Divergent trajectories of antiviral memory after SARS-Cov-2 infection

Adriana Tomic[1]*, Donal T. Skelly[2,3,4]*, Ane Ogbe[2*], Daniel O'Connor[1*], Matthew Pace[2], Emily Adland[2], Frances Alexander[5], Mohammad Ali[2], Kirk Allott[6], M. Azim Ansari[2], Sandra Belij-Rammerstorfer[7], Sagida Bibi[1], Luke Blackwell[1], Anthony Brown[2], Helen Brown[2], Breeze Cavell[5], Elizabeth A. Clutterbuck[1], Thushan de Silva[8], David Eyre[3,9], Amy Flaxman[7], James Grist[10], Carl-Philipp Hackstein[2], Rachel Halkerston[5], Adam C. Harding[11], Jennifer Hill[1,12], Tim James[6], Cecilia Jay[2], Síle A. Johnson[2,3,13], Barbara Kronsteiner[2,14], Yolanda Lie[15], Aline Linder[1,12], Stephanie Longet[5,16], Spyridoula Marinou[1,12], Philippa C. Matthews[2,3,12], Jack Mellors[5], Christos Petropoulos[15], Patpong Rongkard[2], Cynthia Sedik[15], Laura Silva-Reyes[1,12], Holly Smith[7], Lisa Stockdale[1,12], Stephen Taylor[5], Stephen Thomas[5], Timothy Tipoe[2], Lance Turtle[17,18], Vinicius Adriano Vieira[19], Terri Wrin[15], OPTIC Clinical Group, PITCH Study Group, C-MORE Group, Andrew J. Pollard[1,12], Teresa Lambe[7], Chris P. Conlon[20], Katie Jeffery[3], Simon Travis[3,21], Philip Goulder[19], John Frater[2,3], Alex J. Mentzer[3,16], Lizzie Stafford[20], Miles W. Carroll[5,16], William S. James[11], Paul Klenerman[2,3,12]#, Eleanor Barnes[2,3,12]#, Christina Dold[1,12]#, Susanna J. Dunachie[2,3,14,21]#

* These authors contributed equally

# These authors jointly supervised this work and contributed equally

Corresponding authors: Eleanor Barnes, email: ellie.barnes@ndm.ox.ac.uk and Adriana Tomic, email: info@adrianatomic.com

Affiliation

1. Oxford Vaccine Group, Department of Paediatrics, University of Oxford, UK
2. Peter Medawar Building for Pathogen Research, Nuffield Dept. of Clinical Medicine, University of Oxford, UK
3. Oxford University Hospitals NHS Foundation Trust, Oxford, UK
4. Nuffield Dept of Clinical Neuroscience, University of Oxford, UK
5. Public Health England, Porton Down, England
6. Department of Clinical Biochemistry, Oxford University Hospitals NHS Foundation Trust, Oxford, UK
7. Jenner Institute, University of Oxford, UK
8. The Florey Institute for Host-Pathogen Interactions and Department of Infection, Immunity and Cardiovascular Disease, Medical School, University of Sheffield, UK
9. Big Data Institute, Nuffield Dept. of Population Health, University of Oxford, UK
10. Department of Physiology, Anatomy, and Genetics, University of Oxford, UK
11. James & Lillian Martin Centre, Sir William Dunn School of Pathology, University of Oxford, Oxford, UK
12. NIHR Oxford Biomedical Research Centre, Oxford, UK
13. Oxford University Medical School, Medical Sciences Division, University of Oxford, Oxford, UK
14. Oxford Centre For Global Health Research, Nuffield Dept. of Clinical Medicine, University of Oxford, UK
15. Monogram Biosciences LabCorp, San Francisco, CA, USA
16. Wellcome Centre for Human Genetics, University of Oxford, UK
17. HPRU in Emerging and Zoonotic Infections, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, UK
18. Tropical and Infectious Disease Unit, Liverpool University Hospitals NHS Foundation Trust (a member of Liverpool Health Partners), Liverpool, UK
19. Peter Medawar Building for Pathogen Research, Department of Paediatrics, University of Oxford, Oxford, UK

20. Nuffield Department of Medicine, University of Oxford, Oxford, UK
21. Translational Gastroenterology Unit, Nuffield Department of Medicine, University of Oxford, Oxford, UK
22. Mahidol-Oxford Tropical Medicine Research Unit, Bangkok, Thailand

Acknowledgments:

The authors wish to thank all the healthcare worker volunteers who participated in this study, and Suki Kenth for administrative support.

Funding statements

This work was funded by the UK Department of Health and Social Care as part of the PITCH (Protective Immunity from T cells to Covid-19 in Health workers) Consortium, with contributions from UKRI/NIHR through the UK Coronavirus Immunology Consortium (UK-CIC) and from the Huo Family Foundation.

AT is supported by the EU's Horizon2020 Marie Sklodowska-Curie Fellowship (FluPRINT, grant number 796636). DS is supported by the NIHR Academic Clinical Fellow programme in Oxford. MAA is supported by a Wellcome Trust Sir Henry Dale Fellowship (220171/Z/20/Z). DWE is a Robertson Foundation Fellow. PCM is funded by a Wellcome intermediate fellowship, ref. 110110/Z/15/Z. LT is supported by the Wellcome Trust (grant number 205228/Z/16/Z) and the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Emerging and Zoonotic Infections (NIHR200907) at University of Liverpool in partnership with Public Health England (PHE), in collaboration with Liverpool School of Tropical Medicine and the University of Oxford. PK and EB are NIHR Senior Investigators and PK is funded by WT109965MA. SJD is funded by an NIHR Global Research Professorship (NIHR300791).

The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health and Social Care or Public Health England.

Contributions

EB, PK, CD and SJD conceptualised the study. SJD, PK, EB, CPC, DTS and LS designed and oversaw the clinical study. LS, DE, KJ, PM, AJM, PG, SAJ, ST, the OPTIC Study Group and the C-MORE Group contributed to the implementation of the clinical study. AO, AT, CD, DOC, DTS, MP, MWC, WSJ, EB, PK, and SJD designed and oversaw the laboratory studies. AJP, AO, CD, EA, JF, PJG, MWC, WSJ, EB, PK, SJD and TL were responsible for the implementation of the laboratory testing, while AB, ACH, AF, AL, BC, BK, CH, CJ, CP, CS, EAC, FA, HB, HS, JG, JG-J, JH, JM, KA, LB, LS-R, LSt, LT, MAA, MA, MLK, PR, RH, SB, SB-R, SL, SM, ST, STh, TJ, TT, TW, VAV, YL were responsible for laboratory testing and assay development. AT and DOC undertook the advanced data analysis. AO, AT, CD, DOC, DTS, EB, PK, and SJD prepared the manuscript, which was reviewed by all contributing authors. All other authors contributed to the implementation of the study and data collection.

Competing Interests

DWE declares lecture fees from Gilead, outside the submitted work. No other competing interests declared.

Key words: SARS-CoV-2, T cells, B cells, Antibodies, Durability, Healthcare workers, Longevity, variants of concern, COVID-19

Abstract

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection is normally controlled by effective host immunity including innate, humoral and cellular responses. However, the trajectories and correlates of acquired immunity, and the capacity of memory responses months after infection to neutralise variants of concern - which has important public health implications - is not fully understood. To address this, we studied a cohort of 78 UK healthcare workers who presented in April to June 2020 with symptomatic PCR-confirmed infection or who tested positive during an asymptomatic screening programme and tracked virus-specific B and T cell responses longitudinally at 5-6 time points each over 6 months, prior to vaccination. We observed a highly variable range of responses, some of which - T cell interferon-gamma (IFN-γ) ELISpot, N-specific antibody waned over time across the cohort, while others (spike-specific antibody, B cell memory ELISpot) were stable. In such cohorts, antiviral antibody has been linked to protection against re-infection. We used integrative analysis and a machine-learning approach (SIMON - Sequential Iterative Modeling Over Night) to explore this heterogeneity and to identify predictors of sustained immune responses. Hierarchical clustering defined a group of high and low antibody responders, which showed stability over time regardless of clinical presentation. These antibody responses correlated with IFN-γ ELISpot measures of T cell immunity and represent a subgroup of patients with a robust trajectory for longer term immunity. Importantly, this immune-phenotype associates with higher levels of neutralising antibodies not only against the infecting (Victoria) strain but also against variants B.1.1.7 (alpha) and B.1.351 (beta). Overall memory responses to SARS-CoV-2 show distinct trajectories following early priming, that may define subsequent protection against infection and severe disease from novel variants.

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), an RNA virus that causes coronavirus disease 2019 (COVID-19), first emerged in humans in December 2019 and has since spread globally, with more than 3.56 million deaths reported world-wide (June 2021 https://coronavirus.jhu.edu/map.html). Although the majority of infections cause asymptomatic or mild disease, a significant minority develop a severe illness, requiring hospitalisation, oxygen support, and invasive ventilation [1]. Healthcare workers (HCW) have been at the forefront of caring for patients with SARS-CoV-2 infection in community and hospital environments during the pandemic. High exposure rates have meant that a significant proportion of HCW have become infected and HCW most commonly infected are those working on the front line in patient facing roles, predominantly in acute medical specialities [2]. Older age, comorbidities and male sex remain the dominant factors that predispose to severe outcomes [3] – since HCW are predominantly younger and female [2], most have developed mild disease, although deaths are widely reported in this population.

Starting early in the pandemic, we and others have sought to characterise the immune responses during SARS-CoV-2 infection that are associated with viral clearance and disease severity. SARS-CoV-2 infection has been associated with the generation of high magnitude, broad T cell responses and high titres of immunoglobulin G (IgG) targeting SARS-CoV-2 spike and nucleoprotein (NP) antigens, particularly in severe COVID-19 [4]. Asymptomatic infection, that appears more common in younger people, may be associated with discordant T cell and humoral immunity with both the absence of IgG seroconversion in the presence of detectable T cell responses [5, 6] or conversely the presence of IgG in the absence of T cell immune responses [7]. However, more recently critical questions have emerged that include the durability of immune responses following initial infection, the quality of these responses, immune correlates of protection from re-infection, and the capacity of these responses to neutralise new variants of concern (VOC) that have emerged globally. These questions have become paramount following the development of effective vaccines for COVID-19, since deployment of these has been limited by vaccine supply, concerns around adverse events and vaccine hesitancy. Furthermore, to manage limited vaccine resource, people with previous infection are now being offered a single vaccine dose 6 months after infection in many European countries (France, Germany, Spain, and Italy) [8], on the assumption that natural immunity will protect from re-infection.

An in depth understanding of immune responses after SARS-CoV-2 infection, and how these change over time, will be critical to understanding who is susceptible to re-infection and to inform vaccine strategies. Currently, the precise correlates of immune protection from subsequent infection after primary disease, or after vaccination, are unknown. Previous reports suggest SARS-CoV-2 IgG antibodies [9] and previous exposure to seasonal coronaviruses (CoV) [10] are protective against subsequent SARS-CoV-2 infection. However, since the magnitude of T and B cell responses correlate with each other [11], dissecting the role of these immune subsets in protection from re-infection or severe disease on re-exposure is challenging. Several groups have now reported that SARS-CoV-2 specific T and B cells decline after acute disease [12, 13, 14, 15, 16], but there is high heterogeneity between individuals

in the levels of measurable immunity in different compartments it is unclear how or if the kinetics of this decline correlate with protection from subsequent infection. Concerns have been raised that SARS-CoV-2 re-infection associated with waning immunity is plausible, particularly since the seasonal coronaviruses, closely related to SARS-CoV-2, commonly re-infect the same host [17, 18]. However, waning of immune responses following acute infection, or vaccination is well recognised as part of the normal evolution of memory responses, and reports describing decline in immune responses have focused on ex vivo responses that may not reflect the memory recall potential of viral specific T and B cells responses. A particular concern is the identification of SARS-CoV-2 variants of concern (VOC) (B.1.1.7 - alpha, B.1.351 - beta, P.1 - gamma and B.1.617.2 - delta), with mutations which are associated with an increase in transmissibility, severity or escape from vaccine or SARS-CoV-2-induced immunity [19, 20, 21, 22, 23, 24, 25, 26, 27]. Immune escape, with a failure to neutralise the VOC, in live viral assays in vitro, appear following vaccination and after SARS-CoV-2 infection, and is pronounced in the context of lower antibody titres measured against the initial pandemic strain (B/Victoria).

Since April 2020, we have followed a cohort of SARS-CoV-2 infected HCW prospectively over time at Oxford University Hospital NHS Foundation Trust. Seventy-eight HCW infected during the UK's "first wave" (defined by positive PCR and seropositive for anti-spike antibodies) were assessed at up to six timepoints and followed for six months in 2020, pre-vaccination, with multiple immune parameters evaluated in more than 430 blood draws. Our aims are to characterise memory T and B cell responses following infection, and to determine the interactions between clinical presentation and the generation and maintenance of T and B cell responses over time. We assess the association of exposure to seasonal coronaviruses and symptomatic SARS-CoV-2 disease with the durability of SARS-CoV-2 specific responses. We evaluate the predictive value of clinical and immune parameters measured early after infection on the durability of immune responses using an integrative analysis with a machine learning platform (SIMON) [28, 29]. Using this approach, we define a group of high and low antibody responders with a differential capacity to neutralise the VOC.

Methods and materials

Detailed description of methods are included in the Appendix.

HCW volunteer recruitment and ethics

We sampled seventy-eight HCW at five or six time points each, over six months. HCWs were recruited from Oxford University Hospitals NHS Foundation Trust after a positive SARS-CoV-2 PCR test [2] in April-May 2020, including 66 volunteers with symptomatic disease (fever, shortness of breath, cough, loss of taste or smell, sore throat, coryza or diarrhoea) and 12 asymptomatic HCW who did not report any symptoms of COVID-19 in 2020 prior to staff screening or in the seven days following testing positive. The age, sex and ethnicity of the HCW are shown in **Supplementary Table 1**. Blood samples were acquired at multiple timepoints over 6 months (acute[range:1-20], 28 days [21-41], 56 days [42-73], 90 days [74-104], 120 days [110-140], and 180[160-200]) from onset of symptoms in the symptomatic group and from the date of positive PCR test for asymptomatic people diagnosed on screening. Nine hospitalised patients with severe disease were included for comparative analysis. All subjects were seropositive for anti-spike IgG antibodies by ELISA. Mild and asymptomatic participants were recruited under ethics approved by the research ethics committee (REC) at Yorkshire & The Humber - Sheffield (GI Biobank Study 16/YH/0247). Participants with severe disease were recruited after consenting into either the CMORE study protocol (research ethics committee (REC): Northwest – Preston, REC reference 20/NW/0235) and / or Sepsis Immunomics protocol [Oxford Research Ethics Committee C, reference 19/SC/0296]). The study was conducted according to the principles of the Declaration of Helsinki (2008) and the International Conference on Harmonization (ICH) Good Clinical Practice (GCP) guidelines. Written informed consent was obtained for all participants enrolled in the study.

Isolation of peripheral blood mononuclear cells (PBMC), plasma and serum

PBMCs and plasma were isolated by density gradient centrifugation from blood collected in EDTA tubes, and serum was collected in a serum-separating tube (SST, Becton Dickinson) as previously described[5] and detailed in the Appendix.

T cell assays

T cell assays including interferon-gamma (IFN-γ) Enzyme-Linked immunospot (ELISpot) assay, 7-day proliferation assay and intracellular staining were performed [5]. For IFN-γ ELISpot assay we used SARS-CoV-2 peptide pools panning Spike (S1 and S2), membrane (M), nucleocapsid protein (NP), the X-domain of non-structural protein 3 (NSP3B), open reading frames 3 and 8 (ORF3 and ORF8), and cytomegalovirus, Epstein-Barr virus and Flu peptide pools (CEF) (2ug/ml per peptide) in a 16-18hour incubation at 37ºC. ELISpot plates were read using an AID ELISpot Reader (v.4.0) and results were reported as spot-forming units (SFU)/10^6 PBMC. T cell proliferation assay was performed using fresh or cryopreserved PBMC and CellTrace® Violet (CTV, Life Technologies) labelling and stimulated with peptide pools from SARS-CoV-2 spanning Spike (S1 and S2), M, NP, ORF3 and ORF8, and FEC-T (1µg/ml per peptide). On day 7, cells were stained with fluorochrome-conjugated primary human-

specific antibodies for CD3, CD4 and CD8 for analysis on a MACSQuant 10 flow cytometer. For Intracellular cytokine staining, PBMC were stained for CD3, CD4, CD8, CD154, IFN-γ, IL-2 and TNF-α then analysed on a BD LSR II.

Antibody and B cell assays

Standardised total anti-spike IgG ELISA [30] and anti-spike subclass and isotype ELISAs [31, 32] were performed. A multiplexed MSD immunoassay (MSD, Rockville, MD) was used to measure the IgG responses to SARS-CoV-2, severe acute respiratory syndrome coronavirus-1 (SARS-CoV-1), MERS-CoV and seasonal CoVs (human coronavirus (HCoV)-OC43, HcoV-HKU1, HcoV-229E, HcoV-NL63). For Microneutralisation Assay (MNA), the viral isolates used are described in the Appendix, and the assay was performed to determine the concentration of antibody that produces a 50% reduction in infectious focus-forming units of authentic SARS-CoV-2 in Vero CCL81 cells. Infectious foci were enumerated by ELISpot reader and data were analysed using four-parameter logistic regression (Hill equation) in GraphPad Prism 8.3. The Monogram Bioscience pseudotype neutralisation assay (PseudoNA) was performed [30].

For the Spike-specific SARS-CoV-2, OC43, HKU1, 229E and NL63 IgG^+ and IgA^+ B cell memory ELISpot assays, PBMCs were cultured for 3-3.5 days with polyclonal stimulation, and added to Mabtech flurospot plates coated with the relevant spike glycoprotein (SARS-CoV-2 at 10µg/ml, OC43 at 10µg/ml, NL63 at 15µg/ml, HKU1 at 5µg/ml and 229E at 10µg/ml, all diluted in PBS). All cells were incubated for ≥16 hours at 37°C, and following development Spot forming units were enumerated using AID ELISpot 8.0 software on the AID ELR08IFL reader. For antibody-dependent effector functions, the spike-specific antibody-dependent effector functions, natural killer cell activity (ADNKA), neutrophil phagocytosis (ADNP) and monocyte phagocytosis (ADMP) were performed [31], and are detailed in the Appendix alongside the Antibody-dependent complement deposition (ADCD) assay.

Integrative analysis using unsupervised and supervised machine learning in SIMON

The integrative analysis was performed using SIMON (Sequential Iterative Modeling "Over Night") software [28, 29] as detailed in the Appendix. The integrated dataset was generated using the standard extract-transform-load (ETL) procedure to merge total of 29 csv files across 14 assays and clinical data via donor-specific variable (Donor ID) according to the SIMON method. The outcome of immune response durability was calculated based on the titre of the anti-N specific antibodies measured 6 months post symptoms onset (pso), and individuals with anti-N antibody titre ≥ 1.4 were labelled as high responders, while individuals having anti-N antibody titre below 1.4 were low responders. Before integrative analyses, data was pre-processed (centre/scale), missing values were median imputed, features with zero-variance, near-zero-variance and with correlation (cut-off 0.85) were removed using SIMON software. The t-distributed stochastic neighbour embedding (t-SNE) (2,000 iterations, perplexity 30, and theta 0.5) followed by clustering (seed number 1337, number of clusters 3) was performed to analyse the pre-processed integrated dataset (excluding disease severity and timepoint which are used as grouping variables). Principal component analysis (PCA) was performed on multivariate

immunological parameters (continuous variables, excluding features with less than 10% of unique values and grouping variable - disease severity). Pairwise correlations of immunological parameters in the integrated dataset were visualized as a correlogram and Spearman's rank correlation coefficient was computed. Values shown on the correlogram were adjusted for multiple testing using Benjamini-Hochberg correction at the significance threshold (False discover rate, FDR < 0.05). Agglomerative hierarchical clustering was performed on the samples with immunological parameters analysed on day 28 pso and visualized as the dendrogram on heatmap (tightest cluster ordered first). To identify early immunological signature at day 28 pso that can predict if the individual will be high or low responder 6 months pso, we performed SIMON analysis on all immunological parameters (day 28 pso) using 172 ML algorithms . Missing values (29% missingness) were removed using multi-set interaction function ('mulset', SIMON software), resulting in 30 resamples. Each resample was split into train/test partition (75%/25%) preserving the balanced distribution of the outcome class (seed number 1337). The models were evaluated using 10-fold cross-validation on training sets (train AUROC), and additionally on the held-out test sets (test AUROC). The best performing model was built using the Sparse Partial Least Squares (sPLS) algorithm (train AUROC: 0.95 (CI 0.5-1) and test AUROC: 1). In the final step, SIMON calculated the contribution of each feature to the model as variable importance score (scaled to maximum value of 100).

Statistical analyses

Statistical analysis was performed using R (https://www.r-project.org/), integrative analysis was performed using SIMON software [28, 29], figures were made with R using R package ggplot2 [33] and GraphPad Prism 8. Kruskal-Wallis test —unless otherwise specified — was used for comparison of the disease severity groups. Wilcoxon rank-sum test —unless otherwise specified — was employed to compare between study time points. A generalised additive mixed model (GAMM) by restricted maximum likelihood (REML) was used to fit the immunological measures (log10 transformed) using Gaussian process smooth term (R package *gamm4* [34]). ICS cytokine expression analyses was performed using PESTEL v2.0 and SPICE v6.0. Statistical significance was set at $P<0.05$ and all tests were 2-tailed. Machine learning analysis was performed using SIMON software (https://genular.org).

Data Availability

Data relating to the findings of this study are available from a research data repository Zenodo (https://zenodo.org/record/4905965).

Results

Anti-N IgG decline over time and stratify by disease severity, whilst Anti Spike IgG and memory responses are maintained

Anti-nucleocapsid (NP) and spike (S) total IgG (tIgG) responses were assessed by ELISA in both symptomatic and asymptomatic individuals (**Fig. 1A**). The magnitude of the IgG response varied markedly between people in both cohorts, with a proportion of individuals' anti-nucleocapsid tIgG level recorded in the negative or indeterminate range of the assay at all time-points.

Asymptomatic and mild infection induces similar anti-NP responses in the early phase (<20 days post PCR positivity/symptom onset) of observed infection (*P*=0.6125, **Supplementary Fig. 1A**). However, anti-NP tIgG levels in the two disease cohorts separated as higher levels were observed in those with mild infection from the day 28 timepoint onwards (*P* =0.0015 for day 28 comparison, **Supplementary Fig. 1A**). Anti-NP IgG responses waned over time with a significant decrease from approximately day 28 to day 180 timepoints (*P*=0.00071 for asymptomatic and $P=7.2x10^{-9}$ for mild symptomatic individuals, **Fig. 1A**). Most (91.7%) asymptomatic individuals have an indeterminate or negative anti-NP tIgG response to the nucleocapsid antigen at the day 180 timepoint.

Over the time course of observation, anti-spike IgG antibody levels (**Fig. 1B**) in individuals remained consistent in individuals with asymptomatic (*P*=0.35) and severe *(P*=0.44) COVID-19 disease. Similarly, the initial anti-spike tIgG responses increased in individuals with mild disease and remained consistent from day 28 to the 6-month timepoint *(P*=0.12). Furthermore, disease severity was not a significant predictor of anti-spike tIgG levels in those with asymptomatic and mild SARS-CoV-2 infection throughout the 6-month observation (*P* =0.632, GAMM, **Fig. 1B**).

In line with the tIgG antibody binding to spike remaining consistent, we observed a steady number of IgG+ memory B cells following an initial increase (**Fig. 1C**). Anti-SARS-CoV-2 spike-specific IgG+ memory B cells at 6 months following symptom onset were higher than observed during early infection in mild (*P*=0.00042, **Fig. 1C**) and severe (*P*=0.0027, **Fig. 1C**) individuals. For asymptomatic individuals, no change was observed in cell frequencies when comparing the earliest samples collected and 6-month timepoints (*P*=0.54), although we note that the timing of infection onset for asymptomatic individuals cannot be precisely determined. Asymptomatic and mild disease did not predict different kinetics for the IgG memory response (*P*=0.284, GAMM, **Fig. 1C**).

Pseudo-neutralisating antibodies decreased in all disease severities over time

Pseudo-neutralisating antibodies (pseudoNA) were measured in all individuals (**Fig. 1D**) using an assay that incorporates the spike glycoprotein. Disease severity was a significant predictor of pseudoNA (*P*=0.00073, GAMM, **Fig. 1D**) – with higher pseudoNA levels with increasing disease severity at all time points measured (**Fig. 1D and Supplementary Fig. 1D**). Regardless of disease severity, the pseudo-neutralising capacity of circulating antibodies to the Wuhan/B lineage virus decreased over 6 months following the detection of SARS-CoV-2 infection (asymptomatic *P*=0.023; mild $P=4.2\ x10^{-9}$; severe *P*=0.01, **Fig. 1D**). People with severe infection maintained pseudoNA 6 months post symptom onset, and at higher levels than in those with mild or asymptomatic infection (*P*=0.00022, Kruskal-Wallis test, **Supplementary Fig. 1D**). The decline was less marked in asymptomatic individuals with no decrease observed from day 28 to day 180 (*P*=0.41, **Fig. 1D**); however, the difference in the pseudoNA titres in the mild vs asymptomatic groups remained until day 180 (*P*=0.0148). At day 180 post symptom onset or PCR confirmation, one asymptomatic and four symptomatic individuals no longer mounted a positive

result in the pseudoNA assay, one of whom consistently did not mount pseudoNA capacity at all time points measured.

Mild infection induces a more multifunctional antibody profile

A cohort of 30 individuals with mild infection, along with the 9 and 12 participants with severe and asymptomatic infection respectively were selected to comprehensively characterise antibody profiles.

Circulating isotypes and subclasses

Circulating IgM levels decreased over time in those with asymptomatic (*P*=0.021, day <20 vs day 180), mild (*P*=0.0004, day <20 vs day 180) and severe (*P*=0.007, day <20 vs day 180) infection, while IgA levels in participants remained constant in all disease cohorts (asymptomatic: *P*=0.65; mild: *P*=0.59; severe: *P*=0.065), throughout the observed 6-month time course (**Fig. 2A and 2B**) as previously reported[12]. The quantified amounts of IgG1 were consistent over time in asymptomatic (*P*=0.86, day <20 vs day 180) and severe (*P*=0.92, day <20 vs day 180) infection. Despite initial low titres of IgG1 in participants with mild infection, IgG1 circulating antibody titres were maintained from day 28 to 6 months post symptom onset (*P*=0.89, **Fig. 2C**). While circulating IgG3 antibodies in participants with mild infection were maintained at consistent levels throughout the 6-month period (*P*=0.062), levels decreased over this time in asymptomatic (*P*=0.0022, day <20 vs day 180) and severe (*P*=0.021, day <20 vs day 180) individuals (**Fig. 2D**). Notable SARS-CoV-2 spike-specific IgG2 responses were only detected at one or more time-points in a small number of individuals tested (asymptomatic: 3/12; mild: 3/30; severe: 1/8) (**Supplementary Fig. 2B**), while there was no spike-specific IgG4 detected above the LLOQ of the ELISA (data not shown). For all IgG subclasses detected, asymptomatic or mild disease severity were not significant predictors of responses over time (IgG1: *P*=0.36; IgG2: *P*=0.92; IgG3: *P*=0.0519, GAMM, **Figs. 2C–D**). All paired analysis was by Wilcoxon rank sum test.

Diversity of antibody responses

We measured the ability of the anti-spike antibodies in those with severe or asymptomatic infection as well as a selection of individuals with mild infection, to induce innate effector functions: ADNP, ADMP, ADNKA and ADCD.

Asymptomatic and mild disease severity was not a significant predictor of Fc-mediated effector functional responses (ADNKA *P*=0.798; ADMP *P*=0.117; ADNP *P*=0.206) except for ADCD (*P*=0.00314) (**Fig. 2E–H**). Furthermore, normalised ADMP and ADNP scores, as well as the percentage of CD107a-expressing NK cells were stable over time, between 28 days and 180 days post symptom onset or PCR confirmation for those with asymptomatic (ADMP: *P*=0.96; ADNP: *P*=0.48; ADNKA: *P*=0.2) and mild (ADMP: *P*=0.64; ADNP: *P*=0.75; ADNKA: *P*=0.8) infection (**Fig. 2E–H**). Similarly, no decline was observed for these Fc-mediated functions from the acute sampling to 6 months post symptom onset in the severe cohort (ADMP: *P*=0.89; ADNP: increase *P*=0.021; ADNKA: *P*=0.075) with the ADNP increasing over time (*P*=0.021) (**Fig. 2E–H**). ADCD waned dramatically in those with severe disease over the 6-month period (*P*=0.00031) but similarly to the other Fc-mediated functions,

ADCD remained consistent from day 28 to day 180 in asymptomatic (*P*=0.34) and mild (*P*=0.1) infection (**Fig. 2E–H**). Despite waning over time, ADCD responses differed amongst the disease severity groups out until day 180 (*P*=0.0032, Kruskal-Wallis test, **Supplementary Fig. 1L**). All paired analysis were by Wilcoxon rank sum test.

We visualised the relative contribution of each of the anti-SARS-CoV-2 spike antibody feature in **Fig. 2I**. The polar plots demonstrate the diversity of asymptomatic and mild infection-induced antibody characteristics and functions on day 28 and day 180. Each wedge represents an antibody feature, and the size of each wedge is indicative of the magnitude of the response. The consistently high spike-specific IgG and spike-specific IgG+ memory B cells is clearly reflected in these plots for both mild and asymptomatic individuals. For both day 28 and day 180, a more multifunctional response was observed in individuals with mild infection, particularly for the antibody-dependent phagocytosis effector functions, which contribute markedly less to the antibody profile of asymptomatic individuals. Over time, few marked changes were observed in the relative contribution of the SARS-COV-2-specific antibody features in asymptomatic individuals, apart from an increased contribution of IgG1 and ADNKA, and decreased IgG3. Similarly, for individuals with mild infection, substantial relative decreases in IgM, pseudo-neutralising antibodies, IgA and IgG3 were noted, as well as relative increases in ADNKA and ADNP to the antibody profile.

SARS-CoV-2 infection elicits transient cross-reactive antibodies and memory B cells specific for other circulating coronaviruses.

Next, we evaluated the IgG responses to seasonal coronaviruses (229E, HKU-1, NL63-S and OC43-S) severe acute respiratory syndrome (SARS-CoV-1) spike protein and Middle East Respiratory Syndrome (MERS) virus spike protein using the MSD assay (**Fig. 3A**). IgG responses to these viral antigens were detected at the earliest time points. The kinetics of these IgG responses followed those seen to SARS-CoV-2 spike, suggesting that seasonal coronavirus cross-reactive responses were enhanced by SARS-CoV-2 infection. Responses to OC43-S, 229-E and HKU-1 were particularly high and correlated significantly with disease severity at day 180 and at the earliest time point assessed (day <20) **(Supplementary Fig. 2C).** The MSD assay also measured IgG responses against SARS-COV-2 Spike, NP and the RBD antigens, supporting our observations using the ELISA assay (**Supplementary Fig. 2D**).

IgG+ Memory B cells specific for the spike glycoprotein from seasonal coronaviruses (229E, HKU1, NL63 and OC43) were determined at the earliest timepoint available (acute <day 20 or day 28) and the 6-month final sampling (**Fig. 3B**). The lowest responses were observed in 229E and NL63 spike IgG+ ASCs following polyclonal stimulation, which also were consistent over time with the exception of the decreased number of NL63 spike-specific IgG+ memory B cells in individuals with mild infection (*P*=0.0046). Higher responses were detected when testing the specificity of cultured PBMCs to the beta-coronaviruses (HKU1 and OC43) spike glycoprotein. However, the boosted memory response

was transient, particularly in individuals with mild infection (HKU1: P=1x10^{-7}; OC43: P=1.5x10^{-7}) in which the decrease was more marked, which may be due to a higher sample number.

Effector poly-specific SARS-CoV-2 T cells are higher in those with mild symptoms and decline 6 months after infection

We examined the magnitude of the T cell response to SARS-CoV-2 using an ex vivo IFN-γ ELISpot assay at 28 days, 90-120 days and 180 days after SARS-CoV-2 infection (N=64-78 HCW/timepoint, 57 participants at all timepoints (including 12 with asymptomatic infection), and 6 volunteers with severe COVID-19 at day 180 (**Fig. 4A and 4B and Supplementary Table 3).** We have previously shown that this assay is specific for SARS-CoV-2, with negligible responses detected in SARS-CoV-2 pre-pandemic unexposed participants [5].

IFN-γ responses to at least one antigenic pool were seen in 67/70 (96%) volunteers tested 28 days after SARS-CoV-2, with a median total response across the pools of 373 (IQR 201–842) SFC/10^6 PBMC; here a response to spike (S1 and S2) was seen in 61/70 tested (87%) median 180 (IQR 71-364) SFC/10^6 PBMC, for M in 47/70 (67%) median 63 (IQR 25-160) SFC/10^6 PBMC and for NP in 62/70 (89%) median 121 (IQR 73-250) SFC/10^6 PBMC. However, total summed responses declined by a median of 60% after 90 days, and by 75% at 180 days (**Supplementary Table 3**). The majority (61/77 (79%)) of participants had detectable responses to at least one antigenic pool at 180 days, with responses to NP antigen most commonly observed 47/77 (61%) median 40 (IQR 23-73) SFC/10^6 PBMC. Responses to ORF3, ORF8 and NSP3B were less frequent than responses to S1, S2, M and NP at day 28 and lower at day 180.

IFN-γ ELISpot responses to SARS-CoV-2 antigens were higher in the mild symptomatic cohort (n=66), compared to the asymptomatic group (n=12) at 28 days, with median responses to all summed pools 455 (IQR 252-976) SFC/10^6 PBMC for mild disease compared to 196 (IQR 74-243) SFC/10^6 PBMC in the asymptomatic group (**Supplementary Fig. 3A**). There was no significant change in the magnitude of the T cell response in the asymptomatic group in the 6 months after infection (**Fig. 4A**).

We next used ICS to examine the duration of multiple T cell functions and the polyfunctionality of the T cell response over time at 28 and 180 days pso in individuals with ex vivo T cell ELISpot levels >100 SFC/10^6 PBMC for sensitivity reasons (n=18 with n=15 available at both timepoints for paired analysis (Gating strategy in **Supplementary Fig. 3D,** results in **Supplementary Fig. 4 and Supplementary Fig.5**). Similar to the ELISpot data, the majority of T cell responses decreased over time. In terms of functionality, we found that CD4+ T cells were polyfunctional, with the majority of cells expressing >1 and up to all 5 functional markers at both timepoints. Similarly, NSP3B-specific CD8+ T cells were also polyfunctional at both timepoints examined, with most cells expressing >1 functional marker (**Supplementary Fig. 4J**). There were no functional changes between the two timepoints.

T cell memory proliferative responses decline 6 months post SARS-CoV-2

We and others have found the assessment of T cell proliferation to be a sensitive method of detecting antigen-specific recall responses. We used this assay to evaluate the frequency of circulating SARS-CoV-2-specific CD4+ and CD8+ T cell in our longitudinal cohort (n = 54 – 57; gating strategy presented in **Supplementary Fig. 3B)**.

We did not observe any differences in the magnitude of circulating FEC-specific (control) CD4+ or CD8+ T cells within the 6 months period **(Supplementary Fig. 3C).** In the asymptomatic group, at 28 days pso 7/8 (87.5%) made a CD4+ T cell response to at least one SARS-CoV-2 protein (excluding S1 and S2 where have previously reported finding responses in the majority of unexposed volunteers [5]) while 5/8 (62.5%) of them had CD8 T cell response to at least one of M, NP, ORF3 or ORF8 proteins **(Fig. 5A-C Supplementary Table 4)**. Most of this response was targeted to M and NP **(Fig. 5A-C and Supplementary Table 4)**. At 180 days pso, 6/8 (75%) of recovered subjects had a CD4+ or CD8+ T cell response which was mostly focused on M, NP and ORF3. We observed no difference in the proliferative capacity of SARS-CoV-2-specific CD4 and CD8 T cells at 28- and 180-days post disease onset in the group with asymptomatic disease (n = 8) **(Fig. 5A-C and Supplementary table 4 and 5)**.

In the cohort with mild disease, at 28 days, T cell responses to at least one SARS-CoV-2 protein outside of spike region were observed in 42/49 (86%) for CD4+ T cells and 45/49 (91%) for CD8+ T cells. Similar to the asymptomatic cohort, these responses were focused on M, NP and ORF3 regions of SARS-CoV-2 **(Fig. 5A-C, Supplementary Table 4)**. At 180 days after symptom onset, this frequency of people responding to at least one protein as above reduced to 37/49 (75%) within CD4+ T cells and 35/49 (71%) for CD8+ T cells with a focus on M, NP and ORF3 similar to CD4+ T cells **(Fig. 5A-C and supplementary Table 4 and 5)**. In the volunteers with mild disease, we found a significant reduction in the circulating frequencies of SARS-CoV-2-specific CD4+ and CD8+ T cells to all proteins except NP and ORF8 for CD4+ and ORF3 and ORF8 for CD8+ T cells by day 180 **(Fig. 5A-C)**.

When we assessed the difference in the magnitude of the proliferative CD4+ and CD8+ T cell responses at 28- and 180 days pso in both asymptomatic and mild cases (analysed together as one group), we found significantly higher frequencies of SARS-CoV-2 specific CD4+ T cells compared to CD8+ responses at both timepoints in all proteins except NP and ORF8 for 28- and 180-days post symptom onset and ORF3 responses at 28 days post symptom onset only. Our data shows that the bias in antigen-specific responses to SARS-CoV-2 towards CD4+ T cells is maintained in the T cell memory compartment long after recovery from acute infection. Taken together, the results show that at 6 months post infection with SARS-CoV-2, convalescent subjects show diminished but detectable anti-SARS-CoV-2-specific memory T cells in both the CD4 and CD8 T cell compartments, with only 8/56 (14%) showing no proliferative response to any non-spike protein, suggesting durable immune response at least up to 6 months post initial infection.

Integrative analysis to Identify immune and clinical parameters associated with disease severity

To further investigate the trajectory of cellular and humoral adaptive immune responses during SARS-CoV-2 infection and relationship with disease severity, we performed integrative analysis on aggregated immunological and clinical data from 433 samples obtained from 86 donors (12 asymptomatic, 66 mild, 8 severe) on 6 different timepoints (**Fig. 6A)**. We investigated the trajectory of immune responses after SARS-CoV-2 infection and determined whether samples obtained from individuals with asymptomatic infection are more similar to samples obtained at later timepoints after infection in the individuals with mild, symptomatic disease. A t-distributed stochastic neighbour embedding (t-SNE) representation of integrated data revealed heterogeneity of immune responses in infected individuals, irrespective of days post symptom onset when these samples were collected (**Fig. 6B**, *left panel*). Majority of samples were separated between asymptomatic and mild individuals, while there was an overlap in similarity between individuals with mild and more severe disease (**Fig. 6B***, right panel*). To further delineate differences in clinical and immunological parameters of SARS-CoV-2 infected individuals, we performed clustering analysis on the resulting t-SNE representations (**Fig. 6C**) and compared expression of 16 clinical and 49 immunological parameters to identify each of three clusters (**Fig. 6D**). This approach identified heterogeneity within the SARS-CoV-2 positive individuals with mild disease clustered in two groups (**Fig. 6C and 6D**, *clusters 1 and 2*). In cluster 1, the majority of samples displayed increased antibody and T cell responses in comparison to other clusters, and some individuals with mild infection that showed clinical and immunological similarity to severe COVID-19 patients (**Fig. 6C and 6D**, *cluster 1*). In contrast, cluster 2 contained individuals with lower overall antibody and T cell responses and all were from individuals with mild disease (**Fig. 6C and 6D**, *cluster 2*). Clinical parameters were driving a major separation between asymptomatic SARS-CoV-2 positive individuals from those with mild or sever disease (**Fig. 6D**, *cluster 3*).

To gain an insight into immunological differences between individuals with asymptomatic and mild infection, we performed principal component analysis (PCA) on dataset containing only immunological parameters. The immunological parameters alone could explain 38.6% of variance between SARS-CoV-2 positive individuals, while separation was not driven by the disease severity (**Fig. 6E**). Comparable to t-SNE analysis, samples from individuals with mild disease were separated into three major groups having distinct immunophenotype (immunophenotypic group 1) (**Fig. 6E**, *lower right quadrant*) or sharing immunological similarity with samples from individuals with severe (immunophenotypic group 2) (**Fig. 6E**, *upper right quadrant*) or asymptomatic disease (immunophenotypic group 3) (**Fig. 6E**, *center*). To reveal which parameters are driving the separation, we visualized relationship between variables using correlation plot (**Fig. 6F**). T cell parameters were driving the separation of immunophenotypic group 1, while antibody responses separated immunophenotypic group 2 (**Fig. 6F**). The most important variables in explaining the variability between SARS-CoV-2 positive individuals in immunophenotypic group 1 were total IFN-γ ELISpot T cells, S1 and S2-stimulated IFN-γ ELISpot T cells, and anti-S IgG, anti-RBD IgG, ADCD, S-IgG from OC43 and HcoV-HKU1 in immunophenotypic group 2 that were correlated with principal components 1 and 2 (PC1-PC2) (**Fig. 6G and 6H**). The correlation plot revealed positive correlation between antibody responses, and negative correlation between T cell responses with the time when samples were

obtained (**Fig. 6F**). To further examine these associations between immunological parameters, we performed correlation analysis, which confirmed strong positive correlation between antibody and T cells responses (**Fig. 6I**). The antibodies directed against N, S and RBD from SARS-CoV-2, were positively correlated with antibody functionality, such as pseudoneutralising capacity and ADCD, ADNP and ADMP, and positively correlated with IFN-γ ELISpot T cell responses against S1, S2 and N (**Fig. 6I**). The antibody responses to S protein from other circulating coronaviruses, such as SARS-CoV-1, MERS, HcoV-HKU1, 229e and OC43 were also contained in this cluster being positively correlated with antibody and T cell responses (**Fig. 6I**). This cluster was negatively correlated with time, confirming the observations from primary analysis (**Fig. 6I**). Notably, there was a negative correlation between NL63 S antibodies and S and RBD SARS-CoV-2 specific antibodies (**Fig. 6I**). There were other apparent relationships in two other clusters identified, that were not associated with time, including positive correlation between proliferating T cells stimulated with different SARS-CoV-2-specific peptides, and positive correlation between ADNKA and S-IgA and S-IgG1, while negative correlation with S-IgM (**Fig. 6I**).

The integrative analysis revealed three distinct immunophenotypic groups of SARS-CoV-2 infected individuals strongly connected to cellular and humoral immune profiling beyond the disease severity and clinical parameters.

Identifying an early immunological signature associated with a durable immune response to SARS-CoV-2

To elucidate an early immunological signature that could predict whether an individual will mount a durable and protective immunity against SARS-CoV-2 6 months after infection, we stratified SARS-CoV-2 infected individuals into high and low responders, based on the seropositivity status (N IgG titres >=1.4), which has recently been identified as a correlate of protection [35]. We then asked whether the components of cellular or humoral immunity within one month of infection (28 days pso) were predictive of the ability of individuals to develop protective immunity against SARS-CoV-2 (6 months pso). First, using an unsupervised machine learning approach, i.e., hierarchical clustering of integrated immunological data on day 28 pso, we identified two groups of SARS-CoV-2 infected individuals based on the response status 6 months pso (**Fig. 7A**). While the majority of SARS-CoV-2 infected individuals with mild disease would mount protective immunity 6 months pso and become high responders, there was a proportion of individuals with mild disease that failed to mount durable and protective immunity (low responders) (**Fig. 7A**). The majority of individuals with asymptomatic infection were low responders. High responders mounted stronger antibody responses, in particular N-IgG and pseudo-neutralising antibodies, and overall, stronger T cell responses, including IFN-γ-positive and proliferating T cells, than low responders 28 days pso (**Fig. 7A**). Antibody responses to spike protein from 229e and NL63, B cell ELISpot and ADNKA were increased in low responders early after SARS-CoV-2 infection in comparison to high responders (**Fig. 7A**).

To further define the immunological features that can distinguish individuals with durable and protective immunity and predict if the individual is on the trajectory to become a high or low responder, we used

the SIMON supervised machine learning approach [28, 29]. We generated 30 resamples and tested 3,565 models using 172 machine learning algorithms (*Materials and methods*). The best performing model built using Sparse Partial Least Squares (sPLS) algorithm (train AUROC: 0.95 (CI 0.5-1) and test AUROC: 1) used only 8 out of 49 measured parameters on day 28 pso to predict if the individual will become high or low responder 6 months pso (**Fig. 7B**). The features that were contributing the most to this model included antibody responses to N and S, ADCD and pseudo-neutralising antibodies to SARS-CoV-2, and T cell IFN-γ ELISpot (S1/S2, M and total positive T cells) which were significantly increased in high responders 28 days pso compared to low responders (**Fig. 7C and 7D**). Together, these data indicate that early generation of antibodies with high binding, neutralising and effector function, and functional T cell responses following infection can predict the responsiveness potential, i.e., protection and duration of SARS-CoV-2 immunity of the individual. Additionally, these findings suggest that a coordinated action of both T and B cells early after infection is required for establishment of durable and protective immunity.

The generation of durable and functional humoral and cellular immunity in a proportion of SARS-CoV-2 infected individuals (high responders) may provide protection against re-infection, including also against variants of concern (VOCs). Thus, we assessed the neutralising antibody responses in high and low responders against the infecting (Victoria) strain and against variants B.1.1.7 and B.1.351 (**Fig. 7E**). Individuals with durable and protective SARS-CoV2 immunity shown high neutralisation antibody titres against wild-type circulating SARS-CoV-2 (Victoria) strain, and against two novel variants, including B.1.1.7 (alpha) and B.1.351 (beta) (**Fig. 7E**). High responders had significantly higher neutralising antibody titres against B.1.1.7 alpha variant one-month pso, and these higher neutralising antibodies were preserved 6 months pso (**Fig. 7E**).

Altogether, these data suggest that generation of immunity to SARS-COV-2 shows distinct trajectories following early priming, and early antibody responses are important to mediate protective and durable immunity that can also provide protection against novel variants.

Discussion

Key questions on the trajectory of the SARS-CoV-2 specific immune response to natural infection, and the maintenance of immune memory remain highly relevant even as highly effective vaccines are being rolled out worldwide. Firstly, even with high availability of vaccines there will always be a pool of unvaccinated people due to vaccine hesitancy or access difficulties, and this will include people who have had natural infection. Secondly, as of June 2021 only 12% of the world's population is estimated to have received at least one dose of vaccine [36], so for much of the immunity globally is from natural infection, which remains a cornerstone of population-level immunity. Thirdly, measuring immune responses to antigens not included in spike-containing vaccines are used as biomarkers of previous SARS-CoV-2 infection and as such are widely used to stratify immune responses to vaccination, since prior SARS-CoV-2 is known to enhance vaccine responsiveness [37, 38]. Finally, understanding how the early immune response translates into lasting immunity towards emerging variants of concern is crucial

to accelerate predictions of population risk and to drive policy. In this manuscript, we characterise the magnitude, function and maintenance of humoral and cellular T and B cell immunity, and the relationship between clinical and multi parametric immune data. We then evaluate the ability of antibodies to neutralise live SARS-CoV-2 virus 6 months after primary infection to variants of concern and provide insight into the early predictors of durable neutralising antibody after natural infection.

Compatible with other studies [12, 39, 40, 41], our data shows a peak of anti-NP and anti-S binding antibody (IgG) magnitude 28 days after onset of symptoms, with anti-NP responses declining over the next five months, although these responses remain above the threshold of detection in the majority. In contrast, anti-S IgG responses were well maintained, in keeping with the reported longer half-life for decay of anti-S IgG responses compared with anti-NP IgG responses [12], along with maintenance of B cell memory. Neutralisation measured by a pseudo-neutralisation assay showed a decline over time but was generally maintained six months following infection. High levels of neutralisation were seen earlier post symptom onset (from 7 days) compared with the IgG binding assays, which may represent contributions from IgM [42] and IgA [43]. Some of the observed decline in neutralising antibodies over time may represent a threshold effect – NAb are a subset of total IgG such that gradual declines over time are first measurable in NAb, but biologically important neutralisation may still occur below the detection threshold. Fc-mediated functionality including antibody dependent NK activation, phagocytosis and complement deposition was maintained over the 6 months duration which may make an important contribution to protective immunity and was significantly associated with increasing disease severity.

Taken together, B cell polyfunctionality was lower in those with asymptomatic infection, compared with those with mild disease early after infection (day 28), though by 6 months the profiles between the cohorts looked similar. The most notable changes were a reduction in IgM spike responses but a relative maintenance of IgG3 spike responses in the mild cohort that was not seen in the asymptomatic cohort.

Previous studies have shown that early distinct antigenic targets and qualitative features of SARS-CoV-2-specific antibodies are associated with disease trajectory [44, 45], whilst multifunctional antibody responses, and particularly ADCD and ADNP, following adoptive transfer of IgG from convalescent rhesus macaques have been shown to contribute to protection from SARS-CoV-2 challenge [46]. Furthermore, vaccine-induced Fc-mediated polyfunctionality has been observed following administration of efficacious vaccines in both macaque and human studies [31, 47]. While the capacity of Fc receptor binding appears to be lower in convalescent individuals against VOCs, evidence is emerging of maintenance of vaccine-induced Fc-functional antibody properties against VOCs supporting resilience of humoral immunity against VOCs independent of neutralisation [48].

In evaluating SARS-CoV-2 specific effector T cell responses over six months in an IFN-γ ELISpot assay, we showed that there was significant heterogeneity in the magnitude of responses between individuals as previously reported [12, 49, 50]. The majority of people showed robust T cell responses in the first 28 days after infection, though these were significantly lower in the asymptomatic cohort. Within 3 months

of infection there was a marked decline in T cell responses and by 6 months, these were reduced by 75% and were undetectable in approximately 20%. We used a flow cytometry based 7-day proliferation assay to assess memory T responses of both CD4+ and CD8+ T cell subsets to show a dominant CD4+ T cell subset response. Although memory proliferative responses have been shown to "mature" over time, particularly following vaccination [51, 52], we show that proliferative responses (both CD4 and CD8), targeting Spike, M, and NP decline markedly between day 28 and day 180. ICS analysis showed that CD4+ T cells were the dominant subset targeting S1, S2 and M antigens, whilst NP were targeted by both CD4+ and CD8+ T cells, and NSP3B was targeted by CD8+ T cells. Polyfunctional T cells, producing multiple cytokines, were generated at day 28, and although the magnitude of the response declined, polyfunctionality was generally retained out to 6 months.

In our study we show that symptomatic infection is associated with more robust cellular and humoral immune responses compared to the asymptomatic group early after PCR+ confirmed infection. An association between asymptomatic infection and lower antibody responses has been previously reported [53], and we and others have shown a correlation between disease severity and higher levels of antibody and T cell responses in early disease [4, 54]. Similar results have been reported in other disease settings including robust immune responses associated with disease severity in H1N1/09 influenza A [55]. In contrast, a previous prospective SARS-CoV-2 screening study has observed that asymptomatic infection is associated with highly functional cellular immune responses [56]. Either way, humoral and cellular immune responses measured months after primary infection is found at low magnitude following asymptomatic infection. These findings raise the possibility that people with asymptomatic SARS-CoV-2 infection may have less protective immunity months after primary infection. A limitation to our study, is that the timing of infection onset in asymptomatic HCW, (even though PCR+) is not precisely defined. As such, it is theoretically possible that the asymptomatic individuals in our study are later in their disease course at detection, which was further explored by integrative analysis.

To elucidate the trajectory of the immune response of SARS-CoV-2 infected individuals over time and identify signatures associated with the maintenance of protective immunity, we performed an integrative analysis in the cohort of 86 individuals on all 433 samples. The results of the integrative analysis led to several key findings. First is the identification of immunophenotypic groups of SARS-CoV-2 infected individuals beyond disease time course and disease severity. By integrating over 70 immune parameters with clinical data, disease severity and temporal changes, we generated a computational model using t-SNE embedding algorithm that coupled immunological phenotypes of each individual with the disease severity and other clinical parameters. The t-SNE representation of integrated data revealed minimal clustering by time point, suggesting that heterogeneity of the immune response during the SARS-CoV-2 infection is independent of the time course during the infection. While some of the individuals with asymptomatic infection may be later in their disease course at detection, the majority did not cluster with the samples obtained from individuals with mild or severe infection at later timepoints after the infection. The major separation of individuals with asymptomatic disease was driven by clinical parameters, while the mild cohort clustered into 2 immunophenotypic groups (not driven by clinical

parameters), one of which shared phenotype with the severe disease cohort. The PCA analysis provided further support for the heterogeneity of the immune responses in the SARS-CoV-2 infected individuals with mild disease and separation into three immunophenotypic groups, confirming that 38.6% of variance between individuals was explained by the immunological data. The results suggested that immunophenotypic group 1, exhibiting robust binding (anti-N and anti-S) and functional (pseudoneutralising and ADCD/ADMP) antibody responses and memory B cell involvement, shared similarity with individuals with severe disease, while immunophenotypic group 2 composed of functional IFN-γ T cell responses represented an unique proportion of individuals with mild disease, early in the course of the disease (as indicated by negative correlation with time when samples were acquired). The third immunophenotypic group – defined by the lower overall antibody and T cell responses - shared similarities with the asymptomatic cohort, suggesting that some individuals may fail to develop robust antibody and T cell responses despite having mild infection. These results support the magnitude of the immune response being determined by factors beyond disease severity, including viral factors and the individual's immunocompetence. Using correlation analyses, we observed a positive association between spike and nucleocapsid T cell and antibody responses (both decreased with time, confirming the primary analysis) and cross-reactivity to other coronaviruses which correlated with spike and nucleocapsid T cell and antibody responses (NL63 is negatively correlated and OC43 is positively correlated), substantiating the findings that immunity may be defined by immunocompetence and previous exposure to circulating coronaviruses.

To further delineate this observation, we performed integrative analysis using baseline parameters only (measured on day 28 after infection), and this led to the second key finding – identification of an early immunological signature that is associated with durable and protective SARS-CoV-2 immunity. Using hierarchical clustering approach and integrated baseline cellular and humoral immune parameters, we observed distinct clustering of high and low responders at this early time point. High anti-N IgG, along with more robust overall T cell responses (including IFN-γ ELISpot and proliferation) at baseline with a low response to seasonal coronaviruses (NL63 and 229e) dominated in the high responder group, whilst low responders had lower anti-N IgG and overall T cell responses and had more pronounced cross-reactive seasonal CoV responses (NL63 and 229e) at baseline. The final major finding was the ability to predict if the individual will generate durable and protective SARS-CoV-2 immunity 6 months post infection based on the early immunological signature one month after infection. With the use of SIMON data mining tool and generation of more than 3,500 predictive machine learning models, we identified upregulation of antibody responses (spike and NP, with pseudoneutralising and ADCD functions) combined with the more robust T cell responses as predictors of individuals who will generate durable and protective immunity 6 months post infection (high responders). The predictive model built by SIMON suggests a link between both arms of the immune response - cellular and humoral immunity – with the durability of the SARS-CoV-2 protective immunity. Thus, this early immunological signature may determine essential differences of the trajectory that each individual will take after SARS-CoV-2 infection. Importantly, the sera of the individuals who will go on to generate durable and protective SARS-CoV-2 immunity (high responders) 6 months post infection, were better able to neutralise both

744 the Victoria strain (the likely infection strain), and also the VOCs (B.1.1.7 - alpha and B.1.351 - beta)
745 one month after infection, and such protective neutralising antibody responses were durable (as
746 measured 6 months post infection). In contrast, those who were low responders 6 months after infection
747 showed a reduction in the capacity to neutralise the Victoria strain, with a severe loss of neutralisation
748 against both VOC - particularly B1.351.
749
750 Overall, our data reveal the highly variable range of immunity after SARS-CoV-2 infection and suggest
751 that immune events primed during early SARS-CoV-2 infection may define the subsequent trajectories
752 leading to the effective maintenance or loss of long-term SARS-CoV-2 protective immunity as measured
753 by neutralising antibodies. Importantly, previous infection may not give ongoing protection against VOC
754 months later, and people with asymptomatic infection had lower responses at all time points across
755 many of the immune parameters we measured. Maintenance of immune memory over time is critically
756 required for the effective neutralisation of VOC that is most likely to confer sterilising immunity, whilst
757 other immune mechanisms including non-neutralising antibodies and T cells may account for the
758 protection against severe disease, including for VOC [57, 58, 59, 60]. This study provides a basis for more
759 targeted vaccination programme of previously infected individuals based on early immunological
760 signature 28 days after infection.
761
762 **Figure Legends**
763
764 **Figure 1: Longitudinal humoral immune responses in individuals with PCR confirmed SARS-**
765 **CoV-2 asymptomatic, mild or severe infection.**
766 Humoral immune responses were assessed in acute and convalescent by binding antibody ELISA for
767 total IgG specific to the **(A)** Nucleopcapsid and **(B)** Spike glycoprotein, quantification of **(C)** IgG
768 memory B cells specific to the spike glycoprotein, and **(D)** pseudoneutralisation antibody titres. Boxplots
769 represent the median with interquartile range, a Wilcoxon rank-sum test was used to compare between
770 study time points. A generalised additive mixed model (GAMM) by restricted maximum likelihood —
771 right-hand plots — was used to fit the immunological measures (log10 transformed) taken at multiple
772 study time points, using Gaussian process smooth term. Disease severity group was included in the
773 GAMM as a linear predictor and a participant identifier was included as a random effect. See Table S1
774 for number of individuals evaluated per assay.
775
776 **Figure 2: Antibody isotype, subclass and function in individuals with PCR confirmed SARS-**
777 **CoV-2 asymptomatic, mild or severe infection.**
778 SARS-CoV-2 spike-specific antibody isotype and subclasses measured post-infection: **(A)** IgM, **(B)** IgA,
779 **(C)** IgG1 and **(D)** IgG3. Antibody function measure post-SARS-CoV-2 infection: **(E)** antibody-dependent
780 NK cell activation (ADNKA), **(F)** antibody-dependent neutrophil phagocytosis (ADNP), **(G)** antibody-
781 dependent monocyte phagocytosis (ADMP) and **(H)** antibody-dependent complement deposition
782 (ADCD). **(I)** Polar plot of various antibody isotype, subclass and function data, minimum-maximum
783 normalised. Boxplots represent the median with interquartile range, a Wilcoxon rank-sum test was
784 used to compare between study time points. A generalised additive mixed model (GAMM) by restricted

maximum likelihood — right-hand plots — was used to fit the immunological measures (log10 transformed) taken at multiple study time points, using Gaussian process smooth term. Disease severity group was included in the GAMM as a linear predictor and a participant identifier was included as a random effect. See Table S1 for number of individuals evaluated per assay.

Figure 3: Longitudinal specific-IgG and memory B cell responses to spike protein from non-SARS-CoV-2 coronaviruses.

(A) Meso Scale Discovery (MSD) multiplexed immunoassay (MIA) platform measurements of antibody levels to spike protein from non-SARS-CoV-2 coronaviruses. **(B)** Memory B cells responses to spike protein from non-SARS-CoV-2 coronaviruses. See Table S1 for number of individuals evaluated per assay.

Figure 4 Magnitude of SARS-CoV-2 specific Effector T cell Response.

(A) *Ex vivo* IFN-γ ELISpot showing the effector T cell responses to summed SARS-CoV-2 peptide pools spanning spike, accessory and structural proteins (summed total of SARS-CoV-2 proteins tested, S1, S2, NSP3B, M, NP, ORF 3, ORF8 and the CEFT positive control peptides for T cell responses) in 78 individuals 28, 90 and 180 days after mild or asymptomatic SARS-CoV-2 infection (onset of symptoms for mild cases, PCR positive test for asymptomatic participants). **(B)** Heatmap displaying unsupervised hierarchical clustering of the ELISpot data in (A) and disease severity (mild or asymptomatic) for the original SARS-CoV-2 diagnosis. Sfu / million PBMCs = spot forming units per million peripheral blood mononuclear cells, with background subtracted. D28, d90 and d180 = days after SARS-CoV-2 diagnosis. Grey regions on heatmap represent missing data due to insufficient cells. Plots show median with error bars indicating +/- IQR. Friedman test with Dunn's multiple comparisons test was performed.

Figure 5. Proliferative responses to SARS-CoV-2 peptide pools at 1- and 6-months post infection

Proliferative responses against **(A)** SARS-CoV-2 proteins S1, S2, M, NP, ORF3 and ORF8 presented in CD4+ (Left hand panel) and CD8+ (Right hand panel) T cells measured at 28 and 180 days pso for volunteers with mild disease or days post PCR positivity for asymptomatic disease (asymptomatic n = 8, mild disease n = 49). Kruskal Wallis T test, all P values are all stated on plots. **(B)** shows unsupervised hierarchical clustering showing visual representation of SARS-CoV-2 specific responses at day 28 and 180 in both CD4+ and CD8+ T cell compartments and **(C)** comparative analysis of SARS-CoV-2 specific CD4+ and CD8+ T cell responses at day 28 (top panel) and day 180 (bottom panel) in both asymptomatic and mild groups (analysed as one group). Kruskal Wallis T test, all P values are all stated on plots.

Figure 6. Integrative analysis of clinical and longitudinal immunological data reveals distinct immunophenotypic groups of SARS-CoV-2 infected individuals. (A) Clinical study overview. **(B)** t-SNE map of integrated clinical and immunological data color-coded based on timepoint or disease severity. **(C)** Clustered t-SNE analysis. **(D)** Heatmap of clinical and immune parameters across three identified clusters. **(E)** PCA plot representing integrated immunological data, grouped based on the

disease severity. Percentage indicates the variance explained by the principal component (PC). **(F)** Variable correlation plot. Positively correlated variables are grouped together, while negatively correlated variables are positioned on opposite quadrants. The distance between variables and the origin measures the quality of the variables on the factor map, while the colour indicated the quality of representations as cos2. **(G)** Quality of variable representations (color-coded, cos2) and contributions of variables to principal components 1 and 2 (size of the circle). **(H)** Top 10 variables and their contribution to PC 1 and 2. **(I)** Correlations of immunological parameters with time component across samples. Spearman's correlation coefficient (colour coded) and only significant values shown (after adjusted FDR <0.05). Black boxes indicate clusters (hierarchical clustering).

Figure 7. Early signature of durable SARS-CoV2 protective immunity. (A) Hierarchical clustering heatmap of immune parameters on day 28 pso, grouping by responder status 6 months pso and disease severity. Results obtained using complete linkage agglomeration method, dendrogram ordered tightest cluster first. **(B)** Integrative immunological dataset containing 3,626 datapoints (49 features and 74 donors) was used for SIMON analysis to predict if the individual will generate high or low anti-N antibody responses 6 months pso. In total, 184 ML algorithms were tested and 2,556 model built. ROC plot of the best performing model built with the svmPoly algorithm. Train AUROC (black line) is determined using 10-fold cross-validation and test AUROC evaluated on the independent test set (25% of the initial dataset). **(C)** Top variables that contribute to the model and are increased in high relative to low responders. **(D)** Frequency of selected variables on day 28pso (bars show mean with SEM). Mann-Whitney test ($p<0.05$). **(E)** Neutralisation assay against wild-type SARS-CoV2 (Victoria), and two novel variants (B1.1.7 and B1.351) between high and low responders on two timepoints (one and 6 months pso). Plots show mean with SEM. Kruskal-Wallis, with Dunn's multiple comparison test ($p<0.05$) was performed.

References

1. Berlin DA, Gulick RM, Martinez FJ. Severe Covid-19. *N Engl J Med*, (2020).

2. Eyre DW, Lumley SF, O'Donnell D, Campbell M, Sims E, Lawson E, . . . Walker TM. Differential occupational risks to healthcare workers from SARS-CoV-2 observed during a prospective observational study. *eLife* **9**, e60675 (2020).

3. Fan VS, Dominitz JA, Eastment MC, Locke E, Green P, Berry K, . . . Ioannou GN. Risk Factors for testing positive for SARS-CoV-2 in a national US healthcare system. *Clin Infect Dis*, (2020).

4. Peng Y, Mentzer AJ, Liu G, Yao X, Yin Z, Dong D, . . . Investigators IC. Broad and strong memory CD4+ and CD8+ T cells induced by SARS-CoV-2 in UK convalescent individuals following COVID-19. *Nat Immunol* **21**, 1336-1345 (2020).

5. Ogbe A, Kronsteiner B, Skelly DT, Pace M, Brown A, Adland E, . . . Oxford Protective TCIfC-CT. T cell assays differentiate clinical and subclinical SARS-CoV-2 infections from cross-reactive antiviral responses. *Nature Communications* **12**, 2055 (2021).

6. Wang Z, Yang X, Zhong J, Zhou Y, Tang Z, Zhou H, . . . Ran P. Exposure to SARS-CoV-2 generates T-cell memory in the absence of a detectable viral infection. *Nat Commun* **12**, 1724 (2021).

7. Reynolds CJ, Swadling L, Gibbons JM, Pade C, Jensen MP, Diniz MO, . . . Boyton RJ. Discordant neutralizing antibody and T cell responses in asymptomatic and mild SARS-CoV-2 infection. *Sci Immunol* **5**, (2020).

8. Reuters. https://www.reuters.com/business/healthcare-pharmaceuticals/italy-give-just-one-covid-shot-some-patients-eu-struggles-with-inoculations-2021-03-04/ Accessed 13 June 2021.

9. Lumley SF, Wei J, O'Donnell D, Stoesser NE, Matthews PC, Howarth A, . . . Oxford University Hospitals Staff Testing G. The duration, dynamics and determinants of SARS-CoV-2 antibody responses in individual healthcare workers. *Clin Infect Dis*, (2021).

10. Sagar M, Reifler K, Rossi M, Miller NS, Sinha P, White LF, Mizgerd JP. Recent endemic coronavirus infection is associated with less-severe COVID-19. *J Clin Invest* **131**, (2021).

11. Grifoni A, Weiskopf D, Ramirez SI, Mateus J, Dan JM, Moderbacher CR, . . . Sette A. Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell* **181**, 1489-1501 e1415 (2020).

12. Dan JM, Mateus J, Kato Y, Hastie KM, Yu ED, Faliti CE, . . . Crotty S. Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science*, (2021).

13. Ansari A, Arya R, Sachan S, Jha SN, Kalia A, Lall A, . . . Gupta N. Immune Memory in Mild COVID-19 Patients and Unexposed Donors Reveals Persistent T Cell Responses After SARS-CoV-2 Infection. *Front Immunol* **12**, 636768-636768 (2021).

14. Cromer D, Juno JA, Khoury D, Reynaldi A, Wheatley AK, Kent SJ, Davenport MP. Prospects for durable immune control of SARS-CoV-2 and prevention of reinfection. *Nature reviews Immunology* **21**, 395-404 (2021).

15. Wheatley AK, Juno JA, Wang JJ, Selva KJ, Reynaldi A, Tan H-X, . . . Kent SJ. Evolution of immune responses to SARS-CoV-2 in mild-moderate COVID-19. *Nature communications* **12**, 1162-1162 (2021).

16. Zuo J, Dowell AC, Pearce H, Verma K, Long HM, Begum J, . . . Moss P. Robust SARS-CoV-2-specific T cell immunity is maintained at 6 months following primary infection. *Nat Immunol* **22**, 620-626 (2021).

17. Galanti M, Shaman J. Direct observation of repeated infections with endemic coronaviruses. *The Journal of Infectious Diseases*, (2020).

18. Kiyuka PK, Agoti CN, Munywoki PK, Njeru R, Bett A, Otieno JR, . . . Cotten M. Human Coronavirus NL63 Molecular Epidemiology and Evolutionary Patterns in Rural Coastal Kenya. *The Journal of infectious diseases* **217**, 1728-1739 (2018).

19. Davies NG, Abbott S, Barnard RC, Jarvis CI, Kucharski AJ, Munday JD, . . . Edmunds WJ. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science* **372**, (2021).

20. Supasa P, Zhou D, Dejnirattisai W, Liu C, Mentzer AJ, Ginn HM, . . . Screaton GR. Reduced neutralization of SARS-CoV-2 B.1.1.7 variant by convalescent and vaccine sera. *Cell*, (2021).

21. Davies NG, Jarvis CI, Group CC-W, Edmunds WJ, Jewell NP, Diaz-Ordaz K, Keogh RH. Increased mortality in community-tested cases of SARS-CoV-2 lineage B.1.1.7. *Nature* **593**, 270-274 (2021).

22. Zhou D, Dejnirattisai W, Supasa P, Liu C, Mentzer AJ, Ginn HM, . . . Screaton GR. Evidence of escape of SARS-CoV-2 variant B.1.351 from natural and vaccine-induced sera. *Cell* **184**, 2348-2361 e2346 (2021).

23. Tegally H, Wilkinson E, Giovanetti M, Iranzadeh A, Fonseca V, Giandhari J, . . . de Oliveira T. Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature* **592**, 438-443 (2021).

24. Faria NR, Mellan TA, Whittaker C, Claro IM, Candido DDS, Mishra S, . . . Sabino EC. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science* **372**, 815-821 (2021).

25. Dejnirattisai W, Zhou D, Supasa P, Liu C, Mentzer AJ, Ginn HM, . . . Screaton GR. Antibody evasion by the P.1 strain of SARS-CoV-2. *Cell*.

26. Planas D, Veyer D, Baidaliuk A, Staropoli I, Guivel-Benhassine F, Rajah MM, . . . Schwartz O. Reduced sensitivity of infectious SARS-CoV-2 variant B.1.617.2 to monoclonal antibodies and sera from convalescent and vaccinated individuals. *bioRxiv*, 2021.2005.2026.445838 (2021).

27. Skelly DT, Harding AC, Gilbert-Jaramillo J, Knight ML, Longet S, Brown A, . . . James W, S,. Two doses of SARS-CoV-2 vaccination induce more robust immune responses to emerging SARS-CoV-2 variants of concern than does natural infection. *Research Square*, (2021).

28. Tomic A, Tomic I, Rosenberg-Hasson Y, Dekker CL, Maecker HT, Davis MM. SIMON, an Automated Machine Learning System, Reveals Immune Signatures of Influenza Vaccine Responses. *Journal of immunology (Baltimore, Md : 1950)* **203**, 749-759 (2019).

29. Tomic A, Tomic I, Waldron L, Geistlinger L, Kuhn M, Spreng RL, . . . Davis MM. SIMON: Open-Source Knowledge Discovery Platform. *Patterns (New York, NY)* **2**, 100178-100178 (2021).

30. Folegatti PM, Ewer KJ, Aley PK, Angus B, Becker S, Belij-Rammerstorfer S, . . . Oxford CVTG. 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. *Lancet* **396**, 467-478 (2020).

31. Barrett JR, Belij-Rammerstorfer S, Dold C, Ewer KJ, Folegatti PM, Gilbride C, . . . Oxford CVTG. Phase 1/2 trial of SARS-CoV-2 vaccine ChAdOx1 nCoV-19 with a booster dose induces multifunctional antibody responses. *Nat Med* **27**, 279-288 (2021).

32. Frey A, Di Canzio J, Zurakowski D. A statistically defined endpoint titer determination method for immunoassays. *J Immunol Methods* **221**, 35-41 (1998).

33. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York (2016).

34. Wood S, Scheipl F. gamm4: Generalized Additive Mixed Models using 'mgcv' and 'lme4'. R package version 0.2-6.) (2020).

35. Lumley SF, O'Donnell D, Stoesser NE, Matthews PC, Howarth A, Hatch SB, . . . Oxford University Hospitals Staff Testing G. Antibody Status and Incidence of SARS-CoV-2 Infection in Health Care Workers. *N Engl J Med*, (2020).

36. Mathieu E, Ritchie H, Ortiz-Ospina E, Roser M, Hasell J, Appel C, . . . Rodés-Guirao L. A global database of COVID-19 vaccinations. *Nature Human Behaviour*, (2021).

37. Krammer F, Srivastava K, Alshammary H, Amoako AA, Awawda MH, Beach KF, . . . Simon V. Antibody Responses in Seropositive Persons after a Single Dose of SARS-CoV-2 mRNA Vaccine. *N Engl J Med* **384**, 1372-1374 (2021).

38. Prendecki M, Clarke C, Brown J, Cox A, Gleeson S, Guckian M, . . . Willicombe M. Effect of previous SARS-CoV-2 infection on humoral and T-cell responses to single-dose BNT162b2 vaccine. *Lancet* **397**, 1178-1181 (2021).

39. Röltgen K, Powell AE, Wirz OF, Stevens BA, Hogan CA, Najeeb J, . . . Boyd SD. Defining the features and duration of antibody responses to SARS-CoV-2 infection associated with disease severity and outcome. *Science Immunology* **5**, eabe0240 (2020).

40. Wajnberg A, Amanat F, Firpo A, Altman DR, Bailey MJ, Mansour M, . . . Cordon-Cardo C. Robust neutralizing antibodies to SARS-CoV-2 infection persist for months. *Science* **370**, 1227 (2020).

41. Lumley SF, Wei J, O'Donnell D, Stoesser NE, Matthews PC, Howarth A, . . . Oxford University Hospitals Staff Testing G. The duration, dynamics and determinants of SARS-CoV-2 antibody responses in individual healthcare workers. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, ciab004 (2021).

42. Klingler J, Weiss S, Itri V, Liu X, Oguntuyo KY, Stevens C, . . . Hioe CE. Role of Immunoglobulin M and A Antibodies in the Neutralization of Severe Acute Respiratory Syndrome Coronavirus 2. *The Journal of infectious diseases* **223**, 957-970 (2021).

43. Sterlin D, Mathian A, Miyara M, Mohr A, Anna F, Claër L, . . . Gorochov G. IgA dominates the early neutralizing antibody response to SARS-CoV-2. *Sci Transl Med* **13**, eabd2223 (2021).

44. Atyeo C, Fischinger S, Zohar T, Slein MD, Burke J, Loos C, . . . Alter G. Distinct Early Serological Signatures Track with SARS-CoV-2 Survival. *Immunity* **53**, 524-532 e524 (2020).

45. Zohar T, Loos C, Fischinger S, Atyeo C, Wang C, Slein MD, . . . Alter G. Compromised Humoral Functional Evolution Tracks with SARS-CoV-2 Mortality. *Cell* **183**, 1508-1519 e1512 (2020).

46. McMahan K, Yu J, Mercado NB, Loos C, Tostanoski LH, Chandrashekar A, . . . Barouch DH. Correlates of protection against SARS-CoV-2 in rhesus macaques. *Nature* **590**, 630-634 (2021).

47. Gorman MJ, Patel N, Guebre-Xabier M, Zhu A, Atyeo C, Pullen KM, . . . Alter G. Collaboration between the Fab and Fc contribute to maximal protection against SARS-CoV-2 in nonhuman primates following NVX-CoV2373 subunit vaccine with Matrix-M vaccination. *bioRxiv*, (2021).

48. Kaplonek P, Fischinger S, Cizmeci D, Bartsch Y, Kang J, Burke J, . . . Alter G. Resilient Fc-Effector Functions Across SARS-CoV-2 Variants of Concern Following mRNA-1273 Vaccination.). SSRN 10.2139/ssrn.3832979 (2021).

49. Reynolds CJ, Swadling L, Gibbons JM, Pade C, Jensen MP, Diniz MO, . . . Boyton RJ. Discordant neutralizing antibody and T cell responses in asymptomatic and mild SARS-CoV-2 infection. *Science Immunology* **5**, eabf3698 (2020).

50. Zuo J, Dowell AC, Pearce H, Verma K, Long HM, Begum J, . . . Moss P. Robust SARS-CoV-2-specific T cell immunity is maintained at 6 months following primary infection. *Nature Immunology*, (2021).

51. Ahlers JD, Belyakov IM. Memories that last forever: strategies for optimizing vaccine T-cell memory. *Blood* **115**, 1678-1689 (2010).

52. Sallusto F, Lanzavecchia A, Araki K, Ahmed R. From vaccines to memory and back. *Immunity* **33**, 451-463 (2010).

53. Wellinghausen N, Plonné D, Voss M, Ivanova R, Frodl R, Deininger S. SARS-CoV-2-IgG response is different in COVID-19 outpatients and asymptomatic contact persons. *J Clin Virol* **130**, 104542 (2020).

54. Lynch KL, Whitman JD, Lacanienta NP, Beckerdite EW, Kastner SA, Shy BR, . . . Wu AHB. Magnitude and Kinetics of Anti-Severe Acute Respiratory Syndrome Coronavirus 2 Antibody Responses and Their Relationship to Disease Severity. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **72**, 301-308 (2021).

55. Zhao Y, Zhang YH, Denney L, Young D, Powell TJ, Peng YC, . . . Dong T. High levels of virus-specific CD4+ T cells predict severe pandemic influenza A virus infection. *Am J Respir Crit Care Med* **186**, 1292-1297 (2012).

56. Le Bert N, Clapham HE, Tan AT, Chia WN, Tham CYL, Lim JM, . . . Tam CC. Highly functional virus-specific cellular immune response in asymptomatic SARS-CoV-2 infection. *J Exp Med* **218**, (2021).

57. Vasileiou E, Simpson CR, Shi T, Kerr S, Agrawal U, Akbari A, . . . Sheikh A. Interim findings from first-dose mass COVID-19 vaccination roll-out and COVID-19 hospital admissions in Scotland: a national prospective cohort study. *Lancet (London, England)* **397**, 1646-1657 (2021).

58. Lopez Bernal J, Andrews N, Gower C, Robertson C, Stowe J, Tessier E, . . . Ramsay M. Effectiveness of the Pfizer-BioNTech and Oxford-AstraZeneca vaccines on covid-19 related symptoms, hospital admissions, and mortality in older adults in England: test negative case-control study. *BMJ* **373**, n1088 (2021).

59. Sadoff J, Gray G, Vandebosch A, Cárdenas V, Shukarev G, Grinsztejn B, . . . Douoguih M. Safety and Efficacy of Single-Dose Ad26.COV2.S Vaccine against Covid-19. *N Engl J Med* **384**, 2187-2201 (2021).

60. Fischer RJ, van Doremalen N, Adney DR, Yinda CK, Port JR, Holbrook MG, . . . Munster VJ. ChAdOx1 nCoV-19 (AZD1222) protects hamsters against SARS-CoV-2 B.1.351 and B.1.1.7 disease. Preprint at https://doi.org/10.1101/2021.03.11.435000 (2021).

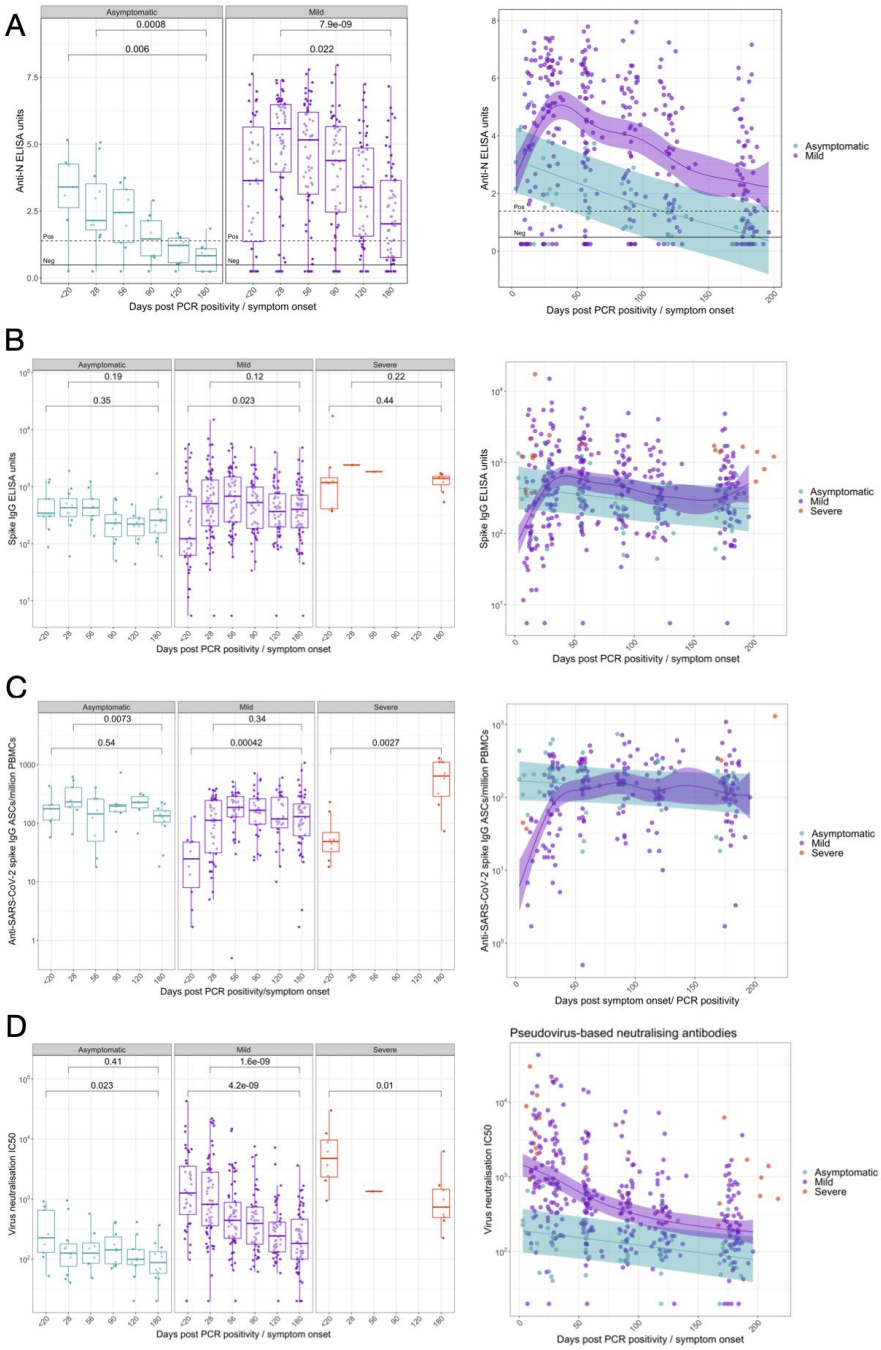

A
Asymptomatic
Mild
0.0008
0.006
7.9e-09
0.022
Anti-N ELISA units
Pos
Neg
Days post PCR positivity / symptom onset
Asymptomatic
Mild
B
Asymptomatic
Mild
Severe
0.19
0.35
0.12
0.023
0.22
0.44
Spike IgG ELISA units
Days post PCR positivity / symptom onset
Asymptomatic
Mild
Severe
C
0.0073
0.54
0.34
0.00042
0.0027
Anti-SARS-CoV-2 spike IgG ASCs/million PBMCs
Days post PCR positivity/symptom onset
Days post symptom onset/ PCR positivity
D
0.41
0.023
1.6e-09
4.2e-09
0.01
Virus neutralisation IC50
Pseudovirus-based neutralising antibodies

Figure 1

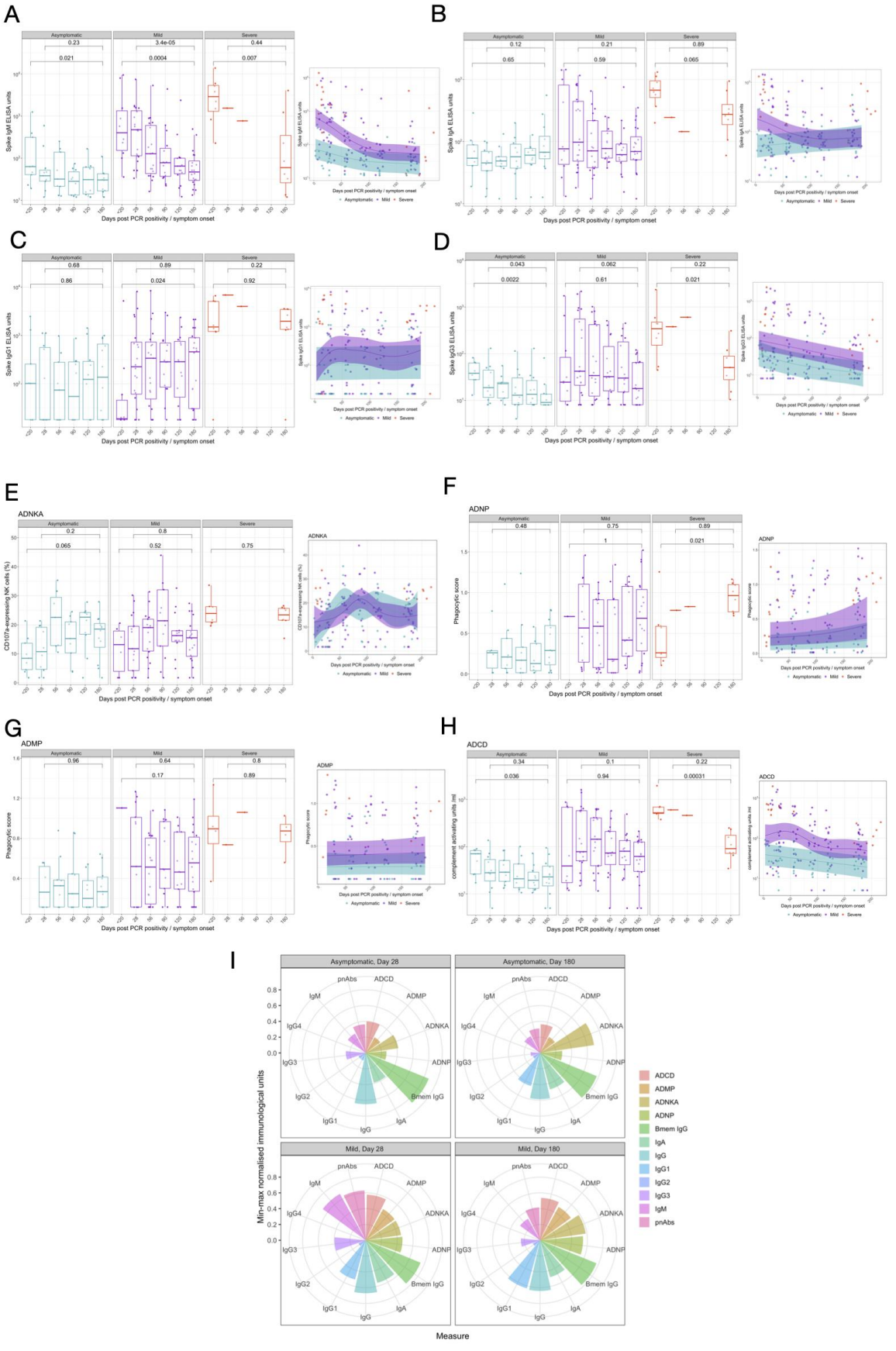
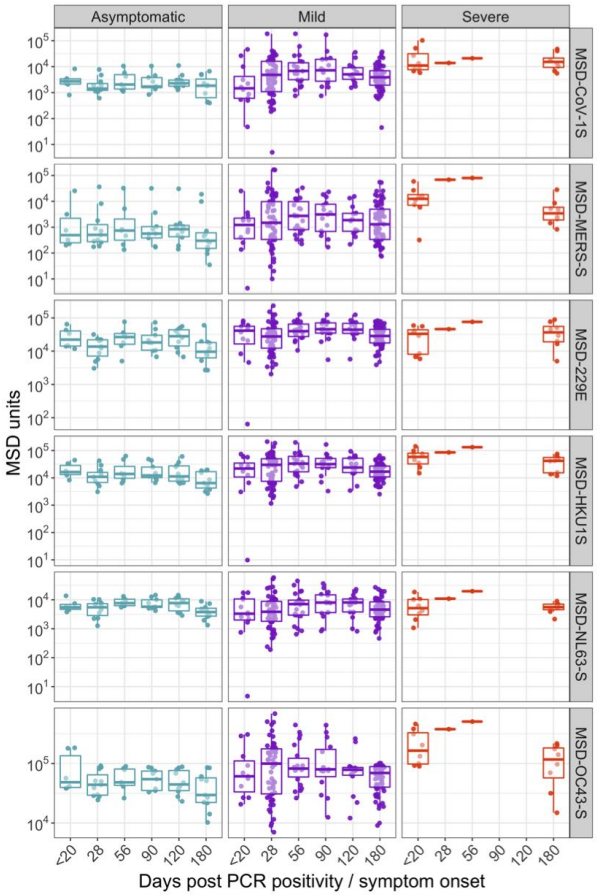

A
Asymptomatic
Mild
Severe
0.23
0.021
3.4e-05
0.0004
0.44
0.007
Spike IgM ELISA units
Days post PCR positivity / symptom onset
B
0.12
0.65
0.21
0.59
0.89
0.065
Spike IgA ELISA units
C
0.68
0.86
0.89
0.024
0.22
0.92
Spike IgG1 ELISA units
D
0.043
0.0022
0.062
0.61
0.22
0.021
Spike IgG3 ELISA units
E
ADNKA
0.2
0.065
0.8
0.52
0.75
CD107a-expressing NK cells (%)
F
ADNP
0.48
0.75
1
0.89
0.021
Phagocytic score
G
ADMP
0.96
0.17
0.64
0.89
0.8
H
ADCD
0.34
0.036
0.1
0.94
0.22
0.00031
complement activating units /ml
I
Asymptomatic, Day 28
Asymptomatic, Day 180
Mild, Day 28
Mild, Day 180
Min-max normalised immunological units
pnAbs
ADCD
IgM
ADMP
IgG4
ADNKA
IgG3
ADNP
IgG2
Bmem IgG
IgG1
IgA
IgG
Measure

Figure 2

A


Asymptomatic
Mild
Severe
MSD-CoV-1S
MSD-MERS-S
MSD-229E
MSD-HKU1S
MSD-NL63-S
MSD-OC43-S
MSD units
<20
28
56
90
120
180
Days post PCR positivity / symptom onset

B

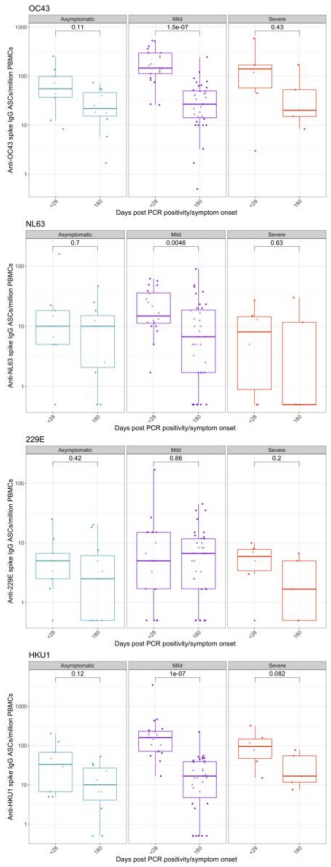

OC43
0.11
1.5e-07
0.43
NL63
0.7
0.0046
0.63
229E
0.42
0.86
0.2
HKU1
0.12
1e-07
0.082
Days post PCR positivity/symptom onset

Figure 3

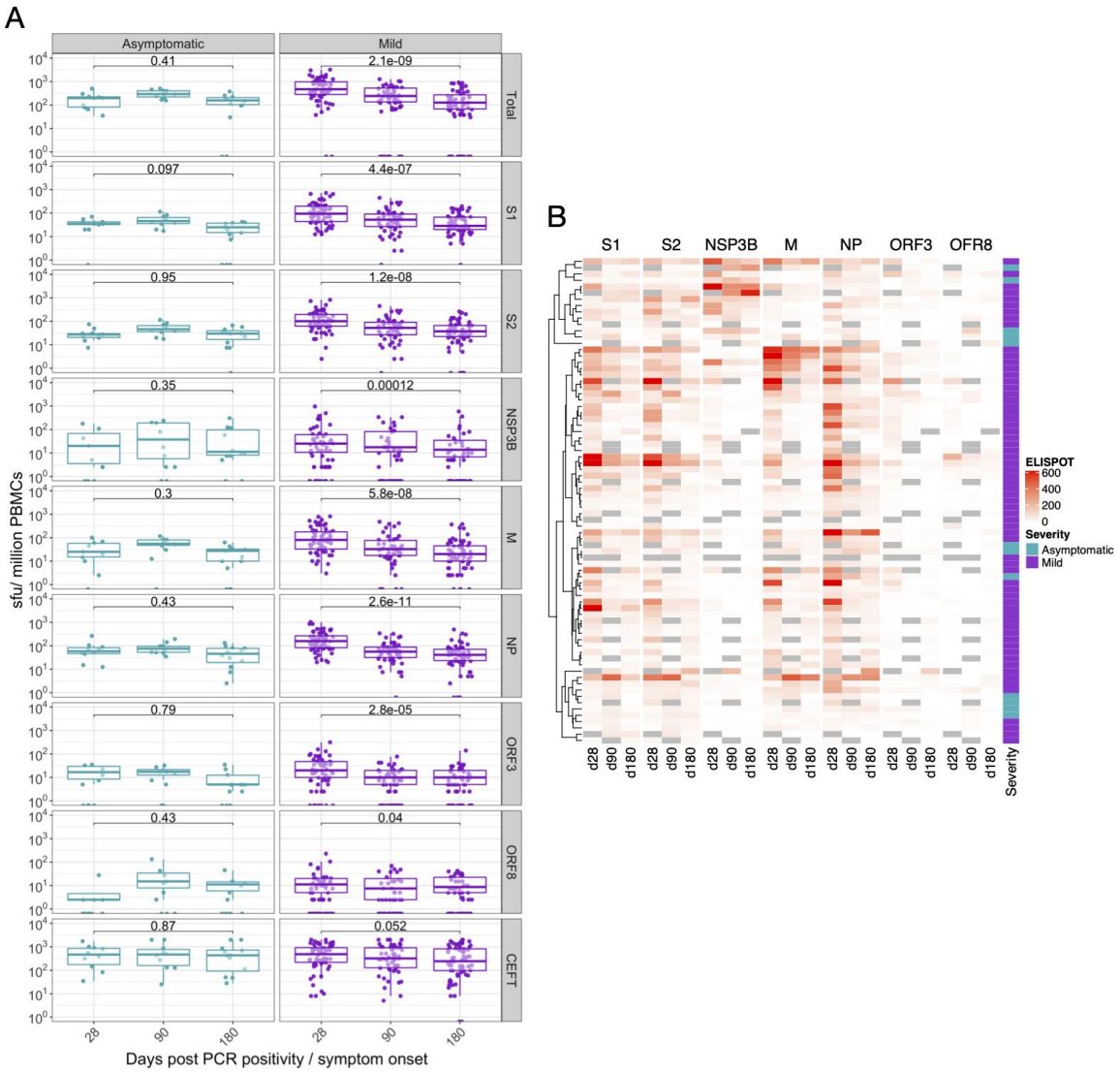

A
Asymptomatic
Mild
Total
S1
S2
NSP3B
M
NP
ORF3
ORF8
CEFT
0.41
2.1e-09
0.097
4.4e-07
0.95
1.2e-08
0.35
0.00012
0.3
5.8e-08
0.43
2.6e-11
0.79
2.8e-05
0.43
0.04
0.87
0.052
sfu/ million PBMCs
28
90
180
Days post PCR positivity / symptom onset
B
S1
S2
NSP3B
M
NP
ORF3
OFR8
d28
d90
d180
Severity
ELISPOT
600
400
200
0
Severity
Asymptomatic
Mild

Figure 4

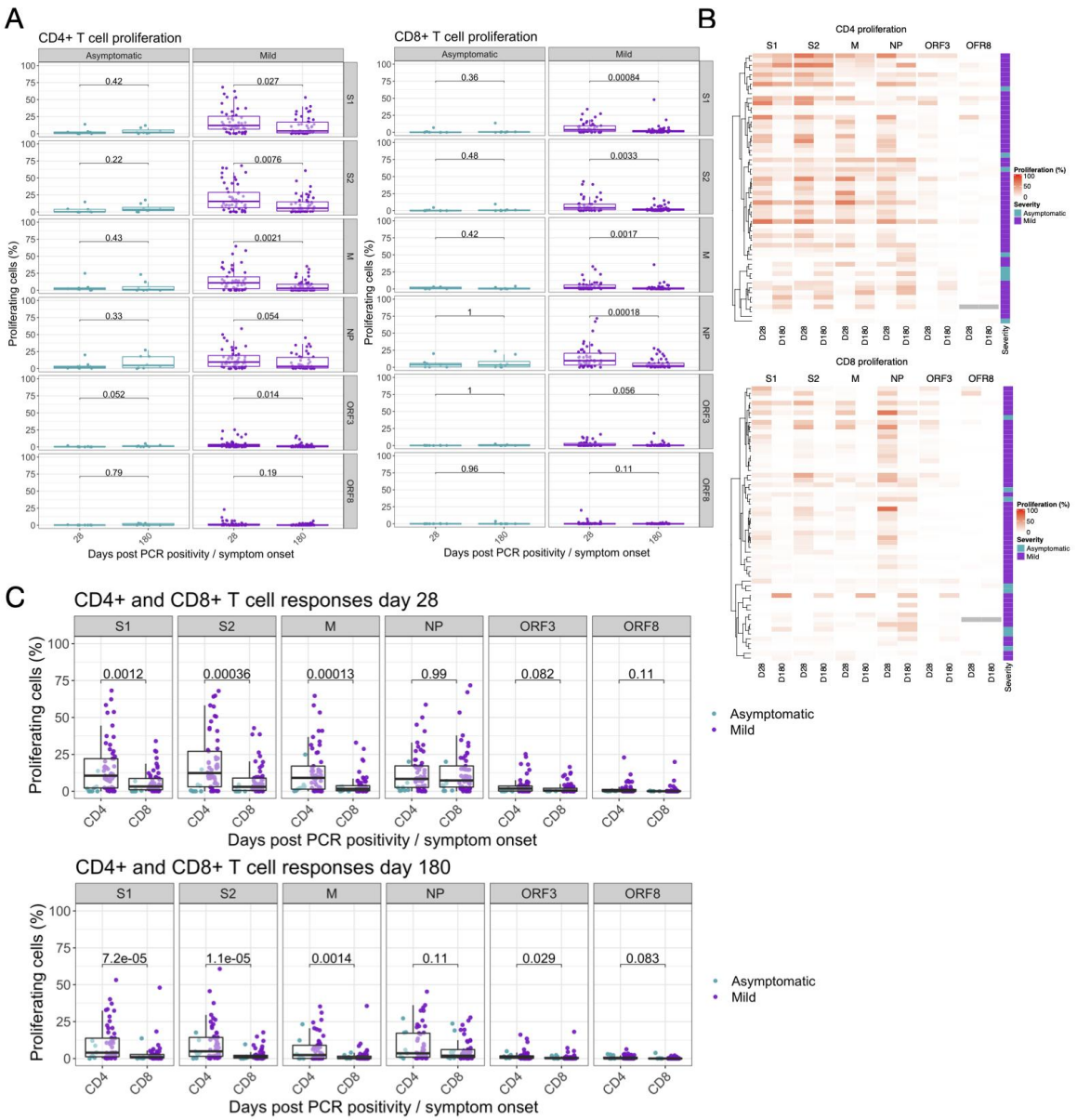

A
CD4+ T cell proliferation
CD8+ T cell proliferation
Asymptomatic
Mild
S1
S2
M
NP
ORF3
ORF8
Proliferating cells (%)
Days post PCR positivity / symptom onset
B
CD4 proliferation
CD8 proliferation
Proliferation (%)
Severity
Asymptomatic
Mild
C
CD4+ and CD8+ T cell responses day 28
CD4+ and CD8+ T cell responses day 180
CD4
CD8

Figure 5

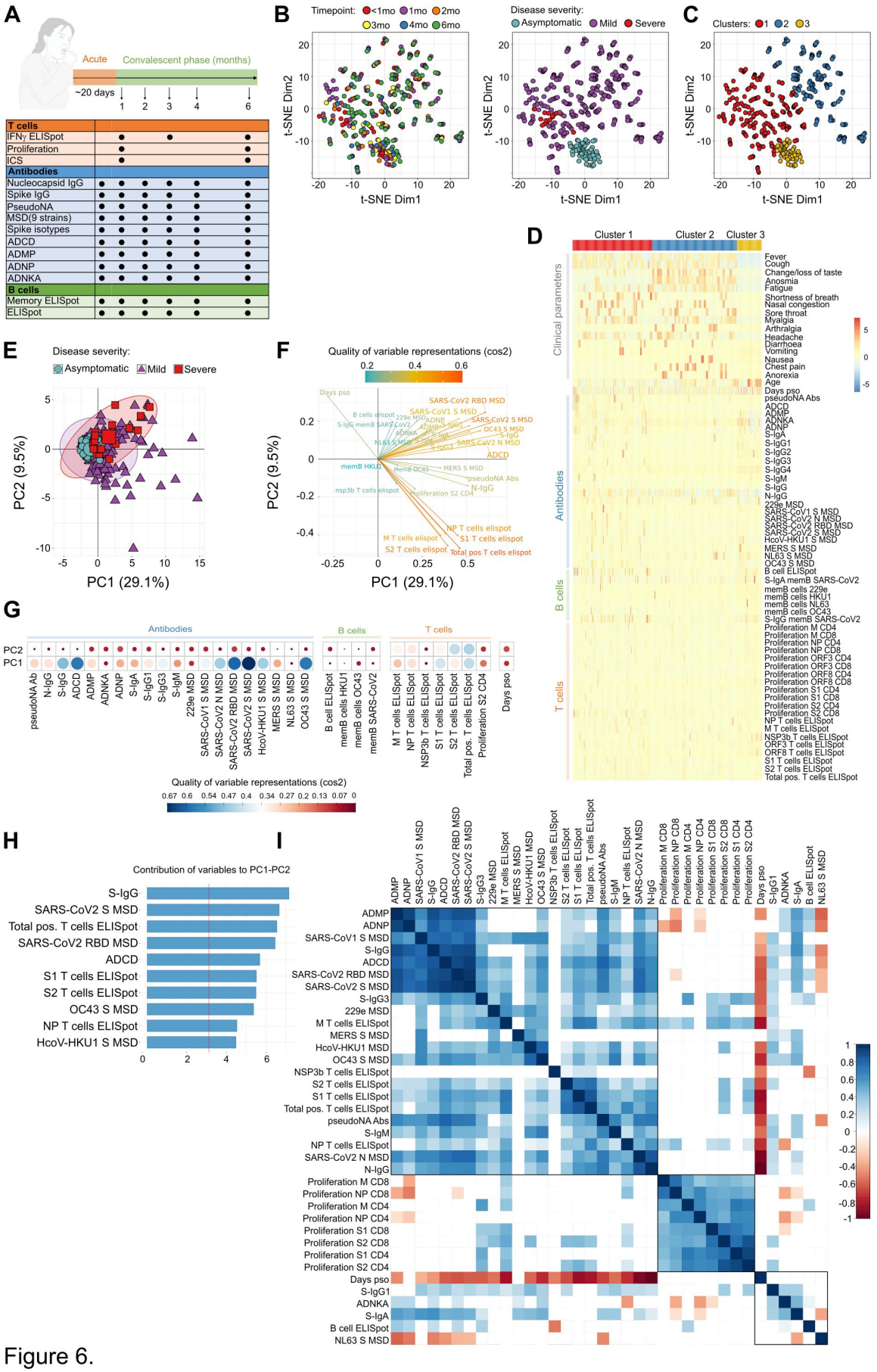

A
Acute
Convalescent phase (months)
~20 days
1
2
3
4
6
T cells
IFNγ ELISpot
Proliferation
ICS
Antibodies
Nucleocapsid IgG
Spike IgG
PseudoNA
MSD(9 strains)
Spike isotypes
ADCD
ADMP
ADNP
ADNKA
B cells
Memory ELISpot
ELISpot
B
Timepoint: <1mo 1mo 2mo 3mo 4mo 6mo
Disease severity: Asymptomatic Mild Severe
t-SNE Dim1
t-SNE Dim2
C
Clusters: 1 2 3
D
Cluster 1
Cluster 2
Cluster 3
Clinical parameters
Antibodies
B cells
T cells
E
Disease severity:
Asymptomatic Mild Severe
PC1 (29.1%)
PC2 (9.5%)
F
Quality of variable representations (cos2)
G
Antibodies
B cells
T cells
PC2
PC1
Quality of variable representations (cos2)
H
Contribution of variables to PC1-PC2
S-IgG
SARS-CoV2 S MSD
Total pos. T cells ELISpot
SARS-CoV2 RBD MSD
ADCD
S1 T cells ELISpot
S2 T cells ELISpot
OC43 S MSD
NP T cells ELISpot
HcoV-HKU1 S MSD
I

Figure 6.

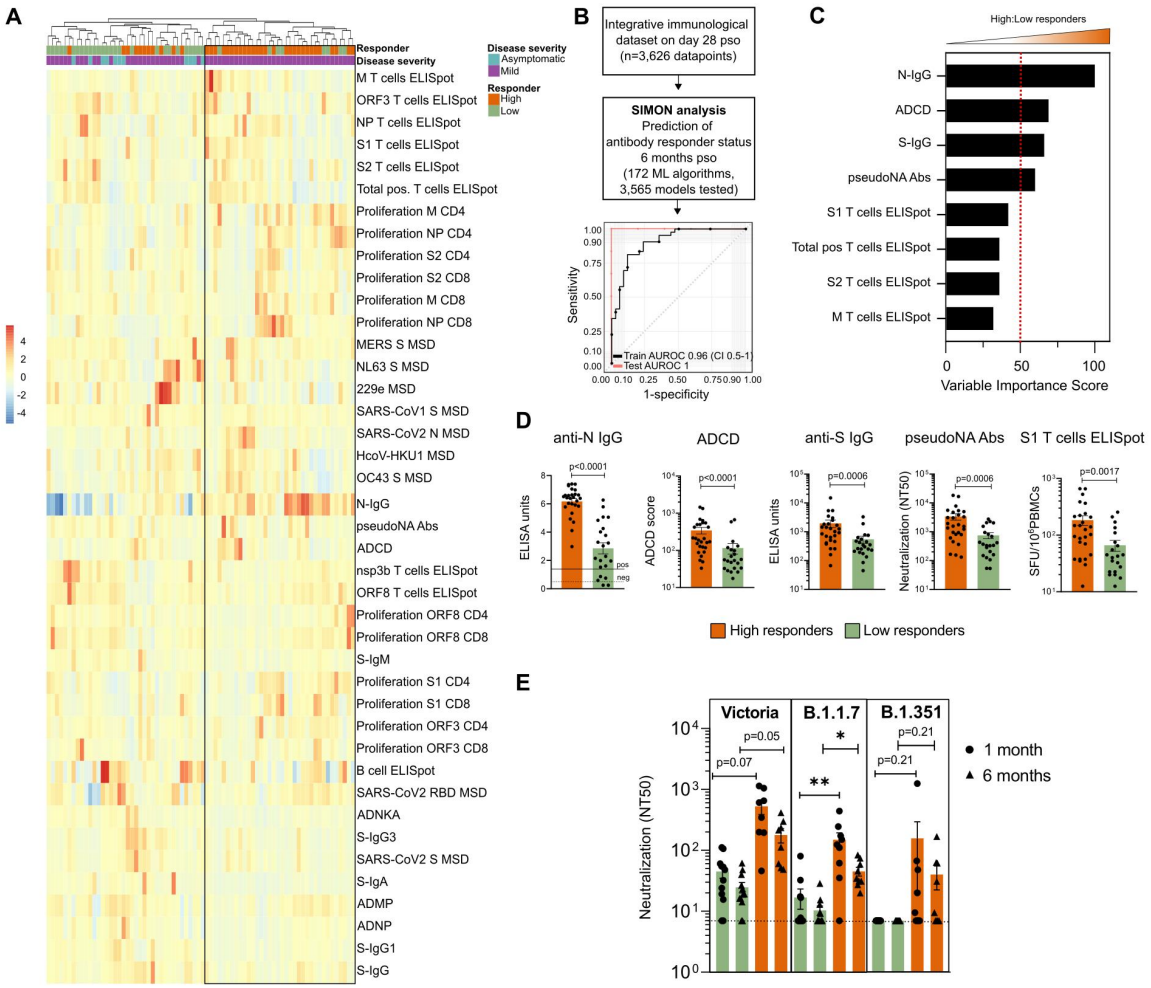

A
Responder
Disease severity
Disease severity
Asymptomatic
Mild
Responder
High
Low
M T cells ELISpot
ORF3 T cells ELISpot
NP T cells ELISpot
S1 T cells ELISpot
S2 T cells ELISpot
Total pos. T cells ELISpot
Proliferation M CD4
Proliferation NP CD4
Proliferation S2 CD4
Proliferation S2 CD8
Proliferation M CD8
Proliferation NP CD8
MERS S MSD
NL63 S MSD
229e MSD
SARS-CoV1 S MSD
SARS-CoV2 N MSD
HcoV-HKU1 MSD
OC43 S MSD
N-IgG
pseudoNA Abs
ADCD
nsp3b T cells ELISpot
ORF8 T cells ELISpot
Proliferation ORF8 CD4
Proliferation ORF8 CD8
S-IgM
Proliferation S1 CD4
Proliferation S1 CD8
Proliferation ORF3 CD4
Proliferation ORF3 CD8
B cell ELISpot
SARS-CoV2 RBD MSD
ADNKA
S-IgG3
SARS-CoV2 S MSD
S-IgA
ADMP
ADNP
S-IgG1
S-IgG
4
2
0
-2
-4
B
Integrative immunological dataset on day 28 pso (n=3,626 datapoints)
SIMON analysis
Prediction of antibody responder status 6 months pso
(172 ML algorithms, 3,565 models tested)
Sensitivity
1-specificity
Train AUROC 0.96 (CI 0.5-1
Test AUROC 1
C
High:Low responders
N-IgG
ADCD
S-IgG
pseudoNA Abs
S1 T cells ELISpot
Total pos T cells ELISpot
S2 T cells ELISpot
M T cells ELISpot
Variable Importance Score
D
anti-N IgG
ADCD
anti-S IgG
pseudoNA Abs
S1 T cells ELISpot
ELISA units
ADCD score
Neutralization (NT50)
SFU/10^6PBMCs
p<0.0001
p<0.0001
p=0.0006
p=0.0006
p=0.0017
pos
neg
High responders
Low responders
E
Victoria
B.1.1.7
B.1.351
Neutralization (NT50)
p=0.05
p=0.07
*
**
p=0.21
p=0.21
1 month
6 months

Figure 7.

Divergent trajectories of antiviral memory after SARS-Cov-2 infection

Tomic et al

Table of Contents

OPTIC Clinical Group

Lizzie Stafford (Lead)
Hibatullah Abuelgasim
Ahmed Alhussni
Carolina V. Arancibia-Cárcamo
Martyna Borak
Joseph Cutteridge
Alexandra Deeks
Lucy Denly
Stavros Dimitriadis
Shayan Fassih
Thomas Foord
Thomas Fordwoh
Jennifer Holmes
Bryn Horsington
Sven Kerneis
David Kim
Katy Lillie
Sheila Lumley
Jordan Morrow
Denise O'Donnell
Thomas G. Ritter
Beatrice Simmons
Adan Taylor
Sarah R. Thomas
Yolanda Warren
Adam J. R. Watson
Esme Weeks
Robert Wilson
Rebecca Young

PITCH Study Group

Susanna J. Dunachie (Lead)
Christopher J.A. Duncan
Shona C. Moore
Rebecca Payne
Alex Richter
Sarah Rowland-Jones

C-MORE Group

Alexander J. Mentzer (Lead)
Mark Philip Cassar
Tao Dong
Anastasia Fries
Javier Gilbert-Jaramillo
Ling-Pei Ho
Julian C. Knight
Stefan Neubauer
Yanchun Peng
Nayia Petousi
Betty Raman
Nick P. Talbot

Table S1 Demographics

		Asymptomatic		Mild		Severe		Total	
Number enrolled		**12**	**% of asymp**	**66**	**% of mild**	**7**	**% of severe**	**85**	**% of total**
Sex	**Female**	11	92%	48	73%	2	29%	61	72%
	Male	1	8%	18	27%	5	71%	24	28%
Age range (at recruitment)	**20-29**	2	17%	21	32%	0	0%	23	27%
	30-39	4	33%	17	26%	0	0%	21	25%
	40-49	2	17%	14	21%	2	29%	18	21%
	50-59	3	25%	14	21%	4	57%	21	25%
	60-69	1	8%	0	0%	1	14%	2	2%
Ethnicity	**Asian**	1	8%	10	15%	1	14%	12	14%
	Black	1	8%	0	0%	1	14%	2	2%
	White	10	83%	48	73%	3	43%	61	72%
	Other	0	0%	4	6%	1	14%	5	6%
	Unknown	0	0%	4	6%	1	14%	5	6%

Table S2. Summary Table of Descriptive Statistics for Humoral Assays

Day	anti-SARS-CoV-2 nucleocapsid tIgG ELISA (ELISA AU)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	7	3.40 (2.17-4.31)	39	3.64 (1.09-5.84)	0	ND
28	11	2.15 (1.61-3.81)	60	5.58 (3.83-6.5)	0	ND
56	9	2.45 (1.24-3.44)	63	5.16 (3.09-6.21)	0	ND
90	9	1.46 (0.80-2.24)	51	4.39 (2.41-5.68)	0	ND
120	9	1.22 (0.55-1.51)	51	3.39 (1.50-4.91)	0	ND
180	12	0.84 (0.33-1.12)	64	2.03 (0.72-3.66)	0	ND
Day	anti-SARS-CoV-2 spike tIgG ELISA (ELISA AU)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	9	341.8 (241.0-897.1)	42	122.4 (61.5-754.8)	8	1182.0 (387.7-1947.0)
28	11	427.2 (260.1-753.0)	64	504.1 (196.6-1361.0)	1	2404.0 (2404.0-2404.0)
56	9	424.4 (287.4-788.3)	62	677.9 (236.1-1532.0)	1	1837.0 (1837.0-1837.0)
90	10	229.8 (114.4-404.9)	51	525.4 (178.5-1006.0)	0	ND
120	10	221.0 (115.9-290.2)	53	365.3 (192.9-765.3)	0	ND
180	12	257.7 (134.7-400.4)	65	399.4 (180.5-710.1)	8	1401.0 (903.6-1616.0)
Day	Pseudoneutralising antibodies (IC50)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	9	226.20 (118.90-701.60)	40	1261.00 (447.90-3813.00)	8	4923.00 (2135.00-30032.00)
28	11	126.10 (53.86-206.00)	64	817.50 (345.00-2954.00)	0	ND
56	8	127.70 (83.88-242.70)	63	441.8 (212.90-913.00)	1	1346.00 (1346.00-1346.00)
90	10	142.60 (80.18-294.00)	51	393.20 (170.20-747.60)	0	ND
120	10	98.96 (71.32-174.50)	51	243.40 (130.80-430.40)	0	ND
180	12	89.76 (52.01-148.70)	65	182.50 (99.29-479.40)	8	765.1 (457.00-1619.00)
Day	anti SARS-CoV-2-spike IgG memory B cell ELISPOT (ASCs per million PBMCs)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	7	176.70 (108.30-207.50)	10	25.80 (5.85-49.15)	8	49.00 (26.50-131.80)
28	9	231.70 (180.00-477.40)	44	112.50 (30.43-246.30)	0	ND
56	8	145.90 (34.18-260.00)	35	186.70 (126.70-293.30)	0	ND
90	8	196.70 (153.30-211.40)	31	166.70 (95.00-270.00)	0	ND
120	6	227.70 (145.50-311.10)	30	119.00 (77.50-289.00)	0	ND
180	12	134.2 (95.00-173.70)	47	130.00 (58.30-230.00)	8	649.00 (236.00-1111.00)
Day	anti-SARS-CoV-2 spike IgA memory B cell ELISPOT (ASCs per million PBMCs)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	7	2.50 (0.00-5.00)	10	0.00 (0.00-10.40)	8	5.00 (2.25-56.50)
28	9	0.00 (0.00-7.50)	44	0.00 (0.00-5.00)	0	ND
56	8	0.00 (0.00-5.03)	35	0.00 (0.00-5.00)	0	ND
90	8	0.00 (0.00-0.00)	31	0.00 (0.00-0.00)	0	ND
120	6	0.00 (0.00-8.75)	30	0.00 (0.00-1.25)	0	ND
180	12	0.00 (0.00-1.46)	47	0.00 (0.00-5.00)	8	0.00 (0.00-12.93)

Day	IgG1 (ELISA AU)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	9	101.50 (18.00-777.50)	10	18.99 (18.00-97.31)	8	1500.00 (1193.00-5106.00)
28	11	18.00 (18.00-612.80)	20	226.9 (54.58-834.30)	1	6835.00 (6835.00-6835.00)
56	9	74.19 (18.00-357.80)	19	338.5 (34.73-826.90)	1	3984.00 (3984.00-3984.00)
90	10	92.07 (18.00-295.90)	15	284.7 (78.74-765.10)	0	ND
120	10	126.40 (18.00-499.60)	15	286 (34.51-818.90)	0	ND
180	12	154.60 (18.00-803.90)	20	459.5 (30.58-941.40)	8	2047.00 (1309.00-3482.00)
Day	IgG2 (OD units)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	9	0.123 (0.118-0.245)	10	0.1145 (0.097-0.1728)	8	0.123 (0.098-0.1485)
28	11	0.099 (0.094-0.125)	20	0.1088 (0.099-0.1449)	1	0.148
56	9	0.097 (0.0905-0.1008)	19	0.117 (0.093-0.15)	1	0.155
90	10	0.097 (0.0935-0.1083)	15	0.1 (0.092-0.14)	0	ND
120	10	0.0985 (0.08725-0.1118)	15	0.112 (0.102-0.148)	0	ND
180	12	0.1015 (0.09175-0.1075)	20	0.109 (0.099-0.1295)	8	0.1255 (0.1058-0.1738)
Day	IgG3 (ELISA AU)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	9	38.08 (27.73-64.09)	10	24.47 (8.00-250.8)	8	254.1 (93.7-553.4)
28	11	18.7 (10.16-49.76)	20	41.99 (16.8-760.4)	1	379.5
56	9	22.76 (10.56-28.34)	19	33.91 (10.59-443.2)	1	605.4
90	10	12.89 (8.00-37.95)	15	32.12 (14.97-256.7)	0	ND
120	10	13.74 (8.00-31.57)	15	29.88 (14.67-217.1)	0	ND
180	12	9.068 (8.00-13.79)	20	17.81 (8.000-65.97)	8	50.39 (19.9-109)
Day	IgG4 (OD units)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	9	0.0605 (0.058-0.064)	10	0.06275 (0.058-0.08113)	8	0.082 (0.08075-0.9225)
28	11	0.08 (0.063-0.0835)	20	0.0815 (0.06088-0.08713)	1	0.0795
56	9	0.0845 (0.08225-0.08575)	19	0.08550 (0.0825-0.087)	1	0.0825
90	10	0.08325 (0.08013-0.08888)	15	0.08550 (0.082-0.0895)	0	ND
120	10	0.083 (0.08038-0.09)	15	0.084 (0.0815-0.0885)	0	ND
180	12	0.08125 (0.078-0.086)	20	0.084 (0.07963-0.08888)	8	0.8725 (0.082-0.093)
Day	ADNKA (% CD107a expressing NK cells)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	9	8.475 (4.188-14.49)	6	13.15 (1.331-19.08)	6	23.91 (20.46-28.12)
28	10	10.76 (3.761-19.68)	17	11.76 (2.77-21.4)	0	ND
56	7	22.54 (12.69-31.59)	18	19.03 (10.2-22.1)	0	ND
90	9	15.25 (7.17-22.02)	13	21.39 (10.62-32.07)	0	ND
120	10	22.62 (14.5-24.4)	15	16.29 (13.79-18.03)	0	ND
180	12	18.54 (8.1-21.44)	20	15.51 (8.6-19.09)	8	23.41 (21.56-25.88)

95

Day	ADNP (phagocytic score)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	0	ND	1	0.7071	7	0.2603 (0.1618-0.6023)
28	6	0.2612 (-0.02452-0.41)	12	0.5654 (0.1024-1.154)	1	0.7829
56	9	0.2087 (0.03429-0.487)	17	0.5863 (0.07072-0.9166)	1	0.8277
90	9	0.1695 (-0.01632-0.4668)	13	0.1808 (-0.006721-0.9186)	0	ND
120	8	0.1283 (0.02477-0.3941)	13	0.416 (0.2192-1.087)	0	ND
180	12	0.2889 (0.06631-0.6443)	18	0.685 (0.2207-1.073)	8	0.9643 (0.7522-1.110)
Day	ADMP (phagocytic score)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	0	ND	1	1.104	4	0.8948 (0.4953-1.231)
28	6	0.2983 (0.04294-0.5566)	13	0.5197 (0.1747-1.087)	1	0.7361
56	9	0.3268 (0.1715-0.5015)	18	0.5135 (0.1884-0.8254)	1	1.06
90	9	0.2479 (0.07806-0.6466)	13	0.4937 (0.2748-0.9776)	0	ND
120	8	0.2539 (0.1134-0.4688)	13	0.4637 (0.2604-0.8632)	0	ND
180	11	0.2683 (0.06291-0.4265)	18	0.5556 (0.2570-0.8358)	4	0.8732 (0.6285-0.9976)
Day	ADCD (complement arbitrary units)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	9	70.71 (26.23-83.65)	8	44.08 (14.62-706.5)	7	521.1 (496.7-702.8)
28	10	29.58 (23.18-63.61)	19	78.25 949.60-581.1)	1	601.5
56	9	28.83 (16.79-66.81)	19	143.3 (38.75-408.9)	1	459.7
90	9	20.88 (12.92-53.86)	14	75.03 (53.85-225.2)	0	ND
120	10	19.63 (12.38-33.38)	14	84.5 (47.04-148.6)	0	ND
180	10	25.76 (17.16-50.46)	18	65.62 (39.92-131.7)	8	92.56 (68.07-213.7)
Day	MSD-CoV-2-N (MSD units)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	6	23874 (10566-63490)	12	22008 (2160-288100)	8	351347 (56110-909112)
28	11	9125 (5550-26640)	59	48980 (11134-150139)	1	759067
56	7	14656 (6139-28209)	22	79304 (25109-122795)	1	1060397
90	9	18875 (6235-24250)	15	66980 (21635-100886)	0	ND
120	10	9668 (4468-18295)	15	41029 (17680-63060)	0	ND
180	11	1450 (800-4315)	65	15270 (4603-34135)	8	66976 (28763-130158)
Day	MSD-CoV-2-S (MSD units)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	6	17508 (13239-35111)	12	2603 (1121-119940)	8	101430 (50745-151061)
28	11	9775 (4675-15650)	59	15934 (6625-58090)	1	172859
56	7	11971 (8063-25715)	22	49465 (12132-103915	1	194617
90	9	10325 (8005-19413)	15	35620 (9060-98421)	0	ND
120	10	11226 (8216-17903)	15	27767 (13731-63222)	0	ND
180	11	5865 (2385-9245)	65	14610 (7848-27165)	8	76010 (50080-83777)

96

Day	MSD-CoV-2-RBD (MSD units)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	6	5162 (3514-13439)	12	1137 (428.5-42463)	8	28094 (10097-40119)
28	11	2640 (1825-5985)	59	6389 (2260-16863)	1	62130
56	7	4724 (2667-8375)	22	16408 (2052-38674)	1	73667
90	9	4110 (2913-5609)	15	21405 (2225-27426)	0	ND
120	10	4292 (2454-5672)	15	15729 (3969-24675)	0	ND
180	11	1280 (575-3015)	65	5470 (2618-11740)8	8	27437 (19479-35138)
Day	MSD-CoV-1-S (MSD units)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	6	2751 (1765-4642)	12	1469 (558.3-8282)	8	11247 (7049-45153)
28	11	1420 (1140-2640)	59	4827 (930-16094)	1	13838
56	7	2049 (1277-5780)	22	6841 (3024-15560)	1	20930
90	9	1710 (1458-4164)	15	7225(2365-19446)	0	ND
120	10	2338 (1670-3455)	15	5045 (3104-9766)	0	ND
180	11	1840 (435-4870)	65	3895 (1908-7018)	8	15296 (7806-34528)
Day	MSD-MERS-S (MSD units)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	6	540.2 (222.1-8731)	12	1325 (291.2-2613)	8	12607 (6286-24051)
28	11	505(260-1790)	59	1485 (326.2-9838)	1	68698
56	7	741.8 (212.1-5195)	22	2738 (735.7-11182)	1	79785
90	9	565 (282.5-2355)	15	3115 (655-8415)	0	ND
120	10	856.2 (353.2-1256)	15	1875 (650-3528)	0	ND
180	11	856.3 (353.2-1256)	65	1300 (335-5815)	8	3635 (1612-7713)
Day	MSD-229E-S (MSD units)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	6	24102 (13468-48192)	12	41196 (11351-59135)	8	33216 (7191-51864)
28	11	13690 (5670-22120)	59	27464 (11980-46933)	1	45764
56	7	26430 (14338-35753)	22	39582 (26185-71234)	1	76423
90	9	18135 (10473-37915)	15	45493 (34515-74800)	0	ND
120	10	28051 (11973-46341)	15	44060 (33901-73586)	0	ND
180	11	9600 (5190-18660)	65	28410 (17525-47678)	8	37402 (17322-70656)
Day	MSD-HKU1-S (MSD units)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	6	16511 (11681-36256)	12	30221 (6930-50917)	8	60514 (27302-92141)
28	11	10980 (6140-19060)	59	33253 (16048-65480)	1	85830
56	7	13990 (8453-47818)	22	32100 (23770-51565)	1	133025
90	9	12306 (9380-34408)	15	24057 (15220-52627)	0	ND
120	10	11651 (7043-30534)	15	24057 (15220-52627)	0	ND
180	11	6490 (3770-18410)	65	16960 (10873-28140)	8	42123 (14326-62614)

97

Day	MSD-NL63-S (MSD units)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	6	5522 (4311-9005)	12	3331 (1781-11318)	8	5252 (2413-10385)
28	11	5505 (2305-7765)	59	3855 (1760-9555)	1	10914
56	7	7673 (6180-11166)	22	7109 (2895-11623)	1	19765
90	9	5759 (5283-11355)	15	8040 (3510-15480)	0	ND
120	10	7819 (3625-11602)	15	8096 (3790-11090)	0	ND
180	11	3780 (2600-5445)	65	4590 (2595-9140)	8	5614 (4137-7214)
Day	MSD-OC43-S (MSD units)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	6	49070 (38844-181033)	12	61023 (26744-131983)	8	168319 (94347-364668)
28	11	44025 (28820-81555)	59	99657 (30725-176897)	1	377370
56	7	48304 (41330-82764)	22	82329 (51781-152439)	1	506545
90	9	54585 (35190-79453)	15	79865 (48805-189773)	0	ND
120	10	44648 (32552-80433)	15	78449 (61302-87133)	0	ND
180	11	29570 (21835-60550)	65	69440 (39608-87133)	8	118990
Day	anti-229E spike IgG memory B cell ELISPOT (ASCs per million PBMCs)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	7	5 (1.7-6.67)	10	5 (0-11.25)	6	6 (2.25-8.5)
28	2	14.17 (3.33-25)	10	10 (1.693-15)	0	ND
56	0	ND	0	ND	0	ND
90	0	ND	0	ND	0	ND
120	0	ND	0	ND	0	ND
180	11	2.5 (0-7.5)	36	6.67 (1.678-11.7)	5	1.67 (0.0-5.835)
Day	anti- HKU1 spike IgG memory B cell ELISPOT (ASCs per million PBMCs)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	7	28.3 (5-67.5)	10	187.7 (93.73-300)	6	97 (33.75-199.3)
28	2	85.85 (46.70-125)	9	105 (55-240)	0	ND
56	0	ND	0	ND	0	ND
90	0	ND	0	ND	0	ND
120	0	ND	0	ND	0	ND
180	11	11 (2.5-31.7)	35	16.7 (3.33-40)	5	16.67 (9.585-65.84)
Day	anti-spike NL63 IgG memory B cell ELISPOT (ASCs per million PBMCs)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	7	6.7 (5-18.3)	10	19.2 (11.28-40.43)	6	9 (0-18)
28	2	84.15 (10-158.3)	10	12.9 (9.575-41.18)	0	ND
56	0	ND	12	12.9 (2.1-22.48)	0	ND
90	0	ND	0	ND	0	ND
120	0	ND	0	ND	0	ND
180	11	10 (1.7-18.3)	24	5 (0.425-15.03)	5	0.00 (0.00-20.84)

98

Day	anti-spike OC43 IgG memory B cell ELISPOT (ASCs per million PBMCs)					
	Asymptomatic		Mild		Severe	
	n	median (IQR)	n	median (IQR)	n	median (IQR)
<20	7	43.4 (12.6-98.3)	11	141.7 (83.3-300)	6	141.5(34.5-272.8)
28	2	163.4 971.7-255)	10	160.9 (106.7-313.1)	0	ND
56	0	ND	0	ND	0	ND
90	0	ND	0	ND	0	ND
120	0	ND	0	ND	0	ND
180	11	21.7 (13.3-53.3)	35	26.7 (13.33-51.7)	5	20 (11.67-110.4)

99

Table S3 Summary statistics for ex vivo interferon-gamma ELISpot assay

	d28 Median (IQR) N = 70	d28 No. +ve (%)	d90 Median (IQR) N = 64	d90 No. +ve (%)	d180 Median (IQR) N = 78	d180 No. +ve (%)	P value d28 VS d90	Sign.	P value d28 VS d180	Sign.	P value d90 VS d180	Sign.
Total positive pools	373 (201-842)	67/70 (96%)	240 (124-430)	56/63 (89%)	105 (40-218)	61/77 (79%)	0.0006	***	<0.0001	****	0.0002	***
Total spike	180 (71-364)	71/70 (87%)	100 (0-175)	41/63 (65%)	43 (0-103)	46/77 (60%)	0.0016	**	<0.0001	****	0.0262	*
S1	70 (37-171)	56/70 (80%)	50 (23-89)	39/63 (62%)	25 (15-60)	32/77. (40%)	0.0001	***	<0.0001	****	0.0508	ns
S2	83 (27-182)	53/70 (76%)	46 (28-84)	39/63 (62%)	30 (18-58)	33/77. (43%)	0.0027	**	<0.0001	****	0.1646	ns
M	63 (25-160)	47/70 (67%)	33 (15-74)	26/63 (41%)	17 (5-43)	24/77 (31%)	<0.0001	****	<0.0001	****	0.03	*
NP	121 (73-250)	62/70 (89%)	56 (33-95)	44/63 (70%)	40 (20-74)	47/77 (61%)	<0.0001	****	<0.0001	****	0.0343	*
ORF3	15 (5-38)	22/70 (31%)	8 (0-18)	6/63 (8%)	5 (0-13)	2/76 (3%)	0.0028	**	<0.0001	****	0.8959	ns
ORF8	3 (0-15)	6/70 (9%)	3 (0-13)	8/63 (13%)	0 (0-9)	3/76 (4%)	0.7126	ns	0.0368	*	0.5576	ns
NSP3	18 (3-48)	23/70 (33%)	8 (0-32)	15/63 (24%)	5 (0-15)	12/76 (16%)	0.0245	*	<0.0001	****	0.2946	ns
CEFT	488 (209-920)	65/70 (93%)	326 (128-908)	56/63 (89%)	288 (91-823)	66/77 (86%)	0.2034	ns	<0.0001	****	0.0508	ns

Ex vivo IFN-γ ELISpot showing the effector T cell responses to summed SARS-CoV-2 peptide pools spanning spike, accessory and structural proteins (S1, S2, M, NP, ORF 3, ORF8, and NSP3B, summed total of SARS-CoV-2 proteins tested and the CEFT positive control peptides for T cell responses) in 78 individuals 28, 90 and 180 days after SARS-CoV-2 (onset of symptoms for mild cases, PCR positive test for asymptomatic participants). SFC / million PBMCs = spot forming cells per million peripheral blood mononuclear cells, with background (negative control wells) subtracted, and a positive response = final count with background subtracted greater than mean + 2 SD of background. D28, d90 and d180 = days after SARS-CoV-2 diagnosis. Significance of differences between timepoints T1, T2 and T3 calculated by paired Friedman test with Dunn's multiple comparisons test for the n = 57 participants with assays at all three timepoints available.

Supplementary table 4 – Summary of tables for proliferation assay based on disease phenotype

		CD4+ T cells															
		CD4_ D28	CD4_ D180	CD4_ D28	CD4_ D180	CD4_ D28	CD4_ D180	CD4_ D28	CD4_ D180	CD4_ D28	CD4_ D180	CD4_ D28	CD4_ D180	CD4_ D28	CD4_ D180	CD4_ D28	CD4_ D180
	Disease Phenotype	**PHA**	**PHA**	**S1**	**S1**	**S2**	**S2**	**M**	**M**	**NP**	**NP**	**ORF3**	**ORF3**	**ORF8**	**ORF8**	**FECT**	**FECT**
Summary (All)	Numbers tested	56.0	56.0	57.0	57.0	57.0	57.0	57.0	57.0	57.0	57.0	57.0	57.0	56.0	56.0	54.0	54.0
	Number positive (>1%)	56.0	56.0	47.0	46.0	45.0	46.0	44.0	32.0	46.0	42.0	30.0	20.0	16.0	14.0	16.0	15.0
	% Positive (>1%)	100.0	100.0	82.5	80.7	78.9	80.7	77.2	56.1	80.7	73.7	52.6	35.1	28.6	25.0	29.6	27.8
Summary (asympto matic)	Numbers tested with peptide pool	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0
	Number positive (>1%)	8.0	8.0	4.0	6.0	3.0	6.0	6.0	4.0	4.0	6.0	1.0	4.0	0.0	3.0	1.0	2.0
	% Positive (>1%)	100.0	100.0	50.0	75.0	37.5	75.0	75.0	50.0	50.0	75.0	12.5	50.0	0.0	37.5	12.5	25.0
Summary (Mild)	Numbers tested with peptide pool	48.0	48.0	49.0	49.0	49.0	49.0	49.0	49.0	49.0	49.0	49.0	49.0	49.0	49.0	49.0	49.0
	Number positive (>1%)	48.0	48.0	43.0	40.0	42.0	40.0	38.0	28.0	42.0	36.0	29.0	16.0	16.0	11.0	15.0	13.0
	% Positive (>1%)	100.0	100.0	87.8	81.6	85.7	81.6	77.6	57.1	85.7	73.5	59.2	32.7	32.7	22.4	30.6	26.5

		CD8+ T cells															
		CD8_ D28	CD8_ D180	CD8_ D28	CD8_ D180	CD8_ D28	CD8_ D180	CD8_ D28	CD8_ D180	CD8_ D28	CD8_ D180	CD8_ D28	CD8_ D180	CD8_ D28	CD8_ D180	CD8_ D28	CD8_ D180
	Disease Phenotype	**PHA**	**PHA**	**S1**	**S1**	**S2**	**S2**	**M**	**M**	**NP**	**NP**	**ORF3**	**ORF3**	**ORF8**	**ORF8**	**FECT**	**FECT**
Summary (All)	Numbers tested	56.0	56.0	57.0	57.0	57.0	57.0	57.0	57.0	57.0	57.0	57.0	57.0	56.0	56.0	54.0	54.0
	Number positive (>1%)	56.0	56.0	44.0	29.0	39.0	32.0	37.0	19.0	49.0	36.0	21.0	10.0	5.0	5.0	29.0	24.0
	% Positive (>1%)	100.0	100.0	77.2	50.9	68.4	56.1	64.9	33.3	86.0	63.2	36.8	17.5	8.9	8.9	53.7	44.4
Summary (asympto matic)	Numbers tested with peptide pool	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0
	Number positive (>1%)	8.0	8.0	2.0	1.0	1.0	2.0	4.0	3.0	5.0	5.0	0.0	3.0	0.0	1.0	2.0	4.0
	% Positive (>1%)	100.0	100.0	25.0	12.5	12.5	25.0	50.0	37.5	62.5	62.5	0.0	37.5	0.0	12.5	25.0	50.0
Summary (Mild)	Numbers tested with peptide pool	48.0	48.0	49.0	49.0	49.0	49.0	49.0	49.0	49.0	49.0	49.0	49.0	49.0	49.0	49.0	49.0
	Number positive (>1%)	48.0	48.0	42.0	28.0	38.0	30.0	33.0	16.0	44.0	31.0	21.0	7.0	5.0	4.0	27.0	20.0
	% Positive (>1%)	100.0	100.0	85.7	57.1	77.6	61.2	67.3	32.7	89.8	63.3	42.9	14.3	10.2	8.2	55.1	40.8

Proliferation assay summary table highlighting the total number of patients in each group as well as percentage of patients that responding (% positive >1%) to any of the different SARS-CoV-2 proteins tested. Data is presents within the separate T cell compartments analysed (CD4+ and CD8+ T cells).

Supplementary table 5 – Summary of tables for proliferation assay based on peptide pool tested

	Median with interquartile range (IQR)							
	CD4				CD8			
	Asymptomatic		Mild		Asymptomatic		Mild	
	Day 28	Day 180	Day 28	Day 180	Day 28	Day 180	Day 28	Day 180
M (%)	2.7 (1.9 - 9.5)	7.6 (1.9 - 20.5)	13.7 (8.4 - 29.6)	7.9 (3.9 - 18.2)	2.7 (1.5 - 3.0)	2.3 (1 - 4)	3.4 (1.6 - 7.9)	2.6 (1.6 - 5.5)
NP (%)	4.2 (2.6 - 16.5)	11.6 (2.9 - 20.1)	12.5 (6.8 - 22.7)	6.7 (2.6 - 21.7)	5.3 (3.7 - 13.5)	5.1 (3.3 - 21.3)	10.2 (5.9 - 22.6)	4.5 (1.8 - 12.1)
ORF3 (%)	1.9 (1.9 - 1.9)	1.9 (1.4 - 4.2)	3.5 (2.5 - 7)	2.9 (1.8 - 5.6)	0	1.9 (1.2 - 2.3)	3.8 (1.7 - 9.2)	4.3 (2 - 7.14)
ORF8 (%)	0	2.6 (2.2 - 2.9)	3.1 (1.6 - 6.4)	2.3 (1.9 - 3.1)	0	3.8 (3.8 - 3.8)	3.7 (2.7 - 13.6)	1.5 (1.1 - 2.0)
Total responses to at least peptide outside spike	87.50%	75%	86%	75.00%	62.50%	75%	91%	71%

Proliferation assay summary table highlighting median responses with IQR to each of the SARS-CoV-2 peptide pool tested in the asymptomatic and mild disease phenotypes. Data is presents within the separate T cell compartments analysed (CD4+ and CD8+ T cells)

Fig S1. Comparison of humoral immune responses in individuals with PCR confirmed SARS-CoV-2 asymptomatic (blue), mild (purple) or severe (red) infection.

Comparison of total IgG specific to the SARS-CoV-2 (A) nucleocapsid and (B) spike glycoprotein, quantification of (C) IgG memory B cells specific to the spike glycoprotein, and (D) pseudoneutralisating antibody titres. Anti-SARS-CoV-2 spike antibody (E-F) isotypes, (G-H) subclasses and (I-L) Fc-mediated effector functions were also compared. See Table S1 for number of individuals evaluated per assay.

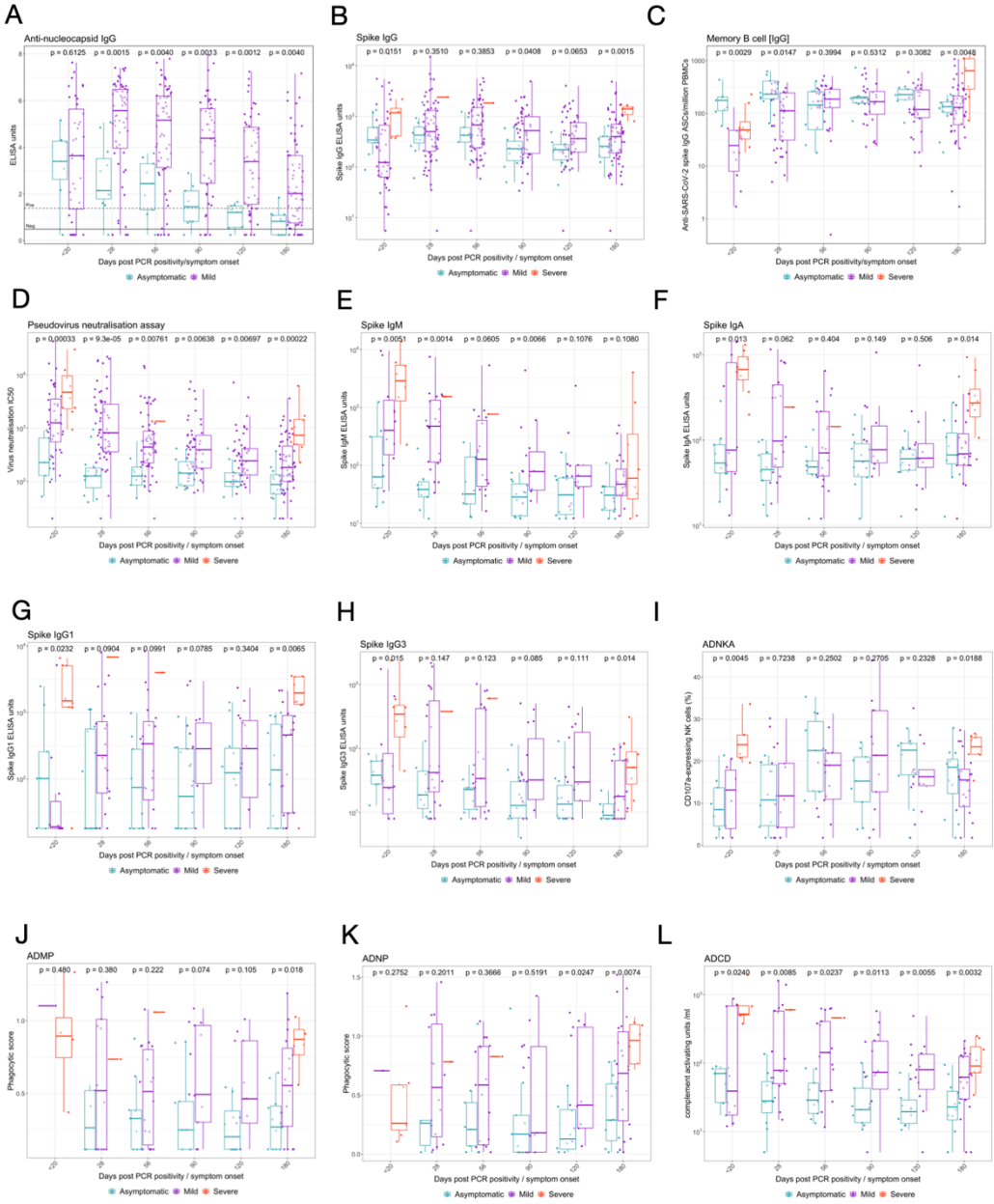


Fig S2. Further characterisation of longitudinal humoral immune responses to SARS-CoV-2 and non-SARS-CoV-2 coronaviruses in individuals with PCR confirmed SARS-CoV-2 asymptomatic (blue), mild (purple) or severe (red) infection.

SARS-CoV-2 spike glycoprotein-specific (A) IgA memory B cells and (B) IgG2 subclass antibodies. Meso Scale Discovery (MSD) multiplexed immunoassay (MIA) platform measurements of antibody levels specific to (C) non-SARS-CoV-2 coronaviruses spike glycoproteins compared across the diseases cohorts at each sampling timepoint, and (D) SARS-CoV-2 spike glycoprotein over time. See Table S1 for number of individuals evaluated per assay.

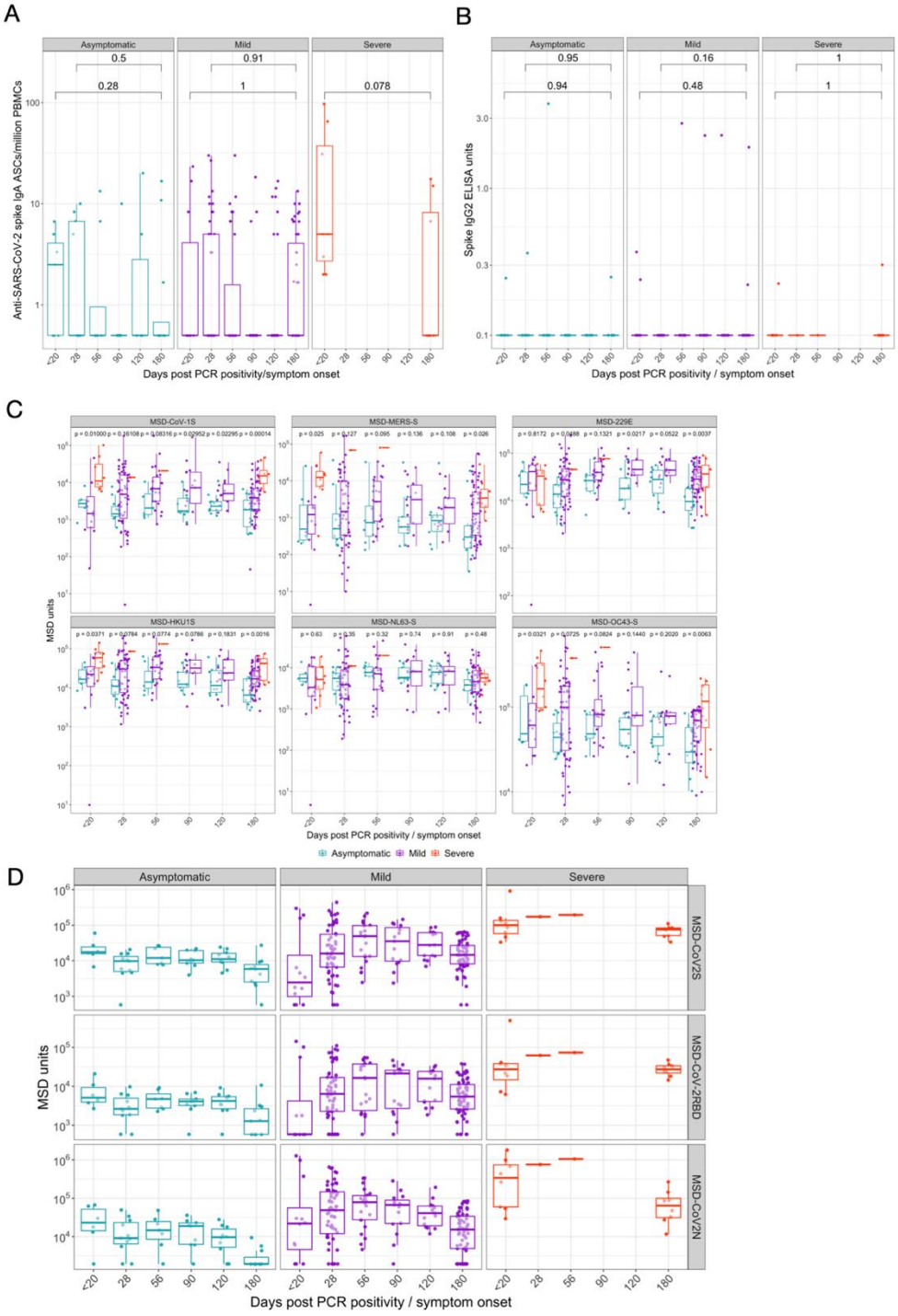


Fig S3. Analysis of T cell responses by clinical disease status, and representative gating strategies

(A) Comparison of *ex vivo* IFN-γ ELISpot responses in healthcare workers post asymptomatic (n=12) vs mild (n=66) SARS-CoV-2 infection. T cell responses to summed SARS-CoV-2 peptide pools spanning spike, accessory and structural proteins (summed total of SARS-CoV-2 proteins, S1, S2, NSP3B, M, NP, ORF 3, ORF8, and the CEFT positive control peptides for T cell responses) at 28, 90 and 180 days post onset of symptoms for mild cases, PCR positive test for asymptomatic participants. Comparison by Mann-Whitney test. **(B) to (D)** PBMC were stimulated with 1ug/ml of peptide pools or 2ug/ml of PHA and 0.2% DMSO for proliferation assay (B and C) or 2ug/mL of peptide pool for ICS assay (D). Representative gating strategy is shown in for T cell proliferation assay (B) . FECT control responses in the proliferation assay is shown in (C) and Representative ICS plots are shown in (D) for both CD4+ and CD8+ T cell responses.

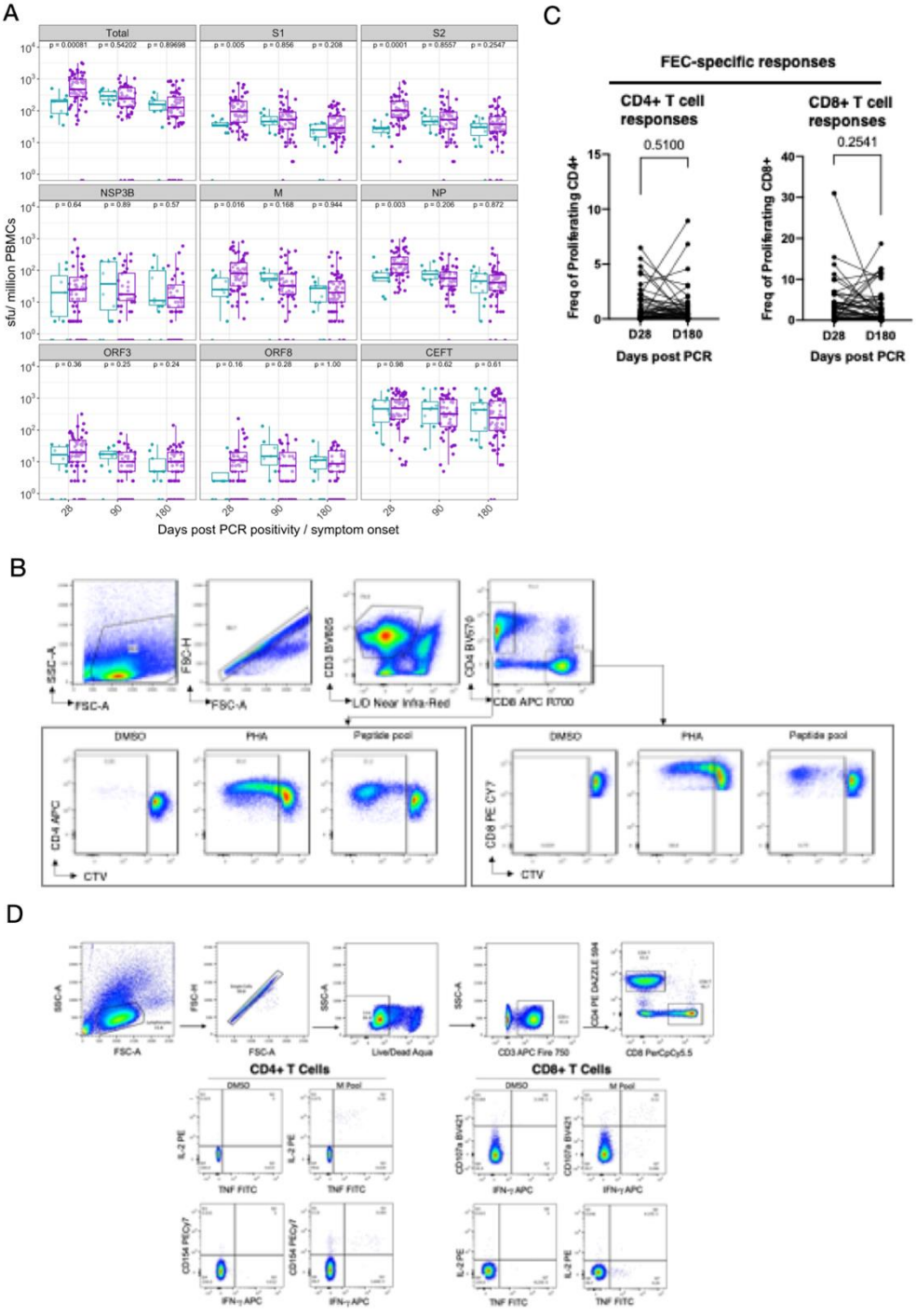


Fig S4 Longitudinal ICS analysis of SARS-CoV-2 T cell responses.
Individuals with ELISpot levels >100 spots/10^6 PBMC for a particular peptide pool at timepoint 2 were studied longitudinally at two timepoints using ICS. A total of n=18 individuals were studied with n=15 providing paired analysis. 1^6 frozen PBMC were stimulated with 2ug/mL of the listed peptide along with 1ug/mL CD28 and CD49d antibodies for 6 hours. Expression levels of IFN-γ, IL-2, and TNF in CD4+ T cells (black) and CD8+ T cells (grey) are shown at two timepoints for S1 **(A)**, S2 **(C)**, M **(E)**, NP **(G)** and NSP3B **(I).** A paired Friedman test was performed with the two-stage step-up method of Benjamini, Kriegar and Yekutieli correction for multiple comparisons. P values <0.1 are shown. Polyfunctional analysis was performed on both timepoints for individuals with sufficient cell populations for polyfunctional analysis. Data for CD4+ T cells are shown at two timepoints for S1 **(B)** (n=7, n=8 timepoints 1 and 2 respectively), S2 **(D)** (n= 5, n=6), M **(F)** (n=4, n=4), and NP pools **(H)** (n= 6, n=5). Data for CD8+ T cells are shown at two timepoints for NSP3B pools **(J)** (n=3, n=5).

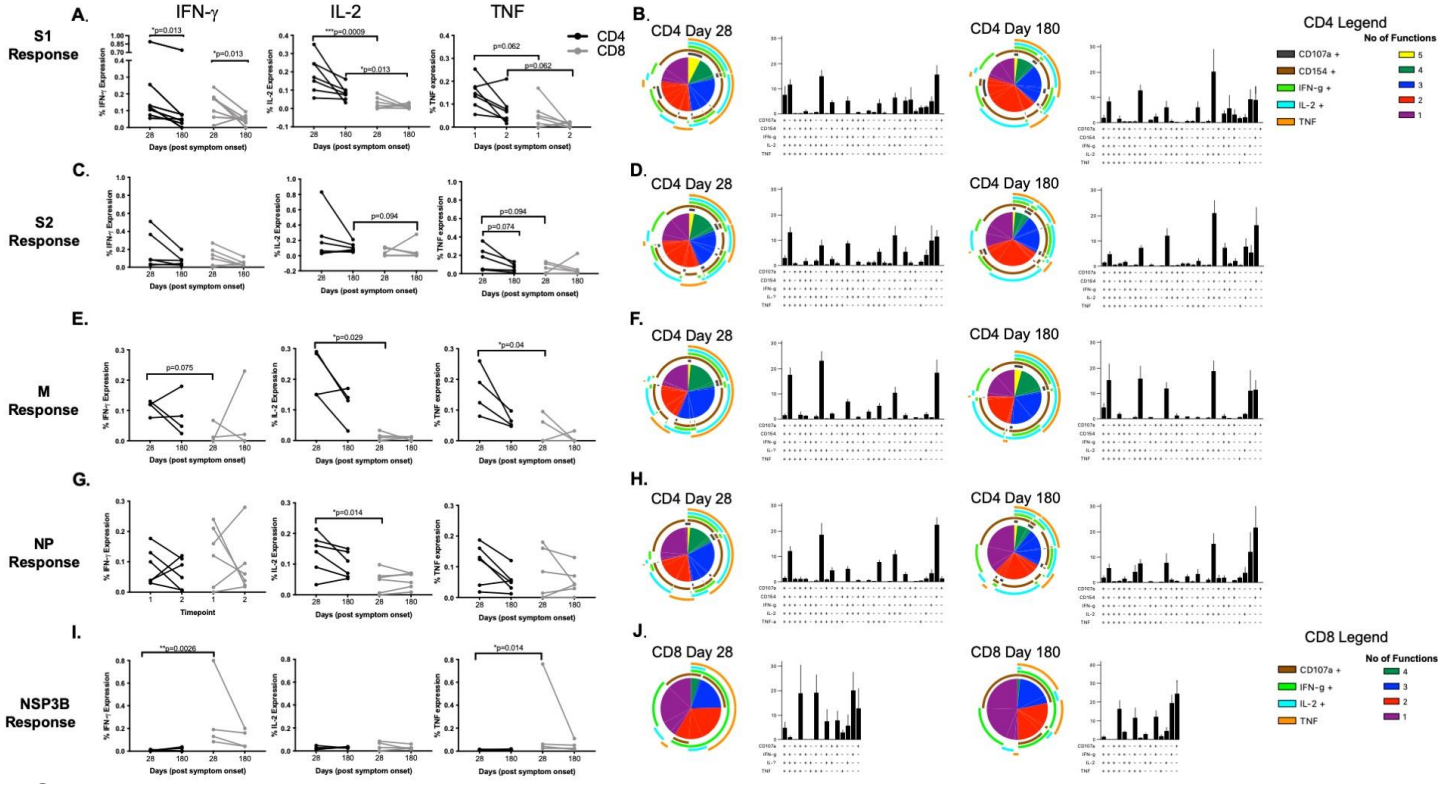
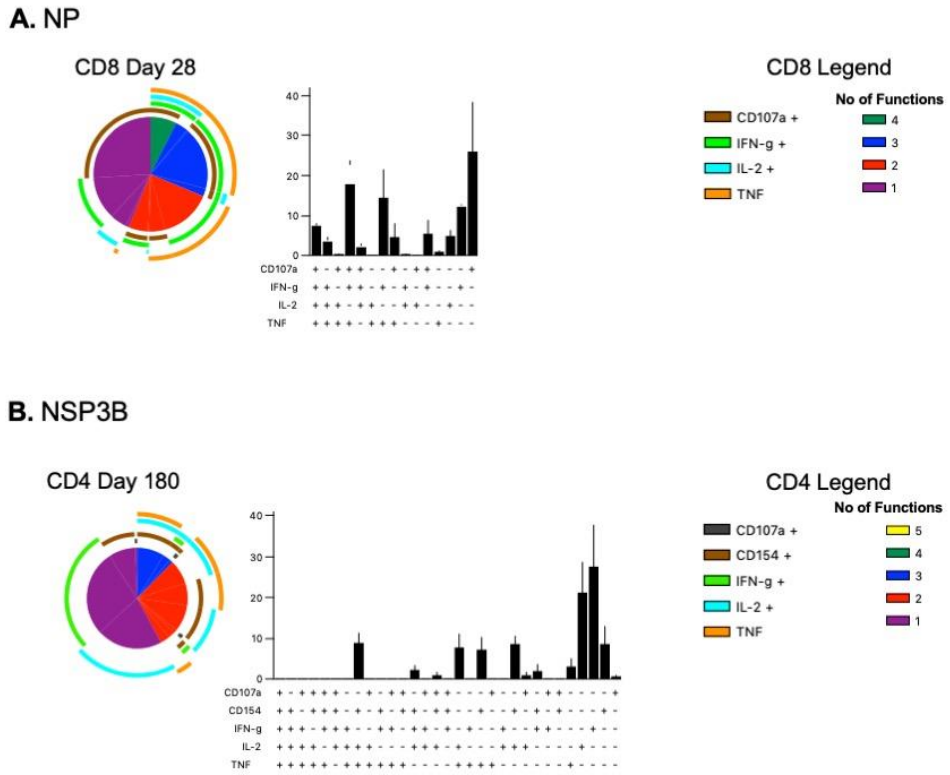


Fig S5. Polyfunctional T cell responses for NP and NSP3B pools.

Polyfunctional analysis was performed as in Supplemental Figure X. **A**) shows CD8+ T cell polyfunctional analysis for the NP pool at day 28 (n=3). **B**) shows CD4+ T cell polyfunctional analysis for the NSP3V pool at day 180 (n=5).



Additional Methods and materials

Peripheral blood mononuclear cells (PBMC) and plasma separation

PBMCs were isolated by density gradient centrifugation using Lymphoprep™ (p=1.077 g/ml, Stem Cell Technologies), washed twice with RPMI 1640 (Sigma, St. Louis, MO, USA) containing 10% heat-inactivated FCS (Sigma), 1mM Pen/Strep (100U/mL) and 2mM L-glutamine (100 ug/mL) (Sigma) or AutoMACS Rinse Buffer and resuspended in R10 or AutoMACs Rinse Buffer and counted using the Guava® ViaCount™ assay on the Muse Cell Analyzer (Luminex Cooperation). PBMCs were frozen and stored in liquid nitrogen. To obtain plasma, the uppermost fraction following the initial Lymphoprep centrifugation above was collected and centrifuged at 2000g for 10 minutes to remove platelets before storage at -80°C.

Serum isolation

Donor blood was collected in a serum-separating tube (SST, Becton Dickinson) which was centrifuged at 2000g for 10 minutes. Serum was removed and stored at -80°C.

IFN-γ Enzyme-Linked immunospot (ELISpot) assay

IFN-γ ELISpot was performed as previously described[1]. In brief, 96-well Multiscreen-I plates (Millipore, UK) were coated for 3 hours with 10 µg/ml clone 1-D1K, Mabtech, AB, Sweden) at room temperature. PBMC were added in duplicate wells at $2x10^5$ cells in 50 µl per well and stimulated with 50 µl of SARS-CoV-2 peptide pools (2ug/ml per peptide). R10 with DMSO (final concentration 0.4%, Sigma) was used as negative control. CMV, EBV, influenza and tetanus antigens (CEFT) peptide pool (2µg/ml, GenScript, Piscataway, NJ, USA were used as positive controls. After 16-18 hours at 37^{o}C PBMC were removed and secreted IFN-γ detected using anti-IFN-γ biotinylated mAb at 1 µg/ml (7-B6-1-biotin, Mabtech) for 2-3 hours, followed by streptavidin alkaline phosphatase at 1 µg/ml for 1-2 hours (SP-3020, Vector Labs). The plates were developed using BCIP/NBT substrate (Pierce) according to the manufacturer's instructions. ELISpot plates were read using an AID ELISpot Reader (v.4.0). Results were reported as spot-forming units (SFU)/10^6 PBMC. Background (mean SFU in negative control wells) was subtracted from antigen stimulated wells to give the final result. Only assays where the background was 50 SFU/10^6 or below were accepted as valid. The cut-off threshold for a positive result was the mean of the negative control well plus 2 times the standard deviation. The lower limit of quantification (LLOQ) for this assay is 2.5 SFU/10^6 PBMC, values below this were assigned a value of 1.

T cell proliferation assay

PBMCs were isolated from blood samples and used fresh or cryopreserved. CellTrace® Violet (CTV, Life Technologies) label was used at a final concentration of 2.5µM to label the cells. Labelling with CTV was done in PBS for 10 minutes at room temperature following which the reaction was stopped using ice-cold fetal bovine serum (FBS). The CTV-labelled PBMC were then plated at 0.25 x 10^6 cells per well of a 96 well round

bottom plate in RPMI supplemented with 10% human blood group type AB serum (Sigma), 1% 1mM Pen/Strep and 1% 2mM L-glutamine and stimulated with peptide pools from SARS-CoV-2 spanning Spike (S1 and S2), M, NP, ORF3 and ORF8, and FEC-T (1µg/ml per peptide). For controls, media containing 0.2% DMSO (Sigma) representing DMSO content in peptide pools was used as a negative control and phytohemagglutinin L (PHA-L, Sigma) at a final concentration of 2ug/ml was used as positive control. Cells were then incubated at 37^{o}C, 5% CO2, 95% humidity for 7 days with hemi-depletion of media on day 4. On day 7, cells were stained for analysis on the flow cytometer. Briefly, PBMC were resuspended in cell staining buffer (Biolegend, San Diego, CA, USA) and incubated for 20min with live/dead near-infrared (Invitrogen, Carlsbad, CA, USA). This was washed off and the cells were incubated with fluorochrome-conjugated primary human-specific antibodies for CD3, CD4 and CD8 in cell staining buffer (Biolegend, San Diego, CA, USA) for 30min at 4^{o}C. This was followed by a wash with cell staining buffer and then fixation with 4% paraformaldehyde (PFA, Sigma). Cells were stored at 4^{o}C in the dark until data acquisition on a MACSQuant 10. Responses above 1% were considered true positive. All data is reported as background subtracted data for each volunteer.

Intracellular cytokine staining

PBMC were thawed and rested overnight in R10 media (1 million cells for peptide stimulation and 500,000 for DMSO and PMA controls) in round bottom plates. Afterwards, cells were stimulated with SARS-CoV-2 peptide pools (2ug/ml), R10 media containing DMSO (0.1%, Sigma) for negative controls and PMA (0.05ug/mL) with ionomycin (0.5ug/mL, Sigma) as a positive control. CD107a BV421 (clone H4A3, BD Biosciences), monensin (Biolegend) and Brefeldin A (MP Biomedicals) were added to cultures at a final concentration of 0.04ug/mL, 0.16uM, and 10ug/mL respectively, and cells were incubated for 6 hours at 37^{o}C, 5% CO_2, 95% humidity. PBMC were then washed with PBS and stained with LIVE/Dead Fixable Aqua stain (Life Technologies) at a 1:400 dilution in PBS and stained for 20 minutes at room temperature. Cells were washed with PBS and resuspended in Cytofix/Cytoperm (BD Biosciences) and incubated for 20 minutes at 4°C. Afterwards, cells were washed twice with BD Perm/Wash buffer (BD Biosciences). Cells were then stained with the following panel of antibodies in Perm/Wash buffer for 20 minutes at room temperature: CD3 APC Fire 750 (clone SK7, Biolegend), CD4 PE Dazzle 594 (clone RPA T4, Biolegend), CD8 PerCpCy5.5 (clone RPA T8, Biolegend), CD154 PE-Cy7 (clone 24-31, Biolegend), IFN-γ APC (clone B27, BD Biosciences), IL-2 PE (clone MQ1-17HI2, Biolegend), TNF-α FITC (clone Mab11, BD Biosciences). Cells were then washed twice with Perm/Wash buffer before resuspending the cells in PBS and running them on a BD LSR II. OneComp compensation beads were used (Life Technologies) as were rainbow fluorescent particles (mid-range intensity) (Biolegend) to calibrate the LSR II before acquisition.

Anti-spike and anti-nucleocapsid total IgG (ELISA/EIA)

Standardised total anti-spike IgG ELISA was performed as described previously[2]. In brief, ELISA plates were coated with 2µg/mL of full-length trimerised SARS-CoV-2 spike glycoprotein protein overnight at 4°C and

blocked with casein in PBS. Plasma samples were diluted in PBS and tested in triplicate. Goat anti-human IgG conjugated to alkaline phosphatase was added as the secondary antibody, and plates were developed using 4-nitrophenyl phosphate in diethanolamine substrate buffer. Plates were read at 405nm, and standardised ELISA units (EU) were determined using a 4-parameter logistic model and various pre-determined control cut-offs (Gen5 v3.09, BioTek). Plate washing in-between each step was undertaken using 0.05% Tween-20 in PBS.Serology for IgG to SARS-CoV-2 nucleocapsid protein was performed using the Abbott Architect i2000 chemiluminescent microparticle immunoassay (Abbott, Maidenhead, UK) and carried out according to manufacturer's instructions using serum. The manufacturer threshold for confirming detection of antibodies is ≥1.40 arbitrary units. Levels between 0.50-1.39 arbitrary units designate equivocal levels (Abbott Diagnostics Product Information Letter PI1060-2020). Values below 0.5 were set to half the LLOQ (i.e. 0.25).

Anti-spike subclass and isotype ELISAs

Both isotype and subclass standardised and OD ELISAs were performed as described previously[3]. In brief, ELISA plates were coated with 5 µg/mL of full-length trimerised SARS-CoV-2 spike protein for overnight incubation at 4°C. Following washing with 0.05% Tween-20 in PBS (PBS/T) plates were blocked with casein in PBS for non-specific binding. In the next step plasma samples were diluted in casein in PBS, as well as positive, negative controls and ten-point standard curve. Plates were incubated for 2h at 37°C with 300 rpm shaking and following washing with PBS/T samples were further incubated with mouse anti-human IgG1 hinge-AP, mouse anti-human IgG3 hinge-AP, goat anti-human IgA-AP and goat anti-human IgM-AP (Southern Biotech) for 1 h at 37 °C with 300 rpm used for detection and the optical density at 405nm was measured until the internal control reached an OD_{405} of 1. For detection of anti-spike IgG2 and IgG4 steps modified as follows: 1) Plates were additionally coated with commercially available human immunoglobulin control (recombinant human IgG2 lambda or recombinant human IgG4 lambda (Bio-Rad)) to serve as internal controls, 2) Mouse anti-human IgG2 Fd-AP or mouse anti-human IgG4 Fc-AP (Southern Biotech) were used, and 3) Optical density at 405 nm was measured using an ELx808 absorbance reader (BioTek) until the immunoglobulin control reached a specified OD405.

Standardised ELISA units (EU) were determined using a 4-parameter logistic model and various pre-determined control cut-offs (Gen5 v3.09, BioTek) while for OD ELISAs negative cut offs were calculated using the formula: mean + 7.858 × standard deviation of the OD405 readings of the pre-pandemic negative-control serum samples, where 7.858 is the standard deviation multiplier with a 99.9% confidence level for n = 5 controls as detailed in [3, 4]. LLOQ were 11, 12, 12, 18, 0.2, 8 and 0.2 EU for total IgG, IgM, IgA, IgG1, IgG2, IgG3 and IgG4, respectively — values below LLOQ were set to half the LLOQ.

MSD Common Cold Coronaviruses

A multiplexed MSD immunoassay (MSD, Rockville, MD) was used to measure the IgG responses to SARS-CoV-2, severe acute respiratory syndrome coronavirus-1 (SARS-CoV-1), MERS-CoV and seasonal CoVs (human

coronavirus (HCoV)-OC43, HcoV-HKU1, HcoV-229E, HcoV-NL63). A MULTI-SPOT® 96-well, 10 Spot Plate was coated with three SARS CoV-2 antigens (S, RBD, N), SARS and MERS-CoV spike trimers, as well as spike proteins from seasonal CoV HCoV-OC43, HCoV-HKU1, HCoV-229E, HCoV-NL63 and bovine serum albumin. Antigens were spotted at 200−400 µg/ml (MSD® Coronavirus Plate 3). Multiplex MSD Assays were performed as per the instructions of the manufacturer. To measure IgG antibodies, 96-well plates were blocked with MSD Blocker A for 30 minutes. Following washing with washing buffer, our samples diluted 1:500-1:5000 in diluent buffer, as well as the reference MSD standard and internal MSD controls were added to the wells. After 2-hour incubation and a washing step, detection antibody (MSD SULFO-TAG™ Anti-Human IgG Antibody, 1/200) was added. Following washing, MSD GOLD™ Read Buffer B was added and plates were read using a MESO® SECTOR S 600 Reader. The standard curve was established by fitting the signals from the standard using a 4-parameter logistic model. Concentrations of samples were determined from the electrochemiluminescence signals by back-fitting to the standard curve. They were multiplied by the dilution factor and expressed in Arbitrary Units/ml. LLOQ were 1160.3, 1169.0 and 3873.5 AU/ml for SARS-CoV-2 spike, SARS-CoV-2 RBD and SARS-CoV-2 NP, respectively — values below the LLOQ were set to half the LLOQ.

Microneutralisation Assay (MNA)

Virus Isolates

Prototype isolate (PANGO lineage B) was Victoria/01/2020 [5] received at Passage(P)3 from Public Health England (PHE) Porton Down (after being supplied by the Doherty Centre Melbourne) in April 2020, passaged in VeroE6/TMPRSS2 cells, used here at P5, and confirmed identical to GenBank MT007544.1, B hCoV-19_Australia_VIC01_2020_ EPI_ ISL_ 406844_ 2020-01-25.

B.1.1.7 (20I/501Y.V1.HMPP1) isolate, H204820430, 2/UK/VUI/1/2020, received in Oxford at P1 from PHE Porton Down in December 2020, passaged in VeroE6/TMPRSS2 cells (NIBSC reference 100978), used here at P4. B.1.351 (20I/501.V2.HV001) isolate was received at P3 from the Centre for the AIDS Programme of Research in South Africa (CAPRISA), Durban, in Oxford in January 2021, passaged in VeroE6/TMPRSS2 cells (NIBSC reference 100978), used here at P4.

For all isolates, identity was confirmed by deep sequencing at the Wellcome Trust Centre for Human Genetics, University of Oxford.

Virus neutralisation

The microneutralisation assay determines the concentration of antibody that produces a 50% reduction in infectious focus-forming units of authentic SARS-CoV-2 in Vero CCL81 cells. Quadruplicate serial dilutions of serum were preincubated with 100-200 FFU (20 µL) of SARS-CoV-2 for 30 minutes at room temperature. After pre-incubation, 100 µL of Vero CCL81 cells (4.5×10^4) were added and incubated at 37°C, 5% CO_2. After 2 hours, 100 µL of a 1.5% carboxymethyl cellulose-containing overlay was applied to prevent satellite focus formation. Eighteen (B.1.351) or 23 hours (B, B.1.1.7) post-infection, the monolayers were fixed with 4%

paraformaldehyde, permeabilized with 2% Triton X-100 and stained for the nucleocapsid antigen or spike (S) antigen using monoclonal antibodies (mAbs) EY 2A and EY 6A, respectively[6]. After development with a peroxidase-conjugated antibody and TrueBlue peroxidase substrate, infectious foci were enumerated by ELISpot reader. Data were analysed using four-parameter logistic regression (Hill equation) in GraphPad Prism 8.3.

Monogram Bioscience pseudotype neutralisation assay (PseudoNA)

The Monogram Biosciences pseudotype neutralisation assay was performed as described previously[2]. Plasma samples were heat inactivated at 56°C for one hour and diluted in a 9 serial three-fold dilution series starting at 1:40 in cell culture medium. Each sample dilution was mixed with 10^5 relative light units of a lentivirus-based SARS-CoV-2 pseudovirus particle virus. As an internal assay control, an irrelevant pseudotyped virus was also incubated with test samples. The pseudotyped virus and sample mixtures were incubated for one hour at 37°C and HEK 293 ACE2-transfected cells were added to each well, which were then incubated for a further 60-80 hours at 37°C. Luciferase expression was determined, and neutralisation titres are reported as the reciprocal of the plasma dilution conferring 50% inhibition (ID50) of pseudovirus infection. %Inhibition = 100% – (((RLU(Vector+Sample+Diluent) – RLU(Background))/(RLU(Vector+Diluent) – RLU(Background))) x 100%). Inter-assay variation was controlled for by monitoring results acquired from one positive control, one negative control and six patient specimens. The LLOQ for this assay is a titre of 1:40, values below this were set to half LLOQ (i.e. 20).

Spike-specific SARS-CoV-2, OC43, HKU1, 229E and NL63 IgG^+ and IgA^+ B cell memory ELISPOT

PBMCs were adjusted to $2x10^6$ cells/ml in complete media and $2x10^5$ cells/per well were added to a 96-well round bottomed plate with an equal volume of complete media supplemented with 1µg/ml R848 and 10ng/ml of recombinant IL-2, each from the Mabtech Memory B-cell Stimpack. The cells were cultured for 3-3.5 days at 37°C in 5% CO_2. Following polyclonal stimulation, the cells were harvested, washed twice in complete media and counted.

Mabtech flurospot plates were activated with 35% ethanol and coated with the relevant spike glycoprotein (SARS-CoV-2 at 10µg/ml, OC43 at 10µg/ml, NL63 at 15µg/ml, HKU1 at 5µg/ml and 229E at 10µg/ml, all diluted in PBS). Control wells were coated with tetanus toxoid (5µg/ml), capture mAbs anti-human IgG (Mabtech MT91/145) and PBS as a negative control. Following incubation for 16-20 hours at 4°C, the plates were washed five times with PBS and blocked for ≥30 minutes with complete media. The harvested PBMCs were adjusted so that $2x10^5$ cells were added to the spike- and tetanus toxoid-coated, and PBS wells while $2x10^4$ cells were added to the IgG positive control wells. All cells were incubated for ≥16 hours at 37°C, plates were washed five times with PBS and detection mAbs IgG-550 (Mabtech MT78/145) and IgA (Mabtech MT20-490) were diluted 1:500 in 0.5% BSA in PBS and added to plates for two hours at room temperature. Following

five washes in PBS, fluorescent enhancer solution was added to each well for 15 minutes at RT in the dark. Plates were decanted and blotted dry and stored in the dark. Spot forming units were enumerated using AID ELISpot 8.0 software on the AID ELR08IFL reader. The LLOQ for these assays is 1 SFU, values below this were set to half LLOQ (i.e. 0.5).

Antibody-dependent effector functions

The spike-specific antibody-dependent effector functions, natural killer cell activity (ADNKA), neutrophil phagocytosis (ADNP) and monocyte phagocytosis (ADMP) were performed as previously described [3].

Bead coupling for ADNP and ADMP: Red fluorescent (580/605) NeutrAvidin-labeled microspheres (Thermo Fisher, F8775) were freshly coupled to biotinylated SARS-CoV-2 spike protein for each assay. Spike protein (at a concentration of 0.388 µl/ml) was added to the beads at a 3:1 ratio and incubated for 2 h at 37 °C. Beads were washed twice with 0.1% BSA and diluted 100-fold in 0.1% BSA. 10 µl was added to each well in the ADNP and ADMP assays.

ADNP assay

Whole donor blood, collected in sodium heparin tubes, was treated with ammonium–chloride–potassium lysing buffer (Thermo Fisher, A1049201) for 5 minutes followed by centrifugation to collect white blood cells. Cells were washed with DPBS (Sigma, D8537), counted and adjusted to 2.5 × 105 cells per ml in medium consisting of RPMI 1640 medium (Sigma, R5886) supplemented with 100 U ml−1 penicillin–streptomycin (Sigma, P4458) and 20 mmol/L l-glutamine (Sigma, G7513).

Serum diluted 100-fold in RPMI was added to antigen-coupled beads in a 96-well plate and incubated for 2 h at 37 °C. All samples were assayed in duplicate, and each plate contained two QC samples in addition to appropriate negative controls. Wells were washed with DPBS, and 50,000 white blood cells were added to each well followed by a further one hour incubation at 37 °C. Cells were then stained using a cocktail of mouse anti-human CD3 Alexa Fluor 700 (BD Pharmingen, clone UCHT1, nos. 557943 and 9185576; 1:80 dilution), mouse anti-human CD14 APC Cy7 (BD Pharmingen, clone MΦP9, nos. 557831 and 0044497; 1:80 dilution) and mouse anti-human CD66b Pacific Blue (BioLegend; clone G10F5, nos. 305112 and B285068; 1:80 dilution) and incubated for 15 minutes at room temperature in the dark. Following washing and fixation using 4% paraformaldehyde (Santa Cruz Biotechnology, SC-281692), cells were analysed by flow cytometry (BD, Fortessa X20).

Data were analysed with FlowJo (BD; version 10), using a gating strategy to select neutrophils. Neutrophils were gated based on forward and side scatter then doublets excluded. Furthermore, T cells and monocytes were excluded using a double-negative gate for CD3 and CD14. The final neutrophil gate was based on CD66b

positivity, after which bead-positive cells were gated. In all cases, there was a clear separation between positive and negative populations.

Normalized phagocytic scores were calculated by multiplying the percentage of bead-positive cells by the MFI of the events within the bead-positive cell gate and normalizing against a QC sample. As multiple plates were run during an experiment, plates failed if any of the QC sample averages were greater than two standard deviations above the mean of that particular QC across plates. In addition, samples were excluded from further analysis if the replicates showed a coefficient of variation of over 25%. All data are derived from one experiment. The LLOQ for this assay is 0.033, values below this were set to half the LLOQ.

ADMP assay

Human monocytic THP-1 cells (American Type Culture Collection) were grown and maintained using supplier instructions. Serum was diluted 1:4,000 in RPMI, added to antigen-coupled beads in a 96-well plate, and incubated for two hours at 37 °C. All samples were assayed in duplicate, and each plate contained two QC samples in addition to appropriate negative controls. At the end of the two hours incubation period, wells were washed with RPMI and 25,000 THP-1 cells diluted in medium consisting of RPMI 1640 medium (Sigma, R5886) supplemented with 100 U/mL penicillin–streptomycin (Sigma, P4458) and 20 mmol/L l-glutamine (Sigma, G7513) were added to each well. Plates were then incubated for 18 hours at 37 °C. Cells were then washed with PBS and fixed using 4% paraformaldehyde before analysis by flow cytometry (BD, Fortessa X20). Data were analysed with FlowJo (BD, version 10). THP-1 cells were gated based on forward and side scatter to exclude debris then doublets excluded and bead-positive cells gated. There was a clear separation between the positive and negative population.

Normalized phagocytic scores were calculated by multiplying the percentage of bead-positive cells with the MFI of the events within the bead-positive cell gate and normalizing against a QC sample. As multiple plates were run during an experiment, plates failed if any of the QC sample averages were greater than two standard deviations above the mean of that particular QC across plates. In addition, samples were excluded from further analysis if the replicates showed a coefficient of variation of over 25%. All data were derived from one experiment. The LLOQ for this assay is 0.23, values below this were set to half the LLOQ.

ADNKA assay

SARS-CoV-2 spike protein in carbonate/bicarbonate solution (2.5µg/ml) was added to 96-well Nunc MaxiSorp ELISA plates and incubated for 16 hours at 4°C. Plates were washed six times with PBS and blocked with 5% BSA in PBS for one hour at 37°C. Plasma samples were added neat and in duplicate, and plates were incubated for two hours at 37°C. Following another wash step, 10^5 natural killer NK-92 cells expressing human CD16 (PTA-8836 cell line, American Type Culture Collection; described by [7] were added to each well with brefeldin

A (10µg/mL; Sigma Aldrich), GolgiStop (BD Biosciences) and CD107a (1:20 dilution; PE, clone H4A3, BD Biosciences). Plated cells were incubated for five hours at 37°C and then transferred to V-bottom plates, incubated with fixable LIVE/DEAD staining (1:500 dilution; R780, BD Biosciences) and fixed. Data was acquired using a BD Fortessa and percentages of CD107a expressing NK cells relative to control wells with spike protein and blocking buffer only were determined using FlowJo Software (version 10.7.1). To assess inter-assay variation, both a pre-pandemic pool of three donors and a pool of six hospitalised SARS-CoV-2-infected individuals were plated in triplicate on each plate. The LLOQ for this assay is 3.5, values below this were set to half the LLOQ.

Antibody-dependent complement deposition (ADCD) assay

SPHERO^{TM} Carboxyl magnetic blue fluorescent beads (Spherotech, USA) were coupled with SARS-CoV-2 whole spike protein (Lake Pharma, USA) using a two-step Sulpho-NHS/EDC process. Briefly, 5 million beads were washed with 82mM sodium phosphate buffer pH 6.2, prior to activation in the same buffer containing 1.24mg each of N-hydroxysulfosuccinimide and 1-ethyl-3-[3- dimethlyaminopropyl]carbodiimide-HCl). After 20 min activation, the beads were washed in coupling buffer of 50mM 2-(N-morpholino) ethanesulfonic acid (MES) pH 5.0 and resuspended in MES buffer containing 14.5µg antigen for 2 h on a rotational mixer. Finally, beads were washed three times with PBS containing 2% BSA and 0.05% sodium azide, pH7.4, and resuspended in the same buffer overnight. Beads were washed and resuspended in a storage buffer of PBS with 0.05% sodium azide, pH7.4, until use.

Heat-inactivated test serum (3µl, in duplicate) was added to 27µl assay blocking buffer (PBS + 2% BSA:BB) and 10µl taken for serial 3-fold dilutions to give final dilutions of 1:20, 1:60, 1:180, 1:540. 20µl of spike-coated magnetic beads (50 beads per µl) was added, and the mixture incubated at 25^{o}C for 30min with shaking at 900rpm. The beads were washed twice in 200µl wash buffer (BB + 0.05% Tween-20: WB) and then resuspended in 50µl BB containing 10% IgG- and IgM-depleted human plasma, prepared as described previously [8] and incubated at 37^{0}C for 15min with shaking at 900rpm. Beads were next washed twice with 200µl WB and resuspended in 100µl FITC-conjugated rabbit anti-human C3c polyclonal antibody (Abcam, UK) and incubated at room temperature in the dark. After two more washes with 200µl WB, the samples were resuspended in 40µl Hank's Balanced Salt Solution and analysed on the IntelliCyt® iQue Screener PLUS platform (Sartorius, Germany) and ForeCyt®t 8.0 software. For each sample, a minimum of 100 beads were collected and complement activation units (CAU) calculated using a 12-point standard curve of the Anti-SARS-CoV-2 Antibody Diagnostic Calibrant (20/162 NIBSC, UK), with the calibrant standard assigned 1000 CAU. The LLOQ for this assay is 10, values below this were set to half the LLOQ.

Generation of the integrated dataset and data pre-processing

The integrated dataset was generated using the standard extract-transform-load (ETL) procedure, as described [9]. Briefly, primary analysis datasets which included total of 29 csv files across 14 assays and clinical data were merged using donor-specific variable (Donor ID). The outcome of immune response durability was calculated based on the titre of the anti-nucleocapsid specific antibodies measured 6 months post symptoms onset (pso). High responders were determined as individuals who are seropositive 6 months pso, i.e., have anti-N antibody titre ≥ 1.4, while low responders are individuals having anti-N antibody titre below 1.4. The responder status was expressed as a binary value: high responders were given a value of 1, whereas low responders a value of 0. Before the integrative analysis, data was pre-processed using transformation methods available in SIMON knowledge discovery software [10] center (mean subtracted) and scale (standard deviation divided) applied before principal component analysis (PCA), t-distributed stochastic neighbour embedding (t-SNE), hierarchical clustering and SIMON analysis, missing values were imputed based on median values (*medianImpute*) (PCA, t-SNE and hierarchical clustering), features with zero-variance (*zv*) and near-zero-variance (*nzv*) were removed (PCA and SIMON), and finally, highly correlated features with cut-off 0.85 (*corr*) were also removed for the supervised machine learning (ML) analysis using SIMON.

Integrative analysis using unsupervised machine learning analysis

<u>High-dimensional analysis using t-SNE and clustering analysis.</u>

The t-distributed stochastic neighbour embedding (t-SNE) followed by clustering was performed to analyse the pre-processed integrated dataset using SIMON software [10]. Disease severity and timepoint were used as grouping variables, and thus, were excluded from the analysis. T-SNE analysis was performed with 2,000 iterations, a perplexity of 30, and a theta of 0.5. Resulting t-SNE maps were used for cluster analysis using model-based clustering algorithm (*mclust*) with seed number 1337 and 3 clusters allowed [11]. To visualize variation of clinical and immunological features across the t-SNE embedding space, we performed hierarchical clustering on t-SNE maps using Euclidean distance, agglomerative hierarchical clustering with Ward and tightest cluster was ordered first.

<u>Principal component analysis.</u>

Principal component analysis (PCA) was performed on multivariate immunological parameters (continuous variables) with pre-filtering to remove all categorical variables and features with less than 10% of unique values, *i.e.*, any column that has number of unique values less than 10% of total number of observations. Disease severity was used as a grouping variable. Quality of variable representations (cos2), variable correlations and contributions (expressed as percentage) of top 10 variables to first two principal components (PCs) were calculated. The correlation between variables and PCs was used as the coordinates of the variables on the PCs. The observations were represented by their projections, while the variables were represented by their correlations [12].

Correlation analysis.

Pairwise correlations of immunological parameters on all analysed samples were calculated and visualized as a correlogram using a SIMON software. Spearman's rank correlation coefficient was computed and indicated on the correlogram by the heat scale. The significance test of correlation coefficients was performed, and values shown on the correlogram were adjusted for multiple testing using false discovery rate (FDR) correction using the Benjamini-Hochberg correction at the significance threshold FDR < 0.05. Following correlation analysis, correlogram map was fed into agglomerative hierarchical clustering with Ward algorithm and three major clusters were identified.

Hierarchical clustering.

Agglomerative hierarchical clustering was performed on the samples with immunological parameters analysed on day 28 pso and visualized as the dendrogram on heatmap using a SIMON software. Cluster analysis was performed using a set of dissimilarities for the number of samples being clustered. Each sample was assigned to its own cluster, then the algorithm iteratively joined the two most similar clusters, continuing until there was just a single cluster. First, the dissimilarity values were computed (dissimilarity matrix calculated using Euclidean method, these values were then fed into hierarchical clustering using complete linkage agglomeration method and finally, dendrogram was plotted (tightest cluster ordered first).

Integrative analysis using SIMON

To identify early immunological signature at day 28 pso that can predict if the individual will be high or low responder 6 months pso, we performed SIMON (Sequential Iterative Modeling Over Nigh) analysis [13, 14] Joint predictive analysis on all immunological parameters at day 28 pso (excluding clinical data) was performed using 172 ML algorithms. The outcome was seropositivity status determined 6 months pso. Initial data was split into train/test partition (75%/25%) preserving the balanced distribution of the outcome class (seed number 1337). The dataset had 29% missing values, and missing values were removed using multi-set interaction function ('mulset, SIMON software) and 30 resamples were used for the SIMON analysis. The models were evaluated using 10-fold cross-validation on the training sets, and additionally to prevent overfitting on the held-out test sets. The best performing model was built using the Sparse Partial Least Squares (sPLS) algorithm (train AUROC: 0.95 (CI 0.5-1) and test AUROC: 1). In the final step, SIMON calculated the contribution of each feature to the model as variable importance score (scaled to maximum value of 100).

Statistical analyses

Statistical analysis was performed using R (https://www.r-project.org/), and figures were made with R using R package ggplot2[15]. Kruskal-Wallis test —unless otherwise specified — was used for comparison of the disease severity groups. Wilcoxon rank-sum test —unless otherwise specified — was employed to compare

between study time points. A generalised additive mixed model (GAMM) by restricted maximum likelihood (REML) was used to fit the immunological measures (log10 transformed) taken at multiple study time points, using Gaussian process smooth term. Disease severity group was included in the GAMM as a linear predictor and a participant identifier was included as a random effect. The R package *gamm4* was utilised for the GAMM analysis [16]. For ICS cytokine expression analyses, data was prepared using PESTEL v2.0 for formatting and baseline subtraction, followed by export of data to SPICE v6.0 for analysis. Statistical significance was set at $P<0.05$ and all tests were 2-tailed. Machine learning analysis was performed using SIMON software (https://genular.org).

References for Additional Methods

1. Ogbe A, Kronsteiner B, Skelly DT, Pace M, Brown A, Adland E, . . . Oxford Protective TCIfC-CT. T cell assays differentiate clinical and subclinical SARS-CoV-2 infections from cross-reactive antiviral responses. *Nature Communications* **12**, 2055 (2021).

2. Folegatti PM, Ewer KJ, Aley PK, Angus B, Becker S, Belij-Rammerstorfer S, . . . Oxford CVTG. 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. *Lancet* **396**, 467-478 (2020).

3. Barrett JR, Belij-Rammerstorfer S, Dold C, Ewer KJ, Folegatti PM, Gilbride C, . . . Oxford CVTG. Phase 1/2 trial of SARS-CoV-2 vaccine ChAdOx1 nCoV-19 with a booster dose induces multifunctional antibody responses. *Nat Med* **27**, 279-288 (2021).

4. Frey A, Di Canzio J, Zurakowski D. A statistically defined endpoint titer determination method for immunoassays. *J Immunol Methods* **221**, 35-41 (1998).

5. Caly L, Druce J, Roberts J, Bond K, Tran T, Kostecki R, . . . Catton MG. Isolation and rapid sharing of the 2019 novel coronavirus (SARS-CoV-2) from the first patient diagnosed with COVID-19 in Australia. *Med J Aust* **212**, 459-462 (2020).

6. Huang KA, Tan TK, Chen TH, Huang CG, Harvey R, Hussain S, . . . Townsend AR. Breadth and function of antibody response to acute SARS-CoV-2 infection in humans. *PLoS Pathog* **17**, e1009352 (2021).

7. Binyamin L, Alpaugh RK, Hughes TL, Lutz CT, Campbell KS, Weiner LM. Blocking NK cell inhibitory self-recognition promotes antibody-dependent cellular cytotoxicity in a model of anti-lymphoma therapy. *J Immunol* **180**, 6392-6401 (2008).

8. Lesne E, Cavell BE, Freire-Martin I, Persaud R, Alexander F, Taylor S, . . . Gorringe A. Acellular Pertussis Vaccines Induce Anti-pertactin Bactericidal Antibodies Which Drives the Emergence of Pertactin-Negative Strains. *Front Microbiol* **11**, 2108-2108 (2020).

9. Tomic A, Tomic I, Dekker CL, Maecker HT, Davis MM. The FluPRINT dataset, a multidimensional analysis of the influenza vaccine imprint on the immune system. *Sci Data* **6**, 214 (2019).

10. Tomic A, Tomic I, Waldron L, Geistlinger L, Kuhn M, Spreng RL, . . . Davis MM. SIMON: Open-Source Knowledge Discovery Platform. *Patterns (N Y)* **2**, 100178 (2021).

11. Fraley C, Raftery AE. Model-Based Clustering, Discriminant Analysis, and Density Estimation. *Journal of the American Statistical Association* **97**, 611-631.

12. Abdi H, Williams LJ. Principal Component Analysis. *Wires Computational Statistics* **2**, 433-459 (2010).

13. Tomic A, Tomic I, Rosenberg-Hasson Y, Dekker CL, Maecker HT, Davis MM. SIMON, an Automated Machine Learning System, Reveals Immune Signatures of Influenza Vaccine Responses. *Journal of immunology (Baltimore, Md : 1950)* **203**, 749-759 (2019).

14. Tomic A, Tomic I, Waldron L, Geistlinger L, Kuhn M, Spreng RL, . . . Davis MM. SIMON: Open-Source Knowledge Discovery Platform. *Patterns (New York, NY)* **2**, 100178-100178 (2021).

15. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York (2016).

16. Wood S, Scheipl F. gamm4: Generalized Additive Mixed Models using 'mgcv' and 'lme4'. R package version 0.2-6.) (2020).

613 **Peptide Sequences**

Peptide Pool	Peptide ID	Peptide Sequence
NSP3B	ORF1a/1ab_207	DWSYSGQSTQLGIEFLKR
	ORF1a/1ab_208	TQLGIEFLKRGDKSVYY
	ORF1a/1ab_209	LKRGDKSVYYTSNPTTF
	ORF1a/1ab_210	VYYTSNPTTFHLDGEVI
	ORF1a/1ab_211	TTFHLDGEVITFDNLKTL
	ORF1a/1ab_212	VITFDNLKTLLSLREVR
	ORF1a/1ab_213	KTLLSLREVRTIKVFTTV
	ORF1a/1ab_214	VRTIKVFTTVDNINL
	ORF1a/1ab_215	VFTTVDNINLHTQVVDM
	ORF1a/1ab_216	INLHTQVVDMSMTYGQQF
	ORF1a/1ab_217	DMSMTYGQQFGPTYL
	ORF1a/1ab_218	YGQQFGPTYLDGADVTKI
	ORF1a/1ab_219	YLDGADVTKIKPHNSHEG
	ORF1a/1ab_220	KIKPHNSHEGKTFYVL
	ORF1a/1ab_221	SHEGKTFYVLPNDDTLRV
	ORF1a/1ab_222	VLPNDDTLRVEAFEYY
	ORF1a/1ab_223	TLRVEAFEYYHTTDPSFL
	ORF1a/1ab_224	YYHTTDPSFLGRYMSAL
	ORF1a/1ab_225	SFLGRYMSALNHTKKWKY
	ORF1a/1ab_226	ALNHTKKWKYPQVNGL
	ORF1a/1ab_227	KWKYPQVNGLTSIKW
	ORF1a/1ab_228	QVNGLTSIKWADNNCYL
	ORF1a/1ab_229	IKWADNNCYLATALLTL
	ORF1a/1ab_230	CYLATALLTLQQIELKF
	ORF1a/1ab_231	LTLQQIELKFNPPAL
	ORF1a/1ab_232	IELKFNPPALQDAYYRAR
	ORF1a/1ab_233	ALQDAYYRARAGEAANF
	ORF1a/1ab_234	RARAGEAANFCALILAY
	ORF1a/1ab_235	ANFCALILAYCNKTVGEL
	ORF1a/1ab_236	AYCNKTVGELGDVRETM
	ORF1a/1ab_237	GELGDVRETMSYLFQHAN
	ORF1a/1ab_238	TMSYLFQHANLDSCKRVL
	ORF1a/1ab_239	ANLDSCKRVLNVVCK
	ORF1a/1ab_240	CKRVLNVVCKTCGQQQTT
	ORF1a/1ab_241	CKTCGQQQTTLKGVEAVM
	ORF1a/1ab_242	TTLKGVEAVMYMGTLSY
	ORF1a/1ab_243	AVMYMGTLSYEQFKKGV
	ORF1a/1ab_244	LSYEQFKKGVQIPCTCGK
	ORF1a/1ab_245	GVQIPCTCGKQATKYLV
	ORF1a/1ab_246	CGKQATKYLVQQESPFVM
	ORF1a/1ab_247	LVQQESPFVMMSAPPAQY
	ORF1a/1ab_248	VMMSAPPAQYELKHGTF
	ORF1a/1ab_249	AQYELKHGTFTCASEY
	ORF1a/1ab_250	HGTFTCASEYTGNYQCGH
	ORF1a/1ab_251	EYTGNYQCGHYKHITSK
	ORF1a/1ab_252	CGHYKHITSKETLYCI
	ORF1a/1ab_253	ITSKETLYCIDGALLTK
	ORF1a/1ab_254	YCIDGALLTKSSEYKGPI
	ORF1a/1ab_255	TKSSEYKGPITDVFYK
Peptide Pool	Peptide ID	Peptide Sequence

NSP3B continued	ORF1a/1ab_256	KGPITDVFYKENSYTTTI
	ORF1a/1ab_257	YKENSYTTTIKPVTYKL
	ORF1a/1ab_258	TTIKPVTYKLDGVVCTEI
	ORF1a/1ab_259	KLDGVVCTEIDPKLDNYY
	ORF1a/1ab_260	EIDPKLDNYYKKDNSYF
	ORF1a/1ab_261	NYYKKDNSYFTEQPIDLV
	ORF1a/1ab_262	YFTEQPIDLVPNQPY
	ORF1a/1ab_263	PIDLVPNQPYPNASFDNF
	ORF1a/1ab_264	PYPNASFDNFKFVCDNIK
	ORF1a/1ab_265	NFKFVCDNIKFADDLNQL
	ORF1a/1ab_266	IKFADDLNQLTGYKK
	ORF1a/1ab_267	DLNQLTGYKKPASRELKV
	ORF1a/1ab_268	KKPASRELKVTFFPDL
	ORF1a/1ab_269	ELKVTFFPDLNGDVVAI
	ORF1a/1ab_270	PDLNGDVVAIDYKHY
	ORF1a/1ab_271	DVVAIDYKHYTPSFKK
	ORF1a/1ab_272	YKHYTPSFKKGAKLLHK
	ORF1a/1ab_273	FKKGAKLLHKPIVWHV
	ORF1a/1ab_274	LLHKPIVWHVNNATNK
	ORF1a/1ab_275	VWHVNNATNKATYKPNTW
	ORF1a/1ab_276	NKATYKPNTWCIRCLW
	ORF1a/1ab_277	PNTWCIRCLWSTKPV
	ORF1a/1ab_278	IRCLWSTKPVETSNSFDV
	ORF1a/1ab_279	PVETSNSFDVLKSEDAQG
	ORF1a/1ab_280	DVLKSEDAQGMDNLACED
	ORF1a/1ab_281	QGMDNLACEDLKPVSEEV
	ORF1a/1ab_282	EDLKPVSEEVVENPTIQK
	ORF1a/1ab_283	EVVENPTIQKDVLECNVK
	ORF1a/1ab_284	QKDVLECNVKTTEVVGDI
	ORF1a/1ab_285	VKTTEVVGDIILKPANNS
	ORF1a/1ab_286	DIILKPANNSLKITEEV
	ORF1a/1ab_287	NNSLKITEEVGHTDLM
	ORF1a/1ab_288	TEEVGHTDLMAAYVDNSS
	ORF1a/1ab_289	LMAAYVDNSSLTIKK
	ORF1a/1ab_290	VDNSSLTIKKPNELSRVL
	ORF1a/1ab_291	KKPNELSRVLGLKTL
	ORF1a/1ab_292	LSRVLGLKTLATHGLAAV
	ORF1a/1ab_293	TLATHGLAAVNSVPWDTI
	ORF1a/1ab_294	AVNSVPWDTIANYAKPFL
	ORF1a/1ab_295	TIANYAKPFLNKVVSTTT
	ORF1a/1ab_296	FLNKVVSTTTNIVTRCL
	ORF1a/1ab_297	TTTNIVTRCLNRVCTNYM
	ORF1a/1ab_298	CLNRVCTNYMPYFFTLLL
	ORF1a/1ab_299	YMPYFFTLLLQLCTFTR
	ORF1a/1ab_300	LLLQLCTFTRSTNSRIK
	ORF1a/1ab_301	FTRSTNSRIKASMPTTI
	ORF1a/1ab_302	RIKASMPTTIAKNTVKSV
	ORF1a/1ab_303	TIAKNTVKSVGKFCL
	ORF1a/1ab_304	TVKSVGKFCLEASFNYLK
Peptide Pool	Peptide ID	Peptide Sequence
NSP3B continued	ORF1a/1ab_305	CLEASFNYLKSPNFSKLI
	ORF1a/1ab_306	LKSPNFSKLINIIIWFLL

Peptide Pool	Peptide ID	Peptide Sequence
ORF3	ORF3a_1	MDLFMRIFTIGTVTLK
	ORF3a_2	IFTIGTVTLKQGEIK
	ORF3a_3	TVTLKQGEIKDATPSDFV
	ORF3a_4	IKDATPSDFVRATATIPI
	ORF3a_5	FVRATATIPIQASLPFGW
	ORF3a_6	PIQASLPFGWLIVGVALL
	ORF3a_7	GWLIVGVALLAVFQSASK
	ORF3a_8	LLAVFQSASKIITLKKRW
	ORF3a_9	SKIITLKKRWQLALSKGV
	ORF3a_10	RWQLALSKGVHFVCNLLL
	ORF3a_11	GVHFVCNLLLLFVTVY
	ORF3a_12	NLLLLFVTVYSHLLLV
	ORF3a_13	VTVYSHLLLVAAGLEAPF
	ORF3a_14	LVAAGLEAPFLYLYALVY
	ORF3a_15	PFLYLYALVYFLQSINFV
	ORF3a_16	VYFLQSINFVRIIMRLWL
	ORF3a_17	FVRIIMRLWLCWKCRSK
	ORF3a_18	LWLCWKCRSKNPLLY
	ORF3a_19	KCRSKNPLLYDANYFLCW
	ORF3a_20	LYDANYFLCWHTNCYDY
	ORF3a_21	LCWHTNCYDYCIPYNSV
	ORF3a_22	YDYCIPYNSVTSSIVI
	ORF3a_23	YNSVTSSIVITSGDGTTS
	ORF3a_24	VITSGDGTTSPISEHDY
	ORF3a_25	TTSPISEHDYQIGGYTEK
	ORF3a_26	DYQIGGYTEKWESGVK
	ORF3a_27	YTEKWESGVKDCVVLHSY
	ORF3a_28	VKDCVVLHSYFTSDYYQL
	ORF3a_29	SYFTSDYYQLYSTQL
	ORF3a_30	DYYQLYSTQLSTDTGV
	ORF3a_31	STQLSTDTGVEHVTFFIY
	ORF3a_32	GVEHVTFFIYNKIVDEPE
	ORF3a_33	IYNKIVDEPEEHVQIHTI
	ORF3a_34	PEEHVQIHTIDGSSGVV
	ORF3a_35	HTIDGSSGVVNPVMEPIY
	ORF3a_36	VVNPVMEPIYDEPTTTTS
	ORF3a_37	PVMEPIYDEPTTTTSVPL

Peptide Pool	Peptide ID	Peptide Sequence
	M(ORF5)_1	MADSNGTITVEELKKLL
	M(ORF5)_2	ITVEELKKLLEQWNLVI
	M(ORF5)_3	KLLEQWNLVIGFLFLTWI
	M(ORF5)_4	VIGFLFLTWICLLQFAY
	M(ORF5)_5	TWICLLQFAYANRNRFLY

M	M(ORF5)_6	AYANRNRFLYIIKLIFLW
	M(ORF5)_7	LYIIKLIFLWLLWPVTL
	M(ORF5)_8	FLWLLWPVTLACFVLAAV
	M(ORF5)_9	TLACFVLAAVYRINWI
	M(ORF5)_10	LAAVYRINWITGGIAIAM
	M(ORF5)_11	WITGGIAIAMACLVGLMW
	M(ORF5)_12	AMACLVGLMWLSYFIASF
	M(ORF5)_13	MWLSYFIASFRLFARTR
	M(ORF5)_14	ASFRLFARTRSMWSF
	M(ORF5)_15	FARTRSMWSFNPETNILL
	M(ORF5)_16	SFNPETNILLNVPLHGTI
	M(ORF5)_17	LLNVPLHGTILTRPLL
	M(ORF5)_18	HGTILTRPLLESELVI
	M(ORF5)_19	RPLLESELVIGAVILR
	M(ORF5)_20	ELVIGAVILRGHLRI
	M(ORF5)_21	AVILRGHLRIAGHHLGR
	M(ORF5)_22	LRIAGHHLGRCDIKDLPK
	M(ORF5)_23	GRCDIKDLPKEITVATSR
	M(ORF5)_24	PKEITVATSRTLSYYKL
	M(ORF5)_25	TSRTLSYYKLGASQRV
	M(ORF5)_26	YYKLGASQRVAGDSGF
	M(ORF5)_27	SQRVAGDSGFAAYSRYRI
	M(ORF5)_28	GFAAYSRYRIGNYKL
	M(ORF5)_29	SRYRIGNYKLNTDHSSSS
	M(ORF5)_30	KLNTDHSSSSDNIALLV
	M(ORF5)_31	KLNTDHSSSSDNIALLVQ

619

Peptide Pool	Peptide ID	Peptide Sequence
ORF8	ORF8_1	MKFLVFLGIITTVAAF
	ORF8_2	LGIITTVAAFHQECSL
	ORF8_3	VAAFHQECSLQSCTQHQP
	ORF8_4	SLQSCTQHQPYVVDDPCP
	ORF8_5	QPYVVDDPCPIHFYSKWY
	ORF8_6	CPIHFYSKWYIRVGARK
	ORF8_7	KWYIRVGARKSAPLIEL
	ORF8_8	ARKSAPLIELCVDEAGSK
	ORF8_9	ELCVDEAGSKSPIQYIDI
	ORF8_10	SKSPIQYIDIGNYTVSCL
	ORF8_11	DIGNYTVSCLPFTINCQE
	ORF8_12	CLPFTINCQEPKLGSLVV
	ORF8_13	QEPKLGSLVVRCSFYEDF
	ORF8_14	VVRCSFYEDFLEYHDVRV
	ORF8_15	FYEDFLEYHDVRVVLDFI
Peptide Pool	Peptide ID	Peptide Sequence
	N(ORF9)_1	MSDNGPQNQRNAPRITF
	N(ORF9)_2	NQRNAPRITFGGPSDSTG
	N(ORF9)_3	TFGGPSDSTGSNQNGER
	N(ORF9)_4	STGSNQNGERSGARSKQR
	N(ORF9)_5	ERSGARSKQRRPQGL
	N(ORF9)_6	RSKQRRPQGLPNNTASWF
	N(ORF9)_7	GLPNNTASWFTALTQHGK

NP	N(ORF9)_8	WFTALTQHGKEDLKFPR
	N(ORF9)_9	HGKEDLKFPRGQGVPI
	N(ORF9)_10	KFPRGQGVPINTNSSPDD
	N(ORF9)_11	PINTNSSPDDQIGYYRR
	N(ORF9)_12	PDDQIGYYRRATRRIR
	N(ORF9)_13	YYRRATRRIRGGDGKMK
	N(ORF9)_14	RIRGGDGKMKDLSPRWYF
	N(ORF9)_15	MKDLSPRWYFYYLGTGPE
	N(ORF9)_16	YFYYLGTGPEAGLPY
	N(ORF9)_17	GTGPEAGLPYGANKDGII
	N(ORF9)_18	PYGANKDGIIWVATEGAL
	N(ORF9)_19	IIWVATEGALNTPKDHI
	N(ORF9)_20	GALNTPKDHIGTRNPANN
	N(ORF9)_21	HIGTRNPANNAAIVLQL
	N(ORF9)_22	ANNAAIVLQLPQGTTLPK
	N(ORF9)_23	QLPQGTTLPKGFYAEGSR
	N(ORF9)_24	PKGFYAEGSRGGSQASSR
	N(ORF9)_25	SRGGSQASSRSSSRSR
	N(ORF9)_26	ASSRSSSRSRNSSRNSTP
	N(ORF9)_27	SRNSSRNSTPGSSRGTSP
	N(ORF9)_28	TPGSSRGTSPARMAGNGG
	N(ORF9)_29	SPARMAGNGGDAALALLL
	N(ORF9)_30	GGDAALALLLLDRLNQL
	N(ORF9)_31	LLLLDRLNQLESKMSGK
	N(ORF9)_32	NQLESKMSGKGQQQQGQT
	N(ORF9)_33	GKGQQQQGQTVTKKSAAE
	N(ORF9)_34	QTVTKKSAAEASKKPRQK
	N(ORF9)_35	AEASKKPRQKRTATKAY
	N(ORF9)_36	RQKRTATKAYNVTQAFGR
	N(ORF9)_37	AYNVTQAFGRRGPEQTQG
	N(ORF9)_38	GRRGPEQTQGNFGDQELI
	N(ORF9)_39	QGNFGDQELIRQGTDYK
	N(ORF9)_40	ELIRQGTDYKHWPQIAQF
	N(ORF9)_41	YKHWPQIAQFAPSASAFF
	N(ORF9)_42	QFAPSASAFFGMSRIGM
	N(ORF9)_43	AFFGMSRIGMEVTPSGTW
	N(ORF9)_44	GMEVTPSGTWLTYTGAIK
	N(ORF9)_45	TWLTYTGAIKLDDKDPNF
	N(ORF9)_46	IKLDDKDPNFKDQVILL
	N(ORF9)_47	PNFKDQVILLNKHIDAYK
	N(ORF9)_48	LLNKHIDAYKTFPPTEPK
	N(ORF9)_49	YKTFPPTEPKKDKKKK
Peptide Pool	Peptide ID	Peptide Sequence
NP continued	N(ORF9)_50	TEPKKDKKKKADETQAL
	N(ORF9)_51	KKKADETQALPQRQKK
	N(ORF9)_52	TQALPQRQKKQQTVTLL
	N(ORF9)_53	QKKQQTVTLLPAADLDDF
	N(ORF9)_54	LLPAADLDDFSKQLQQSM
	N(ORF9)_55	DFSKQLQQSMSSADSTQA

620

621

Peptide Pool	Peptide ID	Peptide Sequence
S1	S_1	MFVFLVLLPLVSSQCVNL
	S_2	PLVSSQCVNLTTRTQL
	S_3	CVNLTTRTQLPPAYTNSF
	S_4	QLPPAYTNSFTRGVYY
	S_5	TNSFTRGVYYPDKVFR
	S_6	GVYYPDKVFRSSVLHSTQ
	S_7	FRSSVLHSTQDLFLPFF
	S_8	STQDLFLPFFSNVTWF
	S_9	LPFFSNVTWFHAIHV
	S_10	NVTWFHAIHVSGTNGTKR
	S_11	HVSGTNGTKRFDNPVLPF
	S_12	KRFDNPVLPFNDGVYF
	S_13	VLPFNDGVYFASTEKSNI
	S_14	YFASTEKSNIIRGWIF
	S_15	KSNIIRGWIFGTTLDSK
	S_16	WIFGTTLDSKTQSLLIV
	S_17	DSKTQSLLIVNNATNVVI
	S_18	IVNNATNVVIKVCEFQF
	S_19	VVIKVCEFQFCNDPFLGV
	S_20	QFCNDPFLGVYYHKNNK
	S_21	LGVYYHKNNKSWMESEFR
	S_22	NKSWMESEFRVYSSANNC
	S_23	FRVYSSANNCTFEYV
	S_24	SANNCTFEYVSQPFLMDL
	S_25	YVSQPFLMDLEGKQGNFK
	S_26	DLEGKQGNFKNLREFVFK
	S_27	FKNLREFVFKNIDGYFKI
	S_28	FKNIDGYFKIYSKHTPI
	S_29	FKIYSKHTPINLVRDL
	S_30	HTPINLVRDLPQGFSAL
	S_31	RDLPQGFSALEPLVDLPI
	S_32	ALEPLVDLPIGINITRF
	S_33	LPIGINITRFQTLLALHR
	S_34	RFQTLLALHRSYLTPGDS
	S_35	HRSYLTPGDSSSGWTAGA
	S_36	DSSSGWTAGAAAYYVGYL
	S_37	GAAAYYVGYLQPRTFLLK
	S_38	YLQPRTFLLKYNENGTI
Peptide Pool	Peptide ID	Peptide Sequence
	S_39	LLKYNENGTITDAVDCAL
	S_40	TITDAVDCALDPLSETK
	S_41	CALDPLSETKCTLKSFTV
	S_42	TKCTLKSFTVEKGIY
	S_43	KSFTVEKGIYQTSNFRV
	S_44	GIYQTSNFRVQPTESIVR
	S_45	RVQPTESIVRFPNITNL
	S_46	IVRFPNITNLCPFGEVF
	S_47	TNLCPFGEVFNATRFASV
	S_48	VFNATRFASVYAWNRKRI
	S_49	SVYAWNRKRISNCVADY
	S_50	KRISNCVADYSVLYNSAS

S1 continued	S_51	DYSVLYNSASFSTFKCY
	S_52	SASFSTFKCYGVSPTKL
	S_53	KCYGVSPTKLNDLCFTNV
	S_54	KLNDLCFTNVYADSFVIR
	S_55	NVYADSFVIRGDEVRQI
	S_56	VIRGDEVRQIAPGQTGKI
	S_57	QIAPGQTGKIADYNYKL
	S_58	GKIADYNYKLPDDFTGCV
	S_59	KLPDDFTGCVIAWNSNNL
	S_60	CVIAWNSNNLDSKVGGNY
	S_61	NLDSKVGGNYNYLYRLFR
	S_62	NYNYLYRLFRKSNLKPF
	S_63	LFRKSNLKPFERDISTEI
	S_64	PFERDISTEIYQAGSTPC
	S_65	EIYQAGSTPCNGVEGF
	S_66	STPCNGVEGFNCYFPL
	S_67	VEGFNCYFPLQSYGF
	S_68	CYFPLQSYGFQPTNGVGY
	S_69	GFQPTNGVGYQPYRVVVL
	S_70	GYQPYRVVVLSFELL
	S_71	RVVVLSFELLHAPATV
	S_72	FELLHAPATVCGPKK
	S_73	APATVCGPKKSTNLVKNK
	S_74	KKSTNLVKNKCVNFNF
	S_75	VKNKCVNFNFNGLTGTGV
	S_76	NFNGLTGTGVLTESNKKF
	S_77	GVLTESNKKFLPFQQFGR
	S_78	KFLPFQQFGRDIADTTDA
	S_79	GRDIADTTDAVRDPQTL
	S_80	TDAVRDPQTLEILDI
	S_81	DPQTLEILDITPCSFGGV
	S_82	DITPCSFGGVSVITPGTN
	S_83	GVSVITPGTNTSNQVAVL
	S_84	TNTSNQVAVLYQDVNCTE
	S_85	VLYQDVNCTEVPVAI
	S_86	VNCTEVPVAIHADQL
	S_87	VPVAIHADQLTPTWRVY

Peptide Pool	Peptide ID	Peptide Sequence
S1 continued	S_88	DQLTPTWRVYSTGSNVF
	S_89	RVYSTGSNVFQTRAGCLI
	S_90	VFQTRAGCLIGAEHV
	S_91	AGCLIGAEHVNNSYECDI
	S_92	HVNNSYECDIPIGAGI
	S_93	ECDIPIGAGICASYQTQT

622

Peptide Pool	Peptide ID	Peptide Sequence
	S_94	GICASYQTQTNSPRRAR
	S_95	TQTNSPRRARSVASQSII
	S_96	ARSVASQSIIAYTMSL
	S_97	QSIIAYTMSLGAENSVAY
	S_98	SLGAENSVAYSNNSIAI

S2	S_99	VAYSNNSIAIPTNFTISV
	S_100	AIPTNFTISVTTEILPV
	S_101	ISVTTEILPVSMTKTSV
	S_102	LPVSMTKTSVDCTMYI
	S_103	KTSVDCTMYICGDSTECS
	S_104	YICGDSTECSNLLLQY
	S_105	TECSNLLLQYGSFCTQL
	S_106	LQYGSFCTQLNRALTGI
	S_107	TQLNRALTGIAVEQDK
	S_108	LTGIAVEQDKNTQEVF
	S_109	EQDKNTQEVFAQVKQIYK
	S_110	VFAQVKQIYKTPPIKDF
	S_111	IYKTPPIKDFGGFNFSQI
	S_112	DFGGFNFSQILPDPSK
	S_113	FSQILPDPSKPSKRSFI
	S_114	PSKPSKRSFIEDLLFNKV
	S_115	FIEDLLFNKVTLADAGFI
	S_116	KVTLADAGFIKQYGDCL
	S_117	GFIKQYGDCLGDIAARDL
	S_118	CLGDIAARDLICAQKF
	S_119	ARDLICAQKFNGLTVL
	S_120	AQKFNGLTVLPPLLTDEM
	S_121	VLPPLLTDEMIAQYTSAL
	S_122	EMIAQYTSALLAGTI
	S_123	YTSALLAGTITSGWTF
	S_124	AGTITSGWTFGAGAALQI
	S_125	TFGAGAALQIPFAMQMAY
	S_126	QIPFAMQMAYRFNGIGV
	S_127	MAYRFNGIGVTQNVLY
	S_128	GIGVTQNVLYENQKLI
	S_129	NVLYENQKLIANQFNSAI
	S_130	LIANQFNSAIGKIQDSL
	S_131	SAIGKIQDSLSSTASAL
	S_132	DSLSSTASALGKLQDVV
	S_133	SALGKLQDVVNQNAQAL
Peptide Pool	Peptide ID	Peptide Sequence
	S_134	DVVNQNAQALNTLVKQL
	S_135	QALNTLVKQLSSNFGAI
	S_136	KQLSSNFGAISSVLNDIL
	S_137	AISSVLNDILSRLDKV
	S_138	NDILSRLDKVEAEVQIDR
	S_139	KVEAEVQIDRLITGRL
	S_140	QIDRLITGRLQSLQTYV
	S_141	GRLQSLQTYVTQQLIR
	S_142	QTYVTQQLIRAAEIR
	S_143	QQLIRAAEIRASANL
	S_144	AAEIRASANLAATKM
	S_145	ASANLAATKMSECVL
	S_146	AATKMSECVLGQSKRVDF
	S_147	VLGQSKRVDFCGKGYHLM
	S_148	DFCGKGYHLMSFPQSAPH
	S_149	LMSFPQSAPHGVVFLHV

S2 continued	S_150	APHGVVFLHVTYVPAQEK
	S_151	HVTYVPAQEKNFTTAPAI
	S_152	EKNFTTAPAICHDGKAHF
	S_153	AICHDGKAHFPREGVFV
	S_154	AHFPREGVFVSNGTHWFV
	S_155	FVSNGTHWFVTQRNFY
	S_156	HWFVTQRNFYEPQII
	S_157	QRNFYEPQIITTDNTFV
	S_158	QIITTDNTFVSGNCDVVI
	S_159	FVSGNCDVVIGIVNNTVY
	S_160	VIGIVNNTVYDPLQPEL
	S_161	TVYDPLQPELDSFKEEL
	S_162	PELDSFKEELDKYFK
	S_163	FKEELDKYFKNHTSPDV
	S_164	YFKNHTSPDVDLGDISGI
	S_165	DVDLGDISGINASVVNI
	S_166	SGINASVVNIQKEIDRL
	S_167	VNIQKEIDRLNEVAKNL
	S_168	DRLNEVAKNLNESLIDL
	S_169	KNLNESLIDLQELGKY
	S_170	LIDLQELGKYEQYIKWPW
	S_171	KYEQYIKWPWYIWLGFI
	S_172	WPWYIWLGFIAGLIAIVM
	S_173	FIAGLIAIVMVTIMLCCM
	S_174	VMVTIMLCCMTSCCSCLK
	S_175	CMTSCCSCLKGCCSCGSC
	S_176	LKGCCSCGSCCKFDEDDS
	S_177	SCCKFDEDDSEPVLKGVK
	S_178	FDEDDSEPVLKGVKLHYT

623

624