

The Galaxy Visual Analysis Framework

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

What is Galaxy?

Web-based GUI for genomics

- ✦ for complete analyses: obtain and integrate data, analyze, visualize, share, publish

A tool integration framework that makes it simple to chain tool usage together step-by-step or create complex workflows

Open source software that makes it simple to

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

<http://usegalaxy.org>

<http://galaxyproject.org>

Goal

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

Goal

distributed, extendable, sharable, *fast*

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

making **highly interactive visual analysis**

tools for **NGS datasets**

Goal

distributed, extendable, sharable, *fast*

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

flexible, customizable

making **highly interactive visual analysis**

tools for **NGS datasets**

Goal

distributed, extendable, sharable, *fast*

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

flexible, customizable

visualization + tools

making **highly interactive visual analysis**

tools for **NGS datasets**

Goal

distributed, extendable, sharable, *fast*

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

flexible, customizable

visualization + tools

making **highly interactive visual analysis**

needs to scale to huge datasets

tools for **NGS datasets**

Demo

Trackster



Paramammonster

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**

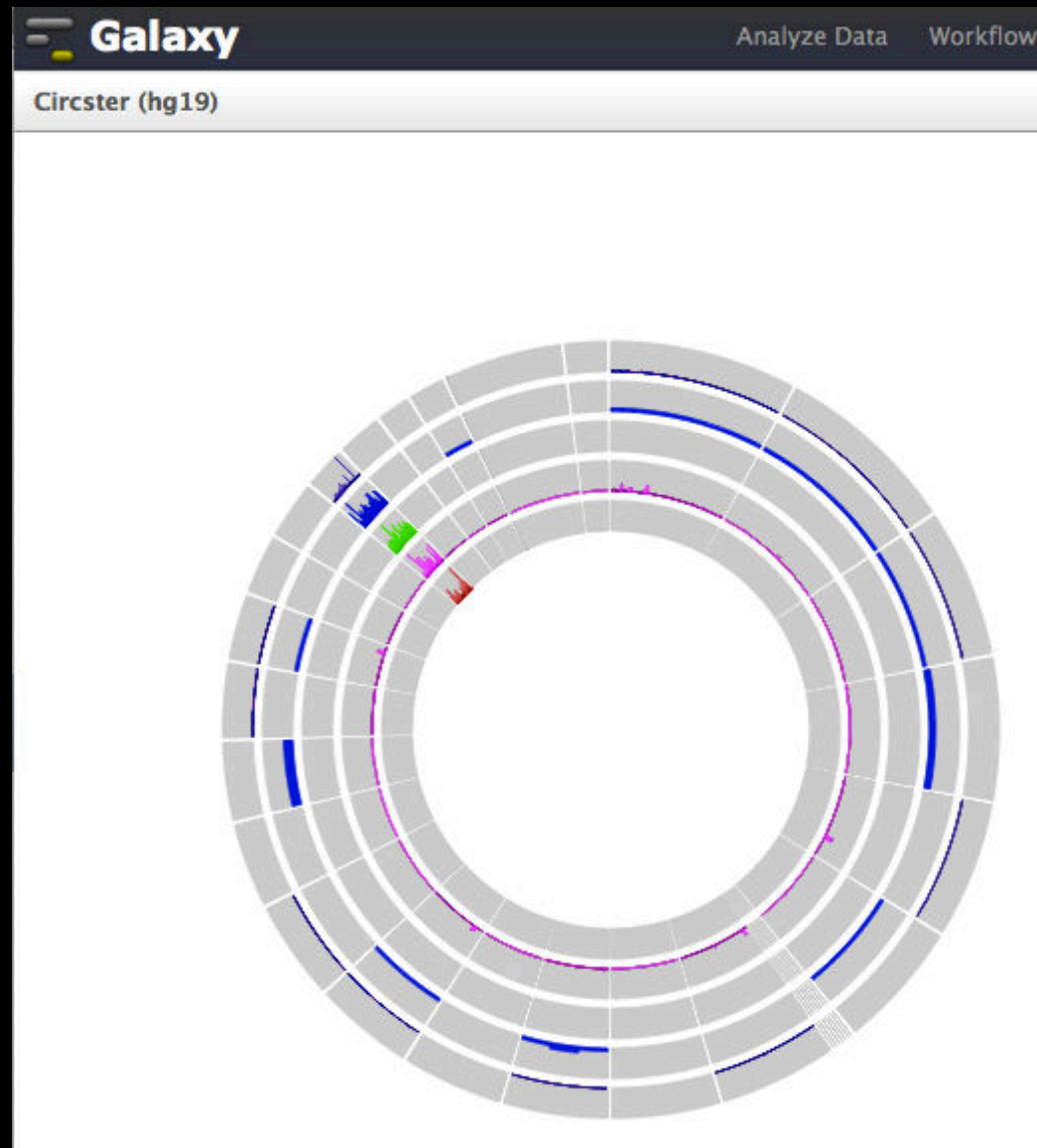
chr19:567970-588681

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

- ✦ no downloads, no add-ons, no Flash

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

Architecture

**Web
browser**

Galaxy
HTML UI

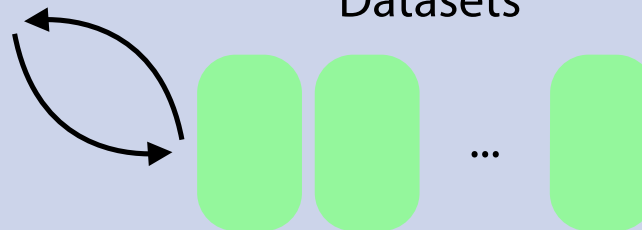


Galaxy

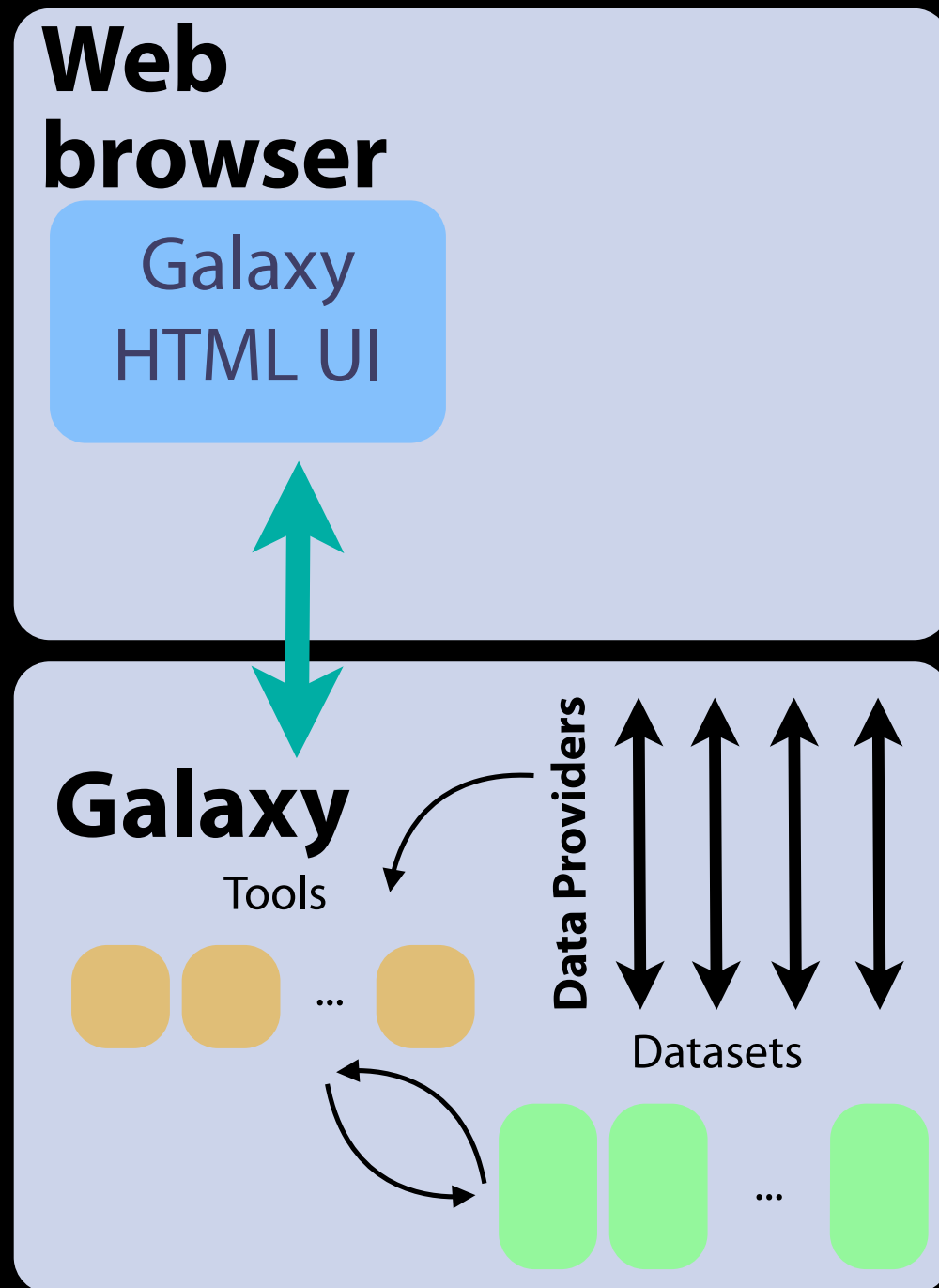
Tools



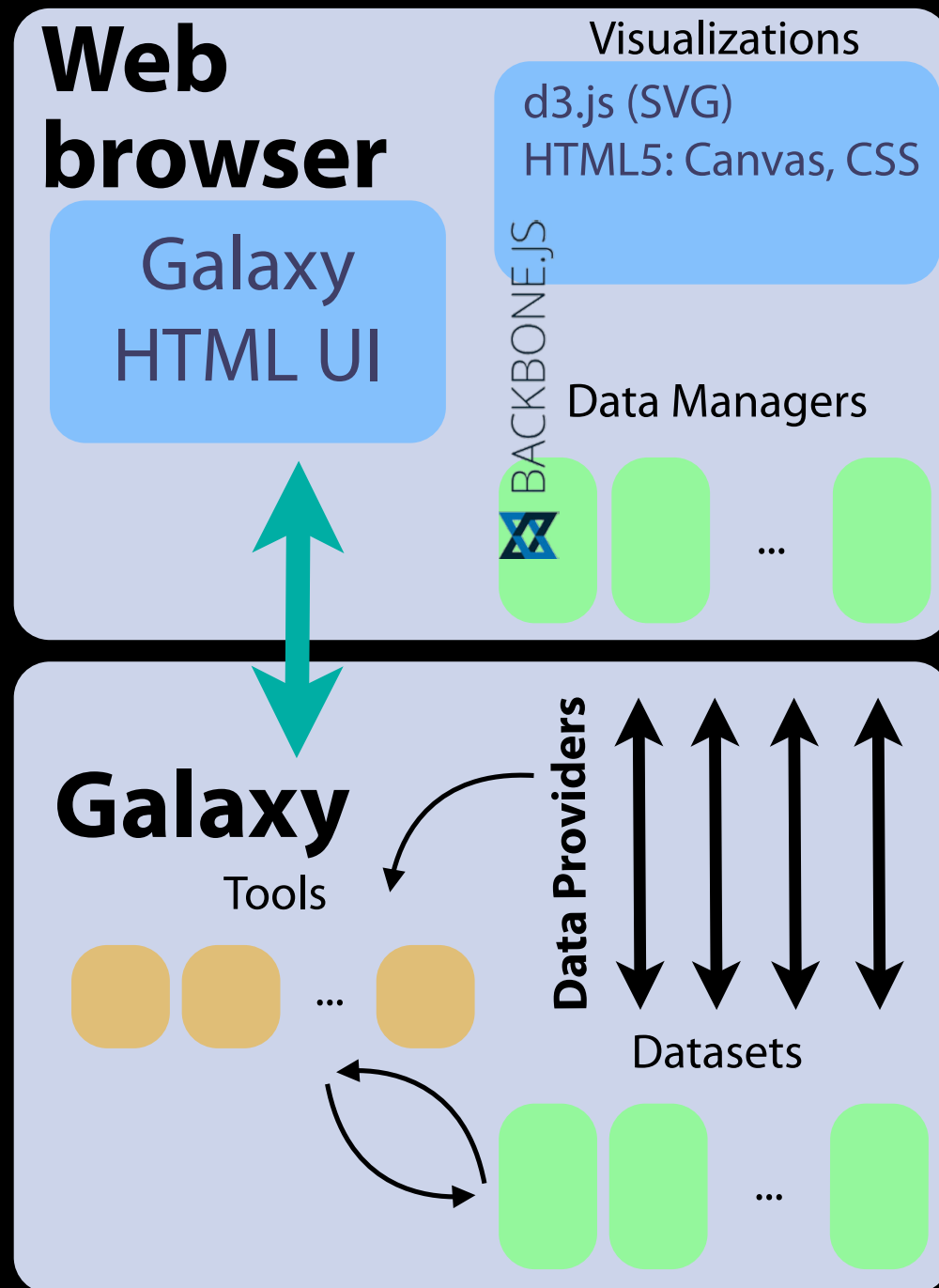
Datasets



Architecture



Architecture



Future Directions

Non-genomic visualizations

- ♦ phylogenetic trees
- ♦ scatterplots

Integration of multiple visualizations

- ♦ multiple views in same visualization
- ♦ views in different visualizations



EMORY

PENNSSTATE.



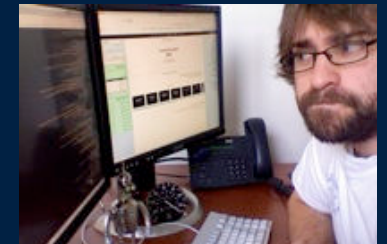
Enis Afgan



Dannon Baker



Dan Blankenberg



Nate Coraor



Dave Clements



Jeremy Goecks



Jennifer Jackson



Greg von Kuster



James Taylor

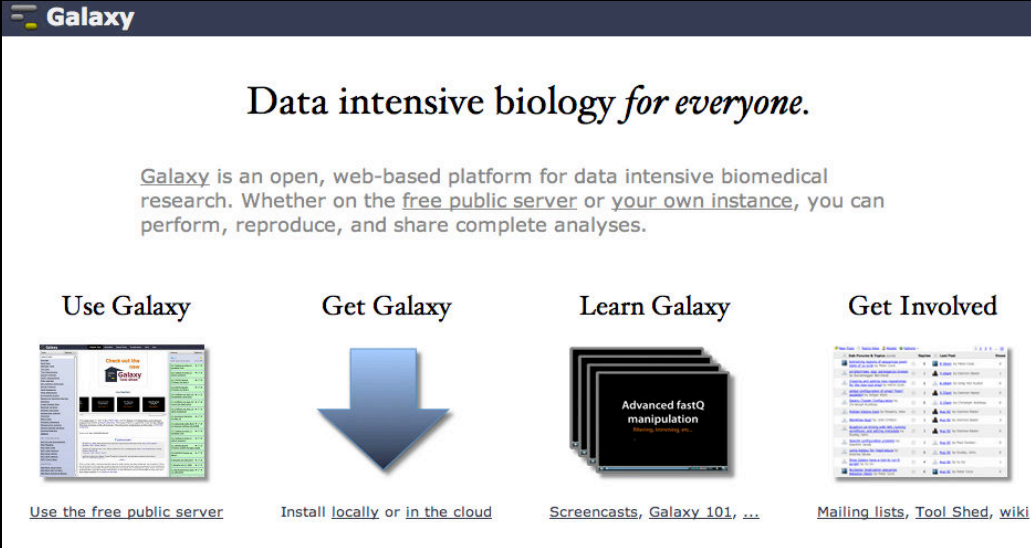


Anton Nekrutenko

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, it says "Galaxy" with a logo. Below that is the tagline "Data intensive biology *for everyone*." A paragraph describes Galaxy as an open, web-based platform for data intensive biomedical research. Below the text are four columns of navigation options: "Use Galaxy" (with a screenshot of the interface and the link "Use the free public server"), "Get Galaxy" (with a large blue downward arrow and the link "Install locally or in the cloud"), "Learn Galaxy" (with a stack of presentation slides titled "Advanced fastQ manipulation" and the link "Screencasts, Galaxy 101, ..."), and "Get Involved" (with a screenshot of a mailing list and the link "Mailing lists, Tool Shed, wiki").

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

Tech Track Talk (TT08): Sunday, 2:30p

jeremy.goecks@emory.edu