CellPress
OPEN ACCESS

Trends in
Immunology

Review

Unraveling B cell trajectories at single cell resolution

Dhakshayini Morgan[1,2] and Vinay Tergaonkar[1,2,3,*]

During adaptive immunity, B cells differentiate either into memory B cells or plasma cells and produce antibodies against foreign antigens to fight infection. Additionally, they behave as antigen-presenting cells and participate in T cell activation during cellular immune responses. However, their functional dysregulation can result in various autoimmune diseases and cancers. With significant breakthroughs in single cell technologies, assessing individual B cell genomics, transcriptomics, and proteomics can give deeper insights into mammalian B cell development, differentiation, antibody repertoire, and responses under conditions of homeostasis, infection, and aberrations during disease. In this review, we discuss the adoption of single cell approaches to identify different B cell gene signatures and biomarkers in normal and diseased tissues, and subsequent benefits for future therapeutic discoveries.

Highlights

Single cell techniques are used to study mammalian B cell signatures required for the differentiation of B cell subtypes within lymphoid organs.

Single cell methods allow the capture of minute changes in B cell gene signatures during health and disease; these allow the translation of newly identified genes into putative biomarkers and/or therapeutic targets.

B cell lineage tracing and signaling mechanisms have been refined due to high-resolution data from single cell methods.

These high-throughput methods, which enable the discovery of new B cell subpopulations, pave the way for studies aiming to understand the functional diversity, plasticity, and heterogeneity of B cells.

The integration of single cell methods and spatial transcriptomics can provide deeper insights into the cell-intrinsic and -extrinsic roles of B cell in healthy and diseased microenvironments.

The role of B cells in adaptive immunity at the single cell level

B cells have important roles in adaptive immune responses. For instance, upon activation by antigens, they differentiate into plasma cells and produce antibodies of different immunoglobulin isotypes [1]. Secreted antibodies neutralize and inactivate pathogens, recruit the **complement system** (see Glossary) to perform antibacterial activities, and trigger **opsonization**, whereby antibody-coated pathogens undergo phagocytosis. Activated B cells also give rise to memory B cells to induce secondary immune responses upon re-encountering the same antigen. Unregulated B cell functions, such as increases in plasma cells or autoantibody production, can result in diseases, including rheumatoid arthritis [2,3], or B cell malignancies, such as lymphomas [4–6]. Understanding various stages of B cell development, activation, and differentiation is vital for determining physiological mechanisms. In addition, understanding how these processes take place in lymphoid organs, such as the bone marrow (BM), spleen, and lymph nodes (LN), is essential for inferring cell–cell communications, the selection of immunoglobulin repertoires, and other changes that might occur during disease. Moreover, even though B cells are found in various locations throughout the mammalian body, a particular state of B cells within each distinctive niche is modified to meet the demands of host tissues [7]. Hence, several single cell methods are being increasingly used for B cell-specific lineage tracing and to resolve phenotypes, interactomes, and plasticity at the single cell level in a high-throughput manner. These methods can be divided into those that analyze (i) transcriptomic data, such as conventional **single cell RNA sequencing** (scRNA-seq) [8], **single cell paired heavy and light chain sequencing** [9], **B cell receptor sequencing** (BCR-seq) [10], and **repertoire and gene expression by sequencing** (RAGE-seq) [11], (ii) transcriptomic and proteomic data together, such as **cellular indexing of transcriptomics and epitope sequencing** (CITE-seq) [12] and **RNA and protein sequencing** (REAP-seq) [7], and lastly (iii) genomic data, such as **VDJ sequencing** (VDJ-seq) [13], **single cell chromatin accessibility sequencing** (scATAC-seq) [14] and **single cell whole-genome sequencing** (sc-WG-seq) [15] (see following sections for a detailed description of these methods). Together with **spatial transcriptomics**, these unbiased methods enable the

[1]Laboratory of NF-κB Signaling, Institute of Molecular and Cell Biology, 61 Biopolis Drive, Proteos, 138673, Singapore
[2]Department of Pathology, Yong Loo Lin School of Medicine, National University of Singapore, 5 Lower Kent Ridge Road, 119 074, Singapore
[3]Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, 117597, Singapore

*Correspondence:
vinayt@imcb.a-star.edu.sg
(V. Tergaonkar).

spatial localization of B cell gene signatures and can unravel physiologically important cell–cell interactions in various anatomical locations. In this review, we summarize recent advances contributing to our understanding of B cell biology using single cell approaches. We focus on the discovery of specific gene expression signatures of B cells in normal and diseased states. We conclude by discussing aberrant patterns of gene expression in B cells populations and their implications for the development of putative therapeutic modalities.

Using single cell and spatial transcriptomic modalities in B cell studies

One of the hallmarks of B cells is that, similar to many immune cells, they comprise a group of heterogeneous populations with a distinctive repertoire of gene expression signatures, characteristic of each developmental state. Thus, scRNA-seq methodologies integrated with novel methods of spatial biology, such as spatial transcriptomics, are among the most suitable approaches for understanding B cell biology. These methods allow inference of the status of the genome, transcriptomic profiles, and protein expression of B cells. Combined with stringent statistical and bioinformatic analyses, scRNA-seq has improved our understanding of B cell population heterogeneity, lineage trajectories, antigen receptor repertoires, and intra- and intercellular communication networks, as well as their spatial localization within tissue architecture [7]. In recent years, scRNA-seq has become a flagship method for studying B cell biology. ScRNA-seq allows the quantification of subtle variations in gene expression among cells during different developmental stages and disease status. Therefore, it significantly improves our capacity to analyze dynamic changes within the cell during its transition from one state to another. However, quantification of RNA transcripts does not provide information on protein expression or protein modifications [16,17]. To overcome this hindrance, scRNAseq is performed in parallel with other assays. These assays include, but are not limited to, flow cytometry, **cytometry by time of flight** (cyTOF), mass spectrometry, fluorescence-activated cell sorting, and histological analysis [16,17]. Nevertheless, these complementary methods probe only a few specific proteins, and do not allow a comprehensive and unbiased study of cell profiles that might reveal previously unrecognized factors. Recent developments have combined the analysis of both RNA and protein expression simultaneously at the single cell level. Some of the most widely used methods include CITE-seq and REAP-seq. The main difference between these two is that CITE-seq uses antibodies conjugated to streptavidin-biotin, while REAP-seq uses antibodies conjugated to animated DNA sequences for the direct reverse transcription of RNA to cDNA [7]. Finally, scATAC-seq has been developed to explore chromatin remodeling and to identify transcriptionally active chromatin regions. These techniques and their application in B cell-related studies are explained in the following sections.

A second hallmark of B cells is that they produce a diverse repertoire of antigen receptors [i.e. B cell receptors (BCRs)]. Sequencing the full BCR heavy and light-chain sequences at the single cell level in B cells and plasma cells has improved our understanding of clonality, **somatic hypermutation** (SHM), **class switch recombination**, and antigen receptor repertoires [18] (Box 1). Combined with scRNA-seq, a growing amount of data has allowed an understanding of B cell biology from different perspectives. However, these studies lack information about the physical localization of B cells in different niches of primary and secondary immune organs. To address this issue, recent progress in spatial biology has provided information about how to infer the spatial localization of cells within a tissue. For example, spatial transcriptomics is a relatively new high-throughput method that measures RNA expression across the entire tissue architecture [19]. It is widely used to illustrate spatial relationships between cell populations, as well as to localize various B cell-gene signatures within a tissue and in specific specialized compartments or niches [20]. The current pipeline uses transcriptomic profiles of cell clusters from scRNA-seq to be integrated with spatial transcriptomics data via deconvolution (using imaging) [3,19]. The fast-speed advances of spatial biology and

Glossary

Activating PI3K8 syndrome (APDS): PIK3CD mutations causing primary immunodeficiency.
B cell receptor sequencing (BCR-seq): sequencing VDJ recombination on BCR from sorted B cells.
Cellular indexing of transcriptomics and epitope sequencing (CITE-seq): DNA-barcoded antibodies bind to surface molecules of a single cell to capture qualitative and quantitative information about RNA and proteins.
Class switch recombination: mutations occurring in B cells at the immunoglobulin heavy chain locus to switch from 1 class (e.g., IgG) to another class (e.g., IgM).
Complement system: cascade of enzymatic reactions leading to plasma membrane attack, as part of an innate immune response against pathogens.
Cytometry by time of flight (cyTOF): flow cytometry method using heavy metals tagged with antibodies for protein quantification
Dark and light regions of the germinal center: proliferating B cells (centroblasts) undergo SHM in the dark zone and migrate to the light zone, where they become centrocytes that undergo selection by T follicular helper cells.
Follicular B cells: found in the follicles of spleen and lymph nodes.
Hidradenitis suppurativa (HS): chronic inflammatory skin condition that causes painful, small lumps to form beneath the skin.
Immune checkpoint blockade (ICB) therapy: type of cancer immunotherapy that uses antibodies to block immune checkpoint receptors of immune cells.
Marginal zone B cells: noncirculating mature B cells present in the marginal zone of lymphoid tissues.
Opsonization: antibodies bind to foreign pathogens and 'tag' them to make them susceptible to phagocytosis.
Plasmablast: differentiation stage of a B cell as it matures from a post-germinal center B cell to a mature plasma cell.
Pseudotime trajectory analysis: measures biological progression of a cell (e.g., cell differentiation) quantitatively.
Repertoire and gene expression by sequencing (RAGE-seq): combines long-read sequencing of BCR mRNA transcripts with short-read transcripts in barcoded single-cell libraries generated

Box 1. Advances in BCR sequencing at single cell resolution

Using scRNA-seq, repertoire profiling has been performed on purified B cells from human peripheral blood mononuclear cells (PMBCs) [9]. However, BCR cDNA is limited in length during scRNA-seq and might not always represent the entire antigen receptor repertoire. Additional capturing of transcription start sites for BCR genes in a single B cell was made possible using the Tn5Prime-modified scRNA-seq method [63]. FB5P-seq, BraCeR, VDJPuzzle, BALDR, VDJView, and BASIC were also designed to assemble the full length of paired heavy and light chains of BCRs using data from scRNA-seq [64–69]. New high-throughput sequencing methods, such as VDJ-seq, BCR-seq, and RAGE-seq, have been developed recently to overcome the limitations associated with identifying entire BCR lengths and sequences.

At the DNA level, VDJ-seq uses biotinylated J gene primers in an extra single-primer extension step to quantitate immunoglobulin diversity in an unbiased manner [10]. Additionally, simultaneous single cell 5′VDJ-seq and 5′RNA-seq of immune cells to generate single cell transcriptome and immunoglobulin repertoire can be performed using chromium immune cell profiling solution [53]. BCR-seq can accurately pair full-length variable regions of light and heavy chains in a massively parallel fashion [11]. RAGE-seq combines long-read sequencing of BCR mRNA transcripts with short-read transcripts in barcoded single-cell libraries generated by droplet-based partitioning to generate the full-length BCR cDNA template [12].

To further understand the antigen specificity of sequenced BCR, B cells are initially mixed with DNA-barcoded antigens, and both antigen and BCR sequences are analyzed [70]. This method is known as LIBRA-seq, which narrows the BCR repertoire to receptors that are specific to a particular antigen and, thus, can be used to accelerate antibody discovery and putative vaccine development.

Recently, the functional landscape of SHMs that accumulate in immunoglobulin genes during development and aging was described using sc-WG-seq [15]. The SHM-associated B cell-specific gene signatures were highly correlated with human B cell-tumor gene signatures, and many of these mutations accumulated with age in a sequential manner, concurring with age being a risk factor for leukemia and a decline in B cell functionality [15]. Other SHM-associated B cell-specific gene signatures could similarly be identified using sc-WG-seq for cancer risk factors, such as environmental factors, pathogens, diet, and family history, and for various other diseases, such as B cell-related autoimmune diseases, to compile a list of SHM-associated B cell-specific gene signatures for future candidate therapeutics [15].

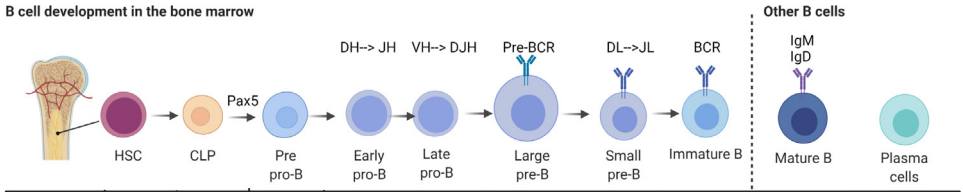
scRNA-seq are expected to improve the discovery of novel markers of disease that might be useful for the early detection or staging of B cell diseases. At the same time, the information obtained using spatial biology methods can open new opportunities for the refinement of novel therapies, such immunotherapy. In the following sections, we describe our current knowledge of B cell development in health and disease and how single cell methods have been used to understand B cell biology.

B cell gene signatures at various stages of differentiation

Bone marrow: B cell development

B cell development comprises a tightly regulated multistep process. It is essential to understand normal B cell transcriptional programs to detect early perturbations that might lead to disease, such as those observed in B cell acute lymphoblastic leukemia (B-ALL) [21]. B cell development starts in the BM, where common lymphoid progenitors (CLP) are committed to the B cell lineage via expression of transcription factors, such as Pax5 (Figure 1) [8,14,22]. Several single cell techniques have been used to study different B cell signatures corresponding to cell differentiation steps within the BM [8,14,22–25]. The first single cell approach used to study developmental trajectories of primary human B cells involved applying **Wanderlust** to mass cytometry [24]. Based on the expression of known canonical markers of B cell development (*CD34*, *CD38*, *CD10*, *CD19*, *IGH*, and *CD20*), the Wanderlust trajectory indicated the specific order of other key developmental markers, including *Tdt*, *CD24*, *CD179B*, *VpreB*, and *IL7RA*, corresponding to the crucial timing of important events, such as heavy and light-chain rearrangements (Figure 1) [24]. However, mass cytometry is biased and limited because it only probes for 44 known protein markers and does not consider the transcriptomic profile of unknown markers. To solve this, in a different study, transcriptomic profiling of B cells from human and mouse BM was performed in an unbiased manner to study the effects of hypoxia-inducible factor-1a

by droplet-based partitioning to generate a full-length BCR cDNA template.
RNA and protein sequencing (REAP-seq): DNA-barcoded antibodies are used to measure proteins and RNA to study their relationship.
Single cell chromatin accessibility sequencing (scATAC-seq): single cell sequencing assay for transposase-accessible chromatin to look at transcriptionally active chromatin sites.
Single cell paired heavy and light chain sequencing: reconstructs immunoglobulin gene sequences of paired heavy and light chains using single cell RNA seq data.
Single cell RNA sequencing (scRNA-seq): individual cells are sequenced using next-generation sequencing to understand the interaction of a single cell with its microenvironment.
Single cell whole-genome sequencing (sc-WG-seq): sequencing DNA sequences in a single cell to identify copy number variations and cell–cell heterogeneity.
Somatic hypermutation (SHM): point mutations occurring in the V-region of heavy and light chains of immunoglobulins, changing their affinity toward antigens.
Spatial transcriptomics: RNA-seq method that provides transcriptome-wide coverage to measure all gene activities within a tissue section and maps the spatial location of this activity.
T cell dependent: antigens requiring T cells for B cell activation.
T cell independent: antigens that do not require T cells for B cell activation.
T follicular helper (T_{FH}): follicular $CD4^+$ T cell helping B cells to produce antibodies.
Tertiary lymphoid structures: formed when lymphocytes and antigen-presenting cells aggregate abnormally during chronic inflammation in solid tissues.
Transitional B cells: developmental B cell stage between a BM immature cell and a mature splenic B cell.
UMAP plots: scatter plots whereby data are clustered accordingly and marked by different colors.
VDJ sequencing (VDJ-seq): quantifies the diversity of immunoglobulins by using biotinylated J gene oligonucleotides for extension with genomic DNA.
Wanderlust: algorithm for elucidating cell lineage trajectories.



Single mass cytometry (human) [8]	*CD34 CD117*	*CD34 CD 38 CD117 CD179B*	*CD34 CD38 TdT CD179B Vpreb CD10 IL7RA*	*CD34 CD38 CD24 TdT*	*CD179B VpreB CD10 CD19 IL7RA*	*CD34 CD38 CD24* / *CD19 IL7RA IGH*	*CD19 CD20 CD24 CD38 CD45*		
scRNA-seq (human) [13,14]	*KIT IL7R ATXN1 FLT3*		*IL7R CD19 CD45 CD43 RAG1*	*RAG2 EBF1 PAX5 SOX4 LEF1*		*IL7R CD19 CD45 CD43* / *RAG1 RAG2 SOX4 LEF1 EIF4EBP1*	*CD19 EBF1 CD24 EIF4EBP1 TNFRSF13C*	*CD20 CD79A CD19 EIF4EBP1 TNFRSF13C*	*IGHM/D TNFRSF13C IGKC/IGLC2 CD20 CD79A*
scRNA-seq (human) [23]				*RAG1 VpreB IGLL1*	*CD19 PAX5*	*CD24 RAG1 Vpreb* / *CD79B PAX5 RSP27* / *CD19*		*CD24 IGHM PAX5* / *CD19 VpreB RSP27*	
scRNA-seq (mouse) [25]				*Dnnt Arpp21 Vpreb Selm Tbxa2r Ttll11* / *Ankrd33b Smtn12 Tmem17b Mgst2 Pax5*		*Bub1 Dlgap5 Neil3 Sgo1 Ncapg Nut2* / *Ckap21 Cep55 Cdca5 Ska1 Pax5*		*CD20 Faim3 Bank1 Ltb Ctsh* / *Ly6a Fcer2a Bcl2 Fcrl1 H2-DM-a*	
scRNA-seq (mouse) [26]			*Runx2 Irf8 Tcf4 Bst2*	*Ebf1 Bok Ifitm2 Ifitm3* / *Vpreb Igll1*		*Vpreb Igll1 Nrgn Ybx3*			

Trends in Immunology

Figure 1. B cell gene signatures in human and mouse bone marrow. Hematopoietic stem cells (HSCs) transition from pro-B through pre-B to immature B cells and their respective transcriptomic profiles can be identified by using various single cell methods, which are listed accordingly [8,13,14,22,24,25]. IL-7Ra and PAX5 are required for B cell lineage commitment [8]. In pre-pro-B cells, transcription factors associated with myeloid lineages (RUNX2, IRF8, BST2, and TCF4) are expressed but silenced when cells proceed to the pro-B cell stage. From the pro-B to pre-B cell stage, LEF1 is required for survival and proliferation, SOX4 is required to produce RAG1 and RAG2 for VDJ recombination at the heavy chain locus (DH➔JH and VH➔DJH), and EBF1 is required to produce Vpreb and IGLL1 for surrogate light chain production [8]. Combinations of heavy chains and surrogate light are presented as pre-B cell receptors (BCRs) on large pre-B cells. Subsequently, expression of SOX4, RAG1, and RAG2 enables VJ chain recombination (VL➔JL) to take place at light chain loci. Eventually, a BCR-expressing immature B cell is formed with help from TdT and undergoes central tolerance [1,57]. CD79A and CD79B are transmembrane proteins that complex with BCRs, whereas CD10, CD19, CD20, CD34, CD43, CD45, CD117, and CD179B are leukocyte antigens present at different B cell development stages. BAFF (TNFRSF13C) is upregulated when both hypoxia-inducible factor-1a (HIF-1α) activity and hypoxia-dependent gene expression are reduced in immature B cells [8]. Created using BioRender (https://biorender.com/).

(HIF-1α) on B cell development [8]. The authors downloaded human and mouse BM scRNA-seq data [22,25] and performed clustering based on B cell canonical markers (*IL7R*, *CD19*, *RAG1*, *RAG2*, *EIF4EBP1*, *TNFRSF13C*, and *CD20*) [8]. They further subclustered B cell populations and produced B cell differentiation trajectories. Additional analysis led to the discovery of a unique group of genes that allowed the differential identification of hematopoietic stem cells (HSCs), pro-B, pre-B, and immature and mature B cells (Figure 1) [8]. Similar analyses were carried out on mouse BM B cell subpopulations and the gene clusters were comparable to those reported in Tabula Muris (a single cell transcriptomic database for *Mus musculus*) (Figure 1) [25]. Subsequently, the varying hypoxia-dependent gene expression patterns across different B cell states (such as upregulation in human pre–pro-B, pro-B, and pre-B cells, and downregulation in immature B cells) revealed that HIF-1α was required to restrict repertoire diversity in pre-immune B cells by limiting immunoglobulin heavy chain – variable genes (IgHV) editing and inducing apoptosis [8]. Thus, this study showed the applicability of scRNA-seq in recognizing key development states of B cells. In addition, the study revealed the respective B cell differentiation transcriptomic profiles, delineating trajectories and organizing differentially expressed genes into signaling and regulatory pathways that could be used to further elucidate mechanisms of differentiation [8]. Similar B cell

gene expression profiles from human fetal liver and BM were reported by another scRNA-seq study in the B cell lineage (Figure 1) [14]. Furthermore, although integrated spatial transcriptomics and scRNA-seq experiments have been conducted on BM niches [13], they have yet to elucidate the spatial relationships and cell–cell communication patterns within B cell compartments during B cell development. This includes the relationship between stromal and B cells, in which B cells depend heavily on IL-7 signals from stromal cells for their growth and survival [24].

Spleen and lymph nodes: B cell maturation

Immature B cells from the BM migrate to secondary lymphoid organs for their maturation into naïve B cells, **follicular B cells** (FOB) or **marginal zone B cells** (MZB). Recently, the switch in transcriptomic and proteomic profiles involved in the commitment of immature B cells to either FOB or MZB subpopulations via **transitional B cells stages** (T1, T2, and T3) was investigated (Figure 2). Initial bulk RNA-seq analyses of human and mouse peripheral mononuclear blood cells (PMBCs) uncovered that an induction of metabolic quiescence, characterized by the downregulation of genes involved in the PI3K-mTOR and oxidative phosphorylation pathways, acted as a transcriptional switch for transitional B cells to become FOB cells [26] (Figure 2). Subsequently, scRNA-seq performed on $IgD^{+}CD27^{-}$transitional B cells isolated from the peripheral blood of healthy individuals and patients with **activating PI3Kδ syndrome** (APDS) further revealed that the activation of the extracellular adenosine salvage pathway resulted in metabolic quiescence during FOB transition [26] (Figure 2). First, scRNA-seq revealed the existence of hyperactive PI3K-mTOR signaling and a developmental block in transitional B cells toward FOB commitment from patients with APDS harboring *PI3KCD* mutations [26]. Second, the overall FOB-related gene signature identified from integrated scRNA-seq analyses presented high amounts of extracellular adenosine salvage pathway-related gene transcripts (*CD73* and *ADK*) and reduced gene expression of metabolic pathways ,such as PI3K-mTOR, aerobic respiration, and protein synthesis [26]. Based on these scRNA-seq results, downstream activation of AMPK in the extracellular adenosine salvage pathway was eventually demonstrated to inhibit mTOR signaling. Additionally, these results contributed to explain the rationale behind using PI3K inhibitors to treat APDS, which aims to normalize transitional B cell counts because of FOB development blockade [26]. However, the activation of the extracellular adenosine salvage pathway for FOB cell maturation has not been observed in mice [26].

The switch from transitional B cells to MZB has been widely studied in mice (Figure 2) [27]. Apart from the fact that transitional B cells become MZB precursors before MZB cells [28,29], little has been known about this maturation process in humans. However, one study analyzed the transcriptomic profiles of IgM^{high} and IgM^{low} human transitional B cells from peripheral blood mononuclear cells (PBMCs) via scRNA-seq [30]. The transcripts for *PLD4*, *MZB1*, *CD1C*, and *SOX4* (MZB-associated genes) were upregulated in IgM^{high} transitional B cells, while *KLF2*, *CCR7*, *IL4R*, and *SELL* (FO-associated genes) were upregulated in IgM^{low} transitional B cells (Figure 2) [30]. This suggested that the amount of IgM expression and signaling in transitional B cells is involved in B cell fate decisions. Of note, the scRNA-seq data were consistent with the scRNA study on transitional-to-FOB B cells mentioned previously [26], which showed that transitional B cells (T2 and T3) that were committed to a FOB cell fate harbored a low IgM gene signature (Figure 2). Moreover, scRNA-seq in combination with CITE-seq of human $CD19^{+}$ B cells presented **UMAP plots** that displayed close juxtaposition of IgM^{high} naïve cells with MZBs [30]. This effect was attributed to the cell subtypes sharing transcriptomic features, such as upregulation of *CD1C* and downregulation of *CD27* [30]. Subsequently, **pseudotime trajectory analysis** showed a linear progression from IgM^{high} transitional B cells to MZBs [30]. Together, these results introduce the idea of an IgM^{high} developmental trajectory for human MZBs (Figure 2) [30].

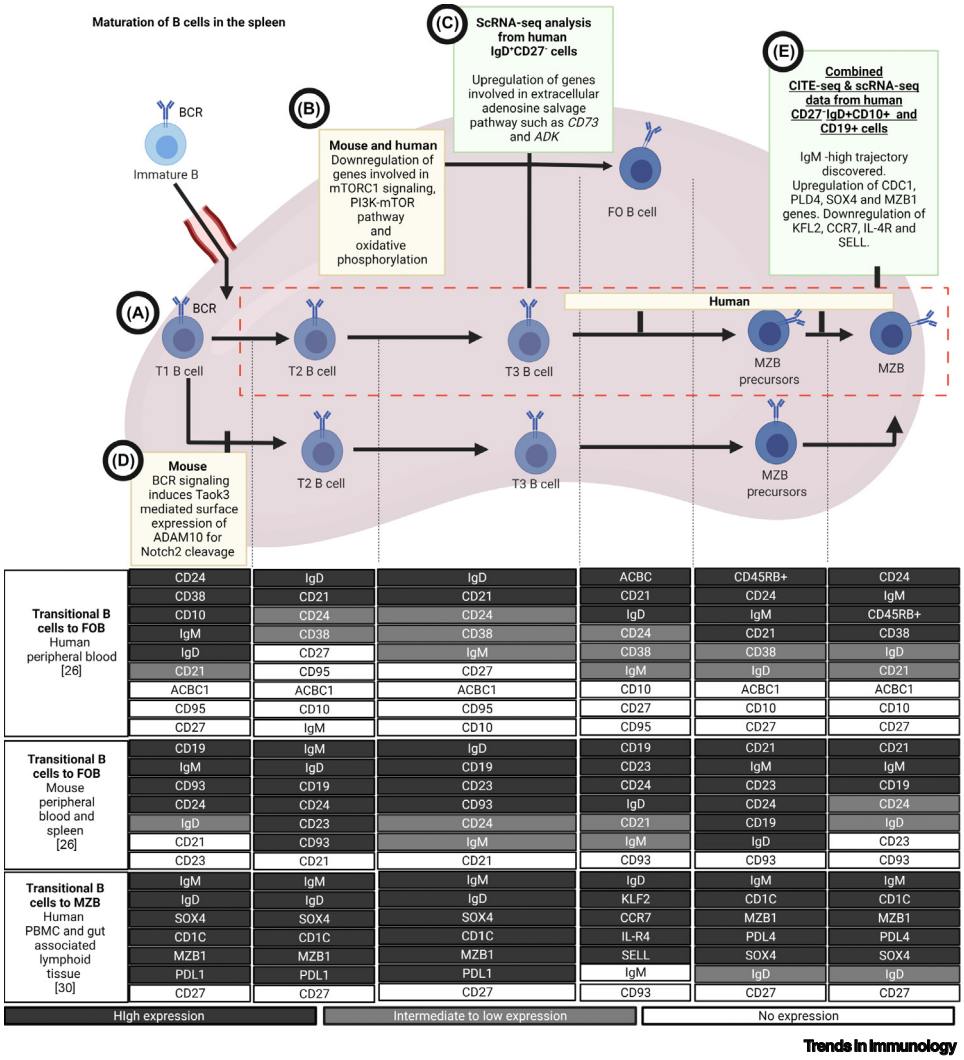


Figure 2. Varying B cell signatures during B cell maturation at secondary lymphoid organs in mice and humans. (A) Immature B cells undergo transitional stages (T1, T2, and T3) at secondary lymphoid organs, including the spleen, to become either follicular B cells (FOB) or marginal zone B cells (MZB) [26,56,58]. Weaker BCR signals combined with Notch 2 signals favor an MZB cell fate, while stronger BCR signals favor FOB cells [58]. Moreover, BAFF survival and NF-κB signals from stromal cells also have a role in lineage decision-making [56]. (B) Bulk RNA-seq analyses of human and mouse peripheral blood mononuclear cells (PBMCs) revealed that the top gene set enrichment pathways in transitional B cells, such as mTORC1 signaling, PI3K-mTOR signaling, and oxidative phosphorylation, exhibit reduced gene expression in FOB cells. This suggests that induction of metabolic quiescence can serve as a switch turning transitional B cells into FOB cells [26]. (C) Single cell RNA sequencing (ScRNA-seq) presenting activation of the extracellular adenosine salvage pathway as an inhibitor of the mTOR pathway is required for metabolic quiescence in transitional B cells during FOB cell fate [26]. (D) As previously reported in mice, for transitional B cells to achieve a MZB cell fate, BCR signaling followed by Toak3-ADAM10-mediated cleavage of Notch 2 in transitional B cells is required. While much remains unknown about MZB cell fates in humans, (E) scRNA-seq has revealed an IgM^{high} developmental trajectory (red lined box) for transitional B cells that commit to becoming MZB cells [30]. Created using BioRender (https://biorender.com/).

Secondary lymphoid organs: naïve B cell activation and differentiation

Depending on the antigen (Ag), **T cell-independent** (TI) or **T cell-dependent** (TD) activation of MZB and FOB leads to plasma or memory B cells, respectively (Figure 3) [31–33]. For TD-Ag responses, T and B cells form long-lived CD40L/CD40R interactions and subsequently differentiate

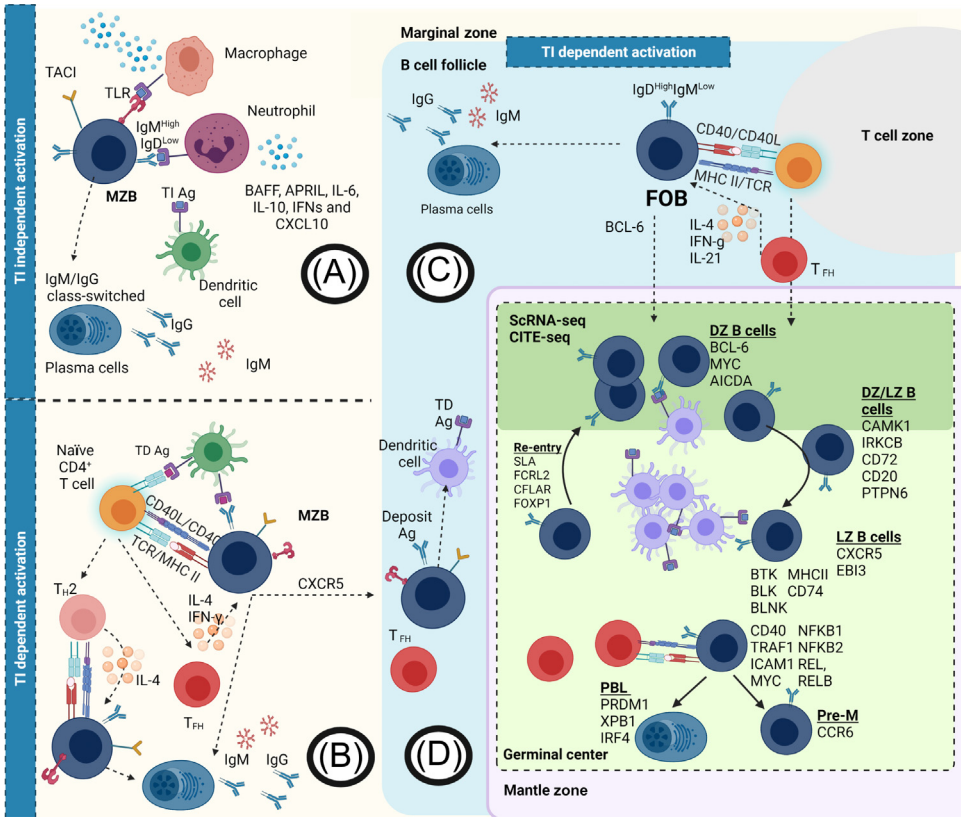


Figure 3. B cell profiles in human germinal centers (GCs) generated via single cell approaches. (A) During T cell-independent (TI) antigen (Ag) responses, innate cells present TI Ag to receptors [TACI, TLR, and B cell receptors (BCRs)] on marginal zone B cells (MZB) and promote their differentiation into IgM/IgG plasma cells [31,32]. (B) Dendritic cells present T-dependent (TD) Ag to activate $CD4^+$ T cells and MZBs. Activated $CD4^+$ T cells differentiate into effector T helper (T_H2) cells and communicate with MZB to promote the differentiation of the latter into short-lived plasma cells. Alternatively, follicular T helper (T_{FH}) cells secrete IL-4, IFN-γ, and IL-21, causing MZB to lose SP1 expression and upregulate CXCR5 for migration of the cells into B cell follicles, where they deposit TD Ag on follicular dendritic cells (FDCs) [31,32]. (C) TD Ag-encountering naïve follicular B cells (FOB) migrate to the edge of follicles and interact with T_{FH} cells [59,60]. They either become short-lived plasma cells or migrate to the center of the follicle to seed an early GC by upregulating BCL-6. (D) Within the GC, FOB cells undergo proliferation and somatic hypermutation (SHM) in the dark zone as centroblasts and move to the light zone as centrocytes for affinity selection, class switch recombination, and differentiation into memory B cells or plasma cells, with help from T_{FH} cells [60]. Light zone cells re-enter the dark zone for several rounds of affinity-based selection, resulting in bi-directional migration. Cellular indexing of transcriptomics and epitope sequencing (CITE-seq) and single cell RNA sequencing (scRNA-seq) have been performed to study the heterogeneity of GC B cells, and some of the major genes listed in Table 1 in the main text are indicated. ScRNA-seq analyses have also revealed that follicular dendritic cells (FDCs) expressing CXCL12 and CXCL13 are present in dark and light zones, respectively, where they aid affinity-based selection [61,62]. Created using BioRender (https://biorender.com/).

into **T follicular helper** (T_{FH}) and plasma cells, respectively**.** Upon germinal center (GC) formation within secondary lymphoid organs (e.g., LNs), a high-affinity BCR repertoire of long-lived plasma and memory B cells ensues.

Different GC B cell subpopulations have been identified based on single cell transcriptomic analyses revealing the heterogeneity of B cell gene expression signatures in the **dark and light regions of the GC** [34] (Table 1). One group developed a pseudo-time inference algorithm based on scRNA-seq analyses of human GC B cell populations in spleen and tonsils that

Table 1. B cell gene signatures in mammalian tissues including human GCs identified via scRNA-seq

Tissue/organ	B cell-specific transcript expression	Refs
Mouse		
Lung	*CD19*$^{+}$	[25]
Adipose tissue	*CD19*$^{+}$	
Limb muscle	*CD19*$^{+}$	
Spleen	Abundant: *CD79a*$^{+}$	
	Marginal zone B cells: *CCR2*$^{+}$	
BM	Early pro-B cells: *Dntt*$^{+}$, *PAX5*$^{-}$, *Rag1*$^{+}$, *Rag2*$^{+}$	
	Late pro-B cells: *Dntt*$^{+/-}$, *Vpreb1*$^{+}$, *PAX5*$^{+}$, *Rag1*$^{+}$, *Rag2*$^{+}$, *CD19*$^{+}$, *CD20*$^{-}$, *CD22*$^{-}$ *Late pro-B cells could be further subdivided into Dntt*$^{+}$ *and Dntt*$^{-}$ *late pro-B cells*	
	Intermediate pre-B cell: *Dntt*$^{-}$, *Vpreb1*$^{-}$, *PAX5*$^{+}$, *Rag1*$^{+}$, *Rag2*$^{+}$, *CD19*$^{+}$, *CHCHD10*$_{high}$, *CD20*$^{-}$, *CD22*$^{-}$, *CD74*$_{low}$	
	Immature B cells: *CHCHD10*$^{+}$, *CD79a*$^{+}$, *CD79b*$^{+}$, *CD19*$^{+}$, *MS4A1*$^{-}$*/low*, *CD74*$^{-}$, *Mki67*$^{+}$, *Stmn1*$^{+}$ Immature B cell subset: *Dntt*$^{-}$, *Vpreb1*$^{-}$, *PAX5*$^{+}$, *Rag1*$^{-}$, *Rag2*$^{-}$, *CD19*$^{+}$, *CHCHD10*$_{high}$, *CD20*$^{+}$, *CD22*$^{+}$, *CD74*$_{high}$	
	Mature (naïve) B cells: *CHCHD10*$^{-}$, *CD79a*$^{+}$, *CD79b*$^{+}$, *CD19*$^{+}$, *MS4A1*$^{+}$, *CD74*$^{+}$, *Mki67*$^{-}$, *Stmn1*$^{-}$ A more mature (naïve) B cell subset: *Dntt*$^{-}$, *Vpreb1*$^{-}$, *PAX5*$^{+}$, *Rag1*$^{-}$, *Rag2*$^{-}$, *CD19*$^{+}$, *CHCHD10*$_{low}$, *CD20*$^{+}$, *CD22*$^{+}$, *CD74*$_{high}$	
CNS/Meninges	B cells: *IGHM, IGKC, IGLC*	
	Pan-B cells: *CD79b, CD19*	
	Mature B cells: *H2-Aa, Ms4a1*	
	Immature B cells: *Rag1*, *CD93*	
	Dura Bc 1: ↓ Antigen presentation (*CD74*, *IFI30*, *RT1-Da*) and migratory potential (*Tmsb4x*, *Pfn1*, *CD37*); ↑ proliferation (*H3f3b*, *Eif1*, *CD24*), NF-κB signaling (*TIFa*, *Nfkbia*), and survival (*CXCR4*) Dura Bc 2: ↑ *CCL2*, *CXCL1/2*, *CCL4*, *ITGB2*	
	Plasma cells: ↑ *IGHA*, *IGHG*, *Sdc1/CD138*, *Prdm1/Blimp-1*, *and JCHAIN*	
Liver	Circulating B cells: *CD79a*$^{+}$, *CD79b*$^{+}$, *CD74*$^{+}$, *CD19*$^{+}$	
Diaphragm	*CD19*$^{+}$	
Human		
Lung	*IGKC*	[22]
Liver	*CD20 (MS4A1), CD79B*	
	Mature inexperienced B cells: *IGHD, CD19, CD20 (MS4A1), CD22, CD52Top differentially expressed genes: MS4A1, LTB, CD37, CD79B, CD52, HLA-DQB1, TNFRSF13C, TCL1A, LINC00926, STAG3, IGHD, BANK1, IRF8, BIRC3, P2RX5, RP11-693J15.5, RP5-887A10.1, VPREB3, CD22, CD74, SELL*	
	Plasma cells: *CD20 (MS4A1)*^{low}, *CD27*, *CD38*, *Ig heavy and light chains* Top differentially expressed genes: *GLC2*, *IGHG1*, *IGKC*, *IGHG2*, *IGHG3*, *IGHGP*, *IGLC3*, *JCHAIN*, *IGHA1*, *IGHG4*, *IGHA2*, *IGHM*, *IGLV3-1*, *IGLC7*, *MZB1*, *CD79A*, *SSR4*, *IL16*	
Blood	*IGKC, MS4A1, CD79A*	[22]
Spleen	B cells: *CD20 (MS4A1), CD79A, CD19*	
	Plasma cells: *JCHAIN*	
Stomach	*CD79A*	

(continued on next page)

Table 1. (continued)

Tissue/organ	B cell-specific transcript expression	Refs
Kidney	*CD79A, CD79b*	
Esophagus	*CD79A, CD19*	
Eye	*CD79A*	
BM	*$CD34^{+}$ pre-B cells*	
	B cells: *CD20 (MS4A1), CD79A, CD19, PRDX2, HBD*	
	Plasma cells: *CD20 (MS4A1), CD79A*	
Method used (cell type/species)	B cell signatures in GC	
scRNA-seq: GC human B cell subpopulations in spleen and tonsils	GC B cells: *BCL6, AICDA* DZ: *STMN1, BIRC5, TOP2A, AURKB, CXCR4, TACC3, CKS1B, CDK1, MKI67, UBE2C, CCNA2, KIAA0101, CDCA3, KIF22, HMGB3, CDKN3, NUSAP1, AICDA, MYBL2, LMNB1* DZ/LZ: *GPX1, CD27.LY86, HMCES, JCHAIN, HSPA4, DNAJC10, FCRLA, INPP5D, QRSL1, REEP5, LILRB1, CD24, WDR83OS, ALOX5, ARHGEF2, SERINC2, RPS19BP1, PYCARD, C2orf88LZ, BCL2A1, CD83, IL4I1, LMO3, EBI3, HLA-DQB2, TRAF4, PLEK, IER2, NFKBIA, FCER2, BCAR3, DUSPSNX11, PLPP5, PHACTR1, TAP1, RAB3GAP2, DHRS9, FCRL5* LZ: *BCL2A, CD83, IL-4I1, LMO2, EBI3, HLA-DQB2, TRAF4, PLEK, IER2, NFKBIA, FCRE2, BCAR3, DUSP2, SNX11, PLPP5, PHACTR1, TAP1, RAB3GAP2, DHRS9, FCRL5* Memory B cells: *NOTCH2, JAM3* Antibody-producing B cells: *PRDM1, SDC1*	[6]
scRNA-seq, CITE-seq; GC human B cell subpopulations in $CD3^{-}IgD^{-}CD38^{+}$ tonsil cells	DZ: *CXCR4 (↑ PCNA, MKI67, CDK1, CDC20, FOXP1, AICDA, MYC, EZH2, E2F1, FOXO1, BCL6)DZ/LZ intermediate: PCNA, MKI67, CDK1, CDC20, CD72, PTPN6, IFNGR1, CAMK1, CD22LZ: CD83, BCL-2A1 (* ↑ BCR signaling, NF-kB signaling, MYC and Ag presentation genes)DZ re-entry: *SLA, FCRL2, CFLAR, FOXP1Memory B cell precursors (pre-M): CCR6, CD44, CD69, CXCR4 (* *↑ BANK1, RASGPR2, CELF2, IFITM1, ITFM2, INFGR1, GPR183, TNFRSF13B, SELL, MYC, FXYD5, STAT1, PAX5; ↓ MEF2B, SIPR2, no cell-cycle genes)Plasmablasts: PRDM1, CD9 (* *↑ IRF4, XPB1, MZB1, TNFRSF17, FKBP11, CAMK1, MEF2B, RGS13; ↓ MS4A1, PAX5)Plasma cells: IRF4, PRDM1, XBP1, FKBP11*	[5]
scRNA seq: mouse splenic GC	LZ: *CD52, CD83, HSPD1, RAN, MIF, ATP5B, MYC, DKC1, LRRC58, H2-AaDZ: GADD45A, AICDA, PAX5, SON, BRD4, TCF3, CECR2, FAM193a, BCL6, CXCR4GZ (second DZ population): RPS15, RPS26, RPS20, SNHG8, BRCA1, MCM5, LIG1, DHFR, CDC6*	[36]
scRNA seq: mouse splenic B cells	Pre-M B cells: BCL6, FAS, CD38, CCR6, BCL2, *MKI67, HIST1H1B, CENPF, 2810417H13RIK, CENPE, INCENP, HIS1H2AP, TOP2A, HIST1H1EMemory B cells: HHEX, TIE3, KIF2, ZEB2, BH1H4E1PML, MNDA1*	[35]

indicated the existence of a large population of cells that comprised dark zone (DZ) and light zone (LZ) intermediates (DZ/LZ), thus changing our perspective of the GC as a two-compartment structure [6]. ScRNA-seq revealed that, as the human GC cells cyclically transitioned between DZs and LZs, the gene expression for the transitional states modeled a cyclic continuum (Figure 3 and Table 1) [6]. Another scRNA-seq study combined with CITE-seq on human GC B cells initially identified 13 clusters [four DZ, six DZ/LZ, one LZ, one memory B cell (pre-M), and one **plasmablast** (PBL)] based on the expression of *CXCR4* (DZ), *CD83* (LZ), *CCR6* (pre-M), and *PRDM1* (PBL) [5]. The cell cycle genes (*PCNA*, *MKI67*, *CDK1*, and *CDC20*) were mostly captured in DZ and DZ/LZ cells via scRNA-seq, validating GC DZ as the proliferative

compartment [5]. However, scRNA-seq analyses also revealed that 20% of LZ GC B cells also expressed these cell cycle genes and, thus, were transcriptionally capable of cell division, a result that had not been previously recognized [5]. Thus, these cell cycle genes were removed for reclustering, which resulted in three DZ, five DZ/LZ intermediate, two LZ (a/b), one pre-M, and two PBL (a/b) clusters instead [5]. Additionally, in this study, DZ/LZ cells constituted more than 50% of GC cells and formed various distinct subpopulations, with gene signatures ranging from relatedness to DZ to LZ, similar to findings from others [7]. The differential switch from LZ to pre-M or plasma cells was further investigated (Figure 3) [5]. Early LZ cells expressing *BACH2*, *CCR6*, *CD44*, *CD69*, *BANK1*, *RASGRP2*, *STAT1*, and *INFGR1* were likely to become pre-M cells, whereas late LZ cells expressing *CD9*, *PRDM1*, *XBP1*, *MZB1*, *TNFRSF17*, and *FKBP11* were likely to become PBL [5]. The transcription factor, Hhex, and co-repressor, Tie3, were also shown to promote memory B cell differentiation through scRNA-seq analyses in mouse and human GC B cells [35]. Lastly, another group described two distinct B cell populations in the DZ of mouse GC that differed in function and location based on their scRNA-seq transcriptomic profiles, and have since separated mouse GC B cells into three populations [DZ, gray zone (GZ), and LZ cells] harboring unique molecular programs [36] (Table 1). Overall, scRNA-seq and CITE-seq have further identified distinct cellular compartments, such as DZ/LZ intermediates, GZ cells, and a subset of LZ cells capable of proliferating within different regions of the GC [36]; thus, further studies are warranted to understand the functions of these newly identified subsets.

B cell populations identified in various tissues via single cell approaches

Single cell transcriptomic data of over 100 000 cells were collected from 20 different mouse organs and tissues [25]. Based on the differential gene expression of various cell populations, including previously uncharacterized populations, it was concluded that 136 transcription factors were sufficient to simultaneously specify all cell populations in all organs [25]. During this process, mouse B cells were found in lung, adipose, limb muscle, spleen, BM, liver, and diaphragm tissues, and produced cell-specific transcripts that corresponded to their protein expression (Table 1) [18]. Of note, a large population of B cells was discovered in the duras layer of the meninges in mice under basal conditions, suggesting that central nervous system (CNS)-associated border tissues harbor additional immunological roles [37] (Box 2).

Similarly, using scRNA seq, 119 cell types were identified in 13 human tissues, which included B cell populations that localized to the lungs, esophagus, stomach, spleen, BM, blood, liver (Box 2), kidney, and eye [22] (Table 1). Furthermore, integrated analysis of human hematopoietic diversity in the BM from eight healthy donors was carried out via scRNA-seq: 18 $CD34^+$ cell populations were resolved, including those with graded gene expression differences, suggesting putative transitional states and rare progenitor states [38]. To investigate lineage dependency, frequencies of various $CD34^+$ cell populations in the BM were compared. $CD34^+$ pre-B cells had the greatest variation in population frequency, ranging from 0.05% to 2.15%, which corresponded to the frequency of unknown $CD34^+$ lymphoid progenitors and pro-B cells; this suggested that distinct cell population frequencies were lineage dependent [38]. Using scATAC-seq, human PBMCs were profiled to identify open chromatin landscapes as well as heterogeneity in chromatin accessibility, which were used to delineate different cell types [39]. The scATAC-seq analysis identified cell type-specific cis-regulatory elements for B cell genes, such as *CD19*, *EBF1*, and *LYN*, in B cells isolated from human PBMCs [39]. Hence with the aid of single cell resolution techniques, it has been possible to determine locations of B cell clusters based on their differential gene expression. However, the trajectories of B cells in nonlymphoid organs remain poorly understood (Box 2), and their interactions with other cells will require more ligand–receptor analyses, as well as the integration of single cell methods with spatial transcriptomics.

Box 2. scRNA-seq revealed B cell roles in nonlymphoid organs

In one study, mouse dura B cells and their progenitors (pro- and early pre-B cells usually found in the BM) were detected in the dura and shown not to interact with skull BM or the periphery [37]. This discredited the possibility of a flux of dura B cells from either of these areas [37]. Moreover, dura B cells gene signatures were revealed, which included the downregulation of antigen presentation and migration genes, as well as the upregulation of chemotaxis, proliferation, and NF-κB signaling genes [37]. Subsequently in this same study, IgA^+ class-switched plasma cells expressing *Sdc1* (CD138), *Prdm1* (Blimp-1), and the *J* chain were identified using scRNA-seq in the multiple sclerosis (MS) mouse model of experimental autoimmune encephalomyelitis (EAE), thus supporting the finding IgA^+ B cells accumulate in the human meninges during MS [37] (see Table 1 in the main text). However, another scRNA-seq study on meningeal B cells yielded contradictory results on the flux of dura B cells from the skull BM and periphery [23]. The calvarium containing a BM niche was shown to supply the CNS with B cells through specialized vascular connections. In addition, infiltration of peripheral age-associated B cells into the meninges and their differentiation into Ig-secreting cells were demonstrated in such aged mice [23]. ScRNA-seq analyses additionally uncovered the differentiation stages of dura B cells, from pro-B to mature B cells, akin to their development in the BM [23]. CXCR4-expressing dura early B cells were found to interact with dura fibroblast-like cells, providing them with CXCL12 and, thus, showing that they were reliant on the CXCL12–CXCR4 axis for their survival and differentiation [1]. Moreover, the sequential expression of *CD93*, *SMARCA4*, *RAG1*, *IL7R*, *MKI67*, *H2-Aa*, and *CD79B* in developing dura B cells was parallel to that in developing BM B cells [23]. To study the immunological roles of the human liver, scRNA-seq was performed to examine hepatic B cells and identified distinct subsets of B cells undergoing varying stages of development [71]. These liver-resident B cells included mature, inexperienced B cells and plasma cells. The naïve matured B cells highly expressed *CD20* (MS4A1) but did not express *CD27* or *CD138*. Plasma cells that had terminally differentiated to secrete antibodies against a particular antigen expressed low amounts of *CD20* and high amounts of *CD27*, *CD138*, and immunoglobulin heavy and light chains. Other differentially expressed genes in both the B cell hepatic populations were also identified (Table 1), and their transcription profiles in the human liver could aid the examination of liver diseases in which B cells might be relevant.

Changes in B cell genetic signatures during disease and putative therapeutic uses

Activated B cells become antibody-secreting cells to fight threats such as foreign pathogens and cancer cells (Table 2, Key table). Moreover, for certain conditions, dysregulated B cell gene signatures and functions can result in pathologies such as certain autoimmune diseases, chronic inflammation, cancer progression, and immunotherapy resistance (Table 2) [4,40–42]. The robustness of single cell methods has been useful in identifying changes in B cell biology during disease states (Table 2) and for translating them into the discovery of potential B cell biomarkers for targeted therapeutic purposes. For instance, scRNA-seq and VDJ-seq have been used to identify 14 potential severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)-neutralizing antibodies (IgG-presenting clonotypes) with high therapeutic and prophylactic efficacy against Coronavirus 2019 (COVID-19) [43]. ScRNA-seq analyses have also led to identifying *MX2*, an IFN-response gene, as a candidate biomarker for naïve B cells during the progression of severe dengue fever [44], as well as the detection of $IgG1^+$ plasma cells with reduced SHM and clonotypic diversity correlating with anti-TNF resistance in patients with inflammatory bowel disease (IBD) [45]. In **hindradenitis suppurativa** (HS), a chronic inflammatory skin disease, scRNA-seq data from skin lesions of patients revealed a predominance of B cells and plasma cells, leading to an increase in immunoglobulin diversity and production [46]. Ligand–receptor analysis from scRNA-seq also revealed that the B cells and plasma cells were interacting with stromal tissue cells and other immune cell components for activation of the complement systems in these patients [46]. Hence, as potential therapeutic targets in B cells and plasma cells, the anti-CD20 antibody drug rituximab and Bruton tyrosine kinase (BTK) inhibitors, respectively, were suggested for HS clinical trials [46]. Thus, single cell approaches are enabling us to better understand disease environments, cellular heterogeneity, and cell interactions. Moreover, the integration of scRNA-seq with spatial transcriptomics has illustrated the reciprocal distribution of B and T cells in human synovial tissue joint sections during rheumatoid arthritis, as evidenced by the spatial localization of plasma cells residing outside of annotated infiltrating T cell regions in joints, while class-switched B cells reside in the border regions of lymphocyte aggregates [3,47].

For various types of cancers, single cell methods have helped to dissect intra- and intertumor heterogeneity [6], identify cells of origin [5], investigate genetic changes in distinct ecosystems

Key table

Table 2. Changes in B cell signatures during different disease states, identified using advanced single cell approaches and spatial transcriptomics

Disease (species)	Single cell approach	B cell signatures observed	Refs
Autoimmunity and chronic inflammation			
Rheumatoid arthritis (human)	scRNA-seq; mass spectrometry; flow cytometry	Four synovial B cell clusters identified: *IGHD*$^{+}$*CD27*$^{-}$ naïve B cells; *IGHG3*$^{+}$*CD27*$^{+}$ memory B cell: *IGM*, *IGG*; *ITGAX*$^{+}$*TBX21*$^{+}$ autoimmune-associated B cells: *ACTB*, *GBP1*, *ISG15*, *IGM*, *IGG*; Plasmablasts: Ig genes and *XBP1*	[2]
	scRNA-seq	Three B cell clusters and common genes identified, including: ↑ *CD74* (HLA invariant chain), *HLA-DRB1*, and paired *HLA-DRA* chain genes, implicating antigen-presenting capacity. ↑ *CD40*, *CD80*, *CD86*, *CD84* for interaction with T cells, ↓ *TNFSF4* (*CD134/OX40L*); Naïve B cells: *IGHD*,*IL4R*, *CCR6*, *CR2* (*CD21*), *FCER2* (*CD23*), *ITGB2* (*CD18*), ↑ *MS4A1* (*CD20*), *ICOSLG*, *IL21R*, *IL16*; Memory B cells: *ITGAM* (*CD11b*), *GPR138* (*EBI2*), *TNFRSF13B* (*TACI*), *CD27* and *CXCR3*, *CCR6*, *CR2* (*CD21*), *FCER2* (*CD23*), *ITGB2* (*CD18*), *MS4A1* (*CD20*); Plasma cells: *XBP1*, *SDC1* (*CD138*), *SLAMF7* and *PRDM1* (BLIMP-1), *TNFRSF13B* (*TACI*), *CD27*, *CXCR3*, *ICAM2*, *ITGA6* (*CD49f*), *CCR2*, *CCR10 CXCR4* (survival), *TNFRSF17*, *IL6R*, *PECAM1* (*CD31*), *SELPLG* (*CD162*), *CD98*, *SLC7A5*; Anti-citrullinated protein antibodies clone-type expansion: *IGHV5-51*01*, *IGHJ4*02/5*02*, *IGKV2-40*01/2D-40*01*, *IGKJ5*01*	[3]
	Spatial transcriptomics	Synovial B cell-rich area contain plasma cells: ↑ *JCHAIN*, *IGH*, *IGL*, *MZB1*, *SDC1*, *XBP1* Plasma cells highly colocalized with memory B cells, *CCR7*$^{+}$ T cells, and *GZMK*$^{+}$ T cells	
		B and T cells reciprocally distributed in synovial tissue joint sections Plasma cells resided outside of annotated infiltrate T cell regions, while class-switched B cells resided in border regions of aggregates	[47]
Systemic lupus erythematous (SLE) (human)	scRNA-seq	Seven B cell subclusters identified SLE-restricted B cell subcluster exhibited: Double-negative switched memory cells (DN2) phenotype: *CD19*$^{+}$*IgD*$^{-}$*CD27*$^{-}$*CXCR5*$^{-}$ ISG expression,↑ *CD19* , *MS4A1* (*CD20*), *TBX21* (T-BET), *ITGAX* (*CD11c*), *FGR*, *TFEC*, *FCR*, *FCRL3*, *FCRL5*, *IL10RA*	[72]
		SLE B cells have reduced heterochromatic modifications, H2AK119Ub and H3K27me3, on inactive X chromosome, causing aberrant X-linked gene expression and XIST RNA interactome genes	[73]
Celiac disease (human)	Single cell sequencing of Ig V_H and V_L genes	Transglutaminase 2-specific *IGHV5-51:IGKV1-5* pairs in BCR autoantibody repertoire indicating biased pairings for V_H:V_L Selective amino acid changes (K45R and Q106H) detected in CDR3 region of V_L *IGHV5-51:IGKV1-5* pairs with preference for CDR3 length; 14-aa-long in H chain and 11-aa-long in κL chain	[74]
HS ((human)	scRNA-seq	B cells and plasma cells predominant in HS lesions Ligand–receptor analysis revealed that B cells and plasma cells interact with stromal tissue cells and other immune cell components Gene signature for B cells: *IGHM*, *MS4AK*, *CD79A*, *IGKC*, BCR- and IFN-γ-mediated signaling Gene signature for plasma cells: *IGHG3*, *IGHG1*, complement activation, B cell activation, antibacterial humoral response Plasma cells as potential therapeutic target in HS using inhibitors for BTK signaling	[46]
Acute lung injury (ALI) (mouse)	scRNA-seq	Four lung B cell clusters identified ALI induced chemokine and inflammatory genes in B cells: *Cd79a*, *Cd79b*, *Ms4a1*, *Cd19*, *Pax5*, *Cd22*, *Cd86*, *Ccl3*, *Ccl4*, *Ifi47*, *Irf1*, *Ifitm1*, *Ifitm2* Upon mesenchymal stem cell treatment, reductions in chemokine and Ig gene expression in lung B cells: *Ccl3*, *Cd86*, *Iglc2*, *Iglc3*	[75]
Inflammatory bowel disease (human)	scRNA-seq	IgG^{+} plasma B cells Correlated with anti-TNF resistance ↓ Clonotypic diversity and SHM, despite their expansion	[45]

(continued on next page)

Table 2. (continued)

Disease (species)	Single cell approach	B cell signatures observed	Refs
Viral infections			
SARS-CoV2 (human)	scRNA-seq	Four clusters of B cell populations from patients with COVID-19: Naïve B cells (B1) expressing *CD19*, *CD20* (*MS4A1*), *IGHD*, *IGHM*, *IL4R*, *TCL1A*; Memory B cells (B2): *CD27*, *CD38*, *IGHG*; Immature B cells (B3): *CD19*, *CD20* (*MS4A1*); Plasma cells (B4): ↑ *XBP1*, *MZB1*	[76]
		All B cells from patients with COVID-19 highly expressed *IL1B*, *CCL3*, *IRF1*, *DUSP1*, *JUN*, *FOS* Memory and plasma B cells: *S100A8*, *IGLL5*, *SSR3*, *IGHA1*, *XBP1*, *MZB1* Memory B cells had undergone high clonal expansion	
		Expanded IgA and IgM^+ BCR clones in patients with COVID-19 Use of *VDJ* genes ↑ *IGHV3* family: *IGHV3-7*, *IGHV3-15*, *IGHV3-21*, *IGHV3-23*, *IGHV3-30* *IGKVs*: *IGKV1-17*, *IGKV2-28*, *IGKV3-15* *IGLVs*: *IGLV1-44*, *IGLV2-8*, *IGLV3-27*	
		B cells communicate with other immune cells: Secrete IL-6, LTA, and LTB, which bind to IL-6R, LTAR, and LTBR, respectively, expressed on monocytes; Secrete IL-6, which binds to IL6R on T cells for production of IFN-γ, IL-1β; Express *IL18RAP*, *TNFRSF13B*, *TNFRSF17*, *TNFRSF13C* to bind to respective ligands on dendritic cells for proliferation	
		Monocytes were predicted to interact with B and plasma cells via cytokine/receptor pairs that include TNF-α/TNFR and IL-6/IL-6R	[77]
		Plasmablasts previously expressing *CD27*, *CD38*, and *TNFRSF17* upregulate *DEF3A*, *ELANE*, *MPO*, *CHI3DL1*, *LCN2*, *LTF*, *MMP8*, *MMP9*, and *CAMP* instead and, thus, possibly differentiate into developing neutrophils characteristic of acute respiratory disease in severe COVID-19 infection	[33]
		Four B cell clusters identified: Naïve B cells: $MS4A1^+IGHG1^{-;}$ Plasma cells: $MZB1^+IGHG1^+$; Cycling plasma cells: $MZB1^+IGHG1^+MKI67^+$; Memory B cells: $MS4A1^+IGHG1^+$	[78]
		B cells from patients with COVID-19: ↑ *PRDM1*, *XBP1*, *IRF4* (differentiation into plasma cells) ↑ *ISG15*, *IFI44L*, *MX1* (interferon-stimulated genes and anti-viral genes) *XAF1*, *TNFSF10*, *FADD IRF1*, *TP53*, *CASP3* (induce apoptosis)	
	scRNA-seq VDJ-seq	8558 $IgG1^+$ antigen-binding clonotypes and 14 SARS-CoV-2 neutralizing antibodies identified Antibody BD-368-2 showed high therapeutic and prophylactic efficacy in SARS-CoV-2-infected mice IgG1-presenting clonotypes with high $CDR3_H$ structuring can be selected as potent neutralizing antibodies	[43]
	scATAC-seq	10 $PAX5^+MS4A1^+$ B-cell clusters: Immature B cells: *SDC1* Naïve B cells : *TCL1A*, *CD19*, *CD20* Memory B cells : *CD27*, *CD38* Plasma cells : *XBP1* Integrated B cell-lineage trajectory based on gene expression of *CL1A*, *CXCR4*, *MEF2C*, *BHLHE41*, *BAFF*, *PLCG2*, *PAX5*, *EBF1*, *RORA* at every B cell development stage. *BCL11A*, *IRF8*, *PAX5*, *REL*, *BATF*, *IRF4*, *EBF1*, *POU2F2*, *TBET*, and *LEF1* promote B cell commitment, differentiation, maintenance, and class-switch recombination	[79]
West Nile virus (human)	Microengraving	WNV-specific B cell clones: Memory B cells: IgA and IgG isotypes; Antigen-secreting cells Four novel WNV-neutralizing antibodies against WNV E protein	[80]

Table 2. (continued)

Disease (species)	Single cell approach	B cell signatures observed	Refs
Influenza virus (human)	scRNA-seq; single cell antibody heavy and light chain sequencing	*Vaccine-induced transcriptional responses* *TCL1A naïve B cells: IL4R and IL21R* CD11c^{+}T-BET^{+} *memory B cell activation: CD86, AICDA, CXCR3*, *EBI3*, *DHRS9*, *ZEB2* *Upon affinity maturation, antibody clone with high affinity to hemagglutinin produced*	[81]
	scATAC-seq	Upon vaccination, CD20^{high} CD38^{int} IgD^{low} population expressing *BCL6* (GC B cells) found in draining lymph nodes Other markers for GC B cells include *CD24^{low}*, *CD27^{int}*, *CD71^{high}*, *CXCR5^{int}*, *Ki-67^{+}* Clonally diverse GC B cell responses due to high levels of SHM	[82]
Dengue (human)	scRNA-seq	Naïve B cells: ↑ *MX2* (IFN-response gene) causes progression to severe dengue (candidate biomarker) Major vRNA-containing naïve IgM B cells: *CD69*, *CXCR4*, *FCRL1*, *IRF1* Non-vRNA-containing B cells: IFN-stimulated genes *IFI6*, *IFI44L*, *IFIT3* Plasmablasts: heavily hypermutated IgG1	[44]
Coronavirus: SARS-CoV-2, MERS-CoV, SARS-CoV (human)	scRNA-seq	119 cell types identified in 13 human tissues B cell populations identified in lung, esophagus, stomach, spleen, BM, blood, liver, kidney, eye	[22]
Genital herpes virus infection (human)	scRNA-seq	Circulating memory B cells migrate to vaginal mucosa via CXCR3 chemotaxis and secrete virus-specific IgG2b, IgG2c, and IgA into lumen	[83]
Flavivirus infection (mouse)	scRNA-seq	Upon *re-encountering antigens from heterologous flaviviruses,* memory B cells expressing *Klf2*, *Zbtb32, Bach2*, *Il9r*, *Ccr6*, *Cd38*, and *S1pr1 differentiate into plasma cells or secondary GC B cells. The latter can become new affinity-matured plasma cells and memory B cells.*	[84]
Antibody-mediated rejection			
Kidney allograft (human)	scRNA-seq	One B cell cluster: few Ig genes, *CD86*, *ID01* Two plasma cell clusters: High Ig genes: *K-chain*-restricted; *K* and *L* gene types B cell plasma lineage is attractive target for therapy	[85]
Renal allograft (human)	scRNA-seq	Intrarenal class-switched B cells within graft result in rejection due to their: Innate cell transcriptional state (*TLR2*, *TLR7*, *HIF* family, *IL15*); Expression of migration and adhesion-related genes (*CD24*, *CD44*, *CD55*, *ITGA4*, *COL4A4*, *CCR6*, *VIM*, *SIPR1*, *SIPR2*)	[86]
Cancer			
Pancreatic cancer (human)	scRNA-seq; droplet-based sequencing	Detected presence of rare CD20^{+} CD19^{+} B cell population in both HGD^{-} intraductual papillary mucinous neoplasms (IPMNs) and LGD-IPMNs that were completely absent in PDAC	[87]
Lymphoma GC (human)	scRNA-seq; CITE-seq	GCB DLBCL: GCB in early LZ or intermediate LZa-like stages ABC DLBCL: late GCB stages in LZb and pre-M clusters. Contrary to current models, most ABC-DLBCLs originate from cells not yet committed to PBL differentiation and that correspond in ~25% of cases to pre-M B cells	[5]
	scRNA-seq	Intratumor heterogeneity of nodal B cell lymphomas in same patient confirmed via transcriptionally distinct malignant subpopulation (CREBBP-mutated cells had lower expression of *HLA-DR* within same tumor)	[6]
		Two populations of tumor-resident B cells: Healthy B cells; Malignant B cells: restricted Ig light chain expression (either *IGκ* or *IGλ*), ↑ *BCL2* (due to *IGH:BCL2* translocation), *FCER2*, *CD52*, and major histocompatibility complex class II genes	[88]
	scRNA-seq; transcriptome-informed flow cytometry	Resistance mechanism against ibrutinib occurs because cancerous B cells receive IL-4 from T cells as form of cell–cell communication signal (inferred from ligand–receptor prediction using Nichenet)	[4,48]
	scRNA-seq	Patients with diffuse large B cell lymphoma administered CD19 chimeric antigen receptor (CAR)-T cells face immune cell activation neurologic syndrome, because CD19-expressing mural cells that maintain the brain–blood barrier are killed by CAR-Ts	[89]

(continued on next page)

Table 2. (continued)

Disease (species)	Single cell approach	B cell signatures observed	Refs
Breast cancer (human)	scRNA-seq; flow cytometry; *myeloid-derived suppressor cells (*MDSCs) in MMTV-PyMT mouse	B cells had lowest MDSC signature score across all immune cells. CD84 serves as homophilic adhesion molecule on B cells and enables cell–cell interactions, activation, and differentiation	[90]
	scRNA-seq	In mouse models, PD1/CTLA4 antibody therapy resulted in expansion of B cells and unique cluster of IgG^+ plasma cells, which helped mediate ICB response. B cell gene expression: *MSA41*, *IGHD*, *CD37*, *H2-DMa*, *H2Oa*, *CD80*, *CD79a*, *EBF1*, *CD72*, *FCMR*, *IGHG1*, *IGHG2*, *IGLV1*, *JCHAIN*, *IGKC*, *XBP1*	[91]
		$ICOSL^+CR2$-high $IL\text{-}10^-CD20^+CD38^+CD27^+IgA^-IgD^-$ B cell subset emerges after chemotherapy that has properties of antitumor T cell immunity and improves therapeutic response by ICOSL	[92]
Nasopharyngeal carcinoma (human)	scRNA-seq	Seven clusters of B cells identified: B cell-1, B cell-2, B cell-3, B cell-4: *MS4A1, CD19, CD79A/B;* $FCRL4^+$ memory B cells: *FCRL4*, *CCR1*; *GC: AICDA, GCSAM, RGS13, BCL6. Activation of pathways*: *ROS detoxification, pyruvate metabolism, oxidative phosphorylation, apical junction, transportation;* *Plasma cells: CD38, SDC1, MZB1, ATF4, glycolysis pathway*	[93]
Lung cancer (human)	scRNA-seq	Co-existence of two B cell types in tumor microenvironment of non small cell lung cancer (NSCLC): Naïve B cells: *MS4A1* (*CD20*), *CD19*, *CD22*, *TCL1A*, *CD83*, ↑ *RACK1*, *JUND*, *CD83*, *ELOB*, *NFKB1A*, *APOE*, *GADD45B*. Reduced naïve B cells in advanced stages of NSCLC, because they inhibit proliferation; Plasma B cells: *CD38*, *TNFRSF17* (*BCMA*), *IGHG1/IGHG4*. $BCMA^+$ IgG^{high} B plasma cells promote cell proliferation in advanced stages of NSCLC Naïve B cells in PBMCs: ↑ *GNB2L1*, *GLTSCR2*, *NBEAL1*, *ALDOA*, *TCEB2*	[94]
		B cells significantly enriched among stromal cells in lung tumors Follicular B cells: ↑ *CD20* (*MS4A1*), *CXCR4*, *HLA-DRs* Plasma cells: IgG Mucosa-associated lymphoid tissue-derived (MALT) B cells: *IGA*, *IGM*, *JCHAIN* Tumor-associated follicular B cells: ↑ coagulation, *KRAS*, *INFA*, ↓*MYC*, *mTOR*, and protein secretion pathways; become exhausted	[95]
Uveal melanoma (human)	scRNA-seq	B cells (*CD19*, *CD79A*, *MS4A1* [*CD20*]) Plasma cells (*IGHG1*, *MZB1*, *SDC1*, *CD79A*) Clonally expanding plasma cells from class 1B UM infiltrate indolent liver metastasis, suggesting antibody-mediated immunity	[96]
Metastatic melanoma (human)	scRNA-seq	Presence of non-malignant B cells: *CD19*, *CD79A*, *CD79B*, *BLK*, *MS4A1*, *BANKL1*, *PAX5*	[97]
	CyTOF; scRNA-seq	Upon ICB therapy, potential predictive biomarkers for immunotherapy-induced irAEs include elevated number of activated memory B cells ($CD21^{low}$ IgD^-CD27^+ $PD1^+$) with enhanced IFN-γ signaling. Lower overall B cell count correlated with less time to onset and severity of irAEs	[51]
	scRNA-seq	TAB population (*CD20*, *CD19*, *CD5*, *CD27*, *CD38*, *CD138*) has signature gene expression Activation-associated genes (*CD69*, *CD72*, *CFLAR*, *FGFR1*, *SELPLG*, *CD86*) Co-stimulatory genes (*ICAM3*, *TNFRSF13C*, *CD40*, *CD72*, *C3*, *CD80*, *CD86*, *CD27*, *CD28*, *ICOS*, *TNFRSF9*, *CD40LG*, *ARHGDIB*) Proinflammatory genes (*TNF*, *IL12B*, *IL18*, *LTA*, *TNFAIP2*, *C3*, *HCK*) Immune checkpoint-associated genes (*CD274*, *PDCD1LG2*, *TNFRSF14/HVEM*, *LGALS9*, *BTLA*, *LAG3*, *HAVCR2/TIM-3*, *ADORA2A*) Immunosuppressive genes (*IL10*, *TGFB1*) B cell exhaustion-associated genes (*PDCD1*, *FCRL4*, *SIGLEC6*, *CD22*) Treg and macrophage chemoattractant genes (*CCL3*, *CCL3L1*, *CCL4*, *CCL5*, *CCL28*, *CXCL16*) Plasmablast-like rich TAB population: ↑CD69, *CD80*, *CD30* (*TNFRSF8*), and *CD137* (*4-1BB/TNFRSF9*) for NF-κB activation ↑ *CD27*, *PAX5*, *CD38*, ↓ *CD19*, *CD20*, *IFR4*, *CD71* Aids $PD\text{-}1^+$ T cell activation upon anti-PD-1 blockade and results in longer overall survival	[52]

Table 2. (continued)

Disease (species)	Single cell approach	B cell signatures observed	Refs
Multiple myeloma (MM; human)	scRNA-seq	MM cells: *CD56*, ↓ *CD27*, *CD45* MM circulating tumor cells and BM MM cells: ↑ *CD38*, *SLAMF7*, *BCMA* (therapeutically relevant molecules) Normal B cells: *CD19*, *CIITA*, *PAX5*	[98]
Acute lymphoblastic leukemia (ALL; human)	scRNA-seq	STING pathway expression in BM DC and B cells from patients with B-ALL based on single cell transcriptomes. BM B cells hardly expressed PD-L1 (*CD274*) and instead expressed *TMEM173* (STING), *NFKB1* (NF-κB), and *GSDMD*. Developmental trajectory of malignant B cells plotted using pseudo-time analysis; *GSDMD* expression gradually increased during development, suggesting STING agonist triggers pyroptosis of B cells in antitumor therapy	[21]
Cutaneous melanoma and renal cancer carcinoma (human)	scRNA-seq	Tertiary lymphoid structures contain tumor-infiltrating $CD20^+$ B cells with prognostic B cell-related gene signature enriched for *FCRL5*, *IDO1*, *IFNG*, and *BTLA* in patients responding to ICB therapy. Potential predictive biomarkers in B cells in response to therapy identified	[49,50]
HPV-associated squamous cell carcinomas (human and mouse)	scRNA-seq; BCR-seq	Prognostic B cell markers for 3-year overall survival: ↑ *CD19*, *IGJ*. Radiotherapy and PD-1 blockade induced development of memory B cells, plasma cells, and antigen-specific B cells and improved survival. BCR-seq revealed enhanced B cell clonality, SHM, and CDR3 length reduction induced upon radiotherapy	[99]
Others			
Atherosclerotic plaques (human)	scRNA-seq; scATAC-seq	Homogenous cluster of B cells expressing *CD79A*, *FCER2*, *CD22*, *CD79B* CXCR4 on B cells interacted with CXCL12 on My.1 cells	[100]
TB (human)	scRNA seq	Identified additional markers for B cells: *BANK1*, *VPREB3*, *FCER2*, *ADAM28* Identified five distinct B cell subsets at various states of B cell development in PBMCs: Follicular B cells (B1 and B5): ↑ *TCL1A*, *CD79A*, *CD79B*, *MS4A1*; ↓ *CD27 CD138* expression; Activated B cells (B2 and B4): ↑ *PMAIP1* (*NOXA*), *ABCA6*, *TCF4*; Mature B cells (B3): ↑ *MS4A1*, *CD79A*, *CRIP1*, *IGJ*; ↓ *TCL1A*, *PMAIP1*; Enrichment of exhausted B1, B2, and B4 cells in PBMCs from patients with TB compared with healthy controls and patients with latent TB.	[101]
Heart failure (human and mouse)	scRNA-seq	In standard mouse nonischemic, pressure-overload heart failure model, $CD79a^+$ $CD69^+CCR7^+$ B cells abundantly present in myocardium, subdivided into four clusters. Two clusters were activated and expanded in heart failure by upregulating *CD69*, *CD86*, *CXCR5*, *CCR7*	[102]
Thrombocytopenia (human)	scRNA-seq	Presence of RTX-resistant splenic $CD19^+$ memory B cells with downregulated genes for B cell-specific factors and proliferation $CD19^+$ memory B cells reactivated upon RTX clearing and differentiated into plasma cells and GC reactions	[103]
Waldenström's macroglobulinemia (human)	scRNA-seq; BCR-seq	Progenitor B cells in BM carrying MYD88 L265P and CXCR4 mutations can give rise to clones resulting in Waldenström's macroglobulinemia	[104]

during metastases, and study anticancer-targeted drug responses and toxicities. For instance, in age-associated B cell (ABC) diffuse large B cell lymphoma (DLBCL), scRNA-seq identified cells of origin as being late GC B and pre-M cells instead of cells that were committed to PBL differentiation, contrary to current models [5]. For B cell-related malignancies and B cell-induced antitumor immunity and resistance to immunotherapy, several studies also used scRNA-seq to characterize various B cell properties and/or candidate biomarkers (Table 1). In lymphomas, based on predictions of ligand–receptor binding from scRNA-seq analyses, resistance mechanisms against ibrutinib were deemed to be the result of cancerous B cells receiving IL-4 from T cells as a form of cell–cell communication signal [4,48]. In human cutaneous melanomas and renal cancer responses to **immune checkpoint blockade (ICB) therapy**, predictive biomarkers, such as *FCRL5*, *IDO1*, *IFNG*, and *BTLA*, were enriched and were characterized as constituting a prognostic

B cell-related gene signature in CD20$^+$ B cells within **tertiary lymphoid structures** [49,50]. In another study of advanced melanoma responses to ICB therapy, a potential predictive biomarker for immunotherapy-induced immune-related adverse events (irAEs) was the elevated number of activated memory B cells in human peripheral blood [51]. However, in response to anti-PD-1 antibody ICB for advanced human melanoma metastases, a plasmablast-like-rich tumor-associated B cell (TAB) population was shown to aid PD-1$^+$ T cell activation and promote longer overall patient survival [52]. This suggested context-dependent contradictory roles of B cells in progression versus suppression of cancers with different immunotherapies. In one study, tumor microenvironment (TME)-intrinsic immune cell responses were studied using air–liquid-interface patient-derived human tumor organoids (PDOs) and single cell technology [53]. Chromium immune cell profiling was performed on fresh human tumors and PDOs (melanoma and renal cell carcinoma) to generate gene transcriptome and immune repertoire analyses at a single cell resolution. PDOs recapitulated the human TMEs, as well as the TME architecture and immune cells present in fresh tumors [53]. This type of finding paves the way for more studies to investigate TME-intrinsic B cell responses, measure clinical outcomes, and monitor clinical responses to immunotherapies. Additionally, to better understand B cell roles in different TMEs, spatial transcriptomics combined with single cell methods should be performed. Other potential biomarkers and therapeutic B cell targets characterized by scRNA-seq in different cancers are listed in Table 2. Thus, dissecting the differential functional roles of various B cell populations in malignancies and/or other pathologies as well as the putative application of such biomarkers for clinical studies warrants robust investigation.

Concluding remarks

Single cell approaches are advancing our understanding of B cell properties and functions by providing information on genomics, transcriptomics, and proteomics during the development, differentiation, and production of antibody repertoires and the responses of these cell populations under homeostasis. Moreover, these approaches have been, and are, useful in detecting changes in B cell biology upon infection and aberrations during autoimmune diseases and cancer. Here, we have summarized examples of B cell signatures that have been uncovered under normal and disease conditions, and we encourage the future discovery of potential B cell biomarkers for investigative and therapeutic purposes. However, even though various groups have used single cell methods to study the phenotypes and functional roles of B cells, the spatial transcriptomics of major lymphoid organs and lymph nodes participating in B cell development and activation remain to be more thoroughly investigated (see Outstanding questions). One group has performed spatial transcriptomics on BM [13], but, to our knowledge, no analyses of B cell development or compartments have been carried out. The integration of scRNA-seq and spatial transcriptomics can help us understand B cell networks with other cells, in conditions of homeostasis versus disease states, and within different tissue architectures. Of note, because long noncoding RNAs (lncRNAs) and circular RNAs (circRNAs) are known to have important roles in NF-kB signaling (a signaling pathway that B cells heavily depend on for their development and activation), their role in B cell differentiation should be further assessed [54–56]; indeed, we posit that single cell approaches might be used to characterize circRNAs and lncRNAs that contribute to B cell signatures and functions and, hence, to ideally result in the future discovery of biological processes and putative therapeutic strategies in which B cells are relevant.

Outstanding questions

lncRNAs and circRNAs have important roles in signaling pathways required for B cell development and activation [54–56], and these transcripts are often missed out in current single cell approaches. Therefore, what future single cell discovery strategies might be devised to determine when these types of RNA are relevant in B cell biology?

Can other SHM-associated B cell-specific gene signatures in humans be identified using sc-WG-seq for cancer risk factors (e.g., environmental factors, pathogens, diet, and family history) and for other diseases, such as B cell-related autoimmune diseases? This might help build a list of SHM-associated B cell-specific gene signatures for future therapeutics.

Can spatial transcriptomics be used to precisely locate the newly identified B cell gray zone areas from scRNA-seq analyses of mouse GCs and identify B cell compartments and niches within various tissue structures of lymphoid and nonlymphoid organs?

Can single cell methods further analyze the trajectories of B cells in nonlymphoid organs, which are still not well understood? For instance, the source of dura B cells and other progenitor B cell in the mouse meninges remains unclear.

What can integrated scRNA-seq and spatial transcriptomics tell us about the location of gene transcripts of B cells and other immune cells, as well as cell–cell communication and interactions using bioinformatic tools, such as Nichenet and CellPhoneDB?

Can integrated scRNA-seq and spatial transcriptomics of the TME fully reveal the spatial distribution of heterogenous cell populations, including B cells with specific transcriptomic signatures? Collective information on intrinsic B cell properties in the TME might be correlated with clinical outcomes in which it might be better to monitor certain responses to immunotherapies.

Acknowledgments

We sincerely thank Eric Meffre for advice. We thank the Agency for Science, Technology and Research (A*STAR) and Singapore National Research Foundation (NRF-CRP17-2017-22) for funding and supporting the V.T. laboratory.

Declaration of interests

None declared by authors.

Resources

[i]https://tabula-muris.ds.czbiohub.org/

References

1. Hoffman, W. *et al.* (2016) B cells, antibodies, and more. *Clin. J. Am. Soc. Nephrol.* 11, 137–154
2. Zhang, F. *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
3. Hardt, U. *et al.* Integrated single cell and spatial transcriptomics reveal autoreactive differentiated B cells in joints of early rheumatoid arthritis. *medRxiv* Published online August 9, 2021. https://doi.org/10.1101/2021.08.09.21260015
4. Roider, T. *et al.* (2020) Dissecting intratumour heterogeneity of nodal B-cell lymphomas at the transcriptional, genetic and drug-response levels. *Nat. Cell Biol.* 22, 896–906
5. Holmes, A.B. *et al.* (2020) Single-cell analysis of germinal-center B cells informs on lymphoma cell of origin and outcome. *J. Exp. Med.* 217
6. Milpied, P. *et al.* (2018) Human germinal center transcriptional programs are de-synchronized in B cell lymphoma. *Nat. Immunol.* 19, 1013–1024
7. Efremova, M. *et al.* (2020) Immunology in the era of single-cell technologies. *Annu. Rev. Immunol.* 38, 727–757
8. Burrows, N. *et al.* (2020) Dynamic regulation of hypoxia-inducible factor-1α activity is essential for normal B cell development. *Nat. Immunol.* 21, 1408–1420
9. Hanamsagar, R. *et al.* (2020) An optimized workflow for single-cell transcriptomics and repertoire profiling of purified lymphocytes from clinical samples. *Sci. Rep.* 10, 1–15
10. Chovanec, P. *et al.* (2018) Unbiased quantification of immunoglobulin diversity at the DNA level with VDJ-seq. *Nat. Protoc.* 13, 1232–1252
11. Goldstein, L.D. *et al.* (2019) Massively parallel single-cell B-cell receptor sequencing enables rapid discovery of diverse antigen-reactive antibodies. *Commun. Biol.* 2, 1–10
12. Singh, M. *et al.* (2019) High-throughput targeted long-read single cell sequencing reveals the clonal and transcriptional landscape of lymphocytes. *Nat. Commun.* 10, 1–13
13. Baccin, C. *et al.* (2020) Combined single-cell and spatial transcriptomics reveal the molecular, cellular and spatial bone marrow niche organization. *Nat. Cell Biol.* 22, 38–48
14. Ranzoni, A.M. *et al.* (2021) Integrative single-cell RNA-seq and ATAC-seq analysis of human developmental hematopoiesis. *Cell Stem Cell* 28, 472–487
15. Zhang, L. *et al.* (2019) Single-cell whole-genome sequencing reveals the functional landscape of somatic mutations in B lymphocytes across the human lifespan. *Proc. Natl. Acad. Sci. USA* 116, 9014–9019
16. Skinnider, M.A. *et al.* (2019) Evaluating measures of association for single-cell transcriptomics. *Nat. Methods* 16, 381–386
17. Stubbington, M.J. *et al.* (2017) Single-cell transcriptomics to explore the immune system in health and disease. *Science* 358, 58–63
18. Nouri, N. and Kleinstein, S.H. (2020) Somatic hypermutation analysis for improved identification of B cell clonal families from next-generation sequencing data. *PLoS Comput. Biol.* 16, e1007977
19. Longo, S.K. *et al.* (2021) Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nat. Rev. Genet.* 22, 627–644
20. Rao, A. *et al.* (2021) Exploring tissue architecture using spatial transcriptomics. *Nature* 596, 211–220
21. Cai, Y. *et al.* (2020) Single-cell transcriptome reveals the role of sting pathway in acute B lymphoblastic leukemia bone marrow cells. *Blood* 136, 47–48
22. Qi, F. *et al.* (2020) Single cell RNA sequencing of 13 human tissues identify cell types and receptors of human coronaviruses. *Biochem. Biophys. Res. Commun.* 526, 135–140
23. Brioschi, S. *et al.* (2021) Heterogeneity of meningeal B cells reveals a lymphopoietic niche at the CNS borders. *Science* 373, eabf9277
24. Bendall, S.C. *et al.* (2014) Single-cell trajectory detection uncovers progression and regulatory coordination in human B cell development. *Cell* 157, 714–725
25. Tabula Muris Consortium (2018) Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature* 562, 367–372
26. Farmer, J.R. *et al.* (2019) Induction of metabolic quiescence defines the transitional to follicular B cell switch. *Sci. Signal.* 12, eaaw5573
27. Hammad, H. *et al.* (2017) Transitional B cells commit to marginal zone B cell fate by Taok3-mediated surface expression of ADAM10. *Nat. Immunol.* 18, 313–320
28. Zhao, Y. *et al.* (2018) Spatiotemporal segregation of human marginal zone and memory B cell populations in lymphoid tissue. *Nat. Commun.* 9, 1–15
29. Su, T.T. *et al.* (2004) Signaling in transitional type 2 B cells is critical for peripheral B-cell development. *Immunol. Rev.* 197, 161–178
30. Tull, T.J. *et al.* (2021) Human marginal zone B cell development from early T2 progenitors. *J. Exp. Med.* 218, e20202001
31. Cerutti, A. *et al.* (2013) Marginal zone B cells: virtues of innate-like antibody-producing lymphocytes. *Nat. Rev. Immunol.* 13, 118–132
32. Martin, F. and Kearney, J.F. (2002) Marginal-zone B cells. *Nat. Rev. Immunol.* 2, 323–335
33. Wilk, A.J. *et al.* (2020) A single-cell atlas of the peripheral immune response in patients with severe COVID-19. *Nat. Med.* 26, 1070–1076
34. Attaf, N. *et al.* (2021) Heterogeneity of germinal center B cells: New insights from single-cell studies. *Eur. J. Immunol.* 51, 2555–2567
35. Laidlaw, B.J. *et al.* (2020) The transcription factor Hhex cooperates with the corepressor Tle3 to promote memory B cell development. *Nat. Immunol.* 21, 1082–1093
36. Kennedy, D.E. *et al.* (2020) Novel specialized cell state and spatial compartments within the germinal center. *Nat. Immunol.* 21, 660–670
37. Schafflick, D. *et al.* (2021) Single-cell profiling of CNS border compartment leukocytes reveals that B cells and their progenitors reside in non-diseased meninges. *Nat. Neurosci.* 24, 1225–1234
38. Hay, S.B. *et al.* (2018) The Human Cell Atlas bone marrow single-cell interactive web portal. *Exp. Hematol.* 68, 51–61
39. Satpathy, A.T. *et al.* (2019) Massively parallel single-cell chromatin landscapes of human immune cell development and intratumoral T cell exhaustion. *Nat. Biotechnol.* 37, 925–936
40. Cildir, G. *et al.* (2017) The transcriptional program, functional heterogeneity, and clinical targeting of mast cells. *J. Exp. Med.* 214, 2491–2506
41. Akıncılar, S.C. *et al.* (2015) Quantitative assessment of telomerase components in cancer cell lines. *FEBS Lett.* 589, 974–984
42. Cildir, G. *et al.* (2019) Genome-wide analyses of chromatin state in human mast cells reveal molecular drivers and mediators of allergic and inflammatory diseases. *Immunity* 51, 949–965
43. Cao, Y. *et al.* (2020) Potent neutralizing antibodies against SARS-CoV-2 identified by high-throughput single-cell sequencing of convalescent patients' B cells. *Cell* 182, 73–84
44. Zanini, F. *et al.* (2018) Virus-inclusive single-cell RNA sequencing reveals the molecular signature of progression to severe dengue. *Proc. Natl. Acad. Sci. USA* 115, E12363–E12369

45. Castro-Dopico, T. *et al.* (2020) Targeting B cells for inflammatory bowel disease treatment: back to the future. *Curr. Opin. Pharmacol.* 55, 90–98
46. Gudjonsson, J.E. *et al.* (2020) Contribution of plasma cells and B cells to hidradenitis suppurativa pathogenesis. *JCI insight* 5, e139930
47. Carlberg, K. *et al.* (2019) Exploring inflammatory signatures in arthritic joint biopsies with spatial transcriptomics. *Sci. Rep.* 9, 1–10
48. Davidson, S. *et al.* (2020) Single-cell RNA sequencing reveals a dynamic stromal niche that supports tumor growth. *Cell Rep.* 31, 107628
49. Gibellini, L. *et al.* (2020) Single-cell approaches to profile the response to immune checkpoint inhibitors. *Front. Immunol.* 11, 490
50. Helmink, B.A. *et al.* (2020) B cells and tertiary lymphoid structures promote immunotherapy response. *Nature* 577, 549–555
51. Das, R. *et al.* (2018) Early B cell changes predict autoimmunity following combination immune checkpoint blockade. *J. Clin. Invest.* 128, 715–720
52. Griss, J. *et al.* (2019) B cells sustain inflammation and predict response to immune checkpoint blockade in human melanoma. *Nat. Commun.* 10, 1–14
53. Neal, J.T. *et al.* (2018) Organoid modeling of the tumor immune microenvironment. *Cell* 175, 1972–1988
54. Mirzaei, S. *et al.* (2021) Regulation of nuclear factor-kappa b (NF-κB) signaling pathway by non-coding RNAs in cancer: Inhibiting or promoting carcinogenesis? *Cancer Lett.* 509, 63–80
55. Akıncılar, S.C. *et al.* (2021) NAIL: an evolutionarily conserved lncRNA essential for licensing coordinated activation of p38 and NFκB in colitis. *Gut* 70, 1857–1871
56. Morgan, D. *et al.* (2020) Pharmacological significance of the non-canonical NF-κB pathway in tumorigenesis. *Biochim. Biophys. Acta BBA Rev. Cancer* 1874, 188449
57. Melchers, F. (2015) Checkpoints that control B cell development. *J. Clin. Invest.* 125, 2203–2210
58. Descatoire, M. *et al.* (2014) Identification of a human splenic marginal zone B cell precursor with NOTCH2-dependent differentiation properties. *J. Exp. Med.* 211, 987–1000
59. Craft, J.E. (2012) Follicular helper T cells in immunity and systemic autoimmunity. *Nat. Rev. Rheumatol.* 8, 337–347
60. De Silva, N.S. and Klein, U. (2015) Dynamics of B cells in germinal centres. *Nat. Rev. Immunol.* 15, 137–148
61. Pikor, N.B. *et al.* (2020) Remodeling of light and dark zone follicular dendritic cells governs germinal center responses. *Nat. Immunol.* 21, 649–659
62. Rodda, L.B. *et al.* (2018) Single-cell RNA sequencing of lymph node stromal cells reveals niche-associated heterogeneity. *Immunity* 48, 1014–1028
63. Cole, C. *et al.* (2018) Tn5Prime, a Tn5 based 5′ capture method for single cell RNA-seq. *Nucleic Acids Res.* 46, e62
64. Attaf, N. *et al.* (2020) FB5P-seq: FACS-based 5-prime end single-cell RNA-seq for integrative analysis of transcriptome and antigen receptor repertoire in B and T cells. *Front. Immunol.* 11, 216
65. Lindeman, I. *et al.* (2018) BraCeR: B-cell-receptor reconstruction and clonality inference from single-cell RNA-seq. *Nat. Methods* 15, 563–565
66. Rizzetto, S. *et al.* (2018) B-cell receptor reconstruction from single-cell RNA-seq with VDJPuzzle. *Bioinformatics* 34, 2846–2847
67. Upadhyay, A.A. *et al.* (2018) BALDR: a computational pipeline for paired heavy and light chain immunoglobulin reconstruction in single-cell RNA-seq data. *Genome Med.* 10, 1–18
68. Samir, J. *et al.* (2020) Exploring and analysing single cell multi-omics data with VDJView. *BMC Med. Genet.* 13, 1–9
69. Canzar, S. *et al.* (2017) BASIC: BCR assembly from single cells. *Bioinformatics* 33, 425–427
70. Setliff, I. *et al.* (2019) High-throughput mapping of B cell receptor sequences to antigen specificity. *Cell* 179, 1636–1646
71. MacParland, S.A. *et al.* (2018) Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. *Nat. Commun.* 9, 1–21
72. Nehar-Belaid, D. *et al.* (2020) Mapping systemic lupus erythematosus heterogeneity at the single-cell level. *Nat. Immunol.* 21, 1094–1106
73. Pyfrom, S. *et al.* (2021) The dynamic epigenetic regulation of the inactive X chromosome in healthy human B cells is dysregulated in lupus patients. *Proc. Natl. Acad. Sci. USA* 118, e2024624118
74. Roy, B. *et al.* (2017) High-throughput single-cell analysis of B cell receptor usage among autoantigen-specific plasma cells in celiac disease. *J. Immunol.* 199, 782–791
75. Feng, B. *et al.* (2020) Immunosuppressive effects of mesenchymal stem cells on lung B cell gene expression in LPS-induced acute lung injury. *Stem Cell Res Ther* 11, 1–9
76. Wen, W. *et al.* (2020) Immune cell profiling of COVID-19 patients in the recovery stage by single-cell sequencing. *Cell Discov.* 6, 1–18
77. Guo, C. *et al.* (2020) Single-cell analysis of two severe COVID-19 patients reveals a monocyte-associated and tocilizumab-responding cytokine storm. *Nat. Commun.* 11, 1–11
78. Zhu, L. *et al.* (2020) Single-cell sequencing of peripheral mononuclear cells reveals distinct immune response landscapes of COVID-19 and influenza patients. *Immunity* 53, 685–696
79. You, M. *et al.* (2021) Single-cell epigenomic landscape of peripheral immune cells reveals establishment of trained immunity in individuals convalescing from COVID-19. *Nat. Cell Biol.* 23, 620–630
80. Tsioris, K. *et al.* (2015) Neutralizing antibodies against West Nile virus identified directly from human B cells by single-cell analysis and next generation sequencing. *Integr. Biol.* 7, 1587–1597
81. Horns, F. *et al.* (2020) Memory B cell activation, broad anti-influenza antibodies, and bystander activation revealed by single-cell transcriptomics. *Cell Rep.* 30, 905–913
82. Turner, J.S. *et al.* (2020) Human germinal centres engage memory and naive B cells after influenza vaccination. *Nature* 586, 127–132
83. Oh, J.E. *et al.* (2019) Migrant memory B cells secrete luminal antibody in the vagina. *Nature* 571, 122–126
84. Wong, R. *et al.* (2020) Affinity-restricted memory B cells dominate recall responses to heterologous flaviviruses. *Immunity* 53, 1078–1094
85. Wu, H. *et al.* (2018) Single-cell transcriptomics of a human kidney allograft biopsy specimen defines a diverse inflammatory response. *J. Am. Soc. Nephrol.* 29, 2069–2080
86. Asano, Y. *et al.* (2021) Innate-like self-reactive B cells infiltrate human renal allografts during transplant rejection. *Nat. Commun.* 12, 1–15
87. Bernard, V. *et al.* (2019) Single-cell transcriptomics of pancreatic cancer precursors demonstrates epithelial and microenvironmental heterogeneity as an early event in neoplastic progression. *Clin. Cancer Res.* 25, 2194–2205
88. Andor, N. *et al.* (2019) Single-cell RNA-Seq of follicular lymphoma reveals malignant B-cell types and coexpression of T-cell immune checkpoints. *Blood* 133, 1119–1129
89. Parker, K.R. *et al.* (2020) Single-cell analyses identify brain mural cells expressing CD19 as potential off-tumor targets for CAR-T immunotherapies. *Cell* 183, 126–142
90. Alshetaiwi, H. *et al.* (2020) Defining the emergence of myeloid-derived suppressor cells in breast cancer using single-cell transcriptomics. *Sci. Immunol.* 5, eaay6017
91. Hollern, D.P. *et al.* (2019) B cells and T follicular helper cells mediate response to checkpoint inhibitors in high mutation burden mouse models of breast cancer. *Cell* 179, 1191–1206
92. Lu, Y. *et al.* (2020) Complement signals determine opposite effects of B cells in chemotherapy-induced immunity. *Cell* 180, 1081–1097
93. Chen, Y.-P. *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
94. Chen, J. *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, 1–21
95. Lambrechts, D. *et al.* (2018) Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat. Med.* 24, 1277–1289

96. Durante, M.A. *et al.* (2020) Single-cell analysis reveals new evolutionary complexity in uveal melanoma. *Nat. Commun.* 11, 1–10
97. Tirosh, I. *et al.* (2016) Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 352, 189–196
98. Lohr, J.G. *et al.* (2016) Genetic interrogation of circulating multiple myeloma cells at single-cell resolution. *Sci. Transl. Med.* 8, 363ra147
99. Kim, S.S. *et al.* (2020) B cells improve overall survival in HPV-associated squamous cell carcinomas and are activated by radiation and PD-1 blockade. *Clin. Cancer Res.* 26, 3345–3359
100. Depuydt, M.A. *et al.* (2020) Microanatomy of the human atherosclerotic plaque by single-cell transcriptomics. *Circ. Res.* 127, 1437–1455
101. Cai, Y. *et al.* (2020) Single-cell transcriptomics of blood reveals a natural killer cell subset depletion in tuberculosis. *EBioMedicine* 53, 102686
102. Martini, E. *et al.* (2019) Single-cell sequencing of mouse heart immune infiltrate in pressure overload-driven heart failure reveals extent of immune activation. *Circulation* 140, 2089–2107
103. Crickx, E. *et al.* (2021) Rituximab-resistant splenic memory B cells and newly engaged naive B cells fuel relapses in patients with immune thrombocytopenia. *Sci. Transl. Med.* 13, eabc3961
104. Rodríguez, S. *et al.* (2020) Waldenström's macroglobulinemia (WM) is preceded by clonal lymphopoiesis including MYD88 L265P in progenitor B cells. *Blood* 136, 42–43