

# User feedback of a Galaxy workflow conception

**Claire Toffano-Nioche**

eBio facility – Université Paris-Sud, Orsay

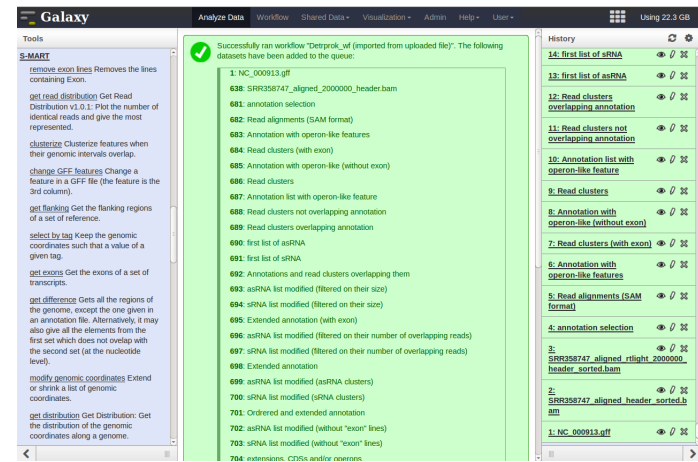
IGM - Institut de Génétique et Microbiologie

# How have I built the Workflow?

step-by-step, biologist  
Data = README file  
s-mart tools, home-made scripts

```
### 1- select reads outside "gene" :
# select reads (-i VIBSP_chrl_smart.gff3) outside
# "annotation" (-j NC_011744_annot.gff3 -c -x), with
# keeping reads overlapping ATG (-e 10)
for i in "NC_011744" "NC_011753" ; do
    compareOverlapping.py -i ${i}_RNAseq_smart.gff3 -f gff3
    clusterize.py -i ${i}_cis_ReadsOutGene.gff3 -f gff3 -e 10 -x -c -o ${i}_cis_ReadsOutGene ;
done ;
### 2- clusterize reads before seed selection :
# 2 clusterization steps: strict + d < 20
for i in "NC_011744" "NC_011753" ; do
    clusterize.py -i ${i}_cis_ReadsOutGene.gff3 -f gff3 -c -o ${i}_cis_e10_clusters ;
    clusterize.py -i ${i}_cis_e10_clusters.gff3 -f gff3 -c -d 20 -o ${i}_cis_e10_clusters20 ;
### 3- crossing cluster with seed :
for i in "NC_011744" "NC_011753" ; do
    seedGff.pl -i ${i}_annot.gff -p 25 -e 15 -o ${i}_cis_seed.gff ; #
    seed creation in -15 -25 bp from ATG
    compareOverlapping.py -i ${i}_cis_e10_clusters20.gff3 -f gff3 -j ${i}_cis_seed.gff -g gff3 -c -o ${i}_cis_e10_cluster20InSeed2515 ; #
    crossing cluster with seed
done ;
#
### 4- filters: nb_reads > 10 reads + length > 50 nt
for i in "NC_011744" "NC_011753" ; do
    selectByTag.py -i ${i}_cis_e10_cluster20InSeed2515.gff3 -f gff -o ${i}_cis_e10_cluster20InSeed2515_nbEUp10 -g "nbElements" -m 10 -d 0 ; # nb_reads > 10 filter
done ;
...
```

Wrap tool, share (test/toolshed)  
Import on my local Galaxy instance  
Manual launch, follow the readme file



The screenshot shows the Galaxy interface with a workflow titled "Detprk\_vf" (imported from uploaded file). The workflow consists of 704 steps, including S-MART, remove exon lines, get read distribution, clusterize, change GFF features, get flanking, select by tag, get exons, get difference, modify genomic coordinates, and get distribution. The workflow is shown as successfully completed.

Extract Workflow from History  
Edit workflow: close or open parameter  
(« set at run time »), add comments,  
change name of the steps  
Share Workflow (test/toolshed)

# Technical issues

- Use « virtual\_env »
- Dissociate « your » directories from the Galaxy distribution (configure the universe\_wsgi.ini file)
  - galaxy-dist
  - galaxy\_env
  - repository\_dependencies
  - toolshed
  - upload\_libraries
- « underlying » database :
  - SQLite => postgres
  - allow concurrent accesses, count DB in the ressources needed
- Install a Galaxy instance on a stratuslab virtual machine during revision process of the publication



# Feedbacks

## User's point of view: local instance vs. Galaxy server ?

User : import workflows, not the tools (admin account).

Ok in a local instance, contact the admin in a server => not so easy !

## Open questions (perhaps have you the answers?):

- How expliciting the need of specific tools? **dependencies?**
- « **Global** » **variables?** Launching a tool many times implies to re-enter the same values for open parameters. Environment variables?
- **Meta-workflow?** Begin = bam file (user chooses the mapper). How add the mapping step? another workflow and combine the two?




## Methods

Volume 63, Issue 1, 1 September 2013, Pages 60–65

Diversity of the non-coding transcriptomes revealed by RNA-seq technologies



## Detection of non-coding RNA in bacteria and archaea using the DETR'PROK Galaxy pipeline ☆

Claire Toffano-Nioche<sup>a</sup>, Yufei Luo<sup>b</sup>, Claire Kuchly<sup>a</sup>, Claire Wallon<sup>a</sup>, Delphine Steinbach<sup>b</sup>, Matthias Zytynski<sup>b</sup>, Annick Jacq<sup>a</sup>, Daniel Gautheret<sup>a</sup>  

<sup>a</sup> Université Paris-Sud, Institut de Génétique et Microbiologie, CNRS UMR 8621, Orsay F-91405, France

<sup>b</sup> URGI, INRA, Versailles F-78026, France



[Open Access](#)

### Galaxy Tool Shed / (sandbox) Repositories Help User

2279 valid tools on Oct 24, 2013

#### Search

- Search for valid tools
- Search for workflows

#### Valid Galaxy Utilities

- Tools
- Custom datatypes
- Repository dependency definitions
- Tool dependency definitions

#### All Repositories

- Browse by category

#### Available Actions

- Login to create a repository

### Repositories with matching workflows

workflow name: **detrprok**  
exact matches only: **False**

<input type="checkbox"/>	Repository name	Synopsis	Revision	Owner
<input type="checkbox"/>	detrprok_wf	ncRNA detection in prokaryote oriented RNAseq	f0ca0981eb86	clairetn
<input type="checkbox"/>	detrprok_wf	ncRNA detection in prokaryote oriented RNAseq	8f7eb127baf3	clairetn
<input type="checkbox"/>	detrprok_wf	ncRNA detection in prokaryote oriented RNAseq	bb71a378053a	clairetn

For 0 selected items:

