Prolonged activation of nasal immune cell populations and development of tissue-resident SARS-CoV-2-specific CD8+ T cell responses following COVID-19

Anna H. E. Roukens[1], Cilia R. Pothast[2,16], Marion König[3,16], Wesley Huisman[3,16], Tim Dalebout[4], Tamar Tak[3], Shohreh Azimi[3], Yvonne Kruize[3], Renate S. Hagedoorn[2], Mihaela Zlei[5], Frank J. T. Staal[5], Fenna J. de Bie[5], Jacques J. M. van Dongen[5], Sesmu M. Arbous[6], Jaimie L. H. Zhang[1], Maaike Verheij[1], Corine Prins[1], Anne M. van der Does[7], Pieter S. Hiemstra[7], Jutte J. C. de Vries[4], Jacqueline J. Janse[3], Meta Roestenberg[1,3], Sebenzile K. Myeni[4], Marjolein Kikkert[4], Maria Yazdanbakhsh[3], Mirjam H. M. Heemskerk[2], Hermelijn H. Smits[3,17], Simon P. Jochems[3,17 ✉], in collaboration with BEAT-COVID group* and in collaboration with COVID-19 LUMC group

Systemic immune cell dynamics during coronavirus disease 2019 (COVID-19) are extensively documented, but these are less well studied in the (upper) respiratory tract, where severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) replicates[1-6]. Here, we characterized nasal and systemic immune cells in individuals with COVID-19 who were hospitalized or convalescent and compared the immune cells to those seen in healthy donors. We observed increased nasal granulocytes, monocytes, CD11c+ natural killer (NK) cells and CD4+ T effector cells during acute COVID-19. The mucosal proinflammatory populations positively associated with peripheral blood human leukocyte antigen (HLA)-DR^low monocytes, CD38+PD1+CD4+ T effector (T_eff) cells and plasmablasts. However, there was no general lymphopenia in nasal mucosa, unlike in peripheral blood. Moreover, nasal neutrophils negatively associated with oxygen saturation levels in blood. Following convalescence, nasal immune cells mostly normalized, except for CD127+ granulocytes and CD38+CD8+ tissue-resident memory T cells (T_RM). SARS-CoV-2-specific CD8+ T cells persisted at least 2 months after viral clearance in the nasal mucosa, indicating that COVID-19 has both transient and long-term effects on upper respiratory tract immune responses.

Although SARS-CoV-2 mainly replicates in the respiratory tract, and lower respiratory tract complications are major drivers of morbidity and mortality, it is unclear to what extent immunological dynamics observed in blood can be translated to the respiratory tract. Nasopharyngeal and peripheral blood cytokines and antibodies do not correlate during COVID-19 (ref. [7]). While the mucosal immune system remains understudied, several studies demonstrated increased neutrophil levels and activated alveolar macrophages/monocytes in bronchoalveolar lavage (BAL)[8] or nasopharyngeal and oropharyngeal swabs[9] from individuals hospitalized with COVID-19. T cell recruitment to the respiratory tract might be beneficial, as increased BAL CD8+ T cell clones[8] or increased tracheal aspirate CD4+ T cells associate with less severe disease or survival[10]. Nasopharyngeal swabs have been mostly used to analyze upper respiratory tract responses[9,11,12] but predominantly sample superficial cells, such as epithelial cells and neutrophils/monocytes, while they incompletely capture other immune cells, such as T cells. Although BAL and tracheal aspirates provide good insight into lower airways[8,10,13,14], they are difficult to collect longitudinally, after recovery or from healthy individuals and individuals that do not require intubation. Therefore, we still have a limited understanding of how COVID-19 affects mucosal immunity.

Here, we performed a prospective observational cohort study to characterize mucosal immune cell dynamics in the upper respiratory tract during acute, early recovery and convalescence stages of COVID-19. Individuals with PCR-confirmed SARS-CoV-2 infection were recruited after hospital admission (Fig. 1a). Longitudinal nasal curettage samples from 20 hospitalized individuals were collected, with up to 4 samples analyzed per individual. Curettage also samples immune cells from the nasal mucosa that are absent from the lumen[15]. Samples were collected from 2 to 61 d after hospital

Figure with panels a, b, c.

Panel a legend:
- Onset of symptoms
- Hospitalization
- ICU
- Individual transferred
- Nasal sampling

56 samples from 29 individuals and 12 healthy donors

39-antibody CyTOF panel

875,567 nasal immune cells clustered in 28 populations

Panel b labels: NK/ILC, CD11c⁺ NK, ILC, CD69⁺ NK, CD16⁻ gran, CD16^dim gran, CD16^hi gran, CD69⁺ gran, pDC, Mast cells, CD16⁺CD127⁺, CD161^hi, T_RM, T_EM, DN, Naive, EMRA, CD8⁺ T, Myeloid, CD163⁺ mono, Mono, CD163⁺CD206⁺ mono, CD206⁻ DC, CD206⁺ DC, Doublets, Dead, B cells, CD4⁺ T, DP, CD161⁺ T_RM, T_RM, T_EM, Naive, T_reg, Doublets

Panel c: heatmap, Median intensity scale 0–5.

---

admission (11–82 d after symptom onset). We stratified hospitalized individuals into those with acute infection (n = 9, 2–11 d since hospital admission) or in early recovery stage (ERS), which is defined by having moved from the intensive care unit (ICU) to ward (n = 11, 15–61 d since hospital admission and with an ICU stay period of 4–55 d). Individual characteristics, comorbidities, outcome and treatment are shown in Supplementary Table 1. Sixteen individuals were also sampled 5–6 weeks after hospital discharge (convalescent

Fig. 1 | Cytometry by time of flight (CyTOF) analysis of nasal immune cells during and after COVID-19 infection. a, Participant timelines. For each included individual, symptom onset (purple cross), hospitalization (black bar) and ICU stay (red bar) are indicated, aligned to the day of hospital admission. The orange x indicates one individual who was included but later transferred to a different hospital. One individual was discharged and then readmitted 1 d later. Blue diamonds indicate nasal curettage samples. The asterisk indicates one individual who was hospitalized for unrelated reasons at the time of positive test and symptom onset, and hospital admission was set at day of symptom onset. **b**, Hierarchical stochastic neighbor embedding (SNE) was used to cluster cellular landmarks on 37 markers into 12 populations. Some of these populations were then further divided into subpopulations in a second t-distributed stochastic neighbor embedding (t-SNE) plot embedding at a data level, as indicated in the rectangles. **c**, Heat map of marker expression per population. Median intensity per population is shown after arcsin transformation; ILCs, innate lymphoid cells; mDCs, myeloid dendritic cells; pDCs, plasmacytoid dendritic cells; T_{EM}, effector memory T cells; DP, double-positive T cells (CD4$^+$CD8$^+$); DN, double-negative T cells (CD4$^-$CD8$^-$); EMRA, effector memory reexpressing CD45RA; T_{reg}, regulatory T cells; gran, granulocytes; mono, monocytes.

individuals with a median of 77 d after symptom onset). Twelve sex- and age-matched healthy individuals with negative SARS-CoV-2 IgG and without symptoms of a respiratory tract infection were included.

In total, 875,564 nasal CD45$^+$ immune cells were analyzed from 56 samples (44 samples from 29 individuals with COVID-19 and 12 samples from healthy donors) using a 39-marker mass cytometry panel (Supplementary Table 2). Nasal CD45$^+$ immune cells were divided into eight main lineages and further subclustered into 28 populations (Fig. 1b,c). At a lineage level, granulocytes dominated the nasal immune profile of individuals with acute COVID-19 (Fig. 2a). This progressively decreased as granulocyte frequencies were slightly lower in ERS and further reduced during convalescence, more similar to healthy age-matched donors. All other lineages, apart from monocytes, appeared decreased during acute infection compared to healthy donors. To understand whether this apparent depletion was related to increased granulocytes and monocytes or to true lymphopenia, we normalized immune cell numbers to epithelial cell numbers for each sample, permitting an independent assessment of immune cell populations while correcting for variable sample yield. In healthy donors and recovered individuals, there was a strong correlation between epithelial and immune cell yields, as expected (Fig. 2b). This association was absent in hospitalized individuals, suggesting that the vastly increased granulocyte numbers are caused by a strong influx into the nasal mucosa. Indeed, when normalizing to epithelial cells, granulocytes and monocytes were highly increased during acute infection and ERS, while a non-significant trend for increased granulocytes was shown during convalescence compared to healthy donors (Fig. 2c and Supplementary Table 3). Other main cell lineages (B cells, NK cells, monocytes, pDCs, mDCs, CD4$^+$ T cells and CD8$^+$ T cells) were not statistically different between individuals with acute infection and healthy individuals. The observation that nasal lymphocyte numbers were unchanged is in contrast with the general lymphopenia detected in peripheral blood[2,6].

On the subclustering level (Fig. 1b,c), 8 of the 28 defined cell clusters were significantly elevated during COVID-19 (Fig. 2d,e and

Supplementary Tables 3 and 4). These showed clear associations with time from hospital admission (Extended Data Fig. 1). Three monocyte/macrophage populations, defined by CD163/CD206 expression, were increased during acute infection and ERS compared to convalescence and healthy donors. The granulocyte subsets CD16^{hi}, CD16^{dim} and CD16$^-$ neutrophils were all elevated to a variable degree during acute stage and in ERS compared to healthy donors. More detailed flow cytometric analysis (Extended Data Fig. 2) showed that nasal CD16$^-$ granulocytes expressed CXCR4 but not Siglec8 and were therefore not eosinophils but neutrophils that had shed CD16. Expressing CXCR4 and losing CD16 is typical for aged neutrophils that are (pre)apoptotic[16,17]. Furthermore, the expression of CD10 indicated that they were fully maturated[18]. Although there was no overall change in CD4$^+$ T cell numbers, effector (CCR7$^-$ CD45RO$^+$) CD4$^+$ T cells were 18-fold increased during acute stage compared to healthy donors. CD8$^+$ EMRA T cells (CCR7$^-$CD45RA$^+$) showed a non-significant trend toward increased numbers during acute infection. This agrees with earlier reports on peripheral blood T cells showing higher induction of antigen-specific CD4$^+$ T cells than of CD8$^+$ T cells, while the majority of SARS-CoV-2-specific CD8$^+$ T cells are EMRA[19,20]. These short-lived effector cells returned during ERS and convalescence to similar numbers as observed in healthy donors. Finally, CD11c$^+$ NK cells were increased in hospitalized individuals (acute and ERS) compared to convalescent individuals and healthy donors. CD11c marks NK cells with increased interferon-producing capacity and effector function[21]. Thus, a dynamic recruitment of various adaptive and innate populations mediating inflammation and antiviral function to the upper respiratory tract was observed during hospitalization that normalized in convalescence to levels resembling those in healthy donors. Of note, nasal B cells were not increased in hospitalized individuals nor did we detect plasmablasts in nasal mucosa, corroborating observations that mucosal antibody levels are reduced compared to systemic titers in hospitalized individuals[7]. Alternatively, B cells might not migrate past the submucosa, and further studies would be required to address mucosal B cell responses early in infection and/or in individuals with mild infection. To understand whether factors like sex,

Fig. 2 | Nasal cell lineage abundance during and after COVID-19. a, Stacked bar charts showing the composition of the nasal immune system in acute COVID-19 (red), during ERS (after ICU but still in the hospital; orange) or in COVID-19 convalescence (5–6 weeks after hospital discharge; pink) or in healthy donors (HD; blue). **b**, Correlation analyses between nasal immune and epithelial cells for hospitalized individuals (left) and convalescent individuals and healthy donors (right) are shown. Ranks of individuals are shown with color corresponding to group as well as a regression line (blue), 95% confidence intervals (shaded area) and results from Spearman correlation analysis. **c**, Ratio of nasal immune cell types normalized to the number of epithelial cells from the same sample. Individuals and box plots are shown, and paired samples between are indicated by gray lines. If a cell type was not detected in at least one sample, half the value of the lowest recorded number was added to each sample before log transformation. **d**, Heat map showing log$_{10}$ relative abundance (RA) of nasal cell clusters scaled per cluster. **e**, Box plots of nasal clusters during acute COVID-19. Box plots depict median and interquartile ranges, with whiskers extending to 1.5× interquartile range or maximum value; *$P < 0.05$ and ***$P < 0.001$ by linear mixed model with group as fixed effect and individuals as random effect with post hoc testing and Tukey multiple testing correction, followed by Benjamini–Hochberg correction for comparing multiple lineages or subsets. See Supplementary Tables 3 and 4 for exact test results. Samples from acute individuals ($n = 9$ individuals, 10 independent samples), ERS individuals ($n = 11$ individuals, 18 independent samples), convalescent individuals ($n = 16$ individuals) and healthy donors ($n = 12$ individuals) are used; Neutro, neutrophils. Only the first sample per donor in a time point (acute or ERS) is shown for **a**, **c** and **e**, but all are included in statistical modeling.

comorbidities and medication were drivers of nasal immune profiles, we performed multidimensional scaling (MDS) using all cell subsets (Extended Data Fig. 3). Acute individuals clustered separately from healthy donors and convalescent individuals, with ERS individuals intermediate. There was no clustering based on any of

these covariates, although larger sample sizes are needed to conclusively exclude such effects.

Next, we analyzed how nasal immune populations associated with systemic immune cells. We measured paired cryopreserved whole-blood samples using CyTOF ($n = 45$) and obtained
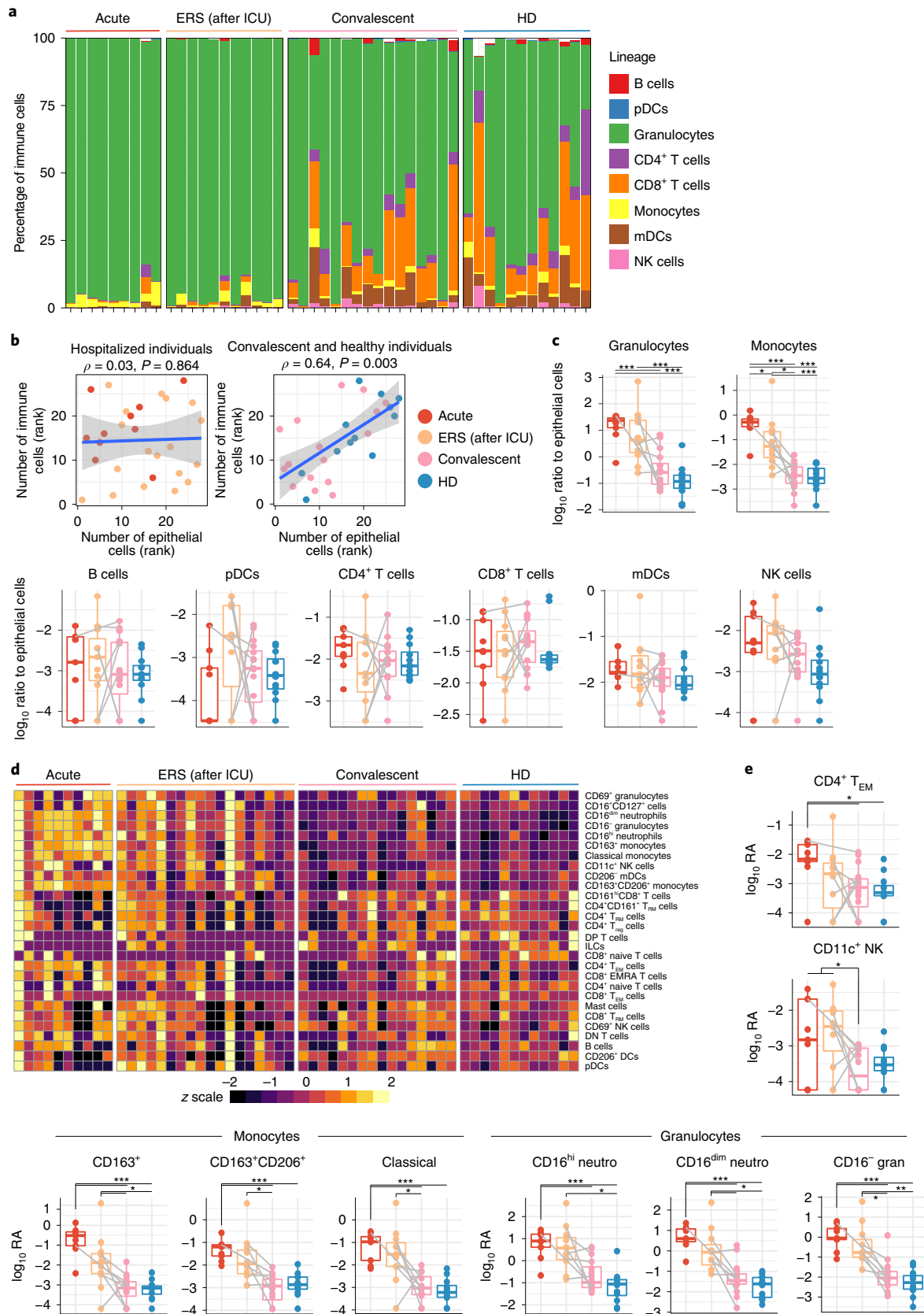
absolute cell counts from fresh whole blood using flow cytometry ($n = 43$) (Extended Data Fig. 4). With CyTOF, we identified 133 clusters and 15 subsets in blood (Extended Data Fig. 5), showing good concordance for most subsets with the freshly acquired flow cytometry data (Extended Data Fig. 6). Of all CyTOF clusters and subsets, 41 were significantly different between the groups after correction for multiple testing (Fig. 3a and Supplementary Tables 5 and 6). Acute COVID-19 was associated with a general lymphopenia, with reduced numbers of naive and central memory CD4+ T (T_{CM}) and CD8+ T cells, mucosal-associated invariant T (MAIT) cells, certain CD4+ T_{eff} cell clusters, pDCs, T cell antigen receptor-γδ (TCRγδ) T cells, NK cells and ILCs. By contrast, HLA-DR^low and HLA-DR^med monocytes, CD163+ monocytes, CD11c+ neutrophils and two clusters of CD38+PD1+CD4+ T_{eff} cells were increased during acute COVID-19. In convalescence, these perturbations were partially reverted, as only pDCs and NK cells remained significantly lower than healthy donors after multiple testing correction (Supplementary Table 6 and Extended Data Fig. 7). Flow cytometry confirmed in absolute counts the depletion of pDCs, NK cells and CD4+ T_{CM} cells (Extended Data Fig. 7). In addition, it showed that eosinophils and basophils were significantly depleted from blood during acute stage, while plasmablasts were increased (Fig. 3b). Our findings in blood thus agree with previous reports[1–5].

To understand how the perturbations in blood and nasal mucosa were linked to each other and clinical variables, we performed MDS using all significant populations (Fig. 3c). Proinflammatory nasal subsets generally clustered together with viral load. Indeed, viral load positively correlated with nasal monocyte and neutrophil levels (Extended Data Fig. 8). Nasal proinflammatory populations were also associated with blood lymphopenia (Fig. 3c and Extended Data Fig. 8). We next evaluated whether mucosal immune cells correlated with their systemic counterparts to assess whether peripheral responses can serve as correlates for mucosal events. Both MDS and hierarchical clustering of the correlation matrix suggested that increased blood monocyte clusters (HLA-DR^low and CD163+) associate with nasal monocytes/neutrophils and viral load. Indeed, HLA-DR^low monocyte numbers in blood were significantly correlated with nasal monocyte numbers. Moreover, CD38+PD1+CD4+ T_{eff} cell clusters 14 (HLA-DR+) and 2 (CD27+) were strongly linked to nasal CD4+ T_{eff} cells (Fig. 3d). Further characterization of nasal CD4+ T_{EM} cells revealed two global subsets, with a group of activated cells expressing CD38, PD1 and CTLA-4 (Extended Data Fig. 9a–c). These activated cells were mostly Tbet+, suggesting a type 1 helper T cell (T_H1) profile. Thus, we found that blood lymphopenia did not associate with the concurrent recruitment of lymphocytes to nasal mucosa, while it was associated with increased proinflammatory populations in nasal mucosa. In addition, we also found populations

in blood, that is, CD38+PD1+CD4+ T_{eff} cells and HLA-DR^low monocytes, that correlated with their nasal counterparts.

Nasal monocyte and granulocyte subsets, blood lymphopenia and perturbed clusters, viral load and time since hospital admission were all associated with each other. To understand which factors associated with clinical features and to disentangle these relationships, we integrated the immunological datasets using MOFA (Fig. 3e). This tool detects hidden LFs explaining variation both within and across different datasets[22]. Here, six LFs were found to which both nasal data and flow cytometric absolute blood counts contributed most. LF1 was strongly associated with time since hospital admission, increasing progressively during recovery and convalescence to levels similar to healthy donors (Fig. 3f). LF1 also was negatively correlated with viral load using marginal correlation, but this was not significant when correcting for time since admission using linear mixed modeling. The nasal monocyte subsets (CD163+ and classical monocytes), blood plasmablasts and CD4+:CD8+ T cell ratio were the most important immune populations negatively driving LF1 (and thus associated with acute infection) (Fig. 3g). Indeed, blood plasmablasts and nasal CD163+ monocytes were positively correlated (Fig. 3d). This links to previous findings showing that antibody development associates with increased nasal and blood inflammatory cytokines during COVID-19 (ref. [7]). Immune populations positively contributing to LF1 (thus associated with longer time since infection, resolving of infection and immune homeostasis in healthy donors) were nasal CD8+ T cells and blood lymphocytes (TCRγδ T cells, CD8+ T terminally differentiated effector cells and NK cells). LF3 was strongly driven by nasal neutrophils, with CD16^hi neutrophils as top loading factor (Fig. 3h). LF3 was negatively associated with blood oxygen saturation, considering time since hospital admission and viral load in linear mixed modeling (Fig. 3i). A similar linear model including CD16^hi neutrophils instead of LF3 confirmed that these cells were independently and inversely associated with oxygen saturation, suggesting that independent of time since admission and viral load, enhanced nasal neutrophils are associated with hypoxia.

Subsequently, we looked more closely at phenotypic expression profiles of abundant nasal cell clusters and to what extent these profiles normalized after hospital discharge. Among all increased monocyte subsets, acute individuals had relatively more CD163+ and fewer CD163+CD206+ monocytes/macrophages than healthy donors, which normalized during recovery (Fig. 4a). These CD206+ cells are likely fully differentiated tissue-resident macrophages, while CD163+ monocytes could represent recently recruited monocytes in individuals with COVID-19 (refs. [10,23]). This hypothesis was supported by trajectory analysis (Fig. 4b). HLA-DR expression was reduced on nasal monocytes/macrophages during hospitalization, which normalized during convalescence (Fig. 4c). This agrees with

Fig. 3 | Integration of blood and nasal immune responses during and following COVID-19. a, Heat map showing significant CyTOF whole-blood (WB) clusters. Individuals are shown in columns ($n = 34$ individuals, 45 samples) ordered by days since hospital admission per group. **b**, Absolute counts of cellular subsets in blood. Acute ($n = 4$ individuals, 5 samples), ERS ($n = 7$ individuals, 10 samples), convalescent ($n = 16$ individuals) and healthy donors ($n = 12$ individuals). Individuals and box plots are shown, and paired samples are indicated by gray lines. Box plots depict median and interquartile ranges, with whiskers extending to 1.5× interquartile range or maximum value. Only the first sample per donor per time point is shown. Results were compared against healthy donors from testing using linear mixed models with group as fixed effect and individuals as random effect with post hoc testing and Tukey multiple testing correction. **c**, MDS of significant nasal mucosa (purple) or whole-blood (red) CyTOF clusters, whole-blood flow subsets (green) and clinical parameters (cyan). Only samples from hospitalized individuals with complete data were included ($n = 3$ acute, $n = 10$ ERS with 15 samples). **d**, Correlation plots between nasal and whole-blood CyTOF subsets. Individuals with paired data ($n = 34$ individuals, 45 samples) are shown, with colors indicating disease status. Estimates and 95% confidence intervals from linear regression analyses are shown as black lines with shaded areas. Pearson correlation analysis results are depicted. **e**, Multiomics factor analysis (MOFA) of all three datasets. Explained variation of all data explained per dataset or per latent factor (LF) is shown in the bar chart and heat map, respectively. **f**, Factor scores of LF1 and LF3 plotted against day of hospital admission. Healthy donors are separated by a dashed line. Loess regression estimate is indicated by the black line, and shaded areas represent 95% confidence intervals. **g,h**, The top five loadings per dominant dataset are indicated for LF1 (**g**) and LF3 (**h**). td, terminally differentiated. **i**, Association between LF3 and oxygen saturation (spO_2) in blood. Symbols indicate hospitalized individuals ($n = 20$ individuals, 28 samples), and the dark line depicts the estimate from a linear mixed model with viral load, days since admission and spO_2 included as fixed effects and individuals as random effects. The model estimate (est) and P value are depicted.

our finding that HLA-DR^low monocytes, but not classical HLA-DR^high monocytes, in blood strongly correlate with nasal monocyte numbers. Low HLA-DR expression is typical for myeloid-derived suppressor cells (MDSC), and MDSC-like cell expansion in peripheral blood during severe COVID-19 was previously reported[24]. Our results suggest that these cells may rapidly seed the upper airway mucosa where they might further differentiate into macrophages.

CD163^+CD206^+ monocytes/macrophages also expressed elevated interleukin-3 (IL-3) receptor (CD123) levels during hospitalization, while CD163^+ and CD163^- monocyte subsets generally lacked CD123 (Fig. 1c). We then investigated more closely CD16^- neutrophils by further grouping them into seven subclusters (Fig. 4d). Subcluster 1, characterized by increased IL-7 receptor (CD127) expression, was significantly increased during hospitalization and
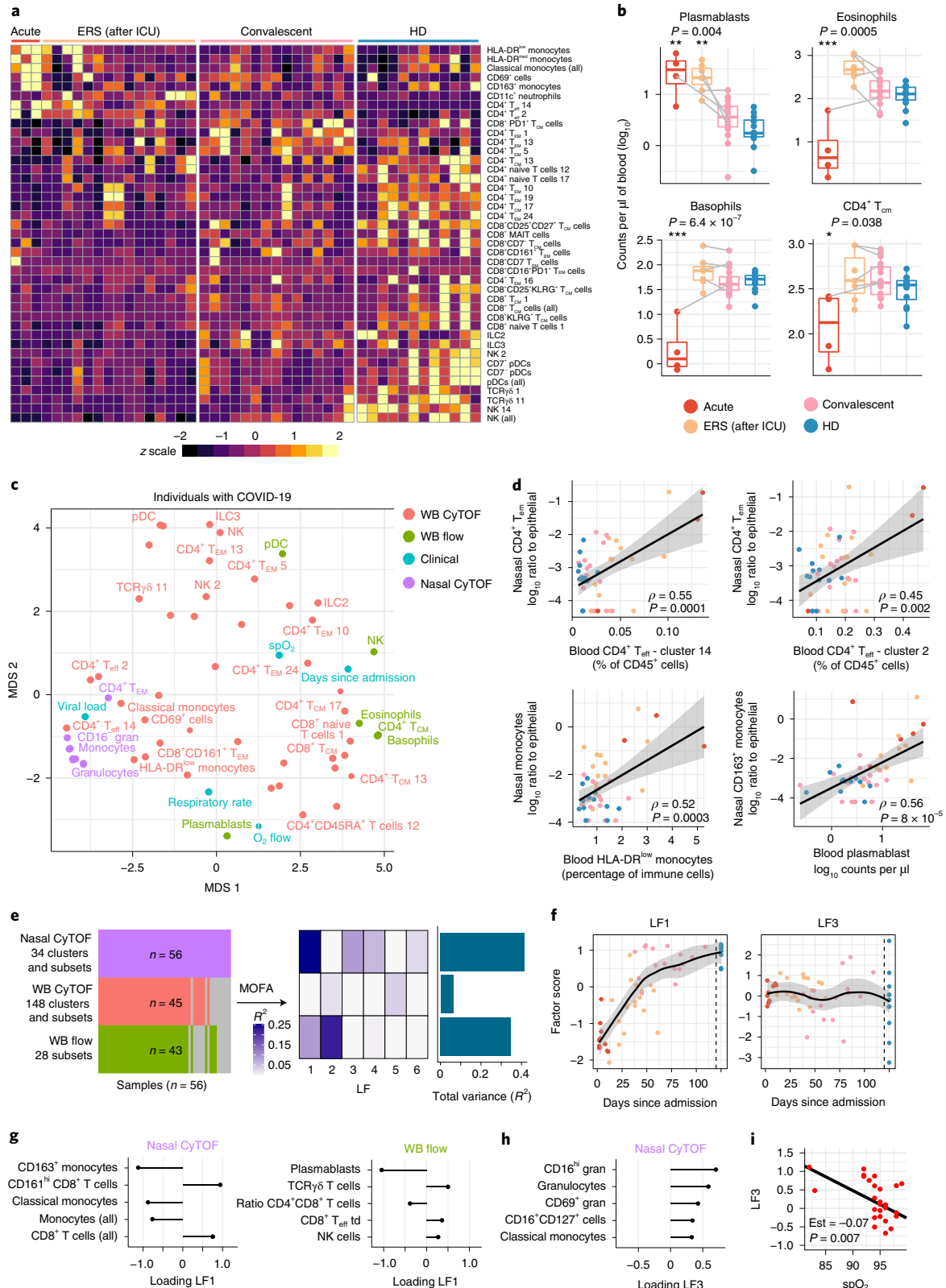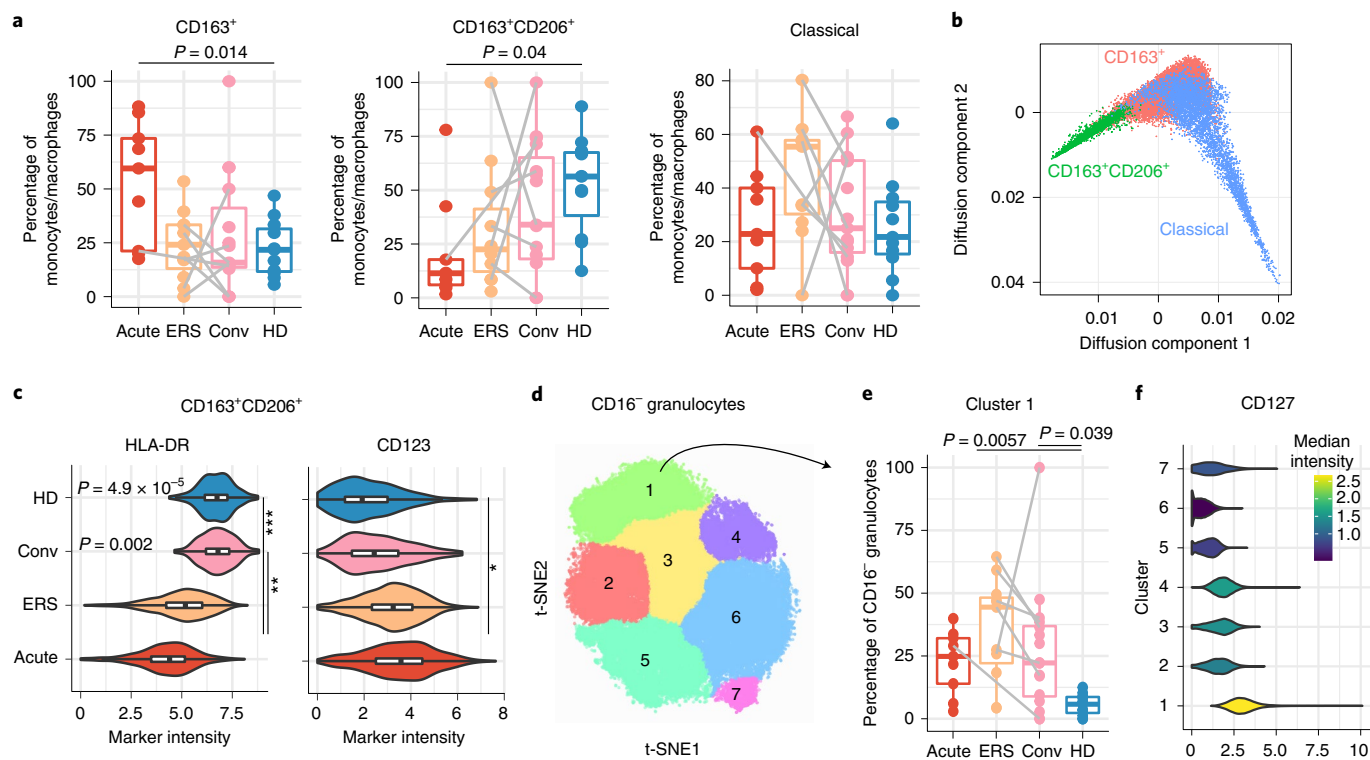
Fig. 4 | Phenotypic changes of innate and adaptive cell subsets during and after COVID-19. a, Percentages of monocytes expressing CD163, CD163 and CD206 or neither (classical). Subset frequencies are shown for individuals with acute COVID-19 (red), during ERS (after ICU but still in the hospital; orange), in convalescent individuals (Conv; pink) and in healthy donors (blue). Box plots and individual data points are depicted, with paired individual samples connected by a gray line. **b**, Diffusion map showing trajectory analysis of monocyte/macrophage subsets. All cells are colored according to their clustered phenotype. **c**, Violin plots showing expression of HLA-DR and CD123 on CD163+CD206+ monocytes/macrophages. All cells are shown; ***$P = 4.9 \times 10^{-5}$, **$P = 0.002$ and *$P = 0.032$ by statistical modeling as described below on the mean signal intensity per sample. **d**, Clustering of CD16⁻ granulocytes using t-SNE and Gaussian mean shift. **e**, Proportion of cluster 1 within CD16⁻ granulocytes. Box plots and individual data points are depicted, with paired individual samples connected by a gray line. **f**, Violin plots showing expression of CD127 per cluster. Box plots depict median and interquartile ranges, with whiskers extending to 1.5× interquartile range or maximum value. Samples from acute individuals ($n = 9$ individuals, 10 independent samples), ERS ($n = 11$ individuals, 18 independent samples), convalescence ($n = 16$ individuals) and healthy donors ($n = 12$ individuals) are used. Only the first sample per donor in a time point (acute or ERS) is shown for **a** and **e**, but all are included in the statistical modeling. Statistical results from a linear mixed model with group as fixed effect and individuals as random effect with post hoc testing and Tukey multiple testing correction are shown.

convalescence compared to healthy donors (Fig. 4e,f). Thus, while total neutrophil numbers normalized during convalescence compared to healthy donors, alterations in their phenotype remained visible. The half-life of granulocytes is relatively low (hours to days)[25], suggesting either ongoing recruitment of altered cells or continued local perturbation in the respiratory mucosa.

In mice, influenza-specific CD8+ T_{RM} cells persist in the nasal mucosa following infection and efficiently control secondary infections[26]. Therefore, we assessed whether long-term protective CD8+ T cell immunity develops in the nasal mucosa of individuals with COVID-19 to act as 'gatekeepers' and protect against reinfection. The majority of nasal CD8+ T cells highly expressed CD69 and very little KLRG1, defining them as T_{RM} cells[27] (Fig. 1b,c). Subclustering of CD8+ T_{RM} cells showed variable expression of activation markers CD38, HLA-DR and Tbet (Fig. 5a). CD8+ T_{RM} cells from acute phase, ERS, convalescent and healthy individuals clustered differentially, indicating altered phenotypes during and following COVID-19 (Fig. 5b). Indeed subcluster frequencies significantly differed between groups (Fig. 5c,d). Subcluster 5 (HLA-DR/Tbet/ CD38 expression) was increased in hospitalized individuals, while subcluster 3 (HLA-DR/CD38 expression) was higher during convalescence than in healthy donors as were all CD38+ T_{RM} cells. Thus CD8+ T_{RM} cells have an increased activation profile, which persisted at least 5–6 weeks after hospital discharge.

To demonstrate antigen specificity, we performed TCR sequencing of nasal samples during convalescence. These sequences were compared with SARS-CoV-2-specific CD8+ and CD4+ T cells sorted by FACS from paired peripheral blood based on either induction of activation markers CD137 or CD154 after stimulation with SARS-CoV-2 peptide pools or following binding to SARS-CoV-2-specific peptide–major histocompatibility complex (pMHC) tetramer complexes (Extended Data Fig. 9d,e). For four convalescent individuals (two with and two without a prior ICU stay), we obtained >10 unique TCRs from both nasal cells and sorted SARS-CoV-2-specific peripheral blood cells. In all individuals, nasal TCRs overlapped with TCRs from FACS-sorted SARS-CoV-2-specific CD8+ T cells from peripheral blood (Fig. 5e and Supplementary Table 7). The number of paired T cell clones ranged from 1 to 12 per individual and included both α- and β-chains that overlapped with CD8+ T cells sorted based on activation following SARS-CoV-2 stimulation and/or tetramer binding. Two individuals had strongly immunodominant nasal SARS-CoV-2-specific TCRs, with each clone accounting for >10% of all TCR reads. Samples were collected 36–70 d after viral clearance, and, on average, 60.7% of the nasal CD8+ T cells in these convalescent samples were of a T_{RM} phenotype (Fig. 5f), indicating that antigen-specific tissue-resident memory was induced. Of note, the number of unique SARS-CoV-2-specific T cell clones detected and their overlap between nasal mucosa and peripheral blood might be underestimated as (1)
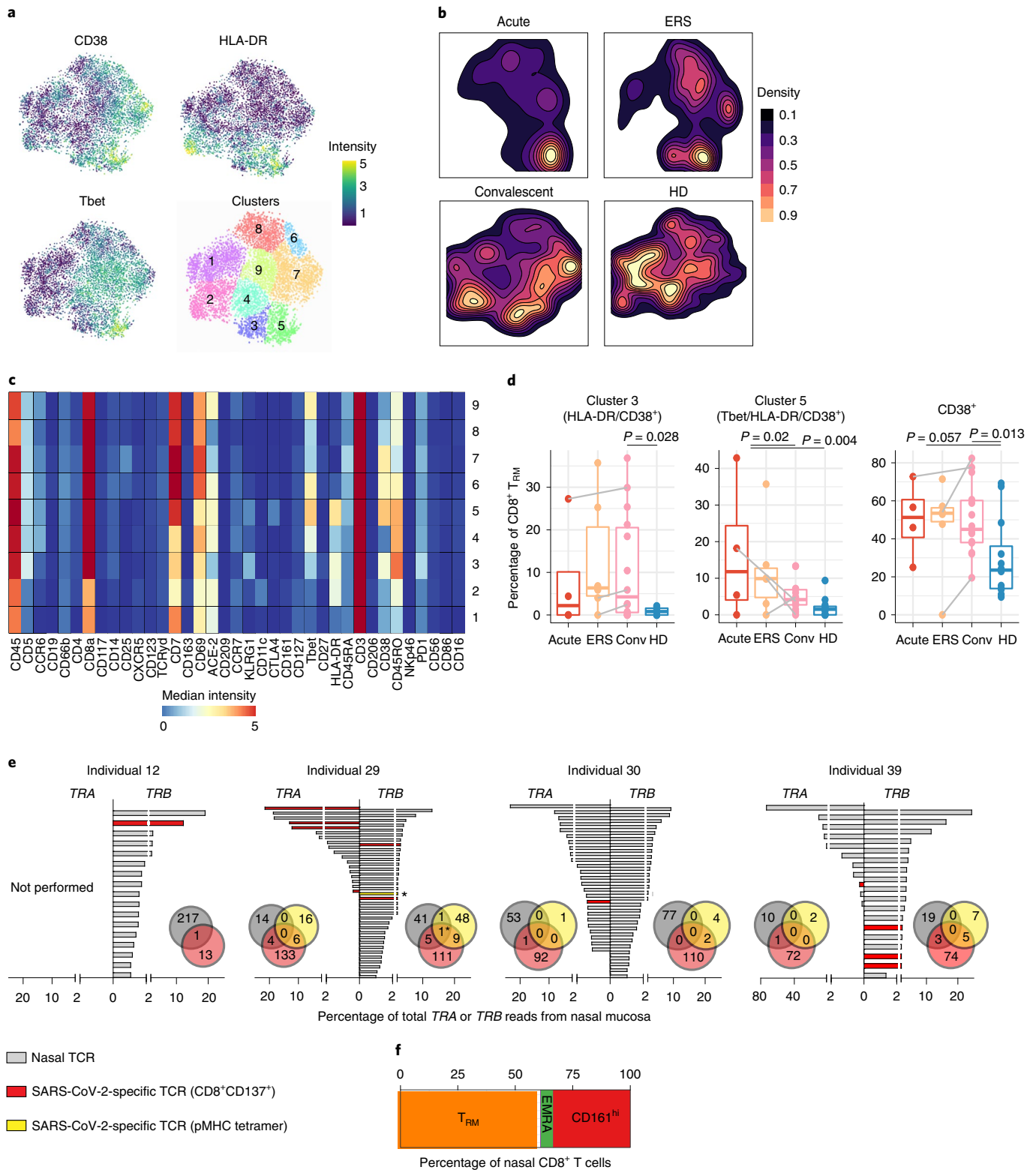
a limited number of pMHC tetramers could be included, and (2) in the activation-induced assay, blood CD8[+] T cells were isolated based on reactivity toward structural proteins, while CD8[+] T cell reactivity is also directed against non-structural proteins in individuals with severe COVID-19 (refs. [28–30]). Indeed, matching nasal TCR sequences to publicly known SARS-CoV-2-specific T cell clones identified another specific T cell clone (Supplementary Table 7). In contrast to CD8[+] T cells, there were no overlapping nasal TCRs and SARS-CoV-2-specific CD4[+] T cells from peripheral blood (Supplementary Table

7), which may reflect a better induction or maintenance of nasal CD8[+] T cells. Alternatively, it may be that too few nasal cells were obtained during sampling, missing nasal SARS-CoV-2-specific CD4[+] T cells during convalescence as CD8[+] T cells are more abundant than CD4[+] T cells in the nasal mucosa. Taken together, we demonstrated that SARS-CoV-2-specific CD8[+] T cells in the nasal mucosa can persist for months after viral clearance. This suggests the establishment of local protective immune memory responses that could rapidly control and attenuate reinfections by SARS-CoV-2.

Fig. 5 | SARS-CoV-2-specific T cells are present in the nasal mucosa in convalescent individuals with COVID-19. a, t-SNE analysis of all 7,826 CD8+ T_{RM} cells. Expression of CD38, HLA-DR or Tbet overlaid onto a t-SNE plot. Bottom right, clustering of t-SNE using Gaussian mean shift. **b**, Two-dimensional kernel density estimation from t-SNE plot of all cells divided per group. **c**, Heat map showing median expression for all markers per T_{RM} cluster. **d**, Percentage of CD8+ T_{RM} cells belonging to either cluster 3 or cluster 5 or all clusters expressing CD38. Box plots and individual data points are depicted, with paired individual samples connected by a gray line. Box plots depict median and interquartile ranges, with whiskers extending to 1.5× interquartile range or maximum value. Only samples with >10 CD8+ T_{RM} cells are included, and samples from acute individuals (red; $n = 4$ individuals, 5 samples), ERS individuals (orange; $n = 7$ individuals, 13 samples), convalescent individuals (pink; $n = 14$ individuals) and healthy donors (blue; $n = 12$ individuals) are used. Only the first sample per donor per time point (acute or ERS) is shown, but all are included in the statistical modeling. P values by linear mixed model with group as fixed effect and individual as random effect followed by post hoc testing and Tukey multiple testing correction are shown. **e**, Bar plots showing the frequency of TCR clonotypes in nasal samples collected from four donors. T cell clones with non-paired TCRα or TCRβ sequences are depicted on the left and right side of the y axis, respectively, and are ranked by their frequency. Only clones with a frequency >1% among reads are shown for TCRβ clones, while all TCRα clones are shown. Red- and yellow-colored bars indicate clones also present in SARS-CoV-2-specific CD8+ T cells sorted by fluorescence-activated cell sorting (FACS) from paired peripheral blood. Venn diagrams depict the total number of clones from nasal samples (gray), SARS-CoV-2-specific CD8+ T cells in peripheral blood mononuclear cells (PBMC) based on activation after peptide stimulation (red) or tetramer binding (yellow), including clones accounting for <1% of total reads. Star indicates that this clone was found in both the tetramer-sorted and stimulation-sorted fraction. **f**, Bar plot showing the average nasal CD8+ T cell composition from the convalescent samples of four individuals with TCR sequencing. Naive and T_{EM} cells were too infrequent to be visible and are not labeled.

In conclusion, we provide an in-depth analysis of how COVID-19 affects nasal mucosal immunity during acute infection, early recovery and convalescence. We confirmed that acute COVID-19 leads to transient increases of respiratory granulocyte/monocyte subsets and extended this to CD11c+ NK cells and CD4+ T_{eff} cells. This proinflammatory response positively correlated with viral load and inversely correlated with time since admission, while specifically nasal neutrophils showed a strong negative association with oxygen saturation in blood. Furthermore, increased blood HLA-DR^low monocyte numbers correlated with their emergence at the nasal mucosa, and blood CD38+PD1+CD4+ T_{eff} cells positively associated with nasal CD4+ T_{eff} abundance. Strikingly, unlike in blood, no general lymphopenia was observed in the nasal mucosa. Simultaneously, lymphocyte populations depleted from blood did not show a concurrent increase in nasal mucosa, suggesting that lymphopenia is not caused by nasal mucosal homing, but other mechanisms might be involved. Although most cell numbers returned to levels comparable to age-matched healthy donors during early and later recovery stages, several phenotypic changes in nasal immune populations persisted. For example, a subset of CD127+ granulocytes remained elevated, and continued increased activation of CD8+ T_{RM} cells was observed during convalescence, while SARS-CoV-2-specific CD8+ T cells persisted for at least 2 months after viral clearance in the nasal mucosa. Altogether, this study provides unique insights into mucosal and systemic immune cell dynamics both during acute infection and recovery of COVID-19.

Online content

Any methods, additional references, Nature Research 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/s41590-021-01095-w.

References

1. Diao, B. et al. Reduction and functional exhaustion of T cells in patients with coronavirus disease 2019 (COVID-19). *Front. Immunol.* **11**, 827 (2020).
2. Lucas, C. et al. Longitudinal analyses reveal immunological misfiring in severe COVID-19. *Nature* **584**, 463–469 (2020).
3. Mann, E. R. et al. Longitudinal immune profiling reveals key myeloid signatures associated with COVID-19. *Sci. Immunol.* **5**, eabd6197 (2020).
4. Maucourant, C. et al. Natural killer cell immunotypes related to COVID-19 disease severity. *Sci. Immunol.* **5**, eabd6832 (2020).
5. Rodriguez, L. et al. Systems-level immunomonitoring from acute to recovery phase of severe COVID-19. *Cell Rep. Med.* **1**, 100078 (2020).
6. Hadjadj, J. et al. Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. *Science* **369**, 718–724 (2020).
7. Smith, N. et al. Distinct systemic and mucosal immune responses during acute SARS-CoV-2 infection. *Nat. Immunol.* **22**, 1428–1439 (2021).
8. Liao, M. et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* **26**, 842–844 (2020).
9. Chua, R. L. et al. COVID-19 severity correlates with airway epithelium–immune cell interactions identified by single-cell analysis. *Nat. Biotechnol.* **38**, 970–979 (2020).
10. Szabo, P. A. et al. Longitudinal profiling of respiratory and systemic immune responses reveals myeloid cell-driven lung inflammation in severe COVID-19. *Immunity* **54**, 797–814 (2021).
11. Trump, S. et al. Hypertension delays viral clearance and exacerbates airway hyperinflammation in patients with COVID-19. *Nat. Biotechnol.* **39**, 705–716 (2020).
12. Ziegler, C. G. K. et al. Impaired local intrinsic immunity to SARS-CoV-2 infection in severe COVID-19. *Cell* **184**, 4713–4733 (2021).
13. Grant, R. A. et al. Circuits between infected macrophages and T cells in SARS-CoV-2 pneumonia. *Nature* **590**, 635–641 (2021).
14. Wauters, E. et al. Discriminating mild from critical COVID-19 by innate and adaptive immune single-cell profiling of bronchoalveolar lavages. *Cell Res.* **31**, 272–290 (2021).
15. Jochems, S. P. et al. Novel analysis of immune cells from nasal microbiopsy demonstrates reliable, reproducible data for immune populations, and superior cytokine detection compared to nasal wash. *PLoS ONE* **12**, e0169805 (2017).
16. Dransfield, I. et al. Neutrophil apoptosis is associated with a reduction in CD16 (Fcγ-RIII) expression. *J. Immunol.* **153**, 1254–1263 (1994).
17. De Filippo, K. & Rankin, S. M. CXCR4, the master regulator of neutrophil trafficking in homeostasis and disease. *Eur. J. Clin. Invest.* **48**, e12949 (2018).
18. Marini, O. et al. Mature CD10+ and immature CD10− neutrophils present in G-CSF-treated donors display opposite effects on T cells. *Blood* **129**, 1343–1356 (2017).
19. Dan, J. M. et al. Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science* **371**, eabf4063 (2021).
20. Zhou, R. et al. Acute SARS-CoV-2 infection impairs dendritic cell and T cell responses. *Immunity* **53**, 864–877 (2020).
21. Burt, B. M. et al. CD11c identifies a subset of murine liver natural killer cells that responds to adenoviral hepatitis. *J. Leukoc. Biol.* **84**, 1039–1046 (2008).
22. Argelaguet, R. et al. Multi-omics factor analysis—a framework for unsupervised integration of multi-omics data sets. *Mol. Syst. Biol.* **14**, e8124 (2018).
23. Taylor, P. R. et al. Macrophage receptors and immune recognition. *Annu. Rev. Immunol.* **23**, 901–944 (2005).
24. Kvedaraite, E. et al. Major alterations in the mononuclear phagocyte landscape associated with COVID-19 severity. *Proc. Natl Acad. Sci. USA* **118**, e2018587118 (2021).
25. Kelly, E. A. et al. Potential contribution of IL-7 to allergen-induced eosinophilic airway inflammation in asthma. *J. Immunol.* **182**, 1404–1410 (2009).
26. Pizzolla, A. et al. Resident memory CD8+ T cells in the upper respiratory tract prevent pulmonary influenza virus infection. *Sci. Immunol.* **2**, eaam6970 (2017).
27. Herndler-Brandstetter, D. et al. KLRG1+ effector CD8+ T cells lose KLRG1, differentiate into all memory T cell lineages, and convey enhanced protective immunity. *Immunity* **48**, 716–729 (2018).

28. Grifoni, A. et al. Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. *Cell* **181**, 1489–1501 (2020).

29. Ferretti, A. P. et al. Unbiased screens show CD8[+] T cells of COVID-19 patients recognize shared epitopes in SARS-CoV-2 that largely reside outside the spike protein. *Immunity* **53**, 1095–1107 (2020).

30. Peng, Y. et al. Broad and strong memory CD4[+] and CD8[+] T cells induced by SARS-CoV-2 in UK convalescent individuals following COVID-19. *Nat. Immunol.* **21**, 1336–1345 (2020).

in collaboration with BEAT-COVID group

M. S. Arbous[6], B. M. van den Berg[8], S. de Bruin-Versteeg[5], S. Cannegieter[9], K. Canté[5], C. M. Cobbaert[10], A. M. van der Does[7], J. J. M. van Dongen[5], J. Eikenboom[11], M. C. W. Feltkamp[4], A. Geluk[1], J. J. Goeman[12], M. Giera[13], R. J. Groenland[5], T. Hankemeier[14], M. H. M. Heemskerk[2], P. S. Hiemstra[7], C. H. Hokke[3], R. van der Holst[5], J. J. Janse[3], S. P. Jochems[3], S. A. Joosten[1], M. Kikkert[4], S. Klaver Flores[5], L. Lamont[14], J. Manniën[12], B. de Mooij[5], T. H. M. Ottenhoff[1], K. Pike-Overzet[5], T. Pongracz[13], M. R. del Prado[6], N. Queralt Rosinach[15], M. Roestenberg[1,3], M. Roos[15], A. H. E. Roukens[1], Alita J. van der Sluijs-Gelling[5], H. H. Smits[3], E. J. Snijder[4], F. J. T. Staal[5], L. A. Trouw[5], R. Tsonaka[12], A. Verhoeven[13], L. G. Visser[1], J. J. C. de Vries[4], D. J. van Westerloo[6], J. Wigbers[6], H. J. van der Wijk[12], R. C. van Wissen[10], M. Wuhrer[13], M. Yazdanbakhsh[3] and M. Zlei[7]

[8]Department of Internal Medicine, Nephrology, Leiden University Medical Center, Leiden, the Netherlands. [9]Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands. [10]Department of Clinical Chemistry, Leiden University Medical Center, Leiden, the Netherlands. [11]Department of Internal Medicine, Thrombosis and Hemostasis, Leiden University Medical Center, Leiden, the Netherlands. [12]Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, the Netherlands. [13]Center for Proteomics and Metabolomics, Leiden University Medical Center, Leiden, the Netherlands. [14]Division of Systems Biomedicine and Pharmacology, Leiden Academic Center for Drug Research, Leiden University, Leiden, the Netherlands. [15]Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands.

in collaboration with COVID-19 LUMC group

Josine A. Oud[2], Meryem Baysan[3,6], Jeanette Wigbers[6], Lieke J. van Heurn[9], Susan B. ter Haar[9], Alexandra G. L. Toppenberg[9], Laura Heerdink[9], Annekee A. van IJlzinga Veenstra[9], Anna M. Eikenboom[9], Julia M. Wubbolts[1], Jonathan Uzorka[1], Willem Lijfering[9], Romy Meier[2], Ingeborg de Jonge[9], Sesmu M. Arbous[6], Mark G. J. de Boer[1], Anske G. van der Bom[9], Olaf M. Dekkers[9] and Frits Rosendaal[9]

Methods

Study design and ethics. In this prospective observational cohort study, adult individuals with PCR-confirmed COVID-19 who were admitted to our academic hospital were recruited. All hospitalized individuals had hypoxia. The study was performed at the Leiden University Medical Center (LUMC), Leiden, the Netherlands, in individuals included from April 2020 to December 2020. All participants provided written informed consent. After informed consent was obtained, longitudinal sampling was performed for the duration of the hospital admission, and one convalescent sample was obtained at the outpatient follow-up appointment, which was scheduled 6 weeks after hospital discharge. Ethical approval was obtained from the Medical Ethical Committee Leiden-Den Haag-Delft (NL73740.058.20). The trial was registered in the Dutch Trial Registry (NL8589). As individuals in the ICU had substantial breathing support, we were unable to collect nasal mucosal cells from individuals in the ICU. Twelve healthy donors were included in the study. These individuals are all 60 years or older and with a male:female ratio of 2:1 to match the participant population. The healthy donors had no recent history of symptoms of airway infection (fever, cough, hypoxia, rhinorrhea, myalgia, anosmia and/or ageusia or fatigue) and were included after confirmed negative SARS-CoV-2 IgG.

Nasal cell collection and storage. Nasal cells were collected by gently scraping the nasal inferior turbinate using curettes (Rhino-Pro, Arlington Scientific) and placing them in a 15-ml Falcon tube (Corning) containing 8 ml of precooled sterile PBS containing 5 mM EDTA (Life Technologies). Per individual and time point, two curettes from one nostril were collected. Cells were dislodged by pipetting liquid up and down the tip of the curette, and cells were centrifuged at 300g for 10 min at 4 °C. The supernatant was completely removed, and cells were resuspended in 500 µl of PBS. For fixation, an equal amount of freshly prepared 8% formaldehyde (Fisher Scientific) was added, followed by a 30-min incubation at room temperature. Cells were then centrifuged at 800g for 10 min. The supernatant was completely removed, and the pellet was resuspended in 1 ml of heat-inactivated fetal bovine serum (FBS) containing 10% DMSO and moved to a cryovial. Cryovials were frozen in a Mr. Frosty freezing container (Thermo Fisher Scientific) at –80 °C and moved to liquid nitrogen within 3 d.

Whole blood cryopreservation. Per sample, 200 µl of whole blood that was collected in an EDTA tube was added to 1 ml of cold Cryostor CS10 (Stemcell Technologies) in a 1.8-ml cryovial and mixed by gently inverting. Cryovials were then placed in a cold Mr. Frosty and moved after 10 min to –80 °C. The next day, cryovials were moved to liquid nitrogen.

CyTOF staining. Samples were barcoded and measured in batches. In every batch, one aliquot of PBMCs from a reference sample was included to normalize staining between batches. Nasal cells were thawed in a water bath at 37 °C, followed by dropwise addition of 2 ml of RPMI + 50% FBS and centrifugation for 10 min at 1,600 r.p.m. at room temperature. Supernatant was discarded by pipetting. Reference PBMCs were washed with 2 ml of PBS and then fixed with 4% formaldehyde for 15 min at room temperature. Reference PBMCs were washed two times with 2 ml of BD Perm/Wash (BD). Nasal cells were washed one time with 1 ml of BD Perm/Wash, and if clumps were visible, cells were filtered through a 100-µm cell strainer (Thermo Fisher Scientific). Nasal cells and reference PBMCs were resuspended in 50 µl of Perm/Wash, and 50 µl of barcode mix targeting β2-microglobulin (B2M) was added to each individual sample in a 6-choose-3 scheme using cadmiums 106, 110, 111, 112, 114 and 116 (refs. [31,32]). Samples were incubated for 30 min at room temperature and then washed with 4 ml of Cell Staining Buffer (Fluidigm). Cells were centrifuged for 5 min at 800g, and the supernatant was removed, resuspended and combined into 3 ml of Perm/Wash. Cells were centrifuged again for 5 min at 800g and were resuspended in 45 µl of Perm/Wash. FcR block (Biolegend; 5 µl) and sodium heparin (0.5 µl, 100 U ml⁻¹) were added to prevent aspecific binding of antibodies, and cells were incubated for 20 min at room temperature[33]. Then, 50 µl of antibody cocktail (Supplementary Table 2) was added, followed by a 45-min incubation at room temperature. Cells were then washed twice with 2 ml of Cell Staining Buffer and centrifuged for 5 min at 800g. DNA was then stained overnight at 4 °C using 1 ml of Fix and Perm buffer (Fluidigm) containing 1,000× diluted Intercalator-Ir (Fluidigm). Cells were then washed with Cell Staining Buffer, counted and divided into tubes of 1×10^6 cells and pelleted. Pellets were washed and resuspended in cell acquisition solution (CAS, Fluidigm) with EQ Four Element Calibration Beads (Fluidigm) and acquired on a Helios mass cytometer (Fluidigm) with CyTOF Software (v7.0.8493) at the Flow Cytometry Core Facility (FCF) of LUMC in Leiden, the Netherlands (https://www.lumc.nl/research/facilities/fcf). Whole-blood samples were thawed in a water bath at 37 °C, followed by dropwise addition of 4 ml of thawing medium (RPMI 1640, penicilin/streptomycin, pyruvate, L-glutamine with 20% heat-inactivated FBS, 2 mM Mg²⁺ and 25 U ml⁻¹ benzonase). Cells were centrifuged at 400g for 10 min and resuspended in 2 ml of 1× RBC lysis buffer (Biolegend). After 10–15 min, the cells were centrifuged at 400g for 10 min, resuspended in medium and filtered through a 100-µm cell strainer if clumps were visible. Staining with barcodes and surface antibodies and acquisition was performed as for nasal cells, except the sodium heparin blocking step was omitted and Cell Staining Buffer was used for whole blood instead of Perm/Wash.

Granulocyte flow cytometry analysis. Cryopreserved whole blood (one donor) or fixed and stored nasal cells (two donors) were thawed at 37 °C. Whole-blood cells were then fixed as described above for nasal cells. Then, 4 ml of BD Perm/Wash buffer was added to each sample, followed by a 5-min centrifugation at 800g. Cells were washed again in 4 ml of BD Perm/Wash and centrifuged again. Cells were resuspended in 100 µl of antibody cocktail containing CD45-BV785 (1:200; clone HI30, Biolegend, 304047), CD66b-APC-Vio770 (1:800; clone REA306, Miltenyi, 130-120-146), CD127-PercpCy5.5 (1:50; clone A019D5, Biolegend, 351321), CD16-BV711 (1:100; clone 3G8, Biolegend, 302043), Siglec8-PEDazzle594 (1:200; clone 7C9, Biolegend, 347109), CD10-PECy7 (1:400; clone HI10a, Biolegend, 312213) and CXCR4-APC (1:3,200; clone 12G5, Biolegend, 306509). Cells were incubated for 15 min and washed with 3 ml of FACS buffer (PBS, 0.5% bovine serum albumin and 2 mM EDTA). Cells were centrifuged for 5 min at 800g, resuspended in 200 µl of FACS buffer and acquired on an Aurora three-laser spectral cytometer (Cytek) using Spectroflo (v1.1) software. Single-stain controls on whole blood were used for unmixing in addition to unstained fixed nasal cells. Exported FCS files were analyzed using Flowjo X (BD).

Peripheral blood flow cytometry. Peripheral blood samples were handled according to a standard sample processing protocol for flow cytometry (for a detailed protocol, see www.EuroFlow.org and ref. [34]). The combination of markers used for cell surface staining (Primary Immunodeficiency Orientation Tube (PIDOT)) has been designed along with a set of fully standardized methods and tools by EuroFlow consortium for the study of primary immunodeficiencies and was modified here by the addition of CD38 for a better separation of the circulating plasmablasts. Briefly, the procedure consisted of the bulk lysis of erythrocytes in up to 2 ml of fresh blood samples and surface staining of 2.5×10^6 white blood cells (100 µl final staining volume) with a reconstituted PIDOT lyophilized antibody cocktail (CYT-PIDOT, Cytognos) containing CD8-FITC (UCHT-4), IgD-FITC (IADB6), IgM-PERCPCy5.5 (clone MHM-88), CD16-PE (3G8), CD56-PE (C5.9), CD4-PerCPCy5.5 (RPA-T4), CD19-PeCy7 (IADB6), TCRγδ-PE-Cy7 (11F2), CD3-APC (IADB6), CD45-APC-C750 (HI30) and a drop-in antibody cocktail containing 2 µl of CD27-BV421 (clone M-T271, 562513, BDBiosciences), 2.5 µl of CD45RA-BV510 (clone HI100, 563031, BDBiosciences), 2 µl of CD38-BV605 (clone HIT-2, 303532, BioLegend) and 0.6 µl of pure CD38 (clone HIT-2, 303502, BioLegend) per test. At least 1 million events were acquired using a three-laser Cytek Aurora instrument (Cytek Biosciences; acquisition software, SpectroFlo, v1.1) from the FCF of LUMC in Leiden, the Netherlands (https://www.lumc.nl/research/facilities/fcf). The main circulating leukocyte subsets identified by flow cytometry were assigned manually using the data analysis software Infinicyt v2.04 (Cytognos SL) based on a standardized gating strategy[35] with an adaptation for the CD38 inclusion (Supplementary Fig. 1). The absolute counts per microliter of fresh blood were determined by a double platform approach using the absolute fresh leukocyte counts determined before sample processing with a hematological analyzer (Sysmex) to the Statistics Configure tool of the Infinicyt software.

CyTOF data preprocessing and clustering. An outline of data preprocessing steps is shown in Extended Data Fig. 10. Debris and normalization beads were filtered from FCS files using the CyTOFclean package (v1.0.1). Single cells were then manually gated based on DNA stain and the CATALYST package (v1.12.2), and single-stain controls were used to compensate data using the non-negative linear least squares method. One-by-one plots were used to confirm correct compensation of data. Epithelial and immune cells were manually gated based on CD45 and EpCAM expression, with exclusion of cPARP⁺ apoptotic cells and immune doublets (CD14⁺CD3⁺, CD66b⁺CD3⁺, CD14⁺CD66b⁺). Subsequently, the CATALYST package (v1.12.2) was used to debarcode immune and epithelial cells individually per batch. FCS files were then normalized using the reference PBMCs and the CyTOFBatchAdjust function with 99th percentile scaling for each marker individually[36]. The marker CD69 was not present in the reference PBMCs at sufficient levels to scale and was thus not normalized. Signal intensity and clustering of reference samples before and after normalization was used to verify appropriate normalization. Clustering of cells into populations was done using hierarchical SNE or t-SNE with Cytosplore software (v2.3.0; https://www.cytosplore.org/) using all markers except EpCAM and cPARP. All t-SNE analyses were performed with complexity = 30. A diffusion map of monocytes was created using the destiny package (v.3.2.0) with $k = 1,000$ using the following markers: HLA-DR, CD11c, CD163, ACE-2, CD45RO, CD14, CD38, CD127, CD206, CD86, CD4, CD123 and CD45RA.

SARS-CoV-2-specific T cell isolation. PBMCs from convalescent individuals with prior COVID-19 were isolated from fresh whole blood using Ficoll-Isopaque and were cryopreserved until further use. PBMCs were thawed, and 80% was used for overnight stimulation assays. The remaining 20% was immediately used for FACS based on pMHC tetramers. For the stimulation assay, ~8×10^6 PBMCs were seeded in 1 ml of IMDM (Lonza) containing 10% heat-inactivated fetal calf serum (FCS; Sigma), 2.7 mM L-glutamine (Lonza), 100 U ml⁻¹ penicillin (Lonza), 100 µg ml⁻¹ streptavidin (Lonza) and in the presence of 1 µg ml⁻¹ SARS-CoV-2 peptide pool or 1% DMSO (negative control). The SARS-CoV-2 peptide pool consisted of 15-mer peptides for all individuals and also consisted of HLA-matched 9-mer peptides

when the HLA type of the individual was known (see Supplementary Table 8). The 15-mer peptides were derived from nucleocapsid (Miltenyi, 130-126-699), membrane (Miltenyi, 130-126-703) and most immunogenic sequences from the spike protein (Miltenyi, 130-126-701). Peptides were dissolved and used according to manufacturer's protocol. Nine-mer SARS-CoV-2 peptides were predicted high binders for the 10 most common European ancestry HLA class 1 alleles and were kindly provided by P. Kvistborg (Netherlands Cancer Institute, Amsterdam, the Netherlands)[37]. See Supplementary Table 9 for a list of all peptides used. After 24 h, the PBMCs were washed and stained for CD4-FITC (1:30; BD, 555346, clone RPA-T4), CD8-PeCy7 (1:320; BD, 557746, clone RPA-T8), CD154-Pacific Blue (1:50; Biolegend, 310820, clone 24–31) and CD137-APC (1:100; BD, 550890, clone 4B4-1) in phenol red-free DMEM (Gibco), 2% FCS and penicillin/streptomycin for 30 min at 4 °C. Cells were sorted with a BD Aria III with BD FACSDiva software (v9.0) directly into lysis buffer and further processed as described below.

For pMHC tetramer FACS, PBMCs (2×10^6) were first incubated with a mix of in-house-produced[38], PE-labeled pMHC complexes for 30 min at 4 °C before adding APC-labeled CD8 (1:64; BD, 555369, clone RPA-T8) and fluorescein isothiocyanate (FITC)-labeled CD4 (1:30; BD, 555346) antibodies at 4 °C for 30 min. The pMHC tetramers used are shown in Supplementary Table 8. Tetramer-positive CD8$^+$CD4$^-$ T cells were sorted with a BD Aria III with BD FACSDiva software (v9.0) into 15-ml Falcon tubes and collected in lysis buffer.

TCR identification. TCRαβ sequences of T cell populations were identified as previously described with minor modifications[39]. Total RNA was isolated from 1×10^2 to 1×10^6 cells using the ReliaPrep RNA cell Miniprep system (Promega). The entire total RNA yield of each sample (10 μl) was mixed with 2 μl of anchored oligo(dT) primer (10 μM; Eurogentec) and 1.7 μl of SS2m_TSO primer (10 μM; Eurogentec), denatured at 72 °C for 3 min and immediately placed on ice afterward (see Supplementary Table 10 for a full list of primers used). To each sample, 7.8 μl of the first-strand reaction mix containing 0.9 μl of SMARTScribe RT enzyme (Takara, Clontech; 100 U μl^{-1}), 0.4 μl of RNAsin (Promega; 40 U μl^{-1}), 5 μl of 5× first-strand SMSRTScribe buffer (Takara, Clontech), 0.9 μl of DTT (100 mM; Invitrogen) and 1.7 μl of dNTPs (Promega; 10 mM) was added. The reverse transcription reaction was performed by incubating at 42 °C for 90 min followed by 10 cycles of 50 °C for 2 min and 42 °C for 2 min. Finally, the reverse transcriptase was inactivated by incubation at 70 °C for 15 min. Preamplification of the cDNA was performed on samples containing RNA from 500 or fewer cells; 12.5 μl of Phusion Flash (Thermo Fisher Scientific), 0.63 μl of IS PCR primer (10 μM; Eurogentec) and 1.87 μl of water (UltraPure water, Invitrogen) were added to 10 μl of the previously synthesized cDNA. The IS primer anneals both to the SA.rt and oligo(dT) IS region to amplify the complete cDNA[40]. The reaction was incubated at 95 °C for 2 min and cycled 18 times between 95 °C for 1 s, 69 °C for 15 s and 72 °C for 2 min, with a final extension at 72 °C for 5 min. Barcoded TCR PCR product was generated in two rounds of PCR. In the first PCR, *TRA* and *TRB* product was generated in separate PCR reactions for α- and β-chains as follows: 25 μl of Phusion Flash (Thermo Fisher Scientific), 1.25 μl of SS2m_For PCR primer (10 μM; Eurogentec), 1.25 μl of TRAC_rev or 1.25 μl of TRBC1+2 mix PCR primer (10 μM; Eurogentec) and 17.5 μl of water (UltraPure water, Invitrogen) added to either 5 μl of cDNA or 5 μl of preamplified cDNA. The reaction was incubated at 98 °C for 2 min and cycled 30 times between 98 °C for 1 s, 67 °C for 15 s and 72 °C for 15 s, with a final extension at 72 °C for 1 min. The PCR product was 96-well plate purified with the Wizard SV 96 PCR Clean-Up System (Promega) and eluted in 70 μl of water. In a second PCR, the first purified PCR product was used to include a two-sided six-nucleotide barcode sequence that allows for discrimination between TCRs of different T cell populations. Then, 20 μl of Phusion Flash (Thermo Fisher Scientific), 1.6 μl of BC_R7xx_For bc PCR primer (10 μM; Eurogentec), 1.6 μl of BC_R7xx_Rev bc PCR primer (10 μM; Eurogentec) and 6.8 μl of water (UltraPure water, Invitrogen) were added to 10 μl of purified PCR product. The reaction was incubated at 98 °C for 2 min and cycled 10 times between 98 °C for 1 s, 65 °C for 15 s and 72 °C for 30 s, with a final extension at 72 °C for 2 min. The barcoded PCR product was also purified with the Wizard SV 96 PCR Clean-Up System (Promega) and eluted in 70 μl of water. PCR products of different T cell populations were pooled, after which TCR sequences were identified by NovaSeq (GenomeScan). NovaSeq data were analyzed using MiXCR software (v3.0.13) to determine the Vα and Vβ family and CDR3 regions using annotation to the IMGT library (http://www.imgt.org; v6). CDR3 regions were analyzed in RStudio, and CDR3 sequences with ≤50 reads that were non-functional or occurred on all samples were excluded from the analysis.

Statistics. Statistical differences in cellular abundance between groups were compared with a linear mixed model in which individuals were included as random effect and groups (acute, ERS or all hospitalized individuals, convalescent and healthy donors) as fixed effect using the lme4 (v.1.1–23) and lmerTest (v3.1–2) packages. Post hoc comparison of all groups included in a linear mixed model was conducted with the emmeans package (v1.4.8) using the Tukey correction for multiple comparisons. For the comparison of multiple subsets or lineages, Benjamini–Hochberg multiple testing correction was subsequently used. Correlation values were calculated with Pearson tests. MDS was performed in R using the dist function in the stats package (4.0.1). Features were normalized before MDS with labels shown for selected clusters. If viral load data were not available for

the time of nasal cell sampling, the closest time point with available data was used. Viral load is calculated as $40 - C_t$ of SARS-CoV-2 PCR, meaning that higher values indicate more virus. MOFA was conducted using the MOFA package (v1.4.0) using standard parameters, with the exception of DropFactorThreshold of 0.02 and Tolerance of 0.01 (ref. [22]). Features were scaled before use in MOFA, and the model was repeated three times to qualitatively assess stability of LFs and defined features. All analyses were performed with R version 4.0.1, except for the cytofclean package performed in R3.6.3 using RStudio (v1.2.5033).

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

Data availability
Raw datasets (TCR sequencing, CyTOF and flow cytometry) have been deposited in the Zenodo repository and are accessible at https://doi.org/10.5281/zenodo.5691167. Researchers will submit their research idea or protocol to the contact person. After approval, data will be made available without restrictions. The reason for restricted access is that clinical data, despite anonymization, can sometimes be identifiable. No charges will be made.

Code availability
All packages, functions and key parameters used for analyses have been included the Methods section. Scripts used are deposited in GitHub at https://github.com/spjochems/COVID_nasal.

References
31. Hartmann, F. J., Simonds, E. F. & Bendall, S. C. A universal live cell barcoding-platform for multiplexed human single cell analysis. *Sci. Rep.* **8**, 10770 (2018).
32. Zunder, E. R. et al. Palladium-based mass tag cell barcoding with a doublet-filtering scheme and single-cell deconvolution algorithm. *Nat. Protoc.* **10**, 316–333 (2015).
33. Rahman, A. H., Tordesillas, L. & Berin, M. C. Heparin reduces nonspecific eosinophil staining artifacts in mass cytometry experiments. *Cytometry A* **89**, 601–607 (2016).
34. van der Velden, V. H. J. et al. Optimization and testing of dried antibody tube: the EuroFlow LST and PIDOT tubes as examples. *J. Immunol. Methods* **475**, 112287 (2019).
35. van der Burg, M. et al. The EuroFlow PID orientation tube for flow cytometric diagnostic screening of primary immunodeficiencies of the lymphoid system. *Front. Immunol.* **10**, 246 (2019).
36. Schuyler, R. P. et al. Minimizing batch effects in mass cytometry data. *Front. Immunol.* **10**, 2367 (2019).
37. Gangaev, A. et al. Identification and characterization of a SARS-CoV-2 specific CD8$^+$ T cell response with immunodominant features. *Nat. Commun.* **12**, 2593 (2021).
38. Burrows, S. R. et al. Peptide–MHC class I tetrameric complexes display exquisite ligand specificity. *J. Immunol.* **165**, 6229–6234 (2000).
39. van Bergen, C. A. et al. Selective graft-versus-leukemia depends on magnitude and diversity of the alloreactive T cell response. *J. Clin. Invest.* **127**, 517–529 (2017).
40. Picelli, S. et al. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* **10**, 1096–1098 (2013).

Author contributions
A.H.E.R., F.J.T.S., J.J.M.v.D., S.M.A., A.M.v.d.D., P.S.H., J.J.C.d.V., J.J.J., M.R., M. Kikkert, M.Y., M.H.M.H., H.H.S., S.P.J. and the BEAT-COVID study group designed the study. J.L.H.Z., M.V., C.P. and the BEAT-COVID study group collected samples. The BEAT-COVID study group and the COVID-19 LUMC group collected the data. M. König, T.D., T.T., S.A., Y.K. and S.K.M. processed the samples. C.R.P., M. König, W.H., R.S.H., M.Z. and F.J.d.B. conducted the experiments. C.R.P., W.H., T.T., R.S.H., M.Z., F.J.d.B., M.H.M.H. and S.P.J. analyzed the data. M.H.M.H., H.H.S. and S.P.J. supervised the project. All authors contributed to, read and approved 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/s41590-021-01095-w.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41590-021-01095-w.

Correspondence and requests for materials should be addressed to Simon P. Jochems.

Peer review information *Nature Immunology* thanks the anonymous reviewers for their contribution to the peer review of this work. Zoltan Fehervari was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team

Reprints and permissions information is available at www.nature.com/reprints.
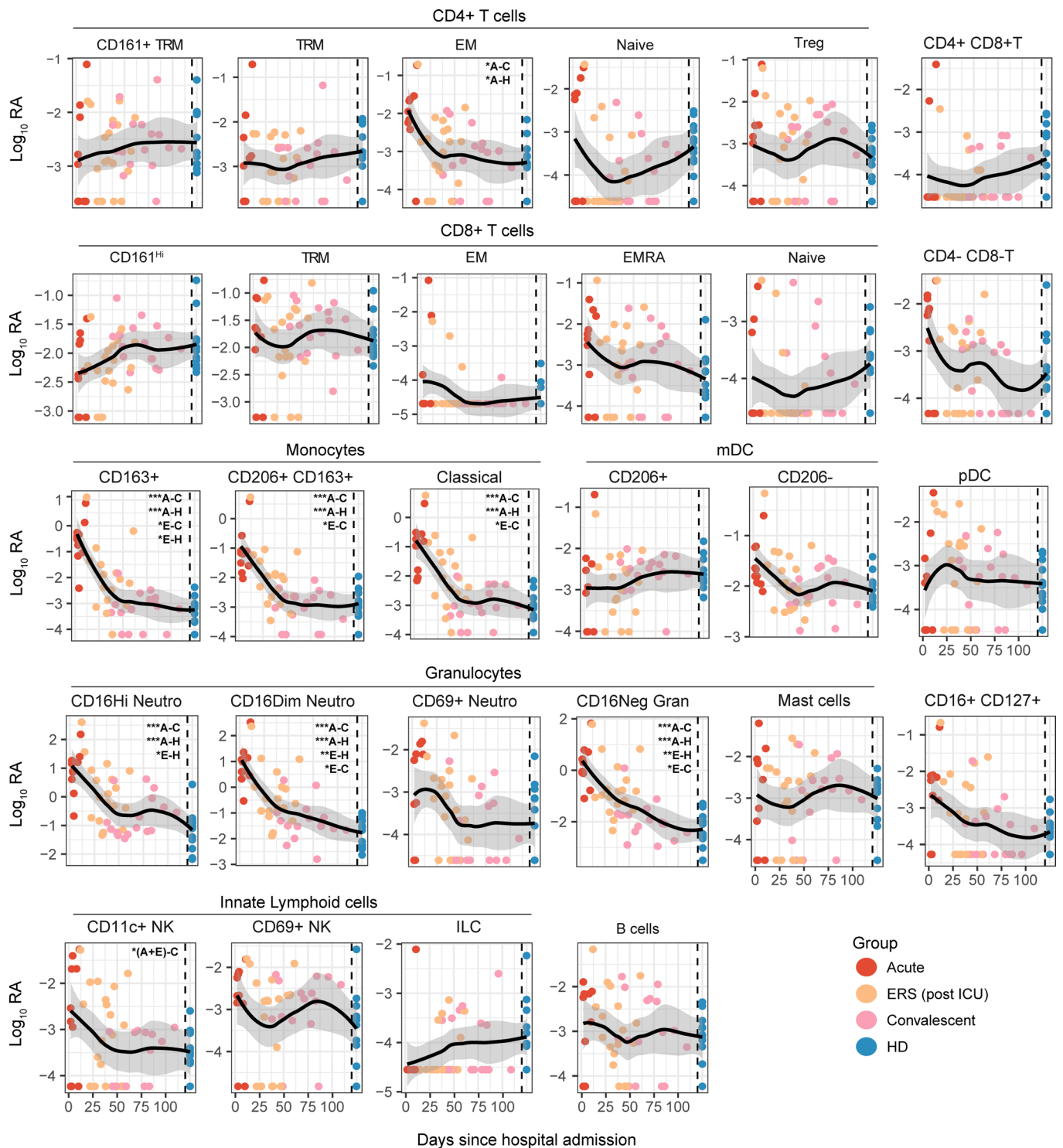
Extended Data Fig. 1 | Time plots of nasal immune subsets relative to hospital admission. Cell subset abundance during acute COVID-19 (red, 10 samples from 9 patients), during the early recovery phase (ERS, post ICU but still in hospital, orange, 18 samples from 11 patients), or in COVID-19 convalescence (5–6 weeks post hospital discharge, pink, n=16) or healthy controls (blue, n=12). Samples are plotted against day of hospital admission, with healthy donors plotted at the right axis separated by a dashed line. *p < 0.05, **p < 0.01, ***p < 0.001, by linear-mixed model with group as fixed effect and individuals as random effect, with post-hoc testing and Tukey multiple testing correction followed by Benjamini-Hochberg correction for comparing multiple subsets. See Supplementary tables 3 and 4 for exact test results. Black lines and shaded areas represent Loess regression results and 95%-confidence intervals, respectively. A = acute (hospitalized), E = early recovery stage, C = convalescent, 5–6 weeks post discharge, H = healthy donor. NK = natural killer cells. ILC = innate lymphoid cells, mDC = myeloid dendritic cells, pDC = plasmacytoid dendritic cells, Neutro = neutrophils, Trm = Tissue-resident memory, EM = effector memory, EMRA = effector memory re-expressing CD45RA. Treg = regulatory T cells.
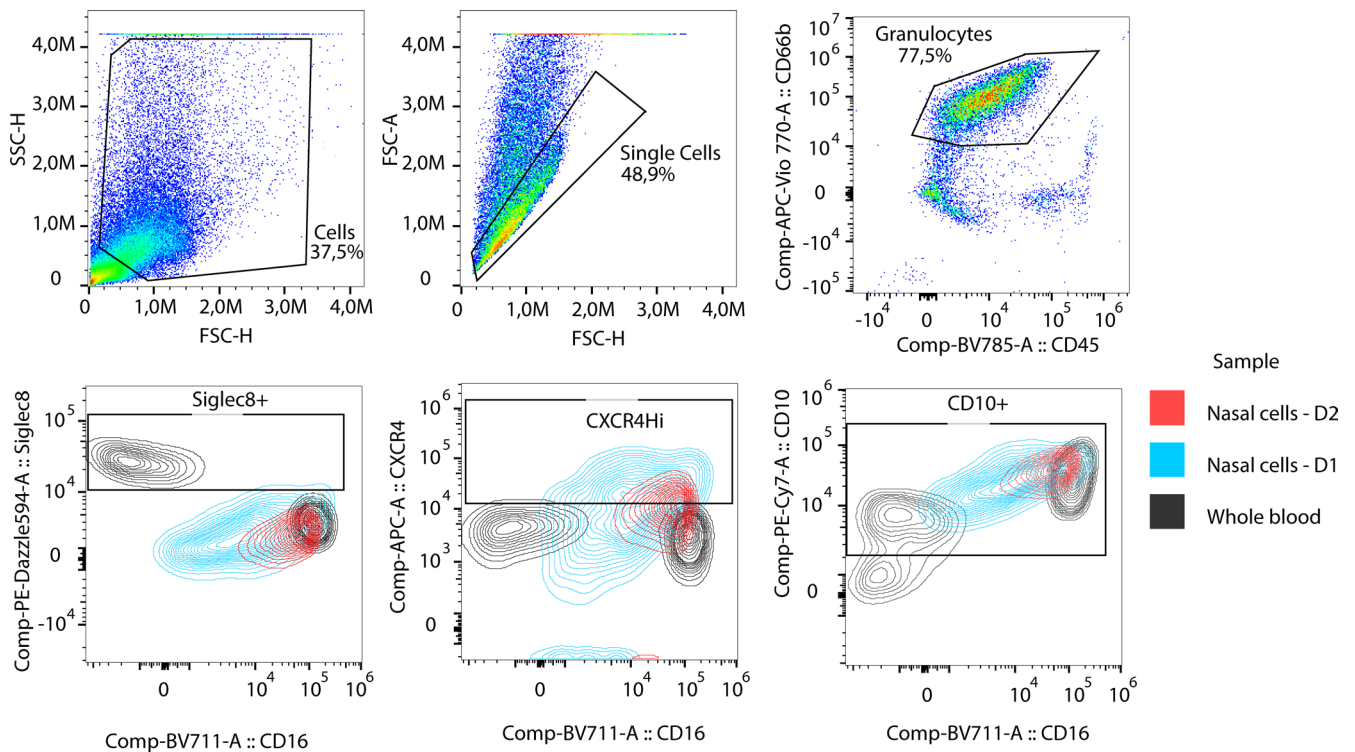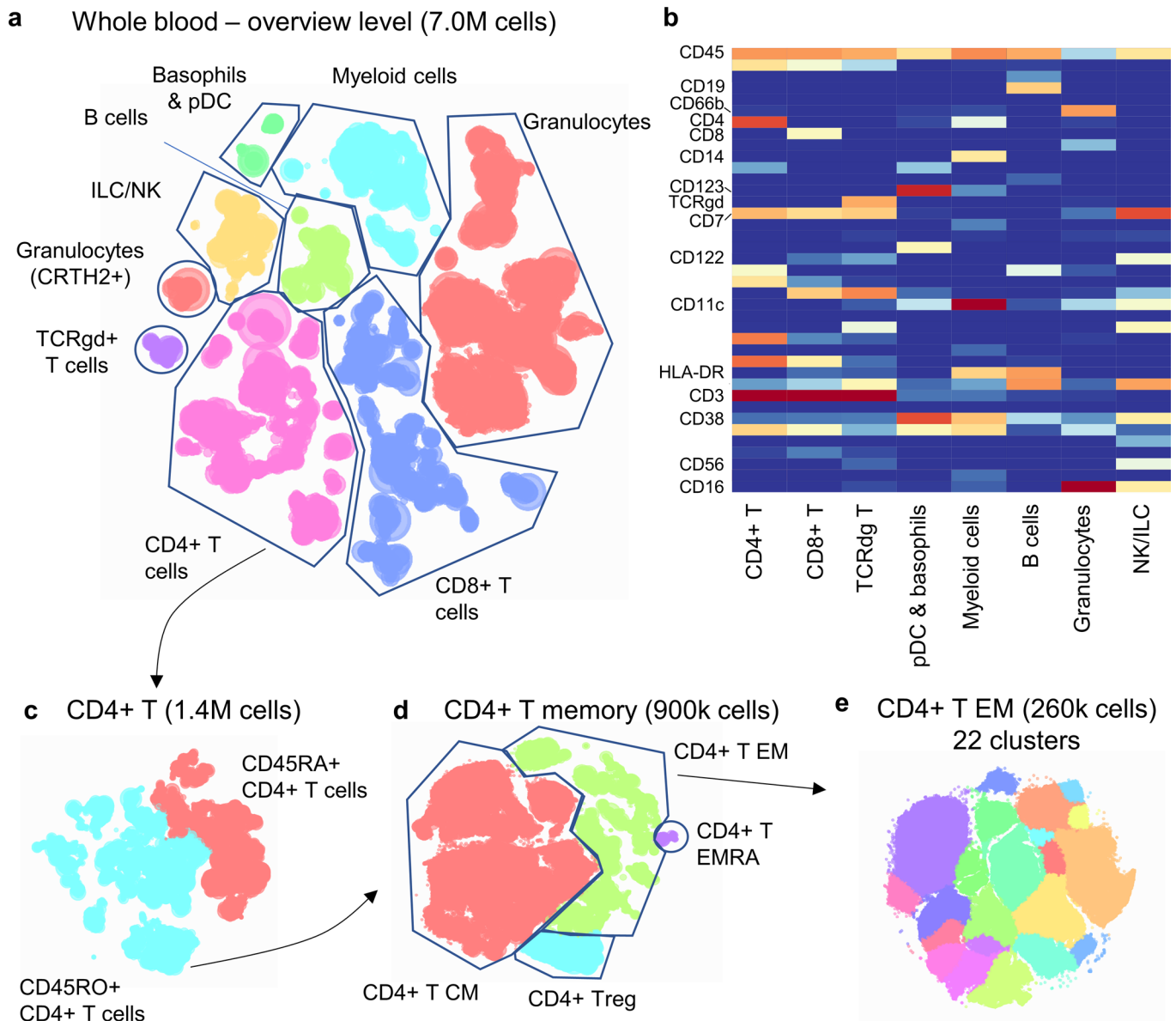
Extended Data Fig. 2 | Flow cytometric analysis of nasal granulocytes. Granulocytes in whole blood from one donor and fixed nasal scrapes from 2 individuals with or without COVID-19 were analysed using additional markers. The top row shows gating strategy to select granulocytes for one of the nasal cell samples. The bottom row shows contour plots of CD16 expression versus Siglec-8, CXCR4 and CD10 expression. Colours correspond to different samples, with black depicting the whole blood.
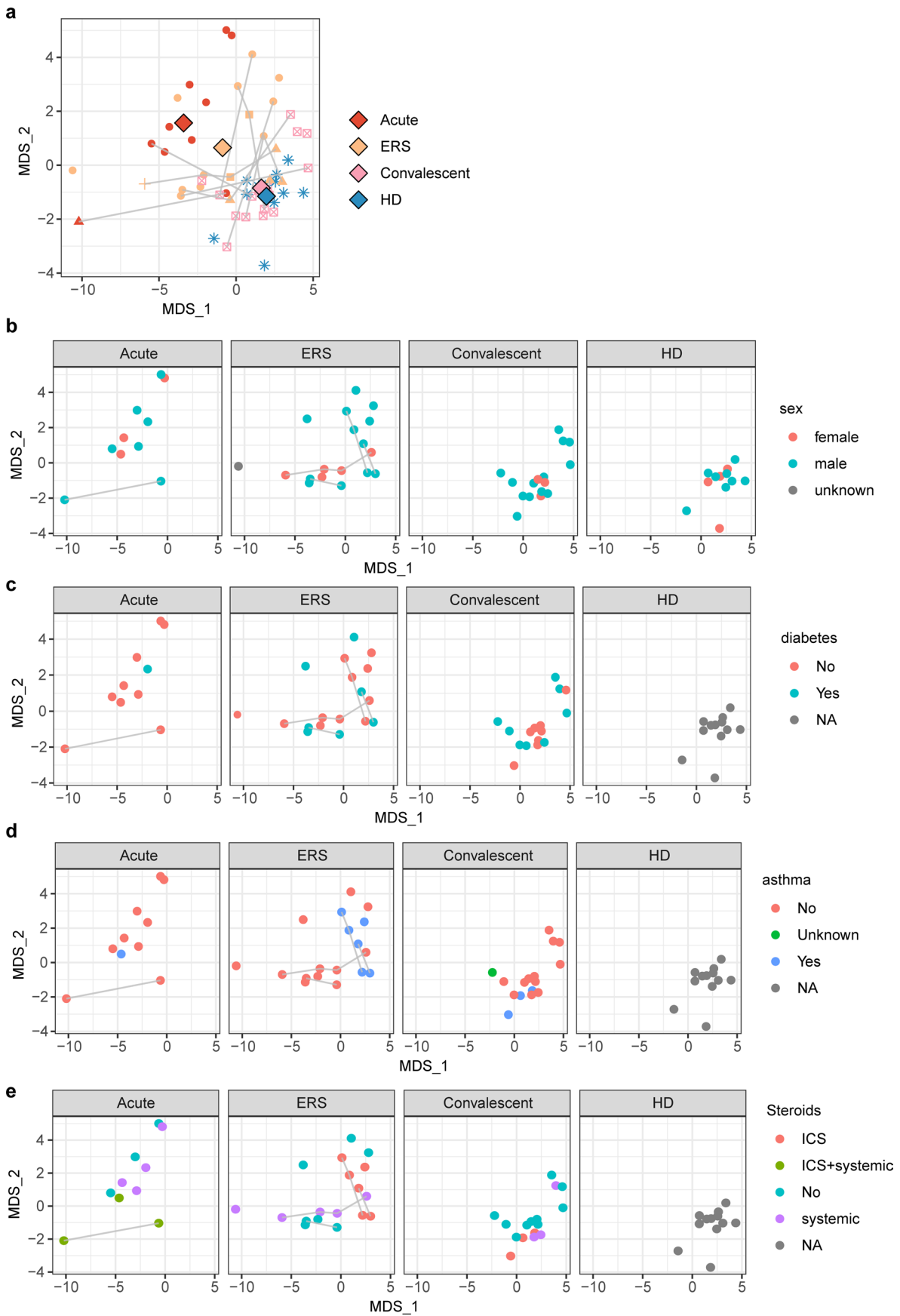
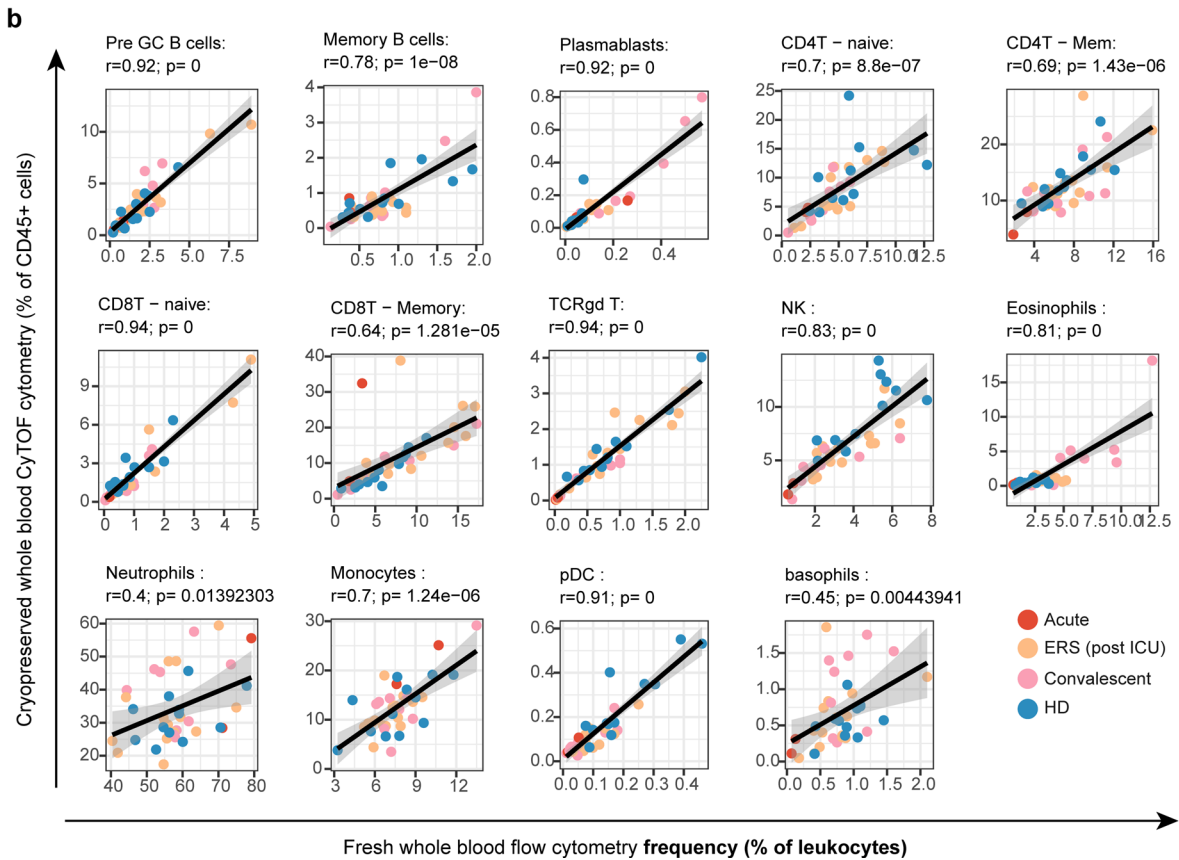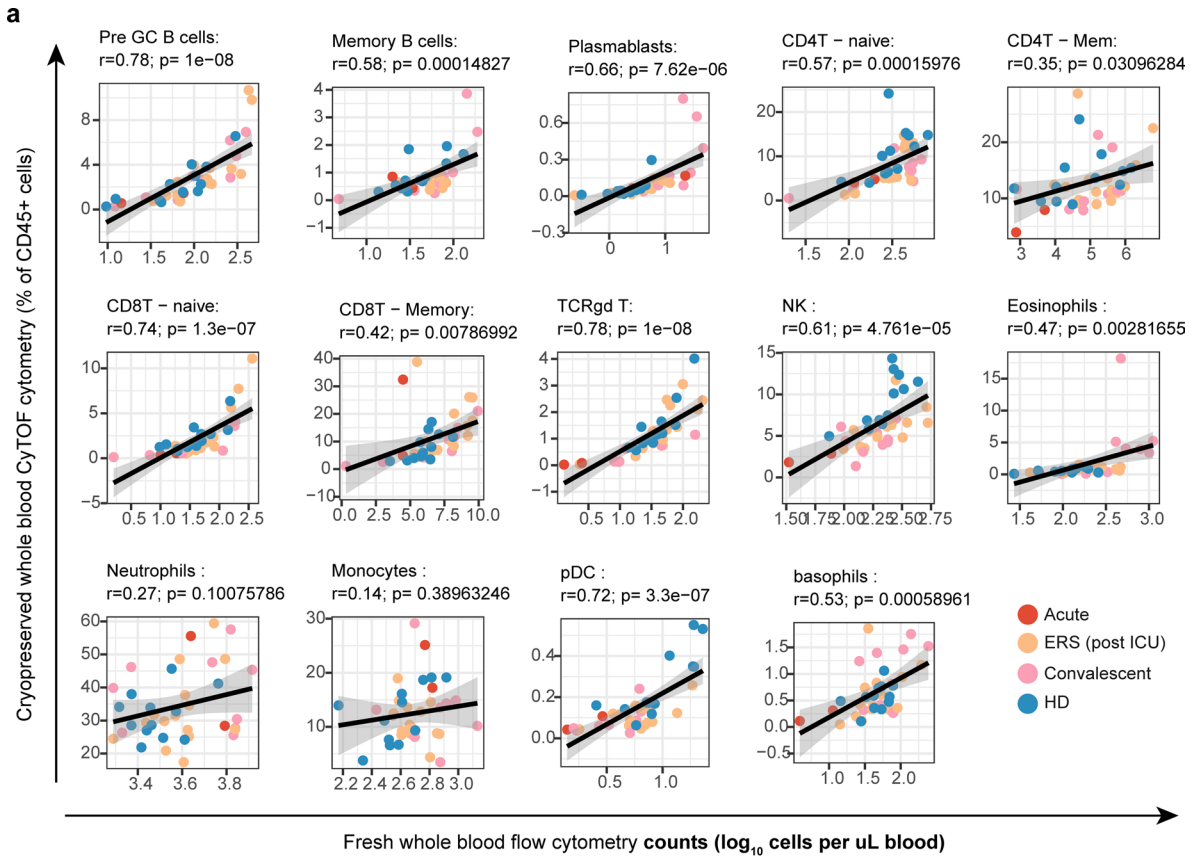**Extended Data Fig. 3 | See next page for caption.**

Extended Data Fig. 3 | Multi-dimensional scaling of nasal immune subsets. a) Clustering of all nasal samples based on abundance of all cellular subsets, colored by group. Acute COVID-19 (red, n=9), during the early recovery phase (ERS, post ICU, orange, n=11), or in COVID-19 convalescence (5–6 weeks post hospital discharge, pink, n=16) or healthy controls (blue, n=12). Individual samples are shown with repeated samples per donor connected by lines. Centroids are shown per group as large diamonds. Per group faceted representation of the same scaling colored by covariates **b**) sex, **c**) diabetes, **d**) asthma, **e**) steroid usage. For these only repeated sample per group are shown and not across groups, for example from ERS to convalescence.

Extended Data Fig. 4 | Whole blood CyTOF and flow cytometry analysis strategy. Pre-processed CyTOF files were clustered using hierarchical SNE (hSNE) and Gaussian mean shift clustering in Cytosplore software. **a**) Landmark clustering of overview level, with color corresponding to clusters. **b**) Heatmap showing marker expression, used for annotation of the lineages. **c**) Clustering of all CD4 + T cells in level 2, with landmarks colored by cluster. **d**) Clustering of memory CD4 + T cells, with colors corresponding to clusters in level 3. **e**) Clustering of CD4 + T EM cells on the data level, with cells colored per cluster. pDC = plasmacytoid dendritic cell. ILC = innate lymphoid cells. NK = natural killer. CM = central memory. EM = effector memory. EMRA = effector memory re-expressing CD45RA.
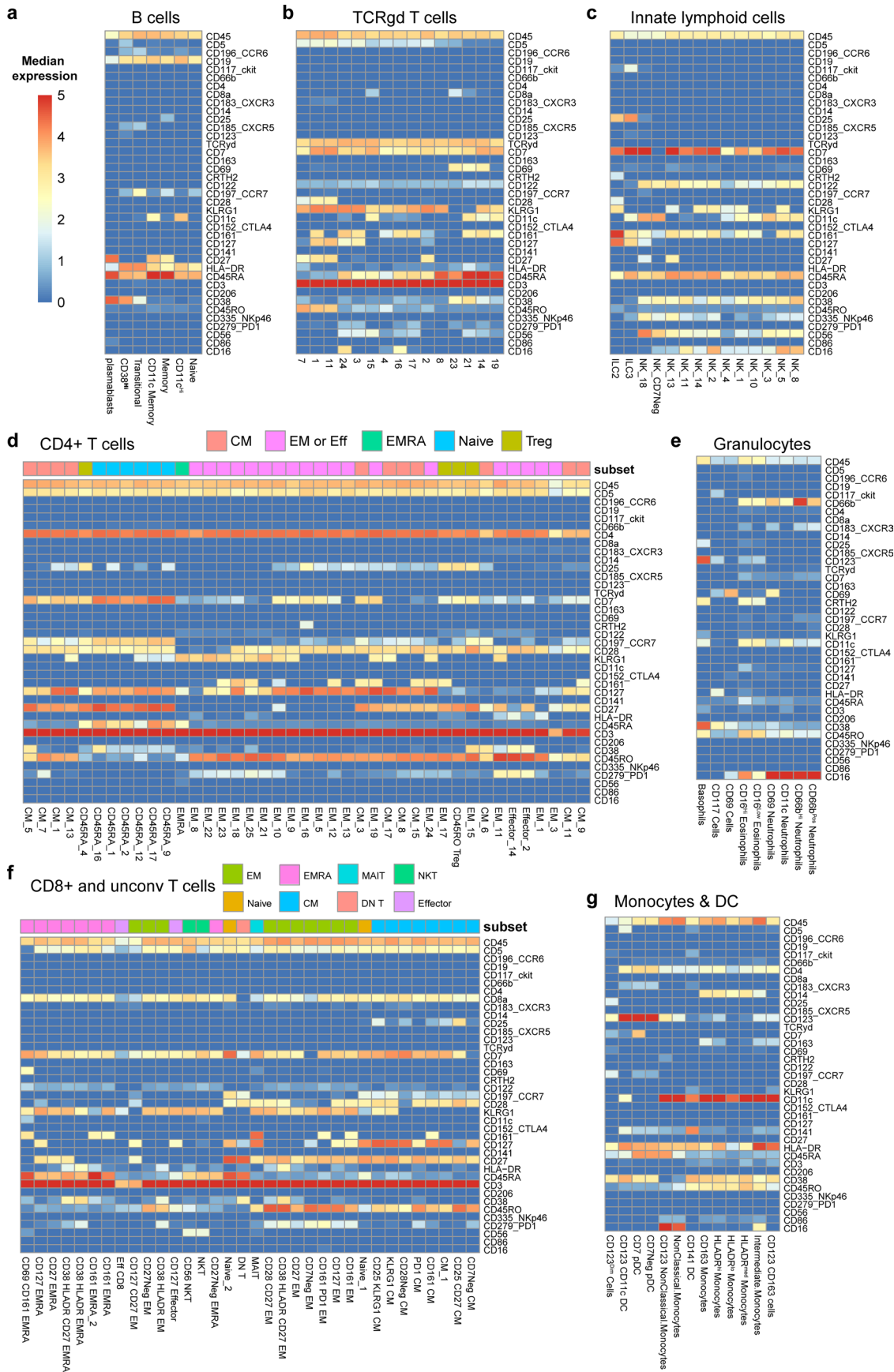
Extended Data Fig. 5 | Phenotypes of whole blood CyTOF immune clusters. Heatmaps show the median expression for each of the 133 defined clusters from whole blood for the included markers. Clusters are shown together per lineage.

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

Extended Data Fig. 6 | Correlations between cryopreserved whole blood analysed by CyTOF and fresh whole blood analysed by flow cytometry on a subset level. Individuals with paired data (n = 43) are shown, with colours indicating disease status. Black lines indicate results from linear regression analyses, with shaded areas corresponding to 95%-confidence intervals. Rho (r) and p-value from Pearson correlation analysis depicted above graphs. For CyTOF data, frequencies of subsets among all CD45 + cells are shown on the y-axis. **a**) shows correlation with fresh whole blood data when using absolute counts of the whole blood data, while **b**) shows correlations when using frequencies of subsets among leukocytes of whole blood data.
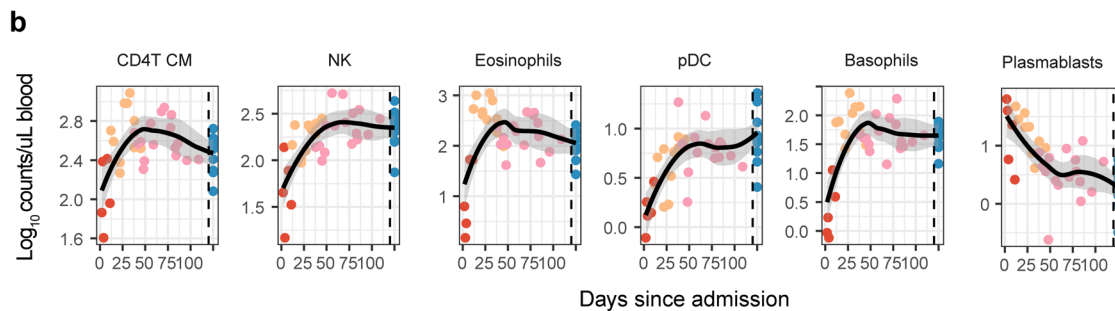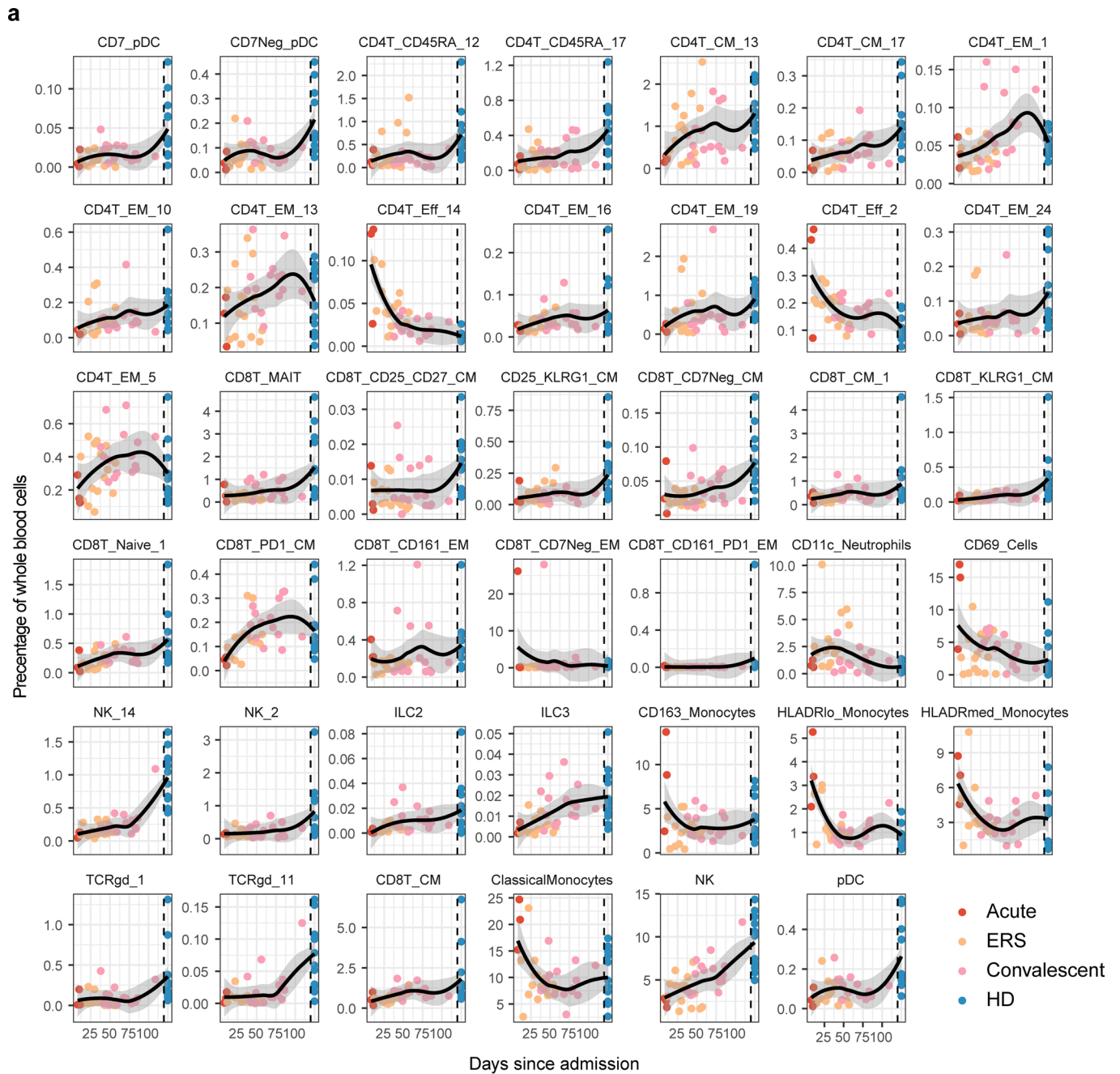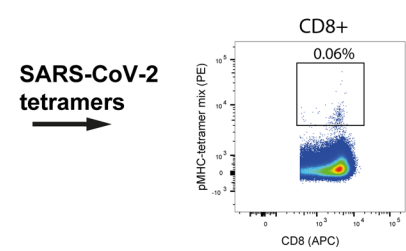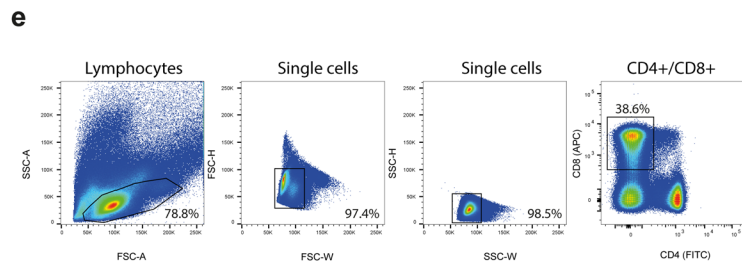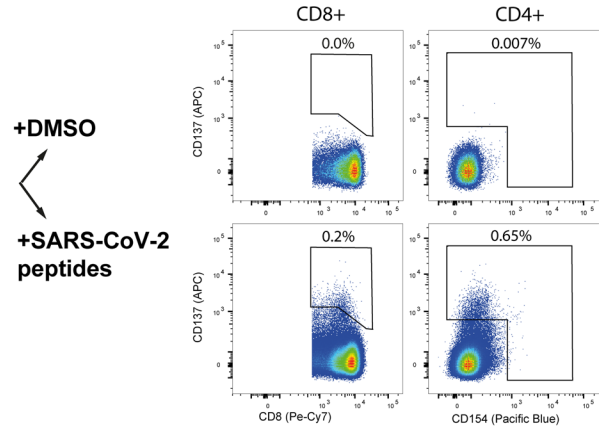
Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Time plots of whole blood immune cell clusters and subsets. a) CyTOF cluster and subset frequencies during acute COVID-19 (red, n = 3), during the early recovery phase (ERS, post ICU but still in hospital, orange, 15 samples from 11 patients), or in COVID-19 convalescence (5–6 weeks post hospital discharge, pink, n = 15) or healthy controls (blue, n = 12). **b**) fresh whole blood flow subset counts during acute COVID-19 (red, 5 samples from 4 patients), during the early recovery phase (ERS, post ICU but still in hospital, orange, 10 samples from 7 patients), or in COVID-19 convalescence (5–6 weeks post hospital discharge, pink, n = 16) or healthy controls (blue, n = 12). Black lines and shaded areas represent Loess regression results and 95%-confidence intervals, respectively. Only clusters and subsets are shown for which $p < 0.05$ by linear-mixed model with group as fixed effect and individuals as random effect, with post-hoc testing and Tukey multiple testing correction followed by Benjamini-Hochberg correction for comparing multiple subsets. NK = natural killer cells. ILC = innate lymphoid cells, pDC = plasmacytoid dendritic cells, CM = central memory. EM = effector memory. Eff = effector.
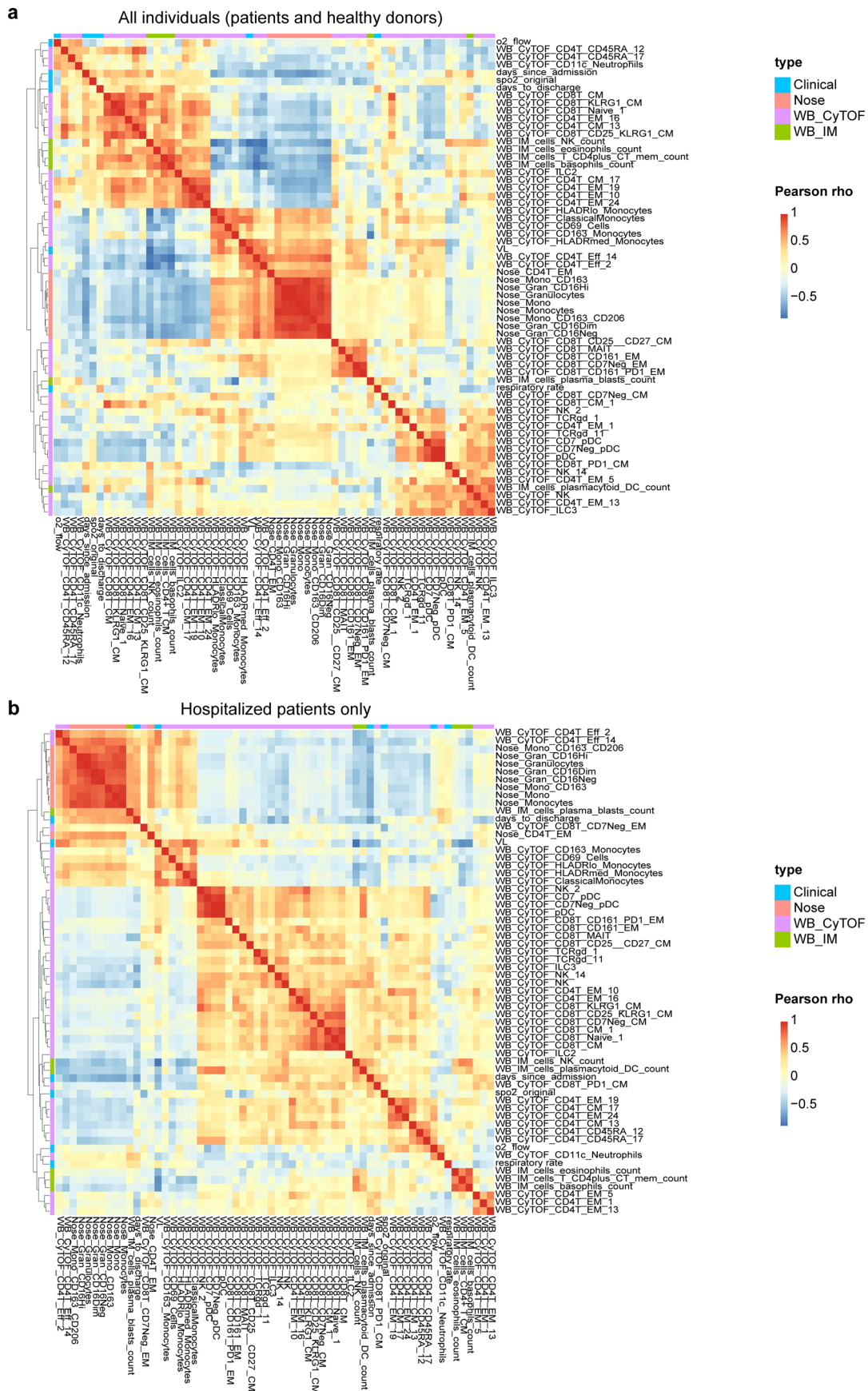
Extended Data Fig. 8 | Correlation heatmaps including all significant immune data and clinical features. a) Heatmap showing the pairwise correlation (Pearson Rho) between significant clusters and subsets from nasal CyTOF, cryopreserved whole blood (WB) CyTOF, fresh WB Flow cytometry, and clinical characteristics. Data from all patients and healthy donors were included in analysis. **b)** Correlation heatmap with only data from hospitalized patients included.

a Nasal CD4+ T EM - 599 cells
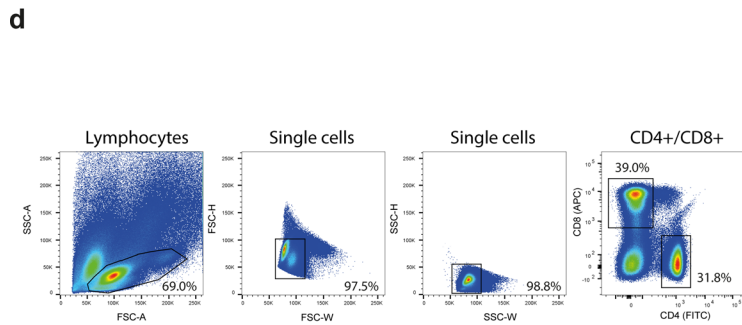

b

c

d

e
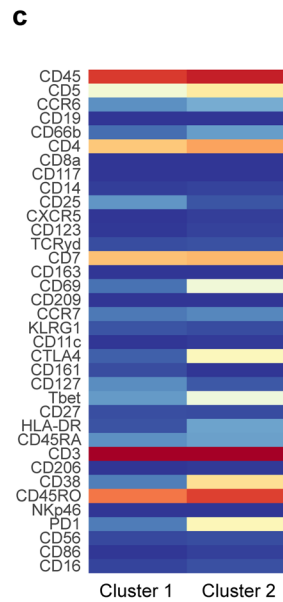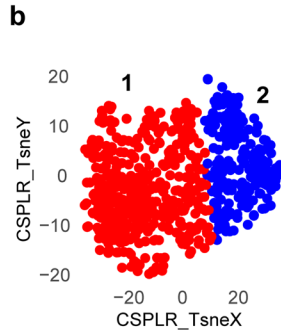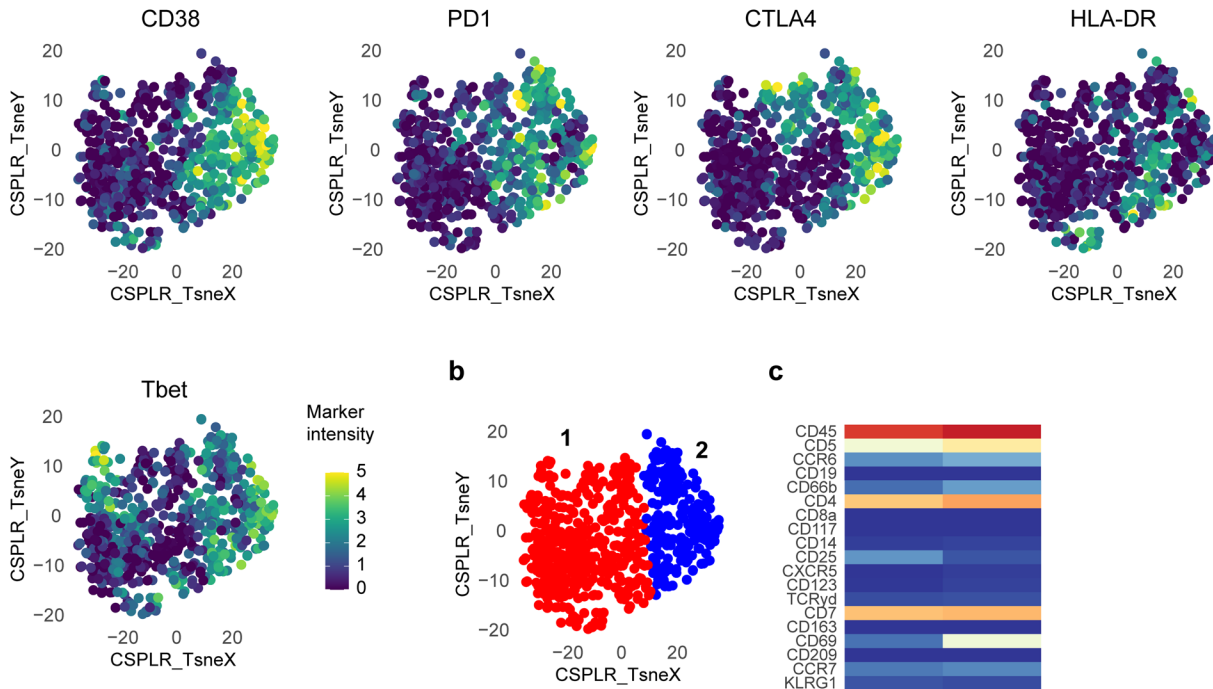
Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Nasal CD4 + T effector memory (EM) phenotype and SARS-CoV-2 specific T cells. a) tSNE analysis of CD4 + T EM cells. Expression of CD38, PD1, CTLA4, HLA-DR or Tbet overlaid onto tSNE embedding. **b**) Clustering of tSNE using Gaussian mean shift, with clusters overlaid onto the embedding. Clusters numbers are indicated. **c**) Heatmap showing median expression for all markers per CD4 + T EM cluster. **d**) Activation induced marker assay strategy. Sequential gating strategy with gates used to sort specific T cells after overnight stimulation with overlapping pool of SARS-CoV-2 peptides. Upregulation of CD137 and/or CD154 after stimulation was used as an indication of specific cells. The frequency per population is shown, and DMSO negative control is shown to indicate background levels of activation **e**) Tetramer sorting strategy. MHC Class 1 tetramers loaded with SARS-CoV-2 peptides were incubated with PBMC in a complementary strategy to identify SARS-CoV-2 specific cells. Gates for sorting, and cell frequencies are shown. HLA class of tetramers were matched to participant HLA type and a pool of immunodominant peptides was used to identify specific cells.

Bead-normalized fcs files

Clean up (Gaussian parameter, beads, DNA)
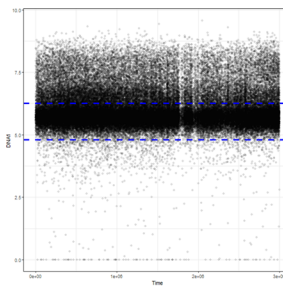
Compensation (Catalyst)
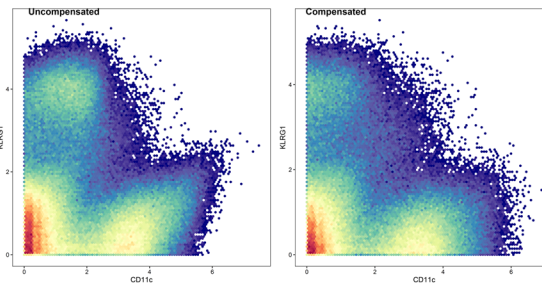
Bi-axial plot of 2 markers before and after compensation

Removal cPARP+ apoptotic cells, split immune cells (CD45+) and epithelial cells (EpCAM+), removal of doublets (CD66b+CD14+, CD66b+CD3+, CD3+CD14+)
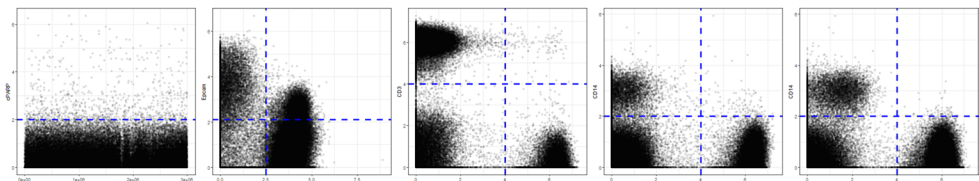
Debarcoding (Catalyst)
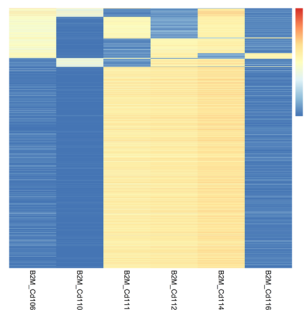
Heatmap of marker expression per barcode

Batch correction using reference sample (BatchAdjust)

Pre batch adjust   Post batch adjust
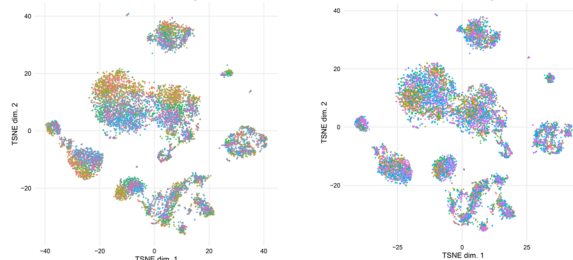
tSNE plots of reference sample before and after batch correction (colours indicate batches)

Hierarchical SNE clustering

Extended Data Fig. 10 | CyTOF pre-processing pipeline. Sequential steps prior to data clustering are shown.

nature research

Corresponding author(s): S.P. Jochems

Last updated by author(s): Nov 14, 2021

Reporting Summary

Nature Research 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 Research 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

CyTOF data were collected on a Helios with manufacturer's software (CyTOF Software, v7.0.8493). Flow cytometry data were acquired on an Aurora flow cytometer with manufacturer's software (SpectroFlo v1.1). FACSsorting was performed on an Aria III (BD) with manufacturer's software (BD Facsdiva v9.0).

Data analysis

Debris and normalization beads were filtered from .fcs files using the 'Cytofclean' package (v1.0.1). Single cells were then manually gated based on DNA stain and the 'CATALYST' package (v1.12.2) and single-stain controls were used to compensate data using the non-negative linear least squares method. One by one plots were used to confirm correct compensation of data. 'CATALYST' package (v1.12.2) was used to debarcode immune and epithelial cells individually per batch. FCS files were then normalized using the reference PBMCs and the CyTOFBatchAdjust function (https://github.com/CUHIMSR/CytofBatchAdjust) with 99 percentile scaling for each marker individually. Clustering of cells into populations was done using hierarchical stochastic neigbour embedding (hSNE) or tSNE with Cytosplore software (v2.3.0, https://www.cytosplore.org/). NovaSeq data were analysed using MiXCR software (v3.0.13) to determine the Vβ family and CDR3 regions Flow cytometry analysis was performed using Flowjo (V10) or using Infinicyt (v2.04, Cytognos SL) software with manual gating. Statistical analysis was done in R (v3.6.3 or v4.0.1), using RStudio (v1.2.5033). Packages used were 'lmerTest' (v3.1-2), 'emmeans' (v1.4.8), 'lme4' (v.1.1-23), 'MOFA' (v1.4.0), 'stats' (4.0.1), 'destiny' (v.3.2.0)

Scripts used have been deposited in github: https://github.com/spjochems/COVID_nasal

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 Research guidelines for submitting code & software for further information.

Data

Raw datasets (TCR sequencing, CyTOF and flow cytometry) have been deposited in the Zenodo repository, and are accessible under doi: 10.5281/zenodo.5691167. This has been included in data availability statement

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 given the exploratory nature of the study, number of individuals included was determined by logistical restrictions, in particular the availability of personnel and laboratory spaces. We analysed 44 patient samples from 29 patients, with 1-4 samples per patient, and 12 healthy donors, which was the maximum we were able to obtain. Based on literature this sample size is of sufficient power to identify relevant differences in mucosal immune responses using cytometric analysis
Data exclusions	Nasal scrape samples from individuals with <100 immune or epithelial cells were excluded as they indicate a 'failed' scrape. This often related to the inability to visualize the inferior turbinate well.
Replication	We analysed the nasal CyTOF data in 5 batches/experiments. These all included a reference sample to be able to compare between batches, allowing to normalize and demonstrate no batch-specific clustering. Blood CyTOF was performed in 2 batches with a similar approach. We included independent individuals (n=29 patients) and 12 healthy donors. with up to 1-4 samples per patient for a total of 44 patient samples. Different datasets also independently verified each other, for example there was very high correlation between whole blood samples analysed by CyTOF and flow cytometry on a subset level (see Supplementary figure 6).
Randomization	randomization was not applicable as this was not an intervention study
Blinding	Clustering of cells was performed including all cells, leading to a de facto blinded assignment of cells to populations.

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
☐	☒ Human research participants
☐	☒ Clinical data
☒	☐ Dual use research of concern

Methods

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

Antibodies

Antibodies used	Flow antibodies: CD4-FITC (BD, cat#555346), CD8-PeCy7 (BD, cat#557746), CD154-Pacific Blue (Biolegend, cat#310820) and CD137-APC (BD, cat#550890). CD45-BV785 (clone HI30, 1:200, Biolegend, cat# 304047), CD66b-APC-Vio770 (clone REA306, Miltenyi, 1:800, cat#130-120-146), CD127-PercpCy5.5 (clone A019D5, Biolegend, cat# 351321), CD16-BV711 (clone 3G8, Biolegend, cat# 302043), Siglec8-PEDazzle594 (clone 7C9, Biolegend, cat# 347109), CD10-PECy7 (clone HI10a, Biolegend, cat# 312213) CXCR4-APC (clone 12G5, Biolegend, cat# 306509). CD8-APC (BD, cat#555369, clone RPA-T8). CD27-BV421 (clone M-T271, 562513, BDBiosciences),

CD45RA-BV510 (clone HI100, 563031, BDBiosciences), CD38-BV605 (clone HIT-2, 303532, BioLegend), CD38-pure (clone HIT-2, 303502, BioLegend)

PIDOT lyophilized antibody cocktail (CYT-PIDOT, Cytognos, Spain) contains CD8-FITC (UCHT-4), IgD-FITC (IADB6), IgM-PerCPCy5.5 (clone MHM-88), CD16-PE (3G8), CD56-PE (C5.9), CD4-PerCPCy5.5 (RPA-T4), CD19-PeCy7 (IADB6), TCRgd-PE-Cy7 (11F2), CD3-APC (IADB6), CD45-APC-C750 (HI30)

For CYTOF:
Label Specificity Clone Vendor CatalogueID SampleType(nose/blood) Dilution
89Y CD45 HI30 Fluidigm 3089003B N / B 1/200
115In CD5 UCHT2 BioLegend 300602 N / B 1/100
141Pr CD196 (CCR6) G034E3 Fluidigm 3141003A N / B 1/200
142Nd CD19 HIB19 Fluidigm 3142001B N / B 1/200
143Nd cPARP F21-852 Fluidigm 3143011A N 1/100
143Nd CD117 (c-Kit) 104D2 BioLegend 313202 B 1/100
144Nd CD66b REA306 Miltenyi Biotech 130-108-019 N / B 1/50
145Nd CD4 RPA-T4 BioLegend
300541 N / B 1/100
146Nd CD8a RPA-T8 BioLegend
301053 N / B 1/200
147Sm CD117 (C-kit) 104D2 BioLegend 313202 N 1/100
147Sm CD183 (CXCR3) G025H7 BioLegend 353733 B 1/100
148Nd CD14 M5E2 BioLegend
301843 N / B 1/100
149Sm CD25 (IL-2Ra) 2A3 Fluidigm 3149010B N / B 1/100
150Nd CD185 (CXCR5) J252D4 BioLegend 356902 N / B 1/100
151Eu CD123 6H6 Fluidigm 3151001B N / B 1/100
152Sm TCRγδ 11F2 Fluidigm 3152008B N / B 1/50
153Eu CD7 CD7-6B7 BioLegend
343111 N / B 1/100
154Sm CD163 GHI/61 BioLegend
333602 N / B 1/100
155Gd CD69 FN50 BioLegend 310939 N / B 1/200
156Gd CD294 (CRTH2) BM16 BioLegend 350102 B 1/100
157Gd ACE-2 AC18F Novus Biologicals NBP2-80035-100UG N 1/100
158Gd CD122 (IL-2Rb) TU27 BioLegend 339015 B 1/100
158Gd CD209 9E9A8 BioLegend
330102 N 1/100
159Tb CD197 (CCR7) G043H7 BioLegend
353237 N / B 1/200
160Dy CD28 CD28.2 Fluidigm 3160003B B 1/200
160Gd Epcam AUA1 Invitrogen MA5-13917 N 1/50
161Dy KLRG1 (MAFA) REA261 Miltenyi Biotech Special order N / B 1/100
162Dy CD11c Bu15 BioLegend
337221 N / B 1/200
163Dy CD152 (CTLA-4) BNI3 BioLegend 369602 N / B 1/100
164Dy CD161 HP-3G10 BioLegend
339919 N / B 1/100
165Ho CD127 (IL-7Ra) AO19D5 Fluidigm 3165008B N / B 1/200
166Er Tbet 4B10 BioLegend 644825 N 1/100
166Er CD141 1A4 BD Biosciences 559780 B 1/100
167Er CD27 O323 BioLegend
302839 N / B 1/200
168Er HLA-DR L243 BioLegend
307651 N / B 1/200
169Tm CD45RA HI100 Fluidigm 3169008B N / B 1/100
170Er CD3 UCHT1 BioLegend
300443 N / B 1/100
171Yb CD206 15-2 BioLegend
321127 N / B 1/200
172Yb CD38 HIT2 BioLegend
303535 N / B 1/200
173Yb CD45RO UCHL1 BioLegend
304239 N / B 1/100
174Yb CD335 (NKp46) 92E BioLegend
331902 N / B 1/100
175Yb PD-1 EH12.2H7 BioLegend
329941 N / B 1/100
176Yb CD56 B159/NCAM16.2 BD Biosciences 555514 N / B 1/100
198Pt CD86 IT2.2 BioLegend 305435 N / B 1/200
209BI CD16 3G8 Fluidigm 3209002B N / B 1/200
106Cd B2M 2M2 BioLegend 316302 N / B 1/50
110Cd B2M 2M2 BioLegend 316302 N / B 1/50
111Cd B2M 2M2 BioLegend 316302 N / B 1/50

112Cd B2M 2M2 BioLegend 316302 N / B 1/50
114Cd B2M 2M2 BioLegend 316302 N / B 1/50
116Cd B2M 2M2 BioLegend 316302 N / B 1/50

Validation	Commercial antibodies were used. Antibodies were tested against fixed and unfixed cells in preliminary experiments using whole blood, PBMC or nasal cells (depending on target) to validate their ability to stain fixed material. All antibodies were titrated in-house per assay to obtain maximum specific signal while minimizing background using PBMC, nasal cells or WB, depending on target used. All antibodies were obtained from commercial sources. In-house conjugations to metals of primary antibodies were validated by staining PBMC, whole blood, nasal cells depending on the target. The following statements are made by the manufacturers: Fluidigm: Each lot of conjugated antibody is quality control tested by CyTOF® analysis of stained cells using the appropriate positive and negative cell staining and/or activation controls by Fluidigm Biolegend: Each lot product is validated by QC testing with a series of titration dilutions by BioLegend. BD: validated for flow cytometry (routinely tested) and QC tested Invitrogen (Epcam, AUA1): This Antibody was verified by Relative expression to ensure that the antibody binds to the antigen stated. Novus biologicals: Flow Cytometry: ACE-2 Antibody (AC18F) [NBP2-80035] - HepG2 cells were stained significantly using anti-ACE2 (human), mAb (AC18F). Miltenyi: CD66b (REA306) Extended validation for CD66b Antibody, anti-human, REAfinity™ based on specificity and sensitivity. Performance comparison, Selected fluorochrome conjugated antibodies from Miltenyi Biotec were compared to commercially available hybridoma clones in flow cytometry analysis.

Human research participants

Policy information about studies involving human research participants

Population characteristics	In this prospective observational cohort study, adult patients with PCR-confirmed COVID-19 who were admitted to our academic hospital were recruited. All hospitalized patients had hypoxia. The study was performed at the Leiden University Medical Center from patients included from April 2020 until December 2020. Common co-morbidities, disease outcome and treatment are given in supplementary table 2. Twelve healthy donors, age and sex-matched with patients were also included. Patients (n=29) Healthy donors (n=12) Female sex (%) 9 (31.0) 4 (33.3) Median age (range) 62 (19-78) 64 (60-72) Median BMI (min-max) 28.1 (17.6-42.3) ND
Recruitment	In this prospective observational cohort study, adult patients with PCR-confirmed COVID-19 who were admitted to our academic hospital were recruited. The study was performed at the Leiden University Medical Center from patients included from April 2020 until December 2020. After informed consent was obtained, longitudinal sampling was performed for the duration of the hospital admission, and one convalescent sample was obtained at the outpatient follow-up appointment, which was scheduled six weeks after hospital discharge. Selection bias could be introduced due to the fact that mostly individuals from Caucasian descend provided informed consent (and also constitute the majority of the population in Leiden and surrounding area), making it impossible to assess the role of genetics or environmental influences. It also affects the ability to extend conclusions to other geographical regions.
Ethics oversight	Ethical approval was obtained from the Medical Ethical Committee Leiden-Den Haag-Delft (NL73740.058.20). The trial was registered in the Dutch Trial Registry (NL8589). Written informed consent was obtained for all study participants. Study participants did not receive compensation for taking part in the study

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

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	https://www.trialregister.nl/trial/8589
Study protocol	The study protocol is available upon request from authors
Data collection	The study was performed at the Leiden University Medical Center from patients included from April 2020 until December 2020. Clinical data was collected using routine diagnostics, and entered into an Opal data warehouse. Data was collected in real-time. Immunological data was collected either in real-time (FACS on fresh blood, Cytek Aurora) or retrospectively (frozen samples for CyTOF (Helios), flow cytometry (Aurora), sorting (Aria)). Nasal CyTOF data was collected between June 2020 and Januari 2021. Whole blood CyTOF data was collected in June 2021.
Outcomes	The outcomes reported in this manuscript were a priori defined in the study protocol. The primary and secondary endpoints defined

| Outcomes | in the first version of the protocol, can be found in the clinical trial registration provided above, and were:

Primary outcome
Biomarker profiles or signature which correlate with future clinical progression of patients infected with SARS-CoV-2 to multi-organ failure and acute severe lung injury requiring mechanical ventilation.

Secondary outcome
The kinetics of:
- Circulating soluble serum biomarkers of innate, adaptive and inflammatory immune responses, in order to decipher and validate biomarker signatures of disease severity and risk of acute disease progression.
- Circulating cellular immune responses, focusing on the distribution of various immune subsets (granulocytes, lymphocytes, monocytic and innate populations) and the innate responses to bacterial or viral motifs (LPS, CpG and PolyIC) and polyclonal and/or specific adaptive immune responses (PHA and SARS-CoV-2).
- Circulating cellular immune responses, focusing on the distribution and quantitation of >250 leukocyte subsets, including 20-25 different innate myeloid cells (granulocyte, monocyte, and dendritic cell subsets, etc.), >85 CD4 T-cell subsets, >45 CD8-NK cell subsets, and >115 B-cell & plasma cell subsets.
Special attention will be given to the B-cell system, particularly to minor clonal subsets and the kinetics of expanded plasma cell subsets, down to levels of 0.1 cell per μL.
- Nasal and lung (using cells from lung aspirates) cellular immune responses, focusing on the distribution of various immune subsets (granulocytes, lymphocytes, monocytic and innate populations) and their activation status based on surface markers by mass cytometry (>40 marker panel). Nasal metabolomics.
- Antibody glycosylation: Total IgG Fc glycosylation and SARS-CoV-19 specific IgG Fc glycosylation profiles (Fc glycosylation as general biomarker of immune activation, SARS-CoV-19 specific IgG Fc glycosylation as co-marker for development of immunity, see parameter "SARS-CoV2 specific serology
- Serum glycan profile, anti-glycan IgG/IgM profiles
- RNA expression profiles in whole blood to allow for pathway analysis and characterize different inflammatory responses. Particularly sepsis response phenotypes (e.g. glucosteroid receptor signaling pathway, T cell exhaustion) for the ICU patients.
- Viral load, focusing on measured cycle-threshold (Ct) value kinetics in consecutive (naso)pharynx swabs from SARS-CoV2 qPCR-positive individuals
- SARS-CoV2 whole genome sequencing (subset of patients)
- SARS-CoV2 specific serology, focusing on seroconversion and relative increase of SARS-CoV2-antigen specific seroreactivity, and neutralizing capacity
- Glycocalyx destruction and heparanase activity, functional glycocalyx assays
- Coagulation activation parameters
- Complement activation parameters
- Obesity-related pro(anti)-inflammatory markers
- Biomarker analysis by upconverting phosphor lateral flow assay (UCP-LFA) |

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 | Nasal cell collection and storage
Nasal cells were collected by gently scraping the nasal inferior turbinate using curettes (Rhino-Pro®, Arlington Scientific), as described previously46, and placing them in a tube containing pre-cooled 8mL sterile PBS containing 5mM EDTA (Life Technologies). Such samples provide immune cells from the mucosa that are not found in the lumen (including lymphoid subsets) 46, and have been used previously to study nasal immune responses during controlled viral and bacterial infections 47,48. Per patient and timepoint, two curettes from one nostril were collected. Cells were dislodged by pipetting liquid up and down the tip of curette and spun down at 300xg for 10' at 4°C. Supernatant was completely removed and cells were resuspended in 500μL of PBS. For fixation, an equal amount of freshly prepared 8% formaldehyde (Fisher Scientific) was then added, followed by 30 minutes incubation at room temperature. Cells were then spun down at 800xg for 10'. The supernatant was completely removed and the pellet resuspended in 1mL heat-inactivated fetal bovine serum containing 10% DMSO and moved to a cryovial. Cryovials were frozen in a Mr. FrostyTM freezing container (ThermoFisher Scientific) at -80°C and moved to liquid nitrogen within three days.

Whole blood cryopreservation
Per sample, 200μL of whole blood, collected in an EDTA tube, was added to 1 mL of cold Cryostor® CS10 (Stemcell Technologies) in a 1.8 mL cryovial and mixed by gently inverting. Then cryovials were placed in a cold Mr. Frosty and after 10 minutes moved to -80°C. The next day, cryovials were moved to liquid nitrogen.

CyTOF staining
Samples were barcoded and measured in batches. In every batch, one aliquot of PBMCs from a reference sample was included to be able to normalize staining between batches. Nasal cells were thawed in 2mL RPMI + 50% FBS and spun down |

for 10' at 1600rpm at room temperature. Supernatant was discarded by pipetting. Reference PBMC were washed with 2mL of PBS and then fixed with 4% formaldehyde for 15' at room temperature. Reference PBMC were washed 2x with 2mL BD Perm/Wash (BD). Nasal cells were washed 1x with 1mL BD Perm/Wash, and if clumps were visible, cells were filtered over a 100µm filter (ThermoFisher Scientific). Nasal cells and reference PBMCs were resuspended in 50µL Perm/Wash and then 50µL barcode mix targeting β2 microglobulin (B2M) was added to each individual sample in a 6-choose-3 scheme using Cadmiums 106, 110, 111, 112, 114 and 116 49,50. Samples were incubated for 30' at room temperature and then washed with 4mL Cell Staining Buffer (Fluidigm). Cells were spun 5' at 800xg, supernatant removed and resuspended and combined into 3mL of Perm/Wash. Cells were spun again 5' at 800xg and were resuspended in 45µL Perm/Wash. FcR block (Biolegend, 5µL) and heparin (0.5µL, 100U/mL) were added to prevent aspecific binding of antibodies and cells were incubated for 20' at room temperature51. Then 50µL of antibody cocktail (Table S2) was added, followed by a 45' incubation at room temperature. Cells were then washed twice with 2mL Cell Staining Buffer and spun down for 5' at 800xg. DNA was then stained overnight at 4°C using 1mL Fix and perm buffer (Fluidigm) containing 1000x diluted Intercalator-Ir (Fluidigm). Cells were then washed with Cell Staining Buffer, counted and divided into tubes of 1x106 cells and pelleted down. Tubes were then washed and resuspended in cell acquisition solution (CAS, Fluidigm) with EQ Four Element Calibration Beads (Fluidigm) and acquired on a Helios mass cytometer (Fluidigm) at the Flow cytometry Core Facility (FCF) of Leiden University Medical Center (LUMC) in Leiden, Netherlands (https://www.lumc.nl/research/facilities/fcf). Whole blood samples were thawed in a water bath at 37°C, followed by dropwise addition of 5mL of thawing medium (RPMI 1640 + pen-strep + pyruvate + L-glutamine with 20% heat-inactivated fetal bovine serum, 2 mM Mg2+ and 25U/mL Benzonase). Cells were spun at 400xg for 10 min, and resuspended in 2mL of 1X RBC lysis buffer (Biolegend). After 10-15 minutes the cells were spun at 400g for 10 minutes and resuspended in medium and filtered through a 100 µm cell strainer in case clumps were present. Staining with barcodes and surface antibodies and acquisition was then done as for nasal cells, except for the addition of sodium heparin and Staining Buffer (Fluidigm) was used for whole blood instead of Perm/Wash.

Granulocyte flow cytometry analysis
Cryopreserved whole blood (one donor) or fixed and stored nasal cells (two donors) were thawed at 37°C. Whole blood cells were then fixed as described above for nasal cells. Then, 4mL BD Perm/Wash buffer was added to each sample, followed by 5 minute centrifugation at 800 x g. Cells were washed again in 4mL BD Perm/Wash and spun down again. Then cells were resuspended in 100uL of antibody cocktail containing CD45-BV785 (clone HI30, 1:200, Biolegend), CD66b-APC-Vio770 (clone REA306, Miltenyi, 1:800), CD127-PercpCy5.5 (clone A019D5, Biolegend, 1:50), CD16-BV711 (clone 3G8, Biolegend, 1:100), Siglec8-PEDazzle594 (clone 7C9, Biolegend, 1:200), CD10-PECy7 (clone HI10a, Biolegend, 1:400) and CXCR4-APC (clone 12G5, Biolegend, 1:3200). Cells were incubated for 15 minutes and washed with 3mL FACS buffer (PBS + 0.5% BSA + 2 mM EDTA). Cells were then spun down for 5 minutes at 800 x g, resuspended in 200uL FACS buffer and acquired on a Aurora 3-laser spectral cytometer (Cytek), using Spectroflo software. Single stain controls on whole blood were used for unmixing in addition to unstained fixed nasal cells. Exported FCS files were analysed using Flowjo X (BD).

Peripheral blood leukocyte subsets assessed by flow cytometry
Peripheral blood samples were handled according to a standard sample processing protocol for flow cytometry (for detailed protocol see www.EuroFlow.org and 52. The combination of markers used for cell surface staining is elsewhere described (Primary Immunodeficiency Orientation Tube: PIDOT, Cytognos, https://www.cytognos.com/products/pidot-primary-immunodeficiency-orientation-tube) and was modified by the addition of CD38, for a better separation of the circulating plasmablasts (Supplementary table 8). Briefly, the procedure consisted in the bulk lysis of erythrocytes in fresh samples and staining of 2,5*106 white blood cells (100 µL final staining volume) with a reconstituted PIDOT lyophilized antibody cocktail (containing CD8 FITC, IgD FITC, CD16 PE, CD56 PE, CD4PerCPCy5.5, CD19 PeCy7, TCRgd PE-Cy7, CD3 APC, CD56 APC-C750) and a drop-in antibody cocktail (containing 2 µL CD27 BV421, 2,5 µL CD45RA BV510, 2 µL CD38 BV605, 0,6 µL pure CD38 per test). At least 1 million events were acquired using a 3-laser Cytek® Aurora instrument (Cytek Biosciences) from the Flow cytometry Core Facility (FCF) of Leiden University Medical Center (LUMC) in Leiden, Netherlands (https://www.lumc.nl/research/facilities/fcf). The main circulating leukocyte subsets identified by flow cytometry were assigned manually using the data analysis software Infinicyt (Cytognos SL) based on a standardized gating strategy 53, with an adaptation for the CD38 inclusion (Supplementary figure 12). The absolute counts per µL fresh blood were determined by a double platform approach, using the absolute fresh leukocyte counts determined prior sample processing with hematological analyzer (Sysmex) to the Statistics Configure tool of the Infinicyt software

Instrument	Sorting: SARS-CoV-2 specific T cells were sorted on a FACSAria 3. Flow analysis: The data was collected on a 3L aurora (Cytek)
Software	FACSDiva (BD) or Spectroflo (Cytek)
Cell population abundance	not done
Gating strategy	Cells were gated on FSC/SSC parameters to select single lymphocytes. Then, CD4+CD154 or CD137 for activated CD4+ T cells and CD137+CD8+ or pMHC-tetramer+ for SARS-CoV2-specific CD8+ T cells.

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