Methods, Models, and Machine Learning Approaches for Understanding Pathogen-Specific Humoral Immunity

by

Tomer Zohar

B.S. Bioengineering
University of Maryland, College Park 2017

Submitted to the department of Biological Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Biological Engineering
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August 2022

Signature of Author.....
Tomer Zohar
Department of Biological Engineering
August 5, 2022

Certified by.....
Douglas A. Lauffenburger
Ford Professor of Engineering, MIT
Thesis Supervisor

Certified by.....
Galit Alter
Professor of Medicine at Harvard Medical School
Thesis Supervisor

Accepted by.....
Katharina Ribbeck
Professor of Biological Engineering
Chair of Graduate Program, Department of Biological Engineering

Thesis Committee Members

Darrell J. Irvine, Ph.D. (Chair)
Professor, Biological Engineering
Massachusetts Institute of Technology

Douglas A. Lauffenburger, Ph.D. (Thesis Supervisor)
Professor, Biological Engineering
Massachusetts Institute of Technology

Galit Alter, Ph.D. (Thesis Supervisor)
Professor of Medicine
Harvard Medical School

Jacquin C. Niles, M.D., Ph.D.
Professor, Biological Engineering
Massachusetts Institute of Technology

Methods, Models, and Machine Learning Approaches for Understanding Pathogen-Specific Humoral Immunity

by

Tomer Zohar

Submitted to the Department of Biological Engineering
on August 5, 2022 in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Biological Engineering

Abstract

The humoral immune response is comprised of vast libraries of polyclonal antibodies capable of recognizing a myriad of targets and directing a spectrum of innate immune functions. The complex heterogeneity in antibody profiles across both populations and diseases makes defining mechanisms of protection difficult. Understanding these mechanisms and the factors that influence them is essential to defining immunity and helps inform the design of vaccines and therapeutics. Thus, in this thesis, I describe five studies that present the development of experimental and computational methods, and machine learning approaches for investigating the mechanisms, dynamics, and determinants of pathogen-specific humoral immunity.

The first study introduces an assay for probing antigen-specific antibody mediated primary monocyte phagocytosis, that is capable of capturing subsequent downstream functions. The second study describes a machine learning approach for defining the correlates of upper and lower respiratory protection against RSV and methods for evaluating vaccine designs. The third study uses machine learning methods to uncover signatures of humoral protection against SARS-CoV-2. The fourth study presents a method for longitudinally modelling humoral immunity that was used to investigate the temporal dynamics of antibody features across individuals with varying COVID-19 severity. Finally, the last study describes a genome-wide association screen of pathogen-specific polyclonal antibody characteristics and functions that was then validated with transcriptomics data. Ultimately, the methods described in this thesis present new approaches for investigating underlying phenomena related to pathogen-specific humoral immunity.

Thesis Supervisors: Douglas A. Lauffenburger | Galit Alter
Title: Ford Professor of Engineering, MIT | Professor of Medicine at Harvard Medical School

Contents

Acknowledgements

I have been immensely fortunate to have incredible mentors, collaborators, colleagues, and friends during my graduate studies.

First and foremost, I would like to thank my thesis advisers Doug Lauffenburger and Galit Alter. To Doug, thank you so much for your guidance, encouragement, and support both scientifically and beyond. From our first interaction you have always made me feel like a valued student in the department and a member of the lab. Thank you for giving me the freedom and flexibility to explore my ideas while always providing guidance when I needed it, and for the connections you have provided me with. To Galit, thank you for all the opportunities you have provided me with and always encouraging me to explore every avenue. Thank you for your infectious enthusiasm, your dedication to making an impact on human health has been a constant source of motivation.

I would also like to thank the members of my thesis committee. To Darrel Irvine, thank you for helping to guide and focus my thesis work, and for your helpful advice and perspectives. To Jacquin Niles, I am greatly appreciative of your scientific inquisitiveness and helpful advice. Our meetings were always a well of intellectual insights which have been immensely helpful in guiding my progress.

Thank you to Jishnu Das, and Madeleine Jennewein for your patience and helping me jump start my research in the lab. Thank you to all my collaborators whom without this thesis would not have been possible, and especially to Carolin Loss, Caroline Atyeo, and Stephanie Fischinger who I worked most closely with.

To Eddie, Carolyn, Krista, Avlant, Brian, Meelim, Chuangqi, Matt Gorham, Yannic, and all the wonderful members of the Lauffenburger and Alter labs thank you for the engaging scientific conversations, your perpetual optimism, and all our memorable moments in lab.

I would like to also thank all my classmates and friends in the BE-2017 cohort. From the moment I moved to Boston you made me feel like family. I have been so fortunate to be a part of such a brilliant group of individuals. Your accomplishments and good nature are inspiring. Additionally, I would like to thank the MIT BE department as a whole for cultivating a supportive, creative, and intellectually stimulating environment and for striving for improvements in diversity and inclusion.

Thank you to the original members of 176 Elm Street: Atti, Josh, and Joseph. I closely cherish our time together; it has been an honor to grow alongside you.

To my parents, thank you for your endless support and guidance. Thank you for always encouraging me to explore my curiosities. Despite our geographic distance your messages and calls have always been a source of great comfort.

To my amazing wife, Stephanie, your strength, kindness, and resilience have been a constant source of inspiration. Thank you for supporting me through long workdays and late worknights. Without your help this work would have not been possible.

To my spectacular daughter, Alma, your presence is always an endless source of joy and happiness that helps to reinvigorate and rejuvenate me. Your constant curiosity about the world around you is inspiring. Perhaps one day you too may be writing your own thesis.

Chapter 1

Introduction

1.1 The immune response against infection and vaccination

The immune response is the mechanism by which the body protects itself from foreign invasion. It has evolved in an 'arms race' alongside pathogens resulting in a complex and elegant system of recognizing, protecting against, and remembering infection[1]. The two branches of the immune system are comprised of innate and adaptive immunity. The innate immune system is the first line of defense against infection and consists of monocytes, macrophages, neutrophils, dendritic cells, natural killer cells, and many other cell types each uniquely responsible for different tasks in the overall response. The general responsibilities of the innate immune system include sensing the presence of foreign pathogens, pathogen killing, coordinating the broader immune response through cytokine release, and triggering the adaptive immune response.

Adaptive immunity is capable of recognizing a vastly larger range of antigens in contrast to the innate immune system that relies on pattern recognition receptors (PPR) that identify pathogen-associated molecular patterns (PAMPs). This results in a tailored response that can optimally eliminate pathogens or pathogen-infected cells, and post-infection is capable of forming a memory response. Adaptive immunity is comprised of both cellular and humoral components. Cellular immunity is optimally capable of eliminating pathogen-infected cells through the help of T cells. The humoral immune system on the other hand protects against infection through the generation of antibodies, proteins capable of targeting different parts of pathogens, and either disrupting their interactions with host cell receptors or triggering innate immune function. A disadvantage of natural active acquired immunity is that the adaptive immune system requires several days in order to generate a substantial response. Alternatively, immunity can be artificially and safely induced with vaccines which

function by exposing the immune system to either an inactivated pathogen or the individual components of a pathogen (i.e., proteins).

1.2 Polyclonal antibody composition impacts disease Outcome

At present, there are 26 pathogenic diseases for which at least one vaccine is licensed[2]. However, despite ongoing efforts to develop vaccines for lower respiratory infections, diarrheal diseases, Tuberculosis, Malaria, and Human immunodeficiency virus (HIV) infectious diseases remain the leading cause of death in low-income countries and the 4th leading cause of death worldwide[3]. Current vaccine approaches are guided by rational design. They rely on preexisting knowledge about pathogenesis and host interactions for the selection of delivery methods and immunogens likely to elicit effective responses[4]. Historical, efficacious humoral responses were evaluated based on high titers and capacity to neutralize. However, neutralization and titer are insufficient to provide immunity to most clinical vaccines, and alone mechanistically cannot drive protection[5,6]. Rather, a broader spectrum humoral features and mediated functions are attributed to conferring protection to HIV[7–11], Influenza[12,13], Herpes Simplex Virus[14,15], Ebola[16], and Malaria[17,18]. Though overlaps exist correlates of immunity and susceptibility can differ across both population and disease. For example, antibody sialyation upon administration of the influenza vaccine determines efficacy through driving affinity maturation[19]. In contrast, antibodies enhanced for FcγRIIIA binding drive disease severity to dengue[20].

Correlates of immunity driven by the properties and functions of polyclonal antibodies are still being revealed in many disease contexts and populations. However, despite insight into the primary implications of certain polyclonal properties the underlying mechanisms that drive separation across populations are not well understood. Mechanistic findings have the ability to further inform vaccine approaches, provide new targets for immunomodulatory therapies, and could be used as diagnostic markers prior to infection.

1.3 Antibody characteristics and mediated functions

Antibodies vary either through affinity or structure. The antigen-binding region (Fab) determines the target epitope (i.e. antigenic region), and the binding affinity. The fragment crystallizable (Fc) region is the antibody's structural component and includes IgG, IgA, IgM, and IgE secreted isotypes. These isotypes are further delineated into subclasses including four IgG 1-4, and IgA 1-2. In addition, antibodies have a single N-linked glycan attached to asparagine 297 (Asn297) on the Fc region. There are 36 possible glycan structures that can occur on Asn297 which result in differences in structure[21]. The structural components of antibodies are important because they act as binding ligands to Fc receptors (FcR) on immune cells. Each FcR propagates signaling through either tyrosine-based activation motifs (ITAM), tyrosine-based inhibitory motifs

(ITIM), or decoy motifs and results in distinct mediated phenotypes. There are three main classes of FcRs ones that bind IgG (Fcγ), IgE (Fcε), and IgA (Fcα).

In contrast to monoclonals, polyclonal pools contain a large diversity of Fab/Fc unique antibodies that range in abundance. This diversity provides advantages over monoclonal antibodies. For one, altered expression and epitope-specific mutations are common pathogen immunoevasion techniques. By targeting different antigenic regions there is a better chance of clearance through mitigating escape. As mentioned before neutralization, a function mediated by affinity alone is insufficient for protection to many diseases. Antibodies mediate many important immune functions such as phagocytosis, complement deposition, natural killer (NK) cell degranulation, and cytokine and chemokine secretion. In addition to FcRs functions can also be mediated by complement receptors (CR) on immune cells like complement receptors CR1 and CR2. Therefore, Fc diversity improves the resolution of immune phenotypes allowing for optimal and custom pathogen-specific responses which can also be altered easily through glycosylation[21].

1.4 Current suite of serological assays for profiling antibody properties

There exist robust methods of measuring polyclonal properties such as isotype and subclass abundancy to different antigens, and enhanced binding to FcRs using Luminex based assays[22]. Mediated functions are measured by incubating antigen covered fluorescent beads with polyclonal antibodies, introducing effectors (i.e. immune cells or complement), and recording phenotypes. However, though it has been shown that phagocytosis and complement assays can be multiplexed more validation and development is required before multiplexing can be fully adopted[23–25].

System serology techniques are composed of fluorescent beads-based assays used to measure isotype and subclasses abundance, binding to various Fc receptors, the ability to trigger phagocytosis, complement deposition, and more. Isotypes and subclasses are measured through carboxy coupling either as pure antigens, a mixture of antigens, lysates, or whole intact pathogens to Luminex beads. Different pathogen coated beads can then be pooled and incubated with serum to form immune complexes. Using secondary fluorescent antibodies isotypes and subclasses are detected and readout with flow cytometry. Fc receptor binding is measured in a similar fashion, but instead of using secondary antibodies as detectors fluorescent recombinant FcRs are used[26]. Collectively, these methods are used to measure IgG1, IgG2, IgG3, IgG4, IgA1, and IgM abundancy and enhanced binding to FcRγIIA, FcRγIIB, FcRγIIIA, and FcRγIIIB. However, all other system serology techniques are not multiplexable which would ultimately accommodate larger cohort sizes.

Antibody-dependent phagocytosis is tested with primary human neutrophils (ADNP), and monocytes (ADMP). Each cell type has unique sets of FcRs and different roles in the overall response. Therefore, it is not uncommon that phagocytosis is mediate better in one cell type over another. Recent work has showed that multiplexing a

phagocytosis assay with THP-1s was possible[23]. However, further characterization is required in order to address concerns about signal loss and cross-reactivity. Furthermore, monocyte phagocytosis is currently only performed with THP-1, a monocytic cell line. In order to capture information that is closer to underlying biology the assay can be performed with human primary monocytes. This would enable the study of phagocytosis in various monocyte subsets such as $CD14^{+}CD16^{-}$ classical, $CD14^{-}CD16^{+}$ non-classical, and $CD14^{+}CD16^{+}$ intermediate monocytes as well as downstream functions.

Antibody-dependent complement deposition (ADCD) like the phagocytosis assays is measured by first forming immune complexes as described above. Then, guinea pig complement is incubated with the complexes and detected with a fluorescent anti-C3 secondary antibody[24]. However, unlike the phagocytosis assays, ADCD is not a cell-based assay and can be performed using the Luminex platform[25]. This would allow samples to be profiled with higher throughput.

1.5 Current computational methods for defining correlates of protection

Both applied machine learning approaches and specific computational techniques have previously been employed in order to understand the role of humoral immunity against many infectious diseases. This analytic framework is composed of both non-supervised and supervised machine learning methods. Customarily, after the data is processed through a normalization method (i.e., z-scoring), exploratory data analysis (EDA) is used to summarize univariate trends. Then, unsupervised methods such as hierarchical clustering and principal component analysis (PCA) are employed to inform on any inherit underlying structure in the data as well as evaluating the existence of any confounding variables. Next, supervised machine learning regression and classification methods are used in order to maximize variance due to a specific outcome of interest (i.e., protection against infection). With systems serology data, full rank datasets are very common, therefore a feature reduction technique is coupled with the machine learning methods of interest. Most commonly, the least absolute shrinkage and selection operator (Lasso) regression is used. Furthermore, popular classifiers include support vector machines (SVM) and partial least squares-discriminant analysis (PLS-DA). The robustness of these modelling approaches are evaluated in a repeated five-fold cross validation and then visualized with PLS-DA. However, there are many multi-level and structured study designs for which this framework does not fit, and therefore the development of additional computational methods is needed.

1.6 Development of the humoral response

B cells are responsible for the composition and quality of polyclonal antibodies. The major source of antibody binding diversity is driven by gene rearrangement[27]. During early B cell development cellular machinery recombines V(D)J segments resulting in cell-specific B cell receptors (BCR). BCRs are membrane anchored immunoglobins that

upon activated triggers the set of events responsible for the overall humoral response[28]. B cells upon maturation leave the bone marrow and traffic to secondary lymphoid tissue in search of stimulation from foreign antigens. B cells with unique BCRs are constantly generated, and if not stimulated within a time window parish.

B cells can be activated in either a T cell independent or dependent manner. In both cases specialized antigen-presenting cells (APC) such as macrophages and follicular dendritic cells (FDC) display antigens either to B cells directly or to CD4+ T helper cells[29]. Activation of the BCRs results in proliferation followed by differentiation into germinal center (GC) B cells, short-lived plasma cells (SLPC), long-lived plasma cells (LLPC), or memory cell (Bmem). Multi signal integration from T helper cells, APCs, and many other signaling events determine the route of differentiation. Within germinal centers B cells undergo affinity maturation and class switching. During this process more SLPC, LLPC, and Bmems are produced as GC B cells leave. The process of generating polyclonal antibodies against pathogens involves dynamic and complex signaling decisions. Mechanistic links between B cell biology and composite humoral properties are still being discovered[30].

1.7 Immunogenetics of the humoral response

Polyclonal pools can vary within an individual to different pathogens as well as across populations for a given pathogen. This range of variability in humoral characteristics and responses gives rise to subpopulations of individuals that control specific diseases differentially[7–18]. These findings lend to the notion that humoral traits developed through selective pressure via the host-pathogen relationship and therefore enriching sets of genes responsible for favorable outcomes. Population genetics studies the genetic and environmental contributions to trait determination (i.e., heritability) as well as allelic variance across populations. Two of the main approaches for studying these effects are through estimating heritability with familial study designs and identifying specific single nucleotide polymorphisms (SNP) that associate with traits using gene wide association studies (GWAS). Increasing evidence from heritability studies and GWAS highlight the influences that genetics can have on both antibody binding and structure.

1.7.1 The genetic influences behind antibody repertories

During the early development of B cells gene recombination occurs in segments V (variable), D (diversity), and J (joining) known as V(D)J recombination. These gene segments encode the antibody repertoire and are considered the main source of antigen-binding diversity. The diversity of repertoires within individuals and across populations has been of great interest, but prohibitively difficult due to the massive sequence size, ability to sample B cells, and small cohorts[31]. Within a single individual, the amount of information encoded by a repertoire exceeds the human genome by four orders of magnitude. However, with improvements in sequencing capabilities, recent

studies are starting to identify important features of repertoires and provide insight on commonality and genetic influence.

The process of gene segment selection during recombination is mostly random, although some segments are used more often than others. Studies have shown multiple mechanisms that account for biases in repertoires including preferred segments, and distance-based bias between V, D, and J segments[32,33]. MZ twin studies have also shown similarities in CDR3 length, and though the mechanisms are unknown genetic contributions have been attributed to the usage of gene segments, clonal expansion, and clonal sequence diversity[34–38]. Additionally, the influence of heritable factors on the recombination process has been confirmed to propagate from naïve to the memory compartments as well[39]. Therefore, though repertoires are highly individualized and very diverse unknown genetic biases underscore how binding repertoires are formed[28].

Studying repertoires informs on the variation of possible sequences against a pathogen but does not necessarily capture all aspects of polyclonal antibodies. For one, sequence does not necessarily inform on the clonal abundance of certain antibodies. There is emergent effort to leverage computational tools to predict antibody-antigen binding complexes from repertoires, but much development is need for this approach can be scaled and implementable[40,41]. Furthermore, binding complexes only have the compacity to inform on neutralization and say nothing about Fc dynamics.

1.7.2 The influence of genetics on antibody Fc regions

Heritability studies are a general approach to estimating the impact of genetics on determining a trait. They are advantageous because they do not make assumptions about the types of genetic interactions and therefore directly to relate to phenotypes. Multiple studies have estimated high heritability of titers to various childhood vaccine (Table S1)[42]. These estimates make no assumptions about sequences like the approaches mentioned above. In addition, the online GWAS Catalog contains over 24 studies of IgG titers to various infectious diseases for which multiple SNPs have been associated. Therefore, there is evidence of highlighting heritability and specific genetic markers for IgG titers to various pathogens. However, despite IgG being the dominant isotype present in the blood other classes can have substantial impact especially in different compartments of the body. With exception to bulk IgG glycosylation, the genetic contributions and molecular components that influence antibody characteristics and mediated functions have yet to be systematically studied across pathogens[43].

1.8 Overview of thesis

This thesis has focused on developing experimental techniques, computational methods, and machine learning approaches for investigating how properties of polyclonal antibodies shape response to infection. Understanding the properties and mechanisms by which antibodies provide protection is essential to defining immunity. Although neutralizing antibodies have been proposed as a potential key mechanism of

protection against many pathogens, antibodies mediate additional immune functions that may have both protective and pathological consequences. The aim of this work has been to expand the repertoire of experimental techniques capable of capturing antibody mediated effector functions, develop machine learning approaches for uncovering antibody correlates of protection and evaluating vaccine designs, develop computational methods to longitudinally model the development of early humoral immunity, and construct a genomic analysis pipeline to define the genetic mechanisms that determine pathogen-specific polyclonal antibody characteristics and functions. Collectively, these works provide novel experimental methods and computational models which have contributed to the body of knowledge of how polyclonal antibody composition shapes disease outcomes. Ultimately, these tools can be applied to future studies and ultimately aid in improving future vaccine designs and informing clinical management.

Chapter 2

A multifaceted high-throughput assay for probing antigen-specific antibody-mediated primary monocyte phagocytosis and downstream functions

The contents of this chapter were published as:

Tomer Zohar, Caroline Atyeo, Caitlin R. Wolf, Jennifer K. Logue, Kiel Shuey, Nicholas Franko, Robert Y. Choi, Anna Wald, David M. Koelle, Helen Y. Chu, Douglas A. Lauffenburger, and Galit Alter. "A multifaceted high-throughput assay for probing antigen-specific antibody-mediated primary monocyte phagocytosis and downstream functions". *J. Immunol. Methods.* (Accepted)

See the online publication for any references to supplementary material not included in this thesis.

Contributions

TZ and GA wrote the manuscript with contributions from all authors. HYC, AW, RYC, and DMK enrolled the clinical cohorts and facilitated access to blood and plasma samples. CRW and JKL facilitated sample selection and analyzed the demographic and clinical data. CRW, JKL, KS, and NF facilitated subject enrollment, including collection and processing of the samples. CA preformed the antibody production of all Fc-variants. TZ preformed all experiments and analysis of the data.

2.1 Abstract

Monocytes are highly versatile innate immune cells responsible for pathogen clearance, innate immune coordination, and induction of adaptive immunity. Monocytes can directly and indirectly integrate pathogen-destructive instructions and contribute to disease control via pathogen uptake, presentation, or the release of cytokines. Indirect pathogen-specific instructions are conferred via Fc-receptor signaling and triggered by antibody opsonized material. Given the tremendous variation in polyclonal humoral immunity, defining the specific antibody-responses able to arm monocytes most effectively remains incompletely understood. While monocyte cell line-based assays have been used previously, cell lines may not faithfully recapitulate the full biology of monocytes. Thus, here we describe a multifaceted antigen-specific method for probing antibody-dependent primary monocyte phagocytosis (ADMP) and secondary responses. The assay not only reliably captures phagocytic uptake of immune complexes, but also detects unique changes in surface markers and cytokine secretions profiles, poorly detected by monocytic cell lines. The assay captures divergent polyclonal-monocyte recruiting activity across subjects with varying SARS-CoV-2 disease severity and also revealed biological nuances in Fc-mutant monoclonal antibody activity related to differences in Fc-receptor binding. Thus, the ADMP assay is a flexible assay able to provide key insights into the role of humoral immunity in driving monocyte phenotypic transitions and downstream functions across many diseases.

2.2 Introduction

Monocytes constitute approximately 5-12% of white blood cells in healthy individuals[44,45]. Monocyte development begins in the bone marrow. Once mature, monocytes enter circulation into the bloodstream, and ultimately migrate into tissues. During homeostasis, monocytes enter tissues without many cellular changes and eventually undergo apoptosis[46]. However, during infection, monocytes drastically alter their basal phenotypes and can differentiate to either monocyte-derived macrophages (MDMs) or monocyte-derived dendritic cells (moDCs). These mononuclear phagocytes (MPs) drive pathogen clearance[47,48], contribute to the induction of adaptive immunity through antigen presentation[49], and coordinate innate immune functions through cytokine release[50]. MPs play a critical role in driving protection against a multitude of bacterial[51], viral[52,53], fungal[54], and parasitic infections[55]. However, dysregulation and

overactivity of MPs can also have detrimental effects against infection leading to chronic infection, pathology, and enhanced disease severity [56–58].

Monocytes mediate their functions either directly following non-specific recognition of pathogens or indirectly following interactions with opsonins such as complement proteins or antibodies that decorate the surface of pathogens. Direct engagement occurs non-specifically through pattern recognition receptors (PRR) such as toll-like receptors (TLR) and C-type lectin receptors (CLR) which recognize pathogen associated molecular patterns (PAMP). In addition to PRRs, monocytes and monocyte-derived cells also express Fc-receptors (FcR) and complement receptors (CR) that allow these cells to bind to opsonized material[59,60]. Thus, antigen-specific antibodies can form immune complexes that can then trigger FcRs or CRs on the surface of monocytes to drive clearance, pathogen control, and cellular activation. Following FcR engagement monocytes are able to engulf foreign material, produce oxygen and nitrogen reactive species, generate anti-bacterial proteases, and release DNA to trap and eliminate pathogens[61,62]. FcRs bind to different antibody isotypes, subclasses, and Fc-glycosylation profiles with different affinities[63]. Monocytes and monocyte-derived cells can express FcαRI, FcγRI, FcγRIIa, FcγRIIb, and FcγRIIIa depending on the level of cellular activation, maturation, phenotypic state, or organ location[64,65]. Therefore, variation in the composition of antigen specific antibodies likely leverage monocytes differentially, contributing to distinct levels of pathogen clustering, uptake, destruction, signaling, and presentation.

Previous efforts to quantify antibody dependent cellular phagocytosis (ADCP) has focused on the use of bead-based assays using THP-1 cells[66], a monocytic cell-line[67]. The assay has been used to investigate the role of ADCP against many diseases[58,68–70]. Although the ADCP assay has been useful for identifying differences in antibody mediated phagocytosis, monocytes can also respond to infection through distinct functions, not currently captured by the ADCP assay, including downstream functions, changes in activation, maturation, cytokine secretion, and antigen presentation. Therefore, an adapted antibody dependent phagocytosis assay, performed with human primary monocytes, could provide key insights into the role of humoral immunity in modulating monocyte activity more broadly.

Here we describe a multifaceted antibody dependent human monocyte phagocytosis (ADMP) assay using antigen-coupled fluorescent beads for investigating ADMP and downstream functions against any given antigen. This high-throughput assay measures antibody mediated phagocytosis by monocytes as well as secondary responses. Comparison of the ADMP assay to the ADCP assay showed superior sensitivity and identified additional unique secondary effects of immune complexes not identifiable by ADCP.

2.3 Results

2.3.1 Developing an Antibody-dependent primary monocyte phagocytosis assay

The ADMP assay can be split into three main parts: 1) the formation of immune complexes, 2) the isolation of monocytes, and 3) the phagocytic uptake of immune complexes by monocytes (Fig. 2-1A). First antigens were biotinylated, and any unbound biotin was washed off. Then biotinylated antigens were coupled to fluorescent neutravidin coated beads, and any excess antigen was washed off. Following antigen-bead coupling, immune complexes were formed by incubating either monoclonal antibodies or plasma with the antigen coated fluorescent beads. In addition, no antibody or other non-specific control antibodies were used as negative controls to capture background phagocytosis mediated by primary monocytes. During the formation of immune complexes primary monocytes were isolated from healthy donors. Following immune complex formation unbound antibodies were washed off and then co-cultured with the primary human monocytes during which phagocytic uptake took place. After phagocytosis, supernatants were collected for downstream analysis, beads that were not phagocytosed were washed off, and cells were stained and fixed.

Antibody-mediated phagocytosis by primary monocytes and relative changes in surface marker abundance were then quantified by flow cytometry. First, single cells were gated (FSC-A/FSC-H) and then monocytes were identified by size (FSC-A/SSC-A) and either CD14 or CD16 expression (Fig. 2-1B). Phagoscores (PS) were then determined by calculating the percent of cells that internalized beads multiplied by the geometric mean fluorescent intensity (gMFI), which corresponds to the average bead uptake. Relative changes in surface expression of maturation/activation markers were also captured by comparing relative changes in sample conditions compared to antibody controls, and thus providing additional insights into subsets of phagocytosing cells and downstream consequences of phagocytosis.

To define potential differences in phagocytic activity between primary monocytes and THP-1 cells, the ADMP assay was first optimized using 27 serum samples from healthy individuals profiled against the influenza HA antigen. First, incubation time was optimized, reported as a phagocytic score (PS) over a no antibody control condition (background) providing an assay signal-to-noise (SNR) ratio (Fig. 2-1C). This revealed that the four-hour time point provided the best SNR for the ADMP assay. Moreover, with viabilities >97%, the 4-hour window showed ideal performance (Fig. 2-S1A). Although significant differences were observed between monocytes incubated with both antigen-coated beads and antibodies compared to those only incubated with antigen-coated beads (no antibody) or uncoated beads (no antigen), effect sizes were negligible (<1%). The optimal serum dilution for Influenza HA1 responses using the ADMP was next determined by performing serial dilutions ranging from 1:256 to 1:32768 (Fig. 2-1D). This revealed the dose dependent manner of ADMP and pointed to the lowest dilution as the optimal dilution for maximizing separation across samples. Serial dilutions with

the ADCP assay showed a narrower range of dilutions that were able to distinguish samples compared to ADMP, and by a dilution of 1:4096 only the ADMP assay was able to distinguish most samples (Fig. 2-1D and Fig. 2-S1B). However, the optimal dilution for the ADMP assay could vary across antigen and sample set and should be predefined accordingly.

Bead and cell concentrations were next optimized using two different serum pools across three donors and two replicates (Fig. 2-1E). The condition with the lowest cell concentrations, 25,000 cells/well, and the highest bead concentration showed the highest signal to noise ratio (SNR) of approximately 15. This same condition when repeated with the ADCP assay yielded an SNR of 5.3, a three-fold decrease (Fig. S-1C). To evaluate the robustness of the assay, PS correlations among donors were determined (Fig. 2-1F and Fig. 2-S1D-E). This revealed that although phagoscores shifted slightly across donors (Fig. 2-1E), their signals were highly concordant. Regardless, a minimum of at least two donors should be used to profile ADMP to ensure that a specific donor does not exhibit a discordant effect. Additionally, from a single buffy coat approximately 4000 samples can be profiled. Although the ADCP assay is not limited by donor cell number constraints, the vast majority of functional primary cell-based assays are limited by the number of effector cells that can be acquired for functional analysis. In contrast, 4000 conditions provide an opportunity to run a considerable number of samples, or multiple dilutions for each sample, and far exceeds the numbers of conditions that can currently be run with traditional primary Natural Killer cell assays. Thus, this assay is considered high-throughput compared to other primary cell based assays, albeit it is medium-throughput compared to cell-line based assays.

Finally, prior to flow cytometry, trypsin-EDTA was used to resuspend monocytes which strongly adhere to plates. Trypsin has previously been shown to digest certain sensitive surface proteins[71]. Therefore, although the treatment of trypsin-EDTA was brief, its effect on CD14 and CD16 was evaluated to ensure that the short incubation would minimally impact the gMFI of tested surface markers (Fig. 2-S1F). However, for additional markers, a rapid analysis should be included to ensure that the trypsin-EDTA does not affect additional surface markers. Ultimately, these results highlight the development and overall performance of the ADMP assay.

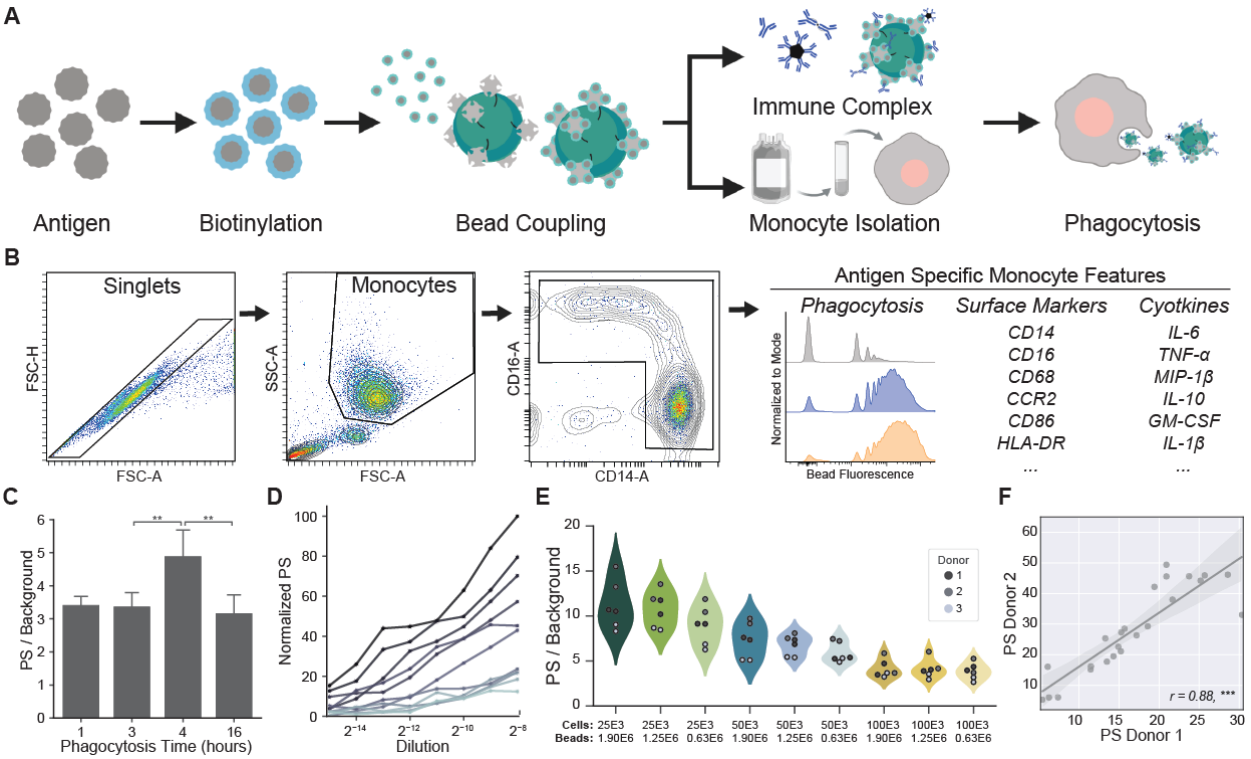


Figure 2-1. Antibody dependent human monocyte phagocytosis assay (ADMP) process and optimization

(A) Schematic illustrating the ADMP assay process. First, antigen is biotinylated, coupled to fluorescent neutravidin beads, and then immune complexes are formed with an antibody sample. In parallel, primary monocytes are isolated from donor buffy coats. Finally, cells and immune complexes are incubated together. (B) Gating strategy and ADMP readouts. Events are first gated for single cells with FSC-A and FSC-H, then monocytes are first gated on FSC-A and SSC-A and then confirmed with either CD14 or CD16 expression. This yields antigen-specific monocyte features which include a phagoscore (PS), and surface expression changes and cytokine secretion due to immune complex stimulation. (C) The effect of different monocyte and immune complexes incubation times on PS. Each bar is comprised of Influenza HA1 responses from 9 healthy individuals which were averaged over 2 technical replicates and 2 donor replicates. The y-axis represents phagoscores over a no serum control (background), error bars depict +1 SD from the mean, and significance was determined by a Mann–Whitney U Test ($p<0.005$: **). (D) Serial dilutions of 9 different samples against HA1 ranging from 1:256 (2-8) to 1:32768 (2-15). All curves were first background subtracted and then min-max normalized. Each color represents a different individual. (E) PS over background of different cell and bead conditions consisting of 2 different sera pools across 3 donors, showing an average of 2 technical replicates. The color of the dots corresponds to the donors, and the number of cells and beads correspond to the amount suspended in each 200µL well of a 96 well plate. (F) Donor correlation of PS from 24 different samples. Each PS is the average of 2 technical replicates. The data was fit with a linear regression model. The shadowed area represents the 96% confidence interval and the Pearson correlation coefficient (ρ) with the corresponding significance are reported ($p<1e-6$: ***).

2.3.2 Anti-spike SARS-CoV-2 antibodies trigger changes in surface marker expression in hospitalized individuals

Beyond the ability to quantify differences in phagocytic activity across plasma samples, primary monocytes also provide an opportunity to analyze the downstream consequences of phagocytosis, via the examination of cellular activation or maturation. To begin to define whether the ADMP assay would be suitable to a more comprehensive monocytic profiling, we focused on a sample set from Coronavirus Disease 2019 (COVID-19) patients, 40 of which experienced mild disease and 20 that experienced severe disease[72]. This cohort was chosen as it had previously been shown to exhibit significantly different ADCP activity across the groups using the traditional THP-1 assay.

Using primary monocytes, the ability of antibodies from COVID-19 patients to drive phagocytosis and alter the surface expression of CD14, CD16, CD68, and CCR2 on monocytes, following immune complex activation, was assessed. As previously observed, PS activity against the SARS-CoV-2 spike antigen was significantly higher in the hospitalized group (Fig. 2-2A), consistent with previous ADCP data[72]. Furthermore, although hospitalized individuals exhibited higher relative antibody levels, they did not correlate with ADMP phagoscores (Fig. 2-S2A), pointing to qualitative changes in the SARS-CoV-2 specific humoral immune response, rather than quantitative differences as a key determinant of ADMP activity. Conversely, antibody levels and phagoscores were correlated among non-hospitalized individuals, pointing to disease-state specific differences in antibody functionality. Additionally, CD14, CD16, and CD68 upregulation was observed in in the presence of antibodies from hospitalized individuals (Fig. 2-2B). However, CCR2 displayed a distinct profile, marked by lower levels in the setting of plasma from hospitalized patients. The data collectively suggest that antibody levels alone represent an incomplete predictor of ADMP and downstream activating potential.

Although, striking differences were noted in phagocytic activity as well as CD16 and CD68 across both the ADMP and ADCP assays, more robust separation in patient-profiles were observed with the ADMP assay (Fig. 2-2C). Yet, only the ADMP assay, using primary monocytes, detected differences in altered CD14 and CCR2 expression across the two groups. Additionally, immune complex stimulation and phagocytosis resulted in opposite shifts in CD14 and CCR2 expression highlighting discordant behavior of THP-1 cells and primary monocytes. Thus, the use of primary monocytes may provide additional insights on immune complex mediated cellular activation, capturing downstream consequences of immune complex-uptake and stimulation.

To further understand the relationships between features captured by each of the assays, pairwise spearman correlations were drawn between PS and alternations in surface marker expression (Fig. 2-2D). THP-1 mediated phagocytosis was strongly correlated to changes in CD68 expression and fairly correlated to CD16 levels. Conversely, THP-1 mediated PS was inversely correlated to CD14 levels and CCR2 expression (Fig. 2-2D). However, monocyte mediated PS, in the ADMP assay, was strongly correlated with changes in CD14, CD16, and CD68 as well as strongly inversely

correlated to CCR2 expression. To summarize the relationship between features across the assays, spearman correlations were used, highlighting concordance across PS, CD68, but more variability across CD14 and CCR2 (Fig. 2-2E). Therefore, the ADMP assay may more sensitively capture coordinated changes in response to antibody-activation on monocytes compared to THP-1 cells.

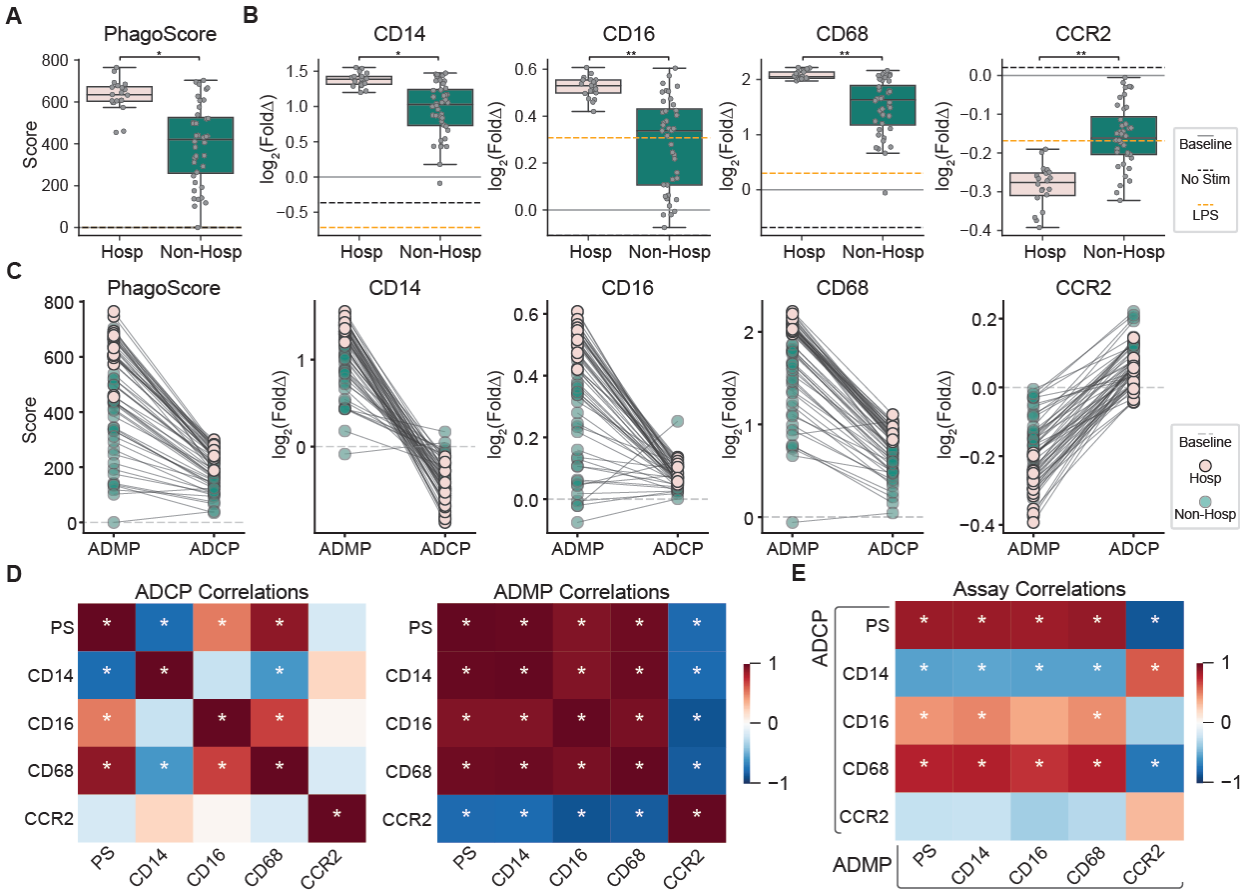


Figure 2-2. Unique surface expression differences in SARS-CoV-2 hospitalized individuals by ADMP

(A) Phagoscores (PS) and (B) shifts in surface expression of CD14, CD16, CD68, and CCR2 from 20 hospitalized and 40 non-hospitalized SARS-CoV-2 infected individuals profiled against the spike protein. Each response is an average across 3 donors and 2 technical replicates. PS were background subtracted using the no serum control, and surface expression differences are represented as log2 fold changes over the no serum control. The solid gray line represents the baseline, the dashed black line a no stimulation control, and the dashed orange line a 100 ng/mL LPS stimulation control. Significance was determined by a Mann–Whitney U Test (p<1e-3: *, and p<1e-6: **). (C) Paired dotted plot illustrating differences in readouts across the ADMP and ADCP assays. The pink and green dots represent hospitalized and non-hospitalized individuals respectively, and the dashed gray line represents the baseline. (D) Heatmap of ADCP (left) and ADMP (right) intra-assay spearman correlations between PS, and changes in surface marker

expression. (E) Spearman correlation heatmap of ADCP and ADMP PS and fold change surface expression across assays. Statistical significance is indicated by the white asterisks and was corrected for multiple hypotheses with the Bonferroni method (p<1e-4: *). Panels C and D were visualized with the data shown in A and B and therefore were normalized in the same manner.

2.3.3 Downstream ADMP activity distinguishes hospitalized SARS-CoV-2 patients

Given the surface marker differences between THP1 cells and primary monocytes, we next aimed to determine whether additional components of downstream activation could be resolved using the ADCP or ADMP assays. Following phagocytosis, immune complexes may trigger a network of Fc-receptors, resulting in signaling cascades, that may lead to cellular activation and cytokine or granule release.

Thus, to determine whether downstream activation could be captured across the monocyte assays we utilized samples from 28 individuals infected with SARS-CoV-2 who experienced mild disease (non-hospitalized), and 12 samples from individuals that experienced severe disease (hospitalized). Analysis of supernatants at 16 hours following co-culture with SARS-CoV-2 spike specific immune complexes revealed little to no cytokine secretion using the THP-1 ADCP assay (Fig. 2-3A). Conversely, THP-1 cells did release cytokines following LPS stimulation (Fig. 2-S2B). In contrast, robust cytokine secretion was noted in the ADMP assay, and revealed significantly higher IL-6 and MIP-1β production in the presence of plasma samples from hospitalized individuals compared to non-hospitalized individuals (Fig. 2-3B). Moreover, the majority, but not all, cytokine levels were enhanced in the presence of antibodies compared to the no antibody controls, albeit the levels were not as strong as those observed with LPS stimulation (Fig. 2-3B). These data ultimately highlight a role for antibodies in monocyte-driven cytokine secretion.

In order to determine whether a multivariate set of features measured by each assay was able to distinguish hospitalized and non-hospitalized individuals a partial least square discriminant analysis (PLS-DA) was used. A minimal set of features consisting of cytokine secretion profiles, surface marker expression, and phagocytosis were able to resolve hospitalized and non-hospitalized COVID-19 patients across both the ADCP and ADMP assay (Fig. 2-3C and D), albeit the data from the ADMP assay provided stronger resolution (Fig. 2-3D). Using the ADCP assay, variable importance in projection (VIP) scores selected PS as the feature which best differentiated the 2 groups, enriched among hospitalized individuals (Fig. 2-3C). Additionally, IL-1β and myeloperoxidase (MPO) were enriched in non-hospitalized patients and were also used for resolving the 2 groups based on data collected from the ADCP assay. Conversely, chemokine receptors, Fc-receptors, and cytokine levels were critical discriminators of patient groups using data collected in the ADMP assay (Fig. 2-3D). To investigate the relationship between PS, surface marker changes, and antibody mediated cytokine secretion pairwise spearman correlations were used and revealed weak correlations among most of the ADCP features (Fig. 2-3E). The ADMP assay was alternatively marked

by robust correlations between several LASSO selected features and cytokine levels in the ADMP assay (Fig. 2-3F).

We next aimed to examine whether particular classes of data from the ADCP or ADMP assay were more discriminatory across the groups. This was achieved through cross validation while maintaining the same training and test set identity for all models during each iteration to determine differences in classification accuracy (Fig. 2-3G). This revealed that ADMP models built on either cytokine or surface marker features both greatly outperformed their ADCP model counterparts. Additionally, the selected model built on the ADMP features outperformed the ADCP assay. Ultimately, these analyses highlighted the utility of the ADMP assay which provided enhanced insight into post phagocytic immune consequences and was able to better distinguish clinical groups.

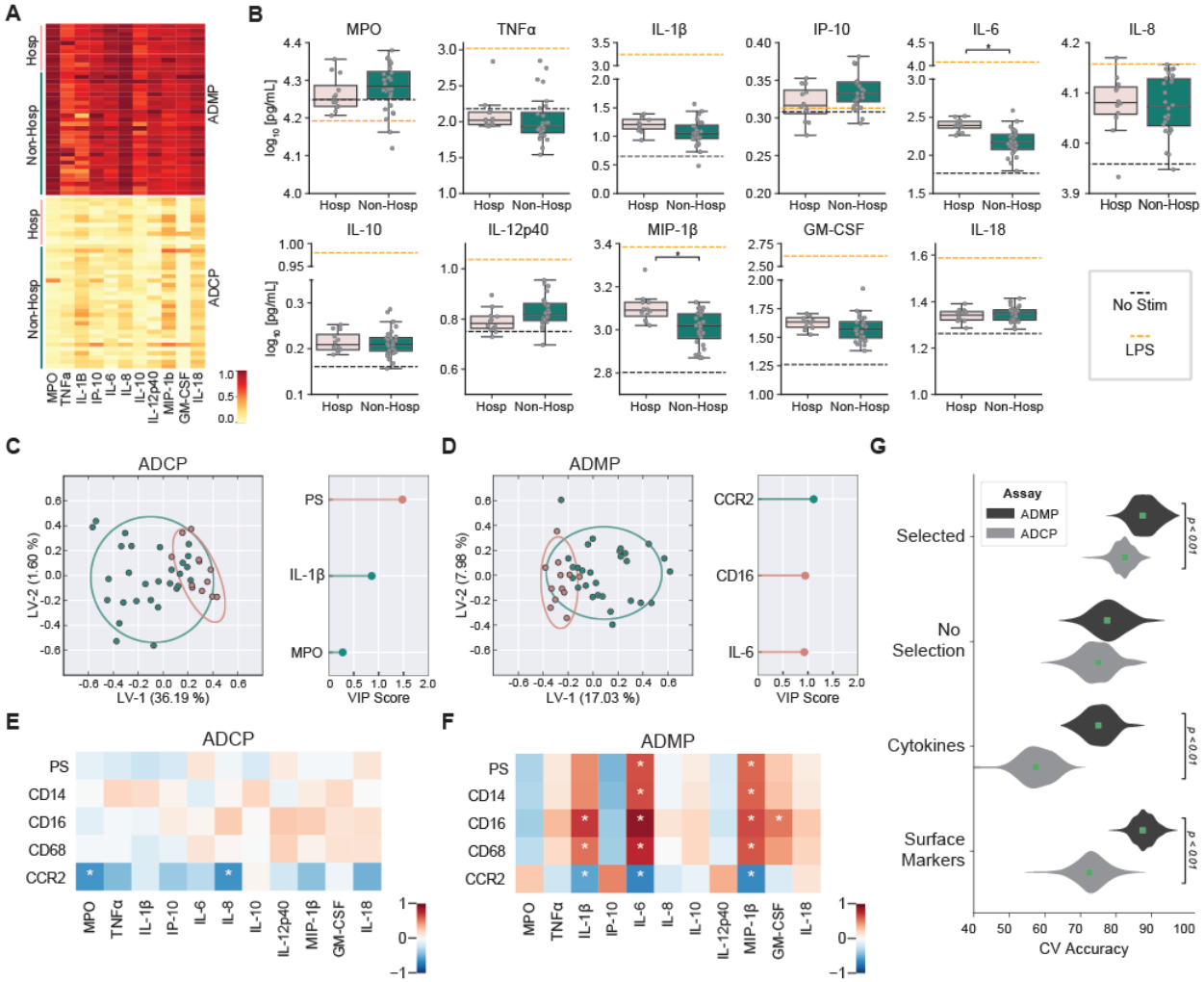


Figure 2-3. ADMP assay provides greater spectrum of post phagocytic detail

Post-phagocytic cytokine analysis of 12 hospitalized and 28 non-hospitalized SARS-CoV-2 infected individuals. (A) Heatmap depicting cytokine secretion of hospitalized and non-hospitalized individuals across the ADCP and ADMP assays, and (B) boxplots depicting the ADMP cytokine differences among hospitalized and non-hospitalized individuals. The dashed black line

represents the no stimulation control and the dashed orange line the 100 ng/mL LPS stimulation control. (C) Cross validated lasso selection models based on ADCP features and (D) ADMP features respectively. The scores plot of the PLS-DA scores plots are based on the final lasso selected features with green corresponding to the non-hospitalized and purple to the hospitalized groups. (E) Spearman correlation heatmap of PS, surface expression, and cytokine features for ADCP and (F) ADMP respectively. Statistical significance is indicated by the white asterisks and was corrected for multiple hypotheses with the Bonferroni method (p<1e-4: *). (G) Cross validated classification accuracy of PLS-DA models built on different sets of features. Models were either built using either lasso selected, all cytokine, all surface marker, or all features. Cross validation sets were synced across models and statistical significance was defined by exact P values of the tail probabilities of the ADMP distributions within the ADCP distributions. All cytokine concentrations were converted using a standard curve, and then log10 transformed. Additionally, A and C-G were z-scored. Each response is an average of 2 technical replicates.

2.3.4 The ADMP assay reveals distinct differences in Fc-variant performance

Recent studies have shown that monoclonal Fc-variants can selectively trigger specific effector functions[73,74], likely due to their ability to bind and signal differentially via Fc-receptors[63]. While attention has been paid to the ability of antibodies to trigger phagocytosis through flow cytometry or microscopy[66,75]. These assays have focused largely on uptake, failing to capture potential additional differences in Fc-induced monocyte and macrophage activation that may also be key in therapeutic activity. Thus, we aimed to determine whether the ADMP assay, with the added downstream readouts, could further define differential monoclonal antibody Fc-activities. To this end, an Fc-engineered library of the SARS1 antibody CR3022, which is cross-reactive to the receptor binding domain of SARS-CoV-2 was used[76], as it has previously been shown to drive ADCP[74].

Comparison of phagocytic activity of the wildtype antibody was first assessed across both the ADCP and ADMP assays and showed stronger phagocytic activity in the ADMP assay (Fig. 2-4A). Moreover, significant differences were observed across surface markers and assays with higher CD14 and CD86 expression and lower overall CCR2, CD16, CD32, CD64, and CCR2 in the ADMP assay compared to the ADCP assay. Additionally, cytokine secretion also revealed additional nuances across the Fc-variants in the ADMP assay. Despite the limited cytokine secretion in the ADCP assay, robust cytokine secretion was still observed in both assays following LPS treatment (Fig. 2-S3). Integration of all phenotypic and cytokine responses in the ADMP assay resolved distinct differences in Fc-variant driven activation of primary monocytes (Fig. 2-4B). Specifically, hierarchical clustering of the variants revealed distinct groups that were able to trigger phagocytosis, change surface marker expression, and alter various cytokine secretion profiles. As expected, LALA PG, a relatively inert variant, clustered with other negative controls, and showed the lowest levels of phagocytosis and cytokine secretion. In addition, clusters of variants with high, medium, and low phagocytic scores with various

cytokine secretion profiles were observed. Of note, SDIEGA-LS showed the strongest phagocytosis and cytokine secretion. Conversely, IgG1, the unaltered variant, showed one the lowest phagoscores but a strong cytokine secretion profile. These data therefore highlight opportunities of exploring Fc-engineering to selectively drive phagocytosis with or without specific activation. Collectively, these profiles provide a map for various responses induced by changes in Fc-structure only observable through the ADMP assay. Ultimately, the ADMP assay provides a distinct assay able to capture monocyte changes that are induced by immune complex stimulation.

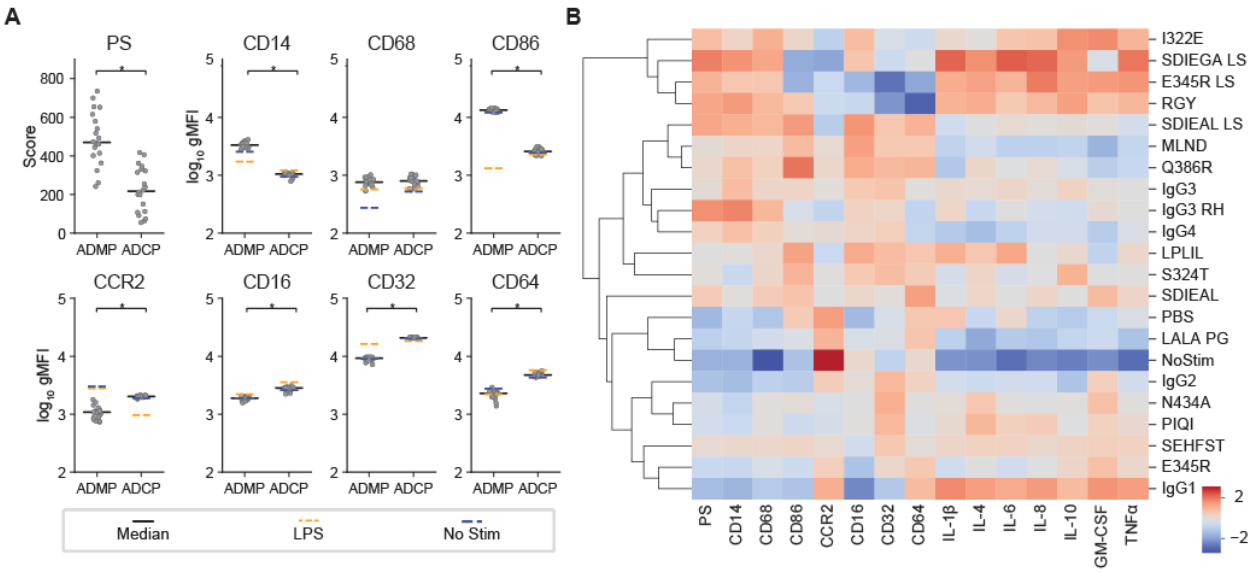


Figure 2-4. CR3022 Fc-variants illicit unique post-phagocytosis responses

(A) Differences in phagocytosis and surface expression of CD14, CD68, CD86, CCR2, CD16, CD32, and CD64 are depicted with each dot representing 1 of 20 different CR3022 monoclonal Fc-variants targeting the receptor binding domain (RBD) of the SARS-CoV-2 virus. Phagoscores were background subtracted using no serum controls, and surface expression is represented as log10 of the geometric mean fluorescent intensity (gMFI). Significance was determined by a Mann–Whitney U Test and corrected for multiple hypotheses with the Bonferroni method (p<1e-3: *). (B) The heatmap and dendrogram depicts the hierarchical clustering of the different variants using phagoscore (PS), surface expression, and cytokine data. Clustering was performed using a correlation distance metric with an 'average' linkage method. Prior to clustering all data was z-scored and is an average of 2 technical replicates.

2.4 Discussion

Fc-effector functions, including antibody dependent cellular cytotoxicity (ADCC)[77], antibody dependent complement deposition[78], and antibody dependent phagocytosis[79], have been implicated in protective immunity across both infectious[80–82] and non-infectious disease[83–85]. Specifically, ADCP has been linked to protection against simian immunodeficiency virus (SIV) in non-human primate studies (NHP)[68], and has been linked to natural resolution of COVID-19[58]. However, human primary monocytes are

heterogeneous and highly plastic and can differ both in phenotype and function under distinct inflammatory conditions as well as in different compartments[86,87]. For simplicity, ADCP assays often exploit monocytic cell lines, including the THP-1 cell line[67] that may not fully recapitulate the heterogeneity of potential immune functions that may emerge from monocytes in vivo. Thus, given the critical role of monocytes, as key innate immune effector cells, an assay using primary cells able to capture monocyte activity more broadly could provide novel insights into mechanisms of protection to guide therapeutic or vaccine design.

The ADMP assay provided several distinct advantages over the traditional ADCP assay. First, primary monocytes showed enhanced sensitivity and greater dynamic range when measuring antibody mediated phagocytosis across samples. This improvement in sensitivity could be utilized to spare both plasma samples and antigenic material. Furthermore, although differences in phagocytosis between hospitalized and non-hospitalized SARS-CoV-2 individuals were detectable by both assays, this may not be the case for cohorts with more nuanced disease phenotypes. Additionally, the surface expression of many markers have previously been shown to differ between THP-1s and primary monocytes[87]. Phagocytosis is a multi-stage process which is initiated by phagocytic receptors which are often engaged cooperatively[88]. Fc-receptors in particular play an important role in shaping cellular responses[89,64,61], and although primary cells and THP-1s have been shown to express FcγRI and FcγRII, the expression and signaling via these markers differ on primary monocytes. While differences in Fc-receptor expression may not necessarily alter the ability to phagocytose (Fig. 2C), changes in Fc-receptor expression and engagement appear to alter the downstream consequences after phagocytosis, including the ability to secrete cytokines, react to stimuli, and alter surface expression of specific markers (Fig. 2C and 3A). Thus, the ADMP assay can both provide information regarding enhancement of antibody mediated phagocytosis, but also drive changes in surface markers and cytokine profiles that more accurately represent mechanisms *in situ*.

In addition to improved sensitivity and the ability to describe downstream consequences of phagocytosis, the use of primary monocytes also allows for the investigation of disease-associated states and whether certain perturbations alter responses to immune complexes. Perturbed monocytic states have been previously associated with many diseases[45,90,91], and although donor variability was low among healthy donors (Fig. 1F and Fig. S1D-E) ADMP assays performed with perturbed monocytic cell states could provide different results. Therefore, monocytes isolated from donors which exhibit disease-specific characteristics could uniquely respond to immune complexes and thus provide additional novel mechanistic insights into disease specific immunopathology.

Beyond, differences across polyclonal sera the ADMP assay also highlighted unique responses to monoclonal Fc-variants. Monoclonal therapies in the past decade have become an important modality in the treatment of cancer[92] and against many

infectious diseases[93–95]. The therapeutic effects of antibodies are linked to both the antigen binding region (Fab) and the constant region (Fc) of the antibodies, the latter that can be engineered to selectively trigger specific Fc-effector functions[96,97]. However, while particular Fc-variants have been identified with enhanced ADCP activity, how these variants affect primary monocyte maturation and activation is not well understood. Using a panel of 20 Fc-variants of the CR3022 monoclonal antibody that recognizes SARS-CoV-2, a number of Fc-variants were identified that could drive enhanced phagocytosis, but also tune monocyte biology (Fig. 4B). Interestingly, the wildtype IgG1 induced relatively poor phagocytosis but potent inflammatory cytokine secretion. This observation highlighted the capacity of certain structural variants to uniquely trigger cytokine profiles and phagocytosis independently. This could prove to be useful in certain cancer therapeutic contexts where monocyte and macrophage function could be tuned to enhance the elimination of tumor cells. Thus, the ADMP assay could provide a new method for evaluating the efficacy of monoclonal therapies.

Additionally, the ADMP assay being a bead-based assay allows for the multiplexing of antigens[98]. While multi-bead uptake has been shown to recapitulate single-plex signals, offering a sample-sparing approach to perform multiplexed ADCP assays, competition assays with antigens of interest are also possible. Therefore, the ADMP assay may be further modified to capture individual antigen responses, multiplexed activity, and even competitive responses. Moreover, the assay may be modified to used cryo-preserved monocytes, that if thawed properly monocyte viability, phagocytic activity, morphology, migration, adherence, and other functions minimally impact monocyte activity[99–103]. Therefore, the use of cryopreserved monocytes could provide additional flexibility when fresh monocytes are not available such as in the case of when profiling perturbed monocytes from non-healthy donors.

Ultimately, the ADMP platform described here provides a method for reproducibly probing antibody mediated primary monocyte function and is distinct from the traditional ADCP assay[66]. This methodology could enable improved vaccine design, the development and evaluation of monoclonal therapeutics, and enhance the understanding of the mechanisms of humoral protection by probing both the level of antibody mediated immune complex clearance as well as the downstream consequences of this activity.

2.5 Methods

2.5.1 Source of samples and monocytes

Primary human monocytes were isolated from the buffy coats of 20 healthy volunteers at the Ragon Institute of MGH, MIT, and Harvard. Serum samples from 27 orthogonal healthy individuals were also obtained from the Ragon Institute of MGH, MIT, and Harvard and were used as a source of antibodies for assay optimization. Additionally, serum samples from 40 non-hospitalized SARS-CoV-2 infected outpatients and 20 previously hospitalized SARS-CoV-2 infected inpatients were profiled. The cohort

consisted of 60 adults, over the age of 18, collected at Harborview Medical Center, University of Washington Medical Center or at Northwest Hospital (Seattle, Washington, USA) between 24- and 74-days following symptom onset[72]. Samples were collected as part of a prospective longitudinal study or as part of a protocol for expanded access to convalescent plasma treatment of COVID-19 (ClinicalTrials.gov, NCT04338360), and is further explained in[104]. This study was conducted in accordance with the World Medical Association's Declaration of Helsinki, was approved by the MGH Institutional Review Board, and all volunteers – across all sample collections - provided informed consent.

2.5.2 Monoclonal antibodies

Twenty CR3022 Fc-variants were created by designing gene blocks containing the Fc domain of IgG1 with Fc point-mutant selected based on known differences in binding to Fc receptors[73,74,97,105,106]. First, pUC donor plasmids encoding various Fc regions, a variable heavy chain, a furin P2A sequence, and variable light chain flanked by BsaI at each site were designed. Then, a destination vector was cloned that included a kappa light chain sequence, an IL-2 secretion signal, and the ccdB suicide gene by BsaI sites. Finally, donor and destination plasmids were digested and ligated with Golden Gate cloning which could produce IgG proteins with the same antigen-binding (Fab) region but differences in Fc regions. Monoclonal Fc variants were produced in 293F suspension cells grown in FreeStyle 293 Expression media (Gibco, Thermo Fisher Scientific). 293F cells were transfected with Polyethylenimine (PEI; Polysciences) at 1 µg/µL in a ratio of 3:1 µg of PEI to DNA. After transfection, cells were incubated for 5 days after which supernatants were collected and then purified with protein G magnetic beads (MilliporeSigma). Assays that were performed using monoclonal antibodies were run at a concentration of 2.5 µg/mL.

2.5.3 Isolation of primary human monocytes

Primary human monocytes were obtained from the buffy coats of healthy donors (MGH blood donor center). Buffy coats were diluted 1:4 in 2% fetal bovine serum (FBS; Sigma) in a phosphate buffered saline (PBS; Sigma-Aldrich) by volume and set at room temperature overnight. Then using sepmate tubes (Stemcell) and following the manufacturer's instructions peripheral blood mononuclear cells (PBMCs) were isolated. PBMCs were then pelleted by centrifugation (300 x g, 8 min) at room temperature and then washed with 2% FBS in PBS. If any residual erythrocytes remained, cells were diluted in 1:10 ratio of ammonium-chloride-potassium (ACK) lysis buffer (150mM NH4Cl, 10mM KHCO3, 0.1 mM Na2EDTA, pH 7.4) by volume and was incubated for 5 min at room temperature. To wash off any residual lysis buffer, cells were pelleted with centrifugation (500 x g, 5 min) at room temperature, and the buffer was poured off and replaced with 2% FBS in PBS. The wash step was repeated to ensure the removal of any residual lysis buffer. PBMCs were then adjusted to a concentration of 50 x 10^6 cells/mL in cold 2% FBS and 1mM EDTA in PBS and using the human monocyte isolation kit

without CD16 depletion (Stemcell) monocytes were isolated. The purified monocytes were then finally diluted in R10 media (RPMI-1640 media (Sigma) with 10% FBS, 2 mM L-Glutamine and 100 U/mL penicillin/streptomycin)

2.5.4 Antigen coupling to fluorescent beads

The SARS-CoV-2 spike, SARS-CoV-2 receptor binding domain (Lake Pharma), or influenza hemagglutinin (HA) (California H1N1 2009, Immunetech) antigens were biotinylated with sulfo-NHS-LC-biotin (Thermo Scientific) using the manufacturer's instructions. Then Zeba spin desalting columns (Thermo Fisher) were used to remove unbound biotin through buffer exchange to PBS. Biotinylated antigens were then separately coupled to blue 1.0 µm fluorescent neutravidin beads (Thermo Fisher) at a ratio of 10µg of biotinylated antigen to 10µL of the neutravidin beads. Beads and biotinylated antigens were coupled at 37ºC for 2 hours in low-binding microcentrifuge tubes (Corning). Beads were then washed twice with 0.1% PBS-BSA through centrifugation (16,000 x g, 5 min) at room temperature to remove unbound antigen. The antigen-coupled fluorescent beads were resuspended 1:200 in 0.1% PBS-BSA by volume and stored for up to no more than a week at 4ºC in the dark.

2.5.5 Immune complex formation

First, to form immune-complexes, in each well of a 96-well ultra low attachment round bottom microplate (Corning) 20µL of samples were diluted 1:200 in 0.1% PBS-BSA and incubated with 20µL of the antigen coupled beads for 2 hours at 37ºC. Profiling of CR3022 Fc-variants was performed using 20µL the monoclonals at 2.5 µg/mL diluted in 0.1% PBS-BSA. Additionally, a PBS no antibody control, a no bead control, and a 100 ng/mL bacterial lipopolysaccharide (LPS) control were included in each experiment. After the formation of immune complexes, 200 µL of 0.1% PBS-BSA was added to each well, pelleted (1000 × g, 10 min), and the supernatant was removed to clear unbound antibodies.

2.5.6 Antibody dependent cellular phagocytosis assay

The antibody dependent cellular phagocytosis (ADCP) assay was performed as previously described in[66]. Antigen-bead conjugation, and immune complex formation was performed in the manner described above. THP-1 cells (ATCC), a monocyte cell line, was adjusted to a concentration of 1.25 x 10^5 cells/mL in R10, and 200µL (25,000 cells/well) of cells were added to each well. The immune complexes and cells were incubated at a ratio of 1:76 (cells/beads), the equivalent of 2.4 million cells per 10µL of the bead stock solution, for 16 hours at 37ºC in a humidified incubator (5% CO2). After phagocytosis cells were pelleted (500 x g, 5 min). Downstream analysis of antibody specific cytokine secretion profiles was performed on supernatants that were collected into a separate 96-well round bottom microplate (Corning), centrifuged at 4ºC (1000 x g, 15 min), and then stored at -80ºC until use. For cellular analysis, cells were first washed with 50µL of Trypsin-EDTA (0.25%, Thermo Fisher) and set for 5 min at 37ºC.

Next, 150µL of 2mM EDTA was added to each well, pelleted (500 x g, 5 min), and washed twice more with 200µL of 2mM EDTA. Cells were then incubated in 20µL of Fc-block (Miltenyi Biotec) which was previously diluted 1:10 in 0.1% PBS-BSA by volume and set for 15 min in darkness. Subsequently, cells were stained for CD14 and CD16 expression at room temperature for 30 min in the dark. Additional stains including maturation markers (CD86 and CD68), Fc-receptors (CD32 and CD64), and a chemokine receptor (CCR2) were included depending on the experiment. Following staining, cells were washed twice with 2mM EDTA in PBS (500 x g, 5 min), fixed with 4% paraformaldehyde (PFA; Alfa Aesar), and washed twice more to remove any residual PFA. Cell analysis was performed on a BD LSR Fortessa flow cytometer, and analyzed using FlowJo (FlowJo, LLC). THP-1 cells were gated by size (FSC/SSC) and singlets (FSC-A/FSC-H). Phagoscores were calculated by taking the geometric mean fluorescence intensity (gMFI) of the bead positive monocytes, multiplying by the percent of bead positive monocytes, and dividing by an arbitrary factor for interpretability. Changes in expression were reported as log2 MFIs over PBS controls. Cytokine secretion by monocytes upon immune complexes stimulation was analyzed by using the previously collected supernatants which were thawed on ice and measured with a customized multiplex luminex bead panel (Thermo Fisher) following the manufacturer's protocol.

2.5.7 Antibody dependent primary monocyte phagocytosis assay

The antibody dependent primary monocyte phagocytosis (ADMP) assay was performed by first purifying primary monocytes, described above. The monocyte concentration was adjusted to a concentration of 1.25 x 105 cells/mL in R10, and 200µL (25,000 cells/well) of cells were added to each well. The immune complexes and cells were incubated at a ratio of 1:76 for 4-16 hours at 37ºC in a humidified incubator (5% CO2). For the analysis of surface marker expression and phagocytosis incubation periods of 4 hours were used, and for cytokine and chemokine secretion incubation times of 16 hours were used.

After cells were pelleted, supernatants were collected as described above. To ensure all cells were detached 50µL of Trypsin-EDTA (0.25%, Thermo Fisher) was added to each well and set for 5 min at 37ºC. Following detachment, 150µL of 2mM EDTA was added to each well, pelleted (500 x g, 5 min), and washed twice more with 200µL of 2mM EDTA. Cells were then incubated in 20µL of Fc-block (Miltenyi Biotec) which was previously diluted 1:10 in 0.1% PBS-BSA by volume and set for 15 min in darkness. Subsequently, cells were stained for CD14 and CD16 expression at room temperature for 30 min in the dark. Additional stains including maturation markers (CD86 and CD68), Fc-receptors (CD32 and CD64), and a chemokine receptor (CCR2) were included depending on the experiment. Following staining cells were washed twice with 2mM EDTA in PBS (500 x g, 5 min), fixed with 4% paraformaldehyde (PFA; Alfa Aesar), and washed twice more to remove any residual PFA. Cellular analysis was performed on a BD LSR Fortessa flow cytometer, and analyzed using FlowJo (FlowJo, LLC). Primary monocytes were first gated by size (FSC/SSC), singlets (FSC-A/FSC-H), and either

expression of CD14 or CD16 (Fig. 1B). Phagoscores were then calculated by taking the geometric mean fluorescence intensity (gMFI) of the bead positive monocytes, multiplying the percent of bead positive monocytes, and dividing by an arbitrary factor for interpretability. Changes in expression were reported as log2 MFIs over PBS controls and were assessed by comparing all gated monocytes. Cytokine secretion by monocytes upon immune complexes stimulation was analyzed in the same manner as described above with a customized multiplex luminex bead panel (Thermo Fisher) following the manufacturer's protocol.

2.5.8 Statistical analyses

All analyses were performed using python version 3.6.8. Classification models were trained to distinguish hospitalized and non-hospitalized SARS-CoV-2 infected individuals. Prior to classification all samples were subtracted for background levels based on the PBS control conditions across all features, and all data was scaled and centered. Correlations between technical and donor replicates were determined using Pearson correlations and between different conditions using Spearman correlations. Statistical differences among different conditions were determined by a Mann–Whitney U Test and were multiple hypothesis corrected with the Bonferroni method. ADCP and ADMP models were separately constructed based on either all features, only surface marker features, only cytokine features, or on features selected by a cross validated least absolute shrinkage and selection operator (LASSO) using a partial least squares discriminant analysis (PLSDA) classifier. Final models were then visualized using scores plots and feature variable of importance in projection (VIP) was determined. Classification accuracy was then determined for all models with multiple iterations of fivefold cross-validation, while maintaining the same training and test set identities for all models during each iteration. Models were generated 100 times, and outcomes were then compared to determine accuracy. Statistical significance between ADMP and ADCP models was defined by exact P values of the tail probabilities of the ADMP distributions within the ADCP distributions. Reported are the median p-values over the course of every repetition[68,107].

2.6 Supplementary figures

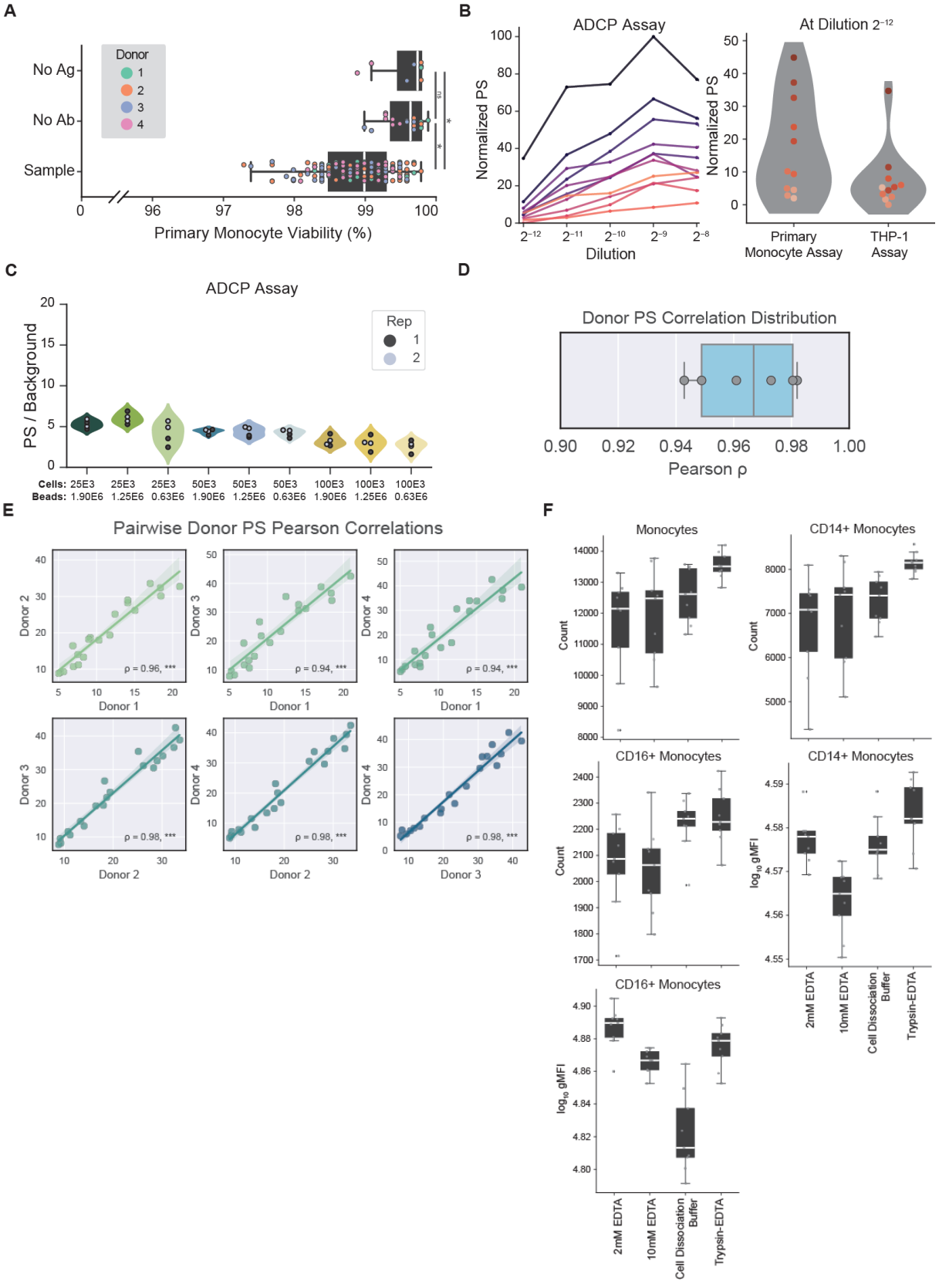


Figure 2-S1. Primary monocyte viability, ADCP benchmarking, and donor correlations

(A) Boxplots depict the viability of primary monocytes following exposure to immune complexes (Sample), to no antibody controls (No Ab), or are unperturbed (No Ag). Each boxplot is composed of 4 donors and profiled across 20 samples and 2 technical replicates, and statistical significance was assessed using the Mann–Whitney U Test (p<0.001: *) (B) (Left) Dilution curves of 9 different samples against HA1 from 1:256 (2-8) to 1:4096 (2-12). All curves were background subtracted and then min-max normalized. Each color represents a different individual and is an average of 2 technical replicates. (Right) Normalized PS of ADCP and ADMP assays across the ADMP and ADCP assays at dilution 1:4096 (2-12). The color of the dots corresponds to a different sample. (C) PS over background of different cell and bead conditions consisting of two different sera pools across 2 technical replicates. The color of the dots corresponds to the replicate, and the number of cells and beads correspond to the amount suspended in each 200µL well of a 96 well plate. (D) Distribution of phagoscore (PS) by Pearson correlations (ρ) among different donors. Each correlation consists of 20 different samples and is an average of 2 technical replicates and profiled against the Influenza HA1 antigen. (E) Visualization of the univariate pairwise correlations among the 4 donors are shown below and their relationship is represented by a linear regression model. The shadowed area represents the 96% confidence interval and the corresponding ρ and significance is reported (p<1e-6: ***). (F) Boxplots depict the count and log10 geometric mean fluorescence intensity (gMFI) of ADMP following treatment with different dissociation agents and is composed of 9 samples.

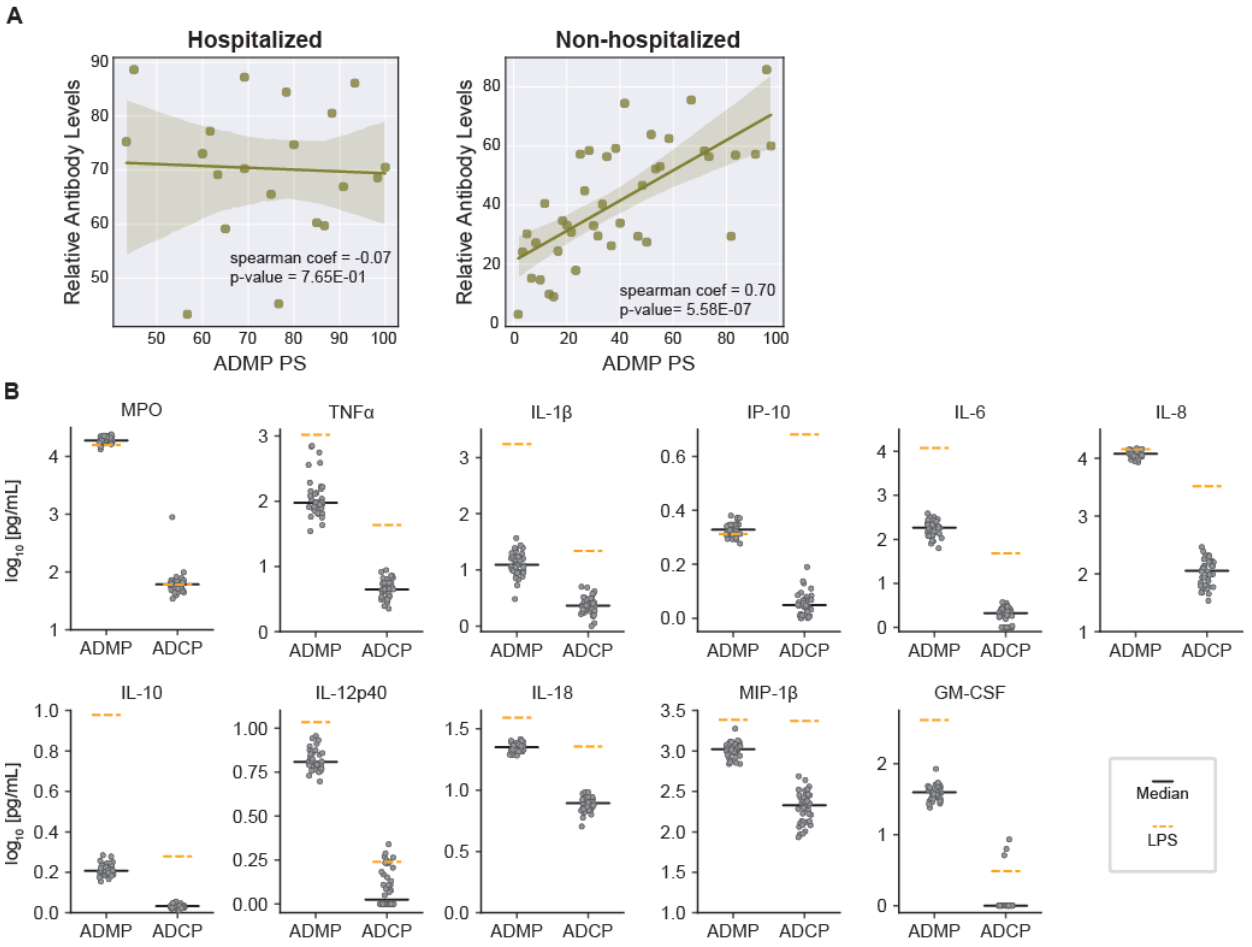


Figure 2-S2. Antibody mediated cytokine secretion against the spike protein of SARS-CoV-2 across ADMP and ADCP assays

(A) Spearman correlations between relative antibody levels and ADMP phagoscores among 20 hospitalized and 40 non-hospitalized SARS-CoV-2 infected individuals. Relative antibody levels were determined by taking the mean percentiles of the isotype and subclass levels across each individual captured by Luminex. Each value is an average of 2 technical replicates, the relationships are represented by a linear regression model, the shadowed area represent the 96% confidence interval, and the corresponding spearman correlation coefficient (ρ) and significance are reported (p<1e-6: ***). (B) The supernatants of 12 hospitalized and 28 non-hospitalized SARS-CoV-2 infected individuals were selected for post-phagocytic cytokine analysis across ADMP and ADCP assays. Distributions depicting the cytokine secretion of hospitalized and non-hospitalized individuals across the ADCP and ADMP assays. Cytokine concentrations were converted by a standard curve, and then log10 transformed. Each response is an average of 2 technical replicates. The solid black line represents the median, and the dashed orange line a 100 ng/mL LPS stimulation control.

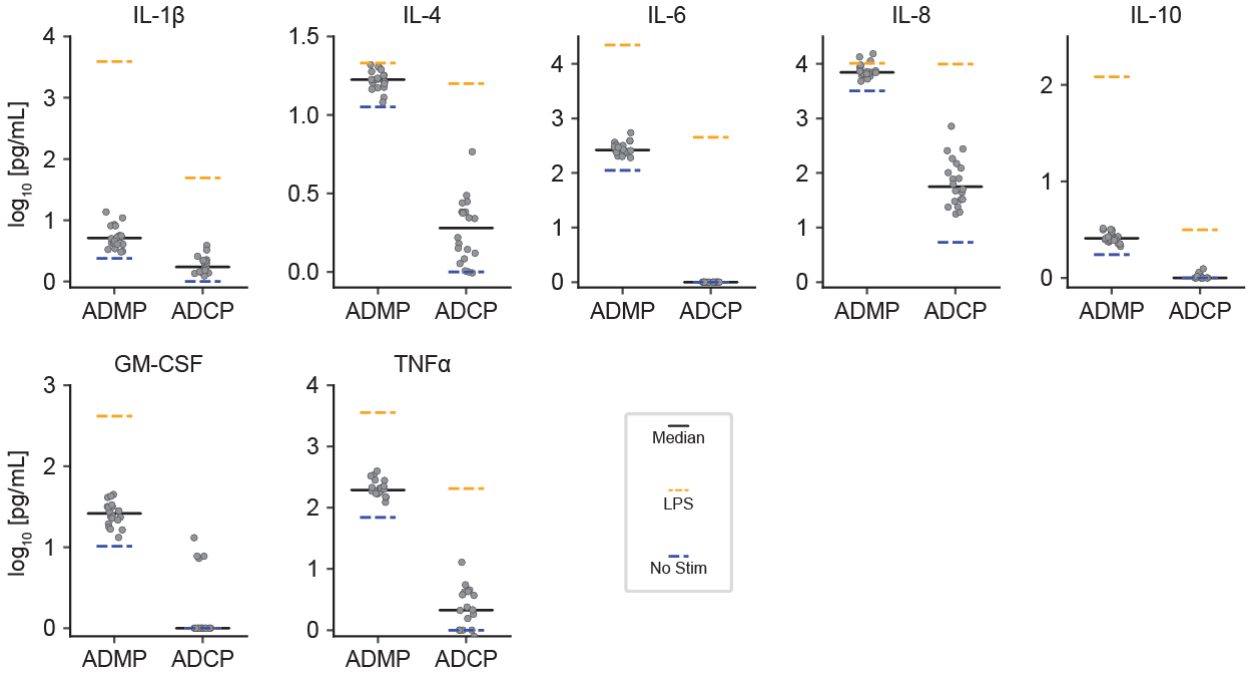


Figure 2-S3. CR3022 Fc-variants induced cytokine secretion across ADMP and ADCP assays

Cytokine secretion following ADMP with monoclonal Fc-variants targeting the receptor binding domain (RBD) of the SARS-CoV-2 are depicted with each dot representing 1 of 20 different CR3022 variants across each assay. Cytokine concentrations were converted by a standard curve, and then log10 transformed. Each response is an average of 2 technical replicates. The solid black line represents the median, the dashed orange line a 100 ng/mL LPS stimulation control, and the dashed blue line the no stimulation control.

2.7 Acknowledgements

We acknowledge support from the Ragon Institute of MGH, MIT, the Massachusetts Consortium on Pathogen Readiness (MassCPR to GA), the Bill and Melinda Gates Foundation (235730 to GA and INV-016575 to HYC), NIAID (U19 AI35995 to GA, and Contract 75N93019C00063 to DMK), and the U.S. Centers for Disease Control and Prevention (CK000490 to GA).

Chapter 3

Upper and lower respiratory tract correlates of protection against respiratory syncytial virus following vaccination of nonhuman primates

The contents of this chapter were published as:

Tomer Zohar, Jeff C. Hsiao, Nickita Mehta, Jishnu Das, Anush Devadhasan, Wiktor Karpinski, Cheryl Callahan, Michael P. Citron, Daniel J. DiStefano, Sinoeun Touch, Zhiyun Wen, Jeffrey R. Sachs, Pedro J. Cejas, Amy S. Espeseth, Douglas A. Lauffenburger, Andrew J. Bett, and Galit Alter. "Upper and lower respiratory tract correlates of protection against respiratory syncytial virus following vaccination of nonhuman primates". *Cell Host Microbe* 30, 41-52. e5 (2022).

See the online publication for any references to supplementary material not included in this thesis.

Contributions

Conceptualization, T.Z., A.J.B., and G.A.; formal analysis, T.Z. and J.C.H.; investigation, N.M., J.D., A.D., W.K., C.C., M.P.C., D.J.D., S.T., and Z.W.; writing–original draft, T.Z.; writing–review & editing, T.Z., J.C.H., J.R.S., A.S.E., A.J.B., and G.A.; visualization, T.Z. and J.C.H.; supervision, D.A.L., A.J.B., and G.A.; funding acquisition, A.J.B. and G.A.

3.1 Abstract

Respiratory syncytial virus (RSV) infection is a major cause of respiratory illness in infants and the elderly. Although several vaccines have been developed, none have succeeded in part due to our incomplete understanding of the correlates of immune protection. While both T cells and antibodies play a role, emerging data suggest that antibody-mediated mechanisms alone may be sufficient to provide protection. Therefore, to map the humoral correlates of immunity against RSV, antibody responses across six different vaccines were profiled in a highly controlled nonhuman primate-challenge model. Viral loads were monitored in both the upper and lower respiratory tracts, and machine learning was used to determine the vaccine platform-agnostic antibody features associated with protection. Upper respiratory control was associated with virus-specific IgA levels, neutralization, and complement activity, whereas lower respiratory control was associated with Fc-mediated effector mechanisms. These findings provide critical compartment-specific insights toward the rational development of future vaccines.

3.2 Introduction

It is estimated that respiratory syncytial virus (RSV) causes approximately 33.8 million new infections annually, with the highest mortalities in developing countries[108]. It is estimated that most children become infected during the first few years of life, and although most cases of RSV result in mild disease, a certain proportion of cases go on to develop bronchiolitis, pneumonia, and broad inflammation of the respiratory tract[109]. RSV is considered to be the most common cause of childhood acute lower respiratory infections and account for one-third of deaths during the first year of life[108,110]. Moreover, emerging data suggest that older adults represent an additionally vulnerable population accounting for nearly 200,000 infections and greater than 10,000 deaths annually in the US alone[111]. Currently, RSV is mainly treated with respiratory supportive care such as supplemental oxygen and ventilators, and though neutralizing monoclonal antibody (mAb) prophylactics exist, use is largely restricted to the treatment of premature infants which are at highest risk. However, despite the clear need to combat RSV infections globally, vaccine development has failed in part due to the limited understanding of the correlates of immunity required to prevent infection and disease, and the challenge of inducing immunogenic responses in newborns with developing immune systems that are in the presence of maternal antibodies.

Substantial evidence highlights the critical role of the humoral immune response in protection against RSV. Prophylaxis with neutralizing monoclonal antibody therapeutics has been associated with a reduction in hospitalizations, with higher neutralizing antibody titers associated with the greatest reduction in disease[112,113]. Additionally, recent meta-analyses integrating data related to titers, neutralization, T cell immunity from various prophylactic mechanisms found neutralization as predictor of protection against RSV, although prediction does not imply mechanistic causality. However, even though several RSV vaccines have been developed over the years that are able to induce neutralizing antibodies, no vaccine is currently available, perhaps related to the strict focus on antibody titers and neutralizing antibodies. Instead, emerging data point to the importance of both neutralizing and non-neutralizing a-ntibody functions, in protection across several infectious diseases[114–118], however, the precise mechanisms of protection against RSV remain incompletely understood.

Recent studies have emphasized the importance of Fc-mediated antibody activity against RSV infection[119]. Reduced NK cell numbers, even in the presence of high-maternally transferred antibodies have been noted in infants that acquire RSV pointing to a potential importance of NK cell mediated ADCC in protection against RSV[120]. Moreover, multiple studies have highlighted the ability of antibodies to the RSV G-antigen, a dominant viral-surface antigen involved in attachment, in inducing efficient ADCC *in vitro*[121] and *in vivo*[122]. Monoclonal antibodies 1812A2B and 1C2, both directed at RSV-G, have been shown to significantly reduce viral load and provide protection *in vivo* in the setting of intact Fc-effector function, suggesting that Fc-dependent innate effector functions in protection *in vivo* is critical[123–125]. Additionally, N-glycan structure variants of palivizumab with improved Fcγ receptor binding have been associated with reduced viral lung titers[126]. Furthermore, enhanced disease has been noted in complement deficient mice[123], pointing to the potential importance of additional Fc-mechanisms in protection against RSV[127]. However, the precise anti-viral humoral mechanisms of protection against RSV remain unknown, hampering the ability to advance a protective vaccine.

Here, we sought to determine potential antibody correlates of protection against RSV in a unique setting of controlled RSV-infection and immunity. Specifically, using non-human primates immunized with disparate vaccine strategies, each showing a unique antibody-vaccine induced profile and striking differences in protection from infection in both the upper and lower respiratory tracts, we were able to perform a unique correlates analysis. Using systems profiling of the antibody response and machine learning, unique antibody features were associated with viral loads in the upper and lower respiratory tracts over time following challenge, with a dominant role of IgA and neutralization in the upper respiratory tract, and a critical role for Fc-effector functions in the lungs. These data highlight important differences in potential immune correlates of protection against RSV infection at different sites within the respiratory

system, agnostic of the vaccine platform, providing a glimpse into highly desirable target immune profiles that provide important insights into vaccine design.

3.3 Results

3.3.1 Striking heterogeneity in RSV restriction across vaccine strategies

Monoclonal antibody studies have clearly illustrated the critical role of antibodies in protection against RSV[112,122,125,127]. Moreover, antibody titers and neutralization have been implicated in protection, but alone represent incomplete correlates of immunity against RSV in both infants[128] and in the elderly population[129]. Conversely, emerging data demonstrate a role for antibody-effector functions, beyond neutralization in the control of RSV[119,123,127], thus, here we aimed at systematically probing the functional role of antibodies against RSV.

We elected to focus on a controlled challenge model, specifically focusing on a group of African Green Monkeys (AGM) that received 1 of 6 different vaccine constructs. In addition, to the six different vaccine platforms, one group was given an empty lipid nanoparticle (LNP) vaccine, delivered intramuscularly (negative control) and a second group was administered an attenuated strain of RSV A2, intranasally (RSV-A2). Three groups were given differently formulated F protein-based DS-Cav1 vaccines[130]. Two of three were administered intranasally, using 20% nanoemulsion at a dose of 125µg in a volume of either 0.1mL (DS-Cav1-LV) or 0.25mL (DS-Cav1-HV), and one was administered intramuscularly with 500 µg of Adju-Phos®, an aluminum phosphate adjuvant, at a dose of 125µg and volume of 0.5mL (DS-Cav1-AD). Two additional groups were given RSV F nanoparticle vaccines intramuscularly at a dose of 135µg in a volume of 0.5 mL either with (RSV-F-AD) or without the Adju-Phos® adjuvant (RSV-F)[131]. The final group was given an mRNA/LNP vaccine encoding the membrane bound version of DS-Cav-1 which was delivered intramuscularly at a dose of 125µg and a volume of 0.5mL (mDS-Cav1)[132] (Figure 3-1A). This mRNA vaccine, also referred to as mRNA-1777 (V171), was recently evaluated in a phase 1, randomized, placebo-controlled study to evaluate the safety and immunogenicity in healthy younger and older adults[133]. All the vaccines delivered the RSV F protein, implicated as a critical target for both neutralizing, and protective antibodies[130]. Three doses of vaccine were administered to each animal at Day 0, 29 and 56 and then the animals were challenged with RSV A2 both intranasally and intratracheally on day 70. Serum samples were collected prior to each vaccination and at the time of challenge.

Following challenge, nasal swabs were collected daily for 15 days and bronchoalveolar lavage fluid (BAL) were collected on days 3, 5, 7, 9 and 14 to assess for viral loads (Figure 3-1B and Figure 3-1C). While differences were noted in BAL viral loads across the groups (Figure 3-1C), complete control of nasal viral loads was only observed with intranasal attenuated RSV-A2 immunization (Figure 3-1B). All other vaccine modalities shifted the kinetics of the nasal viral loads. Furthermore, attenuated RSV-A2 and mDS-Cav1 alone showed complete restriction of viral loads in the BAL, with delayed

and attenuated peak viral loads across the groups. With the exception of vaccine groups DS-Cav1-HV and DS-Cav1-LV variation in viral loads among non-human primates within groups were low for both the nasal cavity and the BAL. RSV-A2 exhibited the best overall control followed by mDS-Cav1, and then DS-Cav1-AD highlighting a clear hierarchy in vaccine induced protection against RSV and pointing to differences in capacity of vaccine induced immune responses to reduce viral load across the upper and lower respiratory tracts.

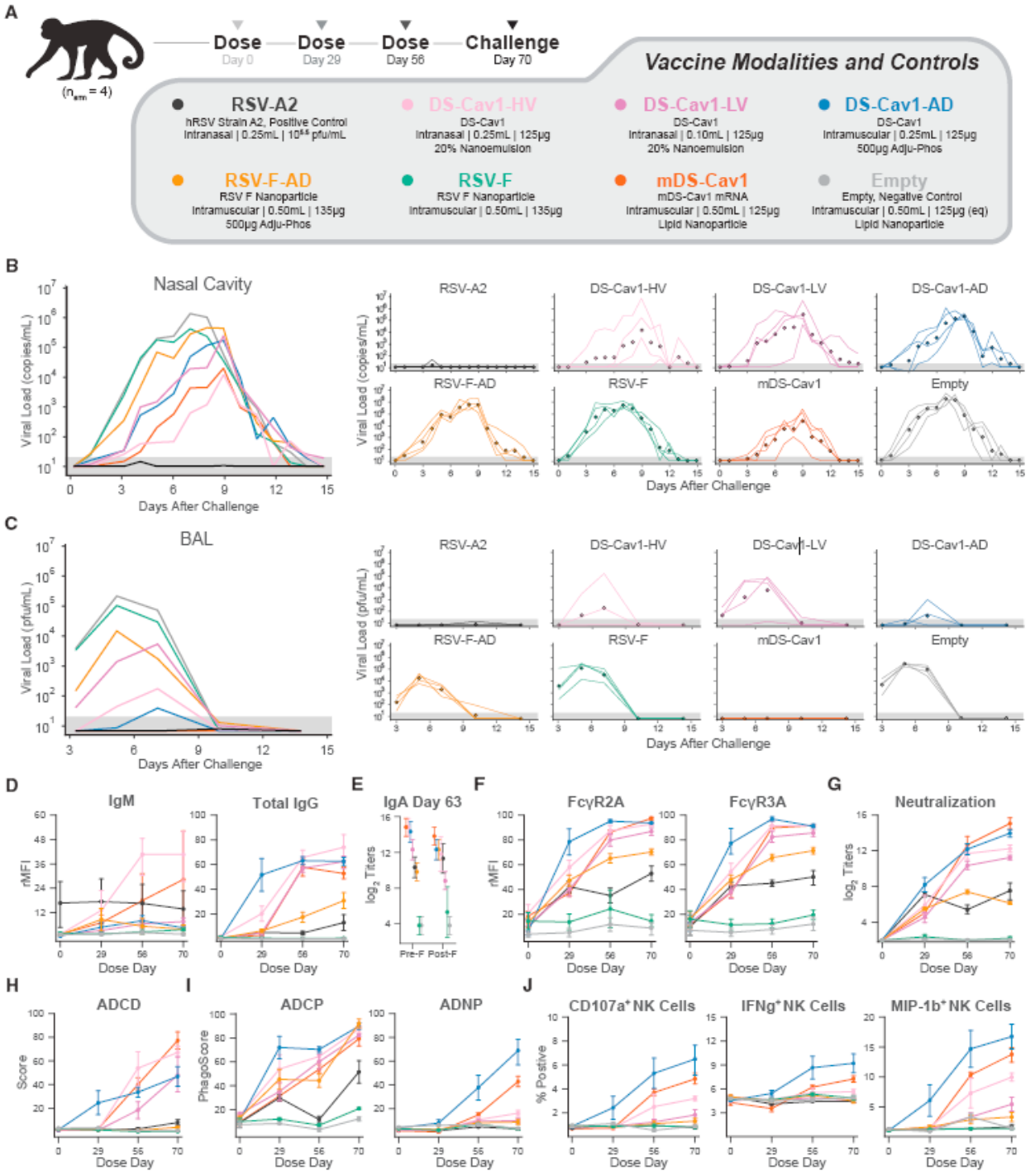


Figure 3-1. Study design, efficacy, and antibody responses with dosage

(A) 32 macaques were split into eight groups, six of which were administrated different vaccine modalities, one administrated an empty lipid nanoparticle control, and one group vaccinated with an attenuated RSV strain. Each group was dosed at day 0, 29, 56, and then challenged at day 70. (B) Viral loads were captured days 0-15 after challenge in the nasal cavity. On the left are the mean trajectories for with every line representing a specific group, and on the right are the individual non-human primate trajectories for each condition. The diamonds at each time point correspond to the mean for that group and corresponding time point. (C) Viral loads were captured days 3-15 after challenge in the BAL and represented in the same manner as panel B for both the graph on the left and the right. (D) Humoral evolution of IgM and total IgG titers, (E) IgA titers against pre and post RSV fusion protein at day 70, and (F) enhanced Fc-receptor binding along three doses and prior to challenge. (G) Development of neutralization titers through three doses and prior to challenge. (H) Functional humoral evolution of antibody mediated responses including complement deposition (ADCD), (I) phagocytosis by monocytes (ADCP) and by neutrophils (ADNP), and (J) natural killer (NK) cell degranulation by CD107a, IFNγ, and MIP-1β. All panel expect D, F, and I were min max normalized providing relative mean fluorescent intensities (rMFI) or corrected scores. Panels D and F are represented as $\log_2$ titers, and panel I as the percent of NK cells positive for each degranulation marker. Error bars show one standard deviation from the mean. Colors correspond to the vaccine strategies and control.

3.3.2 Robust Pre-F vaccine responses observed across different modalities

To define whether differences in vaccine induced antibody profiles were associated with differential challenge outcome, we next compared both the biophysical and functional profiles across the vaccine groups over the course of the study. Limited IgM levels were observed across most vaccine groups except for DS-Cav1-HV and mDS-Cav1 which showed moderate levels. Conversely, RSV IgG and IgA titers were induced across all vaccinated animals with the strongest responses in DS-Cav1-AD, DS-Cav1-HV, DS-Cav1-LV, and mDS-Cav1 (Figure 3-1D and 1E). Intermediate titers were also observed for RSV-F, RSV-F-AD, and RSV-A2. Interestingly, Total IgG and Pre-F IgA levels were similar across the groups, but Post-F levels were higher in the mDS-Cav1 and DS-Cav1-AD groups. Concordant with IgG titer changes, Fcγ-receptor binding antibodies, to both FcγR2A and FcγR3A, followed the same levels as observed with IgG titers, albeit disproportionately higher Fcγ-receptor binding was observed in the RSV-F-AD group early on (Figure 3-1F). Similar to the development of IgG and IgA titers, robust neutralization titers were also observed in DS-Cav1-AD, DS-Cav1-HV, DS-Cav1-LV, and mDS-Cav1 vaccine groups followed by intermediate responses by RSV-A2 and RSV-F-AD. Interestingly, unlike IgG and IgA the mDS-Cav1 and DS-Cav1-AD groups showed even higher neutralization titers than other groups (Figure 3-1G).

In contrast to titers and Fc-receptor binding profiles, antibody-effector function showed greater heterogeneity across vaccine arms. For example, while the DS-Cav1-AD arm exhibited the earliest evolution of antibody dependent complement deposition (ADCD), DS-Cav1-HV and mDS-Cav1 immunized animals exhibited the highest magnitude responses over the study period (Figure 3-1H). Nearly all arms developed antibody dependent cellular phagocytosis (ADCP), with similar levels of ADCP were observed in the mDS-Cav1, DS-Cav1-LV, DS-Cav1-HV, DS-Cav1-AD, and RSV-F-AD

vaccinated animals that exhibited lower antibody titers (Figure 3-1I). These data suggest that the functional quality of the humoral immune response may not be strictly determined by the quantity of the antibody. Interestingly, antibody dependent neutrophil phagocytosis (ADNP) was only observed in the DS-Cav1-AD- and mDS-Cav1-vaccinated animals. Similar trends were observed with antibody dependent NK cell degranulation (CD107a upregulation), cytokine secretion (IFNγ), and MIP-1β secretion (Figure 3-1), with highest levels in the DS-Cav1-AD group, followed by mDS-Cav1 and then DS-Cav1-HV. Ultimately, these data point to qualitatively distinct humoral immune responses across vaccine groups, but to confirm specific quantitative differences across platforms studies with improved sample size are needed.

To explore the unique profiles induced across the vaccine humoral profiles at a multivariate level, we performed a principal component analysis (PCA) (Figure 3-2A). Despite the small numbers of animals in each vaccine arm, each arm separated, highlighting the different antibody-profiles induced by each vaccine (Figure 3-2A, left). The first principal component, PC1, marked substantial quantitative and qualitative differences across weakly immunogenic vaccines and vaccines that elicited robust humoral immune responses, with loadings scaled along PC1 (Figure 3-2A, right). However, loadings along the second principal component, PC2, provided contrast between the quality of the vaccine induced profiles. For example, PC2 segregated vaccines able to induce robust levels of NK cell activation, ADNP, neutralization and IgA (at the top of PC2) from those that induced more robust ADCP, IgM, Fc-receptor binding (at the bottom of PC2), providing comprehensive functional resolution of the overall functional quality of the vaccine induced humoral immune responses induced by distinct vaccine platforms.

Furthermore, to capture temporal differences between the vaccine induced immune responses, a PCA analysis was performed longitudinally, using a phase portrait (Figure 3-2B). The phase portrait highlighted the unique mean path taken by each vaccine following immunization. A unique directionality was observed for the DS-Cav1-AD and mDS-Cav1 vaccine groups. A shared directionality was observed for the DS-Cav1-LV and DS-Cav1-HV vaccines, similar trajectories emerged across the RSV-F-AD and RSV-A2 vaccinated profiles, and common paths were followed by the less-robust antibody inducing RSV-F, that overlapped with the empty control group (Figure 3-2B). Collectively, these data point to the distinct development of antiviral-responses in the setting of differences between qualitative antibody profiles of the vaccine groups.

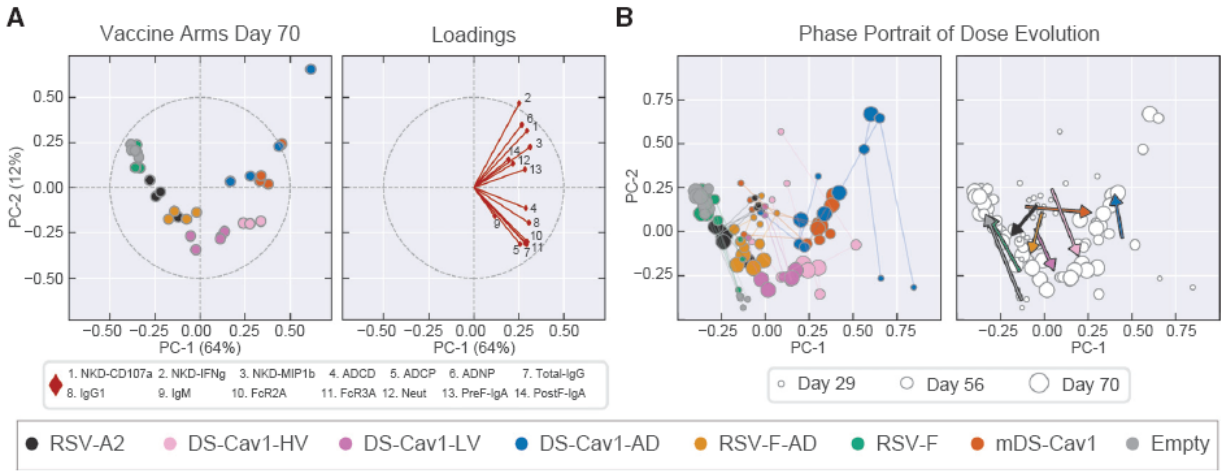


Figure 3-2. Unique composite differences in response dynamics among vaccine strategies

(A) PCA with responses before challenge and the corresponding loadings plot. Colors in all panels correspond to each group. (B) Phase portrait decomposing humoral responses over day 29, 56, and 70 across all NHPs. Each line represents a trajectory of a single NHP with the larger size dots corresponding to later time points. On the right the mean path of each group is depicted over time. All data was scaled and centered within each time point prior to decomposition. Colors correspond to the vaccine strategies and control.

3.3.3 Antibody functions are more strongly linked to protection than neutralization

Given the robust, albeit heterogeneous, induction of functional antibody responses following vaccination with the distinct platforms, we next aimed to determine whether particular vaccine induced antibody properties tracked with protection against RSV. Thus, we explored the individual relationship between antibody functions, titers, Fc-receptor binding capacity, neutralization, and IgA and viral loads at a univariate level (Figure 3-3). After correction for multiple comparisons, fewer significant correlations were noted in the nasal profiles, although IgG titers, FcγR3a binding antibody levels, and Pre-F antibody titers were all significantly correlated with protection at day three, and IgM alone remained significantly associated with protection at day five in the nasal samples. Conversely, many features of the vaccine induced humoral immune response were linked to enhanced viral control in the BAL. Of note, ADCP and IgM correlated poorly with protection in the BAL, post-F IgA titers were only correlated with viral load control at day five, and neutralization was only associated with viral load control at day three. Moreover, the correlation coefficients were stronger for NK cell-activating antibodies, ADNP, Fc-receptor binding and Pre-F IgA levels compared to neutralization, providing novel resolution into the platform-independent antibody properties that may be key to protection against RSV.

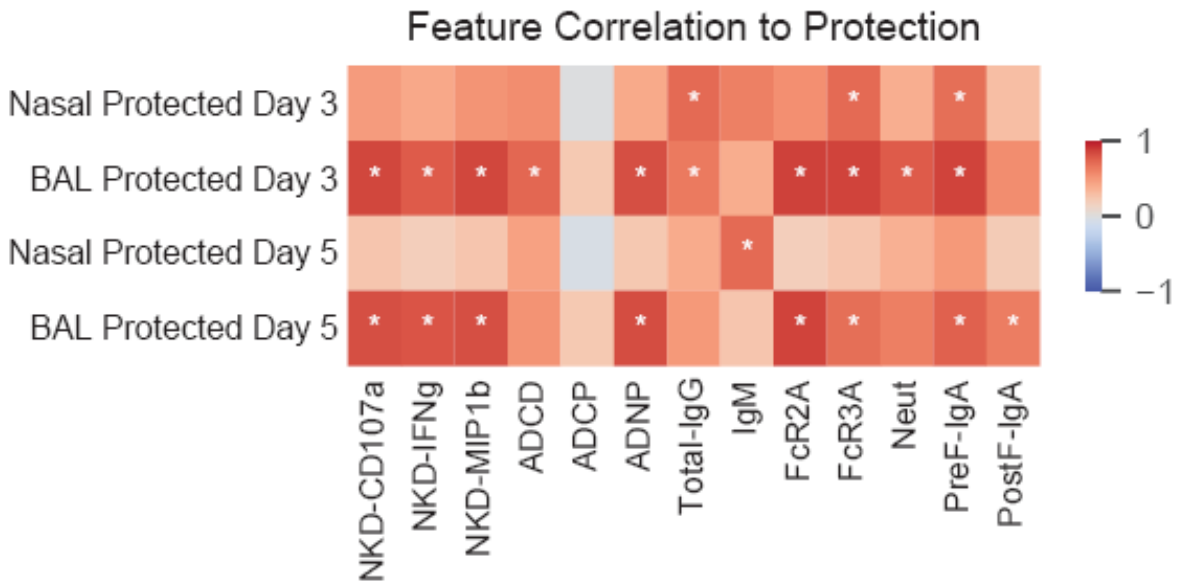


Figure 3-3. Antibody feature correlations to protection

The correlation heatmap depicts Spearman correlations of individual antibody features with viral loads across compartments (nasal and swab) and time points (days 3 and 5). Statistical significance is indicated by white asterisks with Holm-Bonferroni correction for multiple hypothesis testing ($p < 0.001$). Negative correlations are indicated in blue, and positive correlations are denoted in red.

3.3.4 Mechanism of humoral protection in the nasal cavity and BAL

Given the naturally inter-related induction of antibody titers, function, and neutralization, we next aimed to determine the cross-vaccine platform minimal humoral correlates of immunity against RSV. Data from all vaccine groups, excluding control groups, were integrated and partial least square discriminant analyses (PLS-DA) were used to determine a minimal multivariate set of humoral correlates of immunity that tracked with early restriction of viral replication. Control groups were excluded because they did not elicit robust humoral responses. Antibody profiles at day 70 were used for correlates analyses, and models were constructed comparing protected and non-protected NHPs, defined by a viral loads below 20 pfu/mL at day three and five for both the nasal (Figure 3-4A) and lung (Figure 3-4C) compartments following viral challenge.

Both day three and five nasal cavity models were able to accurately predict protection with cross-validation accuracies of 84% and 95% respectively (Figure 3-4A and Figure 3-S1). As few as three antibody features of the total 13 captured, were able to accurately separate animals with viral loads past the lower limit in the nasal cavity. The top features selected in the day three model were all enriched in the protected animals, including complement deposition (ADCD), Pre-F IgA responses, and neutralizing antibody titers. Interestingly, each feature had high univariate predictive power as quantified as the area under (AUC) of the receiving operation curve (ROC) (Figure 3-4B), with early IgA levels ranking as the top predictor of protection followed closely by complement activity that was highly enhanced in protected animals, and

neutralizing titers following closely behind. Day five features changed slightly, with conserved preF-IgA titers still a key correlate of protection, however IgM titers were also included as correlates of protection. In these animals however, ADNP levels were elevated in infected, rather than protected animals (Figure 3-4A). Day five IgM levels possessed complete univariate predictive power, followed closely by IgA and ADNP (Figure 3-4B). However, given the importance of IgM in driving complement activation, these data suggest that IgM and complement as well as IgA are potential correlates of immunity in the upper nasal tract. Thus, strict neutralization with complement activating activity may be sufficient to delay viral infection in the upper respiratory tract.

Since more striking viral control was observed in the lungs (Figure 1B), we next probed whether RSV-specific antibody effector profiles were distinct among protected and unprotected animals. A near complete split was observed in antibody profiles at day three across the animals, linked to just three antibody features that were all enriched in the protected animals, with a cross-validation accuracy of 91% (Figure 3-4C and Figure 3-S1). The three features included antibody dependent NK cell-chemokine secretion (MIP-1β), complement deposition, and vaccine specific IgA levels, each with consistently high individual predictive accuracies (Figure 3-3D). Day five BAL profiles remained distinct across protected and infected animals with a cross-validation accuracy of 85% (Figure 3-4C and Figure 3-S1). The day five model contained five features, including NK cell chemokine (MIP-1β) and cytokine (IFNγ) secretion, ADNP, and IgA and IgM titer. Apart from IgM titers most of the selected features showed high univariate predictive accuracies (Figure 3-4D). This more complex signature points to the critical interaction between multiple functions and isotypes in the control of RSV infection. To confirm whether similar complex signatures were consistent seven days out in BAL a similar analysis was performed (Figure 3-4E). This analysis of protection revealed similar correlates of protection, namely IgA titers and NK degranulation were that were consistent seven days out. Interestingly, ADCD appeared as a new correlate of protection specific to day 7 and may be necessary for suppressing infection seven days out. Although neutralization was not selected as a correlate of immunity by the models, neutralization was clearly associated with reduced viral loads by univariate analysis on day three in the BAL (Figure 3-3), suggesting that neutralization may play a supportive mechanistic role in protection in the lung, particularly early on in infection. Thus, this wholistic agnostic analysis suggests that NK cells, complement, and neutrophil phagocytosis may play a dominant role in containment of RSV in the lower respiratory tract.

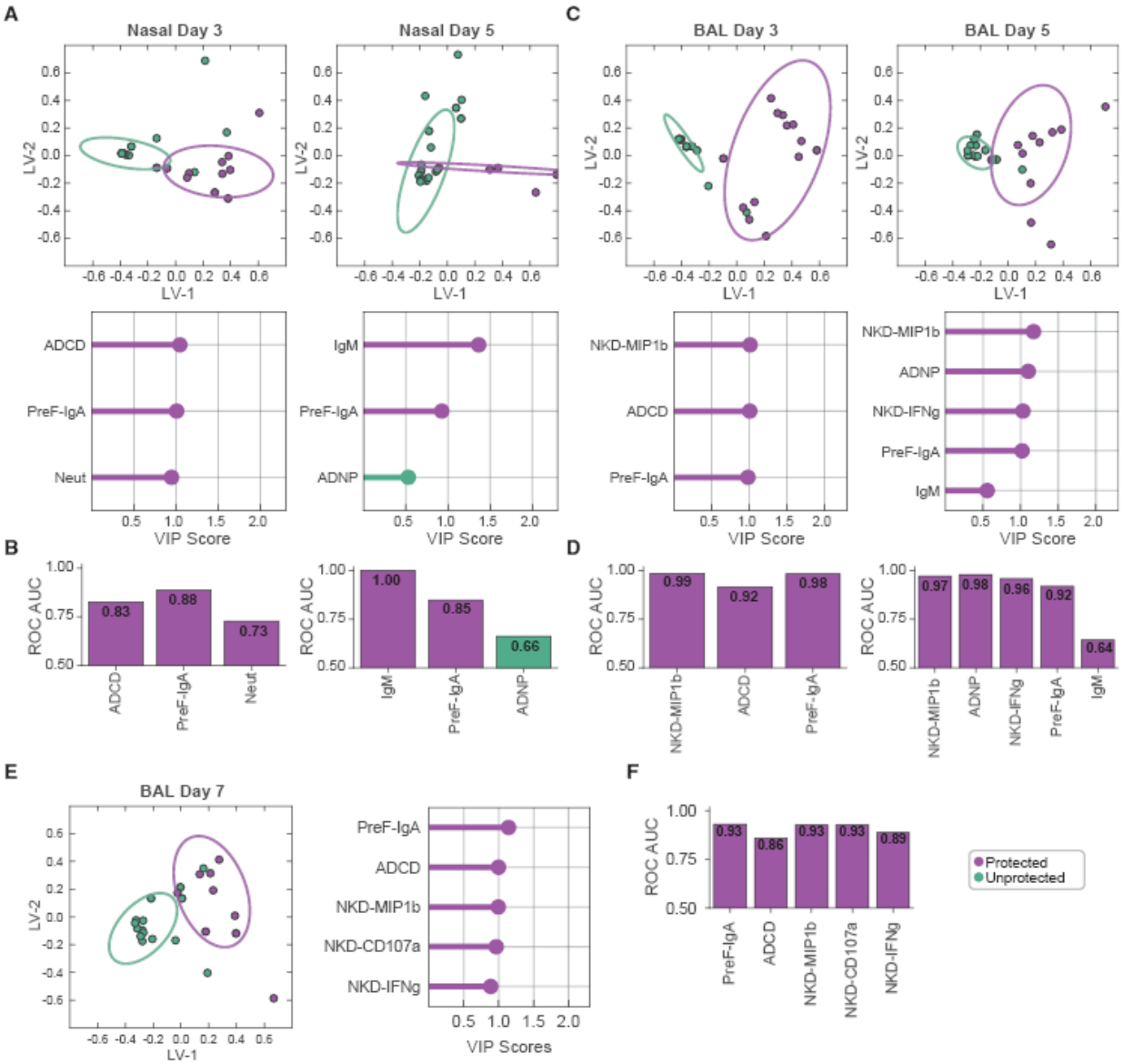


Figure 3-4. CR3022 Fc-variants illicit unique post-phagocytosis responses

A) On top, the scores plots of the PLS-DA classification models, predicting protection in the nasal cavity for days 3 and 5 after challenge, are shown with green corresponding to the unprotected and purple to the protected groups. Below the scores plots, the variable importance in protection (VIP) scores are shown. (B) Area under the receiver operating characteristic curve (ROC) are displayed for the model selected features of days 3 and 5 after challenge in the nasal cavity. (C) Scores plot of the PLS-DA classification models predicting protection in the BAL for days 3 and 5 after challenge are shown on top with the VIP scores below. (D) Area under the ROC of the model selected features of days 3 and 5 after intratracheal challenge in the BAL are displayed. (E) Scores plot of the PLS-DA classification models, predicting protection in BAL for day 7 after challenge, is shown to the left with the VIP scores to the right. (F) Area under the ROC for the model selected feature 7 days after intratracheal challenge is shown.

3.3.5 Collective correlates of immunity are selectively induced by vaccine regimens

To gain further insights into the relationship of the humoral correlates of immunity and additional antibody functions, a correlation network with all the selected features was constructed (Figure 3-5A). The network revealed relationships between all protective features, apart from IgM, highlighting the specific humoral immune functions that co-evolve in animals that exhibited the most robust protection. Specifically, selected features were tightly co-induced with one another, with a critical co-evolution between neutralization, IgA, NK cell activation, and the activation of neutrophils. This tight linkage between Fc-functions and neutralization points to a collaborative role between the Fc and Fab of the antibody in limiting RSV infection.

To further visualize common and distinct features selected by the time-based compartment models the features selected by each model were summarized in a table (Figure 3-5B). IgA titers were important across all compartments and days post challenge. Interestingly, in both compartments, ADCD was selected at day three. Conversely, IgM titers emerged as correlates in the day five models pointing to a potentially expanded role for IgM, able to also drive phagocytic clearance of particles in the lungs. Finally, neutralizing antibody levels to RSV was a robust correlate of immunity early in the nasal cavity but was less apparent over time and in the lower respiratory tract. These data suggest that strict neutralization may only be key to protection early in upper respiratory protection, complement mediated activity may be key on day three across compartments, whereas NK cells and neutrophils appear to act later in the antiviral response, providing protection strictly in the lower respiratory tract.

Finally, we sought to explore vaccine platform specific responses across each of the correlates that were selected in more one of the models (Figure 3-5C). All correlates were robustly induced by the mDS-Cav1 vaccine arm. A more NK and ADNP biased response was observed in the DS-Cav1-AD group. A robust IgA and ADCD shifted response were noted in the DS-Cav1-HV group. Thus, while the distinct regimens showed differences, all vaccine arms exhibited some capacity to elicit immunity (Figure 3-1). However, limited correlates were observed across the other vaccine groups, providing potential mechanistic insights into the lack of protection observed against RSV.

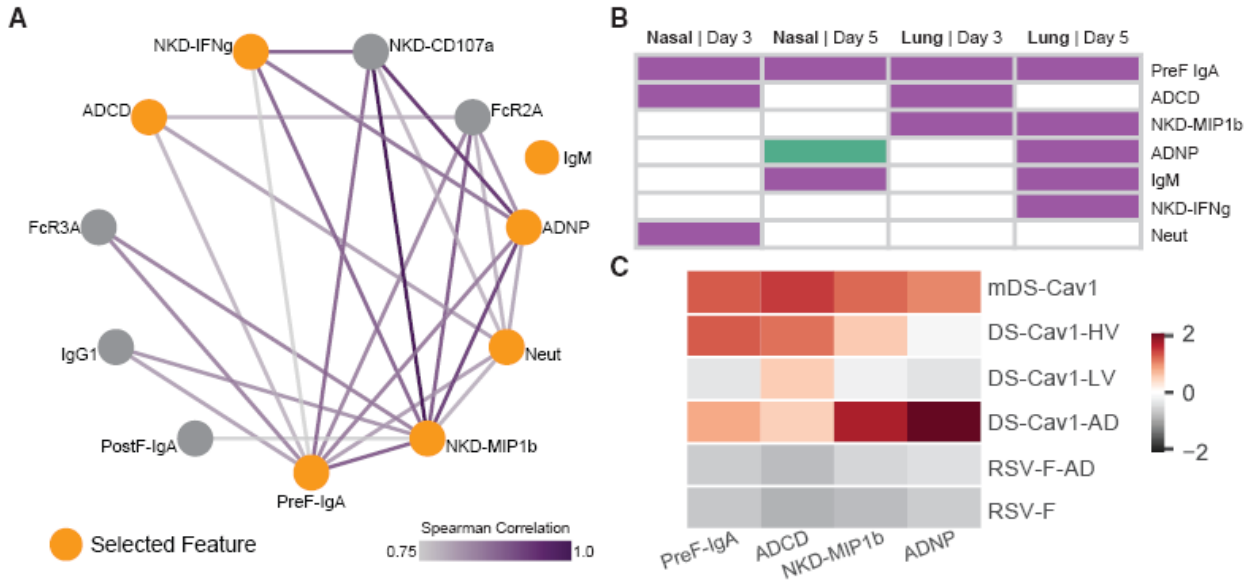


Figure 3-5. Vaccine strategies and humoral feature composition

(A) Spearman correlation network illustrating features that are significantly correlated with biomarkers that track with protection in time point and compartment models. Purple corresponds to high positive correlation while gray corresponds to less positively correlated edges. The orange nodes represent features selected in at least one of the models. (B) Table summarizing the features chosen by each compartment and time post-challenge model. The colors correspond with protection, with green being an average higher value in the unprotected and the purple in the protected group. (C) The heatmap depicts the mean Z scored responses of features that were chosen by more than one model across the different vaccine modalities. First vaccine groups, excluding controls, were Z scored for each feature and the responses were averaged for a given vaccine group. Red corresponds with higher values and gray with lower values.

3.4 Discussion

Despite several attempts to develop a protective vaccine against RSV[134], no vaccine is currently available. Neutralization has been linked to protective immunity in some, but not all studies[135]. Likewise, ADCC activity has been associated with protection against RSV in animal models but has yet to be defined as robust correlates of immunity in humans[122]. While neutralization may provide some protection, antibodies have the capacity to leverage a broader array of antiviral functions beyond neutralization and ADCC that may act in a coordinated manner to confer the most robust protective profile. Thus, to capture wholistic humoral correlates of immunity against RSV and gain mechanistic insight, we profiled the humoral immune response in a highly controlled African Green Monkey (AGM) challenge model, across animals vaccinated with an array of distinct vaccines. AGMs, which are semi-permissive for RSV replication, are one of several animal models that have been used to evaluate RSV vaccine candidates[136]. However, despite many advantages non-human primate models also provide some inherent difficulties due to the small sample sizes. Furthermore, the RSV challenge model used here (RSV introduced into the nose and lung) may represent a stringent

albeit extreme model of infection. Yet despite these differences to natural human infection, striking differences in humoral profiles were observed in the upper and lower respiratory tracts associated with viral restriction. Neutralization and ADCD driven immunity correlated with superior control in the upper respiratory tract. In contrast, cellular based mechanisms were associated with protection in the lower respiratory tract. Ultimately, these data highlight a composite correlate of immunity which may confer a greater level of protection against RSV.

While antibodies represent the primary correlate of immunity following most licensed vaccines, antibody titer alone has been insufficient for providing critical insights needed to guide vaccine development against many pathogens[114–117]. Mounting data highlights the importance of antibody function rather than magnitude as a critical mechanistic correlate of immunity. Consequently, correlates of immunity may be distinct from vaccine to vaccine, providing unique profiles of immunity that can emerge with the use of different adjuvants, platforms, or routes of immunization, but may all leverage a common mechanism of protective immunity [137–139]. Therefore, defining common correlates, rather than vaccine-specific correlates of immunity may reveal key insights required to design an effective vaccine. Here, we assessed vaccines representing different platforms, adjuvants, and routes of administration that provided an opportunity to both examine as well as define common correlates of immunity. Using this agnostic approach, a minimal set of features emerged highlighting the different mechanisms potentially involved in protection in the upper and lower tracts: while neutralization and ADCD corresponded to viral control in the upper respiratory tract, IgG mediated Fc-effector functions were important for lower respiratory viral restriction. These differences in correlates provide interesting clues related to the distinct immunological mechanisms that may be essential to restricting infection in each and both compartments, including early and delayed mechanisms of protection. Thus, vaccine platforms able to deploy these combinations of effector functions may be more likely to provide robust protection against the virus, even if the pathogen escapes early restriction from neutralization.

RSV infection initially begins in the upper respiratory tract, migrating rapidly to the lower respiratory tract, where the virus largely causes disease[140]. Critically, the upper and lower respiratory tracts are composed of distinct cellular components, representing unique immunological landscapes of potential effectors that may be leveraged in response to viral infection[141]. Thus, immune correlates of protection against infection may vary across compartments, representing different antibody effector functions required to restrict protection at the portal of entry compared to the predominant site of viral replication and disease. Therefore, IgA-mediated neutralization with complement activation may relate to the unique local environment of the upper respiratory tract. Given that complement has not only been implicated in driving opsonophagocytic uptake, but also virolysis[142] and enhanced neutralization, these data point to the importance of a collaboration between neutralization and Fc-

effector function as a key early mechanism of pathogen blockade, lysis, and clearance. If restriction does not take place in the upper respiratory tract, neutralization appears to be less critical than cellular functions such as neutrophil phagocytotic activity and NK cell degranulation. This is likely due to mechanisms that rely on opsonized-viral and infected-cell clearance as previously noted against other viral infections[143]. Thus collectively, both initial restriction in the upper respiratory tract and reserved lower respiratory tract control mechanisms may contribute to enhanced restriction against RSV.

PreF-specific IgA emerged as a correlate of immunity across both upper and lower respiratory compartments, and is consistent with previous animal studies[144]. IgA is important for protection against many respiratory illnesses, as individuals with IgA deficiencies are generally more susceptible to RSV re-infection[145]. The intranasally administrated DS-Cav1-HV vaccine induced robust IgA responses. RSV-A2 was also introduced intranasally and led to strong protection, however as with live-attenuated RSV vaccines viral replication levels and the magnitude of humoral responses are linked which could limit durability. The mDS-Cav1 vaccine also showed good viral restriction and robust IgA levels in addition to the correlates of protection. Intranasal immunizations have also previously been proposed to induce higher IgA titers, via the induction of immunity in nasopharyngeal-associated lymphoid tissue where IgA class switching is thought to mainly take place [146,147]. However, whether mRNA vaccination could lead to complete protection if administered intranasally remains unclear.

The mRNA vaccine technology has gained tremendous attention due to its recent application to SARS-CoV-2 vaccine development[148]. The unique profiles observed from the mRNA vaccine platform may be related to its potential self-adjuvating capacity and putative MHC-I and MHC-II presentation[149]. Strikingly, the mRNA vaccine modality provided the most consistent and balanced response and exhibited superior control across both compartments. It is well established that T cells also contribute to the control of RSV, and likely explains how the attenuated A2-virus was able to control infection while inducing lower neutralizing antibody titers[150]. Furthermore, resident T cells have previously been shown to play an important protective role against RSV infection[151,152]. Therefore, it is possible that vaccination could preferentially enhance tissue resident T cell immunity as well. Importantly, while none of the vaccine modalities evaluated in the current study elicited robust RSV F specific CD4+ and CD8+ T cells responses measured in peripheral blood, the mRNA/LNP vaccine did drive the strongest CD4+ T cell responses (data not shown). Thus, improved viral containment with the mRNA platform may also be attributable to enhanced CD8+ and CD4+ T cells responses, which has been observed following mRNA vaccination[149].

Ultimately, these findings highlight the advantage of probing functional antibody differences, beyond neutralization and binding titers, for the selection of protective vaccines. The data suggest that mRNA vaccination, may drive the broad compartment-specific humoral and cellular immune responses potentially required for protection

against RSV. However, both the high volume intranasal nanoemulsion and adjuvated vaccines also illustrated the capacity to induce humoral responses associated with protection. Given that the high volume nanoemulsion vaccine provided substantially stronger responses than the low dose alternative, which has previously has been shown to impact viral restriction in both nasal and lung compartments in both African green monkeys and cotton rats[153], further optimization of antigen dose-levels and/or dose volumes could also provide improved protection. Additionally, future efforts in formulation optimization, able to induce long-lived humoral correlates of immunity across the age-spectrum, could lead to the ultimate development of a highly protective RSV-vaccine. Collectively, these data point to a set of potential humoral correlates of immunity that provide mechanistic insights into the specific immunological functions that may be key to protective immunity against RSV. Therefore, defining compartment specific mechanisms, across vaccine platforms, provides exciting insights into common mechanisms that may guide vaccine development for future vaccines able to fully leverage the humoral immune response against RSV and beyond.

3.5 Limitations of study

There are two main limitations associated with this study. The first main limitation of this study is the need for testing in non-human primates. Although non-human primates are advantageous animal models, they cannot guarantee how and to what extent findings will translate to humans. Furthermore, the need for an RSV vaccine is greatest for infants under 6 months of ages and the elderly. One of the reasons that vaccine has been unsuccessful until now is due to the difficultly of developing a vaccine that is immunogenic in newborns that have unique developing immune systems. Therefore, future studies performed with neonatal non-human primates would aid in attaining a better understanding of correlates of protection specific to neonates. The second limitation of this study revolves around the assays used to capture antibody-dependent functions. The correlates of protection based on antibody-dependent functions were measured ex vivo, and although they attempted to best capture functional responses, they cannot completely account for all the environmental factors present in vivo. Moreover, it is possible that not all humoral correlates of protection are detectable by peripheral blood. For example, is possible that the humoral landscape of mucosal environments is composed of antibodies which contain different characteristics and functions and should be further investigated. Finally, it is also possible both cellular and humoral immunity are necessary in order to fully understand mechanisms of protection against RSV. Therefore, further investigation into correlates of protection with improved animal models and adapted assays are necessary in order to best inform RSV vaccine design.

3.6 Methods

3.6.1 African Green Monkey vaccination

African Green Monkeys were maintained at New Iberia Research Center (NIRC) of University of Louisiana at Lafayette, New Iberia, LA, USA. Monkeys of either sex, with an average weight of 2.4 kg were used in this study. The animal study was approved by the University of Louisiana at Lafayette Institutional Animal Care and Use Committee (IACUC) and conducted in accordance with the US Public Health Service (PHS) Policy on Humane Care and Use of Laboratory Animals. A cohort of 32 African green monkeys that were RSV antibody-negative by RSV neutralization assay were split into eight groups of four and immunized with various RSV vaccine modalities or controls (Figure 3-1A). One group was immunized intramuscularly with 125μg of mDS-Cav1 mRNA vaccine in a lipid nanoparticle (LNP) formulation. The mRNA/LNP vaccine was prepared as described previously[132]. Three groups were given 125μg DS-Cav1 protein formulated with different adjuvants. Two of the three were administered intranasally, using 20% nanoemulsion[154] in a volume of either 0.1mL or 0.25mL, and the third was administered intramuscularly with 500μg of Adju-Phos® (InvivoGen, San Diego, CA), in a volume of 0.5mL. The DS-Cav1 protein was expressed and purified as described previously[155]. Two of the groups were vaccinated intramuscularly with 135μg RSV F nanoparticle vaccine prepared as described in[131] one of which was given with 500μg of Adju-Phos®. In addition to the six vaccine groups, two control groups were included in the study. One group was given 5.5 Log_{10} pfu RSV A2 intranasally in 0.25mL, and the other was given empty lipid nanoparticles at a dose equivalent to that used to deliver 125μg of mRNA, by the intermuscular route. Collectively, vaccines cover different routes of administration, adjuvant use, and vaccine type.

Three doses of vaccine were administered to each animal at Day 0, 29 and 56. Serum samples were collected on Day 0, 29, 56 and 70 to assess immune responses. After receiving a complete vaccination series, each animal was challenged with RSV A2 (2 x 5.5 Log 10 pfu) on day 70 to assess efficacy against RSV infection. The challenge dose was administered both intranasally and intratracheally, resulting in a very stringent challenge, with virus entering both the nose and lung. Nasopharyngeal swabs were collected daily for 15 Days post challenge and lung lavage samples were collected on Days 3, 5, 7, 10 and 14 post challenge to test for viral replication.

3.6.2 Primary Immune Cells

Fresh peripheral blood was collected by the MGH Blood bank from healthy human volunteers. All volunteers gave signed consent and were over 18 years of age, and all samples were de-identified before use. The study was approved by the MGH Institutional Review Board. Human NK cells and Neutrophils were isolated from fresh peripheral blood and maintained at 37 ˚C, 5% CO2 in RPMI with 10% fetal bovine serum, L-glutamine, penicillin/streptomycin.

3.6.3 Cell lines

THP-1 cells (ATCC), a monocytic leukemia cell line, were maintained in RPMI supplemented with 10% fetal bovine serum, L-glutamine, penicillin/streptomycin, HEPES, and beta-mercaptoethanol. THP-1 cells were grown at 37˚C, 5% CO2.

3.6.4 RSV RT-qPCR of African-Green monkey nasopharyngeal samples.

A Maxwell® 16 Viral Total Nucleic Acid Purification Kit (Promega) was used to extract RNA from 300uL of NP samples of African-Green Monkey on an automated extraction instrument called the Maxwell® 16 instrument (Promega). Samples were processed according to the kit's manufacturer's instructions. The kit contains all the necessary reagent in a convenient prefilled cartridge format. Briefly, the samples and lysis solution are mixed together and heated for 10 minutes. The lysates are added to the cartridge and loaded into the Maxwell® 16 instument. Purified RNA was then tested in the RSV Quantitative Reverse Transcriptase PCR assay. RSV RT-qPCR assays were perfomed using the Stratagene® Mx3005P instrument and Qiagen Quantitect® Probe RT-PCR kit. The assay was designed to allow for RSV strain typing and genome quantification by using the Nucleoprotein (N) gene as the target. Primers were designed to the conserved region of the N gene and the probe for RSV A is the probe contained the fluorescent reporter dye 6-carboxyfluorescein (FAM) at the 5′-end and the fluorescent quencher dye 6-carboxytetramethylrhodamin (TAMRA) at the 3′-end. Standard curves for quantification were generated by designing primer pairs to the N gene of RSV_MRKp17. Each sample was tested in duplicate and carried out in the Stratagene® Mx3005P instrument. The sensitivity of the assay is 10 copies.

3.6.5 Plaque assay of African-Green monkey BAL samples

Lung samples were determined via plaque assay on HEp-2 cells. Briefly, samples were series diluted and added in duplicate to 96-well plates as previously described[156]. Virus was incubated with samples for 1h at 37C in 5% CO2 incubator, then Hep2 cells was added into the virus/sample mixture and incubated for another hour. The plates were centrifuged at 300g x 10minutes. 1% methylcellulose was added into wells after centrifugation. Following 3 days incubation at 37 °C with 5% CO2, plaques were visualized by immunostaining using a cocktail of anti-F and anti-N antibodies (Merck, in house). Primary antibodies were incubated with the fixed cells for 1 h before anti-mouse IgG Alex488 conjugated secondary antibodies (Invitrogen) were added. Viral plaques were then imaged and counted using an EnSight imager reader 2.02 (PerkinElmer). Virus titers were expressed as pfu/g of tissue.

3.6.6 Quantification of antibody subclasses, isotypes, and Fc-receptor binding

Antibody subclasses, isotypes, Fc-receptor (FcR) binding levels were measured in a customized multiplexed Luminex assay, as previously described[157]. The assay enabled the assessment relative antibody quantification against VRC04 (provided by Merck &

Co., Inc., Kenilworth, NJ, USA) the RSV pre-fusion F glycoprotein. Briefly, the was first covalently cross-linked to fluorescent carboxyl- modified microspheres (Luminex) via NHS-ester bonds with EDC and Sulfo-NHS (Thermo Scientific). Then the antigen complexed beads were washed and followed by subsequent blocking prior to the addition of plasma at dilutions (IgG1 - 1:500, FcRs - 1:1000, and 1:100 for all other readouts). Following an overnight incubation at 4Cs shaken at 700rpm, immune complexed microspheres were then washed using a Tecan automatic plate washer with 0.1% BSA 0.02% Tween-20. Antigen-specific antibody quantification was detected with a PE-coupled antibody stains detection antibody for each respective subclass and isotype (Southern Biotech). Fc-receptors were quantified through fluorescently labeled with PE prior to the addition of the immune complexes (Duke Protein Production facility). Readouts were acquired using the iQue (Intellicyt) flow cytometry, and gating was preformed using the ForeCyt software. The gating strategy was first gating fluorescent bead regions and determining the median fluorescent intensity (MFI) of the PE channel and is the readout for antigen-specific antibody titers.

IgA titers were determined against prefusion or postfusion F protein essentially as described in[155]. Ninety-six-well ELISA plates (NUNC) were coated with 2 µg/mL purified recombinant RSV F disulfide (DS) cavity-filling (Cav1) protein DS-Cav1 or purified recombinant postfusion RSV F protein as described in[130] and[158], respectively, and incubated at 4 °C overnight. The plates were then washed and blocked for 1 h at room temperature with 3% non-fat milk dissolved in PBS-T. Sera from AGM's was serially diluted 4-fold in blocking buffer, transferred to the RSV F coated plates, and incubated for 2 h at room temperature. The plates were then washed three times with PBS-T. Following the plate wash, HRP conjugated goat anti-AGM IgA secondary antibody (Invitrogen), diluted at 1:3000 in blocking buffer, was added to the plates and incubated for an additional 1 h. Plates were washed again and developed with SuperBlu Turbo TMB (Virolabs). The reaction was stopped after 5 min and absorbance was read at 450 nm on a VersaMax ELISA microplate reader (Molecular Devices). Endpoint titers were defined as the reciprocal of the end point dilution at which the serum sample has an optical density (OD) signal greater than or equal to two (2×) times that of the background.

3.6.7 Effector functional assays

Three bead based functional assays were performed during the study including antibody-dependent cellular phagocytosis (ADCP), antibody-dependent neutrophil phagocytosis (ADNP) and antibody-dependent complement deposition (ADCD)[159–162]. Both phagocytosis assays were performed with yellow fluorescent neutravidin (Thermo Fisher) beads which were first conjugated to biotinylated antigen and followed then incubated with diluted plasma samples (1:100) at 37C for 2 hours which provides the formation of immune complexes. For ADCP THP-1s (ATCC), monocyte cell line, were incubated with immune complexes at 125,000 cells/mL at 37 ˚C for 16 hours. After the

incubation period, cells were fixed with fixed with 4% paraformaldehyde (Alfa Aesar). For ADNP, following the formation of immune complexes white blood cells from healthy donors were added at 125,000 cells/mL at incubated at 37˚C for one hour. White blood cells were isolated from ACD-treated whole blood through lysis of red blood cells by incubating whole blood for 5 min in ACK lysis buffer followed by multiple sequent washes with PBS. Afterwards, neutrophils were stained with an anti-CD66b Pacific blue antibody (Biolegend) and fixed with 4% paraformaldehyde (Alfa Aesar). For quantification of antibody-dependent deposition of C3, first red fluorescently neutravidin beads (Thermo Fisher) were coupled to the antigens and immune complexed were formed in the same process as for ADCP and ANDP. Then, lyophilized guinea pig complement (Cedarlane) was reconstituted according to manufacturer's instructions and diluted 1:50 in gelatin veronal buffer with calcium and magnesium (GBV++) (Boston BioProducts). Afterwards, C3 was detected with an anti-C3 fluorescein-conjugated goat IgG fraction detection antibody (Mpbio).

In addition to the three bead-based assays described above, antibody-dependent NK cell activity was measured through an ELISA-based assay, as previously described[163]. In short, plates were first coated with 3µg/mL of antigen followed by blocking overnight at 4˚C. the following day NK cells were isolated from healthy donors via RosetteSep (Stem Cell Technologies) and set overnight in 1 ng/mL of IL-15 (Stemcell) at 150,000 cells/mL. The following day, diluted plasma samples (1:100) were added to the antigen-coated plates and incubated for 2h at 37˚C. Then, NK cells were stained with CD107a PE-Cy5 (BD), Golgi stop (BD) and Brefeldin A (BFA, Sigma Aldrich) at 250,000 cells/mL for 5h at 37˚C. Next, cells were fixed (Perm A, Life Tech) and stained for with anti-CD16 APC-Cy7 (BD), anti-CD56 PE-Cy7 (BD) and anti-CD3 PacBlue (BD). Following staining, cells were permeabilized using Perm B (Life Tech) and intracellularly stained with an anti-MIP-1β PE (BD).

All effector function assays were acquired with the iQue (Intellicyt) flow cytometer. ADCP, was gated on singlets and bead-positive cells, and neutrophils on singlets, CD66b, and bead positive cells. A phagocytosis score was determined as the percent of bead-positive cells times the MFI of bead-positive cells divided by 10000. ADCD was determined as the MFI of C3 deposition, and NK cells were gated as CD3-, CD16+ and CD56+ and the percentage of cells positive for CD107a, MIP-1β, and IFNγ were reported.

3.6.8 Neutralization

Serum neutralization assays were conducted as described in[164]. Briefly, AGM sera were heat inactivated and serially diluted into a 96-well plate. The sera were combined with RSV (Long) at a final concentration of 100 pfu/well. HEp-2 cells were added to each well and the plates were incubated at 37 °C for 72 h. Cells were then washed and fixed with acetone. Each well was then incubated with a combination of two RSV anti-F mouse mAbs followed by biotinylated horse anti-mouse IgG antibody (Vector laboratories). The

signal was developed by adding a cocktail of IRDye 800CW Streptavidin (Li-Cor Biosciences, 1:1000 final dilution), Sapphire 700 (Li-Cor Biosciences, 1:1000 dilution) and 5 mM DRAQ5 solution (Biostatus Ltd, 1:10,000 dilution) in assay diluent. Plates were read on an Aerius® Automated Imaging System. Titers were calculated by four-parameter curve fit using GraphPad Prism® 7 software.

3.6.9 Quantification and statistical analysis

All analyses were performed using python version 3.6.8. Raw data and custom code are available in Supplementary Information. Networks were visualized in Cytoscape.

3.6.10 Principal component analysis and phase portraits

Principal component analysis for a given period was performed by first scaling and centering all features and then decomposing data into two principal components. Scores were then visualized on both principal component one and two will reporting the variance explained by each component. The loadings plot was constructed using each feature's weight on each component and was visualized as vectors corresponding to their weight in each dimension. The phase portraits were constructed by decomposing each time point separately to ensure sample independence. Prior to the principal component analysis (PCA) features were scaled and centered with respect to a given time point. Then using the first two principal components from each time point the paths for the sample were visualized. The mean path of a vaccine group was determined by averaging the vectors for each sample component-wise.

3.6.11 Classification of compartment specific protection

The classification models were trained to distinguish NHPs that were protected and unprotected at days three and five post challenge with a minimal set of features, to avoid overfitting. Protection was defined as viral loads below 20 pfu/mL. Only vaccinated NHPs were included, and PBS controls were subtracted from all features, and all data was scaled and centered. The models were constructed on the whole dataset using a partial least squares discriminant analysis (PLS-DA) classifier and recursive feature elimination based on feature variable of importance in projection (VIP), and a final model was chosen based on classification accuracy[165,166]. The performance and robustness of the final model was then assessed by contrasting a negative control null model built on permuted data, with multiple iterations of fivefold cross-validation used to generate classification accuracies. Models were generated 100 times, and the permuted control was generated in the same process as above shuffling labels randomly for each repetition. Predicted and true outcomes were then compared to determine accuracy. Performance was defined as the exact P values of the tail probabilities of the true distributions within the control distributions. Reported are the median p-values over the course of every repetition[107].

3.6.12 Spearman correlation network

The correlation network was constructed to visualize the additional humoral immune features that were significantly linked to the selected minimal biomarkers, to provide enhanced insights into the biological mechanisms by which antibodies may provide protection following infection. In brief, antibody features that were significantly correlated with a Holms-Bonferroni correction to the final selected PLS model selected-features were defined as co-correlates. Significant spearman correlations above a threshold of $|r| > 0.75$ were visualized within the networks

3.7 Supplementary figures

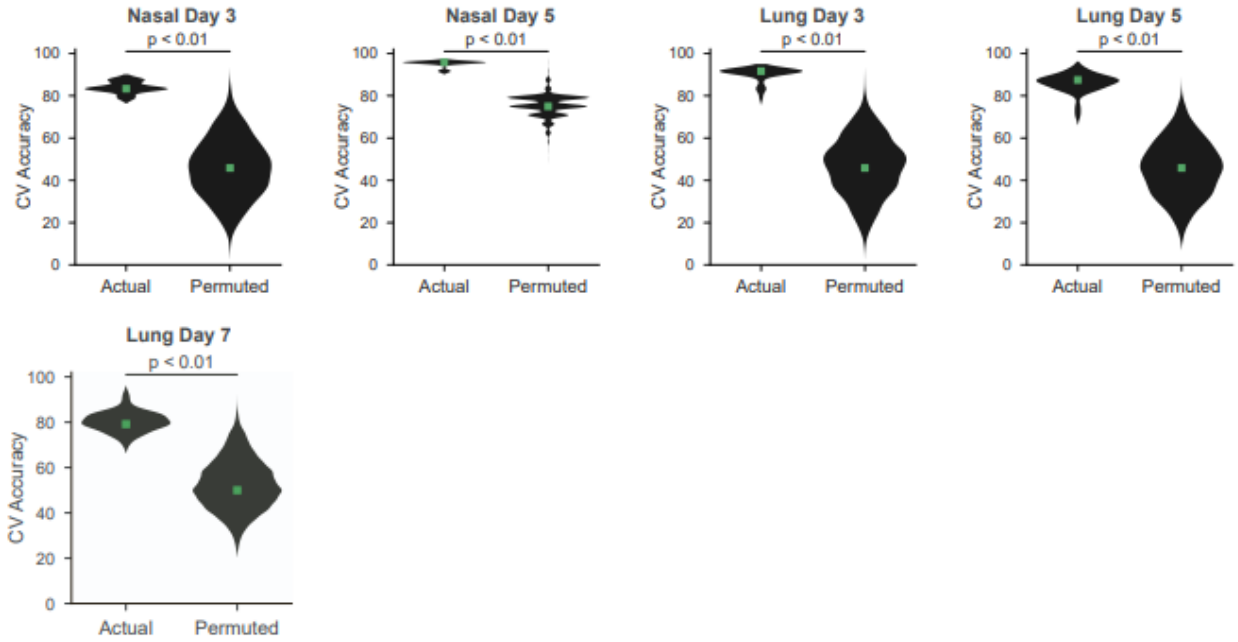


Figure 3-S1. Timepoint and compartment model performance

The violin plots show the distributions of repeated classification accuracy tests for each time point and compartment model using the actual data and shuffled labels highlighting the performance of the model based on the final selected features of that model. The green squares indicate the median accuracies, and exact P values are defined as the tail probabilities of the true distributions within the control distributions.

3.8 Acknowledgements

We thank Terry and Susan Ragon for their support.

Chapter 4

Distinct early serological signatures track with SARS-CoV-2 survival

The contents of this chapter were published as:

Caroline Atyeo*, Stephanie Fischinger*, Tomer Zohar*, Matthew D. Slein, John Burke, Carolin Loos, Denise J. McCulloch, Kira L. Newman, Caitlin Wolf, Jingyou Yu, Kiel Shuey, Jared Feldman, Blake Marie Hauser, and Tim Caradonna, Aaron Schmidt, Todd J Suscovich, Caitlyn Linde, Yongfei Cai, Dan Barouch, Edward T. Ryan, Richelle C. Charles, Douglas Lauffenburger, Helen Chu, and Galit Alter. "Distinct early serological signatures track with SARS-CoV-2 survival". *Immunity* 53, 524-532. e4 (2020)

**These authors contributed equally to this work.*

See the online publication for any references to supplementary material not included in this thesis.

Contributions

S.F., C.A., H.C., and G.A. designed the study. S.F., C.A., M.S., and J.B. performed the immunologic assays. H.C. collected and selected the clinical specimens. T.Z., C.L. and D.L performed all analyses. E.T.R. and R.C.C. participated in sample collection, study design and interpretation of data. A.G.S., J.F., T.C., B.M.H, and Y.C. provided purified proteins. G.A. wrote the paper with all co-authors.

4.1 Abstract

As SARS-CoV-2 infections and death counts continue to rise, it remains unclear why some individuals recover from infection whereas others rapidly progress and die. While the immunological mechanisms that underlie different clinical trajectories remain poorly defined, pathogen-specific antibodies often point to immunological mechanisms of protection. Here, we profiled SARS-CoV-2–specific humoral responses on a cohort of 22 hospitalized individuals. Despite inter-individual heterogeneity, distinct antibody signatures resolved individuals with different outcomes. While no differences in SARS-CoV-2-specific IgG levels were observed, spike–specific humoral responses were enriched among convalescent individuals, whereas functional antibody responses to the nucleocapsid were elevated in deceased individuals. Furthermore, this enriched immunodominant S-specific antibody profile in convalescents was confirmed in a larger validation cohort. These results demonstrate that early antigen-specific and qualitative features of SARS-CoV-2-specific antibodies, point to differences in disease trajectory, highlighting the potential importance of functional antigen-specific humoral immunity to guide patient care and vaccine development.

4.2 Introduction

SARS-CoV-2 is the newest coronavirus to cross into the human population[167,168]. Millions of infections have been diagnosed (WHO); however, the number of asymptomatic carriers is likely to far exceed these numbers[169]. While the rapid spread of SARS-CoV-2, even during the asymptomatic phase of this infection, is alarming, more harrowing is our inability to predict disease trajectories among symptomatic individuals. In the absence of therapeutics and vaccines as countermeasures for this infection, there is an urgent need to begin to map the evolution of immunity to the pathogen to guide patient care and future immune interventions.

Although both antibody responses and T cells have been linked to disease resolution[170], and neutralizing antibodies have been demonstrated to block infection in small animal models[171], little is known about the antibody features that are important for protection. Neutralizing antibodies develop in the majority of SARS- and MERS-infected individuals[172,173]; however, the virus can mutate to overcome these antibody responses[174,175]. Passive immunization studies with both neutralizing and poorly neutralizing antibodies have shown protection in lethal MERS infection in mice[176,177]suggesting that both the neutralizing and extra-neutralizing functions of

antibodies may play a critical role in control and resolution of disease. Moreover, recent studies have found lower neutralization titers in younger individuals and higher neutralization among individuals with severe disease[178,179], suggesting that antibodies may depend on additional mechanisms to clear the virus.

Antibody dynamics during the acute window of infection have been linked to differential outcomes across infections, including HIV[180], influenza[181], and Ebola virus infection[182]. Specifically, the selection of specific antibody subclasses and functional profiles is heavily influenced by inflammatory cascades and may not only forecast disease outcomes but also point to antibody mechanisms of action vital in early pathogen control and clearance. However, whether identifiable antibody functional profiles across SARS-CoV-2 antigen-specificities evolve early following infection that track differentially with disease outcome is unknown. In this study, we assembled two cross-sectional sample sets of SARS-CoV-2-infected individuals at the time of hospital admission to begin to comprehensively profile the evolution of the early SARS-CoV-2 S-specific response and to define antibody features that are predictive of disease outcome. Through this analysis, we found that deceased and convalescent individuals present different humoral profiles, with a more S-focused response in individuals who convalesced and a stronger N-specific response in individuals who succumbed to disease.

4.3 Results

4.3.1 Early SARS-CoV-2 antibody profiles in individuals that ultimately convalesce or pass away

Across infectious diseases, pathogen-specific antibodies can serve as biomarkers of infection and aid in the early control and clearance of infection by blocking host-pathogen interactions and/or recruiting innate immune functions[183]. In order to investigate whether early SARS-CoV-2-specific humoral immune responses differ across individuals that ultimately recover or die from infection, a cohort of 22 hospitalized SARS-CoV-2-infected individuals, of whom 12 recovered and 10 died, was profiled. Samples were collected at hospital admission, all recruited within the first 20 days following symptom onset (Table 1 and Figure S1) at the University of Washington, Seattle, one of the earlier epicenters in the US[184]. Population demographics largely resemble those previously reported[185], including elevated numbers of elderly men in the subset that died.

To profile the SARS-CoV-2-specific humoral immune response, we performed Systems Serology to determine the biophysical and functional characteristics of SARS-CoV-2-specific antibodies that recognize the SARS-CoV-2 spike (S), the S-derived receptor binding domain (RBD), and the nucleocapsid (N). The titers of SARS-CoV-2-specific isotypes and subclasses, the Fcγ-receptor binding profiles, neutralization, as well as antigen-specific innate effector functions were measured. Heterogeneous responses were observed across both populations (Figure 1A and Figure S2), and

convalescents did not appear to possess quantitatively superior immune responses that could explain their different later disease course. Univariate analyses further confirmed that no significant differences were observed in SARS-CoV-2-specific IgG1 or IgA1 titers across S, RBD, and N (Figure 1B-C, and Figure S2). Conversely, subtle distribution differences were observed for SARS-CoV-2-specific IgM responses, with a slight shift towards higher S-specific IgM among survivors, and a trend towards increased N-specific IgM responses among individuals who died (Figure 1C). Functional antibody profiles displayed similar distributions across the cohorts for antibody dependent cellular phagocytosis (ADCP) (Figure 1D) and neutralization (Figure 1G). Surprisingly, RBD-specific antibody-mediated NK cell degranulation (NKD) and antibody dependent neutrophil phagocytosis (ADNP), both driven through related Fcg-receptors Fc□R3A and FcgR3B, respectively, trended towards increases among individuals who died (Figure 1D-F). Antibody measurements were minimally influenced by time since symptom onset (Figure S1), suggesting equivalent evolution of humoral immune responses across groups. However, no single antibody feature could discriminate between the groups.

Characteristics	Convalescent (n = 12)	Deceased (n = 10)
Female sex – no. (%)	4 (33.3)	3 (30)
Age Range – No. (%)		
Younger than 49	3 (25)	1 (10)
50-59	4 (33.3)	0 (0)
60-69	4 (33.3)	2 (20)
70-79	0 (0)	4 (40)
80 and older	1 (8.3)	3 (30)
Race or Ethnic Group – No (%)		
Asian	2 (16.7)	1 (10)
Black	0 (0)	1 (10)
White	9 (75)	7 (70)
Missing data	1 (8.3)	1 (10)
Median days from onset of symptoms to sample collection (IQR)[a]	13.5 (15–8)	7 (12–5)
Median days spent in ICU (IQR)[b]	13 (15–9)	13(14–9)
Median viral load (IQR)[c]	28.3 (30.4–26.5)	26.4 (28.375–21.725)
Interventions – No./Total No. (%)		
Chloroquines	6/12 (50)	7/10 (70)
Remdesivir	9/12 (75)	7/10 (70)
Tocilizumab	3/12 (25)	0/10 (0)
Antibiotics	8/12 (66.7)	8/10 (80)
Consequences of Disease – No./Total No. (%)		
Acute respiratory distress syndrome	5/12 (41.7)	6/10 (60)
Non-ST-elevation myocardial infarction	1/12 (41.7)	5/10 (50)

[a]IQR: interquartile range

[b]For 4 of the deceased individuals, days from symptom onset was unknown.

[c]For half of the recovered individuals, viral load measurements were not available.

Table 4-1. Demographics of SARS-CoV-2 cohort from Seattle

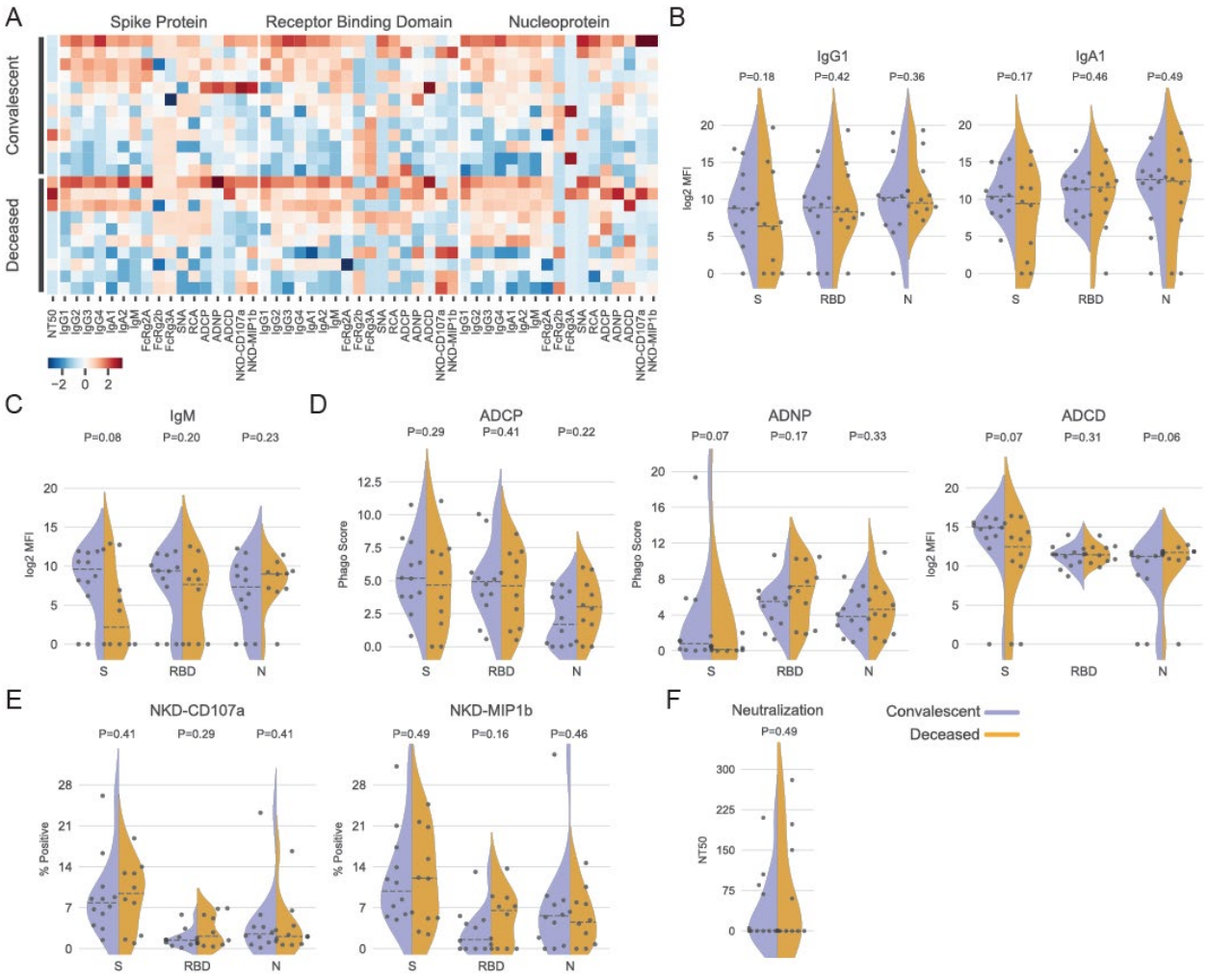


Figure 4-1. Heterogeneity in antibody responses across SARS-CoV-2 antigens in individuals that recover or pass away

22 plasma samples from SARS-CoV-2 infected individuals were profiled at the time of hospitalization against SARS-CoV-2 S, RBD and N antigens. (A) The heatmap shows the humoral immune responses across individuals that later passed away (deceased) or recovered (convalescent). The heat-map is split by SARS-CoV-2 S, RBD and N antigens. Rows correspond to individuals. Columns correspond to antibody features (background subtracted and z-scored) including neutralization, isotype, subclass, and antibody effector functions. High responses are shown in red and low responses are depicted in blue. (B-G) Violin plots show the distribution of each antibody feature split across convalescent (purple) and deceased (orange) individuals across antigens. The dashed gray line indicates the median value of each distribution. A two-sided Mann-Whitney U test was used to calculate uncorrected p- values, shown in the figure. No significance was detected after a Holm-Bonferroni correction for multiple hypothesis testing.

4.3.2 Differences in antibody profile coordination between groups

Beyond univariate differences, emerging data point to a critical role for humoral immune response coordination as a predictor of protection in some infections[137,186]. Given the polyclonal nature of the early humoral immune response, multiple functions or features may simultaneously contribute to differential control and clearance of infection.

Correlation matrices split by group were used to examine the relationships between antibody isotypes or subclasses and antibody-dependent effector functions across the groups (Figure 2A). Within both groups, isotypes and subclasses were highly correlated. Conversely, the relationship between isotype or subclass and functions differed across the two populations. Stronger correlations between titers and functions were observed in convalescent individuals (Figure 2A). Disparities were observed in both NK cell and neutralizing antibody coordination between the two groups. Though not significant, individuals who died exhibited correlated isotype or subclass responses with monocyte and neutrophil phagocytosis, but negative and generally poorer correlations of NK cell activating and complement recruiting antibody responses with all other functions (Figure 2A), suggesting that individuals who pass away develop a functionally biased humoral immune response. While IgG1 responses were associated with all functions across the individuals that later went on to die, diversified isotype and subclass responses were largely inversely correlated with antibody-dependent complement deposition (ADCD) and natural killer cell (NK) NK functions. This observation suggests that these individuals leverage isotype and subclass diversification in a manner that may preclude the full deployment of the humoral immune response.

Conversely, convalescents overall displayed a more uniform correlation profile across subclass and isotype responses and antibody effector function. However, while neutralizing antibody responses were co-induced with isotype and subclass and effector functions among individuals who died, neutralizing antibody responses were largely inversely correlated to all antibody responses among individuals that went on to recover, suggesting a divergent evolution of the antigen-binding and constant domain of the antibody across these populations. These data highlight multiple early functional differences in SARS-CoV-2 specific humoral immunity between the groups.

To further probe the overall humoral profile between groups, the mean percentile of each antibody metric was determined across SARS-CoV-2-antigen specificities for both populations (Figure 2B). The nightingale rose plots reveal that deceased individuals exhibited a more N-focused humoral immune response compared to the S-centric response elicited among convalescents. In particular, higher S-specific ADCD, ADNP, ADCP and enhanced IgG1, IgA1, and IgM responses were observed among survivors. In contrast, S-specific NK cell activating responses were enriched in deceased. Unexpectedly, RBD-specific responses were largely enriched among individuals that went on to pass away, with the exception of RBD-specific monocyte phagocytosis which was enriched among individuals that went on to survive. These data point to both antigen-specific and antibody-effector differences early in infection that differ by clinical trajectory.

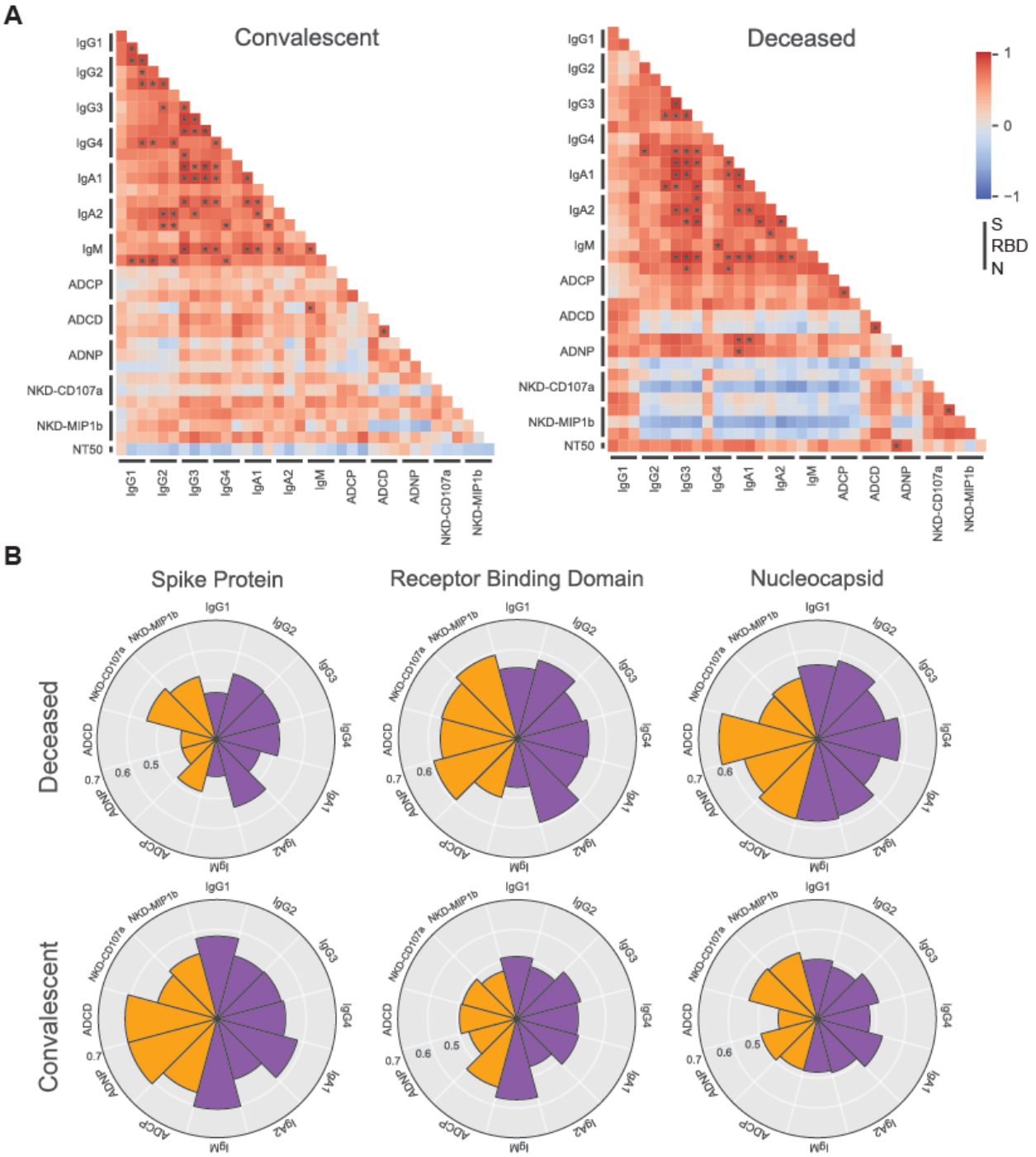


Figure 4-2. Deceased Individuals show less coordinated and N-directed antibody responses.

(A) The correlation heatmap shows pairwise spearman correlation matrices of antigen-specific antibody titers and effector functions for convalescent (left) and deceased (right) patients. For each feature analyzed, the bar covers the S, RBD, and N antigens, shown in the legend on the right. Statistical significance is indicated by gray asterisks with Holm-Bonferroni correction for multiple hypothesis testing ($p<0.001$). Negative correlations are indicated in blue, and positive correlations are denoted in red. B: The

nightingale rose plots show the mean percentile of antibody features within the deceased (top) and convalescent (bottom) groups. Plots represent the S, RBD, and N-specific responses across in deceased (top) and convalescent (bottom) individuals. Each wedge represents a SARS-CoV-2 antibody feature. The size of the wedge depicts the magnitude of the value. The colors represent the type of feature: orange-antibody functions and purple-antibody isotypes and subclasses.

4.3.3 Defining signatures that differentiate disease trajectory

Given the unique correlation and immunodominance profiles across the groups (Figure 2A and B), we next aimed to define whether a minimal set of features could be identified that could segregate individuals with different clinical outcomes. To this end, feature down-selection was performed to avoid overfitting, followed by partial least squares discriminant analysis (PLSDA) to visualize differences (Lau et al., 2011). Despite the small numbers, separation was observed across the groups (Figure 3A). All antibody features as well as sex and interventions (Table 1) were included in the analysis, and as few as 5 features were sufficient to drive separation across the subjects (Figure 3A and B). S-specific IgM and IgA1 responses were enriched in survivors, whereas N-specific complement activity (ADCD), IgM, and IgA1 titers were enriched in individuals who died. These data likely relate to the immunodominant shift towards S in convalescent individuals and towards N in deceased individuals (Figure 2B). Model performance was evaluated using leave-one-out cross validation, to test that significance of the model using different sets of subjects and to test outlier effects. The model clearly outperformed (Cliff's Δ) permuted and size-matched random controls (Figure 3C). Moreover, sensitivity analysis, evaluating model performance with the removal of individual outliers, highlighted the minimal impact of any given individual (Figure S3A). Furthermore, individual model features only possessed modest predictive power in resolving the groups, but collectively, combining all 5 features -in Latent Variable 1 (LV1)- exhibited improved predictive accuracy (Figure 3D). Confounding features, such as days since symptom onset, sex, age, and viral load were also over-layed on the PLSDA scores plot (Figure S3B-F), highlighting the limited capacity of any of these features to distinguish individuals into those who convalesced or died. Furthermore, at individual levels, these demographic factors were poorly predictive of disease outcome, underperforming classification compared to the LV1 classification model (Figure S3G). Thus, a minimal set of SARS-CoV-2 humoral profiles, rather than demographic information, appear to significantly resolve individuals who later went on to die from those who recover.

Given that the feature down-selection algorithm selects a minimal set of features to avoid overfitting, a co-correlates network was used to explore additional features that may distinguish these two groups (Figure 3F). A larger set of co-correlates can help provide mechanistic clues related to the immunologic mechanisms by which antibodies contribute to control and clearance of infection. Thus, a co-correlate network was built highlighting the relationship of model-selected features (large nodes) with additional highly correlated features (smaller nodes). Features enriched among individuals who

later died, included N-specific IgM and IgA2, that were linked to a large number of additional N- and RBD-specific poorly functional antibody features. For example, correlates of risk were linked to the induction of less-functional IgG subclasses, IgG2 and IgG4, pointing to the early rise of dysregulated or less functional humoral immune responses as biomarkers or even drivers of ineffective control or clearance of infection. Conversely, S-specific IgM titers, enriched in convalescent individuals, were correlated with functional S-specific IgG3 responses, RBD-specific IgM, and S-specific monocyte and neutrophil phagocytosis. Moreover, S-specific IgA1 responses, also enriched among convalescents, were linked to RBD-specific complement activation (ADCD) and S-, RBD-, N-specific FcγR2A binding, the Fcγ-receptor involved in phagocytosis. Given our emerging appreciation for the role of complement and phagocytosis in vaccine mediated protection against SARS-CoV-2 (Yasui et al., 2014), these data potentially argue for a similar role for these functions in natural protection against disease. Moreover, the data also highlight the potential importance of a less N-focused, but more functional S-specific phagocytic response as an early correlate of recovery from infection.

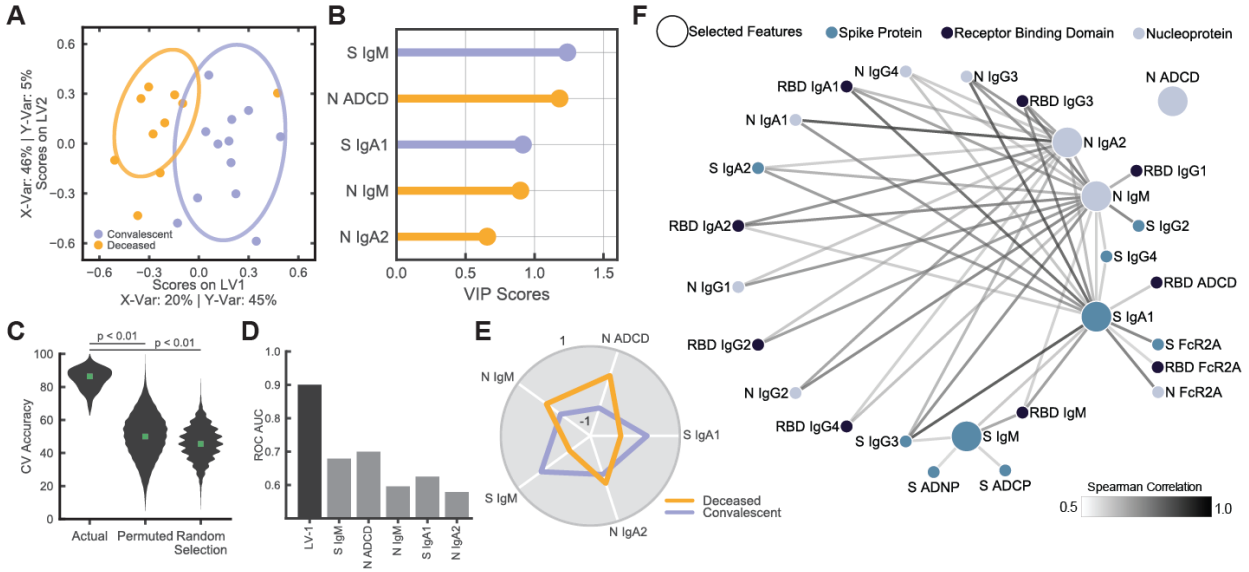


Figure 4-3. Select antibody features distinguish convalescent and deceased individuals

(A) The PLSDA scores plot shows the degree of discrimination that was achievable across the groups following feature-down selection. Each dot represents an individual: convalescent (purple) and deceased (orange). Ellipses correspond to the 95% confidence intervals for each group. (B) The line graph shows the variable importance in projection (VIP) score of the selected features. As few as 5 features were required to separate the groups. The magnitude indicates the importance of the feature in driving separation in the model. The color of the feature corresponds to the group in which the feature is enriched. (C) The violin plots show the distributions of repeated classification accuracy tests using the actual data, shuffled labels, and randomly selected size-matched features illustrating the performance and robustness of the model. Green squares indicate the median accuracies. (D) The predictive power of the model built on the selected features is shown in the LV1 column. In addition, the predictive power of each individual selected feature is represented in grey. The predictive power is illustrated as the AUC of the

ROC curves for the model (LV1) or each feature alone. (E) The radar plot shows the z-scored univariate values of the selected features across both groups. (F) The correlation network illustrates the co-correlated features (small nodes) that are significantly correlated with the Model-selected features (large nodes). Edge transparency corresponds to correlation strength. Antigens are indicated by different colors (S: teal, N: grey, and RBD: black).

4.3.4 Validation of the skewed S-specific response in convalescents

Collectively, the data point to a shift in immunodominance of spike versus nucleocapsid functional antibody responses. To test this hypothesis, we next compared the overall ratio of Spike(S):Nucleocapsid(N)-specific antibody isotype, subclasses, and functions across the groups (Figure 4A and Figure S4A). As expected, several antibody features were selectively biased towards S-immunity in the convalescents compared to individuals that later died, including IgM, ADCP, ADNP, and ADCD. Whether these effects were exclusive to this group of individuals from Seattle or could be generalized was next addressed in a second larger cohort of acutely infected individuals from Boston, of which 20 individuals convalesced and 20 died. Similarly, to the Seattle cohort, Boston samples were profiled in the first 20 days following symptom onset (Table 2). Similar to the Seattle discovery cohort, though differences were observed in S- and N-specific immune responses at a univariate level none passed multiple hypothesis correction (Figure S4B). Yet, when S:N ratios were compared across features, convalescent individuals exhibited a bias towards elevated S-specific humoral immunity compared to N-specific immunity in contrast to individuals who went on to later pass away (Figure 4B and Figure S4C). Thus, to ultimately capture the extent of S:N skewing across the groups, the number of features that had greater S than N responses were summed across convalescents and deceased individuals and compared within each cohort (Figure 4A and B). In both cohorts, a significant enrichment of S:N immunity was observed in convalescents (Figure 4C). Therefore, these findings suggest that a consistent overall shift in S:N immunity early in SARS-CoV-2 infection may have a protective role and aid in recovery from severe disease.

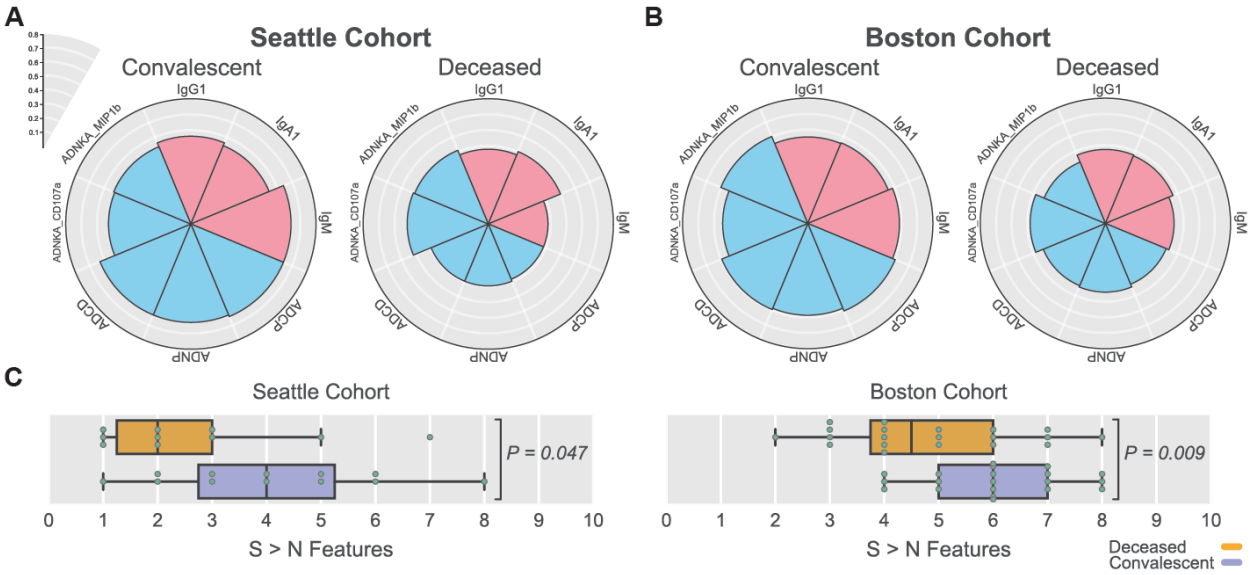


Figure 4-4. Converging shift in immunity across a second acute infection cohort

The nightingale rose plots show the mean percentile of the Spike:Nucleocapsid (S:N) ratio of each read-out are depicted for the (A) Seattle or Discovery cohort and the (B) Boston or Validation cohort for the convalescents (left) and the deceased (right). Titers are shown as pink wedges and functions as blue wedges. (C) The whisker box plots show the number of S-features which are greater than their N-counterparts for all individuals in the Seattle or Discovery cohort (left) and the Boston or Validation cohort (right). Differences across the 2 groups were assessed using a one-sided Mann-Whitney U test.

Characteristics	Convalescent (N = 20)	Deceased (N = 20)
Female sex – no. (%)	6 (30)	6 (30)
Age (IQR)	56 (63–45)	78 (81.5–68)
Median days from onset of symptoms to sample collection (IQR)	9 (14.25–7.5)	8.5 (12–6)

Table 4-2. Demographics of SARS-CoV-2 cohort from Boston

4.4 Discussion

Both cellular and humoral immune responses have been linked to protection against several coronaviruses[187]. Importantly, antibodies represent pathogen-specific markers of exposure, serve as powerful biomarkers of disease activity, and often point to immunological mechanisms of protection able to guide therapeutic or vaccine development[183]. By deeply profiling the SARS-CoV-2 humoral immune response early in infection, here, we defined a unique SARS-CoV-2-specific humoral signature associated with later disease outcomes. A combination of five SARS-CoV-2-specific antibody measurements were sufficient to distinguish individuals with different disease trajectories in a cohort from Seattle, including antibody measurements to S and N, with an overall enhanced S-centric response in individuals who recovered from infection. S-specific phagocytic and complement activity were enriched early in individuals that recovered from infection. This signature was confirmed in a second, larger SARS-CoV-2 infection cohort from Boston, where convalescent individuals exhibited a higher S:N ratio in their humoral immune response. These data point to early diverging humoral immune responses that may mark more effective immunity and suggest that functional antibodies directed against S might be beneficial for SARS-CoV-2 disease trajectory.

In SARS-CoV-1 and SARS-CoV-2 infection, N is highly immunogenic, with N-specific humoral immune responses arising concurrently with S-specific humoral immunity[188–190]. However, immunization of hamsters with a vector expressing N offered no protection against SARS-CoV-2 challenge despite a strong anti-N response, whereas immunization with the same vector expressing S protected hamsters against challenge[191]. It is estimated that 100 copies of S and 1000 copies of N are incorporated into each virion[192], suggesting that 10-fold more N may be produced compared to S during infection to effectively generate viral progeny. Due to the high amounts of

nucleocapsid, N-directed responses may be indicative of higher disease burden and increased antigen exposure. However, the similarity in viral loads between the individuals who recovered and those who died does not support this hypothesis. Rather, the data point to compromised evolution of S-immunity in individuals that later pass away. The potential beneficial role of S-targeted immunity in viral control is reinforced in new studies in non-human primates (NHP), demonstrating elevated and robust functional humoral immune responses to S, rather than RBD and N, following primary infection that were associated with protection upon re-exposure to the virus[193].

It is well known that timing of sampling may influence humoral profiles, where sampling time could result in the comparison of immature versus mature immune responses. Despite the sampling differences in the group, comparable titers were observed across the convalescents and individuals that ultimately passed away. Moreover, similar overall functional profiles were also observed, suggesting that the humoral immune responses were comparable in magnitude across the two groups. Additional analysis of the influence of sampling time on the spread of the antibody profiles in the PLSDA highlighted a minimal influence of time from symptoms on overall antibody profile variation and the time of sampling exhibited a minimal predictive power in classifying individuals into convalescents or deceased. Yet, longitudinal analyses will be illuminating providing further information on the evolving humoral immune response that tracks with protection from infection.

Emerging data point to higher mortality among the elderly and across the genders[194]. Along these lines, individuals who passed away were on average older than those who convalesced. Age can have a profound effect on immune function, and though this study was not suited to explore the relationship between age, outcome, and humoral responses, future larger studies across age groups could provide insights on the differential susceptibility among the elderly. However, the impact of age, sex, and viral loads illustrated a minimal influence of each of these variables on the overall variability of the humoral immune responses. Additionally, the individual predictive power of these demographic variables were lower than the predictive power of the model selected antibody features (LV1).

While S-specific antibodies able to recruit NK cell activity were expanded in individuals that went on to pass away, pointing to a potentially negative influence of NK cells, coordination of NK and phagocytic activity was enriched among convalescents. These seemingly contradictory data point to the potential importance of synergy between innate immune effector functions. While NK cells have been implicated both in protection[119,162,195] and pathology[143], it is possible that the evolution of antibodies able to harness the cytotoxic power of NK cells to eliminate infected or phagocytic cells may play a critical role in elimination and clearance of the infection. Interestingly, this coordination was associated with the synergistic evolution of a broader isotype and subclass specific response among convalescents. However, whether additional changes in antibody-Fc-glycosylation also contribute to this unique functionalization of antibody

isotypes and subclasses, enabling coordination, remains unclear but could point to promising target immune profiles that may confer the greatest level of protection against the virus.

4.5 Limitations of study

There are a number of limitations in this study. First, because these samples were collected early during the COVID-19 pandemic in the US, the Seattle study included a small number of participants, and the groups were not age or sex matched. Confounding factors such as timing of sampling, sex and age are all known to influence SARS-CoV-2 infection and disease trajectory. While antibody profiles clearly segregated individuals that survived compared to that did not survive, more limited variation in antibody profiles were observed across age, sex, viral load and days from symptom onset. However, among the co-morbidities, age was the second major driver of variation in antibody profiles, pointing a potentially critical role for age-associated defects in Fc-variation that may contribute to altered antiviral immunity to SARS-CoV-2 and beyond. The larger validation cohort from Boston identified a similar humoral signature that discriminated survivors from non-survivors, highlighting the conserved nature of this immunological signature, independent of demographic characteristics. Whereas this study only attempted to understand the humoral disparities between convalescent and deceased individuals in a cohort of severly infected individuals, further studies may attempt to define humoral profiles able to further classify individuals aross the clinical trajectory spectrum ranging from asymptomatic to severe disease.

Collectively, the data presented here argue for the evolution of distinct antigen-specific and functional humoral immune responses early in SARS-CoV-2 disease. While further analysis on longitudinal cohorts may provide further mechanistic insights on the specific role of antibodies in control and clearance of infection, here we validated an early functional humoral immune signature that appears to predict disease progression across two distinct cohorts. Linked to emerging animal model experiments, the correlates defined here may provide key mechanistic insights to guide therapeutic and vaccine design efforts.

4.6 Methods

4.6.1 Sample set

Plasma samples from 22 SARS-CoV-2 patients from Seattle were profiled for anti-SARS-CoV-2 antibody responses (Table 1). Patients who tested positive for SARS-CoV-2 by real-time reverse-transcriptase–polymerase-chain-reaction (RT-PCR) of a nasopharyngeal swab were enrolled in the study upon hospital admission, and samples after admission were included in this study (Figure S1). All enrolled participants gave written, informed consent. The enrolled hospitalized 22 individuals were monitored over the course of their stay, and final outcomes were reported. 12 individuals convalesced and were healthy enough to be discharged, whereas 10 individuals died. Demographic information including age, race, and interventions are summarized across the two groups (Table 1 and Data S1).

As a validation cohort, a cohort of 40 individuals from MGH in Boston were enrolled, all participants tested positive for SARS-CoV-2 by RT-PCR and they were monitored over their hospital stay. Samples at time of hospitalization were included in this study. Outcomes were reported as deceased or discharged. Demographics and clinical data for the validation cohort are summarized in Table 2.

All experimental data was performed in two technical and two biological (for primary cell assays) replicates and the average value was used throughout the study. This study was approved by the University of Washington Human Subjects Division Institutional Review Board.

4.6.2 Primary immune cells

Primary immune cells were isolated from fresh peripheral blood from healthy human volunteers collected by the MGH Blood bank or the Ragon institute. The study was approved by the MGH Institutional Review Board. All subjects were over 18 years of age and provided informed consent. All samples were completely de-identified prior to use. Human NK cells and neutrophils isolated from fresh peripheral blood were cultured in RPMI supplemented with 10% fetal bovine serum, L-glutamine, penicillin/streptomycin and maintained at 37˚C, 5% CO2.

4.6.3 Cell lines

THP-1 cells (ATCC) were grown at 37˚C, 5% CO2 in RPMI supplemented with 10% fetal bovine serum, L-glutamine, penicillin/streptomycin and 0.01% b-mercaptoethanol.

4.6.4 Luminex

Antigen-specific antibody subclass, isotype, sialic acid, galactose and Fcγ-receptor (FcγR) binding levels were assessed using a 384-well based customized multiplexed Luminex assay, as previously described[196]. Relative antibody concentration was measured against a panel of SARS-CoV-2 antigens (Data S1). SARS-CoV-2 RBD (kindly provided by Aaron Schmidt), SARS-CoV-2 nucleocapsid (N) protein (Aalto Bio Reagents),

and SARS-CoV-2 spike protein (S) (kindly provided by Bing Chen) were used to profile the SARS-CoV-2-specific humoral immune response. Briefly, antigens were coupled by covalent NHS-ester linkages via EDC and NHS (Thermo Scientific) to fluorescent carboxyl- modified microspheres (Luminex). Antigen-coupled microspheres were then washed with an automated plate washer (Tecan) and incubated with plasma samples at an appropriate sample dilution (1:500 for IgG1 and all Fcγ- receptors, and 1:100 for all other readouts). Detection of antigen-specific antibody titers occurred using a PE-coupled detection antibody for each subclass and isotype (IgG, IgG1, IgG2, IgG3, IgG4, IgA1 and IgM, Southern Biotech), and Fc□-receptors were fluorescently labeled with PE before addition to immune complexes (FcγR2A, 2B, 3A, Duke Protein Production facility). For detection of sialic acid and galactose, fluorescein-labeled plant-based lectin detects, SNA and RCA (Vectorlabs) were added as detection reagents at a 1:100 (SNA) and 1:500 dilution (RCA). Plasma samples were acquired via flow cytometry, using an iQue (Intellicyt) and S-Lab robot (PAA). Analysis was done using ForeCyt software by gating on fluorescent bead regions and PE median fluorescent intensity (MFI) is reported as readout for antigen-specific antibody titers.

4.6.5 Functional profiling

For the functional analysis of plasma samples, bead-based assays were used to quantify antibody-dependent cellular phagocytosis (ADCP), antibody-dependent neutrophil phagocytosis (ADNP) and antibody-dependent complement deposition (ADCD), as previously described [160](Data S1). Fluorescent streptavidin beads (Thermo Fisher) were coupled to biotinylated antigen SARS-CoV-2 RBD, N and S and incubated with diluted plasma (ADCP and ADNP 1:100, ADCD 1:10). For ADCP, THP-1 cells were added to the immune complexes and incubated for 16h at 37˚C. For ADNP, primary neutrophils were isolated via negative selection (Stemcell) from whole blood. After 1h incubation at 37□C, neutrophils were stained with an anti-CD66b PacBlue detection antibody (Biolegend). For the ADCD assay, lyophilized guinea pig complement (Cedarlane) was resuspended according to manufacturer's instructions and diluted in gelatin veronal buffer with calcium and magnesium (Boston BioProducts). Post incubation, C3 was detected with Fluorescein-Conjugated Goat IgG Fraction to Guinea Pig Complement C3 (Mpbio).

For detection of antibody-dependent NK cell activity, an ELISA-based approach was used, as described[197]. Briefly, plates were coated with 2 µg/mL of antigen (as mentioned above) and samples were added at a 1:50 dilution and incubated for 2h at 37˚C. NK cells were isolated the day prior via RosetteSep (Stem Cell Technologies) from healthy buffy coats and rested overnight in 1 ng/ml IL-15 (Stemcell). NK cells were incubated with immune complexes for 5h at 37˚C with a staining cocktail containing CD107a PE-Cy5 (BD), Golgi stop (BD) and Brefeldin A (BFA, Sigma Aldrich). Post NK cell incubation, cells were fixed (Perm A, Life Tech) and stained for surface markers with anti-CD16 APC-Cy7 (BD), anti-CD56 PE-Cy7 (BD) and anti-CD3 PacBlue (BD) while fixing. Post permeabilization with Perm B (Life Tech) and anti-MIP-1β PE (BD) antibodies

were used for intracellular staining. All assays were acquired via flow cytometry with an iQue (Intellicyt) and an S-Lab robot (PAA). For ADCP, events were gated on bead-positive cells, whereas neutrophils were defined as CD66b positive followed by gating on bead-positive neutrophils. A phagocytosis score was calculated for ADCP and ADNP as (percentage of bead-positive cells) x (MFI of bead-positive cells) divided by 10000. ADCD was reported as MFI of C3 deposition. NK cells were defined as CD3-, CD16+ and CD56+. Data were reported as percentage of cells positive for CD107a or MIP-1β.

4.6.6 Pseudovirus neutralization antibody assay

The 2019-nCoV pseudoviruses expressing a luciferase reporter gene were generated as described previously (Data S1)[198]. Briefly, the packaging construct psPAX2 (Cat# 11348, AIDS Reagent), luciferase reporter plasmid pLenti-CMV Puro-Luc (Cat# 17447, Addgene) and Spike protein expressing pcDNA3.1-SARS CoV-2.SΔCT were co-transfected into HEK293T cells at ratio of 1:1:0.5 by Calcium phosphate transfection method. The supernatants containing the pseudotype viruses were collected 48 hours post-transfection and filtered by 0.45-μm filter. The viruses were stored at -80°C freezer till use. To determine the neutralization activity of the antisera from vaccinated animals, HEK293T cells were firstly transfected with pcDNA3.1(-)-hACE2 (Cat# 1786, Addgene). 12 hours post transfection; the HEK293T/hACE2 cells were seeded at 96-well tissue culture plate at density of 2.00E+04 cells/well overnight. Heat (56°C, 30 min) inactivated antisera were twofold serial diluted and mixed with 50μl of pseudoviruses. The mixture was incubated at 37°C incubator for 1 hour before adding into HEK293T/hACE2 cells in 96-well plates. Six hours after infection, the cell culture medium was replenished with fresh DMEM (supplemented with 2% FBS). Forty-eight hours after infection, cells were lysed in Steady-Glo Luciferase Assay (Promega). A standard quantity of cell lysate was used in a luciferase assay with luciferase assay reagent (Promega) according to the manufacturer's protocol.

4.6.7 Quantification and statistical analysis

All analyses were performed using python version 3.6.8 with statistical and machine learning packages[199]. Networks were visualized in Cytoscape. Raw data are available in supplementary information.

4.6.8 Classification of convalescent and deceased groups

The classification models were trained to distinguish convalescent and deceased groups with a minimal set of features, to avoid overfitting. PBS controls was subtracted from all features, Fc array features were log transformed, and all data was scaled and centered. Antibody features including sex and interventions (Table 1) were included the selection process, and covariates were binarized and scaled and center prior to analysis.

The models were built using a backward feature elimination for selection and then classified using the minimal set of features which maximize accuracy[165,166]. Models were trained and tested in a fivefold cross-validation framework using random stratified

sampling to ensure that both groups are represented in each group. Within each fold, samples were further subdivided into four sets for each iterative fold-specific elimination. A partial least squares discriminant analysis (PLS) classifier was then trained using the fold-specific selected features to predict the test set. Multiple iterations of fold specific feature selections were performed to obtain a single model. This process was repeated over twenty replicates and convergent correlates were observed[137].

Performance and robustness of the model was contrasted with negative control models constructed from permuted data and randomly selected size-matched features, with multiple iterations of fivefold cross-validation used to generate classification accuracies. These control models were generated 100 times. The permuted control was generated in the same process as above shuffling labels randomly for each repetition. Size-matched features were chosen at random for each cross-validation step within each repetition. Predicted and true outcomes were compared to determine accuracy. Robustness was defined as the effect size of the distributions (Cliff's Δ), and the exact P values of the tail probabilities of the true distributions within the control distributions. Reported are the median p-values across twenty independent cross-validation replicates[107].

4.6.9 Correlation networks

Correlation networks were constructed to visualize the additional humoral immune features that were significantly linked to the selected minimal biomarkers, to provide enhanced insights into the biological mechanisms by which antibodies may provide protection following infection. In brief, antibody features that were significantly correlated with a Holms-Bonferroni correction to the final selected PLS model selected-features were defined as co-correlates. Significant spearman correlations above a threshold of $|r| > 0.5$ were visualized within the networks.

4.6.10 Sensitivity analysis

Using the selected features from the original model a new PLSDA model was trained excluding a single outlier at a time in a fivefold cross validation framework. This process was repeated three times, each time generating a unique ROC curve as the top 3 individual outliers were removed. Using these cross validated ROC curves the mean performance and variation were assessed and are summarized as area under curve.

4.6.11 Ratio based analyses

In order to evaluate S vs N ratios, first ratios for each feature were defined separately by simply dividing S-responses over N-responses for every given feature. S:N ratios were visualized by log2 transformation for ease of interpretation. Differences across convalescents and deceased were then tested with a one-sided Mann-Whitney U test and a Holm-Bonferroni multiple hypothesis correction criterion. In order to address whether the overall S-response was enriched over N-responses in the convalescents

across all features tested, all data was background corrected and z-score normalized. Then the number of S-features which were greater than their N-counterparts across every feature were summed. This analysis yielded a distribution of individual S greater than N scores for each group and statistical differences were assessed using a one-sided Mann-Whitney U test.

4.7 Supplementary figures

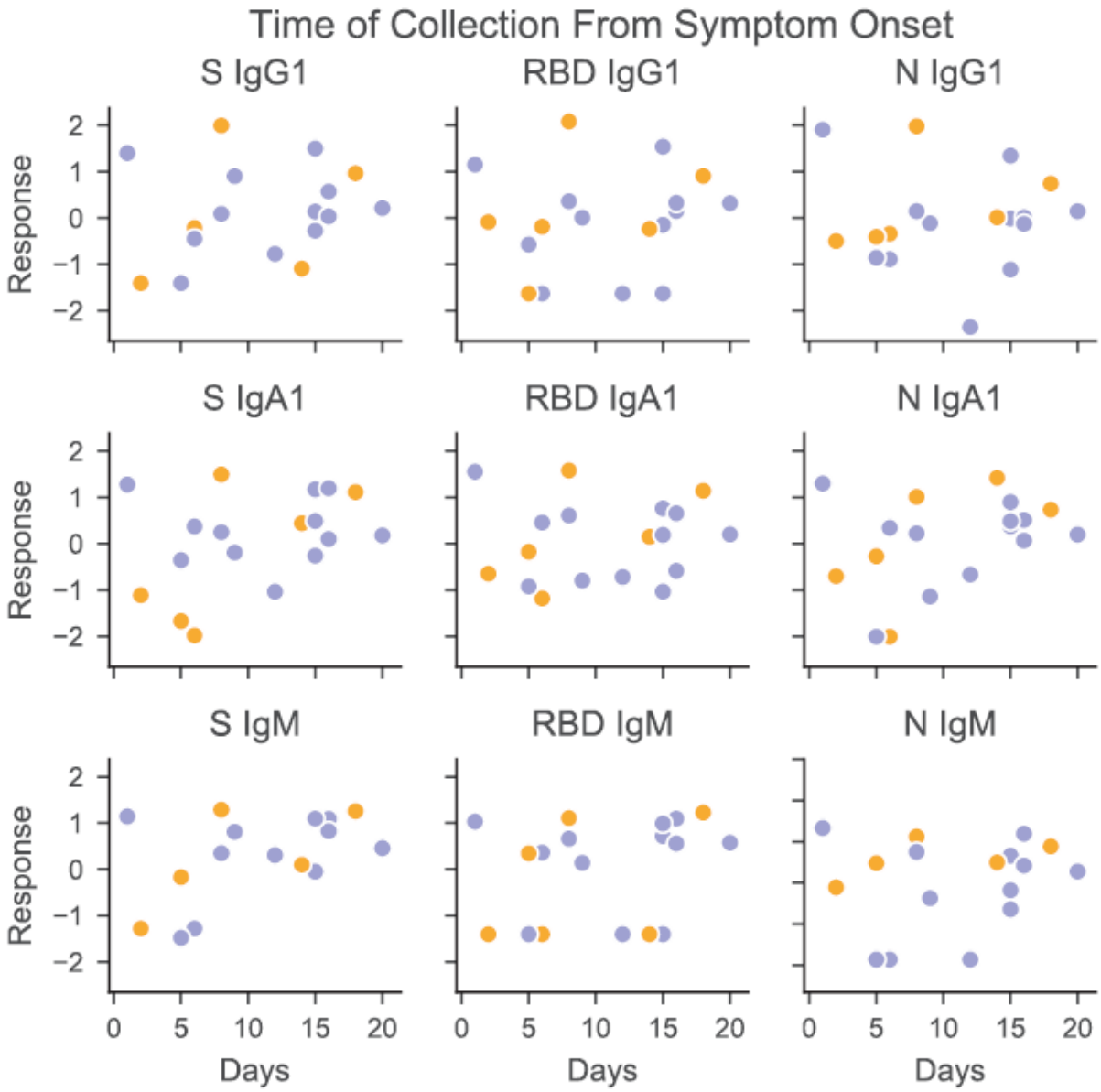


Figure 4-S1. Individual antibody dynamics over time, related to Figure 4-1

The dot plots show the relationship of antibody titers against IgG1, IgA1, and IgM and the estimated time of collection from onset of symptoms across SARS-CoV-2 antigens.

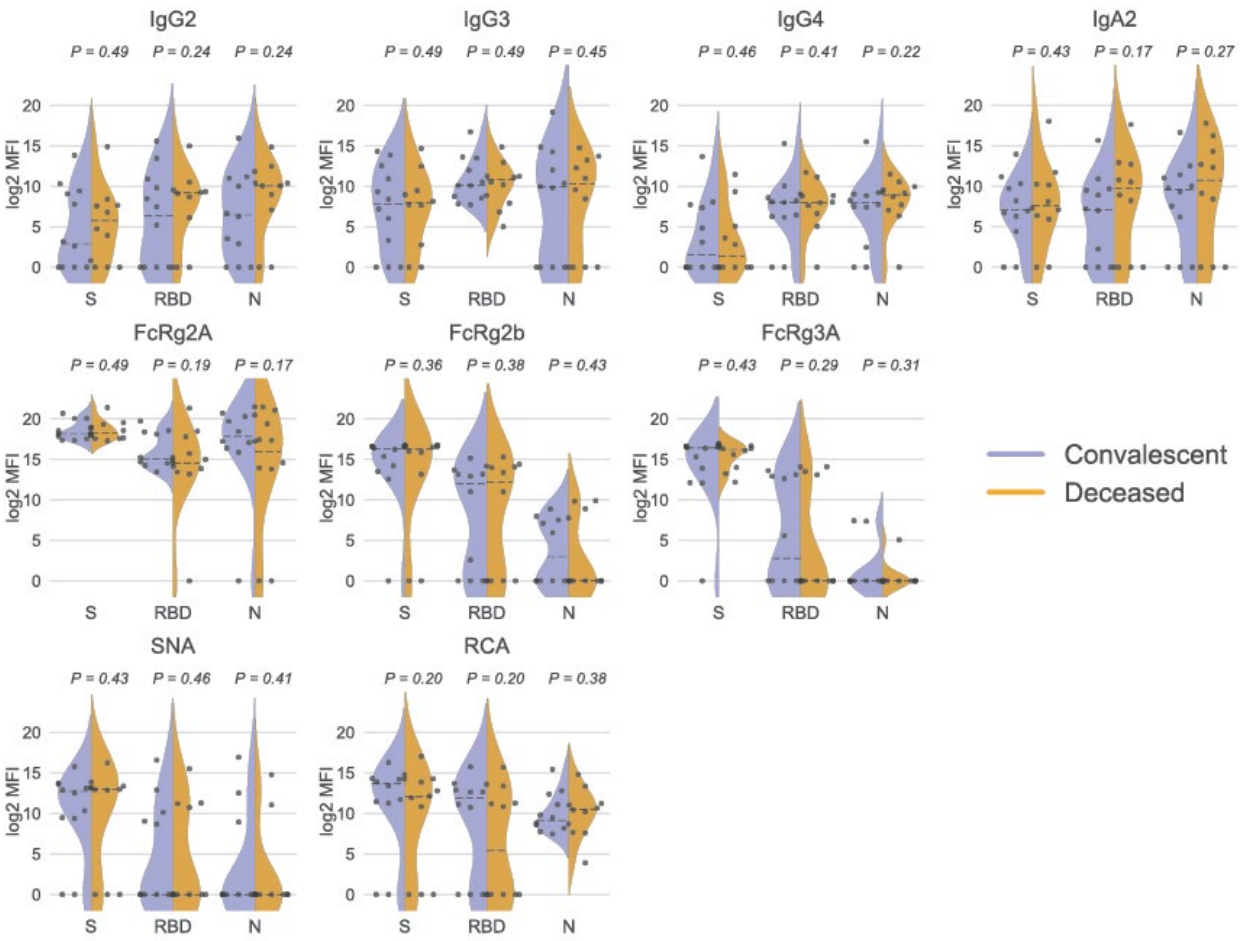


Figure 4-S2 Individual antibody distributions, related to Figure 4-1

The split violin plots show the distribution of subclass, isotype, and Fc-receptor binding profiles across convalescent (purple) and deceased (orange) individuals. A Mann-Whitney U test was used to calculate p values. No significance was detected after a Holm-Bonferroni correction for multiple hypothesis testing.

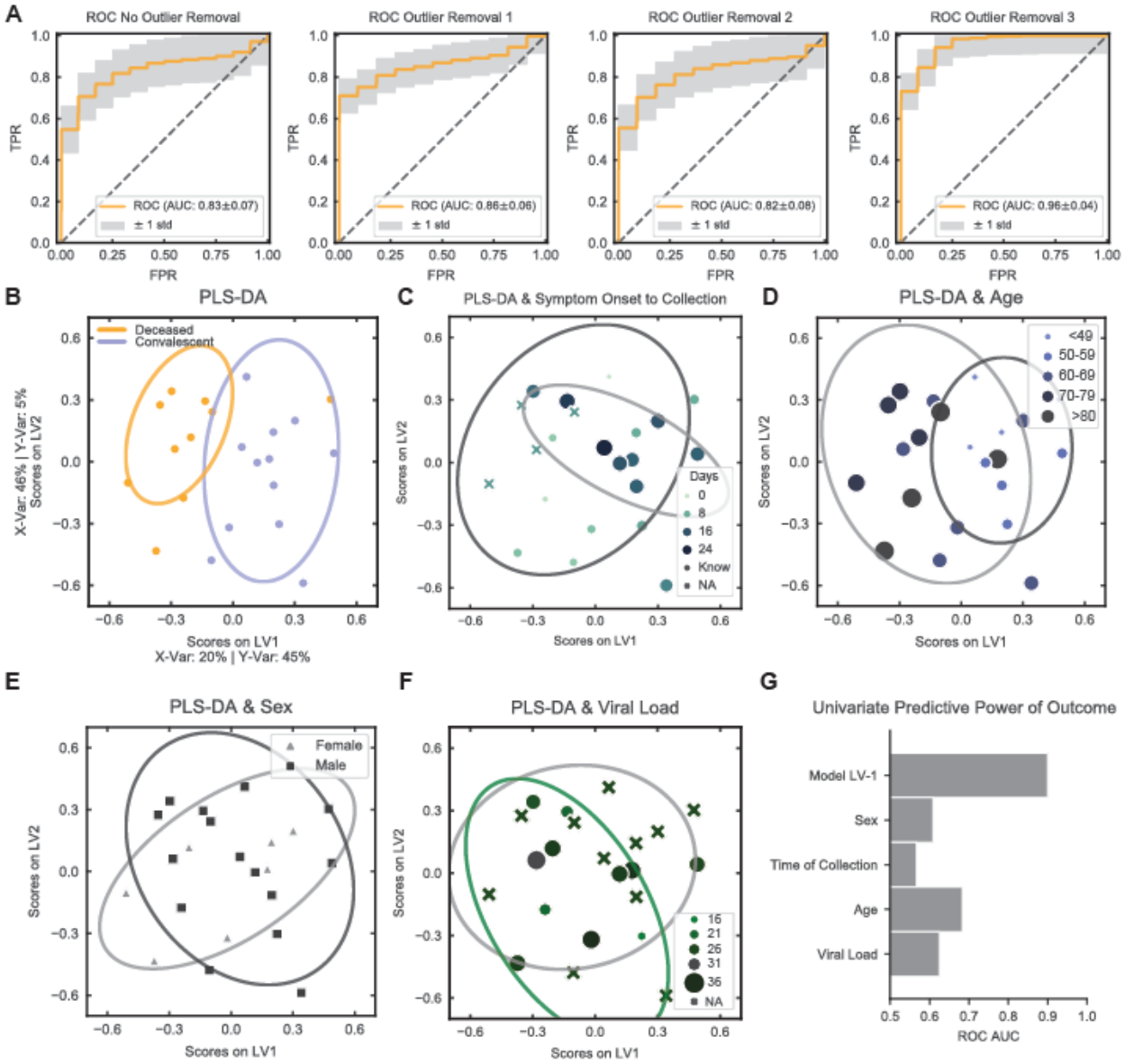


Figure 4-S3. Sensitivity analysis, model performance, and potential influence of confounders, related to Figure 3.

(A) Using the selected features from the original model, outliers were removed one at a time and ROC curves based on scores generated multiple times in a 5-fold cross validation framework were generated. In orange is the average performance and the grey shading represents one standard deviation. The gray dashed line represents the threshold of a random process. The AUC of the ROC curves and uncertainties are listed at the bottom of each graph. (FPR: false positive rate, TPR: true positive rate). (B) The original PLSDA scores plot is plotted based on the final antibody feature selection. (C) The same scores plot was recolored to highlight the distribution of the individuals based on days of symptom onset. The size and intensity of the dots from bright green to dark blue correspond to shorter- to-longer times, and x markers represent samples for which information was not available. (D) The same original scores plot was colored to reflect differences in ages across the subjects, with increased dots size and color (light blue to dark

gray) showing increasing age. (E) The original scores plot was recolored to show the influence of sex on distributions, with males shown as dark squares and females as light gray triangles. (F) The original scores plot was recolored to show viral loads, with larger dots and increasing darkness (green to black) highlighting higher viral loads. (G) The bar graph shows the predictive power (as ROC AUC) of the original model latent variable 1 (LV1) compared to the univariate predictive power of individual demographic confounders, including sex, time to collection, age, and viral loads.

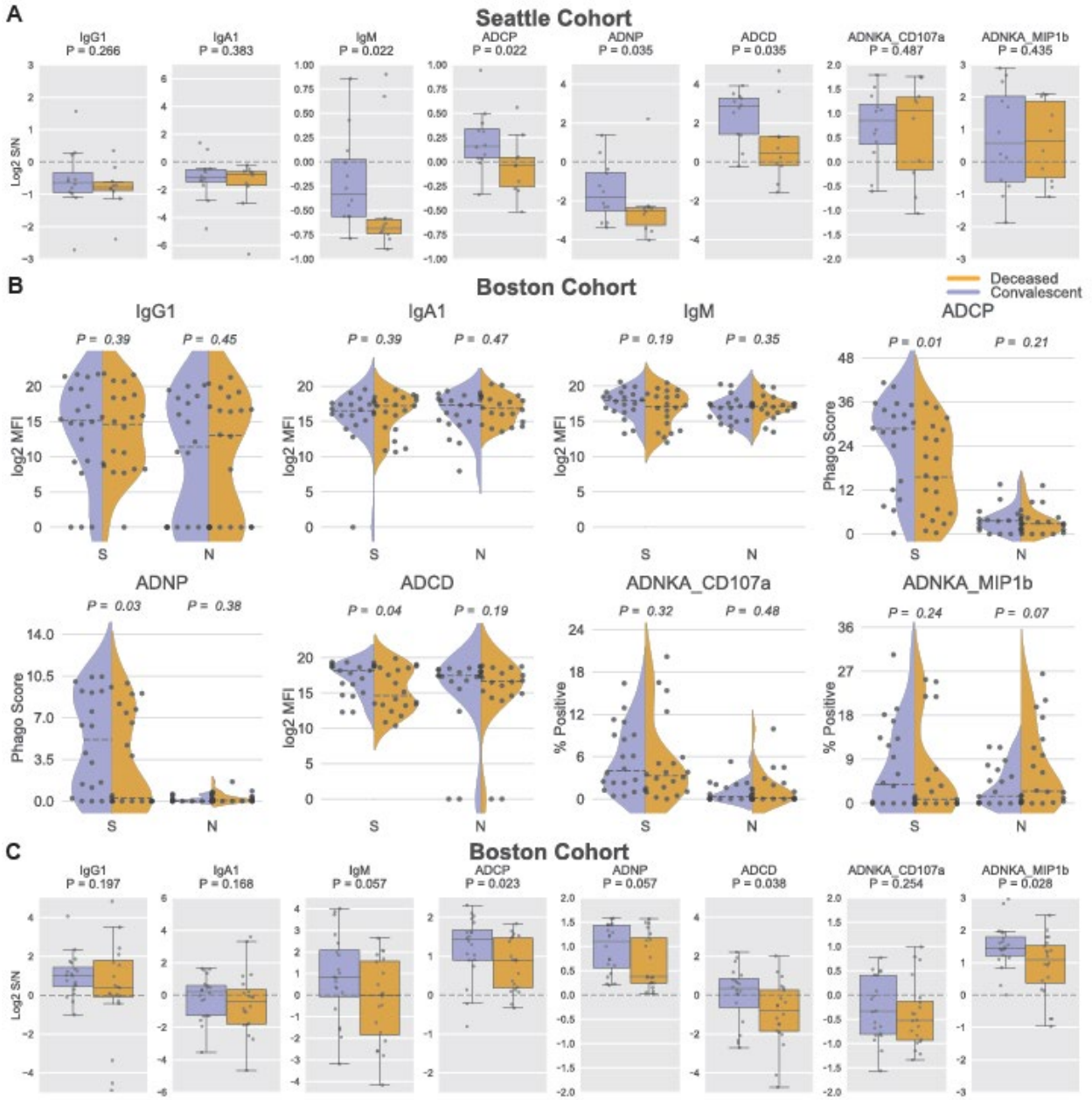


Figure 4-S4. Univariate Spike:Nucleocapsid ratios across the Seattle/Discovery and Boston/Validation cohorts, related to Figure 4.

The whisker plots show the log2 ratio of Spike to Nucleocapsid (S:N) in the (A) Seattle/Discovery cohort for IgG1, IgA1, IgM, antibody dependent cellular phagocytosis (ADCP), antibody

dependent neutrophil phagocytosis (ADNP), antibody dependent complement deposition (ADCD), and antibody dependent NK cell activation (ADNKA) by degranulation (CD107a) and chemokine secretion (MIP1b) (from left to right). (B) The violin plots highlight the distributions of antibody features captured in the Boston/Validation cohort of IgG1, IgA1, IgM titers, ADCD from left to right on the top line; as well as ADCP, ADNP, and ADNKA by degranulation CD107a or MIP1b, from left to right on the bottom row. Responses are shown for Spike (S) and Nucleocapsid (N) for each read out for the convalescents (blue) and the deceased (yellow). The dashed line represents the median of the distribution. (C) The whisker plots show the log2 ratio of Spike to Nucleocapsid (S:N) in the Boston/Validation cohort with the same features as in (A). For both A and C the hatched line represents the point where ratios are enriched in S over N and differences were tested with a one-sided Mann-Whitney U test.

4.8 Acknowledgements

We thank Bruce Walker, Nancy Zimmerman, Mark and Lisa Schwartz, an Anonymous donor, and Terry and Susan Ragon for their support. We would also like to thank Bing Chen for protein production efforts. We acknowledge support from the Ragon Institute of MGH, MIT, the Massachusetts Consortium on Pathogen Readiness (MassCPR), and the Bill & Melinda Gates Foundation (235730), the NIAID (U19 AI35995) and the U.S. Centers for Disease Control and Prevention (CK000490).

Chapter 5

Compromised humoral functional evolution tracks with SARS-CoV-2 mortality

The contents of this chapter were published as:

Tomer Zohar*, Carolin Loos*, Stephanie Fischinger*, Caroline Atyeo*, Chuangqi Wang, Matthew D. Slein, John Burke, Jingyou Yu, Jared Feldman, Blake Marie Hauser, and Tim Caradonna, Aaron Schmidt, Yongfei Cai, Hendrik Streeck, Edward T. Ryan, Dan Barouch, Richelle C. Charles, Douglas Lauffenburger, and Galit Alter. “Compromised humoral functional evolution tracks with SARS-CoV-2 mortality”. *Cell* 183, 1508-1519. e12 (2020)

**These authors contributed equally to this work.*

See the online publication for any references to supplementary material not included in this thesis.

Contributions

Conceptualization, T.Z., C.L., S.F., C.A., C.W., D.L., and G.A.; Methodology, T.Z., C.L., S.F., C.A., C.W., D.L., and G.A.; Formal Analysis, T.Z., C.L., C.W., and D.L.; Investigation, S.F., C.A., M.D.S., J.B., J.Y., J.F., B.M.H., A.S., T.C., and Y.C.; Recourses, R.C.C.; Data Curation, C.L. and R.C.C.; Writing – Original Draft, T.Z., C.L., S.F., C.A., and G.A.; Writing – Review & Editing, T.Z., C.L., S.F., C.A., D.L., E.T.R, R.C.C., and G.A; Visualization, T.Z., C.L., and C.W.; Supervision, D.L. and G.A.

5.1 Abstract

The urgent need for an effective SARS-CoV-2 vaccine has forced development to progress in the absence of well-defined correlates of immunity. While neutralization has been linked to protection against other pathogens, whether neutralization alone will be sufficient to drive protection against SARS-CoV-2 in the broader population remains unclear. Therefore, to fully define protective humoral immunity we dissected the early evolution of the humoral response in 193 hospitalized individuals ranging from moderate-to severe. Although robust IgM and IgA responses evolved in both survivors and non-survivors with severe disease, non-survivors showed attenuated IgG responses, accompanied by compromised Fcɣ-receptor binding and Fc-effector activity, pointing to deficient humoral development rather than disease-enhancing humoral immunity. In contrast, individuals with moderate disease exhibited delayed responses that ultimately matured. These data highlight distinct humoral trajectories associated with resolution of SARS-CoV-2 infection and the need for early functional humoral immunity.

5.2 Introduction

The majority of individuals infected with SARS-CoV-2 develop mild symptoms, however a small but significant proportion of the population develop more severe disease, a fraction of which pass away[200]. The rapid spread of infection and unpredictable evolution of disease severity in some infected individuals has overwhelmed hospitals. While a number of comorbidities have been linked to mortality[201,202], emerging data suggest that inflammatory markers and cellular activation track with severity of disease[203,204] pointing to a generalized activation of the immune response with progressive infection. Moreover, antibody levels and neutralizing antibody activity increase with the rapid expansion of plasmablast populations in the setting of more severe disease [203], raising the possibility that the humoral immune response may contribute to pathology rather than protection[205]. However, recent data from vaccine studies point to protective effects of antibodies against SARS-CoV-2 infection and disease[193]. It is still unclear whether qualitatively distinct antibodies evolve in individuals who survive severe infection or whether antibody functions track with differential disease severity. Understanding how antibody functions, beyond neutralization, evolve during infection and contribute to

recovery rather than pathology may provide key insights for vaccine and therapeutic design to avoid detrimental enhancement and provide highly effective humoral defense.

The acute humoral immune response to SARS-CoV-2 is marked by the rapid evolution of multi-isotype specific humoral immunity[206], likely involved in the highly compartmentalized antiviral response within the respiratory tract. However, whether specific antibody functional profiles explain the variation seen in viral control remains unclear. Thus, here we comprehensively profiled the humoral immune response to SARS-CoV-2 over the first three weeks following symptom onset in depth and used complementary modelling approaches to define whether distinct humoral immune responses evolve among individuals with different degrees of disease severity. A total of 193 hospitalized individuals were included in the study, stratified by disease severity and outcome into three groups: inpatients with moderate infection that recovered, inpatients with severe infection that recovered, and inpatients whose underlying cause of death was COVID-19. Humoral maturation was observed across all three groups, with more rapid and robust evolution in individuals that survived severe infection, with an early and vigorous functional response to S2, just days after symptom onset. Despite comparable evolution of IgM and IgA responses among survivors and non-survivors with severe disease, non-survivors failed to fully deploy a highly functional IgG response able to coordinate Fc-receptor binding and elicit innate immune effector function. While neutralizing antibody activity did not differ across the groups, no signatures of antibody enhancement were noted among individuals with severe infection. Conversely, delayed evolution of functional humoral immunity was also observed in individuals with moderate disease, albeit the IgG response continued to evolve overtime. Thus, the data point to the potential importance of early and continued evolution of IgG-recruiting Fc-effector function in antiviral control and clearance beyond host defense.

5.3 Results

5.3.1 Distinct weekly evolution of antibody features

The humoral immune response to SARS-CoV-2 evolves rapidly following infection in individuals with symptomatic infection, evolving weeks after symptom onset[206]. Significant heterogeneity in the magnitude of the humoral immune response has been observed across individuals who have survived infection[207]. While neutralizing antibodies have been linked to protection following vaccination in non-human primates (NHP)[118,208], antibodies may contribute to antiviral control and clearance through various additional mechanisms, via their capacity to interact and leverage the anti-pathogen functions of the innate immune system[209,210]. Preliminary data point to early differences in the immunodominance of the humoral immune response among individuals with different clinical outcomes[211]. However, whether distinct antibody profiles develop over time in individuals with different clinical trajectories remains unclear but could point to immune mechanisms of convalescence. Thus, here we profiled the humoral immune response across three hospitalized cohorts: a group of 82 individuals who did not require

admission to the intensive care unit (ICU) and were eventually discharged with moderate disease, a group of 76 individuals who required ICU care but survived severe disease, and group of 35 individuals with severe disease that died due to COVID-19 regardless of type of care (Figure 5-1A and Table 5-S1). Individuals were sampled 1-8 times during the first month of infection. These data collectively provided a population level temporal landscape that could be used to define differences in the trajectories of antibody features across disease outcomes (Data 5-S1).

A week after onset of symptoms, similar SARS-CoV-2 antigen specific IgM and IgA titers were observed in all three groups with slightly higher IgA1 S-specific titers in severe disease survivors. (Figure 5-1B and Figure 5-S1A). Similarly, no significant difference was observed in IgG3, the first highly functional IgG subclass selected during acute immune responses [212]. However, the anti-spike (S) IgG1 titers differed significantly between severe and moderate disease and were lower in individuals that died from severe infection. By the second week, significantly higher titers of nearly all the S-specific antibody isotypes and subclasses (Figure 5-1B), and higher titers across all antigens (Figure 5-S1A), were observed in survivors of severe disease compared to those with moderate disease and those who died. Slower S-specific IgG1 development was noted in both those who died and in individuals with moderate disease. By the third week, IgA and IgM were nearly equivalent across all groups (Figure 5-1B and Figure 5-S1A). However, individuals with severe and moderate disease generated higher IgG subclass titers compared to the deceased group by the third week, pointing to a convergence of IgG immunity across survivors but a delayed and incomplete evolution among non-survivors. Thus, despite the early robust evolution of IgA and IgM responses among non-survivors, these data argue for a potential selective defect in IgG development associated with COVID-19 mortality.

5.3.2 Compromised Fc-receptor binding and effector function tracks with COVID-19 mortality

Given the differences in IgG class-switching, we next examined the functional consequences of this defective class-switched response. Specifically, the ability of SARS-CoV-2-specific antibodies to bind to the low-affinity IgG-Fcɣ-receptors (FcɣRs) and the IgA-Fc-α-receptor (FcαR), critical for deploying antibody effector function[210], was assessed. Binding was assessed to the FcαR and the low affinity IgG FcɣR, the activating FcɣR2A and FcɣR3A receptors, the sole human inhibitory receptor FcɣR2B, and the GPI anchored FcɣR3B receptor[210,213]. FcɣR binding profiles mirrored changes in IgG1 and IgG3 titers, with early development of S-specific FcɣR2B and S1- and S2-specific FcɣR2B and FcɣR3B binding antibodies in individuals with severe SARS-CoV-2 infection that survived (Figure 5-1C and Figure 5-S1A). These differences were amplified over time, with FcɣR3A and FcαR binding antibodies reaching relatively similar levels across groups, but overall lower Fc-receptor (FcR) binding antibodies in individuals with moderate disease and non-survivors.

Differences in FcR binding results in changes in signaling cascades and distinct antibody-mediated innate immune effector functions[210]. Therefore, we next probed the consequences of the changes in antibody titer and FcR binding on innate immune effector functions and neutralization (Figure 5-1D). Low but detectable antibody effector functions were observed in all three groups one week following symptoms, with rapid initial development of functional humoral immunity among survivors with severe infection. Notably, by the second week of infection, S-specific antibody dependent neutrophils phagocytosis (ADNP), complement fixation (ADCD), and neutralization were similar among the severe survivors and non-survivors, but were higher than the levels observed in individuals with moderate disease (Figure 5-1D). Conversely, monocyte phagocytosis (ADCP), NK cell activating (ADNKA) S-specific antibodies, measured by MIP-1b secretion, and RBD specific functions, were significantly lower in non-survivors compared to survivors of severe infection but were similar in non-survivors and survivors with moderate disease. In particular, RBD-specific monocyte phagocytic antibodies and complement fixing antibodies remained lower in non-survivors through week three (Figure 5-S1A). Changes in Fc-glycosylation directly influence FcγR binding and effector function[214], Interestingly, by the second week following symptom onset, broad FcγR binding associations were observed in survivors of severe disease (Figure S1B) that were lost in individuals who ultimately succumbed to infection, pointing to potential shifts in Fc-glycosylation across groups, in the setting of continually evolving IgG titers. Thus, two potential temporal shifts in Fc glycosylation may exist during SARS-CoV-2 infection, the first poised to recruit and clear the virus, and a second for clearing infected cells. These data collectively point to a disconnect in the evolution of antibody titer and function, with a slower and milder evolution of antibody effector functions among individuals with moderate infection, a rapid and highly functional humoral immune evolution in individuals with severe disease, and a defect in the evolution of particular IgG functions in individuals that ultimately pass away.

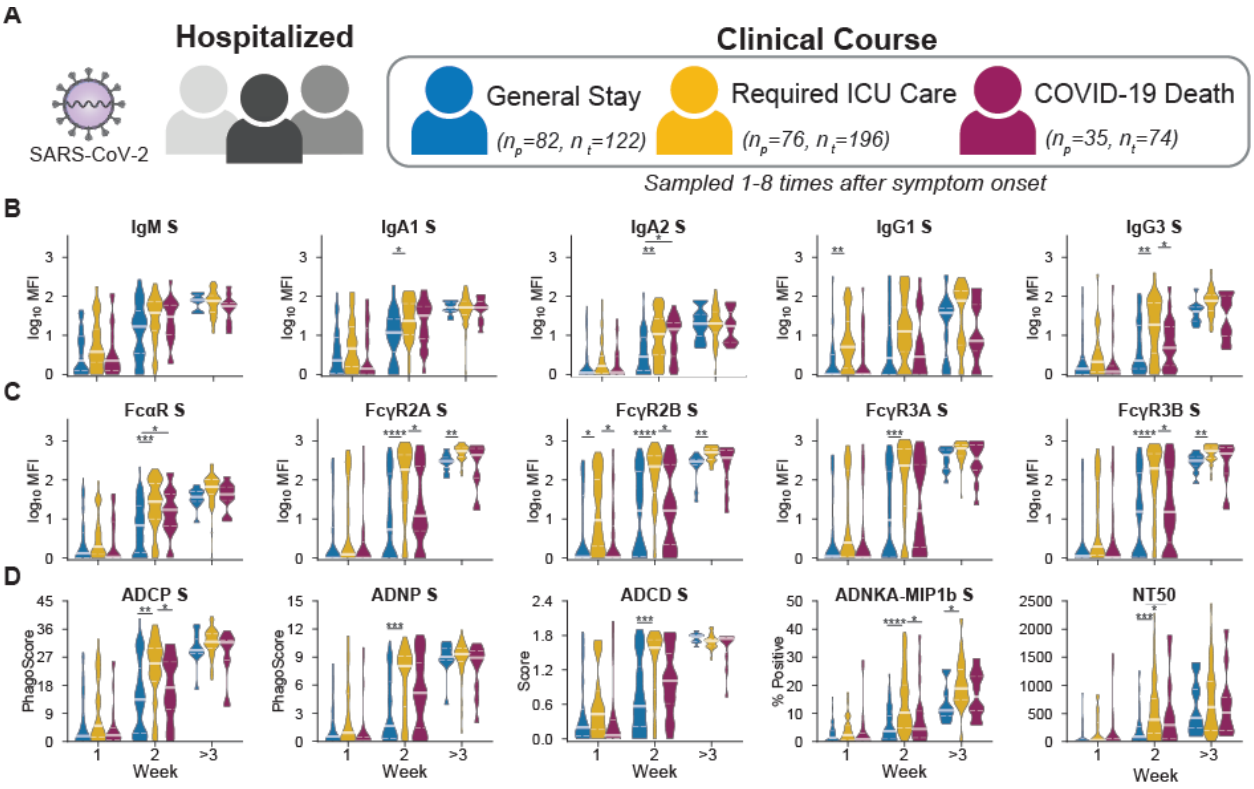


Figure 5-1. Weekly evolution of SARS-CoV-2 specific humoral immune responses following symptom onset across different clinical courses

193 plasma samples from hospitalized SARS-CoV-2 infected individuals were profiled against the SARS-CoV-2 spike antigen (S). A: Of the patients 82 were not admitted to the ICU and were eventually discharged (moderate), 76 required ICU care but did not succumb to infection (severe), and 35 died of COVID-19 (deceased). Patients were sampled from 1-8 times during their hospital stay. np denotes the number of patients in a group, and nt denotes the total number of samples collected across all individuals. Distributions of titers (B), Fc-receptors (C), and functions (D-G) across moderate (blue), severe (yellow), and deceased (red) over the course of 0-7, 8-14, and greater than 14 days against S. The solid white line represents the median and the dotted lines the first and third quartiles. A Kruskal-Wallis test was used to evaluate statistical differences across groups for all intervals and features and was corrected for multiple hypothesis testing using the Benjamini-Hochberg procedure. If statistically significant then a two-sided Mann-Whitney U test was performed for post-hoc comparisons. Significance corresponds to the Mann-Whitney U test p-values (*: p < 5e-2, **: p < 5e-3, ***: p < 5e-4, ****: p < 5e-5, *****: p < 5e-6). Antibody dependent cellular phagocytosis (ADCP), antibody dependent neutrophil phagocytosis (ADNP), antibody dependent complement deposition (ADCD), antibody dependent Natural killer cell activation (ADNKA). See also Figure S1, Table S1, and Data S1.

5.3.3 Developmental differences in humoral architecture

To further investigate whether specific types of antibody properties or antigen specificities differed across individuals that survived or did not survive severe SARS-CoV-2 infection, polar plots were generated to examine potential patterns in humoral features at one, two-, and three-weeks post symptom onset (Figure 5-2). As early as the first week post symptoms, enhanced humoral immunity was observed in severely ill

patients that survived infection (Figure 5-2). While responses holistically increased in both groups, the rise was more uniform and robust among survivors compared to non-survivors. Although there were similar IgA and IgM responses across the two groups, there was an overall trend towards lower IgG3, Fc-receptor binding, and Fc functions in the non-survivors which were observable by the second week post symptoms but were amplified by week three, similar to what was observed in the univariate analysis (Figure 5-1). In particular, survivors evolved overall higher phagocytic responses. To probe the global differences in the humoral immune response across the groups, non-parametric combinations of univariate differences for each feature class further revealed significantly higher overall development of Fc-receptor binding as early as week one and Fc-functional responses by week three in the severely ill recovered individuals compared to those who died (Figure 5-2). Ultimately, the majority of antibody features were stunted in those who died with no evidence of disease enhancing humoral responses linked to mortality. These data highlight more significant qualitative defects in the humoral immune response that track with COVID-19 mortality.

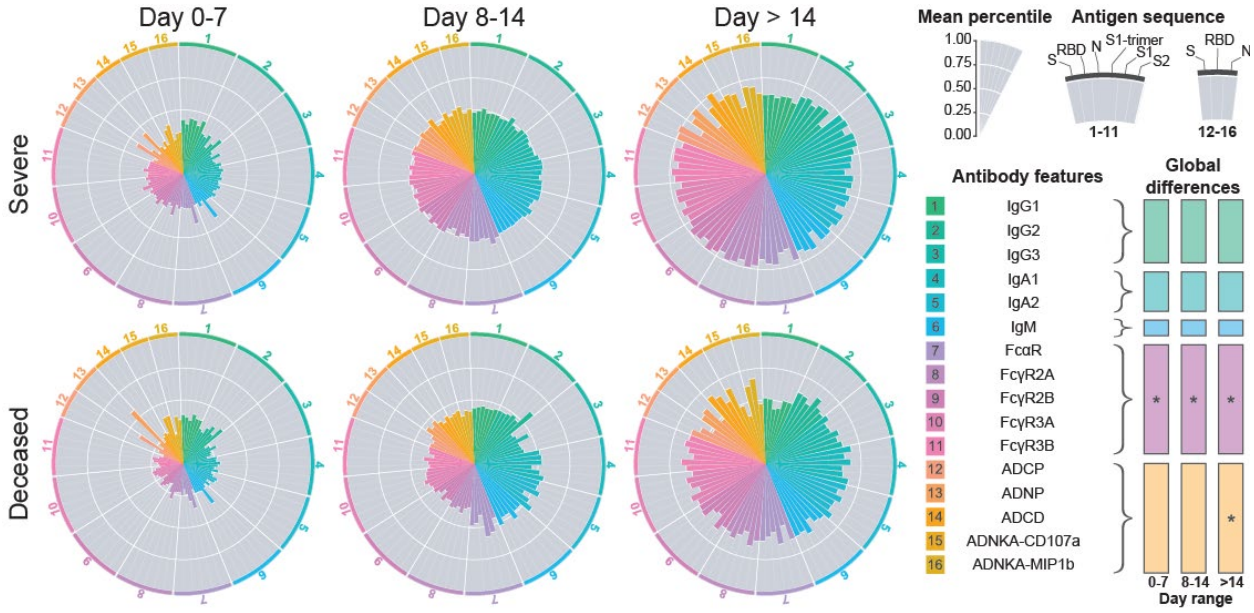


Figure 5-2. Weekly evolution of humoral architecture

The polar plots depict the mean percentile of each antibody feature at each interval across the severe (top) and the deceased (bottom) groups. The major slices 1-6 cover antigen specific isotypes/subclasses, 7-11 antigen-specific antibody Fc-receptor binding, and 12-16 antigen-specific antibody mediated functions. For segments 1-11 antigen specificities repeat in the following order: S, RBD, N, S1 trimer, S1, and S2. For segments 12-16 antigen specificities are repeated S, RBD, and N. The size of the wedge depicts the mean percentile ranging from 0-1. On the right, non-parametric combination global p-values are shown, composed of Mann Whitney U test p-values for partial tests within each feature type, and using the Fisher method for combination (*: $p < 0.05$).

5.3.4 Longitudinal variation across clinical groups

Given the significant differences across the groups over time, we next aimed to gain enhanced temporal granularity related to the nature of the humoral defect in humoral immunity. Whereas the Uniform Manifold Approximation and Projection (UMAP), which represents high-dimensional data reduced to two-dimensional space, showed limited SARS-CoV-2 humoral immune variation by age and sex (Figure 5-3A-B, Figure 5-S2), but variation was observed with respect to time following symptom onset (Figure 5-3C). Temporal changes in individual humoral features across the groups highlighted distinct humoral trajectories across the patient groups and across antigens and isotypes (Figure 5-3D). When analyzed by days following symptom onset, rather than weekly intervals (Figure 5-1 and Figure 5-2), delays in the rise of IgG, IgA, FcγR2A, and antibody functions in non-survivors emerged (Figure 5-3D). All three isotypes appeared to decline more rapidly among the deceased compared to the survivors. Furthermore, no difference was noted in neutralizing antibody evolution between these groups (Figure 5-3D). Thus, temporal analysis revealed both delayed and incomplete evolution of the humoral immune response among non-survivors of SARS-CoV-2 infection.

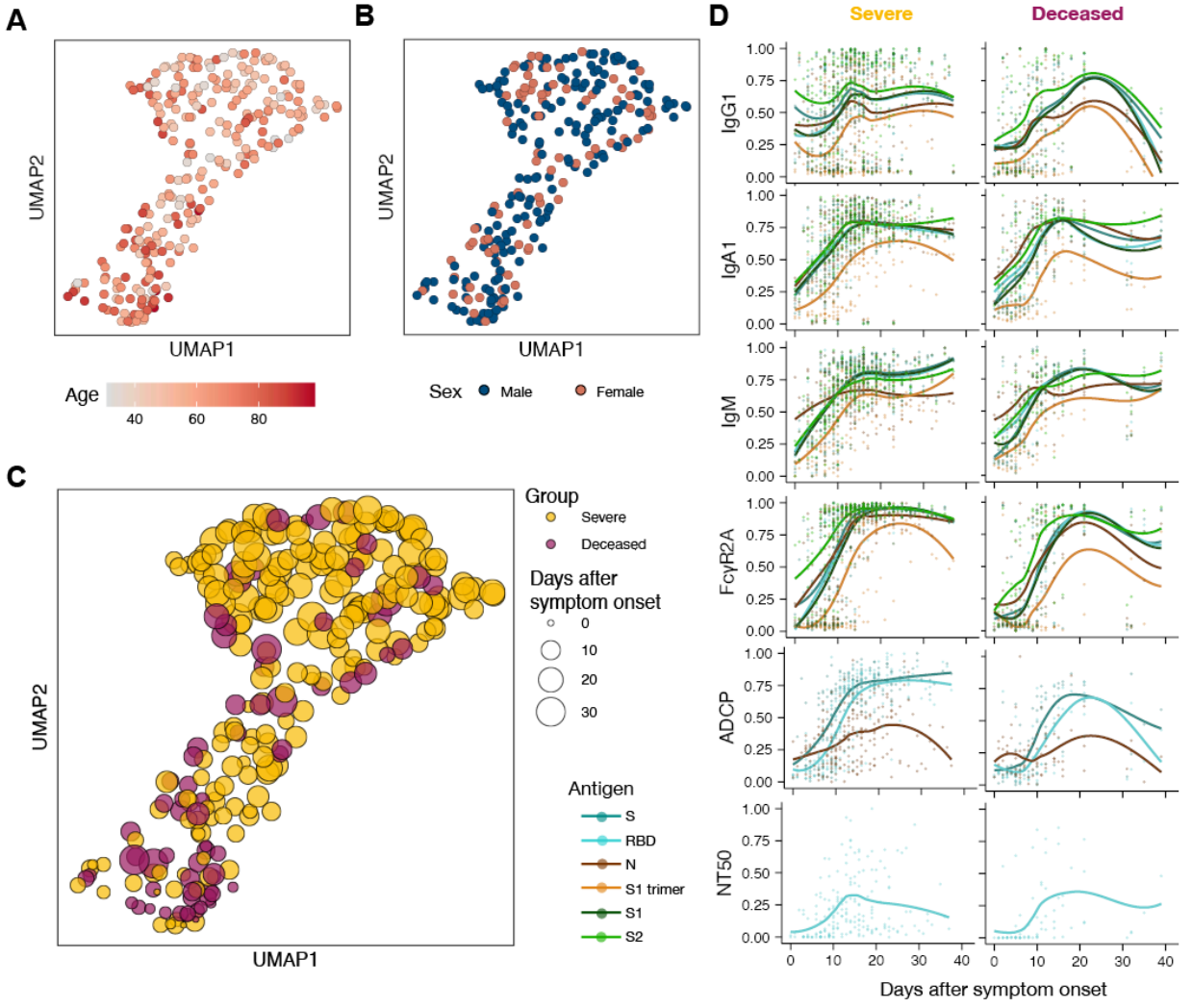


Figure 5-3. Temporal evolution of SARS-CoV-2-specific antibody features

(A-C) Uniform Manifold Approximation and Projection (UMAP) was used to visualize the multivariate data in two dimensions. Each point represents a given individual at a single timepoint and colors indicate age (A), sex (B), and group (C). (D) Normalized antibody levels are shown over time, plotted by days after symptom onset, for the severe and deceased group. Each dot is an individual measurement, the lines show smoothed non-parametric regression models (loess), and the color indicates the antigen-specificity. See also Data S1, and Figure S2.

5.3.5 Dissecting specific temporal differences in early humoral dynamics

To determine the antibody features that differed most across survivors and non-survivors, Akaike Information Criterion (AIC) was used to identify the individual humoral characteristics that showed the greatest variation between the severe and deceased groups, based on quantitative kinetics occurring over the first 17 days following symptom onset (Figure 5-4A). Among the top 5 features, S2-, S- and S1 trimer-specific FcɣR binding were highly divergent across the groups, followed by additional RBD- and S1-specific FcɣR binding and S-specific complement depositing antibody activity (ADCD) (Figure 5-4A). Antibody titers showed less pronounced differences between the groups, highlighting stronger differences in quality, rather than quantity, of antibody evolution during the first 17 days following symptom onset.

To gain a deeper sense of how the features differed over time, curves were fitted for each feature and evaluated based on 4 parameters: 1) “a”=initial levels, 2) “b”=initial seroconversion speed, 3) “c”=seroconversion time, and 4) “d”= endpoint levels (Figure 5-4B). Analyzing the features that were the most divergent across the groups (Figure 5-4A), S2-specific FcɣR3B and S-specific FcɣR2B binding levels differed not only at the time of symptom onset (parameter a), but also seroconverted more rapidly (c) over the first few days following symptom onset in the individuals who survived severe SARS-CoV-2 infection (Figure 5-4C and Figure 5-S3). Conversely, S2-specific FcɣR2B and FcRɣ2A binding were initially higher in survivors compared to non-survivors but reached similar levels in both groups. Further FcR binding antibodies and S-specific ADCD all showed a similar difference in time to seroconversion across the two groups, highlighting the delayed kinetics of this evolution in individuals that did not survive infection. Thus, these data highlight the different temporal changes across the antibody features, pointing to distinct functional consequences in antiviral immunity following infection.

In order to understand generalizable differences in the temporal evolution of the humoral immune response, a composite visual was constructed that summarized kinetic differences in each parameter (a, b, c, d) across each feature and the two group. Early elevated broad IgG1 levels, S1- and S2-specific IgG3, S-specific FcɣR2B, S- and S2-specific FcɣR3A, S2-specific FcɣR3B were noted, with a notable immunodominance of S2-specific immunity among survivors at the time of symptom onset (parameter a) (Figure 5-4D). A consistent, but more abrupt initial conversion speed (parameter b) was observed in the individuals that ultimately passed away across multiple subclasses,

isotypes, FcR binding profiles, and functions, potentially related to their lower early levels. Non-survivors also converted later (parameter c) than survivors across nearly all FcR binding antibodies, with a delay in RBD- and S2-specific FcγR2B and FcγR3B binding antibodies. Final overall magnitudes (parameter d) pointed towards higher levels among survivors. Importantly, no single feature was enhanced early or later in individuals who ultimately passed away, further underscoring that no antibody feature pointed to evidence of disease enhancement in this population.

As mentioned above, in comparison to other targets on S, S2-specific responses were already expanded days after symptom onset in severe survivors (Figure 5-4C-D and Figure 5-S3). Given the emerging appreciation for the more conserved nature of S2 across coronaviruses[215], the early rise in S2-specific FcR binding antibodies may reflect an early evolution of cross-reactive immunity that may be key to disease control. Conversely, no differences were observed in common-coronavirus RBD-specific humoral immune responses at early timepoints across the groups, suggesting that the ability to evolve S2-specific cross-reactive immunity, rather than the level of pre-existing immunity to less cross-reactive RBDs, associated with neutralization (Amanat et al., 2020), may play a more critical role in disease recovery (Figure 5-S4). These data point to both higher initial and overall levels of IgG and FcR binding antibodies among survivors, especially against the S2 domain. In contrast, non-survivors showed lower initial responses that attempted to converge but largely failed to do so.

Finally, to determine the individual antibody features that differed most across the two groups, data were integrated, and an enrichment score was calculated for each antibody Fc-readout (Figure 5-4E), each antigen-specificity (Figure 5-4F), or groups of Fc-features (Figure 5-4G), to define the humoral changes that were most elevated within one group or another. While limited differences were noted in IgG2, IgA, and IgM responses across the two groups, IgG1, IgG3, FcR binding and functional responses differed most across survivors and non-survivors. These differences were observed similarly across all tested SARS-CoV-2 antigens (Figure 5-4F). Moreover, when all feature “types” were collapsed, no enrichment was observed for titers, but FcR binding, and Fc-effector functions were able to resolve individuals across clinical trajectories (Figure 5-4G). These data highlight that cross-antigen differences in antibody effector function, rather than titer, are most divergent between survivors and non-survivors of SARS-CoV-2 infection.

To illustrate whether survivors with severe disease and those who died could be distinguished within the first week following symptom onset, a random forest selection model was constructed. The model recursively chose a minimal set of features that best distinguished the two groups in a cross-validation framework, resulting in the generation of a model able to robustly classify individuals. The model was able to classify survivors or non-survivors with 72% accuracy (Figure 5-4H-I). Many of the top features selected by the model were higher in survivors, including S-specific functions, FcRs, and IgG3. One feature, N-specific FcαR, was higher in non-survivors, in line with previous

observations related to early immunodominance shifts between S- and N- across individuals that ultimately survive or pass away[211]. Thus, early cross-antigen specific antibodies able to drive rapid control and clearance of the virus represent early biomarkers that resolve disease trajectory and provide insights into humoral functions, and dysfunctions, that may be key to early antiviral containment.

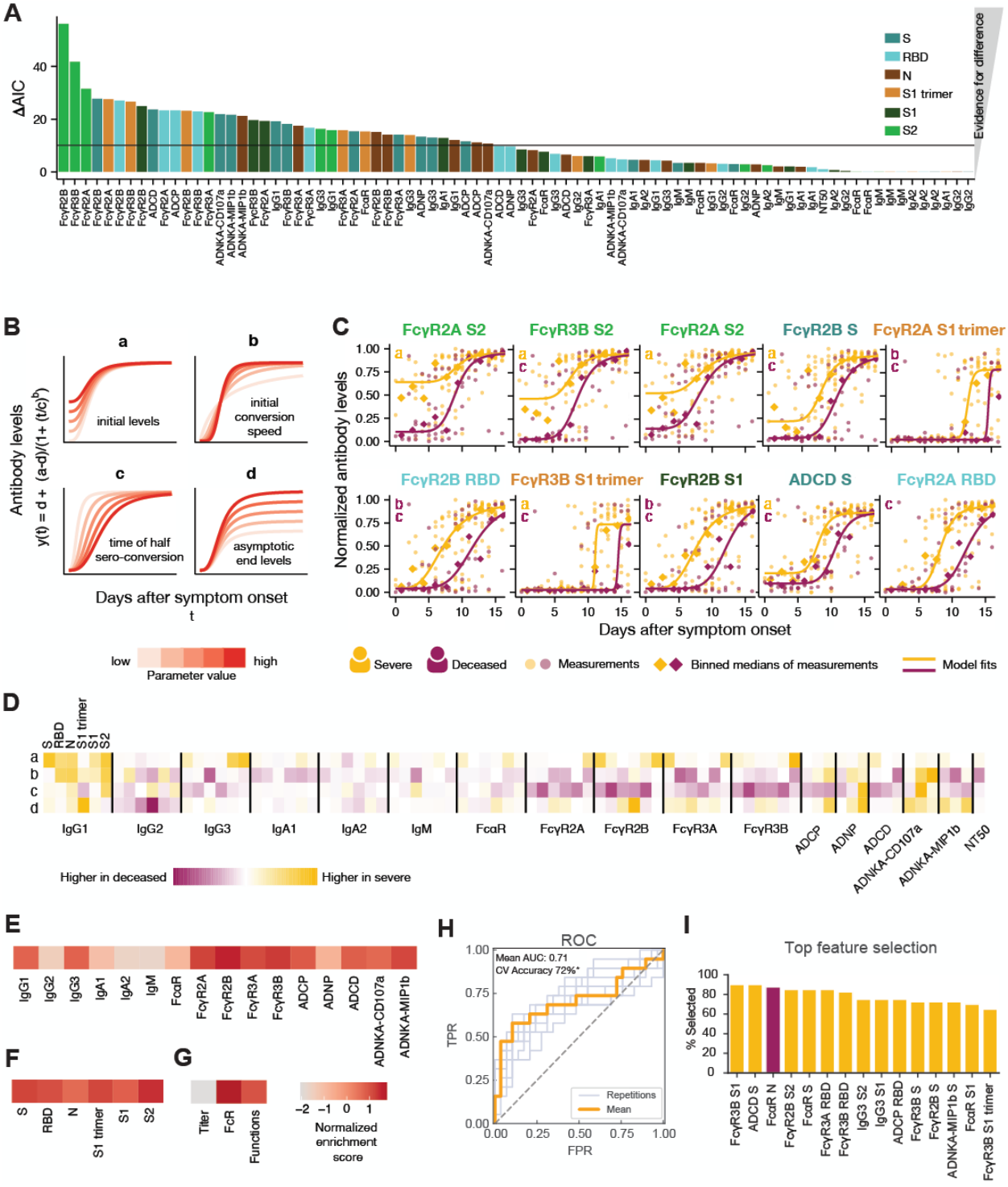


Figure 5-4 Dissecting temporal differences across groups.

(A) The bar plot depicts the ΔAIC of the model without differences between the groups, where the higher the height represents the features that explain trajectory differences best between the groups. The bars are colored according to antigen-specificity, and the vertical line (ΔAIC = 10) indicates the commonly used threshold for rejecting models. (B) Four-parameter logistic growth curves were employed to dissect the specific temporal difference across the groups for each feature. The curves were built by y(t) = d + (a-d)/(1 + (t/c)^b), with y(t) describing the temporal evolution of the antibody levels based on the days after symptom. Differences were then split by: a= defining differences in initial levels, b= the seroconversion speed, c= the seroconversion time and d= the asymptotic end levels. The influence of the parameters on the shape of the curve is shown for varying parameter values indicated by the color. (C) The top 10 different features that differed most between the groups are shown. Dots indicate individual patients, diamonds indicate the binned median, the lines indicate the fitted curves corresponding to the optimal model and the color indicates the group. The specific parameters which differed for the displayed model are indicated in the left corner. The dots and lines are color-coded according to the group. (D) The heatmap shows the Akaike weight averaged parameter differences between the groups. Each row represents a parameter (a, b, c, d) and is normalized across the features, and the color intensity depicts how different the parameter is across the groups, and the color indicates in which group the parameter is higher. Along the x-axis, individual specificities (S, RBD, N, S1 trimer, S1, and S2) are organized in the same repeating order across each Fc-variable that was acquired (subclasses, isotypes, FcR binding, and functions). (E-G) Normalized enrichment scores (a metric of how different the feature is across the two groups) are shown for individual features collapsed by antigen (E), individual antigens (F), and feature "type" (G). The darker the color the more differentially that feature is expressed across the 2 groups. (H) Receiver operating characteristic (ROC) curve shows the model performance in a cross-validation framework. In light blue are the ROC curves for each replicate and the orange is the mean ROC curve showing overall performance. Mean AUC is reported using the mean ROC curve. Classification accuracy was compared to permutated data and significance was assessed using exact p-values of the tail probabilities (*: $p<0.05$). TPR: True positive rate, FPR: False positive rate. (I) Features most often selected during the classification process in yellow are shown and ranked based on the magnitude of the enrichment across severe and deceased individuals. See also Figure S3 and Figure S4.

5.3.6 Antibody profiles evolve more slowly with moderate disease, but mimic survivor profiles

The evolution of early FcR binding and activity in severely infected individuals appeared to emerge as a key correlate of convalescence. However, whether similar antibody profiles developed in individuals with moderate infection, remained unclear. Antibody profiles were therefore compared across individuals with moderate and severe infection who survived. Despite the delayed rise in SARS-CoV-2 antibody levels early in infection (Figure 5-1B and Figure 5-S1) individuals with moderate infection evolved equivalent IgA and IgM levels by the third week following symptoms (Figure 5-1B, Figure 5-5A, and Figure 5-S1). IgG, FcR-binding, and antibody effector functions evolved slowly and remained lower in individuals with moderate disease compared to those with severe disease but continued to develop (Figure 5-5A). Similarly, trajectory analysis demonstrated delayed subclass and isotype titers, FcR binding, and functional responses in individuals with moderate infection (Figure 5-5B). As early as two weeks following symptoms, individuals with a moderate disease trajectory could be resolved from individuals with severe disease based largely on functional antibody features that

were all elevated in individuals with severe disease (Figure 5-5C-D). These data point to similar biophysical, albeit delayed, SARS-CoV-2 antibody profiles among moderately infected individuals that may not require further functional evolution due to early and effective control of the virus.

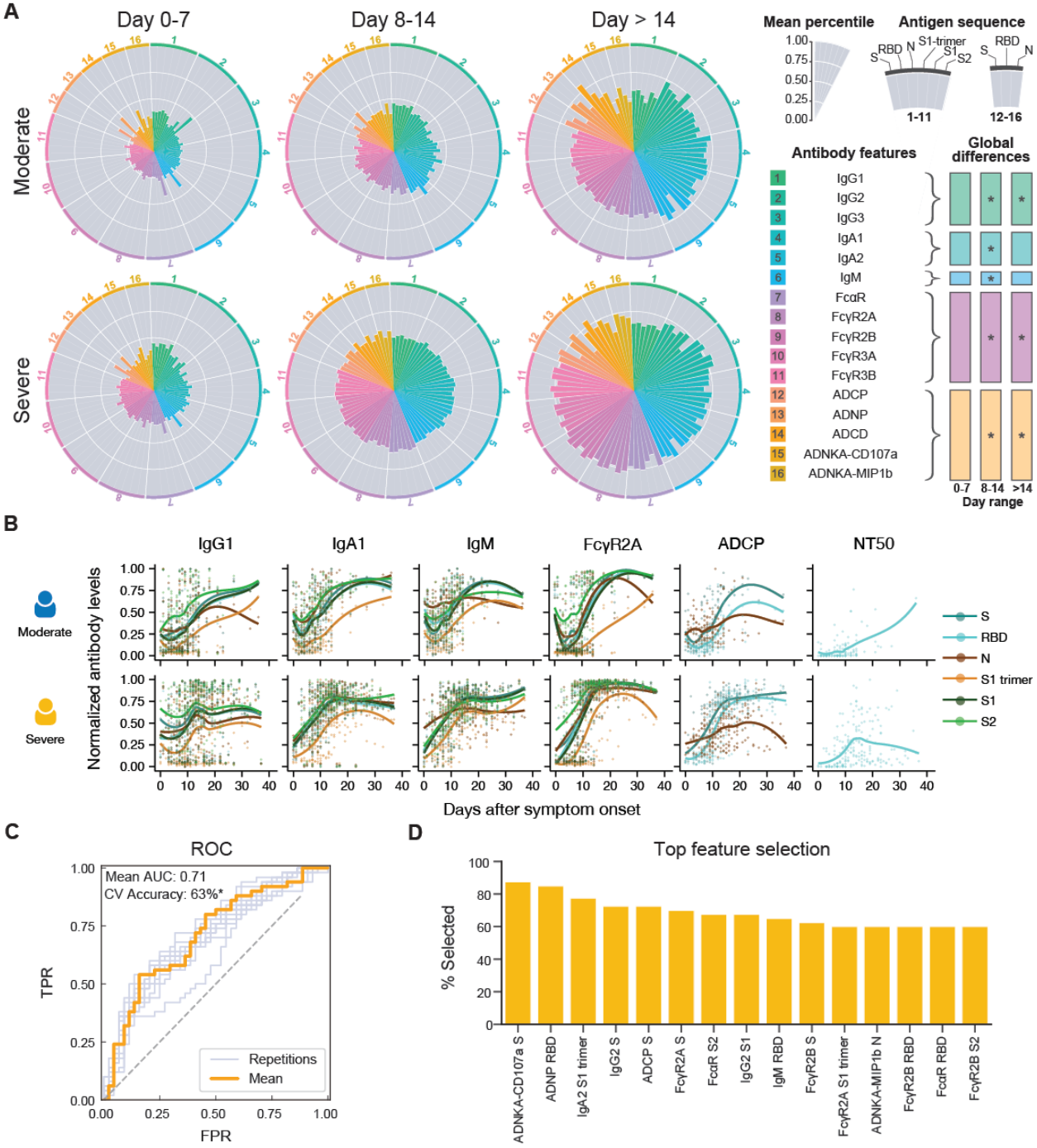


Figure 5-5. Humoral differences between moderate and severe disease

(A) The polar plots depict the mean percentile of each antibody feature at each interval across the moderate (top) and the severely (bottom) infected individuals. The major slices 1-6 cover antigen specific isotypes and subclasses, 7-11 antigen-specific antibody Fc-receptor binding, and 12-16 antigen-specific

antibody mediated functions. For segments 1-11 antigen specificities repeat in the following order: S, RBD, N, S1 trimer, S1, and S2. For segments 12-16 antigen specificities are repeated S, RBD, and N. The size of the wedge depicts the mean percentile ranging from 0-1. On the right, non-parametric combination global p-values are shown, composed of Mann Whitney U test p-values for partial tests within each feature type, and using the Fisher method for combination (*: $p < 0.05$). (B) Normalized antibody levels are shown over time, by days after symptom onset for the moderate and severe groups. Each dot is an individual measurement, the lines show smoothed non-parametric regression models (loess) and the color indicates the antigen-specificity. (C) The receiver operating characteristic (ROC) curve shows the model performance in a cross-validation framework. In light blue are the ROC curves for each replicate and the orange represents the mean ROC curve showing overall performance. Mean AUC is reported using the mean ROC curve. Classification accuracy was compared to permutated data and significance was assessed using exact p-values of the tail probabilities (*: $p<0.05$). TPR: True positive rate, FPR: False positive rate. (D) Features most often selected during the classification process. In yellow are features enriched in the individuals with severe infection, and in blue enriched in the moderates.

5.4 Discussion

In the absence of correlates of immunity, vaccine development efforts have been focused on maximizing antibody titers and neutralization, which have been linked to protection against other pathogens[216–218]. However, once SARS-CoV-2 infection evolves beyond the upper-respiratory tract, dissemination within the lower-respiratory tract, and even across organs, may require more complex immune responses to fully contain and eradicate the infection. Along these lines, emerging vaccine correlates of immunity point to a critical role for both neutralization and Fc-effector functions in protection from infection[118]. Specifically, S- and RBD-specific complement and phagocytosis have been linked to viral control in the bronchoalveolar fluid[118]. These data suggest that a potential synergy is required between the antibody antigen-binding domain (Fab) and constant domain (Fc) functions in immunity against SARS-CoV-2, where neutralization provides a first line of defense, and Fc-effector functions may provide a second line of defense deeper within the respiratory tract. However, whether the same mechanisms are utilized during natural viral clearance or whether spontaneous resolution of infection may be associated with different antibody mechanisms remains to be determined.

Unlike other viral infections, IgM and IgA arise nearly simultaneously in SARS-CoV-2 infection, likely due to the highly compartmentalized nature of this infection. While individuals with moderate infection exhibited lower levels of these isotypes, the similar levels of IgA and IgM among individuals who survived or did not survive severe infection suggest similar early pathogen burden within the groups. However, while the individuals who survived severe infection successfully class switched to IgG antibodies, with high affinity FcR binding capacities, individuals who did not survive showed less robust switching and exhibited a delayed evolution of high affinity FcR binding antibodies and Fc-effector function. This suggests an acute systemic developmental defect in the humoral immune response associated with COVID-19 mortality. Moreover, individuals with moderate disease also exhibited delayed humoral immune evolution, pointing to either non-humoral mechanisms of humoral immune control in moderate

disease or an exposure to less virus, requiring less aggressive immunity for containment and clearance. Furthermore, given the striking perturbations in cellular immunity reported during infection[203], future studies including autologous antibodies and cellular effectors from infected patients could provide enhanced insights into mechanisms of protection or pathology.

Given the staged evolution of antibody isotypes and the time required for affinity maturation, distinct antibody effector functions likely contribute to restriction of infection at different times during infection. Dissecting the trajectory of the humoral immune profiles with respect to time following symptoms and comparing the evolution of humoral features across groups could point to distinct time-specific mechanisms of immunity against SARS-CoV-2. For example, S2-specific FcR binding differed among the groups very early in infection, with S2-specific FcγR2B separating the groups from the first day of symptom onset. Given our emerging appreciation for S2 conservation across coronaviruses, it is plausible that the rapid evolution of S2-specific responses, drawn from pre-existing cross-reactive immunity to other coronaviruses may help facilitate initial viral control[219]. Conversely, S1 trimer-, S-, and RBD-specific humoral immune profiles split between the groups during the second week of infection, highlighting a delayed response to these specificities. Collated, kinetic differences highlighted the unique early and late enrichment of IgG and FcR binding in individuals who survived compared to those who did not, pointing to a critical need for a very early class-switch and maintenance of IgG and FcR binding antibodies for recovery. However, why the deceased class switched to IgA, but not to IgG, early in disease remains unclear. Emerging data point to the aberrant induction of germinal centers among individuals with severe infection[220]. Due to the compartmentalized mucosal nature of the infection, and the ability of T-cell independent IgA-class switching to occur at mucosal sites, it is plausible that equivalent early IgA switching may occur across all severely ill individuals[221], but a lack of sufficient germinal center support may result in poor IgG switching in those who ultimately pass away. T-help is critical for class switching, and T-helper selection biases have been noted with age[222], diabetes[223], and higher body-mass index[224], comorbidities associated with more severe SARS-CoV-2 disease. Additionally, lymphopenia, cytokine dysregulation, and other tissue architectural pathological manifestations may all alter germinal center activity, contributing to this early incomplete class switching. Therefore, future studies considering the dysregulated cellular states observed in the COVID-19 patients, as well as the collaboration of antibodies with cellular immunity, may reveal additional mechanisms critically important for protection.

Antibody responses clearly accrue with more severe disease, raising discussions about a potential pathological role for humoral immunity in disease severity[205]. However, here we did not observe any evidence of higher antibody levels or functions in individuals who ultimately passed away, providing limited evidence of antibody enhancement. It is critical to note that beyond their immunological activities, antibodies also represent

critical biomarkers of the intensity of antigen-exposure. For instance, antibody levels typically increase with antigen-burden in Tuberculosis[225], human immunodeficiency virus[180], and malaria infection[226] but do not contribute to enhanced disease in these settings. Thus, distinguishing the quantitative changes that simply track with pathogen burden from the qualitative changes in antibodies that drive immunity or pathology may be key to unlocking the mechanistic changes that lead to effective immunity.

Neutralization did not differ across the groups in early infection but instead developed with severity of disease. Whereas emerging vaccine studies point to neutralization as a key correlate of immunity[227], after establishment of infection, neutralization may play less of a role in controlling the pathogen. Instead, Fc-effector functions are likely critical for the recognition of infected cells and clearance of new virus. In the context of vaccination, then, neutralization and Fc-effector function are likely to be key collaborative correlates, required to provide first and second line defense in antiviral control, as has been recently observed in vaccinated NHP[228]. However, given that only a small proportion NHPs develop severe disease, like their human counterparts, assessing the impact of these vaccines on attenuating severe disease remains difficult outside of very large primate studies. Nonetheless, harmonizing human pathogenesis studies with NHP vaccine studies offers a unique opportunity to uncover the key correlates of immunity to guide vaccine development.

While no influence was observed in antibody profiles across therapeutic interventions or co-morbid conditions, these data argue for independent influences of lung-disease associated pathophysiological changes in collaboration with SARS-CoV-2-specific antibody profiles in shaping disease outcome. However, collectively, the work here argues for the evolution of a robust, protective functional humoral immunity among individuals who develop severe infection that is perturbed soon after infection among non-survivors. Defining early biomarkers that identify individuals on a deleterious clinical trajectory may provide early opportunities to triage individuals to better and more intense care. Alternatively, these data also highlight the importance of accessing the full range of humoral immune functions to fully provide protection from SARS-CoV-2 infection and disease.

5.5 Limitations of study

There are several limitations in this study. First, given that patients are admitted and discharged at different stages during their disease trajectories, identical temporal sampling was not possible across all samples. However, given the large number of samples, temporal trajectories were constructed across clinical groups. Moreover, complementary modelling approaches were used to ensure that the trajectories were representative of the patient class and that conserved signatures of protection were identified across the groups. Additionally, antibody mediated functional assays were performed with cells from healthy donors, rather than autologous cells from the infected patients. However recent findings suggest that COVID-19 patients, especially those with

severe disease, exhibit unique cellular deficiencies and perturbated cellular states[203]. Therefore, future studies investigating the composite effects of humoral functional immunity linked to altered cell states observed in COVID-19 patients, may reveal additional mechanisms critically important for mechanistically understanding protection. Lastly, peripheral antibodies were analyzed in this study. However, localized production of antibodies may result in the production of localized antibodies with distinct functional properties that may drive unique protective or pathological functions. Thus, future studies focused on compartment specific antibody functional profiles may also provide enhanced resolution on protective or pathological functions of antibodies at the site of viral infection and replication.

Collectively, the data presented here argue for a role for functional humoral immunity in the resolution of severe SARS-CoV-2 infection. Although, additional cohorts may provide future mechanistic insights into the specific signals that result in the generation of these protective humoral immune responses, these data point to specific antibody functions that may be of high value in vaccine or therapeutic design.

5.6 Methods

5.6.1 Sample cohort

Plasma samples from 193 subjects infected with SARS-CoV-2, from Massachusetts General Hospital (MGH), were included in this study. Individuals were tested for SARS-CoV-2 by real-time reverse-transcriptase–polymerase-chain-reaction (RT-PCR) using nasopharyngeal swabs. Subjects that tested positive were enrolled in the study upon hospital admission, and samples at admission were included in this study (Figure 5-1A and Table 5-S1). Patients were admitted to the hospital due to moderate to severe symptoms of COVID-19 and were followed over multiple timepoints (ranging from 1-8 timepoints per individual). Disease outcome was classified as either discharged or deceased. Severity of disease was classified by admission to the intensive care unit (ICU). All enrolled participants gave written, informed consent. Demographic information including age, and whether patients were immunosuppressed are summarized across the groups (Table 5-S1). Plasma samples from 32 hospitalized individuals which tested negative by RT-PCR were used as negative controls throughout the study. All experimental data was captured in two technical replicates and the average value was reported for all assays. This study was approved by the MGH Human Subjects Institutional Review Board.

5.6.2 Primary immune cells

Fresh peripheral blood was collected by the MGH Blood bank from healthy human volunteers. All volunteers gave signed consent and were over 18 years of age, and all samples were de-identified before use. The study was approved by the MGH Institutional Review Board. Human NK cells were isolated from fresh peripheral blood

and maintained at 37˚C, 5% CO2 in RPMI with 10% fetal bovine serum, L-glutamine, penicillin/streptomycin.

5.6.3 Cell lines

HL-60 cells (ATCC), a promyelocytic leukemia cell line, were grown in IMDM supplemented with 20% fetal bovine serum and penicillin/streptomycin at 37°C, 5% CO2. For neutrophil differentiation, the media was supplemented with 1.25% DMSO for THP-1 cells (ATCC), a monocytic leukemia cell line, was maintained in RPMI supplemented with 10% fetal bovine serum, L-glutamine, penicillin/streptomycin, HEPES, and beta-mercaptoethanol. THP-1 cells were grown at 37˚C, 5% CO2.

5.6.4 Luminex

Antigen-specific antibody subclass/isotype and Fc-receptor (FcR) binding levels were measured using a 384-well based customized multiplexed Luminex assay, as previously described[157]. This high-throughput assay allows for the assessment of relative antibody concentration against SARS-CoV-2 RBD, HKU1 RBD, NL63 RBD (all kindly provided by Aaron Schmidt, Ragon Institute), SARS-CoV-2 nucleocapsid (N) protein (Aalto Bio Reagents), and SARS-CoV-2 spike protein (S) (kindly provided by Eric Fischer, Dana Farber) as well as S1 (Sino Biological, 40591-V08B1) , S1 trimer (provided by Bing Chen), S2 (Sino Biological, 40590-V08B), and a mix of HA A/Michigan/45/2015 (H1N1), HA A/Singapore/INFIMH-16-0019/2016 (H3N2), B/Phuket/3073/2013 (Immunetech). In brief, antigens were covalently bound to fluorescent carboxyl- modified microspheres (Luminex) by NHS-ester linkages using EDC and Sulfo-NHS (Thermo Scientific). Antigen-coupled beads were then washed and blocked before adding plasma samples at an appropriate sample dilution (1:500 for IgG1, 1:1000 for all Fc- receptors, and 1:100 for all other isotype/subclass readouts). After an overnight incubation at 4oC while shaking at 700rpm, immune complexed microspheres were washed using an automated plate washer (Tecan) with 0.1% BSA 0.02% Tween-20. Antigen-specific antibody titers were detected using a PE-coupled detection antibody for each subclass and isotype (IgG1, IgG2, IgG3, IgA1 and IgM, Southern Biotech), and Fc-receptors were fluorescently labeled with PE before addition to immune complexes (FcR-2A, -2B, -3A, -3B, Duke Protein Production facility). Plasma samples were acquired via flow cytometry, using an iQue (Intellicyt) and S-Lab robot (PAA). Analysis was done using ForeCyt software by gating on fluorescent bead regions and PE median fluorescent intensity (MFI) was reported as readout for antigen-specific antibody titers.

5.6.5 Effector functional assays

Bead-based assays were used to quantify antibody-dependent cellular phagocytosis (ADCP), antibody-dependent neutrophil phagocytosis (ADNP) and antibody-dependent complement deposition (ADCD) in the MGH SARS-CoV-2 cohort, as previously described[159,229,162,160,161]. Yellow (ADNP and ADCP) as well as red (ADCD) fluorescent neutravidin beads (Thermo Fisher) were coupled to biotinylated SARS-CoV-2 RBD, N

and S antigens and incubated with diluted plasma (ADCP and ADNP 1:100, ADCD 1:10) to allow immune complex formation for 2h at 37˚C. To assess the ability of sample antibodies to induce monocyte phagocytosis, THP-1s (ATCC) were added to the immune complexes at 1.25E5cells/ml and incubated for 16h at 37˚C. For ADNP, HL-60 cells were differentiated into CD11-expressing neutrophils with media including 1.25% DMSO for 5 days as described previously (Worley et al., 2018b), cells were maintained below 1E6 cells/ml. On day 5, 5E5 cells/ml were added per well to immune complexed yellow beads and incubated for 16h at 37˚C. Afterwards, neutrophils were stained with an anti-CD11 BV605 detection antibody (Biolegend) and fixed with 4% paraformaldehyde (Alfa Aesar). In order to measure antibody-dependent deposition of C3, lyophilized guinea pig complement (Cedarlane) was reconstituted according to manufacturer's instructions and diluted in gelatin veronal buffer with calcium and magnesium (GBV++) (Boston BioProducts). Subsequently, C3 was detected with an anti-C3 fluorescein-conjugated goat IgG fraction detection antibody (Mpbio).

Antibody-dependent NK cell activity was measured via an ELISA-based assay, as described previously [163]. Briefly, plates were coated with 3μg/mL of antigen (SARS-CoV-2 RBD, N and S) and blocked overnight at 4˚C. NK cells were isolated the day prior via RosetteSep (Stem Cell Technologies) from healthy buffy coats (MGH blood donor center) and rested overnight in 1 ng/ml IL-15 at 1.5E5 cells/ml (Stemcell). The next day, diluted plasma samples were added to the antigen-coated plates (1:50 dilution) and incubated for 2h at 37˚C. NK cells were mixed with a staining cocktail containing CD107a PE-Cy5 (BD), Golgi stop (BD) and Brefeldin A (BFA, Sigma Aldrich) and 2.5E5 cells/ml were added per well and incubated for 5h at 37˚C. Following, cells were fixed (Perm A, Life Tech) and stained for surface markers with anti-CD16 APC-Cy7 (BD), anti-CD56 PE-Cy7 (BD) and anti-CD3 PacBlue (BD). Subsequently, cells were permeabilized using Perm B (Life Tech) and intracellularly stained with an anti-MIP-1β PE (BD) antibody.

All assays were acquired via flow cytometry with iQue (Intellicyt) and an S-Lab robot (PAA). For ADCP, events were gated on singlets and bead-positive cells, whereas neutrophils were defined as CD11 positive events followed by gating on bead-positive neutrophils. A phagocytosis score was calculated for ADCP and ADNP as (percentage of bead-positive cells) x (MFI of bead-positive cells) divided by 10000. ADCD was reported as median of C3 deposition. NK cells were defined as CD3-, CD16+ and CD56+. Data were reported as the percentage of cells positive for CD107a and MIP-1β.

5.6.6 Neutralization

Neutralization was determined using a SARS-CoV-2 pseudovirus expressing a luciferase reporter gene, as described previously[198]. To generate the pseudovirus, the packaging construct psPAX2 (Cat# 11348, AIDS Reagent), luciferase reporter plasmid pLenti-CMV Puro-Luc (Cat# 17447, Addgene) and Spike protein expressing pcDNA3.1-SARS CoV-2.SΔCT were transfected in HEK293T cells by the calcium phosphate method at a ratio

of 1:1:0.5. Supernatants were collected and filtered with a 0.45-μm filter 48 hours post-transfection. For the neutralization assay, HEK293Ts were transfected with pcDNA3.1(-)-hACE2 (Cat# 1786, Addgene). The hACE2-expressing HEK293T cells were plated in 96-well plates 12 hours after transfection at a density of 20,000 cells/well and rested overnight. Serum was heat inactivated by incubation at 56°C for 30 minutes. Heat inactivated serum was twofold serially diluted, mixed with 50uL of pseudovirus, and incubated at 37°C incubator for 1 hour. After incubation, the serum/pseudovirus mixed was added to the HEK293T/hACE2 cells. Six hours after infection, cell medium was replenished. Cells were lysed in Steady-Glo Luciferase Assay (Promega) 48 hours after infection. A luciferase assay was performed with luciferase assay reagent (Promega) according to the manufacturer's protocol. NT50 was defined as the concentration of serum required to achieve half maximal neutralization.

5.6.7 Quantification and statistical analysis

All analyses were performed using python version 3.6.8, and R version 3.6.1. Raw data and custom code are available in Supplementary Information.

5.6.8 Polar plots

Polar plots summarize the mean percentile of clinical groups across day ranges from symptom onset. First, percentile rank scores were determined for each feature across all time ranges. Samples which were sampled multiple times within an interval were represented by the mean value, and mean percentiles were determined using samples corresponding to intervals and clinical groups.

5.6.9 Non-parametric combination

Global statistical differences of feature types between groups were assessed using non-parametric combination[230,231]. Briefly, for each feature class (i.e. IgG) partial tests consisting of p-values determined by Mann-Whitney U tests were performed for each sub feature (i.e. IgG1 RBD, IgG1 N, etc.), then p-values were combined using the Fisher method. Next, the data was permutated a thousand times, preserving the permutated structure for partial tests, and was used to construct a null distribution of global statistics. Finally, the true global statistic was directly compared the null distribution and the global p-value was determined.

5.6.10 Batch effect evaluation

To evaluate batch effects by confounders including age, sex, body-mass-index (BMI), well plate and past pulmonary disease, UMAP[232] based methods were used to reduce the high-dimensional serological data into a two-dimensional space for qualitative evaluation, and then quantified by the degree of local neighborhood diversity using local inverse Simpson's Index (LISI)[233] (Figure 5-S2A). First, titers and FcR features were log10 transformed. Then, using the first 40 principle components (PCs) that explain more than 95% of the variance, variation was extracted by principal component analysis

(PCA) [234] using the 'prcomp' function in R package 'stats'. Next, the principal components were mapped into a two-dimensional space through the UMAP technique implemented using the R package 'umap' with fine-tuned parameters (neighbor = 30, min. dist = 0.1). Finally, the LISI score was calculated using the R package 'immunogenomics/LISI'. The score ranged from one to the number of categories and was used to evaluate the degree of mixing in the UMAP embedded space. The larger the LISI score, the higher the degree of heterogeneity among the samples and, therefore, the smaller the confounding effect. Unknown samples (for BMI and previous pulmonary diseases) were excluded, and the continuous variables BMI and age were grouped in 4 and 7 categories, respectively (BMI: <25, [25,30), [30, 35), >=35, age: [30,40), [40,50), [50,60), [60,70), [70,80), [80,90), [90,100). The observations for the continuous variables were the same when using a local average score instead of categorizing the samples.

5.6.11 Temporal analysis

First, the Luminex measurements and ADCD were log10 transformed. All measurements were normalized such that the minimal value across groups was 0, and the maximal value was 1. For visualization, a non-parametric regression model was employed to obtain a smoothed line using the R function 'loess' (span = 0.7). It is critical to note, that the late rise of some curves is attributable to a limited number of late timepoints, and not due to a true elevation in antibody levels. To understand and determine differences in the antibody dynamics between the groups, we described the dynamics of each antibody feature y at the group-level using a four-parameter logistic growth curve:

$$y(t) = d + \frac{(a-d)}{\left(1+\frac{t}{c}\right)^b}$$

with t denoting the days after symptom onset, and a, b, c, and d denoting biological parameters for the initial antibody levels at the day of symptom onset (a), the initial seroconversion speed (b), the time of 50% seroconversion (c) and the asymptotic end levels (d). To detect differences between the individuals who survived severe SARS-CoV-2 infection and those who did not, we built models that describe the dynamics of both groups simultaneously, allowing for combinations of parameters to differ between the groups, while the others are shared between the groups. With 4 parameters, there are 16 possible combinations/models for each feature that could potentially explain the antibody feature dynamics. For each feature, each of the 16 models was fitted to the data using maximum likelihood estimation, treating each measurement as an independent data point and assuming that differences in measurements arose due to measurement noise. We employed a Laplacian likelihood function, which has been shown to be robust against outliers in the data[235].

In addition to the parameters a, b, c, and d, also the noise parameter was estimated from the data. Therefore, the simplest model assuming that there is no difference between the two groups has 5 parameters, while the most complex model has 9 parameters and allows all curve parameters to differ. The corresponding likelihood functions were maximized using a multi-start gradient-based optimization[236] with parameter boundaries $a \in [0.01,1], b \in [0.01, 100], c \in [0.01,1000], d \in [0.01, 1.2], \sigma \in [0.01, 1000]$ and 50 starts which were increased to 500 if the maximal value was not found more than 3 times within a log-likelihood threshold of 0.1. Due to improved numerical performance, the parameters were estimated in log10-space[237].

To detect whether there were differences between the groups, and, furthermore, decide which particular differences were most distinct across the groups, we calculated the Akaike Information Criterion (AIC)[238]:

$$AIC_i = 2n_{\theta_i} - 2\log L\left(\hat{\theta}_i\right), i = 1, \dots, 16$$

for each of the 16 models. Here, n_{θ_i} denoted the number of parameters of model i and log-likelihood function $log\ L$ evaluated at the maximum likelihood estimate. The AIC rewards a good fit and penalizes a higher number of parameters to avoid overfitting and, thus, low values of AIC are preferable. The model with the lowest AIC value was then chosen to be the best model and, commonly, models with a difference in AIC values:

$$\Delta AIC_i = AIC_i - \min_{j \in \{1,\dots,16\}} AIC_j$$

that were higher than 10 were rejected[239].To analyze the overall differences in parameters across the groups (Figure 5-4D), the maximum likelihood estimates for all 16 models were combined by weighting the contribution of individual models by the Akaike weight:

$$w_i = \frac{e^{-0.5AIC_i}}{\sum_{j=1}^{16} e^{-0.5AIC_j}}$$

Weights for models that were not plausible were ranked low, and, therefore, did not contribute to the parameter estimate.

5.6.12 Enrichment analysis

Enrichment of features determined to be different between groups was determined using the same framework employed by Gene Set Enrichment Analysis (GSEA)[240]. The R package ‘fgsea’ was used to determine normalized enrichment scores[241]. ΔAIC were used as weights and null distributions were constructed with size matched random selection of features over 10000 times.

5.6.13 Classification of clinical groups

Random forest[199] classification models were trained to distinguish clinical groups using minimal sets of features, to avoid overfitting and identify features that were most predictive. Data were not corrected or transformed prior to analysis but features for which 70% of values fell below one standard deviation above the mean of SARS-CoV-2 negative samples were pruned. Samples which had multiple time points within a time interval were represented as a single mean value.

Models were trained and tested in a fourfold cross-validation framework using random stratified sampling to ensure the groups are represented each set. Within each fold, random forest hyperparameters (number of estimators, max depth, and max features) which best optimize balanced accuracy were selected based on the training set with a random search in a threefold cross validation framework. Balanced subsampling during bootstrapping was used for the construction of each random forest model. Once hyperparameters were determined the minimal set of features which optimizes out of bag accuracy on the training set is selected using recursive feature elimination with fourfold cross validation. Finally, a random forest classifier was fit using both the selected hyperparameters and minimal features and used to predict the test set. Multiple iterations of fold-specific feature selections were performed to determine the features which best distinguish clinical groups and the stability of the results.

Performance was determined using receiver operating characteristic curves (ROC) and summarized with the area under curves (AUC). ROC curves were constructed for each repetition using probability estimates, and the mean ROC curve was determined by using the mean probability for each sample across replicates. Performance and robustness of the model was also contrasted to negative control models built from permuted data. Within each fold of the model the training set labels were shuffled, and classification accuracies were generated using the same process. These control models were generated 50 times for each repetition. Predicted and true outcomes were compared to determine accuracy. Robustness was defined as the exact p-values of the tail probabilities of the true distributions within the control distributions. Reported are the median p-values across ten independent cross-validation repetitions[107].

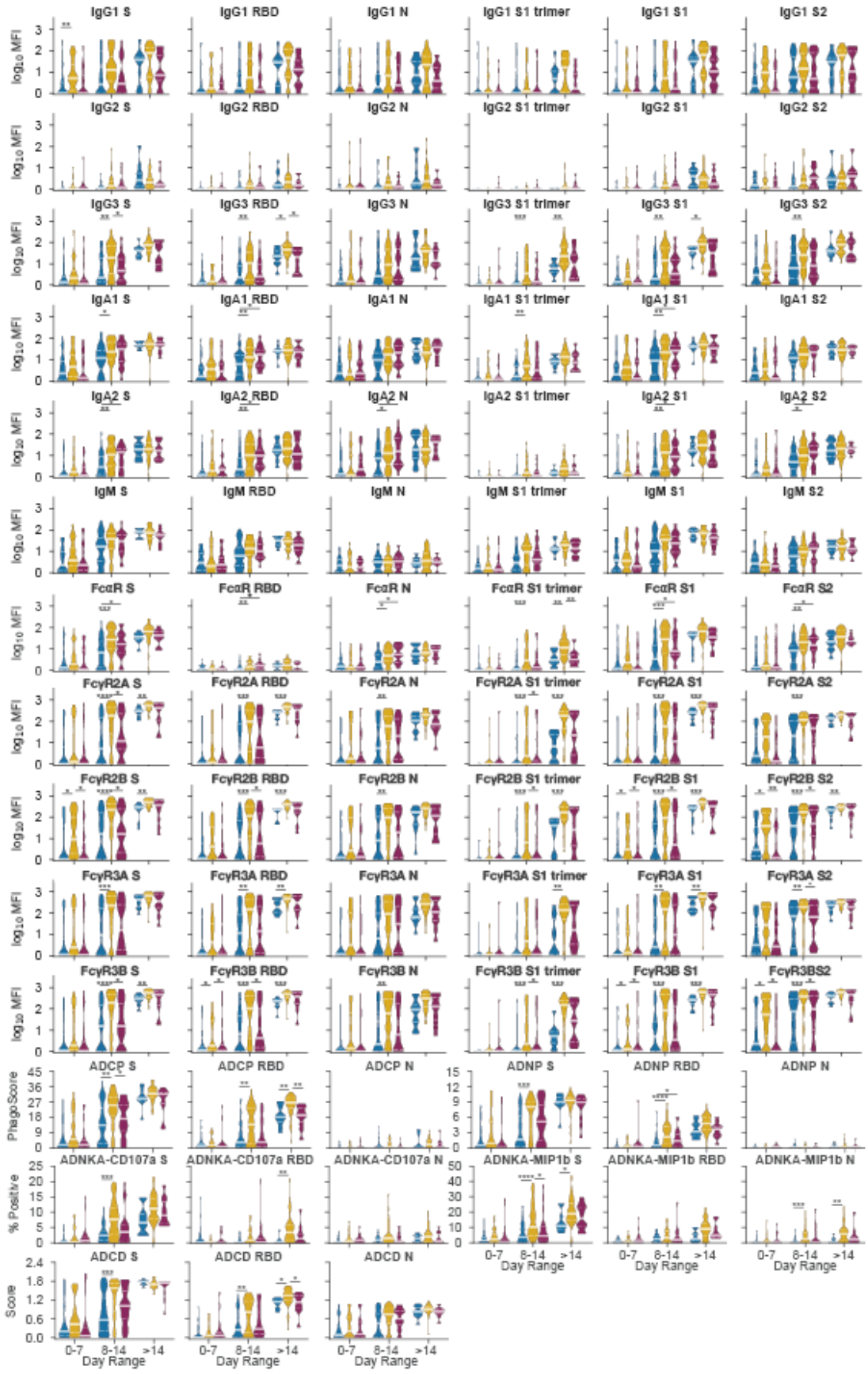
5.7 Supplementary figures and tables

Characteristic	Moderate (n = 82)	Severe (n = 76)	Deceased (n = 35)
Female sex - no. (%)	29 (35.3)	29 (38.2)	8 (22.9)
Age range - no. (%)			
Younger than 50	26 (31.7)	25 (32.9)	3 (8.58)
50-59	11 (13.4)	21 (27.6)	6 (17.1)
60-69	16 (19.5)	17 (22.4)	6 (17.1)
70-79	17 (20.1)	7 (9.21)	7 (0.2)
80 and older	12 (14.6)	6 (7.90)	13 (37.1)
Clinical outcome - no. (%)			
D/C home	59 (72.0)	21 (27.6)	0 (0)
D/C rehab/longterm care facility	21 (26.8)	55 (72.3)	0 (0)
Death	1 (1.22)	0 (0)	35 (100)
Days from onset of symptoms to sample collection - median (IQR)	7 (6.75)	9 (8.25)	6 (8)
Number of time points - median (IQR)	1 (1)	2 (1.25)	2 (1.5)
Acute respiratory distress syndrome - no. (%)[a]	2 (2.44)	43 (70.5)	20 (74..1)
Immunosuppressed - no. (%)	6 (7.32)	3 (3.95)	5 (14.3)
Past pulmonary disease - no (%)[b]	19 (23.5)	17 (23.3)	15 (46.9)
Body-mass-index - median (IQR)[c]	27.8 (8)	30.9 (6.55)	28.5 (8.83)
Treatment - no. (%)			
Remdesivir	4 (4.89)	14 (18.4)	2 (5.71)
Inhaled nitric oxide	4 (4.89)	14 (18.4)	7 (20.0)
Hydroxychloroquine	55 (67.1)	68 (89.5)	33 (94.3)
Tocilizumab	0 (0)	2 (2.63)	0 (0)
Azithromycin	44 (53.7)	69 (90.8)	28 (80.0)
Ceftriaxone	36 (43.9)	60 (78.9)	30 (85.7)
Vancomycin	9 (11.0)	53 (70.0)	27 (77.1)
Sarilumab	2 (2.44)	1 (1.32)	3 (8.57)
Lopinavir/Ritonavir	0 (0)	3 (3.95)	0 (0)

Table 5-S1. Demographics of cohort, related to Figure 1.

[a] Information about acute respiratory distress syndrome was not available for 15 individuals in the severe group, and 8 of the deceased group.
[b] Information about past pulmonary diseases was not available for 1 individual in the moderate group, 3 in the severe group, and 3 in the deceased group.
[c] Information about body-mass-index was not available for 13 individuals in the moderate group, 13 in the severe group, and 3 in the deceased group.

A


IgG1 S
IgG1 RBD
IgG1 N
IgG1 S1 trimer
IgG1 S1
IgG1 S2
IgG2 S
IgG2 RBD
IgG2 N
IgG2 S1 trimer
IgG2 S1
IgG2 S2
IgG3 S
IgG3 RBD
IgG3 N
IgG3 S1 trimer
IgG3 S1
IgG3 S2
IgA1 S
IgA1 RBD
IgA1 N
IgA1 S1 trimer
IgA1 S1
IgA1 S2
IgA2 S
IgA2 RBD
IgA2 N
IgA2 S1 trimer
IgA2 S1
IgA2 S2
IgM S
IgM RBD
IgM N
IgM S1 trimer
IgM S1
IgM S2
FcαR S
FcαR RBD
FcαR N
FcαR S1 trimer
FcαR S1
FcαR S2
FcγR2A S
FcγR2A RBD
FcγR2A N
FcγR2A S1 trimer
FcγR2A S1
FcγR2A S2
FcγR2B S
FcγR2B RBD
FcγR2B N
FcγR2B S1 trimer
FcγR2B S1
FcγR2B S2
FcγR3A S
FcγR3A RBD
FcγR3A N
FcγR3A S1 trimer
FcγR3A S1
FcγR3A S2
FcγR3B S
FcγR3B RBD
FcγR3B N
FcγR3B S1 trimer
FcγR3B S1
FcγR3BS2
log10 MFI
ADCP S
ADCP RBD
ADCP N
ADNP S
ADNP RBD
ADNP N
PhagoScore
ADNKA-CD107a S
ADNKA-CD107a RBD
ADNKA-CD107a N
ADNKA-MIP1b S
ADNKA-MIP1b RBD
ADNKA-MIP1b N
% Positive
ADCD S
ADCD RBD
ADCD N
Score
0-7
8-14
>14
Day Range

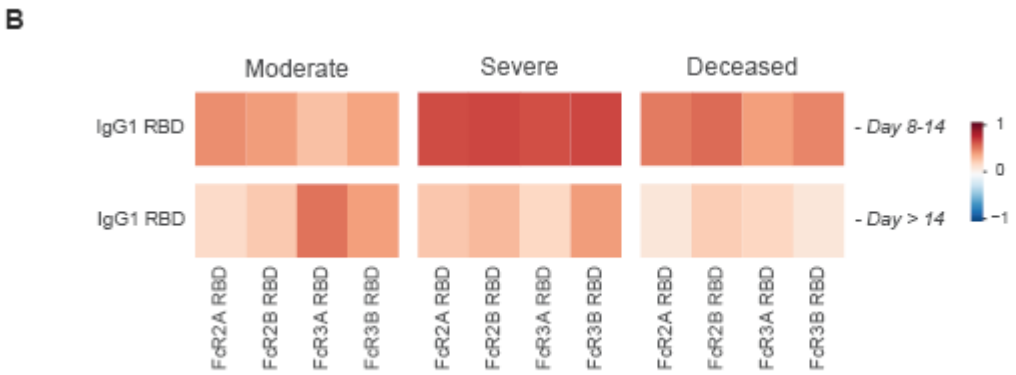


Figure 5-S1. Antibody evolution by week following symptoms and RBD Specific IgG1 Fc-receptor coordination, related to Figure 1.

(A) 193 plasma samples from hospitalized SARS-CoV-2 infected individuals were profiled against SARS-CoV-2 spike antigen (S), receptor binding domain (RBD), nucleocapsid protein (N), subunit 1 of the spike protein as a trimer (S1 trimer), subunit 1 of spike protein (S1) as a monomer, and subunit 2 of the spike protein (S2). Distributions of titers across moderate (blue), severe (yellow), and deceased (red) individuals are shown in the violin plot over the first, second-, and third week following symptom onset. The solid white line represents the median and the dotted lines the first and third quartiles. A Kruskal-Wallis test was used to evaluate statistical differences across groups for all intervals and features and was corrected for multiple hypothesis testing with the Benjamini-Hochberg procedure. If statistically significant then a two-sided Mann-Whitney U test was performed for post-hoc comparisons. Significance shown corresponds to the Mann-Whitney U test p-values (*: $p < 0.5e-1$, **: $p < 0.5e-2$, ***: $p < 0.5e-3$, ****: $p < 0.5e-4$, *****: $p < 0.5e-5$). Antibody dependent cellular phagocytosis (ADCP), antibody dependent neutrophil phagocytosis (ADNP), antibody dependent complement deposition (ADCD), antibody dependent Natural killer cell activation (ADNKA). (B) Spearman correlations were assessed within each clinical group at two- and three-weeks post symptom onset between IgG1 titers and FcγR binding profiles to assess shifts and changes in antibody glycosylation. Red corresponds to higher correlation, white to no correlation, and blue to anti-correlation.

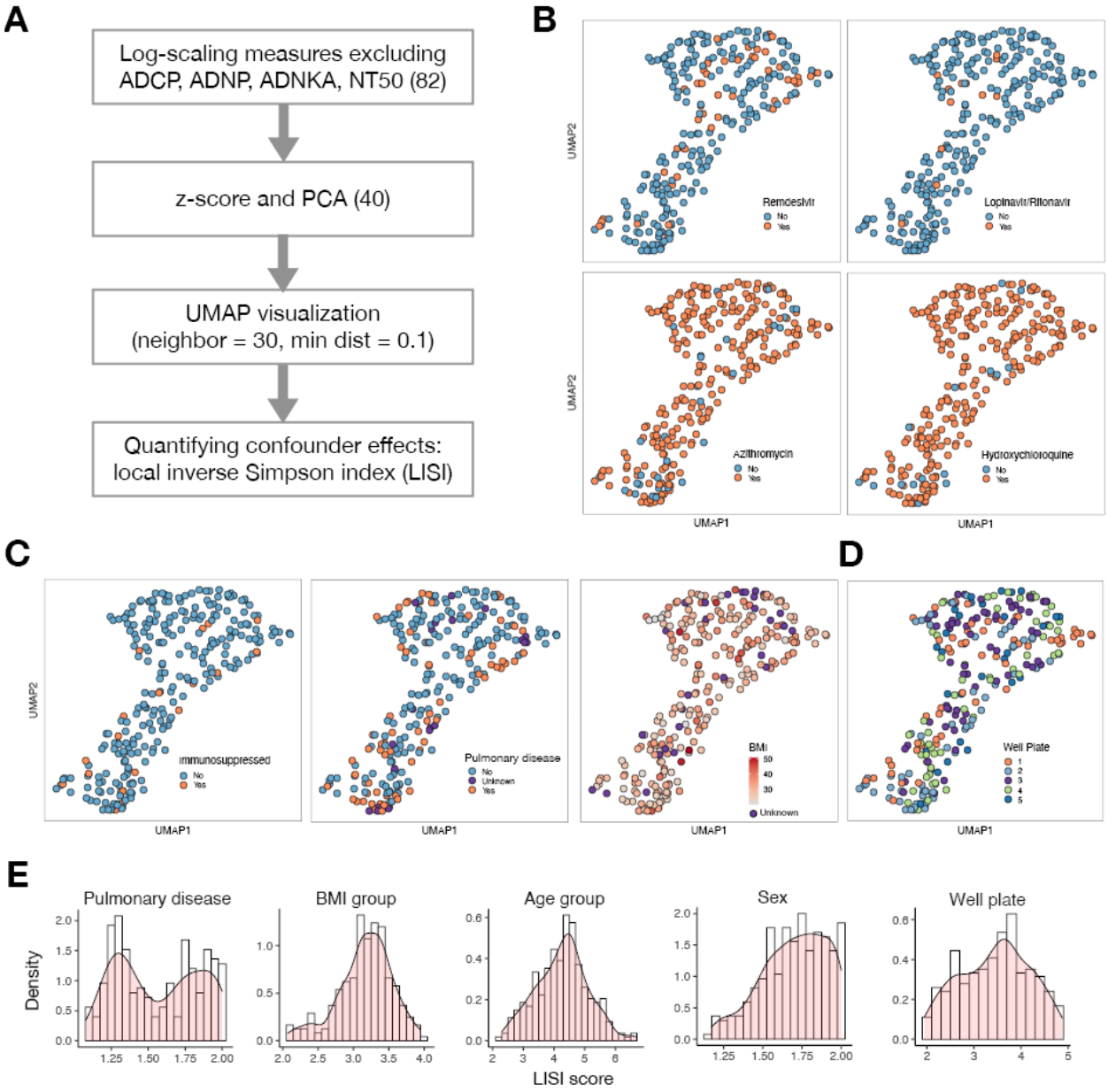


Figure 5-S2. Batch effect evaluation, related to Figure 3.

(A) The algorithm provides an overview of the evaluation pipeline. The first 40 principal components (PCs) explained more than 95% of the variance and were extracted from z-scored measurements. UMAP was applied to map the extracted PCs into two dimensions, in which the local diversity was quantified by the local inverse Simpson index (LISI). (B) UMAP visualizations highlight limited antibody profile differences across four of the treatments that were used in the SARS-CoV-2 patients. (C) UMAP visualizations show the influence of co-morbid conditions – immunosuppression, pulmonary disease, and body-mass-index (BMI) - on antibody profiles. (D) UMAP visualization was used to probe for potential plate-batch effects, where each color represents a different plate run across Systems Serology. (E) The histograms show the distributions of LISI scores for past pulmonary disease, body-mass-index (BMI), age, sex, and well plate. LISI measures the degree of mixing in an embedding ranging from 1 to the number of categories (e.g., 2 for sex), where larger LISI scores indicate less separation and more mixing.

Unknown samples were excluded and the continuous variables BMI and age were grouped in 4 and 7 categories, respectively (BMI: <25, [25,30), [30, 35), >=35, age: [30,40), [40,50), [50,60), [60,70), [70,80), [80,90), [90,100)). Overall, the histograms show no substantial skewing of the antibody profiles.

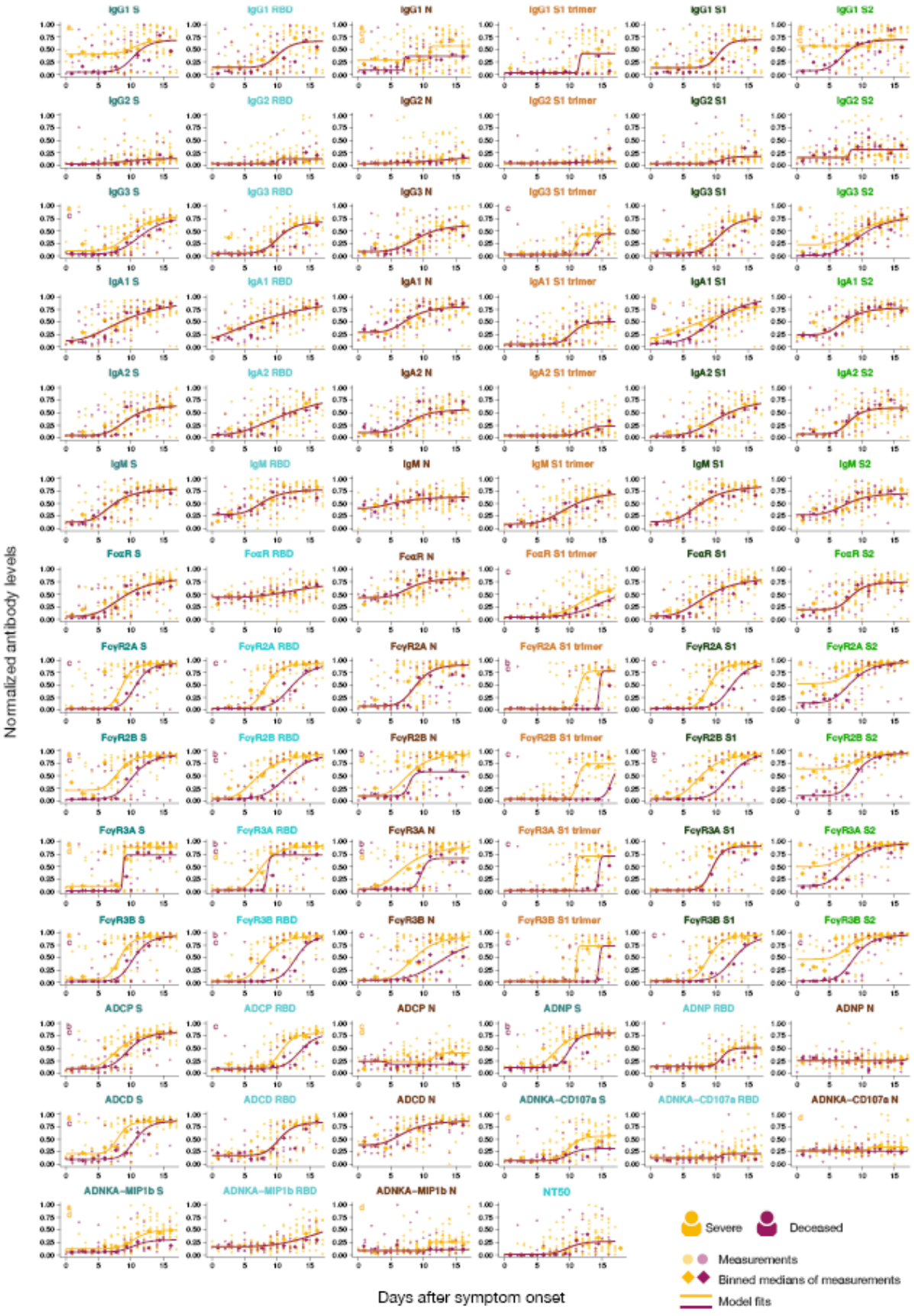

IgG1 S
IgG1 RBD
IgG1 N
IgG1 S1 trimer
IgG1 S1
IgG1 S2
IgG2 S
IgG2 RBD
IgG2 N
IgG2 S1 trimer
IgG2 S1
IgG2 S2
IgG3 S
IgG3 RBD
IgG3 N
IgG3 S1 trimer
IgG3 S1
IgG3 S2
IgA1 S
IgA1 RBD
IgA1 N
IgA1 S1 trimer
IgA1 S1
IgA1 S2
IgA2 S
IgA2 RBD
IgA2 N
IgA2 S1 trimer
IgA2 S1
IgA2 S2
IgM S
IgM RBD
IgM N
IgM S1 trimer
IgM S1
IgM S2
FcαR S
FcαR RBD
FcαR N
FcαR S1 trimer
FcαR S1
FcαR S2
FcγR2A S
FcγR2A RBD
FcγR2A N
FcγR2A S1 trimer
FcγR2A S1
FcγR2A S2
FcγR2B S
FcγR2B RBD
FcγR2B N
FcγR2B S1 trimer
FcγR2B S1
FcγR2B S2
FcγR3A S
FcγR3A RBD
FcγR3A N
FcγR3A S1 trimer
FcγR3A S1
FcγR3A S2
FcγR3B S
FcγR3B RBD
FcγR3B N
FcγR3B S1 trimer
FcγR3B S1
FcγR3B S2
ADCP S
ADCP RBD
ADCP N
ADNP S
ADNP RBD
ADNP N
ADCD S
ADCD RBD
ADCD N
ADNKA–CD107a S
ADNKA–CD107a RBD
ADNKA–CD107a N
ADNKA–MIP1b S
ADNKA–MIP1b RBD
ADNKA–MIP1b N
NT50
Normalized antibody levels
Days after symptom onset
Severe
Deceased
Measurements
Binned medians of measurements
Model fits

Figure 5-S3. Temporal evolutionary curves of antibody features, related to Figure 4.

For each antibody feature, the optimal model fit is shown for each group across each feature. Dots indicate individual patients, diamonds indicate the binned median, the lines indicate the fitted curves corresponding to the optimal model and the color indicates the group. The parameters which are different for the displayed model are indicated in the left corner and color-coded according to the group for which the parameter is higher.

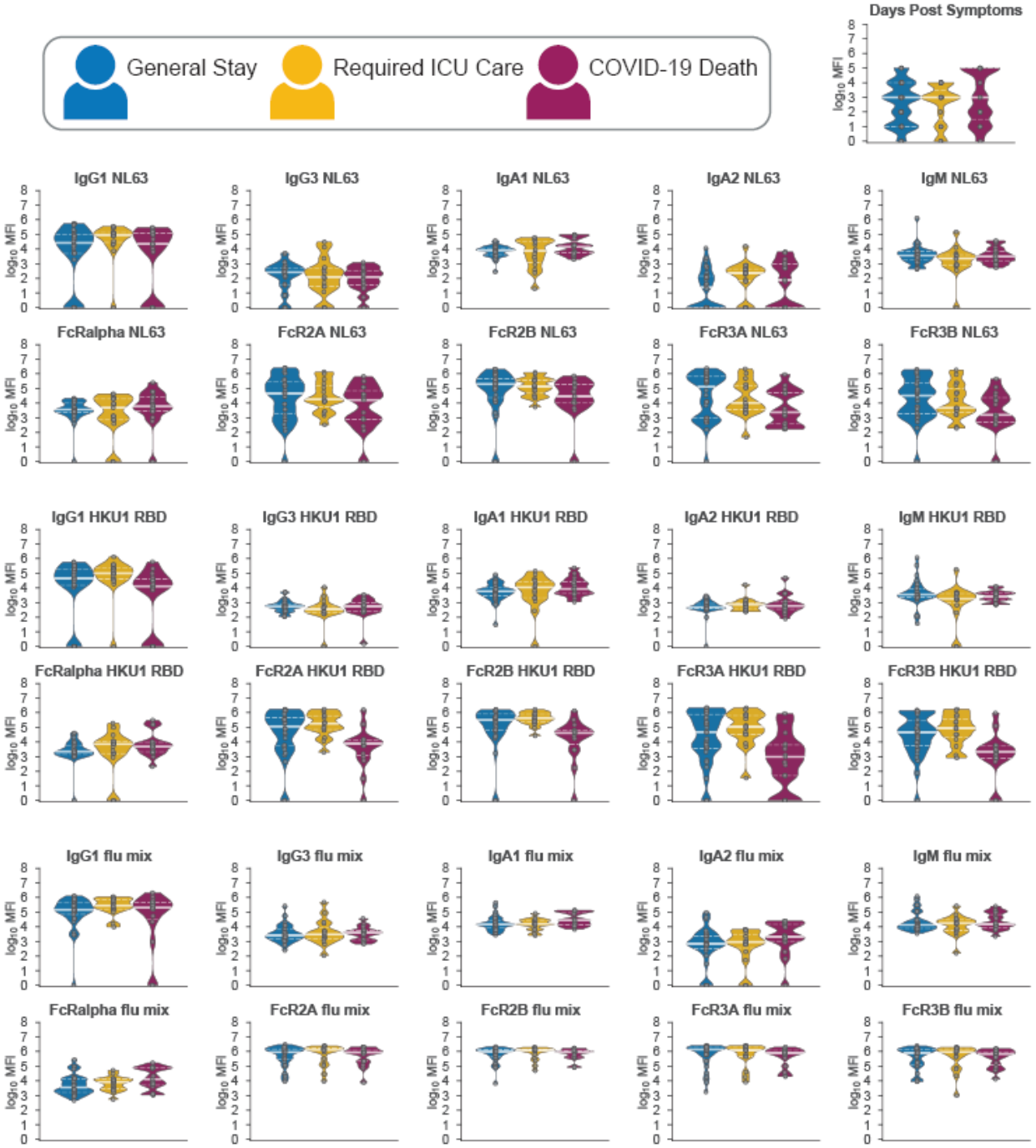


Figure 5-S4. Pre-existing coronavirus immunity, related to Figure 4.

60 plasma samples from hospitalized SARS-CoV-2 infected individuals were profiled against the receptor binding domain of HKU1, NL63, and a mixture of influenza antigens. Distributions of titers across moderate (blue), severe (yellow), and deceased (red) individuals are shown in the violin plot collected within the first five days following symptoms. The solid white line represents the median and the dotted lines the first and third quartiles. A Kruskal-Wallis test was used to evaluate statistical differences across groups for all intervals and features and was corrected for multiple hypothesis testing with the Benjamini-Hochberg procedure. No significant differences were detected.

5.8 Acknowledgements

We thank the SAMANA Kay MGH Research Scholarship, Nancy Zimmerman, Bruce Walker, Mark and Lisa Schwartz, an anonymous donor (financial support), and Terry and Susan Ragon for their support. We would also like to thank Bing Chen for protein production efforts. We acknowledge support from the Ragon Institute of MGH, MIT, the Massachusetts Consortium on Pathogen Readiness (MassCPR), the NIH (3R37AI080289-11S1), the NIAID (U19 AI35995, R37AI80289, R01AI146785), and the U.S. Centers for Disease Control and Prevention (CK000490).

Chapter 6

The Genetic Architecture of Pathogen-Specific Antibody Fc Characteristics and Functions across Twins

The contents of this chapter has been submitted for publication:

Tomer Zohar, Audrey L. Bulter, Tim D. Spector, Douglas Lauffenburger, and Galit Alter. "The Genetic Architecture of Pathogen-Specific Antibody Fc Characteristics and Functions". (Submitted for publication)

Contributions

TZ and GA wrote the manuscript with contributions from all authors. TZ and ALB preformed the quantification of antibody subclasses, isotypes, and Fc-receptors by Luminex. TZ preformed profiling of antibody effector functions. Genotyping, including quality control and imputation was performed by the Twins UK organization. TZ preformed all subsequent analyses.

6.1 Abstract

Antibodies are critically important in protection against disease. Beyond the ability to recognize a myriad of targets, Fc diversification directs immune function, is critically important for combatting infection, and is key for defining mechanisms of protection. While genetic biases in affinity maturation have been demonstrated, little is known about the genetic regulation of Fc biology. Thus, we comprehensively profiled 227 pathogen-specific antibody Fc characteristics and functions to 19 common pathogens and allergens across a cohort of twins (n=510). The majority of subclass, isotype, Fc-receptor binding, and Fc-effector functional responses showed a degree of heritability. GWAS of the top 45 heritable phenotypes identified five genomic regions that were significantly associated to nine humoral traits, and accounted for up to 33% of variance. Subsequent co-expression analysis revealed myeloid cell processes involved in the regulation of Fc-biology. Collectively, these data highlight novel mechanisms that govern Fc-biology and are underscored by host-pathogen interactions.

6.2 Introduction

The humoral immune response has evolved the capacity to uniquely adapt and swiftly combat both environmental and pathogenic threats [242]. This is achieved through the production of vast libraries of polyclonal antibodies, poised to recognize any potential foreign surface and drive rapid immune clearance. Individual antibodies are composed of two regions that during infection mature independently and provide distinct functions. Antigen recognition is mediated by the fragment antigen-binding (Fab) region, diversified through somatic hypermutation, and is capable of binding a near infinite space of targets. Conversely, the fragment crystallizable region (Fc) is diversified through class-switch recombination and post-translational modifications, resulting in distinct constant domains and glycosylation patterns. This alters the affinity of antibodies to Fc receptors which are found on the surface of all immune cells and thus impacts downstream immune function [63]. Thus, antibodies have the ability to not only recognize, but also to regulate inflammation [205], drive cytotoxic destruction or opsonophagocytic clearance of nearly any foreign [209] or altered self-targets [243], and even deliver new-antigens to antigen-presenting cells to promote cellular immunity [244]. Yet, while a great deal is known about the processes that influence binding repertoires, less is known about the biological processes that are involved in regulating Fc-diversification

which is key for the control and clearance of many infectious [6] and non-infectious diseases [245].

Emerging evolutionary genomic studies suggest that selective pressure exerted through host-pathogen interactions have contributed to the enrichment of genes advantageous for reducing pathogen transmissibility across the population [1]. Thus, pathogen-specific humoral immune responses are formed as a result of a combination of both pre-determined genetic and stochastic processes. This is further supported by the enrichment of particular binding repertoires following infection [31] due to preferential usage of gene segments [34] or distance-based biases between V, D, and J segments [33]. Additionally, repertoire twin-based studies have shown similarities in CDR3 lengths among monozygotic twins and have attributed genetic influences to the usage of gene segments, clonal expansion, and clonal sequence diversity [35,246,38,36]. However, the extent of genetic impact and the specific genotypic elements which control isotype and subclass abundance, Fc-receptor binding profiles, and antibody effector functions remain incompletely understood. Ultimately, characterization of humoral genetic architectures could help uncover the mechanisms of Fc-regulation, and aid in next generation vaccine design.

Class switch recombination represents the first key step in tuning antibody effector functions. Specifically, 9 distinct Fc-domains (IgG1-IgG4, IgA1-IgA1, IgM, IgE, and IgD) exist in the human immunoglobulin heavy chain locus (*IGH*) [247]. During B cell differentiation, inflammatory cues lead to genetic excision events that link affinity matured Fab domains to specific Fc-segments [248], each with their own Fc-receptor binding and Fc-effector functional capabilities [249]. For example, Th1 cytokines have been linked to the preferential production of IgG1 and IgG3 subclasses, Th2 cytokine profiles to the production of IgG2 and IgG4 subclasses [250], and mucosal signals to the selection of IgA isotypes [251]. The combinatorial variation in isotypes and subclasses therefore impacts the overall affinity of pathogens coated with antibodies (immune complexes) to the different Fc receptors (FcR) on the surface of innate immune cells. Importantly, because FcγRs bind IgGs with low affinity, triggering multiple FcRs is required to drive immune cell activation and effector function [63]. Emerging data point to the critical role of Fc-effector functions in protecting against a broad array of diseases including human immunodeficiency virus (HIV) [7,9,11], Influenza [12,116], Herpes Simplex Virus [15], and Malaria [17,18]. However, whether Fc-biology is regulated solely via immunological cues, or whether population level genetic variation biases this essential humoral activity remains incompletely understood.

Previous familial-based studies have shown that variation in titers to different childhood vaccines are 36%-91% heritable [252–255], and genome-wide association studies (GWAS) investigating IgG titers against different of infectious diseases have identified multiple strong single nucleotide polymorphisms (SNP) associations [256–259]. Furthermore, robust SNP associations have been identified that drive specific bulk IgG Fc-glycosylation patterns [43]. Collectively, these studies point to a role for genetics in the

regulation of the quantity of the antigen-specific and quality of the bulk IgG humoral immune response. Despite this progress, the genetic contributions and factors which control pathogen-specific antibody isotype and subclass levels, FcR binding, and Fc-effector function remain unknown.

Thus, to begin to define the potential influence of genetics on Fc-biology, we comprehensively profiled 227 humoral traits including isotype and subclass abundance, FcR binding, and antibody mediated effector functions across 18 common viral and bacterial pathogens, and 2 allergens in a cohort of monozygotic and dizygotic twins. Heritability accounted for more than 20% of variation across the majority of humoral traits with IgM levels, antibody dependent neutrophil phagocytosis (ADNP), Varcellia-Zoster (VZV), Measles, and Respiratory Syncytial (RSV) responses showing the highest estimates of heritability. GWAS analysis was performed on 45 of the top heritable humoral traits revealing five genomic regions reaching genome-wide significance across nine humoral traits that explained a median of 30% of the variance. Further transcriptomic analysis of genetic overlap of the hits to previous studies pointed to a critical role in myeloid cell activity as a key regulator of antibody Fc-biology. Ultimately, these data provide a resource for understanding the genetic mechanisms underlying humoral immune diversification.

6.3 Results

6.3.1 Distinct heterogeneity in pathogen-specific humoral characteristics

Antibodies are primary correlates of both naturally occurring and vaccine-induced immunity against many infectious diseases [6]. However, antibodies also represent critical biomarkers of disease activity [260,261], and have been implicated in both protective and pathological consequences across autoimmune [245] and allergic diseases [262]. Previous familial-based studies have found that variation in IgG titers to common infectious diseases [263] and childhood vaccines are heritable [252–255]. Moreover, subsequent GWAS analyses have identified multiple SNPs across many genes including HLA-DQ, HLA-DR, and STING1 that were strongly associated with IgG1 levels to the Epstein-Barr virus (EBV), Varicella zoster virus (VZV), and more [256–259]. However, beyond titers, variation in antibody isotypes, subclasses, Fc-receptor (FcR) binding, and Fc-effector functions also contribute to humoral protection [7,17,13]. Yet, whether genetic factors also control pathogen-specific Fc-biology remains largely unknown.

Thus, we comprehensively profiled pathogen-specific humoral responses across a cohort of 168 monozygotic (MZ) and 342 dizygotic (DZ) twins (Figure 6-1A). All participants were female, between the ages of 41-78 (mean 61), of British ancestry, and 100% Caucasian. Pathogen-specific isotype and subclass levels, binding to FcRs, and effector functions were first systematically measured against 9 viral and 7 bacterial pathogens, and 2 allergens (Table 6-S1). These pathogens and allergens were selected based on the probability of exposure either through infection or vaccination to the participants in this cohort. Protein antigens, polysaccharide antigens, or whole

pathogens were chosen based on previously reported immunodominance or availability. Additionally, Influenza A virus subtypes H1N1 and H3N2, and prefusion and postfusion conformational states of the respiratory syncytial virus (RSV) fusion (F) glycoprotein were also included. On the population scale, a diversity of responses was observed to almost all humoral features (Figure 6-1B). These data collectively provided a humoral landscape that could be used to identify population-level pathogen-specific antibody characteristics and determine the genetic factors responsible for their control.

To first gain a deeper understanding of immune programming across antigen-targets and Fc-readouts, we next examined the mean pathogen-specific isotype and subclass levels, binding to FcRs, and effector functions across the whole cohort (Figure 6-1C). Antigen-specific results for each readout were min-max normalized along feature types. With the exception of allergens Arah2 and Betv1, and Ttox robust IgG1 responses were observed across all pathogens. IgG2, IgG3, and IgG4 levels were more pathogen specific than IgG1 with the strongest mean IgG2, IgG3, IgG4 responses observed against Dip, Polio and Post-RSV, and HBV respectively. Additionally, IgM and IgA also exhibited pathogen specific behavior with the strongest mean responses observed against Dip and Polio, and Polio and Post-RSV respectively. Enhanced binding to FcγR2A followed a similar profile to IgG1 responses. However, FcγR2b, FcγR3A, and FcγR3b showed more nuanced pathogen-specific profiles.

Differences in FcR binding results in alterations in the capacity to drive antibody Fc-effector functions [63]. Thus, population-level mean antibody-dependent effector functions were probed (Figure 6-1C). Antibody-dependent complement deposition (ADCD) was observed across most pathogen-specific target antigens. Conversely, mean antibody-dependent cellular phagocytosis (ADCP) and antibody-dependent neutrophil phagocytosis (ADNP) responses were more pathogen specific. In particular, the strongest mean ADCP responses were observed against Flu and RSV, and the strongest mean ADNP responses against Measles. Collectively, these data highlight striking variation in antibody functionality across antigen-specificities pointing to immunologic functional programming across targets.

Finally, to determine whether variation in antibody responses was driven by potential confounding variables, a principal components analysis (PCA) was performed using all humoral data and was assessed for twin zygosity type, age, alcohol use, and smoking (Figure 6-1D and Figure 6-S1). This revealed limited variation driven by the twin sets (Figure 6-1D) and demographics (Figure 6-S1), pointing to a limited impact of these confounders on polyclonal antibody Fc-profiles.

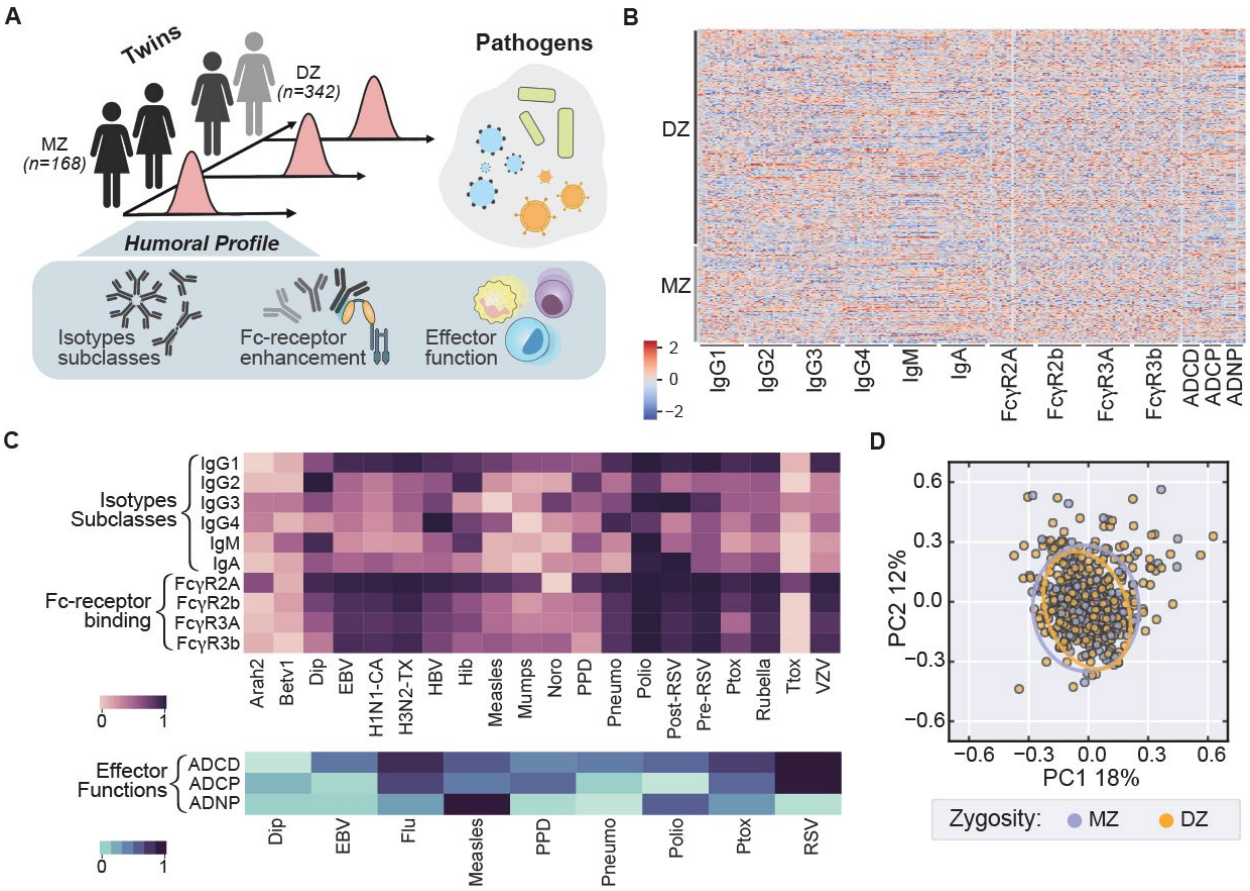


Figure 6-1. Comprehensive Profiling of Pathogen-specific Antibody Features and Population-level Responses.

(A) Pathogen-specific humoral responses to a myriad of viral pathogens, bacterial pathogens, and allergens were profiled in a cohort of 510 individuals. The cohort consisted of 168 monozygotic (MZ) and 342 dizygotic (DZ) matched twins. Humoral profiling of plasma samples consisted of measuring pathogen-specific antibody isotype and subclass levels, levels of binding to Fcγ-receptors (FcγR), and effector functions. (B) Normalized antibody responses across MZ and DZ individuals organized by pairs and antibody characteristics with a repeating sequence of pathogen specificities. Normalization was achieved by first subtracting technical noise followed by rank-based inverse normal transformation. (C) Mean population-level humoral responses to pathogen specificities. Values were determined by first evaluating the normalized mean of each antibody feature followed by min-max normalization along antibody characteristics. (D) Principal components analysis (PCA) scores plot built on normalized humoral features and split by twin zygosity. Ellipses correspond to the 95% confidence intervals for each group. (Betv1: Pollen, Arah2: Peanut, Hib: Haemophilius Influenzae type b, Ptox: Pertussis, Ttox: Tetanus, PPD: Tuberculosis, Pneumo: Streptococcus pneumoniae, Dip: Diphtheria, HBV: Hepatitis B, Noro: Norovirus, Polio: Poliomyelitis, Pre-RSV: Respiratory Syncytial Virus A/B Prefusion F glycoprotein, Post-RSV: Respiratory Syncytial Virus A/B Postfusion F glycoprotein, VZV: Varicella-Zoster virus, EBV: Epstein-Barr Virus, H1N1-CA: Influenza A virus CA/7/2009, H3N2-TX: Influenza A virus Texas/7/2012). Also see Table 6-S1 and Figure 6-S1.

6.3.2 Coordination in Fc-profiles are driven by antigen-specificity and Fc-characteristics

Given the presence of distinct Fc-profiles across antigens, we next aimed to determine whether any coordination existed among humoral features. Pairwise spearman correlations revealed strong relationships within subclasses and isotypes, pathogen FcR binding, and between IgG1 levels and FcR binding (Figure 6-2A). Specifically, cross-antigen relationships were observed within isotypes and subclasses. Furthermore, strong coordination among IgM responses were observed, and pointed to a broad intra-individual bias towards IgM selection. Similar relationships were also observed for other isotypes and subclasses, albeit these relationships were less robust. Conversely, correlations across FcRs or between FcRs and IgG levels were also observed but to specific antigen-specificities. To investigate whether monozygotic and dizygotic twins differed in their coordination to specific features this analysis was extended to each group, and as expected revealed no notable differences (Figure 6-S2). To further quantify differences in the inter-relatedness of antibody characteristics and pathogens specificities, distributions of significant spearman correlation coefficients across Fc-characteristics (Figure 6-2B, top) and antigen-specific antibody features were compared (Figure 6-2B, bottom). Among these readouts, the highest level of coordination was observed across IgM responses and the lowest to FcR-binding levels (Figure 6-2B, top). From an antigen-specific perspective, Polio-specific responses were least coordinated, and Mumps-specific responses the most.

Finer detail regarding the coordination of antigen-specificities and Fc-features was next investigated through hierarchical clustering (Figure 6-2C). This revealed 20 distinct clusters that were further binned into two larger clusters. The first set of clusters (left) included largely individual antigen-specificities composed of multiple Fc-features, and pointed to coordination among antibody isotype, subclass, and FcR binding levels across RSV, Betv1, Arah2, VZV, Noro, Polio, Measles, Mumps specific antibody responses. Therefore, this suggests that individuals with robust IgG responses to these targets also generated robust FcR binding profiles and functions. The second large cluster (right) was composed of groups of antibody Fc-feature driven profiles. These clusters consisted of distinct IgM, IgA, IgG2, IgG3, and IgG4 clusters. These data ultimately point to coordination across both antigen-specificity and Fc-feature.

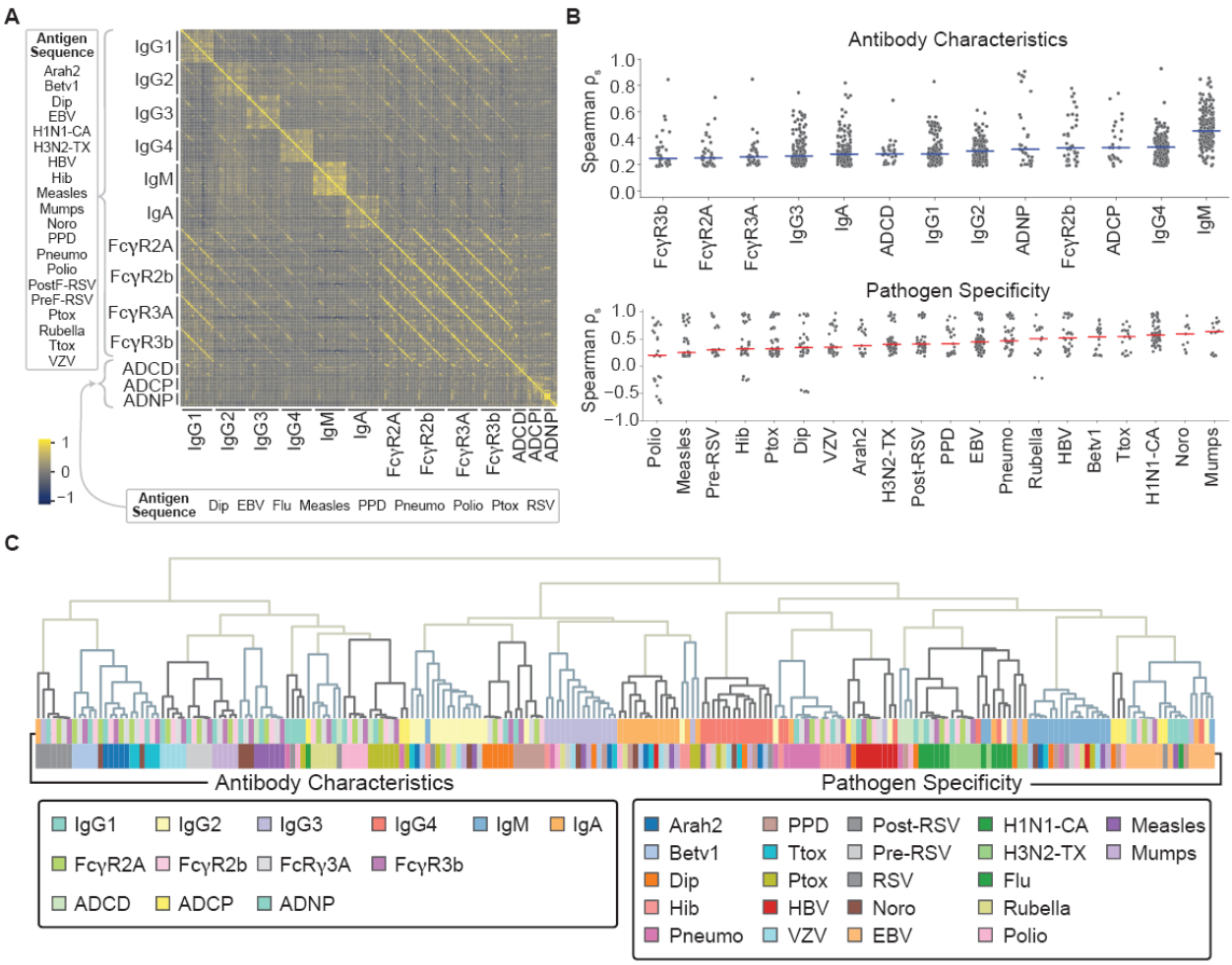


Figure 6-2. Coordination of Antibody Characteristics and Pathogen Specificities

(A) Pairwise spearman correlations organized by type of antibody measurement and with a repeating antigen sequence. (B) Distributions depict significant spearman correlation coefficients (ρs) across antibody characteristics (top) and pathogen specificities (bottom). Bars correspond to the median of each distribution. (C) The dendrogram depicts hierarchical clustering of the pathogen-specific antibody correlations. Agglomerative hierarchal clustering was performed with pairwise distances determined by correlations and 'complete' linkage method. Clusters were defined with a threshold based on half maximum pairwise distance among features and are represented by alternating grey and blue colors. The identity of the antibody characteristics and pathogen specificities in each cluster are summarized by the colors of the bars. Also see Figure 6-S2.

6.3.3 Genetic factors contribute to variation in humoral responses

The architecture observed to the humoral immune response within antigen-specificities and Fc-characteristics across this cohort highlights the potential influence of non-stochastic driving forces. Thus, the familial structure among twin pairs was leveraged to determine the impact of genetics on the diversification of particular antigen-specific antibody profiles. Similar to previous work demonstrating heritability in IgG1-titers [263], here we profiled the impact of heritability on all pathogen-specific Fc-features. The heritability of each feature was estimated using linear mixed effect models, where we

were able to decomposed the variance due to additive genetics (A), common shared environment (C), and unique environment (E) [264]. Heritability was defined as the variance due to additive genetics as a fraction of the total variance, ranging from 0%-suggestive of non-genetic factors to 100%-suggestive of strong genetic influence. Goodness-of-fit was assessed using the Akaike Information Criteria (AIC) [238] and comparison among models with and without genetic parameters (ΔAIC) was used as a measure of confidence, with ΔAIC>0 providing evidence for difference among these models and therefore justifying the existence of the genetic parameter.

Heritability was observed across a number of features including among Fc-characteristics and pathogen-specificities (Figure 6-3A). IgM levels showed the greatest number of heritable features with ΔAIC>0, followed by IgG1 levels and FcRs. In terms of individual features Ptox-IgG1, VZV-IgG4, and PostF-RSV FcγR2b binding levels showed the greatest levels of heritability. Similarly, evidence for heritability was also observed across antibody effector functions, with the highest levels noted among PPD-ADNP and RSV-ADCP levels with estimates of 61% and 68%, respectively. Moreover, EBV and Flu specific ADNP also exhibited ΔAIC levels above 5 (Figure 6-3A and 6-3B). As for pathogen-specificities, high estimates were also observed across FcγRs in the case of Ptox, Measles, and PostF-RSV specific antibodies. Additionally, responses against Measles showed the highest heritability in isotype, subclass, and FcR binding across all pathogens followed by Ptox-, Dip-, PostF-RSV, and VZV-specific antibody responses. Taking a closer look at the distribution of heritability estimates highlighted that the majority of features were more than 20% heritable and exhibited estimates as high as 68% (Figure 6-3C, left). Furthermore, comparing the distributions of the variance captured by each parameter across all features revealed that overall additive genetics accounted for more variance than common environment, albeit not as much as the unique environment which represents exposures that are specific each to individual (Figure 6-3C, right). Collectively, these data highlight the influence genetics has in determining Fc-biology.

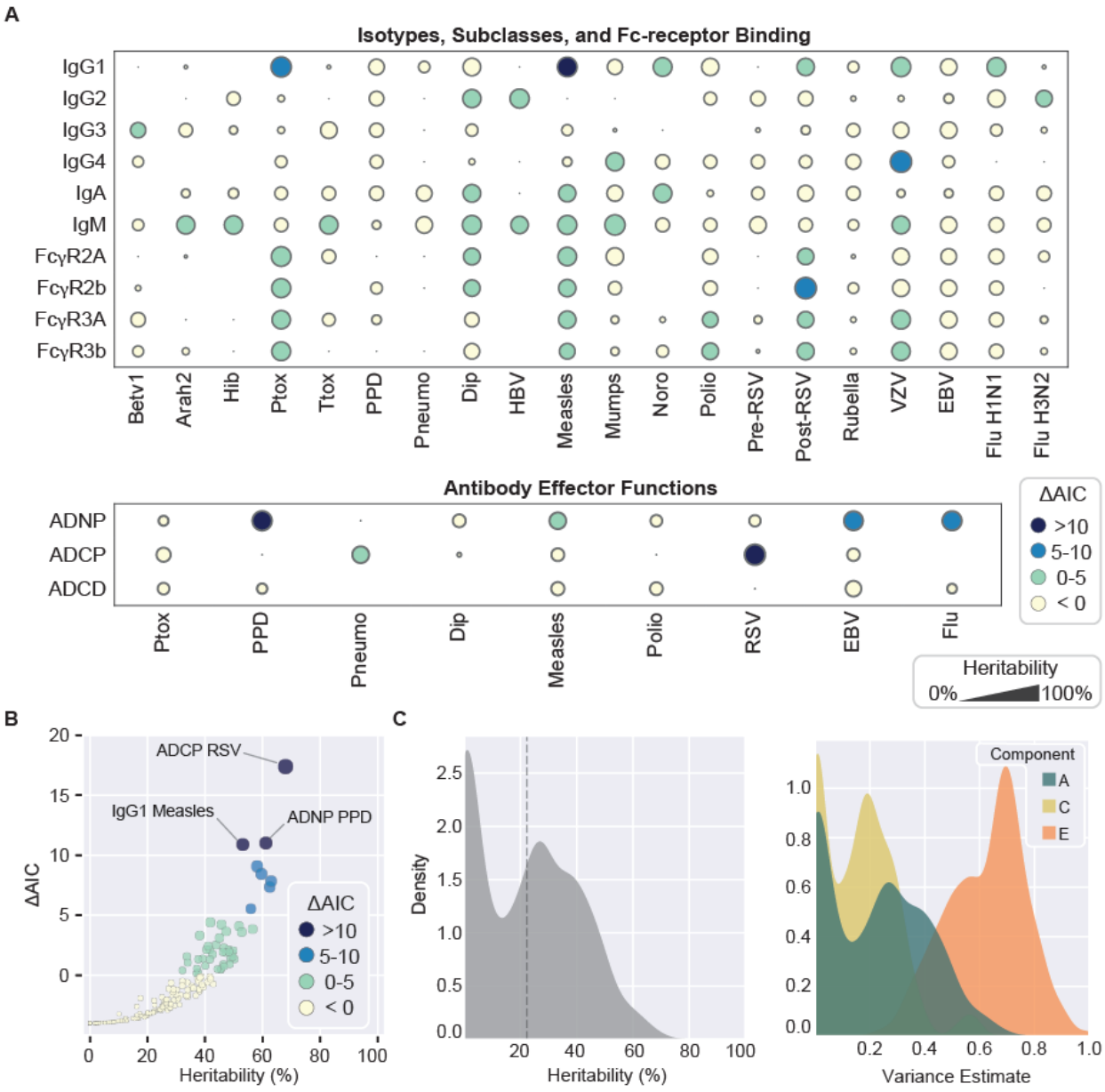


Figure 6-3. Heritability of Pathogen-specific Humoral Phenotypes

(A) Dot plot of heritability estimates for each pathogen-specific antibody measurement. Color corresponds to the ΔAIC and size to the level of heritable influence. (B) Scatter plot depicting the heritability and ΔAIC values of each feature with color corresponding to ΔAIC. (C) The probability density distribution of heritable influences across all measurements and the mean heritability threshold of 0.22. (D) Probability density distributions of the variance explained by each component of the ACE model across all measurements. (A, additive genetics; C, common shared environment; E, unique environment).

6.3.4 Particular classes of antibody features are under greater genetic control

To quantitatively probe global differences in heritable contributions to individual classes of antibody features, across both Fc-characteristics (Figure 6-4A, left) and antigen-specificities (Figure 6-4A, right), non-parametric combination [58] was adapted to the

linear mixed effect model framework. Briefly, by permuting the zygosity labels among twin pairs and calculating the mean heritability estimate for a specific feature type over multiple iterations, a null distribution is formed. Then exact p-values were determined by comparing the true mean heritability estimates within the control distributions. A clear hierarchy appeared in mean heritability across feature types, with FcγR3A binding levels (Figure 6-4A, left) and measles-specific responses (Figure 6-4A, right) exhibiting the highest mean heritability estimates. Conversely, IgM-, ADNP-, and IgG1- levels showed the greatest effect sizes (Figure 6-4A, left). Similarly, antibody responses to VZV, Measles, and PostF-RSV also exhibited high mean heritability estimates, but with reduced levels of significance. Moreover, several additional features, including overall ADCD levels, and Norovirus-, PreF-RSV, Ttox-, H3N2-TX-, Rubella-, Betv1-, Pneumo-, Hib-, Arah2-, and HBV-specific responses all showed low mean heritability estimates. Collectively, Fc-profiles exhibited more evidence for heritability than pathogen specificities.

Given the differences observed across antibody characteristics and pathogen specificities, broader classes of humoral features were explored. First comparing isotype and subclass levels, binding to FcRs, and antibody mediated effector functions revealed that only isotype and subclass levels were globally significantly heritable (Figure 6-4B). Additionally, antibody effector functions showed the greatest mean heritability and was nearly significant. Similarly, consolidation of pathogen-specificities based on viral or bacterial antibody targets pointed to greater mean heritability among viral-specific responses, and with both viral and bacterial responses exhibiting significance compared to their null distributions (Figure 6-4C). Finally, differences in mean heritability were calculated for viruses experienced by infection alone (Viral non-vaccine) and those for which vaccines are used routinely (Viral vaccine). This analysis revealed that antigen-specific antibody responses to viruses that are largely vaccinated against exhibited significant and high mean heritability (Figure 6-4D). Ultimately, these data highlight the specific classes of humoral traits that are most heavily influenced by genetics.

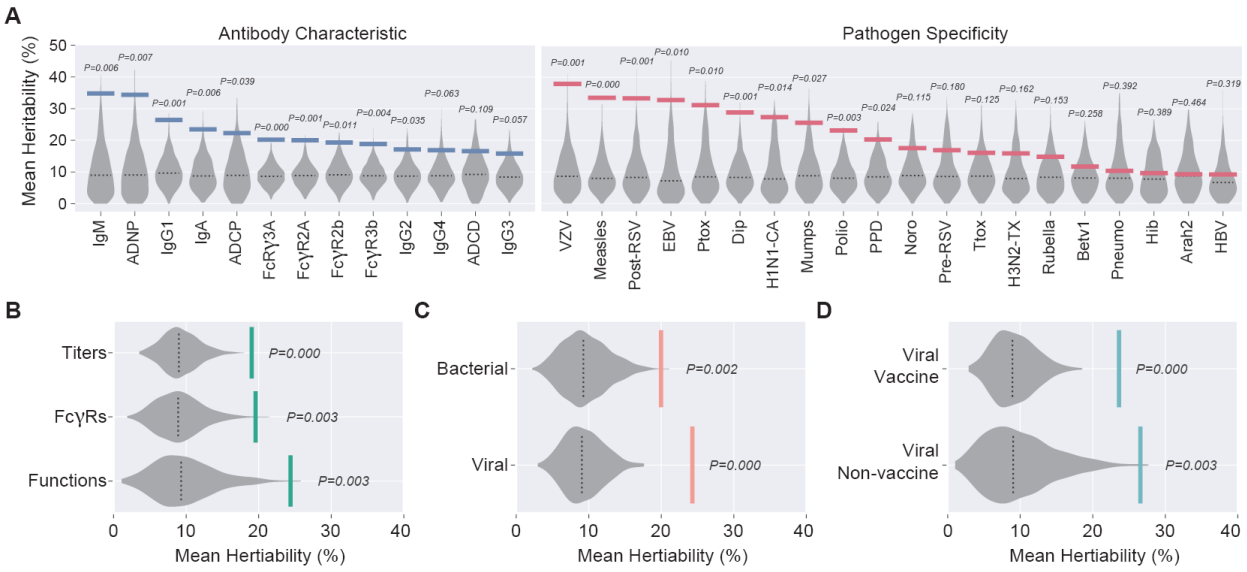


Figure 6-4 Broad Heritable Influences Amongst Specific Antibody Attributes

Mean heritability estimates among (A) antibody characteristics and pathogen specificities, and (B-D) higher-level classes. Violin plots correspond to null distributions determined by non-parametric combination, bars correspond true mean heritability estimates, and exact P values are reported.

6.3.5 Genome-wide association reveals genetic control of pathogen-specific Fc profiles

To determine the specific genetic factors that control the development of particular pathogen-specific Fc-profiles that showed heritability estimates greater than 30% and ΔAIC>0, consecutive GWASs were performed. This included GWASs for 45 antibody features, comprised of 21 isotype and subclass features, 18 FcR-binding features, and 6 effector functions across 16 different pathogens. Following quality control, 4,663,961 variants and 498 individuals were used to screen for associations. This yielded five regions which reached genome-wide significance ($p < 5 \times 10^{-8}$) across nine humoral traits (Figure 6-5A), all of which have not been previously reported. Lead variants had minor allele frequencies (MAF) between 19% and 49% which explained 25% to 33% of the variance among these traits (Table 6-1). Additionally, Post-RSV specific IgG1 levels and FcγR2A-, FcγR2b-, FcγR3A-, and FcγR3b-binding levels all shared the same leading variant, rs679574. Further analyses identified candidate genes for each of the variants [265,266] (Table 6-1). For each lead variant, nearby SNPs were further investigated (Figure 6-5B-K) as well as associations across all select humoral traits (Figure 6-5L).

Single nucleotide polymorphism (SNP) rs62493093T was associated with significantly higher IgG2 levels against HBV (Figure 6-5B). Furthermore, the number of rs62493093T alleles were proportional to IgG2 HBV-specific levels indicative of a dose-dependent response in the presence of this SNP. This variant was nearest to the *DLC1* gene encoding for a Rho GTPase activating protein (Figure 6-5C). Moreover, nominal associations and trends were noted between this HBV-specific rs62493093 SNP and

Measles FcR-binding, Measle-specific ADNP, and RSV-specific ADCP levels suggesting that this SNP may have a broader impact on shaping humoral immune responses across pathogens (Figure 6-5L).

Variant rs10770123C was associated in a dose-dependent manner with FcγR3A-binding levels against Polio and was located in the 3' UTR region of *RNF141* (Figure 6-5D-E). This variant was also in high linkage disequilibrium with other genome-wide significant intron SNPs contained within *RNF141* (Figure 6-5E and Table 6-S2). These variants have also been previously implicated as expression quantitative trait loci (eQTLs), influencing the expression of *RNF141* among blood cells [267] and lymphoblastoid cell lines (LCL) [268], as well as the expression of *MRVI1* and *MRVI1-AS1* in monocytes [269] and neutrophils [270] respectively. Additionally, this SNP was also suggestively associated with Polio-specific FcγR3b binding levels and weakly associated with VZV-specific IgG1, FcγR3b, and FcγR3A (Figure 6-5L). Associations in the opposite direction were also interestingly noted across multiple IgM specific responses.

A significant inverse dose-dependent association was observed between rs1806595T and Flu-specific ADNP levels (Figure 6-5F). This variant was located nearest to the *SPRY2* gene on chromosome 13 and was in high linkage disequilibrium with multiple variants including rs2379 (Figure 6-5G). Although located in an intergenic region, variant rs2379 has previously been identified as a CCCTC-binding factor (CTCF) binding site across many tissues including immune compartments [271] (ENCSR756ZKG). CTCF plays a key role in regulating chromatin structure and maintaining organization on the megabase-scale [272], therefore variants in CTCF binding regions represent potential distal regulatory elements. Moreover, beyond Flu-specific ADNP, the rs1806595 SNP additionally showed a trend to associations with EBV-, PPD-, and Measles-specific ADNP, pointing to a cross-pathogen specific modulation of antibody qualities involved in controlling neutrophil activity (Figure 6-5L).

Subsequently, variant rs10131070G showed a strong association with Mumps-specific IgG4 levels, with higher levels corresponding to the major allele and a complete loss of Mumps-specific IgG4 responses in homozygotes with the rs10131070G minor allele (Figure 6-5H). This SNP was in high linkage disequilibrium with rs10131536 and rs11157248, all of which are located in the T cell receptor alpha locus (*TRA*) (Figure 6-5I). These variants are additionally all located in sites known to be epigenetically modified across various cell types [273]. For example, variant rs10131536 is located at a site of h3k4me1 in T helper cells. These data point to a critical potential role of this SNP in T helper cell function that may bias Mumps and possibly VZV antibodies towards IgG4 (Figure 6-5L).

The strongest associations identified in this study were between variant rs679574 and postfusion F RSV (Post-RSV) specific IgG1, FcγR2A, FcγR2b, FcγR3A, and FcγR3b binding levels (Figure 6-5J). The presence of the minor allele rs679574C, whether heterozygotic or homozygotic, resulted in substantially higher levels of IgG1

and binding to FcRs against the Post-F antigen from RSV (Figure 6-5K). Variant rs679574 was also in perfect linkage disequilibrium with two synonymous variants rs492602 (A68A) and rs681343 (Y83Y) located in exon 2 of *FUT2*, and a nonsense variant rs601338 (W154X) (Figure 6-5K). Variant rs601338A encodes for "non-secretor" status in homozygotes with both null alleles. This results in no functional FUT2 enzyme production and therefore the lack of histo-blood group antigens (HBGA) on mucosal surfaces [274]. Additionally, nominal associations between these SNPs and Ptox specific IgG1 and FcR bindng, Penumo-ADCP, and Measles and VZV-specific IgM levels were also observed but with opposite effects across all *FUT2* variants (Figure 6-5L). These data suggest that these *FUT2* SNPs may have broad implications related to humoral function but with pathogen specific directionality.

To further define the mechanistic basis for SNP mediated antibody Fc-modulation, we next used gene set analysis and Gene Ontology (GO) to identify the pathways in which SNPs were enriched in. Subprocesses of the immune response (GO:0002376) were tested and revealed multiple gene sets that were significantly associated to 16 heritable features. The top significant processes were related to B cell, T cell, and innate immune function (Figure 6-5M and Table 6-S3). Specifically, PPD- and EBV-specific ADNP were linked to lymph node development and germinal center formation whereas Measles-specific IgG1, FcγR2A, and FcγR2b binding levels were associated with T cell proliferation and differentiation. Moreover, although no SNPs that reached genome wide significance to PPD-specific ADNP and Measles-specific IgG1 levels were observed (Figure 6-5A), significant immune processes were detected (Figure 6-5M). Thus, collectively, this genomic screen identified multiple genetic factors which contributed to the modulation of pathogen-specific Fc-biology. These genetic factors were related to both regulatory control and alterations in immune genes that likely influence both adaptive and immune programming.

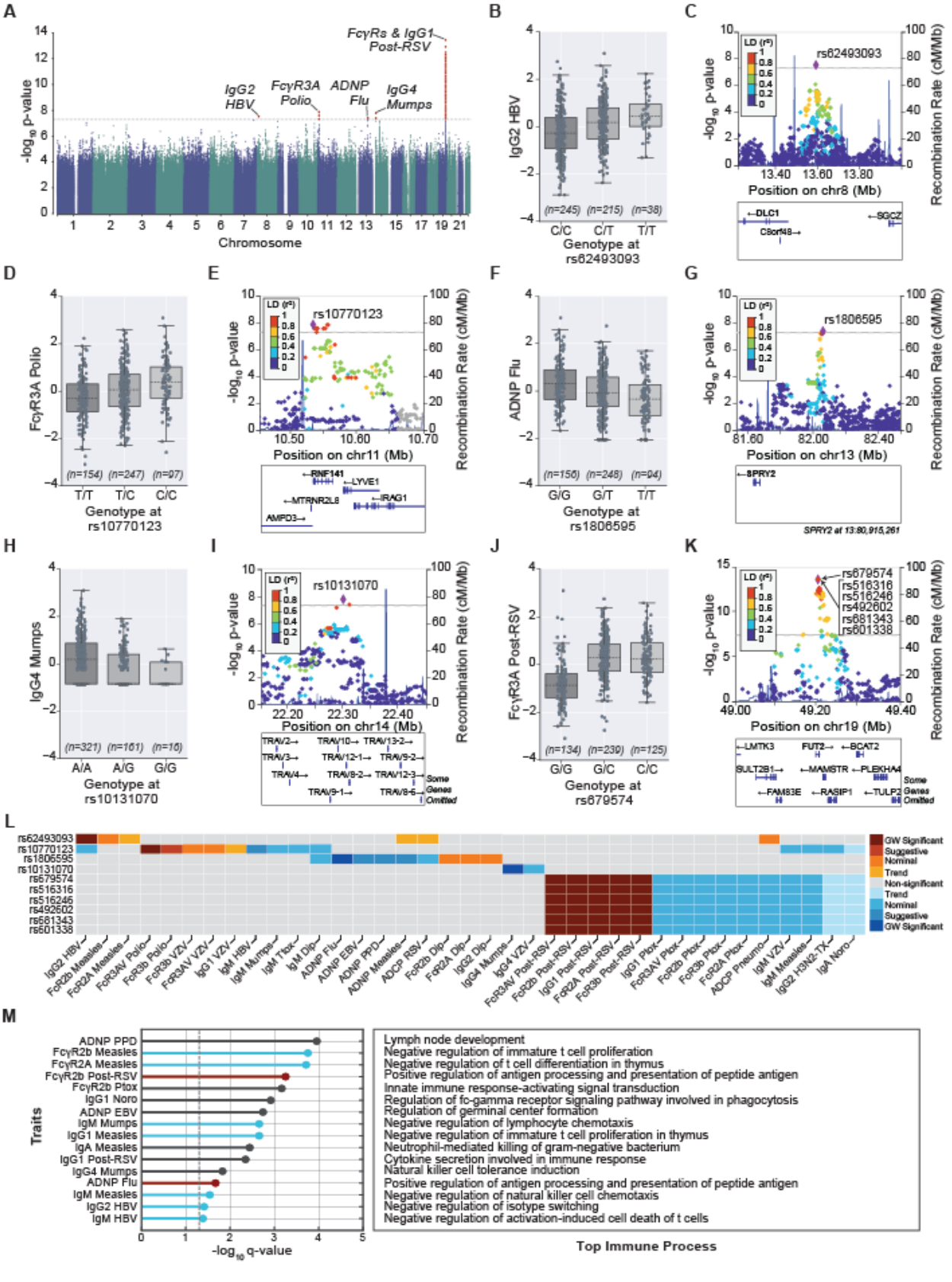

A
FcγRs & IgG1 Post-RSV
IgG2 HBV
FcγR3A Polio
ADNP Flu
IgG4 Mumps
-log10 p-value
Chromosome
B
IgG2 HBV
(n=245) (n=215) (n=38)
C/C C/T T/T
Genotype at rs62493093
C
rs62493093
LD (r²)
-log10 p-value
Recombination Rate (cM/Mb)
Position on chr8 (Mb)
DLC1
SGCZ
C8orf48
D
FcγR3A Polio
(n=154) (n=247) (n=97)
T/T T/C C/C
Genotype at rs10770123
E
rs10770123
Position on chr11 (Mb)
RNF141
LYVE1
MTRNR2L8
IRAG1
AMPD3
F
ADNP Flu
(n=156) (n=248) (n=94)
G/G G/T T/T
Genotype at rs1806595
G
rs1806595
Position on chr13 (Mb)
SPRY2
SPRY2 at 13:80,915,261
H
IgG4 Mumps
(n=321) (n=161) (n=16)
A/A A/G G/G
Genotype at rs10131070
I
rs10131070
Position on chr14 (Mb)
TRAV2 TRAV10 TRAV13-2
TRAV3 TRAV12-1 TRAV9-2
TRAV4 TRAV8-2 TRAV12-3
TRAV9-1 TRAV8-6
Some Genes Omitted
J
FcγR3A Post-RSV
(n=134) (n=239) (n=125)
G/G G/C C/C
Genotype at rs679574
K
rs679574
rs516316
rs516246
rs492602
rs681343
rs601338
Position on chr19 (Mb)
LMTK3 FUT2 BCAT2
SULT2B1 MAMSTR PLEKHA4
FAM83E RASIP1 TULP2
Some Genes Omitted
L
rs62493093
rs10770123
rs1806595
rs10131070
rs679574
rs516316
rs516246
rs492602
rs681343
rs601338
GW Significant
Suggestive
Nominal
Trend
Non-significant
IgG2 HBV
FcγR2b Measles
FcγR2A Measles
FcγR3AV Polio
FcγR3b Polio
FcγR3b VZV
FcγR3AV VZV
IgG1 VZV
IgM HBV
IgM Mumps
IgM Tox
IgM Dip
ADNP Flu
ADNP EBV
ADNP PPD
ADNP Measles
ADCP RSV
FcγR2b Dip
FcγR2A Dip
IgG2 Dip
IgG4 Mumps
IgG4 VZV
FcγR3AV Post-RSV
FcγR2b Post-RSV
IgG1 Post-RSV
FcγR2A Post-RSV
FcγR3b Post-RSV
IgG1 Ptox
FcγR3AV Ptox
FcγR2b Ptox
FcγR3b Ptox
FcγR2A Ptox
ADCP Pneumo
IgM VZV
IgM Measles
IgG2 H3N2-TX
IgA Noro
M
Traits
ADNP PPD
FcγR2b Measles
FcγR2A Measles
FcγR2b Post-RSV
FcγR2b Ptox
IgG1 Noro
ADNP EBV
IgM Mumps
IgG1 Measles
IgA Measles
IgG1 Post-RSV
IgG4 Mumps
ADNP Flu
IgM Measles
IgG2 HBV
IgM HBV
-log10 q-value
Lymph node development
Negative regulation of immature t cell proliferation
Negative regulation of t cell differentiation in thymus
Positive regulation of antigen processing and presentation of peptide antigen
Innate immune response-activating signal transduction
Regulation of fc-gamma receptor signaling pathway involved in phagocytosis
Regulation of germinal center formation
Negative regulation of lymphocyte chemotaxis
Negative regulation of immature t cell proliferation in thymus
Neutrophil-mediated killing of gram-negative bacterium
Cytokine secretion involved in immune response
Natural killer cell tolerance induction
Positive regulation of antigen processing and presentation of peptide antigen
Negative regulation of natural killer cell chemotaxis
Negative regulation of isotype switching
Negative regulation of activation-induced cell death of t cells
Top Immune Process

Figure 6-5. Genome-wide Associations and Processes Involved in Regulating Pathogen-specific Antibody Fc Phenotypes

(A) Summary Manhattan plot demonstrating five regions which reach genome-wide significance and the traits which they are associated with. Variants which cross genome-wide significance ($p < 5 \times 10\text{-}8$) are depicted in red. Genotype-phenotype relationships and regional association plots were then depicted for IgG2 against HBV (B and C), FcγR3A against Polio (D and E), ADNP against Flu (F and G), IgG4 against Mumps (H and I), and FcγR3A against Post-RSV (J and K). The directionality of genotype-phenotype associations correspond to the presence of minor alleles and the number of each set of alleles are reported. The dotted line represents the median and the error bars one standard deviation. Regional plots depict chromosomal position on the x-axis based on GRCh37, negative log10 p-values on the left y-axis, recombination rates on the right y-axis, and RefSeq genes on the bottom. The horizontal line depicts genome-wide significance ($p < 5 \times 10\text{-}8$), the purple diamond the lead variant, and colors to the linkage disequilibrium based on the 1000 Genomes November 2014 European (EUR) database. (L) Levels of association of lead variants to other heritable features. Red colors indicate positive associations toward minor alleles and blue colors indictive negative associations toward major alleles. The intensity of the colors indicates the degree of confidence and ranges from genome-wide significance ($p < 5 \times 10\text{-}8$), suggestive associations ($p < 1 \times 10\text{-}4$), nominal associations ($p < 0.05$), and trends ($p < 0.1$). (M) Top immune related process of heritable traits with at least significant pathway. On the left a lollipop plot with negative log10 q-values corresponding to each trait is depicted. Gene set enrichment analysis was performed using gene ontology (GO) and gene set p-values across all traits were multiple hypothesis corrected using the Benjamini-Hochberg procedure with a false discovery rate of ($\alpha = 0.05$). The vertical dashed line depicts the significance cutoff, and the color of the lollipop corresponds to whether the GO term is a general process (black) or involved in negative (blue) or positive (red) regulation of a process. The corresponding GO terms are shown on the right. Also see Table S3. Also see Table 6-S2 and Table 6-S3.

Chr	Pos	rsID	EA	NEA	EAF	BETA (SE)	% Var	p value	Trait	Gene	Annotation
8	13594961	rs62493093	T	C	0.29	0.27 (0.05)	24.79	2.97×10^{-8}	IgG2 HBV	*DCL1*	*N, T, S*
11	10534866	rs10770123	C	T	0.44	0.27 (0.05)	25.42	1.28×10^{-8}	FcγR3A Polio	*RNF141*	*G, Q, T, S*
13	82061839	rs1806595	T	G	0.44	-0.27 (0.05)	24.58	3.92×10^{-8}	ADNP Flu	*SPRY2*	*N, R, T, S*
14	22301854	rs10131070	G	A	0.19	-0.25 (0.04)	25.23	1.65×10^{-8}	IgG4 Mumps	*TRA cluster*	*G, R*
19	49206108	rs679574	C	G	0.49	0.37 (0.05)	33.42	3.58×10^{-14}	FcγR3A Post-RSV	*FUT2*	*C, Q, S*
19	49206108	rs679574	C	G	0.49	0.36 (0.05)	32.77	1.19×10^{-13}	FcγR3b Post-RSV	*FUT2*	*C, Q, S*
19	49206108	rs679574	C	G	0.49	0.35 (0.05)	32.02	4.51×10^{-13}	IgG1 Post-RSV	*FUT2*	*C, Q, S*
19	49206108	rs679574	C	G	0.49	0.35 (0.05)	31.03	2.51×10^{-12}	FcγR2b Post-RSV	*FUT2*	*C, Q, S*
19	49206108	rs679574	C	G	0.49	0.33 (0.05)	30.36	7.64×10^{-12}	FcγR2A Post-RSV	*FUT2*	*C, Q, S*

Table 6-1. Novel Significant SNPs Associated with Humoral Traits

Lead variants associated to humoral traits which reach genome-wide significance threshold of ($p < 5 \times 10\text{-}8$). Annotations were based on lead variant or variants in linkage disequilibrium with the lead variant ($r2 > 0.8$). These variants were located in the coding regions (C), contained with the gene (G), or were upstream or downstream of a particular gene (N). Whether variants corresponded to previously reported eQTLs (Q), disrupted a regulatory motif (R), was supported by transcriptomics data (T), or was supported by previous research (S) was also reported. (Chr, chromosome; Pos, position in GRCh37; rsID, reference SNP identification; EA, effect allele; NEA, non-effect allele; EAF, effect allele frequency; BETA, weight; SE, standard error; % Var, percent variation explained)

6.3.6 Transcriptomic and pathological consequences of genotype associations to humoral phenotypes

To further explore potential mechanisms related to SNP mediated Fc-biological programming, target gene expression was obtained from single cell transcriptomics data [275]. Expression of candidate genes was first analyzed across immune cell types, revealing robust expression of *DLC1*, *RNF141* and *SPRY2* among common myeloid progenitors (CMP) (**Figure 6-6A**). Specifically, expression of *DLC1* was mainly observed in CMP cells whereas *RNF141* expression was observed broadly among cells of myeloid lineage, and expression of *SPRY2* was restricted to CMP and myeloid cells. Conversely, expression of gene *TRAV8-2,* within the *TRA* cluster, was observed in a subset of regulatory T cells and as expected, *FUT2,* showed no expression among any immune cell types.

In order to investigate the processes which the candidate genes *DLC1*, *RNF141, and SPRY2* influenced, co-expression of transcripts were evaluated among the cell types that showed robust expression of each candidate gene respectively. Co-expressed genes were then attributed to biological processes using gene list enrichment analysis and the gene ontology (GO) database. For example, genes co-expressed with *DLC1*, previously associated with HBV-specific IgG2 levels, were restricted to those with significant correlations ($|\rho| > 0.3$) within CMPs, and genes with the strongest correlations were visualized (**Figure 6-6B**). Enrichment analysis revealed that genes co-expressed with *DLC1* in CMPs were involved in the regulation of myeloid cell differentiation (GO:0045637) (**Figure 6-S2**), and consisted of 11 intersecting genes including *MEF2C*, *PRKCQ*, and *ITGA2B*. Furthermore, to define whether the expression of candidate genes in CMP were associated with changes in the relative abundance of isotype and subclass genes, the mean relative expression of specific immunoglobin heavy constant (*IGH*) genes, expressed in Plasma cells (PCs), were correlated to the mean expression of the candidate genes across donors. This analysis of CMP candidate genes and PC IgG Fc domain profiles revealed strong correlations between *DLC1* expression and the relative abundance of *IGHG2* and *IGHG3*, but reduced levels of *IGHA1* and *IGHA2* (**Figure 6-6C**). This therefore links the expression of *DLC1* within CMPs to the selection of particular isotypes and subclasses.

Similarly, we next examined the co-expression of genes with *RNF141*, previously associated with Polio-specific FcγR3A-binding levels, within CMPs and granulocyte-monocyte progenitors. Both negative and positive relationships were identified (**Figure 6-6D**). Negative correlations were enriched in genes responsible for antigen processing and presentation of exogenous peptide antigen via MHC class II (GO:0019886), and positive correlations in genes related to neutrophil degranulation (GO:0043312) (**Figure 6-S2**). Antigen processing and presentation consisted of 9 overlapping genes comprised of 6 *HLA* genes, *CD74*, and *CTSD* whereas neutrophil degranulation consisted of an overlap of 7 genes including strong correlations to *S100A8*, *GCA*, and *S100A12*. This potentially points to a critical role for antigen-presentation to T cells that ultimately

shapes antibody Fc-biology. Additionally, *RNF141* expression in CMPs was correlated with relative isotype and subclass abundance in PCs and revealed strong correlations between *RNF141* with *IGHG3* and *IGHG4* expression, but reduced *IGHA1* and *IGHA2* levels (**Figure 6-6E**). This once again highlighted a clear correlation between gene expression in myeloid cells and Fc-biological gene profiles in PCs.

Finally, this transcriptomic analysis was extended to SPRY2. Co-expression correlations between SPRY2, previously associated with Flu-specific ADNP levels, in CMPs and myeloid cells revealed strong negative correlations to ribosomal proteins important in protein translation (GO:0006614, GO:0002181, and GO:0006613) and cell-to-cell communication (**Figure 6-6F** and **Figure 6-S2**). Moreover, expression of *SPRY2* in CMPs was strongly associated with elevated *IGHA1* and *IGHA2* in PCs (**Figure 6-6G**), and further supporting the role for specific gene-signatures in CMPs that aid in shaping B cell functional Fc-programming.

Finally, beyond expression analysis, the NHGRI-EBI Catalog of Human Genome-wide Association Studies (https://www.ebi.ac.uk/gwas/) was used to ultimately identify overlap between SNPs which associated with pathogen-specific humoral traits. SNPs, identified in the *FUT2* gene, that were linked to Post-RSV-specific IgG and FcR levels, showed strong associations with population level susceptibility to Mumps, common colds, childhood ear infections, chronic inflammatory diseases, autoimmune diseases, Type 1 Diabetes, antibody levels against BK polyomavirus VP1, and the concentration of breast milk oligosaccharides (**Table 6-S4**). Moreover, among these SNPs the nonsense variant rs601338, encoding for the "non-secretor" phenotype, has previously been reported to provide resistance against respiratory viral pathogens including RSV, Influenza A, and Rhinovirus [276] as well as mucosal pathogens Norovirus (GII.4) [277], Rotavirus (VP8) [278], and HIV-1 [279]. This collectively further validates the link between these SNPs, differential humoral immune functional programming, and resistance/susceptibility to disease. Ultimately, these data collectively provide mechanistic insights into the basis for genetic programming of Fc-biology, that may provide clues for future vaccine design efforts aimed at leveraging the potent role of Fc-biology in the battle against infections, malignancies, and beyond.

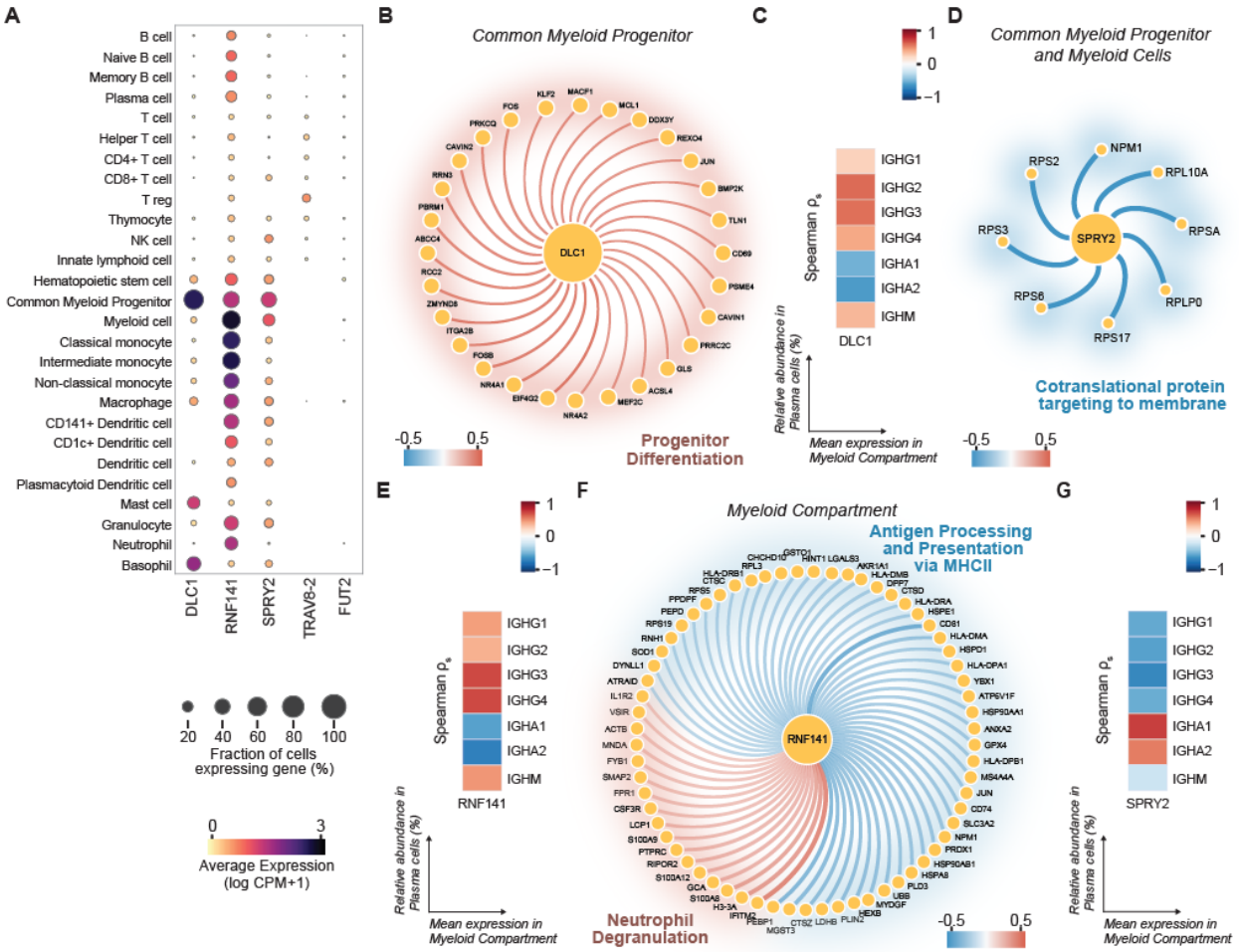


Figure 6-6. Transcriptomic Analysis of Genes Associated with the Development of Pathogen-specific Antibody Responses

(A) Average expression of candidate genes across immune cell populations. Size corresponds to the fraction of cells which expression the gene and color the average expression in log counts per million plus one. Co-expression spearman correlation networks to each candidate genes (B, D, and F), and correlations between mean expression of candidate genes within a cell type or subset of cells and the relative abundance of IGH genes in plasma cells (C, E, and G) across donors were depicted. Co-expression network edge widths and color intensities correspond to the magnitude of the correlation coefficients with red relating to positive and blue negative correlations. All correlations shown are statistically significant past Bonferroni correction for multiple hypotheses. (B) The DLC1 network was built with common myeloid progenitors (CMP), (D) the SPRY2 network on CMP and myeloid cells, and (F) the RNF141 network on CMP and cells of monocyte-granulocyte lineage.

6.4 Discussion

Pathogen specific humoral immunity is driven by polyclonal antibodies, and through diversification of antigenic targeting, antibody structure, and mediated effector functions enable distinctive tailored responses. Compositional differences in these properties have been shown to be important in combatting infection against many of infectious diseases[7,17,13]. Beyond the identification of polyclonal antibodies properties

advantageous against infection, the foundational mechanisms which give rise to resistant or susceptible subpopulations remains poorly understood. Familial-based studies of heritability have previously attributed the influence of genetics in determining pathogen-specific IgG titers against common infectious diseases[263], and subsequent GWAS studies have revealed many strong SNP associations[256–259]. However, the genetic elements which control broader pathogen-specific polyclonal antibody characteristics including isotype and subclass abundance, enhancement to Fc-receptors, and antibody mediated effector functions remain largely unknown. Thus, here using 510 twins a rich dataset of 227 pathogen-specific humoral traits was generated, assessed for heritability, and screened for genetic associations. This revealed the identification of multiple genetic variants spanning five genomic regions with genetics accounting for a up to 68% of total variation. Both pleiotropic and single control of pathogen-specific humoral traits was identified and largely involved innate immune processes. Ultimately, these findings pointed to genetic mechanisms of humoral traits underpinned by both immune processes and host-pathogen relationships.

Infections represent a major selective pressure for humans [280], and variations in host genetics have previously shown the capacity to influence both resistance and susceptibility to infection. For example, mutations in Duffy antigen receptor for chemokines (*DARC*) and CC-chemokine receptor 5 (*CCR5*) have shown to confer resistance against *Plasmodium vivax*[281] and HIV-1[282] respectively. Both of these mutations inhibit host invasion and therefore represent a class of genetic variants related to mechanisms of infection underscored by host-pathogen interactions. More broadly, genetic variation can also impact different stages of humoral development either through direct or indirect influence. Thus, the underlying mechanisms which describe pathogen or allergen specific humoral traits can be best understood through the context of these two delineations.

Variant rs62493093 which was significantly associated to IgG2 levels against HBV related to DLC1 Rho GTPase-activating protein (*DLC1*) and exhibited control influenced by both immune intrinsic and host-pathogen factors. *DLC1* enables the activation of small GTPases which regulate the actin cytoskeleton and consequently induce morphological changes and control cell migration[283] and was robustly expressed of in common myeloid progenitors. Co-expression correlations further revealed an enrichment of genes involved in the regulation of myeloid cell differentiation. Distinct differences in Rho GTPase expression have previously been observed during various stages of myeloid differentiation[284] and presents a potential route in which *DLC1* may influence myeloid differentiation. Elevated levels of *DLC1* were also correlated to a relative increase in IGHG2 expression and validated the IgG2 phenotype. HBV is a non-cytopathic hepatotropic virus[285] and relies on the immune unresponsive nature of the liver[286]. Therefore, effective immune responses against HBV must balance viral control with liver pathology. In this regard IgG2 is advantageous as it has a lesser ability to activate effector functions due to its reduced affinity to FcγRs[212] and thereby potentially

drives opsonization while mitigating inflammation. Beyond its role in CMP differentiation, *DLC1* is also a strong suppressor of *MYC* activation and its down expression is a hallmark of hepatocellular carcinoma (HCC)[287] which is also linked to HBV infection[288]. Furthermore, suppression of *DLC1* through *miR-141* has previously been demonstrated as a requirement for efficient HCV replication[289]. Therefore, control of *DLC1* exhibits both pathogen and host specific attributes which ultimately contribute to the determination of IgG2 levels against HBV.

Contrary to *DLC1*, variants located within ring finger protein 141 (*RNF141*) showed evidence of pleiotropic effects and included significant associations to Polio specific FcγR3A binding as well as strong associations to FcγR3b, Varicella-Zoster Virus (VZV) specific binding to both FcγR3A and FcγR3b, and broad associations to many pathogen-specific IgM isotypes. Strong expression of *RNF141* was observed among cells derived from common myeloid progenitor cells, positively correlated to genes involved in neutrophil degranulation, and negatively correlated to genes involved in antigen processing and presentation via MHCII. At present, the full extinct *RNF141*s molecular function is poorly understood. Being a ring finger protein, *RNF141* exhibits E3 ubiquitin ligase activity resulting in degradations of proteins[290]. The exact substrates *RNF141* targets have yet to be experimentally validated, but binary protein interaction data suggests that *RNF141* may target mevalonate kinase (*MVK*)[291]. *MVK* is an enzyme responsible for the production of non-sterol isoprenoids which are necessary for protein prenylation and can alter protein function[292]. For example, loss of *MVK* function resulting in unprenylated *RhoA* has been shown to lead to the over production of IL-1β, IL-6, and TNF through the activation of *Rac1* [293,294]. Furthermore, impaired prenylation due to *MVK* deficiency causes of hyperimmunoglobulinaemia D[295] and highlights the ability of *MVK* function to influence class switching. Therefore, *RNF141* may present as a potential regulator of protein prenylation altering cytokine production and thereby modulating innate immune function and antibody class switching. This was further supported by the observed correlations between *RNF141* expression and the increase in relative IgG3 abundance. Additionally, FcγR3 binds IgG3 with higher affinity than other subclasses[212,296] and would explain the enhanced binding to FcγR3A and FcγR3b that was observed against Polio and VZV. Ultimately, the mechanism under which *RNF141* influences class switching may be more nuanced and thereby not fully explained by static negative correlations to certain HLA genes.

Sprouty RTK signaling antagonist 2 (*SPRY2*) also exhibited pleiotropic genetic control but with significant and suggestive associations to ADNP function. Expression of *SPRY2* was mainly observed in CMP and myeloid cells. Co-expression correlations further implicated reduced cotranslational protein targeting to the membrane related to negative associations to ribosomal proteins which holistically described regulation of translational activity. *SPRY2* is a feedback regulator which controls the mitogen-activated protein kinase (MAPK) pathway through multiple avenues[297]. The MAPK pathway is important for many cellular processes including ribosomal biogenesis[298].

Additionally, ribosomal levels have previously been shown to regulate lineage commitment in hematopoietic stem and progenitor cells (HSPCs)[299]. Therefore, *SPRY2* through modulation of MAPK signaling influences ribosomal expression and thereby lineage commitment. The expression of *SPRY2* was also correlated to IgA1 and IgA2 expression in plasma cells. This is further supported by reports of rare mutations in *SPRY2* causing IgA nephropathy, a buildup of IgA in kidneys[300]. These mutations exhibited reduced levels of *SPRY2* phosphorylation and therefore enhanced inhibition of the MAPK/ERK pathway and ultimately leading to the over production of IgA. Although the mechanism in which inhibition of the MAPK pathway results in the preference for IgA is not presently known, it is plausible that this occurs due to changes in innate immune cellular phenotypes. Consequentially, IgA isotypes are also potent drivers of ADNP function[301]. Thus, *SPRY2* may influence ADNP function by both effecting CMP differentiation and driving elevated levels of IgA.

Beyond mechanisms related to immune cell function coding variants found in the *FUT2* gene which significantly associated with IgG1 levels and FcγR binding specific to the RSV post-fusion glycoprotein F (Post-RSV) represented mechanisms related to mucosal defense and host-pathogen relationships. *FUT2* encodes for the α1,2-fucosyltransferase 2 enzyme responsible for the attachment of A and B histo-blood group antigens onto proteins and lipids that are either anchored to the surface of mucosal epithelium or secreted[302]. Variant rs601338A which was in complete linkage disequilibrium with lead variant rs679574 is a nonsense variant leading to the enzymatic inactivity of *FUT2* and the lack of histo-blood group antigens (HBGA) on mucosal surfaces known as “non-secretor” status[303]. Non-secretor status has been shown to provide resistance against Norovirus (GII.4)[277] and many respiratory pathogens including RSV[276]. Mechanistically, resistance against norovirus has been attributed to the loss of HBGA which facilitates viral attachment [304]. RSV viral entry relies on attachment usually through the G glycoprotein and membrane fusion which is facilitated by the F glycoprotein which undergoes a large confirmational change in the process[305]. Although less is known about the fusion process it is plausible that HBGA acts as a fusion mediator resulting in the prolonged exposure of the post fusion confirmation state which switches back and forth at a basal rate[306]. In this way non-secretor status may provide resistance as well as aid in the development of humoral response against RSV.

Variants which were significantly associated to various pathogen-specific humoral traits collectively highlighted mechanisms driven by innate immune processes and described by specific host-pathogen interactions. In order to thoroughly understand these relationships, 13 humoral features across 20 pathogens and 510 individuals were profiled. Although this represents a comprehensively large dataset from a systems serology perspective, it is smaller than conventional GWAS studies. Despite the size of the cohort five genomic regions reaching genome-wide significance were identified and accounted for a substantial amount of variation. Furthermore, many significant processes were identified as a result of pathway analysis which with further validation

cohorts in larger cohorts could statistically implicate additional SNPs. Correlates of immunity driven by the properties and functions of polyclonal antibodies are still being revealed in many disease contexts and across many populations. The data generated here can provide a useful roadmap for understanding the humoral landscape as it relates to various pathologies, inform vaccine design, the development of therapeutics, and provide potential diagnostic markers for the prediction disease outcomes.

6.5 Methods

6.5.1 Samples

Plasma samples from 510 participants were obtained from the UK Adult Twin Register, TwinsUK, were included in this study. The TwinsUK cohort consists of twins which was originally formed in order to study the genetic impact of various traits on the female population. The broader cohort includes individuals that are representative of the general British population which are of Caucasian descent, and is further described in greater detail in[307]. The participants selected for this study consisted of 168 monozygotic and 342 dizygotic matched twins. Of the 510 participants, 498 were provided with whole-genome genotyping data. All participants were female between the ages of 41-78 (mean 61), and through the means of self-reporting were 100% Caucasian and of UK ancestry. Regardless, population stratification by principal component analysis was still conducted in order account for any self-reporting discrepancies. All experimental data was captured in two technical replicates and the average value was reported for all assays. This study was approved by the MGH Human Subjects Institutional Review Board and London-Westminster NHS Research Ethics Committee, and all participants provided informed consent.

6.5.2 Primary immune cells

Primary human cells were obtained from healthy volunteers at MGH blood donor center. All donors were over 18 years of age, provided signed consent, and were de-identified before use. This study was approved by the MGH Institutional Review Board. Human primary Neutrophils were isolated from fresh peripheral blood and maintained at 37˚C, 5% CO2 in RPMI with 10% fetal bovine serum, L-glutamine, and penicillin/streptomycin.

6.5.3 Cell lines

THP-1 cells (ATCC), a monocytic leukemia cell line, were maintained in RPMI supplemented with 10% fetal bovine serum, L-glutamine, penicillin/streptomycin, HEPES, and beta-mercaptoethanol. THP-1 cells were grown at 37˚C, 5% CO2.

6.5.4 Quantification of antibody subclasses, isotypes, and Fc-receptors by luminex

Antigen specific subclass and isotypes, and Fc-receptor binding levels were quantified through a customized Luminex assay as described in [308]. This high-throughput assay allows for the assessment of relative antibody concentration against Betv1, Arah2

(Indoor Biotechnology), ActHIB Vaccine, Pneumovax23 Vaccine, IPOL Vaccine (MGH Pharmacy), Pertussis toxin (List Biological Laboratories), Tetanus toxin, Diphtheria toxin, Norovirus VP1 GII.4 (Native Antigen), PPD (Statens Serum Institute), HBV HBsAg Adw (GenWay), EBV gp350/220 (Immune Tech), VZV gE (kindly provided by Lingwood Lab, Ragon Institute of MGH, MIT and Harvard), HA1 (A/California/07/2009) (H1N1) (Immune Tech), HA1 (A/Texas/50/2012/H3N2) (eEnzyme), RSV A/B Pre F Fusion Protein, RSV A/B Post F Fusion Protein (kindly provided by XXX, NIH), Mumps – Enders, Measles – Endmonston, and Rubella HPV-77 (BioRad Antibodies). Antigens were covalently coupled to carboxylated microspheres (Luminex) through EDC and sulfo-NHS crosslinking (Thermo Scientific) as per the manufacture's recommendations. Antigen coated microspheres (~800 of a single microsphere type per well) were added to 384 well plates (Greiner Bio-One). Samples and controls were diluted at 1:100 in PBS except for IgG1, IgG2, IgG4, and IgA1 that were diluted 1:50. Optimal Ig specific dilutions were determined beforehand through titration. Samples and controls were added to the wells and incubated for 16 hrs at 4ºC. Microspheres were washed three times (PBS-0.1% BSA-0.05% tween-20) and were incubated with a PE-coupled detection antibodies for each subclass and isotype (IgG1, IgG2, IgG3, IgA1 and IgM, Southern Biotech), and Fc-receptor (FcγR-2A, -2b, -3AV, -3b, Duke Protein Production facility) over the course of 1 hr. Fc-receptors were fluorescently labeled with PE prior to the addition of immune complexes. This was achieved by using recombinant AviTag FcRs which were biotinylated with the BirA-500 kit (Avidity) according to the manufacturer's instructions. Next, FcRs were incubated with streptavidin-PE (Thermo Scientific) for 10 min, and then quenched with an excess of 20 μM biotin for 10 min. After, incubation with detection antibodies plates were washed three times and acquired through flow cytometry, using the Intellicyt iQue Screener and S-Lab robot (PAA). Analysis was performed using ForeCyt software by gating on fluorescent bead regions and PE median fluorescent intensity (MFI) was reported. Background signal as defined as the MFI of microspheres incubated with human serum minus IgA/IgM/IgG (Millipore) was subtracted, and all samples and controls were run in duplicate.

6.5.5 Effector functional assays

Three bead based functional assays were performed during this study including antibody-dependent cellular phagocytosis (ADCP), antibody-dependent neutrophil phagocytosis (ADNP) and antibody-dependent complement deposition (ADCD) as previously described in[66,161,160].

Phagocytosis assays (ADNP and ADCP) were multiplexed in order to accommodate the throughput needed following the procedure described in [98]. First the Pneumovax23 Vaccine, the IPOL Vaccine (MGH Pharmacy), Pertussis toxin (List Biological Laboratories), Diphtheria toxin (Native Antigen), PPD (Statens Serum Institute), EBV gp350/220 (Immune Tech), a 1:1 mix of HA1 (A/California/07/2009) (H1N1) (Immune Tech) and HA1 (A/Texas/50/2012/H3N2) (eEnzyme), a 1:1 mix of RSV

A/B Pre F Fusion Protein and RSV A/B Post F Fusion Protein (kindly provided by XXX, NIH), and Measles – Endmonston (BioRad Antibodies) were biotinylated with sulfo-NHS-LC-biotin (Thermo Scientific) according to the manufacturer's instructions. Then Zeba spin desalting columns (Thermo Fisher) were used to remove unbound biotin through buffer exchange to PBS. Biotinylated antigens were then separately coupled to either red, yellow, or scarlet fluorescent neutravidin beads (Thermo Fisher) at a ratio of 1:1. Beads and biotinylated antigens were coupled at 37°C for 2 hrs in low-binding microcentrifuge tubes (Corning). Beads were then washed twice with 0.1% PBS-BSA through centrifugation (16,000 x g, 5 min) at room temperature to remove unbound antigen. The antigen-coupled fluorescent beads were resuspended in 1:100 0.1% PBS-BSA and stored for up to no more than a week at 4°C in the dark. Sets unique color antigen coated beads were mixed in equal ratios prior to the beginning of the experiment.

Immune complexes were then formed by the addition of 10µL of samples diluted 1:100 in 0.1% PBS-BSA and 10µL of the antigen coupled bead mixture in each well of a 96-well round bottom microplate (Corning) for 2 hours at 37°C. After the formation of immune complexes, 200 µL of 0.1% PBS-BSA was added to each well, pelleted (1000 × g, 10 min), and the supernatant removed to clear unbound antibodies. For ADCP THP-1s (ATCC), monocyte cell line, were incubated with immune complexes at 125,000 cells/mL at 37˚C for 16 hours. After the incubation period, cells were fixed with fixed with 4% paraformaldehyde (Alfa Aesar). For ADNP, following immune complex formation white blood cells from healthy donors were added at concentration of 125,000 cells/mL and incubated at 37˚C for 1 hr. White blood cells were isolated from ACD-treated whole blood through the lysis of a red blood cell. To remove red blood cells, a 1:10 ratio of ammonium-chloride-potassium (ACK) lysis buffer (150mM NH4Cl, 10mM KHCO3, 0.1 mM Na2EDTA, pH 7.4) was incubated with the cells for 5 min at room temperature followed by sequent washes with PBS. Following the bead-cell incubation period, neutrophils were stained with an anti-CD66b Pacific blue antibody (Biolegend) at concentration of 1:100 in 0.1% PBS-BSA and then fixed with 4% paraformaldehyde (Alfa Aesar).

For quantification of antibody-dependent deposition of C3, an adapted multiplexed protocol was used based off of procedures described in[309]. First, the same antigens that were used for the ADNP and ADCP assays were carboxyl coupled to Luminex beads in the same manner described above, but with four times as more antigen per bead region. Antigen coated microspheres were then combined and diluted 1:300 in PBS-0.1% BSA-0.05% tween-20. 10µL of plasma samples diluted 1:10 in PBS-0.1% BSA were added to each well of a 384 well plate (Greiner Bio-One). Then immune complex formation was then facilitated by adding 45µL of the microsphere bead mixture to each well (~730 bead/well) and was incubated at 37°C for 2 hrs on a plate shaker at 800rpm. Following immune complex formation plates were washed twice in PBS-0.1% BSA. Then, lyophilized guinea pig complement (Cedarlane) was reconstituted according

to manufacturer's instructions, diluted 1:44 in gelatin veronal buffer with calcium and magnesium (GBV++) (Boston BioProducts), 90µL of complement was added to each well, and incubated at 37ºC for 20 minutes while shaking at 800rpm. Following complement deposition plates were washed twice with cold 15mM PBS-EDTA. Afterwards, C3 was detected with an anti-C3 fluorescein-conjugated goat IgG fraction detection antibody (Mpbio). The detection antibody was diluted 1:200 in PBS, 50µL were added to each well, and was incubated at room temperature in the dark while shaking at 800rpm. Plates were then washed twice more before the acquisition of the data.

All effector function assays were acquired with the iQue (Intellicyt) flow cytometer and S-Lab robot (PAA). Analysis was performed using ForeCyt software. THP-1 cells were gated on singlets and bead-positive cells (ADCP), and Neutrophils were gated on singlets, CD66b, and bead positive cells (ADNP). A phagocytosis score was determined as the percent of bead-positive cells times the MFI of bead-positive cells divided by 10000. ADCD was determined by gating microspheres on singlets, fluorescent bead region, and C3 deposition was defined as PE MFI. All samples were averages of two technical replicates. Additionally, each replicated consisted of a different set of antigens which were conjugated on different color beads. Background signal was defined as the MFI of samples incubated with a PBS no antibody control which was subtracted from each score.

6.5.6 Genotyping

Genotyping of the TwinsUK cohort was performed with a combination of Illumina arrays (HumanHap300, HumanHap610Q, 1M-Duo and 1.2MDuo 1M). The normalized intensity data were pooled for each of the three arrays separately, with 1M-Duo and 1.2MDuo 1M being pooled together. For each dataset, Illluminus calling algorithm was used to assign genotypes. No calls were assigned if an individual's most likely genotyped was called with less than a posterior probability threshold of 0.95. Validation of pooling was achieved through the visual inspection of 100 random, shared variants for any overt batch effects. Finally, intensity cluster plots of significant SNPs were visually inspected for over dispersion biased no calling, and/or erroneous genotype assignment. SNPs exhibiting any of these characteristics were discarded.

6.5.7 Heritability estimation

In order to determine the influence of genetics on humoral traits biometrical genetic modeling of twin data was used. Following systems serology profiling of the cohort of twins, contributions to additive genetics (A), common environment (C), and unique environment (E) were decomposed through the use of linear mixed effect models. Prior to analysis, all traits were first rank-based inverse normal transformed. Models were fit with the specific parameterization structure described in[264] and is based on structural equation modelling[310]. Briefly, this approach assumes that MZ twins share 100% of their genetics, DZ twins share approximately 50% of their genetics, all twins share equal

exposure to similar environmental influences, and that each twin has their own unique exposures. Therefore, narrow estimates of heritability can be made by the construction of linear mixed effect models which through the maximization of a log-likelihood function decomposes of covariance matrices into A, C, and E components. Heritability is therefore estimated by evaluating the variance due to additive genetics of particular a trait as a fraction of the total variance. An added benefit of this approach was that age could be accounted for as a potential confounder through its addition as a fixed effect. Thus, this approach was utilized to estimate the genetic influence across all humoral traits.

Following the estimation of heritability, null models were constructed through the exclusion of the genetic parameter. In order to determine whether the inclusion or exclusion of the genetic parameter improved the process of fitting the Akaike Information Criterion (AIC) was used[238]. The AIC rewards goodness of fit while penalizing the addition of parameters which in turn helps avoid overfitting, with lower values of AIC considered as more preferable. Therefore, through contrasting the AIC values between the full genetic model and the null model of specific traits justification for a genetic component can be made. This was further quantified through determining ΔAIC scores which were calculated through the subtraction of a null model's AIC by the AIC of their respective genetic model. Thus, the ΔAIC could be interpreted as a degree of confidence and helped provide justification for the influence of genetics on a trait. All calculation were implemented in R, version 3.4.4 using custom code available in the data availability section.

6.5.8 Non-parametric combination

Statistical evaluation of various classes of humoral traits were assessed using non-parametric combination[230,231]. In brief, for each type of antibody characteristic, pathogen type, or other higher class (i.e., IgG1, Measles, all viral pathogen) class-level mean heritability scores were determined by first estimating the heritability of each sub-feature (i.e., IgG1 H1N1, IgG1 Post-RSV, etc.) through the linear effect modelling framework described above. Next, null distributions for each global mean heritability estimate was constructed through permutating the twin pair zygosity labels a thousand times. This ensured that the permutation structure among the sub-features in each iteration that collectively comprised the null distribution of mean heritability estimates for a specific trait was preserved. Thus, significance was defined as the exact P value of the tail probability of the true mean heritability score within the null distribution.

6.5.9 Genotyping quality control

Quality control (QC) consisted of the removal of samples and variants which did not meet exclusion criteria and was conducted prior to analysis. Samples with genotypic call rates < 98% or excess heterozygosity across all SNPs (≥ 2 standard deviations from the sample mean) were excluded from the analysis. Additionally, ethnic outliers which

showed evidence of non-European ancestry were assessed by principal component analysis in comparison which HapMap3 populations and were therefore removed. Lastly, identity by descent (IBD) probabilities were used in order to identify samples that were suggestive of identity errors which were removed, and misclassifications of MZ and DZ twins were corrected.

Following quality control of samples, variant QC was performed using a set of unrelated samples. This included the removal of SNPs with minor allele frequency (MAF) < 1%, significant deviation from Hardy-Weinberg equilibrium (HWE) ($p < 10^{-6}$), and SNP call rate < 97% (SNPs with MAF ≥ 5%) or < 99% (for 1% ≤ MAF < 5%). Following variant QC, alleles from the genotyping arrays were combined and aligned to build37 forward strand alleles (https://www.well.ox.ac.uk/~wrayner/strand/).

6.5.10 Genotyping imputation

Prior to imputation the "HRC/1KG Imputation Preparation and Checking Tool" (version 4.2.5) (developed by Will Rayner) was used to check input data for accuracy relative to expected HRC (http://www.haplotype-reference-consortium.org/site) or 1000G Phase 3 reference panels (http://www.well.ox.ac.uk/~wrayner/tools/). This process identified errors in the original data, including incorrect REF/ALT designations, strand designations, extreme deviations from expected allele frequencies, and palindromic (A/T and G/C) SNPs with allele frequency near 0.5 which are often the source of imputation errors. The problematic variants identified were then either fixed or removed. Then using the quality controlled genotypes were pre-phased using SHAPEIT2[311] with the 1000 Genomes haplotypes Phase 3 integrated variant set release in NCBI build 37 (hg19) coordinates (https://mathgen.stats.ox.ac.uk/impute/) as a reference panel. Imputations were performed using the Michigan Imputation Server[312].

6.5.11 GWAS analysis

Genome-wide associations between pathogen-specific humoral traits and variants which were either directly genotyped or imputed were performed using the FaST-LMM package[313]. This package was chosen due to its ability to account for relatedness and hidden population structure in an additive genetic model through a realized genetic similarity matrix (GSM) using genotyping data. In order to avoid proximal contamination, GSMs are computed from all variants in other chromosomes other than where the test variant is located and are cross validated in a process called leave out one chromosome (as implemented in FaST-LMM)[314]. Models in the FaST-LMM algorithm are fit with restricted maximum likelihood (REML) method and *P* values between a given variant and trait are computed using the likelihood ratio test. Additionally, age was added as covariate and was corrected for across all associations, and the percentage of the variance explained by each variant was determined using FaST-LMM. Association results were further quality controlled in Plink version 1.9beta[315] by the removal of

variants with minor allele frequency (MAF) > 0.1, Hardy-Weinberg equilibrium < 10^{-8}, and SNP call rate > 90%. This yielded a total 4,663,961 of variants.

GWAS were performed on select pathogen-specific antibody responses (n=45) which provided the best evidence for heritability (heritability > 30%, and ΔAIC > 0). Genomic control inflation factors (λ_{GC}) had a median of 1.009 (range 0.990 - 1.026) and indicated that population stratification had minimal influence on the overall test statistical distribution. Associations were considered genome-wide significant for ($p < 10^{-8}$) and suggestive for ($p < 10^{-4}$). Characterization of lead SNPs was achieved through selection of variants with the smallest p values within a 1-Mb genomic region. Linkage disequilibrium-based clumping was performed on GWAS results for each trait separately using Plink version 1.9 beta for all variants with association of ($p < 10^{-4}$). This process compared variants in LD ($r^2 > 0.5$) within a 100-kb block of the strongest associated variants and pruned them. Lastly, regional association plots were generated using the browser-based application of LocusZoom.js[316] with 1000 Genomes genotypes for European populations on build 37 (hg19) as an LD resource and built-in gene annotation. Genes which were further from variants were omitted for visualization processes when variants were in regions of high gene densities.

6.5.12 Variant annotation

Functional annotation of genetic variants was performed with ANNOVAR using build hg19[265]. Gene-, region-, and filter-based annotations were determined using the refGene, dbnsfp42a, and gwasCatalog databases. The regulatory elements of variants were evaluated using both RegulomeDB v2.03[266] and HaploReg v4.1. This included the search for overlap in eQTL data, regulatory chromatin states and histone modifications, altered motifs, and proteins bound. Both the leading variants and variants in linkage disequilibrium ($r^2 > 0.8$) were annotated (Table 6-S2).

6.5.13 Gene set analysis

Gene set analysis was performed using magma, version v1.09. Traits which showed evidence of heritability (univariate=45, multi-trait=29) were selected for gene set analysis. Prior to analysis, SNPs were mapped to genes based on a genomic window of 10-kb using build 37 (hg19). Gene based p-values and correlations among neighboring genes were then computed using the Top 1 SNP-wise gene analysis model. For each traits variants with phenotypic associations of ($p < 0.05$) were included as inputs. Competitive gene-set analysis was then implemented using gene analysis using the trait specific gene analysis results using a database of immune specific gene ontology biological processes. Gene set p values across all traits were then multiple hypothesis corrected using the Benjamini-Hochberg procedure with a false discovery rate of ($\alpha = 0.05$).

6.5.14 Co-expression correlations

Single cell RNA sequencing data from[275] was used for all expression based analyses and was select for its coverage . Gene counts were normalized to log counts per million (CPM) using Scanpy package version 1.9.1[317]. Co-expression correlation networks were performed using protein coding genes and GWAS candidate genes, and mitochondrial genes and genes that were not expressed in the major of were excluded. Immune cell types selected for co-expression analysis were based on the mean expression of GWAS candidate genes. Spearman correlations were then determined and pruned based on statistical significance with Bonferroni correction for multiple hypotheses and ($|\rho| > 0.3$). Correlation networks were constructed based on top correlations between genes and GWAS candidate genes and visualized using Cytoscape version 3.7.2. Enrichment analysis of co-expressed genes was determined using Enrichr with the GSEAPY package version 0.10.8[318]. Correlations among the mean expression of genes identified by the variant screen and fraction of isotype or subclass abundance were determined through spearman correlations across donors. Isotype or subclass abundance was determined through comparing the expression of immunoglobin heavy constant (IGH) genes (i.e., IGHG1) as a fraction of all IGH isotype and subclass expression. Correlations to class switch recombination genes and were assessed in a similar fashion but with mean expression of genes in B cells.

6.6 Supplementary figures and tables

Disease	Class	Antigens	Antigen Type
Pollen	Allergy	Betv1	Protein
Peanut	Allergy	Arah2	Protein
Haemophilius Influenzae type b	Bacteria	Capsular Polysaccharides	Polysaccharide
Pertussis	Bacteria	Pertussis toxin	Protein
Tetanus	Bacteria	Tetanus toxin	Protein
Tuberculosis	Bacteria	PPD	Mix
Streptococcus pneumoniae	Bacteria	Capsular Polysaccharides	Polysaccharide
Diphtheria	Bacteria	Diphtheria toxin	Protein
Hepatitis B	Virus	HBsAg	Protein
Measles	Virus	Inactivated Virus	Mix
Mumps	Virus	Inactivated Virus	Mix
Norovirus	Virus	VP1	Protein
Poliomyelitis	Virus	IPOL Vaccine (3 Polio Serotypes)	Mix
Respiratory Syncytial Virus A/B	Virus	Pre F Fusion Protein	Protein
Respiratory Syncytial Virus A/B	Virus	Post F Fusion Protein	Protein
Rubella	Virus	Inactivated Virus	Mix
Varicella-Zoster	Virus	gE(Orf68)	Protein
Epstein-Barr Virus	Virus	gp350/220	Protein
Influenza A virus CA/7/2009	Virus	H1N1	Protein
Influenza A virus Texas/7/2012	Virus	H3N2	Protein

Table 6-S1. Summary of Profiled Infectious Diseases and Antigens, related to Figure 6-1

Pathogen classification, antigens, and antigen types which represent the infectious disease or allergens which were profiled in this study.

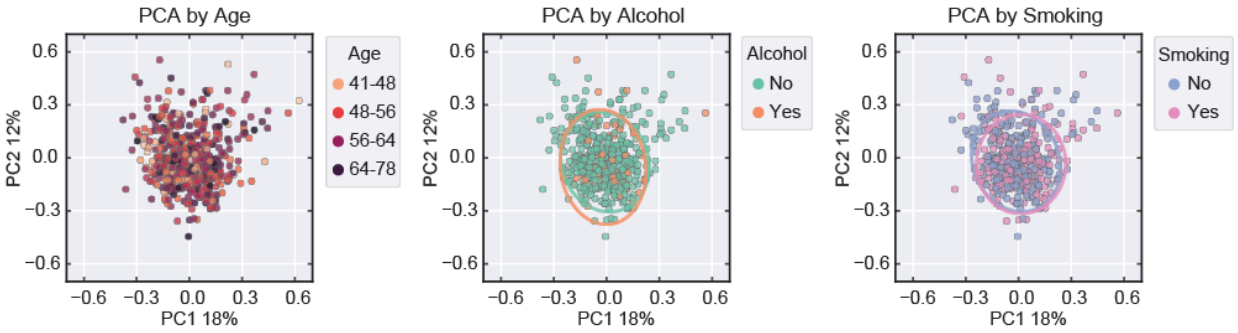


Figure 6-S1. Evaluating of variation driven by potential external effects, related to Figure 6-1

Principal components analysis (PCA) scores plot built on normalized humoral features and split by age ranges, alcohol use, and smoking. Normalization was achieved by first subtracting technical noise followed by rank-based inverse normal transformation. In case of alcohol use and smoking ellipses correspond to the 95% confidence intervals for each group.

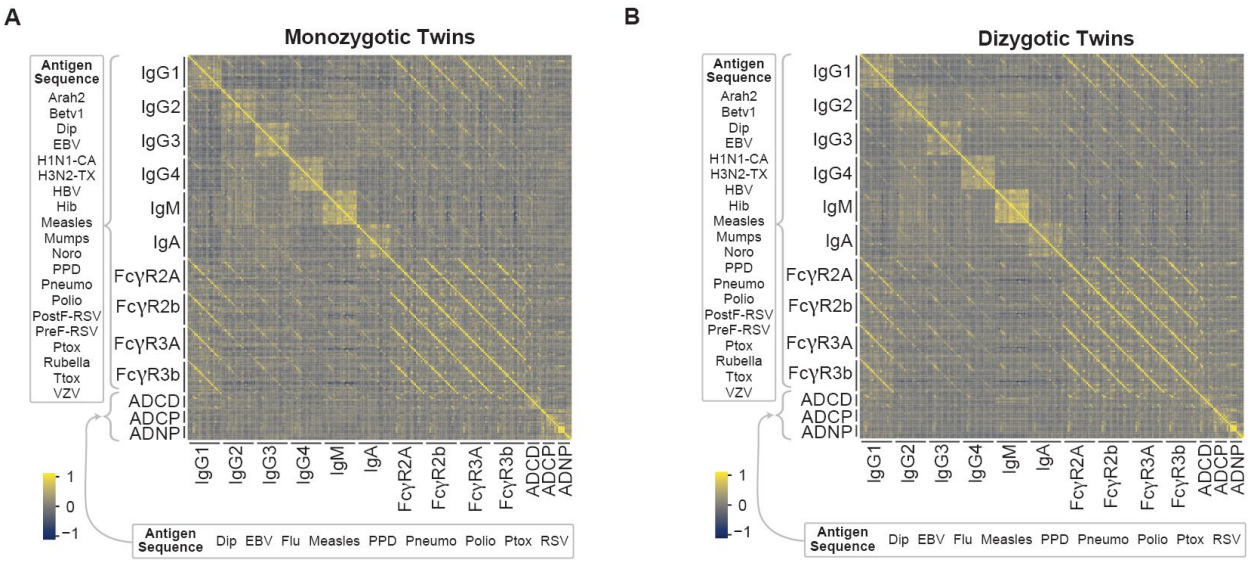


Figure 6-S2. Coordination of Antibody Characteristics and Pathogen Specificities across Monozygotic and Dizygotic Twins, related to Figure 6-2

Pairwise spearman correlations organized by type of antibody measurement and with a repeating antigen sequence determined using only monozygotic twins (A) and dizygotic twins (B) respectively.

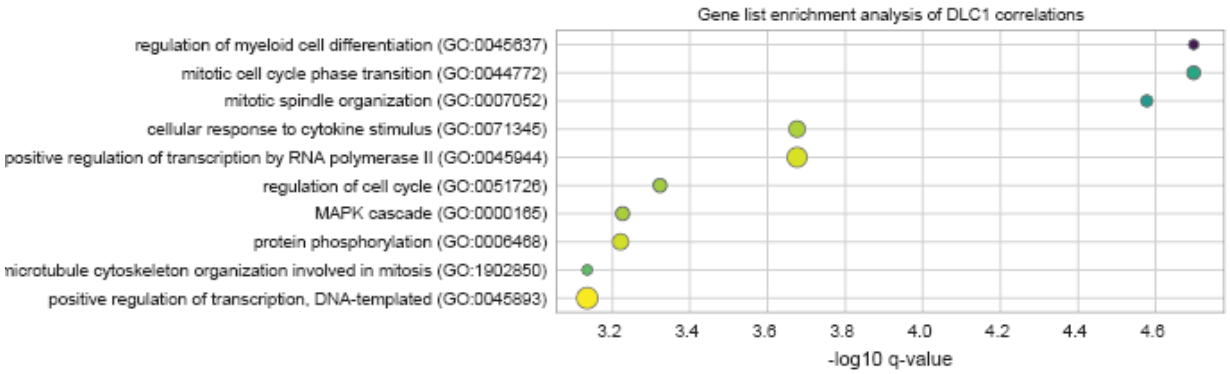
**Avaiable upon request*

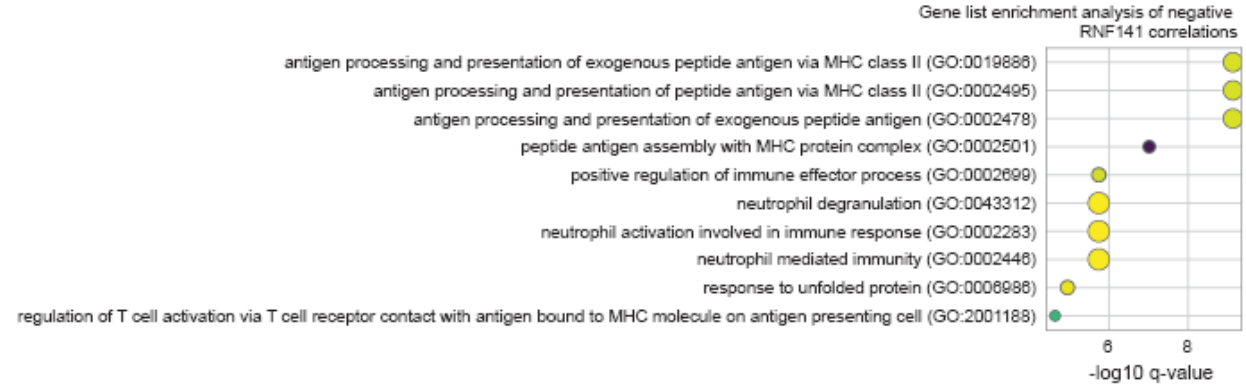
Table 6-S2. SNPs in Linkage Disequilibrium with Genome-wide Significant Hits, related to Figure 6-5

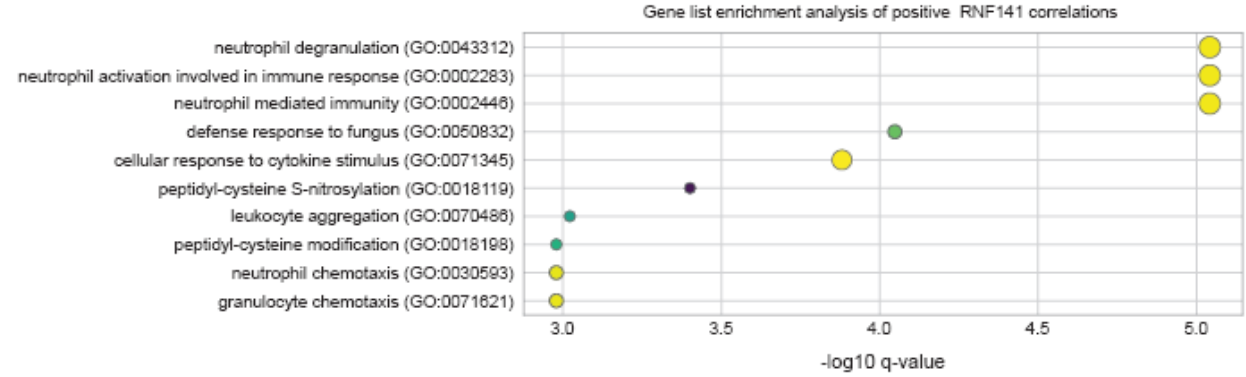
Lead variants associated to humoral traits which reach genome-wide significance threshold of ($p < 5 \times 10^{-8}$). Annotations of variants in linkage disequilibrium with the lead variant ($r2 > 0.8$) are shown.

Trait	q value	Pathway	Top Genes
ADNP PPD	1.10E-04	Lymph node development	LTB, LTA, IL15, CD248, TOX, SPNS2, NKX2-3, CXCR5, RC3H1, PDPN
FcR2b Measles	1.76E-04	Negative regulation of immature t cell proliferation	ERBB2, GNRH1
FcR2A Measles	1.92E-04	Negative regulation of t cell differentiation in thymus	ERBB2, NFKBID, PTPN2
FcR2b Post-RSV	5.60E-04	Positive regulation of antigen processing and presentation of peptide antigen	TREM2
FcR2b Ptox	6.85E-04	Innate immune response-activating signal transduction	MUC7, PSMB4, ICAM3, MUC6, CUL1, CLEC4A, FCER1G, PRKACG, NFKB1, SYK, MUC5B, RELB
IgG1 Noro	1.21E-03	Regulation of fc-gamma receptor signaling pathway involved in phagocytosis	APPL2, PTPRC, CD47, RAPGEF1, RAP1A, APPL1, PTPRJ
ADNP EBV	1.81E-03	Regulation of germinal center formation	MEF2C, TNFSF13B, PKN1
IgG1 Measles	2.20E-03	Negative regulation of immature t cell proliferation in thymus	ERBB2, TMEM131L
IgM Mumps	2.20E-03	Negative regulation of lymphocyte chemotaxis	KLRK1, KLRC4-KLRK1, PADI2, CCL2
IgA Measles	3.61E-03	Neutrophil-mediated killing of gram-negative bacterium	F2RL1, F2, ELANE
IgG1 Post-RSV	4.47E-03	Cytokine secretion involved in immune response	SPHK2, NLRP3
IgG4 Mumps	1.48E-02	Natural killer cell tolerance induction	LGALS9, HLA-E, HAVCR2
ADNP Flu	2.13E-02	Positive regulation of antigen processing and presentation of peptide antigen	PYCARD
IgM Measles	2.87E-02	Negative regulation of natural killer cell chemotaxis	KLRK1, KLRC4-KLRK1, CCL2
IgG2 HBV	3.85E-02	Negative regulation of isotype switching	THOC1, PARP3, NDFIP1, BCL6
IgM HBV	4.11E-02	Negative regulation of activation-induced cell death of t cells	GPAM, FADD

Table 6-S3. Genes Related to Top Immune Process Identified with Gene Set Analysis, related to Figure 6-5


Gene list enrichment analysis of DLC1 correlations
regulation of myeloid cell differentiation (GO:0045637)
mitotic cell cycle phase transition (GO:0044772)
mitotic spindle organization (GO:0007052)
cellular response to cytokine stimulus (GO:0071345)
positive regulation of transcription by RNA polymerase II (GO:0045944)
regulation of cell cycle (GO:0051726)
MAPK cascade (GO:0000165)
protein phosphorylation (GO:0006468)
microtubule cytoskeleton organization involved in mitosis (GO:1902850)
positive regulation of transcription, DNA-templated (GO:0045893)
3.2
3.4
3.6
3.8
4.0
4.2
4.4
4.6
-log10 q-value


Gene list enrichment analysis of negative RNF141 correlations
antigen processing and presentation of exogenous peptide antigen via MHC class II (GO:0019886)
antigen processing and presentation of peptide antigen via MHC class II (GO:0002495)
antigen processing and presentation of exogenous peptide antigen (GO:0002478)
peptide antigen assembly with MHC protein complex (GO:0002501)
positive regulation of immune effector process (GO:0002699)
neutrophil degranulation (GO:0043312)
neutrophil activation involved in immune response (GO:0002283)
neutrophil mediated immunity (GO:0002446)
response to unfolded protein (GO:0006986)
regulation of T cell activation via T cell receptor contact with antigen bound to MHC molecule on antigen presenting cell (GO:2001188)
6
8
-log10 q-value


Gene list enrichment analysis of positive RNF141 correlations
neutrophil degranulation (GO:0043312)
neutrophil activation involved in immune response (GO:0002283)
neutrophil mediated immunity (GO:0002446)
defense response to fungus (GO:0050832)
cellular response to cytokine stimulus (GO:0071345)
peptidyl-cysteine S-nitrosylation (GO:0018119)
leukocyte aggregation (GO:0070486)
peptidyl-cysteine modification (GO:0018198)
neutrophil chemotaxis (GO:0030593)
granulocyte chemotaxis (GO:0071621)
3.0
3.5
4.0
4.5
5.0
-log10 q-value


Gene list enrichment analysis of negative SPRY2 correlations
SRP-dependent cotranslational protein targeting to membrane (GO:0006614)
cytoplasmic translation (GO:0002181)
cotranslational protein targeting to membrane (GO:0006613)
protein targeting to ER (GO:0045047)
nuclear-transcribed mRNA catabolic process, nonsense-mediated decay (GO:0000184)
peptide biosynthetic process (GO:0043043)
nuclear-transcribed mRNA catabolic process (GO:0000956)
translation (GO:0006412)
cellular macromolecule biosynthetic process (GO:0034645)
gene expression (GO:0010467)
10.0
10.5
11.0
11.5
12.0
12.5
13.0
13.5
-log10 q-value

Figure 6-S2. Gene List Enrichment Analysis of Co-expressed Genes Related to Humoral Traits, related to Figure 6-6

Gene list enrichment analysis performed on positive and negative co-expressed genes which were significantly correlated to candidate GWAS genes. The color of the dots corresponds to the combined score, the size of the dot the number of genes related to the process and the negative log10 q-values are depicted on the x-axis.

Chr	Pos	rsID	EA	NEA	Traits	Gene	Overlapping Phenotypes
19	49206108	rs679574	C	G	IgG1 Post-RSV and FcgR Post-RSV	FUT2	Chronic inflammatory diseases (ankylosing spondylitis, Crohns disease, psoriasis, primary sclerosing cholangitis, ulcerative colitis) (pleiotropy),Blood protein levels,LDL cholesterol,Human milk oligosaccharide concentration (2-fucosyllactose)
19	49206145	rs516316	G	C	IgG1 Post-RSV and FcgR Post-RSV	FUT2	Mumps,Blood protein levels,LDL cholesterol levels,Apolipoprotein B levels,Low density lipoprotein cholesterol levels
19	49206172	rs516246	C	T	IgG1 Post-RSV and FcgR Post-RSV	FUT2	Type 1 diabetes,Birth weight,Crohns disease,Blood protein levels,Medication use (agents acting on the renin-angiotensin system),Obesity-related traits,Triglycerides,Cardiovascular disease,Liver enzyme levels (gamma-glutamyl transferase),Inflammatory bowel disease,Human milk oligosaccharide concentration (lacto-N-fucopentaose II),Human milk oligosaccharide concentration (fucodisialyllacto-N-hexaose),Low density lipoprotein cholesterol levels,Total cholesterol levels
19	49206417	rs492602	A	G	IgG1 Post-RSV and FcgR Post-RSV	FUT2	Mammographic density (dense area),Lung function (FEV1/FVC),Tissue factor levels,Vitamin B levels in ischemic stroke,Problematic alcohol use,BK polyomavirus VP1 antibody levels,Cholesterol, total,Total cholesterol levels,Human milk oligosaccharide concentration (lacto-N-fucopentaose I),LDL cholesterol levels,Blood protein levels,Vitamin B12 levels,Serum metabolite levels,Psoriasis,Serum alkaline phosphatase levels,Alcohol use disorder (total score),Triglycerides,Low density lipoprotein cholesterol levels,Blood protein levels in cardiovascular risk
19	49206462	rs681343	C	T	IgG1 Post-RSV and FcgR Post-RSV	FUT2	Lung adenocarcinoma,BK polyomavirus VP1 antibody levels,Primary sclerosing cholangitis,Behcets disease,Blood protein levels,Peptic ulcer disease,Childhood ear infection,Mitochondrial heteroplasmy measurement
19	49206674	rs601338	G	A	IgG1 Post-RSV and FcgR Post-RSV	FUT2	Serum cancer antigen 19.9 levels,Serum carcinoembryonic antigen levels,Hemoglobin levels,C-reactive protein levels,Body size at age 10,Gallstone disease,CCL25 levels,Number of common colds,Blood protein levels,Iron status biomarkers (ferritin levels),Blood metabolite levels,Human milk oligosaccharide concentration (difucosyllactose),C-reactive protein,Blood metabolite ratios,Matrix metalloproteinase-10 levels,Medication use (HMG CoA reductase inhibitors)

Table 6-S4. Overlap of Variants with Disease Outcomes and Immune Related Phenotypes, related to Figure 6-6

6.7 Acknowledgements

We acknowledge support from the Ragon Institute of MGH, MIT, the Massachusetts.

Chapter 7

Conclusions and future perspectives

Humoral immunity is comprised of a complex network of polyclonal antibodies which can vary in both antigen specificity and how they drive effector functions. This thesis described the development of experimental and computational tools to understand how polyclonal antibody composition shapes aspects of human health. This was achieved by developing a unique antibody dependent primary monocyte effector assay, machine learning approaches and novel computational methods for uncovering mechanisms of protection against infectious diseases and evaluating vaccine strategies, and characterizing the genetic architecture of pathogen-specific polyclonal antibody characteristics and functions. These works have collectively led to the discovery of genetic variants that control the development of antibody characteristics and effector functions, aberrant immunity linked to SARS-CoV-2 mortality, and specific vaccine platforms that drive tissue-specific immunity. From a biological perspective much is still unknown regarding the cellular decisions responsible for differences observed in subpopulations of individuals and across different pathogens. In order to gain a more complete picture, development of modelling approaches for integrating multi-omics and antibody data is still needed. This would allow to further characterize the cellular B cell mechanisms that are responsible for generating specific downstream polyclonal pools. Regardless, the works presented here aided in improving the understanding of how polyclonal antibody composition impacts human health and the factors responsible for their development. This has provided a foundational framework which could be used to improve the design of future vaccines and informs the clinical management against specific infectious diseases.

7.1 Extending the ADMP assay

In chapter two of this thesis the development of an antibody dependent primary monocyte phagocytosis assay (ADMP) was described. This assay provides several distinct advantages over the traditional THP-1 cell based assay including enhanced sensitivity and greater dynamic range, and the ability to capture changes in surface markers and cytokine profiles that more accurately represent mechanisms *in situ*. In addition to improved sensitivity and the ability to describe downstream consequences of phagocytosis, the use of primary monocytes also allows for the investigation of disease-associated states and whether certain perturbations alter responses to immune complexes. Perturbed monocytic states have been previously associated with many diseases[45,90,91], and although donor variability was low among healthy donors (Fig. 2-1F and Fig. 2-S1D-E) ADMP assays performed with perturbed monocytic cell states could provide different results. Therefore, monocytes isolated from donors which exhibit disease-specific characteristics could uniquely respond to immune complexes and thus provide additional novel mechanistic insights into disease specific immunopathology.

At present, the ADMP assay readouts are comprised of signals formed by integrating information across all monocyte subsets. Monocyte subsets are divided into three major populations: classical, non-classical, and intermediate. Each subset is defined by CD14 and CD16 expression and is responsible for different functions. Classical monocytes are the dominant subset and are thought to be primed for phagocytosis. Alternatively, intermediate monocytes have been suggested to preferentially participate in antigen presentation, and non-classical monocytes in Fc-driven phagocytosis and anti-viral immunity[91]. Yet, how the composition of polyclonal antibodies and Fc structure impacts the downstream phenotypes and cytokine secretion of each subset has yet to be investigated. This could help better inform how these subsets cooperate to collectively respond to infection and would provide further insight into how the expansion of particular subsets relates to various inflammatory diseases[319].

Additionally, the ADMP assay being a bead-based assay allows for the multiplexing of antigens[98]. While multi-bead uptake has been shown to recapitulate single-plex signals, offering a sample-sparing approach to perform multiplexed ADCP assays, competition assays with antigens of interest are also possible. Therefore, the ADMP assay may be further modified to capture individual antigen responses, multiplexed activity, and even competitive responses. Moreover, the assay may be modified to used cryo-preserved monocytes, that if thawed properly monocyte viability, phagocytic activity, morphology, migration, adherence, and other functions minimally impact monocyte activity[99–103]. Therefore, the use of cryopreserved monocytes could provide additional flexibility when fresh monocytes are not available such as in the case of when profiling perturbed monocytes from non-healthy donors.

7.2 Upper and lower respiratory correlates of humoral protection across other respiratory viruses

In chapter three machine learning approaches and development of computational methods were used to uncover differences in upper and lower correlates of protection against RSV based on different vaccine strategies. RSV infection initially begins in the upper respiratory tract, and ultimately migrates into the lower respiratory tract where it can cause severe disease[140]. The upper and lower respiratory tracts are composed of distinct cellular components and represent unique immunological landscapes[141]. Thus, the immune correlates of protection against RSV infection may vary across different compartments, and thus unique antibody functions maybe required in order to restrict protection upon entry than at the site of viral replication and disease. This is consistent with our findings which highlighted neutralization, IgA levels, and ADCD as correlates of upper respiratory protection, and IgA and ADNP function as correlates of protection in the lower respiratory tract.

Beyond RSV, other respiratory viruses feature similar pathologies. For example, SARS-CoV-2 virus also infects both upper and lower respiratory tracts. Interestingly, the SARS-CoV-2 B.1.1.529, omicron variant, has been shown to preferentially target the upper respiratory tract and result in reduced lethality compared to the delta variant[320]. Yet, the antibody correlates of protection in the lower and upper respiratory tracts against SARS-CoV-2 infection are currently unknown. Therefore, given our previous findings and the reported differences in pathology across SARS-CoV-2 variants, it is likely that extending a similar analysis to study the correlates of compartment-specific protection against SARS-CoV-2 could help further elucidate the etiology of COVID-19. Additionally, defining if the same vaccine platforms induce similar compartment-specific correlates against SARS-CoV-2 could help improve vaccine design especially in the wake of variant immune evasion.

7.3 Approaches for expanding the longitudinal modelling framework

In chapter five, a numerical algorithm was developed in order to temporally model humoral development in a cohort of SARS-CoV-2 infected individuals. To determine differences in dynamics between groups, a four-parameter logistic curve was used and is modelled on the group-level. For a given feature, multiple models are built that describe dynamics of both groups simultaneously, allowing for combinations of parameters to differ while constraining others to be the same. Therefore with 4 parameters, there are 16 possible combinations/models that can be used to potentially explain the dynamics of a feature. Each model is fitted using maximum likelihood estimation, treating each measurement as independent and assuming that differences in measurements are a result of noise. Thus, this framework is best suited for cohorts for which variation is not driven by confounders, and is constrained to study designs without any multi-level structure.

For such applications, a customized nonlinear mixed-effects model could be developed to longitudinally model the dynamics of each individual. This would further enable the ability to cross validate results and thus improve generalizability. Taking this development one step further, the estimation of models which integrate parameters across multiple features would enhance the understanding of whether of the specific aspects of humoral development across multiple features best describes differences among groups. This would thus be representative a multivariate nonlinear mixed effect model.

Both the current and potential future frameworks are well suited for numerous additional applications beyond the one mentioned in this thesis. For example, little is known about the humoral development of antibody characteristics and functions against other infectious diseases. Therefore, the current framework could be easily applied to other pathogens to understand differences between susceptible and resistant subpopulations. Additionally, nonlinear mixed effect models could be used to profile differences in dynamics while accounting for multi-level structure like in the case of maternal-fetal transfer studies and the evaluation of multiple different vaccine designs.

7.4 Methods for validating genetic mechanisms

In chapters six, a genomic pipeline was developed to screen genetic variants which associated to pathogen-specific polyclonal antibody characteristics and functions, and was then functionally validated with transcriptomic data and pathway analysis. This led to the discovery of five significant genomic regions and multiple immune processes which related to humoral phenotypes. These results were then used to carefully construct mechanistic hypotheses. The goal of this study was to propose mechanisms which could then be further validated in future studies. The five mechanisms which were highlighted in the study were not B cell specific but rather involved innate immune and T cells or were related to specific host-pathogen interactions. This indicated that the factors which affect polyclonal antibody properties are more likely due to changes in the external stimuli or environment that B cells experience.

To validate the functional and regulatory consequences of our top five GWAS hits functional genomics methods such as Hi-C for chromatin structure, ChIP-seq for TF binding to DNA, and more can be employed. Then direct effects like change in expression and certain cellular phenotypes (e.g. migration, and antigen presentation) can be tested through genetic perturbations of a cell type *in vitro*. However, the overall induction of humoral immunity involves many cells and secondary interactions. Therefore, to properly evaluate the effects of certain alleles *in vivo* transgenic mouse models are needed. Thus, first a mouse line would be generated to contain our germline mutations with our variant of interest. Then both the transgenic mice and healthy controls would be challenged with the respective pathogen. This allows to investigate associated pathologies and many additional aspects not possible with human donors.

7.5 Enhancing mechanistic insights by multi-omics integration

The majority of SNPs that are identified during GWAS are located in non-coding regions of the genome[321]. In chapter six, several of the SNPs reaching genome-wide significance were located in intergenic regions. Regulatory SNPs can control the downstream expression of genes and ultimately impact phenotypes[322]. SNPs can be located within regulatory elements such as promoters, enhancers, and nuclear structural elements such as CCCTC-binding factors (CTCF). For example, variant rs2379 which was significantly associated with ADNP functional antibodies was located in a CTCF binding region. They can also alter CpG DNA methylation (DNAm) patterns, and has been linked to various diseases[323,324]. In our study, variant rs10131536 which was associated to IgG4 against Mumps was located at a site of h3k4me1 found on T helper cells. Therefore, it is possible that genetic variants located in sites of potential DNAm can impact the composition of polyclonal pools by altering cell function and in turn affect disease outcomes.

Multiple studies over the past few years have uncovered that loci specific DNAm are genetically influenced[325]. DNAm occurs at CpG sites of the genome, and the majority of unmethylated CpGs are within CpG islands (CGI) which are repeats in close proximity, which have been shown to suppress the expression of nearby genes[326]. Interestingly, cis (proximal) SNPs have been estimated to account for ~60% of DNAm variability[327]. These regulatory SNPs are known as methylation quantitative loci (me-QTL) and have been associated with various diseases, processes, and tissue types[328]. me-QTLs are calculated by using linear mixed models to observe whether particular SNPs track with DNAm within a ±1-1000kb range. It can often be the case that a single me-QTL results in multiple cis CpG methylation events[329]. However, due to LD, linking causal SNPs to methylation events is difficult, but Bayesian hierarchal models can be fitted to infer putative causality[330]. In addition, the causality of me-QTLs can be quantified through a two-sample mendelian randomization process. Beyond, integrating methylation data other functional genomics data such as mRNA co-expression and protein-protein interaction data could also further help improve the mechanistic resolution of how our GWAS hits control pathogen-specific antibody characteristics and functions.

References

1. Sironi, M., Cagliani, R., Forni, D. & Clerici, M. Evolutionary insights into host–pathogen interactions from mammalian sequence data. *Nat. Rev. Genet.* **16**, 224–236 (2015).
2. World Health Organization. Vaccines and Diseases. https://www.who.int/immunization/diseases/en/.
3. World Health Organization. The top 10 causes of death. https://www.who.int/en/news-room/fact-sheets/detail/the-top-10-causes-of-death (24 May2018).
4. Sharma, M., Krammer, F., García-Sastre, A. & Tripathi, S. Moving from Empirical to Rational Vaccine Design in the 'Omics' Era. *Vaccines* **7**, 89 (2019).
5. Plotkin, S. A. Vaccines: correlates of vaccine-induced immunity. *Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am.* **47**, 401–409 (2008).
6. Plotkin, S. A. Correlates of protection induced by vaccination. *Clin. Vaccine Immunol. CVI* **17**, 1055–1065 (2010).
7. Forthal, D. N., Landucci, G. & Daar, E. S. Antibody from patients with acute human immunodeficiency virus (HIV) infection inhibits primary strains of HIV type 1 in the presence of natural-killer effector cells. *J. Virol.* **75**, 6953–6961 (2001).
8. Barouch, D. H. *et al.* Protective Efficacy of a Global HIV-1 Mosaic Vaccine against Heterologous SHIV Challenges in Rhesus Monkeys. *Cell* **155**, 531–539 (2013).
9. Bournazos, S. *et al.* Broadly neutralizing anti-HIV-1 antibodies require Fc effector functions for in vivo activity. *Cell* **158**, 1243–1253 (2014).
10. Hessell, A. J. *et al.* Fc receptor but not complement binding is important in antibody protection against HIV. *Nature* **449**, 101–104 (2007).
11. Alter, G. & Barouch, D. Immune Correlate-Guided HIV Vaccine Design. *Cell Host Microbe* **24**, 25–33 (2018).
12. Jegerlehner, A., Schmitz, N., Storni, T. & Bachmann, M. F. Influenza A Vaccine Based on the Extracellular Domain of M2: Weak Protection Mediated via Antibody-Dependent NK Cell Activity. *J. Immunol.* **172**, 5598–5605 (2004).
13. DiLillo, D. J., Tan, G. S., Palese, P. & Ravetch, J. V. Broadly neutralizing hemagglutinin stalk–specific antibodies require FcγR interactions for protection against influenza virus in vivo. *Nat. Med.* **20**, 143–151 (2014).
14. Kohl, S., Loo, L. S. & Pickering, L. K. Protection of neonatal mice against herpes simplex viral infection by human antibody and leukocytes from adult, but not neonatal humans. *J. Immunol.* **127**, 1273–1275 (1981).
15. Kohl, S. & Loo, L. S. Protection of neonatal mice against herpes simplex virus infection: probable in vivo antibody-dependent cellular cytotoxicity. *J. Immunol.* **129**, 370–376 (1982).
16. Warfield, K. L. *et al.* Ebola Virus-Like Particle-Based Vaccine Protects Nonhuman Primates against Lethal Ebola Virus Challenge. *J. Infect. Dis.* **196**, S430–S437 (2007).
17. Shi, J., Mcintosh, R. S. & Pleass, R. J. Antibody- and Fc-receptor-based therapeutics for malaria. *Clin. Sci.* **110**, 11–19 (2006).
18. Bouharoun-Tayoun, H., Oeuvray, C., Lunel, F. & Druilhe, P. Mechanisms underlying the monocyte-mediated antibody-dependent killing of Plasmodium falciparum asexual blood stages. *J. Exp. Med.* **182**, 409–418 (1995).

19. Wang, T. T. *et al.* Anti-HA Glycoforms Drive B Cell Affinity Selection and Determine Influenza Vaccine Efficacy. *Cell* **162**, 160–169 (2015).
20. Wang, T. T. *et al.* IgG antibodies to dengue enhanced for FcγRIIIA binding determine disease severity. *Science* **355**, 395–398 (2017).
21. Jennewein, M. F. & Alter, G. The Immunoregulatory Roles of Antibody Glycosylation. *Trends Immunol.* **38**, 358–372 (2017).
22. Karsten, C. B. *et al.* A versatile high-throughput assay to characterize antibody-mediated neutrophil phagocytosis. *J. Immunol. Methods* **471**, 46–56 (2019).
23. Butler, A. L., Fallon, J. K. & Alter, G. A Sample-Sparing Multiplexed ADCP Assay. *Front. Immunol.* **10**, (2019).
24. Fischinger, S. *et al.* A high-throughput, bead-based, antigen-specific assay to assess the ability of antibodies to induce complement activation. *J. Immunol. Methods* **473**, 112630 (2019).
25. Perez, L. G. *et al.* V1V2-specific complement activating serum IgG as a correlate of reduced HIV-1 infection risk in RV144. *PLOS ONE* **12**, e0180720 (2017).
26. Brown, E. P. *et al.* High-throughput, multiplexed IgG subclassing of antigen-specific antibodies from clinical samples. *J. Immunol. Methods* **386**, 117–123 (2012).
27. Tonegawa, S. Somatic generation of antibody diversity. *Nature* **302**, 575–581 (1983).
28. Briney, B., Inderbitzin, A., Joyce, C. & Burton, D. R. Commonality despite exceptional diversity in the baseline human antibody repertoire. *Nature* **566**, 393–397 (2019).
29. Allen, C. D. C. & Cyster, J. G. Follicular dendritic cell networks of primary follicles and germinal centers: Phenotype and function. *Semin. Immunol.* **20**, 14–25 (2008).
30. Cyster, J. G. & Allen, C. D. C. B Cell Responses: Cell Interaction Dynamics and Decisions. *Cell* **177**, 524–540 (2019).
31. Georgiou, G. *et al.* The promise and challenge of high-throughput sequencing of the antibody repertoire. *Nat. Biotechnol.* **32**, 158–168 (2014).
32. Miles, J. J., Douek, D. C. & Price, D. A. Bias in the αβ T-cell repertoire: implications for disease pathogenesis and vaccination. *Immunol. Cell Biol.* **89**, 375–387 (2011).
33. Hansen, T. Ø., Lange, A. B. & Barington, T. Sterile DJH Rearrangements Reveal that Distance Between Gene Segments on the Human Ig H Chain Locus Influences Their Ability To Rearrange. *J. Immunol.* **194**, 973–982 (2015).
34. Jackson, K. J. L., Kidd, M. J., Wang, Y. & Collins, A. M. The Shape of the Lymphocyte Receptor Repertoire: Lessons from the B Cell Receptor. *Front. Immunol.* **4**, (2013).
35. Glanville, J. *et al.* Naive antibody gene-segment frequencies are heritable and unaltered by chronic lymphocyte ablation. *Proc. Natl. Acad. Sci.* **108**, 20066–20071 (2011).
36. Wang, C. *et al.* B-cell repertoire responses to varicella-zoster vaccination in human identical twins. *Proc. Natl. Acad. Sci.* **112**, 500–505 (2015).
37. Greiff, V. *et al.* A bioinformatic framework for immune repertoire diversity profiling enables detection of immunological status. *Genome Med.* **7**, 49 (2015).
38. Dunand, C. J. H. & Wilson, P. C. Restricted, canonical, stereotyped and convergent immunoglobulin responses. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, 20140238 (2015).
39. Rubelt, F. *et al.* Individual heritable differences result in unique cell lymphocyte receptor repertoires of naïve and antigen-experienced cells. *Nat. Commun.* **7**, 1–12 (2016).

40. Paull, M. L., Johnston, T., Ibsen, K. N., Bozekowski, J. D. & Daugherty, P. S. A general approach for predicting protein epitopes targeted by antibody repertoires using whole proteomes. *PLOS ONE* **14**, e0217668 (2019).
41. Raybould, M. I. J., Wong, W. K. & Deane, C. M. Antibody–antigen complex modelling in the era of immunoglobulin repertoire sequencing. *Mol. Syst. Des. Eng.* **4**, 679–688 (2019).
42. Brodin, P. *et al.* Variation in the Human Immune System Is Largely Driven by Non-Heritable Influences. *Cell* **160**, 37–47 (2015).
43. Shen, X. *et al.* Multivariate discovery and replication of five novel loci associated with Immunoglobulin G N-glycosylation. *Nat. Commun.* **8**, 447 (2017).
44. Tigner, A., Ibrahim, S. A. & Murray, I. Histology, White Blood Cell. in *StatPearls* (StatPearls Publishing, 2021).
45. Karlmark, K. R., Tacke, F. & Dunay, I. R. Monocytes in health and disease – Minireview. *Eur. J. Microbiol. Immunol.* **2**, 97–102 (2012).
46. Jakubzick, C. *et al.* Minimal differentiation of classical monocytes as they survey steady-state tissues and transport antigen to lymph nodes. *Immunity* **39**, 599–610 (2013).
47. Uribe-Querol, E. & Rosales, C. Control of Phagocytosis by Microbial Pathogens. *Front. Immunol.* **8**, (2017).
48. Karaji, N. & Sattentau, Q. J. Efferocytosis of Pathogen-Infected Cells. *Front. Immunol.* **8**, (2017).
49. Jakubzick, C. V., Randolph, G. J. & Henson, P. M. Monocyte differentiation and antigen-presenting functions. *Nat. Rev. Immunol.* **17**, 349–362 (2017).
50. Arango Duque, G. & Descoteaux, A. Macrophage Cytokines: Involvement in Immunity and Infectious Diseases. *Front. Immunol.* **5**, (2014).
51. Serbina, N. V., Jia, T., Hohl, T. M. & Pamer, E. G. Monocyte-Mediated Defense Against Microbial Pathogens. *Annu. Rev. Immunol.* **26**, 421–452 (2008).
52. Salazar-Mather, T. P., Orange, J. S. & Biron, C. A. Early murine cytomegalovirus (MCMV) infection induces liver natural killer (NK) cell inflammation and protection through macrophage inflammatory protein 1alpha (MIP-1alpha)-dependent pathways. *J. Exp. Med.* **187**, 1–14 (1998).
53. Dawson, T. C., Beck, M. A., Kuziel, W. A., Henderson, F. & Maeda, N. Contrasting effects of CCR5 and CCR2 deficiency in the pulmonary inflammatory response to influenza A virus. *Am. J. Pathol.* **156**, 1951–1959 (2000).
54. Hohl, T. M. *et al.* Inflammatory Monocytes Facilitate Adaptive CD4 T Cell Responses during Respiratory Fungal Infection. *Cell Host Microbe* **6**, 470–481 (2009).
55. Ortega-Pajares, A. & Rogerson, S. J. The Rough Guide to Monocytes in Malaria Infection. *Front. Immunol.* **9**, (2018).
56. Dorhoi, A. & Du Plessis, N. Monocytic Myeloid-Derived Suppressor Cells in Chronic Infections. *Front. Immunol.* **8**, (2018).
57. Wong, M. E., Jaworowski, A. & Hearps, A. C. The HIV Reservoir in Monocytes and Macrophages. *Front. Immunol.* **10**, (2019).
58. Zohar, T. *et al.* Compromised Humoral Functional Evolution Tracks with SARS-CoV-2 Mortality. *Cell* **183**, 1508-1519.e12 (2020).
59. Brown, E. J. Complement receptors and phagocytosis. *Curr. Opin. Immunol.* **3**, 76–82 (1991).

60. Nimmerjahn, F. & Ravetch, J. V. Fcγ Receptors: Old Friends and New Family Members. *Immunity* **24**, 19–28 (2006).
61. Uribe-Querol, E. & Rosales, C. Phagocytosis: Our Current Understanding of a Universal Biological Process. *Front. Immunol.* **11**, (2020).
62. Webster, S. J. *et al.* Distinct Cell Death Programs in Monocytes Regulate Innate Responses Following Challenge with Common Causes of Invasive Bacterial Disease. *J. Immunol.* **185**, 2968–2979 (2010).
63. Nimmerjahn, F. & Ravetch, J. V. Fcγ receptors as regulators of immune responses. *Nat. Rev. Immunol.* **8**, 34–47 (2008).
64. Bournazos, S., Wang, T. T. & Ravetch, J. V. The Role and Function of Fcγ Receptors on Myeloid Cells. *Microbiol. Spectr.* **4**, (2016).
65. Rosales, C. Fcγ Receptor Heterogeneity in Leukocyte Functional Responses. *Front. Immunol.* **8**, (2017).
66. 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).
67. Tsuchiya, S. *et al.* Establishment and characterization of a human acute monocytic leukemia cell line (THP-1). *Int. J. Cancer* **26**, 171–176 (1980).
68. Ackerman, M. E. *et al.* Route of immunization defines multiple mechanisms of vaccine-mediated protection against SIV. *Nat. Med.* **24**, 1590–1598 (2018).
69. Ackerman, M. E. *et al.* Polyfunctional HIV-Specific Antibody Responses Are Associated with Spontaneous HIV Control. *PLOS Pathog.* **12**, e1005315 (2016).
70. Ana-Sosa-Batiz, F. *et al.* Influenza-Specific Antibody-Dependent Phagocytosis. *PLOS ONE* **11**, e0154461 (2016).
71. Donnenberg, V. S., Corselli, M., Normolle, D. P., Meyer, E. M. & Donnenberg, A. D. Flow cytometric detection of most proteins in the cell surface proteome is unaffected by trypsin treatment. *Cytometry A* **93**, 803–810 (2018).
72. Yu, K. K. Q. *et al.* Comorbid illnesses are associated with altered adaptive immune responses to SARS-CoV-2. *JCI Insight* **6**, (2021).
73. Gunn, B. M. *et al.* A Role for Fc Function in Therapeutic Monoclonal Antibody-Mediated Protection against Ebola Virus. *Cell Host Microbe* **24**, 221-233.e5 (2018).
74. Atyeo, C. *et al.* Dissecting strategies to tune the therapeutic potential of SARS-CoV-2-specific monoclonal antibody CR3022. *JCI Insight* **6**, (2021).
75. Dunn, P. A. & Tyrer, H. W. Quantitation of neutrophil phagocytosis, using fluorescent latex beads. Correlation of microscopy and flow cytometry. *J. Lab. Clin. Med.* **98**, 374–381 (1981).
76. Tian, X. *et al.* Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody. *Emerg. Microbes Infect.* **9**, 382–385 (2020).
77. Kay, H. D., Bonnard, G. D., West, W. H. & Herberman, R. B. A functional comparison of human Fc-receptor-bearing lymphocytes active in natural cytotoxicity and antibody-dependent cellular cytotoxicity. *J. Immunol. Baltim. Md 1950* **118**, 2058–2066 (1977).
78. Cooper, N. R., Jensen, F. C., Welsh, R. M. & Oldstone, M. B. Lysis of RNA tumor viruses by human serum: direct antibody-independent triggering of the classical complement pathway. *J. Exp. Med.* **144**, 970–984 (1976).
79. Petty, H. R., Hafeman, D. G. & McConnell, H. M. Specific antibody-dependent phagocytosis of lipid vesicles by RAW264 macrophages results in the loss of cell surface Fc but not C3b receptor activity. *J. Immunol. Baltim. Md 1950* **125**, 2391–2396 (1980).

80. Bonsignori, M. *et al.* Antibody-Dependent Cellular Cytotoxicity-Mediating Antibodies from an HIV-1 Vaccine Efficacy Trial Target Multiple Epitopes and Preferentially Use the VH1 Gene Family. *J. Virol.* **86**, 11521–11532 (2012).
81. Jegaskanda, S., Weinfurter, J. T., Friedrich, T. C. & Kent, S. J. Antibody-Dependent Cellular Cytotoxicity Is Associated with Control of Pandemic H1N1 Influenza Virus Infection of Macaques. *J. Virol.* **87**, 5512–5522 (2013).
82. Lu, L. L. *et al.* A Functional Role for Antibodies in Tuberculosis. *Cell* **167**, 433-443.e14 (2016).
83. Richards, J. O. *et al.* Optimization of antibody binding to FcγRIIa enhances macrophage phagocytosis of tumor cells. *Mol. Cancer Ther.* **7**, 2517–2527 (2008).
84. Boyerinas, B. *et al.* Antibody-Dependent Cellular Cytotoxicity Activity of a Novel Anti–PD-L1 Antibody Avelumab (MSB0010718C) on Human Tumor Cells. *Cancer Immunol. Res.* **3**, 1148–1157 (2015).
85. Wang, W., Erbe, A. K., Hank, J. A., Morris, Z. S. & Sondel, P. M. NK Cell-Mediated Antibody-Dependent Cellular Cytotoxicity in Cancer Immunotherapy. *Front. Immunol.* **6**, (2015).
86. Bosshart, H. & Heinzelmann, M. THP-1 cells as a model for human monocytes. *Ann. Transl. Med.* **4**, (2016).
87. Forrester, M. A. *et al.* Similarities and differences in surface receptor expression by THP-1 monocytes and differentiated macrophages polarized using seven different conditioning regimens. *Cell. Immunol.* **332**, 58–76 (2018).
88. Ostrowski, P. P., Grinstein, S. & Freeman, S. A. Diffusion Barriers, Mechanical Forces, and the Biophysics of Phagocytosis. *Dev. Cell* **38**, 135–146 (2016).
89. Swanson, J. A. & Hoppe, A. D. The coordination of signaling during Fc receptor-mediated phagocytosis. *J. Leukoc. Biol.* **76**, 1093–1103 (2004).
90. Yang, J., Zhang, L., Yu, C., Yang, X.-F. & Wang, H. Monocyte and macrophage differentiation: circulation inflammatory monocyte as biomarker for inflammatory diseases. *Biomark. Res.* **2**, 1 (2014).
91. Kapellos, T. S. *et al.* Human Monocyte Subsets and Phenotypes in Major Chronic Inflammatory Diseases. *Front. Immunol.* **10**, (2019).
92. Weiner, L. M., Surana, R. & Wang, S. Monoclonal antibodies: versatile platforms for cancer immunotherapy. *Nat. Rev. Immunol.* **10**, 317–327 (2010).
93. The IMpact-RSV Study Group. Palivizumab, a Humanized Respiratory Syncytial Virus Monoclonal Antibody, Reduces Hospitalization From Respiratory Syncytial Virus Infection in High-risk Infants. *Pediatrics* **102**, 531–537 (1998).
94. Mulangu, S. *et al.* A Randomized, Controlled Trial of Ebola Virus Disease Therapeutics. *N. Engl. J. Med.* **381**, 2293–2303 (2019).
95. Taylor, P. C. *et al.* Neutralizing monoclonal antibodies for treatment of COVID-19. *Nat. Rev. Immunol.* **21**, 382–393 (2021).
96. Lazar, G. A. *et al.* Engineered antibody Fc variants with enhanced effector function. *Proc. Natl. Acad. Sci.* **103**, 4005–4010 (2006).
97. Moore, G. L., Chen, H., Karki, S. & Lazar, G. A. Engineered Fc variant antibodies with enhanced ability to recruit complement and mediate effector functions. *mAbs* **2**, 181–189 (2010).
98. Butler, A. L., Fallon, J. K. & Alter, G. A Sample-Sparing Multiplexed ADCP Assay. *Front. Immunol.* **10**, (2019).

99. De Boer, M., Reijneke, R., Van De Griend, RenéJ., Loos, J. A. & Roos, D. Large-scale purification and cryopreservation of human monocytes. *J. Immunol. Methods* **43**, 225–239 (1981).
100. Myhrvold, V. & Mørland, B. The use of frozen monocytes in phagocytosis studies. *Acta Pathol. Microbiol. Immunol. Scand. [C]* **93**, 43–48 (1985).
101. Hansen, J. B. *et al.* Retention of phagocytic functions in cryopreserved human monocytes. *J. Leukoc. Biol.* **57**, 235–241 (1995).
102. Hori, S. *et al.* Freeze-thawing procedures have no influence on the phenotypic and functional development of dendritic cells generated from peripheral blood CD14+ monocytes. *J. Immunother. Hagerstown Md 1997* **27**, 27–35 (2004).
103. Pardali, E. *et al.* Cryopreservation of primary human monocytes does not negatively affect their functionality or their ability to be labelled with radionuclides: basis for molecular imaging and cell therapy. *EJNMMI Res.* **6**, 77 (2016).
104. Boonyaratanakornkit, J. *et al.* Clinical, laboratory, and temporal predictors of neutralizing antibodies against SARS-CoV-2 among COVID-19 convalescent plasma donor candidates. *J. Clin. Invest.* **131**, (2021).
105. Lazar, G. A. *et al.* Engineered antibody Fc variants with enhanced effector function. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 4005–4010 (2006).
106. Leabman, M. K. *et al.* Effects of altered FcγR binding on antibody pharmacokinetics in cynomolgus monkeys. *mAbs* **5**, 896–903 (2013).
107. Ojala, M. & Garriga, G. C. Permutation Tests for Studying Classifier Performance. in *2009 Ninth IEEE International Conference on Data Mining* 908–913 (IEEE, 2009). doi:10.1109/ICDM.2009.108.
108. Nair, H. *et al.* Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis. *The Lancet* **375**, 1545–1555 (2010).
109. Smyth, R. L. & Openshaw, P. J. Bronchiolitis. *The Lancet* **368**, 312–322 (2006).
110. Lozano, R. *et al.* Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet* **380**, 2095–2128 (2012).
111. Shi, T. *et al.* Global Disease Burden Estimates of Respiratory Syncytial Virus–Associated Acute Respiratory Infection in Older Adults in 2015: A Systematic Review and Meta-Analysis. *J. Infect. Dis.* (2019) doi:10.1093/infdis/jiz059.
112. Claydon, J. *et al.* Respiratory syncytial virus-neutralizing serum antibody titers in infants following palivizumab prophylaxis with an abbreviated dosing regimen. *PLOS ONE* **12**, e0176152 (2017).
113. Piedra, P. A., Jewell, A. M., Cron, S. G., Atmar, R. L. & Glezen, W. P. Correlates of immunity to respiratory syncytial virus (RSV) associated-hospitalization: establishment of minimum protective threshold levels of serum neutralizing antibodies. *Vaccine* **21**, 3479–3482 (2003).
114. Barouch, D. H. *et al.* Protective efficacy of a global HIV-1 mosaic vaccine against heterologous SHIV challenges in rhesus monkeys. *Cell* **155**, 531–539 (2013).
115. Bournazos, S. *et al.* Broadly neutralizing anti-HIV-1 antibodies require Fc effector functions for in vivo activity. *Cell* **158**, 1243–1253 (2014).

116. DiLillo, D. J., Tan, G. S., Palese, P. & Ravetch, J. V. Broadly neutralizing hemagglutinin stalk–specific antibodies require FcγR interactions for protection against influenza virus in vivo. *Nat. Med.* **20**, 143–151 (2014).
117. Kohl, S. & Loo, L. S. Protection of neonatal mice against herpes simplex virus infection: probable in vivo antibody-dependent cellular cytotoxicity. *J. Immunol.* **129**, 370–376 (1982).
118. Yu, J. *et al.* DNA vaccine protection against SARS-CoV-2 in rhesus macaques. *Science* (2020) doi:10.1126/science.abc6284.
119. van Erp, E. A., Luytjes, W., Ferwerda, G. & van Kasteren, P. B. Fc-Mediated Antibody Effector Functions During Respiratory Syncytial Virus Infection and Disease. *Front. Immunol.* **10**, (2019).
120. Welliver, T. P. *et al.* Severe human lower respiratory tract illness caused by respiratory syncytial virus and influenza virus is characterized by the absence of pulmonary cytotoxic lymphocyte responses. *J. Infect. Dis.* **195**, 1126–1136 (2007).
121. Gupta, N. *et al.* Affinity-Purified Respiratory Syncytial Virus Antibodies from Intravenous Immunoglobulin Exert Potent Antibody-Dependent Cellular Cytotoxicity. *PLOS ONE* **8**, e69390 (2013).
122. Haynes, L. M. *et al.* Therapeutic monoclonal antibody treatment targeting respiratory syncytial virus (RSV) G protein mediates viral clearance and reduces the pathogenesis of RSV infection in BALB/c mice. *J. Infect. Dis.* **200**, 439–447 (2009).
123. Corbeil, S., Seguin, C. & Trudel, M. Involvement of the complement system in the protection of mice from challenge with respiratory syncytial virus Long strain following passive immunization with monoclonal antibody 18A2B2. *Vaccine* **14**, 521–525 (1996).
124. Mekseepralard, C., Toms, G. L. & Routledge, E. G. Protection of mice against Human respiratory syncytial virus by wild-type and aglycosyl mouse-human chimaeric IgG antibodies to subgroup-conserved epitopes on the G glycoprotein. *J. Gen. Virol.* **87**, 1267–1273 (2006).
125. Miao, C. *et al.* Treatment with respiratory syncytial virus G glycoprotein monoclonal antibody or F(ab′)2 components mediates reduced pulmonary inflammation in mice. *J. Gen. Virol.* **90**, 1119–1123 (2009).
126. Hiatt, A. *et al.* Glycan variants of a respiratory syncytial virus antibody with enhanced effector function and in vivo efficacy. *Proc. Natl. Acad. Sci.* **111**, 5992–5997 (2014).
127. Bukreyev, A., Yang, L. & Collins, P. L. The secreted G protein of human respiratory syncytial virus antagonizes antibody-mediated restriction of replication involving macrophages and complement. *J. Virol.* **86**, 10880–10884 (2012).
128. Glezen, W. P., Taber, L. H., Frank, A. L. & Kasel, J. A. Risk of Primary Infection and Reinfection With Respiratory Syncytial Virus. *Am. J. Dis. Child.* **140**, 543–546 (1986).
129. Hall, C. B., Walsh, E. E., Long, C. E. & Schnabel, K. C. Immunity to and Frequency of Reinfection with Respiratory Syncytial Virus. *J. Infect. Dis.* **163**, 693–698 (1991).
130. McLellan, J. S. *et al.* Structure-Based Design of a Fusion Glycoprotein Vaccine for Respiratory Syncytial Virus. *Science* **342**, 592–598 (2013).
131. Smith, G. *et al.* Respiratory syncytial virus fusion glycoprotein expressed in insect cells form protein nanoparticles that induce protective immunity in cotton rats. *PloS One* **7**, e50852 (2012).

132. Espeseth, A. S. *et al.* Modified mRNA/lipid nanoparticle-based vaccines expressing respiratory syncytial virus F protein variants are immunogenic and protective in rodent models of RSV infection. *Npj Vaccines* **5**, 1–14 (2020).
133. Aliprantis, A. O. *et al.* A phase 1, randomized, placebo-controlled study to evaluate the safety and immunogenicity of an mRNA-based RSV prefusion F protein vaccine in healthy younger and older adults. *Hum. Vaccines Immunother.* **17**, 1248–1261 (2021).
134. Higgins, D., Trujillo, C. & Keech, C. Advances in RSV vaccine research and development – A global agenda. *Vaccine* **34**, 2870–2875 (2016).
135. Plotkin, S. A. Correlates of Protection Induced by Vaccination. *Clin. Vaccine Immunol. CVI* **17**, 1055–1065 (2010).
136. Taylor, G. Animal models of respiratory syncytial virus infection. *Vaccine* **35**, 469–480 (2017).
137. Ackerman, M. E. *et al.* Route of immunization defines multiple mechanisms of vaccine-mediated protection against SIV. *Nat. Med.* **24**, 1590–1598 (2018).
138. Coffman, R. L., Sher, A. & Seder, R. A. Vaccine Adjuvants: Putting Innate Immunity to Work. *Immunity* **33**, 492–503 (2010).
139. Ols, S. *et al.* Route of Vaccine Administration Alters Antigen Trafficking but Not Innate or Adaptive Immunity. *Cell Rep.* **30**, 3964-3971.e7 (2020).
140. Piedimonte, G. & Perez, M. K. Respiratory Syncytial Virus Infection and Bronchiolitis. *Pediatr. Rev.* **35**, 519–530 (2014).
141. Iwasaki, A., Foxman, E. F. & Molony, R. D. Early local immune defenses in the respiratory tract. *Nat. Rev. Immunol.* **17**, 7–20 (2017).
142. Spear, G. T., Sullivan, B. L., Landay, A. L. & Lint, T. F. Neutralization of human immunodeficiency virus type 1 by complement occurs by viral lysis. *J. Virol.* **64**, 5869–5873 (1990).
143. Cong, J. & Wei, H. Natural Killer Cells in the Lungs. *Front. Immunol.* **10**, (2019).
144. Pierantoni, A. *et al.* Mucosal delivery of a vectored RSV vaccine is safe and elicits protective immunity in rodents and nonhuman primates. *Mol. Ther. - Methods Clin. Dev.* **2**, (2015).
145. Habibi, M. S. *et al.* Impaired Antibody-mediated Protection and Defective IgA B-Cell Memory in Experimental Infection of Adults with Respiratory Syncytial Virus. *Am. J. Respir. Crit. Care Med.* **191**, 1040–1049 (2015).
146. Boyaka, P. N. Inducing mucosal IgA: A challenge for vaccine adjuvants and delivery systems. *J. Immunol. Baltim. Md 1950* **199**, 9–16 (2017).
147. Shikina, T. *et al.* IgA Class Switch Occurs in the Organized Nasopharynx- and Gut-Associated Lymphoid Tissue, but Not in the Diffuse Lamina Propria of Airways and Gut. *J. Immunol.* **172**, 6259–6264 (2004).
148. Jackson, L. A. *et al.* An mRNA Vaccine against SARS-CoV-2 — Preliminary Report. *N. Engl. J. Med.* **383**, 1920–1931 (2020).
149. Pardi, N., Hogan, M. J., Porter, F. W. & Weissman, D. mRNA vaccines — a new era in vaccinology. *Nat. Rev. Drug Discov.* **17**, 261–279 (2018).
150. Jozwik, A. *et al.* RSV-specific airway resident memory CD8+ T cells and differential disease severity after experimental human infection. *Nat. Commun.* **6**, 10224 (2015).
151. de Bree, G. J. *et al.* Selective accumulation of differentiated CD8+ T cells specific for respiratory viruses in the human lung. *J. Exp. Med.* **202**, 1433–1442 (2005).

152. Luangrath, M. A., Schmidt, M. E., Hartwig, S. M. & Varga, S. M. Tissue-Resident Memory T Cells in the Lungs Protect against Acute Respiratory Syncytial Virus Infection. *ImmunoHorizons* **5**, 59–69 (2021).
153. Wang, D. *et al.* A Single-Dose Recombinant Parainfluenza Virus 5-Vectored Vaccine Expressing Respiratory Syncytial Virus (RSV) F or G Protein Protected Cotton Rats and African Green Monkeys from RSV Challenge. *J. Virol.* **91**, (2017).
154. Wong, P. T. *et al.* Formulation and Characterization of Nanoemulsion Intranasal Adjuvants: Effects of Surfactant Composition on Mucoadhesion and Immunogenicity. *Mol. Pharm.* **11**, 531–544 (2014).
155. Zhang, L. *et al.* Design and characterization of a fusion glycoprotein vaccine for Respiratory Syncytial Virus with improved stability. *Vaccine* **36**, 8119–8130 (2018).
156. Wen, Z. *et al.* Development and application of a higher throughput RSV plaque assay by immunofluorescent imaging. *J. Virol. Methods* **263**, 88–95 (2019).
157. Brown, E. P. *et al.* High-throughput, multiplexed IgG subclassing of antigen-specific antibodies from clinical samples. *J. Immunol. Methods* **386**, 117–123 (2012).
158. McLellan, J. S., Yang, Y., Graham, B. S. & Kwong, P. D. Structure of Respiratory Syncytial Virus Fusion Glycoprotein in the Postfusion Conformation Reveals Preservation of Neutralizing Epitopes. *J. Virol.* **85**, 7788–7796 (2011).
159. 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).
160. Fischinger, S. *et al.* A high-throughput, bead-based, antigen-specific assay to assess the ability of antibodies to induce complement activation. *J. Immunol. Methods* **473**, 112630 (2019).
161. Karsten, C. B. *et al.* A versatile high-throughput assay to characterize antibody-mediated neutrophil phagocytosis. *J. Immunol. Methods* **471**, 46–56 (2019).
162. Lu, L. L. *et al.* A Functional Role for Antibodies in Tuberculosis. *Cell* **167**, 433-443.e14 (2016).
163. Chung, A. W. *et al.* Dissecting Polyclonal Vaccine-Induced Humoral Immunity against HIV Using Systems Serology. *Cell* **163**, 988–998 (2015).
164. Citron, M. P. *et al.* A novel method for strict intranasal delivery of non-replicating RSV vaccines in cotton rats and non-human primates. *Vaccine* **36**, 2876–2885 (2018).
165. Guyon, I. & Elisseeff, A. An Introduction to Variable and Feature Selection. *J. Mach. Learn. Res.* 1157–1182 (2003).
166. Pittala, S. *et al.* Antibody Fab-Fc properties outperform titer in predictive models of SIV vaccine-induced protection. *Mol. Syst. Biol.* **15**, (2019).
167. Wu, F. *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265–269 (2020).
168. Zhu, N. *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020).
169. Li, R. *et al.* Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **368**, 489–493 (2020).
170. Chen, G. *et al.* Clinical and immunological features of severe and moderate coronavirus disease 2019. *J. Clin. Invest.* **130**, 2620–2629 (2020).

171. Quinlan, B. D. *et al.* The SARS-CoV-2 receptor-binding domain elicits a potent neutralizing response without antibody-dependent enhancement. 2020.04.10.036418 Preprint at https://doi.org/10.1101/2020.04.10.036418 (2020).
172. Chang, S.-C. *et al.* Longitudinal analysis of Severe Acute Respiratory Syndrome (SARS) coronavirus-specific antibody in SARS patients. *Clin. Diagn. Lab. Immunol.* **12**, 1455–1457 (2005).
173. de Wit, E., van Doremalen, N., Falzarano, D. & Munster, V. J. SARS and MERS: recent insights into emerging coronaviruses. *Nat. Rev. Microbiol.* **14**, 523–534 (2016).
174. He, Y., Li, J. & Jiang, S. A single amino acid substitution (R441A) in the receptor-binding domain of SARS coronavirus spike protein disrupts the antigenic structure and binding activity. *Biochem. Biophys. Res. Commun.* **344**, 106–113 (2006).
175. ter Meulen, J. *et al.* Human monoclonal antibody combination against SARS coronavirus: synergy and coverage of escape mutants. *PLoS Med.* **3**, e237 (2006).
176. Zhao, J. *et al.* Passive immunotherapy with dromedary immune serum in an experimental animal model for Middle East respiratory syndrome coronavirus infection. *J. Virol.* **89**, 6117–6120 (2015).
177. Zhao, Y. *et al.* Passive immunotherapy for Middle East Respiratory Syndrome coronavirus infection with equine immunoglobulin or immunoglobulin fragments in a mouse model. *Antiviral Res.* **137**, 125–130 (2017).
178. Wu, F. *et al.* Neutralizing antibody responses to SARS-CoV-2 in a COVID-19 recovered patient cohort and their implications. 2020.03.30.20047365 Preprint at https://doi.org/10.1101/2020.03.30.20047365 (2020).
179. Wang, X. *et al.* Neutralizing Antibody Responses to Severe Acute Respiratory Syndrome Coronavirus 2 in Coronavirus Disease 2019 Inpatients and Convalescent Patients. *Clin. Infect. Dis.* **71**, 2688–2694 (2020).
180. Tomaras, G. D. & Haynes, B. F. HIV-1-specific antibody responses during acute and chronic HIV-1 infection. *Curr. Opin. HIV AIDS* **4**, 373–379 (2009).
181. Cobey, S. & Hensley, S. E. Immune history and influenza virus susceptibility. *Curr. Opin. Virol.* **22**, 105–111 (2017).
182. Saphire, E. O., Schendel, S. L., Gunn, B. M., Milligan, J. C. & Alter, G. Antibody-mediated protection against Ebola virus. *Nat. Immunol.* **19**, 1169–1178 (2018).
183. Gunn, B. M. & Alter, G. Modulating Antibody Functionality in Infectious Disease and Vaccination. *Trends Mol. Med.* **22**, 969–982 (2016).
184. Holshue, M. L. *et al.* First Case of 2019 Novel Coronavirus in the United States. *N. Engl. J. Med.* **382**, 929–936 (2020).
185. Bhatraju, P. K. *et al.* Covid-19 in Critically Ill Patients in the Seattle Region — Case Series. *N. Engl. J. Med.* **382**, 2012–2022 (2020).
186. Barouch, D. H. *et al.* Protective efficacy of adenovirus/protein vaccines against SIV challenges in rhesus monkeys. *Science* **349**, 320–324 (2015).
187. Li, T. *et al.* Long-term persistence of robust antibody and cytotoxic T cell responses in recovered patients infected with SARS coronavirus. *PloS One* **1**, e24 (2006).
188. Liu, W. *et al.* Evaluation of Nucleocapsid and Spike Protein-Based Enzyme-Linked Immunosorbent Assays for Detecting Antibodies against SARS-CoV-2. *J. Clin. Microbiol.* **58**, e00461-20 (2020).

189. Shi, Y. *et al.* Antibody responses against SARS-coronavirus and its nucleocaspid in SARS patients. *J. Clin. Virol. Off. Publ. Pan Am. Soc. Clin. Virol.* **31**, 66–68 (2004).
190. Timani, K. A. *et al.* Cloning, sequencing, expression, and purification of SARS-associated coronavirus nucleocapsid protein for serodiagnosis of SARS. *J. Clin. Virol. Off. Publ. Pan Am. Soc. Clin. Virol.* **30**, 309–312 (2004).
191. Buchholz, U. J. *et al.* Contributions of the structural proteins of severe acute respiratory syndrome coronavirus to protective immunity. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 9804–9809 (2004).
192. Bar-On, Y. M., Flamholz, A., Phillips, R. & Milo, R. SARS-CoV-2 (COVID-19) by the numbers. *eLife* **9**, e57309 (2020).
193. Chandrashekar, A. *et al.* SARS-CoV-2 infection protects against rechallenge in rhesus macaques. *Science* (2020) doi:10.1126/science.abc4776.
194. Hauser, A. *et al.* Estimation of SARS-CoV-2 mortality during the early stages of an epidemic: A modeling study in Hubei, China, and six regions in Europe. *PLoS Med.* **17**, e1003189 (2020).
195. Jegaskanda, S. *et al.* Generation and Protective Ability of Influenza Virus-Specific Antibody-Dependent Cellular Cytotoxicity in Humans Elicited by Vaccination, Natural Infection, and Experimental Challenge. *J. Infect. Dis.* **214**, 945–952 (2016).
196. Brown, E. P. *et al.* Multiplexed Fc array for evaluation of antigen-specific antibody effector profiles. *J. Immunol. Methods* **443**, 33–44 (2017).
197. Boudreau, C. M. *et al.* Selective induction of antibody effector functional responses using MF59-adjuvanted vaccination. *J. Clin. Invest.* **130**, 662–672 (2020).
198. Yang, Z. *et al.* A DNA vaccine induces SARS coronavirus neutralization and protective immunity in mice. *Nature* **428**, 561–564 (2004).
199. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
200. Wu, Z. & McGoogan, J. M. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA* **323**, 1239–1242 (2020).
201. Richardson, S. *et al.* Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. *JAMA* (2020) doi:10.1001/jama.2020.6775.
202. Williamson, E. J. *et al.* Factors associated with COVID-19-related death using OpenSAFELY. *Nature* 1–7 (2020) doi:10.1038/s41586-020-2521-4.
203. Kuri-Cervantes, L. *et al.* Comprehensive mapping of immune perturbations associated with severe COVID-19. *Sci. Immunol.* **5**, (2020).
204. Vaninov, N. In the eye of the COVID-19 cytokine storm. *Nat. Rev. Immunol.* **20**, 277–277 (2020).
205. Zohar, T. & Alter, G. Dissecting antibody-mediated protection against SARS-CoV-2. *Nat. Rev. Immunol.* **20**, 392–394 (2020).
206. Long, Q.-X. *et al.* Antibody responses to SARS-CoV-2 in patients with COVID-19. *Nat. Med.* **26**, 845–848 (2020).
207. Robbiani, D. F. *et al.* Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* 1–8 (2020) doi:10.1038/s41586-020-2456-9.

208. Folegatti, P. M. *et al.* Safety and immunogenicity of the ChAdOx1 nCoV-19 vaccine against SARS-CoV-2: a preliminary report of a phase 1/2, single-blind, randomised controlled trial. *The Lancet* **0**, (2020).
209. 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).
210. Nimmerjahn, F. & Ravetch, J. V. Fcgamma receptors as regulators of immune responses. *Nat. Rev. Immunol.* **8**, 34–47 (2008).
211. Atyeo, C. *et al.* Distinct Early Serological Signatures Track with SARS-CoV-2 Survival. *Immunity* S1074761320303277 (2020) doi:10.1016/j.immuni.2020.07.020.
212. Vidarsson, G., Dekkers, G. & Rispens, T. IgG Subclasses and Allotypes: From Structure to Effector Functions. *Front. Immunol.* **5**, (2014).
213. Otten, M. A. & van Egmond, M. The Fc receptor for IgA (FcalphaRI, CD89). *Immunol. Lett.* **92**, 23–31 (2004).
214. Jefferis, R., Lund, J. & Pound, J. D. IgG-Fc-mediated effector functions: molecular definition of interaction sites for effector ligands and the role of glycosylation. *Immunol. Rev.* **163**, 59–76 (1998).
215. Braun, J. *et al.* SARS-CoV-2-reactive T cells in healthy donors and patients with COVID-19. *Nature* (2020) doi:10.1038/s41586-020-2598-9.
216. Chen, X. *et al.* Host Immune Response to Influenza A Virus Infection. *Front. Immunol.* **9**, 320 (2018).
217. Murin, C. D., Wilson, I. A. & Ward, A. B. Antibody responses to viral infections: a structural perspective across three different enveloped viruses. *Nat. Microbiol.* **4**, 734–747 (2019).
218. Plotkin, S. A. Correlates of Protection Induced by Vaccination. *Clin. Vaccine Immunol. CVI* **17**, 1055–1065 (2010).
219. Mateus, J. *et al.* Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans. *Science* (2020) doi:10.1126/science.abd3871.
220. Kaneko, N. *et al.* Loss of Bcl-6-expressing T follicular helper cells and germinal centers in COVID-19. *Cell* **0**, (2020).
221. Bergqvist, P., Stensson, A., Lycke, N. Y. & Bemark, M. T Cell-Independent IgA Class Switch Recombination Is Restricted to the GALT and Occurs Prior to Manifest Germinal Center Formation. *J. Immunol.* **184**, 3545–3553 (2010).
222. Haynes, L. & Maue, A. C. Effects of aging on T cell function. *Curr. Opin. Immunol.* **21**, 414–417 (2009).
223. Walker, L. S. K. & Herrath, M. von. CD4 T cell differentiation in type 1 diabetes. *Clin. Exp. Immunol.* **183**, 16–29 (2016).
224. Green, W. D. & Beck, M. A. Obesity altered T cell metabolism and the response to infection. *Curr. Opin. Immunol.* **46**, 1–7 (2017).
225. Kawahara, J. Y., Irvine, E. B. & Alter, G. A Case for Antibodies as Mechanistic Correlates of Immunity in Tuberculosis. *Front. Immunol.* **10**, (2019).
226. Dobbs, K. R. & Dent, A. E. Plasmodium malaria and antimalarial antibodies in the first year of life. *Parasitology* **143**, 129–138 (2016).
227. van Doremalen, N. *et al.* ChAdOx1 nCoV-19 vaccination prevents SARS-CoV-2 pneumonia in rhesus macaques. *bioRxiv* (2020) doi:10.1101/2020.05.13.093195.
228. Yu, J. *et al.* DNA vaccine protection against SARS-CoV-2 in rhesus macaques. *Science* (2020) doi:10.1126/science.abc6284.

229. Worley, M. J. *et al.* Neutrophils mediate HIV-specific antibody-dependent phagocytosis and ADCC. *J. Immunol. Methods* **457**, 41–52 (2018).
230. Pesarin, F. & Salmaso, L. *Permutation Tests for Complex Data*. (John Wiley & Sons, Ltd, 2010). doi:10.1002/9780470689516.
231. Winkler, A. M. *et al.* Non-parametric combination and related permutation tests for neuroimaging. *Hum. Brain Mapp.* **37**, 1486–1511 (2016).
232. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *ArXiv180203426 Cs Stat* (2018).
233. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
234. Wold, S., Esbensen, K. & Geladi, P. Principal Component Analysis. *Chemom. Intell. Lab. Syst.* 16 (1987).
235. Maier, C., Loos, C. & Hasenauer, J. Robust parameter estimation for dynamical systems from outlier-corrupted data. *Bioinformatics* **33**, 718–725 (2017).
236. Raue, A. *et al.* Lessons Learned from Quantitative Dynamical Modeling in Systems Biology. *PLoS ONE* **8**, e74335–e74335 (2013).
237. Hass, H. *et al.* Benchmark problems for dynamic modeling of intracellular processes. *Bioinformatics* **35**, 3073–3082 (2019).
238. Akaike, H. Information theory and an extension of the maximum likelihood principle. *2nd Int. Symp. Inf. Theory Tsahkadsor Armen. SSR Akad. Kiado* **1**, 267–281 (1973).
239. Burnham, K. P. & Anderson, D. R. *Model Selection and Multimodel Inference: a Practical Information-theoretic Approach*. *2nd edn, New York Springer* (2002).
240. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545–15550 (2005).
241. Sergushichev, A. A. An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. *bioRxiv* 060012 (2016) doi:10.1101/060012.
242. Charles A Janeway, J., Travers, P., Walport, M. & Shlomchik, M. J. The Humoral Immune Response. *Immunobiol. Immune Syst. Health Dis. 5th Ed.* (2001).
243. Lo Nigro, C. *et al.* NK-mediated antibody-dependent cell-mediated cytotoxicity in solid tumors: biological evidence and clinical perspectives. *Ann. Transl. Med.* **7**, 105 (2019).
244. DiLillo, D. J. & Ravetch, J. V. Differential Fc-Receptor Engagement Drives an Anti-tumor Vaccinal Effect. *Cell* **161**, 1035–1045 (2015).
245. Silverman, G. J., Vas, J. & Grönwall, C. Protective autoantibodies in the rheumatic diseases: lessons for therapy. *Nat. Rev. Rheumatol.* **9**, 291–300 (2013).
246. Hoi, K. H. & Ippolito, G. C. Intrinsic bias and public rearrangements in the human immunoglobulin Vλ light chain repertoire. *Genes Immun.* **14**, 271–276 (2013).
247. Schroeder, H. W. & Cavacini, L. Structure and Function of Immunoglobulins. *J. Allergy Clin. Immunol.* **125**, S41–S52 (2010).
248. Xu, Z., Zan, H., Pone, E. J., Mai, T. & Casali, P. Immunoglobulin class-switch DNA recombination: induction, targeting and beyond. *Nat. Rev. Immunol.* **12**, 517–531 (2012).
249. Jennewein, M. F. & Alter, G. The Immunoregulatory Roles of Antibody Glycosylation. *Trends Immunol.* **38**, 358–372 (2017).
250. Kawasaki, Y. *et al.* Evaluation of T helper-1/-2 balance on the basis of IgG subclasses and serum cytokines in children with glomerulonephritis. *Am. J. Kidney Dis.* **44**, 42–49 (2004).

251. Cerutti, A. The regulation of IgA class switching. *Nat. Rev. Immunol.* **8**, 421–434 (2008).
252. Tan, P.-L., Jacobson, R. M., Poland, G. A., Jacobsen, S. J. & Pankratz, V. S. Twin studies of immunogenicity — determining the genetic contribution to vaccine failure. *Vaccine* **19**, 2434–2439 (2001).
253. Höhler, T. *et al.* Differential genetic determination of immune responsiveness to hepatitis B surface antigen and to hepatitis A virus: a vaccination study in twins. *Lancet Lond. Engl.* **360**, 991–995 (2002).
254. Lee, Y. C. *et al.* Influence of genetic and environmental factors on the immunogenicity of Hib vaccine in Gambian twins. *Vaccine* **24**, 5335–5340 (2006).
255. Yan, K. *et al.* Genetic effects have a dominant role on poor responses to infant vaccination to hepatitis B virus. *J. Hum. Genet.* **58**, 293–297 (2013).
256. Milet, J. *et al.* Genome-wide association study of antibody responses to Plasmodium falciparum candidate vaccine antigens. *Genes Immun.* **17**, 110–117 (2016).
257. Scepanovic, P. *et al.* Human genetic variants and age are the strongest predictors of humoral immune responses to common pathogens and vaccines. *Genome Med.* **10**, 59 (2018).
258. O'Connor, D. *et al.* Common Genetic Variations Associated with the Persistence of Immunity following Childhood Immunization. *Cell Rep.* **27**, 3241-3253.e4 (2019).
259. Butler-Laporte, G. *et al.* Genetic Determinants of Antibody-Mediated Immune Responses to Infectious Diseases Agents: A Genome-Wide and HLA Association Study. *Open Forum Infect. Dis.* **7**, ofaa450 (2020).
260. Liu, W. *et al.* Autoantibodies to tumor-associated antigens as biomarkers in cancer immunodiagnosis. *Autoimmun. Rev.* **10**, 331–335 (2011).
261. Seeling, M., Brückner, C. & Nimmerjahn, F. Differential antibody glycosylation in autoimmunity: sweet biomarker or modulator of disease activity? *Nat. Rev. Rheumatol.* **13**, 621–630 (2017).
262. Galli, S. J. & Tsai, M. IgE and mast cells in allergic disease. *Nat. Med.* **18**, 693–704 (2012).
263. Rubicz, R. *et al.* Genetic Factors Influence Serological Measures of Common Infections. *Hum. Hered.* **72**, 133–141 (2011).
264. Rabe-Hesketh, S., Skrondal, A. & Gjessing, H. K. Biometrical Modeling of Twin and Family Data Using Standard Mixed Model Software. *Biometrics* **64**, 280–288 (2008).
265. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
266. Boyle, A. P. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* **22**, 1790–1797 (2012).
267. Lepik, K. *et al.* C-reactive protein upregulates the whole blood expression of CD59 - an integrative analysis. *PLOS Comput. Biol.* **13**, e1005766 (2017).
268. Buil, A. *et al.* Gene-gene and gene-environment interactions detected by transcriptome sequence analysis in twins. *Nat. Genet.* **47**, 88–91 (2015).
269. Zeller, T. *et al.* Genetics and beyond--the transcriptome of human monocytes and disease susceptibility. *PloS One* **5**, e10693 (2010).
270. Chen, L. *et al.* Genetic Drivers of Epigenetic and Transcriptional Variation in Human Immune Cells. *Cell* **167**, 1398-1414.e24 (2016).
271. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).

272. Ong, C.-T. & Corces, V. G. CTCF: an architectural protein bridging genome topology and function. *Nat. Rev. Genet.* **15**, 234–246 (2014).
273. Kundaje, A. *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
274. Kelly, R. J., Rouquier, S., Giorgi, D., Lennon, G. G. & Lowe, J. B. Sequence and Expression of a Candidate for the Human Secretor Blood Group α(1,2)Fucosyltransferase Gene (FUT2): HOMOZYGOSITY FOR AN ENZYME-INACTIVATING NONSENSE MUTATION COMMONLY CORRELATES WITH THE NON-SECRETOR PHENOTYPE (∗). *J. Biol. Chem.* **270**, 4640–4649 (1995).
275. THE TABULA SAPIENS CONSORTIUM. The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans. *Science* **376**, eabl4896 (2022).
276. Raza, M. W. *et al.* Association between secretor status and respiratory viral illness. *BMJ* **303**, 815–818 (1991).
277. Currier, R. L. *et al.* Innate Susceptibility to Norovirus Infections Influenced by FUT2 Genotype in a United States Pediatric Population. *Clin. Infect. Dis.* **60**, 1631–1638 (2015).
278. Imbert-Marcille, B.-M. *et al.* A FUT2 gene common polymorphism determines resistance to rotavirus A of the P[8] genotype. *J. Infect. Dis.* **209**, 1227–1230 (2014).
279. Kindberg, E. *et al.* A nonsense mutation (428G-->A) in the fucosyltransferase FUT2 gene affects the progression of HIV-1 infection. *AIDS Lond. Engl.* **20**, 685–689 (2006).
280. Fumagalli, M. *et al.* Signatures of Environmental Genetic Adaptation Pinpoint Pathogens as the Main Selective Pressure through Human Evolution. *PLOS Genet.* **7**, e1002355 (2011).
281. Miller, L. H., Mason, S. J., Clyde, D. F. & McGinniss, M. H. The Resistance Factor to Plasmodium vivax in Blacks. *N. Engl. J. Med.* **295**, 302–304 (1976).
282. Samson, M. *et al.* Resistance to HIV-1 infection in Caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* **382**, 722–725 (1996).
283. Kim, T. Y. *et al.* Effects of Structure of Rho GTPase-activating Protein DLC-1 on Cell Morphology and Migration*. *J. Biol. Chem.* **283**, 32762–32770 (2008).
284. van Helden, S. F. G., Anthony, E. C., Dee, R. & Hordijk, P. L. Rho GTPase Expression in Human Myeloid Cells. *PLoS ONE* **7**, e42563 (2012).
285. Iannacone, M. & Guidotti, L. G. Immunobiology and pathogenesis of hepatitis B virus infection. *Nat. Rev. Immunol.* **22**, 19–32 (2022).
286. Crispe, I. N. Immune Tolerance in Liver Disease. *Hepatol. Baltim. Md* **60**, 2109–2117 (2014).
287. Xue, W. *et al.* DLC1 is a chromosome 8p tumor suppressor whose loss promotes hepatocellular carcinoma. *Genes Dev.* **22**, 1439–1444 (2008).
288. Dong, X. *et al.* Association of DLC1 gene polymorphism with susceptibility to hepatocellular carcinoma in Chinese hepatitis B virus carriers. *Cancer Epidemiol.* **33**, 265–270 (2009).
289. Banaudha, K. *et al.* MicroRNA silencing of tumor suppressor DLC-1 promotes efficient hepatitis C virus replication in primary human hepatocytes. *Hepatology* **53**, 53–61 (2011).
290. Joazeiro, C. A. P. & Weissman, A. M. RING Finger Proteins: Mediators of Ubiquitin Ligase Activity. *Cell* **102**, 549–552 (2000).
291. Luck, K. *et al.* A reference map of the human binary protein interactome. *Nature* **580**, 402–408 (2020).
292. Potter, D. & Miziorko, H. M. Identification of catalytic residues in human mevalonate kinase. *J. Biol. Chem.* **272**, 25449–25454 (1997).

293. van der Burgh, R. *et al.* Unprenylated RhoA Contributes to IL-1β Hypersecretion in Mevalonate Kinase Deficiency Model through Stimulation of Rac1 Activity. *J. Biol. Chem.* **289**, 27757–27765 (2014).
294. Akula, M. K. *et al.* Protein prenylation restrains innate immunity by inhibiting Rac1 effector interactions. *Nat. Commun.* **10**, 3975 (2019).
295. Drenth, J. P. H. *et al.* Mutations in the gene encoding mevalonate kinase cause hyper-IgD and periodic fever syndrome. *Nat. Genet.* **22**, 178–181 (1999).
296. Bruhns, P. *et al.* Specificity and affinity of human Fcγ receptors and their polymorphic variants for human IgG subclasses. *Blood* **113**, 3716–3725 (2009).
297. Brady, S. C. *et al.* Sprouty2 Association with B-Raf Is Regulated by Phosphorylation and Kinase Conformation. *Cancer Res.* **69**, 6773–6781 (2009).
298. Lavoie, H., Gagnon, J. & Therrien, M. ERK signalling: a master regulator of cell behaviour, life and fate. *Nat. Rev. Mol. Cell Biol.* **21**, 607–632 (2020).
299. Khajuria, R. K. *et al.* Ribosome Levels Selectively Regulate Translation and Lineage Commitment in Human Hematopoiesis. *Cell* **173**, 90-103.e19 (2018).
300. Milillo, A. *et al.* A SPRY2 mutation leading to MAPK/ERK pathway inhibition is associated with an autosomal dominant form of IgA nephropathy. *Eur. J. Hum. Genet.* **23**, 1673–1678 (2015).
301. Aleyd, E. *et al.* IgA Enhances NETosis and Release of Neutrophil Extracellular Traps by Polymorphonuclear Cells via Fcα Receptor I. *J. Immunol.* **192**, 2374–2383 (2014).
302. Cooling, L. Blood Groups in Infection and Host Susceptibility. *Clin. Microbiol. Rev.* **28**, 801–870 (2015).
303. Ferrer-Admetlla, A. *et al.* A natural history of FUT2 polymorphism in humans. *Mol. Biol. Evol.* **26**, 1993–2003 (2009).
304. Graziano, V. R., Wei, J. & Wilen, C. B. Norovirus Attachment and Entry. *Viruses* **11**, 495 (2019).
305. Battles, M. B. & McLellan, J. S. Respiratory syncytial virus entry and how to block it. *Nat. Rev. Microbiol.* **17**, 233–245 (2019).
306. Liljeroos, L., Krzyzaniak, M. A., Helenius, A. & Butcher, S. J. Architecture of respiratory syncytial virus revealed by electron cryotomography. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 11133–11138 (2013).
307. Moayyeri, A., Hammond, C. J., Hart, D. J. & Spector, T. D. Effects of age on genetic influence on bone loss over 17 years in women: the Healthy Ageing Twin Study (HATS). *J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res.* **27**, 2170–2178 (2012).
308. Brown, E. P. *et al.* High-throughput, multiplexed IgG subclassing of antigen-specific antibodies from clinical samples. *J. Immunol. Methods* **386**, 117–123 (2012).
309. Perez, L. G. *et al.* V1V2-specific complement activating serum IgG as a correlate of reduced HIV-1 infection risk in RV144. *PloS One* **12**, e0180720 (2017).
310. Rijsdijk, F. V. & Sham, P. C. Analytic approaches to twin data using structural equation models. *Brief. Bioinform.* **3**, 119–133 (2002).
311. O'Connell, J. *et al.* A General Approach for Haplotype Phasing across the Full Spectrum of Relatedness. *PLOS Genet.* **10**, e1004234 (2014).
312. Das, S. *et al.* Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284–1287 (2016).

313. Lippert, C. *et al.* FaST linear mixed models for genome-wide association studies. *Nat. Methods* **8**, 833–835 (2011).
314. Widmer, C. *et al.* Further Improvements to Linear Mixed Models for Genome-Wide Association Studies. *Sci. Rep.* **4**, 6874 (2014).
315. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, s13742-015-0047–8 (2015).
316. Boughton, A. P. *et al.* LocusZoom.js: interactive and embeddable visualization of genetic association study results. *Bioinformatics* **37**, 3017–3018 (2021).
317. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
318. Chen, E. Y. *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).
319. Scott, M. K. D. *et al.* Increased monocyte count as a cellular biomarker for poor outcomes in fibrotic diseases: a retrospective, multicentre cohort study. *Lancet Respir. Med.* **7**, 497–508 (2019).
320. Jung, C. *et al.* Omicron: What Makes the Latest SARS-CoV-2 Variant of Concern So Concerning? *J. Virol.* **96**, e02077-21.
321. Tak, Y. G. & Farnham, P. J. Making sense of GWAS: using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome. *Epigenetics Chromatin* **8**, 57 (2015).
322. Blattler, A. *et al.* Global loss of DNA methylation uncovers intronic enhancers in genes showing expression changes. *Genome Biol.* **15**, 469 (2014).
323. Schalkwyk, L. C. *et al.* Allelic Skewing of DNA Methylation Is Widespread across the Genome. *Am. J. Hum. Genet.* **86**, 196–212 (2010).
324. Chen, L. *et al.* Genetic Drivers of Epigenetic and Transcriptional Variation in Human Immune Cells. *Cell* **167**, 1398-1414.e24 (2016).
325. Smith, A. K. *et al.* Methylation quantitative trait loci (meQTLs) are consistently detected across ancestry, developmental stage, and tissue type. *BMC Genomics* **15**, 145 (2014).
326. Thurman, R. E. *et al.* The accessible chromatin landscape of the human genome. *Nature* **489**, 75–82 (2012).
327. Husquin, L. T. *et al.* Exploring the genetic basis of human population differences in DNA methylation and their causal impact on immune gene regulation. *Genome Biol.* **19**, 222 (2018).
328. Do, C. *et al.* Genetic–epigenetic interactions in cis: a major focus in the post-GWAS era. *Genome Biol.* **18**, 120 (2017).
329. Banovich, N. E. *et al.* Methylation QTLs Are Associated with Coordinated Changes in Transcription Factor Binding, Histone Modifications, and Gene Expression Levels. *PLOS Genet.* **10**, e1004663 (2014).
330. Veyrieras, J.-B. *et al.* High-Resolution Mapping of Expression-QTLs Yields Insight into Human Gene Regulation. *PLOS Genet.* **4**, e1000214 (2008).