Development and Application of Antibody Discovery Technologies For Diverse Viral Pathogens

By

Andrea Raye Shiakolas

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Microbe-Host Interactions

May 13, 2022

Nashville, Tennessee

Approved:

James Crowe Jr., Committee Chair, M.D.

Borden Lacy, Committee Member, Ph.D.

Antonis Rokas, Committee Member, Ph.D.

Spyros Kalams, Committee Member, M.D.

Ivelin Georgiev, Ph.D

Dedicated to my loving family and fiancé

ACKNOWLEDGEMENTS

First, I would like to thank my mentor and friend, Dr. Ivelin Georgiev. Thank you so much for believing in me and pushing me to do great science. I have learned so much from you. I admire your passion for the work we do and have appreciated all the support you have given me through the years, both personally and professionally.

I would also like to thank my committee for the incredible support, encouragement, and constructive help the past four years. Dr. James Crowe Jr., Dr. Borden Lacy, Dr. Antonis Rokas, and Dr. Spyros Kalams – you all have challenged and helped me to grow as a scientist. Thank you so much.

Additionally, I would like to thank the past and present members of the Georgiev laboratory. I would not have made it through the rigorous ups and downs of graduate school without your support, laughter, thoughtfulness, and friendship. First, efharistw to Ian Setliff - working with you on the development of LIBRA-seq was a wonderful experience, and I admire your intellect, attention to detail, and big ideas. Next, thank you to Kevin Kramer for being a great lab partner and spearheading the COVID-19 work together. Thank you for your hard work, sense of humor and for being a great friend. To Kelsey Pilewski, thank you for all the coffee breaks, science conversations, girls bay, and for being an inspirational woman in science. Thank you so much for your honest friendship. To Lauren Walker, thank you for your friendship and levity. I have loved becoming friends - you make the lab a happier place. To Steven Wall, thank you so much for your friendship, support, and admirable work ethic. To Nagarajan Raju, thank you for all of the support over the years on all of my projects – I could not have done it

without you. You are so kind and helped me so many times as a scientist and friend. To Amyn Murji, thank you for being a desk buddy, fantastic friend, and for helping encourage me to join the lab. To Allie Greenplate, Emilee Friedman-Fechter, Juliana Qin, Rohit Venkat, I loved working with each of you and am very thankful our paths crossed. To Alex Abu-Shmais, Matt Vukovich, Clint Holt, Parker Jamieson, Michelle Armstrong, I have truly enjoyed working with each of you as lab members and friends. Good luck with all the spectacular science to come. I would also like to thank the friends I made in graduate school outside of the lab. Specifically, thank you to Jenna Petronglo and Casey Butrico for your support, friendship, and laughter.

Thank you to our collaborators at Vanderbilt and across the world. I am very thankful and grateful to have had the opportunity to work together on all the projects described in this dissertation. Specifically, thank you to the Morris laboratory, the Graham laboratory, the Connors laboratory, the Lingwood laboratory, and the Acharya laboratory for your contributions to the LIBRA-seq work. Thank you to the Ledgerwood laboratory, Graham laboratory, Diamond laboratory, Haynes laboratory, Acharya laboratory, Crowe laboratory, Baric laboratory, Morris laboratory, and McLellan laboratory for your contributions to the cross-reactive coronavirus antibody work. Thank you to the Crowe laboratory, Bukreyev laboratory, and McLellan laboratory for your contributions to the LIBRA-seq with ligand blocking work. These studies would not have come together without the incredible collaborative science – thank you all.

Finally, thank you to my wonderful family for their support over the years. Thank you to my sister and best friend, Evi Shiakolas, and my parents, Panos Shiakolas and Susan Shiakolas. You all have always encouraged me to follow my passions and to

work hard. I would not have made it to this point without you. Evi, thank you for always being there for me, no matter what. To my father, thank you for illustrating selflessness, hard work, and pursuit of knowledge to help others. To my mother, thank you for loving me and encouraging me, every day. Thank you to my grandfathers, Felix Fenter and Stavros Shiakolas, for your continued support and example of perseverance, dedication, and strength. Thank you to my family in Cyprus for your love and support. Thank you to my future family, the Bennicis, for your encouragement and support. And thank you to our new puppy Doug for being a silly addition to the family that helps me to laugh every day.

Lastly, thank you to Anthony Bennici, my partner and future husband. I could not have made it to this point without your never-ending love, support, and encouragement. Thank you for always being there, especially when graduate school became challenging and difficult. You pushed me to do my best, and I am so thankful to have you.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

Introduction

1.1 Thesis Overview

This dissertation is comprised of my research efforts to develop novel antibody discovery technologies and apply them for antibody identification and characterization to a variety of viral pathogens, including HIV-1, influenza, and coronaviruses. The goal of this body of work was to create an antibody discovery tool that could lead to efficient screening of B cells for antigen reactivity and antibody sequence simultaneously, with the ultimate goal of rapid and efficient lead prioritization. This is important because immune repertoire characterization is an integral part of adaptive immunology research, with a special emphasis on studying the antibody response to pathogens, design and evaluation of vaccines, and development of antibody therapeutics. My thesis work has been divided into five chapters, which represent the variety of projects that I have worked on under the mentorship of Ivelin Georgiev, Ph. D, and with the collaborative help of many scientists in the laboratory, at Vanderbilt, and across the world.

In Chapter 1, I describe the immune system and the development of adaptive immune features. I also highlight relevant, important background information about viruses described in this dissertation, including HIV-1, influenza virus, and coronaviruses. Furthermore, I give an overview of existing antibody discovery tools, including the associated limitations and inefficiencies that drove us to pursue the

development of a novel antibody discovery technology termed LIBRA-seq (Linking B cell receptor to antigen specificity through sequencing).

In Chapter 2, I describe the development and application of LIBRA-seq. First, an overview of the LIBRA-seq workflow is included. Briefly, B cells are mixed with DNA-barcoded antigens, which allow for bioinformatic mapping of specific antigen DNA barcodes to paired heavy and light chain sequences of the B cell receptor (BCR) for each cell. Next, LIBRA-seq was validated in a proof-of-principle experiment, where I evaluated the ability of LIBRA-seq to accurately map antigen specificity to B cell receptor sequence using two Ramos cell lines with known B cell receptor sequences and specificities. This experiment demonstrated the ability to use LIBRA-seq to accurately identify antigen binding (including for multiple antigens) and B cell receptor sequence for thousands of cells. In addition to detection of antibody sequence and antigen specificity (including cross-reactivity) in proof-of-principle cell line experiments, I also validated the ability to use LIBRA-seq to analyze the antibody repertoire in multiple HIV-1 infection samples. In these experiments, we identified known HIV-1 broadly neutralizing antibodies and also discovered and characterized additional antibody candidates, including a novel broadly neutralizing antibody that had been missed by previous approaches. Together with the proof-of-principle experiment, this work demonstrated the feasibility and utility of LIBRA-seq as an antibody discovery tool and also contributed to the HIV-1 antibody field.

In Chapter 3, I describe a variety of projects in which I applied LIBRA-seq for antibody discovery. First, a coronavirus antibody discovery project is described, in which we identified antibodies that are cross-reactive to multiple coronaviruses, including

SARS-CoV-2 and SARS-CoV, utilizing LIBRA-seq. Epitope mapping and functional characterization revealed that these antibodies target multiple, distinct epitopes on the SARS-CoV-2 spike, modulate Fc effector functions, and reduce pathological burden *in vivo*. The recurrent nature of novel coronavirus strains, like SARS-CoV-2, highlights the critical need for broadly reactive therapeutics and vaccines, and the discovery of functional, cross-reactive antibodies targeting discrete structural domains has implications for rational vaccine design strategies, both for the current and potential future coronavirus outbreaks. In addition to using LIBRA-seq to study the antibody response to coronaviruses, I used a LIBRA-seq dataset from Chapter 2 to characterize a broadly reactive influenza antibody called 3602-1707, which binds to a variety of influenza hemagglutinin molecules from diverse strains, including hemagglutinin from group I and group II. 3602-1707 targets the hemagglutinin stem and also demonstrates neutralization of a variety of influenza strains. Identification and characterization of broadly neutralizing influenza antibodies has broad applications for therapeutic development and universal flu vaccine design strategies. Lastly, in this chapter, I also describe the identification and characterization of a cross-reactive antibody family that demonstrates binding and effector function activity for both HIV-1 and influenza virus. This antibody family was identified using LIBRA-seq from a dataset described in Chapter 2. Learning more about the observed cross-reactivity of these antibodies will enhance our knowledge of antibody specificity.

In Chapter 4, I describe how I have extended the LIBRA-seq technology to incorporate additional features. In particular, I describe LIBRA-seq with ligand blocking. In these experiments, a ligand and its cognate antigen are each labeled with an

oligonucleotide barcode, enabling the transformation of B cell receptor-antigen and antigen-ligand interactions into sequence-able events. As a proof of concept, I performed antibody discovery for SARS-CoV-2 specific antibodies from individuals who had recovered from SARS-CoV-2 infection by utilizing SARS-CoV-2 spike and its receptor ACE2 in an antigen screening library. LIBRA-seq with ligand blocking demonstrated high rates of neutralizing antibody identification, and the identified antibodies targeted diverse epitopes across the S protein, including a novel antibody epitope bridging the RBD and NTD that was characterized using cryogenic electron microscopy. Several antibodies also neutralized SARS-CoV-2 variants of concern. Inclusion of both SARS-CoV-2 spike and SARS-CoV spike in LIBRA-seq with ligand blocking led to identification of cross-reactive, ACE2 blocking antibodies, as well. The application of LIBRA-seq with ligand blocking can provide critical advantages for rapid development of therapeutic and preventive countermeasures. This technology presents a general platform with applications to areas where targeting the disruption of antigen-ligand interaction is a prime therapeutic goal.

In Chapter 5, I provide a summary of the data presented in this dissertation and also offer some caveats and conclusions. Furthermore, I describe a variety of future directions that could expand upon my thesis work. This includes further characterization of select antibodies, additional extensions of the LIBRA-seq technology, and novel applications of LIBRA-seq for immune repertoire analysis. Overall, I hope that the interrogation of antibody-antigen interactions using a sequencing-based readout will be a transformative tool for vaccine design and antibody discovery efforts against a variety of antigen targets.

1.2 The Immune System

The human immune system is a complex network of cells that work in concert to defend and protect the body from invading pathogens[1]. There are two main features of the immune system: the innate immune response and the adaptive immune response[2]. The innate immune response is comprised of a variety of cells that identify and target pathogens at the site of infection, including neutrophils and macrophages[3]. Once invading microbes are recognized, a variety of sensors are released to advance the immune response or kill infected cells. In addition to the innate immune response, adaptive immunity is another important component of the immune system and is comprised of two main cell types: B cells and T cells[1]. A key feature of adaptive immunity is memory[4-6].

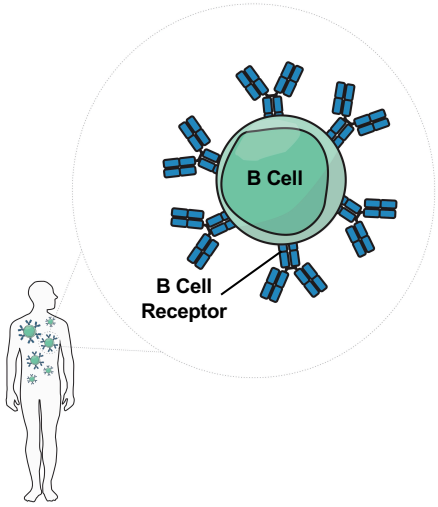


Figure 1.1 B cells. B cells are part of the adaptive immune system, targeting pathogens for destruction. Humans have billions of B cells that contribute to their immune defenses. B cells recognize invading microbes with their B cell receptor, which is highly variable.

1.3 B Cells and Antibodies

B cells are an integral part of the adaptive immune response, helping to fight infectious microbes (**Figure 1.1**)[4, 7]. B cells mature in the bone marrow and later continue to develop in lymph nodes where they encounter antigens. An important feature in B cell development is the generation of the B cell receptor, which recognizes antigens in their native form[8]. Once a B cell engages with an antigen through the B cell receptor, T helper cells can activate the B cell through secondary signals, promoting a process called somatic hypermutation[2, 9-17]. In this process, antibodies will be selected for if they gain beneficial mutations that increase their affinity for their cognate antigen. Once activated, B cells proliferate and can differentiate into plasma cells, which secrete antibodies, or memory B cells[12, 18, 19]. In the case of re-exposure to a pathogen, memory cells are primed for a rapid immune response[4, 20-22].

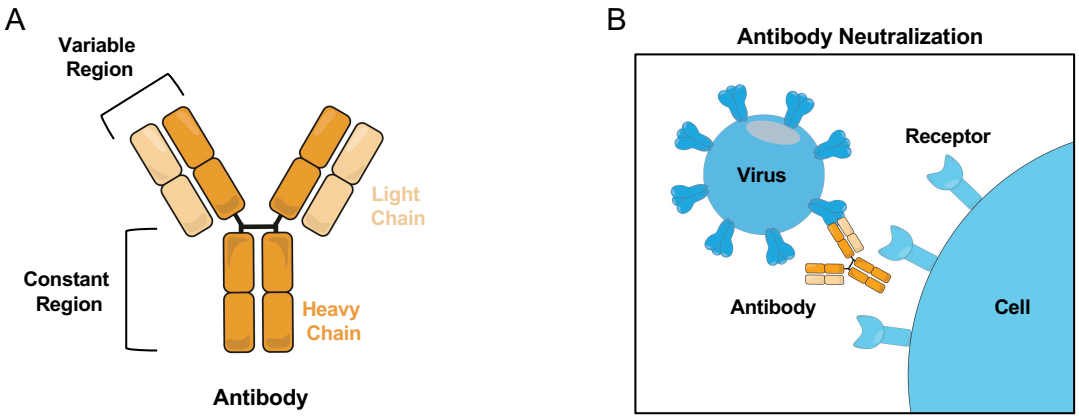


Figure 1.2 Antibody composition and neutralization mechanism. (A) A secreted form of the B cell receptor, antibodies are Y-shaped proteins made by B cells. They are made of a heavy chain (shown in dark orange) and a light chain (shown in light orange). The variable region is responsible for antigen binding, and the constant region designates the isotype and contributes to effector functionality. (B) One main function of antibodies is neutralization. When an antibody binds its cognate antigen (for example, a glycoprotein on the surface of a virus), it can prevent the pathogen from interacting with the target cell and gaining host cell entry.

A secreted form of the B cell receptor, antibodies are proteins that can bind to pathogens and exert a wide array of immune functions. Composed of a paired heavy chain and light chain, antibodies contain a variable region and a constant region (**Figure 1.2A**)[23-25]. The variable region of antibodies is responsible for binding antigens, and in order to counteract the wide diversity of pathogens that an individual might encounter in their lifetime, the sequence space of antibodies is extremely large ($>10^{12}$)[26-29]. There are a number of mechanisms that contribute to this, including VDJ recombination and somatic hypermutation. In VDJ recombination, variable (V), diversity (D), and joining (J) gene segments are randomly recombined in order to form both the heavy chain and the light chain to make up an antibody[10, 30, 31]. This process occurs during B cell development and also occurs for T cells and the assembly of T cell receptor[32, 33]. Additionally, antibodies can acquire mutations to further develop their antigen specificity through somatic hypermutation which occurs in the germinal center[2, 9-14, 34, 35]. The constant region of an antibody determines its isotype, which contributes to effector functions and gives rise a variety of antibody classes[36, 37]. There are five different antibody isotypes, including IgM, IgD, IgG, IgA and IgE. Antibodies exert a variety of functions including neutralization (**Figure 1.2B**), opsonization, complement activation, antibody dependent cellular phagocytosis, and antibody dependent cellular cytotoxicity[36, 38, 39]. Many vaccines rely on immune memory, specifically antibody specificity and functionality, for their utility[4, 40-45].

1.4 Human Immunodeficiency Virus

Human Immunodeficiency Virus (HIV) is an enveloped retrovirus with two copies of positive sense, single strand RNA. Since its discovery in 1981, HIV-1 has posed a significant global health threat, and the search for an HIV-1 vaccine continues to be a top priority. In 2020, 37.7 million people were living with HIV-1, and despite existence of antiretroviral therapy, many people do not have access to therapeutics[46]. In 2020, 27% of people living with HIV-1 did not have access to antiretroviral therapy, and 15% of pregnant women living with HIV-1 did not have access to antiretroviral medicine to prevent mother-to-child transmission[46]. The high levels of diversity and the pronounced mutability of the virus have made the search for an HIV-1 vaccine elusive[47-52]. In an effort to address the difficulty in vaccine design, recent studies have focused on understanding the immune response to HIV-1 infection to inform effective vaccine design[53-55].

HIV-1 attacks the immune system and without treatment, the loss of $CD4^+$ T cells leads to acquired immunodeficiency syndrome (AIDS) and ultimately results in patient death. Although antiretroviral therapy is effective against HIV-1, patients must remain on treatment for life, and many people lack access to or cannot afford these drugs. To fight HIV-1 infection, the immune system must respond to mutations in the virus that allow HIV-1 to evade immune pressures. Understanding the immune response to HIV-1 infection can help with the development of vaccine immunogens and regimens. In particular, antibodies are produced that can recognize and neutralize HIV-1, and the target of these antibodies is the HIV-1 envelope trimer on the surface of the virion[56-59]. The HIV-1 envelope trimer is a trimer of heterodimers (gp41 and gp120) responsible for binding to

receptors on host cells for viral entry. Antibodies develop specificity to the virus through gene recombination and affinity maturation. Some individuals develop broadly neutralizing antibodies that can neutralize diverse HIV-1 viral strains, and many of these antibodies have been shown to be effective in therapy and prevention strategies[59-68]. A goal of reverse vaccinology is to characterize the interactions between naturally elicited antibodies and their targets on HIV-1 in order to develop vaccine strategies to guide the immune system towards elicitation of broadly neutralizing antibodies[53, 69].

1.5 Influenza Virus

Influenza virus has a segmented RNA genome composed of eight single stranded, negative sense RNA molecules[70]. A zoonotic pathogen, influenza virus circulates in birds, pigs, horses, and humans. Worldwide, there are millions of severe cases of influenza and hundreds of thousands of deaths each year[70]. There are two main surface proteins on the influenza virion: hemagglutinin, which is responsible for binding to host sialic acids to gain cell entry, and neuraminidase, which aids in viral budding and release[71-75]. Two key mechanisms contribute to the ever-changing landscape of influenza: antigenic drift and antigenic shift[75, 76] Over time, antigenic drift occurs when immune selection from antibody pressure leads to mutagenic changes in the virus, causing new strains to emerge. For antigenic shift, a viral re-assortment occurs, leading to the creation of new influenza subtypes. To tackle this complicated problem, a focus on universal flu vaccine strategies has continued to increase, including research on broadly neutralizing antibodies[45, 77-81]. A few classes of broadly neutralizing

antibodies have been identified, including group 1 reactive, group 2 reactive, and group 1/2 cross-reactive[71, 77, 82-88].

1.6 Coronaviruses

Coronaviruses are positive sense, single strand RNA viruses with very large (approximately 30 kb) genomes[89]. Coronaviruses are extremely diverse, and infect a variety of mammalian and avian species, posing risks for zoonotic spillover and disease emergence[90-95]. Seven coronaviruses have crossed the zoonotic barrier to infect humans, including endemic strains HCoV-229E, HCoV-NL63, HCoV-HKU1 and HCoV-OC43[92]. These viruses are responsible for “common cold” upper respiratory infections. SARS-CoV, MERS-CoV, and SARS-CoV-2 are highly pathogenic strains that have led to pandemics. Since its emergence in 2019, SARS-CoV-2 has infected hundreds of millions of people, caused over five million deaths, and devastated economies and healthcare systems worldwide. The spike protein on the surface of the virion is responsible for cell entry and viral fusion through its interaction with host receptor ACE2[89, 96]. The spike protein is made of 2 subunits, S1 which is composed of the receptor binding domain and the N-terminal domain, and the S2 subdomain[97]. Because of this, spike is a main target of neutralizing antibodies and is the basis for many vaccine formulations[98, 99].

1.7 Antibody Discovery Technologies

Antibodies are an integral part of the immune response, and studying antibody repertoires can allow for greater understanding of adaptive immunity and also guide vaccine and therapeutic development. Two critical pieces of information needed from a screening experiment include the paired heavy/light chain antibody sequence and the antigen reactivity of a given B cell (**Figure 1.3**)[41, 43, 44, 100-102]. There are a variety of methods that are used to probe the B cell repertoire for discovery of antigen-specific antibodies, including single-cell sorting, screening immortalized B cells, and B cell culture[57, 64, 102-108]. Furthermore, yeast display systems can be used for screening human heavy and light chain pairs[109]. However, these technologies are often limited in the ability to characterize antibody-antigen reactivity for many cells and antigens simultaneously. Advances in next-generation sequencing (NGS) allow for high-throughput examination of the antibody repertoire, including paired heavy and light

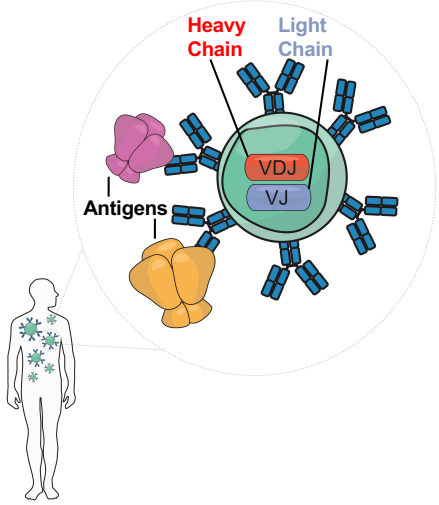


Figure 1.3 Screening B cells. When performing an antibody discovery experiment, there are two main pieces of information needed: the B cell receptor sequence (including the heavy chain and the light chain) and the antigen reactivity of that particular cell. There is a need to glean this information simultaneously from thousands of B cells in a single screening experiment.

chain sequencing[110-115]. However, little is known about the antigen specificity of the antibodies identified in these datasets.

To address these limitations and inefficiencies in the screening process, I helped develop a novel technology called LIBRA-seq (Linking B cell receptor to antigen specificity through sequencing) which enables the high-throughput mapping of paired heavy and light chain antibody sequences to antigen reactivity using single cell sequencing[116]. LIBRA-seq relies on the use of oligo-barcoded antigens, which can be used to bioinformatically map antigen reactivity from single-cell sequencing data. In my thesis work, I have had the opportunity to build and use technologies for analysis of antigen-antibody interactions at the repertoire level, specifically in the context of infectious disease. The development and utilization of the LIBRA-seq technology forms the basis for the following dissertation. The work in this dissertation includes the development and validation of LIBRA-seq technology, its application to antibody discovery for HIV-1, influenza, and coronaviruses, and the extension of the technology for functional antibody discovery in the context of SARS-CoV-2. I have identified, expressed, and characterized antibody candidates to a variety of viral pathogens, including HIV-1, influenza, and SARS-CoV-2. From diverse patient samples, we have discovered novel broadly neutralizing antibodies. Defining the epitopes of these antibodies has broadened our knowledge of antibody-antigen interactions and elucidated novel antibody epitopes. Furthermore, I developed a second generation of the LIBRA-seq assay to incorporate functional readouts from sequencing experiments, and application of LIBRA-seq with ligand blocking resulted in high hit rates for identification of potent, neutralizing antibodies against SARS-CoV-2.

CHAPTER 2

LIBRA-seq Technology Development: High Throughput Mapping of B cell Receptor Sequences to Antigen Specificity

This chapter is adapted from the published manuscript:

Setliff I*, **Shiakolas A***, Pilewski K, Murji AA, Mapengo RE, Janowska K, Richardson S, Oosthuysen C, Raju N, Ronsard L, Kanekiyo M, Qin J, Kramer KJ, Greenplate AR, McDonnell WJ, Graham BS, Connors M, Lingwood D, Acharya P, Morris L, Georgiev IS. (2019). High-throughput mapping of B-cell receptor sequences to antigen specificity, *Cell,* doi: 10.1016/j.cell.2019.11.003. (*Equal Contribution)

Contributions:

Kelsey Pilewski performed cell staining and fluorescence activated cell sorting. Amyn Murji assisted with protein purification. Rutendo Mapengo performed neutralization assays. Katarzyna Janowska performed surface plasmon resonance and binding competition assays. Simone Richardson performed Fc effector function experiments. Charissa Oosthuysen performed neutralization assays. Nagarajan Raju assisted with bioinformatic analysis and sequence analysis. Larance Ronsard created the Ramos cell lines. Masaru Kanekiyo prepared antigen proteins for use in LIBRA-seq experiments. Juliana Qin expressed and tested proteins for antigenicity. Kevin Kramer expressed and

purified antibodies. Allie Greenplate assisted with flow cytometry and graphics. Barney Graham provided antigen proteins for use in LIBRA-seq experiments. Mark Connors provided donor PBMCs for use in LIBRA-seq experiments. Daniel Lingwood provided Ramos cell lines for use in LIBRA-seq experiments. Priyamvada Acharya helped with binding assays and experimental design. Lynn Morris oversaw neutralization experiments and Fc effector function experiments. Ivelin Georgiev oversaw development of LIBRA-seq and accompanying experiments and data analysis. Ian Setliff performed experimental design, data processing, bioinformatic analysis, sequencing analysis, data transformations, creation of phylogenetic trees and figure generation. I performed experimental design, antigen expression, purification, and quality control. I performed oligo-labelling of antigens and quality control, cell line maintenance, cell staining, fluorescence activated cell sorting, antibody expression and purification, ELISA binding assays, and figure generation. Ian Setliff, Ivelin Georgiev, and I wrote the manuscript.

2.1 Introduction

B-cell receptor (BCR) sequencing is a powerful tool for interrogating immune responses to infection and vaccination, but it provides limited information about the antigen specificity of the sequenced BCRs. Here, we present LIBRA-seq (LInking B-cell Receptor to Antigen specificity through sequencing), a technology for high-throughput mapping of paired heavy-/light-chain BCR sequences to their cognate antigen specificities. B cells are mixed with a panel of DNA-barcoded antigens, such that both the antigen barcode(s) and BCR sequence are recovered via single-cell next-generation sequencing. Using LIBRA-seq, we mapped the antigen specificity of thousands of B cells from two HIV-infected subjects. The predicted specificities were confirmed for a number of HIV- and influenza-specific antibodies, including known and novel broadly neutralizing antibodies. LIBRA-seq will be an integral tool for antibody discovery and vaccine development efforts against a wide range of antigen targets.

The antibody repertoire – the collection of antibodies present in an individual – responds efficiently to invading pathogens due to its exceptional diversity and ability to fine-tune antigen specificity via somatic hypermutation[11, 26, 27]. This antibody repertoire is a rich source of potential therapeutics, but its size makes it difficult to examine more than a small cross-section of the total repertoire[43, 100-102]. Historically, a variety of approaches have been developed to characterize antigen-specific B cells in human infection and vaccination samples. The methods most frequently used include single-cell sorting with fluorescent antigen baits[57, 64], screens of immortalized B cells[103, 104], and B cell culture[105-108]. However, these methods to couple functional screens with variable

heavy (V_H) and variable light (V_L) immunoglobulin gene sequences are low throughput; generally, individual B cells can only be screened against a few antigens simultaneously.

Recent advances in next-generation sequencing (NGS) enable high-throughput interrogation of antibody repertoires at the sequence level, including paired heavy and light chains[110, 111, 113]. However, annotation of NGS antibody sequences for their cognate antigen partner(s) generally requires synthesis, production and characterization of individual recombinant monoclonal antibodies[117, 118]. Recent efforts to develop new antibody screening technologies have sought to overcome throughput limitations while still uniting antibody sequence and functional information. For example, natively-paired human BCR heavy and light chain amplicons can be expressed and screened as Fab[102] or scFV[109] in a yeast display system. Although these various antibody discovery technologies have led to the identification of potently neutralizing antibodies, they remain limited by the number of antigens against which single cells can simultaneously be screened efficiently.

Inspired by previous methods combining surface protein marker detection with single-cell RNA sequencing[112, 115], we developed LIBRA-seq (LInking B-cell Receptor to Antigen specificity through sequencing) to simultaneously recover both antigen specificity and paired heavy and light chain BCR sequence. LIBRA-seq is a next-generation sequencing-based readout for BCR-antigen binding interactions that utilizes oligonucleotides (oligos) conjugated to recombinant antigens. Antigen barcodes are recovered during paired-chain BCR sequencing experiments and bioinformatically mapped to single cells. To demonstrate the utility of LIBRA-seq, we applied the method

to peripheral blood mononuclear cell (PBMC) samples from two HIV-infected subjects, and from these, we successfully identified HIV- and influenza-specific antibodies, including both known and novel broadly neutralizing antibody (bNAb) lineages. LIBRA-seq is high-throughput, scalable, and applicable to many targets. This single, integrated assay enables the mapping of monoclonal antibody sequences to panels of diverse antigens theoretically unlimited in number, and facilitates the rapid identification of cross-reactive antibodies that may serve as therapeutics or vaccine templates.

2.2 Results

LIBRA-seq Method and Validation

LIBRA-seq transforms antibody-antigen interactions into sequencing-detectable events by conjugating DNA-barcoded oligos to each antigen in a screening library. All antigens are labeled with the same fluorophore, which enables sorting of antigen-positive B cells by fluorescence-activated cell sorting (FACS) before encapsulation of single B cells via droplet microfluidics. Antigen barcodes and BCR transcripts are tagged with a common cell barcode from bead-delivered oligos, enabling direct mapping of BCR sequence to antigen specificity (**Figure 2.1A**).

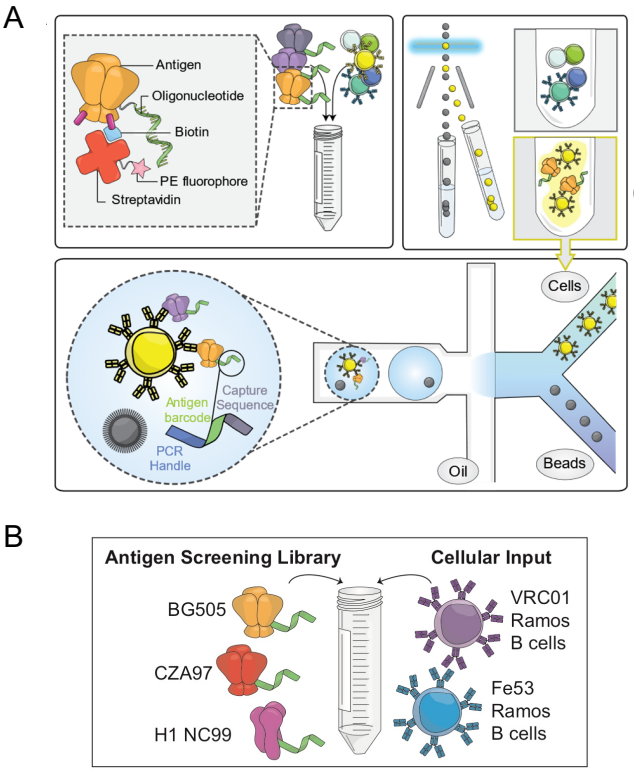
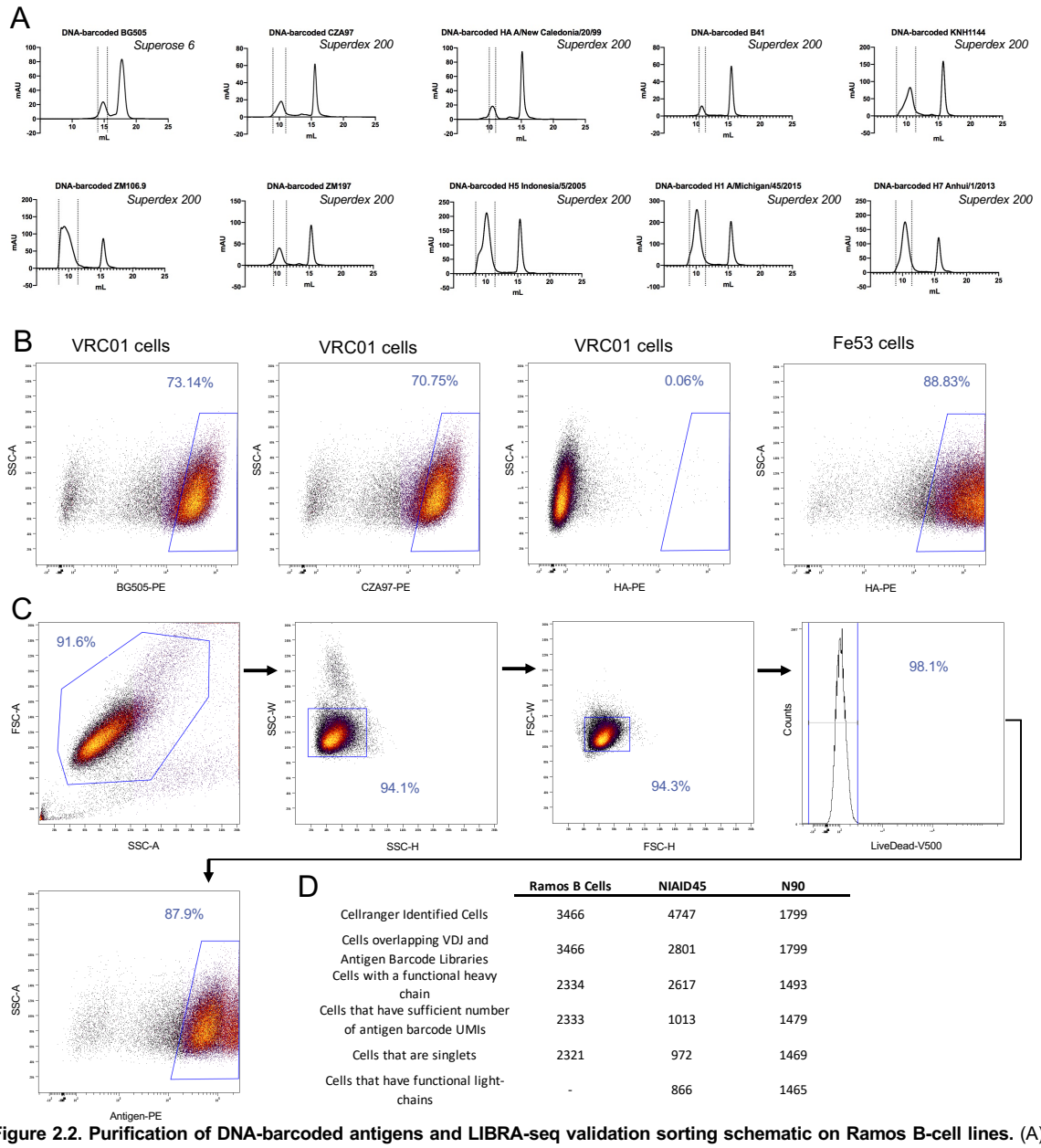


Figure 2.1. LIBRA-seq assay schematic. (A) Schematic of LIBRA-seq assay. (Top left) Fluorescently-labelled, DNA-barcoded antigens are used to (top right) sort antigen-positive B cells before (bottom) co-encapsulation of single B cells with bead-delivered oligos using droplet microfluidics. Bead-delivered oligos index both cellular BCR transcripts and antigen barcodes during reverse transcription, enabling direct mapping of BCR sequence to antigen specificity following sequencing. Note: elements of the depiction are not shown to scale, and the number and placement of oligonucleotides on each antigen can vary. (B) The assay was initially validated on Ramos B-cell lines expressing BCR sequences of known neutralizing antibodies VRC01 and Fe53 with a three-antigen screening library: BG505, CZA97 and H1 A/New Caledonia/20/99. Figure made by Allison Greenplate and adapted from Setliff and Shiakolas et al., *Cell*, 2019.

To test the ability of LIBRA-seq to accurately unite BCR sequence and antigen specificity, we devised a proof-of-principle mapping experiment using two Ramos B-cell lines with different BCR sequences and antigen specificities[119]. These engineered B-cell lines do not display endogenous BCR and instead express specific, user-defined surface IgM BCR sequences[119]. To that end, we chose two well-characterized BCRs: VRC01, a CD4-binding site-directed HIV-1 bNAb[64], and Fe53, a bNAb recognizing the stem of group 1 influenza hemagglutinins (HA)[120]. We mixed these two populations of B-cell lines at a 1:1 ratio and incubated them with three unique DNA-barcoded antigens: two trimeric HIV-1 Env proteins from strains BG505 and CZA97[121-123], and trimeric hemagglutinin from strain H1 A/New Caledonia/20/1999[124] (**Figure 2.1B**; **Figure 2.2A-C**).



D

	Ramos B Cells	NIAID45	N90
Cellranger Identified Cells	3466	4747	1799
Cells overlapping VDJ and Antigen Barcode Libraries	3466	2801	1799
Cells with a functional heavy chain	2334	2617	1493
Cells that have sufficient number of antigen barcode UMIs	2333	1013	1479
Cells that are singlets	2321	972	1469
Cells that have functional light-chains	-	866	1465

Figure 2.2. Purification of DNA-barcoded antigens and LIBRA-seq validation sorting schematic on Ramos B-cell lines. (A) After barcoding each antigen with a unique oligonucleotide, antigen-oligo complexes are run on size exclusion chromatography to remove excess, unconjugated oligonucleotide from the reaction mixture. DNA-barcoded BG505 was run on the Superose 6 Increase 10/300 GL column and all other DNA barcoded antigens were run on the Superdex 200 Increase 10/300 GL on the AKTA FPLC system. For size exclusion chromatography, dotted lines indicate DNA barcoded antigens and fractions taken. The second peak indicates excess oligonucleotide from the conjugation reaction. (B) Binding of VRC01 or Fe53 Ramos B-cell lines to DNA-barcoded, fluorescently labeled antigens via flow cytometry. VRC01 cells bound to DNA-barcoded BG505-PE, DNA-barcoded CZA97-PE, and not DNA barcoded H1 A/New Caledonia/20/99-PE. Fe53 cells bound to DNA-barcoded H1 A/New Caledonia/20/99-PE. (C) Gating scheme for fluorescence activated cell sorting of Ramos B-cell lines. VRC01 and Fe53 Ramos B cells were mixed in a 1:1 ratio and then stained with LiveDead-V500 and a DNA-barcoded antigen screening library consisting of BG505-PE, CZA97-PE, and H1 A/New Caledonia/20/99-PE. Gates as drawn are based on gates used during the sort, and percentages from the sort are listed. (D) For each experiment, the categorization of the number of Cell Ranger-identified (10X Genomics) cells after sequencing is shown. Each category (row) is a subset of cells of the previous category (row). Figure adapted from Setliff and Shiakolas et al., *Cell*, 2019.

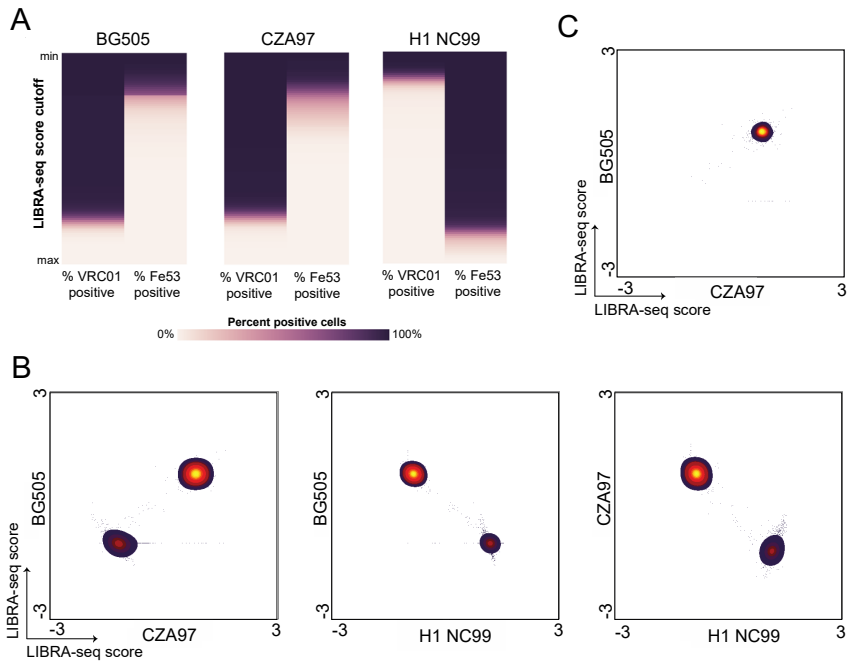


Figure 2.3. LIBRA-seq assay validation. (A) Between the minimum (y-axis, top) and maximum (y-axis, bottom) LIBRA-seq score for each antigen, different cutoffs were tested for their ability to accurately classify each VRC01 cell and Fe53 cell as antigen-positive or -negative, where antigen-positive is defined as having a LIBRA-seq score greater than or equal to the cutoff being evaluated, and antigen-negative is defined as having a LIBRA-seq score below the cutoff. A series of 100 cutoff thresholds between the respective minimum and maximum antigen-specific LIBRA-seq scores were evaluated. At each cutoff, the percent of total VRC01 cells (left column of each antigen subpanel) and percent of total Fe53 cells (right columns) that were classified as positive for a given antigen is represented on a white (0%) to dark purple (100%) color scale. (B) The LIBRA-seq score for each pair of antigens for each B cell was plotted. Each axis represents a range of LIBRA-seq scores for each antigen. Density of total cells is shown, with purple to yellow indicating lowest to highest number of cells, respectively. (C) The LIBRA-seq score for BG505 (y-axis) and CZA97 (x-axis) for each VRC01 B cell was plotted. Each axis represents a range of LIBRA-seq scores for each antigen. Density of total cells is shown, with purple to yellow indicating lowest to highest number of cells, respectively. Figure 2.3.A made by Ian Setliff. Figure adapted from Setliff and Shiakolas et al., *Cell*, 2019.

We recovered 2321 cells with BCR sequence and antigen mapping information, highlighting the high throughput potential of LIBRA-seq (**Figure 2.2D**). For each cell, the LIBRA-seq scores for each antigen in the screening library were computed as a function of the number of unique molecular identifiers (UMIs) for the respective antigen barcode (Methods). The LIBRA-seq scores of each individual antigen reliably categorized Ramos B cells by their specificity (**Figure 2.3A**). Overall, cells fell into two major populations based on their LIBRA-seq scores, and we did not observe cells that were cross-reactive for influenza HA and HIV-1 Env (**Figure 2.3B**). Further, VRC01 Ramos B cells bound both BG505 and CZA97 with a high correlation between the scores for

these two antigens (Pearson's r=0.84), demonstrating that LIBRA-seq readily identifies B cells that bind to multiple HIV-1 antigens (**Figure 2.3C**).

Isolation of Antibodies from a Known HIV bNAb Lineage

We next used LIBRA-seq to analyze the antibody repertoire of donor NIAID45, who had been living with HIV-1 without antiretroviral therapy for approximately 17 years

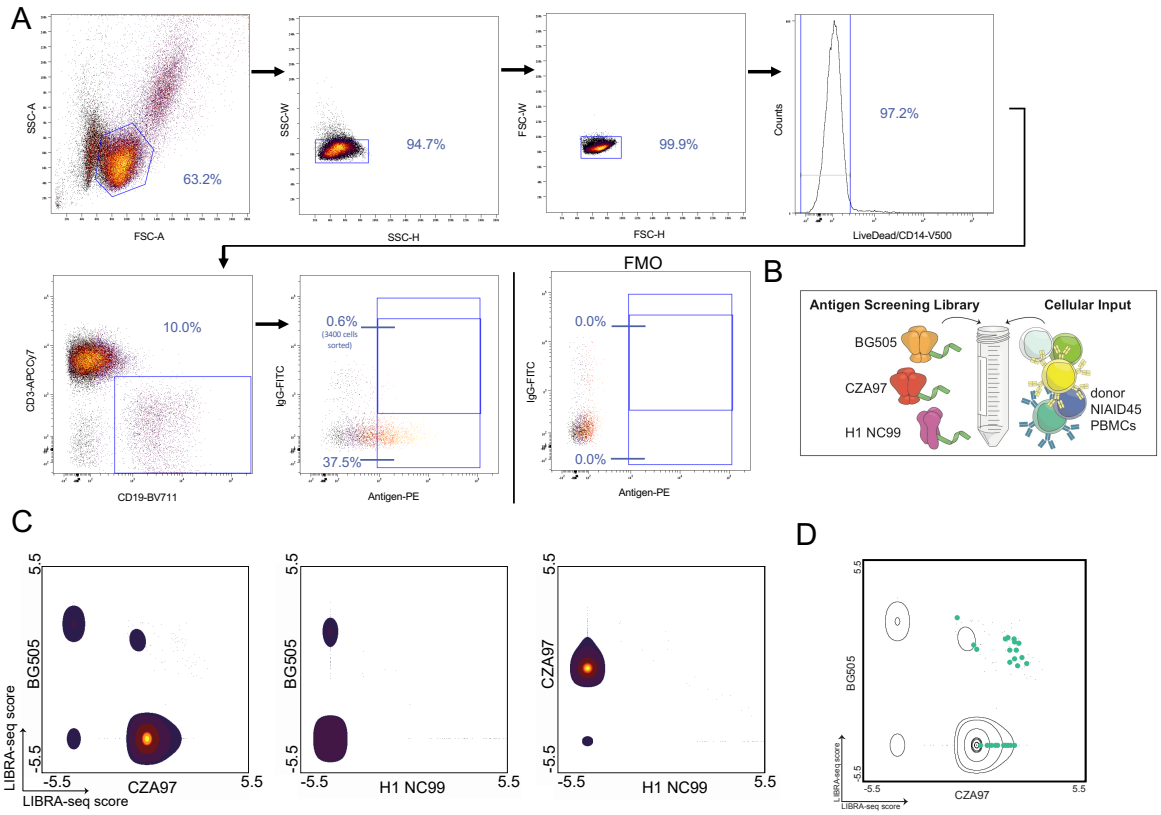


Figure 2.4. LIBRA-seq applied to a human B cell sample from HIV-infected donor NIAID45. (A) Gating scheme for fluorescence activated cell sorting of donor NIAID45 PBMCs. Cells were stained with LiveDead-V500, CD14-V500, CD3-APCCy7, CD19-BV711, IgG-FITC, and a DNA-barcoded antigen screening library consisting of BG505-PE, CZA97-PE, and H1 A/New Caledonia/20/99-PE. Gates as drawn are based on gates used during the sort, and percentages from the sort are listed. These plots show a starting number of 50,187 total events. Due to the visualization parameters, 18 IgG-positive, antigen-positive cells are displayed, but 3400 IgG-positive, antigen-positive cells were sorted and supplemented with 13,000 antigen-positive B cells for single-cell sequencing. A small aliquot of donor NIAID45 PBMCs were used for fluorescence minus one (FMO) staining, and were stained with the same antibody panel as listed above with the exception of the HIV-1 and influenza antigens. (B) LIBRA-seq experiment setup consisted of three antigens in the screening library: BG505, CZA97, and H1 A/New Caledonia/20/99, and the cellular input was donor NIAID45 PBMCs. (C) After bioinformatic processing and filtering of cells recovered from single-cell sequencing, the LIBRA-seq score for each antigen was plotted (total = 866). Each axis represents a range of LIBRA-seq scores for each antigen. Density of total cells is shown, with purple to yellow indicating lowest to highest number of cells, respectively. (D) LIBRA-seq scores for BG505 (y-axis) and CZA97 (x-axis) are shown. Each axis represents the range of LIBRAseq scores for each antigen. Density of total cells is shown. Overlaid on the density plot are the 29 VRC01 lineage members (dots) indicated in light green. Figure adapted from Setliff and Shiakolas et al., *Cell*, 2019.

at the time of sample collection. This sample was selected as an appropriate target for LIBRA-seq analysis because a large lineage of HIV-1 bNAbs had been identified previously from this donor[64, 105, 125]. This lineage

consists of the prototypical bNAb VRC01, as well as multiple clades of clonally related bNAbs with diverse neutralization phenotypes[125]. We used the same BG505, CZA97, and H1 A/New Caledonia/20/99 antigen screening library as in the Ramos B-cell line experiment, and recovered paired V_H:V_L antibody sequences with antigen mapping for 866 cells (**Figure 2.4A-B, Figure 2.2D**). These B cells exhibited a variety of LIBRA-seq scores among the three antigens (**Figure 2.4C**), as can be expected from a polyclonal sample possessing a wide diversity of B cell specificities and antigen affinities. The cells displayed a few discrete patterns based on their LIBRA-seq scores; generally, cells were either (1) $HA^{high}Env^{low}$ or (2) $HA^{low}Env^{high}$ (**Figure 2.4C**). Additionally, we observed cells that were double positive for both HIV-1 Env variants, BG505 and CZA97, suggesting HIV-1 strain cross-reactivity of these B cells (**Figure 2.4C-D**).

To further validate the utility of LIBRA-seq in monoclonal antibody isolation, we next sought to identify members of the VRC01 antibody lineage from the LIBRA-seq-identified antigen-specific B cells. We observed 29 BCRs that were clonally related to previously-identified members of the VRC01 lineage (**Figure 2.5**). All LIBRA-seq-identified BCRs had high levels of somatic hypermutation and utilized *IGHV1-2*02* along with the characteristic five-residue CDRL3 paired with *IGVK3-20* (Figure 2C). These B cells came from multiple known clades of the VRC01 lineage, with sequences with high identity and phylogenetic relatedness to lineage variants VRC01, VRC02,

VRC03, VRC07, VRC08, NIH45-46, and others (**Figure 2.5**). Of these, 25 (87%) had a high LIBRA-seq score for at least 1 HIV-1 antigen, three (10%) had mid-range scores (between 0 and 1) for at least 1 HIV-1 antigen, and only one of the VRC01 lineage B

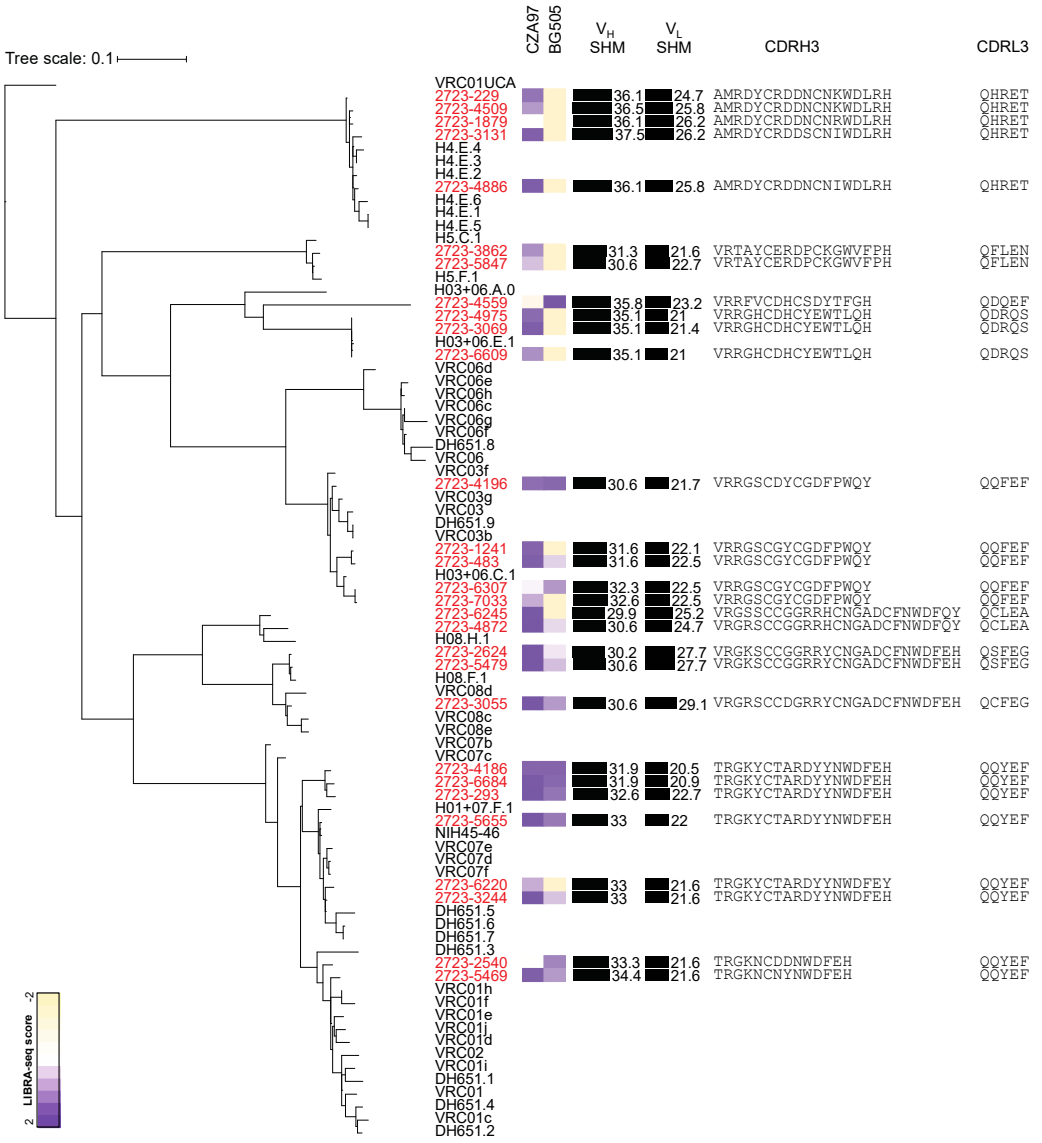
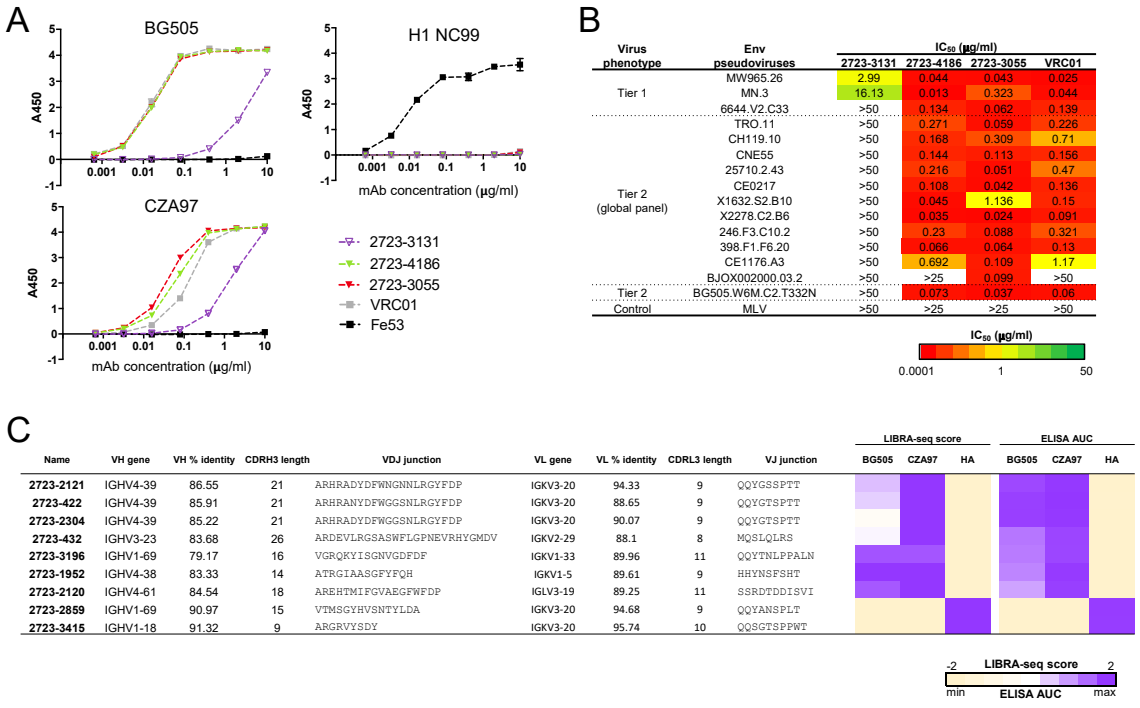


Figure 2.5. Identification of VRC01 Family Members using LIBRA-seq. 29 VRC01 lineage B cells were identified and examined for phylogenetic relatedness to known lineage members and for sequence features, with phylogenetic tree showing relatedness of previously identified VRC01 lineage members (black) and members newly identified using LIBRA-seq (red). Each row represents an antibody. Sequences were aligned using clustalW and a maximum likelihood tree was inferred using maximum likelihood inference. The resulting tree was visualized using an inferred VRC01 unmutated common ancestor (UCA) (accession MK032222) as the root. For each antibody isolated from LIBRA-seq, a heat map of the LIBRA-seq scores for each antigen (BG505, CZA97, and H1 A/New Caledonia/20/99) is shown; a scale of tan-white-purple represents LIBRA-seq scores from -2 to 0 to 2; in this heatmap, scores lower or higher than that range are shown as -2 and 2, respectively. Levels of somatic hypermutation (SHM) at the nucleotide level for the heavy and light chain variable genes as reported by IMGT are displayed as bars, with the numerical percentage value listed to the right of the bar; length of the bar corresponds to level of SHM. Amino acid sequences of the complementarity determining region 3 for the heavy chain (CDRH3) and the light chain (CDRL3) for each antibody are displayed. The tree was visualized and annotated using iTol (Letunic and Bork, 2019). Figure made by Ian Setliff and adapted from Setliff and Shiakolas et al., *Cell*, 2019.

cells had negative scores for both HIV-1 antigens (**Figure 2.5, Figure 2.4D**). We recombinantly expressed three of the LIBRA-seq-identified lineage members, named 2723-3055, 2723-4186 and 2723-3131, to confirm the ability of these antibodies to bind the screening probes. Antibody 2723-3131 showed binding to CZA97 and BG505 by enzyme linked immunosorbent assay (ELISA) (**Figure 2.6A**), and neutralized two Tier 1 viruses but no viruses on a global panel of representative HIV-1 strains (deCamp et al., 2014) (**Figure 2.6B**). Both 2723-3055 and 2723-4186 bound to BG505 and CZA97, and potently neutralized 12/12 and 11/12 viruses on a global panel, respectively (**Figure 2.6A-B**). Together, the results from the donor NIAID45 analysis suggest that the LIBRA-



Virus phenotype	Env pseudoviruses	IC50 (µg/ml) 2723-3131	2723-4186	2723-3055	VRC01
Tier 1	MW965.26	2.99	0.044	0.043	0.025
	MN.3	16.13	0.013	0.323	0.044
	6644.V2.C33	>50	0.134	0.062	0.139
Tier 2 (global panel)	TRO.11	>50	0.271	0.059	0.226
	CH119.10	>50	0.168	0.309	0.71
	CNE55	>50	0.144	0.113	0.156
	25710.2.43	>50	0.216	0.051	0.47
	CE0217	>50	0.108	0.042	0.136
	X1632.S2.B10	>50	0.045	1.136	0.15
	X2278.C2.B6	>50	0.035	0.024	0.091
	246.F3.C10.2	>50	0.23	0.088	0.321
	398.F1.F6.20	>50	0.066	0.064	0.13
	CE1176.A3	>50	0.692	0.109	1.17
	BJOX002000.03.2	>50	>25	0.099	>50
Tier 2	BG505.W6M.C2.T332N	>50	0.073	0.037	0.06
Control	MLV	>50	>25	>25	>50

IC50 (µg/ml): 0.0001 – 1 – 50

Name	VH gene	VH % identity	CDRH3 length	VDJ junction	VL gene	VL % identity	CDRL3 length	VJ junction	LIBRA-seq score BG505	CZA97	HA	ELISA AUC BG505	CZA97	HA
2723-2121	IGHV4-39	86.55	21	ARHRADYDFWNGNNLRGYFDP	IGKV3-20	94.33	9	QQYGSSPTT						
2723-422	IGHV4-39	85.91	21	ARHRANYDFWGGSNLRGYFDP	IGKV3-20	88.65	9	QQYGTSPTT						
2723-2304	IGHV4-39	85.22	21	ARHRADYDFWGGSNLRGYFDP	IGKV3-20	90.07	9	QQYGTSPTT						
2723-432	IGHV3-23	83.68	26	ARDEVLRGSASWFLGPNEVRHYGMDV	IGKV2-29	88.1	8	MQSLQLRS						
2723-3196	IGHV1-69	79.17	16	VGRQKYISGNVGDFDF	IGKV1-33	89.96	11	QQYTNLPPALN						
2723-1952	IGHV4-38	83.33	14	ATRGIAASGFYFQH	IGKV1-5	89.61	9	HHYNSFSHT						
2723-2120	IGHV4-61	84.54	18	ARFHTMIFGVAEGFWFDP	IGLV3-19	89.25	11	SSRDTDDISVI						
2723-2859	IGHV1-69	90.97	15	VTMSGYHVSNTYLDA	IGKV3-20	94.68	9	QQYANSPLT						
2723-3415	IGHV1-18	91.32	9	ARGRVYSDY	IGKV3-20	95.74	10	QQSGTSPPWT						

LIBRA-seq score: -2 (min) – 2 (max); ELISA AUC: min – max

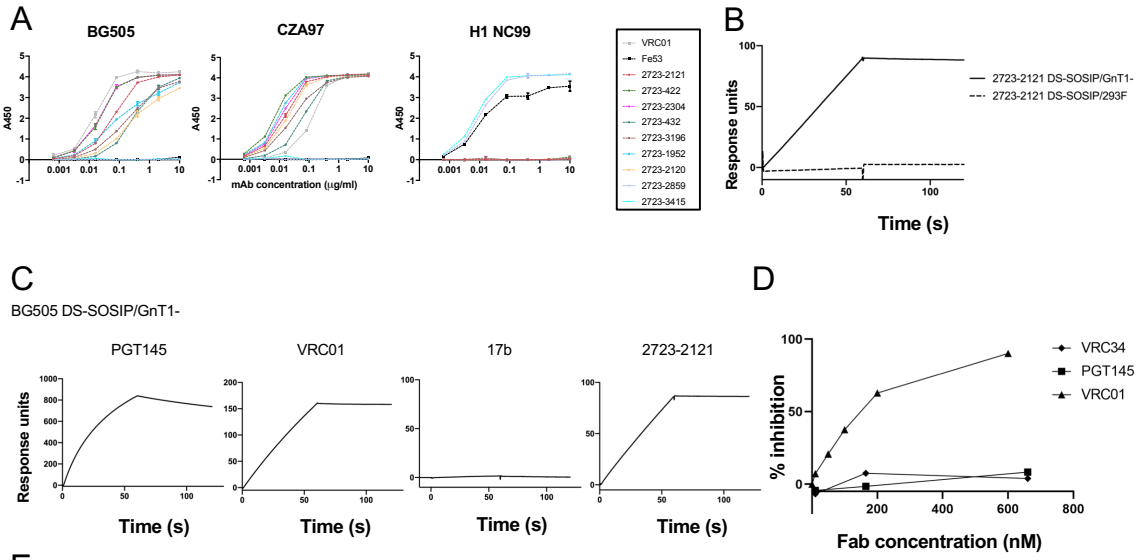
Figure 2.6. Characterization of LIBRA-seq-identified antibodies from donor NIAID45. (A) Antigen specificity as predicted by LIBRA-seq was validated by ELISA for a subset of monoclonal antibodies belonging to the VRC01 lineage. Data are represented as mean ± SEM for one ELISA experiment. ELISAs were repeated 2 or more times. (B) Neutralization of Tier 1, Tier 2, and control viruses by VRC01 and newly identified VRC01 lineage members, 2723-3131, 2723-4186, and 2723-3055. (C) Sequence characteristics and antigen specificity of newly identified antibodies from donor NIAID45. Percent identity is calculated at the nucleotide level, and CDR length and sequences are noted at the amino acid level. LIBRA-seq scores for each antigen are displayed as a heatmap with a LIBRA-seq score of -2 displayed as light yellow, 0 as white, and a LIBRA-seq score of 2 as purple; in this heatmap, scores lower or higher than that range are shown as -2 and 2, respectively. ELISA binding data against BG505, CZA97, and H1 A/New Caledonia/20/99 is displayed as a heatmap of the AUC analysis calculated from the data in Figure 2.7A with AUC of 0 displayed as light yellow, 50% max as white, and maximum AUC as purple. ELISA data are representative from at least two independent experiments. Figure adapted from Setliff and Shiakolas et al., *Cell*, 2019.

seq platform can be successfully used to down-select cross-reactive bNAbs in prospective antibody discovery efforts.

Identification of Additional Broadly-Reactive Anti-HIV and Anti-Influenza Antibodies

To further validate the ability of LIBRA-seq to accurately identify antigen-specific B cells, we produced a number of putative HIV-specific and influenza-specific monoclonal antibodies from donor NIAID45 that did not belong to the VRC01 lineage. In particular, we recombinantly produced seven additional anti-HIV antibodies, three of which were clonally related (2723-2121, 2723-422, and 2723-2304) (**Figure 2.6C**). We selected these seven antibodies because all had high LIBRA-seq scores for at least one HIV-1 antigen. All seven antibodies bound the antigens by ELISA as expected based on the respective LIBRA-seq scores, with high similarity between the patterns of LIBRA-seq scores and ELISA area under the curve (AUC) values (**Figure 2.6C**, **Figure 2.7A**). We further characterized one of these antibodies, 2723-2121, and determined that it bound to a stabilized BG505 trimer[126] by surface plasmon resonance (SPR) (**Figure 2.7B-C**). Antibody 2723-2121 competed for trimer binding with VRC01 (**Figure 2.7D**), neutralized three Tier 1 pseudoviruses and 2/11 Tier 2 pseudoviruses from a global panel (**Figure 2.7E**), and mediated trogocytosis and antibody-dependent cellular phagocytosis (**Figure 2.7F**). In addition to the HIV-specific antibodies, we also characterized two antibodies predicted to have influenza specificity based on their LIBRA-seq scores for H1 A/New Caledonia/20/99 (**Figure 2.6C**). In agreement with the LIBRA-seq scores, antibodies 2723-2859 and 2723-3415 bound H1 A/New

Caledonia/20/99 but not BG505 or CZA97 by ELISA, confirming the ability of LIBRA-seq to simultaneously isolate antibodies to multiple diverse antigens (**Figure 2.6C**).



E

Virus phenotype	Env-pseudoviruses	IC_{50} (μg/ml) 2723-2121	VRC01
Tier 1	MW965.26	0.067	0.025
	MN.3	<0.020	0.044
	6644.V2.C33	2.441	0.139
Tier 2 (sorting antigens)	BG505.W6M.C2	>50	0.035
	BG505.W6M.C2.T332N	>50	0.06
	CZA097	>50	0.09
Tier 2 (global panel)	TRO.11	>50	0.226
	CH119.10	>50	0.71
	CNE55	>50	0.156
	25710.2.43	21.568	0.47
	CE0217	>50	0.136
	X1632.S2.B10	>50	0.15
	X2278.C2.B6	>50	0.091
	246.F3.C10.2	>50	0.321
	398.F1.F6.20	1.76	0.13
	CE1176.A3	>50	1.17
	BJOX002000.03.2	>50	>50
Control	MLV	>50	>50

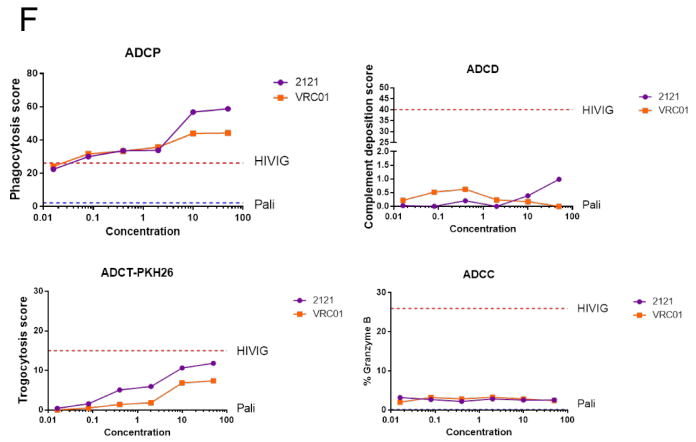


Figure 2.7. Characterization of antibodies from donor NIAID45. (A) Antigen specificity as predicted by LIBRA-seq was validated by ELISA for a variety of antibodies isolated from donor NIAID45. Antibodies were tested for binding to BG505, CZA97, and H1 A/New Caledonia/20/99. Data are represented as mean ± SEM for one ELISA experiment. ELISAs were repeated 2 or more times. (B) Binding of BG505 DS-SOSIP/GnT1- (resulting in Man5-enriched glycans) or BG505 DS-SOSIP/293F cells (complex glycans) to 2723-2121 IgG. (C) Binding of BG505 DS-SOSIP/GnT1- trimer to PGT145 IgG, VRC01 IgG, 17b IgG, and 2723-2121 IgG. (D) Inhibition of BG505 DS-SOSIP/GnT1- binding to 2723-2121 IgG in presence of VRC34 Fab (diamond), PGT145 Fab (square) and VRC01 Fab (triangle). (E) Neutralization of Tier 1, Tier 2, and control viruses by antibody 2723-2121 and VRC01. Results are shown as the concentration of antibody (in mg/ml) needed for 50% inhibition (IC50). (F) Levels of ADCP, ADCD, ADCT-PKH26, and ADCC displayed by antibody 2723-2121 compared to VRC01. HIVIG was used as a positive control and the anti-RSV mAb Palivisumab as a negative control. Figure 2.7B-D made by Acharya laboratory. Figure 2.7F made by Morris laboratory. Figure adapted from Setliff and Shiakolas et al., *Cell*, 2019.

Discovery of an HIV-1 bNAb Using a Nine-Antigen Screening Library

Having validated LIBRA-seq with three antigens on both Ramos B-cell lines and primary B cells from a patient sample, we sought to increase the number of antigens in the screening library. To that end, we screened the B-cell repertoire of NIAID donor N90 against nine antigens (**Figure 2.8A**). We selected this sample because a single broadly neutralizing antibody lineage (VRC38) targeting the V1/V2 epitope was isolated previously from this donor; however, the neutralization breadth of the VRC38 lineage

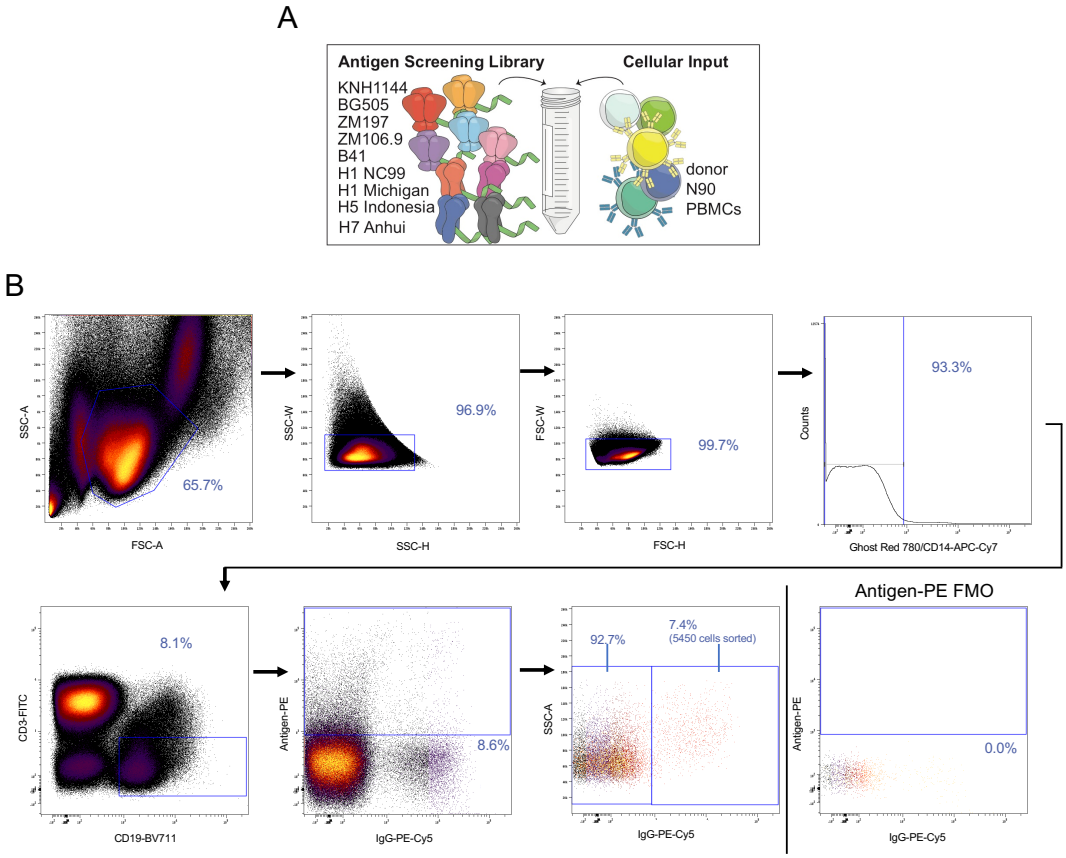


Figure 2.8. LIBRA-seq applied to a sample from NIAID donor N90. (A) LIBRA-seq experiment setup consisted of nine antigens in the screening library: 5 HIV-1 Env (KNH1144, BG505, ZM197, ZM106.9, B41), and 4 influenza HA (H1 A/New Caledonia/20/99, H1 A/Michigan/45/2015, H5 Indonesia/5/2005, H7 Anhui/1/2013), and the cellular input was donor N90 PBMCs. (B) Gating scheme for fluorescence activated cell sorting of donor N90 PBMCs. Cells were stained Ghost Red 780, CD14-APCCy7, CD3-FITC, CD19-BV711, and IgG-PECy5 along with a DNA-barcoded antigen screening library consisting of BG505-PE, KNH1144-PE, ZM197-PE, ZM106.9-PE, B41-PE, H1 A/New Caledonia/20/99-PE, H1 A/Michigan/45/2015-PE, H5 Indonesia/5/2005-PE, H7 Anhui/1/2013-PE. Gates as drawn are based on gates used during the sort, and percentages from the sort are listed. 5450 IgG-positive, antigen-positive cells were sorted and supplemented with 1480 IgG-negative, antigen-positive B cells for single-cell sequencing. A small aliquot of donor N90 PBMCs were used for fluorescence minus one (FMO) staining, and were stained with the same antibody panel as listed above without the antigen screening library. Figure adapted from Setliff and Shiakolas et al., *Cell*, 2019.

could not account for the full serum neutralization breadth[127, 128]. This suggested that there could be additional bNAb lineages present in the B cell repertoire of N90, and we reasoned that utilizing multiple SOSIP probes could help accelerate identification of such antibodies. Thus, we sought to determine whether LIBRA-seq could accomplish two goals: (1) to recover antigen-specific B cells from the VRC38 lineage, and (2) to identify new bNAbs that could neutralize viruses that are resistant to neutralization by the VRC38 lineage.

To increase the number of antigens in our screening library, we utilized a panel that consisted of five HIV-1 Env trimers from a variety of clades, BG505 (clade A), B41 (clade B), ZM106.9 (clade C), ZM197 (clade C) and KNH1144 (clade A)[122, 123, 129-132]; along with four diverse hemagglutinin trimers (H1 A/New Caledonia/20/99, H1 A/Michigan/45/2015, H5 A/Indonesia/5/2005, and H7 A/Anhui/1/2013) (**Figure 2.8A**). After applying LIBRA-seq to donor N90 PBMCs, we recovered paired V_H:V_L antibody sequences with antigen mapping for 1465 cells (**Figure 2.8B, Figure 2.2D**). Within this set of cells, we identified 18 B cells that were members of the VRC38 lineage (**Figure 2.9A**). Of these, 17 had high LIBRA-seq scores for at least one HIV-1 antigen, and one had no high LIBRA-seq scores but had a mid-range score for two SOSIPs (**Figure 2.9A**).

We next focused our analysis on the B cells with the highest LIBRA-seq scores in the N90 sample, focusing on cells that had LIBRA-seq scores for any antigen above one (901 cells) (**Figure 2.9B-C, 2.10**). We observed 32 cells that had high LIBRA-seq scores for three of the four influenza antigens (**Figure 2.9C**); we recombinantly produced one of these, 3602-1707, and confirmed broad influenza recognition, with high

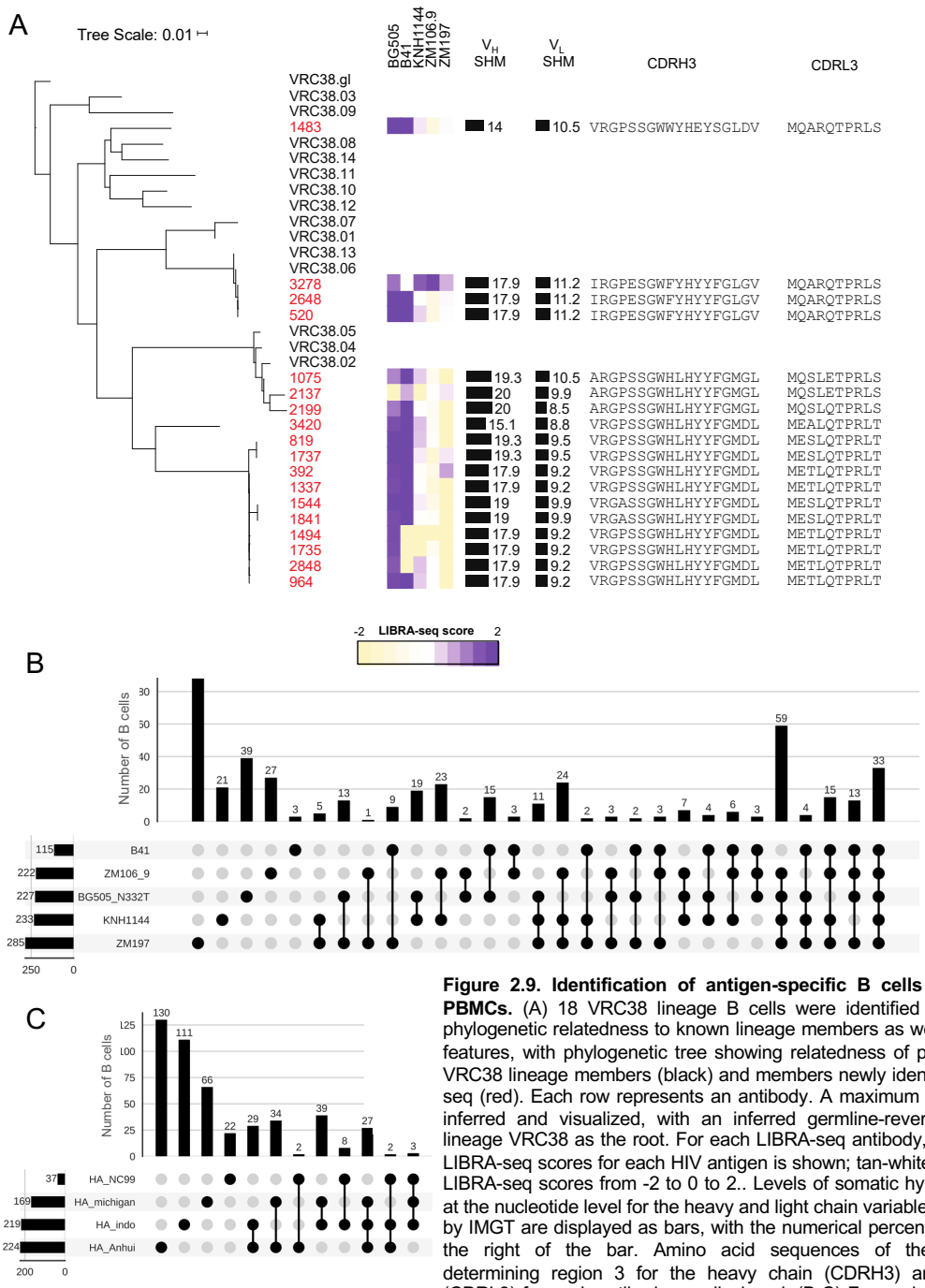


Figure 2.9. Identification of antigen-specific B cells from donor N90 PBMCs. (A) 18 VRC38 lineage B cells were identified and examined for phylogenetic relatedness to known lineage members as well as for sequence features, with phylogenetic tree showing relatedness of previously identified VRC38 lineage members (black) and members newly identified using LIBRA-seq (red). Each row represents an antibody. A maximum likelihood tree was inferred and visualized, with an inferred germline-reverted antibody from lineage VRC38 as the root. For each LIBRA-seq antibody, a heat map of the LIBRA-seq scores for each HIV antigen is shown; tan-white-purple represents LIBRA-seq scores from -2 to 0 to 2.. Levels of somatic hypermutation (SHM) at the nucleotide level for the heavy and light chain variable genes as reported by IMGT are displayed as bars, with the numerical percentage value listed to the right of the bar. Amino acid sequences of the complementarity determining region 3 for the heavy chain (CDRH3) and the light chain (CDRL3) for each antibody are displayed. (B-C) For each combination of (B) HIV SOSIPs or (C) influenza hemagglutinins, the number of B cells with high LIBRA-seq scores (>= 1) is displayed as a bar graph. The combinations of antigens are displayed by filled circles, showing which antigens are part of a given combination. Each combination is mutually exclusive. Total number of B cells with high LIBRA-seq scores for each antigen is shown as a horizontal bar on the left of each subpanel. Figure made by Ian Setliff and adapted from Setliff and Shiakolas et al., *Cell*, 2019.

correlation between LIBRA-seq scores and ELISA AUC (Spearman correlation 0.77, p=0.015) (**Figure 2.11A-B**).

We also observed cells that had high LIBRA-seq scores for each of the different HIV-1 antigens, including 124 cells that had high scores for four or more SOSIPs (**Figure 2.9B**). We then down-selected SOSIP-high B cells based on having high LIBRA-seq scores to at least 3 SOSIP variants. In particular, we identified two members from the same antibody lineage that had high LIBRA-seq scores for BG505, KNH1144, ZM106.9 and ZM197. This lineage utilized the germline genes *IGHV1-46* and *IGK3-20* and was highly mutated in both the heavy- and light-chain V gene. We recombinantly expressed one of the lineage members, 3602-870, which was 28.5% mutated in its heavy chain V gene and 17.0% mutated in its light chain V gene and had a 19 amino acid CDRH3 and 9 amino acid CDRL3 (**Figure 2.11A**). 3602-870 bound all SOSIP probes by ELISA (Spearman correlation of 0.97, $p<0.001$ between LIBRA-seq scores and ELISA AUC) and neutralized 79% of tested Tier 2 viruses (11/14), including four viruses that were not neutralized by VRC38.01 (TRO.11, CH119.10, 25710.2.43, and CE1176.A3)[127] (**Figure 2.11A-C**). Of note, 3602-870 neutralized BG505 and ZM197, both of which were used as probes in the antigen screening library (**Figure 2.11C**). 3602-870 bound BG505 DS-SOSIP by SPR and competed for BG505 DS-SOSIP binding with VRC01 Fab (**Figure 2.11D-E**).

In summary, LIBRA-seq enabled the high-throughput, highly multiplexed screening of single B cells from an HIV-infected subject against a large antigen panel. This resulted in the identification of hundreds of antigen-specific monoclonal antibody

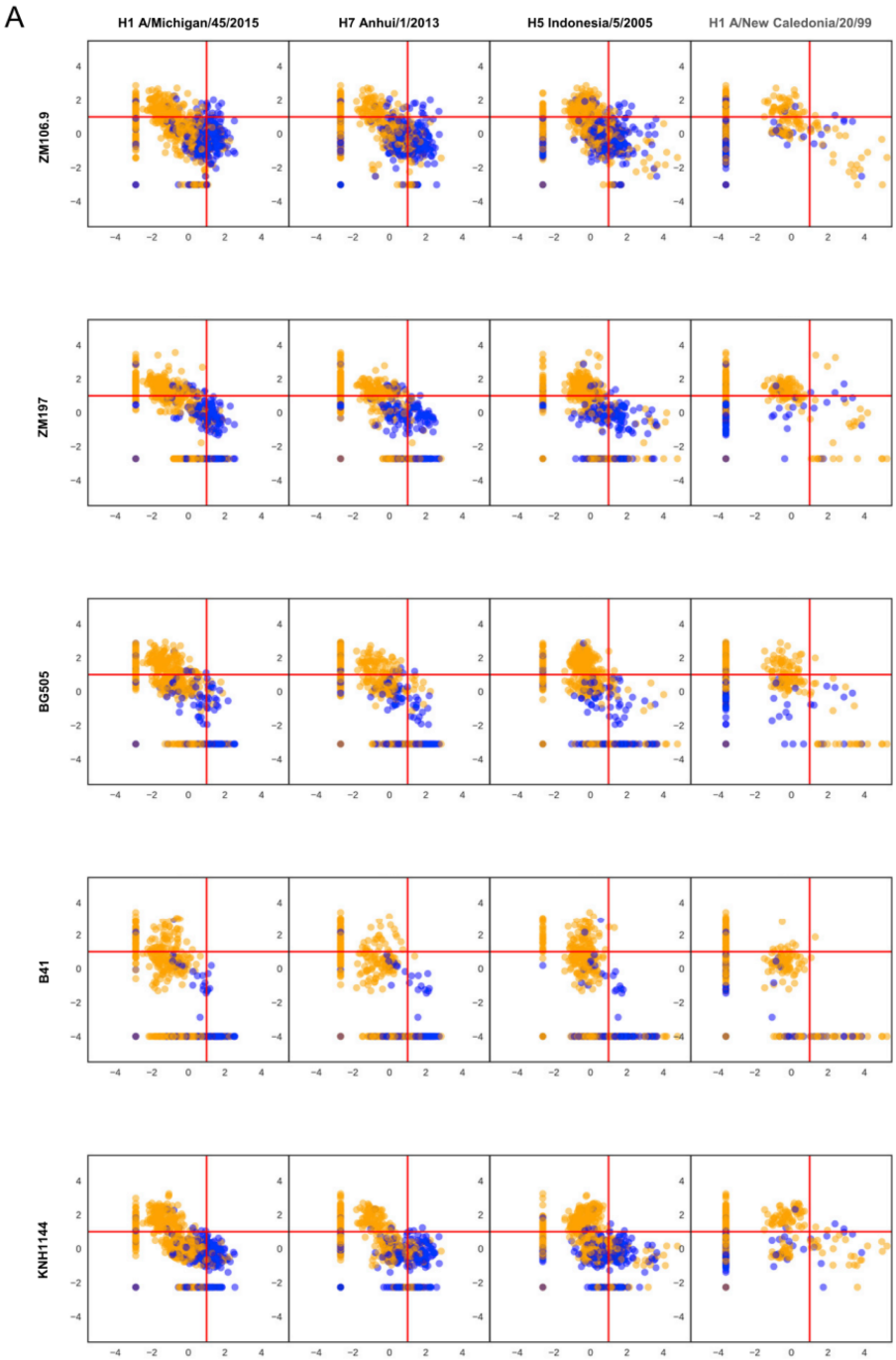


Figure 2.10. Analysis of antigen reactivity for B cells from donor N90. Each graph shows the LIBRA-seq score for an HIV antigen (y-axes) vs. an influenza antigen (x-axes) in the screening library. The 901 cells that had a LIBRA-seq score above one for at least one antigen are displayed as individual dots. IgG cells (591 of 901) are colored orange and cells of all other isotypes are colored blue. Red lines on each axis indicate a LIBRA-seq score of one. Only 9 of the 591 IgG cells displayed high LIBRA-seq scores for at least one HIV-1 antigen and one influenza antigen, confirming the ability of the technology to successfully discriminate between diverse antigen specificities. Figure made by Ian Setliff and adapted from Setliff and Shiakolas et al., *Cell*, 2019.

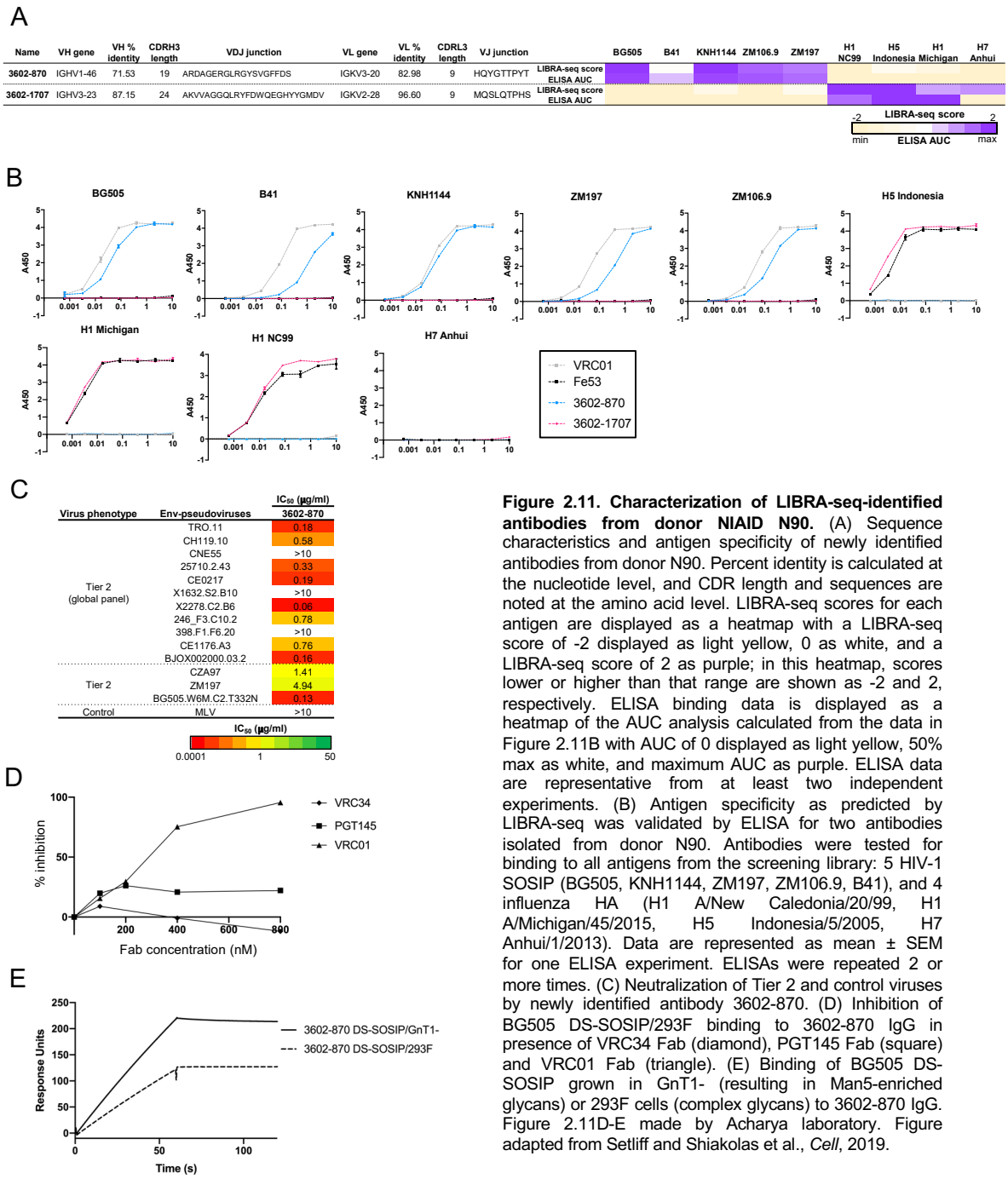


Figure 2.11. Characterization of LIBRA-seq-identified antibodies from donor NIAID N90. (A) Sequence characteristics and antigen specificity of newly identified antibodies from donor N90. Percent identity is calculated at the nucleotide level, and CDR length and sequences are noted at the amino acid level. LIBRA-seq scores for each antigen are displayed as a heatmap with a LIBRA-seq score of -2 displayed as light yellow, 0 as white, and a LIBRA-seq score of 2 as purple; in this heatmap, scores lower or higher than that range are shown as -2 and 2, respectively. ELISA binding data is displayed as a heatmap of the AUC analysis calculated from the data in Figure 2.11B with AUC of 0 displayed as light yellow, 50% max as white, and maximum AUC as purple. ELISA data are representative from at least two independent experiments. (B) Antigen specificity as predicted by LIBRA-seq was validated by ELISA for two antibodies isolated from donor N90. Antibodies were tested for binding to all antigens from the screening library: 5 HIV-1 SOSIP (BG505, KNH1144, ZM197, ZM106.9, B41), and 4 influenza HA (H1 A/New Caledonia/20/99, H1 A/Michigan/45/2015, H5 Indonesia/5/2005, H7 Anhui/1/2013). Data are represented as mean ± SEM for one ELISA experiment. ELISAs were repeated 2 or more times. (C) Neutralization of Tier 2 and control viruses by newly identified antibody 3602-870. (D) Inhibition of BG505 DS-SOSIP/293F binding to 3602-870 IgG in presence of VRC34 Fab (diamond), PGT145 Fab (square) and VRC01 Fab (triangle). (E) Binding of BG505 DS-SOSIP grown in GnT1- (resulting in Man5-enriched glycans) or 293F cells (complex glycans) to 3602-870 IgG. Figure 2.11D-E made by Acharya laboratory. Figure adapted from Setliff and Shiakolas et al., *Cell*, 2019.

leads from donor N90, with high-resolution antigen specificity mapping helping to facilitate rapid lead prioritization to identify a novel bNAb lineage.

2.3 Discussion

Here, we developed a novel method to interrogate antibody-antigen interactions via a sequencing-based readout. After validating the approach on cell lines with known BCRs, we applied LIBRA-seq to prospective antibody discovery. We identified members of two known HIV-specific bNAb lineages from previously characterized human infection samples and a novel bNAb lineage. Additionally, we identified many other candidate broadly-reactive HIV-specific antibodies, and validated specificity for a subset of them. Within both HIV-1 infection samples, we also isolated influenza-specific antibodies using hemagglutinin screening probes, highlighting the utility of LIBRA-seq for simultaneously screening B cell repertoires against multiple, diverse antigen targets. In principle, the NGS-based coupling of antibody sequence and specificity enables screening of potentially millions of single B cells for reactivity to a larger repertoire of epitopes than purely fluorescence-based methods, since sequence space is not hindered by spectral overlap. Using LIBRA-seq may therefore help to maximize lead discovery per experiment, an important consideration when preserving limited sample.

Beyond LIBRA-seq's utility in antibody discovery, the high-throughput coupling of antibody sequence and specificity can enable high-resolution immune profiling. For example, in donor N90 we observed an increase in *IGHV* gene somatic hypermutation between B cells that had a high LIBRA-seq score for a single HIV-1 antigen versus B cells that had high LIBRA-seq scores for multiple HIV-1 antigens (**Figure 2.12A**). We also observed the use of specific germline genes to be more frequent in B cells that exhibited broad, as opposed to strain-specific, HIV-1 antigen reactivity (**Figure 2.12B-**

C). The elucidation of such relationships, enabled by the LIBRA-seq technology, may guide germline-targeting vaccine design efforts[133-136] and can provide insights into the requirements for the acquisition of HIV-1 antigen cross-reactivity.

Future extensions of LIBRA-seq and other applications of molecular genomics technologies to antibody discovery and immune profiling should translate into a rapid accumulation of new data, leading to novel insights into basic and applied immunology.

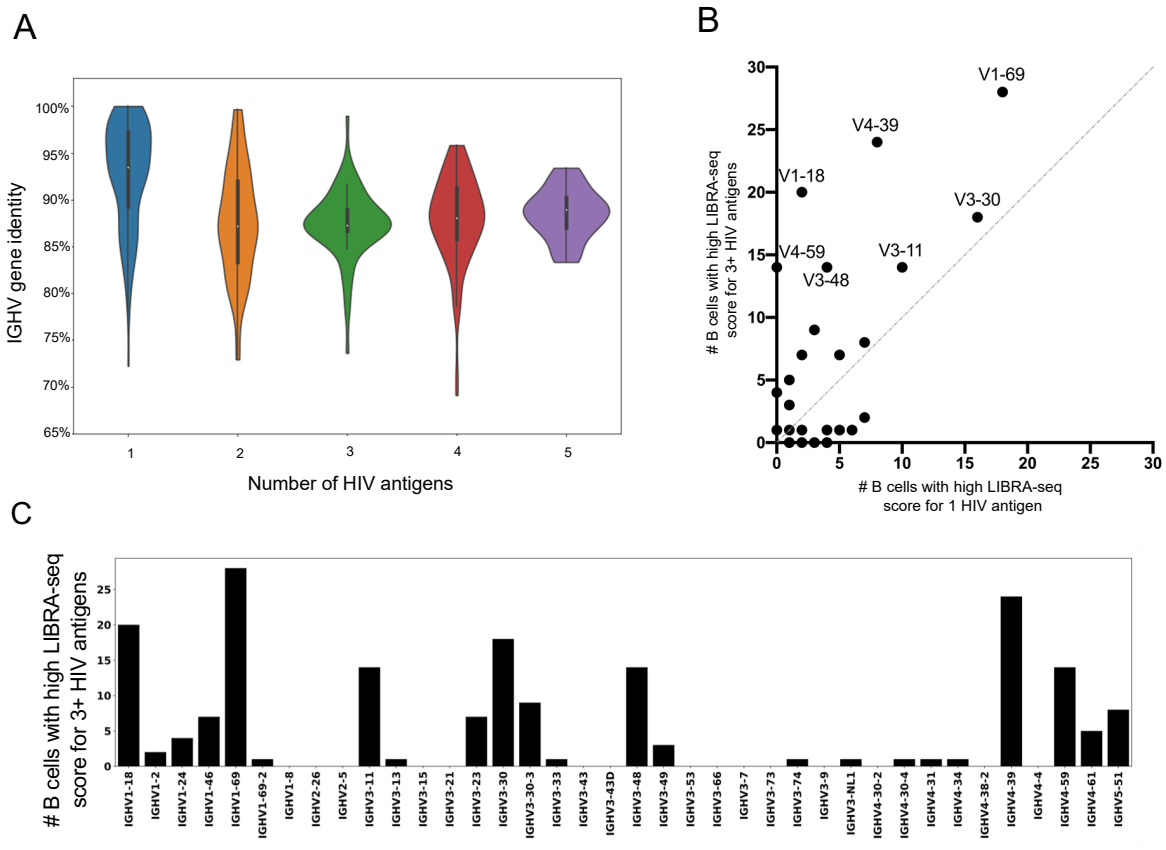


Figure 2.12. Sequence properties of the antigen-specific B cell repertoire. (A) *IGHV* gene identity (y-axis) is plotted for cells with high (>=1) LIBRA-seq scores for any combination of 1 through 5 HIV-1 SOSIP antigens (x-axis). Each distribution is displayed as a kernel density estimation, where wider sections of a given distribution represent a higher probability that B cells possess a given germline identity percentage. The median of each distribution is displayed as a white dot, the interquartile range is displayed as a thick bar, and a thin line extends to 1.5x the interquartile range. The violin ranges were limited to the observed data. Included are cells with IgG or IgA constant heavy genes as determined by Cell Ranger. (B) Each dot represents an *IGHV* germline gene, plotted based on the number of B cells reactive to only 1 HIV-1 SOSIP antigen (x axis) and the number of B cells reactive to 3 or more HIV-1 SOSIP antigens (y axis) that are assigned to that respective *IGHV* germline gene. Only B cells with high (>=1) LIBRA-seq scores for any HIV-1 antigen and with IgG or IgA constant heavy genes as determined by Cell Ranger are shown. (C) V gene usage of broadly HIV-reactive B cells. For each *IGHV* gene, the number of B cells with IgG or IgA constant heavy gene and high (>=1) LIBRA-seq scores for 3 or more HIV-1 SOSIP variants is displayed as a bar. The x-axis shows only *IGHV* genes used by at least 1 B cell with a high LIBRA-seq score for at least 1 HIV-1 antigen and an IgG or IgA C_H gene. Figure made by Ian Setliff and adapted from Setliff and Shiakolas et al., *Cell*, 2019.

Acknowledgements

We thank Angela Jones, Latha Raju and Jamie Robertson of Vanderbilt Technologies for Advanced Genomics for their expertise in next-generation sequencing and library preparation; Rebecca Gillespie for aid in producing the HA probe; Hannah Polikowsky, James Crowe, Jr., and Spyros Kalams for helpful discussions on experimental design; David Flaherty of the Vanderbilt Flow Cytometry Shared Resource for help with flow panel optimization; Carol Crowther of NICD for project management; Peter Kwong and John Mascola for providing the plasmid for BG505; and members of the Georgiev laboratory for providing comments on the manuscript. The Vanderbilt VANTAGE Core provided technical assistance for this work. VANTAGE is supported in part by CTSA Grant (5UL1 RR024975-03), the Vanderbilt Ingram Cancer Center (P30 CA68485), the Vanderbilt Vision Center (P30 EY08126), and NIH/NCRR (G20 RR030956). This work was conducted in part using the resources of the Advanced Computing Center for Research and Education at Vanderbilt University, Nashville, TN. Flow Cytometry experiments were performed in the VUMC Flow Cytometry Shared Resource. The VUMC Flow Cytometry Shared Resource is supported by the Vanderbilt Ingram Cancer Center (P30 CA68485) and the Vanderbilt Digestive Disease Research Center (DK058404). For work described in this manuscript, I.S.G., I.S., A.R.S., N.R. and A.R.G. were supported by NIH grant R01 AI131722, A.R.S. was supported by NIH grant T32 GM008320-30, K.A.P. was supported by NIH grant T32 AI112541, A.A.M. and K.J.K. were supported by the Vanderbilt Trans-Institutional Program (TIPs) “Integrating Structural with Big Data for Next-generation Vaccines”; M.K., B.S.G. and M.C. were supported by the NIH Intramural Research Program; L.M., C.M., and R.Z. were

supported by the South African Medical Research Council and H3A-U01 (NIH grant 1U01AI136677), and P.A. and K.J. were supported by NIH grant R01 AI145687.

Author Contributions

Conception of LIBRA-seq technology, I.S.; Conceptualization, I.S., A.R.S., I.S.G.; Methodology, I.S., A.R.S., I.S.G.; Investigation, I.S., A.R.S., K.A.P., A.A.M., R.Z., S.R., C.M., N.R., A.R.G., K.J., K.J.K., J.Q., D.L., L.R., W.J.M., I.S.G.; Software, I.S., N.R.; Validation, I.S., A.R.S, N.R.; Writing – Original Draft, I.S., A.R.S..; Writing – Review & Editing, All authors; Funding Acquisition, I.S.G., D.L., L.M., P.A., M.C., B.S.G., I.S.; Resources, M.K., P.A., B.S.G., M.C., D.L., L.R., L.M., I.S.G.; Supervision, I.S.G., L.M., P.A.

Declaration of Interests

Vanderbilt University has filed for patent protection of some of the technology and results presented in this study with I.S., A.R.S., and I.S.G. listed as inventors. W.J.M. is an employee and shareholder of 10X Genomics.

2.4 Materials and Methods

Human Subjects

Donor NIAID45: Peripheral blood mononuclear cells were collected from donor NIAID45 on July 12, 2007. Donor NIAID45 was enrolled in investigational review board approved clinical protocols at the National Institute of Allergy and Infectious Diseases

and had been living with HIV-1 without antiretroviral treatment for approximately 17 years at the time of sample collection. *Donor N90:* Peripheral blood mononuclear cells were collected from donor N90 on May 29, 2008. Donor N90 was enrolled in investigational review board approved clinical protocols at the National Institute of Allergy and Infectious Diseases and had been living with HIV-1 without antiretroviral treatment through the timepoint of sample collection since diagnosis in 1985[128].

Cell Lines

Ramos B-cell lines were engineered from a clone of Ramos Burkitt's lymphoma that do not display endogenous antibody, and they ectopically express specific surface IgM B cell receptor sequences. The B cell lines used expressed B cell receptor sequences for HIV-specific antibody VRC01 and influenza-specific antibody Fe53. The cells are cultured at 37°C with 5% CO_2 saturation in complete RPMI, made up of RPMI supplemented with 15% fetal bovine serum, 1% L-Glutamine, and 1% Penicillin/Streptomycin. Although endogenous heavy chains are scrambled, endogenous light chain transcripts remain and are detectable by sequencing. We thus identified and classified single Ramos Burkitt's B cells as either VRC01 or Fe53 based on their heavy chain sequences. These Ramos B-cell lines were validated for binding to our antigen probes by FACS (**Figure 2.2B**).

Antigen Expression and Purification

For the different LIBRA-seq experiments, a total of six HIV-1 gp140 SOSIP variants from strains BG505 (clade A), CZA97 (clade C), B41 (clade B), ZM197 (clade

C), ZM106.9 (clade C), KNH1144 (clade A) and four influenza hemagglutinin variants from strains A/New Caledonia/20/99 (H1N1) (GenBank ACF41878), A/Michigan/45/2015 (H1N1) (GenBank AMA11475), A/Indonesia/5/2005 (H5N1) (GenBank ABP51969), and A/Anhui/1/2013 (H7N9) (GISAID EPI439507) were expressed as recombinant soluble antigens.

The single-chain variants[121] of BG505, CZA97, B41, ZM197, ZM106.9, and KNH1144 each containing an AviTag, were expressed in FreeStyle 293F mammalian cells (ThermoFisher) using polyethylenimine (PEI) transfection reagent and cultured for 5-7 days. FreeStyle 293F were maintained in FreeStyle 293F medium or FreeStyle F17 expression medium supplemented with 1% of 10% Pluronic F-68 and 20% of 200 mM L-Glutamine. These cells were cultured at 37°C with 8% CO_2 saturation and shaking. After transfection and 5-7 days of culture, cultures were centrifuged at 6000 rpm for 20 minutes. Supernatant was filtered with Nalgene Rapid Flow Disposable Filter Units with PES membrane (0.45 μm), and then run slowly over an affinity column of agarose bound *Galanthus nivalis* lectin (Vector Laboratories cat no. AL-1243-5) at 4°C. The column was washed with PBS, and proteins were eluted with 30 mL of 1 M methyl-α-D-mannopyranoside. The protein elution was buffer exchanged 3X into PBS and concentrated using 30kDa Amicon Ultra centrifugal filter units. Concentrated protein was run on a Superose 6 Increase 10/300 GL or Superdex 200 Increase 10/300 GL sizing column on the AKTA FPLC system, and fractions were collected on an F9-R fraction collector. Fractions corresponding to correctly folded antigen were analyzed by SDS-PAGE, and antigenicity by ELISA was characterized with known monoclonal antibodies specific for that antigen.

Recombinant HA proteins all contained the HA ectodomain with a point mutation at the sialic acid-binding site (Y98F), T4 fibritin foldon trimerization domain, AviTag, and hexahistidine-tag, and were expressed in Expi 293F mammalian cells using Expifectamine 293 transfection reagent (Thermo Fisher Scientific) cultured for 4-5 days. Culture supernatant was harvested and cleared as above, and then adjusted pH and NaCl concentration by adding 1M Tris-HCl (pH 7.5) and 5M NaCl to 50 mM and 500 mM, respectively. Ni Sepharose excel resin (GE Healthcare) was added to the supernatant to capture hexahistidine tag. Resin was separated on a column by gravity and captured HA protein was eluted by a Tris-NaCl (pH 7.5) buffer containing 300 mM imidazole. The eluate was further purified by a size exclusion chromatography with a HiLoad 16/60 Superdex 200 column (GE Healthcare). Fractions containing HA were concentrated, analyzed by SDS-PAGE and tested for antigenicity by ELISA with known antibodies. Proteins were frozen in LN_2 and stored at –80Cº until use.

All HIV-1 gp140 SOSIP variant antigens and all influenza hemagglutinin variant antigens included an AviTag modification at the C-terminus of their sequence, and after purification, each AviTag labeled antigen was biotinylated using the BirA-500: BirA biotin-protein ligase standard reaction kit (Avidity LLC, cat no. BirA500).

Oligonucleotide Barcodes

We used oligos that possess a 13-15 bp antigen barcode, a sequence capable of annealing to the template switch oligo that is part of the 10X bead-delivered oligos, and contain truncated TruSeq small RNA read 1 sequences in the following structure: 5’-CCTTGGCACCCGAGAATTCCANNNNNNNNNNNNNCCCATATAAGA*A*A-3’, where

Ns represent the antigen barcode. For the cell line and NIAID45 experiments, we used the following antigen barcodes: CATGATTGGCTCA (BG505), TGTCCGGCAATAA (CZA97), GATCGTAATACCA (H1 A/New Caledonia/20/99). For the N90 experiment, we used longer antigen barcodes (15 bp), as follows: TCCTTTCCTGATAGG (ZM106.9), TAACTCAGGGCCTAT (KNH1144), GCTCCTTTACACGTA (ZM197), GCAGCGTATAAGTCA (B41), ATCGTCGAGAGCTAG (BG505), CAGGTCCCTTATTTC (A/Indonesia/5/2005), ACAATTTGTCTGCGA (A/Anhui/1/2013), TGACCTTCCTCTCCT (A/Michigan/45/2015), AATCACGGTCCTTGT (A/New Caledonia/20/99). Oligos were ordered from Sigma-Aldrich and IDT with a 5' amino modification and HPLC purified.

Conjugation of Oligonucleotide Barcodes to Antigens

For each antigen, a unique DNA barcode was directly conjugated to the antigen itself. In particular, 5'amino-oligonucleotides were conjugated directly to each antigen using the Solulink Protein-Oligonucleotide Conjugation Kit (TriLink cat no. S-9011) according to manufacturer's instructions. Briefly, the oligo and protein were desalted, and then the amino-oligo was modified with the 4FB crosslinker, and the biotinylated antigen protein was modified with S-HyNic. Then, the 4FB-oligo and the HyNic-antigen were mixed together. This causes a stable bond to form between the protein and the oligonucleotide. The concentration of the antigen-oligo conjugates was determined by a BCA assay, and the HyNic molar substitution ratio of the antigen-oligo conjugates was analyzed using the NanoDrop according to the Solulink protocol guidelines. AKTA FPLC

was used to remove excess oligonucleotide from the protein-oligo conjugates, which were also checked using SDS-PAGE with a silver stain.

Fluorescent Labeling of Antigens

After attaching DNA barcodes directly to a biotinylated antigen, the barcoded antigens were mixed with streptavidin labeled with fluorophore phycoerythrin (PE). The streptavidin-PE was mixed with biotinylated antigen at a 5X molar excess of antigen to streptavidin. 1/5 of the streptavidin-oligo conjugate was added to the antigen every 20 minutes with constant rotation at 4°C.

Enrichment of Antigen-Specific B Cells

For a given sample, cells were stained and mixed with fluorescently labeled DNA-barcoded antigens and other antibodies, and then sorted using fluorescence activated cell sorting (FACS). First, cells were counted and viability was assessed using Trypan Blue. Then, cells were washed with DPBS supplemented with 1 % Bovine serum albumin (BSA) through centrifugation at 300 g for 7 minutes. Cells were resuspended in PBS-BSA and stained with a variety of cell markers. For donor NIAID45 PBMCs, these markers included CD3-APCCy7, IgG-FITC, CD19-BV711, CD14-V500, and LiveDead-V500. Additionally, fluorescently labeled antigen-oligo conjugates (described above) were added to the stain. For donor N90 PBMCs, these markers included Ghost Red 780, CD14-APCCy7, CD3-FITC, CD19-BV711, and IgG-PECy5. Additionally, fluorescently labeled antigen-oligo conjugates were added to the stain. After staining in the dark for 30 minutes at room temperature, cells were washed 3 times

with PBS-BSA at 300 g for 7 minutes. Then, cells were resuspended in DPBS and sorted on the cell sorter. Antigen positive cells were bulk sorted and then they were delivered to the Vanderbilt VANTAGE sequencing core at an appropriate target concentration for 10X Genomics library preparation and subsequent sequencing. FACS data were analyzed using Cytobank[137].

10X Genomics Single Cell Processing and Next Generation Sequencing

Single-cell suspensions were loaded onto the Chromium Controller microfluidics device (10X Genomics) and processed using the B-cell Single Cell V(D)J solution according to manufacturer's suggestions for a target capture of 10,000 B cells per 1/8 10X cassette for B cell lines, 9,000 cells for B cells from donor NIAID45, and 4,000 for donor N90, with minor modifications in order to intercept, amplify and purify the antigen barcode libraries. The library preparation followed the CITE-seq protocol (available at https://cite-seq.com), with the exception of an increase in the number of PCR cycles of the antigen barcodes. Briefly, following cDNA amplification using an additive primer (5'-CCTTGGCACCCGAGAATT*C*C-3') to increase the yield of antigen barcode libraries [115], SPRI separation was used to size separate antigen barcode libraries from cellular mRNA libraries, PCR amplified for 10-12 cycles, and purified using 1.6X purification. Sample preparation for the cellular mRNA library continued according to 10X Genomics-suggested protocols, resulting in Illumina-ready libraries. Following library construction, we sequenced both BCR and antigen barcode libraries on a NovaSeq 6000 at the VANTAGE sequencing core, dedicating ~2.5% of a flow cell to each experiment, with a target 10% of this fraction dedicated to antigen barcode libraries.

This resulted in ~334.5 million reads for the cell line V(D)J libraries (~96,500 reads/cell), ~376.3 million reads for donor NIAID45 V(D)J libraries (~79,300 reads/cell), and ~272.4 million reads for the N90 V(D)J libraries (~151,400 reads/cell). The N90 antigen barcode libraries were also sequenced a second time.

Processing of Antigen Barcode Reads and BCR Sequence Contigs

We developed a pipeline that takes paired-end FASTQ files of oligo libraries as input, processes and annotates reads for cell barcode, UMI, and antigen barcode, and generates a cell barcode - antigen barcode UMI count matrix. BCR contigs were processed using Cell Ranger (10X Genomics) using GRCh38 as reference. For the antigen barcode libraries, initial quality and length filtering is carried out by fastp[138] using default parameters for filtering. This resulted in only high-quality reads being retained in the antigen barcode library (**Figure 2.13A**). In a histogram of insert lengths, this resulted in a sharp peak of the expected insert size of 52-54 bp (**Figure 2.13B-D**). Fastx_collapser was then used to group identical sequences and convert the output to deduplicated fasta files. Then, having removed low-quality reads, we proceeded by processing just the R2 sequences, as the entire insert is present in both R1 and R2. Each unique R2 sequence was processed one by one using the following steps: (1) The reverse complement of the R2 sequence was determined. (2) The sequence was screened for possessing an exact match to any of the valid 10X cell barcodes present in the filtered_contig.fasta file output by Cell Ranger during processing of BCR V(D)J FASTQ files. Sequences without a BCR-associated cell barcode were discarded. (3) The 10 bases immediate 3' to the cell barcode were annotated as the read's UMI. (4)

The remainder of the sequence 3' to the UMI is screened for a 13 or 15 bp sequence with a hamming distance of 0, 1, or 2 to any of the antigen barcodes used in the screening library. Following this processing, only sequences with lengths of 51 to 58 were retained, thus allowing for a deletion, an insertion outside the cell barcode, or bases flanking the cell barcode. This general process required that sequences possess all elements needed for analysis (cell barcode, UMI, and antigen barcode), but is

A

Ramos B Cell Lines

Summary	
Sequencing	paired end (150 cycles + 150 cycles)
Insert size peak	52 bp
Before filtering	
Total reads	46.683278 M
Total bases	7.002492 G
Q20 bases	5.890977 G (84.126864%)
Q30 bases	5.343662 G (76.310871%)
After filtering	
Total reads	46.317712 M
Total bases	2.473195 G
Q20 bases	2.438363 G (98.591592%)
Q30 bases	2.358131 G (95.347568%)
Filtering results	
Reads passed filters	46.317712 M
Reads with low quality	253.286000 K (0.542563%)
Reads with too many N	94 (0.000201%)
Reads too short	112.186000 K (0.240313%)

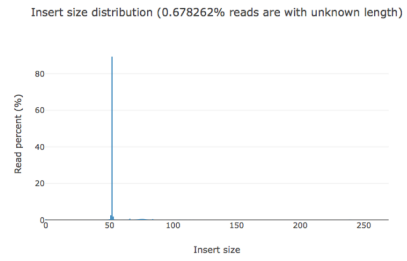
Donor NIH45

Summary	
Sequencing	paired end (150 cycles + 150 cycles)
Insert size peak	52 bp
Before filtering	
Total reads	39.610554 M
Total bases	5.940879 G
Q20 bases	4.959718 G (83.484582%)
Q30 bases	4.463982 G (75.140098%)
After filtering	
Total reads	36.922686 M
Total bases	2.061362 G
Q20 bases	1.992712 G (96.669638%)
Q30 bases	1.910089 G (92.661463%)
Filtering results	
Reads passed filters	36.922686 M
Reads with low quality	1.506144 K (3.802381%)
Reads with too many N	158 (0.000399%)
Reads too short	1.181566 M (2.982958%)

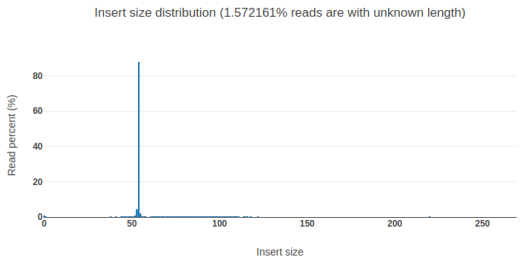
Donor N90

Summary	
Sequencing	paired end (150 cycles + 150 cycles)
Insert size peak	54 bp
Before filtering	
Total reads	82.957516 M
Total bases	8.65364 G
Q20 bases	7.647815 G (88.377127%)
Q30 bases	6.985525 G (80.723789%)
After filtering	
Total reads	80.923090 M
Total bases	6.994344 G
Q20 bases	6.314613 G (90.281709%)
Q30 bases	5.822037 G (83.239223%)
Filtering results	
Reads passed filters	80.923090 M
Reads with low quality	1.494196 M (1.801158%)
Reads with too many N	4.434000 K (0.005333%)
Reads too short	535.806000 K (0.645880%)

B



D



C

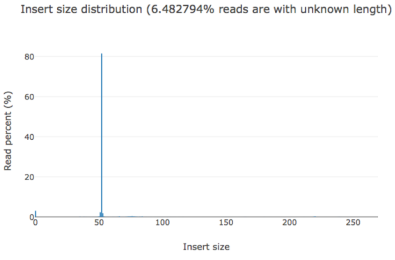


Figure 2.13. Sequencing preprocessing and quality statistics. (A) Quality filtering of the antigen barcode FASTQ files. Fastp (Chen et al., 2018) was used to trim adapters and remove low-quality reads using default parameters. Shown are read and base statistics generated from the output html report from each of the Ramos B cell experiment (left), primary B cell experiment from donor NIAID45 (middle), and primary B cell experiment from donor N90 (right). (B) Shown is a distribution of insert sizes of the antigen barcode reads from the Ramos B-cell line experiment, as output from the fastp html report. (C) Shown is a distribution of insert sizes of the antigen barcode reads from the donor NIAID45 experiment, as output from the fastp html report. (D) Shown is a distribution of insert sizes of the antigen barcode reads from the donor NIH90 experiment, as output from the fastp html report. Figure made by Ian Setliff and adapted from Setliff and Shiakolas et al., *Cell*, 2019.

permissive to insertions or deletions in the TSO region between the UMI and antigen barcode. After processing each sequence one-by-one, we screened for cell barcode - UMI - antigen barcode collisions. Any cell barcode - UMI combination that had multiple antigen barcodes associated with it was removed. We then constructed a cell barcode - antigen barcode UMI count matrix, which served as the basis of subsequent analysis. Additionally, we aligned the BCR contigs (filtered_contigs.fasta file output by Cell Ranger, 10X Genomics) to IMGT reference genes using HighV-Quest[139]. The output of HighV-Quest was parsed using ChangeO[140], and merged with the UMI count matrix.

Determination of LIBRA-seq Score

Starting with the UMI count matrix, we set all counts of 1, 2, or 3 UMIs to 0, with the idea that these low counts could likely be attributed to noise. After this, the UMI count matrix was subset to contain only cells with a count of at least 4 UMIs for at least 1 antigen. We also removed cells that had only non-functional heavy chain sequences as well as cells with multiple functional heavy chain sequences using different *IGHV* genes, reasoning that these may be multiplets. We then calculated the centered-log ratios (CLR) of each antigen UMI count for each cell[114, 115]. Because UMI counts were on different scales for each antigen, possibly due to differential oligo loading during oligo-antigen conjugation, we rescaled the CLR UMI counts using the StandardScaler method in *scikit learn*[141]. Lastly, we performed a correction procedure to the scaled CLRs from UMI counts of 0, setting them to the minimum for each antigen for donor NIAID45 and N90 experiments, and to -1 for the Ramos B-cell line experiment. These CLR-transformed, scaled, corrected values served as the final LIBRA-seq scores.

LIBRA-seq scores were visualized using Cytobank[137] and Matplotlib[142]. Donor NIAID45 and N90 data were subsetted to include only cells with a functional light chain.

Phylogenetic Trees

Phylogenetic trees of antibody heavy chain sequences were constructed in order to assess the relatedness of antibodies within a given lineage. For the VRC01 lineage, the 29 sequences identified by LIBRA-seq and 53 sequences identified from the literature were aligned using clustal within Geneious. We then used the PhyML maximum likelihood[143] plugin in Geneious (available at https://www.geneious.com/plugins/phyml-plugin/) to infer a phylogenetic tree. The resulting tree was then rooted to the inferred unmutated common ancestor[105] (accession MK032222). Names for sequences and their accession include H01+07.F.1 (KP840594); H03+06.C.1 (KP840597); H03+06.E.1 (KP841560); H4.E.6 (KP841696); H4.E.5 (KP841700); H4.E.4 (KP841639); H4.E.3 (KP841608); H4.E.2 (KP841609); H4.E.1 (KP841701); H5.C.1 (KP840607); H5.F.1 (KP840608); H08.F.1 (KP840603); H08.H.1 (KP840835); VRC03b (KP840671); VRC03f (KP840674); VRC03g (KP840675); DH651.1 (MK032223); DH651.3 (MK032225); DH651.9 (MK032231); DH651.8 (MK032230); VRC06c (KP840678); VRC06d (KP840679); VRC06e (KP840680); VRC06f (KP840681); VRC06g (KP840682); VRC06h (KP840683); DH651.2 (MK032224); DH651.4 (MK032226); DH651.5 (MK032227); DH651.6 (MK032228); DH651.7 (MK032229); VRC06 (JX466923.1); VRC03 (GU980706.1); NIH45-46 (HE584543); VRC01 (GU980702); VRC01c (KP840658); VRC01d (KP840659); VRC01e (KP840660); VRC01f (KP840661); VRC01h (KP840663); VRC01i

(KP840664); VRC01j (KP840665); VRC02 (GU980704); VRC07b (KP840666); VRC07c (KP840667); VRC07d (KP840668); VRC07e (KP840669); VRC07f (KP840670); VRC08c (KP840685); VRC08d (KP840686); VRC08e (KP840687); H03+06.A.0 (KP841501); VRC01UCA (MK032222). A similar process was used to build a phylogenetic tree for the VRC38 lineage, with one exception. Rather than using an inferred germline precursor, we germline-reverted framework 1, CDR1, framework 2, CDR2, framework 3, and framework 4 and used the junction nucleotide sequence of the lineage member with the least *IGHV* somatic mutation (VRC38.03). Trees were annotated and visualized in iTol[144]. While trees were constructed based on heavy chains, all VRC01 and VRC38 B cells had a correct light chain transcript, although sometimes additional light chain transcripts were also observed. One LIBRA-seq-identified VRC38 lineage member, 3602-1544, contained a single nucleotide deletion in the Cell Ranger-determined contig sequence in framework 2; this was manually corrected prior to inferring the phylogenetic tree.

Antibody Expression and Purification

For each antibody, variable genes were inserted into plasmids encoding the constant region for the heavy chain (pFUSEss-CHIg-hG1, Invivogen) and light chain (pFUSE2ss-CLIg-hl2, Invivogen and pFUSE2ss-CLIg-hk Invivogen) and synthesized from GenScript. Monoclonal antibodies (mAbs) were expressed in FreeStyle 293F or Expi293F mammalian cells (ThermoFisher) by co-transfecting heavy chain and light chain expressing plasmids using polyethylenimine (PEI) transfection reagent and cultured for 5-7 days. FreeStyle 293F (ThermoFisher) and Expi293F (ThermoFisher)

cells were maintained in FreeStyle 293F medium or FreeStyle F17 expression medium supplemented with 1% of 10% Pluronic F-68 and 20% of 200 mM L-Glutamine. These cells were cultured at 37°C with 8% CO_2 saturation and shaking. After transfection and 5-7 days of culture, cell cultures were centrifuged at 6000 rpm for 20 minutes. Supernatant was 0.45 μm filtered with Nalgene Rapid Flow Disposable Filter Units with PES membrane. Filtered supernatant was run over a column containing Protein A agarose resin that had been equilibrated with PBS. The column was washed with PBS, and then antibodies were eluted with 100 mM Glycine HCl at pH 2.7 directly into a 1:10 volume of 1 M Tris-HCl pH 8. Eluted antibodies were buffer exchanged into PBS 3 times using 10kDa Amicon Ultra centrifugal filter units.

Enzyme Linked Immunosorbent Assay (ELISA)

For hemagglutinin ELISAs, soluble hemagglutinin protein was plated at 2 μg/ml overnight at 4°C. The next day, plates were washed three times with PBS supplemented with 0.05% Tween20 (PBS-T) and coated with 5% milk powder in PBS-T. Plates were incubated for one hour at room temperature and then washed three times with PBS-T. Primary antibodies were diluted in 1% milk in PBS-T, starting at 10 μg/ml with a serial 1:5 dilution and then added to the plate. The plates were incubated at room temperature for one hour and then washed three times in PBS-T. The secondary antibody, goat anti-human IgG conjugated to peroxidase, was added at 1:20,000 dilution in 1% milk in PBS-T to the plates, which were incubated for one hour at room temperature. Plates were washed three times with PBS-T and then developed by adding TMB substrate to each well. The plates were incubated at room temperature for

ten minutes, and then 1 N sulfuric acid was added to stop the reaction. Plates were read at 450 nm.

For recombinant trimer capture for single-chain SOSIPs, 2 µg/ml of a mouse anti-AviTag antibody (GenScript) was coated overnight at 4°C in phosphate-buffered saline (PBS) (pH 7.5). The next day, plates were washed three times with PBS-T and blocked with 5% milk in PBS-T. After an hour incubation at room temperature and three washes with PBS-T, 2 µg/ml of recombinant trimer proteins diluted in 1% milk PBS-T were added to the plate and incubated for one hour at room temperature. Primary and secondary antibodies, along with substrate and sulfuric acid, were added as described above. Data are represented as mean ± SEM for one ELISA experiment. ELISAs were repeated 2 or more times. The area under the curve (AUC) was calculated using GraphPad Prism 8.0.0.

TZM-bl Neutralization Assays

Antibody neutralization was assessed using the TZM-bl assay as described[145]. This standardized assay measures antibody-mediated inhibition of infection of JC53BL-13 cells (also known as TZM-bl cells) by molecularly cloned Env-pseudoviruses. Viruses that are highly sensitive to neutralization (Tier 1) and/or those representing circulating strains that are moderately sensitive (Tier 2) were included, plus additional viruses, including a subset of the antigens used for LIBRA-seq. Murine leukemia virus (MLV) was included as an HIV-specificity control and VRC01 was used as a positive control. Results are presented as the concentration of monoclonal antibody (in µg/ml) required to inhibit 50% of virus infection (IC_{50}).

Surface Plasmon Resonance and Fab Competition

HIV-1 Env BG505 DS-SOSIP was produced either in GnT1- or 293F cells and purified as described previously[126]. The binding of antibodies 2723-2121 and 3602-870 to BG505 DS-SOSIP was assessed by surface plasmon resonance on Biacore T-200 (GE-Healthcare) at 25°C with HBS-EP+ (10 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA, and 0.05% surfactant P-20) as the running buffer. Antibodies VRC01 and PGT145 were tested as positive control, and antibody 17b was tested as negative control to confirm that the trimer was in the closed conformation. Antibodies 2723-2121 and 3602-870 were captured on a flow cell of CM5 chip immobilized with ~9000 RU of anti-human Fc antibody, and binding was measured by flowing over a 200 nM solution BG505-DS SOSIP in running buffer. Similar runs were performed with VRC01, PGT145 and 17b IgGs. To determine their epitopes, antibodies 2723-2121 IgG and 3602-870 were captured on a single flow cell of CM5 chip immobilized with anti-human Fc antibody. Next 200 nM BG505 DS-SOSIP, either alone or with different concentrations of antigen binding fragments (Fab) of VRC01 or PGT145 or VRC34 was flowed over the captured 2723-2121 or 3602-870 flow cell for 60s at a rate of 10 µl/min. The surface was regenerated between injections by flowing over 3M $MgCl_2$ solution for 10 s with flow rate of 100 µl/min. Blank sensorgrams were obtained by injection of same volume of HBS-EP+ buffer in place of trimer with Fabs solutions. Sensorgrams of the concentration series were corrected with corresponding blank curves.

ADCP, ADCD, Trogocytosis, ADCC Assays

Antibody-dependent cellular phagocytosis (ADCP) was performed using gp120 ConC coated neutravidin beads as previously described[146]. Phagocytosis score was determined as the percentage of cells that took up beads multiplied by the fluorescent intensity of the beads. Antibody-dependent complement deposition (ADCD) was performed as in [147] where CEM.NKR.CCR5 gp120 ConC coated target cells were opsonized with mAb and incubated with complement from a healthy donor. C3b deposition was then determined by flow cytometry with complement deposition score determined as the percentage of C3b positive cells multiplied by the fluorescence intensity. Antibody dependent cellular trogocytosis (ADCT) was measured as the percentage transfer of PKH26 dye of the surface of CEM.NKR.CCR5 target cells to CSFE stained monocytic cell line THP-1 cells in the presence of HIV-1 specific mAbs as described elsewhere[148] . Antibody-dependent cellular cytotoxicity (ADCC) was done using a GranToxiLux based assay[149] with gp120 ConC coated CEM.NKR.CCR5 target cells and PBMCs from a healthy donor. The percentage of granzyme B present in target cells was measured by flow cytometry.

Quantification and Statistical Analysis

ELISA error bars (standard error of the mean) were calculated using GraphPad Prism version 8.0.0. The Pearson's r value comparing BG505 and CZA97 LIBRA-seq scores for Ramos B-cell lines was calculated using Cytobank. Spearman correlations and associated p values were calculated using *SciPy* in Python.

Data Availability

Raw sequencing data used in this study are available on the Sequence Read Archive under BioProject accession number PRJNA578389. Identified antibody sequences related to VRC01 have been deposited to GenBank under accession codes MN580550 – MN580578 (heavy chain) and MN580579 – MN580607 (light chain). Identified antibody sequences related to VRC38 have been deposited to GenBank under accession codes MN580608 – MN580625 (heavy chain) and MN580626 – MN580643 (light chain). Other sequences from antibodies identified and recombinantly produced as part of this study have been deposited to GenBank under accession codes MN580644 – MN580654 (heavy chain) and MN580655 – MN580665 (light chain).

CHAPTER 3

Leveraging LIBRA-seq for Antibody Discovery

3.1 Cross-reactive Coronavirus Antibodies with Diverse Epitope Specificities and Fc Effector Functions

This section is adapted from published manuscript:

Shiakolas A*, Kramer K*, Wrapp D, Richardson S, Schäfer A, Wall S, Wang N, Janowska K, Pilewski K, Venkat R, Parks R, Manamela N, Raju N, Friedman Fechter E, Holt C, Suryadevara N, Chen R, Martinez D, Nargi R, Sutton R, Ledgerwood J, Graham B, Diamond M, Haynes B, Acharya P, Carnahan R, Crowe Jr J, Baric R, Morris L, McLellan J, Georgiev I. (2021). Cross-reactive coronavirus antibodies with diverse epitope specificities and Fc effector functions. *Cell Reports Medicine*, doi: 10.1016/j.xcrm.2021.100313. (*Equal Contribution)

Contributions:

Daniel Wrapp expressed and purified antigens. Simone Richardson performed Fc effector function experiments. Alexandra Schäfer performed *in vivo* antibody prophylaxis experiments. Nianshuang Wang provided antigens. Katarzyna Janowska performed surface plasmon resonance experiments. Kelsey Pilewski performed mannose competition experiments. Rohit Venkat created the LIBRA-seq pipeline using R. Rob

Parks performed autoreactivity assays. Nelia Manamela performed Fc effector function assays. Emilee Friedman expressed and purified antibodies. Clinton Holt performed computational epitope analysis of spike proteins. Naveenchandra Suryadevara performed ACE2 inhibition assays. Rita Chen performed RTCA neutralization assays. David Martinez perfomed live virus neutralization tests. Rachel Nargi and Rachel Sutton performed high throughput expression and purification of antibody candidates. Julie Ledgerwood and Barney Graham provided donor PBMCs. Michael Diamond oversaw neutralization experiments. Barton Haynes oversaw autoreactivity experiments. Priyamvada Achyra oversaw surface plasmon resonance experiments. Robert Carnahan and James Crowe Jr. oversaw ACE2 blocking assays and assisted with PBMC acquisition. Ralph Baric oversaw antibody prophylaxis experiments. Lynn Morris oversaw Fc effector function experiments. Jason McLellan oversaw antigen expression and purification. Ivelin Georgiev oversaw experimental design and project execution. Kevin Kramer performed cell staining and FACS, expressed and purified antigen proteins, purified antibodies, and performed ELISA binding experiments. I performed oligo-labeling of antigens for LIBRA-seq, including quality control, cell staining and flow cytometry, bioinformatic processing and sequencing analysis, antibody identification, antibody expression and purification, ELISA binding experiments, competition experiments, and cell-surface display experiments. Kevin Kramer, Ivelin Georgiev, and I wrote the manuscript.

3.1.1 Introduction

The continual emergence of novel coronavirus (CoV) strains, like SARS-CoV-2, highlights the critical need for broadly reactive therapeutics and vaccines against this family of viruses. From a recovered SARS-CoV donor sample, we identify and characterize a panel of six monoclonal antibodies that cross-react with CoV spike (S) proteins from the highly pathogenic SARS-CoV and SARS-CoV-2, and demonstrate a spectrum of reactivity against other CoV. Epitope mapping reveals that these antibodies recognize multiple epitopes on SARS-CoV-2 S, including the receptor binding domain, N-terminal domain, and S2 subunit. Functional characterization demonstrates that the antibodies mediate phagocytosis - and in some cases trogocytosis - but not neutralization *in vitro.* When tested *in vivo* in murine models, two of the antibodies demonstrate a reduction in hemorrhagic pathology in the lungs. The identification of cross-reactive epitopes recognized by functional antibodies expands the repertoire of targets for pan-coronavirus vaccine design strategies.

The emergence of a novel coronavirus (CoV) SARS-CoV-2, the causative agent of COVID-19, has resulted in a worldwide pandemic, threatening the lives of billions and imposing an immense burden on healthcare systems and the global economy. SARS-CoV-2, the seventh coronavirus known to infect humans, is a member of the *Betacoronavirus* genus which includes the highly pathogenic SARS-CoV and MERS-CoV, as well as endemic variants HCoV-OC43 and HCoV-HKU1[92]. Recent coronavirus outbreaks and the threat of future emerging zoonotic strains highlight the need for

broadly applicable coronavirus therapeutic interventions and vaccine design approaches[93].

Coronaviruses utilize the homotrimeric Spike (S) protein to engage with cell-surface receptors and enter host cells. S consists of two functional subunits: S1 and S2. S1 facilitates attachment to target cells and is composed of the N-terminal domain (NTD) and the receptor-binding domain (RBD), whereas S2, which encodes the fusion peptide and heptad repeats, promotes viral fusion[97, 150]. To facilitate cell entry, human coronaviruses employ different host factors; however, SARS-CoV and SARS-CoV-2 both utilize the cell-surface receptor angiotensin converting enzyme 2 (ACE2)[96]. Additionally, SARS-CoV-2 S shares 76% amino acid identity with SARS-CoV S[92]. Furthermore, S serves as a dominant antibody target and is a focus of countermeasure development for the treatment and prevention of COVID-19 infection[98, 99]. S proteins from the *Betacoronavirus* genus share multiple regions of structural homology and thus could serve as targets for a cross-reactive antibody response[89]. Identifying cross-reactive antibody epitopes can inform rational design strategies for vaccines and therapies that target multiple highly pathogenic coronaviruses.

Numerous potent neutralizing antibodies against SARS-CoV-2 have been discovered, including multiple candidates currently in clinical trials or approved for emergency use for prophylactic and acute treatment of COVID-19[151-158]. Investigation of SARS-CoV-2/SARS-CoV cross-reactive antibodies has focused primarily on the RBD epitope, which has resulted in the identification of a number of SARS-CoV-2/SARS-CoV cross-reactive antibody candidates[154, 159, 160]. However, the diversity of epitopes and functions beyond virus neutralization have not been extensively explored for cross-

reactive antibodies[161-163]. Evidence of Fc effector function contributing to protection *in vivo* against SARS-CoV[164] and SARS-CoV-2[165] suggests that the role of antibodies beyond neutralization may be a crucial component of protection and an important consideration in vaccine design strategies for coronaviruses[162, 166-168].

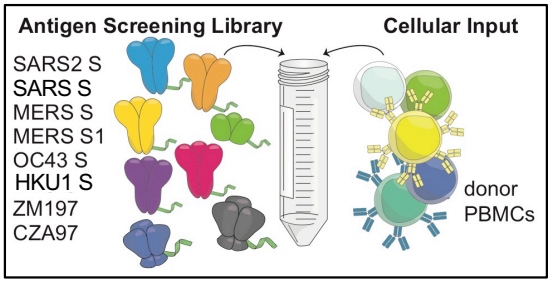
In this study, we investigated antibody cross-reactivity across the *Betacoronavirus* genus at monoclonal resolution. To do this, we applied LIBRA-seq (Linking B Cell receptor to antigen specificity through sequencing[116] to a recovered SARS-CoV donor sample from more than ten years after infection. We identified and characterized SARS-CoV-2/SARS-CoV cross-reactive human antibodies that target multiple, distinct structural domains of S, mediate phagocytosis and trogocytosis, and mitigate pathological burden *in vivo*. A better understanding of the genetic features, epitope specificities, and functional characteristics of cross-reactive coronavirus antibodies may translate into strategies for current vaccine design efforts and additional measures to counteract potential future pandemic strains.

3.1.2 Results

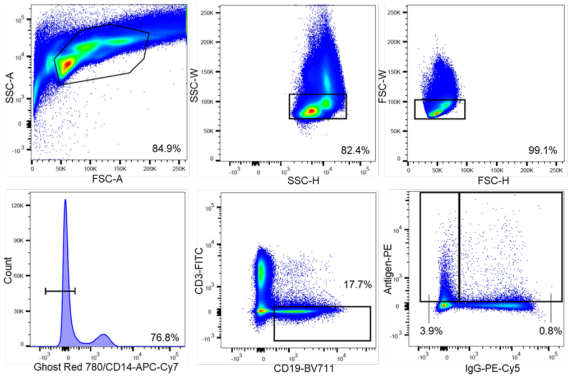
LIBRA-seq Characterization of a SARS-CoV Recovered Donor

To identify cross-reactive antibodies to multiple coronavirus antigens, LIBRA-seq was applied to a PBMC sample from a donor infected with SARS-CoV over ten years prior to sample collection. The antigen screening library consisted of eight oligo-tagged recombinant soluble antigens: six coronavirus trimer antigens (SARS-CoV-2 S, SARS-CoV S, MERS-CoV S, MERS-CoV S1 (with foldon domain), HCoV-OC43 S, HCoV-HKU1 S) and two HIV-1 trimer antigens from strains ZM197 and CZA97 as negative controls (**Figure 3.1.1A**). After the antigen screening library was mixed with donor PBMCs, antigen positive B cells were enriched by fluorescence activated cell sorting and processed for single-cell sequencing (**Figure 3.1.1B**). After bioinformatic processing, we recovered 2625 cells with paired heavy/light chain sequences and antigen reactivity information (**Figure 3.1.1C**), and from these cells, there were 2368 unique VDJ sequences. Overall, LIBRA-seq enabled rapid screening of PBMCs from a patient sample, with recovery of paired heavy/light chain sequences and antigen reactivity for thousands of B cells at the single-cell level.

A



B

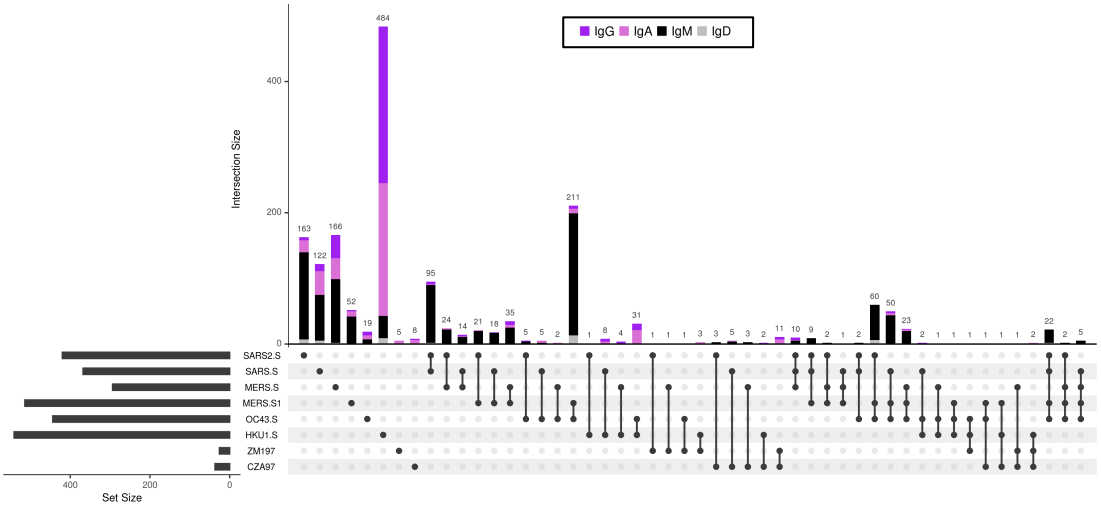


C

	4647-2 Processing
Unique Heavy Chain Barcodes	5113
Unique Heavy Chain Barcodes With Functional Heavy Chains	5072
Unique Heavy Chain Barcodes Without Cells With Multiple Heavy Chains	4875
Unique Light Chain Barcodes	7593
Unique Light Chain Barcodes With Functional Light Chains	7586
Unique Light Chain Barcodes Without Cells With Multiple Light Chains	5914
Unique Paired Heavy-Light Chain Barcodes	3813
Overlapping Barcodes Between Paired Heavy-Light Chains and Antigen Counts	2625

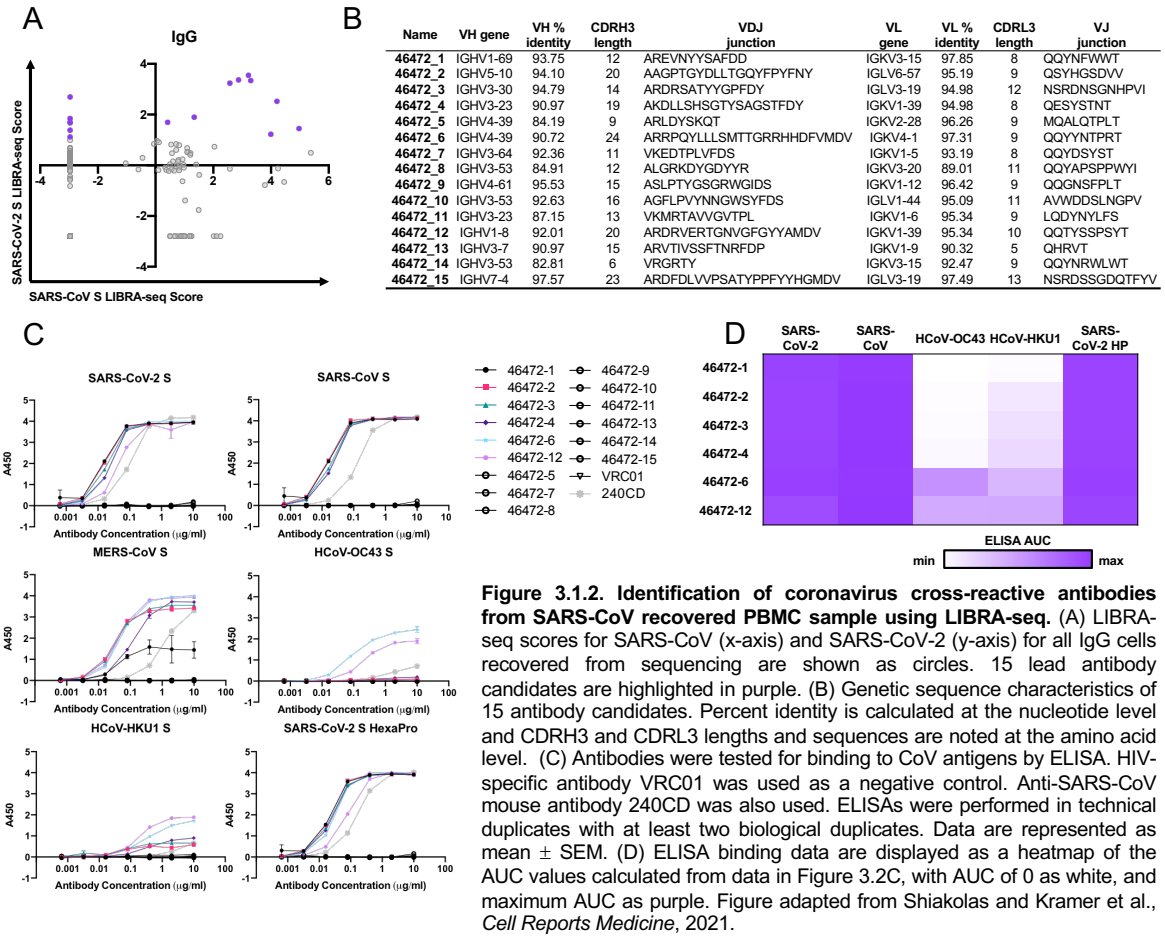
Figure 3.1.1. Application of LIBRA-seq to PBMCs from a recovered SARS-CoV donor sample led to identification of CoV cross-reactive antibodies (A) Schematic of DNA-barcoded antigens used to probe a SARS-CoV donor PBMC sample. (B) Gating scheme for fluorescent-activated cell sorting of recovered SARS-CoV donor. Cells were initially stained with Ghost Red 780, CD14-APC-Cy7, CD3-FITC, CD19-BV711, and IgG-PE-Cy5 along with a DNA-barcoded antigen screening library. To detect antigen-positive B cells, cells were washed and treated with a streptavidin-PE secondary stain. Gates as drawn are based on gates used during the sort, and percentages from the sort are listed. (C) The categorization of processing of Cell Ranger identified cells after sequencing is shown. (D) For each combination of CoV antigens, ZM197 Env, and CZA97 Env, the number of B cells with high LIBRA-seq scores (≥1) is displayed as a bar graph. The combination of antigens are displayed by filled circles, showing which antigens are part of a given combination. Each combination is mutually exclusive. The number of B cells with high LIBRA-seq scores for each antigen is indicated as a horizontal bar at the bottom left of the panel. Figure adapted from Shiakolas and Kramer et al., *Cell Reports Medicine*, 2021.

D



Identification of SARS-CoV-2 and SARS-CoV Cross-reactive Antibodies

With a goal of identifying antibodies that were cross-reactive to multiple coronavirus S proteins, we prioritized lead candidates based on their sequence features and LIBRA-seq scores (**Figure 3.1.1D**). We selected 15 antibody candidates that exhibited diverse sequence features and utilized a number of different variable genes for expression and characterization (**Figure 3.1.2A-B**). These antibodies displayed a broad range of percent identity to germline (83-98%) and a variety of CDRH3 and CDRL3 lengths (6-24 and 5-13 amino acids, respectively) (**Figure 3.1.2B**). By ELISA, SARS-CoV-2 S and SARS-CoV S binding was confirmed for 6/15 of the tested antibodies (46472-1, 46472-2, 46472-3, 46472-4, 46472-6, and 46472-12), indicating



Name	VH gene	VH % identity	CDRH3 length	VDJ junction	VL gene	VL % identity	CDRL3 length	VJ junction
46472_1	IGHV1-69	93.75	12	AREVNYYSAFDD	IGKV3-15	97.85	8	QQYNFWWT
46472_2	IGHV5-10	94.10	20	AAGPTGYDLLTGQYFPYFNY	IGLV6-57	95.19	9	QSYHGSDVV
46472_3	IGHV3-30	94.79	14	ARDRSATYYGPFDY	IGLV3-19	94.98	12	NSRDNSGNHPVI
46472_4	IGHV3-23	90.97	19	AKDLLSHSGTYSAGSTFDY	IGKV1-39	94.98	8	QESYSTNT
46472_5	IGHV4-39	84.19	9	ARLDYSKQT	IGKV2-28	96.26	9	MQALQTPLT
46472_6	IGHV4-39	90.72	24	ARRPQYLLLSMTTGRRHHDFVMDV	IGKV4-1	97.31	9	QQYYNTPRT
46472_7	IGHV3-64	92.36	11	VKEDTPLVFDS	IGKV1-5	93.19	8	QQYDSYST
46472_8	IGHV3-53	84.91	12	ALGRKDYGDYYR	IGKV3-20	89.01	11	QQYAPSPPWYI
46472_9	IGHV4-61	95.53	15	ASLPTYGSGRWGIDS	IGKV1-12	96.42	9	QQGNSFPLT
46472_10	IGHV3-53	92.63	16	AGFLPVYNNGWSYFDS	IGLV1-44	95.09	11	AVWDDSLNGPV
46472_11	IGHV3-23	87.15	13	VKMRTAVVGVTPL	IGKV1-6	95.34	9	LQDYNYLFS
46472_12	IGHV1-8	92.01	20	ARDRVERTGNVGFGYYAMDV	IGKV1-39	95.34	10	QQTYSSPSYT
46472_13	IGHV3-7	90.97	15	ARVTIVSSFTNRFDP	IGKV1-9	90.32	5	QHRVT
46472_14	IGHV3-53	82.81	6	VRGRTY	IGKV3-15	92.47	9	QQYNRWLWT
46472_15	IGHV7-4	97.57	23	ARDFDLVVPSATYPPFYYHGMDV	IGLV3-19	97.49	13	NSRDSSGDQTFYV

Figure 3.1.2. Identification of coronavirus cross-reactive antibodies from SARS-CoV recovered PBMC sample using LIBRA-seq. (A) LIBRA-seq scores for SARS-CoV (x-axis) and SARS-CoV-2 (y-axis) for all IgG cells recovered from sequencing are shown as circles. 15 lead antibody candidates are highlighted in purple. (B) Genetic sequence characteristics of 15 antibody candidates. Percent identity is calculated at the nucleotide level and CDRH3 and CDRL3 lengths and sequences are noted at the amino acid level. (C) Antibodies were tested for binding to CoV antigens by ELISA. HIV-specific antibody VRC01 was used as a negative control. Anti-SARS-CoV mouse antibody 240CD was also used. ELISAs were performed in technical duplicates with at least two biological duplicates. Data are represented as mean ± SEM. (D) ELISA binding data are displayed as a heatmap of the AUC values calculated from data in Figure 3.2C, with AUC of 0 as white, and maximum AUC as purple. Figure adapted from Shiakolas and Kramer et al., *Cell Reports Medicine*, 2021.

LIBRA-seq could successfully identify SARS-CoV-2 reactive B cells, but also suggesting potential differences in antigen binding detection for primary B cells with a sequencing readout vs. recombinant IgG by ELISA (**Figure 3.1.2C-D**). Further, antibodies 46472-6 and 46472-12 bound to S proteins from endemic HCoV-OC43 and HCoV-HKU1, albeit generally at lower levels (**Figure 3.1.2C-D**). Although the six monoclonal antibodies showed reactivity by ELISA to the MERS-CoV antigen probe used in the LIBRA-seq screening library, antibody binding to other independent preparations of this protein was inconsistent, so MERS-CoV S reactivity could not be confirmed definitively (**Figure 3.1.3A-B**). Overall, the application of the LIBRA-seq technology enabled the identification of a panel of cross-reactive antibodies that recognize the S antigen from multiple coronaviruses.

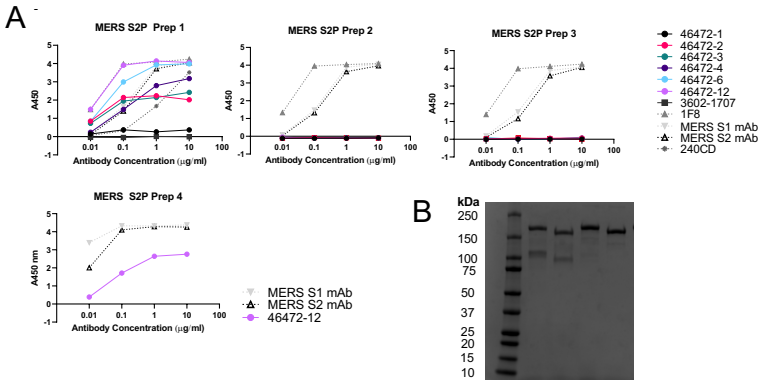


Figure 3.1.3. Recombinant MERS-CoV Spike Protein Characterization and Quality Control. (A) ELISA binding data to independent preparations of MERS-CoV S protein. An influenza HA-specific mAb 3602-1707 was used as a negative control along with positive control antibodies 1F8 (expressed and purified recombinantly) and MERS S1 mAb and MERS S2 mAbs (Sino Biological). (B) SDS-PAGE gels of MERS-CoV S2P S preparations are shown (preparation 1 – lane 1 & 2, preparation 2 – lane 3 & 4) under both non-reducing conditions (lanes 1 & 3) and reducing conditions (lanes 2 & 4). Figure adapted from Shiakolas and Kramer et al., *Cell Reports Medicine*, 2021.

Cross-reactive Coronavirus Antibodies Target Diverse Epitopes on S

To elucidate the epitopes targeted by the cross-reactive antibodies, we performed binding assays to various structural domains of S as well as binding-competition experiments. First, we assessed antibody binding to the S1 and S2 subdomains of SARS-CoV-2. Antibodies 46472-1, 46472-2, 46472-3, and 46472-4

bound to the S2 domain, whereas 46472-6 and 46472-12 recognized the S1 domain but targeted different epitopes, the NTD and RBD, respectively (**Figure 3.1.4A-E**). Although 46472-12 bound to the RBD, it did not compete with ACE2 for binding to SARS-CoV-2 S and showed partial competition with RBD-directed antibody CR3022 (**Figure 3.1.4F-G**). To determine whether the antibodies targeted overlapping or distinct epitopes, we performed competition ELISA experiments and found that the S2-directed antibodies 46472-1, 46472-2, and 46472-4 competed for binding to S (**Figure 3.1.5**). This pattern was observed for both SARS-CoV-2 and SARS-CoV S. Of note, this competition group did not include S2-directed antibody 46472-3, revealing the identification of multiple cross-reactive epitope targets on S2 (**Figure 3.1.5**). Further, antibody binding was not affected by two glycan knockout mutants (N165A or N709A) or mannose competition (**Figure 3.1.6A-B**). Lastly, we measured antibody autoreactivity, and found that with the exception of 46472-6 binding to Jo-1, none of the antibodies showed autoreactivity against the tested antigens (**Figure 3.1.6C**). Together, these data suggest that the identified cross-reactive antibodies are coronavirus-specific and target multiple, diverse epitopes on the S protein (**Figure 3.1.6D**).

Functional Characterization of Cross-reactive Coronavirus Antibodies

Next, we characterized our cross-reactive antibody panel for functional activity. Although none of the antibodies neutralized SARS-CoV or SARS-CoV-2 (**Figure 3.1.7A-B**), all antibodies showed antibody-dependent cellular phagocytosis (ADCP) *in vitro* for SARS-CoV-2 S (**Figure 3.1.8A**). In particular, the RBD-reactive antibody

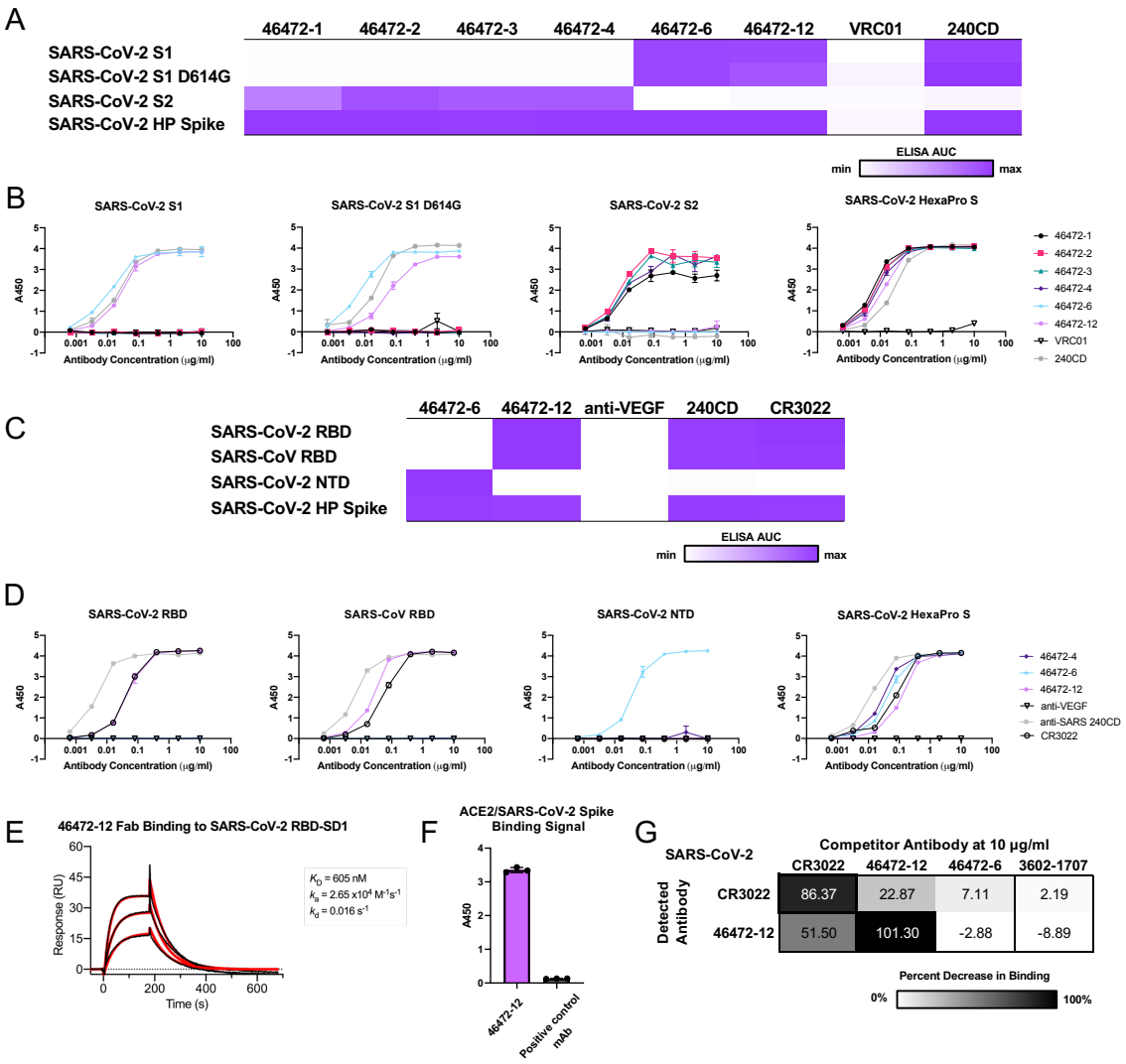


Figure 3.1.4. Epitope mapping of cross-reactive antibodies. (A) For cross-reactive coronavirus antibodies, ELISA data against the antigens are displayed as a heatmap of the AUC values calculated from the data in Figure 3.1.4B. (B) Cross-reactive antibodies were tested for binding to SARS-CoV-2 S1 domain, SARS-CoV-2 S1 domain D614G, SARS-CoV-2 S2 domain, and SARS-CoV-2 S (HexaPro). Anti-HIV antibody VRC01 is shown as a negative control and anti-SARS-CoV antibody 240CD is shown as a positive control. Data are represented as mean ± SEM. (C) For SARS-CoV-2 S1 reactive antibodies, ELISA data against the RBD and NTD are displayed as a heatmap of the AUC values calculated from the data in Figure 3.1.4D. AUC of 0 is displayed as white and maximum AUC as purple. ELISA data are representative of at least two independent experiments. Anti-HIV antibody VRC01 and anti-VEGF antibody are shown as a negative control, and anti-SARS-CoV antibody 240CD is shown as a positive control. (D) S1-directed antibodies 46472-6 and 46472-12 were tested for binding against SARS-CoV-2 RBD, SARS-CoV RBD, SARS-CoV-2 NTD, and SARS-CoV-2 S (HexaPro). Anti-HIV antibody VRC01 is shown as a negative control and anti-SARS-CoV antibody 240CD is shown as positive control. Data are represented as mean ± SEM. (E) Surface plasmon resonance binding of 46472-12 Fab to SARS-CoV-2 RBD. Affinity measurements are shown to the right of the graph. Figure 3.1.4E made by Daniel Lingwood. Figure adapted from Shiakolas and Kramer et al., *Cell Reports Medicine*, 2021.

46472-12 showed greater ADCP activity compared to the other cross-reactive antibodies and the SARS-CoV/SARS-CoV-2 cross-reactive RBD antibody control, CR3022[169] (**Figure 3.1.8A**). Further, we tested and confirmed ADCP activity against

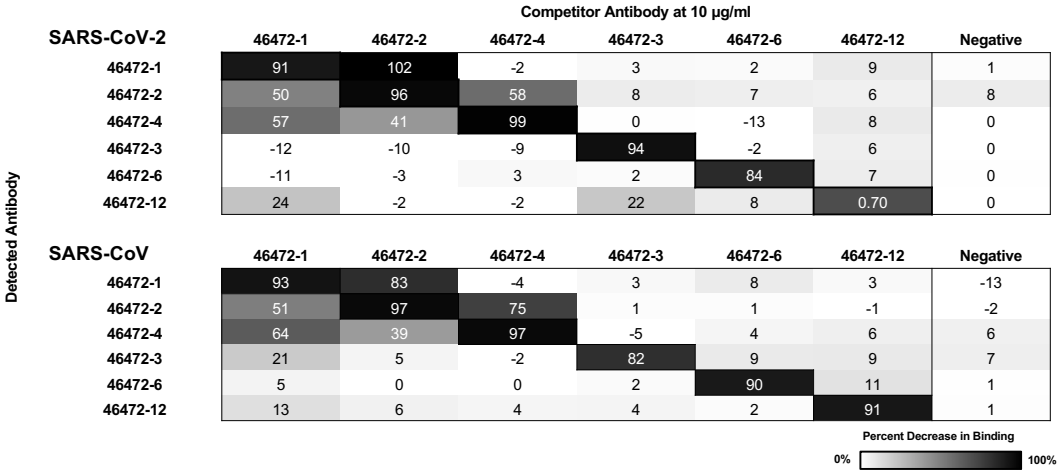


Figure 3.1.5. Cross-reactive CoV antibodies target distinct domains of SARS-CoV-2 S Cross-reactive CoV antibodies target distinct domains of SARS-CoV-2 S. Cross-reactive antibodies were used in a competition ELISA to determine if binding of one antibody affected binding of another. Competitor antibodies were added at 10 µg/ml, and then detected antibodies were added at 0.1 µg/ml. The percent reduction in binding compared to binding without a competitor is shown. An anti-HIV antibody was used as a negative control. ELISAs were performed in technical duplicates with at least two biological duplicates. Figure adapted from Shiakolas and Kramer et al., *Cell Reports Medicine*, 2021.

SARS-CoV for two antibodies that mediated the highest phagocytotic activity against SARS-CoV-2, 46472-4 and 46472-12, illustrating that these antibodies have cross-coronavirus phagocytic ability (**Figure 3.1.8B**). We next tested the antibodies in a trogocytosis assay[148] and found that four antibodies in our panel (46472-1, 46472-2, 46472-3, and 46472-4) mediated trogocytosis (**Figure 3.1.8C**). This warrants further investigation as this is the first description of trogocytosis performed by SARS-CoV-2 specific antibodies. Lastly, none of the antibodies promoted complement deposition (ADCD) (**Figure 3.1.8D**). Together, these results revealed different profiles of Fc effector functionality within the panel of cross-reactive antibodies.

Since non-neutralizing SARS-CoV-2 antibodies with Fc effector function activity have not been extensively characterized *in vivo*, these results prompted us to test antibodies 46472-4 and 46472-12 for prophylaxis in a murine infection model using a mouse-adapted virus strain (SARS-CoV-2 MA)[170, 171] at a non-lethal dose of $1x10^3$ PFU

(**Figure 3.9A**). Although there were no differences in survival and viral load between experimental and control groups, the lung hemorrhage scores (see Methods) for

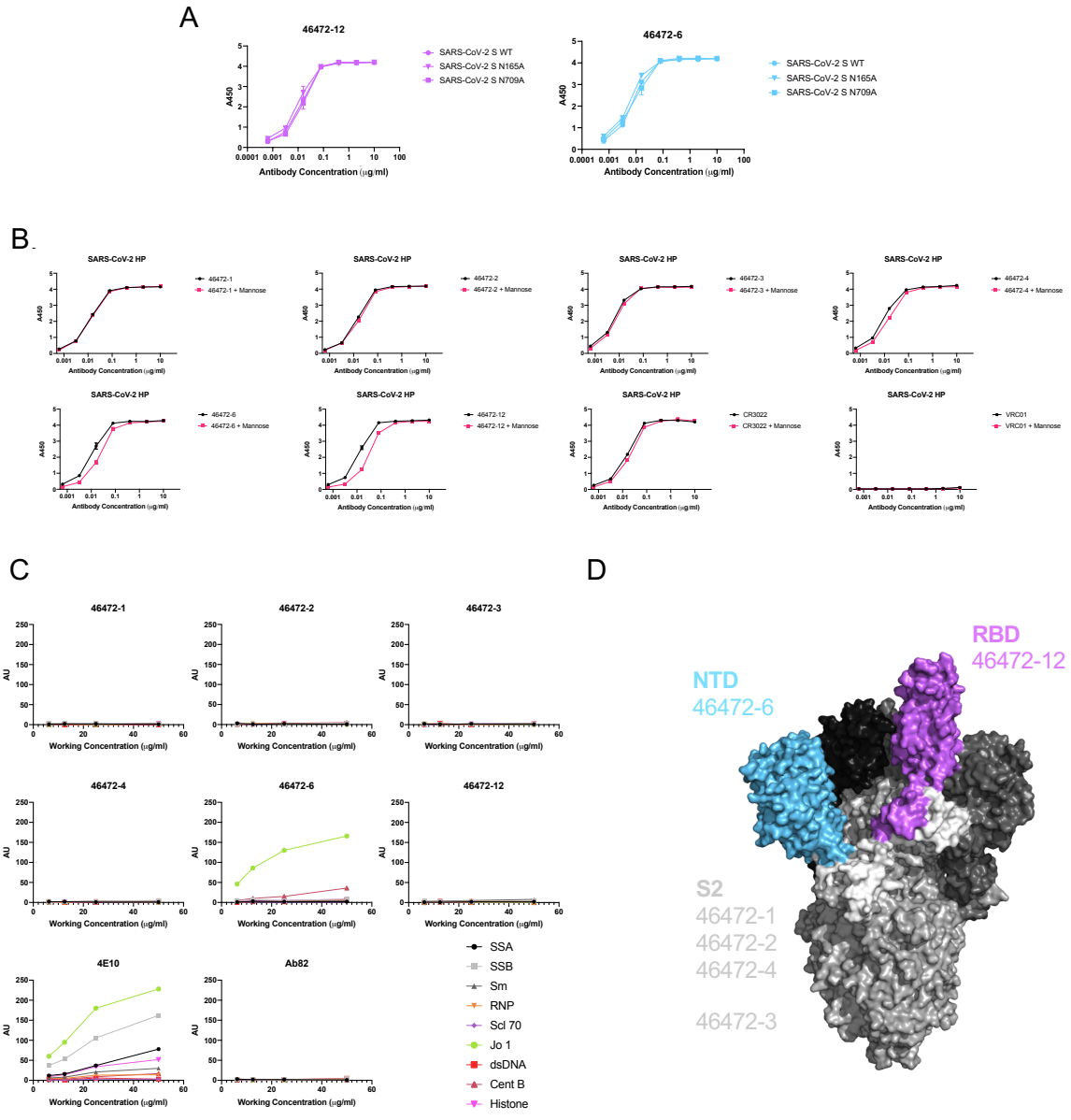


Figure 3.1.6. Antibody Reactivity and Epitope Characterization (A) 46472-6 and 46472-12 were tested for binding to SARS-CoV-2 S (HexaPro) mutants, N165A and N709A, by ELISA. Antibody concentration is shown on the x-axis and absorbance at 450 nm is shown on the y-axis. ELISAs were performed in technical duplicates with at least two biological duplicates. Data are represented as mean ± SEM. (B) Mannose competition ELISA binding assays were performed to determine if cross-reactive antibody binding to SARS-CoV-2 S could be modulated by 1M D-(+)-Mannose. Binding of each antibody with (pink) or without (black) mannose is shown as a separate plot. Antibody concentration is shown on the x-axis and absorbance at 450 nm is shown on the y-axis. Antigen-specific antibody CR3022 and HIV-specific negative control antibody VRC01 are also shown. ELISAs were performed in technical duplicates with at least two biological duplicates. Data are represented as mean ± SEM. (C) Antibodies were tested for autoreactivity against a variety of antigens in the Luminex AtheNA assay. AU stands for Athena Units. Anti-HIV antibody 4E10 was used as a positive control and Ab82 was used as a negative control. (D) Cross-reactive coronavirus antibodies target a variety of epitopes on the SARS-CoV-2 S protein, including the RBD, NTD, and S2 domains, highlighted on the structure (PDB: 6VSB). Figure 3.1.6D made by Clint Holt. Figure adapted from Shiakolas and Kramer et al., *Cell Reports Medicine*, 2021.

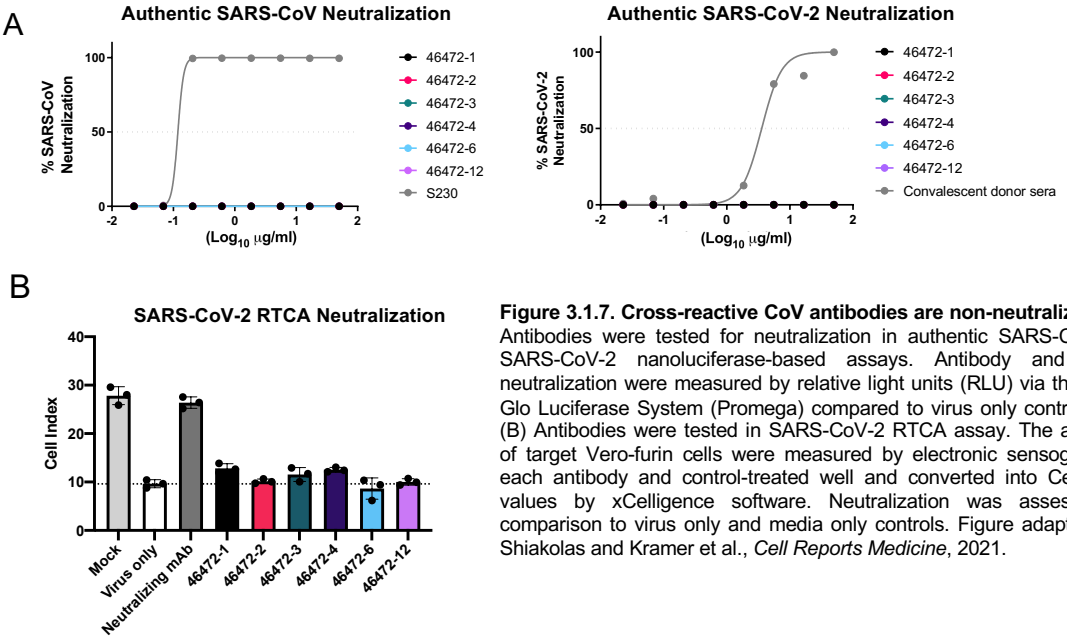


Figure 3.1.7. Cross-reactive CoV antibodies are non-neutralizing. (A) Antibodies were tested for neutralization in authentic SARS-CoV and SARS-CoV-2 nanoluciferase-based assays. Antibody and serum neutralization were measured by relative light units (RLU) via the Nano-Glo Luciferase System (Promega) compared to virus only control wells. (B) Antibodies were tested in SARS-CoV-2 RTCA assay. The adhesion of target Vero-furin cells were measured by electronic sensograms of each antibody and control-treated well and converted into Cell Index values by xCelligence software. Neutralization was assessed by comparison to virus only and media only controls. Figure adapted from Shiakolas and Kramer et al., *Cell Reports Medicine*, 2021.

46472-4 and 46472-12 were similar to antigen-specific control CR3022, and all three groups were significantly lower than the scores for isotype control DENV-2D22 ($p<0.01$, ordinary one-way ANOVA with multiple comparisons) (**Figure 3.1.9B-C**). To evaluate the *in vivo* effect of these antibodies in a more stringent challenge model in 12-month old female BALB/c mice, we increased the viral dose from $1x10^3$ to $1x10^4$ PFU. In this experiment, mice that received antibody 46472-12 exhibited the best survival rate (4/5 at day 4), compared to the other treatment groups that included CR3022 as an antigen-specific control and DENV-2D22 as a negative control, although statistical significance was not achieved (**Figure 3.1.9B, D-E**). There were no significant differences in viral load between groups; however, the surviving animals from the 46472-4 and 46472-12 groups showed significantly lower hemorrhagic pathology scores in harvested mouse lungs compared to the negative control treatment group ($p<0.001$, ordinary one-way ANOVA with multiple comparisons) (**Figure 3.1.9D**). Animals treated with the antigen-

specific control, CR3022, had significantly higher hemorrhage scores than animals treated with 46472-4 and 46472-12 ($p<0.001$, ordinary one-way ANOVA with multiple comparisons), although the statistical analysis may be limited by the small numbers of surviving animals for some of the groups (**Figure 3.1.9D**). While definitive evidence for protection is limited, the data from the *in vivo* experiments suggests that these cross-reactive antibodies could contribute to counteracting coronavirus infection in prophylaxis.

3.1.3 Discussion

Here, we described a set of cross-reactive *Betacoronavirus* antibodies isolated from a recovered SARS-CoV donor. The antibodies targeted diverse epitopes on S, including the S2 subdomain as well as the RBD and NTD on S1, and demonstrated Fc effector function *in vitro*. Additionally, two of these antibodies were tested *in vivo*, and displayed a reduction in lung hemorrhage score, while effects on viral load were not definitive.

Given the similar effect of 46472-4 and 46472-12 on severe disease in the mouse model, their phagocytotic ability along with the inability to mediate neutralization suggests that the former may be a mechanism through which they function, and additional studies are underway to further assess this hypothesis. Phagocytosis has been shown to be associated with protection in a SARS-CoV-2 DNA vaccination in non-human primates as well as survival in natural infection[172] and as such could be an important mechanism for protection by monoclonal antibodies. The role of trogocytosis

in COVID-19 is unknown as are the targets that may be important for this function.

46472-4 was able to mediate this membrane nibbling in contrast to 46472-12,

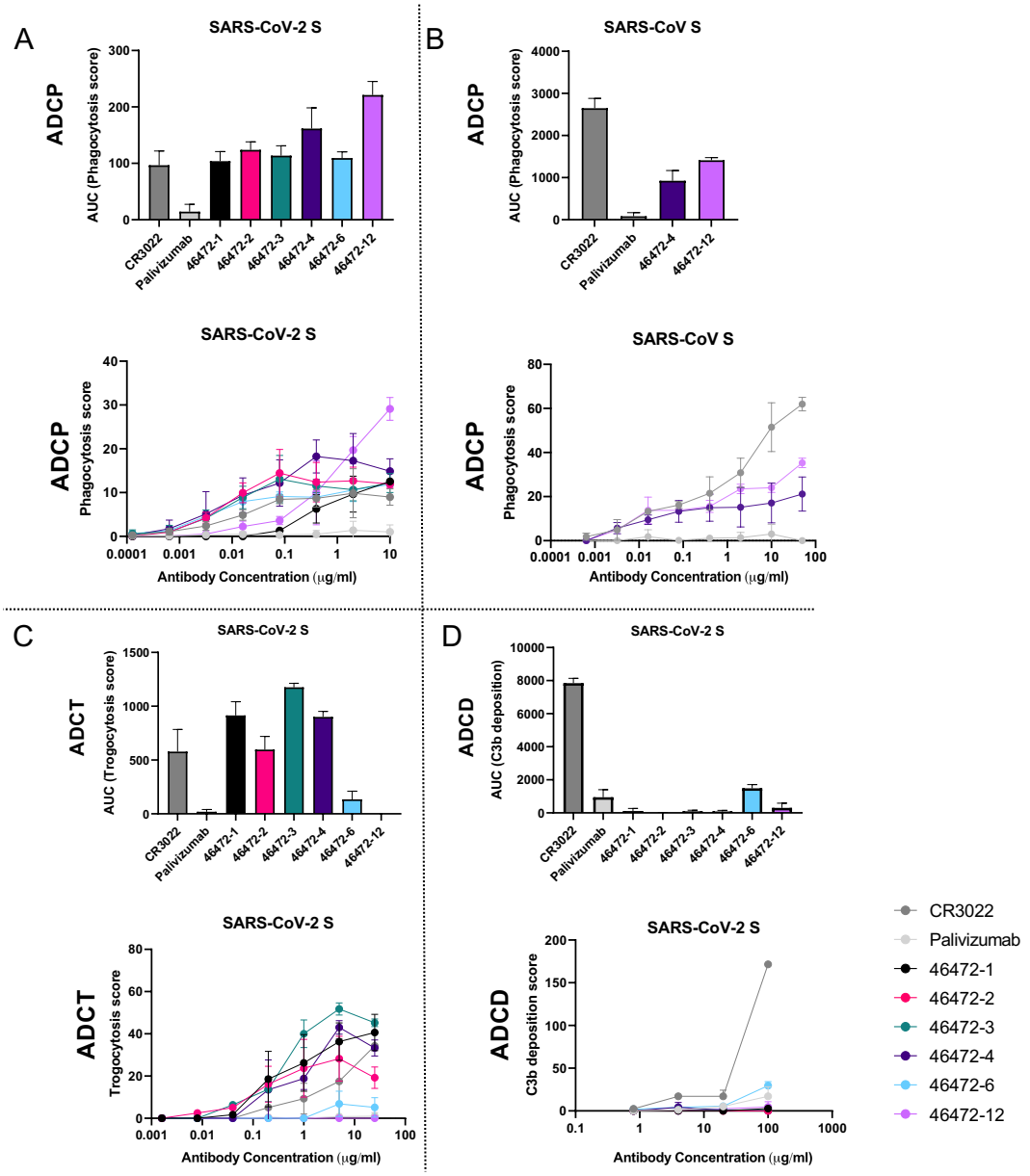
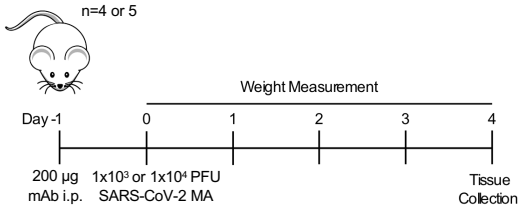


Figure 3.1.8. Functional activity of cross-reactive coronavirus antibodies. (A) Cross-reactive coronavirus antibodies were tested for antibody-dependent cellular phagocytosis activity (ADCP) against SARS-CoV-2 S, compared to positive control CR3022 and negative control Palivizumab, an anti-RSV antibody. AUC of the phagocytosis score is shown, calculated from data shown below. Data are represented as mean ± SD. (B) 46472-4 and 46472-12 were tested for ADCP activity against SARS-CoV S, compared to CR3022 and anti-RSV Palivizumab. AUC of the phagocytosis score is shown, calculated from data shown below. Data are represented as mean ± SD. (C) Cross-reactive coronavirus antibodies were tested for antibody-dependent cellular trogocytosis (ADCT) activity against SARS-CoV-2 S displayed on transfected cells, compared to positive control CR3022 and anti-RSV Palivizumab. AUC of the trogocytosis score is shown, calculated from data shown below. Data are represented as mean ± SD. (D) Cross-reactive coronavirus antibodies were tested for antibody-dependent complement deposition (ADCD) activity against SARS-CoV-2 S, compared to positive control CR3022 and anti-RSV Palivizumab. AUC of the C3b deposition score is shown, calculated from data shown below. Data are represented as mean ± SD. Figure made by Simone Richardson. Figure adapted from Shiakolas and Kramer et al., *Cell Reports Medicine*, 2021.

suggesting that this function in addition to complement activity was not responsible for the *in vivo* effect on severe disease mediated by these antibodies. Although the precise *in vivo* effects of these antibodies have not been elucidated, the identification of multiple, cross-reactive antibodies highlights a potential role for Fc effector function activity, specifically phagocytosis, in coronavirus infection. Evidence of protection associated with Fc effector function in SARS-CoV[164], SARS-CoV-2[165, 166, 173], and other infectious diseases including influenza, Ebola, and HIV-1, motivates further investigation into its contribution for the treatment of COVID-19[36, 38, 39, 174]. Furthermore, the importance of Fc effector functionality of potently neutralizing candidate clinical SARS-CoV-2 antibodies in a therapeutic setting rather than prophylaxis highlights the potential benefit for investigation into non-neutralizing antibodies with phagocytic activity and their administration after infection onset[175]. Elucidation of the functional roles of cross-reactive but non-neutralizing antibodies could have implications for understanding the factors involved in protection or enhancement of disease.

Given the ongoing SARS-CoV-2 pandemic and the potential for future zoonotic coronavirus pathogens to emerge, coronavirus vaccine and therapeutic development is of paramount importance[91, 94, 95, 176]. Antibodies that can cross-react with multiple coronavirus strains are primary targets as potential broadly reactive therapies. Such antibodies can further reveal cross-reactive epitopes that could serve as templates for the development of broadly protective vaccines. Understanding the spectrum of cross-reactive epitopes targeted by human antibodies, as well as the functional role that such antibodies have within coronavirus infection, are therefore a vital element of medical countermeasure development.

A



B

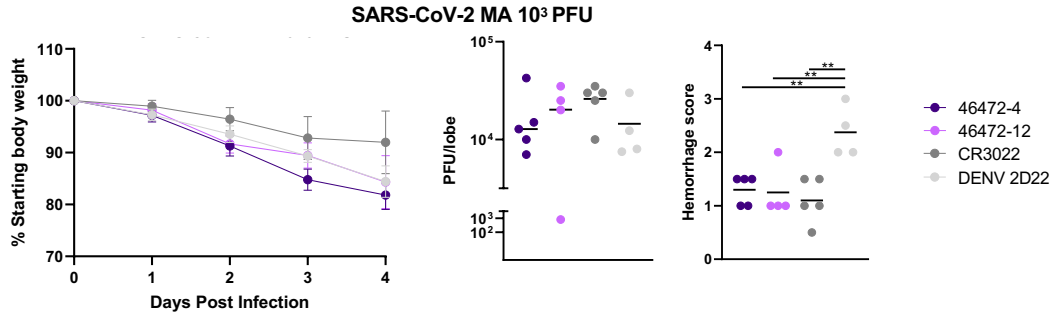
SARS-CoV-2 MA 10^3 PFU

Antibody	Animal Survival, Days Post Infection				
	0	1	2	3	4
46472-12	4/4	4/4	4/4	4/4	4/4
46472-4	5/5	5/5	5/5	5/5	5/5
CR3022	5/5	5/5	5/5	5/5	5/5
DENV-2D22	5/5	5/5	5/5	5/5	4/5

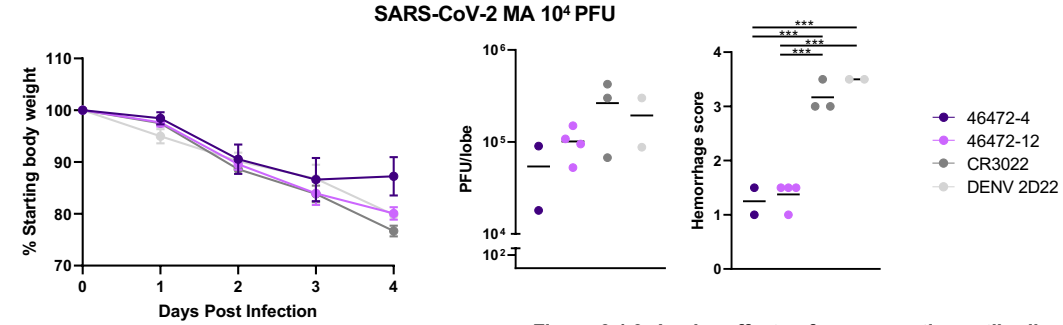
SARS-CoV-2 MA 10^4 PFU

Antibody	Animal Survival, Days Post Infection				
	0	1	2	3	4
46472-12	5/5	5/5	5/5	5/5	4/5
46472-4	5/5	5/5	5/5	4/5	2/5
CR3022	5/5	5/5	5/5	4/5	3/5
DENV-2D22	5/5	5/5	5/5	5/5	2/5

C



D



E

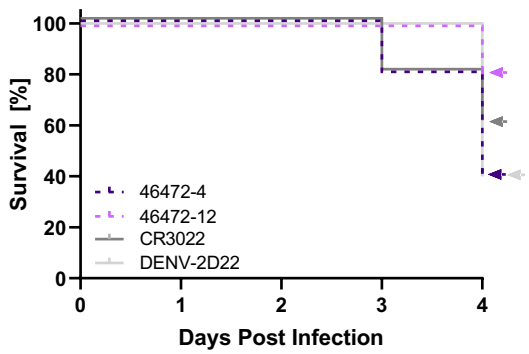


Figure 3.1.9. *In vivo* effects of cross-reactive antibodies (A) Timeline of the prophylactic antibody experiment in SARS-CoV-2 mouse adapted (MA) *in vivo* infection model. (B) For each antibody treatment group for the experiment utilizing 1x10^3 PFU or 1x10^4 PFU of SARS-CoV-2 MA, a table showing the number of animals to survive per group per day is shown. (C, D) For each antibody treatment group for the experiment utilizing (C) 1x10^3 PFU or (D) 1x10^4 PFU of SARS-CoV-2 MA, shown are daily body weight progression, and terminal RT-qPCR quantification of lung viral titer and lung hemorrhage scores of gross pathology. For viral titer values and the lung hemorrhage scores, an ordinary one-way ANOVA test with multiple comparisons was performed. (E) For the experiment with 1x10^4 PFU of SARS-CoV-2 MA, percent survival for each antibody group is shown. Figure made with Alexandra Schäfer and Kevin Kramer. Figure adapted from Shiakolas and Kramer et al., *Cell Reports Medicine*, 2021.

Limitations of the Study

The current study focuses on the characterization of cross-reactive coronavirus antibodies, mostly in the context of SARS-CoV-2. Further characterization of this panel of antibodies against circulating endemic coronavirus strains would enhance the clinical relevance to less severe coronavirus-associated respiratory infections.

The current study utilized a dosing regimen in a prophylactic setting and given the emerging evidence of survival benefit with effector function in antibodies given after infection onset[175], antibody administration in a therapeutic setting may provide further insight into *in vivo* properties. Furthermore, additional effector function characterization such as ADCC and ADNP would strengthen the profile of this panel of non-neutralizing antibodies given their role in both human[177] and mouse SARS-CoV-2 infection studies.

Acknowledgements

We thank Angela Jones, Latha Raju, and Jamie Roberson of Vanderbilt Technologies for Advanced Genomics for their expertise regarding NGS and library preparation; David Flaherty and Brittany Matlock of the Vanderbilt Flow Cytometry Shared Resource for help with flow panel optimization; and members of the Georgiev laboratory for comments on the manuscript. The Vanderbilt VANTAGE Core provided technical assistance for this work. VANTAGE is supported in part by CTSA grant 5UL1 RR024975-03, the Vanderbilt Ingram Cancer Center (P30 CA68485), the Vanderbilt Vision Center (P30 EY08126), and NIH/NCRR (G20 RR030956). This work was conducted in part using the resources of the Advanced Computing Center for Research and Education at Vanderbilt University (Nashville, TN). Flow cytometry experiments

were performed in the VUMC Flow Cytometry Shared Resource. The VUMC Flow Cytometry Shared Resource is supported by the Vanderbilt Ingram Cancer Center (P30 CA68485) and the Vanderbilt Digestive Disease Research Center (DK058404).

For work described in this manuscript, I.S.G., A.R.S., K.J.K., S.W., K.A.P., R.V., N.R., E.F.F., C.M.H. were supported in part by NIH NIAID award R01AI131722-S1, the Hays Foundation COVID-19 Research Fund, Fast Grants, and *CTSA award No.* UL1 TR002243 *from the National Center for Advancing Translational Sciences*. J.S.M and D.W. were supported in part by a National Institutes of Health (NIH)/National Institute of Allergy and Infectious Diseases (NIAID) grant awarded to J.S.M. (R01-AI127521). L.M. and S.I.R. acknowledge research funding from the South African Medical Research Council (MRC) Extramural Unit and SHIP-COVID19 programs and an H3 Africa grant (U01A136677). S.I.R. is supported by the South African Research Chairs Initiative of the Department of Science and Technology and the NRF (Grant No 98341) and is a L'Oreal/UNESCO South Africa Young Talents Awardee. R.B., A.S., D.R.M., were supported by NIH grants (U54CA260543, R01AI157155). P.A. and K.J. were supported by NIH grant R01 AI14567. J.E.C., R.H.C., N.S., R.N.S., and R.E.S., were supported by Defense Advanced Research Projects Agency (DARPA) grants HR0011-18-2-0001 and HR00 11-18-3-0001; NIH contracts 75N93019C00074 and 75N93019C00062; NIH grants U01 AI150739, R01 AI130591 and R35 HL145242; the Dolly Parton COVID-19 Research Fund at Vanderbilt; and NIH grant S10 RR028106 for the Next Generation Nucleic Acid Sequencer, housed in VANTAGE. M.S.D. and R.E.C. were supported by grants from NIH (R01 AI157155) and the Defense Advanced Research Project Agency (HR001117S0019). B.F.H. and R.P. were supported by NC State funding for COVID

research. B.S.G. was supported by intramural funding from the NIAID. C.M.H. was supported in part by NIH grant T32 GM008320-30. D.R.M. was supported by an NIH F32 AI152296, a Burroughs Wellcome Fund Postdoctoral Enrichment Program Award, and was previously supported by an NIH NIAID T32 AI007151.

Author Contributions

Methodology, A.R.S., K.J.K., and I.S.G.; Investigation, A.R.S., K.J.K., D.W., S.I.R., A.S., S.W., N.W., K.J., K.A.P., R.V., R.P., N.P.M., N.R., E.F.F., C.M.H., N.S., R.E.C., D.R.M., R.S.N., R.E.S., J.E.L., B.S.G., M.S.D., B.F.H., P.A., R.H.C., J.E.C., R.S.B., L.M., J.S.M., and I.S.G.; Software, A.R.S., R.V., N.R.; Validation, A.R.S., K.J.K.; Writing - Original Draft, A.R.S., and K.J.K.; Writing -Review & Editing, all authors; Funding Acquisition, I.S.G., B.S.G., M.S.D., B.F.H., P.A., R.H.C., J.E.C., R.S.B., L.M., J.S.M., and A.R.S, and K.J.K.; Resources, B.S.G., M.S.D., B.F.H., P.A., R.H.C., J.E.C., R.S.B., L.M., J.S.M., and I.S.G.; Supervision, I.S.G.

Declaration of Interests

A.R.S. and I.S.G are co-founders of AbSeek Bio. A.R.S., K.J.K, I.S.G., D.W., N.W., and J.S.M are listed as inventors on patents filed describing the antibodies described here. D.W., J.S.M, B.S.G, and N.W. are also listed as inventors on U.S. patent application no. 62/972,886 (2019-nCoV Vaccine). M.S.D. is a consultant for Inbios, Vir Biotechnology, NGM Biopharmaceuticals, and Carnival Corporation and on the Scientific Advisory Boards of Moderna and Immunome. The Diamond laboratory has unrelated sponsored research agreements from Emergent BioSolutions, Moderna and

Vir Biotechnology. J.E.C. has served as a consultant for Eli Lilly, GlaxoSmithKline and Luna Biologics, is a member of the Scientific Advisory Boards of CompuVax and Meissa Vaccines and is Founder of IDBiologics. The Crowe laboratory at Vanderbilt University Medical Center has received sponsored research agreements from IDBiologics and AstraZeneca. R.S.B. has competing interests associated with Eli Lily, Takeda and Pfizer. The Georgiev laboratory at Vanderbilt University Medical Center has received unrelated funding from Takeda Pharmaceuticals.

3.1.4 Materials and Methods

Data and Code Availability

Sequences for antibodies identified and characterized in this study have been deposited to GenBank under GenBank accession numbers MZ126644-MZ126658 (heavy chain) and MZ126659-MZ126673 (light chain). Raw sequencing data used in this study are available on the Sequence Read Archive under BioProject accession number PRJNA727275. Custom scripts used to analyze data in this manuscript are available upon request to the corresponding author.

Human Subjects

The donor had prior SARS-CoV infection during the 2004 outbreak in Hong Kong, and the PBMC sample was collected over 10 years post infection (20 million PBMCs). Additional information about the donor is not available.

Cell Lines

A variety of cell lines were utilized for various assays in this study. Expi293F mammalian cells (ThermoFisher) were maintained in FreeStyle F17 expression medium supplemented at final concentrations of 0.1% Pluronic Acid F-68 and 20% 4mM L-Glutamine. These cells were cultured at 37°C with 8% CO_2 saturation and shaking.

FreeStyle293F cells were grown while shaking at 37 C in 8% CO2 and 80% humidity. Freestyle293F cells are derived from female human embryonic kidney epithelial cells.

THP-1 cells obtained from the AIDS Reagent Program (Division of AIDS, NIAID, NIH contributed by Dr. Li Wu and Vineet N. KewalRamani) were used for both the ADCP and ADCT assays. Cells were cultured at 37°C, 5% CO2 in RPMI containing 10% heat-inactivated fetal bovine serum (Gibco, Gaithersburg, MD), 1% Penicillin Streptomycin (Gibco, Gaithersburg, MD) and 2-mercaptoethanol to a final concentration of 0.05 mM. These cells were not allowed to exceed 4 x 105 cells/ml to prevent differentiation and are from a male donor.

HEK293T cells were obtained from Dr George Shaw and were used for the ADCT assay. These adherent cell lines were cultured at 37°C, 5% CO2, in DMEM containing 10% heat-inactivated fetal bovine serum (Gibco BRL Life Technologies) and supplemented with 50 µg/ml gentamicin (Sigma). Cells were disrupted at confluence with 0.25% trypsin in 1 mM EDTA (Sigma) every 48–72 hours. HEK293F suspension cells were cultured in 293Freestyle media (Gibco BRL Life Technologies) and grown in a shaking incubator at 37°C, 5% CO2, 70% humidity at 125rpm. Cells were diluted twice

a week to between 0.2 and 0.5 million cells/ml. Both HEK293 derived cell lines are from female donors.

Murine Model

12-month old female BALB/c mice (BALB/cAnHsd; Envigo, stock number 047) were used in a murine infection model for SARS-CoV-2 with a mouse adapted strain. Eleven to twelve-month old female BALB/c mice (BALB/c AnNHsd, Envigo, stock# 047) were used for mouse-adapted SARS-CoV-2 (SARS-CoV-2 MA10) *in vivo* protection experiments as described previously[171]. All mouse studies were performed at the University of North Carolina (Animal Welfare Assurance #A3410-01) using protocols (19-168) approved by the UNC Institutional Animal Care and Use Committee (IACUC) and were performed in a BSL3 facility at UNC.

Antigen Purification

A variety of recombinant soluble protein antigens were used in the LIBRA-seq experiment and other experimental assays.

Plasmids encoding residues 1–1208 of the SARS-CoV-2 spike with a mutated S1/S2 cleavage site, proline substitutions at positions 986 and 987, and a C-terminal T4-fibritin trimerization motif, an 8x HisTag, and a TwinStrepTag (SARS-CoV-2 S-2P); residues 1-1190 of the SARS-CoV spike with proline substitutions at positions 968 and 969, and a C-terminal T4-fibritin trimerization motif, an 8x HisTag, and a TwinStrepTag (SARS-CoV S-2P); residues 1-1291 of the MERS-CoV spike with a mutated S1/S2 cleavage site, proline substitutions at positions 1060 and 1061, and a C-terminal T4-

fibritin trimerization motif, an AviTag, an 8x HisTag, and a TwinStrepTag (MERS-CoV S-2P Avi); residues 1-751 of the MERS-CoV spike with a C-terminal T4-fibritin trimerization motif, 8x HisTag, and a TwinStrepTag (MERS-CoV S1); residues 1-1277 of the HCoV-HKU1 spike with a mutated S1/S2 cleavage site, proline substitutions at positions 1067 and 1068, and a C-terminal T4-fibritin trimerization motif, an 8x HisTag, and a TwinStrepTag (HCoV-HKU1 S-2P); residues 1-1278 of the HCoV-OC43 spike with proline substitutions at positions 1070 and 1071, and a C-terminal T4-fibritin trimerization motif, an 8x HisTag, and a TwinStrepTag (HCoV-OC43 S-2P); or residues 319–591 of SARS-CoV-2 S with a C-terminal monomeric human IgG Fc-tag and an 8x HisTag (SARS-CoV-2 RBD-SD1) were transiently transfected into FreeStyle293F cells (Thermo Fisher) using polyethylenimine. The coronavirus trimer spike antigens were in a prefusion-stabilized (S-2P) conformation that better represents neutralization-sensitive epitopes in comparison to their wild-type forms[178]. Two hours post-transfection, cells were treated with kifunensine to ensure uniform glycosylation. Transfected supernatants were harvested after 6 days of expression. SARS-CoV-2 RBD-SD1 was purified using Protein A resin (Pierce), SARS-CoV-2 S-2P, SARS-CoV S-2P, MERS-CoV S-2P Avi, MERS-CoV S1, HCoV-HKU1 S-2P and HCoV-OC43 S-2P were purified using StrepTactin resin (IBA). Affinity-purified SARS-CoV-2 RBD-SD1 was further purified over a Superdex75 column (GE Life Sciences). MERS-CoV S1 was purified over a Superdex200 Increase column (GE Life Sciences). SARS-CoV-2 S-2P, SARS-CoV S-2P, MERS-CoV S-2P Avi, HCoV-HKU1 S-2P and HCoV-OC43 S-2P were purified over a Superose6 Increase column (GE Life Sciences).

For the HIV-1 gp140 SOSIP variant from strain ZM197 (clade C) and CZA97 (clade C), recombinant, soluble antigens contained an AviTag and were expressed in Expi293F cells using polyethylenimine transfection reagent and cultured. FreeStyle F17 expression medium supplemented with pluronic acid and glutamine was used. The cells were cultured at 37°C with 8% CO_2 saturation and shaking. After 5-7 days, cultures were centrifuged and supernatant was filtered and run over an affinity column of agarose bound *Galanthus nivalis* lectin. The column was washed with PBS and antigens were eluted with 30 mL of 1M methyl-a-D-mannopyranoside. Protein elutions were buffer exchanged into PBS, concentrated, and run on a Superdex 200 Increase 10/300 GL Sizing column on the AKTA FPLC system. Fractions corresponding to correctly folded protein were collected, analyzed by SDS-PAGE and antigenicity was characterized by ELISA using known monoclonal antibodies specific to each antigen. Avi-tagged antigens were biotinylated using BirA biotin ligase (Avidity LLC). Non-Avi-tagged antigens were biotinylated using the EZ-Link Sulfo-NHS-Biotin kits using a 50:1 biotin to protein molar ratio.

For binding studies, SARS-CoV-2 HexaPro S, SARS-CoV S, SARS-CoV-2 RBD, SARS-CoV RBD, and MERS-CoV RBD constructs were expressed in the transient expression system previously mentioned. S proteins were purified using StrepTrap HP columns and RBD constructs were purified over protein A resin, respectively. Each resulting protein was further purified to homogeneity by size-exclusion chromatography on a Superose 6 10/300 GL column.

SARS-CoV-2 S1, SARS-CoV-2 S1 D614G, SARS-CoV-2 S2, and SARS-CoV-2 NTD truncated proteins were purchased from the commercial vendor, Sino Biological.

DNA-barcoding of Antigens

We used oligos that possess 15 bp antigen barcode, a sequence capable of annealing to the template switch oligo that is part of the 10X bead-delivered oligos, and contain truncated TruSeq small RNA read 1 sequences in the following structure: 5'-CCTTGGCACCCGAGAATTCCANNNNNNNNNNNNNCCCATATAAGA*A*A-3', where Ns represent the antigen barcode. We used the following antigen barcodes: GCTCCTTTACACGTA (SARS-CoV-2 S), TGACCTTCCTCTCCT (SARS-CoV S), ACAATTTGTCTGCGA (MERS-CoV S), TCCTTTCCTGATAGG (MERS-CoV S1), CAGGTCCCTTATTTC (HCoV-HKU1 S), TAACTCAGGGCCTAT (HCoV-OC43 S), CAGCCCACTGCAATA (CZA97), and ATCGTCGAGAGCTAG (ZM197). Oligos were ordered from IDT with a 5' amino modification and HPLC purified.

For each antigen, a unique DNA barcode was directly conjugated to the antigen itself. In particular, 5'amino-oligonucleotides were conjugated directly to each antigen using the Solulink Protein-Oligonucleotide Conjugation Kit (TriLink cat no. S-9011) according to manufacturer's instructions. Briefly, the oligo and protein were desalted, and then the amino-oligo was modified with the 4FB crosslinker, and the biotinylated antigen protein was modified with S-HyNic. Then, the 4FB-oligo and the HyNic-antigen were mixed together. This causes a stable bond to form between the protein and the oligonucleotide. The concentration of the antigen-oligo conjugates was determined by a BCA assay, and the HyNic molar substitution ratio of the antigen-oligo conjugates was analyzed using the NanoDrop according to the Solulink protocol guidelines. AKTA FPLC was used to remove excess oligonucleotide from the protein-oligo conjugates, which

were also verified using SDS-PAGE with a silver stain. Antigen-oligo conjugates were also used in flow cytometry titration experiments.

Antigen Specific B Cell Sorting

Cells were stained and mixed with DNA-barcoded antigens and other antibodies, and then sorted using fluorescence activated cell sorting (FACS). First, cells were counted and viability was assessed using Trypan Blue. Then, cells were washed three times with DPBS supplemented with 0.1% Bovine serum albumin (BSA). Cells were resuspended in DPBS-BSA and stained with cell markers including viability dye (Ghost Red 780), CD14-APC-Cy7, CD3-FITC, CD19-BV711, and IgG-PE-Cy5. Additionally, antigen-oligo conjugates were added to the stain. After staining in the dark for 30 minutes at room temperature, cells were washed three times with DPBS-BSA at 300 g for five minutes. Cells were then incubated for 15 minutes at room temperature with Streptavidin-PE to label cells with bound antigen. Cells were washed three times with DPBS-BSA, resuspended in DPBS, and sorted by FACS. Antigen positive cells were bulk sorted and delivered to the Vanderbilt Technologies for Advanced Genomics (VANTAGE) sequencing core at an appropriate target concentration for 10X Genomics library preparation and subsequent sequencing. FACS data were analyzed using FlowJo.

Sample and Library Preparation, and Sequencing

Single-cell suspensions were loaded onto the Chromium Controller microfluidics device (10X Genomics) and processed using the B-cell Single Cell V(D)J solution

according to manufacturer's suggestions for a target capture of 10,000 B cells per 1/8 10X cassette, with minor modifications in order to intercept, amplify and purify the antigen barcode libraries as previously described[27].

Sequence Processing and Bioinformatic Analysis

We utilized and modified our previously described pipeline to use paired-end FASTQ files of oligo libraries as input, process and annotate reads for cell barcode, UMI, and antigen barcode, and generate a cell barcode - antigen barcode UMI count matrix[27]. BCR contigs were processed using Cell Ranger (10X Genomics) using GRCh38 as reference. Antigen barcode libraries were also processed using Cell Ranger (10X Genomics). The overlapping cell barcodes between the two libraries were used as the basis of the subsequent analysis. We removed cell barcodes that had only non-functional heavy chain sequences as well as cells with multiple functional heavy chain sequences and/or multiple functional light chain sequences, reasoning that these may be multiplets. Additionally, we aligned the BCR contigs (filtered_contigs.fasta file output by Cell Ranger, 10X Genomics) to IMGT reference genes using HighV-Quest[179]. The output of HighV-Quest was parsed using ChangeO[139] and merged with an antigen barcode UMI count matrix. Finally, we determined the LIBRA-seq score for each antigen in the library for every cell as previously described[116].

Antibody Expression and Purification

For each antibody, variable genes were inserted into custom plasmids encoding the constant region for the IgG1 heavy chain as well as respective lambda and kappa

light chains (pTwist CMV BetaGlobin WPRE Neo vector, Twist Bioscience). Antibodies were expressed in Expi293F mammalian cells (ThermoFisher) by co-transfecting heavy chain and light chain expressing plasmids using polyethylenimine transfection reagent and cultured for 5-7 days. Cells were maintained in FreeStyle F17 expression medium supplemented at final concentrations of 0.1% Pluronic Acid F-68 and 20% 4mM L-Glutamine. These cells were cultured at 37°C with 8% CO_2 saturation and shaking. After transfection and 5-7 days of culture, cell cultures were centrifuged and supernatant was 0.45 µm filtered with Nalgene Rapid Flow Disposable Filter Units with PES membrane. Filtered supernatant was run over a column containing Protein A agarose resin equilibrated with PBS. The column was washed with PBS, and then antibodies were eluted with 100 mM Glycine HCl at 2.7 pH directly into a 1:10 volume of 1M Tris-HCl pH 8.0. Eluted antibodies were buffer exchanged into PBS 3 times using Amicon Ultra centrifugal filter units and concentrated. Antibodies were analyzed by SDS-PAGE. Additionally, antibodies 46472-1, 46472-2, 46472-3, 46472-4, 46472-6 and 46472-12 were assessed by size exclusion chromatography on a Superdex 200 Increase 10/300 GL Sizing column with the AKTA FPLC system.

High-throughput Antibody Expression

For high-throughput production of recombinant antibodies, approaches were used that are designated as microscale. For antibody expression, microscale transfection were performed (~1 ml per antibody) of CHO cell cultures using the Gibco ExpiCHO Expression System and a protocol for deep 96-well blocks (Thermo Fisher Scientific). In brief, synthesized antibody-encoding DNA (~2 µg per transfection) was

added to OptiPro serum free medium (OptiPro SFM), incubated with ExpiFectamine CHO Reagent and added to 800 µl of ExpiCHO cell cultures into 96-deep-well blocks using a ViaFlo 384 liquid handler (Integra Biosciences). The plates were incubated on an orbital shaker at 1,000 r.p.m. with an orbital diameter of 3 mm at 37 °C in 8% CO_2. The next day after transfection, ExpiFectamine CHO Enhancer and ExpiCHO Feed reagents (Thermo Fisher Scientific) were added to the cells, followed by 4 d incubation for a total of 5 d at 37 °C in 8% CO_2. Culture supernatants were collected after centrifuging the blocks at 450*g* for 5 min and were stored at 4°C until use. For high-throughput microscale antibody purification, fritted deep-well plates were used containing 25 µl of settled protein G resin (GE Healthcare Life Sciences) per well. Clarified culture supernatants were incubated with protein G resin for mAb capturing, washed with PBS using a 96-well plate manifold base (Qiagen) connected to the vacuum and eluted into 96-well PCR plates using 86 µl of 0.1 M glycine-HCL buffer pH 2.7. After neutralization with 14 µl of 1 M Tris-HCl pH 8.0, purified mAbs were buffer-exchanged into PBS using Zeba Spin Desalting Plates (Thermo Fisher Scientific) and stored at 4°C until use.

ELISA

To assess antibody binding, soluble protein was plated at 2 µg/ml overnight at 4°C. The next day, plates were washed three times with PBS supplemented with 0.05% Tween-20 (PBS-T) and coated with 5% milk powder in PBS-T. Plates were incubated for one hour at room temperature and then washed three times with PBS-T. Primary antibodies were diluted in 1% milk in PBS-T, starting at 10 µg/ml with a serial 1:5

dilution and then added to the plate. The plates were incubated at room temperature for one hour and then washed three times in PBS-T. The secondary antibody, goat anti-human IgG conjugated to peroxidase, was added at 1:10,000 dilution in 1% milk in PBS-T to the plates, which were incubated for one hour at room temperature. Goat anti-mouse secondary was used for SARS-CoV specific control antibody 240CD (BEI Resources). Plates were washed three times with PBS-T and then developed by adding TMB substrate to each well. The plates were incubated at room temperature for ten minutes, and then 1N sulfuric acid was added to stop the reaction. Plates were read at 450 nm.

Data are represented as mean ± SEM for one ELISA experiment. ELISAs were repeated 2 or more times. The area under the curve (AUC) was calculated using GraphPad Prism 8.0.0. For antibody 240CD, the following reagent was obtained through BEI Resources, NIAID, NIH: Monoclonal Anti-SARS-CoV S Protein (Similar to 240C), NR-616.

Competition ELISA

Competition ELISAs were performed as described above, with some modifications. After coating with antigen and blocking, 25 µl of non-biotinylated competitor antibody was added to each well at 10 µg/ml and incubated at 37°C for 10 minutes. Then, without washing, 75 µl biotinylated antibody (final concentration of 1 µg/ml) was added and incubated at 37°C for 1 hour. After washing three times with PBS-T, streptavidin-HRP was added at 1:10,000 dilution in 1% milk in PBS-T and incubated for 1 hour at room temperature. Plates were washed and substrate and

sulfuric acid were added as described above. ELISAs were repeated at least 2 times. Data is shown as the % decrease in binding.

Autoreactivity

Monoclonal antibody reactivity to nine autoantigens (SSA/Ro, SS-B/La, Sm, ribonucleoprotein (RNP), Scl 70, Jo-1, dsDNA, centromere B, and histone) was measured using the AtheNA Multi-Lyte® ANA-II Plus test kit (Zeus scientific, Inc, #A21101). Antibodies were incubated with AtheNA beads for 30min at concentrations of 50, 25, 12.5 and 6.25 µg/mL. Beads were washed, incubated with secondary and read on the Luminex platform as specified in the kit protocol. Data were analyzed using AtheNA software. Positive (+) specimens received a score >120, and negative (-) specimens received a score <100. Samples between 100-120 were considered indeterminate.

Mannose competition

Mannose competition ELISAs were performed as described above with minor modifications. After antigen coating and washing, nonspecific binding was blocked by incubation with 5% FBS diluted in PBS for 1 hour at RT. Primary antibodies were diluted in 5% FBS-PBST +/- 1M D-(+)-Mannose starting at 10 µg/ml with a serial 1:5 dilution and then added to the plate for 1 hour at RT. After washing, antibody binding was detected with goat anti-human IgG conjugated to peroxidase and added at 1:10,000 dilution in 5% FBS in PBS-T to the plates. After 1 hour incubation, plates were washed

and substrate and sulfuric acid were added as described above. Data shown is representative of three replicates.

Epitope Mapping Visualization

SARS-CoV-2 Spike (PDB-6VSB) was visualized using PyMOL software. Antibody epitopes were visualized on the SARS-CoV-2 spike using a structure of the pre-fusion stabilized SARS-CoV-2 S-2P construct[96] modeled in the molecular graphics software PyMOL (The PyMOL Molecular Graphics System, Version 2.3.5 Schrödinger, LLC).

RTCA Neutralization Assay

To assess for neutralizing activity against SARS-CoV-2 strain 2019 n-CoV/USA_WA1/2020 (obtained from the Centers for Disease Control and Prevention, a gift from N. Thornburg), we used the high-throughput RTCA assay and xCelligence RTCA HT Analyzer (ACEA Biosciences) that has been described previously[153]. After obtaining a background reading of a 384-well E-plate, 6,000 Vero-furin cells[140] were seeded per well. Sensograms were visualized using RTCA HT software version 1.0.1 (ACEA Biosciences). One day later, equal volumes of virus were added to antibody samples and incubated for 1 h at 37°C in 5% CO_2. mAbs were tested in triplicate with a single (1:20) dilution. Virus–mAb mixtures were then added to Vero-furin cells in 384-well E-plates. Controls were included that had Vero-furin cells with virus only (no mAb) and media only (no virus or mAb). E-plates were read every 8–12 h for 72 h to monitor virus neutralization. At 32 h after virus-mAb mixtures were added to the E-plates, cell

index values of antibody samples were compared to those of virus only and media only to determine presence of neutralization.

Nano-luciferase Neutralization Assay

A full-length SARS-CoV-2 virus based on the Seattle Washington isolate and a full-length SARS-CoV virus based on the Urbani isolate were designed to express luciferase and was recovered via reverse genetics and described previously[180, 181]. Viruses were titered in Vero E6 USAMRID cells to obtain a relative light units (RLU) signal of at least 10X the cell only control background. Vero E6 USAMRID cells were plated at 20,000 cells per well the day prior in clear bottom black walled 96-well plates (Corning 3904). Neutralizing antibody serum samples were tested at a starting dilution of 1:40 and were serially diluted 4-fold up to eight dilution spots. Antibody-virus complexes were incubated at 37C with 5% CO2 for 1 hour. Following incubation, growth media was removed and virus-antibody dilution complexes were added to the cells in duplicate. Virus-only controls and cell-only controls were included in each neutralization assay plate. Following infection, plates were incubated at 37C with 5% CO2 for 48 hours. After the 48 hour incubation, cells were lysed and luciferase activity was measured via Nano-Glo Luciferase Assay System (Promega) according to the manufacturer specifications. SARS-CoV and SARS-CoV-2 neutralization titers were defined as the sample dilution at which a 50% reduction in RLU was observed relative to the average of the virus control wells.

SPR

His-tagged SARS-CoV-2 RBD-SD1 was immobilized to a NiNTA sensorchip to a level of ~150 RUs using a Biacore X100. Serial dilutions of purified Fab 46472-12 were evaluated for binding, ranging in concentration from 1 to 0.25 μM. The resulting data were fit to a 1:1 binding model using Biacore Evaluation Software.

Fc Effector function Assays

Antibody-dependent Cellular Phagocytosis (ADCP)

Antibody-dependent cellular phagocytosis (ADCP) was performed using biotinylated SARS-CoV-2 or SARS-CoV S coated fluorescent neutravidin beads as previously described[182]. Briefly, beads were incubated for two hours with antibodies at a starting concentration of 50μg/ml and titrated five fold. CR3022 was used as a positive control while Palivizumab was used as a negative control. Antibodies and beads were incubated with THP-1 cells overnight, fixed and interrogated on the FACSAria II. Phagocytosis score was calculated as the percentage of THP-1 cells that engulfed fluorescent beads multiplied by the geometric mean fluorescence intensity of the population in the FITC channel less the no antibody control.

Antibody-dependent Cellular Trogocytosis (ADCT)

ADCT was performed as described in and modified from a previously described study[148]. HEK293T cells transfected with a SARS-CoV-2 spike pcDNA vector were surface biotinylated with EZ-Link Sulfo-NHS-LC-Biotin as recommended by the manufacturer. Fifty-thousand cells per well were incubated with antibody for 30 minutes

starting at 25µg/ml and titrated 5 fold. CR3022 was used as a positive control with Palivizumab as a negative. Following a RPMI media wash, these were then incubated with carboxyfluorescein succinimidyl ester (CFSE) stained THP-1 cells (5 $X10^4$ cells per well) for 1 hour and washed with 15mM EDTA/PBS followed by PBS. Cells were then stained for biotin using Streptavidin-PE and read on a FACSAria II. Trogocytosis score was determined as the proportion of CFSE positive THP-1 cells also positive for streptavidin-PE less the no antibody control.

Antibody-dependent Complement Deposition (ADCD)

Antibody-dependent complement deposition was performed as previously described[183]. Briefly biotinylated SARS-Cov-2 S protein was coated 1:1 onto fluorescent neutravidin beads for 2 hours at 37 degrees. These beads were incubated with 100ug/ml of antibody for 1 hour and incubated with guinea pig complement diluted 1 in 50 with gelatin/veronal buffer for 15 minutes at 37 degrees. Beads were washed at 2000g twice in PBS and stained with anti-guinea pig C3b-FITC, fixed and interrogated on a FACSAria II. Complement deposition score was calculated as the percentage of C3b-FITC positive beads multiplied by the geometric mean fluorescent intensity of FITC in this population less the no antibody or heat inactivated controls.

Antibody Prophylaxis - Murine Model of Infection

For evaluating the prophylactic efficacy of mAbs, 12-month old female BALB/c mice (BALB/cAnHsd; Envigo, stock number 047) were treated with 200 µg mAb intraperitoneally (i.p.) 12 hours prior to virus inoculation. The next day, mice were

administered intranasally with $1x10^3$ PFU or $1x10^4$PFU of SARS-CoV-2 MA10, respectively. Mice were monitored daily for weight loss, morbidity, and mortality, and after four days, mice were sacrificed and lung tissue was harvested for viral titer as measured by plaque assays. One lung lobe was taken for pathological analysis and the other lobe was processed for qPCR and viral load determination as previously described[171]. For viral plaque assays, the caudal lobe of the right lung was homogenized in PBS, and the tissue homogenate was then serial-diluted onto confluent monolayers of Vero E6 cells, followed by agarose overlay. Plaques were visualized with overlay of Neutral Red dye on day 2 post infection. Gross pulmonary hemorrhage was observed at time of tissue harvest and scored on a scale of 0 (no hemorrhage in any lobe, normal pink healthy lung) to 4 (complete hemorrhage in all lobes of the lung, completely dark red lung). For viral titer and hemorrhage score comparisons, an ordinary one-way ANOVA test with multiple comparisons was performed using Prism software, GraphPad Prism version 8.0.

ACE2 Binding Inhibition Assay

Wells of 384-well microtiter plates were coated with purified recombinant SARS-CoV-2 S-2P ectoprotein at 4°C overnight. Plates were blocked with 2% non-fat dry milk and 2% normal goat serum in DPBS-T for 1 hr. Purified mAbs were diluted two-fold in blocking buffer starting from 10 µg/mL in triplicate, added to the wells (20 µL/well), and incubated at ambient temperature. Recombinant human ACE2 with a C-terminal FLAG tag protein was added to wells at 2 µg/mL in a 5 µL/well volume (final 0.4 µg/mL concentration of ACE2) without washing of antibody and then incubated for 40 min at

ambient temperature. Plates were washed, and bound ACE2 was detected using HRP-conjugated anti-FLAG antibody and TMB substrate. ACE2 binding without antibody served as a control. Experiment was done in biological replicate and technical triplicates, shown is representative of one replicate with positive control mAb COV2-2196[153].

Identification of Residue-level Mutants

Potential cross-reactive epitopes were identified based on sequence and structural homology. Reference sequences for each Coronavirus S were obtained either from NCBI for SARS-CoV-2 (YP_009724390.1) and MERS-CoV (YP_009047204.1) or from Uniprot for SARS-CoV (P59594) of the spikes was then obtained using MUSCLE [184] and the amino acid similarity to SARS-CoV-2 at each residue position was calculated using the BLOSUM-62 scoring matrix[185]. These scores were then used to color each residue position on the SARS-CoV-2 S structure (PDB ID: 6VSB) in PyMOL (Schrodinger, version 2.3.5) in order to visualize surface patches and linear epitopes with structural homology. These conserved regions were then visualized on the other human coronavirus spike structures by retrieving them from the Protein Databank (SARS-CoV: 5X5B, MERS-CoV: 5W9I) and aligning them to the SARS-CoV-2 S structure. Finally, the residue N165 was part of a conserved surface patches and was mutated to alanine and tested for binding with antibodies. The N709A mutant tested was previously described in Acharya et al., BioRxiv (2020).

Quantification and Statistical Analysis

ELISA error bars (standard error of the mean) were calculated using GraphPad Prism version 8.0.0. ANOVA analysis (ordinary one way ANOVA with multiple comparisons) was performed on viral load titers and hemorrhage scores from animal experiments using GraphPad Prism version 8.0.0. Details of the statistical analyses can be found in the main text and figure captions.

3.2 Characterization of a Broadly Neutralizing Influenza Antibody with a Unique Mode of Hemagglutinin Recognition

Contributions:

Ian Setliff helped with identification of 3602-1707. Rob Parks performed autoreactivity assays. Iuliia Gilchuk performed neutralization tests. Elad Binshtein performed negative stain electron microscopy and cryogenic electron microscopy. Giuseppe Sautto performed *in vivo* antibody experiments. I expressed and purified antibodies, performed quality control tests, performed ELISA binding experiments, generated Fab, and expressed and purified antigen proteins including H1 A/New Caledonia/20/99.

3.2.1 Introduction

Broadly reactive influenza antibodies are of interest for use as therapeutic molecules and for rational vaccine design strategies applied to universal flu vaccine development[70, 71, 78, 79]. A select number of neutralizing antibodies targeting the hemagglutinin protein are broadly neutralizing, able to neutralize most or all group I viruses, group II viruses, and even group I/group II viruses. Broadly neutralizing influenza antibodies target a variety of epitopes on the hemagglutinin spike protein, including the stem region, which is less variable than the head domain. As such, stem is a target for universal flu vaccine strategies. Even though the hemagglutinin stem displays more conservation than the head domain, variation in particular amino acid and glycan residues contributes to antibody recognition and neutralization of hemagglutinin strains[186]. Some examples of broadly neutralizing antibodies that target the stem region include MEDI8852 and CR8020, which interact with stem using different angles of approach[186-188]. Overall, identification of additional broadly neutralizing antibodies is of interest, and understanding the molecular mechanism of neutralization for each group of viruses is integral to enhancing vaccine development.

3.2.2 Results

To identify broadly reactive influenza antibodies, I examined a dataset from an experiment described in Setliff et al., 2019[116], where I applied LIBRA-seq to an HIV-positive sample from donor N90. In this experiment, I used a nine-antigen screening

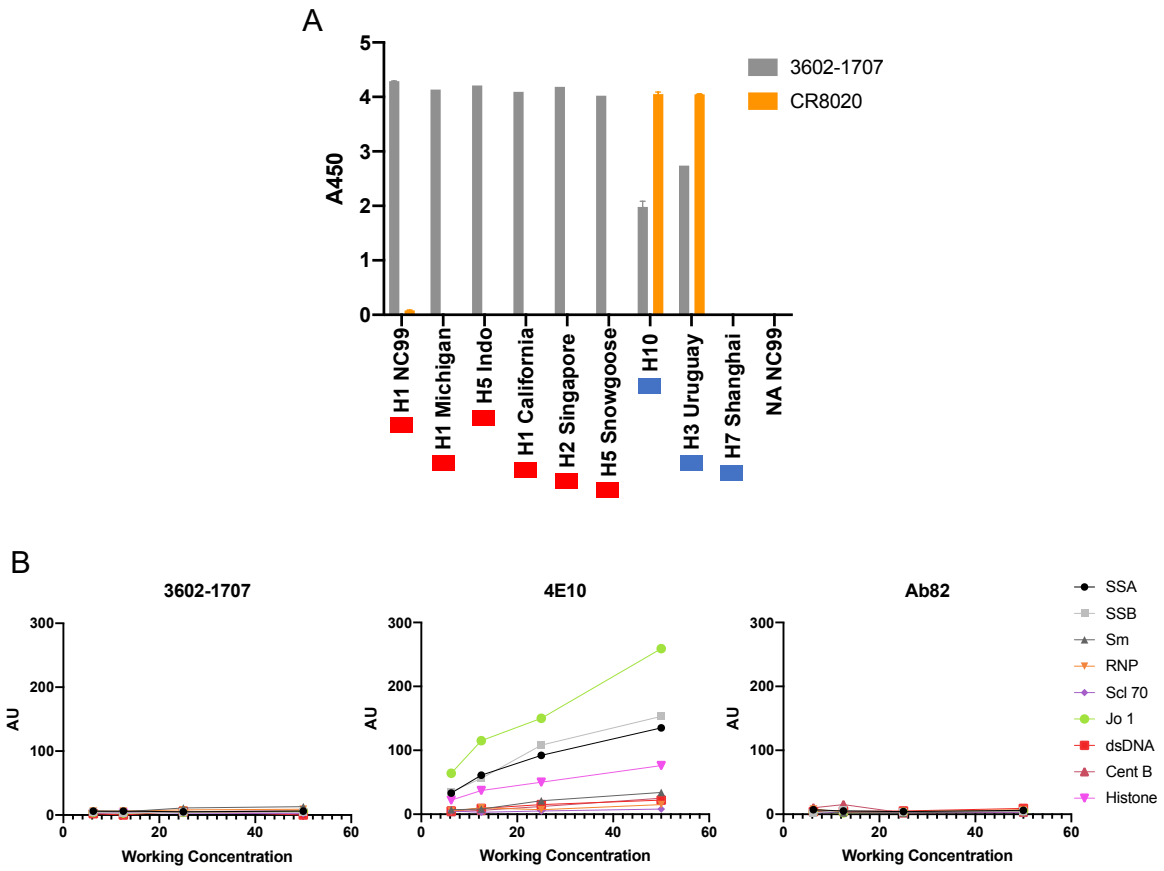


Figure 3.2.1. Characterization of influenza antibody 3602-1707. (A) ELISA binding of 3602-1707 and influenza group 2 control CR8020 against a variety of hemagglutinin antigens (Group 1 shown in red, group 2 shown in blue). All antibodies were tested at a single concentration (10 μg/ml). (B) Antibodies were tested for autoreactivity against a variety of antigens in the Luminex AtheNA assay. AU stands for Athena Units. Anti-HIV antibody 4E10 was used as a positive control and Ab82 was used as a negative control.

library composed of five HIV-1 SOSIP trimers from a variety of clades (BG505, B41, ZM106.9, ZM197 and KNH1144) and four influenza hemagglutinin trimers (H1 A/New Caledonia/20/99, H1 A/Michigan/45/2015, H5 A/Indonesia/5/2005, and H7 A/Anhui/1/2013). After applying the LIBRA-seq workflow, we recovered B cell receptor sequences for functional heavy- and light-chain sequences with antigen mapping information from 1465 cells, and these B cells exhibited a range of LIBRA-seq scores among the nine antigens. We identified one lead candidate called 3602-1707 which was broadly reactive against influenza hemagglutinin (HA) variants (**Figure 2.11A-B**). 3602-1707 utilizes the *IGHV3-23* gene, which is uncommon for broadly reactive influenza

		IC50 µg/ml	
Virus	HA	3602-1707	MEDI8852
gWA14 H5N8	H5	>20	>20
VN04 H5N1	H5	0.36	0.81
PR34 H1N1	H1	0.13	1.14
USSR77 H1N1	H1	0.16	0.9
TX91 H1N1	H1	0.02	0.55
NC99 H1N1	H1	1.08	*ND*
SI06 H1N1	H1	0.77	1.36
CA09 H1N1	H1	0.97	5.2
VI75 H3N2	H3	>20	5.2

Group 1
Group 2

Potency

high low

Figure 3.2.2. Functional characterization of 3602-1707. Influenza micro-neutralization using real time cell analysis (RTCA) assay for 3602-1707 and MEDI8852 control is shown a variety of viral strains. The IC_{50} values are shown.

antibodies (**Figure 2.11A-B**). A select number of antibodies have been identified that utilize this gene, and they have not been extensively characterized[186].

Next, I recombinantly expressed this antibody and confirmed that 3602-1707 binds to a variety of influenza hemagglutinin molecules as predicted by LIBRA-seq (**Figure 2.11A-B**). To further examine the breadth of this antibody, I tested 3602-1707 for binding to additional hemagglutinin antigens from diverse strains, including hemagglutinin proteins from group I and group II. I determined that 3602-1707 is broadly reactive against group I hemagglutinin antigens and also shows reactivity to some group II hemagglutinin proteins (**Figure 3.2.1A**). To assess potential autoreactive

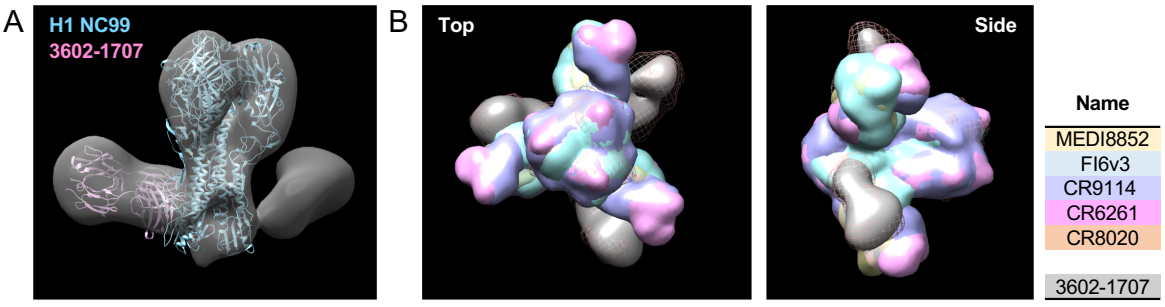


Figure 3.2.3. Structural characterization of 3602-1707. (A) Negative stain EM structure of Fab-trimer complex for 3602-1707 Fab (pink) and H1 NC99 trimer (blue). (B) Fab-trimer complex structure comparison with other antibody/trimer structures (MEDI8852 shown in yellow, FI6v3 shown in blue, CR9114 shown in purple, CR6261 shown in pink, and CR8020 shown in orange). 3602-1707 is depicted in gray. Figure made by Elad Binshtein.

properties of this antibody, we collaborated with the Haynes group at Duke, where Rob Parks tested 3602-1707 for autoreactivity and did not measure autoreactivity against the tested antigens (**Figure 3.2.1B**). To determine functionality of the antibody, I next worked with Iuliia Gilchuk in the Crowe laboratory to assess the neutralization potential of 3602-1707. We found that 3602-1707 is capable of protecting cells from cytopathic effect of influenza H1N1 and H5N1 viruses (**Figure 3.2.2**).

Because of its breadth of reactivity and neutralization properties, we next sought to determine the epitope of 3602-1707. I worked with Elad Binshtein in the Crowe laboratory to evaluate the binding of 3602-1707 Fab to H1 A/New Caledonia/20/99 by negative stain electron microscopy. These experiments demonstrated that 3602-1707 Fab binds to the hemagglutinin stem region (**Figure 3.2.3A**). Interestingly, this antibody targets the stem with a similar manner of approach as antibody CR8020, a group 2 reactive, broadly neutralizing antibody[84, 86, 87, 187, 188] (**Figure 3.2.3B**). We are currently

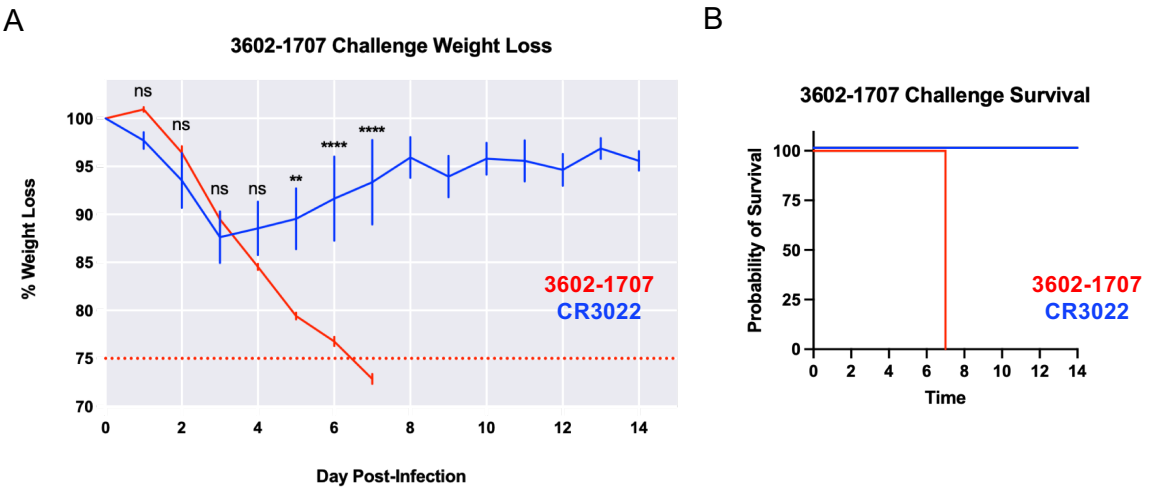


Figure 3.2.4. *In vivo* effects of antibody 3602-1707. (A) Each group of DBA/2J mice (n=4) was infected with challenge strain H1N1 A/New Caledonia/20/1999 with $5x10^5$ PFU/mouse intranasally. 12 hours before challenge, 3602-1707 or control IgG CR3022 were administered prophylactically via the intraperitoneal route. Daily body weight progression is shown for each group. (B) Survival for each antibody group is shown. Figure made by Giuseppe Sautto.

examining atomic level details of the Fab-hemagglutinin interaction using cryogenic electron microscopy.

Additionally, I collaborated with Giuseppe Sautto at the University of Georgia for preliminary animal experiments. In these studies, Giuseppe demonstrated that prophylactic administration of 3602-1707 was able to protect mice from lethal challenge with influenza strain A/New Caledonia/20/1999 compared to control antibody CR3022[169] (**Figure 3.2.4A-B**). Mice administered with the control antibody all succumbed, whereas all mice administered with antibody 3602-1707 survived and recovered (**Figure 3.2.4A-B**). This highlights the functional relevance of the broadly neutralizing antibody 3602-1707.

3.2.3 Discussion

Antibody 3602-1707 is unique due to is genetic features, reactivity pattern, and binding recognition modality. Gaining a more detailed picture of the interactions between the Fab and hemagglutinin stem will help increase understanding into how this particular antibody accommodates group 1 recognition while binding similarly to a group 2 reactive antibody. Future directions include engineering the antibody to accommodate group 1/group 2 cross-reactivity to increase its neutralization breadth. Studying broadly neutralizing influenza antibodies with unique features will increase our understanding of the immune response to this ever-changing pathogen. Adding knowledge to our arsenal will aid in preparing for development of novel influenza strains due to antigenic drift and

antigenic shift. Furthermore, elucidating the epitopes of broadly reactive antibodies can help inform universal flu design strategies.

3.2.4 Materials and Methods

ELISA

To assess antibody binding, soluble protein was plated at 2 µg/ml overnight at 4°C. The next day, plates were washed three times with PBS supplemented with 0.05% Tween-20 (PBS-T) and coated with 5% milk powder in PBS-T. Plates were incubated for one hour at room temperature and then washed three times with PBS-T. Primary antibodies were diluted in 1% milk in PBS-T, starting at 10 µg/ml with a serial 1:5 dilution and then added to the plate. The plates were incubated at room temperature for one hour and then washed three times in PBS-T. The secondary antibody, goat anti-human IgG conjugated to peroxidase, was added at 1:10,000 dilution in 1% milk in PBS-T to the plates, which were incubated for one hour at room temperature. Plates were washed three times with PBS-T and then developed by adding TMB substrate to each well. The plates were incubated at room temperature for ten minutes, and then 1N sulfuric acid was added to stop the reaction. Plates were read at 450 nm.

Autoreactivity

Monoclonal antibody reactivity to nine autoantigens (SSA/Ro, SS-B/La, Sm, ribonucleoprotein (RNP), Scl 70, Jo-1, dsDNA, centromere B, and histone) was measured using the AtheNA Multi-Lyte® ANA-II Plus test kit (Zeus scientific, Inc,

#A21101). Antibodies were incubated with AtheNA beads for 30min at concentrations of 50, 25, 12.5 and 6.25 µg/mL. Beads were washed, incubated with secondary and read on the Luminex platform as specified in the kit protocol. Data were analyzed using AtheNA software. Positive (+) specimens received a score >120, and negative (-) specimens received a score <100. Samples between 100-120 were considered indeterminate.

RTCA Neutralization Assay

To assess for neutralizing activity against influenza strains, we used a high-throughput neutralization assay in which cytopathic effect is monitored on the xCELLigence Real-Time Cell Analysis (RTCA) instrument. Madin-Darby canine kidney (MDCK) cells are first seeded. Antibodies are serially diluted starting at 20 µg/mL and then mixed with influenza virus and incubated at room temperature for 1 hour. Then, the virus antibody mixture is added to the MDCK cell monolayer, incubated, and monitored for cytopathic effect.

Fab Preparation

To generate Fabs, IgGs were incubated with Lys-C at 1:4,000 (weight:weight) overnight at 37 °C. EDTA free protease inhibitor (Roche) was dissolved to 25X and then added to the sample at a final 1X concentration. The sample was passed over a Protein A column. The flow-through was collected run on a Superdex 200 Increase 10/300 GL Sizing column on the AKTA FPLC system. Fabs were visualized on SDS-PAGE.

Electron Microscopy

Purified H1 NC99 (A/New Caledonia/20/1999) and 3602-1707 Fab were combined and incubated on ice for 30 minutes at a 1:3 molar ratio of H1 NC99 trimer: Fab prior to grid deposition. Using 2D class images of 3602-1707 complexed with hemagglutinin, the data was superimposed into an hemagglutinin structure and other antibody/hemagglutinin complexes (PDB: 1ruz, 5jw4, 3ztj, 4fqi, 3gbm, 3sdy).

Murine Model

6-8 week old female DBA/2J mice (Jackson Laboratory) were used in a murine infection model for influenza with strain A/New Caledonia/20/1999 x PR/8 H1N1 virus at $5x10^5$ PFU per mouse. Two groups (n=4 per group) were administered antibody, either CR3022 IgG1 (control) or 3602-1707. 5 mg/kg per mouse per antibody was administered prophylactically via the intraperitoneal route 12 hours before infection. Mice were monitored for weight loss and clinical signs for 14 days. All mouse studies were performed at the University of Georgia.

3.3 Identification of an HIV/influenza Cross-reactive Antibody Family

Contributions:

Ian Setliff helped with identification of the cross-reactive antibody family. Rob Parks performed autoreactivity experiments. Iuliia Gilchuk performed neutralization tests. Simone Richardson performed neutralization tests and Fc effector function experiments. I expressed and purified the antibodies, and performed ELISA binding assays.

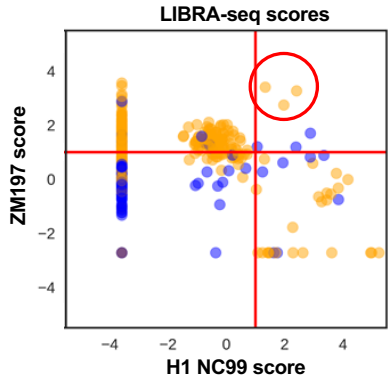
3.3.1 Introduction

Interrogating antibody reactivity is integral to understanding development of antibody specificity to infectious pathogens. An important feature of B cell development is the ability to distinguish self from non-self in order to progress to maturity[1]. Additionally, upon antigen exposure, mutations arise which can help increase B cell receptor affinity for the particular antigen in a process called somatic hypermutation[7]. Along with VDJ recombination, these factors help account for the incredible diversity that the immune system achieves to counteract the wide variety of pathogens encountered by an individual during their lifetime[2, 4]. As such, antibodies are traditionally thought to have a unique, single specificity; however, some antibodies can have polyreactive or autoreactive phenotypes, including some broadly neutralizing HIV-1 antibodies[189-191]. Furthermore, it has been demonstrated that some antibodies that target the gp41 region of the HIV-1 envelope glycoprotein trimer show cross-reactivity with gut microbiota[192, 193]. Additionally, a select number of antibodies have been identified that show binding to multiple epitopes on antigenic pathogens[194, 195]. For example, recent studies demonstrated that 2G12, a glycan targeting, anti-HIV antibody, is able to neutralize HIV-1 and influenza and can also recognize SARS-CoV-2 spike[196-200]. Investigating additional cross-reactive antibodies can shed light on their role in the immune response and highlight potential flexibility, constraints, and considerations for antibody development in the context of pathogens that put sustained immune pressure on the host.

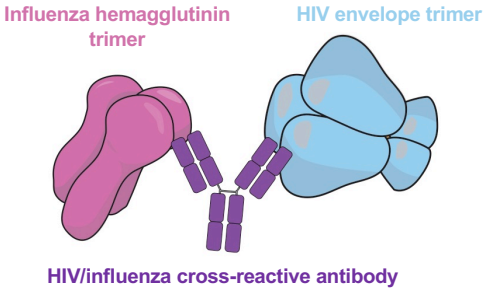
3.3.2 Results

From a LIBRA-seq experiment in Setliff et al., 2019[116], nine antigens were applied to an HIV-positive sample from donor N90. These included HIV-1 antigens (BG505, B41, ZM106.9, ZM197 and KNH1144) and influenza antigens (H1 A/New Caledonia/20/99, H1 A/Michigan/45/2015, H5 A/Indonesia/5/2005, and H7 A/Anhui/1/2013). In addition to identifying HIV-specific and influenza-specific antibodies, we also observed 9 IgG antibodies (out of 591 total IgG) that showed high LIBRA-seq scores for both HIV-1 and influenza antigens. Interestingly, 3 of these antibodies showed reactivity to ZM197 (HIV-1) and H1 NC99 (influenza) and were clonally related, possessing the same *IGHV* gene (*IGHV5-51*), CDRH3 length (14), *IGLV* gene (*IGKV2-28*), and CDRL3 length (9) (**Figure 3.3.1A-C**). This suggested that the HIV/influenza

A



B



C

Name	VH Gene	VH % Identity	CDRH3 Length	CDRH3 Sequence	VL Gene	VL % Identity	CDRL3 Length	CDRL3 Sequence
3602-546	IGHV5-51	92.36	14	ARPPFSGSYEGFDY	IGKV2-28	94.56	9	MQSLQIPST
3602-2526	IGHV5-51	93.06	14	ARPPFSGSYEGLDY	IGKV2-28	95.24	9	MQAIHIPST

Figure 3.3.1. Identification of HIV/influenza cross-reactive antibody family. (A) LIBRA-seq scores for HIV antigen ZM197 and influenza antigen H1 NC99 for analysis of B cells from donor N90, from **Figure 2.10**. Each dot represents a cell that had a LIBRA-seq score of above 1 for at least 1 antigen in the full screening library. IgG (orange) and other isotypes (blue) are shown. Antibodies of interest are circled in red. (B) Cross-reactive antibody schematic. (C) Genetic information for two antibodies from the cross-reactive family that were expressed recombinantly.

cross-reactivity we observed via LIBRA-seq could be true signal rather than an artifact of the assay (**Figure 3.3.1B**).

First, I expressed and purified 2 antibodies from this family, named 3602-2526 and 3602-546, and tested them for binding to HIV-1 and influenza trimer antigens by ELISA (**Figure 3.3.2A-B**). 3602-2526 and 3602-546 showed strong binding to three group I hemagglutinin trimers: H1 NC99, H1 Michigan, and H5 Indonesia (**Figure 3.3.2A-B**). Additionally, 3602-2526 and 3602-546 bound to HIV-1 trimer ZM197, although this binding was lower than for the influenza antigens (**Figure 3.3.2A-B**). 3602-2526 and 3602-546 showed similar patterns of binding, though 3602-546 typically demonstrated less binding to the antigens overall. Of note, at the amino acid level, these antibodies have 1 difference in their CDRH3 sequences and 10 differences in their heavy-chain V gene sequences overall (**Figure 3.3.1C**).

After observing cross-reactivity to the HIV-1 and influenza antigens, we next wanted to assess polyreactivity and autoreactivity of this antibody family. To assess polyreactivity, I tested binding of the antibodies to nuclear antigens and cardiolipin.

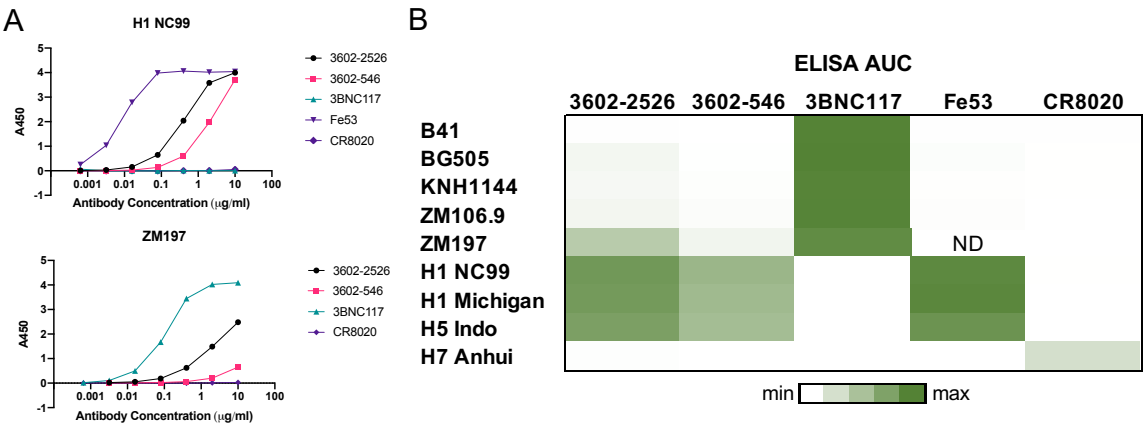


Figure 3.3.2. Characterization of cross-reactive antibody family. (A) ELISA binding data for cross-reactive antibodies, HIV control 3BNC117, and influenza controls Fe53 and CR8020. (B) ELISA area under the curve (AUC) values for binding to HIV SOSIP trimer antigens and influenza hemagglutinin trimer proteins are shown as a heatmap from minimum (white) to maximum (green) binding for cross-reactive antibodies, HIV control 3BNC117, and influenza controls Fe53 and CR8020.

3602-546 and 3602-2526 did not show binding to these antigens by ELISA, indicating these antibodies likely were not polyreactive (**Figure 3.3.3A**). We also collaborated with Rob Parks in the Haynes group at Duke to test the antibodies for autoreactivity, and we did not observe reactivity to any of the tested antigens (**Figure 3.3.3B**). Together, this suggests the observed HIV/influenza cross-reactivity is specific to these particular viruses rather than a polyreactive response. After confirming HIV/influenza cross-reactivity, I sought to determine the epitope of these antibodies. By screening these antibodies against a variety HIV-1 antigens and influenza antigens, I determined that 3602-2526 binds to HIV-1 V1V2 scaffolds (**Figure 3.3.4**).

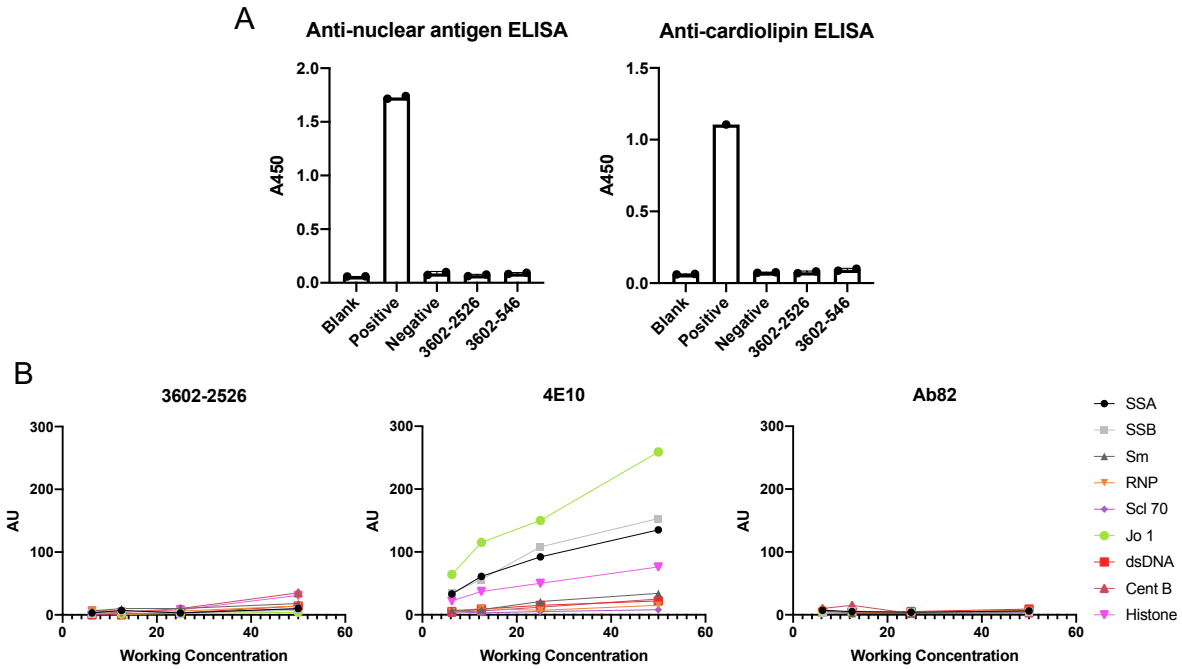


Figure 3.3.3. Assessment of polyreactivity and autoreactivity for cross-reactive antibody family. (A) Binding of antibodies to nuclear antigens and cardiolipin by ELISA. (B) Antibodies were tested for autoreactivity against a variety of antigens in the Luminex AtheNA assay. AU stands for Athena Units. Anti-HIV antibody 4E10 was used as a positive control and Ab82 was used as a negative control.

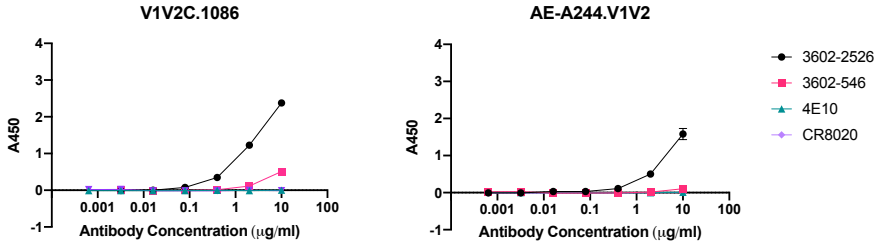


Figure 3.3.4. Epitope mapping. ELISA binding to HIV V1V2 scaffolds for cross-reactive antibodies.

Next, we collaborated with the Crowe lab and the Morris lab to gain insight into the functionality of these antibodies. Though they were non-neutralizing for influenza and HIV-1 (**Figure 3.3.5A-B**), 3602-2526 was able to mediate Fc effector functionality, specifically antibody dependent cellular phagocytosis (**Figure 3.3.5C**). We observed these results for both HIV-1 antigens and influenza antigens, suggesting a cross-reactive functional role for these antibodies.

3.3.3 Discussion

Traditionally, antibodies are thought to have a single, unique antigen specificity, so the identification of HIV/influenza cross-reactive antibodies could have a significant impact on the antibody field. Learning more about the character of these antibodies, including their development, frequency, and the antigenic epitopes they target, will help us learn more about the nature of antibody specificity and antibody-antigen interactions. This will help us to query the determinants of antibody-antigen specificity. Next, we can move toward trying to understand what factors lead to a single, specific response vs. a cross-reactive response vs. a polyreactive, promiscuous response. Learning more

about the nature of these antibodies can help us to investigate if there is interplay between pathogen specific repertoires. Understanding the mechanisms of this cross-reactivity will allow us to have an increased understanding of the antibody response to viral pathogens.

A

Strain	IC50 µg/ml 3602-2526	3602-546
MW965	*ND*	>25
BG505 WT	*ND*	>25
ZM197	*ND*	>25
B41	>25	>25
ZM106.9	>25	>25
KNH1144	>25	>25
X1632	>25	>25
246_F3	>25	>25
HIV-1 25710	>25	>25
X2278	>25	>25
Ce1176	>25	>25
Ce0217	>25	>25
TRO.11	*ND*	>25
CH119	*ND*	>25
CNE55	*ND*	>25
BJOX	*ND*	>25
398 F1	*ND*	>25

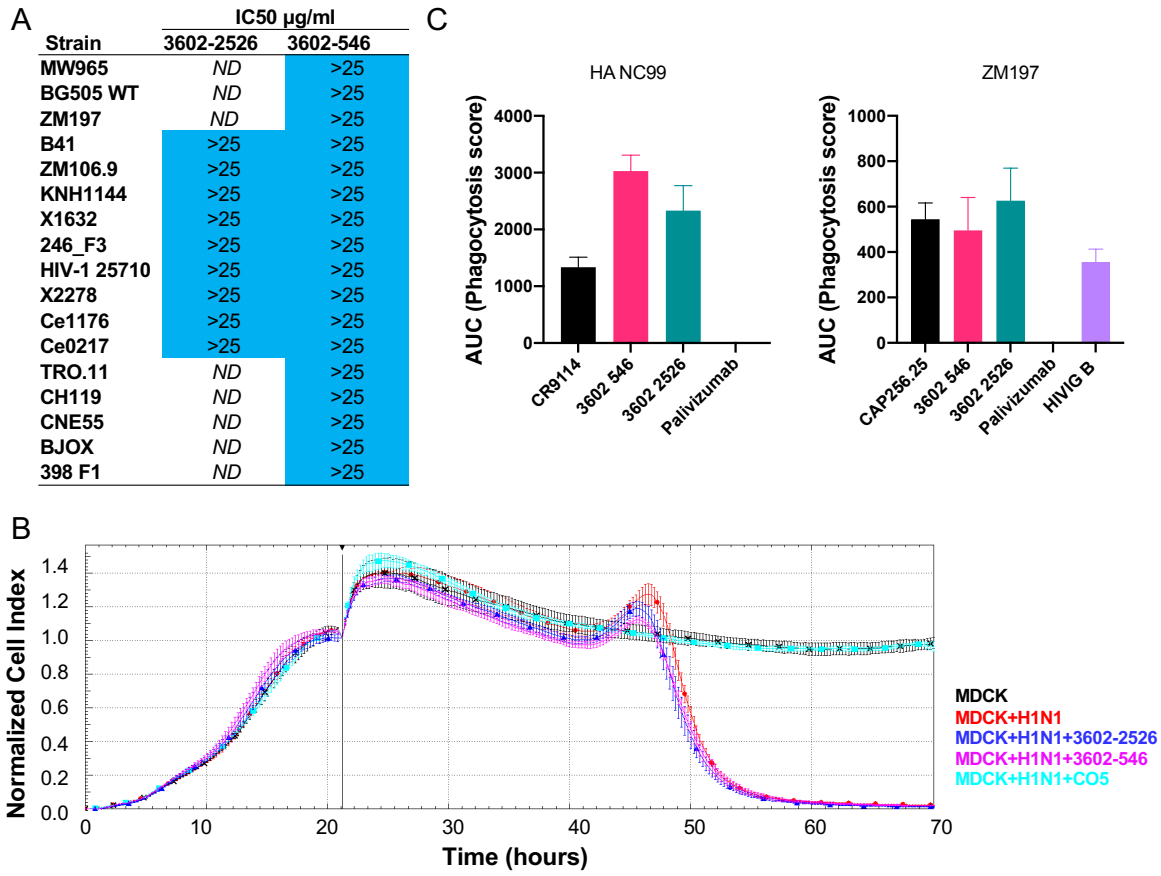


Figure 3.3.5. Functional characterization of cross-reactive antibody family. (A) Neutralization of HIV pseudoviruses by antibodies 3602-2526 and 3602-546. ND = Not done. (B) Influenza micro-neutralization using real time cell analysis (RTCA) assay for 3602-2526, 3602-546 and C05 control is shown for A/New Caledonia/20/1999 H1N1. mAbs 3602-546 and 3602-2526 do not protect MDCK cells from cytopathic effect of A/New Caledonia/20/1999 H1N1. (C) Cross-reactive antibodies were tested for antibody-dependent cellular phagocytosis activity (ADCP) against HIV ZM197 and influenza H1 A/New Caledonia/20/1999, compared to HIV positive control CAP256.25 and HIVIG B or influenza positive control CR9114 and negative control Palivizumab, an anti-RSV antibody. AUC of the phagocytosis score is shown. Data are represented as mean ± SD. Figure 3.3.5B made by Iuliia Gilchuk. Figure 3.3.5C made by Simone Richardson.

3.3.4 Materials and Methods

ELISA

To assess antibody binding, soluble protein was plated at 2 µg/ml overnight at 4°C. The next day, plates were washed three times with PBS supplemented with 0.05% Tween-20 (PBS-T) and coated with 5% milk powder in PBS-T. Plates were incubated for one hour at room temperature and then washed three times with PBS-T. Primary antibodies were diluted in 1% milk in PBS-T, starting at 10 µg/ml with a serial 1:5 dilution and then added to the plate. The plates were incubated at room temperature for one hour and then washed three times in PBS-T. The secondary antibody, goat anti-human IgG conjugated to peroxidase, was added at 1:10,000 dilution in 1% milk in PBS-T to the plates, which were incubated for one hour at room temperature. Plates were washed three times with PBS-T and then developed by adding TMB substrate to each well. The plates were incubated at room temperature for ten minutes, and then 1N sulfuric acid was added to stop the reaction. Plates were read at 450 nm.

Autoreactivity

Monoclonal antibody reactivity to nine autoantigens (SSA/Ro, SS-B/La, Sm, ribonucleoprotein (RNP), Scl 70, Jo-1, dsDNA, centromere B, and histone) was measured using the AtheNA Multi-Lyte® ANA-II Plus test kit (Zeus scientific, Inc, #A21101). Antibodies were incubated with AtheNA beads for 30min at concentrations of 50, 25, 12.5 and 6.25 µg/mL. Beads were washed, incubated with secondary and read

on the Luminex platform as specified in the kit protocol. Data were analyzed using AtheNA software. Positive (+) specimens received a score >120, and negative (-) specimens received a score <100. Samples between 100-120 were considered indeterminate.

RTCA Neutralization Assay

To assess for neutralizing activity against influenza strains, we used a high-throughput neutralization assay in which cytopathic effect is monitored on the xCELLigence Real-Time Cell Analysis (RTCA) instrument (ACEA Biosciences). MDCK cells are seeded in E-Plate wells. Antibodies are serially diluted starting at 20 µg/mL and then mixed with influenza virus and incubated at room temperature for 1 hour. Then, the virus antibody mixture is added to the MDCK cell monolayer, incubated, and monitored for cytopathic effect.

TZM-bl Neutralization Assays

Antibody neutralization was assessed using the TZM-bl assay as described (Sarzotti-Kelsoe et al., 2014). This standardized assay measures antibody-mediated inhibition of infection of JC53BL-13 cells (also known as TZM-bl cells) by molecularly cloned Env-pseudoviruses. Viruses that are highly sensitive to neutralization (Tier 1) and/or those representing circulating strains that are moderately sensitive (Tier 2) were included, plus additional viruses, including a subset of the antigens used for LIBRA-seq. Murine leukemia virus (MLV) was included as an HIV-specificity control and VRC01 was

used as a positive control. Results are presented as the concentration of monoclonal antibody (in μg/ml) required to inhibit 50% of virus infection (IC_{50}).

Fc Effector Function Assay

Antibody-dependent Cellular Phagocytosis (ADCP)

Antibody-dependent cellular phagocytosis (ADCP) was performed using biotinylated H1 NC99 or ZM197 trimer coated fluorescent neutravidin beads as previously described[146]. Briefly, beads were incubated for two hours with antibodies at a starting concentration of 50μg/ml and titrated five-fold. Palivizumab was used as a negative control. Antibodies and beads were incubated with THP-1 cells overnight, fixed and interrogated on the FACSAria II. Phagocytosis score was calculated as the percentage of THP-1 cells that engulfed fluorescent beads multiplied by the geometric mean fluorescence intensity of the population in the FITC channel less the no antibody control.

CHAPTER 4

Next-generation LIBRA-seq for Antibody Discovery: Efficient Discovery of Potent Neutralizing SARS-CoV-2 Antibodies using LIBRA-seq with Ligand Blocking

This chapter is adapted from the manuscript:

Shiakolas A, Kramer K, Johnson N, Wall S, Suryadevara N, Wrapp D, Periasamy S, Pilewski K, Raju N, Nargi R, Sutton R, Walker L, Setliff I, Crowe Jr J, Bukreyev A, Carnahan R, McLellan J, Georgiev I. (2022). Efficient discovery of potent neutralizing SARS-CoV-2 antibodies using LIBRA-seq with ligand blocking, *Nature Biotechnology*, doi: 10.1038/s41587-022-01232-2.

Contributions:

Kevin Kramer expressed and purified antigen proteins. Nicole Johnson performed structural characterization of Fabs in complex with spike. Steven Wall expressed, purified, and characterized antibodies from the cross-reactive LIBRA-seq experiment with ligand blocking using ELISA binding experiments. Naveenchandra Suryadevara performed neutralization tests. Daniel Wrapp performed affinity experiments. Sivakumar Pariasamy performed neutralization tests. Kelsey Pilewski helped with experimental design and performed ELISA binding assays. Nagarajan Raju performed bioinformatic

analysis. Rachel Nargi and Rachel Sutton helped with high throughput antibody expression. Lauren Walker helped with antigen quality control. Ian Setliff helped with experimental design. James Crowe and Robert Carnahan oversaw neutralization tests. Alexander Bukreyev oversaw neutralization tests. Jason McLellan oversaw structural studies and affinity experiments. Ivelin Georgiev oversaw experimental design and project progression. I performed oligo-labeling of antigens and quality control, sequence processing and bioinformatic analysis, data processing, antibody expression and purification, ELISA binding experiments, and ACE2 blocking experiments. Ivelin Georgiev and I wrote the manuscript.

4.1 Introduction

Although discovery of monoclonal antibodies (mAbs) targeting SARS-CoV-2 has led to the approval of several antibodies for COVID-19 therapy, their development was generally inefficient, with lead generation often requiring the production and testing of numerous antibody candidates. Here, we report that the integration of target–ligand blocking with a previously described B-cell receptor sequencing approach (LIBRA-seq) enables the rapid and efficient identification of multiple neutralizing antibodies that prevent the binding of SARS-CoV-2 Spike (S) protein to angiotensin-converting enzyme 2 (ACE2) combination of target–ligand blocking and high-throughput antibody sequencing promises to increase the throughput of programs aimed at discovering novel neutralizing antibodies.

Technologies for developing preventive and therapeutic measures that can counteract potential pandemics are of utmost significance for public health. The COVID-19 pandemic has emphasized the importance of rapid countermeasure development. Through pandemic preparedness initiatives, effective SARS-CoV-2 neutralizing antibodies were discovered and validated within months[99, 151, 152, 155, 201, 202], as were SARS-CoV-2 vaccine candidates[98]. However, even with such unprecedented speed of vaccine and therapeutic development, the pandemic has inflicted devastating worldwide effects. Accelerating actions by weeks or months can make an enormous difference in an exponentially evolving pandemic. Therefore, efficient methods for discovery of effective countermeasures against emerging pathogens can play a critical role in pandemic preparedness for future infectious disease outbreaks.

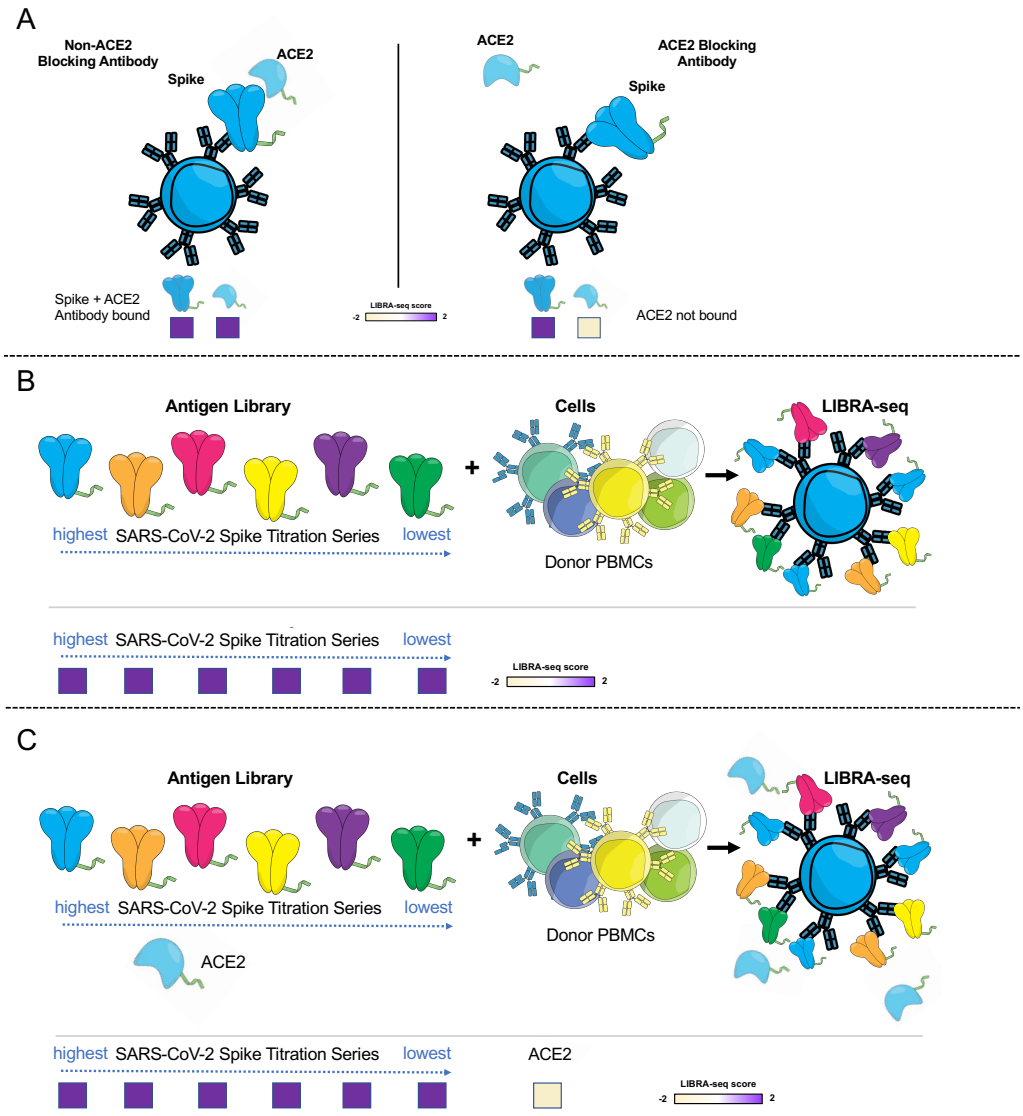


Figure 4.1. Schematic representation of LIBRA-seq experiments. (A) An antigen screening library of oligonucleotide-labeled antigens was generated. This library consisted of SARS-CoV-2 spike antigens and negative controls. Additionally, oligo-labeled ACE2 (the SARS-CoV-2 spike host cell receptor) was included. The antigen screening library was mixed with donor PBMCs. This approach allowed for assessment of B cell ligand blocking functionality from the sequencing experiment. (B) An antigen screening library containing an antigen titration was generated, with a goal of identifying high affinity antibodies from LIBRA-seq. In this experiment, six different amounts of oligo-labeled SARS-CoV-2 S protein, each labeled with a different barcode, were included in a screening library. (C) Schematic of LIBRA-seq with S titrations and ACE2 included for ligand blocking. Figure adapted from Shiakolas et al., *Nature Biotechnology*, 2022.

Antibodies are a major modality for therapy and vaccine design strategies for a wide range of diseases; however, the functional antibody discovery process can be inefficient. Typically, at the screening step, B cells are prioritized based on antigen-recognition, but this often requires time-intensive subsequent monoclonal antibody

validation steps for discovery of functional, neutralizing antibodies. This limitation was exemplified by SARS-CoV-2 antibody discovery initiatives, as testing of large numbers of antibodies (frequently hundreds to thousands) was generally required to identify a small fraction of neutralizing antibodies, with a wide range of hit rates when using Spike (S) as an antigen bait (about 2 to 23%) or when using RBD and/or S1 (about 2-55%)[99, 151, 152, 155, 157, 160, 201-205] in various studies.

To overcome this limitation, we developed LIBRA-seq with ligand blocking, a second-generation LIBRA-seq technology that incorporates a functional readout into the antibody discovery process[116]. LIBRA-seq (Linking B cell receptor to antigen specificity through sequencing) uses DNA-barcoded antigens to map antibody sequence to antigen specificity using next-generation sequencing[116]. For LIBRA-seq with ligand blocking, a ligand and its cognate target antigen(s) are each labeled with a unique oligonucleotide barcode (**Figure 4.1A**), enabling the transformation of antigen-ligand interactions into sequence-able events. In these experiments, B cells that can block antigen-ligand interactions are expected to have high LIBRA-seq scores for the target antigen(s) and low LIBRA-seq scores for the ligand (**Figure 4.1A**). Therefore, a single high-throughput LIBRA-seq with ligand blocking experiment provides both antigen recognition and ligand blocking information simultaneously for many B cells.

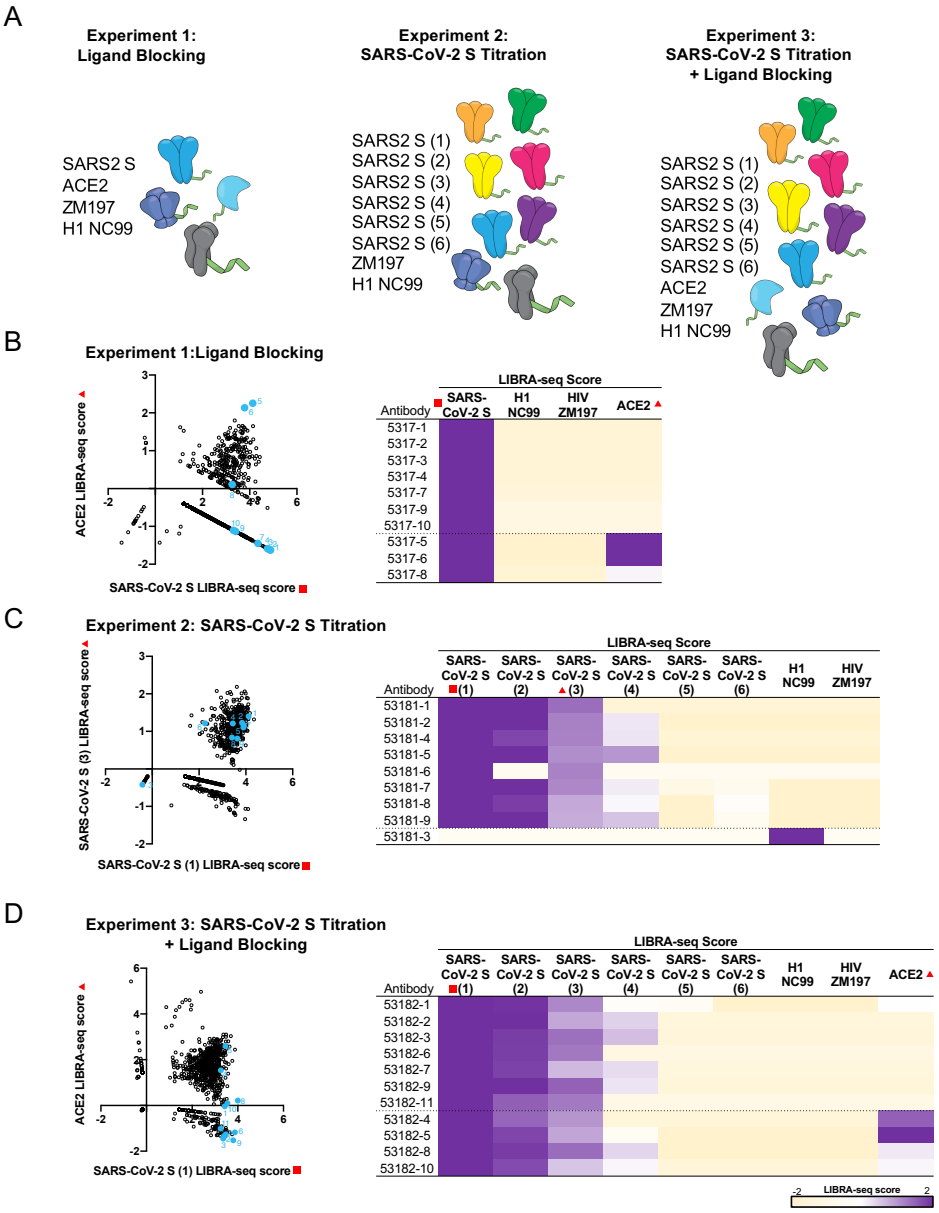


Figure 4.2. Antibody discovery using LIBRA-seq with ligand blocking. (A) Experimental setup of three LIBRA-seq experiments: experiment 1, LIBRA-seq with ligand blocking; experiment 2, LIBRA-seq with a SARS-CoV-2 S titration; and experiment 3, LIBRA-seq with a SARS-CoV-2 S titration and ligand blocking. For experiment 2 and 3, six different aliquots of S protein were added in a titration series (1-6). (B-D) (left) After next-generation sequencing, hundreds of B cells (dots) were recovered that had paired heavy/light chain sequencing information and antigen reactivity information for the three experiments. For experiment 1 (B), 2 (C), and 3 (D), select LIBRA-seq scores for all cells per experiment are shown as open circles (n=828, 829, 957, respectively). Antibodies selected for expression and validation are highlighted and numbered in light blue. (right) LIBRA-seq scores for the selected antibodies for all antigens from each experiment are shown as a heatmap from -2 to 2 (tan to purple); scores outside of this range are shown as the minimum and maximum values. For experiments 1 and 3, antibodies with negative scores for ACE2 are shown above the dotted line while antibodies with positive scores for ACE2 are shown below the dotted line and are controls. For experiment 2, all SARS-CoV-2 reactive antibodies are shown above the dotted line, whereas influenza specific antibody 53181-3 is shown as a control below the dotted line. Figure adapted from Shiakolas et al., *Nature Biotechnology*, 2022.

4.2 Results

To evaluate this technology, we sought to discover SARS-CoV-2-specific antibodies from B cells from subjects with past SARS-CoV-2 infection, since antibodies that block the interactions of the SARS-CoV-2 S protein with its host receptor angiotensin-converting enzyme 2 (ACE2) are among the most potently neutralizing identified to date[151, 152, 155, 157, 160, 201, 202]. We performed three LIBRA-seq experiments, with screening libraries that included: experiment 1, ACE2 and SARS-CoV-2 S; experiment 2, a titration series of different aliquots of SARS-CoV-2 S, each labeled with a unique barcode; and experiment 3, ACE2 and a titration series of S (**Figure 4.1A-C, Figure 4.2A**). The incorporation of a titration series of S antigen in the screening library for experiments 2 and 3 aimed to assess the strength of BCR-antigen interactions (**Figure 4.1B-C, 4.2A**).

The application of LIBRA-seq resulted in 828, 829, and 957 antigen-specific B cells for the three experiments, respectively. We prioritized a set of B cells for monoclonal antibody production and validation based on the following conditions: for experiments 1 and 3 (with ACE2 in the screening library), we selected B cells with high LIBRA-seq scores for S and low scores for ACE2; and for experiment 2, we selected B cells that had positive scores for multiple aliquots of S (**Figure 4.2B-D**). B cells with high S and high ACE2 scores were also selected as controls from experiments 1 and 3, along with an influenza-specific B cell from experiment 2 (**Figure 4.2B-D**). We further aimed to prioritize antibodies with diverse sequence features, although some of the selected antibodies appeared to be clonally related (**Figure 4.3A**).

A

Antibody	V_H Gene	J_H Gene	V_H Identity	J_H Identity	CDRH3 Sequence	CDRH3 Length	V_L Gene	J_L Gene	V_L Identity	J_L Identity	CDRL3 Sequence	CDRL3 Length
5317-1	*IGHV1-58*	*IGHJ4*	0.95	0.88	CAADPFADYW	8	*IGLV1-44*	*IGLJ3*	0.98	0.95	CATWDDSLNAWVF	11
5317-2	*IGHV1-69*	*IGHJ5*	0.95	0.92	CARGLWFGDSETVWFDPW	16	*IGKV1-39*	*IGKJ1*	0.97	0.97	CQQSYSTPPTF	9
5317-3	*IGHV3-64D*	*IGHJ4*	0.93	0.88	CVKGKIQLWLGADYW	13	*IGKV1-39*	*IGKJ1*	0.97	1.00	CQQSYNTPWTF	9
5317-4	*IGHV4-34*	*IGHJ3*	0.93	0.88	CARKPLLHSSVNPGAFDIW	17	*IGKV3-20*	*IGKJ2*	0.96	0.92	CQQYATSPRTF	9
5317-5	*IGHV1-69*	*IGHJ4*	0.97	0.96	CAREKGYSSSSSATYYLDFW	18	*IGLV1-40*	*IGLJ2*	0.98	0.95	CQSYDSSLTALVF	11
5317-6	*IGHV4-39*	*IGHJ6*	0.97	0.84	CARRVPGDYYCLDVW	13	*IGKV1-39*	*IGKJ4*	0.96	0.95	CQQSFSARVPTF	10
5317-7	*IGHV3-7*	*IGHJ4*	0.97	0.88	CARGGLWGTFDYW	11	*IGKV3-20*	*IGKJ2*	0.98	1.00	CQQFAYSLYTF	9
5317-8	*IGHV3-30-3*	*IGHJ6*	0.96	0.87	CARAYGGNYYYGMDVW	14	*IGLV3-1*	*IGLJ1*	0.96	0.97	CQAWDSSTASFVF	11
5317-9	*IGHV1-69*	*IGHJ5*	0.95	0.84	CASLGGDSYISGTHYDRSGYDPW	21	*IGKV3-11*	*IGKJ4*	0.98	0.95	CQRRSNWPPFTF	10
5317-10	*IGHV5-10-1*	*IGHJ4*	0.93	0.81	CARVNRVGDGPDFW	12	*IGKV2-28*	*IGKJ1*	0.97	1.00	CMQALQTPWTF	9
53181-1	*IGHV1-18*	*IGHJ6*	0.97	0.95	CARDPASYYDFWSGYVDYYYYGMDVW	24	*IGKV3-20*	*IGKJ3*	0.99	1.00	CQQYGNSRLTF	9
53181-2	*IGHV1-18*	*IGHJ6*	0.99	0.95	CARDPASYYDLWSGYVDYYYYGMDVW	24	*IGKV3-20*	*IGKJ3*	0.99	1.00	CHHYGSSRLTF	9
53181-3	*IGHV1-69*	*IGHJ4*	0.91	0.75	CARSGGYRLWFGELW	13	*IGKV3-20*	*IGKJ5*	0.94	0.95	CQQYGGSPATF	9
53181-4	*IGHV3-33*	*IGHJ4*	0.98	0.88	CAREGAVGATSGLDYW	14	*IGLV3-10*	*IGLJ3*	0.99	0.89	CYSRDSSGNPLF	10
53181-5	*IGHV4-59*	*IGHJ4*	0.95	0.96	CARGFDYW	6	*IGKV3-20*	*IGKJ1*	0.98	1.00	CQQYGSSPWTF	9
53181-6	*IGHV4-59*	*IGHJ6*	0.94	0.89	CARGAGEQRLVGGLFGVSHFYYYMDVW	25	*IGKV1-5*	*IGKJ1*	0.99	1.00	CQQYNSYPWTF	9
53181-7	*IGHV3-23*	*IGHJ4*	0.94	0.81	CAKSATIVLMVSAIYW	14	*IGLV2-14*	*IGLJ2*	0.97	1.00	CSSYTSTSTLVF	10
53181-8	*IGHV3-72*	*IGHJ6*	0.99	0.94	CARVRGGEWVGDLGWYYYYGMDVW	22	*IGKV2-28*	*IGKJ2*	0.99	0.97	CMQALQTPRTF	9
53181-9	*IGHV3-30*	*IGHJ4*	0.91	0.92	CVKGATKIDYW	9	*IGLV2-14*	*IGLJ1*	0.91	0.83	CFSYTSGGTRVF	10
53182-1	*IGHV3-53*	*IGHJ3*	0.95	0.96	CTRGGWPSGDTFDIW	13	*IGKV1-9*	*IGKJ3*	0.99	0.97	CQQLNSYPEITF	10
53182-2	*IGHV3-48*	*IGHJ5*	0.98	0.92	CAREGGWYSVGWVDPW	14	*IGKV3-20*	*IGKJ2*	0.98	0.97	CQQYGSSRTF	8
53182-3	*IGHV3-66*	*IGHJ6*	0.99	0.87	CARDRRIIGYYFGMDVW	15	*IGLV2-23*	*IGLJ3*	0.97	1.00	CCPYADTWVF	8
53182-4	*IGHV2-5*	*IGHJ3*	0.85	0.86	CARLLIEHDAFDIW	12	*IGKV2-28*	*IGKJ2*	0.92	0.97	CMQALHFPYTF	9
53182-5	*IGHV1-69*	*IGHJ4*	0.85	0.88	CAREEGSGWWKHDYW	13	*IGKV1-9*	*IGKJ2*	0.97	1.00	CQQLSGYPYTF	9
53182-6	*IGHV3-66*	*IGHJ6*	0.95	0.84	CVRDRRIVGYYFGLDVW	15	*IGLV2-23*	*IGLJ3*	0.96	0.95	CCSYATTWVF	8
53182-7	*IGHV3-23*	*IGHJ6*	0.98	0.97	CAKDAFYYGSGSHFYYYYYMDVW	21	*IGKV3-20*	*IGKJ5*	0.99	0.97	CQQYGSSPTF	8
53182-8	*IGHV4-39*	*IGHJ4*	0.99	0.92	CARDRRGGGWTASFDFW	15	*IGKV4-1*	*IGKJ2*	0.99	0.97	CQQHYSTPGYTF	10
53182-9	*IGHV3-53*	*IGHJ3*	0.98	0.98	CARGGWPSGDTFDIW	13	*IGKV1-9*	*IGKJ3*	0.99	0.97	CQQLNSYPEITF	10
53182-10	*IGHV2-5*	*IGHJ4*	0.98	0.83	CAHHTVPTIYDYW	11	*IGLV2-8*	*IGLJ3*	0.98	0.97	CSSYAGSNPLVF	10
53182-11	*IGHV3-9*	*IGHJ3*	0.93	0.92	CAKDIGRYDHYNIFGRVGGAFDIW	22	*IGKV1-33*	*IGKJ3*	0.95	0.94	CQHYDNLPRF	8

B

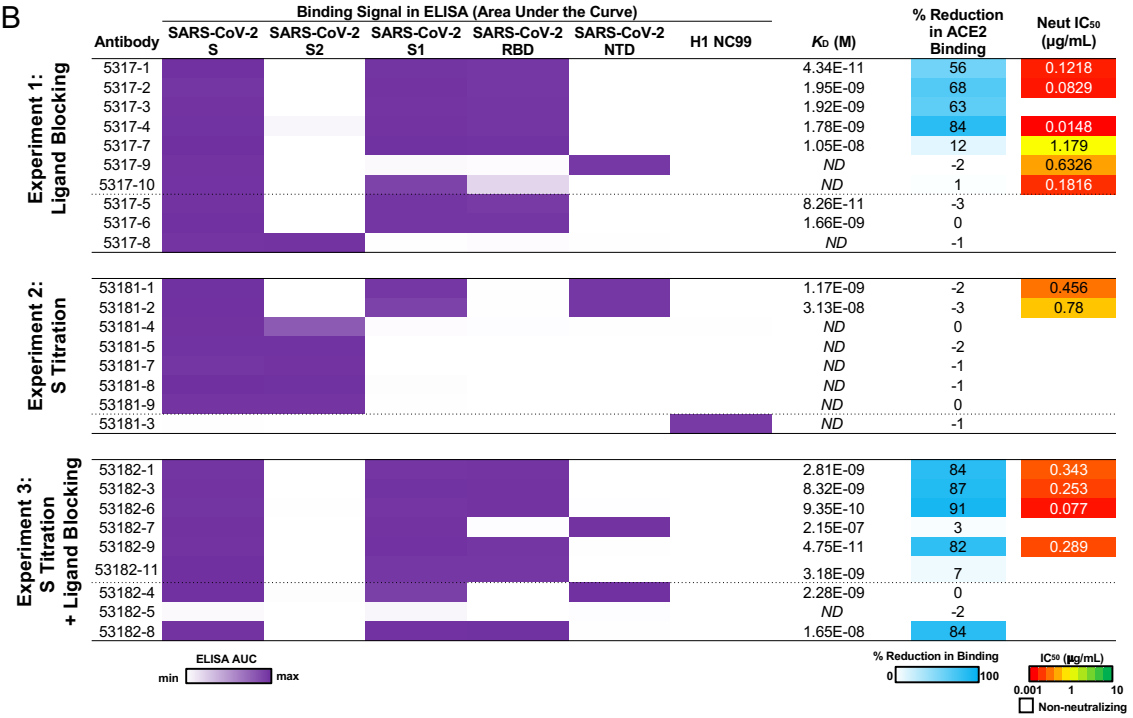


Figure 4.3. Validation and characterization of selected antibodies. (A) Genetic characteristics for monoclonal antibodies prioritized for expression and validation. V_H, J_H, V_L, J_L inferred gene segment identity is shown at the nucleotide level. CDRH3 and CDRL3 amino acid sequence and length are shown. (B) ELISA area under the curve (AUC) values for binding to SARS-CoV-2 recombinant antigen proteins and a negative control influenza hemagglutinin protein are shown for antibodies (rows) in each experiment, calculated from data in Figure 4.4A. Next, shown are K_D (M) of antibodies for SARS-CoV-2 RBD or NTD (based on epitope shown) was determined by biolayer interferometry. ND, not done. Next, percent reduction in ACE2 binding by ELISA is shown as a heatmap from 0 to 100% (white to blue) reduction in binding compared to SARS-CoV-2 binding only. Last, VSV SARS-CoV-2 neutralization IC_{50} values are shown as a heatmap from high potency (red) to low potency (green). Non-neutralizing antibodies are shown as white. Figure adapted from Shiakolas et al., *Nature Biotechnology*, 2022.

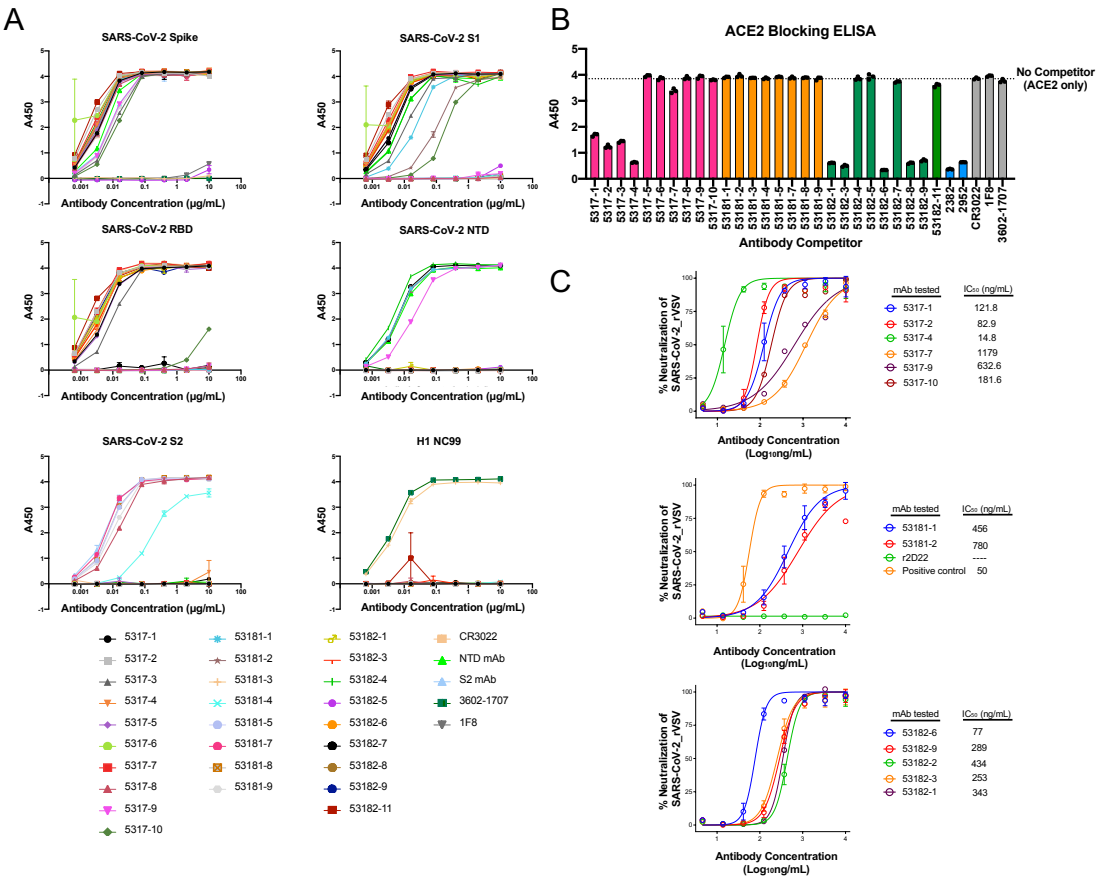


Figure 4.4. Characterization of LIBRA-seq-identified antibodies. (A) ELISA binding of antibodies. Data represented as mean ± SEM of technical duplicates and represent one of at least two independent experiments. (B) ACE2 blocking ELISA. ACE2 binding without competitor shown as dotted line. ELISAs were performed at one antibody concentration, and data are represented as mean ± SEM of technical triplicates and represent one of at least two independent experiments. (C) Antibodies were tested in a VSV SARS-CoV-2 real time cell analysis (RTCA) neutralization assay. Neutralization curves and IC_{50} values are shown. Data are represented as mean ± S.D. of technical triplicates, and represent one of two independent experiments. Figure 4.4C made by Naveenchandra Suryadevara. Figure adapted from Shiakolas et al., *Nature Biotechnology*, 2022.

We confirmed the predicted antigen specificity for 26/27 (96%) antibodies and mapped the general antibody epitope regions by testing antibodies for binding to recombinant SARS-CoV-2 subdomain proteins (**Figure 4.3B, 4.4A**). The majority of antibodies from experiments 1 and 3 (but none from experiment 2) recognized the RBD (**Figure 4.3B, 4.4A**). Further, the antibodies had a wide range of affinities for RBD or NTD, including several antibodies with K_D <1 nM, although we did not observe a correlation between LIBRA-seq spike score and affinity (**Figure 4.3B**). Next, we tested the ability of the antibodies to block ACE2 binding to spike. For antibodies predicted to

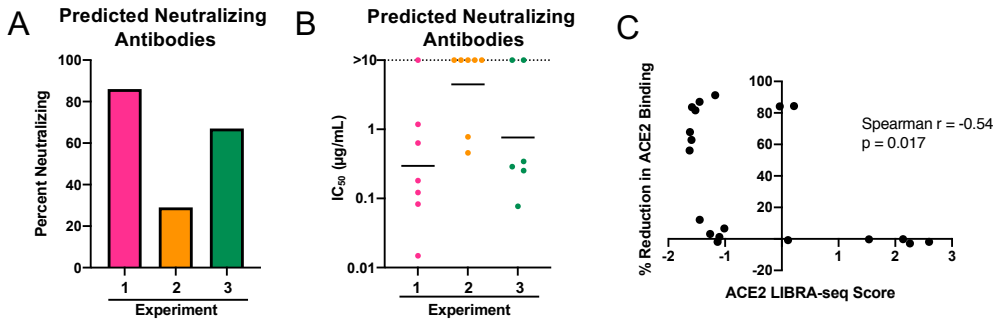


Figure 4.5. Assessment of LIBRA-seq with ligand blocking. (A) "Predicted Neutralizing Antibodies" were defined as the subset of selected antibodies with negative ACE2 LIBRA-seq scores from experiments 1 (n=7 antibodies) and 3 (n=6 antibodies), and all antibodies with high LIBRA-seq scores (>1) for SARS-CoV-2 S from experiment 2 (n=7 antibodies). The percent of neutralizing antibodies from the set of predicted neutralizers is shown for each experiment. (B) The IC_{50} values (µg/mL) for SARS-CoV-2 neutralization by RTCA with VSV-SARS-CoV-2 (IC_{50} value for each antibody shown as single dot) are plotted for the set of predicted neutralizers. Horizontal line shown is geometric mean for each experiment. Non-neutralizing antibodies are shown as >10 µg/mL. (C) Spearman correlation of ACE2 LIBRA-seq score (x-axis) and % Reduction in ACE2 Binding to SARS-CoV-2 (y-axis) for antibodies from experiments 1 and 3. Spearman r = -0.54, p = 0.017 (two-tailed, 95% confidence interval). Figure adapted from Shiakolas et al., *Nature Biotechnology*, 2022.

block ACE2 by LIBRA-seq, 57% from experiment 1 and 67% from experiment 3 demonstrated ACE2 blocking via ELISA, whereas no antibodies from experiment 2 blocked ACE2 binding (**Figure 4.3B, 4.4B**).

Next, we tested the antibodies in a VSV SARS-CoV-2 chimeric virus neutralization assay (**Figure 4.3B, 4.4C**). For antibodies predicted to block ACE2 by LIBRA-seq, 86% from experiment 1 and 67% from experiment 3 were neutralizing, while only two clonally related antibodies (29%) from experiment 2 were neutralizing (**Figure 4.5A-B**). For the antibodies from experiments 1 and 3, the ACE2 LIBRA-seq scores were correlated with the percent reduction in ACE2 binding (**Figure 4.5C**, Spearman r = -0.54, p = 0.017). Furthermore, several antibodies also showed potent neutralization against authentic SARS-CoV-2 virus in a plaque reduction assay, and in some cases against multiple SARS-CoV-2 variants (**Figure 4.6**). Together, these results highlight

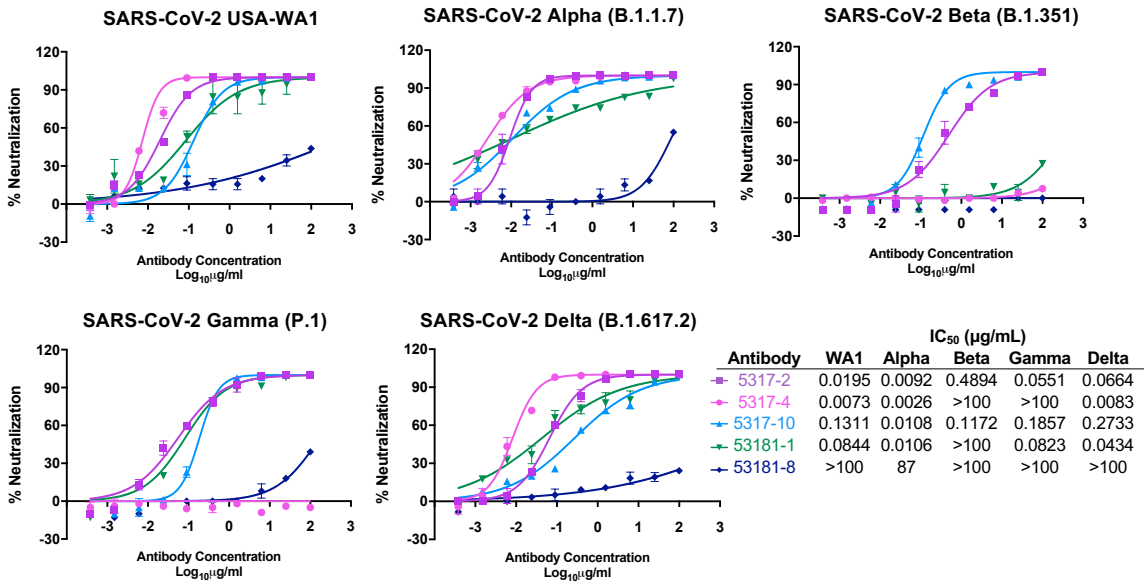


Figure 4.6. Antibody neutralization of SARS-CoV-2 variants. Authentic SARS-CoV-2 neutralization for a panel of antibodies is shown against USA-WA1 and variants (Alpha, Beta, Gamma, and Delta). Data represent the % neutralization as mean ± SD. The IC_{50} values calculated in GraphPad prism software by 4-parameter best-fit analysis are shown to the right of the panel. Figure made with Siva Periasamy. Figure adapted from Shiakolas et al., *Nature Biotechnology*, 2022.

the importance of including ligand blocking in LIBRA-seq for selectively identifying potent neutralizing antibodies.

To investigate antibody recognition of SARS-CoV-2 S, we determined a 9 Å-resolution Cryo-EM structure of the antigen-binding fragments of antibodies 5317-4 and 5317-10 bound to the SARS-CoV-2 S extracellular domain (**Figure 4.7A**). We chose 5317-4 based on its potent neutralization (IC_{50} value of 7.3 ng/mL against authentic SARS-CoV-2, **Figure 4.6**) and ACE2 competition. The 3D reconstruction revealed that 5317-4 binds to RBD in the “up” and “down” conformations, and its epitope partially overlaps the ACE2 binding footprint (**Figure 4.7A-B**). When bound to the RBD in the down conformation, 5317-4 competes with ACE2 binding to the adjacent up RBD (**Figure 4.7B**). We investigated 5317-10 because of its inconclusive epitope, as it bound to S1 but not individual RBD or NTD constructs (**Figure 4.3B, 4.4A**). The map revealed

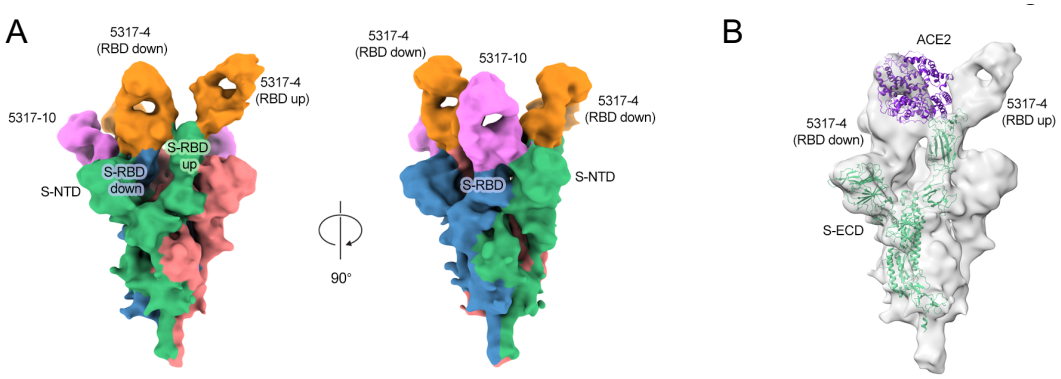


Figure 4.7. Structural characterization of antibodies 5317-4 and 5317-10. (A) 9 Å-resolution cryo-EM structure of Fab-spike complex for 5317-4 Fab (orange) and 5317-10 Fab (pink). Spike protomers are shown in green, blue, and red. (B) Fab-spike complex structure modeled with ACE2 (purple). Figure made by Nicole Johnson. Figure adapted from Shiakolas et al., *Nature Biotechnology*, 2022.

that 5317-10 binds a quaternary epitope that bridges an RBD in the down position and the NTD of an adjacent protomer (**Figure 4.7A**). This mode of recognition may prevent the RBD from transitioning into an ACE2-accessible up position, thereby preventing binding by ACE2.

To further demonstrate the utility of LIBRA-seq with ligand blocking, we sought to identify antibodies that show cross-reactivity between SARS-CoV-2 and SARS-CoV, and that are capable of blocking spike-ACE2 interactions. To that end, we applied LIBRA-seq to B cells from a subject with past SARS-CoV-2 infection, using an antigen library that included SARS-CoV-2 S, SARS-CoV S, and ACE2 (**Figure 4.8A**). This resulted in 120 IgG^+ B cells with high LIBRA-seq scores for both SARS-CoV-2 S and SARS-CoV S (**Figure 4.8B**). Only 8% of these cells were associated with low LIBRA-seq scores for ACE2 (**Figure 4.8C**), highlighting the advantage of including ligand blocking to screen for such rare cells (although we also note that information about B cells that show cross-reactivity but are not ACE2 blocking is also retained, enabling characterization of B cells with alternative phenotypes as well). Based on LIBRA-seq antigen and ligand blocking scores, we produced and validated a set of antibodies,

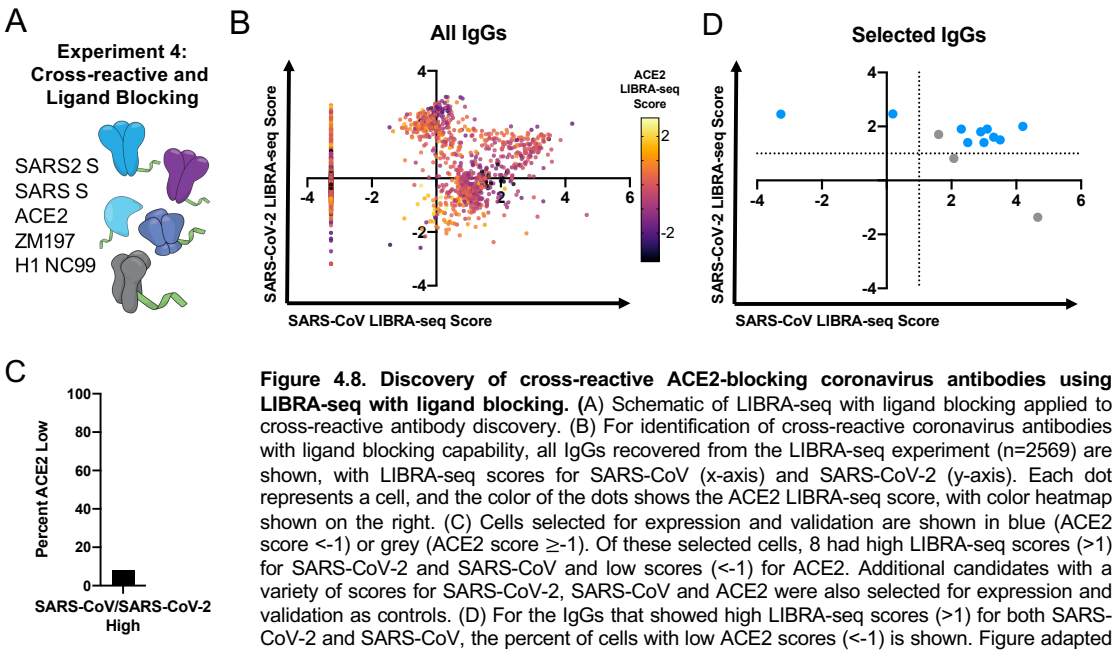


Figure 4.8. Discovery of cross-reactive ACE2-blocking coronavirus antibodies using LIBRA-seq with ligand blocking. (A) Schematic of LIBRA-seq with ligand blocking applied to cross-reactive antibody discovery. (B) For identification of cross-reactive coronavirus antibodies with ligand blocking capability, all IgGs recovered from the LIBRA-seq experiment (n=2569) are shown, with LIBRA-seq scores for SARS-CoV (x-axis) and SARS-CoV-2 (y-axis). Each dot represents a cell, and the color of the dots shows the ACE2 LIBRA-seq score, with color heatmap shown on the right. (C) Cells selected for expression and validation are shown in blue (ACE2 score <-1) or grey (ACE2 score ≥-1). Of these selected cells, 8 had high LIBRA-seq scores (>1) for SARS-CoV-2 and SARS-CoV and low scores (<-1) for ACE2. Additional candidates with a variety of scores for SARS-CoV-2, SARS-CoV and ACE2 were also selected for expression and validation as controls. (D) For the IgGs that showed high LIBRA-seq scores (>1) for both SARS-CoV-2 and SARS-CoV, the percent of cells with low ACE2 scores (<-1) is shown. Figure adapted from Shiakolas et al., *Nature Biotechnology*, 2022.

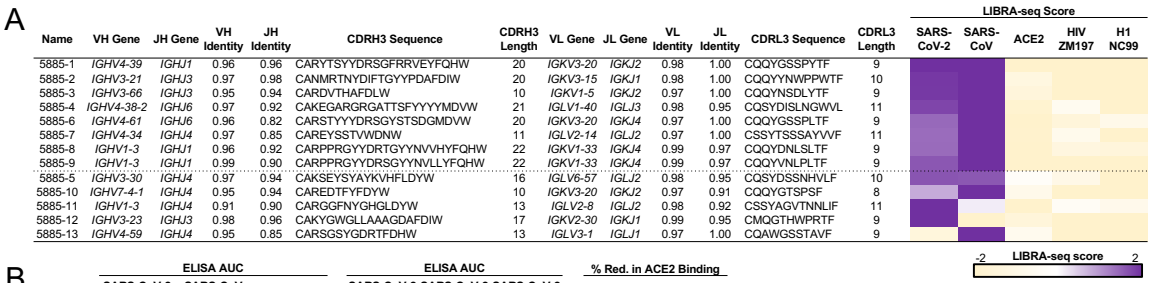
including 8 with high scores for both S antigens and low scores for ACE2 (**Figure 4.8D, 4.9A**). Of these, 100% bound SARS-CoV-2 S, 88% showed the predicted SARS-CoV-2/SARS-CoV cross-reactivity, and 63% demonstrated strong ACE2 blocking ability via ELISA (**Figure 4.9B-E**), confirming that LIBRA-seq with ligand blocking efficiently identified ACE2-blocking antibodies with cross-reactivity between multiple coronaviruses.

4.3 Discussion

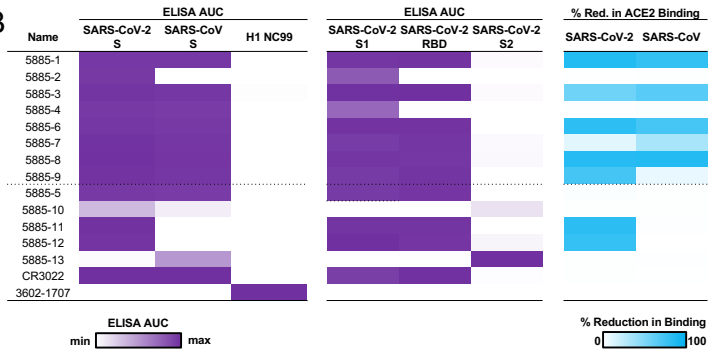
Together, the results from the four LIBRA-seq experiments reported here showcase the advantages of including ligand blocking as part of the sequencing readout. As with most screening tools, there are limitations to the LIBRA-seq with ligand blocking approach, including the prerequisite for a defined antigen-ligand interaction, as well as the potential for identifying false positives. Nevertheless, through a single high-

A

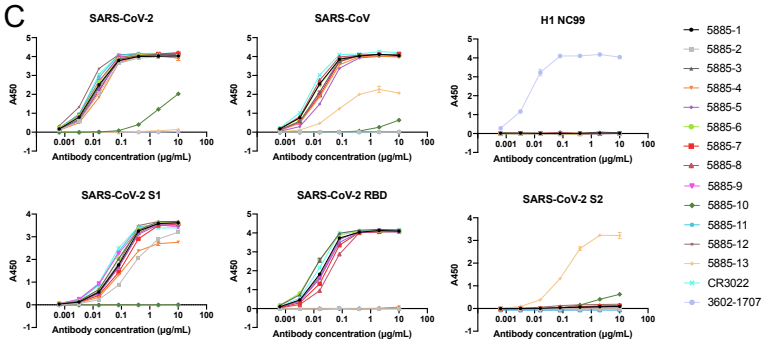
Name	VH Gene	JH Gene	VH Identity	JH Identity	CDRH3 Sequence	CDRH3 Length	VL Gene	JL Gene	VL Identity	JL Identity	CDRL3 Sequence	CDRL3 Length	LIBRA-seq Score: SARS-CoV-2	SARS-CoV	ACE2	HIV ZM197	H1 NC99
5885-1	*IGHV4-39*	*IGHJ1*	0.96	0.96	CARYTSYYDRSGFRRVEYFQHW	20	*IGKV3-20*	*IGKJ2*	0.98	1.00	CQQYGSSPYTF	9					
5885-2	*IGHV3-21*	*IGHJ3*	0.97	0.98	CANMRTNYDIFTGYYPDAFDIW	20	*IGKV3-15*	*IGKJ1*	0.98	1.00	CQQYYNWPPWTF	10					
5885-3	*IGHV3-66*	*IGHJ3*	0.95	0.94	CARDVTHAFDLW	10	*IGKV1-5*	*IGKJ2*	0.97	1.00	CQQYNSDLYTF	9					
5885-4	*IGHV4-38-2*	*IGHJ6*	0.97	0.92	CAKEGARGRGATTSFYYYYMDVW	21	*IGLV1-40*	*IGLJ3*	0.98	0.95	CQSYDISLNGWVL	11					
5885-6	*IGHV4-61*	*IGHJ6*	0.96	0.82	CARSTYYYDRSGYSTSDGMDVW	20	*IGKV3-20*	*IGKJ4*	0.97	1.00	CQQYGSSPLTF	9					
5885-7	*IGHV4-34*	*IGHJ4*	0.97	0.85	CAREYSSTVWDNW	11	*IGLV2-14*	*IGLJ2*	0.97	1.00	CSSYTSSSAYVVF	11					
5885-8	*IGHV1-3*	*IGHJ1*	0.96	0.92	CARPPRGYYDRTGYYNVVHYFQHW	22	*IGKV1-33*	*IGKJ4*	0.99	0.97	CQQYDNLSLTF	9					
5885-9	*IGHV1-3*	*IGHJ1*	0.99	0.90	CARPPRGYYDRSGYYNVLLYFQHW	22	*IGKV1-33*	*IGKJ4*	0.99	0.97	CQQYVNLPLTF	9					
5885-5	*IGHV3-30*	*IGHJ4*	0.97	0.94	CAKSEYSYAYKVHFLDYW	16	*IGLV6-57*	*IGLJ2*	0.98	0.95	CQSYDSSNHVLF	10					
5885-10	*IGHV7-4-1*	*IGHJ4*	0.95	0.94	CAREDTFYFDYW	10	*IGKV3-20*	*IGKJ2*	0.97	0.91	CQQYGTSPSF	8					
5885-11	*IGHV1-3*	*IGHJ4*	0.91	0.90	CARGGFNYGHGLDYW	13	*IGLV2-8*	*IGLJ2*	0.98	0.92	CSSYAGVTNNLIF	11					
5885-12	*IGHV3-23*	*IGHJ3*	0.98	0.96	CAKYGWGLLAAAGDAFDIW	17	*IGKV2-30*	*IGKJ1*	0.99	0.95	CMQGTHWPRTF	9					
5885-13	*IGHV4-59*	*IGHJ4*	0.95	0.85	CARSGSYGDRTFDHW	13	*IGLV3-1*	*IGLJ1*	0.97	1.00	CQAWGSSTAVF	9					



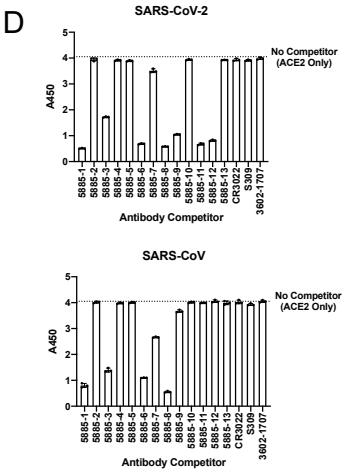
B



C



D



E

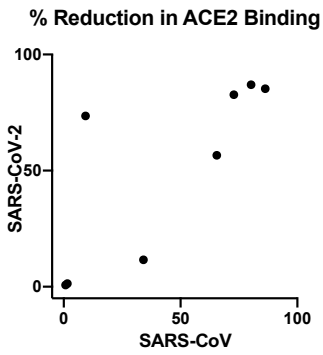


Figure 4.9. Characterization of selected cross-reactive antibodies. (A) The 8 IgGs with high LIBRA-seq scores for SARS-CoV-2 and SARS-CoV and low scores for ACE2 are shown above the dotted line. Control antibodies with other LIBRA-seq score patterns are shown below the dotted line. For each antibody, CDR sequences and lengths are shown at the amino acid level and V-gene and J-gene identity are shown at the nucleotide level. LIBRA-seq scores for antigens included in the screening library are shown as a heatmap low (tan) – white – high (purple). Scores outside of this range are shown as the minimum and maximum values. (B) ELISA area under the curve (AUC) values from binding to coronavirus spike proteins, influenza hemagglutinin H1 NC99 (negative control), and recombinant antigen domains are shown as a heatmap from minimum (white) to maximum (purple) binding. Percent reduction in ACE2 binding by ELISA is shown for SARS-CoV-2 and SARS-CoV spikes, and displayed as a heatmap from 0% (white) to 100% (blue). (C) ELISA binding of antibodies to SARS-CoV-2 spike, SARS-CoV spike, influenza hemagglutinin H1 NC99, SARS-CoV-2 S1, SARS-CoV-2 RBD, and SARS-CoV-2 S2. Data are represented as mean ± SEM of technical duplicates and represent one of at least two independent experiments (n=2). (D) ACE2 blocking ELISA. ACE2 binding without competitor is shown as a dotted line. ELISAs were performed at one antibody concentration, and data are represented as mean ± SEM of technical triplicates and represent one of at least two independent experiments (n=2). (E) For the 8 IgGs with high LIBRA-seq scores for SARS-CoV-2 and SARS-CoV and low scores for ACE2, the percent reduction in ACE2 binding due to antibody blocking by ELISA is shown for SARS-CoV (x-axis) and SARS-CoV-2 (y-axis). Figure adapted from Shiakolas et al., *Nature Biotechnology*, 2022.

throughput sequencing experiment, LIBRA-seq with ligand blocking identified potent SARS-CoV-2 antibodies, requiring the subsequent production and validation of less than a dozen antibodies per experiment. The observed hit rates for the discovery of potently neutralizing antibodies are an improvement over what has been reported in the literature, which also typically required the screening of hundreds to thousands of antibody candidates isolated for their reactivity to antigen alone (recombinant S, S1, or RBD)[99, 151, 152, 155, 157, 160, 201-205]. Further, unlike RBD-only discovery efforts, LIBRA-seq with ligand blocking applied to spike antigens has the potential for more comprehensive coverage of antibody epitopes, as evidenced by the discovery of the RBD-NTD antibody in **Figure 4.7A.** Overall, the application of LIBRA-seq with ligand blocking can provide critical advantages for rapid development of therapeutic and preventive countermeasures and presents a general platform with applications to virtually any area where targeting the disruption of antigen-ligand interaction is a prime therapeutic goal.

Acknowledgements

We thank Angela Jones, Latha Raju, and Jamie Roberson of Vanderbilt Technologies for Advanced Genomics for their expertise regarding NGS and library preparation; David Flaherty and Brittany Matlock of the Vanderbilt Flow Cytometry Shared Resource for help with flow panel optimization; and members of the Georgiev laboratory for comments on the manuscript. The Vanderbilt VANTAGE Core provided technical assistance for this work. VANTAGE is supported in part by CTSA grant 5UL1 RR024975-03, the Vanderbilt Ingram Cancer Center (P30 CA68485), the Vanderbilt Vision Center (P30 EY08126), and NIH/NCRR (G20 RR030956). This work was

conducted in part using the resources of the Advanced Computing Center for Research and Education at Vanderbilt University (Nashville, TN). Flow cytometry experiments were performed in the VUMC Flow Cytometry Shared Resource. The VUMC Flow Cytometry Shared Resource is supported by the Vanderbilt Ingram Cancer Center (P30 CA68485) and the Vanderbilt Digestive Disease Research Center (DK058404).

For work described in this manuscript, I.S.G., A.R.S., K.J.K., S.C.W., K.A.P., N.R., and L.M.W. were supported in part by NIH NIAID award R01AI131722-S1, the Hays Foundation COVID-19 Research Fund, Fast Grants, and CTSA award No. *UL1 TR002243* from the National Center for Advancing Translational Sciences. J.S.M, N.V.J., and D.W. were supported in part by a National Institutes of Health (NIH)/National Institute of Allergy and Infectious Diseases (NIAID) grant awarded to J.S.M. (R01-AI127521), Welch Foundation grant F-0003-19620604, and NIH NIAID award R01AI131722-S1. J.E.C., N.S., R.N., R.E.S., and R.H.C. were supported in part by Defense Advanced Research Projects Agency (DARPA) grant HR0011-18-2-0001, U.S. N.I.H. contract 75N93019C00074, N.I.H. grant R01 AI157155, the Dolly Parton COVID-19 Research Fund at Vanderbilt, and a grant from Fast Grants, Mercatus Center, George Mason University. J.E.C. is a recipient of the 2019 Future Insight Prize from Merck KGaA, which supported this work with a grant.

Author Contributions

Methodology, A.R.S., I.S., and I.S.G.; Investigation, A.R.S., K.J.K., N.V.J., S.C.W., N.S., D.W., S.P., K.A.P., N.R., R.N., R.E.S., L.M.W., J.E.C., A.B., R.H.C., J.S.M., and I.S.G.; Software, A.R.S., N.R.; Validation, A.R.S., K.J.K.; Writing - Original

Draft, A.R.S. and I.S.G.; Writing -Review & Editing, all authors; Funding Acquisition, I.S.G., J.E.C., A.B., R.H.C., J.S.M., A.R.S and K.J.K.; Resources, I.S.G., J.E.C., A.B., R.H.C., and J.S.M..; Supervision, I.S.G.

Declaration of Interest

A.R.S. and I.S.G. are co-founders of AbSeek Bio. I.S.G., A.R.S, and K.J.K. are listed as inventors on antibodies described herein. I.S.G., A.R.S, and I.S. are listed as inventors on patent applications for the LIBRA-seq technology. J.E.C. has served as a consultant for Luna Biologics, is a member of the Scientific Advisory Board Meissa Vaccines and is Founder of IDBiologics. The Crowe laboratory has received funding support in sponsored research agreements from AstraZeneca, IDBiologics, and Takeda. The Georgiev laboratory at Vanderbilt University Medical Center has received unrelated funding from Takeda Pharmaceuticals. The remaining authors declare no competing interests.

4.4 Materials and Methods

Data Availability Statement

All unique reagents generated in this study are available from the corresponding author with a completed Material Transfer Agreement. Sequences for antibodies identified and characterized in this study have been deposited to GenBank (MZ517191-MZ517250, OM001674- OM001699). Raw sequencing data has been deposited to Sequence Read Archive (PRJNA744567, SAMN24369247). Custom scripts used to

analyze data in this manuscript are available on github at https://github.com/lglab-repo/LIBRA-seq-with-ligand-blocking.git.

Donor Information

PBMC samples were purchased from Cellero. The PBMCs were from subjects with past SARS-CoV-2 infection at least 14 days post symptom cessation. For experiment 1, three samples were pooled from donors 523, 527, and 528. For experiments 2 and 3, samples from donor 523 were used for LIBRA-seq. Donor 523 had a plaque reduction neutralization test titer of 1:2,560. For experiment 4 (cross-reactive antibody discovery with ligand blocking), a sample from donor 528 was used for LIBRA-seq.

Antigen Purification

A variety of recombinant soluble protein antigens were used in the LIBRA-seq experiment and other experimental assays.

Plasmids encoding residues 1–1208 of the SARS-CoV-2 spike with a mutated S1/S2 cleavage site, proline substitutions at positions 817, 892, 899, 942, 986 and 987, and a C-terminal T4-fibritin trimerization motif, an 8x HisTag, and a TwinStrepTag (SARS-CoV-2 spike HP); residues 1-1190 of the SARS-CoV spike with proline substitutions at positions 968 and 969, and a C-terminal T4-fibritin trimerization motif, an 8x HisTag, and a TwinStrepTag (SARS-CoV S-2P); and 1−615 of human ACE2 with a C-terminal HRV3C protease cleavage site, a TwinStrepTag and an 8XhisTag (ACE2) were transiently transfected in Expi293F cells using polyethylenimine. Transfected

supernatants were harvested 5 days after expression and purified over a StrepTrap column (Cytiva Life Sciences). Both recombinant SARS-CoV-2 S HP and ACE2 were further purified to homogeneity using a Superose6 Increase column (Cytiva Life Sciences).

For the HIV-1 gp140 SOSIP variant from strain ZM197 (clade C) and hemagglutinin from strain A/New Caledonia/20/99 (H1N1) (GenBank ACF41878), recombinant, soluble antigens contained an AviTag and were expressed in Expi293F cells using polyethylenimine transfection reagent and cultured. FreeStyle F17 expression medium supplemented with pluronic acid and glutamine was used. The cells were cultured at 37°C with 8% CO_2 saturation and shaking. After 5-7 days, cultures were centrifuged and supernatant was filtered and run over an affinity column of agarose-bound *Galanthus nivalis* lectin. The column was washed with PBS and antigens were eluted with 30 mL of 1M methyl-a-D-mannopyranoside. Protein elutions were buffer exchanged into PBS, concentrated, and run on a Superdex 200 Increase 10/300 GL Sizing column on the AKTA FPLC system. Fractions corresponding to correctly folded protein were collected, analyzed by SDS-PAGE and antigenicity was characterized by ELISA using known monoclonal antibodies specific to each antigen. AviTagged antigens were biotinylated using BirA biotin ligase (Avidity LLC).

SARS-CoV-2 S1, SARS-CoV-2 S2, SARS-CoV-2 RBD and SARS-CoV-2 NTD proteins were purchased from the commercial vendor, Sino Biological.

DNA-barcoding of Antigens

We used oligos that possess 15 bp antigen barcode, a sequence capable of annealing to the template switch oligo that is part of the 10X bead-delivered oligos and contain truncated TruSeq small RNA read 1 sequences in the following structure: 5'-CCTTGGCACCCGAGAATTCCANNNNNNNNNNNNNCCCATATAAGA*A*A-3', where Ns represent the antigen barcode[116]. For each antigen, a unique DNA barcode was directly conjugated to the antigen itself. For Experiment 1, the barcodes included SARS-CoV-2 S (GACAAGTGATCTGCA), H1 NC99 (TCATTTCCTCCGATT), ZM197 (TACGCCTATAACTTG), and ACE2 (CTTCACTCTGTCAGG). For Experiment 2, the barcodes included SARS-CoV-2 S aliquot 1 (GACAAGTGATCTGCA), SARS-CoV-2 S aliquot 2 (TGTGTATTCCCTTGT), SARS-CoV-2 S aliquot 3 (GCAGCGTATAAGTCA), SARS-CoV-2 S aliquot 4 (GCTCCTTTACACGTA), SARS-CoV-2 S aliquot 5 (AGACTAATAGCTGAC), SARS-CoV-2 S aliquot 6 (GGTAGCCCTAGAGTA), H1 NC99 (TCATTTCCTCCGATT), and ZM197 (TACGCCTATAACTTG). For Experiment 3, the same barcodes were included as Experiment 2 and also included ACE2 (CTTCACTCTGTCAGG). For Experiment 4, the barcodes included SARS-CoV-2 S (GCAGCGTATAAGTCA), SARS-CoV S (GCTCCTTTACACGTA), ACE2 (TACGCCTATAACTTG), ZM197 (TCATTTCCTCCGATT), and H1 NC99 (CTTCACTCTGTCAGG). In particular, amino-oligonucleotides were conjugated directly to each antigen using the SoluLINK Protein-Oligonucleotide Conjugation Kit (TriLink cat no. S-9011) according to manufacturer's instructions. Briefly, the oligo and protein were desalted, and then the amino-oligo was modified with the 4FB crosslinker, and the biotinylated antigen protein was modified with S-HyNic. Then, the 4FB-oligo and the

HyNic-antigen were mixed. This process causes a stable bond to form between the protein and the oligonucleotide. The concentration of the antigen-oligo conjugates was determined by a BCA assay, and the HyNic molar substitution ratio of the antigen-oligo conjugates was analyzed using the NanoDrop according to the SoluLINK protocol guidelines. AKTA FPLC was used to remove excess oligonucleotide from the protein-oligo conjugates, which were also verified using SDS-PAGE with a silver stain. Antigen-oligo conjugates were also used in flow cytometric titration experiments to determine optimal amounts for antigen-specific B cell sorting.

Antigen-specific B Cell Sorting

Cells were stained and mixed with DNA-barcoded antigens and other antibodies, and then sorted using fluorescence activated cell sorting (FACS). First, cells were counted, and viability was assessed using trypan blue. Then, cells were washed three times with DPBS supplemented with 0.1% bovine serum albumin (BSA). Cells were resuspended in DPBS-BSA and stained with cell markers including viability dye (Ghost Red 780), CD14-APC-Cy7, CD3-FITC, CD19-BV711, and IgG-PE-Cy5. Additionally, antigen-oligo conjugates were added to the stain. For experiment 1, oligo-labeled SARS-CoV-2 S and three-fold molar excess of oligo-labeled ACE2 was added. For experiment 2, six aliquots of S protein that were each labeled with a unique DNA oligonucleotide were added in a titration series from 5 μg to 0.0016 μg (in 5-fold dilutions). For experiment 3, the same titration series of S was added along with three fold molar excess of ACE2. For experiment 4, SARS-CoV-2 S, SARS-CoV S and three-fold molar excess of oligo-labeled ACE2 was added. The antigen screening library for

each of the four experiments also included an influenza virus hemagglutinin and an HIV-1 envelope variant protein as controls.

After staining in the dark for 30 minutes at room temperature, cells were washed three times with DPBS-BSA at 300 x g for five minutes. Cells were then incubated for 15 minutes at room temperature with Streptavidin-PE to label cells with bound antigen. Cells were washed three times with DPBS-BSA, resuspended in DPBS, and sorted by FACS. Antigen positive cells were bulk sorted and delivered to the Vanderbilt Technologies for Advanced Genomics (VANTAGE) sequencing core at an appropriate target concentration for 10X Genomics library preparation and subsequent sequencing. Flow cytometry data were analyzed using FlowJo.

Sample Preparation, Library Preparation, and Sequencing

Single-cell suspensions were loaded onto the Chromium Controller microfluidics device (10X Genomics) and processed using the B-cell Single Cell V(D)J solution according to manufacturer's suggestions for a target capture of 10,000-20,000 B cells, with minor modifications to intercept, amplify and purify the antigen barcode libraries[116]. The 10X Genomics single cell VDJ human B cell assay and target enrichment protocol were completed. cDNA was amplified and additive primers were added to increase the yield of antigen derived transcript products. After cDNA amplification, the antigen derived transcript products were size separated from the mRNA-derived cDNA products using SPRI selection and further purification (per manufacturers protocol). The supernatant fraction contained the antigen-oligo derived cDNA whereas the beads fraction contained the full-length mRNA-derived cDNAs. After purification, the antigen-

derived transcripts sequencing library was prepared using a PCR reaction and purified using SPRI purification. The antigen and VDJ libraries were then analyzed, quantified, and sequenced using the Illumina NovaSeq platform.

Sequence Processing and Bioinformatic Analysis

We used our previously described pipeline to use paired-end FASTQ files of oligo libraries as input, process and annotate reads for cell barcode, UMI, and antigen barcode, and generate a cell barcode - antigen barcode UMI count matrix[206]. BCR contigs were processed using Cell Ranger (10X Genomics) using GRCh38 as reference. Antigen barcode libraries were also processed using Cell Ranger (10X Genomics). The overlapping cell barcodes between the two libraries were used as the basis of the subsequent analysis. We removed cell barcodes that had only non-functional heavy chain sequences as well as cells with multiple functional heavy chain sequences and/or multiple functional light chain sequences, reasoning that these may be multiplets. Additionally, we aligned the BCR contigs (filtered_contigs.fasta file output by Cell Ranger, 10X Genomics) to IMGT reference genes using HighV-Quest[139]. The output of HighV-Quest was parsed using ChangeO and merged with an antigen barcode UMI count matrix[140]. Finally, for experiments 1-3, we determined the LIBRA-seq score for each antigen in the library by calculating the centered-log ratios (CLR) of each antigen UMI count for each cell. A psedo-count of 1 was added to each UMI count and then the CLR was taken for each antigen for each cell. For experiment 4, the LIBRA-seq scores were calculated as previously described[116]. Briefly, the CLR of each

antigen UMI count for each cell was calculated and a Z-score transformation was also performed.

Antibody Expression and Purification

For each antibody, variable genes were inserted into custom plasmids encoding the constant region for the IgG1 heavy chain as well as respective lambda and kappa light chains (pTwist CMV BetaGlobin WPRE Neo vector, Twist Bioscience). Antibodies were expressed in Expi293F mammalian cells (Thermo Fisher Scientific) by co-transfecting heavy chain and light chain expressing plasmids using polyethylenimine transfection reagent and cultured for 5 to 7 days. Cells were maintained in FreeStyle F17 expression medium supplemented at final concentrations of 0.1% Pluronic Acid F-68 and 20% 4 mM L-Glutamine. These cells were cultured at 37°C with 8% CO_2 saturation and shaking. After transfection and 5-7 days of culture, cell cultures were centrifuged and supernatant was 0.45 μm filtered with Nalgene Rapid Flow Disposable Filter Units with PES membrane. Filtered supernatant was run over a column containing Protein A agarose resin equilibrated with PBS. The column was washed with PBS, and then antibodies were eluted with 100 mM Glycine HCl at 2.7 pH directly into a 1:10 volume of 1M Tris-HCl pH 8.0. Eluted antibodies were buffer exchanged into PBS 3 times using Amicon Ultra-centrifugal filter units and concentrated. Antibody plasmids were sequenced. If antibody sequences did not match expected heavy or light chain, antibody was excluded from downstream analysis.

High-throughput Antibody Expression

For high-throughput production of recombinant antibodies, approaches were used that are designated as microscale. For antibody expression, microscale transfection was performed (~1 mL per antibody) of CHO cell cultures using the Gibco ExpiCHO Expression System and a protocol for deep 96-well blocks (Thermo Fisher Scientific). In brief, synthesized antibody-encoding DNA (~2 µg per transfection) was added to OptiPro serum free medium (OptiPro SFM), incubated with ExpiFectamine CHO Reagent and added to 800 µL of ExpiCHO cell cultures into 96-deep-well blocks using a ViaFlo 384 liquid handler (Integra Biosciences). The plates were incubated on an orbital shaker at 1,000 r.p.m. with an orbital diameter of 3 mm at 37°C in 8% CO_2. The next day after transfection, ExpiFectamine CHO Enhancer and ExpiCHO Feed reagents (Thermo Fisher Scientific) were added to the cells, followed by 4 d incubation for a total of 5 d at 37°C in 8% CO_2. Culture supernatants were collected after centrifuging the blocks at 450 x *g* for 5 min and were stored at 4°C until use. For high-throughput microscale antibody purification, fritted deep-well plates were used containing 25 µL of settled protein G resin (GE Healthcare Life Sciences) per well. Clarified culture supernatants were incubated with protein G resin for antibody capturing, washed with PBS using a 96-well plate manifold base (Qiagen) connected to the vacuum and eluted into 96-well PCR plates using 86 µL of 0.1 M glycine-HCL buffer pH 2.7. After neutralization with 14 µL of 1 M Tris-HCl pH 8.0, purified antibodies were buffer-exchanged into PBS using Zeba Spin Desalting Plates (Thermo Fisher Scientific) and stored at 4°C until use.

ELISA

To assess antibody binding, soluble protein was plated at 2 μg/mL overnight at 4°C. The next day, plates were washed three times with PBS supplemented with 0.05% Tween-20 (PBS-T) and coated with 5% milk powder in PBS-T. Plates were incubated for one hour at room temperature and then washed three times with PBS-T. Primary antibodies were diluted in 1% milk in PBS-T, starting at 10 μg/mL with a serial 1:5 dilution and then added to the plate. The plates were incubated at room temperature for one hour and then washed three times in PBS-T. The secondary antibody, goat anti-human IgG conjugated to peroxidase, was added at 1:10,000 dilution in 1% milk in PBS-T to the plates, which were incubated for one hour at room temperature. Plates were washed three times with PBS-T and then developed by adding TMB substrate to each well. The plates were incubated at room temperature for ten minutes, and then 1N sulfuric acid was added to stop the reaction. Plates were read at 450 nm.
Data are represented as mean ± SEM for one ELISA experiment. ELISAs were repeated 2 or more times. If ELISA replicates were inconsistent over more than three experiments, antibody was excluded from *in vitro* characterization analysis. The area under the curve (AUC) was calculated using Prism software version 8.0.0 (GraphPad).

ACE2 Binding Inhibition Assay

96-well plates were coated with 2 μg/mL purified recombinant SARS-CoV-2 at 4°C overnight. The next day, plates were washed three times with PBS supplemented with 0.05% Tween-20 (PBS-T) and coated with 5% milk powder in PBS-T. Plates were incubated for one hour at room temperature and then washed three times with PBS-T.

Purified anti were diluted in blocking buffer at 10 µg/mL in triplicate, added to the wells, and incubated at room temperature. Without washing, recombinant human ACE2 protein with a mouse Fc tag was added to wells for a final 0.4 µg/mL concentration of ACE2 and incubated for 40 minutes at room temperature. Plates were washed three times with PBS-T, and bound ACE2 was detected using HRP-conjugated anti-mouse Fc antibody and TMB substrate. The plates were incubated at room temperature for ten minutes, and then 1N sulfuric acid was added to stop the reaction. Plates were read at 450 nm. ACE2 binding without antibody served as a control. Experiment was done in biological replicate and technical triplicates.

BioLayer Interferometry (BLI)

Purified antibodies were immobilized to AHC sensortips (FortéBio) to a response level of approximately 1.4 nm in a buffer composed of 10 mM HEPES pH 7.5, 150 mM NaCl, 3 mM EDTA, 0.05% Tween 20 and 0.1% (w/v) BSA. Immobilized antibodies were then dipped into wells containing two-fold dilutions of either SARS-CoV-2 RBD-SD1 (residues 306–577) or SARS-CoV-2 NTD, ranging in concentration from 10–0.156 nM, to measure association kinetics. Dissociation kinetics were measured by dipping sensortips into wells containing only buffer. Data were reference subtracted and kinetics were calculated in Octet Data Analysis software v10.0 using a 1:1 binding model.

RTCA Method for Initial Screening of Antibody Neutralizing Activity

To screen for neutralizing activity in the panel of recombinantly expressed antibodies, we used a high-throughput and quantitative RTCA assay and xCelligence

RTCA HT Analyzer (ACEA Biosciences) that assesses kinetic changes in cell physiology, including virus-induced cytopathic effect (CPE). Twenty µL of cell culture medium (DMEM supplemented with 2% FBS) was added to each well of a 384-well E-plate using a ViaFlo384 liquid handler (Integra Biosciences) to obtain background reading. Six thousand (6,000) Vero-furin cells in 20 µL of cell culture medium were seeded per well, and the plate was placed on the analyzer. Sensograms were visualized using RTCA HT software version 1.0.1 (ACEA Biosciences). For a screening neutralization assay, equal amounts of virus were mixed with micro-scale purified antibodies in a total volume of 40 µL using DMEM supplemented with 2% FBS as a diluent and incubated for 1 h at 37°C in 5% CO_2. At ~17–20 h after seeding the cells, the virus–antibody mixtures were added to the cells in 384-well E-plates. Wells containing virus only (in the absence of antibody) and wells containing only Vero cells in medium were included as controls. Plates were measured every 8–12 h for 48–72 h to assess virus neutralization. Micro-scale antibodies were assessed in four 5-fold dilutions (starting from a 1:20 sample dilution), and their concentrations were not normalized. Neutralization was calculated as the percent of maximal cell index in control wells without virus minus cell index in control (virus-only) wells that exhibited maximal CPE at 40–48 h after applying virus–antibody mixture to the cells. An antibody was classified as fully neutralizing if it completely inhibited SARS-CoV-2-induced CPE at the highest tested concentration, while an antibody was classified as partially neutralizing if it delayed but did not fully prevent CPE at the highest tested concentration[153, 207].

Real-time Cell Analysis (RTCA) Neutralization Assay

To determine neutralizing activity of IgG, we used real-time cell analysis (RTCA) assay on an xCELLigence RTCA MP Analyzer (ACEA Biosciences Inc.) that measures virus-induced cytopathic effect (CPE)[208]. Briefly, 50 µL of cell culture medium (DMEM supplemented with 2% FBS) was added to each well of a 96-well E-plate using a ViaFlo384 liquid handler (Integra Biosciences) to obtain background reading. A suspension of 18,000 Vero-E6 cells in 50 µL of cell culture medium was seeded in each well, and the plate was placed on the analyzer. Measurements were taken automatically every 15 min, and the sensograms were visualized using RTCA software version 2.1.0 (ACEA Biosciences Inc). VSV-SARS-CoV-2 (0.01 MOI, ~120 PFU per well) was mixed 1:1 with a dilution of antibody in a total volume of 100 µL using DMEM supplemented with 2% FBS as a diluent and incubated for 1 h at 37°C in 5% CO_2. At 16 h after seeding the cells, the virus-antibody mixtures were added in replicates to the cells in 96-well E-plates. Triplicate wells containing virus only (maximal CPE in the absence of antibody) and wells containing only Vero cells in medium (no-CPE wells) were included as controls. Plates were measured continuously (every 15 min) for 48 h to assess virus neutralization. Normalized cellular index (CI) values at the endpoint (48 h after incubation with the virus) were determined using the RTCA software version 2.1.0 (ACEA Biosciences Inc.). Results are expressed as percent neutralization in a presence of respective antibody relative to control wells with no CPE minus CI values from control wells with maximum CPE. RTCA IC_{50} values were determined by nonlinear regression analysis using Prism software.

Plaque Reduction Neutralization Test (PRNT)

The virus neutralization with live authentic SARS-CoV-2 virus (USA-WA1) was performed in the BSL-3 facility of the Galveston National Laboratory using Vero E6 cells (ATCC CRL-1586) following the standard procedure. Vero E6 cells were cultured in 96-well plates (10^4 cells/well). Next day, 4-fold serial dilutions of antibodies were made using MEM-2% FBS, as to get an initial concentration of 100 µg/mL. Equal volume of diluted antibodies (60 µL) were mixed gently with original SARS-CoV-2 (USA-WA1) (60 µL containing 200 pfu) and incubated for 1 h at 37°C/5% CO_2 atmosphere. The virus-serum mixture (100 µL) was added to cell monolayer in duplicates and incubated for 1 at h 37°C/5% CO_2 atmosphere. Later, the virus-serum mixture was discarded gently, and cell monolayer was overlaid with 0.6% methylcellulose and incubated for 2 days. The overlay was removed, and the plates were fixed in 4% paraformaldehyde twice following BSL-3 protocol. The plates were stained with 1% crystal violet and virus-induced plaques were counted. The percent neutralization and/or NT_{50} of antibody was calculated by dividing the plaques counted at each dilution with plaques of virus-only control. For antibodies, the inhibitory concentration at 50% (IC_{50}) values were calculated in Prism software (GraphPad) by plotting the midway point between the upper and lower plateaus of the neutralization curve among dilutions. The Alpha variant virus incorporates the following substitutions: Del 69-70, Del 144, E484K, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H. The Beta variant incorporates the following substitutions: Del 24, Del 242-243, D80A, D215G, K417N, E484K, N501Y, D614G, H665Y, T1027I. The Gamma variant incorporates the following substitutions: L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, D614G, H655Y, T1027I. The

Delta variant incorporates the following substitutions: T19R, G142D, Del 156-157, R158G, L452R, T478K, D614G, P681R, Del 689-691, D950N; the deletion at positions 689-691 has not been observed in nature, and was identified upon one passage of the virus.

Fab Preparation

To generate Fabs, IgGs were incubated with Lys-C at 1:4,000 (weight:weight) overnight at 37 °C. EDTA free protease inhibitor (Roche) was dissolved to 25X and then added to the sample at a final 1X concentration. The sample was passed over a Protein A column. The flow-through was collected run on a Superdex 200 Increase 10/300 GL Sizing column on the AKTA FPLC system. Fabs were visualized on SDS-PAGE.

Biolayer Interferometry

Purified mAbs were immobilized to AHC sensortips (FortéBio) to a response level of approximately 1.4 nm in a buffer composed of 10 mM HEPES pH 7.5, 150 mM NaCl, 3 mM EDTA, 0.05% Tween 20 and 0.1% (w/v) BSA. Immobilized mAbs were then dipped into wells containing two-fold dilutions of either SARS-CoV-2 RBD-SD1 or SARS-CoV-2 NTD, ranging in concentration from 10-0.15625 nM, to measure association. Dissociation was measured by dipping sensortips into wells containing only running buffer. Data were reference subtracted and kinetics were calculated in Octet Data Analysis software v10.0 using a 1:1 binding model.

Electron Microscopy Sample Preparation and Data Collection

Purified SARS-CoV-2 S HexaPro ectodomain[209] and Fabs 5317-4 and 5317-10 were combined at a final complex concentration of 0.4 mg/mL. Fab 5317-10 was added to spike and incubated on ice for 30 minutes before the addition of Fab 5317-4 immediately prior to grid deposition and freezing. The complex was deposited on Au-300 1.2/1.3 grids that had been plasma cleaned for 4 minutes in a Solarus 950 plasma cleaner (Gatan) with a 4:1 ratio of O_2/H_2. Excess liquid was blotted for 3 seconds with a force of -4 using a Vitrobot Mark IV (Thermo Fisher) and plunge frozen into liquid ethane. 2,655 micrographs were collected from a single grid with the stage at a 30° tilt using a Titan Krios (Thermo Fisher) equipped with a K3 detector (Gatan). Movies were collected using SerialEM[210] at 29,000X magnification with a corresponding calibrated pixel size of 0.81 Å/ pixel.

Cryogenic Electron Microscopy (Cryo-EM)

Motion correction, CTF estimation, particle picking, and 2D classification were performed using cryoSPARC v3.2.0[211] The final iteration of 2D class averaging distributed 17,710 particles into 50 classes using an uncertainty factor of 3. From that, 13,232 particles were selected and an ab inito reconstruction was performed with four classes followed by heterogeneous refinement of those four classes. 6,803 particles from the highest-quality class were used for homogenous refinement of the best volume without imposed symmetry. The resulting volume was used for an additional round of homogenous refinement. To filter out additional junk particles, an *ab initio* reconstruction was performed with three classes followed by heterogeneous refinement of those three

classes. 5,171 particles from the highest-quality class were used for homogenous refinement of the best volume without imposed symmetry, resulting in a final 9 Å map.

Quantification and Statistical Analysis

ELISA error bars (standard error of the mean) were calculated using GraphPad Prism version 8.0.0. Spearman r correlation was performed using GraphPad Prism 8.0.0. ANOVA analysis was performed for neutralization potency comparisons using GraphPad Prism version 8.0.0.

CHAPTER 5

Conclusions and Future Directions

5.1 Summary and Conclusions

Infectious diseases have continued to plague humankind; however, the advent of vaccines and therapies against these pathogens have helped to reduce their burden on society. However, some pathogens, like HIV-1, elusively evade immune system pressures and have proven difficult to fight with traditional methods. Other pathogens pose significant threats, and their emergence could lead to worldwide pandemics, as shown with the introduction of SARS-CoV-2 and the COVID-19 pandemic which began in 2019 yet continues to devastate healthcare systems in 2022. Development and application of antibody discovery tools are integral to improving our ability to rapidly and efficiently characterize the antigen-specific immune repertoire. Antibody discovery can lead to the development of antibodies as therapeutic molecules and can aid in vaccine design strategies. Furthermore, identification and characterization of antigen-specific antibodies, including their B cell receptor sequence, contributes to basic foundational understanding of the immune system and functional antibody development.

In this body of work, I developed a novel antibody discovery technology termed LIBRA-seq (Linking B cell receptor to antigen specificity through sequencing). First, I tested and validated the ability to create antigen-oligo conjugates for use in the experimental assay using HIV-1 and influenza recombinant proteins. I applied proof-of-

principle experiments with these DNA-barcoded antigens and B cells lines with known sequences and specificities. These experiments demonstrated that LIBRA-seq could identify and distinguish distinct antigen specificities for different B cell populations. Furthermore, I created and used different panels of DNA-barcoded HIV-1 and influenza antigens as screening libraries against a variety of HIV-1 positive human samples. In these experiments, I created high resolution antigen specificity maps for thousands of B cells. This allowed me to interrogate the antigen reactivity patterns and sequence features of these B cells in order to rapidly identify broadly neutralizing HIV-1 antibodies from both known and novel lineages. Ultimately, this work demonstrated that LIBRA-seq is a powerful tool for immune repertoire screening and characterization, with many advantages to other antibody discovery technologies, including the ability to create large antigen screening libraries, recovery of paired heavy/light chain sequence information at the screening step, and the rapid speed of the experiment, among others.

Second, after validating LIBRA-seq experimentally and demonstrating its potential as an antibody discovery tool for HIV-1 and influenza, I used LIBRA-seq for antibody discovery and characterization of antigen-specific antibodies to diverse infectious disease pathogens. With the advent of the COVID-19 pandemic, we decided to use LIBRA-seq to study the immune response to coronaviruses with a goal of identifying cross-reactive antibodies. The recurrent nature of novel coronavirus strains, like SARS-CoV-2, highlights the critical need for broadly reactive therapeutics and vaccines. Despite substantial amino acid sequence divergence, coronavirus spike proteins share common structural motifs that may serve as conserved sites that are vulnerable to a cross-reactive antibody response. To study this phenomenon in human

coronavirus infection, I worked to apply LIBRA-seq to a SARS-CoV convalescent donor sample. After bioinformatically processing the data, I identified a set of prioritized antibody candidates for expression and testing, of which six demonstrated cross-reactive antigen binding to both SARS-CoV-2 and SARS-CoV. These antibodies targeted a variety of epitopes on the coronavirus spike protein and mediated Fc effector functions. The elucidation of cross-reactive epitopes has implications for vaccine design strategies for the current pandemic and also potential future zoonotic spillovers.

In addition to using LIBRA-seq to study the coronavirus antibody response, I also prioritized a variety of antibody candidates from existing data sets for deeper characterization. I demonstrate that antibody 3602-1707 was a broadly reactive influenza antibody that targets the hemagglutinin stem and is able to neutralize diverse group 1 viruses. This antibody shares a similar mode of hemagglutinin recognition as a group 2 antibody called CR8020. Gaining a more detailed picture of how 3602-1707 targets a similar epitope will shed light on group 1 vs group 2 antigen reactivity, ultimately helping to inform universal flu vaccine design strategies. Furthermore, this antibody demonstrated efficacy against influenza in animal models in preliminary experiments, suggesting potential for development as a therapeutic candidate. I also identified a cross-reactive antibody family that targets both HIV-1 and influenza viral proteins. I helped characterize this functional antibody family in order to probe our fundamental knowledge of antibody specificity and reactivity. Learning more about cross-reactive, yet specific, antibodies could highlight gaps in our knowledge about antibody development in response to the many pathogens humans encounter in their lifetime.

Third, I extended the LIBRA-seq technology to incorporate a functional element to the sequencing readout. By adding a barcoded antigen and a barcoded ligand to the LIBRA-seq experiment, I was able to interrogate antibody-ligand blocking in a high throughput way. In addition to gaining paired heavy/light chain sequence information and antigen reactivity information, LIBRA-seq with ligand blocking provides functional information for hundreds to thousands of B cells simultaneously. To test this, I applied an antigen screening library of SARS-CoV-2 spike and the ACE2 receptor to PBMCs from convalescent COVID-19 donor samples. LIBRA-seq with ligand blocking demonstrated high rates of neutralizing antibody identification. In these experiments, I also identified a neutralizing antibody, 5317-10, which targets a novel epitope that bridges the receptor binding domain and the N-terminal domain of the SARS-CoV-2 spike protein. I also performed an experiment to identify cross-reactive antibodies that block receptor binding by including SARS-CoV-2 and SARS-CoV in the screening library with ACE2. This experiment also demonstrated the effectiveness of LIBRA-seq with ligand blocking and showed high hit rates for cross-reactive ACE2 blocking antibody identification. Overall, some of the antibodies identified in these datasets could be used as therapeutic options and should be tested as new SARS-CoV-2 variants continue to emerge. Furthermore, their sequence features and epitope reactivities can be used as vaccines are iteratively improved and as pan-coronavirus vaccine designs develop. The LIBRA-seq with ligand blocking technology shows immense promise as a tool for rapid characterization of antigen-ligand interactions in a high throughput way for a variety of areas, including infectious disease and cancer research, among others.

5.2 Caveats

Though LIBRA-seq is an immensely powerful tool for characterization of antigen-specific antibody repertoires, there are some caveats and limitations to the technology and the work described in this dissertation. First, with LIBRA-seq for antibody discovery as described, the antigenic target must be known. This iteration of LIBRA-seq does not allow for prospective antigen discovery. Similarly, for LIBRA-seq with ligand blocking, the antigen-ligand interaction must be known. Additionally, in this dissertation, antibody discovery and characterization were performed with LIBRA-seq using recombinant protein antigens (not other antigenic material such as virus etc.) and human cells (not other species). These limitations could be explored for future iterations of the LIBRA-seq technology; however, as it currently stands, they exemplify potential limitations on the kind of experiments that can be successfully implemented with the present state of the technology.

Additionally, the quality of the proteins used for DNA-barcoded antigen conjugates is of utmost importance for deconvolution of the sequencing data and identifying true positive hits. For example, in Chapter 3, I identified coronavirus cross-reactive antibodies; however, we could not conclusively confirm MERS reactivity due to the inconsistent quality of the MERS protein used in our screening library and for binding characterization.

Furthermore, there are likely experimental restraints on the number of antigens used in a screening library, and we have not empirically determined where this limit is. Currently, we have tested antigen libraries with up to 20-30 antigens, but it is unclear

exactly where the limit lies. Overall, LIBRA-seq is a powerful tool and should be further explored to concretely determine quantitative limits to its applications in the future.

5.3 Future Directions

5.3.1 Further Characterization of Influenza Antibody 3602-1707

3602-1707 is a broadly neutralizing influenza antibody that targets the stem region of the hemagglutinin trimer protein. This antibody has a similar epitope and angle of approach as group 2 reactive antibody CR8020[212]. To understand how the antibody accommodates these features, we are examining the antibody-antigen interaction at atomic level resolution using cryogenic electron microscopy. Identifying key residues (and glycan accommodations) for this interaction on the antibody and antigen will allow us to gain a better understanding of the molecular mechanism of breadth. Furthermore, we can use this information to computationally engineer the antibody with an aim to increase its breadth to accommodate group 1 and group 2 neutralization potential. These experiments can aid in universal flu vaccine design strategies. Furthermore, we will perform additional animal studies in an influenza challenge model. In these experiments, I propose to test the antibody prophylactically and in a therapeutic setting at a variety of concentrations. This will allow us to get a better understanding of the antibody's efficacy and potential as a therapeutic candidate, either alone or in combination with a group 2 broadly neutralizing antibody.

5.3.2 Further Characterization of HIV/influenza Cross-reactive Antibodies

The identification of functionally relevant HIV/influenza cross-reactive antibodies shifts the traditional understanding of the unique specificity of antibodies for one antigen. Gaining insight into the development of these antibodies can highlight the mechanism of cross-reactivity. To do this, I propose creating germline-reverted antibodies and testing them for binding and Fc effector functionality for both HIV-1 and influenza. This will help determine if the antibodies were inherently cross-reactive or if they were mono-reactive and then gained cross-reactivity through somatic hypermutation. To further understand this mechanism, we could perform deep sequencing on another aliquot of the same sample to identify other antibody relatives. These relatives could be expressed and tested in order to gain further understanding about the development of cross-reactivity. Another important next step will be deepening our characterization of the antigenic epitope of these antibodies for both HIV-1 antigens and influenza antigens. To do this, I propose glycan reactivity studies and structural studies, including negative stain electron microscopy, to interrogate the antibody-antigen interaction and elucidate the antibody footprint on each antigen. Epitope determination will give more insight into the cross-reactivity of these antibodies and their biological significance. Other future directions include identifying other putative cross-reactive antibody candidates by examining existing datasets and performing additional LIBRA-seq experiments with HIV-1 and influenza antigens. For example, in the N90 dataset, there are additional antibodies that demonstrate cross-reactive patterns similar to what we observed for 3602-2526 and 3602-546. These antibodies

could be expressed and tested. Learning more about the genetic features, antigen recognition, and functional relevance of cross-reactive antibodies will impact our understanding of antibody development and specificity.

5.3.3 Pan-Coronavirus Antibody Discovery and Pandemic Preparedness

Pan coronavirus vaccine design efforts and pan coronavirus antibody therapeutics are of great interest to the field as there could be potential future zoonotic spillovers[90, 99, 160, 213, 214]. Identifying antibodies that are not only able to recognize all coronavirus strains but are functional with ligand blocking capability would enable application of reverse vaccinology to coronavirus vaccine design. Furthermore, pan-coronavirus antibodies could be used as therapeutic options in the case of future spillover events[215]. To identify these antibodies, I propose assembling an antigen library of spike proteins from fifty diverse coronavirus strains, including the 7 human coronaviruses, and other animal strains. If receptors are known for the strains, then they would also be used in the screening library for LIBRA-seq with ligand blocking. I propose to apply this library of antigens to 10 groups of 10 samples. By screening 100 donors with different exposure histories, there is a greater chance of identifying an extremely unique, broadly reactive antibody. This project could allow for pan-reactive antibodies with ligand blocking capability to be rapidly identified and studied. Additionally, this idea could also be applied to other viruses with diverse strains, including HIV-1, influenza virus, Ebola virus, and paramyxoviruses, for example.

Similarly, many other pathogens are deemed to be pandemic threats[90, 207, 216, 217]. With a goal of stockpiling antibodies to these pathogens for pandemic preparedness, the first step is identifying and characterizing lead candidates. LIBRA-seq could be used to identify antibodies for these pathogens. By labeling viral proteins from these pathogens with oligos and screening against human samples, antibody identification could be utilized for therapy and vaccine design. For example, for Ebola virus disease, the Ebola viral glycoprotein could be labeled. This antigen could be used in a simple experiment for discovery of anti-Ebola antibodies, and this strategy could be applied for other pathogens to identify antibody candidates. However, a large library of one hundred antigens in the pandemic threat category could also be generated. This full library could be applied to a variety of samples, including known infection or vaccination samples and also naïve samples. Application to infection or vaccination samples would give higher probability of antibody discovery; however, application to naive samples could also yield virus-specific antibodies and also give clues for how the immune system might interact with these viruses upon exposure. Furthermore, as described in Chapter 3, our laboratory has begun investigating antibodies that are cross-reactive to diverse viruses, like HIV-1 and influenza virus. Using this panel, antibodies that are cross-reactive to diverse viral families could be identified and used for pandemic preparedness and also enhancement of our basic immunological understanding of antibody specificity.

5.3.4 LIBRA-seq for Evaluation of Germline Targeting Antigens from HIV-Naïve Repertoires and Characterization of Vaccine Responses

To date, there is still no vaccine for HIV-1, although several rationally designed immunogens are in clinical trials. Sparse studies have examined the ability of select immunogens to engage with the HIV-naïve B cell repertoire, but the majority of current lead immunogens in the field have not been screened for reactivity with the HIV-naïve repertoire[14, 42, 44, 135, 136, 218]. It is understood that HIV-reactive B cells within naïve individuals can engage with some vaccine immunogens; however, the specificity and frequency of these B cells have not been characterized extensively. Further, there have not been large-scale examinations of how antibodies from HIV-naïve repertoires react with a variety of HIV-1 immunogens simultaneously, and analyses of HIV-1 vaccine trials have not focused on the influence of pre-vaccination antibody repertoires.

One category of HIV-1 vaccine candidates includes immunogens designed to first engage with germline precursors of known bNAbs, followed by strategically guided affinity maturation towards the development of broad neutralization. Previous studies have examined the frequency of B cells from HIV-naïve samples that can bind to some of these antigen variants; however, the use of a panel of diverse antigens to determine immunogen reactivity within the HIV-naïve repertoire has not been done previously. Further, studies using a large panel of bNAb-class targeting immunogens to probe the HIV-naïve repertoire simultaneously have not been done.

I propose to use an antigen screening library of vaccine candidates, including HIV-1 SOSIP trimer antigens, gp120 proteins, and germline-targeting immunogens including

VRC01-class targeting antigens (eOD-gt8, 426c trimer, 426c gp120 core), PGT121-class targeting antigen (BG505 10-mut trimer), V2-apex targeting antigens (459c trimer, CRF02_AG_250 trimer), and a non-HIV HA negative probe[124, 134, 135, 219-222]. This screening library can be applied to multiple naïve samples. This work would determine the ability of a variety of HIV-1 vaccine candidates, including bNAb class targeting immunogens, to engage with the naïve repertoire. Understanding the frequency of B cells that can engage with each immunogen can guide selection of immunogens for further testing and advancement to clinical trials. Additionally, screening many naïve samples will allow evaluation of the ability of multiple individuals to engage with each immunogen.

Characterization of the properties of B cells from HIV-naïve samples that can engage with vaccine candidates will illustrate the potential of HIV-specific antibody lineages to evolve from the HIV-naïve repertoire and inform the selection of immunogens for further clinical development, helping to guide immunization strategies that can select for B cell progenitors that will develop neutralizing properties (vs. non-neutralizing B cells). The sequence properties of these B cell receptors along with their functional relevance will be helpful in determination of optimal vaccine candidates and their ability to elicit protective and therapeutic antibodies in humans. Ultimately, this information can be used to predict the ability of vaccine candidates to engage HIV-naïve B cells and therefore help prioritization of immunogen selection. Examination of antibody repertoires from HIV-naïve individuals can lead to increased understanding of the evolution of HIV-specific antibody responses and whether or not they can be protective.

Furthermore, this strategy could be used for any vaccine in evaluation. Additionally, in pre-clinical trials, evaluation of the immune response to vaccine

candidates can be used by combining LIBRA-seq with animal studies. Specific primers for animal B cell receptor repertoires can be combined with single cell sequencing platforms to analyze the paired heavy/light chain sequences along with LIBRA-seq antigen reactivity information. This could be done for a variety of animal models, including mice and non-human primates, for example. Overall, rapidly and efficiently characterizing the B cell response to vaccine components can be used for evaluation of pre-clinical vaccine candidates and vaccines in clinical trials.

5.3.5 Antibody Discovery for Neurodegenerative Disorders

Alzheimer's disease affects millions of people worldwide and is a leading cause of death in the United States[223]. This disease leads to memory loss, decline in cognitive function, and affects the ability of people to carry out daily, basic functions. There are limited treatments available that slow progression of the disease – but there is no cure. Some monoclonal antibodies have been shown to target beta-amyloid proteins to prevent and/or reduce plaque formation in the brain; however, many of these antibodies have not been effective[224, 225]. Despite this discouraging data, there could be other antibodies that could help clear amyloid plaques by targeting alternative epitopes or be given to patients at an earlier stage of disease progression. I propose using LIBRA-seq to identify antibodies against the beta-amyloid proteins in two ways. First, I propose oligo-labeling beta-amyloid protein and screening 100 donor samples. Second, I propose vaccinating mice with the amyloid protein and then performing LIBRA-seq on the mouse B cells for antibody identification. Gaining a large set of antibodies to

characterize and test would improve chances of identifying a target candidate to pursue clinically.

5.3.6 Determination of Antibody Potency through Sequencing

In Chapter 4, Next-generation LIBRA-seq for Antibody Discovery: Efficient Discovery of Potent Neutralizing SARS-CoV-2 Antibodies using LIBRA-seq with Ligand Blocking, experiments for the determination of antibody potency through sequencing were described. In these experiments, multiple aliquots of SARS-CoV-2 spike (S) protein were added in a titration series, with a goal of utilizing LIBRA-seq antigen reactivity information to identify antibodies with high affinity for the SARS-CoV-2 spike protein. I hypothesized that antibodies that showed reactivity to the spike added in low amounts would be high affinity. I propose additional experiments to further explore the ability to significantly associate antibody potency determined through sequencing with affinity. To do this, I would label 10 different aliquots of the receptor binding domain of SARS-CoV-2 with DNA barcodes. These would then be added in a titration series to a cell population split into 10 aliquots. Performing single experiments each with a different concentration of the barcoded RBD would allow for determination of the optimal range of detection by sequencing for the assay, from over-saturation down to the limit of detection. After determining the optimal range of RBD to use, I would then add a titration series of 10 different amounts of barcoded spike. By doing this, we can generate a binding curve from the sequencing data and unique molecular identifier counts for each antigen added. After sequencing, I would try a variety of methods for

assessing antibody potency from the sequencing data, including determination of EC50 values and area under the curve values. I would then express 100 antibodies with a variety of patterns RBD reactivity and determine their affinity for the RBD using biolayer interferometry. The experimentally derived affinities can then be associated with the sequencing derived potency measurements to see if there is a significant comparison.

5.3.7 Pairing Antibody Sequence to Antigen Specificity

Datasets generated by the LIBRA-seq technology provide a wealth of information, including antigen reactivity information and antibody sequence. These sequences include the antibody heavy chain and antibody light chain. I propose utilizing the sequences, specific sequence features (such as somatic hypermutation levels and CDR3 length), and heavy/light chain pairing information along with the antigen specificity information to create a dataset that can be used to train a machine learning algorithm for prediction of antigen reactivity from antibody sequence. This supervised learning approach would start with processing and cleaning the dataset, and partitioning the LIBRA-seq dataset into a training dataset for the algorithm and a testing dataset. Next, a model will be selected and fit to the training data (for example, a hidden Markov Model could be developed and used)[226, 227]. Then, the model will be evaluated on the testing dataset, including determination of true positive rates and false positive rates. After evaluating the model, it can be iteratively improved. Upon development of the algorithm, it can be tested on unrelated datasets and then validated experimentally, by expressing and characterizing antibodies *in vitro*. In this way, antibody sequence can be prospectively linked to antigen specificity from datasets where antigen reactivity is

unknown. This can be done systematically, by starting small, with LIBRA-seq data using a small antigen panel. Once successfully implemented, additional antigens can be used to expand the algorithm and antigen reactivities that can be predicted. This kind of analysis can be applied to discriminate antigen reactivity of antibody repertoires from a variety of patient samples, including healthy controls, diseased individuals, and vaccine samples, for example.

References

1. Chaplin, D.D. Overview of the immune response. *J Allergy Clin Immunol* **125**, S3-23 (2010).

2. Flajnik, M.F. & Kasahara, M. Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nat Rev Genet* **11**, 47-59 (2010).

3. Medzhitov, R. & Janeway, C.A., Jr. Decoding the patterns of self and nonself by the innate immune system. *Science* **296**, 298-300 (2002).

4. Dorner, T. & Radbruch, A. Antibodies and B cell memory in viral immunity. *Immunity* **27**, 384-392 (2007).

5. Kurosaki, T., Kometani, K. & Ise, W. Memory B cells. *Nat Rev Immunol* **15**, 149-159 (2015).

6. Weisel, F. & Shlomchik, M. Memory B cells of mice and humans. *Annu Rev Immunol* **35**, 255-284 (2017).

7. Cyster, J.G. & Allen, C.D.C. B cell responses: cell interaction dynamics and decisions. *Cell* **177**, 524-540 (2019).

8. Blattner, F.R. & Tucker, P.W. The molecular biology of immunoglobulin D. *Nature* **307**, 417-422 (1984).

9. McKean, D. et al. Generation of antibody diversity in the immune response of BALB/c mice to influenza virus hemagglutinin. *Proc Natl Acad Sci U S A* **81**, 3180-3184 (1984).

10. Tonegawa, S. Somatic generation of immune diversity. *Biosci Rep* **8**, 3-26 (1988).

11. Rajewsky, K. Clonal selection and learning in the antibody system. *Nature* **381**, 751-758 (1996).

12. Allen, C.D., Okada, T. & Cyster, J.G. Germinal-center organization and cellular dynamics. *Immunity* **27**, 190-202 (2007).

13. Hwang, J.K., Alt, F.W. & Yeap, L.S. Related mechanisms of antibody somatic hypermutation and class switch recombination. *Microbiol Spectr* **3**, MDNA3-0037-2014 (2015).

14. Abbott, R.K. et al. Precursor frequency and affinity determine B cell competitive fitness in germinal centers, tested with germline-targeting HIV vaccine immunogens. *Immunity* **48**, 133-146 e136 (2018).

15. Stewart, I., Radtke, D., Phillips, B., McGowan, S.J. & Bannard, O. Germinal center B cells replace their antigen receptors in dark zones and fail light zone entry when immunoglobulin gene mutations are damaging. *Immunity* **49**, 477-489 e477 (2018).

16. Mayer, C.T. et al. The microanatomic segregation of selection by apoptosis in the germinal center. *Science* **358** (2017).

17. Goodnow, C.C., Vinuesa, C.G., Randall, K.L., Mackay, F. & Brink, R. Control systems and decision making for antibody production. *Nat Immunol* **11**, 681-688 (2010).

18. Liu, W., Wang, H. & Xu, C. Antigen receptor nanoclusters: small units with big functions. *Trends Immunol* **37**, 680-689 (2016).

19. Nutt, S.L., Hodgkin, P.D., Tarlinton, D.M. & Corcoran, L.M. The generation of antibody-secreting plasma cells. *Nat Rev Immunol* **15**, 160-171 (2015).

20. Burnet, F.M. The immunological significance of the thymus: an extension of the clonal selection theory of immunity. *Australas Ann Med* **11**, 79-91 (1962).

21. Burnet, F.M. A modification of Jerne's theory of antibody production using the concept of clonal selection. *CA Cancer J Clin* **26**, 119-121 (1976).

22. Akkaya, M., Kwak, K. & Pierce, S.K. B cell memory: building two walls of protection against pathogens. *Nat Rev Immunol* **20**, 229-238 (2020).

23. Jefferis, R. Nomenclature of V-region markers. *Immunol Today* **16**, 207-208 (1995).

24. Jefferis, R. Rheumatoid factors, B cells and immunoglobulin genes. *Br Med Bull* **51**, 312-331 (1995).

25. Edelman, G.M. Antibody structure and molecular immunology. *Science* **180**, 830-840 (1973).

26. Briney, B., Inderbitzin, A., Joyce, C. & Burton, D.R. Commonality despite exceptional diversity in the baseline human antibody repertoire. *Nature* **566**, 393-397 (2019).

27. Soto, C. et al. High frequency of shared clonotypes in human B cell receptor repertoires. *Nature* **566**, 398-402 (2019).

28. Saphire, E.O. et al. Contrasting IgG structures reveal extreme asymmetry and flexibility. *J Mol Biol* **319**, 9-18 (2002).

29. Kabat, E.A. Antigenic determinants and antibody complementarity. *Folia Allergol (Roma)* **17**, 425 (1970).

30. Early, P., Huang, H., Davis, M., Calame, K. & Hood, L. An immunoglobulin heavy chain variable region gene is generated from three segments of DNA: VH, D and JH. 1980. *J Immunol* **173**, 6503-6514 (2004).

31. Schatz, D.G., Oettinger, M.A. & Baltimore, D. The V(D)J recombination activating gene, RAG-1. *Cell* **59**, 1035-1048 (1989).

32. Davis, M.M. & Bjorkman, P.J. T-cell antigen receptor genes and T-cell recognition. *Nature* **334**, 395-402 (1988).

33. Soto, C. et al. High Frequency of Shared Clonotypes in Human T Cell Receptor Repertoires. *Cell Rep* **32**, 107882 (2020).

34. Bannard, O. & Cyster, J.G. Germinal centers: programmed for affinity maturation and antibody diversification. *Curr Opin Immunol* **45**, 21-30 (2017).

35. Baumgarth, N. How specific is too specific? B-cell responses to viral infections reveal the importance of breadth over depth. *Immunol Rev* **255**, 82-94 (2013).

36. Lu, L.L., Suscovich, T.J., Fortune, S.M. & Alter, G. Beyond binding: antibody effector functions in infectious diseases. *Nat Rev Immunol* **18**, 46-61 (2018).

37. Vidarsson, G., Dekkers, G. & Rispens, T. IgG subclasses and allotypes: from structure to effector functions. *Front Immunol* **5**, 520 (2014).

38. Bournazos, S. et al. Broadly neutralizing anti-HIV-1 antibodies require Fc effector functions for in vivo activity. *Cell* **158**, 1243-1253 (2014).

39. Bournazos, S., DiLillo, D.J., Goff, A.J., Glass, P.J. & Ravetch, J.V. Differential requirements for FcgammaR engagement by protective antibodies against Ebola virus. *Proc Natl Acad Sci U S A* **116**, 20054-20062 (2019).

40. Bonsignori, M. et al. Antibody-virus co-evolution in HIV infection: paths for HIV vaccine development. *Immunol Rev* **275**, 145-160 (2017).

41. Crowe, J.E., Jr. Principles of broad and potent antiviral human antibodies: insights for vaccine design. *Cell Host Microbe* **22**, 193-206 (2017).

42. McGuire, A.T. Targeting broadly neutralizing antibody precursors: a naive approach to vaccine design. *Curr Opin HIV AIDS* **14**, 294-301 (2019).

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

44. Havenar-Daughton, C., Abbott, R.K., Schief, W.R. & Crotty, S. When designing vaccines, consider the starting material: the human B cell repertoire. *Curr Opin Immunol* **53**, 209-216 (2018).

45. Burton, D.R. Antibodies, viruses and vaccines. *Nat Rev Immunol* **2**, 706-713 (2002).

46. UNAIDS (2021).

47. Ndung'u, T. & Weiss, R.A. On HIV diversity. *AIDS* **26**, 1255-1260 (2012).

48. Tomaras, G.D. & Plotkin, S.A. Complex immune correlates of protection in HIV-1 vaccine efficacy trials. *Immunol Rev* **275**, 245-261 (2017).

49. Korber, B. et al. Evolutionary and immunological implications of contemporary HIV-1 variation. *Br Med Bull* **58**, 19-42 (2001).

50. Rerks-Ngarm, S. et al. Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *N Engl J Med* **361**, 2209-2220 (2009).

51. Haynes, B.F. & Burton, D.R. Developing an HIV vaccine. *Science* **355**, 1129-1130 (2017).

52. Haynes, B.F. et al. Immune-correlates analysis of an HIV-1 vaccine efficacy trial. *N Engl J Med* **366**, 1275-1286 (2012).

53. Rappuoli, R., Bottomley, M.J., D'Oro, U., Finco, O. & De Gregorio, E. Reverse vaccinology 2.0: Human immunology instructs vaccine antigen design. *J Exp Med* **213**, 469-481 (2016).

54. Kwong, P.D., Mascola, J.R. & Nabel, G.J. Broadly neutralizing antibodies and the search for an HIV-1 vaccine: the end of the beginning. *Nat Rev Immunol* **13**, 693-701 (2013).

55. Kwong, P.D. & Mascola, J.R. HIV-1 vaccines based on antibody identification, B cell ontogeny, and epitope structure. *Immunity* **48**, 855-871 (2018).

56. Moir, S. & Fauci, A.S. B cells in HIV infection and disease. *Nat Rev Immunol* **9**, 235-245 (2009).

57. Scheid, J.F. et al. Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. *Nature* **458**, 636-640 (2009).

58. Scheid, J.F. et al. Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science* **333**, 1633-1637 (2011).

59. Sok, D. & Burton, D.R. Recent progress in broadly neutralizing antibodies to HIV. *Nat Immunol* **19**, 1179-1188 (2018).

60. Trkola, A. et al. Delay of HIV-1 rebound after cessation of antiretroviral therapy through passive transfer of human neutralizing antibodies. *Nat Med* **11**, 615-622 (2005).

61. Wen, Y., Bar, K.J. & Li, J.Z. Lessons learned from HIV antiretroviral treatment interruption trials. *Curr Opin HIV AIDS* **13**, 416-421 (2018).

62. Gruell, H. & Klein, F. Antibody-mediated prevention and treatment of HIV-1 infection. *Retrovirology* **15**, 73 (2018).

63. Mascola, J.R. et al. Protection of macaques against vaginal transmission of a pathogenic HIV-1/SIV chimeric virus by passive infusion of neutralizing antibodies. *Nat Med* **6**, 207-210 (2000).

64. Wu, X. et al. Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* **329**, 856-861 (2010).

65. Xu, L. et al. Trispecific broadly neutralizing HIV antibodies mediate potent SHIV protection in macaques. *Science* **358**, 85-90 (2017).

66. Lynch, R.M. et al. Virologic effects of broadly neutralizing antibody VRC01 administration during chronic HIV-1 infection. *Sci Transl Med* **7**, 319ra206 (2015).

67. Saunders, K.O. et al. Sustained delivery of a broadly neutralizing antibody in nonhuman primates confers long-term protection against simian/human immunodeficiency virus infection. *J Virol* **89**, 5895-5903 (2015).

68. Pegu, A., Hessell, A.J., Mascola, J.R. & Haigwood, N.L. Use of broadly neutralizing antibodies for HIV-1 prevention. *Immunol Rev* **275**, 296-312 (2017).

69. Burton, D.R. What are the most powerful immunogen design vaccine strategies? Reverse vaccinology 2.0 shows great promise. *Cold Spring Harb Perspect Biol* **9** (2017).

70. Krammer, F. et al. Influenza. *Nat Rev Dis Primers* **4**, 3 (2018).

71. Wu, N.C. & Wilson, I.A. Influenza hemagglutinin structures and antibody recognition. *Cold Spring Harb Perspect Med* **10** (2020).

72. Nobusawa, E. et al. Comparison of complete amino acid sequences and receptor-binding properties among 13 serotypes of hemagglutinins of influenza A viruses. *Virology* **182**, 475-485 (1991).

73. Skehel, J.J. & Wiley, D.C. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu Rev Biochem* **69**, 531-569 (2000).

74. Russell, R.J. et al. H1 and H7 influenza haemagglutinin structures extend a structural classification of haemagglutinin subtypes. *Virology* **325**, 287-296 (2004).

75. Wu, N.C. & Wilson, I.A. A perspective on the structural and functional constraints for immune evasion: insights from influenza virus. *J Mol Biol* **429**, 2694-2709 (2017).

76. Schweiger, B., Zadow, I. & Heckler, R. Antigenic drift and variability of influenza viruses. *Med Microbiol Immunol* **191**, 133-138 (2002).

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

78. Ekiert, D.C. & Wilson, I.A. Broadly neutralizing antibodies against influenza virus and prospects for universal therapies. *Curr Opin Virol* **2**, 134-141 (2012).

79. Krammer, F. & Palese, P. Universal influenza virus vaccines: need for clinical trials. *Nat Immunol* **15**, 3-5 (2014).

80. Yewdell, J.W. To dream the impossible dream: universal influenza vaccination. *Curr Opin Virol* **3**, 316-321 (2013).

81. Zhang, H., Wang, L., Compans, R.W. & Wang, B.Z. Universal influenza vaccines, a dream to be realized soon. *Viruses* **6**, 1974-1991 (2014).

82. Lee, P.S. & Wilson, I.A. Structural characterization of viral epitopes recognized by broadly cross-reactive antibodies. *Curr Top Microbiol Immunol* **386**, 323-341 (2015).

83. Corti, D. et al. Heterosubtypic neutralizing antibodies are produced by individuals immunized with a seasonal influenza vaccine. *J Clin Invest* **120**, 1663-1673 (2010).

84. Corti, D. et al. A neutralizing antibody selected from plasma cells that binds to group 1 and group 2 influenza A hemagglutinins. *Science* **333**, 850-856 (2011).

85. Dreyfus, C., Ekiert, D.C. & Wilson, I.A. Structure of a classical broadly neutralizing stem antibody in complex with a pandemic H2 influenza virus hemagglutinin. *J Virol* **87**, 7149-7154 (2013).

86. Dreyfus, C. et al. Highly conserved protective epitopes on influenza B viruses. *Science* **337**, 1343-1348 (2012).

87. Sui, J. et al. Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses. *Nat Struct Mol Biol* **16**, 265-273 (2009).

88. Wang, T.T. et al. Broadly protective monoclonal antibodies against H3 influenza viruses following sequential immunization with different hemagglutinins. *PLoS Pathog* **6**, e1000796 (2010).

89. Li, F. Structure, function, and evolution of coronavirus spike proteins. *Annu Rev Virol* **3**, 237-261 (2016).

90. Letko, M., Seifert, S.N., Olival, K.J., Plowright, R.K. & Munster, V.J. Bat-borne virus diversity, spillover and emergence. *Nat Rev Microbiol* **18**, 461-471 (2020).

91. Song, Z. et al. From SARS to MERS, thrusting coronaviruses into the spotlight. *Viruses* **11** (2019).

92. Lu, R. et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **395**, 565-574 (2020).

93. Graham, R.L. & Baric, R.S. Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. *J Virol* **84**, 3134-3146 (2010).

94. Menachery, V.D. et al. A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. *Nat Med* **21**, 1508-1513 (2015).

95. Menachery, V.D. et al. SARS-like WIV1-CoV poised for human emergence. *Proc Natl Acad Sci U S A* **113**, 3048-3053 (2016).

96. Wrapp, D. et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260-1263 (2020).

97. Tortorici, M.A. & Veesler, D. Structural insights into coronavirus entry. *Adv Virus Res* **105**, 93-116 (2019).

98. Krammer, F. SARS-CoV-2 vaccines in development. *Nature* **586**, 516-527 (2020).

99. Jiang, S., Hillyer, C. & Du, L. Neutralizing antibodies against SARS-CoV-2 and other human coronaviruses. *Trends Immunol* **41**, 355-359 (2020).

100. Georgiou, G. et al. The promise and challenge of high-throughput sequencing of the antibody repertoire. *Nat Biotechnol* **32**, 158-168 (2014).

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

102. Wang, B. et al. Functional interrogation and mining of natively paired human VH:VL antibody repertoires. *Nat Biotechnol* **36**, 152-155 (2018).

103. Buchacher, A. et al. Generation of human monoclonal antibodies against HIV-1 proteins; electrofusion and Epstein-Barr virus transformation for peripheral blood lymphocyte immortalization. *AIDS Res Hum Retroviruses* **10**, 359-369 (1994).

104. Stiegler, G. et al. A potent cross-clade neutralizing human monoclonal antibody against a novel epitope on gp41 of human immunodeficiency virus type 1. *AIDS Res Hum Retroviruses* **17**, 1757-1765 (2001).

105. Bonsignori, M. et al. Inference of the HIV-1 VRC01 antibody lineage unmutated common ancestor reveals alternative pathways to overcome a key glycan barrier. *Immunity* **49**, 1162-1174 e1168 (2018).

106. Huang, J. et al. Broad and potent HIV-1 neutralization by a human antibody that binds the gp41-gp120 interface. *Nature* **515**, 138-142 (2014).

107. Walker, L.M. et al. Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. *Science* **326**, 285-289 (2009).

108. Walker, L.M. et al. Broad neutralization coverage of HIV by multiple highly potent antibodies. *Nature* **477**, 466-470 (2011).

109. Adler, A.S. et al. Rare, high-affinity mouse anti-PD-1 antibodies that function in checkpoint blockade, discovered using microfluidics and molecular genomics. *MAbs* **9**, 1270-1281 (2017).

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

111. DeKosky, B.J. et al. High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. *Nat Biotechnol* **31**, 166-169 (2013).

112. Peterson, V.M. et al. Multiplexed quantification of proteins and transcripts in single cells. *Nat Biotechnol* **35**, 936-939 (2017).

113. Tan, Y.C. et al. Barcode-enabled sequencing of plasmablast antibody repertoires in rheumatoid arthritis. *Arthritis Rheumatol* **66**, 2706-2715 (2014).

114. Mimitou, E.P. et al. Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat Methods* **16**, 409-412 (2019).

115. Stoeckius, M. et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods* **14**, 865-868 (2017).

116. Setliff, I. et al. High-throughput mapping of B cell receptor sequences to antigen specificity. *Cell* **179**, 1636-1646 e1615 (2019).

117. DeFalco, J. et al. Non-progressing cancer patients have persistent B cell responses expressing shared antibody paratopes that target public tumor antigens. *Clin Immunol* **187**, 37-45 (2018).

118. Setliff, I. et al. Multi-donor longitudinal antibody repertoire sequencing reveals the existence of public antibody clonotypes in HIV-1 infection. *Cell Host Microbe* **23**, 845-854 e846 (2018).

119. Weaver, G.C. et al. In vitro reconstitution of B cell receptor-antigen interactions to evaluate potential vaccine candidates. *Nat Protoc* **11**, 193-213 (2016).

120. Lingwood, D. et al. Structural and genetic basis for development of broadly neutralizing influenza antibodies. *Nature* **489**, 566-570 (2012).

121. Georgiev, I.S. et al. Single-chain soluble BG505.SOSIP gp140 trimers as structural and antigenic mimics of mature closed HIV-1 env. *J Virol* **89**, 5318-5329 (2015).

122. Sanders, R.W. et al. A next-generation cleaved, soluble HIV-1 Env trimer, BG505 SOSIP.664 gp140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies. *PLoS Pathog* **9**, e1003618 (2013).

123. Ringe, R.P. et al. Improving the expression and purification of soluble, recombinant native-like HIV-1 envelope glycoprotein trimers by targeted sequence changes. *J Virol* **91** (2017).

124. Whittle, J.R. et al. Flow cytometry reveals that H5N1 vaccination elicits cross-reactive stem-directed antibodies from multiple Ig heavy-chain lineages. *J Virol* **88**, 4047-4057 (2014).

125. Wu, X. et al. Maturation and diversity of the VRC01-antibody lineage over 15 years of chronic HIV-1 infection. *Cell* **161**, 470-485 (2015).

126. Kwon, Y.D. et al. Crystal structure, conformational fixation and entry-related interactions of mature ligand-free HIV-1 Env. *Nat Struct Mol Biol* **22**, 522-531 (2015).

127. Cale, E.M. et al. Virus-like particles identify an HIV V1V2 apex-binding neutralizing antibody that lacks a protruding loop. *Immunity* **46**, 777-791 e710 (2017).

128. Wu, X. et al. Selection pressure on HIV-1 envelope by broadly neutralizing antibodies to the conserved CD4-binding site. *J Virol* **86**, 5844-5856 (2012).

129. Harris, A. et al. Trimeric HIV-1 glycoprotein gp140 immunogens and native HIV-1 envelope glycoproteins display the same closed and open quaternary molecular architectures. *Proc Natl Acad Sci U S A* **108**, 11440-11445 (2011).

130. Joyce, M.G. et al. Soluble prefusion closed DS-SOSIP.664-Env trimers of diverse HIV-1 strains. *Cell Rep* **21**, 2992-3002 (2017).

131. Julien, J.P. et al. Design and structure of two HIV-1 clade C SOSIP.664 trimers that increase the arsenal of native-like Env immunogens. *Proc Natl Acad Sci U S A* **112**, 11947-11952 (2015).

132. Pugach, P. et al. A native-like SOSIP.664 trimer based on an HIV-1 subtype B env gene. *J Virol* **89**, 3380-3395 (2015).

133. Dosenovic, P. et al. Anti-idiotypic antibodies elicit anti-HIV-1-specific B cell responses. *J Exp Med* **216**, 2316-2330 (2019).

134. Jardine, J. et al. Rational HIV immunogen design to target specific germline B cell receptors. *Science* **340**, 711-716 (2013).

135. Jardine, J.G. et al. HIV-1 broadly neutralizing antibody precursor B cells revealed by germline-targeting immunogen. *Science* **351**, 1458-1463 (2016).

136. Stamatatos, L., Pancera, M. & McGuire, A.T. Germline-targeting immunogens. *Immunol Rev* **275**, 203-216 (2017).

137. Kotecha, N., Krutzik, P.O. & Irish, J.M. Web-based analysis and publication of flow cytometry experiments. *Current protocols in cytometry* **53**, 10.17. 11-10.17. 24 (2010).

138. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884-i890 (2018).

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

140. Gupta, N.T. et al. Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. *Bioinformatics* **31**, 3356-3358 (2015).

141. Pedregosa, F. et al. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research* **12**, 2825-2830 (2011).

142. Hunter, J.D. Matplotlib: A 2D graphics environment. *Computing in science & engineering* **9**, 90-95 (2007).

143. Guindon, S. et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**, 307-321 (2010).

144. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* **47**, W256-W259 (2019).

145. Sarzotti-Kelsoe, M. et al. Optimization and validation of the TZM-bl assay for standardized assessments of neutralizing antibodies against HIV-1. *J Immunol Methods* **409**, 131-146 (2014).

146. Ackerman, M.E. et al. A robust, high-throughput assay to determine the phagocytic activity of clinical antibody samples. *J Immunol Methods* **366**, 8-19 (2011).

147. Richardson, S.I. et al. HIV-specific Fc effector function early in infection predicts the development of broadly neutralizing antibodies. *PLoS Pathog* **14**, e1006987 (2018).

148. Richardson, S.I., Crowther, C., Mkhize, N.N. & Morris, L. Measuring the ability of HIV-specific antibodies to mediate trogocytosis. *J Immunol Methods* **463**, 71-83 (2018).

149. Pollara, J. et al. High-throughput quantitative analysis of HIV-1 and SIV-specific ADCC-mediating antibody responses. *Cytometry A* **79**, 603-612 (2011).

150. Bosch, B.J., van der Zee, R., de Haan, C.A. & Rottier, P.J. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *J Virol* **77**, 8801-8811 (2003).

151. Brouwer, P.J.M. et al. Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science* **369**, 643-650 (2020).

152. Chi, X. et al. A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. *Science* **369**, 650-655 (2020).

153. Zost, S.J. et al. Rapid isolation and profiling of a diverse panel of human monoclonal antibodies targeting the SARS-CoV-2 spike protein. *Nat Med* **26**, 1422-1427 (2020).

154. Pinto, D. et al. Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* **583**, 290-295 (2020).

155. Rogers, T.F. et al. Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease in a small animal model. *Science* **369**, 956-963 (2020).

156. Chen, P. et al. SARS-CoV-2 neutralizing antibody LY-CoV555 in outpatients with Covid-19. *N Engl J Med* **384**, 229-237 (2021).

157. Cohen, M.S. Monoclonal antibodies to disrupt progression of early Covid-19 infection. *N Engl J Med* **384**, 289-291 (2021).

158. Weinreich, D.M. et al. REGN-COV2, a neutralizing antibody cocktail, in outpatients with Covid-19. *N Engl J Med* **384**, 238-251 (2021).

159. Liu, H. et al. Cross-neutralization of a SARS-CoV-2 antibody to a functionally conserved site is mediated by avidity. *Immunity* **53**, 1272-1280 e1275 (2020).

160. Wec, A.Z. et al. Broad neutralization of SARS-related viruses by human monoclonal antibodies. *Science* **369**, 731-736 (2020).

161. Lv, H. et al. Cross-reactive antibody response between SARS-CoV-2 and SARS-CoV infections. *Cell Rep* **31**, 107725 (2020).

162. Zohar, T. & Alter, G. Dissecting antibody-mediated protection against SARS-CoV-2. *Nat Rev Immunol* **20**, 392-394 (2020).

163. Ng, K.W. et al. Preexisting and de novo humoral immunity to SARS-CoV-2 in humans. *Science* **370**, 1339-1343 (2020).

164. Yasui, F. et al. Phagocytic cells contribute to the antibody-mediated elimination of pulmonary-infected SARS coronavirus. *Virology* **454-455**, 157-168 (2014).

165. Schafer, A. et al. Antibody potency, effector function, and combinations in protection and therapy for SARS-CoV-2 infection in vivo. *J Exp Med* **218** (2021).

166. Atyeo, C. et al. Distinct early serological signatures track with SARS-CoV-2 survival. *Immunity* **53**, 524-532 e524 (2020).

167. Loos, C. et al. Evolution of early SARS-CoV-2 and cross-coronavirus immunity. *mSphere* **5** (2020).

168. Ou, X. et al. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat Commun* **11**, 1620 (2020).

169. Yuan, M. et al. A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science* **368**, 630-633 (2020).

170. Dinnon, K.H., 3rd et al. A mouse-adapted model of SARS-CoV-2 to test COVID-19 countermeasures. *Nature* **586**, 560-566 (2020).

171. Leist, S.R. et al. A mouse-adapted SARS-CoV-2 induces acute lung injury and mortality in standard laboratory mice. *Cell* **183**, 1070-1085 e1012 (2020).

172. Zohar, T. et al. Compromised humoral functional evolution tracks with SARS-CoV-2 mortality. *Cell* **183**, 1508-1519 e1512 (2020).

173. Atyeo, C. et al. Dissecting strategies to tune the therapeutic potential of SARS-CoV-2-specific monoclonal antibody CR3022. *JCI Insight* **6** (2021).

174. DiLillo, D.J., Palese, P., Wilson, P.C. & Ravetch, J.V. Broadly neutralizing anti-influenza antibodies require Fc receptor engagement for in vivo protection. *J Clin Invest* **126**, 605-610 (2016).

175. Winkler, E.S. et al. Human neutralizing antibodies against SARS-CoV-2 require intact Fc effector functions for optimal therapeutic protection. *Cell* **184**, 1804-1820 e1816 (2021).

176. Edwards, C.E. et al. Swine acute diarrhea syndrome coronavirus replication in primary human cells reveals potential susceptibility to infection. *Proc Natl Acad Sci U S A* **117**, 26915-26925 (2020).

177. Baca, M., Presta, L.G., O'Connor, S.J. & Wells, J.A. Antibody humanization using monovalent phage display. *J Biol Chem* **272**, 10678-10684 (1997).

178. Walls, A.C. et al. Unexpected receptor functional mimicry elucidates activation of coronavirus fusion. *Cell* **176**, 1026-1039 e1015 (2019).

179. Tang, X.C. et al. Identification of human neutralizing antibodies against MERS-CoV and their role in virus adaptive evolution. *Proc Natl Acad Sci U S A* **111**, E2018-2026 (2014).

180. Madeira, F. et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res* **47**, W636-W641 (2019).

181. Henikoff, S. & Henikoff, J.G. Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci U S A* **89**, 10915-10919 (1992).

182. Pallesen, J. et al. Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. *Proc Natl Acad Sci U S A* **114**, E7348-E7357 (2017).

183. Mukherjee, S. et al. Enhancing dengue virus maturation using a stable furin over-expressing cell line. *Virology* **497**, 33-40 (2016).

184. Scobey, T. et al. Reverse genetics with a full-length infectious cDNA of the Middle East respiratory syndrome coronavirus. *Proc Natl Acad Sci U S A* **110**, 16157-16162 (2013).

185. Yount, B. et al. Reverse genetics with a full-length infectious cDNA of severe acute respiratory syndrome coronavirus. *Proc Natl Acad Sci U S A* **100**, 12995-13000 (2003).

186. Nath Neerukonda, S., Vassell, R. & Weiss, C.D. Neutralizing antibodies targeting the conserved stem region of influenza hemagglutinin. *Vaccines (Basel)* **8** (2020).

187. Kallewaard, N.L. et al. Structure and function analysis of an antibody recognizing all influenza A subtypes. *Cell* **166**, 596-608 (2016).

188. Ekiert, D.C. et al. Antibody recognition of a highly conserved influenza virus epitope. *Science* **324**, 246-251 (2009).

189. Liu, M. et al. Polyreactivity and autoreactivity among HIV-1 antibodies. *J Virol* **89**, 784-798 (2015).

190. Mouquet, H. et al. Polyreactivity increases the apparent affinity of anti-HIV antibodies by heteroligation. *Nature* **467**, 591-595 (2010).

191. Mouquet, H. & Nussenzweig, M.C. Polyreactive antibodies in adaptive immune responses to viruses. *Cell Mol Life Sci* **69**, 1435-1445 (2012).

192. Trama, A.M. et al. HIV-1 envelope gp41 antibodies can originate from terminal ileum B cells that share cross-reactivity with commensal bacteria. *Cell Host Microbe* **16**, 215-226 (2014).

193. Williams, W.B., Han, Q. & Haynes, B.F. Cross-reactivity of HIV vaccine responses and the microbiome. *Curr Opin HIV AIDS* **13**, 9-14 (2018).

194. Wen, X. et al. Structure of the human metapneumovirus fusion protein with neutralizing antibody identifies a pneumovirus antigenic site. *Nature Structural & Molecular Biology* **19**, 461-463 (2012).

195. Tan, J. et al. A public antibody lineage that potently inhibits malaria infection through dual binding to the circumsporozoite protein. *Nat Med* **24**, 401-407 (2018).

196. Calarese, D.A. et al. Antibody domain exchange is an immunological solution to carbohydrate cluster recognition. *Science* **300**, 2065-2071 (2003).

197. Doores, K.J., Fulton, Z., Huber, M., Wilson, I.A. & Burton, D.R. Antibody 2G12 recognizes di-mannose equivalently in domain- and nondomain-exchanged forms but only binds the HIV-1 glycan shield if domain exchanged. *J Virol* **84**, 10690-10699 (2010).

198. Scanlan, C.N., Offer, J., Zitzmann, N. & Dwek, R.A. Exploiting the defensive sugars of HIV-1 for drug and vaccine design. *Nature* **446**, 1038-1045 (2007).

199. Williams, W.B. et al. Fab-dimerized glycan-reactive antibodies are a structural category of natural antibodies. *Cell* **184**, 2955-2972 e2925 (2021).

200. Lee, C.D. et al. A cross-neutralizing antibody between HIV-1 and influenza virus. *PLoS Pathog* **17**, e1009407 (2021).

201. Zost, S.J. et al. Potently neutralizing and protective human antibodies against SARS-CoV-2. *Nature* **584**, 443-449 (2020).

202. Hansen, J. et al. Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. *Science* **369**, 1010-1014 (2020).

203. Ju, B. et al. Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature* **584**, 115-119 (2020).

204. Robbiani, D.F. et al. Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* **584**, 437-442 (2020).

205. Shi, R. et al. A human neutralizing antibody targets the receptor-binding site of SARS-CoV-2. *Nature* **584**, 120-124 (2020).

206. Shiakolas, A.R. et al. Cross-reactive coronavirus antibodies with diverse epitope specificities and Fc effector functions. *Cell Rep Med*, 100313 (2021).

207. Gilchuk, P. et al. Integrated pipeline for the accelerated discovery of antiviral antibody therapeutics. *Nat Biomed Eng* **4**, 1030-1043 (2020).

208. Suryadevara, N. et al. Neutralizing and protective human monoclonal antibodies recognizing the N-terminal domain of the SARS-CoV-2 spike protein. *Cell* **184**, 2316-2331 e2315 (2021).

209. Hsieh, C.L. et al. Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **369**, 1501-1505 (2020).

210. Mastronarde, D.N. Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol* **152**, 36-51 (2005).

211. Punjani, A., Rubinstein, J.L., Fleet, D.J. & Brubaker, M.A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296 (2017).

212. Ekiert, D.C. et al. A highly conserved neutralizing epitope on group 2 influenza A viruses. *Science* **333**, 843-850 (2011).

213. Krammer, F. Pandemic vaccines: how are we going to be better prepared next time? *Med (N Y)* **1**, 28-32 (2020).

214. Giurgea, L.T., Han, A. & Memoli, M.J. Universal coronavirus vaccines: the time to start is now. *NPJ Vaccines* **5**, 43 (2020).

215. Taylor, P.C. et al. Neutralizing monoclonal antibodies for treatment of COVID-19. *Nat Rev Immunol* **21**, 382-393 (2021).

216. Taylor, L.H., Latham, S.M. & Woolhouse, M.E. Risk factors for human disease emergence. *Philos Trans R Soc Lond B Biol Sci* **356**, 983-989 (2001).

217. Woolhouse, M.E. & Gowtage-Sequeria, S. Host range and emerging and reemerging pathogens. *Emerg Infect Dis* **11**, 1842-1847 (2005).

218. Havenar-Daughton, C. et al. The human naive B cell repertoire contains distinct subclasses for a germline-targeting HIV-1 vaccine immunogen. *Sci Transl Med* **10** (2018).

219. Bricault, C.A. et al. HIV-1 neutralizing antibody signatures and application to epitope-targeted vaccine design. *Cell Host Microbe* **26**, 296 (2019).

220. Borst, A.J. et al. Germline VRC01 antibody recognition of a modified clade C HIV-1 envelope trimer and a glycosylated HIV-1 gp120 core. *Elife* **7** (2018).

221. Escolano, A. et al. Sequential immunization elicits broadly neutralizing anti-HIV-1 antibodies in Ig knockin mice. *Cell* **166**, 1445-1458 e1412 (2016).

222. Andrabi, R. et al. Identification of common features in prototype broadly neutralizing antibodies to HIV envelope V2 apex to facilitate vaccine design. *Immunity* **43**, 959-973 (2015).

223. NIA (National Institute on Aging, 2021).

224. Long, J.M. & Holtzman, D.M. Alzheimer disease: an update on pathobiology and treatment strategies. *Cell* **179**, 312-339 (2019).

225. van Dyck, C.H. Anti-Amyloid-beta monoclonal antibodies for Alzheimer's disease: pitfalls and promise. *Biol Psychiatry* **83**, 311-319 (2018).

226. Bystroff, C. & Krogh, A. Hidden Markov Models for prediction of protein features. *Methods Mol Biol* **413**, 173-198 (2008).

227. Durbin, R., Eddy, S.R., Krogh, A. & Mitchison, G. Biological sequence analysis: probabilistic models of proteins and nucleic acids. (Cambridge university press, 1998).