Library Preparation and DNA Sequencing:

150 to 300 ng of total RNA determined by Invitrogen™ Qubit™ high-sensitivity spectrofluorometric measurement was poly-A selected and reverse transcribed using Illumina's TruSeq® stranded mRNA library preparation kit. Each sample was fitted with one of 96 adapters containing a different 8 base molecular barcode for high level multiplexing. After 15 cycles of PCR amplification, completed libraries were sequenced on an Illumina NovaSeq™ 6000, generating 20 million or more high quality 100 base long paired end reads per sample.

RNA-Seq Analysis:

A quality control check on the fastq files was performed using FastQC. Upon passing basic quality metrics, the reads were trimmed to remove adapters and low-quality reads using default parameters in Trimmomatic[1] [Version 0.33].

Alignment, Transcript Abundance and Differential Gene Expression Analysis:

The trimmed reads were then mapped to a reference genome using default parameters with strandness (R for single-end and RF for paired-end) option in Hisat2[2] [Version 2.0.5]. In the next step, transcript/gene abundance was determined using kallisto[3] [Version 0.43.1]. We first created a transcriptome index in kallisto using Ensembl cDNA sequences for the reference genome. This index was then used to quantify transcript abundance in raw counts and transcript per million (TPM).

1. Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics*, btu170.

2. Kim D, Langmead B and Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nature Methods 2015

3. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification, Nature Biotechnology 34, 525-527(2016), doi:10.1038/nbt.3519

Version 3.0 – March 2020