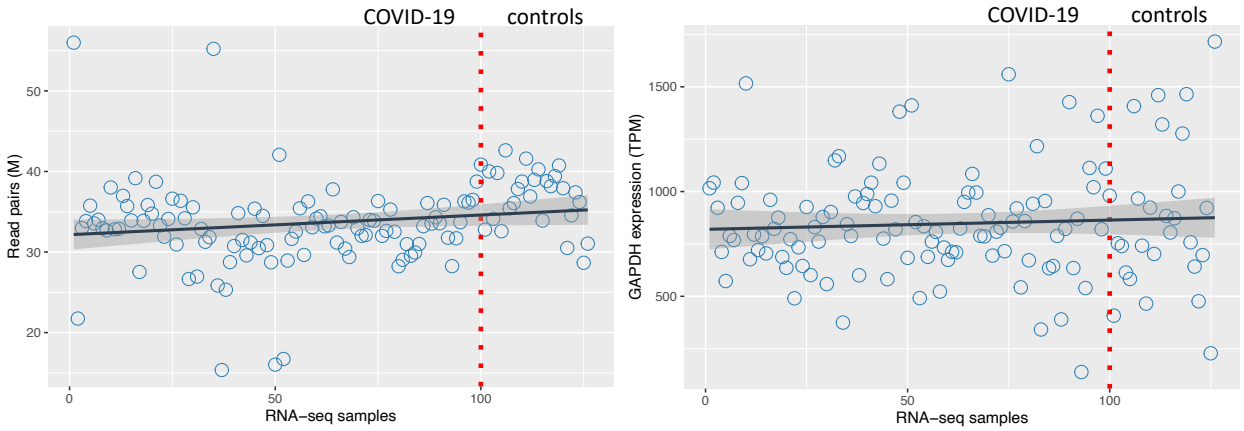
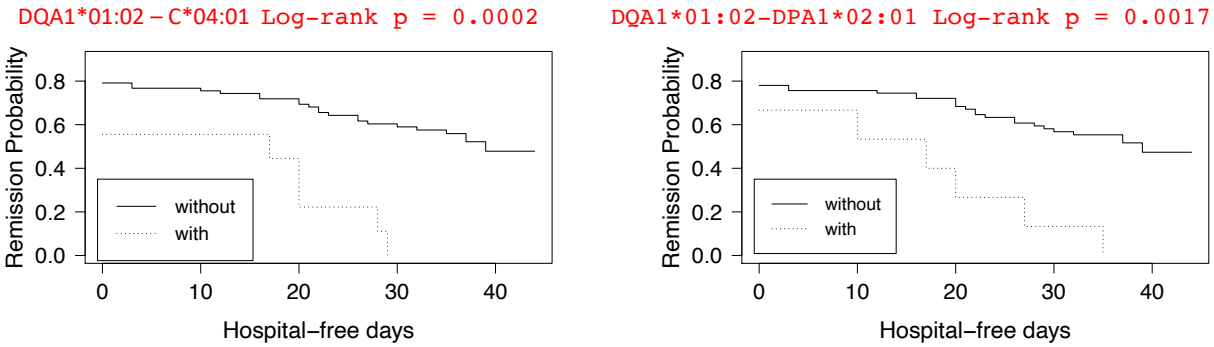
Supplementary data (additionalFile1) for:

Warren RL and Birol I. HLA alleles measured from COVID-19 patient transcriptomes reveal associations with disease prognosis.



Supplementary Fig. S1. COVID-19 New-York cohort (Overmyer [22]) transcriptome data characteristics. Left panel: RNA-Seq read pair count in each patient sample shows sequencing data consistency across. Right panel: We evaluated base RNA expression across all patient RNA-Seq samples by measuring, using Salmon (Patro et al., 2017), the transcript per million (TPM) expression of house-keeping gene Glyceraldehyde 3-phosphate dehydrogenase (abbreviated GAPDH) (EC 1.2. 1.12) in peripheral blood mononuclear cells (PBMC). This was done solely to evaluate the quality of the RNA-Seq prior to running HLA prediction tools. The average GAPDH TPM (847.73 +/- 272.32) is indicated.

Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017 Apr;14(4):417-419. doi: 10.1038/nmeth.4197. PMID: 28263959; PMCID: PMC5600148.



Supplementary fig. S2. HLA haplotypes tested for higher risk of hospitalization in a COVID-19 positive patient cohort. COVID-19 positive patients were split into two groups per haplotype tested, depending on whether they were predicted to have the HLA allele under scrutiny or not. Considering HLA alleles present in 10% or more of the samples (HLA-I n=17 and HLA-II n=11), we tested a total of 378 possible haplotypes for risk associations. We ran the Kaplan-Meier estimator (R package survival) using the Overmyer et al. [22] cohort (2020) HFD-45 metric for estimating the remission probability of patients without or with these haplotypes to estimate the statistical significance. Log-rank p values were calculated for each (R package survminer) and are indicated on the plots. Haplotypes DQA1*01:02 – C*04:01 (left panel) and DQA1*01:02 – DPA1*02:01 (right panel) had nominally significant associations that were not statistically significant after multiple hypothesis testing (Bonferroni p=0.0752 and p=0.6426, respectively, using n=378). We caution that these two haplotypes were observed in only 9% of patients.