O430
1-hour Oral Session
Rising the challenge: novel techniques in microbiology

A novel gene expression diagnostic can robustly distinguish viral from bacterial infections

Timothy Sweeney*[1], Hector Wong[2], Jason Andrews[3], David Spain[4], Purvesh Khatri[5]

[1]*Stanford University, Institute for Immunity, Transplantation and Infections, Stanford, United States*

[2]*Cincinnati Children's Hospital Medical Center , Division of Critical Care Medicine, Cincinnati, Oh, United States*

[3]*Stanford University, Department of Medicine, Stanford, Ca, United States*

[4]*Stanford University, Department of Surgery, Stanford, Ca, United States*

[5]*Stanford University, Institute for Immunity, Transplantation and Infections, Stanford, Ca, United States*

Background: Distinguishing viral from bacterial infections can be difficult. Inappropriate diagnoses lead to increased morbidity, mortality, and antimicrobial resistance. Bacterial cultures can take 24-72 hours to return results, and most viral diagnostics are specific to only one virus. Distinguishing bacterial from viral infections based on host response may be a more effective approach. Here we studied gene expression in viral and bacterial infections using a multi-cohort analysis framework with which we have previously identified diagnostic gene sets in transplant rejection, pulmonary tuberculosis, and sepsis.

Material/methods: We performed a systematic search for public genome-wide expression studies of viral and bacterial infection. All microarray data were re-normalized and log2 transformed. We applied our multi-cohort meta-analysis framework to find statistically differentially expressed genes. A greedy forward search was used to find a diagnostic gene set. The resulting gene set was validated in multiple public gene expression datasets, and in independent patients using a targeted nanoString digital multiplex gene quantitation assay.

Results: We identified 7 datasets (whole blood and PBMCs) composed of 367 patients (101 viral and 266 bacterial infections, including children and adults, medical and surgical patients, and with respiratory and systemic viral infections). Multi-cohort analysis identified 72 genes significantly differentially expressed (FDR<1% and effect size>2 fold) between patients with viral versus bacterial infections. A forward search identified a subset of 7 genes which robustly distinguished viral from bacterial infections in the discovery datasets (summary ROC AUC=0.96, 95% CI 0.88-0.99, Figure 1A). The gene set was directly validated in 4 independent microarray datasets (summary ROC AUC=0.93, 95% CI 0.64-0.99, Figure 1B). We further showed a global ROC AUC=0.90 in 17 additional public microarray datasets that studied either bacterial or viral infections, but not both. Finally, we used targeted nanoString gene expression assays to validate these results in independent whole blood samples from children with sepsis (AUC 0.85 for pediatric sepsis in discovery, AUC 0.86 for pediatric sepsis in validation).

Conclusions: A parsimonious set of 7 genes can robustly differentiate viral from bacterial infections in 7 discovery datasets, and was validated using 21 clinically heterogeneous public microarray datasets,

as well as a targeted nanoString gene expression assay. Prospective validation of a rapid assay will be necessary prior to clinical implementation.