



Community Building, Outreach, and Support in Online Biological Communities

Dave Clements
Galaxy Team
April 6, 2012
<http://galaxyproject.org>

Galaxy Project & Community

Mailing Lists

Wiki

News / Twitter

Social Bookmarking

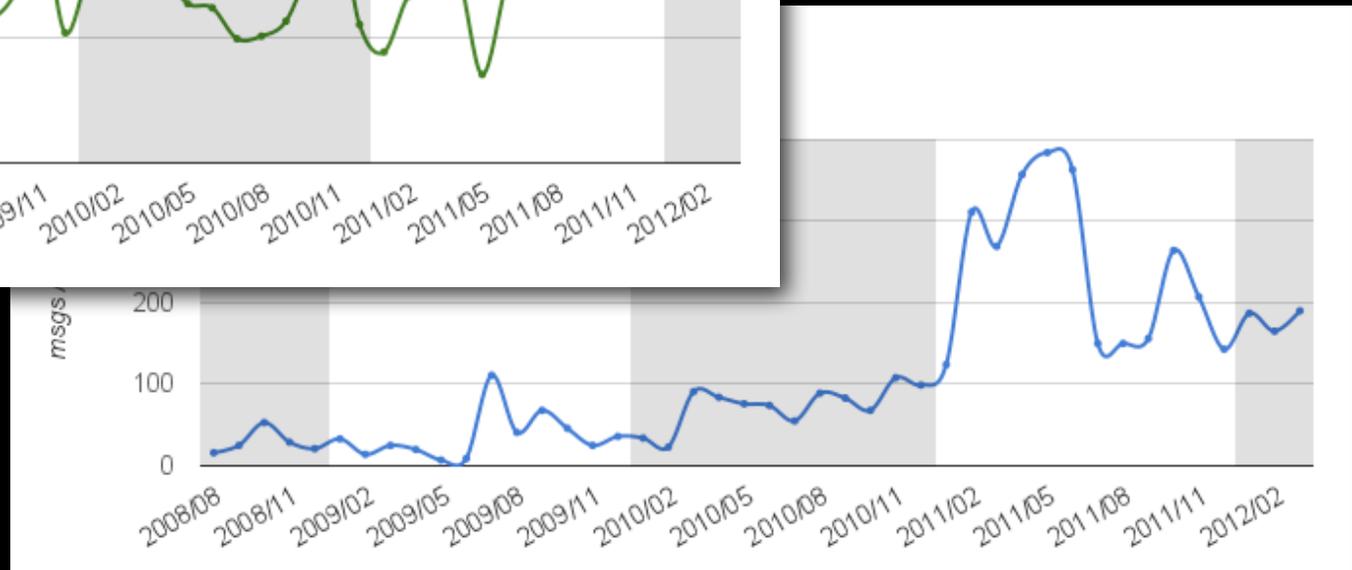
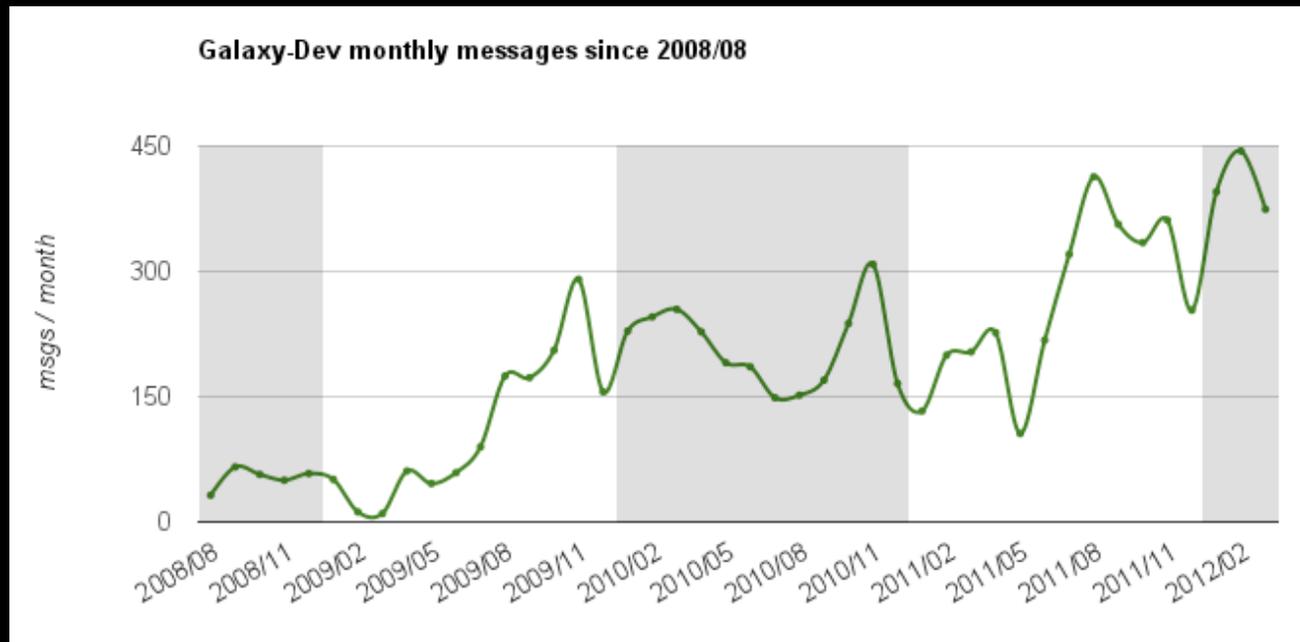
Google Custom Search

Galaxy Tool Shed

Mailing Lists

Very active **galaxy-dev** & **galaxy-user** lists

Led to new low-volume, moderated **galaxy-announce** List



Mailing Lists

Use **Redmine** for tracking

galaxy-dev, galaxy-user threads create an issue

Automatically assigned to last team member to respond

Can close issue by BCC a special email address.

Open issue tracking is done by Jen Jackson

Redmine-Email bridge by Dannon Baker

<http://redmine.org/>

Lists are archived at [Nabble.com](http://nabble.com)

Modified to be an archive rather than a forum

Use Nabble's custom domain name support

Way better search interface than Mailman or Sourceforge

<http://dev.list.galaxyproject.org/>

[Galaxy Development List Archive](#)

[Login](#) [Register](#)

Galaxy Development List Archive

This forum is an archive for the mailing list galaxy-dev@bx.psu.edu ([more options](#)) Messages posted here will be sent to this mailing list.

Archive for the [Galaxy-Dev](#) mailing list. If you have a question about deploying, enhancing, tuning or adding to a [Galaxy](#) instance then this is a good place to find an answer.



[Galaxy](#) is an open, web-based platform for *accessible, reproducible, and transparent* computational biomedical research.

- **Accessibility:** Galaxy enables users without programming experience to easily specify parameters and run tools and workflows.
- **Reproducibility:** Galaxy captures all information necessary so that any user can repeat and understand a complete computational analysis.
- **Transparency:** Galaxy enables users to share and publish analyses via the web and create Pages--interactive, web-based documents that describe a complete analysis.

Galaxy is open source for all organizations. The [public Galaxy service](#) makes analysis tools, genomic data, tutorial demonstrations, persistent workspaces, and publication services available to any scientist that has access to the Internet. Local Galaxy servers can be set up by downloading the Galaxy application and customizing it to meet particular needs. You can also [search the archives](#) of the [Galaxy-User](#) mailing list. Galaxy is a part of the [Generic Model Organism Database \(GMOD\) project](#).

[Subscribe to Galaxy Dev](#)

[People](#) [Options](#) ▾

1 2 3 4 ... 113

Topics (3935)	Replies	Last Post	Views
Can anyone give an example of how to setup the apache2 to handle external user authentication(against postgresql db)? by JIE CHEN	☆ 0	5:27pm by JIE CHEN	0
Workflow API (runtime modification of tool parameters) by Richard Park	☆ 4	4:17pm by Richard Park	10
directories as inputs/dataset file extensions by Aaron Gallagher	☆ 4	2:54pm by Brad Langhorst	8
Best practice about including java.jar in toolshed tool by Joachim Jacob	☆ 1	11:02am by Greg Von Kuster	0
Galaxy loading never ends by Makis Ladoukakis	☆ 0	9:11am by Makis Ladoukakis	0
Auto delete dataset after workflow run? by Praveen Raj Somaraja...	☆ 2	7:55am by Praveen Raj Somaraja...	0

Wiki

Moved from Bibucket.org in 2011

Now using **MoinMoin**

Markup in Creole, native Moin

It's been a hit

Looked at **Mediawiki**, others

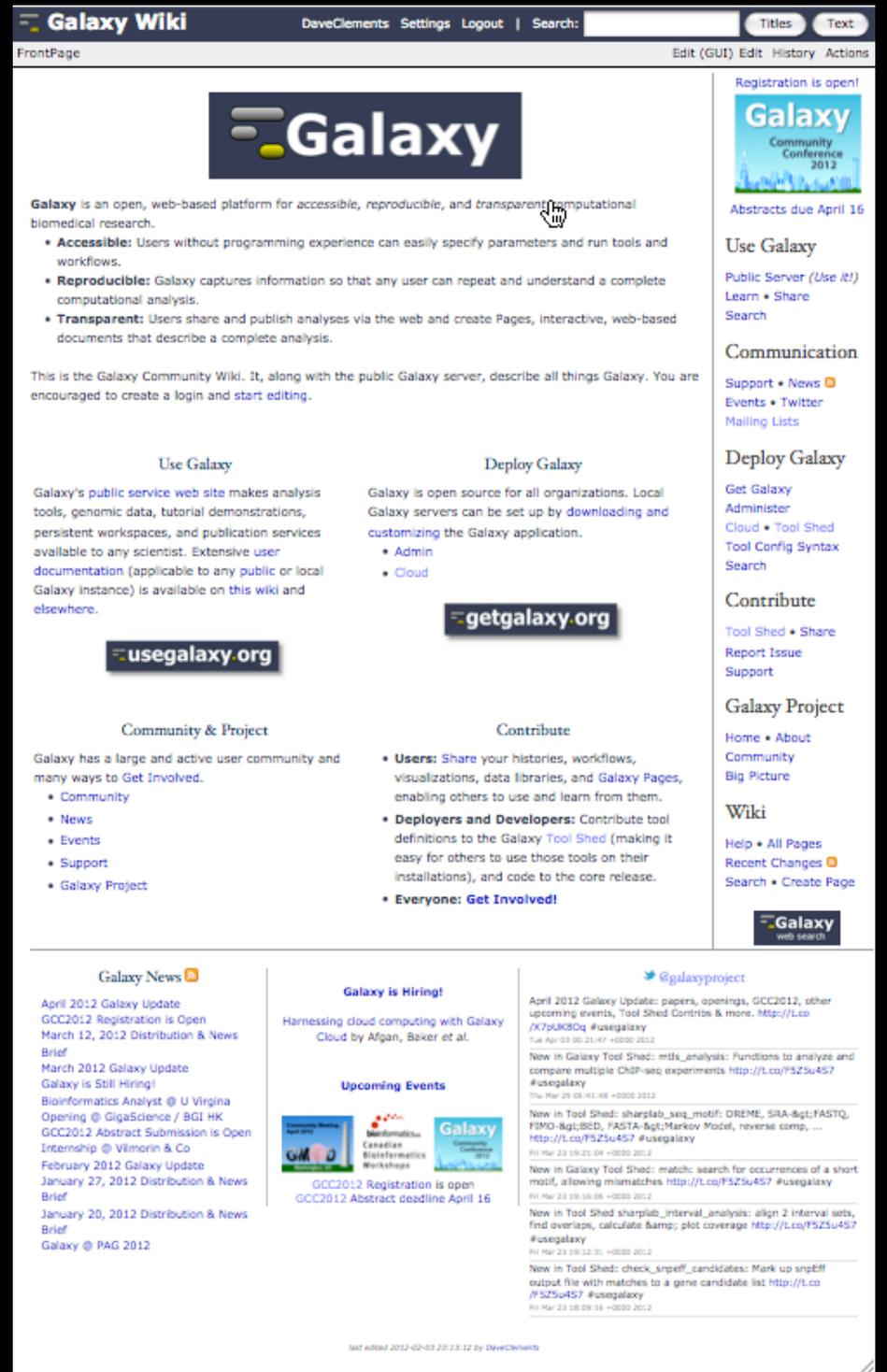
We have a strong Python slant

Mediawiki not built for

projects

<http://galaxyproject.org/wiki>

<http://wikimatrix.org/>



The screenshot shows the Galaxy Wiki homepage. At the top, there is a navigation bar with "Galaxy Wiki" on the left, "DaveClements Settings Logout" in the middle, and a search box on the right. Below the navigation bar, the main content area features the Galaxy logo and a brief description of the platform. A list of key features is provided: Accessible, Reproducible, and Transparent. The page is organized into several columns and sections, including "Use Galaxy", "Deploy Galaxy", "Community & Project", and "Contribute". A sidebar on the right contains links for "Registration is open!", "Use Galaxy", "Communication", "Deploy Galaxy", "Contribute", and "Galaxy Project". At the bottom, there are sections for "Galaxy News", "Galaxy is Hiring!", "Upcoming Events", and a list of recent updates.

Galaxy Wiki DaveClements Settings Logout | Search: [] Titles Text

FrontPage Edit (GUI) Edit History Actions

Galaxy

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, along with the public Galaxy server, describe all things Galaxy. You are encouraged to create a login and start editing.

Use Galaxy

Galaxy's public service web site 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 on [this wiki](#) and elsewhere.

Deploy Galaxy

Galaxy is open source for all organizations. Local Galaxy servers can be set up by downloading and customizing the Galaxy application.

- Admin
- Cloud

usegalaxy.org

getgalaxy.org

Community & Project

Galaxy has a large and active user community and many ways to Get Involved.

- Community
- News
- Events
- Support
- Galaxy Project

Contribute

- **Users:** Share your histories, workflows, visualizations, data libraries, and Galaxy Pages, enabling others to use and learn from them.
- **Deployers and Developers:** Contribute tool definitions to the Galaxy Tool Shed (making it easy for others to use those tools on their installations), and code to the core release.
- **Everyone: Get Involved!**

Galaxy News

April 2012 Galaxy Update
GCC2012 Registration is Open
March 12, 2012 Distribution & News Brief
March 2012 Galaxy Update
Galaxy is Still Hiring!
Bioinformatics Analyst @ U Virginia
Opening @ GigaScience / BGI HK
GCC2012 Abstract Submission is Open
Internship @ Vilmorin & Co
February 2012 Galaxy Update
January 27, 2012 Distribution & News Brief
January 20, 2012 Distribution & News Brief
Galaxy @ PAG 2012

Galaxy is Hiring!

Harnessing cloud computing with Galaxy Cloud by Afgan, Baker et al.

Upcoming Events

Galaxy Community Meeting April 2012
Bioinformatics Canadian Bioinformatics Workshops
Galaxy Community Conference 2012

GCC2012 Registration is open
GCC2012 Abstract deadline April 16

@galaxyproject

April 2012 Galaxy Update: papers, openings, GCC2012, other upcoming events, Tool Shed Contribs & more. [#usegalaxy](http://Lco/X7pUK8Dq)
Tue Apr 03 00:21:47 +0000 2012

New in Galaxy Tool Shed: mts_analysis: Functions to analyze and compare multiple ChIP-seq experiments [#usegalaxy](http://Lco/P5Z5u457)
Thu Mar 29 08:41:48 +0000 2012

New in Tool Shed: sharpseq_motif: DREME, SRA-Seq, FASTQ, FIMO-Seq, BED, FASTA-Seq, Markov Model, reverse comp, ... [#usegalaxy](http://Lco/P5Z5u457)
Thu Mar 23 19:21:04 +0000 2012

New in Galaxy Tool Shed: match_search: search for occurrences of a short motif, allowing mismatches [#usegalaxy](http://Lco/P5Z5u457)
Fri Mar 23 19:20:06 +0000 2012

New in Tool Shed sharpseq_interval_analysis: align 2 interval sets, find overlaps, calculate & plot coverage [#usegalaxy](http://Lco/P5Z5u457)
Fri Mar 23 19:20:06 +0000 2012

New in Tool Shed: check_snpeff_candidates: Mark up snpEff output file with matches to a gene candidate list [#usegalaxy](http://Lco/P5Z5u457)
Fri Mar 23 18:59:16 +0000 2012

last edited 2012-02-02 22:12 by DaveClements

News

Development News Briefs

Accompany each release; about releases

<http://galaxyproject.org/wiki/DevNewsBriefs>

Galaxy News and Monthly Updates

Started in 2011 and 2012; web page & RSS feed

Modeled on GMOD News

<http://galaxyproject.org/wiki/News>

Twitter

Generates a lot of engagement

Use Bit.ly to track what gets interest

Appropriate for items that will only interest some

<https://twitter.com/#!/galaxyproject>

Social Bookmarking of Papers

Use CiteULike

Mirrored at Mendeley (Stephen Turner @ UVA)

Use tags to classify papers:

- ✦ **Project**: papers about Galaxy
- ✦ **Shared**: papers that use Galaxy's sharing and publishing features
- ✦ **Howto**: detailed tutorials that use Galaxy
- ✦ **Methods**: papers that used Galaxy in methods
- ✦ **Workflow**: papers about workflow
- ✦ **Other**: a very large misc.

<http://www.citeulike.org/group/16008>

Google Custom Search

Google Custom Search is very useful when

- ✦ You have information in several places
- ✦ You can identify specific areas of interest and have URLs that go with them

But I can't get Nabble archives to be searched!

<http://galaxyproject.org/search/>

Galaxy Tool Shed

A **Galaxy Tool Shed** is a repository for sharing tools, datatypes, and workflows

Once a tool is defined in one Galaxy instance, it is a **small step to make it available to others**

Built on top of **Mercurial** (similar to Git)

Moving tools out of distribution and into the **Main Tool Shed:**

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

What's Coming: Events



Date	Topic/Event	Venue/Location	Contact
April 4-5	NBIC NGS Exome Sequencing & Variant Calling course Including a session on <i>Read Mapping and Variant Calling with Galaxy</i>	Radboud UMC, Nijmegen	Hailiang "Leon" Mei
April 5-6	April 2012 GMOD Meeting Including a <i>Galaxy Workshop</i> on evening of April 5	Washington DC, immediately following <i>Biocuration 2012</i>	Dave Clements
April 24-26	Leveraging SaaS for Next-Gen Sequencing: Case Study with the Galaxy Community Pre-Conference Workshop	Bio-IT World , Boston, Massachusetts, United States	Ravi Madduri, Elizabeth Bartom
June 6-8	Dynamically Scalable, Accessible Analysis with Galaxy Cloud	Bio-IT World Asia , Singapore	James Taylor
June 11-12	Informatics on High Throughput Sequencing Data Workshop	Toronto, Ontario, Canada	Francis Ouellette
June 11-19	Next-gen Sequencing in Evolutionary Biology Course	US National Evolutionary Synthesis Center (NESCent), Durham, North Carolina, United States	Sergei Kosakovsky Pond
July 11	Reproducible workflows for next generation sequencing analysis	Nowgen, University of Manchester, United Kingdom	✉ Tom Hancocks
July 13-17	Bioinformatics Software Interoperability SIG (BSI-SIG)	ISMB 2012 , Long Beach, California, United States	Anton Nekrutenko
July 25-27	2012 Galaxy Community Conference Early registration and abstract submission are now open.	Chicago, Illinois, United States	Dave Clements
September 27-29	Beyond the Genome 2012	Harvard Medical School, Boston, Massachusetts	James Taylor
November 6-10	Galaxy 101: Data Integration, Analysis and Sharing and Working with High-Throughput Data and Data Visualization workshops	American Society of Human Genetics (ASHG) , San Francisco, California, United States	✉ Galaxy Outreach

<http://galaxyproject.org/wiki/Events>



Galaxy
Community
Conference
2012

July
25-27

UIC
Forum
University of Illinois at Chicago

<http://galaxyproject.org/GCC2012>

UIC  

The Galaxy community's annual gathering

A new Training Day added July 25

- ✦ 7 topics, 3 parallel tracks, 12 sessions
- ✦ 1) Intro, 2) Installing, 3) CloudMan, 4) Integrating Tools & Sources, 5) API, 6) Tool Shed, 7) Ion Torrent SDK

Key Dates

- ✦ **April 16: Abstracts due**
- ✦ **June 11: Early registration ends (early reg is cheap)**





Thanks

Dave Clements
clements@galaxyproject.org

Galaxy Team

April 5, 2012

<http://galaxyproject.org>

Galaxy Analysis Workspace

The screenshot displays the Galaxy Analysis Workspace interface. The browser address bar shows <http://main.g2.bx.psu.edu/>. The navigation menu includes **Analyze Data**, **Workflow**, **Shared Data**, **Visualization**, **Help**, and **User**.

Tools Panel (Left): Lists various tools such as **Get Data**, **Send Data**, **ENCODE Tools**, **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**, **Metagenomic analyses**, **EMBOSS**, **NGS TOOLBOX BETA**, **NGS: QC and manipulation**, **NGS: Mapping**, **NGS: SAM Tools**, **NGS: Indel Analysis**, **NGS: Peak Calling**, **RGENETICS**, **SNP/WGA: Data; Filters**, **SNP/WGA: QC; LD; Plots**, **SNP/WGA: Statistical Models**, and **Workflows**.

Main Panel: Map with Bowtie for Illumina

Will you select a reference genome from your history or use a built-in index?:
 Built-ins were indexed using default options

Select a reference genome:
 if your genome of interest is not listed - contact Galaxy team

Is this library mate-paired?:

Forward FASTQ file:
 Must have Sanger-scaled quality values with ASCII offset 33

Reverse FASTQ file:
 Must have Sanger-scaled quality values with ASCII offset 33

Maximum insert size for valid paired-end alignments (-X):

The upstream/downstream mate orientation for valid paired-end alignment against the forward reference strand (--fr/--rf/--ff):

Bowtie settings to use:
 For most mapping needs use Commonly used settings. If you want full control use Full parameter list

Suppress the header in the output SAM file:
 Bowtie produces SAM with several lines of header information by default

What it does
Bowtie is a short read aligner designed to be ultrafast and memory-efficient. It is developed by Ben Langmead and Cole Trapnell. Please cite: Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* 10:R25.

History Panel (Right): Shows a list of previous jobs with options to view, refresh, or delete. The jobs listed are:
15: Variants from sample E18, consensus different, in RefSeq Genes
14: UCSC mm9 RefSeq Genes
13: Variants from sample E18 where consensus base different than ref. base
10: Variants from sample E18
9: Generate pileup on data 8
8: SAM-to-BAM on data 7
7: Map with Bowtie for Illumina on data 6 and data 5
6: E18 PE.2 Reads Groomed, Trimmed
5: E18 PE.1 Reads Groomed, Trimmed
4: E18 PE.2 Reads Groomed
3: E18 PE.1 Reads Groomed
2: E18 PE.2 Reads
1: E18 PE.1 Reads