Age-Associated B cells predict impaired humoral immunity after COVID-19 vaccination in patients receiving immune checkpoint blockade

Juan Carlos Yam-Puc[1]†*, Zhaleh Hosseini[1]†, Emily C. Horner[1]†, Pehuén Pereyra Gerber[2,3]†, Nonantzin Beristain-Covarrubias[1]†, Robert Hughes[1]†, Maria Rust[1], Rebecca H. Boston[1], Magda Ali[1], Edward Simmons-Rosello[1], Martin O'Reilly[1], Harry Robson[1], Lucy H. Booth[1], Lakmini Kahanawita[1], Sofia Grigoriadou[4], Andrea Correa-Noguera[5], Lourdes Ceron-Gutierrez[5], Thomas Mulroney[1], Aleksei Lulla[6], Katrin Fischer[6], Andrew Craxton[1], Georgina S.F. Anderson[1], Xiao-Ming Sun[1], Anne Elmer[7], Caroline Saunders[7], Areti Bermperi[7], Sherly Jose[7], Mike Chapman[1], Marion MacFarlane[1], Anne E. Willis[1], Kiran R. Patil[1], Florian Hollfelder[6], Marko Hyvönen[6], CITIID-NIHR COVID-19 BioResource Collaboration, Sarah Spencer[1], Emily Staples[1], Matthew S. Buckland[4,5], Rainer Döffinger[8], Christine Parkinson[8§], Sara Lear[8§], Nicholas J. Matheson[2,3,,8,9§], James E. D. Thaventhiran[1,2,8§*].

[1] Medical Research Council Toxicology Unit, School of Biological Sciences, University of Cambridge, Cambridge, UK.
[2] Cambridge Institute of Therapeutic Immunology and Infectious Disease (CITIID), University of Cambridge, UK
[3] Department of Medicine, University of Cambridge, Cambridge, UK
[4] Department of Clinical Immunology, Barts Health, London, UK.
[5] UCL GOSH Institute of Child Health Division of Infection and Immunity, Section of Cellular and Molecular Immunology, London, UK.
[6] Department of Biochemistry, University of Cambridge, Cambridge, UK
[7] NIHR Cambridge Clinical Research Facility
[8] Cambridge University NHS Hospitals Foundation Trust, Cambridge, UK
[9] NHS Blood and Transplant, Cambridge, UK

† These authors contributed equally
§ These authors contributed equally
* Corresponding authors: jedt2@cam.ac.uk, jcy28@cam.ac.uk.

Key words: Inborn error of immunity (IEI), immune-checkpoint blockade, neutralising antibodies, antigen-specific memory B cells, BNT162b2 vaccine, immunodeficiency, cancer

The authors declare no potential conflicts of interest.

Abstract

The effect of immune checkpoint blockade on COVID-19 immunity is unclear. In this study, we determine whether immune checkpoint blockade expanded age-associated B cells (ABCs) are similar to those present in other conditions, and whether they enhance or detract from the COVID-19 vaccine responses. First, we use single cell RNA sequencing (scRNAseq) to show that ABCs arising from distinct aetiologies have common transcriptional profiles and may be further subdivided according to expression of genes associated with different immune functions, including the autoimmune regulator (AIRE). Next, we perform detailed longitudinal profiling of the COVID-19 vaccination response in patients. Finally, we show that high pre-vaccination ABC frequency correlates with decreased levels of antigen-specific memory B cells, and reduced magnitude and longevity of neutralising capacity against authentic SARS-CoV-2 virus. Expansion of ABCs is a biomarker for individuals with cancer requiring additional or more frequent booster immunisation against COVID-19.

Introduction

Despite the success of mRNA-lipid nanoparticle (mRNA-LNP) COVID-19 vaccines in reducing the risk of symptomatic infection, hospitalisation, and death(1,2), vaccinated patients with cancer remain at increased risk of severe outcomes following SARS-CoV-2 infection(3). A cancer therapy that observational studies(4-6) suggests could improve vaccine efficacy is immune checkpoint blockade (ICB). By targeting PD-1 and CTLA-4 checkpoints, ICB non-specifically promotes T cell responses, including those involved in anti-viral and anti-cancer immunity. Furthermore, preclinical evidence indicates immune checkpoint blockade can, via T and B cell interactions, enhance antibody responses(7,8). The neutralising antibodies produced by the humoral response are a vital component of vaccine protection as they inhibit viral replication *in vitro* and correlate with protection against infection *in vivo*(9-11). However, complicating the positive potential of ICB for vaccine enhancement, ICB induces the expansion of a B cell subset termed in other contexts Age-Associated B cells (ABCs)(12), which may have a confounding effect on humoral vaccine responses.

ABCs comprise a naturally occurring population of antigen-experienced B cells which expands continuously with age in healthy individuals but accumulates prematurely in people with certain immune dyscrasias, autoimmunity, and/or infectious diseases(13). They have also been termed $CD11c^+$ B cells, $CD11c^+$T-bet^+ B cells, double negative 2 (DN2) B cells or atypical memory B cells(13-17), and may additionally be generated by external challenges(18,19), including COVID-19 infection and vaccination(20). Although ABCs are associated with disease in autoimmunity, their role in vaccine immunity is uncertain. In mice, these cells are required for optimal antibody responses following influenza vaccination(21), possibly due to enhanced ability to present antigens to T cells relative to follicular B cells(16,22). In patients with cancer treated with ICB, the expansion of ABCs precedes the development of both antibody-mediated and non-antibody-mediated autoimmunity(12). This is in keeping with their role in promoting antibody responses. However, we and others have shown that expansion of ABCs is associated with antibody deficiency in specific cohorts of patients with inborn errors of immunity (IEI)(23-26). These include patients with *NFKB1* haploinsufficiency, in whom the genetic lesion leads predominantly to B cell-intrinsic immunodeficiency(26), and patients with *CTLA4* haploinsufficiency, in whom the genetic-lesion leads predominantly to a B cell-extrinsic functional T cell deficit (24,27). These observations suggest two key questions: are the ABCs found in patients treated with ICB equivalent to those found in other settings? And do these cells, when expanded through distinct mechanisms, have a positive or negative effect on humoral vaccine responses?

The increase in patients with cancer eligible to receive ICB treatment(28) and the survival benefit conferred by ICB treatment means that optimally protecting patients with cancer who are receiving these therapies from infection will be an increasing priority. Here, we answer the above questions regarding the characteristics and activity of ABCs in different patient groups as well as developing an understanding of the crosstalk between ICB and successful vaccination.

Results

ABCs arising from distinct aetiologies have common transcriptional profiles

We first assessed the similarities between ABCs from patients with different causes for ABC expansion, using single cell RNA sequencing (scRNAseq). This included B cells from healthy controls (HC, n=8), patients with cancer treated with ICB (ICB, n=8) and patients with *NFKB1* or *CTLA4* haploinsufficiency (IEI, n=4), together with a published set of single B cell transcriptomes from patients with SLE (n=3)(29). By focussing on patients with rare, monogenic defects leading to well-characterised IEI, we were able to ensure that contrasting B cell intrinsic (*NFKB1*) and extrinsic (*CTLA4*) aetiologies were included. Unsupervised clustering of the 52,402 B cells, displayed in a uniform manifold approximation and projection (UMAP) visualisation, revealed six clusters (**Fig. 1a-c).** Using CITE-seq, we identified ABCs that had low CD21 and high CD11c surface protein expression, in keeping with the definitions used in previous flow cytometry-based studies(12,26,30) (Supplementary **Fig. 1a-b**). Cells were dispersed across all clusters regardless of the patient group or sex from which they originated (**Fig. 1b-d)**. ABC frequencies were higher in all patient groups compared to healthy controls (**Fig. 1d).**

Distinct subsets of ABCs express genes associated with distinct immune function

The clustering and marker gene expression suggested that the ABC population can be subdivided into “Classical ABCs”, “Anergic ABCs” and “CD1c ABCs” (**Fig. 1e).** Classical ABC frequencies were higher in females and patients with SLE (**Fig. 1b-d** and Supplementary **Fig. 1c-e**). ABCs have been described as “anergic B cells” (31,32), interestingly, anergic ABCs (expressing canonical genes associated with anergy, such as EGR1 and NR4A1) clustered separately from Classical ABCs (**Fig. 1a and Fig. 1e** and **Supplementary Table 1**), indicating heterogeneity within ABCs for this phenotype. Indeed, many of the most differentially expressed genes in Classical ABCs were associated with MHC class II-restricted antigen presentation (**Fig. 1f**). Gene ontology enrichment analysis demonstrated selective upregulation of biological processes associated with professional

antigen presentation, such as antigen uptake, processing, class II presentation, and co-stimulation (**Fig. 1g** and **Supplementary Table 2**). This was specific for MHC class-II-restricted antigen presentation, as genes required for MHC class I-restricted antigen presentation were instead differentially expressed in CD1c ABCs, relative to Classical ABCs (Supplementary **Fig. 2**). This supports the functional separation of these two clusters, which were less distinct on UMAP visualisation (**Fig. 1a)**.

Upregulation of AIRE and its target genes is seen in Classical ABCs

We confirmed MHC-II antigen presentation related genes were amongst the most differentially expressed genes of Classical ABCs (**Fig. 2a** and **Supplementary Table 3**). Surprisingly, this analysis revealed the Autoimmune Regulator (AIRE) to be significantly upregulated in Classical ABCs from all subject cohorts (**Fig. 2a-b,** Supplementary **Fig. 3**). AIRE-expressing B lymphocytes were also in other clusters, albeit at a much lower frequency (**Fig. 2c**). This is consistent with the accumulation of AIRE-expressing B cells in the Classical ABC differentiation state during peripheral B cell development, rather than selective upregulation of AIRE after ABC differentiation. Expression of a single gene in 3.85% of cells will make a negligible contribution to the UMAP clustering, therefore enrichment of AIRE-expressing cells within the apex of the Classical ABC cluster (Supplementary **Fig. 3a)** further supports the biological association of AIRE expression and ABC differentiation. AIRE is a transcriptional regulator expressed in both the thymic epithelium and thymic B cells(33,34), which increases expression of tissue-restricted antigens (TRAs). TRA expression in the thymus is a central tolerance mechanism, whereby stimulation of developing T cells by antigens usually expressed in non-thymic tissues directs these cells away from fates with the potential for harmful self-antigen driven autoimmune responses. Previous reports have suggested that AIRE is not functional in the small proportion of peripheral lymphocytes in which it can be detected(35-37), although AIRE function has never been assessed in ABCs (a peripheral B cell subset specifically associated with autoimmune disease(16,38,39)). AIRE-expressing cells also upregulated a set of genes previously defined as AIRE targets in thymic B cells(34) (**Fig. 2d** and Supplementary **Fig. 3**). In addition, Classical ABCs upregulated HLA-G (Supplementary **Fig. 2**), a non-classical class I HLA molecule transactivated by AIRE in thymic epithelial cells(40). Taken together, these data provide functional evidence of AIRE expression in Classical ABCs. Further work will be necessary to evaluate if AIRE expression in ABCs leads to self-antigen expression and productive presentation that can initiate autoimmune responses.

ABC frequency predicts neutralising antibody response to COVID-19 vaccination

To investigate the impact of ABCs on mRNA-LNP vaccination, we next analysed the immune response to the mRNA BNT162b2 COVID-19 vaccine in patients with variable expansion of ABCs, including healthy controls (HC, n=10), patients with cancer treated with ICB (ICB, n=19) and patients with *NFKB1* or *CTLA4* haploinsufficiency (IEI, n=9) (**Fig. 3a** and **Supplementary Table 4**). All subjects were invited for blood sampling prior to their second vaccine dose (day 0), then at early (day 8), mid (day 21) and late (day 105) time points after this dose (**Fig. 3a**). Classical ABC frequencies prior to vaccination (day 0) were determined by multicolour flow cytometry staining for CD11c and CD21 (Supplementary **Fig. 4a**) and found to be elevated in patients compared with healthy controls (**Fig. 3b-c**).

Serum reactivity against SARS-CoV-2 spike (S), receptor-binding domain (RBD) and nucleocapsid (NCP) antigens was evaluated at day 21 after the second dose in the diagnostic immunology laboratory. The majority of patients (18/19 ICB, 5/8 IEI) generated IgG antibodies against S and RBD antigens (encoded by BNT162b2) at levels similar to the healthy control group (Supplementary **Fig. 4b**). These responses were unlikely to reflect previous natural infection with SARS-CoV-2, as patients did not have evidence of NCP reactivity (not encoded by BNT162b2) (Supplementary **Fig. 4b**). The quality of the COVID-19-specific antibody response can be determined by measuring the capacity of serum to neutralise authentic SARS-CoV-2 virus(41). Overall, most subjects displayed a rapid increase in neutralising antibody titres, peaking early (day 8) after their second vaccine dose (**Fig. 3d** and Supplementary **Fig. 4c**). By contrast, 5 patients (4 IEI and 1 ICB) failed to develop detectable neutralising capacity at any time point, despite having detectable anti-S antibodies. These patients are at greater risk of infection with SARS-CoV-2 and may be susceptible to prolonged or refractory COVID-19(42,43).

Critically, the frequency of pre-vaccine ABCs was inversely correlated with neutralising antibody titre at all timepoints (**Fig. 3e**), suggesting that ABCs predict both peak neutralising capacity after vaccination, and the longevity of the neutralising antibody response. Indeed, at day 105, 4/8 patients with IEI and 8/16 patients treated with ICB had neutralising antibody titres which had fallen below the highest titre observed at day 0, compared with 0/7 HCs (**Fig. 3d**). Individuals with neutralising capacity below this threshold showed higher frequencies of ABCs than the individuals above it (**Fig. 3f**). Premature expansion of ABCs is therefore seen both in patients with IEI and patients with cancer treated with ICB, and (regardless of aetiology) correlates with a reduced ability to generate and maintain neutralising antibody responses to COVID-19 vaccination.

Premature expansion of ABCs is associated with lower levels of antigen-specific memory B cells

Whilst circulating antibodies derived from plasma cells wane over time, long-lived immunological memory can persist in expanded clones of antigen-specific memory B cells. We therefore assessed the frequency and differentiation of circulating SARS-CoV-2 RBD-specific B cells elicited after the second dose of vaccine using multicolour flow cytometry (Supplementary **Fig. 5a**). Strikingly, RBD-binding B cells were always significantly less frequent in patients than in healthy controls (**Fig. 4a-b** and Supplementary **Fig. 5B**), and the fraction of RBD-binding B cells amongst all $CD19^+$ B cells correlated consistently with titres of neutralising antibodies (Supplementary **Fig. 5c**). In addition, RBD-specific B cell frequency immediately prior to the second vaccine dose correlated with neutralising capacity at days 8 (peak response), 21 and 105 (**Fig. 4c**). This is consistent with a causal relationship between the level of antigen-specific B cells prior to the second dose of vaccine, and the magnitude and longevity of the neutralising antibody response. Furthermore, similar to neutralising capacity, the levels of RBD-specific B cells at later timepoints could be predicted by the frequency of ABCs at day 0 (**Fig. 4d**). Individuals with the highest frequencies of ABCs showed reduced RBD-specific B cell differentiation (Supplementary **Fig. 5c-d**).

Finally, to analyse the cellular phenotype of rare RBD-binding B cells in a comprehensive, unbiased way, we performed unsupervised clustering and UMAP visualisation of flow cytometry data from all patients and HCs at all timepoints after vaccination (**Fig. 4e**). RBD-binding B cells from both patients and HCs were observed within clusters of memory B cells and plasmablasts (**Fig. 4f**), suggesting a quantitative reduction rather than qualitative difference in the humoral response. Across all participants, the proportion of RBD-specific cells expressing plasmablast markers peaked at day 8, whilst those expressing memory B cell makers peaked a day 21 (**Fig. 4g**). Differentiation of antibody-producing cells therefore occurs contemporaneously with peak neutralisation capacity (day 8). Since increased antibody production cannot be driven by the proliferation of terminally differentiated (non-dividing) plasma cells, a sizeable contribution to the neutralising antibody response after the second dose of vaccine is likely the result of this memory B cell proliferation and differentiation. Taken together, our data therefore suggest that premature expansion of ABCs negatively impacts all aspects of the humoral immune response to COVID-19 vaccination.

Discussion

While the specific causal mechanism for ABC expansion in any single older patient (to which polygenic risk alleles, ageing, therapy and obesity can all contribute) is unknown, in young patients with ultra-rare monogenic inherited disease, expansion is caused by the specific gene disrupted. Our analysis of 4 different patient groups: healthy controls, patients with cancer treated with ICB, patients with autoimmune SLE, and patients with two distinct IEI (CTLA haploinsufficiency and NFKB1 haploinsufficiency) demonstrates a homogeneity of the ABC differentiation states irrespective of the specific cause. Therefore, the pathological consequence of expanded ABCs is likely related to their increased frequency rather than an inherent difference in these cells from patients with distinct diseases. Additionally, the finding of AIRE and tissue restricted antigen expression within a small subset of Classical ABCs, cells that transcriptionally appear as professional antigen presenting cells could provide a mechanistic link between the expansion of these cells and the autoimmune toxicity associated with ICB therapy.

Most clinically relevant, however, is our finding that patients with expanded ABCs have a reduced B cell vaccine response, leading to reduced neutralisation capacity and reduced memory B cell formation. Since memory B frequency before vaccination correlates with the subsequent neutralisation response to immunisation, this predicts that there will be reduced neutralisation capacity to subsequent booster immunisation. Consequently, patients with expanded ABCs will require more frequent booster immunisation to achieve an equal level of longitudinal B cell immune protection. Overall, these results place ABC frequency as a predictive biomarker for reduced vaccine protection that could guide booster vaccination schedules for patients at risk of breakthrough infection.

Methods

Patient recruitment and ethics

Enrolment of patients into this study was based on the deficiency of particular genes of interest (IEI: CTLA4, LRBA, NFkB1, NFkB2) or clinical diagnostics (ICB, n = 19), whilst healthy controls for this study were enrolled based on their clinical healthy status (n = 10). Sample-size calculation was limited due to IEI patients in the study (n = 9) reflect a rare population of individuals within the general population that we were able to recruit (Supplementary Table 4).

The research was conducted in accordance with the principles of Good Clinical Practice and following granted protocols of the National Institute for Health and Care Research (NIHR) Bioresource of England or granted protocols of the Barts and the London Immunology Registry, United Kingdom. Samples were collected with the written informed consent of all

study participants under the NIHR BioResource - Research Tissue Bank (NBR-RTB) ethics (REC:17/EE/0025) or under the Barts and the London Immunology Registry (REC: 11/LO/1689), United Kingdom. Clinical data were collected by Clinical Immunology Consultants at Cambridge University Hospital and Bart's Health at London, United Kingdom via the Electronic Healthcare Record (Epic), or direct patient contact.

The patients and healthy controls were consented under the East of England Cambridge South National research ethics committee (REC) reference 13/EE/0325 or Barts and the London Immunology Registry (REC: 11/LO/1689), United Kingdom.

Blood donation and separation

Patients and controls gave voluntary blood samples prior to second dose and at early (day 8), mid (day 21) and late (day 105) timepoints following their second BNT162b2 vaccine. Where sample size of individual assays was not the same as the total sample size, this was because there was no sample collection at those specific time points. Samples may have also been excluded if there was serological evidence of prior exposure to SARS-COV-2. Peripheral blood samples were acquired in either lithium heparin or serum separating tubes. Peripheral blood mononuclear cells (PBMCs) were isolated by a density gradient centrifugation protocol and stored at -80°C for up to a week before being transferred to liquid nitrogen until use. Whilst serum tubes were centrifuged to separate serum from cell pellet before being aliquoted and stored at -80°C until use.

Single cell library preparation and sequencing

Frozen PBMCs were thawed and B cell population was enriched using magnetic separation (Miltenyi 130-101-638). Samples were next labelled with oligonucleotide-tagged antibodies against CD11c, CD21, CD27 and CD85j (TotalSeq, BioLegend). After labelled, cell counts were adjusted to 1,035 cells per microlitre before being loaded onto a Chromium Next GEM chip K (10X Genomics, 2000182) for the subsequent library's construction according to Single cell 5' v2 protocol (CG000330 RevA, 10X Genomics). The library was quality controlled by Agilent 4150 TapeStation and quantified by RT-PCR using the KAPA library quantification kit for Illumina platforms (Roche, KK4824). Samples were sequenced on an Illumina NovaSeq with a sequencing depth of at least 50, 000 reads per cell.

Software versions

Single cell data was analysed using Cell Ranger software version 6.1.1, citeseq version 1.5.0, R version 4.0.3 and R packages (Seurat 4.1.0, SeuratObject 4.0.4, Scuttle 1.0.4, rstatix 0.7.0, tidyverse 1.3.0 and dplyr 1.0.8). Flow cytometry data and antibody titres were

302 analysed using FlowJo 10.7.1 and GraphPad Prism 9.1.2 (225), respectively. Figures were
303 produced using ggplot2 3.3.5, gridExtra 2.3, ComplexHeatmap 2.6.2, Nebulosa 1.0.2,
304 RColorBrewer 1.1-2, ggrepel 0.9.1, ggpubr 0.4.0, scales 1.1.1, showtext 0.9-5 in R and
305 GraphPad Prism 9.1.2 (225) and FlowJo 10.7.1.
306
307 **Preprocessing of scRNA-seq data**
308 Raw FASTQ files of the gene expression library were analysed using 10x Genomics Cell
309 Ranger software v.6.1.1[1] and aligned to the GRCh38 genome provided by Cell Ranger. All
310 Ig V, D, J and constant genes as well as TCR genes were deleted from the dataset so that
311 downstream analysis is not affected by highly variable clonotype genes. For the
312 quantification of antibody derived tags (ADT) and oligo hashtags, citeseq pipeline v.1.5.0[2]
313 was used. The generated hashtag count matrices as well as raw gene expression count
314 matrices were then used as inputs for the Seurat package v.4.1.0[3] in R v.4.0.3. Seurat
315 objects were created from the corresponding transcript count matrices. ADT assays were
316 added using CreateAssayObject function of Seurat to include expression levels of surface
317 proteins. Samples were then demultiplexed using HTODemux function. Doublets and cells
318 with no assigned hashtag were removed from objects for further analysis. Scuttle package
319 v.1.0.4[4] was used for quality control. 3 x Median absolute deviation (MAD) was considered
320 as the threshold for quality control; number of features and number of read counts were
321 filtered from both sides, whereas percentage of mitochondrial genes was filtered from upper
322 side. The samples were log-normalized for further differential expression analysis. Before
323 integrating all objects, they were normalized based on regularized negative binomial
324 regression using the SCTransform function of the Seurat package, regressing out for cell
325 cycle scores, number of counts and percentage of mitochondrial genes. The objects were
326 then integrated using the integration protocol of the Seurat package to perform batch effect
327 correction. The objects used for integration include: four groups of healthy controls each
328 having two subjects (total of 26900 cells in the final dataset), one group of IEI patients (5345
329 cells in the final dataset) and two groups of ICB patients (total of 9751 cells in the final
330 dataset) each having four subjects, and SLE patients with 3 subjects (10406 cells in the final
331 dataset).
332
333 **SLE data**
334 Data of SLE patients from Bhamidipati et al [5] was downloaded from GSE163121,
335 reprocessed and integrated with other objects using the same procedure as described
336 above.
337

Clustering of data

Principal component analysis (PCA) was performed on the integrated object to reduce the dimensionality of the dataset. First, clustering was done using 30 principal components (PCs) and resolution parameter of 0.1. Clusters with a high expression of non-B cell markers were removed from the dataset. After deletion of all non-B cells, the dataset consists of 52402 cells. PCs were calculated in the new dataset and unsupervised clustering was performed using 20 PCs and resolution parameter of 1.4. Cluster markers were calculated with the default parameters of the FindMarkers function. Clusters with similar markers were manually merged. Final clusters were labelled based on the differentially expressed genes in each cluster.

Gene ontology enrichment analysis

Gene ontology [6, 7] enrichment analysis was performed via PANTHER [8] from http://www.geneontology.org using the list of significantly upregulated genes from classical ABC cluster (Average log2 fold change > 0.25 and percent expressed in each group > 0.25).

Geneset Scores

A list of 74 Aire-induced genes were obtained from Yamano et al., [9] after converting mouse gene IDs to human gene IDs. For calculating gene set score for each cell, AddModuleScore function from Seurat package was used.

Kernel density estimation of gene expression

Nebulosa package v 1.0.2 [10] was used for calculating and plotting gene-weighted density estimation of AIRE expression.

SARS-CoV-2 serology

A Luminex bead-based immunoassay was used to quantify specific antibodies to full-length trimeric spike (S), spike receptor binding domain (RBD) and nucleocapsid (NCP) of SARS-CoV-2 as previously described [11, 12]. Briefly, a multiplex assay was established by covalently coupling recombinant SARS-CoV-2 proteins to distinct carboxylated bead sets (Luminex, Netherlands). The S protein used here was the S-R/PP described in Xiong et al [11], and the RBD protein was described by Stadlbauer et al [13]. The NCP protein used is a truncated construct of the SARS-CoV-2 NCP protein comprising residues 48–365 (both ordered domains with the native linker) with an N terminal uncleavable hexahistidine tag. NCP was expressed in E. Coli using autoinducing media for 7h at 37°C and purified using immobilised metal affinity chromatography (IMAC), size exclusion and heparin chromatography. The S-, RBD- and NCP-coupled bead sets were incubated with patient

sera at 3 dilutions (1/100, 1/1000, 1/10000) for 1 h in 96-well filter plates (MultiScreen HTS; Millipore) at room temperature in the dark on a horizontal shaker. After washes, beads were incubated for 30 min with a PE-labeled anti–human IgG-Fc antibody (Leinco/Biotrend), washed as described above, and resuspended in 100 µl PBS/ Tween. Antibody-specific binding was interpreted using Exponent Software V31 software on the Luminex analyzer (Luminex / R&D Systems) and reported as mean fluorescence intensity (MFI). The diagnostic thresholds used adhered to UK national guidelines.

Neutralising antibodies to SARS-CoV-2

The SARS-CoV-2 used in this study was a wildtype (lineage B) virus (SARS-CoV-2/human/Liverpool/REMRQ0001/2020), a kind gift from Ian Goodfellow (University of Cambridge), isolated early in the COVID-19 pandemic by Lance Turtle (University of Liverpool) and David Matthews and Andrew Davidson (University of Bristol) [14-16] from a patient on the Diamond Princess cruise ship. Sera were heat-inactivated at 56°C for 30 mins before use, and neutralising antibody titres at 50% inhibition (NT50s) measured as previously described [17, 18].

In brief, luminescent HEK293T-ACE2-30F-PLP2 reporter cells (clone B7) expressing SARS-CoV-2 Papain-like protease-activatable circularly permuted firefly luciferase (FFluc) were seeded in flat-bottomed 96-well plates. The next day, SARS-CoV-2 viral stock (MOI=0.01) was pre-incubated with a 3-fold dilution series of each serum for 2 h at 37°C, then added to the cells. 16 h post-infection, cells were lysed in Bright-Glo Luciferase Buffer (Promega) diluted 1:1 with PBS and 1% NP-40, and FFluc activity measured by luminometry. Experiments were conducted in duplicate.

To obtain NT50s, titration curves were plotted as FFluc vs log (serum dilution), then analysed by non-linear regression using the Sigmoidal, 4PL, X is log(concentration) function in GraphPad Prism. NT50s were reported when (1) at least 50% inhibition was observed at the lowest serum dilution tested (1:20), and (2) a sigmoidal curve with a good fit was generated. For purposes of visualisation and ranking, samples for which visual inspection of the titration curve indicated inhibition at low dilutions, but which did not meet criteria (1) and (2) above, were assigned an arbitrary NT50 of 4.

World Health Organisation International Standard 20/136 (WHO IS 20/136) has an NT50 of 1967 against wildtype SARS-CoV-2 when measured in this assay. This standard comprises pooled convalescent plasma obtained from 11 individuals which, when reconstituted, is assigned an arbitrary neutralising capacity of 1000 IU/ml against early 2020 SARS-CoV-2 isolates [19]. NT50s from this study may therefore be converted to IU/ml using a calibration

409 factor of 1000/1967 (0.51), with a limit of quantitation of 10.2 IU/ml (corresponding to an
410 NT50 of 20).
411
412 **Flow cytometric analysis of B cells**
413 Frozen PBMCs were thawed and stained with B cell panel antibody cocktail (Supplementary
414 Table 5). All samples were acquired on the BD LSRFortessa using FACSDIVA software
415 (BD-Biosciences). FCS files were exported and analysed using FlowJo v10.7.2 (BD-
416 Biosciences) software. Additionally, a uniform manifold approximation and projection
417 (UMAP) algorithm for dimensional reduction was performed on a concatenated FCS file
418 comprising events from the CD19+ gate from HC, IEI and ICB samples analysed, utilizing
419 the UMAP FlowJo plugin (v3.1). The composite samples were gated as indicated on
420 Supplementary Fig. 2 for RBD+ B cell identification and then overlaid by study group in the
421 UMAP generated, for additional visualisation. XShift (v1.4.1) and ClusterExplorer (1.5.15)
422 plugins were used for unsupervised cluster generation and exploration, respectively. Class-
423 switched memory B cells and plasmablasts/plasma cells were selected based on the
424 expression of CD19+IgD-IgM-CD38-CD27+ and CD19+IgD-IgM-CD38+/highCD27+/high,
425 respectively.
426
427 **RBD tetramer production**
428 SARS-CoV-2 RBD was expressed with N-terminal fusion of His-Zbasic-TEV module (8×His
429 tag, Zbasic domain [20] and TEV protease cleavage site) and C-terminal Avi-tag in
430 BL21(DE3) cells (Novagen) cultured in 2xYT media at 37°C. RBD containing inclusion
431 bodies were isolated and solubilised. The denatured protein was first purified by immobilized
432 metal affinity chromatography using PureCube Ni-NTA resin, eluted and refolding was
433 allowed to proceed for 72 hours at 4°C. N-terminal fusion tag was cleaved off by NHis-
434 TEVpro (produced in-house). Biotinylation of the RBD protein was carried out using 5 mM
435 MgCl2, 2 mM ATP, 150 µM biotin and 50 µg/ml of biotin ligase BirA-CHis (produced in-
436 house). The excess of biotinylated RBD protein was then incubated with 0.5 mg/ml of
437 fluorescently labelled Streptavidin (BioLegend) for 1 hour at room temperature and, finally,
438 RBD-Streptavidin complexes were separated by size exclusion chromatography using a
439 Superdex 200 Increase 10/300 GL column (Cytiva). The RBD-Streptavidin complexes were
440 then analysed by SDS-PAGE.
441
442 **Statistical analysis**
443 Statistical analysis for FACS and serology data was performed using GraphPad Prism
444 v9.1.1 (GraphPad Prism Software Inc). Two-tailed non-parametric Mann Whitney tests were

used to determine differences between groups. For multiple comparisons ANOVA with Tukey's multiple comparisons test was used. Associations were calculated using two-tailed Spearman's rank correlations and results are shown with linear trend levels. For the single cell data, statistical significance for paired comparisons were performed using Wilcoxon rank sum test in R.

Data and code availability

Sequencing data have been deposited in the Gene Expression Omnibus database [GEO: in preparation]. Additional information and materials and all codes for scRNA-seq analysis will be made available upon request.

References

1. Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R, *et al.* Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. N Engl J Med **2021**;384(5):403-16 doi 10.1056/NEJMoa2035389.
2. Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, *et al.* Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. N Engl J Med **2020**;383(27):2603-15 doi 10.1056/NEJMoa2034577.
3. Schmidt AL, Labaki C, Hsu CY, Bakouny Z, Balanchivadze N, Berg SA, *et al.* COVID-19 vaccination and breakthrough infections in patients with cancer. Ann Oncol **2022**;33(3):340-6 doi 10.1016/j.annonc.2021.12.006.
4. Monin L, Laing AG, Munoz-Ruiz M, McKenzie DR, Del Molino Del Barrio I, Alaguthurai T, *et al.* Safety and immunogenicity of one versus two doses of the COVID-19 vaccine BNT162b2 for patients with cancer: interim analysis of a prospective observational study. Lancet Oncol **2021**;22(6):765-78 doi 10.1016/S1470-2045(21)00213-8.
5. Goshen-Lago T, Waldhorn I, Holland R, Szwarcwort-Cohen M, Reiner-Benaim A, Shachor-Meyouhas Y, *et al.* Serologic Status and Toxic Effects of the SARS-CoV-2 BNT162b2 Vaccine in Patients Undergoing Treatment for Cancer. JAMA Oncol **2021**;7(10):1507-13 doi 10.1001/jamaoncol.2021.2675.
6. Thakkar A, Gonzalez-Lugo JD, Goradia N, Gali R, Shapiro LC, Pradhan K, *et al.* Seroconversion rates following COVID-19 vaccination among patients with cancer. Cancer Cell **2021**;39(8):1081-90 e2 doi 10.1016/j.ccell.2021.06.002.
7. Shi J, Hou S, Fang Q, Liu X, Liu X, Qi H. PD-1 Controls Follicular T Helper Cell Positioning and Function. Immunity **2018**;49(2):264-74 e4 doi 10.1016/j.immuni.2018.06.012.
8. Wang CJ, Heuts F, Ovcinnikovs V, Wardzinski L, Bowers C, Schmidt EM, *et al.* CTLA-4 controls follicular helper T-cell differentiation by regulating the strength of CD28 engagement. Proc Natl Acad Sci U S A **2015**;112(2):524-9 doi 10.1073/pnas.1414576112.
9. Khoury DS, Cromer D, Reynaldi A, Schlub TE, Wheatley AK, Juno JA, *et al.* Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. Nat Med **2021**;27(7):1205-11 doi 10.1038/s41591-021-01377-8.
10. Edara VV, Lai L, Sahoo MK, Floyd K, Sibai M, Solis D, *et al.* Infection and vaccine-induced neutralizing antibody responses to the SARS-CoV-2 B.1.617.1 variant. bioRxiv **2021** doi 10.1101/2021.05.09.443299.
11. Gilbert PB, Montefiori DC, McDermott AB, Fong Y, Benkeser D, Deng W, *et al.* Immune correlates analysis of the mRNA-1273 COVID-19 vaccine efficacy clinical trial. Science **2022**;375(6576):43-50 doi 10.1126/science.abm3425.
12. Das R, Bar N, Ferreira M, Newman AM, Zhang L, Bailur JK, *et al.* Early B cell changes predict autoimmunity following combination immune checkpoint blockade. J Clin Invest **2018**;128(2):715-20 doi 10.1172/JCI96798.
13. Cancro MP. Age-Associated B Cells. Annu Rev Immunol **2020**;38:315-40 doi 10.1146/annurev-immunol-092419-031130.
14. Hao Y, O'Neill P, Naradikian MS, Scholz JL, Cancro MP. A B-cell subset uniquely responsive to innate stimuli accumulates in aged mice. Blood **2011**;118(5):1294-304 doi 10.1182/blood-2011-01-330530.

15. Kim CC, Baccarella AM, Bayat A, Pepper M, Fontana MF. FCRL5(+) Memory B Cells Exhibit Robust Recall Responses. Cell Rep **2019**;27(5):1446-60 e4 doi 10.1016/j.celrep.2019.04.019.
16. Rubtsov AV, Rubtsova K, Fischer A, Meehan RT, Gillis JZ, Kappler JW, *et al.* Toll-like receptor 7 (TLR7)-driven accumulation of a novel CD11c(+) B-cell population is important for the development of autoimmunity. Blood **2011**;118(5):1305-15 doi 10.1182/blood-2011-01-331462.
17. Stewart A, Ng JC, Wallis G, Tsioligka V, Fraternali F, Dunn-Walters DK. Single-Cell Transcriptomic Analyses Define Distinct Peripheral B Cell Subsets and Discrete Development Pathways. Front Immunol **2021**;12:602539 doi 10.3389/fimmu.2021.602539.
18. Myles A, Sanz I, Cancro MP. T-bet(+) B cells: A common denominator in protective and autoreactive antibody responses? Curr Opin Immunol **2019**;57:40-5 doi 10.1016/j.coi.2019.01.002.
19. Rubtsov AV, Marrack P, Rubtsova K. T-bet expressing B cells - Novel target for autoimmune therapies? Cell Immunol **2017**;321:35-9 doi 10.1016/j.cellimm.2017.04.011.
20. Pape KA, Dileepan T, Kabage AJ, Kozysa D, Batres R, Evert C, *et al.* High-affinity memory B cells induced by SARS-CoV-2 infection produce more plasmablasts and atypical memory B cells than those primed by mRNA vaccines. Cell Rep **2021**;37(2):109823 doi 10.1016/j.celrep.2021.109823.
21. Johnson JL, Rosenthal RL, Knox JJ, Myles A, Naradikian MS, Madej J, *et al.* The Transcription Factor T-bet Resolves Memory B Cell Subsets with Distinct Tissue Distributions and Antibody Specificities in Mice and Humans. Immunity **2020**;52(5):842-55 e6 doi 10.1016/j.immuni.2020.03.020.
22. Rubtsov AV, Rubtsova K, Kappler JW, Jacobelli J, Friedman RS, Marrack P. CD11c-Expressing B Cells Are Located at the T Cell/B Cell Border in Spleen and Are Potent APCs. J Immunol **2015**;195(1):71-9 doi 10.4049/jimmunol.1500055.
23. Gamez-Diaz L, August D, Stepensky P, Revel-Vilk S, Seidel MG, Noriko M, *et al.* The extended phenotype of LPS-responsive beige-like anchor protein (LRBA) deficiency. J Allergy Clin Immunol **2016**;137(1):223-30 doi 10.1016/j.jaci.2015.09.025.
24. Kuehn HS, Ouyang W, Lo B, Deenick EK, Niemela JE, Avery DT, *et al.* Immune dysregulation in human subjects with heterozygous germline mutations in CTLA4. Science **2014**;345(6204):1623-7 doi 10.1126/science.1255904.
25. Schwab C, Gabrysch A, Olbrich P, Patino V, Warnatz K, Wolff D, *et al.* Phenotype, penetrance, and treatment of 133 cytotoxic T-lymphocyte antigen 4-insufficient subjects. J Allergy Clin Immunol **2018**;142(6):1932-46 doi 10.1016/j.jaci.2018.02.055.
26. Tuijnenburg P, Lango Allen H, Burns SO, Greene D, Jansen MH, Staples E, *et al.* Loss-of-function nuclear factor kappaB subunit 1 (NFKB1) variants are the most common monogenic cause of common variable immunodeficiency in Europeans. J Allergy Clin Immunol **2018**;142(4):1285-96 doi 10.1016/j.jaci.2018.01.039.
27. Schubert D, Bode C, Kenefeck R, Hou TZ, Wing JB, Kennedy A, *et al.* Autosomal dominant immune dysregulation syndrome in humans with CTLA4 mutations. Nat Med **2014**;20(12):1410-6 doi 10.1038/nm.3746.
28. Haslam A, Gill J, Prasad V. Estimation of the Percentage of US Patients With Cancer Who Are Eligible for Immune Checkpoint Inhibitor Drugs. JAMA Netw Open **2020**;3(3):e200423 doi 10.1001/jamanetworkopen.2020.0423.

29. Bhamidipati K, Silberstein JL, Chaichian Y, Baker MC, Lanz TV, Zia A, *et al.* CD52 Is Elevated on B cells of SLE Patients and Regulates B Cell Function. Front Immunol **2020**;11:626820 doi 10.3389/fimmu.2020.626820.
30. Schwab C, Gabrysch A, Olbrich P, Patino V, Warnatz K, Wolff D, *et al.* Phenotype, penetrance, and treatment of 133 cytotoxic T-lymphocyte antigen 4-insufficient subjects. J Allergy Clin Immunol **2018** doi 10.1016/j.jaci.2018.02.055.
31. Isnardi I, Ng YS, Menard L, Meyers G, Saadoun D, Srdanovic I, *et al.* Complement receptor 2/CD21- human naive B cells contain mostly autoreactive unresponsive clones. Blood **2010**;115(24):5026-36 doi 10.1182/blood-2009-09-243071.
32. Terrier B, Joly F, Vazquez T, Benech P, Rosenzwajg M, Carpentier W, *et al.* Expansion of functionally anergic CD21-/low marginal zone-like B cell clones in hepatitis C virus infection-related autoimmunity. J Immunol **2011**;187(12):6550-63 doi 10.4049/jimmunol.1102022.
33. Gies V, Guffroy A, Danion F, Billaud P, Keime C, Fauny JD, *et al.* B cells differentiate in human thymus and express AIRE. J Allergy Clin Immunol **2017**;139(3):1049-52 e12 doi 10.1016/j.jaci.2016.09.044.
34. Yamano T, Nedjic J, Hinterberger M, Steinert M, Koser S, Pinto S, *et al.* Thymic B Cells Are Licensed to Present Self Antigens for Central T Cell Tolerance Induction. Immunity **2015**;42(6):1048-61 doi 10.1016/j.immuni.2015.05.013.
35. Hubert FX, Kinkel SA, Webster KE, Cannon P, Crewther PE, Proeitto AI, *et al.* A specific anti-Aire antibody reveals aire expression is restricted to medullary thymic epithelial cells and not expressed in periphery. J Immunol **2008**;180(6):3824-32 doi 10.4049/jimmunol.180.6.3824.
36. Mathis D, Benoist C. Aire. Annu Rev Immunol **2009**;27:287-312 doi 10.1146/annurev.immunol.25.022106.141532.
37. Fergusson JR, Morgan MD, Bruchard M, Huitema L, Heesters BA, van Unen V, *et al.* Maturing Human CD127+ CCR7+ PDL1+ Dendritic Cells Express AIRE in the Absence of Tissue Restricted Antigens. Front Immunol **2018**;9:2902 doi 10.3389/fimmu.2018.02902.
38. Rubtsova K, Rubtsov AV, Thurman JM, Mennona JM, Kappler JW, Marrack P. B cells expressing the transcription factor T-bet drive lupus-like autoimmunity. J Clin Invest **2017**;127(4):1392-404 doi 10.1172/JCI91250.
39. Phalke S, Rivera-Correa J, Jenkins D, Flores Castro D, Giannopoulou E, Pernis AB. Molecular mechanisms controlling age-associated B cells in autoimmunity. Immunol Rev **2022**;307(1):79-100 doi 10.1111/imr.13068.
40. Melo-Lima BL, Poras I, Passos GA, Carosella ED, Donadi EA, Moreau P. The Autoimmune Regulator (Aire) transactivates HLA-G gene expression in thymic epithelial cells. Immunology **2019**;158(2):121-35 doi 10.1111/imm.13099.
41. Gerber PP, Duncan LM, Greenwood EJ, Marelli S, Naamati A, Teixeira-Silva A, *et al.* A protease-activatable luminescent biosensor and reporter cell line for authentic SARS-CoV-2 infection. PLoS Pathog **2022**;18(2):e1010265 doi 10.1371/journal.ppat.1010265.
42. Baang JH, Smith C, Mirabelli C, Valesano AL, Manthei DM, Bachman MA, *et al.* Prolonged Severe Acute Respiratory Syndrome Coronavirus 2 Replication in an Immunocompromised Patient. J Infect Dis **2021**;223(1):23-7 doi 10.1093/infdis/jiaa666.
43. Buckland MS, Galloway JB, Fhogartaigh CN, Meredith L, Provine NM, Bloor S, *et al.* Treatment of COVID-19 with remdesivir in the absence of humoral immunity: a case report. Nat Commun **2020**;11(1):6385 doi 10.1038/s41467-020-19761-2.

Figure Legends

Figure 1. Single cell transcriptional signature of ABCs. Magnetically enriched B cells from PBMCs were analysed through droplet-based single cell RNA sequencing technology (n = 52402). **a-c.** UMAPs of cells from all individuals coloured by (**a)** annotated Louvain clusters, (**b)** health condition (HC, healthy controls; ICB, immune-checkpoint blockade treated cancer patients; IEI CTLA-4/NFKB1, inborn errors of immunity, CTLA-4 and NFKB1 mutants respectively; SLE, systemic lupus erythematosus patients), and (**c**) gender. Dots represent a single cell. **d.** Proportion of total cells from each health condition belonging to each cluster, clusters coloured as indicated. **e.** Expression of 46 genes in each B cell cluster which define the different subpopulations. Dot size represents proportion of cells within a cluster expressing the indicated gene and colours represent the average expression level. **f.** Heatmap representing scaled expression values of genes associated with antigen uptake, processing and class-II presentation in each cluster. **g.** Gene ontology enrichment analysis of biological processes associated with upregulated genes in classical ABCs. GO terms with highest fold enrichments are ranked by -log (P value). Dot size is proportional to the fold enrichment.

Figure 2. Classical ABC expression of AIRE and AIRE target genes. a. Volcano plot displaying the differentially expressed genes in ABCs compared to other B cell clusters. Upregulated genes associated with antigen uptake, processing and MHC class II presentation are labelled. AIRE gene is coloured in green. **b.** UMAP of all B cells coloured by kernel density estimation of AIRE expression level across all identified B cell subsets. UMAP showing the different B cell clusters inset. **c.** Percentage of cells which are $AIRE^{+}$ in each cluster. **d.** Violin plot comparing expression score of tissue restricted antigens (gene set from Yamano *et al*) in AIRE expressing cells and all other cells. Statistical testing via Wilcoxon rank sum test.

Figure 3. Correlation between frequency of ABCs and neutralising antibody response to COVID-19 vaccination. a. Cohort details. Samples were collected at days 0, 8, 21 and 105 after the 2nd dose of BNT162b2 vaccine (healthy controls (HC), grey; patients with rare inborn errors of immunity (IEI): NFKB1 red, CTLA-4 and unclassified orange; patients treated with ICB, blue). **b**. Representative FACS contour plots of $CD21^{lo}CD11c^{+}$ ABCs in CD19 B cells. **c.** Frequencies of B cells within total lymphocytes and frequencies ABCs within total B cells, at day 0. **d.** Neutralising antibody titres at 50% inhibition (NT_{50}) against wildtype SARS-CoV-2 at indicated timepoints after 2^{nd} vaccine dose. The limit of detection of

the assay is indicated (grey dotted line at NT_{50} = 20), and an arbitrary threshold at the highest NT_{50} from the HC group at day 0 (brown dashed line). **e**. Correlations between frequencies of ABCs amongst B cells at day 0 and NT_{50}s at days 8, 21 and 105. **f.** Frequencies of ABCs amongst $CD19^+$ B cells for individuals above and below the arbitrary threshold indicated in panel **d** at day 105. Where specified, differences between groups were determined using two-tailed non-parametric Mann Whitney tests. Spearman's rank correlation coefficients (rho) are shown, together with indicative linear regression lines plus, **c-f** each point represents one individual.

Figure 4. Correlation between frequency of RBD-specific B cells and neutralising antibody response to COVID-19 vaccination. a. Representative flow cytometry plots displaying RBD-specific B cell populations amongst total $CD19^+$ B cells for different study groups 0, 8, 21 and 105 days (D0, D8, D21 and D105) after the second dose of BNT162b2 vaccine. **b.** Frequencies of RBD-specific B cells amongst all $CD19^+$ B cells in HC, IEI and ICB patients at days 0, 8, 21 and 105. Each dot represents a single individual (healthy controls (HC), grey; patients with rare inborn errors of immunity (IEI): (NFKB1) red, (CTLA-4 and unclassified) orange; patients treated with ICB, blue). **c.** Correlations between frequencies of RBD-specific B cells amongst all CD19+ B cells at day 0 and NT_{50}s at days 8, 21 and 105. **d.** Correlations between frequencies of ABCs amongst B cells at day 0 and frequencies of RBD-specific B cells at days 8, 21 and 105. **e.** UMAP projection of total B cells from all donors at all time points displaying memory B cell (MBCs, purple) and plasmablast (PBs, green) populations. **f.** RBD-specific B cells with MBC (purple) or PB (green) phenotype are shown at days 0, 8, 21 and 105 (columns) for HC (top), IEI (middle) and ICB (bottom) groups. **g.** Kinetics of RBD-binding cell frequencies amongst MBCs or PBs. Samples with no detectable levels of RBD-specific cells are plotted at an arbitrary value of 10^{-4} in (B) and 10^{-3} in (G). Where specified, statistical significance between groups was determined using one-way ANOVA (G). P values for panel B are shown in Supplementary Fig. 5. Spearman's rank correlation coefficients (rho) are shown, together with indicative linear regression lines where appropriate.

Supplemental Figures and Tables

Supplementary Figure 1. Age-associated B cells converge transcriptionally amongst individuals despite their disparate aetiology. a. Magnetically enriched B cells from PBMCs were labelled using barcode featuring and analysed through droplet-based single cell RNA sequencing technology. Expression of surface protein CD11c against CD21 in all B cells. **b.** Violin plots of CD21 and CD11c surface protein expression in ABCs versus other B cells with CLR-normalised expression values. Statistical testing via Wilcoxon rank sum test with Bonferroni correction. **c.** UMAPs of total B cells per health condition coloured by annotated clusters (HC, healthy controls; ICB, immune-checkpoint blockade treated patients; IEI CTLA-4/NFKB1, immune errors of immunity, CTLA-4 and NFKB1 mutants respectively; SLE, systemic lupus erythematosus patients). **d.** Number of samples from each gender, coloured as indicated, in each health condition group. **e.** Clusters, with percentages of cells from each gender coloured as indicated. **f.** Clusters, with percentages of cells from each health condition coloured as indicated. Dots represent a single cell (a-c).

Supplementary Figure 2. Class-I related genes expression within the different B cell clusters. Heatmap representing scaled expression values of genes associated with antigen uptake, processing and class-I presentation in each cluster.

Supplementary Figure 3. Classical age-associated B cells upregulate AIRE in different health conditions. a. UMAP of total B cells showing AIRE+ cells coloured by density of expression. **b.** UMAPs of total B cells per health condition coloured by kernel density estimation of AIRE expression level (HC, healthy controls; ICB, immune-checkpoint blockade treated patients; IEI CTLA-4/NFKB1, immune errors of immunity, CTLA-4 and NFKB1 mutants respectively; SLE, systemic lupus erythematosus patients).

Supplementary Figure 4. Decline of the neutralisation capacity after second dose of BNT162b2 vaccine. a. Gating strategy to identify age-associated B cell (ABCs, SingletsLiveCD3-CD19+CD21-CD11+) subset from total B cell population (CD19+). **b.** IgG end-point titres specific to the spike (S), receptor-binding domain (RBD) and nucleocapsid (NCP) proteins of SARS-CoV2 at day 21 after the second dose of BNT162b2 vaccine. **c**. Neutralising antibody titres at 50% inhibition (NT_{50}) against wildtype SARS-CoV-2 at indicated timepoints after 2^{nd} dose of BNT162b2 vaccine (healthy controls (HC), grey; patients with rare inborn errors of immunity (IEI): (NFKB1) red, (CTLA-4 and unclassified) orange; patients treated with ICB, blue). The limit of detection of the assays is shown (grey

dotted line at NT_{50} = 20). Where specified, statistical significance between groups was determined using 2-way ANOVA with Tukey's multiple comparisons test.

Supplementary Figure 5. Accumulation of age-associated B cells impairs humoral responses to COVID-19 vaccination. a. Gating strategy to identify RBD-specific B cell (SingletsLiveStreptavidinBV711-CD19+RBDAF488+RBDAF647+) subset after the second dose of BNT162b2 vaccine. **b.** Two-way ANOVA with Tukey's multiple comparisons test for statistical analysis of RBD-binding B cells amongst groups. P values are indicated. **c.** Multiple variable dot plots showing Spearman correlations between RBD-specific B cells and NT_{50} at days 8, 21 and 105 after the second dose of BNT162b2 vaccine. Dot size represents ABC frequency per individual. **d.** Correlation matrix using Spearman's rank correlation amongst ABC and B cell frequencies at day 0; and RBD frequencies and NT_{50} values at days 0, 8, 21 and 105 for the second dose of BNT162b2 vaccine. Data include all samples enrolled in the study; Rho values are shown per correlation.

Supplementary Table 1 – Differentially expressed genes related to Figure 1e

Supplementary Table 2 – Top 20 biological processes from the gene ontology enrichment analysis

Supplementary Table 3 – Differentially expressed genes related to Figure 2a

Supplementary Table 4 – Characteristics of Inborn Errors of Immunity (IEI) patients and Immune Checkpoint blockade treated (ICB) patients

Supplementary Table 5 – Flow cytometry antibodies details

Acknowledgements

This work was funded by the MRC (grant ref. MC_UU_00025/12), the Medical Research Foundation (MRF-057-0002-RG-THAV-C0798) and The Evelyn Trust (grant number 20/40). NM was supported by the MRC (CSF ref. MR/P008801/1 to NJM), NHSBT (grant ref. WPA15-02 to NJM) and Addenbrooke's Charitable Trust (grant ref. 900239 to NJM). The authors thank the Flow Cytometry Services provided by Lucia Pereira Giraldez at the MRC-Toxicology Unit, University of Cambridge and Katarzyna Kania from CRUK-CI-Genomics for advice on single cell RNA sequencing experiments.

Competing interest

The authors declare no competing financial interests.

Corresponding Authors

739 Correspondence and requests for materials should be addressed to James Thaventhiran
740 (jedt2@cam.ac.uk) and Juan Carlos Yam-Puc (jcy28@cam.ac.uk).
741

CITIID-NIHR BioResource COVID-19 Collaboration

Stephen Baker[2,6], John Bradley[1,3,6,11,15] , Patrick Chinnery[3,23,24], Daniel Cooper[11, 25], Gordon Dougan[2,6], Ian Goodfellow[7], Ravindra Gupta[2,6,13,16], Nathalie Kingston[3,4], Paul J. Lehner[2,6,13], Paul A. Lyons[2,6], Nicholas J. Matheson[2,6,13,33], Caroline Saunders[9], Kenneth G. C. Smith[2,6], Charlotte Summers[6,12,26], James Thaventhiran[19], M. Estee Torok [6,13,14], Mark R. Toshner[6,8,26], Michael P. Weekes[2,6,13,34], Gisele Alvio[9], Sharon Baker[9], Areti Bermperi[9], Karen Brookes[9], Ashlea Bucke, Jo Calder, Laura Canna, Cherry Crucusio, Isabel Cruz[9], Ranalie de Jesus[9], Katie Dempsey[9], Giovanni Di Stephano[9], Jason Domingo[9], Anne Elmer[9], Julie Harris, Sarah Hewitt, Heather Jones[9], Sherly Jose[9], Jane Kennet, Yvonne King, , Jenny Kourampa[9], Emily Li, Caroline McMahon[9], Anne Meadows, Vivien Mendoza[9], Criona O'Brien, Charmain Ocaya[9], Ciro Pasquale[9], Marlyn Perales[9], Jane Price, Rebecca Rastall, Carla Ribeiro[9], Jane Rowlands, Valentina Ruffolo, Hugo Tordesillas, Phoebe Vargas[9], Bensi Vergese[9], Laura Watson[9], Jieniean Worsley[9], Julie-Ann Zerrudo[9], Laura Bergamashi[2,6], Ariana Betancourt, Georgie Bower, Ben Bullman, Chiara Cossetti, Aloka De Sa, Benjamin J. Dunmore, Maddie Epping, Stuart Fawke, Stefan Gräf [3,6], Richard Grenfell, Andrew Hinch, Josh Hodgson, Christopher Huang, Oisin Huhn, Kelvin Hunter[2,6], Isobel Jarvis, Emma Jones, Maša Josipović, Ekaterina Legchenko, Daniel Lewis, Joe Marsden, Jennifer Martin, Federica Mescia[2,6], Ciara O'Donnell, Ommar Omarjee, Marianne Perera, Linda Pointon, Nicole Pond, Nathan Richoz, Nika Romashova, Natalia Savoinykh, Rahul Sharma, Joy Shih, Mateusz Strezlecki, Rachel Sutcliffe, Tobias Tilly, Zhen Tong, Carmen Treacy, Lori Turner, Jennifer Wood, Marta Wylot, John Allison[3,4], Heather Biggs[3,18], John R. Bradley[1,3,6,11,15], Helen Butcher[3,5], Daniela Caputo[3,5], Matt Chandler[3,5], Patrick Chinnery[3,23,24], Debbie Clapham-Riley[3,5], Eleanor Dewhurst[3,5], Christian Fernandez[3,4,] Anita Furlong[3,5], Barbara Graves[3,5], Jennifer Gray[3,5], Sabine Hein[3,5], Tasmin Ivers[3,5], Emma Le Gresley[3,5], Rachel Linger[3,5], Mary Kasanicki[3,11], Rebecca King[3,5], Nathalie Kingston[3,4], Sarah Meloy[3,5], Alexei Moulton[3,5], Francesca Muldoon[3,5], Nigel Ovington[3,4], Sofia Papadia[3,5], Christopher J. Penkett[3,4], Isabel Phelan[3,5], Venkatesh Ranganath[3,4], Roxana Paraschiv[3,4], Abigail Sage[3,5], Jennifer Sambrook[3,4], Ingrid Scholtes[3,5], Katherine Schon[3,17,18], Hannah Stark[3,5], Kathleen E. Stirrups[3,4], Paul Townsend[3,4], Neil Walker[3,4], Jennifer Webster[3,5], Mayurun Selvan[35], Petra, Polgarova[12],Sarah L. Caddy[2,6], Laura G. Caller[20,21], Yasmin Chaudhry[7], Martin D. Curran[22], Theresa Feltwell[6], Stewart Fuller[20], Iliana Georgana[7], Grant Hall[7], William L. Hamilton[6,13,14], Myra Hosmillo[7], Charlotte J. Houldcroft[6], Rhys Izuagbe[7], Aminu S. Jahun[7], Fahad A. Khokhar[2,6], Anna G. Kovalenko[7], Luke W. Meredith[7], Surendra Parmar[22], Malte L. Pinckert[7], Anna Yakovleva[7], Emily C. Horner[19], Lucy Booth[19], Alexander Ferreira[19], Rebecca Boston[19], Robert Hughes[19], Juan Carlos Yam Puc[19], Nonantzin Beristain-Covarrubias[19], Maria Rust[19],

Thevinya Gurugama[19], Lihinya Gurugama[19], Thomas Mulroney[19], Sarah Spencer[19], Zhaleh Hosseini[19], Kate Williamson[19].

[1]NIHR Cambridge Biomedical Research Centre, Cambridge Biomedical Campus, Cambridge, UK

[2]Cambridge Institute of Therapeutic Immunology and Infectious Disease (CITIID), Jeffrey Cheah Biomedical Centre, Cambridge Biomedical Campus, Cambridge, UK

[3]NIHR BioResource, Cambridge University Hospitals NHS Foundation Trust, Cambridge Biomedical Campus, Cambridge, UK

[4]Department of Haematology, School of Clinical Medicine, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK

[5]Department of Public Health and Primary Care, School of Clinical Medicine, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK

[6]Department of Medicine, School of Clinical Medicine, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK

[7]Division of Virology, Department of Pathology, University of Cambridge, Cambridge, UK

[8]Royal Papworth Hospital NHS Foundation Trust, Cambridge, UK

[9]Cambridge Clinical Research Centre, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

[10]Intensive Care Unit, Royal Papworth Hospital NHS Foundation Trust, Cambridge, UK

[11]Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge Biomedical Campus, Cambridge, UK

[12]Intensive Care Unit, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge Biomedical Campus, Cambridge, UK

[13]Department of Infectious Diseases, Addenbrooke's Hospital, Cambridge University NHS Hospitals Foundation Trust, Cambridge, UK

[14]Department of Microbiology, Addenbrooke's Hospital, Cambridge University NHS Hospitals Foundation Trust, Cambridge, UK

[15]Department of Renal Medicine, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

[16]Africa Health Research Institute, Durban, South Africa

[17]Clinical Genetics, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

[18]Department of Clinical Neurosciences, School of Clinical Medicine, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK

[19]MRC Toxicology Unit, Gleeson Building, Tennis Court Road, Cambridge, UK

[20]University of Cambridge, Cambridge, UK

[21]The Francis Crick Institute, London, UK

[22]Public Health England, Clinical Microbiology and Public Health Laboratory, Cambridge, UK

[23]Department of Clinical Neurosciences, School of Clinical Medicine, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK

[24]Medical Research Council Mitochondrial Biology Unit, Cambridge Biomedical Campus, Cambridge, UK

[25]Global and Tropical Health Division, Menzies School of Heath Research and Charles Darwin University, Darwin, Northern Territory, Australia

[26]Heart and Lung Research Institute, Cambridge Biomedical Campus, Cambridge, UK

[27]Department of Rheumatology, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

[28]Cambridge Cancer Trials Centre, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

[29]Department of Paediatrics, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK

[30]Patient Safety, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

[31]Clinical Research Network: Eastern, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

[32]Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

[33]NHS Blood and Transplant, Cambridge, UK

[34]Cambridge Institute for Medical Research, Biomedical Campus, Hills Rd, Cambridge UK

[35] Department of Respiratory Medicine, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

Method references

1. Zheng, G.X., et al., Massively parallel digital transcriptional profiling of single cells. Nat Commun, 2017. 8: p. 14049.
2. Roelli, P., B. Flynn, and G. Gui, Hoohm/CITE-seq-Count: 1.5.0. . Zenodo, 2019.
3. Hao, Y., et al., Integrated analysis of multimodal single-cell data. Cell, 2021. 184(13): p. 3573-3587 e29.
4. McCarthy, D.J., et al., Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. Bioinformatics, 2017. 33(8): p. 1179-1186.
5. Bhamidipati, K., et al., CD52 Is Elevated on B cells of SLE Patients and Regulates B Cell Function. Front Immunol, 2020. 11: p. 626820.
6. Ashburner, M., et al., Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet, 2000. 25(1): p. 25-9.
7. Gene Ontology, C., The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Res, 2021. 49(D1): p. D325-D334.
8. Mi, H., et al., PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res, 2019. 47(D1): p. D419-D426.
9. Yamano, T., et al., Thymic B Cells Are Licensed to Present Self Antigens for Central T Cell Tolerance Induction. Immunity, 2015. 42(6): p. 1048-61.
10. Alquicira-Hernandez, J. and J.E. Powell, Nebulosa recovers single cell gene expression signals by kernel density estimation. Bioinformatics, 2021.
11. Xiong, X., et al., A thermostable, closed SARS-CoV-2 spike protein trimer. Nat Struct Mol Biol, 2020. 27(10): p. 934-941.
12. Collier, D.A., et al., Age-related immune response heterogeneity to SARS-CoV-2 vaccine BNT162b2. Nature, 2021. 596(7872): p. 417-422.
13. Stadlbauer, D., et al., SARS-CoV-2 Seroconversion in Humans: A Detailed Protocol for a Serological Assay, Antigen Production, and Test Setup. Curr Protoc Microbiol, 2020. 57(1): p. e100.
14. Daly, J.L., et al., Neuropilin-1 is a host factor for SARS-CoV-2 infection. Science, 2020. 370(6518): p. 861-865.
15. Moore, S.C., et al., Amplicon-Based Detection and Sequencing of SARS-CoV-2 in Nasopharyngeal Swabs from Patients With COVID-19 and Identification of Deletions in the Viral Genome That Encode Proteins Involved in Interferon Antagonism. Viruses, 2020. 12(10).
16. Patterson, E.I., et al., Methods of Inactivation of SARS-CoV-2 for Downstream Biological Assays. J Infect Dis, 2020. 222(9): p. 1462-1467.

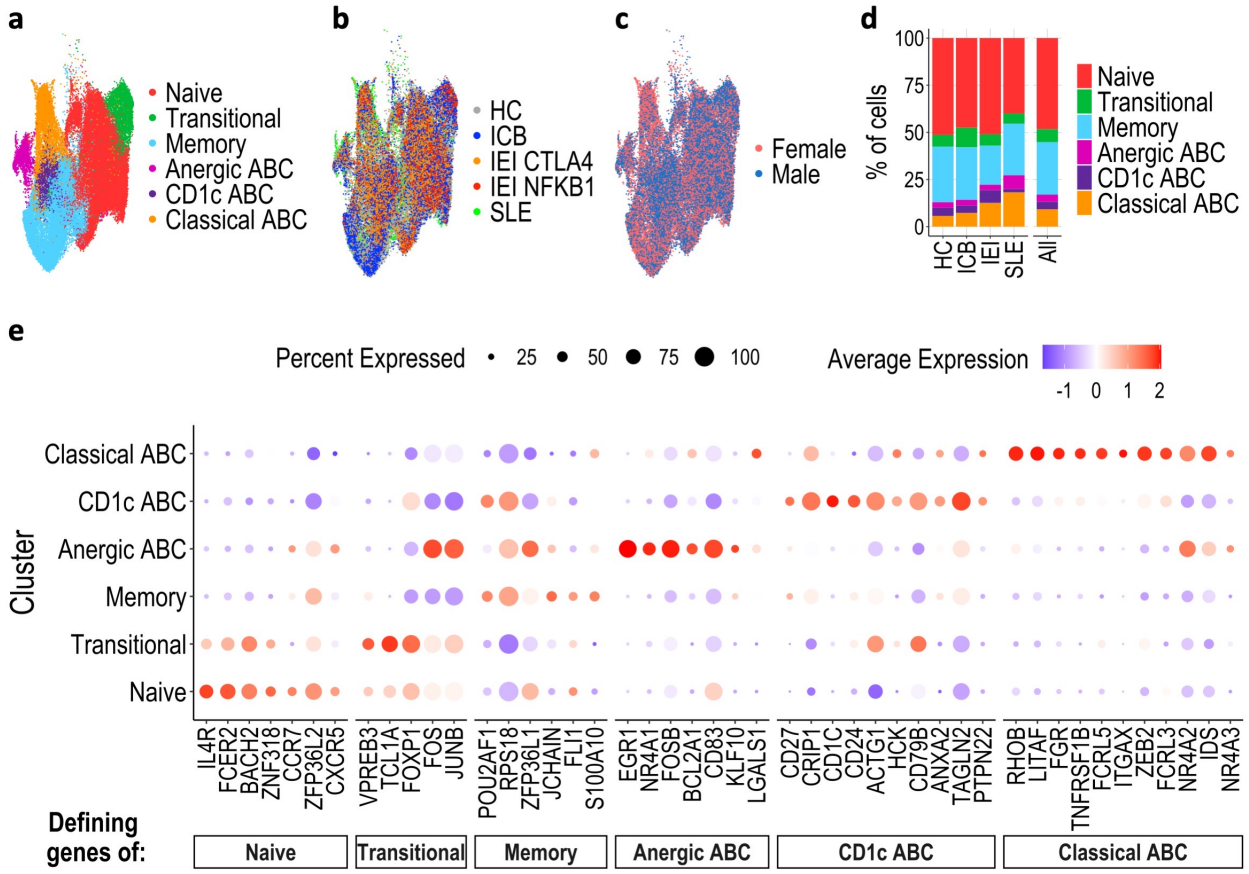
17. Bergamaschi, L., et al., Longitudinal analysis reveals that delayed bystander CD8+ T cell activation and early immune pathology distinguish severe COVID-19 from mild disease. Immunity, 2021. 54(6): p. 1257-1275 e8.

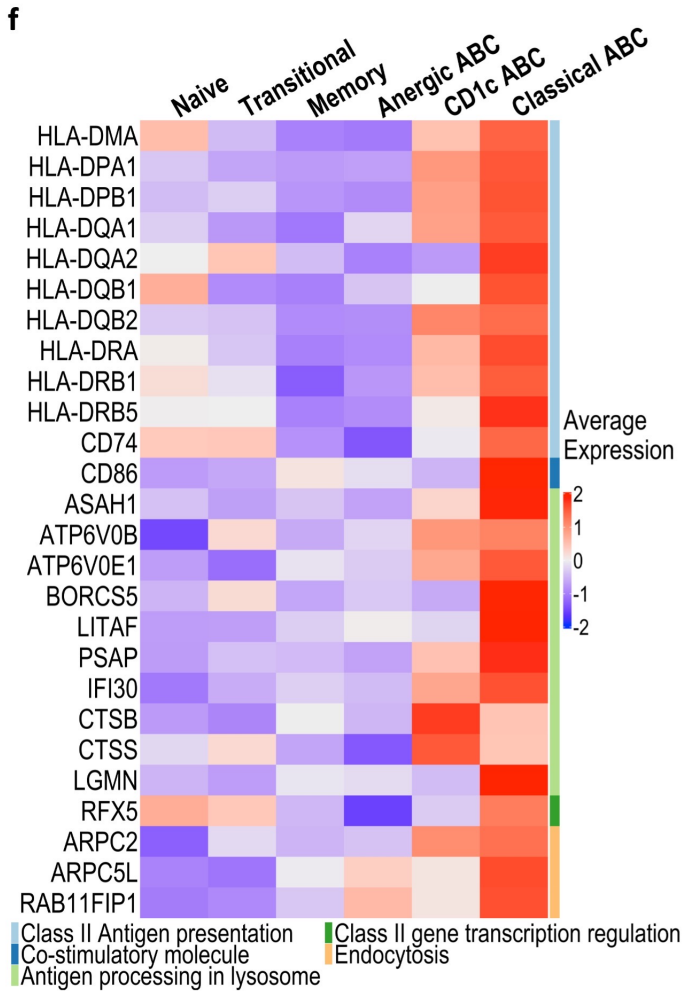
18. Gerber, P.P., et al., A protease-activatable luminescent biosensor and reporter cell line for authentic SARS-CoV-2 infection. PLoS Pathog, 2022. 18(2): p. e1010265.

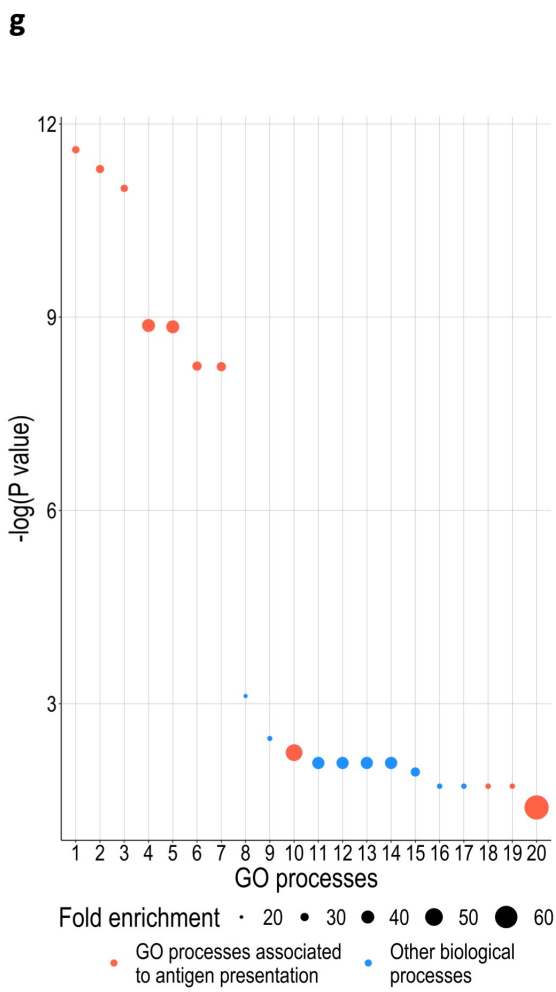
19. Knezevic, I., et al., WHO International Standard for evaluation of the antibody response to COVID-19 vaccines: call for urgent action by the scientific community. Lancet Microbe, 2022. 3(3): p. e235-e240.

20. Hedhammar, M. and S. Hober, Z(basic)--a novel purification tag for efficient protein recovery. J Chromatogr A, 2007. 1161(1-2): p. 22-8.

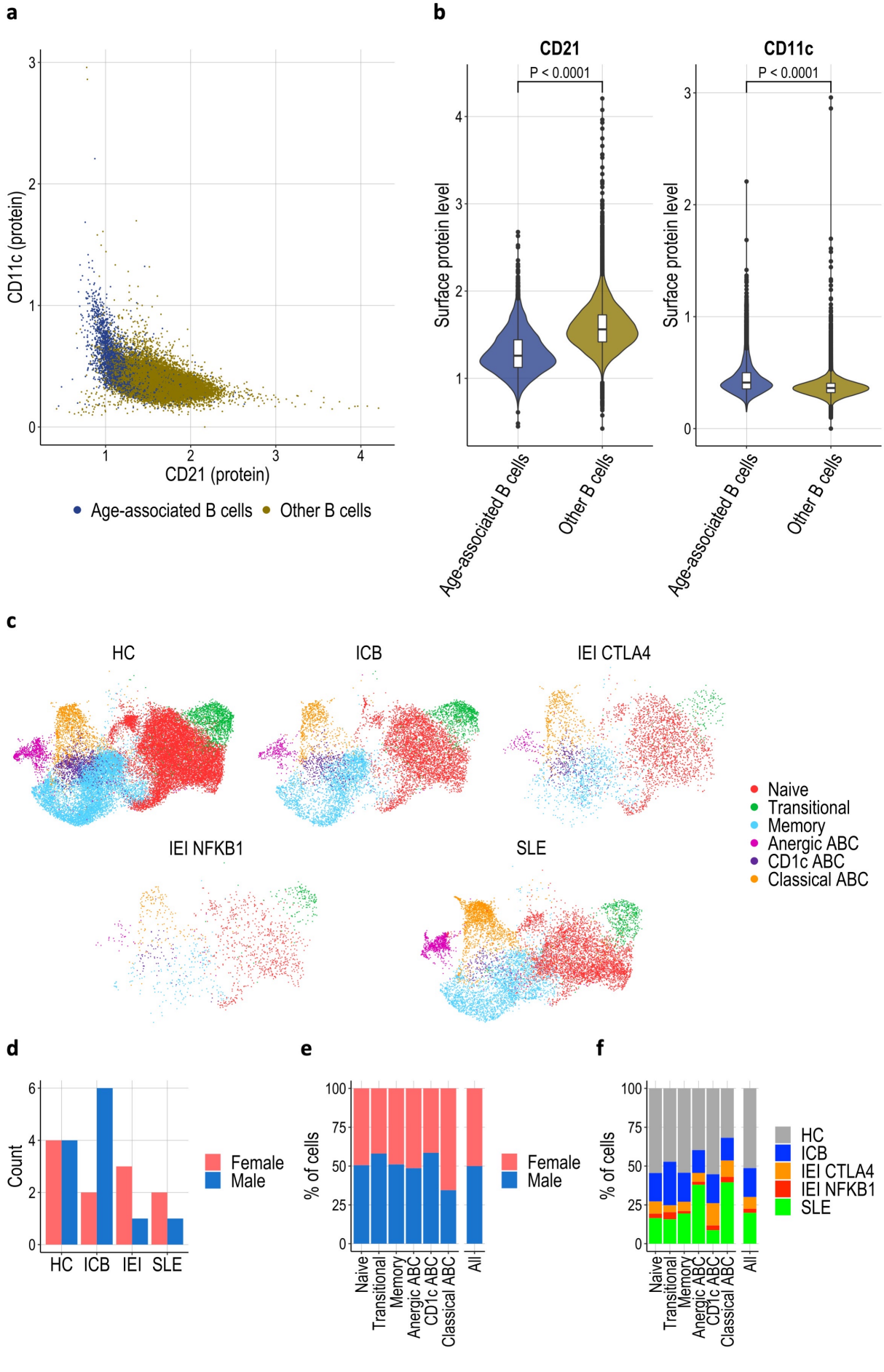
Figure 1







Supplementary Figure 1



Supplementary Figure 2

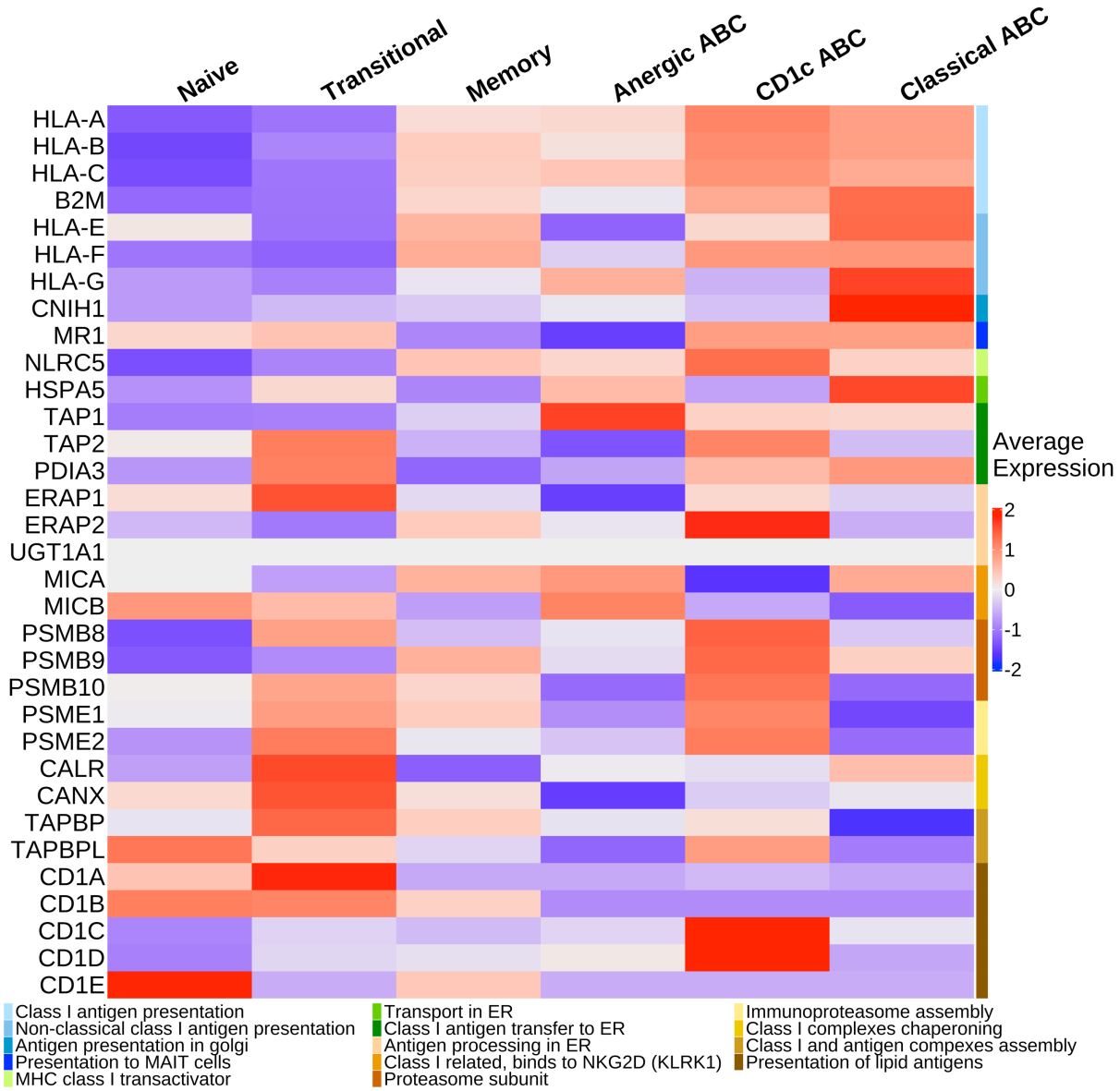
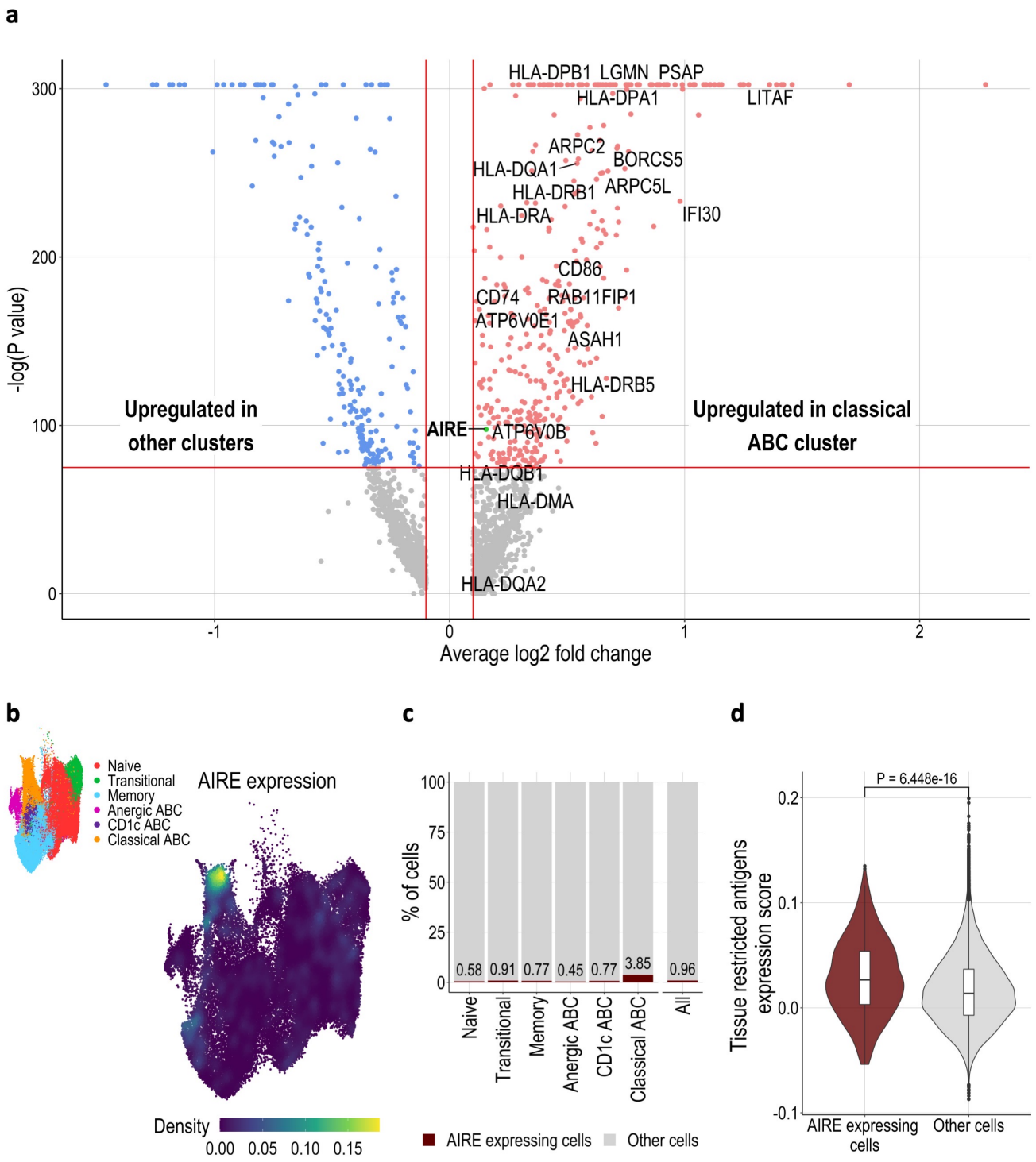


Figure 2



Supplementary Figure 3

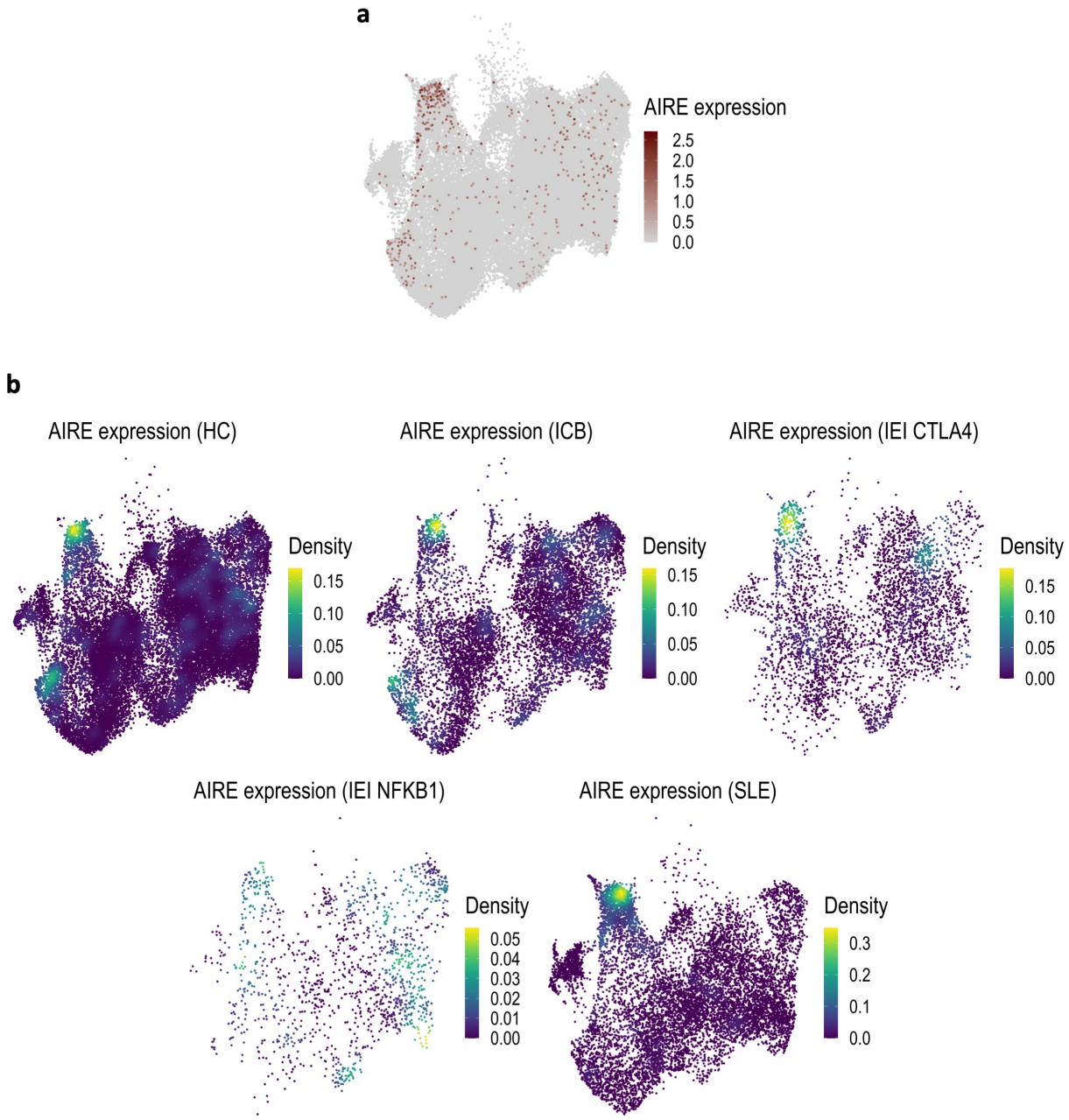
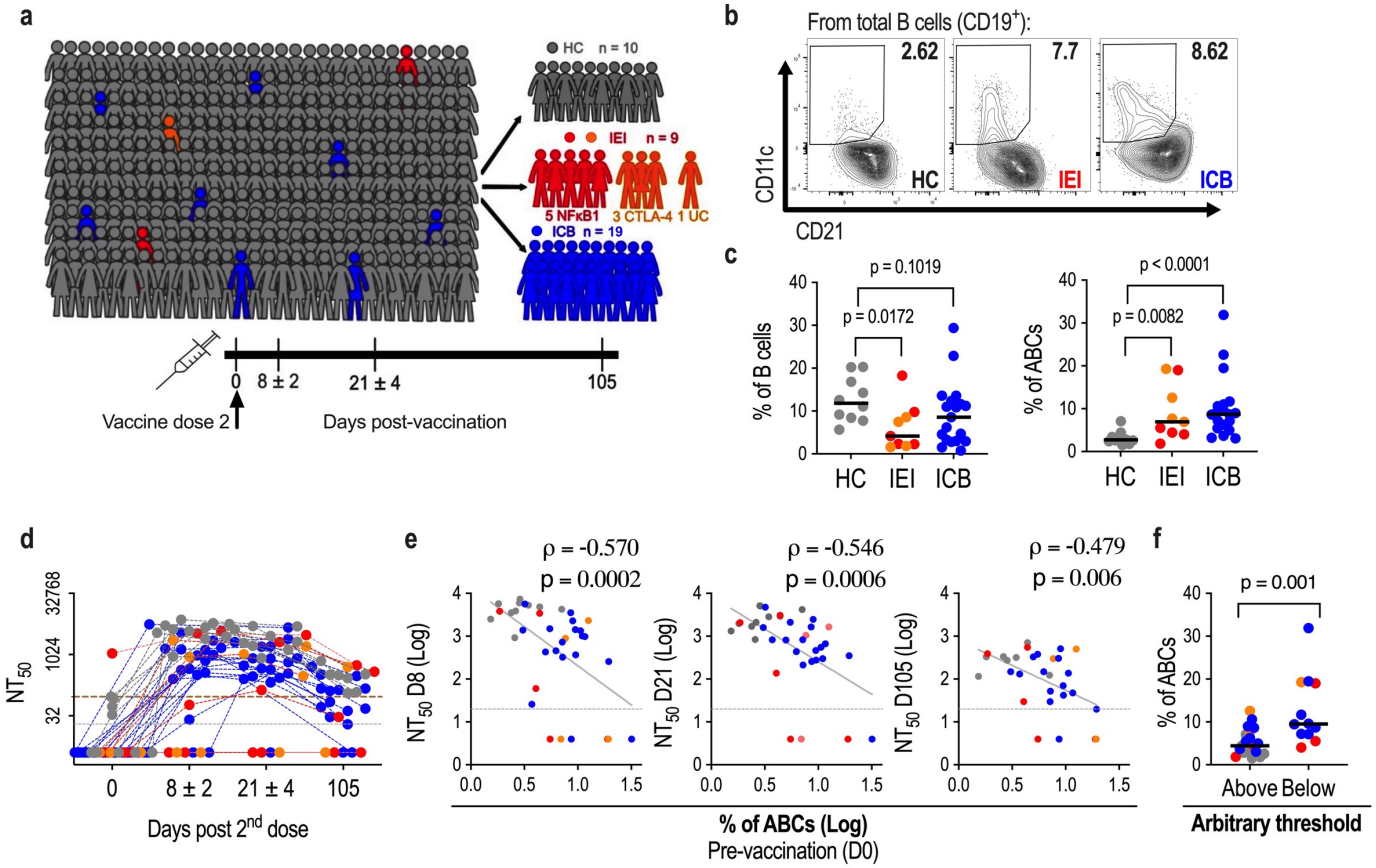


Figure 3



Supplementary Figure 4

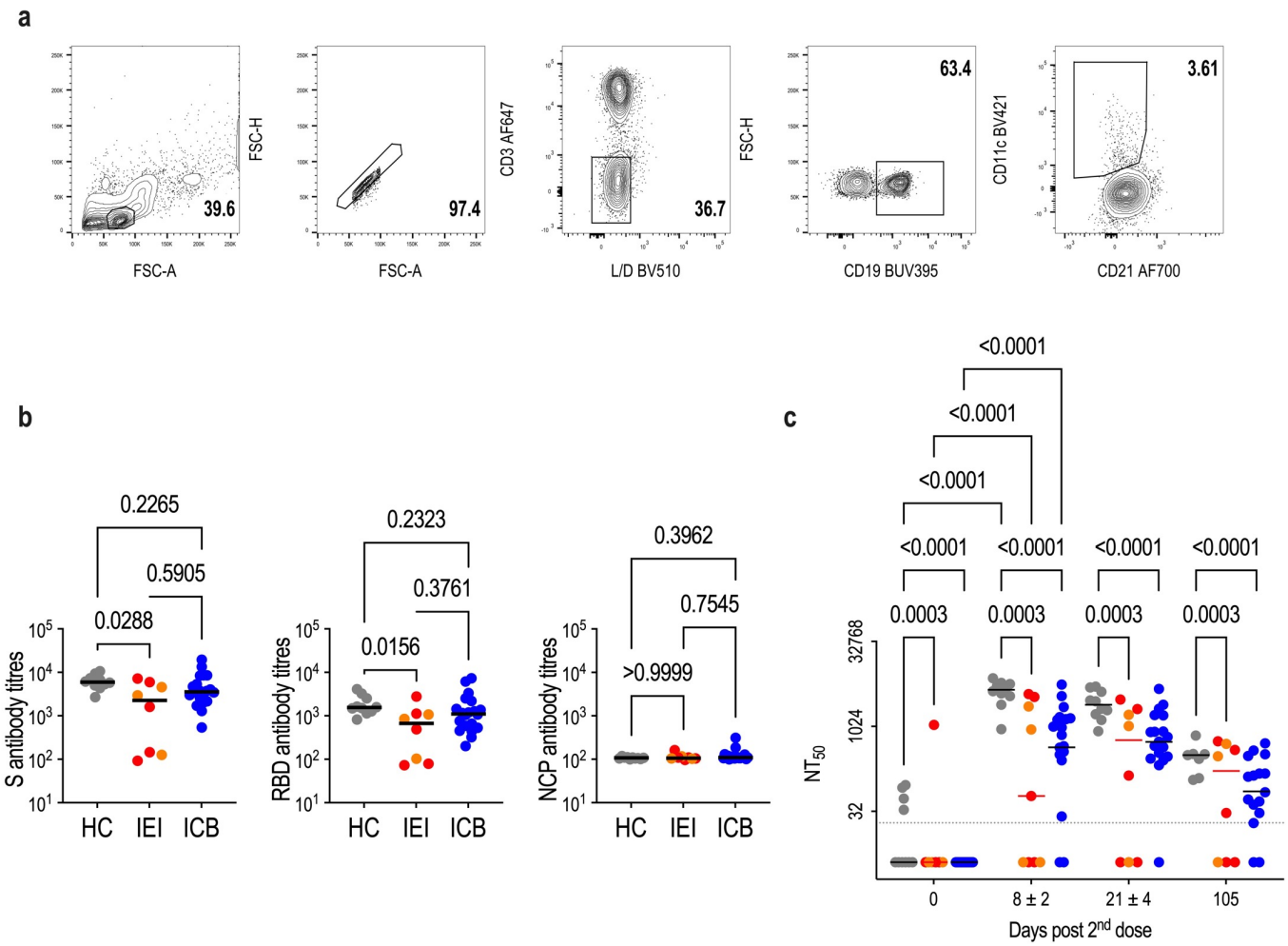
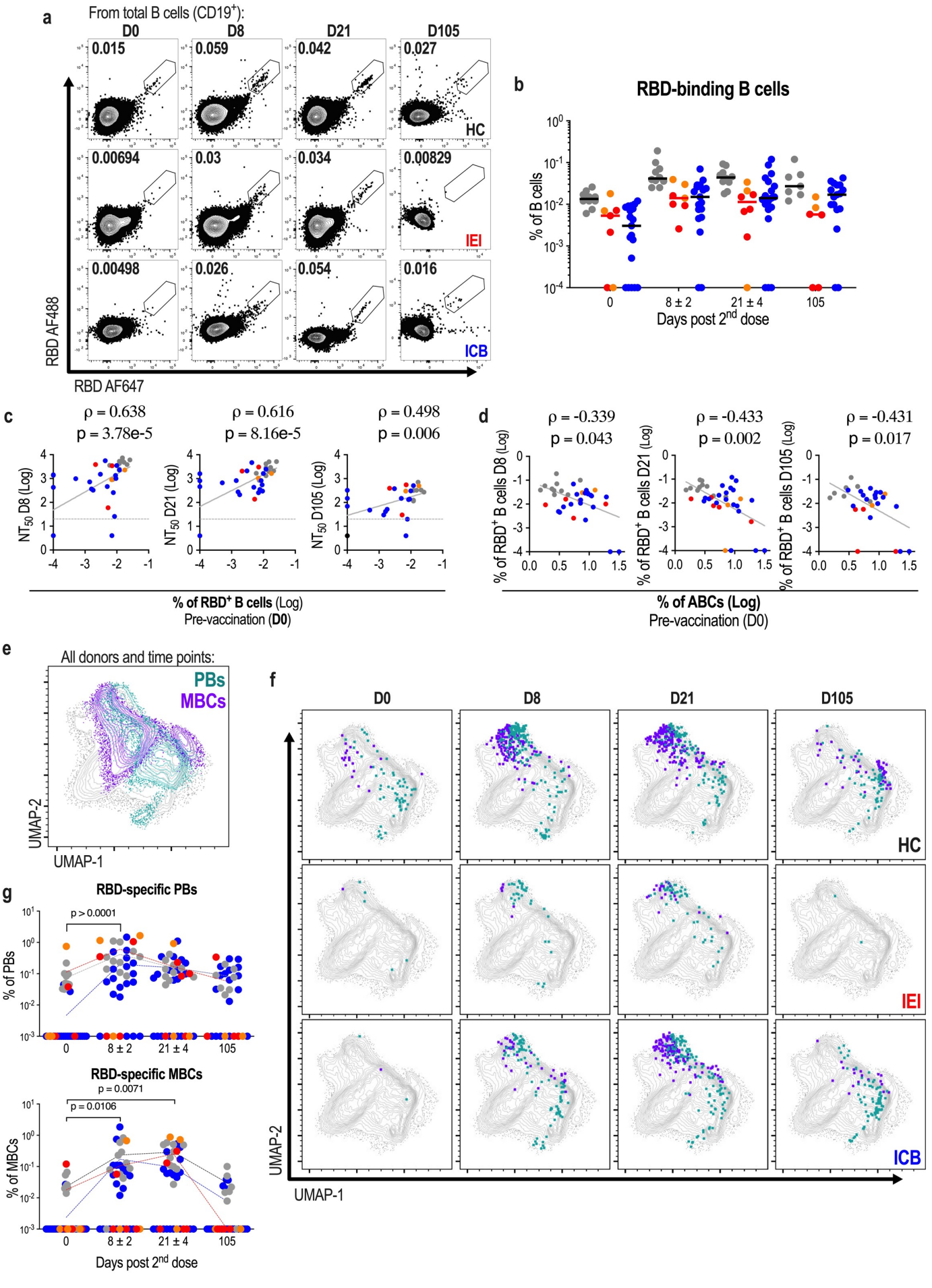


Figure 4



Supplementary Figure 5