LEIDEN UNIVERSITY MEDICAL CENTER

RNA-seq using Galaxy

Jeroen F. J. Laros

Leiden Genome Technology Center

Department of Human Genetics

Center for Human and Clinical Genetics

Introduction

Sequencers.

Figure 1 : HiSeq 2000.

Figure 2 : Ion proton.

High throughput, accurate, cheap.

Introduction

General layout of an RNA-seq pipeline.

1. Pre-alignment.
    - QC.
    - Data cleaning.

Introduction

General layout of an RNA-seq pipeline.

1. Pre-alignment.
    - QC.
    - Data cleaning.
2. Alignment.
    - Use a specialised (RNA) aligner.

Introduction

General layout of an RNA-seq pipeline.

1. Pre-alignment.
    - QC.
    - Data cleaning.
2. Alignment.
    - Use a specialised (RNA) aligner.
3. Expression (gene, transcripts) analysis.
    - Known transcripts.

Introduction

General layout of an RNA-seq pipeline.

1. Pre-alignment.
    - QC.
    - Data cleaning.
2. Alignment.
    - Use a specialised (RNA) aligner.
3. Expression (gene, transcripts) analysis.
    - Known transcripts.
4. Transcript assembly.
    - New transcripts, alternative splicing, etc.

Introduction

Combining tools in a pipeline.

```
1  bwa aln -t 8 $reference $i > $i.sai
2  bwa samse $reference $i.sai $i > $i.sam
3  samtools view -bt $reference -o $i.bam $i.sam
```

Listing 1 : Shell script.

Introduction

Combining tools in a pipeline.

```
1  bwa aln -t 8 $reference $i > $i.sai
2  bwa samse $reference $i.sai $i > $i.sam
3  samtools view -bt $reference -o $i.bam $i.sam
```

Listing 1 : Shell script.

```
1  %.sai: %.fq
2    $(BWA) aln -t $(THREADS) $(call MKREF, $@) $< > $@
3
4  %.sam: %.sai %.fq
5    $(BWA) samse $(call MKREF, $@) $^ > $@
6
7  %.bam: %.sam
8    $(SAMTOOLS) view -bt $(call MKREF, $@) -o $@ $<
```

Listing 2 : Makefile.

The Galaxy user interface

Overview.

Data intensive biology for everyone.

- Open source.
- Web based.
  - No installation required.

http://galaxy.psu.edu/

The Galaxy user interface

Overview.

Data intensive biology for everyone.

- Open source.
- Web based.
    - No installation required.

Wrapper for command line utilities.

- User friendly.
    - Point and click.
- Workflows.
    - Save all the steps you did in your analysis.
    - Rerun the entire analysis on a new dataset.
    - Share your workflow with other people.

http://galaxy.psu.edu/

The Galaxy user interface

Global overview of the GUI.

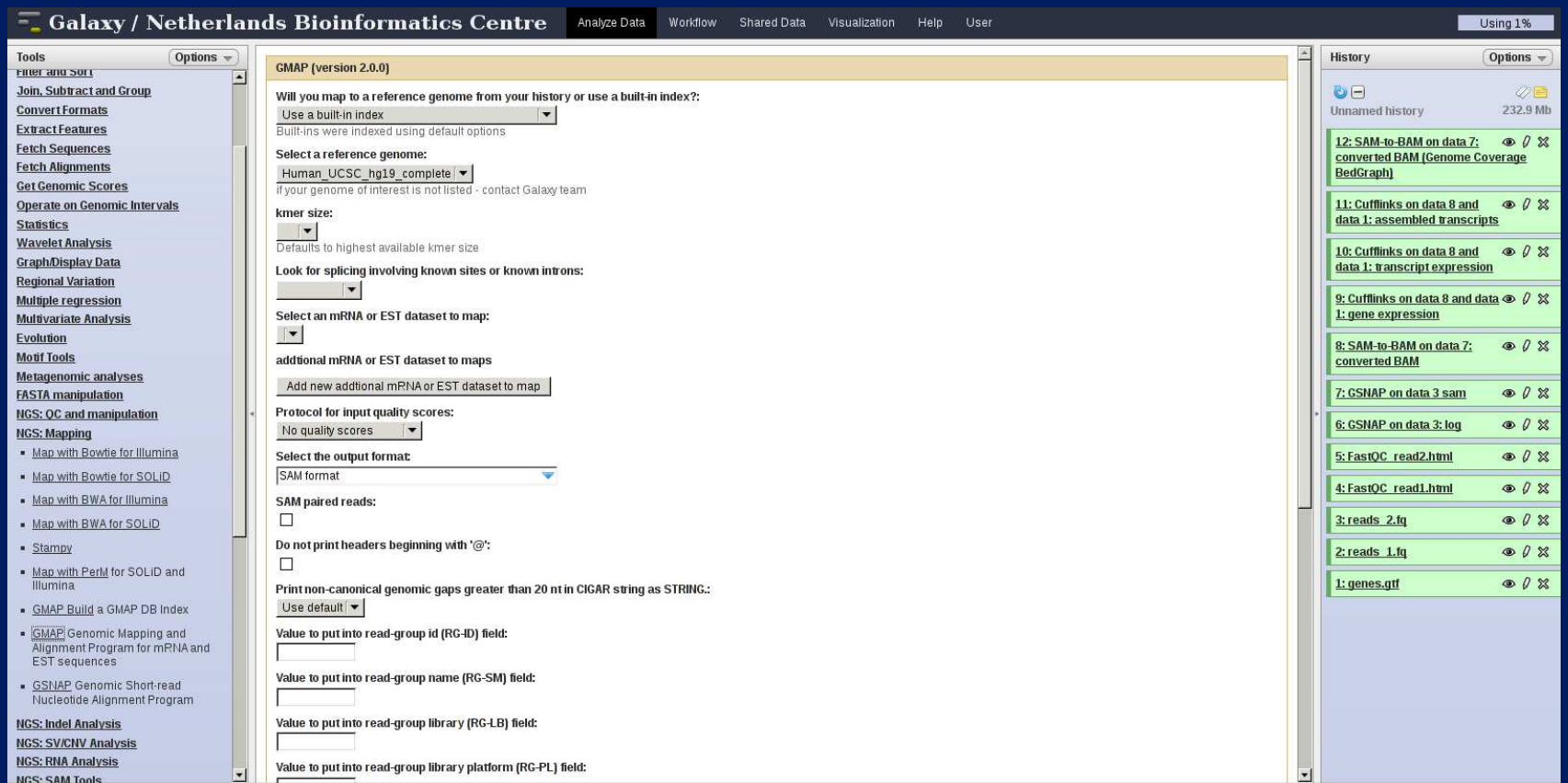

Figure 3 : Galaxy panels.

The Galaxy user interface

The Galaxy tool menu.

Lots of tools installed, especially for NGS data analysis.

Figure 4 : Collapsed tool menu.

The Galaxy user interface

The Galaxy tool menu.

Lots of tools installed, especially for NGS data analysis.

Figure 4 : Collapsed tool menu.

Figure 5 : Tool menu.

The Galaxy user interface

Selecting a tool.

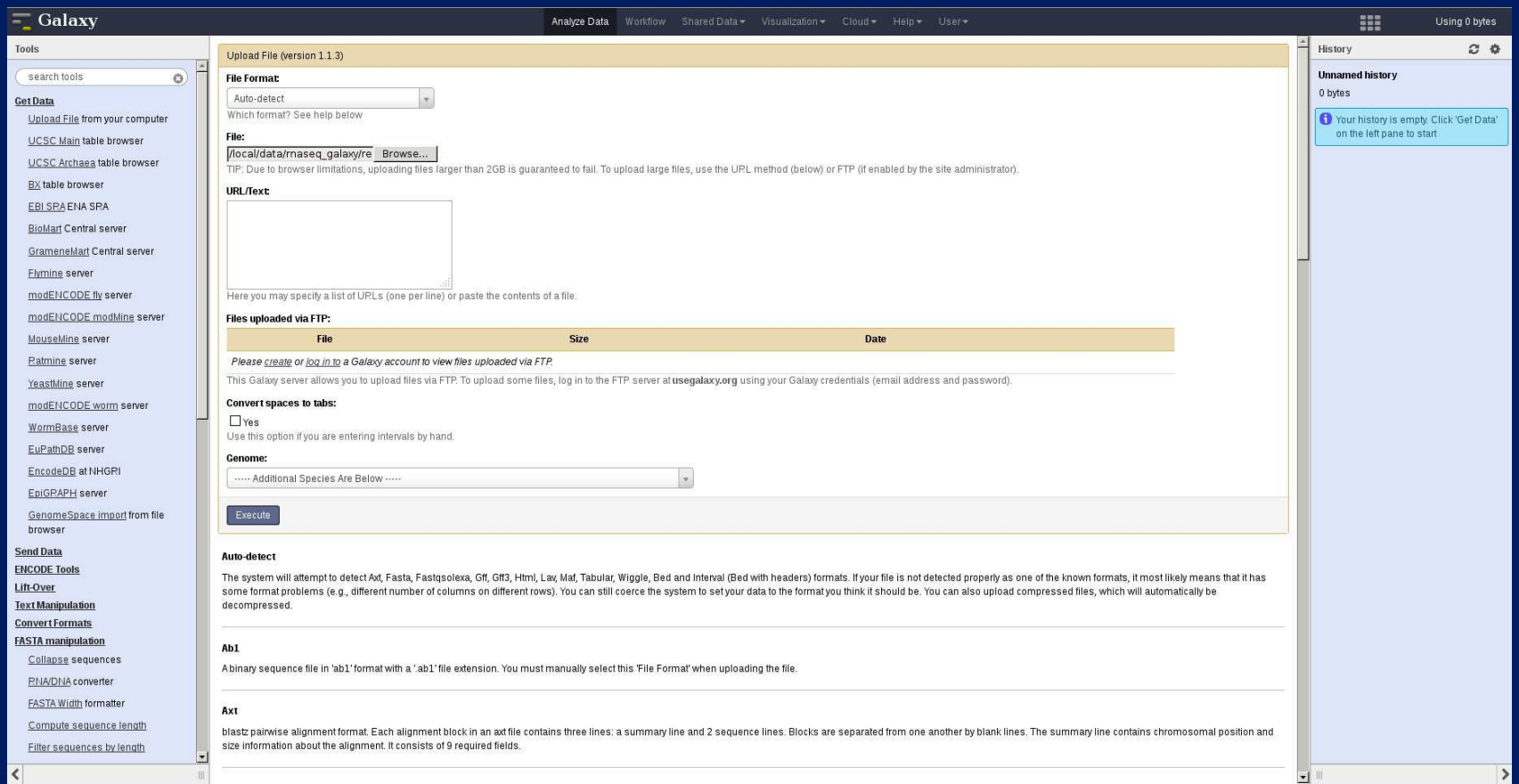

Figure 6 : Uploading a file.

The Galaxy user interface

The history panel.

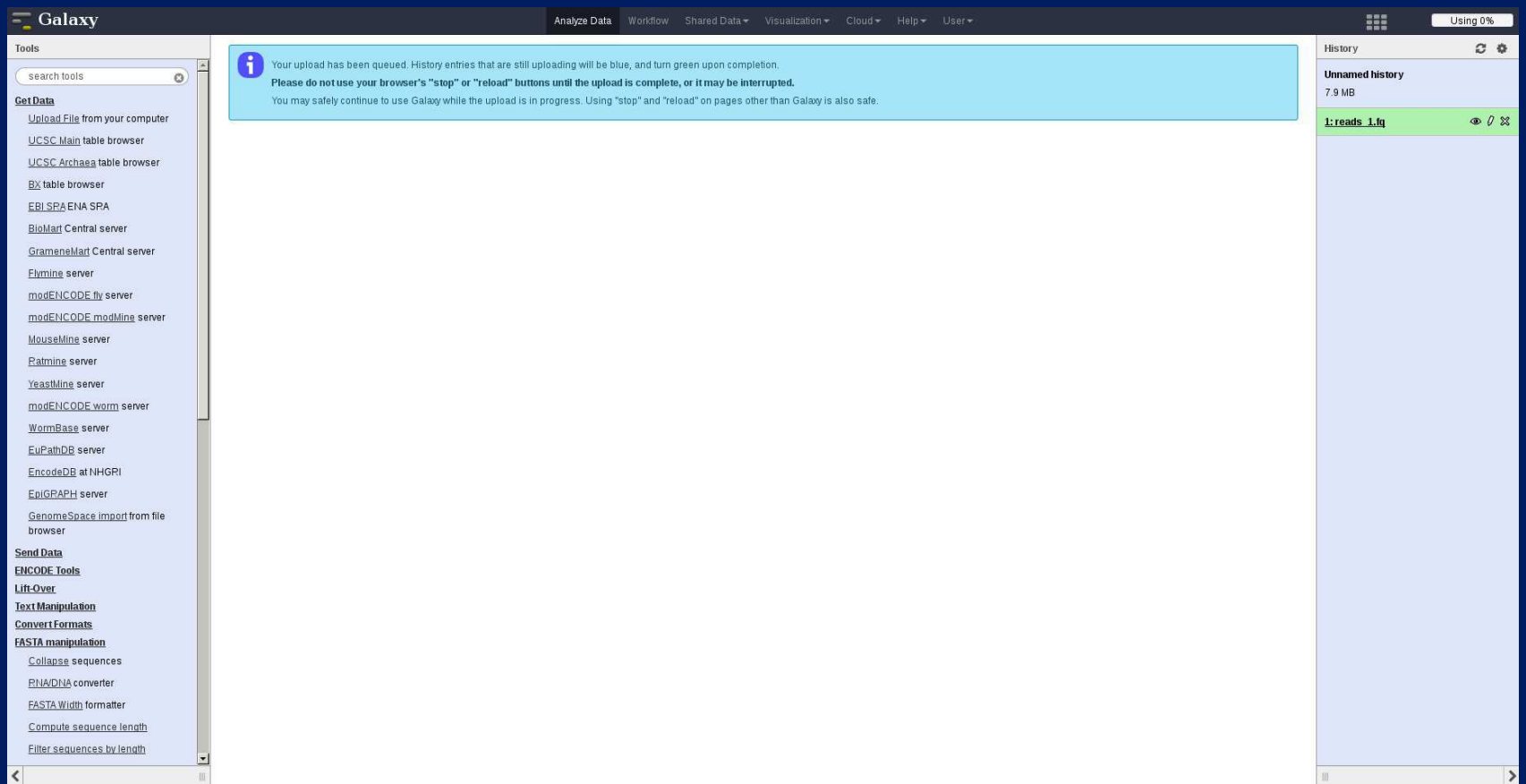

Figure 7 : Uploading completed.

The Galaxy user interface

History item icons.

Figure 8 : Collapsed history item.

Every history item has some basic functions:

- Eye: view.
- Pencil: edit (rename).
- Cross: delete.

The Galaxy user interface

History item icons.

Figure 8 : Collapsed history item.

Every history item has some basic functions:

- Eye: view.
- Pencil: edit (rename).
- Cross: delete.

Click on the title for a more detailed view.

The Galaxy user interface

History item icons.

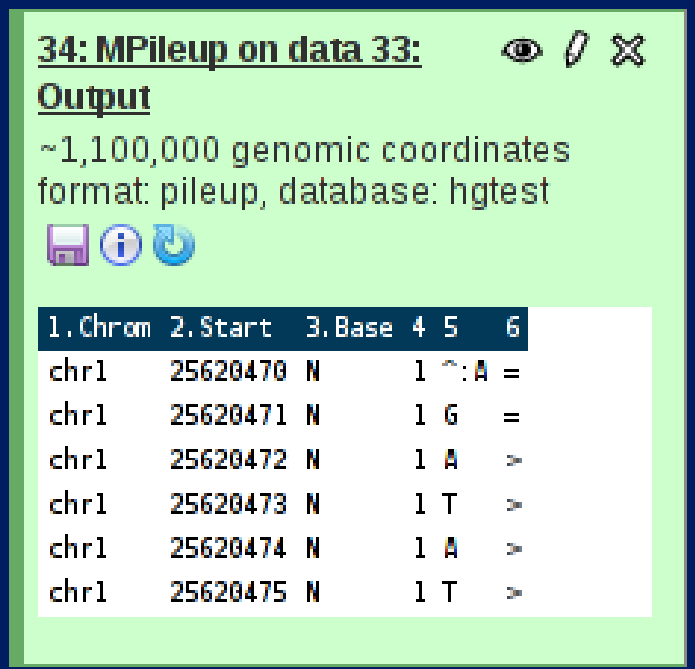

Figure 9 : History item.

- Diskette: save.
- Blue looping arrow: rerun.

The Galaxy user interface

Viewing the data (eye icon).

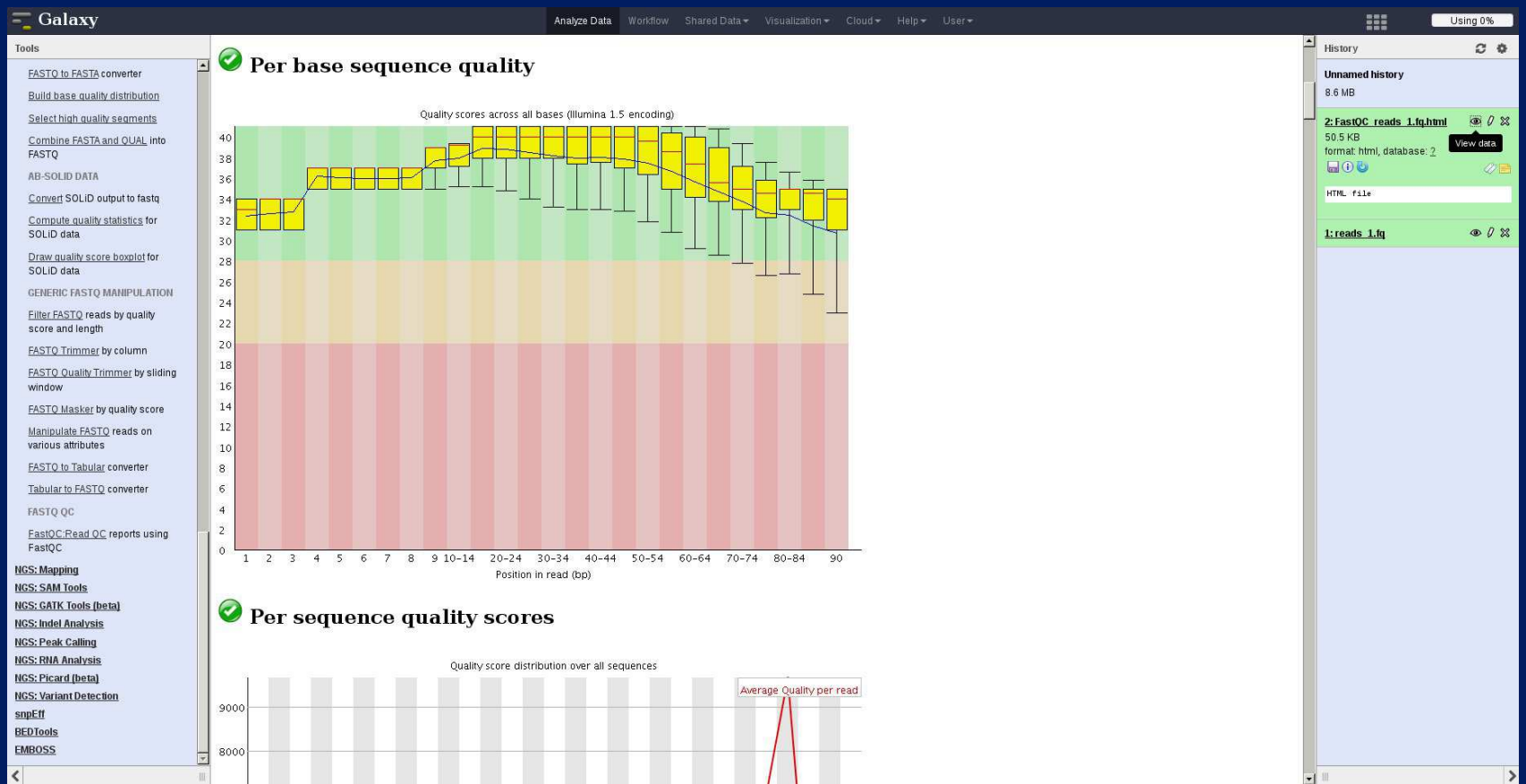

Figure 10 : See your output in the tool panel.

Workflows

Extracting a workflow.

Figure 11 : Extract workflow.

Once you have finished an analysis, you can save it as a workflow.

- Store for later use.
- Share with others.

Workflows

Extracting a workflow.

Figure 11 : Extract workflow.

Once you have finished an analysis, you can save it as a workflow.

- Store for later use.
- Share with others.

Click on the “gear” icon in the history panel.

Workflows

Modifying workflows.

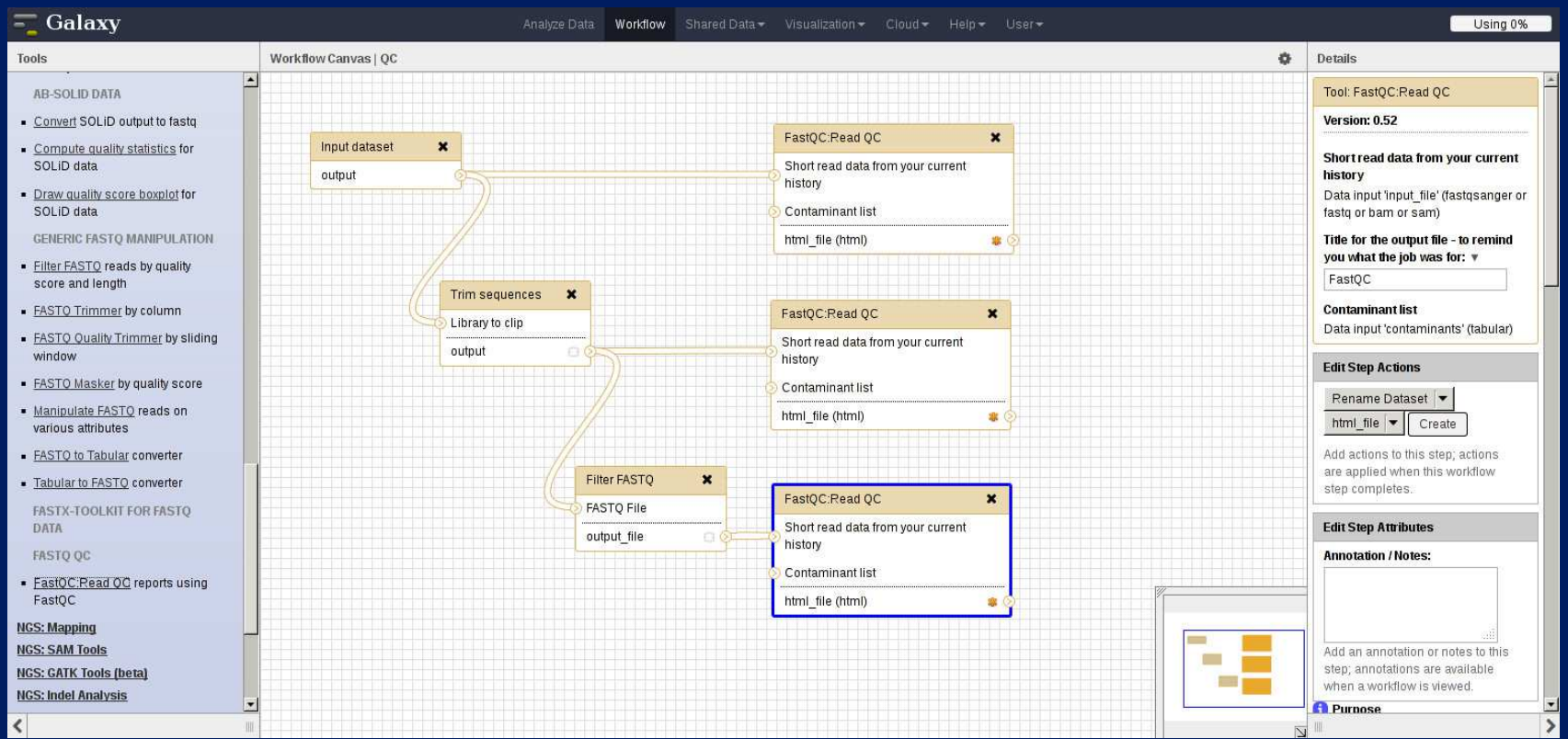

Figure 12 : Create or edit your workflows.

Questions?

Acknowledgements:

Wibowo Arindrarto
Irina Pulyakhina
Mateusz Kuzak
Hailiang Mei
Peter-Bram 't Hoen
Johan den Dunnen

http://galaxy.nbic.nl/