

Accelerated Article Preview

Influenza vaccination reveals sex dimorphic imprints of prior mild COVID-19

Received: 3 February 2022

Accepted: 19 December 2022

Accelerated Article Preview

Cite this article as: Sparks, R. et al. Influenza vaccination reveals sex dimorphic imprints of prior mild COVID-19. *Nature* <https://doi.org/10.1038/s41586-022-05670-5> (2023)

Rachel Sparks, William W. Lau, Can Liu, Kyu Lee Han, Kiera L. Vrindten, Guangping Sun, Milann Cox, Sarah F. Andrews, Neha Bansal, Laura E. Failla, Jody Manischewitz, Gabrielle Grubbs, Lisa R. King, Galina Koroleva, Stephanie Leimenstoll, LaQuita Snow, OP11 Clinical Staff, Jinguo Chen, Juanjie Tang, Amrita Mukherjee, Brian A. Sellers, Richard Apps, Adrian B. McDermott, Andrew J. Martins, Evan M. Bloch, Hana Golding, Surender Khurana & John S. Tsang

This is a PDF file of a peer-reviewed paper that has been accepted for publication. Although unedited, the content has been subjected to preliminary formatting. Nature is providing this early version of the typeset paper as a service to our authors and readers. The text and figures will undergo copyediting and a proof review before the paper is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers apply.

Influenza vaccination reveals sex dimorphic imprints of prior mild COVID-19

Rachel Sparks^{1,10}, William W. Lau^{1,10}, Can Liu^{1,2,10}, Kyu Lee Han³, Kiera L. Vrindten¹, Guangping Sun^{1,4}, Milann Cox¹, Sarah F. Andrews⁵, Neha Bansal¹, Laura E. Failla¹, Jody Manischewitz⁶, Gabrielle Grubbs⁶, Lisa R. King⁶, Galina Koroleva³, Stephanie Leimenstoll⁷, LaQuita Snow⁷, OP11 Clinical Staff⁸, Jinguo Chen³, Juanjie Tang⁶, Amrita Mukherjee³, Brian A. Sellers³, Richard Apps³, Adrian B. McDermott⁵, Andrew J. Martins¹, Evan M. Bloch⁹, Hana Golding⁶, Surender Khurana⁶, John S. Tsang^{1,3,*}

1 Multiscale Systems Biology Section, Laboratory of Immune System Biology, NIAID, NIH, Bethesda, MD, USA

2 Graduate Program in Biological Sciences, University of Maryland, College Park, MD, USA

3 NIH Center for Human Immunology, NIAID, NIH, Bethesda, MD, USA

4 Division of Intramural Research, NIAID, NIH, Bethesda, MD, USA

5 Vaccine Research Center, NIAID, NIH, Bethesda, MD, USA

6 Division of Viral Products, Center for Biologics Evaluation and Research (CBER), FDA, Silver Spring, MD, USA

7 Laboratory of Clinical Immunology and Microbiology, NIAID, NIH, Bethesda, MD, USA

8 See consortium author listing below.

9 Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD, USA

10 These authors contributed equally

*Correspondence (JST): john.tsang@yale.edu

27 **Abstract**

28

29 Acute viral infections can have durable functional impacts on the immune system long after recovery,
30 but how they affect homeostatic immune states and responses to future perturbations remain poorly
31 understood¹⁻⁴. Here we use systems immunology approaches, including longitudinal multimodal single
32 cell analysis (surface proteins, transcriptome, and V(D)J sequences), to comparatively assess baseline
33 immune statuses and responses to influenza vaccination in 33 healthy individuals after recovery from
34 mild, non-hospitalized COVID-19 (mean: 151 days after diagnosis) and 40 age- and sex-matched
35 controls who never had COVID-19. At baseline and independent of time since COVID-19, recoverees
36 had elevated T-cell activation signatures and lower expression of innate immune genes in monocytes.
37 COVID-19-recovered males had coordinately higher innate, influenza-specific plasmablast, and
38 antibody responses after vaccination compared to healthy male and COVID-19-recovered females,
39 partly because male recoverees had monocytes with higher IL-15 responses early after vaccination
40 coupled with elevated pre-vaccination frequencies of “virtual memory” like CD8+ T-cells poised to
41 produce more IFN γ upon IL-15 stimulation. In addition, the expression of the repressed innate immune
42 genes in monocytes increased by day 1 through day 28 post-vaccination in recoverees, thus moving
43 towards the pre-vaccination baseline of healthy controls. In contrast, these genes decreased on day 1
44 and returned to the baseline by day 28 in controls. Our study reveals sex-dimorphic impacts of prior
45 mild COVID-19 and suggests that viral infections in humans can establish new set-points impacting
46 future immune responses in an antigen-agnostic manner.

47 **Introduction**

48
49 Examples of long-term immunological effects of both chronic and resolved viral infections have been
50 described, e.g., following recovery from natural acute measles infection there is marked reduction in
51 humoral immunity and increased susceptibility to non-measles infections for months to years¹. Live
52 vaccines such as BCG and measles can impart “training” effects on innate immune cells such as
53 monocytes and their long-lived progenitors, which could underlie the pathogen non-specific effects of
54 BCG in reducing all-cause mortality in infants^{5,6}. COVID-19 can result in persistent clinical sequelae for
55 months after infection, both in hospitalized and mild cases⁷. While the spectrum of clinical
56 manifestations of post-acute COVID-19 syndrome (a.k.a “long COVID”) is expanding, our understanding
57 of the molecular and cellular immunological changes after recovery from SARS-CoV-2 infection is
58 lacking. A better understanding of functional immune imprints of mild COVID-19 might have
59 particularly important public health implications given that this population constitutes most COVID-19
60 recoverees. More broadly, the fundamental issues of whether and how “homeostatic” baseline
61 immune states may have been altered by viral infections, and whether any such alterations may affect
62 responses to future challenges (e.g., infection or vaccination, with shared or distinct antigens) remain
63 poorly understood.

64 Here we took advantage of a unique opportunity and unprecedented epidemiological
65 environment during the early fall of 2020, months after the first wave of COVID-19, when those with
66 mild COVID-19 had recovered clinically, but before they could be reinfected by SARS-CoV-2 or receive
67 COVID-19 vaccination (which was not available until late 2020); additionally, the prevalence of other
68 respiratory infections was extremely low during this time⁸. We enrolled and comparatively assessed
69 healthy individuals who: 1) recovered from non-hospitalized, mild cases of COVID-19, and 2) age- and

70 sex-matched controls who never had COVID-19, all from the same geographic region. In addition to
71 assessing the post-COVID-19 immunological statuses, we utilized influenza vaccination to evaluate the
72 immune responses of these two populations at the serological, transcriptional, proteomic, and cellular
73 levels. These analyses reveal basic principles regarding what happens to the immune system after two
74 well-defined immunological encounters in humans: mild COVID-19 as a natural infectious perturbation
75 and influenza vaccination as a controlled and timed intervention with non-SARS-CoV-2 antigens.

76 77 **Results**

78
79 Individuals with prior symptomatic SARS-CoV-2 infection (n=31; diagnosed by nasal PCR test) or
80 asymptomatic infection (n=2; by antibody test, see Methods) during early 2020, and age- and sex-
81 matched healthy controls (HC; n=40) with no history of COVID-19 (and negative by antibody test) were
82 recruited from the community during the fall of 2020 and followed longitudinally (Fig. 1a, see
83 Methods). The average time since COVID-19 diagnosis was 151 days for recoverees (COVR) (Extended
84 Data Fig. 1a; Extended Data Table 1), who had clinically mild illness during acute disease that did not
85 require hospitalization (self-reported average length of illness: 16.5 days) and no major medical
86 comorbidities, including infection at the time of enrollment, obesity (BMI > 30) or autoimmune disease
87 (Fig. 1b). None of the participants were enrolled in COVID-19 vaccine trials, nor did they receive recent
88 vaccination of any kind before administration of the seasonal influenza vaccine in this study. A small
89 number of individuals continued to have mild self-reported sequelae from their illness at study
90 enrollment (3 males and 8 females), the most common being loss of taste and/or smell (Extended Data
91 Table 1). Females were more likely to have sequelae (Fisher's exact test $p = 0.09$ for all subjects, $p =$
92 0.03 for those < 65 years of age), at a rate similar to that reported in other large studies⁹.

93

94 **Baseline of mild COVID-19 recoverees**

95 Longitudinal multiomics profiling was performed using whole blood transcriptomics, single cell analysis
96 of 138 surface proteins, transcriptome, and V(D)J sequences via CITE-seq (Cellular Indexing of
97 Transcriptomes and Epitopes by Sequencing¹⁰), serum protein profiling, antibody characterization,
98 peripheral blood immune cell frequencies with hematological parameters from a complete blood count
99 (CBC), as well as clinical and research flow cytometry covering major immune cell lineages and subsets
100 (Fig. 1b, Supplementary Fig. 1). We first assessed baseline, pre-vaccination differences between the
101 recoverees and the age- and sex-matched HCs. As sex-dependent immune responses to COVID-19 have
102 been reported¹¹, our analyses explicitly searched for sex-dependent signatures. Immunological
103 resolution following infection may unfold over time even after symptoms subside, and there were
104 indeed parameters that showed evidence of continued evolution in our cohort—defined as those that
105 were correlated with time since COVID-19 diagnosis (TSD; Supplementary Table 1, see Methods),
106 including, as expected, SARS-CoV-2 neutralizing antibody titers¹² (Extended Data Fig. 1b). However, we
107 were primarily interested in uncovering persistent, TSD-independent post-COVID-19 immune imprints,
108 and thus we focused on “temporally stable” immune states associated with prior mild COVID-19 but
109 *not* correlated with TSD. We thus evaluated differences between 1) COVR females (COVR-F) vs. HC
110 females (HC-F); 2) COVR males (COVR-M) vs. HC males (HC-M); and 3) COVR-M vs. COVR-F after
111 accounting for male-female differences in HCs (herein referred to as “sex differences”; Supplementary
112 Table 2). The frequencies of myeloid cells such as monocytes and conventional/myeloid dendritic cells
113 (cDCs) tended to be higher in COVR-M than HC-M and/or COVR-F (Fig. 1c,d; Extended Data Fig. 1c,d),
114 consistent with reports of myeloid cell disruption in COVID-19, particularly in severe, acute disease¹³.

115 Here male-specific elevation in monocyte frequencies was detected even months after recovery from
116 mild disease.

117 Whole blood transcriptomic (WBT) data also revealed sex-dependent signatures associated
118 with prior COVID-19 (Extended Data Fig. 1e; e.g., the monocyte-related M11.0 and M4.0 from the
119 blood transcriptional module [BTM] collection), including metabolic signatures such as oxidative
120 phosphorylation (Supplementary Table 3). WBT differences can be driven by both cell composition and
121 cell intrinsic transcriptional changes. Indeed, the innate immune, metabolic, and T-cell-related
122 signatures are driven, at least in part, by the increased circulating monocyte and correspondingly lower
123 T-cell frequencies in COVR-M (Fig. 1d and Extended Data Fig. 1f) because these transcriptional
124 enrichment signals became statistically insignificant when monocyte frequencies were taken into
125 account (data not shown).

126 To assess transcriptional alterations independent of cell frequencies, we used CITE-seq to
127 examine cell type-specific contributions underlying the WBT signatures seen above. We clustered
128 single cells and annotated the resulting clusters using surface protein expression profiles (Fig. 1e; see
129 Methods). Cell type-specific transcriptional analysis pointed to both sex-dependent and -independent
130 differences between COVR and HCs (Supplementary Table 4). Among the enriched gene sets from the
131 WBT analysis above (Extended Data Fig. 1e), but now free of cell-frequency confounding, the BTM
132 M11.0/4.0 gene sets exhibit depressed expression in both classical and non-classical monocytes in
133 COVR relative to HCs in both sexes, while the converse is true for genes in the T-cell activation
134 signature (BTM M7.3) in both CD8+ central memory and effector memory (EM) T-cells (Fig. 1f-i;
135 Extended Data Fig. 1g; Supplementary Table 5); the T-cell activation signature in CD8+ EMs was
136 particularly pronounced in COVR-M (Fig. 1i). The genes driving the monocyte repression enrichment

137 [i.e., the so-called “leading edge genes” (LEGs)] include numerous surface receptors, such as those
138 encoding pattern recognition receptors (TLR2, TLR4, and TLR8), the peptidoglycan recognizing receptor
139 NOD2, the high affinity IgE FC receptor FCER1G, and C-type lectin receptor CLEC4A (Fig. 1f,g). This
140 “innate immune receptor” (IIR) signature in the monocytes, as well as the T-cell activation signature,
141 are predominantly not associated with TSD in both males and females (Extended Data Fig. 1h).

142 The T-cell activation signature likely emerged during and persisted after acute COVID-19¹⁴, but
143 this was less clear for the IIR signature. We thus asked whether this signature could be linked to gene
144 expression changes seen in acute COVID-19. Using a previously published CITE-seq dataset we
145 generated from a hospitalized, older, and male-biased severe COVID-19 cohort from Italy¹⁵, we noted
146 that within the classical monocytes, the average expression of the IIR LEGs from above was significantly
147 lower in acute COVID-19 patients than healthy controls and was negatively associated with disease
148 severity (Extended Data Fig. 1i). Thus, this depressed IIR signature could have originated from and
149 stably persisted since the acute response to the infection. Previous studies have reported several
150 (potentially overlapping) types of altered monocytes in acute COVID-19, including those with lower
151 antigen presentation, depressed NF- κ B/inflammation, or myeloid-derived suppressor cell (MDSC)-like
152 phenotypes^{13,16,17}. However, none of these monocyte phenotypes were significantly different in the
153 monocytes of COVR compared to HCs in our cohort at baseline before influenza vaccination
154 (Supplementary Fig. 2), suggesting that our depressed monocyte gene signature involving pattern
155 recognition and IIR genes is distinct from those identified earlier in acute disease. Together, our
156 findings suggest that even mild, non-hospitalized SARS-CoV-2 infections may establish new, temporally
157 stable, sex-dependent immunological imprints.

158 To assess whether other natural respiratory viral infections may leave similar unresolved sex-
159 specific “immune states”, we used a published whole blood transcriptomic dataset assessing two
160 independent cohorts of patients with confirmed community influenza A (predominantly pandemic
161 H1N1) infection during two different seasons (2009-2010 and 2010-2011; Extended Data Fig. 2a)¹⁸. By
162 comparing the WBT profiles before and after each season (i.e., before infection and post-recovery), we
163 found robust post-infection changes consistent between these two independent cohorts in males only
164 (the changes in females were not consistent between these two cohorts; Extended Data Fig. 2b;
165 Supplementary Table 6). The genes with increased expression after recovery in males were also
166 enriched for genes more highly expressed in COVR-M compared to COVR-F in our cohort (after
167 accounting for the expected sex differences present in healthy subjects; Extended Data Fig. 2c). In
168 addition, the genes with lower expression after recovery from influenza infection in males were
169 enriched for the depressed IIR signature above, including TLR5 and VCAN (Fig. 1f,g; Supplementary
170 Table 6). These observations provide independent support that exposure to a respiratory viral
171 pathogen can lead to persistent immunological imprints detectable in blood, even in healthy
172 individuals with mild disease. However, different viral infections are also likely to leave pathogen-
173 dependent imprints with distinct genes and processes; for example, the overlapping signals between
174 post-influenza and post-mild COVID-19 are only a small subset of the sex-specific post-COVID-19
175 changes we detected.

176

177 **Contrasting influenza vaccination responses**

178 We next asked whether prior COVID-19 may impact an individual’s response to non-SARS-CoV-2
179 immunological challenges. Study participants received the seasonal influenza quadrivalent vaccine and

180 were followed longitudinally for up to 100 days, including days 1, 7, and 28, to assess the vaccine
181 response at the serological, molecular, and cellular levels (Fig. 1a, 1b, 2a). This vaccine was selected in
182 part due to its public health importance: the 2020-21 influenza season was approaching at the start of
183 our study and it was not clear whether prior COVID-19 infection would impact influenza vaccine
184 responses. In addition, the responses to seasonal influenza vaccination have been well characterized in
185 healthy adults, including early innate/inflammatory and interferon (IFN) responses on day 1 (D1) after
186 vaccination and a strong but transient plasmablast peak around day 7 (D7) culminating in the
187 generation of influenza-specific antibodies^{19,20}. Thus, influenza vaccination provides an excellent
188 perturbation to probe the functional impacts of prior mild SARS-CoV-2 infection.

189 Blood transcriptomic, peripheral immune cell frequency, CITE-seq, influenza-specific B-cell, and
190 antibody titer analyses [assessing responses on D1, D7 and day 28 (D28) relative to day 0 (D0)]
191 together pointed to coordinated, sex-specific innate and adaptive response differences to the vaccine,
192 with COVR-M generally mounting a more potent response than their healthy counterparts and COVR-F
193 (Fig. 2b-i and Extended Data Fig. 3a, 3c-g; Supplementary Tables 7 and 8). These include stronger
194 innate/inflammatory and particularly IFN-related transcriptional responses (Fig. 2b, Extended Data Fig.
195 3a), with corresponding greater increases in circulating IFN γ protein levels in serum by D1 in COVR-M
196 (Fig. 2c). This systemic increase in IFN γ impacts diverse cell types expressing the IFN γ signaling
197 components as revealed by single cell CITE-seq: most peripheral immune cells had higher IFN response
198 signatures on D1 in COVR-M than the other groups (based on comparing D1 vs. D0; Fig. 2d; Fig. 2e
199 shows CD4⁺ T-cells, B-cells, monocytes and cDCs as examples). Baseline, pre-vaccination IFN-related
200 transcriptional activity was largely indistinguishable between COVR and HC (Extended Data Fig. 3b). In
201 addition, a more robust response was observed for antigen presentation genes including both MHC

202 class I and II genes in classical monocytes of COVR-M (Fig. 2f). Thus, COVR-M mount a stronger
203 circulating IFN γ and corresponding transcriptional response in both innate and adaptive immune cells
204 by D1 following influenza vaccination.

205 Based on previous studies of influenza vaccination in healthy adults and because heightened
206 innate immune responses elicited by adjuvants are known to enhance adaptive responses²¹, we
207 hypothesized that the stronger early inflammatory responses in COVR-M would lead to a more robust
208 humoral response. Indeed, we saw increased D7 B-/plasma-cell related transcriptional signatures in the
209 COVR-M (Extended Data Fig. 3a,c). Furthermore, COVR-M had a greater increase of influenza-specific
210 plasmablasts than HC-M at D7 (Fig. 2g, Supplementary Fig. 3). Consistent with previous observations in
211 healthy adults²² and the hypothesis that the stronger early IFN response in COVR-M could help induce
212 a more robust B-cell response, we detected a positive correlation between those two parameters,
213 including the extent of influenza-specific plasmablast increases (Extended Data Fig. 3d). Consistently,
214 COVR-M also had higher influenza-specific antibody responses than HC-M across all but one of the
215 vaccine strains at D28 relative to baseline (Fig. 2h,i; Extended Data Fig. 3e-g;
216 Supplementary Table 8; see Methods). While influenza infection and vaccination history can influence
217 influenza vaccine responses²³, they alone are unlikely to explain the above findings as the COVR and HC
218 groups had similar baseline antibody titers (Extended Data Fig. 3e,f), were age/sex-matched, and
219 drawn from the same geographic region with very low influenza infection/transmission during the
220 2020-21 season⁸. Additionally, the statistical model used to assess titer response differences
221 incorporated pre-vaccination influenza titers as a covariate (see Methods). The extent of time-
222 dependent immune resolution following COVID-19 was unlikely a factor because TSD and D28 titer
223 responses are not correlated in either sex (data not shown). Together, these observations demonstrate

224 that prior mild infection by SARS-CoV-2 can result in sex-dependent, coordinated changes in both
225 innate and adaptive responses to immunization with non-SARS-CoV-2 antigens months after acute
226 disease.

227

228 **Linking baseline to innate response**

229 Having established that prior mild COVID-19 is associated with new baseline immune states prior to
230 influenza vaccination (Fig. 1 and Extended Data Fig. 1) and COVR-M-specific responses following
231 vaccination (Fig. 2 and Extended Data Fig. 3), we next attempted to link the two and asked what
232 baseline variables and cellular circuits may contribute to the heightened IFN-related responses in
233 COVR-M, which could subsequently contribute to the more robust humoral responses in COVR-M (Fig.
234 3a). Using flow cytometry (Supplementary Fig. 1) and CITE-seq data, we first employed a multivariate
235 linear model to identify baseline/pre-vaccination immune cells whose frequency predicted the D1 IFN-
236 related responses (D1 vs. D0 in serum IFN γ protein levels and IFN transcriptional signature score). A
237 subset of CD8⁺ T-cells with an effector memory (CD8 EM) phenotype (CD45RA⁻ CCR7⁻ CD28⁺ CD27⁻;
238 “early effector-like”) was a top candidate in COVR-M and could thus be a cellular source of IFN γ upon
239 vaccination (Extended Data Fig. 4a,b; Supplementary Fig. 4); the same relationship was not observed in
240 HCs (Supplementary Fig. 5a,b).

241 We next focused on all the CD8⁺ T-cells from clusters with an EM phenotype (CD8 EM) in the
242 CITE-seq data based on both surface protein markers and mRNA expression (see Methods and
243 Supplementary Table 10 for the top cluster protein markers). We searched for differences in average
244 surface marker expression of cells in these CD8 EM clusters across the four subject groups and found
245 that GPR56 was the top differentially expressed marker with increased expression in COVR-M relative

246 to the HC-M and COVR-F (Fig. 3b,c; Supplementary Table 10). This was intriguing because CD4+ EM and
247 TEMRA (EM cells re-expressing CD45RA) T-cells marked by surface GPR56 expression at baseline
248 (before stimulation) have been reported to produce increased amounts of IFN γ upon PMA/ionomycin
249 (PMAI) stimulation²⁴. Consistent with this, GPR56+ CD8 EM cells in our data are enriched for a
250 transcriptional signature (derived in an independent study²⁵) that marks CD8 EM cells poised to secrete
251 higher levels of IFN γ upon PMAI stimulation (Fig. 3d). Thus, GPR56+ CD8 EM cells could be a source of
252 elevated IFN γ production in COVR-M following influenza vaccination. Indeed, the frequency of these
253 cells was elevated in COVR-M relative to both HC-M and COVR-F prior to vaccination (Fig. 3e), but not
254 correlated with the TSD and thus temporally stable (assessed by Spearman's correlation: p value = 0.18
255 in COVR-F and p value = 0.51 in COVR-M). Additionally, IFNG transcripts increased significantly in these
256 cells on D1 following influenza vaccination in COVR-M (Extended Data Fig. 4c,d). These data suggest
257 that prior COVID-19 increases the frequency of GPR56+ CD8 EM cells in males and these cells are
258 poised to make more IFN γ early after influenza vaccination, which together contributed to the higher
259 IFN γ production in COVR-M; consistent with this hypothesis, this was not observed in GPR56- cells
260 (Extended Data Fig. 4d; Supplementary Fig. 5c).

261 Mild, non-hospitalized COVID-19 has been reported to induce "bystander activation" (non-
262 SARS-CoV-2 specific) of CD8+ T-cells²⁶. Interestingly, the GPR56+ cells are also enriched for a
263 transcriptional signature associated with bystander T-cell activation^{26,27} (Fig. 3f). In addition, GPR56+
264 CD8 EM cell frequency is positively correlated with the T-cell activation signature score, which was
265 elevated at baseline in COVR-M as shown above (Fig. 1i, Extended Data Fig. 4e). This suggests that
266 some of these cells may have expanded in a bystander manner during the acute phase of the infection.
267 This prompted us to consider whether these GPR56+ cells are similar to bystander-activated virtual

268 memory (VM) CD8+ T-cells, a feature of which is their ability to be activated rapidly by inflammatory
269 cytokines alone (e.g., IL-12, IL-18, and IL-15) to produce IFN γ without T-cell receptor (TCR)
270 stimulation^{28,29}. VM CD8+ T-cells expand via cytokine stimulation, including IL-15 induced by viral
271 infection (IL-15 concentrations are known to be elevated in acute COVID-19 patients and correlate with
272 disease severity³⁰), and are characterized by a differentiated EM phenotype expressing CD45RA²⁸. We
273 assessed several reported surface markers of these cells²⁸ in GPR56+ vs. GPR56- cells and found that
274 the GPR56+ cells were indeed phenotypically similar to VM cells (Fig. 3g). For example, GPR56+ cells
275 have higher CD122 but lower CD5 surface expression than their GPR56- counterparts; the latter of
276 which has been linked to the extent of prior IL-15 (or potentially other inflammatory cytokine)
277 encounters^{28,31}. Interestingly, based on the surface levels of CD45RA and CD45RO, the GPR56+ cells
278 appear to situate phenotypically between GPR56- and TEMRA cells (Extended Data Fig. 4f).

279 To further test our hypothesis, we performed *in vitro* stimulation experiments to assess
280 whether GPR56+ CD8+ T-cells can produce IFN γ in response to several cytokines known to be induced
281 by vaccination or infection (Supplementary Fig. 6a). Stimulation with IL-15 showed that GPR56+
282 CD45RA+ CD8+ T-cells from COVR-M produced more IFN γ compared to COVR-F (Fig. 3h,i). CD8+ VM-
283 like T-cells were identified using surface markers CD45RA+, KIR+ and/or NKG2A+^{32,33} and COVR-M
284 produced higher IFN γ in these cells (Fig. 3i). Stimulation with IL-12, IL-15, and IL-18 together showed
285 similar trends (Supplementary Fig. 6b). Stimulation with IL-18 alone or IL-12 and IL-18 together also
286 showed similar trends, but these conditions induced less robust IFN γ than IL-15 stimulation (data not
287 shown). We next assessed the cellular source of IL-15 post influenza vaccination using CITE-seq data
288 and found that classical monocytes from COVR-M showed the most significant increases in IL-15 mRNA
289 levels on day 1 after influenza vaccination (Fig. 3j). Together this suggests that the increased IFN γ

290 response in COVR-M following vaccination could be attributed to increased baseline (pre-vaccination)
291 frequencies in cells that are also intrinsically more responsive to inflammatory stimulation, including
292 classical monocytes that produce elevated IL-15 and CD8 VM-like T-cells that mount a more robust
293 IFN γ response to cytokine stimulation alone.

294 Since VM T-cells can be rapidly activated to produce cytokines without clonal, antigen-specific
295 expansion²⁸, we assessed the clonality of the GPR56+ CD8 EM cells at different timepoints after
296 influenza vaccination using V(D)J/TCR data from CITE-seq. The clonality of both the GPR56+ CD8 EM
297 and TEMRA cells remained stable across days 0 (before vaccination), 1 and 28 following influenza
298 vaccination (Extended Data Fig. 4g,h). The frequencies of GPR56+ CD8 EM clones shared across
299 timepoints within individuals were also similar (Extended Data Fig. 4i). Together, these data argue
300 against the notion that the heightened activation of the GPR56+ cells early after influenza vaccination
301 in COVR-M was due solely to TCR-dependent T-cell activation and clonal expansion. As was shown
302 previously^{28,29} and above in our *in vitro* stimulation data, a more plausible explanation is that these
303 CD8+ VM-like cells were activated to produce IFN γ by the inflammatory cytokines elicited by the
304 influenza vaccine in an antigen-independent manner. Despite their resemblance to VM cells, some of
305 the GPR56+ cells could have developed from naïve cells via conventional, non-bystander pathways
306 (e.g., some could be developed during acute COVID-19 and are specific for SARS-CoV-2), although none
307 of these cells had a CDR3 sequence that matches a public clone deemed to be specific for SARS-CoV-2
308 (data not shown). Bona fide, antigen-specific memory CD8+ T-cells developed from naïve cells via TCR
309 stimulation have also been shown to produce IFN γ in response to inflammatory cytokines alone in
310 mice^{34,35}.

311 Our data also revealed other cell types that could have contributed to the increased IFN γ
312 production observed on D1 after vaccination in COVR-M (Supplementary Fig. 7a-c). IFNG transcript
313 increased more in COVR-M than HC-M and COVR-F on day 1 in CD16^{lo} NK cells (Supplementary Fig. 7c,
314 Supplementary Table 4). Additionally, the baseline frequency of CD16^{lo} NK cells was correlated with the
315 extent of D1 increase in both IFNG expression and serum protein levels (Supplementary Fig. 7b).
316 However, the IFN γ response in total NK cells after IL-15 stimulation *in vitro* was not significantly higher
317 in COVR-M (Extended Data Fig. 4j), likely because CD16^{lo} NK cells are a small subset of total NK cells. In
318 contrast, IL-15 stimulation *in vitro* revealed a higher IFN γ response in MAIT cells in COVR-M than both
319 COVR-F and HC-M (Extended Data Fig. 4j), but the IFNG mRNA expression increase on D1 post
320 influenza vaccination was not statistically significant in COVR-M based on CITE-seq data
321 (Supplementary Fig. 7c). CD8⁺ T-cells with a TEMRA (CD45RA⁺ CD45RO⁻ CCR7⁻) phenotype might also
322 play a role as their IFN γ response after IL-15 stimulation *in vitro* was higher in COVR-M than both
323 COVR-F and HC-M (Extended Data Fig. 4j), which is consistent with CITE-seq data (Supplementary Fig.
324 7c).

325 Taken together, we demonstrate a population of CD8 EM T-cells marked by GPR56 expression
326 and VM-like markers with antigen-agnostic pro-inflammatory potential after heterologous vaccination.
327 Importantly, these cells, and potentially CD16^{lo} NK, MAIT, and CD8⁺ TEMRA cells (albeit with less
328 support from our CITE-seq data), emerged in otherwise clinically healthy individuals and are especially
329 elevated and more poised to respond in males who were months recovered from mild SARS-CoV-2
330 infection, providing additional evidence for sex-specific, functionally relevant immune set points linked
331 to prior mild COVID-19.

332

333 **Vaccination shifts monocyte imprints**

334 Given the potential for vaccine-induced “training” effects^{6,36,37}, we next asked whether influenza
335 vaccination can alter some of the post-COVID-19 transcriptional imprints we detected earlier (Fig. 4a).
336 We focused on the monocytes because of the robustly depressed IIR signature reported above (in
337 COVR vs. HC; Fig. 1f,g) and because vaccines can potentially induce long-lasting changes in these
338 cells^{6,36}. By using the HC baseline (D0) as a healthy reference, we used CITE-seq data to assess the
339 average expression of the signature genes (identified above) before and after vaccination in COVR
340 subjects, separately for classical (Fig. 1f) and non-classical monocytes (Fig. 1g) in males and females
341 (Extended Data Fig. 5a,b). As was observed above, these genes had lower average expression in COVR
342 than HC in both sexes at D0 before vaccination. However, their average expression increased towards
343 that of the HCs by D1 and persisted through D28 in both COVR-F and COVR-M, although the effect
344 appeared stronger in COVR-F (Extended Data Fig. 5a,b).

345 Quantifying the average expression (module score) of these sex- and cell type-dependent gene
346 sets (Fig. 1f,g) within individual subjects over time confirmed a similar and significant trend of shift
347 towards the HCs (Fig. 4b,c). This analysis further revealed that the extent of this change in gene
348 expression was more pronounced in the non-classical than the classical monocytes (Fig. 4b,c).
349 Intriguingly, the behavior of these genes was divergent in the HCs: the gene module score trended
350 lower on D1 and reverted to pre-vaccination levels by day 28 in HCs (Fig. 4b,c). While the underlying
351 mechanism of this divergence is unclear, the monocytes in HCs could have responded to the vaccine-
352 induced inflammation by downregulating certain immune receptor and associated signaling genes in a
353 negative feedforward mechanism to avoid over responding, while the “depressed” monocytes in COVR

354 instead responded by increasing the expression of these genes and thus moving towards the normal
355 (healthy baseline) level.

356 We next identified the individual genes within these gene sets that moved towards the HC
357 baseline (see Methods). In both classical and non-classical monocytes, the fraction of reverting genes
358 was significantly higher in females than males (Fig. 4d,e; Extended Data Fig. 5c), although several TLRs
359 (e.g., TLR2, TLR4) and NOD2 were significant in both sexes in one or both monocyte subsets. These
360 changes were unlikely due to continued immune resolution following infection because the baseline
361 (D0) expression of these genes did not correlate with TSD (Extended Data Fig. 1h), and they increased
362 acutely by D1 following vaccination and persisted to D28. Interestingly, unlike this depressed IIR
363 signature (Fig. 1f,g, Extended Data Fig. 1i), other monocyte-related transcriptional signatures known to
364 have lower expression during acute COVID-19, such as genes related to antigen presentation,
365 inflammatory and NF- κ B activation, and myeloid suppressor cells^{13,15-17,38,39}, were similar between
366 COVR and HC at D0/baseline; vaccination also did not consistently elicit longer-lasting changes in these
367 signatures out to D28, although COVR-M tended to have elevated antigen presentation transcriptional
368 responses in non-classical monocytes on D1 that remained mildly elevated by D28 (Extended Data Fig.
369 5d,e).

370 Together, CITE-seq analysis revealed that the early (D1) response to influenza vaccination
371 elevates a set of previously (i.e., before vaccination) depressed IIR genes in the monocytes of COVR
372 subjects out to at least D28 post vaccination. Although the functional relevance of these changes
373 remains to be determined, these results suggest that the early inflammatory responses to influenza
374 vaccination can help to shift the post-COVID-19 immune state of monocytes towards that of healthy,
375 particularly in female recoverees.

376

377 **Discussion**

378 While both acute and longer-term immune perturbations in hospitalized COVID-19 patients have been
379 reported^{13,40–43}, less is known regarding healthy recovered individuals with prior mild, non-hospitalized
380 SARS-CoV-2 infection months after acute illness, without confounding comorbidities such as obesity,
381 autoimmunity, or immunodeficiency. Here we reveal that clinically healthy recoverees of prior non-
382 hospitalized COVID-19 possess sex-specific immune imprints beyond SARS-CoV-2 specific immunity,
383 some of which only become apparent after vaccination with antigens distinct from SARS-CoV-2. Our
384 findings are consistent with the sex dimorphic nature of acute responses to SARS-CoV-2 and other
385 immune challenges¹¹. Healthy females tend to mount heightened inflammatory responses to infections
386 and vaccines⁴⁴; it was therefore surprising to find the qualitative opposite here in which COVR-M were
387 found to have a more “poised” immune status at baseline and stronger innate and adaptive responses
388 to influenza vaccination. While persistent immune state changes (over months) in patients with “long
389 COVID” have been reported⁴¹, most of the individuals in our study reported no or minor post-COVID-19
390 sequelae. Future work could assess whether some of the sex-specific imprints, including differences in
391 vaccination responses, are associated with “long COVID”⁷.

392 Our findings suggest that the “poised” baseline immune states in COVR-M helped establish the
393 more robust IFN, plasmablast, and antibody responses on days 1, 7, and 28, respectively, following
394 influenza vaccination. The early IFN responses may be attributed to monocytes with higher IL-15
395 responses early after vaccination coupled with elevated pre-vaccination frequencies of “virtual
396 memory” like CD8+ T-cells poised to produce more IFN γ upon IL-15 stimulation. The monocyte imprint
397 we described involving poised IL-15 mRNA production in male recoverees and the transcriptionally

398 depressed innate receptor gene signature in both sexes are consistent with the notion of trained
399 innate immunity⁶. Interestingly, while the latter signature could be detected in acute COVID-19
400 patients with severe disease, it is distinct from the depressed antigen presentation or myeloid
401 suppressor cell like states found in earlier studies^{13,15-17,38,39}. As trained innate immunity can be
402 mediated through myriad mechanisms including chromatin and metabolic changes within cells, future
403 studies could explore these potential mechanisms in monocytes, including the influences of
404 sex/gender, acute disease severity, and age among subjects with a range of post-COVID clinical
405 sequelae. Given that the half-life of circulating monocytes is relatively short (and can be shorter than
406 28 days)⁴⁵, the partial reversal we detected is possibly attributable to bone marrow myeloid progenitor
407 cells, as hematopoietic stem and progenitor cells have been shown to exhibit chromatin accessibility
408 changes following SARS-CoV2 infection⁴⁶.

409 Bystander T-cell activation has been reported following natural viral infections⁴⁷, including
410 SARS-CoV-2²⁶. More recently, bystander activated CD8 EM T-cells have been identified as playing an
411 important role in controlling early infection, including VM cells that have no prior antigen exposure or
412 TCR engagement^{28,29}. As these cells can emerge following cytokine stimulation alone, it is possible that
413 a stronger or more prolonged cytokine response to SARS-CoV-2 in males relative to females during
414 acute disease may have resulted in the elevated frequencies of the GPR56+ CD8+ VM-like cells in
415 COVR-M. This hypothesis is consistent with reports that males hospitalized with COVID-19 tend to
416 experience greater innate immune activation (as measured by circulating cytokines) compared to
417 females^{48,49}.

418 Some of the immune imprints we observed could be shared among different types of viral
419 infections, but some are likely unique to SARS-CoV-2, as suggested by our comparison with natural

420 influenza infection. Our findings point to the possibility that any infection or immune challenge may
421 change the immune status to establish new baseline set points encoded by the states of not only a
422 single cell lineage, but a network of interacting cell types such as VM T-cells and monocytes. In
423 addition, while baseline immune statuses predictive of future responses are often different across and
424 temporally stable within individuals over a timescale of months^{50,51}, our results suggest that such
425 baseline immune states could have been established by past infections and are stable up to the next
426 perturbation. Thus, the baseline immune status of an individual, with the potential to impact future
427 responses in both antigen-specific and -agnostic ways, is shaped by a multitude of prior exposures^{2,3}. In
428 addition to revealing underlying principles regarding what happens after two well-defined natural
429 immunological encounters: mild COVID-19 and influenza vaccination in humans, our observations
430 provide a basis for studying more complex scenarios, such as what happens over longer timescales
431 with additional encounters. Our work brings forth the concept that even mild viral infections could
432 establish new immunological set-points impacting future immune responses in an antigen-agnostic
433 manner and illustrates how heterologous vaccination could be used as a tool to reveal such functional
434 imprints.

435 **Limitations of this study** and additional discussion can be found in the Supplementary
436 Information.

437

438 **Main Text References**

439

- 440 1. Mina, M. J. *et al.* Measles virus infection diminishes preexisting antibodies that offer protection
441 from other pathogens. *Science* **366**, 599–606 (2019).
- 442 2. Tomalka, J. A., Suthar, M. S., Diamond, M. S. & Sekaly, R. P. Innate antiviral immunity: how prior
443 exposures can guide future responses. *Trends Immunol.* **43**, 696–705 (2022).
- 444 3. Tsang, J. S. *et al.* Improving Vaccine-Induced Immunity: Can Baseline Predict Outcome? *Trends*
445 *Immunol.* **41**, 457–465 (2020).
- 446 4. Goodridge, H. S. *et al.* Harnessing the beneficial heterologous effects of vaccination. *Nat. Rev.*
447 *Immunol.* **16**, 392–400 (2016).
- 448 5. Aaby, P., Netea, M. G. & Benn, C. S. Beneficial non-specific effects of live vaccines against COVID-19
449 and other unrelated infections. *Lancet Infect. Dis.* S1473-3099(22)00498-4 (2022)
450 doi:10.1016/S1473-3099(22)00498-4.
- 451 6. Netea, M. G. *et al.* Defining trained immunity and its role in health and disease. *Nat. Rev. Immunol.*
452 **20**, 375–388 (2020).
- 453 7. Nalbandian, A. *et al.* Post-acute COVID-19 syndrome. *Nat. Med.* **27**, 601–615 (2021).
- 454 8. Olsen, S. J. *et al.* Changes in Influenza and Other Respiratory Virus Activity During the COVID-19
455 Pandemic - United States, 2020-2021. *MMWR Morb. Mortal. Wkly. Rep.* **70**, 1013–1019 (2021).
- 456 9. Sudre, C. H. *et al.* Attributes and predictors of long COVID. *Nat. Med.* **27**, 626–631 (2021).
- 457 10. Stoeckius, M. *et al.* Simultaneous epitope and transcriptome measurement in single cells. *Nat.*
458 *Methods* **14**, 865–868 (2017).
- 459 11. Ursin, R. L. & Klein, S. L. Sex Differences in Respiratory Viral Pathogenesis and Treatments. *Annu.*
460 *Rev. Virol.* **8**, 393–414 (2021).

- 461 12. Wheatley, A. K. *et al.* Evolution of immune responses to SARS-CoV-2 in mild-moderate COVID-19.
462 *Nat. Commun.* **12**, 1162 (2021).
- 463 13. Schultze, J. L. & Aschenbrenner, A. C. COVID-19 and the human innate immune system. *Cell* **184**,
464 1671–1692 (2021).
- 465 14. Sette, A. & Crotty, S. Adaptive immunity to SARS-CoV-2 and COVID-19. *Cell* **184**, 861–880 (2021).
- 466 15. Liu, C. *et al.* Time-resolved systems immunology reveals a late juncture linked to fatal COVID-19.
467 *Cell* **184**, 1836-1857.e22 (2021).
- 468 16. Paludan, S. R. & Mogensen, T. H. Innate immunological pathways in COVID-19 pathogenesis. *Sci.*
469 *Immunol.* (2022) doi:10.1126/sciimmunol.abm5505.
- 470 17. Reyes, M. *et al.* Plasma from patients with bacterial sepsis or severe COVID-19 induces suppressive
471 myeloid cell production from hematopoietic progenitors in vitro. *Sci. Transl. Med.* (2021)
472 doi:10.1126/scitranslmed.abe9599.
- 473 18. Zhai, Y. *et al.* Host Transcriptional Response to Influenza and Other Acute Respiratory Viral
474 Infections – A Prospective Cohort Study. *PLOS Pathog.* **11**, e1004869 (2015).
- 475 19. Pulendran, B. Systems vaccinology: Probing humanity’s diverse immune systems with vaccines.
476 *Proc. Natl. Acad. Sci.* **111**, 12300–12306 (2014).
- 477 20. Tsang, J. S. Utilizing population variation, vaccination, and systems biology to study human
478 immunology. *Trends Immunol.* **36**, 479–493 (2015).
- 479 21. Pulendran, B., S. Arunachalam, P. & O’Hagan, D. T. Emerging concepts in the science of vaccine
480 adjuvants. *Nat. Rev. Drug Discov.* **20**, 454–475 (2021).
- 481 22. Bucasas, K. L. *et al.* Early Patterns of Gene Expression Correlate With the Humoral Immune
482 Response to Influenza Vaccination in Humans. *J. Infect. Dis.* **203**, 921–929 (2011).

- 483 23. Auladell, M. *et al.* Influenza virus infection history shapes antibody responses to influenza
484 vaccination. *Nat. Med.* **28**, 363–372 (2022).
- 485 24. Truong, K.-L. *et al.* Killer-like receptors and GPR56 progressive expression defines cytokine
486 production of human CD4+ memory T cells. *Nat. Commun.* **10**, 2263 (2019).
- 487 25. Nicolet, B. P. *et al.* CD29 identifies IFN- γ -producing human CD8+ T cells with an increased cytotoxic
488 potential. *Proc. Natl. Acad. Sci.* **117**, 6686–6696 (2020).
- 489 26. Bergamaschi, L. *et al.* Longitudinal analysis reveals that delayed bystander CD8+ T cell activation
490 and early immune pathology distinguish severe COVID-19 from mild disease. *Immunity* **54**, 1257-
491 1275.e8 (2021).
- 492 27. Bangs, S. C. *et al.* Human CD4+ Memory T Cells Are Preferential Targets for Bystander Activation
493 and Apoptosis. *J. Immunol.* **182**, 1962–1971 (2009).
- 494 28. White, J. T., Cross, E. W. & Kedl, R. M. Antigen-inexperienced memory CD8+ T cells: where they
495 come from and why we need them. *Nat. Rev. Immunol.* **17**, 391–400 (2017).
- 496 29. Maurice, N. J., Taber, A. K. & Prlic, M. The Ugly Duckling Turned to Swan: A Change in Perception of
497 Bystander-Activated Memory CD8 T Cells. *J. Immunol.* **206**, 455–462 (2021).
- 498 30. Abers, M. S. *et al.* An immune-based biomarker signature is associated with mortality in COVID-19
499 patients. *JCI Insight* **6**, 144455 (2021).
- 500 31. Herndler-Brandstetter, D. *et al.* Post-thymic regulation of CD5 levels in human memory T cells is
501 inversely associated with the strength of responsiveness to interleukin-15. *Hum. Immunol.* **72**, 627–
502 631 (2011).
- 503 32. Jacomet, F. *et al.* Evidence for eomesodermin-expressing innate-like CD8+ KIR/NKG2A+ T cells in
504 human adults and cord blood samples. *Eur. J. Immunol.* **45**, 1926–1933 (2015).

- 505 33. Jin, J.-H. *et al.* Virtual memory CD8⁺ T cells restrain the viral reservoir in HIV-1-infected patients
506 with antiretroviral therapy through derepressing KIR-mediated inhibition. *Cell. Mol. Immunol.* **17**,
507 1257–1265 (2020).
- 508 34. Yang, J., Zhu, H., Murphy, T. L., Ouyang, W. & Murphy, K. M. IL-18–stimulated GADD45 β required in
509 cytokine-induced, but not TCR-induced, IFN- γ production. *Nat. Immunol.* **2**, 157–164 (2001).
- 510 35. Kastenmüller, W., Torabi-Parizi, P., Subramanian, N., Lämmermann, T. & Germain, R. N. A spatially-
511 organized multicellular innate immune response in lymph nodes limits systemic pathogen spread.
512 *Cell* **150**, 1235–1248 (2012).
- 513 36. Wimmers, F. *et al.* The single-cell epigenomic and transcriptional landscape of immunity to
514 influenza vaccination. *Cell* **184**, 3915–3935.e21 (2021).
- 515 37. Debisarun, P. A. *et al.* Induction of trained immunity by influenza vaccination - impact on COVID-19.
516 *PLoS Pathog.* **17**, e1009928 (2021).
- 517 38. Arunachalam, P. S. *et al.* Systems biological assessment of immunity to mild versus severe COVID-
518 19 infection in humans. *Science* **369**, 1210–1220 (2020).
- 519 39. Schulte-Schrepping, J. *et al.* Severe COVID-19 Is Marked by a Dysregulated Myeloid Cell
520 Compartment. *Cell* **182**, 1419–1440.e23 (2020).
- 521 40. Pérez-Gómez, A. *et al.* Dendritic cell deficiencies persist seven months after SARS-CoV-2 infection.
522 *Cell. Mol. Immunol.* **18**, 2128–2139 (2021).
- 523 41. Phetsouphanh, C. *et al.* Immunological dysfunction persists for 8 months following initial mild-to-
524 moderate SARS-CoV-2 infection. *Nat. Immunol.* 1–7 (2022) doi:10.1038/s41590-021-01113-x.
- 525 42. Utrero-Rico, A. *et al.* Alterations in Circulating Monocytes Predict COVID-19 Severity and Include
526 Chromatin Modifications Still Detectable Six Months after Recovery. *Biomedicines* **9**, 1253 (2021).

- 527 43. You, M. *et al.* Single-cell epigenomic landscape of peripheral immune cells reveals establishment of
528 trained immunity in individuals convalescing from COVID-19. *Nat. Cell Biol.* **23**, 620–630 (2021).
- 529 44. Klein, S. L. & Flanagan, K. L. Sex differences in immune responses. *Nat. Rev. Immunol.* **16**, 626–638
530 (2016).
- 531 45. Patel, A. A. *et al.* The fate and lifespan of human monocyte subsets in steady state and systemic
532 inflammation. *J. Exp. Med.* **214**, 1913–1923 (2017).
- 533 46. Cheong, J.-G. *et al.* Epigenetic Memory of COVID-19 in Innate Immune Cells and Their Progenitors.
534 2022.02.09.479588 Preprint at <https://doi.org/10.1101/2022.02.09.479588> (2022).
- 535 47. Tough, D. F., Borrow, P. & Sprent, J. Induction of Bystander T Cell Proliferation by Viruses and Type
536 I Interferon in Vivo. *Science* **272**, 1947–1950 (1996).
- 537 48. Takahashi, T. *et al.* Sex differences in immune responses that underlie COVID-19 disease outcomes.
538 *Nature* **588**, 315–320 (2020).
- 539 49. Scully, E. P. *et al.* Sex and Gender Differences in Testing, Hospital Admission, Clinical Presentation,
540 and Drivers of Severe Outcomes From COVID-19. *Open Forum Infect. Dis.* **8**, ofab448 (2021).
- 541 50. Tsang, J. S. *et al.* Global Analyses of Human Immune Variation Reveal Baseline Predictors of
542 Postvaccination Responses. *Cell* **157**, 499–513 (2014).
- 543 51. Kotliarov, Y. *et al.* Broad immune activation underlies shared set point signatures for vaccine
544 responsiveness in healthy individuals and disease activity in patients with lupus. *Nat. Med.* **26**,
545 618–629 (2020).
- 546
547
548

549

550 **Figure 1. Study overview and baseline differences.**

551 **a**, Schematic showing the study concept and design.

552 **b**, Data generated in the study. Both COVID-19-recovered (COVR) subjects and healthy controls (HC)
553 were enrolled at seven days before vaccination (Day -7) and sampled at the indicated timepoints
554 relative to the day of influenza vaccination. The number of subjects assayed for each data type is
555 indicated. CBC with diff = complete blood count with differential; TBNK = T- and B-lymphocyte and
556 Natural Killer cell phenotyping; SPR = Surface plasmon resonance.

557 **c**, Bar plots comparing the proportion of CD11c+ dendritic cells (DCs; as the fraction of live cells from
558 flow cytometry) between COVR females (COVR-F; n=15), HC females (HC-F; n=16), COVR males (COVR-
559 M; n=12), and HC males (HC-M; n=11) at day 0 (D0). The statistical significance is determined by two-
560 tailed Wilcoxon test. Error bars indicate the standard error of each group.

561 **d**, Similar to (c) but for monocytes (from CBC; y-axis) between COVR-F (n=17), COVR-M (n=16), HC-F
562 (n=21), and HC-M (n=19) at baseline (average of Day -7 and D0).

563 **e**, UMAP of the CITE-seq single cell data showing clustering of cells based on the expression of cell
564 surface protein markers (632,100 single cells from all timepoints with CITE-seq data: days 0, 1, 28).
565 Colored and boxed cell clusters are further explored in (f-i).

566 **f**, (left) Box plots comparing the innate immune receptor signature scores (see Methods) between HC-F
567 (n=8) and COVR-F (n=12) (left box) and HC-M (n=8) and COVR-M (n=12) (right box) using the CITE-seq
568 classical monocyte pseudobulk expression data at D0. Each point represents a subject. (right) Bubble
569 plot showing the average gene expression of selected genes, including those in the Gene Ontology
570 (GO) “pattern recognition receptor activity” and “immune receptor activity” gene sets.

571 **g**, Similar to (f) but showing the non-classical monocyte population at D0.

572 **h**, Similar to (f) but showing the T-cell activation (BTM-M7.3) module scores of CD8+ central memory T
573 cells at D0. Bubble plot showing the average gene expression of the selected genes shared by male and
574 female from the gene set enrichment analysis (see Methods).

575 **i**, Similar to (h) but showing the CD8+ T-cell effector memory population at D0.

576 All box plots show the median, first and third quantiles (lower and upper hinges) and smallest [lower
577 hinge - 1.5× interquartile range (IQR)] and largest (upper hinge + 1.5× IQR) values (lower and upper

578 whiskers). Unless otherwise noted, statistical significance of difference between groups is determined
579 by two-tailed Wilcoxon test.

ACCELERATED ARTICLE PREVIEW

580 **Figure 2. Sex-specific response differences to influenza vaccination in COVID-19-recovered**
581 **individuals and matching controls.**

582 **a**, Schematic of the sex-specific comparisons of vaccine induced changes from baseline at timepoints
583 post vaccination (D1, D7, and D28) between COVR and HC subjects. Analyses applied to subjects under
584 65 years of age (see Methods).

585 **b**, Box plots of the D1 Interferon Gamma (IFN γ) transcriptional response score (D1 – D0, computed
586 using genes from the Hallmark “Interferon Gamma Response” gene set) for COVR-F (n=15), COVR-M
587 (n=14), HC-F (n=16), and HC-M (n=14).

588 **c**, Box plots of the D1 response (D1 – D0) of serum IFN γ protein level for the subjects shown in **(b)**.

589 **d**, Surface protein expression-based UMAP (as in Fig. 1e) with cells colored by the D1 IFN γ
590 transcriptional response score (D1 – D0; see **(b)** for the gene set used) within each cell subset for HC-F
591 (n=8), COVR-F (n=12), and HC-M (n=8), COVR-M (n=12). Darker color indicates a greater difference
592 between D1 and D0 for the indicated cell subset.

593 **e**, Similar to **(b)**, but for the indicated cell subsets (computed using the CITE-seq pseudobulk mRNA
594 expression data for the cell subset) in HC-F (n=8), COVR-F (n=12), HC-M (n=8) and COVR-M (n=12). cDC
595 = conventional/myeloid dendritic cells.

596 **f**, (left) Box plot showing the D1 transcriptional response score (D1 – D0) of the antigen presentation
597 related genes in classical monocytes for the same subjects in **(e)** (see Methods). (right) Bubble plot
598 showing the averaged expression of individual leading-edge genes (LEGs) from the antigen
599 presentation genes (see Methods) in classical monocytes.

600 **g**, Influenza-specific plasmablast (PB; All HA⁺ CD27⁺CD38⁺CD20^{low}CD21^{low}; see Methods and
601 Supplementary Fig. 3) frequencies at D7 and D0, plotted separately for COVR-F (n=14), HC-F (n=15),
602 COVR-M (n=11), and HC-M (n=9). Lines connect data points from the same subject at D0 and D7.

603 **h**, Analysis of the D28/D0 microneutralization titer fold-change (FC) for each of the four strains in the
604 seasonal influenza vaccine (columns) in COVR-F and HC-F. Each dot represents one individual. The
605 orange and grey lines indicate the average fold change for the HC-F and COVR-F, respectively.
606 Unadjusted p values are derived from generalized linear models accounting for age, race, influenza
607 vaccination history and baseline influenza titers (see Methods).

608 **i**, Similar to **(h)**, but for COVR-M and HC-M.

609 All box plots show the median, first and third quantiles (lower and upper hinges) and smallest (lower
610 hinge – 1.5× interquartile range (IQR)) and largest (upper hinge + 1.5× IQR) values (lower and upper
611 whiskers). Unadjusted p values are shown. Unless otherwise noted, statistical significance of difference
612 between groups is determined by two-tailed Wilcoxon test.

613

614

ACCELERATED ARTICLE PREVIEW

615 **Figure 3. Contributors to increased day 1 IFN γ responses in COVID-19-recovered males.**

616 **a**, Schematic illustrating the study questions regarding why COVR-M had elevated early IFN γ responses.

617 **b**, Box plots comparing the sample means of GPR56 surface expression in CD8⁺ effector memory T-cells
618 (CD8⁺ EM) at D0 for COVR-F (n=12), HC-F (n=8), COVR-M (n=12), and HC-M (n=8).

619 **c**, UMAP of the D0 surface GPR56 protein expression on CD8⁺ EM from all 40 subjects with CITE-seq
620 data. UMAP was derived using the top 60 variable surface proteins within the CD8⁺ EM cells (see
621 Methods).

622 **d**, (top) Same UMAP as (c) but showing the D0 gene-expression signature score computed using genes
623 associated with CD29^{hi} CD8⁺ T-cells identified earlier in an independent study (Nicolet *et al*²³, see
624 Methods). (bottom) Density plot showing the distribution of signature score above in the GPR56⁺ and
625 GPR56⁻ CD8⁺ EM. Dashed line indicates the median of the distribution. Significance of the difference
626 between the medians is determined by two-tailed Wilcoxon test at single-cell level.

627 **e**, Bar plots comparing the proportion of GPR56⁺ cells (as fractions of CD8⁺ EM in the CITE-seq data)
628 between the same subjects as in (b) at D0. Significance is determined by two-tailed Wilcoxon test.
629 Error bars indicate the standard error of each group.

630 **f**, Similar to (d) but showing the bystander T-cell signature score at baseline (D0) (signature genes
631 originated from Bangs *et al*²⁷ and Bergamaschi *et al*²⁶, see Methods).

632 **g**, Box plots comparing the average expression of the indicated cell surface protein markers for the
633 GPR56⁺ versus GPR56⁻ CD8⁺ EM at D0 for the same subjects as in (c). Each point represents a subject.

634 **h**, Representative flow-cytometry contour plots of IFN γ ⁺ and TNF α ⁺ gates within GPR56⁺ CD45RA⁺ CD8⁺
635 T-cells after IL-15 stimulation *in vitro* in the indicated groups. The number shown for each gate denotes
636 the percent of parent (i.e., GPR56⁺ CD45RA⁺ CD8⁺ T-cells).

637 **i**, Boxplots showing the frequencies of IFN γ ⁺ GPR56⁺ CD45RA⁺ VM-like CD8⁺ T-cells (left, as fractions
638 of CD8⁺ T-cells) and IFN γ ⁺ KIR/NKG2A⁺ CD45RA⁺ CD8⁺ T-cells (right, as fractions of CD8⁺ T-cells) in the
639 same subjects as in (b) after IL-15 stimulation *in vitro*.

640 **j**, Box plots comparing D0 and D1 pseudobulk IL-15 mRNA expression (y-axis) in classical monocytes for
641 the same subjects as in (b). Significance is determined by a linear model accounting for age, race, and
642 influenza vaccination history (see Methods).

643 All box plots show the median, first and third quantiles (lower and upper hinges) and smallest (lower
644 hinge – 1.5× interquartile range (IQR)) and largest (upper hinge + 1.5× IQR) values (lower and upper
645 whiskers). Unless otherwise noted, statistical significance of difference between groups is determined
646 by two-tailed Wilcoxon test.

647

648

649

ACCELERATED ARTICLE PREVIEW

650 **Figure 4. Post mild COVID-19 gene expression imprints in monocytes shifted by influenza vaccination.**

651 **a**, Schematic showing the study questions.

652 **b**, Box plots showing the module scores of the innate immune receptor (IIR) signature (see Fig. 1f) in
653 HC-F (n=8), HC-M (n=8), COVR-F (n=12) and COVR-M (n=12) at D0, D1 and D28 using the CITE-seq
654 pseudobulk gene expression data in classical monocytes. The dashed line represents the median D0
655 score of the HCs of the same sex. Lines connect data points from the same subject at different
656 timepoints. Statistical significance of differences is determined by a mixed-effects model accounting
657 for age, race, and influenza vaccination history (see Methods). Unadjusted p values are shown.

658 **c**, Similar to **(b)** but for non-classical monocytes (see Fig. 1g).

659 **d**, Heatmap showing the expression of the “reversal” genes in classical monocytes (row-standardized;
660 see Extended Data Fig. 5c for non-classical monocytes). Reversal genes are defined as those genes in
661 the baseline IIR signature (see also Fig. 1f) whose expression in COVR subjects at D1 and D28 after
662 vaccination moved towards the baseline (pre-vaccination) expression of HCs. COVR-F (top) and COVR-
663 M (bottom) shown separately; HC are also included for comparison. The rows are genes and columns
664 are individual samples (grouped by subject/timepoint) with timepoint and subject group labels shown
665 at the top, including the same subjects as in **(b)** at each timepoint. The names of genes that belong to
666 gene sets of functional interest are shown (FDR-corrected enrichment p values are shown).

667 **e**, Comparison of the proportion of IIR signature genes (see Fig. 1f,g) that show partial reversal in
668 COVR-F versus COVR-M in classical and non-classical monocytes. The mean and 95% confidence
669 intervals (denoted by the bars) are derived from a bootstrapping procedure (see Methods).

670 Significance is determined by the two-tailed Wilcoxon test between the bootstrapped samples.

671 All box plots show the median, first and third quantiles (lower and upper hinges) and smallest (lower
672 hinge – 1.5× interquartile range (IQR)) and largest (upper hinge + 1.5× IQR) values (lower and upper
673 whiskers).

674

675

676 **Methods**

677

678 ***Patient population and sample collection***

679 Subjects at least 18 years of age were recruited between August and December 2020 from the local
680 area (Maryland, Virginia, and the District of Columbia) and enrolled on National Institutes of Health
681 (NIH) protocol 19-I-0126 (Systems analyses of the immune response to the seasonal influenza vaccine).
682 The study was approved by the NIH Institutional Review Board (ClinicalTrials.gov ID: NCT04025580)
683 and complied with all relevant ethical regulations. Informed consent was obtained from all
684 participants. After informed consent, a baseline history and physical examination were performed.
685 Subjects were asked to characterize any present, persistent symptoms of past SARS-CoV-2 infection.
686 Exclusion criteria included obesity (BMI \geq 30); history of or suspicion of any autoimmune,
687 autoinflammatory or immunodeficiency disease; history of any vaccine within the past 30 days (live
688 attenuated) or 14 days (non-live attenuated); history of any experimental vaccine; history of a
689 parasitic, amebic, fungal, or mycobacterial infection in the past year; or current infection. The COVID-
690 19 vaccine was not available at the time of the study, and no study participants participated in any
691 COVID-19 vaccine trials. All study visits occurred at the NIH Clinical Center (CC) in Bethesda, Maryland,
692 USA. Blood samples were collected by phlebotomy staff at the NIH CC. Samples were collected
693 between September 2020 and April 2021.

694

695 Samples were collected on subjects from three groups: 1) those with a prior history of symptomatic
696 SARS-CoV-2 infection (defined as a history positive nasal PCR test and positive Food and Drug
697 Administration (FDA) Emergency Use Authorization (EUA) SARS-CoV-2 antibody test at the time of
698 protocol screening), 2) those with a history of asymptomatic SARS-CoV-2 infection (defined as a
699 positive FDA EUA SARS-CoV-2 antibody test at the time of protocol exam but no history of COVID-like
700 symptoms; no time since COVID-19 infection or diagnosis (TSD) was identifiable for this group and they
701 were excluded from all TSD analyses), and 3) individuals with no history of SARS-CoV-2 infection
702 (defined as a negative FDA EUA SARS-CoV-2 antibody test at the time of the protocol screening).

703

704 Blood for PBMCs, serum, whole blood RNA [Tempus™ Blood RNA Tube (Thermo Fisher Scientific,
705 Waltham, MA)], complete blood count with differential (CBC) and lymphocyte phenotyping was
706 collected at each of the following timepoints relative to seasonal influenza vaccination (day 0): days -7,
707 0, 1, 7, 14, 28, 70, 100. Optional stool was collected at days 0, 28 and 100. Subjects were provided with
708 Cardinal Health Stool Collection kits (Cardinal Health, Dublin, OH) and Styrofoam storage containers
709 with ice packs to collect stool samples at home and return in person to the NIH. Following day 100,
710 subjects had the option to continue to provide monthly blood samples for PBMCs, serum, whole blood
711 RNA, CBC with differential and lymphocyte phenotyping through August 2021.

712

713 At each timepoint following study enrollment, data were collected and managed using REDCap
714 (v8.5.27) electronic data capture tools hosted at the NIH^{52,53}. REDCap (Research Electronic Data
715 Capture) is a secure, web-based software platform designed to support data capture for research
716 studies, providing 1) an intuitive interface for validated data capture; 2) audit trails for tracking data
717 manipulation and export procedures; 3) automated export procedures for seamless data downloads to
718 common statistical packages; and 4) procedures for data integration and interoperability with external

719 sources. REDCap electronic questionnaires were utilized to collect information from participants via
720 two separate IRB-approved surveys. A survey to evaluate vaccine-related adverse events or symptoms
721 was administered on study days 1 and 7 and a separate survey to evaluate for any health changes or
722 new medications was administered at every visit starting on Day 0. Surveys were sent via email to the
723 participants and responses were transferred from the REDCap system to the NIH Clinical Research
724 Information Management System (CRIMSON) system by the study team.

725

726 ***Influenza vaccination***

727 Subjects between ages 18 – 64 years were administered the Flucelvax Quadrivalent seasonal influenza
728 vaccine (2020-2021; Seqirus Inc, Summit, NJ). Subjects 65 years of age and older were administered
729 the high-dose Fluzone Quadrivalent seasonal influenza vaccine (2020-2021; Sanofi Pasteur Inc,
730 Swiftwater, PA).

731

732 ***Influenza microneutralization titers***

733 Virus-neutralizing titers of pre- and post-vaccination sera were determined in a microneutralization
734 assay based on the methods of the pandemic influenza reference laboratories of the Centers for
735 Disease Control and Prevention (CDC) using low pathogenicity vaccine viruses and MDCK cells. The X-
736 179A virus is a 5:3 reassortant vaccine containing the HA, NA, and PB1 genes from
737 A/California/07/2009 (H1N1pdm09) and the 5 other genes from A/PR/8/34 were donated by the high
738 growth virus NYMC X-157. Immune sera were also tested for neutralization titers of the seasonal
739 vaccine strains H1N1 A/Brisbane/59/07, H3N2 A/Uruguay/716/07, and B/Brisbane/60/2001. Internal
740 controls in all assays were sheep sera generated against the corresponding strains at the Center for
741 Biologics Evaluation and Research, FDA, Bethesda, MD. All individual sera were serially diluted (2-fold
742 dilutions starting at 1:10) and were assayed against 100 TCID₅₀ of each strain in duplicates in 96-well
743 plates (1:1 mixtures). The titers represent the highest dilution that completely suppressed virus
744 replication.

745

746 ***SARS-CoV-2 pseudovirus production and neutralization assay***⁵⁴⁻⁵⁶

747 Human codon-optimized cDNA encoding SARS-CoV-2 S glycoprotein (NC_045512) was cloned into
748 eukaryotic cell expression vector pcDNA 3.1 between the *Bam*HI and *Xho*I sites. Pseudovirions were
749 produced by co-transfection of Lenti-X 293T cells with psPAX2(gag/pol), pTrip-luc lentiviral vector and
750 pcDNA 3.1 SARS-CoV-2-spike-deltaC19, using Lipofectamine 3000. The supernatants were harvested at
751 48h post transfection and filtered through 0.45- μ m membranes and titrated using 293T-ACE2 cells
752 (HEK293T cells that express ACE2 protein). The following reagent was obtained through BEI Resources,
753 NIAID, NIH: Human Embryonic Kidney Cells (HEK-293T) Expressing Human Angiotensin-Converting
754 Enzyme 2, HEK-293T-hACE2 Cell Line, NR-52511.

755

756 For the neutralization assay, 50 μ L of SARS-CoV-2 S pseudovirions were pre-incubated with an equal
757 volume of varying dilutions of serum at room temperature for 1 h, then virus-antibody mixtures were
758 added to 293T-ACE2 cells in a 96-well plate. After 3 h incubation, the inoculum was replaced with fresh
759 medium. After 24 hours, cells were lysed and luciferase activity was measured. Controls included cell
760 only control, virus without any antibody control and positive control sera.

761

762 ***SPR based antibody binding kinetics of human serum***⁵⁷⁻⁵⁹

763 Steady-state equilibrium binding of serum was monitored at 25°C using a ProteOn surface plasmon
764 resonance (BioRad). The purified recombinant SARS-CoV-2 or other proteins were captured to a Ni-
765 NTA sensor chip (BioRad, Catalog number: 176-5031) with 200 resonance units (RU) in the test flow
766 channels. The protein density on the chip was optimized such as to measure monovalent interactions
767 independent of the antibody isotype. Serial dilutions (10-, 30- and 90-fold) of freshly prepared sample
768 in BSA-PBST buffer (PBS pH 7.4 buffer with Tween-20 and BSA) were injected at a flow rate of 50
769 µL/min (120 sec contact duration) for association, and disassociation was performed over a 600-second
770 interval. Responses from the protein surface were corrected for the response from a mock surface and
771 for responses from a buffer-only injection. Total antibody binding was calculated with BioRad ProteOn
772 manager software (version 3.1). All SPR experiments were performed twice, and the researchers
773 performing the assay were blinded to sample identity. In these optimized SPR conditions, the variation
774 for each sample in duplicate SPR runs was <5%. The maximum resonance units (Max RU) data shown in
775 the figures were the RU signal for the 10-fold diluted serum sample.

776
777

778 ***PBMC isolation***

779 PBMC samples were isolated from blood collected in Vacutainer EDTA tubes (generic lab
780 supplier) using the SepMate™-50 tubes (STEMCELL Technologies, Cambridge, MA) with following
781 modifications to the manufacturer's protocol: The blood samples were diluted 1:1 with room
782 temperate PBS and mixed by pipetting. The diluted blood was layered on top of 15ml Cytiva™ Ficoll™
783 PAQUE-Plus (Cytiva Life Sciences, Marlborough, MA) layer in SepMate™. The SepMate™ tubes were
784 spun at 1200 g for 10 mins with brake set to 5 at room temperature. Following the spin, the top plasma
785 layer was removed as much as possible without disturbing the PBMC layer. If there were any cells stuck
786 on the wall of the tube, then they were gently scraped from the wall with pipette, so they can be
787 resuspended with rest of the cells. The cells were poured from SepMate™ in to a 50ml conical tube.
788 The tubes containing cells were filled up to 50ml with cold wash buffer (PBS with 2% FBS) and mixed by
789 inverting. The tubes were spun at 300 g for 10 mins with brake set to 5 at room temperature. After the
790 spin, the supernatant was removed without disturbing the cell pellet. After resuspending the pellet
791 with cold wash buffer, the cells were counted using the Guava® Muse® Cell Analyzer (Luminex
792 Corporation, Austin, TX). The tubes were again spun at 300 g for 10 mins with brake set to 5 at room
793 temperature. The supernatant was removed without disturbing the cell pellet.

794

795 Based on the cell count, 6 – 10 million PBMC were frozen per vial for each sample. Since the cells were
796 counted prior to the last spin, a 50% cell loss was assumed and accounted for in the calculations from
797 cell count. The cell pellet was resuspended with $n \times 600 \mu\text{l}$ (n = number of PBMC vials to be frozen)
798 freezing media (RPMI with 10% FBS) by gentle pipetting. After freezing media, $n \times 600 \mu\text{l}$ DMSO freeze
799 (FBS with 15% DMSO) was added drop-by-drop while gently shaking the tube. In other words, for each
800 vial of PBMC that was to be frozen, 600µl of freezing media and 600µl of DMSO freeze was added,
801 bringing the total volume for each vial to 1.2ml. The solution was gently mixed by pipetting before
802 transferring 1.2ml cell solution to each 1.8ml cryovial (general lab supplier). The cell vials were placed
803 in CoolCell Containers (Thomas Scientific, Swedesboro, NJ) and the container was placed in a -80°C
804 freezer. After at least 4 hours, the PBMC vials were transferred to liquid nitrogen.

805

806

807 **RNA isolation**

808 Blood was drawn directly into the Tempus™ Blood RNA Tube (Thermo Fisher Scientific, Waltham, MA)
809 according to manufacturer's protocol. Two Tempus tubes were collected at each study timepoint. The
810 blood sample from each Tempus tube was aliquoted in to two 4.5mL cryovials (General lab supplier).
811 These cryovials were directly stored at -80°C.

812 The RNA samples were isolated in groups of 12-22 samples per batch based on careful batching prior
813 to isolation to reduce confounding factors due to age, gender, and patient group.

814
815 RNA was isolated from tempus blood using the QIA Symphony RNA Kit (Qiagen, Gaithersburg, MD) on
816 QIA Symphony SP instrument (Qiagen, Gaithersburg, MD). Blood samples were thawed on ice before
817 each sample was transferred to a 50ml conical tube. The total volume of the sample was brought to
818 12ml by adding 1x PBS. The tubes were vortexed at full speed for 30 seconds, followed by
819 centrifugation at 3500 g for 1 hour at 4°C. After centrifugation, the supernatant from the tubes was
820 decanted and tubes were placed upside down on clean paper towels for 2 minutes to allow residual
821 liquid to drain. To resuspend the pellet, 800µl of RLT+ buffer was added to the bottom of each tube
822 and vortexed for few seconds. All 800µl of each sample was transferred to 2ml screw cap tubes
823 (Sarstedt, Nümbrecht, Germany). The tubes were placed into #3b adapters (Qiagen, Gaithersburg, MD)
824 to be loaded on to the QIA Symphony.

825
826 On the QIA Symphony, RNA CT 800 protocol was selected and used for RNA isolation. The instrument
827 was set up according to the manufacturer's protocol and the elution volume for RNA samples was set
828 to 100µl. The final volume of the eluted RNA samples ranged from 65 – 95 µl.

829
830 RNA yields were determined using Qubit RNA BR kit or Qubit RNA HS kit (Thermo Fisher Scientific,
831 Waltham, MA) based on the yield. RNA RIN numbers were measured using RNA ScreenTape (Agilent
832 Technologies, Santa Clara, CA). The average RIN was 8.3 and average yield was 81.3 ng/µl for the RNA
833 samples.

834
835 **RNA-seq**

836 RNA-seq libraries were prepared manually using Universal Plus mRNA-Seq with NuQuant, Human
837 Globin AnyDeplete (Tecan Genomics, Redwood City, CA) according to manufacturer's protocol. For
838 each sample, 500ng of total RNA was used to isolate mRNA via poly(A) selection. Captured mRNA was
839 washed, fragmented, and primed with the mix of random and oligo(dT) primers. After cDNA synthesis,
840 ends were repaired and ligated with Unique Dual Index (UDI) adaptor pairs. Unwanted abundant
841 transcripts from rRNA, mtRNA and globin were removed using AnyDeplete module. Remaining library
842 was amplified by 14 cycles of PCR and purified with AMPure XP reagent (Beckman Coulter,
843 Indianapolis, IN).

844
845 Library concentration was determined by Quant-iT™ PicoGreen™ dsDNA Assay kit (Thermo Fisher
846 Scientific, Waltham, MA) on BioTek Synergy H1 plate reader (BioTek Instruments, Winooski, VT) using 2
847 µl sample. Library size distribution was determined using D1000 ScreenTape (Agilent Technologies,
848 Santa Clara, CA) on 4200 TapeStation System (Agilent Technologies, Santa Clara, CA). Thirty-two
849 samples were randomly selected from each plate to measure the library size distribution. To determine

850 fragment size, the region on the electropherogram was set from 200 bp to 700 bp. An average of the
851 fragment sizes was used for the rest of libraries to calculate molarity.

852

853 To create a balanced pool for sequencing, all libraries from one plate were diluted to the same molar
854 concentration by the QIAgility liquid handling robot (Qiagen, Gaithersburg, MD) and equal volumes of
855 normalized samples were pooled. Ninety-six samples were pooled from each plate on Plates 1-4 and 35
856 samples were pooled from Plate 5. For an accurate quantification of the pooled libraries, a qPCR was
857 performed using KAPA Library Quantification Kit (Roche, Wilmington, MA).

858

859 All libraries were sequenced on the NovaSeq 6000 instrument (Illumina, San Diego, CA) at Center for
860 Cancer Research Sequencing Facility, National Cancer Institute. The libraries pooled from Plates 1-4
861 were sequenced using one NovaSeq 6000 S4 Reagent Kit (200 cycles) and NovaSeq XP 4-Lane Kit
862 (Illumina, San Diego, CA) with sequencing parameter as 100 bp paired-end reads. The library pool from
863 Plate 5 was sequenced using a NovaSeq 6000 SP Reagent Kit (300 cycles; Illumina, San Diego, CA) with
864 150 bp paired-end reads as sequencing parameter.

865

866 Additionally, after quality control, 11 samples were re-sequenced as Plate 6 on a NextSeq 500
867 instrument using a NovaSeq 6000 S4 Reagent Kit (200 cycles) with sequencing parameter as 100 bp
868 paired-end reads. Technical replicates were placed on each plate to control for plate variability.

869

870 ***CITE-seq***

871 *a) Single cell CITE-seq processing*

872 Frozen PBMC samples were thawed, recovered, and washed using RPMI media with 10% FBS and
873 10mg/mL DNase I (STEMCELL) and then processed as previously described¹⁵ for CITE-seq staining. In
874 brief, samples from different donors were pooled and different timepoints from the same donor were
875 pooled separately so that each pool contains only one timepoint from one donor. PBMC pools were Fc
876 blocked (Human TruStain FcX, BioLegend) and stained with TotalSeq-C human 'hashtag' antibodies
877 (BioLegend), washed with CITE-seq staining buffer (2% BSA in PBS). Then hashtagged PBMC pools were
878 combined, and cells were stained with a cocktail of TotalSeq-C human lyophilized panel (BioLegend) of
879 137 surface proteins (including 7 isotype controls, refer to Supplementary Table 11) and SARS-CoV-2 S1
880 protein probe. Then, cells were washed, resuspended in PBS, and counted before proceeding
881 immediately to the single cell partition step.

882

883 *b) Single cell CITE-seq library construction and sequencing*

884 PBMC samples were partitioned into single cell Gel-Bead in Emulsion (GEM) mixed together with the
885 reverse transcription (RT) mix using 10x 5' Chromium Single Cell Immune Profiling Next GEM v2
886 chemistry (10x Genomics, Pleasanton, CA), as previously described¹⁵. The RT step was conducted in the
887 Veriti™ Thermal Cycler (ThermoFisher Scientific, Waltham, MA). Single cell gene expression, cell
888 surface protein, T cell receptor (TCR) and B cell receptor (BCR) libraries were prepared as instructed by
889 10x Genomics user guides (<https://www.10xgenomics.com/resources/user-guides/>). All libraries were
890 quality controlled using Bioanalyzer (Agilent, Santa Clara, CA) and quantified using Qubit Fluorometric
891 (ThermoFisher). 10x Genomics 5' Single cell gene expression, cell surface protein tag, TCR and BCR
892 libraries were pooled and sequenced on Illumina NovaSeq platform (Illumina, San Diego, CA) using the
893 following sequencing parameters: read1-100-cycle, i7-10-, i5-10, read2-100.

894

895 ***Serum isolation***

896 Serum was collected directly in Serum Separator Tubes and allowed to clot at room temperature for a
897 minimum of 30 minutes. Within two hours of blood collection, the tubes were spun at 1800 g for 10
898 minutes at room temperature. The top (serum) layer was removed via pipette and stored in individual
899 vials at -80°C.

900

901 ***Complete Blood Counts and lymphocyte phenotyping***

902 Subjects had standard complete blood counts with differential (CBCs) performed at the NIH CC in the
903 Department of Laboratory Medicine. Lymphocyte (T cell, B cell, NK cell) flow cytometry quantification
904 was performed using the BD FACSCanto™ II flow cytometer (BD Biosciences, Franklin Lakes, NJ).

905

906 ***PBMC in vitro stimulation***

907 PBMCs were thawed and cultured in PRMI1640 containing 10% fetal bovine serum, 2 mM glutamine,
908 0.055 mM beta-mercaptoethanol, 1% penicillin/streptomycin, 1 mM sodium pyruvate, 10 mM HEPES,
909 and 1% non-essential amino acids and stimulated with the following conditions: 1) IL-15 (10ng/mL), IL-
910 12 (20ng/mL), IL-18 (20ng/mL) for 48 hours; 2) IL-15 (50ng/mL) for 48 hours; 3) IL-18 (50ng/mL) for 48
911 hours; 4) IL-12 (20ng/mL), IL-18 (20ng/mL) for 48 hours; 5) anti-CD3 (1ug/mL), anti-CD28 (1ug/mL) for
912 24 hours; 6) non-stimulated controls. Protein Transport Inhibitor (BD Biosciences cat# 554724) and
913 Brefeldin A (BFA, Invitrogen, cat# 00-4506-51) were added 4 hours before harvest. Cytokines were
914 purchased from BioLegend: IL-15 (cat# 570304), IL-12 (cat# 573004), IL-18 (cat# 592104).

915

916 ***Flow cytometry***

917 *a) B cell phenotyping panel including influenza HA probes*

918 Thawed PBMC were washed in RPMI culture medium containing 50U/ml benzonase nuclease and then
919 washed by PBS. Cells were incubated with LIVE/DEAD Fixable Blue Dye (Life Technologies, Carlsbad,
920 CA), which was used to exclude dead cells from analysis. Cells were incubated with fluorochrome-
921 conjugated HAs for influenza B (B/Washington/02/2019 and B/Phuket/3073/2013 combined on the
922 same fluorochrome), and Influenza A H1 (A/Hawaii/70/2019) and H3 (A/Hongkong/2671/2019) and
923 fluorochrome-conjugated antibodies against IgM, IgA, CD21, CD85J, FCRL5, CD20, IgG, CD38, CD14,
924 CD56, CD3, CD27, CD71, CD19, IgD for 30 min at 4 C in the dark. The dyes and detailed information of
925 antibodies in the panel (Sarah Andrews, Vaccine Research Center, National Institute of Allergy and
926 Infectious Diseases, NIH) are summarized in Supplementary Table 12. After incubation with antibodies
927 for 30 minutes, cells were washed two times with FACS buffer (0.1%BSA/PBS (pH7.4)) and fixed in 1%
928 paraformaldehyde. Five million cells were acquired on Cytex Aurora spectral cytometer [Cytex
929 Biosciences, Fremont, CA; SpectroFlo® (Version 2.2.0)]. Data were analyzed with FlowJo software
930 version 10 (BD Biosciences).

931

932 *b) General immune phenotyping panel*

933 Thawed PBMC were washed in RPMI culture medium containing 50U/ml benzonase nuclease and then
934 washed by PBS. Cells were incubated with LIVE/DEAD Fixable Blue Dye (Life Technologies, Carlsbad,
935 CA), which was used to exclude dead cells from analysis. Cells were washed in FACS staining buffer (1 X
936 phosphate-buffered saline, 0.5% fetal calf serum, 0.5% normal mouse serum, and 0.02% NaN₃) and
937 incubated with Human Fc block reagent (BD bioscience #564220) at room temperature for 5 min. Cells

938 stained at room temperature for 10 minutes in the dark with fluorochrome-conjugated antibodies
939 against CCR7, CCR6, CXCR5, CXCR3 and TCRgd. Then, stained with fluorochrome-conjugated antibodies
940 against CD45RA, CD16, CD11c, CD56, CD8, CD123, CD161, IgD, CD3, CD20, IgM, IgG, CD28, PD-1,
941 CD141, CD57, CD45, CD25, CD4, CD24, CD95, CD27, CD1c, CD127, HLA-DR, CD38, ICOS, CD21, CD19,
942 CD14 at room temperature for 30 minutes in the dark. Cells were washed two times with FACS staining
943 buffer (1 X phosphate-buffered saline, 0.5% fetal calf serum, 0.5% normal mouse serum, and 0.02%
944 NaN₃) and fixed in 1% paraformaldehyde. Supplementary Table 13 shows the clones and information of
945 antibodies used in the phenotyping panel. A million PBMC were acquired by using Cytex Aurora
946 spectral cytometer [Cytex Biosciences, Fremont, CA; SpectroFlo® (Version 2.2.0)]. The frequency of
947 major populations was analyzed using with FlowJo™ software version 10 (BD Biosciences) based on
948 previously described manual gating strategies⁶⁰⁻⁶².

949

950 *c) In vitro stimulation T cell panel*

951 *In vitro* simulated PBMCs were collected and washed in PBS. Cells were incubated with Zombie UV
952 Fixable Viability Dye (BioLegend) in dark (room temperature) for 20 minutes. Cells were then washed
953 and incubated with Human TruStain FcX (Biolegend) for 10 minutes and subsequently with CCR7
954 antibody for 10 minutes. Fluorochrome-conjugated antibodies CD8, CD4, HLA-DR, CD69, CD45RA,
955 CD11c, CD5, CD3, TCRVa7.2, CD45RO, CD56, CD122, CD158e/k (KIR3DL1/DL2), KIR2D, NKG2A, CD14,
956 CD29, GPR56 cocktail were added and stained for 30 minutes in dark. Cells were washed and fixed with
957 Fixation/Permeabilization (BD Biosciences). Intracellular proteins IFN γ , TNF α , and Ki67 were stained
958 after fixation. Samples were collected using BD FACSymphony flow cytometer (BD Biosciences) and
959 analyzed using FlowJo™ software version 10. Supplementary Table 14 lists the antibodies used in the
960 panel.

961

962 **Data processing and transformation**

963

964 ***Bulk RNA-seq data processing***

965 Sequencing reads from Plate 5 were adaptor- and quality-trimmed to 100 bp using Trimmomatic
966 (v0.38.0)⁶³ to match the read length of the other plates (resulting reads with less than 100 bp were
967 discarded). Reads were then aligned to the human genome hg38 using the STAR (v2.6.0b) aligner.
968 Duplicate reads from PCR amplification were removed based on Unique Molecular Identifiers (UMI)
969 using UMI-tools (v0.5.3). Gene expression quantification was performed using the featureCounts⁶⁴
970 function from *Subread* package (v1.6.2). Samples with less than 5 million assigned reads were re-
971 sequenced and replaced. Reads were normalized and log transformed using *limma voom*⁶⁵. Lowly
972 expressed genes, defined as having fewer than five samples with > 0.5 counts per million reads, were
973 removed. Pre-vaccination (days -7 and 0) samples from the same healthy control (HC) subjects were
974 considered as replicates and were used to estimate latent technical factors by the RUVs function from
975 the *RUVSeq*⁶⁶ R package (v1.18). Four latent variables were included to derive normalized gene
976 expression values used for visualization and when specifically noted. Variable genes based on intra-
977 subject variability of pre-vaccination samples in the HCs and across technical replicates were filtered
978 out, resulting in a total of 10017 remaining genes for downstream analyses.

979

980 ***CITE-seq data processing***

981 *a) Single cell sample demultiplexing and preprocessing*

982 Single cell sequencing data was demultiplexed, converted to FASTQ format, mapped to human hg19
983 reference genome and counted using *CellRanger* (10x Genomics) pipeline. The sample level
984 demultiplex was done based on two levels as previously described¹⁵: 1) Hashtag antibody staining to
985 distinguish different timepoint samples from a same subject; 2) single nucleotide polymorphisms
986 (SNPs) called from the whole blood RNA-seq data to identify different subjects. Specifically, *CellRanger*
987 (v6.0.1) was used for generating count matrix and the software package *demuxlet* (v2, from the
988 'popscl' software suite)⁶⁷ was used to match single cell gene expression data to each donor and
989 identify empty droplets and doublets.

990

991 *b) Single-cell data clustering and cell annotation*

992 Single-cell data were further processed using Seurat (v4.0.3) running in R v4.1.1. We removed cells
993 with less than 200 and greater than 5,000 detected genes, greater than 60% of reads mapped to a
994 single gene, greater than 15% mitochondrial reads, cell surface protein tag greater than 20,000, and
995 hashtag antibody counts greater than 20,000. The protein data was normalized and denoised using the
996 DSB method (v0.3.0)⁶⁸. The following parameters were used in the dsb normalization function:
997 define.pseudocount = TRUE, pseudocount.use = 10, denoise_counts = TRUE, use.isotype.control =
998 TRUE. The DSB-normalized protein data were used to generate the top variable features (n = 100) and
999 principal components (PCs). Then the shared nearest neighbor (SNN) graph followed by k-nearest
1000 neighbors clustering were built using the FindNeighbors and FindClusters functions using first 15 PCs in
1001 Seurat (v4.0.3), respectively. Cell clusters were quality controlled based on their nearest neighbors and
1002 cell surface proteins. Cells were then further clustered within each major cell population using
1003 "weighted-nearest neighbor" (WNN) analysis in Seurat⁶⁹ (v4.1.0) by integrating both cell surface
1004 protein and gene expression modalities. WNN FindMultiModalNeighbors were done using both top 10
1005 PCs for cell surface protein and RNA of variable features. The WNN clusters were manually annotated
1006 and QC using the surface protein together with gene expression.

1007

1008 *c) Effector memory CD8 cell (CD8-EM) annotation for CITE-seq clusters*

1009 All CD8 cells were clustered using WNN as described above. CD8 clusters were annotated based on
1010 their surface markers as reported⁷⁰ together with gene expression profile. RNA expression of CD8 cells
1011 was mapped to external dataset using Seurat Label transfer method^{69,71} (v4.1.0). Clusters annotated as
1012 CD8-EM are surface CD45ROhi, CD45RAlo, CD95+, CD62Llo and CCR7- (mRNA) with most cells (~90%)
1013 mapped to CD8-EM phenotype cells in an external dataset^{69,71}.

1014

1015 *d) Single cell TCR data processing*

1016 *CellRanger* (v6.0.1) was used to assemble V(D)J contigs. The V(D)J assignment and clonotype were from
1017 the CellRanger output of the filtered contig_annotations.csv file for each 10x lane. The data is
1018 combined for all lanes and paired TCR α and TCR β chains for each single cell were combined using
1019 *scRepertoire* R package (v1.4.0)⁷² and integrated with the single-cell CITE-seq Seurat object metadata.
1020 Cells annotated as CD8 T cells and with both α and β chains detected are filtered and analyzed. CD8
1021 subsets and GPR56+ CD8 effector memory cell clonality were visualized with Circos plots using *Circlize*
1022 R package (v0.4.14)⁷³. For visualization purpose, cells from each subset were downsampled with equal
1023 number in each subset (for comparison between subsets, Extended Data Fig. 4g) or in each timepoint
1024 (for comparison between timepoints, Extended Data Fig. 4h,i). Cells were considered as the same clone

1025 with identical CDR3 (both α and β chains). Identical clones were connected within each sample or each
1026 subject across timepoints with lines.

1027

1028 ***OLINK serum proteomics***

1029 Missing values were imputed using k-nearest neighbors approach with k=10 using *impute* R package⁷⁴
1030 (v1.60.0). For each sample, probes targeting the same protein were averaged.

1031

1032 ***Cytek flow cytometry***

1033 Cell frequencies were generated by converting cell counts as fraction of live cells or lymphocytes as
1034 specified. The frequency data were log₂ transformed for linear modeling. For populations with zero
1035 counts in any of the samples, an offset equaling to half of the smallest non-zero value was added
1036 across samples.

1037

1038 ***CBC with diff and TBNK***

1039 Both absolute and relative counts were log₂ transformed for linear modeling. For parameters with zero
1040 values in any of the samples, an offset equaling to half of the smallest non-zero value was added across
1041 samples.

1042

1043 **Statistical Analysis**

1044 ***Baseline differential expression analysis***

1045 Using the *dream*⁷⁵ function in the *variancePartition* R package (v1.16.1), mixed-effects models were
1046 applied to determine differential levels of analytes (i.e., whole-blood gene expression, serum proteins,
1047 cell frequencies, flu titer and SPR, and hematological parameters) between COVID-recovered and HC
1048 subjects in a sex-specific manner as follows:

1049

1050 $\sim 0 + \text{group:sex} + \text{age} + \text{race} + \text{batch.effects} + (1 | \text{subject.id})$

1051

1052 Batch effect-related covariates were added to specific models depending on the assay type. For bulk
1053 RNA-seq, these include the four latent technical factors (see ***Bulk RNA-seq data processing***) and the
1054 timepoint-matched % neutrophils parameter from the CBC panel. For the Cytek and Olink platforms,
1055 sampling batch/plate was included as covariates. In addition to day 0, available samples from day -7 (in
1056 RNA-seq and CBC panel), were included as baseline replicates in the modeling.

1057

1058 Sex-specific group differences were computed from the contrasts covid.Female – healthy.Female and
1059 covid.Male – healthy.Male. Overall COVID vs. HC difference was determined by combining the two
1060 contrasts, i.e. (covid.Female – healthy.Female)/2 + (covid.Male – healthy.Male)/2. Sex difference linked
1061 to SARS-CoV2 infection was derived from the contrast (covid.male – covid.female) – (healthy.male –
1062 healthy.female) to account for normal differences between males and females. P values were adjusted
1063 for multiple testing within each assay type and contrast combination using the Benjamini-Hochberg
1064 (BH) method (Benjamini and Hochberg, 1995).

1065

1066 ***Association with time since COVID-19 diagnosis***

1067 To evaluate whether any of the differences detected at baseline had stabilized or might still be
1068 resolving, a linear model was used to test the association of relevant parameters with the time since
1069 COVID-19 diagnosis (TSD) among COVID-recovered subjects:

1070
1071 $\sim 0 + \text{sex} + \text{sex}:\text{scale}(\text{TSD}) + \text{age} + \text{race} + (1 | \text{subject.id})$

1072
1073 Two asymptomatic subjects without known TSD were excluded from the model. Association was
1074 assessed separately for females and males, and jointly by the combined contrast (Female:TSD +
1075 Male:TSD)/2. Dependent variables were converted to ranks in the model to reduce the effect of
1076 potential outliers.

1077
1078 Using a conservative approach, genes were classified as TSD-associated if they had an unadjusted p
1079 value < 0.05 and were excluded from subsequent analyses as specified. To determine whether any
1080 of the baseline differential gene sets were associated with TSD, leading edge gene (LEG) modules were
1081 derived from the union of all LEGs of the same gene set from different contrasts (see **Bulk RNA-seq**
1082 **gene set module scores**). A gene set was considered stable if none of three contrasts tested in the
1083 association model were significant (using unadjusted p value threshold of 0.05).

1084 1085 **Post-vaccination differential expression analysis**

1086 Similar to the workflow employed in baseline differential expression analysis, mixed-effects models
1087 were created to evaluate changes and group differences at each available timepoint after vaccination.
1088 Subjects aged 65 and above were excluded as they received a different type of vaccine. In addition to
1089 the baseline covariates, the model also accounts for the participants' flu vaccination history within last
1090 10 years as follows:

1091
1092 $\sim 0 + \text{visit}:\text{group}:\text{sex} + \text{age} + \text{race} + \text{flu.vax.count.10yr} + \text{batch.effects} + (1 | \text{subject.id})$

1093
1094 Three types of comparisons were examined using this model:

1095 **1. Timepoint-specific group differences**

1096 Similar to the contrasts in the baseline model, but for individual timepoints post vaccination
1097 (day 1 to day 100).

1098 **2. Vaccine-induced changes in group difference**

1099 Similar to the timepoint-specific contrasts above, but additionally subtracting off the
1100 corresponding baseline contrast to assess changes relative to the baseline. For example,
1101 vaccine-induced changes for female COVID vs. HC differences at D1 is evaluated with the
1102 contrast: (D1.covid.Female – D1.healthy.Female) – (Baseline.covid.Female –
1103 Baseline.healthy.Female).

1104 **3. Reversal of COVID vs. HC difference**

1105 Instead of using the HC subjects at the same corresponding timepoints as reference, post-
1106 vaccination samples from the COVID-recovered subjects were compared to baseline HC with
1107 the contrasts [timepoint].covid.Female – baseline.healthy.Female and [timepoint].covid.Male
1108 – baseline.healthy.Male. These contrasts can inform whether any pre-vaccination differences
1109 observed in the COVID-recovered subjects were reverted towards healthy baseline levels
1110 after vaccination. Reversal is defined as having smaller absolute effect size (using the z.std

1111 value from the dream function) at D1 and D28 after vaccination compared to the baseline
1112 absolute effect size.

1113
1114 P values were adjusted for multiple testing per each timepoint, assay type and contrast combination
1115 using the BH method.

1116 1117 **Gene set enrichment of differentially expressed (DE) genes**

1118 Enriched gene sets were identified using the pre-ranked gene-set enrichment analysis (GSEA) algorithm
1119 implemented in the *clusterProfiler* R package (v3.17.0)⁷⁶. Genes were ranked using signed $-\log_{10} p$
1120 values from differential expression models. Enrichment was assessed with gene set lists from MSigDB's
1121 Hallmark collection⁷⁷, Blood Transcriptomic Modules⁷⁸, and cell type gene signatures⁷⁹. Only gene sets
1122 with 10 to 300 genes were considered. P values were adjusted per gene set list for each contrast using
1123 the BH method and gene sets with FDR < 0.05 were considered significant. Baseline enriched gene sets
1124 were derived by intersecting significant gene sets extracted from DE models using samples
1125 independently from day -7, day 0, and both days combined. Genes associated with time since
1126 diagnosis (TSD) at baseline (see **Association with time since COVID-19 diagnosis**; Supplementary Tale
1127 1) were excluded from the post-vaccination enrichment analyses to help segregate the effect of
1128 vaccination from natural temporal resolution of the SARS-CoV-2 infection.

1129 1130 **Pseudobulk differential expression and gene set enrichment analysis**

1131 Single cells from a given sample were computationally "pooled" according to their cell type assignment
1132 by summing all reads for a given gene. Pseudobulk libraries made up by few cells and therefore likely
1133 not modeled properly by bulk differential expression methods were removed from analysis for each
1134 cell-type to remove samples that contained fewer than 4 cells and less than 35000 library size after
1135 pooling. Lowly expressed genes were removed for each cell type individually using the filterByExpr
1136 function from *edgeR* (v3.26.8)⁸⁰ with min.count = 2. Log counts per million (cpm) of each gene were
1137 calculated with scaling factors for library size normalization provided by the calcNormFactors function.
1138 Differential expression analysis was performed using the same models described in **Post-vaccination**
1139 **differential expression analysis** without running baseline models separately because the entire CITE-
1140 seq cohort was under 65 years of age. Batch assignment and number of barcodes/cells per sample
1141 were included as batch effects in this model.

1142
1143 Similarly, gene set enrichment analysis was carried out for each cell type in the same manner as for the
1144 bulk RNA-seq data (see **Gene set enrichment of differentially expressed (DE) genes**) which particular
1145 focus on the baseline enriched genesets identified by the bulk RNA-seq analysis. The Monaco gene sets
1146 were excluded from the single-cell analysis given the cell clusters were annotated and no further cell
1147 type demultiplex needed.

1148 1149 **Bulk RNA-seq gene set module scores**

1150 Gene set module scores were generated from RUVseq (v1.18) normalized gene expression values (see
1151 **Bulk RNA-seq data processing and transformation**) using gene set variation analysis (GSVA) method in
1152 GSVA R package (v1.30.0)⁸¹. LEG module scores representing enriched pathway activities were
1153 calculated for relevant samples using LEGs identified by GSEA to enhance signal-to-noise ratio. The

1154 average scores between days -7 and 0 were used for calculating post-vaccination changes relative to
1155 baseline.

1156

1157 ***Pseudobulk gene set module score calculation***

1158 Module scores (gene set signature score) representing enriched pathway activities were calculated for
1159 each pseudobulk sample of certain cell types. The pseudobulk gene counts were corrected with
1160 `removeBatchEffect` function in *limma* package (v3.42.2) to remove experimental batch and cell number
1161 effects and then normalized with `voom`⁸². The scores were then generated using gene set variation
1162 analysis (GSVA) method from the *GSVA* R package (v1.42.0)⁸¹. Specifically, for monocyte signatures,
1163 LEGs of BTM modules M4.0 and M11.0 were identified by GSEA from 1) D0.COVR-F vs. D0.HC-F and 2)
1164 D0.COVR-M vs. D0.HC-M models. The union of LEGs were used for the score calculation for female and
1165 male samples, respectively.

1166

1167 For BTM-M7.3 T cell activation signature and other signatures from acute COVID data as indicated in
1168 the figures, LEGs were used from the indicated comparison groups for the score calculation of female
1169 and male separately.

1170

1171 For monocyte antigen presentation signature, the module score was generated using LEGs from the
1172 BTM-M71 enriched in antigen presentation (I) and M95.0 enriched in antigen presentation (II) gene
1173 sets of the comparison: D1-D0 change between COVR-M vs. HC-M (Fig. 2f).

1174

1175 For the HALLMARK IFN γ response module score, all genes from the geneset were used for calculation
1176 of module scores in each celltype, so that the differences between celltypes can be compared.

1177

1178 ***Single-cell module score calculation and visualization***

1179 To visualize the differences between different patient groups in single data of the certain signatures,
1180 the genes from indicated genesets were used to calculate the module scores of each single cell.
1181 Module scores were calculated using `AddModuleScore` function in *Seurat* (v4.1.0) and then visualized
1182 in UMAP plots. For D1 vs. D0 HALLMARK IFN γ response module score differences shown in umaps, cells
1183 from D1.HC-F, D1.COVR-F, D1.HC-M and D1.COVR-M groups were downsampled to the same number
1184 of cells. The UMAP embeddings of cells colored with average differences for each high-resolution cell
1185 subsets are shown.

1186

1187 ***Single-cell module score calculation and test of external acute COVID-19 single-cell CITE-seq data***

1188 Single-cell data from the Brescia cohort of Liu *et al*¹⁵ was downloaded from GEO. Single monocytes
1189 data was extracted and Single-cell data from the Brescia cohort were pooled as described in “c)
1190 Pseudobulk differential expression and gene set enrichment analysis”. The gene set module scores of
1191 BTM modules M4.0 and M11.0 for all samples were generated using the union LEGs of male and
1192 female in “d) Gene set module score calculation”. The pseudobulk gene counts were normalized with
1193 the `varianceStabilizingTransformation` function from *DEseq2* R package (v1.34.0)⁸³. The scores were
1194 then generated using gene set variation analysis (GSVA) method from the *GSVA* R package (v1.42.0)⁸¹.
1195 Given there are multiple samples from each subject, the differences between patient groups (HC, less
1196 severe and more severe, corresponding to HC, DSM-low and DSM-high in Liu *et al*) were tested using
1197 the *Limma* (v3.50.1) linear model, where samples from the same donors were treated as duplicates

1198 using *duplicateCorrelation*. P-values of t statistics from the linear model of indicated contrasts are
1199 shown.

1200

1201 **Visualization of gene expression in heatmaps**

1202 Heatmaps showing pseudo-bulk data was generated using *ComplexHeatmap* R package (v2.10.0)⁸⁴. The
1203 log(CPM) (counts per million) normalized expression for each sample for a given celltype was
1204 calculated by pooling cells as described in “c) Pseudobulk differential expression and gene set
1205 enrichment analysis”. Heatmaps are showing z-score of the normalized expression for each gene in
1206 each sample.

1207

1208 **Data visualization**

1209 Plots were created using ggplot2 (v3.3.5) with ggpubr (v0.4.0) for statistical calculation unless noted.

1210 **Endpoint association**

1211 To evaluate the association of relevant parameters, including gene set module scores and cell
1212 frequencies, with interferon (IFN) or antibody titer fold change endpoints, the following model was
1213 applied:

1214

1215 endpoint ~ group:sex + scale(parameter):group:sex + age + race + flu.vax.count.10yr

1216

1217 The endpoint values were converted to rank to reduce the effects of potential outliers. Replicates from
1218 the same subjects were averaged.

1219

1220 **Serology**

1221 Influenza antibody titers below the detection limit of 1:20 were set to 1:10. Maximum titer across
1222 strains was calculated by normalizing titer levels across all samples from both day 0 and day 28
1223 individually for each of the four strains followed by taking the maximum standardized titer for each
1224 sample.

1225

1226 **Baseline titer difference analysis**

1227 For each of the four strains, a linear model was applied to determine baseline titer differences
1228 between COVID-recovered and HC subjects in a sex-specific manner as follows:

1229

1230 day 0 titer ~ group:sex + age + race

1231

1232 Titer values were log₁₀ transformed in the model, and sex-specific group differences were computed
1233 from the contrasts covid.Female – healthy.Female and covid.Male – healthy.Male. Subjects aged 65
1234 and above were excluded from the analysis.

1235

1236 **Day 28 titer difference analysis**

1237 For post-vaccination titer response, influenza vaccination history and baseline titer were included as
1238 covariates to partly account for prior exposure, similar to the approach used for influenza vaccine
1239 evaluation by the Food and Drug Administration (e.g., see

1240 <https://www.fda.gov/media/135687/download> page 27). Both day 28 titer and day 28/day 0 fold

1241 change (FC) were evaluated as endpoints to determine group differences between COVID-recovered
1242 and HC subjects for each of the four strains:

1243
1244 endpoint ~ group:sex + age + race + flu.vax.count.10yr + day 0 titer

1245
1246 For day 28 FC, a negative binomial model with log link was applied using glm.nb function in the MASS R
1247 package (v7.3-53). A linear model was used to fit the day 28 titers. Strain-specific titer values were
1248 log10 transformed in the model. Group differences were assessed using the same subjects and
1249 contrasts as in the baseline analysis.

1250
1251 Influenza antibody avidity as measured by surface plasmon resonance (SPR) were analyzed in the same
1252 manner as the titer data across HA1 and HA2, with the exception that that a linear model was applied
1253 for the fold changes.

1254
1255 ***Concordance in natural influenza infection cohort***

1256 A prospective cohort study with subjects profiled prior to and at least 21 days after natural influenza
1257 infection in two seasons¹⁸ was utilized to assess residual effects of the infection separately in males
1258 and females. Gene expression data were downloaded from GEO using the accession GSE68310.
1259 Subjects with only influenza A virus infection (n=51 females and 35 males) were identified and included
1260 for this analysis. Lowly expressed probes were removed, and the remaining data were converted to
1261 gene-based expressions. No additional processing steps were performed as the data were already
1262 normalized.

1263
1264 Separately for each season, differential expression analysis between baseline (pre-infection) and spring
1265 (long term post-infection) samples from the same individuals were performed using the dream
1266 function in the *variancePartition* R package (v1.16.1). A mixed-effects model accounting for flu
1267 vaccination history and disease severity (based on fever grade: none, low, and high) was constructed as
1268 follows:

1269
1270 ~ 0 + timepoint:sex + age + num.flu.vaccination + fever.grade + (1 | subject.id)

1271
1272 Differentially expressed (DE) genes were identified using the contrasts Spring.F - Baseline.F and
1273 Spring.M – Baseline.M for females and males, respectively. Sex difference was evaluated by the
1274 contrast (Spring.M – Baseline.M) – (Spring.F - Baseline.F). Concordance of DE results between the two
1275 seasons were evaluated based on correlation of effect size across genes (z.std values generated by
1276 dream).

1277
1278 Enrichment analysis was performed to determine whether the same set of genes were differentially
1279 expressed between pre- and post-influenza infection from this independent cohort and in COVID-
1280 recovered subjects compared to healthy controls prior to vaccination. To better match the age range of
1281 subjects between the two studies, baseline differential gene analysis was performed again with
1282 subjects under 65 years of age in the COVID cohort (see ***Baseline differential expression analysis***).
1283 Given that the males showed stronger concordance between the two flu seasons (Extended Data Fig.

1284 2b), COVID DE genes were ranked by signed $-\log_{10}$ p values and tested against a gene set formed by
1285 the intersect of DE ($p < 0.05$) genes in males from the flu infection cohort.

1286

1287 ***Elastic Net Multivariate Predictive Modeling***

1288 Elastic net models were constructed using the *eNetXplorer* R package (v1.1.3)⁸⁵ to predict day 1 (D1)
1289 INF γ response after influenza vaccination with both CITE-seq and flow cytometry cell frequencies at D0
1290 as predictors. A total of 33 subjects (COVR-F=11, HC-F=8, COVR-M=9, HC-M=5) with both CITE-seq and
1291 flow cytometry data were included. Based on 20 runs of 5-fold cross validation, a grid of regularization
1292 parameters (α and λ) were tested to determine models with best performance and cell subsets with
1293 consistent predictive power. Model performance was assessed by the mean squared error between
1294 the predicted and observed response. Importance of a cell population was determined by the
1295 frequency it was selected by the models (i.e., having non-zero coefficient). P values of the model
1296 performance and feature importance were derived by comparing to null models constructed with
1297 permuted response.

1298

1299 ***TCR diversity metric calculation***

1300 Shannon's entropy (H' index) was calculated as a measure of TCR diversity^{86,87}. Samples for each CD8
1301 subsets with fewer than 50 cells were filtered from the calculation. All samples were downsampled to
1302 50 cells because the diversity metric can be affected by the sample cell numbers. The process was
1303 repeated 1000 times with random downsampling of 50 cells and the median Shannon's index was used
1304 as an estimate of diversity for a given sample. Differences of the diversity metric between different
1305 CD8 subsets or timepoints were tested using two-tailed Wilcoxon test.

1306

1307 ***Reversal genes and bootstrapping to infer significance of difference in reversal of monocyte 1308 repressed signature between COVR-F and COVR-M***

1309 Reversal genes are defined as those whose COVR vs. D0 HC absolute effect size (z.std values from
1310 dream; see ***Post-vaccination differential expression analysis***) are smaller at both D1 and D28
1311 compared to D0.

1312

1313 Bootstrapping was employed to determine the significance of difference between COVR-F and COVR-M
1314 in their proportion of baseline LEGs from the monocyte depressed signature (BTM M4.0 and M11.0)
1315 that moved towards baseline HC. Subjects from each subject group were randomly sampled with
1316 replacement in each round of the bootstrapping and their samples were analyzed as described in ***Post-
1317 vaccination differential expression analysis***. The proportion of LEGs reversed after vaccination were
1318 calculated in each round for COVR-F and COVR-M in classical and non-classical monocytes, separately,
1319 and the p values plotted in Fig. 4e were determined based on 20 rounds of this procedure.

1320

1321

1322

1323 **Data Availability**

1324 Raw and processed data from the whole blood bulk RNA-seq and single-cell CITE-seq are available from
1325 the NCBI Gene Expression Omnibus, accession numbers GEO: GSE194378
1326 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE194378>) and GEO:GSE206265
1327 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE206265>), respectively. Additional datasets,
1328 including clinical, proteomics, flow cytometry, and influenza antibody measurements, are available at:
1329 <https://doi.org/10.5281/zenodo.5935845>. The influenza infection dataset we utilized was downloaded
1330 directly from GEO: GSE68310 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE68310>).
1331 Reprint requests, permissions, and requests for materials may be directed to the corresponding
1332 author, John Tsang (john.tsang@yale.edu).

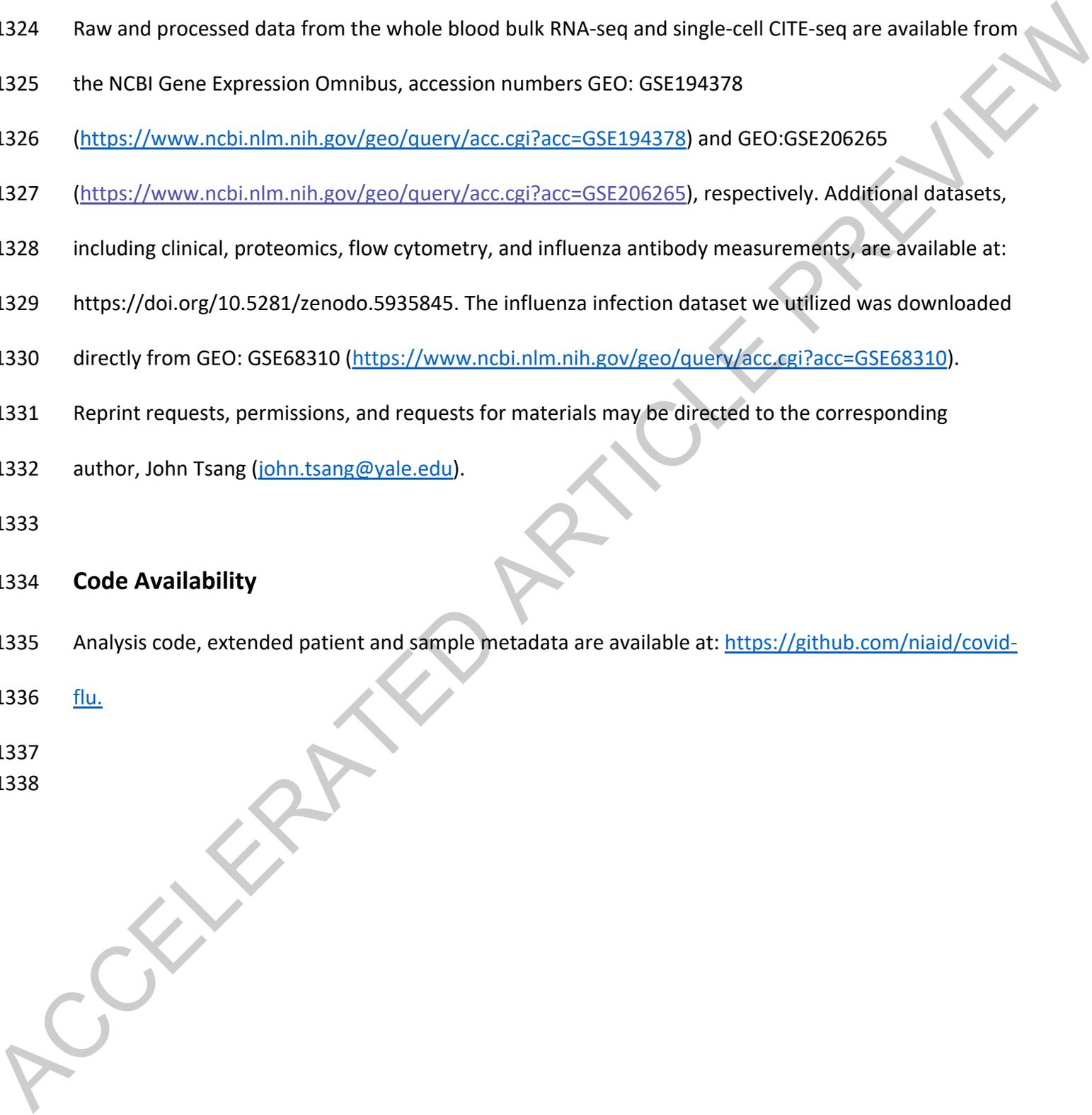
1333

1334 **Code Availability**

1335 Analysis code, extended patient and sample metadata are available at: [https://github.com/niaid/covid-](https://github.com/niaid/covid-flu)
1336 [flu](https://github.com/niaid/covid-flu).

1337

1338



1339 **Methods and Extended Data Figure References**

1340 52. Harris, P. A. *et al.* Research electronic data capture (REDCap)—A metadata-driven methodology
1341 and workflow process for providing translational research informatics support. *J. Biomed. Inform.*
1342 **42**, 377–381 (2009).

1343 53. Harris, P. A. *et al.* The REDCap consortium: Building an international community of software
1344 platform partners. *J. Biomed. Inform.* **95**, 103208 (2019).

1345 54. Ravichandran, S. *et al.* Antibody signature induced by SARS-CoV-2 spike protein immunogens in
1346 rabbits. *Sci. Transl. Med.* **12**, eabc3539 (2020).

1347 55. Ravichandran, S. *et al.* Longitudinal antibody repertoire in ‘mild’ versus ‘severe’ COVID-19 patients
1348 reveals immune markers associated with disease severity and resolution. *Sci. Adv.* **7**, eabf2467
1349 (2021).

1350 56. Tang, J. *et al.* Antibody affinity maturation and plasma IgA associate with clinical outcome in
1351 hospitalized COVID-19 patients. *Nat. Commun.* **12**, 1221 (2021).

1352 57. Khurana, S. *et al.* MF59 adjuvant enhances diversity and affinity of antibody-mediated immune
1353 response to pandemic influenza vaccines. *Sci. Transl. Med.* **3**, 85ra48 (2011).

1354 58. Khurana, S. *et al.* Human antibody repertoire after VSV-Ebola vaccination identifies novel targets
1355 and virus-neutralizing IgM antibodies. *Nat. Med.* **22**, 1439–1447 (2016).

1356 59. Khurana, S. *et al.* Repeat vaccination reduces antibody affinity maturation across different
1357 influenza vaccine platforms in humans. *Nat. Commun.* **10**, 3338 (2019).

1358 60. De Biasi, S. *et al.* Marked T cell activation, senescence, exhaustion and skewing towards TH17 in
1359 patients with COVID-19 pneumonia. *Nat. Commun.* **11**, 3434 (2020).

- 1360 61. Heit, A. *et al.* Vaccination establishes clonal relatives of germinal center T cells in the blood of
1361 humans. *J. Exp. Med.* **214**, 2139–2152 (2017).
- 1362 62. Park, L. M., Lannigan, J. & Jaimes, M. C. OMIP-069: Forty-Color Full Spectrum Flow Cytometry Panel
1363 for Deep Immunophenotyping of Major Cell Subsets in Human Peripheral Blood. *Cytom. Part J. Int.*
1364 *Soc. Anal. Cytol.* **97**, 1044–1051 (2020).
- 1365 63. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data.
1366 *Bioinforma. Oxf. Engl.* **30**, 2114–2120 (2014).
- 1367 64. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning
1368 sequence reads to genomic features. *Bioinforma. Oxf. Engl.* **30**, 923–930 (2014).
- 1369 65. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: Precision weights unlock linear model analysis
1370 tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
- 1371 66. Risso, D., Ngai, J., Speed, T. P. & Dudoit, S. Normalization of RNA-seq data using factor analysis of
1372 control genes or samples. *Nat. Biotechnol.* **32**, 896–902 (2014).
- 1373 67. Kang, H. M. *et al.* Multiplexed droplet single-cell RNA-sequencing using natural genetic variation.
1374 *Nat. Biotechnol.* **36**, 89–94 (2018).
- 1375 68. Mulè, M. P., Martins, A. J. & Tsang, J. S. Normalizing and denoising protein expression data from
1376 droplet-based single cell profiling. *Nat. Commun.* **13**, 2099 (2022).
- 1377 69. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).
- 1378 70. Martin, M. D. & Badovinac, V. P. Defining Memory CD8 T Cell. *Front. Immunol.* **9**, 2692 (2018).
- 1379 71. Zheng, G. X. Y. *et al.* Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.*
1380 **8**, 14049 (2017).

- 1381 72. Borcharding, N., Bormann, N. L. & Kraus, G. scRepertoire: An R-based toolkit for single-cell immune
1382 receptor analysis. *F1000Research* **9**, 47 (2020).
- 1383 73. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize Implements and enhances circular
1384 visualization in R. *Bioinforma. Oxf. Engl.* **30**, 2811–2812 (2014).
- 1385 74. Hastie, T., Tibshirani, R., Narasimhan, B. & Chu, G. impute: impute: Imputation for microarray data.
1386 R package version 1.72.1. (2022).
- 1387 75. Hoffman, G. E. & Roussos, P. Dream: powerful differential expression analysis for repeated
1388 measures designs. *Bioinforma. Oxf. Engl.* **37**, 192–201 (2021).
- 1389 76. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innov. N. Y.*
1390 *N* **2**, 100141 (2021).
- 1391 77. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740
1392 (2011).
- 1393 78. Li, S. *et al.* Molecular signatures of antibody responses derived from a systems biology study of five
1394 human vaccines. *Nat. Immunol.* **15**, 195–204 (2014).
- 1395 79. Monaco, G. *et al.* RNA-Seq Signatures Normalized by mRNA Abundance Allow Absolute
1396 Deconvolution of Human Immune Cell Types. *Cell Rep.* **26**, 1627-1640.e7 (2019).
- 1397 80. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor RNA-Seq
1398 experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012).
- 1399 81. Hänzelmann, S., Castelo, R. & Guinney, J. GSEA: gene set variation analysis for microarray and RNA-
1400 seq data. *BMC Bioinformatics* **14**, 7 (2013).
- 1401 82. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and
1402 microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

- 1403 83. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-
1404 seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
- 1405 84. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in
1406 multidimensional genomic data. *Bioinforma. Oxf. Engl.* **32**, 2847–2849 (2016).
- 1407 85. Candia, J. & Tsang, J. S. eNetXplorer: an R package for the quantitative exploration of elastic net
1408 families for generalized linear models. *BMC Bioinformatics* **20**, 189 (2019).
- 1409 86. Rosati, E. *et al.* Overview of methodologies for T-cell receptor repertoire analysis. *BMC Biotechnol.*
1410 **17**, 61 (2017).
- 1411 87. Morris, E. K. *et al.* Choosing and using diversity indices: insights for ecological applications from the
1412 German Biodiversity Exploratories. *Ecol. Evol.* **4**, 3514–3524 (2014).
- 1413 88. Nakaya, H. I. *et al.* Systems biology of vaccination for seasonal influenza in humans. *Nat. Immunol.*
1414 **12**, 786–795 (2011).

1415
1416
1417
1418

1419 **Acknowledgements**

1420 We are indebted to the study participants whose generosity has expanded our scientific understanding
1421 of COVID-19 and viral infections in general. We thank Kaitlyn Sadtler for assistance with the
1422 development of the Cytex 36-color flow cytometry panel, Qin Xu for assistance with the design of the
1423 cytokine stimulation experiments, Foo Cheung for computational assistance, Matthew Mulè for
1424 providing software (“dsb”), the NCI CCR Sequencing Facility for sequencing support, NIAID HPC Team
1425 for computing support, and Ronald Germain for critical reading of the manuscript. Fig. 1a, 1b, 2a, 3a,
1426 4a, Extended Data Fig. 2a were created using BioRender.com. This research was supported by the

1427 Intramural Research Programs of the NIAID and the Intramural Programs of the NIH Institutes
1428 supporting the NIH Center for Human Immunology. JH was funded in part with federal funds from the
1429 National Cancer Institute, National Institutes of Health, under Contract No. 75N91019D00024. The
1430 content of this publication does not necessarily reflect the views or policies of the Department of
1431 Health and Human Services, nor does mention of trade names, commercial products, or organizations
1432 imply endorsement by the U.S. Government.

1433
1434

1435 **Author Contributions**

1436 Conceptualization: RS, JST; Methodology: WWL, CL, JST; Software and formal analysis: WWL, CL;
1437 Investigation: RS, WWL, CL, KLH, NB, JM, GG, JRK, GK, JC, JT, AJM; Resources: RS, WWL, CL, KLV, GS,
1438 MC, SFA, NB, LEF, SL, LS, PB, DC, AC, JH, AKM, VM, CP, DS, AM, BAS, RA, ABM, EMB, HG, SK, JST; Data
1439 curation: RS, WWL, CL, KLV, MC, NB, LEF; Writing – original draft: RS, JST; Writing – review and editing:
1440 RS, WWL, CL, JST; Visualization: RS, WWL, CL; Supervision: RS, HG, SK, JST; Project administration: RS,
1441 JST; Funding acquisition: RS, JST. Human study and sample acquisition: RS, KLV, GS, MC, NB, LEF, SL, LS,
1442 PB, DC, AC, JH, AKM, VM, CP, DS, AM; Flow cytometry of PBMCs (including influenza-specific cell
1443 analyses): KLH with panel designs, reagents, and help from SFA; Whole blood RNA-seq: GK, NB;
1444 Influenza microneutralization titers, influenza antibody SPR, and SARS-CoV-2 neutralization titers: JM,
1445 GG, LRK, JT; CITE-seq data generation: CL, AJM, JC with help from GK and KLV; PBMC cytokine
1446 stimulation experiments: CL. Integrative analysis of flow cytometry, CITE-seq, blood RNA-seq,
1447 circulating protein, and antibody data: WWL, CL with conceptual contributions from JST.

1448
1449
1450

1450 **Competing Interests**

1451 EMB reports personal fees and non-financial support from Terumo BCT, Abbott Laboratories, Tegus
1452 and UptoDate, outside of the submitted work. EMB is a member of the United States Food and Drug
1453 Administration (FDA) Blood Products Advisory Committee. Any views or opinions that are expressed in
1454 this manuscript are those of the author's, based on his own scientific expertise and professional
1455 judgment; they do not necessarily represent the views of either the Blood Products Advisory
1456 Committee or the formal position of FDA, and do not bind or otherwise obligate or commit either
1457 Advisory Committee or the Agency to the views expressed. No other authors report competing
1458 interests.

1459
1460 **Consortia authors**

1461 OP11 Clinical Staff: Princess Barber¹¹, Daly Cantave¹¹, Anne Carmona¹¹, Jean Hammer¹², Alaina K.
1462 Magnani¹¹, Valerie Mohammed¹¹, Cindy Palmer¹¹, Deitra Shipman¹¹

1463 ¹¹ Clinical Center Nursing Department, NIH, Bethesda, MD, USA

1464 ¹² Clinical Research Directorate, Frederick National Laboratory for Cancer Research, Frederick, MD,
1465 USA

1466 **Extended Data Figure 1. Baseline differences between COVID-19-recovered subjects and matching**
1467 **controls.**

1468 **a**, Box plot showing the distribution of time since diagnosis (TSD; x-axis) in COVR-F (n=16) and COVR-M
1469 (n=15). Two participants with asymptomatic COVID-19 infection and thus unknown TSD are not
1470 included. Significance of group difference is determined by two-tailed Wilcoxon test.

1471 **b**, Scatterplot showing the correlation between the TSD (x-axis) and the SARS-CoV-2 neutralization titer
1472 (WA1 strain; y-axis) for COVR subjects at D0 prior to influenza vaccination. Spearman's rank correlation
1473 and unadjusted p values are shown. Participants with asymptomatic COVID-19 infection not included in
1474 TSD analyses. The shaded area represents 95% confidence interval.

1475 **c**, Similar to **(b)**, but for the percentage of monocytes in peripheral blood as measured by the complete
1476 blood count (y-axis) at D0. Dotted lines represent median level in HC-F and HC-M.

1477 **d**, Similar to **(b)**, but for the proportion of CD11c+ dendritic cells (DCs; as fraction of live cells; y-axis) as
1478 measured by flow cytometry of PBMCs at D0.

1479 **e**, Blood transcriptomic analysis of the stable baseline (pre-vaccination) differences among COVR and
1480 HC groups. Enrichment plot shows the normalized enrichment scores (NES) of selected gene sets of the
1481 different comparisons (GSEA FDR < 0.05; see Methods; see Supplementary Table 3 for all significant
1482 gene sets with FDR < 0.05). The NES are plotted separately for COVR-F versus HC females (HC-F), COVR-
1483 M versus HC males (HC-M), or the difference between the two sets of comparisons (COVR-M versus
1484 COVR-F taking healthy sex differences into account). Positive (negative) NES indicates that gene set
1485 scores are higher (lower) in the first group than the second group listed in the comparison. Only gene
1486 sets not correlated with TSD across COVR subjects at baseline are considered stable.

1487 **f**, Similar to Fig. 1d, but for percent of CD3+ cells (T cells).

1488 **g**, Similar to **(e)**, but for a subset of monocyte and T-cell activation gene sets with significant
1489 enrichment ($p < 0.05$) using the D0 CITE-seq pseudobulk expression for the specified cell types (see
1490 Methods; see Supplementary Table 5 for complete results).

1491 **h**, Scatterplots showing the relationship between the TSD and leading-edge gene (LEG) module scores
1492 [left two boxes: the T-cell activation gene set (BTM-M7.3); right two boxes: the union of the LEGs from
1493 gene sets BTM M4.0 and M11.0; see Methods] in COVR-F (n=12) (top row) and COVR-M (n=12)
1494 (bottom row) at D0 using the CITE-seq pseudobulk data of the indicated cell types. Each dot represents

1495 a COVR subject. The dotted lines represent the median score for the sex-matched HC group at D0 in
1496 the comparison shown. Spearman's rank correlation and p values are shown. The shaded area
1497 represents 95% confidence interval.

1498 i, (left) Box plot comparing the classical monocyte pseudobulk module scores of the LEGs used in Fig. 1f
1499 (union of female (F) and male (M) gene sets) in an acute COVID-19 CITE-seq dataset from Liu *et al*¹⁵.
1500 Both M (n=50) and F (n=9) subjects are included in all three groups (HC n=13, less severe n=21, more
1501 severe n=25). Each dot represents a sample. Unadjusted p values from the indicated two-group
1502 comparisons are shown. P values were generated using the moderated T statistics from a linear model
1503 in which samples from the same donors were treated as duplicates (See Methods). (right) Bubble plot
1504 showing expression of the genes in Fig. 1f right panel within the classical monocyte CITE-seq data from
1505 Liu *et al*. in the same three patient groups shown in the left panel. All box plot elements are the same
1506 as indicated in Fig. 1.

1507

1508

1509 **Extended Data Figure 2. Persistent post-infection gene expression changes following natural**
1510 **influenza infection.**

1511 **a**, Schematic showing the approach used to evaluate changes in blood gene expression between
1512 before (pre-infection baseline) and months after natural influenza infection over two distinct seasons
1513 published in Zhai *et al*¹⁸, and how those gene changes may relate to sex-specific differences resulted
1514 from prior COVID-19 in this study.

1515 **b**, Scatter density plot showing the correlation between the gene expression changes (see Extended
1516 Data Table 7) before (fall) and after (spring) natural influenza A infection in 2009 (x-axis) and 2010 (y-
1517 axis) for females (F; left), males (M; center), and M vs F contrast (right). Shown are Spearman's rank
1518 correlation and unadjusted p values.

1519 **c**, Gene set enrichment plot of the genes that increased in M between fall (pre-infection) and spring
1520 (post-infection) in both 2009 – 2010 and 2010 – 2011 seasons. Genes were ranked by the signed
1521 $\log_{10}(\text{unadjusted p-value})$ in the COVID-19-recovered (COVR)-M vs COVR-F contrast at baseline using
1522 only subjects under 65 years of age. The tick marks denote the location of the genes in the influenza
1523 gene set.

1524

1525

1526 **Extended Data Figure 3. Sex-specific molecular, cellular, and humoral response differences to**
1527 **influenza vaccination in COVID-19-recovered individuals and matching controls.**

1528 **a**, Similar to Extended Data Fig. 1e but here showing enriched gene sets in whole blood comparing the
1529 early [day 1 (D1) and day 7 (D7)] influenza vaccination responses in COVID recovered (COVR) vs.
1530 healthy control (HC) subjects for females (F; Contrast 1), males (M; Contrast 2), and sex differences
1531 (Contrast 2 vs. 1; i.e., COVR-M versus COVR-F taking healthy sex differences into account; see
1532 Methods). Plotted are the gene sets that show significant changes from the baseline [day -7 and day 0
1533 (D0)] within each comparison group [e.g., COVR-F and HC-F for D1] and significant differences between
1534 the two groups at the indicated timepoints (FDR < 0.05; see Supplementary Table 5).

1535 **b**, Similar to Fig. 2e, but showing the D0 Hallmark IFN γ Response module score for the indicated cell
1536 types from the CITE-seq pseudobulk expression data. CD4 = CD4+ T-cells; cDC = conventional/myeloid
1537 dendritic cells; B = B-cells.

1538 **c**, Box plot showing the D7 whole blood signature score from genes identified in Nakaya *et al*⁸⁸ whose
1539 D7/D0 fold-change positively correlated with day 28 (D28) influenza hemagglutination inhibition titers.
1540 Only subjects under 65 years of age [COVR-F (n=15), COVR-M (n=14), and HC-F (n=16), and HC-M
1541 (n=14)] are included. Significance of differences is determined by two-tailed Wilcoxon test.

1542 **d**, Scatter plot showing the correlation of the whole blood D1 – D0 Hallmark Interferon Gamma
1543 Response gene set module score (x-axis) to the whole blood D7 – D0 plasmablast (PB) gene set module
1544 score (left y-axis; Monaco *et al*⁷⁹) and D7 – D0 difference of influenza-specific PB (all HA+ CD27+ CD38+
1545 CD20^{low} CD21^{low}) frequency from flow cytometry (right y-axis; as fraction of CD19+ B-cells). Only study
1546 participants < 65 years of age are included. Spearman's rank correlation and unadjusted p values are
1547 shown.

1548 **e**, Box plots showing the D0 (pre-vaccination) microneutralization titers for each of the four strains in
1549 the seasonal influenza vaccine (columns) in females (COVR-F and HC-F) under the age of 65.
1550 Unadjusted p values are from linear models accounting for age and race (see Methods).

1551 **f**, Similar to (e) but for males (COVR-M and HC-M) under 65 years of age.

1552 **g**, Maximum standardized influenza vaccine titer (among the four strains in the vaccine) at D28 after
1553 vaccination for females (left) and males (right), respectively. Unadjusted p values are from linear
1554 regression models accounting for age, race, influenza vaccination history, and baseline influenza titer

1555 (see Methods). Unadjusted p values are shown. All box plot elements are the same as indicated in Fig.

1556 2.

1557

1558

1559

ACCELERATED ARTICLE PREVIEW

1560 **Extended Data Figure 4. GPR56+ virtual memory-like CD8+ T-cells contribute to increased day 1 IFN γ**
1561 **response in COVID-19-recovered males.**

1562 **a**, Scatterplots showing the correlation between the day 0 (D0) log₂ frequency of early effector-like
1563 CD8+ T-cells measured by flow cytometry (as fractions of live lymphocytes; see Population 50 in
1564 Supplementary Table 9 and Supplementary Fig. 1; x-axis) and the change (D1 – D0) in serum interferon
1565 gamma (IFN γ) protein levels measured by the OLINK platform (y-axis) for COVID-19-recovered (COVR)
1566 females (COVR-F; top, n=14) and COVR males (COVR-M; bottom; n=11). Spearman's rank correlation
1567 and p values are shown.

1568 **b**, Similar to (a) but showing the correlation between the D0 log₂ frequency of early effector-like CD8+
1569 T-cells measured by from flow cytometry (as fraction of live lymphocytes; x-axis) and the change (D1 -
1570 D0) in the whole blood signature score of the Hallmark IFN γ Response gene set (y-axis).

1571 **c**, Box plots comparing D0 and D1 pseudobulk IFNG gene expression (y-axis) in GPR56+ CD8 EM
1572 population for HC-F (n=8), COVR-F (n=12), HC-M (n=8) and COVR-M (n=12). Significance is determined
1573 by a linear model accounting for age, race, and influenza vaccination history (see Methods).

1574 **d**, Bar plot showing the T statistic of the vaccine-induced change (D1 - D0) in IFN γ gene (IFNG)
1575 expression using CITE-seq pseudobulk data (x-axis) within the GPR56+ and GRP56- CD8 EM for HC-F (n=
1576 8), COVR-F (n=12), HC-M (n=8), and COVR-M (n=12). * p < 0.05 with exact value shown in (c).

1577 **e**, Scatter plot showing the correlation between GPR56+ CD8 EM cell frequency (as fractions of total
1578 CD8 EM in the CITE-seq data; x-axis) and BTM-M7.3 T-cell activation signature score of CD8 EM cells
1579 computed using CITE-seq pseudobulk gene expression data (y-axis). Spearman correlation and p values
1580 are shown. The shaded area represents the 95% confidence interval.

1581 **f**, Related to Fig. 3h but showing CD45RA and CD45RO only with CD8+ TEMRA cells included as an
1582 additional comparator.

1583 **g**, (left) Circos plot of T-cell receptor (TCR) clonality for different CD8+ T-cell subsets at D0. Segments in
1584 the outer circle represent different CD8+ T-cell populations. Segments in inner circle represent male
1585 (M) and female (F) for both COVR and HC subjects. Grey lines connect clones sharing identical CDR3
1586 sequences within each individual. Cell subsets are downsampled for visualization (see Methods). (right)

1587 Box plot showing Shannon's entropy index (y-axis) at D0 for each of the indicated CD8+ populations.
1588 Significance of differences is determined by two-tailed Wilcoxon test. Shannon's entropy index

1589 evaluates the TCR repertoire diversity for each sample. Higher indices indicate higher diversity (i.e.,
1590 fewer shared clones shown in Circos plot). EM = effector memory; CM = central memory; TEMRA = EM
1591 cells re-expressing CD45RA.

1592 **h**, (left) Circos plot of TCR clonality for GPR56+ CD8 EM cells at different timepoints. Segments in the
1593 outer circle represent different days in the study (D0, D1, D28). Segments in the inner circle represent
1594 males (M) and females (F) for both COVR and HC subjects. Grey lines connect clones sharing identical
1595 CDR3 sequences within each sample. Timepoints are downsampled for visualization purposes (see
1596 Methods). (right) Box plot showing Shannon's entropy index (y-axis) of TCR clonality at each of the
1597 indicated time points (D0, D1, D28; x-axis) for GPR56+ CD8 EM T-cells (left) and GPR56+ CD8+ TEMRA
1598 (right). Significance of differences is determined by two-tailed Wilcoxon test.

1599 **i**, (left) Similar to (**h**), but showing the shared clones among different timepoints (segments in the outer
1600 circle). Segments in the inner circle represent unique clones for each individual. Clones and lines
1601 connecting shared clones are colored. Samples with less than 30 cells were filtered out for visualization
1602 purposes. (right) Line chart showing frequencies of each clone (y-axis) shown in Circos plot (left) at D0,
1603 D1 and D28 for each subject. P-values of paired Wilcoxon test are shown comparing the clone
1604 frequency differences among D0, D1 and D28.

1605 **j**, Related to Fig. 3i but showing the frequencies of IFN γ^+ NK, IFN γ^+ CD45RA+ CD45RO+ TEMRA CD8+ T-
1606 cells and IFN γ^+ MAIT cells after IL-15 stimulation *in vitro*. All box plot elements are the same as
1607 indicated in Fig. 3. Unadjusted p values are shown.

1608

1609

1610

1611 **Extended Data Figure 5. Changes in immune states in COVID-19-recovered individuals following**
1612 **influenza vaccination.**

1613
1614 **a**, Distributions of gene-level difference of the innate immune receptor (IIR) signature (see Fig. 1f) in
1615 classical monocytes separately for females (F) and males (M) [shown as z-scores, on a per gene level,
1616 capturing the average difference between COVID Recovered (COVR) at the indicated timepoint (top to
1617 bottom: D0, D1, and D28) and healthy control (HC) at D0; see Methods]. Dashed red vertical lines
1618 represent the median of the distribution. Dark tick marks at the bottom represent individual genes and
1619 colored dots highlight specific genes of interest. Significance of differences from D0 is determined by
1620 paired two-tailed Wilcoxon test.

1621 **b**, Similar to (a) but for the non-classical monocytes (see Fig. 1g).

1622 **c**, Similar to Fig. 4d but for non-classical monocytes (see Fig. 1g for the innate receptor signature in
1623 non-classical monocytes).

1624 **d**, Similar to Fig. 4b but for COVR-F (red) and COVR-M (blue) only and gene sets shown on top of each
1625 plot. Box plots showing the classical monocyte LEG module scores (y-axis) of gene sets from
1626 Supplementary Fig. 2: antigen presentation related gene sets, Hallmark Inflammatory response,
1627 Hallmark TNF- α signaling via NF- κ B, and MS-1 signature from Reyes *et al*¹⁷. LEGs from the first three
1628 gene sets were found to be repressed in acute COVID-19 patients in Liu *et al*¹⁵.

1629 **e**, Similar to (d), but for non-classical monocytes. All box plot elements are the same as indicated in Fig.
1630 4. Unadjusted p values are shown.

1631
1632

1633 **Extended Data Table 1. Cohort characteristics.** Demographics (age, sex, self-reported race), influenza
1634 vaccination history, and COVID-19 related information (for recoverees). COVR = COVID-19 recovered.

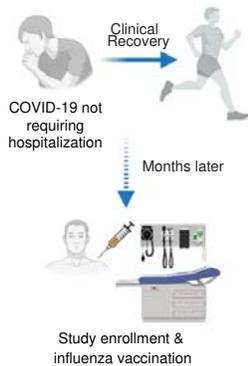
1635

1636

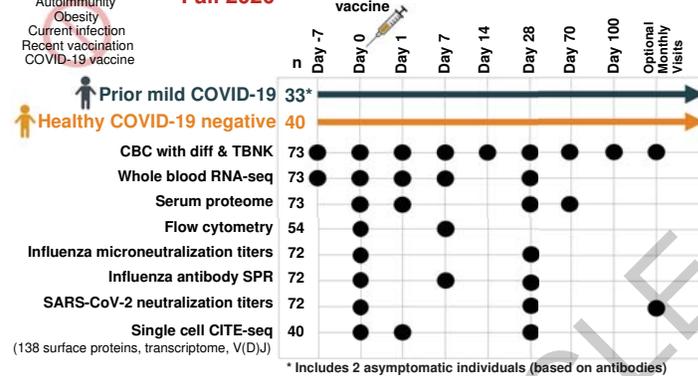
1637

ACCELERATED ARTICLE PREVIEW

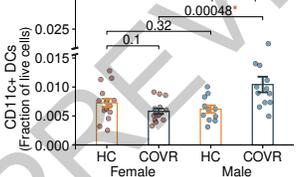
a Non-obese, healthy individuals



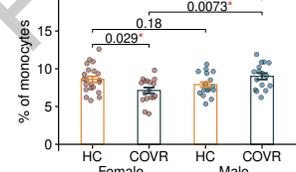
b



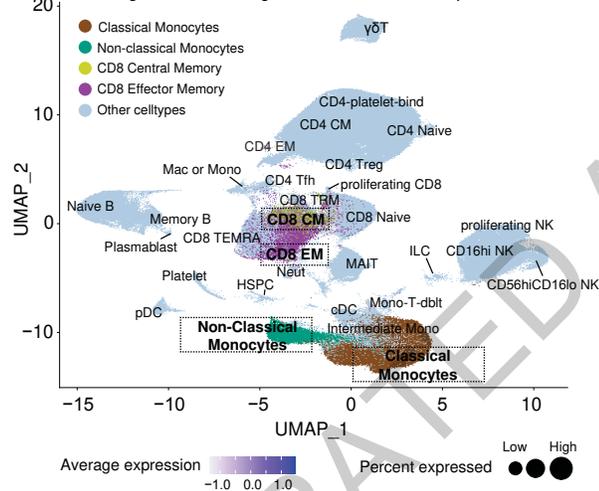
c



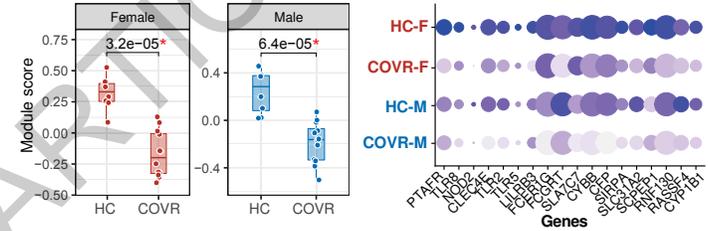
d



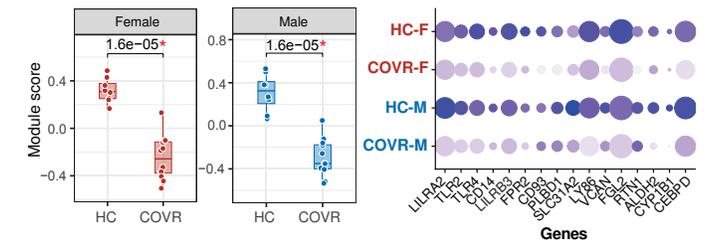
e Single cell clustering based on cell surface protein



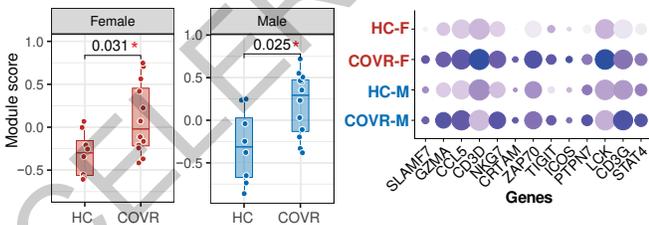
f



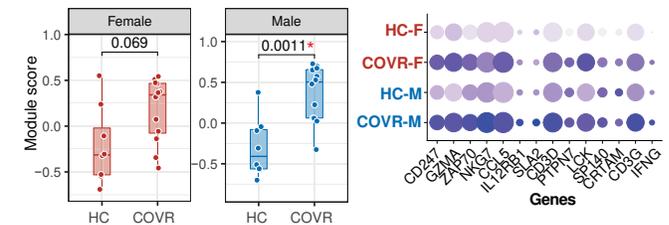
g



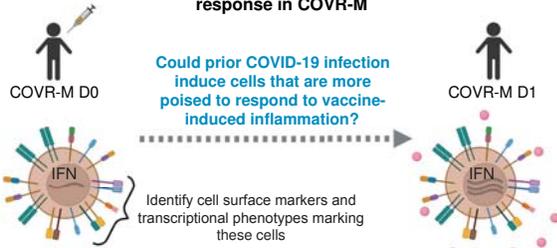
h CD8 Central Memory: BTM-M7.3 T cell activation signature



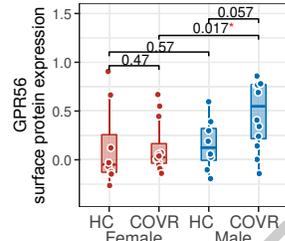
i



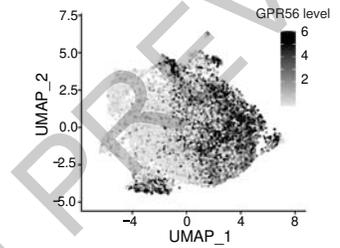
a Identify and characterize cells related to elevated day 1 IFN response in COVR-M



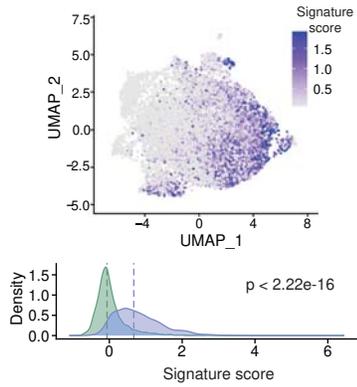
b D0 CD8 EM



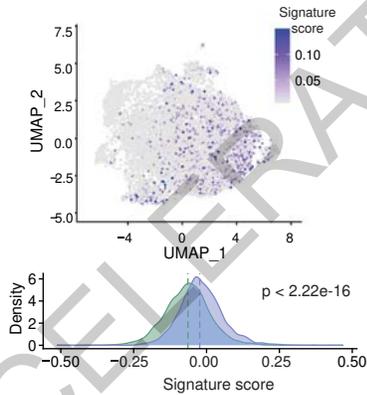
c Surface GRP56 expression of CD8 EM cells



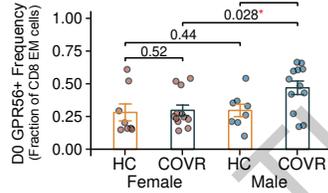
d Score of a transcriptional signature marking pre-stimulation CD8+ T cells that tend to produce more IFN γ after stimulation



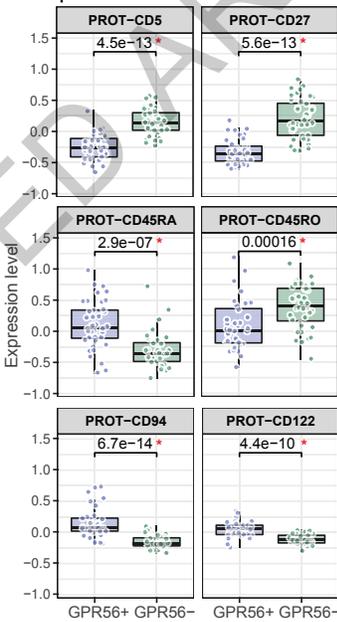
f Score of a transcriptional signature marking bystander T cells up-regulated in acute COVID-19



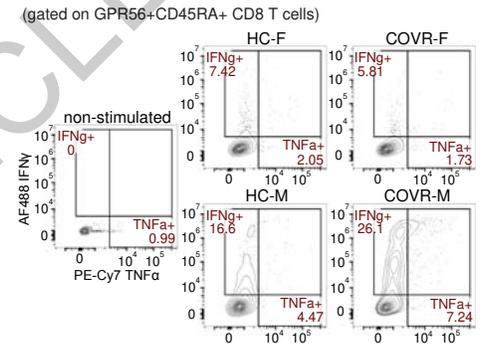
e D0 GPR56+ Frequency (Fraction of CD8 EM cells)



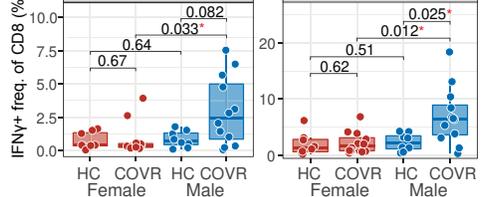
g D0 VM-like CD8 memory surface protein markers



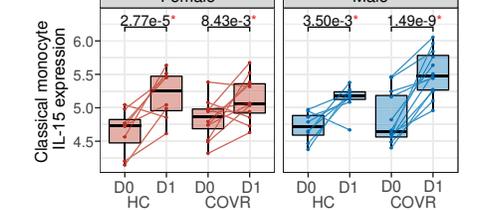
h IL-15 stimulation (gated on GPR56+CD45RA+ CD8 T cells)



i IFN γ + freq. of CD8 (%)

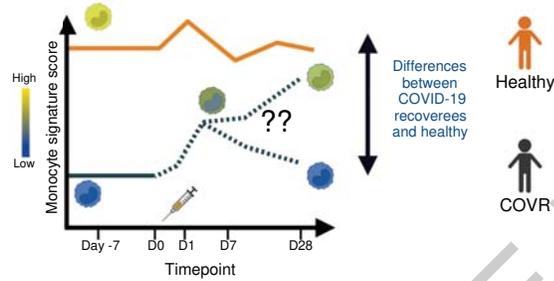


j Classical monocyte IL-15 expression

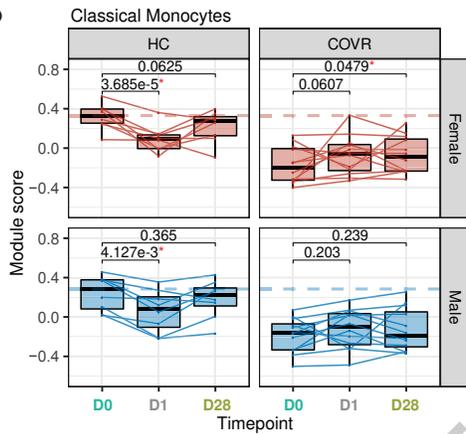


a

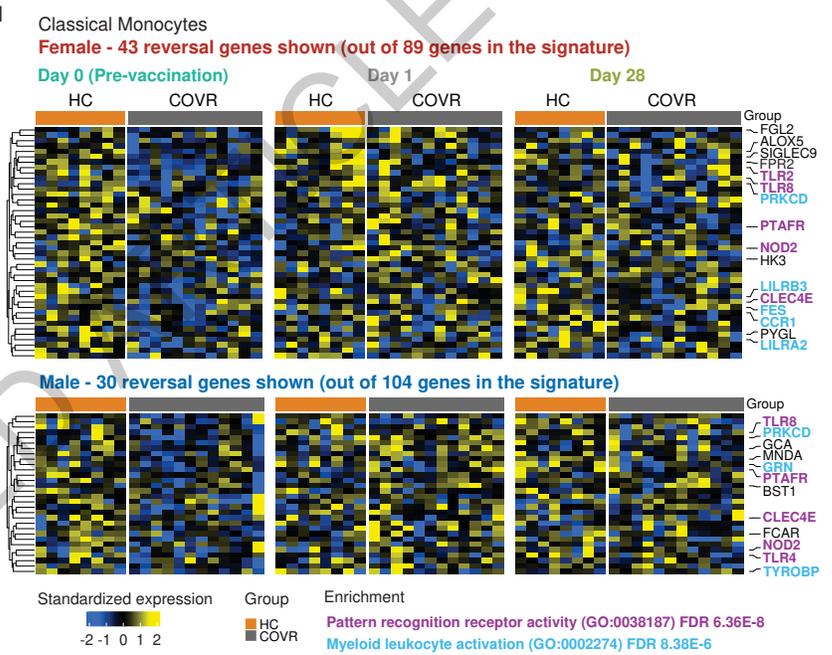
Does the influenza vaccine shift immune states toward normal?



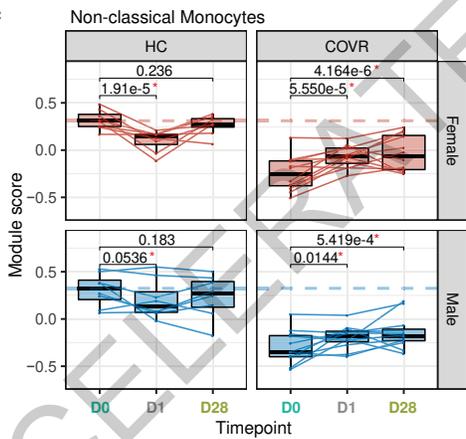
b



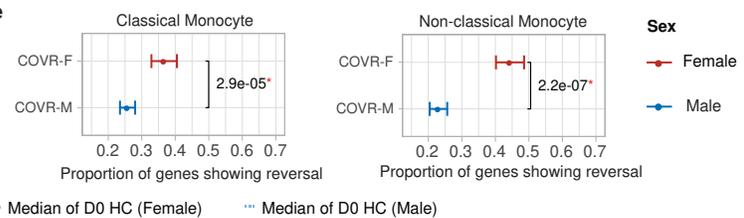
d



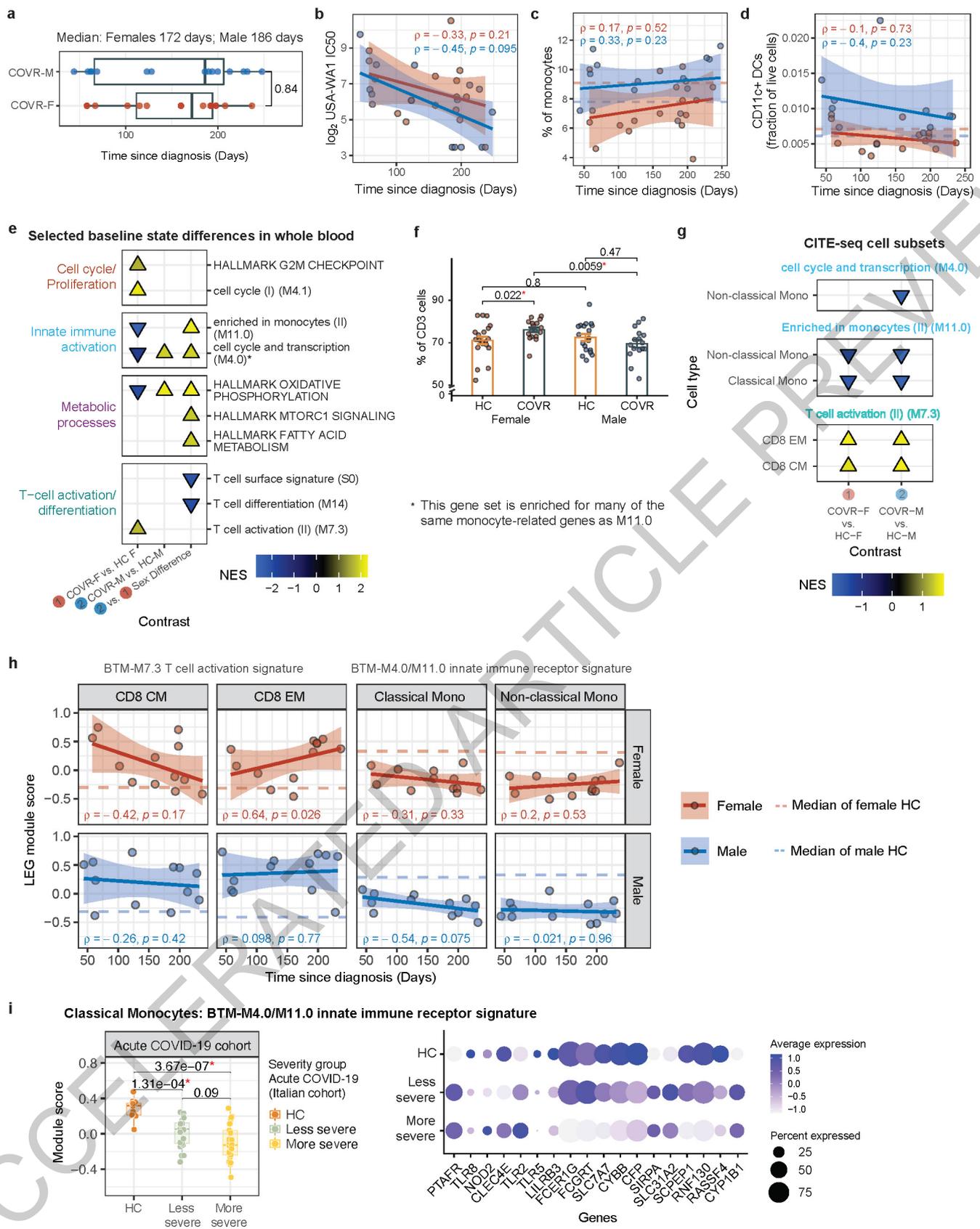
c



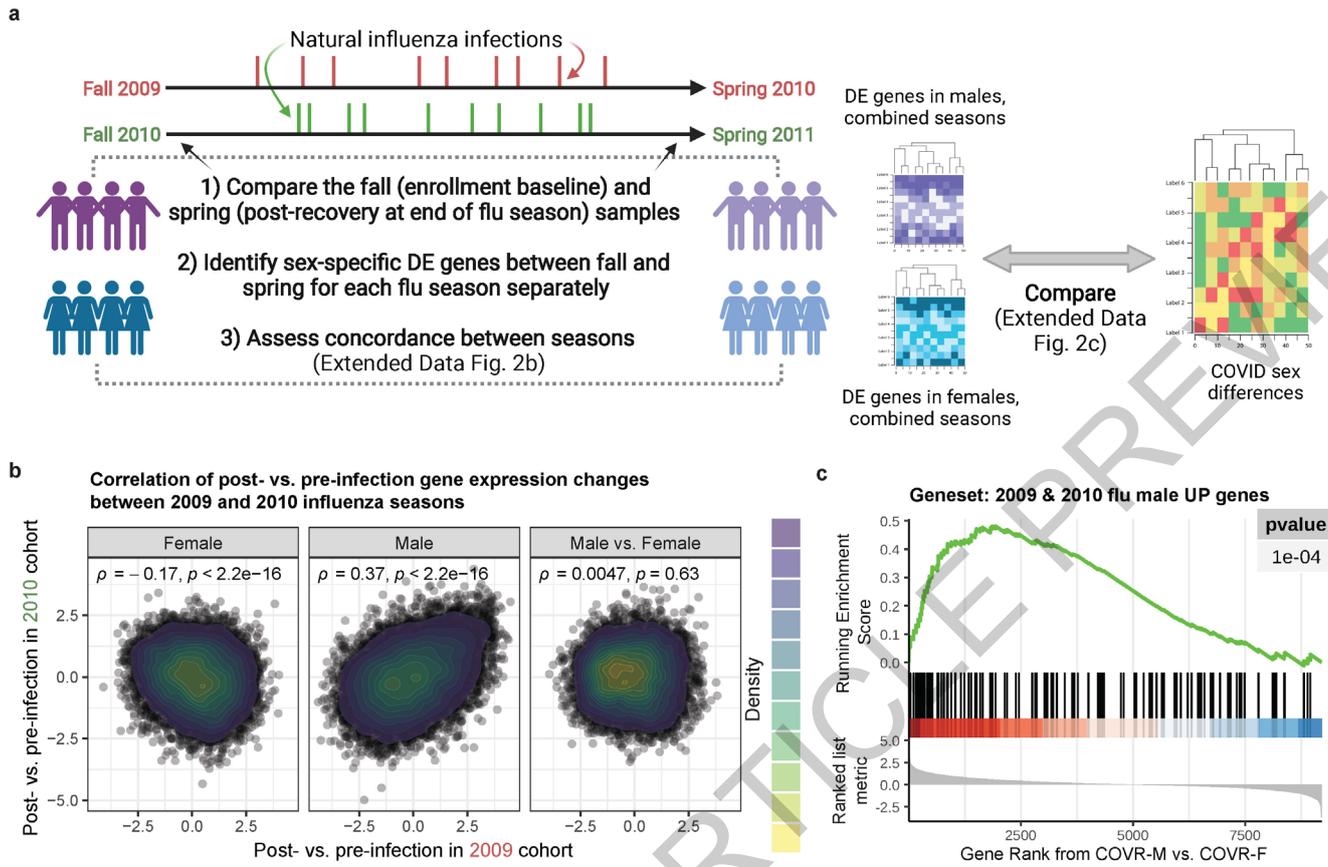
e



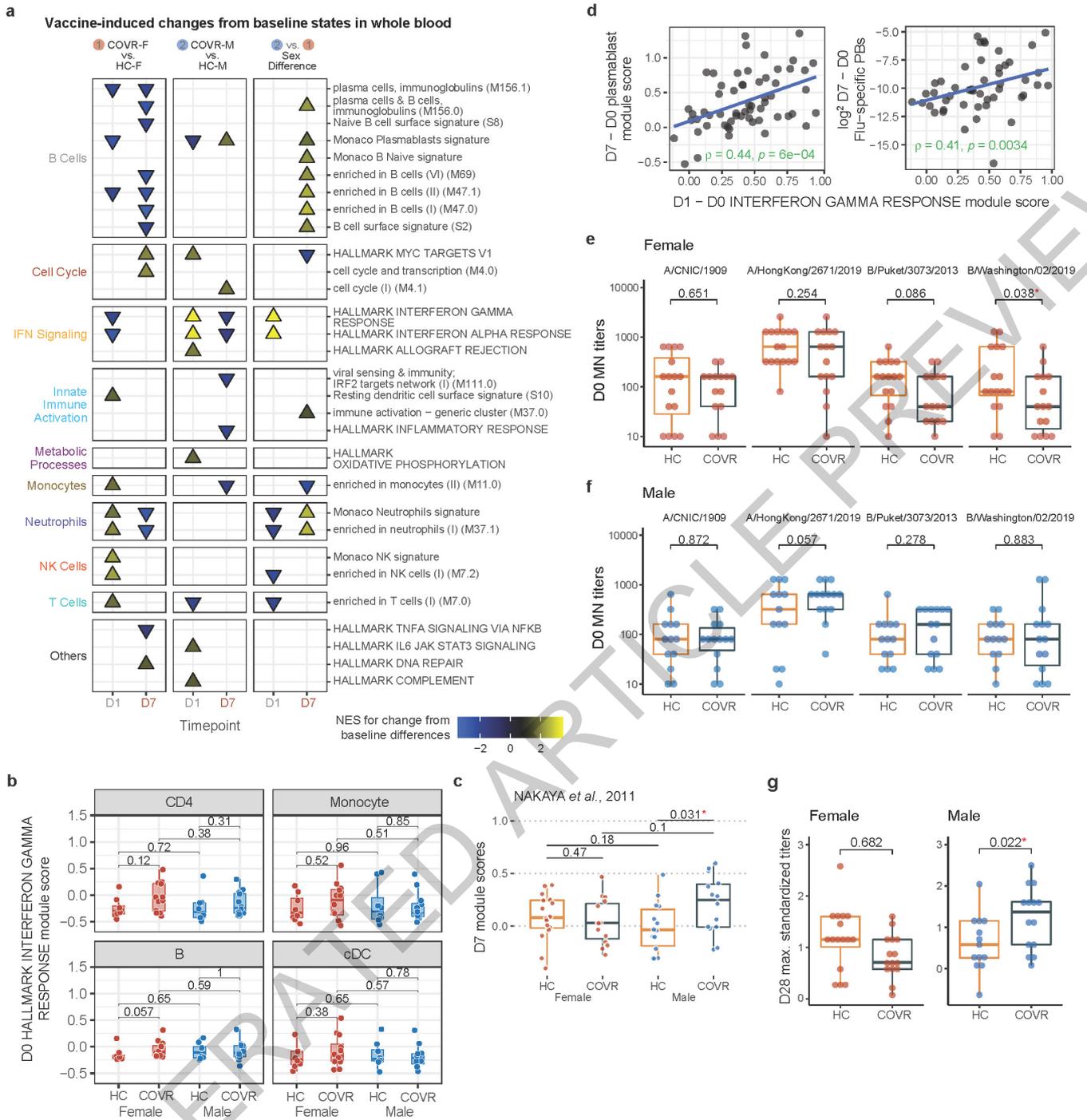
ACCELERATED PREVIEW



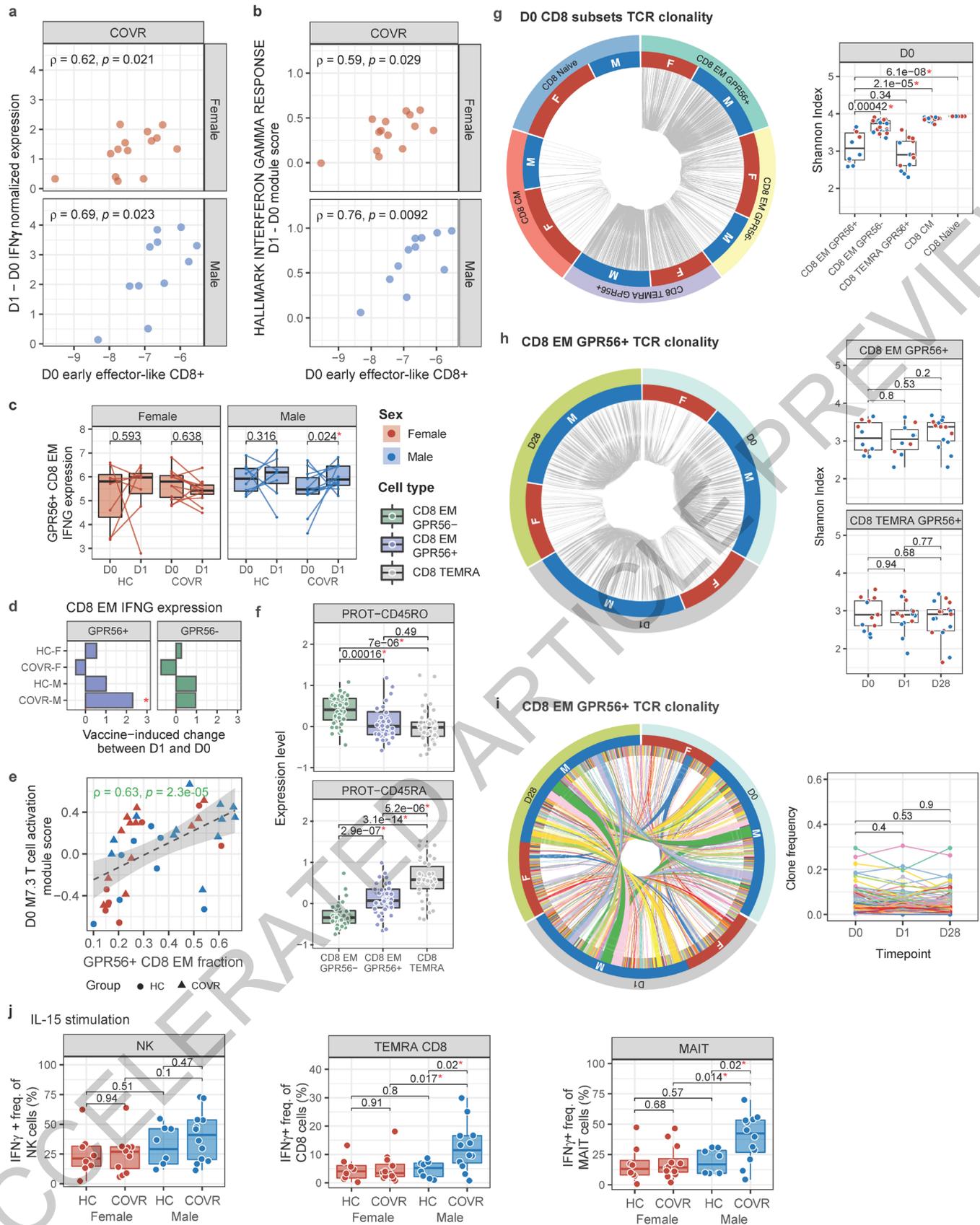
Extended Data Fig. 1



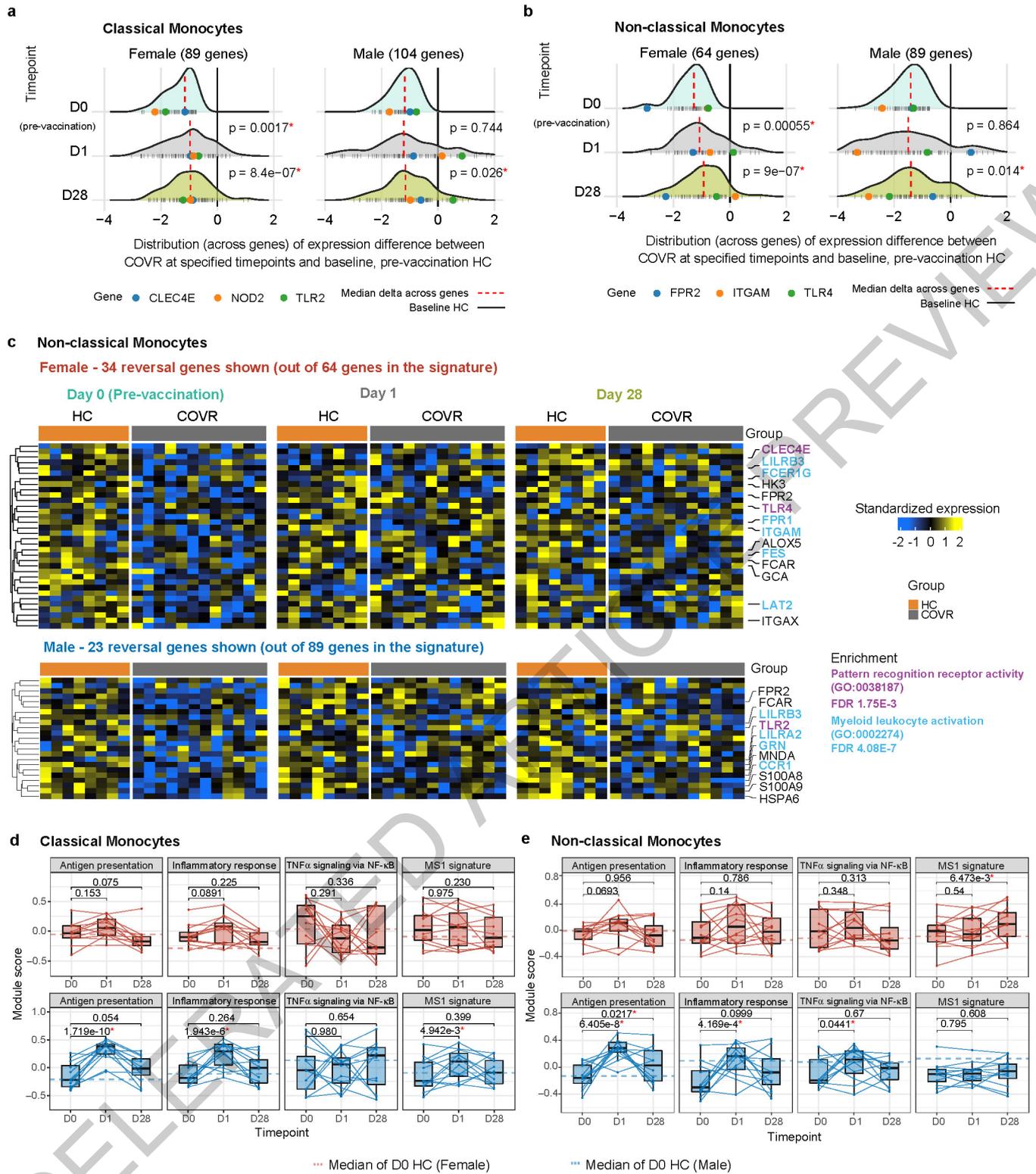
Extended Data Fig. 2



Extended Data Fig. 3



Extended Data Fig. 4



Extended Data Fig. 5

	COVR		Healthy	
	Female	Male	Female	Male
Subject count	17	16	21	19
Age (Years)				
Median	40.2	43.7	52.5	47.6
Mean	44.9	43.7	47.2	47.4
Min	23.4	21.9	22.5	24.0
Max	70.5	67.3	70.4	69.1
Aged > 65	2	2	5	5
Race				
Asian	1	0	2	2
Black	1	0	4	0
Multiple races	0	2	1	3
White	15	14	14	14
Number of Influenza Vaccination in Past 10 Years				
0	0 (0%)	1 (6.25%)	0 (0%)	2 (10.53%)
1	1 (5.88%)	0 (0%)	1 (4.76%)	0 (0%)
2	0 (0%)	1 (6.25%)	2 (9.52%)	2 (10.53%)
3	1 (5.88%)	0 (0%)	1 (4.76%)	0 (0%)
4	0 (0%)	0 (0%)	2 (9.52%)	0 (0%)
5	3 (17.65%)	3 (18.75%)	1 (4.76%)	0 (0%)
6	1 (5.88%)	1 (6.25%)	1 (4.76%)	0 (0%)
7	3 (17.65%)	0 (0%)	1 (4.76%)	1 (5.26%)
8	0 (0%)	1 (6.25%)	1 (4.76%)	3 (15.79%)
9	1 (5.88%)	1 (6.25%)	1 (4.76%)	1 (5.26%)
10	7 (41.18%)	8 (50%)	10 (47.62%)	10 (52.63%)
Experienced side effects after vaccination	16 (94.1%)	9 (56.3%)	17 (90.0%)	15 (78.9%)
COVID-19 Symptoms				
Asymptomatic	1 (5.9%)	1 (6.3%)	-	-
Symptomatic	16 (94.1%)	15 (93.8%)	-	-
Time since COVID-19 Diagnosis (Days) *				
Median	172.0	186.0	-	-
Mean	152.9	149.3	-	-
Min	58.0	44.0	-	-
Max	237.0	248.0	-	-
Duration of acute COVID-19 symptoms (Days) *				
Median	14	11	-	-
Mean	19.62	22.07	-	-
Min	4	1	-	-
Max	87	140	-	-
Experienced COVID-19 residual symptoms at time of screening	8 (47.1%)	3 (18.8%)	-	-
COVID-19 Residual Symptoms				
Brain fog	1 (5.9%)	0 (0%)	-	-
Fatigue	2 (11.8%)	0 (0%)	-	-
Loose stools	0 (0%)	1 (6.3%)	-	-
Reduced sense of taste	1 (5.9%)	1 (6.3%)	-	-
Reduced sense of smell / smell disturbances	5 (29.4%)	2 (12.5%)	-	-
Shortness of breath / Chest pressure	1 (5.9%)	0 (0%)	-	-

* Excluding asymptomatic subjects

Extended Data Table 1

Reporting Summary

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

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | n/a | Confirmed |
|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of all covariates tested |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	REDCap (Research Electronic Data Capture) v8.5.27
Data analysis	Trimomatic (v0.38.0), STAR (v2.6.0b), UMI-tools (v0.5.3), subread (v1.6.2), R (v3.6.1, 4.1.0, 4.1.1), variancePartition (v1.16.1), limma (v3.42.2, 3.50.1), clusterProfiler (v3.17.0), ggplot2(v3.3.5), ggpubr (v0.4.0), RUVSeq (v1.18), GSVA (v1.30.0, 1.42.0), edgeR (v3.26.8), impute (v1.60.0), Cell Ranger (v6.0.1), demuxlet (v2), Seurat (v4.0.3, 4.1.0), dsb (v0.3.0), Circlize (v0.4.14), ComplexHeatmap (v2.10.0), eNetExplorer (v1.1.3), SpectroFlo (v2.2.0), scRepertoire (v1.4.0), DESeq2 (v1.34.0), MASS (v7.3-53), FlowJo (v10)

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

Data

Policy information about [availability of data](#)

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

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

Raw and processed data from the whole blood bulk RNA-seq and single-cell CITE-seq are available from the NCBI Gene Expression Omnibus, accession numbers GEO: GSE194378 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE194378>) and GEO:GSE206265 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE206265>), respectively (will be released to public at time of publication). Additional datasets, including clinical, proteomics, flow cytometry, CITE-seq, and

influenza antibody measurements, are available at: <https://doi.org/10.5281/zenodo.5935845> (will be released to public at time of publication). The influenza infection dataset we utilized was downloaded directly from GEO: GSE68310 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE68310>).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

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

Sample size	n = 73 (COVR-F=17; HC-F=21; COVR-M = 16; HC-M = 19). No sample size calculations were done prior to enrollment, in part because there were no reliable effect size estimates related to the impact of prior COVID-19 infection on vaccine responses. The number of subjects in the study was the number that were able to be recruited during the recruitment period.
Data exclusions	None
Replication	Technical replicates were used to assess technical variability of mRNA expression in bulk RNA-seq and single cell CITE-seq, and genes whose technical variance was higher than a predefined threshold were removed from downstream analysis. In addition, the linear models constructed to assess group differences leveraged technical replicates to control for batch and other technical effects.
Randomization	Because this study looked at the cellular and molecular responses to influenza vaccination, there was no indication to randomize individuals to receive influenza vaccination or placebo. The covariates age, self-reported influenza vaccination history (of the past 10 years) and pre-vaccination (baseline) influenza antibody titers were accounted for in our results as they have been included as co-variates in our statistical model for assessing immune response differences among the groups. We model post-vaccination titer as the outcome variable with influenza vaccination history and pre-vaccination titer included as covariates. This approach has also been used for influenza vaccine evaluation by the Food and Drug Administration (e.g., see https://www.fda.gov/media/135687/download page 27).
Blinding	Blinding was not performed as individuals were recruited based on a known history of SARS-CoV-2 infection. Furthermore, as part of the study, an individual's SARS-CoV-2 infection history was confirmed by antibody test in order to allocate them to the correct subject group. Finally, because this study looked at the cellular and molecular responses to influenza vaccination, we did not administer any placebo vaccine, only an FDA-approved influenza vaccine.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

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

Methods

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

Antibodies

Antibodies used	TotalSeq™-C Human Universal Cocktail, V1.0; Fluorochrome-conjugated antibodies as indicated in Methods Table 1, 2 and 3.
Validation	See Supplementary Tables 11-14 for antibody data, including validation information from the commercial companies who made the products.

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	Lenti-X- 293T cells were obtained from Takara Bio (Cat. No. 632180). 293T-ACE2 cells were obtained from ATCC.
---------------------	---

Authentication	Cell line was checked for expression of ACE2 and validated by FACS analysis. None of the cell lines were authenticated by karyotyping or other genomic techniques.
Mycoplasma contamination	Negative for Mycoplasma
Commonly misidentified lines (See ICLAC register)	No misidentified cell lines were used in the study.

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	COVID-19-recovered males (n=16; ages 21-67; median time since COVID-19 diagnosis: 186 days). COVID-19-recovered females (n=17; ages 23-70; median time since COVID-19 diagnosis: 172 days). Healthy control males (n=19; ages 24-69). Healthy control females (n=21; ages 22-70).
Recruitment	Subjects were recruited from the community using email listserv messages, newspaper advertisements, and follow-up contact from a COVID-19 convalescent plasma study. All subjects were self-selected based on interest in participating in COVID-19-related research, ability to participate in study visits, and exposure to recruitment materials. As influenza vaccine titers and responses are not routinely measured outside of research settings, it is unlikely that individuals who opted to participate in the study introduced systemic bias as a result of prior knowledge about their individual cellular or molecular responses to the influenza vaccine.
Ethics oversight	Institutional Review Board of the National Institutes of Health. Informed consent was obtained from all study participants prior to the onset of study procedures.

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

Clinical data

Policy information about [clinical studies](#)

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

Clinical trial registration	NCT04025580
Study protocol	19-I-0126 (Systems analyses of the immune response to the seasonal influenza vaccine). Trial protocol described on clinicaltrials.gov .
Data collection	Subjects were recruited between August and December 2020. Study visits occurred at the National Institutes of Health (NIH) Clinical Center (CC) in Bethesda, Maryland, USA. Blood samples were collected by phlebotomy staff at the NIH CC. Samples were collected between September 2020 and April 2021.
Outcomes	Primary outcome to be measured was serum microneutralization titers to the four influenza strains in the FDA-approved seasonal influenza vaccines used in the study (Flucelvax and high dose Fluzone).

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Whole blood was collected from study participants, PBMCs were isolated using a modified Ficoll procedure, and cells were frozen and stored at -80C. Thawed PBMC were washed in RPMI culture medium containing 50U/ml benzonase nuclease and then washed by PBS. Cells were incubated with LIVE/DEAD Fixable Blue Dye, which was used to exclude dead cells from analysis. Cells were incubated with fluorochrome-conjugated B, H1, H3 probes and fluorochrome-conjugated antibodies against IgM, IgA, CD21, CD85J, FCRL5, CD20, IgG, CD38, CD14, CD56, CD3, CD27, CD71, CD19, IgD etc as indicated in the Method for 30 min at 4 C in the dark. After incubation with antibodies for 30 minutes, cells were washed two times with FACS buffer (0.1%BSA/PBS (pH7.4)) and fixed in 1% paraformaldehyde.
Instrument	Cytek Auora (Cytek Biosciences), BD Symphony (BD Biosciences)
Software	SpectroFlo version 2.2.0 (Cytek Biosciences) and BD FACSDiva Software were used for collecting the data and FlowJo version 10 (BD Biosciences) was used for flow cytometry data analysis.

Cell population abundance

Post-sort fractions not evaluated.

Gating strategy

Gating strategies provided in Supplementary Information Figures 1, 3 and 6.

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