TITLE: High temporal resolution transcriptomic profiling delineates distinct patterns of interferon response following Covid-19 mRNA vaccination and SARS-CoV2 infection

AUTHORS

Darawan Rinchai[1#*], Sara Deola[1#], Gabriele Zoppoli[2,3§], Basirudeen Syed Ahamed Kabeer[1§], Sara Taleb[4§], Igor Pavlovski[1±], Selma Maacha[1±], Giusy Gentilcore[1], Mohammed Toufiq[1], Lisa Mathew[1], Li Liu[1], Fazulur Rehaman Vempalli[1], Ghada Mubarak[1], Stephan Lorenz[1], Irene Sivieri[3,5,6], Gabriella Cirmena[2], Chiara Dentone[2], Paola Cuccarolo[3], Daniele Giacobbe[2,5], Federico Baldi[5], Alberto Garbarino[3], Benedetta Cigolini[3], Paolo Cremonesi[7], Michele Bedognetti[8], Alberto Ballestrero[2,5], Matteo Bassetti[2,5], Boris P. Hejblum[9], Tracy Augustine[1], Nicholas Van Panhuys[1], Rodolphe Thiebaut[9], Ricardo Branco[1], Tracey Chew[10], Maryam Shojaei[11,12], Kirsty Short[13,14], Carl Feng[15,16], PREDICT-19 consortium, Susu M. Zughaier[17], Andrea De Maria[2,5], Benjamin Tang[12,18], Ali Ait Hssain[19], Davide Bedognetti[1&], Jean-Charles Grivel[1&] and Damien Chaussabel[1&*]

1. Research Branch, Sidra Medicine, PO Box 26999, Doha, Qatar
2. IRCCS Ospedale Policlinico San Martino, Genoa, Italy
3. Department of Internal Medicine and Medical Specialties, University of Genoa, Genoa, Italy
4. Division of Genomics and Translational Biomedicine, College of Health and Life Sciences, Hamad Bin Khalifa University, Doha, Qatar
5. Division of Infectious Diseases, Department of Health Sciences, University of Genoa, Genoa, Italy

25 6. Department of Experimental and Clinical Medicine, School of Internal Medicine,
26 University of Florence, Florence, Italy
27 7. Emergency Department, E.O. Ospedali Galliera Genova, Italy
28 8. Azienda Sanitaria Locale 3 Genovese, Genova, Liguria, Italy
29 9. Univ. Bordeaux, Department of Public Health, Inserm U1219 Bordeaux Population Health
30 Research Centre, Inria SISTM, F-33000 Bordeaux, France
31 10. Sydney Informatic Hub, The University of Sydney, Sydney, Australia
32 11. Nepean Clinical School, University of Sydney, Sydney, New South Wales, Australia;
33 12. Westmead Institute for Medical Research, Westmead, New South Wales, Australia.
34 13. The University of Queensland, School of Chemistry and Molecular Biosciences, St Lucia,
35 Brisbane, Queensland, Australia;
36 14. Australian Infectious Diseases Research Centre, The University of Queensland, Brisbane,
37 Queensland, Australia.
38 15. School of Medical Sciences, Faculty of Medicine and Health, The University of Sydney,
39 Sydney, NSW, Australia;
40 16. Tuberculosis Research Program, Centenary Institute, The University of Sydney, Sydney,
41 NSW, Australia.
42 17. College of Medicine, QU Health, Qatar University, PO Box 2713, Doha, Qatar
43 18. Westmead Institute for Medical Research, Westmead, New South Wales, Australia.
44 19. Medical Intensive Care Unit, Hamad General Hospital, PO BOX 3050, Doha, Qatar
45 #, §, ±, & These authors contributed equally to the work
46 * To whom correspondence may be addressed:
47 drinchai@sidra.org
48 dchaussabel@sidra.org

ABSTRACT:

Knowledge of the factors contributing to the development of protective immunity after vaccination with COVID-19 mRNA vaccines is fragmentary. Thus we employed high-temporal-resolution transcriptome profiling and in-depth characterization of antibody production approaches to investigate responses to COVID-19 mRNA vaccination. There were marked differences in the timing and amplitude of the responses to the priming and booster doses. Notably, two distinct interferon signatures were identified, that differed based on their temporal patterns of induction. The first signature (S1), which was preferentially induced by type I interferon, peaked at day 2 post-prime and at day 1 post-boost, and in both instances was associated with subsequent development of the antibody response. In contrast, the second interferon signature (S2) peaked at day 1 both post-prime and post-boost but was found to be potently induced only post-boost, where it coincided with a robust inflammation peak. Notably, we also observed "post-prime-like" ($S1^{++}$,$S2^{0/+}$) and "post-boost-like" ($S1^{++}$,$S2^{++}$) patterns of interferon response among COVID-19 patients. A post-boost-like signature was observed in most severely ill patients at admission to the intensive care unit and was associated with a shorter hospital stay. Interestingly, severely ill patients who stayed hospitalized the longest showed a peculiar pattern of interferon induction ($S1^{-/0}$,$S2^{+}$), that we did not observe following the administration of mRNA vaccines. In summary, high temporal resolution profiling revealed an elaborate array of immune responses elicited by priming and booster doses of COVID-19 mRNA vaccines. Furthermore, it contributed to the identification of distinct interferon-response phenotypes underpinning vaccine immunogenicity and the course of COVID-19 disease.

73

74 **INTRODUCTION**

75 COVID-19 vaccines are critical to the ongoing efforts to control the SARS-CoV-2 coronavirus
76 pandemic. To date, nine vaccines have received some form of approval for use in humans,
77 and phase III trials are ongoing for an additional 11 vaccines (1). Notable differences exist
78 among the vaccine products in terms of their design and the levels of protection they
79 confer, as well as the type, incidence, and severity of adverse events they may elicit. Gaining
80 a comprehensive understanding of the immunological factors underpinning the different
81 responses to diverse vaccines is a major endeavor. Yet, this knowledge is necessary for
82 guiding timely decisions to modulate vaccination protocols (e.g., the use of different types
83 of vaccines for the priming and booster doses). This information may also assist in matching
84 of individuals with the growing number of available vaccines based on their demographics,
85 health status, or any other relevant clinical/molecular phenotypes. Blood transcriptome
86 profiling, which measures the abundance of transcripts in whole blood and on a system-
87 wide scale, has previously been employed to comprehensively profile the immune
88 responses elicited by vaccines. Notably, this approach identified innate immune signatures
89 arising within hours after administering widely used, as well as experimental, vaccines (2). In
90 a recently published report, Arunachalam et al. described the blood transcriptome profiles
91 measured following BNT162b2 mRNA COVID-19 vaccination (3). They reported the presence
92 of an interferon (IFN) signature 1 day after the priming vaccination that was no longer
93 detectable on day 7. They further found that a more complex IFN/inflammatory response
94 was present 1 day after the administration of the booster dose.

95 However, as the use of mRNA vaccines in humans is new, the precise kinetics of the
96 immune response they elicit have not been characterized. Therefore, through this study, we

97 endeavored to profile the transcriptional changes occurring on a systems-wide scale prior to
98 the administration of the first vaccine dose and for 9 consecutive days thereafter, with
99 subjects also collecting samples for deep serological profiling at three time points. The same
100 sampling and profiling schedule being repeated to assess the response to the second dose
101 of the vaccine. Together, this work permitted the precise delineation of a well-orchestrated
102 immune response to COVID-19 mRNA vaccines and identified marked differences in the
103 magnitude, nature, and timing of the transcriptional signatures elicited by prime and boost
104 vaccination. Most notably, differences in temporal patterns of responsiveness revealed
105 distinct components of the interferon response, which is known to play a key role in
106 controlling SARS-CoV-2 infection (4) and was also found here to associate with the
107 subsequent development of the antibody response post-vaccination.
108
109
110 **RESULTS**
111 **Study design, implementation and serological profiling**
112 We successfully recruited a cohort of volunteers and implemented a high-frequency
113 sampling protocol following the administration of the priming and booster doses of COVID-
114 19 mRNA vaccines. The response to each dose was ascertained via the collection and
115 profiling of samples at 10 consecutive daily timepoints: immediately before vaccination and
116 for 9 days thereafter. We collected samples for serological profiling at three time points:
117 prior to vaccination and on days 7 and 14 post-vaccination. A self-sampling blood collection
118 protocol was implemented so that subjects could extract small volumes (50 µl) of RNA-
119 stabilized blood at the required frequency (the approach is described in the Methods
120 section and an earlier publication (5)). RNA sequencing profiles were generated using a cost-

121 effective 3′-biased library preparation protocol (Lexogen QuantSeq). We also established a
122 self-collection of capillary blood samples for measuring the levels and types of COVID-19-
123 specific antibodies. Antibody profiles were generated from samples collected by Volumetric
124 Absorptive Micro Sampling using a multiplexed Bead array established by our team (see
125 Methods for details). Overall, 23 subjects were enrolled in the study, and the characteristics
126 of this cohort are reported in **Table 1**. In total, 23 subjects received either two doses of the
127 Pfizer/BioNTech mRNA vaccine (BNT162b2, N = 19) or two doses of the Moderna mRNA
128 vaccine (N = 4). Among those 23 subjects, six had recovered from COVID-19 in the months
129 preceding the administration of the first vaccine dose. In total 440 RNA sequencing profiles
130 were generated, and this extensive dataset was shared publicly in GEO with the accession
131 number GSE190001. The serological profiles included reactivity to Spike proteins from seven
132 human coronaviruses (including SARS-CoV-2), as well the Nucleo and Envelope proteins, the
133 receptor-binding domain, and the stabilized S-trimer of SARS-CoV-2. The data are provided
134 in **Supplementary File 1**. The seroreactivity to each of these antigens was dissected by
135 measuring the total IgG, total IgA, and IgM, as well as the finer-scale IgG and IgA subtypes.
136 Serological profiling data showed a rise in the levels of antibodies in the plasma of the
137 subjects post-vaccination (**Figure 1A**), and this included antibodies specific for the SARS-
138 CoV-2 Spike protein, which is targeted by COVID-19 vaccines. No responses to the Envelope
139 protein were measured. Some cross-reactivity was observed with the SARS Spike protein.
140 Notably higher antibody levels were induced after the first dose in individuals who had been
141 previously infected with the virus (**Figure 1B**).

142 Altogether, the implementation of this protocol established the feasibility of
143 obtaining stabilized-RNA blood samples from study-subjects post-vaccination at high
144 temporal frequencies. We generated a large dataset using a cost-effective RNA-sequencing

145 protocol that served as the basis for subsequent analyses presented in this paper and was
146 deposited in a public repository. A detailed map of the serological profiles of the subjects
147 enrolled in the study was obtained that permitted us to explore the possible associations
148 between blood transcriptional responses and vaccine immunogenicity.
149
150 **The post-prime interferon response peaks at day-2 and correlates with antibody response**
151 Vaccines can elicit innate immune responses that are detectable systemically via blood
152 transcriptome profiling. But not all of them do, as for instance in the case of the aluminum-
153 adjuvanted Hepatitis B vaccine (6). Our first question, therefore, was whether
154 transcriptional changes could be observed during the first few days following the
155 administration of COVID-19 mRNA vaccines.
156 Analyses were carried out employing a fixed repertoire of 382 transcriptional modules
157 (BloodGen3) that we had recently established and characterized functionally (7)(see
158 methods section for details). Module responses were determined across all time points, and
159 the differential gene-set enrichment functions of the dearseq R package were run to assess
160 whether changes observed throughout the 9 days post-prime were statistically significant
161 (8). This analysis identified significant temporal changes for 22 of the 382 modules
162 constituting the BlooGen3 repertoire (**Supplementary File 3**).
163 Only seven modules were found to be changed at any given time point during the first 3
164 days following the administration of the priming dose of the vaccine (**Figure 2A**). The
165 abundance of four modules was consistently increased across these time points, and all four
166 belonged to the module aggregate A28. Each "module aggregate" regrouped sets of
167 modules that showed consistent abundance profiles across a reference set of 16 disease
168 cohorts that were employed for the construction of the BloodGen3 repertoire (see methods

169 for details). The module aggregate in question ("Aggregate A28") is comprised of six
170 modules. As described in detail in one of our recent publications, all six are associated with
171 interferon responses (7). The gene composition of the modules and the functional
172 annotations are provided herein (relevant information is provided in **Supplementary File 2**
173 and can be accessed interactively via: https://prezi.com/view/E34MhxE5uKoZLWZ3KXjG/).
174 The responses observed on days 1 and 2 post-prime were mapped onto fingerprint grid
175 plots, where modules occupy a fixed position and are arranged by aggregate, with each
176 aggregate occupying a given row (**Figure 2A**). Time-course gene-set enrichment analysis
177 confirmed that changes observed over time for the four A28 modules were significant. The
178 p-values for the two remaining modules from Aggregate A28 did not reach significance. The
179 response profiles of the A28 modules showed a peak on day 2 post-vaccination. This was
180 also visible on a heatmap showing responses at each timepoint across individual subjects
181 (**Figure 2B**). We next examined whether this signature correlated with antibody responses
182 measured 14-days post -as well as at 14-days post-boost (**Figure 2C**). For this, correlation
183 analyses were run at the module level within Aggregate A28 using, as the endpoint, fold-
184 changes in antibody levels on days 7 and 14 post-prime and days 7 and 14 post-boost
185 relative to the pre-vaccination baseline (immediately prior to the administration of the first
186 dose of COVID-19 mRNA vaccines). "Significance hotspots" were identified when most
187 modules within a given aggregate reached correlation significance thresholds. In the case of
188 the post-prime interferon signature, we identified a significance hotspot on days 2 and 3
189 post-prime for a subset of three interferon modules, M10.1, M15.127, and M83, while a
190 fourth module, M15.86, also displayed significant correlations across all antibody types, but
191 only on day 2.

192 Thus, we found that the administration of the priming dose of mRNA vaccine induces a
193 potent innate response characterized by an increase in the abundance of transcripts
194 associated with interferon responses. It was followed by more marked changes over the
195 subsequent days, as indicated by the rise in the number of responsive modules beyond the
196 third day.
197
198 **A decrease in inflammation is accompanied by an increase in adaptive immune response**
199 **genes on day 5 post-prime**
200 We were next interested in characterizing the changes occurring beyond the first 3 days
201 following administration of the priming dose. In total, 18 modules displayed changes on day
202 4 post-prime, of which 12 showed a decrease in abundance. These modules belonged to
203 three aggregates that have been associated with inflammation (A31, A33, A35). Most
204 changes were observed on day 4, but for some modules, changes were apparent starting on
205 day 3 and continued beyond day 4, day 5, or even day 6 (**Supplementary Figure 1**). In our
206 earlier work, modules within the BloodGen3 Aggregate A35 were associated with systemic
207 inflammation mediated by neutrophils and were found to constitute a common
208 denominator across a wide range of pathologies in which systemic inflammation is present
209 (9). The association of A35 with inflammatory processes was also ascertained based on the
210 results of the functional profiling analyses and the restriction in transcript expression in the
211 reference datasets (7). Detailed functional annotations can be accessed via interactive circle
212 packing charts: https://prezi.com/view/7Q20FyW6Hrs5NjMaTUyW/). Module Aggregate A33 has
213 not been investigated as extensively in any of our prior studies but was clearly associated
214 with inflammation via functional profiling
215 (https://prezi.com/view/VBqKqHuLWCra3OJOIZRR/).

216 The number of responsive modules post-prime peaked on day 5, with a total of 42
217 modules showing differences in comparison to the pre-vaccination baseline (**Figure 3**). At
218 this timepoint, most modules showed an increase in abundance (29 were increased and 13
219 decreased). Some of those modules belonged to aggregates that were associated with
220 adaptive immunity, most notably A27, which is associated with plasmablast responses
221 (three out of five modules were responsive at this timepoint). Others were mapped to
222 module aggregates for which immunological significance is not yet well defined; these
223 included Aggregate A24 (associated with oxidative phosphorylation, in which 6 out of 11
224 modules were responsive) and Aggregate A26 (in which 4 out of 15 modules were
225 responsive, none of which are functionally annotated). Other immune-relevant modules
226 found to be increased at this timepoint are associated with T-cells (M12.6 from aggregate
227 A1). Detailed functional annotations for module aggregate A1 can be accessed via
228 https://prezi.com/view/sxap39tKxkmCNTTNIlVO/). The five modules in Aggregate A27 are
229 functionally associated with plasmablasts, as determined from the restriction of expression
230 observed for the genes forming this aggregate in a reference dataset, including a wide range
231 of cell populations (contributed by Monaco et al. (14)) and by the presence of the
232 plasmablast marker CD38 and other associated genes (*IGJ*, *TNFRSF17*, *TXNDC5*) in one of the
233 modules (M12.15). Detailed annotations for A27 can be accessed via
234 https://prezi.com/view/GgIiA0K9kSFHbpVj2I85/. Notably, these signatures appeared to be
235 transient, and no modules were increased on day 6 post-prime.

236 Taken together, our data showed a peak in the number of responsive modules on
237 day 5 post-prime. The signature at the module repertoire-level appeared relatively modest.
238 A decrease in the abundance of transcripts associated with inflammation was accompanied
239 by an increase in the abundance of transcripts associated with adaptive immune responses.

240 Notably, the latter was more subdued, and it also appeared earlier than typically seen in
241 vaccination responses, in which plasmablast responses are usually observed around day 7
242 post-vaccination (10–12).
243
244 **A post-boost interferon signature peaks on day 1 and correlates with antibody responses**
245 After delineating temporal responses post-prime, we examined changes after the
246 second dose of COVID-19 mRNA vaccines. Time course gene set enrichment analysis
247 identified significant temporal changes for 311 of 382 modules comprising the BloodGen3
248 repertoire (**Supplementary File 4**). After the booster dose, the peak number of responsive
249 modules occurred on day 1, with 261 responsive modules or about two-thirds of the 382
250 modules constituting the BloodGen3 repertoire (**Figure 4A**). This number decreased sharply
251 afterwards, with 115 responsive modules on day 2 and only 9 responsive modules on day 3.
252 The kinetic and amplitude of the post-boost response contrasted markedly with that
253 observed post-prime, when, as described above, the number of responsive modules after
254 the first dose instead peaked on day 4, with changes found in 42 modules at that timepoint.
255 As seen from the fingerprint grid plot, the day 1 post-boost response was extensive
256 and polyfunctional (**Figure 4B**). An overall decrease in abundance was observed for
257 aggregates broadly associated with lymphocytic cells (Aggregates A1-A8) and increased for
258 module aggregates associated with myeloid cells, inflammation, and circulating erythroid
259 cells (Aggregates A33-A38). In addition, a marked increase in the abundance of modules
260 associated with interferon responses was also observed (Aggregate A28). We next
261 compared the day 1 response fingerprint of the COVID-19 mRNA booster vaccine to
262 fingerprints derived from reference datasets encompassing a wide range of infectious
263 diseases (these cohorts are described in previously published work (7,13); the respective

264 blood transcriptome fingerprint collections are accessible via a dedicated web application:
265 https://drinchai.shinyapps.io/BloodGen3Module/). In addition, we more recently employed
266 the BloodGen3 module repertoire in the analysis of two COVID-19 blood transcriptome
267 datasets: one cohort comprising 77 Covid-19 patients with disease severities ranging from
268 mild and moderate to severe (the "PREDICT-19 consortium Italian cohort dataset") (14),
269 while the second cohort comprised 40 COVID-19 patients recruited at the time of admission
270 to the intensive care unit (ICU) ("IMPROVISE cohort whole blood dataset"). These high-level
271 comparisons showed, firstly, that the extent of the changes associated with the day 1
272 response to the second dose of the COVID-19 mRNA vaccine was consistent with that
273 observed in some patient cohorts with acute infections (**Figure 4B**). More specifically, they
274 were found to most resemble the responses seen in a cohort of subjects with influenza
275 infection, with a marked interferon response (A28) and an inflammation signature (A33,
276 A35). At a higher level, these response patterns were also generally consistent with those
277 observed in patients with a COVID-19 infection. However, the changes that occurred in
278 response to the vaccinations were not as extreme as those found, for instance, in patients
279 with sepsis or with the most severe form of COVID-19 (i.e., the IMPROVISE dataset) (most
280 notably for inflammation [A33, A35] and erythroid cell responses [A36-A38]).
281 Overall, the BloodGen3 transcriptome fingerprint observed on day 1 after the
282 second dose contrasted markedly with that observed on day 1 post-prime, which was
283 comparatively muted. The interferon response signature was indeed the only common
284 denominator between the responses to the first and second doses, as it was observed in
285 both cases in the first few days following administration of the vaccine. We therefore began
286 to dissect the post-boost response by examining this interferon response signature in more
287 detail.

288 Following the administration of the booster dose, the interferon response was
289 noticeably sharper in comparison to the post-prime response and peaked on day 1 instead
290 of day 2 (**Figure 5A**). This was illustrated by the difference in the maximum average module
291 response, which was close to 50% of the constitutive transcripts on day 2 post-prime and
292 greater than 80% on day 1 post-boost.

293 We decided to then perform hierarchical clustering to identify the subsets of modules
294 within the A28 aggregates that were grouped together based on patterns of transcript
295 abundance across all subjects and timepoints. Two sets of three modules each were, thus,
296 identified within the A28 aggregate. The first set comprised modules M8.3, M10.1, and
297 M15.127 (referred to as A28/S1), and the second set comprised modules M16.64, M13.17,
298 and M15.86 (referred to as A28/S2). Interestingly, we observed post-prime that, while
299 modules in A28/S1 peaked on day 2, those belonging to A28/S2 peaked on day 1 (**Figure**
300 **5B**). Furthermore, A28/S1 modules showed an extended peak post-boost, with day 2 levels
301 being almost identical to those of the day 1 peak, while A28/S1 modules peaked sharply on
302 day 1, with levels decreasing rapidly thereafter. These findings suggest that both sets of
303 modules measured distinct types of interferon response. Indeed, the reference public
304 datasets in which responses to type 1 interferon were measured in vivo indicated that
305 A28/S1 modules are likely to represent type 1 interferon responses (**Figure 5B**), while we
306 postulated that A28/S2 modules represent a type 2 interferon response. Enrichment of
307 literature terms associated with interferon or anti-viral response was previously found to be
308 more pronounced in the case of modules forming the A28/S1 set (7). Indeed, these
309 comprise some of the better recognized "canonical" interferon response genes, such as
310 Oligoadenylate Synthetase family members (OAS1, OAS2, OAS3, OASL), Interferon Induced
311 Protein family members (IFI6, IFI27, IFI35, IFI44, IFI44L), as well as Interferon Induced

312 Protein With Tetratricopeptide Repeats family members (IFIT1, IFIT3, IFIT5). Modules
313 forming the A28/S2 set comprise instead most notably members of the Nuclear Antigen
314 family members SP100, SP110 and SP140, which are associated with interferon gamma
315 signaling, as well as transcription factors IRF9 and STAT2. Composition and functional
316 annotations for A28 modules can be explored further at:
317 https://prezi.com/view/E34MhxE5uKoZLWZ3KXjG/.
318 Finally, a strong association was found between the post-boost interferon signature and
319 the subsequent development of an antibody response. Indeed, positive correlations were
320 observed for all six A28 modules that reached significance on days 1, 2, and 3 post-boost.
321 Notably, this differed from the post-prime interferon response, for which significance was
322 reached only for four of the six modules and only on days 2 and 3.
323 Taken together, the high temporal resolution profiling results permitted the delineation
324 of distinct patterns of post-prime and post-boost interferon responses. The timing of the
325 responses observed at the individual module level contributed to the definition of the two
326 distinct sets of interferon modules. One set was associated with responses to type I
327 interferon in vivo and dominated the post-prime response, with a peak on day 2. The post-
328 boost response showed a strong induction of both sets and also peaked on day 1.
329
330 **Inflammation and erythroid cell signatures peak sharply on day 1 post-boost**
331 We continued the dissection of the day 1 post-boost signature, focusing this time on
332 responses associated with inflammation and circulating erythroid cell precursors.
333 Aggregates A33 and A35, which are associated with inflammation, tended to
334 decrease from day 4 through day 6 post-prime but displayed a contrastingly sharp and
335 transient increase in abundance post-boost. Indeed, a well-delineated response peak was

336 observed on day 1 post-boost for both the A33 and A35 modules (**Figure 6**), but in contrast
337 to the interferon response, it did not extend beyond the first day. Three distinct response
338 patterns were identified via hierarchical clustering among the 21 modules that formed
339 aggregate A35. The "Boost A35/S1" set comprised five modules, while "Boost A35/S2" and
340 "Boost A35/S3" comprised ten and six modules, respectively. The distinction between those
341 three A35/inflammation module sets was rather more subtle than was the case for the
342 A28/interferon sets. Indeed, all three module sets peaked on day 1 post-boost. Differences
343 were rather apparent in the inflection of changes measured on days 2 and 3 post-boost and
344 in the "recovery phase", as abundances appeared to dip below baseline and progressively
345 rise to reach pre-vaccination levels. The underlying biological factors driving the grouping of
346 the modules to those three distinct sets could not be identified at this time.

347 Modules for three aggregates broadly associated with erythroid cell signatures also
348 displayed a sharp but transient increase in transcript abundance on day 1 post-boost.
349 However, the abundance tended to dip afterwards, with a lower peak on day 4 post-boost,
350 before recovering by day 7. Functionally, this signature was found to be most prominently
351 associated with immunosuppressive states, such as late-stage cancer or pharmacological
352 immunosuppression (13), which is consistent with published functional studies (15,16). We
353 also found such signatures were associated with more severe manifestations in babies
354 infected with Respiratory Syncytial Virus (RSV) (13). Finally, we did not find evidence of an
355 association between the day 1 post-boost inflammation or erythroid cell signatures and the
356 antibody responses.

357

358 **A plasmablast signature peaks on day 4 post-administration of the booster dose and**
359 **correlates with antibody responses**

360 After the booster dose, the number of responsive modules peaked sharply on day 1,
361 then rapidly subsided beyond day 2, with the number of responsive modules on days 3, 4, 5,
362 and 6 being reduced to 8, 11, 3, and 2, respectively. Yet, changes within this later timeframe
363 are meaningful, as they specifically concern the set of five modules comprising aggregate
364 A27, which is associated with the presence of antibody-producing cells in the peripheral
365 blood.
366 Three of the five A27 modules showed significant alterations after the booster dose
367 (M16.60, M13.32, M12.15) (**Figure 6**). The proportion of differentially expressed transcripts
368 in each module was relatively modest (with an average of 15% at the peak of response),
369 especially in comparison with the interferon signatures described above (with an average of
370 >80% for some modules at the response peak). The trajectories of the five A27 modules
371 appeared relatively consistent, with only one of the modules (M15.110) showing a different
372 pattern, i.e., a peak on day 6, slightly above the levels observed on day 4. We also examined
373 the association of this post-boost plasmablast signature with the antibody response and
374 found a significant association starting from about day 3 and lasting until day 7 post-boost
375 (**Figure 6**).
376 In summary, COVID-19 mRNA vaccination induced a plasmablast response that peaked
377 at a post-vaccination timepoint that is not routinely included in sampling protocols (day 4).
378 Instead, plasmablast responses have been reported around day 7 post-vaccine
379 administration (e.g., with influenza or pneumococcal vaccines (11)). However, this type of
380 signature could not be detected following COVID-19 mRNA vaccination at this timepoint.
381 We also found an association between this post-boost plasmablast signature and the
382 development of an antibody response.
383

384 **Patterns of interferon responses consistent with those measured day 1 post-boost are**
385 **found in some COVID-19 patients and associated with shorter intensive care unit stays**
386 An important objective of this study was to identify key determinants of vaccine
387 immunogenicity. We found that only a few of the blood transcriptional signatures that have
388 been identified were associated with the development of antibody responses following
389 COVID-19 mRNA vaccination. Logically, the plasmablast signature that we have just
390 described was one of them, and the other two were the interferon signatures observed
391 post-prime and post-boost. We surmised that if such interferon signatures were observed in
392 COVID-19 patients, they may also be associated with the development of a protective
393 immune response and favorable clinical outcome. This might especially be the case for the
394 post-boost interferon response, since a significant rise in neutralizing antibodies was
395 reportedly observed only after the second dose of the vaccine.
396 Therefore, we sought to next determine whether "post-prime-like" and "post-boost-
397 like" interferon response signatures could be observed among the COVID-19 patients, and
398 whether these were associated with a more favorable course of the disease. We relied on
399 blood transcriptome data from the two cohorts of COVID-19 patients described above,
400 namely the "PREDICT-19 consortium Italian cohort" comprising 77 patients with a wide
401 spectrum of disease severity and the "IMPROVISE Whole Blood Dataset" comprised
402 exclusively of patients admitted to the ICU.
403 Focusing first on the PREDICT-19 Italian cohort, we used the response values for the
404 six interferon modules from Aggregate A28 to map individual COVID-19 patient samples
405 along with post-vaccine samples on the same t-SNE plot (**Figure 7**). First, we confirmed that
406 there was no apparent separation of the vaccination and COVID-19 patient cohorts and that
407 batch correction was therefore not warranted before proceeding with comparative analyses

408 (**Supplementary Figure 2**). This is consistent with the results of meta-analyses we have
409 previously conducted at the module level (13). To help with the interpretation, k-means
410 clustering was performed using the consolidated set of samples, resulting in the formation
411 of eight distinct clusters (**Figure 8**). Next, we examined the distribution of samples from the
412 vaccine and COVID-19 cohorts across the tSNE plot and among the eight clusters.
413 Timepoints at which an interferon response was detectable in vaccinated subjects were of
414 particular interest. Indeed, day 1 and day 2 post-prime samples (P1, P2), while preferentially
415 found in Clusters 1 and 5, appeared to be distributed across the entire t-SNE plot. This is in
416 contrast with day 1 and day 2 post-boost vaccination samples (B1, B2), which were almost
417 exclusively found in Cluster 5. A set of COVID-19 patients also co-localized in Cluster 5, while
418 others were found scattered across clusters, especially Clusters 1, 2, 6, and 3. Interferon
419 responses were detectable in all of these clusters, but with important nuances. For one,
420 samples from Cluster 5 showed by far the most potent responses, with responses seen in
421 most cases across all six interferon modules, which was consistent with the post-boost
422 vaccine response. In comparison, the response was less pronounced in samples from Cluster
423 1, which was dominated by modules associated with type I interferon responses (the
424 A28/S1 set comprising M10.1, M8.3 and M15.127 described above). This A28/S1-high and
425 A28/S2-low phenotype was more pronounced in the day 2 post-prime signature because it
426 corresponded to the peak of the A28/S1 response, which occurred after the peak of the
427 A28/S2 set, and indeed, day 2 post-prime samples were preferentially found in this cluster.
428 Signatures for samples forming Clusters 2 and 6 were not well-defined and were in some
429 cases absent, yet these clusters also included COVID-19 patients. Samples forming Cluster 3
430 displayed a peculiar signature, with an increase in the abundance of modules belonging to
431 the A28/S2 set (M15.64, M13.17, and M15.86) concomitantly with a decrease in modules

432 forming A28/S1. Among the samples forming this cluster, this pattern was most apparent
433 for the COVID-19 patients.

434 To summarize at this juncture, we found that COVID-19 patients can present with
435 distinct interferon response signatures. In subjects of Cluster 5, the pattern is consistent
436 with the response on day 1 to booster doses of COVID-19 mRNA vaccines, which in the
437 vaccination setting, is correlated with the development of a potent antibody response. In
438 subjects from Cluster 1, the response was instead dominated by modules belonging to the
439 A28/S1 set, which are responsive to type I interferons and most consistent with the
440 response on day 2 to the priming dose of the COVID-19 mRNA vaccine. Finally, in COVID-19
441 patients, a third pattern of response was observed that does not have an equivalent in
442 vaccinated individuals or in other pathologies that we have previously examined and was
443 characterized by an increase in A28/S2 modules concomitant with a decrease in A28/S2
444 modules. This latter interferon response pattern was observed in only a small number of
445 COVID-19 subjects, a high proportion of whom presented with a more severe form of the
446 disease.

447 It was difficult to determine whether absent or sub-optimal interferon responses
448 among PREDICT-19 subjects could be due to an impaired response to the infection or simply
449 a milder course of the disease. Thus, we next investigated the patterns of interferon
450 responses among ICU patients enrolled in the IMPROVISE cohort. Employing an approach
451 similar to the one described above, we mapped individual COVID-19 patient samples along
452 with post-vaccine samples on a t-SNE plot on the basis of similarities in the patterns of
453 interferon responsiveness across the six interferon modules (**Figure 8A**). COVID-19 subjects
454 were found to again be distributed throughout multiple clusters. Patients who co-localized
455 with day 1 post-boost vaccine samples tended to have relatively short ICU stays (in Cluster 5

456 = A28/S1^{+},A28/S2^{+}), and only a few patients co-localized with day 2 post-prime samples in
457 Cluster 3, which was characterized by a more prominent A28/S1 signature compared with
458 A28/S2 (A28/S1^{+},A28/S2$^{+/0}$). Contrastingly, distinct patient "hostspots" were found in
459 Clusters 1 and 6 that were characterized by the opposite response pattern, i.e., higher
460 abundance levels for A28/S2 compared with A28/S1 (Cluster 6 = A28/S1$^{0/-}$, A28/S2$^{+/0}$ Cluster
461 1 = A28/S1^{-}, A28/S2^{0}). Notably, these patient clusters tended to have longer ICU stays
462 (**Figure 8B**).
463
464
465
466 **DISCUSSION**
467 Relatively little is known about the types of in vivo immune responses elicited by mRNA
468 vaccines in humans. To address this, we employed bulk blood transcriptomics to map the
469 immune changes taking place in vivo after the administration of priming and booster doses
470 of COVID-19 vaccines in adult volunteers. We did so at a high temporal resolution, collecting
471 small amounts of blood before and 9 consecutive days after the administration of the
472 priming and booster doses of COVID-19 mRNA vaccines. The use of blood transcriptomics
473 eliminated the need to choose a panel of analytes to measure vaccine responses, which is
474 one source of bias. The daily collection and profiling schemes adopted eliminated the need
475 to choose specific timepoints for measuring the response, thus eliminating a second source
476 of bias.
477 Profiling blood transcript abundance post-prime and -booster doses of COVID-19
478 mRNA vaccines at a high temporal resolution revealed a well-orchestrated sequence of
479 immune events (**Figure 9**). The immune signatures elicited following the administration of

480 the two doses of mRNA vaccines differed drastically. Relatively modest changes were
481 observed post-prime that manifested primarily as the induction of interferon-response
482 signatures that were detectable over the first 3 days following the injection of the first dose.
483 This was followed by a more subtle response that could be attributed to the priming of the
484 adaptive response between days 4 and 6. Indeed, a decrease in the abundance of
485 transcripts for modules associated with inflammation was observed over these 3 days,
486 which comprised an increase in transcripts associated with plasma cells and T-cells on day 5.
487 No further changes were detected beyond day 6. After the booster dose, the plasmablast
488 response peaked on day 4. Notably, in studies assessing blood transcriptional responses to
489 vaccines, the peak plasmablast response is typically observed on day 7, as it is for instance
490 with influenza and pneumococcal vaccines (10–12). As a result, sampling schedules in
491 common use are designed to capture changes on day 1, 7, and sometimes day 3, but would
492 miss the peak of the adaptive response to COVID-19 mRNA vaccines observed in our study.
493 In addition to eliminating potential blind spots, high-frequency sampling and profiling also
494 permit the precise resolution of signatures that show the complex kinetics of a response; for
495 instance, the erythroid cell signature peaks sharply post-boost and recedes well below
496 baseline over several days before recovering. The trajectory of this signature may be of
497 significance in the context of vaccination, as we recently described its association with
498 immunosuppressive states, such as late-stage cancer and maintenance therapy in liver
499 transplant recipients (13). In the same work, we found this signature to be strongly
500 associated with the development of more a severe disease in subjects with acute
501 respiratory syncytial virus infection; and we furthermore putatively associated this signature
502 with populations of circulating erythroid cells found to possess immunosuppressive
503 properties (15).

504 Arunachalam et al. previously described the elicitation of qualitatively distinct innate
505 signatures on day 1 following the administration of priming and booster doses of COVID-19
506 mRNA vaccines, with the former inducing an interferon response and the latter a mixed
507 response that also presented an inflammatory component (3). Our findings are consistent
508 with these earlier observations, and employing a high-frequency sampling and profiling
509 protocol permitted us to further dissect those responses. Most notably, while interferon
510 responses appear *a priori* as the common denominator between the post-prime and post-
511 boost responses, the temporal pattern of response that we observed indicates that these
512 are, in fact, qualitatively distinct. This was evidenced by the differences in the timing of the
513 response peak, which corresponded to day 2 post-prime and day 1 post-boost. The kinetics
514 of the response post-boost being therefore most consistent with what is observed following
515 injection of a single dose of influenza vaccine (11). Interestingly, a further investigation of
516 the patterns of response among the six modular components of the interferon responses
517 (module Aggregate A28) identified two distinct sets of modules. The first one (A28/S1),
518 which appeared to be preferentially associated with type 1 interferon responses, peaked on
519 day 2 post-prime. The post-prime response of the second set of interferon modules was
520 more modest and peaked on day 1 instead of day 2. Both sets showed strong responses
521 post-boost, with both peaking on day 1, but again showing distinct kinetics with a more
522 sustained response being observed in the case of A28/S1. Association with antibody
523 responses measured on day 14 post-boost followed a similar pattern, with a preferential
524 association with A28/S1 observed post-prime on days 2 and 3, and an association with both
525 A28/S1 and A28/S2 post-boost on days 1 and 2. We have described, in our earlier report,
526 that distinct interferon modules could be employed to stratify patients with systemic lupus
527 erythematosus (17). Here, we sought to specifically determine whether COVID-19 patients

528 present with “post-prime-like” (i.e., A28/S1+) or “post-boost-like” (i.e., A28/S1+, A28/S2+)
529 patterns of interferon response, since those were associated with the subsequent
530 development of antibody responses. This was even more relevant in the context of COVID-
531 19 disease, as it has been reported that failure to induce interferon responses is associated
532 with worse disease outcomes (4,18–20). In the first patient cohort, which constituted
533 patients with predominantly mild or moderate pathology, both phenotypes were indeed
534 observed, along with a third “atypical” phenotype that was not observed post-vaccination
535 that is dominated instead by A28/S2 (with A28/S1 abundance low or even decreased).
536 Notably, in a second patient cohort made up exclusively of ICU patients, both A28/S1+
537 A28/S2+ (post-boost-like) and A28/S2+ (atypical) phenotypes were also observed, with the
538 latter being clearly associated with extended lengths of stay in the ICU. However, the
539 atypical phenotype did not appear to be preferentially associated with death in this setting,
540 which may be due to the supportive care provided to the patients. While, overall, our
541 observations support the notion that robust interferon responses are associated with a
542 more favorable course of the disease, they also show that the response may be of a
543 different type for other patients, rather than being altogether defective (i.e., A28/S2+). One
544 possibility is that this peculiar response pattern is associated with the presence of
545 endogenously produced autoantibodies that neutralize interferon, as has been previously
546 described (20,21). However, it is not possible to be conclusive on this point and further
547 investigations are warranted.

548 Other points remain to be elucidated. This includes the timing of the adaptive response
549 to mRNA vaccines, which appears to rise and peak several days earlier than is normally
550 observed in responses to other vaccines (± 7 day peak). Also remaining to be determined is
551 the priming mechanism underpinning the robust polyfunctional response observed on day 1

552 post-boost. One of the questions being whether or not such a response, which would
553 typically be considered to be innate, is antigen-specific. Interestingly, in that respect, the
554 subjects who were previously infected but recovered from COVID-19 did not display a
555 noticeable day 1 inflammatory response, and their immune systems behaved like those of
556 naïve individuals. However, the number of recovered subjects was small, and the study was
557 not designed to directly address this question, hence further investigations will also be
558 necessary. Notably, the greater amplitude of responses observed post-boost and the
559 presence of an inflammatory component is also consistent with previous reports of the
560 increase in the incidence of side effects/discomfort following COVID-19 mRNA vaccine
561 booster doses (22,23).

562 The work presented here could help inform the design of studies aiming to investigate
563 vaccine responses. Indeed, a decrease in sequencing costs provides laboratories an
564 opportunity to employ transcriptome profiling approaches in novel ways, such as the
565 implementation of high-temporal frequency protocols. An advantage of the delineation of
566 transcriptome responses at high-temporal resolution is that it is doubly unbiased, i.e., there
567 is no need to select transcripts for inclusion in a panel because RNA sequencing measures all
568 transcript species present in a sample. Similarly, there is no need to select specific
569 timepoints for assessing the vaccine response, as all timepoints were profiled within a given
570 time frame. An obvious advantage of the approach is that it permits the removal of
571 potential blind spots, and the detection of changes that are potentially otherwise missed by
572 more sparse sampling protocols. We have shown here that the peak plasmablast response
573 following COVID-19 mRNA vaccination occurs unexpectedly at around day 4 and wanes by
574 day 7, when such a response would normally be expected after an influenza vaccination.
575 This is significant, as this signature was also shown to correlate with the extent of the

576 antibody response. A somewhat less obvious advantage is that, as exemplified by the
577 interferon signatures, high-frequency profiling data permitted us to more precisely resolve
578 the vaccine response. For instance, in this case, we were able to show that the interferon
579 response peaked on day 2 post-prime but on day 1 post-boost. To be more precise, a
580 component of the post-prime interferon response peaked on day 2, while the other peaked
581 on day 1. Some of the practical elements that may contribute to making the routine
582 implementation of high-temporal resolution transcriptomics approaches viable include, as
583 mentioned earlier, a substantial decrease in the cost of RNA sequencing, especially lower
584 depth, 3'-biased methodologies. Recent publications showed, through down-sampling
585 analysis, that much fewer deep reads than usual are adequate for biomarker discovery
586 projects (24), with the lower costs permitting larger sample sizes or, as in this case, a higher
587 sampling frequency. Another consideration is the availability of solutions for the in-home
588 self-collection of samples. This is the case for the collection of RNA-stabilized blood,
589 although our homegrown method could be improved and interesting solutions are being
590 put forward (25). As we have shown, it is possible to implement the self-collection of
591 samples for serology profiling within a vaccinology study.

592 There were several limitations to our study. While the sample size was adequate for an
593 initial discovery phase, a larger study cohort would help to better resolve inter-individual
594 variations. The dataset we generated, however, has been made available for reuse, and it
595 should be possible to integrate and consolidate this dataset with those generated in follow-
596 on studies by us and others (13). Follow on studies would need to be purposedly designed
597 to formally address specific questions, for instance, comparing responses in individuals who
598 had previously been exposed to SARS-CoV-2 with those in naïve individuals. It would also be
599 interesting to compare responses elicited by the Pfizer/BioNTech and Moderna vaccines,

600 which was not possible in our study due to the small numbers of individuals that received
601 the Moderna vaccine. Indeed, although we hoped it would be possible to obtain more
602 balanced sample sizes for a more detailed comparison, the speed at which the vaccinations
603 were rolled out among our target population of healthcare workers meant we had very little
604 control over the number of volunteers that received the different types of vaccines or their
605 status as naïve or previously exposed individuals. It would also have been particularly
606 interesting to enroll patients from different age categories, especially the elderly
607 population, but this again proved impossible.

608 In conclusion, a considerable number of COVID-19 vaccines have already been approved
609 for use in humans, and others are currently in phase III trials (>20) (26). The data presented
610 herein suggest that high-temporal-frequency blood transcriptomics are an appropriate way
611 to precisely map the types of responses elicited by the different COVID-19 vaccines and,
612 thus, benchmark a wide range of platforms that have been employed for vaccine
613 development. Characterizing and comparing vaccine response profiles in populations that
614 may not respond optimally to vaccines should also prove particularly informative (e.g., in
615 the elderly, immunosuppressed, and during pregnancy). Finally, the most likely roll-out of a
616 systematic third injection of the booster dose should offer opportunities to further dissect
617 the immune responses to mRNA vaccines in individuals who have been previously
618 immunized.

619

620 **METHODS**

621 **Subject recruitment:**

622 COVAX Cohort: We enrolled adult subjects eligible to receive a COVID-19 vaccine and who
623 were willing to adhere to the sampling schedule. The protocol was approved by Sidra

624 Hospital IRB (IRB number 1670047-6), and all participants gave written informed consent.
625 Inclusion criteria matched the clinical eligibility for receiving the vaccine, and the only
626 exclusion criterium was subjects that already had a first dose of any COVID-19 vaccine.
627 Twenty-three subjects were enrolled, and the median age was 38 years (29-57); 20 of the
628 subjects received the Pfizer vaccine and 3 the Moderna vaccine. The demographics, health
629 status at accrual, and vaccination side effects are shown in Table 1. Vaccination and booster
630 intervals were typically 21 days for Pfizer and 29 days for Moderna.
631
632 IMPROVISE cohort: Adult subjects with severe COVID-19 were enrolled to this cohort under
633 the Hamad Medical Corporation IRB approval (MRC-05-007). Blood samples were collected
634 at multiple timepoints during patients' ICU stay (timepoint 1 was taken at ICU admission;
635 timepoints 1 to 4 were 7 days apart). Subjects with burn and trauma, immunological
636 diseases, receiving immunosuppressive treatment, with other immune-related conditions,
637 or with a previous COVID-19 infection were excluded. For this analysis, 40 severe COVID-19
638 patients were included, with a median age of 52 (range = 30 to 92). The clinical parameters
639 of those patients included gender, ICU and hospital stay, mechanical ventilation duration,
640 ECMO initiation, comorbidities, outcomes (death/recovery), nosocomial infection onset, and
641 plasma therapy. Samples were also collected from control subjects who were adults and did
642 not: 1) present with an infectious syndrome during the last 90 days, 2) experience extreme
643 physical stress within the last week, 3) received during the last 90 days a treatment based
644 on: antivirals; antibiotics; antiparasitic; antifungals; 4) received within the last 15 days, a
645 treatment based on non- steroidal anti-inflammatory drugs; 5) received during the last 24
646 months a treatment based on: immunosuppressive therapy; corticosteroids; therapeutic
647 antibodies; chemotherapy and 6) a person with history of: innate or acquired immune

648 deficiency; hematological disease; solid tumor; severe chronic disease; surgery or
649 hospitalization within the last 2 years; pregnancy within the last year; participation to a
650 phase I clinical assay during the last year; participation to a phase I clinical assay during the
651 last year; pregnant or breastfeeding women; person with restricted liberty or under legal
652 protection.

653 PREDICT-19 Cohort:

654 The "Predicting disease progression in severe viral respiratory infections and COVID-19"
655 (PREDICT-19) Consortium is an international consortium formed by a group of researchers
656 who share common interests in identifying, developing and validating clinical and/or
657 bioinformatic tools to improve patient triage in a pandemic such as COVID-19 (14). The
658 PREDICT-19 Italian cohort comprises adult subjects with mild, moderate, or severe COVID-
659 19 diagnosed by real-time PCR on nasopharyngeal swab who were consented and enrolled
660 at E.O. Ospedali Galliera, and IRCCS Ospedale Policlinico San Martino, Genoa, Italy (Ethics
661 Committee of the Liguria Region (N.CER Liguria 114/2020-ID 10420). Blood samples were
662 collected during hospitalization. Subjects with burn and trauma, immunological diseases,
663 receiving immunosuppressive treatment for underlying disorders before COVID-19
664 diagnosis, with other immune-related conditions, or with a previous COVID-19 infection
665 were excluded. For this analysis, 10 healthy subjects and 103 COVID-19 patients were
666 included, with a median age of 61.76 (range = 26 to 86).

667

668 **Sampling protocol:**

669 COVAX Cohort: For transcriptomics applications for the COVAX study, after puncturing the
670 skin with a finger stick, 50 µl of blood was collected in a capillary/microfuge tube assembly
671 supplied by KABE Labortechnik (Numbrecht, Germany) containing 100 µl of tempus RNA-

672 stabilizing solution aliquoted from a regular-sized tempus tube (designed for the collection
673 of 3 ml of blood and containing 6 ml of solution; ThermoFisher, Waltham, MA, USA). This
674 method is described in detail in an earlier report (5), and the collection procedure is
675 illustrated in an uploaded video: https://www.youtube.com/watch?v=xnrXidwg83I. Blood
676 was collected prior to the vaccine being administered (day 0), on the same day, and daily
677 thereafter over the next 10 days. This protocol was followed for both the priming and
678 booster doses.
679 For serology applications, 20 µl of blood was collected using a Mitra blood collection device
680 (Neoteryx, Torrance, CA, USA) prior to the vaccine being administered and on days 7 and 14
681 after vaccination with the priming and booster doses.
682 IMPROVISE Cohort: For the IMPROVISE study, samples were collected using PaxGene Blood
683 RNA tubes (BD Biosciences, Franklin Lakes, NJ, USA) at all timepoints and were frozen at -
684 20C until further processing.
685 PREDICT-19 Cohort: For the Italian cohort of the PREDICT-19 study, blood samples were
686 collected during hospitalization by venipuncture during in tubes containing an RNA
687 stabilizing solution (Tempus™ Blood RNA Tube, ThermoFisher, Waltham, MA, USA, Catalog
688 number: 4342792) and frozen at -20C until further processing.
689
690 **Multiplex serological assay**
691 The concentration of selected proteins in the serum samples was measured with a home-
692 built bead array based on carboxymethylated beads with 6 distinct intensities of a UV-
693 excitable dye. Each was individually coupled to 3 SARS-CoV-2 proteins, envelope,
694 nucleoprotein, S-trimer and fragments thereof and the S1 fragment of SARS-CoV S protein.
695 Therefore, the complete array consisted of 6 antigens, including five SARS-CoV-2 antigens

696 (Full Spike Trimer, RBD, Spike S1, Nucleoprotein, and Envelope), as well as the closely
697 related SARS-CoV-S1 protein. The readout of this assay provided the immunoglobulin
698 isotypic response to each of the antigens, i.e., total IgM, total IgG, total IgA, as well as their
699 individual isotypes, IgG1, IgG2, IgG3, IgA1, and IgA2, containing a total of 42 analytes. The
700 assays were performed on filter plates and acquired on a BD-Symphony A5. Ig-binding to
701 the spike protein antigens differed for each bead region. An average of 300 beads per
702 region was acquired, and the median fluorescence intensity for isotype binding was used for
703 the calculations. An antibody response index was calculated by ratioing form each antigen
704 and isotype, the MFI of polled negative blood controls collected prior to June 2018 (Sidra
705 IRB 1609004823) to the MFI obtained for vaccinated donor samples.
706
707 **RNA extraction and QC**
708 RNA was extracted using the Tempus Spin RNA Isolation Kit (ThermoFisher), which was
709 adapted for the handling of small blood volumes. The methodology has been described
710 previously in detail (27). Contaminating DNA was removed using the TurboDNAse kit
711 (ThermoFisher), and RNA was quantitated on a Qubit instrument (ThermoFisher) and QCed
712 using an Agilent 2100 Bioanalyzer (Agilent, Santa Clara, California, USA).
713
714 **RNA sequencing**
715 <u>COVAX & IMPROVISE Cohorts:</u> mRNA-sequencing was performed using QuantSeq 3' mRNA-
716 Seq Library Prep Kit FWD for Illumina (75 single-end) with a read depth of 8M and average
717 read alignment of 79.60%. Single samples were sequenced across four lanes, and the
718 resulting FASTQ files were merged by sample. Quality trimming is performed to remove
719 adapter sequences and polyA tails. Then trimmed reads were aligned to human genome

720 GRCh38/hg38 (Genome Reference Consortium Human Build 38), INSDC Assembly
721 GCA_000001405.28, Dec 2013) using STAR 2.6.1d and featureCounts v2.0.0 was used to
722 generate the raw counts. Raw expression data were normalized to size factor effects using R
723 package DESeq2. All downstream analyses were performed using R version 4.1, unless
724 otherwise specified. Global transcriptional differences between samples were assessed by
725 principal component analysis using the "prcomp" function. Transcriptome profiling data
726 were deposited, along with detailed sample information, into a public repository, the NCBI
727 Gene Expression Ominibus (GEO), with accession ID GSE190001 and BioProject ID:
728 PRJNA785113
729 PREDICT-19 Cohort: Total RNA was isolated from whole blood lysate using the Tempus Spin
730 Isolation kit (Applied Biosystems) according to the manufacturer's instructions. Globin
731 mRNA was depleted from a portion of each total RNA sample using the GLOBINclear™-
732 Human kit (Thermo fisher). Following removal of globin transcripts transcriptome profiles
733 were generated via mRNA sequencing. Then mRNA-sequencing was performed using
734 Illumina HiSeq 4000 Technology (75 paired-end) with a read depth of 60M. Single samples
735 were sequenced across four lanes, resulting FASTQ files were merged by sample. All FASTQ
736 passed QC and were aligned to reference genome GRCh38 using STAR (2.6.1d). BAM files
737 were converted to a raw count's expression matrix using HTSeq
738 (https://github.com/Sydney-Informatics-Hub/RNASeq-DE). Raw count data was normalized
739 using DEseq2. The ensemble ID targeting multiple genes were collapsed (average) and a
740 final data matrix gene was for modular repertoire analysis.
741
742
743

744 **Statistical Analysis**

745 Analyses were conducted using pre-defined gene sets. Specifically, we employed a fixed
746 repertoire of 382 transcriptional modules that were thoroughly functionally annotated, as
747 described in detail in a recent publication (7). Briefly, this repertoire of transcriptional
748 modules (“BloodGen3") was identified based on co-expression, as measured in a collection
749 of 16 blood transcriptome datasets encompassing 985 individual transcriptome profiles.
750 Sets of co-expressed transcripts were derived from the analysis of a large weighted co-
751 clustering network. Downstream analysis results and visualizations were generated
752 employing a custom R package (28). “Module response” is defined as the percentage of
753 constitutive transcripts with a given abundance that were determined to be different
754 between two study groups, or for the same individual in comparison to a given baseline (in
755 this study, pre-vaccination abundance levels). The values therefore ranged from 100% (all
756 constitutive transcripts increased) to −100% (all constitutive transcripts decreased). Only the
757 dominant trend (i.e., increase or decrease in abundance over control/baseline) was retained
758 for visualization purposes on fingerprint grids or fingerprint heatmaps, with red indicating
759 an increase and blue a decrease in abundance. When performing group comparisons (e.g.,
760 cases vs controls for the disease datasets used as reference), the p-value and false discovery
761 rate cutoffs were applied, which are mentioned in the figure legend. When performing
762 longitudinal analyses, the module response is determined by employing fixed fold-change
763 and expression difference cutoffs. Module response values obtained were used for data
764 visualization. Significance was determined for each module using the differential gene set
765 enrichment function of the dearseq R package (8).

766

767

ACKNOWLEDGMENTS:

We wish to acknowledge Michelle Esblaca and Juvilyn Gusi for their assistance with the training of subjects and collection of blood samples, as well as the research subjects for volunteering and making the study possible.

REFERENCES

1. Tregoning JS, Flight KE, Higham SL, Wang Z, Pierce BF. Progress of the COVID-19 vaccine effort: viruses, vaccines and variants versus efficacy, effectiveness and escape. Nat Rev Immunol. 2021 Aug 9;

2. Hagan T, Nakaya HI, Subramaniam S, Pulendran B. Systems vaccinology: Enabling rational vaccine design with systems biological approaches. Vaccine. 2015 Sep 29;33(40):5294–301.

3. Arunachalam PS, Scott MKD, Hagan T, Li C, Feng Y, Wimmers F, et al. Systems vaccinology of the BNT162b2 mRNA vaccine in humans. Nature. 2021 Jul 12;

4. Hadjadj J, Yatim N, Barnabei L, Corneau A, Boussier J, Smith N, et al. Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. Science. 2020 Aug 7;369(6504):718–24.

5. Rinchai D, Anguiano E, Nguyen P, Chaussabel D. Finger stick blood collection for gene expression profiling and storage of tempus blood RNA tubes. F1000Res. 2016;5:1385.

6. De Mot L, Bechtold V, Bol V, Callegaro A, Coccia M, Essaghir A, et al. Transcriptional profiles of adjuvanted hepatitis B vaccines display variable interindividual homogeneity but a shared core signature. Sci Transl Med. 2020 Nov 11;12(569):eaay8618.

7. Altman MC, Rinchai D, Baldwin N, Toufiq M, Whalen E, Garand M, et al. Development of a fixed module repertoire for the analysis and interpretation of blood transcriptome data. Nat Commun. 2021 Jul 19;12(1):4385.

8. Agniel D, Hejblum BP. Variance component score test for time-course gene set analysis of longitudinal RNA-seq data. Biostatistics. 2017 Oct 1;18(4):589–604.

9. Rawat A, Rinchai D, Toufiq M, Marr AK, Kino T, Garand M, et al. A Neutrophil-Driven Inflammatory Signature Characterizes the Blood Transcriptome Fingerprint of Psoriasis. Front Immunol [Internet]. 2020 [cited 2020 Dec 7];11. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2020.587946/full

10. Nakaya HI, Wrammert J, Lee EK, Racioppi L, Marie-Kunze S, Haining WN, et al. Systems biology of vaccination for seasonal influenza in humans. Nat Immunol. 2011 Jul 10;12(8):786–95.

802 11. Obermoser G, Presnell S, Domico K, Xu H, Wang Y, Anguiano E, et al. Systems scale
803 interactive exploration reveals quantitative and qualitative differences in response to
804 influenza and pneumococcal vaccines. Immunity. 2013 Apr 18;38(4):831–44.

805 12. Nakaya HI, Hagan T, Duraisingham SS, Lee EK, Kwissa M, Rouphael N, et al. Systems
806 Analysis of Immunity to Influenza Vaccination across Multiple Years and in Diverse
807 Populations Reveals Shared Molecular Signatures. Immunity. 2015 Dec 15;43(6):1186–
808 98.

809 13. Rinchai D, Altman MC, Konza O, Hässler S, Martina F, Toufiq M, et al. Definition of
810 erythroid cell-positive blood transcriptome phenotypes associated with severe
811 respiratory syncytial virus infection. Clinical and Translational Medicine.
812 2020;10(8):e244.

813 14. Tang B, Shojaei M, Wang Y, Nalos M, Mclean A, Afrasiabi A, et al. Prospective validation
814 study of prognostic biomarkers to predict adverse outcomes in patients with COVID-
815 19: a study protocol. BMJ Open. 2021 Jan 6;11(1):e044497.

816 15. Elahi S. Neglected Cells: Immunomodulatory Roles of CD71+ Erythroid Cells. Trends
817 Immunol. 2019 Mar;40(3):181–5.

818 16. Elahi S, Ertelt JM, Kinder JM, Jiang TT, Zhang X, Xin L, et al. Immunosuppressive CD71+
819 erythroid cells compromise neonatal host defence against infection. Nature. 2013 Dec
820 5;504(7478):158–62.

821 17. Chiche L, Jourde-Chiche N, Whalen E, Presnell S, Gersuk V, Dang K, et al. Modular
822 transcriptional repertoire analyses of adults with systemic lupus erythematosus reveal
823 distinct type I and type II interferon signatures. Arthritis & Rheumatology (Hoboken,
824 NJ). 2014 Jun;66(6):1583–95.

825 18. Galani I-E, Rovina N, Lampropoulou V, Triantafyllia V, Manioudaki M, Pavlos E, et al.
826 Untuned antiviral immunity in COVID-19 revealed by temporal type I/III interferon
827 patterns and flu comparison. Nat Immunol. 2021 Jan;22(1):32–40.

828 19. Zhang Q, Bastard P, Liu Z, Le Pen J, Moncada-Velez M, Chen J, et al. Inborn errors of type
829 I IFN immunity in patients with life-threatening COVID-19. Science. 2020 Oct
830 23;370(6515):eabd4570.

831 20. Bastard P, Rosen LB, Zhang Q, Michailidis E, Hoffmann H-H, Zhang Y, et al.
832 Autoantibodies against type I IFNs in patients with life-threatening COVID-19. Science.
833 2020 Oct 23;370(6515):eabd4585.

834 21. Bastard P, Gervais A, Le Voyer T, Rosain J, Philippot Q, Manry J, et al. Autoantibodies
835 neutralizing type I IFNs are present in ~4% of uninfected individuals over 70 years old
836 and account for ~20% of COVID-19 deaths. Sci Immunol. 2021 Aug 19;6(62):eabl4340.

837 22. Cheng H, Peng Z, Luo W, Si S, Mo M, Zhou H, et al. Efficacy and Safety of COVID-19
838 Vaccines in Phase III Trials: A Meta-Analysis. Vaccines. 2021 Jun;9(6):582.

23. Mulligan MJ, Lyke KE, Kitchin N, Absalon J, Gurtman A, Lockhart S, et al. Phase I/II study of COVID-19 RNA vaccine BNT162b1 in adults. Nature. 2020 Oct;586(7830):589–93.

24. Milanez-Almeida P, Martins AJ, Germain RN, Tsang JS. Cancer prognosis with shallow tumor RNA sequencing. Nat Med. 2020 Feb;26(2):188–92.

25. Haack AJ, Lim FY, Kennedy DS, Day JH, Adams KN, Lee JJ, et al. homeRNA: A self-sampling kit for the collection of peripheral blood and stabilization of RNA [Internet]. 2021 Feb [cited 2021 Oct 30] p. 2021.02.08.430337. Available from: https://www.biorxiv.org/content/10.1101/2021.02.08.430337v1

26. Sadarangani M, Marchant A, Kollmann TR. Immunological mechanisms of vaccine-induced protection against COVID-19 in humans. Nat Rev Immunol. 2021 Aug;21(8):475–84.

27. Syed Ahamed Kabeer B, Tomei S, Mattei V, Brummaier T, McGready R, Nosten F, et al. A protocol for extraction of total RNA from finger stick whole blood samples preserved with TempusTM solution. F1000Res. 2018 Nov 2;7:1739.

28. Rinchai D, Roelands J, Toufiq M, Hendrickx W, Altman MC, Bedognetti D, et al. BloodGen3Module: Blood transcriptional module repertoire analysis and visualization using R. Bioinformatics. 2021 Feb 24;

Table 1: Subject characteristics

Patient ID	Vaccine name	Gender	Age	Ethnicity	Previous COVID-19	Underlying disease	Drugs	Symptoms at prime (Type)	Symptoms at prime (Grade)	Symptoms at boost (Type)	Symptoms at boost (Grade)
PZB1	Pfizer Biontech	Female	38	Asian	Yes	No	no	Myalgia	G1	Fever/Myalgia	G1
PZB2	Pfizer Biontech	Male	47	Caucasian	Yes	No	no	Myalgia	G1	Myalgia	G1
PZB3	Pfizer Biontech	Male	57	Caucasian	Yes	T2D	Metformin, Insulin	Myalgia	G1	Myalgia	G1
PZB4	Pfizer Biontech	Female	34	Indian	No	No	no	None	NA	Chills/Insomnia/Headache/Myalgia/Fa	G3

										tigue	
PZB5	Pfizer Biontech	Male	38	Indian	No	No	no	Myalgia	G1	Myalgia	G1
PZB6	Pfizer Biontech	Female	48	Caucasian	No	No	no	Myalgia/Headache	G1	Myalgia	G1
PZB7	Pfizer Biontech	Female	34	Caucasia/ Arab	No	Hashimoto thyroiditis	no	None	NA	None	NA
PZB8	Pfizer Biontech	Female	29	Arab	No	No	no	Myalgia/Swelling	G1	Fever/Myalgia	G1
PZB9	Pfizer Biontech	Male	41	Arab	No	Allergic rhinitis	no	Myalgia	G1	Myalgia	G1
PZB10	Pfizer Biontech	Female	35	Arab	No	No	no	Myalgia	G1	Fever/Insomnia/Myalgia/ Fatigue	G2
PZB11	Pfizer Biontech	Male	41	Caucasian	No	No	no	Myalgia	G2	Myalgia	G1
PZB12	Pfizer Biontech	Female	34	Indian	No	Hypothyroidism	Levothyroxine	None	NA	Fever/Myalgia	G1
PZB13	Pfizer Biontech	Male	29	Indian	No	No	no	Fatigue	G1	Fever/Heaviness in arm	G2
PZB14	Pfizer Biontech	Female	38	Arab	No	Allergic	no	Myalgia/Headache	G1	Fatigue	G1
PZB15	Pfizer Biontech	Male	43	Arab	No	Hypothyroidism	Levothyroxine	None	NA	Myalgia/Headache	G2
PZB16	Pfizer Biontech	Female	39	Indian	No	No	no	Heaviness in arm	G1	Fever/Myalgia/Fatigue	G2
PZB17	Pfizer Biontech	Female	42	Indian	Yes	T2D, Hypertension	Methformin, Telmisartan	Fever/Headache/Myalgia/Fatigue	G2	Fatigue/Gastritis	G2
MDA18	Moderna	Male	42	Caucasian	No	Hypertension	Amlodipine, Ramipril	Myalgia	G1	Myalgia	G1
PZB19	Pfizer Biontech	Female	39	Caucasian	No	No	no	Myalgia	G1	Headache/Myalgia/Arthralgia	G3
MDA20	Moderna	Female	36	Arab	Yes	No	no	Chills/Myalgia	G2	Chills/Myalgia	G2
MDA21	Moderna	Female	36	Caucasian	No	No	no	Myalgia	G1	Feveer/Skin rash/Myalgia	G2
PZB25	Pfizer Biontech	Female	39	Caucasian	No	No	no	Myalgia	G1	Myalgia	G1
MDA26	Moderna	Female	30	Caucasian	Yes	Asthma	Seretide, Salbutamol	Fever/Headache/Myalgia/Fatigue	G3	Asthma attack/Fever/Myalgia/Fatigue	G3

Footnotes:

T2D: Type 2 Diabetes

Vaccine type and lot, and subjects' characteristics were recorded, including demographic, biometric data, blood group, underlying diseases, drugs usage, and previous COVID-19 disease.

Every subject recorded and graded the symptoms that occurred after the first and second vaccinations doses, according to the NIH "DAIDS AE Grading Table"

864

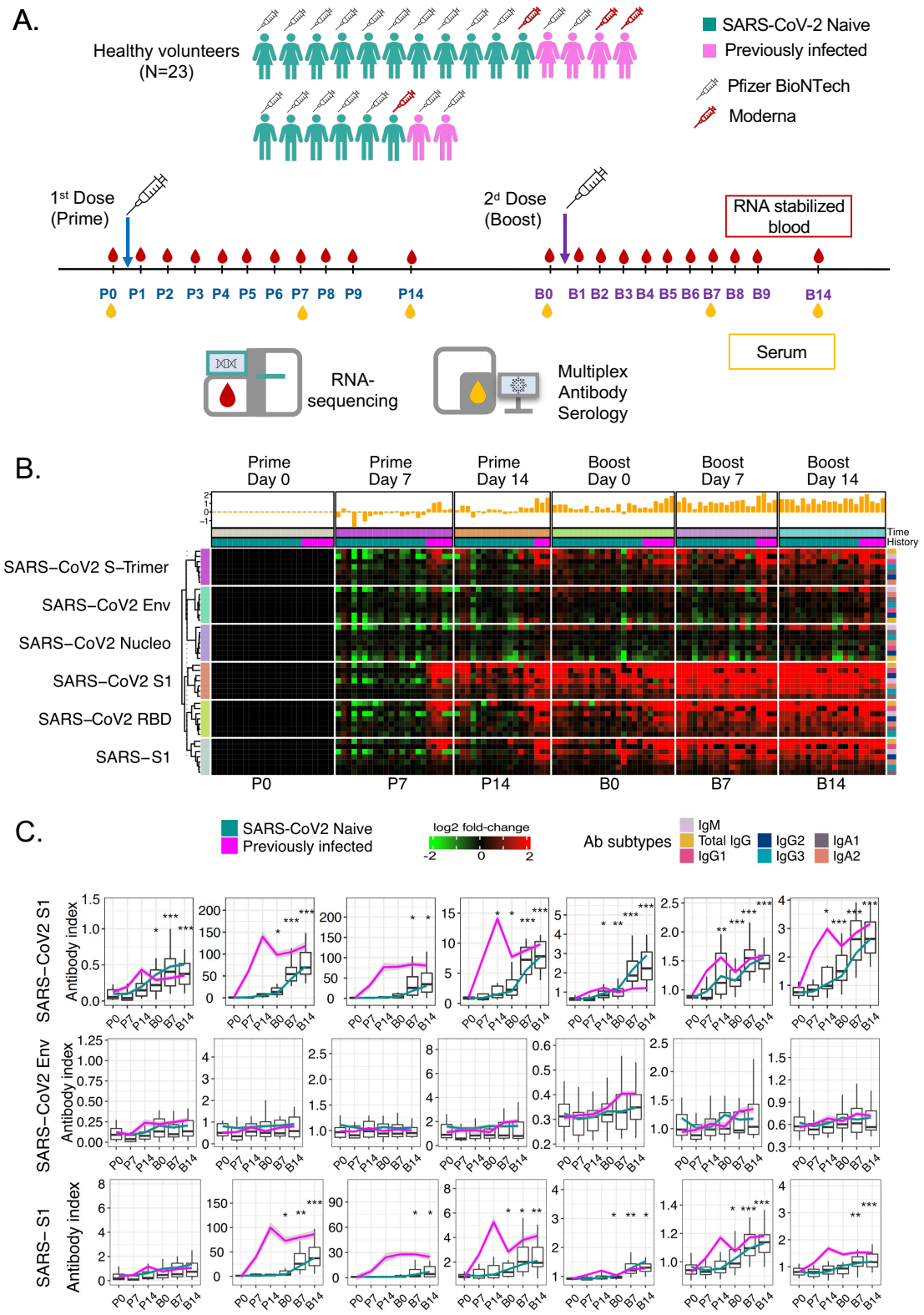

Figure 1: Antibody response to COVID-19 mRNA vaccination. (A) Schematic representation of the study design. (B) The heatmap represents changes in abundance of antibodies specific to several SARS-CoV-2 antigens and control antigens relative to pre-vaccination levels. Red indicates a relative increase and green indicates a relative decrease in abundance. Columns represent subjects, which are arranged by timepoint and have colored tracks at the top indicating whether the subjects were previously infected with SARS-CoV-2. The histogram above represents the average log2 fold-change over baseline for a given column. The rows represent antibodies for which abundances were profiled, which are arranged according to specificity. For each antigen, the different rows represent different types of antibodies. (C) Changes in antibody levels, expressed as an "antibody index", are shown on the box plots, each corresponding to a given antibody type of a given specificity. Lines indicate changes for individuals previously infected with SARS-CoV-2 and who had recovered (in pink) and for naïve individuals (in green). Multiple pairwise tests (paired t-test) were performed comparing antibodies levels to baseline (D0). Asterisks: * represent $p < .01$, **represent $p < .001$, *** represent $p < .0001$.

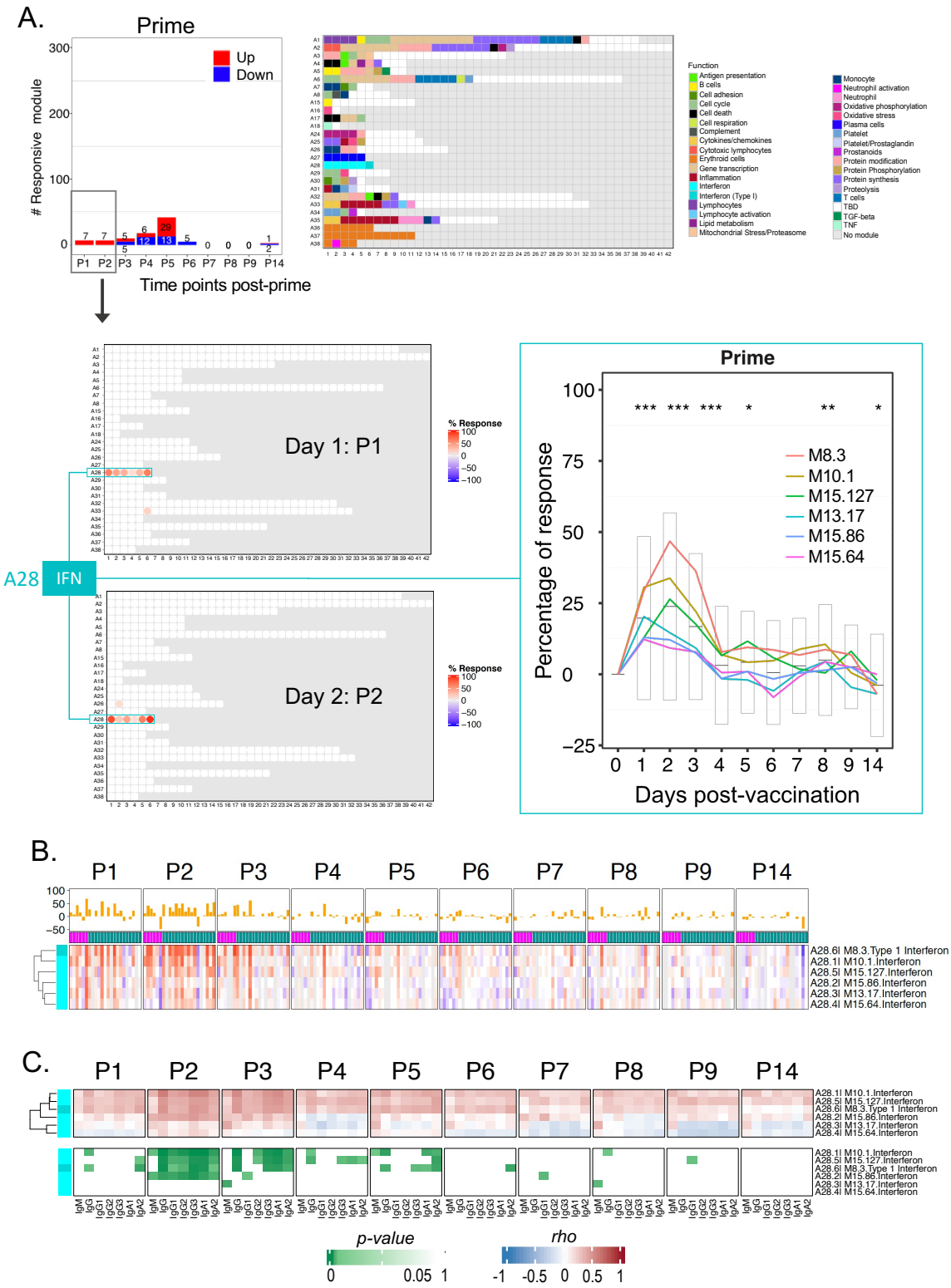

Figure 2: Characterization of the post-prime interferon response signature. (A) The bar graph shows the cumulative module response at the various timepoints following the administration of the priming dose of the vaccine (noted P1-P14). The Y-axis values and numbers on the bars indicate the number of modules meeting the 15% response threshold (out of a total of 382 modules constituting the BloodGen3 repertoire, with percentage response corresponding to the proportion of transcripts predominantly increased or decreased compared to baseline using FDR < 0.1 as the cutoff to determine significance [DESeq2]). The number of modules for which abundance was predominantly increased is shown in red and for those for which abundance was predominantly decreased is shown in blue. The fingerprint grid plots represent the overall module responses on day 1 post-prime (P1) and day 2 post-prime (P2). Modules from the BloodGen3 repertoire occupy fixed positions on the fingerprint grids. They are arranged as rows based on membership to module aggregates (rows A1 through A38). Changes compared to pre-vaccination baseline are indicated on the grid by red and blue spots of varying color intensity, which indicate the

“percentage response” for a given module. The color key at the top indicates the various functions attributed to the modules that are represented on the grid. The response of the six modules comprising aggregate A28 is represented on a line graph that shows the proportion of responsive transcripts for each module across all the post-prime timepoints. For each of module statistical significance for the overall response was determined by time course gene set enrichment analysis. Four of the six A28 modules met significance thresholds FDR < 0.1 (M8.3: p-value = 1.9-e4, FDR = 0.019, M10.1: p-value = 1.9-e4, FDR = 0.019, M15.127: p-value = 1.9-e4, FDR = 0.019, 727 and M15.86: p-value = 3.9-e4, FDR = 0.031) and all six A28 modules p < 0.05 (M13.17: p-value = 1.5-e3, FDR = 0.101 and M15.64: p-value = 0.044, FDR = 0.727). We also ascertained the significance of changes measured post-prime at the level of this module aggregate and at each time point (paired t-test comparing module response at each time point relative to the pre-vaccination baseline; * p<0.01, ** p<0.001, *** p<0.0001). (B) Heatmaps represent proportions of transcripts that changed within the six A28 modules at different timepoints and across different individuals compared to pre-vaccination baseline values. Red indicates that transcripts were predominantly increased over the baseline, and blue indicates that transcripts were predominantly decreased. Rows represent the six A28 modules, which are arranged within an aggregate via hierarchical clustering. Columns represent samples, which are grouped by timepoint and show profiles of individual subjects within each timepoint. (C) The heatmaps represent associations (spearman correlation test) between antibody responses measured 14 days after administration of COVID-19 booster doses and transcriptional responses measured across 9 consecutive days after the priming dose. The heatmap at the top provides the correlation coefficients across multiple days and for each day across multiple subjects, with rows corresponding to the six A28 interferon modules. The heatmap below shows the significance of the correlations shown on the heatmap directly above, with the same ordering of rows and columns.

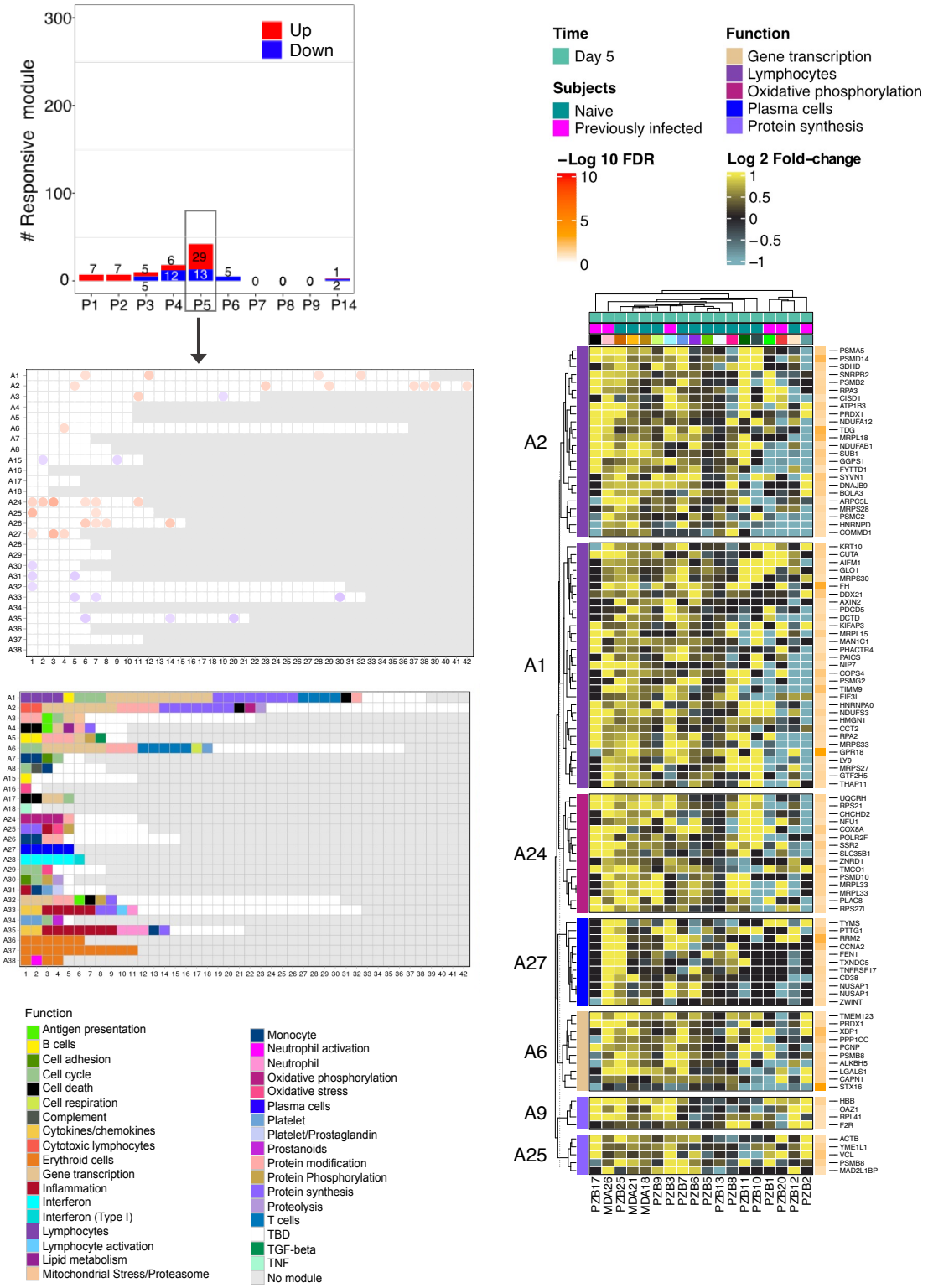

Figure 3: Characterization of responses on day 5 post-prime. (A) The bar graph shows the cumulative module response at the various timepoints following the administration of the priming dose of the vaccine (noted P1-P14). The fingerprint grid plot shows changes elicited by vaccination on day 5 following administration of the priming dose. The position of the modules on the grid is fixed. The percent response of individual modules is represented on the grid by red and blue spots of varying color intensity denoting predominant increase or decrease in abundance, respectively. The percentage response of a given module corresponds to the proportion of transcripts predominantly increased or decreased compared to baseline, meeting a significance cutoff of FDR < 0.1. The color key at the top indicates the various functions attributed to the modules that are represented on the grid. **(B)** The heatmap represents Log2 average fold change in abundance of transcripts constituting sets of modules associated with given functional annotations on day 5 post-vaccination. Rows represent individual transcripts grouped according to the module aggregate they originate from, corresponding to the different rows on the fingerprint grid plot on the left. Each module aggregate is associated with a different function, as indicated by the color key above. The columns on the heatmap represent individual subjects (Pfizer BioNtech = PZB; Moderna = MDA).

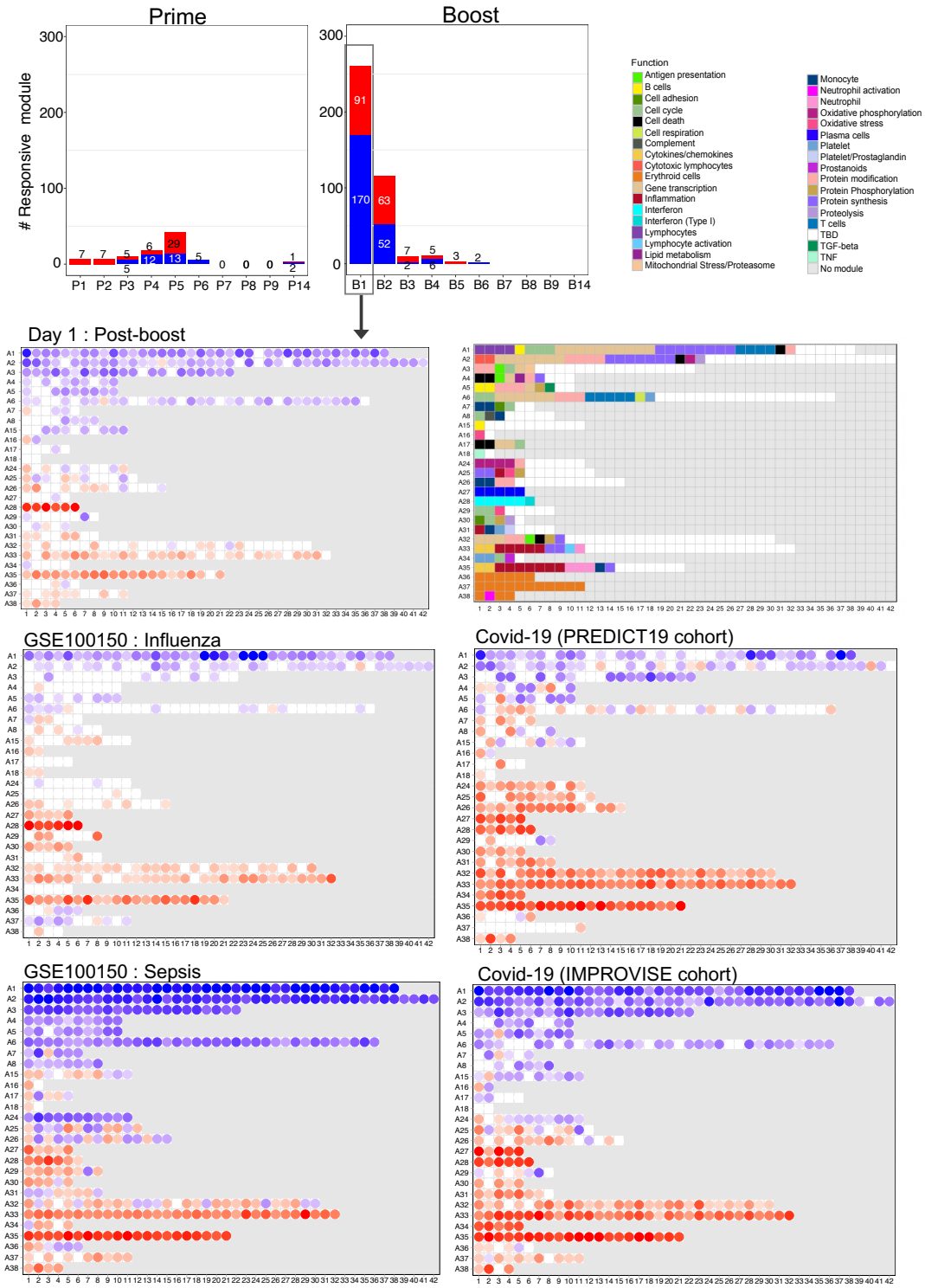

Figure 4: Fingerprint grid plots mapping changes observed on day 1 post-boost and across reference datasets. The bar graphs show the cumulative module response at the various timepoints post-priming and booster doses (noted P1-P14 and B1-B14, respectively). The Y-axis values and numbers on the bars indicate the number of modules meeting the 15% response threshold (out of a total of 382 modules constituting the BloodGen3 repertoire, with percentage response corresponding to the proportion of transcripts predominantly increased or decreased compared to baseline meeting a significance cutoff of DESeq2, FDR < 0.1. The fingerprint grid plots show changes in transcript abundance in a given study group in comparison to baseline (pre-vaccination sample or uninfected control group – with the percent response of individual modules shown by red and blue spots of varying color intensity denoting predominant increase or decrease in abundance, respectively. Changes are shown in the top grid for a group comparison of 1 day after receiving the booster dose of COVID-19 mRNA vaccines with baseline pre-vaccination samples (this study). Grids in the middle and bottom positions show changes for patients with acute infections caused by influenza virus (public dataset) or SARS-CoV-2 (this study) and for patients with bacterial sepsis (public dataset). The color key at the top indicates the various functions attributed to the modules that occupy a fixed position on the grid plot.

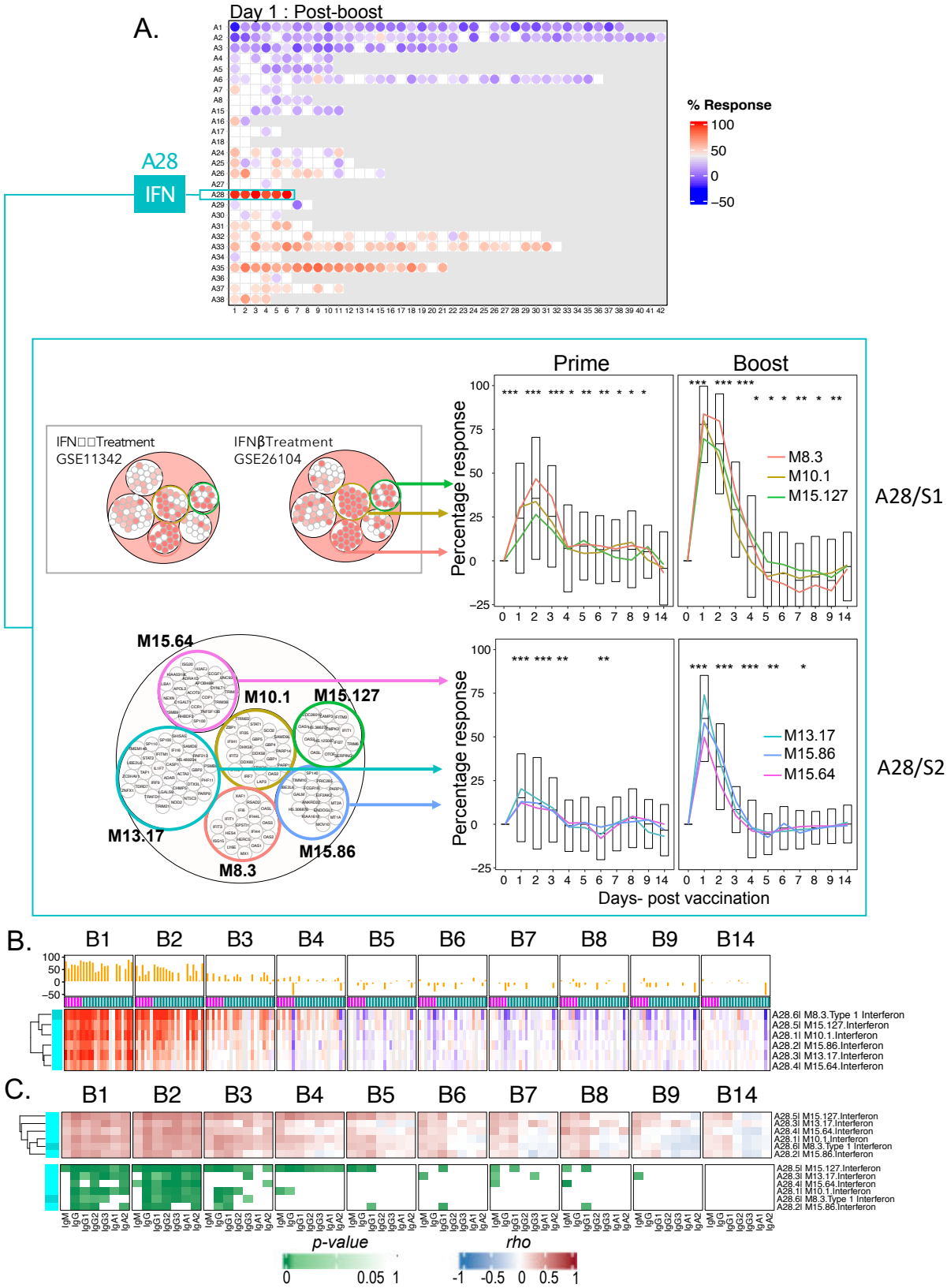

Figure 5: Characterization of the day 1 post-boost interferon response signature. (A) The fingerprint grid plot maps the modular response observed on day 1 post-boost (percent response is determined based on statistical cutoff: DESeq2, FRD < 0.1). The six modules forming the A28 aggregate are highlighted. The line graphs below represent the summarized percentage responses at the module level, encompassing all study subjects (one line per module). Percentage response accounts for the proportion of transcripts for a given module that show differences in abundance post-prime (left) or post-boost (right) compared to baseline pre-vaccination levels. Changes in transcript abundance post-prime and post-boost for two distinct sets of interferon response modules that received the denomination A28/S1 and A28/S2 are plotted on separate graphs. For each of module statistical significance for the overall response was determined by time course gene set enrichment analysis. Significance was reported post-prime in Figure 2. Post-boost all six A28 modules met significance thresholds p < 0.001 and FDR < 0.001 (M8.3: p-

value = 1.9-e4, FDR = 3.6-e4, M10.1: p-value = 1.9-e4, FDR = 3.6-e4, M13.17: p-value = 1.9-e4, FDR = 3.6-e4, M15.127: p-value = 1.9-e4, FDR = 3.6-e4, M15.64: p-value = 1.9-e4, FDR = 3.6-e4 and M15.86: p-value = 1.9-e4, FDR = 3.6-e4). In addition, we ascertained the significance of changes measured post-prime at the level of this module aggregate and at each time point (paired t-test comparing module response at each time point relative to the pre-vaccination baseline; * p<0.01, ** p<0.001, *** p<0.0001). The circle packing plots on the left show module responses at the individual transcript level for two public blood transcriptome datasets. The larger circle below indicates official symbols for the individual transcripts. It also highlights the modules included in A28/S2, which are shown directly on the right. The smaller circles above show changes in abundance of A28 transcripts for two public datasets. One study (GSE11342) measured blood transcriptional response in patients with Hepatitis C infection treated with alpha-interferon (23). The second study (GSE26104) measured transcriptional response in subjects with multiple sclerosis treated with beta-interferon (24). A red circle indicates a significant increase in the abundance of transcripts compared to the pre-treatment baseline (*|fold-change| > 1.5, FDR < 0.1). **(B)** Changes in abundance compared to baseline pre-vaccination levels are represented on a heatmap, with modules as rows and individual samples as columns. The modules are arranged by hierarchical clustering based on abundance patterns across samples. The samples are arranged by timepoints post-prime (top) and post-boost (bottom). **(C)** The heatmaps represent associations (Spearman correlation) between antibody responses measured 14 days after administration of COVID-19 booster doses and transcriptional responses measured across 9 consecutive days after the booster dose. The heatmap on top provides the correlation coefficients across multiple days and for each day across multiple subjects, with rows corresponding to the six A28 interferon modules. The heatmap below shows the significance of the correlations shown on the heatmap on top, with the same ordering of rows and columns.

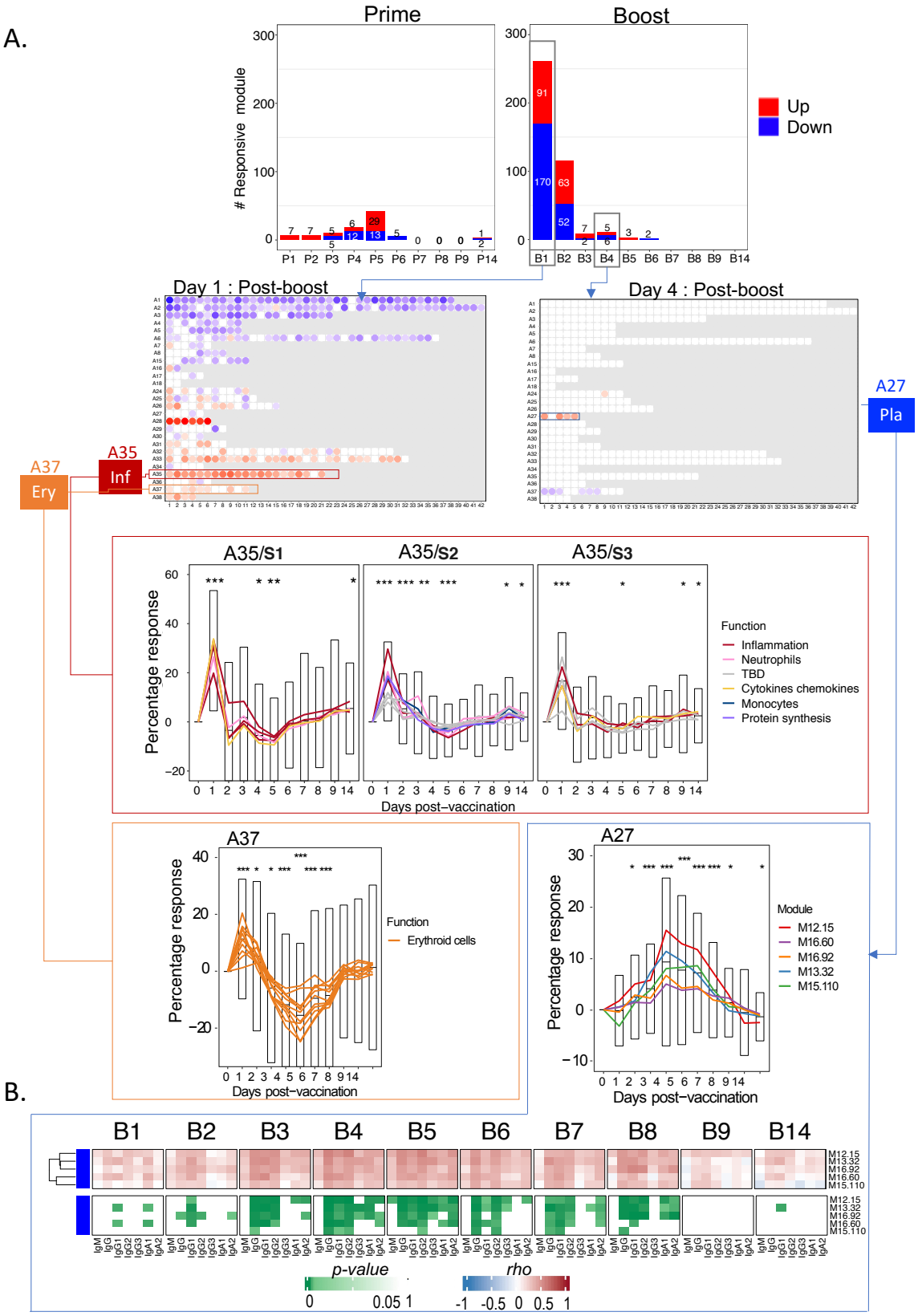

Figure 6: Characterization of post-boost inflammation, erythroid cell, and plasmablast responses. (A) The bar graph at the top represents the number of response modules at any given time point post-prime and post-boost. The fingerprint grid plots show the modules that had changes compared with a fixed visualization and interpretation framework. Changes are shown for the day 1 post-boost timepoint (left) as well as day 5 (right) (percent response is determined based on statistical cutoff: DESeq2, FRD < 0.1). On the left grid, modules belonging to aggregates A35 (associated with inflammation) and A37 (associated with erythroid cells) are highlighted. The profiles of those modules are represented on the line graphs below, which show the average percentage responses of A35 and A37 modules across multiple timepoints. The percentage response for a given module is the proportion of its constitutive transcripts showing significant changes, which ranged from 0% to 100% when transcripts were predominantly

increased to 0% to −100% when transcripts were predominantly decreased. Each line represents the profile of the modules constituting a given aggregate. Line graphs for A35 were split into three sets according to differences in clustering patterns (A35/S1, A35/S2 and A35/S3). On the right grid, modules belonging to aggregates A27 (associated with platelets) are highlighted. The corresponding line graph below represents the changes in abundance of A27 modules over time following administration of the second dose of vaccine. For each module statistical significance for the overall response was determined by time course gene set enrichment analysis using the dearseq R package. For A35, 20 of 21 modules met significance thresholds (p-value < 0.05 and FDR < 0.01). It was also the case in 11 of 11 modules for A37 and 4 of 5 modules for A27 (**Supplementary file 4**). In addition, we ascertained the significance of changes measured post-prime at the level of this module aggregate and at each time point (paired t-test comparing module response at each time point relative to the pre-vaccination baseline; * p<0.01, ** p<0.001, *** p<0.0001). **(B).** The heatmaps represent associations between antibody responses measured 14 days after administration of COVID-19 booster doses and transcriptional responses measured across 9 consecutive days after the booster dose. Specifically, the heatmap at the top represent the correlation coefficients across multiple days and for each day across multiple subjects, with rows corresponding to the five A27 plasmablast modules. The heatmap below shows the significance of the correlations shown on the heatmap at the top, with the same order of rows and columns.

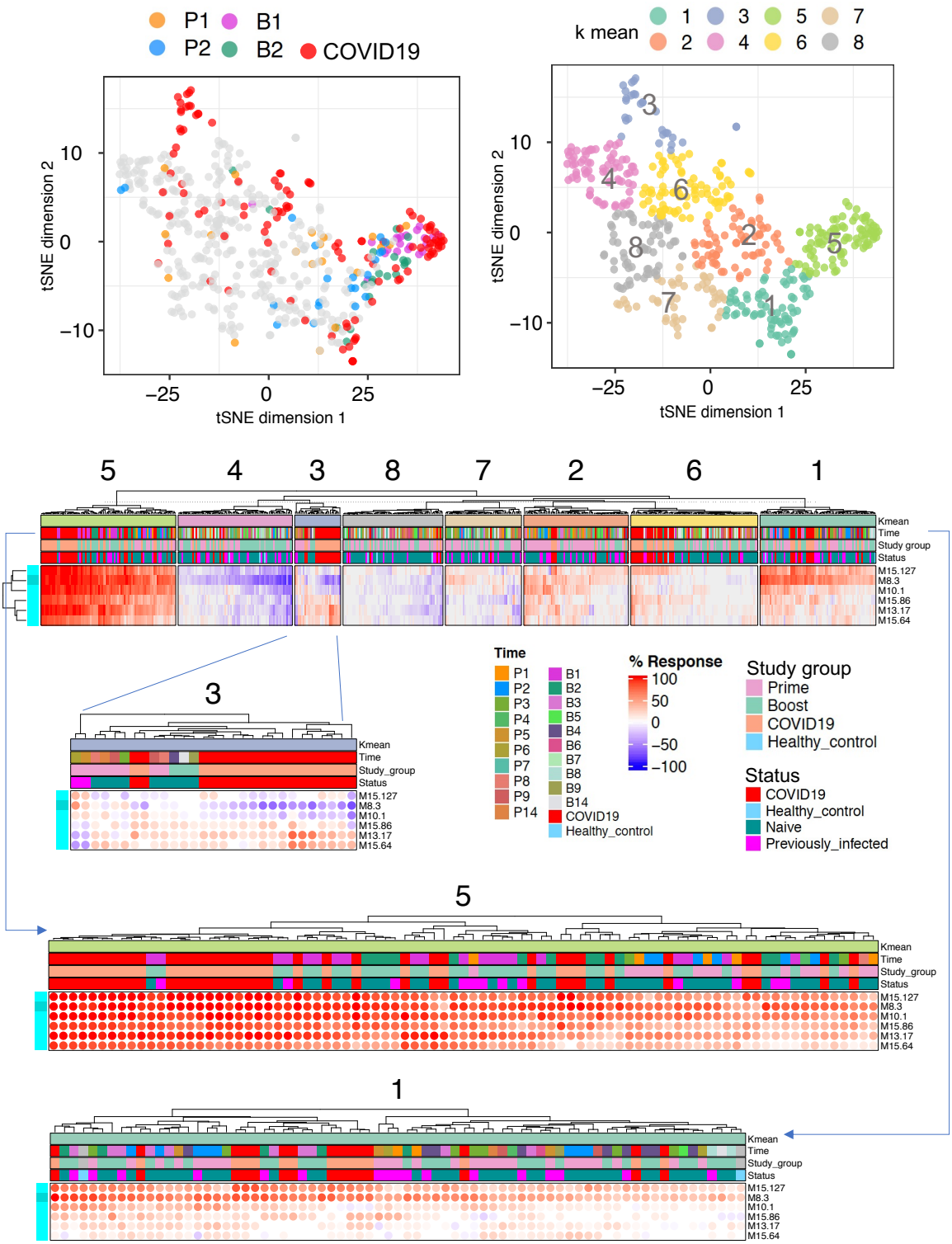

Figure 7: Comparing patterns of interferon response in vaccinated individual and a cohort of COVID-19 patients. Similarities in patterns of interferon response induction across the six modules forming aggregate A28 and among our vaccination cohort and one of our COVID-19 disease cohorts (PREDICT-19 / Italy) are represented on a tSNE plot. COVID-19 samples are shown in red along with specific post-vaccination timepoints (post-prime days 1 and 2 [P1, P2], post-boost days 1 and 2 [B1, B2]). Samples from the consolidated cohorts were partitioned into 8 clusters via k-means clustering, the distribution of which is shown on the tSNE plot on the top right. Patterns of response for the 6 interferon response modules across the 8 sample clusters are shown on a heatmap. The red colors indicate that abundance of transcripts for a given module is predominantly increased with the intensity representing the proportion of constitutive transcripts meeting a given threshold, which at the level of individual samples is a fixed fold change and difference cutoff (|Fold change| > 1.5, and |difference| > 10 in a given sample over its respective pre-vaccination baseline). The blue color denotes a predominant decrease in abundance of constitutive transcripts compared to the same individual's pre-vaccination baseline. Details are shown below for Clusters 3, 5, and 8 in separate heatmaps.

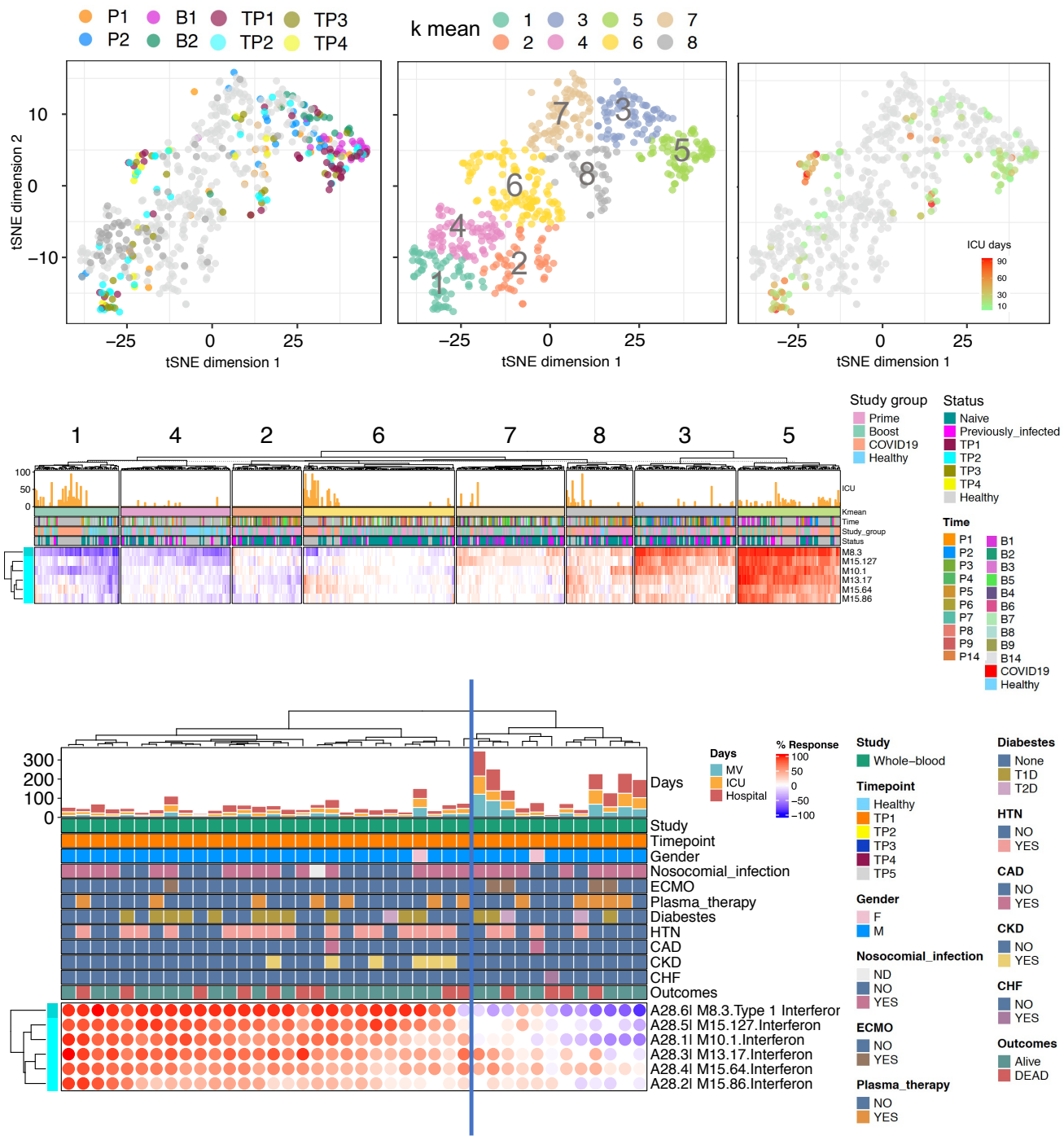

Figure 8: Comparison of interferon response patterns of vaccinated individuals and a cohort of COVID-19 patients with severe disease under intensive care. Similarities in patterns of interferon response induction across the six modules forming aggregate A28 and among our vaccination cohort and one of our COVID-19 disease cohorts (IMPROVISE) are represented in a tSNE plot, onto which disease severity information has been overlaid, along with specific post-vaccination timepoints (post-prime days 1 and 2 [P1, P2], post-boost days 1 and 2 [B1, B2]). Samples from the consolidated cohorts were partitioned into 8 clusters via k-means clustering, the distribution of which is shown on the tSNE plot on the top right. Patterns of response for the 6 interferon response modules across the 8 sample clusters are shown on a heatmap. The red colors indicate that abundance of transcripts for a given module are predominantly increased with the intensity representing the proportion of constitutive transcripts meeting a given threshold, which at the level of individual samples is a fixed fold change and difference cutoff (|Fold change |> 1.5, and |difference| > 10 in a given sample over its respective pre-vaccination baseline). The blue color denotes a predominant decrease in abundance of constitutive transcripts compared to the same individual's pre-vaccination baseline. The heatmap at the bottom shows patterns of interferon responses for COVID-19 patients with severe disease upon admission to the ICU. Multiple clinical parameters are shown on the tracks above. The histogram represents length of stay in the hospital, in the ICU, and time under mechanical ventilation.

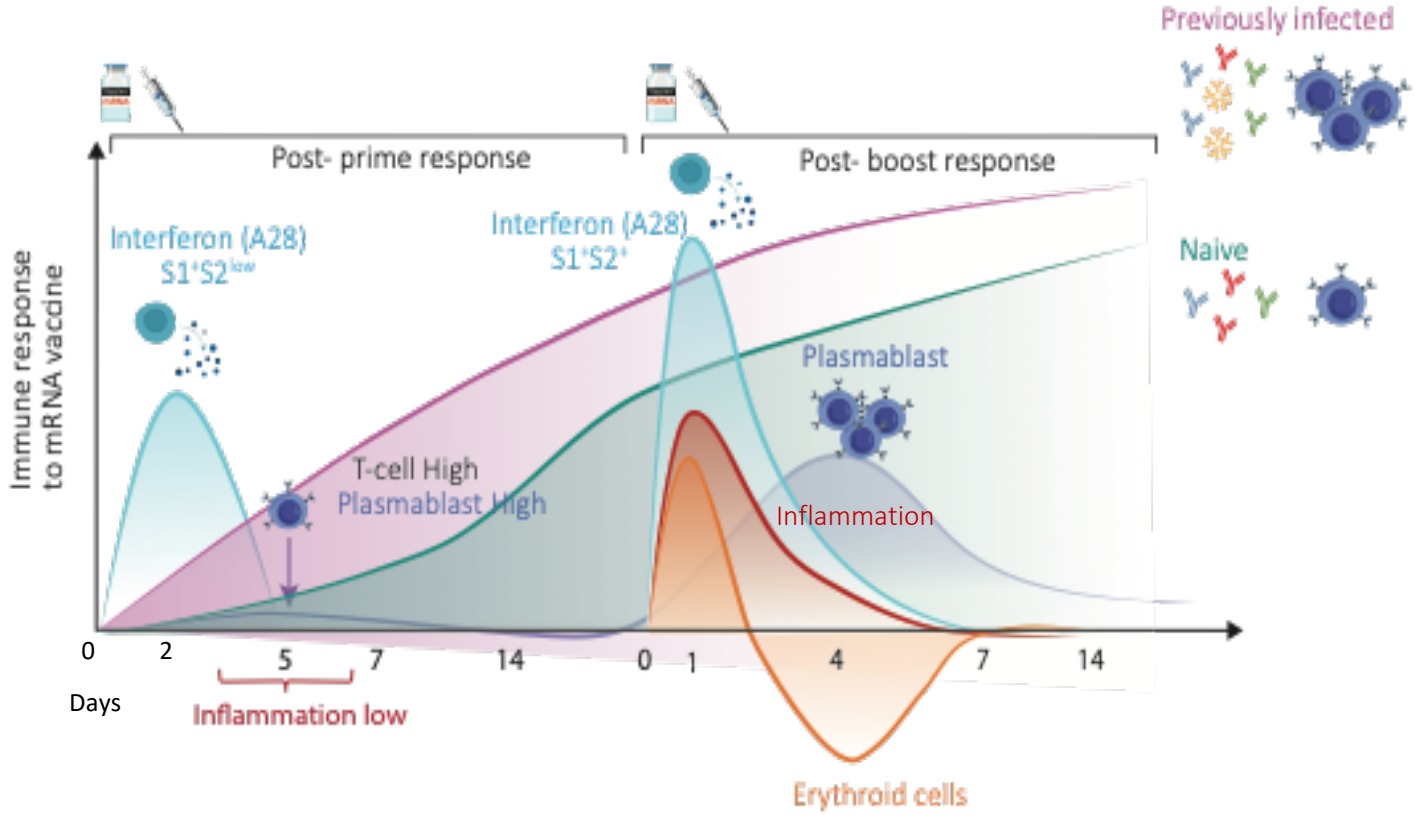

Figure 9: Summary. This diagrammatic representation summarizes the temporal trajectories of blood transcriptional signatures elicited in response to the first and second doses of mRNA vaccines.