

Supplemental Material For:

Microbiome-Transcriptome Interactions Related to Severity of Respiratory Syncytial Virus Infection

Abhijeet R. Sonawane^{1*}, Liang Tian^{2*}, Chinyi Chu³, Xing Qiu⁴, Lu Wang⁴, Jeanne Holden-Wiltse⁴, Alex Grier⁵, Steven R. Gill⁵, Mary T. Caserta³, Ann R. Falsey⁶, David J. Topham⁵, Edward E. Walsh⁶, Thomas J. Mariani³, Scott T. Weiss¹, Edwin K. Silverman¹, Kimberly Glass^{1#}, Yang-Yu Liu^{1,7#}

¹*Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA* ²*Department of Physics, Hong Kong Baptist University, Kowloon Tong, Hong Kong SAR, China* ³*Departments of Pediatrics,* ⁴*Biostatistics and Computational Biology,* ⁵*Microbiology and* ⁶*Medicine, University of Rochester Medical Center, Rochester, NY 14642, USA* ⁷*Center for Cancer Systems Biology, Dana Farber Cancer Institute, Boston, MA 02115, USA*

*These authors contributed equally to this work.

#Correspondence should be addressed to KG (kimberly.glass@channing.harvard.edu) or YYL (yyl@channing.harvard.edu)

Supplemental Figure Legends

Supplemental Figure 1: Analysis of comparison of viral load data and hospitalization against disease severity. We plot the RSV severity score GRSS against viral load values for the samples collected from nasal swab and nasal wash. We did not find any evidence for disease severity being affected by viral load. The points were colored according to hospitalization status (red= hospitalized and black = Not hospitalized). The dashed line represent classification of severity in mild and severe groups based on GRSS.

Supplemental Figure 2: PCA and MDS (multidimensional scaling) on transcriptome data. Principal component analysis was performed on all 80 transcriptomic samples using the log2 normalized RNA-Seq FPKM data and MDS plots for the same samples were obtained using the plotMDS() function in *limma*. For each analysis, points in the plots have been colored to indicate the corresponding sample's severity, sex, enrollment season, race, ethnicity, and visit number. The left hand panel shows results for PCA and MDS performed on the transcriptomic data prior to batch-correction; the right hand panel shows the analyses after batch-correcting the transcriptomic data for enrollment season.

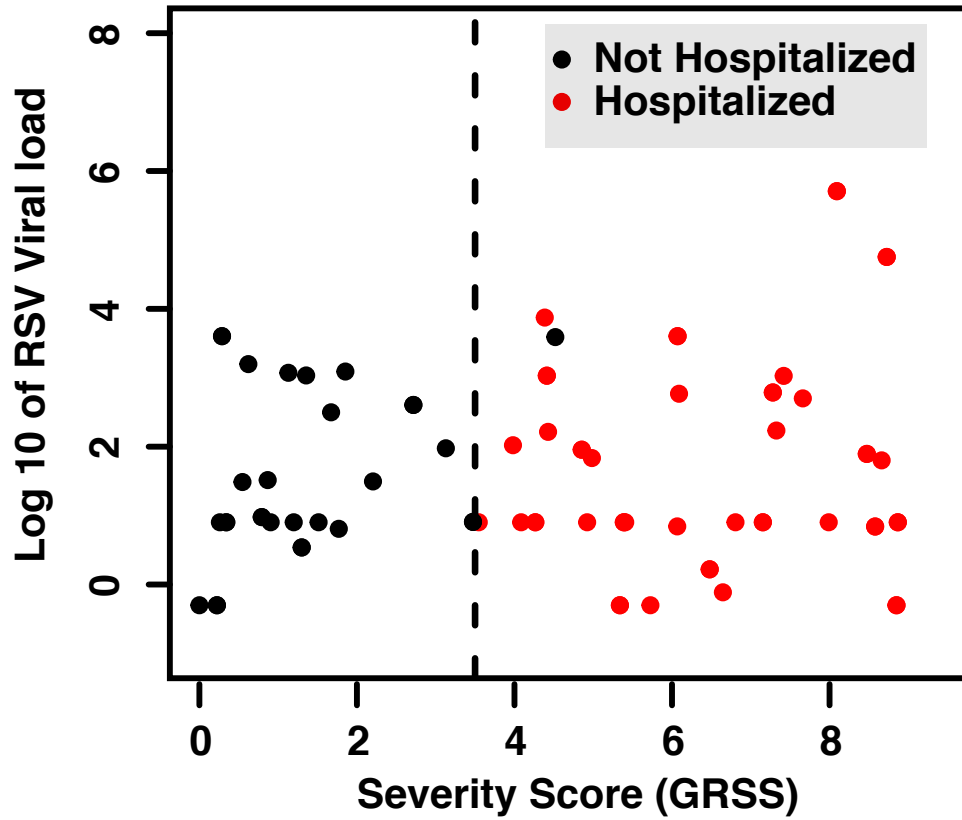
Supplemental Figure 3: Gene enrichment results for 641 genes with significant (FDR<0.05) differentially variable expression levels between mild and severe samples, as measured using the F-test.

Supplemental Figure 4: PCoA on microbiome data. Scatterplots from principal coordinate analysis (PCoA), based on Euclidean, rJSD, unweighted Unifrac, weighted Unifrac distances of the OTUs in nasal microbiome samples from infants with RSV infection are shown and points colored to indicate the corresponding sample's severity, sex, enrollment season, race, ethnicity, and visit number.

Supplemental Figure 5: Sensitivity analysis on single-sample network edges. We plot the percentage of times an edge is identified as significant out of 1000 iterations of *limma* analysis comparing random subsets with 10 mild and 10 severe samples, corrected for sex and race. Color corresponds to direction of fold change for significant edges. The line at 5% indicates the percent occurrence expected by chance.

Supplemental Figure 6: GO term enrichment analysis for the top 100 genes based on the loadings of (A) gPC13, (B) gPC12 and (C) gPC8. The top 25 most significant GO terms, based on their FDR, are shown (all terms meet a significance level of FDR<0.05) (D-H) Plots showing all significant GO terms (FDR<0.05) identified for (D) gPC1, (E) gPC8, (F) gPC3, (G) gPC13, and (H) gPC12. Size of each bubble indicates the number of genes annotated to the respective GO term and the color indicates the percentage of the top 100 genes annotated to that term.

Nasal Wash Titer



Nasal Swab Titer

