Introduction to Galaxy

Johns Hopkins University
February 4, 2013

Dave Clements, Emory University
http://galaxyproject.org/

Mo Heydarian, Johns Hopkins University
http://jhu.edu/

Johns Hopkins University
Center for Computational Biology
CCB

Salzberg Lab

Reddy &
Sollner-Webb Labs

amazon
web services™

Galaxy

Agenda

9:00	Welcome
9:20	Basic Analysis with Galaxy
10:20	Basic Analysis into Reusable Workflows
10:40	Break
11:00	RNA-Seq Example Part I
12:00	Galaxy Project Overview
12:20	Lunch
1:05	RNA-Seq Example Part II
	Cufflinks, Visualization and Visual Analytics
1:55	Sharing, Publishing and Reproducibility
2:15	Break
2:35	Setting up your own Galaxy Cluster on AWS
4:30	Done

Introductions

In 35 seconds or less tell us

- your name
- your affiliation(s)
- something about your research
- something about what you want to learn

Goals

1. Introduce Galaxy
2. Introduce bioinformatics concepts and formats
3. Hands-on experience
    - Load and integrate data
    - Perform bioinformatic analysis with Galaxy
    - Save, share describe and publish your analyses
    - Visualize your results
    - Set up your own Galaxy server in the cloud

This workshop will not cover details of how tools are implemented, or new algorithm designs, or which assembler or mapper or ... is best for you.

Agenda

9:00	Welcome
9:20	Basic Analysis with Galaxy
10:20	Basic Analysis into Reusable Workflows
10:40	Break
11:00	RNA-Seq Example Part I
12:00	Galaxy Project Overview
12:20	Lunch
1:05	RNA-Seq Example Part II
	Cufflinks, Visualization and Visual Analytics
1:55	Sharing, Publishing and Reproducibility
2:15	Break
2:35	Setting up your own Galaxy Cluster on AWS
4:30	Done

Basic Analysis

On human chromosome 22,
which coding exons have the most
repeats in them?

http://cloud1.galaxyproject.org (gold)

http://cloud2.galaxyproject.org (sable)

http://cloud3.galaxyproject.org (black)

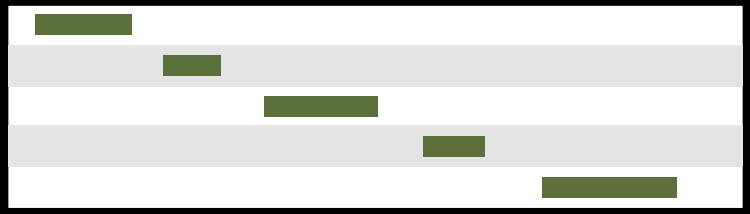
(~ http://usegalaxy.org/galaxy101)

Exons & Repeats: A General Plan

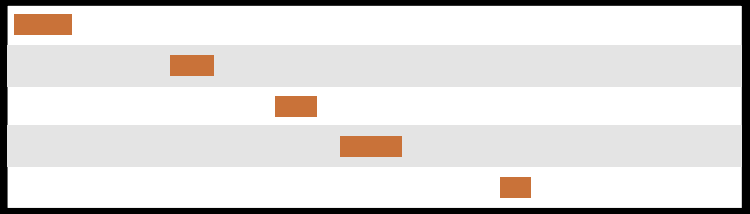
- Get some data
  - Coding exons on chromosome 22
  - Repeats on chromosome 22
- Mess with it
  - Identify which exons have repeats
  - Count repeats per exon
  - Save, download, ... exons with most repeats

http://cloud1.galaxyproject.org (gold)
http://cloud2.galaxyproject.org (sable)
http://cloud3.galaxyproject.org (black)

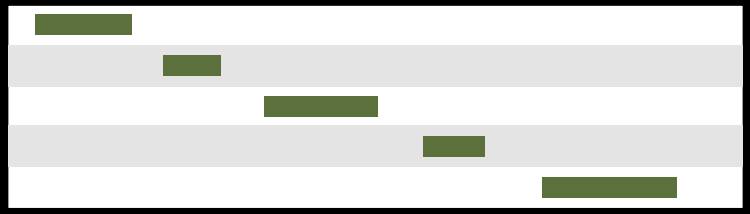
(~ http://usegalaxy.org/galaxy101)



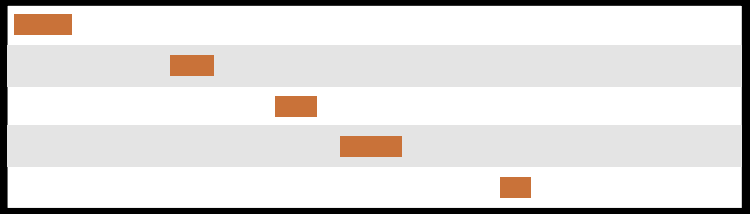
Exons, from UCSC



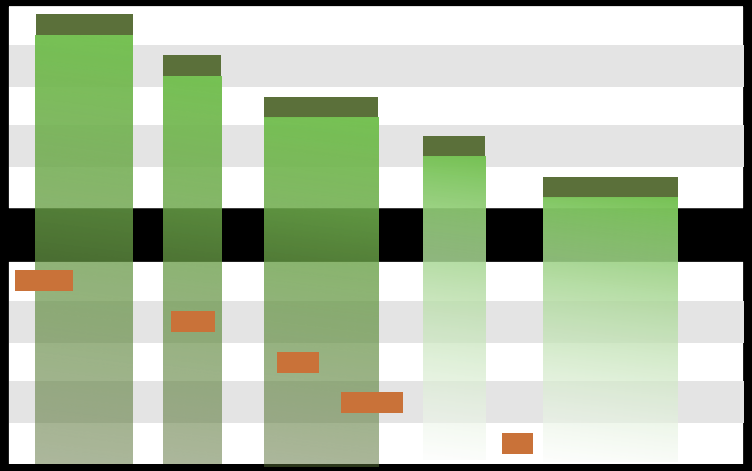
Repeats, from UCSC



Exons, from UCSC



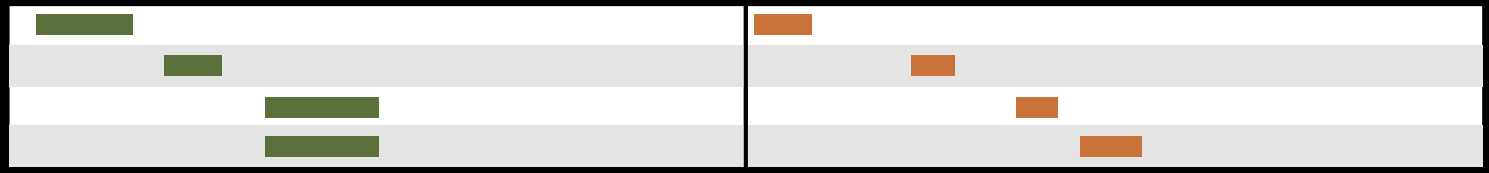
Repeats, from UCSC

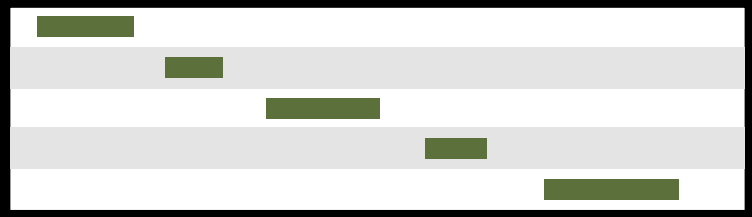


Exons, from UCSC

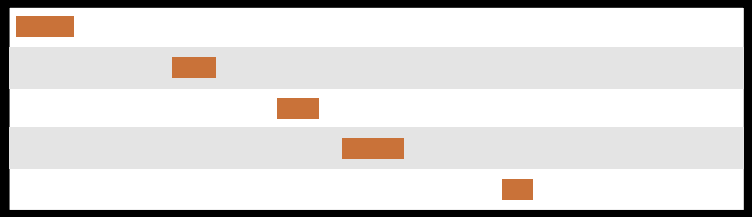
Repeats, from UCSC

Overlap pairings

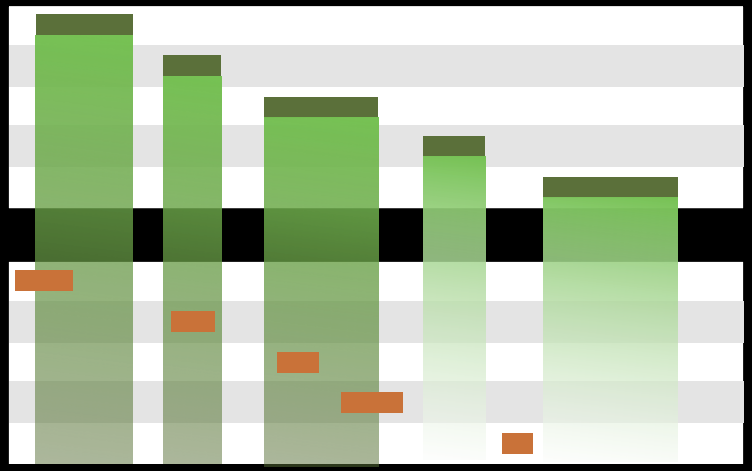




Exons, from UCSC



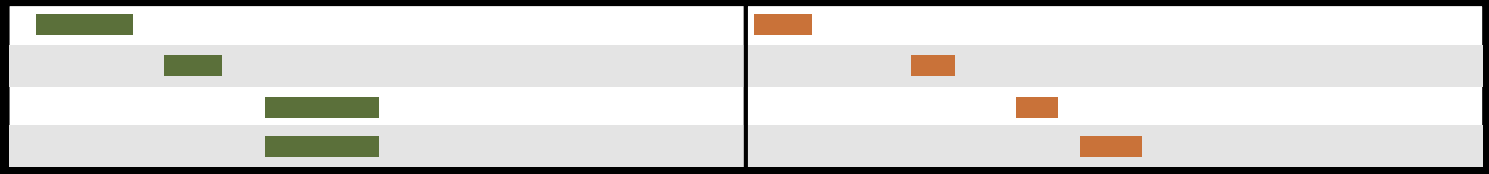
Repeats, from UCSC



Exons, from UCSC

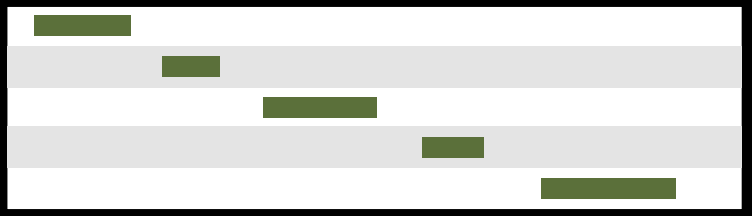
Repeats, from UCSC

Overlap pairings



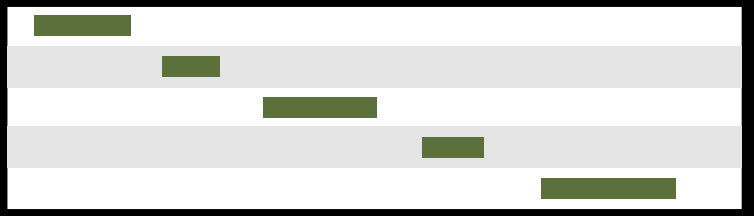
Exon overlap counts

Exon overlap counts



Exons, from UCSC

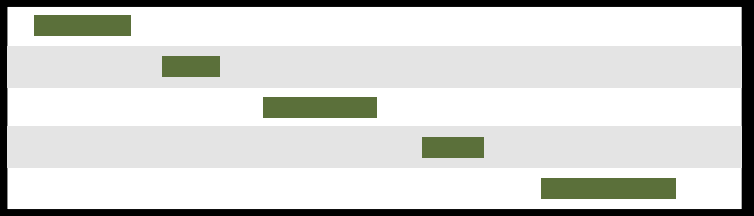
Exon overlap counts



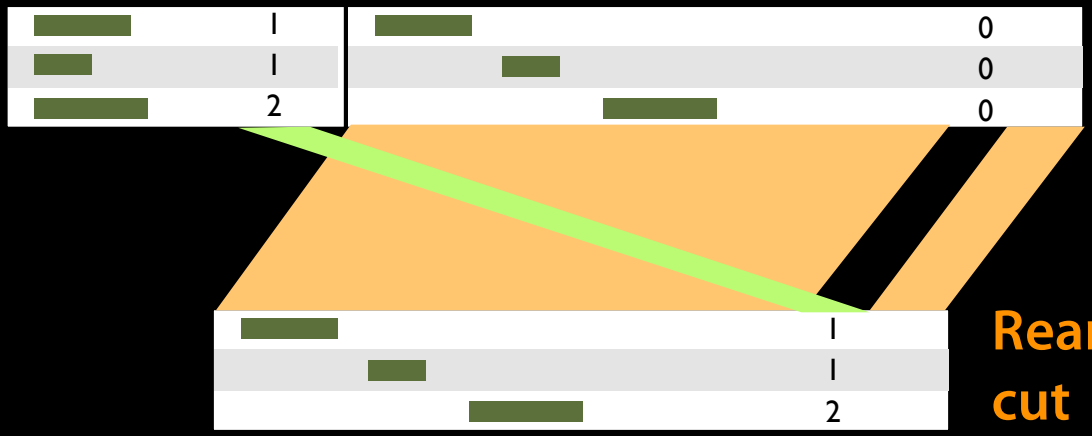
Exons, from UCSC

Join on exon name

Exon overlap counts



Exons, from UCSC



Join on exon name

Rearrange columns w/ cut

Agenda

9:00	Welcome
9:20	Basic Analysis with Galaxy
10:20	Basic Analysis into Reusable Workflows
10:40	Break
11:00	RNA-Seq Example Part I
12:00	Galaxy Project Overview
12:20	Lunch
1:05	RNA-Seq Example Part II
	Cufflinks, Visualization and Visual Analytics
1:55	Sharing, Publishing and Reproducibility
2:15	Break
2:35	Setting up your own Galaxy Cluster on AWS
4:30	Done

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

Exons and Repeats *History* → Reusable *Workflow?*

- The analysis we just finished was about
  - Human chromosome 22
  - Overlap between exons and repeats
- But, ...
  - there is nothing inherently in the analysis about humans, chromosomes, exons or repeats
  - It is a series of steps that sets the score of one set of features to the number of overlaps from another set of features.

Create a generic *Overlap* Workflow

Extract Workflow from history

Create a workflow from this history.

Edit it to make some things clearer.

Run / test it

Guided: rerun with same inputs

On your own:

Count # CpG islands in each exon

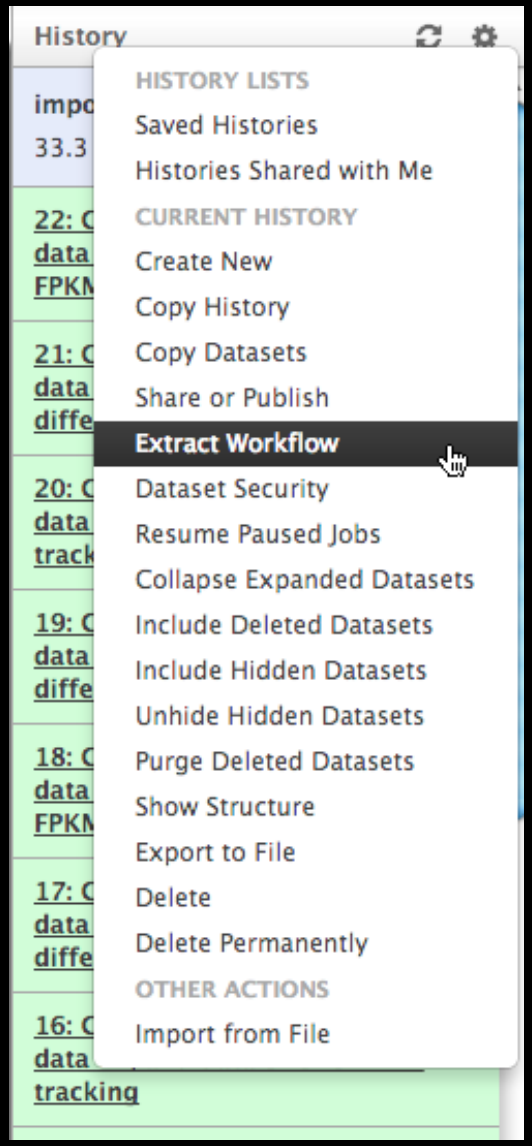
Did that work?

On your own:

Count # of exons in each repeat

Did that work? *Why not?*

Edit workflow: doc assumptions



Agenda

9:00 Welcome
9:20 Basic Analysis with Galaxy
10:20 Basic Analysis into Reusable Workflows
10:40 Break
11:00 RNA-Seq Example Part I
12:00 Galaxy Project Overview
12:20 Lunch
1:05 RNA-Seq Example Part II
Cufflinks, Visualization and Visual Analytics
1:55 Sharing, Publishing and Reproducibility
2:15 Break
2:35 Setting up your own Galaxy Cluster on AWS
4:30 Done

Agenda

9:00 Welcome
9:20 Basic Analysis with Galaxy
10:20 Basic Analysis into Reusable Workflows
10:40 Break
11:00 RNA-Seq Example Part I
12:00 Galaxy Project Overview
12:20 Lunch
1:05 RNA-Seq Example Part II
Cufflinks, Visualization and Visual Analytics
1:55 Sharing, Publishing and Reproducibility
2:15 Break
2:35 Setting up your own Galaxy Cluster on AWS
4:30 Done

RNA-seq Exercise

http://usegalaxy.org/u/jeremy/p/galaxy-rna-seq-analysis-exercise

http://bit.ly/gxyRNASEX

http://cloud1.galaxyproject.org (gold)

http://cloud2.galaxyproject.org (sable)

http://cloud3.galaxyproject.org (black)

RNA-seq Exercise: A Plan

- Get input datasets; hg19, will mostly map to chr19
- Look at quality
- Trim as we see fit.
- Map the reads to the human reference using Tophat
- Run Cufflinks on Tophat output to assemble reads into transcripts
- Visualize it

http://bit.ly/gxyRNASEX

RNA-seq Exercise: A Plan

- Get input datasets; hg19, will mostly map to chr19
  - All datasets are FASTQ and from the Body Map 2.0 project

http://bit.ly/gxyRNASEX

What is FASTQ?

- Specifies sequence (FASTA) and quality scores (PHRED)
- Text format, 4 lines per entry

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

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

```
 SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
 ..............................IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
 ..........................XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
 !"#$%&'()*+,-./0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_`abcdefghijklmnopqrstuvwxyz{|}~
 |                         |    |        |                              |                     |
33                        59   64       73                            104                   126

S - Sanger       Phred+33,  93 values  (0, 93) (0 to 60 expected in raw reads)
I - Illumina 1.3 Phred+64,  62 values  (0, 62) (0 to 40 expected in raw reads)
X - Solexa       Solexa+64, 67 values (-5, 62) (-5 to 40 expected in raw reads)
```

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

RNA-seq Exercise: A Plan

- Get input datasets; hg19, will mostly map to chr19
- Look at quality: Option 1
  - NGS QC and Manipulation → **Compute Quality Statistics**
  - NGS QC and Manipulation → **Draw quality score boxplot**
  - Gives you no control over how it is calculated or presented.

http://bit.ly/gxyRNASEX

RNA-seq Exercise: A Plan

- Get input datasets; hg19, will mostly map to chr19
- Look at quality: Option 2
  - NGS QC and Manipulation → **FastQ Summary Statistics**
  - Graph / Display Data → **Boxplot of quality statistics**
  - Gives you a lot of control over what the box plot looks like, but no additional information

http://bit.ly/gxyRNASEX

RNA-seq Exercise: A Plan

- Get input datasets; hg19, will mostly map to chr19
- Look at quality: Option 3
  - NGS QC and Manipulation → **Fastqc**
  - Gives you a lot a lot more information but little control over how it is calculated or presented.

http://bit.ly/gxyRNASEX

RNA-seq Exercise: A Plan

- Get input datasets; hg19, will mostly map to chr19
- Look at quality
- Trim as we see fit: Option 1
  - **NGS QC and Manipulation → FASTQ Trimmer by column**
  - Trim same number of columns from every record
  - Can specify different trim for 5' and 3' ends

http://bit.ly/gxyRNASEX

RNA-seq Exercise: A Plan

- Get input datasets; hg19, will mostly map to chr19
- Look at quality
- ~~Trim~~ Filter as we see fit: Option 2
  - NGS QC and Manipulation → **Filter FASTQ reads by quality score and length**
  - Keep or discard whole reads at a time
  - Can have different thresholds for different regions of the reads.
  - Keeps original read length.

http://bit.ly/gxyRNASEX

RNA-seq Exercise: A Plan

- Get input datasets; hg19, will mostly map to chr19
- Look at quality
- Trim as we see fit: Option 3
  - NGS QC and Manipulation → **FASTQ Quality Trimmer by sliding window**
  - Trim from both ends, using sliding windows, until you hit a high-quality section.
  - Produces variable length reads

http://bit.ly/gxyRNASEX

RNA-seq Exercise: A Plan

- Get input datasets; hg19, will mostly map to chr19
- Look at quality
- Trim as we see fit.
- Map the reads to the human reference using Tophat
  - *Imagine pages and pages of discussion on the intricacies and pitfalls of RNA-seq mapping here.*

http://bit.ly/gxyRNASEX

Agenda

9:00	Welcome
9:20	Basic Analysis with Galaxy
10:20	Basic Analysis into Reusable Workflows
10:40	Break
11:00	RNA-Seq Example Part I
12:00	Galaxy Project Overview
12:20	Lunch
1:05	RNA-Seq Example Part II
	Cufflinks, Visualization and Visual Analytics
1:55	Sharing, Publishing and Reproducibility
2:15	Break
2:35	Setting up your own Galaxy Cluster on AWS
4:30	Done

What is Galaxy?

- **A free (for everyone) web service** integrating a wealth of tools, compute resources, terabytes of reference data and permanent storage
- **Open source software** that makes integrating your own tools and data and customizing for your own site simple
- These options result in several **ways to use Galaxy**

http://galaxyproject.org

Galaxy is available ...

- As a free (for everyone) web service

  http://usegalaxy.org

  However, ***a centralized solution cannot scale to meet the analysis needs of the entire world.***

Galaxy is available ...

- As a free (for everyone) web service

  http://usegalaxy.org

- As open source software

  http://getgalaxy.org

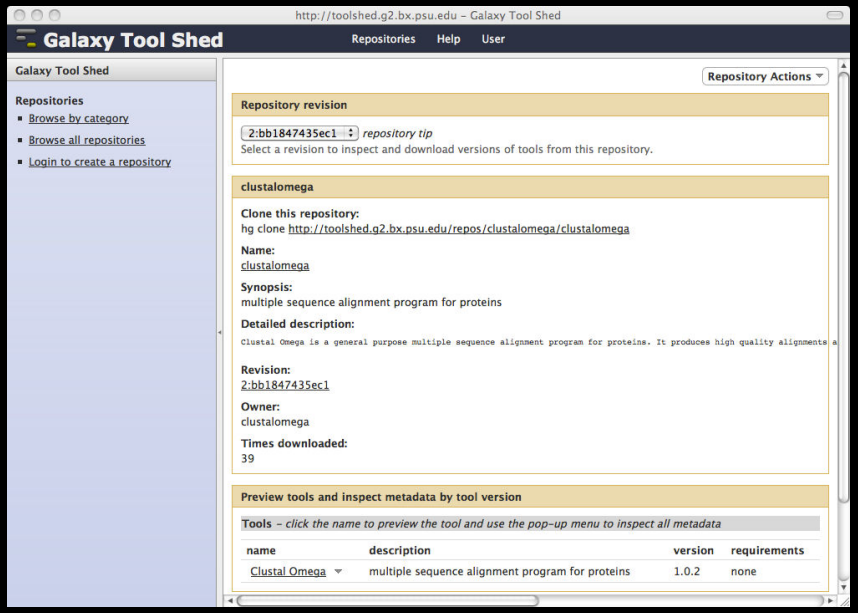
As Open Source Software: Local Galaxy Instances

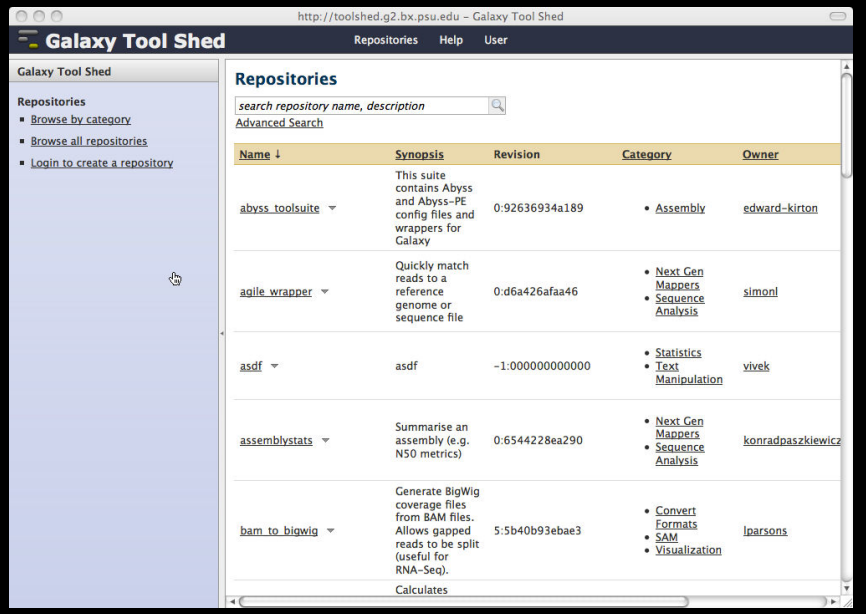
- Galaxy is designed for local installation and customization
  - Easily integrate new tools
  - Easy to deploy and manage on nearly any (unix) system
  - Run jobs on existing compute clusters
- Requires a computational resource on which to be deployed

http://getgalaxy.org

Encourage Local Galaxy Instances

- Encourage and support Local Galaxy Instances
  - Support increasingly decentralized model and improve access to existing resources
  - Focus on building infrastructure to enable the community to integrate and share tools, workflows, and best practices





Galaxy Tool Shed

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

Encourage Public Galaxy Instances

http://wiki.galaxyproject.org/PublicGalaxyServers

Interested in:

ChIP-chip and ChIP-seq?
- ✓ Cistrome

Statistical Analysis?
- ✓ Genomic Hyperbrowser

Protein synthesis?
- ✓ GWIPS-viz

de novo assembly?
- ✓ CBIIT Galaxy

Reasoning with ontologies?
- ✓ OPPL Galaxy

Repeats!
- ✓ RepeatExplorer

Everything?
- ✓ Andromeda

Plus many more

As Open Source Software: Local Galaxy Instances

- Galaxy is designed for local installation and customization
  - Easily integrate new tools
  - Easy to deploy and manage on nearly any (unix) system
  - Run jobs on existing compute clusters
- Requires a **computational resource** on which to be deployed

http://getgalaxy.org

Got your own cluster?

- Control where tool execution happens
- Galaxy works with any DRMAA compliant cluster job scheduler (which is most of them).
- Galaxy is just another client to your scheduler.

Galaxy is available ...

- As a free (for everyone) web service

  http://usegalaxy.org

- As open source software

  http://getgalaxy.org

- *On the Cloud*

  http://usegalaxy.org/cloud

Galaxy CloudMan

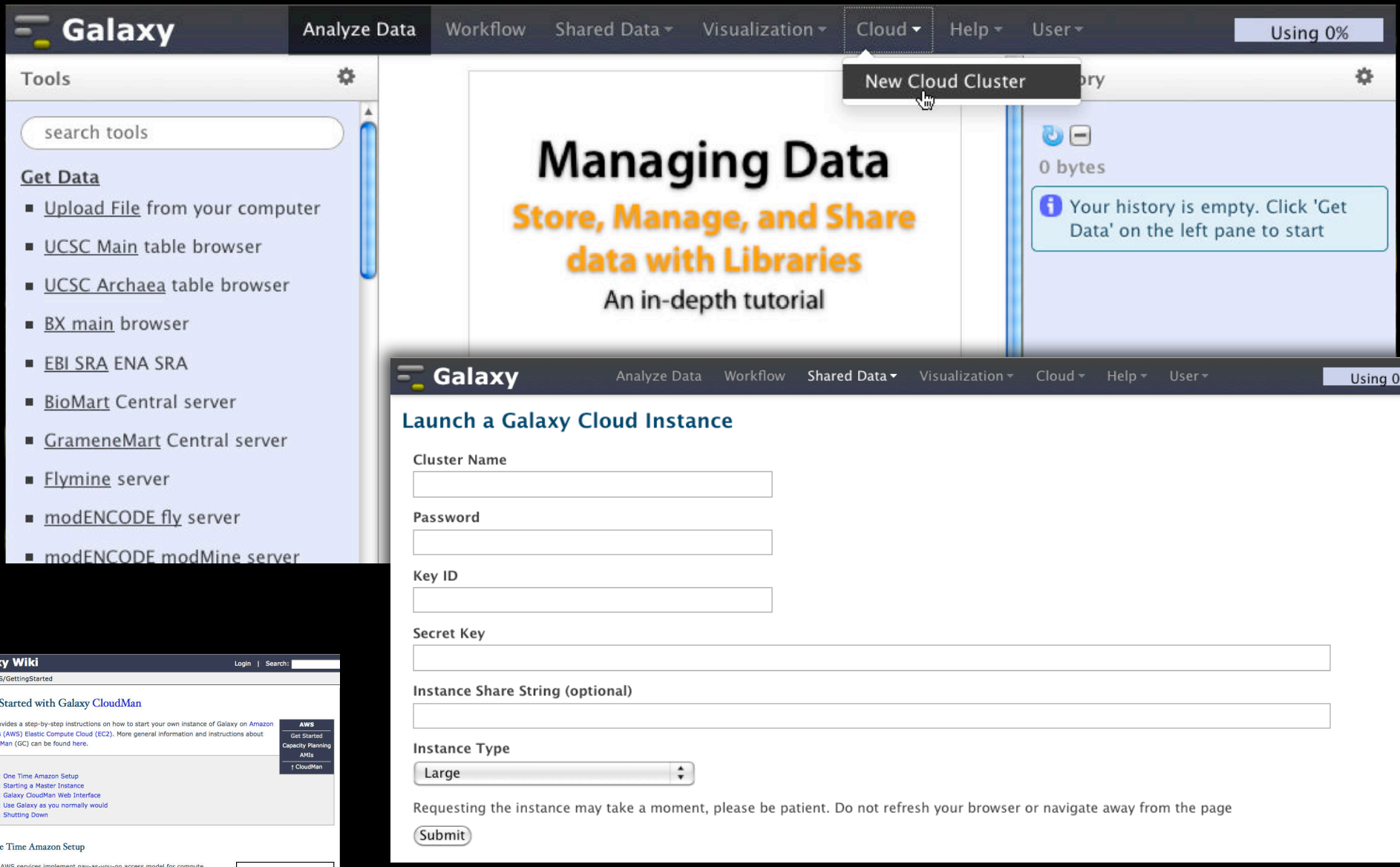
http://usegalaxy.org/cloud

- Start with a **fully configured and populated** (tools and data) Galaxy instance.
- Allows you to scale up and down your compute assets as needed.
- Someone else manages the data center.
- **We are using this today.**

- **You will set up an instance today**

http://aws.amazon.com/education

Instant CloudMan



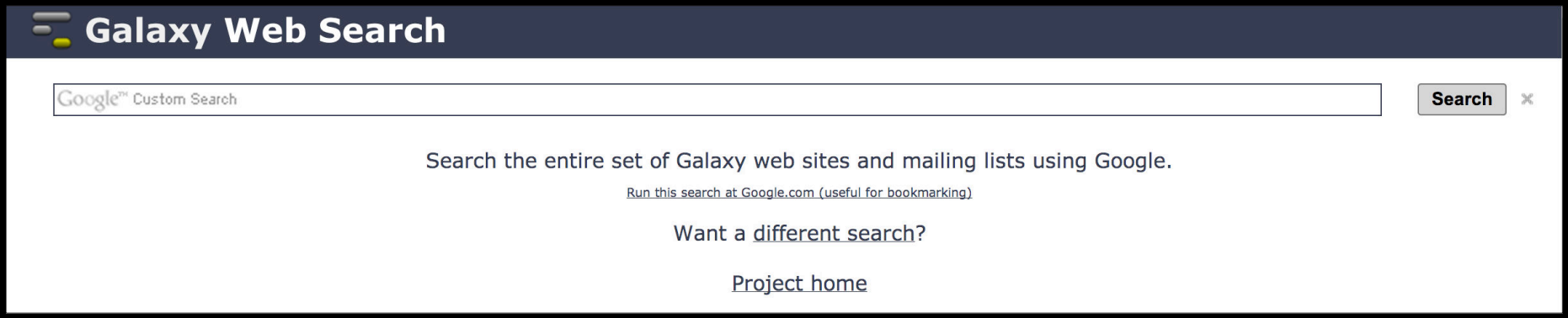
Or, step by step

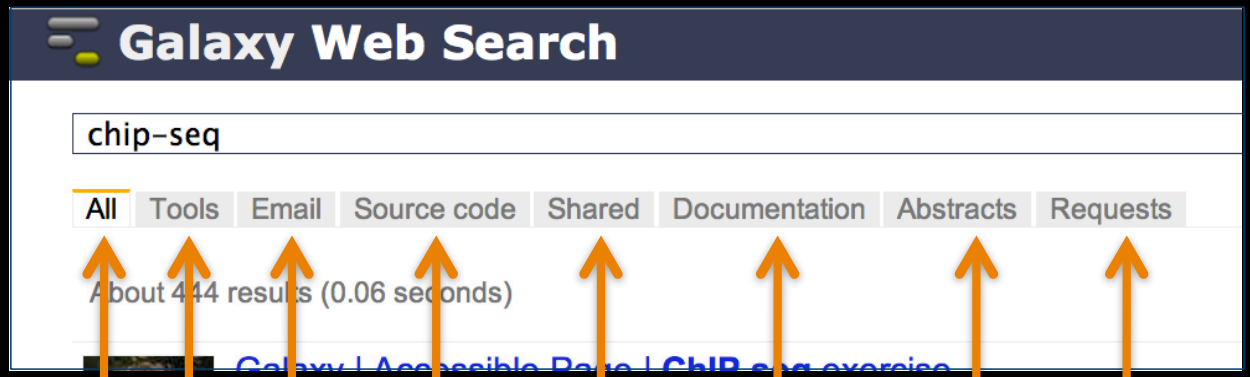
Galaxy Resources and Community

Unified Search
Mailing Lists (very active)
Issues Board
Events Calendar, News Feed
Community Wiki
GalaxyAdmins
Screencasts
Tool Shed
Public Installs
CiteULike group, Mendeley mirror
Annual Community Meting

http://wiki.galaxyproject.org

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

Mailing Lists

http://wiki.galaxyproject.org/MailingLists

Galaxy-Announce

Project announcements, low volume, moderated

Low volume (42 posts, 1600 members in 2012)

Galaxy-User

Questions about using Galaxy and usegalaxy.org

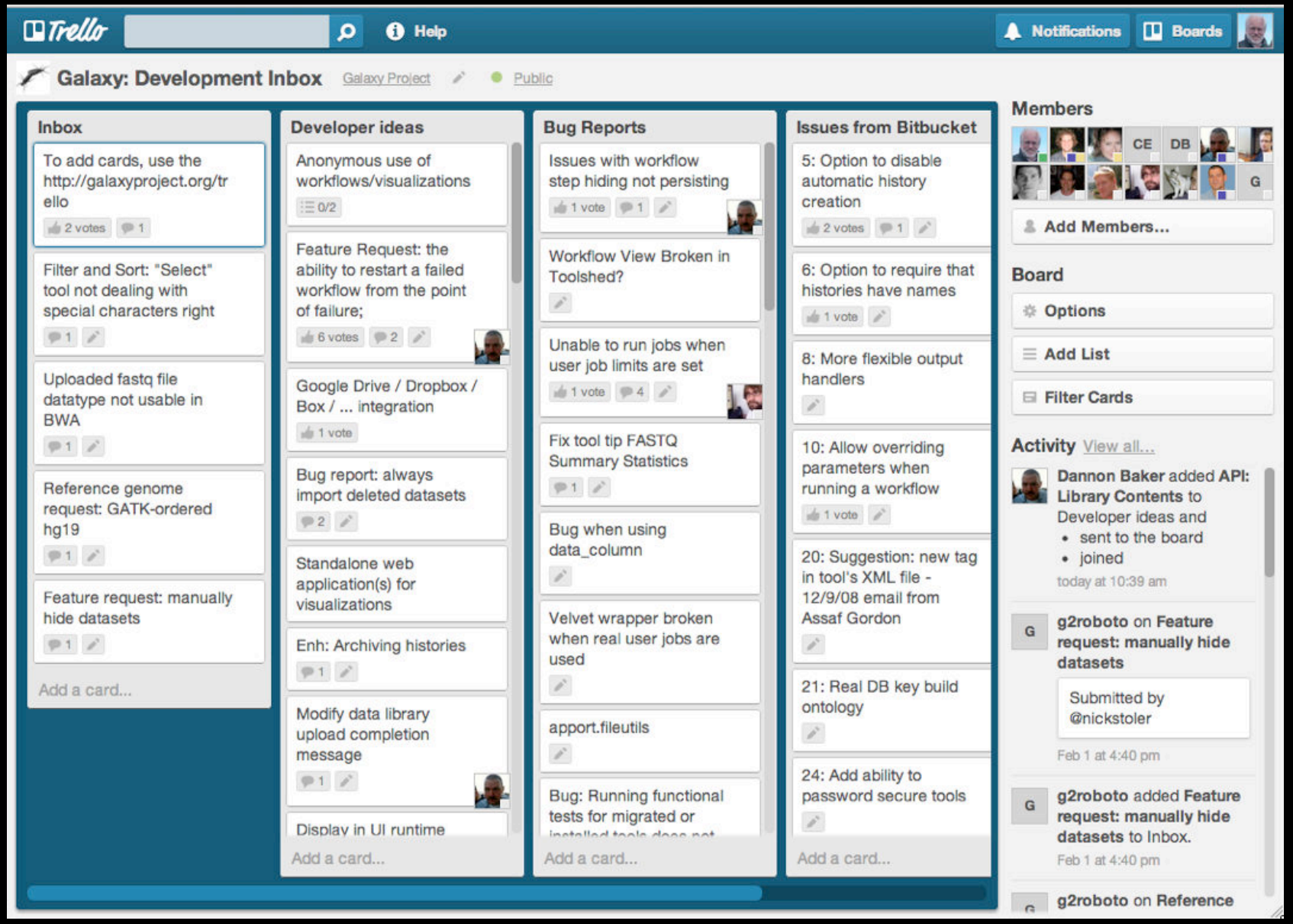
High volume (2900 posts, 2700 members in 2012)

Galaxy-Dev

Questions about developing for and deploying Galaxy

High volume (4500 posts, 850 members in 2012)

Community can create, vote and comment on **issues**



http://bit.ly/gxyissues

Events

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 relevant to the 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, please add it here or send it to outreach@galaxyproject.org .

Upcoming Events

Date	Topic/Event	Venue/Location
February 4	*Introduction to Galaxy Boot Camp*	UC Davis Bioinformatics Core Davis, California, United States
March 2-5	*Accessible, Transparent and Reproducible Analysis With Galaxy*, part of SW1: Application of NGS Platforms for Whole Transcriptome and Genome Analysis *Galaxy for Core Facilities*, part of "W6: Community Resource Solutions to Analyzing Large Genomic Data Sets"	**ABRF 2013** Palm Springs, California, United States
March 26-28	*RNA Technologies and Analysis Workshop*	DOE JGI User Meeting
April 5-6	*2013 GMOD Meeting*	Cambridge, United Kingdom, immediately prior to Biocuration 2013
April 7-10	*GO Galaxy Workshop*	Biocuration 2013, Cambridge, United Kingdom
April 9-11	Workshop: *Integrated Research Data Management for Next Gen Sequencing Analysis Using Galaxy and Globus Online Software-as-a-Service* Talk: *Integrated Research Data management and Analysis in NGS using Globus Online, Galaxy and Amazon Web Services*	BioIT World, Boston, Massachusetts, United States
May 14-16	Tutorial: *Exploring and Enabling Biomedical Data Analysis with Galaxy*	Great Lakes Bioinformatics Conference (GLBIO) 2013, Pittsburgh, Pennsylvania, United States
May 21 May 29	*Initiation à l'utilisation de Galaxy* Les deux ateliers sont maintenant complets	Cycle "Bioinformatique par la pratique" 2013, INRA Jouy-en-Josas, France
May 22 May 30	Analyse de données issues de séquenceurs nouvelle génération sous Galaxy Les deux ateliers sont maintenant complets	
June 6-7	*Informatics on High Throughput Sequencing Data Workshop*	Toronto, Ontario, Cananda

News

News

Announcements of interest to the Galaxy Community. These can include items from the Galaxy Team or the Galaxy community and can address anything that is of wide interest to the community.

The Galaxy News is also available as an RSS feed.

See Add a News Item below for how to get an item on this page, and the RSS feed. Older news items are available in the Galaxy News Archive.

See also

- Distribution News Briefs
- Galaxy Updates
- Galaxy on Twitter
- Events
- Learn
- Support
- About the Galaxy Project

News Items

February 2013 Galaxy Update
GCC2013 Training Day Topics: Vote!
Galaxy Project Openings
Jan 11, 2013 Distribution & News Brief
January 2013 GalaxyAdmins
January 2013 Galaxy Update
Dec 20, 2012 Distribution & News Brief
Galaxy Internships @ EMBL
Nominate GCC2013 Training Topics
Dec 3, 2012 Distribution & News Brief
December 2012 Galaxy Update
Nov 14, 2012 Distribution & News Brief
NGS Analysis by Viz. with Trackster
November 2012 GalaxyAdmins

News Archive

News Items

February 2013 Galaxy Update

The February 2013 Galaxy Update is now available.

Highlights:

- Three new public Galaxy servers
- New papers
- Open Positions at five different institutions
- GCC2013 Training Day Topic voting, Registration, and Sponsorships
- January GalaxyAdmins Web Meetup slides and screencast
- Other Upcoming Events and Deadlines
- Galaxy Distributions
- Tool Shed Contributions
- Other News

If you have anything you would like to see in the March *Galaxy Update*, please let us know.

Dave Clements and the Galaxy Team

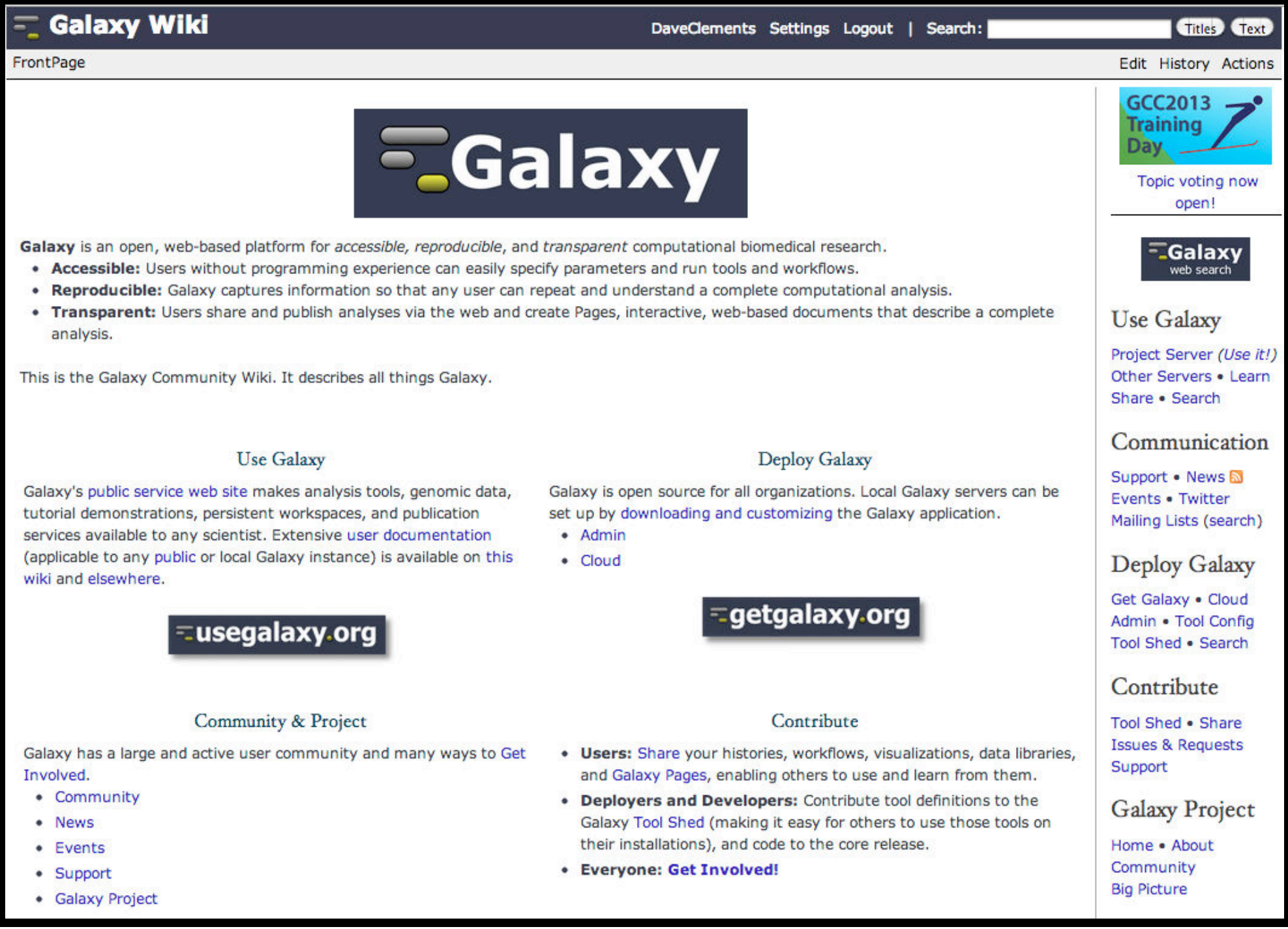
Posted to the Galaxy News on 2013-02-01

GCC2013 Training Day Topics: Vote!

A list of possible topics for the GCC2013 Training Day is now available. Please take a few minutes to review these possibilities and then vote for your favorite three topics.*

Your votes will determine not only the topics that are offered, but also which topics should be offered more than once, assigned to which rooms, and which ones should not be scheduled at the same time. Your vote matters.

http://wiki.galaxyproject.org



Galaxy Community Conference

Registration & abstract submission opens February 22

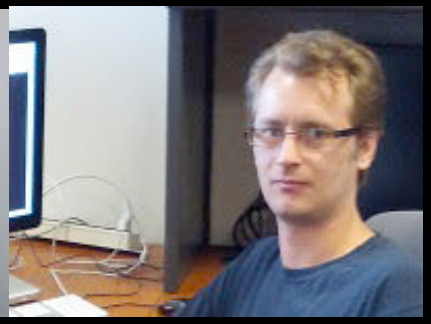
http://galaxyproject.org/GCC2013

Enis Afgan

Guru Ananda

Dannon Baker

Dan Blankenberg



Dave Bouvier

Dave Clements

Nate Coraor

Carl Eberhard

Jeremy Goecks

Jen Jackson

Greg von Kuster

Ross Lazarus

Anton Nekrutenko

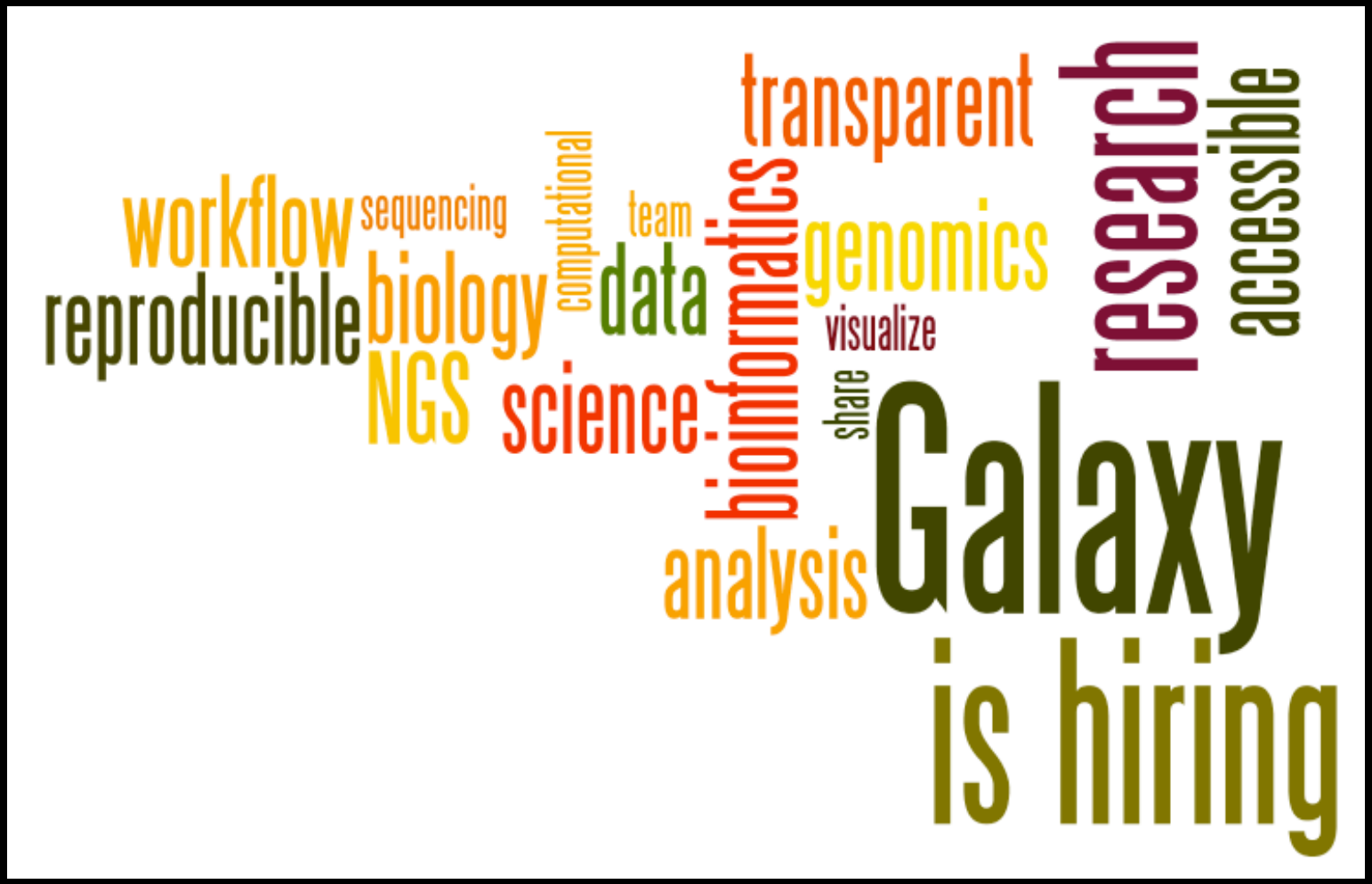
James Taylor

?

The Galaxy Team

http://wiki.galaxyproject.org/GalaxyTeam

Galaxy is hiring post-docs and software engineers at both Emory and Penn State.



Please help.

http://wiki.galaxyproject.org/GalaxyIsHiring

Agenda

9:00	Welcome
9:20	Basic Analysis with Galaxy
10:20	Basic Analysis into Reusable Workflows
10:40	Break
11:00	RNA-Seq Example Part I
12:00	Galaxy Project Overview
12:20	Lunch
1:05	RNA-Seq Example Part II Cufflinks, Visualization and Visual Analytics
1:55	Sharing, Publishing and Reproducibility
2:15	Break
2:35	Setting up your own Galaxy Cluster on AWS
4:30	Done

Agenda

9:00	Welcome
9:20	Basic Analysis with Galaxy
10:20	Basic Analysis into Reusable Workflows
10:40	Break
11:00	RNA-Seq Example Part I
12:00	Galaxy Project Overview
12:20	Lunch
1:05	RNA-Seq Example Part II Cufflinks, Visualization and Visual Analytics
1:55	Sharing, Publishing and Reproducibility
2:15	Break
2:35	Setting up your own Galaxy Cluster on AWS
4:30	Done

RNA-seq Exercise: A Plan

- ...
- Trim as we see fit.
- Map the reads to the human reference using Tophat
- Run Cufflinks on Tophat output to assemble reads into transcripts
  - *Imagine pages and pages of discussion on the intricacies and pitfalls of RNA-seq transcript prediction here.*

http://bit.ly/gxyRNASEX

RNA-seq Exercise: A Plan

- ...
- Map the reads to the human reference using Tophat
- Run Cufflinks on Tophat output to assemble reads into transcripts
  - *Imagine pages and pages of discussion on the intricacies and pitfalls of RNA-seq transcript prediction here.*
  - Visualize it

http://bit.ly/gxyRNASEX

Visualizing Genomics

Supported external browsers

- UCSC
- Ensembl
- GBrowse
- IGB
- IGV

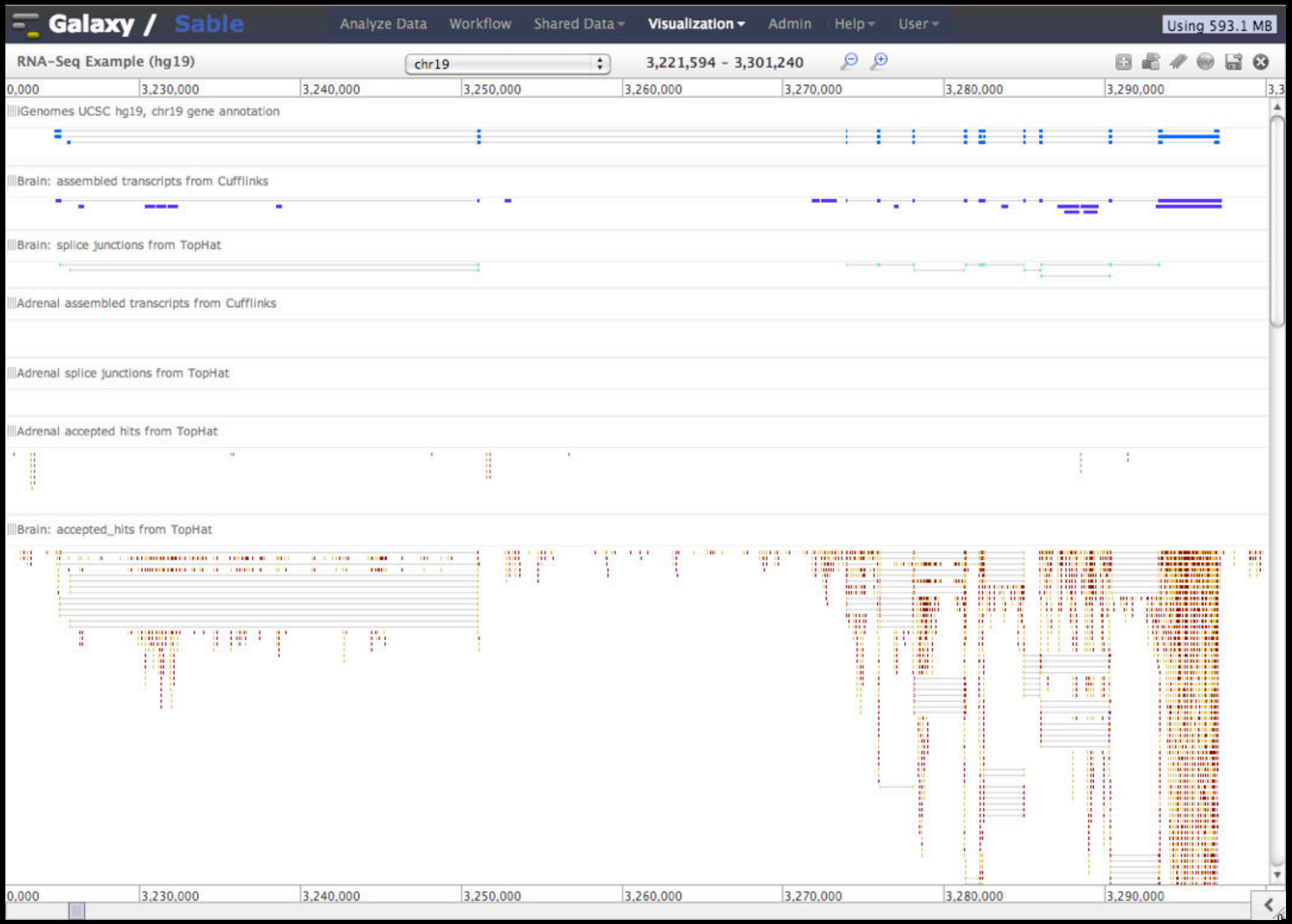
Traditional browser strengths:

- Showing what is nearby
- what else is happening here
- highlighting correlations
- integrating many datasets

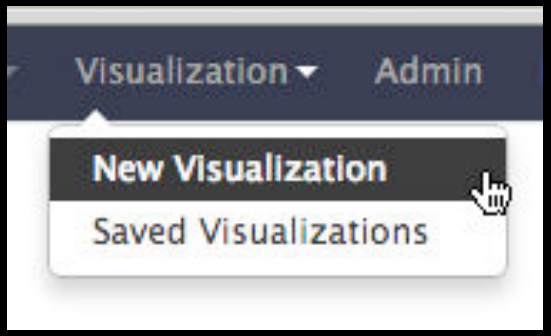
But, *wouldn't it be nice to*

- Use visualization to evaluate and refine analyses?
- Expose some basic analyses in visualization to make it more informative?
- Make that analyze-visualize-refine loop seamless and fast? That is, integrate the two?
- Use visualization to learn tools and explore their parameter space?
- Not be tied to a predefined reference genome?

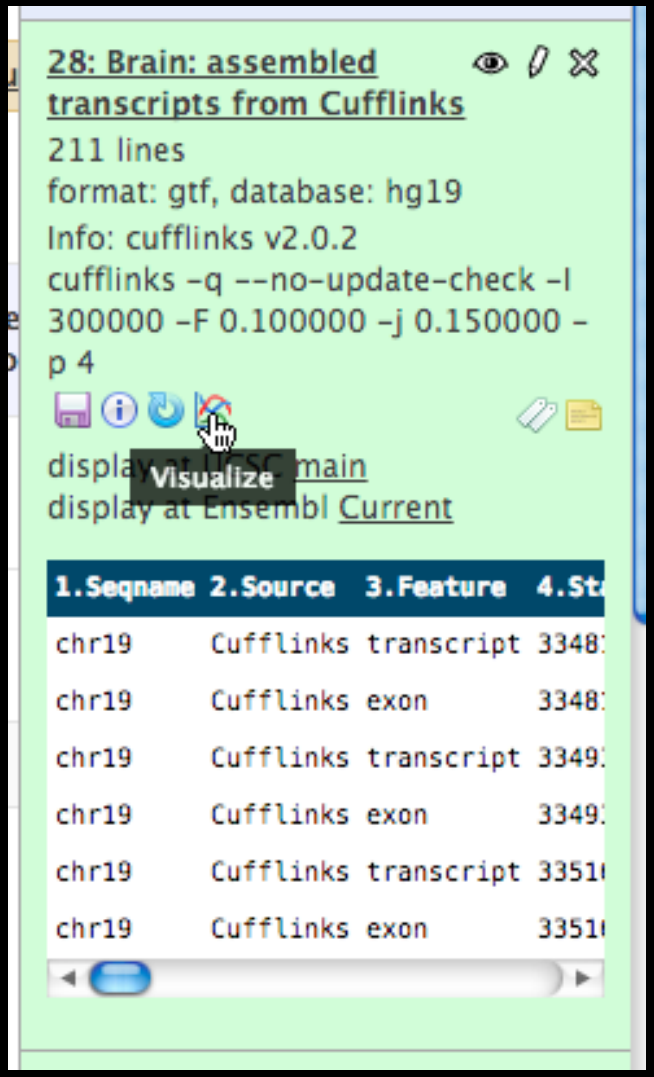
Trackster: Galaxy's embedded track browser



Create a visualization in Galaxy



or



Isn't it nice to

- To do all those things we talked about?
  - Use visualization to evaluate and refine analyses?
  - Expose some basic analyses in visualization to make it more informative?
  - Make that analyze-visualize-refine loop seamless and fast? That is, integrate the two?
  - Use visualization to learn tools and explore their parameter space?
  - Not be tied to a predefined reference genome?

Agenda

9:00 Welcome
9:20 Basic Analysis with Galaxy
10:20 Basic Analysis into Reusable Workflows
10:40 Break
11:00 RNA-Seq Example Part I
12:00 Galaxy Project Overview
12:20 Lunch
1:05 RNA-Seq Example Part II
Cufflinks, Visualization and Visual Analytics
1:55 Sharing, Publishing and Reproducibility
2:15 Break
2:35 Setting up your own Galaxy Cluster on AWS
4:30 Done

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

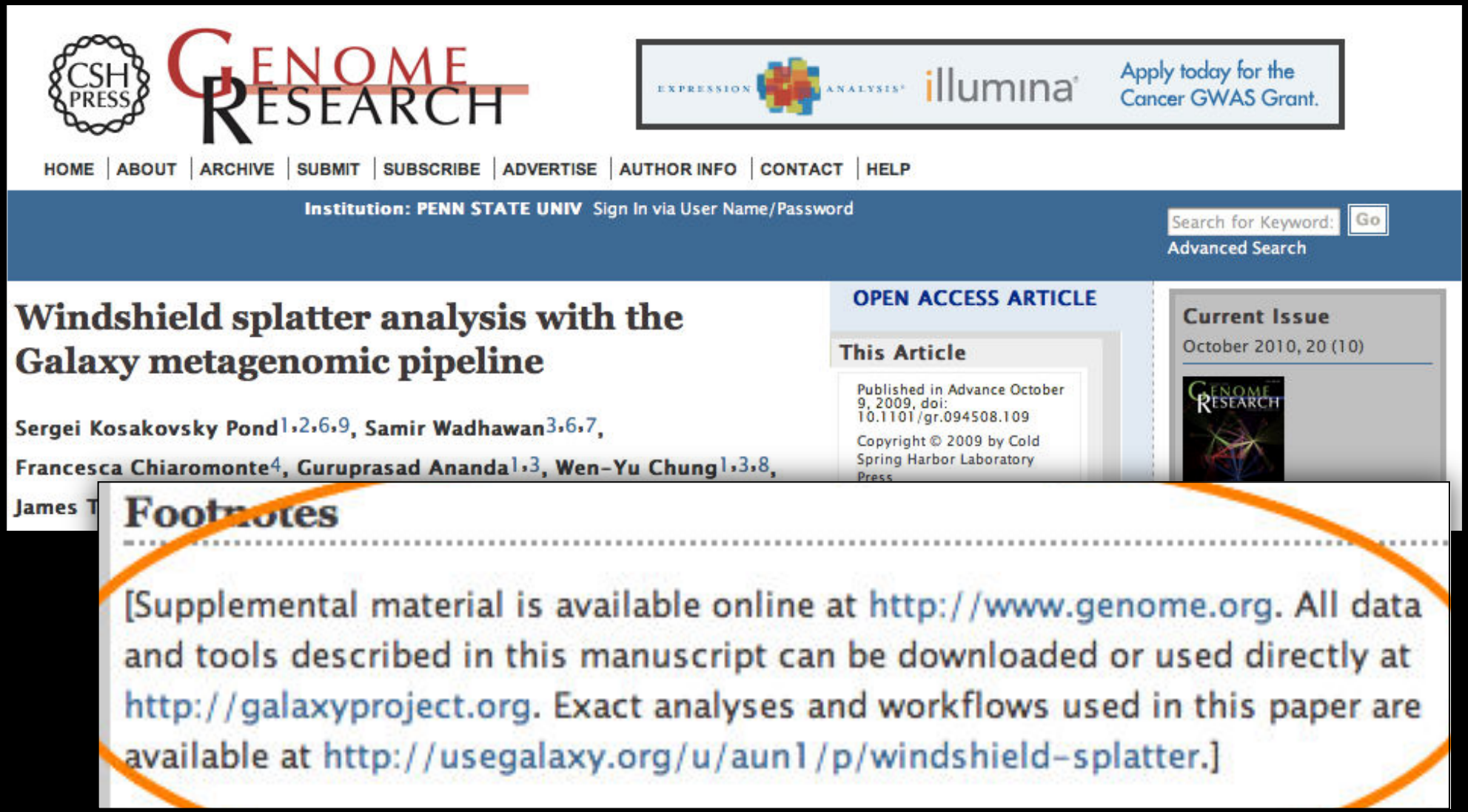
Reproducibility: Everybody talks about it, but ...

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



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

Sharing for Galaxy Administrators Too

Data Libraries

Make data easy to find

Genome Builds

Care about a particular subset of life?

Galaxy Tool Shed

Wrapping tools and datatypes

Agenda

9:00 Welcome
9:20 Basic Analysis with Galaxy
10:20 Basic Analysis into Reusable Workflows
10:40 Break
11:00 RNA-Seq Example Part I
12:00 Galaxy Project Overview
12:20 Lunch
1:05 RNA-Seq Example Part II
Cufflinks, Visualization and Visual Analytics
1:55 Sharing, Publishing and Reproducibility
2:15 Break, *but first* ...
2:35 Setting up your own Galaxy Cluster on AWS
4:30 Done

Acknowledgements

Mo Heyderian
Karen Reddy
Barbara Sollner-Webb

Steven Salzberg
The Galaxy Team
You!

Johns Hopkins University
Biological Chemistry
Center for Computational Biology

AWS Education Grant

NIH NSF Huck Institute
Penn State University Emory University

Feedback Please!

http://bit.ly/JHUfeedback

Thanks

Dave Clements

Galaxy Project
Emory University

clements@galaxyproject.org

Agenda

9:00 Welcome
9:20 Basic Analysis with Galaxy
10:20 Basic Analysis into Reusable Workflows
10:40 Break
11:00 RNA-Seq Example Part I
12:00 Galaxy Project Overview
12:20 Lunch
1:05 RNA-Seq Example Part II
Cufflinks, Visualization and Visual Analytics
1:55 Sharing, Publishing and Reproducibility
2:15 Break
2:35 Setting up your own Galaxy Cluster on AWS
4:30 Done

Agenda

9:00 Welcome
9:20 Basic Analysis with Galaxy
10:20 Basic Analysis into Reusable Workflows
10:40 Break
11:00 RNA-Seq Example Part I
12:00 Galaxy Project Overview
12:20 Lunch
1:05 RNA-Seq Example Part II
Cufflinks, Visualization and Visual Analytics
1:55 Sharing, Publishing and Reproducibility
2:15 Break
2:35 Setting up your own Galaxy Cluster on AWS
4:30 Done

AWS Credentials

http://bit.ly/JHUmon