

Identification of a multi-omics factor predictive of long COVID in the IMPACC study

Gisela Gabernet^{1,21}, Jessica Maciuch^{2,21}, Jeremy P. Gygi^{1,21}, John F. Moore^{3,21}, Annmarie Hoch⁴, Caitlin Syphurs⁴, Tianyi Chu¹, Naresh Doni Jayavelu⁵, David B. Corry⁶, Farrah Kheradmand⁶, Lindsey R. Baden⁷, Rafick-Pierre Sekaly⁸, Grace A. McComsey⁸, Elias K. Haddad⁹, Charles B. Cairns⁹, Nadine Rouphael¹⁰, Ana Fernandez-Sesma¹¹, Viviana Simon¹¹, Jordan P. Metcalf¹², Nelson I Agudelo Higueta¹², Catherine L. Hough¹³, William B. Messer¹³, Mark M. Davis¹⁴, Kari C. Nadeau¹⁴, Bali Pulendran¹⁴, Monica Kraft¹⁵, Chris Bime¹⁵, Elaine F. Reed¹⁶, Joanna Schaeenman¹⁶, David J. Erle¹⁷, Carolyn S. Calfee¹⁷, Mark A. Atkinson¹⁸, Scott C. Brackenridge¹⁸, Esther Melamed³, Albert C. Shaw¹, David A. Hafler¹, Al Ozonoff⁴, Steven E. Bosinger¹⁰, Walter Eckalbar¹⁷, Holden T. Maecker¹⁴, Seunghee Kim-Schulze¹¹, Hanno Steen¹⁹, Florian Krammer¹¹, Kerstin Westendorf²⁰, IMPACC Network[#], Bjoern Peters²⁰, Slim Fourati², Matthew C. Altman⁵, Ofer Levy¹⁹, Kinga K. Smolen¹⁹, Ruth R. Montgomery¹, Joann Diray-Arce⁴, Steven H. Kleinstein^{1,22}, Leying Guan^{1,22}, Lauren I. R. Ehrlich^{3,22,23*}

1. Yale School of Medicine, New Haven, CT 06511, USA
2. Northwestern Feinberg School of Medicine, Chicago, IL 60611, USA
3. The University of Texas at Austin, Austin, TX 78712, USA
4. Clinical and Data Coordinating Center (CDCC), Precision Vaccines Program, Boston Children's Hospital, Harvard Medical School, Boston, MA 02115, USA
5. Benaroya Research Institute, University of Washington, Seattle, WA 98101, USA
6. Baylor College of Medicine and the Center for Translational Research on Inflammatory Diseases, Houston, TX 77030, USA.
7. Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115
8. Case Western Reserve University and University Hospitals of Cleveland, Cleveland, OH 44106, USA
9. Drexel University, Tower Health Hospital, Philadelphia, PA 19104, USA
10. Emory School of Medicine, Atlanta, GA 30322, USA
11. Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
12. Oklahoma University Health Sciences Center, Oklahoma City, OK 73104
13. Oregon Health & Science University, Portland, OR 97239, USA
14. Stanford University School of Medicine, Palo Alto, CA 94305, USA
15. University of Arizona, Tucson AZ 85721, USA
16. David Geffen School of Medicine at the University of California Los Angeles, Los Angeles CA 90095, USA
17. University of California San Francisco, San Francisco, CA 94115, USA
18. University of Florida, Gainesville, FL 32611, USA
19. Precision Vaccines Program, Boston Children's Hospital, Harvard Medical School, Boston, MA 02115, USA
20. La Jolla Institute for Immunology, La Jolla, CA 92037, USA
21. These authors contributed equally
22. These authors contributed equally

23. Lead Contact

*Correspondence: lehrlich@austin.utexas.edu

IMPACC Network group members (listed in Pubmed) are listed at the end of the document

SUMMARY

Following SARS-CoV-2 infection, ~10-35% of COVID-19 patients experience long COVID (LC), in which often debilitating symptoms persist for at least three months. Elucidating the biologic underpinnings of LC could identify therapeutic opportunities. We utilized machine learning methods on biologic analytes and patient reported outcome surveys provided over 12 months after hospital discharge from >500 hospitalized COVID-19 patients in the IMPACC cohort to identify a multi-omics “recovery factor”. IMPACC participants who experienced LC had lower recovery factor scores compared to participants without LC. Biologic characterization revealed increased levels of plasma proteins associated with inflammation, elevated transcriptional signatures of heme metabolism, and decreased androgenic steroids in LC patients. The recovery factor was also associated with altered circulating immune cell frequencies. Notably, recovery factor scores were predictive of LC occurrence in patients as early as hospital admission, irrespective of acute disease severity. Thus, the recovery factor identifies patients at risk of LC early after SARS-CoV-2 infection and reveals LC biomarkers and potential treatment targets.

KEYWORDS

COVID-19, Long COVID, PASC, patient reported outcomes, SARS-CoV-2, Machine Learning, Heme metabolism, androgenic steroids, inflammation, multi-omics

INTRODUCTION

Long COVID (LC) has become a pressing public health concern, affecting an estimated 10-35% of surviving individuals infected with SARS-CoV-2, or 15-20 million individuals in the United States and over 60 million worldwide^{1,2}. In July 2024, the National Academies of Sciences, Engineering, and Medicine (NASEM) released an updated definition of LC, characterizing it as a chronic condition arising after SARS-CoV-2 infection that persists for at least 3 months, irrespective of acute disease severity². LC can encompass a wide range of physical and cognitive symptoms, and can lead to new or worsening neurological, psychiatric, cardiovascular, pulmonary, endocrine, and gastrointestinal conditions, among others¹⁻¹².

Several studies have identified demographic and clinical risk factors for LC¹³⁻¹⁵, including age¹⁶⁻¹⁸, female sex¹⁶⁻²¹, and longer hospital stays^{18,22}. Higher viral loads^{18,19} and lower anti-SARS-CoV-2 antibody titers^{18,23,24} during the acute infection phase have also been associated with LC development. There are multiple, potentially overlapping hypotheses that explain the etiology of this condition, including persistent viral infection^{25,26}, chronic inflammation²⁶⁻²⁹, latent herpesvirus reactivation^{26,30,31}, immune dysregulation^{29,32,33}, complement dysregulation³⁴, and autoimmunity^{35,36}. Despite these efforts, no consensus exists on the mechanisms of LC pathogenesis, and validation of the molecular findings across cohorts has been challenging. Additionally, most existing studies rely on measurements from a single or limited number of assays and are confined to restricted sampling time points during the acute or convalescent

phases of the disease. Therefore, a multi-omics longitudinal study spanning both acute infection and convalescent disease phases could help elucidate the molecular mechanisms underlying LC.

The Immunophenotyping Assessment in a COVID-19 Cohort (IMPACC) study³⁷ offers a unique opportunity to investigate the temporal dynamics of multi-omics immune profiles during the acute and convalescent COVID-19 infection phases in a clinically well-characterized cohort of hospitalized patients from across the United States. Data from the IMPACC study can be leveraged to identify molecular correlates of post-acute symptom development or resolution for one year after hospital discharge. This cohort has previously been studied to characterize multi-omics determinants of COVID-19 severity and mortality during the acute phase of disease³⁸⁻⁴⁰. Another study of the IMPACC cohort identified LC participants who experienced patient-reported outcome deficits up to 12 months after COVID-19 hospital discharge¹⁸. Clinical characteristics such as female sex, a higher respiratory SARS-CoV-2 viral burden, and lower antibody titers against the SARS-CoV-2 Spike protein during acute disease were associated with persistent deficits after hospital discharge. B cell lymphopenia and elevated fibroblast growth factor 21 (FGF21) during the acute disease phase were also characteristics of participants who developed LC¹⁸. However, longitudinal immune profiles of IMPACC participants experiencing LC during the convalescent phase of disease have not yet been compared to those of IMPACC participants who experienced minimal deficits during convalescence, an analysis that could reveal LC biomarkers and uncover biological processes underlying the disease.

In the current work, we applied supervised multi-omics integration methods to develop interpretable models that differentiate participants with LC from recovered individuals based on their longitudinal immunophenotyping profiles during the convalescent disease phase. We identified key biological programs and biomarkers driving LC classification. Our findings highlight persistent inflammation, dysregulation of heme metabolism, and altered androgenic steroid profiles as characterizing features of LC, independent of acute disease severity or SARS-CoV-2 vaccination status post hospital discharge. Notably, these molecular profiles were already detectable during the acute phase of disease, suggesting their potential value as early predictive biomarkers for identifying patients at risk of developing LC. Additionally, despite a general lack of consensus about the definition of LC or consistency in the timing of sampling across different studies, we validated dysregulation of the heme metabolism signature in an independent LC cohort. These findings provide valuable insights into the molecular underpinnings of LC and offer a foundation for future research aimed at improving diagnostics and developing targeted interventions.

RESULTS

Longitudinal multi-omics profiling of long COVID

The IMPACC study included 1,164 participants admitted to 20 US hospitals for COVID-19 infection between May 2020 and March 2021³⁷. Clinical data collection and immunophenotyping were performed longitudinally during the acute disease phase within 72h of hospital admission and 4, 7, 14, 21 and 28 days after hospital admission (Visits 1-6, respectively)². Surviving

participants were contacted 3, 6, 9, and 12 months after hospital discharge (Visits 7-10, respectively) to complete patient-reported outcome and symptom surveys during the convalescent phase, and to provide biosamples for immunophenotyping assays. Of the 702 participants who could be reached by the study team after discharge, 513 were included in the IMPACC Convalescent cohort. These participants were selected because they completed at least one survey and provided at least one biosample during the convalescent period (Figure 1A and Table S1)¹⁸. IMPACC core labs performed immunophenotyping both in the acute and convalescent phases, including measurements of inflammatory mediators in blood serum via Olink (SO), global blood plasma metabolomics (PMG), global and targeted blood plasma proteomics (PPG and PPT), peripheral blood mononuclear cell (PBMC) transcriptomics (PGX), whole blood cell frequencies measured by mass cytometry by time of flight (CyTOF), and CyTOF mean marker signal intensity measurements (BCT).

LC status was defined in this cohort according to the participant's response to post discharge surveys that captured symptoms and Patient-Reported Outcome (PRO) measures that evaluated general health and deficits in specific domains. Participants who responded to at least one set of post discharge surveys were assigned to PRO clusters according to latent class modeling and clustering using standardized scores of the PRO measures, as previously reported¹⁸ (details on the PRO measures used can be found in the Supplementary Methods). PRO clusters were classified as participant clusters with no or minimal deficits (MIN), or with deficits attributed to LC in several domains: physical predominant (PHY), mental/cognitive predominant (COG), and multi/pan domain (MLT)¹⁸ (Figure 1B).

In this study, we utilized multi-omics immunophenotyping profiles to develop interpretable models for predicting LC and exploring the underlying molecular mechanisms. To assess the model performance, we split the Convalescent cohort into an 80% Train and 20% Test cohort, maintaining the proportions of participants in each PRO cluster (Figure 1C), with no noticeable imbalance in other clinical characteristics or biosample availability between the cohorts (Figure S1). We then used Signature-based multiPle-omics intEgration via lAtent factoRs (SPEAR)⁴¹, a supervised Bayesian factor model for the identification of multi-omics features, to integrate the high dimensional data and construct multi-omics predictive factors from immune profiles obtained during the convalescent phase in the Train cohort. We assessed their predictive performance by repeated cross-validation on the Train cohort and validated the selected model performance on the Test cohort (Figure 1D and 1E). To identify the immune programs captured in the predictive factors, we conducted in depth analyses of enriched biological pathways and analytes identified as highly relevant for the prediction performance by the model and performed associations with assay data not included in model training, such as blood CyTOF cell frequencies.

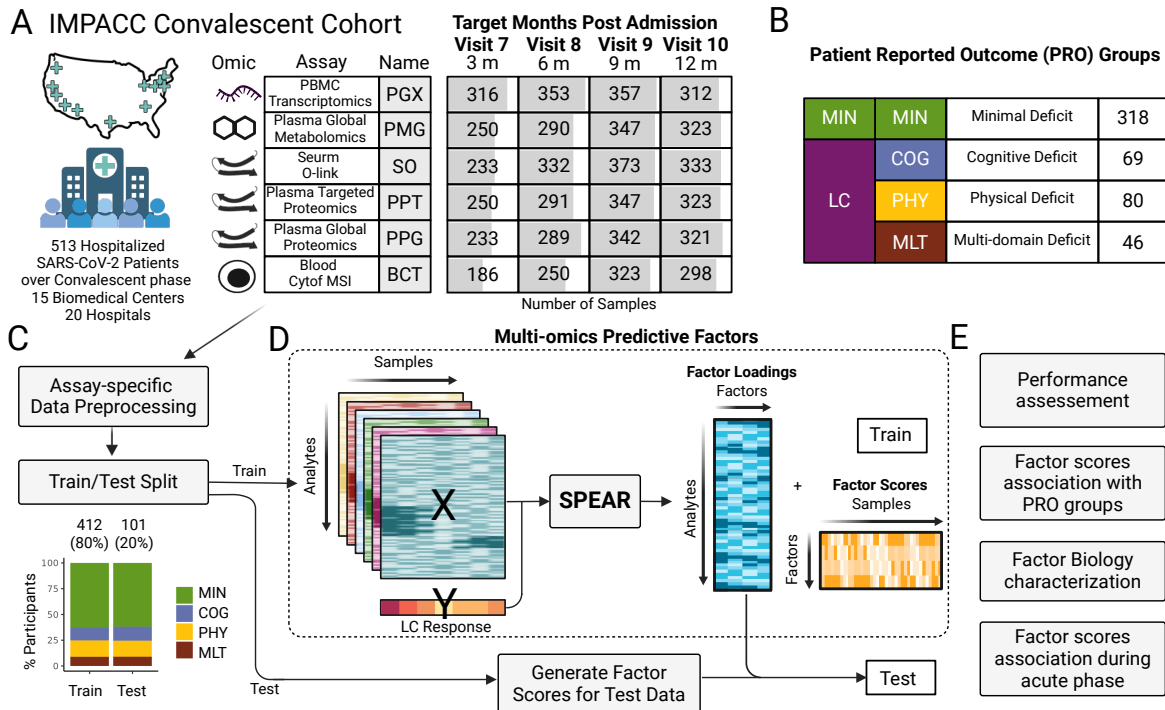


Figure 1. Multi-omics data overview and generation of a predictive LC factor. (A) Number of samples used in the multi-omics data integration strategy by assay (rows) and scheduled time of collection (columns). Shading indicates the frequency of samples with data availability at the indicated visit. (B) Patient classification in Patient Reported Outcome (PRO) clusters according to the PRO measure survey scores¹⁸. (C) Individual assay data were preprocessed and split into Train and Test cohorts by participant in an 80/20 split, maintaining the proportion of PRO cluster participants in each partition. (D) Preprocessed assay data and LC response outcomes for the Train cohort were used to identify multi-omics predictive factors with SPEAR. Factor scores were then calculated for the Test cohort. (E) The performance of the multi-omics predictive factors to classify patients into presence and absence of LC was assessed on the Train cohort via cross-validation and then validated on the Test cohort. The predictive factor scores were confirmed to be associated with LC after correcting for possible confounding variables. In depth analysis of enriched biological pathways and significant analytes relevant for the prediction was performed. Factor scores were computed for the acute infection immune profiles, and association analysis with LC at these early time points was performed. See also Figure S1.

Multi-omics factors are predictive of long COVID

We focused on predicting LC in the Convalescent cohort from the multi-omics immune profiling data collected during the convalescent phase. Since the binary LC labels (presence or absence of LC) per participant could omit valuable information captured by the numeric PRO measures at each participant visit, we constructed separate SPEAR models to generate supervised factors including PRO measure scores (SPEAR Physical, SPEAR Cognitive, SPEAR Mental, SPEAR Impact, SPEAR Dyspnea) or the LC binary labels (SPEAR LC) as response variables (Figure S2A). We evaluated the predictive performance of the SPEAR multi-omics factors trained on the individual PRO measure scores to reconstruct the score values of unseen data in the training

cohort by cross-validation using a lasso regression model (Figure S2B). The SPEAR Physical model was the best performing, with lowest prediction error, and was selected for further evaluation. We then utilized the same procedure to assess the ability of the SPEAR LC factors and the best performing SPEAR model trained on individual PRO measure scores (SPEAR Physical) to predict binary LC status. The lasso model trained on the SPEAR Physical multi-omics factor achieved the highest predictive performance as evaluated with the area under the receiver-operating characteristic curve (AUROC) (Figure 2A). Additionally, we compared the predictive performance of the SPEAR multi-omics supervised factors to the performance of equivalent unsupervised multi-omics factors, obtained with the Multi-Omics Factor Analysis (MOFA)⁴² framework, that do not consider a response variable during the factor construction step. Both lasso classifiers trained on the SPEAR Physical and SPEAR LC predictive factors outperformed the classifier using MOFA unsupervised factors on the Train cohort (Figure 2A, Figure S2B). The SPEAR Physical model achieved an AUROC of 0.69 for predicting LC presence or absence in the Test cohort (Figure 2B). The SPEAR Physical Factor, learned by the SPEAR physical model, was significantly associated with LC in the Test cohort after correcting for sex and age ($p=0.00098$), two variables which have been previously associated with LC in our cohort¹⁸. The SPEAR Physical Factor scores were significantly higher for participants in the MIN group compared to the LC group, so we termed this factor as the “recovery factor” (Figure 2C, Figure S4A). Recovery factor scores were significantly associated with PRO clusters ($p=0.0009$); however, they showed a differential ability to identify individual LC deficit domains, with significant differences between MIN vs COG and MIN vs MLT PRO clusters, but not MIN vs PHY PRO clusters (Figure 2D, Figure S4B). Sparse lasso regression models to reconstruct recovery factor scores utilizing all analytes included in the model or analytes from individual omics assays showed that the model including all assays was best at reconstructing the factor scores, indicating that the recovery factor scores captured contributions from multiple omics layers (Figure S2C). Taken together, the recovery factor is a multi-omics model comprised of biologic analyte levels during the convalescent phase of COVID-19 that is able to distinguish MIN from LC over a 12 month period post hospital discharge in the IMPACC cohort.

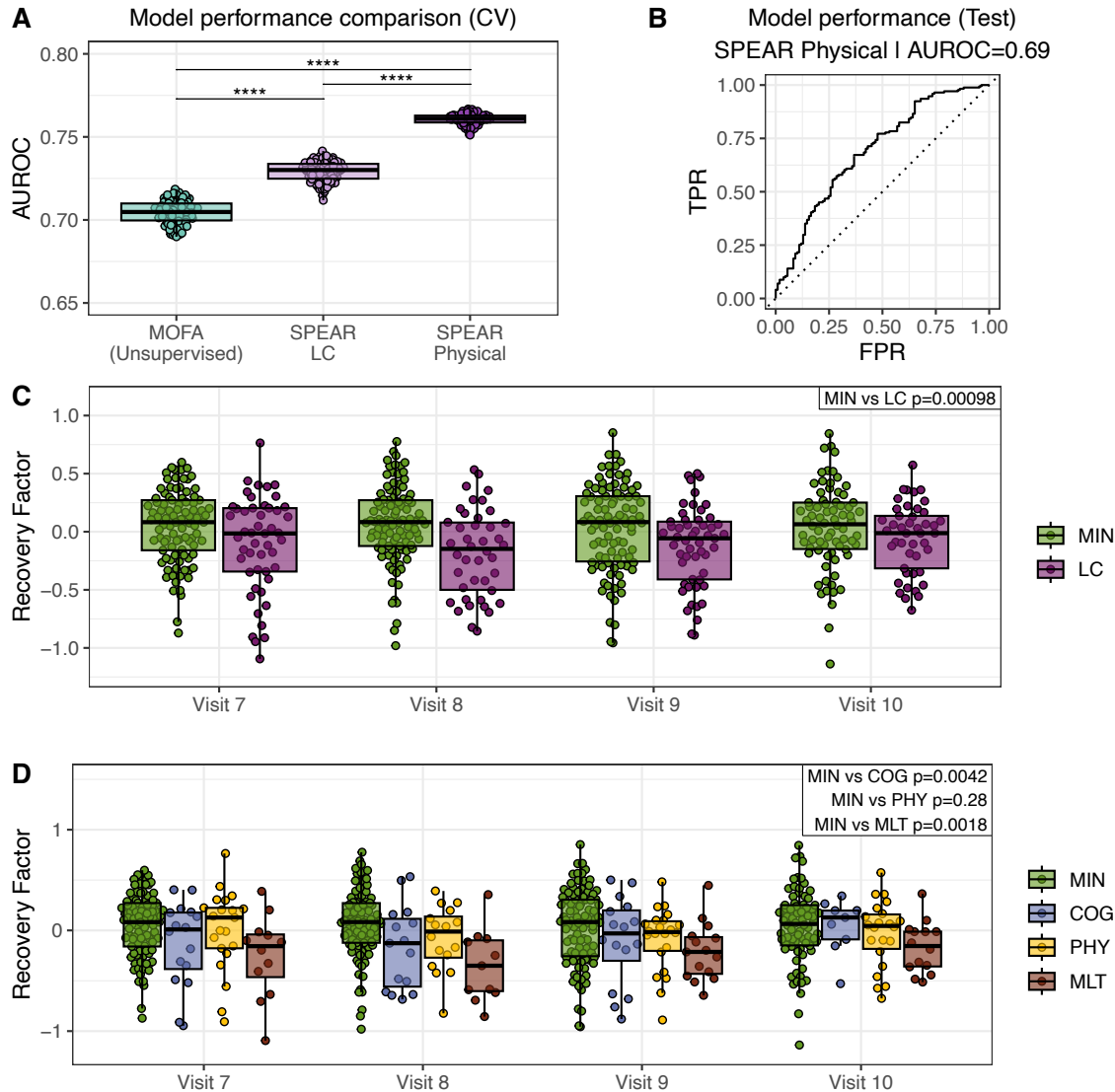


Figure 2. Identification of a convalescent multi-omics recovery factor that discriminates long COVID. (A) Predictive performance of a lasso model trained on the MOFA and SPEAR factors to discriminate LC vs MIN at the event level. The mean AUROC of a 10-fold cross-validation on the Train Cohort, for 100 bootstrapped model training repetitions are shown. Significance was calculated by standard normal approximation of bootstrapped differences between models (t-test, ****adj. p-value ≤ 0.0001) (B) Predictive performance of the SPEAR Physical model to discriminate LC vs MIN on the Test cohort. ROC curve of model (solid line), random classifier (dashed line), and AUROC value are shown. TPR: true positive rate, FPR: false positive rate. (C) Recovery factor scores for the Test cohort of the MIN and LC groups at 3 months (Visit 7), 6 months (Visit 8), 9 months (Visit 9) and 12 months (Visit 10) after hospital discharge. (D) Recovery factor scores of the individual PRO clusters by visit for the Test cohort. P-values in C and D show the significance of the recovery factor score association with MIN vs LC and pairwise PRO cluster combinations, respectively (see methods for association details). See also Figure S2, S3, and S4.

Functional characterization of the recovery factor

To characterize the biologic processes underlying the recovery factor, we performed gene set enrichment analysis (GSEA) for each of the multi-omics assays based on the SPEAR model's internal ranking of the relative importance of each feature for predicting the PRO Physical score. Surveyed pathways included gene sets from the Molecular Signatures Database (MSigDB) Hallmark⁴³ and KEGG⁴⁴ resources as well as metabolite sets from the Subpathway resource⁴⁵⁻⁴⁷. Joint p-values were computed to assess analyte set enrichment across the multi-omics assays. The Hallmark Heme Metabolism transcriptomic pathway was negatively associated with the recovery factor, indicating upregulation in LC participants, whereas the androgenic steroids metabolite set was positively associated with the recovery factor, indicating downregulation in LC participants (Figure 3A). Evaluated individually, several leading edge analytes in the Hallmark Heme Metabolism gene set and androgenic steroids Subpathway metabolite set showed significant associations with LC status (Figure S5 A and B).

SPEAR performs internal significance testing to determine the importance of each analyte in predicting the response variable. The SPEAR Physical model identified 26 analytes across four assays that were significant in the recovery factor (SPEAR Bayesian posterior selection probability ≥ 0.95), and we performed individual associations of these features with LC status in the test cohort, adjusting for age and sex (Figure 3B, Figure S5). Nine of these 26 analytes were from the serum Olink assay. Of these, DNER (Delta And Notch-Like Epidermal Growth Factor-Related Receptor), a non-canonical Notch ligand that has been implicated in promoting tumor growth and metastasis and in supporting wound healing^{48,49} was significantly reduced in LC participants, consistent with a prior study of plasma proteomics in LC subjects²⁸. The remaining serum Olink analytes were negatively associated with the recovery factor. In particular, they included proteins and cytokines associated with chronic inflammatory conditions⁵⁰⁻⁵⁵, particularly endothelial/vascular inflammation (FGF23, FGF21, CXCL9, TNFRSF11B and TNFRSF9 (CD137)), as well as inflammation-associated myeloid regulators⁵⁶⁻⁵⁸ (MMP10 and CSF1). Elevated levels of IL10RB have been previously associated with worse outcomes in acute COVID-19 infection⁵⁹, consistent with elevation under inflammatory conditions. LRG1, a protein elevated in LC participants, is induced by IL-6 and other inflammatory cytokines and has been implicated in angiopathic activity⁶⁰⁻⁶². Phenylacetylglutamate and phenylacetylglutamine are gut microbiota-derived metabolites associated with vascular inflammation and thrombosis⁶³. Finally, the OSBP2 (ORP4) transcript, which encodes an oxysterol binding protein⁶⁴, was a leading edge gene in the Hallmark Heme Metabolism gene set that was elevated in LC participants.

Several metabolites from the androgenic steroids pathway were represented in the 26 significant analytes and were positively associated with the recovery factor, indicating higher levels correlate with better physical function. When we tested these metabolites for their individual association with LC status, five (DHEA-S, epiandrosterone sulfate, androsterone sulfate, 5alpha-androstan-3beta,17beta-diol monosulfate (2), 5alpha-androstan-3beta,17alpha-diol disulfate) were significantly lower in LC participants, adjusting for age and sex (Figure 3B). Androgens can suppress inflammation⁶⁵, suggesting that the higher level of androgenic steroids in participants of

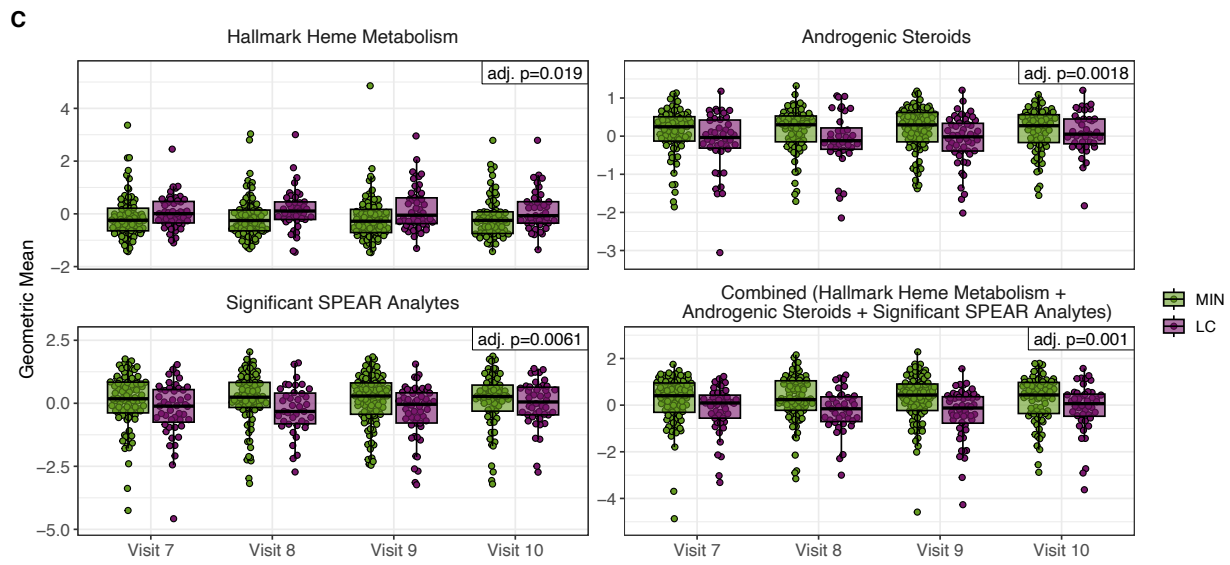
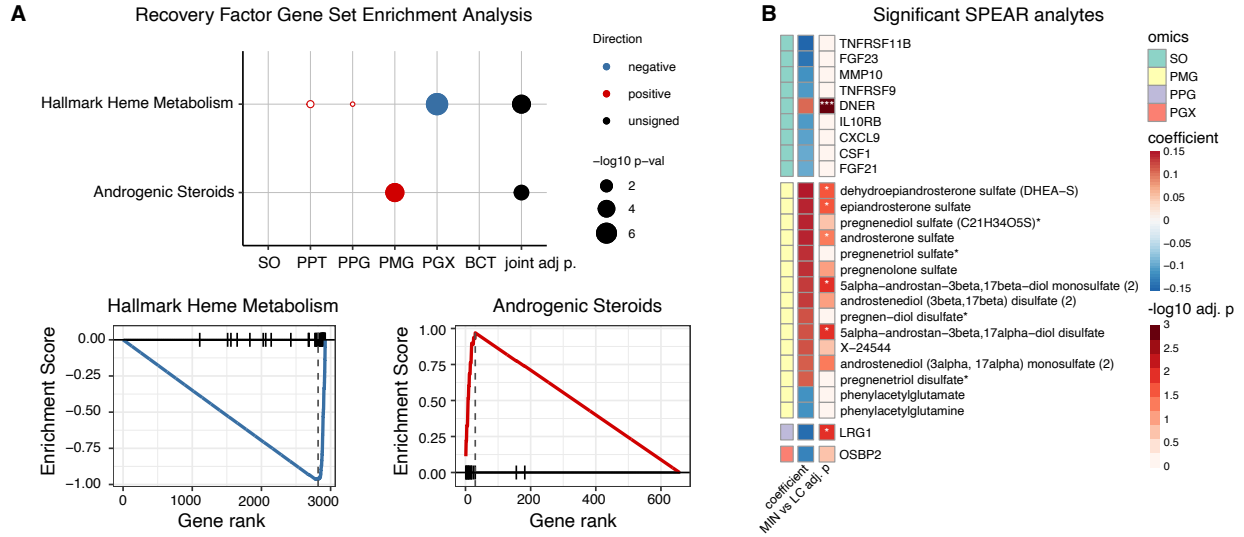
the MIN group could reflect better control of chronic inflammation. These findings are consistent with prior reports showing lower levels of sex hormones in LC³¹. Five metabolites related to pregnenolone were also represented in the significant SPEAR analytes (Figure 3B). Pregnenolone is synthesized from cholesterol as the first step of the steroid hormone biosynthesis pathway and is known to have potent effects as an inhibitor of inflammation⁶⁶ and as a neurosteroid⁶⁷. Altogether, these findings are consistent with a prominent role for persistent inflammation in LC with dysregulation of key analytes that may contribute to symptoms in LC, including elements that drive angiopathy, reduce wound healing, and alter heme metabolism.

The feature sets from heme metabolism and androgenic steroids identified by GSEA analysis combined with the significant SPEAR analytes represent 73 unique features that potentially condense the predictive power of the recovery factor into a smaller feature set. To test this hypothesis, we calculated the geometric mean of the 43 leading edge heme metabolism and 12 androgenic steroid features, as well as the 26 significant SPEAR analytes. All three geometric mean scores were independently significantly associated with LC in the test cohort (Figure 3C). Furthermore, the combined score that includes analytes from all three feature sets discriminates MIN and LC participants with even greater significance (Figure 3C). Thus, while the recovery factor is comprised of weighted contributions from 6,807 features, we have identified a smaller set of 73 unique features that discriminates participants according to LC status in the convalescent period.

Consistent with our finding, the Hallmark Heme Metabolism pathway was previously reported by Hanson et al.²⁹ as an enriched pathway in participants with persisting symptoms 1-3 months after acute SARS-CoV-2 infection compared to participants without persisting symptoms. This cohort comprised 102 participants, including non-hospitalized and hospitalized individuals²⁹. To determine whether the same heme metabolism-related genes were dysregulated in LC participants in the IMPACC and Hanson et al. cohorts, we used the leading edge genes from the significant Hallmark Heme Metabolism pathway in our GSEA analysis (Figure S5A) and calculated the geometric mean gene expression in PBMCs from the Hanson et al. cohort²⁹. We found that our heme metabolism leading edge genes significantly differentiated participants with persistent vs. resolved symptoms after COVID-19 infection at multiple time points in the independent cohort (Figure S5C), validating the reproducibility of the gene expression datasets and underscoring the importance of this subset of heme metabolism genes.

Prior studies have identified altered leukocyte frequencies as a feature of LC^{18,26,29,31,33}. To determine whether similar cellular changes were associated with the recovery factor, we analyzed whole blood CyTOF cell frequencies for 15 parent and 46 child immune cell types in our cohort during convalescence (Figure 3D). We found several cell subsets that were significantly associated with the recovery factor. B cells and CD161⁺ MAIT cells were positively associated with the recovery factor. In contrast, polymorphonuclear leukocytes (PMN) and monocytes, specifically the CD14⁺CD16⁻ classical monocyte subset, were negatively associated with the recovery factor. Together, these findings suggest that a persistent elevation in monocytes and neutrophils, along with a deficit in B cells, is associated with prolonged inflammation during LC. These findings are consistent with a previous report that monocytes are elevated in males with

LC³¹. The decrease in MAIT cells with LC could be another effect of sustained inflammation as lower levels of circulating MAIT cells have been associated with chronic HIV⁶⁸ and hepatitis C⁶⁹ viral infections.



D Recovery Factor association with whole blood CyTOF

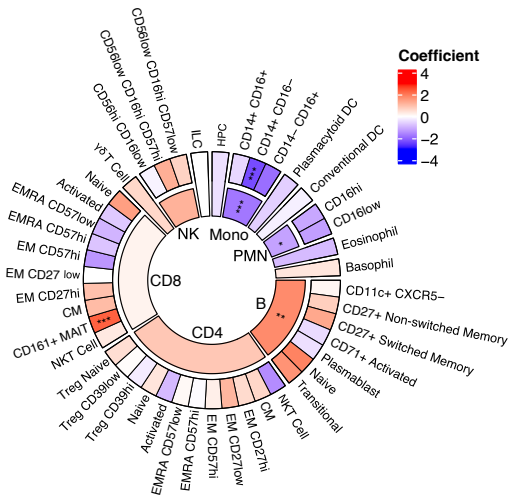


Figure 3. Heme metabolism and androgenic steroid pathways, inflammation-associated serum factors, and altered immune cell composition are associated with the recovery factor during convalescence. (A) GSEA identifies heme metabolism and androgenic steroid pathways as significantly associated with the recovery factor, with significance shown per assay, as well as across assays (joint adj. $p < 0.05$). (B) 26 significant analytes (SPEAR Bayesian posterior selection probability ≥ 0.95) in the recovery factor across different assays (left) were identified using SPEAR factor loadings (middle; coefficient in the factor), and each was tested for association in the test cohort with MIN vs LC groups (right; adj. intercept p -value). (C) Geometric means of analytes from the significantly enriched gene and metabolite sets and/or significant SPEAR analytes are shown per sample at each convalescent visit in the test cohort. The p -values indicate significance of the association with MIN vs LC. (D) Association in the full cohort of whole blood cell counts determined by CyTOF with the recovery factor for parent and child immune cell types. NK: Natural Killer cells, Mono: Monocytes, PMN: polymorphonuclear neutrophils, B: B lymphocytes, CD4: CD4+ T lymphocytes, CD8: CD8+ T lymphocytes. For a full list of the child populations see Table S3. (* adj. p -value < 0.05 , ** adj. p -value < 0.01 , *** adj. p -value < 0.001). See also Figure S5.

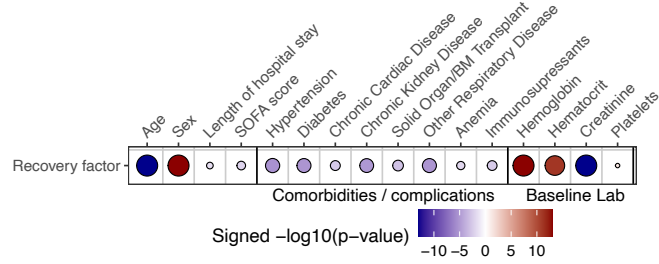
The recovery factor is associated with clinical characteristics and multiple patient reported outcomes in the convalescent period

We next evaluated whether the recovery factor was associated with clinical features and additional clinical outcomes. We tested the association of recovery factor scores with clinical features at hospital admission (i.e., Visit 1), including demographics, comorbidities, complications, and baseline lab measurements (Figure 4A). Several demographic and clinical measures were significantly associated with recovery factor scores, including age, sex, length of hospital stay and the Sequential Organ Failure Assessment (SOFA) score. Notably, anemia as a complication was negatively associated with the recovery factor, whereas hemoglobin and hematocrit baseline measurements showed a significant positive association (Figure 4A). We additionally conducted association testing with Patient Reported Outcome (PRO) measures from surveys conducted at the same visit at which the recovery score was assessed in participants across the convalescent period, correcting for age and sex. The recovery factor score was significantly associated in the test cohort with the PROMIS Physical score, on which the model was trained (Figure 4B), and the EQ-5D-5L score, both of which contained similarly worded questions assessing physical function (Figure 4B). Interestingly, the recovery factor was also correlated with PROMIS Mental and PROMIS Psychosocial Impact scores, although these associations did not remain significant after p -value correction (Figure 4B). We also tested whether recovery factor scores associated with whether participants reported clinical symptoms in the 7 days prior to each visit but found no significance with any symptom group (Figure 4C).

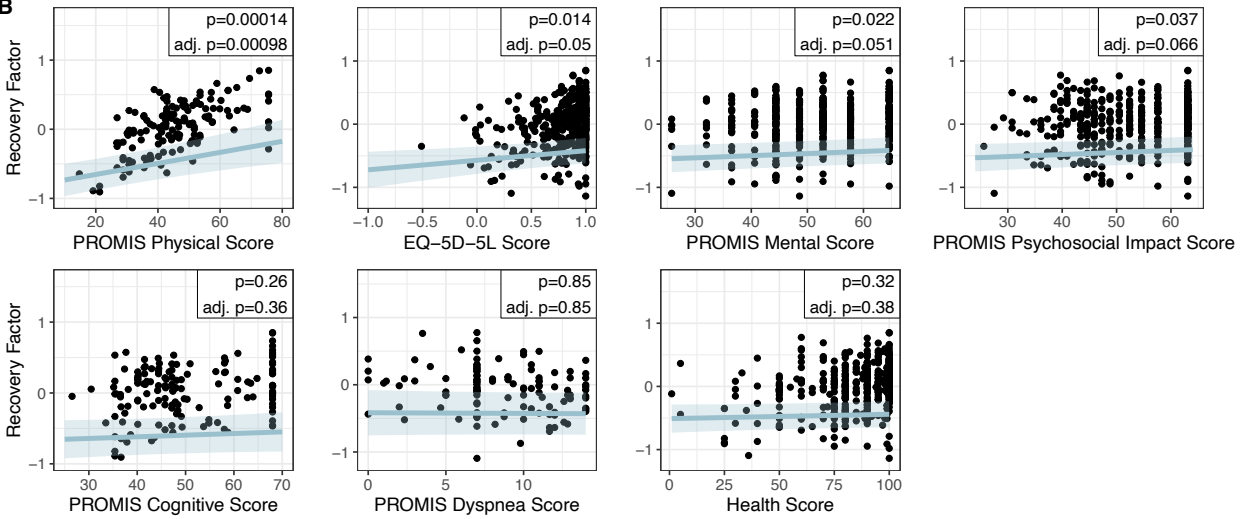
There is a general lack of consensus about whether LC is associated with the severity of acute disease. A previous analysis of clinical features from the IMPACC cohort showed no association between the severity of acute infection, as assessed by clinical trajectory groups, and LC development^{6,18}. However, other studies have found an association⁶. Thus, we sought to determine whether acute disease severity contributed to the association between recovery factor scores and LC status in our cohort. Clinical severity in the IMPACC cohort during the acute phase

was defined by unsupervised clustering of respiratory ordinal score over time, taking discharge status and limitations into account, with trajectory group 1 (TG1) representing the mildest disease course and TG4 representing the most severe disease among participants who survived for at least 28 days post hospitalization⁷⁰. After correcting for acute phase trajectory group assignment, recovery factor scores remain significantly associated with LC at the first three convalescent time points (Figure 4D), indicating that acute clinical severity does not contribute to the association between participant recovery factor scores in the convalescent phase of disease and LC status.

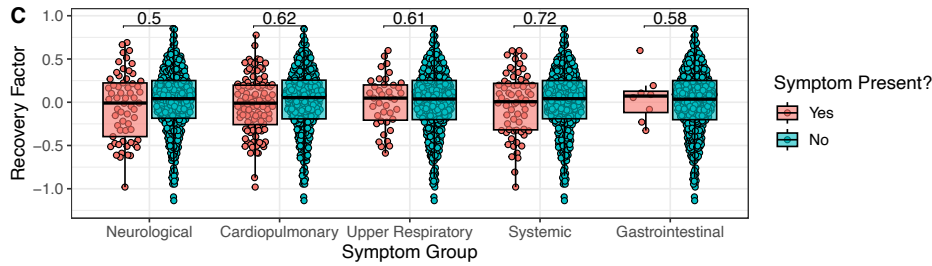
A



B



C



D

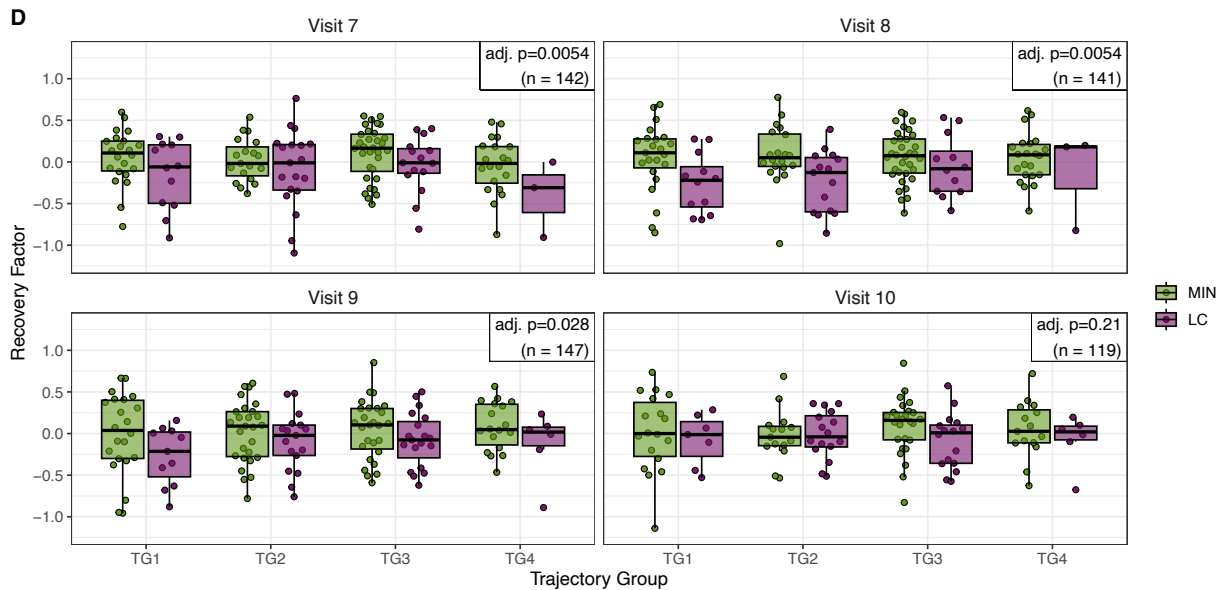


Figure 4. Associations of clinical measurements with recovery factor scores. (A) Association of recovery factor scores with clinical features (demographics, comorbidities, complications and baseline lab measurements). Dot plot shows the signed adjusted p-values indicating the clinical feature term significance from a linear mixed-effect model with enrollment site and participant as random effects to explain the convalescent phase recovery factor scores. Sex and discretized age were further adjusted as fixed effects for clinical features other than sex and age. Only significant associations (adj. p-value <0.05) are shown. **(B)** Associations of recovery factor scores with individual PRO survey scores (PROMIS scale scores, EQ-5D-5L and health score) in the test cohort. Raw and adjusted p-values indicated the PRO score term significance in linear mixed effect models. **(C)** Associations of recovery factor scores with each indicated symptom group in the test cohort. Numbers are the uncorrected significance (p-values) of the symptom group term in linear mixed effect models. **(D)** Recovery factor scores per participant in the test cohort, separated into MIN and LC groups by acute phase trajectory groups, stratified by visit. P-values for panels B-D show the endpoint term of a linear mixed effect model with sex, discretized admit age, and trajectory group as fixed effects and enrollment site as random effect. No individual MIN vs. LC comparisons were significant after p-value correction. (* adj. p-value < 0.05, ** adj. p-value < 0.01, *** adj. p-value < 0.001)

Sex impacts recovery factor scores

The incidence of LC is higher in females than males, despite a higher percentage of males with severe COVID-19 acute disease courses^{21,71}. In the IMPACC Convalescent cohort, nearly half of the female participants presented with long term deficits compared to only ~30% of male participants (Fig. S6A). Assignment to clinical subtypes of LC was not influenced by sex, with similar proportions and numbers of male and female LC participants assigned to COG, PHY and MLT PRO clusters (Figure S6A). However, consistent with the known influence of sex on LC status, sex was a statistically significant covariate in the association of the recovery factor with LC status from Figure 2C ($p=3.6e-7$) and with PRO clusters from Figure 2D (adj. $p<0.001$ in all pairwise comparisons). Thus, we tested if the recovery factor was able to discriminate LC in both males and females by repeating our associations with LC status in the Test cohort separated by sex. Recovery factor scores were significantly associated with the binary assignment to LC vs. MIN groups in females but not in males after p-value adjustment (Figure S6B), although the trend of lower scores in LC participants persisted in males. When considering individual PRO groups, recovery factor scores discriminated between MIN vs. COG and MIN vs. MLT groups for females and MIN vs. MLT PRO groups for males (Figure S6C). Given that the incidence of LC is lower in males, it is notable that recovery factor scores were generally higher in males versus females regardless of LC status.

We next investigated whether the gene and metabolic pathways or top analytes associated with the recovery factor (Figure 3A-B) were differentially represented between sexes, where sex was also a statistically significant covariate (Figure 3C, adj. $p<0.001$). The geometric mean score of the heme metabolism pathway approached significance but was no longer significantly associated with LC when the cohort was divided into male and female subsets after p-value correction (Figure S6D). Geometric mean scores for androgenic steroids and significant SPEAR analytes (which contained several top metabolites from the androgenic steroids pathway) were significantly

associated with LC only in males (Figure S6D). The androgenic steroid scores were higher in males than females irrespective of LC status (Figure S6D), corresponding to the higher overall recovery factor score in males (Figure S6B). Notably, the combined score of the top 73 unique features from Hallmark Heme Metabolism, androgenic steroids, and significant SPEAR analytes remained significantly associated with LC in both sexes (Figure S6D).

Vaccination is not associated with altered recovery factor scores

Our cohort was enrolled prior to the national rollout of SARS-CoV-2 vaccines for the general population. During the longitudinal post-hospitalization follow-up period, as vaccines became broadly available, close to 75% of the participants in the IMPACC Convalescent cohort received a SARS-CoV-2 vaccination (Figure S7A-B). To assess the potential influence of the vaccine response on the immune profiling data and thus the recovery factor, we compared recovery factor scores per visit for events occurring before and after the first vaccination dose, as well as events occurring within a three-week period after any vaccination dose, when vaccine responses have been shown to impact immune profiles^{72,73}. No significant difference was found in recovery factor scores across these comparisons, indicating a negligible effect of vaccination on immune profiles related to LC in our patient cohort (Figure S7 C-D).

Recovery factor scores during the acute disease phase associate with LC status during convalescence

We next investigated whether the immune elements identified in the recovery factor were detectable during the acute infection phase, prior to development of LC. We computed recovery factor scores using immune profiling data from all participants in the Convalescent cohort during their acute phase visits (Visits 1 to 6, spanning hospital admission through 26-35 days post-admission). Remarkably, we found that recovery factor scores were significantly higher in MIN versus LC participants as early as hospital admission (Visit 1) and consistently during the acute period (Figure 5A; Figure S8A). Recovery factor scores were also significantly higher in MIN versus COG groups and MIN versus PHY groups in the acute phase when assessed across the 28-day time course (Figure 5B; Figure S8B). Geometric means of heme metabolism and androgenic steroid pathway analytes from the recovery factor, as well as the 26 significant SPEAR recovery factor analytes were also significantly associated with LC status during the acute phase. The combined geometric mean score of analytes from these three feature groups in acute phase data associated most significantly with MIN versus LC status (Figure 5C), as it did previously in the convalescent phase (Figure 3C).

We further assessed whether altered circulating immune cell composition in the acute phase of disease could contribute to acute-phase recovery factor scores. Association testing of recovery factor scores with whole blood CyTOF measurements during the acute disease phase showed that CD4+ and CD8+ T cells, conventional and plasmacytoid dendritic cells, eosinophils, basophils, and CD56hi CD16low natural killer (NK) cells were significantly positively associated with the recovery factor scores. Within the CD4+ T cell compartment, naïve, central memory (TCM), and effector memory (TEM) subsets, as well as non-naïve regulatory T cells (Treg) were significantly associated with recovery factor scores, while activated CD4+ T cells were inversely correlated. Within the CD8+ T cell compartment, naïve, TCM, and TEM subsets were positively

associated with recovery factor scores, as were NKT cells and CD161+ MAIT cells. In contrast, monocytes, neutrophils, B cells, and plasmablasts in the acute phase were negatively associated with recovery factor scores (Figure 5D). These findings are consistent with a previous study that found higher plasmablast counts and lower total CD4+ T, total CD8+ T, CD4+ TEM, CD8 TEM, Treg, NK, and dendritic cell counts in immune-cell populations sampled at days 0-14 after infection in COVID-19 patients who experienced persisting symptoms at days 91-180 after infection²⁹. The similarities across both studies are indicative of an acute blood immune cell type signature of LC that is robust to variance in patient cohorts and LC definition.

In summary, our findings indicate that the major biologic signatures of the recovery factor that stratify LC from recovered participants in the convalescent phase – elevated heme metabolism gene signatures, reduced androgenic steroids, increased circulating inflammatory mediators, and increased monocytes and neutrophils – are evident early in the acute phase of disease.

in enriched gene and metabolic sets and/or significant SPEAR analytes during the acute phase. No individual per-visit comparisons were significant after p-value correction. P-values in top-right box in A-C show the significance of the recovery factor score or geometric mean signature association with MIN vs LC or pairwise PRO cluster combinations. Bars above the boxplots show the pairwise significance across groups in a per-visit comparison (** $p < 0.01$, * $p < 0.05$). **(D)** Recovery factor scores association with whole blood CyTOF immune cell populations during the acute phase (* adj. p-value < 0.05 , ** adj. p-value < 0.01 , *** adj. p-value < 0.001). See also Figure S8.

Acute-phase recovery factor scores distinguish acute disease severities and predict LC risk irrespective of acute severity

We investigated the full IMPACC study cohort ($n = 1,148$ participants with at least one sample measurement for the omics modalities included in our model) to assess whether recovery factor scores determined from acute phase data would associate with patient severity trajectory group assignments. For this analysis we included participants who did not survive beyond 28 days post-hospital admission and participants without biospecimens and/or surveys during the convalescent phase. Recovery factor scores were significantly associated longitudinally with acute disease trajectory groups and were highest in participants with milder disease courses (TG1-TG3) and lowest in participants with the most severe acute disease trajectories (TG4 and TG5) (Figure 6A). Acute-phase recovery factor scores increased over time for participants in all trajectory groups except TG5, the most severe group in which participants died by day 28 after hospital admission (Figure 6A). To assess whether the association between acute recovery factor scores and convalescent LC status was simply due to acute recovery factor scores being an indicator of acute disease severity, we repeated the association test including trajectory group as a covariate at each visit (Figure 6B) and longitudinally (Figure S8C). LC status was still significantly associated with acute recovery factor scores even after taking trajectory group into account. These findings suggest that recovery factor scores in the acute phase contain valuable information for predicting convalescent LC status beyond its correlation with acute disease severity.

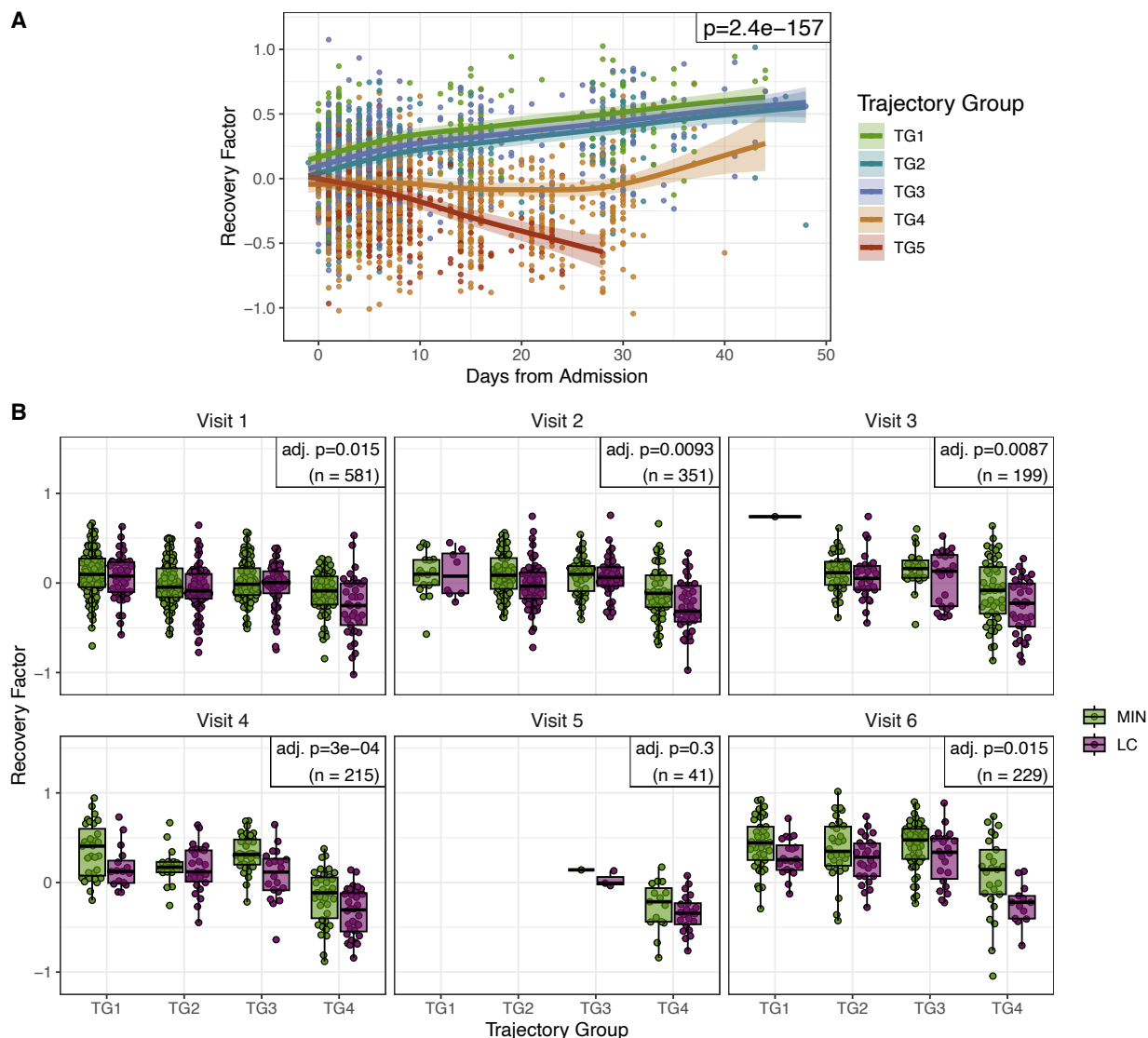


Figure 6. Recovery factor scores associate with acute disease phase trajectory groups, but identify LC irrespective of acute severity. (A) Longitudinal analysis of acute recovery factor scores for the full IMPACC cohort stratified by trajectory group ($N=1,148$ participants). P-value shows the significance of the trajectory group term in a longitudinal model correcting for age and sex as fixed effects and enrollment site and participant ID as random effects. **(B)** Recovery factor scores in the acute phase by convalescent MIN/LC label, stratified by acute trajectory group and visit number. P-values show significance in distinguishing MIN vs. LC labels in linear mixed models with sex, discretized admit age, and trajectory group as fixed effects and enrollment site and participant ID as random effects, performed separately for each acute visit and corrected across all visits.

Machine learning models based on the recovery factor scores at the acute or the convalescent phase together with clinical features predict LC status during the convalescent phase

Building on the strength of the recovery factor to predict LC even from the acute phase of disease, we assessed whether a combination of the recovery factor scores and clinical characteristics could improve the predictive performance. We selected clinical features including the routinely recorded elements of age at enrollment, sex, body mass index (BMI), length of hospital stay, Sequential Organ Failure Assessment (SOFA) score, along with Spike IgG antibody titers and viral load (SARS-CoV-2 N1 PCR) during the acute phase of disease, which were previously associated with LC status in the IMPACC cohort¹⁸ (Figure S9), as well as the presence of comorbidities including hypertension, diabetes, chronic cardiac disease, chronic kidney disease, malignant neoplasms, chronic neurological disorders, liver dysfunction/failure, history of transplants, smoking/vaping, asthma, respiratory diseases other than asthma, substance use, HIV infection, and the total number of comorbidities (Figure S10A). Lasso models trained exclusively on the mean recovery factor scores during the acute phase (SPEAR acute, mean cross-validation AUROC 0.64) outperformed models trained exclusively on clinical features (Clinical, mean cross-validation AUROC 0.63) (Figure S10B). The performance in predicting LC was further improved when training models on the combination of baseline and acute-phase clinical features and recovery factor scores (SPEAR acute + clinical, mean cross-validation AUROC 0.66) (Figure S10B). The lasso model trained on the recovery factor scores during the convalescent phase (mean cross-validation AUROC 0.74) also outperformed the clinical-only model and was slightly improved when combined with clinical features (mean cross-validation AUROC 0.75) (Figure S10C). In addition, we assessed whether a reduced set of analytes based on the 26 SPEAR significant analytes could achieve similar predictive performance to the models trained on the full recovery factor (Figure S10C). Such streamlined models would be advantageous in a clinical setting or to develop clinical diagnostics tests to identify individuals with LC more easily. Even though the best performing reduced analyte set models (SPEAR significant analytes) had a lower predictive power than the models trained on the full convalescent recovery factor scores, this sparse model achieved a mean cross-validation AUROC of 0.69, suggesting its utility.

We further evaluated whether single time point measurements of the recovery factor scores could achieve similar performance as the models with averaged scores over the acute or convalescent phases. Interestingly, we observed a slowly increasing trend in predictive performance for each individual time point with increased visit numbers during earlier visits, which peaked at Visit 8 (6 months after hospital discharge) with a mean cross-validation AUROC of 0.74 (Figure S10D). The 6-month time point was also the strongest individual time point predictor when the model included the clinical features (mean cross-validation AUROC of 0.76) (Figure S10D). Notably, recovery factor scores as early as Visit 1 provided predictive performance (mean cross-validation AUROC 0.63), indicating that the recovery factor captures early predictive features of LC during the acute disease phase, albeit the signal at this early time point is not as strongly predictive as later in the convalescent phase.

DISCUSSION

In this study, we applied supervised multi-omics integration methods to identify biologic features associated with LC in 513 participants from the IMPACC cohort. We took advantage of data availability from this cohort that was followed longitudinally after hospitalization with COVID-19 through the first 28 days of the acute phase of disease and subsequently for up to one year after hospital discharge^{18,70}. The IMPACC cohort is unique in its comprehensive inclusion of clinical data, biospecimens, and quarterly patient-reported outcome surveys, combined with multi-omics immunophenotyping at multiple time points throughout the acute and convalescent phases of the disease. A previous study of this cohort identified demographic and clinical risk factors associated with the development of LC, which included female sex, comorbidities such as chronic heart, lung, and neurologic diseases, and a longer length of hospital stay¹⁸. Here, we report the first study of the IMPACC cohort to analyze biologic data collected during the convalescent phase of disease, allowing us to identify a multi-omics “recovery factor” capable of discriminating participants who recover with minimal deficits from those who experience a variety of clinical LC symptoms. Notably, we find that as early as 72h after hospital admission for COVID-19, recovery factor scores predict which patients, irrespective of their acute disease severity, will go on to experience LC. Biologic features associated with the recovery factor score indicate that reduced androgenic steroid levels, increased heme metabolism signatures, and persistent elevation of inflammation-associated serum proteins are hallmarks of LC that both identify individuals with LC during convalescence and predict which acute COVID-19 patients will experience LC.

Increased levels of androgenic steroids in serum were positively associated with recovery factor scores and with participants who recovered from COVID-19 with minimal deficits. Seven of the twelve leading-edge androgenic steroid metabolites, as well as Oxysterol Binding Protein 2 (OSBP2), a leading-edge gene in the Hallmark Heme Metabolism pathway, were also included within the list of 26 analytes that were statistically significant within the recovery factor. Limited studies have elucidated the role of reduced androgenic steroids in LC, but in agreement with our findings, lower testosterone levels have been associated with increased LC symptomatology in both males and females³¹. We observed that several intermediate metabolites in the canonical steroid hormone biosynthesis pathway were associated with the recovery factor and were decreased in LC participants, including sulfated forms of testosterone precursors (pregnenolone and DHEA) and downstream metabolites (androsterone, epiandrosterone, and 5alpha-androstan-3beta,17beta-diol). Notably, our list of top androgenic steroid metabolites has strong overlap with the androgenic steroid signature from an independent all-female cohort of healthy controls compared to myalgic encephalomyelitis / chronic fatigue syndrome (ME/CFS) patients⁷⁴. All six androgenic steroid metabolites that were significantly elevated in healthy controls in this study (DHEA-S, androstenediol (3alpha, 17alpha) monosulfate (2), androstenediol (3beta,17beta) disulfate (2), 5alpha-androstan-3beta,17alpha-diol disulfate, androsterone sulfate, epiandrosterone sulfate)⁷⁴ were also statistically significant in our recovery factor, strongly implicating this signature in the shared symptomatology, such as fatigue, post exertional malaise, and sleep disturbances, between ME/CFS and LC⁷⁵. In our study, a geometric mean score consisting of leading-edge metabolites from the androgenic steroids pathway significantly differentiated MIN vs. LC in the entire cohort. Nonetheless, we note that when the cohort was

subset into males and females, the geometric mean scores for the androgenic steroid pathway only significantly stratified LC from recovered male participants, although levels of androgenic steroid scores also trended lower in females with LC.

The fact that pregnenolone (sulfate) and many of its downstream metabolites were all significantly elevated in the recovery factor, despite being generated by distinct enzymes, suggests that a rate-limiting enzymatic step upstream of pregnenolone may explain the observed signature. We observed that cholesterol, which is cleaved by the mitochondrial enzyme CYP11A1 to produce pregnenolone in the first step of the steroid hormone biosynthesis process, was detected in the metabolomics assay but was not identified as a significant metabolite in the recovery factor. CYP11A1 was not detected in our proteomics or transcriptomic assays, likely due to localized activity in the adrenal gland. Further experimentation is needed to test if impaired generation of pregnenolone from cholesterol in LC is responsible for decreased levels of intermediate androgenic steroid metabolites. Alternatively, the observed decrease of sulfated androgenic steroid levels in LC may reflect altered activity of sulfatases or sulfotransferase enzymes. Notably, some sulfonated steroids, such as DHEAS, a leading-edge gene in the androgenic steroid pathway and a top significant analyte in the SPEAR recovery factor, have been shown to exhibit immunosuppressive and anti-inflammatory effects, particularly in neuroinflammation⁷⁶. In addition, testosterone has shown to play an immunomodulatory role and is often reduced in patients with other critical illnesses⁷⁷. Thus, elevated sulfated steroids in the recovery factor may contribute to the reduction in inflammation associated with recovery from COVID-19 with minimal deficits⁷⁶.

In addition to altered androgenic steroids, the transcriptional signature of heme metabolism in PBMCs was inversely associated with the recovery factor, such that LC participants had elevated expression of genes associated with the heme metabolism pathway during the convalescent period. As with androgenic steroids, we found the leading-edge genes of the heme metabolism pathway were also higher during the acute phase of disease in participants who would later experience LC. Notably, overexpression of a heme metabolism signature in blood was recently reported for a separate cohort of 102 participants, including non-hospitalized and hospitalized COVID-19 patients evaluated 1-3 months after infection, which found this pathway to be enriched in participants experiencing persisting symptoms²⁹. The authors related the elevated heme metabolism signature to stress erythropoiesis induced in response to inflammation-associated anemia driven by IL-6-mediated hepcidin upregulation^{29,78}. In that study, participants who experienced LC had reduced iron and HGB 2 weeks to 1 month after COVID-19 infection. This iron restriction was proposed not only to induce anemia, but also to impair lymphocyte function. The subsequent delayed resolution of acute infection might result in sustained inflammation and persistent anemia up to at least 6 months after infection, which may partially explain the systemic symptomatology of acute COVID-19 and LC. Anemia of inflammation, also known as anemia of chronic disease, is a common complication associated with chronic inflammatory illnesses such as chronic kidney disease, congestive heart failure, as well as ICU admission⁷⁹. Anemia as a complication was also significantly negatively associated with the recovery factor and was shown to be associated with PRO clusters in our previous study¹⁸, whereas hemoglobin and hematocrit baseline measurements showed a significant positive association with the recovery factor. Our findings are largely consistent with this study, as we also find evidence of persistent inflammation

in LC participants: inflammation-associated serum factors, such as CXCL9, CSF1, and FGF21, were identified by SPEAR as significant analytes in the recovery factor, albeit they did not reach significance when associated individually with LC status. Of note, FGF21 measured in the acute phase was previously associated with cognitive and multidomain deficit PRO clusters relative to MIN in this cohort¹⁸, and has been proposed as a biomarker for chronic inflammation in myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS)⁶⁰, a complex chronic disease that overlaps clinically with LC. Furthermore, LRG1, which is activated by the inflammatory IL-6/STAT3 pathway is significantly elevated in LC participants, perhaps contributing to vascular pathology in LC⁶⁰⁻⁶². Although IL-6 (SPEAR Bayesian posterior selection probability = 0.86) did not reach significance in the SPEAR factor, it was within the top 50 analytes of the factor. Along with these persistent signatures of inflammation, we also identified elevated heme metabolism gene expression signatures as a key feature of LC. In fact, expression levels of the leading-edge heme metabolism genes we identified in our study could also predict LC participants in another cohort²⁹, demonstrating that the same heme metabolism signature during the acute phase of COVID-19 can predict which patients will experience LC when tested in participants from distinct studies. Together, these findings point to persistent inflammation, driving anemia and stress erythropoiesis as key drivers of LC.

Across all time points, consistent immune cell subsets were associated with the recovery factor. In both the acute and convalescent disease phases, CD161+ MAIT cell frequencies were positively associated with the recovery factor, while monocyte, neutrophil, and CD14+ CD16-classical monocyte cell frequencies were negatively associated with the recovery factor. However, the B cell population is negatively associated with the recovery factor during the acute disease phase, but becomes positively associated during convalescence. We also found several points of agreement for immune cell frequencies in the recovery factor with previous reports of LC subjects. While comparing immune composition across cohorts can be challenging due to differences in cohort construction and cell type annotation, we found that CD4+ and CD8+ T cells, plasmacytoid DCs, and conventional DCs are positively associated with the recovery factor during the acute phase of disease, while monocytes, neutrophils, and B cells are negatively associated, suggesting reduced T cell immunity relative to inflammatory innate immunity in the acute phase of disease in individuals most susceptible to LC. This finding is in keeping with cellular trends observed in Hanson et al.²⁹ in the acute phase of disease and in males during convalescence by Silva and colleagues³¹. Consistent with Klein et al.²⁶, we do not find significant associations of the recovery score with naive CD4+ or naive CD8+ T cell subsets in the convalescent phase of COVID-19.

A previous study from IMPACC identified a 'severity factor' that significantly associated with clinical outcomes during the acute phase of disease³⁹. Given that both the severity factor and the recovery factor are associated with inflammatory signatures during acute COVID-19, we compared immune cell types associated with the recovery factor to those associated with the severity factor at this stage of disease. There were notable similarities: for example, the combined monocyte, B cell, and neutrophil populations are negatively associated with the recovery factor and positively associated with the severity factor. Likewise, CD4+ T cell and CD8+ T cell populations are positively associated with the recovery factor and negatively associated with the severity factor. These cellular associations suggest that inefficient adaptive immunity during the

acute phase of disease, with elevated frequencies of inflammatory innate cells contribute to LC susceptibility. This model is consistent with our past¹⁸ and current findings in the IMPACC cohort that reduced levels of anti-SARS-CoV-2 antibodies and increased viral titers in patients within the first 72 hours of hospital admission are associated with participants who will experience LC (Figure S9). These findings are also consistent with other reports that hospitalized individuals are more susceptible to LC than non-hospitalized individuals^{14,81}. Nonetheless, our recovery factor predicts which patients will experience LC, irrespective of acute disease severity, indicating that the model has learned features of COVID-19 beyond inflammation that are associated with COVID-19 severity.

Although symptom groups, such as respiratory symptoms, were found to be significantly associated with LC in this cohort¹⁸, the multi-omics recovery factor does not associate with a particular clinical symptom group. Instead, it captures biomarkers predictive of global physical deficits, as reported by patients after acute COVID-19 disease. While assessing the entire multi-omics SPEAR factor in convalescent patients is impractical, our findings indicate that assessing the 26 significant SPEAR analytes would aid in LC diagnoses (Fig S10c).

This study has several limitations. The reliance on self-reported survey data to identify symptoms and classify participants into MIN/LC and individual PRO clusters may introduce potential biases. To address this, population-normalized PRO measure scores were utilized, and comparisons were made to pre-illness health status when possible. Additionally, the surveys were developed early in the pandemic (March 2020), before the full spectrum of LC symptoms was characterized, and thus did not capture current commonly-recognized manifestations such as brain fog, fatigue, sleep disturbances, neuropathy, and dysautonomia. Self-selection bias may also be present, as patients with severe LC symptoms might have been less likely to respond to the surveys. As the study cohort was recruited during early phases of the pandemic (May 2020 through March 2021), it consists of individuals infected with the original SARS-CoV-2 strain and does not include data on subsequent variants of concern, limiting the generalization of the findings to other later dominant strains. Vaccination data were self-reported and limited to the post-acute phase since enrollment was largely completed prior to vaccine rollout, and exact vaccination dates were unavailable for some participants. Latent virus reactivation was not considered in this model. Furthermore, as part of the study design, all participants in the IMPACC cohort were hospitalized for COVID-19. Consequently, the multi-omics factors were constructed without incorporating profiles from individuals with COVID-19 who did not require hospitalization or healthy controls, potentially introducing a bias toward those with more severe disease. However, we were able to discover in our hospitalized cohort molecular signatures that were previously observed in cohorts that included non-hospitalized and asymptomatic SARS-CoV-2 infected participants²⁹.

Despite these limitations, the study possesses significant strengths. In addition to the prospective design, with acute and convalescent longitudinal multi-omics profiling, enrollment through multiple sites across the United States enhances the diversity and broad representation of the cohort and mitigates potential participant recruitment biases, contributing to the robustness of the findings.

In conclusion, supervised multi-omics factor construction of immune profiling data from SARS-CoV-2 infected participants who recovered with minimal deficits or experienced LC indicates that immune cell types and serum factors associated with inflammation, reduced androgenic steroids, and an elevated heme metabolism signature predict which participants will experience LC, irrespective of acute disease severity. Moreover, these signatures are maintained into convalescence, indicating that persistent inflammation is likely a key driver of LC. Further studies will be needed to determine why inflammation persists in some COVID-19 patients. We did not assess persistent SARS-CoV-2 viral loads or viral reactivation in this study; however, a recent study of IMPACC data⁸² suggests reactivation of latent viruses could contribute, consistent with other studies^{26,30,31}. Altogether, our data, paired with prior congruent reports, suggest that impaired lymphocyte function early in COVID-19 reduces cellular and humoral adaptive immunity and contributes to high SARS-CoV-2 viral loads. Elevated viral loads can trigger innate immune cell responses that increase inflammatory cytokines, driving inflammation-associated anemia that further reduces lymphocyte function, which could enable reactivation of latent viruses. Such persistent inflammation that is not successfully resolved likely leads to LC pathology. Strategies to break the cycle of inflammation and correct the inflammation-associated anemia may promote recovery from LC, and merit further investigation.

RESOURCE AVAILABILITY

Lead Contact

Requests for further information should be directed to Lauren Ehrlich (lehrlich@austin.utexas.edu)

Materials availability

This study did not generate new, unique reagents.

Data and Code availability

Data used in this study is available on the ImmPort repository under accession number SDY1760 and in the Database of Genotypes and Phenotypes (dbGaP) under the accession number phs002686.v2.p2.

All code used in this study is deposited on Bitbucket (<https://bitbucket.org/kleinsteini/impacc-public-code/src/master/multiomics-longcovid/>).

CONSORTIA

The members of the IMPACC Network are:

Clinical and Data Coordinating Center (CDCC), Precision Vaccines Program, Boston Children's Hospital, Harvard Medical School, Boston, MA 02115, USA:

Al Ozonoff, Joann Diray-Arce, Jing Chen, Alvin T. Kho, Carly E. Milliren, Annmarie Hoch, Ana C. Chang, Kerry McEnaney, Caitlin Syphurs, Brenda Barton, Claudia Lentucci, Maimouna D. Murphy, Mehmet Saluvan, Tanzia Shaheen, Shanshan Liu, Marisa Albert, Arash Nemati Hayati, Robert Bryant, James Abraham, Mitchell Cooney, Meagan Karoly

Benaroya Research Institute, University of Washington, Seattle, WA 98101, USA:

Matthew C. Altman, Naresh Doni Jayavelu, Scott Presnell, Bernard Kohr, Tomasz Jancsyk, Azlann Arnett

La Jolla Institute for Immunology, La Jolla, CA 92037, USA:

Bjoern Peters, James A. Overton, Randi Vita, Kerstin Westendorf

Knocean Inc. Toronto, ON M6P 2T3, Canada:

James A. Overton

Precision Vaccines Program, Boston Children's Hospital, Harvard Medical School, Boston, MA 02115, USA:

Ofer Levy, Hanno Steen, Patrick van Zalm, Benoit Fatou, Kinga K. Smolen, Arthur Viode, Simon van Haren, Meenakshi Jha, David Stevenson, Sanya Thomas, Boryana Petrova, Naama Kanarek

Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA:

Lindsey R. Baden, Kevin Mendez, Jessica Lasky-Su, Alexandra Tong, Rebecca Rooks, Michael Desjardins, Amy C. Sherman, Stephen R. Walsh, Xhoi Mitre, Jessica Cauley, Xiaofang Li, Bethany Evans, Christina Montesano, Jose Humberto Licon, Jonathan Krauss, Nicholas C. Issa, Jun Bai Park Chang, Natalie Izaguirre

Metabolon Inc, Morrisville, NC 27560, USA:

Scott R. Hutton, Greg Michelotti, Kari Wong

Prevention of Organ Failure (PROOF) Centre of Excellence, University of British Columbia, Vancouver, BC V6T 1Z3, Canada:

Scott J. Tebbutt, Casey P. Shannon

Case Western Reserve University and University Hospitals of Cleveland, Cleveland, OH 44106, USA:

Rafick-Pierre Sekaly, Slim Fourati, Grace A. McComsey, Paul Harris, Scott Sieg, George Yendewa, Mary Consolo, Heather Tribut, Susan Pereira Ribeiro

Drexel University, Tower Health Hospital, Philadelphia, PA 19104, USA:

Charles B. Cairns, Elias K. Haddad, Michele A. Kutzler, Mariana Bernui, Gina Cusimano, Jennifer Connors, Kyra Woloszczuk, David Joyner, Carolyn Edwards, Edward Lee, Edward Lin, Nataliya Melnyk, Debra L. Powell, James N. Kim, I. Michael Goonewardene, Brent Simmons, Cecilia M. Smith, Mark Martens, Brett Croen, Nicholas C. Semenza, Mathew R. Bell, Sara Furukawa, Renee McLin, George P. Tegos, Brandon Rogowski, Nathan Mege, Kristen Ullring, Pam Schearer, Judie Sheidy, Crystal Nagle

MyOwnMed Inc., Bethesda, MD 20817, USA:

Vicki Seyfert-Margolis

Emory School of Medicine, Atlanta, GA 30322, USA:

Nadine Roupheal, Steven E. Bosinger, Arun K. Boddapati, Greg K. Tharp, Kathryn L. Pellegrini, Brandi Johnson, Bernadine Panganiban, Christopher Huerta, Evan J. Anderson, Hady Samaha, Jonathan E. Sevransky, Laurel Bristow, Elizabeth Beagle, David Cowan, Sydney Hamilton, Thomas Hodder, Amer Bechnak, Andrew Cheng, Aneesh Mehta, Caroline R. Ciric, Christine Spainhour, Erin Carter, Erin M. Scherer, Jacob Usher, Kieffer Hellmeister, Laila Hussaini, Lauren Hewitt, Nina Mcnair, Susan Pereira Ribeiro, Sonia Wimalasena

Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA:

Ana Fernandez-Sesma, Viviana Simon, Florian Krammer, Harm Van Bakel, Seunghee Kim-Schulze, Ana Silvia Gonzalez-Reiche, Jingjing Qi, Brian Lee, Juan Manuel Carreño, Gagandeep Singh, Ariel Raskin, Johnstone Tcheou, Zain Khalil, Adriana van de Guchte, Keith Farrugia, Zenab Khan, Geoffrey Kelly, Komal Srivastava, Lily Q. Eaker, Maria C. Bermúdez-González, Lubbertus C.F. Mulder, Katherine F. Beach, Miti Saksena, Deena Altman, Erna Kojic, Levy A. Sominsky, Arman Azad, Dominika Bielak, Hisaaki Kawabata, Temima Yellin, Miriam Fried, Leeba Sullivan, Sara Morris, Giulio Kleiner, Daniel Stadlbauer, Jayeeta Dutta, Hui Xie, Manishkumar Patel, Kai Nie, Brian Monahan

Immunai Inc., New York, NY 10016, USA:

Adeeb Rahman

Oregon Health & Science University, Portland, OR 97239, USA:

William B. Messer, Catherine L. Hough, Sarah A.R. Siegel, Peter E. Sullivan, Zhengchun Lu, Amanda E. Brunton, Matthew Strand, Zoe L. Lyski, Felicity J. Coulter, Courtney Micheletti

Stanford University School of Medicine, Palo Alto, CA 94305, USA:

Holden Maecker, Bali Pulendran, Kari C. Nadeau, Yael Rosenberg-Hasson, Michael Leipold, Natalia Sigal, Angela Rogers, Andrea Fernandes, Monali Manohar, Evan Do, Iris Chang, Alexandra S. Lee, Catherine Blish, Henna Naz Din, Jonasel Roque, Linda N. Geng, Maja Artandi, Mark M. Davis, Neera Ahuja, Samuel S. Yang, Sharon Chinthrajah, Thomas Hagan, Tyson H. Holmes, Koji Abe

David Geffen School of Medicine at the University of California Los Angeles, Los Angeles CA 90095, USA:

Elaine F. Reed, Joanna Schaeenman, Ramin Salehi-Rad, Adreanne M. Rivera, Harry C. Pickering, Subha Sen, David Elashoff, Dawn C. Ward, Jenny Brook, Estefania Ramires-Sanchez, Megan Llamas, Claudia Perdomo, Clara E. Magyar, Jennifer Fulcher

University of California San Francisco, San Francisco, CA 94115, USA:

David J. Erle, Carolyn S. Calfee, Carolyn M. Hendrickson, Kirsten N. Kangelaris, Viet Nguyen, Deanna Lee, Suzanna Chak, Rajani Ghale, Ana Gonzalez, Alejandra Jauregui, Carolyn Leroux, Luz Torres Altamirano, Ahmad Sadeed Rashid, Andrew Willmore, Prescott G. Woodruff, Matthew F. Krummel, Sidney Carrillo, Alyssa Ward, Charles R. Langelier, Ravi Patel, Michael Wilson, Ravi Dandekar, Bonny Alvarenga, Jayant Rajan, Walter Eckalbar, Andrew W. Schroeder, Gabriela K. Fragiadakis, Alexandra Tsitsiklis, Eran Mick, Yanedth Sanchez Guerrero, Christina Love, Lenka Maliskova, Michael Adkisson, Aleksandra Leligdowicz, Alexander Beagle, Arjun Rao, Austin Sigman, Bushra Samad, Cindy Curiel, Cole Shaw, Gayelan Tietje-Ulrich, Jeff Milush, Jonathan Singer, Joshua J. Vasquez, Kevin Tang, Legna Betancourt, Lekshmi Santhosh, Logan Pierce, Maria Tecero Paz, Michael Matthay, Neeta Thakur, Nicklaus Rodriguez, Nicole Sutter, Norman Jones, Pratik Sinha, Priya Prasad, Raphael Lota, Saurabh Asthana, Sharvari Bhide, Tasha Lea, Yumiko Abe-Jones

Yale School of Medicine, New Haven, CT 06510, USA:

David A. Hafler, Ruth R. Montgomery, Albert C. Shaw, Steven H. Kleinstein, Jeremy P. Gygi, Dylan Duchon, Shrikant Pawar, Anna Konstorium, Ernie Chen, Chris Cotsapas, Xiaomei Wang, Charles Dela Cruz, Akiko Iwasaki, Subhasis Mohanty, Allison Nelson, Yujiao Zhao, Shelli Farhadian, Hiromitsu Asashima, Omkar Chaudhary, Andreas Coppi, John Fournier, M. Catherine Muenker, Khadir Raddassi, Michael Rainone, William Ruff, Syim Salahuddin, Wade L. Shulz,

Pavithra Vijayakumar, Haowei Wang, Esio Wunder Jr., H. Patrick Young, Albert I. Ko, Gisela Gabernet

Yale School of Public Health, New Haven, CT 06510, USA:

Denise Esserman, Laying Guan, Anderson Brito, Jessica Rothman, Nathan D. Grubaugh, Kexin Wang, Leqi Xu

Baylor College of Medicine and the Center for Translational Research on Inflammatory Diseases, Houston, TX 77030, USA:

David B. Corry, Farrah Kheradmand, Li-Zhen Song, Ebony Nelson

Oklahoma University Health Sciences Center, Oklahoma City, OK 73104, USA:

Jordan P. Metcalf, Nelson I. Agudelo Higueta, Lauren A. Sinko, J. Leland Booth, Douglas A. Drevets, Brent R. Brown

University of Arizona, Tucson AZ 85721, USA:

Monica Kraft, Chris Bime, Jarrod Mosier, Heidi Erickson, Ron Schunk, Hiroki Kimura, Michelle Conway, Dave Francisco, Allyson Molzahn, Connie Cathleen Wilson, Ron Schunk, Trina Hughes, Bianca Sierra

University of Florida, Gainesville, FL 32611, USA:

Mark A. Atkinson, Scott C. Brakenridge, Ricardo F. Ungaro, Brittany Roth Manning, Lyle Moldawer

University of Florida, Jacksonville, FL 32218, USA:

Jordan Oberhaus, Faheem W. Guirgis

University of South Florida, Tampa FL 33620, USA:

Brittney Borresen, Matthew L. Anderson

The University of Texas at Austin, Austin, TX 78712, USA:

Lauren I. R. Ehrlich, Esther Melamed, Cole Maguire, Dennis Wylie, Justin F. Rousseau, Kerin C. Hurley, Janelle N. Geltman, Nadia Siles, Jacob E. Rogers, Pablo Guaman Tipan

ACKNOWLEDGMENTS

This research was supported by the following grants: NIH (3U01AI167892-03S2, 3U01AI167892-01S2, 5R01AI135803-03, 5U19AI118608-04, 5U19AI128910-04, 4U19AI090023-11, 4U19AI118610-06, R01AI145835-01A1S1, 5U19AI062629-17, 5U19AI057229-17, 5U19AI057229-18, 5U19AI125357-05, 5U19AI128913-03, 3U19AI077439-13, 5U54AI142766-03, 5R01AI104870-07, 3U19AI089992-09, 3U19AI128913-03, and 5T32DA018926-18); NIAID, NIH (3U19AI1289130, U19AI128913-04S1, and R01AI122220); NCATS, NIH UM1TR004528 and National Science Foundation (DMS2310836).

We thank the participants of the study for their voluntary enrollment and contribution of samples for this work. See supplement information for details on the IMPACC Network. We acknowledge the assistance of the following individuals: Tyson H. Holmes (Human Immune Monitoring Center, Stanford University), Sanya Thomas, Mitchell Cooney, Shun Rao, Sofia Vignolo, and Elena Morrocchi (all from the CDCC); Arash Naeim, Marianne Bernardo, Sarahmay Sanchez, Shannon Intluxay, Clara Magyar, Jenny Brook, Estefania Ramires-Sanchez, Megan Llamas, Claudia

Perdomo, Clara E. Magyar, and Jennifer A. Fulcher (all from the David Geffen School of Medicine at UCLA); members of the UCLA Center for Pathology Research Services and the Pathology Research Portal; M. Catherine Muenker, Dimitri Duvilaire, Maxine Kuang, William Ruff, Khadir Raddassi, Denise Shepherd, Haowei Wang, Omkar Chaudhary, Syim Salahuddin, John Fournier, Michael Rainone, and Maxine Kuang (all from the Yale School of Medicine). We thank the leadership of Boston Children's Hospital including Drs. Wendy Chung, Gary Fleisher and Kevin Churchwell for their support for the Precision Vaccines Program.

DECLARATION OF INTERESTS

The Icahn School of Medicine at Mount Sinai has filed patent applications relating to SARS-CoV-2 serological assays, NDV-based SARS-CoV-2 vaccines influenza virus vaccines and influenza virus therapeutics which lists Florian Krammer as co-inventor. Mount Sinai has spun out a company, Kantaro, to market serological tests for SARS-CoV-2 and another company, Castlevax, to develop SARS-CoV-2 vaccines. Florian Krammer is co-founder and scientific advisory board member of Castlevax. Florian Krammer has consulted for Merck, Curevac, Seqirus, GSK and Pfizer and is currently consulting for 3rd Rock Ventures, Sanofi, Gritstone and Avimex. The Krammer laboratory is also collaborating with Dynavax on influenza vaccine development and with VIR on influenza virus therapeutics development. Viviana Simon is a co-inventor on a patent filed relating to SARS-CoV-2 serological assays (the "Serology Assays"). Ofer Levy is a named inventor on patents held by Boston Children's Hospital relating to vaccine adjuvants and human in vitro platforms that model vaccine action. His laboratory has received research support from GlaxoSmithKline (GSK) and Pfizer, and he is a co-founder of and advisor to Ovax, Inc that develops opioid vaccines. Charles Cairns serves as a consultant to bioMerieux and is funded by a grant from Bill & Melinda Gates Foundation. James A Overton is a consultant at Knocean Inc. Jessica Lasky-Su serves as a scientific advisor of Precion Inc. Scott R. Hutton, Greg Michelloti and Kari Wong are employees of Metabolon Inc. Vicki Seyfer- Margolis is a current employee of MyOwnMed. Nadine Rouphael reports grants or contracts with Merck, Sanofi, Pfizer, Vaccine Company and Immorna, and has participated on data safety monitoring boards for Moderna, Sanofi, Seqirus, Pfizer, EMMES, ICON, BARDA, and CyanVan, Imunon Micron. N.R. has also received support for meetings/travel from Sanofi and Moderna and honoraria from Virology Education and Krog consulting. Chris Cotsapas is a current employee of Vesalius Therapeutics. Adeeb Rahman is a current employee of Immunai Inc. Steven Kleinstein is a consultant related to ImmPort data repository for Peraton. Nathan Grabaugh is a consultant for Tempus Labs and the National Basketball Association. Akiko Iwasaki is a consultant for 4BIO, Blue Willow Biologics, Revelar Biotherapeutics, RIGImmune, Xanadu Bio, Paratus Sciences. Monika Kraft receives research funds paid to her institution from NIH, ALA; Sanofi, Astra-Zeneca for work in asthma, serves as a consultant for Astra-Zeneca, Sanofi, Chiesi, GSK for severe asthma; is a co-founder and CMO for RaeSedo, Inc, a company created to develop peptidomimetics for treatment of inflammatory lung disease. Esther Melamed received research funding from Babson Diagnostics and honorarium from Multiple Sclerosis Association of America and has served on the advisory boards of Genentech, Horizon, Teva, and Viela Bio. Carolyn Calfee receives research funding from NIH, FDA, DOD, Roche-Genentech and Quantum Leap Healthcare Collaborative as well as consulting services for Janssen, Vasomune, Gen1e Life Sciences, NGMBio, and Cellenkos. Wade Schulz was an investigator for a research agreement, through Yale University, from the

Shenzhen Center for Health Information for work to advance intelligent disease prevention and health promotion; collaborates with the National Center for Cardiovascular Diseases in Beijing; is a technical consultant to Hugo Health, a personal health information platform; cofounder of Refactor Health, an AI-augmented data management platform for health care; and has received grants from Merck and Regeneron Pharmaceutical for research related to COVID-19. Grace A McComsey received research grants from Rehdhill, Cognivue, Pfizer, and Genentech, and served as a research consultant for Gilead, Merck, Viiv/GSK, and Janssen. Linda N. Geng received research funding paid to her institution from Pfizer, Inc. Catherine Hough receives support through her institution from the NIH and CDC. David Hafler has received research funding from Bristol-Myers Squibb, Novartis, Sanofi, and Genentech. He has been a consultant for Bayer Pharmaceuticals, Repertoire Inc, Bristol Myers Squibb, Compass Therapeutics, EMD Serono, Genentech, Novartis Pharmaceuticals, and Sanofi Genzyme.

SUPPLEMENTAL INFORMATION TITLES AND LEGENDS

1. Supplemental Figures 1-10 and Legends
2. Experimental Model, Study Participant Details and Method Details

AUTHOR CONTRIBUTIONS

Conceptualization: IMPACC Network, PMB, SK, LIRE, LG

Data Curation: AH, CS, BP, JD-A

Formal analysis: GG, JM, JPG, JFM, AH, CS, TC, NDJ

Funding Acquisition: IMPACC Network

Methodology: GG, JM, JPG, JFM, SHK, LG

Software: GG, JM, JPG, JFM, SHK, LG

Resources: IMPACC Network

Supervision: SF, MCA, OL, KKS, THH, RRM, JD-A, SHK, LG, LIRE

All authors wrote, edited, and reviewed the manuscript.

REFERENCES

1. Ely, E.W., Brown, L.M., and Fineberg, H.V. (2024). Long Covid Defined. *New England Journal of Medicine* 391, 1746–1753. <https://doi.org/10.1056/NEJMsb2408466>.
2. National Academies of Sciences, E., and Medicine (2024). *A Long COVID Definition: A Chronic, Systemic Disease State with Profound Consequences* (The National Academies Press) <https://doi.org/10.17226/27768>.
3. Groff, D., Sun, A., Ssentongo, A.E., Ba, D.M., Parsons, N., Poudel, G.R., Lekoubou, A., Oh, J.S., Ericson, J.E., Ssentongo, P., et al. (2021). Short-term and Long-term Rates of Postacute Sequelae of SARS-CoV-2 Infection: A Systematic Review. *JAMA Netw Open* 4, e2128568. <https://doi.org/10.1001/jamanetworkopen.2021.28568>.
4. Davis, H.E., Assaf, G.S., McCorkell, L., Wei, H., Low, R.J., Re'em, Y., Redfield, S., Austin, J.P., and Akrami, A. (2021). Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. *eClinicalMedicine* 38. <https://doi.org/10.1016/j.eclinm.2021.101019>.

5. Melamed, E., Rydberg, L., Ambrose, A.F., Bhavaraju-Sanka, R., Fine, J.S., Fleming, T.K., Herman, E., Phipps Johnson, J.L., Kucera, J.R., Longo, M., et al. (2023). Multidisciplinary collaborative consensus guidance statement on the assessment and treatment of neurologic sequelae in patients with post-acute sequelae of SARS-CoV-2 infection (PASC). *PM R* 15, 640–662. <https://doi.org/10.1002/pmrj.12976>.
6. Greenhalgh, T., Sivan, M., Perlowski, A., and Nikolich, J.Ž. (2024). Long COVID: a clinical update. *The Lancet* 404, 707–724. [https://doi.org/10.1016/S0140-6736\(24\)01136-X](https://doi.org/10.1016/S0140-6736(24)01136-X).
7. Whiteson, J.H., Azola, A., Barry, J.T., Bartels, M.N., Blitshteyn, S., Fleming, T.K., McCauley, M.D., Neal, J.D., Pillarisetti, J., Sampsel, S., et al. (2022). Multi-disciplinary collaborative consensus guidance statement on the assessment and treatment of cardiovascular complications in patients with post-acute sequelae of SARS-CoV-2 infection (PASC). *PM&R* 14, 855–878. <https://doi.org/10.1002/pmrj.12859>.
8. Blitshteyn, S., Whiteson, J.H., Abramoff, B., Azola, A., Bartels, M.N., Bhavaraju-Sanka, R., Chung, T., Fleming, T.K., Henning, E., Miglis, M.G., et al. (2022). Multi-disciplinary collaborative consensus guidance statement on the assessment and treatment of autonomic dysfunction in patients with post-acute sequelae of SARS-CoV-2 infection (PASC). *PM&R* 14, 1270–1291. <https://doi.org/10.1002/pmrj.12894>.
9. Cheng, A.L., Anderson, J., Didehbani, N., Fine, J.S., Fleming, T.K., Karnik, R., Longo, M., Ng, R., Re'em, Y., Sampsel, S., et al. (2023). Multi-disciplinary collaborative consensus guidance statement on the assessment and treatment of mental health symptoms in patients with post-acute sequelae of SARS-CoV-2 infection (PASC). *PM&R* 15, 1588–1604. <https://doi.org/10.1002/pmrj.13085>.
10. King, L.R. (2024). Gastrointestinal manifestations of long COVID. *Life Sci* 357, 123100. <https://doi.org/10.1016/j.lfs.2024.123100>.
11. Bandara, T., Deshmukh, H.A., Abdalla, M., and Sathyapalan, T. (2023). Metabolic and Endocrine Complications of Long-COVID-19: A Review. *Exp Clin Endocrinol Diabetes* 131, 367–374. <https://doi.org/10.1055/a-2063-8697>.
12. Kobusiak-Prokopowicz, M., Fułek, K., Fułek, M., Kaaz, K., Mysiak, A., Kurpas, D., Beszlej, J.A., Brzecka, A., and Leszek, J. (2022). Cardiovascular, Pulmonary, and Neuropsychiatric Short- and Long-Term Complications of COVID-19. *Cells* 11, 3882. <https://doi.org/10.3390/cells11233882>.
13. Proal, A.D., and VanElzakker, M.B. (2021). Long COVID or Post-acute Sequelae of COVID-19 (PASC): An Overview of Biological Factors That May Contribute to Persistent Symptoms. *Front Microbiol* 12, 698169. <https://doi.org/10.3389/fmicb.2021.698169>.
14. Davis, H.E., McCorkell, L., Vogel, J.M., and Topol, E.J. (2023). Long COVID: major findings, mechanisms and recommendations. *Nat Rev Microbiol* 21, 133–146. <https://doi.org/10.1038/s41579-022-00846-2>.
15. Wang, Y., Alcalde-Herraziz, M., Güell, K.L., Chen, L., Mateu, L., Li, C., Ali, R., Wareham, N., Paredes, R., Prieto-Alhambra, D., et al. (2025). Refinement of post-COVID condition core symptoms, subtypes, determinants, and health impacts: a cohort study integrating real-world

data and patient-reported outcomes. *eBioMedicine* 111.
<https://doi.org/10.1016/j.ebiom.2024.105493>.

16. Sudre, C.H., Murray, B., Varsavsky, T., Graham, M.S., Penfold, R.S., Bowyer, R.C., Pujol, J.C., Klaser, K., Antonelli, M., Canas, L.S., et al. (2021). Attributes and predictors of long COVID. *Nat Med* 27, 626–631. <https://doi.org/10.1038/s41591-021-01292-y>.
17. Subramanian, A., Nirantharakumar, K., Hughes, S., Myles, P., Williams, T., Gokhale, K.M., Taverner, T., Chandan, J.S., Brown, K., Simms-Williams, N., et al. (2022). Symptoms and risk factors for long COVID in non-hospitalized adults. *Nat Med* 28, 1706–1714. <https://doi.org/10.1038/s41591-022-01909-w>.
18. Ozonoff, A., Jayavelu, N.D., Liu, S., Melamed, E., Milliren, C.E., Qi, J., Geng, L.N., McComsey, G.A., Cairns, C.B., Baden, L.R., et al. (2024). Features of acute COVID-19 associated with post-acute sequelae of SARS-CoV-2 phenotypes: results from the IMPACC study. *Nat Commun* 15, 216. <https://doi.org/10.1038/s41467-023-44090-5>.
19. Lui, D.T.W., Lee, C.H., Chow, W.S., Lee, A.C.H., Tam, A.R., Pang, P., Ho, T.Y., Fong, C.H.Y., Law, C.Y., Leung, E.K.H., et al. (2021). Long COVID in Patients With Mild to Moderate Disease: Do Thyroid Function and Autoimmunity Play a Role? *Endocr Pract* 27, 894–902. <https://doi.org/10.1016/j.eprac.2021.06.016>.
20. Tleyjeh, I.M., Saddik, B., Ramakrishnan, R.K., AlSwaidan, N., AlAnazi, A., Alhazmi, D., Aloufi, A., AlSumait, F., Berbari, E.F., and Halwani, R. (2022). Long term predictors of breathlessness, exercise intolerance, chronic fatigue and well-being in hospitalized patients with COVID-19: A cohort study with 4 months median follow-up. *Journal of Infection and Public Health* 15, 21–28. <https://doi.org/10.1016/j.jiph.2021.11.016>.
21. Hamlin, R.E., Pienkos, S.M., Chan, L., Stabile, M.A., Pinedo, K., Rao, M., Grant, P., Bonilla, H., Holubar, M., Singh, U., et al. (2024). Sex differences and immune correlates of Long Covid development, symptom persistence, and resolution. *Science Translational Medicine* 16, eadr1032. <https://doi.org/10.1126/scitranslmed.adr1032>.
22. Becker, C., Beck, K., Zumbrunn, S., Memma, V., Herzog, N., Bissmann, B., Gross, S., Loretz, N., Mueller, J., Amacher, S.A., et al. (2021). Long COVID 1 year after hospitalisation for COVID-19: a prospective bicentric cohort study. *Swiss Med Wkly* 151, w30091. <https://doi.org/10.4414/smw.2021.w30091>.
23. Augustin, M., Schommers, P., Stecher, M., Dewald, F., Gieselmann, L., Gruell, H., Horn, C., Vanshylla, K., Cristanziano, V.D., Osebold, L., et al. (2021). Post-COVID syndrome in non-hospitalised patients with COVID-19: a longitudinal prospective cohort study. *Lancet Reg Health Eur* 6, 100122. <https://doi.org/10.1016/j.lanepe.2021.100122>.
24. García-Abellán, J., Padilla, S., Fernández-González, M., García, J.A., Agulló, V., Andreo, M., Ruiz, S., Galiana, A., Gutiérrez, F., and Masiá, M. (2021). Antibody Response to SARS-CoV-2 is Associated with Long-term Clinical Outcome in Patients with COVID-19: a Longitudinal Study. *J Clin Immunol* 41, 1490–1501. <https://doi.org/10.1007/s10875-021-01083-7>.
25. Swank, Z., Senussi, Y., Manickas-Hill, Z., Yu, X.G., Li, J.Z., Alter, G., and Walt, D.R. (2023). Persistent Circulating Severe Acute Respiratory Syndrome Coronavirus 2 Spike Is Associated

- With Post-acute Coronavirus Disease 2019 Sequelae. *Clinical Infectious Diseases* 76, e487–e490. <https://doi.org/10.1093/cid/ciac722>.
26. Klein, J., Wood, J., Jaycox, J.R., Dhodapkar, R.M., Lu, P., Gehlhausen, J.R., Tabachnikova, A., Greene, K., Tabacof, L., Malik, A.A., et al. (2023). Distinguishing features of long COVID identified through immune profiling | *Nature*. *Nature* 623, 139–148. <https://doi.org/10.1038/s41586-023-06651-y>.
 27. Talla, A., Vasaikar, S.V., Szeto, G.L., Lemos, M.P., Czartoski, J.L., MacMillan, H., Moodie, Z., Cohen, K.W., Fleming, L.B., Thomson, Z., et al. (2023). Persistent serum protein signatures define an inflammatory subcategory of long COVID. *Nat Commun* 14, 3417. <https://doi.org/10.1038/s41467-023-38682-4>.
 28. Liew, F., Efstathiou, C., Fontanella, S., Richardson, M., Saunders, R., Swieboda, D., Sidhu, J.K., Ascough, S., Moore, S.C., Mohamed, N., et al. (2024). Large-scale phenotyping of patients with long COVID post-hospitalization reveals mechanistic subtypes of disease. *Nat Immunol* 25, 607–621. <https://doi.org/10.1038/s41590-024-01778-0>.
 29. Hanson, A.L., Mulè, M.P., Ruffieux, H., Mescia, F., Bergamaschi, L., Pelly, V.S., Turner, L., Kotagiri, P., Göttgens, B., Hess, C., et al. (2024). Iron dysregulation and inflammatory stress erythropoiesis associates with long-term outcome of COVID-19. *Nat Immunol* 25, 471–482. <https://doi.org/10.1038/s41590-024-01754-8>.
 30. Peluso, M.J., Deveau, T.-M., Munter, S.E., Ryder, D., Buck, A., Beck-Engeser, G., Chan, F., Lu, S., Goldberg, S.A., Hoh, R., et al. (2023). Chronic viral coinfections differentially affect the likelihood of developing long COVID. *J Clin Invest* 133, e163669. <https://doi.org/10.1172/JCI163669>.
 31. Silva, J., Takahashi, T., Wood, J., Lu, P., Tabachnikova, A., Gehlhausen, J.R., Greene, K., Bhattacharjee, B., Monteiro, V.S., Lucas, C., et al. (2024). Sex differences in symptomatology and immune profiles of Long COVID. *medRxiv*, 2024.02.29.24303568. <https://doi.org/10.1101/2024.02.29.24303568>.
 32. Phetsouphanh, C., Darley, D.R., Wilson, D.B., Howe, A., Munier, C.M.L., Patel, S.K., Juno, J.A., Burrell, L.M., Kent, S.J., Dore, G.J., et al. (2022). Immunological dysfunction persists for 8 months following initial mild-to-moderate SARS-CoV-2 infection. *Nat Immunol* 23, 210–216. <https://doi.org/10.1038/s41590-021-01113-x>.
 33. Yin, K., Peluso, M.J., Luo, X., Thomas, R., Shin, M.-G., Neidleman, J., Andrew, A., Young, K.C., Ma, T., Hoh, R., et al. (2024). Long COVID manifests with T cell dysregulation, inflammation and an uncoordinated adaptive immune response to SARS-CoV-2 | *Nature Immunology*. *Nature Immunology*. <https://doi.org/10.1038/s41590-023-01724-6>.
 34. Cervia-Hasler, C., Brüningk, S.C., Hoch, T., Fan, B., Muzio, G., Thompson, R.C., Ceglarek, L., Meledin, R., Westermann, P., Emmenegger, M., et al. (2024). Persistent complement dysregulation with signs of thromboinflammation in active Long Covid. *Science* 383, eadg7942. <https://doi.org/10.1126/science.adg7942>.
 35. Wallukat, G., Hohberger, B., Wenzel, K., Fürst, J., Schulze-Rothe, S., Wallukat, A., Hönicke, A.-S., and Müller, J. (2021). Functional autoantibodies against G-protein coupled receptors in

- patients with persistent Long-COVID-19 symptoms. *J Transl Autoimmun* 4, 100100. <https://doi.org/10.1016/j.jtauto.2021.100100>.
36. Su, Y., Yuan, D., Chen, D.G., Ng, R.H., Wang, K., Choi, J., Li, S., Hong, S., Zhang, R., Xie, J., et al. (2022). Multiple early factors anticipate post-acute COVID-19 sequelae. *Cell* 185, 881-895.e20. <https://doi.org/10.1016/j.cell.2022.01.014>.
 37. IMPACC Manuscript Writing Team and IMPACC Network Steering Committee (2021). Immunophenotyping assessment in a COVID-19 cohort (IMPACC): A prospective longitudinal study. *Sci Immunol* 6, eabf3733. <https://doi.org/10.1126/sciimmunol.abf3733>.
 38. Diray-Arce, J., Fourati, S., Doni Jayavelu, N., Patel, R., Maguire, C., Chang, A.C., Dandekar, R., Qi, J., Lee, B.H., van Zalm, P., et al. (2023). Multi-omic longitudinal study reveals immune correlates of clinical course among hospitalized COVID-19 patients. *Cell Reports Medicine* 4, 101079. <https://doi.org/10.1016/j.xcrm.2023.101079>.
 39. Gygi, J.P., Maguire, C., Patel, R.K., Shinde, P., Konstorum, A., Shannon, C.P., Xu, L., Hoch, A., Jayavelu, N.D., Haddad, E.K., et al. (2024). Integrated longitudinal multiomics study identifies immune programs associated with acute COVID-19 severity and mortality. *J Clin Invest* 134. <https://doi.org/10.1172/JCI176640>.
 40. Wang, K., Nie, Y., Maguire, C., Syphurs, C., Sheen, H., Karoly, M., Lapp, L., Gygi, J.P., Jayavelu, N.D., Patel, R.K., et al. (2024). Unraveling SARS-CoV-2 Host-Response Heterogeneity through Longitudinal Molecular Subtyping. *bioRxiv*, 2024.11.22.624784. <https://doi.org/10.1101/2024.11.22.624784>.
 41. Gygi, J.P., Konstorum, A., Pawar, S., Aron, E., Kleinstein, S.H., and Guan, L. (2024). A supervised Bayesian factor model for the identification of multi-omics signatures. *Bioinformatics* 40, btae202. <https://doi.org/10.1093/bioinformatics/btae202>.
 42. Argelaguet, R., Velten, B., Arnol, D., Dietrich, S., Zenz, T., Marioni, J.C., Buettner, F., Huber, W., and Stegle, O. (2018). Multi-Omics Factor Analysis—a framework for unsupervised integration of multi-omics data sets. *Molecular Systems Biology* 14, e8124. <https://doi.org/10.15252/msb.20178124>.
 43. Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* 1, 417–425. <https://doi.org/10.1016/j.cels.2015.12.004>.
 44. Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28, 27–30. <https://doi.org/10.1093/nar/28.1.27>.
 45. DeHaven, C.D., Evans, A.M., Dai, H., and Lawton, K.A. (2010). Organization of GC/MS and LC/MS metabolomics data into chemical libraries. *Journal of Cheminformatics* 2, 9. <https://doi.org/10.1186/1758-2946-2-9>.
 46. Evans, A.M., O'Donovan, C., Playdon, M., Beecher, C., Beger, R.D., Bowden, J.A., Broadhurst, D., Clish, C.B., Dasari, S., Dunn, W.B., et al. (2020). Dissemination and analysis of the quality assurance (QA) and quality control (QC) practices of LC-MS based untargeted metabolomics practitioners. *Metabolomics* 16, 113. <https://doi.org/10.1007/s11306-020-01728-5>.

47. Evans, A.M., DeHaven, C.D., Barrett, T., Mitchell, M., and Milgram, E. (2009). Integrated, Nontargeted Ultrahigh Performance Liquid Chromatography/Electrospray Ionization Tandem Mass Spectrometry Platform for the Identification and Relative Quantification of the Small-Molecule Complement of Biological Systems. *Anal. Chem.* *81*, 6656–6667. <https://doi.org/10.1021/ac901536h>.
48. Wang, L., Wu, Q., Li, Z., Sun, S., Yuan, J., Li, J., Zhang, Y., Yu, D., Wang, C., and Sun, S. (2019). Delta/notch-like epidermal growth factor-related receptor promotes stemness to facilitate breast cancer progression. *Cell Signal* *63*, 109389. <https://doi.org/10.1016/j.cellsig.2019.109389>.
49. Wang, Z., Li, Z., Wu, Q., Li, C., Li, J., Zhang, Y., Wang, C., Sun, S., and Sun, S. (2020). DNER promotes epithelial–mesenchymal transition and prevents chemosensitivity through the Wnt/ β -catenin pathway in breast cancer. *Cell Death Dis* *11*, 1–16. <https://doi.org/10.1038/s41419-020-02903-1>.
50. Zhang, D., Xu, C., Zhang, J., Zeng, R., Qi, Q., Xu, J., Pan, Y., Liu, X., Shi, S., Zhang, J., et al. (2023). Plasma TNFRSF11B as a New Predictive Inflammatory Marker of Sepsis-ARDS with Endothelial Dysfunction. *J Proteome Res* *22*, 3640–3651. <https://doi.org/10.1021/acs.jproteome.3c00576>.
51. Seo, D.H., Corr, M., Patel, S., Lui, L.-Y., Cauley, J.A., Evans, D., Mau, T., and Lane, N.E. (2024). Chemokine CXCL9, a marker of inflammaging, is associated with changes of muscle strength and mortality in older men. *Osteoporos Int* *35*, 1789–1796. <https://doi.org/10.1007/s00198-024-07160-y>.
52. Sayed, N., Huang, Y., Nguyen, K., Krejciova-Rajaniemi, Z., Grawe, A.P., Gao, T., Tibshirani, R., Hastie, T., Alpert, A., Cui, L., et al. (2021). An inflammatory aging clock (iAge) based on deep learning tracks multimorbidity, immunosenescence, frailty and cardiovascular aging. *Nat Aging* *1*, 598–615. <https://doi.org/10.1038/s43587-021-00082-y>.
53. Czaya, B., and Faul, C. (2019). The Role of Fibroblast Growth Factor 23 in Inflammation and Anemia. *Int J Mol Sci* *20*, 4195. <https://doi.org/10.3390/ijms20174195>.
54. Kwon, B. (2012). Regulation of Inflammation by Bidirectional Signaling through CD137 and Its Ligand. *Immune Netw* *12*, 176–180. <https://doi.org/10.4110/in.2012.12.5.176>.
55. Geng, T., Yan, Y., Zhang, Y., Xu, L., Zang, G., and Yan, J.C. (2020). CD137 Signaling Promotes Endothelial Apoptosis by Inhibiting Nrf2 Pathway, and Upregulating NF- κ B Pathway. *Mediators Inflamm* *2020*, 4321912. <https://doi.org/10.1155/2020/4321912>.
56. Lee, H.S., and Kim, W.J. (2022). The Role of Matrix Metalloproteinase in Inflammation with a Focus on Infectious Diseases. *Int J Mol Sci* *23*, 10546. <https://doi.org/10.3390/ijms231810546>.
57. McMahan, R.S., Birkland, T.P., Smigiel, K.S., Vandivort, T.C., Rohani, M.G., Manicone, A.M., McGuire, J.K., Gharib, S.A., and Parks, W.C. (2016). Stromelysin-2 (MMP10) Moderates Inflammation by Controlling Macrophage Activation. *J Immunol* *197*, 899–909. <https://doi.org/10.4049/jimmunol.1600502>.
58. Hamilton, J.A. (2008). Colony-stimulating factors in inflammation and autoimmunity. *Nat Rev Immunol* *8*, 533–544. <https://doi.org/10.1038/nri2356>.

59. Voloudakis, G., Vicari, J.M., Venkatesh, S., Hoffman, G.E., Dobrindt, K., Zhang, W., Beckmann, N.D., Higgins, C.A., Argyriou, S., Jiang, S., et al. (2022). A translational genomics approach identifies IL10RB as the top candidate gene target for COVID-19 susceptibility. *npj Genom. Med.* 7, 1–15. <https://doi.org/10.1038/s41525-022-00324-x>.
60. Dritsoula, A., Camilli, C., Moss, S.E., and Greenwood, J. (2024). The disruptive role of LRG1 on the vasculature and perivascular microenvironment. *Frontiers in Cardiovascular Medicine* 11, 1386177. <https://doi.org/10.3389/fcvm.2024.1386177>.
61. Dritsoula, A., Dowsett, L., Pilotti, C., O'Connor, M.N., Moss, S.E., and Greenwood, J. (2022). Angiopathic activity of LRG1 is induced by the IL-6/STAT3 pathway. *Sci Rep* 12, 4867. <https://doi.org/10.1038/s41598-022-08516-2>.
62. Camilli, C., Hoeh, A.E., De Rossi, G., Moss, S.E., and Greenwood, J. (2022). LRG1: an emerging player in disease pathogenesis. *Journal of Biomedical Science* 29, 6. <https://doi.org/10.1186/s12929-022-00790-6>.
63. Krishnamoorthy, N.K., Kalyan, M., Hediya, T.A., Anand, N., Kendaganna, P.H., Pendyala, G., Yelamanchili, S.V., Yang, J., Chidambaram, S.B., Sakharkar, M.K., et al. (2024). Role of the Gut Bacteria-Derived Metabolite Phenylacetylglutamine in Health and Diseases. *ACS Omega* 9, 3164–3172. <https://doi.org/10.1021/acsomega.3c08184>.
64. Charman, M., Colbourne, T.R., Pietrangelo, A., Kreplak, L., and Ridgway, N.D. (2014). Oxysterol-binding Protein (OSBP)-related Protein 4 (ORP4) Is Essential for Cell Proliferation and Survival*. *Journal of Biological Chemistry* 289, 15705–15717. <https://doi.org/10.1074/jbc.M114.571216>.
65. Pace, S., and Werz, O. (2020). Impact of Androgens on Inflammation-Related Lipid Mediator Biosynthesis in Innate Immune Cells. *Front Immunol* 11, 1356. <https://doi.org/10.3389/fimmu.2020.01356>.
66. Murugan, S., Jakka, P., Namani, S., Mujumdar, V., and Radhakrishnan, G. (2019). The neurosteroid pregnenolone promotes degradation of key proteins in the innate immune signaling to suppress inflammation. *J Biol Chem* 294, 4596–4607. <https://doi.org/10.1074/jbc.RA118.005543>.
67. Vallée, M. (2016). Neurosteroids and potential therapeutics: Focus on pregnenolone. *J Steroid Biochem Mol Biol* 160, 78–87. <https://doi.org/10.1016/j.jsbmb.2015.09.030>.
68. Leeansyah, E., Ganesh, A., Quigley, M.F., Sönnnerborg, A., Andersson, J., Hunt, P.W., Somsouk, M., Deeks, S.G., Martin, J.N., Moll, M., et al. (2013). Activation, exhaustion, and persistent decline of the antimicrobial MR1-restricted MAIT-cell population in chronic HIV-1 infection. *Blood* 121, 1124–1135. <https://doi.org/10.1182/blood-2012-07-445429>.
69. Hengst, J., Strunz, B., Deterding, K., Ljunggren, H.-G., Leeansyah, E., Manns, M.P., Cornberg, M., Sandberg, J.K., Wedemeyer, H., and Björkström, N.K. (2016). Nonreversible MAIT cell-dysfunction in chronic hepatitis C virus infection despite successful interferon-free therapy. *European Journal of Immunology* 46, 2204–2210. <https://doi.org/10.1002/eji.201646447>.
70. Ozonoff, A., Schaenman, J., Jayavelu, N.D., Milliren, C.E., Calfee, C.S., Cairns, C.B., Kraft, M., Baden, L.R., Shaw, A.C., Krammer, F., et al. (2022). Phenotypes of disease severity in a

- cohort of hospitalized COVID-19 patients: Results from the IMPACC study. *EBioMedicine* 83, 104208. <https://doi.org/10.1016/j.ebiom.2022.104208>.
71. Takahashi, T., Ellingson, M.K., Wong, P., Israelow, B., Lucas, C., Klein, J., Silva, J., Mao, T., Oh, J.E., Tokuyama, M., et al. (2020). Sex differences in immune responses that underlie COVID-19 disease outcomes. *Nature* 588, 315–320. <https://doi.org/10.1038/s41586-020-2700-3>.
 72. Fourati, S., Tomalin, L.E., Mulè, M.P., Chawla, D.G., Gerritsen, B., Rychkov, D., Henrich, E., Miller, H.E.R., Hagan, T., Diray-Arce, J., et al. (2022). Pan-vaccine analysis reveals innate immune endotypes predictive of antibody responses to vaccination. *Nat Immunol* 23, 1777–1787. <https://doi.org/10.1038/s41590-022-01329-5>.
 73. Hagan, T., Gerritsen, B., Tomalin, L.E., Fourati, S., Mulè, M.P., Chawla, D.G., Rychkov, D., Henrich, E., Miller, H.E.R., Diray-Arce, J., et al. (2022). Transcriptional atlas of the human immune response to 13 vaccines reveals a common predictor of vaccine-induced antibody responses. *Nat Immunol* 23, 1788–1798. <https://doi.org/10.1038/s41590-022-01328-6>.
 74. Germain, A., Barupal, D.K., Levine, S.M., and Hanson, M.R. (2020). Comprehensive Circulatory Metabolomics in ME/CFS Reveals Disrupted Metabolism of Acyl Lipids and Steroids. *Metabolites* 10, 34. <https://doi.org/10.3390/metabo10010034>.
 75. Dehlia, A., and Guthridge, M.A. (2024). The persistence of myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) after SARS-CoV-2 infection: A systematic review and meta-analysis. *Journal of Infection* 89, 106297. <https://doi.org/10.1016/j.jinf.2024.106297>.
 76. Vitku, J., Hill, M., Kolatorova, L., Kubala Havrdova, E., and Kancheva, R. (2022). Steroid Sulfation in Neurodegenerative Diseases. *Front. Mol. Biosci.* 9. <https://doi.org/10.3389/fmolb.2022.839887>.
 77. Stevenson, R., Bishop, D.G., and Rodseth, R.N. (2024). A review of the role of testosterone in the care of the critically ill patient. *South Afr J Crit Care* 40, e1303. <https://doi.org/10.7196/SAJCC.2024.v40i1.1303>.
 78. Marques, O., and Muckenthaler, M.U. (2024). Alterations of iron homeostasis as a potential druggable driver of long COVID. *Nat Immunol* 25, 387–389. <https://doi.org/10.1038/s41590-024-01759-3>.
 79. Marques, O., Weiss, G., and Muckenthaler, M.U. (2022). The role of iron in chronic inflammatory diseases: from mechanisms to treatment options in anemia of inflammation. *Blood* 140, 2011–2023. <https://doi.org/10.1182/blood.2021013472>.
 80. Domingo, J.C., Cordobilla, B., Ferrer, R., Giralt, M., Alegre-Martín, J., and Castro-Marrero, J. (2021). Are Circulating Fibroblast Growth Factor 21 and N-Terminal Prohormone of Brain Natriuretic Peptide Promising Novel Biomarkers in Myalgic Encephalomyelitis/Chronic Fatigue Syndrome? *Antioxid Redox Signal* 34, 1420–1427. <https://doi.org/10.1089/ars.2020.8230>.
 81. Taquet, M., Dercon, Q., Luciano, S., Geddes, J.R., Husain, M., and Harrison, P.J. (2021). Incidence, co-occurrence, and evolution of long-COVID features: A 6-month retrospective

cohort study of 273,618 survivors of COVID-19. *PLOS Medicine* 18, e1003773. <https://doi.org/10.1371/journal.pmed.1003773>.

82. Maguire, C., Chen, J., Roupael, N., Pickering, H., Phan, H.V., Glascock, A., Chu, V., Dandekar, R., Corry, D., Kheradmand, F., et al. (2024). Chronic Viral Reactivation and Associated Host Immune Response and Clinical Outcomes in Acute COVID-19 and Post-Acute Sequelae of COVID-19. Preprint at bioRxiv, <https://doi.org/10.1101/2024.11.14.622799> <https://doi.org/10.1101/2024.11.14.622799>.