The Galaxy Visual Analysis Framework

Jeremy Goecks & The Galaxy Team
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
Goal

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

making highly interactive visual analysis

tools for NGS datasets
Goal

An open, Web-based approach for distributed, extendable, sharable, fast flexible, customizable visualization + tools making highly interactive visual analysis needs to scale to huge datasets tools for NGS datasets
Not everything for everyone, but pretty sweet nonetheless.
Demo
Trackster
Sweepster
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
5C data in Trackster
5C data in Trackster
Circster

Circos-like visualization that provides genome-wide views

Complements Trackster

Very much a work in progress
Sweepster

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
Sharing!
Architecture

Web browser

Galaxy

HTML UI

Galaxy

Tools

Data Providers

Datasets
Technical Highlights

Rough pluggable support for generic JavaScript visualizations + data providers

Data providers = fast, random access to data in Python, JS

API: run tools, run them on data subsets

Backbone + HTML5 objects for Web-based genomic visualizations
  * e.g. data managers, linear and circular views

JS binding to Galaxy API (blendJS?)
  * visualizations, tools, datasets
  * custom Galaxy UIs
// -- Viz set up. --

var genome = new Genome(JSON.parse('${ h.to_json_string( genome ) }'));
visualization = new GenomeVisualization(JSON.parse('${ h.to_json_string( viz_config ) }'));
viz_view = new CircsterView({
    width: 600,
    height: 600,
    // Gap is difficult to set because it very dependent on chromosome size and organization.
    total_gap: 2 * Math.PI * 0.2,
    genome: genome,
    model: visualization,
    radius_start: 100,
    dataset_arc_height: 15
});

// -- Render viz. --

viz_view.render();
$('#vis').append(viz_view.$el);
Future Directions

Easy access to outside resources/tracks

More tools + workflows

Incorporate visual analysis into main analysis UI

Non-genomic visualizations

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