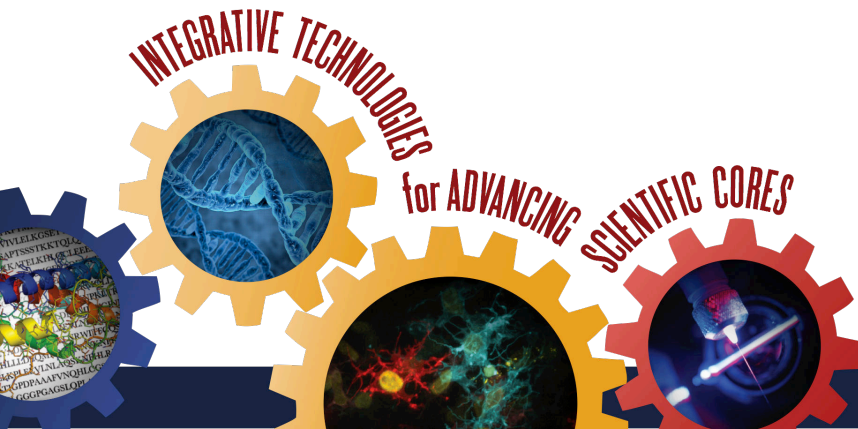


# Visualization Methods within Galaxy

**Aysam Guerler**

Software Engineer – Taylor Lab.

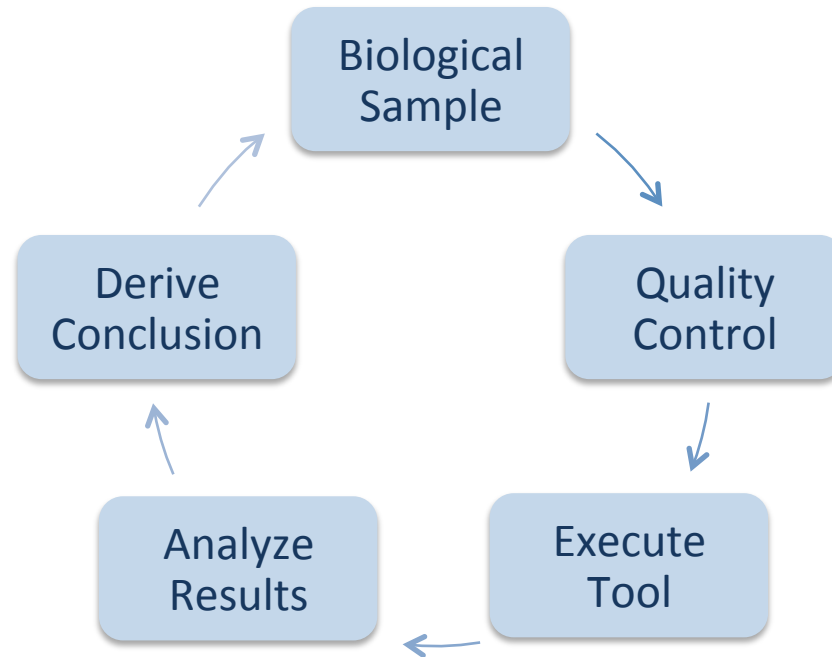


**JOHNS HOPKINS**  
UNIVERSITY



**Galaxy**  
PROJECT

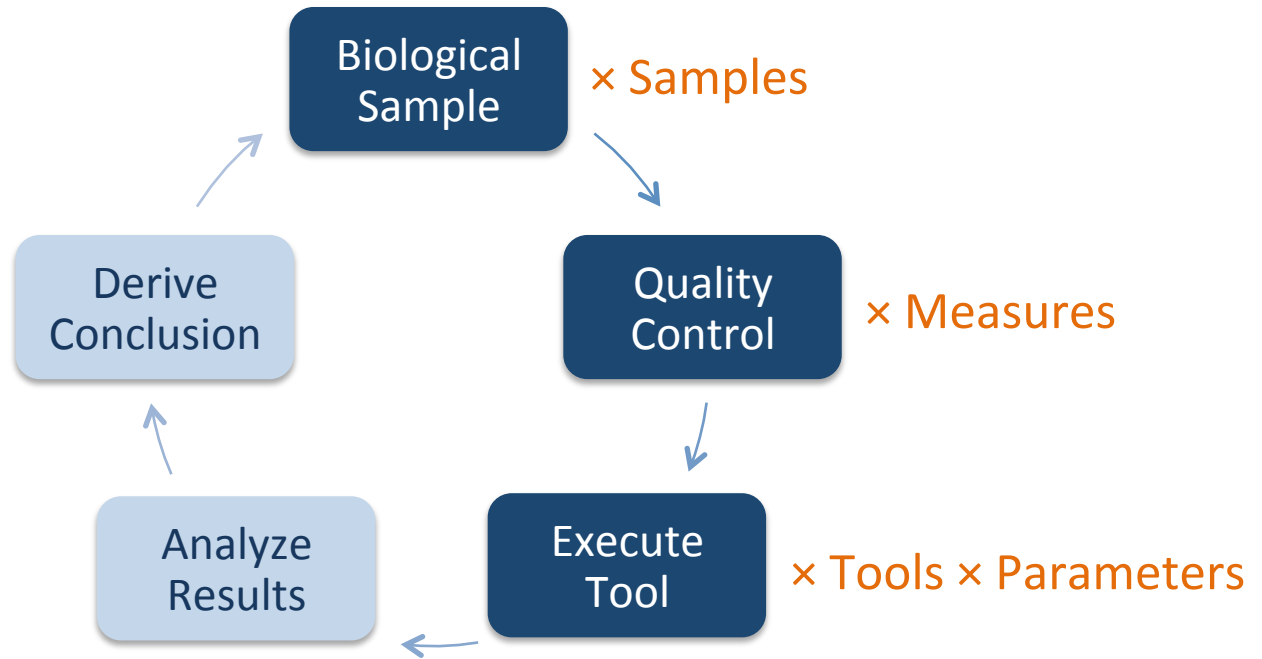
# Why Galaxy?



**INTEGRATIVE TECHNOLOGIES**  
for **ADVANCING SCIENTIFIC CORES**

# Why Galaxy?

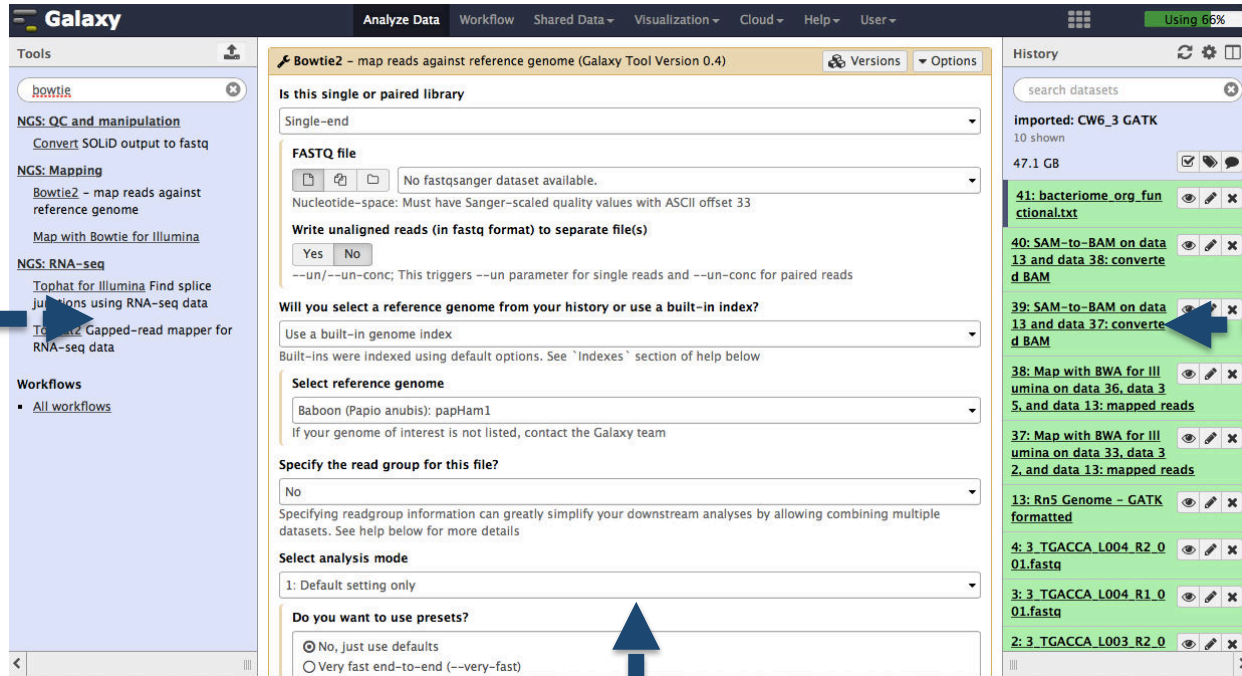
More realistic:



INTEGRATIVE TECHNOLOGIES  
for ADVANCING SCIENTIFIC CORES

# Why Galaxy?

Tools



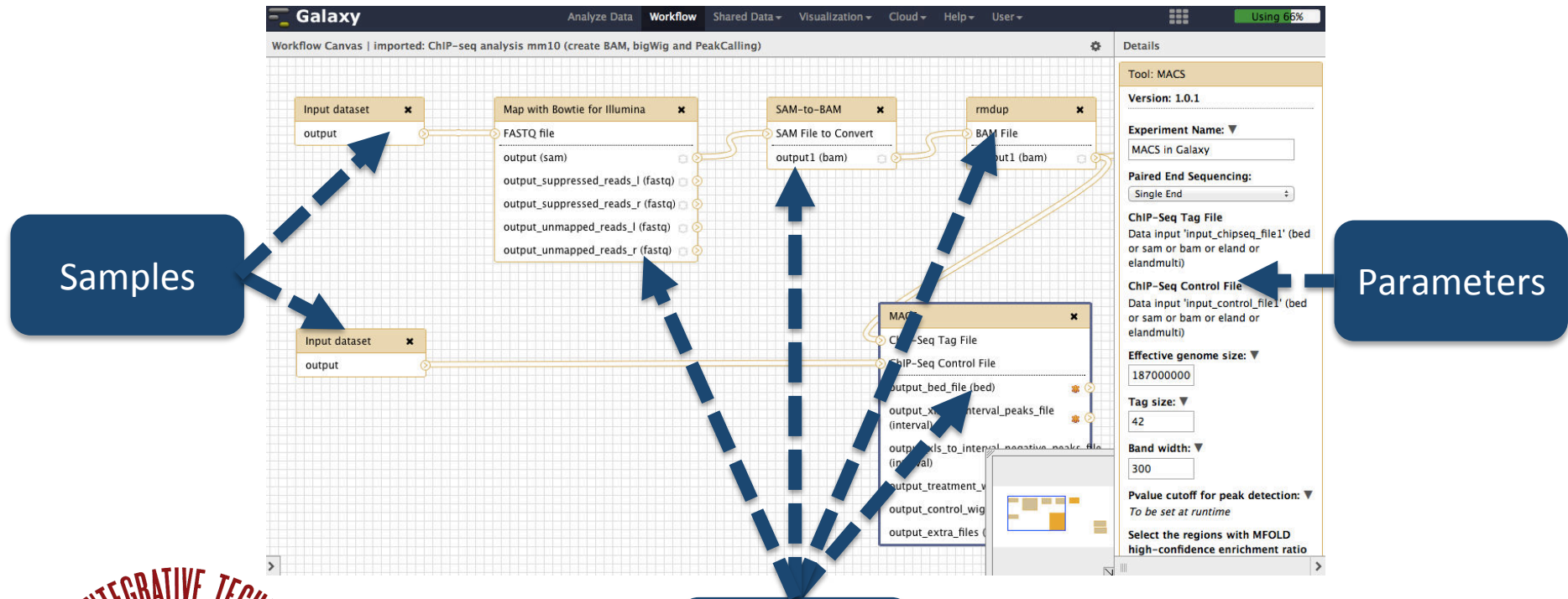
Samples

Parameters



**Galaxy**  
www.usegalaxy.org

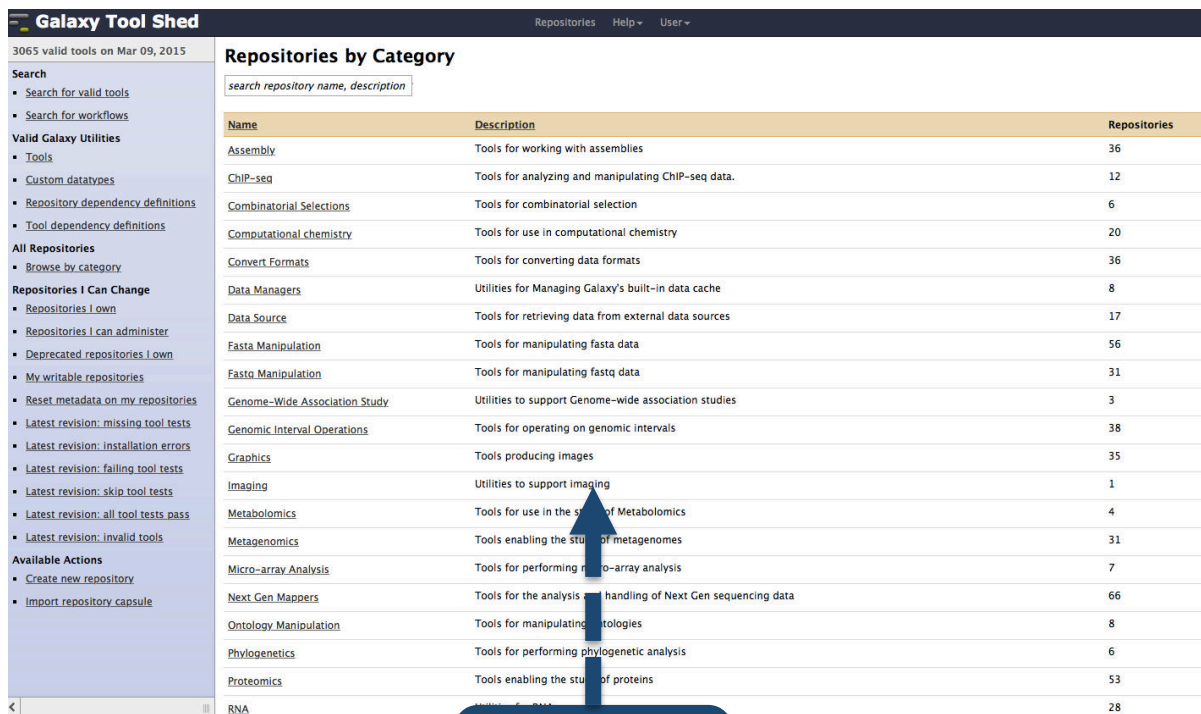
# Why Galaxy?



INTEGRATIVE TECHNOLOGIES  
for ADVANCING SCIENTIFIC CORES

**Galaxy**  
[www.usegalaxy.org](http://www.usegalaxy.org)

# Why Galaxy?



**Galaxy Tool Shed**    Repositories    Help    User

3065 valid tools on Mar 09, 2015

**Repositories by Category**

search repository name, description

Name	Description	Repositories
Assembly	Tools for working with assemblies	36
ChIP-seq	Tools for analyzing and manipulating ChIP-seq data.	12
Combinatorial Selections	Tools for combinatorial selection	6
Computational chemistry	Tools for use in computational chemistry	20
Convert Formats	Tools for converting data formats	36
Data Managers	Utilities for Managing Galaxy's built-in data cache	8
Data Source	Tools for retrieving data from external data sources	17
Fasta Manipulation	Tools for manipulating fasta data	56
Fasta Manipulation	Tools for manipulating fastq data	31
Genome-Wide Association Study	Utilities to support Genome-wide association studies	3
Genomic Interval Operations	Tools for operating on genomic intervals	38
Graphics	Tools producing images	35
Imaging	Utilities to support imaging	1
Metabolomics	Tools for use in the study of Metabolomics	4
Metagenomics	Tools enabling the study of metagenomes	31
Micro-array Analysis	Tools for performing micro-array analysis	7
Next Gen Mappers	Tools for the analysis and handling of Next Gen sequencing data	66
Ontology Manipulation	Tools for manipulating ontologies	8
Phylogenetics	Tools for performing phylogenetic analysis	6
Proteomics	Tools enabling the study of proteins	53
RNA		28

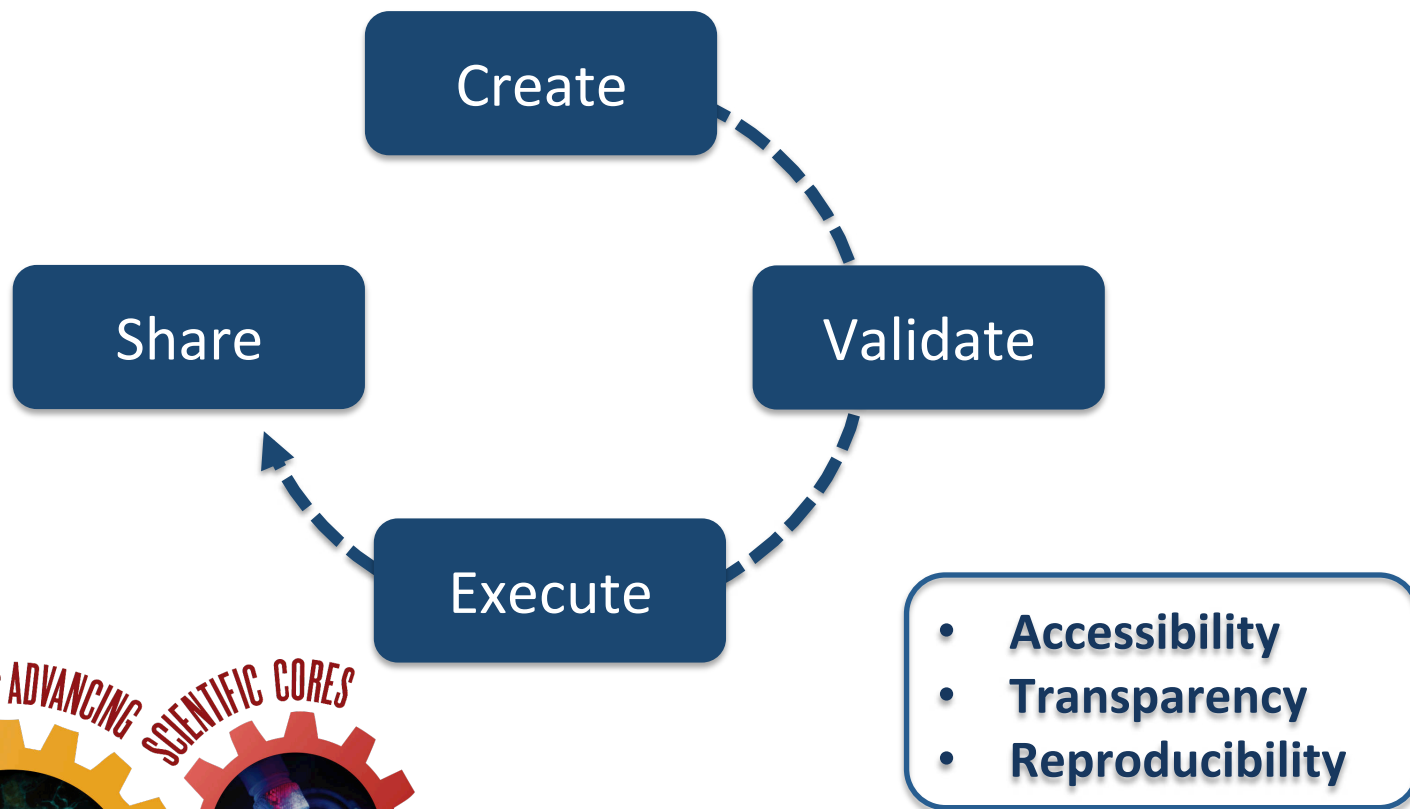
**Over 3k  
Tools**

**Galaxy**  
 www.usegalaxy.org



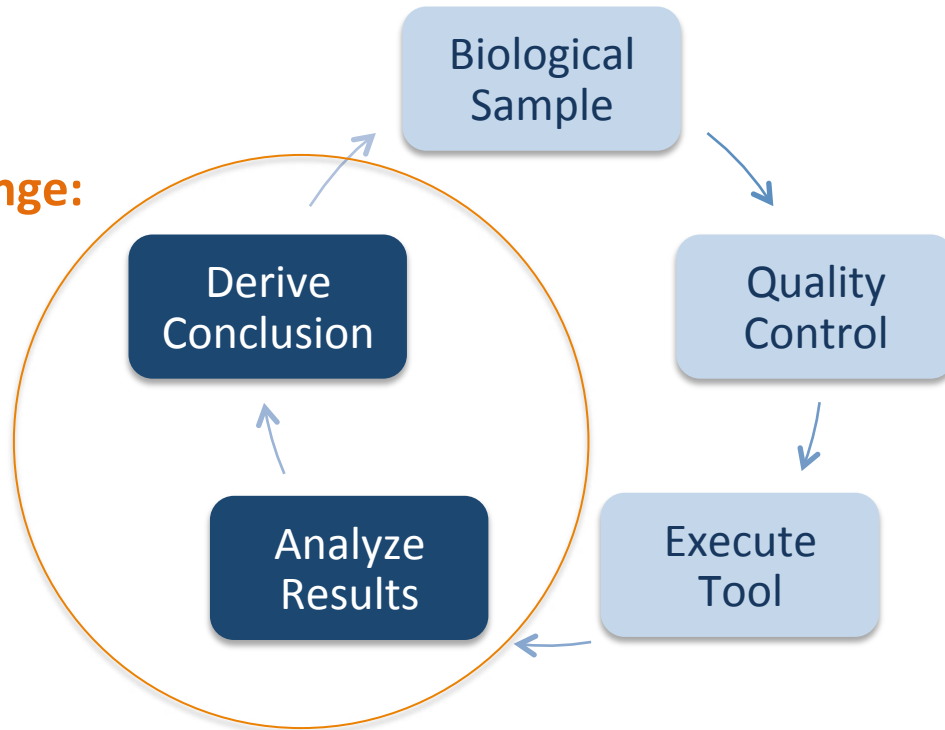
# Why Galaxy?

Community-driven HPC  
**Workflows for Researchers**



# What's missing?

**The Challenge:**



**INTEGRATIVE TECHNOLOGIES**  
for **ADVANCING SCIENTIFIC CORES**

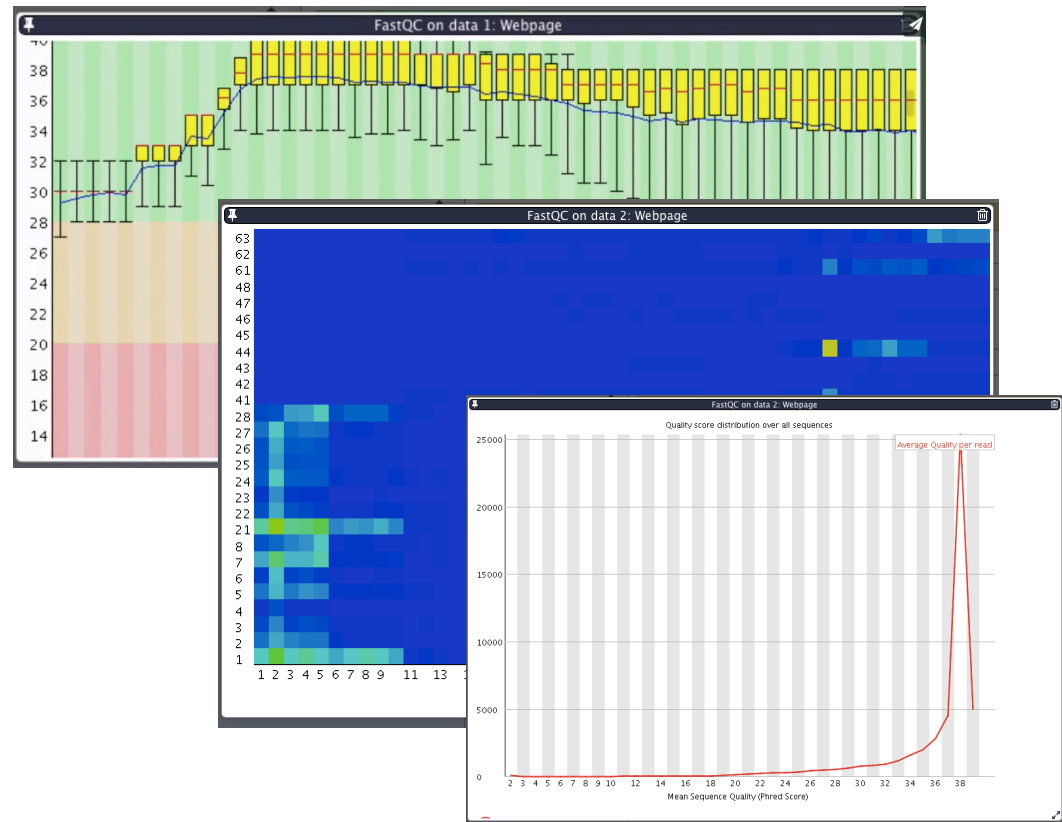


# FastQC

FastQC aims to provide a simple way to do some quality control checks on raw sequence data coming from high throughput sequencing pipelines.

## Reference

[www.bioinformatics.bbsrc.ac.uk](http://www.bioinformatics.bbsrc.ac.uk)



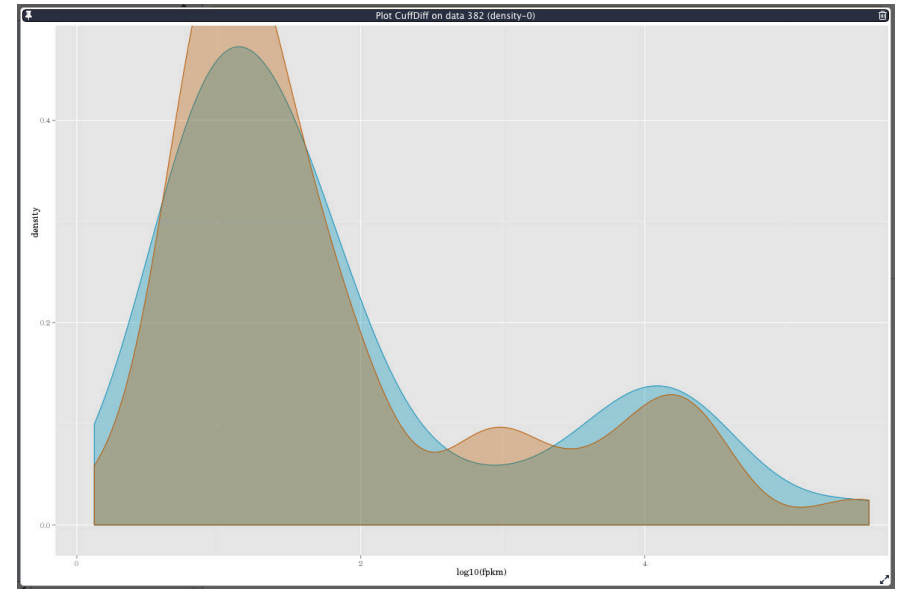
INTEGRATIVE TECHNOLOGIES  
for ADVANCING SCIENTIFIC CORES

# cummeRbund

cummeRbund is a visualization package for Cufflinks sequencing data. It assists in the analysis of RNA-Seq differential expression analysis.

## Reference

<http://compbio.mit.edu>

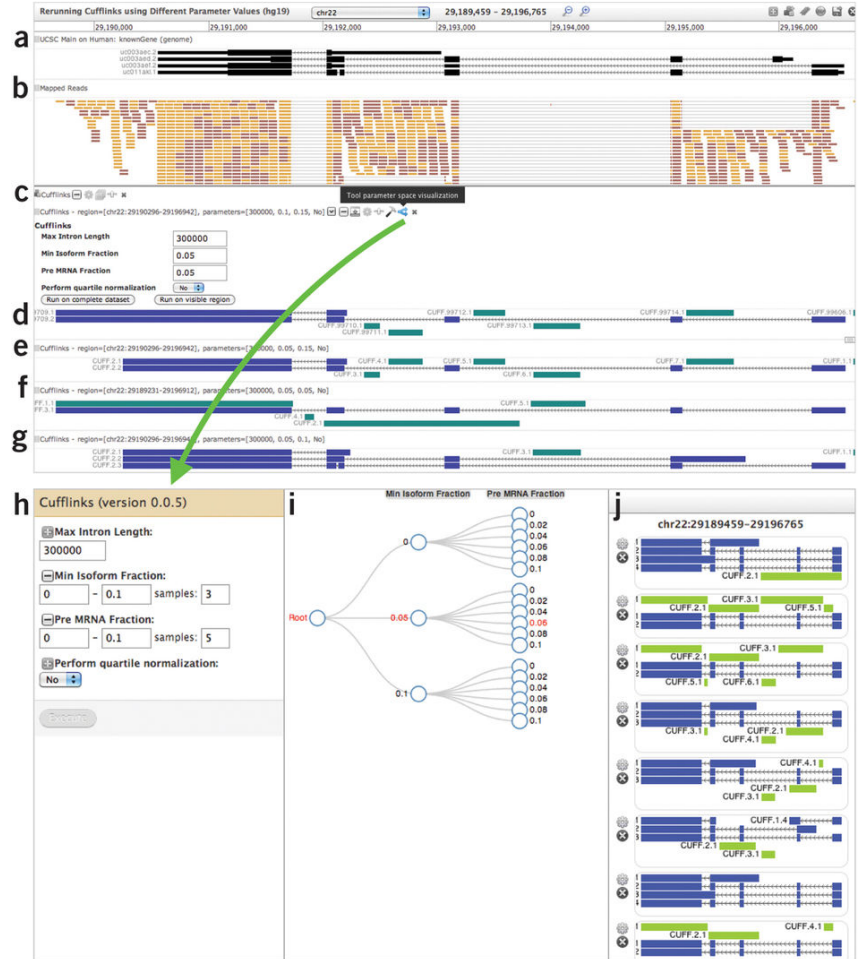


# Trackster

A genome browser with the ability to sample parameters and re-run tools on specific genome regions

## Reference

*NGS analyses by visualization with Trackster, Goecks et al., Nat. Biotech. 30(11), 1036-9*



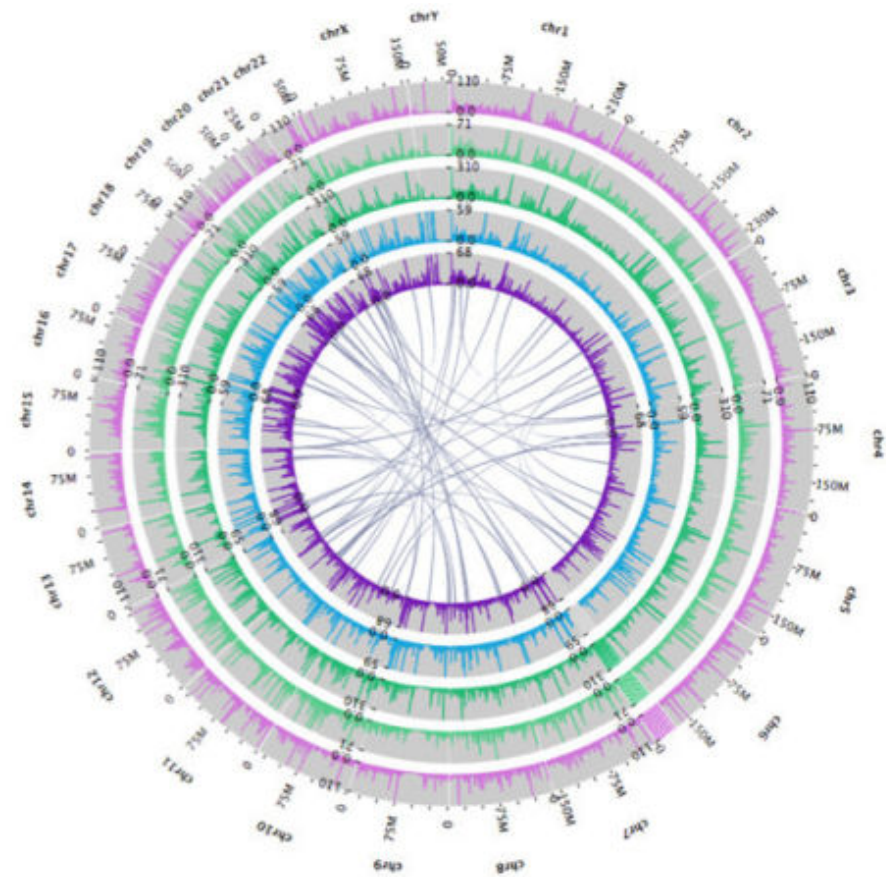
INTEGRATIVE TECHNOLOGIES  
 for ADVANCING SCIENTIFIC CORES

# Circster

Circster is a Circos-style viewer for rendering genome-wide data. Position-based data is laid out in concentric circles representing chromosome positions.

## Reference

*Web-based visual analysis for high-throughput genomics, Goecks et al., BMC Genomics 2013, 14:397*



**INTEGRATIVE TECHNOLOGIES**  
for **ADVANCING SCIENTIFIC CORES**

# Charts

An interface for a range of common visualization types i.e. bar diagrams, pie charts, scatter plots, heat maps and others.

## Reference

- [www.nvd3.com](http://www.nvd3.com)
- [www.jqplot.com](http://www.jqplot.com)

### • Bar diagrams



Regular (NVD3)



Stacked (NVD3)

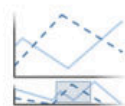


Horizontal (NVD3)



Stacked horizontal (NVD3)

### • Others



QLine with focus (NVD3)



QLine chart (NVD3)



QScatter plot (NVD3)



QHeatmap (Custom)

### • Area charts



QRegular (NVD3)



QExpanded (NVD3)



QStream (NVD3)



Pie chart (NVD3)

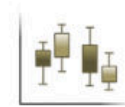
### • Data processing (requires 'charts' tool from Toolshed)



Histogram (NVD3)



QDiscrete Histogram (jqPlot)



QBox plot (jqPlot)



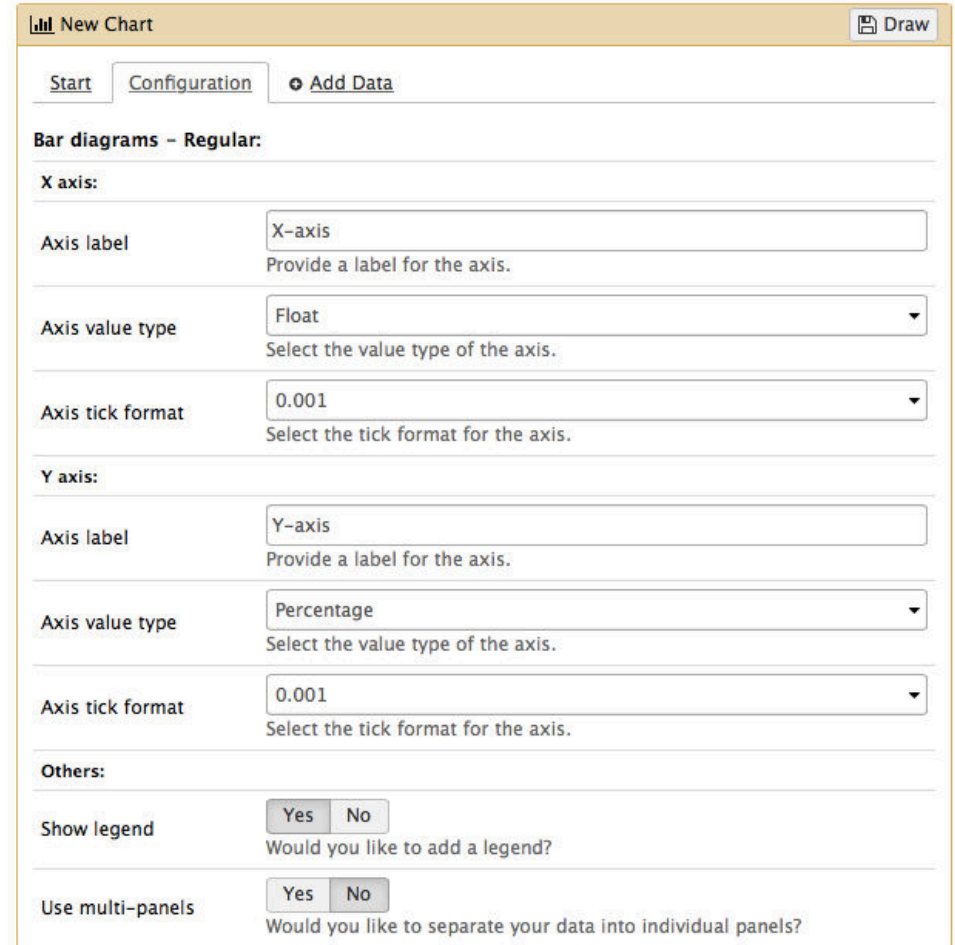
QClustered Heatmap (Custom)

INTEGRATIVE TECHNOLOGIES  
for ADVANCING SCIENTIFIC CORES

# Charts

## Customization

Visualizations can be customized and data columns can be associated from tabular datasets.



**New Chart** Draw

Start Configuration [Add Data](#)

**Bar diagrams – Regular:**

**X axis:**

Axis label: X-axis  
Provide a label for the axis.

Axis value type: Float  
Select the value type of the axis.

Axis tick format: 0.001  
Select the tick format for the axis.

**Y axis:**

Axis label: Y-axis  
Provide a label for the axis.

Axis value type: Percentage  
Select the value type of the axis.

Axis tick format: 0.001  
Select the tick format for the axis.

**Others:**

Show legend:  Yes  No  
Would you like to add a legend?

Use multi-panels:  Yes  No  
Would you like to separate your data into individual panels?

INTEGRATIVE TECHNOLOGIES  
for ADVANCING SCIENTIFIC CORES

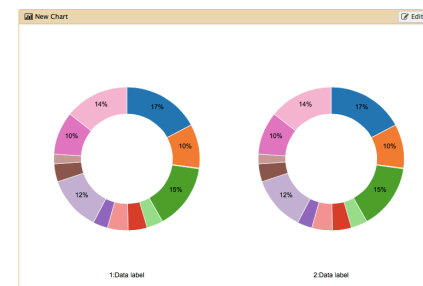
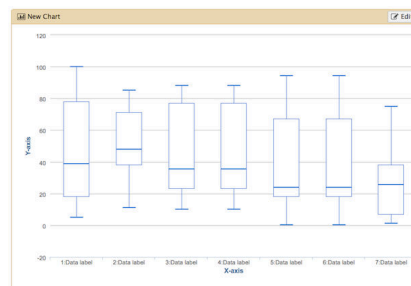
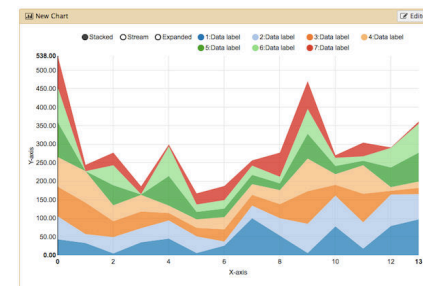
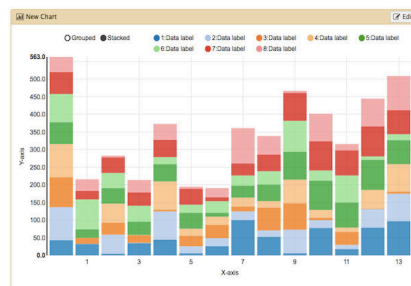
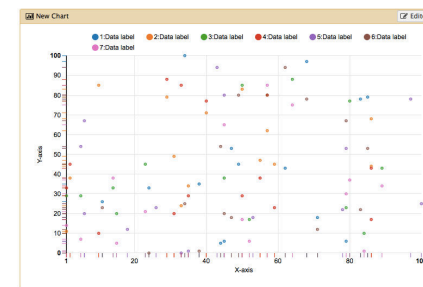
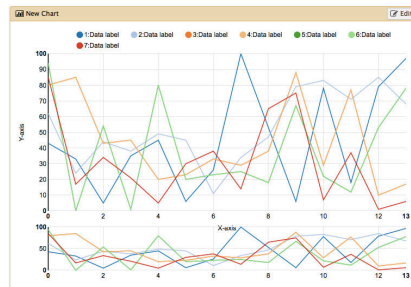
# Charts

## Examples

In principal, Charts is parsing data from Galaxy to 3<sup>rd</sup> party plugins.

## Open-source libraries

- [www.nvd3.com](http://www.nvd3.com)
- [www.jqplot.com](http://www.jqplot.com)



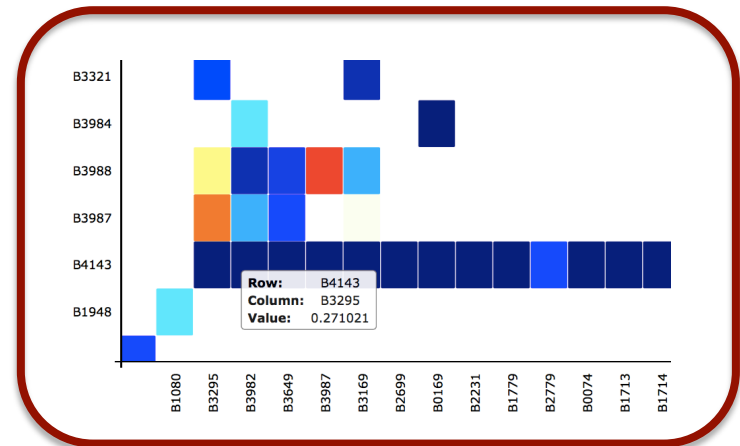
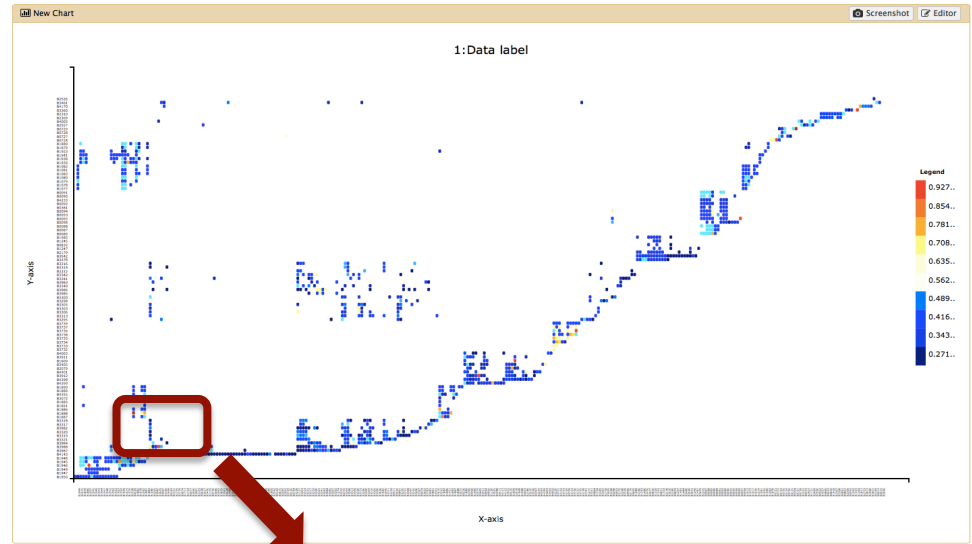
# Charts

## Clustered Heat Map

If a plugin is unavailable, custom visualization types like this heat map here can be added. Charts is able to pre-process large-scale data behind the scenes.

**Protein-Protein-Interaction data**

[www.bacteriome.org](http://www.bacteriome.org)





# deepTools

## ARTICLE

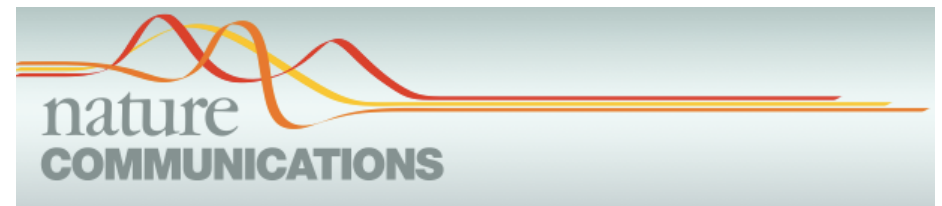
Received 15 May 2014 | Accepted 17 Sep 2014 | Published 22 Oct 2014

DOI: 10.1038/ncomms6288

OPEN

# Dynamic DNA methylation orchestrates cardiomyocyte development, maturation and disease

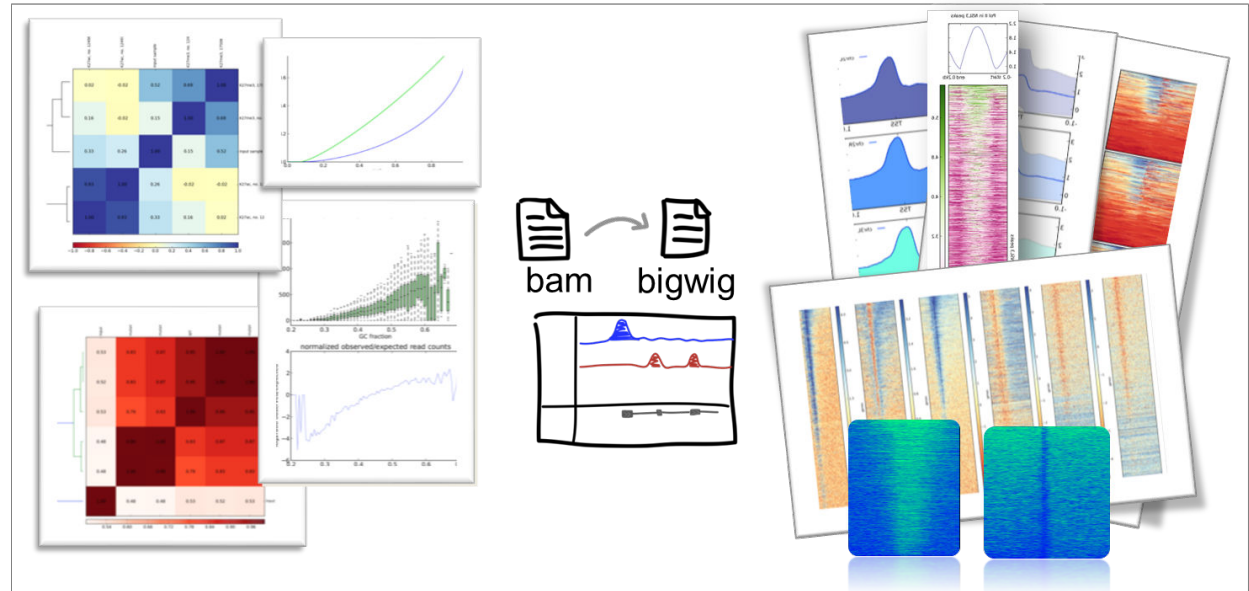
Ralf Gilsbach<sup>1,\*</sup>, Sebastian Preissl<sup>1,2,\*</sup>, Björn A. Grüning<sup>3,4</sup>, Tilman Schnick<sup>1,5</sup>, Lukas Burger<sup>6,7</sup>, Vladimir Benes<sup>8</sup>, Andreas Würch<sup>9</sup>, Ulrike Bönisch<sup>9</sup>, Stefan Günther<sup>4</sup>, Rolf Backofen<sup>3</sup>, Bernd K. Fleischmann<sup>10</sup>, Dirk Schübeler<sup>6,11</sup> & Lutz Hein<sup>1,12</sup>



# deepTools

## QUALITY CHECKS – FORMAT CONVERSION & NORMALIZATION – PLOTTING

A suite of user-friendly tools for the visualization, quality control and normalization of data from deep-sequencing experiments.



### Reference

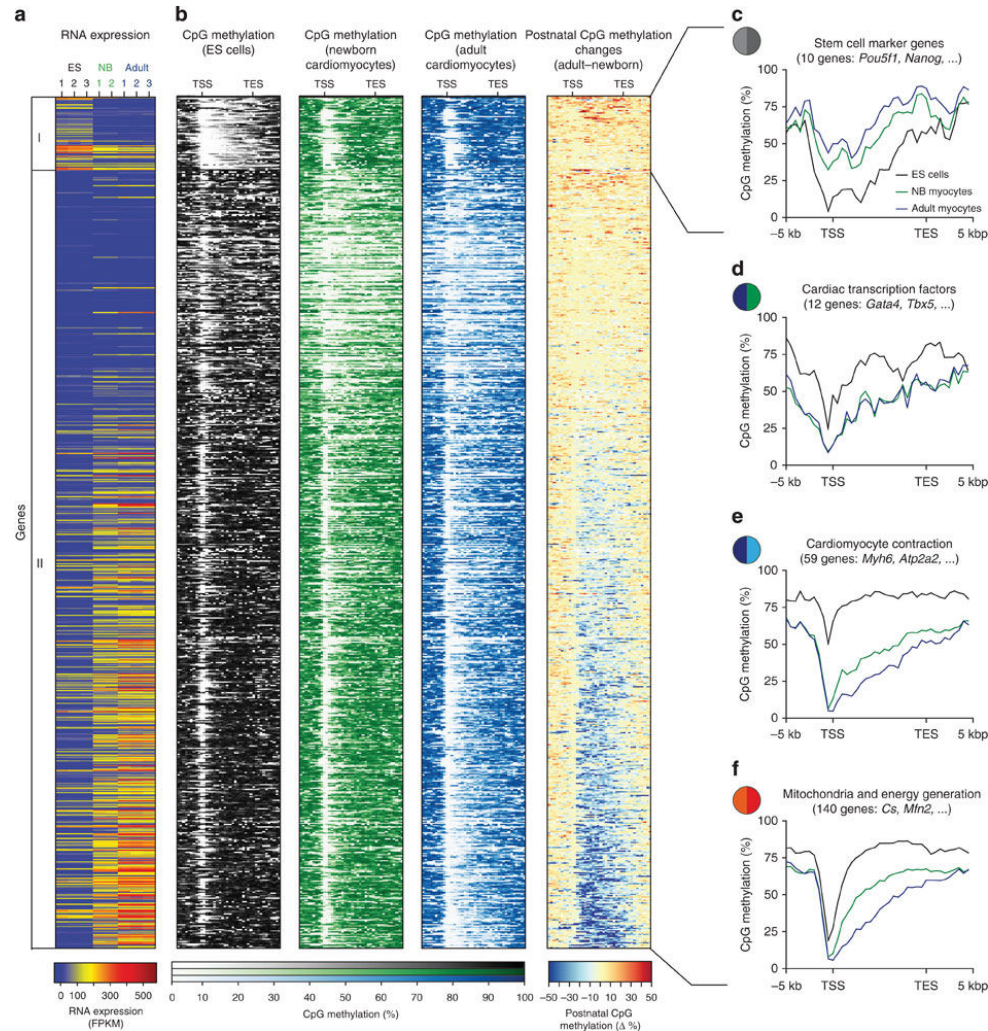
Ramírez et al. *deepTools: a flexible platform for exploring deep-sequencing data*. NAR, 2014

INTEGRATIVE TECHNOLOGIES  
for ADVANCING SCIENTIFIC CORES

# deepTools

## Heat Map and Line Plots

Heat maps of gene expression and CpG methylation of gene bodies and flanking regions ( $\pm 5$  kbp) with differential methylation in neonatal and adult cardiomyocytes versus ES cells.



INTEGRATIVE TECHNOLOGIES  
 for ADVANCING SCIENTIFIC CORES

**ABRF 2015**  
**ANNUAL MEETING**  
**MARCH 28-31, 2015**

**Thank you.**



**AMERICA'S CONVENTION CENTER • ST. LOUIS, MISSOURI**