

Raw RNA-Seq Processing

⚠ Download and install the UGENE [FULL](#) or [NGS package](#) to use this pipeline.

Use this workflow sample to process raw RNA-seq next-generation sequencing (NGS) data from the Illumina platform. The processing includes:

- *Filtration:*
 - Filtering of the NGS short reads by the CASAVA 1.8 header;
 - Trimming of the short reads by quality;
- *[Optionally] Mapping:*
 - Mapping of the short reads to the specified reference sequence (the TopHat tool is used in the sample);

The result output of the workflow contains the filtered and merged FASTQ files. In case the TopHat mapping has been done, the result also contains the TopHat output files: the accepted hits BAM file and tracks of junctions, insertions and deletions in BED format. Other intermediate data files are also output by the workflow.

✓ How to Use This Sample

If you haven't used the workflow samples in UGENE before, look at the ["How to Use Sample Workflows"](#) section of the documentation.

✓ What's Next?

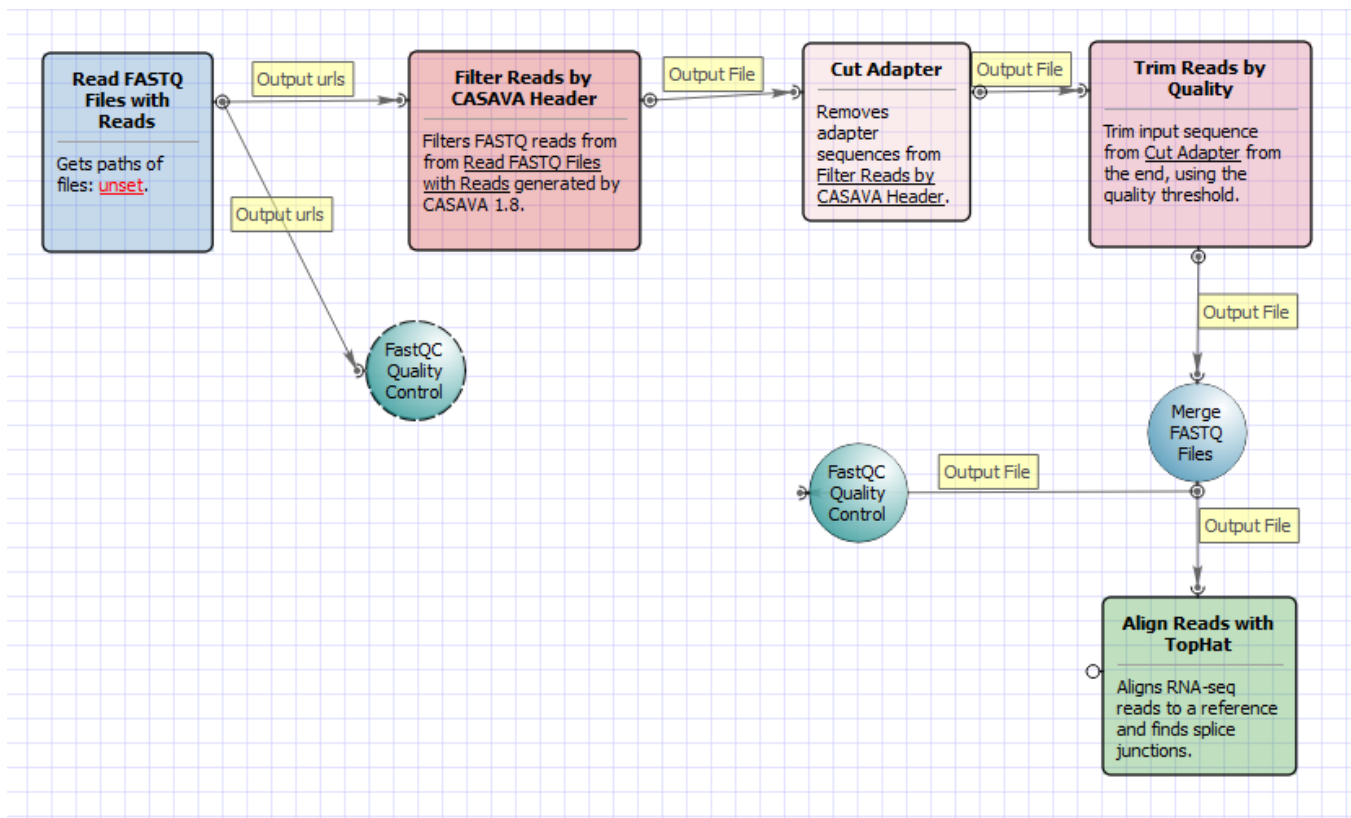
The [Tuxedo workflow](#) can be used to analyze the filtered RNA-seq data. In this case the mapping step of this workflow can be skipped, as it also present in the Tuxedo pipeline.

Workflow Sample Location

The workflow sample "Raw DNA-Seq processing" can be found in the "NGS" section of the Workflow Designer samples.

Workflow Image

There are four versions of the workflow available. The workflow with mapping for single-end reads looks as follows:



The workflow with mapping for paired-end short appearance is the following:

Raw RNA-Seq Processing Wizard

Pre-processing

Filtration

Quality threshold: 20

Min length: 10

Adapters: ork/ugene/data/adapters/adapters.fasta

Buttons: Defaults, < Back, Next >, Cancel

3. Mapping: On this page you must input reference and optionally modify advanced parameters.

Raw RNA-Seq Processing Wizard

Mapping

TopHat input

Bowtie index directory: Required

Bowtie index basename: Required

Bowtie version: Bowtie1

Buttons: Select bowtie index file

Parameters

Known transcript file: ...

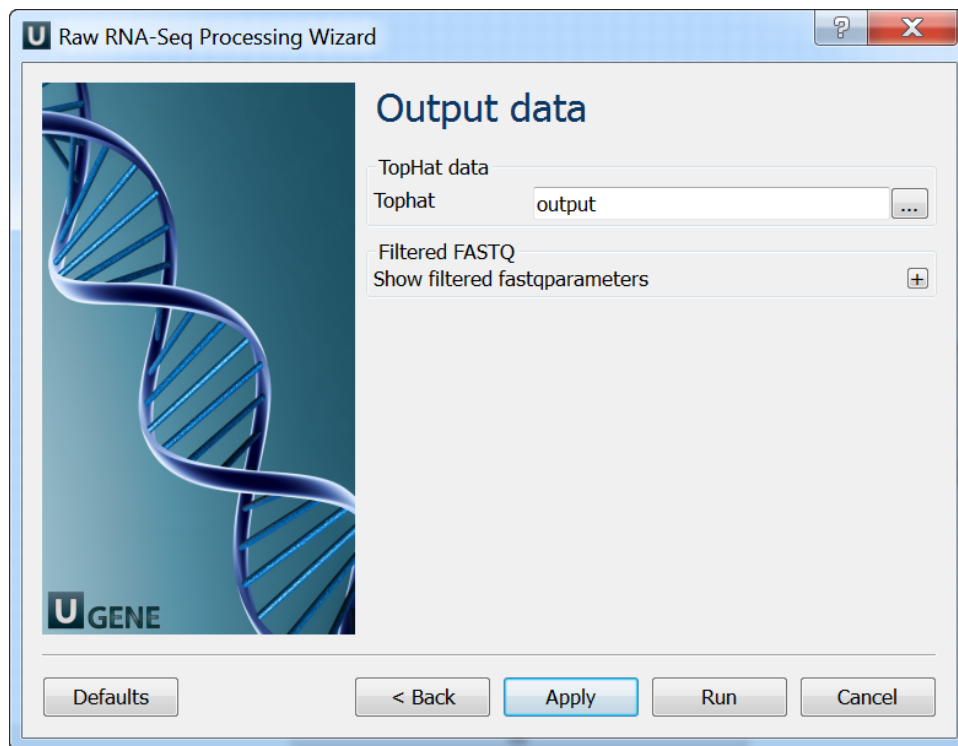
Raw junctions: ...

Additional

Show additionalparameters: +

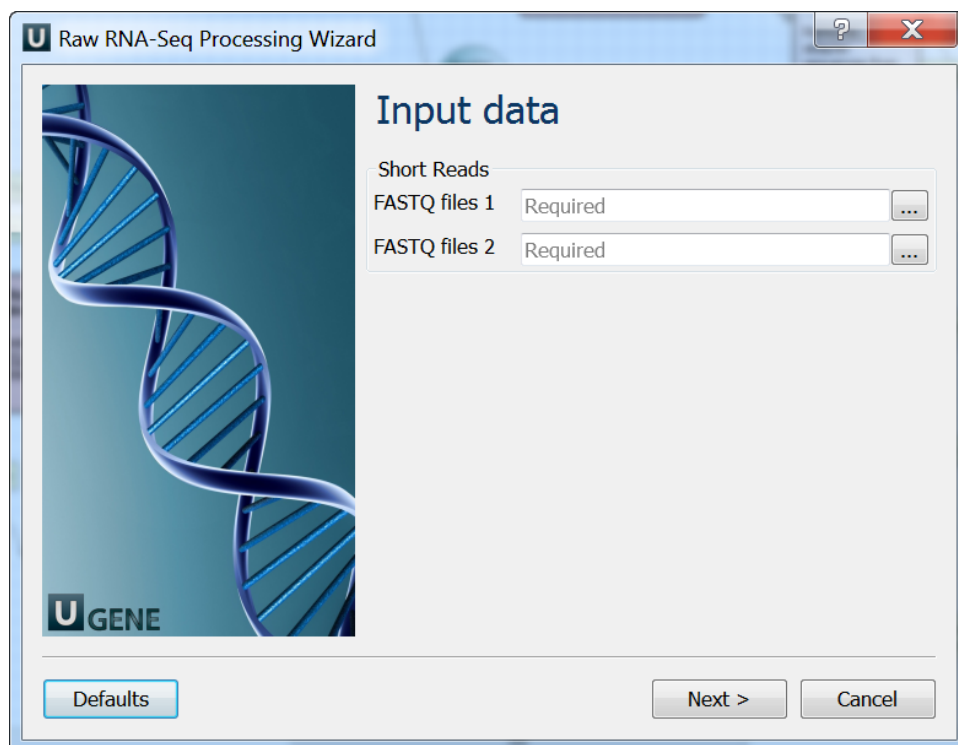
Buttons: Defaults, < Back, Next >, Cancel

4. Output data: On this page you must input output parameters.



The wizard for paired-end reads has 5 page.

1. Input data: On this page you must input FASTQ file(s).



2. Pre-processing: On this page you can modify filtration parameters.

Raw RNA-Seq Processing Wizard

Pre-processing

Filtration 1

Quality threshold: 20

Min length: 10

Adapters: ork/ugene/data/adapters/adapters.fasta ...

Filtration 2

Quality threshold: 20

Min length: 10

Adapters for paired: ork/ugene/data/adapters/adapters.fasta ...

Defaults < Back Next > Cancel

3. Mapping: On this page you must input reference and optionally modify advanced parameters.

Raw RNA-Seq Processing Wizard

Mapping

TopHat input

Bowtie index directory: Required

Bowtie index basename: Required

Bowtie version: Bowtie1

Select bowtie index file

Parameters

Known transcript file: ...

Raw junctions: ...

Additional

Show additional parameters: +

Defaults < Back Next > Cancel

4. Output data: On this page you must input output parameters.

U

Raw RNA-Seq Processing Wizard

?

X



U GENE

Output data

TopHat data

Tophat...

Filtered FASTQ 1

Show filtered fastq 1parameters

+

Filtered FASTQ 2

Show filtered fastq 2parameters

+

Defaults

< Back

Apply

Run

Cancel