

# Assembly Transcripts with Cufflinks Element

Cufflinks accept aligned RNA-Seq reads and assemble the alignments into a parsimonious set of transcripts. Cufflinks then estimate the relative abundances of these transcripts based on how many reads support each one, taking into account biases in library preparation protocols.

**Element type:** cufflinks

## Parameters

| Parameter                   | Description  | Default value     | Parameter in Workflow File  | Type           |
|-----------------------------|--|-------------------|-----------------------------|----------------|
| <b>Output directory</b>     | Directory to save MACS output files.   |                   | <b>out-dir</b>              | <i>string</i>  |
| <b>Reference annotation</b> | Tells Cufflinks to use the supplied reference annotation to estimate isoform expression. Cufflinks will not assemble novel transcripts and the program will ignore alignments not structurally compatible with any reference transcript.   |                   | <b>ref-annotation</b>       | <i>string</i>  |
| <b>RABT annotation</b>      | Tells Cufflinks to use the supplied reference annotation to guide Reference Annotation Based Transcript (RABT) assembly. Reference transcripts will be tiled with faux-reads to provide additional information in an assembly. The output will include all reference transcripts as well as any novel genes and isoforms that are assembled.   |                   | <b>rabt-annotation</b>      | <i>string</i>  |
| <b>Library type</b>         | Specifies RNA-Seq protocol.  | Standart Illumina | <b>library-type</b>         | <i>numeric</i> |
| <b>Mask file</b>            | Ignore all reads that could have come from transcripts in this file. It is recommended to include any annotated rRNA, mitochondrial transcripts other abundant transcripts you wish to ignore in your analysis in this file. Due to variable efficiency of mRNA enrichment methods and rRNA depletion kits, masking these transcripts often improves the overall robustness of transcript abundance estimates.   |                   | <b>mask-file</b>            | <i>string</i>  |
| <b>Multi-read correct</b>   | Tells Cufflinks to do an initial estimation procedure to more accurately weight reads mapping to multiple locations in the genome.   | False             | <b>multi-read-correct</b>   | <i>boolean</i> |
| <b>Min isoform fraction</b> | After calculating isoform abundance for a gene, Cufflinks filters out transcripts that it believes are very low abundance, because isoforms expressed at extremely low levels often cannot reliably be assembled, and may even be artifacts of incompletely spliced precursors of processed transcripts. This parameter is also used to filter out introns that have far fewer spliced alignments supporting them.   | 0.1               | <b>min-isoform-fraction</b> | <i>numeric</i> |
| <b>Frag bias correct</b>    | Providing Cufflinks with a multifasta file via this option instructs it to run the bias detection and correction algorithm which can significantly improve the accuracy of transcript abundance estimates.   |                   | <b>frag-bias-correct</b>    | <i>string</i>  |
| <b>Pre-mRNA fraction</b>    | Some RNA-Seq protocols produce a significant amount of reads that originate from incompletely spliced transcripts, and these reads can confound the assembly of fully spliced mRNAs. Cufflinks use this parameter to filter out alignments that lie within the intronic intervals implied by the spliced alignments. The minimum depth of coverage in the intronic region covered by the alignment is divided by the number of spliced reads, and if the result is lower than this parameter value, the intronic alignments are ignored. | 0.15              | <b>pre-mrna-fraction</b>    | <i>numeric</i> |
| <b>Cufflinks tool path</b>  | The path to the Cufflinks external tool in UGENE.  | default           | <b>path</b>                 | <i>string</i>  |
| <b>Temporary directory</b>  | The directory for temporary files.   | default           | <b>tmp-dir</b>              | <i>string</i>  |

## Input/Output Ports

The element has 1 *input port*:

**Name in GUI:** Input reads

**Name in Workflow File:** in-assembly

**Slots:**

| Slot In GUI          | Slot in Workflow File | Type            |
|----------------------|-----------------------|-----------------|
| <b>Assembly data</b> | assembly              | <i>assembly</i> |
| <b>Source url</b>    | url                   | <i>string</i>   |

And 1 *output port*:

**Name in GUI:** Output annotations

**Name in Workflow File:** out-annotations

**Slots:**

| Slot In GUI                     | Slot in Workflow File | Type             |
|---------------------------------|-----------------------|------------------|
| Isoform-level expression values | isolevel.slot         | <i>ann_table</i> |