Salmon.Alevin.Quant Documentation

Description: Quantification and analysis of 3' tagged-end single-cell sequencing data using Salmon's Alevin (v1.5.2) function. See: The Alevin Documentation for detailed usage guidelines.

Author: Avi Srivastava, COMBINE Lab, Department of Computer Science, University of Maryland. Wrapped as a module by Anthony S. Castanza, Mesirov Lab, UCSD School of Medicine.

Contact: Algorithm and scientific questions: https://github.com/COMBINE-lab/salmon/issues. Module specific issues: genepattern.org/help

Summary: Quantify gene expression from single-cell sequencing experiments using the Alevin module of the pseudo-alignment based method “Salmon”.

Source Publication: Srivastava A, Malik L, Smith T, Sudbery I, Patro R. Alevin efficiently estimates accurate gene abundances from dscRNA-seq data. Genome Biol. 2019 Mar 27;20(1):65. doi: 10.1186/s13059-019-1670-y. PMID: 30917859; PMCID: PMC6437997.

Basic Parameters:

Name	Description
Barcodes*	Cellular barcodes+UMI file(s), the FASTQ file containing CB+UMI raw sequences. Alevin also supports parsing of data from multiple files as long as the order is the same as in the "Reads" parameter. Barcodes and reads should share the same base file name in order to ensure order matching, barcodes should have the suffix _R1 (SampleID_R1.fastq.gz). File should be gzipped.
Reads*	The FASTQ file containing raw read-sequences. Alevin also supports parsing of data from multiple files as long as the order is the same as in the "Barcodes" parameter. Barcodes and reads should share the same base file name in order to ensure order matching, barcodes should have the suffix _R2 (SampleID_R1.fastq.gz) File should be gzipped.
Transcriptome Index*	The indexed transcriptome output from the Salmon.Indexer module (or comparable pipeline). File must be .tar.gz
Chemistry*	The chemistry used by the single-cell sequencing platform. Drop-Seq, 10x Chromium v2, 10x Chromium v3, CITE-Seq, CEL-Seq, CEL-Seq2, and Quartz-Seq2 are currently supported.
Library Type*	The orientation of the Barcodes and Reads. The Alevin authors recommend using ISR (default) for both Drop-seq and Chromium chemistry.
TgMap*	A two column tsv (tab-separated) file with no header containing the transcript to gene map file. The first column lists each transcript present in the reference, the second column lists the corresponding gene. Alternatively, a .GTF file can be supplied to

	automatically create a transcript to gene map.
Output Basename*	The base name to use for naming the alevin results file (default: alevin.output)

*required

Advanced Parameters: See: The Alevin Documentation for details on additional parameters offered in this module.

Output Files:

Name	Description
<alevin.output>.tar.gz	Gzipped files containing the alevin quantification results.

Module Language: Shell script
Source Repository: https://github.com/genepattern/Salmon.Alevin.Quant/releases/tag/v0.6
Docker image: genepattern/salmon-alevin-quant:beta

Version	Comment
0.6	Initial beta release.