

Galaxy Community Update

PAG XXVI

January 17, 2018

San Diego, California, United States

Dave Clements

Johns Hopkins University

Galaxy Team / Galaxy Community



#usegalaxy @galaxyproject

bit.ly/gxy-pag2018-upd



This talk complements yesterday's talks

Galaxy: An Open Platform for Data Analysis and Integration

Plant & Animal Genome XXVI (PAG 2018)

January 16, 2018

San Diego, California,
United States



#usegalaxy @galaxyproject

Using Galaxy for Genome Annotation

Jeremy Goecks
Assistant Professor, Oregon Health and Science University



<http://www.galaxyproject.org>

@galaxyproject /
#usegalaxy



Today's talk focuses on what's happened since last year, and what's coming

Some New Features from the past year

Switched to a 4 month release cycle

17.01, 16.05, 17.09

18.01 is in the pipe

Singularity Support (17.09)

HPC-friendly containers

Collection improvements (all releases)

Continue to improve support for large, multi-sample
analysis

Propagating Dataset Tags (17.05)

Tags make it easy to track dataset (and collection)
relationships.

New Features in past year

Drag & Drop Datasets into tool inputs (17.05)

Drag datasets from history panel into tool forms.

Conda package management (17.01)

Became the default for tools and dependencies

Plus hundreds more features large and small

New Releases of lots of Software

BioBlend	a Python library for interacting with CloudMan and Galaxy's API
galaxy-lib	a subset of the Galaxy core code base designed to be used as a library.
Planemo	set of command-line utilities to assist in building tools for the Galaxy project.
StarForge	help build Galaxy things in virtualization
sequence_utils	set of Python modules for reading, analyzing, and converting sequence formats.
Pulsar	Python server application that allows a Galaxy server to run jobs on remote systems (including Windows)

New Releases of lots of Software

ephemeris	Python library and set of scripts for managing the bootstrapping of Galaxy plugins - tools, index data, and workflows.
CloudLaunch	New platform for launching multiple appliances on multiple cloud platforms
nebulizer	Command-line utilities for managing users, data libraries and tools in a Galaxy instance
CloudBridge	provides a simple layer of abstraction over different cloud providers
Galaksio	New Galaxy user interface that emphasizes workflows over tools

<https://github.com/galaxyproject/>

Plans for the near future

usegalaxy.org

Better use of Jetstream (IU & TACC) Bridges (PSC), and
Stampede (TACC)

Interactive Environment support: RStudio, Jupyter

Dataset Collections & Semantic Scalability

More, more, more

Workflows

CWL support

Running time feedback to users

<https://github.com/galaxyproject/galaxy/issues/1928>

Conversion from Wiki to Hub

Latest Tutorials and Trainings

The Galaxy community maintains a wide variety of trainings at <https://training.galaxyproject.org>. Check out the latest [metagenomics tutorial](#) by [@bebatut](#) and [@shiltemann](#).



News

[January 2018 News of the Galaxy!](#) – New events, blog posts, publications, jobs, ...

[All-new CloudLaunch release](#) – Manages your credentials, multiple clouds, multiple apps,...

[GCCBOSC2018 Training Topic Nominations](#) – Choose

Events

[Plant & Animal Genome XXVI \(PAG 2018\)](#) – Galaxy is for Plants and Animals too...

[Packaging your bioinformatics tool with Bioconda and Galaxy](#) – from Melbourne Bioinformatics

[Developing bioinformatics training materials](#) – from

@galaxyproject



Galaxy Project @galaxyproject

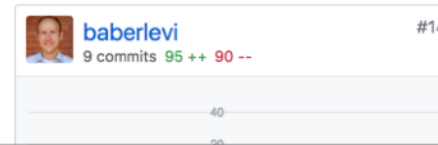
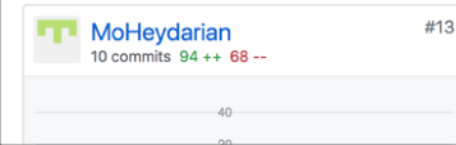
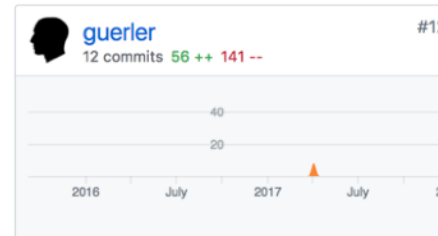
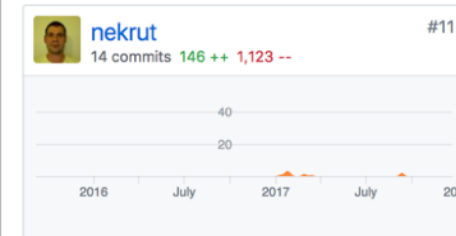
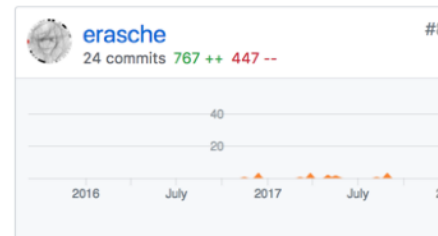
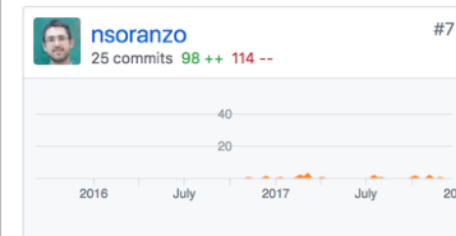


We are delighted to have @thePeerJ as a #GCCBOSC 2018 sponsor! (and they bring goodies - watch the Feb newsletter :-)) peerj.com
galaxyproject.org/events/gccbosc...

OPEN CHAT

<http://galaxyproject.org>

Contributions to master, excluding merge commits



Converted from Wiki to Hub

From MoinMoin based to
GitHub, Markdown, Metalsmith based

First edits became an order of magnitude
(or two) harder
Subsequent small edits only slightly harder

Big edits became an order of magnitude
easier.

github.com/galaxyproject/galaxy-hub

Relaunched Blog in October

The Galactic Blog

Welcome to the *Galactic Blog*, a list of Galaxy related blog posts from community members. This blog contains longer form posts or pointers to longer form posts that are elsewhere on the web.

The Galactic Blog complements the [Galxy News feed](#).

Add your voice!

See the hub [CONTRIBUTING page](#) in general, and the [blog section](#) in particular for directions on how to add a blog entry.



Here's what people are saying

Date	Author(s)	Topic
2018 Jan 08	Ming Chen <i>from Staton Lab</i>	Galaxy R Markdown Tools Using R Markdown as a framework to develop Galaxy tools
2017 Dec 12	Tomas Klingström <i>from the Galactic Blog</i>	Galaksio: a Galaxy user interface focused on running prepared workflows Make your workflows more accessible to non-bioinformaticians
2017 Dec 11	Tomas Klingström <i>from Freiburg Galaxy Team Blog</i>	Bioinformatic training with the B3Africa project Training using the Galaksio interface
2017 Dec 01	Vimalkumar Velayudhan <i>from Vimalkumar Velayudhan's Blog</i>	Setup your own instance of the Galaxy Bioinformatics platform on DigitalOcean Setup a DigitalOcean droplet with Galaxy, Apache and PostgreSQL
2017 Nov 08	B. Grüning <i>from the Galactic Blog</i>	Galaxy Docker 17.09 CVMFS and Singularity support

New Public Galaxy servers

General Purpose / Genomics

ABiMS

Galaxy@GenOuest

GVL Melbourne

Domain

BIPAA (Bioinformatics Platform
for Agroecosystem Arthropods)

EuPathDB

GrAPPA (Graph Algorithms
Pipeline for Pathway Analysis)

Tool Publishing Servers

Antigen Receptor (ARGalaxy)

BF2I-Map

ChimeRScope

Codon Harmonizer

Langille Lab PICRUSt server

USMI Galaxy

12 new servers in past year. Now 99 public servers

Publications: Hit 5,000 pubs in October

Year	# Pubs	Methods	Work-bench	Use-Public	Use-Main	Tools	Ref-Public	Use-Local	Is-Galaxy	Cloud	Repro-ducibility	Oth-er	Shar-ed	Un-known	How-To	Pro-ject	Visu-aliza-tion	Use-Cloud
2005	2		1													1		
2006	4		3												1			
2007	12	2	7							1					2	2		
2008	32	15	12		1				2					2		1	1	
2009	52	26	18			3			2			1	1	4	1			
2010	107	50	36			1			5			1	1	7	3	5		
2011	205	93	70		1	8			16	3	6		8	3	4	6	1	
2012	396	196	128	1	3	29		3	15	13	7	9	12	10	12	10	2	
2013	499	262	150	16	91	37	10	28	27	22	8	9	22	13	7	6	3	2
2014	733	329	224	59	97	67	28	42	47	40	25	40	23	7	11	7	8	1
2015	920	468	232	137	116	67	51	57	48	48	23	34	23	14	8	11	7	2
2016	1104	567	243	206	113	70	109	72	46	35	43	49	20	14	19	7	9	3
2017	929	541	156	207	107	57	75	71	53	19	53	22	16	18	5	5	4	1
2018	5	4		3			1	1										
Total	5000	2553	1280	629	529	339	274	274	261	181	165	165	126	92	73	61	35	9

<https://galaxyproject.org/blog/2017-10-5000-pubs/>

Publications: Some trends

Year	# Pubs	Meth-ods	Work-bench	Use-Public	Use-Main	Tools	Ref-Public	Use-Local	Is-Galaxy	Cloud	Repro-ducibility	Oth-er	Shar-ed	Un-known	How-To	Pro-ject	Visu-aliza-tion	Use-Cloud
2005	2		1													1		
2006	4		3												1			
2007	12	2	7							1					2	2		
2008	32	15	12		1				2					2		1	1	
2009	52	26	18			3			2			1	1	4	1			
2010	107	50	36			1			5			1	1	7	3	5		
2011	205	93	70		1	8			16	3	6		8	3	4	6	1	
2012	396	196	128	1	3	29		3	15	13	7	9	12	10	12	10	2	
2013	499	262	150	16	91	37	10	28	27	22	8	9	22	13	7	6	3	2
2014	733	329	224	59	97	67	28	42	47	40	25	40	23	7	11	7	8	1
2015	920	468	232	137	116	67	51	57	48	48	23	34	23	14	8	11	7	2
2016	1104	567	243	206	113	70	109	72	46	35	43	49	20	14	19	7	9	3
2017	929	541	156	207	107	57	75	71	53	19	53	22	16	18	5	5	4	1
2018	5	4		3			1	1										
Total	5000	2553	1280	629	529	339	274	274	261	181	165	165	126	92	73	61	35	9

Year	# Pubs	Meth-ods	Work-bench	Use-Public	Use-Main	Tools	Re Pu
2005	2		1				
2006	4		3				
2007	12	2	7				
2008	32	15	12		1		
2009	52	26	18			3	
2010	107	50	36			1	
2011	205	93	70		1	8	
2012	396	196	128	1	3	29	
2013	499	262	150	16	91	37	
2014	733	329	224	59	97	67	
2015	920	468	232	137	116	67	
2016	1104	567	243	206	113	70	
2017	929	541	156	207	107	57	
2018	5	4		3			
Total	5000	2553	1280	629	529	339	

Publications: Some trends

In 2013-2014, there are **2.5x** as many **UseMain** pubs as **UsePublic** pubs.

In 2016-2017 there are nearly **2x** as many **UsePublic** pubs as there are **UseMain** pubs.

[Home](#) > [Groups](#) > Galaxy

Galaxy

[Group Library](#)

Recently Added Items

Title	Added By	Date Modified
 Development of a High Resolution Virulence Allelic Profiling...	Galaxy Project	1/7/2018, 12:46:12
 Feed efficiency and the microbiota of the alimentary tract	Galaxy Project	12/29/2017, 20:05:30
 TruXy: Trusted Storage Cloud for Scientific Workflows	Galaxy Project	12/29/2017, 20:00:25
 Integrating Scientific Workflows with Scientific Gateways: A...	Galaxy Project	12/29/2017, 19:59:23
 ENISI SDE: A New Web-Based Tool for Modeling Stochastic Proc...	Galaxy Project	12/29/2017, 19:56:33
 Body Size Regulation via BMP Signaling in Caenorhabditis ele...	Galaxy Project	12/29/2017, 19:55:46
 Omics Informatics: From Scattered Individual Software Tools ...	Galaxy Project	12/29/2017, 19:54:13
 RNA-Seq Analysis Pipeline Based on Oshell Environment	Galaxy Project	12/29/2017, 19:53:37
 Computational prediction and qPCR validation of miRNA in Cer...	Galaxy Project	12/29/2017, 19:52:31
 Improving Integration Effectiveness of ID Mapping based Biol...	Galaxy Project	12/29/2017, 19:51:12

See all 5346 items for this group in the [Group Library](#).

Recent Group Discussion

No recent group discussions.



Publication library for the [Galaxy Project](#). Contains publications that use, reference, implement, and extend Galaxy.

Galaxy is an open, web based platform for data integration and analysis, primarily in the life sciences.

Group [tags are explained here](#).

Want to contribute? [Here's how](#).

<https://galaxyproject.org/>

Owner: [Galaxy Project](#)

Registered: 2017-09-18

Type: Public

Membership: Closed

Library Access: You can only view

[Log in](#) or [Register](#) to join groups

Members (1)



Move from CiteULike to Zotero

In general, a big win.

Galaxy Training Network



[Fork me on GitHub](#) [Help](#) ▾

Welcome to Galaxy Training!

Collection of tutorials developed and maintained by the worldwide Galaxy community.

Galaxy for Scientists

Topic	Tutorials
Introduction to Galaxy	10
Assembly	3
ChIP-Seq data analysis	2
Epigenetics	1
Metagenomics	2
Proteomics	8
Sequence analysis	6

Galaxy for Developers and Admins

Topic	Tutorials
Galaxy Server administration	8
Development in Galaxy	13
Train the trainers	6

<http://galaxyproject.github.io/training-material/>

Galaxy Training Network

Transcriptomics















Training material for all kinds of transcriptomics analysis.

Requirements

Before diving into this topic, we recommend you to have a look at:

- [Galaxy introduction](#)
- [Quality control](#)

Material

Lesson	Slides	Hands-on	Input dataset	Workflows	Galaxy tour
Introduction to Transcriptomics					
De novo transcriptome reconstruction with RNA-Seq					
Reference-based RNA-Seq data analysis					
Differential abundance testing of small RNAs					
Visualization of RNA-Seq results with CummeRbund					

<http://galaxyproject.github.io/training-material/>

Galaxy Training Network

De novo transcriptome reconstruction with RNA-Seq

Overview

❓ Questions

- What genes are differentially expressed between G1E cells and megakaryocytes?
- How can we generate a transcriptome de novo from RNA sequencing data?

🎯 Objectives

- Analysis of RNA sequencing data using a reference genome
- Reconstruction of transcripts without reference transcriptome (de novo)
- Analysis of differentially expressed genes

✅ Requirements

- [Galaxy introduction](#)
- [Quality control](#)

🕒 Time estimation: 6h

Introduction

The data provided here are part of a Galaxy tutorial that analyzes RNA-seq data from a study published by *Wu et al.* in 2014 [DOI:10.1101/gr.164830.113](https://doi.org/10.1101/gr.164830.113). The goal of this study was to investigate “the dynamics of occupancy and the role in gene regulation of the transcription factor Tal1, a critical regulator of hematopoiesis, at multiple stages of hematopoietic differentiation.” To this end, RNA-seq libraries were constructed from multiple mouse cell types including G1E, a TAL1 null immortalized cell line derived from targeted disruption of TAL1 in mouse embryonic stem cells, and megakaryocytes. This

<http://galaxyproject.github.io/training-material/>

Galaxy Admin Training

At this time last year, there had been one week long Galaxy Admin training.

We've now had 3 of these.

Most recent one ended last Friday in Oslo.

Safe to say they are probably here to stay.

Cool upcoming events

ELIXIR Galaxy Community Kickoff and Meeting



The ELIXIR Galaxy community will have a [kickoff meeting](#) 14-16th March 2018, in Freiburg/Germany. We will combine this with a Galaxy User Conference and the **official launch of the usegalaxy.eu server.**

This event will also be coordinated with the [EMBL Australia](#) to also **launch usegalaxy.au at the same time.**



Wednesday (the 14th) will be dedicated to the ELIXIR Galaxy community, to discuss and plan our roadmap in 2018 and 2019. Thursday and Friday (15th, 16th) will be a conference where Galaxy users talk about their research and use-cases. However, a few spots we will reserve for official talks from ELIXIR.

[Register now](#) as space is limited.

2018 Big Genomics Data Skills Training course

The [Jackson Lab \(JAX\)](#) is pleased to open the application period for the 2018 Big Genomics Data Skills Training course, funded by the NIH BD2K initiative: [Big Genomic Skills Training info and online application](#). The program provides training and hands-on experience working with genomic scale data and will enable participants to integrate genomic data analysis into their courses or launch new courses. Topics covered include RNAseq, Whole Exome Sequencing and variant detection, ChIPseq and network analysis. Participants will gain experience using Galaxy, R and Python.



As a grant funded activity housing is included; however you will need to cover your own travel. If you teach at an HBCU or Minority-serving institution (by Dept. of Education criteria) JAX may be able to provide a modest travel award.

The course is for faculty who primarily teach undergraduate students and will be held at JAX Genomic Medicine in Farmington CT, May 21-25.

<https://galaxyproject.org/events>



Just
announced
yesterday

galaxyafrika
.sanbi.ac.za

Cape Town, 3-5 April 2018

Data-intensive biology for the African continent



National
Research
Foundation



GCCBOSC 2018



Galaxy's Community Conference &
OBF's Bioinformatics Open Source Conference
are colocating in 2018

Please join us in Portland, Oregon, United States
June 25-30, 2018

<https://gccbosc2018.sched.com/>

The Galaxy Team



Enis Afgan



Dannon Baker



Dan Blankenberg



Dave Bouvier



Marten Cech



John Chilton



Dave Clements



Nate Coraor



Jeremy Goecks



Sam Guerler



Jen Jackson



Vahid Jalili



Delphine Lariviere



Anton Nekrutenko



Nick Stoler



James Taylor

<https://galaxyproject.org/galaxy-team>

Acknowledgements

You

Scott Cain

The Galaxy Community

NIH NSF

Johns Hopkins University*

Penn State University

Oregon Health Sciences University

Cleveland Clinic*



Dave Clements

Galaxy Project

Johns Hopkins University

clements@galaxyproject.org

Thanks