

# The Galaxy Visual Analysis Framework

Jeremy Goecks, The Galaxy Team, Anton  
Nekrutenko, and James Taylor

# Vision

Galaxy is an **open, Web-based platform** for accessible, reproducible, and transparent computational biomedical research

# What is Galaxy?

## GUI for genomics

- ✦ for complete analyses: analyze, visualize, share, publish

## A free and open Website (<http://usegalaxy.org>)

integrating a wealth of tools, compute resources, terabytes of reference data and permanent storage

## Open source software that makes it simple to

- ✦ integrate your own tools and data
- ✦ customize and run on your own resources

# Goal

An **open, Web-based** approach for  
making **highly interactive visual analysis**  
tools for **your NGS datasets**

# Goal

distributed, extendable, sharable, *fast*

An **open, Web-based** approach for

making **highly interactive visual analysis**

tools for **your NGS datasets**

# Goal

distributed, extendable, sharable, *fast*

An **open, Web-based** approach for

flexible, customizable

making **highly interactive visual analysis**

tools for **your NGS datasets**

# Goal

distributed, extendable, sharable, *fast*

An **open, Web-based** approach for

flexible, customizable

visualization + tools

making **highly interactive visual analysis**

tools for **your NGS datasets**

# Goal

distributed, extendable, sharable, *fast*

An **open, Web-based** approach for

flexible, customizable

visualization + tools

making **highly interactive visual analysis**

your own huge datasets

tools for **your NGS datasets**



**Let's go!**

# Trackster



# Paramamonster

Galaxy Analyze Data Workflow Shared Data Visualization Admin Help User Using 1.5 Gb

### Cufflinks (version 0.0.5)

**Max Intron Length:**  
300000

**Min Isoform Fraction:**  
0 - 0.1 samples: 3

**Pre MRNA Fraction:**  
0 - 0.1 samples: 3

**Perform quartile normalization:**  
No

Execute

Min Isoform Fraction      Pre MRNA Fraction

```
graph LR; Root((Root)) --- MIF_0((0)); Root --- MIF_005((0.05)); Root --- MIF_01((0.1)); MIF_0 --- PMRF_0_0((0)); MIF_0 --- PMRF_0_005((0.05)); MIF_0 --- PMRF_0_01((0.1)); MIF_005 --- PMRF_005_0((0)); MIF_005 --- PMRF_005_005((0.05)); MIF_005 --- PMRF_005_01((0.1)); MIF_01 --- PMRF_01_0((0)); MIF_01 --- PMRF_01_005((0.05)); MIF_01 --- PMRF_01_01((0.1));
```

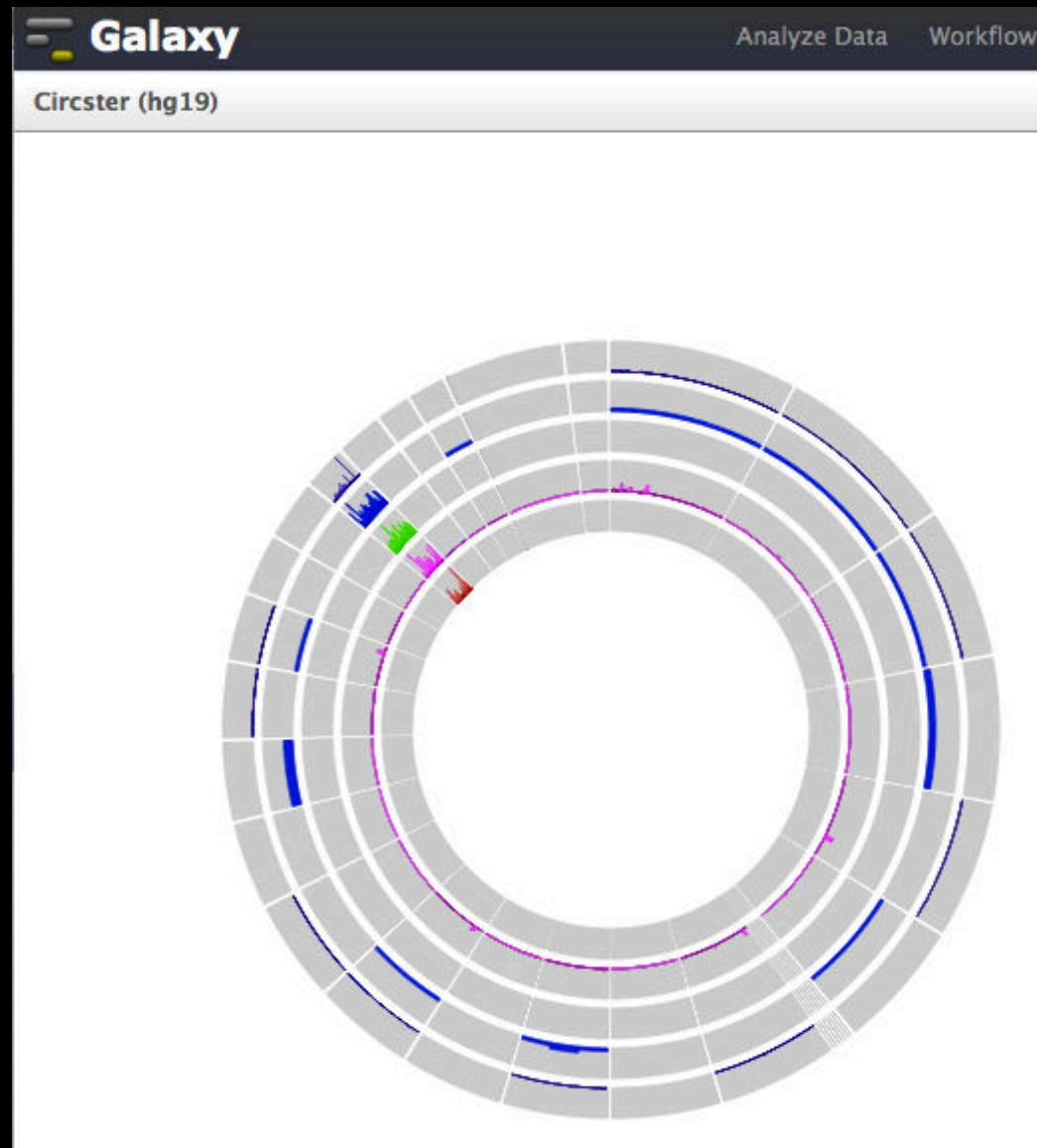
#### chr19:567970-588681

UFF.1.1  
CUFF.1.2  
CUFF.1.3  
CUFF.1.4  
CUFF.5.1      CUFF.2.1  
                  CUFF.3.1  
                  CUFF.4.1

UFF.1.1  
CUFF.1.2  
CUFF.1.3  
                  CUFF.2.1  
                  CUFF.3.1  
                  CUFF.4.1  
                  CUFF.5.1

UFF.1.1  
CUFF.1.2  
CUFF.1.3  
                  CUFF.2.1  
                  CUFF.3.1  
                  CUFF.4.1  
                  CUFF.5.1

# Circster



# Trackster

## Completely Web-based

- ♦ only a modern Web browser is needed

## Supports arbitrarily large NGS datasets

- ♦ SAM/BAM, BED, GFF/GTF, VCF, WIG

## Highly flexible

- ♦ e.g. custom rainbow tracks

## Integrated with Galaxy tool framework

- ♦ dynamic filtering
- ♦ re-running tools

# “Paramammonster”

Visualization for

- ✦ tool parameter space
- ✦ outputs from different settings

Can easily find good settings by visual inspection

- ✦ for many settings, across multiple regions

Can explore parameter space systematically or  
ad-hoc

# Circster

Circos-like visualization that provides genome-wide views

Complements Trackster

Very much a work in progress

# Summary

Galaxy is a Web-based platform for doing all facets of genomics analysis

Galaxy visualizations are

- ✦ Web-based
- ✦ highly interactive and dynamic
- ✦ integrated with Galaxy tools
- ✦ sharable and publishable

Only a Web browser is required for using any and all Galaxy features and visualizations





EMORY

PENNSTATE.



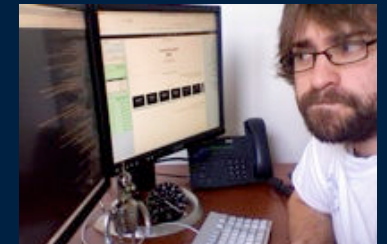
Enis Afgan



Dannon Baker



Dan Blankenberg



Nate Coraor



Dave Clements



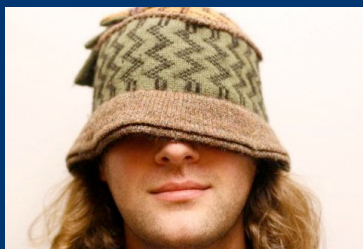
Jeremy Goecks



Jennifer Jackson



Greg von Kuster



James Taylor

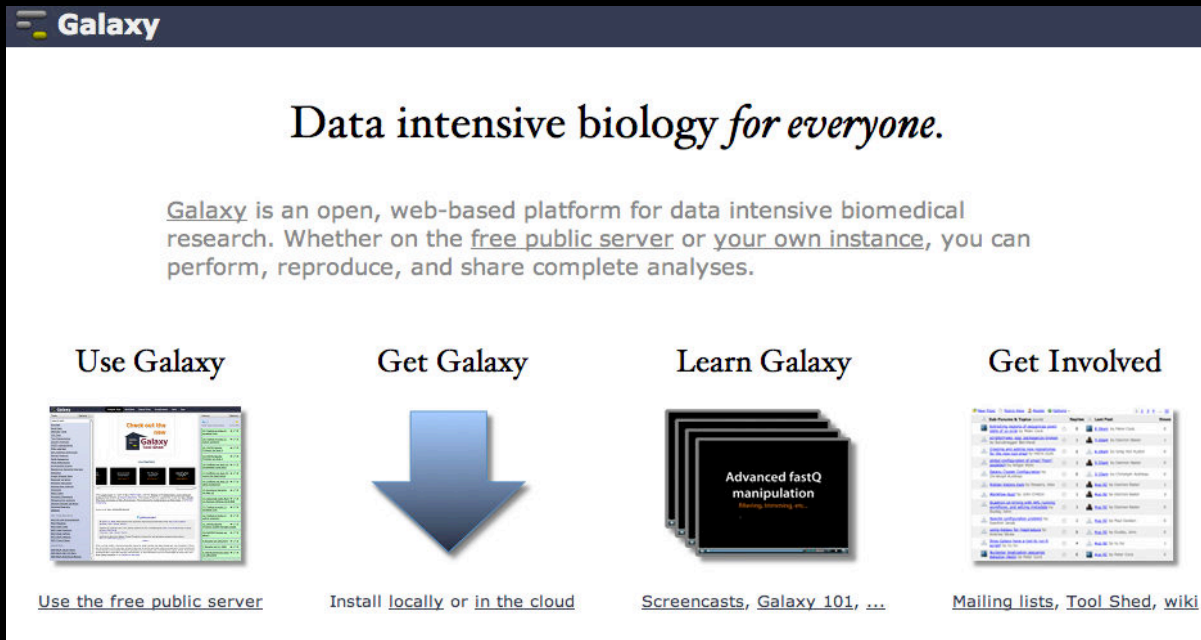


Anton Nekrutenko

Supported by the **NHGRI** (HG005542, HG004909, HG005133, HG006620), **NSF** (DBI-0850103), Penn State University, Emory University, and the Pennsylvania Department of Public Health

# Thanks! Questions?

<http://galaxyproject.org>



The screenshot shows the Galaxy project website homepage. At the top, the word "Galaxy" is displayed in a dark blue header. Below the header, the text "Data intensive biology *for everyone.*" is centered. Underneath, a paragraph describes Galaxy as an open, web-based platform for data intensive biomedical research, available on either a free public server or a user's own instance. The page is divided into four main sections: "Use Galaxy" with a screenshot of the web interface and the text "Use the free public server"; "Get Galaxy" with a large blue downward-pointing arrow and the text "Install locally or in the cloud"; "Learn Galaxy" with a stack of video thumbnails and the text "Screencasts, Galaxy 101, ..."; and "Get Involved" with a screenshot of a mailing list and the text "Mailing lists, Tool Shed, wiki".

Galaxy publications: <http://galaxyproject.org/wiki/Citing>

[jeremy.goecks@emory.edu](mailto:jeremy.goecks@emory.edu)