Efficacy of Parallel Assessment of Host and Pathogen RNA-seq from an In Vitro Cell Infection Assay

Natalie E. Underwood [1], Tina Z. Akondeng-Azise [2], Stuart Cantlay [3], Nicole L. Garrison [3]

[1] Department of Natural Sciences, Fairmont State University; [2] Department of Biology, Shepherd University; [3] Department of Biomedical Sciences, West Liberty University

Abstract

To identify changes in gene expression of host cells in the presence of culturable and VBNC (viable but nonculturable) *Francisella tularensis* (subspecies *holarctica* LVS), RNA was extracted from RAW 264.7 mouse macrophages exposed to either form of the bacteria. Cells were sampled at three time points for each treatment; prior to exposure (T0), after 4 hours of exposure (T4), and after 8 hours (T8). The bioinformatic challenge addressed in this project was to evaluate the efficacy of using this type of sequencing to discover transcriptional activity of *F. tularensis* from a small proportion of total RNA reads, and to assess the baseline similarity of macrophage expression.. Few reads mapped to the bacterial genome, those that were recovered aligned to the three 16S/23S genomic islands present in the *F. tularensis* genome.

Background

Little is known about the mechanism of persistence outside of host cells of *F. tularensis*, an intracellular pathogen that has been categorized as a Tier 1 select agent due to its high virulence and capacity to cause severe disease. Under laboratory conditions, it spontaneously enters a Viable But Non-Culturable (VBNC) state, which may have important implications for pathogenicity and persistence. The aim of the in vitro cell infection assay analyzed here was to identify whole transcriptome changes in gene expression of host macrophage cells exposed to VBNC *F. tularensis*; however, the bioinformatic pipeline presented here was designed to 1) assess the possibility of analyzing bacterial expression using sequences generated primarily from host cells and to 2) evaluate baseline expression of host macrophages.

Methods

Quality Control: fastqc/multiqc

↓

Alignment: Hisat2, *Mus musculus*

↓

Filter and Sort: samtools

↓

Alignment: BWA-MEM, *F. tularensis*

↓

Quantify Reads: bash script

↓

Visualize Reads: IGV

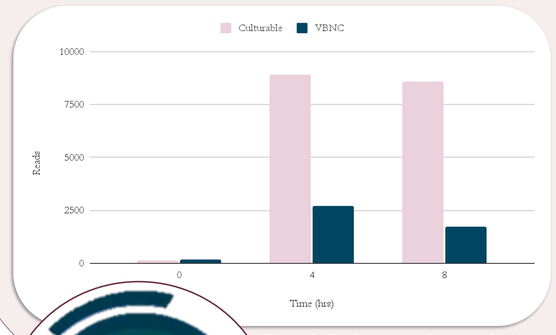

Figure 2: *F. tularensis* reads mapped at three time points for both treatments.

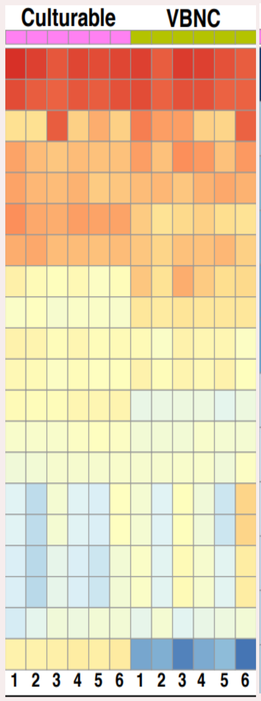

Figure 3: Top 20 significant genes heat map

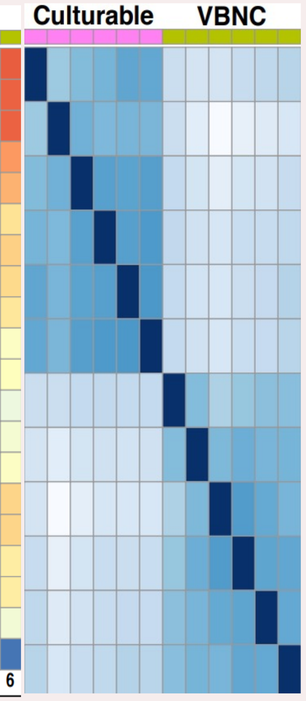

Figure 4: Heat map of sample to sample count distances

Results

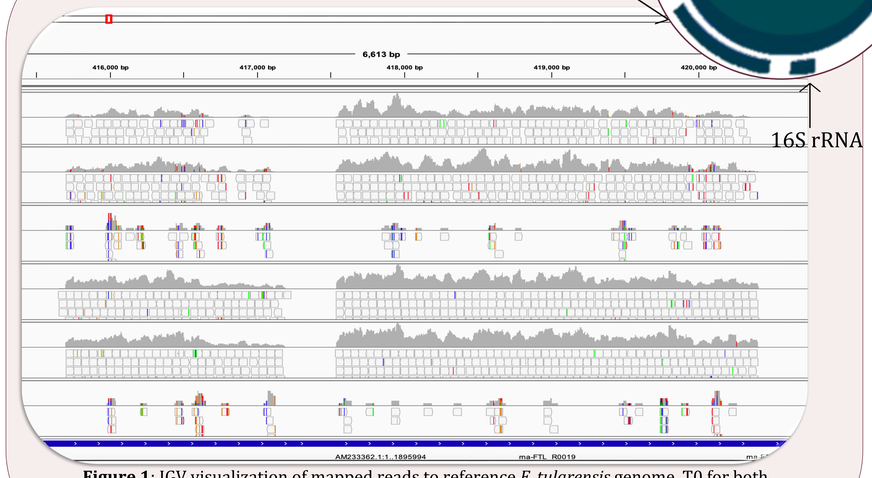

Figure 1: IGV visualization of mapped reads to reference *F. tularensis* genome. T0 for both treatment groups had very few and possibly erroneous mappings. T4 and T8 for both groups had the most reads, with Culturable samples generating more reads total.

Results/Conclusion

Very few reads were recovered that originated from the bacterial pathogen *F. tularensis* (Fig. 2), most reads (>94%) in each sample mapped to the mouse reference genome. *F. tularensis* reads that were detected mapped primarily to the three 16S/23S rRNA genomic islands present in the bacterial reference genome (Fig. 1). Overall patterns of expression for the first time point (T0) of mouse macrophages appears similar in direction and magnitude when examining the top 20 most significant genes (Fig. 3); however, there were broader patterns of clustering by treatment group seen when looking at the pairwise count comparisons (Fig. 4). The next step in analyzing this data on gene expression of host cells in the presence of culturable and VBNC *F. tularensis* will be to 1) identify differentially expressed gene identities for the T0 replicates and 2) compare additional time points for each treatment group using T0 as the reference condition.

Acknowledgments: Supported by NIH Grant P20GM103434 to the West Virginia IDeA Network for Biomedical Research Excellence and and WV-CTSI NIGMS.NIH U54GM10