RESEARCH ARTICLE

Figures and figure supplements

Seroconversion stages COVID19 into distinct pathophysiological states

Matthew D Galbraith *et al*

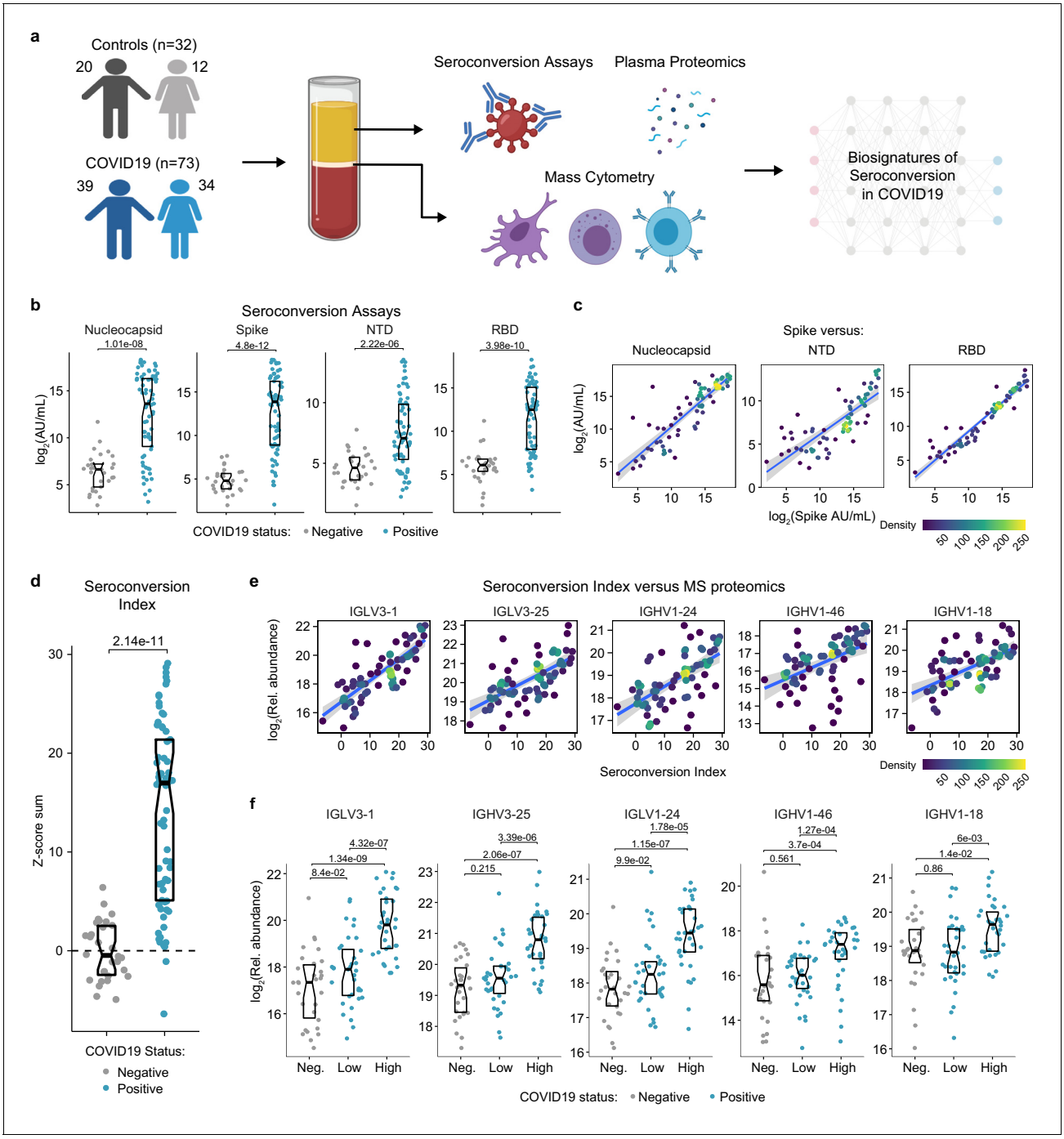


Figure 1. Highly variable seroconversion status among hospitalized COVID19 patients. (**a**) Overview of experimental approach. Blood samples from 105 research participants, 73 of them with COVID19, were analyzed by matched multiplex immunoassays for detection of antibodies against SARS-CoV-2, plasma proteomics using mass spectrometry (MS), SOMAscan proteomics, and cytokine profiling using Meso Scale Discovery (MSD) technology. Data was then analyzed to define biosignatures of seroconversion. (**b**) Multiplex immunoassays were used to measure antibodies against the SARS-CoV-2 nucleocapsid and spike proteins, as well as specific peptides encompassing the N-terminus domain (NTD) and receptor-binding domain (RBD) of the spike protein. Data are presented as modified Sina plots with boxes indicating median and interquartile range. Numbers above brackets are p-values for Mann–Whitney tests. (**c**) Scatter plots showing correlations between antibodies against the full-length spike protein versus antibodies against the nucleocapsid, NTD, and RBD domains. Points are colored by density; lines represent linear model fit with 95% confidence interval. (**d**) Seroconversion indices were calculated for each research participant by summing the Z-scores for each of the four seroconversion assays. Z-scores were calculated from the adjusted concentration values for each epitope in each sample, based on the mean and standard deviation of COVID19-negative samples. (**e**) Scatter plots displaying the top five correlations between seroconversion indices and proteins detected in the MS proteomics data set among COVID19

Figure 1 continued on next page

Figure 1 continued

patients. Points are colored by density; lines represent linear model fit with 95% confidence interval. (**f**) Sina plots showing values for the top five proteins correlated with seroconversion comparing the control cohort (Negative, Neg.) to COVID19 patients divided into seroconversion low and high status. Data are presented as modified Sina plots with boxes indicating median and interquartile range. Numbers above brackets are q-values for Mann–Whitney tests. See also ***Figure 1—figure supplement 1***.

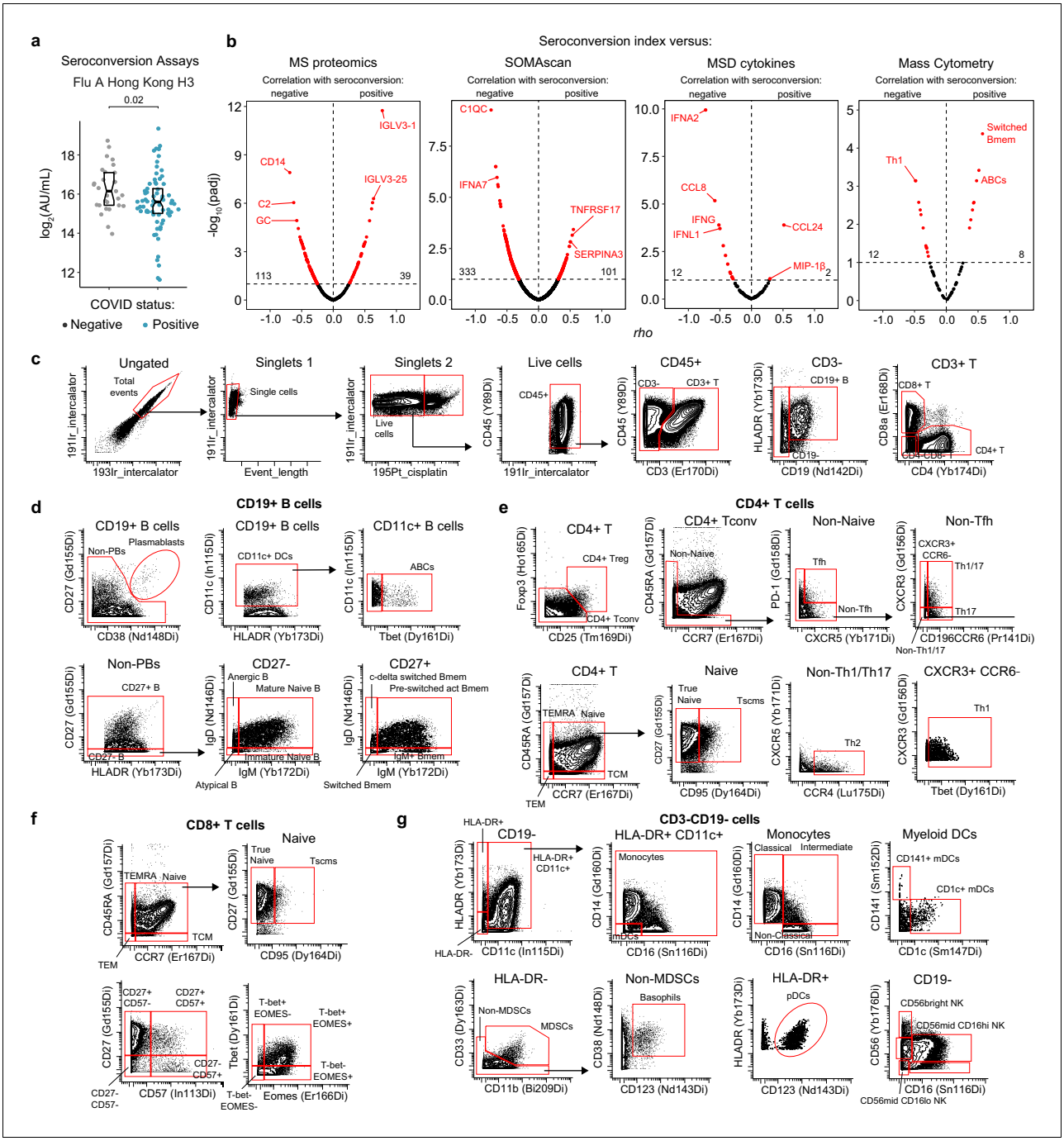


Figure 1—figure supplement 1. Biosignatures of seroconversion among hospitalized COVID19 patients. (**a**) Meso Scale Discovery (MSD) assays show no elevation in levels of circulating antibodies against the Flu A Hong Kong H3 virus strain among COVID19 patients. Data are presented as modified Sina plots with boxes indicating median and interquartile range. Number above brackets is p-value for Mann–Whitney test. (**b**) Volcano plots for Spearman correlations between seroconversion indices and circulating plasma proteins detected by mass spectrometry (MS), SOMAscan assays, MSD assays, as well as immune cell subsets detected by mass cytometry (MC) among all live cells. X axes show Spearman *rho* values. Y axes show $-\log_{10}$ p-values adjusted with Benjamini–Hochberg method. Dashed vertical line indicates *rho* = 0. Dashed horizontal line indicates the statistical cut off of false discovery rate (FDR) = 10 (q = 0.1). (**c–g**) Representation of gating strategy employed during MC analysis of peripheral immune cell lineages. In (**c**), single live cells were gated for CD45+ staining followed by gating into T cells (CD3+), B cells (CD3- CD19+), and CD4+ and CD8+ T cells subsets. In (**d**), B cells were further gated into the indicated subsets. In (**e**) and (**f**), CD4+ and CD8+ T cell lineages were further characterized in the indicated subsets. Panel (**g**) shows gating for the indicated myeloid subsets and natural killer (NK) cells.

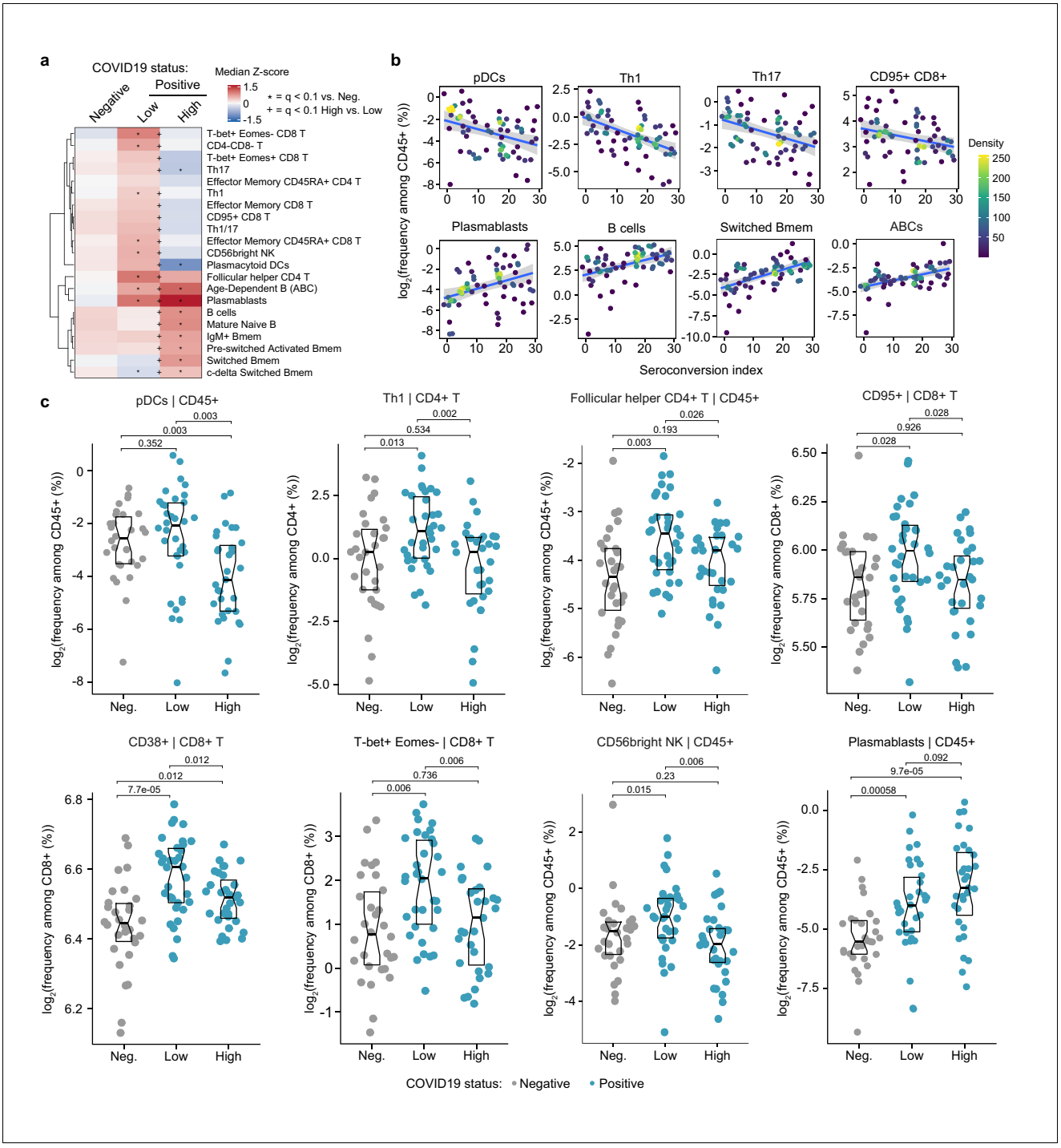


Figure 2. Seroconversion associates with significant changes in peripheral immune cell frequencies. (**a**) Heatmap representing changes in the frequency of immune cell subsets that are significantly correlated, either positively or negatively with seroconversion status. Values displayed are median Z-scores, derived from cell frequencies among all CD45+ cells, for each cell subset for controls (negative, Neg.) versus COVID19 patients divided into seroconversion low (Low) and high (High) status. Z-scores were calculated from the adjusted frequency values for each cell type in each sample, based on the mean and standard deviation of COVID19-negative samples. Asterisks indicate a significant difference relative to the control COVID19-negative group, and the + symbols indicate a significant difference between sero-low and sero-high groups after multiple hypothesis correction ($q < 0.1$, Mann–Whitney test). (**b**) Scatter plots for indicated immune cell types significantly correlated with seroconversion indices among COVID19 patients. Points are colored by density; lines represent linear model fit with 95% confidence interval. (**c**) Sina plots showing values for indicated immune cell types significantly correlated with seroconversion indices among COVID19 patients. The parent cell lineage is indicated in the header and Y axis label for each plot. Data are presented as modified Sina plots with boxes indicating median and interquartile range. Numbers above brackets are q-values for Mann–Whitney tests. See also ***Figure 2—figure supplement 1***.

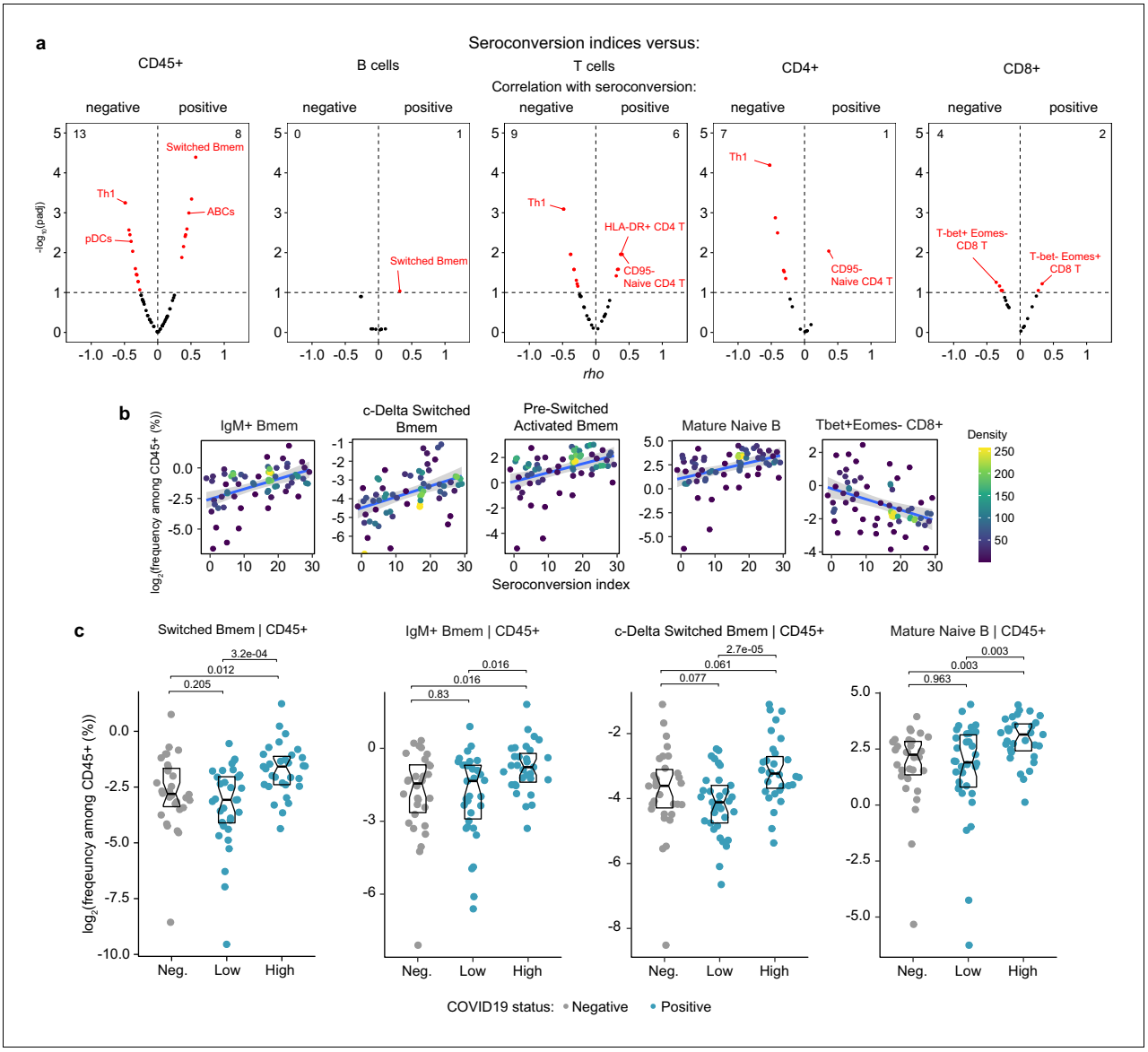


Figure 2—figure supplement 1. Immune cell signatures of seroconversion. (**a**) Volcano plots displaying the correlations between seroconversion indices and circulating levels of immune cell subsets detected by mass cytometry (MC). X axes show Spearman *rho* values. Y axes show $-\log_{10}$ p-values adjusted with Benjamini–Hochberg method. Dashed vertical line indicates *rho* = 0. Dashed horizontal line indicates a false discovery rate (FDR) threshold of 10% (q = 0.1). Correlations were calculated for immune cell subsets measured as frequencies among all live CD45+ cells (far left), all B cells, all T cells, CD4+ T cells, and CD8+ T cells (far right). (**b**) Scatter plots for indicated cell types against seroconversion indices. Values shown are derived from frequency among all CD45+ cells. Points are colored by density; lines represent linear model fit with 95% confidence interval. (**c**) Sina plots for indicated immune cell types comparing controls (Negative, Neg.) to COVID19 patients divided into seroconversion low (Low) and high (High) status. Data are presented as modified Sina plots with boxes indicating median and interquartile range. Numbers above brackets are q-values for Mann–Whitney tests.

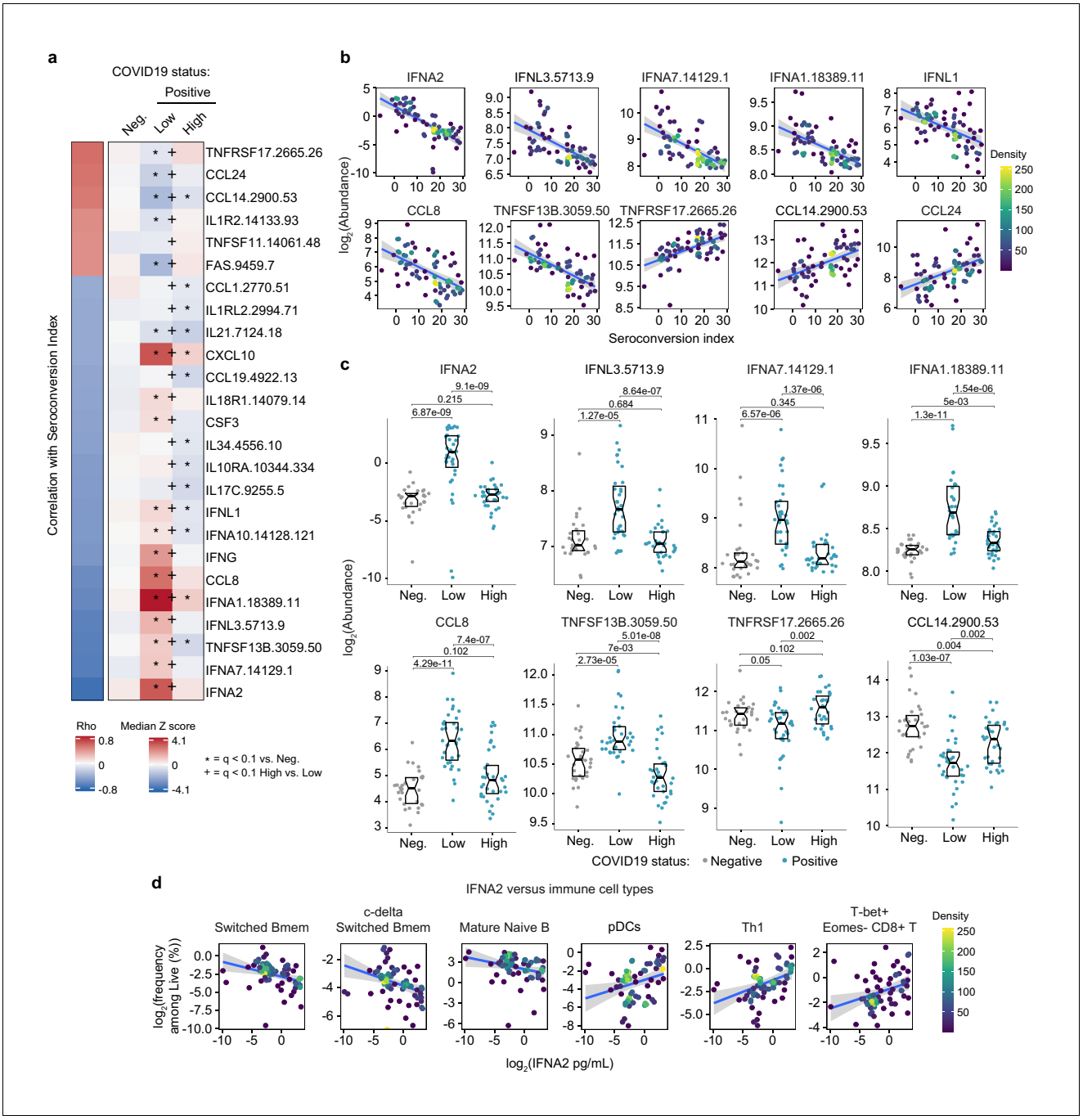


Figure 3. Seroconversion is associated with decreased interferon signaling. (**a**) Heatmap displaying changes in circulating levels of immune factors that are significantly correlated, either positively or negatively, with seroconversion status. The left column represents Spearman *rho* values, while the right columns display median Z-scores for each immune factor for controls (negative, Neg.) versus COVID19 patients divided into seroconversion low (Low) and high (High) status. Factors are ranked from most positively correlated (top, high *rho* values) to most anti-correlated (bottom, low *rho* values) with seroconversion index. Z-scores were calculated from the adjusted concentration values for each immune factor in each sample, based on the mean and standard deviation of COVID19-negative samples. Asterisks indicate a significant difference relative to the control COVID19-negative group, and the + symbols indicate a significant difference between sero-low and sero-high groups ($q < 0.1$, Mann–Whitney test). (**b**) Scatter plots for indicated immune factors significantly correlated with seroconversion indices among COVID19 patients. Points are colored by density; lines represent linear model fit with 95% confidence interval. (**c**) Sina plots showing values for immune factors correlated with seroconversion comparing controls (Neg.) to COVID19 patients divided into seroconversion low and high status. Data are presented as modified Sina plots with boxes indicating median and interquartile range. Numbers above brackets are q-values for Mann–Whitney tests. (**d**) Scatter plots showing correlations between circulating levels of IFNA2 measured by MSD and the indicated cell types measured by mass cytometry. Values for immune cells correspond to frequency among all live cells. Points are colored by density; lines represent linear model fit with 95% confidence interval. See also ***Figure 3—figure supplement 1***.

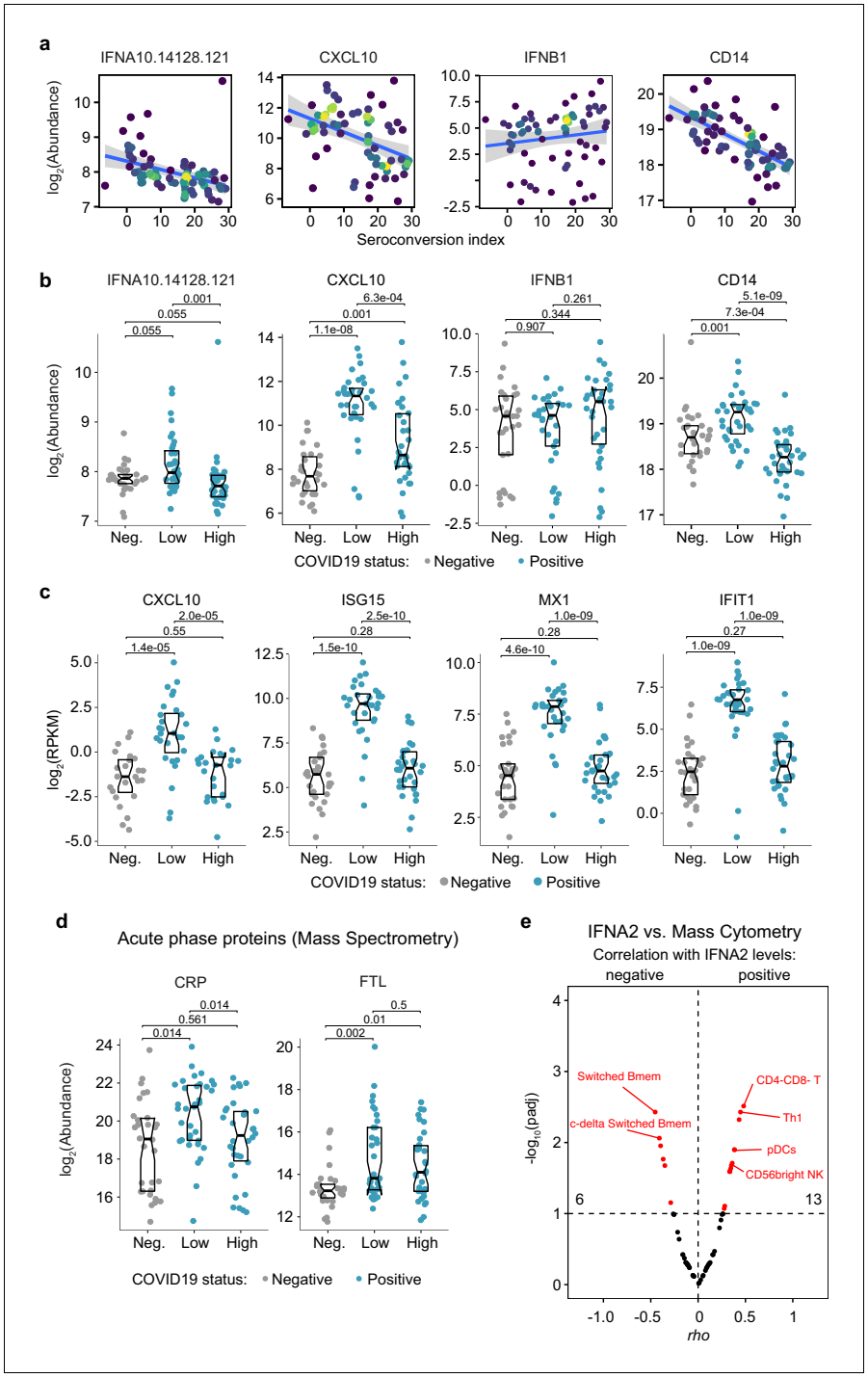


Figure 3—figure supplement 1. Seroconversion associates with differential abundance of circulating immune factors. (**a and b**) XY scatter plots (**a**) and Sina plots (**b**) for select circulating immune factors. Points in (**a**) are colored by density; lines represent linear model fit with 95% confidence interval. Data in (**b**) are presented as modified Sina plots with boxes indicating median and interquartile range. Numbers above brackets are q-values for Mann–Whitney tests. (**c**) Sina plots showing expression of select IFN-inducible mRNAs in a whole blood transcriptome analysis comparing controls (Neg., negative) to COVID19 patients divided into seroconversion low (Low) and high (High) status. Data are presented as in **b**. (**d**) Sina plots for the acute phase proteins CRP and FTL (Ferritin Light Chain) detected by MS comparing controls (Neg., negative) to COVID19 patients divided into seroconversion low (Low) and high (High) status. Data are presented as in **b**. (**e**) Volcano plot showing associations between circulating levels of IFNA2 measured by MSD versus immune cell subsets among all live peripheral blood

Figure 3—figure supplement 1 continued on next page

Figure 3—figure supplement 1 continued

mononuclear cells. X axes show Spearman *rho* values. Y axes show $-\log_{10}$ p-values adjusted with Benjamini–Hochberg method. Dashed vertical line indicates *rho* = 0. Dashed horizontal line indicates a false discovery rate (FDR) threshold of 10% ($q < 0.1$).

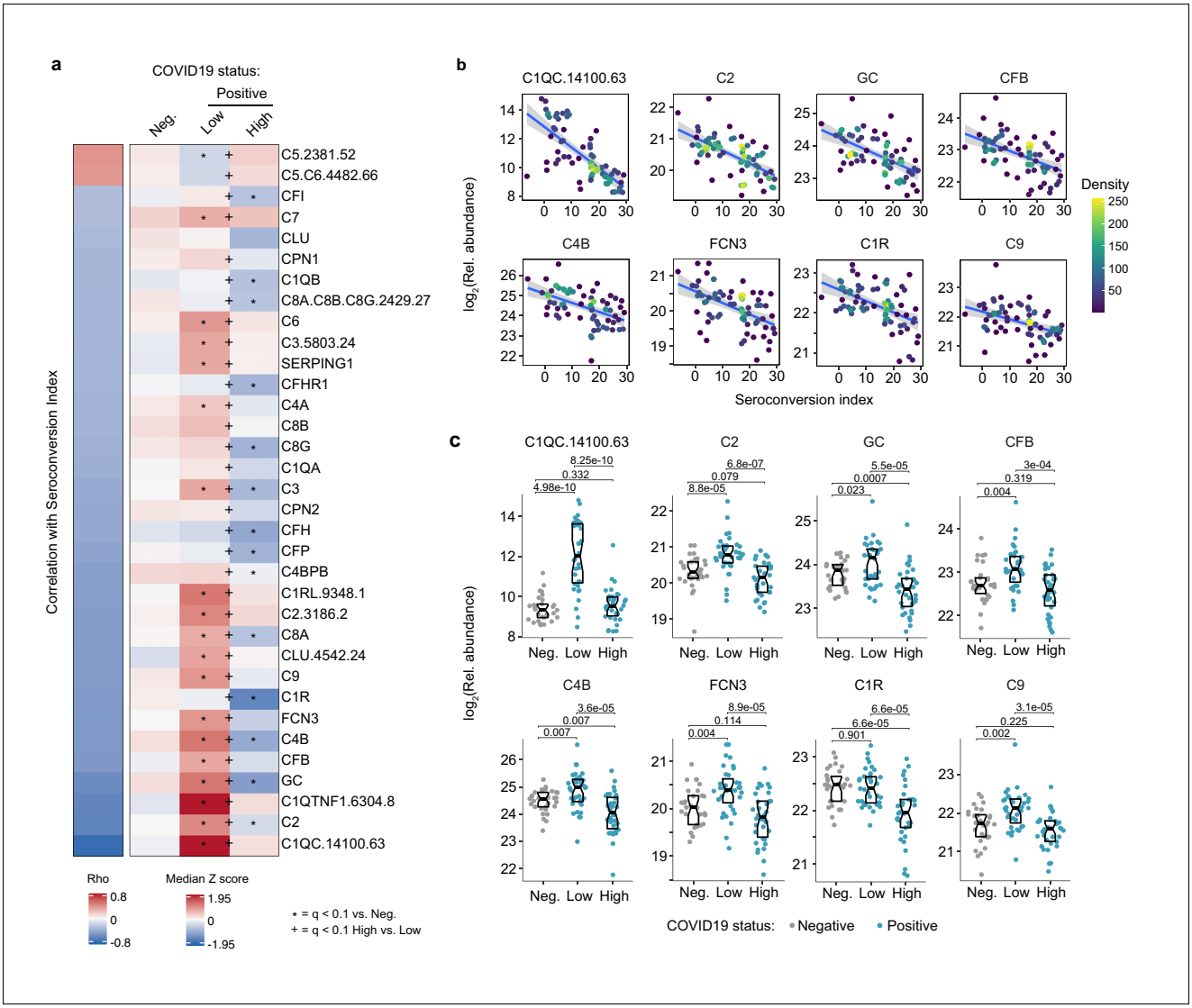


Figure 4. Seroconversion correlates with decreased markers of systemic complement activation. (**a**) Heatmap displaying changes in circulating levels of components of the various complement pathways that are significantly correlated, either positively or negatively, with seroconversion status. The left column represents Spearman *rho* values, while the right columns display median Z-scores for each complement factor for controls (negative, Neg.) versus COVID19 patients (positive) divided into seroconversion low (Low) and high (High) status. Factors are ranked from most positively correlated (top, high *rho* values) to most anti-correlated (bottom, low *rho* values) with seroconversion status. Z-scores were calculated from the adjusted concentration values for each analyte in each sample, based on the mean and standard deviation of COVID19-negative samples. Asterisks indicate a significant difference relative to the control COVID19-negative group, and the + symbols indicate a significant difference between sero-low and sero-high groups ($q < 0.1$, Mann–Whitney test). (**b**) Scatter plots for indicated complement factors significantly correlated with seroconversion indices among COVID19 patients. Points are colored by density; lines represent linear model fit with 95% confidence interval. (**c**) Sina plots showing values for complement factors correlated with seroconversion comparing controls (Negative, Neg.) to COVID19 patients divided into seroconversion low (Low) and high (high) status. Data are presented as modified Sina plots with boxes indicating median and interquartile range. Numbers above brackets are q-values for Mann–Whitney tests. See also ***Figure 4—figure supplement 1***.

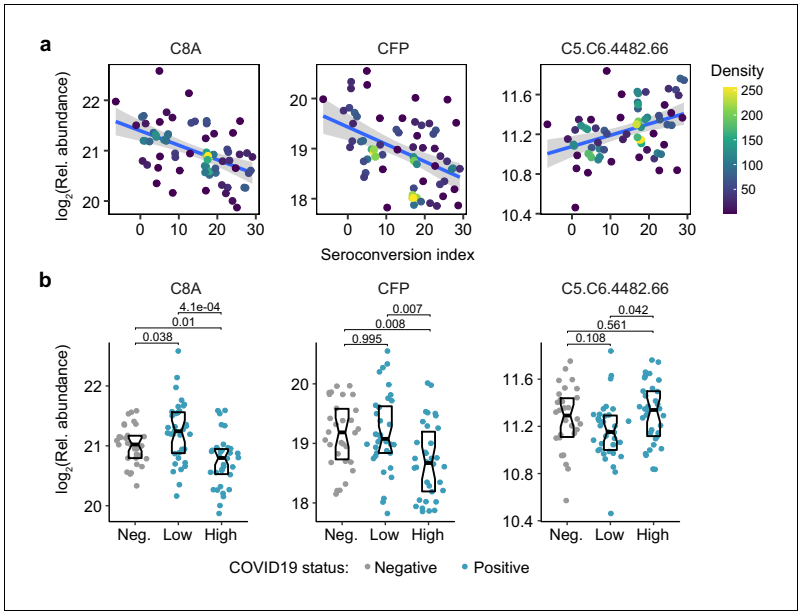


Figure 4—figure supplement 1. Seroconversion associates with decreased markers of systemic complement activation. (**a and b**) XY scatter plots (**a**) and Sina plots (**b**) for select complement factors significantly associated, either positively or negatively, with seroconversion indices among COVID19 patients. Points in (**a**) are colored by density; lines represent linear model fit with 95% confidence interval. Data in (**b**) are presented as modified Sina plots with boxes indicating median and interquartile range comparing controls (Neg., negative) to COVID19 patients divided into seroconversion low (Low) and high (High) status. Numbers above brackets are q-values for Mann–Whitney tests.

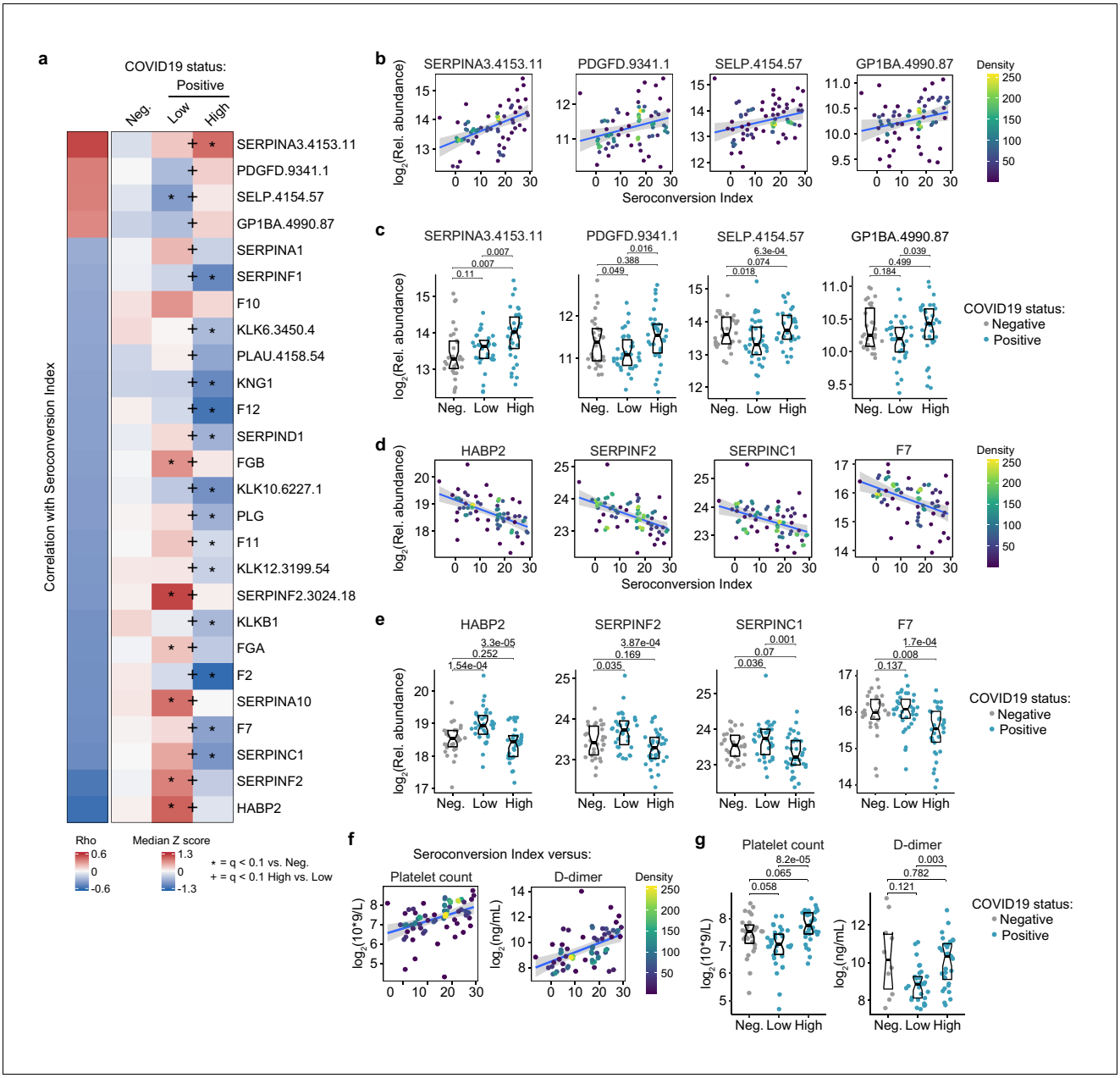


Figure 5. Seroconversion associates with remodeling of the hemostasis control network. (**a**) Heatmap displaying changes in circulating levels of known modulators of hemostasis that are significantly correlated, either positively or negatively, with seroconversion status. The left column represents Spearman *rho* values, while the right columns display row-wise Z-scores for each factor for controls (negative, Neg.) versus COVID19 patients divided into seroconversion low (Low) and high (High) status. Factors are ranked from most positively correlated (top, high *rho* values) to most anti-correlated (bottom, low *rho* values) with seroconversion status. Z-scores were calculated from the adjusted concentration values for each analyte in each sample, based on the mean and standard deviation of COVID19-negative samples. Asterisks indicate a significant difference relative to the control COVID19-negative group, and the + symbols indicate a significant difference between sero-low and sero-high groups (q < 0.1, Mann–Whitney test). (**b and c**) Scatter plots (**b**) and Sina plots (**c**) for factors positively correlated with seroconversion indices. (**d and e**) Scatter plots (**d**) and Sina plots (**e**) for factors negatively correlated with seroconversion indices. (**f and g**) Scatter plot (**f**) and Sina plots (**g**) displaying the correlations between seroconversion index and platelet counts and D-dimer values obtained from clinical laboratory testing. Points in (**b**), (**d**), and (**f**) are colored by density; lines represent linear model fit with 95% confidence interval. Data in (**c**), (**e**), and (**g**) are presented as modified Sina plots with boxes indicating median and interquartile range. Numbers above brackets are q-values for Mann–Whitney tests. See also ***Figure 5—figure supplement 1***.

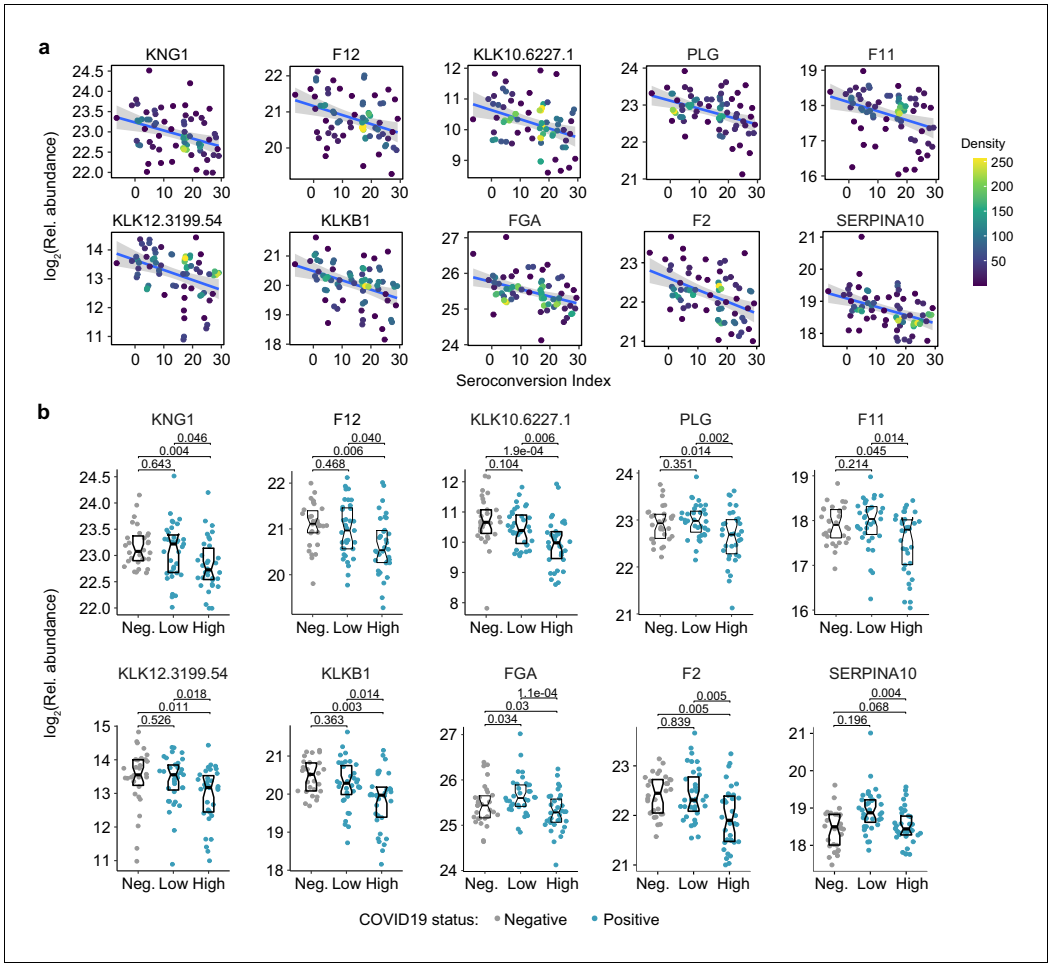
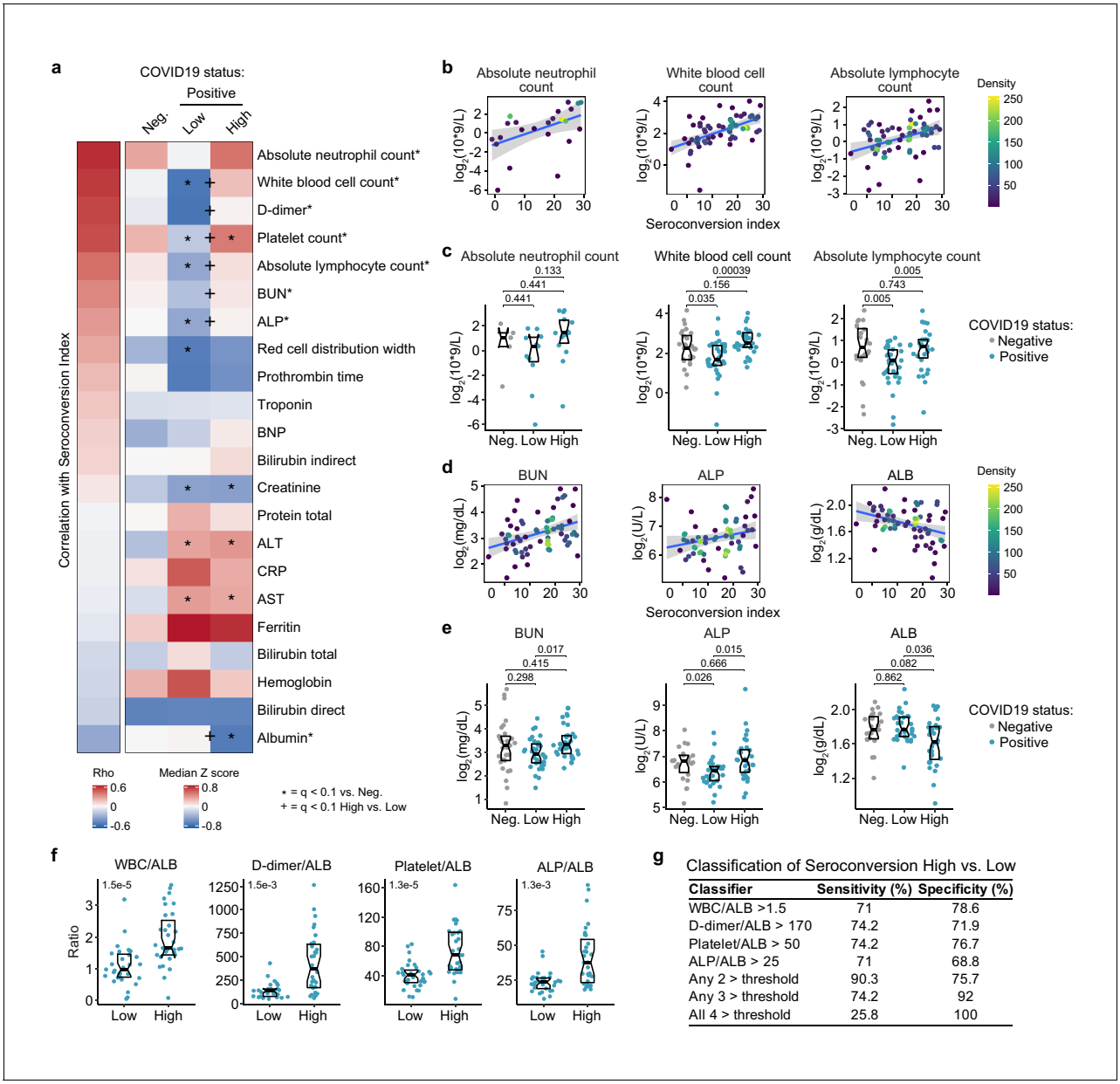


Figure 5—figure supplement 1. Seroconversion associates with remodeling of the hemostasis network. (**a and b**) XY scatter plots (**a**) and Sina plots (**b**) for select factors involved in control of hemostasis significantly associated, either positively or negatively, with seroconversion indices among COVID19 patients. Points in (**b**) are colored by density; lines represent linear model fit with 95% confidence interval. Data in (**b**) are presented as modified Sina plots with boxes indicating median and interquartile range comparing controls (Neg., negative) to COVID19 patients divided into seroconversion low (Low) and high (High) status. Numbers above brackets are q-values for Mann–Whitney tests.



Classifier	Sensitivity (%)	Specificity (%)
WBC/ALB >1.5	71	78.6
D-dimer/ALB > 170	74.2	71.9
Platelet/ALB > 50	74.2	76.7
ALP/ALB > 25	71	68.8
Any 2 > threshold	90.3	75.7
Any 3 > threshold	74.2	92
All 4 > threshold	25.8	100

Figure 6. Seroconversion associates with recovery in blood cell numbers and hypoalbuminemia. (**a**) Heatmap displaying correlations between clinical laboratory values and seroconversion status. The left column represents Spearman *rho* values, while the right columns display row-wise Z-scores for each variable for controls (negative, Neg.) versus COVID19 patients divided into seroconversion low (Low) and high (High) status. Measures are ranked from most positively correlated (top, high *rho* values) to most anti-correlated (bottom, low *rho* values) with seroconversion status. Asterisks after the clinical parameter name indicate a significant correlation. Z-scores were calculated from the adjusted concentration values for each analyte in each sample, based on the mean and standard deviation of COVID19-negative samples. Asterisks indicate a significant difference relative to the control COVID19-negative group, and the + symbols indicate a significant difference between sero-low and sero-high groups (q < 0.1, Mann–Whitney test). (**b–e**) Scatter plots (**b** and **c**) and Sina plots (**c** and **e**) for indicated clinical laboratory values significantly correlated with seroconversion indices among COVID19 patients. In **b** and **c**, points are colored by density; lines represent linear model fit with 95% confidence interval. In **c** and **e**, Sina plots show values for clinical laboratory tests correlated with seroconversion comparing controls (Negative, Neg.) to COVID19 patients divided into seroconversion low (Low) and high (High) status. Data are presented as modified Sina plots with boxes indicating median and interquartile range. Numbers above brackets are q values for Mann–Whitney tests. (**f**) Differences in the indicated ratios of clinical laboratory values between sero-low and sero-high COVID19 patients. Data are presented as modified Sina plots with boxes indicating median and interquartile range. Numbers at upper left of each plot are p-values for Mann–Whitney tests. (**g**) Table showing how the indicated ratios of the specified clinical values could be potentially used to gauge the seroconversion status of a hospitalized patient with moderate pathology. The units employed for calculating these ratios are 10^3/mcL for white blood cells (WBC) and platelets; g/dL for albumin (ALB); ng/mL for D-dimer; and U/L for alkaline phosphatase (ALP). ALT: alanine aminotransferase, AST:

Figure 6 continued on next page

Figure 6 continued

aspartate aminotransferase; BUN: blood urea nitrogen; BNP: brain natriuretic peptide; CRP: C-reactive protein. See also ***Figure 6—figure supplement 1***.

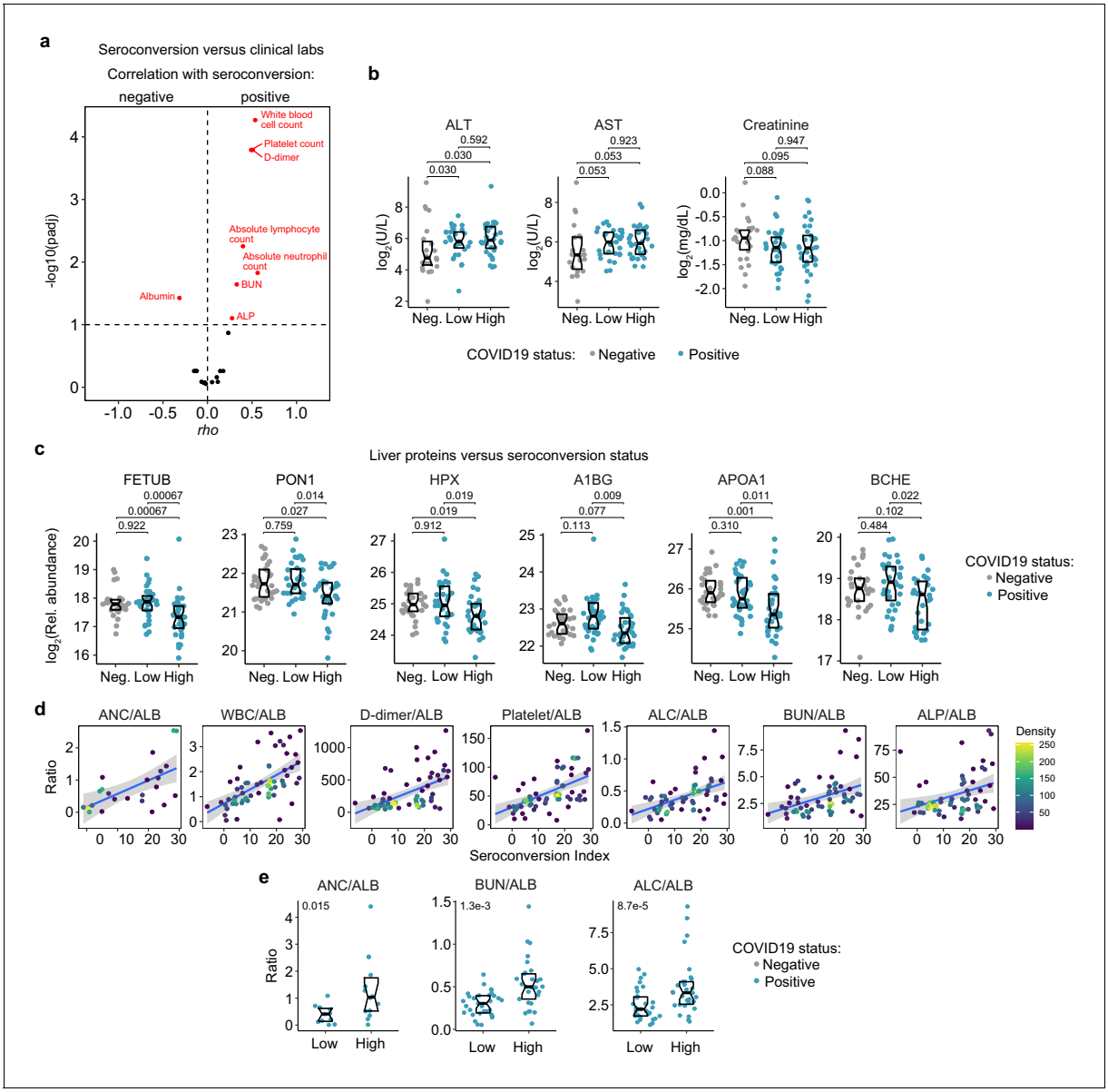


Figure 6—figure supplement 1. Seroconversion associates with recovery of blood cell counts and hypoalbuminemia. (**a**) Volcano plot displaying the correlations between seroconversion indices and clinical laboratory values. X axes show Spearman *rho* values. Y axes show $-\log_{10}$ p-values adjusted with Benjamini–Hochberg method. Dashed vertical line indicates *rho* = 0. Dashed horizontal line indicates a false discovery rate (FDR) threshold of 10% ($q < 0.1$). (**b**) Sina plots for select clinical laboratory values. Data are presented as modified Sina plots with boxes indicating median and interquartile range comparing controls (Neg., negative) to COVID19 patients divided into seroconversion low (Low) and high (High) status. Numbers above brackets are q-values for Mann–Whitney tests. (**c**) Sina plots showing values for liver proteins detected by MS that are significantly correlated with seroconversion comparing controls (Negative, Neg.) to COVID19 patients divided into seroconversion low (Low) and high (High) status. Data are presented as modified Sina plots with boxes indicating median and interquartile range. Numbers at upper left of each plot are p-values for Mann–Whitney tests. (**d**) Scatter plots showing the correlation between the indicated ratios of clinical values and the seroconversion index. ALB: albumin; ALC: absolute lymphocyte count; ANC: absolute neutrophil count; ALT: alanine aminotransferase; AST: aspartate aminotransferase; BUN: blood urea nitrogen, WBC: white blood cell count. (**e**) Sina plots for indicated clinical laboratory ratios. Data are presented as modified Sina plots with boxes indicating median and interquartile range comparing controls (Neg., negative) to COVID19 patients divided into seroconversion low (Low) and high (High) status. Numbers at upper right of each plot are q-values for Mann–Whitney tests.

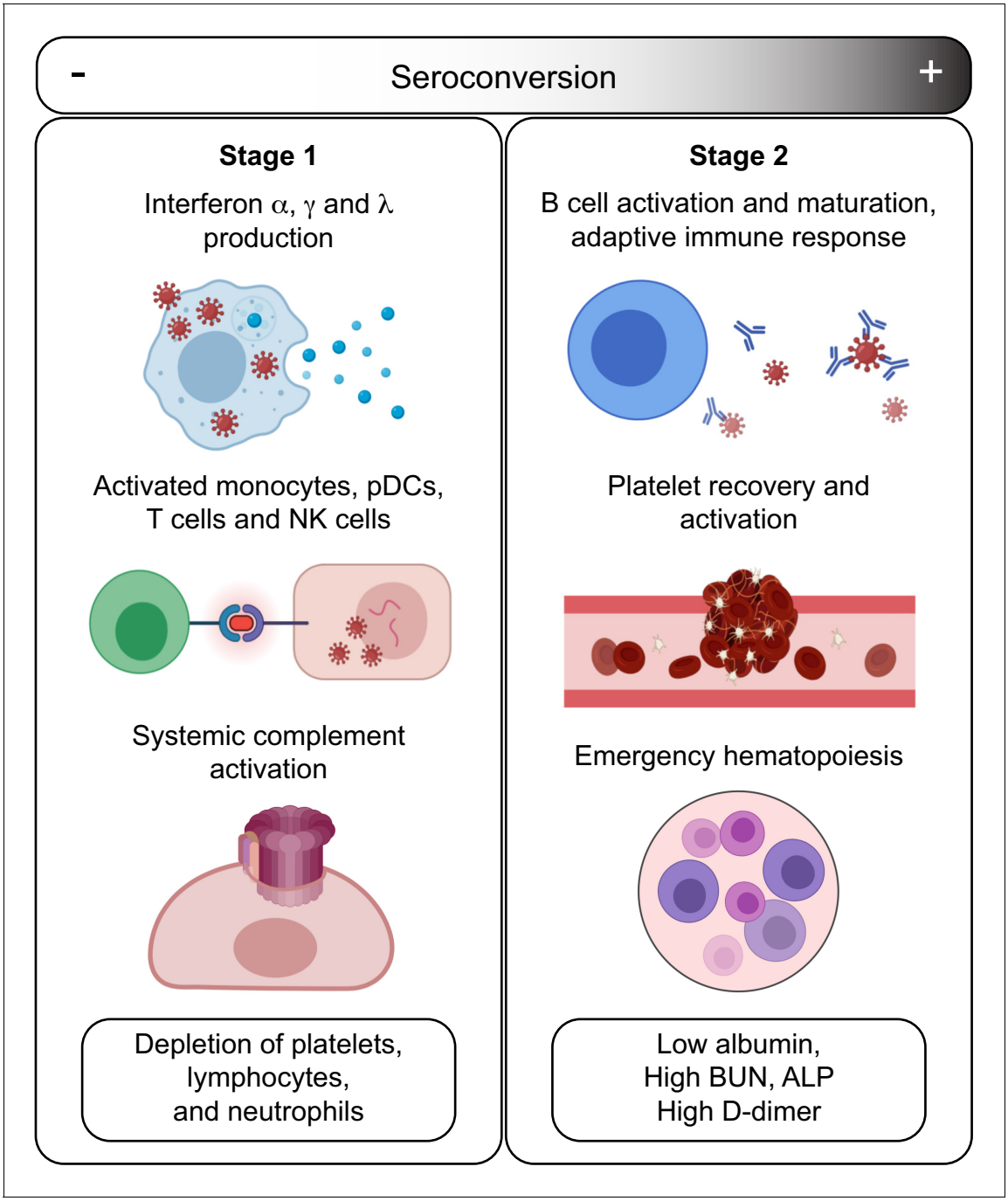


Figure 7. Model for staging COVID19 pathophysiology based on seroconversion status. Stage 1 applies to COVID19 patients with low degree of seroconversion and involves high levels of circulating IFNs, signs of strong systemic complement activation, hyperactive T cells, activated monocytes, and cytokine-producing NK cells, as well as depletion of key blood cell types. Stage 2 applies to COVID19 patients with high degree of seroconversion and is characterized by increased blood cell numbers, increased levels of markers of platelet degranulation, elevated D-dimer, and markers of increased liver dysfunction and/or interstitial leakage.