

A panel of European Public Galaxy Instances

**ECCB Demo track
Strasbourg, September 8/9 2014.**

Part I

this Demo is presented by

- Sep. 8 { **Part I** **Hans-Rudolf Hotz**
Friedrich Miescher Institute for Biomedical Research
- Part II** **Björn Grüning**
Albert Ludwigs Universität Freiburg

- Sep. 9 { **Part III** **Alban Lermine**
Institut Curie
- Part IV** **Christophe Blanchet**
French Institute of Bioinformatics



Friedrich Miescher Institute
for Biomedical Research

A panel of European Public Galaxy Instances

Sep. 8 { **Part I** Galaxy Project Introduction

Part II deepTools: user-friendly tools for data visualization

Sep. 9 { **Part III** The Nebula ChIP-Seq web service

Part IV The French Institute of Bioinformatics academic cloud infrastructure for life sciences

the Galaxy Project



Friedrich Miescher Institute
for Biomedical Research

Galaxy is an open,
web-based platform for
accessible, reproducible, and transparent
computational biomedical research.

[*http://galaxyproject.org/*](http://galaxyproject.org/)

FMI

Friedrich Miescher Institute
for Biomedical Research

the Galaxy Project



Accessible: Users without programming experience can easily specify parameters and run tools and workflows.

Reproducible: Galaxy captures information so that any user can repeat and understand a complete computational analysis.

Transparent: Users share and publish analyses via the web and create Pages, interactive, web-based documents that describe a complete analysis.

<http://galaxyproject.org/>

FMI

Friedrich Miescher Institute
for Biomedical Research

using Galaxy



Galaxy's public web site

The screenshot shows the Galaxy public web site interface. At the top, there is a navigation bar with the Galaxy logo and a search bar. Below the navigation bar, the main content area is divided into three sections: Tools, History, and Tweets. The Tools section on the left lists various tools such as Get Data, Lift-Over, Text Manipulation, Convert Formats, FASTA manipulation, Filter and Sort, Join, Subtract and Group, Extract Features, Fetch Sequences, Fetch Alignments, Get Genomic Scores, Operate on Genomic Intervals, Statistics, Graph/Display Data, Regional Variation, Multiple regression, Multivariate Analysis, Evolution, Motif Tools, Multiple Alignments, Metagenomic analyses, Genome Diversity, NGS TOOLBOX BETA, Phenotype Association, NGS: QC and manipulation, NGS: Mapping, and NGS: SAM Tools. The History section on the right shows a list of recent jobs, including '4: Select on data 2', '3: Select on data 1', '2: BED-to-GFF on data 1', and '1: UCSC Main on Human: knownGene (chr1:1-249250621)'. The Tweets section in the center displays tweets from GigaScience and the Galaxy Project. At the bottom of the page, there are logos for Penn State, Johns Hopkins University, TACC, and iPlant Collaborative.

Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy [start here](#) or consult our [help resources](#).

Tools

- Get Data
- Lift-Over
- Text Manipulation
- Convert Formats
- FASTA manipulation
- Filter and Sort
- Join, Subtract and Group
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Regional Variation
- Multiple regression
- Multivariate Analysis
- Evolution
- Motif Tools
- Multiple Alignments
- Metagenomic analyses
- Genome Diversity
- NGS TOOLBOX BETA
- Phenotype Association
- NGS: QC and manipulation
- NGS: Mapping
- NGS: SAM Tools

History

- Unnamed history
- 5.7 MB
- 4: Select on data 2
- 3: Select on data 1
- 2: BED-to-GFF on data 1
- 1: UCSC Main on Human: [knownGene \(chr1:1-249250621\)](#)

Tweets

GigaScience @GigaScience 3h
#openscience folks in Melbourne, sign up for our free workshop next month
[eventbrite.com.au/e/the-great-gi...](#)
For more see:
[blogs.biomedcentral.com/bmcblog/20](#)
Retweeted by Galaxy Project
Expand

Galaxy Project @galaxyproject 1h
September 2014 Newsletter: Lots of new releases, Galaxy-UK, upcoming events, plus new jobs, papers, tools
[bit.ly/1pphMf1](#) #usegalaxy
Expand

Tweet to @galaxyproject

Penn State **Johns Hopkins University** **TACC** **iPlant Collaborative**

<http://usegalaxy.org>

FMI

Friedrich Miescher Institute
for Biomedical Research

using Galaxy



Analyze Data
Workflow
Shared Data
Visualization
Cloud
Help
User

Using 0%

chr1	11873	14409	uc001aaa.3	0	+	11873	11873	0	3	354,109,1189,
chr1	11873	14409	uc010nrx.1	0	+	11873	11873	0	3	354,52,1189,
chr1	11873	14409	uc010nxq.1	0	+	12189	13639	0	3	354,127,1007,
chr1	14361	16765	uc009vis.3	0	-	14361	14361	0	4	468,69,147,159,
chr1	14361	19759	uc009vit.3	0	-	14361	14361	0	9	468,69,152,159,198,510,147,99,847,
chr1	14361	19759	uc009viu.3	0	-	14361	14361	0	10	468,69,152,159,198,510,147,102,54,847,
chr1	14361	19759	uc001aae.4	0	-	14361	14361	0	10	468,69,152,159,198,136,137,147,99,847,
chr1	14361	29370	uc001aah.4	0	-	14361	14361	0	11	468,69,152,159,198,136,137,147,99,154,5,
chr1	14361	29370	uc009vir.3	0	-	14361	14361	0	10	468,69,152,159,198,510,147,99,154,50,
chr1	14361	29370	uc009viq.3	0	-	14361	14361	0	7	468,152,159,198,456,154,50,
chr1	14361	29370	uc001aac.4	0	-	14361	14361	0	11	468,69,152,159,198,110,137,147,102,154,
chr1	14406	29370	uc009viv.2	0	-	14406	14406	0	7	2359,198,136,137,147,154,50,
chr1	14406	29370	uc009viw.2	0	-	14406	14406	0	7	2359,198,510,147,99,154,50,
chr1	15602	29370	uc009vix.2	0	-	15602	15602	0	7	345,159,198,136,147,154,50,
chr1	15795	18061	uc009vjd.2	0	-	15795	15795	0	5	152,159,198,136,456,
chr1	16606	29370	uc009viy.2	0	-	16606	16606	0	9	159,198,136,137,147,95,58,154,50,
chr1	16606	29370	uc009viz.2	0	-	16606	16606	0	8	159,202,136,137,147,112,154,50,
chr1	16857	17751	uc009vjc.1	0	-	16857	16857	0	2	198,519,
chr1	16857	19759	uc001aai.1	0	-	16857	16857	0	6	198,136,137,147,112,847,
chr1	16857	29370	uc010nxs.1	0	-	16857	16857	0	8	198,136,137,147,99,227,154,50,
chr1	16857	29961	uc009vjb.1	0	-	16857	16857	0	7	198,136,137,147,112,154,138,
chr1	17232	29370	uc009vje.2	0	-	17232	17232	0	4	510,147,99,50,
chr1	17605	29370	uc009vjf.2	0	-	17605	17605	0	7	137,147,95,58,227,154,50,
chr1	34610	36081	uc001aak.3	0	-	34610	34610	0	3	564,205,361,

History

Unnamed history
5.7 MB



1: UCSC Main on Human:
knownGene
(chr1:1-249250621)
7,967 regions
format: bed, database: hg19
display in IGB View
display at Ensembl Current
display at RViewer main
display at UCSC main



FMI


Friedrich Miescher Institute
for Biomedical Research

using Galaxy



BED-to-GFF (version 2.0.0) Help from Biostar  

Convert this dataset:  

1: UCSC Main on Human: knownGene (chr1:1-249250621) 

Execute

FMI

Friedrich Miescher Institute
for Biomedical Research

using Galaxy



BED-to-GFF (version 2.0.0) Help from Biostar

Convert this dataset:

1: UCSC Main on Human: knownGene (chr1:1-249250621)

Execute

Seqname	Source	Feature	Start	End	Score	Strand	Frame	Group
##gff-version 2								
##bed_to_gff_converter.py								
chr1	bed2gff	mRNA	11874	14409	0	+	.	mRNA uc001aaa.3;
chr1	bed2gff	exon	11874	12227	0	+	.	exon uc001aaa.3;
chr1	bed2gff	exon	12613	12721	0	+	.	exon uc001aaa.3;
chr1	bed2gff	exon	13221	14409	0	+	.	exon uc001aaa.3;
chr1	bed2gff	mRNA	11874	14409	0	+	.	mRNA uc010nxr.1;
chr1	bed2gff	exon	11874	12227	0	+	.	exon uc010nxr.1;
chr1	bed2gff	exon	12646	12697	0	+	.	exon uc010nxr.1;
chr1	bed2gff	exon	13221	14409	0	+	.	exon uc010nxr.1;
chr1	bed2gff	mRNA	11874	14409	0	+	.	mRNA uc010nxq.1;
chr1	bed2gff	exon	11874	12227	0	+	.	exon uc010nxq.1;
chr1	bed2gff	exon	12595	12721	0	+	.	exon uc010nxq.1;

History

Unnamed history

5.7 MB

2: BED-to-GFF on data 1



1: UCSC Main on Human: knownGene (chr1:1-249250621)



FMI

Friedrich Miescher Institute
for Biomedical Research

using Galaxy




Select (version 1.0.1) Help from Biostar  

Select lines from:  

1: UCSC Main on Human: knownGene (chr1:1-249250621)

2: BED-to-GFF on data 1

that:

Matching 

the pattern:

uc010nxs

here you can enter text or regular expression (for syntax check lower part of this frame)

Execute


FMI

Friedrich Miescher Institute
for Biomedical Research


using Galaxy




Seqname	Source	Feature	Start	End	Score	Strand	Frame	Group
chr1	bed2gff	mRNA	16858	29370	0	-	.	mRNA uc010nxs.1;
chr1	bed2gff	exon	16858	17055	0	-	.	exon uc010nxs.1;
chr1	bed2gff	exon	17233	17368	0	-	.	exon uc010nxs.1;
chr1	bed2gff	exon	17606	17742	0	-	.	exon uc010nxs.1;
chr1	bed2gff	exon	17915	18061	0	-	.	exon uc010nxs.1;
chr1	bed2gff	exon	18268	18366	0	-	.	exon uc010nxs.1;
chr1	bed2gff	exon	18913	19139	0	-	.	exon uc010nxs.1;
chr1	bed2gff	exon	24738	24891	0	-	.	exon uc010nxs.1;
chr1	bed2gff	exon	29321	29370	0	-	.	exon uc010nxs.1;

History  


Unnamed history

5.7 MB  

4: Select on data 2   

3: Select on data 1   

2: BED-to-GFF on data 1   




1: UCSC Main on Human: knownGene (chr1:1-249250621)   




FMI




Friedrich Miescher Institute
for Biomedical Research

a few more highlights









4: Select on data 2   

3: Select on data 1   

2: BED-to-GFF on data 1   




80,918 lines, 3 comments
format: **gff**, database: **hg19**

2967 lines converted to GFF version 2.

display at Ensembl [Current](#)
display at UCSC [main](#)

1. Seqname	2. Source	3. Feature	4. Score
##gff-version 2			
##bed_to_gff_converter.py			
chr1	bed2gff	mRNA	118
chr1	bed2gff	exon	118
chr1	bed2gff	exon	126

1: UCSC Main on Human: knownGene   
([chr1:1-249250621](#))




re-run a tool with
the same parameters
(or different parameters)









Friedrich Miescher Institute
for Biomedical Research

a few more highlights








4: Select on data 2   

3: Select on data 1   

2: BED-to-GFF on data 1   




80,918 lines, 3 comments
format: **gff**, database: **hg19**

2967 lines converted to GFF version 2.

display at Ensembl Current
display at UCSC main

1. Seqname	2. Source	3. Feature	4. Score
##gff-version 2			
##bed_to_gff_converter.py			
chr1	bed2gff	mRNA	118
chr1	bed2gff	exon	118
chr1	bed2gff	exon	126

1: UCSC Main on Human: knownGene   
(chr1:1-249250621)

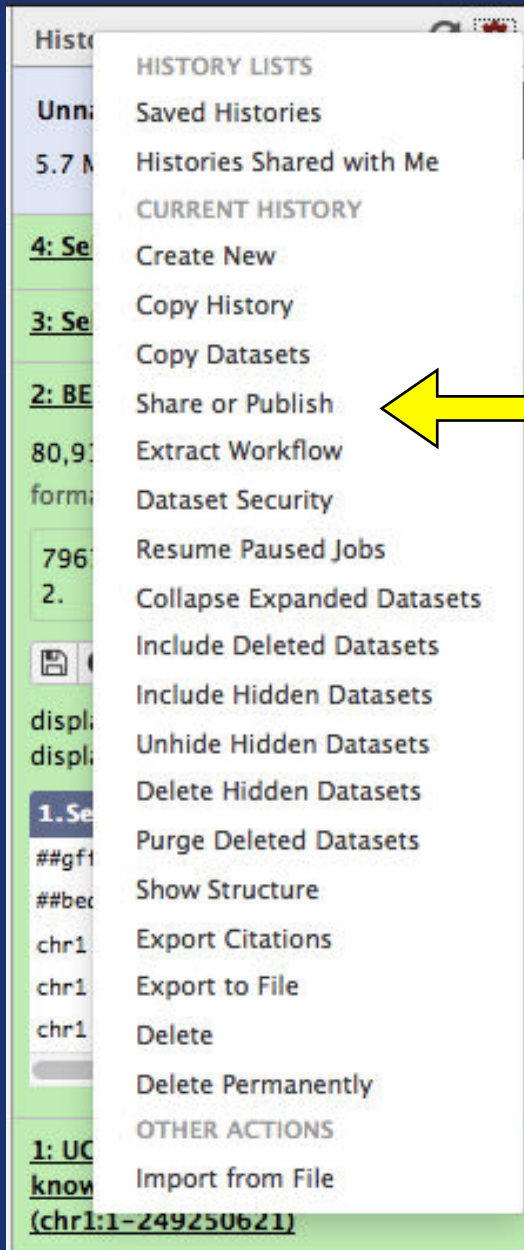
re-run a tool with
the same parameters
(or different parameters)

display in the built in or
a web genome browser

FMI

Friedrich Miescher Institute
for Biomedical Research

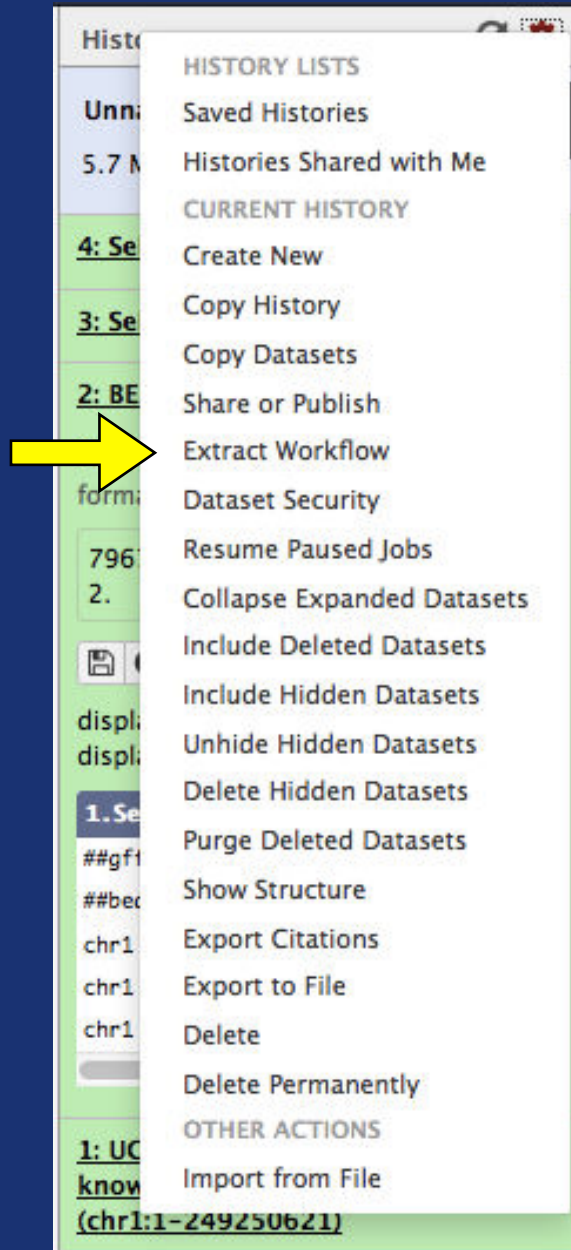
a few more highlights



FMI

Friedrich Miescher Institute
for Biomedical Research

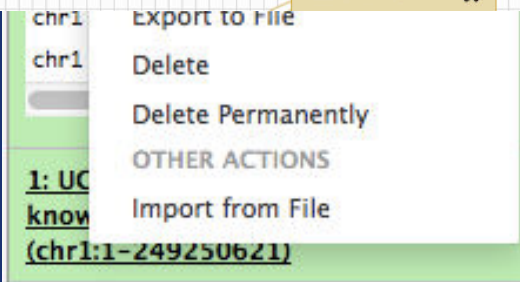
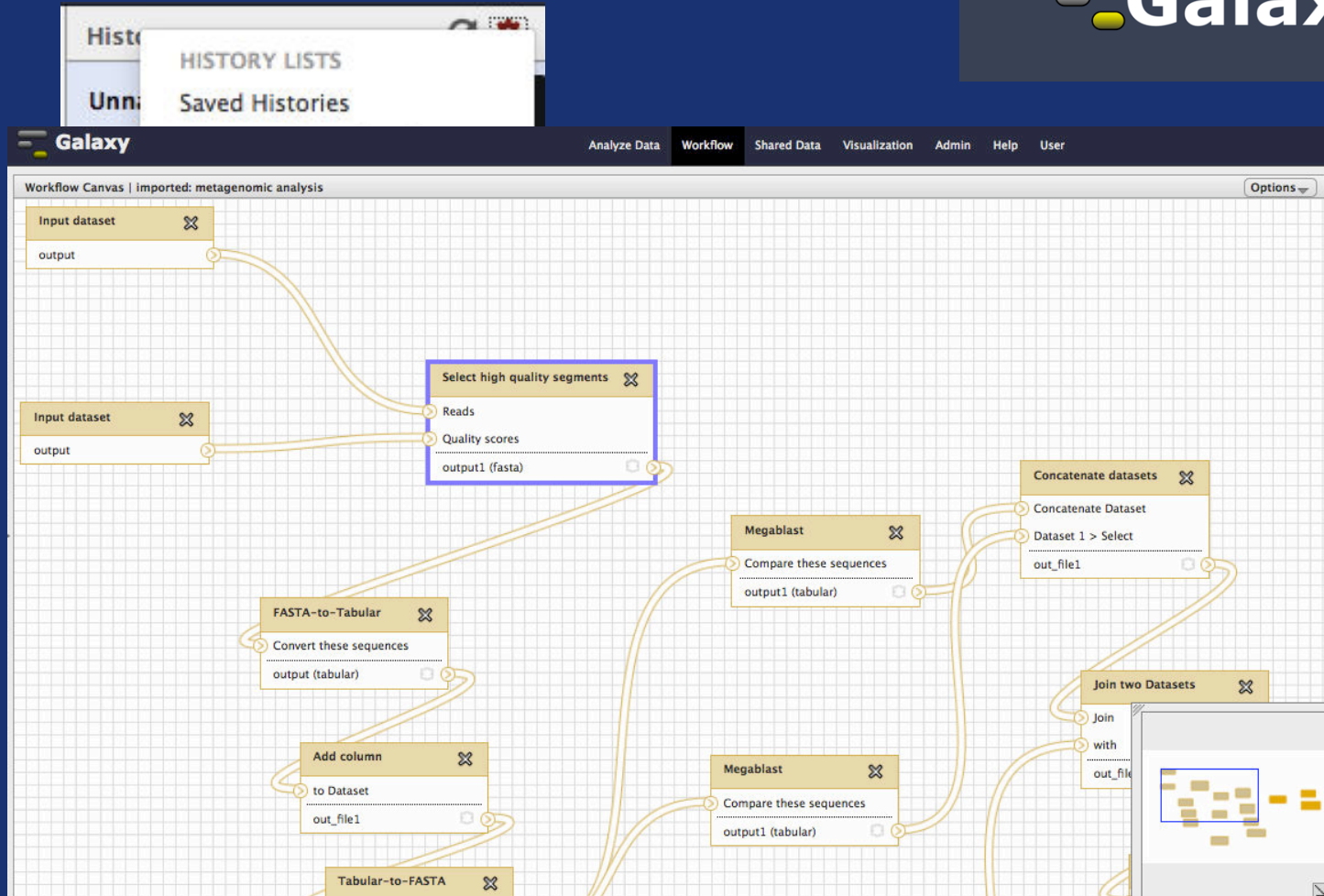
a few more highlights



FMI

Friedrich Miescher Institute
for Biomedical Research

a few more highlights



FMI

Friedrich Miescher Institute
for Biomedical Research

<https://wiki.galaxyproject.org/Learn>

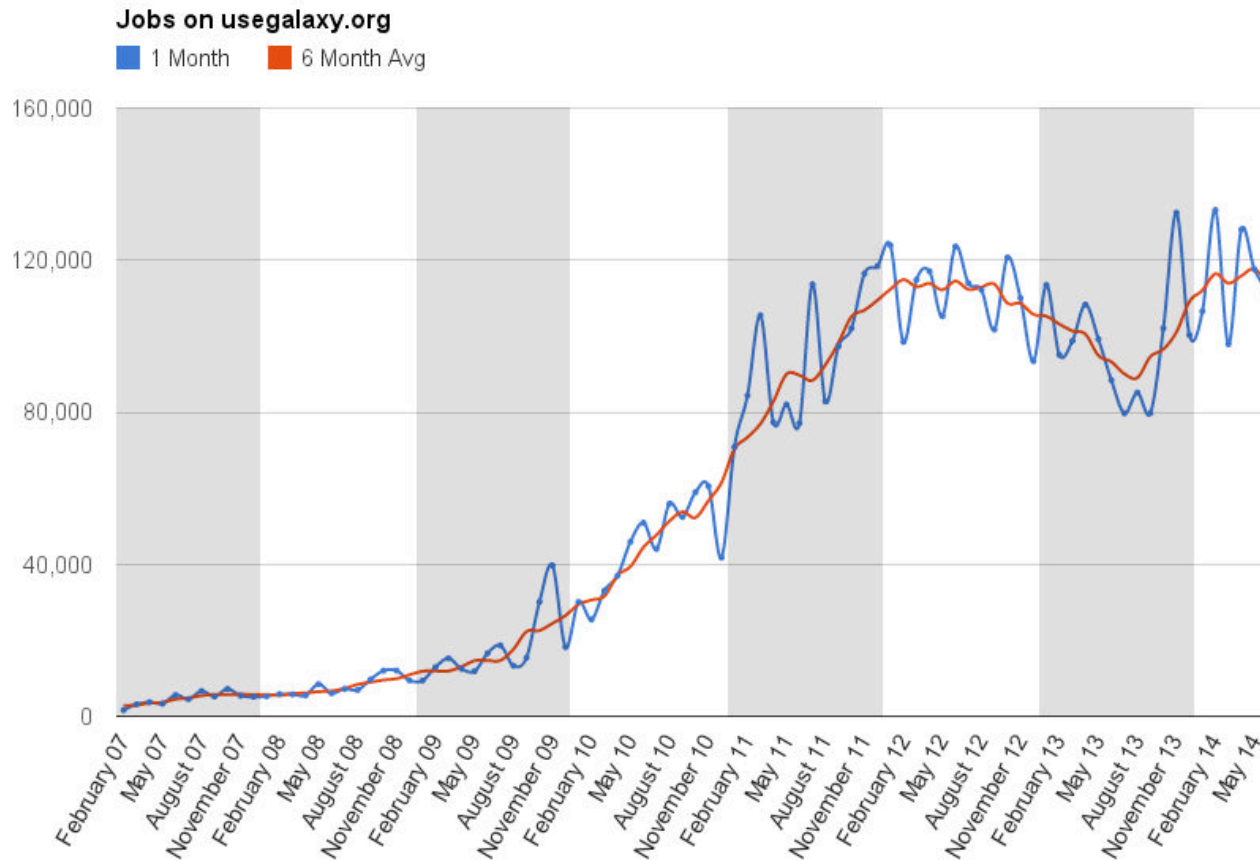
tutorials

screencats

shared histories & workflows

<https://wiki.galaxyproject.org/CitingGalaxy>

and Galaxy is used



FMI

Friedrich Miescher Institute
for Biomedical Research

Galaxy is a community



FMI

Friedrich Miescher Institute
for Biomedical Research

Galaxy is a community



in addition to *<http://usegalaxy.org>* there are 60 public instances offered by institutes from all over the world

FMI

Friedrich Miescher Institute
for Biomedical Research

Galaxy is a community



in addition to *<http://usegalaxy.org>* there are 60 public instances offered by institutes from all over the world

plus an unknown number of private Galaxy servers

FMI

Friedrich Miescher Institute
for Biomedical Research

Galaxy is a community



in addition to *<http://usegalaxy.org>* there are 60 public instances offered by institutes from all over the world

plus an unknown number of private Galaxy servers

~4000 code commits from over 50 people per year

FMI

Friedrich Miescher Institute
for Biomedical Research

2570 'tools' published from ~220 developers

<u>Name</u>	<u>Description</u>	<u>Repositories</u>
<u>Assembly</u>	Tools for working with assemblies	33
<u>ChIP-seq</u>	Tools for analyzing and manipulating ChIP-seq data.	7
<u>Computational chemistry</u>	Tools for use in computational chemistry	23
<u>Convert Formats</u>	Tools for converting data formats	30
<u>Data Managers</u>	Utilities for Managing Galaxy's built-in data cache	6
<u>Data Source</u>	Tools for retrieving data from external data sources	17
<u>Fasta Manipulation</u>	Tools for manipulating fasta data	41
<u>Fastq Manipulation</u>	Tools for manipulating fastq data	17
<u>Genome-Wide Association Study</u>	Utilities to support Genome-wide association studies	1
<u>Genomic Interval Operations</u>	Tools for operating on genomic intervals	37
<u>Graphics</u>	Tools producing images	22
<u>Imaging</u>	Utilities to support imaging	1
<u>Metabolomics</u>	Tools for use in the study of Metabolomics	3
<u>Metagenomics</u>	Tools enabling the study of metagenomes	23
<u>Micro-array Analysis</u>	Tools for performing micro-array analysis	8

2570 'tools' published from ~220 developers

Name	Description	Repositories
Assembly	Tools for working with assemblies	33
ChIP-seq	Tools for analyzing and manipulating ChIP-seq data.	7
Computational chemistry		
Convert Formats		
Data Managers		
Data Source	Tools for retrieving data from external data sources	17
Fasta Manipulation	Tools for manipulating fasta data	41
Fastq Manipulation		
Genome-Wide Association Study		
Genomic Interval Operations	Tools for operating on genomic intervals	37
Graphics	Tools producing images	22
Imaging	Utilities to support imaging	1
Metabolomics	Tools for use in the study of Metabolomics	3
Metagenomics	Tools enabling the study of metagenomes	23
Micro-array Analysis	Tools for performing micro-array analysis	8

Galaxy wrappers for existing command line tools

Tools especially developed for Galaxy

deploy Galaxy



Galaxy is a free and open source project.

install Galaxy locally

use Galaxy on the Cloud

get the SlipStream Galaxy appliance

FMI

Friedrich Miescher Institute
for Biomedical Research

Demo / Training Day
30 Sep 2014, Bern

2nd Swiss Workshop
1 Oct 2014, Bern

German Developer Day
2 Oct 2014, Freiburg



<https://wiki.galaxyproject.org/Events/Switzerland2014>

<https://wiki.galaxyproject.org/Events/Germany2014>

next:



Björn Grüning: "deepTools"

and please come back tomorrow

FMI

Friedrich Miescher Institute
for Biomedical Research