NKI: expression - Illumina HiSeq 2000, RNA-Seq

Analysis Description:

The reads were trimmed using Cutadapt (Martin, 2011) to remove any remaining adapter sequences, filtering reads shorter than 20 bp after trimming to ensure good mappability. The trimmed reads were aligned to the GRCh38 reference genome using STAR (version 2.5.2b; Dobin et al., 2013). Mouse reads were filtered out by AstraZeneca's tool disambiguate (Ahdesmäki et al., 2016). QC statistics from Fastqc (Andrews S., 2010). FastQC: a quality control tool for high throughput sequence data. and the above-mentioned tools were collected and summarized using Multiqc (Ewels et al., 2016). Gene expression counts were generated by featureCounts (Liao et al., 2014) using gene definitions from Ensembl GRCh38 version 89. This pipeline was created by Julian de Ruiter and is available at: https://github.com/jrderuiter/snakemake-rnaseq. Normalized expression values were obtained by correcting for differences in sequencing depth between samples using DESeqs median-of-ratios approach (Anders and Huber, 2010) and then log-transforming the normalized counts