An Introduction to Using Galaxy for Genetic Data Analysis

TAGC 2016
July 16, 2016

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

Other Features and Galaxy Ecosystem Update

http://galaxyproject.org

What is Galaxy?

Keith Bradnam's definition:

"A web-based platform that provides a simplified interface to many popular bioinformormatics 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

Galaxy | Analyze Data | Workflow | Shared Data | Visualization | Admin | Help | User | Using 384.0 GB

We have updated Galaxy to the forthcoming 16.04 release and made some configuration and deployment changes, so you may encounter a few problems. We are working on these, and please feel free to report any errors you encounter.

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

Galaxy Project Retweeted

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…
11h

Galaxy Project @galaxyproject
GCC2016 student/postdoc #underrepresentedgroups scholarships. Apply by May 1 bit.ly/gcc2016scholar… #usegalaxy

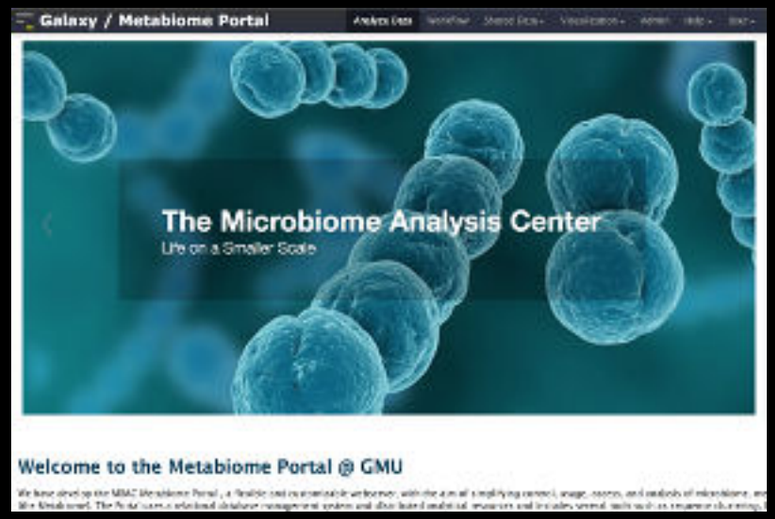
Embed View on Twitter

PENNSTATE 1855 | JOHNS HOPKINS UNIVERSITY | TACC | CYVERSE

Explore the Galaxy with

RNA-Rocket

PATHOGENPORTAL
THE BIOINFORMATICS RESOURCE CENTERS PORTAL



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

(GIGA)^n Galaxy
by CBIIT

Integrated publishing of workflows from (GIGA)^n SCIENCE

Cistrome

A Galaxy Server dedicated to ChIP-* analysis

080+

Public Galaxy Servers
and *still* counting

The Genomic HyperBrowser

Powered by Galaxy

SCDE
STEM CELL DISCOVERY ENGINE

Experiments Connected

South Green
bioinformatics platform

Genomic analysis tools for southern and Mediterranean plants

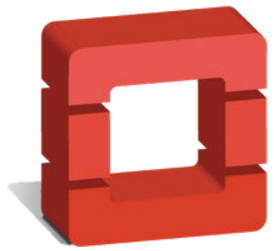
bit.ly/gxyServers

Galaxy is available as Open Source Software

Galaxy is installed in locations around the world.

http://getgalaxy.org

Galaxy is available on the Cloud



openstack™
CLOUD SOFTWARE

amazon
web services™

globus
genomics

OpenNebula.org
The Open Source Toolkit for Cloud Computing

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

Other Features and Galaxy Ecosystem Update

http://galaxyproject.org

Tutorial

http://bit.ly/TGAC_GXY_PDF

http://galaxyproject.org

Pronephros and Kidney

Particular phenotype we are looking into are dilated renal tubules.

(Why?)

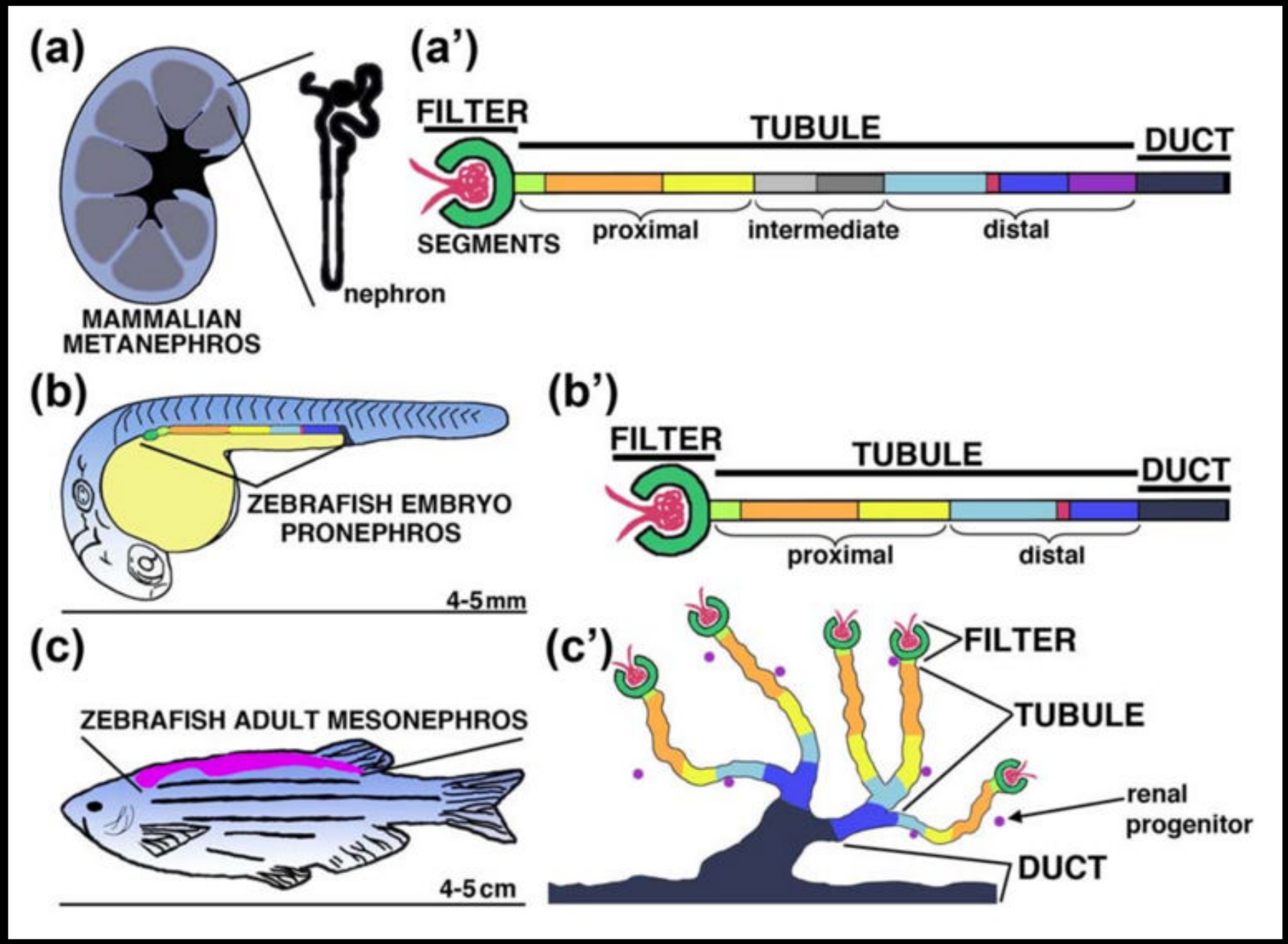


Fig. 1. Kroeger, P. T. and Wingert, R. A. (2014), Using zebrafish to study podocyte genesis during kidney development and regeneration. Genesis, 52: 771–792. doi: 10.1002/dvg.22798

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

Slides & posters are now online. Video will be shortly

26 - 30 June France

GCC 2017 Montpellier

Le Corum
Conference centre

gcc2017.sciencesconf.org

November 7-11

2016

GALAXY

ADMIN TRAINING

SLC

Salt Lake City, Utah

#GAMe2017
@EMBL_ABR

GAMe 2017

Galaxy Australasia Meeting

3 - 9 February, Melbourne, Australia

www.embl-abr.org.au/GAMe2017

Galaxy Community Resources: Galaxy Biostar

Tens of thousands of users leads to a lot of questions.

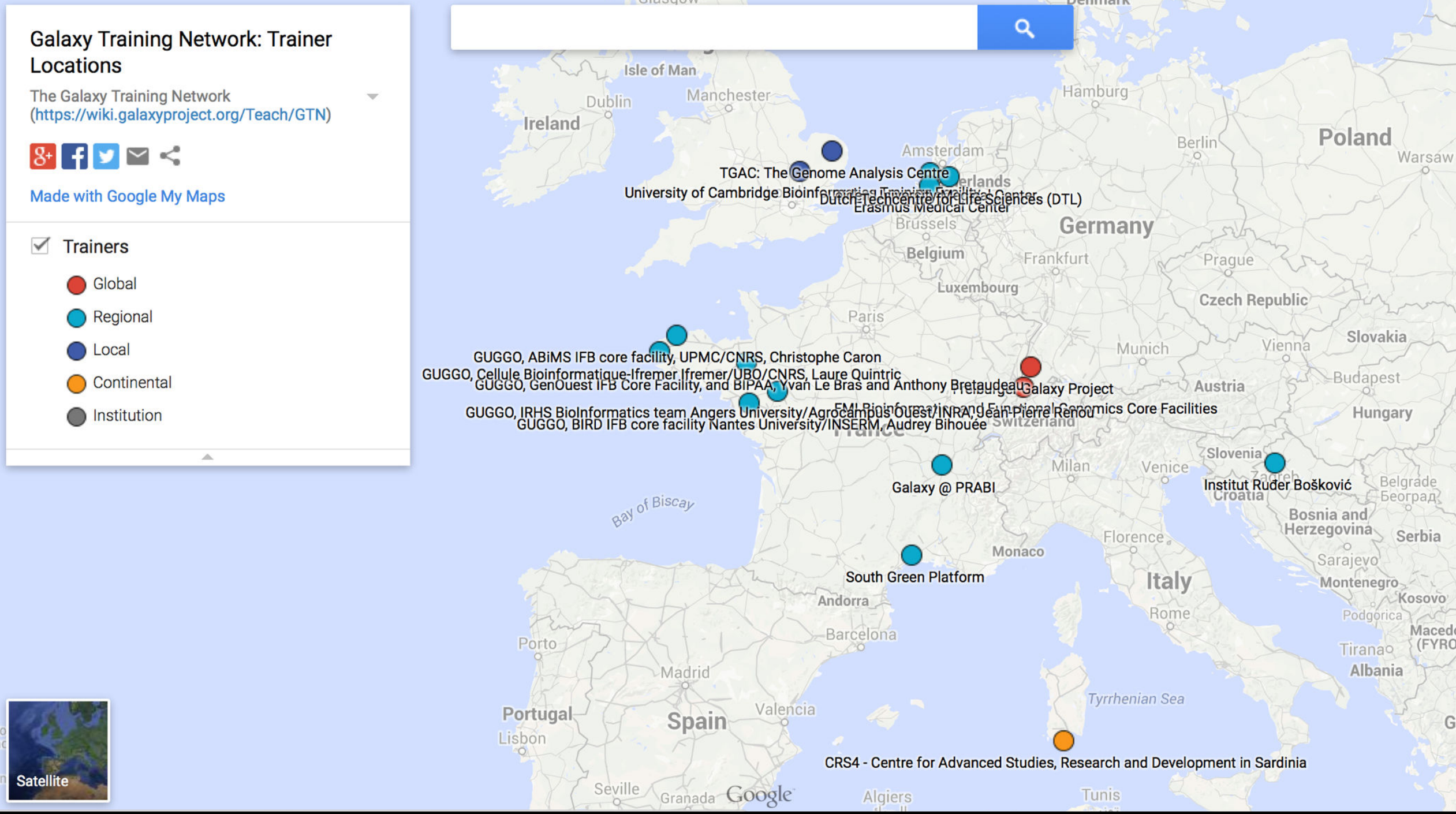
Absolutely have to encourage community support.

Project traditionally used mailing list

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

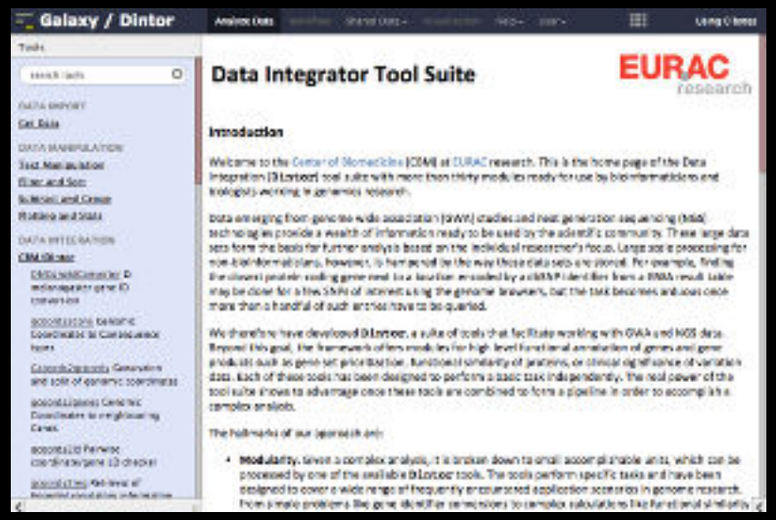
https://biostar.usegalaxy.org/

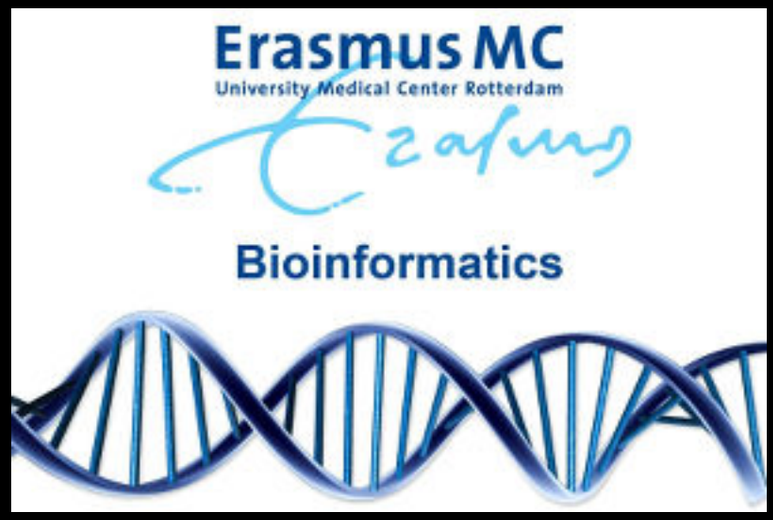
Scaling Training

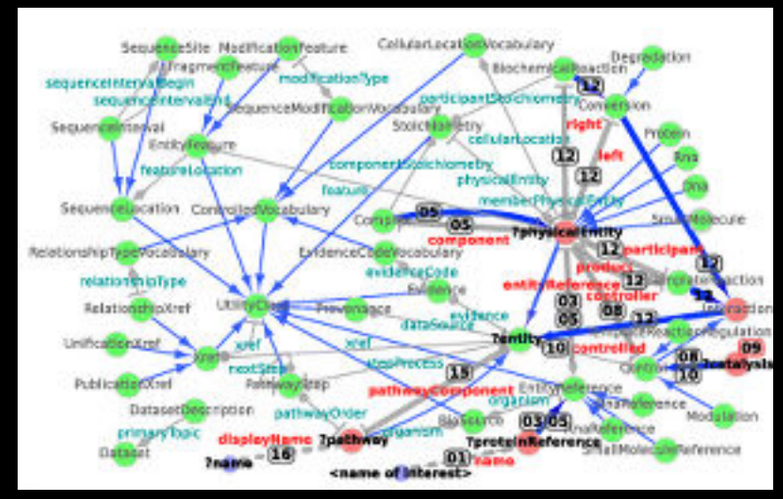


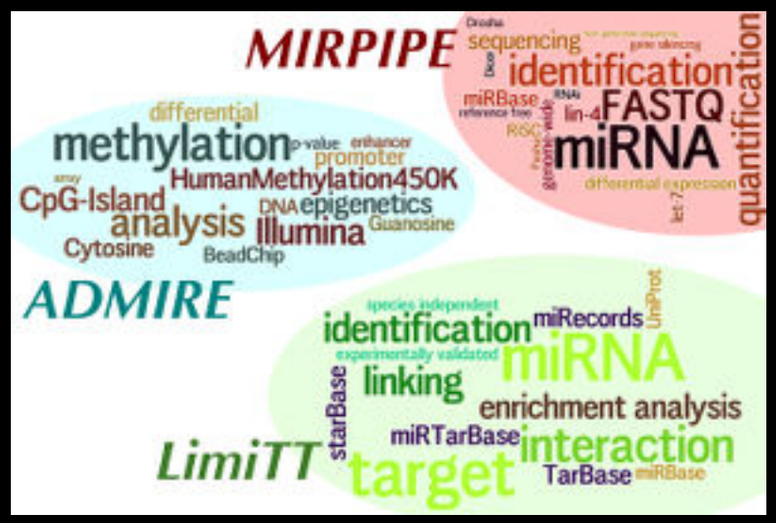
Galaxy Training Network

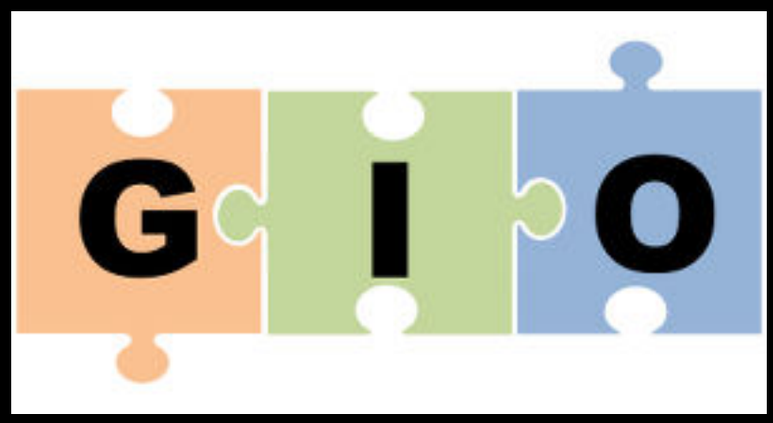
bit.ly/gxygtn

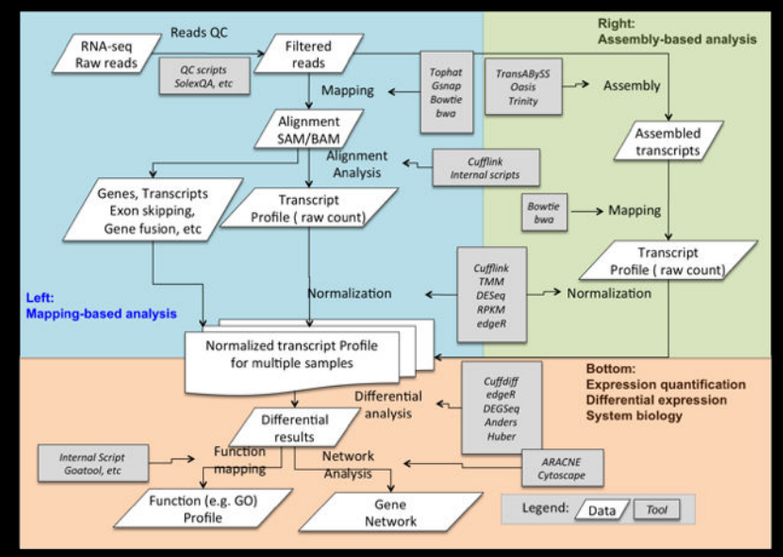




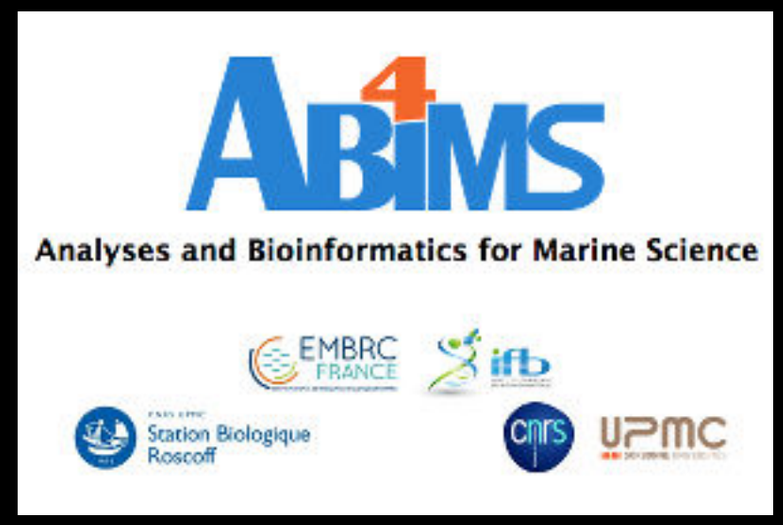








Proteomics
Metabolomics
Natural Language
Image Analysis
Climate Change
Social Science
Cosmology



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

TAGC 2016 and GSA

NIH
Johns Hopkins University
Penn State University

Thanks