Lawrence Berkeley National Laboratory
Climate & Ecosystems

Title

Long-term perturbation of the peripheral immune system months after SARS-CoV-2 infection

Permalink

https://escholarship.org/uc/item/5tt880cb

Journal

BMC Medicine, 20(1)

ISSN

1741-7015

Authors

Ryan, Feargal J
Hope, Christopher M
Masavuli, Makutiro G
et al.

Publication Date

2022-12-01

DOI

10.1186/s12916-021-02228-6

Copyright Information

Peer reviewed

BMC Medicine

RESEARCH ARTICLE   **Open Access**

Long-term perturbation of the peripheral immune system months after SARS-CoV-2 infection

Feargal J. Ryan[1†], Christopher M. Hope[2,3†], Makutiro G. Masavuli[4†], Miriam A. Lynn[1†], Zelalem A. Mekonnen[4], Arthur Eng Lip Yeow[4], Pablo Garcia-Valtanen[4], Zahraa Al-Delfi[4], Jason Gummow[5], Catherine Ferguson[6], Stephanie O'Connor[7], Benjamin A. J. Reddi[7], Pravin Hissaria[6], David Shaw[6], Chuan Kok-Lim[6,8], Jonathan M. Gleadle[9,10], Michael R. Beard[11], Simon C. Barry[2,3*†], Branka Grubor-Bauk[4*†] and David J. Lynn[1,10*†]

Abstract

Background: Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a highly infectious respiratory virus which is responsible for the coronavirus disease 2019 (COVID-19) pandemic. It is increasingly clear that recovered individuals, even those who had mild COVID-19, can suffer from persistent symptoms for many months after infection, a condition referred to as "long COVID", post-acute sequelae of COVID-19 (PASC), post-acute COVID-19 syndrome, or post COVID-19 condition. However, despite the plethora of research on COVID-19, relatively little is known about the molecular underpinnings of these long-term effects.

Methods: We have undertaken an integrated analysis of immune responses in blood at a transcriptional, cellular, and serological level at 12, 16, and 24 weeks post-infection (wpi) in 69 patients recovering from mild, moderate, severe, or critical COVID-19 in comparison to healthy uninfected controls. Twenty-one of these patients were referred to a long COVID clinic and > 50% reported ongoing symptoms more than 6 months post-infection.

---

* Correspondence: simon.barry@adelaide.edu.au;
branka.grubor@adelaide.edu.au; david.lynn@sahmri.com
†Feargal J. Ryan, Christopher M. Hope, Makutiro G. Masavuli, Miriam A. Lynn
contributed equally as first authors. Simon C. Barry, Branka Grubor-Bauk and
David J. Lynn contributed equally as senior authors.
²Women's and Children's Health Network, North Adelaide, SA, Australia
⁴Viral Immunology Group, Adelaide Medical School, University of Adelaide
and Basil Hetzel Institute for Translational Health Research, Adelaide, SA,
Australia
¹Precision Medicine Theme, South Australian Health and Medical Research
Institute, Adelaide, SA 5001, Australia
Full list of author information is available at the end of the article

Results: Anti-Spike and anti-RBD IgG responses were largely stable up to 24 wpi and correlated with disease severity. Deep immunophenotyping revealed significant differences in multiple innate (NK cells, LD neutrophils, CXCR3+ monocytes) and adaptive immune populations (T helper, T follicular helper, and regulatory T cells) in convalescent individuals compared to healthy controls, which were most strongly evident at 12 and 16 wpi. RNA sequencing revealed significant perturbations to gene expression in COVID-19 convalescents until at least 6 months post-infection. We also uncovered significant differences in the transcriptome at 24 wpi of convalescents who were referred to a long COVID clinic compared to those who were not.

Conclusions: Variation in the rate of recovery from infection at a cellular and transcriptional level may explain the persistence of symptoms associated with long COVID in some individuals.

Keywords: SARS-CoV-2, COVID-19, Immunity, RNA-Seq, T cell, Antibody responses, Convalescent patients, Immunophenotyping, Long COVID, Post-acute sequelae of COVID-19 (PASC), Post COVID-19 condition, Infection

Background

Coronavirus disease 2019 (COVID-19) is caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), a highly infectious respiratory virus responsible for the ongoing global pandemic. COVID-19 usually presents as an asymptomatic or mild to moderate respiratory infection in previously healthy individuals with symptoms that include fever, cough, headache, fatigue, myalgia, diarrhoea, and anosmia [1, 2]. However, in older individuals or in those with prior co-morbidities such as obesity or cardiovascular disease, COVID-19 can quickly develop into a severe and life-threatening disease requiring urgent intensive care support. While the death toll from COVID-19 has been devastating (> 4.8 million as of 5 October 2021 according to the Johns Hopkins University Coronavirus Resource Center [3]), the vast majority of those infected fortunately do recover, with case fatality rates in most countries falling below 3%. It is now increasingly clear, however, that recovered individuals, even those who had mild COVID-19, can suffer from persistent symptoms for many months after infection [4], which is commonly referred to as long COVID. For example, a cohort study of COVID-19 patients (median age 57) discharged from hospital in Wuhan, China, 6 months prior, reported that 63% of patients presented with fatigue or muscle weakness, 23% sleep difficulties, and 23% anxiety or depression [5]. Individuals who were previously severely ill during their hospital stay have ongoing impaired pulmonary function and abnormal chest imaging. Similar reports continue to pour in from around the world [6–11]. While the majority of these reports involve patients who were hospitalised with COVID-19, persistent, albeit milder and less-frequent, symptoms have also been reported in non-hospitalised individuals months after recovery [12]. These reports resemble similar post-infectious syndromes after other infections, such as Ebola [13] and SARS-CoV-1 [14], and suggest that there may be a long-lasting dysregulation of the immune response in individuals recovering from COVID-19.

Flow cytometric analysis of peripheral blood samples collected from convalescents in the USA (median 29 days post-infection) has revealed altered frequencies of innate and adaptive immune cell populations including CD4$^+$ and CD8$^+$ T cell activation and exhaustion marker expression in recovered individuals [15]. A similar study in Singapore (median 34 days post-infection) found increased levels of circulating endothelial cells and effector T cells in those recovering from active disease [16]. Single-cell RNA sequencing (scRNA-Seq) of peripheral blood mononuclear cells (PMBC) from a small (*n* = 10) cohort of patients that were 7–14 days post-recovery also found an increased ratio of classical CD14$^+$ monocytes with high inflammatory gene expression, decreased CD4$^+$ and CD8$^+$ T cells, and significantly increased plasma B cells [17]. scRNA-Seq profiling of PBMC gene expression in a larger cohort of recovering individuals (*n* = 95) found those with severe disease (*n* = 36) had decreased plasmacytoid dendritic cells (pDCs) and increased levels of proliferative effector memory CD8$^+$ T cells, relative to healthy controls [18]. A potential limitation of this study, however, was that samples from recovered individuals were not collected at uniform timepoints during recovery, instead samples were collected between 9 and 126 days post-infection (on average 44.5 days). Longitudinal profiling of the transcriptome of PBMC collected from individuals (*n* = 18) during treatment, convalescence, and recovery phases of infection (up to 10 weeks post-infection) revealed that relative to acute disease, recovery from COVID-19 was marked by decreased expression of genes involved in the interferon response, humoral immunity, and increased signatures indicative of *T* cell activation and differentiation [19]. However, these responses were not compared with healthy controls. Another recent study longitudinally profiled immune cell populations and the blood transcriptome in > 200 SARS-CoV-2-infected patients over 12 weeks from symptom onset to recovery [20]. They compared the

blood transcriptome in 2-time bins (0–24 and 25–48 days from symptom onset) and found substantial changes relative to uninfected controls in immune cell populations and increased expression of genes involved in immunometabolism and inflammation, which persisted after infection.

Here, we have performed anti-receptor-binding domain (RBD) and anti-Spike serology, comprehensive multi-parameter immunophenotyping, and transcriptome-wide RNA sequencing on blood collected from individuals recovering from mild/moderate or severe/critical COVID-19 at 12, 16, and 24 weeks after their first positive SARS-CoV-2 PCR test, as well as age-matched healthy controls (HCs). Our analyses reveal robust but heterogenous humoral immunity in convalescents until at least 6 months post-infection. Deep immunophenotyping highlighted profound changes in immune cell populations in COVID-19 convalescents compared with HCs, particularly at 12 and 16 weeks post-infection (wpi). Furthermore, RNA sequencing revealed significant changes in whole blood gene expression for up to 24 wpi, even in individuals that had mild disease without hospitalisation. Significant differences in gene expression were also identified at 24 wpi in convalescent individuals who were referred to a long COVID clinic compared to those who were not. These data suggest that SARS-CoV-2 infection leads to persistent changes to the peripheral immune system long after the infection is cleared, which has important potential implications for understanding symptoms associated with long COVID. These changes to the peripheral immune system could have implications for how individuals recovering from infection respond to vaccination or other challenges encountered in this period and persistent immune activation may also exacerbate other chronic conditions.

Methods
Patient recruitment
Study participants were recruited via the Central Adelaide Health Network (CALHN). The study was performed in accordance with the ethical principles consistent with the latest version of the Declaration of Helsinki (version Fortaleza 2013) and Good Clinical Practice (GCP) and according to the National Health and Medical Research Council (NHMRC) Guidelines for Research published in the National Statement on the Ethical Conduct in Human Research (2007; updated 2018). The protocol was approved by CALHN Human Research Ethics Committee, Adelaide, Australia (Approval No. 13050). Inclusion criteria were PCR-confirmed SARS-CoV-2 infection from nasopharyngeal swabs (which occurred in March & April of 2020 for all participants), the ability to attend study follow-up visits, and voluntary informed consent. The study size was determined in a pragmatic fashion by opportunistically recruiting as many participants as possible. A total of 69 COVID-19 convalescent individuals (35

male, 36 female) representing a range of prior mild, moderate, severe, and critical COVID-19 cases were recruited (Table S1). COVID-19 disease severity was scored as per NIH descriptors [21] where 5 = "asymptomatic", 4 = "mild", 3 = "moderate", 2 = "severe", and 1 = "critical" (Table S1). Blood samples were collected from convalescents at 12, 16, and 24 weeks (± 14 days) post the date of their initial PCR-positive test. Participation at each timepoint was determined by availability to attend follow-up sample collection clinics. Convalescent patients were requested to complete a retrospective questionnaire detailing self-reported symptoms related to long COVID at each of the sampling timepoints in this study (Additional file 10). The survey was administered at approximately 18 months (mean 70.4 weeks, min 61 weeks, max 74 weeks) post-infection. Additionally, we obtained clinical data indicating which convalescent individuals were referred to a long COVID clinic run by the South Australian State health service (SA Health). The long COVID clinic provides a pathway of care focused on comprehensive clinical care, psychological support, and rehabilitation to patients living with long-term sequelae of COVID-19 disease. Patients were offered referral to this clinic at their 18-month study visit, as the clinic was only established around that time. Clinical assessments from this clinic are ongoing. Healthy controls (*n* = 14) in the same ranges of age and sex as the COVID-19 convalescent cohort were also recruited. Healthy controls had no respiratory disease, no positive COVID-19 PCR test in 2020/21, no known significant systemic diseases, and negative anti-Spike and anti-RBD serology. Blood (54 ml/individual) was collected in serum separator (acid citrate dextrose (ACD)) tubes or ethylenediaminetetraacetic acid (EDTA) tubes and processed for serum, peripheral blood mononuclear cells (PBMCs), and plasma isolation. 2.5 mL of blood for RNA sequencing was collected into PAXgene® tubes (762165 BD, North Ryde, Australia) and stored at − 80 °C until processing. C-reactive protein (CRP) titres were assayed by a National Association of Testing Authorities Australia-certified commercial pathology service (SA Pathology, Adelaide, Australia).

SARS-CoV-2 PCR testing
Extraction of RNA was achieved from nasopharyngeal swabs using the Automated MagMAX nucleic acid extraction protocol (Thermo Fisher) and RNA subjected to a one-step qRT-PCR using a Roche light cycler LC408II using cycle conditions described by Corman et al. [22].

SARS-CoV-2 protein purification and ELISA
Prefusion SARS-CoV-2 ectodomain (isolate WHU1, residues1-1208) with HexaPro mutations [23] (kindly provided by Adam Wheatley) and SARS-Cov-2 receptor-

binding domain (RBD) with C-terminal His-tag [24] (residues 319-541; kindly provided by Florian Krammer) were overexpressed in Expi293 cells and purified by Ni-NTA affinity and size-exclusion chromatography. Recombinant proteins were analysed via a standard SDS-PAGE gel to check protein integrity. Gels were stained with Comassie Blue (Invitrogen) for 2 h and de-stained in distilled water overnight. MaxiSorp 96-well plates were coated overnight at 4 °C with 5 μg/mL of recombinant RBD or S proteins. After blocking with 5% w/v skim milk in 0.05% Tween-20/PBS (PBST) at room temperature, serially diluted (heat inactivated) sera were added and incubated for 2 h at room temperature. Plates were washed 4 times with 0.05% PBST and secondary antibodies added. Secondary antibodies were diluted in 5% skim milk in PBST as follows: Goat anti-Human IgG (H + L) Secondary Antibody, HRP (1:30,000; Invitrogen); Mouse Anti-Human IgG1 Fc-HRP (1:5000, Southern Biotech), Mouse Anti-Human IgG3 Hinge-HRP (1:5000; Southern Biotech); goat anti-human IgM HRP (1:5000; Sigma): anti-human IgA HRP antibody (1:5000; Sigma) and incubated for 1 hour at room temperature. Plates were developed with 1-Step™ Ultra TMB Substrate (Thermo Fisher) and stopped with 2 M sulphuric acid. OD readings were read at 450 nm on a Synergy HTX Multi-Mode Microplate Reader. AUC calculation was performed using Prism GraphPad, where the *X*-axis is half log10 of sera dilution against OD450 on *Y*-axis.

PBMC isolation

Post plasma centrifugation, the white blood cell pack was harvested, pooled into 1×50 ml falcon tube, diluted in 2% FCS/PBS up to 35 ml and overlaid onto 15 ml Ficoll, centrifuged for 20 min, $1000 \times g$, RT, no brake. The PBMC were isolated, washed in 2% FCS/PBS, centrifuged at $480 \times g$ for 10 min at RT, PBMC resuspended in 50 ml 2% FCS/PBS, manually counted using trypan blue exclusion assay. For deep immunophenotyping 2×10^6 cells were plated across 4 wells (5×10^5 per well) of a 96-well plate. The 50-ml tube was then spun at $300 \times g$ for 10 min, the pellet was resuspended in ½ volume of FCS with ½ volume of 20% DMSO/80% FCS added dropwise to final cell concentration of 1×10^7 per ml. The samples were stored 800 μL–1.8 ml per vial placed in a CoolCell at − 80 °C. The frozen PBMC tubes were transferred to liquid nitrogen for long-term storage within 1–7 days.

Flow cytometry staining

The 96-well plate was centrifuged at $300 \times g$ for 4 min, the plate was inverted on paper towel, and the PBMC pellets were stained with 30 μL of 1 of 3 master-mixes of antibodies (lineage, 15 color; memory, 8 color; T helper/Treg, 14 color) for 20 min RT, in dark, which included a co-stain of BD LIVE/DEAD fixable dye (stained at 1:

1000). The stained PBMC were washed with 200 μL of FACS wash, centrifuged $300 \times g$ for 4 min and fixed with 200 μL FACS Fix for 20 min, RT, in dark. Fixed cells were then centrifuged $300 \times g$ for 4 min, washed in 200 μL FACS wash then spun $300 \times g$ for 4 min, and resuspended in 50 μL FACS wash. The cells were resuspended and transferred to tubes before being analysed using a BD FACS Symphony within 3 days of staining/fixing.

Flow cytometry data acquisition and analysis

To control for batch effects, the BD FACS symphony lasers were calibrated with dye conjugated standards (Cytometer Set &Track beads) run every day. All samples were acquired with all 28 PMTs recording events. All PMT voltages were adjusted to unstained negative control baseline typically log scale 10^2. Antibodies were titrated for optimal signal over background so that single positive stains sat within log scale 10^3–10^5 of designated PMT. Compensation was set with beads matched to each panel antibody combination using spectral compensation using FlowJo Software V10. Exported FCS files had compensation values adjusted manually post-acquisition on a file-by-file basis in FCS express v6. Once compensated, low data quality events were excluded based upon time acquired (at the sample acquisition start and before sample exhaustion), with further time exclusion gates based on blockages or unexplained loss of events for a period of time during acquisition. Events that were highly positive for LIVE/DEAD staining were removed from subsequent analysis, to prevent exclusion of live monocytes, which take up more live/dead dye per cell than T cells, giving a high background. Events were gated for FCS-H/A as well as SSC-H/A linearity, and restricted FSC-W and SSC-W values for doublet discrimination. Live single cells were then broadcast on SSC-A / FSC-A plot to determine size and complexity. Lymphocyte, monocyte, and granulocyte gates were based on physical parameters (unless a lineage-specific antibody was added). Populations of cells were expressed as a proportion of the highest order lineage gate (namely: lymphocytes, monocytes, and granulocytes). See Additional file 11 for representative gating strategy. For parameters measured in healthy controls and COVID-19 samples at 12 weeks, 16 weeks and 24 weeks, a Wilcoxon rank sum test was used to assess statistical significance with the Benjamini and Hochberg method employed to correct for multiple comparisons. Statistical significance was determined as FDR < 0.05.

RNA extraction and library preparation

RNA extraction and genomic DNA elimination was carried out using the PAXgene® Blood RNA kit (762164,

Qiagen, Feldbachstrasse, Germany) as per the manufacturer's instructions. Final elution was done into 80 μL RNase-free water. A further RNA precipitation reaction was carried out. Briefly, RNA was resuspended 2.5 × 100% ethanol and 10% sodium acetate and spun at 12,000×*g* for 30 min at 4 °C. Samples were washed in 75% ethanol. Pellets were air dried and resuspended in 40 μL RNase-free water and total RNA yield was determined by analysis of samples using a TapeStation (Agilent) and Qubit (Thermo Fisher Scientific, Australia). Total RNA was converted to strand-specific Illumina compatible sequencing libraries using the Nugen Universal Plus Total RNA-Seq library kit from Tecan (Mannedorf, Switzerland) as per the manufacturer's instructions (MO1523 v2) using 12 cycles of PCR amplification for the final libraries. An Anydeplete probe mix targeting both human ribosomal and adult globin transcripts (HBA1, HBA2, HBB, HBD) was used to deplete these transcripts. Sequencing of the library pool (2 × 150 bp paired-end reads) was performed using 2 lanes of an S4 flowcell on an Illumina Novaseq 6000.
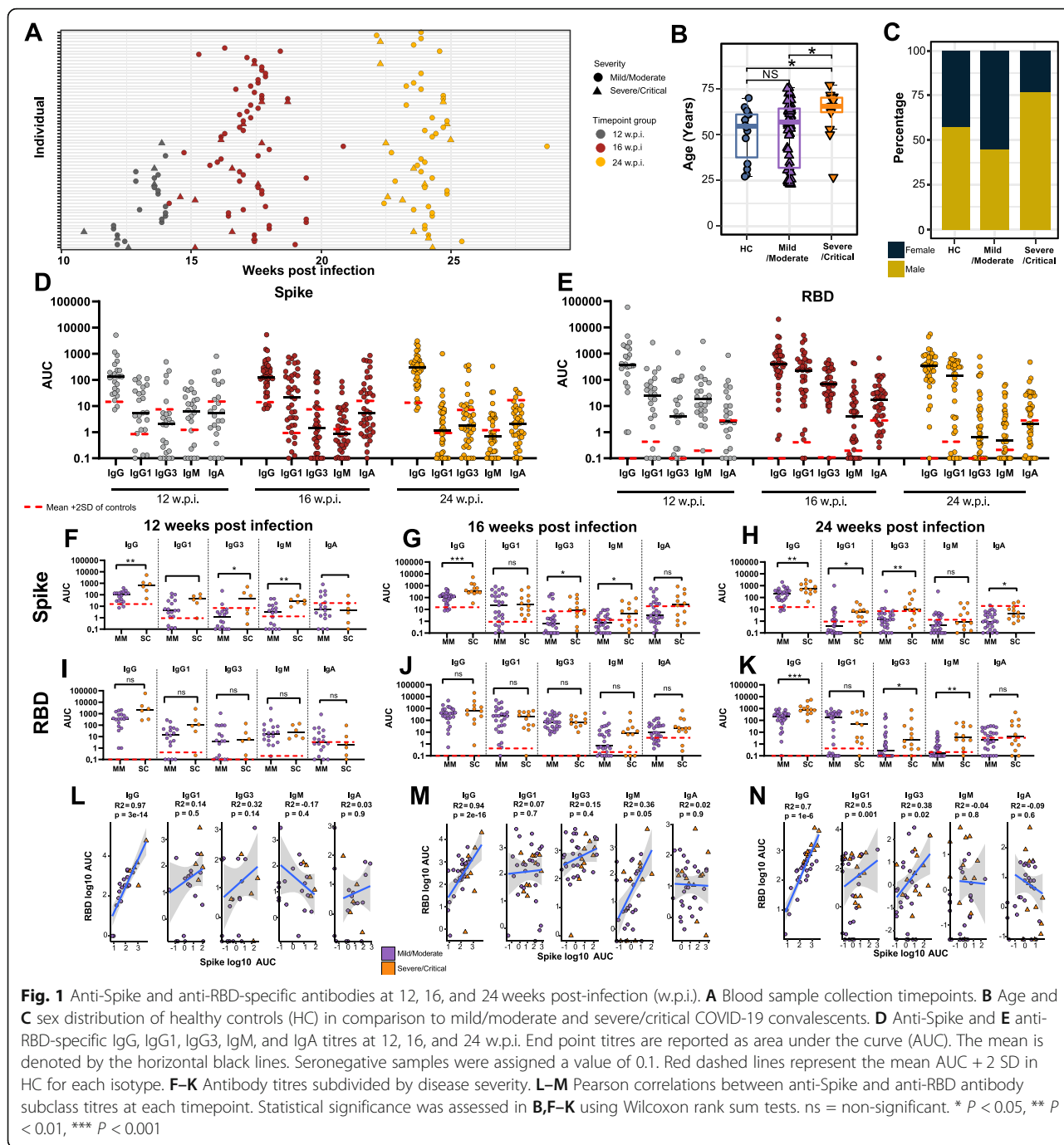
RNA-Seq analysis

Sequence read quality was assessed using FastQC version 0.11.4 [25] and summarised with MultiQC version 1.8 [26] prior to quality control with Trimmomatic version 0.38 [27] with a window size of 4 nucleotides and an average quality score of 25. Following this, reads which were < 50 nucleotides after trimming were discarded. Reads that passed all quality control steps were then aligned to the Human genome (GRCh38 assembly) using HISAT2 version 2.1.0 [28]. The gene count matrix was generated with FeatureCounts version 1.5.0-p2 [29] using the union model with Ensembl version 101 annotation. The count matrix was then imported into R version 4.0.3 for further analysis and visualisation in ggplot2 v2.3.3. Counts were normalised using the trimmed mean of M values (TMM) method in EdgeR version 3.32 and represented as counts per million (cpm) [30]. Prior to multidimensional scaling analysis and generation of heatmaps, svaseq v3.38 was applied to remove batch effects and other unwanted sources of variation in the data [31]. Differential gene expression analysis was performed using the glmLRT function in EdgeR adjusting for sex and batch (run) in the model. Genes with < 3 cpm in at least 15 samples were excluded from the differential expression analysis. Pathway and Gene Ontology (GO) overrepresentation analysis was carried out in R using a hypergeometric test. To assess if differential gene expression was primarily driven by differences in the proportion of any major immune cell population (i.e. LD granulocytes, LD neutrophils, CXCR3$^+$ neutrophils, monocytes, lymphocytes, CD56^{++} NK cells, CD19$^+$ B cells, CD3$^+$ T cells, NKT cells, CD4$^+$

T cells, or CD8$^+$ T cells), we additionally fit the frequency of each population in each individual into the EdgeR model and reperformed the differential gene expression and pathway overrepresentation analysis. Gene Set Enrichment Analysis (GSEA) was carried out using the camera function in the EdgeR library with the Molecular Signatures Database (MSigDB) R package (msigdbr v7.4.1). Blood transcriptional module (BTM) analysis was carried using a pre-defined set of modules defined by Li et al. as an alternative to pathway-based analyses [32]. Gene Set Variation Analysis (GSVA) [33] was used to calculate a per sample activity score for each of the modules (excluding unannotated modules labelled as "TBA"). *limma* v3.46.0 was used to identify modules that were differentially active. Pearson correlation analysis was performed using the Hmisc v4.4-2 package in R to determine correlations between anti-Spike and anti-RBD antibody titres, flow cytometry data, and BTM activity scores. Correlation networks were exported to Cytoscape v3.8.1 for visualisation.

Results

To assess the long-term effects of SARS-CoV-2 infection on the peripheral immune system, blood samples were collected from 69 recovering/convalescent COVID-19 individuals at 12, 16, and 24 weeks (± 14 days) postinfection (wpi) (Fig. 1A). Blood samples were also collected from *n* = 14 seronegative healthy controls (HCs) with no history of prior SARS-CoV-2 infection. COVID-19 convalescent individuals were classified according to the NIH classification of disease severity [21] as mild (*n* = 50), moderate (*n* = 6), severe (*n* = 7), or critical (*n* = 6) (Table S1). All participants were discharged from hospital prior to sample collection (58% of participants were hospitalised during acute COVID-19). Convalescents with mild illness spent a median of 1.5 days hospitalised (min = 0, max = 10), moderate illness 2 days (min = 1, max = 7), severe illness 12 days (min = 3, max = 16), and critical illness 13 days (min = 8, max = 82). HCs were age- and sex-matched with mild/moderate convalescent individuals; however, as expected based on previous reports, severe/critical convalescent individuals were older and mostly male (Fig. 1B,C). The presence of post-acute sequelae of COVID-19 (PASC) was assessed retrospectively through a symptom questionnaire (Additional file 10) which was completed by 83% (*n* = 57) of the convalescent individuals. Fatigue was the most commonly reported symptom (40.4%) followed by dyspnea (35.1%), worsened memory/concentration (35.1%), and decreased muscle strength (33.3%) (Table 1). The prevalence and type of self-reported symptoms was consistent with long COVID symptoms reported in other studies [34] and aligned with which individuals were referred to a dedicated long COVID clinic (*n* = 21, 30.4% of patients).

Ryan *et al. BMC Medicine*   (2022) 20:26

Page 6 of 23

Fig. 1 Anti-Spike and anti-RBD-specific antibodies at 12, 16, and 24 weeks post-infection (w.p.i.). **A** Blood sample collection timepoints. **B** Age and **C** sex distribution of healthy controls (HC) in comparison to mild/moderate and severe/critical COVID-19 convalescents. **D** Anti-Spike and **E** anti-RBD-specific IgG, IgG1, IgG3, IgM, and IgA titres at 12, 16, and 24 w.p.i. End point titres are reported as area under the curve (AUC). The mean is denoted by the horizontal black lines. Seronegative samples were assigned a value of 0.1. Red dashed lines represent the mean AUC + 2 SD in HC for each isotype. **F–K** Antibody titres subdivided by disease severity. **L–M** Pearson correlations between anti-Spike and anti-RBD antibody subclass titres at each timepoint. Statistical significance was assessed in **B,F–K** using Wilcoxon rank sum tests. ns = non-significant. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$

Referral to a long COVID clinic was not significantly associated with subject age (Wilcox rank sum test, $P = 0.1$), sex (Fisher's exact test, $P = 0.9$), or COVID-19 severity (Fisher's exact test, $P = 0.09$). All samples in this study were collected in South Australia where early and strict international and interstate border control measures eliminated community transmission of the virus during the sample collection period [35]. None of the participants had received a COVID-19 vaccine at the time of sample collection. This cohort was therefore uniquely placed for the assessment of immune responses in COVID-19 convalescents due a negligible risk of re-infection or changes to the immune system induced by vaccination.

COVID-19 convalescents have robust anti-Spike and anti-RBD antibody responses for at least 6 months post-infection

Anti-SARS-CoV-2 Spike and receptor-binding domain (RBD) total IgG, IgG1, IgG3, IgM, and IgA responses

Table 1. The proportion of patients self-reporting long COVID-associated symptoms
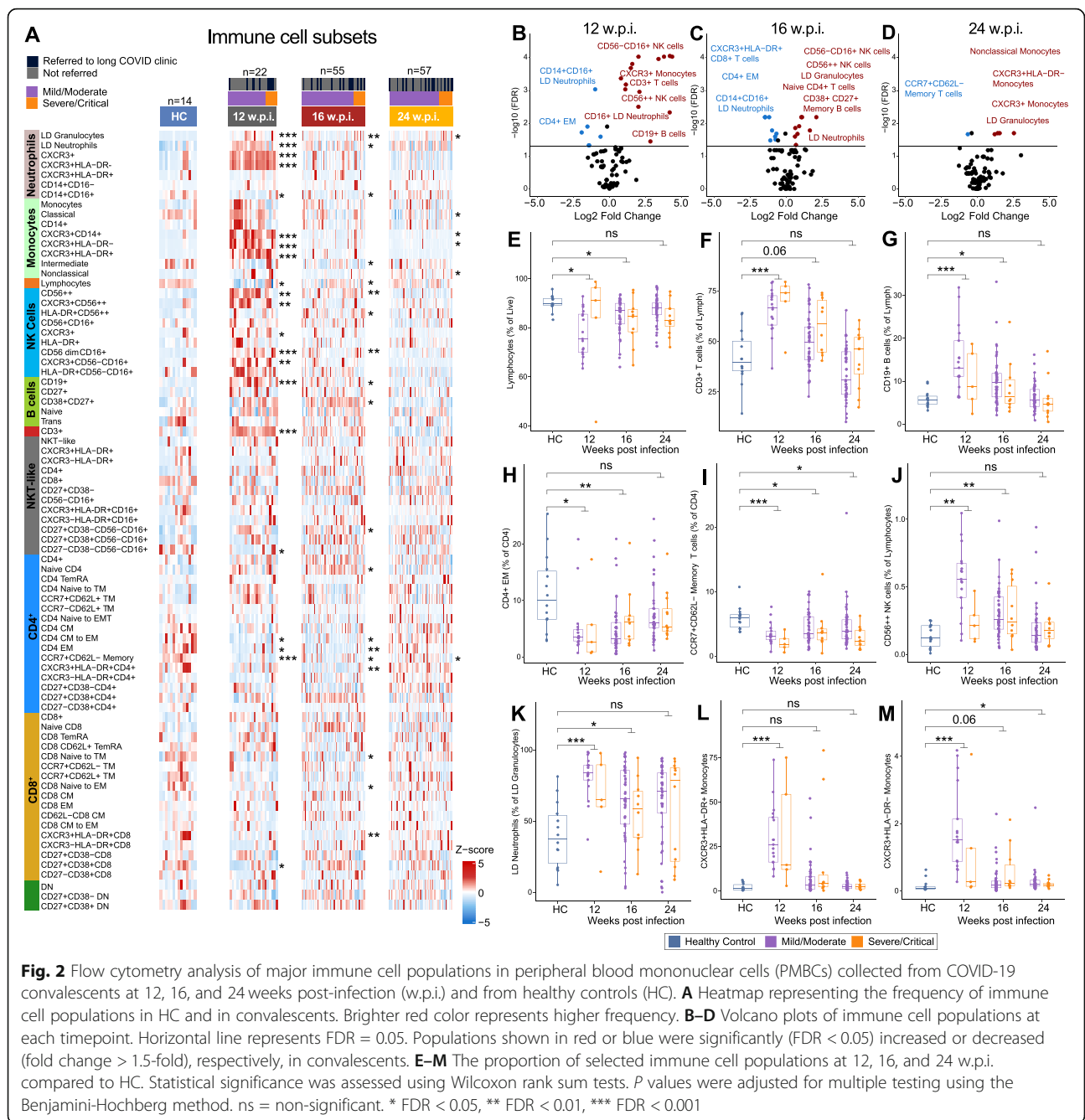
		Number of Subjects	Percent of subjects
General	Referred to long COVID clinic	21	30.4
	Any reported symptoms?	33	57.9
Non-specific	Fatigue	23	40.4
	Dyspnoea	20	35.1
	Decreased muscle strength	19	33.3
	Worsened sleep quality	16	28.1
	Palpitations	4	7.0
	Dizziness	4	7.0
	Hair loss	6	10.5
Neuropsychiatric	Anxiety/depression	14	24.6
	Worsened memory/concentration	20	35.1
Pain	Headache	10	17.5
	Chest pain	7	12.3
	Joint pain	6	10.5
	Myalgia	6	10.5
Oral/pharyngeal	Worsened sense of smell/taste	11	19.3
	Difficulty swallowing /sore throat	5	8.8
Gastrointestinal	Diarrhoea	4	7.0
	Nausea	3	5.3
	Vomiting	1	1.8

were evaluated in convalescent individuals at 12, 16, and 24 wpi (Fig. 1D,E, Table S1). The titres of Spike-specific IgG were diverse but largely stable over time (Fig. S1A-C), although there was a trend for anti-Spike IgG1 titres to decline over time (Fig. S1B). The seropositivity of Spike-specific serum IgM and IgA gradually diminished over time (Fig. S1D-E). Overall, the kinetics for anti-RBD antibodies were similar to those observed for anti-Spike antibodies (Fig. S1F-J), though anti-RBD IgG3 and IgM appeared to decline more rapidly than anti-Spike antibodies. We also compared the levels of anti-Spike and anti-RBD circulating antibodies between individuals recovering from mild/moderate versus severe/critical COVID-19. Anti-Spike total IgG and IgG3 levels at 12, 16, and 24 wpi were significantly higher in severe/critical convalescents compared to those with previous mild/moderate disease (Fig. 1F–H). Anti-Spike IgG1 and IgA levels were significantly higher in severe/critical convalescents at 24 wpi only (Fig. 1H). There was no detectable difference in RBD-specific antibody responses between individuals recovering from mild/moderate or severe/critical disease at 12 or 16 wpi (Fig. 1I,J). At 24 wpi, severe/critical convalescent individuals maintained significantly higher anti-RBD IgG, IgG3, and IgM levels compared to individuals with previous mild/moderate COVID-19 disease (Fig. 1K). Anti-Spike and anti-RBD total IgG levels (but not other antibody subclasses) were

significantly correlated at all timepoints (Fig. 1L–N). Anti-Spike and anti-RBD total IgG1 and IgG3 levels were significantly correlated at 24 wpi only. In summary, anti-Spike and anti-RBD antibody titres were generally positively correlated with COVID-19 disease severity, in accordance with previous observations [36–38].

Deep immunophenotyping reveals persistent alterations in immune cell populations in COVID-19 convalescents up to 24 weeks post-infection

We used a multi-parameter flow cytometry approach to identify and enumerate ∼ 130 different immune cell subpopulations in PBMC collected from COVID-19 convalescent individuals at 12, 16, and 24 wpi and from HCs (Table S2; Additional file 11). Flow cytometry was performed on PBMC rather than whole blood to enrich for rarer immune cell populations and to facilitate comparison with the majority of other studies assessing immune phenotypes in COVID-19 patients or convalescents. Our analysis included deep immunophenotyping of the CD4 and CD8 compartments, interrogating their maturation status, and in the CD4 compartment, interrogation of T helper (Th) lineage subsets, T regulatory (Treg) subsets, and T follicular helper (Tfh) subsets using a combination of chemokine receptor expression patterns to resolve Th-lineages (Th1, 2, 17, 1/7, 9, 22, 2/22). Immune cell populations were first categorised into 10 major lineages (Fig. 2A). Each cell type was further

Ryan *et al. BMC Medicine*   (2022) 20:26

Page 8 of 23

Fig. 2 Flow cytometry analysis of major immune cell populations in peripheral blood mononuclear cells (PMBCs) collected from COVID-19 convalescents at 12, 16, and 24 weeks post-infection (w.p.i.) and from healthy controls (HC). **A** Heatmap representing the frequency of immune cell populations in HC and in convalescents. Brighter red color represents higher frequency. **B–D** Volcano plots of immune cell populations at each timepoint. Horizontal line represents FDR = 0.05. Populations shown in red or blue were significantly (FDR < 0.05) increased or decreased (fold change > 1.5-fold), respectively, in convalescents. **E–M** The proportion of selected immune cell populations at 12, 16, and 24 w.p.i. compared to HC. Statistical significance was assessed using Wilcoxon rank sum tests. *P* values were adjusted for multiple testing using the Benjamini-Hochberg method. ns = non-significant. * FDR < 0.05, ** FDR < 0.01, *** FDR < 0.001

segregated based on functional marker characteristics including activation or maturation status. Differences in these major lineages, compared with HCs, were most strongly evident at 12 wpi, but some populations were still significantly different at 24 wpi (Fig. 2A–D, Table S2). While there was significant lymphopenia evident in convalescent individuals at 12 and 16 wpi (Fig. 2E), CD3+ T cells were significantly increased at 12 wpi, when expressed either as a percentage of lymphocytes (Fig. 2F) or as a percentage of live cells (data not shown). CD19+ B cells were also significantly increased at 12 and 16 wpi (Fig. 2G). We also

observed significantly increased CD38+CD27+ memory B cells at 16 wpi (Fig. 2A). When interrogating CD4+ T cell maturation, we observed a significant reduction in both the CD4+ and CD8+ compartments at 12 and 16 wpi (Fig. 2B,C). CD4+ effector memory (EM) pools were significantly reduced (Fig. 2H), and we also observed a significant reduction in migratory central memory (CM) CD4+ T cells, defined as CCR7+CD62L−, at all timepoints (Fig. 2I).

The NK cell compartment was also altered in convalescents at 12 and 16 wpi (Fig. 2A–C) with CD56++ NK cells, for example, significantly elevated at 12 and 16 wpi
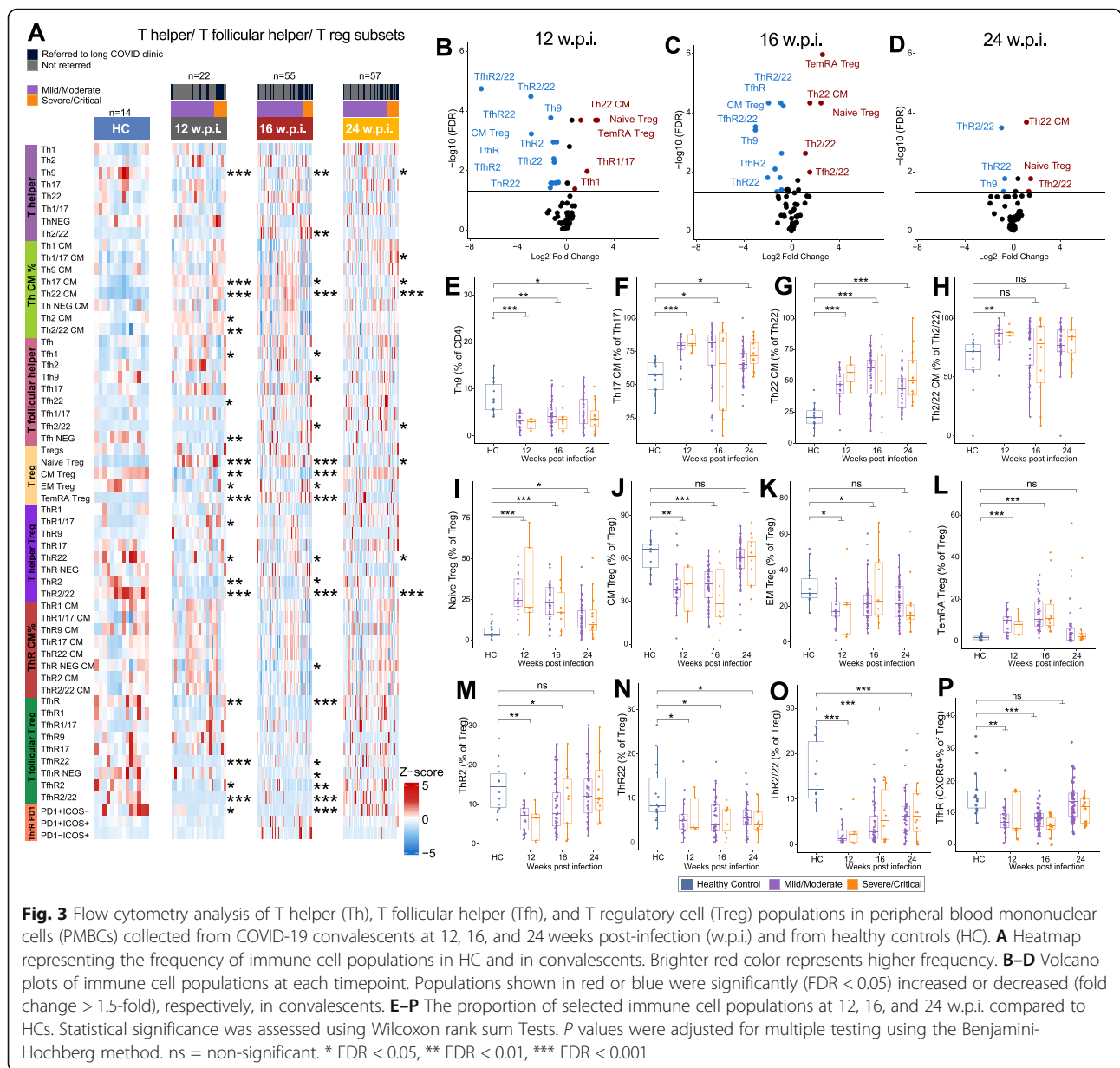
(Fig. 2J). We also observed a significant increase in total granulocytes at all 3 timepoints post-infection (Fig. 2A–D), and for low-density (LD) neutrophils at 12 and 16 wpi (Fig. 2K). CXCR3[+] LD neutrophils, which are actively recruited to sites of tissue damage [39], were elevated in convalescents at 12 wpi but returned to baseline by 16 wpi (Fig. 2A). Interestingly, CD14[+]CD16[+] neutrophils were significantly decreased at 12 and 16 wpi (Fig. 2A). While total monocyte proportions were not significantly altered, two subsets of tissue-homing CXCR3[+] monocytes (HLA-DR[+], activated antigen-presenting pro-inflammatory monocytes and HLA-DR[−], regulatory monocytes) were significantly increased in convalescent individuals at 12 wpi (Fig. 2L–M). We also investigated differences in immune cell populations between mild/moderate and severe/critical convalescents; however, after correction for multiple testing, there were no statistically significant differences (Table S2), most likely due to the small sample size of severe/critical convalescent samples, particularly at 12 wpi.

Next, we assessed correlations between immune cell populations (at 12, 16, or 24 wpi) and both anti-Spike and anti-RBD IgG, IgM, and IgA responses at 24 wpi (Table S3). Significant positive correlations were observed between the frequency of granulocytes, CD16[+] NK, and NKT-like cells at 12 wpi and anti-Spike IgG1 and anti-RBD IgG titres at 24 wpi. These data may be reflective of the correlation between disease severity and antibody responses. Previous work has suggested an association between increased percentage of neutrophils and lower anti-RBD IgG responses [40], which we did not detect in our analysis (Table S3). Components of the CD4 compartment were also significantly associated with anti-Spike IgG1 and anti-RBD IgG titres at 24 wpi. For example, there was a positive correlation between the proportion of CD4[+] cells in transition from naïve to CM, CM, to EM, and activated (HLA-DR[+] or CD38[+]) CD4[+] T cells, and anti-Spike and anti-RBD IgG/G1 titres at 24 wpi, suggesting each of these CD4 populations might contribute to robust T cell help. Significant correlations between immune cell populations at 16 and 24 wpi and anti-Spike or anti-RBD antibody responses were also observed (Table S3).

To interrogate CD4 Th responses in more depth, we applied a chemokine receptor-based gating strategy to characterise the Th effector phenotypes in both Th and Tfh subsets [41, 42]. We also used CD45RO[+] and CD62L[+] staining as a marker of T cell memory formation in the Th subsets. In addition, we applied the same strategy to T regulatory (Treg) subsets, which are functionally paired with their Th and Tfh counterparts in vivo [41]. Th and Tfh lineages were categorised into 8 functional subsets (Fig. 3A), and significant differences were observed for multiple subsets in COVID-19

convalescents (Fig. 3B–D, Table S2). We observed a significant decrease in Th9 cells at all timepoints (Fig. 3E). Th9 cells are implicated in autoimmune pathologies including airway damage and are predicted to home to sites of inflammation including the lung. It is plausible that decreased proportion in the PBMC reflect homing to the lung, or transdifferentiating into pathologic Th17 cells [43]. There was also a significant increase in Th2/22 cells at 16 wpi (Fig. 3A). We observed that while the proportion of Th17 and Th22 cells was not significantly different between groups, there was an increased proportion of Th17 and Th22 CM cells at all timepoints (Fig. 3F,G). This may indicate a role for these subsets in recovery after viral infection. In addition, there was evidence of increased formation of Th2/22 memory at 12 wpi (Fig. 3H), suggesting establishment of memory focused on tissue repair [44]. In the Tfh compartment, we observed significant differences in Tfh1, 9, 22, and 2/22 cells at different timepoints post-infection (Fig. 3A), with Tfh1 cells significantly elevated in convalescents at 12 and 16 wpi (Fig. 3A).

As with CD8[+] and CD4[+] effector T cells, Tregs segregate into naïve and mature populations depending on antigen exposure. While we found no difference in total Tregs (Table S2), we observed a significant increase in naïve Tregs at all timepoints post-infection (Fig. 3I), accompanied by a significant decrease in CM and EM Tregs at 12 and 16 wpi, and a significant increase in TEMRA Tregs (effector memory with acquired CD45RA) at 12 and 16 wpi (Fig. 3J–L). These data suggest either a block in maturation, or an increase in formation of naïve Treg cells in convalescents. The dual role of Treg cells in immune suppression and tissue repair suggests the potential for more than one mechanism of action in recovering individuals, so we examined functionally paired helper lineages in the Treg compartment, as they are likely to respond to the same pathogen-triggered homing cues as their Th effector counterparts. We observed a significant decrease in the proportion of ThR2 Tregs at 12 and 16 wpi, and a significant decrease in ThR22 and ThR2/22 Tregs at all timepoints (Fig. 3M–O), suggesting a block in commitment of these lineages. Finally, we also examined the follicular regulatory T cell lineages (TfhR), as they serve a similar regulatory role in germinal centres, controlling Tfh function and B cell help. We observed a significant decrease in total TfhR at 12 and 16 wpi (Fig. 3P) suggesting that follicular help is less restrained by TfhR in individuals recovering from COVID-19. Specifically, TfhR2, 22, and 2/22 subsets were all significantly reduced at 12 and 16 wpi but returned to baseline by 24 wpi. This is consistent with the regulatory follicular arm licencing a Tfh1 response early in infection, but later, skewing to Tfh2/22-driven B cell help in germinal

Fig. 3 Flow cytometry analysis of T helper (Th), T follicular helper (Tfh), and T regulatory cell (Treg) populations in peripheral blood mononuclear cells (PMBCs) collected from COVID-19 convalescents at 12, 16, and 24 weeks post-infection (w.p.i.) and from healthy controls (HC). **A** Heatmap representing the frequency of immune cell populations in HC and in convalescents. Brighter red color represents higher frequency. **B–D** Volcano plots of immune cell populations at each timepoint. Populations shown in red or blue were significantly (FDR < 0.05) increased or decreased (fold change > 1.5-fold), respectively, in convalescents. **E–P** The proportion of selected immune cell populations at 12, 16, and 24 w.p.i. compared to HCs. Statistical significance was assessed using Wilcoxon rank sum Tests. *P* values were adjusted for multiple testing using the Benjamini-Hochberg method. ns = non-significant. * FDR < 0.05, ** FDR < 0.01, *** FDR < 0.001

centres, both of which are required to drive an effective B cell response.

We also sought to determine links between T cell help and antibody responses to COVID-19, given that priming and durable immunity are underpinned by the interaction of T and B cells. To do this, we performed a correlation analysis between CD4[+] T cell subsets at 12, 16, and 24 wpi and antibody responses at 24 wpi (Table S3). We observed a number of interesting statistically significant correlations. For example, we observed a significant positive correlation between anti-Spike IgG1 levels and both ThR2/22 and TfhR2/22 subsets, suggesting that the effector function of this epithelial tissue-homing lineage may regulate antibody responses. Similar correlations between these subsets and anti-RBD IgG responses were also evident.

Whole blood RNA sequencing reveals significant perturbations to gene expression in COVID-19 convalescents until at least 6 months post-infection

To assess the potential long-term effects of SARS-CoV-2 infection on the peripheral blood transcriptome, total RNA sequencing was performed on 138 blood samples collected from individuals recovering from mild (*n* = 47), moderate (*n* = 6), severe (*n* = 7), or critical (*n* = 6) COVID-19 at 12, 16, and 24 wpi (Fig. 1A). RNA sequencing was also performed on blood collected from age-matched HCs (*n* = 14) with negative serology for the

SARS-CoV-2 Spike and RBD proteins. Approximately 9 billion 2 × 150 bp read pairs (mean 68.2 million per sample) were sequenced (Table S4).
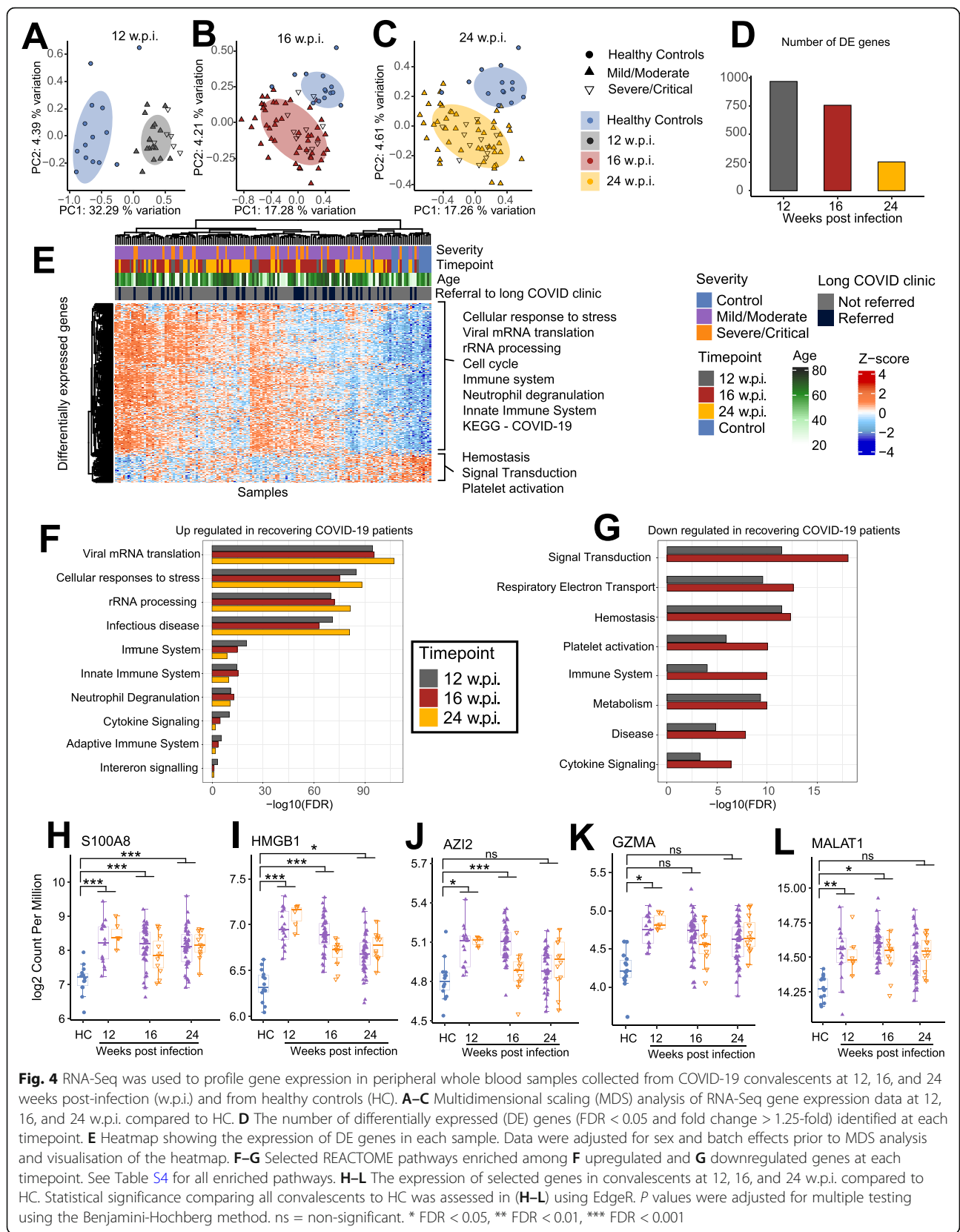
After adjusting for sex and batch effects, MDS analysis of the gene expression data revealed a clear separation between HCs and convalescent individuals at each timepoint (Fig. 4A–C). Consistent with these data, differential gene expression analysis identified > 950 genes that were significantly (FDR < 0.05, fold change > 1.25) differentially expressed (738 upregulated genes; 230 downregulated) in convalescent individuals at 12 wpi compared to HCs (Fig. 4D, Table S4). Similar results were observed when only mild/moderate convalescents were included in the analysis, indicating that differential gene expression was not driven solely by convalescents recovering from severe disease. Fewer differentially expressed genes (DEGs) were identified at 16 and 24 wpi, but there were still > 250 DEGs identified at 24 wpi (Fig. 4D, Table S4). Unsupervised hierarchical clustering analysis of DEGs did not reveal an obvious clustering by disease severity, suggesting that even individuals with mild COVID-19 have long-lasting changes to their blood transcriptome (Fig. 4E). There was a tendency for samples from the earlier timepoints to cluster together, consistent with a decrease in the number of DEGs over time, but clearly there was a spectrum in the recovery in gene expression among convalescent individuals, with some recovering more quickly (clustering with HCs).

Pathway and Gene Ontology (GO) analysis revealed a very strong enrichment for pathways related to transcription, translation, and ribosome biosynthesis among genes upregulated in convalescents, at all 3 timepoints (Fig. 4E,F, Table S4). In many cases, these signatures were predominantly driven by the upregulation of ribosomal RNA (rRNA) genes. Viral polypeptide synthesis is reliant upon host ribosomes and many viruses have been reported to stimulate rRNA synthesis upon infection [45, 46], although both the SARS-CoV-1 and SARS-CoV-2 Nsp1 protein has been shown to act a strong inhibitor of translation [47, 48]. Interestingly, a recent study has surprisingly shown that rRNA accumulation positively regulates antiviral innate immune responses against human cytomegalovirus infection [49], raising the possibility that the continued upregulation of rRNAs in individuals recovering from COVID-19 is a cellular defence mechanism. Consistent with this, the Reactome pathway "innate immune system" was significantly enriched among genes upregulated in convalescents (Fig. 4E,F, Table S4). Other statistically enriched pathways among upregulated genes included neutrophil degranulation, antimicrobial peptides, immune system, pathways related to other viral infections, cell cycle related pathways, and pathways related to the citric acid (TCA) cycle and respiratory electron transport/oxidative phosphorylation (Table S4).

Among downregulated genes at 12 and 16 wpi, there was a strong enrichment for metabolic related pathways such as oxidative phosphorylation as well as pathways related to platelet activation, signaling, and aggregation (Fig. 4G, Table S4). Platelet aggregation has previously been identified as a marker of severe SARS-CoV-2 infection [50], so it is interesting that genes involved in this process appear to be downregulated in recovering individuals (Fig. S2A, Table S4). Interestingly, we identified oxidative phosphorylation to be enriched among upregulated genes as well as downregulated genes. Increased expression of genes involved in oxidative phosphorylation has recently been reported in another study assessing COVID-19 convalescents [20]. Further examination of our data revealed that downregulated oxidative phosphorylation genes were encoded by the mitochondria, whereas upregulated ones were nuclear encoded (Fig. S2B). Differential expression of nuclear versus mitochondrially encoded oxidative phosphorylation genes has been reported in a number of other contexts [51]. Interestingly, given the memory and concentration issues frequently reported by long COVID patients, including in our study (Table 1), mitochondrial dysfunction in PBMC has previously been associated with cognitive impairment in other contexts [52]. As mentioned, at 24 wpi there were considerably fewer DEGs (~ 250) in convalescents compared to HCs; consistent with this, only one pathway, "complement activation", was identified as being enriched among genes downregulated at 24 wpi (Table S4).

Many of the most strongly upregulated genes in COVID-19 convalescents encoded known biomarkers of inflammation and innate immunity including S100 calcium-binding protein A8 (S100A8), and high-mobility group protein 1 (HMGB1), 5-azacytidine induced 2 (AZI2), and granzyme A (GZMA) (Fig. 4H–K). C-reactive protein (CRP) levels in serum were not, however, significantly different in convalescents compared to healthy controls (Fig. S2C). As we performed total RNA sequencing, we were also able to identify many differentially expressed long non-coding RNAs (Table S4) including metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) (Fig. 4L), which has been found upregulated in response to flavivirus and SARS-CoV-2 infection [53, 54] and is an important regulator of immunity and the cell cycle [55, 56].

As detailed above, flow cytometry analysis revealed significant changes in the proportion of multiple immune cell populations in convalescent individuals compared with HCs (Figs. 2 and 3). As we performed RNA-Seq on whole blood samples, it was therefore possible that the differences we observed in the transcriptome of recovering individuals simply reflected changes in immune cell populations, rather than differences in gene

Fig. 4 RNA-Seq was used to profile gene expression in peripheral whole blood samples collected from COVID-19 convalescents at 12, 16, and 24 weeks post-infection (w.p.i.) and from healthy controls (HC). **A–C** Multidimensional scaling (MDS) analysis of RNA-Seq gene expression data at 12, 16, and 24 w.p.i. compared to HC. **D** The number of differentially expressed (DE) genes (FDR < 0.05 and fold change > 1.25-fold) identified at each timepoint. **E** Heatmap showing the expression of DE genes in each sample. Data were adjusted for sex and batch effects prior to MDS analysis and visualisation of the heatmap. **F–G** Selected REACTOME pathways enriched among **F** upregulated and **G** downregulated genes at each timepoint. See Table S4 for all enriched pathways. **H–L** The expression of selected genes in convalescents at 12, 16, and 24 w.p.i. compared to HC. Statistical significance comparing all convalescents to HC was assessed in (**H–L**) using EdgeR. *P* values were adjusted for multiple testing using the Benjamini-Hochberg method. ns = non-significant. * FDR < 0.05, ** FDR < 0.01, *** FDR < 0.001

Ryan *et al. BMC Medicine*   (2022) 20:26

Page 13 of 23

expression. To assess this, we repeated the differential expression analysis multiple times, each time adjusting for changes in a major immune cell population. We found that accounting for changes in specific immune cell populations in our differential gene expression analysis models resulted in a decrease in the number of genes identified as differentially expressed (mean reduction in the number of genes with FDR < 0.05 was 50.47%); however, the statistical enrichment of immune, rRNA processing, cell cycle, and transcription/translation signatures identified among DEGs was robust to correction for differences in the proportion of any specific immune cell populations (Fig. S3, Table S4). These data indicate that the long-term perturbation of the blood transcriptome that we observe in convalescents compared to HCs is not solely explained by changes in the frequency of any single immune cell population.

Persistent changes in the transcriptome are associated with post-acute COVID-19 syndrome.

To assess whether differences observed in convalescent patients were associated with long COVID symptoms, we compared cellular immunophenotypes and transcriptional responses in convalescent individuals ($n = 21$) referred to a "long COVID" clinic compared to those convalescent individuals who were not ($n = 48$). Referral to the clinic was closely aligned with self-reported on-going symptoms (Fig. 5A) and was not significantly associated with disease severity (Fischer's exact test, P value = 0.1). Analysis of anti-Spike and anti-RBD antibody titres found no significant differences between individuals referred to the clinic and those who were not, indicating that humoral responses are similar in both groups of convalescent individuals. Similarly, none of the immune cell subsets we found to be persistently changed in convalescents compared to healthy controls were significantly altered in individuals referred to the long COVID clinic compared to those who were not. Interestingly, however, there were 446 genes that were significantly differentially expressed at 24 wpi in convalescent individuals referred to the long COVID clinic compared to those who were not (FDR < 0.05 and > 1.25× fold change (Fig. 5B,C, Table S5). No differentially expressed genes were identified between convalescent individuals referred to the long COVID clinic and those who were not at 12 or 16 wpi (although all convalescent had substantial changes in gene expression compared to HCs at these timepoints). These data suggest that changes to the blood transcriptome persist in individuals referred to a long COVID clinic, whereas they tend to resolve in convalescent individuals not suffering from long COVID symptoms.

Genes involved in transcription and translation and the cell cycle were enriched among genes upregulated in long COVID convalescents. Furthermore, there was a very strong enrichment for platelet-related pathways among downregulated genes (Fig. 5D,E, Table S5). Given none of the cell types we measured were found to be significantly different in long COVID convalescents in our flow cytometry data, we used the Molecular Signatures Database (MSigDB) cell type collection to attempt to identify which cell type(s) drove the observed transcriptional differences. Consistent with our pathway analyses, we identified a strong downregulation of platelet and megakaryocyte gene sets among individuals referred to a long COVID clinic (Fig. 5F, Table S5). Plotting the expression of well-known platelet genes encoding, for example, platelet factor 4 (PF4), platelet glycoprotein IX (GP9), thrombopoietin receptor (MPL), and coagulation factor XIII A chain (F13A1) (Fig. 5G–J), confirmed the downregulation of platelet genes in long COVID patients. There are a growing number of case reports of thrombocytopenia in COVID-19 patients [57–59] and interestingly, given that fatigue is one of the most commonly reported side-effects of long COVID, a common symptom of thrombocytopenia is also fatigue. There are also numerous case reports of thrombocytopenia after infection with a range of pathogens, including severe acute respiratory syndrome coronavirus 1 (SARS-CoV-1), influenza, and Zika virus [60–62].

Aside from platelet-related signatures, monocyte- and myeloid-related cells were also significantly enriched among downregulated genes, whereas CD4⁺ T cells were enriched among upregulated genes (Fig. 5F). Interestingly, given the important role of type I interferon (IFN-I) in COVID-19 [18], we also found decreased expression of multiple IFN-I-inducible genes including *MX1*, *OAS3*, and *OASL* (Fig. 5K–M). Of further interest given the neurological symptoms associated with long COVID including those reported in our cohort (Table 1), we found that the expression of S100B, a biomarker of neurological damage, was significantly increased in patients referred to a long COVID clinic [63].

Blood transcriptional module analysis highlights variable rates of recovery in the transcriptome of COVID-19 convalescents and correlations with antibody responses

We next sought to investigate individual-specific transcriptional changes in COVID-19 convalescents using pre-defined blood transcriptional modules (BTMs) [32]. To do this, we used Gene Set Variation Analysis (GSVA) [33] to reduce variation captured across > 20,000 genes in our gene expression data to an "activity score" for 256 BTMs in each individual (Fig. 6A and Table S6). Using *limma*, we identified 80 of these BTMs that were differentially active in convalescents (Table S6). The annotation of these BTMs was broadly consistent with our pathway analysis identifying multiple modules related to
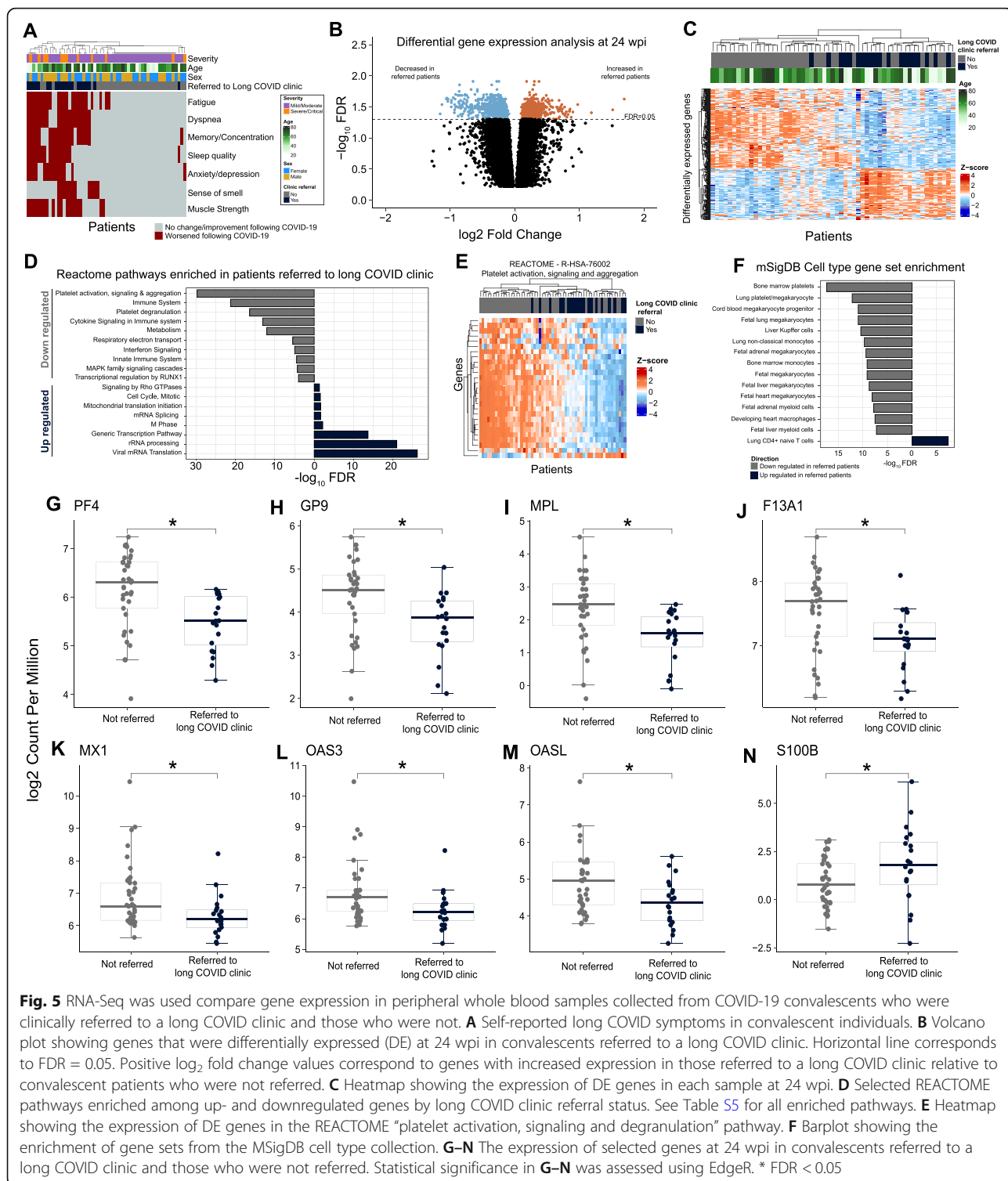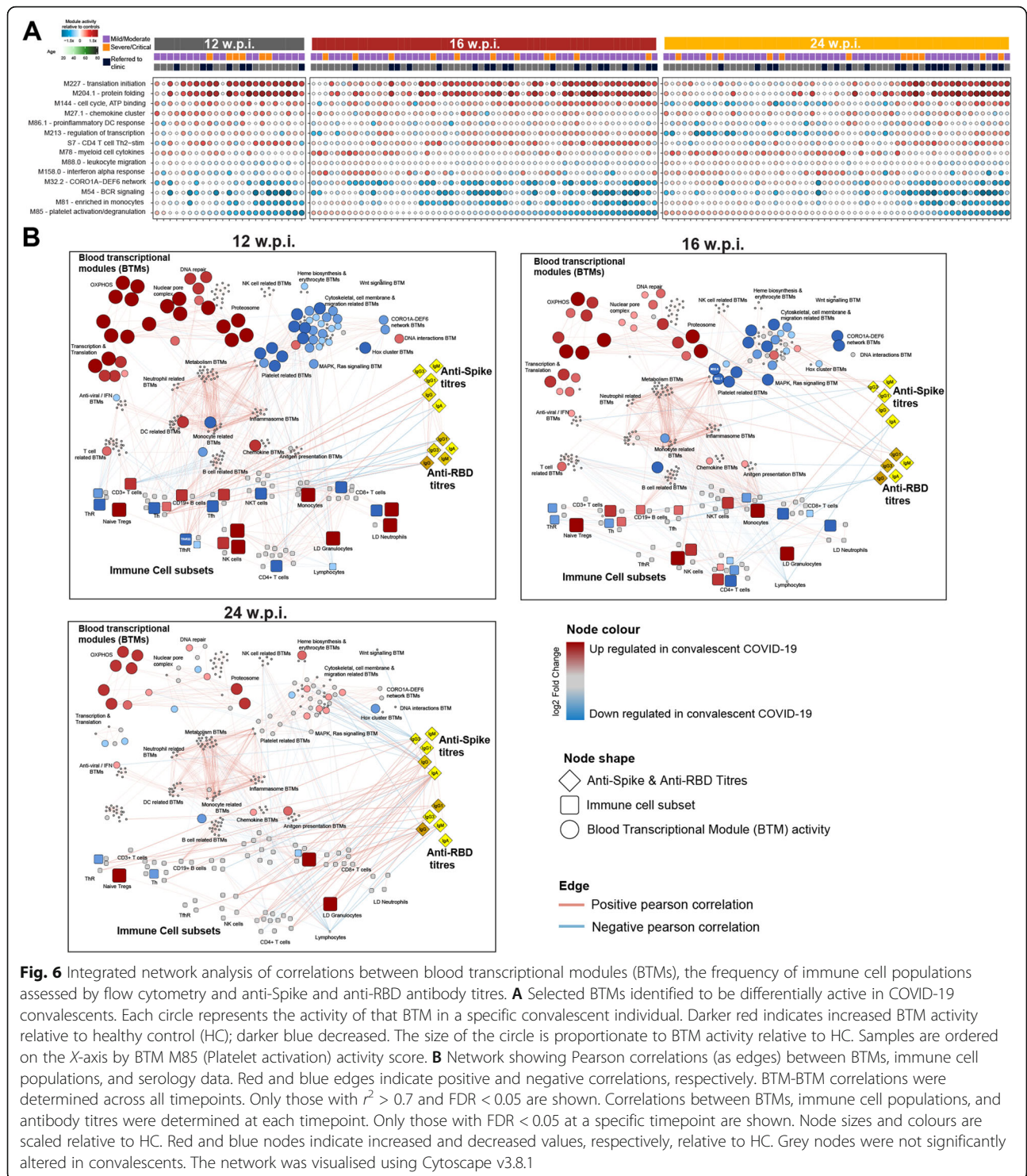
Fig. 5 RNA-Seq was used compare gene expression in peripheral whole blood samples collected from COVID-19 convalescents who were clinically referred to a long COVID clinic and those who were not. **A** Self-reported long COVID symptoms in convalescent individuals. **B** Volcano plot showing genes that were differentially expressed (DE) at 24 wpi in convalescents referred to a long COVID clinic. Horizontal line corresponds to FDR = 0.05. Positive log$_2$ fold change values correspond to genes with increased expression in those referred to a long COVID clinic relative to convalescent patients who were not referred. **C** Heatmap showing the expression of DE genes in each sample at 24 wpi. **D** Selected REACTOME pathways enriched among up- and downregulated genes by long COVID clinic referral status. See Table S5 for all enriched pathways. **E** Heatmap showing the expression of DE genes in the REACTOME "platelet activation, signaling and degranulation" pathway. **F** Barplot showing the enrichment of gene sets from the MSigDB cell type collection. **G–N** The expression of selected genes at 24 wpi in convalescents referred to a long COVID clinic and those who were not referred. Statistical significance in **G–N** was assessed using EdgeR. * FDR < 0.05

transcription/translation, the cell cycle and specific immune cell populations, and pathways as being significantly enriched in convalescents (Fig. 6A, Table S6). Interestingly, this analysis highlighted that while the proportion of recovering COVID-19 convalescents with

"healthy-like" BTM activity increased over time (consistent with a recovery to baseline over time), there were still a subset of convalescents with persistent transcriptional dysregulation at 24 wpi, which was associated with referral to a long COVID clinic (red and blue modules

Fig. 6 Integrated network analysis of correlations between blood transcriptional modules (BTMs), the frequency of immune cell populations assessed by flow cytometry and anti-Spike and anti-RBD antibody titres. **A** Selected BTMs identified to be differentially active in COVID-19 convalescents. Each circle represents the activity of that BTM in a specific convalescent individual. Darker red indicates increased BTM activity relative to healthy control (HC); darker blue decreased. The size of the circle is proportionate to BTM activity relative to HC. Samples are ordered on the *X*-axis by BTM M85 (Platelet activation) activity score. **B** Network showing Pearson correlations (as edges) between BTMs, immune cell populations, and serology data. Red and blue edges indicate positive and negative correlations, respectively. BTM-BTM correlations were determined across all timepoints. Only those with $r^2 > 0.7$ and FDR < 0.05 are shown. Correlations between BTMs, immune cell populations, and antibody titres were determined at each timepoint. Only those with FDR < 0.05 at a specific timepoint are shown. Node sizes and colours are scaled relative to HC. Red and blue nodes indicate increased and decreased values, respectively, relative to HC. Grey nodes were not significantly altered in convalescents. The network was visualised using Cytoscape v3.8.1

in Fig. 6A). Consistent with this, we identified 48 BTMs at 24 wpi that were differentially active by long COVID clinic referral status including multiple platelet, cell cycle, and immune-related BTMs (Table S6).

Finally, we undertook a systems-level integration of BTM activity scores, anti-Spike and anti-RBD antibody data, and flow cytometry data at 12, 16, and 24 wpi (Fig. 6B, Table S6, Additional files 12, 13 and 14). To do this, we constructed a network of significant correlations between BTMs, antibody titres, and the frequency of immune cell populations in each individual. Many BTMs, including those differentially active in convalescents,

were strongly correlated with each other. For example, monocyte-, DC-, neutrophil-, and inflammasome-related BTMs were strongly correlated with each other and, interestingly, with metabolism-related BTMs (Fig. 6B). DC-related BTMs also correlated with antiviral and interferon-related BTMs. At 12 wpi, there were also significant correlations identified between BTMs and > 30 different immune cell populations. For example, the proportion of CD19+ B cells, CD3+ T cells, and NK cells were strongly correlated with the activity scores of BTMs independently annotated to be related to these cell types (Fig. 6B, Table S6). Multiple different immune-related BTMs were also strongly positively correlated with the proportion of Tfh22-like Tregs at 12 wpi. At 16 wpi, BTMs correlated with fewer immune cell populations; however, multiple strong correlations were identified between BTMs and monocytes, B cells, and lymphocytes at this timepoint. Similar correlations were also identified at 24 wpi. Despite the strong correlations between many different BTMs and immune cell populations, interestingly, there were relatively few significant correlations between immune cell populations and differentially active BTMs, particularly at 16 and 24 wpi. These data suggest that the majority of differentially active BTMs we observed in convalescent individuals are not explained by differences in the frequency of immune cell populations in these individuals. For example, oxidative phosphorylation-related BTMs were differentially active at all timepoints; however, these BTMs were not significantly correlated with any immune cell population. Anti-Spike and/or anti-RBD antibody titres were also significantly correlated with BTMs at 16 and 24 wpi, but not 12 wpi (Fig. 6B, Table S6). Many of the BTMs that were correlated with antibody titres were downregulated in convalescents. For example, two platelet activation BTMs (M32.0 and M32.1) were significantly correlated with anti-Spike IgM responses at 16 wpi, while multiple cell adhesion-related BTMs were significantly negatively correlated with anti-Spike IgG responses at 24 wpi. We also identified multiple different immune cell populations that correlated with antibody titres at each timepoint. These relationships were particularly evident at 24 wpi. For example, the proportion of LD granulocytes, CD16+ NK cells, and CCR7−CD62L+ transitional memory T cells were significantly positively correlated with anti-Spike and anti-RBD IgG titres at 24 wpi. In summary, our integrated network analysis reveals a complex interplay of relationships between circulating immune cell populations, transcriptional dysregulation, and humoral immune responses in COVID-19 convalescent patients and provides a resource for further exploration and investigation of these relationships.

Discussion

Recovery from SARS-CoV-2 infection is frequently associated with persistent symptoms months after infection including fatigue, muscle weakness, sleep impairment, and anxiety or depression [4, 5, 64]. These data suggest ongoing immune dysregulation in COVID-19 convalescents which has been supported by several recent studies profiling the immune system in individuals recovering from COVID-19 using multi-parameter flow cytometry, bulk and single-cell transcriptomics, and other approaches [15, 16, 65–70]. Our study extends on these recently published studies, which have mostly assessed immune responses at 2–12 weeks post-infection. Here, we report an integrated analysis of immune responses at a transcriptional, cellular, and serological level, in individuals recovering from mild/moderate or severe/critical COVID-19 at 12, 16, and 24 weeks post-infection, in comparison to age-matched HCs.

Anti-Spike and anti-RBD serology data demonstrated heterogeneity of antibody responses to SARS-CoV-2 consistent with previously published reports showing long-lasting IgG and IgG1 antibody responses to at least 6 months post-infection which were correlated with disease severity [36, 71, 72]. Our cohort is particularly well-suited to the assessment of the durability of antibody responses due to the negligible risks of re-infection in South Australia where, due to strict border restrictions and public health measures, community transmission was eliminated during the sample collection period. Despite the anticipated decay in IgA and IgM [73–75], a large percentage of convalescents remained seropositive for both RBD- and Spike-specific Ig (all isotypes) for the duration of the study. This decay was less pronounced at 24 wpi in the severe COVID-19 convalescents compared to the mild cohort, with significant differences in RBD-specific IgM and IgG3 isotypes between the two groups. Recently, declining levels of SARS-CoV-2 Spike-specific IgM in mild COVID-19 convalescents were found to strongly correlate with serum virus neutralisation activity [76], findings that were further confirmed in experiments with purified IgM fractions and IgM-depleted sera from similar patients [40, 77]. In COVID-19 convalescents, IgM, similarly to IgG1, preferentially targets the S1 domain of the Spike protein [78], the region that contains the RBD and N-terminus domains and the target of most neutralising antibodies and regions of high interest for developing passive immunotherapies to deal with new SARS-CoV-2 variants of concern [79]. Conversely, less abundant SARS-CoV-2-specific IgG3 targets the S2 domain more efficiently [78], which suggests that its ability to neutralise the virus is, by comparison, reduced. Yet, S2 contains the sequences that allow SARS-CoV-2 membrane fusion with the cell host membrane, a key step in virus entry [2]. In fact, the ability of

antibodies targeting S2 regions involved in membrane fusion to block Spike protein-mediated cell-cell fusion has been confirmed experimentally [80]. In the future, it will be necessary to elucidate the particular roles of IgM and IgG3 in neutralising SARS-CoV-2 but, perhaps too, blocking virus infection by other mechanisms such as blockade of membrane fusogenic regions of the Spike protein. This will provide further insights into the over-all importance of specific Ig isotypes in determining disease severity and outcomes.

In addition to our serological analysis of COVID-19 convalescents, we extensively and longitudinally profiled immune cell populations in the same individuals using a multi-panel approach that enabled the identification and enumeration of ~ 130 different sub-populations including deep phenotyping of the CD4 and CD8 compartments. Differences in immune cell populations compared with HCs were most strongly evident at 12 wpi, but some populations were still significantly different at 24 wpi. $CD56^{++}$ NK cells, granulocytes, LD neutrophils, and tissue-homing $CXCR3^+$ monocytes were significantly increased in convalescents at 12 wpi. Many of these changes persisted until at least 16 or 24 weeks. Consistent with our data, increased NK cells [65] and granulocytes [68] have been reported in other cohorts of convalescents and scRNA-Seq has revealed that increased non-classical monocytes are associated with more severe disease during active infection [81]. In contrast to our study, a study of 109 Austrian convalescents at 10 weeks post-infection did not find neutrophils, monocytes, $CD3^+$ T cells, $CD56^+$ NK cells, or $CD19^+$ B cells to be significantly different in convalescents [68]. Other studies have also reported significant decreases in the frequencies of invariant NKT and NKT-like cells [66], which we and others [20] did not observe.

Several previous studies have reported that T and B cell activation/exhaustion markers remain elevated following SARS-CoV-2 infection [15]. Furthermore, $CD4^+$ and $CD8^+$ EM T cells have been reported to be significantly higher in convalescents at 10 wpi [68]. Consistent with reports in active infection and convalescence [15], convalescent individuals in our study had lymphopenia until at least 16 wpi; however, $CD3^+$ T cells were significantly increased at 12 wpi. We also observed significantly increased $CD19^+$ B cells at 12 and 16 wpi and $CD38^+CD27^+$ memory B cells at 16 wpi in convalescents. Recent studies have shown that increased activation and exhaustion of memory B cells observed during COVID-19 correlates with $CD4^+$ T cell functions [82], and consistent with this, we observed reduced $CD4^+$ EM cell proportions in COVID-19 convalescents at 12 wpi. We were particularly interested in the role of regulatory T cells (Tregs) in COVID-19, as there have been conflicting reports of Tregs being either increased or decreased in convalescents. Significantly increased $Foxp3^+$

Tregs were observed in 49 convalescents from Wuhan at ~ 112 days post-recovery [66]; however, another study observed that $CD25^+Foxp3^+$ Tregs were significantly reduced 10 weeks after COVID-19 [68]. A more recent study has also reported that Tregs in severe COVID-19 patients have a distinct transcriptional signature with similarities to tumour-infiltrating Tregs, which persist in convalescent patients [83]. We observed no significant difference in the total ($CD4^+CD25^+CD127^{low}$) Treg pool at any timepoint, but when we interrogated Tregs for their memory/maturation status, we observed that the naïve and TEMRA Treg proportions were significantly increased at 12 and 16 wpi, while EM and CM Tregs were significantly reduced, mirroring a similar reduction in the proportion of $CD4^+$ EM and CM pools at 12 and 16 wpi. Interestingly, a number of the Th lineage subsets including Th2, Th22, Th2/22, and Th17 had increased proportions of CM vs EM, revealing subtle skewing of the Th memory formation. The expansion of naïve Tregs could be an attempt to restore the balance in the Treg pool in the face of both inflammation and tissue damage, which is supported by emerging evidence of a dual role for Tregs in suppressing immune responses and promoting tissue repair [84]. Increased TEMRA Tregs, which are often associated with exhaustion, but are in fact a poly-functional effector Treg population with characteristics of cytotoxic cells, migratory T cells, and tissue repair cells [85, 86], further suggest a competition between classical immune suppression and tissue repair by these cells in response to tissue damage in COVID-19 convalescents.

Each Th subset has a paired regulatory subset [41], and this includes Tfh subsets, as B cell help in germinal centres also requires regulation in the steady state [87]. In a stereotypical antiviral immune response, Th1 cells migrate to sites of viral infection to establish an adaptive response, and regulatory cells co-migrate to limit chronic inflammation once the pathogen levels decline; however, there is an emerging function of tissue-resident Treg cells in tissue repair [84, 88]. We did not observe increased Th1 cells, but we did observe a reduction of Th9 cells potentially suggesting a diversion of Th9 cells to other sites. We also observed that the maturation of Th pools was enhanced in both Th17 and Th22 subsets, where CM marker proportions were increased at all timepoints post-infection. This may suggest that epithelial homing and tissue damage trigger activation and form part of the COVID-19 T cell recall response. It is intriguing that the Treg partners of these lineages, including ThR2, ThR22, and ThR2/22 were all significantly reduced over the same time course post-infection, suggesting that the signal recruiting Th cells to tissue locations are persistent long after COVID infection. A similar imbalance in follicular help vs follicular

regulation was also observed, whereby Tfh1 and Tfh2/22 cells were significantly elevated post COVID-19, but total TfhR, TfhR2, TfhR22, and TfhR2/22 cells were reduced. Other studies have demonstrated that CXCR5$^+$ Tfh populations are significantly elevated in individuals recovering from COVID-19 and correlate with robust humoral immunity [89]; however, this previous study did not analyse the regulatory arm in this compartment. Another previous study has reported a decline in Tfh cells at 4 months post-infection [74]. Interestingly, another previous study has suggested that germinal centre formation is impaired in acute infection [90]. This previous study was based on the analysis of post-mortem lymph nodes and spleen in patients that succumbed to SARS-CoV-2 infection, whereas, in our study, we have assessed antibody responses in convalescent survivors, who clearly have strong humoral responses. Our data would suggest germinal centre formation is sufficient in convalescents.

In addition to immunophenotyping by flow cytometry, we performed RNA sequencing of total RNA from 138 blood samples collected from convalescent individuals at 12, 16, and 24 wpi, as well as HCs. To our knowledge, no other study has profiled transcriptome-wide changes in COVID-19 convalescents for such a long period post-infection. We found that the blood transcriptome of convalescents was significantly perturbed compared to HCs, with the largest numbers of DEGs being identified at 12 wpi. Transcriptional dysregulation persisted until at least 24 weeks. There was a very strong enrichment for pathways and BTMs related to transcription, translation, and ribosome biosynthesis among genes upregulated in recovering individuals, at all 3 timepoints. Many viruses upregulate rRNA synthesis during infection [42, 43], but why rRNA gene expression remains upregulated months after infection is currently unknown. Other statistically enriched pathways among upregulated genes included neutrophil degranulation, antimicrobial peptides, immune system, and pathways related to other viral infections. These data suggest ongoing inflammatory responses and immune dysregulation in COVID-19 convalescents weeks-to-months after infection. Consistent with these data, neutrophil degranulation has reported to be significantly upregulated in active infection [91, 92], suggesting that certain signatures of active infection persist well into convalescence. We also found evidence for dysregulated expression of genes involved in oxidative phosphorylation, a signature which has also been identified in one other recent study of convalescents to occur irrespective of whether elevated inflammatory markers persist or not [20], but whose functional significance is currently unknown. Interestingly, mitochondrial dysfunction in PBMC has previously been associated with cognitive impairment in other contexts [52]. This warrants further investigation given the frequent reports of cognitive issues in long COVID sufferers.

While some changes in gene expression were associated with variation in specific immune cell populations between individuals, differences in gene expression were not solely explained by changes in the frequency of any single immune cell population. A patient-specific analysis of the gene expression activity of pre-annotated BTMs enabled a more thorough assessment of the variation in gene expression responses. There was a broad spectrum in the recovery of gene expression responses in both mild/moderate and severe/critical convalescents. Variation in the rate of recovery from infection at a cellular and transcriptional level may explain the persistence of symptoms, such as fatigue, associated with long COVID in some convalescent individuals. We observed a strong association between our transcriptional signature of convalescence and referral to a dedicated long COVID clinic. While the majority of convalescent individuals in this cohort returned to a transcriptional baseline by 24 wpi, those referred to a long COVID clinic did not. More than 400 genes were identified to be differentially expressed in those convalescent individuals referred to a long COVID clinic compared to those convalescents who were not. Interestingly, these differences were only evident at 24 wpi, suggesting that while transcriptional dysregulation in many convalescents begins to resolve around 6 months post-infection, it persists in those individuals suffering from long COVID symptoms.

Of particular interest given known associations with symptoms such as fatigue, we identified several transcriptional signatures among long COVID convalescents that suggested a mild thrombocytopenia. There was, for example, a very strong enrichment for platelet-related pathways among downregulated genes and cell type enrichment analysis revealed a strong downregulation of platelet and megakaryocyte gene sets among individuals referred to a long COVID clinic. Consistent with our data, there are reports of thrombocytopenia in COVID-19 patients [57–59]. Furthermore, SARS-CoV2-2 infection has also been shown to induce changes in platelet gene expression and function [93, 94]. Unfortunately, we did not measure platelet levels in these individuals, so this is something that requires further assessment in future studies. Interestingly, a link between gene expression in peripheral blood and fatigue following infectious mononucleosis has been previously reported [95], with at least some of the same genes differentially expressed in COVID-19 convalescents. These data may point towards common mechanisms regulating long COVID and post-viral infection fatigue more generally. Finally, we also uncovered significant inverse correlations between dysregulated BTMs and anti-Spike and anti-RBD antibody responses suggesting that prolonged transcriptional

dysregulation may be associated with reduced antibody responses with potential consequences for the durability of protective immunity. Further work is now needed to assess whether dysregulated immunity following COVID-19 has implications for responses to other infections, vaccination, or in the management of chronic diseases.

While our study provides a high-resolution, multi-level insight into the immune dysregulation experienced post COVID-19, we recognise that our study also has some important limitations. While comparable to or larger than most other studies to date, the sample size is still relatively limited, particularly in the case of patients with more severe disease. This is particularly important given the apparently highly heterogenous recovery in immune dysregulation over time. Further larger studies will be needed to more fully assess differences due to disease severity, treatment, and other confounders and validate that the observed transcriptional changes are reflected at the protein level. Other single-cell approaches may also provide further resolution of the immune dysregulation experienced by convalescents and the transcriptional signatures we find associated with long COVID. We chose to perform the high-resolution immunophenotyping on freshly isolated PBMC in order to enrich for the rare lymphocyte subsets that are functionally important but not found in large number in whole blood, and this has some limitations when calculating the proportion of all cells found in the blood. Similarly, there are known biases introduced due to the removal of mature granulocytes from whole blood. Importantly, this is an approach that has also been successfully applied to other published COVID-19 cohort studies [20, 65]). It is important to acknowledge the limitations associated with examination of cell proportions versus absolute cell counts, specifically that a lowered proportion does not always equate to a lowered absolute cell count. With regard to presenting the data as absolute cell counts or proportion of a reference cell pool, we selected a proportion analysis to reflect changes in the balance between multiple rare but clinically important lymphocyte subsets using parameters such as maturation status or homing potential. In addition, we have normalised the staining protocols to a fixed PBMC count (5×10^5) at input for each sample, to minimise batch effects or donor cell count differences, ensuring the data are comparable between multiple donors over multiple time points.

While our flow cytometry analyses enabled the assessment of ~ 130 parameters, it did not include markers for dendritic cells (DC), which have been found to be altered in COVID-19 convalescents in previous studies [96]. Our BTM analysis, however, supports the dysregulation of DC populations in convalescents. Finally, while we assessed the relationships between immune dysregulation and anti-Spike and anti-RBD antibody responses,

we did not assess T cell immunity in our study [97, 98]. Further studies should also assess the effects of SARS-CoV-2 variants on long-term immune dysregulation in convalescents and comparative studies assessing differences between post-infectious immune dysregulation following SARS-CoV-2 infection in comparison to other infections would be highly beneficial. Due to the global impact of the pandemic, multiple protocols for separating and analysing the immune compartment have been used in multiple studies, and we acknowledge the limitation that in order to directly compare data between multiple cohorts, an international clinical protocol would need to be established with standardised cohort clinical inclusion criteria, standardised cell isolation and flow cytometry protocols, and standardised data analysis.

Conclusions

In conclusion, this study found persistent changes to the peripheral immune system of SARS-CoV-2 convalescents until at least 6 months post-infection and identified a subset of these changes that were associated with long COVID. These changes to the peripheral immune system could have implications for how individuals recovering from SARS-CoV-2 infection respond to other infections encountered in this period and persistent immune activation may also exacerbate other chronic conditions.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12916-021-02228-6.

Additional file 1: Figure S1: Stability of anti-Spike and anti-RBD antibody titres over time. **(A-E)** Anti-Spike and (F-J) anti-RBD IgG, IgG1, IgG3, IgM and IgA titres plotted as a function of time. End point titres are reported as $\log_{10}$ area under the curve (AUC). The blue line represents the line of best fit from a linear regression analysis. The shaded areas represent the 95% confidence interval. The *P* value shown is from the linear regression. Red dashed lines represent the mean AUC + 2 SD in healthy controls for each isotype.

Additional file 2: Figure S2: Expression of genes in two pathways identified as downregulated in COVID-19 convalescents and healthy controls. **(A)** Reactome pathway R-HSA-76002 "Platelet activation, signaling and aggregation" and **(B)** KEGG pathway hsa00190 "Oxidative phosphorylation". Oxidative phosphorylation genes are sub-divided into nuclear and mitochondrially encoded, with the same x axis order of samples in each panel. Only differentially expressed genes (FDR < 0.05 and fold change > 1.25-fold) within each pathway are shown. **(C)** Serum CRP levels in samples collected from healthy controls (HC) and COVID-19 convalescent individuals at 12, 16 and 24 weeks post infection.

Additional file 3: Figure S3: Adjusting for differences in immune cell populations. We repeated the differential expression analysis multiple times, each time adjusting for differences in the frequency of major immune cell populations among individuals. Each panel shows the enrichment of selected pathways (same as those shown in Fig. 4F) among genes identified as being significantly up-regulated in each analysis (FDR < 0.05 and fold change > 1.25-fold). None = no immune cell population adjustment.

Additional file 4: Table S1. Subject meta-data, and antibody titres for COVID-19 convalescents and healthy controls.

Ryan *et al. BMC Medicine*   (2022) 20:26

Page 20 of 23

Additional file 5: Table S2. Flow cytometry data showing the fold-change and FDR values for each immune cell population in COVID-19 convalescents compared to healthy controls.

Additional file 6: Table S3. Pearson correlations between the frequency of immune cell populations at 12, 16, and 24 weeks post-infection (wpi) and anti-Spike and anti-RBD antibody responses.

Additional file 7: Table S4. RNAseq sample statistics, differentially expressed genes (DEGs), pathway analyses, and adjusting for immune cell populations.

Additional file 8: Table S5. Differentially expressed genes and pathways at 24 weeks post infection in convalescent individuals that were clinically referred to a dedicated long COVID clinic, compared to those convalescent individuals who were not.

Additional file 9: Table S6. BTM and correlation analysis.

Additional file 10. Symptom questionnaire sent to patients to assess presence of long COVID associated symptoms.

Additional file 11. Flow cytometry gating strategy.

Additional file 12. Correlation network at 12-weeks post infection in Simple Interaction Format. The first column is a source node, second the spearman correlation coefficient and third the target node. File related to Fig. 6B.

Additional file 13. Correlation network at 16-weeks post infection in Simple Interaction Format. The first column is a source node, second the spearman correlation coefficient and third the target node. File related to Fig. 6B.

Additional file 14. Correlation network at 24-weeks post infection in Simple Interaction Format. The first column is a source node, second the spearman correlation coefficient and third the target node. File related to Fig. 6B.

Authors' contributions
The study was conceived and designed by B.G.B, S.C.B, and D.J.L.; C.M.H. and S.C.B. generated the flow cytometry data. F.J.R. performed the RNA-Seq analysis and the majority of statistical analyses and visualisations in the paper under the supervision of D.J.L. M.A.L. extracted the RNA for RNA-Seq. A.E.L.Y., M.G.M., Z.A.M., P.G-V., Z.A-D., and B.G.B. generated the serology data. J.G. and C.M.H. did the sample collection and processing. C.F. S.O'C., B.A.J.R., D.S., C.K-L., J. G., and M.R.B. did the patient enrolment and management and sample collection. P.H. did the CRP analysis. The manuscript was written by F.J.R., B.G.B, S.C.B, and D.J.L with contributions from all of the authors. All authors read and approved the final manuscript

Availability of data and materials
RNA-Seq data have been deposited in the Gene Expression Omnibus (GEO) under accession GSE169687 [99]. Count tables, metadata, and R code for all analyses (serology, flow cytometry and RNA-Seq) have been uploaded to the Lynn Laboratory BitBucket (https://bitbucket.org/lynnlab/covid-sa).

Declarations

Ethics approval and consent to participate
The study was performed in accordance with the ethical principles consistent with the latest version of the Declaration of Helsinki (version Fortaleza 2013) and Good Clinical Practice (GCP) and according to the National Health and Medical Research Council (NHMRC) Guidelines for Research published in the National Statement on the Ethical Conduct in Human Research (2007; updated 2018). The protocol was approved CALHN Human Research Ethics Committee, Adelaide, Australia (Approval No. 13050). Written informed consent was obtained from all participants enrolled in the studies.

Consent for publication
Informed consent was collected from all subjects to publish de-identified data.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Precision Medicine Theme, South Australian Health and Medical Research Institute, Adelaide, SA 5001, Australia. [2]Women's and Children's Health Network, North Adelaide, SA, Australia. [3]Molecular Immunology, Robinson Research Institute, University of Adelaide, Adelaide, SA, Australia. [4]Viral Immunology Group, Adelaide Medical School, University of Adelaide and Basil Hetzel Institute for Translational Health Research, Adelaide, SA, Australia. [5]Gene Silencing and Expression Core Facility, Adelaide Health and Medical Sciences, Robinson Research Institute, University of Adelaide, Adelaide, SA, Australia. [6]Infectious Diseases Department, Royal Adelaide Hospital, Central Adelaide Local Health Network, Adelaide, SA, Australia. [7]Intensive Care Unit, Royal Adelaide Hospital, Central Adelaide Local Health Network and Adelaide Medical School, University of Adelaide, Adelaide, SA, Australia. [8]Microbiology and Infectious Diseases Department, SA Pathology, Adelaide, SA, Australia. [9]Department of Renal Medicine, Flinders Medical Centre, Flinders University, Bedford Park, SA 5042, Australia. [10]Flinders Health and Medical Research Institute, Flinders University, Bedford Park, SA 5042, Australia. [11]Research Centre for Infectious Diseases, School of Biological Sciences, University of Adelaide, Adelaide, SA 5005, Australia.

References
1. Marshall M. COVID's toll on smell and taste: what scientists do and don't know. Nature. 2021;589(7842):342–3. https://doi.org/10.1038/d41586-021-00055-6.
2. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan. China Lancet. 2020;395(10223):497–506. https://doi.org/10.1016/S0140-6736(20)30183-5.
3. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis. 2020;20(5):533–4. https://doi.org/10.1016/S1473-3099(20)30120-1.
4. Carfì A, Bernabei R, Landi F, Gemelli Against C-P-ACSG. Persistent symptoms in patients after acute COVID-19. JAMA. 2020;324(6):603–5. https://doi.org/10.1001/jama.2020.12603.
5. Huang C, Huang L, Wang Y, Li X, Ren L, Gu X, et al. 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. Lancet. 2021;397(10270):220–32. https://doi.org/10.1016/S0140-6736(20)32656-8.
6. Bellan M, Soddu D, Balbo PE, Baricich A, Zeppegno P, Avanzi GC, et al. Respiratory and psychophysical sequelae among patients with COVID-19 four months after hospital discharge. JAMA Netw Open. 2021;4(1):e2036142. https://doi.org/10.1001/jamanetworkopen.2020.36142.
7. Taboada M, Carinena A, Moreno E, Rodriguez N, Dominguez MJ, Casal A, et al. Post-COVID-19 functional status six-months after hospitalization. J Infect. 2020;82(4):e31–e3. https://doi.org/10.1016/j.jinf.2020.12.022.
8. Moreno-Perez O, Merino E, Leon-Ramirez JM, Andres M, Ramos JM, Arenas-Jimenez J, et al. Post-acute COVID-19 Syndrome. Incidence and risk factors: a Mediterranean cohort study. J Infect. 2021;82(3):378–83. https://doi.org/10.1016/j.jinf.2021.01.004.
9. Yelin D, Wirtheim E, Vetter P, Kalil AC, Bruchfeld J, Runold M, et al. Long-term consequences of COVID-19: research needs. Lancet Infect Dis. 2020; 20(10):1115–7. https://doi.org/10.1016/S1473-3099(20)30701-5.
10. Mandal S, Barnett J, Brill SE, Brown JS, Denneny EK, Hare SS, et al. 'Long-COVID': a cross-sectional study of persisting symptoms, biomarker and

imaging abnormalities following hospitalisation for COVID-19. Thorax. 2020;
76(4):396–8. https://doi.org/10.1136/thoraxjnl-2020-215818.

11. Arnold DT, Hamilton FW, Milne A, Morley AJ, Viner J, Attwood M, et al.
Patient outcomes after hospitalisation with COVID-19 and implications for
follow-up: results from a prospective UK cohort. Thorax. 2020;76(4):399–401.
https://doi.org/10.1136/thoraxjnl-2020-216086.

12. Stavem K, Ghanima W, Olsen MK, Gilboe HM, Einvik G. Persistent symptoms
1.5-6 months after COVID-19 in non-hospitalised subjects: a population-
based cohort study. Thorax. 2020;76(4):405–7. https://doi.org/10.1136/thora
xjnl-2020-216377.

13. Clark DV, Kibuuka H, Millard M, Wakabi S, Lukwago L, Taylor A, et al. Long-
term sequelae after Ebola virus disease in Bundibugyo, Uganda: a
retrospective cohort study. Lancet Infect Dis. 2015;15(8):905–12. https://doi.
org/10.1016/S1473-3099(15)70152-0.

14. Ngai JC, Ko FW, Ng SS, To KW, Tong M, Hui DS. The long-term impact of
severe acute respiratory syndrome on pulmonary function, exercise capacity
and health status. Respirology. 2010;15(3):543–50. https://doi.org/10.1111/
j.1440-1843.2010.01720.x.

15. Files JK, Boppana S, Perez MD, Sarkar S, Lowman KE, Qin K, et al.
Sustained cellular immune dysregulation in individuals recovering from
SARS-CoV-2 infection. J Clin Invest. 2021;131(1):e140491. https://doi.org/1
0.1172/JCI140491.

16. Chioh FW, Fong S-W, Young BE, Wu K-X, Siau A, Krishnan S, et al.
Convalescent COVID-19 patients are susceptible to endothelial dysfunction
due to persistent immune activation. Elife. 2021;10:e64909. https://doi.org/1
0.7554/eLife.64909.

17. Wen W, Su W, Tang H, Le W, Zhang X, Zheng Y, et al. Immune cell profiling
of COVID-19 patients in the recovery stage by single-cell sequencing. Cell
Discov. 2020;6(1):31. https://doi.org/10.1038/s41421-020-0168-9.

18. Ren X, Wen W, Fan X, Hou W, Su B, Cai P, et al. COVID-19 immune features
revealed by a large-scale single-cell transcriptome atlas. Cell. 2021;184(23):
1815–913. https://doi.org/10.1016/j.cell.2021.10.023.

19. Zheng HY, Xu M, Yang CX, Tian RR, Zhang M, Li JJ, et al. Longitudinal transcriptome
analyses show robust T cell immunity during recovery from COVID-19. Sig
Transduct Target Ther. 2020;5(1):294. https://doi.org/10.1038/s41392-020-00457-4.

20. Bergamaschi L, Mescia F, Turner L, Hanson AL, Kotagiri P, Dunmore BJ, et al.
Longitudinal analysis reveals that delayed bystander CD8+ T cell activation and
early immune pathology distinguish severe COVID-19 from mild disease.
Immunity. 2021;54(6):1257–75 e8. https://doi.org/10.1016/j.immuni.2021.05.010.

21. NIH. Clinical Spectrum of SARS-CoV-2 Infection 2021 [updated 17/12/2020].
Available from: https://www.covid19treatmentguidelines.nih.gov/overview/
clinical-spectrum. Accessed 18 Dec 2020.

22. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, et al.
Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR.
Eurosurveillance. 2020;25(3):2000045. https://doi.org/10.2807/1560-7917.ES.2
020.25.3.2000045.

23. Hsieh C-L, Goldsmith JA, Schaub JM, DiVenere AM, Kuo H-C, Javanmardi K,
et al. Structure-based design of prefusion-stabilized SARS-CoV-2 spikes.
Science. 2020;369(6510):1501–5. https://doi.org/10.1126/science.abd0826.

24. Amanat F, Stadlbauer D, Strohmeier S, Nguyen TH, Chromikova V, McMahon
M, et al. A serological assay to detect SARS-CoV-2 seroconversion in humans.
Nat Med. 2020;26(7):1033–6. https://doi.org/10.1038/s41591-020-0913-5.

25. Andrews S. FastQC: a quality control tool for high throughput sequence
data; 2010. Available online at: http://www.bioinformatics.babraham.ac.uk/
projects/fastqc/.

26. Ewels P, Magnusson M, Lundin S, Kaller M. MultiQC: summarize analysis
results for multiple tools and samples in a single report. Bioinformatics.
2016;32(19):3047–8. https://doi.org/10.1093/bioinformatics/btw354.

27. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina
sequence data. Bioinformatics. 2014;30(15):2114–20. https://doi.org/10.1093/
bioinformatics/btu170.

28. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low
memory requirements. Nat Methods. 2015;12(4):357–60. https://doi.org/10.1
038/nmeth.3317.

29. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose
program for assigning sequence reads to genomic features. Bioinformatics.
2014;30(7):923–30. https://doi.org/10.1093/bioinformatics/btt656.

30. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for
differential expression analysis of digital gene expression data.
Bioinformatics. 2010;26(1):139–40. https://doi.org/10.1093/bioinformatics/
btp616.

31. Leek JT. svaseq: removing batch effects and other unwanted noise from
sequencing data. Nucleic Acids Res. 2014;42(21):e161.

32. Li S, Rouphael N, Duraisingham S, Romero-Steiner S, Presnell S, Davis C,
et al. Molecular signatures of antibody responses derived from a systems
biology study of five human vaccines. Nat Immunol. 2014;15(2):195–204.
https://doi.org/10.1038/ni.2789.

33. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for
microarray and RNA-Seq data. BMC Bioinformatics. 2013;14(1):7. https://doi.
org/10.1186/1471-2105-14-7.

34. Nalbandian A, Sehgal K, Gupta A, Madhavan MV, McGroder C, Stevens JS,
et al. Post-acute COVID-19 syndrome. Nat Med. 2021;27(4):601–15. https://
doi.org/10.1038/s41591-021-01283-z.

35. Australia GoS. Government of South Australia; 2021 Available from: https://
www.covid-19.sa.gov.au/home/dashboard. Accessed 15 Sept 2021.

36. Dan JM, Mateus J, Kato Y, Hastie KM, Yu ED, Faliti CE, et al. Immunological
memory to SARS-CoV-2 assessed for up to 8 months after infection. Science.
2021;371(6529):eabf4063.

37. Piccoli L, Park Y-J, Tortorici MA, Czudnochowski N, Walls AC, Beltramello M,
et al. Mapping neutralizing and immunodominant sites on the SARS-CoV-2
spike receptor-binding domain by structure-guided high-resolution
serology. Cell. 2020;183(4):1024–42. e21.

38. Robbiani DF, Gaebler C, Muecksch F, Lorenzi JC, Wang Z, Cho A, et al.
Convergent antibody responses to SARS-CoV-2 in convalescent individuals.
Nature. 2020;584(7821):437–42. https://doi.org/10.1038/s41586-020-2456-9.

39. Metzemaekers M, Gouwy M, Proost P. Neutrophil chemoattractant receptors
in health and disease: double-edged swords. Cell Mol Immunol. 2020;17(5):
433–50.

40. Klingler J, Weiss S, Itri V, Liu X, Oguntuyo KY, Stevens C, et al. Role of IgM
and IgA Antibodies in the Neutralization of SARS-CoV-2. medRxiv. 2020.
https://doi.org/10.1101/2020.08.18.20177303.

41. Duhen T, Duhen R, Lanzavecchia A, Sallusto F, Campbell DJ. Functionally
distinct subsets of human FOXP3+ Treg cells that phenotypically mirror
effector Th cells. Blood. 2012;119(19):4430–40. https://doi.org/10.1182/
blood-2011-11-392324.

42. Hope CM, Welch J, Mohandas A, Pederson S, Hill D, Gundsambuu B, et al.
Peptidase inhibitor 16 identifies a human regulatory T-cell subset with
reduced FOXP3 expression over the first year of recent onset type 1
diabetes. Eur J Immunol. 2019;49(8):1235–50. https://doi.org/10.1002/eji.201
948094.

43. Kaplan MH. Th9 cells: differentiation and disease. Immunol Rev. 2013;252(1):
104–15. https://doi.org/10.1111/imr.12028.

44. Eyerich S, Eyerich K, Pennino D, Carbone T, Nasorri F, Pallotta S, et al. Th22
cells represent a distinct human T cell subset involved in epidermal
immunity and remodeling. J Clin Invest. 2009;119(12):3573–85. https://doi.
org/10.1172/JCI40202.

45. Mohr I, Sonenberg N. Host translation at the nexus of infection and
immunity. Cell Host Microbe. 2012;12(4):470–83. https://doi.org/10.1016/j.
chom.2012.09.006.

46. Stern-Ginossar N, Thompson SR, Mathews MB, Mohr I. Translational control
in virus-infected cells. Cold Spring Harbor Perspect Biol. 2019;11(3):a033001.
https://doi.org/10.1101/cshperspect.a033001.

47. Schubert K, Karousis ED, Jomaa A, Scaiola A, Echeverria B, Gurzeler LA, et al. SARS-
CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation. Nat Struct
Mol Biol. 2020;27(10):959–66. https://doi.org/10.1038/s41594-020-0511-8.

48. Huang C, Lokugamage KG, Rozovics JM, Narayanan K, Semler BL, Makino S.
SARS coronavirus nsp1 protein induces template-dependent
endonucleolytic cleavage of mRNAs: viral mRNAs are resistant to nsp1-
induced RNA cleavage. PLoS Pathogens. 2011;7(12):e1002433. https://doi.
org/10.1371/journal.ppat.1002433.

49. Bianco C, Mohr I. Ribosome biogenesis restricts innate immune responses to
virus infection and DNA. Elife. 2019;8:e49551. https://doi.org/10.7554/eLife.49551.

50. Rampotas A, Pavord S. Platelet aggregates, a marker of severe COVID-19
disease. J Clin Pathol. 2020;74(11):750–1. https://doi.org/10.1136/jclinpath-2
020-206933.

51. Reinecke F, Smeitink JA, Van Der Westhuizen FH. OXPHOS gene expression
and control in mitochondrial disorders. Biochimica et Biophysica Acta (BBA)-
Mol Basis Dis. 2009;1792(12):1113–21.

52. Apaijai N, Sriwichaiin S, Phrommintikul A, Jaiwongkam T, Kerdphoo S,
Chansirikarnjana S, et al. Cognitive impairment is associated with mitochondrial
dysfunction in peripheral blood mononuclear cells of elderly population. Sci Rep.
2020;10(1):21400. https://doi.org/10.1038/s41598-020-78551-4.

Ryan *et al. BMC Medicine*   (2022) 20:26

Page 22 of 23

53. Bhattacharyya S, Vrati S. The Malat1 long non-coding RNA is upregulated by signalling through the PERK axis of unfolded protein response during flavivirus infection. Sci Rep. 2015;5(1):17794. https://doi.org/10.1038/srep17794.

54. Huang K, Wang C, Vagts C, Raguveer V, Finn PW, Perkins DL. LncRNAs NEAT1 and MALAT1 Differentiate Inflammation in Severe COVID-19 Patients. medRxiv. 2021. https://doi.org/10.1101/2021.03.26.21254445.

55. Hewitson JP, West KA, James KR, Rani GF, Dey N, Romano A, et al. Malat1 suppresses immunity to infection through promoting expression of Maf and IL-10 in Th cells. J Immunol. 2020;204(11):2949–60.

56. Gutschner T, Hämmerle M, Eissmann M, Hsu J, Kim Y, Hung G, et al. The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. Cancer Res. 2013;73(3):1180–9. https://doi.org/10.1158/0008-5472.CAN-12-2850.

57. Bhattacharjee S, Banerjee M. Immune thrombocytopenia secondary to COVID-19: a systematic review. SN Compr Clin Med. 2020;2(11):2048–58. https://doi.org/10.1007/s42399-020-00521-8.

58. Chen W, Yang B, Li Z, Wang P, Chen Y, Zhou H. Sudden severe thrombocytopenia in a patient in the recovery stage of COVID-19. Lancet Haematol. 2020;7(8):e624. https://doi.org/10.1016/S2352-3026(20)30175-7.

59. Bomhof G, Mutsaers PGNJ, Leebeek FWG, te Boekhorst PAW, Hofland J, Croles FN, et al. COVID-19-associated immune thrombocytopenia. Br J Haematol. 2020;190(2):e61–e4. https://doi.org/10.1111/bjh.16850.

60. Yang M, Ng MH, Li CK. Thrombocytopenia in patients with severe acute respiratory syndrome (review). Hematology (Amsterdam, Netherlands). 2005;10(2):101–5.

61. Joob B, Wiwanitkit V. Thrombocytopenia: an important presentation of new emerging H7N9 influenza. Platelets. 2014;25(4):308.

62. Sharp TM, Muñoz-Jordán J, Perez-Padilla J, Bello-Pagán MI, Rivera A, Pastula DM, et al. Zika virus infection associated with severe thrombocytopenia. Clin Infect Dis. 2016;63(9):1198–201.

63. Korfias S, Stranjalis G, Papadimitriou A, Psachoulia C, Daskalakis G, Antsaklis A, et al. Serum S-100B protein as a biochemical marker of brain injury: a review of current concepts. Curr Med Chem. 2006;13(30):3719–31. https://doi.org/10.2174/092986706779026129.

64. Al-Aly Z, Xie Y, Bowe B. High-dimensional characterization of post-acute sequalae of COVID-19. Nature. 2021;594(7862):259–64. https://doi.org/10.1038/s41586-021-03553-9.

65. Koutsakos M, Rowntree LC, Hensen L, Chua BY, van de Sandt CE, Habel JR, et al. Integrated immune dynamics define correlates of COVID-19 severity and antibody responses. Cell Rep Med. 2021;2(3):100208. https://doi.org/10.1016/j.xcrm.2021.100208.

66. Liu J, Yang X, Wang H, Li Z, Deng H, Liu J, et al. Analysis of the long-term impact on cellular immunity in COVID-19-recovered individuals reveals a profound NKT cell impairment. mBio. 2021;12(2):e00085–21.

67. Wilk AJ, Lee MJ, Wei B, Parks B, Pi R, Martinez-Colon GJ, et al. Multi-omic profiling reveals widespread dysregulation of innate immunity and hematopoiesis in COVID-19. J Exp Med. 2021;218(8):e20210582. https://doi.org/10.1084/jem.20210582.

68. Kratzer B, Trapin D, Ettel P, Kormoczi U, Rottal A, Tuppy F, et al. Immunological imprint of COVID-19 on human peripheral blood leukocyte populations. Allergy. 2021;76(3):751–65. https://doi.org/10.1111/all.14647.

69. Peluso MJ, Deitchman AN, Torres L, Iyer NS, Munter SE, Nixon CC, et al. Long-term SARS-CoV-2-specific immune and inflammatory responses in individuals recovering from COVID-19 with and without post-acute symptoms. Cell Rep. 2021;36(6):109518. https://doi.org/10.1016/j.celrep.2021.109518.

70. Cheon IS, Li C, Son YM, Goplen NP, Wu Y, Cassmann T, et al. Immune signatures underlying post-acute COVID-19 lung sequelae. Sci Immunol. 2021;6(65):eabk1741.

71. Pradenas E, Trinite B, Urrea V, Marfil S, Avila-Nieto C, de la Concepcion ML R, et al. Stable neutralizing antibody levels 6 months after mild and severe COVID-19 episodes. Med (N Y). 2021;2(3):313–20 e4.

72. Wu J, Liang B, Chen C, Wang H, Fang Y, Shen S, et al. SARS-CoV-2 infection induces sustained humoral immune responses in convalescent patients following symptomatic COVID-19. Nat Commun. 2021;12(1):1813.

73. Bonifacius A, Tischer-Zimmermann S, Dragon AC, Gussarow D, Vogel A, Krettek U, et al. COVID-19 immune signatures reveal stable antiviral T cell function despite declining humoral responses. Immunity. 2021;54(2):340–54 e6. https://doi.org/10.1016/j.immuni.2021.01.008.

74. Wheatley AK, Juno JA, Wang JJ, Selva KJ, Reynaldi A, Tan HX, et al. Evolution of immune responses to SARS-CoV-2 in mild-moderate COVID-19. Nat Commun. 2021;12(1):1162. https://doi.org/10.1038/s41467-021-21444-5.

75. Bilich T, Nelde A, Heitmann JS, Maringer Y, Roerden M, Bauer J, et al. T cell and antibody kinetics delineate SARS-CoV-2 peptides mediating long-term immune responses in COVID-19 convalescent individuals. Sci Transl Med. 2021;13(590):eabf7517.

76. Harrington WE, Trakhimets O, Andrade DV, Dambrauskas N, Raappana A, Jiang Y, et al. Rapid decline of neutralizing antibodies is associated with decay of IgM in adults recovered from mild COVID-19. Cell Rep Med. 2021;2(4):100253. https://doi.org/10.1016/j.xcrm.2021.100253.

77. Gasser R, Cloutier M, Prevost J, Fink C, Ducas E, Ding S, et al. Major role of IgM in the neutralizing activity of convalescent plasma against SARS-CoV-2. Cell Rep. 2021;34(9):108790.

78. Yates JL, Ehrbar DJ, Hunt DT, Girardin RC, Dupuis AP 2nd, Payne AF, et al. Serological analysis reveals an imbalanced IgG subclass composition associated with COVID-19 disease severity. Cell Rep Med. 2021;15(7):100329. https://doi.org/10.1016/j.xcrm.2021.100329.

79. Haslwanter D, Dieterle ME, Wec AZ, O'Brien CM, Sakharkar M, Florez C, et al. A combination of RBD and NTD neutralizing antibodies limits the generation of SARS-CoV-2 spike neutralization-escape mutants. bioRxiv. 2021. https://doi.org/10.1128/mBio.02473-21.

80. Wang L, Zhao J, Nguyen LNT, Adkins JL, Schank M, Khanal S, et al. Blockade of SARS-CoV-2 spike protein-mediated cell-cell fusion using COVID-19 convalescent plasma. Sci Rep. 2021;11(1):5558. https://doi.org/10.1038/s41598-021-84840-3.

81. Stephenson E, Reynolds G, Botting RA, Calero-Nieto FJ, Morgan MD, Tuong ZK, et al. Single-cell multi-omics analysis of the immune response in COVID-19. Nat Med. 2021;27(5):904–16.

82. Pusnik J, Richter E, Schulte B, Dolscheid-Pommerich R, Bode C, Putensen C, et al. Memory B cells targeting SARS-CoV-2 spike protein and their dependence on CD4(+) T cell help. Cell Rep. 2021;35(13):109320. https://doi.org/10.1016/j.celrep.2021.109320.

83. Galván-Peña S, Leon J, Chowdhary K, Michelson DA, Vijaykumar B, Yang L, et al. Profound Treg perturbations correlate with COVID-19 severity. Proc Natl Acad Sci. 2021;118(37):e2111315118. https://doi.org/10.1073/pnas.2111315118.

84. Boothby IC, Cohen JN, Rosenblum MD. Regulatory T cells in skin injury: at the crossroads of tolerance and tissue repair. Sci Immunol. 2020;5(47):eaaz9631.

85. Patil VS, Madrigal A, Schmiedel BJ, Clarke J, O'Rourke P, de Silva AD, et al. Precursors of human CD4(+) cytotoxic T lymphocytes identified by single-cell transcriptome analysis. Sci Immunol. 2018;3(19):eaan8664.

86. Tian Y, Babor M, Lane J, Schulten V, Patil VS, Seumois G, et al. Unique phenotypes and clonal expansions of human CD4 effector memory T cells re-expressing CD45RA. Nat Commun. 2017;8(1):1473. https://doi.org/10.1038/s41467-017-01728-5.

87. Linterman MA, Pierson W, Lee SK, Kallies A, Kawamoto S, Rayner TF, et al. Foxp3+ follicular regulatory T cells control the germinal center response. Nat Med. 2011;17(8):975–82. https://doi.org/10.1038/nm.2425.

88. Lund JM, Hsing L, Pham TT, Rudensky AY. Coordination of early protective immunity to viral infection by regulatory T cells. Science. 2008;320(5880):1220–4. https://doi.org/10.1126/science.1155209.

89. Juno JA, Tan HX, Lee WS, Reynaldi A, Kelly HG, Wragg K, et al. Humoral and circulating follicular helper T cell responses in recovered patients with COVID-19. Nat Med. 2020;26(9):1428–34. https://doi.org/10.1038/s41591-020-0995-0.

90. Kaneko N, Kuo HH, Boucau J, Farmer JR, Allard-Chamard H, Mahajan VS, et al. Loss of Bcl-6-expressing T follicular helper cells and germinal centers in COVID-19. Cell. 2020;183(1):143–57.e13.

91. Overmyer KA, Shishkova E, Miller IJ, Balnis J, Bernstein MN, Peters-Clarke TM, et al. Large-scale multi-omic analysis of COVID-19 severity. Cell Syst. 2021;12(1):23–40 e7. https://doi.org/10.1016/j.cels.2020.10.003.

92. Aschenbrenner AC, Mouktaroudi M, Kramer B, Oestreich M, Antonakos N, Nuesch-Germano M, et al. Disease severity-specific neutrophil signatures in blood transcriptomes stratify COVID-19 patients. Genome Med. 2021;13(1):7. https://doi.org/10.1186/s13073-020-00823-5.

93. Manne BK, Denorme F, Middleton EA, Portier I, Rowley JW, Stubben C, et al. Platelet gene expression and function in patients with COVID-19. Blood. 2020;136(11):1317–29. https://doi.org/10.1182/blood.2020007214.

94. Koupenova M, Corkrey HA, Vitseva O, Tanriverdi K, Somasundaran M, Liu P, et al. SARS-CoV-2 initiates programmed cell death in platelets. Circ Res. 2021;129(6):631–46. https://doi.org/10.1161/CIRCRESAHA.121.319117.

95. Cameron B, Galbraith S, Zhang Y, Davenport T, Vollmer-Conna U, Wakefield D, et al. Gene expression correlates of postinfective fatigue syndrome after infectious mononucleosis. J Infect Dis. 2007;196(1):56–66. https://doi.org/10.1086/518614.

96. Saichi M, Ladjemi MZ, Korniotis S, Rousseau C, Ait Hamou Z, Massenet-Regad L, et al. Single-cell RNA sequencing of blood antigen-presenting cells in severe COVID-19 reveals multi-process defects in antiviral immunity. Nat Cell Biol. 2021;23(5):538–51. https://doi.org/10.1038/s41556-021-00681-2.

97. Tarke A, Sidney J, Kidd CK, Dan JM, Ramirez SI, Yu ED, et al. Comprehensive analysis of T cell immunodominance and immunoprevalence of SARS-CoV-2 epitopes in COVID-19 cases. Cell Rep Med. 2021;2(2):100204. https://doi.org/10.1016/j.xcrm.2021.100204.

98. Sekine T, Perez-Potti A, Rivera-Ballesteros O, Stralin K, Gorin JB, Olsson A, et al. Robust T cell immunity in convalescent individuals with asymptomatic or mild COVID-19. Cell. 2020;183(1):158–68 e14. https://doi.org/10.1016/j.cell.2020.08.017.

99. Gene Expression Omnibus https://identifiers.org/geo/GSE169687. (2021).

Publisher's Note