The Dissertation Committee for Bo Wang Certifies that this is the approved version of the following dissertation:

High-throughput analysis of native antibody repertoires for therapeutics discovery

Committee:

Andrew D. Ellington, Co-Supervisor

George Georgiou, Co-Supervisor

Lauren I. Ehrlich

Hal S. Alper

Jennifer A. Maynard

High-throughput analysis of native antibody repertoires for therapeutics discovery

by

Bo Wang, B.E.; M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2016

Dedication

To my mom and dad, for their support, and my wife, for her encouragement and company, and my son

Acknowledgements

I would like to express my appreciation to Andy, who offered me a position in his lab, and allowed me to explore my interest for the past five and half years. And I am quite proud to say that I am the only chemical engineer who survived in Ellington lab. I am also indebted to George, who always gave excellent directions and helped me not to go too far away in the scientific world. Both Andy and George are very insightful, and I have learned a lot from them, not only their scientific visions, but also their enthusiasm and love for science. They are great role models for my future career.

It is quite a unique journey for a Ph.D. student. Ellington lab and Georgiou lab are quite different, and I have benefited from both. I would like to thank all the present and past Ellington and Georgiou lab members with whom I have worked with, not only for their invaluable help, but also for their influence on me about seeking dreams. Their ways of doing science also greatly influenced me.

Special thanks to Alex, who talked me into working on proteins instead of aptamers on my first day in Ellington lab, and opened the door of the wonderful world of protein computational design and protein engineering to me; Christien, who taught me how to do animal immunization, and lent me Janeway's immunology book, which made me a real immunologist; Oana, who taught me how to use Rosetta, the best protein computational design software; Scott, who unselfishly shared his tips for job hunting and his interview experience, which is invaluable; Chang-Han, who helped a lot with initial yeast display setup; Brandon, whom I learned V_H:V_L pairing from, which is the foundation for a lot of work here, and I collaborated with on many projects; and Greg, who is the 'pubmed man' in the lab that can always answer my immunology questions.

Also, I would like to thank my wife Ying, for her support and encouragement. I would not be able to get here without your company. And my mom and dad, for their support. And my son David, for the joy he brings to me.

High-throughput analysis of native antibody repertoires for therapeutics discovery

Bo Wang, Ph.D.

The University of Texas at Austin, 2016

Co-Supervisors: Andrew D. Ellington, George Georgiou

Adaptive immunity is the foundation of recognition and protection against a diverse array of pathogens. The humoral arm of adaptive immunity whose most significant effector mechanism relies on antibodies is critical for protection against many viral and bacterial pathogens. Additionally, antibodies are extremely important for clinical medicine either as therapeutics or for detection purposes. As a result, there is great interest in the development of technologies for isolating therapeutically or diagnostically useful antibodies from immunized animals or from patients.

In this work I describe several high throughput approaches for the "mining" of the mammalian adaptive antibody response for the purpose of isolating antibodies with high affinity and selectivity towards desired antigens. The first technology capitalizes on the fact that the antibody repertoire encoded by B cells in the draining lymph node from the immunization site is highly enriched for antibodies specific to the immunogen. The second technology takes advantage of repertoire analysis by next-generation sequencing and yeast display for antibody discovery. The third technology integrates high-throughput V_H:V_L pairing with yeast surface display to enable rapid, high-throughput screening of all native antibodies in the repertoire. These three technologies have been used to identify a

large panel of antibodies against several different pathogens, including Ebola virus, ricin, influenza virus, and HIV-1 virus.

These technologies will continue to play a critical role in adaptive immunity exploration for new therapeutics discovery, and in characterization of immune responses elicited by vaccination or natural infection to guide design of more effective vaccines.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

The immune system is specialized to recognize and fight against various arrays of exogenous and self pathogens (in autoimmunity). As the first defense, innate immunity, including complement pathways, phagocytosis by neutrophils, macrophages, and dendritic cells, and cytotoxicity by natural killer cells (NK) and eosinophils, helps to clear most common pathogens directly. However, when innate immunity fails (Kumar et al., 2011), adaptive immune responses that are tailored to recognize molecular features of the pathogen that is posing a challenge to the organism takes over. Adaptive immunity provides the basis for recognition of, and protection from, numerous pathogenic cells and viruses. There are two branches of adaptive immunity: B cell based humoral immunity and T cell based cellular immunity (Boehm, 2011). The effector arm of B cell immunity is mediated by the production of antibodies, while T cells destroy cells with intracellular pathogens and help with B cell activation.

ANTIBODY STRUCTURES

Antibodies are a class of protein molecules secreted by B cells. They can not only recognize pathogens, but also opsonize them to initiate pathogen clearance by other effector cells. Hence, antibodies are very important effector molecules that protect host organisms from pathogenesis, and they connect adaptive immunity with innate immunity. The modular structures of antibodies are indispensable for their functions, which also make them convenient to manipulate. This modularity also forms part of the foundation of antibody discovery and engineering.

Antibodies are 'Y' shaped symmetric homodimers of heterodimers, and each heterodimer is made of 1 heavy chain and 1 light chain. In humans, there are 5 different kinds of heavy chains (isotypes), IgG, IgA, IgE, IgM, and IgD, , which are encoded by

constant γ (including 4 subtypes, γ1, γ2, γ3, and γ4, with γ1 being the most abundant), α (including 2 subtypes, α1 and α2), ε, μ, and δ regions, respectively (Fig. 1-1). There are also 2 different kinds of light chains, kappa (κ) and lambda (λ), which can both pair with heavy chains. The heavy chain has 4 immunoglobulin superfamily (IgSF) domains (5 for IgM and IgE), namely V_H (variable heavy), C_{H1} (constant heavy), C_{H2}, and C_{H3} (C_{H4} for IgM and IgE), while the light chain has 2 IgSF domains, V_L (variable light) and C_L (constant light). Each IgSF domain is a 70-110 amino acid (aa) long, sandwich-like structure formed by antiparallel β-sheets that have a hydrophobic inner core and conserved cysteines that form disulfide bonds. Each heterodimer of 1 heavy chain and 1 light chain is also connected by 1 conserved disulfide bond between the C_{H1} and C_L domains, The two heterodimers are also connected by conserved disulfide bonds in the hinge region. IgG, IgD, and IgE antibodies are monomeric, while IgM are pentameric, and IgA are dimeric or monomeric. The multimerization of IgM and IgA is mediated by a small polypeptide called the J chain. Multimerization can significantly enhance the apparent affinity of IgM or IgA and also facilitates transport, especially of IgA across epithelial cells for secretion to mucosal surfaces.

The entire antibody molecule can be separated into 3 parts: 2 antigen binding fragments (Fab) and 1 crystallizable fragment (Fc). The Fab fragment binds antigen with contributions from both the V_H and V_L, while the Fc fragment binds to Fc receptors expressed on different cells. This interaction is very important for antibody effector functions, extending the antibody half-life, and linking innate immunity to adaptive immunity. Different antibody isotypes have diverse effector functions. IgG antibodies usually elicit cytotoxicity and phagocytosis by binding to different Fcγ receptors expressed on various innate immune cells (including natural killer cells (NK), neutrophils, macrophages, and dendritic cells), IgE antibodies induce inflammation by

binding to Fcε receptors expressed on mast cells, while IgA antibodies elicit phagocytosis by binding to Fcα receptors expressed on certain cellular subsets (including neutrophils, macrophages, and Kuppfer cells). Another important Fc receptor is the neonatal Fc receptor (FcRn), which only IgG can bind to. The binding of IgG to FcRn provides the molecular basis for the long half-life of IgG in the serum, which explains why most therapeutic antibodies are of IgG isotype. Different FcRs have cytoplasmic domains that contain different signaling motifs, either activating immunoreceptor tyrosine-based activation motifs (ITAMs) or inhibiting immunoreceptor tyrosine-based inhibitory motifs (ITIMs), which determine downstream signaling pathways after antibody binding to the FcR. Notably, the Asn-Ser-Thr motif in the C_{H2} domain where Asn is glycosylated is indispensable for FcR binding, and the glycan composition has large effects on FcR specificity and affinity.

Previous research has shown that there are 3 regions within the V_H and V_L domains that show very high variability, namely heavy chain complementary determining regions (CDRH) 1, 2, 3, and light chain CDRL 1, 2, 3 (Johnson and Wu, 2004) (Fig. 1-2). These 6 CDRs are supported by 4 framework regions (FR) on each chain, and they determine much of the specificity and affinity of the antibody to the antigen by forming the main contact interface between the two. This provides the structural basis of antibody recognition of diverse antigens. Within these 6 CDRs, CDRH3 is the most variable and is typically thought to contribute the most to antigen binding. It is noteworthy to mention that these 6 CDRs, especially CDRH3, are not only variable in the amino acid compositions, but also in their lengths. The discovery of HIV-1 broadly neutralizing antibodies (bnAbs) that have longer than 30 aa CDRH3s shows the effectiveness of long CDRs to penetrate viral glycan shields to target conserved epitopes (Doria-Rose et al., 2014).

On the other hand, whether such a diverse amino acid composition in CDRs is essential for binding is under investigation, as previous research has shown that high-affinity binding to some antigens can still be achieved with just Tyr and Ser in CDRs (Fellouse et al., 2004; Persson et al., 2013). However, a more diverse aa composition provides a larger sequence space for the immune system to search through, which ensures a higher success rate to identify antibodies specific for an antigen.

As a result of the large size of full-length antibodies, several different formats of small antibody fragments have been engineered. The most common format is that of the single chain fragment variable (scFv), which is a V_H linked to a V_L with a flexible linker, usually $(Gly_4Ser)_3$, and the other common format is that of the Fab (Fig. 1-3). These fragments can be easily functionally expressed in *E. coli* and yeast with high yields. However, the reconstitution of a scFv as an IgG antibody after the identification of antigen-specific clones is usually necessary, and sometimes this may result in a loss of binding. It is noteworthy to mention that some species, including lamas, produce antibodies that only have heavy chains, from which nanobodies (V_H) have been designed (Muyldermans, 2013). Nanobodies are the smallest antigen binding fragment.

Since the Fc fragment is not involved in antigen binding, only Fab fragments need to be engineered to select for antibodies that bind different antigens. Furthermore, as CDRs contribute the most to antigen binding, the antibody selection procedure can be focused more specifically on CDR engineering.

Finally, the modular structure of antibodies also inspired the design and engineering of novel antibody formats, with the most well-known one being that of a bispecific antibody, which has two different Fab fragments in one single antibody molecule (Chan and Carter, 2010). Several different techniques have been developed for bispecific antibody production, including Fab arm exchange (Labrijn et al., 2013), and

the introduction of mutations to favor heterodimer formation and prevent homodimer formation (Lewis et al., 2014). Spiess *et al*. recently demonstrated co-culture of E. coli strains that express two half-antibodies (1 heavy chain and 1 light chain), and purification and combination of the half-antibodies to make the entire bispecific antibodies (Spiess et al., 2013). The US Food and Drug Administration (FDA) approval of the bispecific Blinatumomab (CD19 and CD3 targeting) demonstrated the therapeutic potential of bispecific antibodies and encouraged more effort towards their development (Goebeler et al., 2016).

Other formats of antibodies include the minibody, diabody, tetrabody, and the IgG-scFv (Chan and Carter, 2010). However, the therapeutic potential of these antibody formats are yet to be tested, and a large potential problem for *in vivo* use of these antibody formats is their immunogenicity, which is a result of their 'irregular' structures compared to that of normal antibodies.

IMMUNOGLOBULIN LOCI AND B CELL DEVELOPMENT

Based on the well-known International Immunogenetics Information System (IMGT) database (Giudicelli and Lefranc, 1999; Giudicelli et al., 2006; Lefranc et al., 2009), there are 56 V (variable), 23 D (diversity), and 6 J (junction) segments for the V_H domain, 38 V and 5 J for V_κ domain, and 35 V and 7 J for V_λ domain in the human genome. These genes are located on chromosomes 14 (14q32.33), 2 (2p11.2), and 22 (22q11.2), respectively. During the development of B cells, these genes need to be joined together to form productive V_H and V_L domains.

The unique developmental process of B cells provides the foundation for the generation of a diverse repertoire of antibodies (Dudley et al., 2005; Li et al., 2004; Nadel

and Feeney, 1997; Tonegawa, 1983). There are two different sequence diversification mechanisms, somatic recombination and somatic hypermutation.

The first stage of B cell development occurs in the bone marrow and ends with the formation of IgM^+ IgD^+ naïve B cells. First, pro-B cells are derived from common lymphoid progenitors (CLP). During this process D and J genes are recombined first, followed by V genes recombining with DJ segments. After VDJ recombination, pre-B cell receptors (BCR) composed with rearranged productive heavy chain paired with a surrogate light chain consisting of VpreB and λ5 will be expressed on the surface of large pre-B cells. As a result of errors in gene rearrangement, cells with an unproductive pre-BCR will become apoptotic and be eliminated. Hence, this serves as the first checkpoint for B cell development.

After the large pre-B cell stage, V and J alleles for the light chain are recombined at the small pre-B cell stage. Productive VJ rearrangement results in the successful generation of a BCR , which replaces the pre-BCR on the surface. At this stage, only IgM is expressed on immature B cell surface. These IgM^+ immature B cells are tested for auto-reactivity, which serves as the second checkpoint for B cell development. Auto-reactive B cells can undergo either receptor editing (further gene rearrangement of light chain in a B cell that already has a productive light chain gene rearrangement) to have another chance to generate another BCR that is not auto-reactive, or become apoptotic and are eliminated. After this, the immature B cells start to express IgD on the surface and migrate to secondary lymphoid organs (lymph nodes, spleen). The IgM^+IgD^+ cells are now mature naïve B cells.

Somatic recombination of V(D)J alleles is catalyzed by the recombination-activating gene (RAG) complex composed of RAG-1 and RAG-2, the Ku70:80 complex, DNA-dependent protein kinase (DNA-PK), the Artemis complex, DNA ligase IV, and

terminal deoxynucleotidyl transferase (TdT). The RAG complex first recognizes the recombination signal sequences (RSS) that flank V(D)J alleles and brings them in proximity to excise parts of the chromosome. The to-be-joined segments with palindromic overhangs (P nucleotides) are then brought together, where TdT may then introduce non-templated nucleotides (N nucleotides). After that, the other complexes correct and complete the complementary strand, which leaves behind palindromic sequences. As a result of N/P nucleotide addition, the theoretical diversity of an antibody repertoire is estimated to be more than 10^{12}. However, some studies using next-generation sequencing (NGS) showed that the actual diversity is only about 10^{7} (Arnaout et al., 2011; Boyd et al., 2009; Glanville et al., 2009). N/P nucleotides addition results in higher diversity in the CDRH3 and CDRL3 than that expected from V(D)J recombination only.

Recent studies have shown that V(D)J recombination can also occur with other genes (Tan et al., 2016). Tan *et al*. isolated broadly reactive malaria antibodies from two *P. falciparum*-negative donors, and found these antibodies have unusually long CDRH3s that are recombined not only from V(D)J alleles on chromosome 14, but also with LAIR1, an IgSF inhibitory receptor on chromosome 19 (Tan et al., 2016). The mechanism for this rare recombination event is still under investigation, but the genome instability of B cells after malaria infection may contribute to recombination of V(D)J and LAIR1 on different chromosomes (Robbiani et al., 2015).

Before migration to secondary lymphoid organs, some mature naïve B cells develop into B-1 cells, which reside in pleural cavities and the peritoneum. B-1 cells produce natural antibodies, which usually do not undergo somatic hypermutation, and unlike B-2 cells that further develop in secondary lymphoid organs, they don't need T cell help. Natural antibodies can usually only protect from pathogens with repeated molecular patterns (for example, phospholipids and carbohydrates). In contrast, B-2 cells

need T cell help for activation. In secondary lymphoid organs, these B cells reside in the follicles, and once activated through the binding to cognate antigens by their BCRs, they rapidly proliferate and differentiate into plasmablasts, and eventually short-lived plasma cells, both of which can secrete antibodies. A portion of activated B cells undergo germinal center reactions, during which they affinity mature their BCR for cognate antigens. In germinal centers, B cells travel back and forth between a light zone and dark zone. In the light zone, B cells with distinct BCRs that have different affinities to cognate antigens compete for the antigens, as well as for follicular dendritic cell and follicular T cell help. Those B cells with high affinities to antigens are then selected and stimulated. These stimulated B cells travel to dark zone and proliferate there. During proliferation, activation-induced cytidine deaminase introduces random mutations into the V_H and V_L domains. These B cells with mutated BCR sequences then travel back to the light zone, where a small portion of them that have beneficial affinity-improving mutations can get further help. Then they travel back again to dark zone to proliferate. The repeated cycling of B cells between the light zone and dark zone leads to BCR affinity maturation. Somatic hypermutation is not only the basis for the secondary diversification of antibody repertoires, but also controls antibody isotype switching (class-switch recombination).

Perhaps the most significant example of somatic hypermutation is the identification of heavily mutated HIV-1 bnAbs (Corti and Lanzavecchia, 2013). The selection pressure that the constantly mutating HIV-1 virus exerts on the immune system, especially on B cells, forces them to constantly refine their BCR affinity and specificity. Different lineages of bnAbs that target different vulnerable sites on HIV-1 envelopes have been isolated after a long period of infection (usually several years) but not after a short time (Corti and Lanzavecchia, 2013; Doria-Rose et al., 2014; Liao et al., 2013). A large number of these HIV-1 bnAbs have more than 20% amino acid divergence from the

germline sequence. Specifically, the CD4 binding site targeting bnAb VRC01 has more than a 35% amino acid divergence (Wu et al., 2010). The long period of time required for bnAb generation and the high levels of somatic hypermutation indicate the importance of a high mutation load in the generation of bnAbs.

Affinity-matured B cells can differentiate into either memory B cells, or plasmablasts. Memory B cells are usually long lived, and upon subsequent cognate antigen stimulation, they can either differentiate into plasmablasts and rapidly start antibody secretion, or undergo further affinity maturation in germinal centers. However, plasmablasts are short lived, but a fraction of them can home to specific niches in the bone marrow, where they become long-lived plasma cells, and can persist for up to several decades (Amanna et al., 2007; Höfer et al., 2006; Slifka et al., 1998).

A detailed understanding of B cell development and adaptive immune responses constitutes the foundation of antibody repertoire characterization and antibody discovery.

CHARACTERIZATION OF ANTIBODY REPERTOIRES

As a result of somatic recombination and hypermutation, a diverse antibody repertoire is produced after antigenic challenge, thus making characterization of such a diverse repertoire a formidable task. Much effort has been made to advance techniques for antibody repertoire characterization, both at the cellular (antibody transcripts) level and at the serological (antibody proteins) level (Calis and Rosenberg, 2014; Corti et al., 2011; Huang et al., 2012, 2014; Lavinder et al., 2014; Liao et al., 2013; Scheid et al., 2009, 2011; Walker et al., 2009, 2011; Wine et al., 2013; Wrammert et al., 2008; Wu et al., 2010; Zhu et al., 2013a, 2013b).

Cellular antibody repertoire analysis

Single-cell RT-PCR

For the analysis of antibody transcripts, the single-cell reverse transcription polymerase chain reaction (RT-PCR) has been most widely used since its development (Wardemann et al., 2003). B cells from different sources (for example, from peripheral blood mononuclear cells (PBMCs) when analyzing human antibody repertoires, or from secondary lymphoid organs (lymph nodes and spleens) when analyzing mouse antibody repertoires) are first isolated using fluorescence activated cell sorting (FACS) and sorted into multi-well plates, with one cell in each well. Isolated single cells are lysed *in situ*, and RT-PCR is performed to amplify the V_H and V_L from single cells separately. The cognate V_H:V_L pairs from single B cells are sequenced, cloned, expressed and assayed.

Wardemann *et al*. first used single-cell RT-PCR to investigate the origin of auto-reactive B cells in human bone marrow and the mechanism for their elimination (Wardemann et al., 2003). Since cloning, expressing, and testing many native V_H:V_L pairs from a large number of B cells is a laborious and expensive task, different modifications have been made at the B cell isolation step to minimize the numbers of B cells that need to be analyzed.

As a result of the ease for B cell sorting using FACS, antigen-specific sorting technologies have been developed to isolate B cells of interest. Wu *et al*. demonstrated the isolation of B cells that target CD4 binding site in HIV-1 virus (Wu et al., 2010). They computationally designed an HIV-1 envelope probe that has all other epitopes substituted with those from SIV but an intact CD4 binding site, and a CD4 binding site knockout HIV-1 envelope probe, then fished out B cells that showed binding to the former but not the latter probe. From these B cells, they isolated the most potent HIV-1 bnAbs to date by singe-cell RT-PCR. Similar antigen-specific sorting (with single-cell

RT-PCR) approaches have been used to isolate antibodies targeting rotavirus (Di Niro et al., 2010), transglutaminase-2 (Di Niro et al., 2012), and influenza virus (Whittle et al., 2014).

Alternatively, Epstein-Barr Virus (EBV) mediated B cell immortalization has been used to interrogate a large number of B cells to identify those secreting antigen-specific antibodies (Traggiai et al., 2004). Compared to antigen-specific B cell sorting, B cell immortalization enables the functional characterization of antibodies (for example, neutralization) rather than just binding. Traggiai *et al*. first reported the improvement of EBV immortalization by Toll-like receptor 9 (TLR9) stimulation with CpG DNA or other polyclonal stimulants and the subsequent isolation of neutralizing antibodies against SARS coronavirus (Traggiai et al., 2004). Antibodies against influenza virus (Corti et al., 2011), HIV-1 (Corti et al., 2010), dengue virus (de Alwis et al., 2012; Dejnirattisai et al., 2010), malaria (Tan et al., 2016), Ebola virus (Corti et al., 2016), and Zika virus (Stettler et al., 2016) have all been isolated with similar approaches.

Although there has been much success with this technique, a major limitation of B cell immortalization is the restriction to memory B cells, due to the resistance of plasmablasts/plasma cells to EBV immortalization. The same limitation also restricts the use of antigen-specific B cell sorting, as plasmablasts/plasma cells have very low (none) surface BCR expression. The use of plasmablasts/plasma cells for antigen-specific antibody isolation was first demonstrated by Wrammert *et al*., who isolated influenza virus specific antibodies from day 7 plasma cells after booster vaccination (Wrammert et al., 2008). The success of isolating antigen-specific antibodies from plasmablasts/plasma cells heavily depends on the quality of immune response and the time plasmablasts/plasma cells are collected.

In general, single-cell RT-PCR is a very powerful technique, but it is severely limited by being low-throughput, with usually less than 100,000 cells processed per experiment, along with the heavy workload and cost with antibody cloning and expression. On the other hand, NGS provides a much higher throughput, which enables the processing of more than 1,000,000 cells in a single experiment. This allows much deeper analysis of antibody repertoires than with single-cell RT-PCR.

Next-generation Sequencing

Since its release in 2005, NGS has been used more and more for cellular antibody repertoire analysis, as a result of the rapidly increasing throughput and decreasing cost. The inherent complexity of antibody repertoires also makes NGS a superior tool for interrogation.

NGS for antibody repertoires usually uses antibody mRNA as templates. From isolated B cells, mRNA is extracted after cell lysis, and the V_H and V_L are amplified separately. This pool of V_H and V_L sequences encoded by different B cells can be sequenced on various platforms (the most popular is Illumina). Bioinformatic analysis is performed to obtain the antibody sequences in the repertoire (Fig. 1-4). Important features of the antibody repertoire, including unique clonotypes, CDR3 length distribution, and somatic hypermutation rates can also be obtained. NGS has been used for in-depth analysis of antibody repertoires after vaccination (Jiang et al., 2013) and natural infection (Doria-Rose et al., 2014; Liao et al., 2013), which provides important insights into how B cells respond to pathogens longitudinally (Doria-Rose et al., 2014; Liao et al., 2013). It has also been used to interrogate antibody repertoires in neonatals and the elderly (Rechavi et al., 2015), which helps to better understand the development of B cells. Moreover, the comparison of antibody repertoires in twins by NGS enables the

investigation of genetic and environmental effects on antibody repertoire development (Wang et al., 2015b).

The selection of promising high-affinity, antigen-specific antibodies from millions of sequences requires special consideration. The most common way is to select antibodies within the same lineages as known functional antibodies (for example, HIV-1 bnAbs) (Zhu et al., 2013a, 2013b). If CDRH3 sequences are available from proteomic analysis, they can be used to select antibodies with the same CDRH3 (Lavinder et al., 2014). However, of millions of sequences obtained from NGS, only a very minor portion of them, namely, tens or hundreds can be reconstructed, expressed, and characterized due to laborious and expensive gene synthesis and cloning. Likewise, the intrinsically higher error rate of NGS compared to that of Sanger sequencing used in single-cell RT-PCR makes antibody repertoire analysis nontrivial (Georgiou et al., 2014a). In order to distinguish between sequencing errors and true sequence variations, careful experimental design and bioinformatic analysis to minimize sequencing errors are critical. Several techniques are used to improve NGS accuracy, with barcoding being the most widely used (Georgiou et al., 2014b). For sequence analysis and annotation, IMGT (the International Immunogenetics Information System) is the most commonly used tool (Lefranc et al., 2009).

The major limitation of NGS is the bulky processing of B cells makes native V_H:V_L pairing lost. Furthermore, only up to hundreds of antibodies can be made with as a result of the cost for gene synthesis. In general, these repertoire analysis methods have been widely used and shown to identify potent antibodies against complex human pathogens, such as HIV-1, Ebola and influenza virus, malaria, and Zika virus (Corti et al., 2011, 2016; Huang et al., 2012, 2014; Stettler et al., 2016; Tan et al., 2016).

Serological antibody repertoire analysis

Antibodies in the serum are those that protect from pathogens, hence, the characterization of serological repertoires is of great importance. Recent advances in mass spectrometry have enabled detailed than ever, and accurate deconvolution of serological antibody repertoires (Cheung et al., 2012; Wine et al., 2013). Wine *et al*. and Cheung *et al*. first reported the characterization of rabbit/mouse antibody repertoires with liquid chromatography-tandem mass spectrometry after immunization (Cheung et al., 2012; Wine et al., 2013). Lavinder *et al*. and Sato *et al*. then reported the characterization of human antibody repertoires after tetanus toxoid and hepatitis B virus vaccination and human cytomegalovirus infection (Lavinder et al., 2014; Sato et al., 2012).

In these cases, antigen-specific antibodies in the serum are enriched by antigen-affinity capture and digested with pepsin to remove Fc fragments. The Fab fragments are then digested with trypsin to produce CDRH3 containing peptides (most informational peptides) (Fig. 1-5). These peptides are analyzed by mass spectrometry and mapped back to a database of antibody sequences (usually generated by NGS of antibodies isolated from B cells collected at the same time with the serum) to retrieve the full-length sequences. Promising antibodies are then selected, reconstructed, and assayed. Although proteomic analysis of serological antibody repertoires is very powerful, it has to rely on the full-length antibody sequence database generated by NGS to predict and infer the fragmented peptides.

ADVANCING HIGH-THROUGHPUT CHARACTERIZATION OF NATIVE ANTIBODY V_H:V_L PAIRS

As a result of the importance of antibody V_H:V_L chain pairing, and the major limitations of single-cell RT-PCR (low throughput) and NGS (loss of native V_H:V_L pairing), DeKosky *et al*. developed a high-throughput pairing and sequencing method for

native V_H:V_L pairs, that combines the advantages of single-cell RT-PCR (preservation of native V_H:V_L pairing) and next-generation sequencing (high-throughput) (DeKosky et al., 2013, 2015) (Fig. 1-6). By sequestering single B cells in micro emulsion droplets, lysing the cells and capturing mRNA using oligo(dT) beads, and linking and amplifying captured antibody coding sequences in single bead emulsion through RT-PCR, amplicons composed of native V_H: V_L pairs are obtained. Illumina paired end sequencing is used to determine antibody sequences in the amplicons.

Alternatively, Busse *et al*. used a two-dimensional, bar-coded primer matrix to integrate single-cell RT-PCR of V_H:V_L pairs with NGS (Busse et al., 2014). After next-generation sequencing, they were able to recover pairs from up to 50,000 cells. Although it can process a lot more B cells than single-cell RT-PCR, the throughput is still low compared to that of emulsion based pairing and sequencing.

Zhu *et al*. developed a 'pseudo' V_H:V_L pairing method which relies on phylogenetic reconstruction for antibody lineage characterization (Zhu et al., 2013a, 2013b). They first observed the matching patterns of V_H and V_L phylogenetic trees during the co-evolution of HIV-1 virus and bnAbs, then paired V_H and V_L from the matching branches of the phylogenetic trees based on this observation to obtain new bnAbs. This phylogenetics based V_H:V_L pairing is useful for the analysis of antibody repertoires with high somatic hypermutation levels. However, unlike the physical linkages between V_H and V_L generated in emulsions, and barcoded PCRs used to retain pairing information, there are no real physical linkages produced in phylogenetic pairing, so the V_H:V_L pairs obtained may not exist.

These high-throughput V_H:V_L pairing methods greatly expanded the number of native antibody pairs that can be obtained in one experiment, but the number of antibodies that can be functionally characterized is still very limited.

ANTIBODY DISCOVERY AND ENGINEERING

B cell immortalization

The most commonly used method for antibody discovery is B cell immortalization. Depending on origins of B cells, two different methods can be used. Köhler and Milstein first demonstrated B cell immortalization by fusing mouse spleen cells with myeloma cells (Kohler and Milstein, 1975). The fused hybridomas can secrete antibodies in *in vitro* culture. Hybridoma fusion has been widely used for antibody discovery. In many cases, after immunization of mice with the target antigen, isolation of antibody-producing cells, and fusion with myeloma cells, immortalized antibody-producing cells can be generated and screened to identify those that secrete antigen-specific antibodies. However, as a result of the low cell viability after fusion, a large number of fused cells have to be made to select antigen-specific antibodies. Moreover, as a result of the general difficulty of and regulations for human immunization and antibody-producing cell isolation, hybridomas are often produced using mouse cells.

The success for mouse antibody discovery using hybridomas inspired efforts to adapt the hybridoma approach for human antibody discovery. However, human B cells were found to be more resistant to myeloma fusion than mouse B cells (Traggiai et al., 2004). As a result, EBV mediated immortalization has been developed (Traggiai et al., 2004). As discussed for single-cell RT-PCR, EBV mediated human B cell immortalization has enabled the isolation of antibodies against many human pathogens (Corti et al., 2010, 2011, 2016; Stettler et al., 2016; Tan et al., 2016).

The development of humanized mouse strains provided easy-to-fuse mouse B cells that can produce human antibodies. A few previously developed humanized mouse strains had mouse Ig loci replaced with only part of the human Ig loci (Fishwild et al., 1996; Green et al., 1994; Lonberg et al., 1994; Taylor et al., 1992). Thanks to the

development of mouse genetics, transgenic mouse strains with the complete human Ig loci integrated into the genome have been reported (Lee et al., 2014; Macdonald et al., 2014; Murphy et al., 2014). In both reports, bacterial artificial chromosomes that have different parts of human IgH, Igκ, and Igλ loci respectively were used to target mouse embryonic stem cells iteratively to insert the complete human Ig loci in multiple steps. These modified embryonic stem cells were then injected into blastocysts to generate transgenic mouse.

The use of human antibodies identified from humanized mouse strains for therapeutics also minimizes the immunogenicity problem. As a result of sequence differences between mouse and human antibodies, mouse antibodies could elicit strong immune responses in humans (for example, human-anti-mouse antibodies). The immunogenicity greatly reduces the efficacy of mouse antibodies for therapeutics. On the other hand, human antibodies recovered from humanized mouse strains do not have this problem and can thus have maximal efficacy.

***In vitro* display technologies**

In addition to B cell immortalization, *in vitro* selection and screening of antibody libraries is also widely used for antibody discovery and engineering (Bradbury et al., 2011). Among the various techniques for antibody library selection, phage display (Winter et al., 1994), *E. coli* surface display (Mazor et al., 2007), and yeast surface display (Chao et al., 2006) are the three most widely used. All three techniques share the commonality that they link the antibody genotype to phenotype (binding) by expressing and displaying vector encoded antibody variants on the cell surface.

The antibody libraries used in these selection methods are usually combinatorially assembled from naïve, immunized, or synthetic V_H and V_L. These combinatorial libraries

provide rich pools of antibodies to select from to identify high-affinity, antigen-specific hits. The selected antibodies can be further optimized for functions.

With recent advance for protein sequence space sampling algorithms (Kaufmann et al., 2010), loop modeling algorithms (Mandell et al., 2009), and the improvement of energy functions (Kaufmann et al., 2010), computational protein design is used more and more to guide antibody library construction. A commonly used approach for computation-guided library design and selection is to first identify promising antibodies by computational design and assemble them into an antibody library, then select the best hits. The designed antibody usually needs further optimization, which includes constructing and selecting error-prone or site-directed libraries (Lapidoth et al., 2015).

In phage display, antibody fragments (usually scFv is used) are fused with phage coat proteins (for example, pIII of M13 phage), and displayed on the surface of phages. By panning these antibody-displaying phages toward target antigens, antigen-specific antibodies can be selected. Phage display has played an important role in mouse antibody humanization.

Similarly, in *E. coli* display, antibody fragments are covalently anchored on the inner membrane. After spheroplasting, these antibody fragments can be screened with desired antigens.

In yeast display, scFv variants are displayed on the cell surface as fusion proteins with Aga2, which is covalently linked to Aga1 (Boder and Wittrup, 1997; Chao et al., 2006). After staining with fluorescently labeled antigens, cells carrying variants that possess better binding properties will have higher fluorescence, and can then be sorted using FACS. The use of FACS permits high-throughput screening of very large antibody libraries.

Compared to phage and *E. coli* display, yeast surface display is a eukaryotic system that better mimics the native mammalian conditions under which the antibodies express, fold, and exert their functions. Moreover, by using 2-color FACS, it allows normalization of binding signal with antibody display levels (Boder and Wittrup, 1997). This prevents the recovery of antibodies that bind tighter as a result of higher expression level instead of improved affinity.

A major limitation of *in vitro* antibody library screening is the non-cognate (non-native) pairing of V_H and V_L in the combinatorial libraries may result in many problems for downstream development (Liu et al., 2014; Ponsel et al., 2011; Tiller et al., 2013; Wang et al., 2014; Wörn and Plückthun, 2001; Xu et al., 2013). Biophysical properties of these non-natively paired antibodies are sometimes mediocre (Liu et al., 2014; Ponsel et al., 2011; Tiller et al., 2013; Wang et al., 2014; Wörn and Plückthun, 2001; Xu et al., 2013), which requires extensive improvement for use as drugs. Moreover, an scFv cannot be used directly under most circumstances and would need to be reformatted to a Fab format for structural stability (Quintero-Hernández et al., 2007).

Another limitation of *in vitro* antibody library screening is it only selects for binding capacity. After high-affinity antibodies are obtained, functional assays such as neutralization still need to be performed to identify those that not only bind, but also exert desired functions. In contrast, in B cell immortalization, functional assays can be performed with secreted antibodies, which allows direct identification of antibodies that both bind and function as expected.

The recent report of function-based selection for combinatorial libraries using a mammalian autocrine system by Lerner and colleagues demonstrates the progress toward (Xie et al., 2013). The combinatorial antibody libraries were first polished by phage display to enrich target binders, then the enriched libraries were transfected into

mammalian cells, which expressed antibodies either in the cytoplasm or on the surface. By carefully designing and performing selection (for example, cell proliferation or cell death), antibodies with rare functions (for example, cell death prevention, stem cell transdifferentiation, tumor cell reprogramming) were selected. This is the first demonstration of functional selection of antibody libraries rather than just binding.

SIGNIFICANCE OF HIGH-THROUGHPUT ANTIBODY REPERTOIRE ANALYSIS AND ANTIBODY DISCOVERY

As a result of their unique features, including high specificity, affinity, stability, and their ability to link adaptive immunity to innate immunity, antibodies have been used extensively in clinics. For example, monoclonal antibodies (mAbs) based treatment has been successfully developed for cancers (including leukemia, lymphoma, and breast cancer), graft-versus-host diseases, autoimmune diseases (including systemic lupus erythematosus and rheumatoid arthritis), and infectious diseases (RSV and Ebola) (Chan and Carter, 2010; Topalian et al., 2016). The year-by-year increasing sale of antibody drugs highlights the need of more antibodies for disease treatment and novel techniques for rapid high-throughput identification of promising antibodies.

Antibodies are also widely used in research. Western blots, flow cytometry, and chromatin immunoprecipitation are just three of many techniques where antibodies are essential. The variable performance of different antibodies and lack of antibodies for some targets highlight the need for facile discovery of more antibodies for research use.

Both traditional antibody discovery and engineering methods and antibody repertoire characterization techniques are widely used for novel antibody identification. Trastuzumab (HER2 targeting), Rituximab (CD20 targeting), Cetuximab (EGFR targeting), Ipilimumab (CTLA-4 targeting) and Nivilomab (PD-1 targeting) are a few examples of FDA approved antibodies identified by traditional antibody discovery and

engineering, while VRC01 and 3BNC117 (HIV-1 CD4 binding site targeting) are two examples of repertoire analysis identified antibodies that are in clinical trials (Scheid et al., 2011; Wu et al., 2010).

Furthermore, antibody repertoire analysis can help vaccine design to elicit more robust immune responses. By analyzing a large number of antibodies generated after infection or vaccination, the dominant and protective (sub)-dominant epitopes that are recognized by these antibodies can be identified. Then stable, immunogenic antigens can be rationally designed and engineered to display protective (sub)-dominant epitopes only on the protein surface. For some chronic, highly mutating virus infections, the longitudinal analysis of antibody repertoires together with analysis of the co-evolving virus at different times after infection can provide clues about how a virus can evade antibody recognition, and how antibodies go through germinal center reactions to recognize constantly mutating viral variants (Doria-Rose et al., 2014; Liao et al., 2013). With this longitudinal information, Jardine *et al*. developed stepwise vaccines to mimic the mutating HIV-1 virus for the priming naïve B cells that express predicted germline BCRs and the boosting of these selected B cells for affinity maturation to recognize the native virus (Briney et al., 2016; Escolano et al., 2016).

Although many technologies have been developed for antibody discovery and repertoire analysis, and a lot of success has been achieved, a high-throughput technology that enables native antibody repertoire analysis and antibody discovery at single cell level has yet to be fully realized. Moreover, the development of protein therapeutics other than antibodies has yet to be shown.

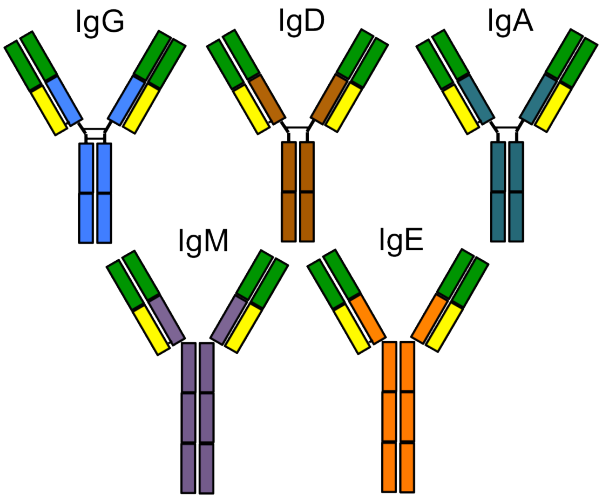

Figure 1-1: Schematic diagrams of IgM, IgD, IgG, IgA, and IgE antibodies.

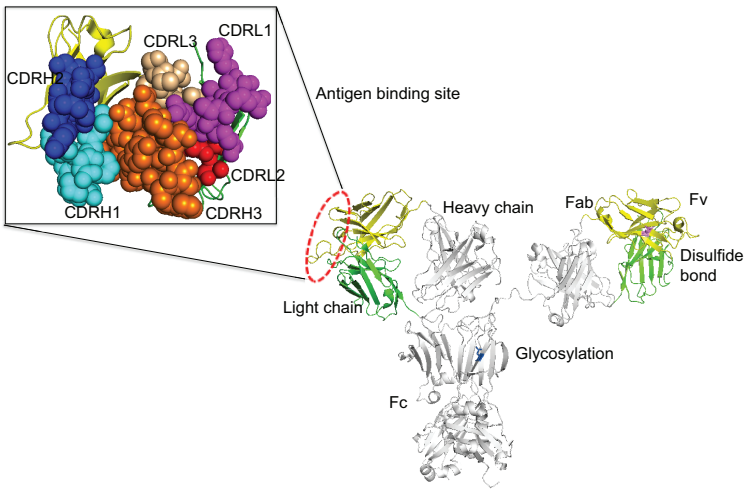

Figure 1-2: Structure of a typical human IgG showing different features in the molecule (PDB code 1HZH).

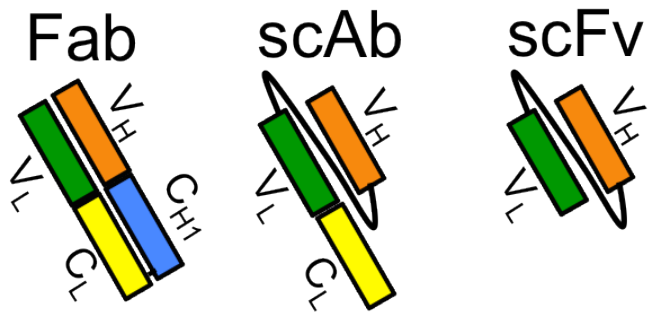

Figure 1-3: Schematic diagrams of Fab, scAb, and scFv.

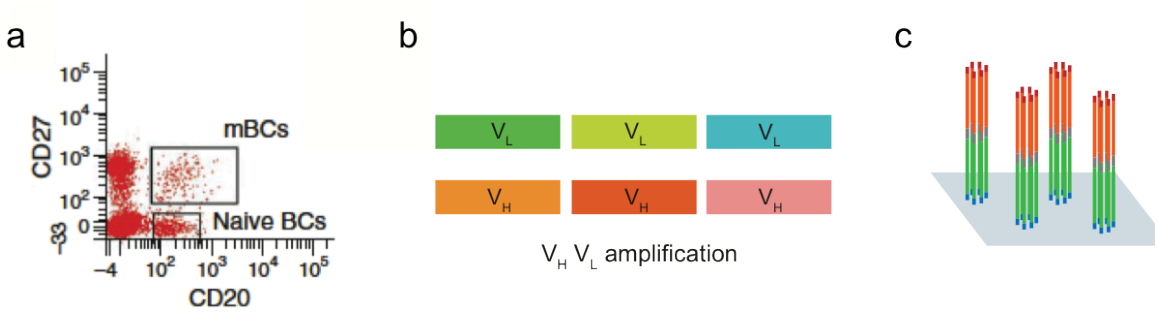

Figure 1-4: NGS workflow for antibody repertoire analysis.

(a) Isolation of B cell populations of interest. (b) Cell lysis and bulk amplification of V_H and V_L transcripts. (c) Next-generation sequencing of the V_H and V_L fragments.

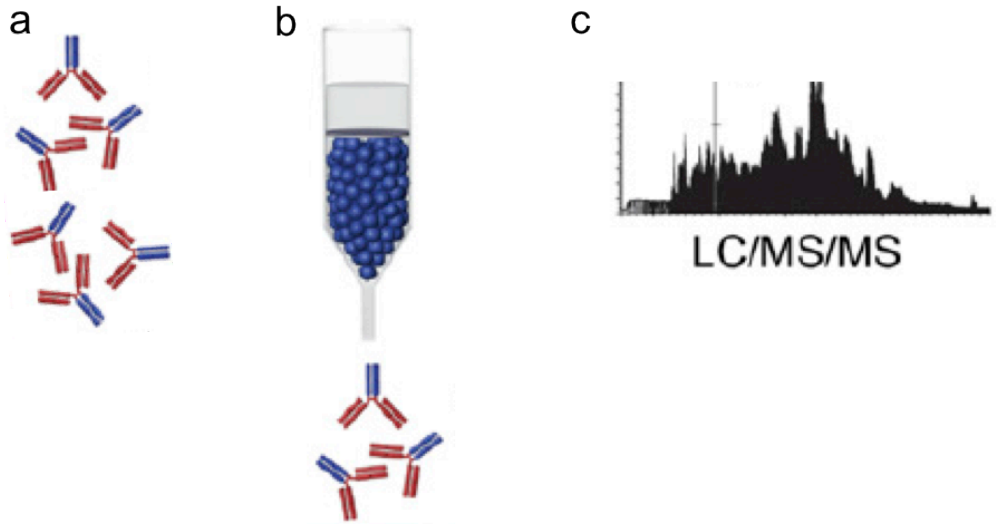

Figure 1-5: Proteomic workflow for serum antibody repertoire analysis.

(a) Serum antibodies are pooled by protein A purification. (b) Antigen-specific antibodies are purified by antigen-affinity chromatography. (c) After protease digestion, peptides are resolved by high-resolution LC-MS/MS.

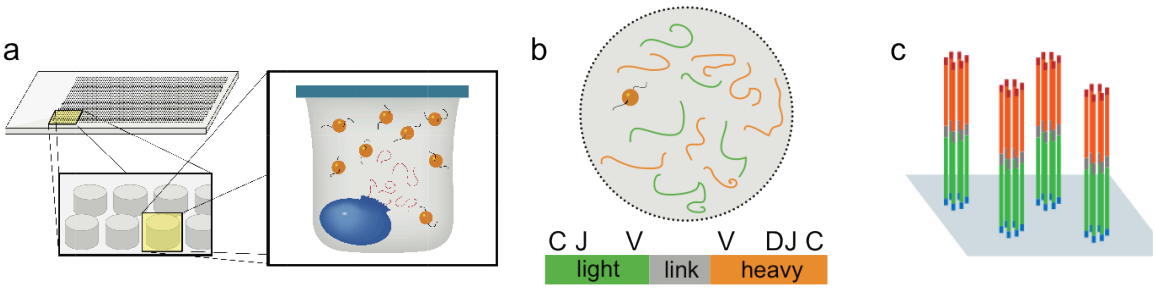

Figure 1-6: High-throughput V_H:V_L pairing workflow.

(a) Isolated B cells are deposited into 125 pL wells on PDMS slides that also contain poly(dT) beads. Cells are lysed *in situ* and mRNA is captured on the poly(dT) beads. (b) The poly(dT) beads are emulsified and V_H:V_L amplicons are generated following reverse transcription and overlap extension PCR. (c) V_H:V_L amplicons are sequenced using Illumina 2x250 MiSeq and the V_H:V_L pairs are identified via bioinformatics analysis.

Chapter 2: Facile Discovery of a Diverse Panel of Anti-Ebola Virus Antibodies by Immune Repertoire Mining[1]

The ongoing evolution of Ebolaviruses poses significant challenges to the development of immunodiagnostics for detecting emergent viral variants. There is a critical need for the discovery of monoclonal antibodies with distinct affinities and specificities for different Ebolaviruses. We developed an efficient technology for the rapid discovery of a plethora of antigen-specific monoclonal antibodies from immunized animals by mining the V_H:V_L paired antibody repertoire encoded by highly expanded B cells in the draining popliteal lymph node (PLN). This approach requires neither screening nor selection for antigen-binding. Specifically we show that mouse immunization with Ebola VLPs gives rise to a highly polarized antibody repertoire in $CD138^+$ antibody-secreting cells within the PLN. All highly expanded antibody clones (7/7 distinct clones/animal) were expressed recombinantly, and shown to recognize the VLPs used for immunization. Using this approach we obtained diverse panels of antibodies including: (i) antibodies with high affinity towards GP; (ii) antibodies which bound Ebola VLP Kissidougou-C15, the strain circulating in the recent West African outbreak; (iii) non-GP binding antibodies that recognize wild type Sudan or Bundibugyo viruses that have 39% and 37% sequence divergence from Ebola virus, respectively and (iv) antibodies to the Reston virus GP for which no antibodies have been reported.

[1]This chapter is reproduced with minor modifications from its initial publication:
Wang, B., Kluwe, C.A., Lungu, O.I., DeKosky, B.J., Kerr, S.A., Johnson, E.L., Jung, J., Rezigh, A.B., Carroll, S.M., Reyes, A.N., Bentz, J.R., Villanueva, I., Altman, A.L., Davey, R.A., Ellington, A.D., Georgiou, G. (2015). Facile Discovery of a Diverse Panel of Anti-Ebola Virus Antibodies by Immune Repertoire Mining. Sci. Rep. *5*, 13926. Wang, B. designed and performed the experiments, and analyzed the data.

INTRODUCTION

Ebolaviruses are negative-sense RNA filamentous viruses that cause very high morbidity and mortality (Bowen et al., 1977). Host cell entry is mediated first by the attachment of the heavily glycosylated glycoprotein (GP) on the viral envelope to the host cell encoded T-cell immunoglobulin and mucin domain 1 (TIM-1) (Kondratowicz et al., 2011). Following cathepsin cleavage in the lysosome, GP mediates cellular entry by binding the host cell encoded Niemann-Pick C1 (NPC1) (Carette et al., 2011). Five antigenically distinct ebolaviruses exhibiting 35-45% genome sequence divergence have been discovered (Towner et al., 2008): Ebola virus (abbreviated as EBOV, formerly designated as Zaire ebolavirus); Sudan virus (SUDV); Bundibugyo virus (BDBV); Reston virus (RESTV, for which no zoonotic infections have been reported to date) (Miranda and Miranda, 2011); and Taï Forest virus (TAFV, one incident of human infection) (Le Guenno et al., 1995). The recent EBOV outbreak in West Africa, centered in Guinea, Sierra Leone, and Liberia with isolated outbreaks in Nigeria and Mali, was the largest ever with a mortality rate estimated at 70% of recorded definitive clinical outcomes (http://www.cdc.gov/vhf/ebola/outbreaks/2014-west-africa/index.html) (Team, 2014). Phylogenetic comparison of isolates from the recent outbreak (Baize et al., 2014) with 20 Ebolavirus genomes from earlier outbreaks suggested that the 2014 West African virus likely spread from central Africa within the past decade, having diverged from a common ancestor around 2004 (Gire et al., 2014). The five Ebolavirus species have varying rates of molecular evolution, with the highest of 8.21×10^{-4} nucleotide substitutions/site/year for Reston virus (Carroll et al., 2013). The ongoing evolution of Ebolaviruses poses significant challenges to the development of immunodiagnostics. Specifically, there is a critical need for the discovery of panels of monoclonal antibodies with distinct affinities and specificities for different Ebolaviruses.

Antibodies to EBOV and SUDV have been produced from hybridomas (Dias et al., 2011; Qiu et al., 2011); by *in vitro* screening of synthetic scFv libraries (Chen et al., 2014; Koellhoffer et al., 2012), and from human immune antibody libraries constructed from infected individuals (Maruyama et al., 1999). However additional monoclonal antibodies to Ebolaviruses are urgently needed both for diagnostic purposes and as therapeutics (Parren et al., 2002; Qiu et al., 2014). Specifically, the generation of diagnostic antibodies to Ebolaviruses is complicated by the structural complexity of the GP, which is heavily glycosylated in a host cell-specific manner (Heinz Feldmann, Stuart T. Nichol, Hans-Dieter Klenk, Clarence J. Peters, 1994; Lin et al., 2003) and subjected to proteolytic cleavage during entry (Hood et al., 2010), as well as by the sequence diversity of the Ebolaviruses. Finally, characterization of useful antibodies to Ebolaviruses is limited by the safety concerns associated with handling the live virus.

Antibody discovery has relied either on the immortalization (Kohler and Milstein, 1975) or cloning of antibodies isolated from individual B cells obtained from an antigen-challenged host (Corti et al., 2011; Ghosh and Campbell, 1986; Lane and Koprowski, 1982; Reddy et al., 2010; Saggy et al., 2012; Wu et al., 2010) or, alternatively, on the *in vitro* isolation from combinatorial libraries using a variety of screening techniques (Bradbury et al., 2011). The current collection of antibody technologies is predicated on the isolation of clones that display high antigen binding. However, animal immunization induces the stimulation and expansion of a highly diverse population of B cells encoding an antibody repertoire with a wide range of antibody affinities (Reddy et al., 2010; Saggy et al., 2012). Antibodies with low affinity nonetheless may exhibit other highly desirable properties, including broad cross-reactivity or heteroclitic specificity, i.e. stronger binding reaction to a different antigen other than the one used for immunization (Ghosh and Campbell, 1986; Lane and Koprowski, 1982; Van Regenmortel, 2014).

Unfortunately, there is no straightforward way to identify such interesting antibodies. For example, while the isolation of antibodies that bind to multiple antigens (e.g. to different flu hemagglutinins) or that neutralize rapidly evolving pathogens such as HIV-1 or flu has been accomplished by B cell cloning, the process typically requires the screening of many thousands of B cells and therefore is very laborious and expensive (Corti et al., 2011; Walker et al., 2009, 2011; Wu et al., 2010).

In order to satisfy the need for a wider variety of antibodies to Ebolaviruses, we developed a novel approach to comprehensively mine the full suite of antibody diversity, shaped by *in vivo* selective mechanisms and generated within the boundary of reactive secondary lymphoid tissues in immunized animals. We reasoned that antibodies encoded by antigen-stimulated B cells that had undergone the greatest degree of expansion within the confinement of a secondary lymphoid organ are most likely to display desirable antigen recognition properties including heteroclite recognition of diverse Ebolaviruses. Briefly, mice were first immunized in the footpad with Ebola virus-like particles (VLPs). Footpad immunization triggers a strong and highly-focused immune reaction in the PLN, especially for particles <40 nm such as VLPs[32]. Antigen experienced, $CD138^+$ B cells (plasmablasts) from the PLN were isolated and the natively paired V_H:V_L repertoire encoded by these cells was determined by NextGen sequencing[33]. Antibodies corresponding to the highest frequency V_H:V_L pairs, and thus likely arising from the most clonally expanded and highly-transcribing $CD138^+$ B cells within the PLN, were expressed recombinantly and their binding properties were characterized in detail. In each of two mice tested, 7/7 antibodies encoded by the highest frequency antigen-draining PLN plasmablast sequences recognized the antigen (Ebola VLPs) with several binding to recombinant GPs with up to nM affinities. Interestingly, even though animals had been immunized with EBOV VLPs, mining of the expanded native B cell repertoire within the

PLN resulted in antibodies that also recognized SUDV and BDBV. In a separate experiment we also employed this technique to generate the first RESTV GP-specific antibodies. Thus, the antibodies reported here collectively constitute a panel of reagents for the detection of most Ebolaviruses.

RESULTS

Immunization of the PLN yields antigen-specific antibodies

Fig. 2-1 summarizes our approach for mining the antibody repertoire encoded by the most highly expanded, antigen-experienced B cells following antigen stimulation. Footpad immunization leads to a strong inflammatory response in the draining popliteal lymph node (Fig. 2-1a). Unlike lymph nodes that drain sites of frequent extracorporeal interaction (e.g. the oral cavity or lungs), the germinal centers of the popliteal lymph node are normally relatively unstimulated (Kamala, 2007). Footpad immunization results in a marked increase in cellularity in the ipsilateral popliteal lymph node relative to the unstimulated contralateral lymph node (Gleichmann, 1981; Ravel and Descotes, 2005), and a large fraction of the constituent antibody-secreting B cells were expected to be antigen-specific.

Ebola VLPs were produced by co-transfection of HEK293FT cells with plasmids encoding the three major virus structural proteins: nucleoprotein (NP), VP40 and GP of the EBOV Mayinga strain and were purified by sucrose density centrifugation. Electron microscopy and gel staining confirmed that the VLPs displayed a morphology and consistency characteristic of EBOV virions. Three mice were immunized with Ebola VLPs in emulsified adjuvant in the left hind footpad, followed by boost immunization in the lateral hock to minimize pain and discomfort. An anti-VLP titer of $>1{:}10^4$ was observed after the second boost (Fig. 2-2a) in all three mice, and two mice (denoted as

ZM1 and ZM2) were chosen for further analysis. A final boost was administered and 6 days later the popliteal lymph nodes were extracted. As expected, the ispilateral PLN was observed to be hypertrophic (Fig. 2-2b).

$CD45R^{-}CD19^{-}CD138^{+}$ antibody secreting B cells were enriched by magnetic sorting and the paired V_H:V_L repertoire from single cells was determined following sequestration of the cells into 125 pL wells on PDMS slides (DeKosky et al., 2013) (Fig. 2-1b**).** Approximately $1x10^5$ $CD45R^{-}CD19^{-}CD138^{+}$ plasmablasts were isolated by magnetic sorting, of which $2.5\text{-}3.5x10^4$ cells were processed to create natively paired V_H:V_L amplicons. Linked V_H:V_L amplicons of approximately 850 bp were generated and then sequenced using Illumina MiSeq technology (Fig. 2-1c). High quality reads were clustered based on the CDRH3:CDRL3 sequences (Fig. 2-1d).

The PLN plasmablast IgG repertoire elicited by immunization with EBOV

283 and 333 unique V_H:V_L pairs (represented by ≥2 sequence reads per pair each), comprising the PLN plasmablast repertoires were identified in mouse ZM1 and mouse ZM2, respectively. The repertoires from both Ebola VLPs immunized mice investigated were heavily skewed, with the top ten most abundant V_H:V_L pairs representing 53.9% and 48.4% of the total sequence counts in each mouse, respectively (Fig. 2-3a). We observed a strong bias in germline V-gene usage in immunized mice (Fig. 2-3b,c). Biased usage of IGHV subgroups 1–3 and 5 (Kaushik and Lim, 1996; Lu et al., 2014), as well as IGKV1, 3, 4 and 6 have been previously observed in the repertoires of unimmunized mice (Lu et al., 2014). Likewise, in Ebola VLPs immunized mice, IGHV1, IGHV5, as well as IGKV1, IGKV3, IGKV4, and IGKV6 heavy and light V genes were most strongly represented. Importantly, however, the IGHV8, IGHV14, and IGKV5 families were also strongly overrepresented in both mice immunized with EBOV VLPs. The enrichment of

IGHV8 in the plasmablast repertoire from PLN of Ebola VLPs immunized animals is particularly noteworthy as this germline family is expressed at a very low level in mice (Kaushik and Lim, 1996; Lu et al., 2014) and additionally it was shown to be utilized at very low frequencies in $CD138^+$ lymphocytes from animals immunized with various other antigens (Reddy et al., 2010).

The repertoire of Ebola VLPs immunized mice displayed an average CDRH3 length distribution similar to that observed in $CD138^+$ repertoires previously reported (Reddy et al., 2010), although there was a slight skewing toward shorter CDRH3 lengths. Shorter CDRH3s are commonly found among antibodies that bind carbohydrates and thus the skewing observed here likewise may have reflected the elicitation of antibodies to the glycan component of the heavily glycosylated GP (Schoonbroodt et al., 2008). The plasmablast repertoire from EBOV VLPs immunized mice displayed a higher level of somatic hypermutation level in the framework 3 (FR3) heavy chain region compared to plasmablast repertoires reported for mice hyperimmunized with various other protein antigens (Reddy et al., 2010).

Construction and characterization of anti-EBOV VLP antibodies

The CDRH3 and CDRL3 amino acid sequences of the 7 highest frequency V_H:V_L clonotypes from each mouse, together with the respective V(D)J gene segments are listed in Table 2-1. Antibody genes for these 14 most prevalent V_H:V_L antibody clonotypes were synthesized and cloned as mouse V region-human constant domain chimeric antibodies, expressed in HEK293 cells, and characterized for antigen binding (Fig. 2-1e,f). KZ52, a very well characterized Ebola virus recognizing antibody isolated from a survivor of the 1995 Kikwit outbreak (Lee et al., 2008; Maruyama et al., 1999), was expressed as a positive control. All 14 antibodies gave an ELISA signal above

background on plates coated with the immunizing antigen, i.e. EBOV (Mayinga strain) VLPs (Fig. 2-4a). ELISA titer analysis revealed that the majority of the antibodies bound VLPs at a titer >1:1,000. Interestingly, the ELISA signals did not correlate with the frequency of the respective antibodies in the PLN $CD138^+$ repertoire.

5/14 antibodies recognized the recombinant uncleaved form of the EBOV Mayinga strain GP (Fig. 2-4b). One antibody bound to GP with a single-digit nM equilibrium dissociation constant (ZM1.3, K_D=7.7 nM), as determined by SPR analysis (Table 2-2, Fig. 2-5), while two others (ZM1.1 and ZM2.1) exhibited K_D values in the 10 nM range. Finally, 2/5 GP specific IgGs, ZM1.2 and ZM1.6, exhibited lower affinities (K_D =156 and 635 nM, respectively (Table 2-2, Fig. 2-5)), consistent with the lower titer of those antibodies for Ebola VLPs.

The three highest affinity, EBOV GP-specific antibodies (ZM1.1, ZM1.3 and ZM2.1) were tested for binding to VLPs encoding GPs from Ebola strains isolated from two different outbreaks: 034-KS (Democratic Republic of Congo, 2008, NCBI Accession number HQ613402) and Kissidougou-C15 (Kissidougou, Guinea, 2014, NCBI Accession number KJ660346). Both ZM1.3 and ZM2.1 showed higher binding to Ebola 034-KS and Kissidougou-C15 VLPs than the well-studied KZ52 antibody (K_D for purified Mayinga GP=1.55 nM). We observed that KZ52 failed to recognize VLPs containing a Mayinga GP N550K variant in which Arg 550 residue was replaced with Lys, a mutation observed in Marburg GP (Manicassamy et al., 2007), while ZM1.1, ZM1.3 and ZM2.1 were still able to bind the Mayinga GP N550K variant. This finding indicates that ZM1.1, ZM1.3 and ZM2.1 likely recognize a different epitope than KZ52. Interestingly, ZM1.3 displayed heteroclitic specificity in that it bound better to EBOV 034-KS and Kissidougou-C15 VLPs than to the immunization Mayinga strain VLPs (EC_{50} of 0.334 nM and 0.39 nM, respectively compared to 1.7 nM for Mayinga VLPs). These antibodies

also bound to live wild type Ebola virus, indicating that they should be useful for developing diagnostic assays against primary biological samples.

Finally, encouraged by the results detailed above, we used the strategy described in Fig. 2-1 to develop antibodies that recognized Reston Ebola virus (RESTV, Reston, 1996, NCBI Accession number AB050936), an Ebolavirus for which no anti-GP monoclonal antibodies are available. RESTV VLPs were generated, mice were immunized as above, and the repertoire encoded by PLN plasmablasts was determined. We identified three antibodies designated RM2.4, RM3.2, and RM3.3 that bound to both RESTV VLPs as well as to RESTV recombinant GP (Fig. 2-6).

Diagnostic utility of non-GP binding antibodies

While 9/14 anti-EBOV VLP antibodies did not show binding to the GP and thus presumably recognized other VLP proteins (NP, VP40) they nonetheless are of diagnostic utility. Specifically, in addition to binding to EBOV Mayinga VLPs, antibodies ZM1.4, ZM1.7, ZM2.2, ZM2.3, ZM2.5, and ZM2.6 displayed measurable binding to Bundibugyo Virus (BDBV, Bundibugyo, 2007, NCBI Accession number FJ217161), while ZM1.7, ZM2.2, ZM2.5, and ZM2.6 also showed binding activity to Sudan Virus (SUDV, Gulu, 2000, NCBI Accession number AY729654) (Representative ELISA data is shown in Fig. 2-7a-c). Thus, antibodies elicited by immunization with EBOV VLPs bound differentially to phylogenetically diverse variants.

DISCUSSION

Here we report a facile and rapid approach for generating large panels of distinct monoclonal antibodies that, unlike existing antibody discovery platforms, does not rely on screening for antigen binding; and further describe application of this approach by developing panels of diagnostic antibodies for Ebola virus strains. We found that in

multiple animals immunized with EBOV or RESTV VLPs, the repertoire of $CD138^+$ plasmablasts in the draining PLN was dominated by highly expanded, antigen-specific, antibody sequences. We find that these highly expanded sequences include antibodies that display high affinity to GP, bind to live, wild-type virus and, somewhat surprisingly, display diverse specificities to VLPs from different Ebolaviruses. Among the antibodies isolated, six recognized live Bundibugyo or both Bundibugyo and Sudan viruses in addition to EBOV. These results demonstrate the power of mining the antibody repertoire of highly expanded B cells after immunization for discovery of antibodies with interesting properties.

The ability to isolate 7 (and possibly more) distinct antibodies with very different CDR3 sequences per animal as well as a much larger number of somatic variants whose sequences are also available in the V_H:V_L sequence database provides a rich source of antibodies for practical purposes. Undoubtedly, many PLN B cells encoding antigen-specific antibodies are likely to have been subject to more limited expansion and thus are present at a lower abundance within the repertoire. However, such medium or low abundance antibody sequences within the repertoire are present at comparable levels to those elicited by environmental stimuli and thus recognizing unrelated antigens. Therefore the low abundance antigen-specific antibody sequences in the repertoire cannot be identified directly without significant additional effort.

The method we have pioneered should prove particularly useful for the facile development of diagnostic (and possibly therapeutic) antibodies. Starting from antibody-secreting B cells, it took us only 3 weeks to produce and characterize a diverse set of antigen-specific antibodies with a variety of useful specificities. This method should be particularly valuable for assessing and combating fast spreading pandemics.

In case of rapidly spreading emerging diseases such as Ebola, rapid and robust diagnostics are critical for treatment and disease control. For Ebola in particular, validated PCR based assays suitable for field work in third world countries are not available. Highly sensitive antibody based immunodiagnostics are extremely important and easy to implement (Towner et al., 2004). However, a dearth of monoclonal antibodies for emergent Ebolaviruses limits the ability to use antigen-capture technologies for viral identification of early-stage infections. The panel of antibodies identified and characterized here should be useful for diagnostic applications including discriminating SUDV, BDBV, and EBOV. By using multiplex immunoassay platforms such as the Luminex MagPlex® technology, it should be possible to multiplex up to 50 different antibodies with varying specificities for a broad range of epitopes, thus ensuring wide coverage of Ebolaviruses variants. Studies to incorporate the antibodies we have described here onto the Luminex MagPlex® diagnostic platform for field applications are on-going.

METHODS

VLP Production and Characterization

VLPs were produced by co-transfection of HEK293FT (Invitrogen) cells with plasmids encoding the three major virus structural proteins, NP, VP40 and GP. All open reading frames for virus structural proteins were obtained from NCBI and were codon optimized for mammalian cell expression using Gene Designer (DNA 2.0), synthesized (Epoch Life Science) and inserted into either pcDNA3 (for Ebola GP) or pCAGGS (all other genes) mammalian expression plasmids. Sequences of the structural protein ORFs were verified and are available upon request to R. A. D. Cells were transfected with each of 5 mg NP, 5 mg VP40 and 1 mg GP encoding plasmids by the calcium chloride/BES

transfection method. After 24 h cells were washed with DMEM, which was replaced with DMEM containing 10% (v/v) FBS. After an additional 24 h, the culture supernatant was collected and clarified by centrifugation at 3750 rpm for 30 min at 4ºC to remove cell debris. The supernatants were then overlaid onto a 5 mL 20% (w/v) sucrose cushion in 20 mM NaCl, 20 mM HEPES, pH 7.4 in an SW28 ultracentrifuge tube (Beckman). The VLPs were then pelleted by centrifugation at 28,000 rpm for 2 h at 4ºC in an SW28 rotor. To further purify VLPs, the pellet was resuspended in PBS and overlaid onto a 20 to 60% sucrose step gradient (5% increments) in 20 mM NaCl, 20 mM HEPES, pH 7.4 in a SW55 rotor tube (Beckman). The gradient was centrifuged at 38,000 rpm for 2 h at 4ºC after which an opaque band corresponding to VLPs was visible. Fractions were collected corresponding to the band as well as directly above and below it and analyzed by SDS-PAGE, staining for total protein with Krypton stain (Thermo Scientific) as well as immunoblotting. The middle fraction containing the peak of the VLPs was stored at -80°C until required. For immunoblotting, proteins were transferred to nitrocellulose membranes and stained using broadly reactive polyclonal antibodies against GP (gift from Dr. Andrew Hayhurst, Texas Biomedical Research Inst.), VP40 (gift from Dr. Ricardo Carrion, Texas Biomedical Research Inst.) or NP (IBT Bioservices). A specific monoclonal antibody against Zaire Ebolavirus GP was also used (4F3, IBT Bioservices). Appropriate secondary antibodies were purchased from LiCor Biosciences. Blots were imaged using a LiCor Odyssey SA imager. Electron microscopy of VLPs was performed at the University of Texas Health Sciences Center, San Antionio, Department of Pathology electron microscopy facility. VLPs were adhered to copper grids. The samples were fixed in glutaraldehyde and osmium tetroxide was used as contrast agent. Images were captured on a Philips 208S digital imaging electron transmission microscope.

Immunizations and Serum Titer Determination

The study was approved by the University of Texas Institutional Animal Care and Use Committee (AUP-2013-00015). All animal experiments were carried out in accordance with the approved protocol. VLPs in PBS pH 7.4 were emulsified in a 1:1 ratio with TiterMax Gold adjuvant (Sigma). For footpad immunizations, 20 μL (containing a total of 5 mg VLPs) antigen/adjuvant mixture was injected into the subcutaneous space of three BALB/c mice; for lateral hock injections, up to 50 μL was injected into the subcutaneous space just proximal to the lateral aspect of the ankle. Mice were immunized at footpad on day 0 for primary immunizations, and in the lateral hock on days 21, 35, and 77 for secondary immunizations. At days 10, 28, and 42 mice were bled for titration of the antigen-specific response. In order to determine serum antibody titer, mice were restrained in a tube restrainer, the tail wiped with isopropyl alcohol, and small incisions made with a fresh scalpel blade to nick the tail vein. 20-50 μL blood was obtained and allowed to coagulate at room temperature (RT) for 30 min, followed by centrifugation at 13,000 g for 15 min to pellet the clot. The serum was then used for ELISA assays. High binding ELISA plates (Corning) were coated overnight (O/N) at 4°C with 50 μL of 4 μg/mL Zaire Ebolavirus VLPs in PBS pH 7.4. Antigen solution was decanted and plates were then blocked at RT for 2 h in 2% milk (w/v) in PBS. Blocking solution was then decanted and plates were then incubated with 50 μL of serum diluted three-fold from 1:100 to 1:218,700 in 2% milk (w/v) in PBS for 1 h. Plates were then aspirated and washed 3 times with PBS containing 0.05% tween-20 (PBST), then incubated with 50 μL of 1:5000 diluted goat anti-mouse HRP secondary antibody (Jackson ImmunoResearch) for 1 h. Plates were washed 3 times with PBST, and incubated with 50 μL TMB-Ultra (Thermo Scientific) for 15 min. The reaction was

quenched with 50 μL 2 M H_2SO_4 and absorbance was read at 450 nm on a Tecan M200 plate reader.

Tissue Collection, Cell Isolation, and Subtype Purification

After determination of significant titer for Ebolavirus VLPs (signal evident above background at dilution > 1:10,000), mice were administered a final boost at day 77, and lymph nodes were collected 6 days later. For lymph nodes collection, mice were injected with 5-10 μL of 2% Evans Blue (Sigma) in PBS into the footpad. 30 min post-injection, mice were sacrificed by carbon dioxide asphyxiation followed by cervical dislocation. The skin and fur around the leg was removed to reveal the blue-stained popliteal lymph node (Fig. 2-1a), located just behind the knee. The lymph node was isolated and and stored in PBS pH 7.4 supplemented with 0.1% (w/v) BSA, 2 mM EDTA in a 6-well plate (Corning). Lymph node was homogenized by mechanical disruption using two 18G needles. The cells were then passed through a 70 μm cell strainer (Corning), with additional disruption using the plunger from a 3 mL syringe to aid passage of single cells. Cells were then spun down at 500 g for 10 min in a swinging bucket rotor. The cell pellets were then resuspended in 2 mL red blood cell lysis buffer (155 mM NH_4Cl, 12 mM $NaHCO_3$, 0.1 mM EDTA) and incubated at room temperature for 3.5 min. The lysis reaction was quenched by adding 20 mL PBS buffer followed by centrifugation at 500 g for 10 min at RT. Cells were washed again with 5 mL PBS buffer and resuspended in a final volume of 1 mL buffer. Plasma cells were then isolated using the Miltenyi Plasma Cell Isolation kit (Miltenyi Biotec). Briefly, non-plasma cells were depleted by magnetic labeling of CD49b and CD45R followed by enrichment of magnetically-labeled $CD138^+$ cells. CD45R is a pan-B cell marker expressed on naïve and activated B lymphocytes, but not on antibody-secreting cells. Conversely, CD138 is expressed on pre-B and

immature B-lymphocytes in the bone marrow, lost upon emigration into secondary lymphoid tissues, and re-expressed upon differentiation into plasma cells.

Single Cell V_H:V_L Sequencing

Sorted cells were analyzed by single B cell V_H:V_L sequencing as previously described[33]. Briefly, single cells were isolated into 125 pL wells printed in PDMS along with poly(dT) conjugated magnetic beads. Cell lysis and capture of mRNA was performed *in situ*, and beads were collected and emulsified to serve as template for emulsion overlap extension RT-PCR. A follow-up nested PCR resulted in 850 bp amplicons containing linked genetic information for V_H and V_L genes. 850 bp amplicons were analyzed using the Illumina MiSeq 2x250 platform. V_H and V_L genes were amplified separately for full-length V_H and V_L analysis using the Illumina MiSeq platform as previously described (DeKosky et al., 2013).

Sequence Analysis

Raw MiSeq data was analyzed as previously described (DeKosky et al., 2013). Briefly, raw data were filtered for a minimum Phred quality score of 20 over 50% of nucleotides to ensure high read quality in the CDR3 regions of heavy and light genes. Sequence data were submitted to the IMGT information system for V-D-J germline gene mapping. Sequences were filtered for in-frame V-D-J junctions and V_H:V_L pairs were compiled by exact CDRH3:CDRL3 nucleotide match. CDRH3 junction nucleotide sequences were clustered to 96% identity and resulting clusters with >=2 V_H:V_L reads were ranked by MiSeq read counts (DeKosky et al., 2013). Due to read length limitations of current next-generation sequencing technology, the complete V_H and V_L genes were also sequenced and analyzed separately. Full-length V_H and V_L genes were filtered for a minimum Phred quality score of 20 over 50% of nucleotides and were compiled by

CDRH3 and CDRL3 exact nucleotide match. Consensus sequences of V_H and V_L genes (i.e. from all reads passing quality filters and that contained exact matches to the CDRH3:CDRL3 pair of interest) were used for antibody gene synthesis, expression, and *in vitro* analysis (DeKosky et al., 2013).

IgG Synthesis, Expression, and Purification

Consensus V_H and V_L genes were designed and purchased as gBlocks (Integrated DNA Technologies) and cloned into the pcDNA3.4 vector (Invitrogen) containing *Oryctolagus cuniculus* IgG leader peptide as fusions to human IgG1 and kappa constant regions, respectively. Sequences of both the heavy and the light chain for each antibody variant were confirmed by Sanger sequencing. Plasmids for each antibody variant were transfected into Expi293 cells (Invitrogen) at a 1:3 heavy:light ratio. After incubating at 37°C with 8% CO_2 at 125 rpm for 6 days, the supernatant containing secreted antibodies was collected by centrifugation at 500 g for 15 min at 25°C. Supernatant was passed over a column of 0.5 mL Protein A agarose resin (Thermo Scientific) three times to ensure efficient binding. After washing with 20 column volumes of PBS, antibodies were eluted with 3 mL 100 mM citric acid pH 3.0 and immediately neutralized with 500 μL 1M Tris pH 8.0. Antibodies were buffer exchanged into PBS, pH 7.4 utilizing Amicon Ultra-30 centrifugal spin columns (Millipore) for storage and subsequent use.

ELISA

Costar 96-well ELISA plates (Corning) were coated with 50 μL of 4 μg/mL recombinant Ebola Glycoprotein (a gift from Dr. Erica O. Saphire, The Scripps Research Institute) or Ebola virus VLPs. The coated plates were incubated at 4°C O/N, after which they were decanted and blocked with 2% milk in PBS for 2 h at RT. After blocking, 1:5 serially diluted antibodies were applied to the plates for 1h, after which 1:5000 diluted

donkey anti-human IgG HRP-conjuated secondary antibodies were applied (Jackson ImmunoResearch) for 1 h. For detection, 50 μL TMB-Ultra substrate was applied for 15 min before quenching with 50 μL 2 M H_2SO_4. Absorbance was measured at 450 nm using a Tecan M200 plate reader. Data were analyzed and fitted for EC_{50} using a 4-parameter logistic nonlinear regression model in the Prism software.

Surface Plasmon Resonance

Antibody affinity to recombinant Ebola GP protein was measured by surface plasmon resonance using a BIAcore 3000 biosenor (Biacore). In order to fit the responses to 1:1 Langmuir binding model for more accurate affinity determination, antibodies were immobilized on the CM5 sensor chip (GE Healthcare) using the amine coupling chemistry. All binding experiments were done in HBS-EP buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 3.4 mM EDTA, and 0.005% P20 surfactant) (GE Healthcare). GP was injected in triplicates at concentrations 80, 100, 200, 300, 400, 500, and 600 nM with a flow rate of 60 µL/min for 2 min and a dissociation time of 10 min. Regeneration of the antibody was performed by a single injection of 100 mM citric acid, pH 3.0. The response generated by flowing GP over a bovine serum albumin (BSA) coupled surface was used as control and was consequently subtracted. All kinetic parameters were determined in BIAevaluation 3.0 software and were reported as the average of three technical replicates.

Wild type virus ELISA assays

All work with wild type virus was performed at BSL4 at Texas Biomedical Research Institute. All virus stocks were cultivated on Vero-E6 cells in DMEM with 2% FBS and antibiotics. When 80% of cells began showing a cytopathic effect, the culture supernatant containing virus was collected. Virus was purified as for VLPs by pelleting

cell debris and then pelleting virus from the culture supernatants through 20% sucrose in 20 mM NaCl and 20 mM HEPES, pH 7.4. The virus pellets were resuspended in PBS and stored in aliquots at -80°C until needed. Virus titers were determined by conventional plaque assay using Vero-E6 cells. For ELISA assays, an aliquot of virus was thawed and the equivalent of 10^6 PFU of virus was diluted 1:3 into RIPA buffer. This was then diluted 1:100 into 10 mM sodium phosphate buffer, pH 7.4. After coating O/N, plates were washed with PBS containing 0.1% Tween-20 and incubated with each antibody starting at 1:100 of a 1 mg/mL stock and then over serial 4-fold dilutions on the plates. The secondary antibody was anti-human IgG HRP conjugate from Pierce. TMB Ultra substrate (Life Technologies) was used to detect antibody binding on plates. All assays were performed at least in duplicate and repeated 3 times. ELISAs were analyzed and fitted for EC_{50} using a 4-parameter logistic nonlinear regression model in the Prism software.

ACKNOWLEDGEMENTS

This study was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under Award Number R01AI096228 and by the Defense Threat Reduction Agency grant HDTRA1-12-C-0105 and HDTRA1-12-C-007.

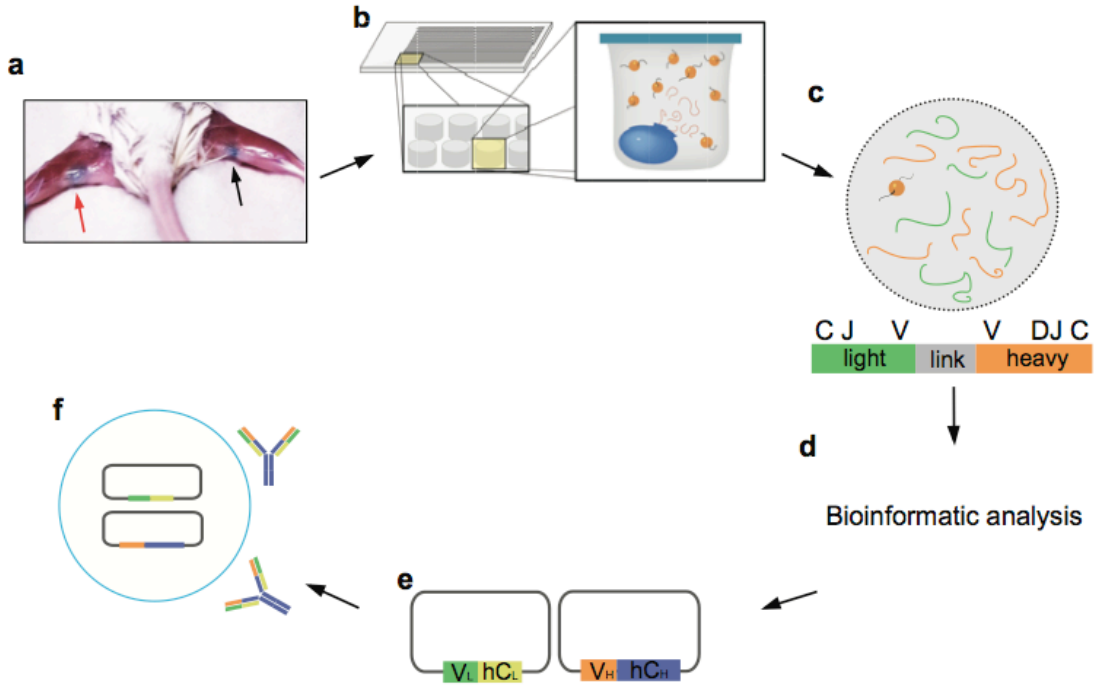

Figure 2-1: Isolation of antibodies by mining the paired V_H:V_L repertoire of draining popliteal lymph node (PLN) antibody-secreting B cells.

(a) Footpad immunization leads to a marked increase in cellularity within the ipsilateral popliteal lymph node relative to the contralateral lymph node (red and black arrows respectively). (b) PLN CD138$^+$ cells isolated by magnetic sorting are deposited into 125 pL wells on PDMS slides that also contain poly(dT) beads. Cells are lysed *in situ* and mRNA is captured on the poly(dT) beads. (c) The poly(dT) beads are emulsified and V_H:V_L amplicons are generated following reverse transcription and overlap extension PCR. (d) V_H:V_L amplicons are sequenced using Illumina 2x250 MiSeq and the highest frequency V_H:V_L pairs are identified via bioinformatics analysis. (e) Highest frequency V_H (orange) and V_L (green) genes are synthesized and cloned into IgH and IgL expression vectors containing human IgG1 (blue) and human kappa (yellow) constant regions, respectively. (f) Following co-transfection into Expi293 cells, recombinant IgG antibodies are expressed and purified.

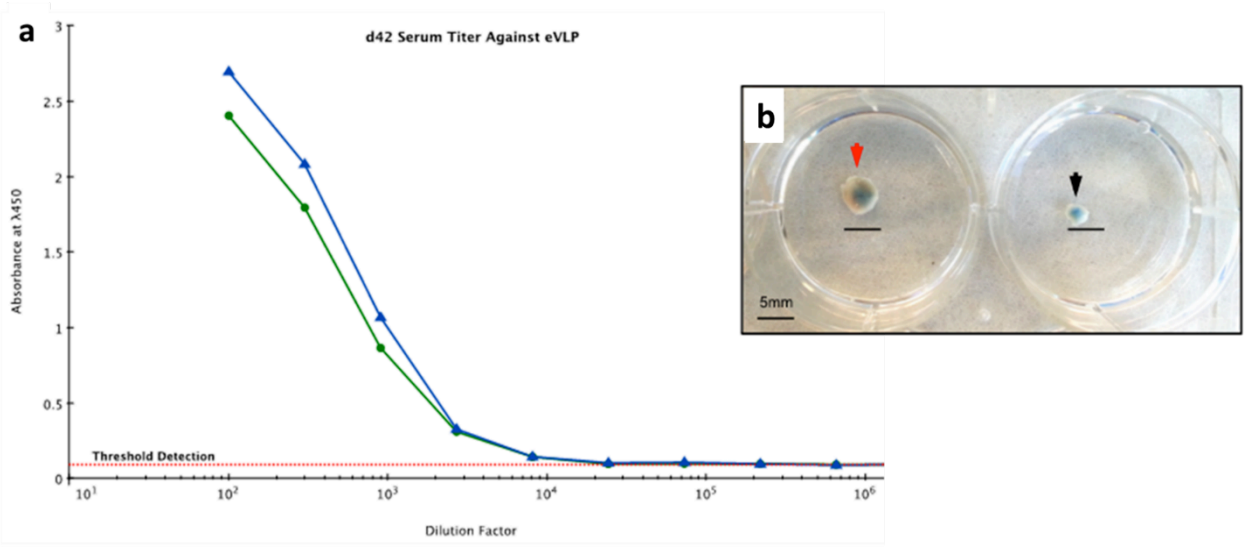

Figure 2-2: Footpad immunization with EBOV VLPs.

BALB/c mice were immunized with 5 μg Ebola VLPs in 10 μL PBS pH 7.4 emulsified in a 1:1 ratio with TiterMax Gold adjuvant with boosts on days 21 and 35. (a) Tail-vein bleeds were performed 7 days post day 35 boost, and antibody titers against VLPs measured in Mouse 1 (blue) and Mouse 2 (green). (b) Size comparison of the dissected ipsilateral (red arrow) and contralateral (black arrow) LNs. Ipsilateral LN: ~5mm in diameter; contralateral LN: ~2mm in diameter.

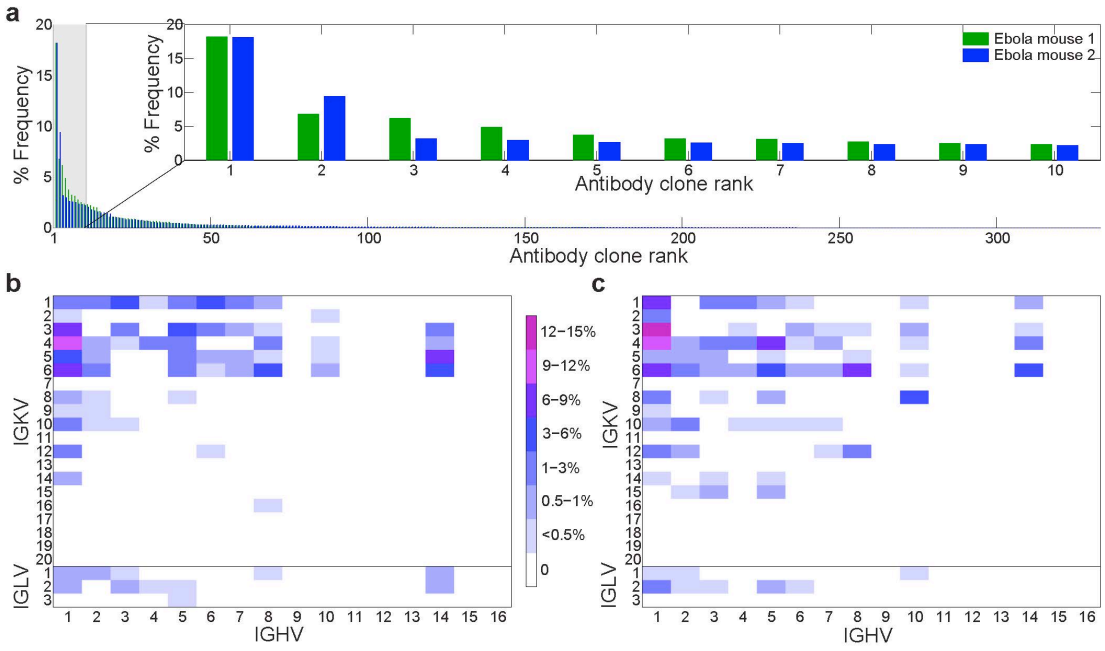

Figure 2-3: Characteristics of the PLN V_H:V_L repertoire in CD138$^+$ antibody secreting cells.

(a) Polarization of V_H:V_L repertoire after immunization with EBOV VLPs. The frequency of each unique CDRH3:CDRL3 antibody clonotype is shown as a percentage of total sequencing read counts. CDRH3 sequencing reads having at least 96% identity at the nucleotide level were clustered and compiled, then analyzed for corresponding CDRL3 per pair. Sequences identified in <2 reads were excluded to minimize sequencing error. Inset: frequency of the ten most frequently observed CDRH3:CDRL3 antibody clonotypes from each mouse. (b, c) V_H:V_L gene family usage of unique CDRH3:CDRL3 clonotypes in mouse ZM1 and ZM2, respectively.

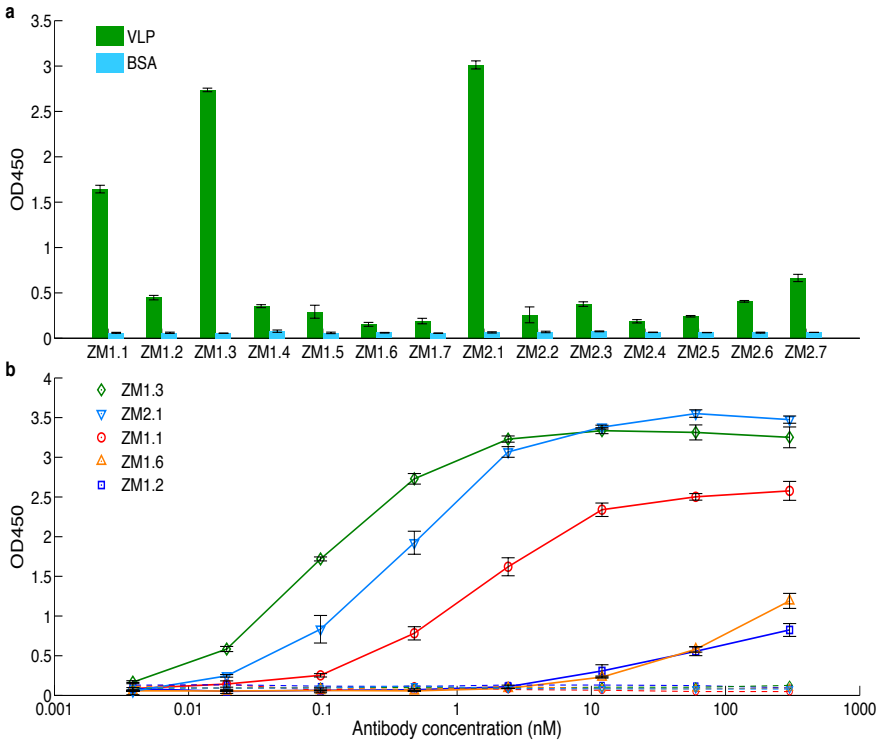

Figure 2-4: Functional characterization of IgG antibodies isolated via mining of the PLN CD138$^+$ B cell repertoire.

(a) Binding to EBOV VLPs for antibodies encoded by the seven most frequently observed CDRH3:CDRL3 clonotypes from the sequenced PLN CD138$^+$ B cell repertoires of each mouse. (b) Binding to purified EBOV recombinant GP for select antibodies as determined by ELISA. Binding to BSA as a control is shown in correspondingly colored dashed lines. Error bars represent the standard error of the mean for three technical replicates.

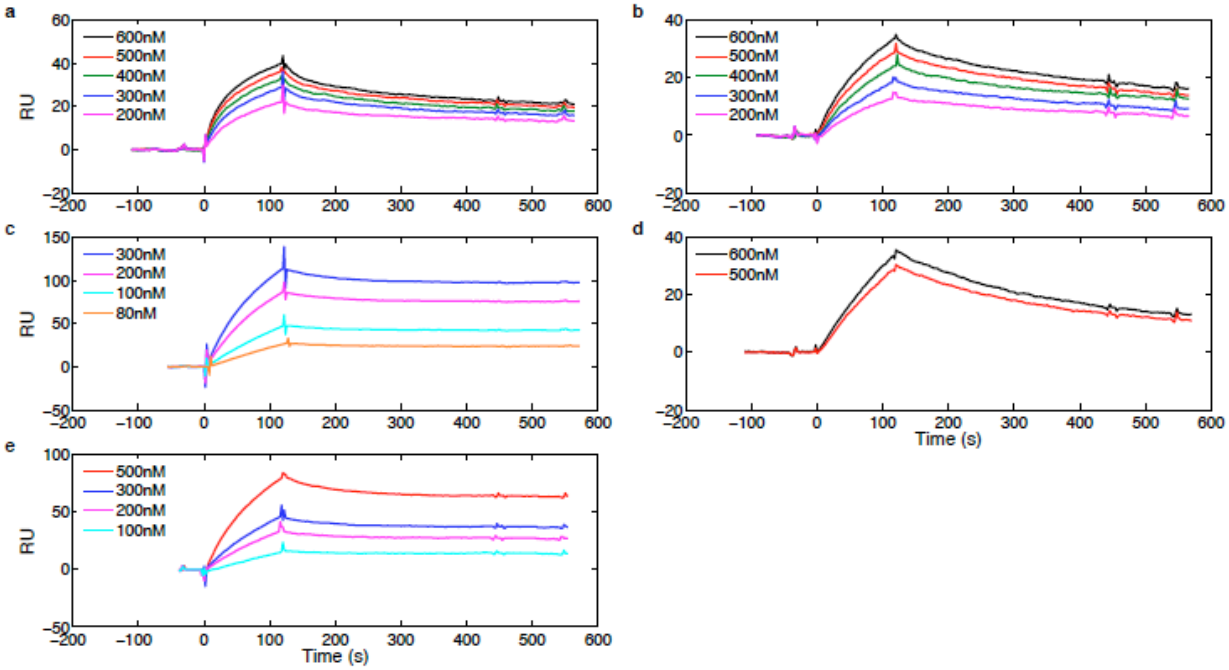

Figure 2-5: BIACore sensorgrams for selected antibodies binding to recombinant EBOV GP.

(a) ZM1.1, (b) ZM1.2, (c) ZM1.3, (d) ZM1.6, (e) ZM2.1. Antibodies were immobilized on CM5 sensor chips and varying concentrations of GP were injected using a flow rate of 60 μL/min for 2 min. Experiments were performed in three technical replicates, and all curves were fit to a 1:1 Langmuir binding model using BIAevaluation software.

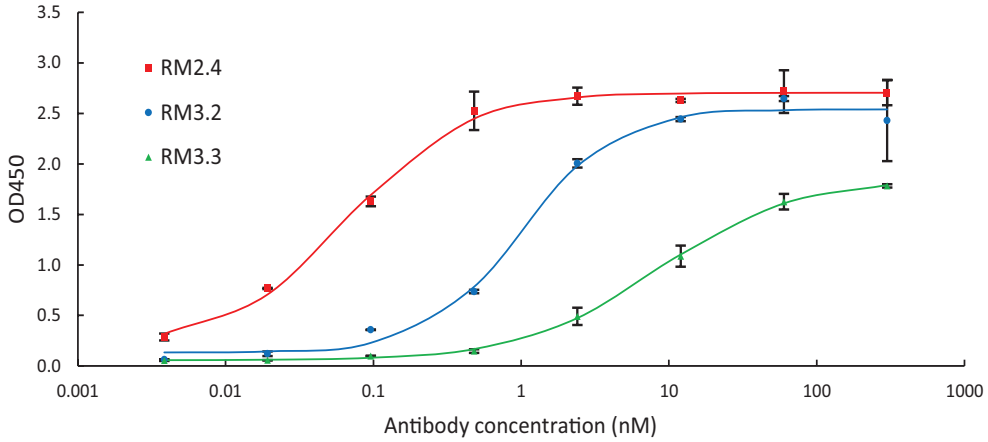

Figure 2-6: Functional characterization of IgG antibodies isolated via mining of RESTV immunized PLN CD138^{+} B cell repertoire.

Binding to RESTV recombinant GP for select antibodies as determined by ELISA. Curves were fitted using 4-parameter logistic non-linear regression. Error bars represent the standard error of the mean for three technical replicates.

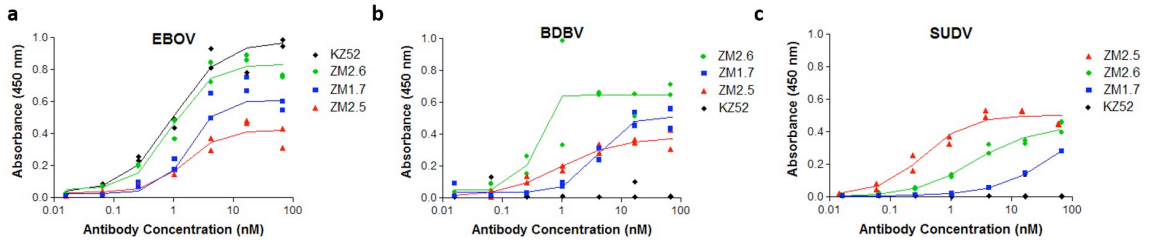

Figure 2-7: Binding of cross-reactive antibodies isolated via mining of the PLN CD138^{+} B cell repertoire to wild type EBOV, BDBV and SUDV viruses.

ELISA assays using antibodies ZM1.7, ZM2.5, ZM2.6, and KZ52 for the detection of (a) wild type Ebola virus; (b) wild type Bundibugyo virus; and (c) wild type Sudan virus. Assays were performed in two technical replicates. Lines represent measurements fitted via 4-parameter logistic nonlinear regression for EC_{50}.

Mouse	Rank (Name)	CDRH3 Sequence	CDRL3 Sequence	Gene Usage
1	1 (ZM1.1)	ARSFAY	QQSNEDPYTF	IGHV1-IGHJ3:IGKV3-IGKJ2
	2 (ZM1.2)	TGDGYYGFAY	FQGSHVPFT	IGHV6-IGHD2-IGHJ3:IGKV1-IGKJ4
	3 (ZM1.3)	ARGIGY	WQGTHFPFT	IGHV3-IGHJ3:IGKV1-IGKJ4
	4 (ZM1.4)	ARSTTATLDC	QQSDSWPTLT	IGHV14-IGHD1-IGHJ2:IGKV5-IGKJ5
	5 (ZM1.5)	ATISTATFPY	QQSDSWPTLT	IGHV1-IGHD1-IGHJ3:IGKV5-IGKJ5
	6 (ZM1.6)	ARRAMITTEGVDFDY	QQSRKVPWT	IGHV3-IGHD2-IGHJ2:IGKV3-IGKJ1
	7 (ZM1.7)	AREGYRYDWYFDV	QQRSSYPLT	IGHV1-IGHD2-IGHJ1:IGKV4-IGKJ5
2	1 (ZM2.1)	TRSVSDY	WQGTHFPHT	IGHV1-IGHD2-IGHJ2:IGKV1-IGKJ5
	2 (ZM2.2)	ARRTYRYDRFDY	QQWSSDPLT	IGHV1-IGHD2-IGHJ2:IGKV4-IGKJ5
	3 (ZM2.3)	TRRSNFPYYFDF	QQSIEDPFT	IGHV1-IGHD2-IGHJ2:IGKV3-IGKJ4
	4 (ZM2.4)	ARSELGATGFAY	QQGQSYPIFT	IGHV5-IGHD3-IGHJ3:IGKV15-IGKJ4
	5 (ZM2.5)	ARQKYGNYVLYWYFDV	QQWNSNPPT	IGHV5-IGHD2-IGHJ1:IGKV4-IGKJ4
	6 (ZM2.6)	TGMVTSY	LQHWNYPYT	IGHV6-IGHD2-IGHJ3:IGKV6-IGKJ2
	7 (ZM2.7)	VREGLGSYFDY	QQYYNYPRT	IGHV10-IGHD5-IGHJ2:IGKV8-IGKJ1

Table 2-1: List of characterized EBOV antibodies sequenced from PLN CD138^{+} cells.

For each antibody, CDRH3:CDRL3 clonotypes and their V(D)J gene assignment are provided.

Antibody ID	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	K_D (nM)
ZM1.1	(3.12 ± 0.55) $x10^4$	(1.16 ± 0.07) $x10^{-3}$	37.7 ± 0.55
ZM1.2	(1.17 ± 0.41) $x10^4$	(1.71 ± 0.12) $x10^{-3}$	156 ± 47.4
ZM1.3	(3.12 ± 0.53) $x10^4$	(2.39 ± 0.62) $x10^{-4}$	7.71 ± 1.77
ZM1.6	(3.95 ± 0.7) $x10^3$	(2.46 ± 0.12) $x10^{-3}$	635 ± 93.8
ZM2.1	(1.47 ± 0.23) $x10^4$	(2.97 ± 0.07) $x10^{-4}$	20.5 ± 3.44

Table 2-2: SPR Binding kinetics and equilibrium dissociation constants (K_D) towards uncleaved EBOV GP.

Experiments were performed in three technical replicates and data were fit to a 1:1 Langmuir binding model.

Chapter 3: Discovery of high affinity anti-ricin antibodies by B cell receptor sequencing and by yeast display of combinatorial V_H:V_L libraries from immunized animals[2]

Ricin is a toxin that could potentially be used as a bioweapon. We identified anti-ricin A chain antibodies by sequencing the antibody repertoire from immunized mice and by selecting high affinity antibodies using yeast surface display. These methods led to the isolation of multiple antibodies with high (sub-nanomolar) affinity. Interestingly, the antibodies identified by the two independent approaches are from the same clonal lineages, indicating for the first time that yeast surface display can identify native antibodies. The new antibodies represent well-characterized reagents for biodefense diagnostics and therapeutics development.

INTRODUCTION

Ricin is a toxin derived from the castor bean *Ricinus communis*, and it is classified as a Category B Agent by the Centers for Disease Control and Prevention in part because of its high lethality (LD_{50} about 22 μg/kg body weight for human) and ease of production (Lord et al., 1994). Ricin is composed of an enzymatic A chain, which deactivates eukaryotic ribosomes by depurinating adenine 4324 in the 28S rRNA of the 60S ribosomal subunit, and a lectin B chain, which binds carbohydrates on the cell surface (Audi et al., 2005; Lord et al., 1994). As a result, ricin is a potent biological weapon, and past cases of malicious exposure to ricin highlight the need for both an increase in the sensitivity of diagnostics and for treatment after poisoning (Knight, 1979; Mayor, 2003). Antibodies to both the A chain and B chain of ricin have been produced

[2]This chapter is reproduced with minor modifications from its initial publication:
Wang, B., Lee, C.H., Johnson, E.L., Kluwe, C.A., Cunningham, J.C., Tanno, H., Crooks, R.M., Georgiou, G., and Ellington, A.D. (2016). Discovery of high affinity anti-ricin antibodies by B cell receptor sequencing and by yeast display of combinatorial VH:VL libraries from immunized animals. MAbs *8*, 1035–1044. Wang, B. designed and performed the experiments, and analyzed the data.

using hybridoma-based technology (Maddaloni et al., 2004; Mantis et al., 2006; Neal et al., 2010; Prigent et al., 2011), and by phage display screening of immune libraries reconstituted from the bone marrow of immunized cynomolgus macaques (Pelat et al., 2009), resulting in antibodies with affinities ranging from 40 pM to 5 nM. However, there are currently no U.S. Food and Drug Administration-approved treatments for ricin poisoning, and the diagnostic potential of these reported antibodies has yet to be tested.

In vitro screening of large combinatorial libraries using display technologies that rely on phage, bacteria, yeast, mammalian cells or even in vitro transcription/translation systems are widely employed for antibody discovery (Bradbury et al., 2011; Georgiou et al., 2014a; Hoogenboom, 2005). Combinatorial antibody libraries are constructed either by mining the natural diversity of immunoglobulin genes in immunized or antigen-naïve animals (Feldhaus et al., 2003; Hayhurst et al., 2003; Sheets et al., 1998), or by diversifying the complementarity-determining regions (CDRs) within one or more "scaffold" antibodies (Knappik et al., 2000). Variable heavy (V_H) and variable light (V_L) chain genes are then joined combinatorially, yielding, at least in theory, libraries that contain combinations of heavy chains joined with all possible light chains (Feldhaus et al., 2003; Hayhurst et al., 2003; Knappik et al., 2000; Sheets et al., 1998). Typically, immune libraries constructed from mRNA obtained from the spleen, bone marrow or from peripheral blood mononuclear cells (PBMCs, primarily in the case of human donors) are more likely to encode a significant fraction of antigen-specific antibodies and thus represent the most reliable route to high affinity antibodies (provided that the antigen is immunogenic) (Hayhurst et al., 2003; Maruyama et al., 1999). Isolation of high affinity antibodies from immune and other libraries is most readily accomplished by taking advantage of the quantitative nature of fluorescence-activated cell sorting (FACS)-based

library screening methods, and yeast display has been established as the dominant technology for screening libraries by FACS (Angelini et al., 2015).

The power of combinatorial library screening has been validated by the identification and development of therapeutic antibodies that are now entering clinical trials (Brekke and Sandlie, 2003). However, certain antibodies isolated from combinatorial libraries may express at low yields in mammalian cells and display poor biophysical properties in vitro (Liu et al., 2014; Ponsel et al., 2011; Wang et al., 2014; Wörn and Plückthun, 2001; Xu et al., 2013), which can hamper their development into therapeutics. The random pairing of heavy and light chains in combinatorial libraries results in antibodies with non-natively paired V_H and V_L genes, and this is potentially one cause of poor antibody expression and stability (Tiller et al., 2013). Moreover, a recent report has suggested that several mouse and human V_H germ-line genes exhibit strong preferential pairing with specific V_L chains (Jayaram et al., 2012), and the random pairing of non-preferential V_H and V_L chains could lead to conformational incompatibilities, again affecting expression and stability. Enormous efforts have been made to advance screening techniques to remove antibodies with mediocre biophysical properties and on investigating methods to improve them in the past decade (Liu et al., 2014; Mcconnell et al., 2014; Rouet et al., 2014; Wang et al., 2014; Xu et al., 2013).

Our lab has pioneered methods for the discovery of high affinity antigen-specific antibodies directly via mining of the immunoglobulin repertoire by capitalizing on next-generation sequencing technologies without the need for screening (DeKosky et al., 2013, 2015; Reddy et al., 2010; Wang et al., 2015a). Specifically, we developed methods for high-throughput determination of the natively paired V_H:V_L repertoire from single B cells (DeKosky et al., 2013, 2015). More recently, we have shown that antibody secreting B cells ($CD138^+$ plasmablasts) within the draining lymph node are overwhelmingly antigen-

specific, and that antibodies derived from these cells exhibit high binding affinity (Wang et al., 2015a).

In this report, we sought to compare the isolation of high-affinity anti-ricin antibodies via mining of the draining lymph node repertoire and via yeast display of immune combinatorial libraries constructed from antibody mRNAs obtained from spleen or bone marrow cells. Overall, both approaches yielded strong ricin A chain binders (the lowest K_d values were 0.97 and 0.58 nM, respectively, for these two methods). Interestingly, we found that antibodies isolated by yeast display from combinatorial libraries in which the V_H and V_L from spleen or bone marrow had been randomly paired were clonal relatives of antibodies identified via mining of the draining lymph node repertoire, and comprised authentic, natively paired, V_H and V_L sequences. Thus, in hyperimmune animals where antigen-specific antibodies comprise a significant portion of the repertoire (Becker et al., 2010; Kuroiwa et al., 2009; Kurosawa et al., 2012; Reddy et al., 2010; Wang et al., 2015a), flow cytometric screening of libraries constructed via the random pairing of V_H and V_L genes can nonetheless result in the isolation of native antibodies.

RESULTS

Overview of the experimental approach

As we have previously shown, mouse footpad immunization results in a strong immune response in the popliteal draining lymph node (DLN), characterized by the robust expansion of antibody secreting B cells ($CD138^+$ plasmablasts) (Wang et al., 2015a). Analysis of the DLN $CD138^+$ B cell antibody repertoire overwhelmingly revealed that the most highly represented sequences, i.e., those present in the largest number of reads, are antigen-specific (Wang et al., 2015a). Therefore, we sought to

explore the antibody repertoire in DLN antibody-secreting B cells, and to compare antigen-specific antibodies with those selected from bone marrow and spleen combinatorial libraries, which are commonly used in antibody discovery efforts following intraperitoneal or subcutaneous immunization (Hayhurst et al., 2003; Pelat et al., 2009; Saggy et al., 2012). In contrast, samples from the DLN could only support high-throughput V_H:V_L sequencing and repertoire analysis, as we were limited by the number of cells. Four mice were immunized by footpad injection with ricin A chain in TiterMax Gold adjuvant (Fig. 3-1A), followed by two booster immunizations after 14 and 28 days, after which all four mice generated significant titers (>1:10,000) against ricin A chain (Fig. 3-2A). One more booster immunization was performed, after which even higher titers were reached (Fig. 3-2B). Mice were sacrificed six days after the final booster immunization, at which point the draining lymph node was observed to be enlarged compared to the contralateral lymph node. Spleen and bone marrow were collected to screen for antigen-specific antibodies (Fig. 3-1B), and the draining lymph node was collected as well for V_H:V_L mining (Fig. 3-1C).

Isolation of high affinity ricin A chain antibodies by yeast display of combinatorial libraries

Single-chain variable fragment (scFv) libraries were constructed using mRNA from ~ 5 x 10^6 total splenocytes, and separately from femur bone marrow cells (Fig. 3-1B). Briefly, V_H and V_L genes were amplified separately, paired combinatorially by overlap extension-polymerase chain reaction that also introduced a $(Gly_4Ser)_3$ linker (Hayhurst et al., 2003), and then subcloned into the yeast display vector pCTCON2 via homologous recombination (Chao et al., 2006).

Two libraries of ~ 2 x 10^6 transformants each (bone marrow-derived and splenocyte-derived) were obtained. Screening was performed by carrying out one round

of magnetic-activated cell sorting (MACS), and then three rounds of FACS with progressively more stringent gates that collected cells positive both for expression (using Alexa Fluor 488-labeled anti-c-Myc) and for ricin A chain binding (using Alexa Fluor 633-labeled ricin A chain) (Fig. 3-3). For the first two rounds of selection, ricin A chain was used at a 1 μM concentration. Prior to selection, both libraries showed little binding to ricin A chain (Fig. 3-3A and B, before selection). Significant enrichment of cells that stained positively both with fluorescent anti-c-Myc antibody and ricin A chain was observed from both libraries after the second round (Fig. 3-3A and B, after 1st FACS). Subsequently, ricin A chain concentrations of 200 nM and 40 nM were used for the third and fourth rounds of sorting, respectively. After the fourth round, almost all of the cells that expressed scFv also showed binding to ricin A chain (Fig. 3-3A and B, after 3rd FACS).

Twenty colonies from each library were sequenced, yielding a total of 5 unique clones (3 from the bone marrow-derived library and 2 from the splenocyte-derived library), indicating that screening had led to convergence of the libraries to a small number of scFv sequences (Table 3-1). Apparent equilibrium binding constants (K_d) for these 5 clones were estimated by incubating yeast cells displaying the respective antibodies with different concentrations of ricin A chain, ranging from 0.09 to 600 nM, and then measuring the mean fluorescence intensity of binding at each concentration by FACS. Binding constants were calculated using a Langmuir 1:1 binding model (Chao et al., 2006). All 5 antibodies demonstrated low nanomolar affinities binding to ricin A chain (Table 3-2). These 5 antibodies were then expressed in *E. coli* as single-chain antibodies (scAbs, scFv fused with human kappa light chain at the C-terminus (Hayhurst et al., 2003)), and purified to >95% homogeneity by Ni-NTA chromatography (as determined by SDS-PAGE). The antibodies were further purified by size-exclusion

chromatography, which also revealed that all antibodies were monomeric. Surface plasmon resonance (SPR) analyses indicated that, consistent with the flow cytometric analyses (Chao et al., 2006), all 5 antibodies displayed high affinity binding to ricin A chain (Table 3-2, Fig. 3-4).

The question of whether in vitro screening technologies can identify native antibodies in a repertoire has generally not been addressed; the two methods often return different antigen-specific antibodies. Different screening platforms lead to the identification of different antibodies, but some platforms seem to return antibodies that are of low abundance or apparently not present at all (Bowley et al., 2007; Saggy et al., 2012). The identification of antibodies via yeast display presented an opportunity to compare the results with the natural repertoire, and thus we sequenced both the bone marrow and spleen V_H and V_L repertoires using an Illumina 2 x 250 Miseq platform. From the same 5 x 10^6 cells of bone marrow and spleen from a given mouse that were used for yeast library construction, about 10^6 and 7 x 10^5 reads, respectively, were obtained for V_H repertoires, from which 8735 and 7057 unique clonotypes (defined as the group of V_H sequences that share the same germ-line V and J segments, and have >90% amino acid identity in their CDRH3s (Lavinder et al., 2014) were identified (Table 3-3). Similar to previous reports (Lu et al., 2014; Reddy et al., 2010), both repertoires were observed to have skewed germ-line V gene usage. For example, both repertoires showed biases for IGHV1, 5, 14, and IGKV1, 4, 6 families (Fig. 3-5A, B), with IGKV4 as the most commonly used light chain family in the bone marrow, and IGKV6 as the most commonly used light chain in the spleen. We noticed the preferential usage of the IGKV1 family in bone marrow relative to spleen, and of the IGKV6 family in spleen relative to bone marrow. It will be interesting to determine if similar repertoire biases towards specific germline light chain families are observed with other antigens and adjuvant

combinations. Both repertoires showed a CDRH3 length distribution comparable to those previously reported from animals immunized with different antigens (Reddy et al., 2010), with a median CDRH3 length of 13 amino acids in the bone marrow repertoire and 11 amino acids in the spleen repertoire (Fig. 3-5C).

Importantly, we now find that the antibodies identified by yeast surface display are amongst the most abundant antibodies in the sequenced repertoire (Table 3-1). This indicates for the first time that yeast surface display of combinatorial libraries can identify native antibodies in the repertoire.

Isolation of high affinity anti-ricin A chain antibodies by next-generation sequencing of DLN B cells

$CD138^+$ antibody secreting B cells were isolated from the draining lymph node and the paired V_H:V_L repertoire was determined, as previously described (Wang et al., 2015a). Briefly, a total of 90,000 $CD138^+$ antibody secreting cells were isolated by magnetic sorting by first depleting $CD45R^+$, $CD49b^+$, and $CD19^+$ cells, and then enriching for $CD138^+$ cells. Approximately 40,000 antibody secreting cells were deposited into microfabricated nanowell plates such that >98% of wells contained a single cell, after which poly(dT) magnetic beads were added to capture mRNAs, the cells were lysed, and the poly(dT) beads were collected. The beads were emulsified, and overlap extension RT-PCR was performed to synthesize linked V_H:V_L amplicons (Fig. 3-1C) (DeKosky et al., 2013). The region of the linked amplicons comprising CDRH3 and CDRL3 was then sequenced in an Illumina MiSeq 2 x 250 run. Full-length V_H and V_L sequences were identified in separate MiSeq 2 x 250 runs. After bioinformatics analysis (DeKosky et al., 2013), 212 unique V_H:V_L pairs were identified in the draining lymph node repertoire, and these showed great polarization, indicative of clonal expansion. Specifically, the top 10 most abundant V_H:V_L pairs constituted 65.8% of the total reads

(Fig. 3-6A). The virtual absence of antibodies using IGLV families in the repertoire was expected, as only 5% of all mouse antibodies use the lambda light chain (Lu et al., 2014).

Comparisons with mouse antibody repertoires generated by immunization with other antigens showed similar germ-line gene usage between the anti-ricin A chain antibody repertoire and other repertoires (Fig. 3-6B) (Reddy et al., 2010; Wang et al., 2015a). The CDRH3 length distribution in $CD138^+$ antibody secreting cells from the draining lymph node showed three peaks at 12, 13, and 15 amino acids (Fig. 3-6C).

Synthetic genes for the top 10 highest frequency paired V_H:V_L sequences from the DLN repertoire were constructed and the respective mouse-human chimeric antibodies (mouse V_H fused with the human IgG1 constant domain, and mouse V_L fused with the human kappa constant domain) were expressed in Expi293 cells and purified. 4/10 recombinant antibodies displayed binding by ELISA (Table 3-4, Fig. 3-7). SPR analysis was used to obtain antigen-binding kinetics for these 4 clones (Table 3-5, Fig. 3-4). All but one antibody showed high affinity binding to the ricin A chain. Interestingly, the affinities of the antibodies that bound ricin A chain correlated with their abundance in the V_H:V_L repertoire, with the antibody displaying the lowest K_d (RAM1.2) being the most abundant.

We compared the native antibodies recovered from the DLN to yeast display-selected antibodies from the bone marrow and spleen combinatorial libraries. Interestingly, BM1, BM3, BM17 and SP19 were found to share germ-line V and J segments with RAM1.5, and SP1 had the same germ-line V and J segments as RAM1.4. The two groups of antibodies also had the same V_H:V_L pairing. Furthermore, antibodies recovered from the DLN without selection had very similar CDRH3s with those isolated by yeast display from bone marrow and spleen (Table 3-1, 3-4), as well as additional

amino acid substitutions, indicating that they correspond to somatic variants derived from the same antibody lineage during clonal expansion (Fig. 3-8).

DISCUSSION

In vitro screening of antibody libraries constructed from different sources (synthetic, naïve, or immune) is one of the most commonly used techniques for antibody discovery (Bradbury et al., 2011; Hoogenboom, 2005). In this work, we identified anti-ricin A chain antibodies by eliciting an immune response in mice and then exploring the resultant antibody repertoires via both yeast display of combinatorial libraries from the bone marrow and spleen and paired V_H:V_L sequencing from the draining lymph node.

While both approaches generated high affinity clones (K_d ranging from 0.55 nM to 6.12 nM), as has previously been observed (Bowley et al., 2007), unexpectedly the antibodies generated by these two very different approaches were found to be clonal relatives. Different amino acids encoded by different nucleotides in the isolated antibodies from different lymphoid organs indicated that BM1, BM3, BM17 (derived from bone marrow using yeast display), SP19 (derived from the spleen using yeast display) and RAM1.5 (derived from the draining lymph node using paired V_H:V_L sequencing), and separately SP1 and RAM1.4 were clonal variants (Fig. 3-8). A previous report has shown that while naïve mouse spleen and lymph node repertoires share some common CDRH3s, the scFvs isolated from these lymphoid organs following intraperitoneal and subcutaneous immunization were unique (Venet et al., 2013). In contrast, when we probed different lymphoid organs, clonal variants were obtained. These differing conclusions may reflect differences in the immunization routes used, the numbers of boost immunizations, and significant differences introduced by the expression biases in phage compared to yeast display. The clonal variants we isolated

from different lymphoid organs indicate that these antibodies were likely encoded by the same ancestral B cells that homed to different lymphoid organs and experienced clonal expansion there.

We have previously shown that by dissecting the draining lymph node antibody repertoire with paired V_H:V_L sequencing, the highly abundant clones were usually antigen specific as a result of clonal expansion (Wang et al., 2015a). Here, we again showed that the highly abundant clones in the draining lymph node probed with paired V_H:V_L sequencing encoded high-affinity, antigen-specific antibodies. Comparison of the antibodies isolated using these two approaches supports our hypothesis that yeast display can recover native antibodies in a repertoire, as antibodies isolated by yeast display are clonal variants of those isolated by paired V_H:V_L sequencing, which identifies native V_H:V_L pairings. Given the relative abundance of the V_H and V_L sequences of antibodies in the bone marrow and spleen repertoires, the calculated probability of obtaining these native antibodies by combinatorially random pairing would be about 0.2%. Thus, in hyperimmunized animals the proportion of randomly paired V_H and V_L genes is sufficiently high that natively paired heavy and light chains occur. Moreover, high affinity native pairs can be isolated following yeast surface expression and FACS. To our knowledge this is the first report that yeast display selections can identify native antibodies from the immune repertoire.

The fact that an in vitro combinatorial library screening method can identify native V gene pairs is relevant for antibody discovery. Numerous successes in developing antibody drugs from combinatorial libraries generally validate this approach (Brekke and Sandlie, 2003). However, some high-affinity antibodies identified by library screening have mediocre biophysical properties, such as low yield and poor stability (Liu et al., 2014; Ponsel et al., 2011; Wang et al., 2014; Wörn and Plückthun, 2001; Xu et al., 2013).

Such deficiencies are thought to emanate from random or non-native pairings of V_H and V_L that lead to unfavorable conformations that in turn promote poor biophysical properties (Jayaram et al., 2012; Tiller et al., 2013). Another contributor may be that expression in a non-native host (i.e., mouse antibody libraries screened by phage display) leads to selection at the expression level against certain clones (Bowley et al., 2007). In contrast, antibodies with natively paired heavy and light chains, such as those isolated from hybridomas, express better, and in general show much better stability (Dessain et al., 2008; Jayaram et al., 2012; Tiller et al., 2013). Our methods can therefore potentially streamline the huge amount of effort put into the discovery process (Liu et al., 2014; Mcconnell et al., 2014; Rouet et al., 2014; Wörn and Plückthun, 2001; Xu et al., 2013).

Our results support the utility of next-generation sequencing in combinatorial library screening (Georgiou et al., 2014a; Reddy et al., 2010; Wang et al., 2015a). Hyperimmunization leads to a high degree of polarization of the B-cell repertoire in a particular compartment, such that the dominant antibodies are highly represented and combinatorial screening by a technique with low expression bias, such as yeast display, can further yield high-affinity native antibodies. The high-affinity anti-ricin A chain antibodies we identified may provide a greater diversity of reagents for both diagnostics and therapeutics.

MATERIALS AND METHODS

Cell line and media

The yeast strain EBY100 (MATa *AGA1::GAL1-AGA1::URA3 ura3-52 trp1 leu2Δ200 his3Δ200 pep4::HIS3 prb11.6R can1 GAL*) was used for library construction and screening. Yeast cells were maintained in YPD medium (20 g/l dextrose, 20 g/l peptone, and 10 g/l yeast extract); after library transformation, they were maintained in

SDCAA medium (20 g/l dextrose, 6.7 g/l yeast nitrogen base, 5 g/l casamino acids, 8.56 g/l $NaH_2PO_4.H_2O$, and 10.2 g/l $Na_2HPO_4.7H_2O$). SGCAA medium (identical to SDCAA except 20 g/l galactose is used instead of dextrose) was used for library induction. *E. coli* strain DH10β was used for subcloning, and *E.coli* strain Jude-1 was used for soluble scAb expression. Expi293 cells (Invitrogen) were used for IgG expression, and were maintained in Expi293 expression medium (Invitrogen).

Antigen and antibodies

Ricin A chain was purchased from Sigma-Aldrich (cat# L9514). It was biotinylated using an EZ-Link Sulfo-NHS-LC-Biotin kit (Thermo Scientific). Chicken anti-c-Myc IgY, Alexa Fluor 488-goat anti-chicken IgG (GaC-488), and streptavidin-Alexa Fluor 633 (SA-633) were obtained from Invitrogen (cat# A-21281, A-11039, and S21375, respectively). Anti-biotin microbeads were purchased from Miltenyi Biotec (cat# 130-090-485).

Mouse immunizations

This study was approved by the University of Texas Institutional Animal Care and Use Committee under protocol# AUP-2013-00009. Ricin A chain was mixed with TiterMax Gold adjuvant (Sigma-Aldrich, cat# T2684) at a 1:1 ratio and pipetted several times to obtain stable emulsions for immunization. On day 1, 4 female BALB/c mice at 6 weeks of age (Jackson Laboratory) were injected subcutaneously at the left hind footpad with 5 μg of ricin A chain in 25 μl antigen adjuvant mixture. A booster immunization was performed on day 14. 7 days after the booster immunization, the serum antibody titers against ricin A chain were determined. About 30 μl of blood was collected from each mouse at a small tail vein incision made with a scalpel blade, and coagulated at room temperature (RT) for 30 min. Following centrifugation at 13000 rpm for 15 min,

the supernatant (serum) was used for ELISAs. The serum was first serially diluted with phosphate-buffered saline (PBS) +2% milk (PBSB) at a 1:3 ratio from 1:100 to 1:218,700. The diluted serum was applied in triplicate onto ELISA plates (Corning) that had been coated with 50 µl of 4 µg/ml of ricin A chain overnight (O/N) at 4ºC and then blocked with PBSB at RT for 2 hours. After incubation at RT for 1 hour, plates were washed with PBS+0.05% Tween-20 (PBST), followed by adding 50 µl of a 1:5000 diluted horseradish peroxidase (HRP)-conjugated goat anti-mouse antibody (Jackson ImmunoResearch, cat# 115-035-166). After 1 hr incubation, plates were washed again with PBST, after which 50 µl of TMB Ultra ELISA substrate (Thermo Scientific, cat# 34028) was added. The reaction was quenched after 10 min with 50 µl of 2 M H_2SO_4. Absorbance at 450 nm was determined using a Tecan M200 plate reader, and the serum titer was calculated as the dilution at which the absorbance was 3 times higher than background. The second booster was performed on day 28, after which significant titers (>1:10,000) were generated in all mice. The final booster was performed on day 42, and 6 d later, mice were sacrificed for lymphoid organ collection.

Lymphoid organ collection

Bone marrow, spleen, and draining lymph node tissues were collected separately. 30 min before sacrifice, 10 µl 2% Evans Blue in PBS (Sigma-Aldirch, w/v, cat# E2129) was injected into the left hind footpad. After CO_2 asphyxiation and cervical dislocation, the blue-stained popliteal lymph node behind the knee was collected. The spleen was also collected. For bone marrow collection, after clipping the ends of the tibias and femurs, bone marrow was flushed out with PBS+0.1% bovine serum albumin (BSA)+2 mM EDTA (PBSM) using a syringe. Each lymphoid organ tissue sample was first mechanically disrupted using a needle, then passed through a 70 µm cell strainer

(Corning) to collect single cells. The single cell suspension of each lymphoid organ was washed with 20 ml of PBSM buffer and resuspended in 2 ml red blood cell lysis buffer (155 mM NH_4Cl, 12 mM $NaHCO_3$, and 0.1 mM EDTA). After incubation at RT for 5 min, cells were diluted with 20 ml of PBS and then spun down. Cells were washed twice more with PBSM buffer, then cells from the bone marrow and spleen were used for yeast library construction, and cells from draining lymph node were used for high-throughput V_H:V_L pairing.

Yeast library construction for bone marrow and spleen repertoires

Cells from the bone marrow and spleen were resuspended in TRI Reagent (Invitrogen), and RNA was extracted from each sample. cDNA was generated using 200 ng of RNA as a template with the SuperScript III First-Strand synthesis kit (Invitrogen). V_H and V_L sequences were amplified using mouse V gene specific primers. For scFv library construction, 50 ng of V_H and V_L DNA were used as templates for overlap extension PCR, and amplified with primers that contained 50 nucleotides at each end that were in common with the display vector, which should be sufficient to promote homologous recombination upon yeast transformation. The display vector pCTCON2 was linearized by NheI and BamHI digestion (Chao et al., 2006) and the digested backbone was purified and co-transformed with 5-fold mass excess of scFv fragments into electrocompetent yeast cells (Benatuil et al., 2010). The libraries were cultured in SDCAA medium at 30ºC for 2 d.

Yeast library screening

Cells were induced in SGCAA medium at 20ºC for 2 d. For the 1st round of selection, 2 x 10^7 cells (10-fold coverage relative to the number of transformants) were incubated with 1 μM biotinylated ricin A chain at RT for 1 hr and washed thoroughly

with PBS+0.5% BSA+2 mM EDTA (PBSA) to remove any unbound ricin A chain protein. Cells were then mixed with 400 μl of anti-biotin microbeads and incubated at 4ºC for 30 min, and washed again. Cells that bound ricin A chain were selected using MidiMACS system (Miltenyi Biotec), and recovered in SDCAA medium at 30ºC. For the 2nd round of selection, 10^7 cells in 200 μl PBS+0.1% BSA (PBSF) were labeled with 1 μM biotinylated ricin A chain and 4 μg/ml chicken anti-c-Myc IgY at RT for 1hr. After washing with PBSF, cells were incubated with 5 μg/ml SA-633 and 5 μg/ml GaC-488 at 4ºC for 30 min. After washing, cells were resuspended in PBSF buffer and sorted using FACS Aria (BD Bioscience). Gated cells were collected from the double positive quadrant as shown in Fig. 3-3. The subsequent 2 rounds of selection were performed in a similar way, except that the concentration of ricin A chain used was decreased to 200 nM and 40 nM, respectively. After 4 rounds of selection, 20 random colonies from each library were picked, and plasmids were isolated using Zymoprep Yeast Plasmid Miniprep II kit (Zymo Research). Plasmids were amplified in *E. coli* DH10β and sequenced. Three and two unique clones were thus isolated from the bone marrow and spleen libraries, respectively.

Plasmids encoding the 5 unique antibody clones above were re-transformed into strain EBY100. Cells were cultured and induced for scFv expression as described above. 10^6 cells were labeled with 1:3 serially diluted ricin A chain, ranging from 0.09 nM to 600 nM. The labeling process was the same as that used for library sorting. After labeling, cells were analyzed on BD FACS Aria and the mean fluorescence intensity (MFI) values were used to calculate apparent equilibrium binding constant (K_d), as described (Chao et al., 2006).

scAb expression and characterization

For soluble antibody purification, antibody genes were subcloned into pMopac16 vector and expressed as scAb fragments in *E.coli* Jude-1 cells (Hayhurst et al., 2003) and then purified by Ni-NTA chromatography according to the manufacturer's instructions (Thermo Scientific). After elution, proteins were dialyzed in PBS, then loaded onto a Hiload 16/600 Superdex 75 pg column (GE Healthcare) and antibody-containing fractions were pooled and concentrated to 2 mg/ml.

For affinity validation, SPR was performed for each antibody clone using a BIAcore 3000 biosenor (Biacore). About 400 RU (response units) of scAbs were immobilized on the CM5 sensor chip (GE Healthcare) using amine coupling chemistry. All binding experiments were done in HBS-EP buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 3.4 mM EDTA, and 0.005% P20 surfactant) (GE Healthcare). Ricin A chain was injected at concentrations of 25, 50, 100, 200, and 400 nM with a flow rate of 30 μL/min for 2 min and a dissociation time of 10 min. The chip was regenerated after each injection by 100 mM citric acid, pH 3.0. The response generated by flowing ricin A chain over a BSA-coupled surface was used as control and consequently subtracted. Experiments were carried out in triplicates. All kinetic parameters were determined in BIAevaluation 3.0 software using a 1:1 Langmuir model and were reported as the average of the three technical replicates.

V gene repertoire sequencing and analysis

V_H and V_L cDNA for bone marrow and spleen samples (the same cDNAs used for yeast library construction) were prepared and sequenced using the Illumina 2 x 250 paired end Miseq platform as described previously (DeKosky et al., 2013). Briefly, raw sequences were first filtered for sequences containing at least half of the nucleotides at a minimum Phred quality score of 20 to obtain high quality sequences, then assembled to

obtain full length V_H and V_L sequences, respectively. Sequences represented by at least 2 reads were submitted to IMGT (the international immunogenetics information system, www.imgt.org) High V-Quest analysis (Lefranc et al., 2009) and those with productive V_H or V_L, and in-frame V(D)J junctions (CDRH3 or CDRL3) were analyzed further.

Paired V_H:V_L sequencing of antibody secreting cells from the draining lymph node

From single cell suspension of the draining lymph node, $CD138^+$ antibody secreting B cells were isolated using a mouse $CD138^+$ Plasma Cell Isolation kit (Miltenyi Biotec) (Wang et al., 2015a). The paired V_H:V_L gene repertoire was determined as described (DeKosky et al., 2013). Briefly, isolated single cells were deposited into 125 pL wells in PDMS slides together with poly(dT) magnetic beads (Invitrogen). After cell lysis *in situ*, beads with captured mRNA from the same cell were collected, washed, and emulsified. Mouse V gene specific primers were used to link endogenous V_H and V_L in the subsequent emulsion overlap extension RT-PCR. A second nested PCR was used to amplify the linked V_H:V_L pairs. Sequencing was performed by Illumina 2 x 250 paired end Miseq. The raw Miseq reads were filtered for reads where at least half of the nucleotides had a minimum Phred quality score of 20 to remove low-quality reads, and were then submitted to IMGT for V(D)J germ-line alignment (Lefranc et al., 2009). Sequence data were again filtered for productive V_H or V_L, and for in-frame V(D)J junctions (CDRH3 or CDRL3), and were then matched using their Illumina read IDs. CDRH3 nucleotide sequences were clustered to 96% identity, and CDRH3:CDRL3 pairs represented by 2 or more reads were rank-ordered. In two separate sequencing runs the complete V_H and V_L sequences encoded within the linked V_H:V_L amplicons were determined. The sequences for the V_H:V_L junction and the separate V_H and V_L sequences

were assembled to yield natively paired full length V_H and V_L sequences, as described (DeKosky et al., 2013).

Antibody expression and characterization

Consensus V_H and V_L sequences of the top 10 most abundant antibodies from the draining lymph node were synthesized using gblocks (IDT) and subcloned into pcDNA3.4 vectors (Invitrogen) as fusions with human IgG1 and kappa constant domains, respectively. After sequence verification, the heavy and light chain plasmids were mixed at 1:3 ratio, and transfected into Expi293 cells for antibody expression. Six days after transfection, the supernatant was collected and antibodies were purified using protein A chromatography (Thermo Scientific).

To verify the specificity of these antibodies, ELISAs were carried out with both ricin A chain and BSA as a control. ELISA plates were coated with 50 µl of 4 µg/ml ricin A chain or BSA at 4°C O/N. On the next day, plates were blocked with PBSB at RT for 2 hr. After that, 1:5 serially diluted antibodies with concentrations ranging from 300 nM to 3.84 pM were added to the plates and the plates were incubated at RT for 1 hr. After washing with PBST, 50 µl of 1:5000 diluted HRP-conjugated donkey anti-human antibody (Jackson ImmunoResearch, cat# 709-035-149) was added and the plates were again incubated at RT for 1hr. After washing with PBST, 50 µl of TMB substrate was added and the reaction was stopped after 10 min by adding 50 µl of 2 M H_2SO_4. The absorbance at 450 nm was read and EC_{50} was calculated. For each antibody that showed binding to ricin A chain, affinity determinations were carried out via SPR analyses, as described above for scAbs.

ACKNOWLEDGEMENTS

This work was funded by the Defense Threat Reduction Agency under grant No. HDTRA-1-13-1-0031.

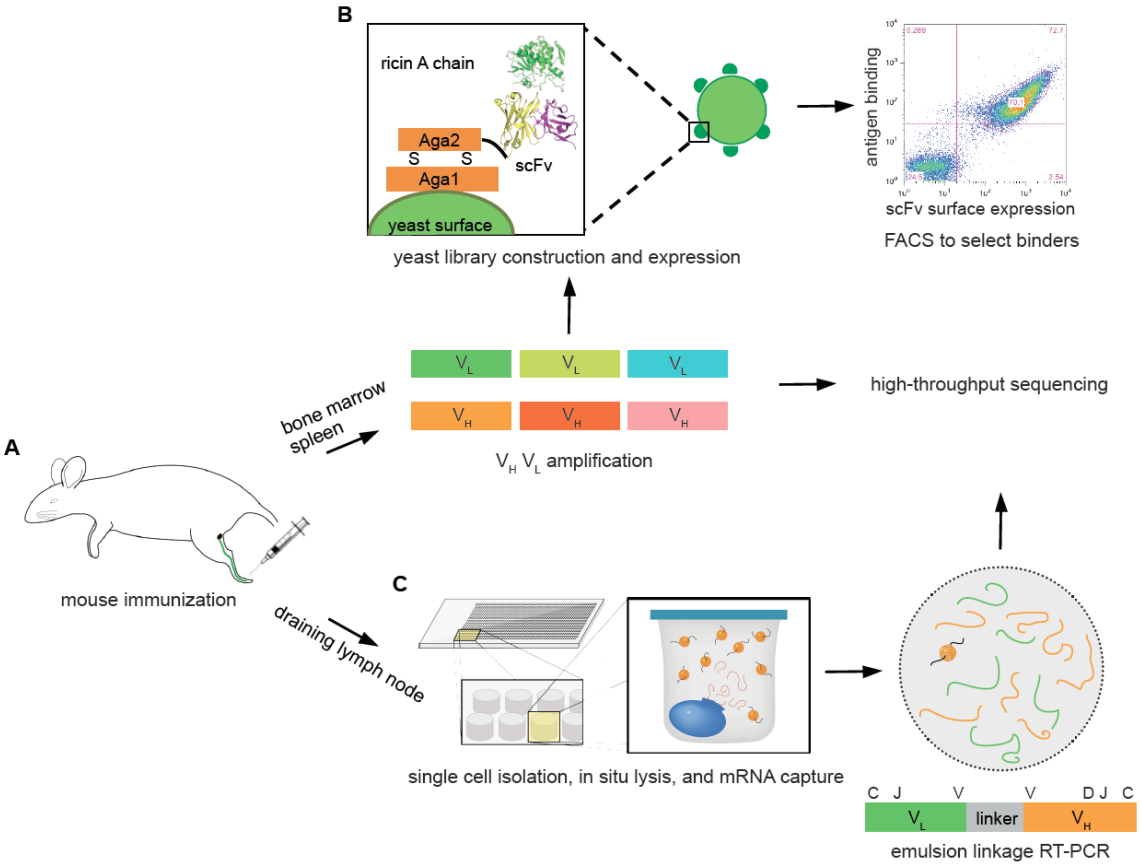

Figure 3-1: Overview of the experimental approach.

(A) Mouse is immunized at footpad with ricin A chain, and peripheral lymphoid organs, including bone marrow, spleen, and draining lymph node are isolated after three booster immunizations. (B) Total cells in bone marrow and spleen are collected, and V_H and V_L mRNA are reverse transcribed and amplified, which are used to construct scFv libraries, respectively. These are selected against ricin A chain using yeast surface display to isolate high-affinity binders. V_H and V_L cDNAs are also sequenced on an Illumina platfrom. (C) $CD138^+$ antibody secreting cells are isolated from draining lymph node, and processed through our high-throughput V_H:V_L pairing platform.

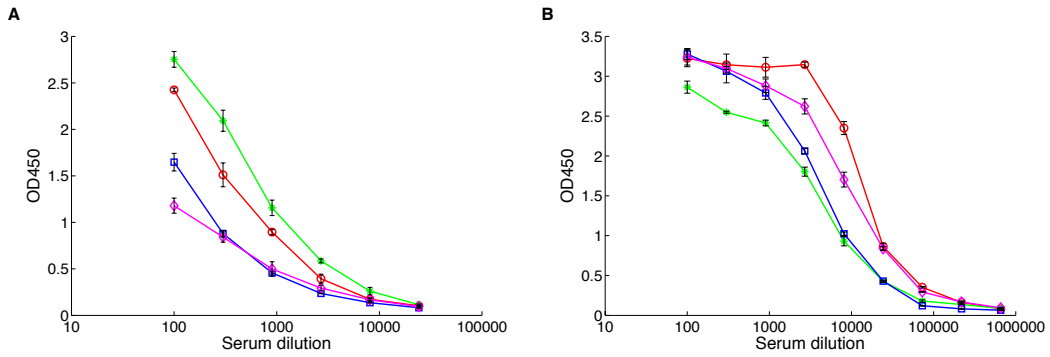

Figure 3-2: Serum titers of mice immunized with ricin A chain.

Titers after the 2nd booster immunization (A) and after the 3rd booster immunization (B) are shown. (mouse 1 (used for antibody discovery): red circle; mouse 2: green star; mouse 3: blue squaure; mouse 4: magenta diamond).

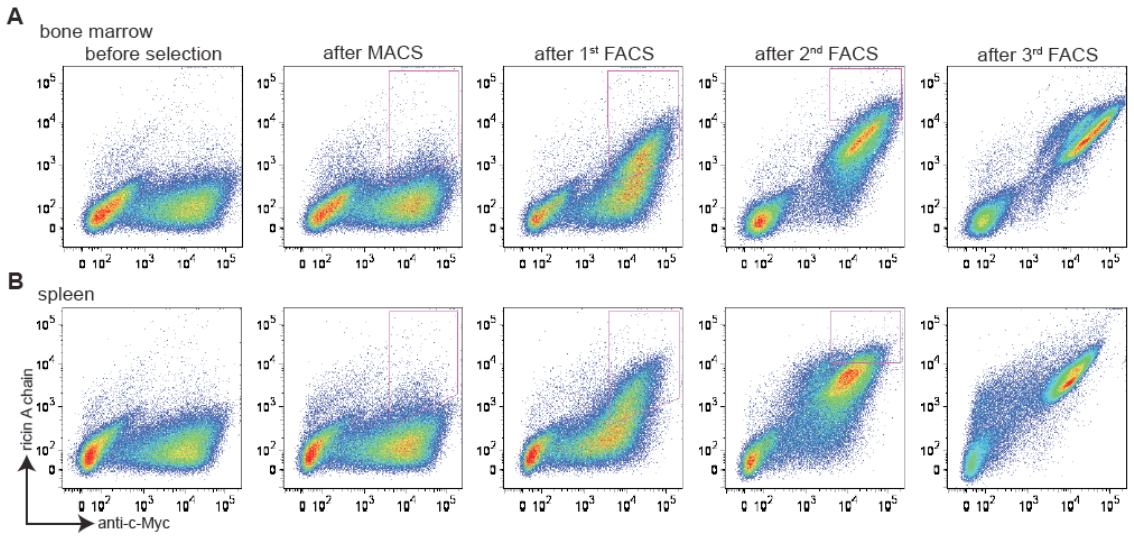

Figure 3-3: Selection of libraries constructed from bone marrow and spleen antibody repertoires.

Cells are doubly stained with chicken anti-c-Myc IgY/GaC-488 for scFv surface expression and biotinylated ricin A chain/SA-633 for antigen binding. (A) Bivariate plots are shown for bone marrow antibody library stained with 100 nM ricin A chain before selection, after MACS, after 1st FACS, after 2nd FACS, and after 3rd FACS, with x axis being surface expression, and y axis being antigen binding. (B) Bivariate plots are shown for spleen antibody library stained with 100 nM ricin A chain before and after each round of selection. For both libraries, cells in the upper-right quadrant (both express scFv on the surface and bind ricin A chain) are sought. Cells falling within strict FACS sort gates designed to ensure enrichment of clones showing increased binding to ricin A chain are collected for the next round of selection.

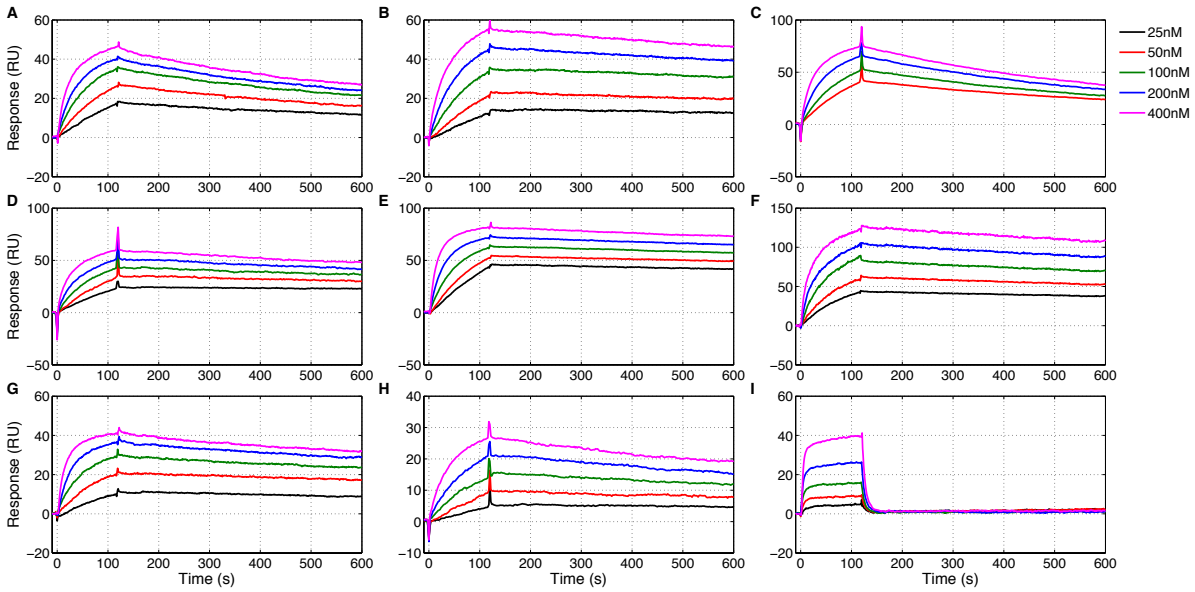

Figure 3-4: SPR sensorgrams for BM1(A), BM3(B), BM17(C), SP1(D), and SP19(E) scAbs and RAM1.2(F), RAM1.4(G), RAM1.5(H), and RAM1.10(I) IgGs.

One representative binding curve with soluble ricin A chain from three replicates is shown.

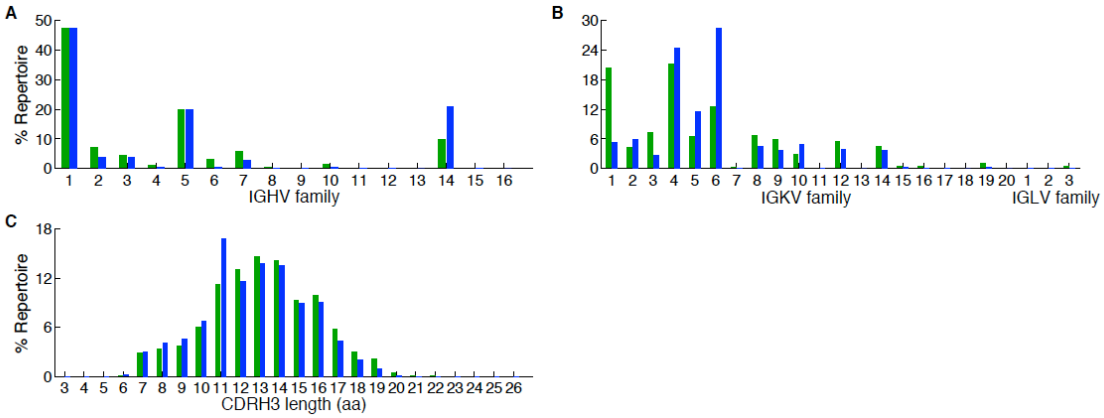

Figure 3-5: Characteristics of bone marrow and spleen antibody repertoires from the same mouse immunized with ricin A chain.

(A) Germ-line V_H gene usage in bone marrow and spleen repertoires. (B) Germ-line V_L gene usage in the two repertoires. (C) CDRH3 length distribution in the two repertoires. Green denotes bone marrow repertoire, and blue denotes spleen repertoire.

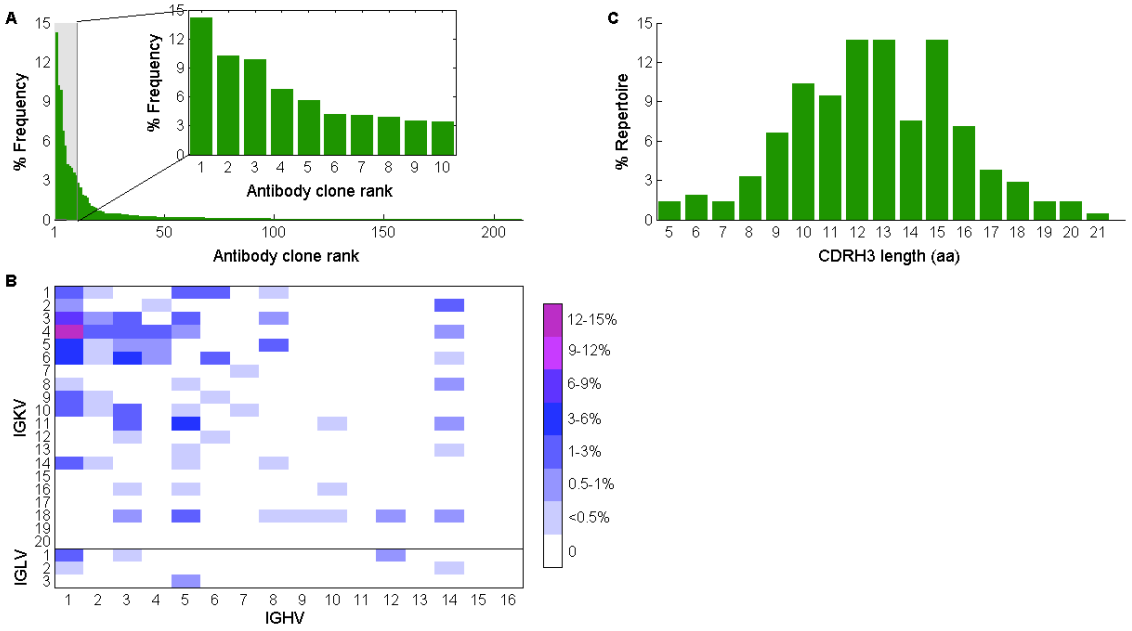

Figure 3-6: Characteristics of CD138^{+} antibody-secreting cells repertoire in draining lymph node from mouse immunized with ricin A chain.

(A) The frequency of each unique antibody clone in the repertoire, which is calculated as the percentage of its read counts in the read counts of all clones, is shown with its rank, which is ordered by the number of read counts of each unique clone. Clones with only 1 read are removed, and CDRH3 nucleotide sequences are clustered to 96% identity. Inset shows the distribution of the top 10 most abundant clones in the repertoire. (B) Germ-line V gene family usage in the same repertoire is shown. (C) CDRH3 length (calculated as amino acid length) distribution of antibodies in the repertoire is shown as the percentage of antibodies that have the denoted CDRH3 length in the whole repertoire.

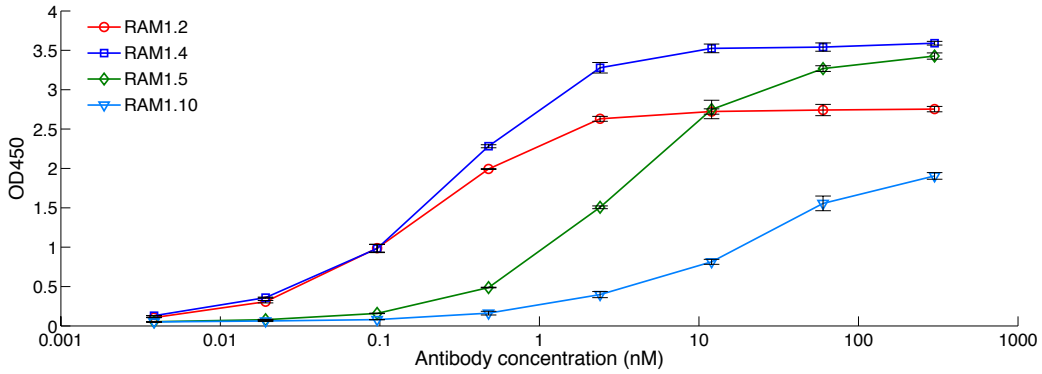

Figure 3-7: Characterization of the 4 antibodies isolated from $CD138^+$ antibody secreting cells in the draining lymph node.

Reactivity of the antibodies to ricin A chain is determined by ELISA. EC_{50} for the 4 antibodies are shown in Table 3-5. Error bar represents standard error mean of three technical replicates.

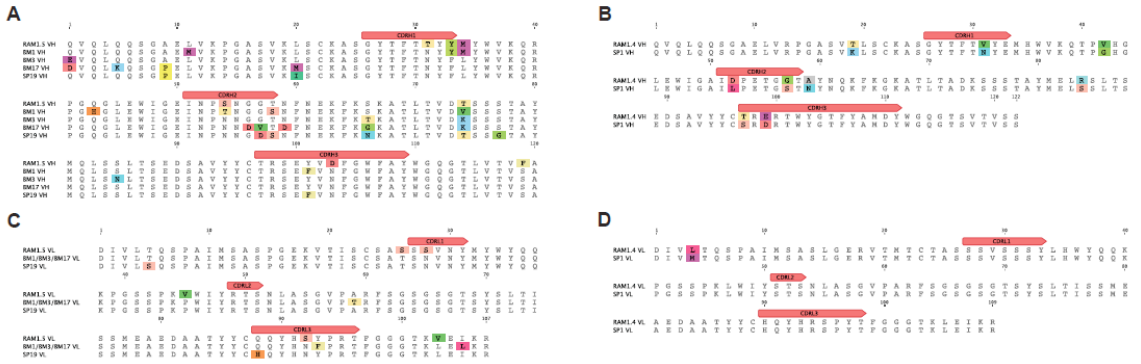

Figure 3-8: Sequence comparison between yeast surface display and next-generation sequencing discovered antibodies.

(A), (B) V_H alignments of BM1, BM3, BM17, SP19 and RAM1.5 and of SP1 and RAM1.4 showed mutations through the complete V_H sequences, indicating they are somatic variants. (C), (D) V_L alignments of BM1/BM3/BM17, SP19 and RAM1.5 and of SP1 and RAM1.4 also showed mutations. Mutations are indicated by different colors.

Antibody	Source	CDRH3: CDRL3	Gene usage	V_H:V_L abundance (percentage and rank)
BM1	bone marrow	CTRSEFVNFGWFAYW: CQQYHNFPRTF	IGHV1-IGHD2-IGHJ3: IGKV4-IGKJ1	2.6%, 3rd: 3.1%, 3rd
BM3	bone marrow	CTRSEYVNFGWFAYW: CQQYHNFPRTF	IGHV1-IGHD2-IGHJ3: IGKV4-IGKJ1	5.1%, 2nd: 3.1%, 3rd
BM17	bone marrow	CTRSEYVNFGWFAYW: CQQYHNFPRTF	IGHV1-IGHD2-IGHJ3: IGKV4-IGKJ1	1.2%, 5th: 3.1%, 3rd
SP1	spleen	CSRDRTWYGTFYAMDYW: CHQYHRSPYTF	IGHV1-IGHD2-IGHJ4: IGKV4-IGKJ2	0.9%,5th: 2.8%,3rd
SP19	spleen	CTRSEFVNFGWFAYW: CHQYHNYPRTF	IGHV1-IGHD2-IGHJ3: IGKV4-IGKJ1	5.3%,2nd: 1.2%,4th

Table 3-1: List of the antibodies isolated from the bone marrow and spleen combinatorial libraries by yeast display and abundance of their V_H and V_L sequences in the corresponding repertoires assessed by next-generation sequencing.

Note that BM3 and BM17 have the same CDRH3 sequences, but differ in sequence outside of CDRH3 (Fig. 3-8).

Antibody	K_d (nM, determined by yeast display titration)	Binding kinetics determined by SPR		
		k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	K_d (nM)
BM1	10.3	$(2.09\pm0.04)\times 10^{5}$	$(9.7\pm0.52)\times10^{-4}$	4.64±0.33
BM3	2.76	$(1.13\pm0.03)\times10^{5}$	$(3.41\pm0.23)\times10^{-4}$	3.03±0.12
BM17	2.96	$(2.5\pm0.13)\times10^{5}$	$(1.36\pm0.006)\times10^{-3}$	5.44±0.27
SP1	4.25	$(2.28\pm0.07)\times10^{5}$	$(3.72\pm0.06)\times10^{-4}$	1.64±0.08
SP19	8.22	$(3.91\pm0.32)\times10^{5}$	$(2.25\pm0.04)\times10^{-4}$	0.58±0.05

Table 3-2: The binding kinetics of the antibodies isolated from bone marrow and spleen as measured by yeast surface titration and surface plasmon resonance.

Tissue	Region	Total sequence reads	unique productive gene	unique clonotypes
bone marrow	V_H	1078639	849735	8735
	V_L	542905	385398	
spleen	V_H	690554	370357	7057
	V_L	696061	410954	

Table 3-3: Next-generation sequencing statistics of bone marrow and spleen repertoires.

Antibody	Rank in the repertoire	CDRH3:CDRL3	Gene usage
RAM1.2	2	CARPTLLYGSSPCFDYW: CQQWSSSPTF	IGHV1-IGHJ2:IGKV4-IGKJ4
RAM1.4	4	CTRERTWYGTFYAMDYW: CHQYHRSPYTF	IGHV1-IGHD2-IGHJ4:IGKV4-IGKJ2
RAM1.5	5	CTRSEYVDFGWFAYW: CQQYHSYPRTF	IGHV1-IGHD2-IGHJ3:IGKV4-IGKJ1
RAM1.10	10	CARSRDYDGYGDYW: CQQSNRWPLTF	IGHV1-IGHD2-IGHJ2:IGKV5-IGKJ5

Table 3-4: List of the antibodies isolated from CD138^{+} antibody-secreting cells in the draining lymph node.

Antibody	EC_{50} (nM,determined by ELISA)	SPR binding kinetics		
		k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	K_d (nM)
RAM1.2	0.186	$(3.88\pm0.36)x10^{5}$	$(3.73\pm0.13)x10^{-4}$	0.97±0.11
RAM1.4	0.284	$(1.28\pm0.12)x10^{5}$	$(4.93\pm0.11)x10^{-4}$	3.89±0.35
RAM1.5	3.21	$(9.84\pm0.51)x10^{4}$	$(6.01\pm0.12)x10^{-4}$	6.12±0.19
RAM1.10	20.9	$(5.46\pm2.4)x10^{4}$	0.149±0.01	$(3.12\pm1.36)x10^{3}$

Table 3-5: The binding kinetics of the antibodies isolated from the draining lymph node as measured by ELISA and surface plasmon resonance.

Chapter 4: High-throughput mining of native antibody repertoires

In chapters 2 and 3, I have described the development and adaptation of high-throughput V_H:V_L pairing and sequencing technology for antibody discovery without screening, and then compared the antibodies obtained from mining the V_H:V_L paired repertoire with those obtained by yeast surface display. Here, I describe a novel antibody repertoire analysis and antibody discovery platform that is based on screening libraries of natively paired V_H:V_L sequences coupled with yeast display. This technology holds promise for interrogating the antibody specificity of the human repertoire isolated from peripheral B cells and is being used for the isolation of antibodies to influenza, Ebola (following vaccination) and to HIV-1 infection. This platform also enables epitope-specific antibody discovery and affinity binning of the whole repertoire. The platform should be widely applicable for antibody repertoire analysis and antibody discovery for other pathogens.

INTRODUCTION

The high-throughput analysis and "mining" of the native antibody repertoires for isolation of antibodies that bind to desired antigens are yet to be demonstrated. Existing technologies for sequencing and expressing antibodies from B cells all have major limitations: single B cell cloning suffers from very low throughput (usually at most a few hundred antibodies are expressed and characterized) (Georgiou et al., 2014b; Wilson and Andrews, 2012), while existing combinatorial library screening technologies rely on the screening of V_H or V_L libraries, and thus the native V_H:V_L pairing information found in B cells is not maintained (Bradbury et al., 2011). On the other hand, although high-throughput V_H:V_L pairing and sequencing can generate sequence information about most native antibodies in the repertoire, only a very small number (< 100) of those antibodies

can be produced and assayed as a result of the cost for gene synthesis, and the extensive workload for cloning and expression (DeKosky et al., 2013; Georgiou et al., 2014b; Wang et al., 2015a, 2016). Furthermore, it is not a trivial task to select which antibodies to make (DeKosky et al., 2013; Wang et al., 2015a, 2016). So far, the three demonstrations that use high-throughput V_H:V_L paring for antibody discovery have relied on proteomic identification of the antibodies in the serum (DeKosky et al., 2013), or direct selection of highly expanded antibodies after immunization without screening (Wang et al., 2015a, 2016). Hence, a technology that enables high-throughput screening of native antibody repertoires without any special requirements is highly desirable. Furthermore, such a technology should enable linkage of antibody sequence-function relationships at the single clone level.

RESULTS

Implementation of novel yeast display platform for native antibody repertoire exploration

In order to rapidly screen large antibody repertoires (libraries), one of the *in vitro* display methods has to be used, as recombinant expression of all antibodies in the repertoire in a soluble format with concomitant characterization is unrealistic. In contrast, after library transformation, every cell in the antibody display library expresses a single antibody from the repertoire, and this therefore allows detailed examination of every antibody.

The first consideration for selection of a screening technology is the host for antibody library construction. Of all the display technologies, phage and *E. coli* display use *E. coli* as the host for antibody expression; yeast display uses yeast; while mammalian display uses mammalian cells (for example, HEK293). However, as antibody

library size is usually very small in mammalian cells (less than 10^5), I chose yeast display, as antibody library size can be as large as 10^{10} (similar to in *E. coli*). Furthermore, the similar protein production, folding, and secretion apparatus, and quality control mechanisms of yeast compared to those of mammalian cells make yeast a better host than *E. coli*. The use of FACS in yeast display also enables quantitative control of screening and normalization of affinity by surface expression level. Previous studies have also shown that yeast display, but not phage display, can recover native antibodies in the repertoire (Saggy et al., 2012; Wang et al., 2016), and compared to phage display, yeast display can identify more antibodies (Bowley et al., 2007).

However, as high-throughput V_H:V_L pairing produces physically linked amplicons in which the 5' end of V_H is adjacent to the 5' end to V_L, the V_H:V_L paired repertoire can not be simply transferred into the display vector, in which the 3' end of V_H is adjacent to the 5' end of V_L (scFv). Hence, a new yeast display platform was designed.

Design of the new platform

In order to express the V_H:V_L pairs together, a galactose inducible bidirectional promoter was used to drive expression of both chains. C_{H1} (IgG1 isotype) and C_L were fused at the C-terminus of V_H and V_L, respectively. To facilitate library construction, 4 restriction enzyme sites that are very rarely found in immunoglobulin variable sequences were used (Fig. 4-1). As shown in Fig. 4-1, upon induction, the light chain was covalently linked to the heavy chain through a C-terminal disulfide bond, and the Fab fragment was displayed on the yeast cell surface as a N-terminal fusion of Aga2, which was linked to Aga1 by 2 disulfide bonds. C-terminal c-Myc and Flag tags were used to monitor expression of heavy and light chains, respectively. As antibodies have two different kinds

of light chains (κ and λ), whole repertoires were separated into two parts, and cloned into the κ and λ platforms (which used C_κ and $C_{\lambda 2}$ domains), respectively.

Compared to the scFv used in the original yeast display, the Fab fragment, which only requires fusion with the Fc domain, is much closer to a native IgG. Once desired clones are identified, they can be expressed as IgGs very easily without reformatting.

In order to clone intact V_H:V_L pairs into the platform, NotI (which cuts at the C-terminus of V_L) and AscI (which cuts at the C-terminus of V_H) were used first to digest the pairs, which were then ligated into similarly digested vectors. This ligation produces vectors with intact V_H:V_L pairs inserted. These vectors were then digested with NcoI (which cuts at the N-terminus of V_L) and NheI (which cuts at the N-terminus of V_H), and ligated with a similarly digested bidirectional promoter fragment. The vectors with intact V_H:V_L pairs were eventually transformed into yeast, and could be screened by FACS after induction.

Verification of the new platform

Initially, 13 anti-influzenza hemagglutinin antibodies (Lee, *et al*. in press) were tested in the new platform. Only 7 of them expressed well on the surface and showed binding to hemagglutinin. We hypothesized that inferior protein folding and secretion in yeast compared to human cells accounted for the poor expression of some antibodies. Moreover, the single disulfide bond between heavy and light chains may make Fab assembly inefficient in yeast compared to humans, and the lack of light chain expression for several antibodies on the yeast cell surface further suggested this. Thus, two methods were used to improve folding and assembly of the Fab.

First, instead of EBY100, the strain that is commonly used for yeast display, we used the yeast strain AWY101 that overexpresses protein disulfide-isomerase (PDI)

(Wentz and Shusta, 2007). PDI is a protein folding chaperone that plays a critical role in antibody expression and folding (Feige et al., 2010). In AWY101, the overexpression of PDI is driven by the GPD promoter (a very strong constitutive promoter in yeast), and the whole cassette is integrated into the genome. We found that this innovation improved expression and binding of 3 antibodies that didn't express well initially.

Second, we fused a pair of leucine-zipper dimerization tags at the C-terminus of the heavy and light chains, respectively. At the C-terminus of the heavy chain, we put a basic leucine-zipper tag, and at the C-terminus of the light chain, we put an acidic leucine-zipper tag. The use of leucine-zipper tags should force better assembly of the heavy and light chains, and the strong interaction of basic and acidic leucine-zipper tags should promote heterodimerization of the heavy and light chains instead of homodimerization of heavy chains or light chains. The use of leucine-zipper tags improved expression and binding of the other 3 antibodies.

We then combined the two methods. As shown in Fig. 4-2a, all of the tested antibodies, including 13 anti-influenza hemagglutinin antibodies (2 shown), 3 anti-Ebola GP antibodies (1 shown) (Corti et al., 2016; Lee et al., 2008), and 3 anti-HIV-1 bnAbs (1 shown) (Doria-Rose et al., 2014; Kong et al., 2016; Wu et al., 2010), showed good surface expression and binding to their cognate proteins hemagglutinin, GP, and fusion peptide probes, respectively. The antibodies VRC01 and VRC26 were not tested for antigen binding due to the unavailability of the HIV-1 envelope trimer.

In order to make sure that leucine-zipper mediated enhancement of Fab assembly didn't interfere with binding, we made 3 chimeric antibodies that were composed of their native heavy chains and non-cognate light chains of other antibodies. As shown in Fig. 4-2b, although the chimeric antibodies were still expressed on the surface, they didn't bind to cognate antigens. The loss of binding for non-native antibodies, together with binding

to cognate but not to non-cogate antigens by native antibodies, indicated that leucine-zipper mediated enhancement of the Fab assembly didn't induce nonspecific binding.

We also dissected the effects of combining chaperone overexpression and leucine-zipper mediated assembly enhancement. As shown in Fig. 4-2c, both chaperone overexpression and leucine-zipper enhanced assembly improved expression and binding, and had additive effects in which the largest enhancement was seen when both of them were present.

While this work was going on, Ojima-Kato *et al*. also demonstrated the effect of leucine-zipper fusion on Fab expression and antigen binding in *E. coli* (Ojima-Kato et al., 2016). This work serves as an independent verification of our incorporation of leucine-zippers into Fab assembly.

Exploration of native antibody repertoires

Next, we analyzed 4 native antibody repertoires, including one of total B cells from PBMCs isolated 7 days after TIV01 vaccination (Lee et. al, in press), one of leukapheresis B cells collected 9 months after TIV01 vaccination, one of plasmablasts isolated 6 days after Ebola vaccine boost, and one of memory B cells collected 2 years after HIV-1 infection (Kong et al., 2016). The V_H:V_L paired libraries were separately amplified to construct V_H:V_κ and V_H:V_λ sub-libraries, which were then cloned into the κ and λ platforms, respectively. These libraries were then stained with fluorescently-labeled hemagglutinin, GP, and fusion peptide probes, and screened by FACS. Screening stringency was increased round-by-round by decreasing antigen concentrations: for influenza antibody repertoire screening, hemagglutinin was used at 1 μM, 200 nM, and then 40 nM concentrations; for the Ebola virus antibody repertoire, GP was used at 1 μM,

100 nM, and then 10 nM concentrations; and for the HIV-1 antibody repertoire, the fusion peptide probe was used at 1 μM, 100 nM, and then 10 nM concentrations.

As shown in Fig. 4-3, in the HIV-1 V_H:V_κ library, we saw a small population of cells that showed binding to the fusion peptide probe after 1 round of screening. After the 2nd round of screening, significant enrichment of cells showing binding was observed. Similar enrichment of binding cells was also seen in the other repertoires after 3 rounds of screening.

Toward functional screening: epitope-specific selection

Another major limitation for most of the current antibody library screening/antibody discovery technologies is no function other than binding can be selected (Wilson and Andrews, 2012). The lack of functional selection methods in part results from a lack of simple and straightforward ways to link antibody function (for example, neutralization) to sequence. We therefore have also worked towards functional screening of antibody libraries.

We have started by performing epitope-specific screening. We hypothesized that there is a higher chance to identify antibodies with desired functions (such as neutralization) from the pool of antibodies that target specific epitopes. This is supported by the fact that a large number of functional antibodies target specific, subdominant epitopes, especially for complex pathogens that have many immunodominant epitopes (such as Ebola and HIV-1 virus) (Corti and Lanzavecchia, 2013).

We devised a strategy to fish out antibodies that target specific epitopes. We used the cloned anti-Ebola virus GP antibodies in a proof-of-principle study. In this set of antibodies, mAb 114, mAb 166, and c13c6 recognize similar epitopes (on the top of GP), while KZ52 and mAb 100 recognize similar epitopes (at the base of GP) (Corti et al.,

2016; Lee et al., 2008). As expected (Fig. 4-4), mAb 114 binding to GP is blocked by itself, but not by KZ52, and c13c6 binding to GP is blocked by mAb166 and mAb 114, but not by KZ52, while KZ52 binding to GP is blocked by mAb 100 and itself, but not by mAb 114. This demonstrated that epitope-specific antibodies could be selected by first blocking undesired epitopes on the antigens.

We screened the Ebola virus antibody repertoire using the epitope-specific approach. Fluorescently-labeled GP was blocked with KZ52, and then used for screening. We observed enrichment of cells binding to non-KZ52 epitopes after the 2^{nd} round of screening.

DISCUSSION

Here, we developed a novel platform that enables high-throughput analysis of native antibody repertoires, and used it to explore antibody repertoires elicited after influenza and Ebola virus vaccination, and after HIV-1 infection. Compared to other antibody repertoire characterization and discovery methods, the new platform has several advantages: 1) it allows characterization of almost every antibody in the repertoire, and these antibodies have preserved their native V_H:V_L pairings; 2) use of FACS provides a high-throughput, quantitative way to characterize antibodies; and 3) NGS can still be used for sequence determination.

Furthermore, we achieved epitope-specific antibody discovery with this platform by using antigens that were blocked with antibodies that targeted undesired epitopes. Starting from a pool of epitope-specific antibodies, we hypothesize there should be a higher chance to identify antibodies with desired functions (such as neutralization).

Given these advantages, we anticipate the wide application of this new platform for antibody repertoire analysis and antibody discovery.

MATERIALS AND METHODS

Cell line and media

The yeast strain AWY101 (MATα *AGA1::GAL1-AGA1::URA3 PDI1::GAPDH-PDI1::LEU2 ura3-52 trp1 leu2Δ1 his3Δ200 pep4::HIS3 prb1Δ1.6R can1 GAL)* was used for library construction and screening. Yeast cells were maintained in YPD medium (20 g/l dextrose, 20 g/l peptone, and 10 g/l yeast extract); after library transformation, they were maintained in SDCAA medium (20 g/l dextrose, 6.7 g/l yeast nitrogen base, 5 g/l casamino acids, 8.56 g/l $NaH_2PO_4.H_2O$, and 10.2 g/l $Na_2HPO_4.7H_2O$). SGDCAA (SGCAA with 2 g/l dextrose) medium was used for library induction.

Antigens and antibodies

Hemagglutinins (A/California/07/2009, A/Solomon Islands/3/2006, A/Wisconsin/67/2005, and B/Malaysia/2508/2004) were provided by Dr. Stephen Harrison at Harvard University. Full-length and mucin-like domain deleted Ebola virus GP were provided by Dr. Nancy Sullivan at NIH. HIV-1 fusion peptide probe and knock-out probe were provided by Dr. John Mascola at NIH. Anti-flag FITC antibody was purchased from Sigma-Aldrich (cat# F4049). Hemagglutinins were biotinylated using an EZ-Link Sulfo-NHS-LC-Biotin kit (Thermo Scientific). Chicken anti-c-Myc IgY, Alexa Fluor 633-goat anti-chicken IgG (GaC-633), streptavidin-PE (SA-PE) and streptavidin-v450 (SA-450) were obtained from Invitrogen and BD Biosciences (cat# A-21281, A-21103, S866, and 560797, respectively). Anti-biotin and streptavidin microbeads were purchased from Miltenyi Biotec (cat# 130-090-485 and 130-048-101).

Yeast library construction for influenza and Ebola vaccinees and HIV-1 elite controller

V_H:V_L pairing was done as previously described (DeKosky et al., 2015). For library construction, 10 ng of V_H:V_L paired DNA was used as template for PCR to introduce NotI and AscI sites with primers that primes at 3' of both chains. The product was digested, ligated with similarly digested vector, and transformed into *E. coli*. The libraries were minipreped, digested with NcoI and NheI, and ligated with similarly digested PCR product of bidirectional promoter. After transformation into *E. coli*, the libraries were minipreped and transformed into AWY101. The libraries were cultured in SDCAA medium at 30^{o}C for 2 d.

Library screening

Cells were induced in SGDCAA medium at 20^{o}C for 2 d. For the 1st round of selection, 10^9 cells (10-fold coverage relative to the number of transformants) were incubated with 1 μM biotinylated hemagglutinins at RT for 40 min and washed thoroughly with PBS+0.5% BSA+2 mM EDTA (PBSA) to remove any unbound hemagglutinins. Cells were then mixed with 400 μl of anti-biotin microbeads and incubated at 4^{o}C for 15 min, and washed again. Cells that bound hemagglutinins were selected using MidiMACS system (Miltenyi Biotec), and recovered in SDCAA medium at 30^{o}C. For the 2nd round of selection, 10^8 cells in 1ml PBSA were labeled with 1 μM biotinylated hemagglutinins and 4 μg/ml chicken anti-c-Myc IgY and anti-flag FITC at RT for 40 min. After washing with PBSA, cells were incubated with 5 μg/ml SA-450 and 5 μg/ml GaC-633 at 4^{o}C for 15 min. After washing, cells were resuspended in PBSA buffer and sorted using FACS Aria (BD Bioscience). The subsequent rounds of selection were performed in a similar way, except that the concentration of hemagglutinins used was decreased to 200 nM and 40 nM, respectively.

Ebola and HIV-1 antibody libraries were screened similarly.

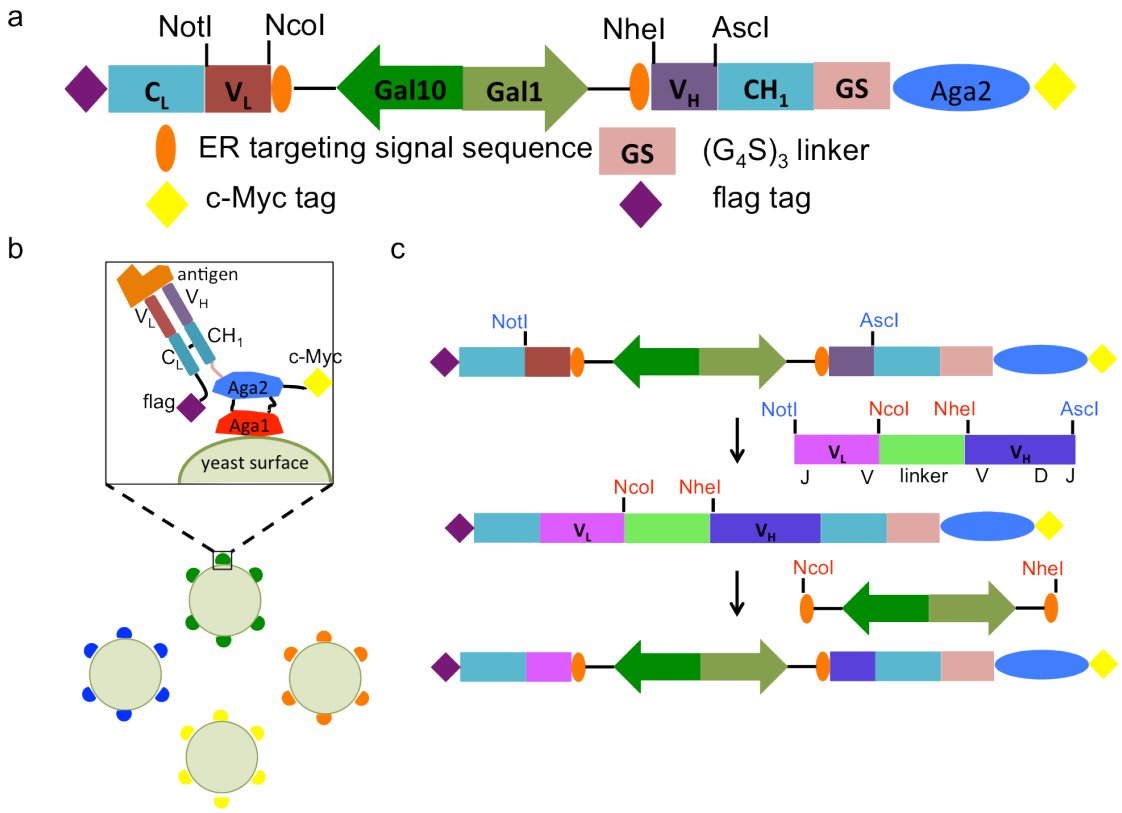

Figure 4-1: Design of the new platform.

(a) Structure of the new vector. (b) Fab fragments displayed on yeast cell surface. (c) Cloning strategy for paired library construction. In the first step, NotI and AscI digested V_H:V_L pairs is ligated with similarly digested vectors. In the second step, NcoI and NheI digested promoter is ligated with similarly digested V_H:V_L vectors.

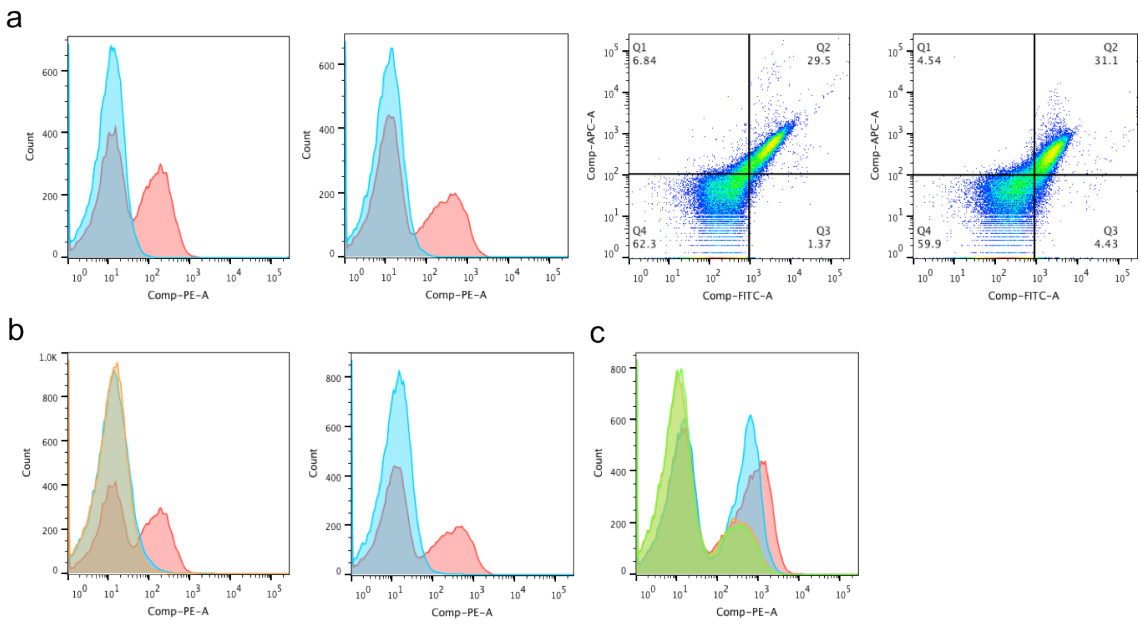

Figure 4-2: Verification of the new platform.

(a) Binding of hemagglutinin (left two panels), Ebola GP (the third panel), and HIV-1 fusion peptide (rightmost panel) antibodies to their cognate antigens. In left two panels, binding to cognate hemagglutinin (A/California/07/2009 and B/Malaysia/2508/2004, respectively) is shown in red, and to noncogate hemagglutinin (B/Malaysia/2508/2004 and A/California/07/2009, respectively) is shown in blue. In right two panels, binding is monitored in APC channel, and light chain expression is monitored in FITC channel. (b) Binding of native and chimeric hemagglutinin antibodies to cognate antigens. Native antibody is shown in red, and chimeric antibodies are shown in blue and orange. (c) Effects of chaperone overexpression and leucine-zipper fusion on binding of a hemagglutinin antibody to cognate antigen. Red, blue, orange, and green denotes binding with both, with leucin-zipper fusion only, with chaperone overexpression only, and without both, respectively.

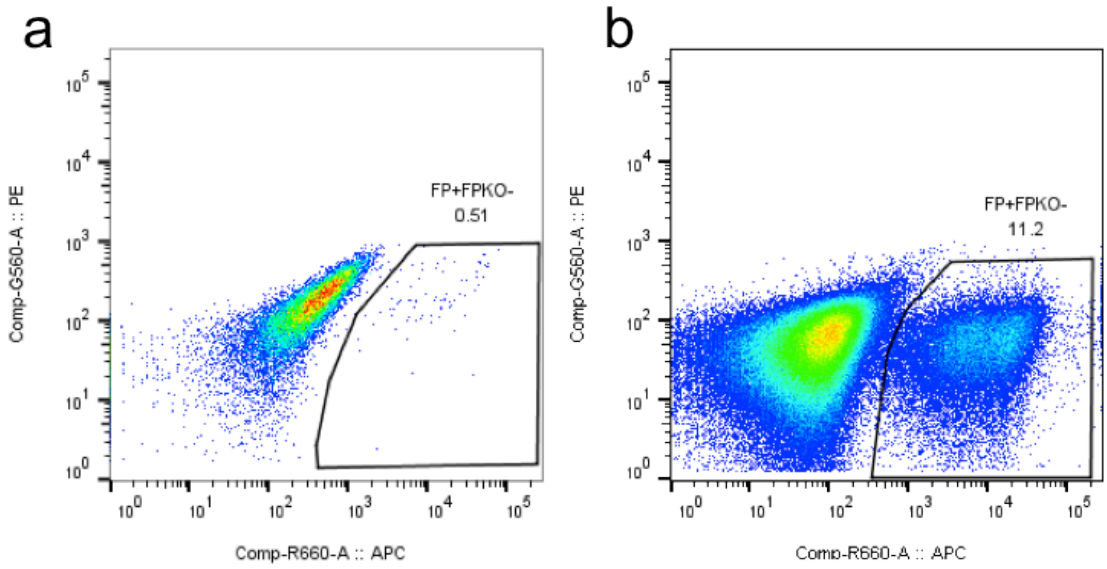

Figure 4-3: HIV-1 elite controller library screening.

(a) V_H:V_κ library after 1st round of screening with fusion peptide probe. (b) V_H:V_κ library after 2nd round of screening.

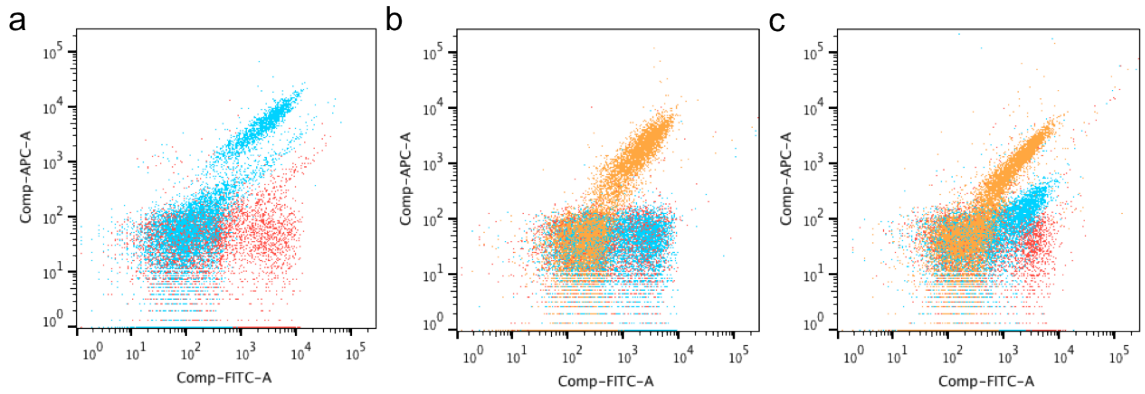

Figure 4-4: Epitope specific staining of antibodies.

(a) Ebola GP antibody 114 was stained with 114 blocked GP (red), and KZ52 blocked GP (blue). (b) c13c6 was stained with 114 blocked GP (red), 166 blocked GP (blue), and KZ52 blocked GP (orange). (c) KZ52 was stained with KZ52 blocked GP (red), 100 blocked GP (blue), and 114 blocked GP (orange).

Chapter 5: Conclusion

The immune system plays a very important role in human health, as it can not only protect from astronomical number of pathogens, but also keep from self-destruction (autoimmunity). Adaptive immunity is one of the two branches of the immune system, which is the basis of the diversity of recognition and protection. Antibodies are an important class of mediators in adaptive immunity, which not only neutralize pathogens, but also destroy them by recruiting innate immune effectors (for example, NK cells, macrophages, and complement). As a result of their high specificity, affinity, and stability, antibodies are widely used in basic research and clinics. Moreover, characterization of the antibody repertoires elicited after vaccination or infection is also of great importance, which can provide insights about how to design more effective vaccines. Thus, antibody discovery and repertoire analysis technologies are critical.

The work described here presents high-throughput technologies for antibody discovery and repertoire analysis, and protein engineering for rationally manipulating adaptive immunity. Compared to existing technologies, these novel approaches have multiple advantages and shed light on further efforts for improvement, which will be detailed below. The applications and implications of these approaches will also be discussed.

Advantages of Native Antibody Library Screening Coupled with NGS over Other Existing Antibody Repertoire Mining and Discovery Technologies

An important consideration in antibody repertoire analysis is throughput, as analyzing all sequences in a large repertoire one by one is laborious and expensive. Another important consideration is intactness of native repertoire (native V_H:V_L pairing), which is lost in most cases. Single-cell RT-PCR can keep the native repertoire intact, but

its throughput is relatively low, which only permits interrogation of a fraction of the complete repertoire. On the other hand, NGS has very high throughput, but it disrupts the native repertoire during library preparation. Compared to the two technologies, native antibody library screening by yeast display coupled with NGS is like an ultra fast single-cell RT-PCR, which combines the advantages of both methods. First, native pairing of V_H:V_L is achieved in high-throughput emulsion based pairing, and conserved during antibody library construction and screening, which outcompetes normal NGS. Second, the use of yeast display provides a means to rapidly screen the library quantitatively (with FACS). Evolutionarily, yeast is closer to human than *E. coli*, which serves as a better host for antibody expression as a result of its protein folding and secretory, and quality control pathways. Previous studies have also shown that yeast display can recover native antibodies from the repertoire (Wang et al., 2016), and allow identification of more diverse antibodies than phage display (Bowley et al., 2007). Moreover, yeast is more manipulative than human cell lines (for example, HEK293), which allows construction of large enough library for repertoire analysis (sizes of the libraries described here are usually between 10^8 and 10^9, which is more than enough for characterization of common antibody repertoires, and 10^{10} is achievable with some efforts). The use of FACS for library screening is also advantageous, as FACS is robust and fast, which allows screening of 10^8 library members in several hours. Moreover, multiple compatible colors used in FACS nowadays allows fine tune of screening, which is critical for some applications. In regards to throughput of screening, the platform also outcompetes both methods. Third, both NGS and Sanger sequencing can be used in the platform. When antibody repertoire analysis is needed, NGS can be used to probe all the antibodies in the library after each round of screening, while when hits are needed (for antibody discovery), Sanger sequencing can be used to sequence the highly enriched clones.

Compared to existing antibody discovery technologies, the platform is also advantageous. The increase in throughput makes the platform superior than B cell immortalization, and the use of yeast as antibody expression host makes it better than phage display, while Fab used in screening avoids the trouble of reformatting scFv into IgG. An even larger improvement is the preservation of native V_H:V_L pairing in the library. Non-cognate V_H:V_L pairing has been shown to result in unfavorable properties in engineered antibodies (for example, instability) (Jayaram et al., 2012; Ponsel et al., 2011; Tiller et al., 2013), which requires additional efforts to fix. Thus, the platform should allow identification of promising antibodies with superior biophysical properties.

One limitation of the platform is it doesn't enable functional selection besides binding. Functional selection is highly desirable, as many antibodies do not neutralize or protect besides just binding. However, functional selection is very difficult to achieve, as there is no simple method to link antibody sequences with their functions. The only demonstration of functional selection is B cell immortalization, in which antibodies secreted by single B cells can be functionally screened (for example, neutralization). On the other hand, functional selection by the platform can be achieved by smart design of the selection. For example, when isolating antibodies against HIV-1 or influenza virus, specific antigens that only display broadly neutralizing epitopes or intact antigens that are blocked with antibodies targeting undesired epitopes can be used, which allows identification of broadly neutralizing epitope specific antibodies. There is a higher chance that bnAbs can be isolated from these broadly neutralizing epitope specific antibodies.

FUTURE APPLICATION AND DIVERSIFICATION OF THE PLATFORM

In the work described here, the platform is used to analyze antibody repertoires elicited after Ebola and flu vaccination and HIV-1 infection. I foresee application of the

platform for analysis of more antibody repertoires and discovery of promising therapeutic antibodies. Specifically, a naïve human V_H:V_L natively paired antibody library can be constructed, which can be used to isolate antibodies against various pathogens.

Moreover, as a result of the similarity between TCR and antibody, the platform can be applied for TCR repertoire analysis and discovery. T cell is indispensible in adaptive immunity as it provides help for B cell and executes cytotoxicity. Thus, a high-throughput technology for analyzing native TCR repertoire is invaluable. Moreover, as there are more and more needs for cancer neoantigen specific TCRs to be used in TCR transgenic T cells for cancer immunotherapy, the platform can be used to isolate these TCRs. Similar to naïve antibody library, a naïve human V_α:V_β natively paired TCR library can be constructed for antigen specific TCR mining, and as T cells don't undergo somatic hypermutation, these isolated TCRs can be directly used without further affinity maturation.

A potential diversification of the platform is making it dual mode: selection and secretion. By incorporating an amber codon (TAG) at the end of heavy chain and introducing a suppressor tRNA with its cognate aminoacyl-tRNA synthetase, the platform can be switched between two modes. In selection mode, a noncanonical amino acid is supplied, which will be incorporated at the TAG codon by tRNA and its synthetase. This will anchor the Fab on the surface of yeast for selection. In secretion mode, without the noncanonical amino acid, translation will stop there, which makes soluble Fab for characterization.

References

de Alwis, R., Smith, S. a., Olivarez, N.P., Messer, W.B., Huynh, J.P., Wahala, W.M.P.B., White, L.J., Diamond, M.S., Baric, R.S., Crowe, J.E., et al. (2012). Identification of human neutralizing antibodies that bind to complex epitopes on dengue virions. Proc. Natl. Acad. Sci. *109*, 7439–7444.

Amanna, I.J., Carlson, N.E., and Slifka, M.K. (2007). Duration of humoral immunity to common viral and vaccine antigens. N. Engl. J. Med. *357*, 1903–1915.

Angelini, A., Chen, T.F., Picciotto, S. De, Yang, N.J., Tzeng, A., Santos, M.S., Deventer, J.A. Van, Traxlmayr, M.W., and Wittrup, K.D. (2015). Protein Engineering and Selection Using Yeast Surface Display. In Yeast Surface Display: Methods, Protocols, and Applications, Methods in Molecular Biology, B. Liu, ed. (New York, NY: Springer New York), pp. 3–36.

Arnaout, R., Lee, W., Cahill, P., Honan, T., Sparrow, T., Weiand, M., Nusbaum, C., Rajewsky, K., and Koralov, S.B. (2011). High-resolution description of antibody heavy-chain repertoires in humans. PLoS One *6*.

Audi, J., Belson, M., Patel, M., Schier, J., and Osterloh, J. (2005). Ricin Poisoning A Comprehensive Review. J. Am. Med. Assoc. *294*, 2342–2351.

Baize, S., Pannetier, D., Oestereich, L., Rieger, T., Koivogui, L., Magassouba, N., Soropogui, B., Sow, M.S., Keïta, S., De Clerck, H., et al. (2014). Emergence of Zaire Ebola Virus Disease in Guinea. N. Engl. J. Med. *371*, 1418–1425.

Becker, P.D., Legrand, N., van Geelen, C.M.M., Noerder, M., Huntington, N.D., Lim, A., Yasuda, E., Diehl, S.A., Scheeren, F.A., Ott, M., et al. (2010). Generation of human antigen-specific monoclonal IgM antibodies using vaccinated "human immune system" mice. PLoS One *5*, 1–10.

Benatuil, L., Perez, J.M., Belk, J., and Hsieh, C.M. (2010). An improved yeast transformation method for the generation of very large human antibody libraries. Protein Eng. Des. Sel. *23*, 155–159.

Boder, E.T., and Wittrup, K.D. (1997). Yeast surface display for screening combinatorial polypeptide libraries. Nat. Biotechnol. *15*, 553–557.

Boehm, T. (2011). Design principles of adaptive immune systems. Nat. Rev. Immunol. *11*, 307–317.

Bowen, E.T.W., Platt, G.S., Lloyd, G., Baskerville, A., Harris, W.J., and Vella, E.E. (1977). Viral Haemorrhagic Fever in Southern Sudan and Northern Zaire Rreliminary Studies on the Aetiological Agent. Lancet *309*, 571–573.

Bowley, D.R., Labrijn, A.F., Zwick, M.B., and Burton, D.R. (2007). Antigen selection from an HIV-1 immune antibody library displayed on yeast yields many novel antibodies compared to selection from the same library displayed on phage. Protein Eng. Des. Sel. *20*, 81–90.

Boyd, S.D., Marshall, E.L., Merker, J.D., Maniar, J.M., Zhang, L.N., Sahaf, B., Jones, C.D., Simen, B.B., Hanczaruk, B., Nguyen, K.D., et al. (2009). Measurement and clinical monitoring of human lymphocyte clonality by massively parallel {VDJ} pyrosequencing. Sci Transl Med *1*, 12ra23.

Bradbury, A.R.M., Sidhu, S., Dübel, S., and McCafferty, J. (2011). Beyond natural antibodies: the power of in vitro display technologies. Nat. Biotechnol. *29*, 245–254.

Brekke, O.H., and Sandlie, I. (2003). Therapeutic antibodies for human diseases at the dawn of the twenty-first century. Nat. Rev. Drug Discov. *2*, 52–62.

Briney, B., Sok, D., Jardine, J.G., Kulp, D.W., Skog, P., Menis, S., Jacak, R., Kalyuzhniy, O., de Val, N., Sesterhenn, F., et al. (2016). Tailored Immunogens Direct Affinity Maturation toward HIV Neutralizing Antibodies. Cell *166*, 1459–1470.e11.

Busse, C.E., Czogiel, I., Braun, P., Arndt, P.F., and Wardemann, H. (2014). Single-cell based high-throughput sequencing of full-length immunoglobulin heavy and light chain genes. Eur. J. Immunol. *44*, 597–603.

Calis, J.J.A., and Rosenberg, B.R. (2014). Characterizing immune repertoires by high throughput sequencing: strategies and applications. Trends Immunol. *35*, 581–590.

Carette, J.E., Raaben, M., Wong, A.C., Herbert, A.S., Obernosterer, G., Mulherkar, N., Kuehne, A.I., Kranzusch, P.J., Griffin, A.M., Ruthel, G., et al. (2011). Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. Nature *477*, 340–343.

Carroll, S.A., Towner, J.S., Sealy, T.K., McMullan, L.K., Khristova, M.L., Burt, F.J., Swanepoel, R., Rollin, P.E., and Nichol, S.T. (2013). Molecular evolution of viruses of the family Filoviridae based on 97 whole-genome sequences. J. Virol. *87*, 2608–2616.

Chan, A.C., and Carter, P.J. (2010). Therapeutic antibodies for autoimmunity and inflammation. Nat. Rev. Immunol. *10*, 301–316.

Chao, G., Lau, W.L., Hackel, B.J., Sazinsky, S.L., Lippow, S.M., and Wittrup, K.D. (2006). Isolating and engineering human antibodies using yeast surface display. Nat. Protoc. *1*, 755–768.

Chen, G., Koellho, J.F., Zak, S.E., Frei, J.C., Liu, N., Long, H., Ye, W., Nagar, K., Pan, G., Chandran, K., et al. (2014). Synthetic Antibodies with a Human Framework That Protect Mice from Lethal Sudan Ebolavirus Challenge. ACS Chem. Biol. *9*, 2263–2273.

Cheung, W.C., Beausoleil, S. a, Zhang, X., Sato, S., Schieferl, S.M., Wieler, J.S., Beaudet, J.G., Ramenani, R.K., Popova, L., Comb, M.J., et al. (2012). A proteomics approach for the identification and cloning of monoclonal antibodies from serum. Nat. Biotechnol. *30*, 447–452.

Corti, D., and Lanzavecchia, A. (2013). Broadly neutralizing antiviral antibodies. Annu. Rev. Immunol. *31*, 705–742.

Corti, D., Langedijk, J.P.M., Hinz, A., Seaman, M.S., Vanzetta, F., Fernandez-Rodriguez, B.M., Silacci, C., Pinna, D., Jarrossay, D., Balla-Jhagjhoorsingh, S., et al. (2010). Analysis of memory B cell responses and isolation of novel monoclonal antibodies with neutralizing breadth from HIV-1-infected individuals. PLoS One *5*.

Corti, D., Voss, J., Gamblin, S.J., Codoni, G., Macagno, A., Jarrossay, D., Vachieri, S.G., Pinna, D., Minola, A., Vanzetta, F., et al. (2011). A Neutralizing Antibody Selected from Plasma Cells That Binds to Group 1 and Group 2 Influenza A Hemagglutinins. Science (80-.). *333*, 850–856.

Corti, D., Misasi, J., Mulangu, S., Stanley, D.A., Kanekiyo, M., Wollen, S., Ploquin, A., Doria-Rose, N.A., Staupe, R.P., Bailey, M., et al. (2016). Protective monotherapy against lethal Ebola virus infection by a potently neutralizing antibody. Science (80-.). *351*, 1339–1342.

Dejnirattisai, W., Jumnainsong, A., Onsirisakul, N., Fitton, P., Vasanawathana, S., Limpitikul, W., Puttikhunt, C., Edwards, C., Supasa, T., Sunpetchuda, D., et al. (2010). Cross-Reacting Antibodies Enhance Dengue Virus Infection in Humans. Science (80-.). *328*, 745–748.

DeKosky, B.J., Ippolito, G.C., Deschner, R.P., Lavinder, J.J., Wine, Y., Rawlings, B.M., Varadarajan, N., Giesecke, C., Dörner, T., Andrews, S.F., et al. (2013). High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. Nat. Biotechnol. *31*, 166–169.

DeKosky, B.J., Kojima, T., Rodin, A., Charab, W., Ippolito, G.C., Ellington, A.D., and Georgiou, G. (2015). In-depth determination and analysis of the human paired heavy- and light-chain antibody repertoire. Nat Med *21*, 86–91.

Dessain, S.K., Adekar, S.P., and Berry, J.D. (2008). Exploring the native human antibody repertoire to create antiviral therapeutics. Curr. Top. Microbiol. Immunol. *317*, 155–183.

Dias, J.M., Kuehne, A.I., Abelson, D.M., Bale, S., Wong, A.C., Halfmann, P., Muhammad, M. a, Fusco, M.L., Zak, S.E., Kang, E., et al. (2011). A shared structural solution for neutralizing ebolaviruses. Nat. Struct. Mol. Biol. *18*, 1424–1427.

Doria-Rose, N.A., Schramm, C.A., Gorman, J., Moore, P.L., Bhiman, J.N., DeKosky, B.J., Ernandes, M.J., Georgiev, I.S., Kim, H.J., Pancera, M., et al. (2014). Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. Nature *509*, 55–62.

Dudley, D.D., Chaudhuri, J., Bassing, C.H., and Alt, F.W. (2005). Mechanism and control of V(D)J recombination versus class switch recombination: similarities and differences. Adv. Immunol. *86*, 43–112.

Escolano, A., Steichen, J.M., Dosenovic, P., Kulp, D.W., Golijanin, J., Sok, D., Freund, N.T., Gitlin, A.D., Oliveira, T., Araki, T., et al. (2016). Sequential Immunization Elicits Broadly Neutralizing Anti-HIV-1 Antibodies in Ig Knockin Mice. Cell *166*, 1445–1458.e12.

Feige, M.J., Hendershot, L.M., and Buchner, J. (2010). How antibodies fold. Trends Biochem. Sci. *35*, 189–198.

Feldhaus, M.J., Siegel, R.W., Opresko, L.K., Coleman, J.R., Feldhaus, J.M.W., Yeung, Y.A., Cochran, J.R., Heinzelman, P., Colby, D., Swers, J., et al. (2003). Flow-cytometric isolation of human antibodies from a nonimmune Saccharomyces cerevisiae surface display library. Nat. Biotechnol. *21*, 163–170.

Fellouse, F.A., Wiesmann, C., and Sidhu, S.S. (2004). Synthetic antibodies from a four-amino-acid code: a dominant role for tyrosine in antigen recognition. Proc. Natl. Acad. Sci. U. S. A. *101*, 12467–12472.

Fishwild, D.M., O'Donnell, S.L., Bengoechea, T., Hudson, D. V, Harding, F., Bernhard, S.L., Jones, D., Kay, R.M., Higgins, K.M., Schramm, S.R., et al. (1996). High-avidity human IgG kappa monoclonal antibodies from a novel strain of minilocus transgenic mice. Nat. Biotechnol. *14*, 845–851.

Georgiou, G., Ippolito, G.C., Beausang, J., Busse, C.E., Wardemann, H., and Quake, S.R.

(2014a). The promise and challenge of high-throughput sequencing of the antibody repertoire. Nat. Biotechnol. *32*, 158–168.

Georgiou, G., Ippolito, G.C., Beausang, J., Busse, C.E., Wardemann, H., and Quake, S.R. (2014b). The promise and challenge of high-throughput sequencing of the antibody repertoire. Nat. Biotechnol. *32*, 158–168.

Ghosh, S., and Campbell, A.M. (1986). Multispecific monoclonal antibodies. Immunol. Today *7*, 217–222.

Gire, S.K., Goba, A., Andersen, K.G., Sealfon, R.S.G., Park, D.J., Kanneh, L., Jalloh, S., Momoh, M., Fullah, M., Dudas, G., et al. (2014). Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. Science (80-.). *345*, 1369–1372.

Giudicelli, V., and Lefranc, M.-P. (1999). Ontology for immunogenetics: the IMGT-ONTOLOGY. BIoinformatics *15*, 1047–1054.

Giudicelli, V., Duroux, P., Ginestoux, C., Folch, G., Jabado-Michaloud, J., Chaume, D., and Lefranc, M.-P. (2006). IMGT/LIGM-DB, the IMGT comprehensive database of immunoglobulin and T cell receptor nucleotide sequences. Nucleic Acids Res. *34*, D781–D784.

Glanville, J., Zhai, W., Berka, J., Telman, D., Huerta, G., Mehta, G.R., Ni, I., Mei, L., Sundar, P.D., Day, G.M.R., et al. (2009). Precise determination of the diversity of a combinatorial antibody library gives insight into the human immunoglobulin repertoire. Proc. Natl. Acad. Sci. *106*, 20216–20221.

Gleichmann, H. (1981). Studies on the Mechanism of Drug Sensitization: T-cell-Dependent Popliteal Lymph Node Reaction to Diphenylhydantoin. Clin. Immunol. Immunopathol. *18*, 203–211.

Goebeler, M.E., Knop, S., Viardot, A., Kufer, P., Topp, M.S., Einsele, H., Noppeney, R., Hess, G., Kallert, S., Mackensen, A., et al. (2016). Bispecific T-cell engager (BiTE) antibody construct Blinatumomab for the treatment of Patients with relapsed/refractory non-Hodgkin lymphoma: Final results from a phase I study. J. Clin. Oncol. *34*, 1104–1111.

Green, L.L., Hardy, M.C., Maynard-Currie, C.E., Tsuda, H., Louie, D.M., Mendez, M.J., Abderrahim, H., Noguchi, M., Smith, D.H., Zeng, Y., et al. (1994). Antigen-specific human monoclonal antibodies from mice engineered with human Ig heavy and light chain YACs. Nat. Genet. *7*, 13–21.

Le Guenno, B., Formentry, P., Wyers, M., Gounon, P., Walker, F., and Boesch, C. (1995). Isolation and partial characterisation of a new strain of Ebola virus. Lancet *345*, 1271–1274.

Hayhurst, A., Happe, S., Mabry, R., Koch, Z., Iverson, B.L., and Georgiou, G. (2003). Isolation and expression of recombinant antibody fragments to the biological warfare pathogen Brucella melitensis. J. Immunol. Methods *276*, 185–196.

Heinz Feldmann, Stuart T. Nichol, Hans-Dieter Klenk, Clarence J. Peters, A.S. (1994). Characterization of filoviruses based on differences in structure and antigenicity of the virion glycoprotein. Virology *199*, 469–473.

Höfer, T., Muehlinghaus, G., Moser, K., Yoshida, T., Mei, H.E., Hebel, K., Hauser, A., Hoyer, B., Luger, E.O., Dörner, T., et al. (2006). Adaptation of humoral memory. Immunol. Rev. *211*, 295–302.

Hood, C.L., Abraham, J., Boyington, J.C., Leung, K., Kwong, P.D., and Nabel, G.J. (2010). Biochemical and structural characterization of cathepsin L-processed Ebola virus glycoprotein: implications for viral entry and immunogenicity. J. Virol. *84*, 2972–2982.

Hoogenboom, H.R. (2005). Selecting and screening recombinant antibody libraries. Nat. Biotechnol. *23*, 1105–1116.

Huang, J., Ofek, G., Laub, L., Louder, M.K., Doria-Rose, N. a, Longo, N.S., Imamichi, H., Bailer, R.T., Chakrabarti, B., Sharma, S.K., et al. (2012). Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. Nature *491*, 406–412.

Huang, J., Kang, B.H., Pancera, M., Lee, J.H., Tong, T., Feng, Y., Imamichi, H., Georgiev, I.S., Chuang, G.-Y., Druz, A., et al. (2014). Broad and potent HIV-1 neutralization by a human antibody that binds the gp41-gp120 interface. Nature *515*, 138–142.

Jayaram, N., Bhowmick, P., and Martin, A.C.R. (2012). Germline VH/VL pairing in antibodies. Protein Eng. Des. Sel. *25*, 523–529.

Jiang, N., He, J., Weinstein, J.A., Penland, L., Sasaki, S., He, X.-S., Dekker, C.L., Zheng, N.-Y., Huang, M., Sullivan, M., et al. (2013). Lineage structure of the human antibody repertoire in response to influenza vaccination. Sci. Transl. … *5*, 1–16.

Johnson, G., and Wu, T. Te (2004). The Kabat Database and a Bioinformatics Example. Methods Mol. Biol. *248*, 11–25.

Kamala, T. (2007). Hock immunization: a humane alternative to mouse footpad injections. J. Immunol. Methods *328*, 204–214.

Kaufmann, K.W., Lemmon, G.H., Deluca, S.L., Sheehan, J.H., and Meiler, J. (2010). Practically useful: What the R osetta protein modeling suite can do for you. Biochemistry *49*, 2987–2998.

Kaushik, A., and Lim, W. (1996). The primary antibody repertoire of normal, immunodeficient and autoimmune mice is characterized by differences in V gene expression. Res. Immunol. *147*, 9–26.

Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wölle, J., Plückthun, A., and Virnekäs, B. (2000). Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. J. Mol. Biol. *296*, 57–86.

Knight, B. (1979). Ricin-a potent homicidal poison. Br. Med. J. *1*, 350–351.

Koellhoffer, J.F., Chen, G., Sandesara, R.G., Bale, S., Saphire, E.O., Chandran, K., Sidhu, S.S., and Lai, J.R. (2012). Two synthetic antibodies that recognize and neutralize distinct proteolytic forms of the ebola virus envelope glycoprotein. Chembiochem *13*, 2549–2457.

Kohler, G., and Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. Nature *256*, 495–497.

Kondratowicz, A.S., Lennemann, N.J., Sinn, P.L., Davey, R. a, Hunt, C.L., Moller-Tank, S., Meyerholz, D.K., Rennert, P., Mullins, R.F., Brindley, M., et al. (2011). T-cell immunoglobulin and mucin domain 1 (TIM-1) is a receptor for Zaire Ebolavirus and Lake Victoria Marburgvirus. Proc. Natl. Acad. Sci. U. S. A. *108*, 8426–8431.

Kong, R., Xu, K., Zhou, T., Acharya, P., Lemmin, T., Liu, K., Ozorowski, G., Soto, C., Taft, J.D., Bailer, R.T., et al. (2016). Fusion peptide of HIV-1 as a site of vulnerability to neutralizing antibody. Science (80-.). *352*, 423–426.

Kumar, H., Kawai, T., and Akira, S. (2011). Pathogen recognition by the innate immune system. Int. Rev. Immunol. *30*, 16–34.

Kuroiwa, Y., Kasinathan, P., Sathiyaseelan, T., Jiao, J., Matsushita, H., Sathiyaseelan, J., Wu, H., Mellquist, J., Hammitt, M., Koster, J., et al. (2009). Antigen-specific human polyclonal antibodies from hyperimmunized cattle. Nat. Biotechnol. *27*, 173–181.

Kurosawa, N., Yoshioka, M., Fujimoto, R., Yamagishi, F., and Isobe, M. (2012). Rapid production of antigen-specific monoclonal antibodies from a variety of animals. BMC Biol. *10*, 1–14.

Labrijn, A.F., Meesters, J.I., de Goeij, B.E.C.G., van den Bremer, E.T.J., Neijssen, J., van Kampen, M.D., Strumane, K., Verploegen, S., Kundu, A., Gramer, M.J., et al. (2013). Efficient generation of stable bispecific IgG1 by controlled Fab-arm exchange. Proc. Natl. Acad. Sci. U. S. A. *110*, 5145–5150.

Lane, D., and Koprowski, H. (1982). Molecular recognition and the future of monoclonal antibodies. Nature *296*, 200–202.

Lapidoth, G.D., Baran, D., Pszolla, G.M., Norn, C., Alon, A., Tyka, M.D., and Fleishman, S.J. (2015). AbDesign: An algorithm for combinatorial backbone design guided by natural conformations and sequences. Proteins Struct. Funct. Bioinforma. *83*, 1385–1406.

Lavinder, J.J., Wine, Y., Giesecke, C., Ippolito, G.C., Horton, A.P., Lungu, O.I., Hoi, K.H., DeKosky, B.J., Murrin, E.M., Wirth, M.M., et al. (2014). Identification and characterization of the constituent human serum antibodies elicited by vaccination. Proc. Natl. Acad. Sci. U. S. A. *111*, 2259–2264.

Lee, E.-C., Liang, Q., Ali, H., Bayliss, L., Beasley, A., Bloomfield-Gerdes, T., Bonoli, L., Brown, R., Campbell, J., Carpenter, A., et al. (2014). Complete humanization of the mouse immunoglobulin loci enables efficient therapeutic antibody discovery. Nat. Biotechnol. *32*, 356–363.

Lee, J.E., Fusco, M.L., Hessell, A.J., Oswald, W.B., Burton, D.R., and Saphire, E.O. (2008). Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor. Nature *454*, 177–182.

Lefranc, M.-P., Giudicelli, V., Ginestoux, C., Jabado-Michaloud, J., Folch, G., Bellahcene, F., Wu, Y., Gemrot, E., Brochet, X., Lane, J., et al. (2009). IMGT, the international ImMunoGeneTics information system. Nucleic Acids Res. *37*, D1006–D1012.

Lewis, S.M., Wu, X., Pustilnik, A., Sereno, A., Huang, F., Rick, H.L., Guntas, G., Leaver-Fay, A., Smith, E.M., Ho, C., et al. (2014). Generation of bispecific IgG antibodies by structure-based design of an orthogonal Fab interface. Nat Biotechnol *32*, 191–198.

Li, Z., Woo, C.J., Iglesias-ussel, M.D., Ronai, D., and Scharff, M.D. (2004). The generation of antibody diversity through somatic hypermutation and class switch recombination. Genes Dev. *18*, 1–11.

Liao, H.-X., Lynch, R., Zhou, T., Gao, F., Alam, S.M., Boyd, S.D., Fire, A.Z., Roskin, K.M., Schramm, C. a, Zhang, Z., et al. (2013). Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. Nature *496*, 469–476.

Lin, G., Simmons, G., Pohlmann, S., Baribaud, F., Ni, H., Leslie, G.J., Haggarty, B.S., Bates, P., Weissman, D., Hoxie, J.A., et al. (2003). Differential N-Linked Glycosylation of Human Immunodeficiency Virus and Ebola Virus Envelope Glycoproteins Modulates Interactions with DC-SIGN and DC-SIGNR. J. Virol. *77*, 1337–1346.

Liu, Y., Caffry, I., Wu, J., Geng, S.B., Jain, T., Sun, T., Reid, F., Cao, Y., Estep, P., Yu, Y., et al. (2014). High-throughput screening for developability during early-stage antibody discovery using self-interaction nanoparticle spectroscopy. MAbs *6*, 483–492.

Lonberg, N., Taylor, L.D., Harding, F.A., Trounstine, M., Higgins, K.M., Schramm, S.R., Kuo, C.C., Mashayekh, R., Wymore, K., and McCabe, J.G. (1994). Antigen-specific human antibodies from mice comprising four distinct genetic modifications. Nature *368*, 856–859.

Lord, J.M., Roberts, L.M., and Robertus, J.D. (1994). Ricin: structure, mode of action, and some current applications. J. Fed. Am. Soc. Exp. Biol. *8*, 201–208.

Lu, J., Panavas, T., Thys, K., Aerssens, J., Naso, M., Fisher, J., Rycyzyn, M., and Sweet, R.W. (2014). IgG variable region and VH CDR3 diversity in unimmunized mice analyzed by massively parallel sequencing. Mol. Immunol. *57*, 274–283.

Macdonald, L.E., Karow, M., Stevens, S., Auerbach, W., Poueymirou, W.T., Yasenchak, J., Frendewey, D., Valenzuela, D.M., Giallourakis, C.C., Alt, F.W., et al. (2014). Precise and in situ genetic humanization of 6 Mb of mouse immunoglobulin genes. Proc. Natl. Acad. Sci. U. S. A. *111*, 5147–5152.

Maddaloni, M., Cooke, C., Wilkinson, R., Stout, A. V, Eng, L., and Pincus, S.H. (2004). Immunological Characteristics Associated with the Protective Efficacy of Antibodies to Ricin. J. Immunol. *172*, 6221–6228.

Mandell, D.J., Coutsias, E. a, and Kortemme, T. (2009). Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling. Nat. Methods *6*, 551–552.

Manicassamy, B., Wang, J., Rumschlag, E., Tymen, S., Volchkova, V., Volchkov, V., and Rong, L. (2007). Characterization of Marburg virus glycoprotein in viral entry. Virology *358*, 79–88.

Mantis, N.J., McGuinness, C.R., Sonuyi, O., Edwards, G., and Farrant, S.A. (2006). Immunoglobulin A antibodies against ricin A and B subunits protect epithelial cells from ricin intoxication. Infect. Immun. *74*, 3455–3462.

Maruyama, T., Rodriguez, L.L., Jahrling, P.B., Sanchez, A., Khan, A.S., Nichol, S.T., Peters, C.J., Parren, P.W.H.I., Burton, D.R., and Khan, A.L.I.S. (1999). Ebola Virus Can Be Effectively Neutralized by Antibody Produced in Natural Human Infection. J. Virol. *73*, 6024–6030.

Mayor, S. (2003). UK doctors warned after ricin poison found in police raid. Br. Med. J. *326*, 126.

Mazor, Y., Van Blarcom, T., Mabry, R., Iverson, B.L., and Georgiou, G. (2007). Isolation of engineered, full-length antibodies from libraries expressed in Escherichia coli. Nat. Biotechnol. *25*, 563–565.

Mcconnell, A.D., Zhang, X., Macomber, J.L., Chau, B., Joseph, C., Rahmanian, S., Hare, E., Spasojevic, V., Horlick, R.A., King, D.J., et al. (2014). A general approach to antibody thermostabilization. MAbs *6*, 1274–1282.

Miranda, M.E.G., and Miranda, N.L.J. (2011). Reston ebolavirus in humans and animals in the Philippines: a review. J. Infect. Dis. *204 Suppl*, S757–S760.

Murphy, A.J., Macdonald, L.E., Stevens, S., Karow, M., Dore, A.T., Pobursky, K., Huang, T.T., Poueymirou, W.T., Esau, L., Meola, M., et al. (2014). Mice with megabase humanization of their immunoglobulin genes generate antibodies as efficiently as normal mice. Proc. Natl. Acad. Sci. U. S. A. *111*, 5153–5158.

Muyldermans, S. (2013). Nanobodies: Natural Single-Domain Antibodies. Annu. Rev. Biochem. *82*, 775–797.

Nadel, B., and Feeney, A.J. (1997). Nucleotide Deletion and P Addition in V(D)J Recombination: a Determinant Role of the Coding-End Sequence. Mol. Cell. Biol. *17*, 3768–3778.

Neal, L.M., O'Hara, J., Brey, R.N., and Mantis, N.J. (2010). A monoclonal immunoglobulin G antibody directed against an immunodominant linear epitope on the ricin A chain confers systemic and mucosal immunity to ricin. Infect. Immun. *78*, 552–

561.

Di Niro, R., Mesin, L., Raki, M., Zheng, N.-Y., Lund-Johansen, F., Lundin, K.E.A., Charpilienne, A., Poncet, D., Wilson, P.C., and Sollid, L.M. (2010). Rapid Generation of Rotavirus-Specific Human Monoclonal Antibodies from Small-Intestinal Mucosa. J. Immunol. *185*, 5377–5383.

Di Niro, R., Mesin, L., Zheng, N.-Y., Stamnaes, J., Morrissey, M., Lee, J.-H., Huang, M., Iversen, R., du Pré, M.F., Qiao, S.-W., et al. (2012). High abundance of plasma cells secreting transglutaminase 2–specific IgA autoantibodies with limited somatic hypermutation in celiac disease intestinal lesions. Nat. Med. *18*, 441–445.

Ojima-Kato, T., Fukui, K., Yamamoto, H., Hashimura, D., Miyake, S., Hirakawa, Y., Yamasaki, T., Kojima, T., and Nakano, H. (2016). "Zipbody" leucine zipper-fused Fab in *E. coli in vitro* and *in vivo* expression systems. Protein Eng. Des. Sel. *29*, 149–157.

Parren, P.W.H.I., Geisbert, T.W., Maruyama, T., Jahrling, P.B., and Burton, D.R. (2002). Pre- and Postexposure Prophylaxis of Ebola Virus Infection in an Animal Model by Passive Transfer of a Neutralizing Human Antibody. J. Virol. *76*, 6408–6412.

Pelat, T., Hust, M., Hale, M., Lefranc, M.-P., Dübel, S., and Thullier, P. (2009). Isolation of a human-like antibody fragment (scFv) that neutralizes ricin biological activity. BMC Biotechnol. *9*, 1–13.

Persson, H., Ye, W., Wernimont, A., Adams, J.J., Koide, A., Koide, S., Lam, R., and Sidhu, S.S. (2013). CDR-H3 diversity is not required for antigen recognition by synthetic antibodies. J. Mol. Biol. *425*, 803–811.

Ponsel, D., Neugebauer, J., Ladetzki-Baehs, K., and Tissot, K. (2011). High affinity, developability and functional size: The holy grail of combinatorial antibody library generation. Molecules *16*, 3675–3700.

Prigent, J., Panigai, L., Lamourette, P., Sauvaire, D., Devilliers, K., Plaisance, M., Volland, H., Créminon, C., and Simon, S. (2011). Neutralising antibodies against ricin toxin. PLoS One *6*, e20166.

Qiu, X., Alimonti, J.B., Melito, P.L., Fernando, L., Ströher, U., and Jones, S.M. (2011). Characterization of Zaire ebolavirus glycoprotein-specific monoclonal antibodies. Clin. Immunol. *141*, 218–227.

Qiu, X., Wong, G., Audet, J., Bello, A., Fernando, L., Alimonti, J.B., Fausther-Bovendo, H., Wei, H., Aviles, J., Hiatt, E., et al. (2014). Reversion of advanced Ebola virus disease

in nonhuman primates with ZMapp. Nature *514*, 47–53.

Quintero-Hernández, V., Juárez-González, V.R., Ortíz-León, M., Sánchez, R., Possani, L.D., and Becerril, B. (2007). The change of the scFv into the Fab format improves the stability and in vivo toxin neutralization capacity of recombinant antibodies. Mol. Immunol. *44*, 1307–1315.

Ravel, G., and Descotes, J. (2005). Popliteal lymph node assay: facts and perspectives. J. Appl. Toxicol. *25*, 451–458.

Rechavi, E., Lev, A., Lee, Y.N., Simon, A.J., Yinon, Y., Lipitz, S., Amariglio, N., Weisz, B., Notarangelo, L.D., and Somech, R. (2015). Timely and spatially regulated maturation of B and T cell repertoire during human fetal development. Sci. Transl. Med. *7*, 276ra25.

Reddy, S.T., Swartz, M.A., and Hubbell, J.A. (2006). Targeting dendritic cells with biomaterials: developing the next generation of vaccines. Trends Immunol. *27*, 573–579.

Reddy, S.T., Ge, X., Miklos, A.E., Hughes, R. a, Kang, S.H., Hoi, K.H., Chrysostomou, C., Hunicke-Smith, S.P., Iverson, B.L., Tucker, P.W., et al. (2010). Monoclonal antibodies isolated without screening by analyzing the variable-gene repertoire of plasma cells. Nat. Biotechnol. *28*, 965–969.

Van Regenmortel, M.H. V (2014). Specificity, polyspecificity, and heterospecificity of antibody-antigen recognition. J. Mol. Recognit. *27*, 627–639.

Robbiani, D.F., Deroubaix, S., Feldhahn, N., Oliveira, T.Y., Callen, E., Wang, Q., Jankovic, M., Silva, I.T., Rommel, P.C., Bosque, D., et al. (2015). Plasmodium Infection Promotes Genomic Instability and AID-Dependent B Cell Lymphoma. Cell *162*, 727–737.

Rouet, R., Lowe, D., and Christ, D. (2014). Stability engineering of the human antibody repertoire. FEBS Lett. *588*, 269–277.

Saggy, I., Wine, Y., Shefet-Carasso, L., Nahary, L., Georgiou, G., and Benhar, I. (2012). Antibody isolation from immunized animals: comparison of phage display and antibody discovery via V gene repertoire mining. Protein Eng. Des. Sel. *25*, 539–549.

Sato, S., Beausoleil, S. a, Popova, L., Beaudet, J.G., Ramenani, R.K., Zhang, X., Wieler, J.S., Schieferl, S.M., Cheung, W.C., and Polakiewicz, R.D. (2012). Proteomics-directed cloning of circulating antiviral human monoclonal antibodies. Nat. Biotechnol. *30*, 1039–1043.

Scheid, J.F., Mouquet, H., Feldhahn, N., Seaman, M.S., Velinzon, K., Pietzsch, J., Ott, R.G., Anthony, R.M., Zebroski, H., Hurley, A., et al. (2009). Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. Nature *458*, 636–640.

Scheid, J.F., Mouquet, H., Ueberheide, B., Diskin, R., Klein, F., Oliveira, T.Y.K., Pietzsch, J., Fenyo, D., Abadir, A., Velinzon, K., et al. (2011). Sequence and Structural Convergence of Broad and Potent HIV Antibodies That Mimic CD4 Binding. Science (80-.). *333*, 1633–1637.

Schoonbroodt, S., Steukers, M., Viswanathan, M., Frans, N., Timmermans, M., Wehnert, a., Nguyen, M., Ladner, R.C., and Hoet, R.M. (2008). Engineering Antibody Heavy Chain CDR3 to Create a Phage Display Fab Library Rich in Antibodies That Bind Charged Carbohydrates. J. Immunol. *181*, 6213–6221.

Sheets, M.D., Amersdorfer, P., Finnern, R., Sargent, P., Lindquist, E., Schier, R., Hemingsen, G., Wong, C., Gerhart, J.C., and Marks, J.D. (1998). Efficient construction of a large nonimmune phage antibody library: the production of high-affinity human single-chain antibodies to protein antigens. Proc. Natl. Acad. Sci. U. S. A. *95*, 6157–6162.

Slifka, M.K., Antia, R., Whitmire, J.K., and Ahmed, R. (1998). Humoral immunity due to long-lived plasma cells. Immunity *8*, 363–372.

Spiess, C., Merchant, M., Huang, A., Zheng, Z., Yang, N.-Y.Y., Peng, J., Ellerman, D., Shatz, W., Reilly, D., Yansura, D.G., et al. (2013). Bispecific antibodies with natural architecture produced by co-culture of bacteria expressing two distinct half-antibodies. Nat. Biotechnol. *31*, 753–758.

Stettler, K., Beltramello, M., Espinosa, D.A., Graham, V., Cassotta, A., Bianchi, S., Vanzetta, F., Minola, A., Jaconi, S., Mele, F., et al. (2016). Specificity, cross-reactivity and function of antibodies elicited by Zika virus infection. Science (80-.). *353*, 1–9.

Tan, J., Pieper, K., Piccoli, L., Abdi, A., Foglierini, M., Geiger, R., Maria Tully, C., Jarrossay, D., Maina Ndungu, F., Wambua, J., et al. (2016). A LAIR1 insertion generates broadly reactive antibodies against malaria variantTan, J., Pieper, K., Piccoli, L., Abdi, A., Foglierini, M., Geiger, R., … Lanzavecchia, A. (2016). A LAIR1 insertion generates broadly reactive antibodies against malaria varia. Nature *529*, 105–109.

Taylor, L.D., Carmack, C.E., Schramm, S.R., Mashayekh, R., Higgins, K.M., Kuo, C.

chi, Woodhouse, C., Kay, R.M., and Lonberg, N. (1992). A transgenic mouse that expresses a diversity of human sequence heavy and light chain immunoglobulins. Nucleic Acids Res. *20*, 6287–6295.

Team, W.E.R. (2014). Ebola Virus Disease in West Africa - The First 9 Months of the Epidemic and Forward Projections. N. Engl. J. Med. *371*, 1481–1495.

Tiller, T., Schuster, I., Deppe, D., Siegers, K., Strohner, R., Herrmann, T., Berenguer, M., Poujol, D., Stehle, J., Stark, Y., et al. (2013). A fully synthetic human Fab antibody library based on fixed VH/VL framework pairings with favorable biophysical properties. MAbs *5*, 445–470.

Tonegawa, S. (1983). Somatic generation of antibody diversity. Naure *302*, 575–581.

Topalian, S.L., Taube, J.M., Anders, R.A., and Pardoll, D.M. (2016). Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. Nat. Rev. Cancer *16*, 275–287.

Towner, J.S., Rollin, P.E., Bausch, D.G., Sanchez, a., Crary, S.M., Vincent, M., Lee, W.F., Spiropoulou, C.F., Ksiazek, T.G., Lukwiya, M., et al. (2004). Rapid Diagnosis of Ebola Hemorrhagic Fever by Reverse Transcription-PCR in an Outbreak Setting and Assessment of Patient Viral Load as a Predictor of Outcome. J. Virol. *78*, 4330–4341.

Towner, J.S., Sealy, T.K., Khristova, M.L., Albariño, C.G., Conlan, S., Reeder, S. a, Quan, P.-L., Lipkin, W.I., Downing, R., Tappero, J.W., et al. (2008). Newly discovered ebola virus associated with hemorrhagic fever outbreak in Uganda. PLoS Pathog. *4*, e1000212.

Traggiai, E., Becker, S., Subbarao, K., Kolesnikova, L., Uematsu, Y., Gismondo, M.R., Murphy, B.R., Rappuoli, R., and Lanzavecchia, A. (2004). An efficient method to make human monoclonal antibodies from memory B cells: potent neutralization of SARS coronavirus. Nat. Med. *10*, 871–875.

Venet, S., Kosco-Vilbois, M., and Fischer, N. (2013). Comparing CDRH3 diversity captured from secondary lymphoid organs for the generation of recombinant human antibodies. MAbs *5*, 690–698.

Walker, L.M., Phogat, S.K., Wagner, D., Phung, P., Goss, J.L., Wrin, T., Simek, M.D., Fling, S., Mitcham, J.L., Lehrman, J.K., et al. (2009). Broad and Potent Neutralizing Antibodies from an African Donor Reveal a New HIV-1 Vaccine Target. Science (80-.). *326*, 285–289.

Walker, L.M., Huber, M., Doores, K.J., Falkowska, E., Pejchal, R., Julien, J.-P., Wang, S.-K., Ramos, A., Chan-Hui, P.-Y., Moyle, M., et al. (2011). Broad neutralization coverage of HIV by multiple highly potent antibodies. Nature *477*, 466–470.

Wang, B., Kluwe, C.A., Lungu, O.I., DeKosky, B.J., Kerr, S.A., Johnson, E.L., Jung, J., Rezigh, A.B., Carroll, S.M., Reyes, A.N., et al. (2015a). Facile Discovery of a Diverse Panel of Anti-Ebola Virus Antibodies by Immune Repertoire Mining. Sci. Rep. *5*, 13926.

Wang, B., Lee, C.H., Johnson, E.L., Kluwe, C.A., Cunningham, J.C., Tanno, H., Crooks, R.M., Georgiou, G., and Ellington, A.D. (2016). Discovery of high affinity anti-ricin antibodies by B cell receptor sequencing and by yeast display of combinatorial VH:VL libraries from immunized animals. MAbs *8*, 1035–1044.

Wang, C., Liu, Y., Cavanagh, M.M., Le Saux, S., Qi, Q., Roskin, K.M., Looney, T.J., Lee, J.-Y., Dixit, V., Dekker, C.L., et al. (2015b). B-cell repertoire responses to varicella-zoster vaccination in human identical twins. Proc. Natl. Acad. Sci. U. S. A. *112*, 500–505.

Wang, S., Liu, M., Zeng, D., Qiu, W., Ma, P., Yu, Y., Chang, H., and Sun, Z.W. (2014). Increasing stability of antibody via antibody engineering: Stability engineering on an anti-hVEGF. Proteins Struct. Funct. Bioinforma. *82*, 2620–2630.

Wardemann, H., Yurasov, S., Schaefer, A., Young, J.W., Meffre, E., and Nussenzweig, M.C. (2003). Predominant autoantibody production by early human B cell precursors. Science (80-.). *301*, 1374–1377.

Wentz, A.E., and Shusta, E. V. (2007). A novel high-throughput screen reveals yeast genes that increase secretion of heterologous proteins. Appl. Environ. Microbiol. *73*, 1189–1198.

Whittle, J.R., Wheatley, A.K., Wu, L., Lingwood, D., Kanekiyo, M., Ma, S.S., Narpala, S.R., Yassine, H.M., Frank, G.M., Yewdell, J.W., et al. (2014). Flow cytometry reveals that H5N1 vaccination elicits cross-reactive stem-directed antibodies from multiple Ig heavy-chain lineages. J Virol *88*, 4047–4057.

Wilson, P.C., and Andrews, S.F. (2012). Tools to therapeutically harness the human antibody response. Nat. Rev. Immunol. *12*, 709–719.

Wine, Y., Boutz, D.R., Lavinder, J.J., Miklos, A.E., Hughes, R. a, Hoi, K.H., Jung, S.T., Horton, A.P., Murrin, E.M., Ellington, A.D., et al. (2013). Molecular deconvolution of the monoclonal antibodies that comprise the polyclonal serum response. Proc. Natl.

Acad. Sci. U. S. A. *110*, 2993–2998.

Winter, G., Griffiths, A.D., Hawkins, R.E., and Hoogenboom, H.R. (1994). Making antibodies by phage display technology. Annu. Rev. Immunol. *12*, 433–455.

Wörn, A., and Plückthun, A. (2001). Stability engineering of antibody single-chain Fv fragments. J. Mol. Biol. *305*, 989–1010.

Wrammert, J., Smith, K., Miller, J., Langley, W. a, Kokko, K., Larsen, C., Zheng, N.-Y., Mays, I., Garman, L., Helms, C., et al. (2008). Rapid cloning of high-affinity human monoclonal antibodies against influenza virus. Nature *453*, 667–671.

Wu, X., Yang, Z.-Y., Li, Y., Hogerkorp, C.-M., Schief, W.R., Seaman, M.S., Zhou, T., Schmidt, S.D., Wu, L., Xu, L., et al. (2010). Rational Design of Envelope Identifies Broadly Neutralizing Human Monoclonal Antibodies to HIV-1. Science (80-.). *329*, 856–861.

Xie, J., Zhang, H., Yea, K., and Lerner, R. a (2013). Autocrine signaling based selection of combinatorial antibodies that transdifferentiate human stem cells. Proc. Natl. Acad. Sci. U. S. A. *110*, 8099–8104.

Xu, Y., Roach, W., Sun, T., Jain, T., Prinz, B., Yu, T.Y., Torrey, J., Thomas, J., Bobrowicz, P., Vásquez, M., et al. (2013). Addressing polyspecificity of antibodies selected from an in vitro yeast presentation system: A FACS-based, high-throughput selection and analytical tool. Protein Eng. Des. Sel. *26*, 663–670.

Zhu, J., Ofek, G., Yang, Y., Zhang, B., Louder, M.K., Lu, G., McKee, K., Pancera, M., Skinner, J., Zhang, Z., et al. (2013a). Mining the antibodyome for HIV-1-neutralizing antibodies with next-generation sequencing and phylogenetic pairing of heavy/light chains. Proc. Natl. Acad. Sci. U. S. A. *110*, 6470–6475.

Zhu, J., Wu, X., Zhang, B., Mckee, K., O'Dell, S., Soto, C., Zhou, T., Casazza, J.P., Mullikin, J.C., Kwong, P.D., et al. (2013b). De novo identification of VRC01 class HIV-1–neutralizing antibodies by next-generation sequencing of B-cell transcripts. Proc. Natl. Acad. Sci. U. S. A. *110*, 6470–6475.