

1 **Autoantigen profiling reveals a shared post-COVID signature in fully recovered and Long**
2 **COVID patients**

3

4 Aaron Bodansky,¹ Chung-Yu Wang,² Aditi Saxena,² Anthea Mitchell,² Saki Takahashi,³ Khamal
5 Anglin,⁴ Beatrice Huang,³ Rebecca Hoh,³ Scott Lu,⁴ Sarah A. Goldberg,⁴ Justin Romero,³ Brandon
6 Tran,³ Raushun Kirtikar,³ Halle Grebe,³ Matthew So,³ Bryan Greenhouse,³ Matthew S.
7 Durstenfeld,⁵ Priscilla Y. Hsue,⁵ Joanna Hellmuth,⁶ J. Daniel Kelly,⁴ Jeffrey N. Martin,⁴ Mark S.
8 Anderson,^{7,8} Steven G. Deeks,³ Timothy J. Henrich,⁹ Joseph L. DeRisi,² Michael J. Peluso³

9

10 1 Division of Pediatric Critical Care Medicine, University of California, San Francisco, San
11 Francisco, CA, USA

12 2 Chan Zuckerberg Biohub, San Francisco, CA, USA

13 3 Division of HIV, Infectious Diseases, and Global Medicine, University of California, San
14 Francisco, San Francisco, CA, USA

15 4 Department of Epidemiology and Biostatistics, University of California, San Francisco, San
16 Francisco, CA, USA

17 5 Division of Cardiology, University of California, San Francisco, San Francisco, CA, USA

18 6 Department of Neurology University of California, San Francisco, San Francisco, CA, USA

19 7 Diabetes Center, University of California, San Francisco, San Francisco, CA, USA

20 8 Department of Medicine, University of California, San Francisco, San Francisco, CA, USA

21 9 Division of Experimental Medicine, University of California, San Francisco, San Francisco, CA,
22 USA

23

24

25

26 Corresponding Author: Michael Peluso, MD
27 Division of HIV, Infectious Diseases, and Global Medicine
28 University of California, San Francisco
29 San Francisco, CA, USA
30 Email: michael.peluso@ucsf.edu

31

32 **Keywords:** SARS-CoV-2, COVID-19, post-acute sequelae of SARS-CoV-2 infection (PASC),
33 Long COVID, autoreactivity, autoimmunity, autoantigen

34 **ABSTRACT**

35 Some individuals do not return to baseline health following SARS-CoV-2 infection, leading to a
36 condition known as Long COVID. The underlying pathophysiology of Long COVID remains
37 unknown. Given that autoantibodies have been found to play a role in severity of COVID infection
38 and certain other post-COVID sequelae, their potential role in Long COVID is important to
39 investigate. Here we apply a well-established, unbiased, proteome-wide autoantibody detection
40 technology (PhIP-Seq) to a robustly phenotyped cohort of 121 individuals with Long COVID, 64
41 individuals with prior COVID-19 who reported full recovery, and 57 pre-COVID controls. While a
42 distinct autoreactive signature was detected which separates individuals with prior COVID
43 infection from those never exposed to COVID, we did not detect patterns of autoreactivity that
44 separate individuals with Long COVID relative to individuals fully recovered from SARS-CoV-2
45 infection. These data suggest that there are robust alterations in autoreactive antibody profiles
46 due to infection; however, no association of autoreactive antibodies and Long COVID was
47 apparent by this assay.

48

49 **BACKGROUND**

50 Some individuals do not fully return to baseline health following SARS-CoV-2 infection and
51 experience ongoing morbidity following the acute phase of COVID-19 (1, 2). There is now intense
52 interest in determining the underlying mechanisms of Long COVID, one type of post-acute
53 sequelae of SARS-CoV-2 infection (PASC) characterized by symptoms that newly developed or
54 worsened following infection that cannot be clearly attributed to another cause (3). Immune
55 dysregulation, including the generation of antibodies against self antigens, has been suggested as
56 one potential driver of Long COVID that warrants further investigation (4, 5).

57
58 Acute SARS-CoV-2 infection is associated with the generation of autoreactive antibodies,
59 particularly among individuals with severe disease requiring hospitalization (6–8). For example,
60 one study of 147 individuals hospitalized with COVID-19 found that autoantibodies associated with
61 connective tissue diseases and anti-cytokine antibodies were identified in 50% of samples, and
62 tracked with the humoral response to SARS-CoV-2 infection (7). Furthermore, multiple studies
63 have described the contribution of likely pre-existing anti-interferon antibodies to vaccine
64 breakthrough infections and severe manifestations of COVID-19, including death (9–12).

65
66 Although much work has been done exploring the potential virologic and immunologic factors
67 driving Long COVID, the evaluation of autoantibodies in this condition has been more limited.
68 Measurement of anti-nuclear antibodies on standard clinical tests has yielded mixed findings, with
69 some studies identifying a high prevalence of antinuclear antibodies (ANAs) among those with
70 Long COVID (13–15) and other studies finding low prevalence consistent with ANA positivity in
71 the general population (16–18). Although perturbations in interferon signaling pathways have been
72 suggested as one potential mechanism of Long COVID (4, 15), anti-interferon antibodies have not
73 been identified in most individuals outside of those who had severe acute infection (16). This is

74 consistent with other post-acute sequelae of COVID such as multisystem inflammatory syndrome
75 in children (MIS-C), which also has no association with anti-interferon antibodies (19).

76

77 The identification of previously reported autoantibodies can be performed using targeted assays.
78 However, identifying the full range of novel autoreactive antibodies, whether they are pathological
79 or not, requires technologies capable of high-throughput, unbiased, proteome-wide screens. In
80 this study, we screened a cohort of individuals with prior SARS-CoV-2 infection, many of whom
81 met clinical criteria for Long COVID, to determine whether a consistent pattern of autoreactivity
82 could be identified. This same technology has been previously utilized to discover novel
83 autoantibodies in a wide range of disease contexts (20–23).

84

85 **RESULTS**

86 **Distinct set of autoreactive antibodies in individuals with prior SARS-CoV-2 infection**

87 We employed a previously published proteome-wide approach using a T7 phage-display assay
88 with immunoprecipitation and next-generation sequencing (PhiP-Seq) (20–24). We tested sera
89 from 185 otherwise healthy individuals with prior SARS-CoV-2 infection in parallel to sera from 57
90 otherwise healthy individuals collected prior to the known existence of COVID-19 (pre-COVID).
91 Using an unbiased analysis, we identified a distinct pattern of autoreactivity which effectively
92 classifies individuals with prior SARS-CoV-2 infection from individuals not yet exposed to the virus
93 with a logistic regression AUC of 0.90 (Figure 1A). The protein targets from which these enriched
94 immunoprecipitated peptides were derived are widely varied and lack any apparent shared
95 biological functions or cell type (Figure 1B). Among the identified targets, a single peptide derived
96 from ARHGAP31 (Figure 2A) displayed the greatest amount of enrichment with 22% of individuals
97 with prior SARS-CoV-2 infection yielding enrichment greater than 6 standard deviations of the

98 mean of the pre-COVID controls. No difference in enrichment was observed for those diagnosed
99 with Long COVID with respect to Post-COVID (Figure 2B).

100

101 To investigate whether the ARHGAP31 peptide enrichment could be the result of cross-reactivity
102 with antibodies directed against SARS-CoV-2 proteins, we performed a multiple sequence
103 alignment of this 49-amino acid fragment against the full SARS-CoV-2 proteome. A region within
104 the SARS-CoV2 Orf1a polyprotein was identified with considerable physico-chemical similarity to
105 a portion of the autoreactive fragment in ARHGAP31(Jalview Version 2;Figure 2C), supporting the
106 notion that the observed human peptidome peptide enrichments in post-COVID samples are being
107 driven by anti-SARS-CoV-2 antibodies.

108

109 ARHGAP31 is most highly expressed in neuronal cells, Langerhans cells, and endothelial cells
110 (Human Protein Atlas, [proteinatlas.org](https://www.proteinatlas.org)). We were unable to identify any clinical differences
111 between individuals with and without ARHGAP31 autoreactivity, such as differences in frequency
112 of neurologic symptoms.

113

114 **Post-COVID autoreactivities are not enriched in individuals with Long COVID**

115 To assess whether autoreactive antibodies present in individuals following SARS-CoV-2 infection
116 are associated with Long COVID, we compared the distribution of the enriched post-COVID
117 peptides amongst the 121 individuals with Long COVID and the 64 individuals with prior SARS-
118 CoV-2 infection but without Long COVID (Convalescent COVID). The 20 most enriched proteins
119 with at least 5-fold greater than background (defined as fold-change over mock-IP with protein
120 A/G beads) were compared (Figure 3). Seventeen of the 20 enriched proteins were present in
121 both Long COVID and Convalescent COVID, and none of these enrichments was observed in any
122 of the pre-COVID controls. Overall, there was no significant differences in enrichment between

123 Long COVID and Convalescent COVID. Peptides derived from three proteins, TMED10, FUCA1,
124 and POL2RK, were only observed in Long COVID, however these were infrequently observed.

125

126 Subgroup analyses were performed to determine whether any of these observed autoreactivities
127 were enriched in particular symptom-defined phenotypes of Long COVID. These include
128 cardiopulmonary (cough, shortness of breath, chest pain, palpitations, and fainting), central
129 neurologic symptoms (problems with vision, headache, difficulty with concentration or memory,
130 dizziness, and difficulty with balance), any neurologic symptom (problems with vision, headache,
131 difficulty with concentration or memory, dizziness, difficulty with balance, trouble with smell or
132 taste, phantosmia, or paresthesia), gastrointestinal (diarrhea, constipation, nausea, vomiting, loss
133 of appetite, and abdominal pain), musculoskeletal (back pain, muscle pain, pain in the arms, legs,
134 or joints), and upper respiratory (runny nose and sore throat).

135

136 The peptide enrichments from individuals with particular Long COVID phenotypes were compared
137 to the enrichments from individuals with Convalescent COVID (no Long COVID symptoms) using
138 a one-sided Kolmogorov-Smirnov test. None of the top 20 autoantibodies were enriched in severe
139 Long COVID, and only 3 autoantibodies were statistically increased in any phenotype: TTF2 and
140 KDM3B in those with cardiopulmonary symptoms and FUCA1 in those with upper respiratory
141 symptoms (Figure 4A). However, using a strict cutoff of 6 standard deviations above the pre-
142 COVID controls to determine positivity, none of these autoantibodies was phenotype specific. By
143 looking at the distribution of these antibodies in Long COVID patients with given sub-phenotype
144 relative to the remaining Long COVID patients and all convalescent COVID patients, it was
145 apparent that the statistical significance in a particular phenotype was driven by either a single
146 individual with extremely high autoantibody signal in the case of KDM3B and TTF2, or higher
147 group signal but not meeting the positive threshold cutoff in the case of FUCA1 (Figure 4B).

148 **Absence of Long COVID specific autoreactivities**

149 In addition to analyzing the distribution of post-COVID autoreactive peptide enrichments, we also
150 performed additional analyses to detect those enrichments which might be present only in Long
151 COVID or particular subcategories of Long COVID. To feature-weight enriched peptides, we
152 applied logistic regressions to the degree of enrichment for each individual with Long COVID and
153 each individual Long COVID symptom phenotype versus enrichment from convalescent COVID
154 patients who did not have the particular symptoms. We included all previously discussed symptom
155 phenotypes, as well as individuals with difficulty concentrating, worsened quality of life, new
156 depression, and generalized anxiety disorder. We were unable to identify a set of enriched
157 proteins specific to any of these groups of patients which could effectively distinguish the cohort
158 from controls, as the best receiver operating characteristic (ROC) area under the curve (AUC) was
159 0.67, and the mean AUC was 0.44 (Supplement 1).

160

161 **DISCUSSION**

162 Autoimmunity has been proposed as one potential mechanism driving Long COVID. We applied
163 an unbiased, proteome-wide, validated approach to assess associations between antibody
164 autoreactivity and clinical phenotype. A clear and robust difference in autoreactivity was detected
165 between those infected with SARS-CoV-2 and pre-COVID controls. This difference was
166 constituted by peptides from diverse and varied proteins, most of which are intracellular,
167 suggesting that the origin of the differential enrichment is due to cross-reactivity with SARS-CoV-2
168 directed antibodies in those that were exposed. A sequence comparison between a peptide from
169 the most enriched protein, ARHGAP31, and Orf1a of SARS-CoV-2 supports this notion, but
170 orthogonal validation through fine-scale epitope mapping and antibody cloning would be required
171 to demonstrate this conclusively. While the clinical significance of incidental autoreactivity due to
172 the humoral immune response to SARS-CoV-2 remains largely unknown, prior studies have

173 identified cross reactive autoantibodies in severe sequelae of SARS-CoV-2 infection, including
174 those that develop severe neurological symptoms (Song et al., 2021). Understanding the clinical
175 consequences of SARS-CoV-2 driven autoreactivity deserves further attention, perhaps through
176 long-term longitudinal studies.

177

178 We found no association to support the hypothesis that autoreactivity, as detected in this assay,
179 contributes to Long COVID. Despite numerous successes, PhIP-seq possesses a number of
180 limitations. Because the T7 phage displayed peptides are only 49 amino acids, the assay
181 inherently detects mostly linear epitopes. Therefore, complex conformational, post translationally
182 modified, or multimeric protein configurations are not predicted to be detectable by this assay, and
183 thus these results do not completely rule out autoimmune interactions in Long COVID that are
184 beyond the scope of PhIP-seq.

185

186 In addition, our classification of Long COVID and its symptom phenotypes was based exclusively
187 on participant self-report of symptoms. It is possible that future analyses in more homogenous
188 cohorts, particularly those with objectively measured physiologic perturbations now associated
189 with certain Long COVID phenotypes (e.g., postural orthostasis/tachycardia, neurocognitive
190 function deficits, abnormalities on cardiopulmonary exercise testing) may yet reveal a role of
191 autoantibodies in at least a subset of individuals experiencing Long COVID.

192

193 Long COVID remains a complex clinical entity. Its causes are likely multifactorial, and there is
194 growing consensus that different phenotypes are driven by different pathophysiologic mechanisms
195 (2, 3). Additional work characterizing SARS-CoV-2 specific and autoreactive immune responses in
196 large, well-characterized cohorts over time, during both the acute and post-acute phases of the
197 illness, will be necessary to delineate the biology of Long COVID and other post-acute sequelae of

198 SARS-CoV-2 infection, and to lead to the development of potential interventions to treat the
199 millions of individuals currently affected by this condition.

200

201 **METHODS**

202 **Study participants and measurements**

203 Participants were volunteers in the UCSF Long-term Impact of Infection with Novel Coronavirus
204 (LIINC) study (NCT04362150). The details of study design and measurement have been reported
205 previously (25). For the current analysis, we included 185 individuals with a history of nucleic acid-
206 confirmed SARS-CoV-2 infection who had a plasma sample collected between 60 and 240 days
207 following initial symptom onset or, if asymptomatic, first positive SARS-CoV-2 test (Supplemental
208 Table 1). Long COVID was defined using study instruments, which have been described in detail
209 elsewhere (25). Briefly, each participant was queried regarding the presence, severity, and
210 duration of 32 physical and mental health symptoms and quality of life. Details of their medical
211 history and COVID-19 infection and vaccination and treatment history were also recorded.
212 Symptoms that pre-dated SARS-CoV-2 infection and that were not changed following infection, as
213 well as those obviously attributed to another cause (e.g. ankle fracture), were not considered to
214 represent Long COVID. All samples included were collected prior to the subject ever receiving a
215 SARS-CoV-2 vaccine.

216

217 Because of challenges in objectively defining Long COVID and to thoroughly explore the data in
218 an unbiased manner, we utilized a number of predefined case definitions in our analysis. We
219 constructed case definitions based on symptom presentation and quality-of-life (QOL) responses.
220 Symptoms case definitions include: the presence of any new or worsening symptoms since
221 SARS-CoV-2 infection (Long COVID); presence of 5 or more symptoms (severe Long COVID);
222 specific symptom groups according to organ system involvement or phenotypic cluster; and

223 individual symptoms when at least 25 individuals experienced the symptom. For individual and
224 grouped symptom outcomes, we developed 3 potential comparisons between the symptomatic
225 group and (1) all individuals who reported absence of the symptoms of interest, regardless of
226 Long COVID status (2) only individuals who were consistently asymptomatic and (3) individuals
227 with Long COVID but not the symptoms of interest. Severe Long COVID was compared only to
228 those without any reported symptoms.

229

230 We defined 6 groups of symptoms (symptom phenotypes) based on either organ system cluster.
231 These include cardiopulmonary (cough, shortness of breath, chest pain, palpitations, and fainting),
232 CNS-specific (problems with vision, headache, difficulty with concentration or memory, dizziness,
233 and difficulty with balance), any neurologic symptom (problems with vision, headache, difficulty
234 with concentration or memory, dizziness, difficulty with balance, trouble with smell or taste,
235 phantosmia, or paresthesia), gastrointestinal (diarrhea, constipation, nausea, vomiting, loss of
236 appetite, and abdominal pain), musculoskeletal (back pain, muscle pain, pain in the arms, legs, or
237 joints), and upper respiratory (runny nose and sore throat).

238

239 Quality-of-life was assessed using the EuroQoL-5D (EQ-5D), Patient Health Questionnaire
240 depression scale (PHQ-8), and Generalized Anxiety Disorder scale (GAD-7). Individuals with
241 Long COVID and the lowest overall quality-of-life (QOL) score measured via the visual analogue
242 scale of the EQ-5D were compared to individuals with the highest overall QOL scores among
243 those with and without Long COVID. Individuals with responses categorized as “moderate
244 depression” on the PHQ-8 (score higher than 10) and “moderate anxiety” on the GAD-7 (score
245 higher than 9) were compared to all participants with scores indicating less severe classifications
246 than “moderate depression” and “moderate anxiety” in the following groups: (1) all participants

247 regardless of Long COVID status (2) all participant without Long COVID and (3) all participants
248 with Long COVID.

249 In addition, we recently demonstrated associations between Long COVID and other chronic latent
250 viral infections, including serologic evidence suggesting recent Epstein-Barr virus (EBV)
251 reactivation (26). For this reason, we also used binary variables to create groups indicating the
252 presence of this condition.

253

254 **Biospecimen collection**

255 At each visit, whole blood was collected in EDTA tubes. Plasma was isolated and stored in -80F
256 until the time of analysis.

257

258 **Phage Immunoprecipitation and Sequencing (PhIP-Seq)**

259 PhIP-Seq was performed following our previously published vacuum-based PhIP-Seq protocol
260 (23) (<https://www.protocols.io/view/scaled-high-throughput-vacuum-hip-protocol-ewov1459kvr2/v1>).

262

263 **PhIP-Seq analysis**

264 All analysis (except when specifically stated otherwise) was performed at the gene-level, in which
265 all reads for all peptides mapping to the same gene were summed, and 0.5 reads were added to
266 each gene to allow inclusion of genes with zero reads in mathematical analyses. Within each
267 individual sample, reads were normalized by converting to the percentage of total reads. To
268 normalize each sample against background non-specific binding, a fold-change(FC) over mock-IP
269 was calculated by dividing the sample read percentage for each gene by the mean read-
270 percentage of the same gene for the AG bead only controls. This FC signal was then used for
271 side-by-side comparison between samples and cohorts. Samples which had a FC of 5 or greater

272 were considered enriched for an antibody, and samples with a FC of 6 standard deviations above
273 the mean of pre-COVID controls were considered positive for an autoantibody. FC values were
274 also used to calculate z-scores for each disease category sample by using each respective control
275 (as specified in figures and results), and for each control sample by using all remaining controls.
276 These z-scores were used for the logistic-regression feature weighting. In the case peptide-level
277 analysis, raw reads were normalized by calculating the number of reads per 100,000 reads.

278

279 **Statistical methods**

280 All statistical analysis was performed in Python using the Scipy Stats package. A two-way
281 Kolmogorov-Smirnoff(KS) test was used for comparisons of FC PhIP-Seq data between groups of
282 samples, except in the case of specifically looking for those genes with increased signal only in
283 the disease-cohort in which a one-way KS test was employed. The logistic regression machine-
284 learning classifiers were performed using our recently described methods (23). Utilizing the Scikit-
285 learn package, logistic regression classifiers were applied to z-scored PhIP-Seq values from
286 individuals with a designated disease category versus the designated control. A liblinear solver
287 was used with L1 regularization, and the model was evaluated using a five-fold cross-validation (4
288 of the 5 for training, 1 of the 5 for testing).

289

290 **Informed consent**

291 Participants provided written informed consent. The study was approved by the UCSF Institutional
292 Review Board.

293

294 **FOOTNOTES**

295 **Acknowledgements**

296 We are grateful to the study participants and their medical providers. We acknowledge current and
297 former LIINC clinical study team members Tamara Abualhsan, Andrea Alvarez, Khamal Anglin,
298 Urania Argueta, Mireya Arreguin, Kofi Asare, Melissa Buitrago, Monika Deswal, Nicole DelCastillo,
299 Emily Fehrman, Halle Grebe, Heather Hartig, Yanel Hernandez, Rebecca Hoh, Beatrice Huang,
300 Marian Kerbleski, Raushun Kirtikar, James Lombardo, Monica Lopez, Michael Luna, Sadie
301 Munter, Lynn Ngo, Enrique Martinez Ortiz, Antonio Rodriguez, Justin Romero, Dylan Ryder, Ruth
302 Diaz Sanchez, Matthew So, Celina Chang Song, Viva Tai, Alex Tang, Cassandra Thanh, Fatima
303 Ticas, Leonel Torres, Brandon Tran, Daisy Valdivieso, Deepshika Varma, and Meghann Williams;
304 and LIINC laboratory team members Amanda Buck, Tyler-Marie Deveau, Joanna Donatelli, Jill
305 Hakim, Nikita Iyer, Owen Janson, Brian LaFranchi, Christopher Nixon, Isaac Thomas, and
306 Keirstinne Turcios. We thank Jessica Chen, Aidan Donovan, Carrie Forman, and Rania Ibrahim
307 for assistance with data entry and review. We thank the UCSF AIDS Specimen Bank for
308 processing specimens and maintaining the LIINC biospecimen repository. We are grateful to
309 Elnaz Eilkhani and Monika Deswal for regulatory support.

310

311 **Author Contributions**

312 AB, ST, BG, JDT, JNM, SGD, TJH, JLD, and MJP designed the study. JDT, JNM, SGD, TJH, and
313 MJP designed and oversaw the LIINC cohort, with major contributions from BG, MSD, PYH, and
314 JH. KA, BH, and RH managed cohort operations. Clinical data were collected by JR, BT, RK, HG,
315 and MS. SL and SG managed the data and developed the clinical dataset. ST, SL, SG, and MJP
316 selected participants for inclusion in this study. Biospecimens were processed in the laboratories
317 of BG and TJH. CW, AS, and AM performed the measurements in the laboratories of MSA and
318 JLD. AB performed the data analysis. AB, TJH, JLD, and MJP wrote the first draft of the paper,
319 with input from all other authors. All authors edited the manuscript and approved the final version.

320

321 **Funding**

322 This work was supported by NIH/NIAID 3R01AI141003-03S1 and NIH/NIAID R01AI158013 (to M
323 Gandhi and M Spinelli), Pediatric Scientist Development Program and the Eunice Kennedy
324 Shriver National Institute of Child Health and Human Development (K12-HD000850 to A.
325 Bodansky), and Chan Zuckerberg Biohub support for J. DeRisi.

326

327 **Conflicts of Interest**

328 TJH reports grants from Merck & Co, Gilead Sciences, and Bristol-Myers Squibb, and has
329 provided consulting for Roche, outside the submitted work. SGD reports grants and/or personal
330 fees from Gilead Sciences, Merck & Co., Viiv, AbbVie, Eli Lilly, ByroLogyx, and Enochian
331 Biosciences, outside the submitted work. JD reports paid compensation for consulting for the
332 Public Health Company and Allen & Co. MJP reports consulting fees from Gilead Sciences and
333 AstraZeneca, outside the submitted work. All other authors report no potential conflicts.

334

335

336 **FIGURE LEGENDS**

337 **Figure 1: PhIP-Seq autoreactivities distinguish post-COVID sera from pre-COVID controls.**

338 Logistic regression comparing PhIP-Seq autoreactivities in all individuals with prior COVID
339 infection compared to pre-COVID controls. Barplot showing autoreactivities with the top 20 logistic
340 regression coefficients.

341

342 **Figure 2: Post-COVID anti-ARHGAP31 autoreactivities target specific region with similarity**
343 **to SARS-CoV-2**

344 (A) Stripplots showing distribution of ARHGAP31 autoreactivities in Long COVID, individuals with
345 prior COVID infection but without Long COVID, and pre-COVID controls. Dotted line at 6 standard
346 deviations above mean of pre-COVID controls. (B) Distribution of anti-ARHGAP31 autoreactivity
347 signal within ARHGAP31 full length protein. One specific fragment is targeted. (C) Amino-acid
348 sequence of the autoreactive region of ARHGAP31 and amino-acid sequence of a region of
349 SARS-CoV-2 Orf1a with similarity. Shown below is the multiple sequence
350 alignment(ClustalOmega; asterisk=identical amino acid; colon=strongly similar properties with
351 Gonnet PAM 250 matrix score >0.5; period=weakly similar with Gonnet PAM 250 matrix score
352 between 0 and 0,5) and strong physical-chemical conservation (JalView; amino acid physical-
353 chemical conservation scored on a scale of 1-11, Asterisk=score of 11 and identical amino acid,
354 Plus=10, all properties conserved).

355

356 **Figure 3: Post-COVID autoreactivities are similarly distributed among Long COVID and**
357 **controls.**

358 Hierarchically clustered (Pearson) heatmaps showing the PhIP-Seq enrichment for the top 20
359 autoreactivities ranked by logistic regression coefficient in each Long COVID patient, each COVID
360 convalescent patient, and each pre-COVID control.

361 **Figure 4: Few significantly increased autoreactivities in Long COVID symptom phenotypes.**

362 (A) Heatmap with p-values (Kolmogorov-Smirnov testing) of differences in autoantigen enrichment
363 for all individuals with prior COVID infection with and without additional clinical factors. Top-row
364 compares those with and without Long COVID. Lower rows show subcategories of Long COVID.
365 (B) Stripplots showing the three autoantibodies with statistically significant enrichment in a post-
366 COVID clinical phenotype. Dotted-lines show 6 standard-deviations above the mean of pre-
367 COVID signal.

368

369 **Supplemental Table 1: Characteristics of study participants.**

370 Characteristics of all study participants, stratified by Long COVID status.

371

372 **Supplemental Figure 1: PhIP-Seq is unable to distinguish Long COVID symptom**
373 **phenotypes from controls.**

374 (A) Logistic regression receiver operating characteristic (ROC) curves for Long COVID patients
375 with different specified symptom phenotypes relative to patients previously infected with COVID
376 without the phenotype, or (B) with and without EBV laboratory findings.

377

378

379 **REFERENCES**

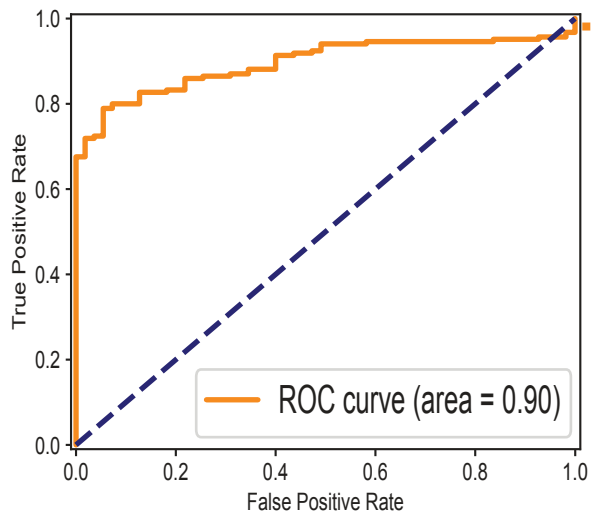
- 380 1. Nalbandian A et al. Post-acute COVID-19 syndrome [Internet]. *Nat. Med.* [published online
381 ahead of print: March 22, 2021]; doi:10.1038/s41591-021-01283-z
- 382 2. Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: major findings, mechanisms and
383 recommendations. *Nat. Rev. Microbiol.* 2023;1–14.
- 384 3. Peluso MJ, Deeks SG. Early clues regarding the pathogenesis of long-COVID. *Trends*
385 *Immunol.* 2022;43(4):268–270.
- 386 4. Proal AD, VanElzakker MB. Long COVID or Post-acute Sequelae of COVID-19 (PASC): An
387 Overview of Biological Factors That May Contribute to Persistent Symptoms [Internet]. *Frontiers in*
388 *Microbiology* 2021;12. doi:10.3389/fmicb.2021.698169
- 389 5. Peluso MJ, Donatelli J, Henrich TJ. Long-term immunologic effects of SARS-CoV-2 infection:
390 leveraging translational research methodology to address emerging questions. *Transl. Res.*
391 2022;241:1–12.
- 392 6. Zuo Y et al. Prothrombotic autoantibodies in serum from patients hospitalized with COVID-19
393 [Internet]. *Sci. Transl. Med.* 2020;12(570). doi:10.1126/scitranslmed.abd3876
- 394 7. Chang SE et al. New-onset IgG autoantibodies in hospitalized patients with COVID-19. *Nat.*
395 *Commun.* 2021;12(1):5417.
- 396 8. Woodruff MC et al. Dysregulated naïve B cells and de novo autoreactivity in severe COVID-19
397 [Internet]. *Nature* [published online ahead of print: August 31, 2022]; doi:10.1038/s41586-022-
398 05273-0
- 399 9. Bastard P et al. Autoantibodies against type I IFNs in patients with life-threatening COVID-19

- 400 [Internet]. *Science* 2020;370(6515). doi:10.1126/science.abd4585
- 401 10. Bastard P et al. Autoantibodies neutralizing type I IFNs are present in ~4% of uninfected
402 individuals over 70 years old and account for ~20% of COVID-19 deaths [Internet]. *Sci Immunol*
403 2021;6(62). doi:10.1126/sciimmunol.abl4340
- 404 11. Vazquez SE et al. Neutralizing Autoantibodies to Type I Interferons in COVID-19
405 Convalescent Donor Plasma. *J. Clin. Immunol.* 2021;41(6):1169–1171.
- 406 12. van der Wijst MGP et al. Type I interferon autoantibodies are associated with systemic
407 immune alterations in patients with COVID-19. *Sci. Transl. Med.* 2021;13(612):eabh2624.
- 408 13. Son K et al. Circulating anti-nuclear autoantibodies in COVID-19 survivors predict long-COVID
409 symptoms [Internet]. *Eur. Respir. J.* [published online ahead of print: September 22, 2022];
410 doi:10.1183/13993003.00970-2022
- 411 14. Seeßle J et al. Persistent symptoms in adult patients one year after COVID-19: a prospective
412 cohort study [Internet]. *Clin. Infect. Dis.* [published online ahead of print: July 5, 2021];
413 doi:10.1093/cid/ciab611
- 414 15. Su Y et al. Multiple Early Factors Anticipate Post-Acute COVID-19 Sequelae [Internet]. *Cell*
415 2022;0(0). doi:10.1016/j.cell.2022.01.014
- 416 16. Peluso MJ, Thomas IJ, Munter SE, Deeks SG, Henrich TJ. Lack of Antinuclear Antibodies in
417 Convalescent Coronavirus Disease 2019 Patients With Persistent Symptoms. *Clin. Infect. Dis.*
418 2022;74(11):2083–2084.
- 419 17. Schultheiß C et al. The IL-1 β , IL-6, and TNF cytokine triad is associated with post-acute
420 sequelae of COVID-19. *Cell Rep Med* 2022;3(6):100663.

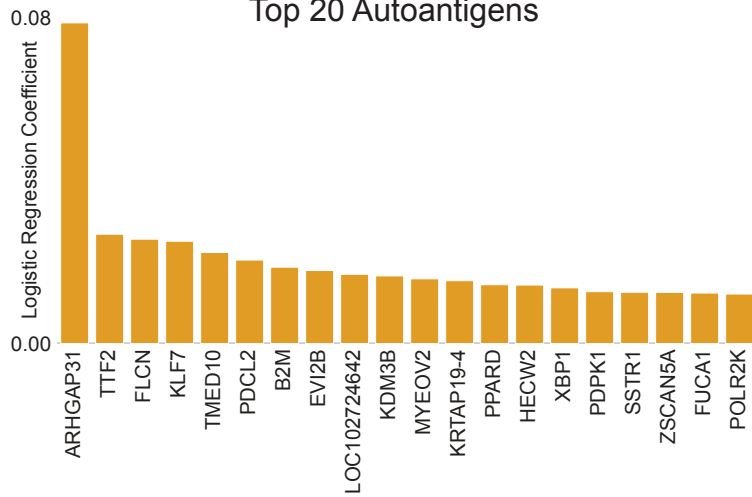
- 421 18. Klein J et al. Distinguishing features of Long COVID identified through immune profiling
422 [Internet]. *medRxiv* [published online ahead of print: August 10, 2022];
423 doi:10.1101/2022.08.09.22278592
- 424 19. Bodansky A et al. NFKB2 haploinsufficiency identified via screening for IFN α 2 autoantibodies
425 in children and adolescents hospitalized with SARS-CoV-2-related complications [Internet]. *J.*
426 *Allergy Clin. Immunol.* [published online ahead of print: December 9, 2022];
427 doi:10.1016/j.jaci.2022.11.020
- 428 20. Mandel-Brehm C et al. Kelch-like Protein 11 Antibodies in Seminoma-Associated
429 Paraneoplastic Encephalitis. *N. Engl. J. Med.* 2019;381(1):47–54.
- 430 21. O'Donovan B et al. High-resolution epitope mapping of anti-Hu and anti-Yo autoimmunity by
431 programmable phage display. *Brain Commun* 2020;2(2):fcaa059.
- 432 22. Vazquez SE et al. Identification of novel, clinically correlated autoantigens in the monogenic
433 autoimmune syndrome APS1 by proteome-wide PhIP-Seq [Internet]. *Elife* 2020;9.
434 doi:10.7554/eLife.55053
- 435 23. Vazquez SE et al. Autoantibody discovery across monogenic, acquired, and COVID-19-
436 associated autoimmunity with scalable PhIP-seq [Internet]. *Elife* 2022;11. doi:10.7554/eLife.78550
- 437 24. Larman HB et al. Autoantigen discovery with a synthetic human peptidome. *Nat. Biotechnol.*
438 2011;29(6):535–541.
- 439 25. Peluso MJ et al. Persistence, magnitude, and patterns of postacute symptoms and quality of
440 life following onset of SARS-CoV-2 infection: cohort description and approaches for measurement.
441 In: *Open forum infectious diseases*. Oxford University Press US; 2022:ofab640
- 442 26. Peluso MJ et al. Impact of pre-existing chronic viral infection and reactivation on the

443 development of long COVID [Internet]. *J. Clin. Invest.* [published online ahead of print: December
444 1, 2022]; doi:10.1172/JCI163669

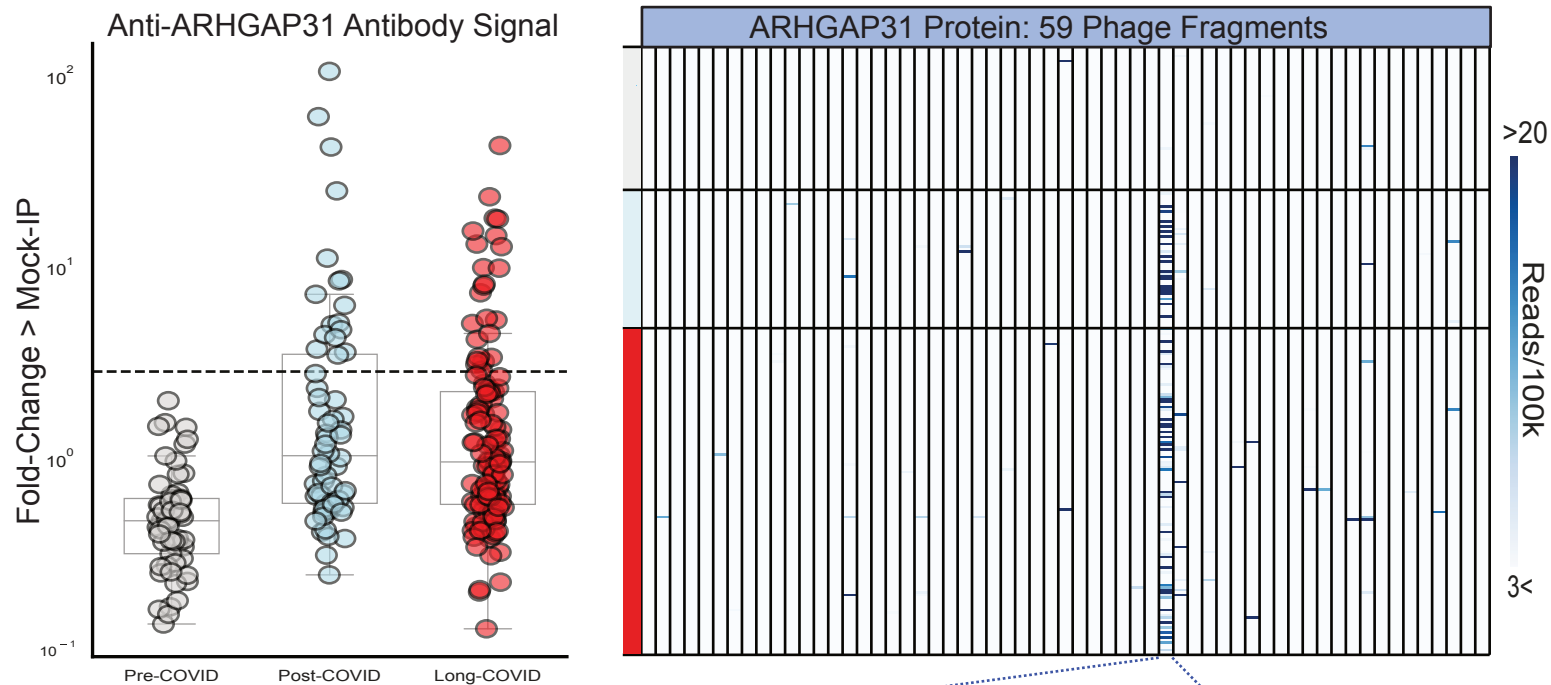
Post-COVID vs Pre-COVID



Top 20 Autoantigens



A



C

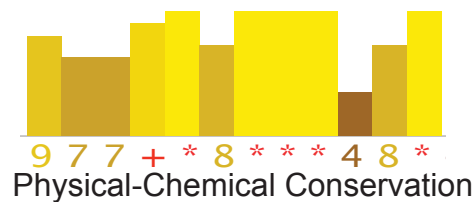
ARHGAP31 Targeted Sequence

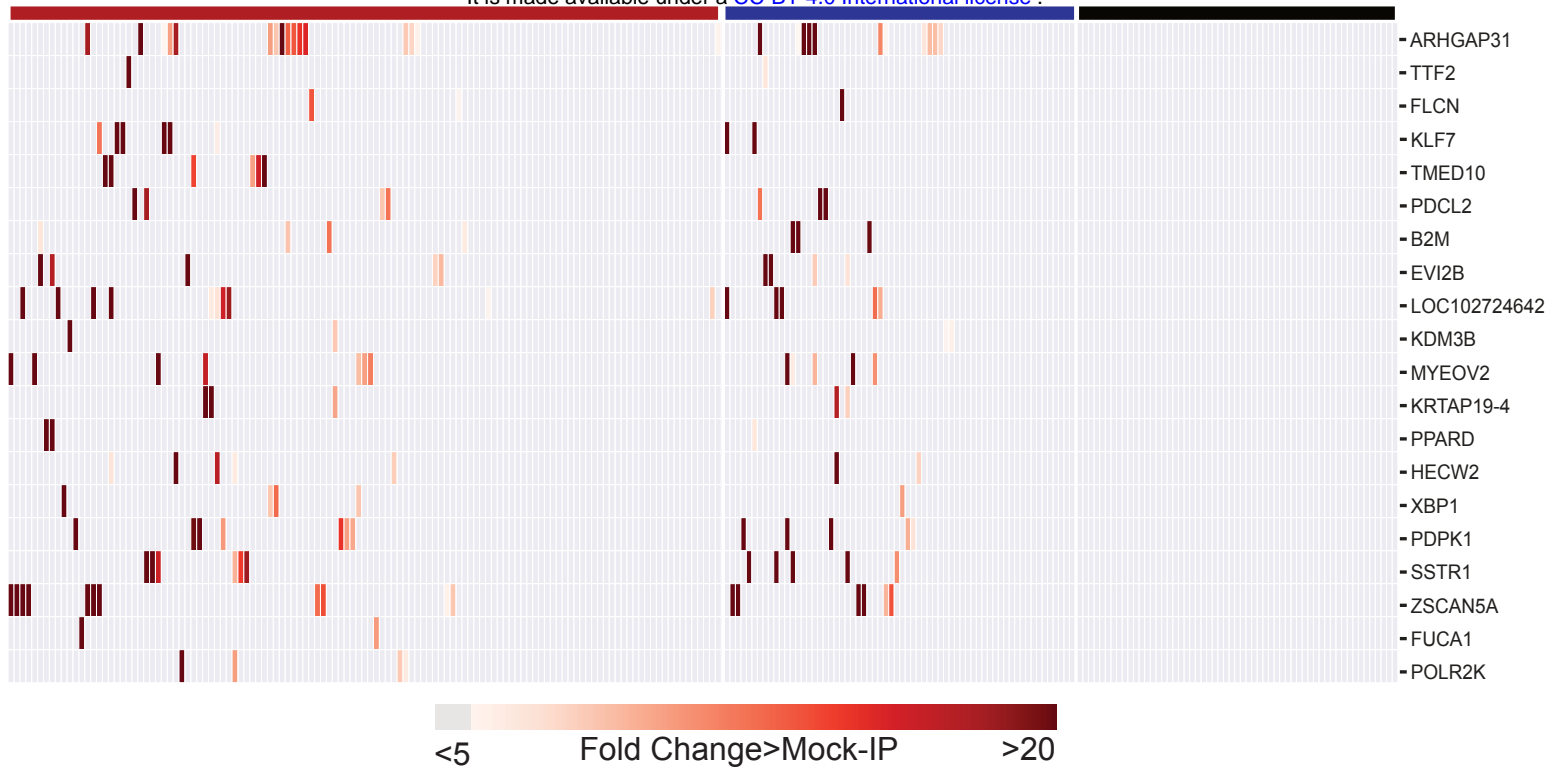
885- PSDCDEDDT**VTDIAQHGLE**MPWEEPQWVTSPLHSPTLKDAHKAQVQG -933

SARS-CoV-2 Orf1a Similar Sequence

2119- **LKTLATHGLAAV**NSVPWDTIANYAKPFLNKVVSTTTNIVTRC -2161

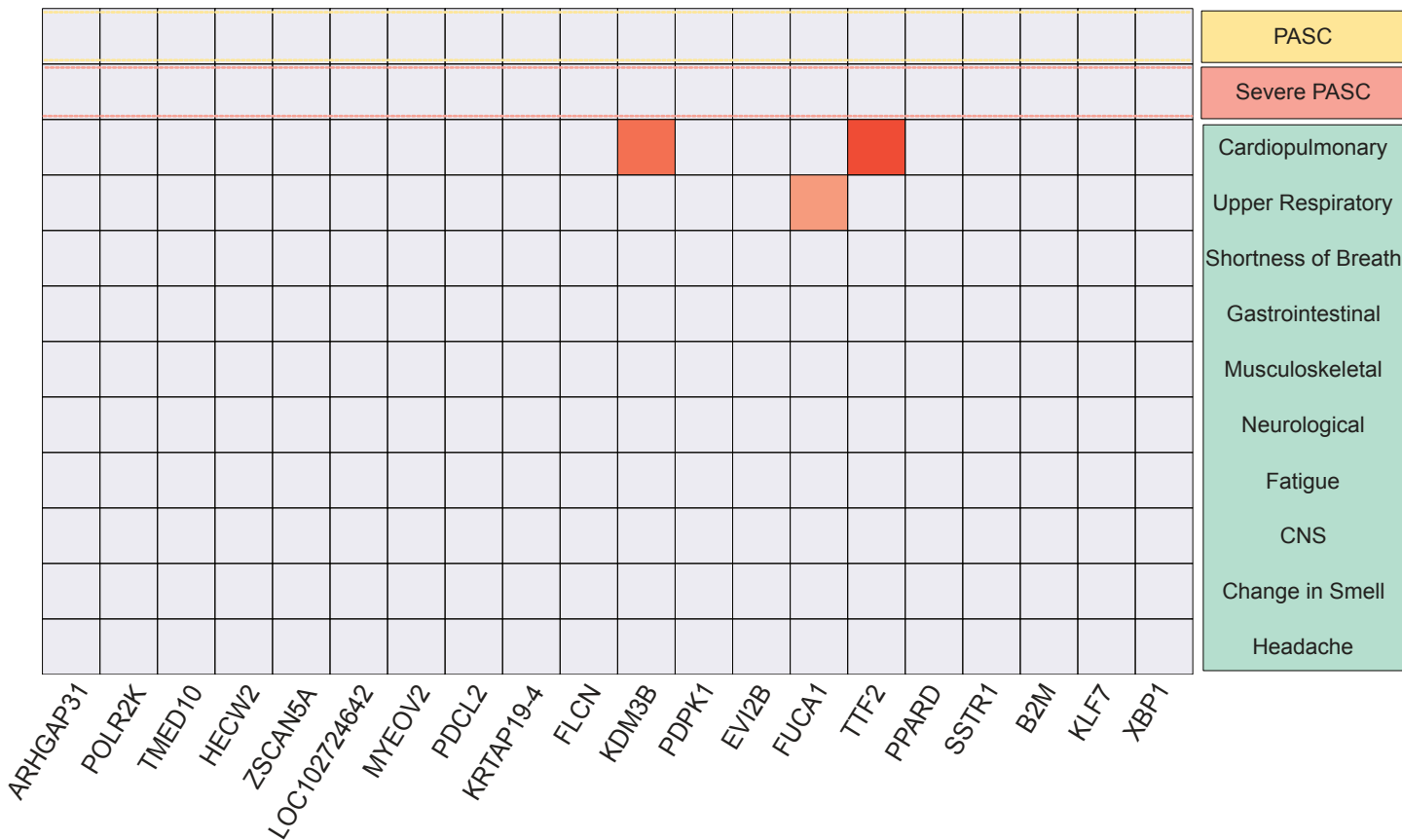
VTDIAQHGLEMP
LKTLATHGLAAV
: . : * * * *





A

>0.05 | **p-value** | 0.00

B

