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

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

distributed, extendable, sharable, fast
An open, Web-based approach for distributed, extendable, sharable, fast 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

An open, Web-based approach for making highly interactive visual analysis tools for NGS datasets needs to scale to huge datasets.
Demo
Trackster
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
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
Architecture

Web browser

Galaxy HTML UI

Galaxy

Tools

Datasets

Tools...Datasets...
Future Directions

Non-genomic visualizations
- phylogenetic trees
- scatterplots

Integration of multiple visualizations
- multiple views in same visualization
- views in different visualizations
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

Tech Track Talk (TTo8): Sunday, 2:30p
jeremy.goecks@emory.edu