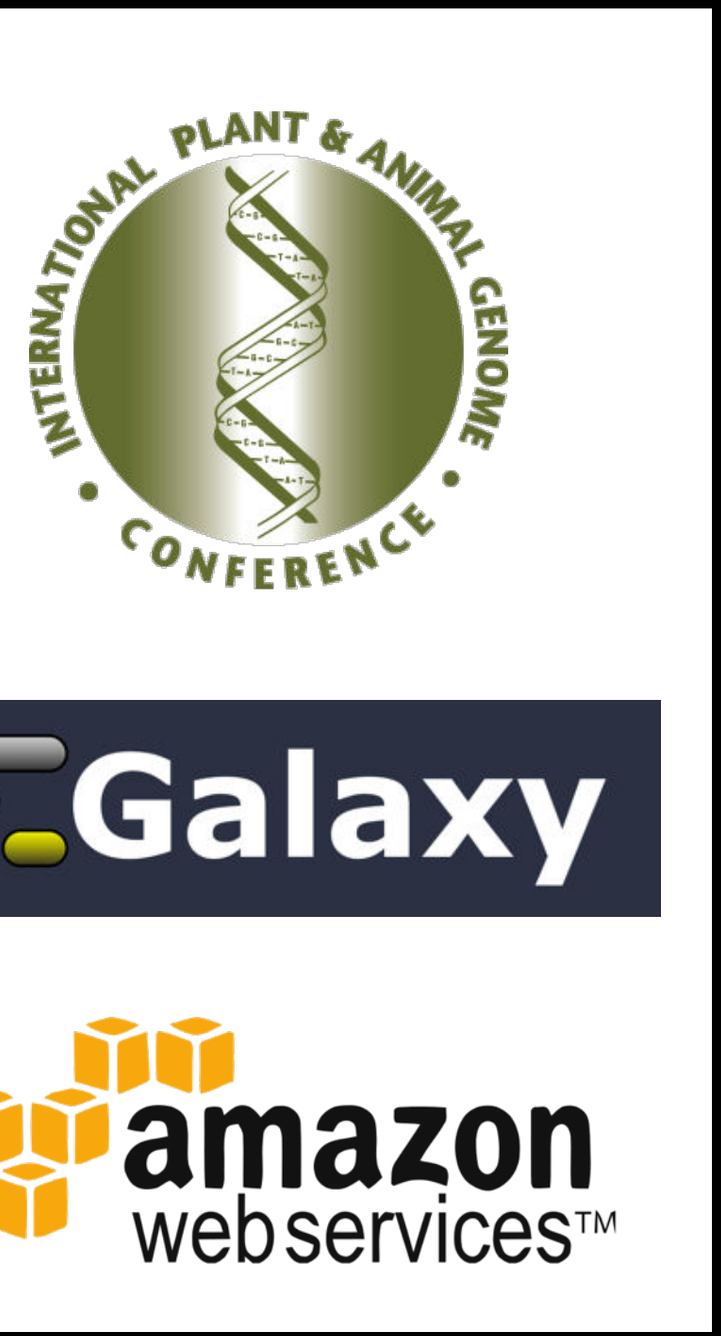
Galaxy for SNP and Variant Data Analysis

Plant and Animal Genome XXIII
San Diego, California, United States
January 13, 2015

Dave Clements
Galaxy Project
Johns Hopkins University

Galaxy Team
http://bit.ly/gxyteam



Outline

4:00 Introduction: The Galactic Landscape

4:10 Variant Analysis: A worked example

5:40 Variant Analysis: Other options

5:50 Galaxy: Resources and Community

6:10 Done

Slides at http://bit.ly/

Goals

Provide a basic introduction to using Galaxy for bioinformatic analysis using SNP calling as the driving example.

Demonstrate how Galaxy can help you explore and learn options, perform analysis, and then share, repeat, and reproduce your analyses.

Not Goals

This workshop will *not* cover

- details of how tools are implemented, or
- new algorithm designs, or
- which SNP caller or mapper or peak caller or ... is best for you.

While this workshop does cover Galaxy you won't become a Galaxy expert in the next two hours.

What is Galaxy?

Data integration and analysis platform that emphasizes accessibility, reproducibility, and transparency

A free (for everyone) web server

Open source software

These options result in several ways to use Galaxy

http://galaxyproject.org

Galaxy is available ...

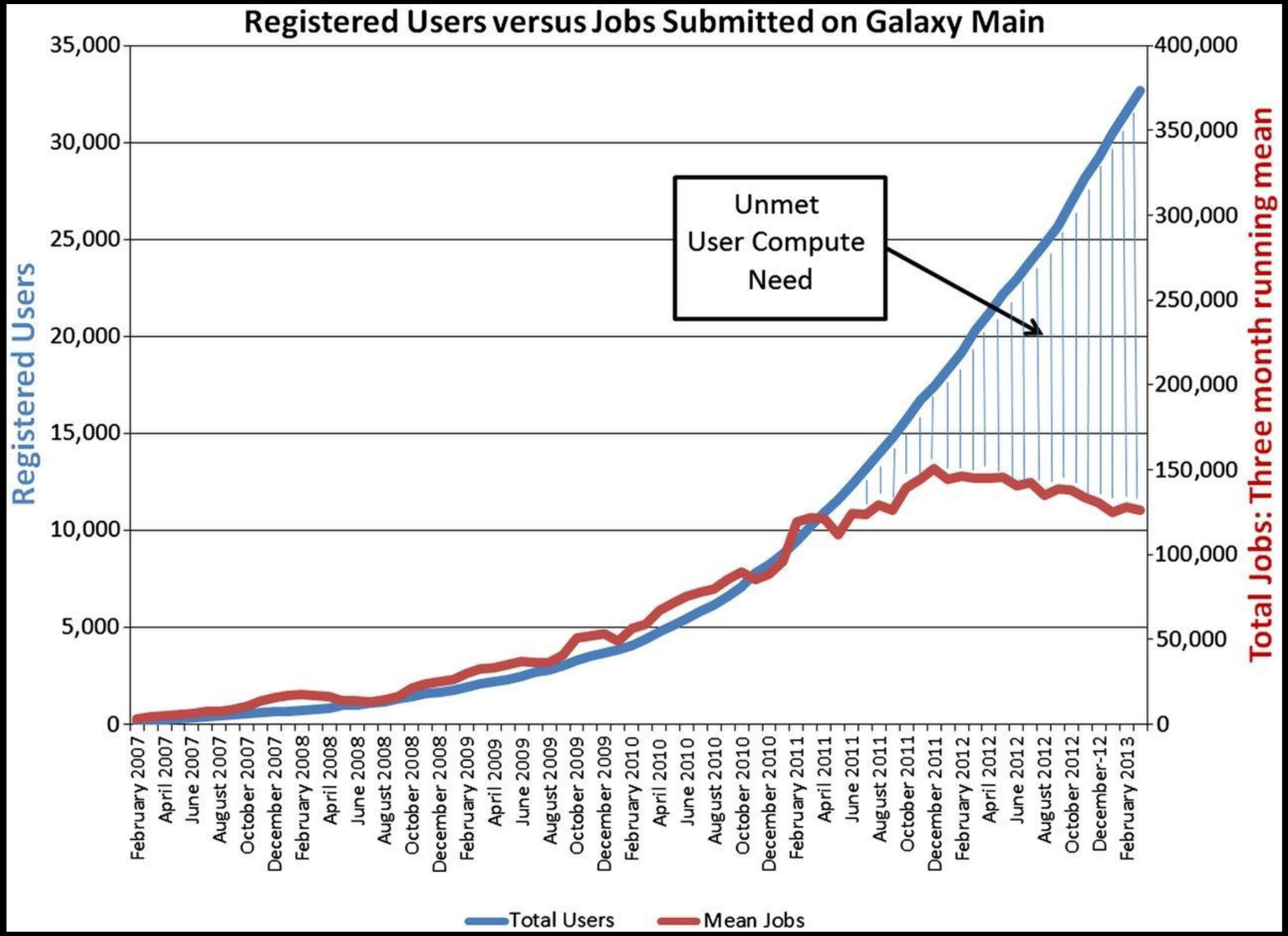
As a free (for everyone) web server integrating a wealth of tools, compute resources, petabytes of reference data and permanent storage

http://usegalaxy.org

TACC

iPlant Collaborative™

However, ***a centralized solution cannot support the different analysis needs of the entire world.***



Leveraging the national cyberinfrastructure for biomedical research

LeDuc, *et al. J Am Med Inform Assoc doi:10.1136/amiajnl-2013-002059*

Galaxy is available ...

- As a free (for everyone) web service

  http://usegalaxy.org

- As open source software

  http://getgalaxy.org

It is installed in locations around the world

Galaxy is available ...

http://aws.amazon.com/education
http://globus.org/
http://wiki.galaxyproject.org/Cloud

We are using the cloud today.

Galaxy is available: With Commercial Support

A ready-to-use appliance
(BioTeam)

Cloud-based solutions
(ABgenomica, AIS, GenomeCloud)

Consulting & Customization
(Arctix, BioTeam, Deena Bioinformatics)

SLIPSTREAM APPLIANCE GALAXY EDITION
Galaxy made easy.

abgenomica

GenomeCloud

AIS™
ENTERPRISE-CLASS DATA CENTER SERVICES

Galaxy Project: Further reading & Resources

http://galaxyproject.org

http://usegalaxy.org

http://getgalaxy.org

http://wiki.galaxyproject.org/Cloud

http://bit.ly/gxychoices

Outline

4:00 Introduction: The Galactic Landscape

4:10 Variant Analysis: A worked example

5:40 Variant Analysis: Other options

5:50 Galaxy: Resources and Community

6:10 Done

SNP Calling

Basic SNP Calling Exercise
Tutorial created by Dan Bolser
of EMBL-EBI

Datasets from Institute of
Population Genetics, University
of Veterinary Medicine, Vienna
(Vetmeduni Vienna)

http://bit.ly/bolserSNP2014

Variant Calling: A General Plan

- Get to Galaxy
- Get data to look for variation in
- Get some reference
- Align reads to reference
- Call variants

http://bit.ly/bolserSNP2014

Get started with Galaxy

usegalaxy.org

http://bit.ly/bolserSNP2014

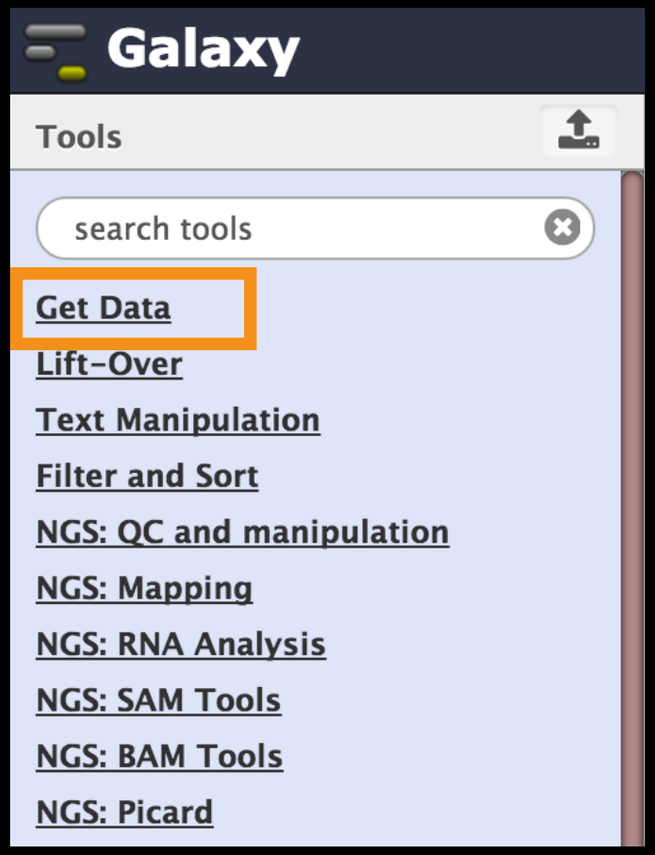
Variant Calling: A General Plan

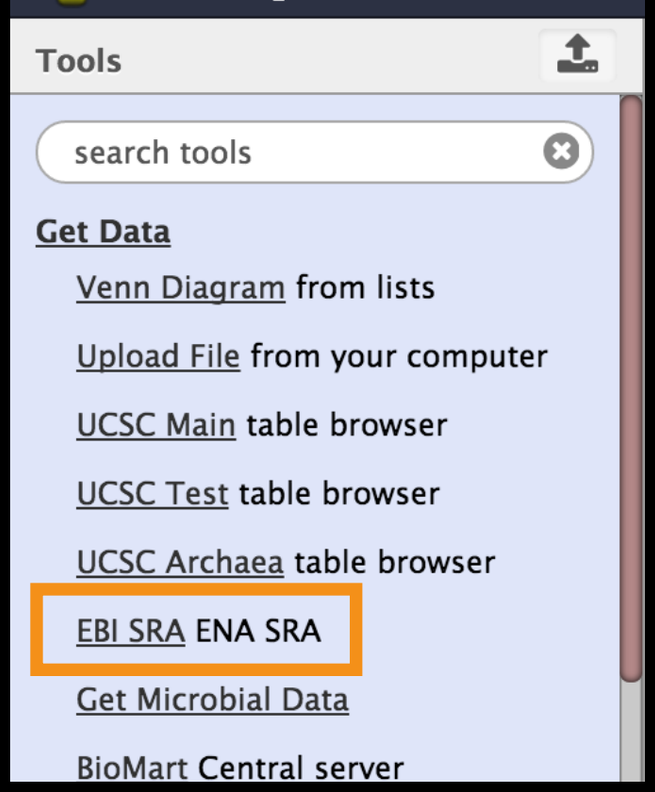
- Get to Galaxy
- Get data to look for variants in
- Align reads to reference
- Call variants

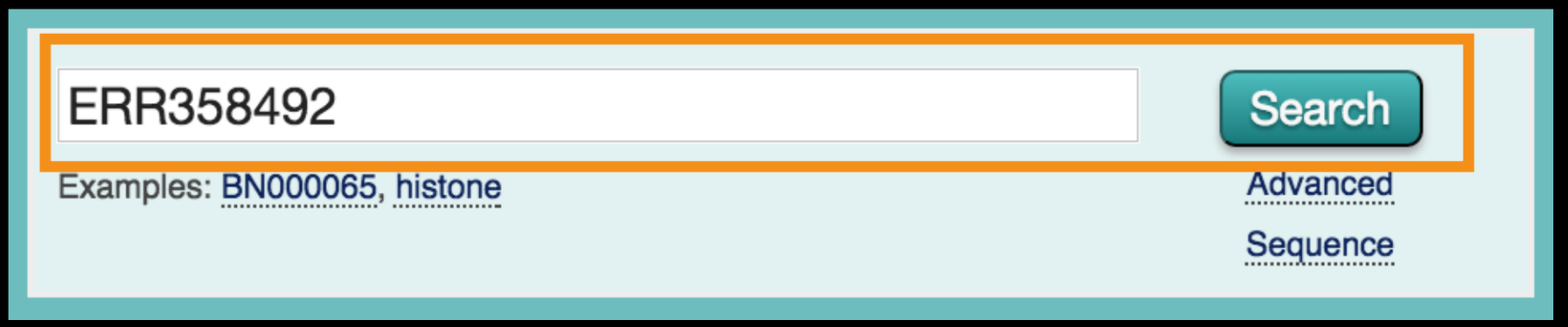
http://bit.ly/bolserSNP2014

Get Some Data *(but don't)*

Using ENA read set ERR358492







SNP Calling: Get some data

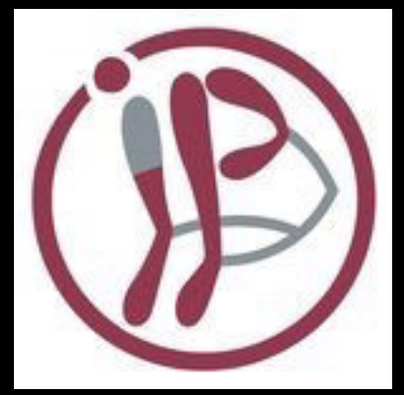
ENA read set ERR358492

Wolbachia endosymbiont of Drosophila melanogaster; wol.F15ColdReplicate d(7)

Illumina HiSeq 2000 paired end sequencing

http://www.ebi.ac.uk/ena/data/view/ERR358492

http://bit.ly/ERR358492



Wolbachia endosymbiont of Drosophila melanogaster; wol.F15ColdReplicate d(7)

Illumina HiSeq 2000 paired end sequencing

Illumina HiSeq 2000 paired end sequencing

View: XML

Send Feedback

Download: XML

http://bit.ly/ERR358492

Submitting Centre	Run Date	Platform	Model	Read Count	Base Count
Institute of Population Genetics, University of Veterinary Medicine, Vienna (Vetmeduni Vienna)		ILLUMINA	Illumina HiSeq 2000	522,080	103,658,531

Library Layout	Library Strategy	Library Source	Library Selection	Library Name
PAIRED	WGS	GENOMIC	RANDOM PCR	unspecified

Navigation | Read Files

This table contains the files for run ERR358492

Download files

View: TEXT

Select columns

Showing results 1 - 1 of 1 results

Study accession	Secondary study accession	Sample accession	Secondary sample accession	Experiment accession	Run accessi
PRJEB4648	ERP003956	SAMEA2242609	ERS362605	ERX331264	ERR358

Fastq files (ftp)	Fastq files (galaxy)	Submitted files (ftp)	Submitted files (galaxy)
File 1	File 1	BAM File 1	BAM File 1
File 2	File 2		
File 3	File 3		

But we'll just cheat.

Create new history

 (cog) → Create New, name it

Import:

Shared Data → Data Libraries

→ **Wolbachia Example**

→ **Reference (select all)**

→ **Reads (select all)**

Import to current history → Go

http://bit.ly/SRR488253

NGS Data Quality Control

- FASTQ format
- Examine quality
- Trim/filter as we see fit

Quality Control is not sexy.

It is vital.

What is FASTQ?

Specifies sequence (FASTA) and quality scores (PHRED)

```
@SEQ_ID
GATTTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTTGTTCAACTCACAGTTT
+
!''*((((***+))%%%++)(%%%%).1***-+*''))**55CCF>>>>>>CCCCCCC65
```

FASTQ is such a cool standard, there are 3 (or 5) of them!

```
SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS.....................................
.........................XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX.....................
.............................IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII.....................
.............................JJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJ.....................
LLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLL.....................................
!"#$%&'()*+,-./0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_`abcdefghijklmnopqrstuvwxyz{|}~
|                         |    |        |                                  |                     |
33                        59   64       73                                 104                   126
 0........................26...31.......40
                          -5....0........9.............................40
                               0........9.............................40
                                  3.....9.............................40
 0.2......................26...31.........41

S - Sanger        Phred+33,  raw reads typically (0, 40)
X - Solexa        Solexa+64, raw reads typically (-5, 40)
I - Illumina 1.3+ Phred+64,  raw reads typically (0, 40)
J - Illumina 1.5+ Phred+64,  raw reads typically (3, 40)
    with 0=unused, 1=unused, 2=Read Segment Quality Control Indicator (bold)
    (Note: See discussion above).
L - Illumina 1.8+ Phred+33,  raw reads typically (0, 41)
```

http://en.wikipedia.org/wiki/FASTQ_format

FastQC: NGS Data Quality Assessment

NGS QC and Manipulation → FastQC

Gives you a lot of information but little control over how it is calculated or presented.

http://bit.ly/FastQCBoxPlot

FASTQ Groomer: FASTQ Translation

NGS QC and Manipulation → FASTQ Groomer

→ Input FASTQ: Illumina 1.3 - 1.7

→ Execute

Variant Calling: A General Plan

- Get to Galaxy
- Get data to look for SNPs in
- Align reads to reference
- Call Variants

http://bit.ly/bolserSNP2014

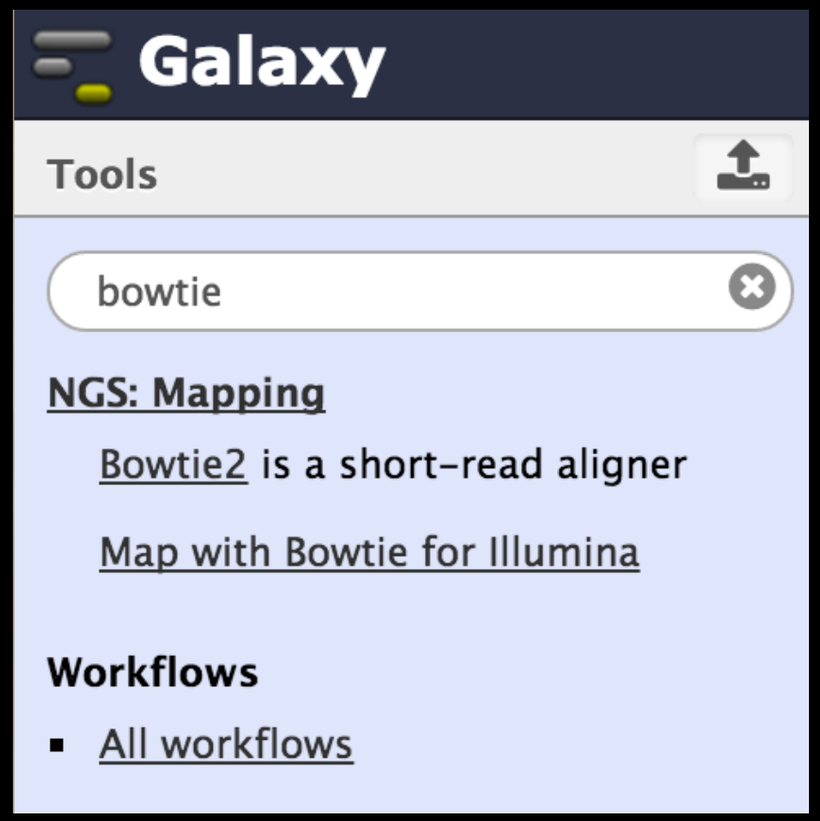
Mapping

Two most (?) common choices:

BWA and Bowtie(2)

Both supported by many Galaxy instances

I'll use Bowtie2



But first a word about naming ...

... and tediousness

Chromosome?

chr?

ENA|AE017196|AE017196.1 ???

GFF3: Chromosome → chr

Text manipulation → Add column

Text manipulation → Cut

ENA|AE017196|AE017196.1 ???

Reference: ENA|AE017196|AE017196.1 → chr

Get data → Upload File → Paste/Fetch data

Paste:

```
chr ENA|AE017196|AE017196.1 Wolbachia
endosymbiont of Drosophila melanogaster, complete
genome
```

Reference: ENA|AE017196|AE017196.1 → chr

Text manipulation

→ Line/Word/Character count

Count number of lines in sequence

→ Select last

Select the last *n-1* lines.

→ Concatenate datasets

Concatenate the pasted sequence name

with

the last n-1 lines of the reference sequence

→ Give the dataset a meaningful name

Bowtie2

From FASTQC
Reads are 50-100 bp

From the SRA Sample Accession:
average fragment size is ... ?

Is this library mate paired?
→ Paired-end

Minimum insert size
→ 0

Maximum insert size
→ 5000 ?

Reference Genome?
→ Use one from the history

Remove PCR duplicates

NGS: SAM Tools → rmdup

Find variation

NGS: Variant Analysis → Naive Variant Caller

Source for the reference list → History

Minimum number of reads to consider a REF/ALT → 20

Minimum base quality → 20

Minimum mapping quality → 20

...

Find variation

NGS: Variant Analysis → Naive Variant Caller

- ...

Ploidy → 1

- ✓ **Only write out positions with alternate alleles**
- ✓ **Report counts by strand**

Additional Filtering: VCFfilter

NGS: VCF Manipulation → VCFfilter

Which ones have enough evidence to be called?

Try several tests:

-f "DP > 20" Read depth is more than 20

-f "QUAL > 30" Phred score likelihood of variant

-f "QUAL > 30" -f "DP > 20"

Filter on quality *and* read depth

VCF Filtering and Annotation w/ features

NGS: VCF Manipulation

VCFannotate

Add information about overlapping features

VCF-BEDintersect

Filter out variation that does not overlap with a genomic feature.

Some Galaxy Terminology

Dataset:

Any input, output or intermediate set of data + metadata

History:

A series of inputs, analysis steps, intermediate datasets, and outputs

Workflow:

A series of analysis steps
Can be repeated with different data

Wolbachia Variation *History* → Reusable *Workflow?*

If we are running 20 or 50 or 2000 of these experiments, we can create a workflow to rerun this entire analysis on each of those datasets.

Create a Workflow from a History

Extract Workflow from history

Create a workflow from this history.

Edit it to make some things clearer.

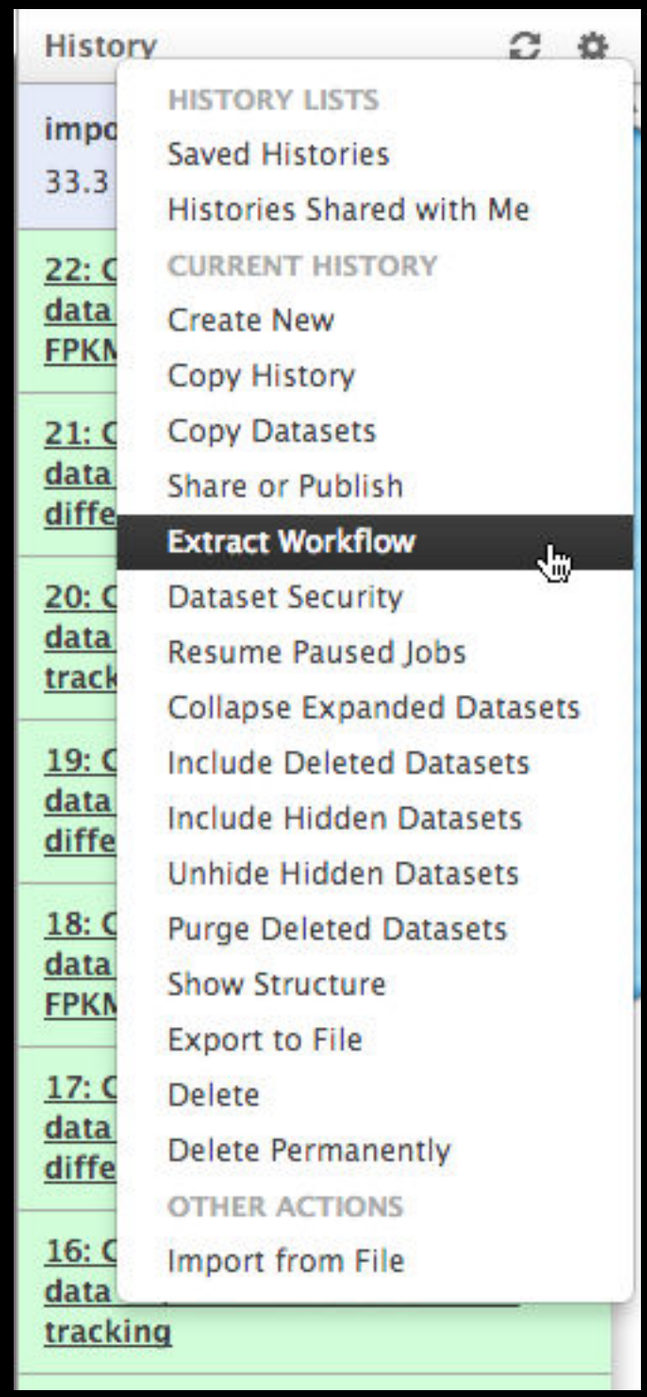
(cog) → Extract Workflow

Run / test it

Rerun with same inputs

Did that work?

When might this not work?



More Galaxy Terminology

Share:

Make something available to someone else

Publish:

Make something available to everyone

Galaxy Page:

Analysis documentation within Galaxy; easy to embed any Galaxy object

Let's all share...

Sharing & Publishing enables Reproducibility

Reproducibility: Everybody talks about it, but ...

Galaxy aims to push the goal of reproducibility from the bench to the bioinformatics realm

All analysis in Galaxy is recorded without any extra effort from the user.

Histories, workflows, visualizations and ***pages*** can be shared with others or published to the world.

Sharing & Publishing enables Reproducibility

CSH PRESS GENOME RESEARCH

Windshield splatter analysis with the Galaxy metagenomic pipeline

Sergei Kosakovsky Pond[1,2,6,9], Samir Wadhawan[3,6,7], Francesca Chiaromonte[4], Guruprasad Ananda[1,3], Wen-Yu Chung[1,3,8], James Taylor[1,5,9], Anton Nekrutenko[1,3,9] and The Galaxy Team[1]

OPEN ACCESS ARTICLE

This Article

Published in Advance October 9, 2009, doi: 10.1101/gr.094508.109

» Abstract *Free*

Current Issue
October 2010, 20 (10)

Footnotes

[Supplemental material is available online at http://www.genome.org. All data and tools described in this manuscript can be downloaded or used directly at http://galaxyproject.org. Exact analyses and workflows used in this paper are available at http://usegalaxy.org/u/aun1/p/windshield-splatter.]

Windshield splatter analysis with the Galaxy metagenomic pipeline: A live supplement

SERGEI KOSAKOVSKY POND[1,2,*], SAMIR WADHAWAN[3,6*], FRANCESCA CHIAROMONTE[4], GURUPRASAD ANANDA[1,3], WEN-YU CHUNG[1,3,7], JAMES TAYLOR[1,5], ANTON NEKRUTENKO[1,3] and THE GALAXY TEAM[1*]

Correspondence should addressed to SKP, JT, or AN.

How to use this document

This document is a live copy of supplementary materials for the manuscript. It provides access to the **exact** analyses and workflows discussed in the paper, so you can play with them by re-running, changing parameters, or even applying them to your own data. Specifically, we provide the two histories and one workflow found below. You can view these items by clicking on their name to expand them. You can also import these items into your Galaxy workspace and start using them; click on the green plus to import an item. To import workflows you must create a Galaxy account (unless you already have one) – a hassle-free procedure where you are only asked for a username and password.

This is the Galaxy history detailing the comparison of our pipeline to MEGAN:

Galaxy History | Galaxy vs MEGAN
Comparison of Galaxy vs. MEGAN pipeline.

This is the Galaxy history showing a generic analysis of metagenomic data. (This corresponds to the "A complete metagenomic pipeline" section of the manuscript and **Figure 3A**):

Galaxy History | metagenomic analysis

This is the Galaxy workflow for generic analysis of metagenomic data. (This corresponds to the "A complete metagenomic pipeline" section of the manuscript and **Figure 3B**):

Galaxy Workflow | metagenomic analysis
Generic workflow for performing a metagenomic analysis on NGS data.

Accessing the Data

Windshield Splatter datasets analyzed in this manuscript can be accessed through this Galaxy Library. From

Author

aun1

Related Pages

All published pages
Published pages by aun1

Rating

Community
(6 ratings, 5.0 average)

Tags

Community:
paper galaxy megan

http://usegalaxy.org/u/aun1/p/windshield-splatter

Outline

4:00 Introduction: The Galactic Landscape

4:10 Variant Analysis: A worked example

5:40 Variant Analysis: Other options

5:50 Galaxy: Resources and Community

6:10 Done

Galaxy

Tools

NGS: Variant Analysis

BamLeftAlign indels in BAM datasets

FreeBayes - Bayesian genetic variant detector

Slice VCF to get data from selected regions

MAF boxplot Minor Allele Frequency Boxplot

Phylorelatives Relatedness of minor alelle sequences in NJ tree

CloudMap: in silico complementation Perform in silico complementation analysis on multiple tabular snpEff output files

CloudMap: Variant Discovery Mapping with WGS data Map a mutation using in silico bulk segregant linkage analysis using variants that are already present in the mutant strain of interest (rather than those introduced by a cross to a polymorphic strain).

CloudMap: Hawaiian Variant Mapping with WGS data Map a mutation by plotting recombination frequencies resulting from crossing to a highly polymorphic strain

CloudMap: EMS Variant Density

Galaxy

Tools

CloudMap: EMS Variant Density Mapping Map a mutation by linkage to regions of high mutation density using WGS data

Variant Annotator process variant counts

FASTA from allele counts Generate major and minor allele sequences from alleles table

Naive Variant Caller - tabulate variable sites from BAM datasets

Annotate a VCF dataset with custom filters

Varscan for variant detection

ANNOVAR Annotate VCF with functional information using ANNOVAR

NGS: VCF Manipulation

VCFfixup: Count the allele frequencies across alleles present in each record in the VCF file

VCFprimers: Extract flanking sequences for each VCF record

VCFcombine: Combine multiple VCF datasets

VCFbreakCreateMulti: Break multiple alleles into multiple records, or combine overallpoing alleles into a single record

Galaxy

Tools

VcfAllelicPrimitives: Split alleleic primitives (gaps or mismatches) into multiple VCF lines

VCFfilter: filter VCF data in a variety of attributes

VCFrandomSample: Randomly sample sites from VCF dataset

VCFgenotype-to-haplotype: Convert genotype-based phased alleles into haplotype alleles

VCFcommonSamples: Output records belonging to samples commong between two datasets

VCFselectsamples: Select samples from a VCF dataset

VCFleftAlign: Left-align indels and complex variants in VCF dataset

VCFtoTab-delimited: Convert VCF data into TAB-delimited format

VCFaddinfo: Adds info fields from the second dataset which are not present in the first dataset

VCFannotate: Intersect VCF records with BED annotations

VCF-VCFintersect: Intersect two VCF datasets

VCFannotateGenotypes:

Galaxy

Tools

VCFannotateGenotypes: Annotate genotypes in a VCF dataset using genotypes from another VCF dataset

VCFcheck: Verify that the reference allele matches the reference genome

VCFsort: Sort VCF dataset by coordinate

VCFgenotypes: Convert numerical representation of genotypes to allelic

VCFhetHomAlleles: Count the number of heterozygotes and alleles, compute het/hom ratio

VCF-BEDintersect: Intersect VCF and BED datasets

VCFdistance: Calculate distance to the nearest variant

snpEff

CloudMap: Check snpEff Candidates Marks up a snpEff output file with matches to a gene candidate list.

SnpSift Filter Filter variants using arbitrary expressions

SnpEff Variant effect and annotation

BEDTools

EMBOSS

https://toolshed.g2.bx.psu.edu

Galaxy Tool Shed

Repositories Help User

2857 valid tools on Dec 16, 2014

Search

- Search for valid tools
- Search for workflows

Valid Galaxy Utilities

- Tools
- Custom datatypes
- Repository depend
- Tool dependency

All Repositories

- Browse by categor

Available Actions

- Login to create a r

Repositories by Category

search repository name, description

Name	Description	Repositories
Assembly	Tools for working with assemblies	36
ChIP-seq	Tools for analyzing and manipulating ChIP-seq data.	10
Transcriptomics	Tools for use in the study of Transcriptomics.	6
Variant Analysis	Tools for single nucleotide polymorphism data such as WGA	147
Visualization	Tools for visualizing data	39
		20
		38
	data cache	5
Data Source	Tools for retrieving data from external data sources	17
Fasta Manipulation	Tools for manipulating fasta data	42

Getting help on Variant Calling

- http://biostars.org/
- http://seqanswers.com/
- http://galaxyproject.org/search

Outline

4:00 Introduction: The Galactic Landscape

4:10 SNP Analysis: A worked example

5:40 SNP Analysis: Other options

5:50 Galaxy: Resources and Community

6:10 Done

Galaxy Community Resources: Galaxy Biostar

Tens of thousands of users leads to a lot of questions.

Absolutely have to encourage community support.

Project traditionally used mailing list

Moved the user support list to Galaxy Biostar, an online forum, that uses the Biostar platform

https://biostar.usegalaxy.org/

Galaxy Community Resources: Mailing Lists

http://wiki.galaxyproject.org/MailingLists

Galaxy-Dev

Questions about developing for and deploying Galaxy
High volume (5200 posts in 2013, 900+ members)
(3246 posts in 2014, 1000+ members)

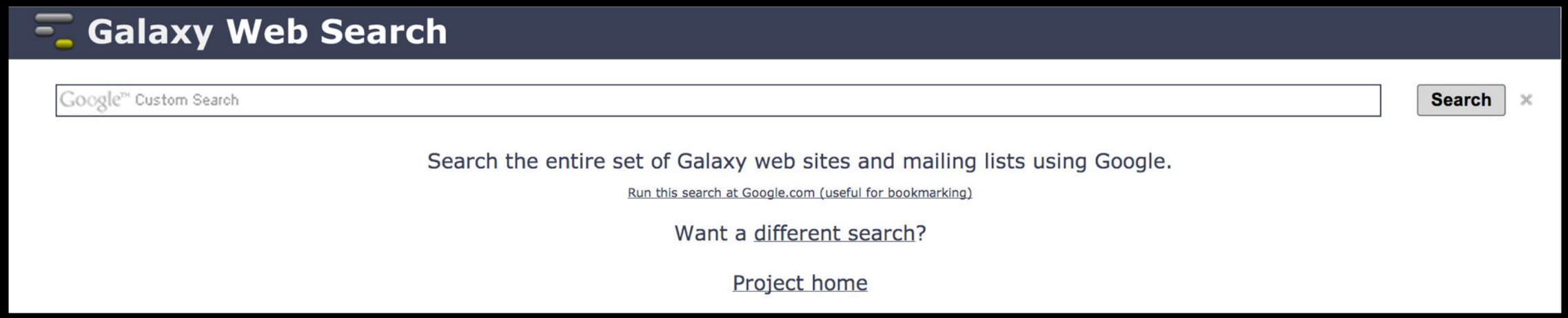
Galaxy-Announce

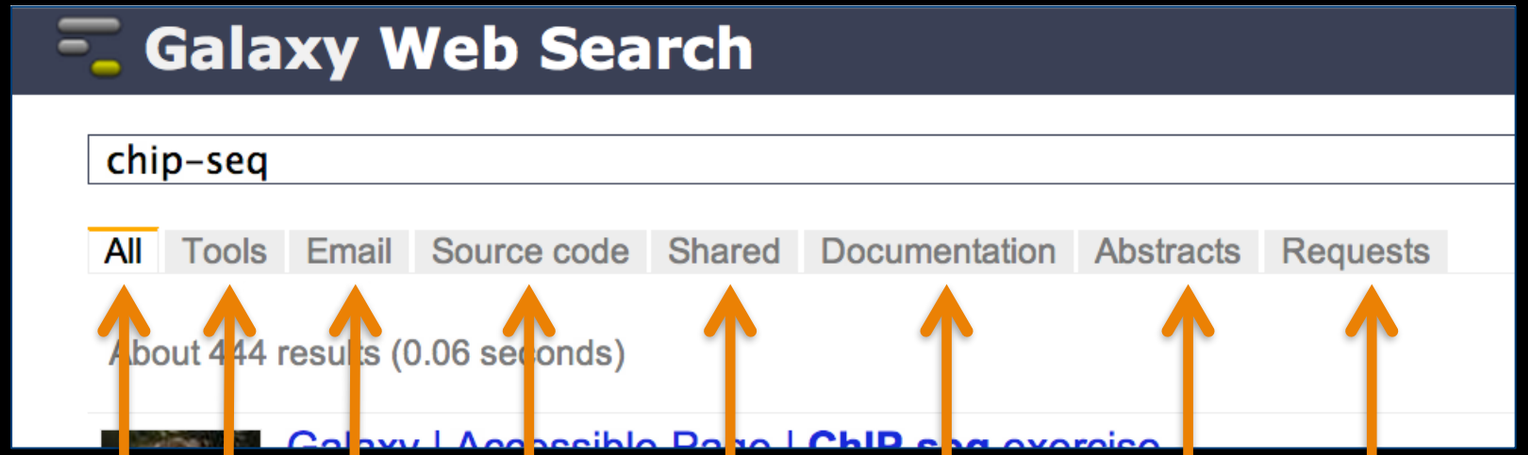
Project announcements, low volume, moderated
Low volume (47 posts in 2013, 3400+ members)
(34 posts in 2014, 4400+ members)

Galaxy-User (discontinued 2014/05)

Questions about using Galaxy and usegalaxy.org
High volume (1328 posts in 2013, 2600+ members)
(358 posts in 2014, 2600+ members)

Unified Search: http://galaxyproject.org/search





Find

Everything on …

Tools for …

Email about …

Source code for …

Published Histories, Pages, Workflows, about …

Documentation on …

Papers using Galaxy for …

Related feature requests

http://wiki.galaxyproject.org

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

- **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.

This is the Galaxy Community Wiki. It describes all things Galaxy.

Use Galaxy

Galaxy's public web server usegalaxy.org makes analysis tools, genomic data, tutorial demonstrations, persistent workspaces, and publication services available to any scientist. Extensive user documentation applicable to any public or local Galaxy instance is available.

Deploy Galaxy

Galaxy is a free and open source project available to all. Local Galaxy servers can be set up by downloading the Galaxy application.

- Admin
- Cloud

Community & Project

Galaxy has a large and active user community and many ways to get involved.

- Community

Contribute

- **Users:** Share your histories, workflows, visualizations, data libraries, and Galaxy Pages, enabling others to use and learn from them.

Galaxy web search

Use Galaxy

Servers • Learn
Main • Choices
Share • Search

Communicate

Support • Biostar
Events • Mailing Lists
News • Twitter

Deploy Galaxy

Get Galaxy • Cloud
Admin • Tool Config
Tool Shed • Search

Contribute

Develop • Tools
Issues & Requests
Logs • Deployments
Teach

Galaxy Project

Home • About • Cite
Community
Big Picture

Events

News

Galaxy Wiki

DaveClements Settings Logout | Search: Titles Text

Events

Edit History Actions

Galaxy Event Horizon

Events with Galaxy-related content are listed here.

Also see the Galaxy Events Google Calendar for a listing of events and deadlines that are Galaxy Community. This is also available as an RSS feed.

If you know of any event that should be added to this page and/or to the Galaxy Event Calendar, send it to outreach@galaxyproject.org.

For events prior to this year, see the **Events Archive**.

Upcoming Events

Date	Topic/Event	Venue/Location
December 12	*Introduction to Galaxy Workshop*	Virginia State University, Petersburg, Virgin
December 16-19	*RNA-Seq and ChIP-Seq Analysis with Galaxy*	UC Davis, California, United States
		2015
January 10-14	*Galaxy for SNP and Variant Data Analysis*	Plant and Animal Genome XXIII (PAG2014), States
January 19-20	*NGS pipelines with Galaxy*	e-Infrastructures for Massively Parallel Sequ Sweden
February 9-13	*Analyse bioinformatique de séquences sous Galaxy*	Montpellier, France
February 16-18	*Accessible and Reproducible Large-Scale Analysis with Galaxy*	Genome and Transcriptome Analysis, pa Conference, San Francisco, Cali
	Large-Scale NGS data Analysis on Amazon Web Services Using Globus Genomic	Genomics & Sequencing Data Integration, of Molecular Medicine Tri-Conference, Sa States
	iReport: An Integrative "omics"	

News Items

Opening at McMaster University

The McArthur Lab in the McMaster University Department of Biochemistry & Biomedical Sciences is seeking a Systems Administrator / Information Technologist to help establish a new bioinformatics laboratory at McMaster, plus develop the next generation of the Comprehensive Antibiotic Resistance Database (CARD).

From the job announcement on EvolDir:

> The candidate will configure BLADE and other hardware for general bioinformatics analysis, development of a GIT version control system, **construction of an in house Galaxy server (usegalaxy.org)**, and development of a new interface, stand-alone tools, APIs, and algorithms for the CARD (based on Chado).

See the full announcement for details.

Posted to the Galaxy News on 2014-12-05

December 2014 Galaxy Newsletter

As always there's a lot going on in the Galaxy this month. "Like what?" you say. Well, read the dang December Galaxy Newsletter we say! Highlights include:

- Galaxy Day! In Paris! This Wednesday!
- Near Richmond, Virginia? There's a Galaxy Workshop at Virginia State U on December 12.
- GCC2015 needs sponsors!
- Other upcoming events on two continents
- **96 new papers**, including 6 highlighted papers, referencing, using, extending, and implementing Galaxy.
- Job openings at 7+ organizations
- A new mailing list: Galaxy-Training
- 15 new ToolShed repositories from 10 contributors
- And, 10 other juicy (well maybe not *juicy*, but certainly not *crunchy*) bits of news

Dave Clements and the *crisp* Galaxy Team

Posted to the Galaxy News on 2014-12-01

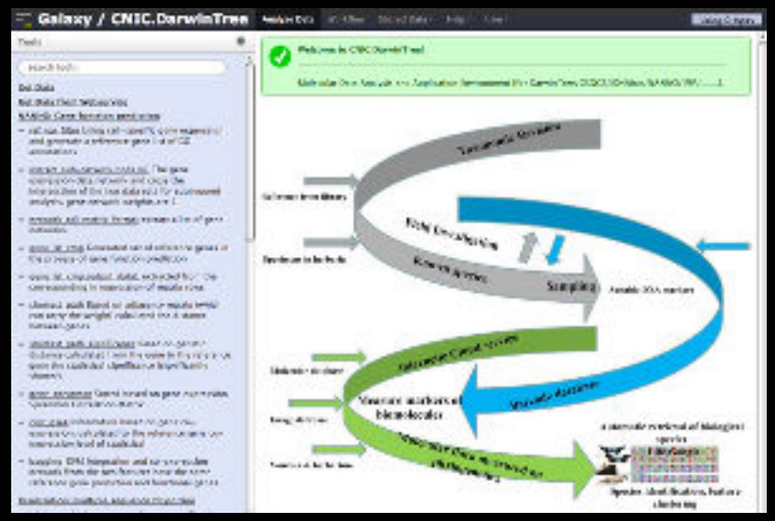
Bioinformaticians, Freiburg

Max Planck Institute of Immunobiology and Epigenetics in Freiburg, Germany has an opening for a Bioinformatician for an initial period of two years. The successful candidate will work at the interface between an in-house deep-sequencing facility (HiSeq-2500) and the various research groups at the institute. Main responsibilities include

- primary analysis of deep-sequencing data and quality controls

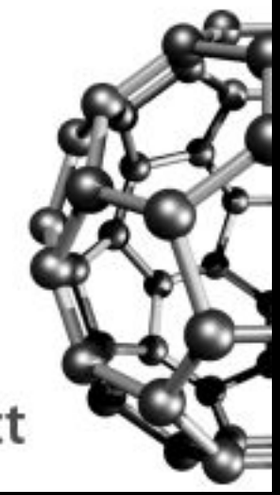
Cistrome

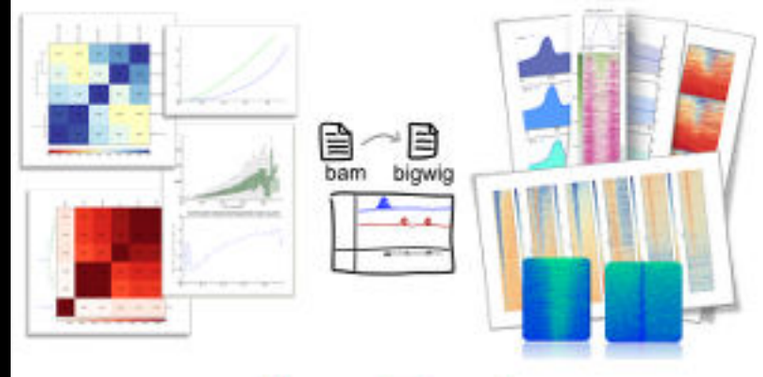
A Galaxy Server dedicated to ChIP-* analysis



ballaxy

Powered by the Biochemical Algorithms Library Project





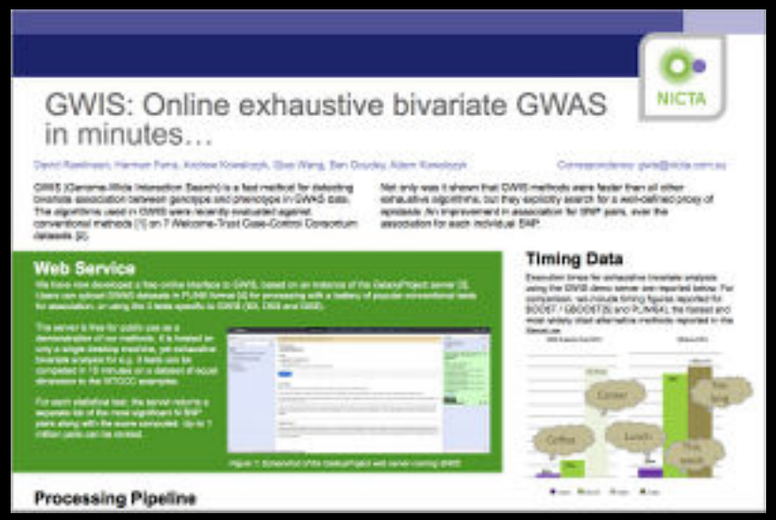
deepTools

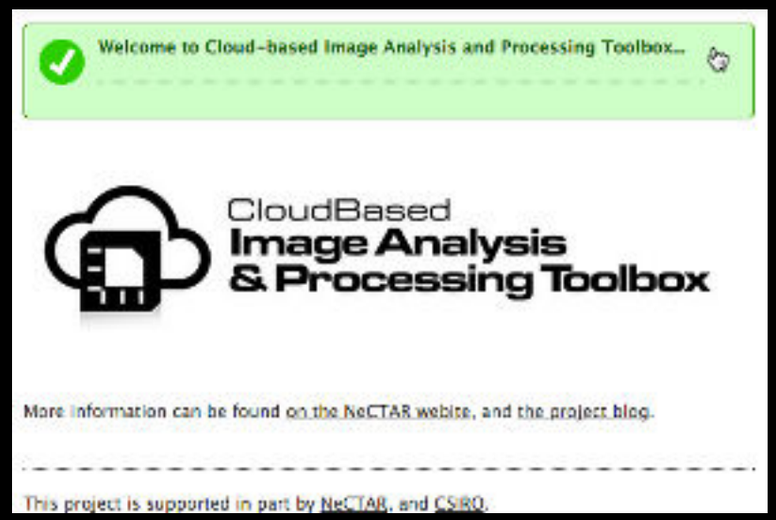
CoSSci

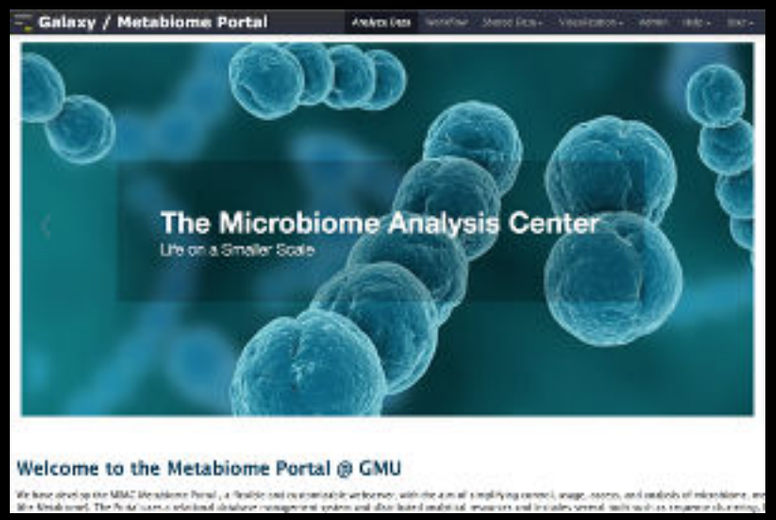
Galaxy for Complex Social Sciences

GENEONTOLOGY
Unifying Biology

galaxy.berkeleybop.org







bit.ly/gxyServers

Community can create, vote and comment on issues

HOME TOUR GOLD BUSINESS CLASS BLOG Trello Sign Up Log In

Want to subscribe, vote or comment on these cards? Sign up for free or learn more about Trello

Galaxy: Development

Public

Inbox

To add cards, use http://galaxyproject.org/trello
4 votes 2

To request reference genome, comment on this card.
1 vote 5 0/6

Toolshed installation fails silently
3 votes 1

Handle cluster job preemption
2 votes 1

Return code 271 causes traceback for PBS torque
1 vote 2

BUG: Tool shed repository export to capsule does not always capture all dependencies
1 vote 1

Remove manual_builds.txt from source control and replace with a .sample version
1 vote 1

Tool Requests

595: Add SAMTools "Sort"
4 votes 13

601: SAM-to-BAM tool enhancements
2 votes 1

Tools: Add tool to generate simulated reads to Main
3 votes 1

default max insert size of Bowtie2 should be increased
2 votes 5

307: A tool to produce a set of random intervals.
2 votes 2

Converter Tool: SAM to BAM enhancements
2 votes

New Tool: convert IUPAC chars to N
1 vote 7 1

Bug Reports

Usability: expanding datasets near the bottom of panel
CE

Bug: SICER on Main dependency issue
2 votes 20 3/5

Profile Annotations bad values when "select all"
1 vote 5

Filter pileup tool doesn't recognize pileup output data
1 vote 2

Bug: Odd Fetch Taxonomy tool behavior
1 vote 1

Strip message after pause jobs resumed
1 vote 1

Ideas

697: Workflow job control functions
10 votes 9

User Metrics and Analytics
3 votes 3 1/2
CE

Tuxedo RNA-seq tools: report command-line
2 votes 3

Tools: Incorporate key Cuffdiff output files for Cummerbund
2 votes 1 0/3

Moving objects between Galaxy instances, data federation, distributed storage, and data locality
2 votes

Workflow Editor: Provide explicit access to implicit datatype converter tools
1 vote

Pull Re

665: P issue

Custo
2 vo

Tools: Reque
3 vo

add rn downs
2 vo

Please wrapp mappi
2 vo

[galax libxml
1 vo

Pull Re manag

Menu

Members

CE NS G

Activity

Lance Parsons on Add or update wrappers for SamTools 1.0

I see that @peterjc has a wrapper for idxstats already and that it's listed on this card as "done" but I don't see it in the github repo. Will idxstats become part of this devteam collection or should I just start using the wrapper from @peterjc (Thanks Peter!)

today at 3:52 pm

G **g2roboto** added Pull Request #606 - [STABLE] Escape instances of message passed in through kwd before pushing them back out to

http://bit.ly/gxytrello

GALAXY COMMUNITY CONFERENCE

BALTIMORE, MD | JUNE 30 - JULY 2, 2014

Slides, posters & videos now online
http://bit.ly/gcc2014

GCC 2015

Galaxy Community Conference

6-8th July 2015

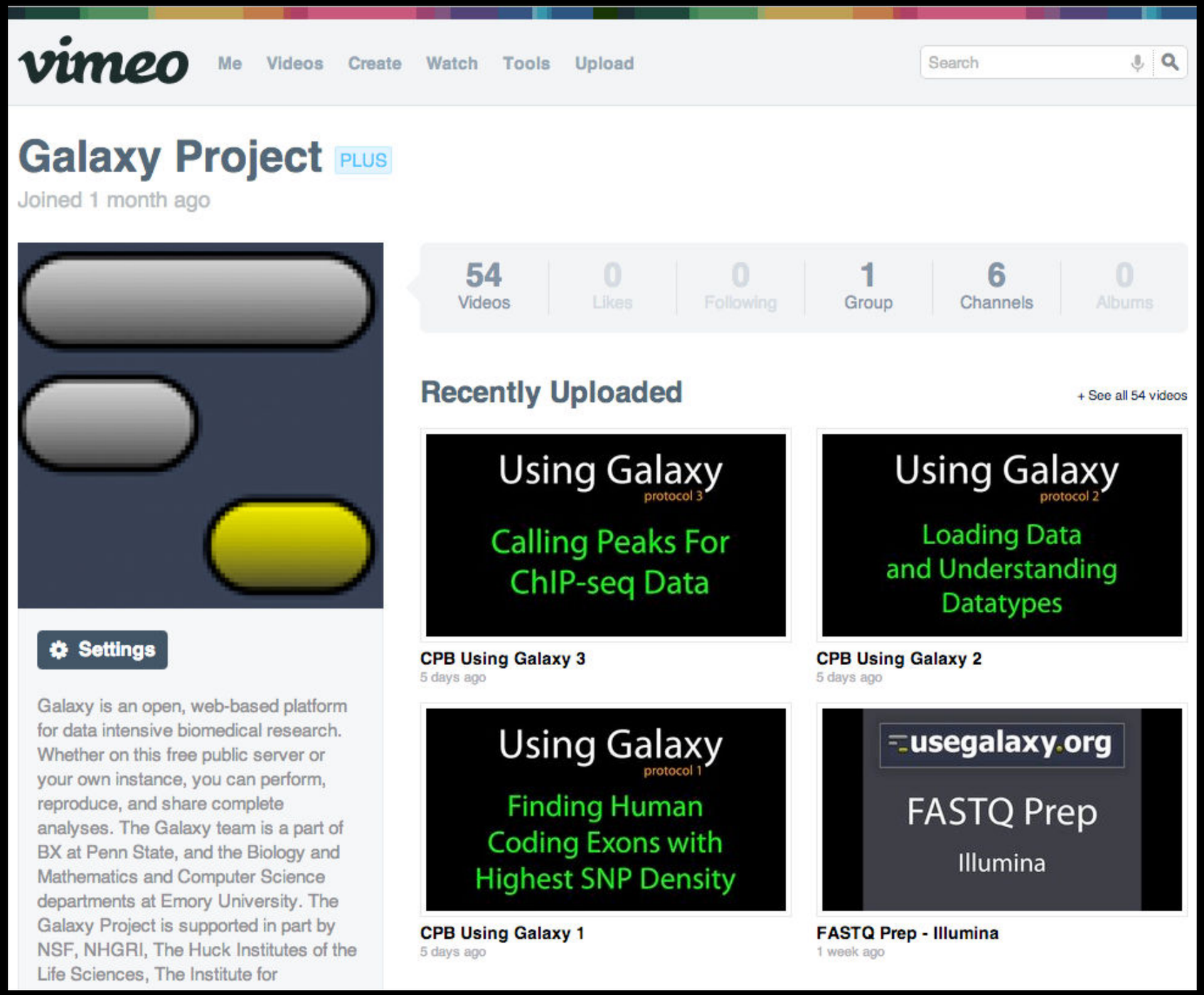
The Sainsbury Laboratory
Norwich, UK

galaxyproject.org

Galaxy Australasia Workshop 2014

We also support community organized efforts and events.

Galaxy Resources & Community: Videos

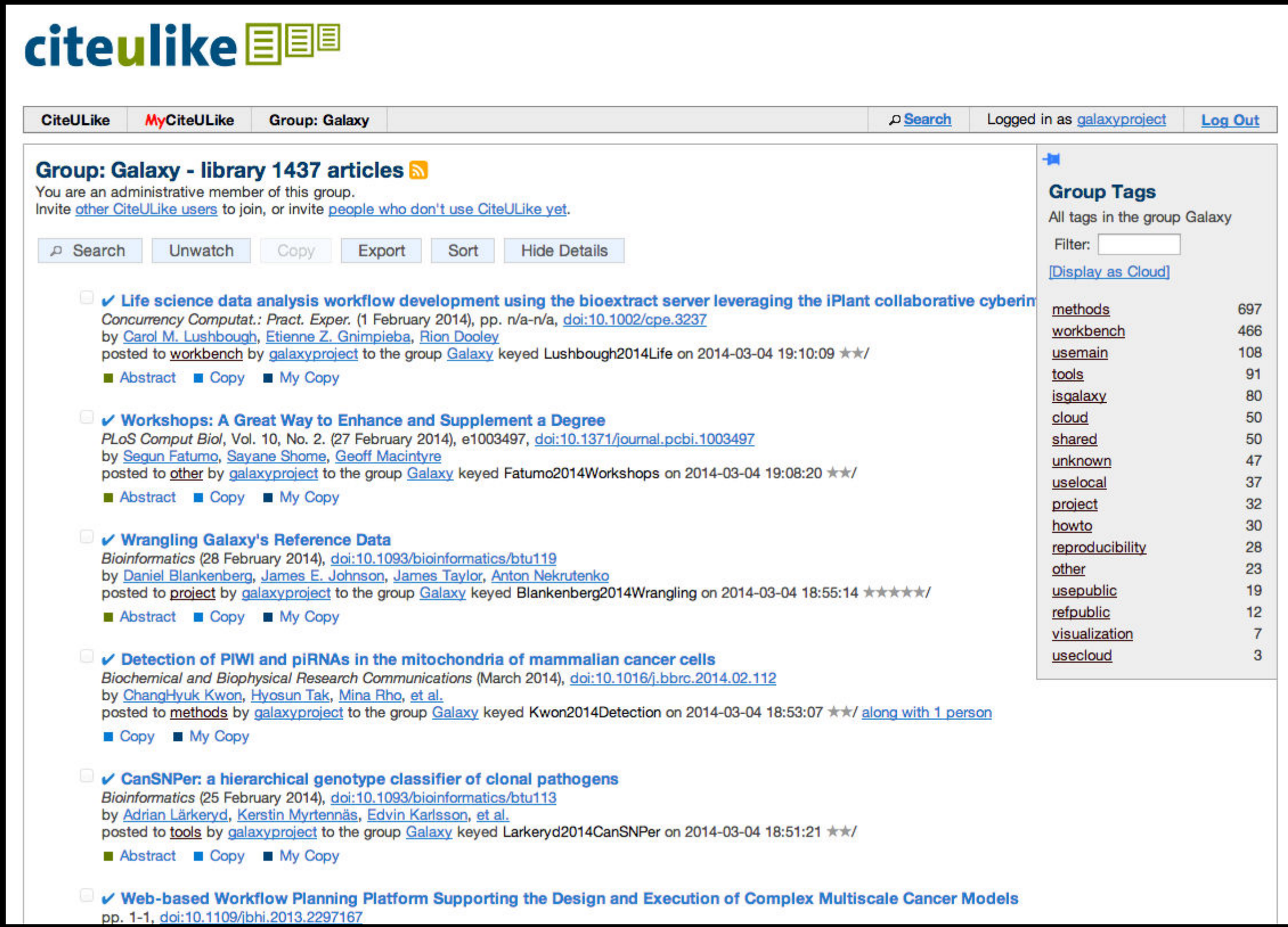


"How to" screencasts on using and deploying Galaxy

Talks from previous meetings.

http://vimeo.com/galaxyproject

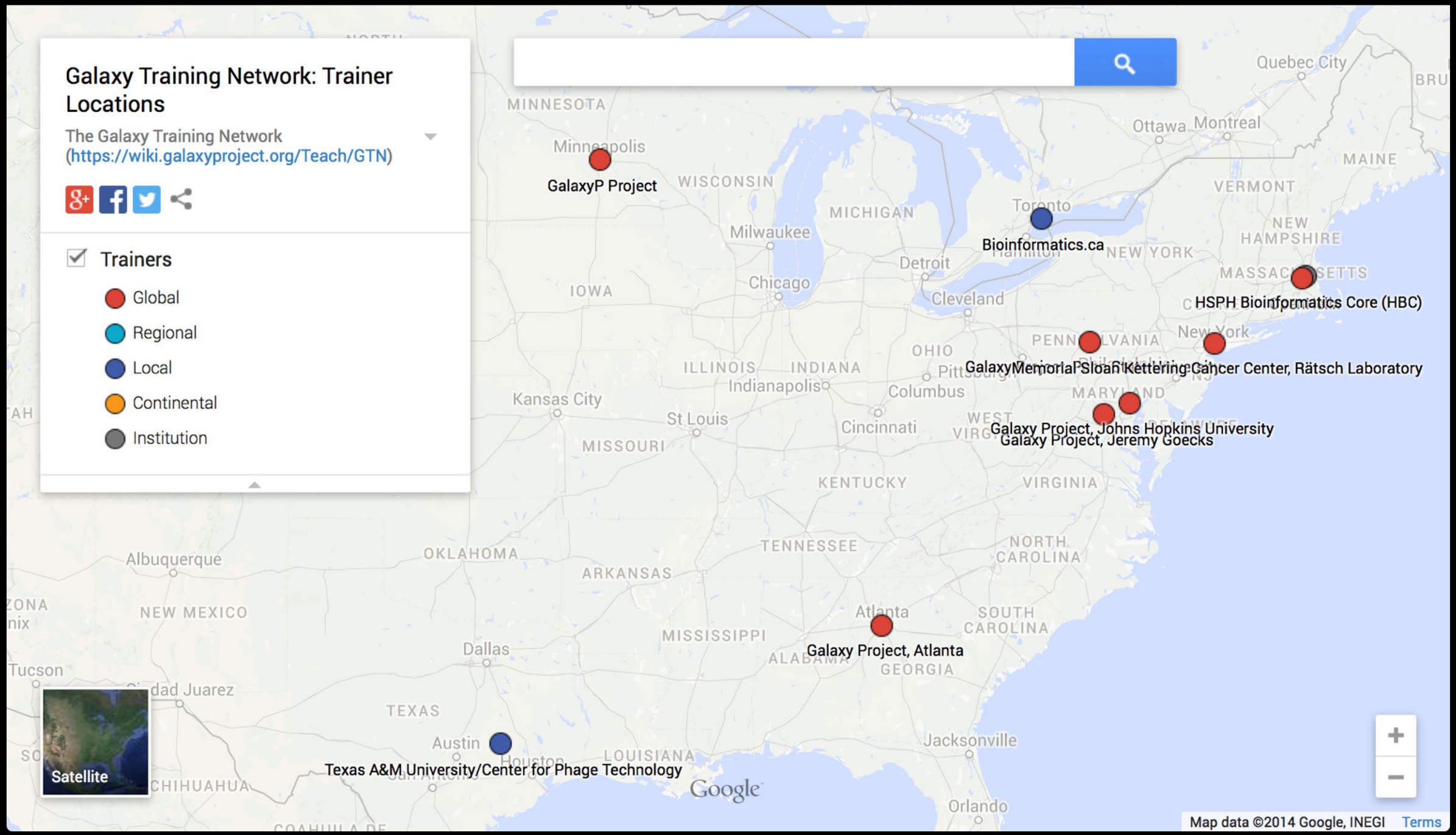
Galaxy Resources & Community: CiteULike Group



Over 2000 papers

http://bit.ly/gxycul

Scaling the Project: Training



Galaxy Training Network launched In October.

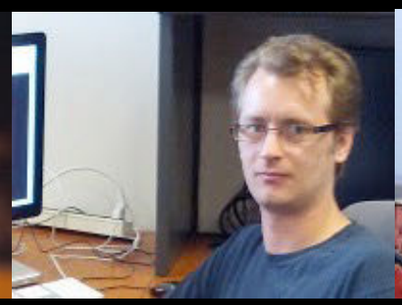
bit.ly/gxygtn

The Galaxy Team

Enis Afgan

Dannon Baker

Dan Blankenberg



Dave Bouvier

Marten Cech

John Chilton

Dave Clements

Nate Coraor

Carl Eberhard

Jeremy Goecks

Sam Guerler

Jen Jackson

Ross Lazarus

Anton Nekrutenko

Nick Stoler

James Taylor

Nitesh Turaga

http://wiki.galaxyproject.org/GalaxyTeam

Galaxy is hiring post-docs and software engineers

Please help.

http://wiki.galaxyproject.org/GalaxyIsHiring

Thanks

Dave Clements

Galaxy Project

Johns Hopkins University

outreach@galaxyproject.org