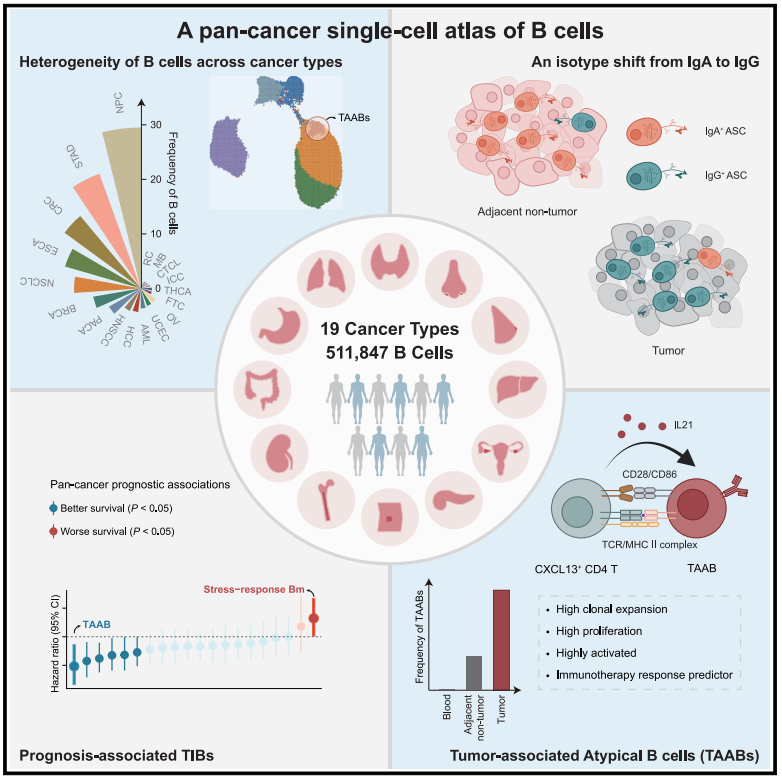
Cell

Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes

Graphical abstract



Highlights

- A pan-cancer single-cell atlas of tumor-infiltrating B cells (TIBs) is constructed
- Abundance and subtype composition of TIBs exhibit heterogeneity across 19 cancer types
- Tumor-associated atypical B cells (TAABs) are associated with favorable prognosis
- TAABs are clonally expanded and activated, appearing to interact with CD4 T cells

Authors

Yu Yang, Xueyan Chen, Jieying Pan, ..., Dongfang Wang, Min-Min Chen, Zemin Zhang

Correspondence

wangdf19@pku.edu.cn (D.W.),
chenmm@szbl.ac.cn (M.-M.C.),
zemin@pku.edu.cn (Z.Z.)

In brief

Analyses of tumor-infiltrating B cells from patients across 19 cancer types depict their abundance and subtype heterogeneity at the pan-cancer level and identify tumor-associated atypical B cells that can communicate with CD4 T cells and have prognostic potential.

Yang et al., 2024, Cell *187*, 1–22
August 22, 2024
https://doi.org/10.1016/j.cell.2024.06.038

CellPress

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

Cell

CellPress
OPEN ACCESS

Resource

Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes

Yu Yang,[1,4] Xueyan Chen,[1,4] Jieying Pan,[2] Huiheng Ning,[2] Yaojun Zhang,[3] Yufei Bo,[1] Xianwen Ren,[1] Jiesheng Li,[1] Shishang Qin,[1] Dongfang Wang,[1,*] Min-Min Chen,[2,*] and Zemin Zhang[1,5,*]

[1]Biomedical Pioneering Innovation Center (BIOPIC), Academy for Advanced Interdisciplinary Studies, and School of Life Sciences, Peking University, Beijing 100871, China

[2]Institute of Cancer Research, Shenzhen Bay Laboratory, Shenzhen 518132, China

[3]State Key Laboratory of Oncology in South China, Department of Liver Surgery, Sun Yat-sen University Cancer Center, Guangzhou 510060, China

[4]These authors contributed equally

[5]Lead contact

*Correspondence: wangdf19@pku.edu.cn (D.W.), chenmm@szbl.ac.cn (M.-M.C.), zemin@pku.edu.cn (Z.Z.)

https://doi.org/10.1016/j.cell.2024.06.038

SUMMARY

Characterizing the compositional and phenotypic characteristics of tumor-infiltrating B cells (TIBs) is important for advancing our understanding of their role in cancer development. Here, we establish a comprehensive resource of human B cells by integrating single-cell RNA sequencing data of B cells from 649 patients across 19 major cancer types. We demonstrate substantial heterogeneity in their total abundance and subtype composition and observe immunoglobulin G (IgG)-skewness of antibody-secreting cell isotypes. Moreover, we identify stress-response memory B cells and tumor-associated atypical B cells (TAABs), two tumor-enriched subpopulations with prognostic potential, shared in a pan-cancer manner. In particular, TAABs, characterized by a high clonal expansion level and proliferative capacity as well as by close interactions with activated CD4 T cells in tumors, are predictive of immunotherapy response. Our integrative resource depicts distinct clinically relevant TIB subsets, laying a foundation for further exploration of functional commonality and diversity of B cells in cancer.

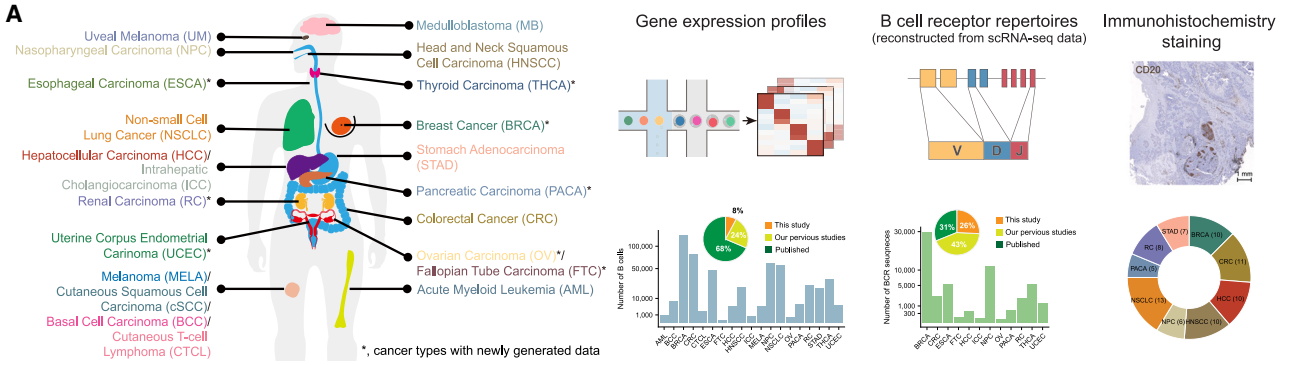
INTRODUCTION

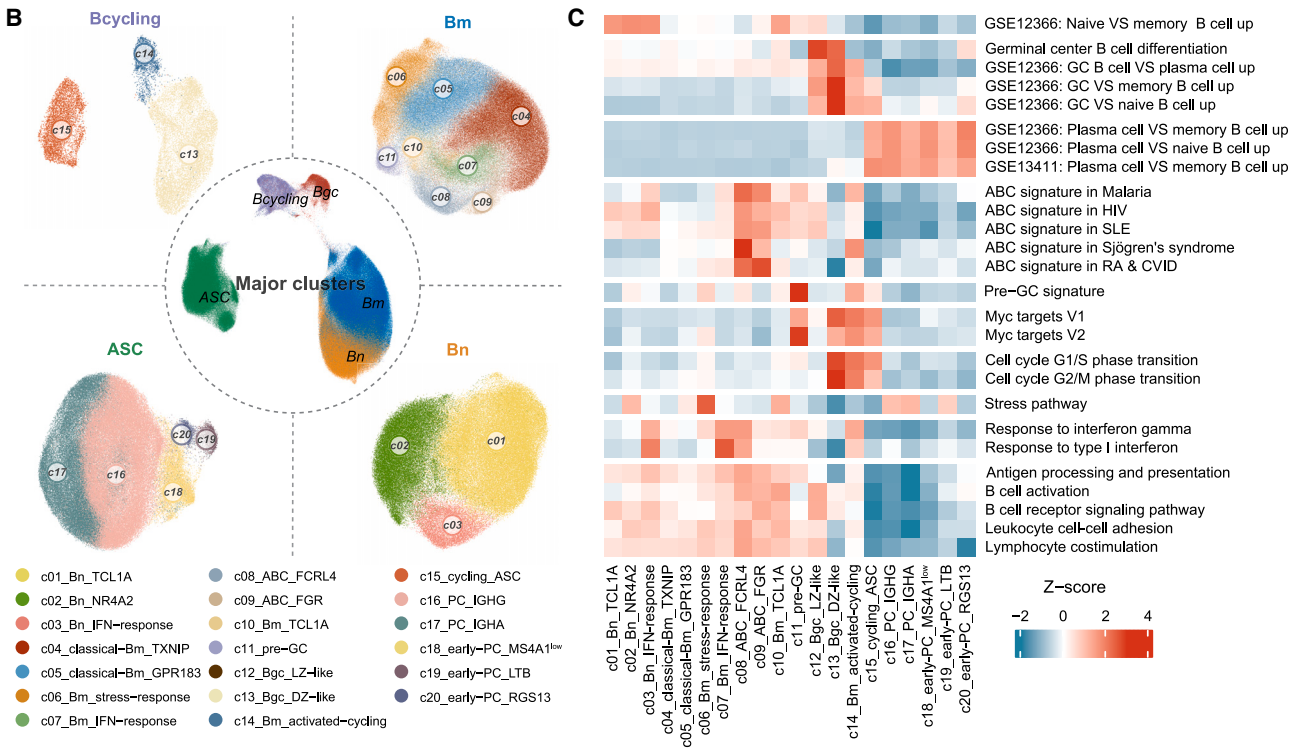
The success of cancer immunotherapy has inspired the in-depth exploration of the holistic tumor ecosystem, especially the immune aspect of the tumor microenvironment (TME).[1] The comprehensive chart of the functional states and heterogeneity of tumor-infiltrating T and myeloid cells[2,3] has enlightened multiple novel therapeutic strategies. Despite these, only some patients achieve durable response, and the gap in response rates among different cancer types is substantial.[4] A complete understanding of the TME at the pan-cancer level is warranted to fully unleash the potential of immunosurveillance in human cancers. B cells, as a central component of the immune system, exert important effects in immunity,[5,6] but they have received disproportionately less attention, with the compositional and functional heterogeneity of tumor-infiltrating B cells (TIBs) not systematically examined.

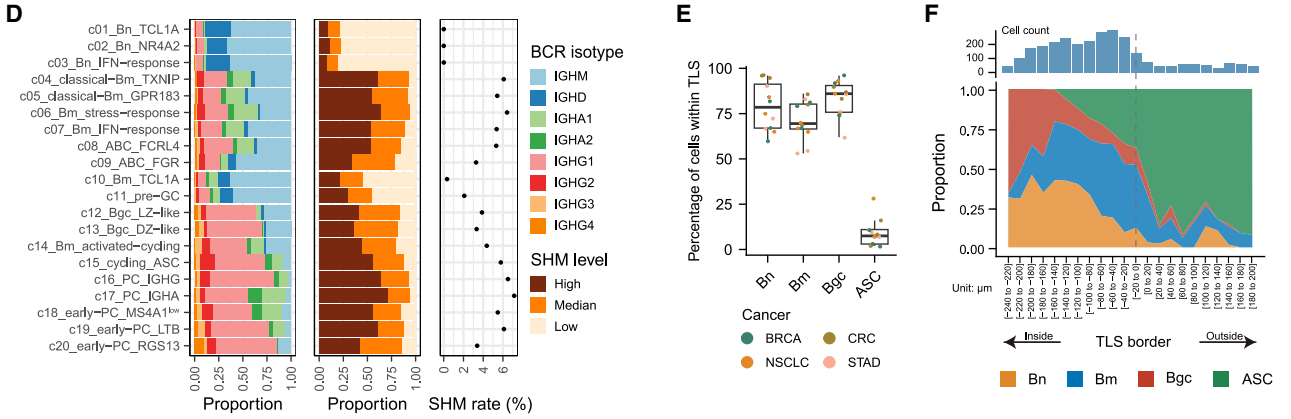
Emerging studies have begun to decipher the functional roles of B cells in anti-tumor immunity. In ovarian carcinoma (OV) and renal carcinoma (RC), the terminally differentiated B cells, plasma cells (PCs), have been reported to produce anti-tumor antibodies, which may tag cancer cells for subsequent elimination.[7,8] In addition, TIBs can closely interact with other immune cells in tumors.[9] Such interactions are supported by observations that aggregated B cells in tumors typically mark the spatial enrichment of various immune cell types and the formation of tertiary lymphoid structures (TLSs),[10] which function as regional hubs for B cell maturation and contribute to enhanced anti-tumor immune response.[11] Furthermore, growing evidence suggests B cells as a prognostic marker in various cancer types, but with notable cancer type heterogeneity.[12–14] The clinical significance of B cells has been further underscored by their associations with immunotherapy response.[15] Particularly, the presence of TLSs has been identified as a robust predictive biomarker of response to cancer immunotherapy.[16–18] Although these findings underscore the important role of B cells in anti-tumor immunity, their pan-cancer-shared and cancer-type-specific characteristics, and especially the potential B cell subsets underlying these processes, remain less understood.

TIBs comprise several major lineages including naive B (Bn) cells, memory B (Bm) cells, germinal center B (Bgc) cells, and antibody-secreting cells (ASCs).[10,19] The single-cell RNA sequencing (scRNA-seq) technologies have enabled the fine-grained characterization of TIBs, and the spectrum of TIBs is starting to emerge.[20,21] For example, an FCRL4^{+} Bm subset has been identified to be associated with improved patient

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038







(legend on next page)

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

survival in nasopharyngeal carcinoma (NPC),[22,23] and linked with immunotherapy response in non-small cell lung cancer (NSCLC).[24] However, it is unclear whether such findings from certain cancer types are generalizable. Further, the inadequate number of sequenced B cells per dataset may impede the detection of rare subsets. Akin to the construction of pan-cancer T and myeloid cell atlases,[2,3] the large-scale integration of pan-cancer TIB datasets is expected to illuminate perspectives to address the aforementioned challenges.

In this study, we collected a broad scope of published scRNA-seq datasets of the TME and integrated them with our newly generated scRNA-seq data to construct a comprehensive human B cell atlas. We additionally reconstructed the B cell receptor (BCR) repertoire *in silico* and performed immunohistochemistry (IHC) staining to illustrate the spatial distribution patterns of TIBs. Such a resource enabled the systematic inspection of the transcriptional, clonotypic, and spatial heterogeneity of TIBs across cancer types. We also uncovered prognosis-associated TIB subsets and, specifically, defined TIB cells that can closely communicate with other immune components to engage in tumor-associated immune responses.

RESULTS

Construction of a pan-cancer single-cell atlas of B cells

Our newly generated pan-cancer scRNA-seq dataset was assembled with 54 additional published studies to elucidate the pan-cancer characteristics of TIBs (Table S1). We conducted rigorous quality control and filtering and then distinguished B cells from other cell types through unsupervised clustering in each dataset separately, with clusters highly expressing canonical B cell marker genes retained. A high-quality single-cell transcriptome atlas was then compiled, containing 511,847 B cells derived from 948 samples of 649 patients across 19 major cancer types, primarily covering tumor tissues, adjacent non-tumor tissues (ANTs), and peripheral blood (Figures 1A and S1A; Table S1; STAR Methods). In addition, for those 59,592 cells with available raw sequencing data (Figure 1A), we computationally reconstructed their BCR sequences, leveraging a repertoire sequence calling tool, TRUST4.[25] Furthermore, to study the spatial localization of TIBs, we performed IHC of multiple TIB-related markers on 80 paraffin-embedded tissue specimens covering nine cancer types with six different panels (Figures 1A and S1B; Table S2).

All collected scRNA-seq datasets were integrated with batch effects removed (Figure S1C) through Harmony,[26] and systematic unsupervised clustering was then performed to characterize the transcriptional subsets of B cells. As expected, five major clusters, including four B cell major lineages—Bn cells, Bm cells, Bgc cells, and ASCs—as well as one cluster of cycling B cells (Figure 1B), were discriminated based on the high expression of their canonical markers—*IGHD/TCL1A, CD27/TNFRSF13B, BCL6/AICDA*, *MZB1*, and *MKI67*, respectively (Figure S1D). Independent analyses of single cancer types generated consistent clustering results (Figure S1E). We then separately clustered the aforementioned five major clusters, discovering a total of 20 distinct fine-grained B cell subtypes, each with specific signature genes (Figures 1B and S1F; Table S3; STAR Methods). In particular, cycling B cells were divided into three subsets, which differed in their expression of major lineage-specific markers and were thereby reassigned to Bgc, Bm, and ASC lineages, respectively.

We next employed gene set-based analyses to further elucidate the B cell subset molecular phenotyping (Figure 1C; Table S4; STAR Methods). For Bm cells, we identified eight subsets, among which c08_ABC_FCRL4 and c09_ABC_FGR resembled the transcriptional phenotypes of atypical B cells (ABCs). ABCs have been reported in studies of infections, autoimmune diseases, and vaccination,[27–29] and both c08 and c09 showed a high expression of ABC signature genes from multiple studies. Specifically, they harbored high expression of *FCRL5*, *ITGAX*, and *TBX21*, as well as low expression of *CR2* (Figure S1F); however, compared with c09, c08 showed a unique up-regulation of *FCRL4*, and around 10% of c08 cells expressed *PDCD1* (Figure S1F). In addition, we observed a rare population of c11, highly expressing the previously reported pre-germinal center (pre-GC) B cell signature.[30] These cells also displayed elevated Myc pathway activity (Figures 1C and S1F), required for germinal center initiation.[31] Further considering that genes mechanistically linked with class switch recombination (CSR) were highly expressed in c11 cells (Figure S1G), we defined c11 as pre-GC cells. Furthermore, we identified c10_Bm_TCL1A as a transitional state between Bn and Bm cells, as the markers of these two major lineages were highly expressed in this subset simultaneously. Two Bm subsets, c06_Bm_stress-response and c07_Bm_IFN-response, exhibited molecular phenotypes influenced by environmental cues, with c06 featuring a high expression of stress-related heat shock genes including *HSPA1A, HSPA1B*, and *DNAJB1*, and c07 showing elevated expression of IFN-stimulated genes (Figures 1C and S1F). All of the ASC subsets (c15-c20) demonstrated heightened expression of known PC signatures and ASC differentiation-related transcription factors,[32] such as *PRDM1*, *IRF4*, and *XBP1* (Figure S1F). As expected, a substantial proportion of immunoglobulin (Ig) gene counts was detected within these cells (Figure S1H). Among these subsets, c15, c18, c19, and c20 maintained the expression of major histocompatibility complex (MHC)-II genes (Figure S1F), indicative of a transitional state before becoming

Figure 1. Pan-cancer single-cell profiling of B cells

(A) Schematics overview of our atlas.

(B) Uniform manifold approximation and projection (UMAP) plots showing the major clusters and subsets of B cells.

(C) Heatmap showing expression of gene signatures in B cell subsets.

(D) Distribution of IgH isotype (left) and SHM levels (middle) across TIB subsets, with the median SHM rate shown for each subset (right).

(E) The percentage of cells within TLSs across B cell major lineages quantified based on mIHC data.

(F) Spatial distribution patterns of B cell major lineages inside or outside of a representative TLS (as shown in Figure S2D).

See also Figures S1 and S2 and Tables S1, S2, S3, and S4.

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

fully matured PCs.[33] Especially, *MS4A1* was still moderately expressed in c19_early-PC_LTB and c20_early-PC_RGS13 cells (Figure S1F), suggesting an early stage of PC differentiation.[34] The c15_cycling_ASC was characterized by the specific expression of cell cycle-related genes, such as *MKI67*, *STMN1*, and *TOP2A* (Figure S1F). The two mature PC clusters, c16_PC_IGHG and c17_PC_IGHA, were distinguished with the aid of BCR evidence (STAR Methods), corresponding to the IgG and IgA isotypes, respectively. In brief, we established a comprehensive transcriptome atlas of B cells, providing a detailed profile of their fine-grained subsets.

The reconstructed BCR sequences allowed us to examine various clonotypic characteristics of TIBs (STAR Methods). To exemplify the utility of the BCR repertoire, we identified the Ig heavy-chain isotypes of TIBs based on their Ig constant regions, finding diverse isotype distribution patterns among B cell major lineages (Figure 1D). Bn clusters were predominated by IgM and IgD isotypes, and ASC clusters showed a high enrichment of IgA and IgG isotypes, whereas Bm and Bgc clusters exhibited a mixture of various isotypes. Then, the somatic hypermutation (SHM) level of each cell quantified as the point mutation rate of the Ig variable region was stratified into three groups: low SHM (mutation rate < 1%), median SHM (1%–5%), and high SHM (> 5%). We observed a low level of SHM in the BCRs of Bn cells and a medium or high level of SHM in the BCRs of Bm cells, Bgc cells, and ASCs (Figure 1D), corresponding to their progressive maturation stages, a pattern consistently observed across different cancer types (Figure S2B). We further stratified the Bgc cells, Bm cells, and ASCs into two groups according to their respective median SHM levels and found that the B cell maturation marker *CD27* exhibited higher expression levels in the SHM-high groups, while *IGHM* and *IGHD* were expressed higher in the SHM-low groups (Figure S2C).

To map the spatial layout of B cell major lineages in tumors, we performed multiplex IHC (mIHC) with four panels on twelve tumor tissues (Figure S1B). We first defined TLSs as organized structures with clear B cell and T cell zones (STAR Methods).[11,18,19] Through quantification (STAR Methods), we illustrated that Bn cells ($CD20^+IgD^+$), Bm cells ($CD20^+CD27^+$ $Bcl\text{-}6^-$), and Bgc cells ($CD20^+Bcl\text{-}6^+$) were mainly located in TLSs, with Bgc cells primarily occupying the central region of TLSs (Figures 1E, 1F, and S2D), consistent with the previous findings that Bm cells and Bgc cells are enriched in TLSs in renal cell cancer.[8] By contrast, ASCs ($CD79a^+CD27^+CD138^+$) were detected in the perimeter of TLSs and were found to be the most abundant B cell major lineage in the vicinity of cancer cells. IHC staining on tumors covering nine cancer types confirmed the aggregated distribution of $CD20^+$ B cells and the presence of $CD138^+$ PCs in close proximity to TLSs, within the tumor, as well as among the fibrotic stroma (Table S2; STAR Methods). Thus, our data constituted a useful resource to investigate the transcriptional phenotypes, clonotypic states, and spatial distributions of human B cells at the pan-cancer level. We also developed a web-based tool (http://pan-b.cancer-pku.cn/) to facilitate further investigation of our data.

Heterogeneity of B cells across cancer types

The pan-cancer atlas enabled us to systematically examine the heterogeneity of TIBs across cancer types. In tumors, the proportions of B cells within the immune compartment varied dramatically across cancer types (Figure 2A), ranging from close to 0% in uveal melanoma (UM) to 29.5% in NPC. Intriguingly, for certain cancer types harboring abundant B cells in tumors, such as colorectal cancer (CRC), stomach adenocarcinoma (STAD), and esophageal carcinoma (ESCA), high B cell prevalence rates were also detected in their ANTs (Figure S3A). In addition, within these cancer types originating from mucosal tissues, positive correlations of B cell population sizes between the ANTs and tumors were observed (Figure 2B), reflecting that the presence of inflammation states or TLS in these tissues may contribute to the observed higher levels of B cells. By contrast, in breast cancer (BRCA), NSCLC, and pancreatic carcinoma (PACA) that were also characterized by a large pool of B cells, we identified elevated B cell abundances in tumors compared with ANTs (Figure S3B), aligning with a previous report based on TCGA bulk RNA-seq data.[13] Such changes were also observed by IHC staining of the ANT and tumor specimens (Figure S3C). Given the variability of TIB levels even among patients with the same cancer type (Figure 2A), we hypothesized that the TME-specific intercellular communication could also influence the B cell infiltration. Indeed, compared with tumor-infiltrating T and myeloid cells, TIBs rarely existed on their own at the pan-cancer level (Figure 2C; STAR Methods), expanding the previous observations in certain cancer types.[16,35,36] Furthermore, the CellChat-based[37] analysis revealed that TIBs were predicted to be attracted by rather than to recruit other cell types (Figure S3D; STAR Methods). Among the chemokine pathways, the CXCL12–CXCR4 axis, which has been reported to mediate the localization of B cells[38] and the dissemination of PCs in tumors,[8] was recurrently observed across various cancer types (Figure S3E). Thus, the organ contexture and TME-associated factors could both affect the cancer-type-dependent B cell infiltration levels.

We then examined the compositional heterogeneity of TIBs in terms of major lineage proportions (Figure 2D). Generally, Bgc cells only accounted for less than 5% of TIBs, representing a relatively small population at the tumor site. Bn cells exhibited stable (~10%–15%) prevalence across cancer types, whereas higher variabilities of the Bm and ASC proportions were observed. Especially, the major lineage compositions in tumors might be influenced by the inherent characteristics of host organs. For example, CRC tumors and their ANTs featured a high abundance of Bgc cells and ASCs (Figures 2D and S3F). In addition, B cell subset-level heterogeneity was depicted through cancer type stratification based on subset proportions (Figure 2E). Of note, CRC and STAD exhibited similar TIB compositions and were clustered together, indicative of common factors in shaping their B cell-associated anti-tumor immunity. Overall, our integrative datasets provided insights into the complexity and heterogeneity of TIBs at the pan-cancer level.

Recapitulation of germinal center reactions within human tumors

Our atlas also provided opportunities to analyze rare cell types. Taking Bgc cells as an example, we observed their transcriptional characteristics and transition dynamics in the TME. Conventionally, B cell affinity maturation in secondary lymphoid organs (SLOs) primarily occurs in germinal centers, where Bgc cells cycle between the light zone and dark zone and iteratively

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

CellPress
OPEN ACCESS

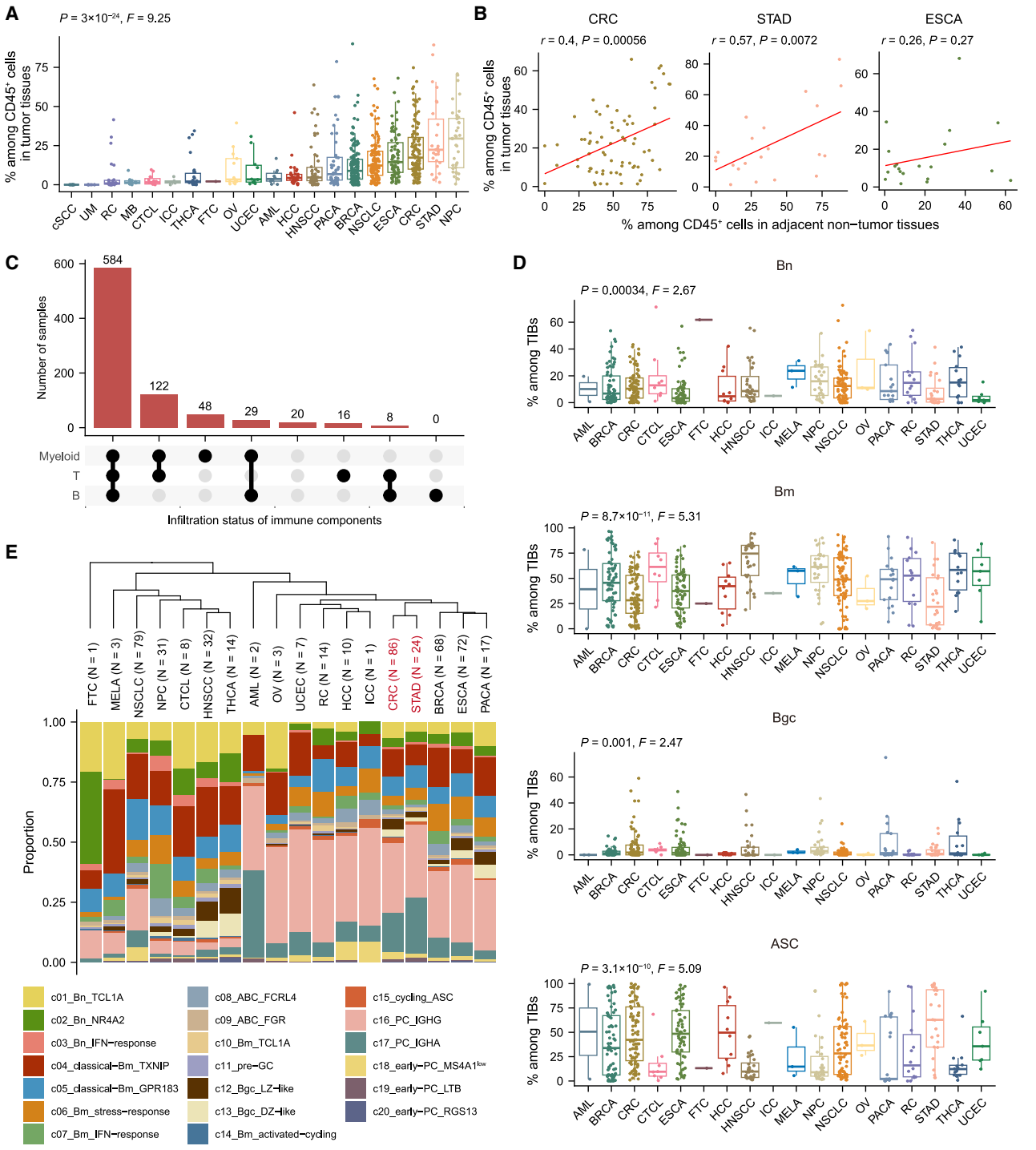


Figure 2. Heterogeneity of B cells across cancer types

(A) Proportions of TIBs in CD45⁺ cells across cancer types. One-way ANOVA test.

(B) Pearson correlation of B cell proportions between tumors and ANTs in CRC, STAD, and ESCA.

(C) UpSet plot showing the infiltration status of three major immune components within tumors.

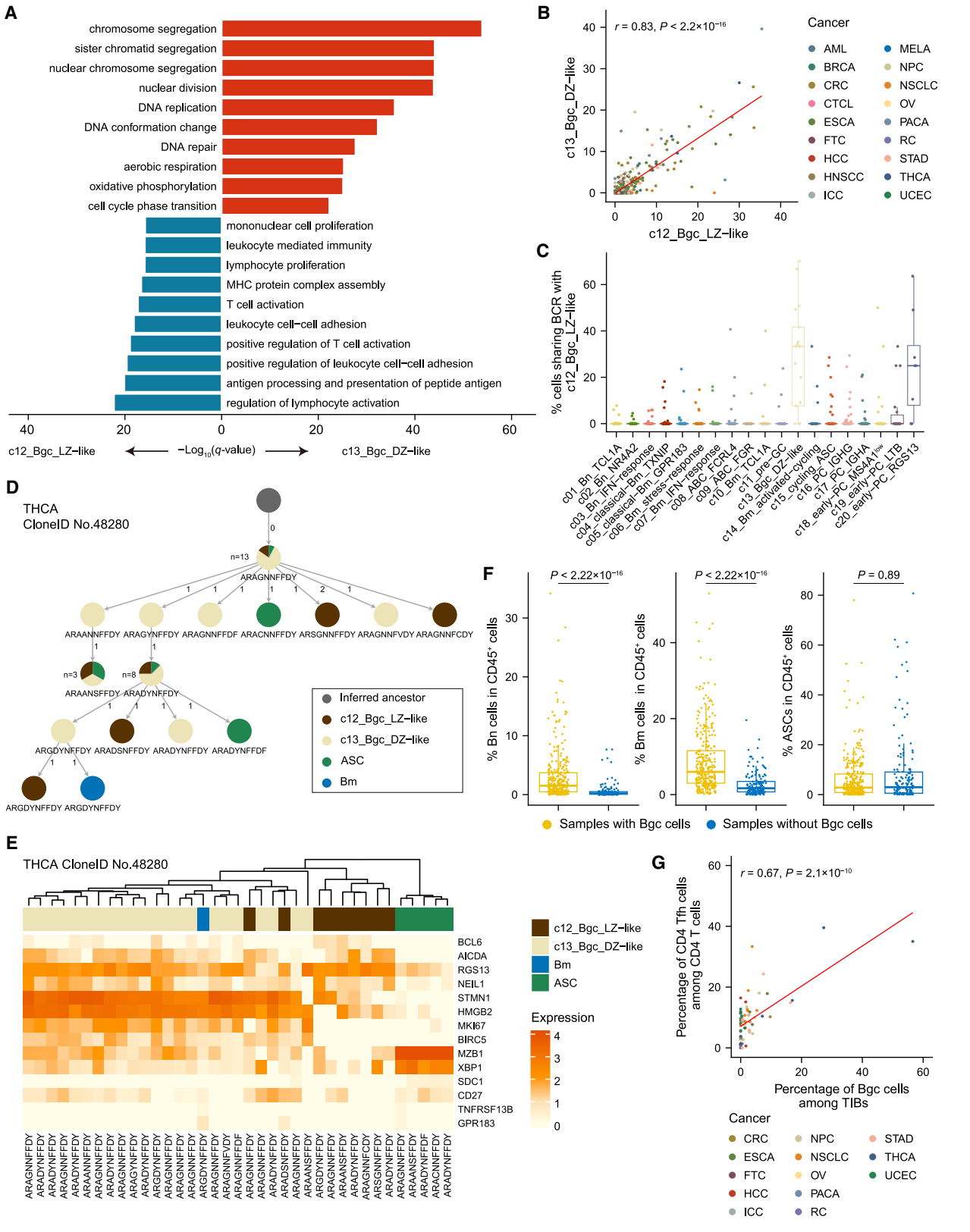
(D) TIB major lineage compositions across cancer types. One-way ANOVA test.

(E) Cancer type stratification by hierarchical clustering of B cell subset proportions, with sample numbers annotated.

Only samples with CD45⁺ cells > 100 and an immune compartment unaffected by FACS are shown in (A)–(C). Only samples with B cells > 50 are shown in (D) and (E). All samples are from treatment-naive patients.

See also Figure S3.

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038



(legend on next page)

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

accumulate BCR mutations, eventually giving rise to antigen-specific Bm cells and PCs.[19] In tumors, we identified two Bgc subsets, c12_Bgc_LZ-like and c13_Bgc_DZ-like, and confirmed their presence in almost all analyzed cancer types (Figures S4A and S4B). Those c12 cells expressed genes involved in pathways of regulation of T cell activation, antigen processing, and presentation, whereas cell proliferation-associated pathways were upregulated in c13 cells (Figure 3A). Of note, in tumors, the abundances of c12 and c13 cells were correlated, indicative of their frequent co-occurrence (Figure 3B). Accordingly, large fractions of cells from c12 and c13 shared BCR clones (Figures 3C and S4C), indicating state transitions between them. To further elucidate the evolutionary dynamics of Bgc cells, for each TIB clone containing Bgc cells, we computationally reconstructed the lineage tree based on the third complementarity determining region (CDR3) sequences of corresponding cells (STAR Methods). For several representative lineage trees, multiple rounds of alternate transition between c12 and c13 cells manifested with SHMs gradually accumulated, and Bm cells and PCs emerged accompanying such process (Figures 3D, 3E, S4D, and S4E). Indeed, certain Bm and ASC subsets shared clones with c12 and c13, with notable BCR sharing recognized in c20_early-PC_RGS13 (Figures 3C and S4C), which might be newly differentiated and primed to exit the germinal center.

We next explored the associations of Bgc cells, both with other B cell subsets and with other immune subtypes within tumors. The abundances of Bn and Bm cells appear to be significantly higher in tumors with Bgc cells than those without, whereas no significant difference was observed in the prevalence of PCs (Figure 3F). Among other immune subtypes, CD4 T follicular helper (Tfh) cells exhibited the strongest positive correlation with Bgc cells (Figures 3G and S4F). As CD4 Tfh cells can assist the clonal selection of B cells during the germinal center reactions in SLO,[39] we hypothesized that germinal center reactions in tumors might also engage CD4 Tfh cells to facilitate TIB affinity maturation. Collectively, our data supported that, similar to SLOs, T cell-dependent germinal center reactions also took place within tumors, in which those two Bgc subsets underwent cycles of transition and BCR sequence evolution to produce Bm cells and ASCs.

Identifying potential B cell subsets associated with immune responses in tumors

Using the cross-tissue transcriptomics and BCR data, we next focused on pinpointing critical B cell subsets potentially involved in the tumor-associated immune response from three perspectives, including the degree of tumor enrichment, clonal expansion, and proliferation. Analysis of B cell subset distribution among different tissues revealed dramatic differences between circulating and tissue-infiltrating B cells. In the blood, ASCs featured much lower abundances compared with tissues (Figure S5A), while a few B subsets were still predominantly enriched, including c01_Bn_TCL1A, c04_classical-Bm_TXNIP, c09_ABC_FGR, and c18_early-PC_MS4A1^{low} (Figures 4A and S5B; STAR Methods). In fact, the blood showed a lower B cell compositional diversity than tumors or ANTs, with B cells in tumors featuring the highest diversity (Figure S5C). In addition, c17_PC_IGHA cells were enriched in ANTs, while c06_Bm_stress-response, c07_Bm_IFN-response, c08_ABC_FCRL4, and c16_PC_IGHG displayed a strong distribution preference in tumors (Figure 4A).

Antigen-driven clonal expansion and proliferation are typical characteristics of B cell response.[40] For the clonal expansion level, analyses of our assembled BCR repertoire revealed that the tumor tissues harbored a much higher proportion of B cell clones with clone size > 1, suggesting the occurrence of tumor-related B cell immune response (Figure S5D). As anticipated, the Bgc and ASC lineages exhibited higher levels of clonal expansion in tumors than Bn and Bm cells (Figure 4B), consistent with their status as the expansion phase of a clonotype. Within the ASC lineage, we additionally compared the clonal expansion capacity of different isotypes. Notably, whereas ASCs with IgA and IgG isotypes showed similar clonal expansion levels in ANTs, IgG ASCs from tumors were characterized with greater clonal expansion than those from ANTs (Figure 4C). Relatively limited Bm cells were clonally expanded (Figure 4B), but the c08_ABC_FCRL4 subset still displayed markedly higher clonal expansion levels within tumors (Figures 4D and S5E). Additionally, such tumor-derived c08_ABC_FCRL4 cells were much more expanded than those in ANTs (Figure 4E), implying that intratumoral stimuli might be associated with their clonal expansion.

Focusing on the transcriptional characteristics of cycling B cells, we observed that cycling Bm cells (c14) highly expressed multiple marker genes of c08_ABC_FCRL4 cells (Figure S1F). Further leveraging CellTypist (STAR Methods),[41] we found that among all cycling Bm cells, approximately 52.6% were predicted to be c08_ABC_FCRL4 (Figure 4F), an exceptionally high level given the low proportion of c08 cells in all Bm cells. Additionally, the proportion of cycling cells within FCRL4$^+$ Bm cells was notably elevated in tumor tissues compared with ANTs (Figure 4G). The mIHC staining corroborated the existence of Ki67$^+$FCRL4$^+$ Bm cells in human tumors from multiple cancer

Figure 3. Recapitulation of germinal center reactions within human tumors

(A) Top pathways enriched in intratumoral c12 and c13 Bgc cells compared with each other.

(B) Pearson correlation between the abundance of c12 and c13 Bgc cells in treatment-naive tumor samples with B cells > 50.

(C) Proportions of cells sharing BCR with intratumoral c12_Bgc_LZ-like cells for each B cell subset.

(D) Representative lineage tree of TIBs from a THCA tumor, with cell number and CDR3 amino acid sequence annotated for each node and point mutation count annotated for each edge. The cell type composition of each heterogeneous node is shown as a pie chart.

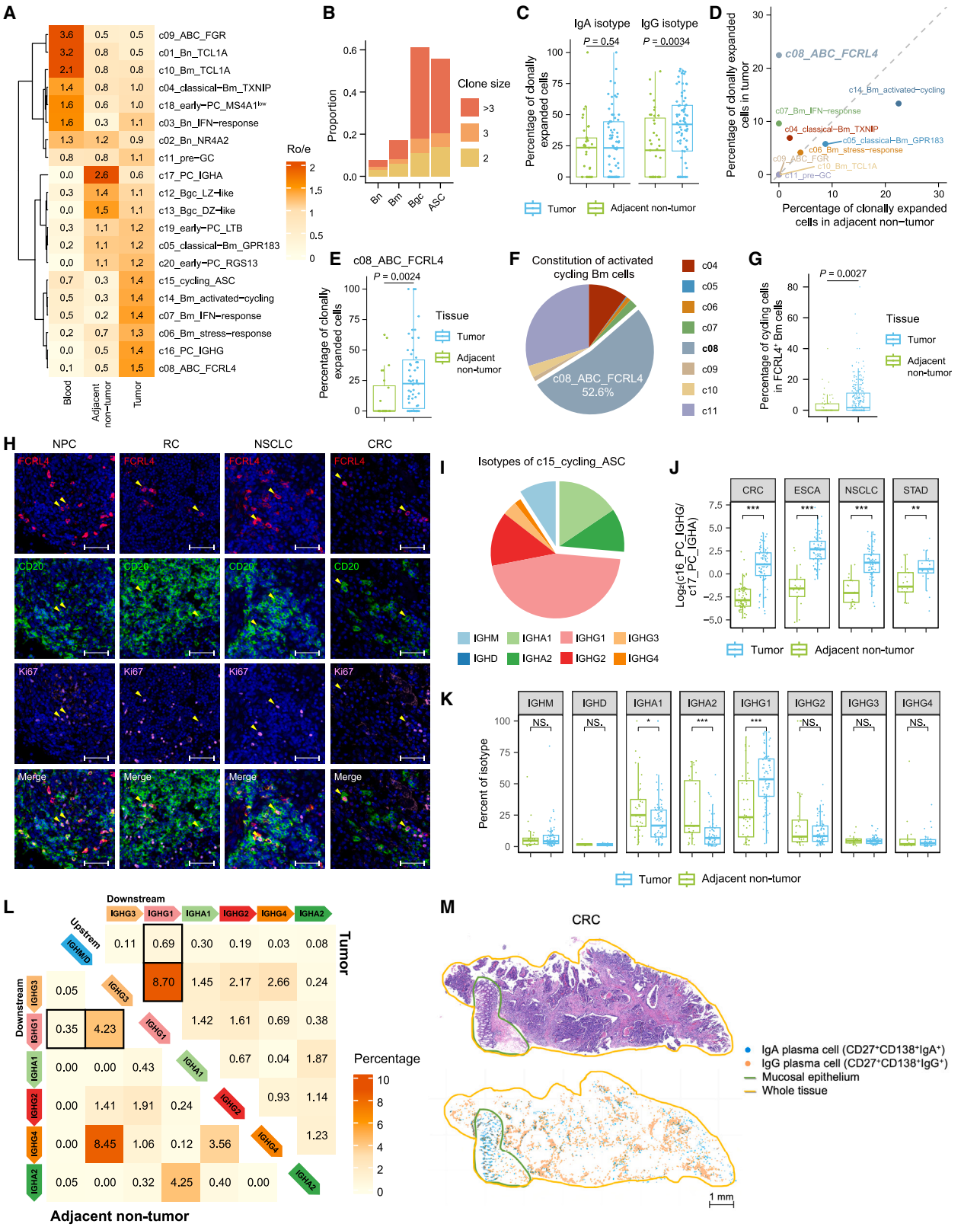
(E) Heatmap showing B cell major lineage marker expression among B cells from the lineage tree in (D).

(F) Boxplots comparing the abundances of Bn cells, Bm cells, and ASCs between tumors with and without Bgc cells. Only treatment-naive tumor samples with B cells > 50 are shown. Two-sided unpaired Wilcoxon test.

(G) Pearson correlation between the abundance of Bgc cells and CD4 Tfh cells in tumors.

See also Figure S4.

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038



(legend on next page)

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

types (Figure 4H). These observations demonstrated the proliferative capacity of FCRL4$^+$ Bm cells, indicating their sustained presence within the TME. Furthermore, the isotypes of cycling ASCs were mainly from the IgG class (Figure 4I). Briefly, these analyses highlighted certain B cell subsets, including the FCRL4$^+$ Bm cells and IgG PCs, might actively engage in tumor-associated immune responses.

An isotype shift from IgA to IgG in intratumoral ASC response

We then investigated the characteristics of the ASC lineage with a focus on the aforementioned tumor-enriched IgG PCs. ASCs tended to harbor the highest variability across cancer types, which could be mainly attributed to their cancer-type-dependent preferential expression of Ig genes (Figures S5F and S5G; STAR Methods). Although c16_PC_IGHG exhibited a stronger enrichment in tumor tissues at the pan-cancer level compared with c17_PC_IGHA (Figures 4A and S5H), varied enrichment degrees existed among cancer types (Figure S5I). In CRC, STAD, ESCA, and NSCLC, which mainly originate from gastrointestinal and respiratory tissues where the normal mucosal immunity is actively maintained by abundant IgA antibodies,[42] the greatest heterogeneity of the abundance of c16 to c17 PCs between the tumors and ANTs was observed (Figure 4J). The tissue preference of IgG and IgA PCs might be in part explained by their chemokine receptor expression patterns. For example, compared with c16_PC_IGHG, c17_PC_IGHA expressed a higher level of *CCR10* (Figure S5J), which is involved in the recruitment of IgA PCs to healthy mucosal tissues.[43]

We further compared the isotype composition of ASCs between tumors and ANTs. Importantly, IgG1 was the most frequent isotype in ASCs from tumor samples and experienced a striking increase from the ANTs, but minimal alterations were identified for other IgG isotypes (Figure 4K). Consistently, we observed an elevated percentage of upstream antibody isotypes (IgM/D and IgG3) that shared clonotypes with the IgG1 isotype in tumors (Figure 4L; STAR Methods). Thus, Ig class switching to IgG1 isotype might occur more frequently in tumors. Additionally, the prevalence of both IgA1 and IgA2 isotypes was substantially decreased in tumors (Figure 4K). Such changes were validated by IHC staining of paired CRC tumor and ANT specimens (Figure S5K). Furthermore, within the CRC tumors, IgA$^+$ cells were primarily aggregated in mucosal regions near the tumor region, while IgG$^+$ cells were predominantly found within the tumor region (Figure 4M). Finally, we probed the interaction coordinated by the fragment crystallizable (Fc) regions of their antibodies from these IgG and IgA PCs, with cues provided by the expression of Fc receptors in TME immune subtypes (Figure S5L). Compared with Fc alpha receptors (FcαRs) that selectively bind IgA, the Fc gamma receptors (FcγRs), specific for IgG, were highly expressed in cytotoxic natural killer (NK) cells and multiple subsets of macrophages. Collectively, the isotypes of ASCs skewed from IgA to IgG isotype in tumors, and we speculated that IgG PCs might be more likely to engage in antibody-associated responses.

FCRL4$^+$ and stress-response Bm cells were linked to patient prognosis

To gain further insights into the role of B cell subsets in tumors, we leveraged the TCGA data to probe the association of B cell subset signature with the overall survival of cancer patients after correcting the effect of total B cell abundance (Figure 5A; STAR Methods). Notably, B cell subsets demonstrated cancer type-dependent correlations with patient prognosis. Patients with a relatively higher expression of the FCRL4$^+$ Bm signature tended to display prolonged survival at the pan-cancer level. However, the cancer type heterogeneity should be noted, as in certain cancers, such as head and neck squamous cell carcinoma (HNSC) and colon adenocarcinoma (COAD), the FCRL4$^+$ Bm signature trended with a slightly worse prognosis (Figure 5B).

By contrast, another B cell subset, c06_Bm_stress-response, was linked to poor survival at the pan-cancer level (Figure 5A). Specifically, its signature was predictive of unfavorable outcomes in ESCA, lung squamous cell carcinoma (LUSC), and pancreatic adenocarcinoma (PAAD) (Figure 5C). The molecular phenotype of c06, highly expressing stress-response-associated genes, resembled the previously reported tumor-associated dysfunctional NK and stress-response T cells.[44,45] Thus, stress response may represent a common molecular characteristic across different cell lineages in tumors, which was linked with poor clinical outcomes. However, to disentangle the effect of specific

Figure 4. Identification of potential B cell subsets associated with immune responses in tumors

(A) Tissue preference of each B cell subset evaluated by the Ro/e index.
(B) Clonal expansion levels of B cell major lineages in tumors, with cells categorized by the clone size of their corresponding clones.
(C) Boxplot comparing the percentage of clonally expanded cells (defined as cells from clonotypes with clone size > 1) in IgA and IgG isotype ASCs between tumors and ANTs.
(D) Scatter plot showing the median percentage of clonally expanded cells in each Bm subset from tumors and ANTs.
(E) Boxplot comparing the percentage of clonally expanded cells in c08_ABC_FCRL4 cells between tumors and ANTs.
(F) Distribution of CellTypist-assigned subset labels in activated cycling Bm cells.
(G) Boxplot comparing the proportions of cycling cells in all FCRL4$^+$ Bm cells between tumors and ANTs.
(H) Representative examples of tumor samples stained by mIHC showing the existence of Ki67$^+$FCRL4$^+$ B cells (arrows). Scale bars, 50 μm.
(I) Distribution of IgH isotypes among c15_cycling_ASC.
(J) Boxplots comparing the abundance ratio of c16 to c17 PCs between tumors and ANTs.
(K) Boxplots comparing the composition of ASC IgH isotypes between tumors and ANTs.
(L) BCR repertoire overlaps across B cells with different IgH isotypes in tumors (top right) and ANTs (bottom left). Heatmap showing the proportions of BCR clonotypes belonging to an upstream isotype (each row) that are shared with a downstream isotype (each column).
(M) Hematoxylin-eosin staining of a CRC tumor sample (top), with the spatial distribution of IgA and IgG PCs shown (bottom).
In (C), (E), (G), (J), and (K), two-sided unpaired Wilcoxon test was used. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, NS. $p \geq 0.05$.
Only treatment-naive tumor samples with B cells > 50 are shown in (G) and (J). See also Figure S5 and Table S2.

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

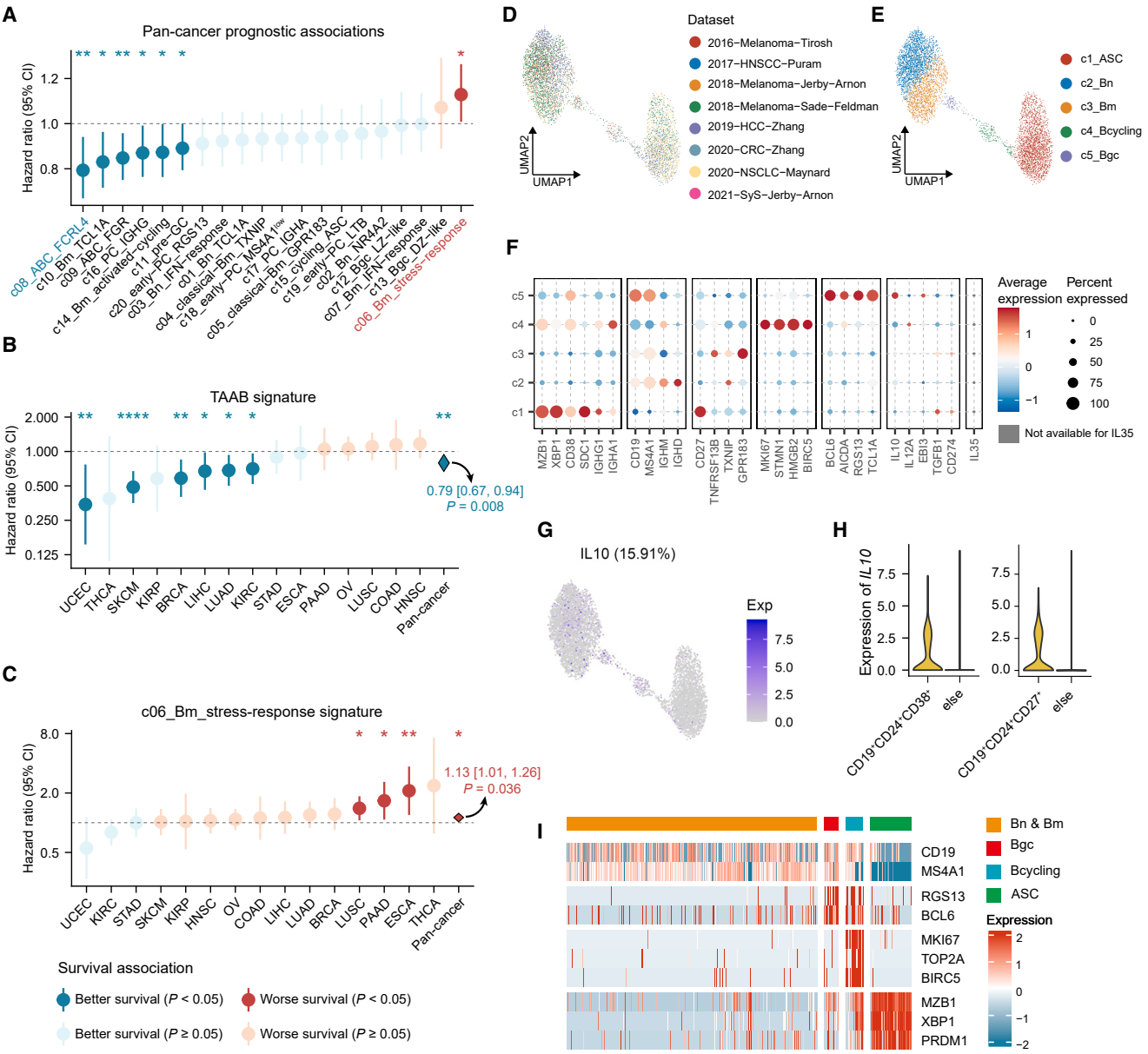


Figure 5. FCRL4$^+$ and stress-response memory B cells were linked to patient prognosis

(A–C) Prognostic associations of TIB subset signatures, including pan-cancer associations of each subset (A), per-cancer associations of TAABs (B), and c06_Bm_stress-response (C). *$p < 0.05$, **$p < 0.01$, ****$p < 0.0001$, p values were derived from Cox proportional hazards model for each cancer type, and from pan-cancer meta-analysis.

(D and E) UMAP plots of the SMART-seq2 B cell atlas, showing the distribution of datasets (D) and major clusters (E).

(F) Expression of signature genes across B cell clusters in the SMART-seq2 atlas. IL35 expression was defined as the concurrent expression of its subunits *IL12A* and *EBI3*.

(G) UMAP plot showing the expression pattern of *IL10* in the SMART-seq2 atlas.

(H) *IL10* expression in B cells from the SMART-seq2 atlas selected by *in silico* FACS according to reported Breg surface markers.

(I) Heatmaps showing major lineage markers expression on *IL10*$^+$ B cells from the SMART-seq2 atlas. Each column represents an *IL10*$^+$ B cell.

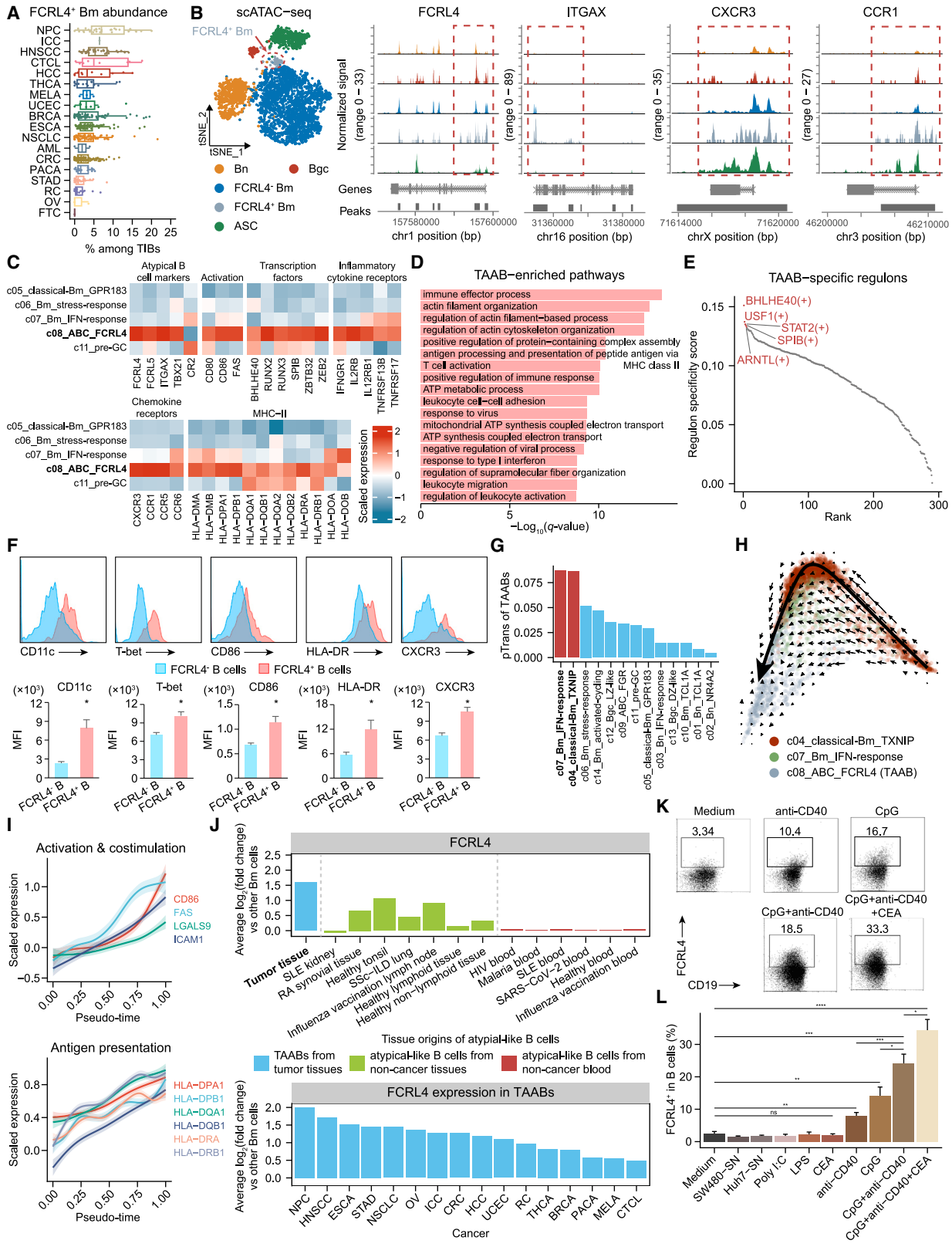
See Table S5.

stress-response-associated cell types, more single-cell level data with paired survival information are required.

We further noted that regulatory B cells (Bregs), which have been reported to play pro-tumor roles and are associated with poor prognosis,[46] were not uniquely discriminated by the unsupervised clustering of transcriptome profiles. We then examined the expression of cytokines used to distinguish Bregs, such as interleukin-10 (IL-10), IL-35, and transforming growth factor β

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

CellPress
OPEN ACCESS



(legend on next page)

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

(TGF-β).[47] Notably, *IL10* and *IL12A/EBI3* (encoding IL35) were rarely detected at the transcriptional level, while *TGFB1* expression was found in almost all TIB subsets without clear patterns (Figure S1F). To reduce the potential technical bias, such as the high drop-out ratio, we then established another atlas using only scRNA-seq datasets from the SMART-seq2 platform (Figures 5D–5F; Table S5A; STAR Methods). As expected, in this SMART-seq2-based atlas, a higher expression rate of *IL10* was observed (Figures 5F and 5G). Cells conforming to previously reported Breg surface markers, including $CD19^+CD24^+CD38^+$ and $CD19^+CD24^{high}CD27^+$,[48,49] exhibited increased *IL10* expression (Figure 5H). However, the expression of *IL10* was dispersed across different B cell major clusters (Figure 5G). Examining those $IL10^+$ B cells also revealed that these cells could variably express different markers indicative of different B cell major lineages (Figure 5I). Briefly, these results supported a previous view that Bregs, or at least $IL10^+$ B cells might not represent a distinct lineage of B cells, but rather potentially arise from various stages of B cell differentiation in a context-dependent manner.[50]

Activated phenotype of tumor-associated ABCs

The above analyses collectively pointed to the pivotal role of $FCRL4^+$ Bm cells in tumors. Although with varying abundances, $FCRL4^+$ Bm cells were identified in tumors of almost all cancer types we analyzed (Figure 6A). We then investigated the epigenome landscape of TIBs by analyzing a single-cell sequencing assay for transposase-accessible chromatin (scATAC-seq) dataset of BRCA patients.[51] At the epigenetic level, all B cell major lineages could be readily identified through unsupervised clustering (STAR Methods). A small cluster of Bm cells was distinguished as $FCRL4^+$ Bm cells (Figure 6B). Delving into the specifically accessible chromatin regions of this subset revealed enhanced openness in DNA segments of transcriptional marker genes of c08 cells (Figure 6B), further confirming the epigenetic distinctiveness of $FCRL4^+$ Bm cells. Thus, $FCRL4^+$ Bm cells might be a common, stable cell type across cancer types. Considering the transcriptional resemblance of tumor-infiltrating $FCRL4^+$ Bm cells with previously reported ABCs (Figures 1C and 6C),[27–29] we thereby named them tumor-associated ABCs (TAABs).

We next sought to interrogate the molecular properties of TAABs in tumors. Of note, such cells exhibited enhanced activities in multiple processes pertaining to BCR signaling and B cell activation, including actin cytoskeleton reconstruction, cell-cell adhesion, and antigen presentation (Figure 6D). Consistently, compared with other tissue-enriched Bm cells, TAABs exhibited an elevated expression of B cell activation markers including *CD80*, *CD86*, and *FAS* (encoding CD95) (Figure 6C). In addition, certain transcription factors, including *BHLHE40* and *SPIB* (Figure 6C), which are known to regulate the BCR signaling pathway and impact the B cell activation, respectively,[52,53] were highly expressed in TAABs. The SCENIC-based regulon analysis[54] also confirmed the activity of these transcription factors (Figure 6E; STAR Methods). These observations, together with their aforementioned clonal expansion state, indicated that TAABs might experience antigen stimulation-based activation in tumors. Furthermore, TAABs upregulated MHC-II genes, suggestive of their antigen presentation potentials (Figure 6C). Moreover, we observed higher expression of multiple cytokine receptors including *IFNGR1*, *CXCR3*, and *CCR1* (Figure 6C), which are known to be involved in the response to inflammatory signals,[55–57] in TAABs. Flow cytometry confirmed that $FCRL4^+$ B cells sorted from hepatocellular carcinoma (HCC) and intrahepatic cholangiocarcinoma (ICC) fresh tumor tissues displayed higher expression of ABC markers CD11c and T-bet, activation marker CD86, MHC-II molecule HLA-DR, and chemokine receptor CXCR3 when compared with $FCRL4^-$ B cells (Figures 6F and S6A; Table S6; STAR Methods). CODEX staining in a STAD tumor specimen also corroborated the expression of CD86 and CXCR3 in $FCRL4^+$ B cells (Figure S6B; Table S2; STAR Methods).

We then examined the developmental dynamics of TAABs by checking their lineage relationship with other Bm cells and computationally inferring the evolutionary trajectory in tumors. First, quantifying the degree of BCR sharing of TAABs with other clusters by the STARTRAC transition index[58] revealed that two Bm subsets, c04_classical-Bm_TXNIP and c07_Bm_IFN-response, exhibited notably higher transition propensity with TAABs than others (Figure 6G; STAR Methods). We thereby focused on these three subsets, c04, c07, and c08, and used

Figure 6. Epigenetic and transcriptional characteristics of tumor-associated atypical B cells

(A) Proportions of $FCRL4^+$ Bm cells among TIBs across cancer types. Points represent treatment-naive tumor samples with B cells > 50.

(B) t-distributed stochastic neighbor embedding (t-SNE) plot showing B cell annotations in a scATAC-seq dataset of BRCA patients (left). ATAC-seq tracks showing chromatin accessibilities at four TAAB marker gene loci (right).

(C) Expression of selected genes differentially expressed in TAABs compared with other tissue-enriched Bm subsets in tumors (BH-adjusted p values < 0.01, two-sided unpaired Wilcoxon test).

(D) Top pathways enriched in TAABs compared with other Bm cells.

(E) TAAB-specific regulons ranked by regulon specificity scores in descending order. The top five regulons are highlighted.

(F) Flow cytometry analysis of $FCRL4^+$ and $FCRL4^-$ B cells isolated from tumor samples (n = 5). MFI, median fluorescence intensity. Data are represented as mean ± SEM. *p < 0.05, two-sided paired t test.

(G) The pairwise transition index (pTrans) of TIB subsets with TAABs, with the top two subsets highlighted.

(H) The RNA velocity field visualized on the diffusion map of intratumoral c04, c07, and c08 cells.

(I) The relationship between selected gene expression and pseudo-time with a 95% confidence interval.

(J) Average $\log_2$ fold change of *FCRL4* expression between ABCs and other Bm cells. TAABs from the entire TIB atlas along with ABCs from each non-cancer condition (top), TAABs from each cancer type of the TIB atlas (bottom) were examined.

(K) Induction experiments of $FCRL4^+$ B cells. CEA, carcinoembryonic antigen.

(L) The percentage of $FCRL4^+$ B cells among total B cells cultured under different conditions. $n \geq 3$. Data are represented as mean ± SEM. *p < 0.05, **p < 0.01, ****p < 0.0001, two-sided unpaired t test.

See also Figure S6 and Tables S2, S5, and S6.

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

CellPress
OPEN ACCESS

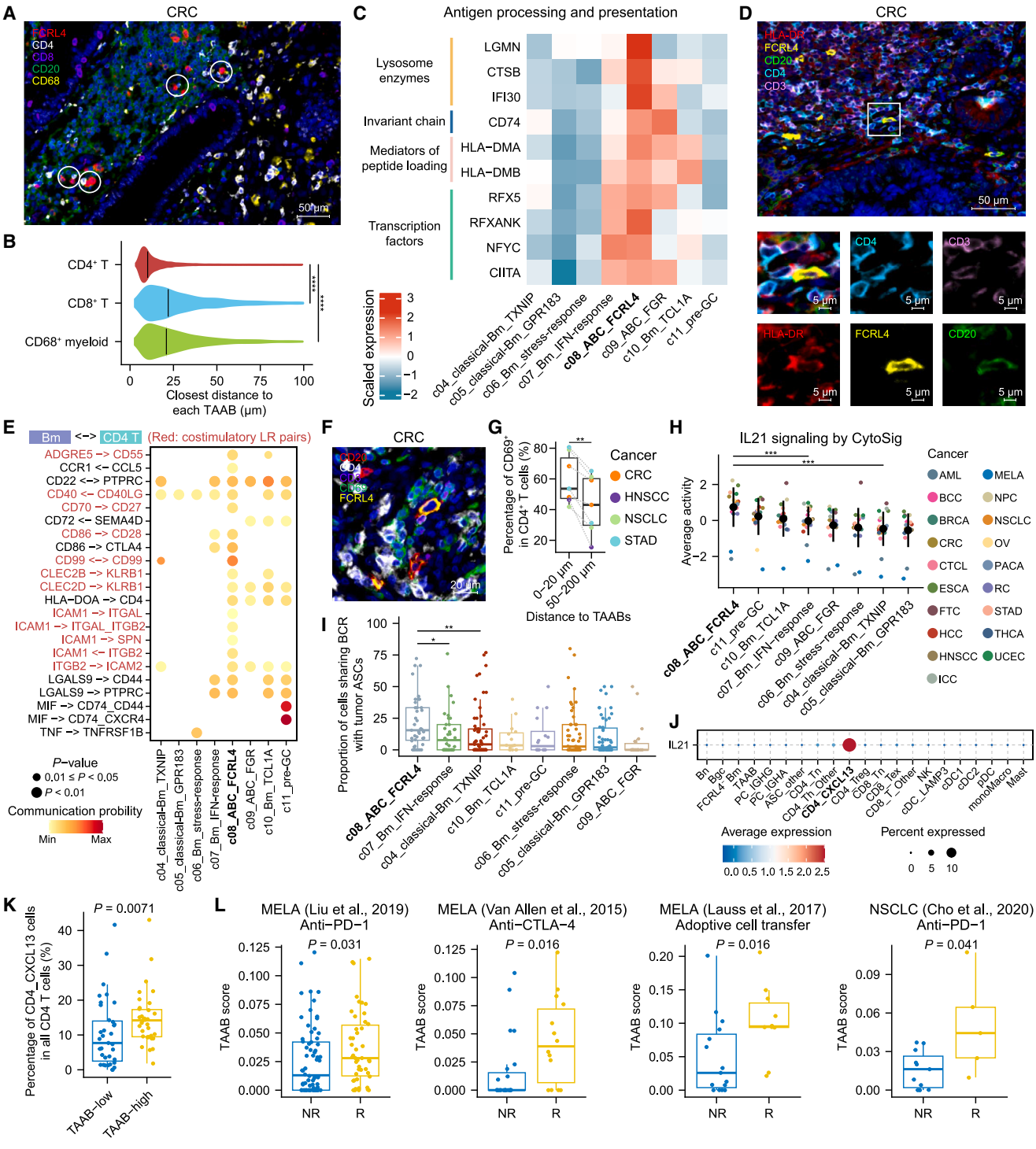


Figure 7. Cellular communication and clinical associations of TAABs

(A) mIHC of a CRC tumor to show the proximity relationship between TAABs (CD20⁺FCRL4⁺) and CD4⁺ T cells (CD4⁺CD68⁻), CD8⁺ T cells (CD8⁺), or CD68⁺ myeloid cells (CD68⁺). White circles indicate juxtaposition between TAABs and CD4⁺ T cells.

(B) Distances between each TAAB and their nearest CD4⁺ T, CD8⁺ T, or CD68⁺ myeloid cells in all stained tumors.

(C) Expression of antigen processing and presentation-associated genes across intra-tumoral Bm subsets. All genes here are highly expressed in TAABs (BH-adjusted p values < 0.01).

(D) mIHC of a CRC tumor to show the juxtaposition between TAABs and CD4⁺ T cells with HLA-DR expression at their junction.

(legend continued on next page)

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

the RNA velocity analysis[59,60] to decipher their transition directions, with a directional flow, starting from c04 and c07 to c08 cells, observed (Figure 6H; STAR Methods). Another single-cell fate-mapping algorithm, CellRank,[61] consistently predicted TAABs as the terminal state among these three subsets (Figure S6C; STAR Methods). We further inferred the pseudo-time of each cell to reflect their developmental order (STAR Methods), finding that the density of TAABs peaked at the region near the largest pseudo-time value (Figure S6D). Correlating the pseudo-time with the gene expression profile showed that the increase of pseudo-time was accompanied by a gradual upregulation of ABC marker genes, T cell costimulation-associated genes, MHC-II genes, and inflammatory cytokine receptors (Figures 6I and S6E; STAR Methods). Together, our results suggested a differentiated and activated state of TAABs.

To further delineate the distinctive features of TAABs, we compared them with ABCs in other immune contexts. We first constructed an integrative scRNA-seq B cell atlas from 16 published studies under several non-cancer conditions,[28–30,41,62–73] encompassing health, autoimmune diseases, infections, and vaccination (Figures S6F–S6H; Table S5B; STAR Methods). The atypical-like cluster was identified by unsupervised clustering, highly expressing *FCRL5*, *ITGAX*, and *TBX21* (Figure S6G). Notably, while *FCRL4* was widely detected in TAABs in our atlas, it exhibited only a marginal detection rate in ABCs from the non-cancer atlas. Examining the expression fold change of the ABC markers between ABCs and other Bm cells unveiled that *FCRL4* displayed a selectively higher level in tumors compared with non-cancer conditions (Figures 6J and S6I). This observation suggested *FCRL4* as a context-dependent ABC marker, with a preference for intratumoral TAABs. To explore the potential mechanisms behind the high expression of *FCRL4* in TAABs, we conducted *in vitro* induction of FCRL4$^+$ B cells and found that FCRL4 expression could be effectively induced in B cells through a combination of carcinoembryonic antigen, CpG, and anti-CD40 (Figures 6K and 6L; STAR Methods).

TAABs interact with CD4 T cells and could act as an immunotherapy response predictor

Considering that TIBs have been reported to engage other immune cells, especially T cells, we next explored the physical neighborhood and molecular regulators of TAABs in the TME. First, on seven tumor samples from three cancer types (Table S2), we performed mIHC to locate TAABs and other immune cell types (Figures 7A and S7A). Throughout the entire tumor sections, we identified the physically nearest CD4$^+$ T, CD8$^+$ T, and CD68$^+$ myeloid cell to each FCRL4$^+$ TAAB, respectively (STAR Methods). Quantitative assessments revealed a closer spatial relationship between TAABs and the nearest CD4$^+$ T cells (median distance ~10 μm) than CD8$^+$ T cells and CD68$^+$ myeloid cells (both median distances ~25 μm) (Figure 7B). Such proximity, combined with the elevated MHC-II gene expression on TAABs (Figure 6C) led us to hypothesize that TAABs might interact with CD4 T cells.

At the transcriptional level, TAABs were observed to upregulate multiple genes involved in various steps of antigen processing and presentation (Figure 7C).[74,75] Consistently, additional mIHC on tumor samples revealed clear HLA-DR expression at the junctions between TAABs and CD4$^+$ T cells (Figures 7D and S7B; Table S2; STAR Methods), suggesting that TAABs could potentially provide signal for CD4 T cell activation via MHC-II antigen presentation. Notably, the aforementioned high expression of *CD80* and *CD86* in TAABs reflected their potential to provide the costimulatory signal (Figure 6C). The CellChat-based analyses further showed that TAABs exhibited the strongest interaction potential with CD4 T cells among TIBs (Figure S7C; STAR Methods). Specifically, TAABs harbored more costimulatory ligand-receptor pairs that could mediate the interaction with CD4 T cells than other Bm subsets, with prominent examples including CD86–CD28,[76] CD99–CD99,[77] ADGRE5–CD55,[78] and ICAM1/2 related adhesion molecules[79,80] (Figure 7E). Notably, in tumor samples across various cancer types, TAABs uniformly displayed higher overall expression of MHC-II genes and those costimulation-related molecules than other Bm cells (Figure S7D). Further, mIHC on tumor samples revealed that CD4 T cells situated in close proximity (< 20 μm) to TAABs exhibited higher expression of the activation marker CD69, compared with those located 50–200 μm away (Figures 7F and 7G; Table S2; STAR Methods). Together, TAABs might be specialized for antigen presentation, potentially providing two signals via MHC-II and costimulatory molecules, respectively, and thereby were associated with the activation of CD4 T cells.

TAABs might simultaneously receive stimulatory signals from CD4 T cells. Indeed, costimulatory ligand-receptor pairs typically undergo bidirectional co-signaling.[81] For example, the CD86–CD28 engagement has been reported to promote B cell activation and IgG secretion via direct signaling in murine models.[82] To further elucidate the TAAB cell surface molecule signaling events implicated in their interaction with CD4 T cells, we

(E) Significant ligand-receptor pairs between CD4 T cells and each Bm subset within tumors. Red indicates pairs with costimulatory functions. Pairs that are significant in all Bm subsets are not shown.
(F) mIHC of a CRC tumor to show the distribution of activated CD4$^+$ T cells (CD3$^+$CD4$^+$CD69$^+$) around TAABs (CD20$^+$FCRL4$^+$).
(G) The relationship between the distance of CD4$^+$ T cells to TAABs and the percentage of activated CD4$^+$ T cells among total CD4$^+$ T cells. Two-sided paired t test.
(H) IL-21 signaling activity in each Bm subset within tumors. Dots represent cancer types. Black lines with midpoints represent mean ± SD.
(I) Proportions of cells sharing BCR with intratumoral ASCs for each Bm subset.
(J) *IL21* expression in TME immune subtypes.
(K) Comparison of the percentage of *CXCL13*$^+$ CD4 T cells in all CD4 T cells between two groups of tumor samples, stratified by the median TAAB abundance among B cells.
(L) Comparison of the TAAB signature expression between responders and non-responders of cancer immunotherapy.
In (B), (C), (H), (I), (K), and (L), two-sided unpaired Wilcoxon test was used. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.
See also Figure S7 and Table S2.

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

additionally performed the gene set enrichment analysis (GSEA) of TAABs compared with other Bm cells within tumor tissues (STAR Methods). TAABs appeared to highly express genes in multiple pathways associated with B cell activation, including CD40 signaling, integrin-mediated signaling,[83] and semaphorin-plexin signaling pathway[84] (Figure S7E). In addition, as CD4 T cells can also secrete miscellaneous cytokines to regulate other immune cell types,[85] we then leveraged CytoSig[86] to analyze the cytokine signaling in TAABs (STAR Methods). IL-21, which is known to promote B cell activation and expansion and function as the most potent inducer of PC differentiation,[87,88] was ranked among the top cytokines (Figure S7F) and demonstrated higher signaling activity in TAABs than all other Bm subsets in tumors (Figure 7H). Accordingly, we observed that TAABs possessed a higher BCR sharing degree with ASCs in tumors than other Bm cells (Figure 7I). ASCs sharing BCR with TAABs exhibited an IGHG1-biased isotype preference compared with other ASCs (Figure S7G). We then examined the expression pattern of *IL21* in TME immune subtypes, finding that *IL21* was exclusively expressed by a *CXCL13*$^+$ subset of CD4 T cells (Figure 7J). *CXCL13*$^+$ CD4 T cells have been referred to as follicular helper,[89,90] tumor-reactive,[2,91,92] or exhausted CD4 T cells[93,94] and can promote effective adaptive immunity as well as predict immunotherapy response in human cancers.[95,96] We then further investigated the association between TAABs and *CXCL13*$^+$ CD4 T cells within the CD4 T compartment. CD4 T cells from tumors harboring higher TAAB prevalence featured elevated expression of multiple marker genes of *CXCL13*$^+$ CD4 T cells, including *CXCL13*, *SOX4*, and *LAG3*[2,91,92] (Figure S7H), as well as a higher proportion of *CXCL13*$^+$ CD4 T cells (Figure 7K), implying the co-occurrence of TAABs and tumor-reactive CD4 T cells in the TME. In brief, CD4 T cells, especially tumor-reactive CD4 T cells, might cooperate with TAABs in their activation and differentiation toward PCs.

Taken together, our analyses revealed that TAABs, through reciprocally interacting with CD4 T cells, might manifest two facets of TIB-associated anti-tumor activities,[19] as antigen-presenting cells to engage in T cell response and as PC precursors to mediate humoral response, aligning with their associations with favorable prognosis in multiple cancers independent of the overall B cell presence (Figures 5B and S7I). We also collected datasets with treatment-naive tumor samples from individuals who underwent anti-PD-1, anti-CTLA-4, or adoptive cell transfer therapy, finding that in NSCLC and melanoma, responders tended to harbor higher expression of the TAAB signature when compared with non-responders (Figures 7L and S7J; STAR Methods).

DISCUSSION

Here, the comprehensive single-cell atlas of human B cells we built enabled a systematic delineation of the population landscape of TIBs at the pan-cancer level, as well as their spatial distribution patterns, clonal states, and developmental dynamics. TIBs exhibited a clear preference for certain cancer types, with their abundances affected by both the organ contexture and TME-associated factors. Our atlas also enabled the identification of critical transitional subsets amidst B cell differentiation stages, such as pre-GC B cells and plasmablasts, within tumors. Particularly, we demonstrated the occurrence of germinal center reactions within tumors akin to SLOs, with the process of the iterative BCR sequence evolution and the genesis of both Bm cells and ASCs recapitulated. These observations underscore the TME as a dynamic ecosystem allowing *in situ* B cell activation and differentiation. In our BCR analysis, BCR sharing between two individuals was barely detectable. The signatures of LZ/DZ-like Bgc cells did not show distinct associations with favorable prognosis at the pan-cancer level. Considering the Bgc cells are rare and typically present only in a portion of tumors with mature TLS, thus marginally contributing to the bulk-level expression, it is still challenging to quantify their subtle appearance through computational strategies in TCGA bulk data.

TIBs primarily function as tumor-associated antigen-presenting cells and antibody producers, serving as a link between the innate and adaptive mechanisms of the cognate antigen-driven immune response.[19] However, limited attention has been given to the investigation of specific B cell subsets responsible for carrying out these activities. Our fine-grained transcriptome atlas of TIBs allowed the subset-level depiction of TIB phenotypes. In particular, we identified a tumor-enriched FCRL4$^+$ Bm subset as a potential contributor to the TIB-engaged antigen presentation during the tumor-associated immune response. In addition, a skewness toward IgG, especially IgG1, in the isotypes of intratumoral PCs was demonstrated, indicating that TME-specific stimulation or selection pressure might influence the antibody repertoire. Previous studies have demonstrated both anti-tumor and pro-tumor roles of PCs in different contexts. We observed a broad spectrum of macrophage subsets that expressed IgG receptors, and except for the role in guiding anti-tumor antibody-dependent cell-mediated cytotoxicity (ADCC) and antibody-dependent cellular phagocytosis (ADCP), IgG-mediated macrophage activation may also be implicated in angiogenesis and immunosuppression, or other pro-tumor activity.[46] Meanwhile, IgA antibodies can also engage in T cell-mediated cancer elimination to undertake anti-tumor roles.[97] Given that IgG and IgA PCs have been reported to communicate with other immune components to coordinate anti- or pro-tumor activities in different cancer types,[98,99] the investigation into their prognostic association should further consider the context of multi-cellular modules in tumors. Of note, although we identified evidence for tumor-associated B cell immune responses in tumors, our data cannot distinguish immune responses specific to tumor-specific antigens.

Two Bm subsets, c08_ABC_FCRL4 and c09_ABC_FGR, displayed a shared gene signature with previously reported ABCs, as well as those variably termed double-negative, exhausted, tissue-like memory, and age-associated B cells under different immune contexts including infections, autoimmune diseases, vaccination, and aging.[27–29,100,101] In contrast to the blood-enriched c09 cells, the tumor-enriched FCRL4$^+$ Bm cells (c08) featured high clonal expansion levels and proliferative capacity, as well as a highly activated transcriptional state within tumors, and were thereby named TAABs. The expression pattern of *CR2* and *CD86* in TAABs was reminiscent of the previously reported CD21^{low}CD86$^+$ B cells in CRC, which can present antigens to enhance the IFN-γ secretion of autologous CD3$^+$

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

T cells.[102] In addition, alongside the expression of *CXCR5* similar to other CD20$^+$ B cell subsets, TAABs also possessed a notably high expression of *CXCR3*, suggesting distinctive chemotactic behaviors. Given the wide expression of *CXCR3* in CD4 T cells,[103] we postulated that TAABs and CD4 T cells might share common migration patterns.

The TAAB signature exhibited the strongest association with favorable prognosis among all B cell subsets at the pan-cancer level, yet the context-dependent nature of its prognostic significance is present. Our results aligned with previous reports that FCRL4$^+$ B cells or ABCs have been linked to favorable survival in NSCLC[24] and HCC[104] but poorer outcomes in CRC.[105] Interestingly, in renal cancer, where B cells have been associated with unfavorable prognosis,[13] the TAAB signature was predictive of improved patient outcomes. Notably, the limited presence of TAABs in most renal cancer patients (Figure 6A), when compared with other cancer types, might constrain their beneficial impact. Conversely, in NPC, which harbored relatively high TAAB abundances across all cancer types, FCRL4$^+$ B cells have been reported to confer positive prognostic values.[22,23] In addition, TAABs might serve as a predictor of response to multiple immunotherapy strategies. *CXCL13*$^+$ CD4 T cells, known to promote effective adaptive immunity and serve as a biomarker of immunotherapy response in human cancers,[95,96] appear to communicate with TAABs. Further research is needed to explore how the anti-tumor potential of TAABs can be harnessed for clinical practices.

In our attempt to pinpoint TIB subsets associated with unfavorable patient outcomes, we found that Bregs, or at least *IL10*$^+$ B cells, remained elusive in the single-cell clustering analysis and potentially arose from various stages of B cell differentiation. Such observations aligned with previous observations that Breg cells are found among B cell populations of different maturation and differentiation stages, including early transitional B cells,[49,106] as well as Bm cells[106] and plasmablasts.[107,108] Nevertheless, the roles of Bregs should be further explored beyond transcriptome-based analyses. Apart from Bregs, a tumor-enriched B cell subset with a stress phenotype appeared to be associated with unfavorable prognosis. Previously, we and others found that within the tumor-infiltrating T cell and NK cell compartments, the enrichment of cells with a stress-response signature was also linked with immunotherapy resistance or poor survival.[44,45] We hypothesize that this stressed state observed across various immune cell types could be induced by shared cues within the TME. Although c06_Bm_stress-response cells did not exhibit elevated expression of immunosuppressive molecules or reduced expression of activation markers and antigen presentation-associated genes, they were found to downregulate BCR signaling-related genes (Figure 1C) and exhibit a low potential of interaction with CD4 T cells (Figure S7C).

During the final stages of our revision process, a newly published study also highlighted the role of ABCs in human cancers.[109] Unlike their treating of these cells as one homogeneous population, our analyses disentangled them into two subclusters, c08_ABC_FCRL4 and c09_ABC_FGR. Importantly, only c08 was enriched in tumors, while c09 cells were primarily present in blood. Re-analysis of their defined ABC population (B09.DUSP4+ AtM) confirmed a similar division into such two subsets (Figure S7K). This difference in the level of granularity might account for the discrepancies observed between these two studies. While their study claimed ABCs in tumors were derived from extrafollicular responses, our data suggested that the possible origins from the germinal center should not be ruled out. The SHM rates of c08 cells were higher than Bgc cells, while c09 cells harbored SHMs below the Bgc level (Figure 1D). Additionally, the CSR rates of c08 cells were markedly higher than c09 cells. Given the role of germinal centers as the primary site for SHM and CSR,[50] these differences indicated varying levels of GC dependency during the development of these two ABC subsets. Whether ABCs have germinal center origins[110,111] or are derived from the extrafollicular response[112,113] remains unresolved and requires further investigation.

To summarize, our study improves the current understanding of B cells from a pan-cancer perspective, shedding light on the heterogeneity of B cells and their anti-tumor immune response. We hope that our large-scale data can further enable the development of B cell-associated immunotherapy.

Limitations of the study

While our results reveal the heterogeneity of B cells across cancer types, certain cancer types, including NPC, acute myeloid leukemia (AML), and cutaneous T cell lymphoma (CTCL), were composed of biopsy samples, which might suffer from greater sampling bias compared with surgical sections. Although the functional and compositional variations of B cells could be associated with clinical measures, such as the reported variation in CD20$^+$ B cell proportions across tumor stages in NSCLC,[114] the limited availability of comprehensive clinical information from public datasets hinders our further investigation. Future studies with a larger cohort with matched clinical information are needed. Additionally, the limited antibody panels of our CODEX analyses and the lack of paired information about the presence of TLS in our scRNA-seq datasets prevented us from delving deeper into the impact of TLS on anti-tumor immunity and B cell immune activity across cancer types. Nevertheless, this pan-cancer atlas provides a valuable resource for B cell studies.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Human participants
- METHOD DETAILS
  - Immunohistochemistry and multiplex Immunofluorescence
  - CODEX (now called PhenoCycler-Fusion)
  - Induction of FCRL4$^+$ B cells *in vitro*
  - Flow cytometry analysis of tumor-associated atypical B cells isolated from tumor tissues
  - Quantification analysis of whole-tissue scans using HALO
- QUANTIFICATION AND STATISTICAL ANALYSIS

- Single-cell RNA-seq data collection, preprocessing and the identification of B cells
- Data integration, dimension reduction, and unsupervised clustering
- Scoring cells using gene expression signatures
- Reconstruction and analysis of the BCR repertoires of TIBs
- Assessment of infiltration status for major immune components in tumors
- Cellular interaction analysis by CellChat
- Tissue distribution preference of TIB subsets
- Cell type assignment of activated cycling Bm cells by CellTypist
- Evaluation of the transcriptional heterogeneity of B cell clusters
- TCGA data analysis
- Construction of a SMART-seq2-based single-cell transcriptome atlas of B cells
- Analysis of scATAC-seq data
- SCENIC regulon analysis
- RNA velocity analysis and pseudo-time trajectory inference
- Construction and analyses of an integrative scRNA-seq B cell atlas under non-cancer conditions
- GSEA of cell surface receptor signaling pathways
- CytoSig analysis of cytokine signaling activity within cells
- Analysis of immunotherapy datasets

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2024.06.038.

ACKNOWLEDGMENTS

We thank M. Selby and L. Wang for their valuable consultation and Z. Li and C. Li for their insightful discussions. We thank Z. Ye, Y. Huang, and Z. Deng for their assistance in bioinformatic data analysis. Part of the analysis was performed on the High Performance Computing Platform of the Center for Life Sciences (Peking University). We thank the Bioimaging Core of Shenzhen Bay Laboratory for providing imaging support. This project was supported by funding from the National Natural Science Foundation of China (81988101, 62203019, 91959000, and 31991171), the Beijing Municipal Science and Technology Commission (Z221100007022002), the National Key Research and Development Program of China (2023YFF1204700), Major Program of Shenzhen Bay Laboratory (S201101004), and the Open Program of Shenzhen Bay Laboratory (SZBL2020090501005).

AUTHOR CONTRIBUTIONS

Z.Z., D.W., and M.-M.C. designed this study. Y.Y. and X.C. performed bioinformatic data analysis, and J.L., S.Q., and X.R. assisted. M.-M.C., P.Y., H.N., and Y.B. performed the experiments. Y.Z. collected the surgical samples for flow cytometry analyses. The manuscript was written by Y.Y., X.C., M.-M.C., D.W., and Z.Z. with contributions from all authors.

DECLARATION OF INTERESTS

Z.Z. is a founder of Analytical Bioscience and also serves on the Advisory Board of *Cell*. All financial interests are unrelated to this study.

Received: June 9, 2023
Revised: April 25, 2024
Accepted: June 26, 2024
Published: July 23, 2024

REFERENCES

1. Wang, D.F., Liu, B.L., and Zhang, Z.M. (2023). Accelerating the understanding of cancer biology through the lens of genomics. Cell *186*, 1755–1771. https://doi.org/10.1016/j.cell.2023.02.015.
2. Zheng, L., Qin, S., Si, W., Wang, A., Xing, B., Gao, R., Ren, X., Wang, L., Wu, X., Zhang, J., et al. (2021). Pan-cancer single-cell landscape of tumor-infiltrating T cells. Science *374*, abe6474. https://doi.org/10.1126/science.abe6474.
3. Cheng, S., Li, Z., Gao, R., Xing, B., Gao, Y., Yang, Y., Qin, S., Zhang, L., Ouyang, H., Du, P., et al. (2021). A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. Cell *184*, 792–809.e23. https://doi.org/10.1016/j.cell.2021.01.010.
4. Sharma, P., Goswami, S., Raychaudhuri, D., Siddiqui, B.A., Singh, P., Nagarajan, A., Liu, J., Subudhi, S.K., Poon, C., Gant, K.L., et al. (2023). Immune checkpoint therapy—current perspectives and future directions. Cell *186*, 1652–1669. https://doi.org/10.1016/j.cell.2023.03.006.
5. Morgan, D., and Tergaonkar, V. (2022). Unraveling B cell trajectories at single cell resolution. Trends Immunol. *43*, 210–229. https://doi.org/10.1016/j.it.2022.01.003.
6. Hoffman, W., Lakkis, F.G., and Chalasani, G. (2016). B Cells, Antibodies, and More. Clin. J. Am. Soc. Nephrol. *11*, 137–154. https://doi.org/10.2215/Cjn.09430915.
7. Mazor, R.D., Nathan, N., Gilboa, A., Stoler-Barak, L., Moss, L., Solomonov, I., Hanuna, A., Divinsky, Y., Shmueli, M.D., Hezroni, H., et al. (2022). Tumor-reactive antibodies evolve from non-binding and autoreactive precursors. Cell *185*, 1208–1222.e21. https://doi.org/10.1016/j.cell.2022.02.012.
8. Meylan, M., Petitprez, F., Becht, E., Bougoüin, A., Pupier, G., Calvez, A., Giglioli, I., Verkarre, V., Lacroix, G., Verneau, J., et al. (2022). Tertiary lymphoid structures generate and propagate anti-tumor antibody-producing plasma cells in renal cell cancer. Immunity *55*, 527–541.e5. https://doi.org/10.1016/j.immuni.2022.02.001.
9. Cui, C., Wang, J., Fagerberg, E., Chen, P.-M., Connolly, K.A., Damo, M., Cheung, J.F., Mao, T., Askari, A.S., Chen, S., et al. (2021). Neoantigen-driven B cell and CD4 T follicular helper cell collaboration promotes anti-tumor CD8 T cell responses. Cell *184*, 6101–6118.e13. https://doi.org/10.1016/j.cell.2021.11.007.
10. Fridman, W.H., Meylan, M., Petitprez, F., Sun, C.-M., Italiano, A., and Sautès-Fridman, C. (2022). B cells and tertiary lymphoid structures as determinants of tumour immune contexture and clinical outcome. Nat. Rev. Clin. Oncol. *19*, 441–457. https://doi.org/10.1038/s41571-022-00619-z.
11. Schumacher, T.N., and Thommen, D.S. (2022). Tertiary lymphoid structures in cancer. Science *375*, eabf9419. https://doi.org/10.1126/science.abf9419.
12. Iglesia, M.D., Parker, J.S., Hoadley, K.A., Serody, J.S., Perou, C.M., and Vincent, B.G. (2016). Genomic Analysis of Immune Cell Infiltrates Across 11 Tumor Types. J. Natl. Cancer Inst. *108*, djw144. https://doi.org/10.1093/jnci/djw144.
13. Laumont, C.M., Banville, A.C., Gilardi, M., Hollern, D.P., and Nelson, B.H. (2022). Tumour-infiltrating B cells: immunological mechanisms, clinical impact and therapeutic opportunities. Nat. Rev. Cancer *22*, 414–430. https://doi.org/10.1038/s41568-022-00466-1.
14. Wouters, M.C.A., and Nelson, B.H. (2018). Prognostic Significance of Tumor-Infiltrating B Cells and Plasma Cells in Human Cancer. Clin. Cancer Res. *24*, 6125–6135. https://doi.org/10.1158/1078-0432.Ccr-18-1481.
15. Patil, N.S., Nabet, B.Y., Müller, S., Koeppen, H., Zou, W., Giltnane, J., Au-Yeung, A., Srivats, S., Cheng, J.H., Takahashi, C., et al. (2022). Intratumoral plasma cells predict outcomes to PD-L1 blockade in non-small cell lung cancer. Cancer Cell *40*, 289–300.e4. https://doi.org/10.1016/j.ccell.2022.02.002.
16. Cabrita, R., Lauss, M., Sanna, A., Donia, M., Skaarup Larsen, M., Mitra, S., Johansson, I., Phung, B., Harbst, K., Vallon-Christersson, J., et al. (2020). Tertiary lymphoid structures improve immunotherapy and survival in melanoma. Nature *577*, 561–565. https://doi.org/10.1038/s41586-019-1914-8.

17. Helmink, B.A., Reddy, S.M., Gao, J., Zhang, S., Basar, R., Thakur, R., Yizhak, K., Sade-Feldman, M., Blando, J., Han, G., et al. (2020). B cells and tertiary lymphoid structures promote immunotherapy response. Nature *577*, 549–555. https://doi.org/10.1038/s41586-019-1922-8.

18. Petitprez, F., De Reyniès, A., Keung, E.Z., Chen, T.W.-W., Sun, C.-M., Calderaro, J., Jeng, Y.-M., Hsiao, L.-P., Lacroix, L., Bougoüin, A., et al. (2020). B cells are associated with survival and immunotherapy response in sarcoma. Nature *577*, 556–560. https://doi.org/10.1038/s41586-019-1906-8.

19. Laumont, C.M., and Nelson, B.H. (2023). B cells in the tumor microenvironment: Multi-faceted organizers, regulators, and effectors of anti-tumor immunity. Cancer Cell *41*, 466–489. https://doi.org/10.1016/j.ccell.2023.02.017.

20. Hu, Q., Hong, Y., Qi, P., Lu, G., Mai, X., Xu, S., He, X., Guo, Y., Gao, L., Jing, Z., et al. (2021). Atlas of breast cancer infiltrated B-lymphocytes revealed by paired single-cell RNA-sequencing and antigen receptor profiling. Nat. Commun. *12*, 2186. https://doi.org/10.1038/s41467-021-22300-2.

21. Hao, D.P., Han, G.C., Sinjab, A., Gomez-Bolanos, L.I., Lazcano, R., Serrano, A., Hernandez, S.D., Dai, E.Y., Cao, X.Y., Hu, J., et al. (2022). The Single-Cell Immunogenomic Landscape of B and Plasma Cells in Early-Stage Lung Adenocarcinoma. Cancer Discov. *12*, 2626–2645. https://doi.org/10.1158/2159-8290.Cd-21-1658.

22. Gong, L., Kwong, D.L.-W., Dai, W., Wu, P., Li, S., Yan, Q., Zhang, Y., Zhang, B., Fang, X., Liu, L., et al. (2021). Comprehensive single-cell sequencing reveals the stromal dynamics and tumor-specific characteristics in the microenvironment of nasopharyngeal carcinoma. Nat. Commun. *12*, 1540. https://doi.org/10.1038/s41467-021-21795-z.

23. Chen, Y.-P., Yin, J.-H., Li, W.-F., Li, H.-J., Chen, D.-P., Zhang, C.-J., Lv, J.-W., Wang, Y.-Q., Li, X.-M., Li, J.-Y., et al. (2020). Single-cell transcriptomics reveals regulators underlying immune cell diversity and immune subtypes associated with prognosis in nasopharyngeal carcinoma. Cell Res. *30*, 1024–1042. https://doi.org/10.1038/s41422-020-0374-x.

24. Hu, J., Zhang, L., Xia, H., Yan, Y., Zhu, X., Sun, F., Sun, L., Li, S., Li, D., Wang, J., et al. (2023). Tumor microenvironment remodeling after neoadjuvant immunotherapy in non-small cell lung cancer revealed by single-cell RNA sequencing. Genome Med. *15*, 14. https://doi.org/10.1186/s13073-023-01164-9.

25. Song, L., Cohen, D., Ouyang, Z., Cao, Y., Hu, X., and Liu, X.S. (2021). TRUST4: immune repertoire reconstruction from bulk and single-cell RNA-seq data. Nat. Methods *18*, 627–630. https://doi.org/10.1038/s41592-021-01142-2.

26. Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nat. Methods *16*, 1289–1296. https://doi.org/10.1038/s41592-019-0619-0.

27. Gao, X., and Cockburn, I.A. (2022). The development and function of CD11c+ atypical B cells - insights from single cell analysis. Front. Immunol. *13*, 979060. https://doi.org/10.3389/fimmu.2022.979060.

28. Holla, P., Dizon, B., Ambegaonkar, A.A., Rogel, N., Goldschmidt, E., Boddapati, A.K., Sohn, H., Sturdevant, D., Austin, J.W., Kardava, L., et al. (2021). Shared transcriptional profiles of atypical B cells suggest common drivers of expansion and function in malaria, HIV, and autoimmunity. Sci. Adv. *7*, eabg8384. https://doi.org/10.1126/sciadv.abg8384.

29. Horns, F., Dekker, C.L., and Quake, S.R. (2020). Memory B Cell Activation, Broad Anti-influenza Antibodies, and Bystander Activation Revealed by Single-Cell Transcriptomics. Cell Rep. *30*, 905–913.e6. https://doi.org/10.1016/j.celrep.2019.12.063.

30. King, H.W., Orban, N., Riches, J.C., Clear, A.J., Warnes, G., Teichmann, S.A., and James, L.K. (2021). Single-cell analysis of human B cell maturation predicts how antibody class switching shapes selection dynamics. Sci. Immunol. *6*, eabe6291. https://doi.org/10.1126/sciimmunol.abe6291.

31. Calado, D.P., Sasaki, Y., Godinho, S.A., Pellerin, A., Köchert, K., Sleckman, B.P., de Alborán, I.M., Janz, M., Rodig, S., and Rajewsky, K. (2012). The cell-cycle regulator c-Myc is essential for the formation and maintenance of germinal centers. Nat. Immunol. *13*, 1092–1100. https://doi.org/10.1038/ni.2418.

32. Nutt, S.L., Hodgkin, P.D., Tarlinton, D.M., and Corcoran, L.M. (2015). The generation of antibody-secreting plasma cells. Nat. Rev. Immunol. *15*, 160–171. https://doi.org/10.1038/nri3795.

33. Duan, M., Nguyen, D.C., Joyner, C.J., Saney, C.L., Tipton, C.M., Andrews, J., Lonial, S., Kim, C., Hentenaar, I., Kosters, A., et al. (2023). Understanding heterogeneity of human bone marrow plasma cell maturation and survival pathways by single-cell analyses. Cell Rep. *42*, 112682. https://doi.org/10.1016/j.celrep.2023.112682.

34. Sanz, I., Wei, C., Jenks, S.A., Cashman, K.S., Tipton, C., Woodruff, M.C., Hom, J., and Lee, F.E.-H. (2019). Challenges and Opportunities for Consistent Classification of Human B Cell and Plasma Cell Populations. Front. Immunol. *10*, 2458.

35. Garaud, S., Buisseret, L., Solinas, C., Gu-Trantien, C., de Wind, A., Van den Eynden, G., Naveaux, C., Lodewyckx, J.-N., Boisson, A., Duvillier, H., et al. (2019). Tumor infiltrating B-cells signal functional humoral immune responses in breast cancer. JCI Insight *5*, e129641. https://doi.org/10.1172/jci.insight.129641.

36. Nielsen, J.S., Sahota, R.A., Milne, K., Kost, S.E., Nesslinger, N.J., Watson, P.H., and Nelson, B.H. (2012). CD20+ tumor-infiltrating lymphocytes have an atypical CD27- memory phenotype and together with CD8+ T cells promote favorable prognosis in ovarian cancer. Clin. Cancer Res. *18*, 3281–3292. https://doi.org/10.1158/1078-0432.CCR-12-0234.

37. Jin, S.Q., Guerrero-Juarez, C.F., Zhang, L.H., Chang, I., Ramos, R., Kuan, C.H., Myung, P., Plikus, M.V., and Nie, Q. (2021). Inference and analysis of cell-cell communication using CellChat. Nat. Commun. *12*, 1088. https://doi.org/10.1038/s41467-021-21246-9.

38. Barinov, A., Luo, L., Gasse, P., Meas-Yedid, V., Donnadieu, E., Arenzana-Seisdedos, F., and Vieira, P. (2017). Essential role of immobilized chemokine CXCL12 in the regulation of the humoral immune response. Proc. Natl. Acad. Sci. USA *114*, 2319–2324. https://doi.org/10.1073/pnas.1611958114.

39. Nakagawa, R., and Calado, D.P. (2021). Positive Selection in the Light Zone of Germinal Centers. Front. Immunol. *12*, 661678.

40. Cyster, J.G., and Allen, C.D.C. (2019). B Cell Responses: Cell Interaction Dynamics and Decisions. Cell *177*, 524–540. https://doi.org/10.1016/j.cell.2019.03.016.

41. Domínguez Conde, C., Xu, C., Jarvis, L.B., Rainbow, D.B., Wells, S.B., Gomes, T., Howlett, S.K., Suchanek, O., Polanski, K., King, H.W., et al. (2022). Cross-tissue immune cell analysis reveals tissue-specific features in humans. Science *376*, eabl5197. https://doi.org/10.1126/science.abl5197.

42. Li, Y., Jin, L., and Chen, T.X. (2020). The Effects of Secretory IgA in the Mucosal Immune System. BioMed Res. Int. *2020*, 2032057. https://doi.org/10.1155/2020/2032057.

43. Muthuswamy, R.V., Sundström, P., Börjesson, L., Gustavsson, B., and Quiding-Järbrink, M. (2013). Impaired migration of IgA-secreting cells to colon adenocarcinomas. Cancer Immunol. Immunother. *62*, 989–997. https://doi.org/10.1007/s00262-013-1410-1.

44. Tang, F., Li, J., Qi, L., Liu, D., Bo, Y., Qin, S., Miao, Y., Yu, K., Hou, W., Li, J., et al. (2023). A pan-cancer single-cell panorama of human natural killer cells. Cell *186*, 4235–4251.e20. https://doi.org/10.1016/j.cell.2023.07.034.

45. Chu, Y., Dai, E., Li, Y., Han, G., Pei, G., Ingram, D.R., Thakkar, K., Qin, J.-J., Dang, M., Le, X., et al. (2023). Pan-cancer T cell atlas links a cellular stress response state to immunotherapy resistance. Nat. Med. *29*, 1550–1562. https://doi.org/10.1038/s41591-023-02371-y.

46. Fridman, W.H., Petitprez, F., Meylan, M., Chen, T.W.-W., Sun, C.-M., Roumenina, L.T., and Sautès-Fridman, C. (2021). B cells and cancer:

To B or not to B? J. Exp. Med. *218*, e20200851. https://doi.org/10.1084/jem.20200851.

47. Michaud, D., Steward, C.R., Mirlekar, B., and Pylayeva-Gupta, Y. (2021). Regulatory B cells in cancer. Immunol. Rev. *299*, 74–92. https://doi.org/10.1111/imr.12939.

48. Iwata, Y., Matsushita, T., Horikawa, M., Dilillo, D.J., Yanaba, K., Venturi, G.M., Szabolcs, P.M., Bernstein, S.H., Magro, C.M., Williams, A.D., et al. (2011). Characterization of a rare IL-10-competent B-cell subset in humans that parallels mouse regulatory B10 cells. Blood *117*, 530–541. https://doi.org/10.1182/blood-2010-07-294249.

49. Blair, P.A., Noreña, L.Y., Flores-Borja, F., Rawlings, D.J., Isenberg, D.A., Ehrenstein, M.R., and Mauri, C. (2010). CD19(+)CD24(hi)CD38(hi) B cells exhibit regulatory capacity in healthy individuals but are functionally impaired in systemic Lupus Erythematosus patients. Immunity *32*, 129–140. https://doi.org/10.1016/j.immuni.2009.11.009.

50. Downs-Canner, S.M., Meier, J., Vincent, B.G., and Serody, J.S. (2022). B Cell Function in the Tumor Microenvironment. Annu. Rev. Immunol. *40*, 169–193. https://doi.org/10.1146/annurev-immunol-101220-015603.

51. Zhang, Y., Chen, H., Mo, H., Hu, X., Gao, R., Zhao, Y., Liu, B., Niu, L., Sun, X., Yu, X., et al. (2021). Single-cell analyses reveal key immune cell subsets associated with response to PD-L1 blockade in triple-negative breast cancer. Cancer Cell *39*, 1578–1593.e8. https://doi.org/10.1016/j.ccell.2021.09.010.

52. Cook, M.E., Jarjour, N.N., Lin, C.C., and Edelson, B.T. (2020). Transcription Factor Bhlhe40 in Immunity and Autoimmunity. Trends Immunol. *41*, 1023–1036. https://doi.org/10.1016/j.it.2020.09.002.

53. Willis, S.N., Tellier, J., Liao, Y., Trezise, S., Light, A., O'Donnell, K., Garrett-Sinha, L.A., Shi, W., Tarlinton, D.M., and Nutt, S.L. (2017). Environmental sensing by mature B cells is controlled by the transcription factors PU.1 and SpiB. Nat. Commun. *8*, 1426. https://doi.org/10.1038/s41467-017-01605-1.

54. Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.-C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nat. Methods *14*, 1083–1086. https://doi.org/10.1038/Nmeth.4463.

55. Domeier, P.P., Chodisetti, S.B., Soni, C., Schell, S.L., Elias, M.J., Wong, E.B., Cooper, T.K., Kitamura, D., and Rahman, Z.S.M. (2016). IFN-γ receptor and STAT1 signaling in B cells are central to spontaneous germinal center formation and autoimmunity. J. Exp. Med. *213*, 715–732. https://doi.org/10.1084/jem.20151722.

56. Jenh, C.H., Cox, M.A., Hipkin, W., Lu, T., Pugliese-Sivo, C., Gonsiorek, W., Chou, C.C., Narula, S.K., and Zavodny, P.J. (2001). HUMAN B CELL-ATTRACTING CHEMOKINE 1 (BCA-1; CXCL13) IS AN AGONIST FOR THE HUMAN CXCR3 RECEPTOR. Cytokine *15*, 113–121. https://doi.org/10.1006/cyto.2001.0923.

57. Marques, R.E., Guabiraba, R., Russo, R.C., and Teixeira, M.M. (2013). Targeting CCL5 in inflammation. Expert Opin. Ther. Targets *17*, 1439–1460. https://doi.org/10.1517/14728222.2013.837886.

58. Zhang, L., Yu, X., Zheng, L., Zhang, Y., Li, Y., Fang, Q., Gao, R., Kang, B., Zhang, Q., Huang, J.Y., et al. (2018). Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. Nature *564*, 268–272. https://doi.org/10.1038/s41586-018-0694-x.

59. La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lönnerberg, P., Furlan, A., et al. (2018). RNA velocity of single cells. Nature *560*, 494–498. https://doi.org/10.1038/s41586-018-0414-6.

60. Bergen, V., Lange, M., Peidli, S., Wolf, F.A., and Theis, F.J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. Nat. Biotechnol. *38*, 1408–1414. https://doi.org/10.1038/s41587-020-0591-3.

61. Lange, M., Bergen, V., Klein, M., Setty, M., Reuter, B., Bakhti, M., Lickert, H., Ansari, M., Schniering, J., Schiller, H.B., et al. (2022). CellRank for directed single-cell fate mapping. Nat. Methods *19*, 159–170. https://doi.org/10.1038/s41592-021-01346-6.

62. Nehar-Belaid, D., Hong, S., Marches, R., Chen, G., Bolisetty, M., Baisch, J., Walters, L., Punaro, M., Rossi, R.J., Chung, C.-H., et al. (2020). Mapping systemic lupus erythematosus heterogeneity at the single-cell level. Nat. Immunol. *21*, 1094–1106. https://doi.org/10.1038/s41590-020-0743-0.

63. Perez, R.K., Gordon, M.G., Subramaniam, M., Kim, M.C., Hartoularos, G.C., Targ, S., Sun, Y., Ogorodnikov, A., Bueno, R., Lu, A., et al. (2022). Single-cell RNA-seq reveals cell type-specific molecular and genetic associations to lupus. Science *376*, eabf1970. https://doi.org/10.1126/science.abf1970.

64. Arazi, A., Rao, D.A., Berthier, C.C., Davidson, A., Liu, Y., Hoover, P.J., Chicoine, A., Eisenhaure, T.M., Jonsson, A.H., Li, S., et al. (2019). The immune cell landscape in kidneys of patients with lupus nephritis. Nat. Immunol. *20*, 902–914. https://doi.org/10.1038/s41590-019-0398-x.

65. Bhamidipati, K., Silberstein, J.L., Chaichian, Y., Baker, M.C., Lanz, T.V., Zia, A., Rasheed, Y.S., Cochran, J.R., and Robinson, W.H. (2020). CD52 Is Elevated on B cells of SLE Patients and Regulates B Cell Function. Front. Immunol. *11*, 626820. https://doi.org/10.3389/fimmu.2020.626820.

66. Zhang, F., Wei, K., Slowikowski, K., Fonseka, C.Y., Rao, D.A., Kelly, S., Goodman, S.M., Tabechian, D., Hughes, L.B., Salomon-Escoto, K., et al. (2019). Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. Nat. Immunol. *20*, 928–942. https://doi.org/10.1038/s41590-019-0378-1.

67. Penkava, F., Velasco-Herrera, M.D.C., Young, M.D., Yager, N., Nwosu, L.N., Pratt, A.G., Lara, A.L., Guzzo, C., Maroof, A., Mamanova, L., et al. (2020). Single-cell sequencing reveals clonal expansions of pro-inflammatory synovial CD8 T cells expressing tissue-homing receptors in psoriatic arthritis. Nat. Commun. *11*, 4767. https://doi.org/10.1038/s41467-020-18513-6.

68. Tsukui, T., Sun, K.-H., Wetter, J.B., Wilson-Kanamori, J.R., Hazelwood, L.A., Henderson, N.C., Adams, T.S., Schupp, J.C., Poli, S.D., Rosas, I.O., et al. (2020). Collagen-producing lung cell atlas identifies multiple subsets with distinct localization and relevance to fibrosis. Nat. Commun. *11*, 1920. https://doi.org/10.1038/s41467-020-15647-5.

69. Gao, X., Jia, G., Guttman, A., DePianto, D.J., Morshead, K.B., Sun, K.-H., Ramamoorthi, N., Vander Heiden, J.A., Modrusan, Z., Wolters, P.J., et al. (2020). Osteopontin Links Myeloid Activation and Disease Progression in Systemic Sclerosis. Cell Rep. Med. *1*, 100140. https://doi.org/10.1016/j.xcrm.2020.100140.

70. Wang, S., Zhang, Q., Hui, H., Agrawal, K., Karris, M.A.Y., and Rana, T.M. (2020). An atlas of immune cell exhaustion in HIV-infected individuals revealed by single-cell transcriptomics. Emerg. Microbes Infect. *9*, 2333–2347. https://doi.org/10.1080/22221751.2020.1826361.

71. Kazer, S.W., Aicher, T.P., Muema, D.M., Carroll, S.L., Ordovas-Montanes, J., Miao, V.N., Tu, A.A., Ziegler, C.G.K., Nyquist, S.K., Wong, E.B., et al. (2020). Integrated single-cell analysis of multicellular immune dynamics during hyperacute HIV-1 infection. Nat. Med. *26*, 511–518. https://doi.org/10.1038/s41591-020-0799-2.

72. Sokal, A., Chappert, P., Barba-Spaeth, G., Roeser, A., Fourati, S., Azzaoui, I., Vandenberghe, A., Fernandez, I., Meola, A., Bouvier-Alias, M., et al. (2021). Maturation and persistence of the anti-SARS-CoV-2 memory B cell response. Cell *184*, 1201–1213.e14. https://doi.org/10.1016/j.cell.2021.01.050.

73. Turner, J.S., Zhou, J.Q., Han, J., Schmitz, A.J., Rizk, A.A., Alsoussi, W.B., Lei, T., Amor, M., McIntire, K.M., Meade, P., et al. (2020). Human germinal centres engage memory and naive B cells after influenza vaccination. Nature *586*, 127–132. https://doi.org/10.1038/s41586-020-2711-0.

74. Honey, K., and Rudensky, A.Y. (2003). Lysosomal cysteine proteases regulate antigen presentation. Nat. Rev. Immunol. *3*, 472–482. https://doi.org/10.1038/nri1110.

75. Jabrane-Ferrat, N., Nekrep, N., Tosi, G., Esserman, L.J., and Peterlin, B.M. (2002). Major histocompatibility complex class II transcriptional platform: Assembly of nuclear factor Y and regulatory factor X (RFX) on DNA requires RFX5 dimers. Mol. Cell. Biol. *22*, 5616–5625. https://doi.org/10.1128/Mcb.22.15.5616-5625.2002.

76. Greenwald, R.J., Freeman, G.J., and Sharpe, A.H. (2005). The B7 family revisited. Annu. Rev. Immunol. *23*, 515–548. https://doi.org/10.1146/annurev.immunol.23.021704.115611.

77. Oh, K.I., Kim, B.K., Ban, Y.L., Choi, E.Y., Jung, K.C., Lee, I.-S., and Park, S.H. (2007). CD99 activates T cells via a costimulatory function that promotes raft association of TCR complex and tyrosine phosphorylation of TCR ζ. Exp. Mol. Med. *39*, 176–184. https://doi.org/10.1038/emm.2007.20.

78. Capasso, M., Durrant, L.G., Stacey, M., Gordon, S., Ramage, J., and Spendlove, I. (2006). Costimulation via CD55 on Human CD4+ T Cells Mediated by CD97. J. Immunol. *177*, 1070–1077. https://doi.org/10.4049/jimmunol.177.2.1070.

79. Dubey, C., Croft, M., and Swain, S.L. (1995). Costimulatory Requirements of Naive Cd4(+) T-Cells - Icam-1 or B7-1 Can Costimulate Naive Cd4 T-Cell Activation but Both Are Required for Optimum Response. J. Immunol. Baltim. *155*, 45–57.

80. Damle, N.K., Klussman, K., and Aruffo, A. (1992). Intercellular adhesion molecule-2, a second counter-receptor for CD11a/CD18 (leukocyte function-associated antigen-1), provides a costimulatory signal for T-cell receptor-initiated activation of human T cells. J. Immunol. *148*, 665–671.

81. Chen, L., and Flies, D.B. (2013). Molecular mechanisms of T cell co-stimulation and co-inhibition. Nat. Rev. Immunol. *13*, 227–242. https://doi.org/10.1038/nri3405.

82. Rau, F.C., Dieter, J., Luo, Z., Priest, S.O., and Baumgarth, N. (2009). B7-1/2 (CD80/CD86) Direct Signaling to B Cells Enhances IgG Secretion. J. Immunol. *183*, 7661–7671. https://doi.org/10.4049/jimmunol.0803783.

83. Arana, E., Harwood, N.E., and Batista, F.D. (2008). Regulation of integrin activation through the B-cell receptor. J. Cell Sci. *121*, 2279–2286. https://doi.org/10.1242/jcs.017905.

84. Takamatsu, H., Okuno, T., and Kumanogoh, A. (2010). Regulation of immune cell responses by semaphorins and their receptors. Cell. Mol. Immunol. *7*, 83–88. https://doi.org/10.1038/cmi.2009.111.

85. Propper, D.J., and Balkwill, F.R. (2022). Harnessing cytokines and chemokines for cancer therapy. Nat. Rev. Clin. Oncol. *19*, 237–253. https://doi.org/10.1038/s41571-021-00588-9.

86. Jiang, P., Zhang, Y., Ru, B., Yang, Y., Vu, T., Paul, R., Mirza, A., Altan-Bonnet, G., Liu, L., Ruppin, E., et al. (2021). Systematic investigation of cytokine signaling activity at the tissue and single-cell levels. Nat. Methods *18*, 1181–1191. https://doi.org/10.1038/s41592-021-01274-5.

87. Kuchen, S., Robbins, R., Sims, G.P., Sheng, C., Phillips, T.M., Lipsky, P.E., and Ettinger, R. (2007). Essential Role of IL-21 in B Cell Activation, Expansion, and Plasma Cell Generation during CD4+ T Cell-B Cell Collaboration. J. Immunol. *179*, 5886–5896. https://doi.org/10.4049/jimmunol.179.9.5886.

88. Moens, L., and Tangye, S.G. (2014). Cytokine-Mediated Regulation of Plasma Cell Generation: IL-21 Takes Center Stage. Front. Immunol. *5*, 65. https://doi.org/10.3389/fimmu.2014.00065.

89. Gu-Trantien, C., Loi, S., Garaud, S., Equeter, C., Libin, M., de Wind, A., Ravoet, M., Le Buanec, H., Sibille, C., Manfouo-Foutsop, G., et al. (2013). CD4$^+$ follicular helper T cell infiltration predicts breast cancer survival. J. Clin. Invest. *123*, 2873–2892. https://doi.org/10.1172/JCI67428.

90. Gu-Trantien, C., Migliori, E., Buisseret, L., de Wind, A., Brohée, S., Garaud, S., Noël, G., Dang Chi, V.L., Lodewyckx, J.-N., Naveaux, C., et al. (2017). CXCL13-producing TFH cells link immune suppression and adaptive memory in human breast cancer. JCI Insight *2*, e91487. https://doi.org/10.1172/jci.insight.91487.

91. Lowery, F.J., Krishna, S., Yossef, R., Parikh, N.B., Chatani, P.D., Zacharakis, N., Parkhurst, M.R., Levin, N., Sindiri, S., Sachs, A., et al. (2022). Molecular signatures of antitumor neoantigen-reactive T cells from metastatic human cancers. Science *375*, 877–884. https://doi.org/10.1126/science.abl5447.

92. Liu, B., Zhang, Y., Wang, D., Hu, X., and Zhang, Z. (2022). Single-cell meta-analyses reveal responses of tumor-reactive CXCL13+ T cells to immune-checkpoint blockade. Nat. Cancer *3*, 1123–1136. https://doi.org/10.1038/s43018-022-00433-7.

93. Cohen, M., Giladi, A., Barboy, O., Hamon, P., Li, B., Zada, M., Gurevich-Shapiro, A., Beccaria, C.G., David, E., Maier, B.B., et al. (2022). The interaction of CD4+ helper T cells with dendritic cells shapes the tumor microenvironment and immune checkpoint blockade response. Nat. Cancer *3*, 303–317. https://doi.org/10.1038/s43018-022-00338-5.

94. Oliveira, G., Stromhaug, K., Cieri, N., Iorgulescu, J.B., Klaeger, S., Wolff, J.O., Rachimi, S., Chea, V., Krause, K., Freeman, S.S., et al. (2022). Landscape of helper and regulatory antitumour CD4+ T cells in melanoma. Nature *605*, 532–538. https://doi.org/10.1038/s41586-022-04682-5.

95. Noël, G., Fontsa, M.L., Garaud, S., De Silva, P., de Wind, A., Van den Eynden, G.G., Salgado, R., Boisson, A., Locy, H., Thomas, N., et al. (2021). Functional Th1-oriented T follicular helper cells that infiltrate human breast cancer promote effective adaptive immunity. J. Clin. Invest. *131*, e139905. https://doi.org/10.1172/JCI139905.

96. Goubet, A.-G., Lordello, L., Alves Costa Silva, C., Peguillet, I., Gazzano, M., Mbogning-Fonkou, M.D., Thelemaque, C., Lebacle, C., Thibault, C., Audenet, F., et al. (2022). Escherichia coli-Specific CXCL13-Producing TFH Are Associated with Clinical Efficacy of Neoadjuvant PD-1 Blockade against Muscle-Invasive Bladder Cancer. Cancer Discov. *12*, 2280–2307. https://doi.org/10.1158/2159-8290.CD-22-0201.

97. Biswas, S., Mandal, G., Payne, K.K., Anadon, C.M., Gatenbee, C.D., Chaurio, R.A., Costich, T.L., Moran, C., Harro, C.M., Rigolizzo, K.E., et al. (2021). IgA transcytosis and antigen recognition govern ovarian cancer immunity. Nature *591*, 464–470. https://doi.org/10.1038/s41586-020-03144-0.

98. Shalapour, S., Lin, X.-J., Bastian, I.N., Brain, J., Burt, A.D., Aksenov, A.A., Vrbanac, A.F., Li, W., Perkins, A., Matsutani, T., et al. (2017). Inflammation-induced IgA+ cells dismantle anti-liver cancer immunity. Nature *551*, 340–345. https://doi.org/10.1038/nature24302.

99. Leader, A.M., Grout, J.A., Maier, B.B., Nabet, B.Y., Park, M.D., Tabachnikova, A., Chang, C., Walker, L., Lansky, A., Le Berichel, J., et al. (2021). Single-cell analysis of human non-small cell lung cancer lesions refines tumor classification and patient stratification. Cancer Cell *39*, 1594–1609.e12. https://doi.org/10.1016/j.ccell.2021.10.009.

100. Courey-Ghaouzi, A.-D., Kleberg, L., and Sundling, C. (2022). Alternative B Cell Differentiation During Infection and Inflammation. Front. Immunol. *13*, 908034. https://doi.org/10.3389/fimmu.2022.908034.

101. Gjertsson, I., McGrath, S., Grimstad, K., Jonsson, C.A., Camponeschi, A., Thorarinsdottir, K., and Mårtensson, I.-L. (2022). A close-up on the expanding landscape of CD21-/low B cells in humans. Clin. Exp. Immunol. *210*, 217–229. https://doi.org/10.1093/cei/uxac103.

102. Wennhold, K., Thelen, M., Lehmann, J., Schran, S., Preugszat, E., Garcia-Marquez, M., Lechner, A., Shimabukuro-Vornhagen, A., Ercanoglu, M.S., Klein, F., et al. (2021). CD86+ Antigen-Presenting B Cells Are Increased in Cancer, Localize in Tertiary Lymphoid Structures, and Induce Specific T-cell Responses. Cancer Immunol. Res. *9*, 1098–1108. https://doi.org/10.1158/2326-6066.Cir-20-0949.

103. Karin, N. (2020). CXCR3 Ligands in Cancer and Autoimmunity, Chemoattraction of Effector T Cells, and Beyond. Front. Immunol. *11*, 976. https://doi.org/10.3389/fimmu.2020.00976.

104. Shi, J.-Y., Gao, Q., Wang, Z.-C., Zhou, J., Wang, X.-Y., Min, Z.-H., Shi, Y.-H., Shi, G.-M., Ding, Z.-B., Ke, A.-W., et al. (2013). Margin-infiltrating

CD20(+) B cells display an atypical memory phenotype and correlate with favorable prognosis in hepatocellular carcinoma. Clin. Cancer Res. *19*, 5994–6005. https://doi.org/10.1158/1078-0432.CCR-12-3497.

105. Sorrentino, C., D'Antonio, L., Fieni, C., Ciummo, S.L., and Di Carlo, E. (2021). Colorectal Cancer-Associated Immune Exhaustion Involves T and B Lymphocytes and Conventional NK Cells and Correlates With a Shorter Overall Survival. Front. Immunol. *12*, 778329. https://doi.org/10.3389/fimmu.2021.778329.

106. Khoder, A., Sarvaria, A., Alsuliman, A., Chew, C., Sekine, T., Cooper, N., Mielke, S., de Lavallade, H., Muftuoglu, M., Fernandez Curbelo, I., et al. (2014). Regulatory B cells are enriched within the IgM memory and transitional subsets in healthy donors but are deficient in chronic GVHD. Blood *124*, 2034–2045. https://doi.org/10.1182/blood-2014-04-571125.

107. Matsumoto, M., Baba, A., Yokota, T., Nishikawa, H., Ohkawa, Y., Kayama, H., Kallies, A., Nutt, S.L., Sakaguchi, S., Takeda, K., et al. (2014). Interleukin-10-producing plasmablasts exert regulatory function in autoimmune inflammation. Immunity *41*, 1040–1051. https://doi.org/10.1016/j.immuni.2014.10.016.

108. de Masson, A., Bouaziz, J.-D., Le Buanec, H., Robin, M., O'Meara, A., Parquet, N., Rybojad, M., Hau, E., Monfort, J.-B., Branchtein, M., et al. (2015). CD24(hi)CD27^{+} and plasmablast-like regulatory B cells in human chronic graft-versus-host disease. Blood *125*, 1830–1839. https://doi.org/10.1182/blood-2014-09-599159.

109. Ma, J., Wu, Y., Ma, L., Yang, X., Zhang, T., Song, G., Li, T., Gao, K., Shen, X., Lin, J., et al. (2024). A blueprint for tumor-infiltrating B cells across human cancers. Science *384*, eadj4857. https://doi.org/10.1126/science.adj4857.

110. Gao, X., Shen, Q., Roco, J.A., Dalton, B., Frith, K., Munier, C.M.L., Ballard, F.D., Wang, K., Kelly, H.G., Nekrasov, M., et al. (2024). Zeb2 drives the formation of CD11c+ atypical B cells to sustain germinal centers that control persistent infection. Sci. Immunol. *9*, eadj4748. https://doi.org/10.1126/sciimmunol.adj4748.

111. Levack, R.C., Newell, K.L., Popescu, M., Cabrera-Martinez, B., and Winslow, G.M. (2020). CD11c+ T-bet+ B Cells Require IL-21 and IFN-γ from Type 1 T Follicular Helper Cells and Intrinsic Bcl-6 Expression but Develop Normally in the Absence of T-bet. J. Immunol. *205*, 1050–1058. https://doi.org/10.4049/jimmunol.2000206.

112. Jenks, S.A., Cashman, K.S., Zumaquero, E., Marigorta, U.M., Patel, A.V., Wang, X., Tomar, D., Woodruff, M.C., Simon, Z., Bugrovsky, R., et al. (2018). Distinct Effector B Cells Induced by Unregulated Toll-like Receptor 7 Contribute to Pathogenic Responses in Systemic Lupus Erythematosus. Immunity *49*, 725–739.e6. https://doi.org/10.1016/j.immuni.2018.08.015.

113. Song, W., Antao, O.Q., Condiff, E., Sanchez, G.M., Chernova, I., Zembrzuski, K., Steach, H., Rubtsova, K., Angeletti, D., Lemenze, A., et al. (2022). Development of Tbet- and CD11c-expressing B cells in a viral infection requires T follicular helper cells outside of germinal centers. Immunity *55*, 290–307.e5. https://doi.org/10.1016/j.immuni.2022.01.002.

114. Chen, J., Tan, Y., Sun, F., Hou, L., Zhang, C., Ge, T., Yu, H., Wu, C., Zhu, Y., Duan, L., et al. (2020). Single-cell transcriptome and antigen-immunoglobin analysis reveals the diversity of B cells in non-small cell lung cancer. Genome Biol. *21*, 152. https://doi.org/10.1186/s13059-020-02064-6.

115. Liu, D., Schilling, B., Liu, D., Sucker, A., Livingstone, E., Jerby-Arnon, L., Zimmer, L., Gutzmer, R., Satzger, I., Loquai, C., et al. (2019). Integrative molecular and clinical modeling of clinical outcomes to PD1 blockade in patients with metastatic melanoma. Nat. Med. *25*, 1916–1927. https://doi.org/10.1038/s41591-019-0654-5.

116. Van Allen, E.M., Miao, D., Schilling, B., Shukla, S.A., Blank, C., Zimmer, L., Sucker, A., Hillen, U., Foppen, M.H.G., Goldinger, S.M., et al. (2015). Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. Science *350*, 207–211. https://doi.org/10.1126/science.aad0095.

117. Lauss, M., Donia, M., Harbst, K., Andersen, R., Mitra, S., Rosengren, F., Salim, M., Vallon-Christersson, J., Törngren, T., Kvist, A., et al. (2017). Mutational and putative neoantigen load predict clinical benefit of adoptive T cell therapy in melanoma. Nat. Commun. *8*, 1738. https://doi.org/10.1038/s41467-017-01460-0.

118. Cho, J.-W., Hong, M.H., Ha, S.-J., Kim, Y.-J., Cho, B.C., Lee, I., and Kim, H.R. (2020). Genome-wide identification of differentially methylated promoters and enhancers associated with response to anti-PD-1 therapy in non-small cell lung cancer. Exp. Mol. Med. *52*, 1550–1563. https://doi.org/10.1038/s12276-020-00493-8.

119. Gide, T.N., Quek, C., Menzies, A.M., Tasker, A.T., Shang, P., Holst, J., Madore, J., Lim, S.Y., Velickovic, R., Wongchenko, M., et al. (2019). Distinct Immune Cell Populations Define Response to Anti-PD-1 Monotherapy and Anti-PD-1/Anti-CTLA-4 Combined Therapy. Cancer Cell *35*, 238–255.e6. https://doi.org/10.1016/j.ccell.2019.01.003.

120. Auslander, N., Zhang, G., Lee, J.S., Frederick, D.T., Miao, B., Moll, T., Tian, T., Wei, Z., Madan, S., Sullivan, R.J., et al. (2018). Robust prediction of response to immune checkpoint blockade therapy in metastatic melanoma. Nat. Med. *24*, 1545–1549. https://doi.org/10.1038/s41591-018-0157-9.

121. Prat, A., Navarro, A., Paré, L., Reguart, N., Galván, P., Pascual, T., Martínez, A., Nuciforo, P., Comerma, L., Alos, L., et al. (2017). Immune-Related Gene Expression Profiling After PD-1 Blockade in Non-Small Cell Lung Carcinoma, Head and Neck Squamous Cell Carcinoma, and Melanoma. Cancer Res. *77*, 3540–3550. https://doi.org/10.1158/0008-5472.Can-16-3556.

122. R Core Team (2023). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).

123. Van Rossum, G., and Drake, F.L. (2009). Python 3 Reference Manual (CreateSpace).

124. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.

125. Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. *19*, 15.

126. Gupta, N.T., Vander Heiden, J.A., Uduman, M., Gadala-Maria, D., Yaari, G., and Kleinstein, S.H. (2015). Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. Bioinformatics *31*, 3356–3358. https://doi.org/10.1093/bioinformatics/btv359.

127. Ye, J., Ma, N., Madden, T.L., and Ostell, J.M. (2013). IgBLAST: an immunoglobulin variable domain sequence analysis tool. Nucleic Acids Res. *41*, W34–W40. https://doi.org/10.1093/nar/gkt382.

128. Bodenhofer, U., Bonatesta, E., Horejš-Kainrath, C., and Hochreiter, S. (2015). msa: an R package for multiple sequence alignment. Bioinformatics *31*, 3997–3999. https://doi.org/10.1093/bioinformatics/btv494.

129. Dolgalev, I. (2022). msigdbr: MSigDB Gene Sets for Multiple Organisms in a Tidy Data Format. GitHub. https://igordot.github.io/msigdbr/.

130. Stuart, T., Srivastava, A., Madad, S., Lareau, C.A., and Satija, R. (2021). Single-cell chromatin state analysis with Signac. Nat. Methods *18*, 1333–1341. https://doi.org/10.1038/s41592-021-01282-5.

131. Liu, B., Li, C., Li, Z., Wang, D., Ren, X., and Zhang, Z. (2020). An entropy-based metric for assessing the purity of single cell populations. Nat. Commun. *11*, 3155. https://doi.org/10.1038/s41467-020-16904-3.

132. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888–1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.

133. Moir, S., Ho, J., Malaspina, A., Wang, W., DiPoto, A.C., O'Shea, M.A., Roby, G., Kottilil, S., Arthos, J., Proschan, M.A., et al. (2008). Evidence for HIV-associated B cell exhaustion in a dysfunctional memory B cell compartment in HIV-infected viremic individuals. J. Exp. Med. *205*, 1797–1805. https://doi.org/10.1084/jem.20072683.

134. Wu, C.M., Fu, Q., Guo, Q., Chen, S., Goswami, S., Sun, S.H., Li, T., Cao, X.J., Chu, F.Y., Chen, Z.C., et al. (2019). Lupus-associated atypical memory B cells are mTORC1-hyperactivated and functionally dysregulated. Ann. Rheum. Dis. *78*, 1090–1100. https://doi.org/10.1136/annrheumdis-2019-215039.

135. Verstappen, G.M., Ice, J.A., Bootsma, H., Pringle, S., Haacke, E.A., De Lange, K., Van Der Vries, G.B., Hickey, P., Vissink, A., Spijkervet, F.K.L., et al. (2020). Gene expression profiling of epithelium-associated FcRL4+ B cells in primary Sjögren's syndrome reveals a pathogenic signature. J. Autoimmun. *109*, 102439. https://doi.org/10.1016/j.jaut.2020.102439.

136. Isnardi, I., Ng, Y.-S., Menard, L., Meyers, G., Saadoun, D., Srdanovic, I., Samuels, J., Berman, J., Buckner, J.H., Cunningham-Rundles, C., and Meffre, E. (2010). Complement receptor 2/CD21− human naive B cells contain mostly autoreactive unresponsive clones. Blood *115*, 5026–5036. https://doi.org/10.1182/blood-2009-09-243071.

137. Thorsson, V., Gibbs, D.L., Brown, S.D., Wolf, D., Bortone, D.S., Ou Yang, T.-H., Porta-Pardo, E., Gao, G.F., Plaisier, C.L., Eddy, J.A., et al. (2018). The Immune Landscape of Cancer. Immunity *48*, 812–830.e14. https://doi.org/10.1016/j.immuni.2018.03.023.

138. Haghverdi, L., Buettner, F., and Theis, F.J. (2015). Diffusion maps for high-dimensional single-cell analysis of differentiation data. Bioinformatics *31*, 2989–2998. https://doi.org/10.1093/bioinformatics/btv325.

139. Haghverdi, L., Büttner, M., Wolf, F.A., Buettner, F., and Theis, F.J. (2016). Diffusion pseudotime robustly reconstructs lineage branching. Nat. Methods *13*, 845–848. https://doi.org/10.1038/nmeth.3971.

140. Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. Nature *566*, 496–502. https://doi.org/10.1038/s41586-019-0969-x.

141. Shannon, C.E. (1948). A Mathematical Theory of Communication. Bell Syst. Tech. J. *27*, 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Purified antibodies	Various	See Table S7 for details
Biological samples		
FFPE tissue samples	National Human Genetic Resources Sharing Service Platform	See Table S2A for details
Fresh tissue samples	Sun Yat-sen University Cancer Center	See Table S6 for details
Critical commercial assays		
Tumor Dissociation Kit, human	Miltenyi Biotec	Cat#130-095-929
eBioscience™ Foxp3 / Transcription Factor Staining Buffer Set	Invitrogen	Cat#00-5523-00
PanoPANEL Kits	Panovue	Cat#10234100050
PhenoCycler-Fusion Conjugation Kit	Akoya Biosciences	Cat#7000009
Deposited data		
Data files for human scRNA-seq dataset	This study	GEO: GSE233236
Raw data for human scRNA-seq dataset	This study	GSA for human: HRA000321
Public human cancer scRNA-seq datasets	Various public studies	See Table S1A for details
SMART-seq2 scRNA-seq datasets	Various public studies	See Table S5A for details
scRNA-seq datasets from non-cancer conditions	Various public studies	See Table S5B for details
Human BCRA scATAC-seq dataset	Zhang et al.[51]	GEO: GSE169246
Human melanoma anti-PD-1 treatment dataset	Liu et al.[115]	dbGAP: phs000452.v3.p1
Human melanoma anti-CTLA-4 treatment dataset	Van Allen et al.[116]	dbGaP: phs000452.v2.p1
Human melanoma adoptive cell transfer treatment dataset	Lauss et al.[117]	GEO: GSE100797
Human lung cancer anti-PD-1 treatment dataset	Cho et al.[118]	GEO: GSE126044
Human melanoma anti-PD-1 treatment dataset	Gide et al.[119]	ENA: PRJEB23709
Human melanoma anti-CTLA-4 treatment dataset	Auslander et al.[120]	GEO: GSE115821
Human lung cancer anti-PD-1 treatment dataset	Prat et al.[121]	GEO: GSE93157
Human pan-cancer B cell scRNA-seq dataset	Ma et al.[109]	http://pancancer.cn/B/
Software and algorithms		
ImageScope v12.4.3.5008	Leica Biosystems	https://www.leicabiosystems.com/digital-pathology/manage/aperio-imagescope/
OlyVIA 3.3	OLYMPUS	https://olyvia.software.informer.com/
HALO 3.5	Indica Labs	https://www.indicalab.com/halo
HighPlex FL 4.2.3	Indica Labs	https://indicalab.com/products/high-plex-fl/
Cell Ranger 3.0.0	10x Genomics	https://10xgenomics.com/
R 4.1.2	R Core Team[122]	https://www.r-project.org/
Python 3.8.0	Python Software Foundation[123]	https://www.python.org/

(*Continued on next page*)

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Seurat 4.3.0	Hao et al.[124]	https://satijalab.org/seurat
Scanpy 1.8.2	Wolf et al.[125]	https://scanpy.readthedocs.io
Harmony 0.1.1	Korsunsky et al.[26]	https://github.com/immunogenomics/harmony
TRUST4 v1.0.6	Song et al.[25]	https://github.com/liulab-dfci/TRUST4
Change-O 1.3.0	Gupta et al.[126]	https://changeo.readthedocs.io/
IgBLAST 1.18.0	Ye et al.[127]	https://ncbi.github.io/
SHazaM 1.1.2	Gupta et al.[126]	https://shazam.readthedocs.io/
STARTRAC 0.1.0	Zhang et al.[58]	https://github.com/Japrin/STARTRAC
msa 1.26.0	Bodenhofer et al.[128]	https://github.com/UBod/msa
Alakazam 1.2.1	Gupta et al.[126]	https://alakazam.readthedocs.io/
msigdbr 7.5.1	Dolgalev[129]	https://github.com/igordot/msigdbr
CellChat 1.5.0	Jin et al.[37]	https://github.com/sqjin/CellChat
CellTypist 0.1.9	Domínguez Conde et al.[41]	https://github.com/Teichlab/celltypist
Signac 1.13.0	Stuart et al.[130]	https://stuartlab.org/signac/
SCENIC 1.3.1	Aibar et al.[54]	https://github.com/aertslab/SCENIC
scVelo 0.2.5	Bergen et al.[60]	https://github.com/theislab/scvelo
CellRank 1.5.1	Lange et al.[61]	https://github.com/theislab/cellrank
CytoSig 0.0.2	Jiang et al.[86]	https://github.com/data2intelligence/CytoSig
ROGUE	Liu et al.[131]	https://github.com/PaulingLiu/ROGUE
Other		
Interactive explorer of human B cells	This study	http://pan-b.cancer-pku.cn/
IHC and mIHC imaging data	This study	http://pan-b.cancer-pku.cn/downloads

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Zemin Zhang (zemin@pku.edu.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The raw sequencing data can be accessed from the Genome Sequence Archive (GSA) using the accession number HRA000321. The gene expression data have been deposited at the Gene Expression Omnibus (GEO) with the accession number GSE233236. Visualization of the scRNA-seq data from our atlas is accessible on the open-access website http://pan-b.cancer-pku.cn. IHC and mIHC imaging data are available in the download section at http://pan-b.cancer-pku.cn/downloads. This paper analyzed existing, publicly available data. These accession numbers for the datasets are listed in the key resources table.
- The analysis code has been deposited on GitHub (https://github.com/yuyang3/pan-B).
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Human participants

Our self-generated scRNA-seq dataset, which was reported and analyzed in our previous studies,[2,3] was composed of surgical specimens of fresh ANT (2-3 cm from the edge of matched tumor tissues) and primary tumor samples, collected from 52 treatment-naïve cancer patients and each measuring 2-3 cm^3. Of note, before library preparation and sequencing, single-cell suspensions were labeled with anti-CD45 antibodies and 7AAD for FACS, performed on a BD Aria SORP instrument. Written informed

consent was provided by every patient. This study was approved by the Research and Ethical Committee of Peking University Cancer Hospital and complied with relevant ethical regulations. The ages of the patients ranged from 20 to 77, with a mean age of 54. Among these patients, 17 were male and 35 were female. All of them were East Asians. The detailed information of these patients was summarized in Table S1B.

For flow cytometry analysis of B cells isolated from tumor tissues, five patients diagnosed with liver cancer were enrolled through the Sun Yat-Sen University Cancer Center in Guangzhou, China. Written informed consents were obtained from all participants before tissue sample collection for research according to regular principles. Ethical approval was obtained from the Ethics Committee of Sun Yat-Sen University Cancer Center. Patients were pathologically diagnosed with HCC, except P20230911 and P20230925, who were diagnosed with ICC. All patients had no prior chemotherapy or targeted systemic therapy. The ages of these patients ranged from 37 to 68, with a mean age of 58. Among these patients, four were male and one was female. All of them were East Asians. Detailed clinical metadata of these patients were provided in Table S6.

METHOD DETAILS

Immunohistochemistry and multiplex Immunofluorescence

Human tissue specimens were provided by the National Human Genetic Resources Sharing Service Platform (10234100050) (Table S2). Formalin-fixed paraffin-embedded (FFPE) tissues were cut into 4-μm sections and placed on poly-lysine-coated slides. Before staining, sections were deparaffinized in xylene and were then rehydrated in 100%, 90%, and 70% alcohol successively, followed by microwave-based antigen retrieval, endogenous peroxidase inactivation and nonspecific site blocking. For IHC, primary antibodies were incubated at 4°C overnight, followed by HRP-labeled secondary antibody incubation and DAB visualization. Nuclei were stained with hematoxylin. Slides were enclosed using neutral balsam mounting medium, scanned using the Versa 200 (Leica), and analyzed with ImageScope software. For mIHC, PANO 6-plex IHC kits (the maximum number of protein labels is 5, plus DAPI, Panovue, 10234100050) were used according to the manufacturer's instructions. In some cases, just two or three protein markers were labeled in a panel (see Table S2B). Briefly, different primary antibodies (Table S7) were sequentially applied, followed by HRP-conjugated secondary antibody incubation and tyramide signal amplification. The sections were microwave heat-treated after each TSA operation. Nuclei were stained with DAPI after all the human antigens had been labeled. Following enclosure by ProLongTM Diamond Antifade Mountant (Invitrogen), slides were scanned using the SLIDEVIEW VS200 (Olympus) and analyzed with OlyVIA software. The slides used for the statistical analysis in Figure 1E included both serial and non-serial sections. Detailed information regarding serial and non-serial sections is available on the website http://pan-b.cancer-pku.cn/downloads/.

CODEX (now called PhenoCycler-Fusion)

Before staining, custom conjugated antibodies (Table S7) were prepared using the PhenoCycler-Fusion Conjugation Kit (PN# 7000009). Briefly, the antibody was reduced to expose thiol ends of the antibody heavy chains, and then conjugated with a PhenoCycler barcode. Purified PhenoCycler custom-conjugated antibodies were validated and tested before use.

Tumor tissue samples were prepared and stained following PhenoCycler-Fusion User Manual Rev K (https://www.akoyabio.com). FFPE sections provided by the National Human Genetic Resources Sharing Service Platform (10234100050) were deparaffinized and rehydrated, followed by antigen retrieval in AR9 (Panovue, 0019020500) for 20min. Slides were washed in ddH_2O and then equilibrated in staining buffer. The conjugated antibody cocktail solution including 30 antibodies (Table S7) was added to slides in a humidity chamber and incubated for 3h at room temperature. After incubation, the samples were washed and fixed following the PhenoCycler User Manual. For data acquisition, sample slides were mounted on a microscope stage and the images were acquired using at ×20 objective.

Induction of FCRL4$^+$ B cells *in vitro*

Human peripheral blood mononuclear cells (PBMCs) were isolated by density gradient centrifugation with ficoll paque plus (Cytiva,17-1440-02). Total B cells were isolated by magnetic activated cell sorting with CD19 Beads (Miltenyi Biotec, 130-050-301) and cultured in 96-well round bottom plates. B cells were left untreated or treated with 5ng/mL LPS (Sigma-Aldrich, L4391), 5ng/mL poly I:C (Sigma-Adrich, P0913), 50% culture supernatant from SW480 or Huh7 cell line, 5ng/mL CpG (Invivogen, tlrl-2006), 0.3ug/mL CD40 antibody (R&D, AF632-SP), 0.01 mM CEACAM5 (CEA) Protein (SINO BIOLOGICAL,11077-H08H), the combination of CpG and anti-CD40, or the combination of CpG, anti-CD40 and CEACAM5. After 20-hour culturing, the phenotype of B cells was examined by flow cytometry.

Flow cytometry analysis of tumor-associated atypical B cells isolated from tumor tissues

To identify TAABs *in vivo*, we collected fresh tumor samples from patients with HCC or ICC (Table S6). Tissues were dissociated with Tumor Dissociation Kit (Miltenyi Biotec, 130-095-929) for one hour in 37°C under continuous rotation using the gentleMACS™ Octo Dissociator with Heaters. Mononuclear cells were isolated by density gradient centrifugation with ficoll paque plus. The cell suspension was filtered through the 70μm mesh cell strainer (Falcon, 352350) to remove cell debris before being incubated with antibodies (Table S7) at 4°C for 30 minutes. B cell phenotypes were examined by flow cytometry (Beckman Coulter CytoFLEX).

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

Quantification analysis of whole-tissue scans using HALO

For the analysis of the spatial distribution of four B cell major lineages in tumors, we conducted mIHC staining with four panels: panel 3 (CD20+IgD), panel 4 (CD20+CD27+Bcl-6), panel 5 (CD79a+CD27+CD138), and panel 6 (CD20+CD4+CD8) (Table S2B). Based on the digital images of tumor sections stained with panel 6, we defined TLSs as organized structures with clear B cell and T cell zones.[11,18,19] The surface area of the identified TLSs ranged from 20,000 μm^2 to 1,000,000 μm^2. For less organized B cell aggregates without a well-defined T cell zone, considering the potential spatial variation depending on the plane of sectioning, we think it is possible the presence of the T cell zone was not captured in the sectioned plane. Thus, such B cell aggregates were regarded as ambiguous regions and excluded from downstream quantification analyses when comparing TLS and non-TLS regions. These two types of regions were then mapped across sections manually, facilitated by the registration module to fit one slide on the other (Halo 10 software).

We next identified and quantified Bn cells ($CD20^+IgD^+$), Bm cells ($CD20^+CD27^+Bcl\text{-}6^-$), Bgc cells ($CD20^+Bcl\text{-}6^+$) and ASCs ($CD79a^+CD27^+CD138^+$) across the whole-tissue scans leveraging the HighPlex FL v4.2.3 algorithm from HALO software v3.5.3577.285 (Indica Labs). In brief, after nuclei identification, a positive threshold was set for each individual marker and the combination of markers was employed to recognize target cells. The exact numbers of Bn cells, Bm cells, Bgc cells and ASCs residing in each type of region were then obtained. For a representative TLS from the lung cancer slide #LC-E05A5472, the Infiltration Analysis module was used to quantify the numbers of Bn cells, Bm cells, Bgc cells and ASCs within a sequential distance range from the TLS border, either inside or outside of the TLS (Figure 1F). The accuracy of region division, nuclei segmentation, and positive cell recognition was validated by two independent pathologists.

To explore the spatial proximity between TAABs and other immune cells, TAABs ($CD20^+FCRL4^+$), $CD4^+$ T Cells ($CD4^+CD68^-$), $CD8^+$ T Cells ($CD8^+$) and $CD68^+$ myeloid cells ($CD68^+$) were first identified throughout the whole tissue section. The Nearest Neighbor Analysis module was then used to define the nearest $CD4^+$ T, $CD8^+$ T and $CD68^+$ myeloid cell to each TAAB and quantify the physical distance between them, respectively.

To investigate the relationship between the activation status of $CD4^+$ T cells and their distance to TAABs, we utilized the Proximity Analysis module to quantify the numbers of activated $CD4^+$ T cells ($CD3^+CD4^+CD69^+$) and total $CD4^+$ T cells ($CD3^+CD4^+$) located within a specific distance range (0–20 μm; 50–200 μm) from TAABs ($CD20^+FCRL4^+$).

QUANTIFICATION AND STATISTICAL ANALYSIS

Single-cell RNA-seq data collection, preprocessing and the identification of B cells

The Cell Ranger Single-Cell Software Suite was employed to align the newly generated scRNA-seq data against the GRCh38 human reference genome, and the resulting unique molecular identifier (UMI) matrix was subsequently analyzed using Seurat (version 4.3.0).[132] Low-quality cells, defined as those with fewer than 400 detected genes or more than 20% mitochondrial UMI counts, were excluded.

Our newly generated pan-cancer scRNA-seq dataset consisted of 92 samples from 52 patients diagnosed with one of eight cancer types (Table S1B). We additionally collected publicly available datasets by conducting a search using the keyword (scRNA-seq AND cancer) on PubMed at the time of research. Subsequently, studies about B cell lymphoma were not included and datasets not derived from human cancer patients were excluded. In addition, we filtered out samples in which the composition of the B cell compartment was affected by the isolation or sorting strategies, such as FACS sorting $CD3^+$ cells. For published datasets, raw count matrix and metadata tables (including the tissue location, patient identifier, isolation or sorting strategy, tissue distance to tumor edge, tumor stage, treatment status, percentage of cancer cells, sex, age, sample type, tissue area, etc.) were obtained from the original publications. In cases where count data was not available, the CPM or TPM matrix was used as an alternative. The same quality control metrics were applied to these collected scRNA-seq datasets. Initially, a total of 1,622 samples from 1,034 patients were collected (Table S1C).

After quality control, following the standard protocol of Seurat, the count data were normalized using the *NormalizeData* function and the logarithm-transformed normalized matrix was used for downstream analyses. For the identification of B cells, we performed dimension reduction and unsupervised clustering with default parameters in Seurat for each dataset. Then, we coarsely annotated each cell cluster according to canonical cell markers, and identified $CD20^+$ B cells, ASCs, CD4 T cells, CD8 T cells, NK cells, monocytes/macrophages, dendritic cells, mast cells, endothelial cells, epithelial cells, and fibroblasts.

Data integration, dimension reduction, and unsupervised clustering

Considering the substantial expression of Ig genes in the transcriptome of B cells, particularly in PCs where the median proportion can reach up to 60%, we excluded the Ig genes to prevent the underrepresentation of non-Ig gene expression variabilities and subsequently normalized the expression data by logarithm transformation. For better data integration, we excluded samples with fewer than 50 B cells and datasets with no more than three samples.

We utilized the *SelectIntegrationFeatures* function in Seurat to identify the top 2000 highly variable genes that are repeatedly variable across datasets to avoid the selection of dataset-unique genes. The mitochondrial genes, ribosomal genes, mitochondrial ribosomal genes and heat-shock protein genes were further filtered out to eliminate unexpected noise. Then, principal component analysis (PCA) was performed on the variable gene matrix. Harmony was applied to remove the batch effects, and the top 30

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

components were used for downstream analyses. The *RunHarmony* function was applied to correct batch effects among datasets and among samples, with the parameters theta and lambda both set as one for these two variables. The *FindNeighbors* function of Seurat was employed to construct the Shared Nearest Neighbor Graph, based on which unsupervised clustering was performed using the *FindClusters function in Seurat*, with the parameter ''resolution = 0.2''. For visualization, the dimensionality was further reduced using Uniform Manifold Approximation and Projection (UMAP) implemented in the Seurat function *RunUMAP* with parameter ''umap.method = 'umap-learn'; dims = 1:30; reduction = 'harmony'''. Notably, we removed clusters expressing signatures of other major cell types and the above steps were repeated to ensure that we have filtered out these cell doublets. For any unspecified parameter, we used the default settings.

Following the first round of unsupervised clustering, we annotated each cell cluster according to the expression of canonical B cell markers, and identified the major clusters including naïve, memory, germinal center, cycling, and antibody-secreting B cells. The second-round clustering procedure was the same as the first-round clustering, both of which started from the logarithm-transformed matrix, and then identified the top 1000 highly variable genes, calculated PCA matrix and corrected batch effects using Harmony. For each specific B cell major cluster, the batch effect correction parameters used were as follows: Bn cells: batch_ID = "DatasetID'', harmony_theta = 3, harmony_lambda = 1; Bm: batch_ID = "DatasetID'', harmony_theta = 3, harmony_lambda = 1; proliferative B cells: batch_ID = "DatasetID'', harmony_theta = 2, harmony_lambda = 1; ASCs: batch_ID = "DatasetID'', harmony_theta = 0, harmony_lambda = 1.

We then detected cell clusters and performed dimension reduction for visualization. For each major cluster, the resolution parameter used was as follows: Bn cells: 0.1, Bm cells: 0.5, cycling B cells: 0.1, ASC: 0.6. The selection of the resolution used for major clusters was guided by the ROGUE statistic.[131] Specifically, in the case of ASCs, we examined cluster purity across different resolutions and observed that the ROGUE value reached saturation at a resolution of 0.6 (Figure S2A). Notably, for ASCs, a full investigation of underlying biological or clinical indications of 16 subclusters remains challenging, and for the mature PCs, our primary focus in this study was to understand the differences across PC subsets with different Ig isotypes. To accomplish this, we leveraged the transcriptional phenotypes of those cells with available BCR sequences and a reliable tool CellTypist[41] to predict the isotype of the remaining cells. Then, clusters were aggregated into the IgA and IgG meta-clusters based on the isotype majority of their cells, yielding the c16_PC_IGHG and c17_PC_IGHA clusters. To accurately define the cell types and cellular states, differentially expressed genes were identified for each cell cluster using the *FindAllMarkers* function in Seurat (Table S3).

Scoring cells using gene expression signatures

Signature sets of ABCs were collected from studies under various chronic disease settings, including malaria,[28] HIV infection,[133] systemic lupus erythematosus (SLE),[134] Sjögren's syndrome,[135] rheumatoid arthritis (RA) and common variable immunodeficiency (CVID).[136] The pre-GC signature, as well as the list of genes mechanistically linked with CSR, was from a study of human tonsillar B cells.[30] Other gene signatures used to characterize B cells in Figure 1C were obtained from MSigDB (R package msigdbr, version 7.5.1). The compiled sources of all used gene signatures were deposited in Table S4. The *AddModuleScore* function of Seurat was applied with default parameters to score each signature in each B cell, and the mean score of all cells in each B cell subset was calculated.

Reconstruction and analysis of the BCR repertoires of TIBs

For 59,592 B cells with raw sequencing data available in our collected scRNA-seq data, we *in silico* reconstructed their BCR sequences using TRUST4.[25] For the assembled IgH sequence of each cell, the C, V, D, J gene and CDR3 sequence was next annotated by the *AssignGenes.py* from the Change-O toolkit,[126] with the alignment performed by IgBLAST.[127] We excluded out-of-frame or partial sequences to keep only productive sequences for downstream analyses. To cluster BCR sequences into clonal groups, according to the standard workflow of Change-O, we first automatically determined the clustering threshold via a gamma/gamma mixture model of the nearest neighbor distances by the *findThreshold* function in the R package SHazaM.[126] The *DefineClones.py* from Change-O was then applied to group similar IgH sequences into clonal groups based on IGHV alleles, IGHJ alleles, junction length and junction distance (below the threshold 0.144 calculated by SHazaM) in each patient. Germline sequences were reconstructed by *CreateGermlines.py* from Change-O. Cells from clonal groups with at least two cells were defined as clonally expanded. The SHM level of each B cell was quantified as the point mutation rate in the IGHV sequence, and classified into three categories, low-SHM (< 1%), median-SHM (1%–5%), and high-SHM (> 5%). Of note, the SHM levels were only calculated for cells from the 10x 5' platform, as the short length of IGHV sequences assembled from the 10x 3' platform may cause inaccurate estimation of SHM levels.

For the construction of the lineage trees of TIB clonal groups containing Bgc cells (Figures 3D and S4D), we first used the function *msaConsensusSequence* from the R package msa[128] to obtain a consensus CDR3 nucleotide acid sequence for each clonal group as the root sequence in the tree building. Next, lineage trees were constructed via maximum parsimony employing the *buildPhylipLineage* function from the R package Alakazam.[126] We annotated the cell number and the amino acid sequence of CDR3 for each node, as well as the point mutation count for each edge. The cell type composition of each node was depicted by a pie chart. For comparison of the CSR events of TIBs between tumor and ANTs (Figure 4L), we ordered the Ig subclasses according to their genomic coordinates (5′ to 3′) and for each combination of two isotypes, calculated the fraction of BCR clones belonging to an upstream isotype that are shared with a downstream isotype to represent the frequency of class switching between them. For the

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

exploration of the developmental dynamics of TAABs within tumors (Figure 6G), we evaluated the transition tendency between TAABs and non-ASC TIB subsets using the pTrans index derived from STARTRAC analysis.[58]

Assessment of infiltration status for major immune components in tumors

For tumor samples containing a minimum of 100 cells, either without prior FACS treatment or where the FACS process did not impact the composition of the CD45 compartment, we established the infiltration status of three immune components—B cells, T cells, and myeloid cells. If one immune component in a tumor encompassed more than 20 cells, we considered its infiltration status as "True".

Cellular interaction analysis by CellChat

We utilized CellChat[37] to investigate TIB-involved cellular interactions. The workflow started with inputting pre-processed expression profiles from the data slot of the Seurat object with corresponding annotations to create a CellChat object. CellChatDB.human was set as the ligand-receptor interaction database. The expression data then underwent default preprocessing in CellChat.

Especially, for the investigation of the role of TIBs in the chemokine signaling network within the TME (Figure S3D), we used the coarse-grained annotations of all datasets in our scRNA-seq atlas, including CD20$^+$ B cells, ASCs, CD4 T cells, CD8 T cells, NK cells, monocytes/macrophages, dendritic cells, mast cells, endothelial cells, epithelial cells and fibroblasts. Tumor samples that contained at least 100 cells and did not undergo FACS were retained. For computational consideration, we randomly sampled 5,000 cells for each cell type as the input of CellChat. We aggregated the CXCL, CCL, XCR and CX3C pathways in CellChat into one chemokine signaling pathway. We further computed the network centrality score and identified the signaling role of each cell type in the chemokine signaling pathway. To uncover consistent chemokine–chemokine receptor interactions across cancer types (Figure S3E), we further performed CellChat analysis in each cancer type separately, with randomly sampled 3,000 cells for each cell type in each cancer type as the input.

In addition, for comparison of the communication between CD4 T cells and each TIB subset (Figures 7F and S7C), we constructed a dataset cohort with fine-grained CD45$^+$ cell annotations by incorporating the subset-level annotations of T cells, NK cells and myeloid cells from our previous pan-cancer single-cell atlases,[2,3,44] retaining only the datasets intersecting with our atlas. We performed the CellChat analysis according to the standard procedure, using only cells derived from tumor samples. Of note, this dataset cohort was also used when examining the expression of Fc receptors (Figure S5L) or *IL21* (Figure 7J) in the CD45$^+$ compartment, and when calculating the abundance correlation of Bgc cells with other immune subtypes (Figures 3G and S4F).

Tissue distribution preference of TIB subsets

To assess the tissue distribution preference of TIB subsets, we calculated the ratio of observed to expected cell numbers (Ro/e) for each annotated cluster in different tissues, where the expected cell numbers of clusters in a given tissue were derived from the chi-square test.[58] An Ro/e value greater than one indicates the enrichment of a cluster in a specific tissue.

Cell type assignment of activated cycling Bm cells by CellTypist

For the cell type assignment of activated cycling Bm cells (c14), we employed CellTypist,[41] a semi-automatic cell-type classification and annotation tool. In brief, we used the non-cycling Bm cells (c04–c11) in our atlas as the training set to generate a custom model, which was then applied to assign a cell type label for each c14_Bm_activated-cycling cell. The resulting labels were retrieved from the "predicted_labels" column in the output.

Evaluation of the transcriptional heterogeneity of B cell clusters

The transcriptional heterogeneity of TIB clusters across cancer types was calculated as follows:

1. For each B cell subset (S) in each cancer type (T_i), calculate the average expression profile of tumor-infiltrating B cells using the *AverageExpression* function of Seurat. Two kinds of averaged expression profiles were obtained, with one of them considering all genes ($\overline{X}^{All}_{S,T_i}$) and the other excluding immunoglobulin genes ($\overline{X}^{NoIg}_{S,T_i}$).
2. For each B cell subset (S), calculate the Pearson correlation coefficient of their average expression profiles between each pair of cancer types (T_i and T_j):

$$r^{All}_{S,T_i,T_j} = \mathrm{cor}\left(\overline{X}^{All}_{S,T_i}, \overline{X}^{All}_{S,T_j}\right)$$

$$r^{NoIg}_{S,T_i,T_j} = \mathrm{cor}\left(\overline{X}^{NoIg}_{S,T_i}, \overline{X}^{NoIg}_{S,T_j}\right)$$

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

3. For each B cell subset (S), define the heterogeneity (H) between each pair of cancer types (T_i and T_j) as

$$H^{All}_{S,T_i,T_j} = 1 - r^{All}_{S,T_i,T_j}$$

$$H^{NoIg}_{S,T_i,T_j} = 1 - r^{NoIg}_{S,T_i,T_j}$$

4. For each B cell subset (S), calculate the overall heterogeneity across cancer types as the average heterogeneity between all pairs of cancer types:

$$\overline{H}^{All}_{S} = \frac{1}{\binom{n}{2}} \sum_{T_i \in TT_j \in T, i \neq j} \sum H^{All}_{S,T_i,T_j}$$

$$\overline{H}^{NoIg}_{S} = \frac{1}{\binom{n}{2}} \sum_{T_i \in TT_j \in T, i \neq j} \sum H^{NoIg}_{S,T_i,T_j}$$

in which n is the total number of cancer types.

TCGA data analysis

The TCGA Toil re-computed expression data and patient metadata were downloaded from the TCGA Pan-Cancer cohort on the UCSC Xena website (https://xenabrowser.net/). We kept only samples from primary tumors. To reflect the relative abundance of a certain cell type, we calculated the enrichment score of the top 20 differentially expressed genes of this cell type among the immune compartment. To mitigate the impact of cell cycle, genes satisfying at least one of the following criteria were regarded as cell cycle-related and excluded from each signature: (1) being present in the "KEGG_CELL_CYCLE" pathway from the MSigDB database (R package msigdbr version 7.5.1); (2) being highly expressed (average $\log_2$ fold change > 0.5 and BH-adjusted P value < 0.05) in all three cycling B cell subsets (c13–c15) when compared with non-cycling B cells. The scoring was performed by AUCell[54] in each cancer type separately.

To model the predictive value of TAABs on the ten-year overall survival of cancer patients (Figure 5B), in each cancer type, we first stratified the patients into high and low groups based on the median AUCell enrichment score of the TAAB signature. To correct for the prognostic impact of the B cell abundance and probe how the presence of TAABs refined the prognosis of B cells, we additionally obtained the total B cell abundance for each TCGA tumor sample as described in Thorsson et al.[137] We next built a Cox proportional-hazards model to investigate the survival difference between these two groups, with age, sex, tumor stage and the total B cell abundance as covariates. The model provides a hazard ratio (HR) for each predictor variable, representing the prognostic value, with HR less than 1 indicating the variable associated with better survival, and greater than 1 representing an increased risk of death related to the variable. Then the per-cancer-type models were combined into a pan-cancer model by meta-analysis using a random effect model implemented in the R package meta. Kaplan–Meier curves were plotted to visualize the survival difference using the R package survival and survminer. When examining the prognostic value of the c06_Bm_stress-response signature (Figure 5C), similar procedures were conducted.

For the comparison of prognostic values among all B cell subsets (Figure 5A), we constructed a pan-cancer Cox proportional-hazards model for each B cell subset, employing the same procedures as for TAABs. The B cell subset with the lowest pan-cancer HR was recognized as possessing the strongest association with favorable prognosis.

Construction of a SMART-seq2-based single-cell transcriptome atlas of B cells

We assembled SMART-seq2-generated scRNA-seq datasets of B cells from eight published studies (Table S5A). B cells were isolated based on literature-provided annotations, resulting in a single-cell cohort of 5,995 B cells across six cancer types. We then integrated these datasets and performed unsupervised clustering following the same workflow employed in the construction of our pan-cancer single-cell B cell atlas. Similarly, five major clusters, including Bn cells, Bm cells, Bgc cells, ASCs, and cycling B cells, were discriminated based on the high expression of their canonical markers (Figures 5D and 5E).

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

Analysis of scATAC-seq data

The processed scATAC-seq data of immune cells from five pre-treatment and post-treatment BRCA tumor samples were obtained from Zhang et al.[51] Downstream analyses were performed according to Signac (version 1.13.0),[130] an extension to the Seurat package. The peak count matrix was first filtered, during which only peaks detected in at least 10 cells and cells with at least 200 peaks were retained. The GRCh38 gene annotations were subsequently added for each peak. A series of additional quality control metrics were applied to obtain high-quality cells with (1) a nucleosome banding signal < 4, (2) transcriptional start site enrichment score > 2, (3) > 3,000 and < 20,000 fragments in peaks, (4) > 15% fragments in peaks, and (5) < 5% of reads in genomic blacklist regions. The peak count matrix after QC then underwent normalization and dimension reduction as described in the Signac workflow.

In order to annotate cell clusters identified in scATAC-seq data, we calculated a gene activity matrix with the *GeneActivity* function. Unsupervised clustering based on the peak matrix was next conducted, identifying B cell clusters according to the high gene activity of canonical B cell marker genes (*CD79A*, *CD79B*, *CD19*, *MS4A1*, *MZB1* and *XBP1*). To obtain B cell subtype annotations for the B cell clusters, we performed cross-modality integration and label transfer following the procedure of Seurat.[130] We next calculated the confusion matrix based on the predicted labels, and annotated each cluster as the label with the highest occurrence, yielding five B cell subsets, Bn cells, Bgc cells, FCRL4$^-$ Bm cells, FCRL4$^+$ Bm cells and ASCs. The cellular identities of each cluster were further corroborated with their marker gene activities. The DNA accessibility of FCRL4$^+$ Bm marker gene regions, including the gene body and 2 kb upstream/downstream regions, was visualized by the *CoveragePlot* function from Signac.

SCENIC regulon analysis

Activated regulons in each Bm subset were analyzed using SCENIC[54] with default settings. The regulon specificity score (Figure 6E) was calculated by the *calcRSS* function in the R package SCENIC.

RNA velocity analysis and pseudo-time trajectory inference

The RNA velocity analysis and pseudo-time trajectory inference were both performed on our newly generated 10x scRNA-seq data. First, the diffusion map algorithm[138] was applied to model the cell state transition and infer the differentiation trajectory. We input the scaled expression matrix and principal component matrix from the integrated Seurat object into the Scanpy pipeline.[125] A neighborhood graph was constructed based on the top 10 principal components using the *scanpy.pp.neighbors* function. We next applied the *scanpy.tl.diffmap* function to build the diffusion map, and the first two diffusion components were used for visualization.

For RNA velocity analysis, the spliced and unspliced UMIs for each gene in each cell were counted using the Python package velocyto.[59] The subsequent analyses were performed by Scanpy, scVelo[60] and CellRank.[61] Specifically, the count matrices were filtered to retain only genes detectable in over 20 cells for both spliced and unspliced matrices and normalized by the library size. The top 200 genes exhibiting the highest variability were selected for downstream analyses. PCA was performed on the log-transformed spliced matrix and a k nearest-neighbor graph (k=30) was built using the top 30 principal components. For each cell, we computed the moments (means and uncentered variances) of normalized spliced/unspliced counts using the 30 nearest neighbors by the *scv.pp.moments* function. These moments facilitated the RNA velocity estimation implemented in the *scv.tl.velocity* function, with the mode set to "dynamical". Based on the estimated velocities, a velocity graph representing the transition probabilities among cells was constructed by the *scvelo.tl.velocity_graph* function. The velocity graph was further projected onto the diffusion map and visualized as arrows upon a grid by the *scv.pl.velocity_embedding_grid* function. Finally, cells in the initial and terminal states were inferred by the *cr.tl.initial_states* and *cr.tl.terminal_states* functions from CellRank respectively, using the *cr.tl.estimators.CFLARE* estimator with default settings.

The RNA velocity and CellRank analysis guided us to select the cell with the maximum value of the first diffusion component as the root cell for the computation of pseudo-time. The diffusion pseudo-time[139] was calculated using the *scanpy.tl.dpt* function. The original diffusion pseudo-time values were further converted to percentile rank values as described in Cao et al.[140] To find the potential genes driving the differentiation process, we fitted a generalized additive model (the *gam* function in the R package gam) to explain the expression of each gene with the pseudo-time.

Construction and analyses of an integrative scRNA-seq B cell atlas under non-cancer conditions

We assembled scRNA-seq datasets of B cells from 16 published studies (Table S5B), encompassing blood or tissue samples from a wide spectrum of non-cancer conditions, including health,[30,41] autoimmune diseases—SLE,[62–65] RA,[66] psoriatic arthritis (PA)[67] and systemic sclerosis-associated interstitial lung disease (SSc-ILD)[68,69]—as well as HIV infection,[28,70,71] malaria infection,[28] SARS-CoV-2 infection[72] and influenza vaccination.[29,73] Following the same procedures employed in the construction of our pan-cancer single-cell B cell atlas, we obtained a high-quality single-cell atlas, comprising 370,451 B cells from 476 samples of 390 donors. We integrated these datasets and performed unsupervised clustering mirroring the workflow of our pan-cancer atlas construction. Seven B cell subclusters were identified, including Bn cells, ISG$^+$ Bn cells, Bm cells, Bgc cells, ASCs, cycling B cells, as well as ABCs.

To compare the expression of ABC markers (*FCRL4*, *FCRL5*, *ITGAX* and *TBX21*) across different conditions, for each condition, atypical-like B cells and other Bm cells were first selected as two populations. Specifically, for tumor tissues, TAABs and other Bm cells were selected from those treatment-naïve tumors from all cancer types in our atlas. Next, the Seurat function *FoldChange* was used to calculate the average log fold change of gene expression between the two cell populations, and the resulting log fold changes

Please cite this article in press as: Yang et al., Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes, Cell (2024), https://doi.org/10.1016/j.cell.2024.06.038

for all conditions were visualized as a bar plot. For the comparison of *FCRL4* expression across different cancer types, similar procedures were conducted for each cancer type (with at least five TAABs).

GSEA of cell surface receptor signaling pathways

We collected the gene sets of all child terms under the gene ontology term "cell surface receptor signaling pathway" but not under "cytokine-mediated signaling pathway" from the gene ontology database (R package org.Hs.eg.db, version 3.14.0). Based on these gene sets, we performed GSEA of tumor-infiltrating TAABs compared with other Bm cells within tumor tissues.

CytoSig analysis of cytokine signaling activity within cells

For inspection of the cytokine signaling activity within intratumoral Bm cells, we fed their count expression matrix into the CytoSig prediction model.[86] As the quality control was performed when constructing the atlas, we set the parameter minimum_read_count as 0 and max_dropout_ratio as 1. A scaled activity score was obtained for each curated cytokine in each cell. For each Bm subset, the average activity score for each cytokine was calculated and ranked.

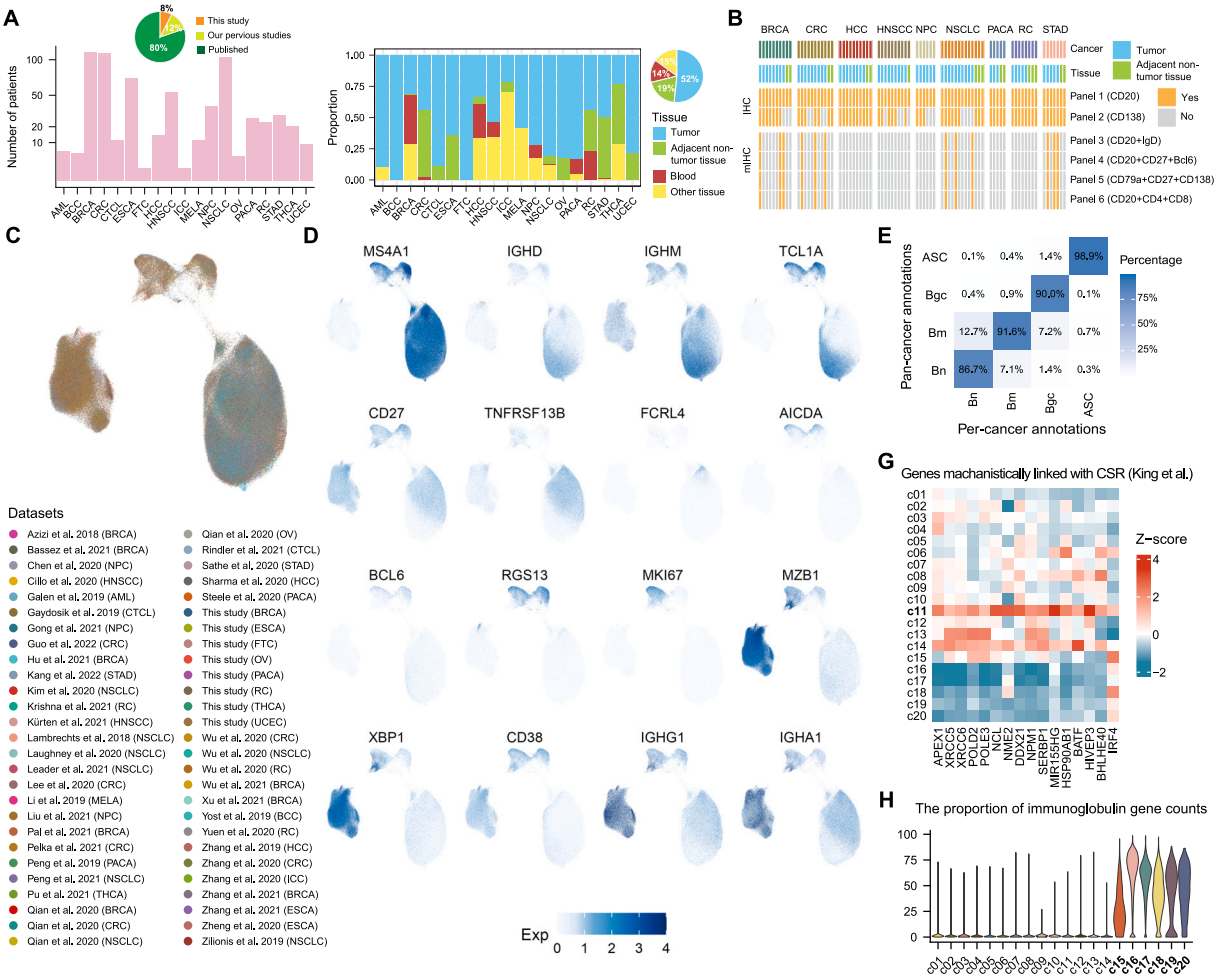
Analysis of immunotherapy datasets

For the exploration of whether the expression of the TAAB signature was associated with immunotherapy response, we collected seven bulk RNA-seq datasets of cancer patients receiving anti-PD-1, anti-CTLA-4 or adoptive cell transfer therapy.[115–121] We kept only samples from primary tumors before immunotherapy. To resolve the confusion in the nomenclature of response evaluation from different studies, we broadly defined CR/PR patients as responders (R), and SD/PD patients as non-responders (NR). The enrichment score of the TAAB signature was calculated as aforementioned in each sample and compared between samples from R and NR.

Cell
Resource

CellPress
OPEN ACCESS

Supplemental figures



(legend on next page)

Figure S1. Basic information of the atlas and B cell integration, related to Figure 1

(A) Bar plot showing the number of patients collected for each cancer type, and pie chart showing the proportion of patients from different data sources (left). The composition of tissue sources among B cells across cancer types (right, bar plot) or at the pan-cancer level (right, pie chart).

(B) IHC or mIHC staining panels (bottom) applied to each tissue specimen (top).

(C) UMAP plot showing the distribution of datasets in the integrated B cell atlas.

(D) UMAP plots showing the expression patterns of major cluster marker genes in the integrated B cell atlas.

(E) Heatmap showing the consistency of annotation for four B cell major lineages between individual cancer types and the integrated atlas.

(F) Expression of representative signature genes across B cell subsets. IL35 expression was defined as the concurrent expression of its subunits *IL12A* and *EBI3*.

(G) Heatmap showing the expression of genes mechanistically linked with CSR.

(H) Proportions of immunoglobulin gene count across B cell subsets.

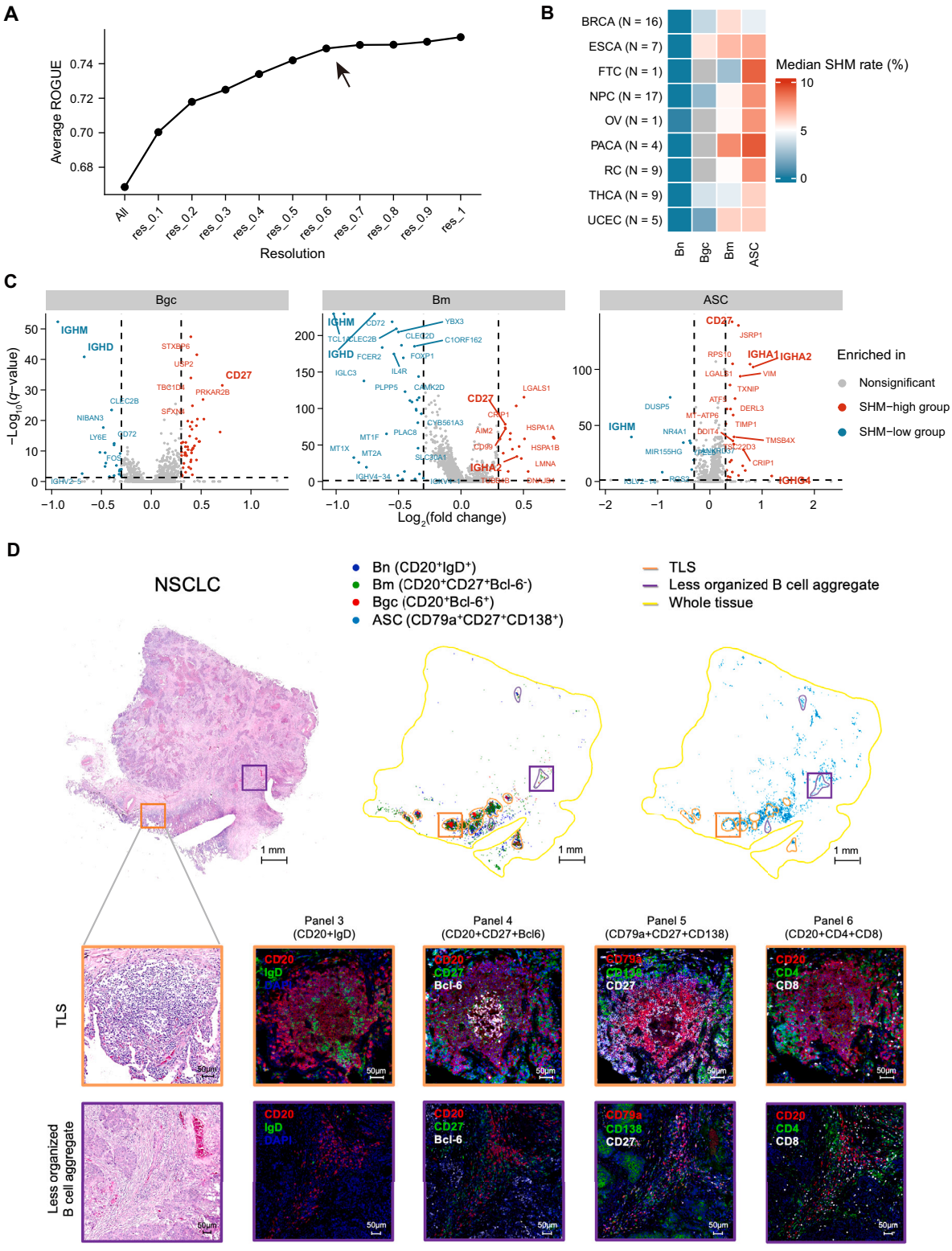


Figure S2. Clonotypic and spatial characteristics of B cell major lineages, related to Figure 1

(A) The cluster purity (assessed by average ROGUE value) of ASCs under different resolutions.

(B) Heatmap showing the median SHM rate for each TIB major lineage across cancer types. The sample number for each cancer is annotated.

(legend continued on next page)

(C) Differentially expressed genes between the SHM-high (red) and SHM-low (blue) groups (stratified by the median level) within Bgc cells, Bm cells, and ASCs. Each red or blue point indicates a significant gene with BH-adjusted p value < 0.05, two-sided unpaired Wilcoxon test, and $|\log_2(\text{fold change})| > 0.3$.
(D) Hematoxylin-eosin staining of an NSCLC tumor sample (top left), with the representative views of a TLS and a less organized B cell aggregate shown at higher magnification (orange line and purple line, respectively, bottom left). mIHC staining of four separate panels showing the distribution of B cell major lineages in these two views (bottom right). Serial sections were used for the staining. The spatial distribution of four B cell major lineages in serial sections (top right, panel-CD20 + IgD and panel-CD20 + Bcl-6 + CD27 were merged by HALO software).

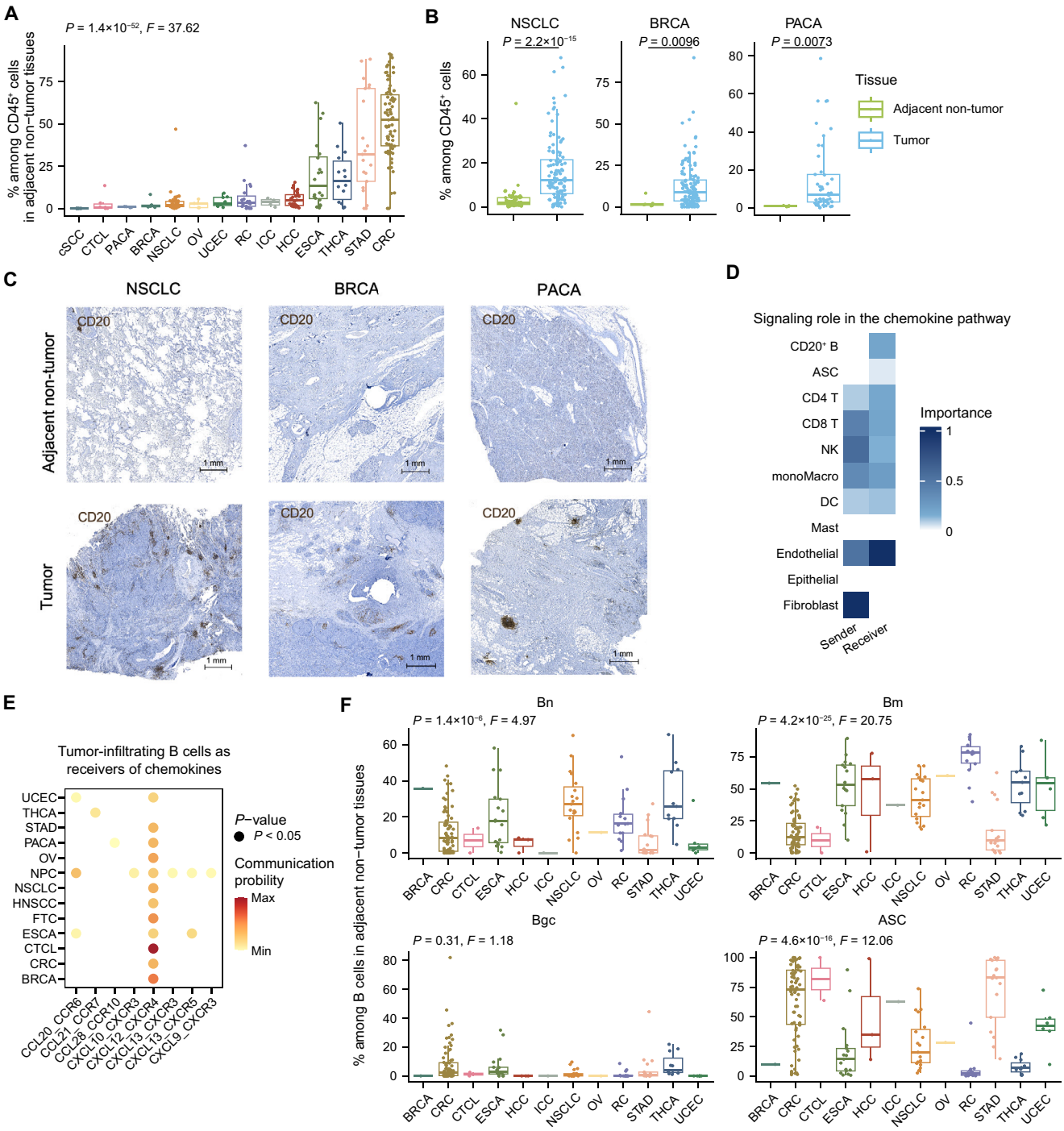


Figure S3. Heterogeneity of B cells across cancer types, related to Figure 2

(A) Proportions of B cells in CD45$^+$ cells from ANTs across cancer types. One-way ANOVA test.

(B) Boxplots comparing the proportions of B cells in CD45$^+$ cells between tumors and ANTs in NSCLC, BRCA, and PACA. Two-side unpaired Wilcoxon test.

(C) IHC staining of CD20 to compare CD20$^+$ B cell abundances between tumors and ANTs in NSCLC, BRCA, and PACA.

(D) CellChat-predicted signaling roles of TME immune components in the chemokine signaling network.

(E) Significant chemokine–chemokine receptor pairs between TIBs and other TME immune subtypes across cancer types. Only pairs significant in at least one cancer type, along with cancer types featuring at least one significant pair, are shown.

(F) B cell major lineage compositions in ANTs across cancer types. Only samples with B cells > 50 are shown. One-way ANOVA test.

Only samples with CD45$^+$ cells > 100 and an immune compartment unaffected by FACS are shown in (A) and (B). All samples are from treatment-naive patients.

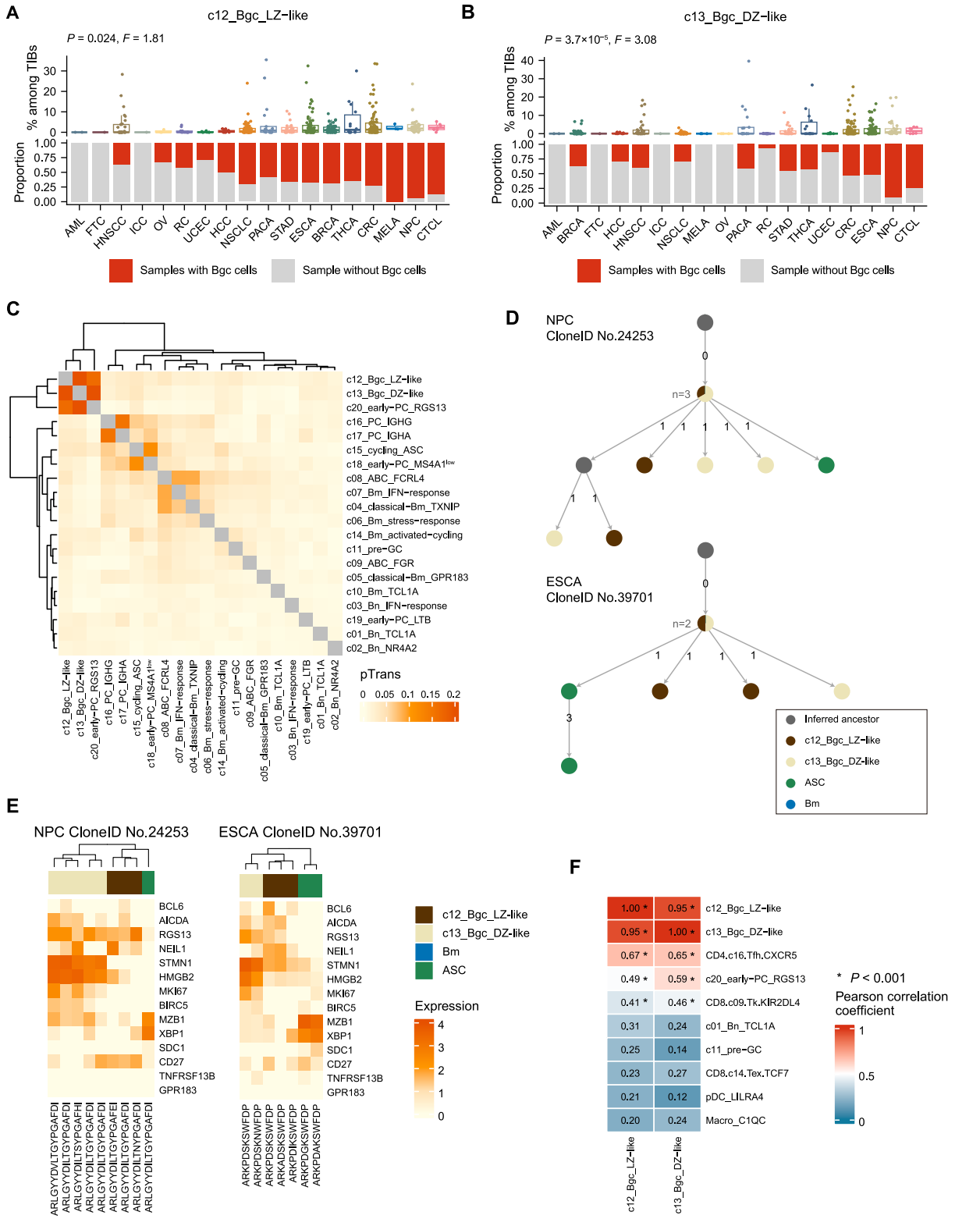


Figure S4. Characteristics of germinal center B cell subsets in tumors, related to Figure 3

(A and B) Boxplots showing the proportions of c12 (A) and c13 (B) Bgc cells among TIBs across cancer types. Only treatment-naive tumor samples with B cells > 50 are shown. One-way ANOVA test.

(legend continued on next page)

(C) Heatmap showing the pTrans index for every combination of TIB subsets.
(D) Two representative lineage trees of TIBs from an NPC tumor (top) and an ESCA tumor (bottom).
(E) Heatmaps showing B cell major lineage marker expression among B cells from the lineage trees in (D).
(F) Heatmap showing the Pearson correlation between the frequencies of cell clusters in tumors.

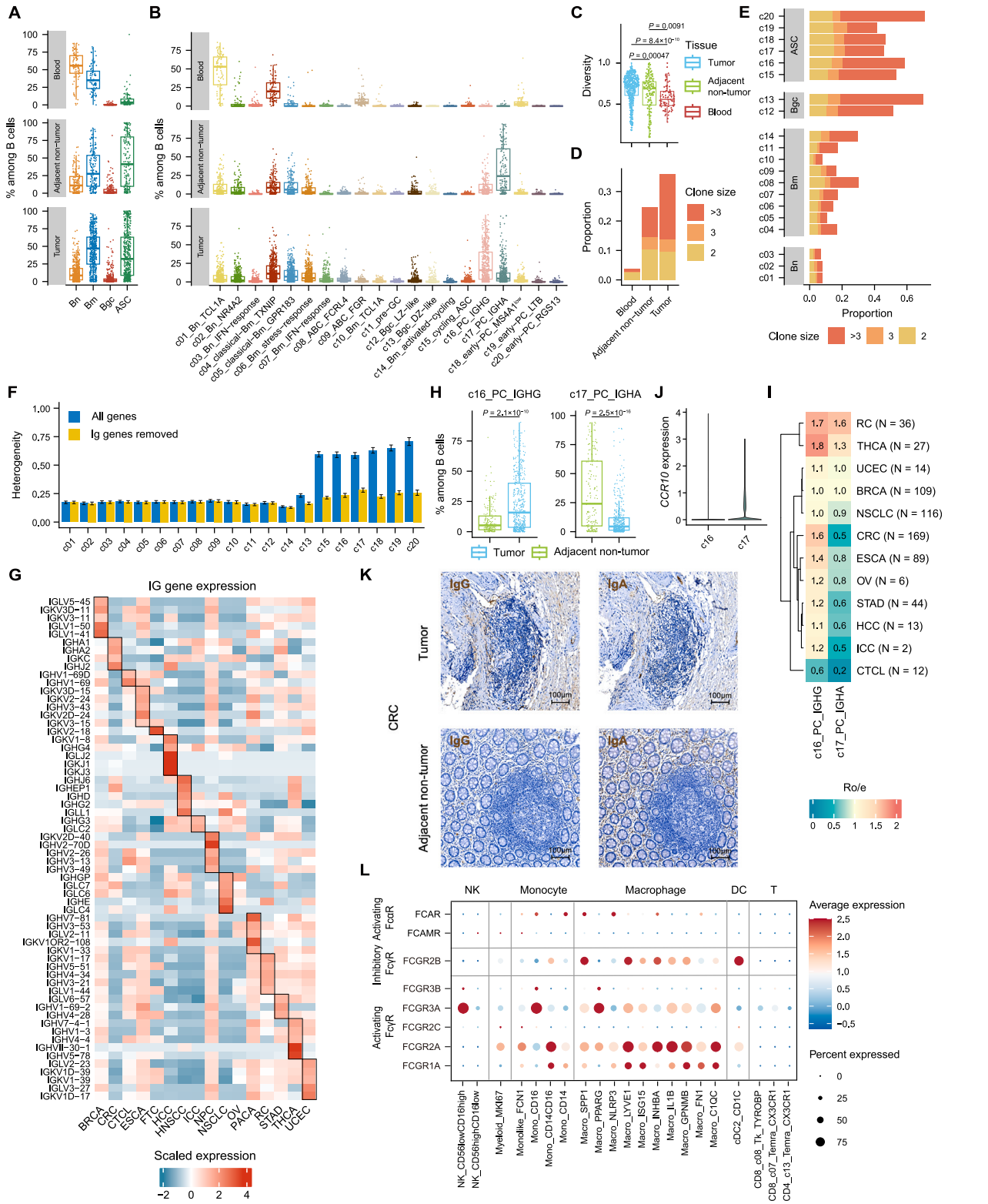


Figure S5. Identification of potential B cell subsets associated with immune responses in tumors, related to Figure 4

(A) B cell major lineage compositions in the blood, ANTs, and tumors.

(B) B cell subset compositions in the blood, ANTs, and tumors.

(legend continued on next page)

(C) Boxplot comparing the compositional diversity of B cells in the blood, ANTs, and tumors, measured by Shannon equitability index.[141] Two-sided unpaired Wilcoxon test.
(D) Clonal expansion levels of B cells in the blood, ANTs, and tumors, with cells categorized by the clone size of their corresponding clones.
(E) Clonal expansion levels of TIB subsets, with cells categorized by the clone size of their corresponding clones.
(F) Transcriptional heterogeneity of TIB subsets across cancer types. Data are represented as mean ± SEM.
(G) Differentially expressed Ig genes among tumor-infiltrating ASCs across cancer types. For the ASCs of each cancer type, Ig genes that met the following criteria were first selected: (1) significant differential expression (BH-adjusted p value < 0.05); (2) higher expression than the ASCs of any other cancer type. For each cancer type, the top five Ig genes with the highest average log fold change are shown. If less than five Ig genes satisfy the demands, all qualifying genes are shown.
(H) Boxplots comparing the proportions of c16 and c17 PCs in total B cells between tumors and ANTs. Only treatment-naive samples with B cells > 50 are shown. Two-sided unpaired Wilcoxon test.
(I) Tissue preference of c16 and c17 PCs across cancer types evaluated by the Ro/e index. The sample number for each cancer is annotated.
(J) *CCR10* expression in c16 and c17 PCs.
(K) IHC staining of IgG and IgA showing the abundances of IgG and IgA PCs in a CRC tumor and the paired ANT.
(L) Expression of Fc receptors in TME immune subtypes, excluding TIB subsets. Clusters in which less than 15% of cells expressing any of the eight Fc receptors are not shown.

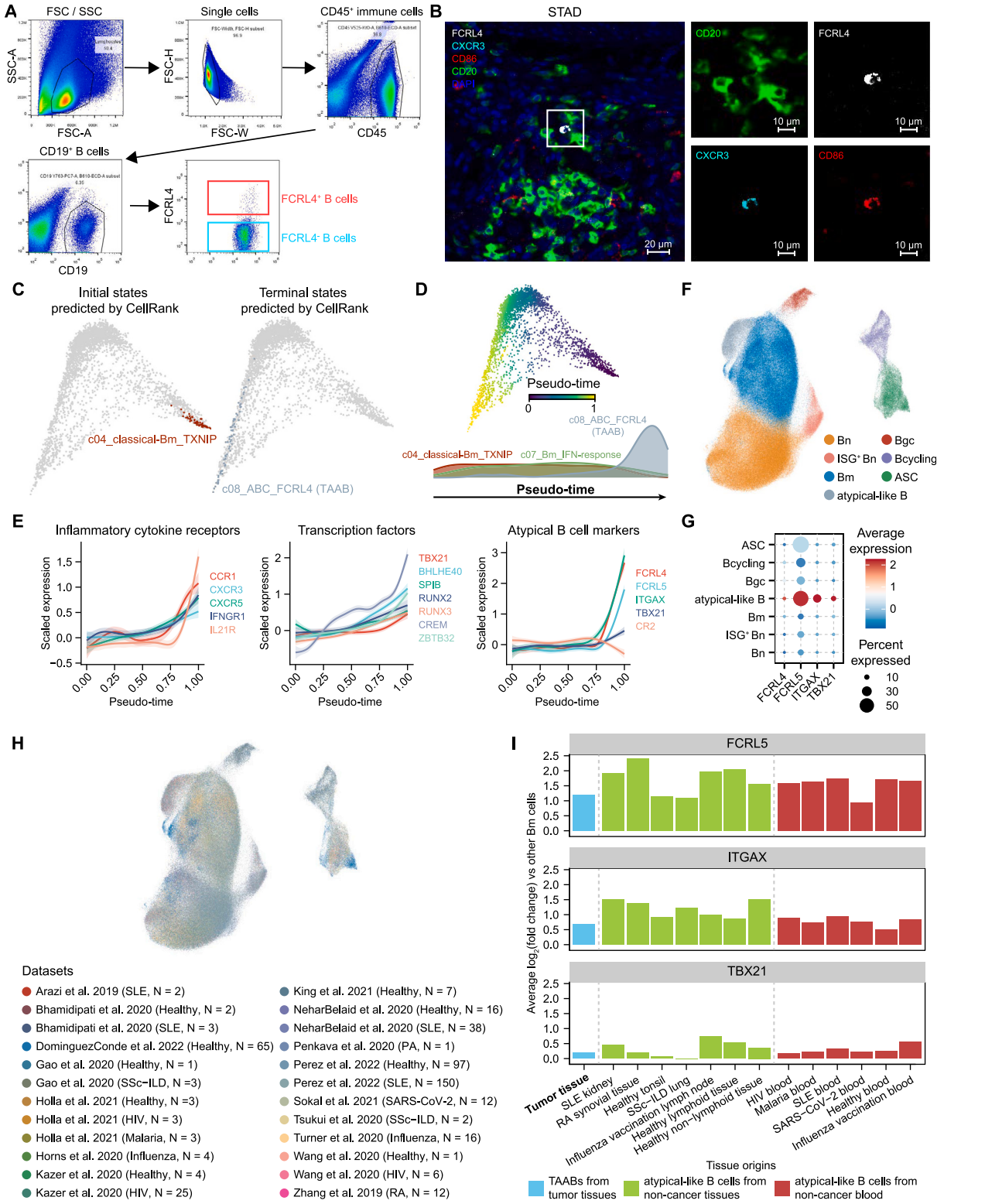


Figure S6. Characteristics of TAABs, related to Figure 6

(A) The gating strategy of flow cytometry for FCRL4^{+} and FCRL4^{-} B cells from tumors.

(B) CODEX staining in a STAD tumor showing CXCR3 and CD86 expression in FCRL4^{+} B cells.

(C) CellRank-predicted cells in developmentally initial (left) and terminal (right) states visualized on the diffusion map of intratumoral c04, c07, and c08 cells.

(legend continued on next page)

(D) Pseudo-time visualized on the diffusion map of intratumoral c04, c07, and c08 cells (top). Density curves of these three subsets along the pseudo-time (bottom).
(E) The relationship between selected gene expression and pseudo-time with a 95% confidence interval.
(F–H) Overview of the integrated scRNA-seq B cell atlas under non-cancer conditions, with UMAP plots showing the distribution of identified B cell subclusters (F) and datasets (H), as well as a bubble heatmap showing the expression of ABC signature genes across B cell subclusters (G). In (H), the sample number for each dataset is annotated.
(I) Average $\log_2$ fold change of ABC signature gene expression between ABCs and other Bm cells. TAABs from the entire TIB atlas along with ABCs from each non-cancer condition were examined.

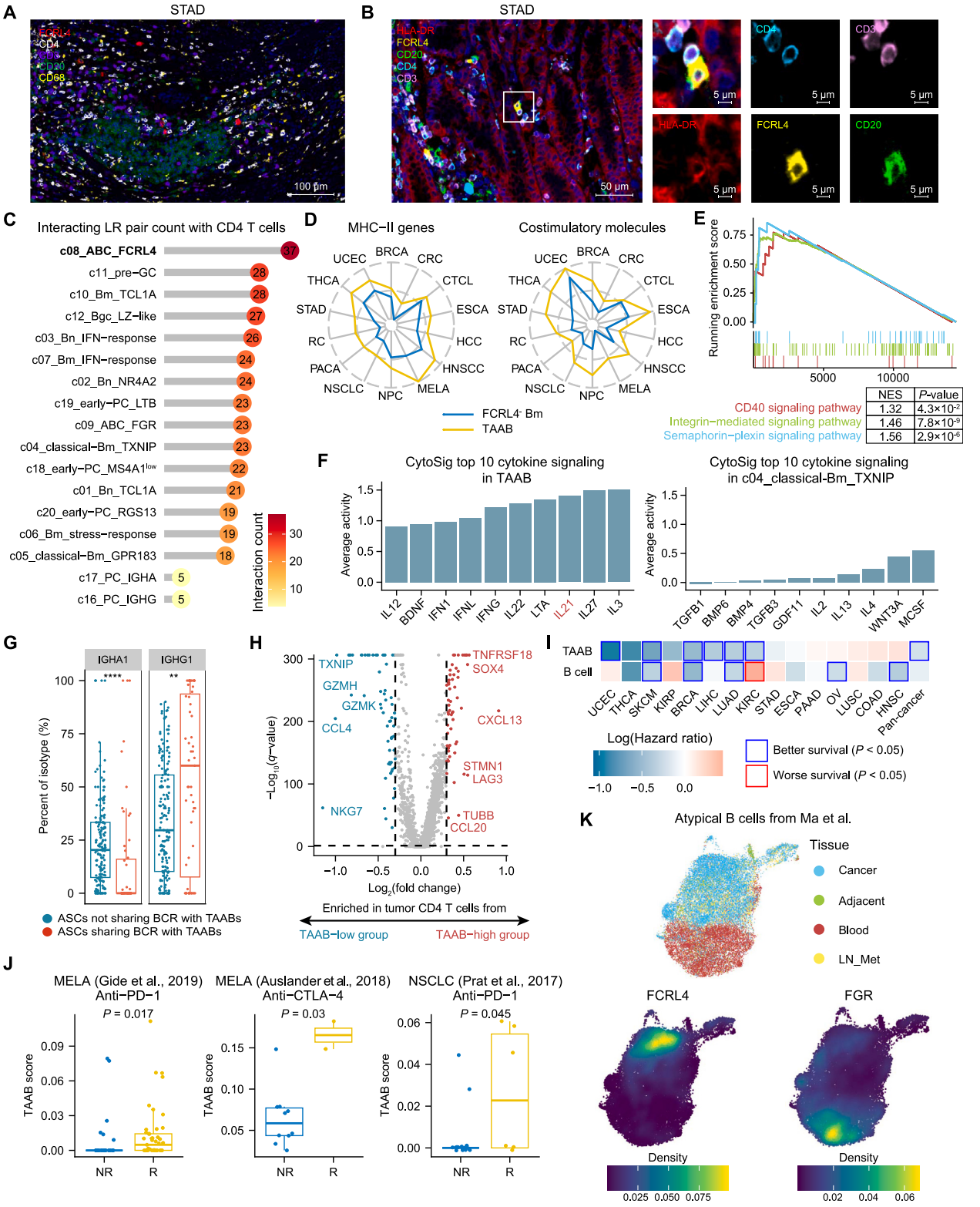


Figure S7. Cellular communication and clinical associations of TAABs, related to Figure 7

(A) mIHC of a STAD tumor to show the proximity relationship between TAABs ($CD20^+FCRL4^+$) and $CD4^+$ T cells ($CD4^+CD68^-$), $CD8^+$ T cells ($CD8^+$), or $CD68^+$ myeloid cells ($CD68^+$).

(B) mIHC of a STAD tumor to show the juxtaposition between TAABs and $CD4^+$ T cells with HLA-DR expression at their junction.

(legend continued on next page)

(C) CellChat-predicted significant ligand-receptor pair numbers between CD4 T cells and each TIB subset.
(D) Radar charts showing the scaled overall expression level of MHC-II genes (left) and costimulatory molecules (right) in intratumoral FCRL4^{-} Bm cells and TAABs across cancer types. The average AUCell score is shown and scaled to 0–1.
(E) GSEA enrichment plot for representative cell surface signaling pathways enriched in TAABs compared with other Bm cells within tumors. NES, normalized enrichment score. *p* values were determined by a one-tailed permutation test by GSEA.
(F) Top 10 CytoSig-predicted cytokines with the highest signaling activity in TAABs (left) and c04_classical-Bm_TXNIP (right) within tumors.
(G) Boxplots comparing the percentage of IGHA1 and IGHG1 isotypes between ASCs sharing BCR with TAABs and other ASCs. $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$, ns, $p \geq 0.05$.
(H) Differentially expressed genes of CD4 T cells in tumors with high (red) and low (blue) TAAB abundances, stratified by the median level. Each red or blue point indicates a significant gene with BH-adjusted *p* value < 0.05 and $|\log_2(\text{fold change})| > 0.3$.
(I) Heatmap showing the hazard ratios of the TAAB signature and the total B cell abundance across cancer types.
(J) Comparison of the TAAB signature expression between responders and non-responders of cancer immunotherapy.
(K) UMAP plots showing the tissue distribution (top) and expression patterns of *FCRL4* and *FGR* (bottom) in the ABCs (B.09.DUSP4+AtM) from the scRNA-seq dataset by Ma et al.,[109] after reintegration using Harmony.
In (G), (H), and (J), two-sided unpaired Wilcoxon test was used.