

Galaxy Workflow ' transcripts-assembly-pe-dataset-collection'

Step

Annotation

Step 1: Input dataset

V_corymbosum_scaffold_May_2013.fa
select at runtime

Step 2: Input dataset collection

Input Dataset Collection
select at runtime

Step 3: HISAT2

Input data format

FASTQ

Single end or paired reads?

Collection of paired reads

Paired reads

Output dataset 'output' from step 2

Paired-end options

Use default values

Write unaligned reads (in fastq format) to separate file(s)

False

Write aligned reads (in fastq format) to separate file(s)

False

Source for the reference genome to align against

Use a genome from history

Select the reference genome

Output dataset 'output' from step 1

Primary alignments

Not available.

Maximum number of seeds that will be extended

Not available.

Report secondary alignments

False

Alignment options

Use default values

Input options

Use default values

Scoring options

Use default values

Spliced alignment parameters

Use default values

Paired alignment parameters

Use default values

Step 4: Flagstat

BAM File to report statistics of

Output dataset 'output_alignments' from step 3

Step 5: StringTie

Mapped reads to assemble transcripts from

Output dataset 'output_alignments' from step 3

Use GFF file to guide assembly

Do not use GFF/GTF

Options

Use defaults

Step 6: StringTie merge

Transcripts

Output dataset 'output_gtf' from step 5

Reference annotation to include in the merging

select at runtime

Minimum input transcript length to include in the merge

50

Minimum input transcript coverage to include in the merge

0

Minimum input transcript FPKM to include in the merge

1.0

Minimum input transcript TPM to include in the merge

1.0

Minimum isoform fraction

0.01

Gap between transcripts to merge together

250

Keep merged transcripts with retained introns

False