Reads were downloaded in gzipped fastq format from Illumina’s BaseSpace via their CLI to scratch space on Amarel. BaseSpace log files show that the basecall files were demultiplexed and adapter trimmed to produce the fastq files. Furthermore, adapter and end trimming are not needed for our purposes (Liao & Shi 2020), but the cleanup done on BaseSpace will not adversely impact downstream analyses. Each pair of files was processed by process\_reads.sh. The Jupyter notebook RNAseq0\_py0.ipynb reads the index of files and traverses the download on Amarel to produce the pipeline scripts; these pipeline scripts ultimately call process\_reads.sh. Quality analysis is performed on the reads via fastqc (Andrews 2010). Reads are then aligned to the appropriate genome with bowtie2 (Langmead & Salzberg 2012). Samples from *Mycobacterium tuberculosis* were aligned to the genome and annotation set produced and used in (Chitale et al. 2022), and samples from *Escherichia coli* were aligned to the MG1655 reference (Hayashi et al. 2006). Counts of aligned reads were compiled with featureCounts (Liao et al. 2014). Performance logs and runtime diagnostics were aggregated with multiqc (Ewels et al. 2016). All samples passed quality control. Read pair alignment rates were generally good (>90%), as were rates of assignment of read pairs to features (>85%). Counts of reads were compiled by organism and sorted by sample number in the Jupyter notebook RNAseq0\_py1.ipynb. For each sample, transcripts per million (tpm) as well as log2(tpm) were calculated and saved.

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data.

Chitale, P., Lemenze, A. D., Fogarty, E. C., Shah, A., Grady, C. A., Odom-Mabey, A. R., ... & Alland, D. (2022). A comprehensive update to the Mycobacterium tuberculosis H37Rv reference genome. *bioRxiv*.

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, *32*(19), 3047-3048.

Hayashi, K., Morooka, N., Yamamoto, Y., Fujita, K., Isono, K., Choi, S., ... & Horiuchi, T. (2006). Highly accurate genome sequences of Escherichia coli K‐12 strains MG1655 and W3110. *Molecular systems biology*, *2*(1), 2006-0007.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods*, *9*(4), 357-359.

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*(7), 923-930.

Liao, Y., & Shi, W. (2020). Read trimming is not required for mapping and quantification of RNA-seq reads at the gene level. *NAR genomics and bioinformatics*, *2*(3), lqaa068.