

**Med, Volume 4**

## **Supplemental information**

### **Cytotoxic CD4<sup>+</sup> tissue-resident memory T cells are associated with asthma severity**

**Sara Herrera-De La Mata, Ciro Ramírez-Suástegui, Heena Mistry, Francisco Emmanuel Castañeda-Castro, Mohammad A. Kyyaly, Hayley Simon, Shu Liang, Laurie Lau, Clair Barber, Monalisa Mondal, Hongmei Zhang, Syed Hasan Arshad, Ramesh J. Kurukulaaratchy, Pandurangan Vijayanand, and Grégory Seumois**

## Supplementary Items

### Supplementary Figures:

Figure S1. Single-cell clustering analysis using Seurat (Related to Figure 1).

Figure S2. Single-cell cluster proportions, flow cytometry gating strategy to isolate CD4<sup>+</sup> T cells and subsets, and correlation of single-cell cluster proportions with clinical features (Related to Figure 2).

Figure S3. Expression of differentially expressed genes in CD103<sup>+</sup> T<sub>RM</sub> subset, and bulk RNA-seq and TCR analysis of sorted airway CD4<sup>+</sup> T cells in resting condition (Related to Figure 3).

Figure S4. Qualitative changes in gene expression of airway CD4<sup>+</sup> T cell subsets in relation to disease severity and sex (Related to Figure 4).

Figure S5. Single-cell analysis of CD4<sup>+</sup> T cells upon stimulation (Related to Figure 5).

### Supplementary Tables:

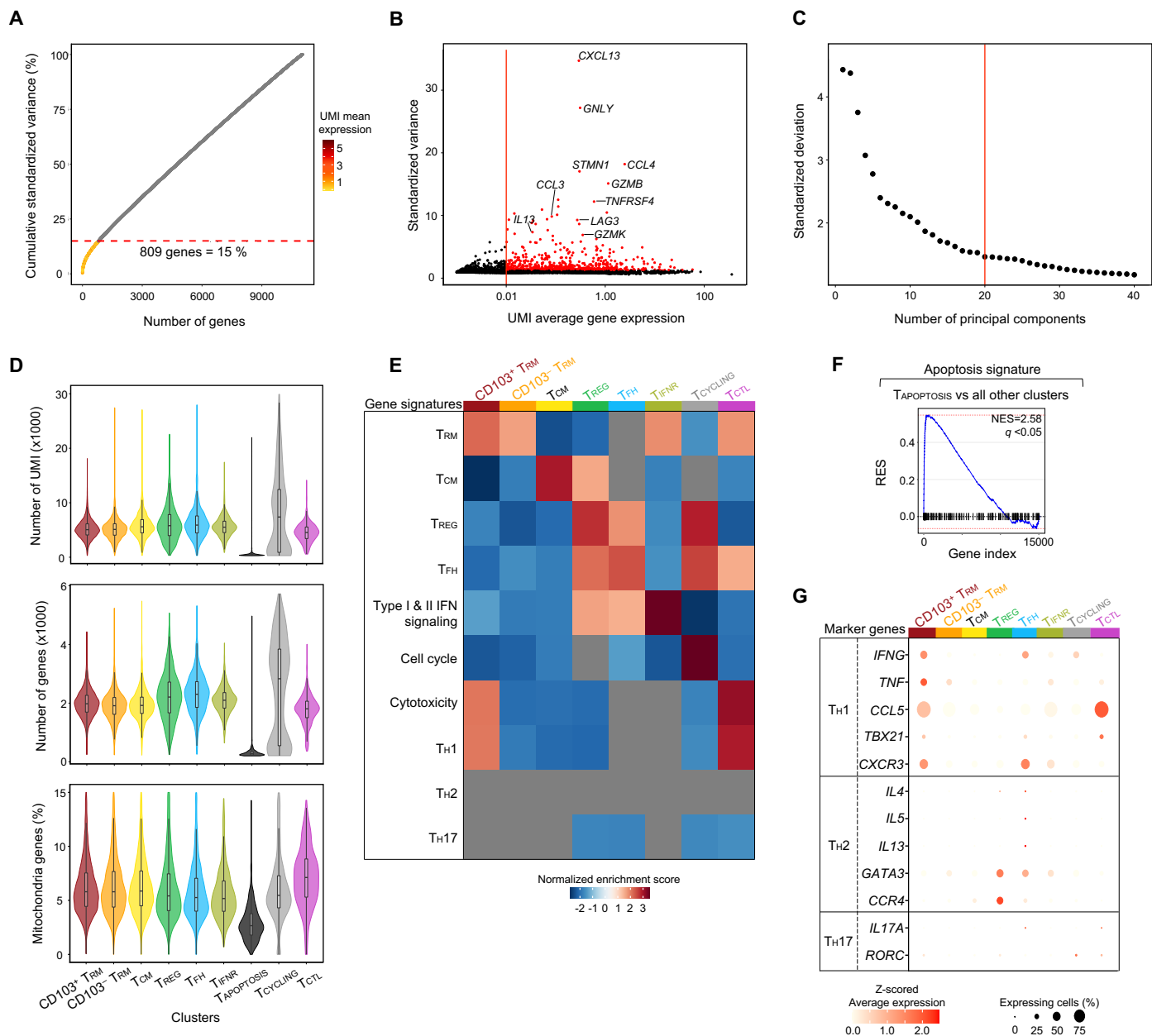
Provided as separate supplementary files:

Table S1. Study design and experimental details (Excel spreadsheet) (Related to Figures 1, 2, and 5; and Figures S2 and S3).

Table S2. Single-cell analysis (Excel spreadsheet) (Related to Figures 1-5; and Figures S1-S5).

Table S3. Gene signatures (Excel spreadsheet) (Related to Figures 1 and 3-5; and Figures S1, S3, and S4).

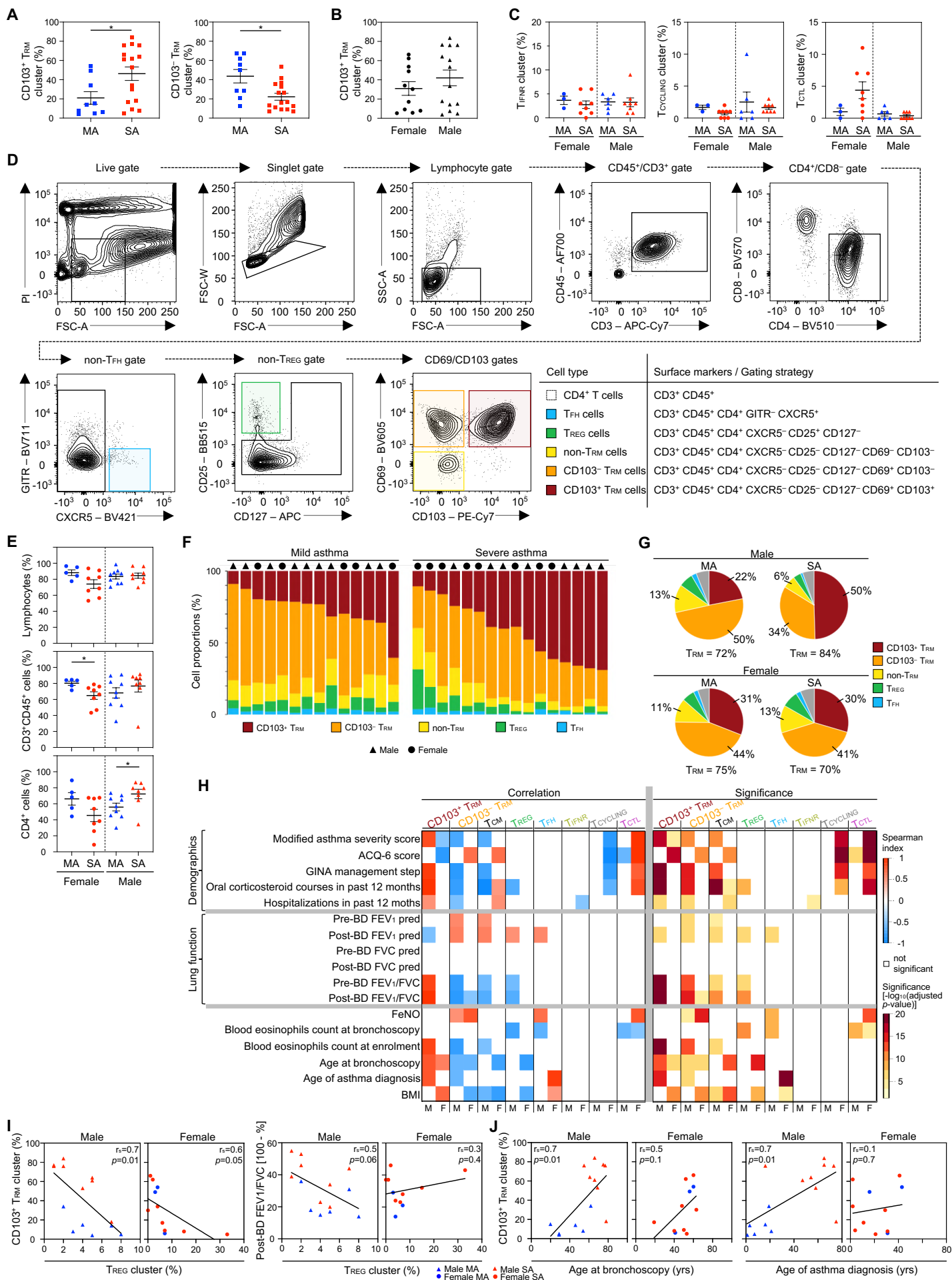
Table S4. Differential gene expression analysis in bulk populations (Excel spreadsheet) (Related to Figure 3; and Figures S3 and S5).



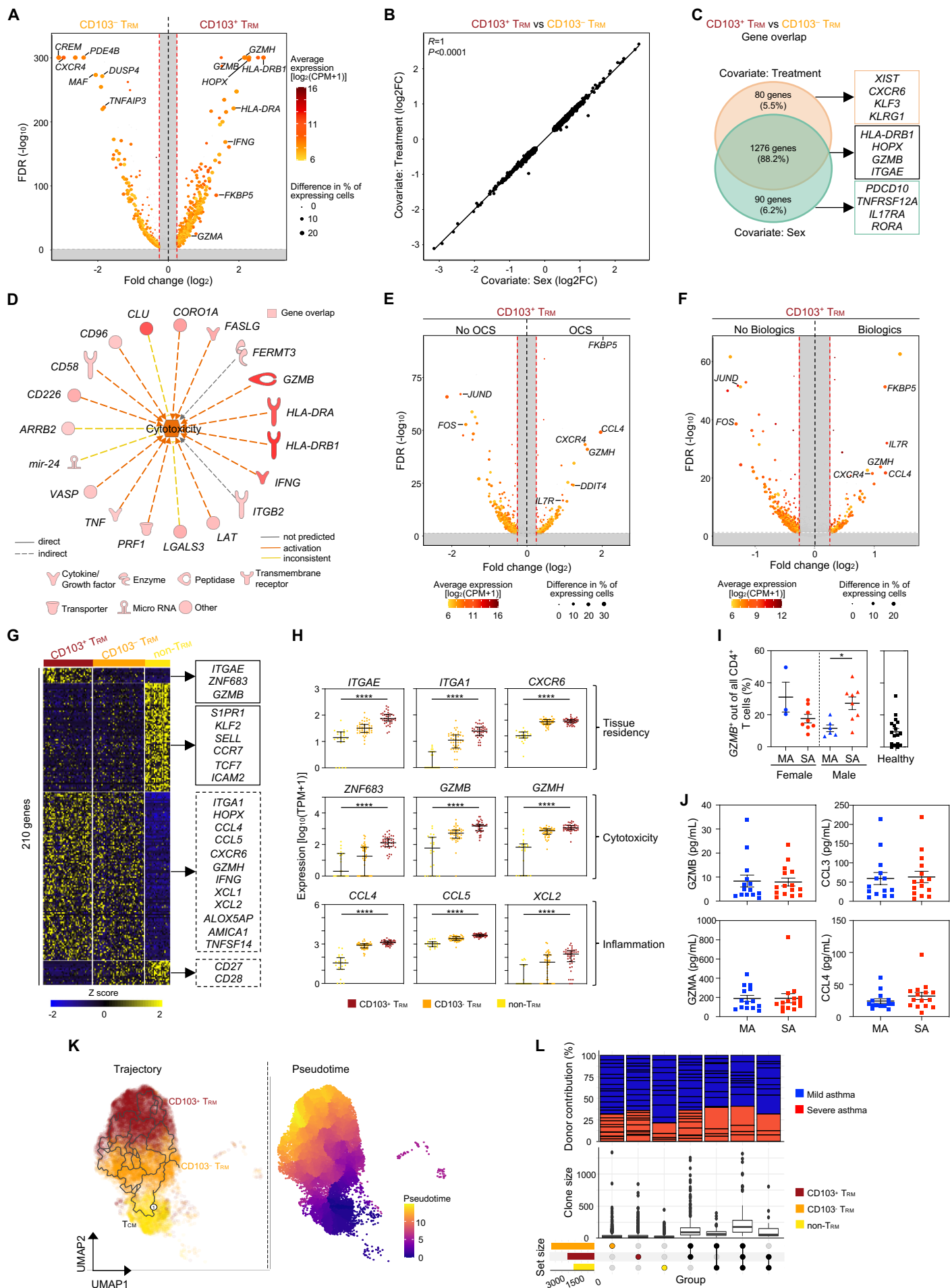
**Figure S1. Single-cell clustering analysis using Seurat** (Related to Figure 1). **(A)** Scatter plot represents cumulative standardized variance for genes ( $n=11,055$ ) expressed in more than 0.1% of the  $CD4^+$  T cells ( $n=27,771$ ) and with a UMI mean  $> 0.01$ . Genes are ranked from highest to lowest standardized variance values. Genes below the red dotted line were selected ( $n=809$ ) as part of the highest variable genes explaining 15% of the cumulative standardized variance. Genes are colored based on the UMI mean expression given by Seurat. **(B)** Scatter plot represents the distribution of genes ordered based on their UMI mean expression values in function of their individual standardized variance values. Genes colored in red are the 809 highest variable genes with a UMI mean expression value  $> 0.01$ . **(C)** Scatter plot shows the standardized variation for first 40 principal components (PCs) using the 809 most variable genes. Red lines indicate the number of PCs selected for clustering analysis in respect of Seurat methodology. **(D)** Violin plots show distribution of number of UMI (threshold  $< 30,000$ ) (top), number of genes (thresholds: lower = 200, upper = 6,000) (middle), and the percentage of mitochondrial genes detected (threshold  $< 15\%$ ) (bottom) per cell for each cluster. Colors are based on cluster-type. **(E)** Heatmap with normalized gene set enrichment scores between each cluster against the rest of cells for each cluster. Color scale corresponds to normalized enrichment scores for each gene list

(rows) and cluster (columns) (see Table S3A). Gray color indicates no statistical significance (adjusted  $P$ -value  $> 0.05$ ). **(F)** GSEA for apoptosis signature in  $T_{\text{APOPTOSIS}}$  cluster *versus* all clusters (see Table S3A). An enrichment of more than 0.5 was threshold value of exclusion of cluster from analysis. NES, normalized enrichment score; q, false discovery rate. **(G)** Row-wise z-score-normalized mean expression (color scale) and percent of expressing cells (size scale) plot for a selection of canonical marker genes for  $T_{\text{H}}1$ ,  $T_{\text{H}}2$ , and  $T_{\text{H}}17$  in each cluster.

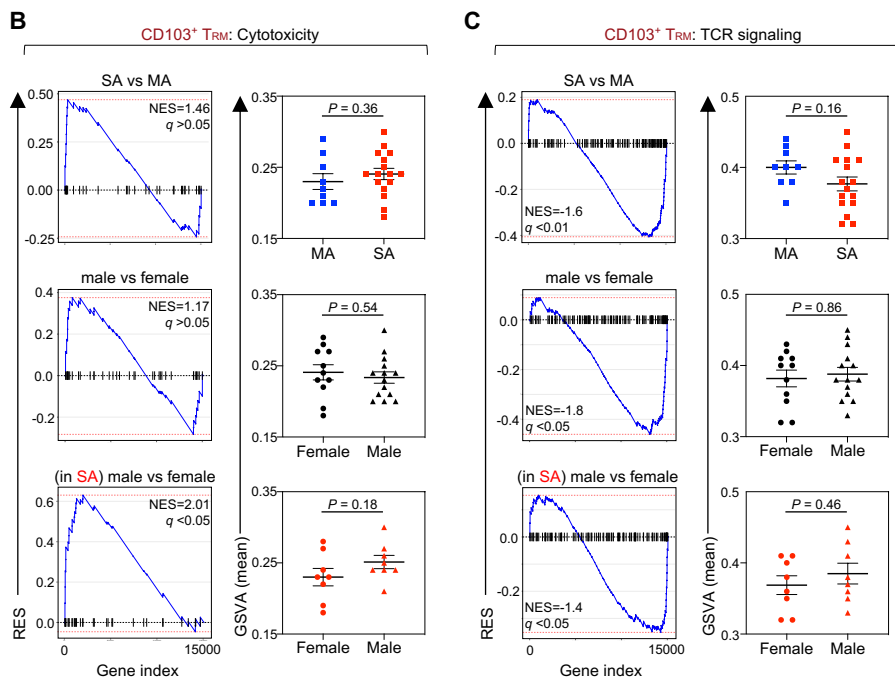
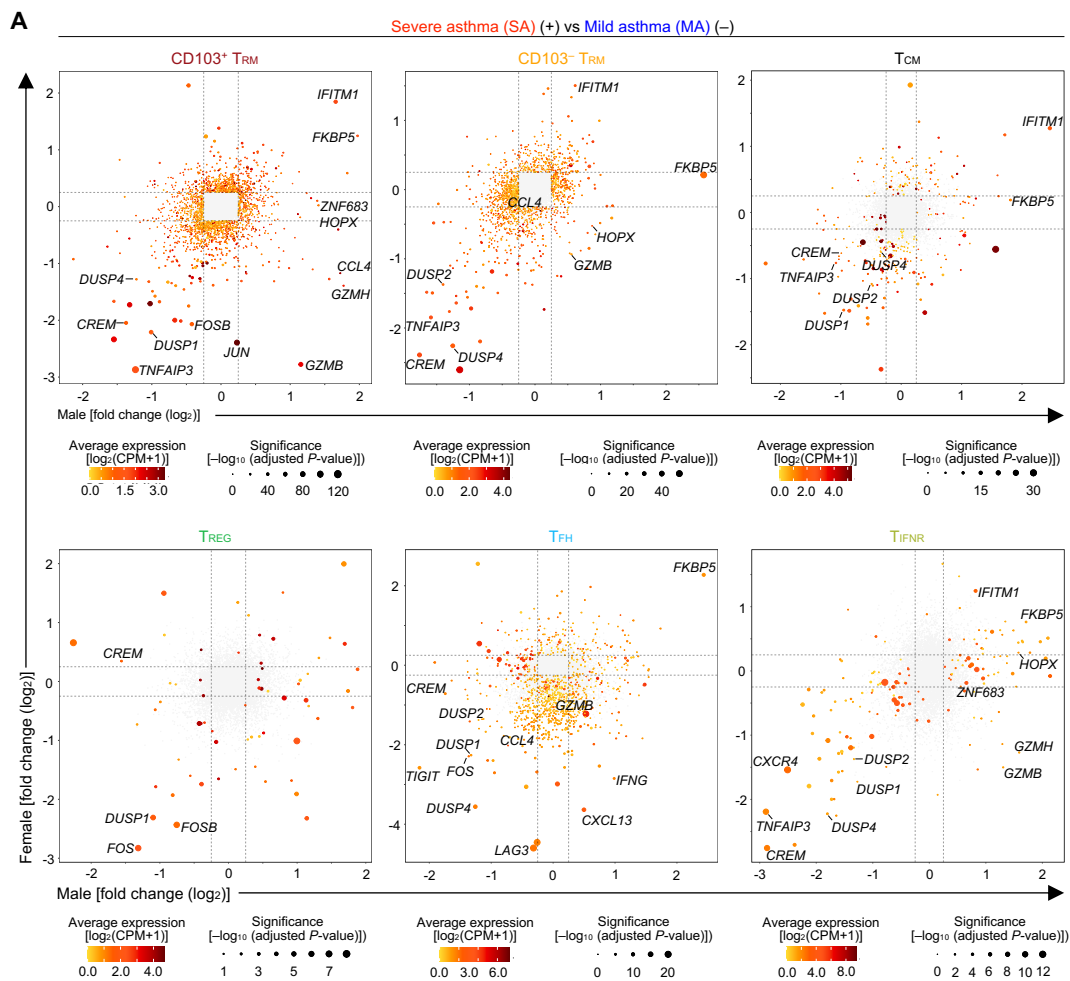




**Figure S2. Single-cell cluster proportions, flow cytometry gating strategy to isolate CD4<sup>+</sup> T cells and subsets, and correlation of single-cell cluster proportions with clinical features** (Related to Figure 2). **(A)** Dot plots show proportions of CD103<sup>+</sup> T<sub>RM</sub> (left) and CD103<sup>-</sup> T<sub>RM</sub> (right) cluster cells for each donor separated in mild and severe asthma groups (\*,  $P < 0.05$ ; Mann-Whitney U test). **(B)** Dot plot shows proportions of CD103<sup>+</sup> T<sub>RM</sub> cluster cells for each donor separated by sex (Mann-Whitney U test). **(C)** Dot plots show the proportions of T<sub>IFNR</sub>, T<sub>CYCLING</sub>, and T<sub>CTL</sub> cluster cells for each donor, grouped per disease severity and sex (Mann-Whitney U test). **(D)** Flow cytometry gating strategy to sort live (Propidium iodide (PI), singlets (Width vs Area forward scatter (FSC)), lymphocyte size (Side scatter vs Forward scatter), CD3<sup>+</sup>CD45<sup>+</sup>, CD4<sup>+</sup>CD8<sup>-</sup>, CXCR5<sup>-</sup>, CD25<sup>-</sup>CD127<sup>+</sup>, CD69<sup>+/-</sup>CD103<sup>+/-</sup>. **(E)** Dot plots showing percentage of lymphocytes, CD3<sup>+</sup>CD45<sup>+</sup>, and CD4<sup>+</sup> T cell subsets per patient distributed between sex and disease groups (\*,  $P < 0.05$ ; Mann-Whitney U test). **(A, B, C, E)** Each dot represents data from a single patient, bars represent the mean, and t-lines represent SEM. **(F)** Normalized stacked bar charts represent proportions of the different CD4<sup>+</sup> T cell subsets-based on cell surface marker expression per donor grouped as mild and severe asthma. Sex is indicated on top of each bar (triangle=male, circle=female). Colors correspond to cell subset type. **(G)** Pie charts represent average proportions of CD4<sup>+</sup> T cell subsets in mild and severe asthma patients separated by sex (MA=mild asthma, SA=severe asthma). Colors correspond to cell subset type. **(H)** Heatmap shows Spearman correlation indices calculated between cluster proportions separated by sex (columns) with clinical features (rows). Color code is based on correlation coefficient  $r$  (left) and the adjusted  $P$ -value (right), which were computed using Spearman correlation (Table S1E). **(I)** Correlation scatter plots between proportions of CD103<sup>+</sup> T<sub>RM</sub> cluster and proportions of T<sub>REG</sub> cluster (left) as well as proportions of T<sub>REG</sub> cluster and post-BD FEV1/FVC (right), for each patient in males and females separately. **(J)** Correlation scatter plots between proportions of CD103<sup>+</sup> T<sub>RM</sub> cluster for each donor and specific clinical features (age at bronchoscopy and age of asthma diagnosis) in males and females separately. **(I-J)** Dots shaped based on patient sex, and colored based on disease severity status. Correlation coefficient  $r$  and exact  $P$  value were computed using Spearman correlation (Table S1E).

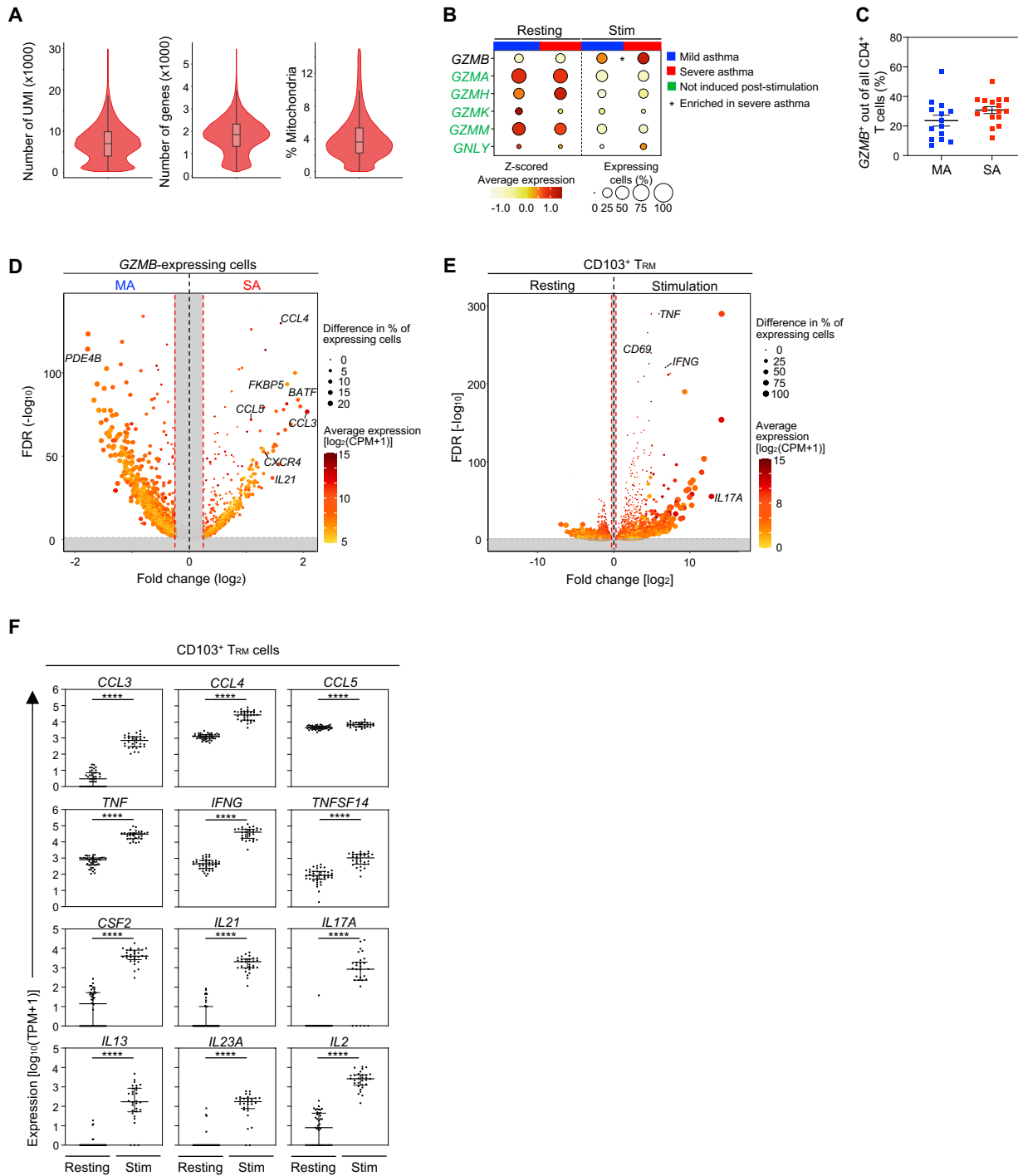


**Figure S3. Expression of differentially expressed genes in CD103<sup>+</sup> T<sub>RM</sub> subset, and bulk RNA-seq and TCR analysis of sorted airway CD4<sup>+</sup> T cells in resting condition** (Related to Figure 3). **(A)** Volcano plot shows false discovery rate (FDR) ( $-\log_{10}$  adjusted  $P$ -value) and  $\log_2$  (fold change) in expression for genes differentially expressed in CD103<sup>+</sup> T<sub>RM</sub> *versus* CD103<sup>-</sup> T<sub>RM</sub> clusters using treatment as covariate. Dots are colored according to the mean of expression ( $\log_2$ ) and sized based on the difference in the percentage of cells expressing the given gene, both derived from the group in which the gene is up-regulated. Gray dotted lines represent the statistical threshold values:  $\log_2(\text{fold change}) > 0.25$  and  $-\log_{10}(\text{FDR}) > 1.3$  (adjusted  $P$ -value  $< 0.05$ ). **(B)** Scatter plot shows the  $\log_2$  (fold change) in expression for genes differentially expressed in CD103<sup>+</sup> T<sub>RM</sub> *versus* CD103<sup>-</sup> T<sub>RM</sub> clusters using sex (x-axis) or treatment (y-axis) as covariate (Spearman correlation coefficient  $R$  and  $P$  value were computed). **(C)** Venn diagram shows overlap between differentially expressed genes in CD103<sup>+</sup> T<sub>RM</sub> *versus* CD103<sup>-</sup> T<sub>RM</sub> clusters when using treatment (top) or sex (bottom) as covariate. Example genes are listed on the right. **(D)** Ingenuity pathway analysis (IPA) network of cytotoxicity-linked genes with increased expression (adjusted  $P$ -value  $< 0.01$  and  $\log_2$  (fold change)  $> 0.5$ ) in CD103<sup>+</sup> T<sub>RM</sub> cluster compared to CD103<sup>-</sup> T<sub>RM</sub> cluster. Arrows represent type of interaction between molecules and network biology, molecules are colored based on their relative expression, and shaped based on molecule function. **(E, F)** Volcano plots show false discovery rate (FDR) ( $-\log_{10}$  adjusted  $P$ -value) and  $\log_2$  (fold change) in expression for genes differentially expressed in CD103<sup>+</sup> T<sub>RM</sub> of patients on oral corticosteroids (OCS) *versus* off OCS (No OCS) **(E)** or on biologics (Biologics) *versus* off biologics (No Biologics) **(F)** using sex as covariate. Dots are colored according to the mean of expression ( $\log_2$ ) and sized based on the difference in the percentage of cells expressing the given gene, both derived from the group in which the gene is up-regulated. Gray dotted lines represent the statistical threshold values:  $\log_2(\text{fold change}) > 0.25$  and  $-\log_{10}(\text{FDR}) > 1.3$  (adjusted  $P$ -value  $< 0.05$ ). **(G)** Heatmap of sorted bulk RNA-seq samples shows row-wise z-scored expression of 210 differentially expressed genes between CD103<sup>+</sup> T<sub>RM</sub>, CD103<sup>-</sup> T<sub>RM</sub>, and non-T<sub>RM</sub> subsets. Adjusted  $P$ -value  $< 0.05$  and  $\log_2$  (fold change)  $> 1$ . **(H)** Dot plots show normalized expression for example genes differentially up-regulated in CD103<sup>+</sup> T<sub>RM</sub> cells linked to tissue residency, cytotoxicity, and inflammation (\*\*\*\*,  $P < 0.0001$ ; Kruskal-Wallis one-way test followed by Dunn's post-hoc test). **(I)** Dot plots show proportions of GZMB<sup>+</sup> cells from total CD4<sup>+</sup> T cells per patient distributed between sex (shapes) and disease (color) groups (left) and in healthy subjects (from published datasets)<sup>[S1-S4]</sup> (right) (\*,  $P < 0.05$ ; Mann-Whitney U test). **(J)** Dot plots show concentration of GZMB, GZMA, CCL3, and CCL4 proteins measured in BAL supernatants using multiplex ELISA for each donor separated by disease (Mann-Whitney U test). **(H-J)** Each dot represents data from a single subject, bars represent the mean, and t-lines represent SEM. **(K)** Single-cell pseudotime trajectory analysis of airway CD4<sup>+</sup> T<sub>CM</sub> and T<sub>RM</sub> cell subsets. Trajectory constructed using the Monocle3 algorithm. **(L)** Bar chart shows the proportion of TCR clones per donor (donor contribution) separated by disease (top) and the TCR clone size specific or shared (connecting black dotted line) between bulk samples from CD103<sup>+</sup> T<sub>RM</sub>, CD103<sup>-</sup> T<sub>RM</sub>, and non-T<sub>RM</sub> cells (bottom). Total number of clones in each sample group (left bottom corner) is shown.



**Figure S4. Qualitative changes in gene expression of airway CD4<sup>+</sup> T cell subsets in relation to disease severity and sex** (Related to Figure 4). **(A)** Crater plots show log<sub>2</sub> (fold change) of gene expression between mild and severe asthma in male (x-axis) and female (y-axis) patients in each cluster. Dotted lines indicate the

statistical threshold values of fold change for gene filtering. **(B, C)** (Left) GSEA shows enrichment of genes linked to cytotoxicity **(B)** or TCR signaling **(C)** in CD103<sup>+</sup> T<sub>RM</sub> cluster when comparing severe asthma (SA) *versus* mild asthma (MA) (top), male *versus* female (middle), and male *versus* female within severe asthma (bottom). NES, normalized enrichment score; q, false discovery rate. On the right, GSVA shows cytotoxicity **(B)** or TCR signaling **(C)** enrichment scores per donor for the corresponding comparisons mentioned here above (Mann-Whitney U test). Gene lists in Table S3A. Each dot represents data from a single patient, bars represent the mean, and t-lines represent SEM.



**Figure S5. Single-cell analysis of CD4<sup>+</sup> T cells upon stimulation** (Related to Figure 5). **(A)** Violin plots show distribution of single-cell dataset passing quality checks: number of UMI per cell (threshold = 30,000) (left), number of genes per cell (thresholds: lower = 200, upper = 6,000) (middle), and percentage of mitochondrial genes per cell (< 15 %) (right). **(B)** Plot shows row-wise z-score-normalized mean expression (color scale) and percent of expressing cells (size scale) for indicated genes in resting and stimulation conditions per disease. **(C)** Dot plot shows proportions of *GZMB*<sup>+</sup> cells from total stimulated CD4<sup>+</sup> T cells per patient distributed between disease (color) groups (Mann-Whitney U test). **(D)** Volcano plot shows statistical significance [ $-\log_{10}$  (adjusted *P* value)] in function of the fold change ( $\log_2$ ) in expression between severe versus mild asthma in *GZMB*-expressing cells after stimulation. **(E)** Volcano plot shows statistical significance [ $-\log_{10}$  adjusted *P* value] in function of the  $\log_2$ -fold change in expression for differentially expressed genes when comparing expression

between bulk RNA-seq stimulation versus resting within the CD103<sup>+</sup> T<sub>RM</sub> cells. Equal numbers of cells were sampled in each group. **(D, E)** Dots are colored according to the average of expression ( $\log_2$ ) and sized on the basis of the fraction of cells expressing the given gene, both derived from the group in which the gene is up-regulated. Gray dotted lines represent the threshold value for fold change [y-axis,  $\log_2(|\text{fold change}|) > 1$ ] and significance [x-axis,  $-\log_{10}(\text{adjusted P value}) > 2$ ]. FDR, false discovery rate. **(F)** Dot plots of genes differentially upregulated in stimulated bulk CD103<sup>+</sup> T<sub>RM</sub> cells (Mann-Whitney U test). **(C, F)** Each dot represents data from a single patient, bars represent the mean, and t-lines represent SEM.



## REFERENCES

- S1. Liao, M., Liu, Y., Yuan, J., Wen, Y., Xu, G., Zhao, J., Cheng, L., Li, J., Wang, X., Wang, F., et al. (2020). Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nature Medicine* 26, 842-844. 10.1038/s41591-020-0901-9.
- S2. Grant, R.A., Morales-Nebreda, L., Markov, N.S., Swaminathan, S., Querrey, M., Guzman, E.R., Abbott, D.A., Donnelly, H.K., Donayre, A., Goldberg, I.A., et al. (2021). Circuits between infected macrophages and T cells in SARS-CoV-2 pneumonia. *Nature* 590, 635-641. 10.1038/s41586-020-03148-w.
- S3. Morse, C., Tabib, T., Sembrat, J., Buschur, K.L., Bittar, H.T., Valenzi, E., Jiang, Y., Kass, D.J., Gibson, K., Chen, W., et al. (2019). Proliferating SPP1/MERTK-expressing macrophages in idiopathic pulmonary fibrosis. *Eur Respir J* 54. 10.1183/13993003.02441-2018.
- S4. Mould, K.J., Moore, C.M., McManus, S.A., McCubbrey, A.L., McClendon, J.D., Griesmer, C.L., Henson, P.M., and Janssen, W.J. (2021). Airspace Macrophages and Monocytes Exist in Transcriptionally Distinct Subsets in Healthy Adults. *Am J Respir Crit Care Med* 203, 946-956. 10.1164/rccm.202005-1989OC.