



OPEN ACCESS

EDITED BY

Juan Bautista De Sanctis,
Palacký University Olomouc, Czechia

REVIEWED BY

Alexis Hipólito García,
Central University of Venezuela, Venezuela
Alexandra Vujkovic,
Institute of Tropical Medicine Antwerp,
Belgium

*CORRESPONDENCE

David S. Gibson

✉ d.gibson@ulster.ac.uk

Priyank Shukla

✉ p.shukla@ulster.ac.uk

[†]These authors have contributed
equally to this work and share
first authorship

RECEIVED 21 December 2024

ACCEPTED 20 February 2025

PUBLISHED 02 April 2025

CITATION

Markey G, McLaughlin J, McDaid D,
Lynch SM, English A, Alexander HD, Kelly M,
Bhavsar M, McGilligan V, Zhang S-D,
Murray EK, Rai TS, Walsh C, Bjourson AJ,
Shukla P and Gibson DS (2025) Distinct
Omicron longitudinal memory T cell profile
and T cell receptor repertoire associated with
COVID-19 hospitalisation.
Front. Immunol. 16:1549570.
doi: 10.3389/fimmu.2025.1549570

COPYRIGHT

© 2025 Markey, McLaughlin, McDaid, Lynch,
English, Alexander, Kelly, Bhavsar, McGilligan,
Zhang, Murray, Rai, Walsh, Bjourson, Shukla
and Gibson. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Distinct Omicron longitudinal memory T cell profile and T cell receptor repertoire associated with COVID-19 hospitalisation

Gavin Markey^{1†}, Joseph McLaughlin^{1,2†}, Darren McDaid¹,
Seodhna M. Lynch¹, Andrew English^{1,3}, H. Denis Alexander¹,
Martin Kelly⁴, Manav Bhavsar⁴, Victoria McGilligan¹,
Shu-Dong Zhang¹, Elaine K. Murray¹, Taranjit Singh Rai¹,
Colum Walsh^{5,6}, Anthony J. Bjourson¹, Priyank Shukla^{1*}
and David S. Gibson^{1*}

¹Personalised Medicine Centre, School of Medicine, Ulster University, Londonderry, United Kingdom, ²Medical Directorate, Clinical and Translational Research and Innovation Centre, Londonderry, United Kingdom, ³School of Health and Life Sciences, Teesside University, Middlesborough, United Kingdom, ⁴Intensive Care Unit, Western Health Social Care Trust, Londonderry, United Kingdom, ⁵Genomic Medicine Research Group, School of Biomedical Science, Ulster University, Coleraine, United Kingdom, ⁶Biomedical and Clinical Sciences Division, Department for Cell and Neurobiology, Faculty of Medicine, Linköping University, Linköping, Sweden

SARS-CoV-2 has claimed more than 7 million lives worldwide and has been associated with prolonged inflammation, immune dysregulation and persistence of symptoms following severe infection. Understanding the T cell mediated immune response and factors impacting development and continuity of SARS-CoV-2 specific memory T cells is pivotal for developing better therapeutic and monitoring strategies for those most at risk from COVID-19. Here we present a comprehensive analysis of memory T cells in a convalescent cohort (n=20), three months post Omicron infection. Utilising flow cytometry to investigate CD4⁺CD45RO⁺ and CD8⁺CD45RO⁺ memory T cell IL-2 expression following Omicron (B.1.1.529/BA.1) peptide pool stimulation, alongside T cell receptor repertoire profiling and RNA-Seq analysis, we have identified several immunological features associated with hospitalised status. We observed that while there was no significant difference in median CD4⁺CD45RO⁺ IL-2⁺ and CD8⁺ CD45RO⁺ IL-2⁺ memory T cell count between subgroups, the hospitalised subgroup expressed significantly more IL-2 per cell following Omicron peptide pool exposure in the CD8⁺CD45RO⁺ population (p <0.03) and trended towards significance in CD4⁺CD45RO⁺ cells (p <0.06). T cell receptor repertoire analysis found that the non-hospitalised subgroup had a much higher number of circulating clonotypes, targeting a wider range of predominantly MHC-I epitopes across the SARS-CoV-2 genome. Several immunodominant epitopes, conserved between both subgroups, were observed, however hospitalised individuals were less likely to express putative HLA alleles responsible for pMHC presentation which may impact TCR affinity. We observed a bias towards shorter CDR3 segments in TCRβ repertoire analysis within the hospitalised subgroup, alongside lower rates of repertoire overlap in CDR3 sequences compared to the non-hospitalised subgroup. We found a significant proportion of TCRs targeted epitopes along the SARS-CoV-2 genome including

non-structural proteins, responsible for viral replication and immune evasion. These findings highlight how the continuity of T cell based protective immunity is impacted by both the viral replication cycle of SARS-CoV-2 upon intracellular and innate immune responses, and HLA-type upon TCR affinity and clonotype formation. Our novel Epitope Target Analysis Pipeline (Epi-TAP) could prove beneficial in development of new therapeutic strategies through rapid identification of shared immunodominant epitopes across non-hospitalised and hospitalised subgroups.

KEYWORDS

SARS-CoV-2, COVID-19, T cell immunity, T cell receptor (TCR) recognition, antigen presentation, adaptive immunity, immuno-informatics

Introduction

COVID-19 is a highly infectious and potentially fatal disease that has claimed more than 7 million lives worldwide over the past five years (1). In severe cases, COVID-19 is known to cause prolonged viral pneumonia, acute respiratory distress syndrome (ARDS), and potentially fatal multi-organ failure. Severity of infection is often linked to dysregulation of immune responses following activation of NF- κ B and MAPK signalling pathways. SARS-CoV-2 viral proteins have been shown to disrupt type I and III interferon (IFN) production, cleaving transcription factor IRF3 and inhibiting mRNA transcription. During infection activation of inflammasome sensor NLRP3 leads to cleavage of gasdermin D (GSDMD) inducing pyroptosis, releasing IL-1 β and IL-18 leading to further cytokine production. Elevated levels of IL-1 β and IL-18 correlate with disease severity and poorer clinical outcomes (2, 3).

In the early stages of infection macrophages, dendritic cells and natural killer (NK) cells are recruited to the site of infection, inducing the production of pro-inflammatory cytokines and chemokines that attract monocytes, macrophages and T cells. Recruited CD4⁺ and CD8⁺ T cells in turn produce pro-inflammatory proteins, resulting in the recruitment and proliferation of various T helper (Th) and cytotoxic T cells (CTLs) leading to the production of antibodies by B cells working in synergy with Th cells, and memory T cells that target and eliminate the virus should reinfection occur (4). Although several studies have reported on the duration, quality and factors influencing T cell based protective immunity to SARS-CoV-2, there is limited longitudinal evidence of these features association with Omicron (B.1.1.529) acute phase infection severity and hospitalisation. We therefore investigated if differences existed in IL-2 expression as a key cytokine promoting memory T cell persistence, CD158b expression as a key modulator of immune response, and T cell receptor repertoire in hospitalised and non-hospitalised cases. Interleukin-2 (IL-2) is a critical cytokine primarily produced by activated T cells, particularly CD4⁺ T-

helper cells. It plays a central role in T cell proliferation, differentiation, and function, essential for a strong immune response (5). IL-2 promotes the expansion, survival, and effectiveness of T cells after they recognise an antigen, ensuring a sustained immune response (6). IL-2 is elevated along with other pro-inflammatory cytokines in severe COVID-19 (7). Rapid IL-2 production following T cell activation serves as a key marker for evaluating immune responses, particularly during infections such as SARS-CoV-2. When T cells encounter SARS-CoV-2, they secrete IL-2, triggering the activation and proliferation of other immune cells. This early production of IL-2 is crucial for assessing the magnitude and quality of the immune response, providing insights into T cell based immune memory (8, 9). IL-2 levels can indicate the presence of memory T cells, which are vital for quick and effective responses to future infections (9). Elevated IL-2 production is crucial for promoting the proliferation and persistence of memory T cells, enhancing the body's capacity to manage persistent infection or inflammation (10, 11). Furthermore, assessing IL-2 production is useful for evaluating polyfunctional T cells following vaccination and natural exposure to SARS-CoV-2 (12, 13).

CD158b, also known as KIR2DL2/L3/S2, is a receptor typically associated with natural killer (NK) cells but is also expressed on a subset of T cells, where it plays a role in modulating immune responses, often through inhibitory signals (14). These receptors play a crucial role in regulating immune responses by providing inhibitory signals that help to balance immune activity and prevent excessive tissue damage (15, 16). Pioneering research on SARS-CoV-1 infection revealed that both the number of circulating NK cells and the expression of inhibitory KIR CD158b were lower compared to those observed in healthy individuals and patients with *Mycoplasma pneumoniae* infection. This reduction was linked to greater disease severity and the presence of antibodies specific to SARS-CoV-1 (17). In the context of COVID-19, CD158b and its associated receptors are of interest, as SARS-CoV-2 induces intricate immune responses that engage both innate and adaptive immunity. NK cells, which express KIRs like CD158b, are vital to the innate immune system and are instrumental in the initial

response to viral infections (18). CD158b has been found at lower frequency in acute phase patients with severe COVID-19 (18). However, the longitudinal role of CD158b receptors in T cells post SARS-CoV-2 infection has not been investigated.

During infection, antigen presenting cells (APCs) present short peptide sequences known as epitopes to T cell receptors (TCRs) via class I and class II major histocompatibility complex (MHC) proteins. Class I MHC proteins are found on the surface of all nucleated cells and present epitopes of 8-10 amino acids (aa) in length to cytotoxic CD8⁺ T cells. Class II MHC proteins are only present on specialised antigen presenting cells (dendritic cells, macrophages, B cells) and predominantly present to CD4⁺ T cells (19). In both cases the peptide MHC (pMHC) interacts with TCRs in a conserved diagonal docking modality that is believed to have arisen from co-evolution of TCR and MHC genes (20, 21).

TCRs are dimeric structures comprised of either alpha-beta ($\alpha\beta$) or gamma-delta ($\gamma\delta$) chains. $\alpha\beta$ chains are the dominant subgroup and are found on $\geq 95\%$ of all T lymphocytes and are formed through somatic recombination of gene segments during thymocyte development (22, 23). There are a total of 70 variable (V) and 61 joining (J) genes involved in α chain rearrangement and 52 V and 13 J genes alongside 2 diversity (D) genes in β chain rearrangement (24). This process, known as V(D)J recombination can theoretically give rise to more than 2.5×10^7 unique TCRs (25), yet due to MHC-Restriction we often identify overlaps in TCR repertoires for a variety of foreign and self-antigens that can provide valuable insights into previous exposure and persistence of protective immunity.

In this paper we investigate persistence of T cell based protective immunity by restimulation with SARS-CoV-2 peptide pools and comprehensive TCR repertoire analysis, in hospitalised and non-hospitalised subgroups. Here we describe a novel immunoinformatics pipeline for TCR repertoire analysis using publicly

available tools, for rapid identification and investigation of epitope presentation and clonotype formation. Our novel Epitope Target Analysis Pipeline (Epi-TAP) is capable of discerning unique biological features between subgroups including the presence of immunodominant epitopes, meta-clonotype analysis, severity associated HLA-types, differential expressed genes (DEGs) and pathway activation following viral peptide stimulation. This comprehensive approach could aid in developing future diagnostic and therapeutic strategies against SARS-CoV-2 and future pathogens (Figure 1).

Methods

Participant inclusion criteria and ethical approval

The following inclusion criteria were used for the current study: 1) Participants were ≥ 18 years of age at time of recruitment, with any body mass index (BMI) or ethnicity, 2) All participants had a positive PCR test for SARS-CoV-2 at recruitment, 3) Non-hospitalised participants recovered from acute infection without admission to hospital, 4) Hospitalised participants were admitted to hospital during the acute infection, 5) Participants were excluded if < 18 years of age or with intellectual disabilities or mental health illness. Ethical approval was obtained from the Health and Care Research Wales Ethics committee (20/WA/0179; 14/07/2020). Informed consent was obtained for all participants in the study, allowing for publication of anonymised demographic and clinical data. The COVID-19 Research Study (COVRES) protocol has further details on study design, data and sample collection and analyses performed (ClinicalTrials.gov NCT0548829) (26).

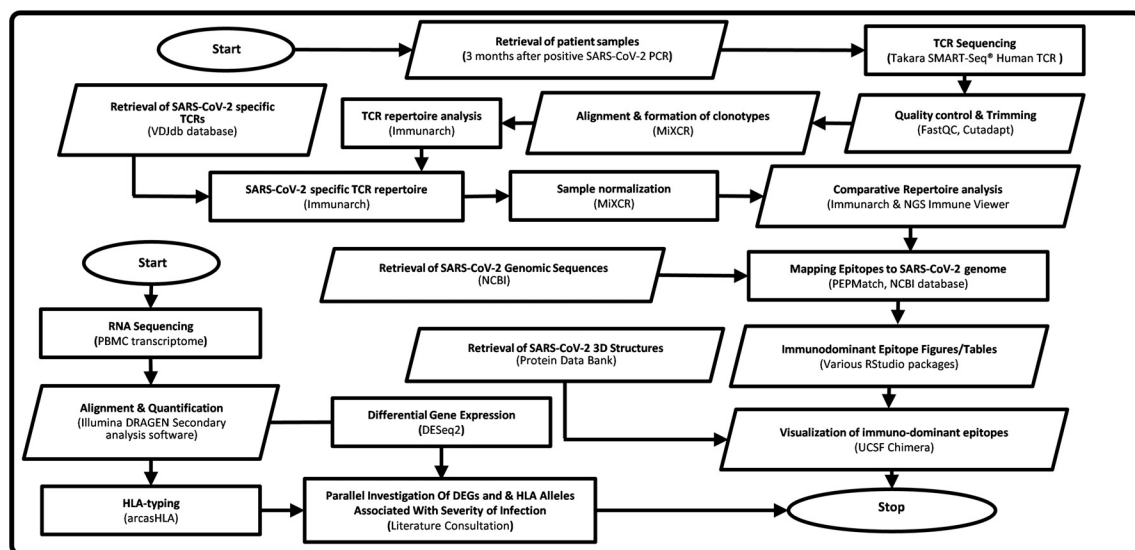


FIGURE 1

Flowchart of Epi-TAP: Epitope Target Analysis Pipeline (Version 1.0). Oval shapes represent start/stop of the pipeline. Parallelogram boxes represent input/output. Rectangular boxes represent processing steps. Servers/tools/software and databases used are mentioned in parenthesis.

Participant recruitment

Participants (n=20) were recruited between October 2021 - May 2022 during which the predominant variant of concern detected in UK during this period was Omicron B.1.1.529 (which includes sub lineages BA.1, BA.2, BA.4 and BA.5).

Sample collection and processing

Blood was collected in EDTA tubes using 21G Vacuette® safety needles (Greiner Bio-One Ltd, UK) and processed no more than 2 hours post blood draw in a biosafety level 2 laboratory. Peripheral blood mononuclear cells (PBMCs) were isolated using the ficoll gradient separation method. Whole blood collected in EDTA tubes, 18ml was separated using ficoll-plaque plus (Cytiva, UK) by spinning at 400 x g for 30 min, and PBMCs were collected at the interface layer. Cells were washed and counted for recovery and viability using trypan blue, the Countess Automated Cell Counter and slides as per manufacturer instructions (Thermo Scientific, UK). Aliquots of PBMCs were cryopreserved in human serum + 10% DMSO (Sigma-Aldrich; UK) at -80°C for 24 h then transferred to liquid nitrogen for long term storage.

Full blood counts

A sample of whole blood (WB) was also used to measure full blood counts (FBC) with a standard flow detection Sysmex XE-instrument in the pathology laboratories at Altnagelvin Area Hospital, Londonderry. Clinical data were recovered from the Northern Ireland Electronic Care Records (NIECR).

Omicron peptide pool and peripheral blood mononuclear cell culture

Peripheral blood mononuclear cells (PBMCs) were cultured and stimulated *in vitro*. Briefly, the cryopreserved PBMCs were thawed and washed, followed by seeding at 1×10^6 cells per well in RPMI + 5% human AB serum (Gibco, Thermo Scientific, UK). Cells were stimulated with a cocktail of peptides consisting of SARS-CoV Omicron variant Prot_S B.1.1.529/BA.1 Mutation pool (#130-129-928) and SARS-CoV Wuhan wild-type Prot_N (#130-126-698) (Miltenyi Biotec, UK) at a concentration of 1 µg/ml of each peptide pool in RPMI + 5% AB serum for 24 hrs, after which cells were harvested and analysed by immunocytochemistry. Cells treated with no peptide (unstimulated cells) were used for baseline negative control readings. Cells treated with phytohemagglutinin (PHA) were included as a positive control. Cell viability was assessed using trypan blue staining method.

Flow cytometry cell staining and analysis

PBMCs harvested from *in vitro* experiments were stained with fluorophore tagged monoclonal antibodies. Two antibody panels (TBNK and COVRES2) were used for this analysis (Ig isotype, dye, clone and product # are listed for each). TBNK: FITC mouse IgG2a anti-human CD3 (clone OKT3, #566783), PE-Cy7 mouse IgG1 anti-human CD4 (clone SK3 #557852), APC-Cy7 mouse IgG1 anti-human CD8 (clone SK1, #557834), PE mouse IgG1 anti-human CD16 (clone 3G8 #555407), APC mouse IgG1 anti-human CD19 (clone HIB19 #555415), PerCP-Cy5.5 mouse IgG1 anti-human CD45 (clone HI30, #564105), PE mouse IgG1 anti-human CD56 (clone B159 #555516). COVRES2: FITC mouse IgG2a anti-human CD3 (clone OKT3, #566783), APC mouse IgG1 anti-human CD4 (clone SK3, #566915), BV605 mouse IgG1 anti-human CD8 (clone SK1, #564116), R718 mouse IgG1 anti-human CD16 (clone 3G8, #566970), PE-Cy7 mouse IgG1 anti-human CD56 (clone B159, #557747), BV421 mouse IgG2b anti-human CD45RA (clone HI100, #562885), APC-H7 mouse IgG2b anti-human CD45RO (clone UCHL1, #561137), BB700 mouse IgG2b anti-human CD158b (clone CH-L, #561137), PE mouse IgG1 anti-human IL-2 (clone 5344.111, #569370) antibodies, before analysis by a FACS Aria III instrument (all BD Biosciences, UK). The FACS Aria III was operated using FACSDiva software (v9.4; BD). One million (1×10^6) cells in staining buffer were stained with above antibodies according to the manufacturer's instructions. For surface staining, cells were incubated with relevant fluorochrome-labelled antibodies for 30 min at room temperature in the dark. For cells to be both surface and intracellular stained, prior to staining, these cells were also incubated with a transport inhibitor, BD GolgiPlug™ containing brefeldin A, for 4 hrs at 37°C, BD Cytofix/Cytoperm Plus kit (BD Biosciences, UK), following surface staining cells were fixed and permeabilised using the BD Cytofix/Cytoperm Plus kit (BD Biosciences, UK) and stained with PE-anti-IL-2 for 30 min at 4°C in the dark. All samples were analysed immediately after staining. In each instance, 10,000 events were recorded in triplicate for each individual well/sample. Antibody isotype control, unstained control, cell viability control (see cell culture above) were included in each experiment. Isotype controls were Ig isotype and fluorophore matched antibodies raised against non-human target; viability control was cell samples stained by trypan blue.

Flow cytometry gating strategy

The gating strategy for flow cytometry analysis is shown in [Figure 2](#). Lymphocyte populations were gated based on forward scatter area (FSC-A) and low side scatter area (SSC-A). Where the FCS-A event size was less than 80,000 the population was predominantly cell debris and excluded from analysis. The

lymphocyte population was further gated by duplet exclusion (FSC-A, area vs FSC-W, width and SSC-A, area vs SSC-W, width). The T cell subpopulations were identified based on T cells specific markers, T helper ($CD3^+CD4^+$) and cytotoxic T cell ($CD3^+CD8^+$) markers. The T helper and cytotoxic T cell populations were further gated for $CD45RO^+$ (memory) T cells. $CD45RO^+$ T helper and cytotoxic T cell populations were further gated for $CD158b^+$ cell surface and intracellular IL-2⁺ expression, as shown in **Figure 2**.

Procedure for fluorochrome spillover compensation

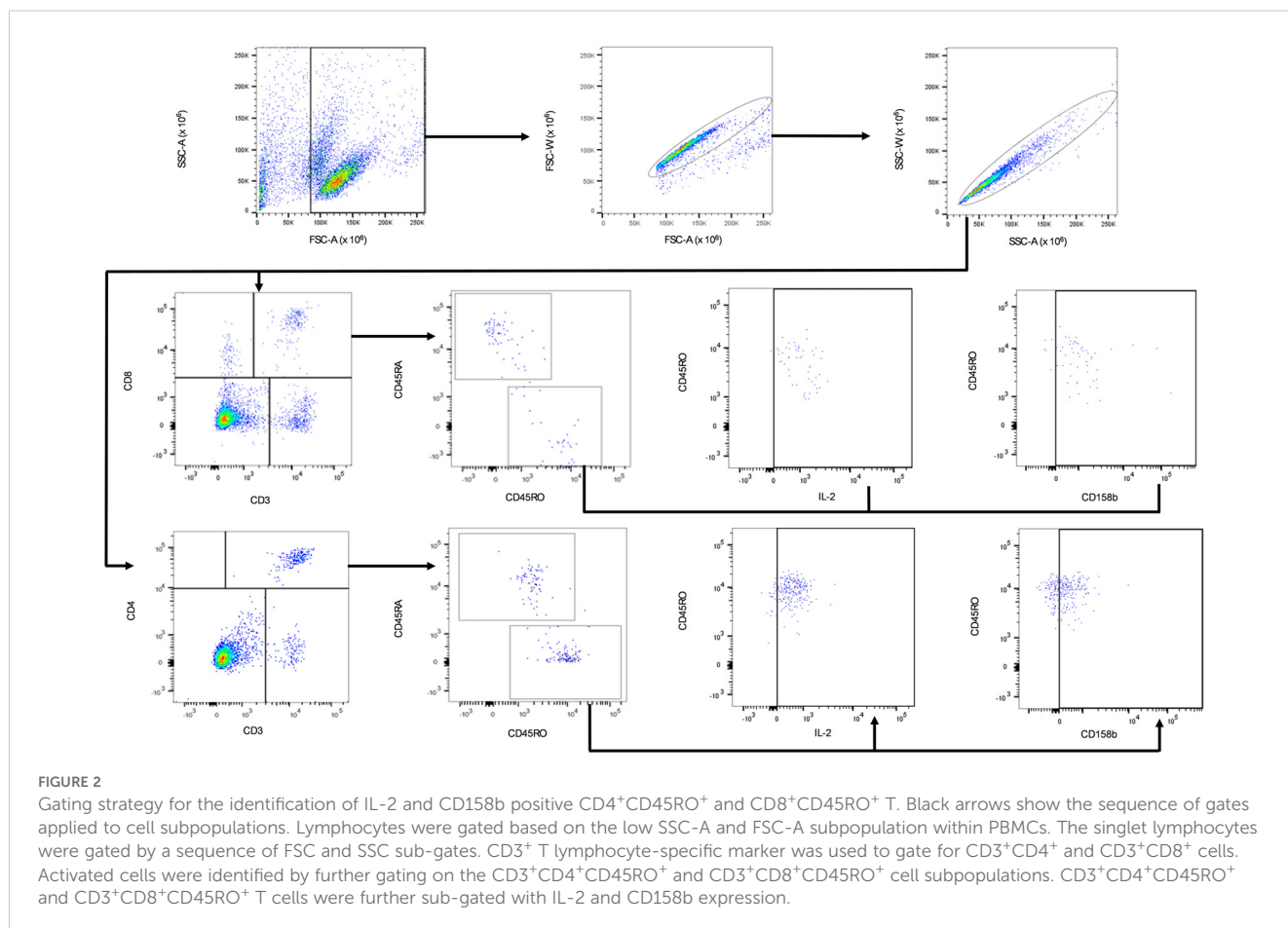
To calculate the compensation matrix and to correct the reciprocal spillover among fluorochromes with overlapping emission spectra, anti-mouse (#552843) Ig-Kappa Comp Beads were stained accordingly to the manufacturer's conditions with the same mouse Igk anti-human antibodies used in the panels and immediately acquired using BD FACSDiva software (BD Biosciences version 9.4). The compensation matrix was automatically calculated by BD FACSDiva software. FACSARIA III flow cytometer performance was validated daily with BD FACSDiva CS&T research beads (#655051) following the manufacturer's instructions (BD Biosciences, UK).

T cell receptor sequencing

Cryopreserved PBMCs were thawed and washed in Mg^{2+} and Ca^{2+} free PBS prior to RNA extraction (Qiagen RNeasy RNA extraction kit). Bulk unpaired TCR sequencing was performed using an Illumina platform with the TakaraBio SMART-Seq Human TCR (with UMIs) and Unique Dual Index (UDI) kits (Cat. 634779, 634725). Total RNA concentrations were calculated using QubitTM RNA High Sensitivity (HS) kits (Q10210). Sample integrity was investigated using Agilent Fragment Analyzer - High Sensitivity RNA kit (DNF-472-0500). Total RNA concentrations between 2-40ng were used for library preparation using the Illumina Stranded mRNA Prep Ligation kit. PCR amplification was run for a total of 20 cycles after which libraries were normalised to account for variations in concentration prior to pooling.

Quality control & adapter trimming

FastQC (v 0.11.9) (27) was first used to check the quality of sequencing output files and identify the threshold required to remove poor quality reads and truncated sequences. Cutadapt (v 4.8) (28) was then used to remove adapters flagged by FastQC, trim sequences based on their length and quality and remove poor



quality reads and truncated sequences. Sequences <200bp long were removed as downstream analysis requires full coverage of the V(D)J region for alignment and clonotype assembly. Following trimming FastQC was once again used to ensure all unwanted adapters and poor-quality reads were removed.

TCR repertoire assembly

Repertoire assembly was performed using MiXCR immune profiling software (v 4.4.2) (29). Paired-end fastq files were analysed using the built-in analyse preset for the SMART-Seq Human TCR (with UMIs) kit. The preset leverages the SMART (Switching mechanism at 5' end of RNA Template) and unique molecular identifiers (UMIs) to capture complete V(D)J variable regions for α and β chains ($\alpha\beta$ TCRs) for accurate clonotype assembly and quantification. MiXCR default parameters were used for alignment, PCR error correction and clonotype assembly. For samples with off target (non TCR) reads totalling >10%, alignment was repeated with mix-in options to keep non CDR3 alignments and export non-aligned sequences. Seqtk (v 1.4) (30) was then used to convert output fastq files to fasta format for compatibility with National Center for Biotechnology Information (NCBI) databases (31). Fasta sequences were analysed using the basic local alignment search tool (BLAST) to investigate the origin of non TCR reads. After implementing the above described post-alignment quality control checks, productive clones were exported for $\alpha\beta$ TCRs in tsv format for downstream analysis.

SARS-CoV-2 TCR repertoire analysis

A variety of software packages and databases were used to investigate the presence of SARS-CoV-2 specific $\alpha\beta$ TCRs (CoV-TCRs) and explore shared and unique features between subgroups (Figure 1). A number of databases were considered prior to analysis including ImmuneCODETM, Pan immune repertoire database (PIRD), McPAS-TCR and VDJdb (32–35). ImmuneCODETM was disregarded as it focuses solely on TCR β repertoires and was not natively compatible with our immunoinformatics pipeline. PIRD and McPAS-TCR datasets were compatible, however PIRD lacked SARS-CoV-2 specific TCRs and McPAS-TCR contained <200 CoV-TCRs at the time of analysis. VDJdb was both compatible with our pipeline and contained a total of 10,080 CoV-TCRs at the time of analysis. First, publicly available CoV-TCRs were retrieved from the VDJdb browser and imported into RStudio (36) using the Immunarch package (37). Output files from MiXCR analysis were also imported using Immunarch and matched against VDJdb using complementary determining region 3 (CDR3) for both α and β chains to identify the presence of CoV-TCRs. Repertoires that reported CoV-TCRs were taken forward for further analysis. Samples were first downsampled in MiXCR to remove redundant data and account for sequencing bias between samples. Automatic downsampling (with no threshold set in MiXCR) was performed

using the number of reads from the smallest sample, with productive clones exported for further analysis. Following normalisation samples were again imported into RStudio for comparative analysis. TCR repertoires were mapped against VDJdb separately for α and β chains using a custom function that reduced each individual's repertoire to include only TCRs that had an exact match with CDR3 sequences retrieved from VDJdb. An initial investigation of clone counts, clonotype abundance, repertoire overlap, CDR3 length distribution and Kmer analysis of CoV-TCRs was then preformed in RStudio using the Immunarch package.

Repertoire tables were updated in RStudio using a custom function that appended the epitope targeted by the CDR3 region to each. A custom function was then used to combine the repertoire results of each subgroup and extract relevant data to investigate the possibility of unique repertoire features, including chain distribution and pairing, meta-clonotype analysis and presence of immunodominant epitopes. Clonotypes were defined as having a unique clonal sequence at the nucleotide level. If a clonal sequence was observed across multiple individuals, sample data was retained and read counts were combined to accurately calculate the number of individuals targeting each epitope, alongside the overall fraction of the repertoire each epitope represented. The number of individuals with CoV-TCRs targeting each epitope, alongside the number of clonotypes generated for that epitope and percentage of the repertoire they represent were used to investigate the presence of immunodominant epitopes (>5% of total repertoire for either subgroup). Immunodominant epitope tables were exported from RStudio in NGS Immune Viewer (38) compatible format to create chord diagrams for V/J gene pairing and distribution charts. We further investigated differences in gene pairing and distribution by performing two sample t-tests for reported V J genes, using the Benjamini & Hochberg method for false discovery rate correction.

The Protein Data Bank (39) was used to retrieve structural files (Nucleocapsid:8FG2, NSP8:7JLT, NSP3:6WUU, pre-fusion spike:7TGW, post-fusion spike:8FDW) to visualise immunodominant epitopes using UCSF Chimera and obtain the position of structural domains for downstream analysis. PEPMatch (40) was used to determine the position of each epitope in the Omicron BA.2 variant genome (NCBI Accession: OR575624.1) (31) to chart the proteins and structural domains targeted by CoV-TCRs. Results tables then were compiled, detailing the number of individuals and clonotypes targeting each epitope along the SARS-CoV-2 genome and imported into RStudio (36) to visualise the data in bar-chart format.

RNA-Seq analysis

RNA sequencing was performed on PBMC samples using the Illumina Stranded mRNA Prep Ligation kit. Illumina DRAGEN software (V3.10.12) was used for upstream analysis of fastq files. Sequences were first aligned (human GRCh38_alt_aware reference genome) and annotated (Gencode release 44 GRCh38.p14) for

comprehensive gene annotation of reference chromosomes, scaffolds, assembly patches and alternate loci (haplotypes).

DESeq2 (41) was used to identify differentially expressed genes (DEGs). Raw counts and metadata were imported to RStudio with the non-hospitalised subgroup used as the reference level. Following normalisation differential expression analysis was carried out on protein coding genes. Genes were classified as differentially expressed if they had a p-adjusted value <0.05 . Volcano plots and heatmaps for DEGs were then plotted in RStudio using Pheatmap (42) and EnhancedVolcano (43) packages, incorporating hierarchical clustering of both samples and DEGs. Gene ontology (GO) analysis was performed on DEGs using clusterProfiler (44) with a relaxed p-adjusted value <0.10 for biological processes. Enriched GO terms were then plotted using Enrichplot (45) to visualise the relationship between enriched pathways using pathway enrichment and directed acyclic graphs. Output fastq files from RNA-Seq analysis were also subjected to HLA typing using arcasHLA (46). The frequency of HLA genes was then calculated to assess variance between subgroups and how it might affect antigen presentation and T cell differentiation.

Data and statistical analysis

Demographic data statistical analyses were performed using IBM SPSS statistics 28.0.1.1. Chi square test was used to compare differences between groups.

GraphPad Prism software (version 9.4.0) was used to analyse full blood count data. T-tests with Welch's t-tests were used to test differences between groups.

Flow cytometry data analysis and TBNK counts were performed using FlowJo Software (version 10.8.1) (BD Biosciences, UK) on live cells after the exclusion of cell doublets (Figure 2). Manual gating for flow cytometry memory T cell data analysis was performed using FlowJo Software (version 10.8.1) (BD Biosciences, UK) on live cells after the exclusion of cell doublets (Figure 2), data for each participant was harvested and GraphPad Prism software (version 9.4.0) was used to generate scatter dot plots. For each of $n=2$ biological replicates per patient, one million PBMCs were isolated from whole blood and antibody labelled, then analysed by flow cytometry with ten thousand events recorded per technical replicate ($n=3$). The median and IQR event counts were averaged for each patient from the $n=3$ technical replicates from each of the $n=2$ biological replicates. Data from stimulated biological replicates were normalised to respective non-stimulated biological replicates, and the normalised values were used to generate the dot plots in Figures 3, 4. Differences between groups were assessed using Welch's t-tests.

MiXCR immune profiling software (v 4.4.2) was used for TCR repertoire assembly. Downsampling was performed by calculating the 20th quantile across all clonesets (Q20) and normalising the data based on the minimal sample which is above $0.5 \times Q20$.

TCR repertoire analysis was performed using R (v 4.4.1) and RStudio (2024.04.2). SARS-CoV-2 specific repertoire data was exported in csv format for clonotype and meta-clonotype analysis. Chain pairing and CDR3 length distribution analysis was performed using Cogent NGS Immune Viewer available from: <https://www.takarabio.com/products/next-generation-sequencing/bioinformatics-tools/cogent-ngs-immune-viewer>.

Differential expression analysis of RNA-Seq data was performed using DESeq2 package using R (v 4.4.1) and RStudio (2024.04.2).

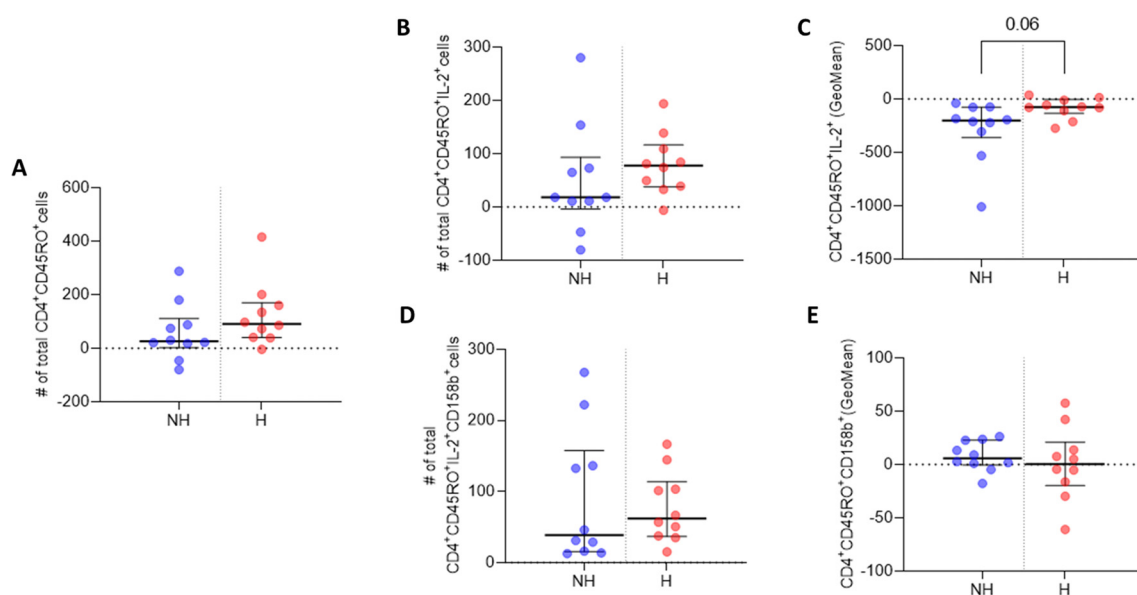


FIGURE 3

IL-2 and 158b expression in $CD4^+CD45RO^+$ T cells. (A) Median number of $CD4^+CD45RO^+$ T cells. (B) Median number of $CD4^+CD45RO^+IL-2^+$ T cell events. (C) GeoMean of IL-2 expression in $CD4^+CD45RO^+IL-2^+$ T cells. (D) Median number of $CD4^+CD45RO^+158b^+$ T cell events. (E) GeoMean of 158b expression in $CD4^+CD45RO^+158b^+$ T cells. All data are presented as Delta values with median and IQR for a given number of observations ($n=10$ non-hospitalised, $n=10$ hospitalised). Welch's t-tests were used for statistical analysis. Data was considered to be significant if $p < 0.05$.

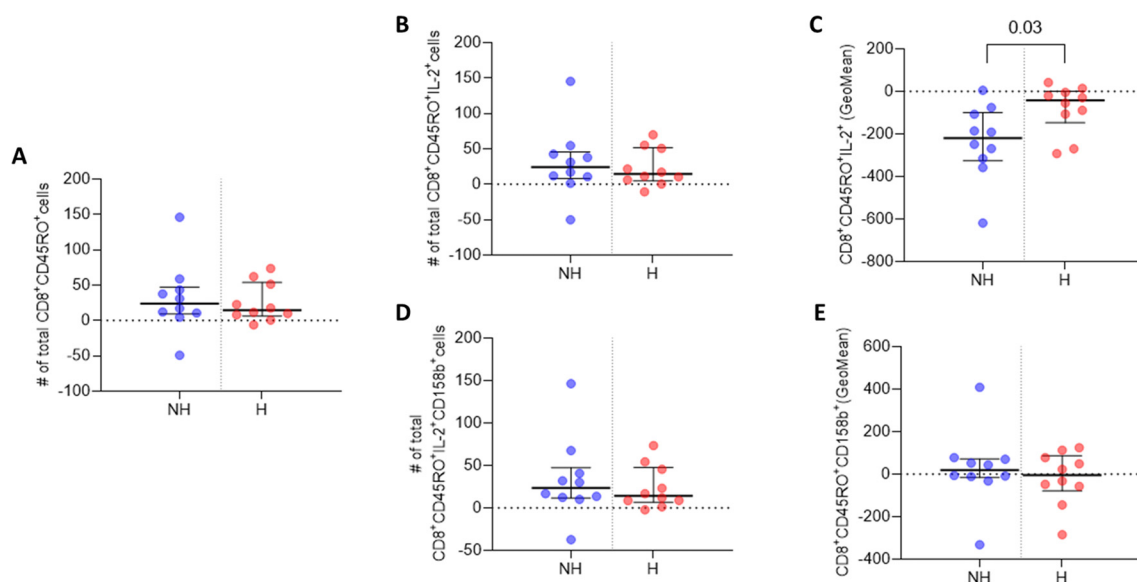


FIGURE 4

IL-2 and 158b expression in CD8⁺CD45RO⁺ T cells. (A) Median number of CD8⁺CD45RO⁺ T cells. (B) Median number of CD8⁺CD45RO⁺IL-2⁺ T cell events. (C) GeoMean of IL-2 expression in CD8⁺CD45RO⁺IL-2⁺ T cells. (D) Median number of CD8⁺CD45RO⁺158b⁺ T cell events. (E) GeoMean of 158b expression in CD8⁺CD45RO⁺158b⁺ T cells. All data are presented as Delta values with median and IQR for a given number of observations (n=10 non-hospitalised, n=10 hospitalised). Welch's t-tests were used for statistical analysis. Data was considered to be significant if $p < 0.05$.

DESeq2 utilises negative binomial regression for differential expression analysis, Wald test for significance (p value) and Benjamini-Hochberg for multiple testing correction (p -adjusted value). Z-score normalisation and hierarchical clustering was used to produce heatmaps from differentially expressed genes and hypergeometric distribution was used to test for significance in gene ontology analysis. HLA gene frequency and distribution analysis was performed using arcasHLA (v 3.24.0) which utilises read alignment and quantification to infer HLA alleles from RNA-Seq data.

A detailed list of the libraries and the statistical methods employed by their functions along with the code is available in the following public repository: <https://github.com/ShuklaLab/Epi-TAP>.

Results

Demographics

The study cohort consisted of non-hospitalised (n=10) and hospitalised (n=10) participants; 50% were female, the median age was 46.5 years, ranging between 35 to 73 years. Hospitalised participants had a significantly higher median age of 53 years, versus the non-hospitalised median age of 46 years ($p < 0.001$). There was no significant difference in vaccination status between subgroups, with non-hospitalised (n=9) and hospitalised (n=9) participants reporting vaccination prior to infection. No significant difference in comorbidities were reported between non-hospitalised and hospitalised participants (Table 1).

Full blood counts

Initially, baseline full blood counts revealed a significant decrease in haemoglobin ($p < 0.02$), haematocrit ($p < 0.02$), and red blood cells (RBC) ($p < 0.01$) among the hospitalised subgroup compared to the non-hospitalised subgroup. However, all other baseline blood markers exhibited no significant differences between the two groups (Table 2). After three months, no significant difference in blood counts were observed between the two subgroups (Table 2).

Flow cytometry analysis, TBNK cell counts

Upon examining the total counts of immune cells in both subgroups at three months, it was observed that the median cell counts of all cell subpopulations aside from B cells were higher in the non-hospitalised subgroup in comparison to the hospitalised subgroup. However, none of these differences were statistically significant (Table 3).

Flow cytometry analysis, memory T cells

CD4⁺ T cell analysis revealed a higher median number of CD4⁺CD45RO⁺ memory T cells three months post-infection in the hospitalised subgroup, compared to non-hospitalised individuals, and a higher median number of CD4⁺CD45RO⁺IL-2⁺ memory T cells was also observed upon Omicron peptide pool

TABLE 1 Demographics of study subgroup. Values are median with interquartile range (IQR) or number (n), or percentage (%), where indicated.

Prospective Omicron cohort demographics	Non-hospitalised (n = 10)	Hospitalised (n = 10)	Total (n = 20)	P-value
Female, n (%)	5 (50)	5 (50)	10 (50)	1
Age at diagnosis: Median (IQR)	46 (4.5)	53 (17.5)	46.5 (12.5)	*<0.001
Over 50 years old, n (%)	1 (10)	6 (60)	7 (35)	0.019
Vaccine status, n (%)	9 (90)	9 (90)	18 (90)	1
Comorbidity				
1. Autoimmune, n (%)	1 (10)	2 (20)	3 (15)	0.531
2. Metabolic, n (%)	3 (30)	3 (30)	6 (30)	1
3. Respiratory, n (%)	2 (20)	3 (30)	5 (25)	0.606
4. Cardiovascular, n (%)	1 (10)	4 (40)	5 (25)	0.121
5. Cancer, n (%)	0 (0)	2 (20)	2 (10)	0.136
6. Gastrointestinal, n (%)	1 (10)	1 (10)	2 (10)	1
7. Musculoskeletal, n (%)	1(10)	2 (20)	3 (15)	0.531

P-values were calculated using chi-square (twin paired) test, with the exception of continuous variables* that used a paired Welch's t-test.

stimulation *in vitro* in the hospitalised subgroup (Figures 3A, B). The median subgroup values of IL-2 geometric mean fluorescence intensity (GeoMean) in these CD4⁺ memory T cells trended towards a statistically significant increase in per cell IL-2 expression in the hospitalised subgroup (p <0.06; Figure 3C).

The hospitalised patient subgroup had a higher median number of CD4⁺CD45RO⁺CD158b⁺IL-2⁺ memory T cells, compared to

non-hospitalised individuals (Figure 3D). However, the hospitalised subgroup median GeoMean value for CD158b expression per cell was marginally lower than that observed in non-hospitalised individuals (Figure 3E).

CD8⁺ T cell analysis indicated that the non-hospitalised subgroup had a higher median number of CD8⁺CD45RO⁺ memory T cells three months post infection, relative to the

TABLE 2 Full blood counts at baseline and three months for non-hospitalised and hospitalised: Values are median number of events counted with IQR for each patient.

Cell type (unit)	Baseline					3 months				
	Non-hospitalised		Hospitalised		P-value	Non-hospitalised		Hospitalised		P-value
	Parameter	IQR	Parameter	IQR		Parameter	IQR	Parameter	IQR	
Haemoglobin (g/L)	141.50	13.75	131.00	14.50	*0.02	142.00	13.00	139.00	6.00	0.33
Haematocrit	0.42	0.02	0.39	0.02	*0.02	0.42	0.03	0.41	0.01	0.30
RBC (x10 ¹²)	4.78	0.55	4.39	0.40	*0.01	4.75	0.68	4.35	0.11	0.13
MCH (pg)	29.90	1.65	30.95	1.40	0.39	30.80	1.50	31.95	1.90	0.19
MCHC (g/L)	33.7	8.25	33.65	9.75	0.39	33.7	6.00	33.6	7.25	0.83
MCV (fl)	88.65	4.70	91.55	3.78	0.39	90.50	5.20	93.70	5.30	0.18
RDW (fl)	12.70	0.88	13.45	1.25	0.06	12.80	0.60	13.65	1.98	0.90
Platelets (x10 ⁹ /L)	278.50	66.50	240.50	136.75	0.65	302.00	49.00	261.00	64.75	0.74
Total WCC (x10 ⁹ /L)	6.47	2.17	6.11	5.14	0.31	7.19	2.04	7.22	2.93	0.36
Neutrophils (x10 ⁹ /L)	3.44	1.79	3.58	2.57	0.70	4.22	1.62	3.77	1.54	0.92
Lymphocytes(x10 ⁹ /L)	1.92	0.40	1.94	1.75	0.33	2.05	0.72	1.86	1.45	0.40
Monocytes (x10 ⁹ /L)	0.57	0.28	0.51	0.22	0.90	0.49	0.22	0.52	0.18	0.68
Basophils (x10 ⁹ /L)	0.03	0.03	0.02	0.02	0.20	0.04	0.02	0.04	0.02	0.23
Eosinophils (x10 ⁹ /L)	0.17	0.19	0.12	0.25	0.93	0.15	0.11	0.17	0.19	0.56

Statistical comparisons between groups were done using two tailed Welch's t-test. Data was considered to be significant *(highlighted in bold) if p <0.05.

TABLE 3 TBNK cell counts at three months for non-hospitalised and hospitalised.

Cell type	3 months				
	Non-hospitalised		Hospitalised		P-value
	Median cell counts	IQR	Median cell counts	IQR	
T cells	330.7	766.5	251	537.6	0.964
Cyto T cells	196	205.5	128.7	399.8	0.438
Helper T cells	195.7	483.5	155.3	317.2	0.516
B cells	114.8	144.5	148.8	335.3	0.4
NK cells	164.2	134.2	135	121.3	0.523
NK T cells	27.2	81.1	16	80.3	0.848

Statistical comparisons between groups were done using two tailed Welch's t-test. Data was considered to be significant if $p < 0.05$. Values are median number of events counted with IQR for each patient.

hospitalised subgroup (Figure 4A). A higher median number of CD8⁺CD45RO⁺IL-2⁺ memory T cells was observed in the non-hospitalised subgroup upon Omicron peptide pool stimulation *in vitro* (Figure 4B). The median subgroup GeoMean values for IL-2 per cell expression in these CD8⁺ memory T cells was significantly increased in the hospitalised subgroup ($p < 0.03$; Figure 4C). The non-hospitalised individuals had a marginally higher median number of CD8₊CD45RO₊CD158b₊ memory T cells, compared to the non-hospitalised subgroup (Figure 4D). The non-hospitalised subgroup median GeoMean value for CD158b expression per cell was marginally higher than that observed in hospitalised individuals (Figure 4E).

TCR repertoire assembly

Following repertoire assembly alignment and chain usage, alignment rates were investigated to assess sample quality. Two hospitalised subgroup individuals (H7, H10) had alignment rates <90% and were subjected to further investigation. BLAST alignment uncovered that in most instances sequences aligned to TCR genes but were discarded by MiXCR due to partial coverage and not the product of off-target events or foreign contamination. These sequences were subsequently disregarded with only productive clones exported for further analysis. Following alignment, chain usage rates ranged between 15-30% for TCR α and 70-85% for TCR β (Supplementary Figure 1). Prior to CoV-TCR repertoire analysis we investigated the total clonal abundance for $\alpha\beta$ TCRs between subgroups. The non-hospitalised subgroup had a 2.41-fold increase in TCR α clones and 2.28-fold increase in TCR β (Supplementary Figure 2).

Characteristics of CoV-TCRs

Comparative analysis of CoV-TCRs indicated a disparity in both the abundance and diversity of clonotypes between subgroups (Figure 5A). Repertoire overlap (Figure 5B), calculated using

Jaccard similarity index, also showed a higher degree of overlap in the non-hospitalised subgroup (bottom right quadrant) compared to the hospitalised (top left quadrant) for both $\alpha\beta$ TCR repertoires indicating the presence of meta-clonotypes with shared V genes and CDR3 segments. CDR3 length distribution (Figure 5C) for TCR α repertoires is similarly distributed in both subgroups, with peaks at 11 aa in length. The non-hospitalised subgroup sequences ranged between 7-17 aa in length with 75.12% of sequences ranging between 10-12 aa, whereas the hospitalised length ranged between 9-15 with 68.81% of sequences between 10-12 aa in length. CDR3 length ranged between 10-15 aa in the non-hospitalised and 11-17 aa in the hospitalised for TCR β repertoires. CDR3 length was normally distributed in the non-hospitalised subgroup with 72.03% of sequences in the 11-13 aa range, while the hospitalised subgroup showed a bias for shorter sequence length with 90.35% of sequence between 11-13 aa in length. Kmer analysis (Figure 5D) of the top 20 6mers for CDR3 sequence found that the 55% of Kmers overlapped in $\alpha\beta$ TCR repertoires. This analysis found a disproportionate amount of Kmers for the hospitalised subgroup were reported by two hospitalised subgroup individuals (H1, H3), whereas the non-hospitalised subgroup had a more even distribution of Kmers reported for $\alpha\beta$ TCR repertoires.

CoV-TCRs epitope targeting and repertoire diversity

An investigation into the epitopes targeted by CoV-TCRs (Figure 6) found a noticeable increase in both number of epitopes targeted and circulating clonotypes for the non-hospitalised subgroup, with a 1.56-fold increase in targeted epitopes and 3.01-fold increase in clonotypes for the TCR α repertoire and a respective 1.44 and 2.46-fold increase in TCR β repertoire. An increase in the number of individuals targeting epitopes was observed in the non-hospitalised subgroup. On average 1.94 individuals generated at least 1 clonotype per epitope across $\alpha\beta$ TCR repertoires compared to an average of 1.39 in the hospitalised subgroup. Overall, the

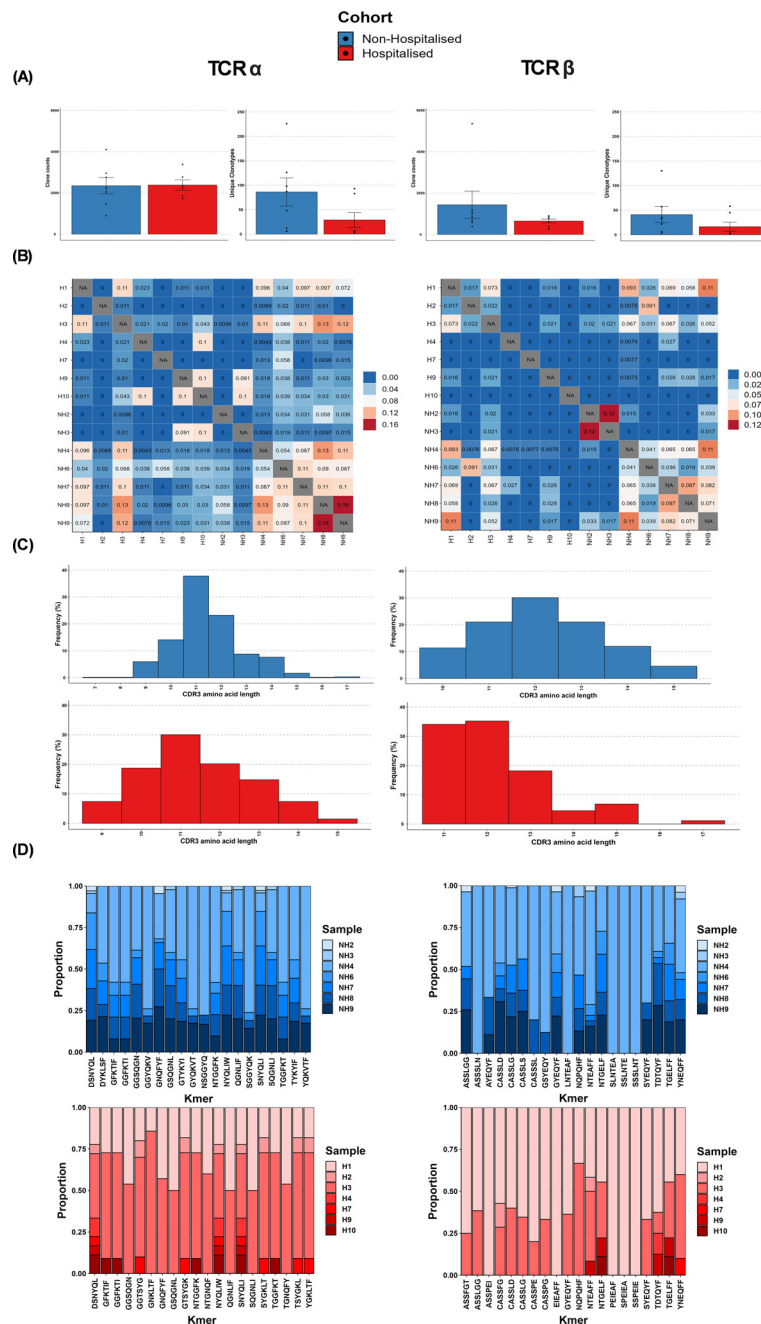


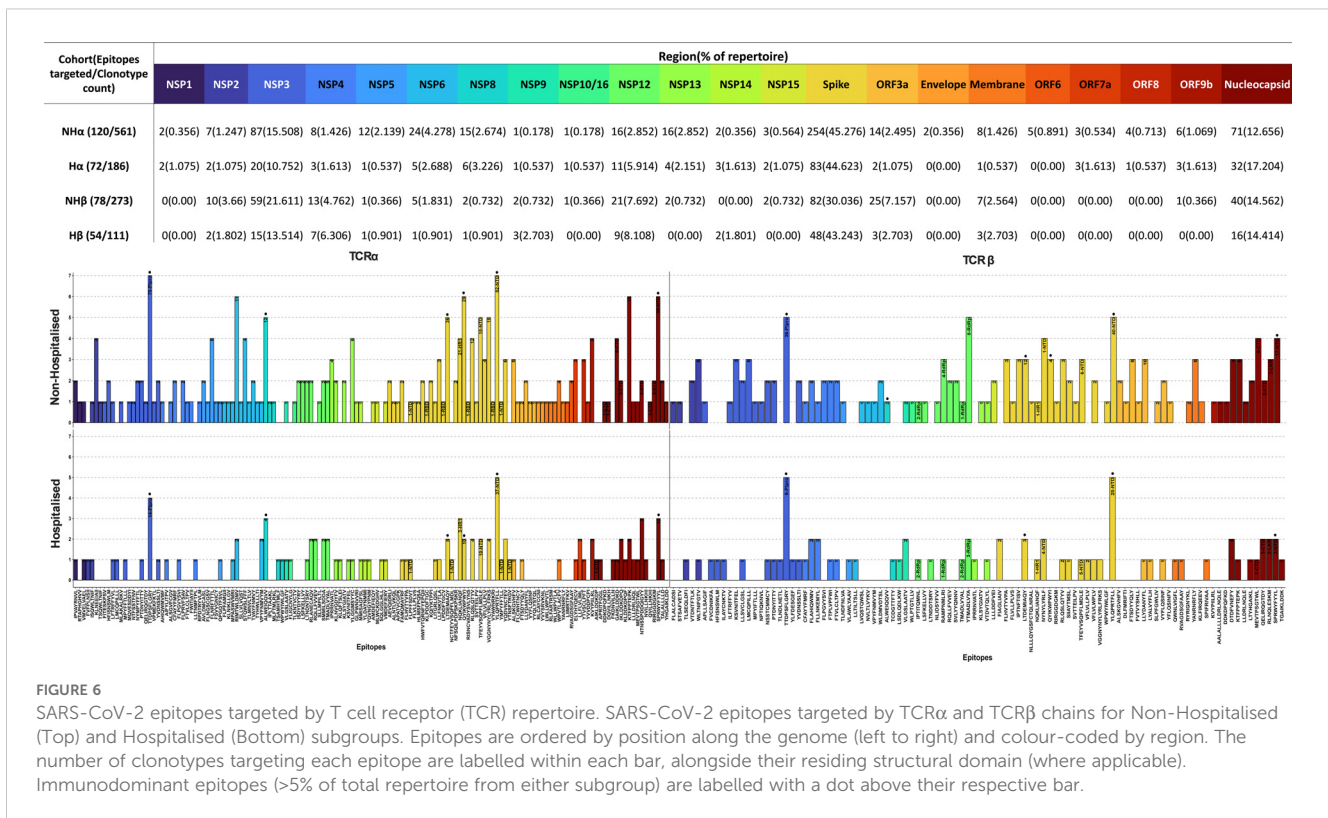
FIGURE 5

T cell receptor (TCR) repertoire overview. (A) Average abundance of SARS-CoV-2 specific clones following immune repertoire analysis. (B) Average diversity of SARS-CoV-2 Clonotypes (defined by unique V/J genes CDR3 sequence) (C) TCR repertoire overlap (Jaccard similarity index) (D) Distribution of CDR3 amino acid length.

distribution of clonotypes across the SARS-CoV-2 genome was relatively similar between both subgroups, with the majority of clonotypes targeting epitopes in the spike, nucleocapsid and NSP3 regions. The largest disparity in clonotype distribution was found in the spike and NSP3 regions of TCRβ repertoires. The non-hospitalised subgroup had a 1.60-fold increase in distribution of clonotypes for NSP3 whereas the hospitalised subgroup had a 1.44-fold increased distribution of clonotypes for the spike region. Interestingly, despite both subgroups having a heavy bias towards

TCRβ chain usage (Supplementary Figures 1, 2) when investigating their complete repertoire, TCRα chain usage was more prevalent in CoV-TCRs.

Several immunodominant epitopes (Supplementary Figure 3), conserved between both subgroups, were responsible for a large portion of repertoire diversity. Although the immunodominant epitopes were conserved between subgroups, a stark contrast in circulating clonotypes and stereotypy was noted in the non-hospitalised subgroup, with individuals whose HLA-genotype



matched the putative HLA alleles for immunodominant epitopes producing a larger variety of clonotypes (Table 4, Supplementary Table 1). In this context stereotypy is defined as unrelated clones sharing (quasi) identical VDJ rearrangements and CDR3 motifs. V/J gene pairing and distribution (Figure 7) further demonstrated that a greater number of individuals producing higher number of clonotypes resulted in convergent recombination of V and J gene segments with a bias towards particular gene pairings for immunodominant epitopes including TRBV27/TRBJ1-1 and TRBV7-8/TRBJ2-7 for epitope 1 (TTDPSFLGRY). Distribution of V/J genes was also markedly different in both subgroups, with the non-hospitalised subgroup having higher proportions of TRAJ1-1, TRAJ2-7, TRAJ49, TRBV7-8 and TRBV27 while the hospitalised subgroup reported higher prevalence of TRAJ9, TRAJ15, TRAJ20, TRAV13-2, TRAV20 and TRAV25. Despite this disparity we found no significant difference in these genes following p-value adjustment of two-sided student’s t-test results.

RNA-Seq analysis

RNA-Seq analysis reported a total of 26 protein coding DEGs predominantly activating pathways involved in wound healing, immune cell development and differentiation (p <0.05; Supplementary Figures 4A, C). Hierarchical clustering of individual samples uncovered that while the non-hospitalised subgroup clustered in close proximity the hospitalised subgroup had groups at either end of the clustering dendrogram with vastly different expression of certain DEGs (Supplementary Figure 4B).

Interestingly samples H1 and H3 which clustered with the non-hospitalised subgroup also shared the highest degree of repertoire overlap and Kmer similarity with our non-hospitalised subgroup following TCR repertoire analysis. HLA typing analysis (Supplementary Table 1) found that each member of the non-hospitalised subgroup expressed at least one of the putative HLA alleles associated with immunodominant epitopes (Table 4) which was only the case for 5/7 participants from the hospitalised subgroup. The overall frequency of expression (Supplementary Figure 4D) for MHC class I alleles, including HLA-A and HLA-B genes on which all immunodominant epitopes are presented was found to be higher in the non-hospitalised subgroup. In contrast the hospitalised subgroup reported higher frequencies of MHC class II alleles including HLA-DR alleles.

Discussion

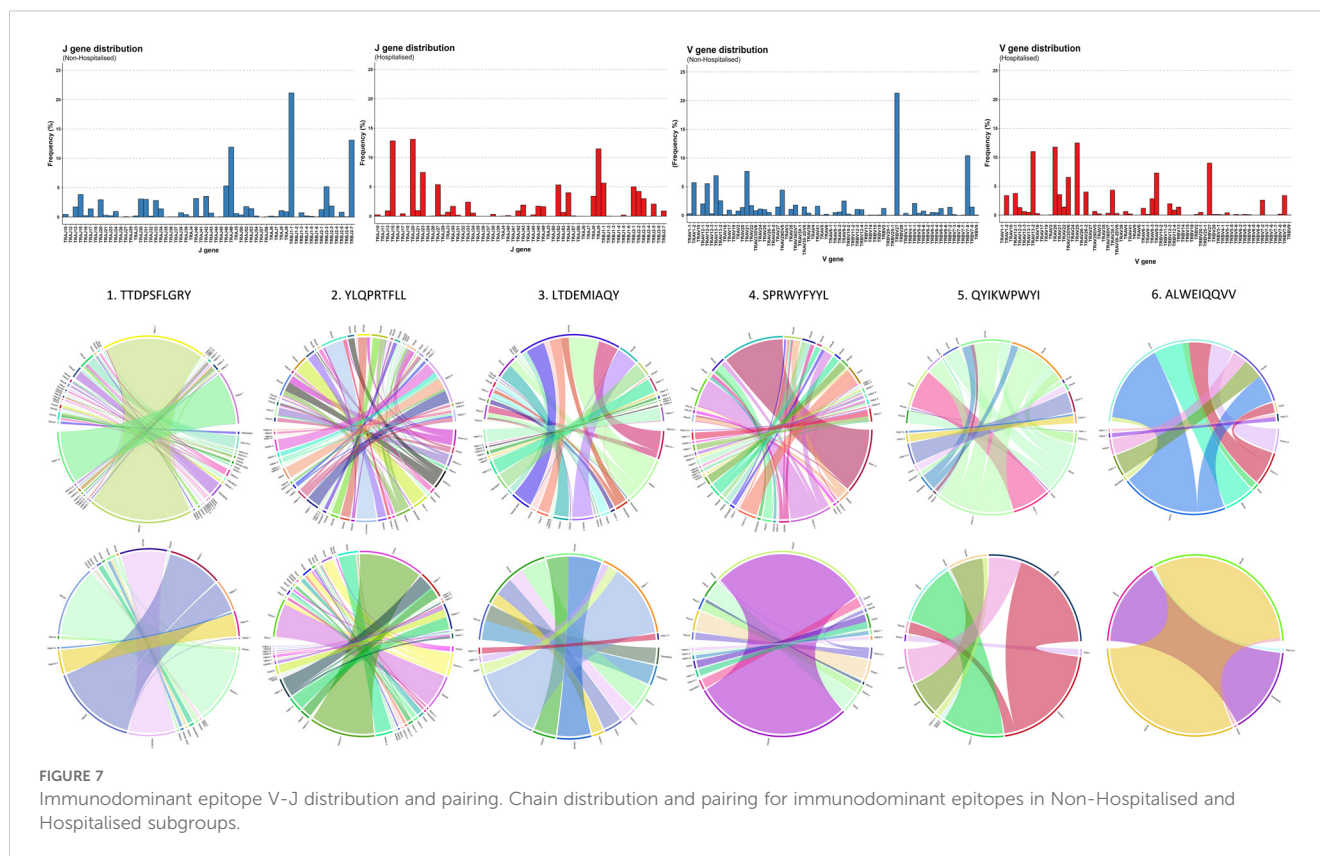
A comprehensive analysis of CD4⁺ and CD8⁺ memory T cell responses to Omicron spike and nucleocapsid antigens three months after SARS-CoV-2 infection, revealed distinct patterns in protective immune responses across hospitalisation subgroups. In hospitalised COVID-19 patients, an increased frequency of IL-2 producing CD4⁺CD45RO⁺ memory T cells was observed, indicating a persistent CD4⁺ T cell-mediated immune response. Furthermore, within these Omicron specific memory T cells, IL-2 per cell expression was stronger in hospitalised patients. In contrast, CD8⁺CD45RO⁺ T cells producing IL-2 were more frequent in non-hospitalised individuals after re-exposure to Omicron peptide

TABLE 4 Immunodominant epitope results.

Epitope	Cohort								Epitope Information		
	Non-hospitalised				Hospitalised				Putative HLA-Alleles	Protein (region)	Conservation (VOI)
	Unique clonotypes	Individuals targeting	% of Repertoire	Dominant pairing	Unique clonotypes	Individuals targeting	% of Repertoire	Dominant pairing			
(1) TTDPSFLGRY	76	7/7	26.11	TRBV27/ TRBJ1-1	20	5/7	11.13	TRAV9- 2/TRAJ23	HLA-A*01:01	NSP3	91.79% (123/134)
(2) YLQPRTFLL	100	7/7	13.78	TRAV13- 2/TRAJ42	57	5/7	20.06	TRAV13- 2/TRAJ9	HLA-A*02:01	Spike(NTD)	100.00% (15/15)
(3) LTDEMQAY	43	5/7	7.04	TRAV21/TRAJ49	12	2/7	1.53	TRBV12-4/ TRBJ1-1	HLA-A*01:01	Spike	100.00% (15/15)
(4) SPRWYFYLY	51	4/7	5.75	TRAV1- 2/TRAJ31	15	3/7	4.41	TRAV20/TRAJ27	HLA-B*07:02	Nucleocapsid (NTD)	100.00% (125/125)
(5) QYIKWPWYI	25	6/7	3.31	TRAV1- 2/TRAJ33	10	2/7	5.01	TRAV38- 1/TRAJ52	HLA-A*24:02	Spike	100.00% (15/15)
(6) ALWEIQQVV	12	5/7	2.22	TRAV21/TRAJ15	4	3/7	9.55	TRAV25/TRAJ20	HLA-A*02:01	NSP8	100.00% (134/134)

Epitopes >5% of the total repertoire from either cohort were considered immunodominant.

Unique clonotypes and dominant pairing were calculated in RStudio using the immunarch package. Putative HLA alleles were retrieved from the VDJdb database. Conservation analysis was performed on sequences retrieved from NCBI database.



pools. However, per cell IL-2 expression was significantly higher in hospitalised patients. Although hospitalised patients had fewer IL-2 producing memory CD8⁺ T cells, those present produced considerably more IL-2 per cell.

This analysis shows key differences in immune memory responses between subgroups, T cells previously infected and subsequently re-exposed to Omicron peptide pools. Hospitalised individuals exhibited a robust CD4⁺ T cell-mediated response, marked by increased IL-2 production, crucial for memory T cell persistence and immune resilience. These nuanced findings underscore the complexity of memory T cell immunity in response to Omicron acute phase severity and hospitalisation. Weiskopf et al. (9), observed a similar trend, reporting significantly elevated IL-2 levels in the supernatant of *in vitro* peptide-stimulated PBMCs isolated from hospitalised patients 14 days post-admission, though they did not specify which cell subsets were responsible for IL-2 production. Cohen et al. (12), employed comparable methods to investigate CD4⁺ and CD8⁺ T cell memory, using SARS-CoV-2 peptide pools to stimulate PBMCs from COVID-19 patients. Their findings, based on intracellular cytokine analysis, including IL-2, revealed that both T cell subsets retained long-term memory to the virus post-infection. In an acute phase comparison of moderate hospitalised and severe hospitalised cases, Sandberg et al. (47), observed lower absolute numbers of both CD4⁺ and CD8⁺ memory T cells in severe cases. SARS-CoV-2 specific memory T cell frequency to a particular variant appears to peak at 3 months in mild (non-hospitalised) cases (48). Contrary to our observations, at the later time point of 4 months post symptom

onset Dan et al. (49), observed that CD4⁺ memory T cell frequencies trended lower in hospitalised cases and CD8⁺ memory T cell frequencies did not differ significantly between hospitalisation subgroups. It is possible that post-acute phase sample timing, sample size and alternate methodologies between studies have contributed to the differences observed. Similar observations were made in the acute phase frequency of CD4⁺ central and effector memory T cells, which were significantly lower in severe (hospitalised) versus mild (non-hospitalised) COVID-19 cases, whereas CD8⁺ compartment frequencies didn't differ significantly between subgroups (50).

The analysis of CD158b expression revealed no significant differences in the numbers of CD158b-positive CD4⁺CD45RO⁺ and CD8⁺CD45RO⁺ memory T cells, as well as in single-cell levels of CD158b expression, between the hospitalisation subgroups. However, there was a modest increase in the frequency of CD4⁺CD45RO⁺CD158b⁺ cells in the hospitalised group and a slight increase in CD8⁺CD45RO⁺CD158b⁺ cells in the non-hospitalised group. These findings aligned with the broader patterns observed in the CD4⁺CD45RO⁺ and CD8⁺CD45RO⁺ cell populations.

TCR repertoire analysis found that 7 out of 10 participants from both subgroups retained memory of known CoV-TCRs retrieved from VDJdb (35). The non-hospitalised subgroup had notably higher numbers of circulating clones, with a broader variety of clonotypes and a higher degree of convergent recombination, as shown by our repertoire overlap and Kmer analysis. These findings mirror that which have been reported in similar investigations that have associated improved clinical outcome and persistence of protective immunity

with increased clonal diversity towards specific epitopes with particular variable and CDR3 segments (51, 52). We also reported a bias towards shorter CDR3 lengths in the hospitalised subgroup in our TCR β analysis. Germline V(D)J recombination of particular TRBV and TRBJ segments can result in shortening of the CDR3 loop and is in part influenced by MHC haplotype impacting specificity of pMHC presentation to thymocytes (53, 54). Although we found no statistical difference in TRBV and TRBJ genes we observed a higher prevalence of TRBV27 in the non-hospitalised subgroup which has been associated with longer CDR3 segments targeting immunodominant epitopes in recent studies (55). We found the hospitalised bias towards shorter CDR3 lengths was largely driven by an epitope-specific response to immunodominant epitope YLQPRTFLL, which accounted for over half of the clonotypes generated for the hospitalised TCR β repertoire. Several COVID-19 studies indicate pMHC presentation of particular epitopes can lead to restricted V gene usage impacting CDR3 motif length (56, 57). Additionally Sureshchandra et al. (58), reported increased CDR3 lengths following both vaccination and natural SARS-CoV-2 infection, with preferentially enriched V genes: TRAV29/DV5, TRBV5-1, TRBV6-5, TRBV11-2, TRBV7-9, TRAV12-2 and TRBV6-2. Our analysis found all but TRBV7-9 were enriched in the non-hospitalised subgroup (Figure 7), further highlighting how CDR3 length and chain usage characteristics can be clinically relevant and serve as a potential biomarker for at risk populations.

In our findings we reported that a small number of CD8⁺ immunodominant epitopes conserved between both subgroups were responsible for the vast majority of generated clonotypes. Individuals in the non-hospitalised subgroup generated higher numbers of unique and convergent clonotypes, with a fundamentally different distributions of epitope specific V/J gene pairings. Three of the six immunodominant epitopes resided on spike glycoprotein, one on the nucleocapsid protein, and two on NSP3 and NSP8. Two of these epitopes TTDPSFLGRY (NSP3) and YLQPRTFLL (Spike-NTD) were targeted by all members of the non-hospitalised subgroup, accounting for 39.89% of the entire CoV-TCR repertoire. The same epitopes were heavily targeted by the hospitalised subgroup, accounting for 31.19% of their CoV-TCR repertoire, however only five members of the hospitalised subgroup generated clonotypes with decreased clonal diversity. HLA-typing analysis found more individuals in the non-hospitalised subgroup expressed putative HLA alleles for immunodominant epitopes and also reported higher expression of MHC class I alleles overall. In contrast, two of the hospitalised subgroup lacked HLA alleles associated with immunodominant epitopes and had higher expression of MHC class II HLA-DR alleles which have been associated with poorer clinical outcomes and long covid in recent studies (59, 60). Recently Du et al. (61), found persistently high HLA-DR⁺CD38⁺ CD8⁺ T cell counts were associated with immune dysregulation, systemic inflammation and impaired killing potential in severe COVID-19. Similarly, Santopaolo et al. (62), observed significantly higher HLA-DR and CD38 expression, alongside elevated levels of numerous pro-inflammatory cytokines (IL-4, IL-7, IL-17, TNF- α) in CD4⁺ and CD8⁺ T cells in severe COVID-19 compared to mild and moderate patients, 3 months post infection. These results align with findings from our flow

cytometry analysis of increased IL-2 expression in both CD4⁺CD45RO⁺ and CD8⁺CD45RO⁺ memory T cells, alongside RNA-seq analysis which indicated increased HLA-DR allele expression of following in the hospitalised subgroup.

It is important to acknowledge the limitations of the study, including the modest cohort sample size, timing of sample collection, the demographics of the study participants, comorbidities and reinfection history (which was not possible to reliably discern without PCR test data). Additional factors for example could contribute to T cell numbers between hospitalisation subgroups such as age, comorbidities and vaccines or treatments received. TCR repertoire analysis is limited by both the number of COVID-19 sequences available from the VDJdb database at the time of the study and strict inclusion criteria using exact CDR3 matches. Widening the scope of the study to include other databases, such as ImmuneCODETM and exploring CDR3 sequences with small changes in aa sequence using tools such as TCRmatch (63) could further elucidate links between TCR repertoire and hospitalisation, and potentially identify new SARS-CoV-2 clonotypes and HLA relationships.

At 3 months post COVID-19, we have observed strong and persistent Omicron specific CD4⁺ memory T cells in hospitalised individuals in particular, and a significant proportion of CoV-TCRs targeting epitopes along the SARS-CoV-2 genome. The accentuated per cell IL-2 expression in both CD4⁺ and CD8⁺ memory T cells of hospitalised patients and altered TCR diversity may indicate the regulation and integrity of T cell based immune protection is compromised over the longer term in individuals previously hospitalised by COVID-19, though this requires confirmation by further longitudinal sampling and analyses. A study of the same cohort revealed that hospitalised patients more frequently experience persistent symptoms in the 3-12 months after Omicron infection (64). The most prevalent persisting symptoms in hospitalised patients were muscle pain and shortness of breath, and they display increased mental health difficulties over time. Furthermore, we observed that hospitalised patients have lower frequency of virus neutralising antibodies and higher proinflammatory protein levels 3 months after infection. This study therefore highlights the importance of considering how thymocyte development and antiviral immune responses are impacted by both HLA-type of patients and the viral lifecycle of SARS-COV-2. The identification of shared immunodominant T cell epitopes in hospitalised and non-hospitalised subgroups that reside outside of the spike protein further highlights the importance of targeting replication proteins in future vaccine strategies.

Data availability statement

The data generated in this study has been deposited in the European Genome-phenome Archive (EGA) and are made findable through the EGA web portal (<https://ega-archive.org>, study number: EGAD50000001356). Access to the data will be granted upon approval by the Data Access Committee (DAC) and in accordance with the conditions outlined in the Data Access Agreement (DAA).

Ethics statement

The studies involving humans were approved by Health and Care Research Wales Ethics committee (20/WA/0179; 14/07/2020). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

GM: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. JM: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. DM: Writing – review & editing. SL: Writing – review & editing. AE: Writing – review & editing. HA: Writing – review & editing. MK: Writing – review & editing. MB: Writing – review & editing. VM: Writing – review & editing. SZ: Writing – review & editing. EM: Writing – review & editing. TR: Writing – review & editing. CW: Writing – review & editing. AB: Writing – review & editing. PS: Funding acquisition, Investigation, Supervision, Writing – original draft, Writing – review & editing. DG: Funding acquisition, Investigation, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by an Opportunity-Led Research Award to DG (PI) and co-authors AB, EM, TS, S-DZ, and PS from Health and Social Care, Research and Development Division (HSC R&D), Public Health Agency (COM/5631/20) and National Core Study: Immunity (NCSi4P program) funding from United Kingdom Research and Innovation (UKRI) to DG and PS. PS and DG wish to acknowledge funding of a PhD studentship to GM from the Department for the Economy (NI). The funding sources had no role

References

1. WHO. COVID-19 epidemiological update – 15 March 2024(2024). Available online at: <https://www.who.int/> (Accessed March 19, 2024).
2. Grant RA, Morales-Nebreda L, Markov NS, Swaminathan S, Querrey M, Guzman ER, et al. Circuits between infected macrophages and T cells in SARS-CoV-2 pneumonia. *Nature*. (2021) 590:635–41. doi: 10.1038/s41586-020-03148-w
3. Diamond MS, Kanneganti T. Innate immunity: the first line of defense against SARS-CoV-2. *Nat Immunol*. (2022) 23:165–76. doi: 10.1038/s41590-021-01091-0
4. Tay MZ, Poh CM, Rénia L, MacAry PA, Ng LFP. The trinity of COVID-19: immunity, inflammation and intervention. *Nat Rev Immunol*. (2020) 20:363–74. doi: 10.1038/s41577-020-0311-8
5. Abbas AK, Lichtman AH, Pillai S. *Cellular and Molecular Immunology*. 10th Edition. Philadelphia, Pennsylvania: Elsevier Health Sciences (2021).
6. Bachmann MF, Oxenius A. Interleukin 2: from immunostimulation to immunoregulation and back again. *EMBO Rep*. (2007) 8:1142–1148; 1148. doi: 10.1038/sj.embor.7401099
7. Qudus MS, Tian M, Sirajuddin S, Liu S, Afaq U, Wali M, et al. The roles of critical pro-inflammatory cytokines in the drive of cytokine storm during SARS-CoV-2 infection. *J Med Virol*. (2023) 95:e28751. doi: 10.1002/jmv.28751
8. Braun J, Loyal L, Frensch M, Wendisch D, Georg P, Kurth F, et al. SARS-CoV-2-reactive T cells in healthy donors and patients with COVID-19. *Nature*. (2020) 587:270–4. doi: 10.1038/s41586-020-2598-9

in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Acknowledgments

The authors would like to extend their sincere gratitude to the participants of this study for their invaluable contribution and to the Western Health & Social Care Trust (WHSC), Northern Ireland, for facilitating access to patients and clinics.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2025.1549570/full#supplementary-material>

9. Weiskopf D, Schmitz KS, Raadsen MP, Grifoni A, Okba NMA, Endeman H, et al. Phenotype and kinetics of SARS-CoV-2-specific T cells in COVID-19 patients with acute respiratory distress syndrome. *Sci Immunol.* (2020) 5:eabd2071. doi: 10.1126/sciimmunol.abd2071
10. Kalia V, Sarkar S. Regulation of effector and memory CD8 T cell differentiation by IL-2—A balancing act. *Front Immunol.* (2018) 9. doi: 10.3389/fimmu.2018.02987
11. Boyman O, Sprent J. The role of interleukin-2 during homeostasis and activation of the immune system. *Nat Rev Immunol.* (2012) 12:180–90. doi: 10.1038/nri3156
12. Cohen KW, Linderman SL, Moodie Z, Czartoski J, Lai L, Mantus G, et al. Longitudinal analysis shows durable and broad immune memory after SARS-CoV-2 infection with persisting antibody responses and memory B and T cells. *medRxiv.* (2021) 2(7). doi: 10.1101/2021.04.19.21255739
13. Sedegah M, Porter C, Goguet E, Ganeshan H, Belmonte M, Huang J, et al. Cellular interferon-gamma and interleukin-2 responses to SARS-CoV-2 structural proteins are broader and higher in those vaccinated after SARS-CoV-2 infection compared to vaccinees without prior SARS-CoV-2 infection. *PLoS One.* (2022) 17:e0276241. doi: 10.1371/journal.pone.0276241
14. Lanier LLY2. NK cell recognition. *Annu Rev Immunol.* (2005) 23:225–74. doi: 10.1146/annurev.immunol.23.021704.115526
15. Uhrberg M. The KIR gene family: life in the fast lane of evolution. *Eur J Immunol.* (2005) 35:10–5. doi: 10.1002/eji.200425743
16. Parham P, Moffett A. Variable NK cell receptors and their MHC class I ligands in immunity, reproduction and human evolution. *Nat Rev Immunol.* (2013) 13:133–44. doi: 10.1038/nri3370
17. National Research Project for SARS, Beijing Group. The involvement of natural killer cells in the pathogenesis of severe acute respiratory syndrome. *Am J Clin Pathol.* (2004) 121:507–11. doi: 10.1309/WPK7Y2XKNF4CBF3R
18. Littera R, Chessa L, Deidda S, Angioni G, Campagna M, Lai S, et al. Natural killer-cell immunoglobulin-like receptors trigger differences in immune response to SARS-CoV-2 infection. *PLoS One.* (2021) 16:e0255608. doi: 10.1371/journal.pone.0255608
19. Wiczorek M, Abualrous ET, Sticht J, Álvaro-Benito M, Stolzenberg S, Noé F, et al. Major histocompatibility complex (MHC) class I and MHC class II proteins: conformational plasticity in antigen presentation. *Front Immunol.* (2017) 8:292. doi: 10.3389/fimmu.2017.00292
20. Deng L, Langley RJ, Wang Q, Topalian SL, Mariuzza RA. Structural insights into the editing of germ-line-encoded interactions between T-cell receptor and MHC class II by V α CDR3. *Proc Natl Acad Sci.* (2012) 109:14960–5. doi: 10.1073/pnas.1207186109
21. Castro CD, Luoma AM, Adams EJ. Coevolution of T-cell receptors with MHC and non-MHC ligands. *Immunol Rev.* (2015) 267:30–55. doi: 10.1111/imr.12327
22. Glusman G, Rowen L, Lee I, Boyesen C, Roach JC, Smit AFA, et al. Comparative genomics of the human and mouse T cell receptor loci. *Immunity.* (2001) 15:337–49. doi: 10.1016/S1074-7613(01)00200-X
23. Zhao Y, Niu C, Cui J. Gamma-delta ($\gamma\delta$) T cells: friend or foe in cancer development? *J Transl Med.* (2018) 16:3–2. doi: 10.1186/s12967-017-1378-2
24. Dupic T, Marcou Q, Walczak AM, Mora T. Genesis of the $\alpha\beta$ T-cell receptor. *PLoS Comput Biol.* (2019) 15:e1006874. doi: 10.1371/journal.pcbi.1006874
25. Zarnitsyna V, Evavold B, Schoettle L, Blattman J, Antia R. Estimating the diversity, completeness, and cross-reactivity of the T cell repertoire. *Front Immunol.* (2013) 4. doi: 10.3389/fimmu.2013.00485
26. English A, McDaid D, Lynch SM, McLaughlin J, Cooper E, Wingfield B, et al. Genomic, proteomic, and phenotypic biomarkers of COVID-19 severity: protocol for a retrospective observational study. *JMIR Res Protoc.* (2024) 13:e50733. doi: 10.2196/50733
27. Andrews S. *FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]*. (2010). Available at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
28. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal; Vol 17 No 1: Next Generation Sequencing Data Anal.* (2011) 17:10–12. doi: 10.14806/ej.17.1.200
29. Bolotin DA, Poslavsky S, Mitrophanov I, Shugay M, Mamedov IZ, Putintseva EV, et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat Methods.* (2015) 12:380–1. doi: 10.1038/nmeth.3364
30. Li H. seqtk, Toolkit for processing sequences in FASTA/Q formats (2023). Available online at: <https://github.com/lh3/seqtk> (Accessed June 5, 2024).
31. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, et al. Database resources of the national center for biotechnology information. *Nucleic Acids Res.* (2022) 50:D20–6. doi: 10.1093/nar/gkab112
32. Nolan S, Vignali M, Klinger M, Dines JN, Kaplan IM, Svejnova E, et al. A large-scale database of T-cell receptor beta (TCR β) sequences and binding associations from natural and synthetic exposure to SARS-CoV-2. *Res Sq.* (2020). doi: 10.21203/rs.3.rs-51964/v1
33. Zhang W, Wang L, Liu K, Wei X, Yang K, Du W, et al. PIRD: pan immune repertoire database. *Bioinformatics.* (2020) 36:897–903. doi: 10.1093/bioinformatics/btz614
34. Tickotsky N, Sagiv T, Prilusky J, Shifrut E, Friedman N. McPAS-TCR: a manually curated catalogue of pathology-associated T cell receptor sequences. *Bioinformatics.* (2017) 33:2924–9. doi: 10.1093/bioinformatics/btx286
35. Goncharov M, Bagaev D, Shcherbinin D, Zvyagin I, Bolotin D, Thomas PG, et al. VDJdb in the pandemic era: a compendium of T cell receptors specific for SARS-CoV-2. *Nat Methods.* (2022) 19:1017–9. doi: 10.1038/s41592-022-01578-0
36. Team R. *RStudio: integrated development for R* Vol. 700. Boston, MA: RStudio, Inc. (2015). p. 879.
37. Team I. immunarch: an R package for painless bioinformatics analysis of T-cell and B-cell immune repertoires. *Zenodo.* (2019) 10:5281. doi: 10.5281/zenodo.3367200
38. Takara Bio. NGS-immune-viewer(2024). Available online at: <https://www.takarabio.com/products/next-generation-sequencing/bioinformatics-tools/cogent-ngs-immune-viewer> (Accessed June 18, 2024).
39. Parasuraman S. Protein data bank. *J Pharmacol pharmacotherapeutics.* (2012) 3:351–2. doi: 10.4103/0976-500X.103704
40. Marrama D, Chronister W, Sternberg L, Vita R, Kosaloglu Z, Sette A, et al. PEPMatch: a tool to identify short peptide sequence matches in large sets of proteins. *BMC Bioinf.* (2023) 24:485. doi: 10.1186/s12859-023-05606-4
41. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* (2014) 15:550. doi: 10.1186/s13059-014-0550-8
42. Kolde R. *Pheatmap: pretty heatmaps*. R package version, Vol. 1. (2019). p. 726. Available at: <https://github.com/raivokolde/pheatmap>
43. Blighe K, Rana S, Lewis M. *EnhancedVolcano*. EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling (2023). R package version 1.20.0. doi: 10.18129/B9.bioc.EnhancedVolcano Available at: <https://github.com/kevinblighe/EnhancedVolcano>
44. Yu G, Wang L, Han Y, He Q. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omic: J Integr Biol.* (2012) 16:284–7. doi: 10.1089/omi.2011.0118
45. Yu G. *Enrichplot: visualization of functional enrichment result*. R package version, Vol. 1. (2021). doi: <https://yulab-smu.top/biomedical-knowledge-mining-book/>
46. Orenbuch R, Filip I, Comito D, Shaman J, Pe'er I, Rabadan R. arcasHLA: high-resolution HLA typing from RNAseq. *Bioinformatics.* (2020) 36:33–40. doi: 10.1093/bioinformatics/btz474
47. Sandberg JT, Varnaité R, Christ W, Chen P, Muvva JR, Maleki KT, et al. SARS-CoV-2-specific humoral and cellular immunity persists through 9 months irrespective of COVID-19 severity at hospitalisation. *Clin Transl Immunol.* (2021) 10:e1306. doi: 10.1002/cti2.1306
48. Gurevich M, Zilkha-Falb R, Sonis P, Magalashvili D, Menascu S, Flechter S, et al. SARS-CoV-2 memory B and T cell profiles in mild COVID-19 convalescent patients. *Int J Infect Dis.* (2022) 115:208–14. doi: 10.1016/j.ijid.2021.12.309
49. 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:eabf4063. doi: 10.1126/science.abf4063
50. Rajamanickam A, Pavan Kumar N, Pandiaraj AN, Selvaraj N, Munisankar S, Renji RM, et al. Characterization of memory T cell subsets and common γ -chain cytokines in convalescent COVID-19 individuals. *J Leukoc Biol.* (2022) 112:201–12. doi: 10.1002/JLB.5COVA0721-392RR
51. Chang C, Feng P, Wu T, Alachkar H, Lee K, Chang W. Profiling of T cell repertoire in SARS-CoV-2-infected COVID-19 patients between mild disease and pneumonia. *J Clin Immunol.* (2021) 41:1131–45. doi: 10.1007/s10875-021-01045-z
52. Zornikova KV, Khmelevskaya A, Sheetikov SA, Kiryukhin DO, Shcherbakova OV, Titov A, et al. Clonal diversity predicts persistence of SARS-CoV-2 epitope-specific T-cell response. *Commun Biol.* (2022) 5:1351. doi: 10.1038/s42003-022-04250-7
53. Sankar K, Hoi KH, Hötzel I. Dynamics of heavy chain junctional length biases in antibody repertoires. *Commun Biol.* (2020) 3:207–3. doi: 10.1038/s42003-020-0931-3
54. Matsutani T, Ohmori T, Ogata M, Soga H, Kasahara S, Yoshioka T, et al. Comparison of CDR3 length among thymocyte subpopulations: impacts of MHC and BV segment on the CDR3 shortening. *Mol Immunol.* (2007) 44:2378–87. doi: 10.1016/j.molimm.2006.10.026
55. Yang G, Wang J, Sun P, Qin J, Yang X, Chen D, et al. SARS-CoV-2 epitope-specific T cells: Immunity response feature, TCR repertoire characteristics and cross-reactivity. *Front Immunol.* (2023) 14:1146196. doi: 10.3389/fimmu.2023.1146196
56. Choy C, Chen J, Li J, Gallagher DT, Lu J, Wu D, et al. SARS-CoV-2 infection establishes a stable and age-independent CD8+ T cell response against a dominant nucleocapsid epitope using restricted T cell receptors. *Nat Commun.* (2023) 14:6725. doi: 10.1038/s41467-023-42430-z
57. Zhou D, Luo Y, Ma Q, Xu Y, Yao X. The characteristics of TCR CDR3 repertoire in COVID-19 patients and SARS-CoV-2 vaccine recipients. *Virulence.* (2024) 15:2421987. doi: 10.1080/21505594.2024.2421987
58. Sureshchandra S, Lewis SA, Doratt BM, Jankeel A, Coimbra Ibraim I, Messaoudi I. Single-cell profiling of T and B cell repertoires following SARS-CoV-2 mRNA vaccine. *JCI Insight.* (2021) 6(24). doi: 10.1172/jci.insight.153201

59. Silva MJA, Ribeiro LR, Lima KVB, Lima LNGC. Adaptive immunity to SARS-CoV-2 infection: A systematic review. *Front Immunol.* (2022) 13:1001198. doi: 10.3389/fimmu.2022.1001198
60. Klein J, Wood J, Jaycox JR, Dhodapkar RM, Lu P, Gehlhausen JR, et al. Distinguishing features of long COVID identified through immune profiling. *Nature.* (2023) 623:139–48. doi: 10.1038/s41586-023-06651-y
61. Du J, Wei L, Li G, Hua M, Sun Y, Wang D, et al. Persistent high percentage of HLA-DR(+)CD38(high) CD8(+) T cells associated with immune disorder and disease severity of COVID-19. *Front Immunol.* (2021) 12:735125. doi: 10.3389/fimmu.2021.735125
62. Santopaolo M, Gregorova M, Hamilton F, Arnold D, Long A, Lacey A, et al. Prolonged T-cell activation and long COVID symptoms independently associate with severe COVID-19 at 3 months. *Elife.* (2023) 12. doi: 10.7554/eLife.85009
63. Chronister WD, Crinklaw A, Mahajan S, Vita R, Koşaloğlu-Yalçın Z, Yan Z, et al. TCRMatch: predicting T-cell receptor specificity based on sequence similarity to previously characterized receptors. *Front Immunol.* (2021) 12:640725. doi: 10.3389/fimmu.2021.640725
64. McLaughlin J, McDaid D, Lynch SM, English A, McLaughlin J, Kelly M, et al. Longitudinal sequelae of severe COVID-19 3-12 months post omicron infection. *Preprints Lancet.* (2024). doi: 10.2139/ssrn.4881312