

## Accessible and Reproducible Data Analysis for Bench Scientists

Meetups

Galaxy is an open-source, web-based, x. Galaxy enables bench scientists to create, share, and publish sophisticated, reproducible bioinformatic analyses without requiring researchers to learn command line interfaces, or Unix system management skills. Galaxy can be accessed through the project's public server, or on one of the over

60 publicly accessible Galaxy servers. Galaxy can also be installed locally, and on cloud infrastructures.

This talk will introduce the Galaxy platform and discuss the project's recent work and plans going forward. Time allowing, there will also be a brief demonstration.

Galaxy is ... Galaxy is ... Galaxy is ... Galaxy is a ... web-based, data platform for life science research d analysis scientists to ceres, and publish sepristic teo, reproducible high formationaly of the science system researchers to learn command line interfaces, or Unix system management skills. False can be accessed through the project's public server, or on one of the over 60 publicly accessible Galaxy servers. Galaxy can also be installed locally, and of his taskuwis introduce the

This talk will **Grand Axy platform** discuss the project's recent work and plans going forward. there will also be a brief demonstration.

platform for life science research. scientists to create reproducible Galaxy enables beinch scientists management skills. Galaxy can be accessed through the projecto creater on bioinformatic accessible Galaxy servers. Galaxy can also be installed locally, and on cloud infrastanalyses ...

#### **Basic Analysis**

# Which exons have most overlapping repeats in 3 spine stickelback, chromosome XXI?

(~ http://usegalaxy.org/galaxy101 )

#### Galaxy is analysis platform for life science research. Galaxy enables bench scientists to create,

reproducible

researchers to learn command line interfaces, or Unix system management skills. Galaxy can be accessed through the project's public serve processible Galaxy servers. Galaxy can also be installed locally, and on cloud infrastructures.

•••

platform for life science research. Galaxy enables bench scientists to create,

bioinformatic analyses

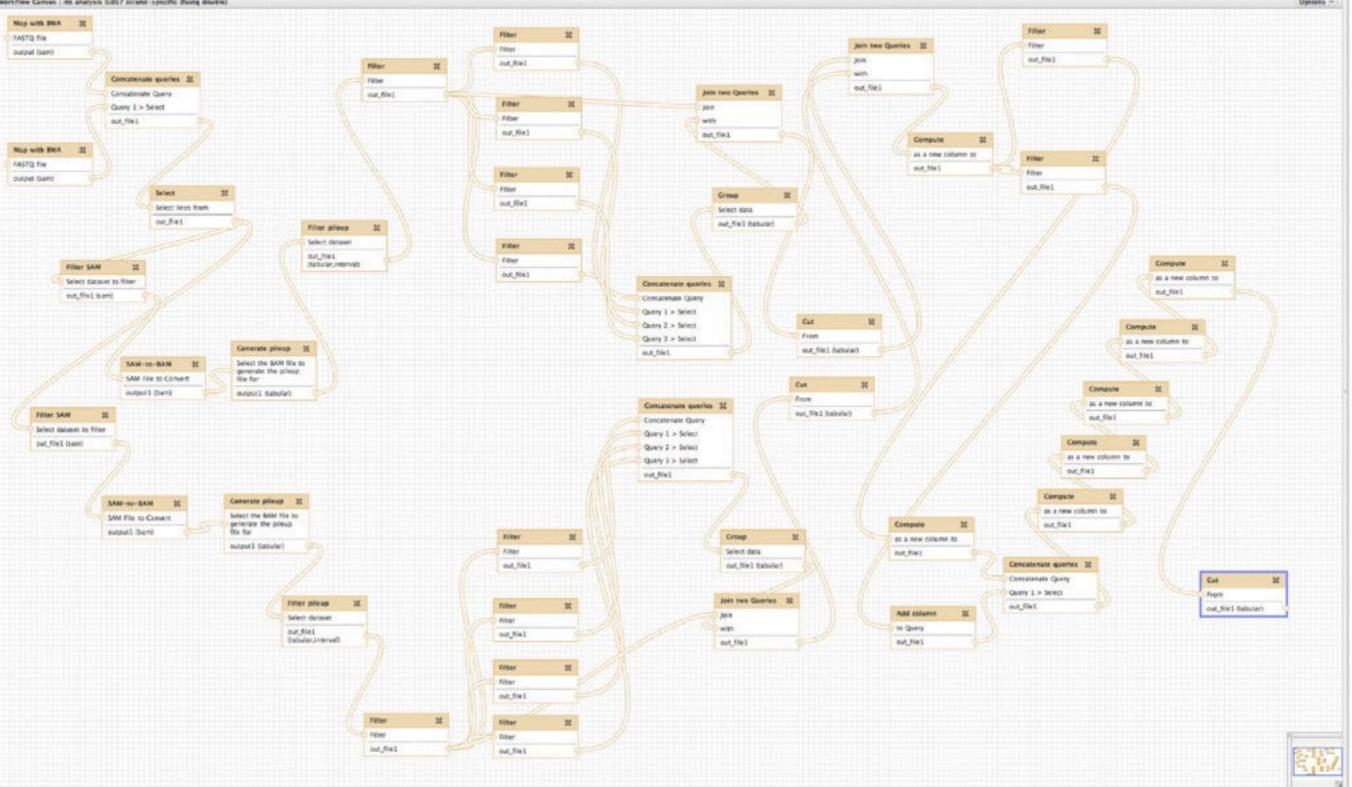
command line interfaces, or Unix system management skills. Galaxy can be accessed through the project's public server, or on one of the **Share**blian **Capture blis provers**. Galaxy can also be installed locally, and on cloud infrastructures.

•••

platform for life science research. Galaxy enables bench scientists to create, share, and publish bioinformatic analyses command line interfaces, or Unix system management skills. Galaxy can be accessed through the project's public server, or on one of the over so phils to galte daxy servers. Galaxy can also be installed locally, and on cloud

infrastructures.

Workflow Canvas | mt analysis 0.017 strand-specific dastg double)



Dynamics of mitochondrial heteroplasmy in three families investigated via a repeatable re-sequencing study, Goto et al. Genome Biology 2011, 12:R59 http://genomebiology.com/2011/12/6/R59

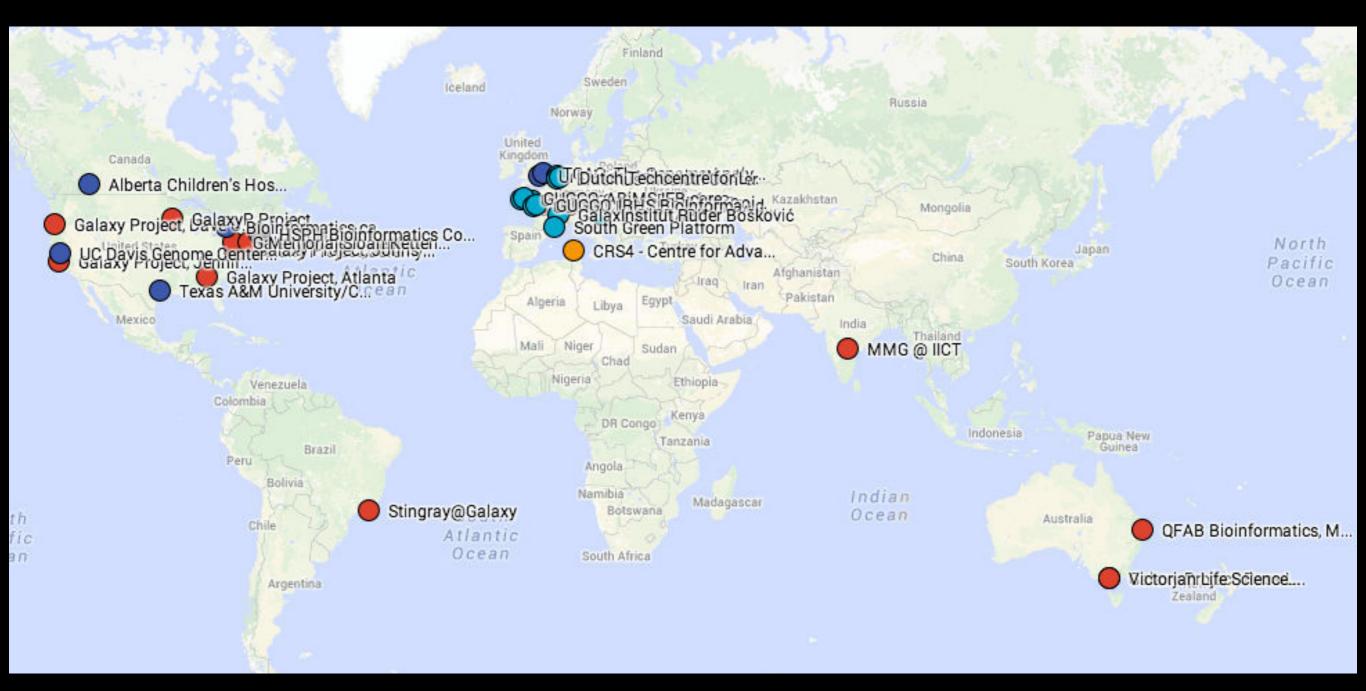
Options

•••

platform for life science research. Galaxy enables bench scientists to create, share, and publish sophisticated, reproducible bioiwithout requiring researchers command line interfaces, or Unix system management skills. Galaxy ctoelearn toommands lineserver, or on one of the over 60 publicly accessible Galaxy servers. Galaxinterfaces, or Unix system infrastructures.

**management skills.** This talk will introduce the Galaxy platform and discuss the project's recent work and plans going forward. there will also be a brief demonstration.

#### **Galaxy Training Network**





•••

platform for life science research. Galaxy enables bench scientists to create, share, and publish sophisticated, reproducible bioinform Gialaxy can be accessed command line interfaces, or Unix system management skills. Galathrough the project's public. Galaxy can also be instage of the project of a public of the public of the project of a public of the project of the public of the project of the public of t



•••

platform for life science research. Galaxy enables bench scientists to create, share, and publish sophisticated, reproducible bioinformatic and OPEN-SOURCE ... command line interfaces, or Unix system management skills. Galax can be accessed through the project's public rever one Galaxy Capitol accessible Calaxy servers. can also be installed Include Interfaces.

#### getgalaxy.org

•••

platform for life science research. Galaxy enables bench scientists to create, share, and publish sophisticated, reproducible bioinformatic analyses command fine interfaces, or Unix system management skills. Galaxy can be installed locally accessible Galaxy servers.

#### bit.ly/gxyServers



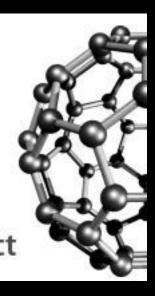


A Galaxy Server dedicated to ChIP-\* analysis

Galaxy / CNIC.DarwinTree	Anderbes wither Scottale Ange Lief	CARGO TANA
n • •	Petros & Officiants and Reference of the Constant and Reference of the Constant and Statistical Systems (Statistical Systems) (Stati	Make 197 _ 1
the Gran Incolor and size		
al non Man Units of Harrin's core reasons of an generate a reference para 1 et al CE constantes	Ton and A low and	
International Accession of the service of the servi		
static all many long some alle of pro-	Invinution PI	S
te and the second	Interitate Kan	
ensite consistent data, why did for the meaning in manual or of made rows	Investitiete Konstantes Steples and	Alastan
lentara anak Barri ne aliana-o-manta ketiki ne arne derarigisi zaini nei rozzi danen mani-pinen		
Index and enclosed land to share a dense of database and to see to the set of one to sadded definition they from to hence to be added to be from the set of the database of the set of the database of database of	Manure and and and	
ner senamar Ganni krani se pro nyrmnia. pisnina tarraleva staro	Preserve autoricet Strands Assesse	
ner o las relativadors tanto en generala- neral tercadoridad de Tre estancia en arra de menantes fend el castellad	ting diving the mercent	sectoral of Malegies) species
onaire UNI himpation and site are noted	tanitation and the standar	and Budyon but or p

#### ballaxy

Powered by the Biochemical Algorithms Library Project





•••

platform for life science research. Galaxy enables bench scientists to create, share, and publish sophisticated, reproducible bioinformatic analyses command line interfaces, or Unix system management skills. Galaanoeonselouchin frastructureson

one of the over 60 publicly accessible Galaxy servers. Galaxy can also be installed locally,

wiki.galaxyproject.org/Cloud









The Open Source Toolkit for Cloud Computing

http://aws.amazon.com/education

http://globus.org/

http://wiki.galaxyproject.org/Cloud



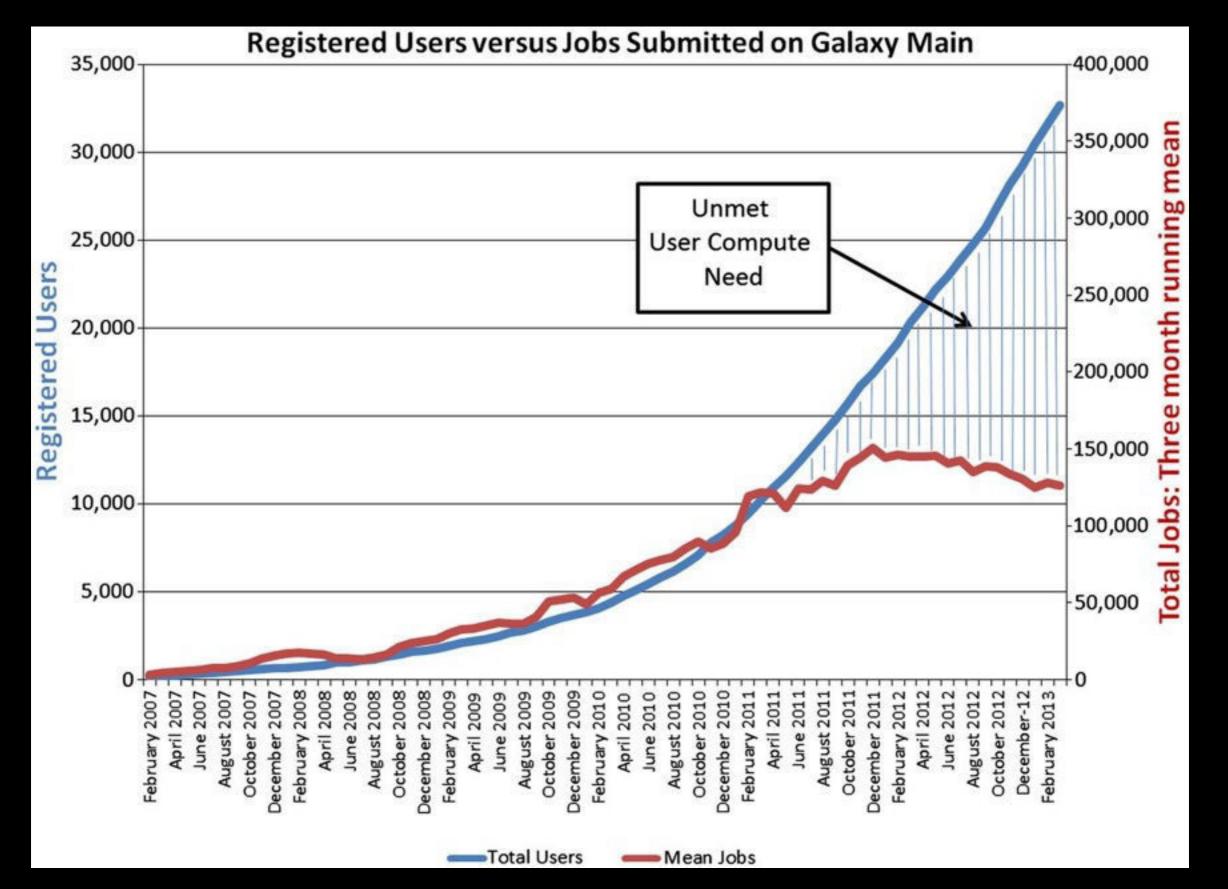
## Scalability ...



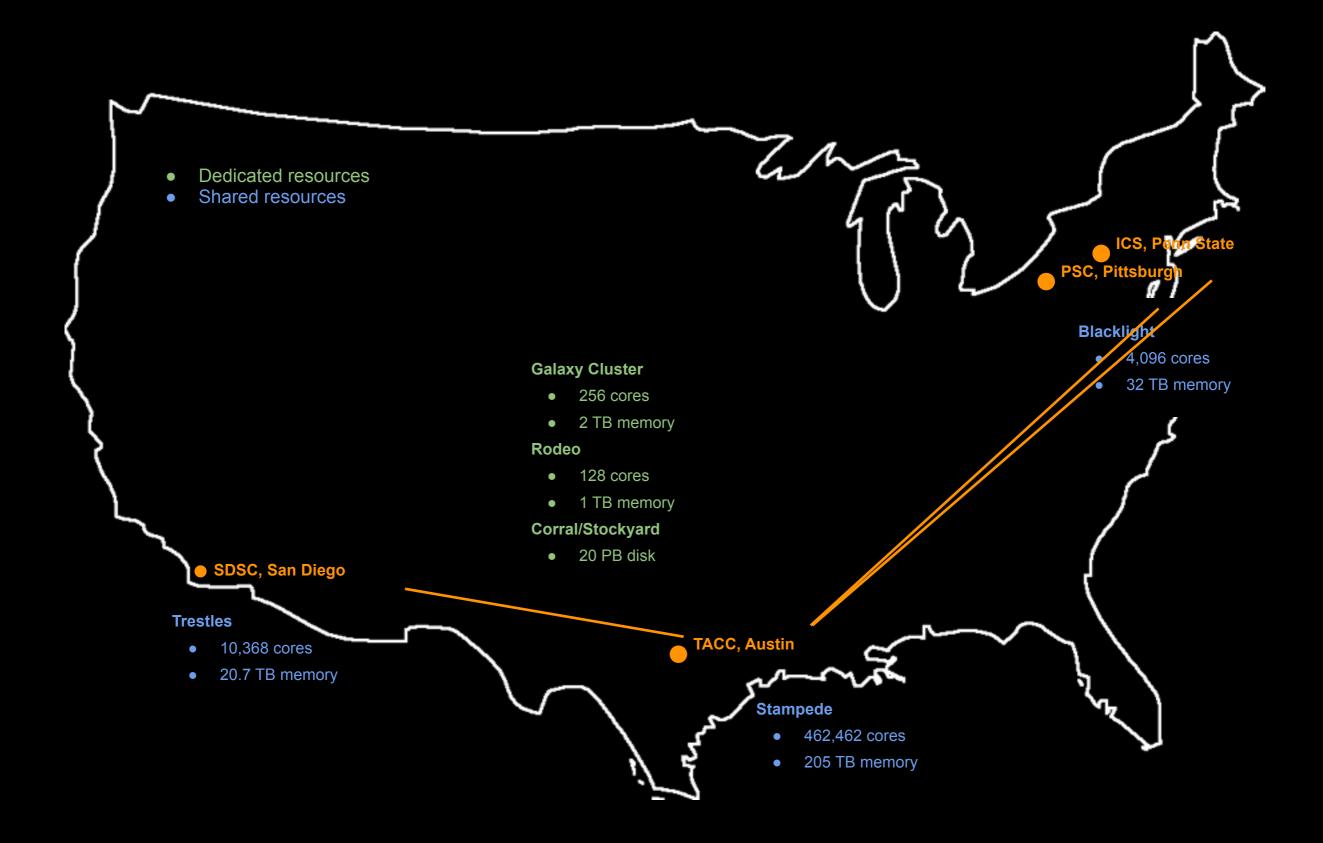
#### Scalability

Data generation is cheap and will stay cheap. Scale & complexity of analysis will continue to grow. More researchers are running bioinformatics analyses of all scales and complexities.

Galaxy needs to scale to the next few orders of magnitude.



Leveraging the national cyberinfrastructure for biomedical research LeDuc, et al. J Am Med Inform Assoc doi:10.1136/amiajnl-2013-002059



#### The Galaxy Team



Enis Afgan

Dannon Baker Dan

Dan Blankenberg

**Dave Bouvier** 

Marten Cech

John Chilton



**Dave Clements** 

Nate Coraor

or

Carl Eberhard

Jeremy Goecks

Sam Guerler



Jen Jackson

**Ross Lazarus** 

Anton Nekrutenko

Nick Stoler

**James** Taylor

Nitesh Turaga

http://wiki.galaxyproject.org/GalaxyTeam

#### Scaling the Project: Community Gatherings

## Galaxy Community Conference

6-8th July 2015

The Sainsbury Laboratory Norwich, UK

## galaxyproject.org

1

1