

# ICBO Workshop #W03

## Integrating Ontological and Genomic Analysis with Galaxy

Organizer: Dave Clements

# Integrating Ontological and Genomic Analysis with Galaxy

---

## ICBO 2016

August 1, 2016  
Corvallis, Oregon

Dave Clements  
Galaxy Team  
Johns Hopkins University  
<http://galaxyproject.org/>



#usegalaxy @galaxyproject

# Agenda

**A quick stroll through the Galaxy**

Demonstrate Galaxy by addressing a specific question

Galaxy Ecosystem

<http://galaxyproject.org>

# What is Galaxy?

Keith Bradnam's definition:

**"A web-based platform that provides a simplified interface to many popular bioinformatics tools."**

From

**"13 Questions You May Have About Galaxy"**


<http://bit.ly/13questions>

<http://galaxyproject.org>

**Galaxy is available several ways ...**

<http://galaxyproject.org>

# As a free for everyone service on the web: [usegalaxy.org](http://usegalaxy.org)



Analyze DataWorkflowShared DataVisualizationAdminHelpUser

Using 384.0 GB

Tools

search tools





[Get Data](#)  
[Send Data](#)  
[Lift-Over](#)  
[Text Manipulation](#)  
[Datamash](#)  
[Convert Formats](#)  
[Filter and Sort](#)  
[Join, Subtract and Group](#)  
[Fetch Alignments/Sequences](#)  
[NGS: QC and manipulation](#)  
[NGS: DeepTools](#)  
[NGS: Mapping](#)  
[NGS: RNA Analysis](#)  
[NGS: SAMtools](#)  
[NGS: BamTools](#)  
[NGS: Picard](#)  
[NGS: VCF Manipulation](#)  
[NGS: Peak Calling](#)  
[NGS: Variant Analysis](#)  
[NGS: RNA Structure](#)  
[NGS: Du Novo](#)  
[NGS: Gemini](#)  
[Operate on Genomic Intervals](#)  
[Statistics](#)  
[Graph/Display Data](#)  
[CloudMap](#)  
[Phenotype Association](#)

Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy [start here](#) or consult our [help resources](#). You can install your own Galaxy by following the [tutorial](#) and choose from thousands of tools from the [Tool Shed](#).

## Running Your Own


### Understanding how Galaxy works

An in-depth tutorial



## Tweets


by @galaxyproject



Hans-Rudolf Hotz @hrhotz

check out this useful resource for #usegalaxy training and teaching material from @bjoerngruening and his team: [wiki.galaxyproject.org/Teach/Resource...](http://wiki.galaxyproject.org/Teach/Resource...)

111



Galaxy Project @galaxyproject

GCC2016 student/postdoc #underrepresentedgroups scholarships. Apply by May 1 [bit.ly/gcc2016scholar...](http://bit.ly/gcc2016scholar...) #usegalaxy

Embed

View on Twitter





Explore the  
Galaxy with  
**RNA-Rocket**

PATHOGENPORTAL  
THE BIOINFORMATICS RESOURCE CENTERS PORTAL

Galaxy / Metabiome Portal



The Microbiome Analysis Center  
Life on a Smaller Scale

Welcome to the Metabiome Portal @ GMU

We have developed the MMC Metabiome Portal, a flexible and extensible web browser, with the ability to simplify, control, integrate, compare, and analyze all microbiome and metagenomic data. The portal is a unified database management system and data-based analytical framework and includes several tools such as: taxonomic clustering,...

香港中文大學 - 華大基因跨組學創新研究院  
CUHK-BGI Innovation Institute of Trans-Omics

**BGI**

(GIGA)<sup>n</sup> Galaxy  
by CBIIT

Integrated publishing of workflows from GIGA<sup>n</sup> SCIENCE

**Cistrome**



A Galaxy Server  
dedicated to  
ChIP-\* analysis




**Public Galaxy Servers**  
and *still* counting



The Genomic  
HyperBrowser

**Powered by Galaxy**

SCDE  
STEM CELL DISCOVERY ENGINE



**Experiments  
Connected**



Whale Shark Galaxy! 

**South Green**  
bioinformatics platform

**Genomic analysis tools  
for southern and  
Mediterranean plants**

[bit.ly/gxyServers](http://bit.ly/gxyServers)

**Galaxy is available as Open Source Software**

**Galaxy is installed in locations around the world.**

**<http://getgalaxy.org>**



# Galaxy is available on the Cloud



We are using this today

<http://aws.amazon.com/education>

<http://globus.org/>

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

# Agenda

A quick stroll through the Galaxy

**Demonstrate Galaxy by addressing a specific question**

Galaxy Ecosystem Update

<http://galaxyproject.org>

# Quick Poll: Are you ...

1. A bioinformatics novice

2. A bioinformatics apprentice

3. A bioinformatics guru

Yes, those are your only choices.

<http://galaxyproject.org>

# Tutorial

[http://bit.ly/ICBO\\_GXY\\_PDF](http://bit.ly/ICBO_GXY_PDF)

[cloud1.galaxyproject.org](http://cloud1.galaxyproject.org)

[cloud2.galaxyproject.org](http://cloud2.galaxyproject.org)

[cloud3.galaxyproject.org](http://cloud3.galaxyproject.org)

<http://galaxyproject.org>

# Demonstrate Galaxy by addressing a specific question

Why?

Goal is to demonstrate how to use Galaxy in concert with ontological tools.

*Warning: No meaningful science will happen in this workshop.*

[http://bit.ly/ICBO\\_GXY\\_PDF](http://bit.ly/ICBO_GXY_PDF)



## **Specific Question: Repeats in coding regions**

Which genes physically overlap with large numbers of repeats?

Once we have that we'll look for GO term enrichment and then use Galaxy to analyze the reported enrichment

[http://bit.ly/ICBO\\_GXY\\_PDF](http://bit.ly/ICBO_GXY_PDF)

# Repeats in Coding Regions: A General Plan

- Get some data
  - Get Data → UCSC Table Browser
- Identify which genes have overlapping repeats
- Count repeats per gene
- Look for enrichment

[http://bit.ly/ICBO\\_GXY\\_PDF](http://bit.ly/ICBO_GXY_PDF)



# Agenda

A quick stroll through the Galaxy

Demonstrate Galaxy by addressing a specific question

**Other Features and Galaxy Ecosystem Update**

<http://galaxyproject.org>

# 2016 Galaxy Community Conference (GCC2016)

June 25-29, 2016

Bloomington, Indiana

[galaxyproject.org/GCC2016](http://galaxyproject.org/GCC2016)

Slides & posters are now  
online. Video will be shortly



Join us in beautiful

*Bloomington, Indiana*

for the 2016 Galaxy  
Community Conference  
and pre-conference activities!

June 25-29, 2016



Considered one of the five  
prettiest campuses in the US,  
Indiana University is one of  
the major public research  
universities in the nation, and  
home to the National Center  
for Genome Analysis Support.



[galaxyproject.org/gcc2016](http://galaxyproject.org/gcc2016)





Le Corum  
Conference centre

[gcc2017.sciencesconf.org](http://gcc2017.sciencesconf.org)

November 7-11



Salt Lake City, Utah



#GAMe2017  
@EMBL\_ABR

# GAMe 2017

## Galaxy Australasia Meeting

3 - 9 February, Melbourne, Australia

[www.embl-abr.org.au/GAMe2017](http://www.embl-abr.org.au/GAMe2017)

# 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

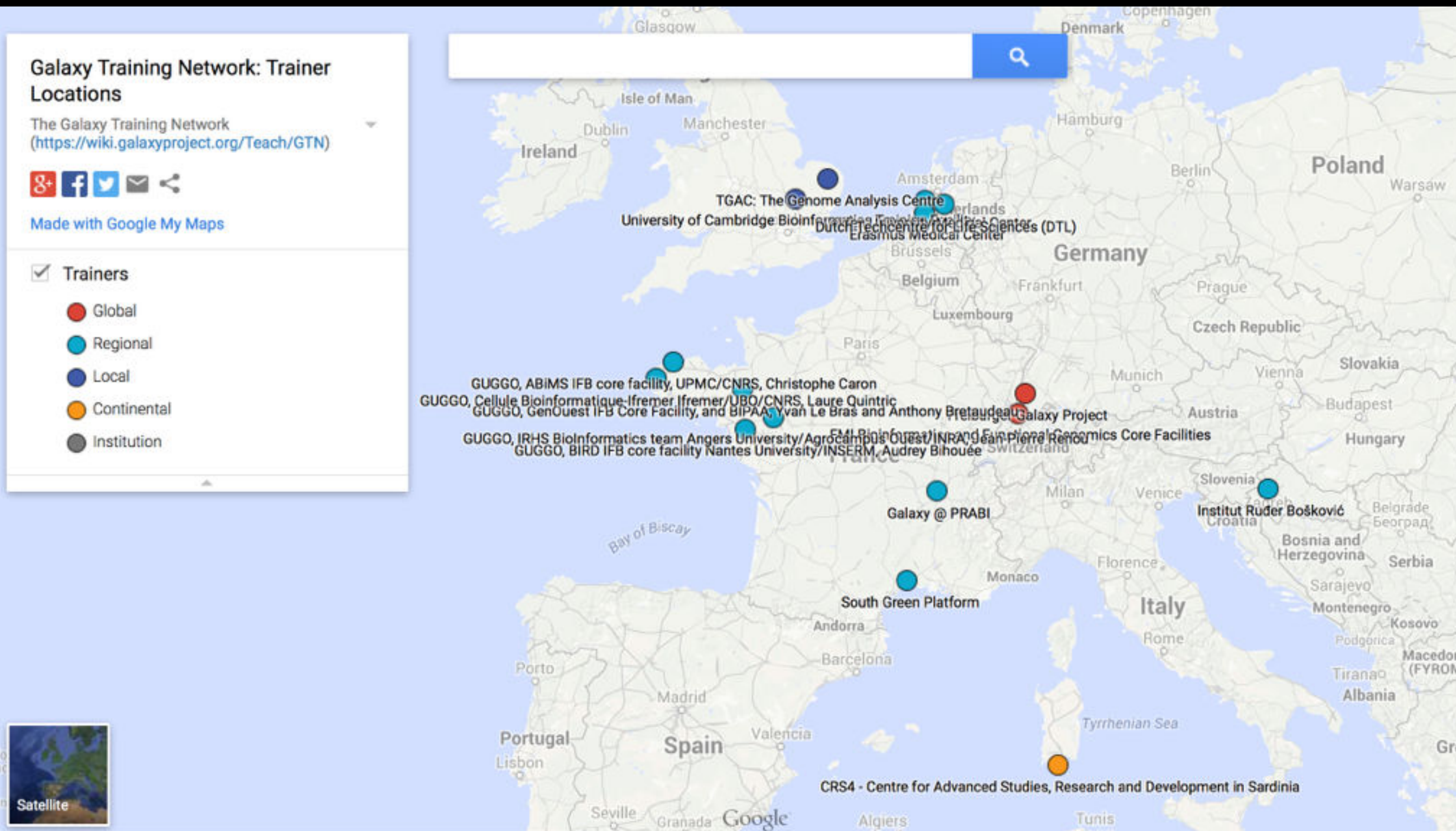
Moved the **user support list** to **Galaxy Biostar**, an online **forum**, that uses the Biostar platform



<https://biostar.usegalaxy.org/>



# Scaling Training



Galaxy Training Network  
[bit.ly/gxygtn](https://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>

# Acknowledgements

You

ICBO 2016

Pankaj Jaiswal

Chris Mungall

Seth Carbon

Nathan Dunn

NIH

Johns Hopkins University

Penn State University





Thanks