-Cre experiments, these mice were mostly female, with the exception of ATAC-seq data in which 80% were female and 20% male.
Field-collected samples	No field-collected samples were used.
Ethics oversight	The study was approved by the University of Pittsburgh Institutional Animal Care and Use Committee.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Splenocyte single cell suspensions were generated by crushing spleens between glass microscope slides in 5 ml R10 (RPMI, 10% FCS, 1% ml Pen/Strep, 1% ml L-glutamine, 1% ml HEPES, 0.1% ul 1000x 2-mercaptoethanol). Cells were filtered through 80 uM nylon mesh and washed at 1500 RPM, 4 °C, 5 minutes. Pellets were resuspended in 3 ml of RT ACK lysis buffer/spleen and incubated RT for 2 minutes, followed by addition of 12 ml of R10 and washing at 1500 RPM, 4 °C, 5 minutes. For flow cytometry, cells were resuspended in staining buffer (PBS, 3% FCS, 1 mM EDTA, 0.02% sodium azide). For bead enrichment followed by cell sorting, cells were resuspended in staining buffer without sodium azide. Cells were blocked at a concentration of 100x10^6 cells/ml with anti-CD16/32 (clone 2.4G2), rat serum, and mouse serum (when mouse-derived antibodies were used) on ice for 10 minutes. If mouse-derived antibodies or mouse serum was required, anti-IgM or anti-IgG antibodies were stained first. For flow cytometry, cells were washed and stained with antibody cocktails for 20 minutes on ice, followed by washing and staining in a live/dead discriminator for 15 minutes on ice. For samples that required a streptavidin-conjugated fluorophore, it was also added with the live/dead stain. Cells were washed and fixed in 1% PFA for 30 minutes, washed again, and resuspended in staining buffer for analysis on BD LSR II, BD LSR Fortessa, or Cytek Aurora flow/spectral cytometers. Analysis of cytometry data was performed using FlowJo v10.
Instrument	Cells were collected using a BD LSRII, BD LSRFortessa, or Cytek Aurora.
Software	Data were collected using BD FACSDiva or SpectroFlo softwares. All analysis was performed in FlowJo v10.
Cell population abundance	We identified rare cell populations, and at minimum attempted to acquire ~200 events of these. Most samples had enough cells to collect thousands of rare cell events.
Gating strategy	All gating strategies are shown and described in the main or supplemental figures. We first always gated on lymphocytes vis FSC-A and SSC-, then excluded doublets via FSC-W/FSC-H and SSC-W/SSC-H, and lastly used a live/dead discriminator to gate on live cells. For B cell populations, we almost always used antigen specificity, except in the case of the Balb/c transfer system where >90% of cells are antigen-specific and we required untouched cells. B cells were gated on CD19+ (or B220+ when CD19 was not available), CD138+ cells were excluded, and the resulting cells were separated by CD38+ Fas low (non-GCs, MBCs) and CD38- Fas hi (GCs) popiulations. In some cases, CD4 T cells were gated using CD3e and CD4. Total T cells were gated using TCRb. Myeloid cells were gated using TCRb- CD19-, followed by CD11b, CD11c, and MHC II markers.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.