Review

B-cell receptor repertoire sequencing: Deeper digging into the mechanisms and clinical aspects of immune-mediated diseases

Bohao Zheng,[1,2,3] Yuqing Yang,[2,3] Lin Chen,[2,3] Mengrui Wu,[2] and Shengtao Zhou[2,*]

SUMMARY

B cells play an essential role in adaptive immunity and are intimately correlated with pleiotropic immune-mediated diseases. Each B cell occupies a unique B cell receptor (BCR), and all BCRs throughout our body form "BCR repertoire." With the development of sequencing technology and coupled bioinformatics, accumulating evidence indicates that BCR repertoire largely varies under physiological and pathological conditions. Therefore, comprehensive grasp of BCR repertoire will provide new insights into the pathogenesis of immune-mediated diseases and help exploit efficient diagnostic and treatment strategies. In this review, we start with an overview of BCR repertoire and related sequencing technologies and summarize their current applications in immune-mediated diseases. We also underscore the challenges of this emerging field and propose promising future directions in advancing BCR repertoire exploration.

INTRODUCTION

In order to recognize a wide variety of antigens from the outside world, individuals must generate a large number of B cell receptors (BCRs). The tremendous variation of BCRs is generated through the rearrangement of variable (V), diversity (D), and joining (J) genes. BCR has three complementarity determining regions (CDRs), namely CDR1, CDR2, and CDR3, among which diversity mainly arises from complementarity determining region 3 (CDR3) (Pineda et al., 2019). Advances in high-throughput sequencing continue to improve our understanding of diseases, especially in the application of BCR repertoire (Yaari and Kleinstein, 2015). In recent years, BCR repertoire sequencing has been reported to be applied to dozens of immune-mediated diseases, especially in systemic lupus erythematosus (SLE) (Bashford-Rogers et al., 2019; Liu et al., 2017; Zheng et al., 2021), COVID-19 (Bieberich et al., 2021; Galson et al., 2020; Jin et al., 2021; Montague et al., 2021; Niu et al., 2020; Paschold et al., 2021; Schultheiβ et al., 2020; Xiang et al., 2022; Zhang et al., 2020a, 2022; Zhou et al., 2021) and vaccination (Lee et al., 2016a; Miyasaka et al., 2019; Schneikart et al., 2020; Strauli and Hernandez, 2016; Zhao et al., 2021a, 2021b). With the advent of bioinformatics and the development of a variety of bioinformatics analysis tools, we have grasped a more comprehensive understanding of diseases, and meanwhile, disease diagnosis and treatment have become increasingly diversified (Liu et al., 2021).

B CELL RECEPTOR ASSEMBLY AND REPERTOIRE DIVERSIFICATION

BCRs are composed of membrane-bound immunoglobulins that can identify and bind to specific antigens (Hoehn et al., 2016). The diversification of BCR is formed by frequent somatic changes of germline DNA, which generates unique sequences and facilitates our body's defense against all different kinds of antigens (Nielsen and Boyd, 2018). BCR sequencing has raised our awareness of the diversity and dynamics of BCR repertoires, offering novel insights into the development of diagnostic and therapeutic strategies in immune-mediated diseases. In this section, we primarily introduce the structure and variety of human BCR repertoires and emphasize its major functions in fighting against exogenous or pathogenic stimuli.

BCRs constitute important mediators in adaptive immune responses, and they are heterodimers composed of two immunoglobulin light chains (IgLs) and two immunoglobulin heavy chains (IgHs) (Minervina et al., 2019). The IgH and IgL of BCRs are produced in sequence, controlled by sequential gene rearrangements, at subsequent cellular stages of lymphocyte development (Melchers, 2005). IgH and IgL comprise a constant region and variable region, and three constant regions are included in IgH and one

[1]Wuxi School of Medicine, Jiangnan University, Wuxi, P. R. China

[2]Department of Obstetrics and Gynecology, Key Laboratory of Birth Defects and Related Diseases of Women and Children of MOE and State Key Laboratory of Biotherapy, West China Second Hospital, Sichuan University and Collaborative Innovation Center, Chengdu, P. R. China

[3]These authors contributed equally

*Correspondence: shengtaozhou@scu.edu.cn, taotaovip2005@163.com

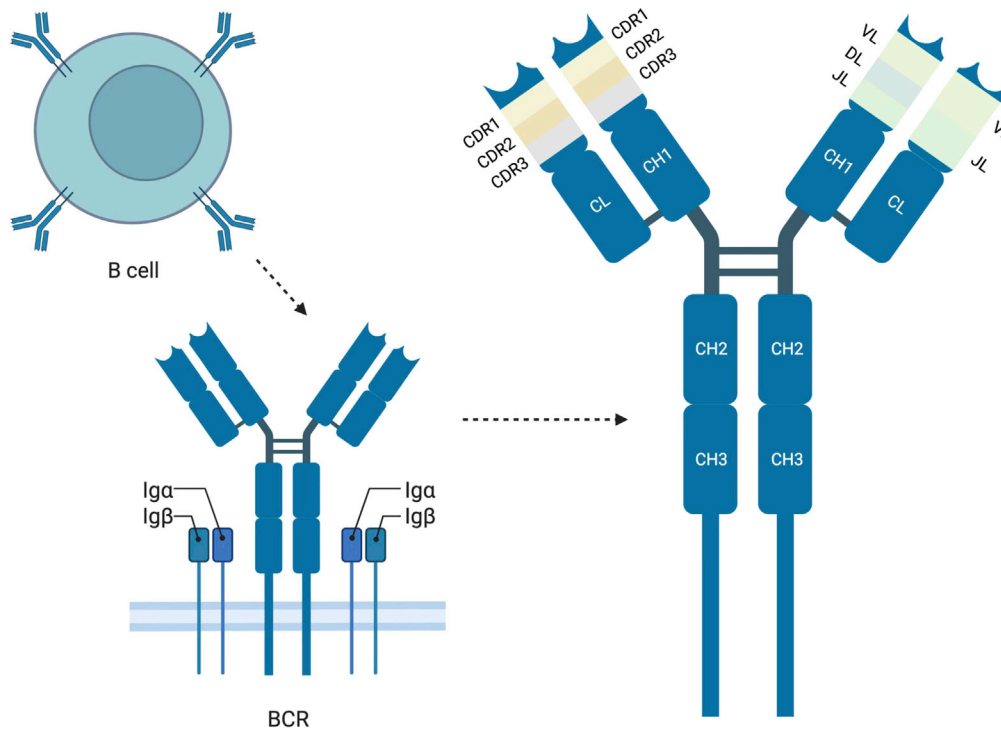https://doi.org/10.1016/j.isci. 2022.105002

Figure 1. The general structure of BCR

in IgL, which determines the species of produced Igs and coupled Igκ or Igγ of IgL with IgH, respectively. The general structure of BCR has been summarized in Figure 1. Notably, the variable region is the basis of antigen binding and adaptive immune defenses.

Somatic recombination or V(D)J recombination (VDJ recombination for IgH and VJ recombination for IgL) is the basis of BCR repertoire diversification. The variable (V), diversity (D), and joining (J) segments locus on the long arm (q) of chromosome 14 perform gene arrangements to form a specific exon of IgH, and simultaneously the V and J segments locus on the short arm (p) of chromosome 2 (for Igκ) and 22 (Igγ) perform gene arrangements to generate IgL. There are multiple open reading frames (ORFs) in V, D, and J sequence. V composes of 44 segments, 25 for D and 6 for J to encode IgH (Li et al., 2004). Clearly, B cells can never take full advantage of all V, D, and J segments. Instead, only one segment of V, D, and J can be selected to form a particular exon to finish BCR assembly under physiological and pathological conditions. Generally, during cell development, different B cells exhibit diverse BCR patterns, controlled mainly by gene arrangements, and the total number of BCR is predicted by a model with results of more than 10^{18} (Elhanati et al., 2015). The somatic recombination perfectly couples variable sequence regions that determine specific antigen identification (the complementarity determining regions; CDRs) with relatively conserved framework regions (FWRs) to form a complete variable region of IgH.

In addition to this natural process, random deletion or insertion of nucleotides in segment junctions can also contribute to final BCR repertoire diversity (Hoehn et al., 2016). After V(D)J recombination, progenitor B cells switch into mature B cells and enter peripheral blood from the bone marrow. When experiencing foreign stimuli, naïve B cells will begin cell division and somatic hypermutation (SHM) to generate B cell subclones with various BCRs and confront different threats. SHM alters the BCR sequence by diversifying constant regions and variable regions on IgL and IgH. Class-switch recombination occurs during the changes of constant regions in IgH, which facilitates the production of multiple Igs with specific functions and high affinity, such as IgM, IgG, IgA, IgD, and IgE. It is estimated that SHM involves athe dramatically increased mutation rate, which generates approximately one mutation per B cell subclone in the relevant locus (Teng and Papavasiliou, 2007; Victora and Nussenzweig, 2012). The alterations in variable regions lay the foundation of antigen-binding properties, allowing the immune system to recognize and respond to foreign antigens.
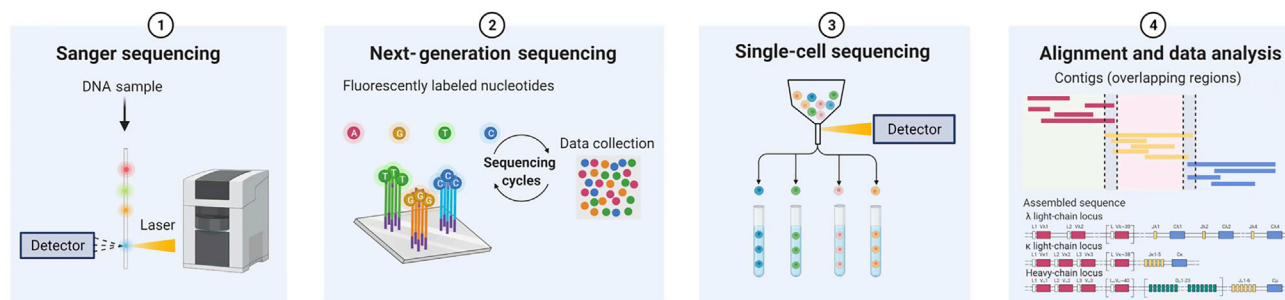
Figure 2. The workflow of three mainstream BCR-seq strategies
2.1 Sanger sequencing; 2.2 Next-generation sequencing; 2.3 Single-cell sequencing; 2.4 Alignment and data analysis.

Under certain pathological conditions, BCR repertoire varies dramatically compared with physiological status, helping us improve our knowledge of immune-mediated diseases. A typical example is that B-cell lymphoma may only exhibit a single BCR repertoire owing to the proliferation of identical B cell sub-clones, leading to a better understanding of the disease. Therefore, developing BCR repertoire sequencing technology will greatly help establish a comprehensive landscape of BCR repertoire under physiological and pathological conditions.

OVERVIEW OF B CELL RECEPTOR REPERTOIRE SEQUENCING: MAINSTREAM TECHNOLOGY AND PERFORMANCE COMPARISON

With the advances in sequencing technology, several different strategies for BCR repertoire sequencing have emerged, including Sanger sequencing, next-generation sequencing, and single-cell sequencing (Figure 2). Herein, we briefly introduce different technologies of BCR repertoire sequencing and summarize their respective performance.

Sanger sequencing

Sanger sequencing has been used in antibody display library analysis isolating primary clones and is still the gold standard with regard to DNA sequencing for clinical applications (Tiller et al., 2009). This technique is widely used in B cells or CDR3 spectratyping, only allowing access to the fundamental information of BCR repertoire. As early as 20 years ago, it was used for sequencing Ig heavy-chain genes of chronic lympho-blastic leukemia patients (Rosenwald et al., 2001). In addition, the fusion of BCR e14 and ABL a3 with or without any extra insertions or deletions could be learned through cloning and Sanger sequencing (Cai et al., 2018). Clinically, Sanger sequencing has been widely used to detect mutations such as BCR-ABL1 mutations associated with treatment resistance for a long time (Deininger et al., 2016). However, even though Sanger sequencing is already greatly improved in its first generation, the throughput is still low as large DNA fragments cannot be sequenced quickly.

Next-generation sequencing

By contrast, several next-generation sequencing platforms are available, capable of reading up to 10^7 base pairs, which are more cost-effectively and faster than traditional Sanger sequencing (Metzker, 2010). NGS platforms enable detailed examination of T and B cell receptors (TCRs and BCRs) at the nucleotide level (Carlson et al., 2013; Six et al., 2013). It is now widely used to analyze the immune repertoire because NGS can sequence millions of V(D)J sequences simultaneously, shedding light on the assessment of diver-sity, distribution, and mutation rate of BCR genes across isotypes. Several main applications of NGS in im-munogenetics include: clonality assessment, detection of minimal residual disease (MRD), and repertoire analysis of BCR IG and TR gene sequences (Rawstron et al., 2016; Rodriguez-Vicente et al., 2017; Vardi et al., 2017). However, owing to the limitation in read length and PCR-based amplification, novel gross chromo-somal aberrations are hard to recognize among enormous sequencing data.

Single-cell sequencing

In recent years, single-cell RNA sequencing technology (scRNA-seq) has been developed, which enables the analysis of RNA expression differences between individual cells. It has many advantages over bulk RNA

sequencing (Islam et al., 2011; Tang et al., 2009), providing full-length sequences of both IgH and IgG genes as a byproduct (Stubbington et al., 2016). For a long time, the lack of necessary strategies for assembling BCR sequences from scRNA-seq experiments is a persistent obstacle to B-cell biology advancement. Recent strategies focus on the reconstruction of immune repertoire from the scRNA-seq data (Afik et al., 2017, 2019; Rizzetto et al., 2018). Stefan et al. developed the BASIC (BCR assembly from single cells) method, which determines the BCR full-length sequence in B cells from single scRNA-seq data. They performed in nearly 200 individual human B cells, performed full-length heavy and light chains assembling, and confirmed the result by nested PCR and Sanger sequencing experiments based on single-cell primers (Canzar et al., 2017).

Third-generation sequencing, which is also referred to as single molecule sequencing, for instance, PacBio and Oxford Nanopore, produces longer sequence reads than traditional NGS and has developed to impressive levels in combination with scRNA-seq (Magi et al., 2018; Rang et al., 2018). It has been found to be useful for determining the diverse landscape of isoforms, as full-length transcripts can be obtained to provide information about single-nucleotide polymorphisms, and allow VDJ region assembly of T and B cell receptor sequences (Haque et al., 2017; James et al., 2020; Vieira Braga et al., 2019). Lira et al. developed a robotic protocol for full-length scRNA sequencing, which is a combination of the Smart-SeQ2 protocol and a commercial kit-based workflow (Mamanova et al., 2021). Through mapping pairs of heavy and light chain BCR sequences with their homologous antigen specificity, Ian and coworkers reported LIBRA-Seq (linking B cell receptors to antigen specificity by sequencing), which was used in mapping antigen specificity of thousands of B cells from two patients with HIV against masses of antigen targets (Setliff et al., 2019). Noudjoud et al. used FB5P-seq, a 5′ end scRNA-seq workflow, to generate the single-cell gene count matrix and the single-cell repertoire sequence of BCRs and TCRs in the analyses of FACS-sorted B cells or T cells (Attaf et al., 2020).

BIOINFORMATIC PIPELINES AND REPERTOIRE ANALYSIS: COUPLING WITH CURRENT B CELL RECEPTOR REPERTOIRE SEQUENCING

High-throughput sequencing of BCR sequencing produces vast, highly complex datasets that interrogate specialized analysis strategies. Several existing tools, such as IMGT/V-Quest (Alamyar et al., 2012) and IgBLAST (Ye et al., 2013) are commonly used for alignment against V(D)J germline databases, CDR3 identification, and characterization (Table 1).

Sara et al. (D'Angelo et al., 2014) proposed a free open-source application, antibody mining toolbox, which is able to analyze antibody repertoires and identify heavy chain CDR3s quickly and easily. An open-source platform called Vidjil was also developed for analyzing high-throughput sequencing reads from BCR IG gene rearrangements (Duez et al., 2016). Several ways are proposed on the platform to utilize complementary software (IMGT/V-Quest, IgBLAST, MiXCR). IMonitor (Zhang et al., 2015), which identifies V(D)J genes and alleles after common local alignment, is unique among other BCR repertoire monitoring programs as it can export comprehensive statistics and graphs. MiXCR is another software application for the identification of BCR repertoire profiles (Bolotin et al., 2015) and is designed to track BCRs somatic hypermutations.

However, transforming the resulting germline-annotated repertoire into the research spotlight is a tedious, error-prone, and time-consuming process. Change-O (Gupta et al., 2015) was introduced with the integration of a series of complex analyses, including identifying new Ig variable regions carried by individuals, dividing them into clonal populations, performing pedigree trees, inferring target models of somatic hypermutations, measuring system diversity, quantifying selection pressure, and calculating sequence chemistry. Automated novel germline alleles detection and genotypes inference can be applied using TIgGER for sequences with detectable junction sequences, which relies on the ability to assign BCR sequences using alleles from an IgGRdb (Gadala-Maria et al., 2015). Restricted to the output format and limitations of sequences number, the R package bcRep provides comprehensive analyses of the receptor repertoire of B cells (Bischof and Ibrahim, 2016). Input data for bcRep are tables generated by IMGT/HighV-QUEST. Additionally, VDJPipe (Christley et al., 2017) is a high-powered tool optimized for pre-processing enormous immune repertoire sequencing data. Processing times can be slashed as multiple processing steps are performed sequentially.

Several platforms combine the analysis of V(D)J genes, clonotypes, CDR3, diversity, somatic high mutation, and Ag together such as ARGalaxy (H et al., 2017). IGoR (Marcou et al., 2018) (Inference and

Table 1. Several specialized analysis tools coupled with BCR repertoire sequencing

Tools	Year	Language	Application	Reference
IMGT/HighVQuest	2012	–	a highly customized and integrated online tool for the standardized analysis of the BCR rearranged nucleotide sequences	(Alamyar et al., 2012)
IgBLAST	2013	C++	an Ig variable domain sequence analysis tool	(Ye et al., 2013)
MiXCR	2015	Java	a general framework that handles large immunome data from raw sequences to quantitated clonotypes	(Bolotin et al., 2015)
Change-O	2015	Python	a repertoire clonal assignment toolkit	(Gupta et al., 2015)
TIgGER	2015	R	an Ig Geno-type Elucidation tool via Rep-Seq	(Gadala-Maria et al., 2015)
IMonitor	2015	Perl/R	an efficient analytical pipeline to identify V(D)J genes and alleles	(Zhang et al., 2015)
Vidjil	2016	Java	an open-source tool for the interactive analysis of high-throughput sequencing data from lymphocyte recombination	(Duez et al., 2016)
bcRep	2016	R	a platform for advanced analyses of B cell receptor repertoires	(Bischof and Ibrahim, 2016)
VDJviz	2016	Java	a web-based tool that can be used for quality control analysis of Rep-Seq results	(Bagaev et al., 2016)
VDJPipe	2017	C++	a pipelined tool for pre-processing analysis of immune repertoire sequencing data	(Christley et al., 2017)
ClonoCalc and ClonoPlot	2017	R	ClonoCalc prepares the raw data for downstream analyses. ClonoPlot performs self-developed plots for the descriptive comparative investigation of immune repertoires.	(Fahnrich et al., 2017)
ARGalaxy	2017	R	a web-based tool for of BCR sequencing data analysis and visualization	(H et al., 2017)
IGoR	2018	C++	a comprehensive tool that takes BCR or TCR sequencing data and quantifies and characterizes receptor generation from both cDNA and gDNA	(Marcou et al., 2018)
sumrep	2019	R	an R package that takes a wide variety of repertoire summaries and comparisons, and performs model validation.	(Olson et al., 2019)
Platypus	2021	R	an open-source platform providing a user-friendly interface for B cell receptor investigation from scRNA-seq experiments	(Yermanos et al., 2021)

Generation of Repertoires) could obtain raw immunome sequence reads from cDNA or gDNA and characterize the statistics of V(D)J recombination and SHM. Sumrep (Olson et al., 2019) could summarize, compare, and visualize the summary distributions for BCR-seq datasets and process repertoire comparisons according to the summaries, and could also be performed in model validation. It can be challenging for researchers without computing expertise in BCR repertoire to analyze high-throughput sequencing datasets. In ClonoPlot and ClonoCalc (Fahnrich et al., 2017), scientists and less experienced users have access to a GUI, reducing barriers and providing assistance. Likewise, VDJviz (Bagaev et al., 2016) also implements a graphical user interface (GUI) for bioinformatics programs, simplifying the experience for end-users.

Recently, a growing number of bioinformatic tools emerged for the development of scRNA-Seq. Platypus (Yermanos et al., 2021), an open-source R package, with an automated pipeline for integrating scRNA-seq and transcriptome data, reveals clonal expansion of B cells that can identify different patterns of somatic high mutation, amino acid utilization, transcriptional heterogeneity, and clonal convergence. Using platypus, full-length BCR and TCR sequences can be acquired, which greatly accelerates the exploration of adaptive immunotherapy and gives quantitative insights into the interactions between the single-cell immune systems and heterogeneous gene expression.

B CELL RECEPTOR REPERTOIRE IN IMMUNE-MEDIATED DISEASES - DIAGNOSTIC AND THERAPEUTIC IMPLICATIONS

In view of the significance of BCR repertoire in physiological and pathological conditions, BCR repertoire sequencing has been applied in a variety of immune-mediated diseases (Figure 3). In this section, we primarily describe the functions of BCR repertoire sequencing in different diseases and emphasize the great potential of BCR repertoire sequencing in medical practice.
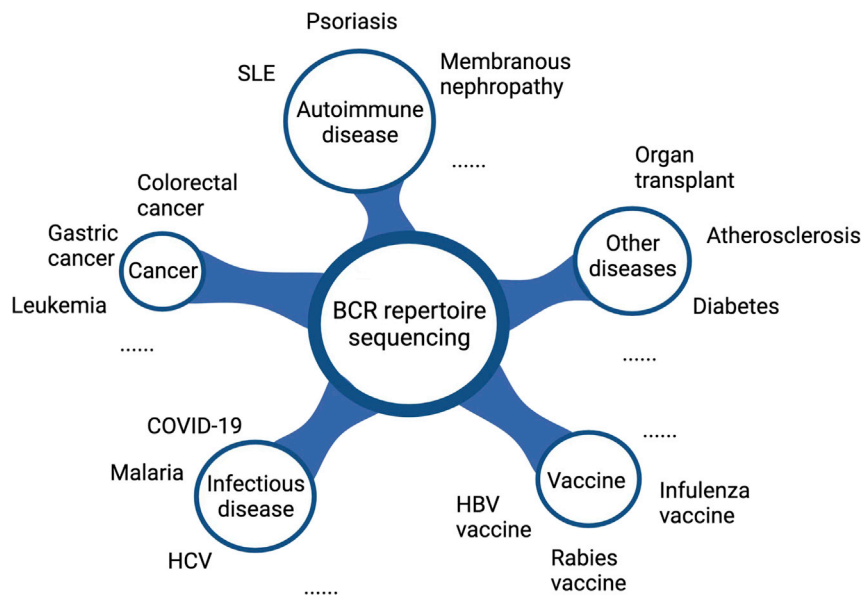
Figure 3. The application of BCR repertoire sequencing in immune-mediated diseases

Autoimmune diseases

Autoimmune diseases are a general term that includes at least 80 diseases, which occur increasingly frequently, approximately 20 million people in the US suffer hardships from autoimmune diseases and are predominantly female (Rose, 2016). With the advancement of sequencing technology and the emergence of bioinformatics, humans have a more comprehensive understanding of autoimmune diseases. Current diagnostic factors for systemic lupus erythematosus (SLE), as a common autoimmune disease, include ANA, anti-DSDNA, complement C3/4, blood routine, urine routine, 24-h urine protein, ESR, and CRP, but the performance is far from satisfactory (González et al., 2021). Increasing evidence suggests that B cells play a pivotal role in SLE development (Shlomchik et al., 2001). The emergence of BCR sequencing is a new direction for the clinical diagnosis of SLE (Liu et al., 2017). Using single-cell 5′ RNA sequence and single-cell BCR sequencing to analyze 20 samples (ten patients with SLE and the paired normal controls), It was found that patients with SLE had more BCR clonotypes than the normal control group, and the use of BCR V(D)J genes was different in the two groups (Zheng et al., 2021). Rogers et al. developed a new program by amplifying and sequencing BCR repertoires from RNAs that encode the antigen-binding (IgH (VDJ)) and constant regions of the BCR heavy chain. It was observed that significantly increased IgA levels and length of IgG and IgA CDR3 were found in patients with SLE compared with normal controls (Bashford-Rogers et al., 2019). The average length of CDR3 in the SLE group was significantly higher than that in the healthy control group and the ratio of CDR3 amino acid arginine in the SLE group was greatly increased (Liu et al., 2017). Systemic sclerosis (SSc) is another common autoimmune disease, and the advent of high-throughput sequencing and informational biology has led to a better understanding of this disease. The mean length of CDR3 in patients with SSc was significantly reduced and the diversity of CDR3 was significantly increased compared to normal controls (Shi et al., 2020). A similar situation has been observed in SSc patients with pulmonary arterial hypertension (SSc-PAH). Immunoglobulin heavy chain (IGH) sequencing suggests that the frequency of V(D)J rearrangement in patients with SSc-PAH has changed (de Bourcy et al., 2017).

Peanut allergy is a common type of allergy and the mechanism of its occurrence remains unclear. Currently, there are a variety of methods for the diagnosis of food allergy and some detection schemes are expensive. The emergence of BCR sequencing enriches the detection methods for food allergy (Ehlers et al., 2021). The new analysis platform constructed by using the NGS platform and algorithm "Bcrip" provides great convenience for the study of peanut allergy and other immune diseases (Kiyotani et al., 2018).

BCR sequencing has also been applied to other autoimmune diseases. Anti-*n*-methyl-*d*-aspartate receptor (NMDAR) encephalitis, an autoimmune disease discovered in recent years, has been found that the heavy

chain cloning of B cell receptor is common in most patients but not in healthy people (Feng et al., 2020). Mutations in IL10 or IL10 receptor cause infantile inflammatory bowel disease (IBD), and next-generation sequencing of TCR and BCR provides positive implications for the diagnosis of this disease (Werner et al., 2020). Membranous nephropathy (MN) is a common kidney disease, and high-throughput sequencing revealed that some of the indicators in patients with MN were significantly abnormal, such as CDR-H3, hydrophobicity, and somatic hypermutation (SHM). More importantly, these indicators demonstrate the potential to predict the prognosis of the disease (Su et al., 2021).

Cancer

Cancer has increasingly become a global burden. The latest data show that 19.29 million new patients with cancer merged globally in 2020 (Sung et al., 2021). Tumor diagnosis methods are diverse, including physical examination, hematology examination, imaging examination, pathological diagnosis, genetic diagnosis, and so forth. In addition, BCR repertoire sequencing is also increasingly applied to a variety of tumor diagnosis and treatment. RNA sequencing analysis of 936 different types of cancer cell lines and 462 Epstein-Barr virus (EBV) transformed normal B lymphocyte lines showed that lymphocyte-derived cell lines were more likely to have BCR/TCR rearrangement, and subclonicity or potential biclonicity was found in many blood cancer cell lines (Tan et al., 2018). Ultra-deep sequencing of BCR repertoire of tumor tissue, paratumoral normal tissue, and blood of 7 patients with esophageal squamous cell carcinoma revealed that the oligonicity of BCR spectrum in tumor tissue was significantly higher than that of adjacent normal tissue or peripheral blood, and the clonal amplification of B cells in multiple tumor regions had significant heterogeneity (Zhang et al., 2018). As the most common type of thyroid cancer, high-throughput sequencing provides a better understanding of the pathogenesis of Papillary thyroid carcinoma (PTC). Shannon's entropy showed a decreased diversity of PTC and a higher expression of high-amplification clone (HEC) of PTC than that of the peri-cancer group (Sun et al., 2018). Leukemia is a common malignant tumor of the blood system. The pathogenesis of leukemia is generally believed to be mainly related to environmental and genetic factors (Hutter, 2010). Previous study has found that B cells play a crucial role in the development of leukemia (Messmer et al., 2004). The emergence of BCR sequencing helps us better understand leukemia. At present, BCR sequencing focuses on the use preference of the VDJ gene, changes in CDR3 length, and SHM level. Patients with chronic lymphocytic leukemia (CLL) have longer CDR3 and higher SHM level than healthy controls (Petrova et al., 2018). Another study found that CLL-associated stereotypic BCR increased with age, which may account for the high incidence of CLL in the elderly population (Muggen et al., 2019). Single-cell sequencing revealed heterogeneity of CD5 expression on the surface of B cells in patients with CLL (Bashford-Rogers et al., 2017). Single-cell RNA-Seq revealed that CDR3 in acute myeloid leukemia (AML), another common leukemia, increased in length but decreased in diversity compared with normal controls (Zhang et al., 2019). Based on the development of a novel undifferentiated amplification of IgH repertoire, it was found that advanced adenoma (AD) has a IgH profile similar to that of adjacent normal mucosa. In patients with colorectal cancer (CRC), IgH distribution is slightly different. Repertoire properties different from AD and normal mucosa are found in CRC, which may have great potential significance for early diagnosis of CRC (Zhang et al., 2017). Furthermore, BCR repertoire is increasingly used to evaluate the drug efficacy of treatments, analysis of blood samples from 30 patients with NSCLC before and after anti-PD-1 antibody treatment for 6 weeks also showed that the changes in peripheral TCR and BCR spectrum diversity after anti-PD-1 treatment have clinical significance (Nakahara et al., 2021). A similar phenomenon was also found in animal experimental models, BCR repertoire diversity of tumor samples decreased significantly after drug treatment (Zhang et al., 2020b). Based on the combination of machine learning and next-generation sequencing, 90 patients with gastric cancer were analyzed. It was found that BCR immunoglobulin repertoire of gastric cancer tissue had obvious sequence characteristics compared with normal gastric tissue, and this combination mode may provide new ideas for cancer diagnosis (Konishi et al., 2019).

Infectious diseases

Infectious diseases refer to diseases caused by a variety of biological pathogens (e.g. bacteria, viruses, fungi, parasites, and so forth) that attack or infect the human body. The novel Coronavirus outbreak in 2020 has once again attracted intensive attention to these diseases. The emergence of BCR repertoire sequencing enables us to have a clearer understanding of this kind of disease. Adaptive immune response is crucial to clear viral infection, and BCR repertoire has been widely used to investigate COVID-19 pathology. Single-cell BCR repertoire sequencing revealed a significant decrease in BCR diversity in patients with COVID-19 and a higher frequency of highly amplified BCR clones in younger patients. Another finding was

that patients with COVID-19 had different IGHV3 and IGHV4 rearrangements and CDR3H were significantly longer than healthy controls, but no similar phenomenon was observed in the light chain, these findings suggest that the heavy chain may play a more important role in virus clearance. (Jin et al., 2021). Single-cell sequencing of lymphocyte immune reserve and transcriptome from convalescent patients with COVID-19 of different ages revealed that clonally amplified B cells from young patients were predominantly IgA isotypes and had higher levels of somatic mutations in BCR compared with elderly patients (Bieberich et al., 2021). Mild, severe, convalescent, or retesting-positive (RTP) statuses may exist from COVID-19 infection to recovery. Early identification of patients with RTP by high-throughput sequencing can prevent re-transmission of novel coronavirus. At the same time, analysis of immune characteristics of different courses of disease may become an effective indicator of the prognosis of COVID-19 (Zhou et al., 2021). Detection of the immune status of patients with COVID-19 using next-generation sequencing technology found that IGH of patients' immune systems would change dramatically after the onset of symptoms. In addition, there was a significant increase in clone type overlap and lineage extension of B cell clones 2-3 weeks after onset, which is critical for the immune responses of B cells (Xiang et al., 2022). Single-cell RNA sequencing showed that patients with severe disease had a more unique V(D)J rearrangement than normal, moderate, and convalescence patients (Zhang et al., 2020a). The BCR repertoire varies among patients with different severities, and next-generation sequencing also shows that patients with high SHM may be at higher risk for severe disease states and require more stringent monitoring (Schultheiß et al., 2020). The next-generation sequencing technology may also be used to predict the recovery period of COVID-19. Expression profiles of TCRs and BCRs are distinct at different stages of the disease, especially the expression levels of T cell receptor alpha and beta (TRA and TRB). The progress of COVID-19 can be judged by the expression levels of these indicators (Niu et al., 2020). In addition, BCR repertoire sequencing also plays a role in other infectious diseases. The CDR3 length of memory B cells was significantly higher in patients with hepatitis C than in the normal control population, and the analysis of somatic mutations of the IGHV gene showed no significant difference in the level of memory B cell subsets mutation compared with the normal control (Tucci et al., 2018). Malaria is a global disease caused by infection with plasmodium parasites. A study shows that infection with Plasmodium alters the length of the CDR3 of the H chain and the preference of specific V genes, as well as the level of SHM in atypical memory B cells(atMBC) (Braddom et al., 2021). Similar phenomenon of increased B cell cloning and reduced BCR diversity was also observed in patients with pneumocystis infection (Sun et al., 2021).

Vaccine/antibody

Hepatitis B is a kind of viral hepatitis that does great harm to the human population. The emergence of hepatitis B vaccine can effectively avoid infection with hepatitis B virus. It was found that CDR3 diversity of the TCRβ chain was significantly increased but that of the BCR H chain was significantly decreased after the second dose of hepatitis B vaccination (Miyasaka et al., 2019), Zhao et al. also made a similar study (Zhao et al., 2021a). Rabies is a zoonotic disease caused by rabies virus. Once infected, the mortality rate is almost 100%, and the emergence of rabies vaccine has greatly avoided the occurrence of death. High-throughput technology gives us a better understanding of what happens to the immune system before and after vaccination, recent study has found that the CDR3 diversity of BCR is significantly reduced after vaccination, suggesting that vaccination can reduce the BCR nucleotide sequence amplified by cloning (Zhao et al., 2021b). High-throughput sequencing also revealed the specific mechanism by which tetravalent influenza vaccines provide additional benefits over trivalent influenza vaccines (Lee et al., 2016a). The immune responses of different age groups were also significantly different. IGH repertoires of young people who had received live yellow fever vaccine had higher selective strength than those of middle-aged people (Davydov et al., 2018). Existing study has, indeed, found that the diversity of BCR repertoire in the bone marrow and peripheral blood of the elderly presents a decreasing trend (Tabibian-Keissar et al., 2016)

Other diseases

In addition to the above-mentioned diseases, BCR repertoire sequencing gradually plays a crucial role in the research of other types of diseases. Organ transplantation is one of the few effective treatments for many end-stage diseases. Immune rejection is an important factor affecting the outcome of transplantation. The emergence of BCR repertoire sequencing is beneficial to the prediction and management of immune rejection after organ transplantation. It has been reported that the B cell diversity of non-progressors (NP), progressors with rejection (PR), and progressors with no rejection (PNR) after kidney transplantation is different, in which the diversity of NP increases. The diversity of PR decreased while the diversity of PNR remained unchanged (Pineda et al., 2019) Another study on acute rejection after kidney transplantation

found that BCR repertoire diversity decreased significantly after transplantation compared with that before transplantation (Lai et al., 2019). BCR seq is also used in the management of immune rejection after heart transplantation (Vollmers et al., 2015). Myasthenia gravis (MG) can be divided into MuSK-MG and AChR-MG clinically. Compared with healthy subjects, the physical and chemical properties of VH CDR3 were changed in AChR-MG and the distance of V-J segment in the MuSK-MG recombination sequence was shortened (Vander Heiden et al., 2017). DOCK8 immunodeficiency syndrome (DIDS) is an extremely rare immune deficiency disorder (Zhang et al., 2016), and NGS analysis showed no significant difference in the frequency of total IGHV, IGHD, and IGHJ gene usage between patients with DIDS and healthy individuals (Tang et al., 2019).

Cardiovascular disease is the main cause of death in human beings, and it is gradually becoming younger. BCR repertoire sequencing technology provides a new way to analyze the causes of diseases. CDR3 AA diversity was significantly lower in patients with acute myocardial infarction (AMI) and unstable angina (UA) than in healthy controls. Different patients with acute coronary syndrome (ACS) had different preferences for V and J genes (Weng et al., 2022). It was found that the diversity of B cell repertoire decreases significantly in atherosclerotic plaques, but common and dominant clones increase greatly (Zhang et al., 2021). Furthermore, in-stent restenosis (ISR) and type 2 diabetes mellitus (T2DM) both alter the diversity of the BCR repertoire and clonal distribution in patients with coronary artery disease (CAD) (Weng et al., 2021).

Hepatobiliary diseases are characterized by diversity and complexity, and the pathogenesis of some hepatobiliary diseases is still unknown. Next-generation sequencing provides a new avenue to explore the etiology of diseases. Acute chronic liver failure (ACLF) is a severe and fatal syndrome associated with immune disorders. Compared with healthy controls, patients with ACLF show more B cell clone amplification. In addition, BCR CDR3 V, D, J, and V-J gene segments were also used preferentially in patients with ACLF (Yan et al., 2019). Immunoglobulin G4 (IgG4)-associated cholangitis is a disease manifestation associated with IgG4, and IgG4$^+$ clones are more prevalent in patients with active IAC than in healthy controls. It also offers new insights into the diagnosis and treatment of patients with IAC (Maillette de Buy Wenniger et al., 2013). BCR repertoire sequencing has also been used to study cholestatic liver disease (Thapa et al., 2020).

End-stage renal disease (ESRD) is a disease associated with immune disorders, and high-throughput sequencing showed similar results. Compared with healthy controls, patients with ESRD had more B cell clones and skewed use of BCR CDR3 V, D, J, and V-J gene fragments (Wang et al., 2019). The Wiskott–Aldrich Syndrome (WAS) is an X chromosome-linked disease, and NGS revealed the use preference of IGHV, IGHJ, and IGHC in patients with WAS (O'Connell et al., 2014). Common variable immunodeficiency (CVID) is a primary immunodeficiency disorder in which B cell differentiation is impaired and immunoglobulin production is defective. But the results show that BCR repertoire does not find quality defects in formation or specification (van Schouwenburg et al., 2018). Myalgic encephalomyelitis (ME)/chronic fatigue syndrome (CFS) is a complex disease with unknown etiology associated with motor and nervous systems. Bansal et al. found that B cell subsets are altered in patients with ME/CFS (Bradley et al., 2013). Next-generation sequencing also revealed up-regulation of V, D, and J gene expression in B cells from patients with ME/CFS compared with healthy subjects (Sato et al., 2021). Lack of *RAG* gene will cause severe immune deficiency in patients and the Shannon's H index of IGH repertoire of patients with RAG deficiency will be lower than that of healthy people (Lee et al., 2016b). Dietary habits and environmental factors also affect B-cell repertoire (Nielsen et al., 2019; Pham et al., 2017). Likewise, the B cell repertoire also changes with age (H et al., 2016; Pickman et al., 2013).

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

BCR repertoire as a vital property of B cells takes an important part in the adaptive immune system and is the basis of identifying and defending against different pathogens. It covers wide-spectrum immunology information, which contributes to a better understanding of acquired immunity and immune-mediated diseases and provides a novel perspective to develop new-generation diagnostic and therapeutic strategies.

B-cell receptor repertoire sequencing, for its part, sprouting up over the decades, enables us to understand diseases from an unprecedented perspective. Coupled with advanced bioinformatics analysis, BCR repertoire sequencing has achieved great success in a wide spectrum of diseases accompanied by discovering specific immunological characteristics under physiological and pathological conditions. However, there

are still some issues that inevitably need to be pondered: (1) Uni-form standards: Different research groups use different sequencing platforms and data processing methods, thus there is an urgent need to develop an industry consensus to unify standards and pave the way for new-comers (Liu and Wu, 2018). (2) Clinical interpretation: Functional investigation of the sequencing data is still far less concerned, which need more specialists to devote their efforts. (3) Curb costs: Taking single-cell sequencing, which is pricey with all but a low total number of cells, makes itself limited in the development of this field.

We urgently need to profile a whole BCR repertoire of healthy populations that covers enough people from different backgrounds, such as age, sex, race, and living environment. Only when we fully understand how the BCR repertoire works in normal conditions can we better identify different diseases based on the BCR repertoire and exploit more efficient targeted treatments. Under this status, more advanced BCR reper-toire sequencing technologies can exert a more prominent role in helping overcome human diseases. Therefore, though some challenges remain, I believe that BCR repertoire sequencing is still promising and with technological advancements, more interest will be focused on this field and its diagnostic and therapeutic value will be revealed.

AUTHOR CONTRIBUTIONS

B.Z. and S.Z. completed the conceptual design. B.Z., Y.Y., and L.C. performed the literature review. M.W. created the table and figures. All authors contributed to writing and final editing of the article.

DECLARATION OF INTERESTS

No conflict of interest is declared.

REFERENCES

Afik, S., Raulet, G., and Yosef, N. (2019). Reconstructing B-cell receptor sequences from short-read single-cell RNA sequencing with BRAPeS. Life Sci Alliance 2, e201900371.

Afik, S., Yates, K.B., Bi, K., Darko, S., Godec, J., Gerdemann, U., Swadling, L., Douek, D.C., Klenerman, P., Barnes, E.J., et al. (2017). Targeted reconstruction of T cell receptor sequence from single cell RNA-seq links CDR3 length to T cell differentiation state. Nucleic Acids Res. 45, e148.

Alamyar, E., Duroux, P., Lefranc, M.P., and Giudicelli, V. (2012). IMGT((R)) tools for the nucleotide analysis of immunoglobulin (IG) and T cell receptor (TR) V-(D)-J repertoires, polymorphisms, and IG mutations: IMGT/V-QUEST and IMGT/HighV-QUEST for NGS. Methods Mol. Biol. 882, 569–604.

Attaf, N., Cervera-Marzal, I., Dong, C., Gil, L., Renand, A., Spinelli, L., and Milpied, P. (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.

Bagaev, D.V., Zvyagin, I.V., Putintseva, E.V., Izraelson, M., Britanova, O.V., Chudakov, D.M., and Shugay, M. (2016). VDJviz: a versatile browser for immunogenomics data. BMC Genom. 17, 453.

Bashford-Rogers, R.J., Palser, A.L., Hodkinson, C., Baxter, J., Follows, G.A., Vassiliou, G.S., and Kellam, P. (2017). Dynamic variation of CD5 surface expression levels within individual chronic lymphocytic leukemia clones. Exp. Hematol. 46, 31—37.e10.

Bashford-Rogers, R.J.M., Bergamaschi, L., McKinney, E.F., Pombal, D.C., Mescia, F., Lee, J.C., Thomas, D.C., Flint, S.M., Kellam, P., Jayne, D.R.W., et al. (2019). Analysis of the B cell receptor repertoire in six immune-mediated diseases. Nature 574, 122–126.

Bieberich, F., Vazquez-Lombardi, R., Yermanos, A., Ehling, R.A., Mason, D.M., Wagner, B., Kapetanovic, E., Di Roberto, R.B., Weber, C.R., Savic, M., et al. (2021). A single-cell atlas of lymphocyte adaptive immune repertoires and transcriptomes reveals age-related differences in convalescent COVID-19 patients. Front. Immunol. 12, 701085.

Bischof, J., and Ibrahim, S.M. (2016). bcRep: R package for comprehensive analysis of B cell receptor repertoire data. PLoS One 11, e0161569.

Bolotin, D.A., Poslavsky, S., Mitrophanov, I., Shugay, M., Mamedov, I.Z., Putintseva, E.V., and Chudakov, D.M. (2015). MiXCR: software for comprehensive adaptive immunity profiling. Nat. Methods 12, 380–381.

Braddom, A.E., Bol, S., Gonzales, S.J., Reyes, R.A., Musinguzi, K., Nankya, F., Ssewanyana, I., Greenhouse, B., and Bunnik, E.M. (2021). B cell receptor repertoire analysis in malaria-naive and malaria-experienced individuals reveals unique characteristics of atypical memory B cells. mSphere 6, e0072621.

Bradley, A.S., Ford, B., and Bansal, A.S. (2013). Altered functional B cell subset populations in patients with chronic fatigue syndrome compared to healthy controls. Clin. Exp. Immunol. 172, 73–80.

Cai, H., Yang, L., Shen, K., Zhang, W., Xiong, J., Zhang, M., Mao, X., Wang, Y., and Xiao, M. (2018). A rare e14a3 BCR/ABL fusion transcript in acute lymphoblastic leukemia patient treated with CAR-modified T-cell therapy. Oncol. Lett. 15, 2491–2494.

Canzar, S., Neu, K.E., Tang, Q., Wilson, P.C., and Khan, A.A. (2017). BASIC: BCR assembly from single cells. Bioinformatics 33, 425–427.

Carlson, C.S., Emerson, R.O., Sherwood, A.M., Desmarais, C., Chung, M.W., Parsons, J.M., Steen, M.S., LaMadrid-Herrmannsfeldt, M.A., Williamson, D.W., Livingston, R.J., et al. (2013). Using synthetic templates to design an unbiased multiplex PCR assay. Nat. Commun. 4, 2680.

Christley, S., Levin, M.K., Toby, I.T., Fonner, J.M., Monson, N.L., Rounds, W.H., Rubelt, F., Scarborough, W., Scheuermann, R.H., and Cowell, L.G. (2017). VDJPipe: a pipelined tool for

pre-processing immune repertoire sequencing data. BMC Bioinf. *18*, 448.

D'Angelo, S., Glanville, J., Ferrara, F., Naranjo, L., Gleasner, C.D., Shen, X., Bradbury, A.R., and Kiss, C. (2014). The antibody mining toolbox: an open source tool for the rapid analysis of antibody repertoires. mAbs 6, 160–172.

Davydov, A.N., Obraztsova, A.S., Lebedin, M.Y., Turchaninova, M.A., Staroverov, D.B., Merzlyak, E.M., Sharonov, G.V., Kladova, O., Shugay, M., Britanova, O.V., and Chudakov, D.M. (2018). Comparative analysis of B-cell receptor repertoires induced by live yellow fever vaccine in young and middle-age donors. Front. Immunol. *9*, 2309.

de Bourcy, C.F.A., Dekker, C.L., Davis, M.M., Nicolls, M.R., and Quake, S.R. (2017). Dynamics of the human antibody repertoire after B cell depletion in systemic sclerosis. Sci. Immunol. *2*, eaan8289.

Deininger, M.W., Hodgson, J.G., Shah, N.P., Cortes, J.E., Kim, D.W., Nicolini, F.E., Talpaz, M., Baccarani, M., Muller, M.C., Li, J., et al. (2016). Compound mutations in BCR-ABL1 are not major drivers of primary or secondary resistance to ponatinib in CP-CML patients. Blood *127*, 703–712.

Duez, M., Giraud, M., Herbert, R., Rocher, T., Salson, M., and Thonier, F. (2016). Vidjil: a web platform for analysis of high-throughput repertoire sequencing. PLoS One *11*, e0166126.

Ehlers, A.M., den Hartog Jager, C.F., Knulst, A.C., and Otten, H.G. (2021). Distinction between peanut allergy and tolerance by characterization of B cell receptor repertoires. Allergy 76, 2753–2764.

Elhanati, Y., Sethna, Z., Marcou, Q., Callan, C., Mora, T., and Walczak, A. (2015). Inferring processes underlying B-cell repertoire diversity. Philos. Trans. R. Soc. Lond. B Biol. Sci. *370*, 20140243.

Fahnrich, A., Krebbel, M., Decker, N., Leucker, M., Lange, F.D., Kalies, K., and Moller, S. (2017). ClonoCalc and ClonoPlot: immune repertoire analysis from raw files to publication figures with graphical user interface. BMC Bioinf. *18*, 164.

Feng, J., Fan, S., Sun, Y., Zhang, Z., Ren, H., Li, W., Cui, L., Peng, B., Ren, X., Zhang, W., et al. (2020). Study of B cell repertoire in patients with anti-N-Methyl-D-Aspartate Receptor encephalitis. Front. Immunol. *11*, 1539.

Gadala-Maria, D., Yaari, G., Uduman, M., and Kleinstein, S.H. (2015). Automated analysis of high-throughput B-cell sequencing data reveals a high frequency of novel immunoglobulin V gene segment alleles. Proc. Natl. Acad. Sci. USA *112*, E862–E870.

Galson, J.D., Schaetzle, S., Bashford-Rogers, R.J.M., Raybould, M.I.J., Kovaltsuk, A., Kilpatrick, G.J., Minter, R., Finch, D.K., Dias, J., James, L.K., et al. (2020). Deep sequencing of B cell receptor repertoires from COVID-19 patients reveals strong convergent immune signatures. Front. Immunol. *11*, 605170.

González, L.A., Ugarte-Gil, M.F., and Alarcón, G.S. (2021). Systemic lupus erythematosus: the

search for the ideal biomarker. Lupus *30*, 181–203.

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.

H, I.J., van Schouwenburg, P.A., van Zessen, D., Pico-Knijnenburg, I., Driessen, G.J., Stubbs, A.P., and van der Burg, M. (2016). Evaluation of the antigen-experienced B-cell receptor repertoire in healthy children and adults. Front. Immunol. *7*, 410.

H, I.J., van Schouwenburg, P.A., van Zessen, D., Pico-Knijnenburg, I., Stubbs, A.P., and van der Burg, M. (2017). Antigen receptor galaxy: a user-friendly, web-based tool for analysis and visualization of T and B cell receptor repertoire data. J. Immunol. *198*, 4156–4165.

Haque, A., Engel, J., Teichmann, S.A., and Lonnberg, T. (2017). A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. Genome Med. *9*, 75.

Hoehn, K., Fowler, A., Lunter, G., and Pybus, O.G. (2016). The diversity and molecular evolution of B-cell receptors during infection. Mol. Bio. Evol. *33*, 1147–1157.

Hutter, J.J. (2010). Childhood leukemia. Pediatr. Rev. *31*, 234–241.

Islam, S., Kjallquist, U., Moliner, A., Zajac, P., Fan, J.B., Lonnberg, P., and Linnarsson, S. (2011). Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. Genome Res. *21*, 1160–1167.

James, K.R., Gomes, T., Elmentaite, R., Kumar, N., Gulliver, E.L., King, H.W., Stares, M.D., Bareham, B.R., Ferdinand, J.R., Petrova, V.N., et al. (2020). Distinct microbial and immune niches of the human colon. Nat. Immunol. *21*, 343–353.

Jin, X., Zhou, W., Luo, M., Wang, P., Xu, Z., Ma, K., Cao, H., Xu, C., Huang, Y., Cheng, R., et al. (2021). Global characterization of B cell receptor repertoire in COVID-19 patients by single-cell V(D)J sequencing. Brief Bioinform. *22*.

Kiyotani, K., Mai, T.H., Yamaguchi, R., Yew, P.Y., Kulis, M., Orgel, K., Imoto, S., Miyano, S., Burks, A.W., and Nakamura, Y. (2018). Characterization of the B-cell receptor repertoires in peanut allergic subjects undergoing oral immunotherapy. J. Hum. Genet. *63*, 239–248.

Konishi, H., Komura, D., Katoh, H., Atsumi, S., Koda, H., Yamamoto, A., Seto, Y., Fukayama, M., Yamaguchi, R., Imoto, S., and Ishikawa, S. (2019). Capturing the differences between humoral immunity in the normal and tumor environments from repertoire-seq of B-cell receptors using supervised machine learning. BMC Bioinf. *20*, 267.

Lai, L., Zhou, X., Chen, H., Luo, Y., Sui, W., Zhang, J., Tang, D., Yan, Q., and Dai, Y. (2019). Composition and diversity analysis of the B-cell receptor immunoglobulin heavy chain complementarity-determining region 3 repertoire in patients with acute rejection after kidney transplantation using high-throughput sequencing. Exp. Ther. Med. *17*, 2206–2220.

Lee, J., Boutz, D.R., Chromikova, V., Joyce, M.G., Vollmers, C., Leung, K., Horton, A.P., DeKosky, B.J., Lee, C.H., Lavinder, J.J., et al. (2016a). Molecular-level analysis of the serum antibody repertoire in young adults before and after seasonal influenza vaccination. Nat. Med. *22*, 1456–1464.

Lee, Y.N., Frugoni, F., Dobbs, K., Tirosh, I., Du, L., Ververs, F.A., Ru, H., Ott de Bruin, L., Adeli, M., Bleesing, J.H., et al. (2016b). Characterization of T and B cell repertoire diversity in patients with RAG deficiency. Sci. Immunol. *1*, eaah6109.

Li, A., Rue, M., Zhou, J., Wang, H., Goldwasser, M., Neuberg, D., Dalton, V., Zuckerman, D., Lyons, C., Silverman, L., et al. (2004). Utilization of Ig heavy chain variable, diversity, and joining gene segments in children with B-lineage acute lymphoblastic leukemia: implications for the mechanisms of VDJ recombination and for pathogenesis. Blood *103*, 4602–4609.

Liu, H., Pan, W., Tang, C., Tang, Y., Wu, H., Yoshimura, A., Deng, Y., He, N., and Li, S. (2021). The methods and advances of adaptive immune receptors repertoire sequencing. Theranostics *11*, 8945–8963.

Liu, S., Hou, X.L., Sui, W.G., Lu, Q.J., Hu, Y.L., and Dai, Y. (2017). Direct measurement of B-cell receptor repertoire's composition and variation in systemic lupus erythematosus. Genes Immun. *18*, 22–27.

Liu, X., and Wu, J. (2018). History, applications, and challenges of immune repertoire research. Cell Biol. Toxicol. *34*, 441–457.

Magi, A., Semeraro, R., Mingrino, A., Giusti, B., and D'Aurizio, R. (2018). Nanopore sequencing data analysis: state of the art, applications and challenges. Brief Bioinform. *19*, 1256–1272.

Maillette de Buy Wenniger, L.J., Doorenspleet, M.E., Klarenbeek, P.L., Verheij, J., Baas, F., Elferink, R.P., Tak, P.P., de Vries, N., and Beuers, U. (2013). Immunoglobulin G4+ clones identified by next-generation sequencing dominate the B cell receptor repertoire in immunoglobulin G4 associated cholangitis. Hepatology *57*, 2390–2398.

Mamanova, L., Miao, Z., Jinat, A., Ellis, P., Shirley, L., and Teichmann, S.A. (2021). High-throughput full-length single-cell RNA-seq automation. Nat. Protoc. *16*, 2886–2915.

Marcou, Q., Mora, T., and Walczak, A.M. (2018). High-throughput immune repertoire analysis with IGoR. Nat. Commun. *9*, 561.

Melchers, F.J.n.r.I. (2005). The pre-B-cell receptor: selector of fitting immunoglobulin heavy chains for the B-cell repertoire. Nat. Rev. Immunol. *5*, 578–584.

Messmer, B.T., Albesiano, E., Efremov, D.G., Ghiotto, F., Allen, S.L., Kolitz, J., Foa, R., Damle, R.N., Fais, F., Messmer, D., et al. (2004). Multiple distinct sets of stereotyped antigen receptors indicate a role for antigen in promoting chronic lymphocytic leukemia. J. Exp. Med. *200*, 519–525.

Metzker, M.L. (2010). Sequencing technologies - the next generation. Nat. Rev. Genet. *11*, 31–46.

Minervina, A., Pogorelyy, M., and Mamedov, I. (2019). T-cell receptor and B-cell receptor

repertoire profiling in adaptive immunity. Transpl. Int. *32*, 1111–1123.

Miyasaka, A., Yoshida, Y., Wang, T., and Takikawa, Y. (2019). Next-generation sequencing analysis of the human T-cell and B-cell receptor repertoire diversity before and after hepatitis B vaccination. Hum. Vaccin. Immunother. *15*, 2738–2753.

Montague, Z., Lv, H., Otwinowski, J., DeWitt, W.S., Isacchini, G., Yip, G.K., Ng, W.W., Tsang, O.T., Yuan, M., Liu, H., et al. (2021). Dynamics of B cell repertoires and emergence of cross-reactive responses in patients with different severities of COVID-19. Cell Rep. *35*, 109173.

Muggen, A.F., de Jong, M., Wolvers-Tettero, I.L.M., Kallemeijn, M.J., Teodósio, C., Darzentas, N., Stadhouders, R., Ijspeert, H., van der Burg, M., van IJcken, W.F.J., et al. (2019). The presence of CLL-associated stereotypic B cell receptors in the normal BCR repertoire from healthy individuals increases with age. Immun. Ageing *16*, 22.

Nakahara, Y., Matsutani, T., Igarashi, Y., Matsuo, N., Himuro, H., Saito, H., Yamada, K., Murotani, K., Hoshino, T., Azuma, K., and Sasada, T. (2021). Clinical significance of peripheral TCR and BCR repertoire diversity in EGFR/ALK wild-type NSCLC treated with anti-PD-1 antibody. Cancer Immunol. Immunother. *70*, 2881–2892.

Nielsen, S.C.A., and Boyd, S.D. (2018). Human adaptive immune receptor repertoire analysis-Past, present, and future. Immunol. Rev. *284*, 9–23.

Nielsen, S.C.A., Roskin, K.M., Jackson, K.J.L., Joshi, S.A., Nejad, P., Lee, J.Y., Wagar, L.E., Pham, T.D., Hoh, R.A., Nguyen, K.D., et al. (2019). Shaping of infant B cell receptor repertoires by environmental factors and infectious disease. Sci. Transl. Med. *11*.

Niu, X., Li, S., Li, P., Pan, W., Wang, Q., Feng, Y., Mo, X., Yan, Q., Ye, X., Luo, J., et al. (2020). Longitudinal analysis of T and B cell receptor repertoire transcripts reveal dynamic immune response in COVID-19 patients. Front. Immunol. *11*, 582010.

O'Connell, A.E., Volpi, S., Dobbs, K., Fiorini, C., Tsitsikov, E., de Boer, H., Barlan, I.B., Despotovic, J.M., Espinosa-Rosales, F.J., Hanson, I.C., et al. (2014). Next generation sequencing reveals skewing of the T and B cell receptor repertoires in patients with wiskott-Aldrich syndrome. Front. Immunol. *5*, 340.

Olson, B.J., Moghimi, P., Schramm, C.A., Obraztsova, A., Ralph, D., Vander Heiden, J.A., Shugay, M., Shepherd, A.J., Lees, W., and Matsen, F.A.t. (2019). Sumrep: a summary statistic framework for immune receptor repertoire comparison and model Validation. Front. Immunol. *10*, 2533.

Paschold, L., Simnica, D., Willscher, E., Vehreschild, M.J., Dutzmann, J., Sedding, D.G., Schultheiß, C., and Binder, M. (2021). SARS-CoV-2-specific antibody rearrangements in prepandemic immune repertoires of risk cohorts and patients with COVID-19. J. Clin. Invest. *131*.

Petrova, V.N., Muir, L., McKay, P.F., Vassiliou, G.S., Smith, K.G.C., Lyons, P.A., Russell, C.A., Anderson, C.A., Kellam, P., and Bashford-Rogers, R.J.M. (2018). Combined influence of B-cell

receptor rearrangement and somatic hypermutation on B-cell class-switch fate in health and in chronic lymphocytic leukemia. Front. Immunol. *9*, 1784.

Pham, T.D., Chng, M.H.Y., Roskin, K.M., Jackson, K.J.L., Nguyen, K.D., Glanville, J., Lee, J.Y., Engleman, E.G., and Boyd, S.D. (2017). High-fat diet induces systemic B-cell repertoire changes associated with insulin resistance. Mucosal Immunol. *10*, 1468–1479.

Pickman, Y., Dunn-Walters, D., and Mehr, R. (2013). BCR CDR3 length distributions differ between blood and spleen and between old and young patients, and TCR distributions can be used to detect myelodysplastic syndrome. Phys. Biol. *10*, 056001.

Pineda, S., Sigdel, T.K., Liberto, J.M., Vincenti, F., Sirota, M., and Sarwal, M.M. (2019). Characterizing pre-transplant and post-transplant kidney rejection risk by B cell immune repertoire sequencing. Nat. Commun. *10*, 1906.

Rang, F.J., Kloosterman, W.P., and de Ridder, J. (2018). From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy. Genome Biol. *19*, 90.

Rawstron, A.C., Fazi, C., Agathangelidis, A., Villamor, N., Letestu, R., Nomdedeu, J., Palacio, C., Stehlikova, O., Kreuzer, K.A., Liptrot, S., et al. (2016). A complementary role of multiparameter flow cytometry and high-throughput sequencing for minimal residual disease detection in chronic lymphocytic leukemia: an European Research Initiative on CLL study. Leukemia *30*, 929–936.

Rizzetto, S., Koppstein, D.N.P., Samir, J., Singh, M., Reed, J.H., Cai, C.H., Lloyd, A.R., Eltahla, A.A., Goodnow, C.C., and Luciani, F. (2018). B-cell receptor reconstruction from single-cell RNA-seq with VDJPuzzle. Bioinformatics *34*, 2846–2847.

Rodriguez-Vicente, A.E., Bikos, V., Hernandez-Sanchez, M., Malcikova, J., Hernandez-Rivas, J.M., and Pospisilova, S. (2017). Next-generation sequencing in chronic lymphocytic leukemia: recent findings and new horizons. Oncotarget *8*, 71234–71248.

Rose, N.R. (2016). Prediction and prevention of autoimmune disease in the 21st century: a review and preview. Am. J. Epidemiol. *183*, 403–406.

Rosenwald, A., Alizadeh, A.A., Widhopf, G., Simon, R., Davis, R.E., Yu, X., Yang, L., Pickeral, O.K., Rassenti, L.Z., Powell, J., et al. (2001). Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. J. Exp. Med. *194*, 1639–1647.

Sato, W., Ono, H., Matsutani, T., Nakamura, M., Shin, I., Amano, K., Suzuki, R., and Yamamura, T. (2021). Skewing of the B cell receptor repertoire in myalgic encephalomyelitis/chronic fatigue syndrome. Brain Behav. Immun. *95*, 245–255.

Schneikart, G., Tavarini, S., Sammicheli, C., Torricelli, G., Guidotti, S., Andreano, E., Buricchi, F., D'Oro, U., Finco, O., and Bardelli, M. (2020). The respiratory syncytial virus fusion protein-specific B cell receptor repertoire reshaped by post-fusion subunit vaccination. Vaccine *38*, 7916–7927.

Schultheiß, C., Paschold, L., Simnica, D., Mohme, M., Willscher, E., von Wenserski, L., Scholz, R., Wieters, I., Dahlke, C., Tolosa, E., et al. (2020). Next-generation sequencing of T and B cell receptor repertoires from COVID-19 patients showed signatures associated with severity of disease. Immunity *53*, 442–455.e4.

Setliff, I., Shiakolas, A.R., Pilewski, K.A., Murji, A.A., Mapengo, R.E., Janowska, K., Richardson, S., Oosthuysen, C., Raju, N., Ronsard, L., et al. (2019). High-throughput mapping of B cell receptor sequences to antigen specificity. Cell *179*, 1636–1646.e5.

Shi, X., Shao, T., Huo, F., Zheng, C., Li, W., and Jiang, Z. (2020). An analysis of abnormalities in the B cell receptor repertoire in patients with systemic sclerosis using high-throughput sequencing. PeerJ *8*, e8370.

Shlomchik, M.J., Craft, J.E., and Mamula, M.J. (2001). From T to B and back again: positive feedback in systemic autoimmune disease. Nat. Rev. Immunol. *1*, 147–153.

Six, A., Mariotti-Ferrandiz, M.E., Chaara, W., Magadan, S., Pham, H.P., Lefranc, M.P., Mora, T., Thomas-Vaslin, V., Walczak, A.M., and Boudinot, P. (2013). The past, present, and future of immune repertoire biology - the rise of next-generation repertoire analysis. Front. Immunol. *4*, 413.

Strauli, N.B., and Hernandez, R.D. (2016). Statistical inference of a convergent antibody repertoire response to influenza vaccine. Genome Med. *8*, 60.

Stubbington, M.J.T., Lönnberg, T., Proserpio, V., Clare, S., Speak, A.O., Dougan, G., and Teichmann, S.A. (2016). T cell fate and clonality inference from single-cell transcriptomes. Nat. Methods *13*, 329–332.

Su, Z., Jin, Y., Zhang, Y., Guan, Z., Li, H., Chen, X., Xie, C., Zhang, C., Liu, X., Li, P., et al. (2021). The diagnostic and prognostic potential of the B-cell repertoire in membranous nephropathy. Front. Immunol. *12*, 635326.

Sun, G., Qiu, L., Cheng, Z., Pan, W., Qiu, J., Zou, C., Xie, N., Liu, S., Zhu, P., Zeng, J., and Dai, Y. (2018). Association of the characteristics of B- and T-cell repertoires with papillary thyroid carcinoma. Oncol. Lett. *16*, 1584–1592.

Sun, H., Yang, H.Q., Zhai, K., and Tong, Z.H. (2021). Signatures of B cell receptor repertoire following pneumocystis infection. Front. Microbiol. *12*, 636250.

Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., and Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN Estimates of incidence and mortality Worldwide for 36 cancers in 185 Countries. CA Cancer J. Clin. *71*, 209–249.

Tabibian-Keissar, H., Hazanov, L., Schiby, G., Rosenthal, N., Rakovsky, A., Michaeli, M., Shahaf, G.L., Pickman, Y., Rosenblatt, K., Melamed, D., et al. (2016). Aging affects B-cell antigen receptor repertoire diversity in primary and secondary lymphoid tissues. Eur. J. Immunol. *46*, 480–492.

Tan, K.-T., Ding, L.-W., Sun, Q.-Y., Lao, Z.-T., Chien, W., Ren, X., Xiao, J.-F., Loh, X.Y., Xu, L., Lill, M., et al. (2018). Profiling the B/T cell receptor

repertoire of lymphocyte derived cell lines. BMC Cancer 18, 940.

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A., et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. Nat. Methods 6, 377–382.

Tang, W., Dou, Y., Qin, T., Ding, Y., Tang, X., Zhao, X., and An, Y. (2019). Skewed B cell receptor repertoire and reduced antibody avidity in patients with DOCK8 deficiency. Scand. J. Immunol. 89, e12759.

Teng, G., and Papavasiliou, F.N. (2007). Immunoglobulin somatic hypermutation. Annu. Rev. Genet. 41, 107–120.

Thapa, M., Tedesco, D., Gumber, S., Elrod, E.J., Han, J.H., Kitchens, W.H., Magliocca, J.F., Adams, A.B., and Grakoui, A. (2020). Blockade of BAFF reshapes the hepatic B cell receptor repertoire and attenuates autoantibody production in cholestatic liver disease. J. Immunol. 204, 3117–3128.

Tiller, T., Busse, C.E., and Wardemann, H. (2009). Cloning and expression of murine Ig genes from single B cells. J. Immunol. Methods 350, 183–193.

Tucci, F.A., Kitanovski, S., Johansson, P., Klein-Hitpass, L., Kahraman, A., Dürig, J., Hoffmann, D., and Küppers, R. (2018). Biased IGH VDJ gene repertoire and clonal expansions in B cells of chronically hepatitis C virus-infected individuals. Blood 131, 546–557.

van Schouwenburg, P.A., IJspeert, H., Pico-Knijnenburg, I., Dalm, V.A.S.H., van Hagen, P.M., van Zessen, D., Stubbs, A.P., Patel, S.Y., and van der Burg, M. (2018). Identification of CVID patients with defects in immune repertoire formation or specification. Front. Immunol. 9, 2545.

Vander Heiden, J.A., Stathopoulos, P., Zhou, J.Q., Chen, L., Gilbert, T.J., Bolen, C.R., Barohn, R.J., Dimachkie, M.M., Ciafaloni, E., Broering, T.J., et al. (2017). Dysregulation of B cell repertoire formation in myasthenia gravis patients revealed through deep sequencing. J. Immunol. 198, 1460–1473.

Vardi, A., Vlachonikola, E., Karypidou, M., Stalika, E., Bikos, V., Gemenetzi, K., Maramis, C., Siorenta, A., Anagnostopoulos, A., Pospisilova, S., et al. (2017). Restrictions in the T-cell repertoire of chronic lymphocytic leukemia: high-throughput immunoprofiling supports selection by shared antigenic elements. Leukemia 31, 1555–1561.

Victora, G., and Nussenzweig, M.C. (2012). Germinal centers. Annu. Rev. Immunol. 30, 429–457.

Vieira Braga, F.A., Kar, G., Berg, M., Carpaij, O.A., Polanski, K., Simon, L.M., Brouwer, S., Gomes, T., Hesse, L., Jiang, J., et al. (2019). A cellular census of human lungs identifies novel cell states in health and in asthma. Nat. Med. 25, 1153–1163.

Vollmers, C., De Vlaminck, I., Valantine, H.A., Penland, L., Luikart, H., Strehl, C., Cohen, G., Khush, K.K., and Quake, S.R. (2015). Monitoring pharmacologically induced immunosuppression by immune repertoire sequencing to detect acute allograft rejection in heart transplant patients: a proof-of-concept diagnostic accuracy study. PLoS Med. 12, e1001890.

Wang, L., Dai, Y., Liu, S., Lai, L., Yan, Q., Chen, H., Zhang, J., and Sui, W. (2019). Assessment of variation in B-cell receptor heavy chain repertoire in patients with end-stage renal disease by high-throughput sequencing. Ren. Fail. 41, 1–13.

Weng, R., Liu, S., Gu, X., and Zhong, Z. (2021). Clonal diversity of the B cell receptor repertoire in patients with coronary in-stent restenosis and type 2 diabetes. Open Life Sci. 16, 884–898.

Weng, R., Liu, S., Gu, X., and Zhong, Z. (2022). Characterization of the B cell receptor repertoire of patients with acute coronary syndrome. Genes Genomics 44, 19–28.

Werner, L., Lee, Y.N., Rechavi, E., Lev, A., Yerushalmi, B., Ling, G., Shah, N., Uhlig, H.H., Weiss, B., Somech, R., et al. (2020). Alterations in T and B cell receptor repertoires patterns in patients with IL10 signaling defects and history of infantile-onset IBD. Front. Immunol. 11, 109.

Xiang, H., Zhao, Y., Li, X., Liu, P., Wang, L., Wang, M., Tian, L., Sun, H.X., Zhang, W., Xu, Z., et al. (2022). Landscapes and dynamic diversifications of B-cell receptor repertoires in COVID-19 patients. Hum. Immunol. 83, 119–129.

Yaari, G., and Kleinstein, S.H. (2015). Practical guidelines for B-cell receptor repertoire sequencing analysis. Genome Med. 7, 121.

Yan, Q., Wang, L., Lai, L., Liu, S., Chen, H., Zhang, J., Dai, Y., and Sui, W. (2019). Next generation sequencing reveals novel alterations in B-cell heavy chain receptor repertoires associated with acute-on-chronic liver failure. Int. J. Mol. Med. 43, 243–255.

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.

Yermanos, A., Agrafiotis, A., Kuhn, R., Robbiani, D., Yates, J., Papadopoulou, C., Han, J., Sandu, I., Weber, C., Bieberich, F., et al. (2021). Platypus: an open-access software for integrating lymphocyte single-cell immune repertoires with transcriptomes. NAR Genom. Bioinform. 3, lqab023.

Zhang, C., Huang, H., Miao, Y., Xiong, H., and Lu, Z. (2018). Clonal distribution and intratumour heterogeneity of the B-cell repertoire in oesophageal squamous cell carcinoma. J. Pathol. 246, 323–330.

Zhang, J., Hu, X., Wang, J., Sahu, A.D., Cohen, D., Song, L., Ouyang, Z., Fan, J., Wang, B., Fu, J., et al. (2019). Immune receptor repertoires in

pediatric and adult acute myeloid leukemia. Genome Med. 11, 73.

Zhang, J.Y., Wang, X.M., Xing, X., Xu, Z., Zhang, C., Song, J.W., Fan, X., Xia, P., Fu, J.L., Wang, S.Y., et al. (2020a). Single-cell landscape of immunological responses in patients with COVID-19. Nat. Immunol. 21, 1107–1118.

Zhang, L., Wang, I.M., Solban, N., Cristescu, R., Zeng, G., and Long, B. (2020b). Comprehensive investigation of T and B cell receptor repertoires in an MC38 tumor model following murine anti-PD-1 administration. Mol. Med. Rep. 22, 975–985.

Zhang, Q., Jing, H., and Su, H.C. (2016). Recent advances in DOCK8 immunodeficiency syndrome. J. Clin. Immunol. 36, 441–449.

Zhang, S., Zhang, S., Lin, Z., Zhang, X., Dou, X., Zhou, X., Wang, X., Wang, Z., and Zhang, Q. (2021). Deep sequencing reveals the skewed B-cell receptor repertoire in plaques and the association between pathogens and atherosclerosis. Cell. Immunol. 360, 104256.

Zhang, W., Du, Y., Su, Z., Wang, C., Zeng, X., Zhang, R., Hong, X., Nie, C., Wu, J., Cao, H., et al. (2015). IMonitor: a robust pipeline for TCR and BCR repertoire analysis. Genetics 201, 459–472.

Zhang, W., Feng, Q., Wang, C., Zeng, X., Du, Y., Lin, L., Wu, J., Fu, L., Yang, K., Xu, X., et al. (2017). Characterization of the B cell receptor repertoire in the intestinal mucosa and of tumor-infiltrating lymphocytes in colorectal adenoma and carcinoma. J. Immunol. 198, 3719–3728.

Zhang, Y., Yan, Q., Luo, K., He, P., Hou, R., Zhao, X., Wang, Q., Yi, H., Liang, H., Deng, Y., et al. (2022). Analysis of B cell receptor repertoires reveals key signatures of the systemic B cell response after SARS-CoV-2 infection. J. Virol. 96, e0160021.

Zhao, M., Li, X., Xie, S., Gong, M., Yan, R., Zheng, A., Xu, Y., Wu, H., and Wang, Z. (2021a). The dynamics and association of B and T cell receptor repertoires upon antibody response to hepatitis B vaccination in healthy adults. Hum Vaccin Immunother 17, 3203–3213.

Zhao, P., Guo, S., Zhong, Z., Yang, S., and Xia, X. (2021b). Quantitative characterization of the B cell receptor repertoires of human immunized with commercial rabies virus vaccine. Hum Vaccin Immunother 17, 2538–2546.

Zheng, F., Xu, H., Zhang, C., Hong, X., Liu, D., Tang, D., Xiong, Z., and Dai, Y. (2021). Immune cell and TCR/BCR repertoire profiling in systemic lupus erythematosus patients by single-cell sequencing. Aging 13, 24432–24448.

Zhou, Y., Zhang, J., Wang, D., Wang, D., Guan, W., Qin, J., Xu, X., Fang, J., Fu, B., Zheng, X., et al. (2021). Profiling of the immune repertoire in COVID-19 patients with mild, severe, convalescent, or retesting-positive status. J. Autoimmun. 118, 102596.