Tools

Options

Your user name: smaman
Your file path : /work/smaman/

1 - UPLOAD YOUR DATA

Get Data

2 - FILES MANIPULATION

Text Manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats

3 - SEQUENCES MANIPULATION

FASTA manipulation
FASTQ manipulation
SAM/BAM manipulation : Picard (beta)
SAM/BAM manipulation : SAM Tools

4 - MAPPING

BWA - Bowtie

5 - INDEL ET SNP

Indel Analysis
RNA-Seq
GATK Tools (beta)

6 - SRNASEQ

Analyse des miRNA
Annotations
Alignement sur reference

WELCOME ON SIGENAE GALAXY WORKBENCH

Galaxy is a workbench available for biologists from Sigenae Platform. Galaxy objectives are:

- Make bioinfo Linux tools accessible to biogists.
- Hide the complexity of the infrastructure.
- Allow creation, execution and sharing of workflows.

History

Options

TP FastQC — 54.0 Mb

8: FastQC_data 5.html
6: GM.fastqsanger
5: h1.fastqsanger
4: FastQC_data 18.html
3: FASTQ Summary Statistics on data 18

2: FASTQ Summary Statistics on data 18

76 lines, 1 comments
format: tabular, database: ?
Info: 99115 fastq reads were processed.
Based upon quality values and sequence characters, the input data is valid for: sanger
Input ASCII range: '#'(35) - 'C'(67)
Input decimal range: 2 - 34
Epilog : job finished at ven mai 11 10:36:43 CEST 2012

1	2	3	4	5	6
#column	count	min	max	sum	mean
1	99115	2	33	3194703	32.2
2	99115	2	34	3156652	31.8
3	99115	2	34	3145060	31.7
4	99115	2	34	3120431	31.4
5	99115	2	34	3096075	31.2

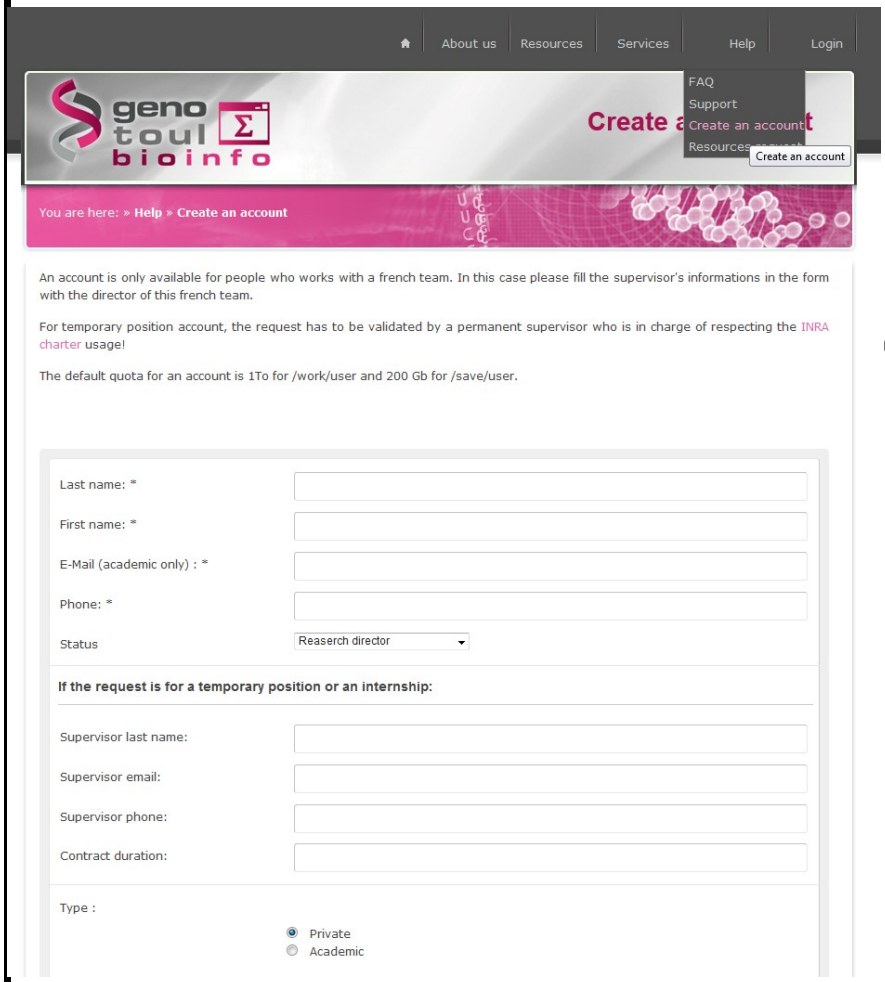
Etat des lieux Galaxy & Evolutions récentes

Sarah Maman – Ibouniyamine Nabihoudine
24 mars 2014 – GenPhySE

http://galaxy-workbench.toulouse.inra.fr

Aperçu général

Comment se connecter ?



1 – Ouvrir un compte sur Genotoul :

Formulaire de demande de compte:
http://bioinfo.genotoul.fr

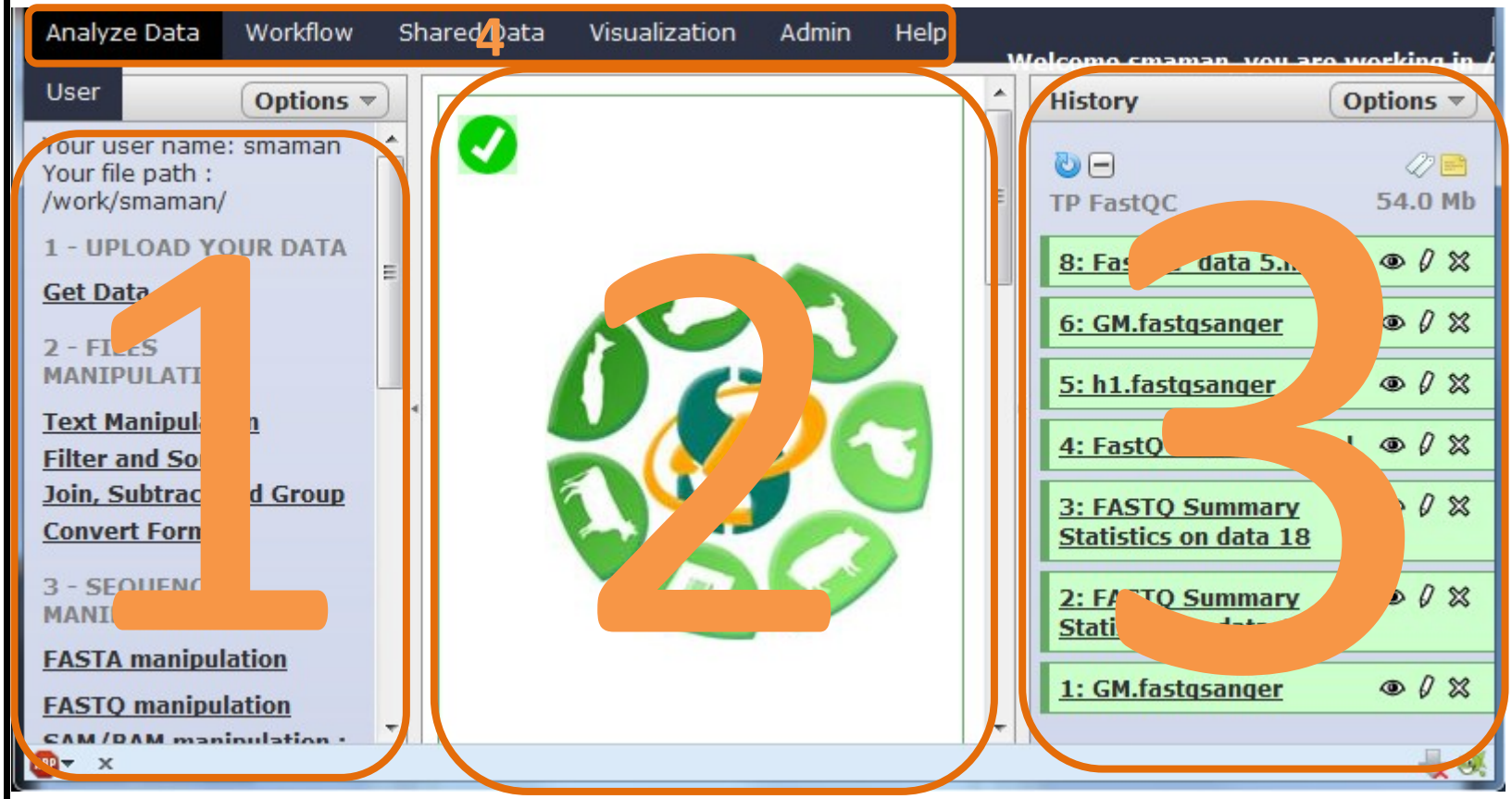
Utiliser un mail académique

2 – Accéder à Galaxy à l'aide du login/mot de passe obtenus :

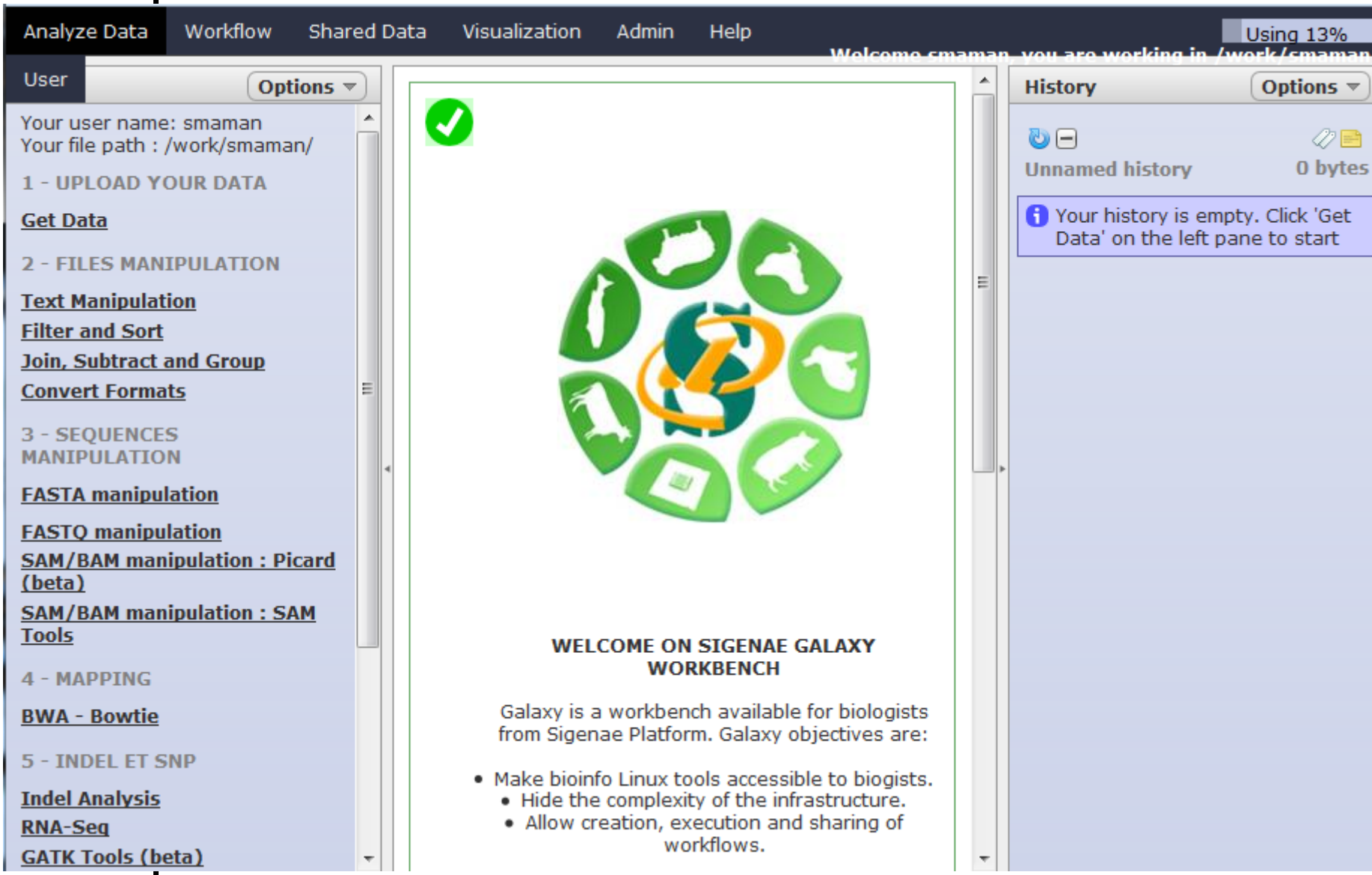
http://sigenae-workbench.toulouse.inra.fr/

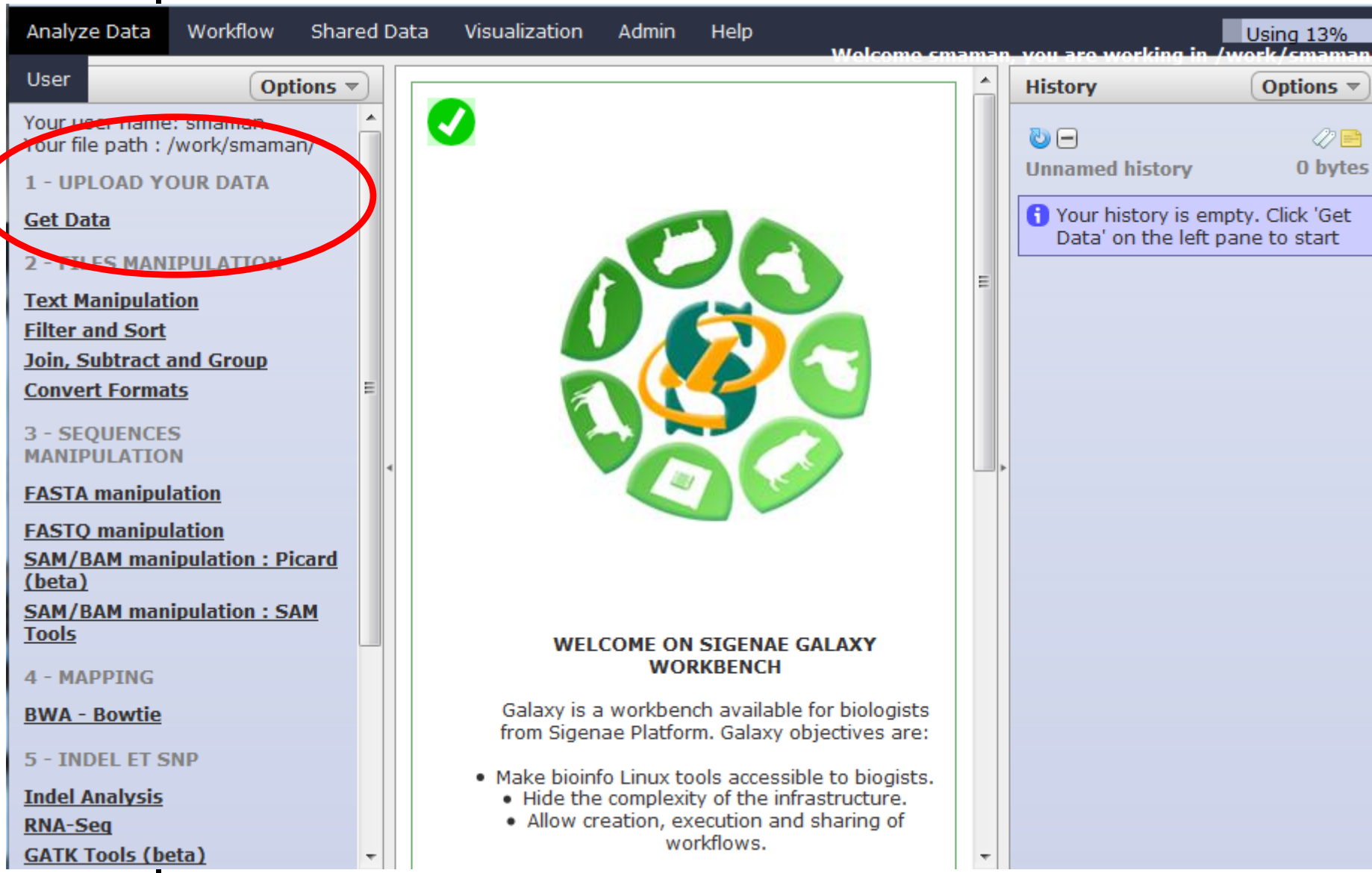
Une interface simplifiée et intuitive

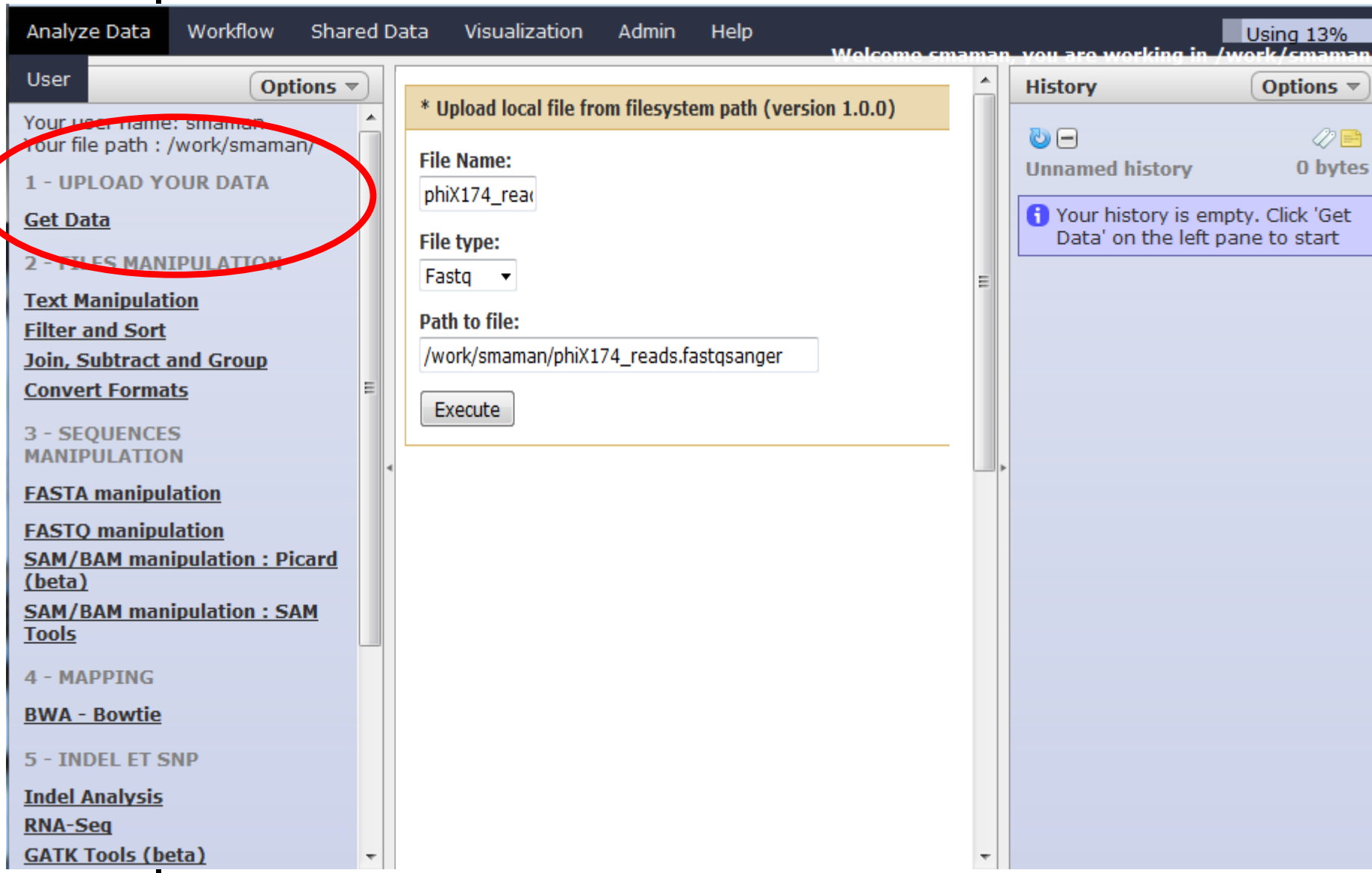
Votre écran est divisé en 4 zones principales :

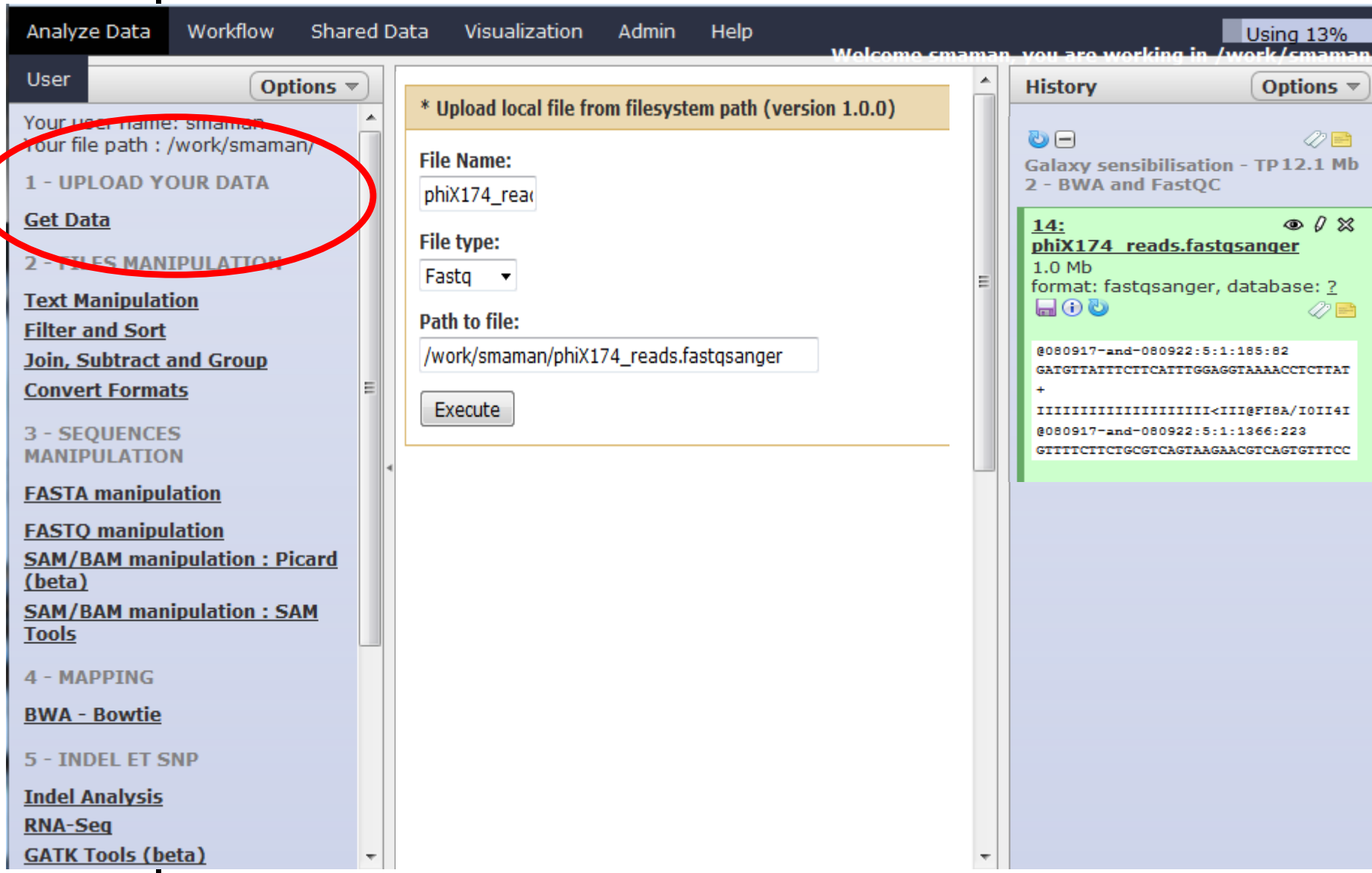


Comment utiliser un outil ?



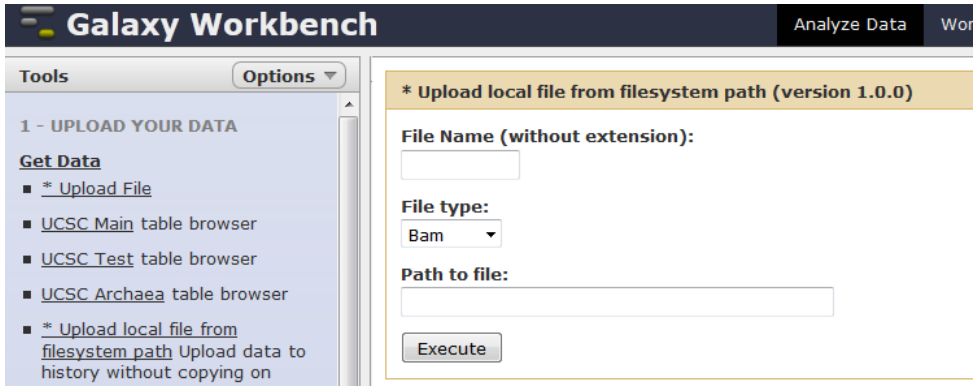

Analyze Data
Workflow
Shared Data
Visualization
Admin
Help
Using 13%
User
Options
Your file path : /work/smaman/
1 - UPLOAD YOUR DATA
Get Data
2 - FILES MANIPULATION
Text Manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
3 - SEQUENCES MANIPULATION
FASTA manipulation
FASTQ manipulation
SAM/BAM manipulation : Picard (beta)
SAM/BAM manipulation : SAM Tools
4 - MAPPING
BWA - Bowtie
5 - INDEL ET SNP
Indel Analysis
RNA-Seq
GATK Tools (beta)
WELCOME ON SIGENAE GALAXY WORKBENCH
Galaxy is a workbench available for biologists from Sigenae Platform. Galaxy objectives are:
Make bioinfo Linux tools accessible to biogists.
Hide the complexity of the infrastructure.
Allow creation, execution and sharing of workflows.
History
Options
Unnamed history
0 bytes
Your history is empty. Click 'Get Data' on the left pane to start


Analyze Data
Workflow
Shared Data
Visualization
Admin
Help
Using 13%
User
Options
Your file path : /work/smaman/
1 - UPLOAD YOUR DATA
Get Data
2 - FILES MANIPULATION
Text Manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
3 - SEQUENCES MANIPULATION
FASTA manipulation
FASTQ manipulation
SAM/BAM manipulation : Picard (beta)
SAM/BAM manipulation : SAM Tools
4 - MAPPING
BWA - Bowtie
5 - INDEL ET SNP
Indel Analysis
RNA-Seq
GATK Tools (beta)
* Upload local file from filesystem path (version 1.0.0)
File Name:
phiX174_rea
File type:
Fastq
Path to file:
/work/smaman/phiX174_reads.fastqsanger
Execute
History
Options
Unnamed history
0 bytes
Your history is empty. Click 'Get Data' on the left pane to start

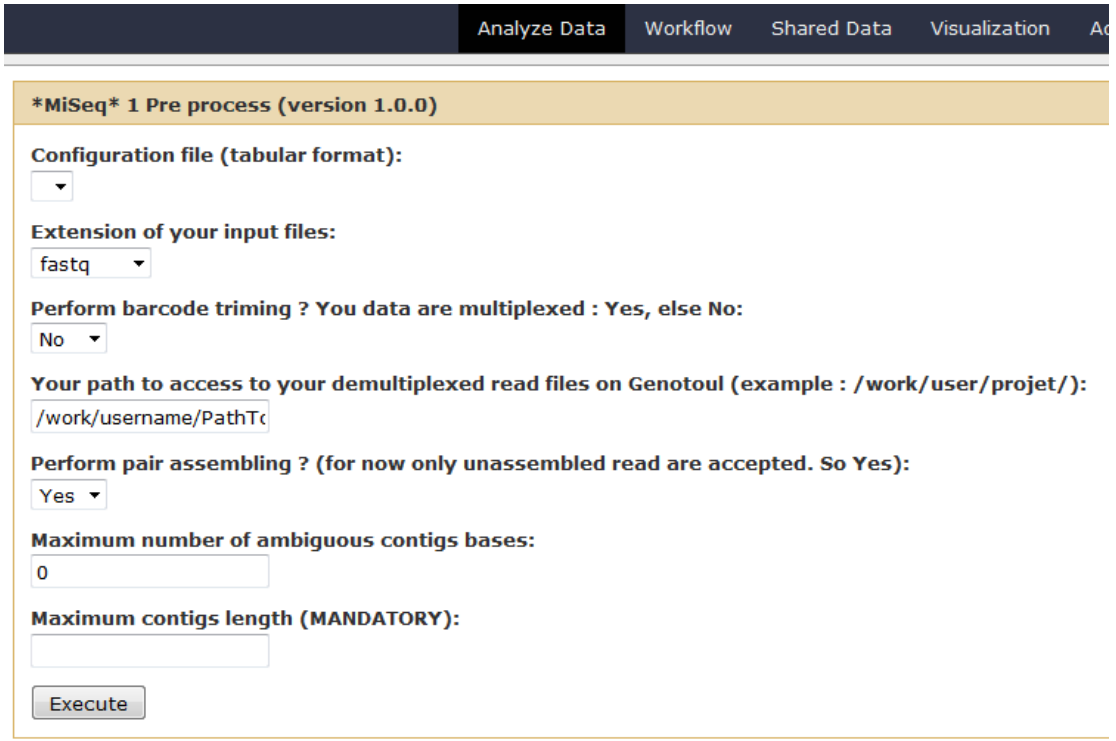

Analyze Data
Workflow
Shared Data
Visualization
Admin
Help
Using 13%
User
Options
Your file path : /work/smaman/
1 - UPLOAD YOUR DATA
Get Data
2 - FILES MANIPULATION
Text Manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
3 - SEQUENCES MANIPULATION
FASTA manipulation
FASTQ manipulation
SAM/BAM manipulation : Picard (beta)
SAM/BAM manipulation : SAM Tools
4 - MAPPING
BWA - Bowtie
5 - INDEL ET SNP
Indel Analysis
RNA-Seq
GATK Tools (beta)
* Upload local file from filesystem path (version 1.0.0)
File Name:
phiX174_rea
File type:
Fastq
Path to file:
/work/smaman/phiX174_reads.fastqsanger
Execute
History
Options
Galaxy sensibilisation - TP 12.1 Mb
2 - BWA and FastQC
14: phiX174_reads.fastqsanger
1.0 Mb
format: fastqsanger, database: ?
@080917-and-080922:5:1:185:82
GATGTTATTTCTTCATTTGGAGGTAAAACCTCTTAT
+
IIIIIIIIIIIIIIIIIIII<III@FI8A/I0II4I
@080917-and-080922:5:1:1366:223
GTTTTCTTCTGCGTCAGTAAGAACGTCAGTGTTTCC

Liens avec NG6

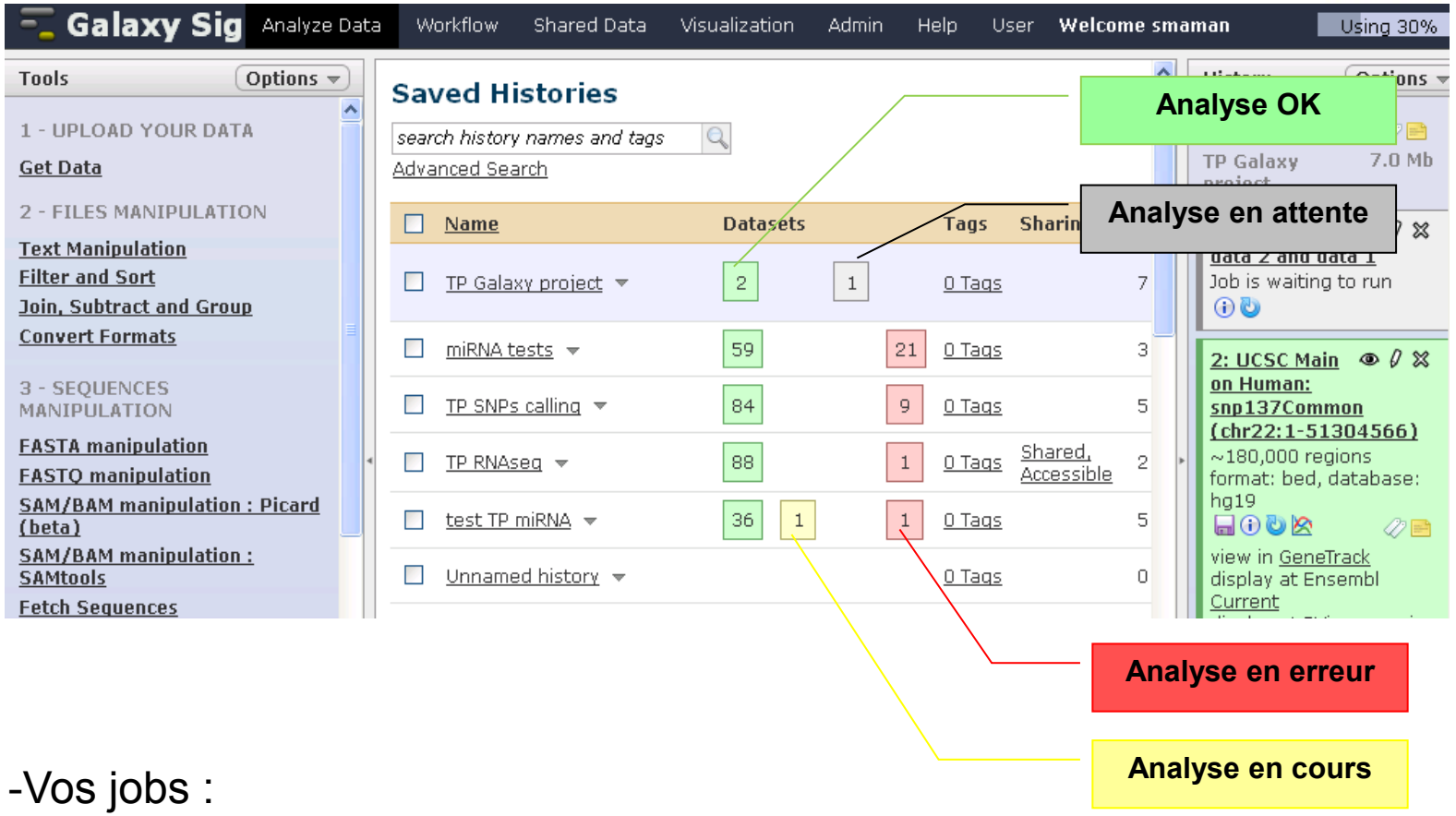
-Soit fichier par fichier :



-Soit un répertoire de fichiers :



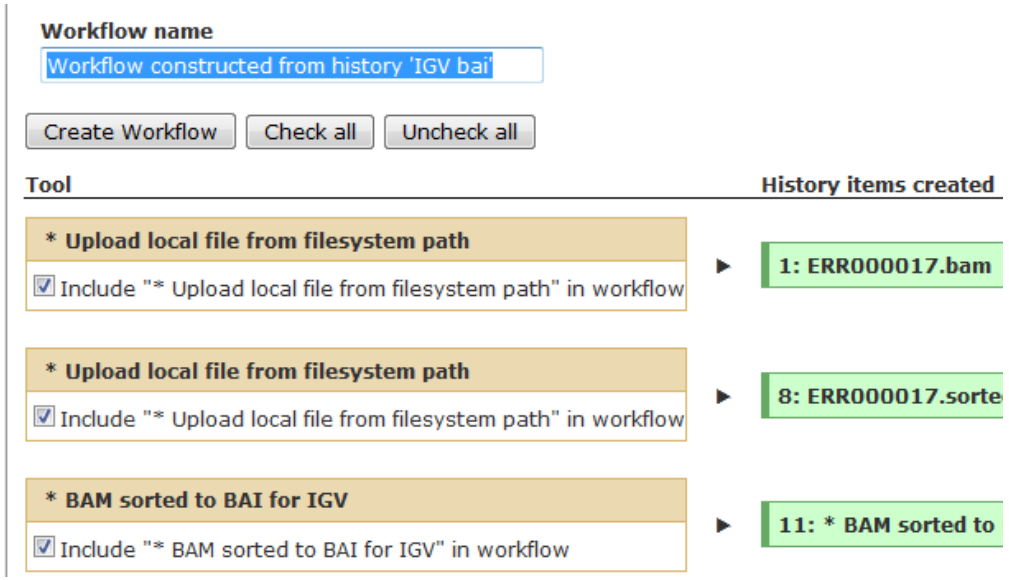
De l'outil à l'historique.

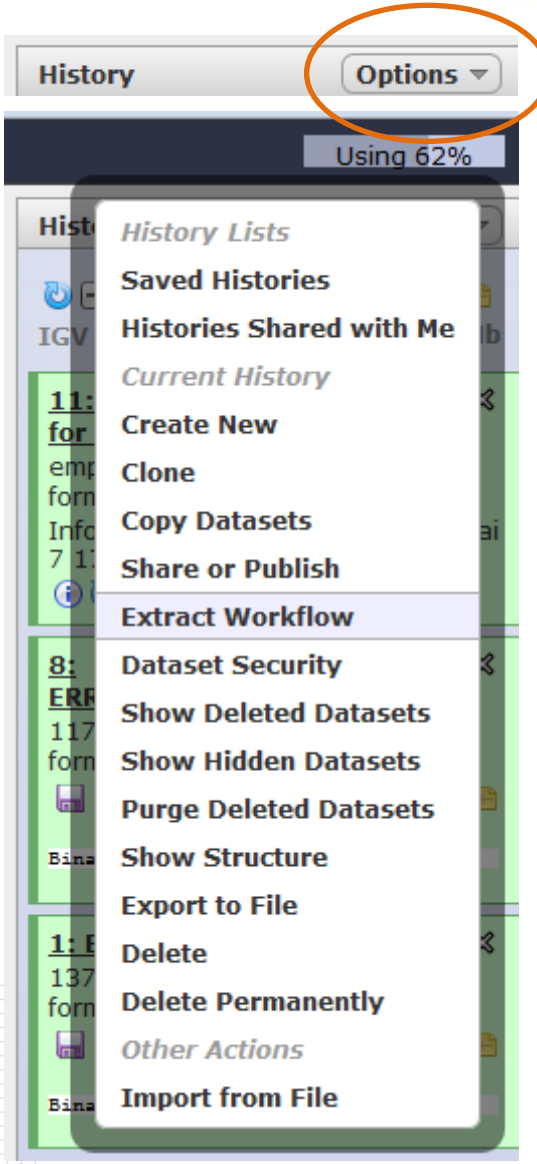


-Vos jobs :
    -Se poursuivent même si vous êtes déconnectés du Galaxy.
    -Leurs états sont indiqués par la couleur du dataset.
    -Peuvent être lancés en parallèle & sans connaître linux.
    -Peuvent être partagés.
-Votre données :
    -N'encombrent pas votre disque dur.
    -Sont protégées (accès LDAP).

Comment générer une chaîne de traitement ?

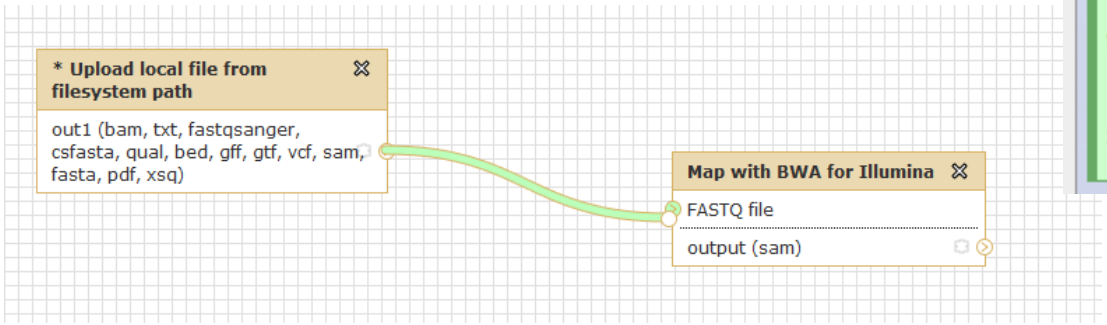
-Soit depuis un historique :





-Soit depuis une page blanche :

Les résultats produits sont typés,
il n'est donc pas possible de brancher une dataset sur un mauvais tool !



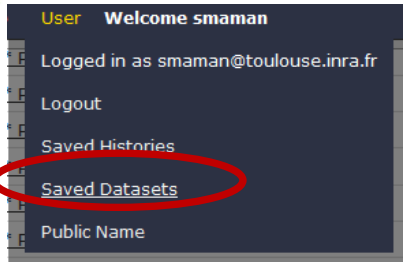
Evolutions récentes

Des outils plus complets

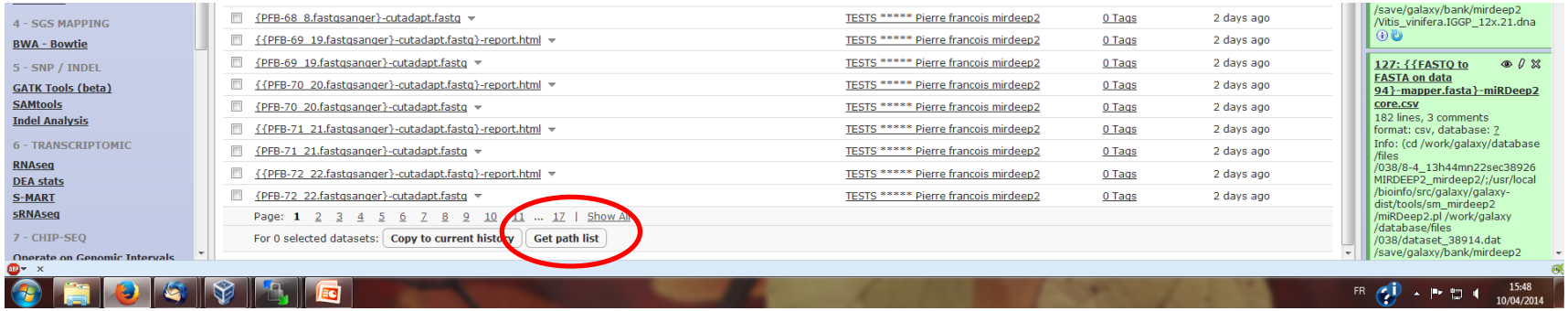
http://galaxy-workbench.toulouse.inra.fr/

Pour sauver vos datasets Galaxy dans votre /work

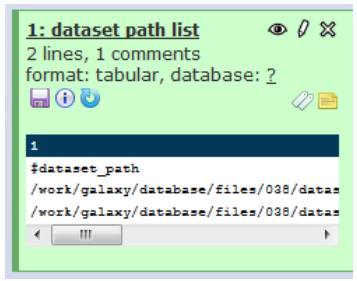
1



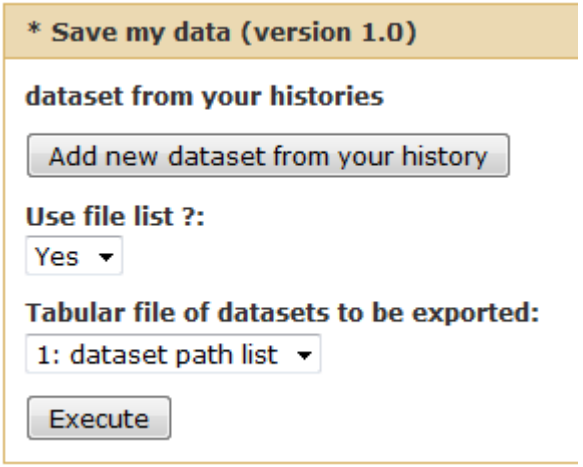
2



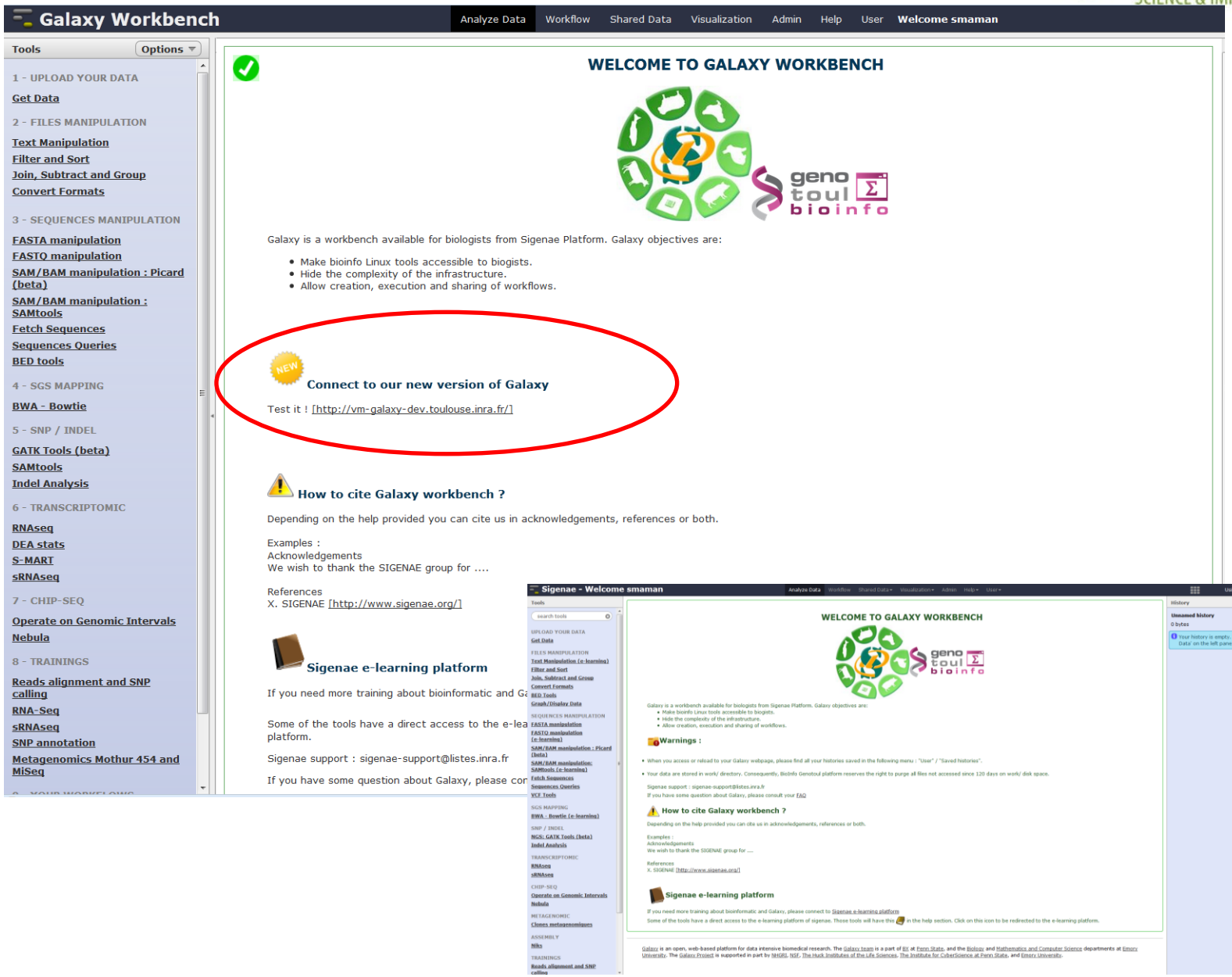
3



4



Nouvelle machine a venir



Améliorations de la nouvelle instance

•~30 nouveaux outils et 5 chaînes de traitements.

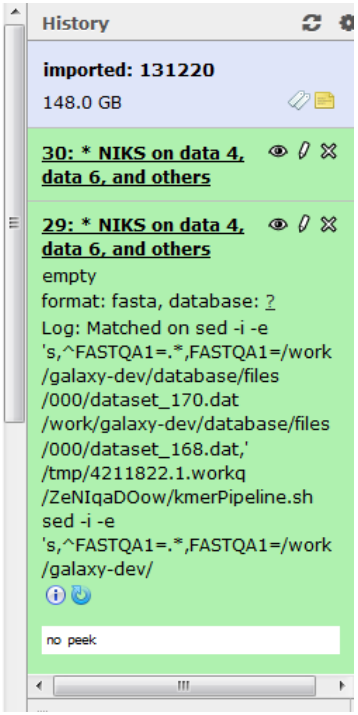
•Matériel & méthode : Commandes / versions d'outils Galaxy et bioinfo.

•Machine :

- ✓Sur la nouvelle infrastructure (genovcenter)
- ✓Plus rapide d'accès sur site distant (multithreading)
- ✓Envoi des jobs au nom du user LDAP (workq ou autre q.)
- ✓Paramétrage plus fin des jobs en fonction des ressources.

Tool: * NIKS

Name:	* NIKS on data 4, data 6, and others
Created:	Jan 13, 2014
Filesize:	0 bytes
Dbkey:	?
Format:	fasta
Galaxy Tool Version:	1.0.0
Tool Version:	
Tool Standard Output:	stdout
Tool Standard Error:	stderr
Tool Exit Code:	0
API ID:	aa03efb04bdce6b0
Full Path:	/work/galaxy-dev/database/files/000/dataset_621.dat
Job Command-Line:	perl /usr/local/bioinfo/src/galaxy-dev/galaxy-dist/tools/sm_niks /sm_niks.pl --pipeline-type KMER --fastqa1 /work/galaxy-dev/database /files/000/dataset_170.dat --fastqa1 /work/galaxy-dev/database/files /000/dataset_168.dat --fastqb1 /work/galaxy-dev/database/files /000/dataset_169.dat --fastqb1 /work/galaxy-dev/database/files /000/dataset_171.dat --insert-size 100 --output-fasta-a /work/galaxy-dev/database/job_working_directory/000/304/galaxy_dataset_620.dat --output-fasta-b /work/galaxy-dev/database/job_working_directory /000/304/galaxy_dataset_621.dat --output-csv /work/galaxy-dev/database/job_working_directory/000/304/galaxy_dataset_622.dat

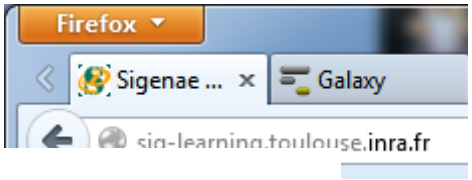


Site E-Learning

http://sig-learning.toulouse.inra.fr

Plateforme d'autoformation en ligne

« sig-learning » : **http://sig-learning.toulouse.inra.fr/**



Enquête 2013 :
Accès ponctuels à sig-learning

Quelques statistiques

Mise en production

Formations PF BioInfo

Formations PF BioInfo

Formations Roscoff

Formations Roscoff

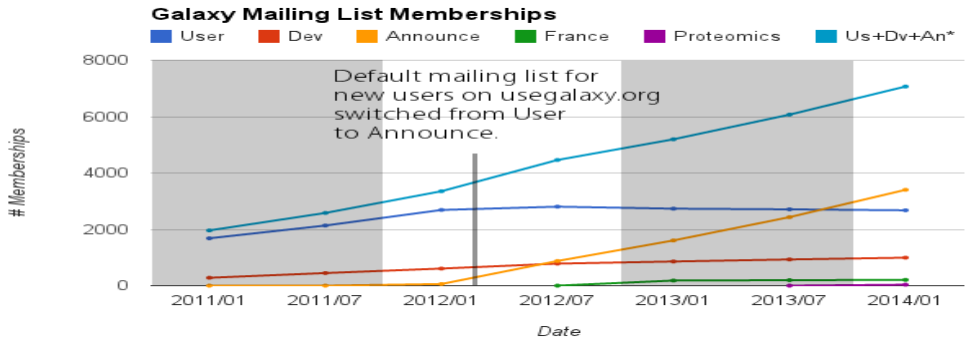
Logins (Month)			
2012 February		103	14,5%
2012 March		34	4,8%
2012 April		33	4,6%
2012 May		92	13,0%
2012 June		103	14,5%
2012 July		15	2,1%
2012 August		13	1,8%
2012 September		14	2,0%
2012 October		42	5,9%
2012 November		11	1,5%
2012 December		12	1,7%
2013 January		29	4,1%
2013 February		19	2,7%
2013 March		6	0,8%
2013 April		4	0,6%
2013 May		11	1,5%
2013 June		6	0,8%
2013 July		18	2,5%
2013 August		6	0,8%
2013 September		22	3,1%
2013 October		5	0,7%
2013 November		13	1,8%
2013 December		71	10,0%
2014 January		21	3,0%
2014 February		7	1,0%
		Total: 710	

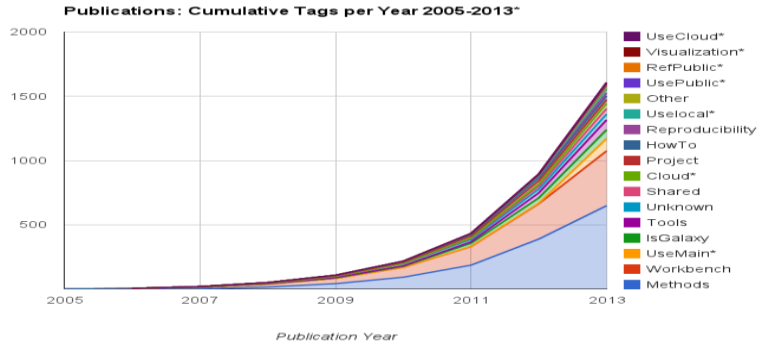
Enquête 2013 auprès des ~40 utilisateurs locaux

Utilisation des instances Galaxy

Statistiques du Galaxy Project

https://wiki.galaxyproject.org





Une communauté internationale vivante.
Une communauté française grandissante.

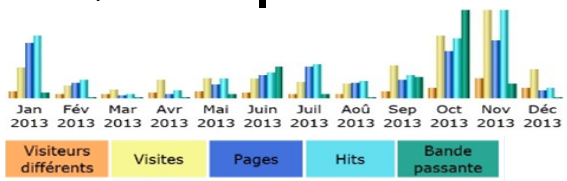
Belle augmentation des citations Galaxy dans les publications

Statistiques de l'instance Galaxy Sigenae / BioInfo Genotoul

http://galaxy-workbench.toulouse.inra.fr

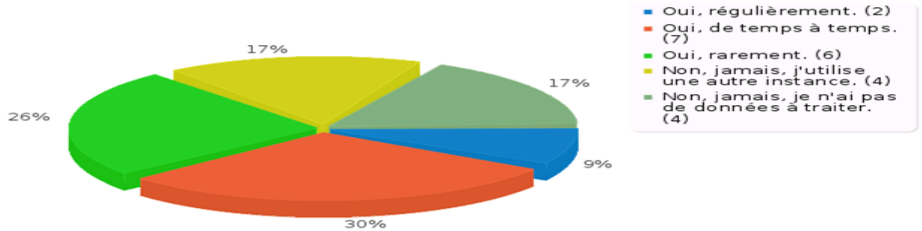
23 réponses sur ~40 utilisateurs

Historique mensuel des visites



~40 utilisateurs
Les 10 plus gros utilisateurs (hors tests Sigenae) utilisent Galaxy dans le cadre de leur projet.

Utilisez-vous l'instance Sigenae de Galaxy ?



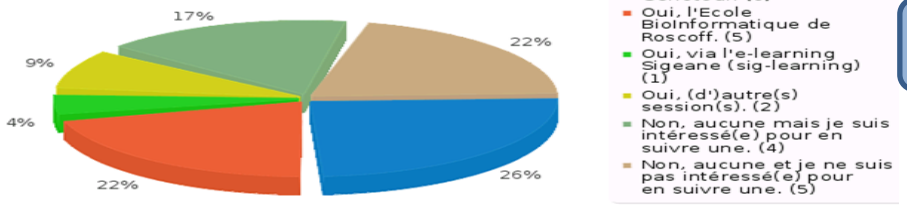
65% des utilsateurs Galaxy ayant répondus à l'enquête utilisent l'instance Sigenae

L'instance Sigenae est intégrée au groupe de travail Galaxy IFB France http://www.ifb-galaxy.org/

Formation & utilisation de Galaxy

Importance de la formation et du e-learning

Avez-vous déjà suivi une session de formation Galaxy ?



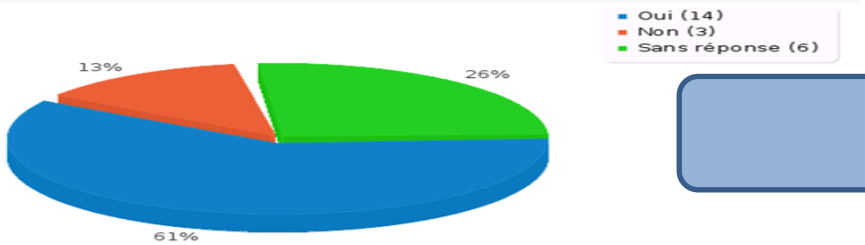
6% des utilisateurs seulement ont déjà suivi une formation en e-learning.

60% des utilisateurs :

•ont déjà suivi une formation Galaxy -> Besoin d'une 'mise en route',
•souhaitent plus de support en bioinfo.
•créent et utilisent des workflows.

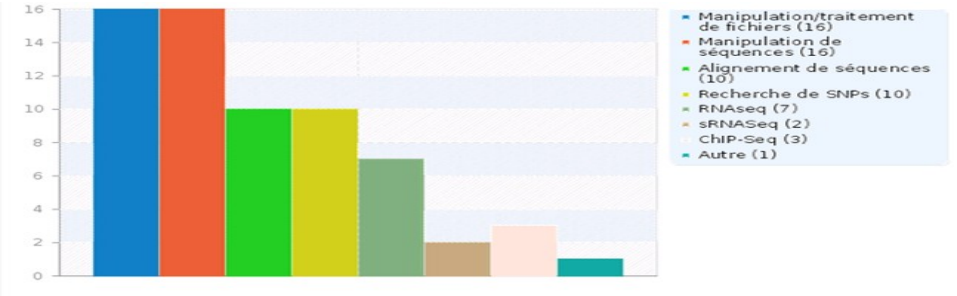
Les traitements Galaxy sont aussi lancés dans le cadre de projets

Galaxy vous permet-il de traiter vos données bio-informatiques ?



~60% traitent leurs données dans Galaxy.
… Et 60% de ces 60% utilisent ces données dans leur projet.

Quels sont les principaux outils utilisés ?



Principalement des applications de type SNP et RNAseq

Contact : sigenae-support@listes.inra.fr

Les formations avec Galaxy

http://galaxy-workbench.toulouse.inra.fr

Formations Galaxy ou utilisant Galaxy

- Formations PF BioInfo Genotoul
- UPS Master 2 BioInformatique

Galaxy training :

- Galaxy : First step
- Galaxy : Reads alignment and SNP calling
- Galaxy : RNAseq alignment and transcripts assemblies
- Galaxy : sRNAseq
- Galaxy : metagenomic : sequence analysis of 16S DNA reads

http://bioinfo.genotoul.fr

Training calendar :

The table below gives you information on the training subject, session, date, and the number of remaining free places. To register **fill out the registration form** (Unix, Cluster, ...). You will then be in a **waiting list** and you will be informed if you have been selected for the next session or scheduled in an other one.

Title	Date	Organizers	Vacancies	Registration
Metagenomic: 16S DNA reads	05.02.2014	Jerome Mariette, Christophe Klopp, Maria Bernard, Sarah Maman, Laurent Cauquil	fully booked	
Unix / Linux	07.04.2014	Didier Laborie	3	Register now
Cluster	08.04.2014	Didier Laborie, Celine Noirot	5	Register now
Reads alignment and SNP calling	09.04.2014	Jerome Mariette, Philippe Bardou	3	Register now
RNAseq alignment and transcripts assemblies	10.04.2014	Celine Noirot, Delphine Labourdette	7	Register now
Galaxy training days	12.-14.05.2014	Jerome Mariette, Philippe Bardou, Sarah Maman, Celine Noirot, Delphine Labourdette	fully booked	Register for the waiting list (0 registrations on the waiting list)

Formations Galaxy ou utilisant Galaxy

aviesan
alliance nationale
pour les sciences de la vie et de la santé

Objectifs & Missions | Instituts thématiques multi-organismes | Membres d'Aviesan | Actions & Initiatives

Accueil › Toute l'actualité › Ecole de bioinformatique :... Taille du texte English

Ecole de bioinformatique : Initiation au traitement des données de génomique obtenues par séquençage à haut débit

Ecoutez 14 jan 2013

Du 14 au 18 Janvier 2013 à la Station Biologique de Roscoff
Cette école organisée par l'ITMO Génétique, génomique et bioinformatique avec le support de 9 plateformes de bioinformatique, accueillera 40 participants.

Consulter la plaquette de l'école (397,7 ko)

Contact : ecole-bioinfo@aviesan.fr
Consulter le programme (école bioinformatique)

- Pré-traitement et alignement des données exome-seq (M. Bernard, R. Daveau, S. Gallina, E. Girard, B. Job, Y. Luo, N. Servant)
- Détection et annotation fonctionnelle de variants (M. Bernard, R. Daveau, S. Gallina, E.Girard, B. Job, Y. Luo, S. Marthey, N. Servant)
- Prédiction de miRNA aves MirDeep2 et autres outils. Annotation des reads : miRNA et autres ncRNA (C. Gaspin, S. Maman, O. Rué)
- Analyse des données des participants.

Les liens avec les communautés et groupes de travail

Galaxy Project

Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences

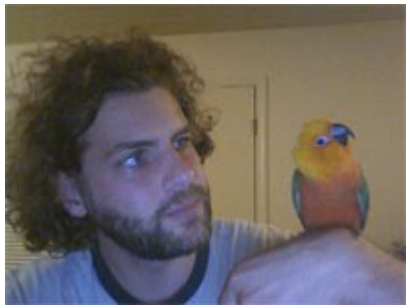
Jeremy Goecks[1], Anton Nekrutenko[2,*], James Taylor[1,*] and The Galaxy Team

Equipe "Galaxy project" :

- Le Center for Comparative Genomics and Bioinformatics - Penn State,
- Des départements "Biology" et "Mathematics and Computer Science" de l'Université d'Emory.

Anton Nekrutenko
Penn State

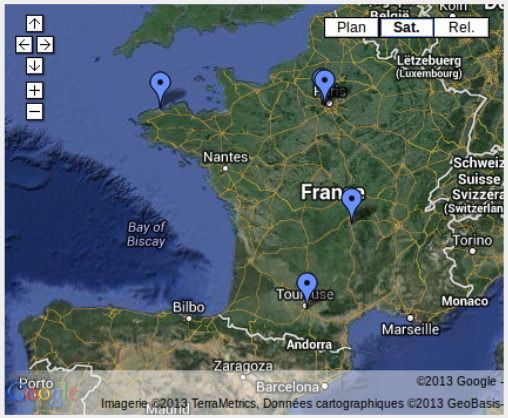
Nate Coraor
Penn State



James Taylor
Emory

Groupe de travail GALAXY IFB http://www.ifb-galaxy.org/

- ✓Documentation collaborative (wiki)
- ✓Formations (mise en commun agenda PF)
- ✓Architecture
- ✓Intégration d'outils (Tool Shed)



Afficher Galaxy IFB France sur une carte plus grande

Liste des instances

ABiMS Roscoff	Initiation, NGS Cleaning, RNASeq Differential Expression	http://galaxy.sb-roscoff.fr/	Christophe Caron - Alexandre Cormier - Gildas Lecorguille - Pierre Pericard
Institut Curie	ChIP-Seq Analysis	http://nebula.curie.fr/	Alban Lermine
Genotoul / Sigenae	Initiation to Galaxy, SNP calling, RNASeq, sRNASeq	http://galaxy-workbench.toulouse.inra.fr/	Sarah Maman
INRA URGI	Differential expression analysis, Variant detection	http://urgi.versailles.inra.fr/galaxy2	Olivier Inizan
INRA MIGALE	Initiation to Galaxy, NGS Galaxy	http://migale.jouy.inra.fr/galaxy/.	Sandra Derozier - Franck Samson
Southgreen	Generalist platform, and crop breeding	gohelle.cirad.fr/galaxy/root/	Jean-Francois Dufayard
INRA PFEM / MetaboHUB	Metabolomics data analysis	https://pfem-galaxy/	Franck Giacomoni

Merci à toutes celles / tous ceux qui travaillent sur Galaxy !

✓**Les wrappers boys & girls**
Maria Bernard
Frédéric Escudié
Ibouniyamine Nabihoudine
Olivier Rué
Sabrina Rodriguez

✓**Les admins**
Patrice Dehais
Didier Laborie
Marie-Stephane Trotard

✓**Les formateurs**
Philippe Bardou
Christophe Klopp
Delphine Labourdette
Jérôme Mariette
Céline Noirot

✓**Les personnes ayant utilisés Galaxy pour leurs projets**
Maria Bernard
Sandrine Laguerre
Olivier Monestier
Jérôme Montfort

…

Merci pour votre écoute

sigenae-support@listes.inra.fr