Pipeline overview

RAMPAGE (RNA Annotation and Mapping of Promoters for the Analysis of Gene Expression) is a very accurate sequencing approach to identify transcription start sites (TSSs) at base-pair resolution, the quantification of their expression and the characterization of their transcripts. This assay uses direct cDNA evidence to link specific genes and their regulatory TSSs.

The ENCODE RAMPAGE pipeline can be used for libraries generated using rRNA-depleted total RNA >200 nucleotides in size. The pipeline inputs paired-end, gzipped, RNA-sequencing reads (fastqs) and a STAR genome index tar.gz file and produces a bam file of aligned reads. These bams are then processed to generate:

- Minus/plus strand signal files of unique reads (bigWig) for easy display on a genome browswer
- Minus/plus strand signal files of all reads (bigWig)
- Transcription start sites available in the following formats: bed, bigBed, gff)
- Irreproducible Discovery Rate (IDR) files: compares two peak files (bed), typically from a pair of replicates of the same experiment, allowing validation of the experiment methods and reducing noise in the final results.
- Quality control files (png): compares two TSS quantification files and calculates the Mean Absolute Deviation and correlations.

ENCODE2 CAGE (Cap Analysis Gene Expression) data is also processed using the RAMPAGE pipeline.

Genomic References Used in this Pipeline

These pipelines require both assembly information for the species of interest and a gene reference. Each of the main programs, TopHat, STAR, and RSEM create an index for use in subsequent steps. The current reference files and indexes can be found on this site:

- STAR Indexes: https://www.encodeproject.org/references/ENCSR314WMD/
- TopHat Indexes: https://www.encodeproject.org/references/ENCSR641UDW/
- RSEM Indexes: https://www.encodeproject.org/references/ENCSR219BJA/
- Unmodified Genome References and Chromosome Sizes: https://www.encodeproject.org/references/ENCSR425FOI/
- GENCODE References: https://www.encodeproject.org/references/ENCSR884DHJ/

References

Batut, Philippe et al. "High-fidelity promoter profiling reveals widespread alternative promoter usage and transposon-driven developmental gene expression." *Genome Research* 23.1 (2013): 169–180. *PMC*. Web. 9 Feb. 2016. PMCID: PMC3530677