The fastq to counts conversion was done using the *indropIndex* and the *indropCounts* functions which are part of the rCASC framework (Alessandri et al. Gigascience 2019 Sep 1;8(9):giz105. PMID: 31494672).

[indropIndex](https://kendomaniac.github.io/rCASC/reference/indropIndex.html)(group="docker", index.folder=[getwd](https://rdrr.io/r/base/getwd.html)(),

ensembl.urlgenome="ftp.ensembl.org/pub/release-97/fasta/mus\_musculus/dna/Mus\_musculus.GRCm38.dna.primary\_assembly.fa.gz",

ensembl.urlgtf="ftp.ensembl.org/pub/release-97/gtf/mus\_musculus/Mus\_musculus.GRCm38.85.gtf.gz")

The mouse genome assembly used for the index file was GRCm38 (mm10). Genome fasta file was retrieved form ENSEMBL ftp site, and annotation was done using ENSEMBL GTF version 85, also retrieved from the ENSEMBL ftp site.

[indropCounts](https://kendomaniac.github.io/rCASC/reference/indropCounts.html)(group="docker", scratch.folder="/data/scratch", fastq.folder=[getwd](https://rdrr.io/r/base/getwd.html)(), index.folder="/data/genomes/indropMm10",

sample.name="C1", split.affixes="S0\_L001", bowtie.index.prefix="genome",

M=10, U=2, D=400, low.complexity.mask="False")

The M parameter indicates to ignore reads with more than M alignments, after filtering on distance from transcript end; U parameter indicates to ignore counts from UMI that should be split among more than U genes; D parameter indicates the maximal distance from transcript end expressed in nucleotides; low.complexity.mask parameter indicates that low complexity regions are masked.

inDrop pipeline version 20170126

Bowtie version 1.1.1.1

Samtools version 1.3.1

RSEM version 1.3.0

Java version 1.8.0