* Hello and thank you for allowing me to present our ongoing research on behalf of the COVIRT microbial subgroup. I have no conflicts of interest of financial disclosures. **[NEXT SLIDE]**
* This study began early in the SARS-CoV-2 outbreak, when scientists began openly publishing metatranscriptome sequences from BALF samples of patients with COVID-19 disease, prompting us to investigate **microbially derived transcriptomic changes surrounding COVID-19 moderate to severe disease and progression**, despite limitations in experimental study design.
* This study began early in the SARS-CoV-2 outbreak, \*\*\* when scientists began openly publishing metatranscriptome sequences derived from human bronchoalveolar lavage fluid (BALF) from patients with COVID-19 disease, prompting us to investigate the potential relationship between **microbially derived transcriptomic changes surrounding COVID-19 moderate to severe disease and progression** from a hologenome standpoint. We began our analysis using metatranscriptomes sample sequences sourced from 8 different publications made available from public repositories. \*\*\* Samples in this analysis comprised 3 main cohorts categorized by case type, consisting of an uninfected control cohort, a community acquired pneumonia or CAP patient cohort, and a COVID-19 patient cohort \*\*\* with and additional stratified comparison based of disease severity in the COVID-19 cohort broken down by survival outcome. \*\*\* The objectives of the study were to compare the BALF metatranscriptomes in the **COVID19 cohort** amongst **uninfected** and **CAP** patient cohorts and identify \*\*\* taxonomic changes in **microbial** derived **community dynamics** and functional changes derived from **gene ontologies** associated with **COVID19 moderate to severe disease or its treatment \*\*\* as well as identify predictors of disease outcome in the COVID19** cohort, \*\*\* With the overarching hypothesis that there is a potential informative relationship between the BALF microbiome and the severity of COVID-19 disease onset, progression, and outcome. **[NEXT SLIDE]**
* The bioinformatic analysis pipeline began by downloading Raw sequencing reads download from either SRA or CRA followed by Read preprocessing, which consisted of adapter trimming,QA/QC, and filtering of human and low complexity reads using Trimmomatic, fastQC, Kraken2, FastP, followed by \*\*\*
* Taxonomic classification via Kraken2 \*\*\* and functional annotation via Seqscreen. \*\*\* Taxonomic classifications were decontaminated against negative controls when applicable using the library decontam in R, followed by statistical analysis and visualization using the bioinformatic software package phyloseq , vegan, and metacoder in R.
* Functional annotation gene ontologies counts derived from the seqscreen outputs were parent propagated \*\*\* using the covirt-micro conda package created by Mike Lee at NASA's Ames Research Center. Parent propagated Gene ontology terms counts \*\*\* were then subjected to a Dirichlet mixture modeling (or Dmm clustering) and \*\*\* comparative analysis using multivariable associations with linear models Maaslin2 in R. \*\*\*
* All Scripts, intermediate files, and results from this analysis can be found at either the covert microbial subgroup OSF website, the covirt-micro githubs repos, or the covert microbial conda package. **[NEXT SLIDE]**
* After read filtering and batch effect sample removal, sample cohorts consisted of 29 Uninfected sampled, 25 CAP samples, and 32 COVID19 samples, bringing the total n to 86. \*\*\* Amongst the COVID19 cohort with known survival outcomes, 10 were deceased and 15 were survived. **[NEXT SLIDE]**
* Results from the Maaslin2 Multivariable Associations with Linear Models comparison of COVID19 moderate to severe disease cohort to the Uninfected and community acquired viral pneumonia cohort revealed 35 out of 13,534 Significant Gene ontologies using a pvalue cutoff of 0.05 with Benjamini hochberq multiple test correction and controlling for random effects of publication and patient. Significant GO Terms of interest were comprised of \*\*\* 6 Depth 1 Parents involving:\*\*\* catalytic activity, \*\*\*Ion binding, \*\*\* metabolic and cellular processes, biological regulation, \*\*\*and interspecies interaction between organisms. **[NEXT SLIDE]**
* Of notable interest, GO Terms associated the Depth 1 Parent include catalytic activity \*\*\* include hydrolase activity, transferase activity, transferase activity transferring phosphorus, and nucleotidyltransferase activity. Other Significant Terms of interest that detail the metatranscriptomes profile of COVID19 include \*\*\* Ion binding, \*\*\* nitrogen /organonitrogen compound metabolic processes,\*\*\* and the viral modulation of host cellular processes interactions. **[NEXT SLIDE]**
* Additionally, Results from the Dirichlet Multinomial Mixtures clustering analysis using all 13,534 Gene ontologies counts resulted [\*\*\* \*\*\*] in a best model fit of 3 distinct dmm clusters that were significantly associated [\*\*\* \*\*\*] with case type and disease outcome with a p value of <0.0001. **[NEXT SLIDE]**
* A Taxonomic comparative analysis of the kraken2 derived taxonomy counts revealed a statistically significant decrease in log2 median ration of several species \*\*\* belonging to the genus *Sphingomonas* when compared to both the uninfected (p<0.0001, q <0.001) and Community acquired pneumonia cohorts (p<0.005,q <0.05) cohorts (table X). This finding supports previous reports regarding an association with *Sphingomonas*, which is commonly known as an opportunistic pathogen found in healthcare-associated pneumonia. Sig. GO Terms derived from *Sphingomonas* proteins in the COVID19 samples were hydrogen peroxide catabolic process, response to oxidative stress, catalase activity, heme binding and metal ion binding.
* The catalase protein decomposes hydrogen peroxide into water and oxygen; serves to protect cells from the toxic effects of hydrogen peroxide. So as a discussion point perhaps Sphingomonas is responding to COVID-19 conditions in the patient by expressing genes that help it to survive well under the COVID-19 disease conditions. **[NEXT SLIDE]**
* We then conducted a stratified Maaslin2 analysis amongst COVID19 samples with known survival outcomes analysis revealing notable functional profiles associated Phosphate / phosphorylation, Metal ion binding (mg,zn,etc) Nucleotide terms (DNA/RNA) Lytic activity (hydrolase, endopeptidase,etc). **[NEXT SLIDE]**
* These findings were similar in nature to the GO Terms associated with COVID19 versus the uninfected and community acquired pneumonia patient cohorts **[NEXT SLIDE]**
* Taxonomic analysis revealed a statistically significant increase in log2 median ratio of the family Comanomonadacea, belonging to the genus *Variovorax* \*\*\*and decreases in the family Bacteriodales when comparing the deceased to the survive (p<0.0001, q <0.001) (table X).
* In conclusion,\*\*\* we observed unique and taxonomic and functional discriminant features in the brochoalveolar lavage metatranscriptomes distinctive of COVID19 moderate to serve disease or its treatment, and predictors of COVID19 mortality. \*\*\*
* Functionally annotated Gene ontologies of interest included associated with:Phosphate / phosphorylation, metal ion binding (mg,zn,etc), nucleotide terms (DNA/RNA), Lytic activity (hydrolase, endopeptidase,etc) \*\*\*
* Distinct Taxonomic features of COVID19 disease and mortality include increases in log2 median ratios of genera Sphingomonas and Variovorax belonging to the Sphingomonadacae and Comonomonadacea families, and decreases in the class Bacteroidia belonging to the order Bacteroidiales.
* Collectively, while this data does cannot speak to causality or directionality of the association, it does demonstrate a significant relationship between the human microbiome and severity of COVID-19, rendering further testable hypotheses that warrant further investigation. **[NEXT SLIDE]**
* With that I would like to thank the COVIRT microbial subgroup team members and give special acknowledgment to John Fonner and the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing HPC resources that have contributed to the research results reported. I would also like to thank my PI Dr. Kjersti Aagaard, Mike Lee at NASA, viktoria zakas at the University of Chicago, Kristen Curry and Dr. Todd Trainjen, and the entirety of the Trainjen lab at Rice, and finally the COVIRT microbial subgroup leader Dr. Krista Ternus at Signature science.
* And with that I would like to thank you for listening and am happy to take any questions.