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

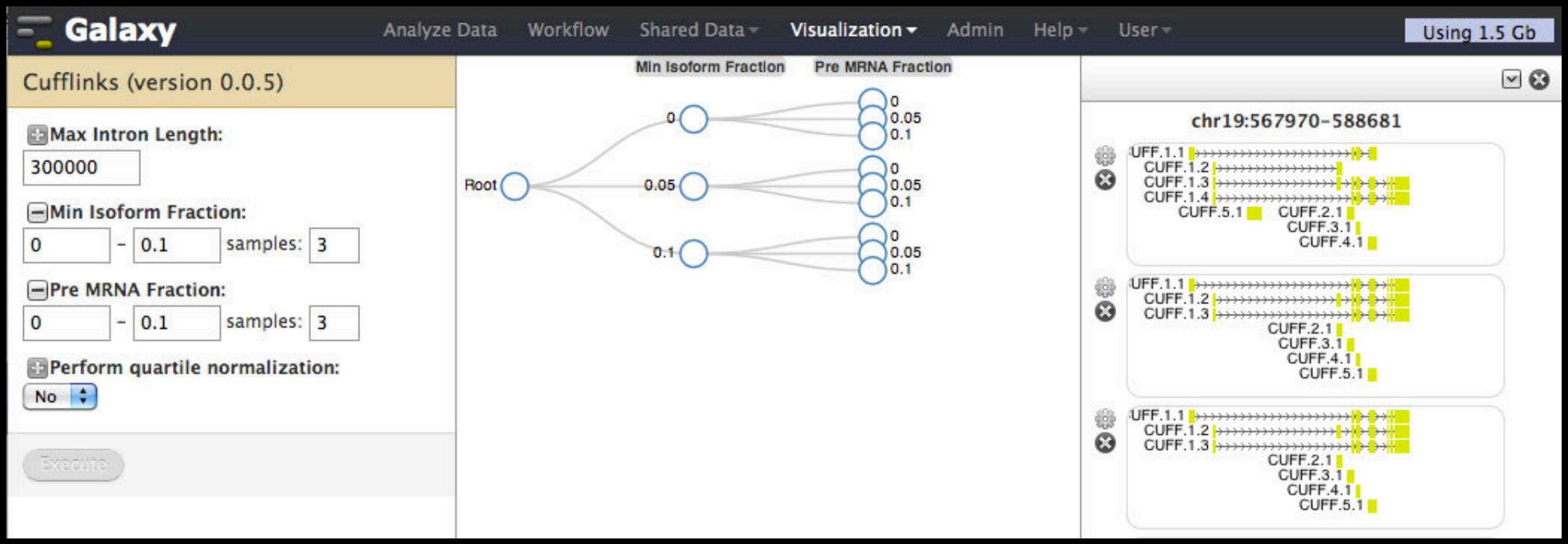
tools for your NGS datasets

Let’s go!

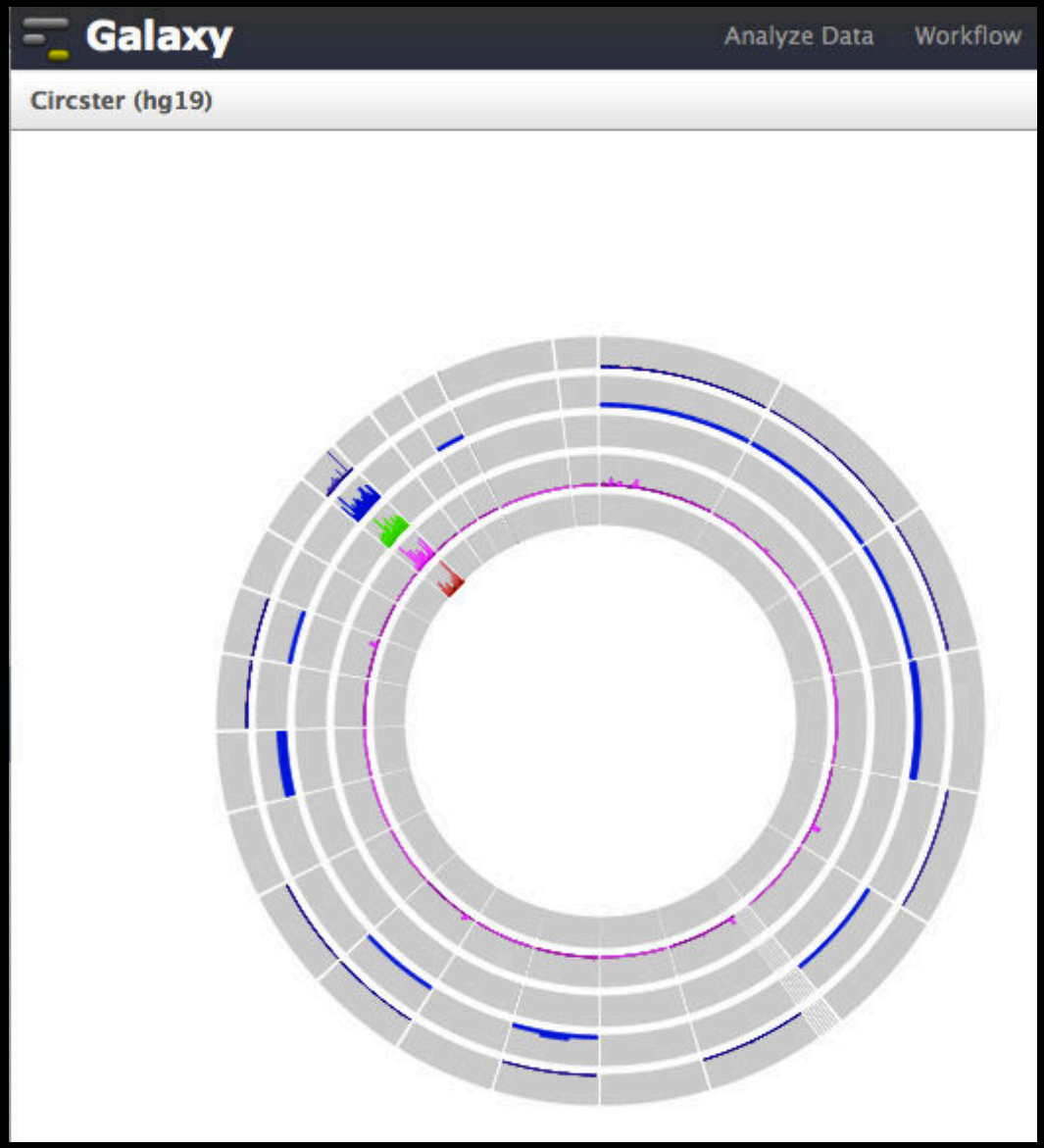
Trackster

Galaxy
Analyze Data
Workflow
Shared Data
Visualization
Admin
Help
User
Using 1.5 Gb
Demo (hg19)
chr19
0 - 11,255,991
0
10,000,000
GM12878 Cufflinks assembled transcripts
GM12878 Tophat Mapped Reads
2000
h1-hESC Cufflinks assembled transcripts
h1-hESC Tophat mapped reads
5219
UCSC Main on Human: phyloP44wayPrimate
1
-2
UCSC Main on Human: all_est
3000
UCSC Main on Human: knownGene

Paramamonster



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

“Paramamonster”

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

Enis Afgan

Dannon Baker

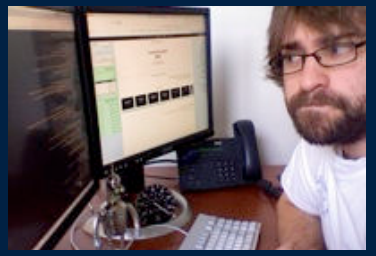
Dave Clements

Jeremy Goecks

James Taylor

PENNSTATE. 1855

Dan Blankenberg



Nate Coraor

Jennifer Jackson

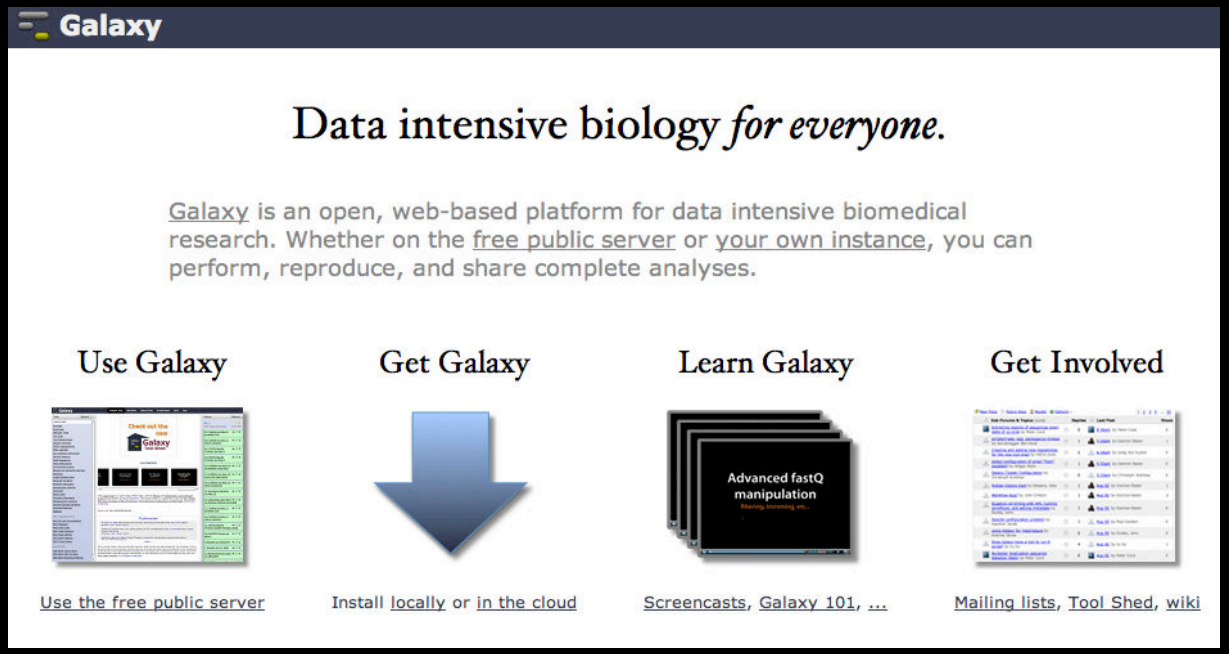
Greg von Kuster

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



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

jeremy.goecks@emory.edu