

DNA Sequence Bioinformatics Analysis with the Galaxy Platform

University of São Paulo, Brazil
28 July - 1 August 2014

Dave Clements
Johns Hopkins University

Robson Francisco de Souza
University of São Paulo

José Belazario
University of São Paulo



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

Just 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)

Galaxy-Announce

Project announcements, low volume, moderated


Low volume (47 posts in 2013, 3400+ members)

Galaxy-User (discontinued 2014/06)


Questions about using Galaxy and usegalaxy.org

High volume (1328 posts in 2013, 2600+ members)

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

 **Galaxy Web Search**

Google™ Custom Search

Search 

Search the entire set of Galaxy web sites and mailing lists using Google.

[Run this search at Google.com \(useful for bookmarking\)](#)

Want a [different search](#)?

[Project home](#)

Find

Everything on ...

Tools for ...

Email about ...


Source code for ...

Published Histories, Pages, Workflows, about ...

Documentation on ...

Papers using Galaxy for ...

Related feature requests

 **Galaxy Web Search**

chip-seq

All Tools Email Source code Shared Documentation Abstracts Requests

About 444 results (0.06 seconds)

[Galaxy | Accessible Page | ChIP-seq exercise](#)

Community: Public Galaxy Instances

<http://bit.ly/gxyServers>

Interested in:

ChIP-chip and ChIP-seq?

✓ Cistrome, Nebula

Statistical Analysis?

✓ Genomic Hyperbrowser

Protein synthesis?

✓ GWIPS-viz

de novo assembly?

✓ CBIIT Galaxy

Reasoning with ontologies?

✓ GO Galaxy

Repeats!

✓ RepeatExplorer

Over 60 public Galaxy servers

Community can create, vote and comment on issues

The screenshot displays the Trello interface for the 'Galaxy: Development' board. The board is organized into several columns, each representing a different category of issues or tasks. The columns are: Inbox, Tool Requests, Bug Reports, Ideas, Pull Requests / Patches, and Project in Planning. Each card in the board contains a title, a description, and a progress bar. The cards are sorted by their position in the column, with the most recent or important cards at the top. The right sidebar shows the 'Menu' section, which includes 'Members' (a list of user avatars), 'Activity' (a list of recent actions), and 'Tools' (a list of tools used in the board). The top navigation bar includes links to 'HOME', 'TOUR', 'GOLD', 'BUSINESS CLASS', and 'BLOG'. The top right corner has 'Sign Up' and 'Log In' buttons. The bottom of the board has a 'Want to subscribe, vote or comment on these cards?' section with a 'Sign up for free' button and a link to 'learn more about Trello'.

Galaxy: Development • Public

Want to subscribe, vote or comment on these cards? [Sign up for free](#) or [learn more about Trello](#)

Inbox

- To add cards, use <http://galaxyproject.org/trello> (4 votes, 1 comment)
- add ma seq metrics and downsample sam to picard tools (3 votes, 1 comment)
- Reference genomes (2 votes, 2 comments)
- Please merge patch to bowtie2 wrapper (adds support for mapping fasta files) (1 vote, 1 comment)
- R 3.0.2 woes on test.g2.bx.psu.edu - libgomp.so not found when job runs: [Rscript error while loading shared libraries: libgomp.so.1: cannot open shared object file: No such file or directory\n] - possibly execution node missing gfortran? (1 vote, 1 comment)
- unhandled exception when installing metaphlan from source repo (1 vote, 1 comment)

Tool Requests

- 595: Add SAMTools "Sort" (5 votes, 1 comment)
- 601: SAM-to-BAM tool enhancements (1 vote, 1 comment)
- Bug: some characters not permitted in 'add column' tool (2 votes, 5 comments)
- 307: A tool to produce a set of random intervals. (2 votes, 2 comments)
- Tool: Add tool to generate simulated reads to Main (2 votes, 1 comment)
- default max insert size of Bowtie2 should be increased (1 vote, 4 comments)
- Wrapper for bigWigToWig (1 vote, 1 comment)
- Converter Tool: SAM to BAM enhancements (1 vote, 1 comment)
- 607: Create new tool to "trim" coordinates to ref chrom lengths (1 vote, 1 comment)
- New Tool: convert IUPAC chars to N (5 votes, 1 comment)
- Optimize FASTQ tools. (1 vote, 1 comment)
- Tool 'Extract Genomic DNA' should parse GFF/GTF better so to include gene_id or transcript_id attributes (1 vote, 1 comment)
- Enh: tabular-to-fasta should let you choose how to concatenate the id string (1 vote, 1 comment)

Bug Reports

- Impersonate a user admin option broken when using external authentication (13 votes, 1 comment)
- Bug: SICKER on Main dependency issue (2 votes, 18 comments, 3/5)
- Toolshed: Installing multiple versions of the same tool results in separate entries in the tool panel. (1 vote, 14 comments)
- Profile Annotations bad values when "select all" (1 vote, 2 comments)
- The option from_file="internal.log" is broken. (1 vote, 1 comment)
- 68: Apparent bug in Intersect intervals, overlapping pieces (5 votes, 1 comment)
- 106979439 108792355 1 problem: SAMTOOLS is broken? (1 vote, 1 comment)
- Bug: Returning Bitset error 536670912 (4 votes, 1 comment)
- EMBOSS: several tools fail with default options (4 votes, 3 comments)
- Tools: Cloudmap reference files not found (4 votes, 2 comments)
- Run: Patch taxonomi (1 vote, 1 comment)

Ideas

- Implement JavaScript build process (1 vote, 6 comments, 0/13)
- Tools: Incorporate key Cuffdiff output files for Cumberbund (1 vote, 1 comment, 0/3)
- Workflow Editor: Provide explicit access to implicit datatype converter tools (1 vote, 1 comment, 0/3)
- Google Drive / Dropbox / Box / ... integration (6 votes, 3 comments)
- 720: Capture and report time taken to run each job (8 votes, 0/2)
- Allow administrators to "trust" certain HTML outputs based on tool producing them. (4 votes, 4 comments)
- Workflows: highlight the noodles in the workflow editor upon hovering (4 votes, 2 comments, CE)
- 5: Option to disable automatic history creation (4 votes, 2 comments, CE)
- Allowing workflow step dependencies when no input/output files exist (4 votes, 1 comment)
- Assistive UI (4 votes, 1 comment, 0/4, CE)
- For sensible output. Add input name to Son_string (4 votes, 0/3)
- RFC: Implement sophisticated user behavior analysis tool (1 vote, 1 comment)

Pull Requests / Patches

- 685: Patch for FASTQ paired-end issue (1 vote, 1 comment)
- Tools: Bowtie Wrapper Pull Requests from Community (2 votes, 6 comments, 1 comment)
- Pull Request #343 - Need to traverse the other_value dict to find dependencies for ParamValueFilter in dynamic_options when the dependencies are scoped in a conditional. Error was noted attempting to run iuc SnpEff 3.4 in a workflow. (0/3)
- Pull Request #336 - Patch to expose the actual dataset id in the LDDA and HDA to_dict calls (in addition to the instance id). (0/3)
- Pull Request #336 - Traverse context for SelectToolParameter need_jate_validation. (0/3)
- Pull Request #334 - Trello Card #1437: Optional Input Datasets Not Compatible with Parallelism Tag (0/3)
- Pull Request #281 - tools/fastq/fastq_pair_end_joiner: added support for recent Illumina headers (0/3)

Project in Planning

- 308: Demystifying the first ever Galaxy login experience - make tools offer test data if empty history? (3 votes, 2 comments)
- Data Manager: Genome Builds / dbkeys: Make adding builds accessible by Data Manager tools (2 votes, 0/3)
- resetting the password deactivates the user (1 vote, 4 comments)
- Tools: Moving to BAM format as primary representation of sequence data (1 vote, 2 comments)
- Libraries: Role selection (1 vote, 1 comment)
- Data Manager: Rsync version (1 vote, 1 comment)
- UI enhancements (1 vote, 0/7)
- BWA aln -n param update (2 votes, 2 comments)
- Show placement in queue / throughput (1 vote, 1 comment)
- Core: Make the jobs admin interface not suck (6 votes, 1 comment)
- Deleting history using the API does not delete/stop jobs (6 votes, 7 comments, CE)
- Tool Shed (and Galaxy?) should have user profiles. (3 votes, 5 comments)

Members

NS, CE, DF, G, etc.

Activity

- Peter Cock on Tool Shed (and Galaxy?) should have user profiles. (2 hours ago)
- I like the CRCID idea from John, might help in reverse for recognising ToolShed repositories as scientific output? (2 hours ago)
- martenson removed dorine francheteau from 623: picard index indicates failure, but it is successful. (2 hours ago)
- martenson joined Tool Shed (and Galaxy?) should have user profiles.. (2 hours ago)
- martenson on Tool Shed (and Galaxy?) should have user profiles. (2 hours ago)
- Ideas don't bring harm. I am merely trying to determine the demand for / priority of this (2 hours ago)
- Björn Grüning on Tool Shed (and Galaxy?) should have user profiles. (2 hours ago)
- Why not? I'm not that social web guy, but it does not harm, or? (2 hours ago)
- martenson on Tool Shed (and Galaxy?) should have user profiles. (2 hours ago)
- Social Logins? Persona, ResearchGate, LinkedIn, Twitter, G+, FB ? (2 hours ago - edited 2 hours ago)
- John Chilton on Tools: Dataset Collections - (2 hours ago)

<http://bit.ly/gxytrello>

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 service web site 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 on this wiki for your convenience.

 usegalaxy.org

Community & Project

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

- [Community](#)

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](#)
- [Galaxy Appliance](#)

 getgalaxy.org

Contribute

- **Users:** [Share](#) your histories, workflows, visualizations, data libraries, and [Galaxy Pages](#), enabling others to use and learn from them



Use Galaxy

[Servers](#) • [Learn](#)
[Main](#) • [Share](#) • [Search](#)

Communicate

[Support](#) • [Biostar](#)
[Events](#) • [Mailing Lists](#)
[News](#)  • [Twitter](#)

Deploy Galaxy

[Get Galaxy](#) • [Cloud](#)
[Admin](#) • [Tool Config](#)
[Tool Shed](#) • [Search](#)



Contribute

[Develop](#) • [Share](#)
[Issues & Requests](#)
[Teach](#) • [Support](#)

Galaxy Project

[Home](#) • [About](#)

Events

News

Galaxy Event Horizon

Events with Galaxy-related content are listed here.

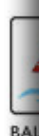


Also see the [Galaxy Events Google Calendar](#) for a listing of events and deadlines. This is also available as an [RSS feed](#).

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

For events prior to this year, see the [Events Archive](#).

Upcoming Events



Date	Topic/Event	Venue
May 6-7	Scaling Galaxy for Big Data	NGS TGAC
May 9	Introduction to Galaxy Workshop	The C (TGA)
May 12	Galaxy Workshop	Unive UK
	Galaxy Project Update	5th E Meet Edinb
May 13	Galaxy Workshop	Instit Medic
May 12-14	Short course on RNA-seq and ChIP-seq	Unive Norw
May 16	Galaxy Initiation	Form Plate Biolog
May 19	Initiation au traitement et à l'analyse des données métaboliques sur la plateforme scientifique web Galaxy IFB-MetaboHUB	8e Jo Lyon

News Items

May 2014 Galaxy News



The [May 2014 Galaxy Update Newsletter](#) is out! There's a lot going on in the project and the community right now. The big news in the past month is the move from the [Galaxy-User mailing list](#) to [Galaxy Biostar](#) for user support. This has been running for a week now, and has been very well received.



The other big news is upcoming events. [Early registration for GCC2013 closes May 23](#). Register now and save more than 70% on registration costs, and [Training Day](#) registration is an additional 55% off if you register for both at the same time. We are also pleased to [announce this year's keynote speaker](#) and the first ever [GCC Hackathon](#).



There's also a [Galaxy UK Tour](#) which is visiting Norwich and Edinburgh in May, and [there are at least 17 other Galaxy related events](#) in the next 70 days in Norway, France, *online*, Croatia, Thailand, Canada, the US, the Netherlands, and Australia



As always, there are [new papers](#) (47 of them, including four we highlighted), [new public Galaxy Servers](#) (Globus Genomics Proteomics and SunLab Galaxy), [new jobs](#) (7 postings in 6 countries), [new tools](#) in the project ToolShed (um, lots), and a [new public ToolShed](#) (at the [Dutch Techcentre for Life Sciences \(DTL\)](#)).

[Dave Clements](#) and the [Galaxy Team](#)

Posted to the [Galaxy News](#) on 2014-04-30

Galaxy Biostar Launched

Galaxy has teamed up with [Biostar](#) to create a [Galaxy User support forum](#) at <https://biostar.usegalaxy.org>!



We want to create a space where researchers using Galaxy can come together and share both scientific advice and practical tool help. Whether on [usegalaxy.org](#), a [CloudMan](#) instance, or any other Galaxy ([public](#) or [local](#)), if you have something to say about *Using Galaxy*, this is the place to do it!

Current integration with [usegalaxy.org](#)

- We imported the **whole history** of the [galaxy-user@bx.psu.edu](#) mailing list into [Galaxy Biostar](#). Your prior posts are automatically claimed when you login!
- If you access [Galaxy Biostar](#) from <http://usegalaxy.org> (Menu: **Help** → **Galaxy Biostar**) you will be automatically logged in. A Galaxy Biostar account will be created for you if it did not previously exist. To obtain this account's password please use the [password reset feature](#) of Galaxy Biostar.
- When you have a question, search Galaxy Biostar directly from any Galaxy tool page.

Read more about how to get started on the [Biostar wiki page](#).



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

galaxypproject.org

Galaxy Australasia Workshop

2
0
1
4

We also support
community
organized efforts
and events.



Galaxy Resources & Community: Videos

The screenshot shows the Vimeo channel for the Galaxy Project. The header includes the Vimeo logo and navigation links: Me, Videos, Create, Watch, Tools, Upload. A search bar is located in the top right. The channel name "Galaxy Project" is displayed with a "PLUS" badge and a note "Joined 1 month ago". Below this, a statistics bar shows: 54 Videos, 0 Likes, 0 Following, 1 Group, 6 Channels, and 0 Albums. A "Recently Uploaded" section features four video thumbnails. Each thumbnail has a title, a subtitle, and a timestamp. The first two videos are "Using Galaxy protocol 3" and "Using Galaxy protocol 2", both titled "CPB Using Galaxy" and uploaded 5 days ago. The next two are "Using Galaxy protocol 1" and "FASTQ Prep - Illumina", both titled "CPB Using Galaxy" and uploaded 1 week ago. A "Settings" button is visible on the left side of the channel page.

Galaxy Project PLUS
Joined 1 month ago

54 Videos | 0 Likes | 0 Following | 1 Group | 6 Channels | 0 Albums

Recently Uploaded + See all 54 videos

- Using Galaxy protocol 3**
Calling Peaks For ChIP-seq Data
CPB Using Galaxy 3
5 days ago
- Using Galaxy protocol 2**
Loading Data and Understanding Datatypes
CPB Using Galaxy 2
5 days ago
- Using Galaxy protocol 1**
Finding Human Coding Exons with Highest SNP Density
CPB Using Galaxy 1
5 days ago
- FASTQ Prep**
Illumina
FASTQ Prep - Illumina
1 week ago

Settings

Galaxy is an open, web-based platform for data intensive biomedical research. Whether on this free public server or your own instance, you can perform, reproduce, and share complete analyses. The Galaxy team is a part of BX at Penn State, and the Biology and Mathematics and Computer Science departments at Emory University. The Galaxy Project is supported in part by NSF, NHGRI, The Huck Institutes of the Life Sciences, The Institute for

“How to”
screencasts on
using and
deploying
Galaxy

Talks from
previous
meetings.

<http://vimeo.com/galaxyproject>

Galaxy Resources & Community: CiteULike Group



[CiteULike](#) [MyCiteULike](#) [Group: Galaxy](#) [Search](#) Logged in as [galaxyproject](#) [Log Out](#)

Group: Galaxy - library 1437 articles

You are an administrative member of this group.
Invite [other CiteULike users](#) to join, or invite [people who don't use CiteULike yet](#).

[Search](#) [Unwatch](#) [Copy](#) [Export](#) [Sort](#) [Hide Details](#)

☐ **✓ Life science data analysis workflow development using the bioextract server leveraging the iPlant collaborative cyberin**
Concurrency Computat.: Pract. Exper. (1 February 2014), pp. n/a-n/a, [doi:10.1002/cpe.3237](#)
by [Carol M. Lushbough](#), [Etienne Z. Gnimpieba](#), [Rion Dooley](#)
posted to [workbench](#) by [galaxyproject](#) to the group [Galaxy](#) keyed Lushbough2014Life on 2014-03-04 19:10:09 ★★/
[Abstract](#) [Copy](#) [My Copy](#)

☐ **✓ Workshops: A Great Way to Enhance and Supplement a Degree**
PLoS Comput Biol, Vol. 10, No. 2. (27 February 2014), e1003497, [doi:10.1371/journal.pcbi.1003497](#)
by [Segun Fatumo](#), [Sayane Shome](#), [Geoff Macintyre](#)
posted to [other](#) by [galaxyproject](#) to the group [Galaxy](#) keyed Fatumo2014Workshops on 2014-03-04 19:08:20 ★★/
[Abstract](#) [Copy](#) [My Copy](#)

☐ **✓ Wrangling Galaxy's Reference Data**
Bioinformatics (28 February 2014), [doi:10.1093/bioinformatics/btu119](#)
by [Daniel Blankenberg](#), [James E. Johnson](#), [James Taylor](#), [Anton Nekrutenko](#)
posted to [project](#) by [galaxyproject](#) to the group [Galaxy](#) keyed Blankenberg2014Wrangling on 2014-03-04 18:55:14 ★★★★★/
[Abstract](#) [Copy](#) [My Copy](#)

☐ **✓ Detection of PIWI and piRNAs in the mitochondria of mammalian cancer cells**
Biochemical and Biophysical Research Communications (March 2014), [doi:10.1016/j.bbrc.2014.02.112](#)
by [ChangHyuk Kwon](#), [Hyosun Tak](#), [Mina Rho](#), [et al.](#)
posted to [methods](#) by [galaxyproject](#) to the group [Galaxy](#) keyed Kwon2014Detection on 2014-03-04 18:53:07 ★★/ [along with 1 person](#)
[Copy](#) [My Copy](#)

☐ **✓ CanSNPer: a hierarchical genotype classifier of clonal pathogens**
Bioinformatics (25 February 2014), [doi:10.1093/bioinformatics/btu113](#)
by [Adrian Lärkeryd](#), [Kerstin Myrtenäs](#), [Edvin Karlsson](#), [et al.](#)
posted to [tools](#) by [galaxyproject](#) to the group [Galaxy](#) keyed Larkeryd2014CanSNPer on 2014-03-04 18:51:21 ★★/
[Abstract](#) [Copy](#) [My Copy](#)

☐ **✓ Web-based Workflow Planning Platform Supporting the Design and Execution of Complex Multiscale Cancer Models**
pp. 1-1, [doi:10.1109/jbhi.2013.2297167](#)

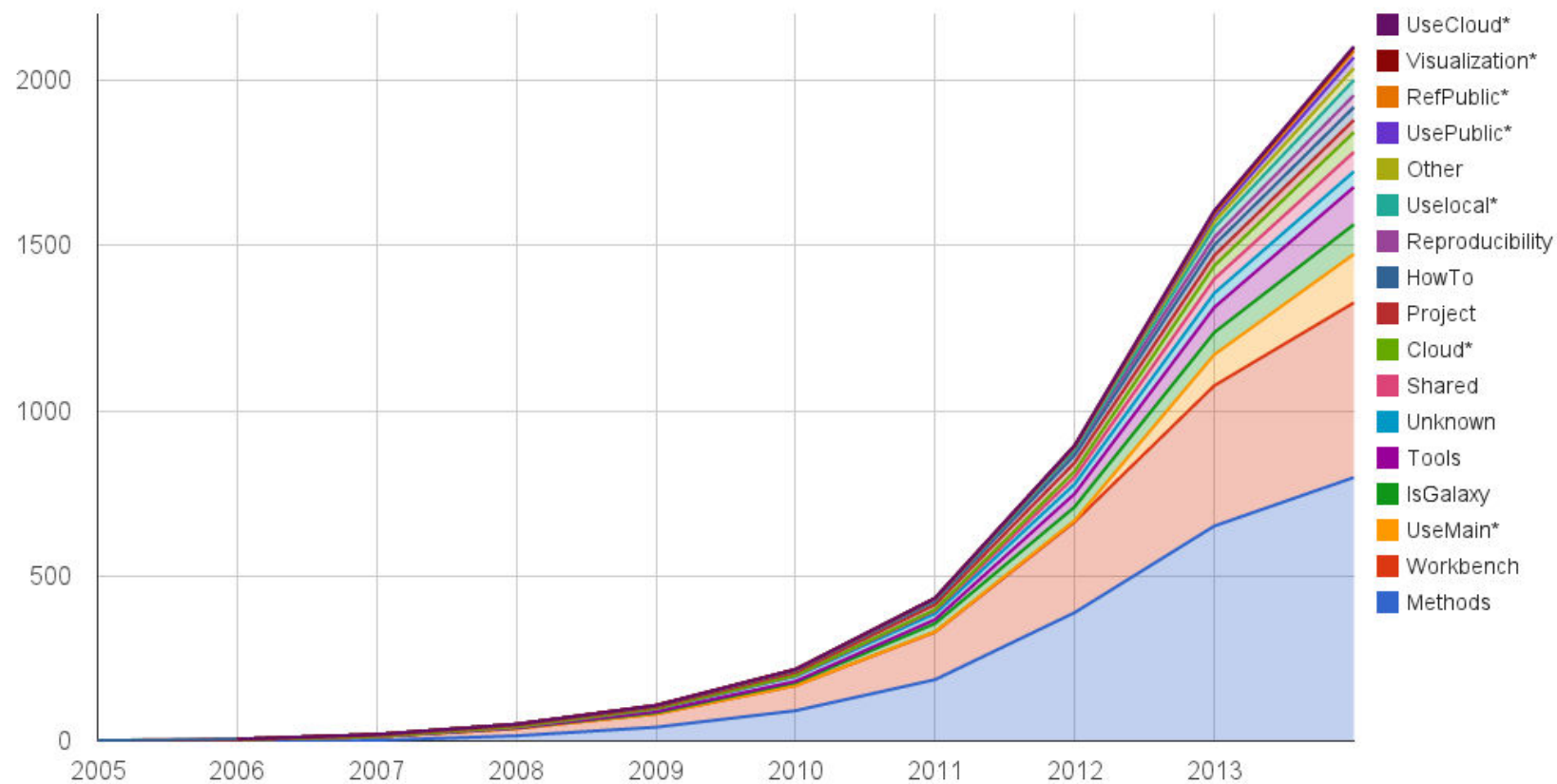
Group Tags
All tags in the group Galaxy
Filter:
[Display as Cloud](#)

methods	697
workbench	466
usemain	108
tools	91
isgalaxy	80
cloud	50
shared	50
unknown	47
uselocal	37
project	32
howto	30
reproducibility	28
other	23
usepublic	19
refpublic	12
visualization	7
usecloud	3

Over
1600
papers

<http://bit.ly/gxycul>

Publications: Cumulative Tags per Year (through July 2014)

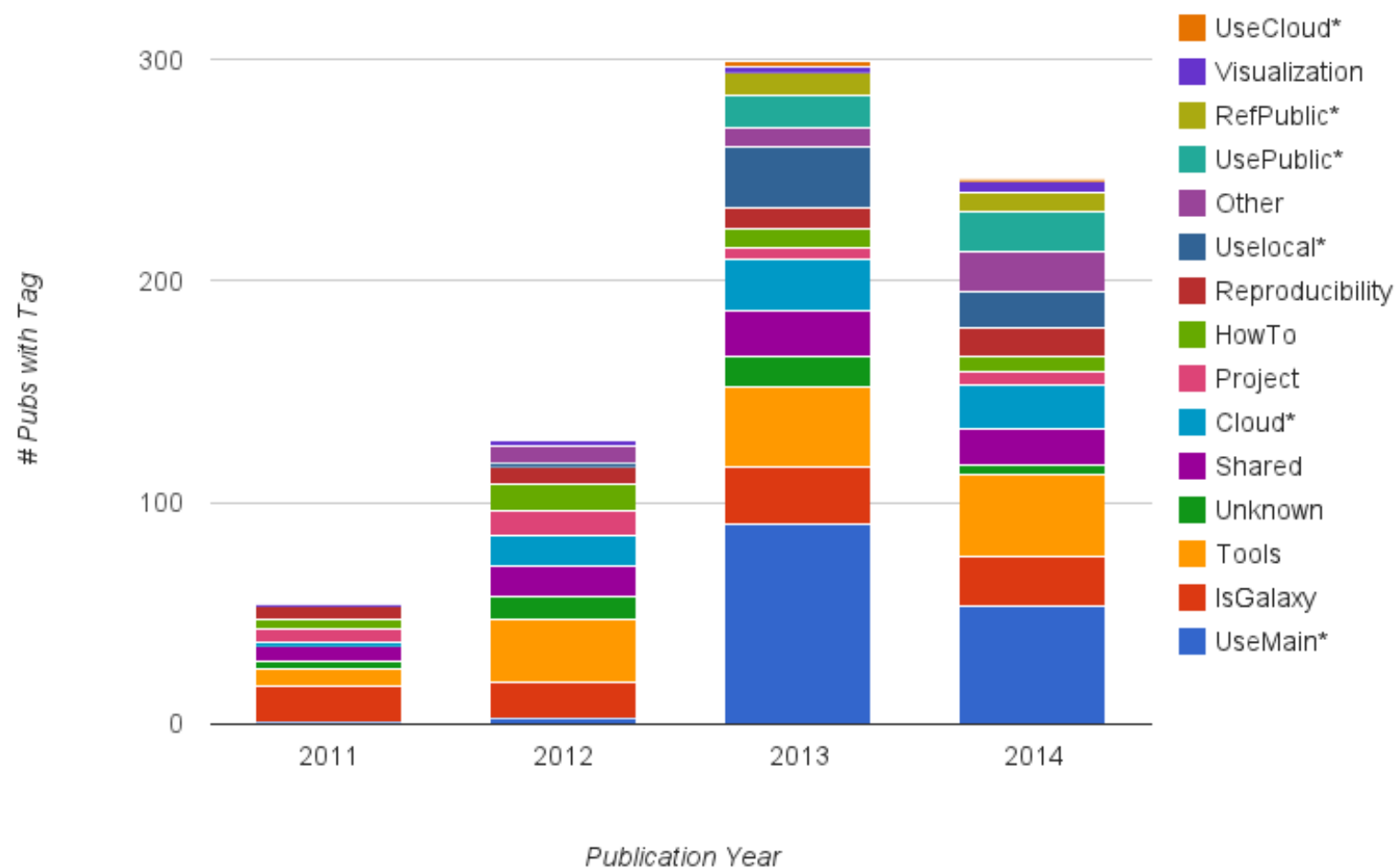


17 tags

Publication Year

Several new tags added in 2013.

Tags excluding Methods and Workbench: 2011-2014 (July 2014)



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



Greg von Kuster



Ross Lazarus



Anton Nekrutenko



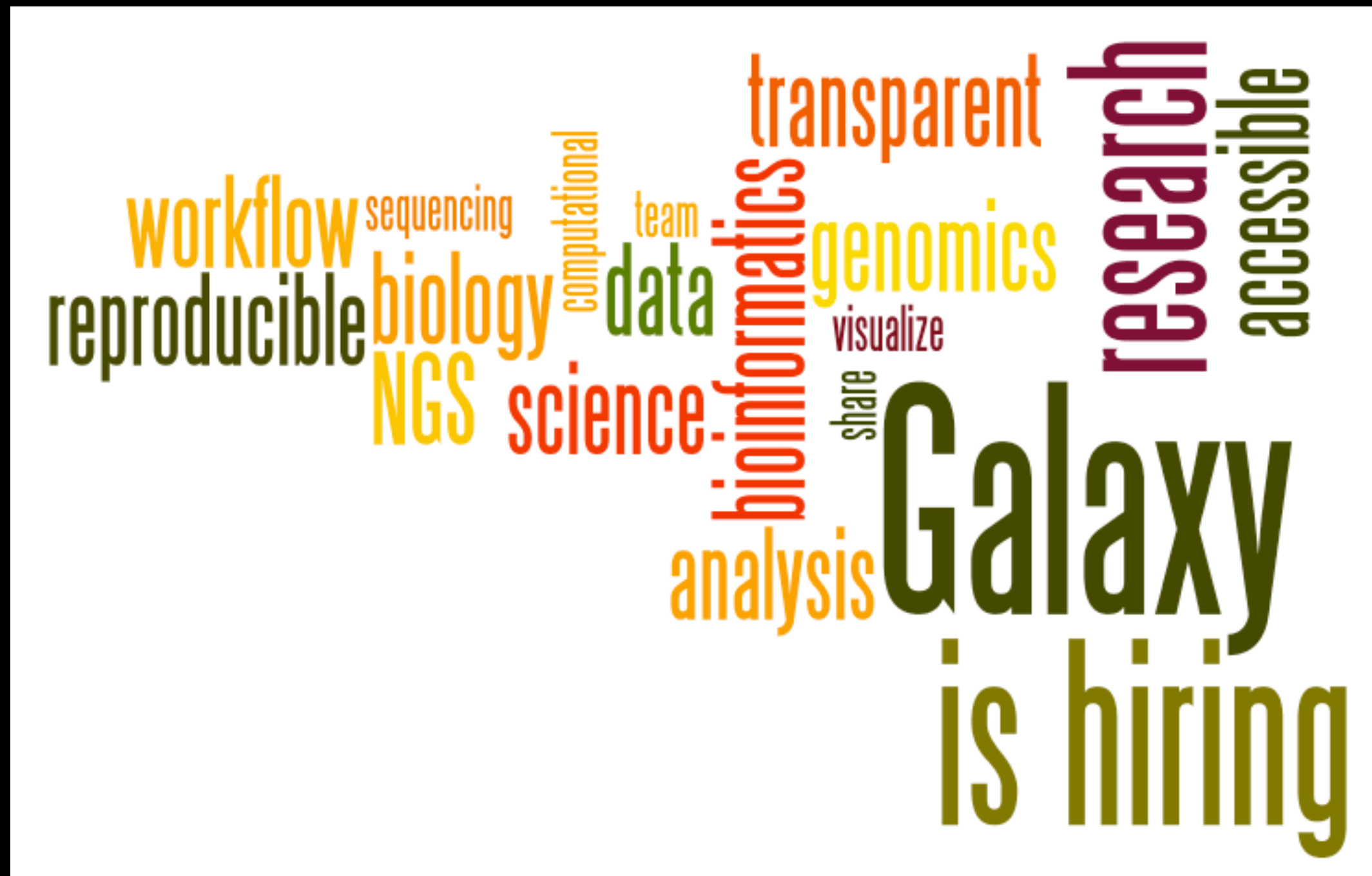
Nick Stoler



James Taylor

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

Galaxy is hiring post-docs and software engineers



Please help.

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

