# Supplemental Information

Large-Scale Multi-omic Analysis

of COVID-19 Severity

Katherine A. Overmyer, Evgenia Shishkova, Ian J. Miller, Joseph Balnis, Matthew N. Bernstein, Trenton M. Peters-Clarke, Jesse G. Meyer, Qiuwen Quan, Laura K. Muehlbauer, Edna A. Trujillo, Yuchen He, Amit Chopra, Hau C. Chieng, Anupama Tiwari, Marc A. Judson, Brett Paulson, Dain R. Brademan, Yunyun Zhu, Lia R. Serrano, Vanessa Linke, Lisa A. Drake, Alejandro P. Adam, Bradford S. Schwartz, Harold A. Singer, Scott Swanson, Deane F. Mosher, Ron Stewart, Joshua J. Coon, and Ariel Jaitovich

Large-scale Multi-omic Analysis of COVID-19 Severity
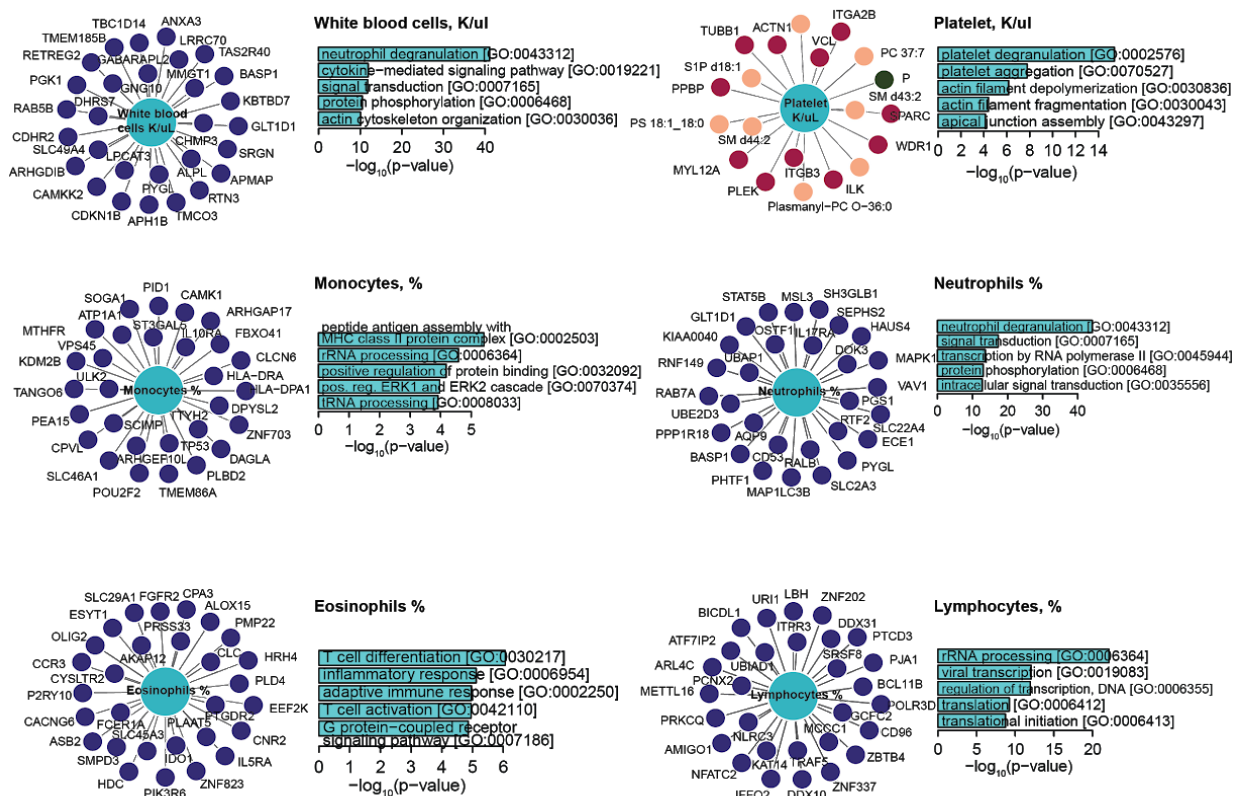

Supplementary Figures
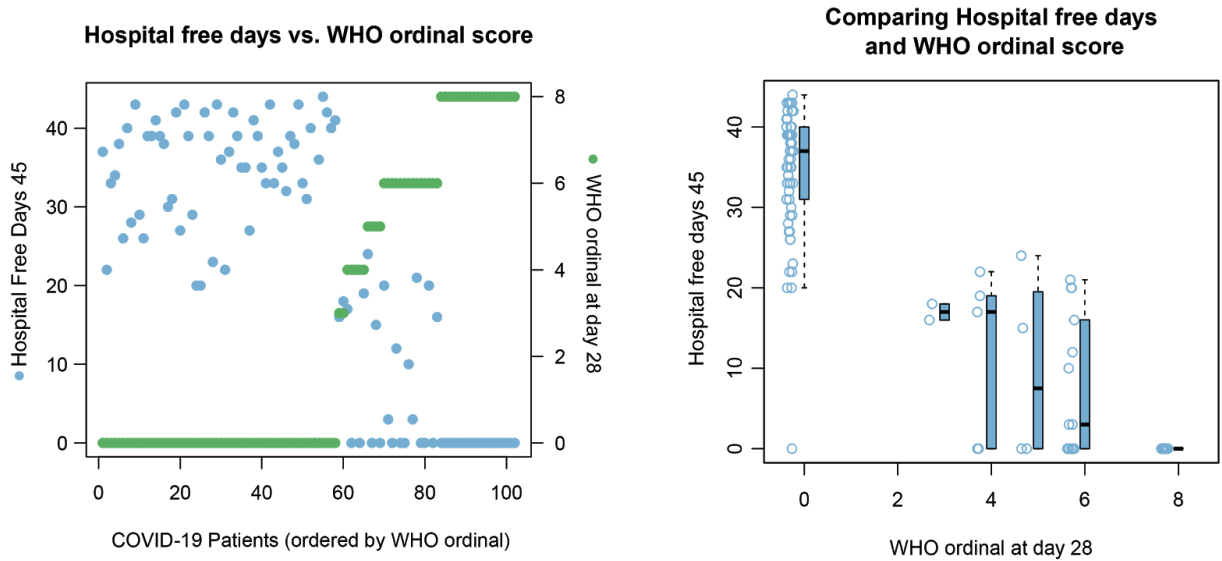
Supplementary Figure 1. Comparing the WHO ordinal score at 28 days to hospital free days at 45 days, related to Figure 1. Both the WHO ordinal score (green points, scale 0-8) and the number of hospital free days at 45 days (blue points, 0-45 days) represent outcome metrics. Hospital free days at 45 days was used for the main analysis while WHO ordinal score was provided as an alternative metric.
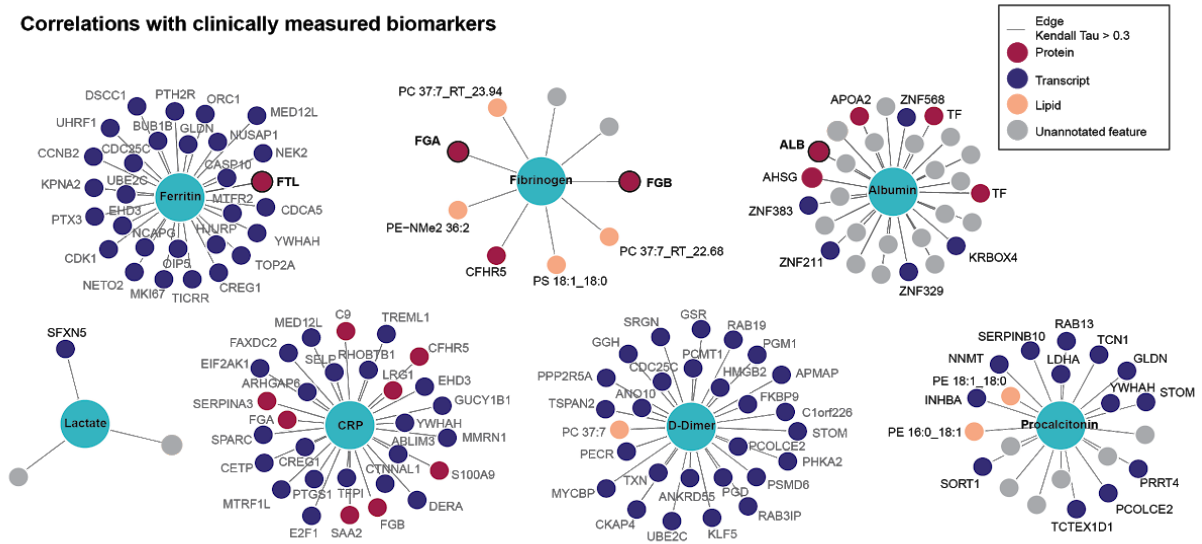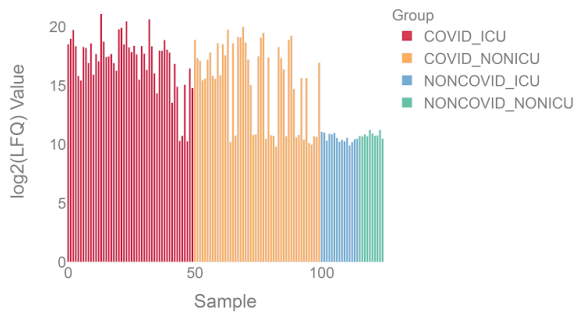
Supplementary Figure 2. Correlations between clinical metrics and biomolecules, related to Table 1. Kendall Tau coefficients were calculated between clinically measured features (light blue circles) and biomolecules measured by mass spectrometry and RNAseq. Top 30 positively correlated features, minimum kendall tau > 0.3, are shown as connected nodes in the networks above. For correlations to blood cell counts/percentages, all features with Kendall Tau > 0.3 were input for enrichment analysis using GO-term biological processes.
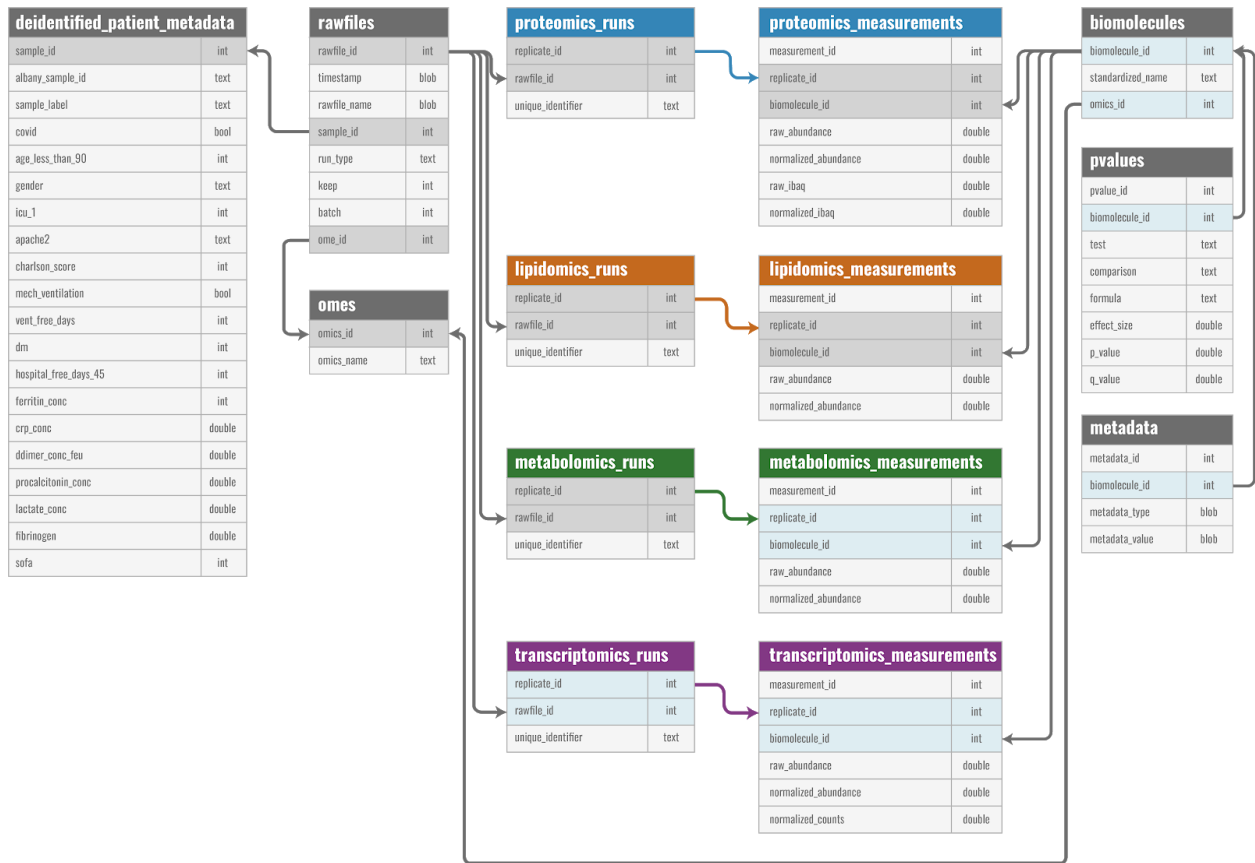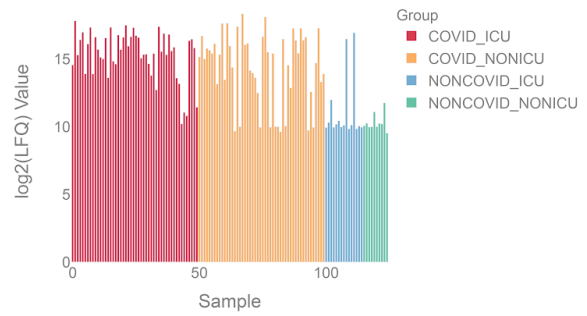
Supplementary Figure 3. SQLite Database Schema; related to Figure 1. Experimental data from both the high-throughput sequencing of patient leukocytes and the mass spectrometry analysis of patient plasma is compiled into a SQLite database. Deidentified clinical metrics are linked to profiled transcripts, proteins, lipids, and metabolites. Additionally, results from precalculated statistical analyses are stored to quickly facilitate data exploration through the manuscript's companion resource `covid-omics.app`.
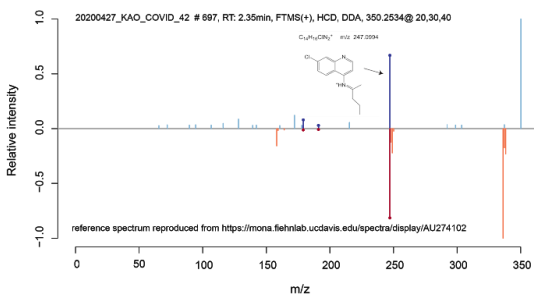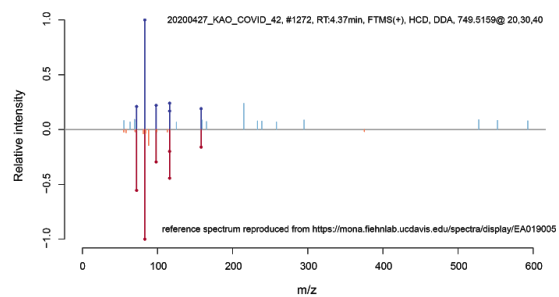
a [L] Unknown_mz_350.16278_+_RT_2.553

b [L] Unknown_mz_749.5155_+_RT_4.421

Spectral match to Hydroxychloroquine

20200427_KAO_COVID_42 # 697, RT: 2.35min, FTMS(+), HCD, DDA, 350.2534@ 20,30,40

reference spectrum reproduced from https://mona.fiehnlab.ucdavis.edu/spectra/display/AU274102

Spectral match to Azithromycin

20200427_KAO_COVID_42, #1272, RT:4.37min, FTMS(+), HCD, DDA, 749.5159@ 20,30,40

reference spectrum reproduced from https://mona.fiehnlab.ucdavis.edu/spectra/display/EA019005

Supplementary Figure 4. Spectral matching to annotate two mass spectrometry features as drugs administered in the clinic, related to Figure 2. Feature abundances for two unidentified features acquired using the lipidomics platform were significantly associated with COVID-19 status and greater than 3 fold higher in COVID-19 samples vs. non-COVID-19 samples. **a** One feature with mass-to-charge (*m/z*) of 350.16278, had no direct match to any features in our spectra databases, but had a fragment ion (*m/z* 247.1) which matched fragments of the drug hydroxychloroquine. The *m/z* difference between the observed features and hydroxychloroquine (expected m/z 336.1837) is 13.979 and could match the mass of an additional oxygen and loss of a hydrogen. **b** The feature with a *m/z* of 749.5155 resulted in a good spectral match to Azithromycin.
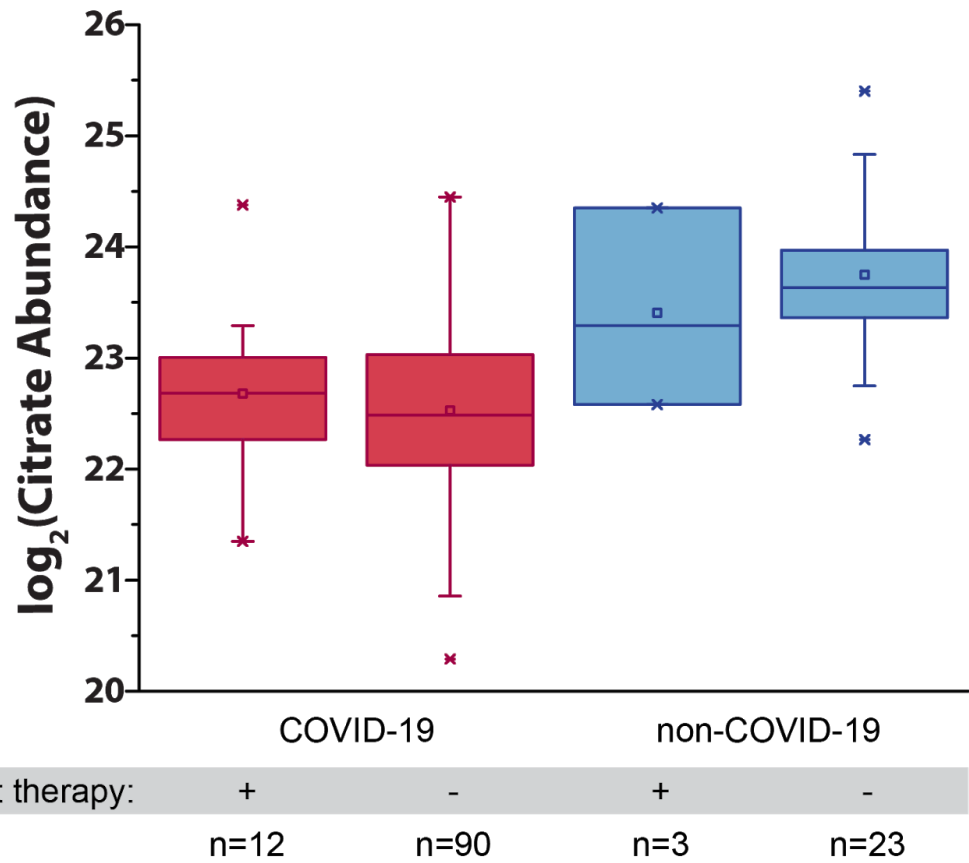
Supplementary Figure 5. Citrate administration as part of renal replacement therapy did not have a significant effect on citrate abundance, related to Figure 3. Citrate abundances (measured via mass spectrometry) were not modulated due to renal replacement therapy occurring prior to sample collection. Horizontal lines are at the median and the box extends to the first and third quartiles of the distribution.
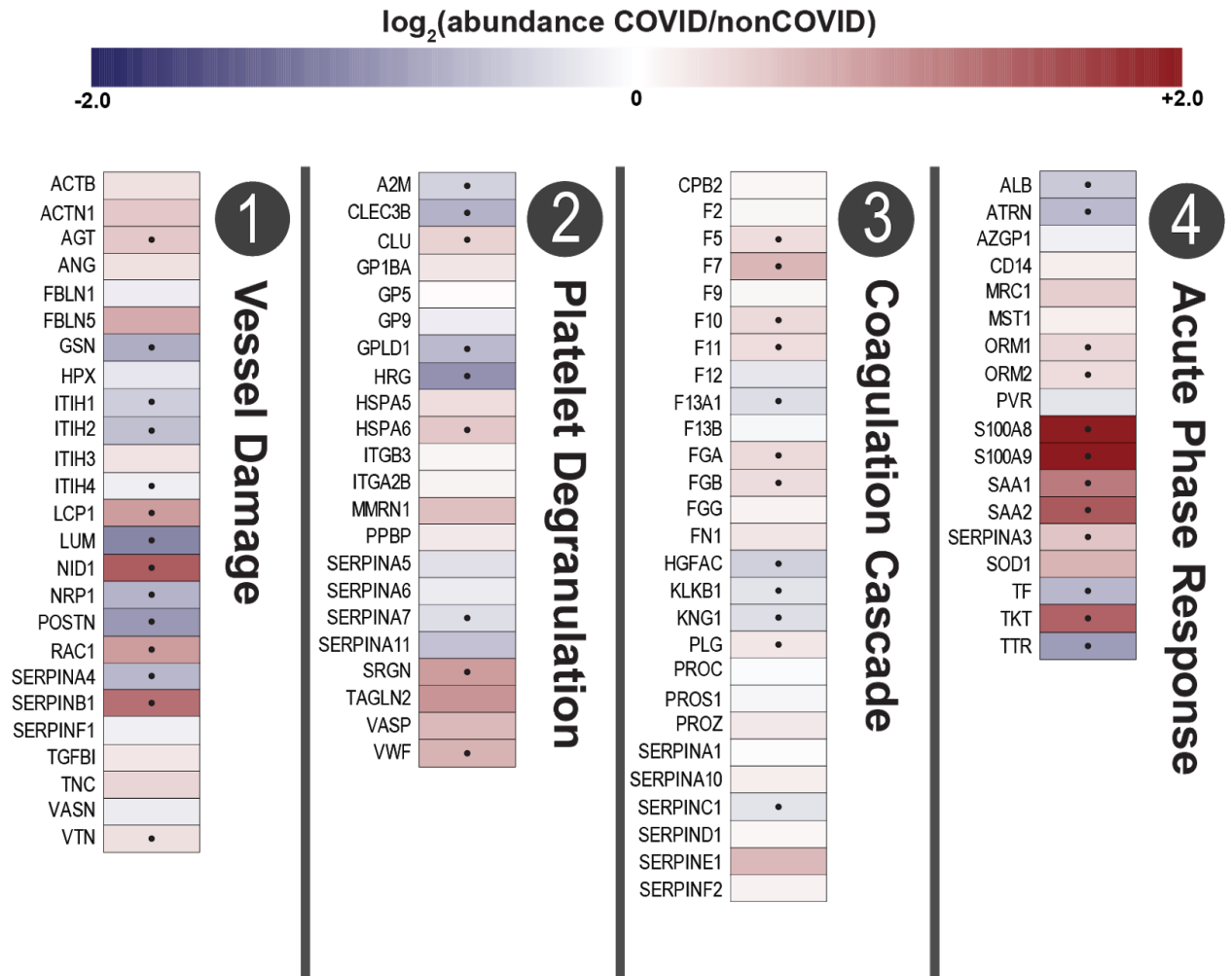
Supplementary Figure 6. The log2 fold changes (COVID-19/non-COVID-29) for features in the four categories listed in Figure 4b. The fold changes of means are denoted by color scale and the significance of the changes ($p < 0.05$) is denoted by *. The p-values were calculated using the webtool linear regression for COVID-19 status (https://covid-omics.app)
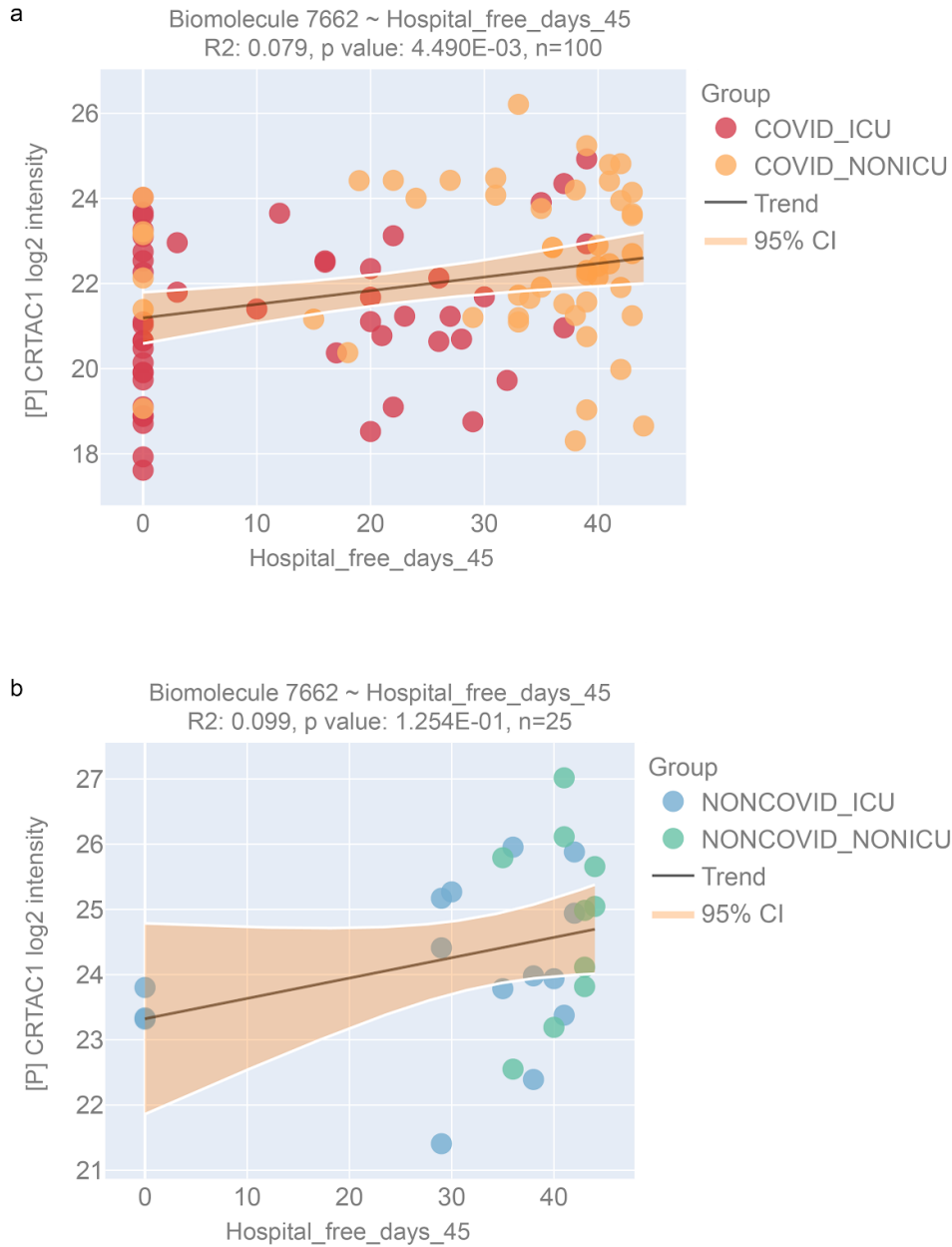
Supplementary Figure 7. Linear regression analysis of Cartilage acidic protein 1 measured abundance (log2(intensity)) against hospital free days 45, related to Figure 5. Cartilage acid protein 1 (CRTAC1) a protein [P] whose abundance was significantly lower in COVID-19 plasma samples (**Figure 5**) also had significant positive association with hospital free days 45 in COVID samples (a). This association was less pronounced in the non-COVID-19 cohort (**b**). These figures were obtained using the webtool (https://covid-omics.app).
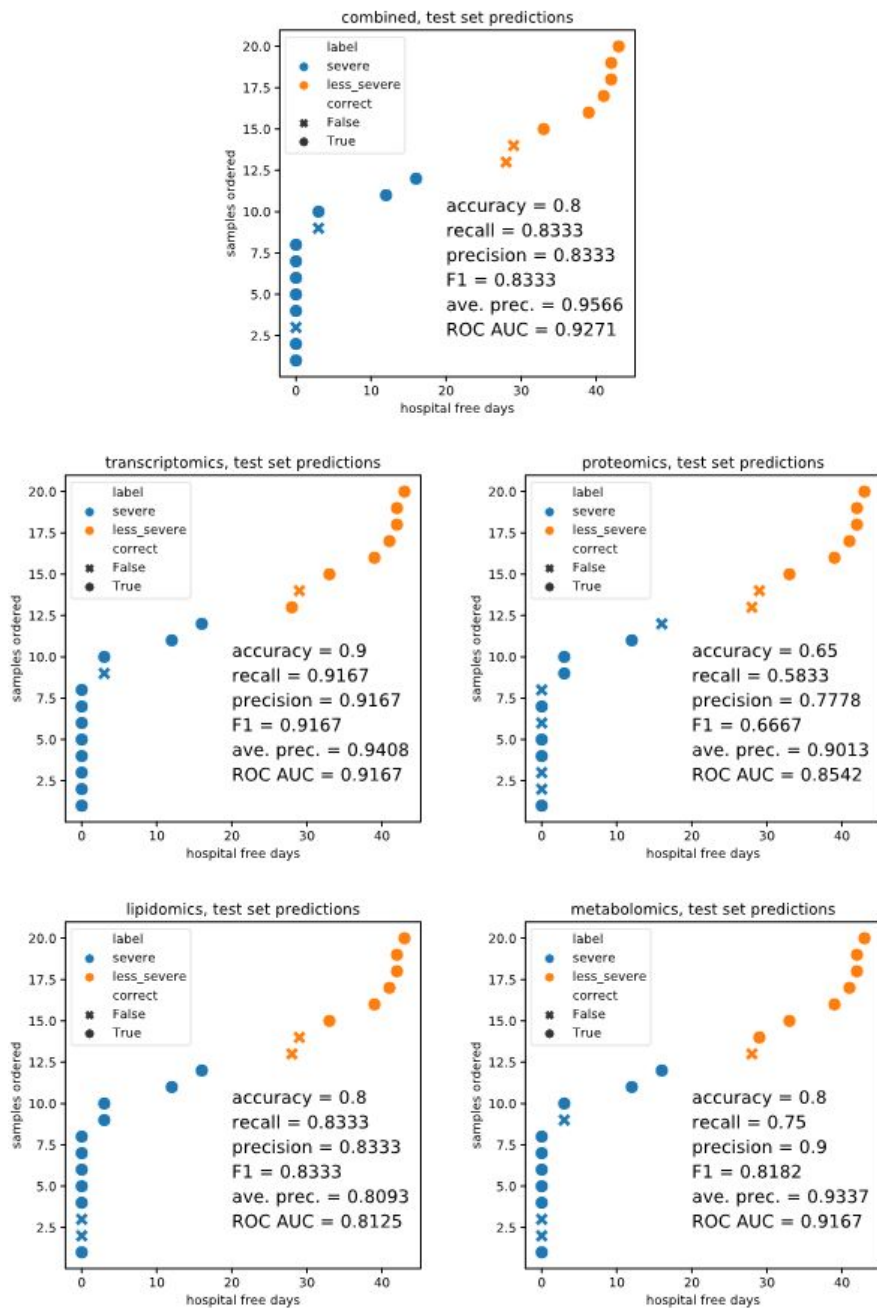
Supplementary Figure 8: Performance metrics of ExtraTrees classification models trained to predict COVID-19 "severe" or "less severe" using all omic data or each omic data subset. These data correspond to the analysis presented in **Figure 6**. "Severe" was defined as less than the median hospital free days (26 days) and "less severe" was defined as greater than the median hospital free days. For each panel, the prediction accuracy is annotated for the held out test-set samples (n=20).