### Code for quality control, genome alignment, and read counts for murine PDAC cell line RNA-seq data

# Packages used: FastQC (v0.11.5), MultiQC (v1.9), RSEM (v1.3.3)

### Run FastQC and MultiQC on FASTQ files for quality control

cd /home/RNAseq\_tumors/raw\_fastq/

fastqc\_0.11.5/FastQC/fastqc \*fastq.gz

multiqc .

### Generate an RSEM reference file with the GENCODE *Mus musculus* M27 genome assembly (downloaded from https://www.gencodegenes.org/mouse/release\_M27.html)

conda activate RNAseq\_STAR

cd /home/RNAseq\_tumors/M27\_reference/

rsem-prepare-reference --gtf gencode.vM27.primary\_assembly.annotation.gtf --star GRCm39.primary\_assembly.genome.fa RSEM\_M27index

### Align reads to mouse genome reference (using STAR aligner option) and count features with RSEM

cd /home/RNAseq\_tumors/raw\_fastq/

# Unzip fastq.gz files

gunzip \*fastq.gz

# Align files to RSEM index with STAR and quantify

list1=( $(find . -name "\*R1\_001.fastq" | sort))

list2=( $(find . -name "\*R2\_001.fastq" | sort))

for i in {0..17};

do

rsem-calculate-expression --star --num-threads 10 --paired-end --estimate-rspd --append-names ${list1[i]} ${list2[i]} /home/RSEM\_M27index/RSEM\_M27index /home/RNAseq\_tumors/RSEM\_align\_quant/${list1[i]}\_align\_quant

done

conda deactivate