nature immunology

Article

https://doi.org/10.1038/s41590-024-02064-9

Temporal profiling of human lymphoid tissues reveals coordinated defense against viral challenge

Received: 14 February 2023

Accepted: 10 December 2024

Published online: 31 January 2025

Check for updates

Matthew L. Coates[1,2,5], Nathan Richoz[1,2,5], Zewen K. Tuong[1,2,3,5], Georgina S. Bowyer[1,2,5], Colin Y. C. Lee[1,2,3], John R. Ferdinand[1], Eleanor Gillman[1,2], Mark McClure[1,2], Lisa Dratva[3,4], Sarah A. Teichmann[3,4], David R. Jayne[2], Rafael Di Marco Barros[1], Benjamin J. Stewart[1,2] & Menna R. Clatworthy[1,2,3]

Adaptive immunity is generated in lymphoid organs, but how these structures defend themselves during infection in humans is unknown. The nasal epithelium is a major site of viral entry, with adenoid nasal-associated lymphoid tissue (NALT) generating early adaptive responses. In the present study, using a nasopharyngeal biopsy technique, we investigated longitudinal immune responses in NALT after a viral challenge, using severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection as a natural experimental model. In acute infection, infiltrating monocytes formed a subepithelial and perifollicular shield, recruiting neutrophil extracellular trap-forming neutrophils, whereas tissue macrophages expressed pro-repair molecules during convalescence to promote the restoration of tissue integrity. Germinal center B cells expressed antiviral transcripts that inversely correlated with fate-defining transcription factors. Among T cells, tissue-resident memory CD8 T cells alone showed clonal expansion and maintained cytotoxic transcriptional programs into convalescence. Together, our study provides unique insights into how human nasal adaptive immune responses are generated and sustained in the face of viral challenge.

Secondary lymphoid organs orchestrate the spatial and temporal arrangement of different immune cell subsets to enable appropriate cellular interactions and the generation of timely pathogen-specific responses. T and B lymphocytes are largely segregated into distinct areas within lymphoid tissues, except within germinal centers (GCs), highly organized structures with light and dark zones, where GC B cells acquire antigen from follicular dendritic cells (FDCs) and present it to PD^+CD4^+ T follicular helper (T_{FH}) cells[1]. This interaction leads to iterative rounds of GC B cell proliferation, during which there is somatic hypermutation and the emergence of both memory B cells and antibody-secreting plasmablasts or cells that defend against current and future infection[1,2–7]. As lymphoid tissues are critical for the generation of adaptive immunity, they need robust defense mechanisms. These processes are well studied in animal models, for example, in mice, subcapsular sinus macrophages contribute to lymph node defense, preventing the spread of lymph-borne viruses and bacteria, the latter via interactions with innate lymphocytes[8,9]. In contrast, profiling of human lymphoid tissue responses during infection is limited, resulting partly from the challenges of sampling these structures in humans.

[1]Department of Medicine, Molecular Immunity Unit, University of Cambridge, Cambridge, UK. [2]Cambridge Institute for Therapeutic Immunology and Infectious Diseases, Cambridge, UK. [3]Cellular Genetics, Wellcome Sanger Institute, Hinxton, UK. [4]Cambridge Stem Cell Institute, Cambridge, UK. [5]These authors contributed equally: Matthew L. Coates, Nathan Richoz, Zewen K. Tuong, Georgina S. Bowyer. e-mail: mrc38@medschl.cam.ac.uk

The nasopharynx contains the pharyngeal tonsils (also known as adenoids), which form part of Waldeyer's ring of lymphoid tissue and are the major site of nasal-associated lymphoid tissue (NALT). Structurally, they have a surface crypt epithelium with antigen-transporting M cells and underlying B cell follicles and T cell-containing interfollicular areas, which together generate protective antibody-secreting and memory cells after a nasal challenge[10,11]. Adenoid tissue typically macroscopically involutes in adolescence and the extent to which NALTs contribute to local mucosal adaptive immune responses in adult humans is unclear. $CD68^+$ macrophages have been identified in human adenoid NALT[12], but macrophage heterogeneity and contributions to lymphoid tissue defense or repair post-infection in humans are completely unknown. The posterior nasal space can be readily visualized in humans using an endoscope, raising the possibility that endoscopic adenoid NALT biopsies might offer a feasible method to profile adaptive immune responses in secondary lymphoid organs after an upper respiratory tract immune challenge. To test this, we undertook a 'natural experiment', sampling adult participants after severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection—a virus that uses nasal epithelium as a point of entry and replication[13,14]—to understand how NALTs defend themselves in the face of local infection to ensure that the cells and infrastructure required to support an ongoing adaptive immune response are maintained.

Although SARS-CoV-2 has already been well studied in humans, much of this work has used peripheral blood or epithelial sampling in live participants, which fails to capture lymphoid follicular components, such as GC B cells[15–17]. Lymphoid tissue sampling in SARS-COV-2 infection has been conducted in children undergoing adenoidectomy, suggesting a robust GC response in pediatric NALT whereas, in contrast, autopsy samples from the spleen and lymph nodes in adults with fatal SARS-CoV-2 infection show minimal or absent GCs[18–20].

In the present study, we applied single-cell RNA sequencing (scRNA-seq), multi-parameter flow cytometry and confocal imaging to adult NALT biopsies, to understand the cellular molecular adaptations that support, polarize and defend human secondary lymphoid tissue in adults in the face of a viral immune challenge.

Results

Postnasal space biopsy enables lymphoid tissue profiling

We developed a well-tolerated endoscopic technique for collecting NALT biopsies from live humans using topical local anesthesia (Extended Data Fig. 1a,b). 23 participants underwent sampling (aged 19–91 years), including 10 healthy controls, 8 patients with acute COVID-19 (sampled within 1 week of a first positive SARS-CoV-2 reverse transcriptase–polymerase chain reaction (RT–PCR) test) and 5 convalescent patients with COVID-19 (clinically asymptomatic participants sampled 3–5 weeks after the first positive SARS-CoV-2 RT–PCR; Fig. 1a and Supplementary Table 1). Postnasal space biopsies (with additional nasal brushing or curettage in some cases) were collected, alongside paired blood samples and processed for flow cytometry, scRNA-seq and confocal microscopy (Supplementary Table 1).

After quality control, we generated data on 162,738 cells and identified 20 major cell clusters across peripheral blood mononuclear cells (PBMCs) and nasal tissues, including epithelial cells, stromal cells and a variety of immune cell populations (Fig. 1b and Extended Data Fig. 1c–e). There was a significant overrepresentation of GC B cells and T_{FH} cells in nasal samples (Fig. 1c,d), reflecting the presence of GC-containing lymphoid structures. GC B cells were particularly abundant in convalescent NALT samples (Fig. 1d), consistent with an expansion of virus-specific B cells. NALT samples also differed from adult nasal cavity and bronchial brushings[17], in which T_{FH} cells and GC B cells were scarce (Extended Data Fig. 1f–h). Confocal imaging of NALT biopsies confirmed the presence of B cell follicles containing CD21-expressing FDCs and spatially distinct T cell-rich areas, with overlying surface $EPCAM^+$ epithelial cells (Fig. 1e). Notably, the number of differentially expressed genes (DEGs) in COVID-19 samples (both acute or active (A) and convalescent (C)), relative to controls, was greater in NALT than in blood (Fig. 1f). Similarly, gene-set enrichment analysis (GSEA) of these DEGs showed a greater magnitude, significance and number of enriched gene sets in NALT compared with peripheral blood, including interferon (IFN)α response pathway genes (Fig. 1g).

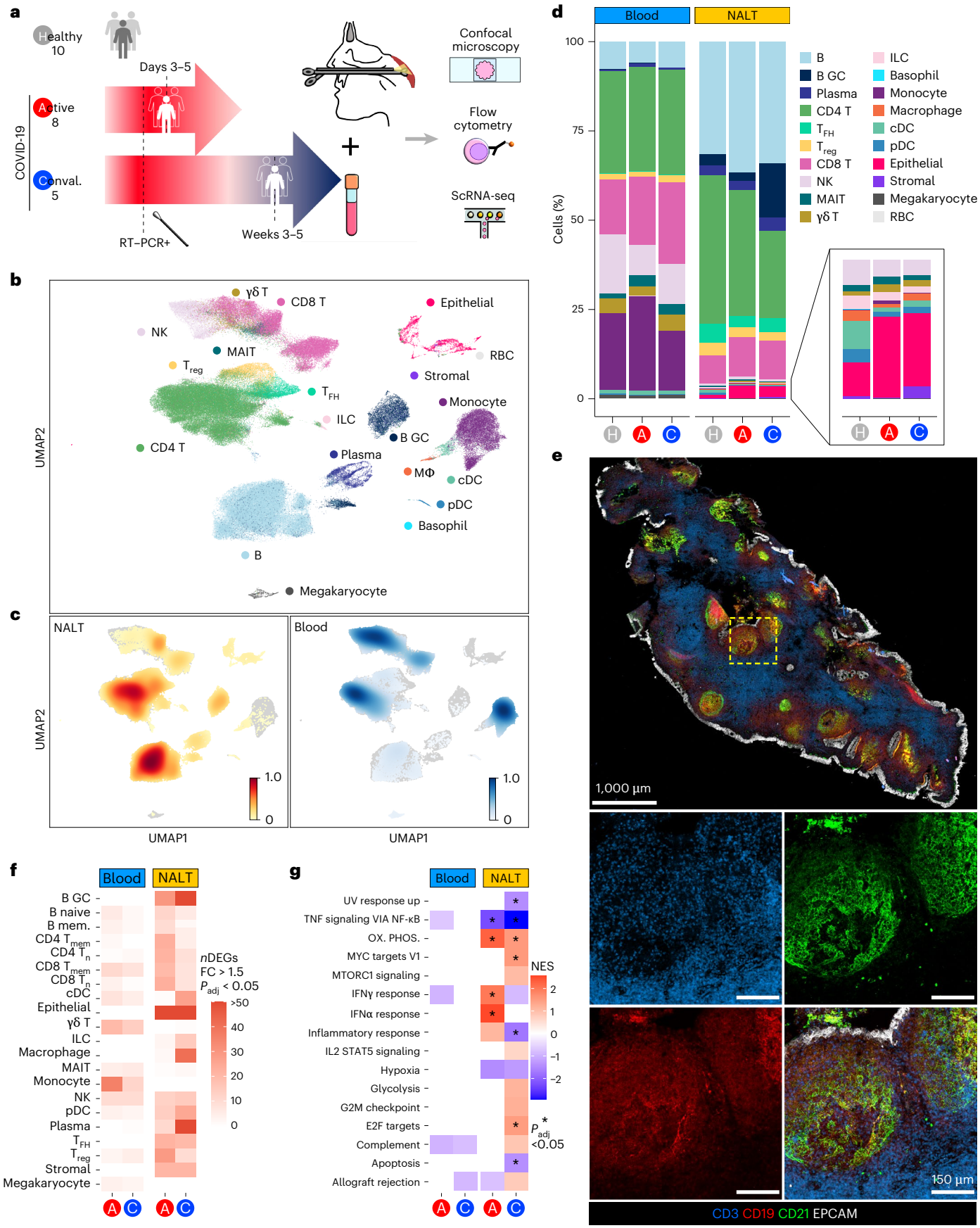
Overall, these data show that nasopharyngeal biopsies can be used to profile functional lymphoid tissue in adults, with capture of follicular cell types not found in blood or nasal cavity brushings. NALT sampling presents a feasible method to interrogate the cellular molecular processes occurring during the generation of adaptive immune responses to a nasal viral infection in humans.

MNPs contribute to lymphoid tissue defense and repair

Postnasal space sampling 3–5 d after diagnostic testing for SARS-CoV-2 provided a unique opportunity to understand how early innate immune responses act to defend NALT from local viral attack. Among mononuclear phagocytes (MNPs), we identified 11 cell types: conventional dendritic cells (cDCs) (*CD141*$^+$ cDC1 and *CD1C*$^+$ cDC2), plasmacytoid dendritic cells (pDCs), six monocyte clusters (classical, intermediate, nonclassical, *SIGLEC1*$^+$, *CD163*^{hi} and *C1Q*$^+$), basophils and macrophages (Fig. 2a and Extended Data Fig. 2a). The proportion of classical monocytes in both nasal tissue and peripheral blood increased in acute COVID-19 compared with healthy controls and convalescent COVID-19 (Fig. 2b,c), consistent with tissue infiltration (confirmed by flow cytometry; Extended Data Fig. 2b). These monocytes showed increased expression of several self- and neutrophil-recruiting chemokines (*CCL2* and *CXCL2*, *CXCL3* and *CXCL8*, respectively) (Fig. 2d). *S100A8/9* transcripts were also increased in classical monocytes in active and convalescent COVID-19 samples (Fig. 2d,e and Extended Data Fig. 2c). *S100A8*/9 dimerizes to form calprotectin, which sequesters metal ions, limiting their availability to bacteria, including upper respiratory tract-dwelling organisms such as staphylococci[21]. SARS-CoV-2 is known to directly infect respiratory tract epithelium and, indeed, we found

Fig. 1 | Experimental overview and cellular landscape of blood and NALT in SARS-CoV-2 infection. a, Schematic of experimental design with number of participants and timing of sampling in each group, sample types (paired NALT and peripheral blood) taken from participants and the techniques used to analyze processed samples. Conval., convalescent. **b**, Uniform Manifold Approximation and Projection (UMAP) of all cells in participants with COVID-19 and healthy controls, with major cell annotations. MΦ, macrophage; γδ T, gamma–delta T cell; RBC, red blood cell. **c**, UMAP of Scanpy embedding density (scanpy.tl.embedding_density), represented as a scaled Gaussian kernel density estimation, of all cells by sample type. NALT embedding density is shown in orange and blood embedding density in blue. **d**, Stacked bar charts showing proportional split (as percentage) of NALT and blood samples by cell type and disease type, with nested bar chart showing proportions, by disease type, of nasal NK cells, MAIT cells, γδ T cells, ILCs, monocytes, macrophages, cDCs, pDCs and epithelial and stromal cells. H, healthy control. **e**, Representative confocal images of a section of a postnasal space biopsy sample in a 19-year-old female from the convalescent COVID-19 group. The yellow dashed box indicates the region magnified in the four boxes in the lower part of the image panel ($n = 1$). **f**, Number of DEGs for disease versus healthy control samples in NALT and blood samples by fine cell-type cluster and disease type, where expression fold-change (FC) is >1.5 and Bonferroni's adjusted *P* value (P_{adj}) is <0.05. Two-sided Wilcoxon's rank-sum test was used for computing DEGs. mem., memory. **g**, Hallmark GSEA of DEGs between disease participants and healthy controls for all (combined) cells by sample type and disease state. Ox. phos., oxidative phosphorylation. NES, normalized enrichment score: *Benjamini–Hochberg $P_{adj} < 0.05$. NESs were calculated in the fgsea package with *P*-value estimation using an adaptive, multilevel, split Monte-Carlo scheme.

https://doi.org/10.1038/s41590-024-02064-9



https://doi.org/10.1038/s41590-024-02064-9

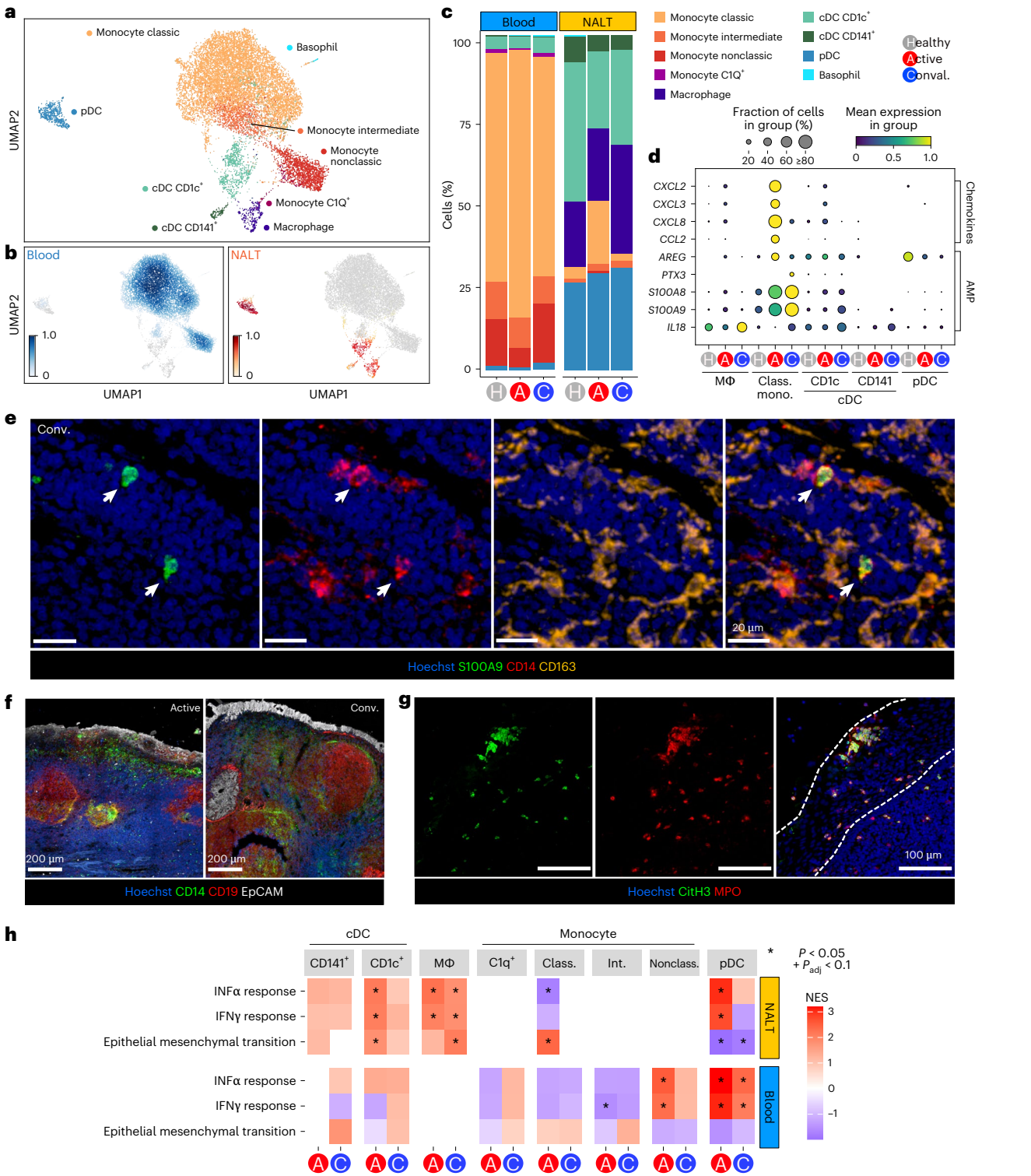


Fig. 2 | Recruited monocytes and inflammatory DCs form a defensive shield in the subepithelium. a, UMAP of MNPs reclustered in isolation, with assigned subset cell-type cluster labels shown. **b**, Scanpy embedding density (Gaussian kernel estimation) UMAP of reclustered MNP subset cell types showing scaled density of cells by sample type (blood or NALT) across the UMAP. NALT embedding density is shown in orange and blood embedding density in blue. **c**, Stacked bar chart showing proportional representation of MNP cell types in blood and nasal samples by disease status. **d**, Selected chemokine and antimicrobial peptide-scaled expression in NALT MNP subsets. Class. mono., classical monocyte. The size of the dot indicates the fraction of cells in the group expressing the gene and the color the scaled mean gene expression. **e**, Representative confocal imaging of section of NALT from a convalescent patient with COVID-19. The white arrowheads highlight the co-expression of S100A9 and CD14 ($n = 1$). **f**, Representative confocal imaging of sections of NALT from an active (left) and convalescent (right) patient ($n = 2$ per group). **g**, Representative confocal imaging of section of NALT from a patient with active COVID-19. The white dotted lines show the edges of the epithelium ($n = 1$). **h**, Heatmap of GSEA Hallmark gene-set enrichment for DEGs in MNP subsets by disease group against healthy control samples. Class., classical; Int., intermediate; Nonclass., nonclassical. The color of the tile shows the NES. Red shows increased gene-set enrichment and blue decreased gene-set enrichment relative to healthy controls, and white or no color shows no enrichment of the gene set. *Statistical significance $P < 0.05$ and Benjamini–Hochberg's $P_{adj} < 0.1$. NESs were calculated in the fgsea package with P-value estimation using an adaptive, multi-level, split Monte-Carlo scheme[αγ].

SARS-CoV-2 spike protein in epithelial cells overlying NALT in acute COVID-19 (Extended Data Fig. 2d) and in some subepithelial T cells (Extended Data Fig. 2e). CD14$^+$ monocytes predominantly localized to the subepithelium in acute COVID-19 (Fig. 2f and Extended Data Fig. 3a–c) and surrounded some B cell follicles in convalescent disease (Fig. 2f), forming a shield around them. In addition, neutrophils were also present in the epithelial and subepithelial areas, some with histone citrullination, required for formation of neutrophil extracellular trap (NET) structures that bind and kill invading bacteria[22] (Fig. 2g). Therefore, calprotectin-producing monocytes and NET-losing neutrophils may work together to defend the underlying lymphoid tissues after epithelial cell damage by invading virus.

Macrophages were found in NALT but not in blood and were increased in convalescent COVID-19 samples. The number of DEGs in NALT-resident macrophages was also greater in convalescent COVID-19 compared with acute infection (Fig. 1f), including interleukin (IL)-18 (Fig. 2d), reminiscent of studies in murine lymph nodes showing macrophage production of IL-18 in the context of bacterial challenge[9]. It is interesting that a population of circulating *C1Q*$^+$ monocytes was also expanded in convalescent COVID-19 (Fig. 2c and Extended Data Fig. 2a), with trajectory analysis suggesting that these represent tissue macrophage precursors (Extended Data Fig. 3d).

GSEA[23] showed a greater magnitude of enrichment in nasal cell types compared with their circulating counterparts (Fig. 2h and Extended Data Fig. 3e). Considering the majority myeloid populations in NALT, 'IFNα response' and 'INFγ response' genes were significantly enriched in macrophages, *CD1C*$^+$ dendritic cells (DCs) and pDCs in acute COVID-19 compared with controls (Fig. 2h and Extended Data Fig. 3e). In convalescent COVID-19, this IFN response gene enrichment became less prominent in cDCs and pDCs but was sustained in macrophages (Fig. 2h). Furthermore, 'epithelial-to-mesenchymal transition' (EMT) gene-set enrichment increased in macrophages in convalescent disease and, indeed, was exclusively enriched in macrophages at this time point (Fig. 2h and Extended Data Fig. 3e), consistent with a role in the restoration of tissue homeostasis.

To further explore myeloid cell heterogeneity in NALT, we reclustered nasal MNPs in isolation from those in blood (Fig. 3a). This enabled the further identification of so-called 'inflammatory' or 'monocyte-derived' DCs (moDCs, based on the expression of *CD14*, *S100A8/9* and *FCGR1A*[24,25] and two distinct clusters of macrophages based on *C1Q* and *FOLR2* expression; Fig. 3a and Extended Data Fig. 4a). The *FOLR2*$^+$ macrophages expressed additional markers associated with tissue-resident macrophages, including *TIMD4* (ref. 26) (Extended Data Fig. 4a). Monocytes and moDCs were increased in acute COVID-19 (Fig. 3b). CD14$^+$ moDCs localized to the subepithelial region (Fig. 3c) and expressed ficolin (*FCN1*) (Fig. 3d), a membrane c-type lectin that mediates CXCL8 upregulation on pathogen encounter[27]. A small proportion of moDCs also expressed *CXCL8* (Fig. 3d) and showed reduced enrichment of antigen presentation genes in acute COVID-19, in contrast to NALT cDCs (Fig. 3e). Altogether, this is consistent with a functional switch in moDCs in acute COVID-19 away from antigen presentation and toward neutrophil recruitment.

C1Q^{hi} macrophage representation in NALT remained stable in health, acute and convalescent COVID-19, but *FOLR2*$^+$ macrophages significantly increased in convalescent COVID-19 (Fig. 3b). Spatially, macrophages were predominantly located in the T cell zone, including the FOLRB$^+$ subset (Fig. 3f and Extended Data Fig. 4b). These *FOLR2*$^+$ macrophages also showed temporal changes in their transcriptional program between acute and convalescent COVID-19, demonstrating enrichment of IL-10- and IL-4-stimulated macrophage gene signatures[28] in convalescent disease (Fig. 3g and Extended Data Fig. 4c). EMT pathway genes highly expressed in *FOLR2*$^+$ macrophages included *COL6A* (Fig. 3h), an important structural component in many tissues[29]. These *FOLR2*$^+$ macrophages also expressed the tolerogenic receptor *LILRB4*, matrix metalloproteinases and *CCL18* (Fig. 3h), a profibrotic cytokine associated with the differentiation of 'M2'-like macrophages[30]. Flow cytometric analysis confirmed increased expression of CD206, a marker of M2 macrophages, in nasal CD14$^+$ cells in convalescent COVID-19 (Extended Data Fig. 4d).

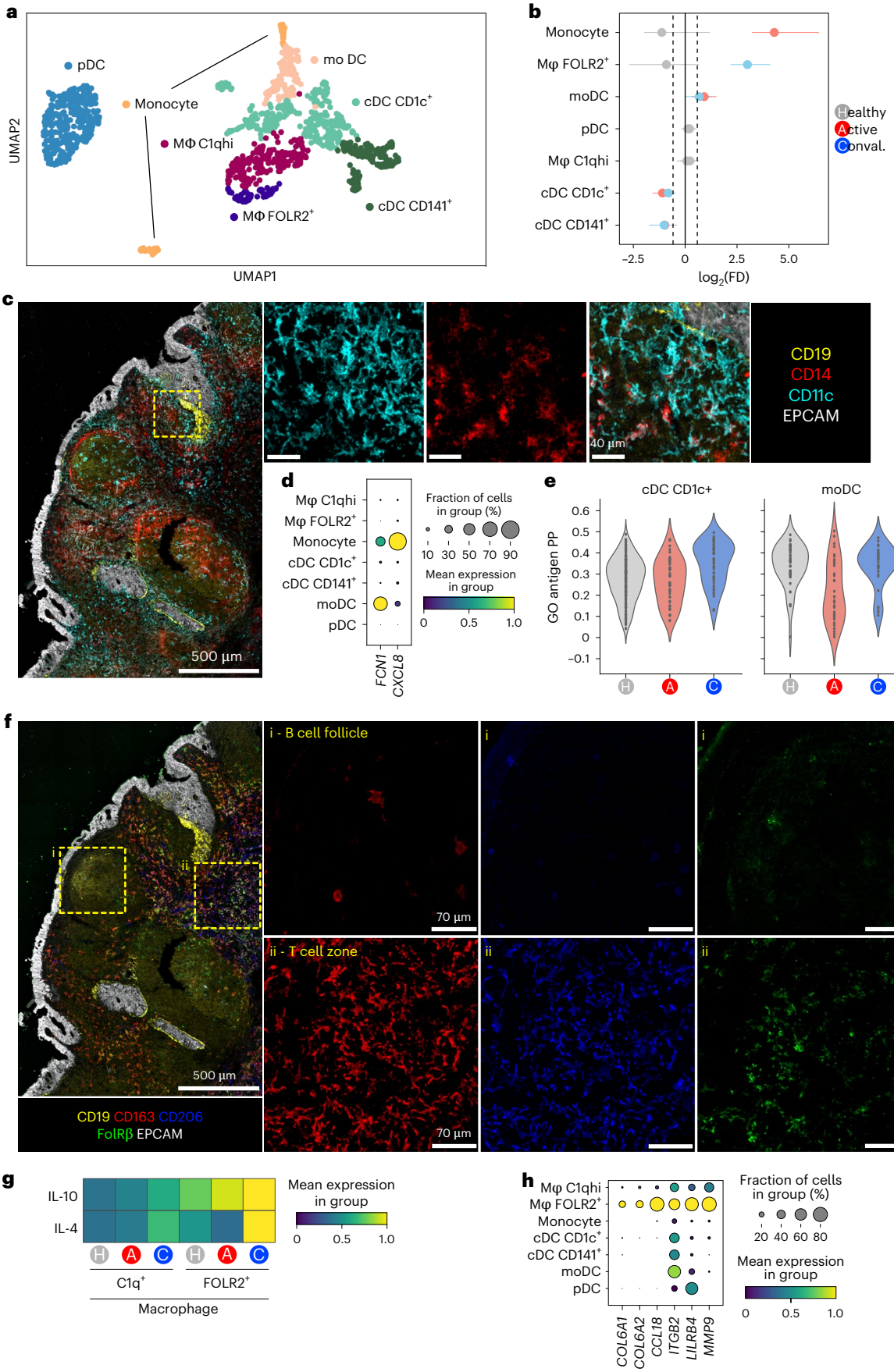
Altogether, our analysis suggests that NALT myeloid cells play differing roles during the course of SARS-CoV-2 infection; infiltrating monocytes and moDCs produce self- and neutrophil-recruiting chemokines to set up a proinflammatory positive feedback loop in acute COVID-19 and localize to the periphery of NALT, forming a defensive shield. Meanwhile, NALT macrophages reside predominantly within the T cell zone and include a *FOLR2*$^+$ subset that is expanded and transcriptionally activated in convalescent COVID-19, adopting a pro-repair, anti-inflammatory transcriptome, promoting tissue integrity and ensuring that the infrastructure for the generation of local adaptive immune responses is maintained.

Clonally expanded NALT CD8 T_{RM} cells show prolonged activation

We next considered T cell, natural killer (NK) cell and innate-like lymphocytes in isolation and identified 19 cell types (Fig. 4a and Extended Data Fig. 5a). We found distinct clusters of naive and memory CD4 and CD8 T cells, regulatory T cells (T_{reg} cells), T_{FH} cells, innate lymphoid cells (ILCs), γδ T cells, mucosal-associated invariant T cells (MAIT cells) and three subsets of NK cells marked by expression of CD16, CD56 and/or KLRC3 (Fig. 4a and Extended Data Fig. 5a). Overall, CD4 T cells were enriched in nasal samples, comprising 80% of CD3$^+$ cells in NALT (Fig. 4b, Extended Data Fig. 5b, confirmed by flow cytometry, and Extended Data Fig. 5c). Tissue-resident memory (T_{RM}) CD8 T cells, T_{FH} cells and ILCs were almost exclusively found in NALT samples, whereas CD16$^+$ NK cells, CD8 T effector cytotoxic (CTL) and CD8 T effector memory cells were predominantly found in blood (Fig. 4b,c and Extended Data Fig. 5d,e). Spatially, programmed cell death protein 1 (PD-1)-expressing T_{FH} cells were located within B cell follicles (Fig. 4d), whereas CD8 T cells dominated within the epithelium (Fig. 4e), including CD8$^+$CD103$^+$ T_{RM} cells, which were also scattered in the B cell follicles and interfollicular regions (Fig. 4f,g and Extended Data Fig. 5f).

Fig. 3 | Tissue-resident macrophages take on pro-repair phenotype in convalescence. a, UMAP of NALT MNPs in isolation for participants with COVID-19 and healthy controls reclustered in isolation, with assigned subset cell-type cluster labels shown. **b**, Proportional difference in abundance of NALT MNP cell types in active COVID-19 (red) and convalescent COVID-19 (blue), relative to healthy controls. Nonsignificant values ($P > 0.05$) relative to healthy controls, following permutation testing, are shown in gray. The bars indicate the bootstrapped 95% CI. FD, fold distribution. **c**, Representative confocal imaging of a section of NALT from a convalescent COVID-19 participant. The yellow dashed box indicates a region magnified in the three boxes on the right-hand side ($n = 1$). **d**, Dotplot showing expression of *FCN1* and *CXCL8* in NALT MNP subsets. The size of the point indicates the fraction of cells in each group expressing the corresponding gene and the color the scaled mean expression of the corresponding gene in each named cell-type group. **e**, Violin plot showing expression of antigen processing and presentation gene ontology (GO) term genes (GO antigen PP) in conventional CD1c$^+$ cDC compared with moDC cell subsets. Each point represents a cell. **f**, Representative confocal imaging of a section of NALT from a convalescent participant with COVID-19. The yellow dashed boxes indicate the regions magnified in the boxes on the right-hand side ($n = 1$). **g**, Heatmap indicating expression of genes associated with IL-4 or IL-10 stimulation in macrophage subsets, calculated as scaled gene-set expression scores[60] using reference IL-10- or IL-4-stimulated macrophage gene sets[28]. **h**, Dotplot showing NALT MNP expression of *COL6A1*, *COL6A2*, *CCL18*, *ITGB2*, *LILRB4* and *MMP9*. The size of the point indicates the fraction of cells in each group expressing the corresponding gene and the color the scaled mean expression of the corresponding gene in each named cell-type group.

https://doi.org/10.1038/s41590-024-02064-9



https://doi.org/10.1038/s41590-024-02064-9

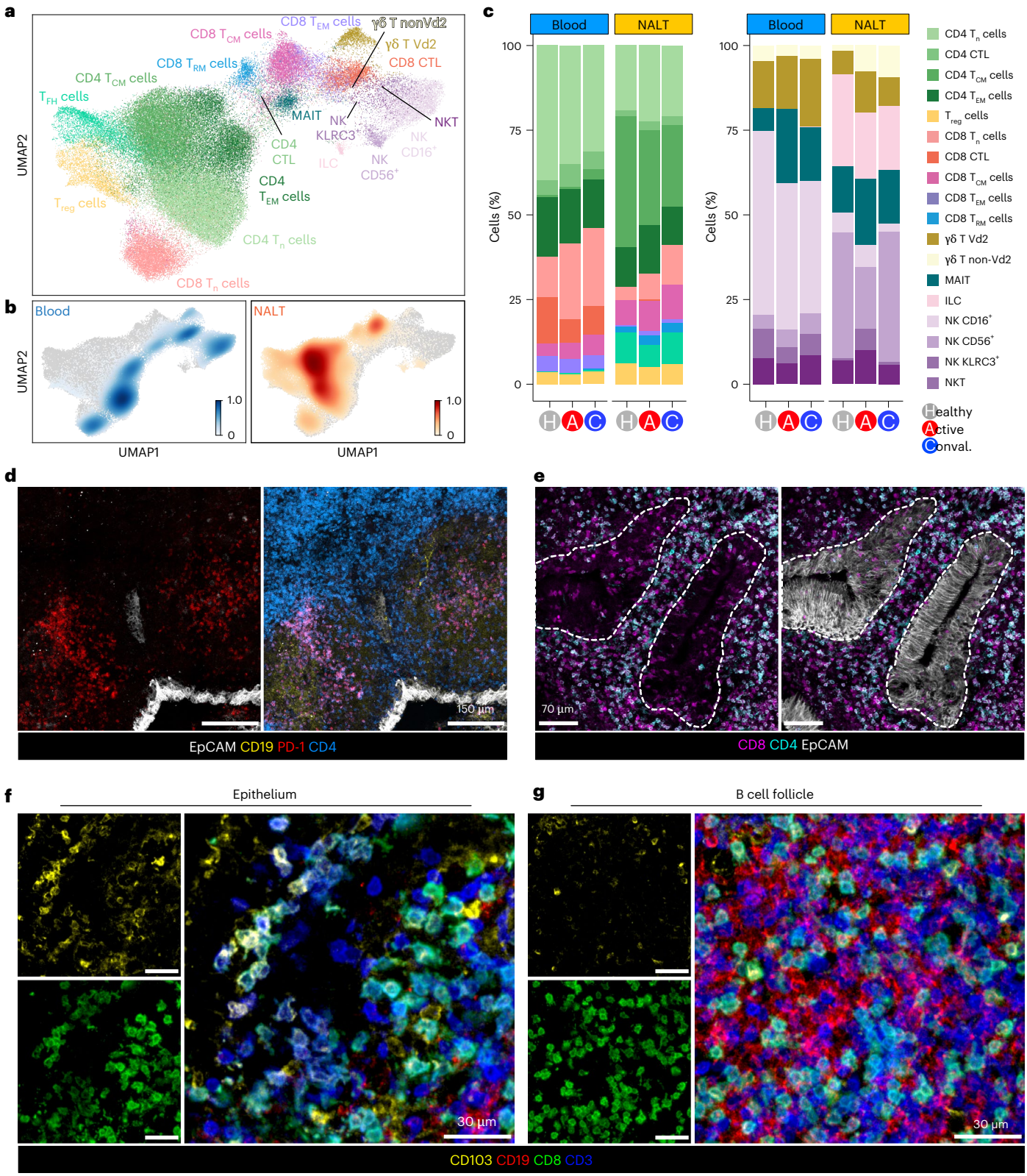


Fig. 4 | NALT CD8 T_{RM} cells exhibit clonal expansion and prolonged activation. **a**, UMAP of T cells and innate lymphocytes from participants with COVID-19 and healthy controls, reclustered in isolation, with assigned subset cell-type cluster labels shown. **b**, Scanpy embedding density (Gaussian kernel estimation) UMAP of T cells and innate lymphocytes showing scaled density of cells by sample type (blood or nasal) across the UMAP where 1 is a high density of cells of the indicated sample type. **c**, Stacked bar charts showing proportional representation of CD4 and CD8 T cells (left, paired bar chart) and innate lymphocytes (right, paired bar chart) divided by sample type (blood or NALT) and disease status. **d**, Representative confocal imaging of subepithelial B cell follicles on a section of NALT from a convalescent COVID-19 participant ($n = 1$). **e**, Representative confocal imaging of section of NALT from a convalescent COVID-19 participant ($n = 1$). **f**,**g**, Representative confocal imaging of the epithelium (**f**) and a B cell follicle (**g**) on a section of NALT from a convalescent COVID-19 participant. The arrows indicate co-expression of CD8 and CD103 ($n = 1$).

https://doi.org/10.1038/s41590-024-02064-9

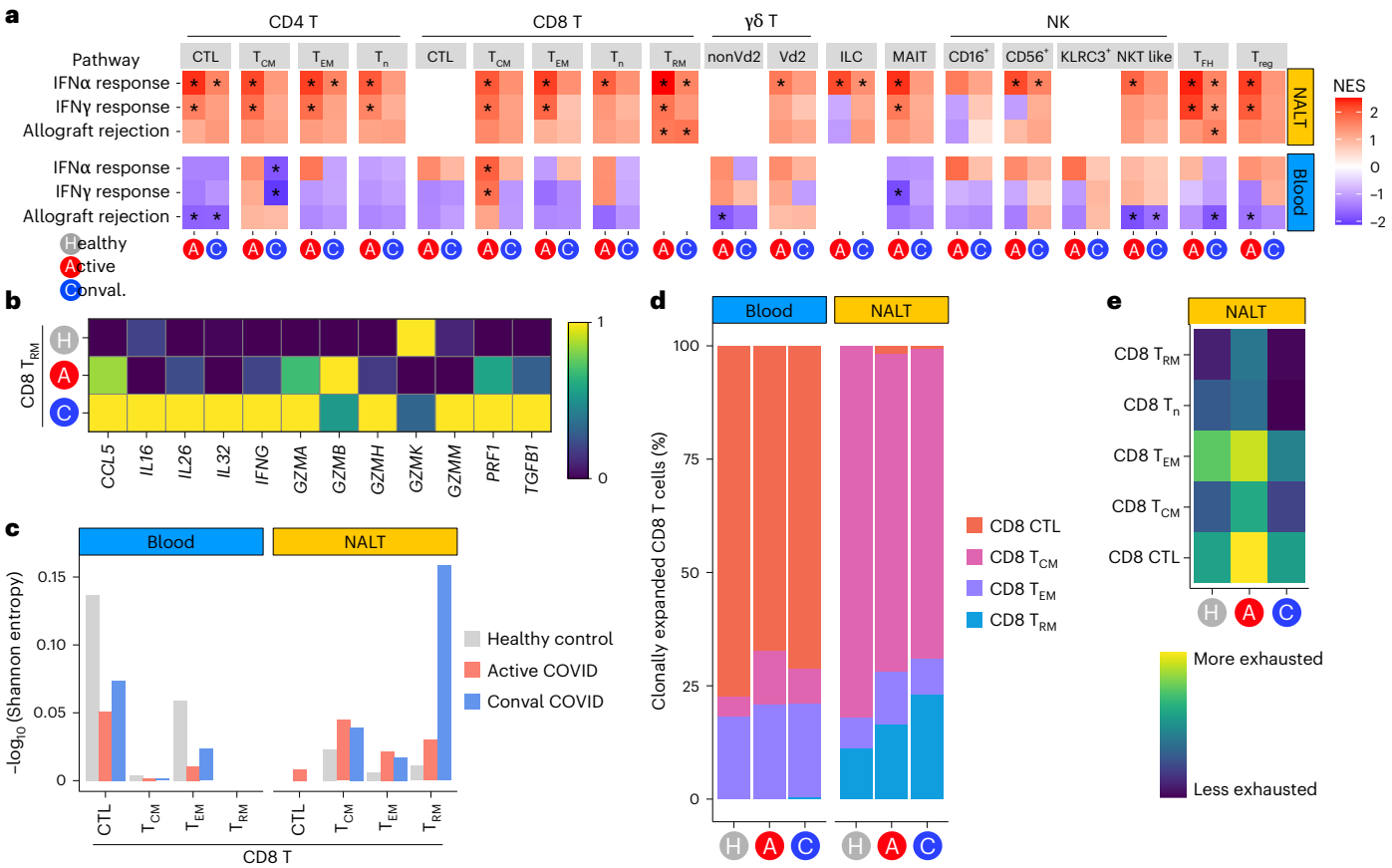


Fig. 5 | NALT CD8 T_{RM} cells exhibit clonal expansion and prolonged activation. **a**, Heatmap of a DEG enrichment of selected GSEA Hallmark gene sets in T cell and innate lymphocyte subsets by COVID status against healthy control samples. The color of the tile shows the NES. Red shows increased gene-set enrichment and blue decreased gene-set enrichment relative to healthy controls. *Statistical significance $P < 0.05$ and Benjamini–Hochberg $P_{adj} < 0.1$. **b**, Scaled heatmap of mean NALT CD8 T_{RM} lymphocyte (CD8 T_{RH} cell) gene expression by disease group, showing selected gene transcripts highly expressed in CD8 T_{RM} cells in active and convalescent COVID-19 participants compared with healthy controls. **c**, TCR clonality assessment in CD8 T cell subtypes. CD8 T cell TCR diversity is indicated by a Shannon entropy index and shown by cell subset and sample type, in active COVID-19 (red) and convalescent COVID-19 (blue) compared with healthy controls (gray). The Shannon entropy value has been $-\log_{10}$(transformed) so that higher values indicate lower TCR diversity or greater TCR clonality. **d**, Bar chart showing proportional representation of clonally expanded (clone size ≥2) CD8 T cell subsets in active and convalescent COVID-19 participants compared with healthy controls, as determined by single-cell TCR analysis, split by sample type. **e**, Heatmap showing expression (calculated as AddModuleScore) of human CD8 T cell viral 'exhaustion' signature[38] in NALT CD8 T cell subsets.

In COVID-19, there was an expansion of NALT CD8 T cells compared with healthy controls, particularly naive (T_n), central memory (T_{CM}) and T_{RM} cells (Fig. 4c). The last is consistent with studies of T_{RM} cells in murine lymph nodes, demonstrating that, although most tissue CD8 T_{RM} cells remain indefinitely resident, a small number of emigrants re-locate to local draining lymph nodes, becoming broadly distributed to enable rapid expansion and defense on local rechallenge[31].

As with myeloid cells, nasal T cell or innate lymphocyte subsets demonstrated greater infection-associated transcriptional changes compared with their circulating counterparts in acute COVID-19, particularly 'INFα response' and 'INFγ response' gene sets (Fig. 5a and Extended Data Fig. 5g). In convalescent COVID-19, the INFα response gene-set expression returned to healthy control levels in all NALT CD8 T cell subsets except for CD8 T_{RM} cells (Fig. 5a and Extended Data Fig. 6a). Indeed, CD8 T_{RM} cells showed evidence of ongoing activation in convalescent participants with significant enrichment of several immune-relevant gene sets (Fig. 5a and Extended Data Fig. 6b), including cytotoxicity-effector programs, with *IFNG* and *IL26* among notable leading-edge genes (Fig. 5b and Extended Data Fig. 6c,d). Of relevance, IL-26 binds to DNA released from damaged cells to promote damage-associated inflammation[32] and may also act directly on epithelial cells, promoting CXCL8 production[33]. In convalescence, NALT T_{RM} cells also expressed tissue repair-associated molecules including *TGFB1* and *IL32* (Fig. 5b). Single-cell T cell receptor (TCR) analysis showed the lowest levels of diversity in the nasal CD8 T_{RM} cell population in convalescent COVID, as measured by Shannon entropy (Fig. 5c), suggestive of clonal expansion. Indeed, comparing healthy control, acute and convalescent COVID-19, a progressive increase in clonally expanded T cells (defined as TCR clonotype with ≥2 cells) was uniquely detectable in T_{RM} cells among nasal CD8 T cells (Fig. 5d and Extended Data Fig. 6e,f). Expression of canonical CD8 T_{RM} cell transcripts was also restricted to the clonally expanded population (Extended Data Fig. 6g). Using 44,139 unique major histocompatibility complex (MHC)-I-matched SARS-COV-2-specific TCR CDR3 regions from a combination of the ImmunoCODE[34] and VDJdb[35] databases, we used GLIPH2 (ref. 36) as an in silico method to identify SARS-CoV-2-specific CD8 T cells through CDR3 sequence matching. This indicated an increase in virus-specific CD8 T cells in peripheral blood in acute COVID samples, but a later expansion in convalescent NALT CD8 T_{RM} cells (Extended Data Fig. 6h). In murine models, expression of ID2 and ID3 has been linked to the establishment of short- and long-term memory CD8 T cells, respectively[37]. CD8 T cells in both NALT and peripheral blood upregulated ID2 in acute and convalescent COVID-19, but upregulation of the long-term

memory marker ID3 was significant only in convalescent patients (Extended Data Fig. 6i).

Persistent CD8 T cell activation can lead to exhaustion. Overall, exhaustion-associated genes[38] and inhibitory molecules, such as *TIGIT*, *LAG3*, *VSIR* and *PDCD1* (encoding PD-1), were upregulated in convalescent compared with acute COVID-19 (Fig. 5e and Extended Data Fig. 6e,j). In the present study, CD8 T_{RM} cells again differed from other cytotoxic nasal CD8 T cell subsets, with less enrichment of exhaustion gene-set enrichment and lower expression of *TIGIT and PDCD1* (Fig. 5e and Extended Data Fig. 6e).

Together, our data reveal marked induction of IFN-induced antiviral programs in NALT T cells in active COVID-19 infection that rapidly subsides in convalescence. The exception to this was CD8 T_{RM} cells, which alone showed clonal expansion and evidence of ongoing cytotoxic activity in convalescence, as well as production of IFNγ, antimicrobial peptides and pro-repair molecules. Remarkably, despite this persistent activation, T_{RM} cells had reduced transcriptional evidence of exhaustion compared with other CD8 T cell subsets. Although T_{RM} cells have previously been described in human secondary lymphoid organs[39], our study provides a unique insight into how NALT T_{RM} cells respond during viral challenge, suggesting that they play parallel roles in defense, the coordination or recruitment of other T cell subsets and tissue repair.

Antiviral GC responses in SARS-CoV-2 infection

In the B cell compartment, we identified naive, nonswitched and switched memory, FCRL4$^+$ mucosa-associated memory B cells, immunoglobulin (Ig)A/IgG plasma cells (PCs) and early plasmablasts, as well as distinct clusters of GC-associated populations, including dark zone, light zone and memory precursors[4,40,41] (Fig. 6a and Extended Data Fig. 7a–c). Cells annotated as B GCs showed expression of canonical GC markers (*AICDA* and *BCL6*), but failed to enrich for either a light or dark zone signature specifically.

Most B cells profiled originated from nasal samples (Fig. 6b,c), confirmed by flow cytometry (Extended Data Fig. 7d,e). The nasal B cell compartment was proportionally expanded in both acute and convalescent COVID-19 compared with healthy controls, with a prominent increase in GC populations in convalescent samples (Fig. 6b). FCRL4$^+$ memory cells decreased in COVID-19 and there was also a proportional reduction in nasal switched memory B cells in convalescence compared with acute disease (Fig. 6b), confirmed by flow cytometry (Extended Data Fig. 7e).

Spatially, PCs were layered along the entire subepithelium in convalescence, with many IgG$^+$ and IgA$^+$ PCs co-localizing here, along with occasional IgD$^+$ PCs (Fig. 6d and Extended Data Fig. 7f,g), the latter previously described in tonsils and the nasopharynx[42]. Expression of the polyclonal Ig receptor (PIGR), required for transepithelial shuttling of IgA[43], was evident in NALT epithelium and increased in convalescent COVID-19 (Fig. 6e), augmenting the capacity to transport IgA generated by adjacent PCs into the nasal space. Transepithelial transport of IgG requires the neonatal Fc receptor (FcRn)[44]. *FCGRT* transcripts were detectable in nasal epithelial cells but did not show an increased post-COVID-19 infection and, indeed, even decreased (Fig. 6e), potentially indicating that IgG produced in NALT may bolster systemic rather than nasal IgG. Most IgG$^+$ PCs, as well as many IgA$^+$ PCs, expressed *CXCR4* (Fig. 6f), enabling localization to a *CXCL12*-expressing niche. The remaining IgA$^+$ PCs expressed *CCR10 and CXCR3* (Fig. 6f), with NALT stromal cells the only detectable source of *CCL28* (the ligand for CCR10), but myeloid cells the major source of the CXCR3 ligands *CXCL9*, *CXCL10* and *CXCL11* (Fig. 6g and Extended Data Fig. 7h). Indeed, macrophages and DCs also expressed *TNFSF13* and *TNSF13B* (encoding APRIL and BAFF, respectively; Fig. 6g and Extended Data Fig. 7h), survival factors for PCs[45], and co-localized with PC clusters (Fig. 6h).

Nasal B and plasmablast cells showed induction of IFNα response pathway genes in acute COVID-19 (Fig. 6i and Extended Data Fig. 8a). Notably, GC B cells showed substantial expression of several type I IFN-dependent antiviral genes, including *MX1*, *ISG15*, *IRF7* and *IFITM1/2* (Fig. 6j). MX1 directly inhibits viral ribonucleoprotein complexes and has been implicated in SARS-CoV-2 defense[46] and its expression by GC B cells indicates that these specialized adaptive immune cells retain cell-autonomous antiviral capacity while simultaneously differentiating to produce progeny with class-switched, somatically mutated B cell receptors (BCRs). Indeed, among immune cells, in acute COVID-19, light zone B cells showed the highest expression of IFN-induced protein 20 (*ISG20*), an RNA exonuclease with broad antiviral properties[47] (Fig. 6j).

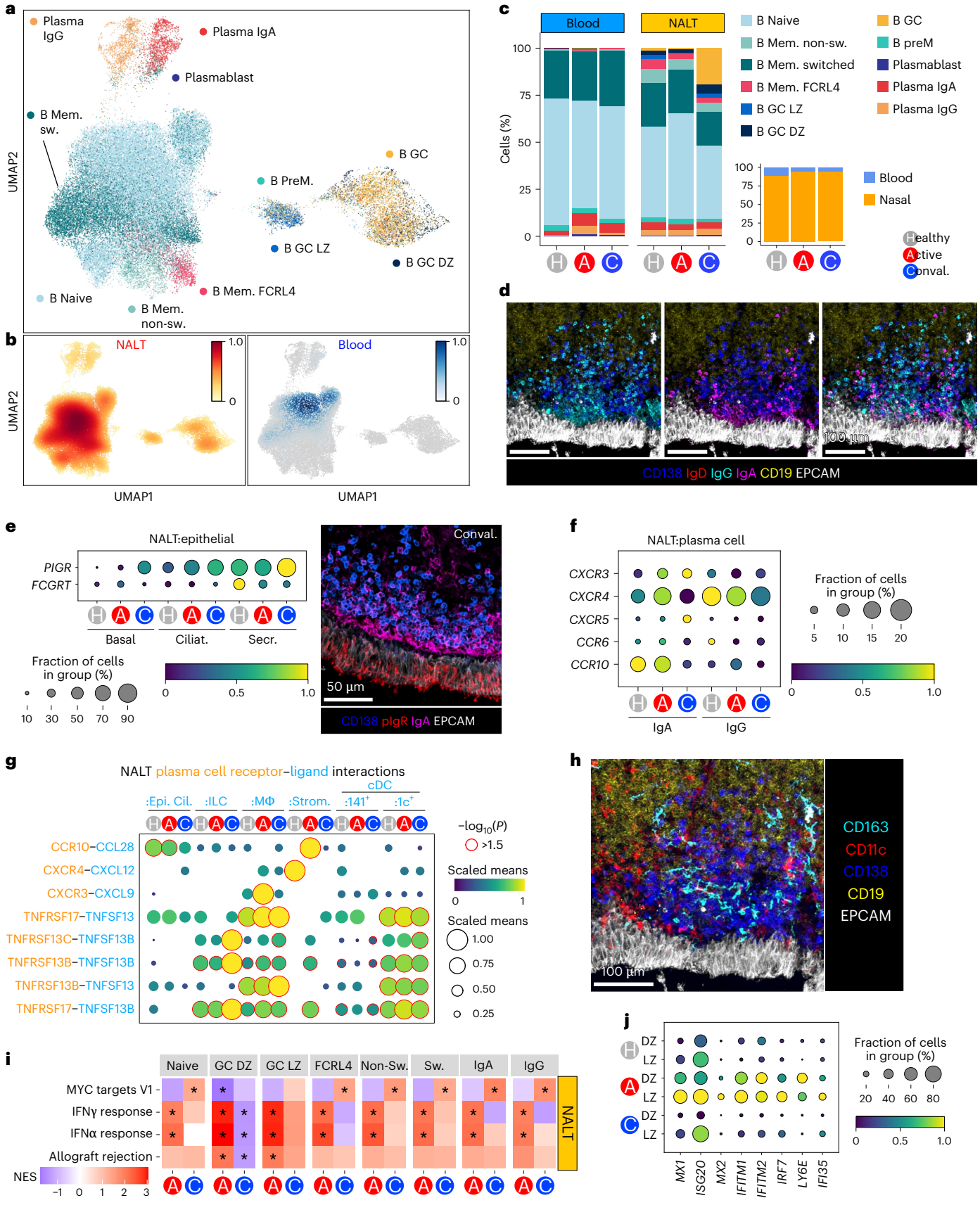
BCR analysis showed an increase in CDR3 junctional length in GC B cell populations in convalescent COVID-19 compared with other groups, suggesting increased V(D)J rearrangement and selection (Extended Data Fig. 8b). The Gini index (a measure of clonal selection) was also significantly increased in convalescent COVID-19 (Fig. 7a). Analysis of class switching in expanded clones (more than two) showed a marked expansion of IgM$^+$ clones (predominantly nonswitched memory B cell clones) and IgG1 and IgG2 clones (predominantly GC B cells in active COVID-19, with IgG clones more numerous in convalescence) (Fig. 7b). Thus, class switching to IgG rather than IgA appears to be the favored local mucosal response to SARS-CoV-2 infection. In acute COVID, we found an expansion of memory B cell clones, with expanded GC clones more prominent in convalescence (Fig. 7c).

Next, we sought to identify SARS-CoV-2-specific BCRs, comparing our single-cell BCR data to a reference COVID antibody database (CoV-AbDab)[48]. First, using GLIPH2 (Grouping of Lymphocyte Interactions by Paratope Hotspots v.2) to perform heavy chain matching with

Fig. 6 | GC and PC expansion in NALT after COVID infection. a, UMAP of B cells and PCs from patients with COVID-19 and healthy controls, reclustered in isolation, with assigned subset cell-type cluster labels shown. **b**, Scanpy embedding density (Gaussian kernel estimation) UMAP of B cells and PCs showing scaled density of cells by sample type. NALT embedding density is shown in orange and blood embedding density in blue. **c**, Stacked bar chart showing proportional representation of B cells and PCs in blood and NALT samples by disease status. The nested bar chart on the right shows the proportional split of B cells and PCs by disease type and sample type. Mem., memory; non-sw., nonswitched; GC, germinal center; LZ, light zone; DZ, dark zone; preM, prememory. **d**, Representative confocal imaging of subepithelial space on a section from a convalescent COVID-19 participant ($n = 1$). **e**, Dotplot showing PIGR and FCGRT expression in epithelial cell subset populations in NALT by disease state (left) and representative confocal imaging of epithelium on a section from a convalescent COVID-19 participant (right) ($n = 1$). Ciliat., ciliated; Secr., secretory. **f**, Dotplot showing selected chemokine receptor expression in PC populations in NALT, split by isotype. **g**, Dotplot showing CellPhoneDB cell–cell interaction prediction analysis for NALT PCs and ciliated epithelial cells, ILCs, macrophages, cDCs and stromal cells. The size and color of the dot indicate scaled mean, where the mean value refers to the total mean of the individual partner average expression values in the corresponding interacting pairs of cell types. The red dot border highlight indicates significance ($-\log_{10}(P) > 1.5$). The P values were calculated in the CellPhoneDB package using 1,000 permutations. **h**, Representative confocal imaging of the subepithelial space on a section from a convalescent COVID-19 patient ($n = 1$). **i**, Heatmap of DEG enrichment of GSEA Hallmark gene sets in B cell and PC subsets by COVID status against healthy control samples. The color of the tile shows the NES (red is increased gene-set enrichment and blue decreased gene-set enrichment relative to healthy controls), *Statistical significance $P < 0.05$ and $P_{adj} < 0.1$. NESs were calculated in the fgsea package with P-value estimation using an adaptive, multi-level, split Monte-Carlo scheme. **j**, Dotplot showing expression of selected, highly expressed, type I IFN response genes in B cell subsets in NALT. DZ, dark zone; LZ, light zone. In **e**, **f** and **j**, the size of the point indicates the fraction of cells in each group expressing the corresponding gene and the color of the point indicates the scaled mean expression of the corresponding gene in each named cell-type group.

https://doi.org/10.1038/s41590-024-02064-9

CoV-AbDab, SARS-CoV-2-binding BCR sequences were defined as cells with an identical CDR3-L sequence and a paired CDR3-H with significant motif similarity to an antibody within the CoV-AbDab (Extended Data Fig. 9a,b). Putative SARS-CoV-2-binding BCRs (n = 768) comprised around 5% of GC B cell BCRs in patients with COVID-19, but were not detectable in control GC B cell BCRs (Fig. 7d,e). Clonal expansion of SARS-CoV-2-specific B cells was also principally evident among GC B cells in convalescent COVID-19 participants (Fig. 7f and Extended



https://doi.org/10.1038/s41590-024-02064-9

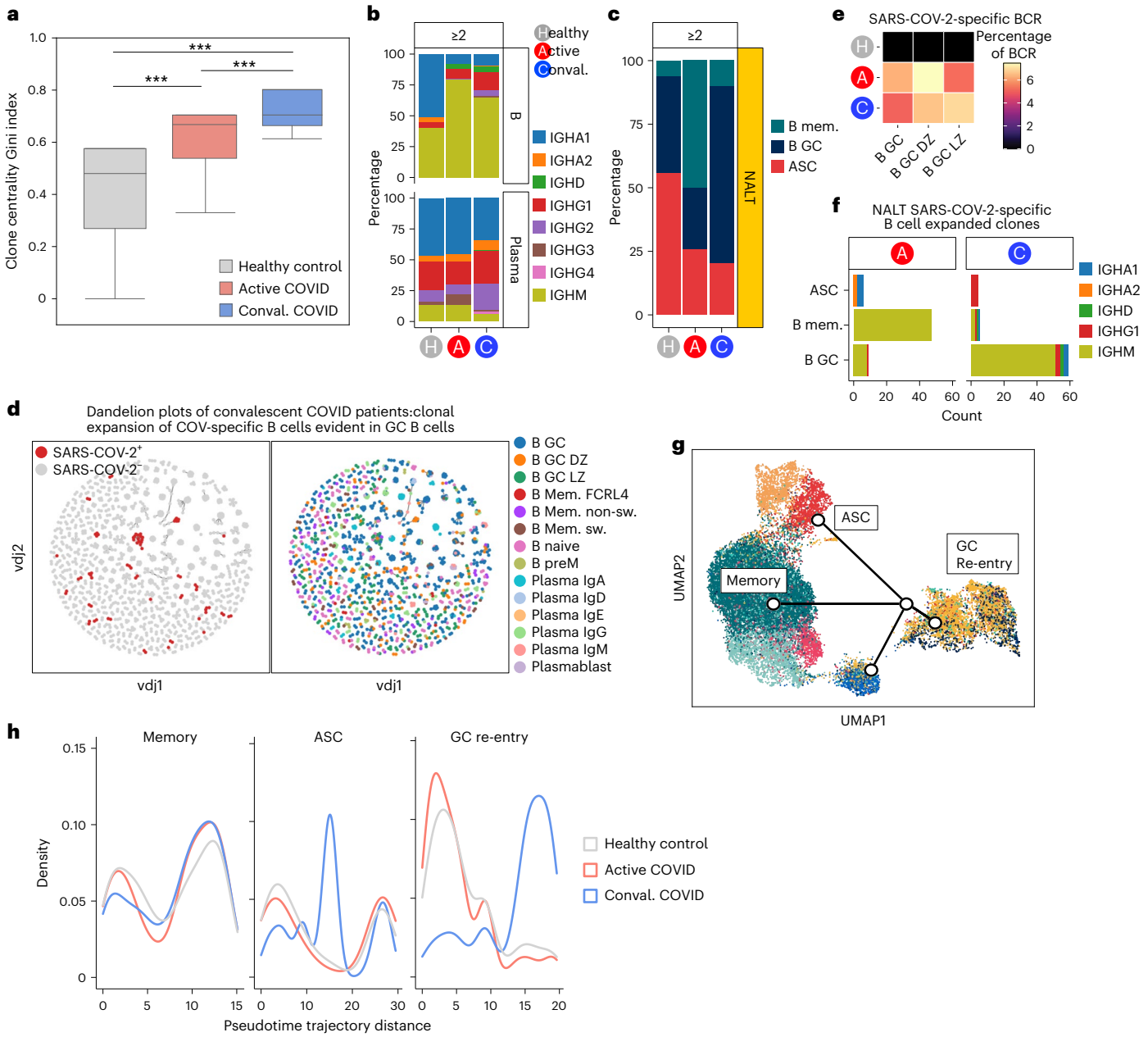


Fig. 7 | Type I IFN response and GC cell fate and progression. a, Box plot of NALT BCR Gini centrality index, as a measure of B cell clonality, split by disease state. Gini index values are on a scale from 0 to 1, where 1 indicates a monoclonal, highly mutated clonal response and 0, conversely, a polyclonal and unmutated response. Data are shown as box plots (median, box as 25th and 75th percentiles and whiskers as 1.5× the interquartile range). $^{***}P < 0.001$, two-sided Wilcoxon's rank-sum test ($n = 10$ healthy, $n = 8$ active, $n = 5$ convalescent). **b**, Bar chart showing proportional representation of Ig heavy chain isotypes in expanded NALT BCR clones (clone size ≥2), split by cell type and disease state. **c**, Bar chart showing proportional representation of memory, GC cells and ASCs in expanded NALT BCR clones. **d**, Single-cell BCR network plots for convalescent COVID-19 participants. Each circle or node corresponds to a single B cell with a corresponding set of BCR(s). Each clonotype is presented as a minimally connected graph with edge widths scaled to 1 per $d + 1$ for edge weight d, where d corresponds to the total (Levenshtein) edit distance of BCRs between two cells. The left-hand plot shows SARS-COV-2-specific clones (red) in the context of all B cell clones and the right-hand plot the assigned cell-type labels derived from gene expression data. **e**, Heatmap showing SARS-COV-2-specific BCRs as a percentage of total sequenced single-cell BCRs in GC B cells, by condition. **f**, Stacked bar chart showing counts of expanded (≥2) SARS-COV-2-specific B cell clones in NALT, by disease state, cell type and isotype. The x axis shows the cell type and the y axis the count; the color shows the isotype. B mem, memory B cell. **g**, B lineage cell Slingshot pseudotime trajectories, plotted in UMAP (corresponding to Fig. 6a). The start point is pinned on to B_GZ_LZ (light zone GC, royal blue). **h**, Cell density map along pseudotime trajectory distance for memory, GC re-entry and ASC lineages, split by disease state. The x axis indicates the pseudotime trajectory distance and the y axis the density of cells.

Data Fig. 9c,d), supporting the conclusion that this lymphoid tissue generates early local antiviral antibodies in COVID-19. As GLIPH2 is typically used for TCR matching, we also matched SARS-CoV-2-specific B cells identified by GLIPH2 with SARS-CoV-2-specific cells identified by Levenshtein distance similarity to CoV-AbDab BCRs. Although this identified a smaller number of SARS-CoV-2-specific BCRs, the pattern of expansion of SARS-COV-2-specific B cells mirrored the results generated by GLIPH2 (Extended Data Fig. 9e,f).

The cues that determine which cellular fate a GC B cell adopts and what signals initiate GC exit to a memory B cell or PC fate in humans

https://doi.org/10.1038/s41590-024-02064-9

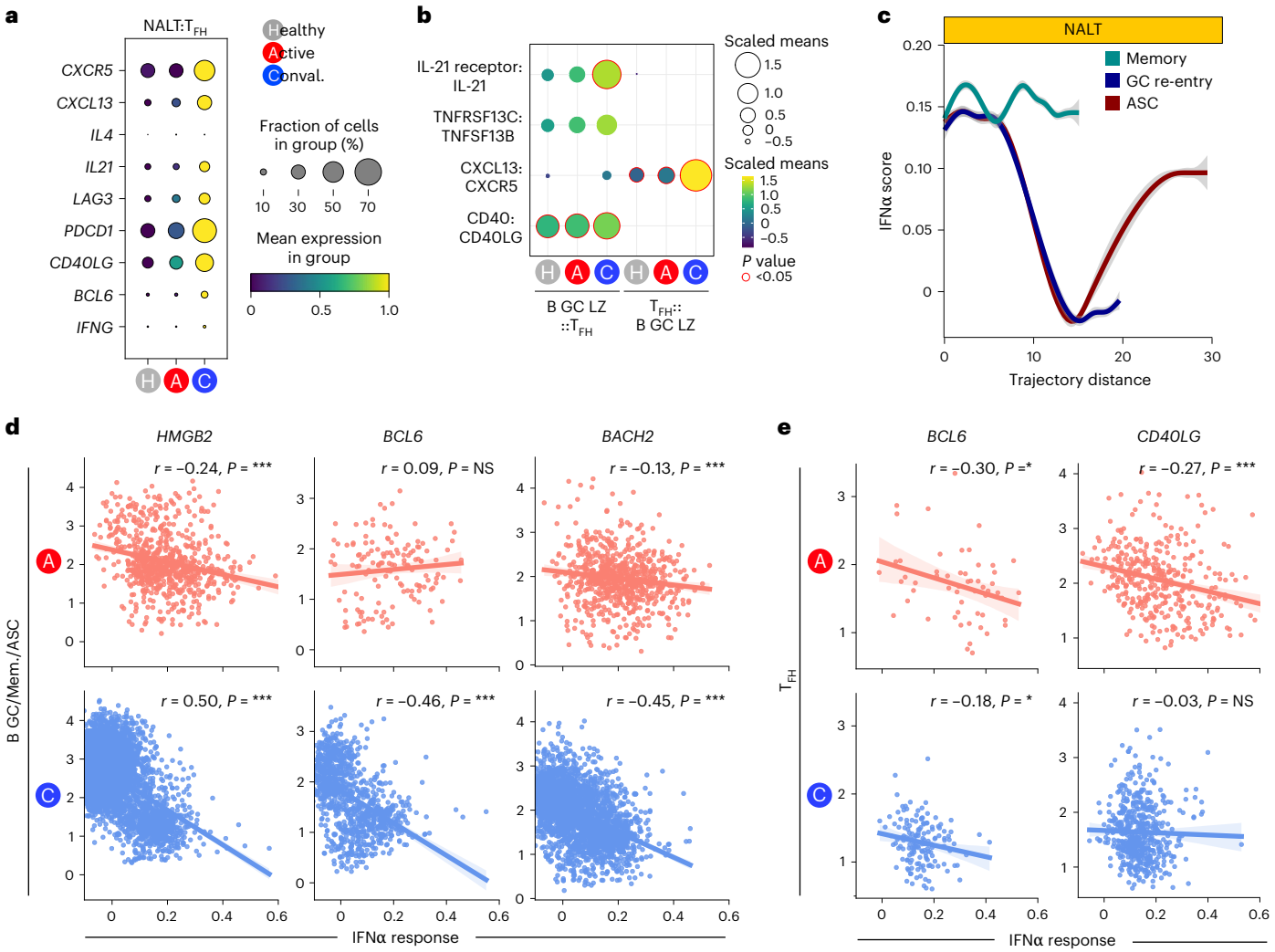


Fig. 8 | Type I IFN response and GC cell fate and progression. a, Dotplot showing gene expression in NALT T_{FH} cells, by disease state, of selected core genes associated with T_{FH} cell polarization and function. The size of the point indicates the fraction of cells in each group expressing the corresponding gene and the color of the point the scaled mean expression of the corresponding gene in each disease group. **b**, Dotplot showing CellPhoneDB cell–cell interaction prediction analysis for NALT light zone GC B cells (B GC LZ) and T_{FH} cells. The size and color of the dot indicate the scaled mean, where the mean value refers to the total mean of the individual partner average expression values in the corresponding interacting pairs of cell types. The red dot border highlight indicates the significance ($-\log_{10}(P) > 1.5$). **c**, Line graph of type I IFN response gene-set (GSEA Hallmark IFNα response) AddModuleScore in memory, GC and antibody-secreting B lineage cells along a pseudotime trajectory distance for memory, GC re-entry and ASC trajectories (trajectories shown in Fig. 6d). Memory is teal, GC re-entry dark blue and ASC red. **d**, Scatter plot with linear regression line plotting expression of type IFN response (GSEA Hallmark IFNα response) AddModuleScore against HMGB2, BCL6 and BACH2 expression in GC and memory B cells and PCs. The linear regression line is shown with a shaded area representing the 95% CI. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, NS (nonsignificant) $P > 0.05$. **e**, Scatter plot with linear regression line plotting expression of the type IFN response (GSEA Hallmark IFNα response) AddModuleScore against *BCL6* and *CD40L* in T_{FH} cells. The linear regression line is shown with a shaded area representing the 95% CI. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, NS $P > 0.05$.

remain unclear. Murine studies suggest that this decision is influenced by a combination of BCR antigen affinity, local cytokine environment and the provision of co-receptor signals by T_{FH} cells. A temporal model has been proposed, where the GC response undergoes a switch in its output as it matures, with memory B cells preferred at early time points and PCs later[49]. Temporal studies of GC progression in lymphoid organs in humans are lacking, so, to investigate this, we used Slingshot, pseudotime, cell-fate trajectory analysis[50]. To specify the root of this trajectory analysis, we used cellular entropy, a marker of transcriptional diversity that has been proposed as inversely proportional to cellular differentiation[51]. After dark zone GC cells, which are known to increase their genetic diversity through somatic hypermutation, this identified the B GC light zone cluster (B_GC_LZ) as the least differentiated (Extended Data Fig. 10a). Pseudotime trajectory analysis rooted in the light zone GC cluster revealed three distinct cell trajectories, one terminating in the antibody-secreting cell (ASC) cluster, the second in the switched memory B cell cluster and the third in the GC dark zone population, which we termed a 'GC re-entry' fate (Fig. 7g). When comparing disease groups and control, we found a paucity of cells in the terminal stages of the GC re-entry trajectory in acute COVID-19, in contrast to convalescent COVID-19, where many cells were evident in the late GC re-entry trajectory. Cells from convalescent COVID-19 patients also showed an early peak in density in the ASC trajectory compared with acute COVID-19 and controls (Fig. 7h). Taken together with our analysis of expanded clones, our pseudotime data are consistent with a temporal model of GC progression.

Murine studies suggest that the switch in GC output over time may be mediated by changes in the transcriptional characteristics of T_{FH} cells, with the ratio IL-21:IL-4 decreasing over time, and a switch toward IL-4 production. In our dataset, T_{FH} cells showed increased expression

of *IL21* in convalescent COVID-19 compared with acute disease and control samples, with little *IL4* detectable (Fig. 8a). T_{FH} cell–light zone GC B cell interaction prediction using CellPhoneDB also showed a temporal increase in several interactions that would promote GC progression, including *CXCL13 CXCR5*, *TNFRSF13C–TNFSF13B*, *IL21R–IL21* and *CD40–CD40L* in convalescent COVID (Fig. 8b). Using cell2TCR[52], we matched T cell subsets in the NALT to an experimentally derived T cell activation signature generated from a SARS-CoV-2 human infection model[52], and used this to plot T cell subset activation profiles in NALT over the course of infection. This showed a significant increase in T_{FH} and follicular regulatory T (T_{FR}) cell activation in convalescent disease, the latter T cell expressing increased levels of *TNFRSF18* (GITR), CTLA-4 and PD-1 (Extended Data Fig. 10c).

We hypothesized that, during viral infection, type I IFNs may provide an environmental cue that influences GC output. Assessment of gene expression showed persistent induction of type I IFN response genes in cells across the memory B cell trajectory, in contrast to cells progressing along a GC re-entry or ASC trajectory (Fig. 8c). It is interesting that, in GC B cells, there was a negative correlation between type I IFN response genes and *HMGB2* and *BACH2* in SARS-CoV-2 infection, particularly in convalescent disease (Fig. 8d and Extended Data Fig. 10d,e). In NALT T_{FH} cells, type I IFN response genes showed a negative correlation with *BCL6* in COVID-19 (Fig. 8e). Altogether, this is consistent with the conclusion that type 1 IFN may inhibit the progression of the GC reaction, potentially providing an explanation as to why severe COVID-19 is associated with limited GC formation.

Discussion

Although there is macroscopic involution of the pharyngeal tonsils with age, our study showed that the postnasal space (PNS) contains lymphoid tissue in adults that is responsive to viral infection and can be sampled to profile the cellular molecular processes occurring during the generation of adaptive immune responses in living participants. We showed that these NALT biopsies can be collected in an outpatient clinic setting, providing a means of longitudinally assessing secondary lymphoid tissue in humans. NALT immune responses are particularly relevant to pathogens that enter via the nasal mucosa, including SARS-CoV-2, and to nasal vaccines[10,11]. We demonstrated that NALT profiling enables the study of cell types that are central to adaptive immunity, but not present in blood, particularly GCs and stromal cells. A recent study using PNS swabs similarly captured adenoid GC populations in the context of SARS-CoV-2 infection[53]. This methodology has the advantage of enabling longitudinal sampling but does not provide spatial information or include stromal or significant numbers of myeloid cells.

Lymphoid tissues contain several myeloid cell populations. We found that, in acute COVID-19, monocytes formed a superficial lining in NALT, expressing the antibacterial protein calprotectin[21], as well as self- and neutrophil-recruiting chemokines. Neutrophils were also evident in this zone, some of which had histone citrullination, required for chromatin decondensation and NET formation[22]. NETs can also entrap and kill bacteria[54], potentially acting together with monocytes to defend the underlying follicles and T cells from nasal bacteria during a period of vulnerability when overlying epithelial cells are damaged or killed by invading virus.

Murine studies show that subcapsular sinus macrophages play important roles in viral defense[8]. Furthermore, lymphatic sinus-lining macrophages produce IL-18 in response to bacterial infection, stimulating innate lymphocytes to produce IFNγ, which enhanced macrophage antibacterial function[9]. The pharyngeal tonsils do not have afferent lymphatic drainage via a subcapsular sinus, but rather antigens and pathogens come directly from the overlying nasal space. In the present study, we found that human NALT contains a dense network of macrophages, including a *FOLR2*$^+$ population, which expanded in convalescent COVID-19, expressing *IL18*, and pro-repair molecules, including type VI collagen.

Murine models of upper respiratory tract influenza challenge have shown CD8 T_{RM} cell expansion in the nasal tissue[55] and SARS-CoV-2-specific CD8 T_{RM} cells have also been described in human nasal mucosa[56]. Our study confirmed their presence in NALT and showed that, unlike other CD8 T cell subsets, T_{RM} cells clonally expand and show prolonged cytotoxic gene expression in convalescence, with little evidence of exhaustion. This is in contrast to previous reports suggesting T_{RM} cell exhaustion in active COVID-19 infection[57]. It is interesting that NALT T_{RM} cells also expressed lymphocyte-recruiting chemokines and *TGFB1*, suggesting that they play additional roles beyond immediate antiviral defense.

We found reciprocal changes in the expression of IgG and IgA transporters in NALT epithelium, which suggest an increase in IgA transport to the nasal mucosa, but that NALT IgG contributes to systemic immunity. We did not take nasal secretion or saliva samples to directly measure Ig levels; however, studies delineating secretory IgA, IgA and IgG responses in SARS-CoV-2 support this finding[58,59].

Our longitudinal analysis of GC progression in SARS-CoV-2 infection provides unique insights into the dynamics of GC responses in humans. Although studies in mice suggest a temporal model where memory B cells emerge earlier in responses, this has been difficult to confirm in humans. We found increased representation of memory B cells in expanded clones in acute COVID-19, which decreased in convalescence, and our pseudotime analysis showed a peak in cell density in the ASC trajectory in convalescence. Together, these analyses are supportive of a temporal model of GC progression. We also found that type I IFN may provide an environmental cue that influences GC output, with memory cells showing the highest expression of type I IFN response genes throughout their trajectory.

In the present study, we showed that adults with mild COVID-19 disease do form GCs in response to SARS-CoV-2. This enabled us to probe potential mechanisms of GC collapse described in fatal disease[19]. We found that high IFN response gene expression was associated with reduced expression of *BACH2* and *BCL6* in GC B cells and reduced expression of *CD40L* in T_{FH} cells. This implies that an excess of type I IFNs, as has been described in severe disease, may inhibit the progression of the GC reaction.

A limitation to our study is that the acute and convalescent samples were not taken from the same participants and two out of eight of the acute COVID-19 patients were on corticosteroids, one on corticosteroids and baricitinib and one on corticosteroids and tocilizumab, compared with the five convalescent participants, none of whom had severe disease or were on immunomodulating medications, which may influence the acute and convalescent comparisons.

In summary, our study provides a unique insight into how nasal adaptive immune responses are generated and defended during SARS-CoV-2 infection. Beyond COVID-19, our study provides proof of principle that this sampling strategy will enable long-sought efforts to interrogate longitudinal changes in lymphoid organs in humans in disease and after therapeutic intervention.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41590-024-02064-9.

References

1. Liu, Y. J., Zhang, J., Lane, P. J., Chan, E. Y. & MacLennan, I. C. Sites of specific B cell activation in primary and secondary responses to T cell-dependent and T cell-independent antigens. *Eur. J. Immunol.* **21**, 2951–2962 (1991).

2. Han, S. et al. Cellular interaction in germinal centers. Roles of CD40 ligand and B7-2 in established germinal centers. *J. Immunol.* **155**, 556–567 (1995).
3. Linterman, M. A. et al. IL-21 acts directly on B cells to regulate Bcl-6 expression and germinal center responses. *J. Exp. Med.* **207**, 353–363 (2010).
4. Victora, G. D. et al. Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter. *Cell* **143**, 592–605 (2010).
5. Gitlin, A. D., Shulman, Z. & Nussenzweig, M. C. Clonal selection in the germinal centre by regulated proliferation and hypermutation. *Nature* **509**, 637–640 (2014).
6. Hase, H. et al. BAFF/BLyS can potentiate B-cell selection with the B-cell coreceptor complex. *Blood* **103**, 2257–2265 (2004).
7. Cyster, J. G. et al. Follicular stromal cells and lymphocyte homing to follicles. *Immunol. Rev.* **176**, 181–193 (2000).
8. Junt, T. et al. Subcapsular sinus macrophages in lymph nodes clear lymph-borne viruses and present them to antiviral B cells. *Nature* **450**, 110–114 (2007).
9. Kastenmuller, W., Torabi-Parizi, P., Subramanian, N., Lammermann, T. & Germain, R. N. A spatially-organized multicellular innate immune response in lymph nodes limits systemic pathogen spread. *Cell* **150**, 1235–1248 (2012).
10. Gallichan, W. S., Johnson, D. C., Graham, F. L. & Rosenthal, K. L. Mucosal immunity and protection after intranasal immunization with recombinant adenovirus expressing herpes simplex virus glycoprotein B. *J. Infect. Dis.* **168**, 622–629 (1993).
11. Wu, H. Y., Nikolova, E. B., Beagley, K. W. & Russell, M. W. Induction of antibody-secreting cells and T-helper and memory cells in murine nasal lymphoid tissue. *Immunology* **88**, 493–500 (1996).
12. Boyaka, P. N. et al. Human nasopharyngeal-associated lymphoreticular tissues. Functional analysis of subepithelial and intraepithelial B and T cells from adenoids and tonsils. *Am. J. Pathol.* **157**, 2023–2035 (2000).
13. Ahn, J. H. et al. Nasal ciliated cells are primary targets for SARS-CoV-2 replication in the early stage of COVID-19. *J. Clin. Invest.* https://doi.org/10.1172/jci148517 (2021).
14. Guan, W.-J. et al. Clinical characteristics of coronavirus disease 2019 in China. *N. Engl. J. Med.* **382**, 1708–1720 (2020).
15. Ren, X. et al. COVID-19 immune features revealed by a large-scale single-cell transcriptome atlas. *Cell* **184**, 1895–1913.e1819 (2021).
16. Liao, M. et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* **26**, 842–844 (2020).
17. Yoshida, M. et al. Local and systemic responses to SARS-CoV-2 infection in children and adults. *Nature* **602**, 321–327 (2022).
18. Xu, Q. et al. Adaptive immune responses to SARS-CoV-2 persist in the pharyngeal lymphoid tissue of children. *Nat. Immunol.* **24**, 186–199 (2023).
19. Kaneko, N. et al. Loss of Bcl-6-expressing T follicular helper cells and germinal centers in COVID-19. *Cell* **183**, 143–157 e113 (2020).
20. Trevelin, S. C. et al. Disrupted Peyer's patch microanatomy in COVID-19 including germinal centre atrophy independent of local virus. *Front. Immunol.* **13**, 838328 (2022).
21. Kehl-Fie, T. E. et al. Nutrient metal sequestration by calprotectin inhibits bacterial superoxide defense, enhancing neutrophil killing of *Staphylococcus aureus*. *Cell Host Microbe* **10**, 158–164 (2011).
22. Wang, Y. et al. Histone hypercitrullination mediates chromatin decondensation and neutrophil extracellular trap formation. *J. Cell Biol.* **184**, 205–213 (2009).
23. Liberzon, A. et al. The molecular dignatures database Hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
24. Iwabuchi, R. et al. Development of an inflammatory $CD14^+$ dendritic cell subset in humanized mice. *Front. Immunol.* **12**, 643040 (2021).
25. Langlet, C. et al. CD64 expression distinguishes monocyte-derived and conventional dendritic cells and reveals their distinct role during intramuscular immunization. *J. Immunol.* **188**, 1751–1760 (2012).
26. Dick, S. A. et al. Three tissue resident macrophage subsets coexist across organs with conserved origins and life cycles. *Sci. Immunol.* **7**, eabf7777 (2022).
27. Zhang, J. et al. Secreted M-ficolin anchors onto monocyte transmembrane G protein-coupled receptor 43 and cross talks with plasma C-reactive protein to mediate immune signaling and regulate host defense. *J. Immunol.* **185**, 6899–6910 (2010).
28. Xue, J. et al. Transcriptome-based network analysis reveals a spectrum model of human macrophage activation. *Immunity* **40**, 274–288 (2014).
29. von der Mark, H., Aumailley, M., Wick, G., Fleischmajer, R. & Timpl, R. Immunochemistry, genuine size and tissue localization of collagen VI. *Eur. J. Biochem.* **142**, 493–502 (1984).
30. Schraufstatter, I. U., Zhao, M., Khaldoyanidi, S. K. & Discipio, R. G. The chemokine CCL18 causes maturation of cultured monocytes to macrophages in the M2 spectrum. *Immunology* **135**, 287–298 (2012).
31. Beura, L. K. et al. T Cells in nonlymphoid tissues give rise to lymph-node-resident memory T cells. *Immunity* **48**, 327–338 e325 (2018).
32. Larochette, V. et al. IL-26, a cytokine with roles in extracellular DNA-induced inflammation and microbial defense. *Front. Immunol.* **10**, 204 (2019).
33. Hor, S. et al. The T-cell lymphokine interleukin-26 targets epithelial cells through the interleukin-20 receptor 1 and interleukin-10 receptor 2 chains. *J. Biol. Chem.* **279**, 33343–33351 (2004).
34. Nolan, S. et al. A large-scale database of T-cell receptor beta (TCRβ) sequences and binding associations from natural and synthetic exposure to SARS-CoV-2. Preprint at *Research Square* https://doi.org/10.21203/rs.3.rs-51964/v1 (2020).
35. Goncharov, M. et al. VDJdb in the pandemic era: a compendium of T cell receptors specific for SARS-CoV-2. *Nat. Methods* **19**, 1017–1019 (2022).
36. Huang, H., Wang, C., Rubelt, F., Scriba, T. J. & Davis, M. M. Analyzing the *Mycobacterium tuberculosis* immune response by T-cell receptor clustering with GLIPH2 and genome-wide antigen screening. *Nat. Biotechnol.* **38**, 1194–1202 (2020).
37. Yang, C. Y. et al. The transcriptional regulators Id2 and Id3 control the formation of distinct memory $CD8^+$ T cell subsets. *Nat. Immunol.* **12**, 1221–1229 (2011).
38. Quigley, M. et al. Transcriptional analysis of HIV-specific $CD8^+$ T cells shows that PD-1 inhibits T cell function by upregulating BATF. *Nat. Med.* **16**, 1147–1151 (2010).
39. Sathaliyawala, T. et al. Distribution and compartmentalization of human circulating and tissue-resident memory T cell subsets. *Immunity* **38**, 187–197 (2013).
40. Victora, G. D. et al. Identification of human germinal center light and dark zone cells and their relationship to human B-cell lymphomas. *Blood* **120**, 2240–2248 (2012).
41. Holmes, A. B. et al. Single-cell analysis of germinal-center B cells informs on lymphoma cell of origin and outcome. *J. Exp. Med.* **217**, e20200483 (2020).
42. Gutzeit, C., Chen, K. & Cerutti, A. The enigmatic function of IgD: some answers at last. *Eur. J. Immunol.* **48**, 1101–1113 (2018).
43. Kaetzel, C. S., Robinson, J. K., Chintalacharuvu, K. R., Vaerman, J. P. & Lamm, M. E. The polymeric immunoglobulin receptor (secretory component) mediates transport of immune complexes across epithelial cells: a local defense function for IgA. *Proc. Natl Acad. Sci. USA* **88**, 8796–8800 (1991).
44. Israel, E. J. et al. Expression of the neonatal Fc receptor, FcRn, on human intestinal epithelial cells. *Immunology* **92**, 69–74 (1997).

45. Lindquist, R. L., Niesner, R. A. & Hauser, A. E. In the right place, at the right time: spatiotemporal conditions determining plasma cell survival and function. *Front. Immunol.* **10**, 788 (2019).
46. Bizzotto, J. et al. SARS-CoV-2 infection boosts MX1 antiviral effector in COVID-19 patients. *iScience* **23**, 101585 (2020).
47. Wu, N. et al. The interferon stimulated gene 20 protein (ISG20) is an innate defense antiviral factor that discriminates self versus non-self translation. *PLoS Pathog.* **15**, e1008093 (2019).
48. Raybould, M. I. J., Kovaltsuk, A., Marks, C. & Deane, C. M. CoV-AbDab: the coronavirus antibody database. *Bioinformatics* **37**, 734–735 (2021).
49. Weisel, F. J., Zuccarino-Catania, G. V., Chikina, M. & Shlomchik, M. J. A temporal switch in the germinal center determines differential output of memory B and plasma cells. *Immunity* **44**, 116–130 (2016).
50. Street, K. et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genom.* **19**, 477 (2018).
51. Gulati, G. S. et al. Single-cell transcriptional diversity is a hallmark of developmental potential. *Science* **367**, 405–411 (2020).
52. Lindeboom, R. G. H. et al. Human SARS-CoV-2 challenge uncovers local and systemic response dynamics. *Nature* **631**, 189–198 (2024).
53. Ramirez, S. I. et al. Immunological memory diversity in the human upper airway. *Nature* **632**, 630–636 (2024).
54. Brinkmann, V. et al. Neutrophil extracellular traps kill bacteria. *Science* **303**, 1532–1535 (2004).
55. Pizzolla, A. et al. Resident memory CD8$^+$ T cells in the upper respiratory tract prevent pulmonary influenza virus infection. *Sci. Immunol.* **2**, eaam6970 (2017).
56. Roukens, A. H. E. et al. Prolonged activation of nasal immune cell populations and development of tissue-resident SARS-CoV-2-specific CD8$^+$ T cell responses following COVID-19. *Nat. Immunol.* **23**, 23–32 (2022).
57. Rha, M.-S. & Shin, E.-C. Activation or exhaustion of CD8+ T cells in patients with COVID-19. *Cell. Mol. Immunol.* **18**, 2325–2333 (2021).
58. Dowell, A. C. et al. Nasal mucosal IgA levels against SARS-CoV-2 and seasonal coronaviruses are low in children but boosted by reinfection. *J. Infect.* **87**, 403–412 (2023).
59. Bladh, O. et al. Comparison of SARS-CoV-2 spike-specific IgA and IgG in nasal secretions, saliva and serum. *Front. Immunol.* https://doi.org/10.3389/fimmu.2024.1346749 (2024).
60. Goel, R. R. et al. mRNA vaccines induce durable immune memory to SARS-CoV-2 and variants of concern. *Science* **374**, abm0829 (2021).

Methods

The list of reagents used in the present study can be found in Supplementary Table 2.

Human participant recruitment

COVID-19 participants were recruited from a single site—Addenbrooke's Hospital, Cambridge, UK—between May 2020 and July 2021. Healthy control individuals were recruited between August 2019 and November 2019 from Cambridge, UK. Ethical approval was given through the Cambridge Central Research Ethics Committee (reference no. 08/H0308/267, IRAS project no. 194217), administered through Cambridge University Hospitals NHS Foundation Trust. All participants provided informed written consent and, thus, patients unable to provide this, including those undergoing invasive ventilation, were excluded. The following participants were also excluded: aged <18 years, with a chronic systemic infection (for example, HIV, viral hepatitis), with a diagnosis of active malignancy, previous head and neck radiotherapy, a diagnosis of systemic autoimmune disease (for example, rheumatoid arthritis, ANCA-associated vasculitis) or a known nasal or postnasal anomaly precluding access or safe biopsy (for example, severe septal deviation, nasal vascular malformation). All active COVID-19 participants, with the exception of participant no. 1, were recruited within 7 d of their first positive swab (quantitative RT–PCR) result and within 14 d of symptom onset. Participant no. 1 was recruited on the basis of radiological changes consistent with COVID-19, in the presence of positive COVID-19 serology and the onset of symptoms consistent with COVID-19 within the past 14 d. All convalescent COVID-19 participants were recruited between 21 d and 28 d after a positive COVID-19 swab result. No COVID-19 participants had received a COVID-19 vaccine at the time of recruitment.

Sample collection

Nasal samples were collected by an otolaryngologist under direct endoscopic visualization. Briefly, the nasal cavity was anesthetized with topical local anesthetic and epinephrine (lidocaine 5%:phenylephrine 0.5%) administered via spray and neurosurgical patties (Codman). Nasal brushings and microcurettage samples were collected from the inferior turbinate under direct vision with the aid of a Thudichum nasal speculum. The PNS was inspected transnasally with a 0° nasendoscope and one to three biopsies taken from the PNS mucosa using Tilley–Henkel, Takahashi fine-cupped or Blakesley forceps, and placed directly into Roswell Park Memorial Institute (RPMI)-1640 medium on ice. Blood samples were collected contemporaneously into sodium citrate tubes using the vacutainer system (BD)

Sample processing to single-cell suspension

All samples were processed fresh, within 1 h of collection. PNS samples were washed in RPMI-1640 medium, diced and placed into a 15-ml Falcon tube with 5 ml of RPMI-1640 medium and a digestion solution of 62.5 µl of Liberase (stock concentration 1 mg ml^{-1}; Sigma-Aldrich) and 250 µl of DNase I (stock concentration 2.5 mg ml^{-1}; Roche). The sample tube was placed into a shaking incubator at 37 °C for 30 min at 220 r.p.m. The sample was then pushed through a 70-µm cell strainer (Falcon) into a 50-ml Falcon tube and washed in 50 ml of RPMI-1640 medium at 500*g* for 6 min with the brake on. The cell pellet was resuspended in 10 ml of 44% Percoll solution and centrifuged at 800*g* for 20 min and 20 °C with no brake. The supernatant was aspirated. If the cell pellet had visible blood staining, 2 ml of ACK red cell lysis buffer was added for 2 min, before quenching with 20 ml of 1× phosphate-buffered saline (PBS) and centrifuging at 500*g* and 4 °C for 6 min (brake on). After aspirating the supernatant, the cell pellet was resuspended in 100 µl of 1× PBS. Nasal brushing heads were vigorously agitated 15× in 5 ml of RMPI-1640 collection fluid. The brush head was cut off and left in the collection fluid while it was centrifuged at 500*g* for 6 min at 4 °C (brake on). The brush head and supernatant were removed and the cell pellet resuspended in 100 µl. Nasal microcurettes were agitated in 5 ml of RPMI-1640 collection fluid before centrifugation at 500*g* for 6 min (4 °C, brake on). The resulting cell pellet was resuspended in a digestion fluid of Liberase or DNase I at the same concentration as the PNS samples and placed into a shaking incubator for 20 min at 37 °C and 220 r.p.m. After digestion, samples were disassociated by pipette mixing and washed at 500*g* for 6 min (brake on) and resuspended in 100 µl of 1× PBS. Then, 18 ml of blood from the sodium citrate collection tubes was diluted 1:1 with 1× PBS and 15 ml of Histopaque 1077 (Sigma-Aldrich) was added to each of two 50-ml Falcon tubes, with the blood–PBS mixture layered on top. Each tube was centrifuged at 800*g* for 20 min at room temperature (low brake). The resulting PBMC layer was aspirated with a Pasteur pipette and washed twice in 50 ml of PBS before being resuspended in 500 µl of 1× PBS.

The samples were counted with Trypan Blue staining to assess cell counts and viability, before downstream applications.

ScRNA-seq library generation

PNS samples were pooled with curettage and brushing cells at a ratio of 2:1:1. Where nasal curettage and/or nasal brushing cell counts were insufficient or unavailable, only PNS samples were used. Blood and nasal samples were diluted to a concentration of 2×10^6 cells per ml and loaded on to a 10x Chromium controller in parallel lanes with a targeted cell recovery of 14,000 cells. Single-cell libraries + TCR and BCR VDJ libraries were created using Chromium Next GEM (Gel Beads in Emulsion) Single cell 5′ kit v.1.1 (10x Genomics), in accordance with the manufacturer's user guide. Samples were sequenced using a Novaseq 6000 S4 system (Illumina).

Flow cytometry

After scRNA-seq GEM creation, the remaining single-cell suspensions were diluted to a concentration of 2×10^6 cells in 100 µl of 1× PBS. If this concentration was not achievable, all remaining cells were used. Cells were blocked with 5 µl of FcR blocking reagent (Miltenyi-Biotech) before undergoing surface staining with fluorophore-labeled antibodies. The antibodies used in the present study are listed in Supplementary Table 2. The samples were washed and resuspended in 1× PBS and run on a Fortessa II flow cytometer (BD). Data were gated into populations using FlowJo (Treestar/BD). All populations used in this publication were gated on a starting population of CD45$^+$ live singlets. Gating strategies are available in Supplementary Fig. 1.

Immunofluorescence microscopy

PNS biopsies were collected as described above and fixed for 15–30 min (depending on tissue size) in AntigenFix (DiaPath, cat. no. P0016), followed by 8 h in 30% sucrose in PBS before embedding in OCT (CellPath, cat. no. KMA-0100-00A). Then, 15-µm sections were mounted on Polysine slides (Thermo Fisher Scientific, cat. no. J2800AMNZ) and permeabilized and blocked in 0.1 M Tris, pH 7.4 containing 0.1% Triton (Sigma-Aldrich, cat. no. 93426-100ML), 1% normal mouse serum (Invitrogen, cat. no. 10410), 1% normal donkey serum (Abcam, cat. no. ab7475) and 1% bovine serum albumin (R&D Systems, cat. no. DY995) for 1 h at 20 °C. Samples were stained for 2 h at 20 °C (primary antibodies) or 1 h at 20 °C (secondary antibodies) in a humid chamber with the appropriate antibodies, washed 3× in PBS and mounted in Fluoromount-G (Southern Biotech, cat. no. 0100-01). Images were captured using a TCS SP8 (Leica) inverted microscope, on a ×40, 1.1 numerical aperture, water immersion objective. Raw imaging data were processed using Imaris (Bitplane/Oxford Instruments). All antibodies used in the present study are listed in Supplementary Table 2.

Iterative staining of PNS biopsies was performed as described by Radtke et al.[61,62]. Staining and imaging were performed as described above. After imaging, the coverslip was removed and the slides washed 3× in PBS to remove any mounting medium. Bleaching of the fluorochromes was achieved using 1 mg ml^{-1} of a solution of lithium

borohydride (Acros Organics, cat. no. 206810050) in water for 15 min at 20 °C. The slides were then washed 3× in PBS before staining as described above with a different set of antibodies. The process was repeated up to 7×. Raw imaging data were processed using Imaris with either Hoechst or CD19 as fiducial for alignment of subsequent images.

Quantification and statistical analysis

ScRNA-seq data processing. Sequenced data were aligned to hu38 human reference genome using Cellranger v.6 (10x Genomics). Removal of ambient RNA and assigning of droplets were undertaken using Cellbender[63]. Automated doublet detection and filtering were performed using Scrublet[64] ($P < 0.1$), as well as manually by gene co-expression. Quality control and filtering were performed in Scanpy[65], using the sc-dandelion pre-processing module (dandelion.pp,recipe_scanpy_qc) with max. genes = 6000, min. genes = 200 and a Gaussian mixture model distribution to determine mitochondrial filtering values. Batch correction was performed with Harmony[66] and Leiden clustering was also performed. Cell labels were assigned manually by gene expression profiles (Supplementary Fig.1c), with iterative rounds of subclustering to assign cell subtype labels. Assigned cluster labels were checked against signatures of flow-sorted, sequenced immune cells[67] using SingleR[68]. Differential gene expression was performed by cell type and condition using FindMarkers in Seurat[69] (logfc.threshold = 0.1, Min.pct = 0.1, min.cells.group = 10). GSEA[70] was performed on lists of DEGs using fgsea[71] (minSize = 15) and plotted in ggplot2, with *P* values calculated by permutation using the fgsea package intrinsic functions. Gene-set expression scores were calculated as the difference between the average expression levels of each gene set and randomly sampled pool of all (control) genes for each cell, using scanpy.tl.score_genes: an implementation of Seurat's AddModuleScore function[72]. INFα response gene expression correlations were calculated using Pearson's correlation coefficient implemented through the scipy package, with two-sided *P* values reported. Correlations were calculated on named cell subtypes with named gene expression values >0. Cell proportion plots were performed using ggplot2, with calculation of *P* values and confidence intervals (CIs) using a permutation and bootstrapping approach from the scProportionTest package. Trajectory analysis was performed using Slingshot[50]. For B cell trajectories, memory B cells, GC B cells and PCs were subset from the main object and reclustered. B cell Slingshot lineages were inferred with B GC light zone (B_GC_LZ) assigned as the starting cluster. For MNP cell trajectories, monocytes, macrophages and cDCs were subset from the main object and reclustered. MNP cell Slingshot lineages were inferred with classic monocytes assigned as the starting cluster. CellPhoneDB analysis was performed as previously described[73], with additional plotting of results using ktplots. Single-cell TCR analysis was performed using Scirpy[74]. Single-cell BCR analysis was performed using sc-Dandelion[75] with V(D)J gene reannotation and parsing to AIRR format performed in the sc-Dandelion singularity container (sc-dendelion_latest.sif). Gini's index was calculated in sc-Dandelion, with *P* values calculated using Tukey's honestly significant difference test implemented through the scikit-posthocs package. For identification of putative BCR sequences with capacity to bind SARS-CoV-2 surface antigens, we downloaded CDR3 sequences and germline assignments of known SARS-CoV-2-binding antibodies from the Coronavirus Antibody Database (CoV-AbDab, updated 26 July 2022)[48] and filtered for antibodies of B cell origin. We defined putative SARS-CoV-2-binding BCR sequences in our dataset as B cells with complete CDR3-L match and significant motif similarity for CDR3-H compared with the CoV-AbDab database. Motif enrichment analysis between BCR CDR3-H sequences in our data and CoV-AbDab was performed using GLIPH2 (ref. 36). First, we compiled a reference database of naive BCR CDR3 sequences using published CDR3 sequence from COVID-naive B cells from multiple sources, for the GLIPH2 pipeline. Briefly, 1,437,743 CDR3-H sequences from bulk BCR sequencing (downloaded from iReceptor gateway 30 July 2022)[60] and 55,444 CDR3-H sequences from three single-cell GEX + V(D)J-seq studies[75–77] of control COVID-naive B cells were included in the reference. For GLIPH2 analysis, we considered only motifs that were more than six residues in length. We assessed motifs biased toward the target data (CoV-AbDab + BCR-seq V(D)J data combined) versus the naive reference data (Fisher's exact test $P < 0.05$), to identify sequences enriched in our cohort. Shared motifs were visualized using ggseqlogo[78]. We recombined CDR3-H and CDR3-L sequences using their single-cell barcodes and filtered for single cells where both CDR3-H and CDR3-L sequences matched the corresponding CoV-AbDab antibody.

Harmonization and label transfer with a publicly available, airway brushing dataset were performed with scANVI[79], using seed labels from the annotated NALT and PBMC sample cell-type labels[80–84].

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All single-cell gene expression and V(D)J sequencing data have been uploaded to the Gene Expression Omnibus, accession no. GSE287808 and will be freely available on publication. Source data are provided with this paper.

Code availability

Details of packages and versions can be found in the online methods key resources table. Code to reproduce the analysis in the manuscript has been deposited in the Clatworthy lab GitHub repository (https://github.com/clatworthylab/COVID_analysis/COVID_NALT) and will be freely available on publication.

References

61. Radtke, A. J. et al. IBEX: an iterative immunolabeling and chemical bleaching method for high-content imaging of diverse tissues. *Nat. Protoc.* **17**, 378–401 (2022).
62. Radtke, A. J. et al. IBEX: a versatile multiplex optical imaging approach for deep phenotyping and spatial analysis of cells in complex tissues. *Proc. Natl Acad. Sci. USA* **117**, 33455–33465 (2020).
63. Fleming, S. J. et al. *Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender* (Cold Spring Harbor Laboratory, 2019).
64. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst.* **8**, 281–291.e289 (2019).
65. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
66. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
67. Monaco, G. et al. RNA-seq signatures normalized by mRNA abundance allow absolute deconvolution of human immune cell types. *Cell Rep.* **26**, 1627–1640.e1627 (2019).
68. Aran, D. et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* **20**, 163–172 (2019).
69. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e1821 (2019).
70. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
71. Korotkevich, G. et al. *Fast gene set enrichment analysis* (Cold Spring Harbor Laboratory, 2016).
72. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).

73. Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat. Protoc.* **15**, 1484–1506 (2020).
74. Sturm, G. et al. Scirpy: a Scanpy extension for analyzing single-cell T-cell receptor-sequencing data. *Bioinformatics* **36**, 4817–4818 (2020).
75. Stephenson, E. et al. Single-cell multi-omics analysis of the immune response in COVID-19. *Nat. Med.* **27**, 904–916 (2021).
76. Consortium, C. A blood atlas of COVID-19 defines hallmarks of disease severity and specificity. *Cell* **185**, 916–938.e958 (2022).
77. Zhang, J.-Y. et al. Single-cell landscape of immunological responses in patients with COVID-19. *Nat. Immunol.* **21**, 1107–1118 (2020).
78. Wagih, O. ggseqlogo: a versatile R package for drawing sequence logos. *Bioinformatics* **33**, 3645–3647 (2017).
79. Xu, C. et al. Probabilistic harmonization and annotation of single-cell transcriptomics data with deep generative models. *Mol. Syst. Biol.* **17**, e9620 (2021).
80. Korotkevich, G. et al. Fast gene set enrichment analysis. Preprint at *bioRxiv* https://doi.org/10.1101/060012 (2021).
81. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).
82. Waskom, M. L. Seaborn: statistical data visualization. *J. Open Source Softw.* **6**, 3021 (2021).
83. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e3529 (2021).
84. Amezquita, R. A. et al. Orchestrating single-cell analysis with Bioconductor. *Nat. Methods* **17**, 137–145 (2020).

Acknowledgements

M.L.C. was funded by a Wellcome Clinical PhD Fellowship, Z.K.T. and M.R.C. by a Medical Research Council Research Project grant (no. MR/S035842/1), C.Y.C.L. by a Gates Scholarship and B.J.S. by a National Institute for Health and Care Research (NIHR) Clinical Lecturership. L.D., S.A.T. and M.R.C. acknowledge funding from the Wellcome Human Cell Atlas Strategic Science Support (grant no. WT211276/Z/18/Z). M.R.C. and R.D.M.B. were supported by an NIHR Research Professorship (no. RP-2017-08-ST2-002), M.R.C. and E.G. by a Wellcome Trust Investigator Award (no. 220268/Z/20/Z), J.R.F. and M.R.C. by the NIHR Blood and Transplant Research Unit in Organ Donation (grant no. NIHR203332), a partnership between NHS Blood and Transplant, University of Cambridge and Newcastle University, and N.R. and M.R.C. by the NIHR Cambridge Biomedical Research Centre (grant no. NIHR203312). M.R.C. is also supported by a Wellcome Discovery Award (no. 227890/Z/23/Z). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Author contributions

M.L.C. and M.R.C. designed the study. M.L.C., N.R., Z.K.T., G.S.B., C.Y.C.L., J.R.F., E.S.G., M.M.C., L.D. and R.D.M.B. performed experiments and/or analyzed data, S.A.T. and B.J.S. advised on data analysis, D.R.J. and M.R.C. supervised the study. M.L.C. and M.R.C. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

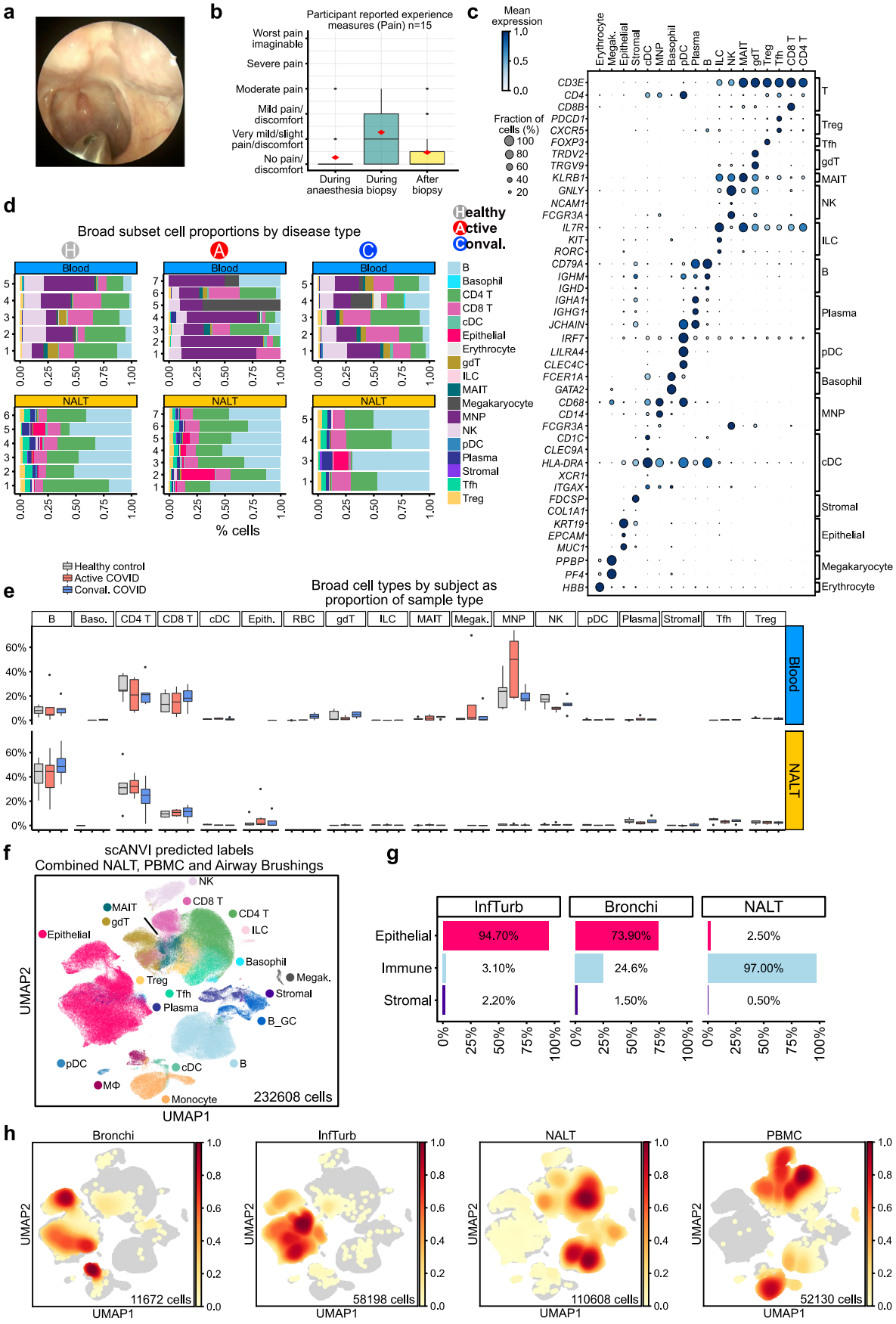
Extended data is available for this paper at https://doi.org/10.1038/s41590-024-02064-9.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41590-024-02064-9.

Correspondence and requests for materials should be addressed to Menna R. Clatworthy.

Peer review information *Nature Immunology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editor: L. A. Dempsey, in collaboration with the *Nature Immunology* team.

https://doi.org/10.1038/s41590-024-02064-9

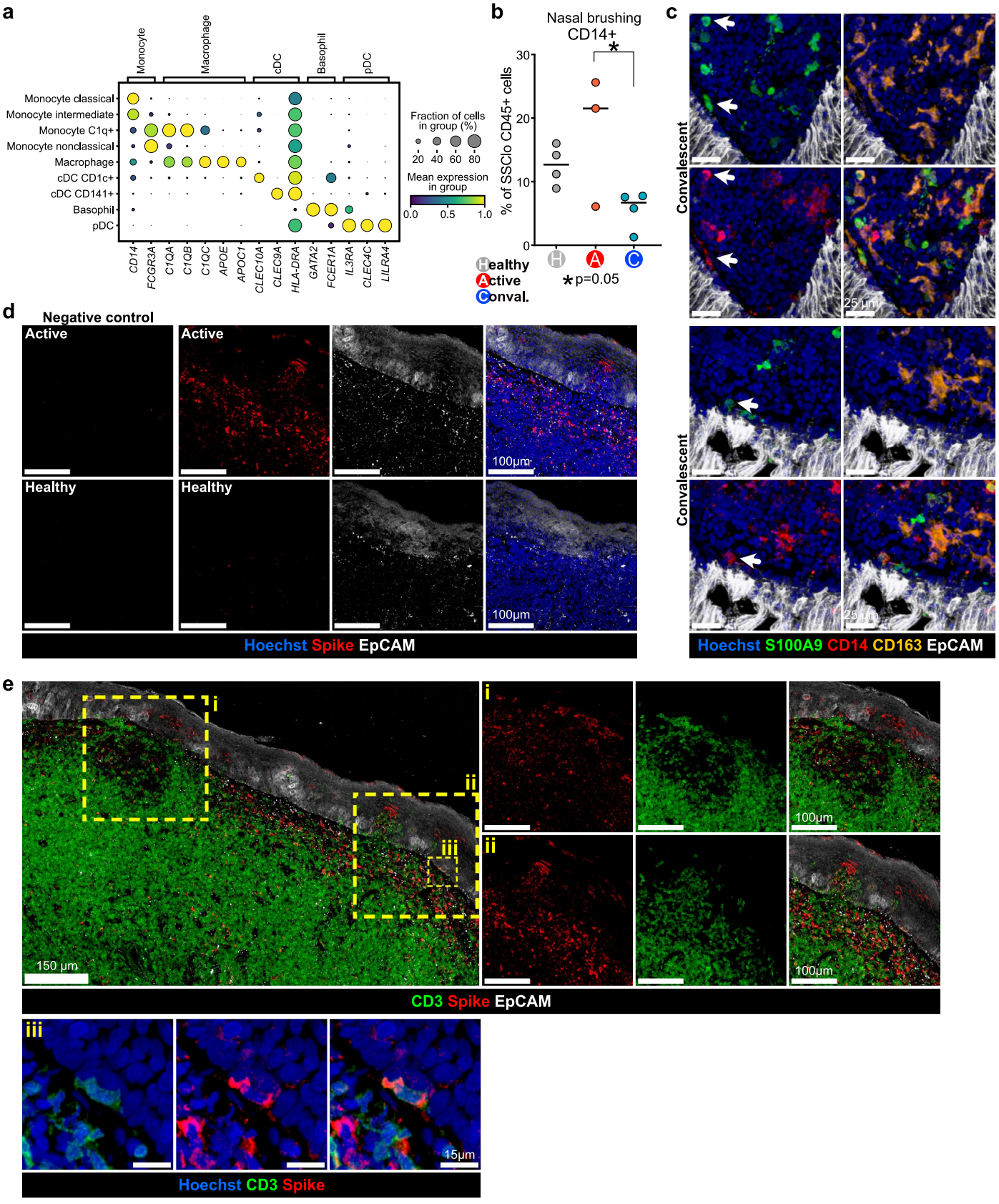


Extended Data Fig. 1 | See next page for caption.

https://doi.org/10.1038/s41590-024-02064-9

Extended Data Fig. 1 | Experimental overview and cellular landscape of blood and NALT in SARS-CoV-2 infection. a, Endoscopic nasal view during an office-based postnasal space biopsy procedure in an awake human subject under topical local anaesthesia. Tilley-Henkel forceps are seen towards the bottom of the image. **b**, Boxplot of subject reported experience data during stages of local anaesthetic NALT biopsy procedure in 15 subjects. Median score represented by line, with red diamond showing mean score. **c**, Expression of canonical markers by assigned cell type label. Horizontal axis shows assigned broad cell labels, with genes grouped by canonical expression groups on vertical axis. Dot size shows percentage of gene expression by cells in group, with colour indicating normalised mean gene expression by cells in group. **d**, Stacked barplots showing cell type proportions by subject, condition and sample type. **e**, Serial boxplot showing cell type proportions by disease state, plotted by subject and sample. **f**, UMAP of scANVI integration of data and cell type labels from this paper with adult inferior turbinate and bronchial brushing data in SARS-CoV-2 subjects and controls previously published[2]. **g**, Proportional representation of airway sample data from e presented by sampling location and predicted cell type following scANVI integration. Cell type percentages are of total cells by sample location. **h**, Scanpy embedding density UMAP showing showing density of cells by sampling location on UMAP projection in **e**. Cell numbers by sample location are printed in the bottom right of the projection.

https://doi.org/10.1038/s41590-024-02064-9

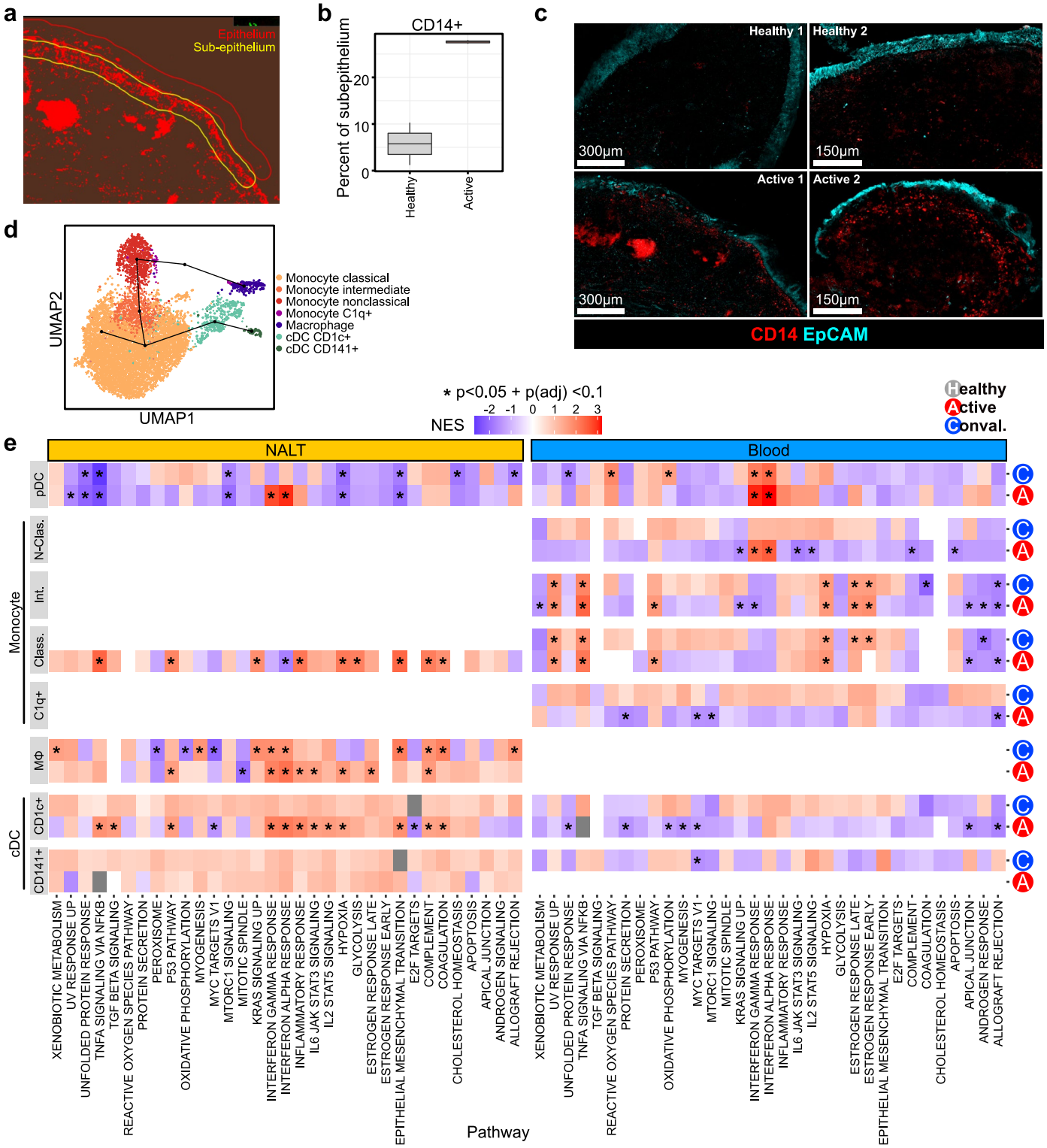


Extended Data Fig. 2 | See next page for caption.

https://doi.org/10.1038/s41590-024-02064-9

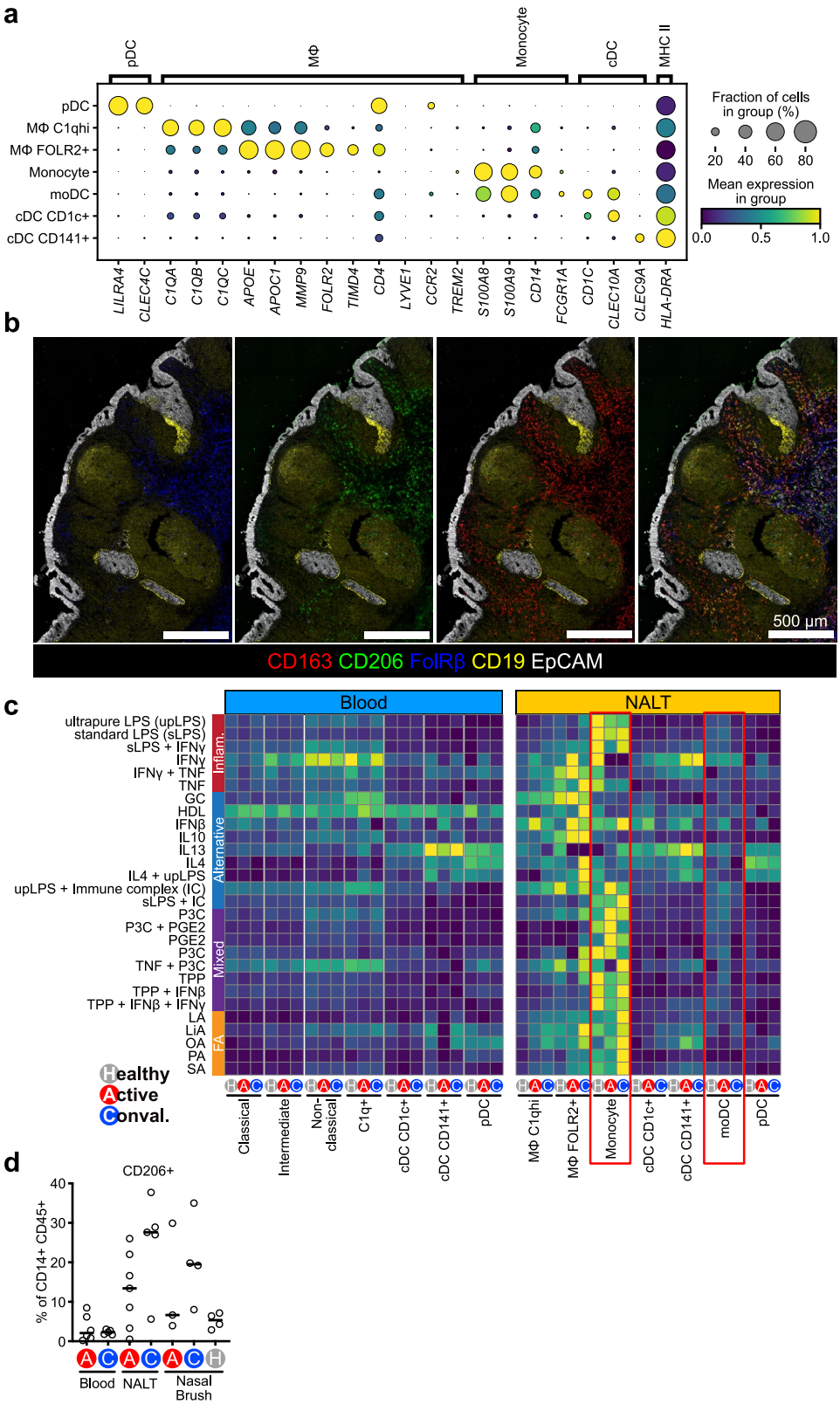
Extended Data Fig. 2 | MNPs contribute to lymphoid tissue defence and repair. a, Expression of canonical markers by assigned MNP cell label for MNP cell subset. Vertical axis shows assigned broad cell labels, with genes grouped by canonical expression groups on horizontal axis. Size of point indicates fraction of cells in each group expressing corresponding gene, colour of point indicates scaled mean expression of corresponding gene in each named cell type group. **b**, Flow cytometry on inferior turbinate nasal brushing showing % of CD14+ cells in SSClo CD45+ live single cells. Each point represents a subject. Horizontal lines indicate median value. * p = 0.05 (Tukey HSD test). **c**, Confocal imaging of section of NALT from a convalescent COVID-19 subject. White arrowheads highlight co-expression of S100A9 and CD14. Related to Fig. 2e. n = 1. **d**, Representative confocal imaging of section of NALT from a healthy control and an active COVID-19 subject. n = 1 each. **e**, Confocal imaging of section of NALT from an active COVID-19 subject with indicated regions shown at higher magnification. n = 1.

https://doi.org/10.1038/s41590-024-02064-9



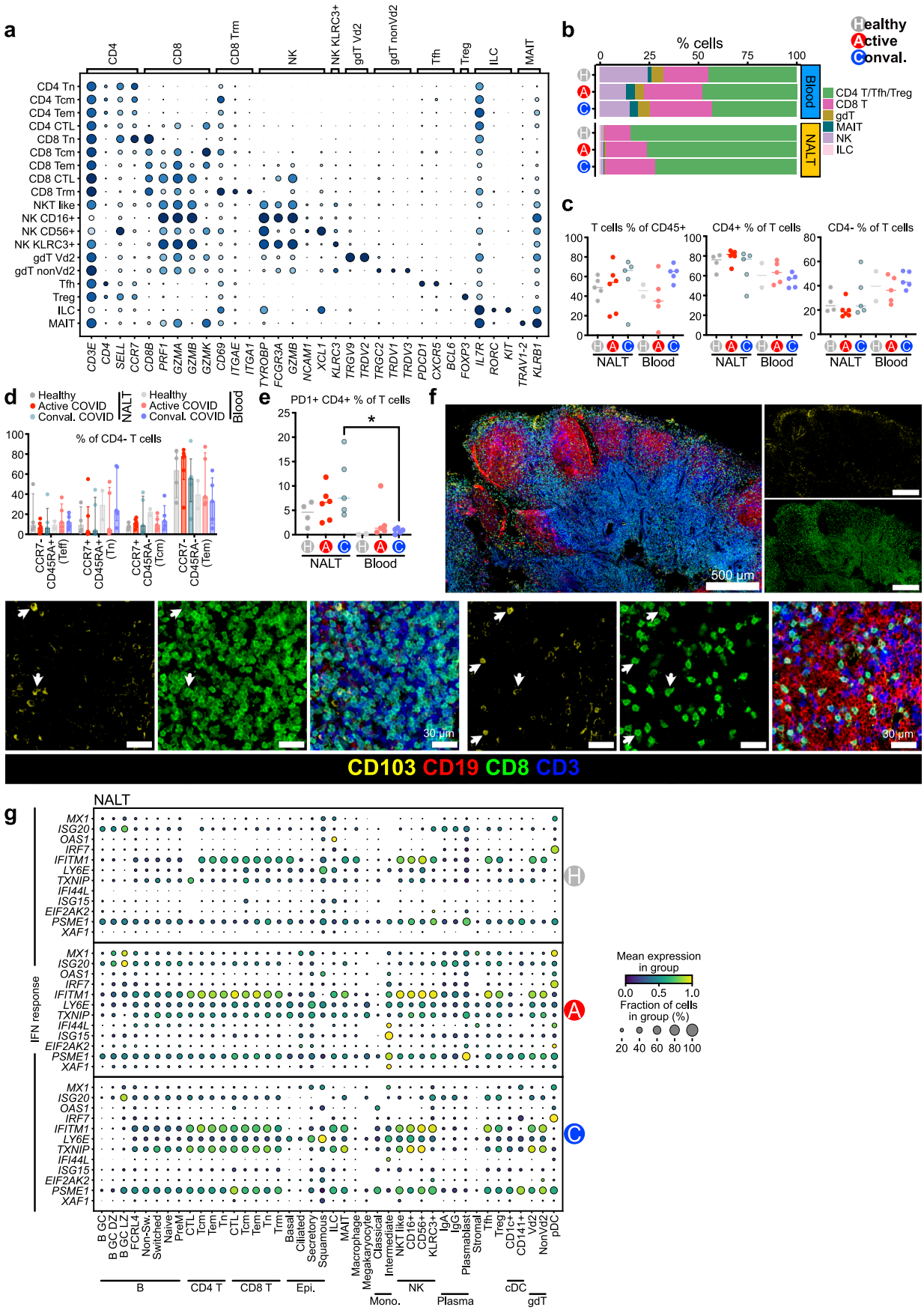
Extended Data Fig. 3 | Monocytes localize at the subepithelium during acute SARS-COV-2 infection. a, Example of subepithelial marking strategy used in **b-c**. A 25 pt intensity thresholder has been applied to CD14 staining. Epithelium marked in dotted red line using EpCAM staining with fixed subepithelial area width marked in yellow dotted line. **b**, Barplot showing comparison of CD14+ staining in samples from **c** as a percentage of subepithelial region. **c**, Confocal imaging of sections of NALT from two separate healthy controls and active COVID-19 subjects. **d**, Slingshot pseudotime trajectories for MNP subset cell types, plotted in UMAP space. Black lines indicate Slingshot lineage trajectories, starting from Classical monocyte cluster. **e**, Heatmap showing all significantly enriched Hallmark pathways in MNP cell types following Gene set enrichment analysis of differentially expressed genes between COVID-19 disease groups and healthy control subjects. Only pathways containing at least one significant enrichment ($p < 0.05$ and Benjamini-Hochberg adjusted p value < 0.1) in MNP cell types are shown. Colour indicates normalised enrichment score (NES), with red indicating greater pathway enrichment in disease group compared with healthy controls, and blue indicating increased pathway enrichment in healthy controls than disease group.

https://doi.org/10.1038/s41590-024-02064-9



Extended Data Fig. 4 | Tissue resident macrophages take on pro-repair phenotype in convalescence. a, Expression of canonical markers by assigned nasal MNP subset cell label. Horizontal axis shows assigned broad cell labels, with genes grouped by canonical expression groups on vertical axis. Size of point indicates fraction of cells in each group expressing corresponding gene, colour of point indicates scaled mean expression of corresponding gene in each named cell type group. pDC, plasmacytoid Dendritic Cell; Mφ, macrophage, moDC, monocyte-derived DC; cDC, conventional DC. **b**, Representative confocal imaging of section of NALT from a convalescent COVID-19 subject. Related to Fig. 3f. n = 1. **c**, Heatmap showing expression of genes associated with macrophage polarisation in all MNP subset cell types, split by disease type and sample type, calculated as scaled geneset expression scores (AddModuleScore[3]) using reference experimentally derived macrophage stimulation gene-sets[4]. Stimulation agent used for each gene-set shown on y-axis label, in correspondence to the original publication. Inflam., Inflammatory; FA, Fatty Acids. **d**, Column scatter graph showing percentage of CD206+ cells in live CD45 + CD14+ cells as determined by flow cytometry in blood, NALT and inferior turbinate nasal brushing samples. *, p(adj)<0.05 (Tukey HSD).

https://doi.org/10.1038/s41590-024-02064-9

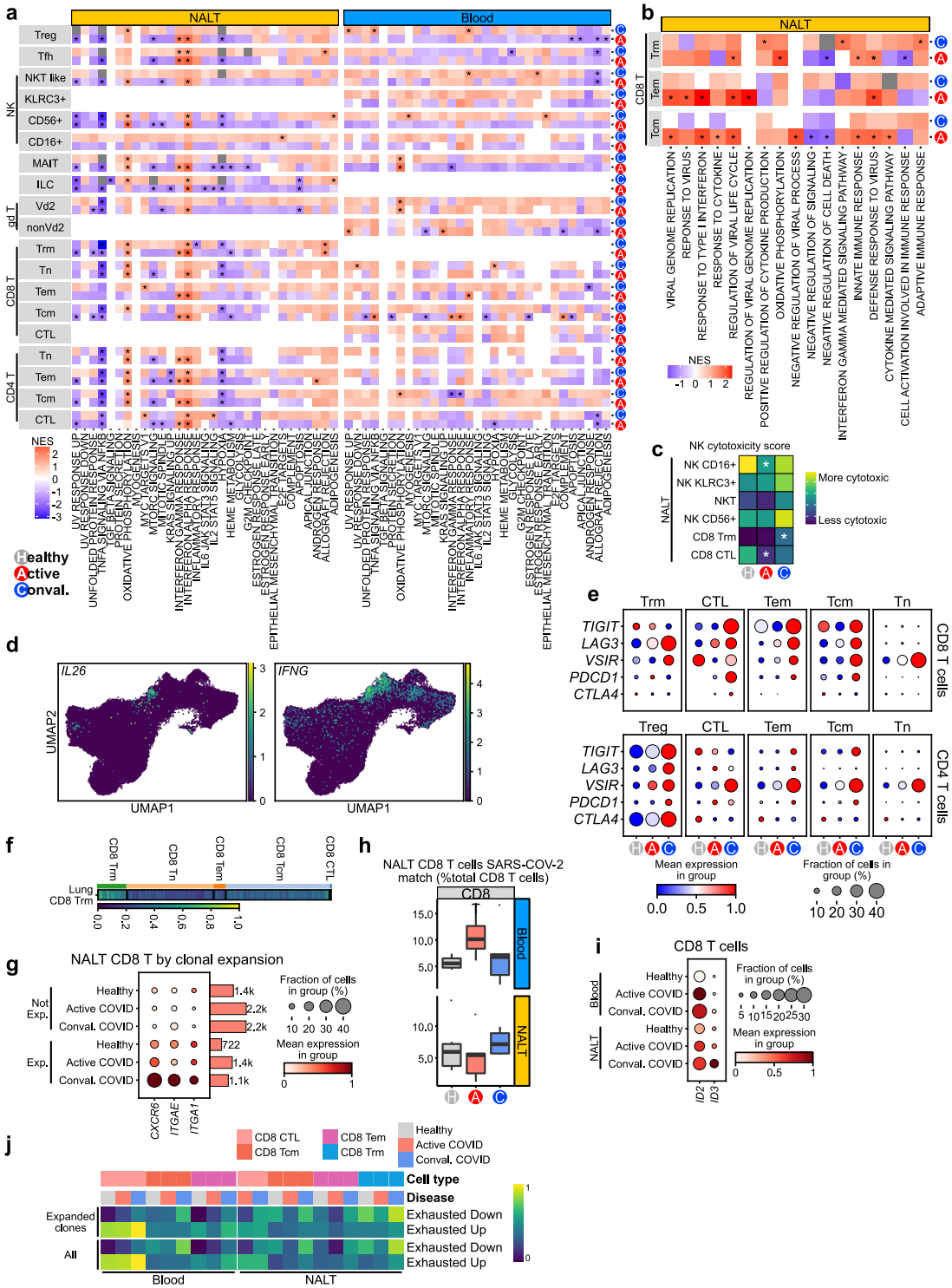


Extended Data Fig. 5 | See next page for caption.

https://doi.org/10.1038/s41590-024-02064-9

Extended Data Fig. 5 | Temporal characterization of T cells and innate-like lymphocytes. a, Expression of canonical markers by assigned nasal T/innate lymphocyte subset cell label. Vertical axis shows assigned broad cell labels, with genes grouped by canonical expression groups on horizontal axis. Size of point indicates fraction of cells in each group expressing corresponding gene, colour of point indicates scaled mean expression of corresponding gene in each named cell type group. **b**, Stacked barcharts showing combined proportional representation of CD4 T cells, CD8 T cells and innate lymphocytes, grouped into broad cell type categories and split by sample type and disease status. **c**, Grouped scatter plot of T cell proportions, as determined by flow cytometry, split by disease type and sample type. Left, percentage of T cells in live CD45+ cells; center and right, percentage of CD4+ (centre) and CD4- (right) T cells. Each point represents an individual subject sample. Lines represent group median value. **d**, Grouped scatter and bar plot showing percentage T cell subsets based on CCR7 and CD45RA expression assessed by flow cytometry. Each point represents an individual subject sample. Bars represent median values, with lines showing interquartile range. Teff, T effector; Tn, T naïve; Tcm, T central memory; Tem, T effector memory. **e**, Grouped scatter plot showing percentage of CD4 + PD1 + T cells by tissue and condition. Each point represents an individual subject sample. Lines represent group median value. *, adjusted $p < 0.05$ (Šidák multiple comparisons test). **f**, Representative confocal imaging of section of NALT from a convalescent COVID-19 subject with higher magnification shown below. Related to Fig. 4f, g. n = 1. **g**, Scaled expression of selected leading edge genes in NALT cell types that were enriched in Hallmark Interferon Alpha Response and Hallmark Interferon Gamma Response (IFN Response) gene sets following gene set enrichment analysis of differentially expressed genes in COVID-19 subjects compared with healthy controls. Size of point indicates fraction of cells in each group expressing corresponding gene, colour of point indicates scaled mean expression of corresponding gene in each named cell type group. Epi, Epithelial; Mono., Monocyte.

https://doi.org/10.1038/s41590-024-02064-9

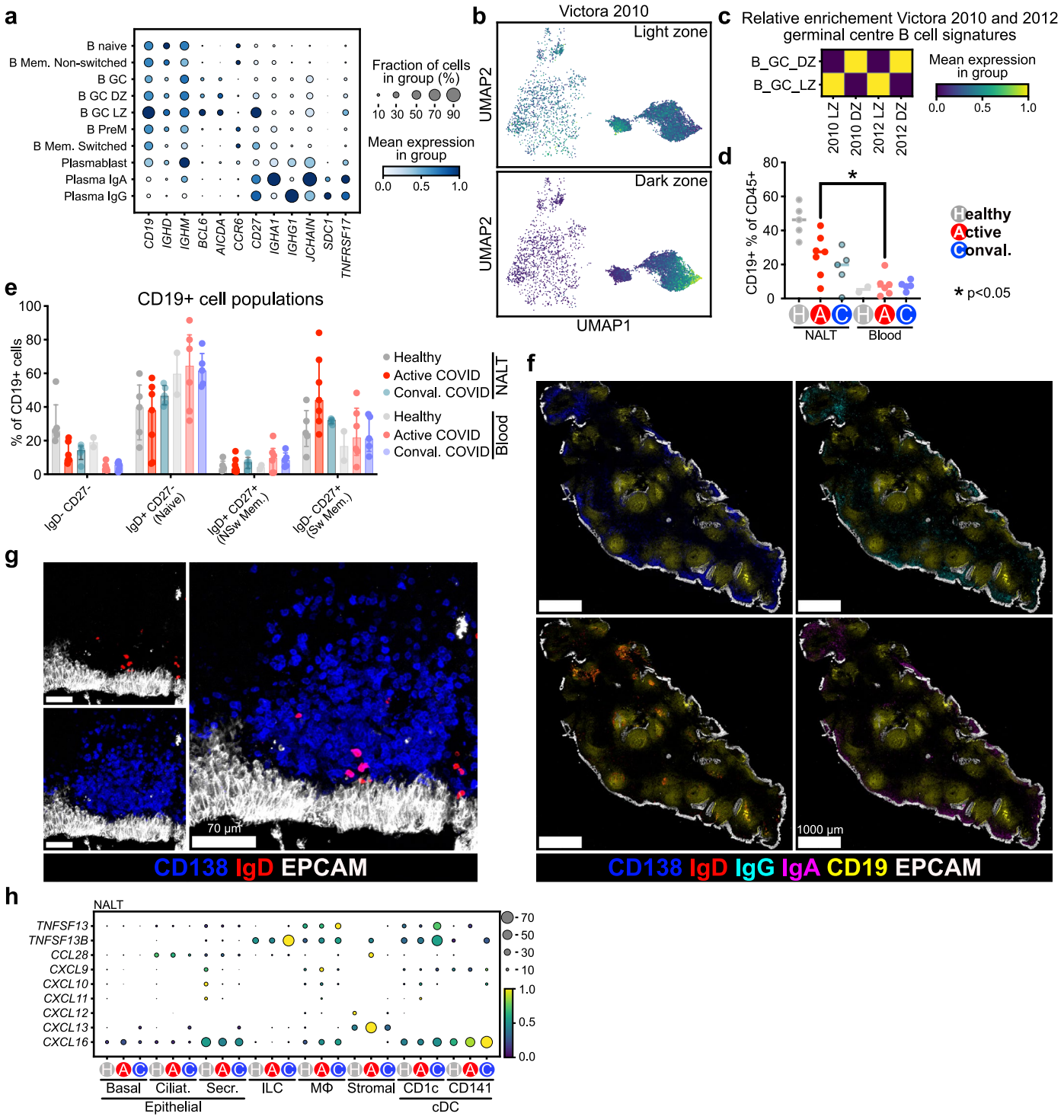


Extended Data Fig. 6 | See next page for caption.

https://doi.org/10.1038/s41590-024-02064-9

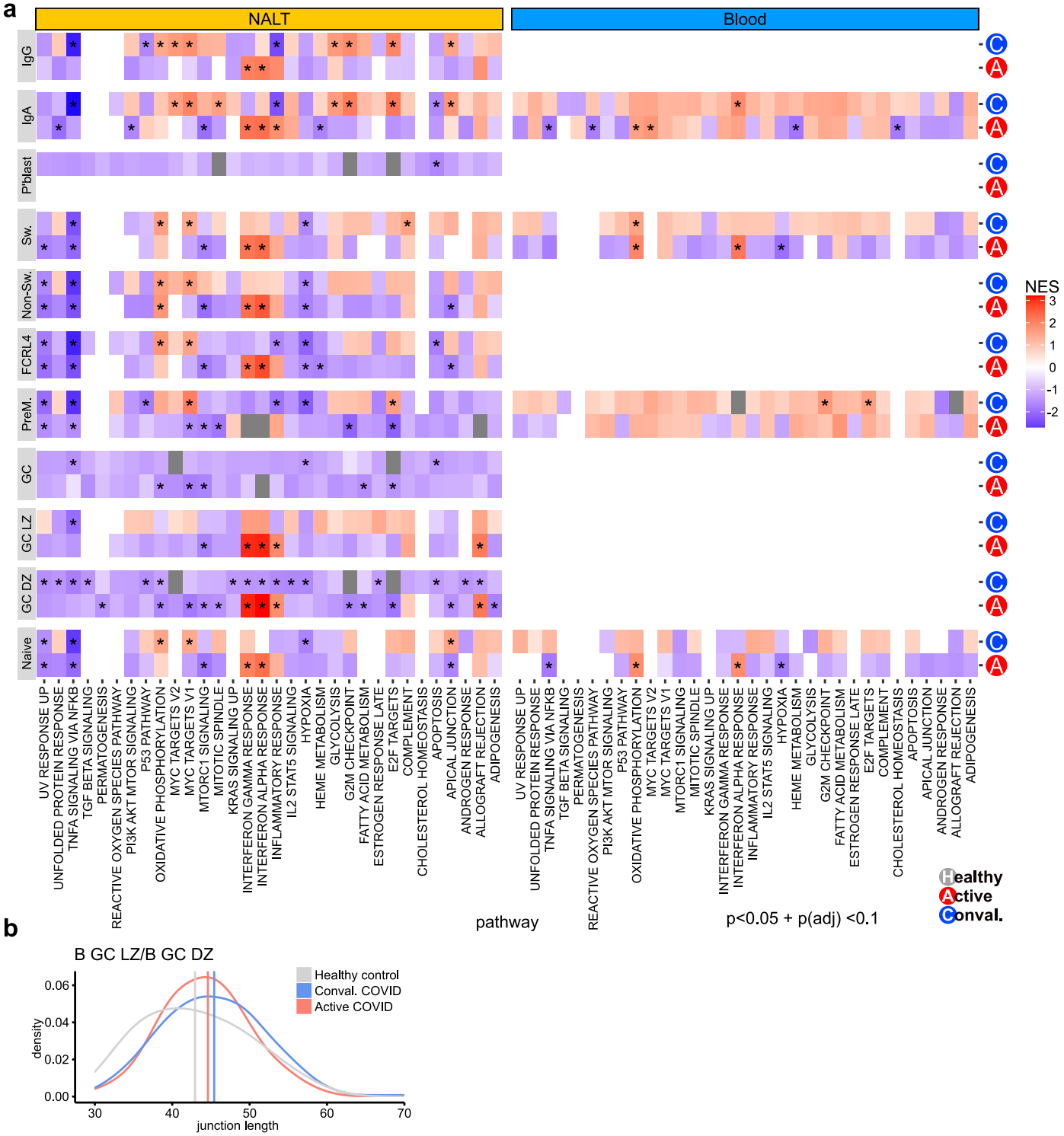
Extended Data Fig. 6 | Clonally expanded NALT CD8 Trm cells show prolonged activation. a, Heatmap showing all all significantly enriched Hallmark pathways in T/innate lymphocyte cell types following gene set enrichment analysis of differentially expressed genes between COVID-19 disease groups and healthy control subjects. Only pathways containing at least one significant enrichment ($p < 0.05$ and Benjamini-Hochberg adjusted p value < 0.1) in T cell types are shown. Colour indicates normalised enrichment score (NES), with red indicating greater pathway enrichment in disease group compared with healthy controls, and blue indicating increased pathway enrichment in healthy controls than disease group. Results are split by sample type (NALT/Blood). **b**, Selected significantly enriched Gene Ontology (GO) terms in NALT CD8 T cell memory subsets, following Geneset Enrichment Analysis of differentially expressed genes in COVID-19 disease states compared with healthy controls. NES, Normalised Enrichment Score, *, p value < 0.05 + Benjamini-Hochberg adjusted p value < 0.1. **c**, Heatmap showing Gene-set scoring (AddModuleScore) of KEGG Natural killer cell mediated cytotoxicity gene set in NALT CD8 T and NK cell subsets. *, p value < 0.05. **d**, UMAP showing expression of *IL26* and *IFNG* within subset re-clustered T/innate lymphocyte UMAP. Scale bars indicate level of gene expression. **e**, Dotplot showing expression of canonical T cell exhaustion marker genes in CD4 and CD8 T cell subsets. Size of point indicates fraction of cells in each group expressing corresponding gene, colour of point indicates scaled mean expression of corresponding gene in each named cell type group. **f**, Heatmap of single cell transcriptomic data showing enrichment of human lung CD8 Trm signature[6] in CD8 T cell subsets by assigned label. **g**, Single cell gene expression dotplot of NALT CD8 T cells showing expression of canonical CD8 Trm marker genes by disease type and clonal expansion status, with expanded clones of size >=2. Bars to right of plot indicate absolute number of cells in each group. **h**, Boxplot showing GLIPH2 identified SARS-CoV-2 specific CD8 T cells, as a proportion of total CD8 T cell population by subject and sample type and plotted by disease status. SARS-COV-2 cells were identified through TCRb similarity matching with MHC-I identified SARS-COV-2 reactive T cell sequences in the ImmunoCODE[7] and VDJdb[8] databases. **i**, Single cell gene expression dotplot of all CD8 T cells showing selected short- and long-term memory associated genes by disease type and sample type location. **j**, Heatmap showing enrichment of Human HIV exhaustion signatures[5] split by cell type, sample type, disease type and clonal expansion (TCR clone size ≥2) status. Exhausted Up indicates genes upregulated in exhausted vs non-exhausted CD8 T cells and Exhausted Down indicated genes downregulated in exhausted vs non-exhausted CD8 T cells. Top two rows indicate cell type and disease status for each column.

https://doi.org/10.1038/s41590-024-02064-9



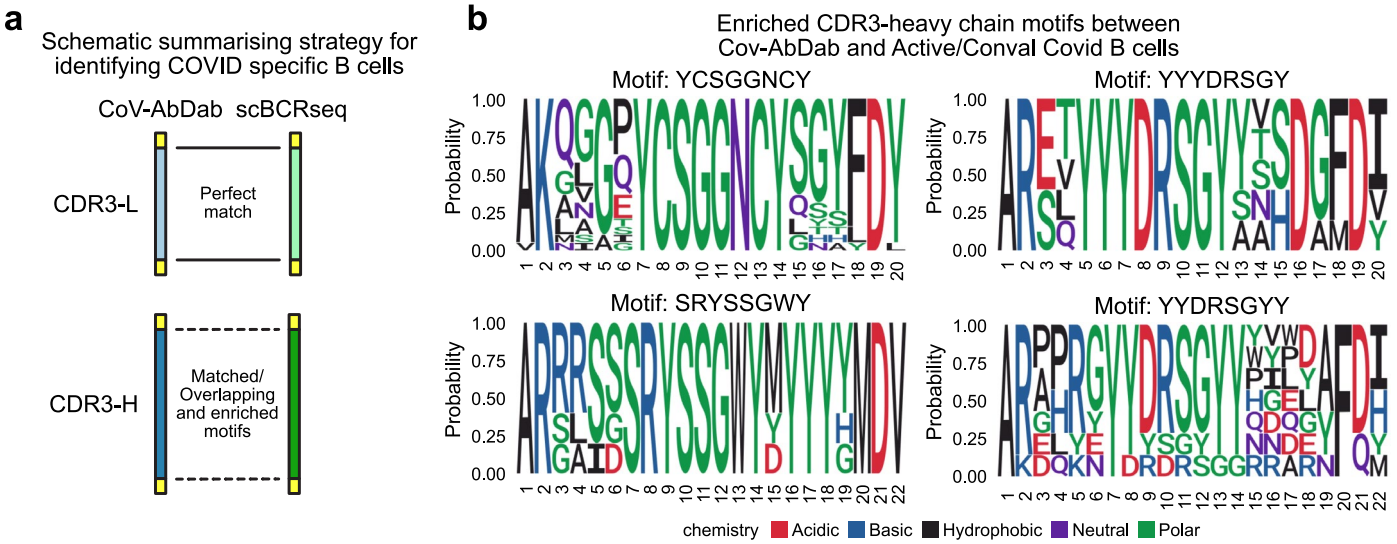
Extended Data Fig. 7 | Local anti-viral germinal centre response in NALT. **a**, Expression of canonical markers by assigned B lymphocyte subset cell label. Vertical axis shows assigned broad cell labels, with genes grouped by canonical expression groups on horizontal axis. Size of point indicates fraction of cells in each group expressing corresponding gene, colour of point indicates scaled mean expression of corresponding gene in each named cell type group. **b**, UMAP showing gene set enrichment score (AddModuleScore) of experimentally derived Germinal centre B cell signatures[9] in germinal centre B cell clusters. **c**, Single cell gene expression matrixplot of germinal centre B cell light zone (B_GC_LZ) and dark zone (B_GC_DZ) labelled cells, showing relative enrichment of experimentally derived mouse (2010[9]) and human (2012[10]) light zone (LZ) and dark zone (DZ) germinal cell B cell specific cell signatures. **d**, Grouped scatter plot showing percentage of CD19+ cells in live CD45+ cells, as determined by flow cytometry. Each point represents an individual subject sample. Lines show group median value. Data is divided by disease and sample type. *, $p < 0.05$ (Wilcoxon test). **e**, Grouped scatter and bar plot of CD19+ cell subpopulation proportions by IgD and CD27 expression, as determined by flow cytometry. Each point represents an individual subject sample. Bars represent median values, with lines showing interquartile range. NSw Mem, Non-switched memory; Sw. Mem, Switched memory. **f-g**, Representative confocal imaging of section of NALT from a convalescent COVID-19 subject (**f**) with higher magnification for three of the markers (**g**). n = 1. **h**, Dotplot showing selected chemokine expression across NALT Epithelial cells, ILCs, Macrophages (MΦ), Stromal cells and conventional dendritic cells. Size of point indicates fraction of cells in each group expressing corresponding gene, colour of point indicates scaled mean expression of corresponding gene in each named cell type group.

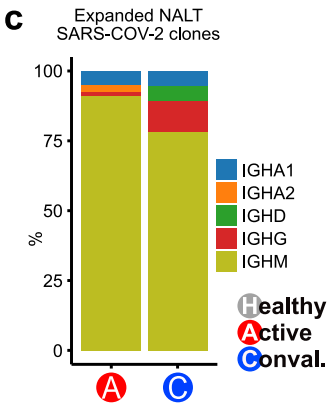
https://doi.org/10.1038/s41590-024-02064-9

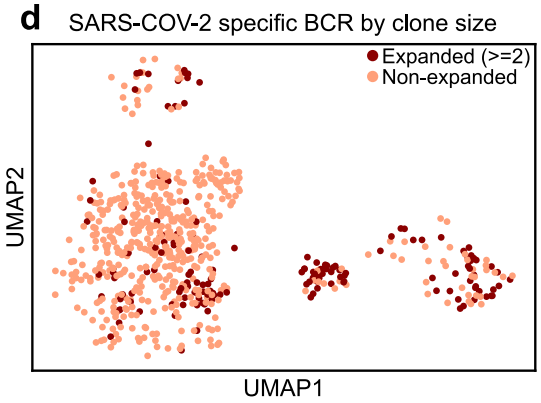


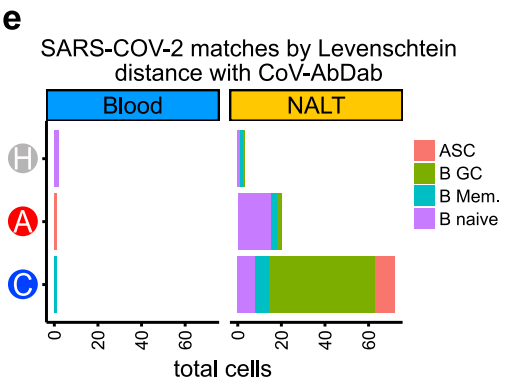
Extended Data Fig. 8 | V(D)J rearrangement and selection in NALT. a, Heatmap showing all significantly enriched Hallmark pathways in B lymphocyte/plasma cell types following gene set enrichment analysis of differentially expressed genes between COVID-19 disease groups and healthy control subjects. Only pathways containing at least one significant enrichment ($p < 0.05$ and Benjamini-Hochberg adjusted p value < 0.1) in B cell types are shown. Colour indicates normalised enrichment score (NES), with red indicating greater pathway enrichment in disease group compared with healthy controls, and blue indicating increased pathway enrichment in healthy controls than disease group. Results are split by sample type (NALT/Blood). **b**, Density plot of BCR junction length in light zone and dark zone germinal centre B cells, split by condition. Vertical lines indicate mean junction length for each disease group.

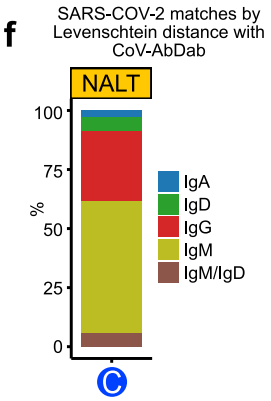
https://doi.org/10.1038/s41590-024-02064-9





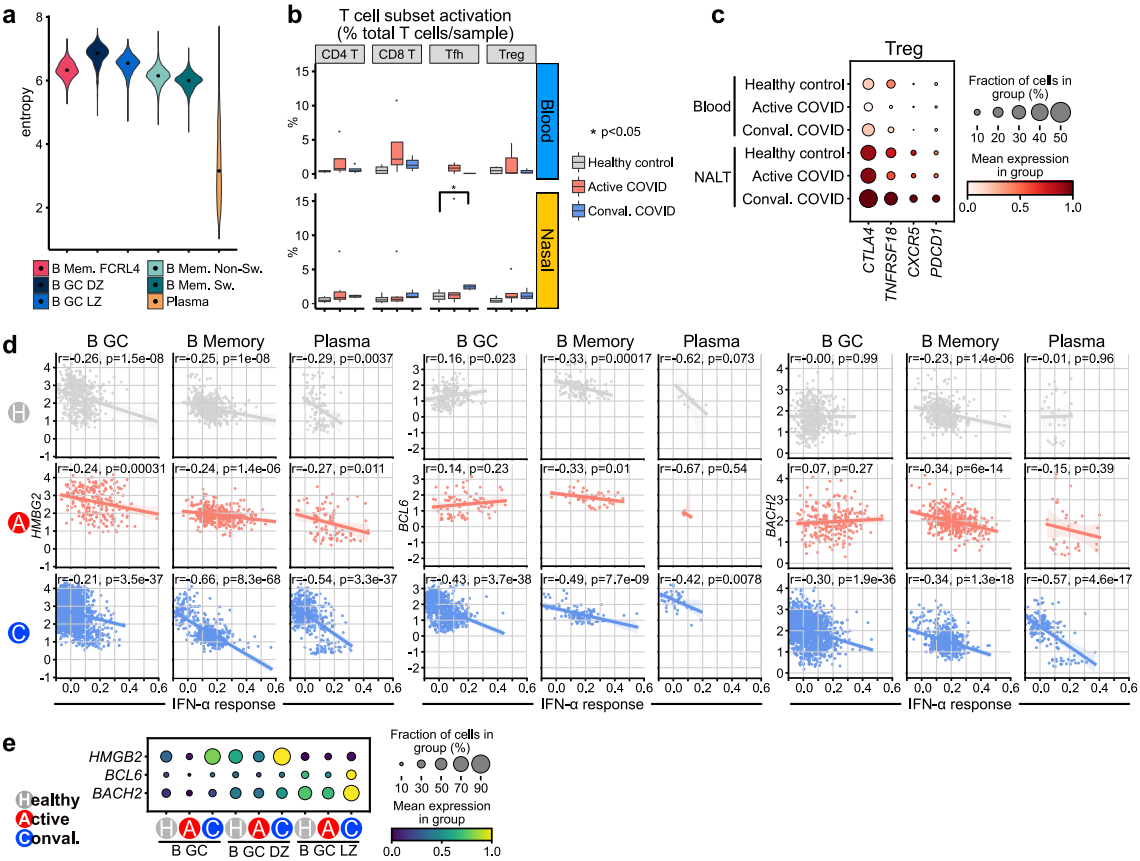






Extended Data Fig. 9 | Local production of anti-SARS-COV-2 antibodies in NALT following infection. a, Schematic summarising the strategy for identifying SARS-COV-2 specific B cells from the single cell CDR3-L and CDR3-H sequences shared with CoV-AbDab. **b**, Enriched CD3-heavy chain motifs between Cov-AbDab and B cells from Active/Conval COVID-19 subjects. **c**, Stacked bar plot showing proportional split by isotype of expanded (clone size >=2) SARS-COV-2 specific B cell clones by disease state. **d**, UMAP showing SARS-COV-2 specific B cell clones plotted using co-ordinates from B/Plasma UMAP seen in Fig. 6a. Cells forming part of expanded BCR clones (clone size >=2) are shown in dark red, with clonally unexpanded SARS-COV-2 specific B cells shown in pink. **e**, Stacked barplot showing absolute numbers of SARS-COV-2 specific B cells identified by both Gliph2 and Levenschtein distance heavy chain matching with CDR3 sequences from Cov-AbDab11 presented by sample type, disease type and assigned cell type label. **f**, Stacked barplot showing proportional representation of isotypes in SARS-COV-2 specific B cells from Convalescent COVID-19 subecjts, identified by both Gliph2 and Levenschtein distance heavy chain matching with CDR3 sequences from Cov-AbDab[11].

https://doi.org/10.1038/s41590-024-02064-9



Extended Data Fig. 10 | Type I interferon may inhibit the progression of the GC reaction during infection. a, Violin plot of per-cell transcriptomic entropy for B cells by subset cell type label. **b**, Boxplot of enrichment of a T cell activation signature induced by controlled SARS-COV-2 infection[12] in T cells from NALT and peripheral blood by disease status. Two sided significance is indicated by *. **c**, Single cell gene expression dotplot showing expression of selected inhibitory and lymphoid homing receptors in T regulatory cells by sample type and disease status. **d**, Extended figures from Fig. 6k, showing plots split by cell type and disease type for interferon alpha response scores plotted against expression of HMGB2 (upper left), BCL6 (upper right), and BACH2 (lower left). Linear regression line is shown with a shaded area representing the 95% confidence interval. r, Pearson correlation statistic; p, p-value. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ns, $p > 0.05$. **e**, Dotplot showing expression of genes shown in Fig. 8d in NALT germinal centre B cells, split by cell type and disease type. Size of point indicates fraction of cells in each group expressing corresponding gene, colour of point indicates scaled mean expression of corresponding gene in each cell type and disease group.

nature portfolio

Corresponding author(s): Menna Clatworthy

Last updated by author(s): Oct 12, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection.
Data analysis	Details of all packages and software provided in manuscript. All notebooks will be uploaded to Clatworthy lab github repository and made freely available, with a link provided in the code availability statement.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Data upload to GEO in progress and will be made available to reviewers once this is complete. GEO data will be made public and freely available on publication

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	This data is contained in Supplementary Table 1
Population characteristics	This data is contained in Supplementary Table 1
Recruitment	COVID-19 subjects were recruited from a single site - Addenbrooke's hospital, Cambridge, UK, between May 2020 and July 2021. Healthy control subjects were recruited between August 2019 and November 2019 from Cambridge, UK. Inclusion and exclusion criteria are listed in Methods:Human Subject Recruitment
Ethics oversight	Ethical approval was given through Cambridge Central Research Ethics Committee (REC reference 08/H0308/267, IRAS project ID: 194217), administered through Cambridge University Hospitals NHS Foundation Trust. All subjects provided informed written consent.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculation was performed. Small sample sizes are typical of single cell RNA sequencing data, given the large number of individual data points taken for each subject. We have ensured a sample size of >=5 subjects/group, which is typical for contemporary studies of this nature.
Data exclusions	Single cell RNA sequencing data was filtered and doublets excluded as described in the manuscript. Filtered 'cells' and doublets were not included in the final analysis.
Replication	Single cell RNA sequencing findings have been validated by microscopy and/or flow cytometry as stated in the text. As microscopy was performed primarily as a validation tool for single cell RNA sequencing data findings, microscopy replicates were not included in the manuscript unless stated. We are happy to undertake these, on a patient or technical level replicate basis, on request.
Randomization	No randomisation
Blinding	No blinding

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).*
Research sample	*State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.*
Sampling strategy	*Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.*
Data collection	*Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.*

Timing	*Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Non-participation	*State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.*
Randomization	*If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.*

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.*
Research sample	*Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.*
Sampling strategy	*Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.*
Data collection	*Describe the data collection procedure, including who recorded the data and how.*
Timing and spatial scale	*Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Reproducibility	*Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.*
Randomization	*Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.*
Blinding	*Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.*

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	*Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).*
Location	*State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).*
Access & import/export	*Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).*
Disturbance	*Describe any disturbance caused by the study and how it was minimized.*

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	All antibodies, sources and research resource identifiers listed in supplementary table 2
Validation	Negative/Isotype controls were undertaken for SARS-CoV-2 spike protein staining and are listed in supplementary table 2 and shown in extended data 2d.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	After Single cell RNA-seq GEM creation, remaining single cell suspensions were diluted to a concentration of 2x106 cells in 100µL 1x PBS. If this concentration was not achievable, all remaining cells were used. Cells were blocked with 5µL FcR blocking reagent (Miltenyi-Biotech) before undergoing surface staining with fluorophore labelled antibodies as outlined in methods table. Samples were washed and resuspended in 1x PBS.
Instrument	Fortessa II (BD)
Software	Flowjo v10
Cell population abundance	Cells were not sorted before flow cytometry was undertaken
Gating strategy	All cells were gated on a starting population of CD45+ Live singlets. Individual population gating strategies are described in figures.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.