# Systems vaccinology of the BNT162b2 mRNA vaccine in humans

**Prabhu S. Arunachalam**[1,14], **Madeleine K. D. Scott**[1,2,14], **Thomas Hagan**[3,4,14], **Chunfeng Li**[1], **Yupeng Feng**[1], **Florian Wimmers**[1], **Lilit Grigoryan**[1], **Meera Trisal**[1], **Venkata Viswanadh Edara**[5], **Lilin Lai**[5], **Sarah Esther Chang**[1,6], **Allan Feng**[1,6], **Shaurya Dhingra**[1,6], **Mihir Shah**[7], **Alexandra S. Lee**[7], **Sharon Chinthrajah**[7], **Sayantani B. Sindher**[7], **Vamsee Mallajosyula**[1], **Fei Gao**[1], **Natalia Sigal**[1], **Sangeeta Kowli**[1], **Sheena Gupta**[1], **Kathryn Pellegrini**[5], **Gregory Tharp**[5], **Sofia Maysel-Auslender**[1], **Sydney Hamilton**[5], **Hadj Aoued**[5], **Kevin Hrusovsky**[8], **Mark Roskey**[8], **Steven E. Bosinger**[4,9], **Holden T. Maecker**[1], **Scott D. Boyd**[7,10], **Mark M. Davis**[1,10,11], **Paul J. Utz**[1,6], **Mehul S. Suthar**[5], **Purvesh Khatri**[1,2,✉], **Kari C. Nadeau**[7,12,13,✉], **Bali Pulendran**[1,10,11,✉]

[1]Institute for Immunity, Transplantation and Infection, Stanford University, Stanford, CA, USA.

[2]Center for Biomedical Informatics, Department of Medicine, Stanford University School of Medicine, Stanford University, Stanford, CA, USA.

[3]Division of Infectious Diseases, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA.

[4]Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, USA.

[5]Yerkes National Primate Research Center, Atlanta, GA, USA.

[6]Department of Medicine, Division of Immunology and Rheumatology, Stanford University School of Medicine, Stanford University, Stanford, CA, USA.

[7]Sean N. Parker Center for Allergy and Asthma Research, Stanford University, Stanford, CA, USA.

[8]Quanterix, Billerica, MA, USA.

[9]Department of Pathology, Emory University School of Medicine, Atlanta, GA, USA.

[10]Department of Pathology, Stanford University School of Medicine, Stanford University, Stanford, CA, USA.

[11]Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford University, Stanford, CA, USA.

[12]Howard Hughes Medical Institute, Stanford University, Stanford, CA, USA.

[13]Department of Medicine, Division of Pulmonary, Allergy and Critical Care Medicine, Stanford University School of Medicine, Stanford University, Stanford, CA, USA.

[14]These authors contributed equally: Prabhu S. Arunachalam, Madeleine K. D. Scott, Thomas Hagan.

## Abstract

The emergency use authorization of two mRNA vaccines in less than a year from the emergence of SARS-CoV-2 represents a landmark in vaccinology[1,2]. Yet, how mRNA vaccines stimulate the immune system to elicit protective immune responses is unknown. Here we used a systems vaccinology approach to comprehensively profile the innate and adaptive immune responses of 56 healthy volunteers who were vaccinated with the Pfizer–BioNTech mRNA vaccine (BNT162b2). Vaccination resulted in the robust production of neutralizing antibodies against the wild-type SARS-CoV-2 (derived from 2019-nCOV/USA_WA1/2020) and, to a lesser extent, the B.1.351 strain, as well as significant increases in antigen-specific polyfunctional CD4 and CD8 T cells after the second dose. Booster vaccination stimulated a notably enhanced innate immune response as compared to primary vaccination, evidenced by (1) a greater frequency of $CD14^+CD16^+$ inflammatory monocytes; (2) a higher concentration of plasma IFNγ; and (3) a transcriptional signature of innate antiviral immunity. Consistent with these observations, our single-cell transcriptomics analysis demonstrated an approximately 100-fold increase in the frequency of a myeloid cell cluster enriched in interferon-response transcription factors and reduced in AP-1 transcription factors, after secondary immunization. Finally, we identified distinct innate pathways associated with CD8 T cell and neutralizing antibody responses, and show that a monocyte-related signature correlates with the neutralizing antibody response against the B.1.351 variant. Collectively, these data provide insights into the immune responses induced by mRNA vaccination and demonstrate its capacity to prime the innate immune system to mount a more potent response after booster immunization.

BNT162b2 has demonstrated 95% efficacy in preventing severe COVID-19[1]. Although adaptive immunity to BNT162b2 has been characterized in humans[2,3], little is known about the innate immune response to this vaccine (or to any mRNA vaccine). Systems immunology enables the comprehensive characterization of the cellular and molecular networks that drive innate and adaptive immunity to vaccines and infections[4–8]. Here we used systems tools to analyse immune responses in 56 volunteers (Extended Data Table 1) who received two doses of BNT162b2 (Extended Data Fig. 1a). More volunteers reported mild side effects after secondary than after primary vaccination (Extended Data Table 2).

## Antibody and T cell responses

Primary vaccination induced binding antibody and neutralizing antibody responses in all but three individuals; these responses were boosted significantly after the secondary vaccination (Fig. 1a, b, Extended Data Fig. 1b–d). The neutralizing antibody response reduced about 2-fold by days 90–120, this being comparable to the response to the Moderna mRNA-1273 vaccine[9] (Fig. 1b, Extended Data Fig. 1c). There was no gender-associated difference in antibody responses; however, the neutralizing antibody response inversely correlated with age (Extended Data Fig. 1e). Four participants with previous confirmed SARS-CoV-2 infection did not have high baseline titres, but after the first immunization three of these individuals had titres that were greater than 30-fold the geometric mean titres of uninfected individuals following the first dose, but did not increase further after the boost[10] (filled black circles in Extended Data Fig. 1b, c). BNT162b2 vaccination also induced a neutralizing antibody response against the B.1.351 variant of concern, albeit at a tenfold-lower magnitude than against the wild-type WA1/2020 (WA1) strain (Fig. 1c, Extended Data Fig. 1f), consistent with previous studies[11]. The cross-neutralization potential, calculated as a ratio of neutralizing antibody responses between B.1.351 and WA1 strains, also showed a negative association with age (Extended Data Fig. 1g). Vaccination also stimulated spike-specific T cell responses, which were more readily detectable seven days after the secondary immunization (Fig. 1d, e, Extended Data Fig. 2a–e). Consistent with previous studies[3], the CD4 T cell responses were primarily of the T-helper-1 type, although there was a low-level T-helper-2 (IL-4) response (Extended Data Fig. 2b). IFNγ and TNF were the dominant responses in CD8 T cells; three individuals with no known exposure to SARS-CoV-2 responded even at baseline, suggestive of CD8 T cells that are cross-reactive to related viruses, as has previously been shown[12] (Extended Data Fig. 2d). There was no significant correlation between T cell responses and age or neutralizing antibodies (Extended Data Fig. 2f–k).

## Lack of autoantibodies or anticytokine antibodies

Several studies have demonstrated the presence of serum autoantibodies[13,14] and anticytokine antibodies[15] in individuals infected with SARS-CoV-2, and the development of new-onset antibodies in a subset of patients who are hospitalized with COVID-19[16]. We screened sera of 31 vaccinated individuals for IgG autoantibodies and anticytokine antibodies on days 0, 21 and 42 using a 55-plex antigen and 58-plex cytokine arrays. We included sera of 17 patients with autoimmune and immunodeficiency disorders as positive controls. Five vaccinated individuals had pre-existing autoantibodies (suggestive of autoimmune thyroiditis, primary biliary cirrhosis or connective tissue disease) (Extended Data Fig. 3a). Anticytokine antibodies were largely absent or were observed at a low mean fluorescence intensity with no significant changes (Extended Data Fig. 3b). Two individuals had anti-IL-21 autoantibodies, and two additional individuals had anti-IL-1 antibodies (Extended Data Fig. 4). Importantly, none of the individuals with pre-existing autoantibodies or anticytokine antibodies experienced adverse events, nor did levels of pre-existing autoantibodies or anticytokine antibodies change in response to vaccination.

## Innate immune responses

We first assessed whole-blood samples of 27 individuals using cytometry by time of flight (CyTOF). Unsupervised clustering identified 14 major cell types (Fig. 2a, Extended Data Fig. 5a, b), which we further subtyped manually (Extended Data Fig. 5c). The frequency of intermediate monocytes ($CD14^+CD16^+$ monocytes) increased significantly two days after the primary vaccination, and was substantially higher two days after the secondary vaccination (Fig. 2b, Extended Data Figs. 5d, 6). In addition, there were enhanced levels of phosphorylated (p)STAT3 and pSTAT1 in multiple cell types on day 1 after secondary vaccination, relative to day 1 after primary vaccination (Fig. 2c, d). These data suggested that BNT162b2 vaccination induced a heightened innate immune response after secondary immunization relative to primary immunization.

To further investigate this phenomenon, we measured plasma cytokines in 31 vaccinated individuals using Olink (https://www.olink.com/products/target/inflammation/). Of the 67 cytokines detected, the concentration of 2 cytokines (IFNγ and CXCL10) were increased significantly on days 1 and 2 after primary vaccination (Fig. 2e left). Similar to our observations on intermediate monocytes and pSTAT1 and pSTAT3, the concentrations of these cytokines were increased even further after the secondary immunization (Fig. 2e right). The concentration of plasma IFNγ was 11.3-fold higher at day 22 relative to day 1 (Fig. 2f). CXCL10 peaked on day 23 (Extended Data Fig. 7a). The anti-inflammatory cytokine IL-10 showed a similar trend with enhanced response following secondary immunization, although this did not reach statistical significance (Extended Data Fig. 7b). The concentration of IFNα2 (type I interferon) was not significantly increased on days 1 and 2 after vaccination, and the responses were just over the assay limit of the highly sensitive single-molecule array (SIMOA) (16 fg ml$^{-1}$) (Extended Data Fig. 7c, d). Furthermore, there was a correlation between plasma IFNγ levels and pSTAT1 and pSTAT3 expression levels across several cell types (Fig. 2g, h, Extended Data Fig. 7e). Collectively, these data demonstrate that vaccination with BNT162b2 stimulates modest innate immune responses after primary immunization, which increase notably after the secondary immunization.

## Transcriptional signatures of vaccination

Next, we performed bulk mRNA sequencing of whole blood from 31 participants. Six of 185 samples did not pass quality control and were removed from the analysis (Extended Data Fig. 8a, b). Secondary vaccination generated a greater transcriptional response (as has previously been seen in a recent study of adjuvanted hepatitis B vaccine[17]), with nearly a fourfold increase of differentially expressed genes at day 22 as compared to day 1 (Fig. 3a); this is consistent with the increased markers of innate immunity detected by both CyTOF and Olink (Fig. 2). Gene-set enrichment analysis (GSEA) revealed that both doses of BNT162b2 stimulated antiviral and interferon response modules[18] (Fig. 3b). However, the booster immunization led to a broader innate response. In addition to the induction of antiviral pathways, secondary vaccination led to increases in signatures of dendritic cell activation and the upregulation of Toll-like receptor signalling, monocyte and neutrophil modules on days 22 and 23 (Fig. 3b, Extended Data Fig. 8c, d). The results were consistent

regardless of the baseline time point (that is, day 0 or 21) that we used (Extended Data Fig. 8e, f, Supplementary Table 1).

As mortality to COVID-19 is highest among individuals who are elderly and older populations are known to mount suboptimal responses to many vaccines[19], we examined whether there were age-associated differences in response to mRNA vaccination. On day 22, younger participants tended to have greater changes in monocyte and inflammatory modules, whereas older individuals had increased expression of B and T cell modules (Extended Data Fig. 8g). Given that the plasma IFNγ concentration was significantly higher after secondary vaccination, we asked whether there was an association between IFNγ and the increased innate responses after the boost. Both interferon and inflammatory modules were significantly enriched by GSEA when using genes ranked by correlation with IFNγ on day 22 (Fig. 3c). The average fold changes of these modules also correlated with IFNγ (Extended Data Fig. 8h, i), which suggests that IFNγ may have a role in driving enhanced innate and antiviral responses after the boost.

## Single-cell transcriptional response

Bulk transcriptomics signatures could reflect changes in cell composition as well as alterations in transcriptional activity within cells. We performed cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq) of 45 peripheral blood mononuclear cell (PBMC) samples from 6 individuals (Extended Data Fig. 1a) to disentangle these two effects and to examine transcriptional changes at the single-cell level. We enriched dendritic cells and mixed them with total PBMCs at a ratio of 1:2 to represent minor subsets of dendritic cells sufficiently[4]. After quality control, we obtained 242,479 high-quality transcriptomes that segregated into 18 cell clusters (Fig. 4a, Extended Data Fig. 9a–c). Notably, cluster C8 (which expressed *CD14*, *VCAN*, *CD1C*, *FCGR1A* and *CD274* mRNA or protein) emerged on day 22, one day after secondary vaccination (Fig. 4b). These cells uniquely expressed interferon-stimulated genes, including *WARS* (also known as *WARS1*), *GBP1*, *IFI30* and *IFITM3* (Extended Data Fig. 9d), and constituted about 0.01% of the Lin⁻HLA-DR⁺ population on day 1 after primary vaccination but increased almost 100-fold to about 1% one day after secondary vaccination (Fig. 4c). Iterative removal of each cluster from a pseudobulk score showed that C8 contributed to the increased interferon as well as monocyte blood transcriptional modules (BTMs) that we observed in the bulk transcriptomics data on day 22 (Extended Data Fig. 9e). To examine whether these cells uniquely emerge in response to mRNA vaccination or whether they are seen in natural SARS-CoV-2 infection (because they expressed *CD14*, *CD1C* and *CD274*, which are known to be expressed by myeloid-derived suppressor cells observed at elevated frequencies in the blood of patients infected with SARS-CoV-2[4,20–22]), we combined and analysed innate immune cells (myeloid cells and plasmacytoid dendritic cells) from this study with data from two previous studies[4,23] after batch correction using Harmony[24]. Well-annotated cell types such as monocytes and dendritic cells overlapped between the datasets, but the C8 cluster did not overlap (Extended Data Fig. 10a, b). The cluster closest to C8 was IFN-experienced CD14⁺ monocytes, previously defined in patients with COVID-19[4] (Extended Data Fig. 10b). Although the partial overlap was due to expression of interferon-stimulated genes in both clusters, C8 expressed higher levels of HLA-DR and

other activation molecules and lower levels of S100 genes that are known to be expressed in myeloid-derived suppressor cells induced in patients with COVID-19 (Extended Data Fig. 10c). These data suggest that cluster C8 is not present in COVID-19 infection, and does not represent myeloid-derived suppressor cells.

To further delineate the cellular composition of C8, we re-embedded C8 with uniform manifold approximation and projection using participant-corrected principal component analyses. Using Louvain community detection, we resolved seven distinct clusters within the original C8 cluster (Fig. 4d). C8 was a heterogeneous mix of classical monocytes (C8_0, C8_1 and C8_3), a classical dendritic cell subtype (cDC2; C8_2) and intermediate monocytes (C8_4), as evidenced by proximity to the original clusters calculated by Euclidean distance (Fig. 4e, Extended Data Fig. 10d–f). We analysed subclusters C8_1 and C8_2, which are closer to CD14$^+$ monocytes and cDC2, respectively. In addition to expressing higher interferon-stimulated genes as compared to their parent clusters, they had a reduced expression of the AP-1 transcription factors *FOS* and *JUN* (Fig. 4f). This C8 population is similar to an epigenetically remodelled monocyte population in the blood of humans, 21 days after vaccination with two doses of an AS03-adjuvanted H5N1 pandemic influenza vaccine (H5N1 + AS03)[25]. These monocytes demonstrated an enhanced chromatin accessibility of interferon-stimulated genes and reduced accessibility of AP-1 transcription factors, and showed heightened resistance to infection with unrelated blood-borne viruses, such as dengue virus and Zika virus[25]. We asked whether C8 represents an analogous cell type at the transcriptional level. C8 had a relatively higher expression of *IRF1*, *STAT1*, *STAT2*, *STAT3* and *IRF8* and reduced levels of *FOS*, *JUNB*, *JUND* and *ATF3*—the same transcription factors that defined the monocyte population in the previous study[25] (Fig. 4g). We confirmed this using an extended set of genes for which the chromatin accessibility profile was higher 21 days after H5N1 + AS03 vaccination (Fig. 4h). Notably, the emergence of C8 correlated with plasma IFNγ levels (Extended Data Fig. 10g, h). In vitro stimulation of purified healthy monocytes with IFNγ or day-22 plasma also induced a C8 signature, which suggests that IFNγ has a key role in inducing cluster C8, in response to mRNA vaccination (Extended Data Fig. 10i, j).

In addition to the emergence of C8, our CITE-seq analysis demonstrated that the interferon signature was broadly induced across cell types on day 1 and day 22, and the higher magnitude of response on day 22 was more evident (Fig. 4i). Furthermore, there was activation of natural killer cells on day 22, as observed by downregulation of natural-killer-cell-associated gene modules, and upregulation of AP-1 transcription factors[26] on day 22 (Extended Data Fig. 10k, l).

## Comparison with other vaccines

As mRNA vaccines have only recently received approval for use in humans, the degree to which these vaccines induce similar or distinct immune responses compared to other vaccine types (such as inactivated or live-attenuated vaccines) is unknown. To address this, we performed a comparative analysis of a set of published vaccine trials with BNT162b2 (Extended Data Table 3) by generating similarity matrices through pairwise correlations of mean gene fold changes between vaccines at days 1 and 7 after vaccination. Although

the day-1 response to the first dose of BNT162b2 showed little overlap with that of other vaccines, the response at day 1 after the boost was broadly similar to the response induced by vaccination with adjuvanted vaccines (H5N1 + AS03), live viral vectors (Ebola and HIV vaccines), or inactivated influenza (which stimulates a recall response) (Extended Data Fig. 11a). At the BTM level, the shared signature consisted of innate immunity modules, including interferon signalling, dendritic cell activation and inflammatory responses (Fig. 5). Meanwhile, day-7 responses to both the prime and boost doses of BNT162b2 exhibited weak correlation both between themselves and with other vaccines (Extended Data Fig. 11b), with cell-cycle-related transcriptional modules after the prime dose being the signature shared with many vaccines (Extended Data Fig. 11c). With most vaccines, this cell cycle signature is also associated with upregulation of B and plasma cell modules, reflecting the expansion of antibody-secreting cells[18,27]. However, this induction of B and plasma cell modules was absent in BNT162b2 (Extended Data Fig. 11d, e). Given that BNT162b2 successfully promoted a robust antibody response (Fig. 1), the lack of detectable plasma cell or B cell signature on day 7 (particularly after the boost) was surprising. Consistent with this, we observed less than a twofold increase in plasmablast responses by CyTOF (Extended Data Fig. 11f), in contrast to seasonal influenza or other vaccines that induce a higher frequency of plasmablasts[27].

## Transcriptional correlates of adaptive immunity

As cellular and humoral immunity are the chief functional components that mediate protection from infection, we used GSEA to identify early transcriptional responses correlated with day-42 neutralizing antibody or day-28 $CD8^+IFN\gamma^+$ T cell responses. On day 22 (one day after the boost), monocyte-related modules correlated with neutralizing antibody responses, whereas interferon and antiviral signatures were associated with the CD8 T cell response (Extended Data Fig. 12a, b). The emergence of SARS-CoV-2 variants poses a serious challenge for the success of ongoing vaccination efforts. We asked whether there are early transcriptional correlates of the cross-neutralization potential induced by BNT162b2. We defined a cross-neutralization index (a ratio of variant-to-WA1 neutralizing antibody titres) and, using GSEA, found that monocyte and inflammatory modules were highly associated with this index (Extended Data Fig. 12c). In addition, the peak frequency of classical monocytes at two days after the boost correlated with cross-neutralization index (Extended Data Fig. 12d, e). Consistent with this, a gene score that defines C3 (the classical monocyte cluster in the CITE-seq data) also correlated with cross-neutralization (Extended Data Fig. 12f).

## Discussion

Our study represents, to our knowledge, the first systems-level analysis of innate and adaptive immunity to an mRNA vaccine. In contrast to the dysregulated innate immune responses to SARS-CoV-2 infection[4,20,28,29], BNT162b2 vaccination stimulated antiviral immunity with little type I IFN response after the first dose, but a notably enhanced innate response after the secondary immunization. Cluster C8, defined by single-cell transcriptional profiling, is a heterogenous mixture of myeloid cells that are uniquely induced by mRNA vaccination and are distinct from the IFN-experienced HLA-DR[low] myeloid cells

observed in natural infection[4,20]. A previous study has shown that vaccination with H5N1 + AS03 induces a similar transcriptional signature (enriched interferon-stimulated genes and diminished expression of AP-1 transcription factors) in monocytes and myeloid dendritic cells[25]. In that study, vaccination induced epigenetic reprogramming of these myeloid cells, leading to enhanced resistance against heterologous viruses such as dengue virus and Zika virus, even several weeks later[25]. In the case of BNT162b2, whether epigenetic reprogramming underlies the enhanced interferon-stimulated gene response in C8 after secondary immunization, and whether this confers enhanced resistance to viruses, remains an open question. Alternatively, it is conceivable that the enhanced myeloid cell response after secondary immunization reflects the response of these cells to a systemic cytokine response. Consistent with this, plasma IFN$\gamma$ concentration was significantly higher one day after secondary immunization and associated with the emergence of C8. Thus, these data provide a model in which 'cytokine feedback' regulates the enhanced innate immune responses to secondary vaccination (Extended Data Fig. 10h–j). Whether or not T cells provide this feedback warrants further investigation, but we did not detect IFN$\gamma$ within PBMCs (data not shown). Natural killer cells, tissue-resident T cells or innate lymphoid cells at the site of vaccination or draining lymphoid tissues could be potential sources of the rapidly induced circulating IFN$\gamma$. This model does not preclude complementary mechanisms, such as persistent epigenetic changes[25,30]. Finally, our analysis of transcriptional signatures of BNT162b2 vaccination relative to those induced by six other vaccines provides a useful benchmark to assess human immunity to mRNA vaccination in the broader context of immune responses to other vaccines.

## 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/s41586-021-03791-x.

## Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized, and investigators were not blinded to allocation during experiments and outcome assessment.

### Human subjects and experimentation

Fifty-six healthy volunteers were recruited for the study under informed consent. The study was approved by Stanford University Institutional Review Board (IRB 8269) and was conducted within full compliance of Good Clinical Practice as per the Code of Federal Regulations. The demographics of all participants are provided in Extended Data Table 1.

### Anti-spike binding enzyme-linked immunosorbent assay

SARS-CoV-2 spike protein was purchased from Sino Biologicals. Ninety-six-well high binding plates were coated with 100 ng of spike protein diluted at a concentration of 2 μg ml$^{-1}$ in PBS. The next morning, the plates were washed once, blocked with 3%

non-fat milk in PBS containing 0.1% Tween 20 (PBST) for 1 h at room temperature. Sera samples serially diluted in 1% non-fat milk containing PBST were added to the plates and incubated at 37 °C for 1 h. The plates were washed 3× with PBST, horseradish peroxidase conjugated goat anti-monkey IgG (γ-chain specific, Alpha Diagnostics, 1:4,000 dilution), in PBS-T containing 1% non-fat milk was added and incubated for 1 h at room temperature. Wells were washed 3× with PBST before addition of 3,3′,5,5′-tetramethylbenzidine (TMB) substrate solution. The reaction was stopped after 12 min by addition of 0.16 M sulfuric or 1 M hydrochloric acid. The optical density at 450 nanometres was measured with a Biorad microplate reader.

### Focus reduction neutralization titre assay

Neutralization assays with authentic SARS-CoV-2 virus (infectious clone SARS-CoV-2 derived from 2019-nCoV/USA_WA1/2020 strain) were performed as previously described[31]. Sera samples were serially diluted (threefold) in serum-free Dulbecco's modified Eagle's medium (DMEM) in duplicate wells and incubated with 100–200 focus-forming units infectious clone derived SARS-CoV-2-mNG virus[32] at 37 °C for 1 h. The antibody–virus mixture was added to Vero E6 cell (C1008, ATCC, no. CRL-1586) monolayers seeded in 96-well blackout plates and incubated at 37 °C for 1 h. After incubation, the inoculum was removed and replaced with pre-warmed complete DMEM containing 0.85% methylcellulose. Plates were incubated at 37 °C for 24 h. After 24 h, methylcellulose overlay was removed, cells were washed twice with PBS and fixed with 2% paraformaldehyde in PBS for 30 min at room temperature. Following fixation, plates were washed twice with PBS and foci were visualized on a fluorescence ELISPOT reader (CTL ImmunoSpot S6 Universal Analyzer) and enumerated using Viridot[33]. The neutralization titres were calculated as follows: 1 − (ratio of the mean number of foci in the presence of sera and foci at the highest dilution of respective sera sample). Each specimen was tested in two independent assays performed at different times. The focus reduction neutralization titre (FRNT)–mNeonGreen 50% ($mNG_{50}$) titres were interpolated using a four-parameter nonlinear regression in GraphPad Prism 8.4.3. Samples with an FRNT–$mNG_{50}$ value that was below the limit of detection were plotted at 10. For these samples, this value was used in fold reduction calculations.

### FRNT assay against the variants of concern

The wild-type infectious clone SARS-CoV-2, derived from the 2019-nCoV/USA_WA1/2020 strain, was propagated in Vero E6 cells (ATCC) and sequenced[32]. The B.1.351 variant was isolated as previously described[34]. Our laboratory plaque-isolated the virus on VeroE6 cells followed by a single round of propagation on Vero E6 cells (multiplicity of infection of 0.05), aliquoted to generate a working stock and sequenced. Viral titres were determined by focus-forming assay on Vero E6 cells. Viral stocks were stored at −80 °C until use.

FRNT assays were performed as described for the wild-type FRNT assay. The assay with each variant was performed simultaneously with wild-type controls. The samples were diluted at 3-fold in 8 serial dilutions using DMEM in duplicates with an initial dilution of 1:10 in a total volume of 60 μl. Serially diluted samples were incubated with an equal volume of SARS-CoV-2 (wild type or the variant) (100–200 foci per well) at 37 °C for 1

h in a round-bottomed 96-well culture plate. The antibody–virus mixture was then added to Vero cells and incubated at 37 °C for 1 h. After incubation, the antibody–virus mixture was removed and 100 μl of prewarmed 0.85% overlay was added to each well. Plates were incubated at 37 °C for 24 h. After 24 h, methylcellulose overlay was removed, and cells were washed 3 times with PBS. Cells were then fixed with 2% paraformaldehyde in PBS (Electron Microscopy Sciences) for 30 min. Following fixation, plates were washed twice with PBS and 100 μl of permeabilization buffer (0.1% bovine serum albumin (BSA), saponin in PBS), was added to the fixed Vero cells for 20 min. Cells were incubated with an anti-SARS-CoV spike primary antibody directly conjugated to biotin (CR3022–biotin) for 1 h at room temperature. Next, the cells were washed 3 times in PBS and avidin–HRP was added for 1 h at room temperature followed by 3 washes in PBS. Foci were visualized using TrueBlue HRP substrate (KPL, no. 5510–0050) and imaged on an ELISPOT reader (CTL).

### Intracellular cytokine staining assay

Antigen-specific T cell responses were measured using the intracellular cytokine staining assay as previously described[35]. Live frozen PBMCs were revived, counted and resuspended at a density of 2 million live cells per ml in complete RPMI (RPMI supplemented with 10% FBS and antibiotics). The cells were rested overnight at 37 °C in a $CO_2$ incubator. The next morning, the cells were counted again, resuspended at a density of 15 million cells per ml in complete RPMI and 100 μl of cell suspension containing 1.5 million cells was added to each well of a 96-well round-bottomed tissue culture plate. Each sample was treated with two conditions, no stimulation and a peptide pool spanning the spike protein at a concentration of 1 μg ml$^{-1}$ of each peptide in the presence of 1 μg ml$^{-1}$ of anti-CD28 (clone CD28.2, BD Biosciences) and anti-CD49d (clone 9F10, BD Biosciences) as well as anti-CXCR3 and anti-CXCR5. The peptides were custom-synthesized to 90% purity using GenScript, a commercial vendor. All samples contained 0.5% v/v DMSO in total volume of 200 μl per well. The samples were incubated at 37 °C in $CO_2$ incubators for 2 h before addition of 10 μg ml$^{-1}$ brefeldin-A. The cells were incubated for an additional 4 h. The cells were washed with PBS and stained with Zombie UV fixable viability dye (Biolegend). The cells were washed with PBS containing 5% FCS, before the addition of surface antibody cocktail. The cells were stained for 20 min at 4 °C in 100-μl volume. Subsequently, the cells were washed, fixed and permeabilized with cytofix/cytoperm buffer (BD Biosciences) for 20 min. The permeabilized cells were stained with intracellular cytokine staining antibodies for 20 min at room temperature in 1× perm/wash buffer (BD Biosciences). Cells were then washed twice with perm/wash buffer and once with staining buffer before acquisition using the BD Symphony Flow Cytometer and the associated BD FACS Diva software. All flow cytometry data were analysed using Flowjo software v.10 (TreeStar).

### Bead-based antigen arrays

We used an existing bead-based autoantigen array, and a cytokine array with expanded content that was based on recent COVID-19 studies[16]. A complete list of all antigens, vendors and catalogue numbers can be found in Supplementary Tables 2, 3. The 'COVID-19 Autoantigen Array' included 55 commercial protein antigens associated with connective tissue diseases (Supplementary Table 2). The 'COVID-19 Cytokine Array' comprised 58 proteins including cytokines, chemokines, growth factors, acute phase proteins and cell

surface proteins (Supplementary Table 3). Antigens were coupled to carboxylated magnetic beads (MagPlex-C, Luminex) such that each antigen was linked to beads with unique barcodes, as previously described[16,36,37]. Prototype human plasma samples derived from participants with autoimmune diseases with known reactivity patterns were purchased from ImmunoVision or were obtained from Stanford rheumatology clinics and had previously been characterized[16]. Serum samples from individuals with autoimmune polyendocrine syndrome type 1 (APS1), pulmonary alveolar proteinosis (PAP) or atypical mycobacterial infection (AMI) were used for validation of anticytokine antibodies[16]. Serum samples were tested at 1:100 dilution in 0.05% PBS-Tween supplemented with 1% (w/v) BSA. Bound antibody was detected using R-phycoerythrin (R-PE) conjugated Fcγ-specific goat anti-human IgG F(ab′)2 fragment (Jackson ImmunoResearch) before analysis using a FlexMap3D instrument (Luminex). A minimum of 100 events per bead identifier were counted, and samples were studied in duplicate. Binding events were displayed as mean fluorescence intensity (MFI). All data analysis and statistics were performed using R and various R packages[38]. For normalization, average MFI values for 'bare bead' identifier were subtracted from average MFI values for antigen-conjugated bead identifiers.

### CyTOF analysis of whole-blood samples

Fresh whole-blood samples collected in sodium citrate cell preparation tubes (CPT) were fixed in proteomic stabilizer buffer. Two hundred and seventy µl of whole-blood samples were mixed with 420 µl of smart buffer, mixed and incubated at room temperature for 12 min and frozen at −80 °C until processing. Fixed frozen cells were thawed by gentle resuspension in CSM (PBS supplemented with 2% BSA, 2 mM EDTA and 0.1% sodium azide), washed twice with CSM and counted. Cells were permeabilized and barcoded using Cell-IDTM 20-Plex Pd Barcoding Kit (Fluidigm). The samples were washed with CSM, pooled and counted. One pooled sample containing a mix of all barcoded PBMC samples was stained for 30 min with surface antibody cocktail at room temperature. The sample was then fixed with 4% freshly prepared paraformaldehyde (Alfa Aesar) for 10 min at room temperature, washed with CSM, permeabilized with 100% methanol (Sigma) and kept at −80 °C overnight. The next day, the cells were washed with CSM, counted and stained with pre-titrated intracellular antibody cocktail for 30 min at room temperature. Cells were then washed with CSM, stained with iridium-containing DNA intercalator (Fluidigm), washed with MilliQ water and acquired on Helios mass cytometer (Fluidigm) in MilliQ water supplemented with 1× EQ four element calibration beads (Fluidigm).

The FCS files were bead-normalized before data export. The data were processed for debarcoding in Flowjo software v.10 (TreeStar). In brief, the bead-normalized file was used to gate single cells on the basis of DNA content and event length using FlowJo. The single cells were reimported and debarcoded using Helios software version 7.0.5189. The debarcoded samples were analysed using FlowJo or R version 1.2.1335 for analysis and visualization.

### CyTOF data analysis

High-dimensional analysis of phospho-CyTOF data was performed using a previously described R-based pipeline[39]. In brief, the raw .fcs files were imported into R and the data

were transformed to normalize marker intensities using arcsinh with a cofactor of 5. For visualization, another transformation was applied that scales the expression of all values between 0 and 1 using percentiles as the boundary. Cell clustering was performed with 4,000 cells randomly selected from each sample using FlowSom and ConsensusClusterPlus. The transformed matrix was used as an input for FlowSom and cells were separated into 20 clusters. To obtain reproducible results (avoid random start), a seed was set for each clustering. The 20 clusters were manually annotated on the basis of the lineage marker expression, and were merged to produce the final clusters. The clusters were visualized in two-dimensional space using UMAP. The abundance of cell populations was determined using Plotabundance function. In parallel, the data were manually gated to identify 25 immune cell subpopulations that were not well-distinguished in UMAP and used for all quantification purposes.

### Plasma protein profiling using multiplex Olink panel

We measured cytokines in plasma using Olink multiplex proximity extension assay (PEA) inflammation panel (Olink proteomics: www.olink.com) according to the manufacturer's instructions. The PEA is a dual-recognition immunoassay, in which two matched antibodies labelled with unique DNA oligonucleotides simultaneously bind to a target protein in solution. This brings the two antibodies into proximity, allowing their DNA oligonucleotides to hybridize, serving as template for a DNA polymerase-dependent extension step. This creates a double-stranded DNA 'barcode' that is unique for the specific antigen and quantitatively proportional to the initial concentration of target protein. The hybridization and extension are immediately followed by PCR amplification and the amplicon is then finally quantified by microfluidic qPCR using Fluidigm BioMark HD system (Fluidigm).

### Bulk transcriptomics

RNA was isolated from blood samples stored in Paxgene tubes at the Yerkes Genomics Core (http://www.yerkes.emory.edu/nhp_genomics_core/). RNA quality was assessed using an Agilent 4200 TapeStation and concentration via the RNA HS assay on the Qubit. Globin transcripts in blood RNA were blocked with the FastSelect Globin Reagent (Qiagen) before library preparation. Libraries were prepared using the Clontech SMART-Seq v.4 Ultra Low Input RNA kit (Takara Bio) in combination with the NexteraXT DNA Library Preparation kit to append dual-indexed adaptor sequences (Illumina). Libraries were validated by capillary electrophoresis on an Agilent 4200 TapeStation, pooled at equimolar concentrations, and sequenced on an Illumina NovaSeq6000 at 100SR, yielding 25–30 million reads per sample.

### Bulk transcriptomics analysis

Gene-level counts were filtered to remove those with a median expression less than 32. Principal component analysis (PCA) was performed on baseline samples to identify outliers. Three samples were more than 1.5 s.d. away from the mean and were removed from the analysis. Relative log expression (RLE) plots were generated with EDAseq[40]; samples with an RLE > 0.6 were removed from the analysis. Differential gene analysis was performed using DESeq2[41] (v.1.26.0), incorporating participant identifier into the model to account for inter-participant bias. Genes were ranked by the Wald statistic as reported by DESeq2

for GSEA using the BTMs[18]. Per-participant fold changes were computed by dividing the DESeq2 normalized expression data for the day of interest by either day 0 (for day 1, day2 and day 7) or day 21 (for day 22, 28 and 42). To obtain BTM correlates with age, the age of each participant was compared against the per-participant fold changes for day 22. The resulting correlation values were ranked by $t$-statistic and analysed with GSEA. The same method was employed to obtain BTM correlates with IFN$\gamma$. IFN scores were computed by taking the per-participant mean fold change on day 22 of the unique set of genes present in the 5 interferon BTMs (M75, M111.1, M150, M127 and M68) that significantly correlated with day-22 IFN$\gamma$ fold change. Similarly, the per-participant M16 gene score was computed using average fold change on day 22 of the genes present in M16.

### Vaccine dataset meta-analysis

Datasets were obtained from Gene Expression Omnibus (GEO) via the accession identifiers in Supplementary Table 4. CEL files of all the samples belonging to the same trial were grouped and normalized in Bioconductor by RMA[42], which includes global background adjustment and quantile normalization. Probes mapping to multiple genes were discarded, and the remaining probes were collapsed to gene level in each dataset by selecting the probe for each gene with the highest mean expression across all subjects. The only non-microarray dataset was GSE97590, for which the normalized count matrix from GEO was used. Genes not present in all datasets were removed. Baseline normalized $log_2$-transformed fold changes were then computed per subject for all genes. GSEA was then performed to identify enriched BTMs using gene lists for each dataset ranked by $t$-statistic from two-sided Student's $t$-tests on the post-vaccination $log_2$-transformed fold changes.

### CITE-seq

CITE-seq analysis of PBMCs were assayed exactly as previously described[4]. In brief, live frozen PBMCs were thawed and $2\times$ washed with RPMI supplemented with 10% FBS and 20 μg ml$^{-1}$ DNase I (Sigma Aldrich). Dendritic cells were enriched using the Dynabeads DC Enrichment Kit (Invitrogen, 11308D) according to manufacturer's instructions with 3–4 million PBMCs as starting material. The enriched cells were mixed with total PBMCs at a ratio of 1:2 and mixed cells were stained with a cocktail of TotalSeq-A antibodies in PBS supplemented with 5% FBS, 2 mM EDTA and 5 mg ml$^{-1}$ human IgG, washed twice with PBS supplemented with 5% FBS, and 2 mM EDTA, and resuspended in PBS supplemented with 1% BSA (Miltenyi), and 0.5 U μl$^{-1}$ RNase Inhibitor (Sigma Aldrich). About 9,000 cells were targeted for each experiment.

Cells were mixed with the reverse transcription mix and subjected to partitioning along with the Chromium gel-beads using the 10X Chromium system to generate the gel-bead in emulsions (GEMs) using the 3′ V3 chemistry (10X Genomics). The reverse transcription reaction was conducted in the C1000 touch PCR instrument (BioRad). Barcoded cDNA was extracted from the GEMs by post-GEM reverse transcription cleanup and amplified for 12 cycles. Before amplification, the cDNA amplification mix was spiked in with ADT additive primer (0.2 μM stock) to amplify the antibody barcodes. Amplified cDNA was subjected to $0.6\times$ SPRI beads cleanup (Beckman, B23318). Amplified antibody barcodes were recovered from the supernatant and were processed to generate TotalSeq-A libraries as

instructed by the manufacturer (BioLegend, TotalSeq-A antibodies with 10x Single Cell 3′ Reagent Kit v.3 3.1 protocol). The rest of the amplified cDNA was subjected to enzymatic fragmentation, end repair, A tailing, adaptor ligation and 10X-specific sample indexing as per manufacturer's protocol. Libraries were quantified using Bioanalyzer (Agilent) analysis.

10x Genomics scRNA-seq and TotalSeq-A libraries were pooled and sequenced on an Illumina HiSeq 4000 using the recommended sequencing read lengths of 28 bp (read 1), 8 bp (i7IndexRead) and 91 bp (read 2). CellRanger v.3.1.0 (10xGenomics) was used to demultiplex raw sequencing data and quantify transcript levels against the 10x Genomics GRCh38 reference v.3.0.0.

### CITE-seq analysis

10x Genomics scRNA-seq and TotalSeq-A libraries were pooled and sequenced on a Novaseq S4. Cell Ranger v.3.1.0 (10x Genomics) was used to quantify transcript levels against the 10x Genomics GRCh38 reference (v.3.0.0.) Raw count data were filtered to remove cells with a mitochondrial RNA fraction greater than 20% of total RNA counts per cell, cells with fewer than 100 unique features and cells with fewer than 200 total reads. The filtered count matrix was used to create a Seurat[43] (v.3.1.4) object. Filtered read counts were scaled by a factor of 10,000 and log-transformed. The antibody-derived tag matrix was normalized per feature using centre log normalization. Doublets were identified with scds[44] (v.1.2.0); cells with a doublet score in the top decile were removed. The remaining 242,479 cells were processed with the default Seurat pipeline. Specifically, the most variable 2,000 RNA features were used to perform PCA on the log-transformed counts. The first 25 principal components were used further downstream analyses, including clustering and UMAP projections. Clusters were identified with Seurat SNN graph construction followed by Louvain community detection on the resultant graph with a resolution of 0.2, yielding 18 clusters. Differential expression across time points was calculated with MAST[45] (v.1.12.0) to account for inter-participant heterogeneity.

Pseudobulk profiles were constructed by taking the average expression across all cells in each participant, per day. When computing fold changes across time points, the pseudobulk profile of each participant was compared to their baseline profile to reduce participant-specific biases. To calculate the effect of removing a cluster, each cluster across all time points was iteratively removed and resulting fold changes were recomputed.

C8 was re-embedded and reclustered with UMAP and Louvain community detection, respectively. Distances from each subcluster to the other clusters was calculated as the Euclidean distance between the average expression of all genes of each cluster. The Euclidean distances were calculated in the original data space. Specifically, the Euclidean distance was calculated using all genes as input to the dist function in R. The dist function calculates Euclidean distance $d(x, y)$ as:

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2}$$

in which $x$ is the value for gene $i$ in cluster A and $y$ is the value for gene $i$ in cluster B.

Complexheatmap (v.2.2.0) was used for all heat maps. All analysis was performed in R (v.3.6.3).

### Combined analysis of single-cell RNA sequencing

Data from ref. [23] were downloaded from https://covid19cellatlas.org/ as an .h5ad file and converted to a Seurat object in R. Both the resulting Seurat object and the vaccine data were subset to include only myeloid cells and combined using Harmony. Similarly, the data from ref. [4] were integrated with the myeloid cells from the vaccine study using Harmony[24]. Lymphoid cells from ref. [4] were removed after integration. For both integrations, UMAP was performed on the Harmony-corrected embeddings.

### Monocyte purification and stimulation

Monocytes were negatively enriched from healthy PBMCs using Dynabeads Untouched Human Monocyte kit (Invitrogen, cat. no. 11350D) following manufacturer's instructions. In brief, 50 million live PBMCs were stained with the antibody cocktail for 20 min at 4 °C. The cells were washed and mixed with 0.5 ml of premixed Dynabeads. The samples were incubated in a hulamixer for 15 min at 4 °C. The tubes were placed on the magnet and the unbound fraction containing purified CD14$^+$ monocytes was aspirated using a pipette. The purified monocytes were washed thoroughly and resuspended at a density of 5 million per ml for stimulation. The purity of monocytes was estimated by flow cytometry and was over 95% in all the samples.

Monocytes (0.5 million) were stimulated per condition in 96-well round-bottomed plate for 24 h in 100 μl complete RPMI. Different concentrations of IFNγ, as shown in Extended Data Fig. 7j, were added in 100 μl complete RPMI. Day 0, 1 or 22 plasma samples were from the participant 2055 with <10 pg ml$^{-1}$, <10 pg ml$^{-1}$ and 300 pg ml$^{-1}$ IFNγ, respectively, as measured by enzyme-linked immunosorbent assay. Fifty μl of plasma samples was added to appropriate wells. Fifty μl of complete medium was added to make up the volume to 0.2 ml in total. The plates were incubated at 37 °C, 5% CO$_2$ cell culture incubators.

### RNA isolation and qPCR

RNA was isolated using Aurum Total RNA minikit (Biorad, cat. no. 7326820) following the manufacturer's protocol. cDNA was synthesized using iScript Advanced cDNA synthesis kit (Biorad, cat. no. 1725038) using 150 ng total RNA in 20 μl volume. The cDNA samples were diluted 5-fold by adding 80 μl sterile nuclease-free water and 5 μl of cDNA was used for PCR reaction. The PCRs were carried out using Biorad Prime PCR reagents and SYBR green chemistry (SsoAdvanced Universal SYBR Green Supermix (cat. no. 1725272)) in Biorad CFX384 real-time PCR.

### Reporting summary

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

**Data availability**

CITE-seq and bulk RNA data are publicly accessible in the GEO under accession numbers GSE171964 and GSE169159, respectively. Any other relevant data are available from the corresponding authors upon reasonable request. Source data are provided with this paper.

**Code availability**

The codes used in the study are available in GitHub (https://github.com/scottmk777/PfizerCovid). Some codes used for meta-analysis can be obtained from the corresponding authors upon reasonable request.

## Extended Data
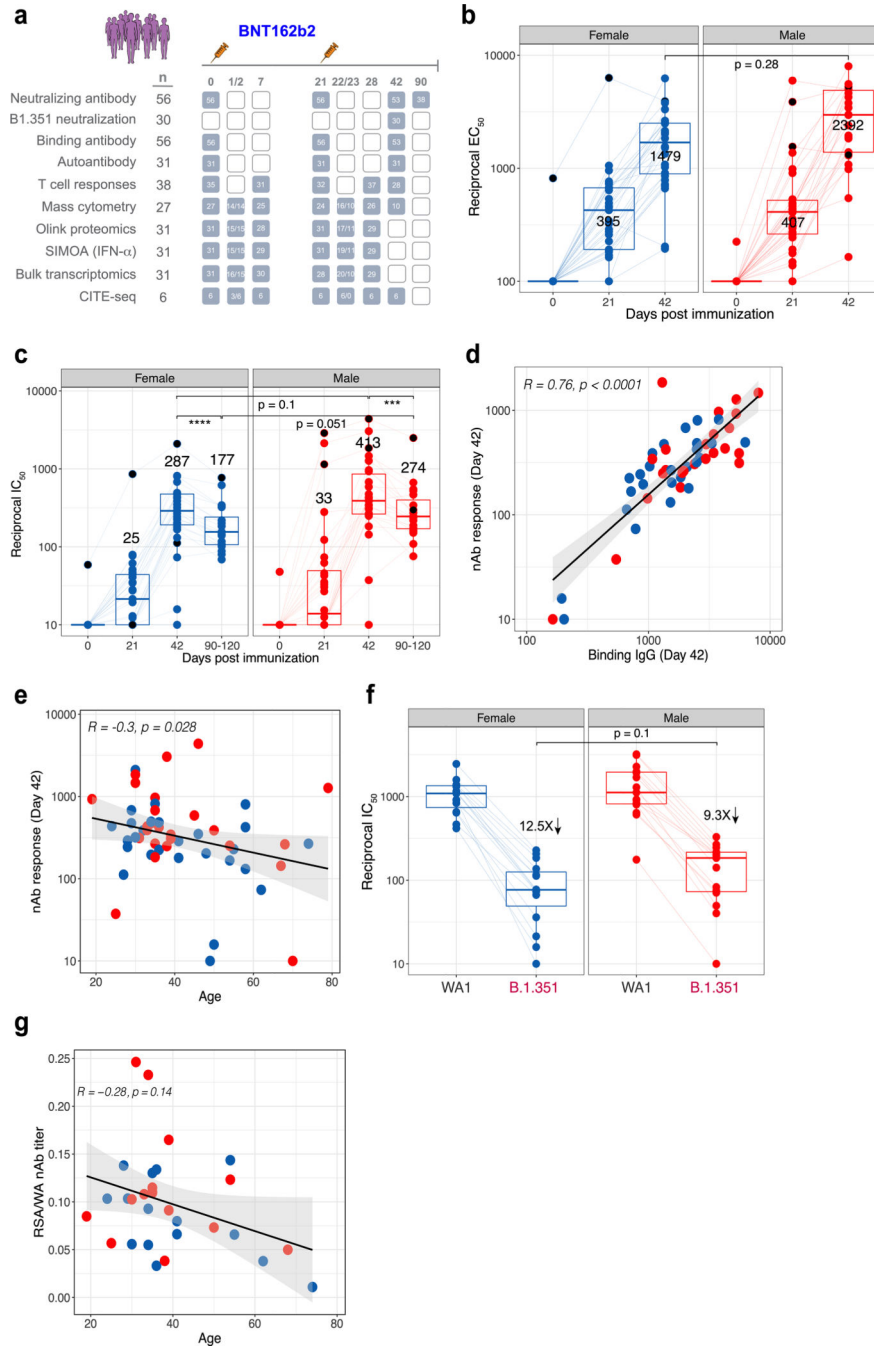


**Extended Data Fig. 1 |. Antibody responses to BNT162b2 vaccination.**
**a**, Schematic of the study design and the number of participants used in various assays. **b**,
**c**, Binding (**b**) and neutralizing (**c**) antibody responses to BNT162b2 vaccination in female
and male participants (*n* = 56). **d**, **e**, Correlation between binding antibody and neutralizing
antibody titres (**d**) and neutralizing antibody responses and age (**e**). **f**, Neutralizing antibody
response to B.1.351 variant of concern in female and male participants (*n* = 30). **g**,

Correlation between age and cross-neutralization index, defined as the ratio of neutralizing antibody response to B.1.351 versus WA1 strains. Each dot represents an individual in all plots. Blue and red colour denote female and male participants, respectively. Boxes show median and 25th–75th percentiles, and whiskers show the range in all box plots. All correlations are two-sided Spearman's correlations. The error bands represent 95% confidence limits. The statistical difference between time points is calculated using two-sided Wilcoxon matched-pairs signed-rank test and the statistical differences between groups were calculated using two-sided Mann–Whitney rank-sum test. ***$P < 0.001$, **** $P < 0.0001$.
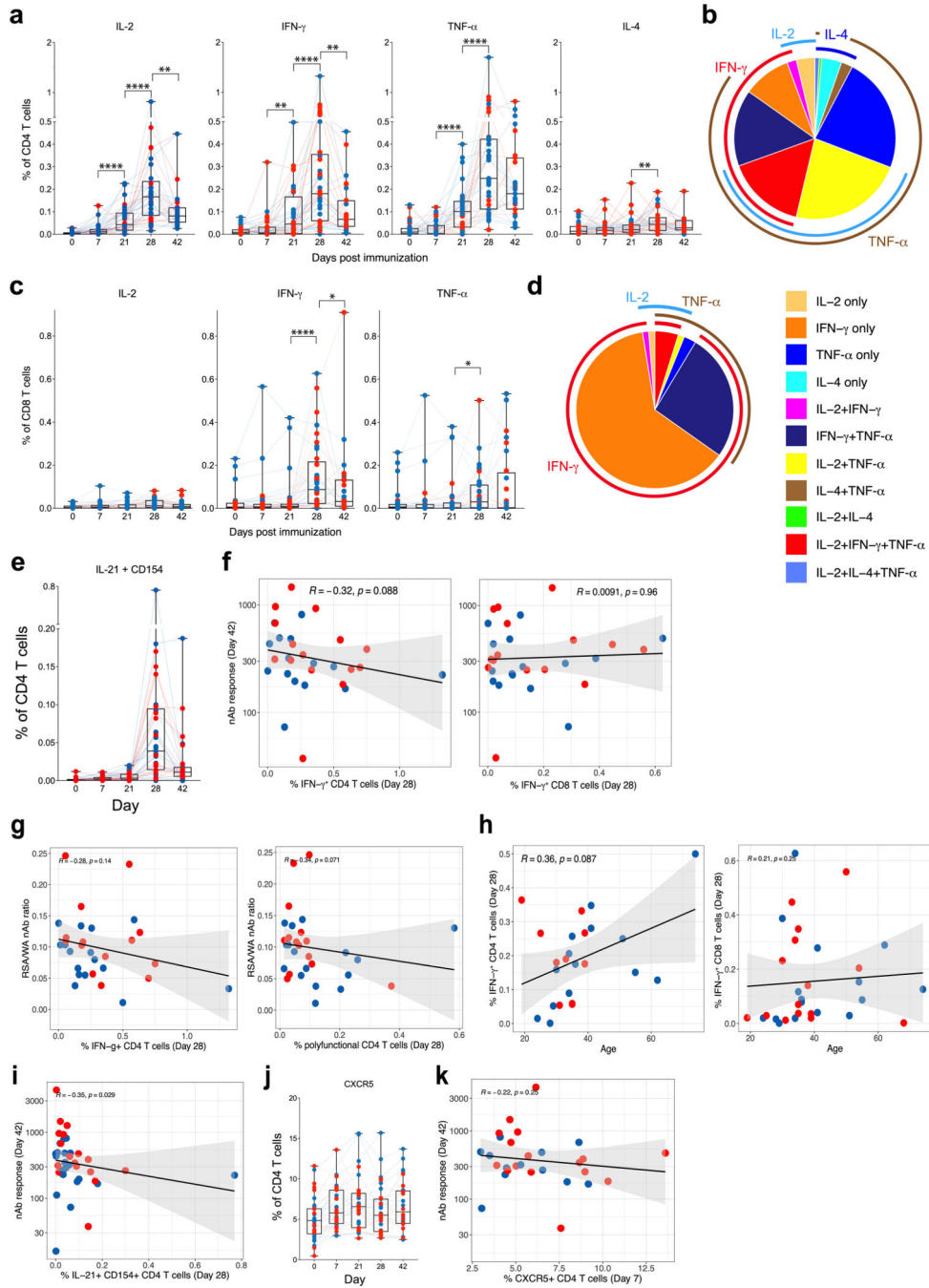
**Extended Data Fig. 2 |. T cell responses to BNT162b2 vaccination.**
**a**, Frequency of spike-specific CD4 T cell responses measured in blood at time points indicated on the *x* axis. **b**, Polyfunctional profiles of CD4 T cells. **c**, Frequency of spike-specific CD8 T cell responses measured in blood at time points indicated on the *x* axis. **d**, Polyfunctional profiles of CD8 T cells. **e**, Frequency of spike-specific CD4 T cells secreting IL-21 and CD154 at time points indicated on *x*-axis. **f**, Correlation between spike-specific CD4 (left) and CD8 (right) T cell frequencies and neutralizing antibody responses. **g**, Correlation between cross-neutralization index, ratio between neutralizing antibody

responses against B.1.351 to WA1 strains, and spike-specific CD4 T cell frequencies, IFN$\gamma^+$ (left) or polyfunctional CD4 T cells expressing IL-2, IFN$\gamma$ and TNF (right). **h**, Correlation of spike-specific CD4 (left) and CD8 (right) T cell responses with age. **i**, Correlation of spike-specific IL-21$^+$CD154$^+$ T follicular helper-like cells on day 28 and neutralizing antibody response on day 42. **j**, Frequency of CXCR5$^+$ CD4 T cells in PBMCs in DMSO-stimulation condition. **k**, Correlation of peak (day 7) CXCR5$^+$ CD4 T cells and neutralizing antibody response on day 42. Each dot represents an individual in all plots. Blue and red colour denote female and male participants, respectively. Boxes show median and 25th–75th percentiles, and whiskers show the range in all box plots. The IFN$\gamma$ response plots in CD4 (**a**) and CD8 (**c**) T cells are from Fig. 1d, e repeated here for completeness. The pie charts in **b**, **d** represent the proportion of T cells expressing one, two or three cytokines as shown in the legend. All correlations are two-sided Spearman's correlations. The error bands represent 95% confidence limits. The statistical difference between time points is calculated using two-sided Wilcoxon matched-pairs signed-rank test. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $**** P < 0.0001$. $n = 38$. Number of samples differ between time points as shown in Extended Data Fig. 1a.
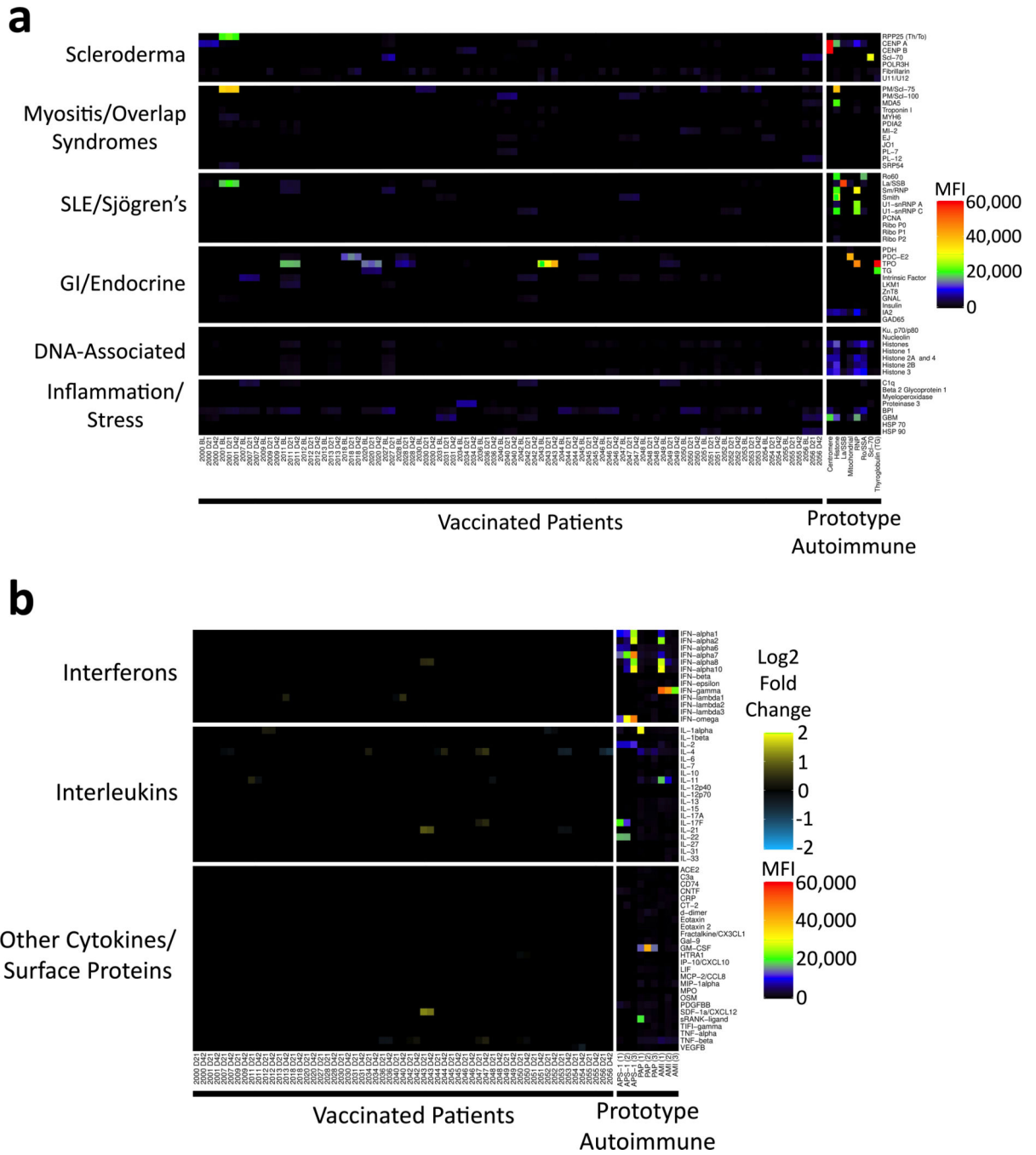
**Extended Data Fig. 3 |. Autoantibodies and anticytokine antibodies in vaccinated individuals.**
**a**, Heat map depicting serum IgG antibodies discovered using a 55-plex bead-based protein array containing the indicated autoantigens (*y* axis). Autoantigens are grouped on the basis of disease (for example, scleroderma, myositis and overlap syndromes such as mixed connective tissue disease, systemic lupus erythematosus (SLE) and Sjögren's, and gastrointestinal and endocrine disorders), DNA-associated antigens and antigens associated with tissue inflammation or stress responses. Vaccinated individuals are shown on the left (*n* = 30 individuals, on day 0, day 21, and day 42 and *n* = 1 on day 0 and day 21), and

eight prototype autoimmune disorders are shown on the right. Colours correspond to the MFI values shown at far right. **b**, Heat map using a 58-plex array of cytokines, chemokines, growth factors and receptors. Cytokines are grouped on the *y* axis by category (interferons, interleukins and other cytokines, growth factors and receptors), and serum samples are shown on the *x* axis. Vaccinated individuals are shown on the left ($n = 30$ individuals, on day 0, day 21 and day 42 and $n = 1$ on day 0 and day 21). Data are displayed as fold change over baseline. Prototype samples from patients with immunodeficiency disorders are shown on the right, and include three patients with AMI, three patients with PAP and three patients with APS1. Colours for the prototype samples correspond to the MFI values shown at far right.
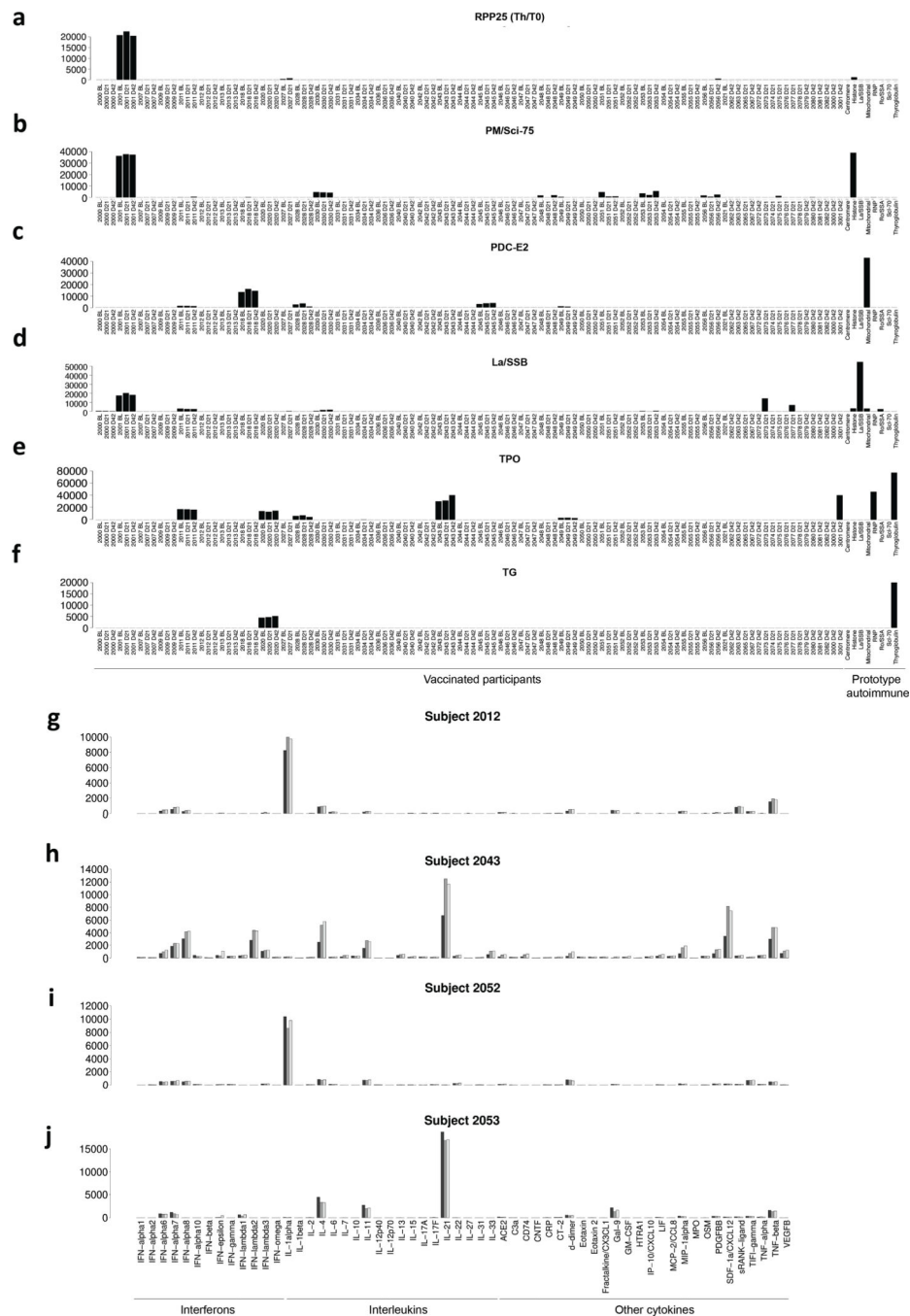
**Extended Data Fig. 4 |. Pre-existing autoantibodies and autocytokine antibodies do not change in vaccinated individuals.**

**a**–**f**, Bar plots (mean of two technical replicates) of representative patients with high baseline MFI autoantibodies for select antigens. **a**, Anti-RPP25 (Th/To). **b**, Anti-PM/Scl-75. **c**, Anti-SSB/La. **d**, Anti-PDC-E2. **e**, Anti-thyroperoxidase, TPO. **f**, Anti-thyroglobulin, TG. **g**–**j**, Bar plots (mean of two technical replicates) of representative patients with high baseline MFI autocytokine antibodies. Group bars represent antigens for baseline (black), day 21 (grey), and day 42 (white) time points. **g**, Individual 2012, anticytokine antibody measurements.

**h**, Individual 2043, anticytokine antibody measurements. **i**, Individual 2052, anticytokine antibody measurements; **j**, Individual 2053, anticytokine antibody measurements.
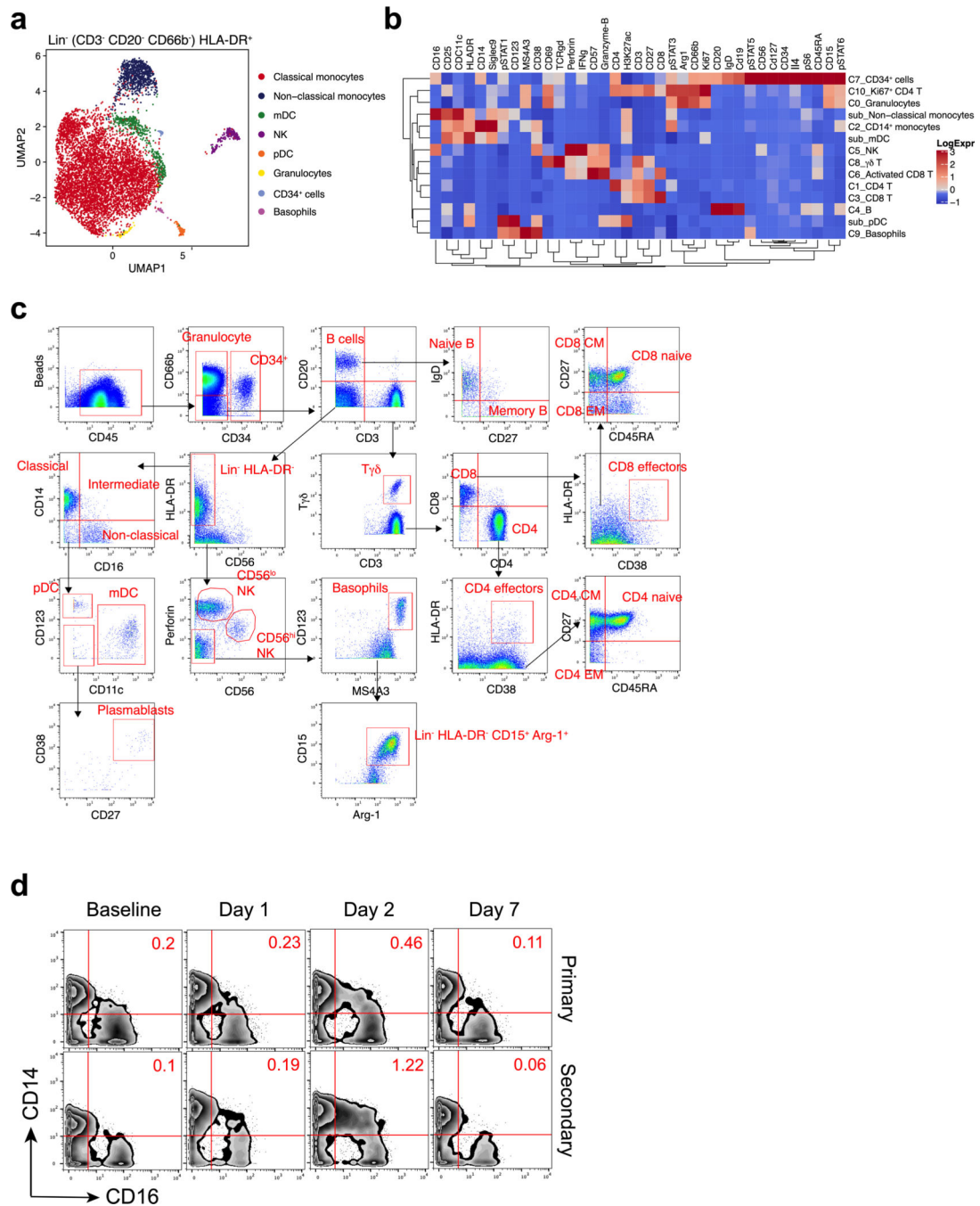


**Extended Data Fig. 5 |. Innate immune responses to BNT162b2 vaccination.**
**a**, Representation of CyTOF-identified cell clusters in Lin⁻ (CD3⁻CD20⁻CD66b⁻) HLA-DR⁺ cells visualized by UMAP in two-dimensional space. **b**, Heat map showing expression of markers in the different cell clusters identified in the CyTOF dataset. **c**, Gating strategy

identifying immune cell populations by the CyTOF panel. **d**, Mass cytometry plots showing CD14 and CD16 expression of Lin$^-$HLA-DR$^+$ population. $n = 27$.
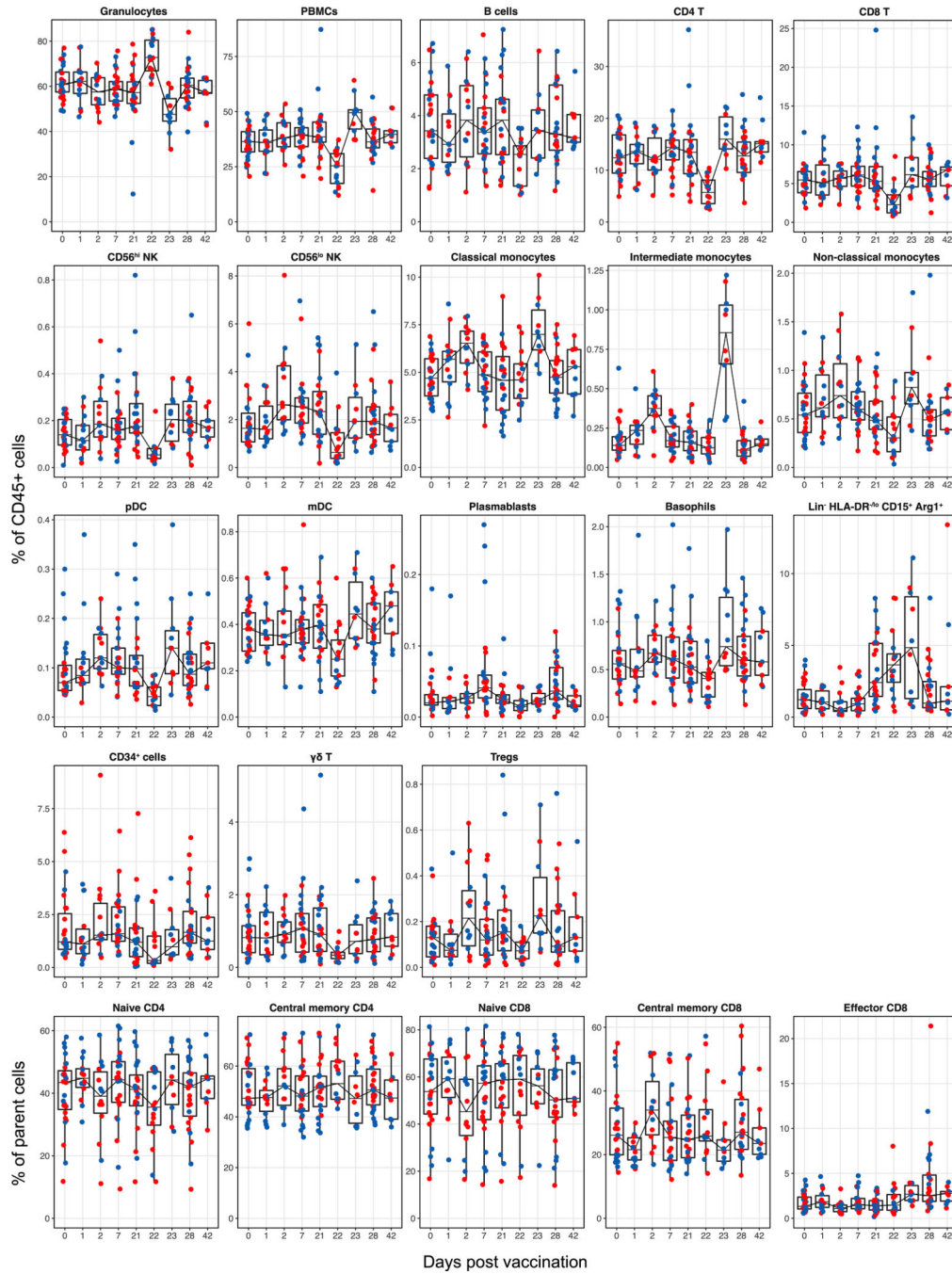


**Extended Data Fig. 6 |. Frequency of all cell types identified by CyTOF.**

Frequency of all major cell types measured in whole blood by CyTOF. Each dot represents an individual in all plots. Blue and red colour denote female and male participants, respectively. Boxes show median and 25th–75th percentiles, and whiskers show the range in

all box plots. $n = 27$; some time points have a smaller $n$ value (as shown in Extended Data Fig. 1a).
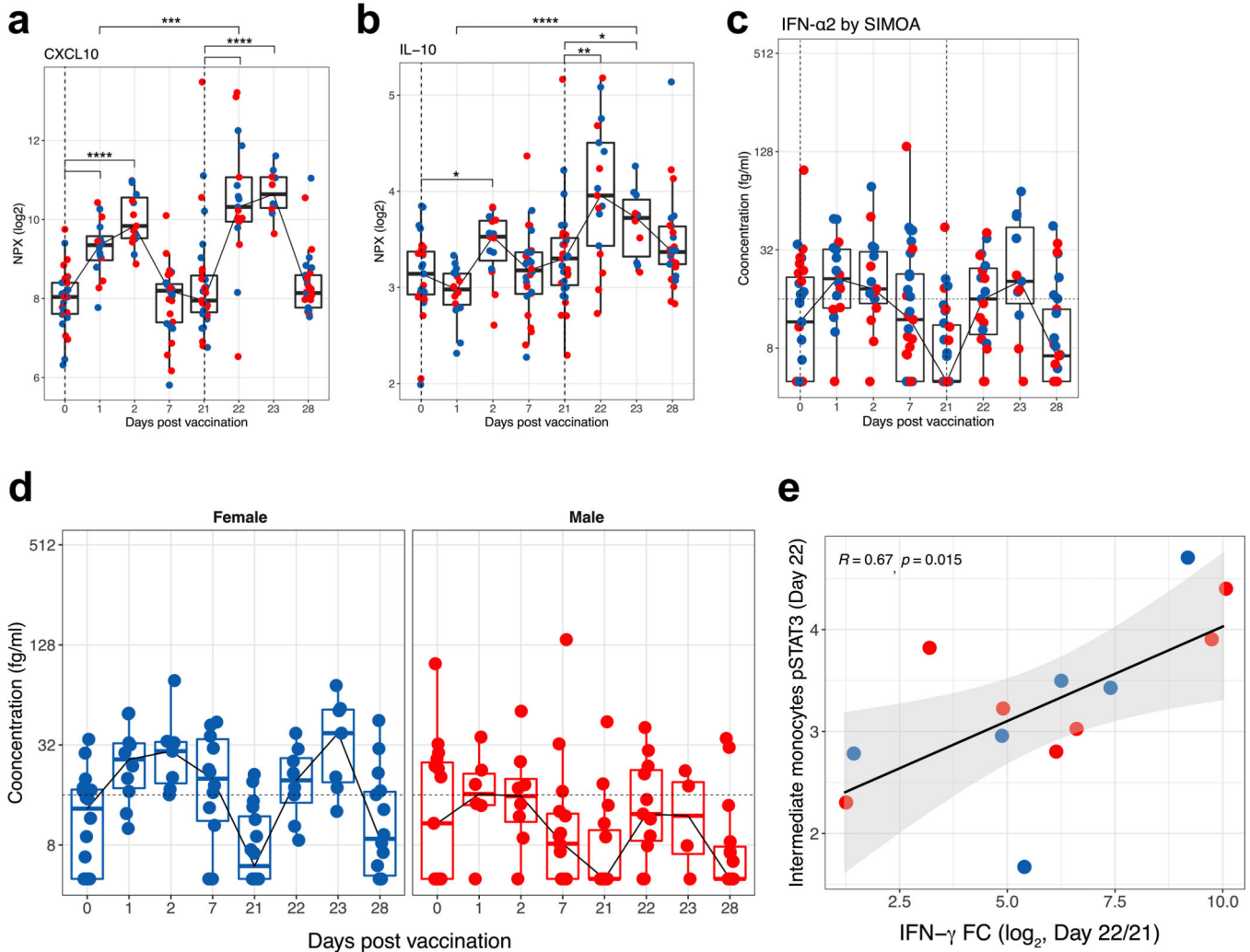


**Extended Data Fig. 7 |. Plasma cytokine analysis.**
**a–c,** Plasma levels of CXCL10 (**a**) and IL-10 (**b**) determined by Olink, and IFNα2 measured by SIMOA (**c**). **d,** Kinetic plots of IFNα2 in plasma of female or male participants measured by SIMOA. **e,** Scatter plots showing two-sided Spearman's correlation between pSTAT3 level in intermediate monocytes and plasma IFNγ levels. The error bands represent 95% confidence limits. Each dot represents an individual in all plots. Blue and red colour denote female and male participants, respectively. Boxes show median and 25th–75th percentiles, and whiskers show the range in all box plots. The statistically significant differences between the peak and baseline time points were measured using two-sided Wilcoxon matched-pairs signed-rank test. The differences between peak time points were measured using two-sided Mann–Whitney rank-sum test. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ and ****$P < 0.0001$. $n = 31$.
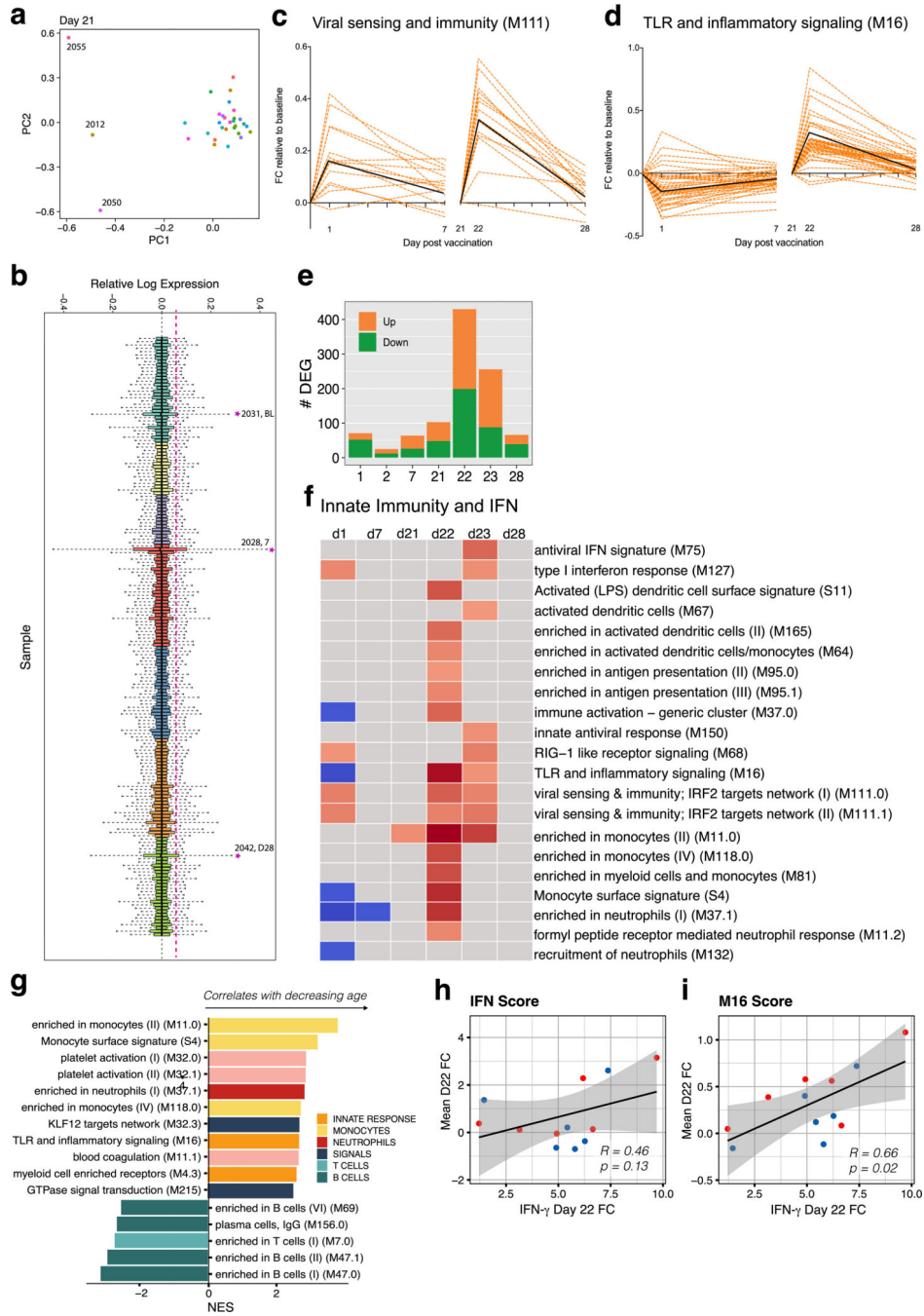
**Extended Data Fig. 8 |. Transcriptional signatures induced by BNT162b2 vaccination.**
**a**, PCA analysis of bulk RNA-sequencing samples. **b**, RLE analysis of bulk RNA-sequencing data. **c**, **d**, Temporal expression patterns of genes within modules M111.0 (**c**) or M16 (**d**). Black lines represent the median fold change of all genes. **e**, Number of genes differentially expressed (absolute $\log_2$-transformed fold change > 0.2 and Wald $P < 0.01$) at each time point. All time points were compared to universal baseline, day 0. Number of upregulated and downregulated genes are shown in orange and green, respectively. **f**, BTMs that were significantly enriched (FDR < 0.05, absolute NES > 2) after vaccination. GSEA

was used to identify increased (red) or decreased (blue) enrichment of BTMs within gene lists ranked by Wald statistic between before and after vaccination at each time point. **g**, BTMs on day 22 relative to day 21 that were significantly associated with age. **h, i**, Scatter plots showing two-sided Pearson correlation of the mean fold change between day 22 and day 21 of genes in interferon modules M127, M75, M150, M111.1 and M68 (**h**) or module M16 (**i**) with IFNγ. Each dot represents an individual in all plots. Blue and red colour denote female and male participants, respectively.
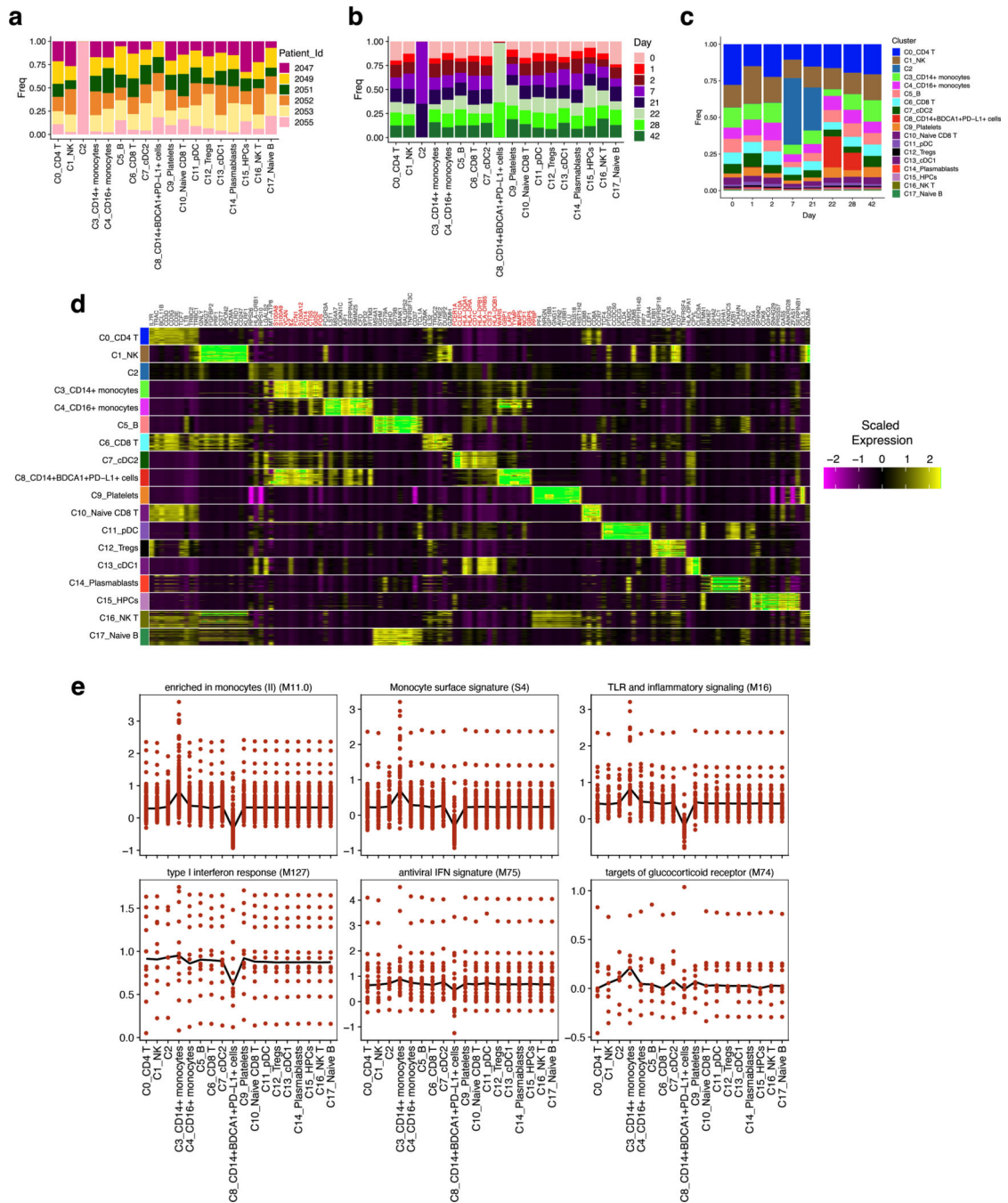
**Extended Data Fig. 9 |. CITE-seq analysis of PBMCs.**

**a**, **b**, Fraction of cells in each cluster of CITE-seq data classified by subject (**a**) or time point (**b**), determined from all single cells that passed quality control. **c**, Proportion of cells in each cluster at different time points indicated on *x* axis. **d**, Heat map showing top 10 cluster-defining genes in each cluster. **e**, Pseudobulk gene expression score showing the contribution from each cluster after iterative removal of each cluster.
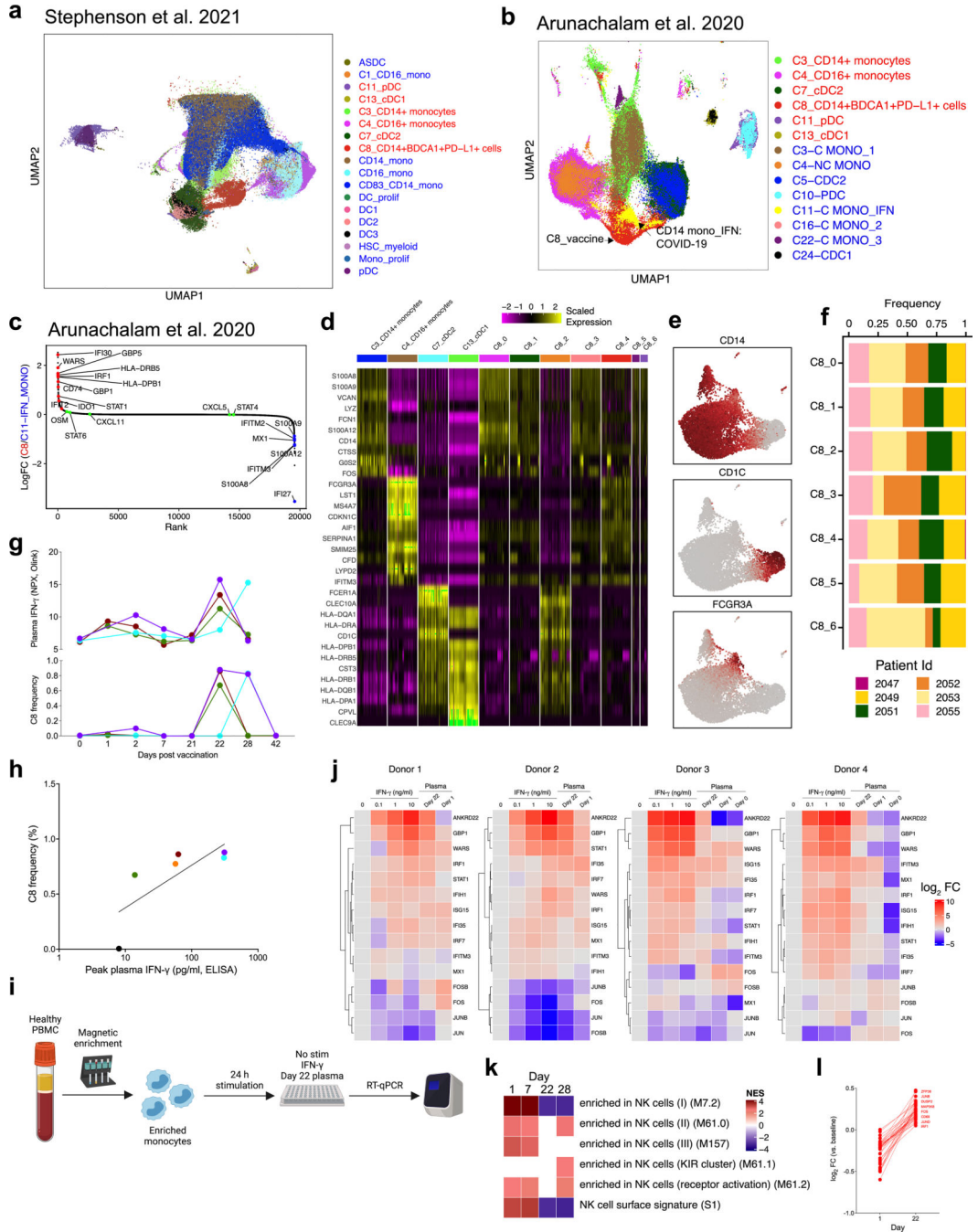


**Extended Data Fig. 10 |. Analysis of gene expression within cluster C8.**

**a**, **b**, UMAP representation of 71,276 innate immune cells (myeloid cells + pDCs) from this study combined with innate cells from ref. [23] (**a**, 154,108 cells) or ref. [4] (**b**, 18,926 cells). Red and blue fonts on the legends denote cell clusters defined in the current study versus previously published studies. **c**, Rank plot showing genes overexpressed in cluster C8 from this study in comparison to IFN-experienced monocyte cluster ('C11_C Mono_IFN'). Genes expressed more in C8 are in red font; genes in blue font are expressed at a higher level in C11_C Mono_IFN. **d**, Heat map showing genes in C3 CD14$^+$ monocytes, C4 CD16$^+$ monocytes, C7 cDC2 and C13 cDC1 that are shared with C8 subclusters and are closest with the parental clusters on the basis of Euclidean distance. **e**, Feature plots showing expression of CD14, CD1c or CD16 within C8 subclusters. **f**, Fraction of cells in each subcluster of C8 classified by subject. **g**, Frequency of C8 as a proportion of Lin$^-$ HLA-DR$^+$ population (bottom) or plasma IFN$\gamma$ levels (top) as measured in Olink. **h**, Correlation between C8 frequency and plasma IFN$\gamma$ levels measured by enzyme-linked immunosorbent assay. **i**, Schematic of the experimental set up for in vitro stimulation of purified healthy monocytes with IFN$\gamma$ and day-22 plasma from vaccinated participants. **j**, Heat maps showing expression of genes measured by quantitative real-time PCR (qRT–PCR) relative to the 'no stimulation' condition. Each column shows the condition as shown on the top. Each row represents a gene quantified by qRT–PCR. Five hundred thousand PBMCs from four donors were treated with different concentration of IFN$\gamma$, 1:4 diluted plasma sample from day 0, 1 or 22. **k**, Significantly enriched natural killer cell modules (FDR < 0.05, absolute NES>2) within the natural killer cell cluster. Days 1 and 7 were compared against day 0; days 22 and 28 were compared against day 21. GSEA was used to identify increased (red) or decreased (blue) enrichment of interferon BTMs. **l**, log$_2$-transformed fold change of top 50 DEGs between natural killer cells from day 22 samples versus day 1 samples.
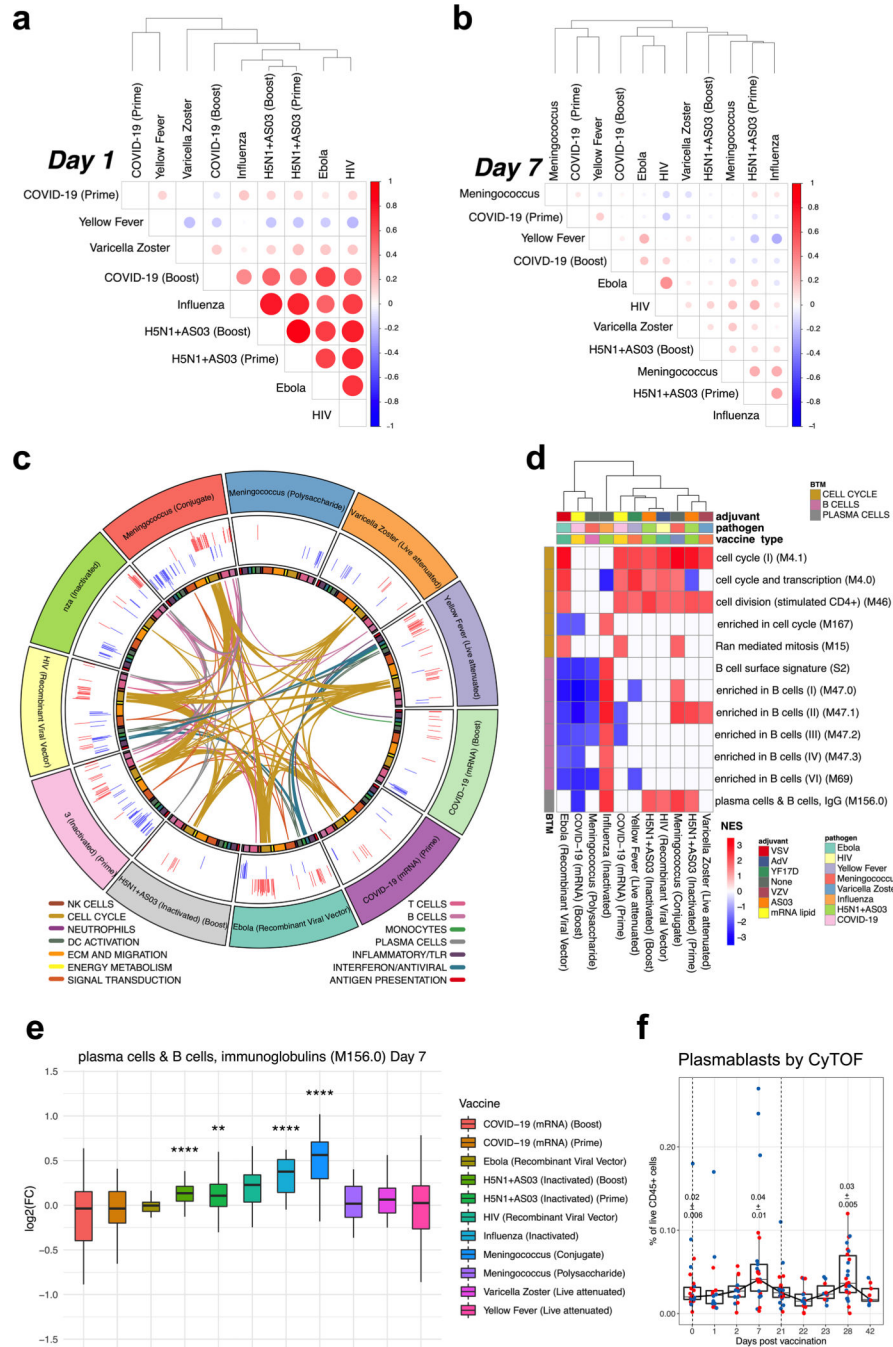
**Extended Data Fig. 11 |. Comparison of transcriptional responses with other vaccines.**
**a**, Correlation matrix of vaccines on day 1. Spearman's correlation was computed using mean fold changes over all genes between each pair of vaccines. Circle size and colour represents the correlation coefficient. **b**, Correlation matrix of vaccines on day 7. **c**, Circos plot of the overlap across vaccines in enriched BTMs on day 7. GSEA was performed on genes ranked by day 7 versus baseline *t*-statistic in each vaccine. Each segment of the circle represents one vaccine, and each point in a segment represents a single BTM. Bars in outer circle represent the NES of significantly enriched BTMs (FDR < 0.05). Lines

connect BTMs with a significant positive enrichment shared between vaccines. Inner circle boxes and line colours represent the functional groups of the BTMs. **d**, Heat map of cell cycle, B cell and plasma cell BTMs on day 7. Cell cycle, B cell and plasma cell BTMs that were significantly enriched (FDR < 0.05) on day 7 after either dose of BNT162b2 are shown. Colour represents significant NES. **e**, Expression of genes in the plasma cell module M156 across different vaccines on day 7 after vaccination. The asterisks denote statistical significance measured by two-sided Wilcoxon test between baseline and day-7 samples within each vaccine group. **f**, Plasmablast frequencies measured by CyTOF in whole-blood samples. The numbers within plots indicate geometric mean ± s.e.m. at the time points as indicated. Each dot represents an individual in all plots. Blue and red colour denote female and male participants, respectively. Boxes show median and 25th–75th percentiles, and whiskers show the range in all box plots. The statistically significant differences between the peak and baseline time points were measured using two-sided Wilcoxon matched-pairs signed-rank test.
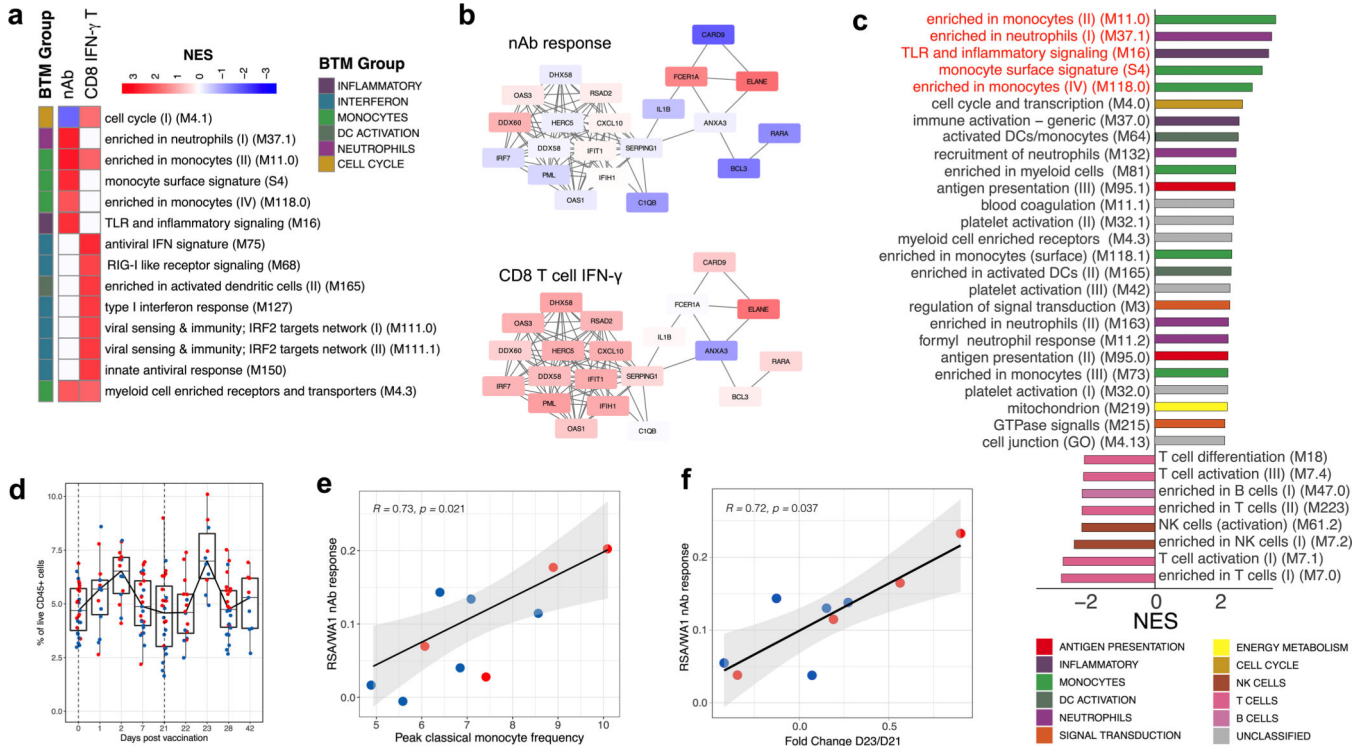


**Extended Data Fig. 12 |. Transcriptional correlates of neutralizing antibody and T cell responses.**
**a**, BTMs (day 22 relative to day 21) associated with the neutralizing antibody or CD8 IFNγ T cell response to BNT162b2. GSEA was run using BTMs on gene lists ranked by correlation with either day-42 neutralizing antibody titres or day-28 antigen-specific CD8⁺IFNγ⁺ T cell frequencies. Modules shown are those with NES > 2 and FDR < 0.05. **b**, Correlation of genes in M75, an antiviral BTM, on day 22 with the day-42 neutralizing antibody response (top) or day-28 antigen-specific CD8⁺IFNγ⁺ T cell frequencies (bottom). Colour represents two-sided Pearson correlation coefficient. Each edge (grey line) represents a co-expression relationship, as previously described[18]. **c**, BTMs

correlated with cross-neutralization index, ratio of B.1.351:WA1 neutralizing antibody titres, analysed as in **a**. Modules shown are those with NES > 2 and FDR < 0.05. **d**, Frequency of classical monocytes (Lin⁻HLA-DR⁺CD14⁺CD16⁻ cells) in whole-blood samples analysed by CyTOF. **e**, **f**, Scatter plots of two-sided Spearman's correlation between cross-neutralization index and peak (day 23) classical monocyte frequency (**e**) or a gene score created in the bulk RNA-sequencing data using the cluster-defining genes of the classical monocyte cluster, C3, in CITE-seq (**f**). The error bands represent 95% confidence limits. Each dot represents an individual in all plots. Blue and red colour denote female and male participants, respectively in **d**–**f**. Boxes show median and 25th–75th percentiles, and whiskers show the range in **d**.

**Extended Data Table 1 |**

Participant demographics

| Variable | N (%) |
|---|---|
| **Age (median [range])** | 36 (19–79) |
| **Sex** | |
| Female | 28 (50) |
| Male | 28 (50) |
| **Race** | |
| White | 27 (48.2) |
| Asian | 22 (39.3) |
| Black | 4 (7.1) |
| Native American | 1 (1.8) |
| Other | 2 (3.6) |
| **Ethnicity** | |
| Not of Hispanic, Latinx, or Spanish origin | 54 (96.4) |
| Hispanic, Latinx, Spanish origin | 2 (3.6) |
| **Comorbidities** | |
| Lung disease | 8 (14.2) |
| Diabetes mellitus | 2 (3.6) |
| Hypertension | 1 (1.8) |
| Cardiovascular | 2 (3.6) |
| Liver disease | 0 (0) |
| Renal diseases | 1 (1.8) |
| Cancer | 1 (1.8) |
| Hematological disorder | 1 (1.8) |
| Pregnancy | 0 (0) |
| Neurological | 4 (7.1) |
| HIV | 0 (0) |
| Solid organ transplant recipient | 0 (0) |
| Bone marrow transplant recipient | 0 (0) |
| Hyperlipidemia | 1 (1.8) |
| Gastrointestinal | 4 (7.1) |

| Variable | N (%) |
|---|---|
| Psychiatric | 1 (1.8) |
| Genitourinary disorder | 3 (5.3) |
| Hypersensitivity | 6 (10.7) |
| Autoimmune disorder | 4 (7.1) |
| **Confirmed SARS-CoV-2 infection** | 4 (7.1) |
| Time from SARS-CoV-2 infection to baseline visit in days (median [range]) | 180 (48–270) |

**Extended Data Table 2 |**

Vaccine side effects and symptoms

| Variable | N (%) | Variable | N (%) | p-value |
|---|---|---|---|---|
| **First Dose** | | **Second Dose** | | |
| None | 11 (19.6) | None | 5 (8.9) | 0.177 |
| Fever | 2 (3.6) | Fever | 7 (12.5) | 0.154 |
| Site tenderness | 39 (69.6) | Site Tenderness | 40 (71.4) | 0.881 |
| Site swelling | 6 (10.7) | Site Swelling | 16 (28.6) | 0.029 |
| Site redness | 4 (7.1) | Site Redness | 14 (25) | 0.018 |
| Muscle aches | 15 (26.8) | Muscle aches | 29 (51.8) | 0.015 |
| Fatigue | 15 (26.8) | Fatigue | 21 (37.5) | 0.28 |
| Headache | 10 (17.9) | Headache | 17 (30.4) | 0.167 |
| Chills | 2 (3.6) | Chills | 10 (17.9) | 0.03 |
| Joint pain | 1 (1.8) | Joint Pain | 4 (7.1) | 0.349 |
| Nausea | 1 (1.8) | Nausea | 2 (3.6) | 0.49 |
| Difficulty breathing | 1 (1.8) | Difficulty Breathing | 1 (1.8) | 1 |
| Hives | 0 (0) | Hives | 1 (1.8) | 0.993 |
| Swelling | 0 (0) | Swelling | 2 (3.6) | 0.468 |
| Tachycardia | 0 (0) | Tachycardia | 1 (1.8) | 0.993 |
| Rash | 0 (0) | Rash | 2 (3.6) | 0.468 |
| Dizziness | 2 (3.6) | Dizziness | 5 (8.9) | 0.317 |
| Sweating | 2 (3.6) | Sweating | 1 (1.8) | 1 |
| Brain fog | 1 (1.8) | Brain Fog | 0 (0) | 0.993 |
| Loss of appetite | 1 (1.8) | Loss of appetite | 0 (0) | 0.993 |
| Rhinitis | 1 (1.8) | Rhinitis | 0 (0) | 0.993 |
| Paresthesia | 1 (1.8) | Paresthesia | 1 (1.8) | 1 |
| Diarrhea | 1 (1.8) | Diarrhea | 0 (0) | 0.993 |
| Vertigo | 0 (0) | Vertigo | 1 (1.8) | 0.993 |
| **Site of injection** | | **Site of injection** | | |
| Right arm | 5 (8.9) | Right arm | 8 (14.3) | |
| Left arm | 51 (91.1) | Left arm | 48 (85.7) | |

The statistical differences were determined using $\chi^2$ test for categorical variables and two-sided $t$-test for continuous variables considering each parameter individually (no multiple corrections).

**Extended Data Table 3 |**

Vaccine meta-analysis datasets

| Vaccine | Pathogen | Vaccine Type | Adjuvant/ Vector | Timepoints used | N | GEO/ ImmPort |
|---|---|---|---|---|---|---|
| BNT162b2 | SARS-CoV-2 | mRNA | mRNA-LNP | 0,1,7,21,22,28 | 31 | GSE169159 |
| TIV | Seasonal Influenza | Inactivated | None | 0,1,7 | 19 | GSE74813/ SDY56 |
| VZV | Varicella zoster | Live attenuated | VZV | 0,1,7 | 31 | GSE79396/ SDY984 |
| YF17D | Yellow fever | Live attenuated | YF17D | 0,1,7 | 25 | GSE13486/ SDY1264 |
| rVSV-ZEBOV (high dose) | Ebola | Recombinant viral vector | VSV | 0,1,7 | 7 | GSE97590/ SDY1373 |
| MRKAd5/HIV | HIV | Recombinant viral vector | Ad5 | 0,1,7 | 10 | GSE22768/ SDY1291 |
| H5N1+AS03 | H5N1 Influenza | Inactivated | AS03 | 0,1,7,21,22,28 | 33 | GSE102012 |
| MPSV4 | Meningococcus | Polysaccharide | None | 0,7 | 13 | GSE52245/ SDY1260 |
| MCV4 | Meningococcus | Conjugate | None | 0,7 | 17 | GSE52245/ SDY1260 |

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

## References

1. Polack FP et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. N. Engl. J. Med. 383, 2603–2615 (2020). [PubMed: 33301246]

2. Mulligan MJ et al. Phase I/II study of COVID-19 RNA vaccine BNT162b1 in adults. Nature 586, 589–593 (2020). [PubMed: 32785213]

3. Sahin U et al. COVID-19 vaccine BNT162b1 elicits human antibody and TH1 T cell responses. Nature 586, 594–599 (2020). [PubMed: 32998157]

4. Arunachalam PS et al. Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans. Science 369, 1210–1220 (2020). [PubMed: 32788292]

5. Nakaya HI et al. Systems biology of vaccination for seasonal influenza in humans. Nat. Immunol. 12, 786–795 (2011). [PubMed: 21743478]

6. Querec TD et al. Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans. Nat. Immunol. 10, 116–125 (2009). [PubMed: 19029902]

7. Gaucher D et al. Yellow fever vaccine induces integrated multilineage and polyfunctional immune responses. J. Exp. Med. 205, 3119–3131 (2008). [PubMed: 19047440]

8. Tsang JS et al. Global analyses of human immune variation reveal baseline predictors of postvaccination responses. Cell 157, 499–513 (2014). [PubMed: 24725414]

9. Doria-Rose N et al. Antibody persistence through 6 months after the second dose of mRNA-1273 vaccine for Covid-19. N. Engl. J. Med. 384, 2259–2261 (2021). [PubMed: 33822494]

10. Samanovic MI et al. Poor antigen-specific responses to the second BNT162b2 mRNA vaccine dose in SARS-CoV-2-experienced individuals. Preprint at 10.1101/2021.02.07.21251311 (2021).

11. Liu Y et al. Neutralizing activity of BNT162b2-elicited serum. N. Engl. J. Med. 384, 1466–1468 (2021). [PubMed: 33684280]

12. Sette A & Crotty S Pre-existing immunity to SARS-CoV-2: the knowns and unknowns. Nat. Rev. Immunol. 20, 457–458 (2020). [PubMed: 32636479]

13. Woodruff MC, Ramonell RP, Lee FE-H & Sanz I Clinically identifiable autoreactivity is common in severe SARS-CoV-2 infection. Preprint at 10.1101/2020.10.21.20216192 (2020).

14. Zuo Y et al. Prothrombotic autoantibodies in serum from patients hospitalized with COVID-19. Sci. Transl. Med. 12, eabd3876 (2020). [PubMed: 33139519]

15. Wang EY et al. Diverse functional autoantibodies in patients with COVID-19. Nature 595, 283–288 (2021). [PubMed: 34010947]

16. Chang SE et al. New-onset IgG autoantibodies in hospitalized patients with COVID-19. Preprint at 10.1101/2021.01.27.21250559 (2021).

17. De Mot L et al. Transcriptional profiles of adjuvanted hepatitis B vaccines display variable interindividual homogeneity but a shared core signature. Sci. Transl. Med. 12, eaay8618 (2020). [PubMed: 33177181]

18. Li S et al. Molecular signatures of antibody responses derived from a systems biology study of five human vaccines. Nat. Immunol. 15, 195–204 (2014). [PubMed: 24336226]

19. Boraschi D & Italiani P Immunosenescence and vaccine failure in the elderly: strategies for improving response. Immunol. Lett. 162, 346–353 (2014). [PubMed: 24960535]

20. Schulte-Schrepping J et al. Severe COVID-19 is marked by a dysregulated myeloid cell compartment. Cell 182, 1419–1440 (2020). [PubMed: 32810438]

21. Wilk AJ et al. A single-cell atlas of the peripheral immune response in patients with severe COVID-19. Nat. Med. 26, 1070–1076 (2020). [PubMed: 32514174]

22. Bakdash G et al. Expansion of a BDCA1+CD14+ myeloid cell population in melanoma patients may attenuate the efficacy of dendritic cell vaccines. Cancer Res. 76, 4332–4346 (2016). [PubMed: 27325645]

23. Stephenson E et al. Single-cell multi-omics analysis of the immune response in COVID-19. Nat. Med. 27, 904–916 (2021). [PubMed: 33879890]

24. Korsunsky I et al. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat. Methods 16, 1289–1296 (2019). [PubMed: 31740819]

25. Wimmers F et al. The single-cell epigenomic and transcriptional landscape of immunity to influenza vaccination. Cell 184, 3915–3935 (2021). [PubMed: 34174187]

26. Ponti C et al. Role of CREB transcription factor in c-fos activation in natural killer cells. Eur. J. Immunol. 32, 3358–3365 (2002). [PubMed: 12432566]

27. Nakaya HI et al. Systems analysis of immunity to influenza vaccination across multiple years and in diverse populations reveals shared molecular signatures. Immunity 43, 1186–1198 (2015). [PubMed: 26682988]

28. Hadjadj J et al. Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. Science 369, 718–724 (2020). [PubMed: 32661059]

29. Mathew D et al. Deep immune profiling of COVID-19 patients reveals distinct immunotypes with therapeutic implications. Science 369, eabc8511 (2020). [PubMed: 32669297]

30. Domínguez-Andrés J & Netea MG The specifics of innate immune memory. Science 368, 1052–1053 (2020). [PubMed: 32499424]

31. Vanderheiden A et al. Development of a rapid focus reduction neutralization test assay for measuring SARS-CoV-2 neutralizing antibodies. Curr. Protoc. Immunol 131, e116 (2020). [PubMed: 33215858]

32. Xie X et al. An infectious cDNA clone of SARS-CoV-2. Cell Host Microbe 27, 841–848 (2020). [PubMed: 32289263]

33. Katzelnick LC et al. Viridot: an automated virus plaque (immunofocus) counter for the measurement of serological neutralizing responses with application to dengue virus. PLoS Negl. Trop. Dis. 12, e0006862 (2018). [PubMed: 30356267]

34. Tegally H et al. Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. Preprint at 10.1101/2020.12.21.20248640 (2020).

35. Arunachalam PS et al. Adjuvanting a subunit COVID-19 vaccine to induce protective immunity. Nature 594, 253–258 (2021). [PubMed: 33873199]

36. Ayoglu B et al. Anoctamin 2 identified as an autoimmune target in multiple sclerosis. Proc. Natl Acad. Sci. USA 113, 2188–2193 (2016). [PubMed: 26862169]

37. Degn SE et al. Clonal evolution of autoreactive germinal centers. Cell 170, 913–926 (2017). [PubMed: 28841417]

38. R Core Team. R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, 2013).

39. Nowicka M et al. CyTOF workflow: differential discovery in high-throughput high-dimensional cytometry datasets. F1000Res. 6, 748 (2017). [PubMed: 28663787]

40. Risso D, Schwartz K, Sherlock G & Dudoit S GC-content normalization for RNA-seq data. BMC Bioinformatics 12, 480 (2011). [PubMed: 22177264]

41. Love MI, Huber W & Anders S Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550 (2014). [PubMed: 25516281]

42. Irizarry RA et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4, 249–264 (2003). [PubMed: 12925520]

43. Stuart T et al. Comprehensive integration of single-cell data. Cell 177, 1888–1902 (2019). [PubMed: 31178118]

44. Bais AS & Kostka D scds: computational annotation of doublets in single-cell RNA sequencing data. Bioinformatics 36, 1150–1158 (2020). [PubMed: 31501871]

45. Finak G et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. Genome Biol. 16, 278 (2015). [PubMed: 26653891]
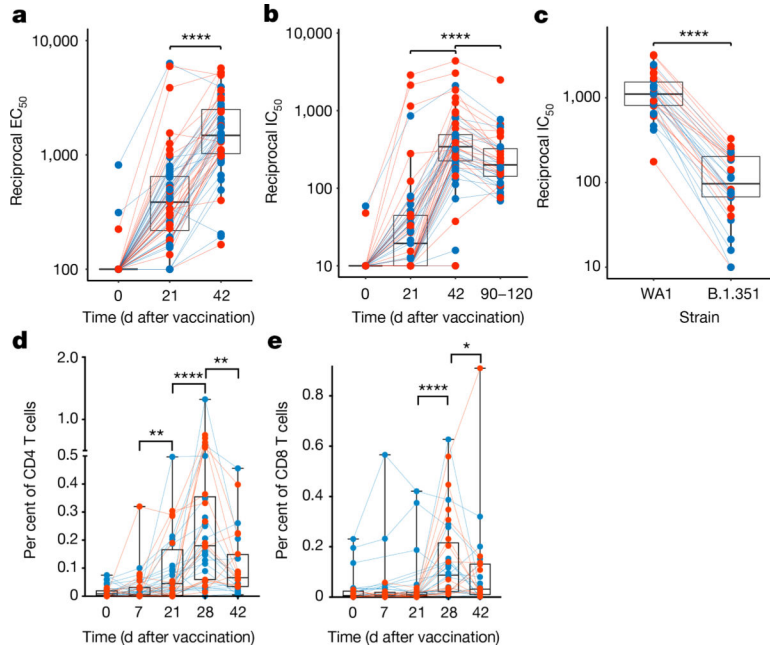
**Fig. 1 |. BNT162b2 vaccination induces robust antibody and T cell responses.**
**a**, **b**, Serum anti-S IgG titres (**a**) (half-maximal effective concentration ($EC_{50}$)) and live-virus
neutralizing antibody titres (**b**) (half-maximal inhibitory concentration ($IC_{50}$)), measured
by enzyme-linked immunosorbent assay and focus reduction neutralization titre assays,
respectively. **c**, Live-virus neutralizing antibody response against the wild-type (WA1) or
B.1.351 variant measured in day-42 sera. **d**, **e**, Spike-specific CD4 (**d**) and CD8 (**f**) T
cell IFN$\gamma$ response in blood. Each dot represents a participant ($n = 56$ (**a**, **b**), 30 (**c**) and
38 (**d**, **e**)). Statistical differences were calculated using two-sided Wilcoxon matched-pairs
signed-rank test. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$. Boxes show median
and 25th–75th percentiles, and whiskers show the range.
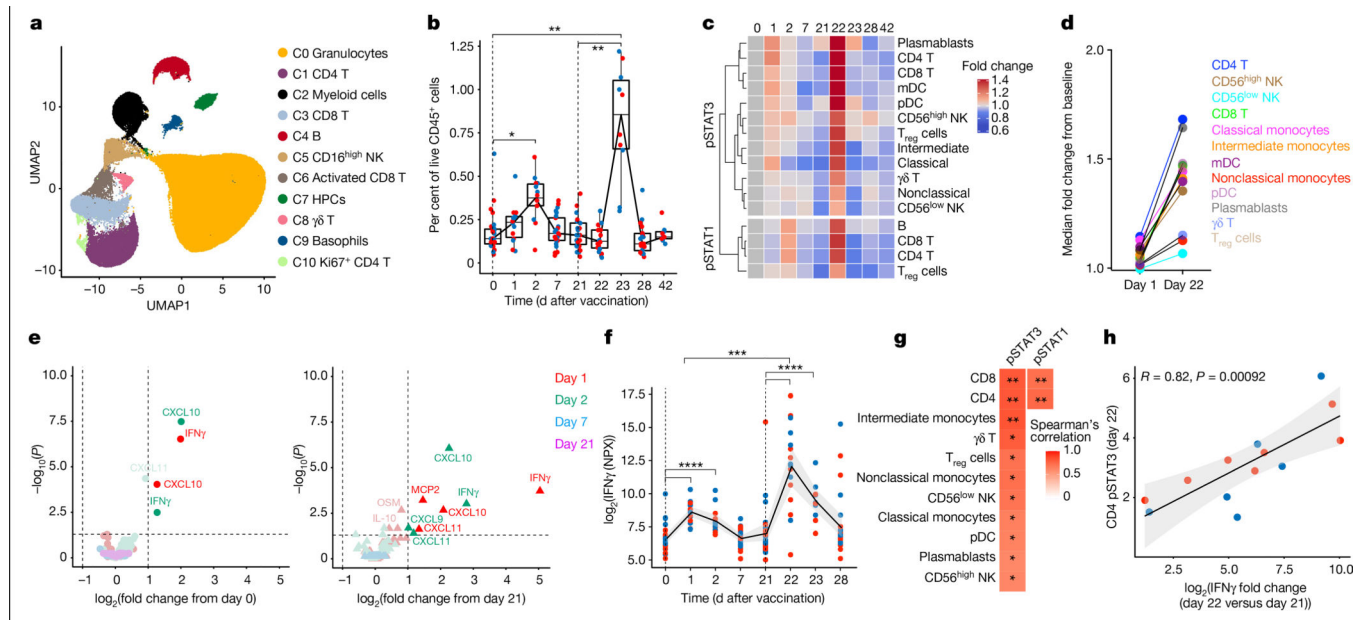
**Fig. 2 |. Innate immune responses induced by BNT162b2 vaccination.**
**a**, CyTOF-identified cell clusters from whole blood visualized by uniform manifold approximation and projection (UMAP). HPCs, haematopoietic progenitor cells; NK, natural killer. **b**, Frequency of inflammatory monocytes ($CD14^+CD16^+$ monocytes) as a proportion of live $CD45^+$ cells. Boxes show median and 25th–75th percentiles, and whiskers show the range. **c**, Heat map of fold change of pSTAT3 and pSTAT1 levels in comparison to baseline in the cell types indicated on the $y$ axis. Only statistically significant changes between the increase on day 1 after primary and day 1 after secondary immunizations, as measured using two-sided Mann–Whitney rank-sum test ($P < 0.05$), were plotted. mDC, myeloid dendritic cell; pDC, plasmacytoid dendritic cell; $T_{reg}$, regulatory T. **d**, Fold change in pSTAT3 levels on days 1 and 22, compared to primary and secondary baselines, respectively. **e**, Volcano plots showing plasma cytokines significantly increased after primary (left) and secondary (right) vaccinations versus day 0 and day 21, respectively. **f**, Plasma IFNγ levels measured by Olink. NPX, normalized protein expression. **g**, Heat map of two-sided Spearman's correlation between increase in plasma IFNγ and pSTAT3 or pSTAT1 levels in different cell types, shown on the $y$ axis, on day 1 after secondary vaccination. The $P$ values were corrected for multiple testing. **h**, Spearman's correlation between pSTAT3 levels in CD4 T cells and plasma IFNγ levels. The error bands represent 95% confidence limits. In **b**, **f**, statistical differences between the peak and baseline time points were measured using two-sided Wilcoxon matched-pairs signed-rank test. The differences between peak time points were measured using two-sided Mann–Whitney rank-sum test. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$. Blue and red dots indicate female and male participants, respectively. $n = 27$ and 31 participants for CyTOF and Olink.
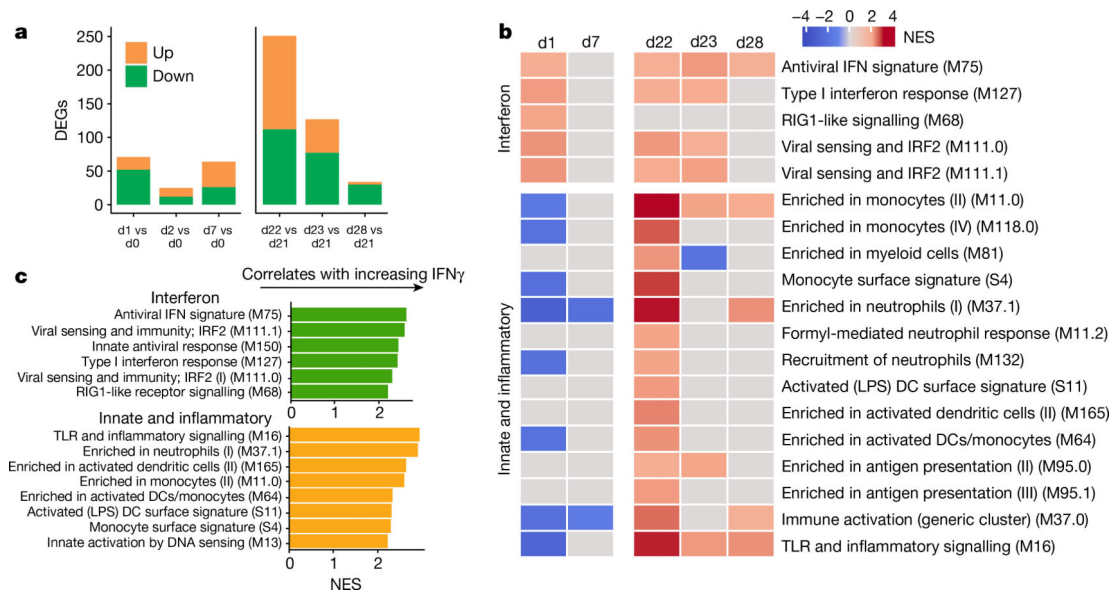
**Fig. 3 |. Transcriptional signatures of BNT162b2 vaccination.**

**a**, Number of differentially expressed genes (DEGs) (absolute $\log_2$-transformed fold change > 0.2 and Wald $P$ < 0.01) at each time point after prime (left) and boost (right). d, day. **b**, Interferon and innate BTMs that were significantly enriched (false discovery rate (FDR) < 0.05, absolute normalized enrichment score (NES) > 2) after prime (left) and boost (right) vaccination. Days 1, 2 and 7 were compared against day 0; days 22, 23 and 28 were compared against day 21 in **a**, **b**. **c**, BTMs on day 22 that were significantly associated with plasma IFNγ. GSEA was used to identify enrichment of BTMs within gene lists ranked by correlation with fold change in IFNγ between days 22 and 21.
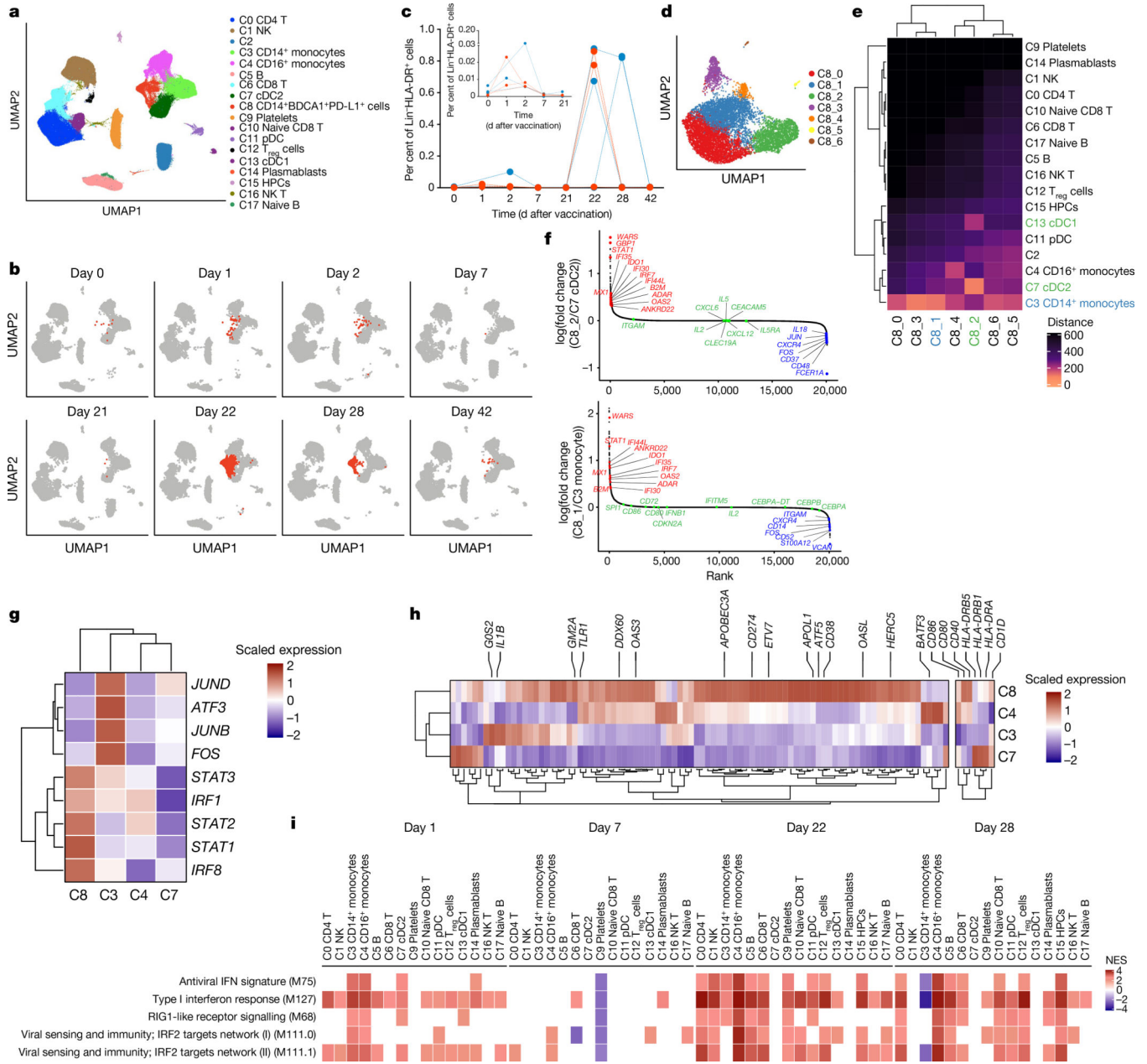
**Fig. 4 |. Single-cell transcriptional response to BNT162b2 vaccination.**

**a**, UMAP representation of cell types identified by single-cell transcriptional profiling.
**b**, Feature plots across time points, showing cluster C8 in red. **c**, Frequency of C8 as a
proportion of Lin⁻HLA-DR⁺ cell clusters (C3, C4, C7, C8, C11 and C13). Each dot or line
represents an individual. Blue and red indicate female and male participants, respectively.
**d**, UMAP representation of subclusters within C8 resolved using Louvain clustering. **e**,
Heat map showing Euclidean distance between C8 subclusters and the rest of the cell
types identified by single-cell profiling. **f**, DEGs determined between C8_2 and C8_1, and
their closest cell clusters ranked by fold change between C8 and their parental clusters,
are plotted. Genes in red, green and blue represent genes upregulated, unchanged and

downregulated, respectively, in the C8 subcluster as compared to its closest parental cluster. **g**, Heat map showing scaled expression of key interferon-response and AP-1 transcription factors. **h**, Heat map showing an extended set of genes from ref. [25] that showed enhanced accessibility on day 42 after H5N1 + AS03 vaccination. **i**, Significantly enriched interferon BTMs (FDR < 0.05, absolute NES > 2) across clusters over time. Days 1 and 7 were compared against day 0; days 22 and 28 were compared against day 21.
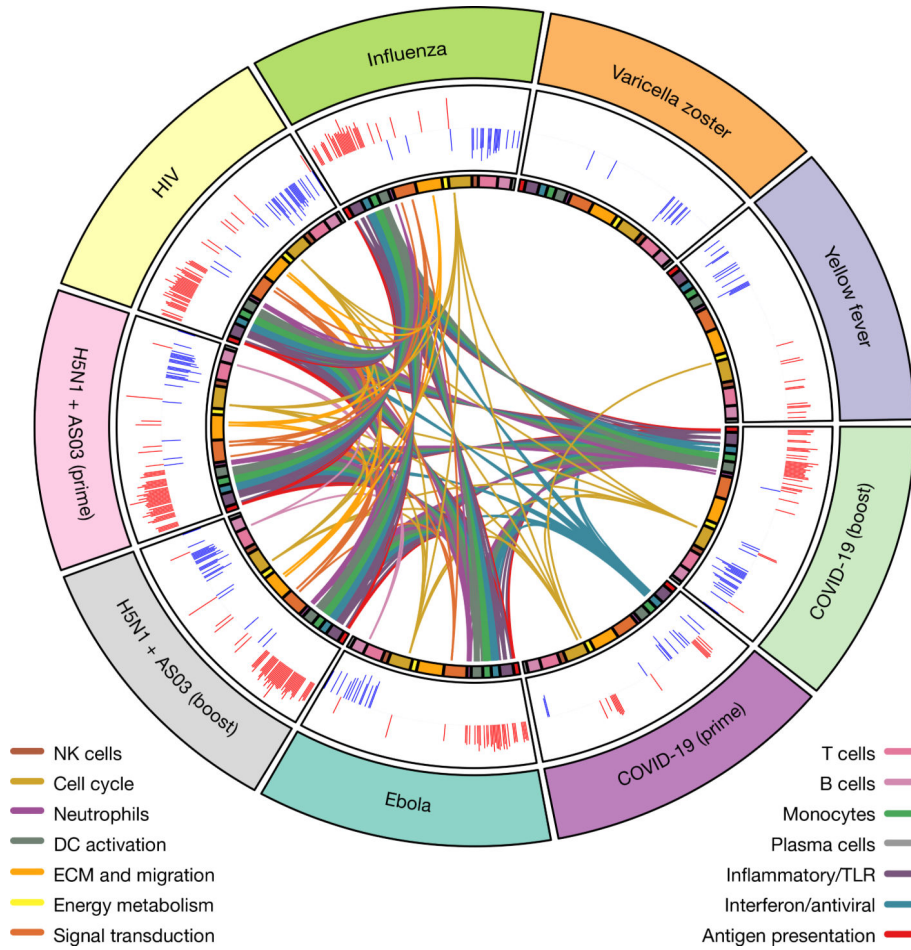
**Fig. 5 |. Comparison of transcriptional responses with other vaccines.**
Circos plot of the overlap across vaccines in BTMs enriched on day 1. GSEA was performed on genes ranked by day 1 versus baseline *t*-statistic in each vaccine. Each segment of the circle represents one vaccine, and each point in a segment represents a single BTM. Bars in outer circle represent the NES of significantly enriched BTMs (FDR < 0.05). Lines connect BTMs with a significant positive enrichment shared between vaccines. Inner circle boxes and line colours represent the functional groups of the BTMs. ECM, extracellular matrix.