

Common human genetic variants of *APOE* impact murine COVID-19 mortality

<https://doi.org/10.1038/s41586-022-05344-2>

Received: 25 September 2021

Accepted: 13 September 2022

Published online: 21 September 2022

 Check for updates

Benjamin N. Ostendorf^{1,2,3,4}✉, Mira A. Patel^{1,7}, Jana Bilanovic^{1,7}, H.-Heinrich Hoffmann⁵, Sebastian E. Carrasco⁶, Charles M. Rice⁵ & Sohail F. Tavazoie¹✉

Clinical outcomes of severe acute respiratory syndrome 2 (SARS-CoV-2) infection are highly heterogeneous, ranging from asymptomatic infection to lethal coronavirus disease 2019 (COVID-19). The factors underlying this heterogeneity remain insufficiently understood. Genetic association studies have suggested that genetic variants contribute to the heterogeneity of COVID-19 outcomes, but the underlying potential causal mechanisms are insufficiently understood. Here we show that common variants of the apolipoprotein E (*APOE*) gene, homozygous in approximately 3% of the world's population¹ and associated with Alzheimer's disease, atherosclerosis and anti-tumour immunity²⁻⁵, affect COVID-19 outcome in a mouse model that recapitulates increased susceptibility conferred by male sex and advanced age. Mice bearing the *APOE2* or *APOE4* variant exhibited rapid disease progression and poor survival outcomes relative to mice bearing the most prevalent *APOE3* allele. *APOE2* and *APOE4* mice exhibited increased viral loads as well as suppressed adaptive immune responses early after infection. In vitro assays demonstrated increased infection in the presence of *APOE2* and *APOE4* relative to *APOE3*, indicating that differential outcomes are mediated by differential effects of *APOE* variants on both viral infection and antiviral immunity. Consistent with these in vivo findings in mice, our results also show that *APOE* genotype is associated with survival in patients infected with SARS-CoV-2 in the UK Biobank (candidate variant analysis, $P = 2.6 \times 10^{-7}$). Our findings suggest *APOE* genotype to partially explain the heterogeneity of COVID-19 outcomes and warrant prospective studies to assess *APOE* genotyping as a means of identifying patients at high risk for adverse outcomes.

SARS-CoV-2 has caused the COVID-19 pandemic with more than 580 million confirmed infections and 6 million deaths so far worldwide. Clinical presentations of SARS-CoV-2 infection show pronounced variation, ranging from asymptomatic infection to lethal disease. Several epidemiological factors have been identified that associate with adverse outcome, including male sex, advanced age, select comorbidities and genetic ancestry⁶. However, these factors only partially explain the wide interindividual clinical spectrum of SARS-CoV-2 infection. There is thus a major need to identify the factors underlying susceptibility to poor outcome in COVID-19. Major efforts have shown germline genetics to correlate with disease severity in COVID-19 (reviewed in ref. ⁷). Among these, candidate gene approaches have revealed rare autosomal inborn errors of type I interferon (IFN) immunity to alter type I IFN signalling in vitro^{8,9}. In addition, genome-wide association studies have identified several genomic loci to be significantly associated with critical COVID-19 (refs. ¹⁰⁻¹⁵). However, it remains unknown whether common germline variants causally affect the course of COVID-19.

APOE is a secreted protein with canonical roles in lipid metabolism. Importantly, *APOE* has also been shown to modulate immunity in different contexts, including infection and anti-tumour immunity^{2,16,17}. Two single nucleotide polymorphisms give rise to three highly prevalent variants of *APOE*, termed *APOE2*, *APOE3* and *APOE4*. The proteins encoded by these alleles differ by one or two amino acids. Forty per cent of the world's population carry at least one copy of either the *APOE2* or *APOE4* allele, and approximately 3% are homozygous for either *APOE2* or *APOE4* (ref. ¹). The *APOE4* variant is the strongest monogenetic risk factor for Alzheimer's disease^{18,19}. *APOE* variants also modulate several immune-related processes, including atherosclerosis⁴ and anti-tumour immunity³, prompting us to determine whether *APOE* causally modulates SARS-CoV-2 infection. Using genetic mouse models of *APOE* human genetic variation as well as supportive clinical association studies, we found that the *APOE2* and *APOE4* variants confer adverse outcomes in SARS-CoV-2 infection in vivo including reduced survival.

¹Laboratory of Systems Cancer Biology, The Rockefeller University, New York, NY, USA. ²Department of Hematology, Oncology, and Tumor Immunology, Charité-Universitätsmedizin Berlin, Berlin, Germany. ³Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Berlin, Germany. ⁴Berlin Institute for Medical Systems Biology (BIMSB), Max Delbrück Center for Molecular Medicine, Berlin, Germany. ⁵Laboratory of Virology and Infectious Disease, The Rockefeller University, New York, NY, USA. ⁶Laboratory of Comparative Pathology, Weill Cornell Medicine, Memorial Sloan Kettering Cancer Center, and Rockefeller University, New York, NY, USA. ⁷These authors contributed equally: Mira A. Patel, Jana Bilanovic.

✉e-mail: bostendorf@rockefeller.edu; sohail.tavazoie@rockefeller.edu

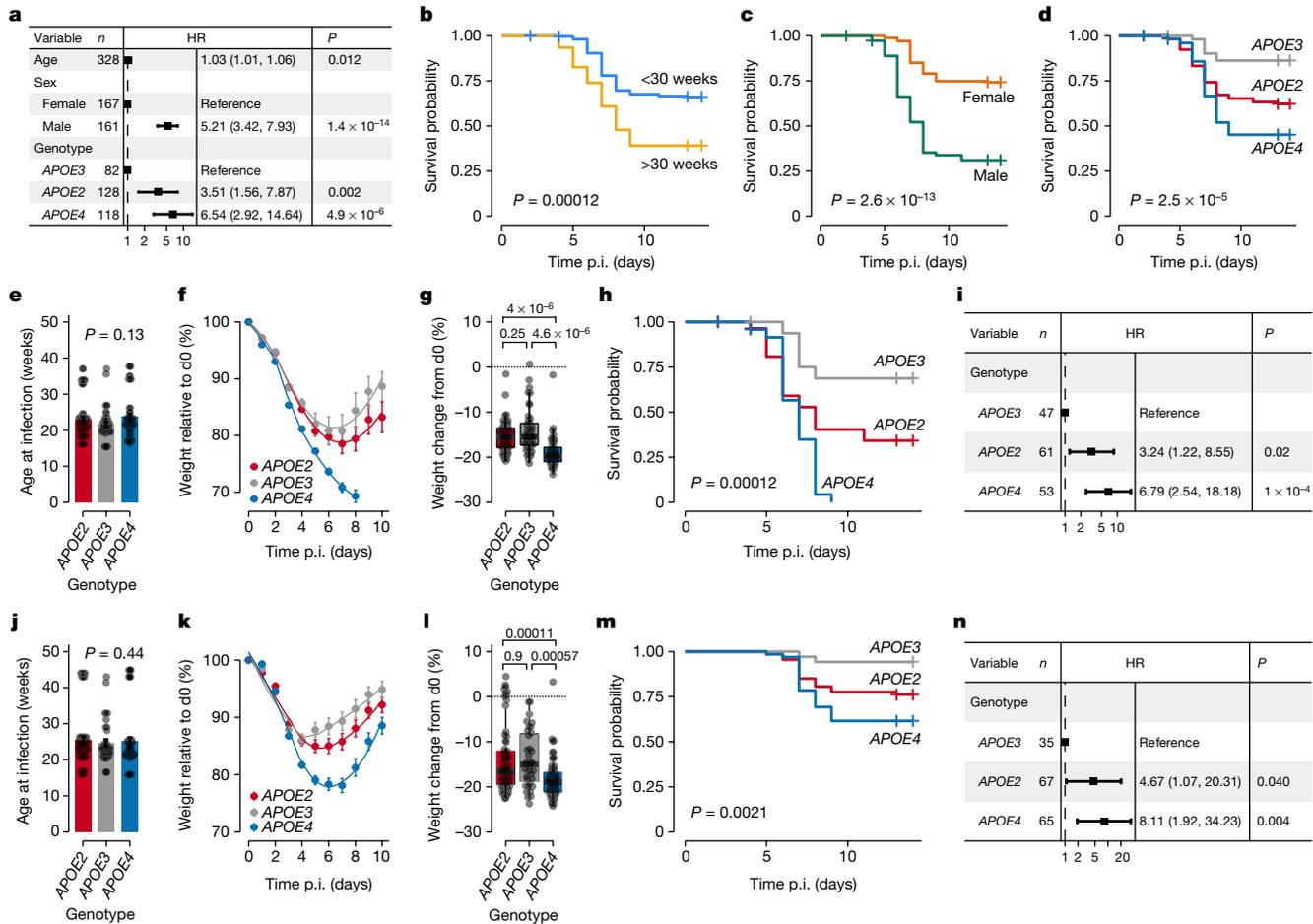


Fig. 1 | *APOE* variants modulate outcome of mouse SARS-CoV-2 MA10

infection. **a**, Multivariate analysis of the impact of age, sex and *APOE* genotype on survival of SARS-CoV-2 MA10-infected *APOE*-knock-in mice (*P* values according to multivariable Cox proportional hazards model; *n* = 128, 82 and 118 for *APOE2*, *APOE3* and *APOE4*, respectively; data pooled from 13 independent experiments). *n*, sample size; HR, hazard ratio for death. **b–d**, Survival of combined male and female SARS-CoV-2 MA10-infected *APOE*-knock-in mice stratified by age (cutoff: 30 weeks) (**b**), sex (**c**) and *APOE* genotype (**d**) (*P* values according to log-rank tests). p.i., post infection. **e–n**, Age distribution (**e**), weight course (**f**), weight on day 4 post infection (**g**), survival (**h**, **m**) and hazard ratios (**i**, **n**) of male

(**e–i**) versus female (**j–n**) *APOE*-knock-in mice from **a** stratified by *APOE* genotype (*P* values according to Kruskal–Wallis test (**e**, **j**), two-sided *t*-tests (**g**, **l**), log-rank test (**h**, **m**) and Cox proportional hazard models (**i**, **n**); note that **f** and **k** show group averages, but some animals died or were censored for tissue collection during the course of the experiment). d0, day 0. The error bars in **a**, **i**, **n** indicate 95% confidence intervals. The error bars in **e**, **f**, **j**, **k** indicate the standard error of the mean. The boxplot whiskers in **g**, **l** extend to the smallest and largest value within 1.5 times the interquartile ranges of the hinges, and the box centre and hinges indicate the median and first and third quartiles, respectively.

APOE causally affects mouse COVID-19 outcomes

To assess the impact of *APOE* germline variation on SARS-CoV-2 infection, we infected 328 *APOE*-knock-in mice across different ages and of both sexes with SARS-CoV-2 MA10, a mouse-adapted strain of SARS-CoV-2 (ref. ²⁰) (Extended Data Fig. 1a–c). In *APOE*-knock-in mice, the mouse *Apoe* gene is replaced with one of the three human *APOE* alleles. Multivariate analysis revealed that this mouse model recapitulated the increased risk for poor survival conferred by male sex and those of advanced age as previously shown in humans (Fig. 1a–c). Remarkably, *APOE* genotype also significantly affected survival, with both the *APOE2* and *APOE4* variants conferring poor survival outcomes relative to the *APOE3* variant (Fig. 1a, d). In both male and female age-matched mice, *APOE4* mice exhibited accelerated weight loss relative to the other variants (Fig. 1e–g, j–l and Extended Data Fig. 1d). Although female mice showed a higher level of survival overall, *APOE2* and *APOE4* conferred worse survival outcomes in both male and female mice (Fig. 1h, i, m, n). The impact was particularly pronounced in male mice, with 100% of *APOE4* mice succumbing to COVID-19 in contrast to approximately 30% mortality in *APOE3* mice. We observed a significant

interaction between *APOE* genotype and age, with the impact of *APOE* on survival being more pronounced in younger mice. No significant interaction was observed for *APOE* genotype and sex (Extended Data Fig. 1e–h). Of note, no spontaneous deaths were detected in similarly aged and non-infected *APOE*-knock-in mice over a comparable period, indicating that the known impact of *APOE* genotype on longevity does not confound these results (Extended Data Fig. 1i–k). Thus, *APOE* variants causally and markedly affect the outcome of mouse COVID-19.

APOE2 and *APOE4* mice exhibit accelerated COVID-19

To assess viral load, we carried out TaqMan quantitative real-time PCR on lungs from *APOE*-knock-in mice on day 4 post infection. Consistent with faster disease progression, elevated viral loads were present in *APOE2* and *APOE4* relative to *APOE3* mice (Fig. 2a). These differences were already evident on day 2 post infection (Extended Data Fig. 2a) and validated by SARS-CoV-2 nucleocapsid immunofluorescence staining (Fig. 2b). Histopathological analyses on day 4 post infection revealed pronounced lung injury in *APOE4* mice with increased bronchiolar necrosis, alveolar damage and fibrin deposition (Fig. 2c–f and Extended

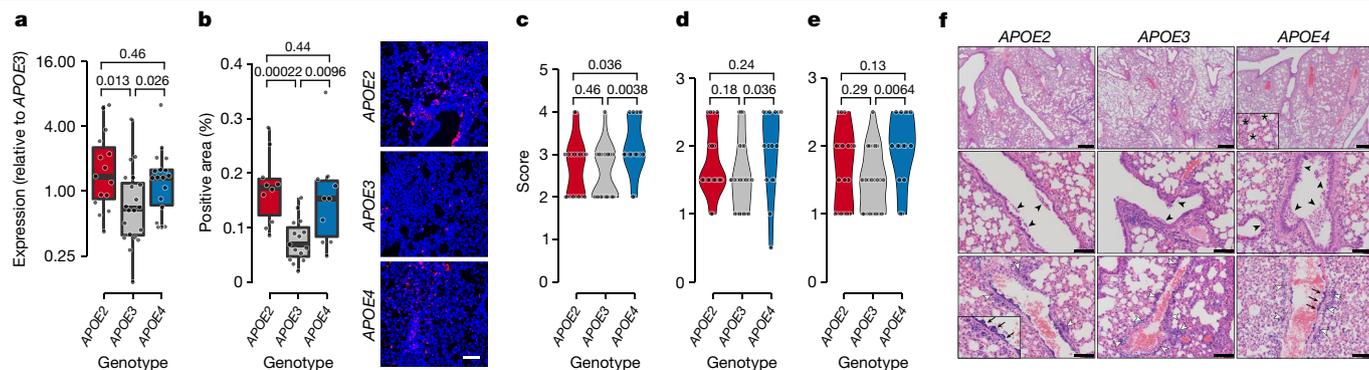


Fig. 2 | *APOE2* and *APOE4* mice exhibit accelerated progression of COVID-19 relative to *APOE3* mice. **a**, TaqMan quantitative PCR for SARS-CoV-2 N1 in homogenized lungs from *APOE*-knock-in mice on day 4 post infection with SARS-CoV-2 MA10 (data pooled from two experiments; *P* values according to two-sided Mann–Whitney test; *n* = 15, 20 and 18 for *APOE2*, *APOE3* and *APOE4*, respectively). **b**, Left: quantification of immunofluorescence staining for SARS-CoV-2 nucleocapsid in lungs of *APOE*-knock-in mice on day 4 after infection with SARS-CoV-2 MA10 (*P* values according to two-sided Mann–Whitney tests; *n* = 10, 15 and 10 for *APOE2*, *APOE3* and *APOE4*, respectively). Right: images showing representative sections (with SARS-CoV-2 N1 stained red, and nuclei stained with 4',6-diamidino-2-phenylindole (DAPI; blue)); scale bar, 100 μ m. **c–e**, Histopathologic scoring of bronchiolar necrosis (**c**), alveolar

damage (**d**) and fibrin deposition (**e**) in lungs from *APOE*-knock-in mice on day 4 post infection with SARS-CoV-2 MA10 (data pooled from two independent experiments; *n* = 18, 22 and 15 for *APOE2*, *APOE3* and *APOE4*, respectively; *P* values according to two-sided Mann–Whitney tests). **f**, Representative images for **c–e**. Black arrowheads, bronchiolar epithelial necrosis; asterisks, fibrin; white arrowheads, interstitial and perivascular inflammation; arrows, endothelialitis. Scale bars, 1,000 μ m (top row) and 400 μ m (middle and bottom rows). Insets magnify pathological findings as labelled. Boxplot whiskers in **a** and **b** extend to the smallest and largest value within 1.5 times the interquartile ranges of the hinges, and box centre and hinges indicate median and first and third quartiles, respectively.

Data Fig. 2b,c). No differences were observed for inflammatory infiltrates in the pulmonary interstitium and vessels in *APOE2* or *APOE4* mice in comparison with *APOE3* mice (Extended Data Fig. 2d–h). These results indicate accelerated COVID-19 progression in *APOE2* and *APOE4* mice relative to *APOE3*, with histopathologic features evident by day 4 primarily in *APOE4* mice.

APOE impacts antiviral immunity and viral infection

We next carried out transcriptional profiling of homogenized lungs of non-infected *APOE*-knock-in mice and those of *APOE*-knock-in mice on days 2 and 4 post infection with SARS-CoV-2 MA10 (Fig. 3a). To identify clusters of highly correlated genes and relate their expression to genotype and time point relative to infection, we used weighted gene coexpression network analysis²¹ (Methods). This analysis revealed five modules of coexpressed genes that were significantly correlated with *APOE* genotype, and ten modules correlated with time point relative to infection (Fig. 3b). Assessment of the trajectories of the eigengene of these modules (a metric summarizing the weighted overall expression of a module) revealed modules that became specifically upregulated or downregulated in *APOE3* relative to *APOE2* and *APOE4* mice during disease progression. This was most pronounced for modules greenyellow and midnightblue (upregulated in *APOE3* relative to *APOE2* and *APOE4*) and black (downregulated in *APOE3* relative to *APOE2* and *APOE4*; Fig. 3c and Extended Data Fig. 3a), and these results were validated in an independent cohort of mice (Extended Data Fig. 3b,c). Correlations of the greenyellow, midnightblue and pink modules with *APOE* genotype were also significant in an independent third cohort of young female mice (Extended Data Fig. 3d,e). Pathway analysis of the black module that exhibited a higher level of expression in *APOE2* and *APOE4* relative to *APOE3* mice on day 4 revealed enrichment of genes implicated in blood coagulation and haemostasis, abnormalities of which are frequent in severe COVID-19 (ref.²²) (Extended Data Fig. 3f,g). Notably, analysis of the modules that exhibited downregulation in *APOE2* and *APOE4* mice relative to *APOE3* on day 4 (greenyellow, midnightblue and yellow) showed enrichment of genes implicated in T and B cell activation as well as positive immune response regulation (Extended Data Fig. 3h–o). Immunofluorescence staining indicated overall similar

levels of CD45⁺ leukocytes in *APOE2* and *APOE4* relative to *APOE3* mice (Extended Data Fig. 3p). These data are consistent with dampened adaptive immunity during early response to COVID-19 in *APOE2* and *APOE4* relative to *APOE3* mice.

Consistent with these findings, the results of flow cytometry on dissociated lungs on day 4 post infection confirmed an expansion of myeloid cells and relative depletion of lymphoid cells in the lungs of both *APOE2* and *APOE4* relative to *APOE3* mice (Fig. 3d and Extended Data Fig. 4a,b). In humans with severe COVID-19, depletion of lymphoid subsets has also been observed in the peripheral blood^{23,24}. To assess whether these changes were recapitulated by our animal model, we carried out flow cytometry on peripheral blood of *APOE*-knock-in mice. Although total leukocyte numbers were not significantly different between *APOE* genotypes, both *APOE2* and *APOE4* mice showed expansion of myeloid cells mainly driven by Ly6G⁺ neutrophils with concomitant contraction of all major lymphoid populations (Extended Data Fig. 4c–g). These data are consistent with the reported elevation of myeloid/lymphoid ratios in patients with adverse COVID-19 outcomes^{23,24} and suggest that adaptive immune responses are blunted in *APOE2* and *APOE4* mice during early COVID-19 progression.

To further profile the immunological response in *APOE*-knock-in mice during COVID-19, we carried out single-cell RNA sequencing on a total of 41,500 cells (post-filtering) from 29 mice across all 3 genotypes with and without COVID-19 (Extended Data Fig. 5a–d). Infected mice showed a marked expansion of myeloid cells, which, consistent with our flow cytometry data, was more prominent in *APOE2* and *APOE4* relative to *APOE3* mice (Fig. 3e,f and Extended Data Fig. 6). To assess changes in the functional status of cell clusters, we performed gene set enrichment analysis. Notably, *APOE2* mice showed more pronounced enrichment of various immune-related pathways relative to *APOE3* in comparison to *APOE4* relative to *APOE3* mice (Fig. 3g). In humans, hyperactivation of proinflammatory signalling has been implicated in adverse outcomes^{23,25}. We therefore reasoned that despite a similar change in immune subset abundances during early infection, antiviral immune responses might diverge between *APOE2* and *APOE4* mice over the course of infection. To test this, we assessed the generation of virus spike-specific CD8⁺ T cells during infection (Fig. 3h). The fraction of virus spike-specific CD8⁺ T cells as

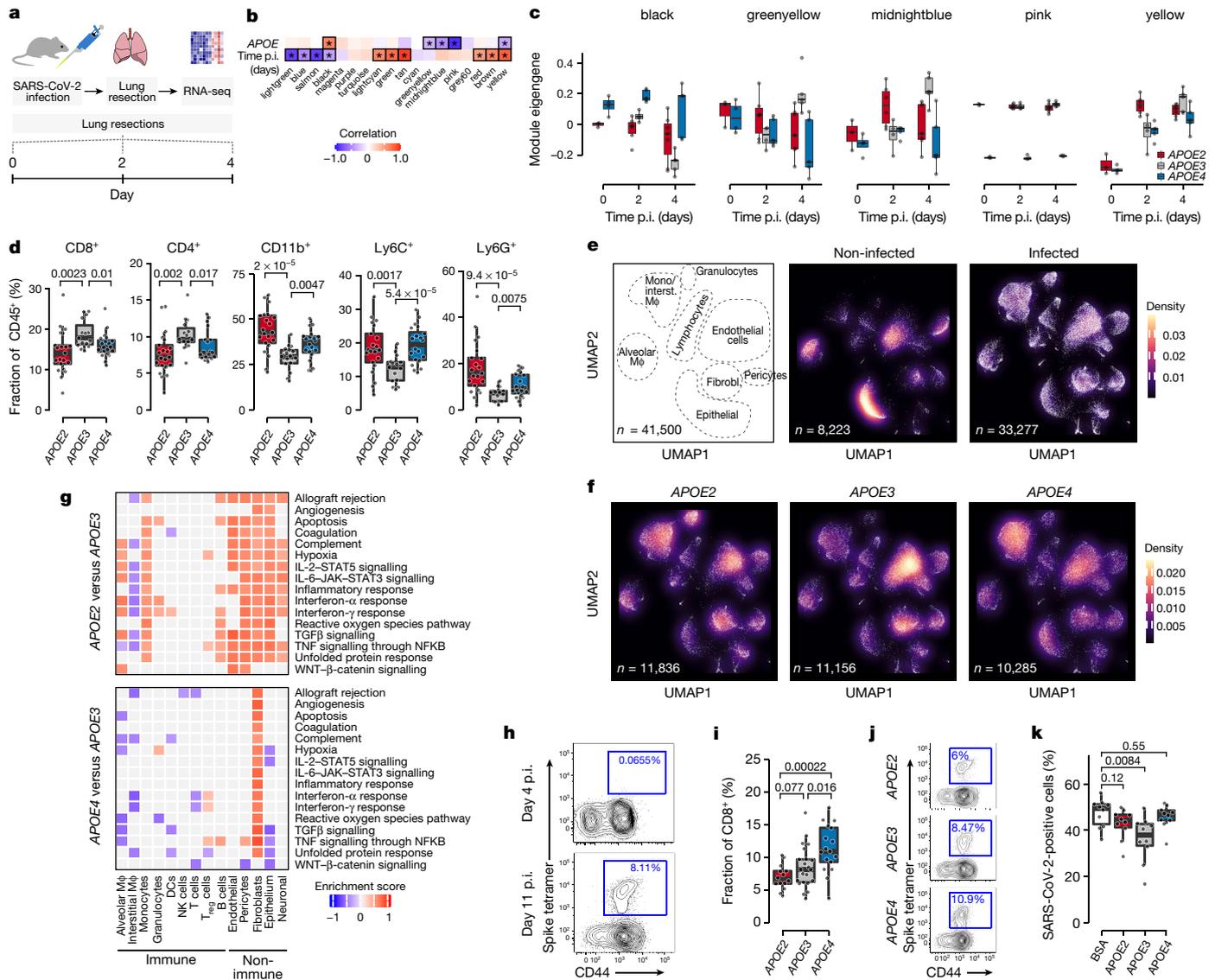


Fig. 3 | *APOE* genotype affects COVID-19 progression through immune modulation and altered viral infection. **a**, Method for transcriptional profiling of lungs from SARS-CoV-2 MA10-infected mice. RNA-seq, RNA sequencing. **b**, Correlation of module eigengenes with time after infection and *APOE* genotype ordered by its impact on COVID-19 survival (*APOE3* > *APOE2* > *APOE4*); asterisks indicate significant correlations (Pearson correlation tests). **c**, Module eigengene trajectories for modules significantly correlating with *APOE* genotype ($n = 4$ (*APOE2*, day 0); *APOE3*, day 2), 3 (*APOE4*, day 0), 6 (*APOE2*, day 2), 5 (*APOE4*, day 2); *APOE3*, day 4; *APOE4*, day 4) and 7 (*APOE2*, day 4). **d**, Flow cytometry for indicated cells in lungs of *APOE*-knock-in mice on day 4 post infection ($n = 21, 15$ and 20 for *APOE2*, *APOE3* and *APOE4*, respectively; data pooled from two independent experiments; P values according to one-tailed t -tests). **e, f**, Density plots of 41,500 RNA-sequenced lung cells from *APOE*-knock-in mice stratified by infection status (**e**) or *APOE* genotype in infected mice (**f**). M ϕ , macrophage; UMAP, uniform manifold approximation and projection. **g**, Gene set enrichment

assessed by tetramer staining was significantly larger in *APOE4* relative to *APOE3* and *APOE2* mice, consistent with *APOE4* mice eventually mounting more effective adaptive antiviral immunity than *APOE2* mice (Fig. 3i,j). These data indicate that although both *APOE2* and *APOE4* mice initially exhibited blunted adaptive immune responses, *APOE4* mice generated more robust antiviral T cell responses in later stages of infection, which emerged after pathological tissue damage had occurred.

analysis for grouped clusters from **f**; grey boxes indicate no significant enrichment. DCs, dendritic cells; NK cells, natural killer cells; T_{reg} cells, regulatory T cells. **h**, Representative flow cytometry plots of tetramer-positive CD8⁺ T cells on days 4 and 11 post infection (independent experiments). **i**, Proportion of tetramer-positive CD8⁺ T cells on day 11 post infection ($n = 14, 22$ and 17 for *APOE2*, *APOE3* and *APOE4*, respectively; data pooled from two independent experiments, P values according to two-tailed t -tests; note that some mice died during the course of infection). **j**, Representative samples for **i**. **k**, Fraction of infected cells after incubation with SARS-CoV-2 in the presence of the indicated proteins ($n = 10$ per group; representative of three independent experiments; P values according to two-tailed t -tests). Boxplot whiskers in **c, d, i, k** extend to the smallest and largest value within 1.5 times the interquartile ranges of the hinges, and box centre and hinges indicate median and first and third quartiles, respectively.

We next assessed whether *APOE* directly affects viral infection, potentially explaining the emergence of differences in viral titre and immune responses early following infection. Remarkably, recombinant *APOE3*, but not recombinant *APOE2* or *APOE4*, significantly suppressed infection of Huh-7.5 cells in vitro (Fig. 3k). In sum, these data indicate that adverse outcomes in *APOE2* and *APOE4* mice are driven by both enhanced viral infection and dampened adaptive antiviral immunity.

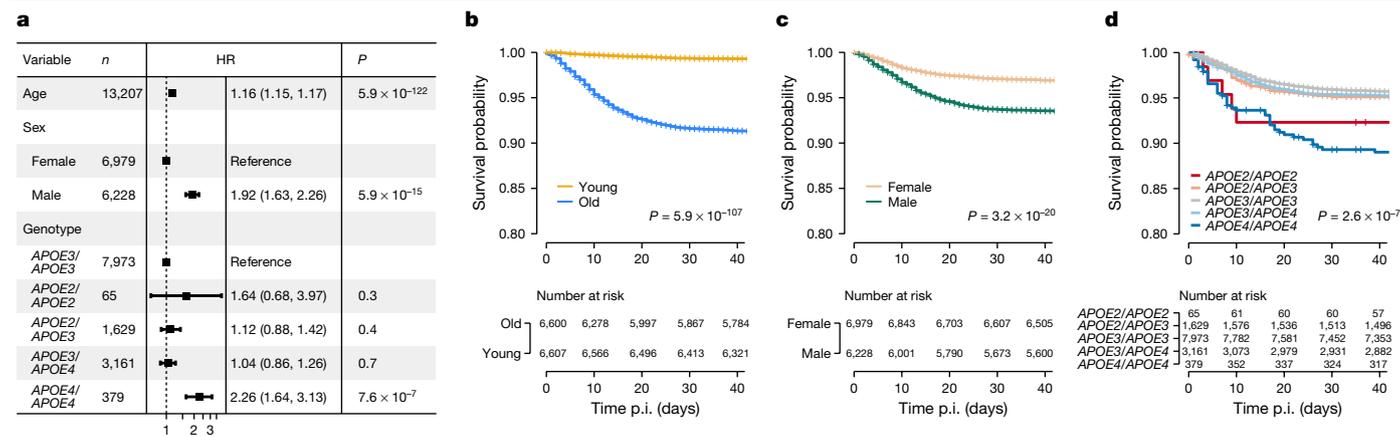


Fig. 4 | APOE germline variants are associated with survival in humans with SARS-CoV-2 infection. **a**, Multivariate analysis of the impact of age, sex and APOE genotype on survival of patients with SARS-CoV-2 infection in the UK Biobank (P values according to multivariable Cox proportional hazards model,

error bars indicate 95% confidence intervals, $n = 13,207$). **b–d**, Survival of patients from **a** stratified by age below or above median (**b**), sex (**c**) and APOE genotype (**d**). P values in **b–d** according to log-rank tests. n , sample size; HR, hazard ratio for death.

APOE genotype and survival in patients with COVID-19

To assess the impact of APOE genotype on COVID-19 outcome in humans, we analysed participants of the UK Biobank²⁶. The overall distribution of APOE genotype in 402,763 UK Biobank participants was comparable to that of similarly aged individuals in the Atherosclerosis Risk in Communities study²⁷, with approximately 40% carrying at least one copy of the APOE2 or APOE4 allele (Extended Data Fig. 7a,b). Consistent with previous reports carried out at earlier times during the pandemic^{28,29}, our observations show a moderate enrichment of APOE4 homozygosity in participants with positive versus negative test results and in participants with positive test results versus the remaining participants (Extended Data Fig. 7c,d). There was no significant difference in APOE genotype distribution between patients with a positive test regarding the test origin (inpatient versus outpatient; Extended Data Fig. 7e).

We next carried out survival analysis of patients with confirmed SARS-CoV-2 infection. Consistent with known epidemiological observations, multivariate analysis confirmed male sex and advanced age to confer adverse survival outcomes (Fig. 4a–c). Notably, patients homozygous for APOE4 also exhibited poor survival with a more than twofold increased hazard ratio for death relative to APOE3 homozygous patients (Fig. 4a,d). Patients homozygous for APOE2 also experienced an increased hazard ratio for death that did not reach statistical significance (Fig. 4a,d). The association between APOE genotype and survival remained significant on adjustment for the first ten principal components of genetic variation, indicating population structure to be unlikely to account for this association (Extended Data Fig. 8a–c). Consistently, the association of APOE with COVID-19 was maintained on restriction of the analysis to individuals of European ancestry (Extended Data Fig. 8d–g). No significant association of APOE genotype with survival was detected over a similar period before the start of the COVID-19 pandemic, indicating that the known association of APOE genotype with longevity also does not confound these results (Extended Data Fig. 8h). Overall, these results are consistent with our animal studies that demonstrate a causal role of APOE genotype in modulating mouse COVID-19 outcome. While the present work was in revision, an independent study validated the epidemiologic association of APOE4 with adverse outcomes in COVID-19 in the large FinnGen cohort³⁰.

Discussion

The COVID-19 pandemic has had a devastating impact on public health, but individual outcomes are markedly heterogeneous. Comprehensive

efforts have been made to uncover the genetic basis of COVID-19 outcome. These efforts were carried out using either genome-wide or candidate gene approaches and identified genetic variants and regions epidemiologically associated with COVID-19 outcome^{8,10–12,14,15,31}. However, whether common germline variants could causally modulate COVID-19 outcomes in vivo is unknown. In this work, we undertook a reverse genetic approach and specifically focused on APOE variants given their previously established roles in modulating immunity. By using genetic mouse models of human APOE germline variation, we established a causal link between APOE genotype and COVID-19 outcome in mice, supported by clinical association data in humans. Notably, our focused genetic and biochemical studies of these APOE variants led us to assess their epidemiological associations with human outcomes. Although previous genome-wide association studies for COVID-19 critical illness have not detected associations with variants in APOE that reached genome-wide significant threshold levels, our data on APOE variant association with survival in patients with COVID-19 in a candidate analysis are supported by the reverse genetic approach in mice, suggesting a potential causal relationship between APOE4 genotype and COVID-19 outcome in human disease.

We uncovered two mechanisms underlying APOE-genotype-dependent differences in mouse COVID-19 outcomes: both APOE2 and APOE4 mice showed impaired immune responses during early infection. Single-cell transcriptional profiling indicated hyperactivation of proinflammatory signalling in APOE2 relative to APOE3 and APOE4 mice. In addition, APOE4 mice exhibited increased expansion of virus-specific CD8⁺ T cells during later stages of infection, indicating that antiviral T cell responses diverge between APOE2 and APOE4 during later infection stages. In addition to these effects on antiviral immunity, we found that recombinant APOE3, but not recombinant APOE2 or APOE4, inhibited viral infection in vitro. These findings are consistent with a previous study demonstrating increased infection of APOE4 relative to APOE3 neurons and astrocytes³². Although this past study's findings could be interpreted as APOE4 enhancing infection of neurons and astrocytes relative to APOE3, we interpret our findings as APOE3 repressing infection in contrast to APOE2 and APOE4. Our data indicate that adverse outcomes in APOE2 and APOE4 mice may be mediated by both enhanced viral infection and maladaptive immunity during early infection, with APOE4 mice ultimately generating more robust antiviral T cell immunity than APOE2 mice.

It will be important to further dissect the mechanistic basis of how these variants exert detrimental effects on COVID-19 outcome at a molecular level in future studies. APOE has been shown to directly

modulate both innate and adaptive immune responses^{2,33,34}, providing potential clues towards its molecular mechanism of action in immune modulation. In addition, a genetic screen identified cholesterol metabolism to affect SARS-CoV-2 infection^{35–37}, and SARS-CoV-2 may bind directly to APOE³⁸, providing starting points for further mechanistic studies focused on how APOE affects viral infection. It is important to note that the effects of APOE variants seem to be disease-context specific, with *APOE2* and *APOE4* conferring beneficial and/or detrimental outcomes depending on phenotype^{3,18,19,39–41}. Moreover, the dual impact of *APOE* genetic variation on COVID-19 and Alzheimer's outcomes has implications for understanding the neurocognitive changes imparted by both disorders.

Our findings have several potential clinical implications. First, prospective clinical studies are warranted to determine whether *APOE* genotyping could be used for risk stratification in SARS-CoV-2 and perhaps other virus infections. Such genotyping may allow future patients to benefit from more aggressive preventative and therapeutic approaches, including early booster vaccinations, antiviral drugs and monoclonal antibody therapies. The impact of vaccination or prior infection history on *APOE* genotype dependence of COVID-19 outcomes will need to be determined. Additionally, it will be important to assess vaccination efficacy in individuals of distinct *APOE* genotypes. More generally, our work confirms that common genetic variation can give rise to heterogeneous outcomes 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/s41586-022-05344-2>.

- Farrer, L. A. et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease: a meta-analysis. *JAMA* **278**, 1349–1356 (1997).
- Tavazoie, M. F. et al. LXR/ApoE activation restricts innate immune suppression in cancer. *Cell* **172**, 825–840 (2018).
- Ostendorf, B. N. et al. Common germline variants of the human *APOE* gene modulate melanoma progression and survival. *Nat. Med.* **26**, 1048–1053 (2020).
- Mahley, R. W. Apolipoprotein E: from cardiovascular disease to neurodegenerative disorders. *J. Mol. Med.* **94**, 739–746 (2016).
- Shi, Y. & Holtzman, D. M. Interplay between innate immunity and Alzheimer disease: APOE and TREM2 in the spotlight. *Nat. Rev. Immunol.* **18**, 759–772 (2018).
- Williamson, E. J. et al. Factors associated with COVID-19-related death using OpenSAFELY. *Nature* **584**, 430–436 (2020).
- Zhang, Q., Bastard, P., Cobat, A. & Casanova, J.-L. Human genetic and immunological determinants of critical COVID-19 pneumonia. *Nature* <https://doi.org/10.1038/s41586-022-04447-0> (2022).
- Zhang, Q. et al. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science* **370**, eabd4570 (2020).
- Asano, T. et al. X-linked recessive TLR7 deficiency in ~1% of men under 60 years old with life-threatening COVID-19. *Sci. Immunol.* **6**, eabl4348 (2021).
- The Severe Covid-19 GWAS Group. Genomewide association study of severe Covid-19 with respiratory failure. *New Engl. J. Med.* **383**, 1522–1534 (2020).
- Pairo-Castineira, E. et al. Genetic mechanisms of critical illness in COVID-19. *Nature* **591**, 92–98 (2021).
- COVID-19 Host Genetics Initiative. Mapping the human genetic architecture of COVID-19. *Nature* **600**, 472–477 (2021).
- Nakanishi, T. et al. Age-dependent impact of the major common genetic risk factor for COVID-19 on severity and mortality. *J. Clin. Investig.* **131**, e152386 (2021).

- Kousathanas, A. et al. Whole-genome sequencing reveals host factors underlying critical COVID-19. *Nature* **607**, 97–103 (2022).
- Pathak, G. A. et al. A first update on mapping the human genetic architecture of COVID-19. *Nature* **608**, E1–E10 (2022).
- Ludewig, B. et al. Hypercholesterolemia exacerbates virus-induced immunopathologic liver disease via suppression of antiviral cytotoxic T cell responses. *J. Immunol.* **166**, 3369–3376 (2001).
- Toledo, A., Monzón, J. D., Coleman, J. L., Garcia-Monco, J. C. & Benach, J. L. Hypercholesterolemia and ApoE deficiency result in severe infection with Lyme disease and relapsing-fever *Borrelia*. *Proc. Natl Acad. Sci. USA* **112**, 5491–5496 (2015).
- Strittmatter, W. J. et al. Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proc. Natl Acad. Sci. USA* **90**, 1977–1981 (1993).
- Corder, E. H. et al. Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. *Nat. Genet.* **7**, 180–184 (1994).
- Leist, S. R. et al. A mouse-adapted SARS-CoV-2 induces acute lung injury and mortality in standard laboratory mice. *Cell* **183**, 1070–1085 (2020).
- Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* **9**, 559 (2008).
- Perico, L. et al. Immunity, endothelial injury and complement-induced coagulopathy in COVID-19. *Nat. Rev. Nephrol.* **17**, 46–64 (2021).
- Lucas, C. et al. Longitudinal analyses reveal immunological misfiring in severe COVID-19. *Nature* **584**, 463–469 (2020).
- Mathew, D. et al. Deep immune profiling of COVID-19 patients reveals distinct immunotypes with therapeutic implications. *Science* **369**, eabc8511 (2020).
- Lee, J. S. et al. Immunophenotyping of COVID-19 and influenza highlights the role of type I interferons in development of severe COVID-19. *Sci. Immunol.* **5**, eabd1554 (2020).
- Sudlow, C. et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
- Blair, C. K. et al. APOE genotype and cognitive decline in a middle-aged cohort. *Neurology* **64**, 268–276 (2005).
- Del Ser, T. et al. Residence, clinical features, and genetic risk factors associated with symptoms of COVID-19 in a cohort of older people in Madrid. *Gerontology* **67**, 281–289 (2021).
- Kuo, C.-L. et al. APOE e4 genotype predicts severe COVID-19 in the UK Biobank community cohort. *J. Gerontol. A* **75**, 2231–2232 (2020).
- Kurki, S. N. et al. APOE e4 associates with increased risk of severe COVID-19, cerebral microhaemorrhages and post-COVID mental fatigue: a Finnish biobank, autopsy and clinical study. *Acta Neuropathol. Commun.* **9**, 199 (2021).
- Kosmicki, J. A. et al. Pan-ancestry exome-wide association analyses of COVID-19 outcomes in 586,157 individuals. *Am. J. Hum. Genet.* **108**, 1350–1355 (2021).
- Wang, C. et al. ApoE-isoform-dependent SARS-CoV-2 neurotropism and cellular response. *Cell Stem Cell* **28**, 331–342 (2021).
- Bonacina, F. et al. Myeloid apolipoprotein E controls dendritic cell antigen presentation and T cell activation. *Nat. Commun.* **9**, 3083 (2018).
- Laskowitz, D. T., Lee, D. M., Schmechel, D. & Staats, H. F. Altered immune responses in apolipoprotein E-deficient mice. *J. Lipid Res.* **41**, 613–620 (2000).
- Wang, R. et al. Genetic screens identify host factors for SARS-CoV-2 and common cold coronaviruses. *Cell* **184**, 106–119 (2021).
- Schneider, W. M. et al. Genome-scale identification of SARS-CoV-2 and pan-coronavirus host factor networks. *Cell* **184**, 120–132 (2021).
- Hoffmann, H.-H. et al. Functional interrogation of a SARS-CoV-2 host protein interactome identifies unique and shared coronavirus host factors. *Cell Host Microbe* **29**, 267–280 (2021).
- Yin, Y., Sheng, Y., Wang, M., Ma, Y. & Ding, H. Interaction of serum proteins with SARS-CoV-2 RBD. *Nanoscale* <https://doi.org/10.1039/D1NR02687A> (2021).
- Wozniak, M. A. et al. Apolipoprotein E-c4 protects against severe liver disease caused by hepatitis C virus. *Hepatology* **36**, 456–463 (2002).
- van Exel, E. et al. Effect of APOE e4 allele on survival and fertility in an adverse environment. *PLoS ONE* **12**, e0179497 (2017).
- Burt, T. D. et al. Apolipoprotein (apo) E4 enhances HIV-1 cell entry in vitro, and the APOE e4/e4 genotype accelerates HIV disease progression. *Proc. Natl Acad. Sci. USA* **105**, 8718–8723 (2008).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2022

Article

Methods

Cell lines

VeroE6 cells (*Chlorocebus sabaues*; sex: female, kidney epithelial) obtained from the ATCC (CRL-1586) and Ralph Baric (University of North Carolina at Chapel Hill), Caco-2 cells (*Homo sapiens*, sex: male, colon epithelial) obtained from the ATCC (HTB-37) and Huh-7.5 hepatoma cells (*H. sapiens*; sex: male, liver epithelial)⁴² were cultured in Dulbecco's modified Eagle medium supplemented with 1% nonessential amino acids and 10% fetal bovine serum (FBS) at 37 °C and 5% CO₂. All cell lines were tested negative for contamination with mycoplasma.

Virus propagation and titration

The SARS-CoV-2 MA10 was provided by Ralph Baric (University of North Carolina at Chapel Hill). A P1 stock was amplified in VeroE6 cells obtained from the ATCC that were engineered to stably express TMPRSS2 (VeroE6-TMPRSS2). To generate a P2 working stock, VeroE6-TMPRSS2 cells were infected at a multiplicity of infection (MOI) of 0.1 plaque-forming units (PFUs) per cell and incubated at 37 °C for 4 days. The virus-containing supernatant was subsequently collected, clarified by centrifugation (3,000g for 10 min) and filtered using a disposable vacuum filter system with a 0.22-µm membrane. Virus stock titres were measured by a standard plaque assay on Huh-7.5 cells that stably express ACE2 and TMPRSS2 (Huh-7.5-ACE2/TMPRSS2) and on VeroE6 cells obtained from Ralph Baric (referred to as VeroE6-UNC). In brief, 500 µl of serial tenfold virus dilutions in Opti-MEM were used to infect 4×10^5 cells seeded the day before into wells of a 6-well plate. After 90 min of adsorption, the virus inoculum was removed, and cells were overlaid with Dulbecco's modified Eagle medium containing 10% FBS with 1.2% microcrystalline cellulose (Avicel). Cells were incubated for 4 days at 33 °C, followed by fixation with 7% formaldehyde and crystal violet staining for plaque enumeration. SARS-CoV-2, strain USA-WA1/2020, was obtained from BEI Resources and amplified in Caco-2 cells. Caco-2 cells were infected at a MOI = 0.05 PFU per cell and incubated for 6 days at 37 °C. The virus-containing supernatant was subsequently collected, clarified by centrifugation (3,000g for 10 min) and stored at -80 °C. Viral titres were measured on Huh-7.5 cells by standard plaque assay as described above. All SARS-CoV-2 and SARS-CoV-2 MA10 experiments were performed in a biosafety level 3 (BSL-3) laboratory.

To confirm virus identity and evaluate for unwanted mutations that were acquired during the amplification process, RNA from virus stocks was purified using TRIzol reagent (Thermo Fisher, number 15596026). In brief, 200 µl of each virus stock was added to 800 µl TRIzol reagent, followed by 200 µl chloroform, which was then centrifuged at 12,000g for 5 min. The upper aqueous phase was transferred to a new tube, mixed with an equal volume of isopropanol, and then added to RNeasy Mini Kit columns (Qiagen, number 74014) to be further purified following the manufacturer's instructions. Viral stocks were subsequently confirmed through next-generation sequencing using libraries for Illumina MiSeq.

Animal studies

All animal experiments were conducted in accordance with a protocol approved by the Institutional Animal Care and Use Committee at The Rockefeller University, including the use of SARS-CoV-2 MA10 virus under BSL-3 conditions. Human *APOE2* (strain no. 1547), *APOE3* (no. 1548) and *APOE4* (no. 1549) targeted replacement (knock-in) mice on C57Bl/6 background were obtained from Taconic Biosciences.

SARS-CoV-2 MA10 in vivo infections

All infection experiments were performed in a dedicated BSL-3 facility at The Rockefeller University at negative pressure. Staff performing experiments were protected by wearing Tyvek suits connected to powered air-purifying respirators. Mice were intranasally infected with

14,700 PFU (based on titration in VeroE6-UNC cells) of SARS-CoV-2 MA10 in a volume of 30 µl under anaesthesia with a combination of ketamine and xylazine. *APOE*-knock-in mice were infected between 7 and 45 weeks of age as indicated in the figures. Experimental cohorts were age matched. Mice were monitored daily for weight loss and general condition. Mice were recorded as dead when found dead in the cage or when meeting criteria for euthanasia as defined in the animal protocol, including when falling below 70% initial body weight. All infected mice were included in survival and weight analyses. Some mice were selected before infection for tissue collection on the days as indicated in the figure legends and censored for survival and weight analyses on the respective days. Mice were gently twirled before weighing to prevent measurement inaccuracies due to mouse movements. In addition, weight measurements were performed with investigators blinded for the genotype in two independent experiments that recapitulated the results of the overall large cohort.

RNA isolation from homogenized lungs

The right lung lobe was resected and homogenized in TRIzol (Thermo Fisher, number 15596026) in a gentleMACS dissociator (Miltenyi) according to the manufacturer's instructions (program RNA_01). Debris was removed by centrifugation (2,000g for 1 min), and RNA was isolated using the Direct-zol RNA purification kit (Zymo Research, no. R2050) including DNase digestion according to the manufacturer's instructions.

TaqMan quantitative real-time PCR

For quantification of SARS-CoV-2 MA10 titres from homogenized lungs, RNA was isolated as described above, reverse-transcribed and quantified using the TaqMan Fast Virus One Step Master Mix (Thermo Fisher, number 4444436) on a QuantStudio 5 system running QuantStudio Design and Analysis v1.4.3 (Thermo Fisher) according to the manufacturer's instructions. Primers for viral nucleocapsid were as recommended in the US Centers for Disease Control and Prevention diagnostic N1 assay (IDT, number 10006713), and 18S rRNA was used as housekeeping control (Thermo Fisher, number 4319413E).

Bulk RNA sequencing

For preparation of RNA-sequencing libraries, 250–500 ng of RNA isolated from homogenized lungs as outlined above was used as input for the Quantseq 3' FWD library preparation kit (Lexogen, number 015). For cohorts 1 and 2, age-matched 17–23-week-old male mice were used. For cohort 3, 7-week-old female mice were used. Libraries were sequenced on an Illumina NovaSeq sequencer (single end, 100-base-pair read length), and polyA and adapter sequences were trimmed using the BBDuk utility (v38.9; options $k = 13$, $ktrim = r$, $forcetrimleft = 11$, $use-shortkmers = t$, $mink = 5$, $qtrim = t$, $trimq = 10$, $minlength = 20$). As genome references, mouse (assembly GRCm38) and SARS-CoV-2 MA10 (Genbank accession number MT952602; ref.²⁰) genomes were concatenated, and trimmed reads were aligned using STAR aligner (v2.7.8a) with default settings, apart from `--outFilterMismatchNoverLmax 0.1`, as recommended by Lexogen (personal communication). STAR was also used for counting reads mapping to genes. Further analysis was performed using R (v4.1.0). Two samples (out of 88 samples total) were removed from analysis because of their identification as outliers on the basis of principal component analysis and/or SARS-CoV-2 MA10 transcript abundance.

Weighted gene coexpression network analysis and pathway analysis

Weighted gene coexpression network analysis⁴³ was carried out using the WGCNAR package (v1.70) to identify modules of coexpressed genes. Weighted gene coexpression network analysis identifies clusters of genes whose expression correlates with each other and relates these clusters to traits, such as *APOE* genotype and time point relative to

infection in our study. The module eigengene represents the first principal component of the expression matrix and can be used to summarize the (weighted average) expression of a module. Gene expression data were subjected to library size normalization and variance-stabilizing transformation using DESeq2 (v1.32.0), and the top 30% genes in terms of variance of expression were used as input for WGCNA. To compute the adjacency matrix for a signed coexpression network, a soft threshold power of 10 was used. To calculate correlations between traits and module eigengenes, *APOE* genotype was assigned values based on its impact on survival as shown in Fig. 1a, and the time point trait was assigned values in terms of days relative to infection. Hub genes were identified as the genes exhibiting the highest connectivity within a given model.

To assess enrichment of gene sets listed in the Gene Ontology biological processes, the clusterProfiler package for R (v4.0.0) was used to perform overrepresentation analysis based on a hypergeometric model.

Single-cell RNA sequencing

For single-cell RNA sequencing of lung-resident cells, mice were anaesthetized, and the pulmonary circulation was flushed with 5–10 ml ice-cold phosphate-buffered saline (PBS). The right lung lobe was dissociated using the lung dissociation kit (130-095-927, Miltenyi Biotec) with a gentleMACS dissociator according to the manufacturer's instructions (program 37C_m_LDK_1). Cells were strained using a 70- μ m filter, washed and pelleted, and red blood cells were lysed by incubation in ACK buffer (A10492, Gibco) for 2 min before neutralization with PBS. Cells were then strained again with a 40- μ m filter and processed using the cell fixation (SB1001) and single-cell whole-transcriptome (SB2001) kits from Parse Biosciences according to the manufacturer's instructions. This single-cell RNA-sequencing approach is based on combinatorial barcoding, which enabled us to multiplex lungs from a total of 29 mice representing each of the 3 *APOE* genotypes and conditions in the absence and presence of SARS-CoV-2 MA10 infection (Extended Data Fig. 5a). One of the eight resulting sublibraries was sequenced on an Illumina Nextseq 500 sequencer, and the other seven sublibraries were pooled and sequenced on an Illumina Novaseq sequencer (S2 flowcell) to an average depth of 65,256 reads per cell.

For data processing, the ParseBioscience processing pipeline (v0.9.6p) was used with default settings to align sequencing reads to the GRCh38 mouse genome and to demultiplex samples. In brief, each of the eight sublibraries was first processed individually using the command `split-pipe -mode all`, and the output of the eight sublibraries was combined using `split-pipe -mode combine`. Downstream processing was performed using the R package Seurat (v4.0.2) at default settings unless otherwise noted. Cells with fewer than 150 or more than 7,500 detected unique genes, more than 40,000 unique molecular identifiers, or more than 15% mitochondrial reads were excluded from analysis. The resulting gene–cell matrix was normalized and scaled using Seurat's `NormalizeData` and `ScaleData` functions and principal component analysis was performed with Seurat's `RunPCA` function; cells were clustered using the `FindNeighbors` (30 dimensions of reduction) and `FindClusters` (resolution = 1.4) functions; for visualizing clusters, `RunUMAP` (30 dimensions) was run. Wilcoxon rank-sum tests were performed to determine differentially expressed genes between clusters using the `FindAllMarkers` function (minimal fraction of 25% and log-transformed fold-change threshold of 0.25). The identity of cell clusters was determined by cross-referencing top differentially expressed transcripts with previous studies reporting on single-cell transcriptomes of the lung^{44–46}. Ambiguous cells with expression of distinct lineage markers were deemed to be likely multiplets and were excluded. Three clusters expressing T cell markers were characterized further using Seurat's `subset` function and reanalysed similarly to the main dataset, including running the `RunPCA`, `FindNeighbors` (20 dimensions), `FindClusters` (resolution = 0.5) and `RunUMAP` functions. Ambiguous cells from the subset were removed, and annotations

for the remaining clusters were added to the main dataset. For summary analyses, clusters were grouped as follows: alveolar macrophages A and B and proliferating alveolar macrophages as alveolar macrophages; monocytes A and B as monocytes; T cells naive, T cells and T cells proliferating as T cells; myofibroblasts, lipofibroblasts and *Col14a1*-expressing fibroblasts as fibroblasts; capillary endothelial cells, vascular endothelial cells A and B, other endothelial cells, and *Vcam1*-expressing endothelial cells A and B as endothelial cells; alveolar type 1, alveolar type 2, ciliated cells, airway epithelial A and B, and mesothelial cells as epithelial cells. In total, filtering low-quality and ambiguous cells resulted in 41,500 cells for analysis (of 50,104 cells before filtering).

For gene set enrichment analysis of the samples from infected mice, differentially expressed genes between either *APOE2* and *APOE3* or *APOE4* and *APOE3* were identified according to Wilcoxon rank-sum tests using Seurat's `FindMarkers` function. Genes were ranked using the metric $(-\log_{10}[P \text{ value}])/(\text{sign of } \log_2[\text{fold change}])$. The ranked gene lists were used as input for the GSEA function of the clusterProfiler R package (v4.0.0) to assess enrichment of selected immune-related pathways of the Hallmark gene set of the MSigDB database (<http://www.gsea-msigdb.org>).

Histological analysis and immunofluorescence staining

The left lung lobe was resected and fixed by submersion in 4% paraformaldehyde for 24 h at room temperature. Fixed lungs were embedded in paraffin and sectioned in 5- μ m-thick slices. Sections were dewaxed and rehydrated by incubation with xylene and descending ethanol concentrations and then either stained with haematoxylin–eosin for histological analysis or processed for immunofluorescence staining.

For immunofluorescence staining, samples were permeabilized with 0.1% Triton X-100 for 15 min. Antigen retrieval was performed by microwaving samples in Tris-EDTA buffer (Abcam, number ab93684) for 20 min. Samples were blocked by incubation with 5% goat serum in PBS with 0.1% Tween-20 (PBST) for 1 h. Subsequently, sections were stained with anti-CD45 (polyclonal, Abcam, number ab10558; 1:750) or anti-SARS nucleocapsid (polyclonal, Novus Biologicals, number 56576, 1:1,000) at 4 °C overnight. All antibodies were diluted in PBST with 5% goat serum. Slides were washed three times with PBS and stained with AF555-conjugated anti-rabbit antibody (1:200 in PBST, Thermo Fisher) for 45 min. Slides were washed with PBS and nuclei were counterstained with DAPI (1 μ g ml⁻¹, Roche) before mounting with Prolong Gold (Thermo Fisher). Images of lung sections were acquired using a Nikon A1R confocal microscope at 20 \times magnification using Nikon NIS elements software (v5.20.02). Images were quantified using CellProfiler (v4.2.1). Three to four randomly sampled fields of view per lung were analysed and averaged.

For histological analysis, haematoxylin–eosin-stained lung sections were evaluated and scored by a board-certified veterinary pathologist (S.E.C.) using a semiquantitative histopathology scoring system used in mouse models of SARS-CoV-2 (refs. 47,48). In brief, five random fields of the lung lobe at 200 \times total magnification were chosen and scored in a blinded manner for histopathological changes. Ordinal scores for lesion parameters were assigned using the following tiers: 0, within expected limits; 1, uncommon, <5%; 2, detectable in 5–33%; 3, detectable in 34–66%; and 4, detectable in >66% of lung fields. Tissues were graded for the presence of edema, haemorrhage, fibrin and/or necrotic debris in alveoli, bronchiolar epithelial necrosis, perivascular and interstitial inflammation and mononuclear cell infiltrates. Endothelial inflammation (endothelialitis) was evaluated by the extent of the lesion using the following ordinal scoring: 0, absent; 1, minor, solitary to loose adhesion or aggregation of leukocytes to the vascular endothelium with or without infiltration of leukocytes in the vascular wall (up to five blood vessels affected); 2, moderate, small to medium adhesion/aggregates and infiltration (six to ten blood vessels affected); and 3, severe, robust leukocytic aggregates

Article

and infiltrates around pulmonary vessels (more than ten blood vessels affected). Neutrophil cell infiltration (200–600× objective magnification) was scored as follows: 0, within normal limits; 1, scattered neutrophils sequestered in septa and/or infiltrating blood vessels; 2, no. 1 plus solitary neutrophils extravasated in alveolar spaces; 3, no. 2 plus small aggregates in blood vessels, alveolar spaces, and perivascular and peribronchiolar interstitium. An Olympus BX45 light microscope was used to capture images with a DP26 camera using cellSens Dimension software (v1.16). Lungs in SARS-CoV-2 MA10-infected mice exhibited multifocal areas of airway epithelial damage in bronchioles. Bronchioles had focal to multifocal changes characterized by segmental attenuation of bronchiolar epithelium with an accumulation of necrotic cellular debris, fibrin and sloughed epithelial cells, and occasional foamy macrophages in the airway lumina (Fig. 2f). Peribronchiolar interstitium was multifocally infiltrated by increased numbers of neutrophils and lymphocytes. The adjacent alveolar sacs and septae exhibited multifocal to coalescing areas of alveolar damage. Histological changes included hypercellular thickening of the alveolar septae caused by infiltrating leukocytes and congestion of alveolar capillaries, pneumocyte degeneration and necrosis, edema, fibrin strands and increased numbers of macrophages and scattered neutrophils and lymphocytes in alveolar spaces. Often, the vascular endothelium of pulmonary vessels was reactive with adherence and aggregation of leukocytes to the endothelium and transmigrating within vessel walls, indicative of endothelialitis.

Flow cytometry

All steps were performed on ice and under protection from light unless stated otherwise. Peripheral blood was obtained by submandibular bleedings, and red blood cells were lysed by incubation in ACK buffer (A10492, Gibco) for 3 min at room temperature before addition of PBS for neutralization. For flow cytometry of dissociated lungs, mice were anaesthetized, and the pulmonary circulation was flushed with 5–10 ml ice-cold PBS. The right lung lobe was then dissociated using the lung dissociation kit (130-095-927, Miltenyi Biotec) with a gentleMACS dissociator according to the manufacturer's instructions (program 37C_m_LDK_1). Cells were strained through a 70- μ m filter, washed and pelleted, and red blood cells were lysed by incubation in ACK buffer as indicated above before addition of PBS for neutralization. Cells were pelleted by centrifugation at 200g for 5 min and resuspended in staining buffer (25 mM HEPES, 2% FBS, 10 mM EDTA (351-027, Quality Biological) and 0.1% sodium azide (7144.8-16, Ricca) in PBS). To block Fc receptors, cells were incubated with 2.5 μ g ml⁻¹ anti-CD16/32 antibody in staining buffer (clone 93; 101320, BioLegend) before incubation with antibodies diluted in staining buffer for 20 min. After washing with PBS, cells were incubated with Zombie NIR Fixable Live/Dead stain (423105, BioLegend; 1:10,000 in PBS) for 15 min at room temperature, washed with staining buffer and fixed in 4% PFA. CountBright counting beads (C36950, Thermo Fisher) were added to the peripheral blood samples before analysis on an LSR Fortessa (BD Biosciences). For compensation, single-colour controls with UltraComp beads (01-2222-42, Thermo Fisher) for antibodies and amine-reactive beads (A10628, Thermo Fisher) for Zombie live–dead stain were used. The following anti-mouse fluorophore-conjugated antibodies were used: CD45–BV785 (clone 30-F11, catalogue number 103149, BioLegend, dilution: 1:3,000), CD11b–FITC (M1/70, 101206, BioLegend, 1:1,000), Ly6G–PerCP/Cy5.5 (1A8, 127616, BioLegend, 1:1,000), Ly6C–BV711 (HK1.4, 128037, BioLegend, 1:10,000), I-A/I-E-PE (M5/114.15.2, 107607, BioLegend, 1:10,000), CD19–PB (6D5, 115526, BioLegend, 1:500), CD19–BV421 (6D5, 115549, BioLegend, 1:500), NK1.1–APC (PK136, 17-5941-82, eBiosciences, 1:500), CD4–BV605 (GK1.5, 100451, BioLegend, 1:300), CD8 α –AF700 (53-6.7, 100730, BioLegend, 1:1,000). For staining of SARS-CoV-2 spike-specific CD8⁺ T cells, BV421-labelled SARS-CoV-2 S 539–546 tetramer was used (NIH Tetramer Core Facility, 1:200).

SARS-CoV-2 in vitro infections

The day before infection, Huh-7.5 cells were seeded into 96-well plates at a density of 7.5×10^3 cells per well. The next day, recombinant APOE2, APOE3, APOE4 (21-9195, 21-9189, 21-9190; Tonbo Biosciences) or BSA (A9576, Sigma) as control was added to the wells at a concentration of 10 μ g ml⁻¹, followed by infection with SARS-CoV-2 (WA1/2020) at an MOI of 0.01 PFU per cell. Cells were then incubated at 33 °C for 48 h. Next, they were fixed by adding an equal volume of 7% formaldehyde to the wells and subsequently permeabilized with 0.1% Triton X-100 for 10 min. After extensive washing, SARS-CoV-2-infected cells were incubated for 1 h at room temperature with blocking solution of 5% goat serum in PBS (005-000-12L, Jackson ImmunoResearch). A rabbit polyclonal anti-SARS-CoV-2 nucleocapsid antibody (GTX135357, GeneTex) was added to the cells at 1:1,000 dilution in blocking solution and incubated at 4 °C overnight. A goat anti-rabbit AlexaFluor 594 (A-11012, Life Technologies) was used as a secondary antibody at a 1:2,000 dilution. Nuclei were stained with Hoechst 33342 (62249, Thermo Fisher) at a 1 μ g ml⁻¹ dilution. Images were acquired with a fluorescence microscope and analysed using ImageXpress Micro XLS (Molecular Devices). All SARS-CoV-2 experiments were carried out in a BSL-3 laboratory.

Analysis of the UK Biobank

APOE genotyping results as determined by the rs7412 and rs429358 single nucleotide polymorphisms were downloaded from the UK Biobank⁴⁹. Clinical data, including SARS-CoV-2 test results and survival data, were downloaded from the UK Biobank data portal on 22 June 2021. For survival analyses, in patients with multiple tests the earliest positive test result was used as day zero of infection and COVID-19-associated death was recorded if the death cause was ICD10-coded as U07.1 or U07.2. Out of 502,619 patients, *APOE* genotype could be determined in 413,219 patients. A total of 77,221 participants had SARS-CoV-2 test results available, and 16,562 patients of these were tested positive at least once (Extended Data Fig. 7a). *APOE2/APOE4* heterozygous patients ($n = 10,456$) were excluded from analyses except for summary statistics shown in Extended Data Fig. 7a. For visualization purposes, survival data were truncated at 40 days. To account for genetic ancestry, the first ten genetic principal components as provided by the UK Biobank were included in a multivariate analysis. To restrict the analyses to individuals of European genetic ancestry, field 22006 provided by UK Biobank was used.

Statistical analysis

R v4.1.0 was used for data visualization and analyses. Statistical tests and sample sizes are listed in the respective figure legends. Unless otherwise noted, data are expressed as mean \pm standard error of the mean. For boxplots, hinges represent the first and third quartiles, whiskers extend to the smallest and largest value within 1.5 times the interquartile ranges of the hinges, and points represent individual mice. Survival analyses were performed using the R packages *survival* and *survminer*; summary tables were compiled using the *gtsummary* package. Multivariate analyses were performed according to a Cox proportional hazards model using the *survival* package and visualized with the *forestmodel* package. A significant difference was concluded at $P < 0.05$ in all figures.

Reporting summary

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

Data availability

Bulk RNA-sequencing and single-cell RNA-sequencing data have been deposited at the Gene Expression Omnibus under accession numbers

GSE184289 and GSE199498, respectively. All data from the UK Biobank is publicly available at <https://www.ukbiobank.ac.uk>. MSigDB is publicly available at <http://www.gsea-msigdb.org>. Source data are provided with this paper.

Code availability

Code is publicly available at https://github.com/benostendorf/ostendorf_et al_2022.

42. Blight, K. J., McKeating, J. A. & Rice, C. M. Highly permissive cell lines for subgenomic and genomic hepatitis C virus RNA replication. *J. Virol.* **76**, 13001–13014 (2002).
43. Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* **4**, <https://doi.org/10.2202/1544-6115.1128> (2005).
44. Angelidis, I. et al. An atlas of the aging lung mapped by single cell transcriptomics and deep tissue proteomics. *Nat. Commun.* **10**, 963 (2019).
45. Hurskainen, M. et al. Single cell transcriptomic analysis of murine lung development on hyperoxia-induced damage. *Nat. Commun.* **12**, 1565 (2021).
46. Melms, J. C. et al. A molecular single-cell lung atlas of lethal COVID-19. *Nature* **595**, 114–119 (2021).
47. Zheng, J. et al. COVID-19 treatments and pathogenesis including anosmia in K18-hACE2 mice. *Nature* **589**, 603–607 (2021).
48. Sun, J. et al. Generation of a broadly useful model for COVID-19 pathogenesis, vaccination, and treatment. *Cell* <https://doi.org/10.1016/j.cell.2020.06.010> (2020).
49. Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).

Acknowledgements We are grateful to members of the laboratory of S.F.T. and to D. Mucida for comments on previous versions of the manuscript. We thank Rockefeller University Resource Centers for assistance: V. Francis and other veterinary staff of the Comparative Bioscience

Center for animal husbandry and care, C. Zhao, C. Lai and staff at the Genomics Resource Center for assistance with RNA sequencing, and G. McNab for advice on biosafety measures. We thank the Laboratory of Comparative Pathology for histopathology support (with funding from NIH Core Grant P30 CA008748), the group of R. Baric for generating the mouse-adapted SARS-CoV-2 MA10 strain and making it available to the scientific community, and R. Yamin for experimental advice. We thank the NIH Tetramer Core Facility for providing BV421-labelled SARS-CoV-2 S 539–546 tetramer (H-2K(b)). This work was supported by the Pershing Square Innovation Fund. B.N.O. was supported by a Max Eder grant of the German Cancer Aid (reference 70114327) and the David Rockefeller Graduate Program in Bioscience, and is a fellow of the Digital Clinician Scientist Program at BIH-Charité. J.B. was supported by a scholarship from the German National Academic Foundation. M.A.P. was supported in part by the National Center for Advancing Translational Sciences, NIH, through Rockefeller University, Grant UL1 TR001866. S.F.T. was supported by the Black Family Foundation, The Rockefeller Center for Human Metastasis Research and The Breast Cancer Research Foundation. C.M.R. was supported by the G. Harold and Leila Y. Mathers Charitable Foundation, the Bawd Foundation and Fast Grants, a part of Emergent Ventures at the Mercatus Center, George Mason University. Human data analyses were conducted using the UK Biobank Resource under application number 62709.

Author contributions B.N.O. and S.F.T. conceived the study and wrote the manuscript. B.N.O., C.M.R. and S.F.T. supervised all research. B.N.O., J.B., M.A.P. and H.-H.H. conducted experiments. S.E.C. performed pathological analysis of lung sections. All authors edited and approved the manuscript.

Competing interests The authors declare no competing interests.

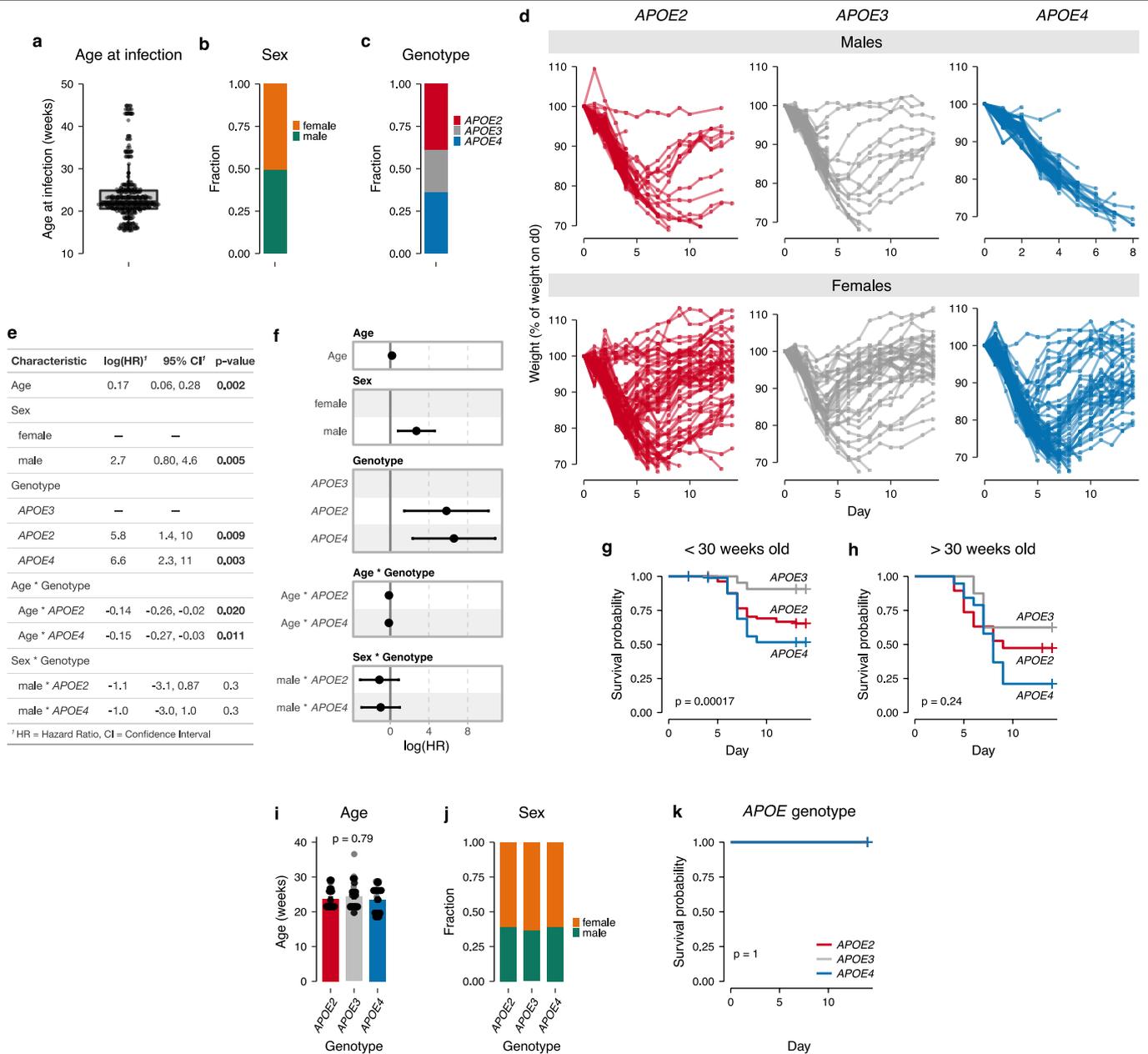
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-022-05344-2>.

Correspondence and requests for materials should be addressed to Benjamin N. Ostendorf or Sohail F. Tavazoie.

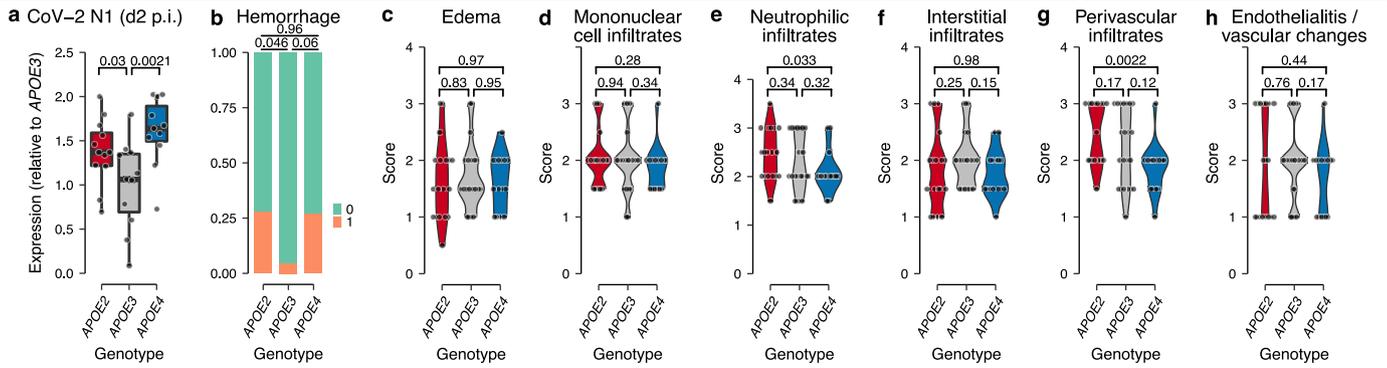
Peer review information *Nature* thanks Stanley Perlman and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



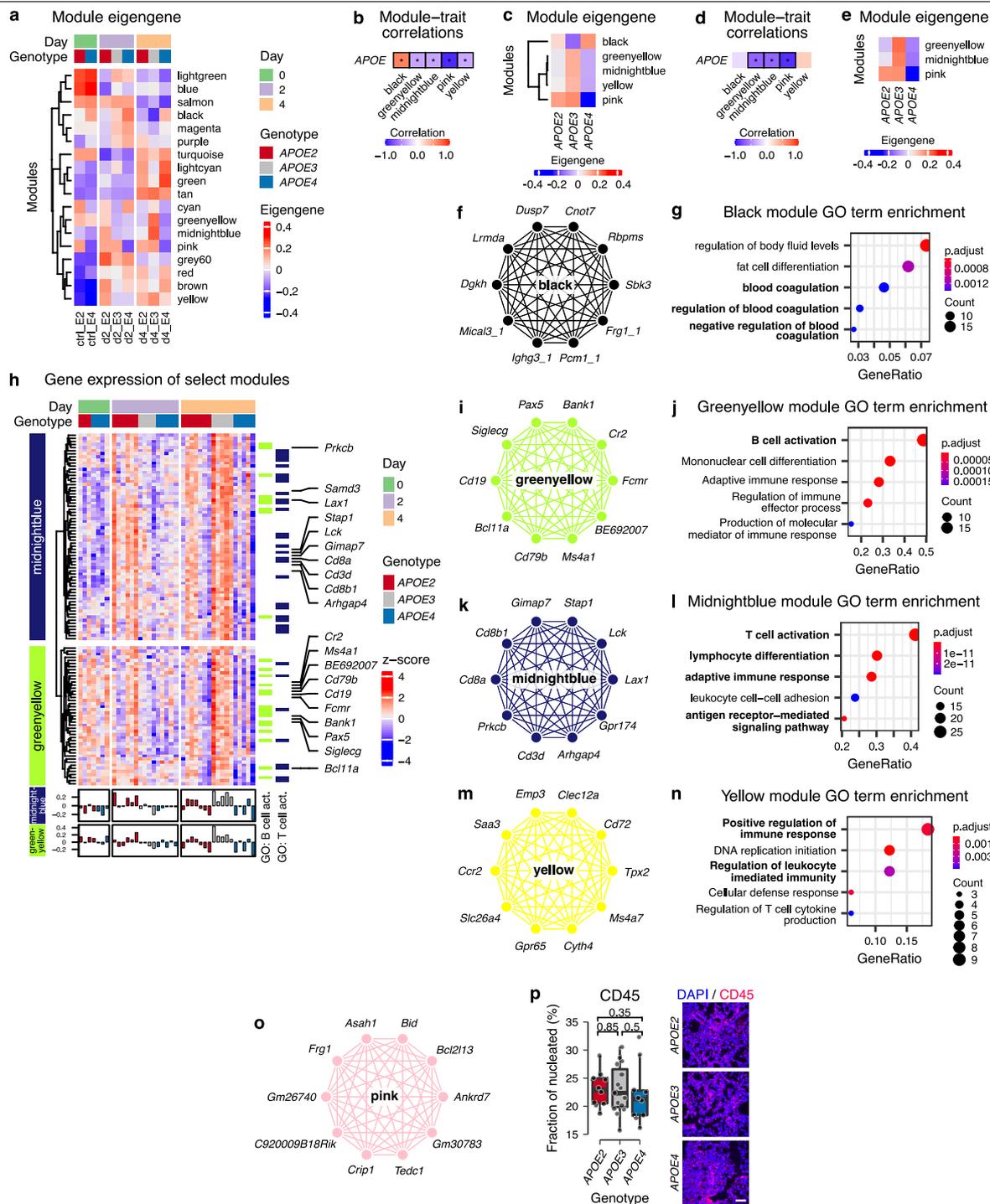
Extended Data Fig. 1 | Expanded characteristics of *APOE* knock-in mice infected with SARS-CoV-2 MA10. **a–c**, Distribution of age at infection (**a**), sex (**b**), and *APOE* genotype (**c**) of *APOE* knock-in mice infected with SARS-CoV-2 MA10 (n = 328; data pooled from 13 independent experiments). **d**, Individual weight course of male and female *APOE* knock-in mice infected with SARS-CoV-2 MA10 from (a). **e–f**, Multivariate analysis of the impact of age, sex, *APOE* genotype, and the interaction of age/*APOE* and sex/*APOE* on survival of SARS-CoV-2 MA10-infected *APOE* knock-in mice from (a) (P values according to multivariable Cox proportional hazards model; error bars in (f) denote 95% confidence intervals; n = 128, 82, and 118 for *APOE2*, *APOE3*, and *APOE4*,

respectively). **g–h**, Survival of young (< 30 weeks old) (**g**) and old (> 30 weeks old) (**h**) SARS-CoV-2 MA10-infected *APOE*-knock-in mice from (a) stratified by *APOE* genotype; P values according to log-rank tests. **i–k**, Age (**i**), sex distribution (**j**), and survival of non-infected *APOE* knock-in mice over a two-week period (**k**) (n = 67, 55, 67 for *APOE2*, *APOE3*, and *APOE4*, respectively; P values according to Kruskal-Wallis (**i**) and logrank (**k**) tests). Boxplot whiskers in (a) extend to the smallest and largest value within 1.5 × interquartile ranges of the hinges, and box centre and hinges indicate median and first and third quartiles, respectively.



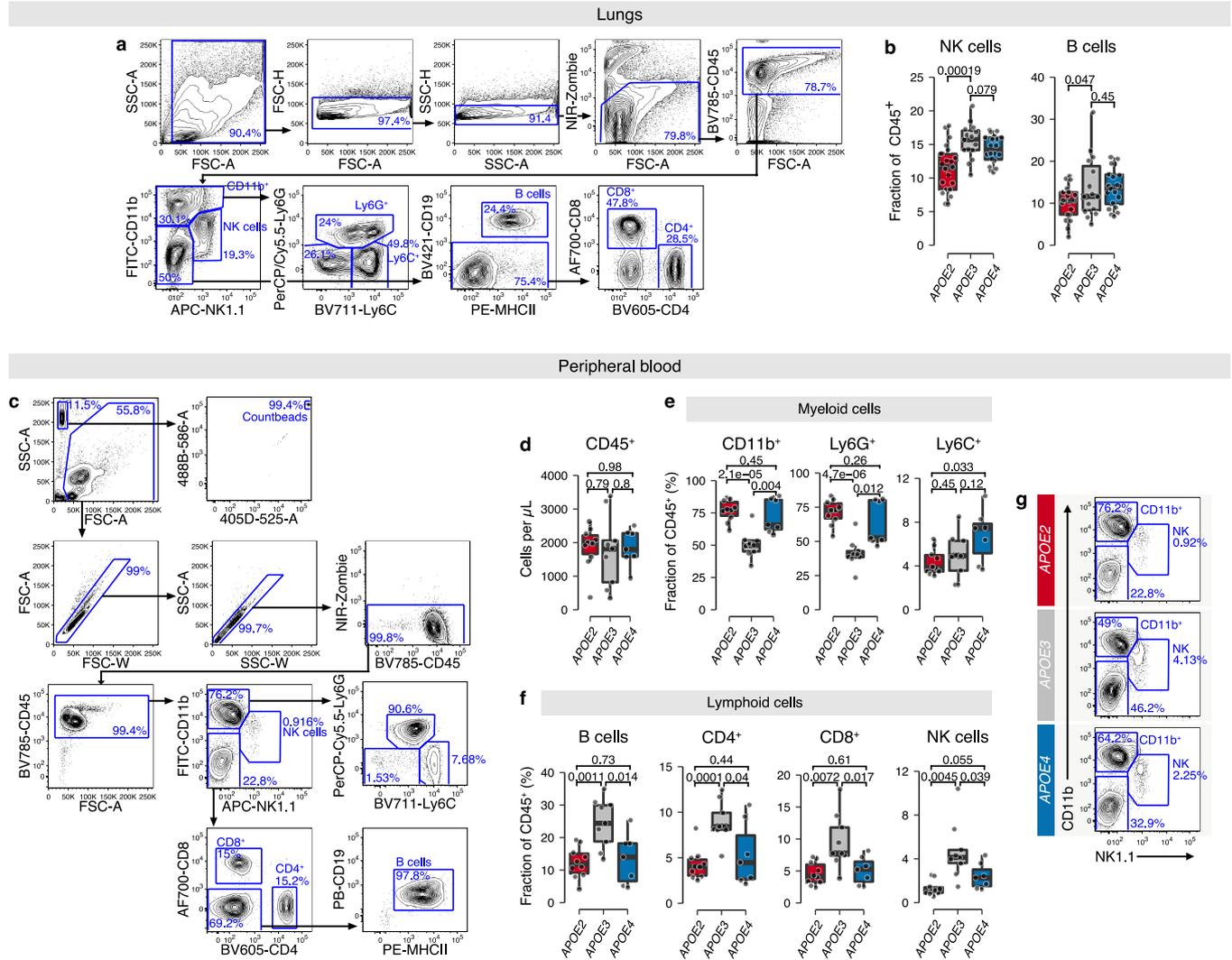
Extended Data Fig. 2 | Viral load early post infection and extended histopathologic analysis of lungs from SARS-CoV-2 MA10-infected *APOE* knock-in mice. **a**, TaqMan qPCR for SARS-CoV-2 N1 in homogenized lungs from *APOE* knock-in mice on day 2 post infection with SARS-CoV-2 MA10 (data pooled from two experiments; P values according to one-tailed Mann-Whitney test; n = 12, 11, 11 for *APOE2*, *APOE3*, and *APOE4*, respectively; boxplot whiskers extend to the smallest and largest value within $1.5 \times$ interquartile ranges of the

hinges, and box centre and hinges indicate median and first and third quartiles, respectively). **b-h**, Histopathologic scoring of lungs from *APOE* knock-in mice on day 4 post infection with SARS-CoV-2 MA10 for hemorrhage (**b**), edema (**c**), mononuclear cell infiltrates (**d**), neutrophilic cell infiltrates (**e**), interstitial infiltrates (**f**), perivascular infiltrates (**g**), and endothelialitis/vascular changes (**h**); P values according to two-sided Mann-Whitney-tests, n = 18, 22, 15 for *APOE2*, *APOE3*, and *APOE4*, respectively.



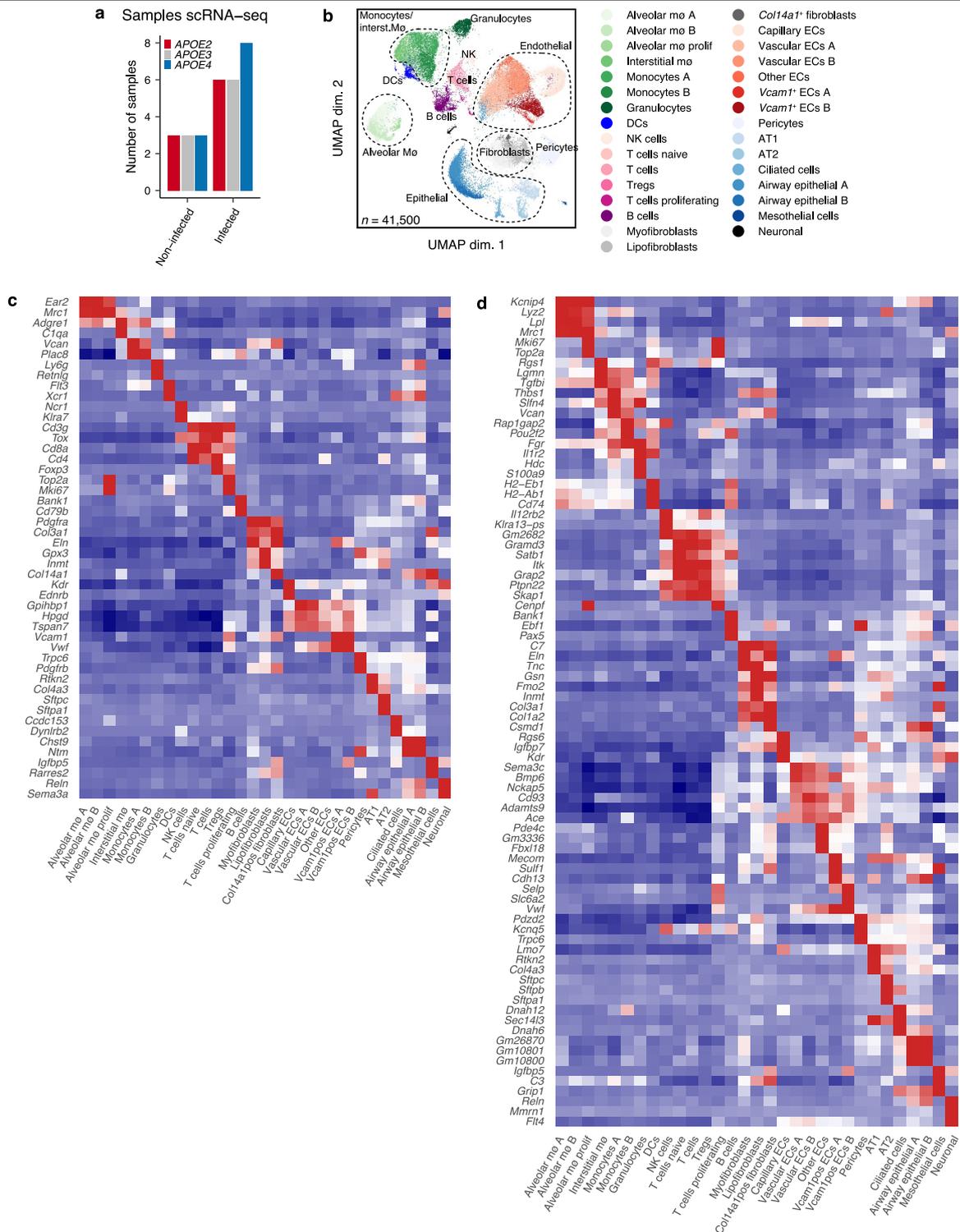
Extended Data Fig. 3 | Extended analysis of transcriptional profiles and immune cell infiltration in APOE knock-in mice during COVID-19. **a**, Module eigengenes averaged per condition. **b**, Independent validation of correlation of specific gene modules with APOE genotype in lungs of male APOE knock-in mice on day 4 post infection with SARS-CoV-2 MA10. Red indicates positive correlation of module eigengenes with APOE genotype ordered by its impact on COVID-19 survival ($E3 > E2 > E4$); stars indicate significant correlations (one-sided Pearson correlation tests). **c**, Averaged module eigengenes in the validation experiment. **d**, Correlation of gene modules with APOE genotype in lungs of 7-weeks old female APOE knock-in mice on day 4 post infection. Red indicates positive correlation of module eigengenes with APOE genotype ordered by its impact on COVID-19 survival ($E3 > E2 > E4$); stars indicate significant correlations (one-sided Pearson correlation tests). **e**, Averaged module eigengenes of modules significantly associated with APOE genotype in mice from (d). **f-g**, Network plot of the top ten hubgenes (genes with highest

intramodular connectivity) (f) and the top five GO pathways enriched in the 348 genes of the black module (g) (P values according to hypergeometric tests adjusted for FDR). **h**, Expression of genes constituting the midnightblue and greenyellow modules. Hubgenes are annotated by name. **i-o**, Network plots of the top ten hubgenes (i,k,m,o) and the top five GO pathways enriched in 45, 67, and 58 genes of the greenyellow, midnightblue, and yellow modules, respectively (j,l,n) (P values in j, l, and n according to hypergeometric tests adjusted for FDR). No pathways were enriched in the 24 genes making up the pink module. **p**, Immunofluorescence staining for CD45⁺ cells in lungs of APOE knock-in mice on day 4 post infection (n = 10, 15, 10 for APOE2, APOE3, and APOE4, respectively; P values according to two-sided Mann Whitney tests; . interquartile ranges of the hinges, and box centre and hinges indicate median and first and third quartiles, respectively). Images on the right show representative sections; scale bar, 100 μ m.



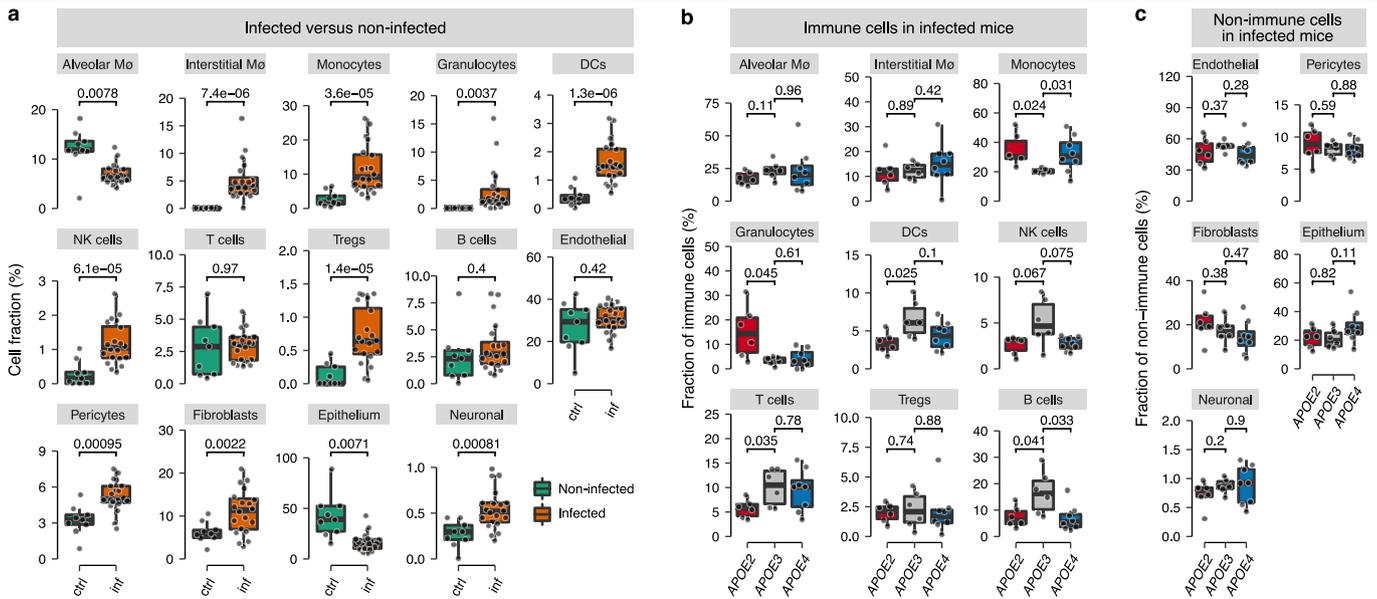
Extended Data Fig. 4 | Immune cell profiling of lungs and peripheral blood of *APOE* knock-in mice with COVID-19. a–b. Gating strategy to delineate leukocyte subsets (a) and assessment of the proportion of leukocyte subsets (b) in dissociated lungs of *APOE* knock-in mice on day 4 post infection with SARS-CoV-2 MA10 (n = 21, 15, 20 for *APOE2*, *APOE3*, and *APOE4*, respectively; data pooled from two independent experiments; P values according to one-tailed t-tests). **c.** Gating strategy to delineate leukocyte subsets in peripheral blood of *APOE* knock-in mice with COVID-19. **d–g.** Concentration of CD45⁺

leukocytes (d) and proportion of myeloid (e) and lymphoid (f) subsets in the peripheral blood of *APOE* knock-in mice on day 4 post infection with SARS-CoV-2 MA10 as assessed by flow cytometry (n = 10, 9, 7 for *APOE2*, *APOE3*, and *APOE4*, respectively; P values according to two-sided t tests). **g.** Representative flow cytometry plots for (e–f). Boxplot whiskers in b and d–f extend to the smallest and largest value within 1.5 × interquartile ranges of the hinges, and box centre and hinges indicate median and first and third quartiles, respectively.



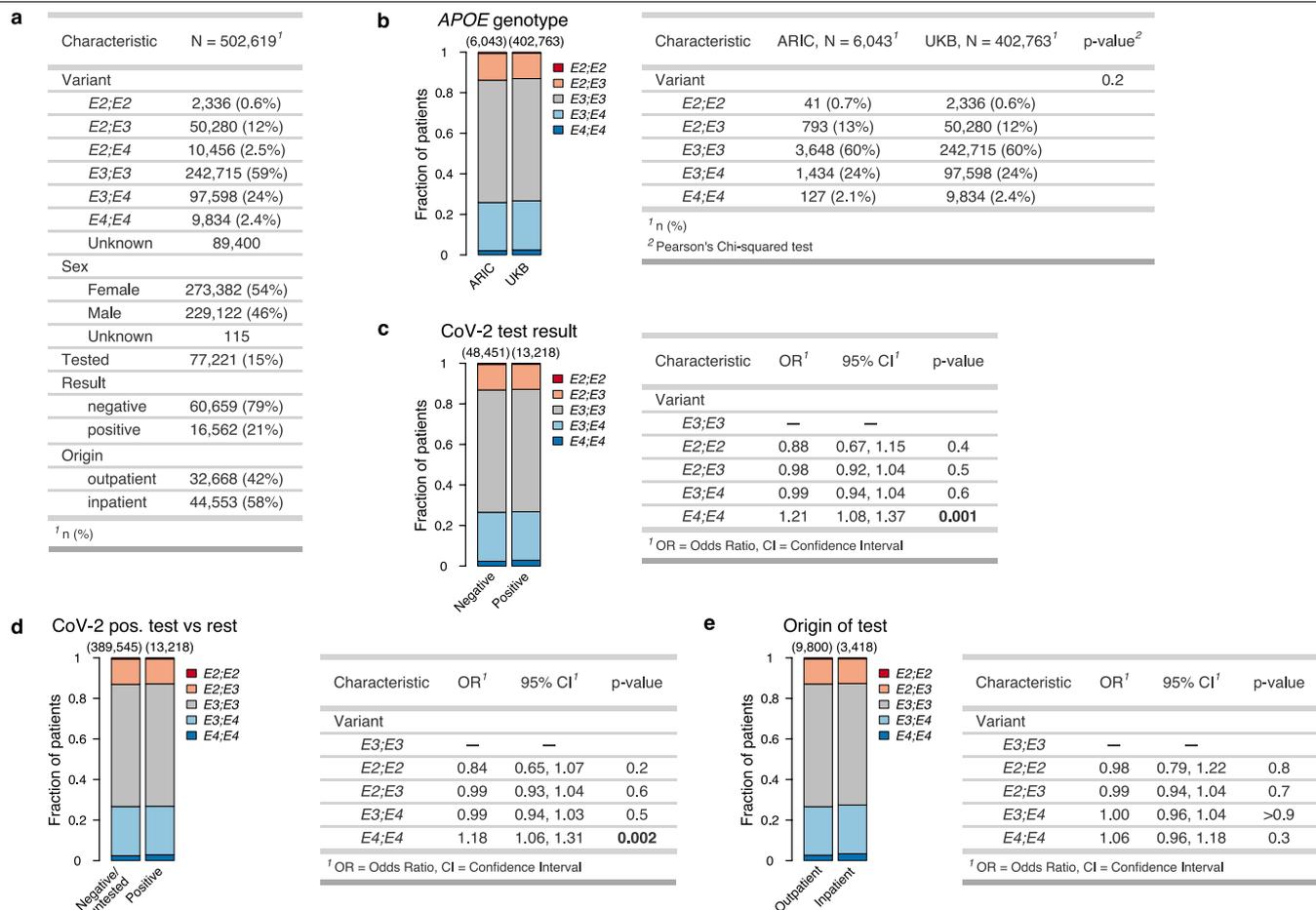
Extended Data Fig. 5 | Extended single cell RNA-sequencing data. a, Number of samples per genotype and condition for single cell RNA-sequencing (scRNA-seq). **b**, Uniform manifold approximation and projection (UMAP) plot

of 41,500 RNA-sequenced cells from *APOE* knock-in mice with or without COVID-19. **c-d**, Heatmaps of manually curated marker genes (**c**) and of top three differentially expressed genes per cluster (**d**) for cells from (**b**).



Extended Data Fig. 6 | Cellular composition in lungs from *APOE* knock-in mice with or without COVID-19. **a**, Fraction of grouped clusters of RNA-sequenced lung cells in *APOE* knock-in mice with or without COVID-19 (n = 9 and 20 for non-infected and infected, respectively; P values according to two-tailed t tests). **b–c**, Fraction of clusters in immune (**b**) and non-immune (**c**) cells from

infected mice from (a) (n = 6, 6, and 8 for *APOE2*, *APOE3*, and *APOE4*, respectively; P values according to two-tailed t tests). Boxplot whiskers in a–c extend to the smallest and largest value within 1.5 × interquartile ranges of the hinges, and box centre and hinges indicate median and first and third quartiles, respectively.



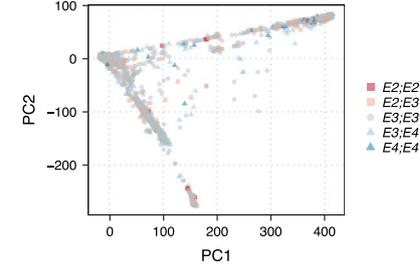
Extended Data Fig. 7 | *APOE* genotyping and SARS-CoV-2 test results in participants of the UK Biobank. **a**, General characteristics of the UK Biobank population. **b**, Distribution of *APOE* genotype in participants of the UK Biobank versus the ARIC study (Blair et al., Neurology, 2015) ($P = 0.2$, Chi-squared test). **c–e**, Distribution of *APOE* genotype in UK Biobank patients with positive versus negative SARS-CoV-2 test (**c**), with positive test versus negative test or untested

(**d**), and with SARS-CoV-2 test performed in- versus outpatient (**e**). Tables in (**c–e**) show odds ratios for testing positive versus negative, having a positive versus negative or no test, and having at least one inpatient versus only outpatient tests, respectively; P values are based on binomial general linearized models. Numbers on top of bars indicate sample sizes.

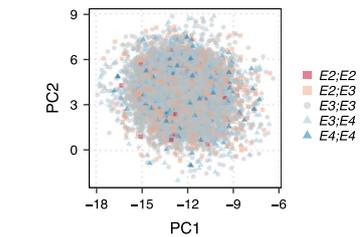
a Multivariate analysis UK Biobank

Variable	n	HR	p
Age	13207	1.163 (1.148, 1.177)	2.3e-123
Sex			
Female	6979	Reference	
Male	6228	1.926 (1.635, 2.268)	4.2e-15
PC1	13207	1.002 (1.001, 1.003)	0.00028
PC2	13207	0.999 (0.996, 1.002)	0.44
PC3	13207	1.002 (0.997, 1.006)	0.43
PC4	13207	0.997 (0.991, 1.004)	0.44
PC5	13207	0.989 (0.978, 1.001)	0.07
PC6	13207	0.987 (0.958, 1.016)	0.38
PC7	13207	0.990 (0.976, 1.005)	0.18
PC8	13207	1.006 (0.985, 1.027)	0.58
PC9	13207	0.998 (0.980, 1.016)	0.82
PC10	13207	0.988 (0.973, 1.003)	0.10
Genotype			
E3:E3	7973	Reference	
E2:E2	65	1.629 (0.673, 3.941)	0.28
E2:E3	1629	1.101 (0.865, 1.401)	0.44
E3:E4	3161	1.036 (0.856, 1.252)	0.72
E4:E4	379	2.220 (1.602, 3.076)	1.7e-06

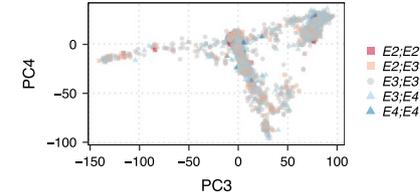
b Principal components 1-2



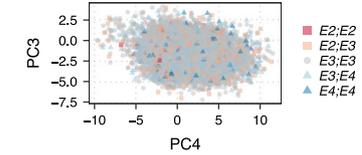
d Principal components 1-2



c Principal components 3-4



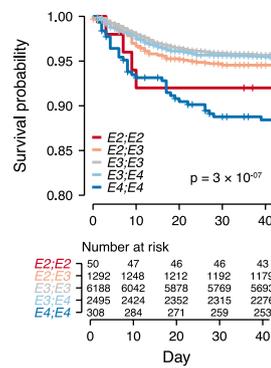
e Principal components 3-4



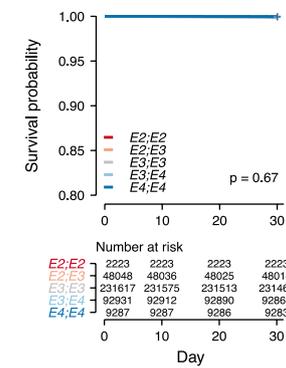
f Multivariate analysis UK Biobank

Variable	n	HR	p
Age	10333	1.164 (1.147, 1.181)	6.05e-93
Sex			
Female	5475	Reference	
Male	4858	1.967 (1.637, 2.365)	5.7e-13
Genotype			
E3:E3	6188	Reference	
E2:E2	50	1.561 (0.581, 4.190)	0.4
E2:E3	1292	1.244 (0.961, 1.610)	0.1
E3:E4	2495	1.000 (0.805, 1.242)	1.0
E4:E4	308	2.313 (1.626, 3.289)	3.1e-06

g APOE genotype



h Survival pre-pandemic



Extended Data Fig. 8 | The impact of APOE genotype on COVID-19 outcome is not confounded by population structure or its impact on longevity.

a, Multivariate analysis of the impact of age, sex, the first ten genetic principal components, and APOE genotype on survival of patients with SARS-CoV-2 infection in the UK Biobank (P values according to multivariable Cox proportional hazards model, n = 13,207). **b-c**, Dot plot of the genetic principal components 1-2 (**b**) and 3-4 (**c**) colored by APOE genotype of SARS-CoV-2-positive patients of the UK Biobank. PC, principal component. **d-e**, Dot plot of the genetic principal components 1-2 (**d**) and 3-4 (**e**) colored by APOE genotype of SARS-CoV-2-positive

patients with European ancestry in the UK Biobank. **f**, Multivariate analysis of the impact of age, sex, and APOE genotype on survival of patients with European ancestry and SARS-CoV-2 infection in the UK Biobank (P values according to multivariable Cox proportional hazards model, n = 10,333). **g**, Survival of patients from (**f**) stratified by APOE genotype (P value according to log-rank test). **h**, Survival of UK Biobank participants over a 30 day observation period in January 2019. The start of the observation period was Jan 1, 2019, and data were censored on Jan 31, 2019 (P value according to log-rank test; n = 384,106). Error bars in **a** and **f** indicate 95% confidence intervals.

Reporting Summary

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

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study. [For final submission](#): please carefully check your responses for accuracy; you will not be able to make changes later.

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 Nikon NIS elements v5.20.02, QuantStudio Design & Analysis (v1.4.3), FACSDiva v8 (BD Biosciences), cellSens Dimension software (v1.16).

Data analysis Weight and survival of mice were analyzed in RStudio v1.4 using R v4.1.0 and the R packages 'survival' (v3.2), 'survminer' (v0.4.9), 'forestmodel' (v0.6.2), 'ggpubr' (v0.4.0), and 'tidyverse' (v1.3.1). qPCR data were analyzed in QuantStudio Design & Analysis (v1.4.3); immunofluorescence images were acquired using Nikon NIS elements (v5.20.02) and analyzed using CellProfiler (v4.2.1). RNA-seq data were analyzed using BBDuk (v38.9), STAR (v2.7.8a), R (v4.1.0), and DESeq2 (v1.32.0). For weighted gene co-expression analysis the R package WGCNA (v1.70) was used. scRNAseq data were analyzed using ParseBiosciences pipeline v0.9.6p and Seurat v4.0.2. GSEA was performed using the clusterProfiler package (v4.0.0). Analysis of human data was performed using PLINK v1.9 and downstream analysis was performed using the same tools as for mouse data described above. Summary tables were compiled using the R package 'gtsummary' (v1.4.1). Flow cytometry data were analyzed using Flowjo v9.3.

Computer code: All custom code is publicly available at https://github.com/benostendorf/ostendorf_et al_2022

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

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

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

Bulk RNA-seq and scRNA-seq data have been deposited at the Gene Expression Omnibus (GEO) under accession numbers GSE184289 and GSE199498, respectively. All data from the UK Biobank is publicly available under www.ukbiobank.ac.uk. MSigDB is publicly available under <http://www.gsea-msigdb.org>. Source data is provided with this paper.

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

Life sciences study design

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

Sample size	No statistical methods were used to predetermine sample sizes. The number of samples per group was empirically chosen based on estimates of intra-group variation (Leist et al., Cell, 2020) and expected effect size (Ostendorf et al., Nature Medicine, 2020).
Data exclusions	All mouse data were included in the manuscript. For bulk RNA-sequencing, two out of 88 samples were identified as outliers based on principal component and sample distance analysis (large euclidean distance from all other samples of the same condition) and SARS-CoV-2 MA10 transcript abundance (outlying low virus transcript detection) and excluded from analysis.
Replication	Experiments were generally performed at least twice and pooled where appropriate as outlined in the individual figure legends.
Randomization	Samples were allocated randomly if possible (no infection versus infection groups). For experiments with genetically modified mice, allocation was performed according to genotype and mice were sex- and age-matched. No other covariates systematically differed between groups.
Blinding	Investigators were blinded for data collection and analysis of histology and immunofluorescence stainings. Two experiments on mouse survival and mouse weight courses were performed with investigators blinded with regards to genotype, yielding comparable results to the whole animal cohort. No blinding was performed for the remaining in-vivo experiments due to cage labeling requirements.

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
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used

IF stainings: CD45 (polyclonal, Abcam, #ab10558, 1:750), SARS nucleocapsid (polyclonal, Novus Biological, 56576, 1:1000), SARS-CoV2 nucleocapsid (polyclonal, GTX135357, 1:1000), goat anti-rabbit Alexa Fluor 555 (polyclonal, ThermoFisher Scientific, A-21428, 1:200). Flow cytometry: CD45-BV785 (clone: 30-F11, cat#: 103149, supplier: BioLegend, dilution: 1:3,000), CD11b-FITC (M1/70, 101206, BioLegend, 1:1,000), Ly6G-PerCP/Cy5.5 (1A8, 127616, BioLegend, 1:1,000), Ly6C-BV711 (HK1.4, 128037, BioLegend, 1:10,000), I-A/I-E-PE (M5/114.15.2, 107607, BioLegend, 1:10,000), CD19-PB (6D5, 115526, BioLegend, 1:500), CD19-BV421 (6D5, 115549, BioLegend,

1:500), NK1.1-APC (PK136, 17-5941-82, eBiosciences, 1:500), CD4-BV605 (GK1.5, 100451, BioLegend, 1:300), CD8 α -AF700 (53-6.7, 100730, BioLegend, 1:1,000), BV421-labeled SARS-CoV-2 S 539-546 tetramer (NIH Tetramer Core Facility, 1:200).

Validation

Validation data of the antibodies listed above was performed by the manufacturers and is available at each manufacturer's website by searching under the provided catalog numbers. Antibodies for IF were validated by the manufacturers either by assessing cells known to express or not to express the target protein and cross-referencing the expression pattern with the available literature or by orthogonal validation using an antibody-independent strategy. Antibodies for flow cytometry were validated by specificity testing on 1-3 target cell types with either single- or multi-color analysis including positive and negative cell types. New lots were validated to perform with similar intensity (MFI) as assessed on both positive and negative populations and each lot product was validated by QC testing with a series of titration dilutions.

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	VeroE6: ATCC (CRL-1586) and Ralph Baric (University of North Carolina at Chapel Hill); Huh-7.5: Charles Rice (Rockefeller University; Blight et al, J Virol, 2002, developed in the laboratory of Charles Rice); Caco-2: ATCC (HTB-37).
Authentication	No authentication was performed on the cell lines.
Mycoplasma contamination	Mycoplasma contamination was ruled out in all cell lines.
Commonly misidentified lines (See ICLAC register)	None.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	Human APOE2, APOE3, and APOE4 mice were obtained from Taconic Biosciences. Both female and male mice were used between 7-46 weeks of age as indicated in the figures. Mice were housed in a controlled ambient temperature (20-25°C) and humidity (30-70%) environment with a 12 hour light-dark cycle.
Wild animals	No wild animals were used in this study.
Field-collected samples	No field-collected samples were used in this study.
Ethics oversight	All animal experiments were conducted in accordance with a protocol approved by the Institutional Animal Care and Use Committee at The Rockefeller University.

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

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	PBMCs were obtained from mice by submandibular bleedings and red blood cells were lysed in ACK buffer as outlined in the methods section. Lungs were dissociated into single-cell suspensions using the lung dissociation kit and a gentleMACS dissociator (Milteny biotec) as outlined in the methods section.
Instrument	BD LSR Fortessa
Software	BD DIVA software v8 was used for data collection and Flowjo software v9.3 was used for data analysis.
Cell population abundance	No cell sorting was performed.
Gating strategy	Gating strategies were followed as outlined in Extended Data Fig 4. In brief, an initial gate based on basal scatter characteristics served to exclude debris followed by singlet gates based on FSC-H and SSC-H. Compensation was calculated using single color controls using Ultracomp compensation beads (ThermoFisher) for antibodies and amine-reactive beads for Zombie (ThermoFisher).

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