British Journal of Medicine & Medical Research
18(12): 1-11, 2016; Article no.BJMMR.27765
ISSN: 2231-0614, NLM ID: 101570965

SCIENCEDOMAIN *international*
www.sciencedomain.org

High-throughput Sequencing and Profiling of Antibody Repertoire and the Promises for Diagnostic and Therapeutic Applications

Anteneh Getachew[1*] and Malaiyarasa Pandian[1]

[1]Department of Biotechnology, College of Natural and Computational Sciences, Wolkite University, P.O.Box -07, Wolkite, Ethiopia.

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/BJMMR/2016/27765
Editor(s):
(1) Faris Q. B. Alenzi, Department of Medical Laboratories, College of Applied Medical Sciences Salman bin Abdulaziz University (Al-Kharj), Saudi Arabia.
Reviewers:
(1) Wagner Loyola, Concordia University, Brazil.
(2) Anonymous, Instituto de Investigaciones Biomédicas de la UNAM, Spain.
(3) Arturo Solís-Herrera, Center for Studies of Human Photosynthesis, México.
Complete Peer review History: http://www.sciencedomain.org/review-history/17049

Review Article

Received 18th June 2016
Accepted 20th August 2016
Published 29th November 2016

ABSTRACT

High-throughput DNA sequencing technologies are advancing at rapid state and transforming our indulgent about humoral immune modulation. The antibody repertoire encoded by B cells in the blood or lymphoid organs and the information obtained from high-throughput DNA sequencing of immunoglobulin genes (Ig-seq) can be applied for various disciplines. These are basically used to detect B-cell malignancies with high sensitivity and discovering specific antibodies for specific antigens. In advance, it plays significant roles in vaccine development process and it increases our understanding about autoimmunity. The broader applications of Ig-seq in clinical sets highly necessitate the development of a consistent experimental design framework that will enable the sharing and meta-analysis of sequencing data generated by many scientists. The present review attempts to give backgrounds about high-throughput antibody repertoires gene sequencing as well as the ways in which Ig-seq might be applied to characterize immune responses and identify antibodies of therapeutic, diagnostic or mechanistic relevance to autoimmune diseases.

**Corresponding author: E-mail: antutensu@gmail.com;*

Keywords: High-throughput; Ig-seq; sequencing; antibody repertoire; diagnostic; therapeutic.

1. INTRODUCTION

Antibodies are a major component of the adaptive immune system and have critical roles in protective and pathogenic immune responses. In response to microbial infection, vaccination, autoimmune disease or cancer, the immune system generates distinct antibody repertoires. Analysis of these antibody repertoires, particularly those contributing to functional immune responses, can provide important information on protective and pathogenic immunity. In autoimmune diseases, including autoimmune rheumatic diseases, antibody characterization has enabled the identification of autoantigens and has provided insights into the underlying mechanisms of disease; furthermore, detection of autoantibodies has become a cornerstone of modern diagnostics [1,2,3].

There are extensive trails to determine the antibody repertoire encoded by B cells in the blood or lymphoid organs using high-throughput DNA sequencing technologies. Efforts have been advancing tremendously in rapid state and are transforming our perceptions of humoral immune responses. Evidences gained from high-throughput DNA sequencing of immunoglobulin genes (Ig-seq) can be useful to detect B-cell malignancies with high sensitivity, to discover antibodies, autoimmunity and to guide vaccine development [4,5,6]. Speedy growth in the development of experimental protocols and informatics analysis tools is plateful to reduce sequencing artifacts, to accomplish more accurate quantification of clonal diversity and to extract the most relevant biological information. A powerful adaptive immune system is primarily dependent upon the generation of a varied repertoire of B-lymphocyte antigen receptors (BCRs) which entail the membrane-bound form of antibodies expressed on the surface of B cells. The BCRs are principally assembled by somatic recombination of a large number of immunoglobulin gene fragments and the repertoire of BCRs articulated and continuously shaped by exposure to external antigens and internal host factors. The mechanisms to be had for BCR diversification can give up an enormous number of possible BCRs (in theory, >1013 in humans) [7,8]. There is a possibility that, this number can exceed from the total number of B lymphocytes in the human body (~1–2 × 1011). As such, considering labor and cost, it is totally impossible to analyze such a diverse BCR repertoire using the conventional type of Sanger sequencing technique. Even though, the Ig-seq has permitted us to find out antibody gene repertoires at an unmatched depth [7,8]. The information has been gained by Ig-seq is demonstrating the ways helpful for understanding antibody responses in health and disease conditions as well as for diagnostic purposes. Furthermore, the Ig-seq can be combined and practical with other relevant techniques, such as expression of antigen-specific antibodies, sequencing of multiple RNAs [8], proteomic analyses of antibodies in blood or secretions [5], the properties of antibodies that mediate protection against infectious diseases [6] and alternatively that mediate autoimmune responses [6]. This review illustrate high-throughput antibody gene sequencing as well as the ways in which Ig-seq might be applied to characterize immune responses and identify antibodies of therapeutic, diagnostic or mechanistic relevance to autoimmune diseases.

2. NATURE OF THE GENERATION OF ANTIBODY REPERTOIRE

Antibodies are circulating proteins that are produced in vertebrates in response to exposure to foreign structures known as antigens. Antibodies are extremely diverse and specific in their skill to recognize foreign molecular structures and are the primary mediators of humoral immunity against all classes of microbes [9,10]. The family of circulating proteins that mediate these protective responses was initially called antitoxins. When it was appreciated that similar proteins could be generated against many substances, not just microbial toxins, these proteins were given the general name antibodies. The substances that generated or were recognized by antibodies were then called antigens [10]. Antibodies are produced by a developmentally ordered series of somatic gene rearrangement events that occur exclusively in developing B cells and continue throughout the life of an organism. Antibodies are having chains that are tremendously expressed in all cases. The chains are expressed as heavy such as, μ, α, γ, δ, ε) and light such as, chains κ, λ, which are connected by disulfide bonds. The integral antibody having variable and constant domains and antigen binding would occurs in the variable domain which is produced by recombination of a predetermined place of tandemly arranged variable (V), diversity (D) and joining (J) germline

gene segments. This process is known as the VDJ recombination process and, often results in the addition and deletion of nucleotides at the junctions between ligated gene segments. Strictly taking, during recombination of gene segments DNA exonucleases can trim the ends of the gene segments, and DNA polymerases and transferases can randomly insert templated palindromic or nontemplated nucleotides, respectively [11]. Initially during B-cell development, immunoglobulin heavy (IgH) chain gene recombination usually takes place before immunoglobulin light (IgL) chain. These both chains typically IgH and IgL genes are prolifically rearranged and entirely assembled antibody heterodimer is expressed on the surface of the B cell. B cells bearing prolifically reorganized antibodies the process of allelic exclusion ensures that, each B cell expresses a single antibody [5]. The aforementioned process and after passage of all the way through developmental checkpoints, recently produced mature $IgM^{+}IgD^{+}$ B cells from the naive B cell repertoire. On the other way, the majority of the diversity in the naive antibody repertoire is concerted on the location of IgH VDJ gene segment ligation and which is known as the IgH complementarity-determining region 3 (CDR-H3)[11].

Because of the combinatorial and non templated nature of the mechanisms that generate the CDR-H3, it is the most diverse component in terms of length and sequence of the antibody H-chain repertoire and is a principal determinant of antibody specificity [12,13,14]. On the other hand, there are occasions where antigen specificity is ordered exclusively or primarily by the L chain. An activated helper T-cell with receptors for the same antigen binds to the B-cell, and, the helper T-cell sends a signal to the B-cell, activating the B-cell. The activated B-cell starts to divide by mitosis to form a clone of plasma cells. This process, known as B-cell proliferation and plasma expansion, occurs primarily in highly structured parts of secondary lymphoid organs such as spleen, lymph nodes and Peyer's patches [15,16]. Next, the clonal expansion is followed by somatic hypermutation of the variable domains of antibodies. This process easily mediated by activation-induced cytidine deaminase activity. B cells expressing BCRs bearing somatic mutations that increase affinity for antigen out competes other B cells for access to antigen. As such, the B cells bearing the highest-affinity antibodies undertake special expansion and survival, a process known as affinity maturation. The mutation results in sequence diversification of the CDR-H1 and

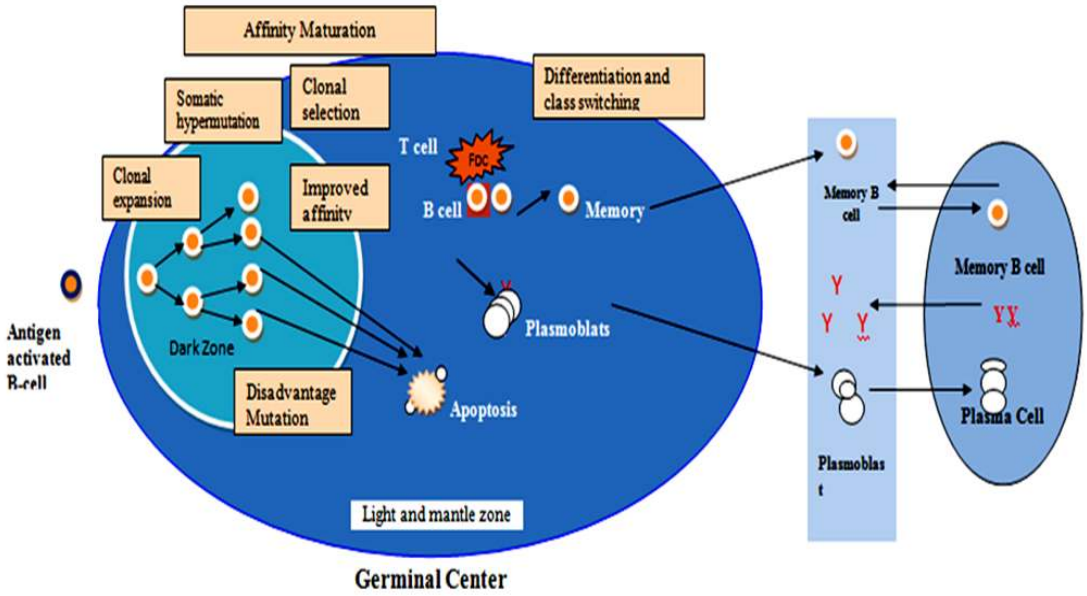

Fig. 1. Steps in the development of antigen-specific B cells. The steps of normal B-call differentiation and diversification of the antibody repertoire are indicated in black text. Normal B cells are generated in the bone marrow, migrate to the periphery and, following developmental checkpoint selection, comprise the population of $IgM^{-}IgD^{-}$ mature naïve B cells

CDR-H2 hypervariable regions and of the framework 3 (FR3) regions [17]. Activation-induced cytidine deaminase also mediates class-switch recombination, which generates antibodies bearing different constant regions. B cells expressing somatically mutated, high-antigen-affinity BCRs can differentiate into long-lived memory B cells, capable of mediating rapid recall responses to the same antigen, or into terminally differentiated plasma cells; the latter down regulate BCR expression, establish residency in the bone marrow, gut lamina propia and secrete protective antibodies at extremely high rates estimated at 10,000–20,000 antibody molecules per second [18]. Antibody production by long-lived plasma cells in the bone marrow is postulated to proceed for very long times, possibly throughout the entire lifetime of the organism. Diversity in the primary antibody repertoire stems from the allelic diversity in immunoglobulin gene segments, combinatorial diversity introduced during somatic recombination. Junctional diversity caused by lack of fit during recombination process such as pairing of IgH and IgL polypeptide chains and receptor editing, in which the offered V-gene segment is substituted with another. Moreover, VH substitution is a progression resulting from the existence of a cryptic recombination signal sequence in FR3. This might influence as much as 5–12% of the human primary B-cell antibody repertoire [19,20,21].

3. HIGH-THROUGHPUT SEQUENCING AND PROFILING OF THE ANTIBODY REPERTOIRE

3.1 Small-scale Sequencing and Profiling Antibody Repertoire

There are significant developments in the area of sequencing of antibody repertoire. Earlier, there were approaches used in the 1990s and 2000s for the analysis of antibody repertoires which involves reverse transcriptase polymerase chain reaction (RT-PCR) and Sanger sequencing of antibodies expressed by individual B cells. The experimental steps in this technique require manual handling of sequences and this approach can only characterize limited sequences of tens to hundreds of B cells per experiment [22]. This sequencing technique which is referred as low throughput has been used to recognize antibodies that neutralize clinically important pathogens such as HIV-1 and influenzae viruses [22,23,24] or autoantibodies that contribute to autoimmune disease [25]. However, other approaches have been used for sequencing of single-cell RNA sequencing including sequencing of coexpressed genes and use next generation sequencing technology [26]. However, these assay platforms and short-read sequencing technologies read significantly fewer base pairs than are essentially (>500) for sequencing the entire IgL and IgH variable regions [27].

3.2 Large-scale Sequencing and Profiling of Antibody Repertoire

There are a number of high-throughput next-generation sequencing technologies have been developed and applied [21,27] and are renovating the investigation of B cells and their antibody repertoires [4,28,29]. The methodological approaches have been applied for each of these technologies is in variety ways and parameters. The methodologies using the B-cell populations characterized, the clinical characteristics of the individuals analysed, the approach to bioinformatic analysis and the objective of the experiment. The following approaches demonstrate large scale sequences for profiling of antibody repertoire.

3.3 B Cells Monitoring

Several investigations have been reported and a number of research groups have developed approaches for deep-sequencing IgH (or in some cases the IgL), or solely the CDR3, in genomic DNA or cDNA generated from bulk RNA isolated from individuals with autoimmune disease, [30,31] infections [32] or cancer, [33,34] or from vaccine recipients [30]. In addition, molecular ‘barcodes’ are used to enhance single-chain or CDR3 sequencing, that is, random hexamers are used as primers [28,35,36] or adapters with a unique molecular identifier [25,37] to tag each cDNA. Base-calling errors essentially corrected by the molecular bar coding which is inherent to the sequencing platform or of PCR bias. Due to the arrangements of 50% of B cells having multiple immunoglobulin gene loci a result of nonproductive rearrangements and receptor editing, and sequencing of immunoglobulin genes in genomic DNA is challenging [38,39]. Additionally, receptor editing is a vital attribute in B-cell maturation that assists to avoid autoimmunity, which can lead to allelic inclusion at the immunoglobulin IgL loci and the development of B cells that co express two different immunoglobulin IgLs [38,39]. These methods generally entail the sequencing of only

one immunoglobulin chain and, therefore, afford restricted information about the antibodies produced during an immune response. Although both IgHs and IgLs can be sequenced simultaneously in bulk mRNA, pairing of the IgH and IgL sequences generated in this manner can only be stabbed according to resemblances in sequence frequency [40]. Without the precise pairing of the cognate IgHs and IgLs, the information obtained about an antibody response is partial. Although not optimal for analysis of antibody repertoires, single-chain sequencing from bulk RNA can be used to track malignant or clonal populations of B cells by detecting their unique IgH, IgL or CDR3 sequences [41]. This technique is verifying to be influential for distinguishing and monitoring the return of B-cell malignancies and has enhanced sensitivity for detecting residual disease or disease reappearance than flow cytometry approaches [42,43]. The IgH CDR3 sequencing was additionally used in the search for biomarkers of autoimmune diseases and characterizing B-cell responses both for infection and vaccination. Furthermore, it has significant potential to immunomodulation in patients with cancer. Results from bulk IgH sequencing of unsorted B cells implies that direct examination of isolated B-cell subsets might be essential for attaining the most informative data regarding diseases of immune function [41].

3.4 Functional Antibody Repertoire

There are many challenges that have been involved during study of antibody repertoire. These challenges must be overcome to recognize and classify useful antibody repertoires. Wide-ranging surveys of the B-cell repertoire necessitate sequences that allow precise recreation of endogenous, bioinformatically selected antibodies. So that, their antigen specificity, binding and functional properties can be easily characterized. To recreate accurately the recombinant versions of endogenous antibodies, one have to sequence and properly pair the cognate IgHs and IgLs expressed by individual B cells. This will ensure that, the sequencing wrap the complete variable regions encoding the CDR1, CDR2, CDR3 and framework regions of both the IgHs and IgLs to identify error-free sequence [27,42,43]. The challenges for analysis of antibody repertoires are high-process fidelity and quality. In addition, to allow bioinformatic discovery of important functional antibodies, antibody repertoire sequencing must be applied to the suitable B-cell subsets obtained from individuals with an immunological phenotype of interest. There are several types of new single-cell sequencing methods can preserve the endogenous pairing of IgHs and IgLs. These methods include linkage-PCR based, barcode-based, bead-based, microwell-based and droplet-based methods, with most approaches integrating more than one of these methods [20].

3.5 Applications of High-through Put Sequencing

3.5.1 Generation of recombinant antibodies

Recombinant antibody technology is opening new point of view for the advances of novel therapeutic and diagnostic agents. Recent development of methods for preparation of recombinant DNA libraries and in silico bioinformatics approaches for protein structure analysis makes possible antibody preparation using gene engineering approaches. The development of gene engineering methods allowed creating recombinant antibodies and improving characteristics of existing antibodies; this significantly extends the applicability of antibodies. Modifying biochemical and immunochemical properties of antibodies by changing their amino acid sequences, it is possible to create antibodies with properties optimal for certain tasks. Monoclonal antibodies are a mainstay of diagnostic tests, therapeutics and research tools. High-throughput sequencing of antibody repertoires is a powerful technique to generate recombinant monoclonal antibodies directly from humans or other animals during a functional immune response. Therapeutic monoclonal antibodies have revolutionized care for patients with autoimmune diseases (anti-TNF, anti-IL-6, anti-IL-1 and anti-IL-12p40 monoclonal antibodies), infection (anti-respiratory syncytial virus antibody) and cancer (anti-CD20, anti-CTLA4, and anti-programmed-death 1 [PD-1] antibody). Investigators in my laboratory used barcode-enabled antibody repertoire sequencing of plasma blasts isolated from *Staphylococcus aureus*-infected humans to generate recombinant antibodies that bind and mediate killing of the bacteria [41]. Others have isolated individual B cells in microwells, then performed linkage RT-PCR and sequencing to generate tetanus toxoid (TT) reactive recombinant antibodies from TT-sorted peripheral blood plasmablasts isolated from a TT-vaccinated human [44].

3.5.2 Autoimmunity

There is barcode-enabled antibody repertoire sequencing used for peripheral blood plasmablasts. The report indicates sequence derived from patients with RA to identify sequences encoding autoantibodies that target citrullinated fibrinogen and citrullinated enolase [45] and IgH (single-chain) sequencing of bulk RNA. The sample isolated from synovium of patients with RA to identify dominant clones utilizing the IGHV4–34 gene segment and having CDR3s longer than those of antibodies expressed by naive B cells [46]. Sequences encoding pathogenic autoantibodies have the potential to be used as analytical biomarkers to discover individuals likely to develop disease [47]. There is a suggestion that critically describes that, mucosal sites including the lung and oral cavity, might be vital in the instigation of ACPA reaction that lead to the development of RA [48]. Some antibody repertoire sequencing technologies [45] generates sequencing reads that extend sufficiently far into the antibody constant region to identify the antibody isotype and subclasses, thereby enabling IgA ACPA-expressing B cells to be pinpointed to gain insight into the potential environmental exposures and microbial infections that initiate their production in RA. These sequences also potentially to be used as pharmacodynamic biomarkers during drug response mechanism. This can be usual to monitor the response to an immunomodulatory drug, or to expedite clinical development by demonstrating the activity of these drugs in proof-of-concept studies [47].

3.5.3 Vaccine development

Vaccine development is a long, complex process, often lasting 10-15 years and involving a combination of public and private involvement. The current system for developing, testing, and regulating vaccines developed during the 20th century as the groups involved standardized their procedures and regulations. There are great needs for the expansion of vaccines for diversified types of pathogens exists. Antibody repertoire sequencing is used to identify the microbial antigens and epitopes targeted by effective antimicrobial antibody responses that naturally control infection in humans and other animals [4,5,49,39,50,51] thereby enabling the development of vaccines based on these antigens. In addition, antibody repertoire sequencing is used in clinical proof-of-concept trials to demonstrate that a candidate vaccine, adjuvant or vaccination regimen can induce protective immune responses [52].

3.5.4 Immunomodulatory drug development

Immunomodulatory drugs (IMiDs) are thalidomide analogues, which possess pleiotropic anti-myeloma properties including immune-modulation, anti-angiogenic, anti-inflammatory and anti-proliferative effects. Their development was facilitated by an improved understanding in myeloma (MM) biology and initiated a profound shift in the therapeutic approach towards MM. The Immunomodulatory 'checkpoint inhibitors' are revolutionizing the care of patients with autoimmune diseases and cancer cases. The Abatacept (CTLA4-Ig) blocks CD28-mediated activation of T cells and is effective in the cure of RA [30]. In patients with metastatic melanoma and other cancers, ipilimumab (anti-CTLA4 antibody) and pembrolizumab (anti-PD-1 antibody) are therapeutically effective by blocking inhibition of T cells [53,54]. Antibody repertoire analysis and sequencing provides an approach to monitor and characterize the immune responses induced by candidate immunomodulatory drugs [55] or vaccines [56]. Checkpoint inhibitor therapy in cancer results in the initiation of autoimmunity in a subset of patients, and antibody repertoire sequencing has the potential to be used to identify individuals at risk for developing autoimmune disease and to characterize the autoimmune response in individuals that do [53,55,57,58].

3.5.5 B-cell repertoire development

There is determination in the lineage development of B cells and evolution of the corresponding germline antibody genes by comparing the sequence of antibody repertoires in an individual. As such, the lineage development of B cells and evolution of antibody responses can be tracked [27,14,24]. Using this analysis, it's possible to investigate the evolution of functional antibody responses in health and disease, reveal mechanisms underlying the development of naive B-cell repertoires [57,58] and the activation, development and trafficking of B cells in immune responses of interest [50,58,59]. Several academic laboratories are in the process of utilizing antibody repertoire sequencing to investigate the development of anti-citrullinated-protein-reactive B cells in RA and anti-nuclear-antigen-reactive B cells in SLE. Such studies might identify autoantigens and

provide insight into disease initiation and progression in these and other diseases [60].

3.5.6 Infectious diseases

Considering infectious disease the Ig-seq is also offering insights towards the development adaptive immune responses elicited by pathogen challenge or vaccination. This pathogen challenge can affect the BCR repertoire of responding B as well as the naive repertoire. There are phenomenons that, many pathogens can be produce superantigens. These are proteins that bind to certain antibody V domains, resulting in BCR cross-linking and subsequent B-cell deletion [60,61]. One study reported that the depletion of V genes bound by superantigens was not observed in the naive repertoire of transgenic mice constitutively expressing superantigens [62]. Surprisingly, skewed naive B-cell antibody repertoire was also reported for patients with chronically evolving hepatitis C infection [63]. Furthermore, any changes regarding the overall antibody repertoire are also evident following vaccination or infection have been reported [64,65,66]. The report of Boyd, Fire and coworkers [65] surprisingly observed that the convergent antibody signatures (stereotyped CDR-H3 sequences) in patients developing acute dengue infection. This observation gives insight the possibility that Ig-seq intended at detecting stereotypical responses and it can be used as a diagnostic tool for predicting infectious disease severity. On the other hand, tracing the evolutionary paths that lead to the generation of bNAbs is also critical for the design of immunogens and vaccination schedules. This will have the probability to elicit an immune response by first activating naive of antibody-expressing B cells and then steering B clonal selection toward an affinity maturation pathway that leads to the production of bNAbs [48].

3.5.7 Cancer research and treatment

For the past several years there was seen progress in cancer immunotherapy and in recent time two cancer immunotherapy drugs which are a cell-based vaccine and an anti-CTLA-4 monoclonal antibody were approved [67,68]. In recent study, Vonderheide and Glennie [69] reported agonistic CD40 monoclonal antibodies (mAb) offer a new therapeutic option which has the potential to generate anticancer immunity by various mechanisms. The study depicted agonistic CD40 mAb have been shown to activate APC and promote antitumor T-cell responses and to foster cytotoxic myeloid cells with the potential to control cancer in the absence of T-cell immunity. There are different approaches yet discovered for cancer research and treatment. More relevant methods focused towards genomics that provides quantitative and economic advantages for biomarker analysis and shifts toward genomics to bolster with broadly quantitative, accuracy, sensitive and high throughput (Table 1).

Table 1. Methodologies used for analysis of cancer

Requirement	Technology/Method
Somatic mutation analysis	DNA deep Sequencing, Breakpoint analysis and Fusion Detection.
Gene expression analysis	RNA-Seq, Arrays, RNA panels.
Germ line variant analysis	CNV, Structural and Small Variant Detection
HLA characterization	HLA Calling (DNA, RNA, Arrays).
Antigen-specific immune response	B/T Cell Repertoire (DNA/RNA), VDJ mutation (RNA).

During antibody repertoire B-cell leukemias, lymphomas and multiple myeloma are malignancies that arise at different stages of B-cell development. As such, BCRs on malignant B cells represent a biomarker for the plenty of the malignant cell population. For disease detection and delineation of the degree to which antibody evolution in malignant cells correlates with disease progression, the Ig-seq of the V-gene repertoire in peripheral B cells, bone marrow, tumors and blood-borne free DNA samples has been used [69,70,71]. For example, Ig-seq of V genes in peripheral blood facilitated detection of cancerous cells and minimal residual disease following treatment of B-cell chronic lymphocytic leukemia (CLL) [66,70]. Moreover, the Ig-seq was also used to identify negligible residual disease in pediatric patients with B-cell acute lymphoblastic leukemia (B-ALL) [31]. On the other hand it can be also used as a marker of non-Hodgkins lymphoma [65]. The identified portion of Ig-seq was also naked that, the B-ALL patients exhibit a variety of clonotypic diversity, which arises predominantly from VH gene replacement [31,72]. The degree of heterogeneity of B-cell clones observed on CLL is small, with the degree varying based on whether the disease derived from unmutated or

somatically hypermutated B cells [68,73]. In opposite way, the malignant clonotype in various myelomas, a disease which occur from terminally differentiated plasma cells in the bone marrow that lack active mechanisms of antibody diversification, displayed little evidence of heterogeneity [74]. Study conducted by Hargreaves and his co-workers showed that, they have evaluated a suite of assays that has been employed in high throughput manner for genomic analysis of the FcγR locus that is scalable for application in large clinical trials of mAb therapy [75]. The result obtained from this study has promises for establishing the importance of FcγR genetics in predicting response to antibody therapeutics. The broader application of high-throughput sequencing for the detection of malignancy-stereotypic clonotypes, which may be present at a low frequency in subjects that do yet not display clinical disease, may prove to be a useful, early diagnostic tool. Moreover, the Ig-seq may also be practical to appreciate gammopathies conditions that result in peculiarly high levels of antibodies in serum [76,77].

4. CONCLUSION

The present review has illustrated the high-throughput sequencing and profiling of antibody repertoire for diagnostic and therapeutic applications. Antibodies are the major component of the adaptive immune system and have vital roles in protective and pathogenic immune responses. In response to microbial infection, vaccination, autoimmune disease or cancer, the immune system generates distinct antibody repertoires. Analysis of these antibody repertoires, particularly those contributing to functional immune responses, can provide important information on protective and pathogenic immunity. Antibody repertoire analysis will transform our understanding of immune responses to autoimmunity, vaccination, infection and cancer, providing new biomarkers and diagnostic tools, and enabling efficient generation of therapeutic antibodies. The advent of high-throughput DNA sequencing has enabled the determination of the antibody gene repertoire at extraordinary depth that was unimaginable. Technologies to improve sequence precision and data analysis are being developed at a breakneck pace, reshaping our understanding of many important aspects of B-cell immunology and increasingly affecting clinical diagnosis, antibody drug discovery and vaccine development.

CONSENT

It is not applicable.

ETHICAL APPROVAL

It is not applicable.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Leadbetter EA, et al. Chromatin-IgG complexes activate B cells by dual engagement of IgM and Toll-like receptors. Nature. 2002;416:603–607.
[PubMed: 11948342]
2. Rosen A, Casciola-Rosen L. Autoantigens as substrates for apoptotic proteases: Implications for the pathogenesis of systemic autoimmune disease. Cell Death Differ. 1999;6:6-12.
[PubMed: 10200542]
3. Tan EM. Autoantibodies, autoimmune disease, and the birth of immune diagnostics. J. Clin. Invest. 2012;122: 3835–3836.
[PubMed: 23154275]
4. Wilson PC, Andrews SF. Tools to therapeutically harness the human antibody response. Nat. Rev. Immunol. 2012;10:709–719.
[PubMed: 23007571]
5. Burton DR, et al. A blueprint for HIV vaccine discovery. Cell Host Microbe. 2012;12:396–407.
[PubMed: 23084910]
6. Haynes BF, Kelsoe G, Harrison SC, Kepler TB. B-cell-lineage immunogen design in vaccine development with HIV-1 as a case study. Nat. Biotech. 2012;30:423–433.
7. Schroeder HWJR. Similarity and divergence in the development and expression of the mouse and human antibody repertoires. Dev. Comp. Immunol. 2006;30:119–135.
8. Weinstein JA, Zeng X, Chien YH, Quake, SR. Correlation of gene expression and genome mutation in single B-cells. Plos One. 2013;8:e67624.
9. Danilova N, Amemiya CT. Going adaptive: The saga of antibodies. Annals of the New

York Academy of Sciences. 2009;1168: 130-155.

10. Fagarasan S. Evolution, development, mechanism and function of IgA in the gut. Current Opinion in Immunology. 2008;20: 170-177.
11. Brady BL, Steinel NC, Bassing CH. Antigen receptor allelic exclusion: An update and reappraisal. J. Immunol. 2010; 185:3801–3808.
12. Georgiou G, et al. The promise and challenge of high-throughput sequencing of the antibody repertoire. Nat. Biotech. 2014;32:158–168.
DOI:10.1038/nbt.2782
13. Xu JL, Davis MM. Diversity in the CDR3 region of V(H) is sufficient for most antibody specificities. Immunity. 2000;13: 37–45.
14. Ippolito GC, et al. Forced usage of positively charged amino acids in immunoglobulin CDR-H3 impairs B cell development and antibody production. J. Exp. Med. 2006;203:1567–1578.
15. McHeyzer-Williams M, Okitsu S, Wang N, McHeyzer-Williams L. Molecular programming of B cell memory. Nat. Rev. Immunol. 2012;12:24–34.
16. Victora GD, Nussenzweig MC. Germinal centers. Annu. Rev. Immunol. 2012;30: 429–457.
17. Capra JD, Kehoe JM. Variable region sequences of five human immunoglobulin heavy chains of the VH3 subgroup: Definitive identification of four heavy chain hypervariable regions. Proc. Natl. Acad. Sci. USA. 1974;71:845–848.
18. Hibi T, Dosch HM. Limiting dilution analysis of the B cell compartment in human bone marrow. Eur. J. Immunol. 1986;16:139–145.
19. Zhang Z. VH replacement in mice and humans. Trends Immunol. 2007;28:132–137.
20. Klein F, et al. HIV therapy by a combination of broadly neutralizing antibodies in humanized mice. Nature. 2012;492:118–122.
[PubMed: 23103874]
21. Robinson WH. Sequencing the functional antibody repertoire-diagnostic and therapeutic discovery. Nat Rev Rheumatol. 2015;11(3):171–182.
DOI:10.1038/nrrheum.2014.220
22. Scheid JF. et al. Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. Nature. 2009;458:636–640.
23. Corti D, et al. A neutralizing antibody selected from plasma cells that binds to group 1 and group 2 influenza A hemagglutinins. Science. 2011;333:850–856.
[PubMed: 21798894]
24. Amara K, et al. Monoclonal IgG antibodies generated from joint-derived B cells of RA patients have a strong bias toward citrullinated autoantigen recognition. J. Exp. Med. 2013;210:445–455.
[PubMed: 23440041]
25. Jaitin DA, et al. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. Science. 2014;343:776–779.
[PubMed: 24531970]
26. Finn JA, Crowe JE Jr. Impact of new sequencing technologies on studies of the human B cell repertoire. Curr. Opin. Immunol. 2013;25:613–618.
[PubMed: 24161653]
27. Georgiou G, et al. The promise and challenge of high-throughput sequencing of the antibody repertoire. Nat. Biotech. 2014;32:158–168.
28. Palanichamy A, et al. Immunoglobulin class-switched B cells form an active immune axis between CNS and periphery in multiple sclerosis. Sci. Transl. Med. 2014;6:248ra106.
29. Parameswaran P, et al. Convergent antibody signatures in human dengue. Cell Host Microbe. 2013;13:691–700.
[PubMed: 23768493]
30. Wu D, et al. High-throughput sequencing detects minimal residual disease in acute T lymphoblastic leukemia. Sci. Transl. Med. 2012;4:134ra163.
31. Faham M, et al. Deep-sequencing approach for minimal residual disease detection in acute lymphoblastic leukemia. Blood. 2012;120:5173–5180.
[PubMed: 23074282]
32. Boyd SD, et al. Measurement and clinical monitoring of human lymphocyte clonality by massively parallel VDJ pyrosequencing. Sci. Transl. Med. 2009;1:12ra23.
33. De Vlaminck I, et al. Temporal response of the human virome to immunosuppression and antiviral therapy. Cell. 2013;155: 1178–1187.
[PubMed: 24267896]
34. Vollmers C, Sit RV, Weinstein JA, Dekker CL, Quake SR. Genetic measurement of

memory B-cell recall using antibody repertoire sequencing. Proc. Natl Acad. Sci. USA. 2013;110:13463–13468.
[PubMed: 23898164]

35. Islam S, et al. Quantitative single-cell RNA-seq with unique molecular identifiers. Nat. Methods. 2014;11:163–166.
[PubMed: 24363023]
36. Meffre E, Wardemann H. B-cell tolerance checkpoints in health and autoimmunity. Curr. Opin. Immunol. 2008;20:632–638.
[PubMed: 18848883]
37. Halverson R, Torres RM, Pelanda R. Receptor editing is the main mechanism of B cell tolerance toward membrane antigens. Nat. Immunol. 2004;5:645–650.
[PubMed: 15156139]
38. Reddy ST, et al. Monoclonal antibodies isolated without screening by analyzing the variable-gene repertoire of plasma cells. Nat. Biotech. 2010;28:965–969.
39. Lu DR, Robinson WH. Street-experienced peripheral B cells traffic to the brain. Sci. Transl. Med. 2014;6:248fs231.
40. Loman NJ, et al. Performance comparison of benchtop high-throughput sequencing platforms. Nat. Biotech. 2012;30:434–439.
41. Nguyen P, et al. Identification of errors introduced during high throughput sequencing of the T cell receptor repertoire. BMC Genomics. 2011;12:106.
[PubMed: 21310087]
42. DeKosky BJ, et al. High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. Nat. Biotech. 2013;31:166–169.
43. Tan YC, et al. Barcode-enabled sequencing of plasmablast antibody repertoires in rheumatoid arthritis. Arthritis Rheumatol. 2014;66:2706–2715.
[PubMed: 24965753]
44. Doorenspleet ME, et al. Rheumatoid arthritis synovial tissue harbours dominant B-cell and plasma-cell clones associated with autoreactivity. Ann. Rheum. Dis. 2014; 73:756–762.
[PubMed: 23606709]
45. Robinson WH, Lindstrom TM, Cheung RK, Sokolove J. Mechanistic biomarkers for clinical decision making in rheumatic diseases. Nat. Rev. Rheumatol. 2013; 9:267–276.
[PubMed: 23419428]
46. Demoruelle MK, Deane KD, Holers VM. When and where does inflammation begin in rheumatoid arthritis? Curr. Opin. Rheum. 2014;26:64–71.
47. Kwong PD, Mascola JR, Nabel GJ. Broadly neutralizing antibodies and the search for an HIV-1 vaccine: the end of the beginning. Nat. Rev. Immunol. 2013;13: 693–701.
[PubMed: 23969737]
48. Laserson U, et al. High-resolution antibody dynamics of vaccine-induced immune responses. Proc. Natl Acad. Sci. USA. 2014;111:4928–4933.
[PubMed: 24639495]
49. Wiley SR, et al. Targeting TLRs expands the antibody repertoire in response to a malaria vaccine. Sci. Transl. Med. 2011; 3:93ra69.
50. Couzin-Frankel J. Breakthrough of the year cancer immunotherapy. Science. 2013; 342:1432–1433.
[PubMed: 24357284]
51. Riley JL. Combination checkpoint blockade taking melanoma immunotherapy to the next level. N. Engl. J. Med. 2013;369:187–189.
[PubMed: 23724866]
52. Cha E, et al. Improved survival with T cell clonotype stability after anti-CTLA-4 treatment in cancer patients. Sci. Transl. Med. 2014;6:238ra270.
53. Garren H, et al. Phase 2 trial of a DNA vaccine encoding myelin basic protein for multiple sclerosis. Ann. Neurol. 2008; 63:611–620.
[PubMed: 18481290]
54. Hodi FS, et al. Improved survival with ipilimumab in patients with metastatic melanoma. N. Engl. J. Med. 2010; 363:711–723.
[PubMed: 20525992]
55. Watson CT, et al. Complete haplotype sequence of the human immunoglobulin heavy-chain variable, diversity, and joining genes and characterization of allelic and copy-number variation. Am. J. Hum. Genet. 2013;92:530–546.
[PubMed: 23541343]
56. Von Budingen H, et al. Frontline: epitope recognition on the myelin/oligodendrocyte glycoprotein differentially influences disease phenotype and antibody effector functions in autoimmune demyelination. Eur. J. Immunol. 2004;34:2072–2083.
[PubMed: 15259004]
57. Stern JN, et al. B cells populating the multiple sclerosis brain mature in the draining cervical lymph nodes. Sci. Transl. Med. 2014;6:248ra107.

58. Quach H, et al. Mechanism of action of immunomodulatory drugs (IMiDS) in multiple myeloma 2010;24(1):22-32.
59. Wardemann H, et al. Predominant autoantibody production by early human B cell precursors. Science. 2003;301:1374–1377.
[PubMed: 12920303]
60. Gronwall, C, Kosakovsky Pond SL, Young JA, Silverman GJ. *In vivo* VL-targeted microbial superantigen induced global shifts in the B cell repertoire. J. Immunol. 2012;189,850–859.
61. Aoki-Ota M, Torkamani A, Ota T, Schork N, Nemazee D. Skewed primary Igkappa repertoire and V-J joining in C57BL/6 mice: implications for recombination accessibility and receptor editing. J. Immunol. 2012; 188:2305–2315.
62. Racanelli V. et al. Antibody V(h) repertoire differences between resolving and chronically evolving hepatitis C virus infections. PLoS ONE 2011;6:e25606.
63. Jiang N. et al. Lineage structure of the human antibody repertoire in response to influenza vaccination. Sci. Transl. Med. 2013;5:171ra119.
64. Parameswaran P. et al. Convergent antibody signatures in human dengue. Cell Host Microbe. 2013;13:691–700.
65. Krause JC, et al. Epitope-specific human influenza antibody repertoires diversify by B cell intraclonal sequence divergence and interclonal convergence. J. Immunol. 2011; 187:3704–3711.
66. Boyd SD. et al. Measurement and clinical monitoring of human lymphocyte clonality by massively parallel VDJ pyrosequencing. Sci. Transl. Med. 2009;1:12ra23.
67. Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. N Engl J Med. 2012;366:2443-54.
68. Brahmer JR, Tykodi SS, Chow LQ, Hwu WJ, Topalian SL, Hwu P, et al. Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. N Engl J Med. 2012;366:2455–65.
69. Robert H. Vonderheide, Martin J. Glennie. Agonistic CD40 Antibodies and Cancer Therapy. American Association for Cancer Research; 2013.
DOI:10.1158/1078-0432.CCR-12-2064
70. Campbell PJ, et al. Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing. Proc. Natl. Acad. Sci. USA 2008;105:13081–13086.
71. He J, et al. IgH gene rearrangements as plasma biomarkers in Non- Hodgkin's lymphoma patients. Oncotarget. 2011;2: 178–185.
72. Gawad C. et al. Massive evolution of the immunoglobulin heavy chain locus in children with B precursor acute lymphoblastic leukemia. Blood. 2012;120, 4407–4417.
73. Logan AC, et al. High-throughput VDJ sequencing for quantification of minimal residual disease in chronic lymphocytic leukemia and immune reconstitution assessment. Proc. Natl. Acad. Sci. USA. 2011;108:21194–21199.
74. Darzentas N, Stamatopoulos K. Stereotyped B cell receptors in B cell leukemias and lymphomas. Methods Mol. Biol. 2013;971:135–148.
75. Hargreaves CE, Iriyama C, Rose-Zerilli MJJ, Nagelkerke SQ, Hussain K, Ganderton R, et al. Evaluation of high-throughput genomic assays for the Fc gamma receptor locus. Plos One. 2015; 10(11):e0142379.
DOI:10.1371/journal. pone.0142379
76. Varettoni M. et al. Clues to pathogenesis of Waldenstrom macroglobulinemia and immunoglobulin M monoclonal gammopathy of undetermined significance provided by analysis of immunoglobulin heavy chain gene rearrangement and clustering of B-cell receptors. Leuk. Lymphoma. 2013;54:2485–2489.
77. Kyle RA. et al. Prevalence of monoclonal gammopathy of undetermined significance. N. Engl. J. Med. 2006;354:1362–1369.

Peer-review history:
The peer review history for this paper can be accessed here:
http://sciencedomain.org/review-history/17049