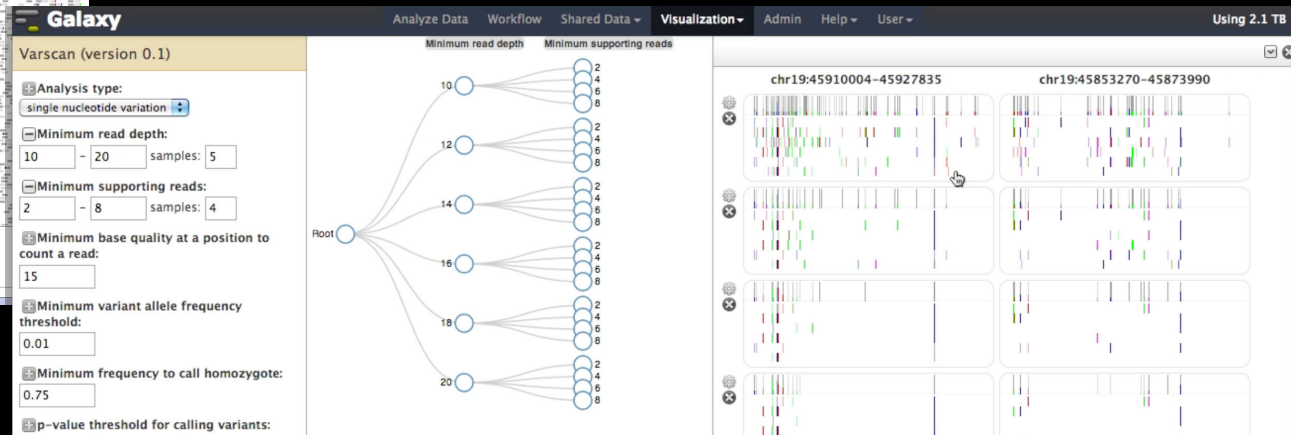
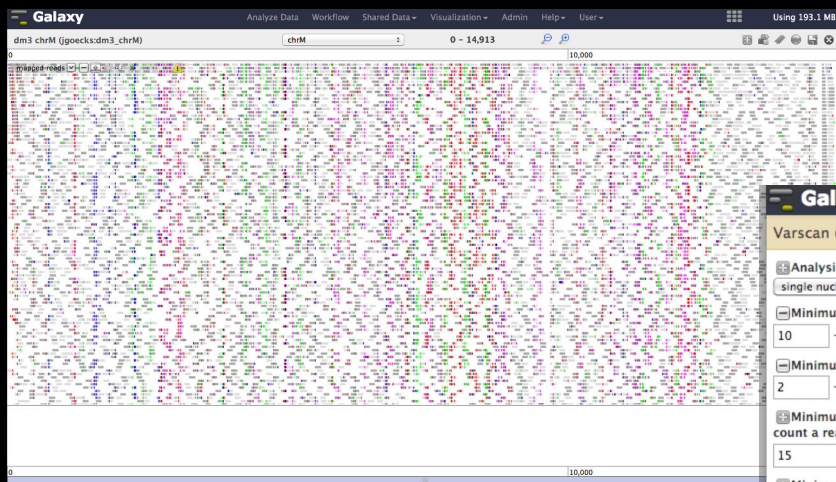




Galaxy Community Conference 2014: Visualization Workshop

Jeremy Goecks

Assistant Prof. of Comp. Biology

THE GEORGE WASHINGTON UNIVERSITY

WASHINGTON, DC

Sam Guerler

Research Software Engineer



JOHNS HOPKINS
UNIVERSITY

Topics

Visualization history and introduction

Biological Visualizations

Numerical Visualizations

Adding your own visualizations

Why Visualize?

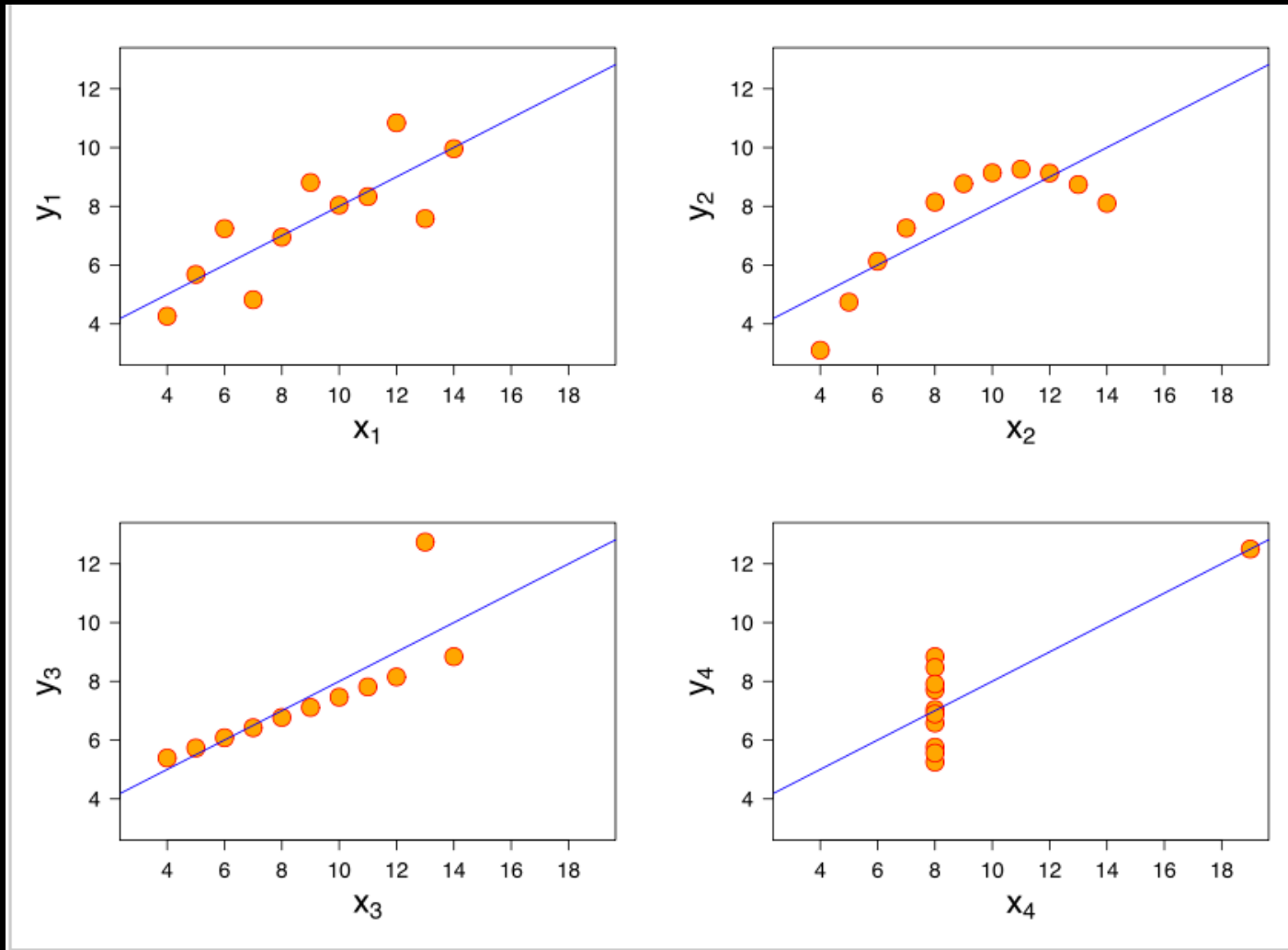
Why Visualize?

Quick check: did it work?

Exploration and hypothesis generation

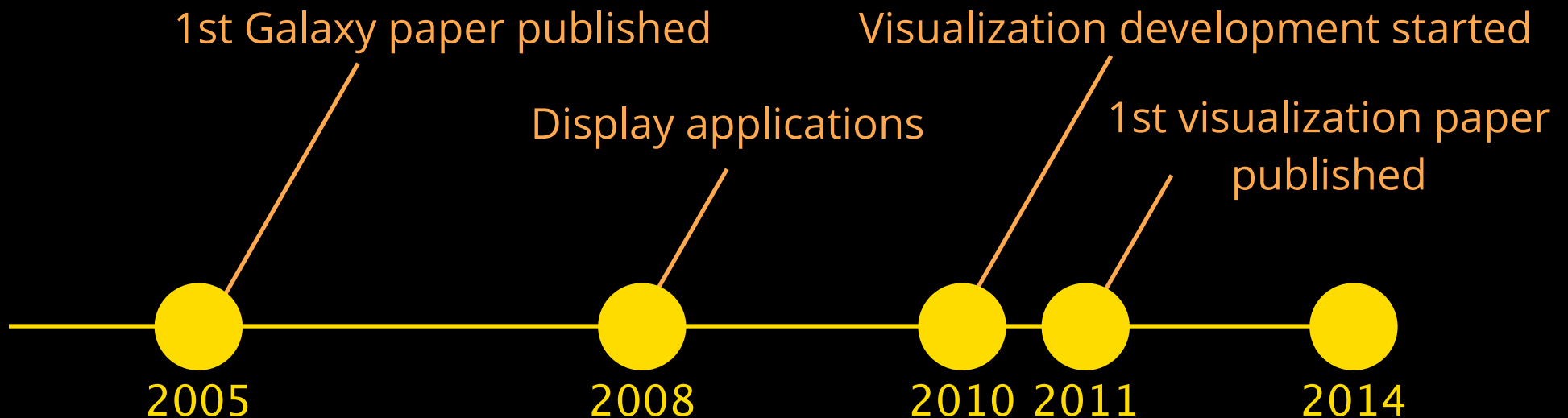
Sharing/publishing

Anscombe's Quartet

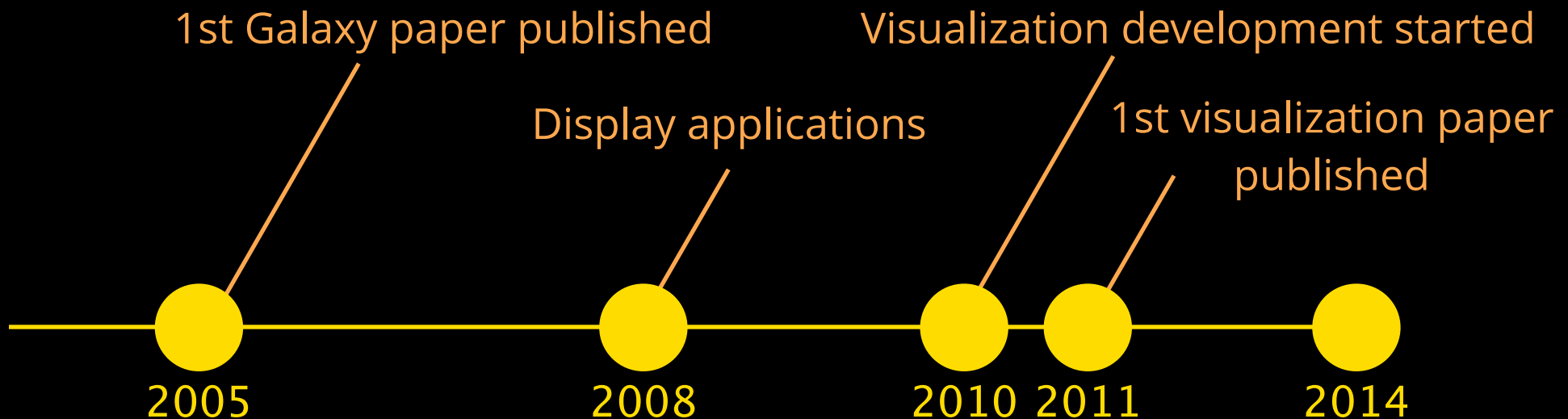


http://en.wikipedia.org/wiki/Anscombe's_quartet

Timeline of Visualization in Galaxy



Timeline of Visualization in Galaxy



- 1. *visualization in Galaxy is nascent***
- 2. *you will be working with awesome new features***
- 3. *there may be bugs — help us fix them!***

Workshop Goals

Participants: learn about how to visualize your data in Galaxy

- ✦ biological visualizations
- ✦ numerical visualizations
- ✦ what Galaxy is doing underneath the covers

Instructors: feedback from you about what you like, don't like, and where to go next

Galaxy Visualizations

Visualizations are first-class objects in Galaxy, just like tools

A visualization can be added to Galaxy via a configuration file that specifies:

- ✦ datasets that can be used
- ✦ location of visualization code (client-side or on server)

Galaxy handles visualization integration and data management, so users can focus on analyzing data (and developers can focus on creating visualizations)

Visualizations are 1st class Galaxy objects

Can be saved and versioned for reproducibility

Have a human-readable URL for sharing a fully interactive visualization:

`http://usegalaxy.org/u/jgoecks/v/tumor-mutations`

Can embed interactive visualizations in online supplementary materials via Galaxy Pages

Visualization

Visual Analysis



Visualization Architecture

Client-server architecture



Lots of moving pieces

- ✦ prepare/process data on server
- ✦ send to client
- ✦ render on client

Topics

Visualization history and introduction

Biological Visualizations

Numerical Visualizations

Adding your own visualizations

Analysis goal: what similarities and differences can be found in cancer cell lines using exome and transcriptome sequencing?

Sequencing and Analysis

Profiled 3 pancreatic cancer cell lines using a 26-gene panel targeting known oncogenic driver mutations

- ✦ MiaPaCa2, HPAC, and PANC-1

Data available:

- ✦ exome: mapped reads, removed dups, called variants
- ✦ transcriptome: mapped reads, assembled transcripts, computed expression

Display Applications

11: MiaPaCa2: Varscan v   

ariants

~150,000 lines

format: **vcf**, database: **hg19**

Log: tool progress
Log: tool progress
Picked up _JAVA_OPTIONS: -
Djava.io.tmpdir=/tmp
Got the following sample list:
MiaPaCa2-exome
Only variants will be reported
Min coverage: 8
Min reads2: 2
Min var freq: 0.01
Min avg qual: 15
P-value thresh: 0.99
Read

display at UCSC [main](#)
display with IGV [web](#) [current](#) [local](#)
display at RViewer [main](#)

Display Applications

11: MiaPaCa2: Varscan v   

ariants

~150,000 lines
format: **vcf**, database: **hg19**

Log: tool progress
Log: tool progress
Picked up _JAVA_OPTIONS: -
Djava.io.tmpdir=/tmp
Got the following sample list:
MiaPaCa2-exome
Only variants will be reported
Min coverage: 8
Min reads2: 2
Min var freq: 0.01
Min avg qual: 15
P-value thresh: 0.99
Read

display at UCSC main
display at web current local
display at RViewer main

Display Applications

Advantages

- ✦ use tool that is familiar to you
- ✦ easy to view your data alongside public datasets

Disadvantages

- ✦ cannot save/share/version visualization
- ✦ many more visualizations than display applications in Galaxy
- ✦ no data processing or visual analysis, only visualization

Trackster—Galaxy's Genome Browser



Trackster—Galaxy's Genome Browser

Genome browsers remain a (the most?) powerful genome visualization

- ✦ foundational tool

Trackster is for the high-throughput sequencing era

- ✦ very large datasets, numerous simultaneous tracks
- ✦ maximum flexibility for customization (e.g. rainbow tracks)
- ✦ 2-3 indices per dataset for fast visualization

SAM/BAM, BED, GFF/GTF, VCF, Wiggle, BigWig, BigBed, BedGraph

Let's visualize our data in Trackster

1. Create visualization
2. Add gene annotation (RefSeq)
3. Save visualization
4. Exit
5. Reopen visualization

Let's visualize our data in Trackster

1. Create visualization

Let's visualize our data in Trackster

1. Create visualization
2. Add gene annotation (RefSeq)

Let's visualize our data in Trackster

1. Create visualization
2. Add gene annotation (RefSeq)
3. Save visualization

Let's visualize our data in Trackster

1. Create visualization
2. Add gene annotation (RefSeq)
3. Save visualization
4. Exit

Let's visualize our data in Trackster

1. Create visualization
2. Add gene annotation (RefSeq)
3. Save visualization
4. Exit
5. Reopen visualization

Behind the Scenes

Galaxy is indexing datasets for

- ✦ viewing large genomic regions (coverage plots)
- ✦ viewing small genomic regions (getting individual data points)
- ✦ feature names and locations

Indexes is the primary way that big datasets are visualized quickly

Modes and Searching

Tracks can be displayed differently

- ✦ coverage
- ✦ individual features

Let's try different modes

- ✦ this is fast because data is sent from Galaxy server and rendered in your Web browser

Let's try searching for a gene: ERBB2

Let's Call Variants

VarScan

- ✦ Sample names: MiaPaCa2, PANC1, HPAC
- ✦ Run

Rename output: "Cell line variants"

Let's Assemble Transcripts

Cufflinks

- ✦ input dataset is #7
- ✦ run

Rename output: "MiaPaCa2
Assembled Transcripts"

Let's add data to Trackster

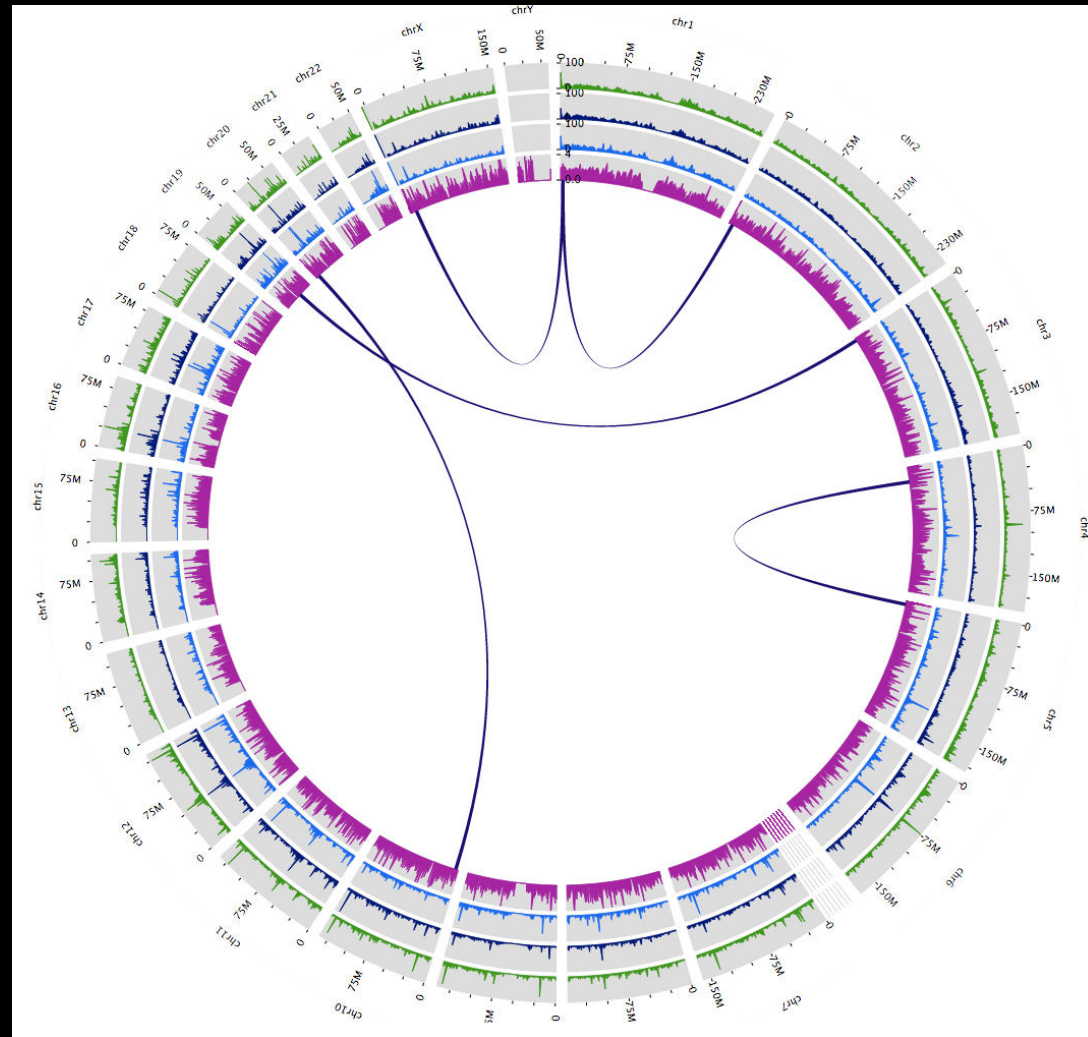
Add exome data for all cell lines...

...but where is our data?

Circster

Interactive Circos
plot

Whole genome
view with structural
variation



Let's view our data in Circster

Double-click or use trackpad to zoom
in

change track min/max

what do we see?

Let's add data to Circster and adjust options

1. Add transcriptome coverage data

Let's add data to Circster and adjust options

1. Add transcriptome coverage data
2. Change arc dataset height

Let's add data to Circster and adjust options

1. Add transcriptome coverage data
2. Change arc dataset height
3. Change max for tracks

Let's add data to Circster and adjust options

1. Add transcriptome coverage data
2. Change arc dataset height
3. Change max for tracks
4. Save visualization

Back to Trackster: Rainbow Track for Coverage

1. Navigate to ERBB2 gene
2. Create group
3. Add transcriptome coverage tracks to group
4. Create composite track
5. Adjust max
6. *what do we see?*

Add More Data!

Add RNA-seq mapped reads, variants, and assembled transcripts

Look at ERBB2

- ✦ bookmark!

Look at STK11

- ✦ bookmark!

Look at KRAS —> LYRM5

- ✦ bookmark!

Visual Analysis

KRAS and Variants

Sweepster

Topics

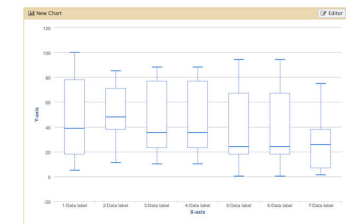
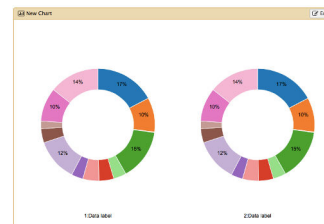
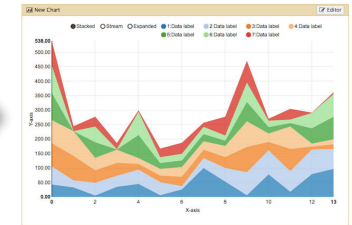
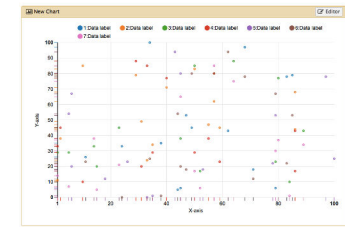
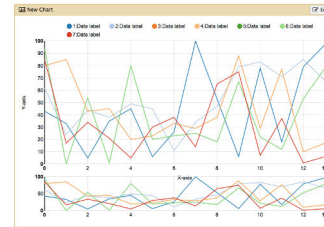
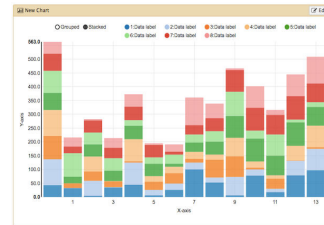
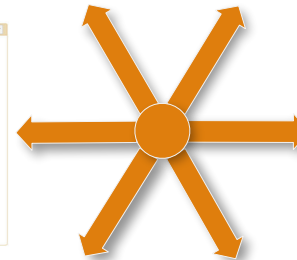
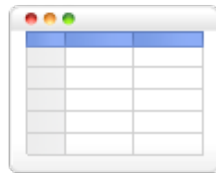
Visualization history and introduction

Biological Visualizations

Numerical Visualizations

Adding your own visualizations

What is Galaxy Charts?



Use Galaxy

Create Tabular Results

Visualize with Galaxy Charts

Import data files

Galaxy Analyze Data W **1** **Shared Data** Visualization Help User Using 2.1

Data Library "Charts"

3 ☐ Name

☐ amino_acid_features.txt

☐ http://www.compsysbio.org/bacteriome/dataset/functional_i

For selected datasets: **4**

2

- Data Libraries
- Data Libraries Beta
- Published Histories
- Published Workflows
- Published Visualizations
- Published Pages

Data type	Date uploaded	File size
tabular	Mon Jun 30 04:13:31 2014 (UTC)	974 bytes
tabular	Mon Jun 30 16:33:33 2014 (UTC)	81.6 KB

i TIP: You can download individual library datasets by selecting "Download this dataset" from the context menu (triangle) next to each dataset's name.

i TIP: Several compression options are available for downloading multiple library datasets simultaneously:

- gzip: Recommended for fast network connections
- bzip2: Recommended for slower network connections (smaller size but takes longer to compress)
- zip: Not recommended but is provided as an option for those who cannot open the above formats

Click on **Shared Data** and select **Data Libraries**. Navigate to the **Chart** library and import it into your history (*data reference: <http://dna.cs.byu.edu/treesaap> and bacteriome.org*).

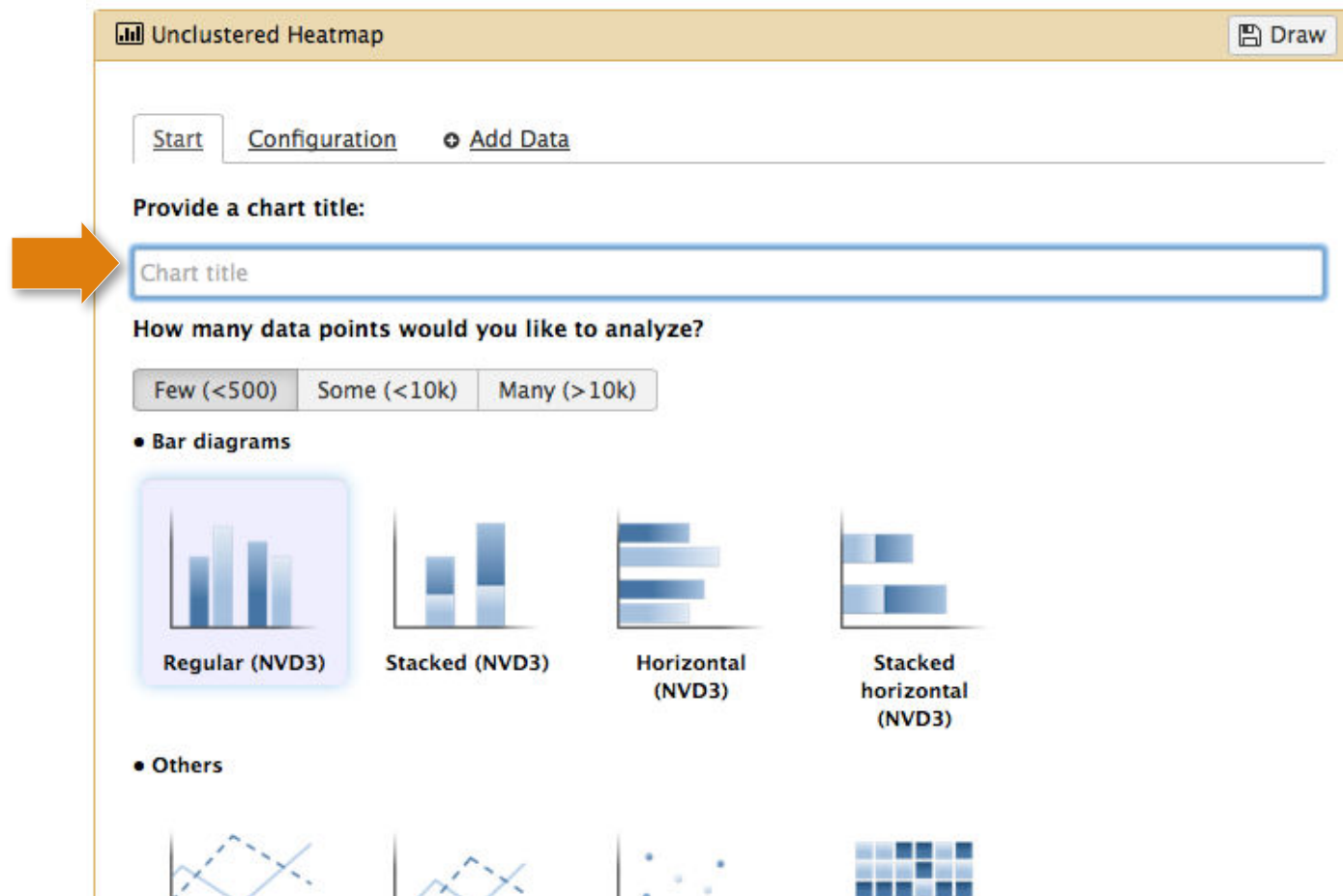
Make a new chart (1 of 4)

The screenshot shows a web interface for uploading and visualizing data. An orange arrow labeled '1' points to the URL bar containing the text: `S1: http://www.compsy`. Below the URL bar, the text `sbio.org/bacteriome/dataset/functional_interactions.txt` is displayed. Further down, it shows `3,989 lines` and `format: tabular, database: ?`. A green button labeled `uploaded tabular file` is visible. Below this, a row of icons includes a save icon, an information icon, a refresh icon, and a visualization icon. An orange arrow labeled '2' points to the visualization icon. A dropdown menu is open, showing the options `Charts`, `Scatterplot`, and `Trackster`. An orange arrow labeled '3' points to the `Charts` option. Below the dropdown, a table of data is visible with columns for gene IDs and a numerical value.

B4200	B4202	0.933934
B0779	B4058	0.933934
B0032	B0033	0.933183

Wait for the upload to complete. Select your **Dataset** and click on the **Visualization Icon** then select **Charts**.

Give your chart a name



Unclustered Heatmap Draw

[Start](#) [Configuration](#) [Add Data](#)

Provide a chart title:

Chart title

How many data points would you like to analyze?

Few (<500) Some (<10k) Many (>10k)

• Bar diagrams

Regular (NVD3) Stacked (NVD3) Horizontal (NVD3) Stacked horizontal (NVD3)

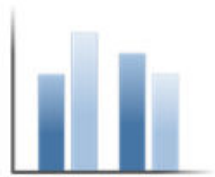
• Others

Name your chart **Unclustered Heatmap**.

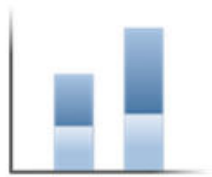
Select a chart type

Few (<500) Some (<10k) Many (>10k)

• Bar diagrams



Regular (NVD3)



Stacked (NVD3)

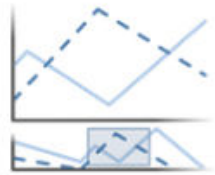


Horizontal
(NVD3)

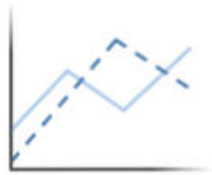


Stacked
horizontal
(NVD3)

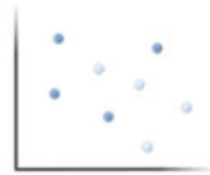
• Others



Line with focus
(NVD3)



Line chart
(NVD3)



Scatter plot
(NVD3)



Heatmap
(Custom)



Double click on the **Heatmap** icon.

Select data columns

New Chart Cancel Draw

Start Configuration 1: Data label + Add Data

Provide a label:

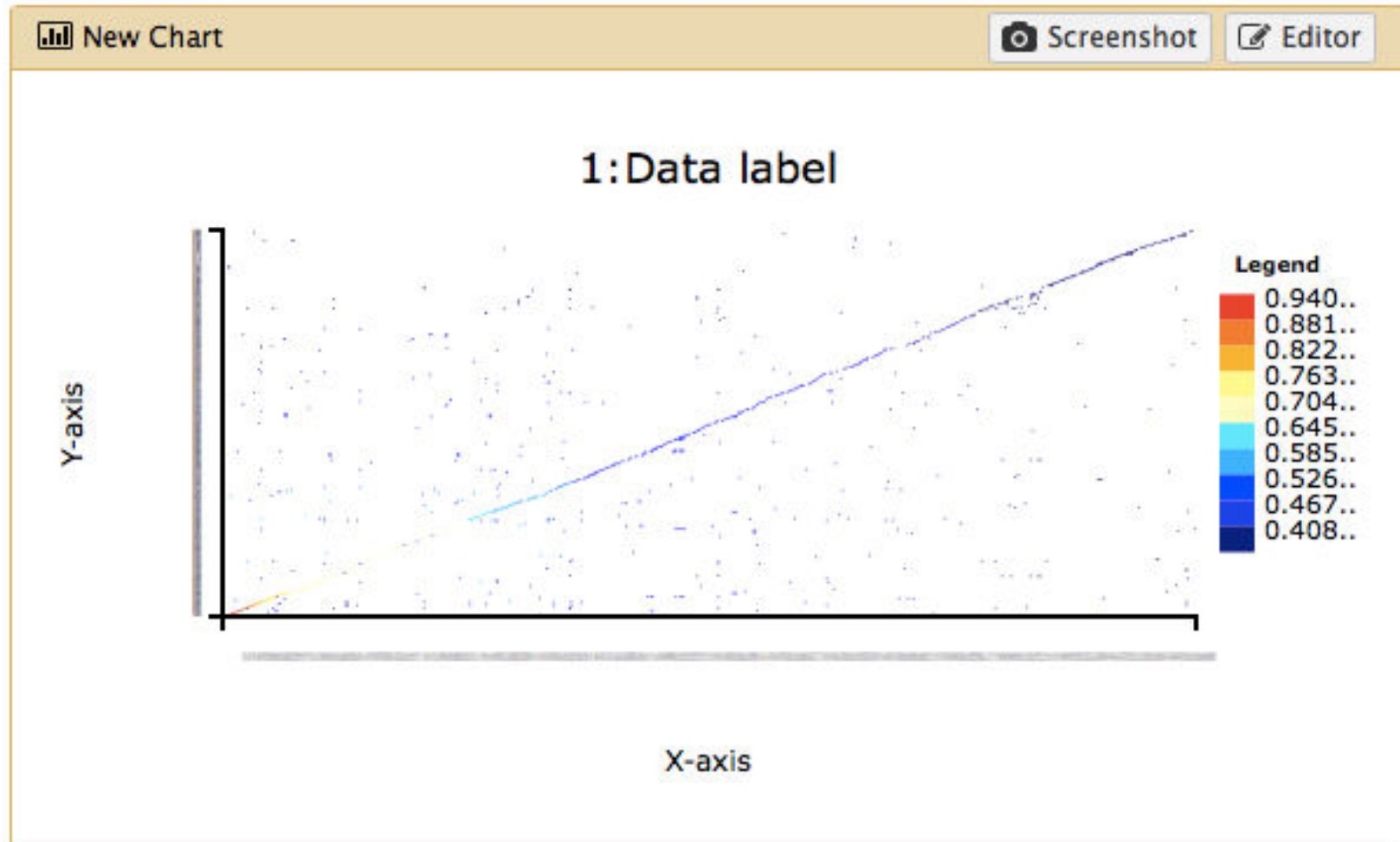
Data label

Select columns:

Column labels (all data labels)	Column: 1 [str]	▼
Row labels (all data labels)	Column: 2 [str]	▼
Observation	Column: 3 [float]	▼

At first click on **Row labels** and select **Column 2**. Then, click on **Draw**.

Unclustered Heatmap



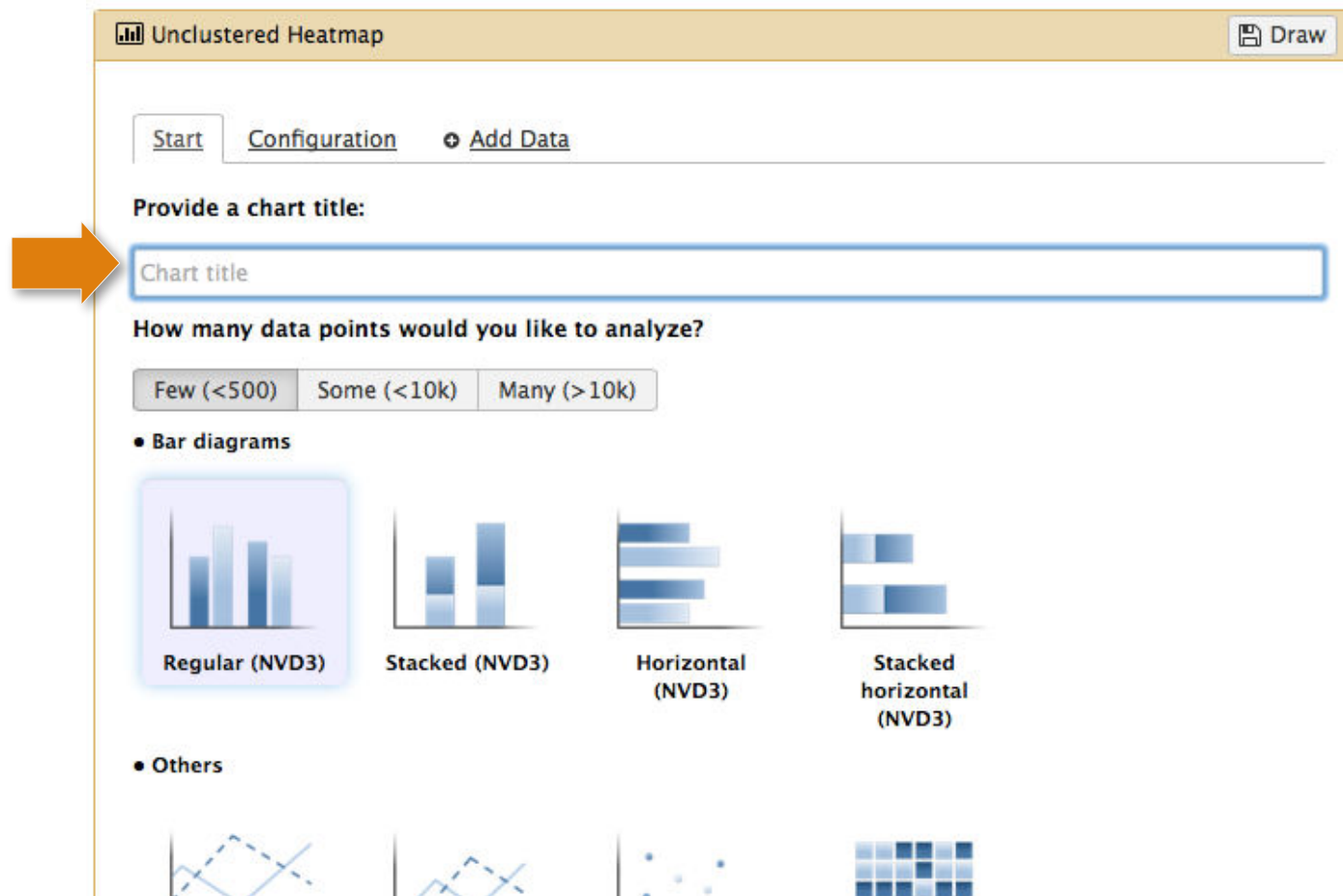
Make a new chart (2 of 4)

The screenshot shows a web interface with a green background. An orange arrow labeled '1' points to the URL bar containing the text: `S1: http://www.compsy`. Below the URL bar, the text `sbio.org/bacteriome/dataset/functional_interactions.txt` is displayed. Underneath, it says '3,989 lines' and 'format: tabular, database: ?'. A light green box contains the text 'uploaded tabular file'. Below this, there is a row of icons: a save icon, an information icon, a refresh icon, and a visualization icon. An orange arrow labeled '2' points to the visualization icon. A dropdown menu is open, showing the options 'Charts', 'Scatterplot', and 'Trackster'. An orange arrow labeled '3' points to the 'Charts' option. Below the dropdown menu, a table of data is visible with columns for gene IDs and a numerical value.

B4200	B4202	0.933934
B0779	B4058	0.933934
B0032	B0033	0.933183

Select your **Dataset** and click on the **Visualization Icon** then select **Charts**.

Give your chart a name



Unclustered Heatmap Draw

[Start](#) [Configuration](#) [Add Data](#)

Provide a chart title:

Chart title

How many data points would you like to analyze?

Few (<500) Some (<10k) Many (>10k)

• Bar diagrams

Regular (NVD3) Stacked (NVD3) Horizontal (NVD3) Stacked horizontal (NVD3)

• Others

Name your chart **Clustered Heatmap**.

Select a new chart type

- 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)



Double click on the **Clustered Heatmap** icon.

Select data columns

New Chart [Cancel] [Draw]

Start Configuration 1: Data label [Add Data]

Provide a label:

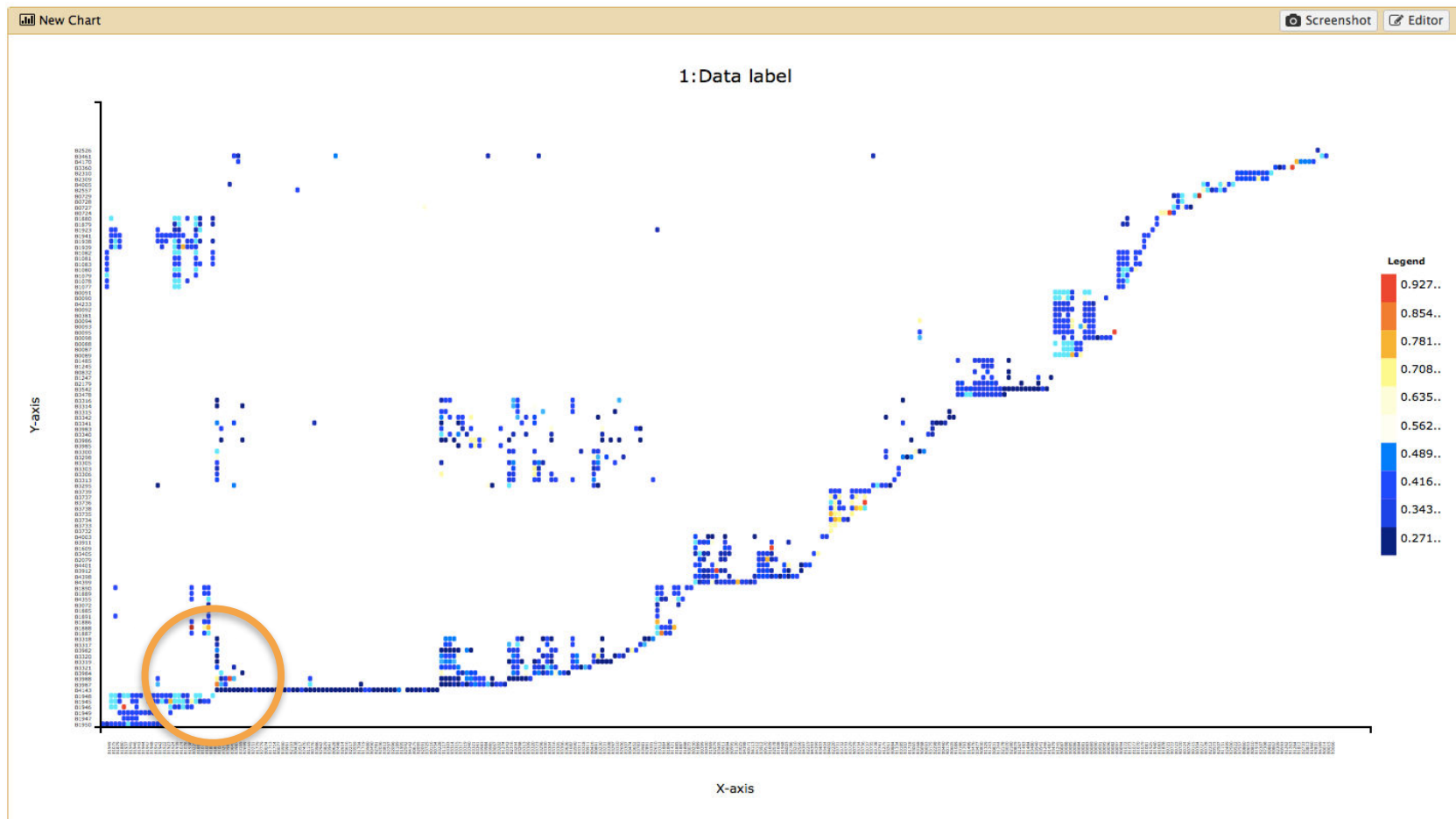
Data label

Select columns:

Column labels (all data labels)	Column: 1 [str]	▼
Row labels (all data labels)	Column: 2 [str]	▼
Observation	Column: 3 [float]	▼

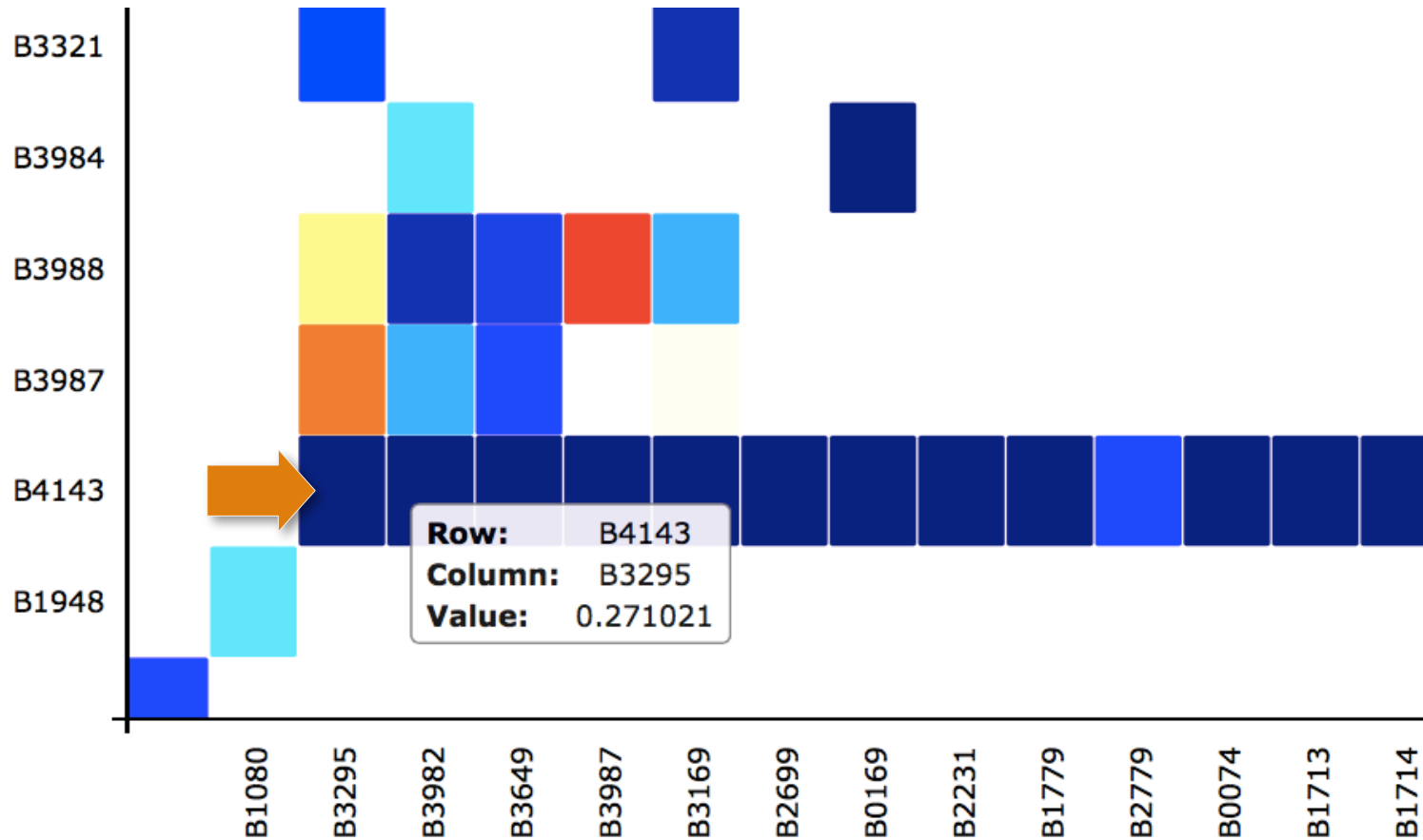
At first click on **Row labels** and select **Column 2**. Then, click on **Draw**.

Clustered Heatmap



Use the **mouse wheel** or **your touch pad** to zoom into the **highlighted area**.

Enlarged view



Tooltips popup if you move the mouse pointer over a box. Here the interaction between **B4143** and **B3295** is highlighted. Click on **Editor** again to further customize this chart.

Chart configuration

 New Chart

Start

Configuration

1: Data label 

 Add Data

Provide a chart title:

New Chart

How many data points would you like to analyze?

Few (<500)Some (<10k)Many (>10k)

• Bar diagrams



Go to the **Configuration** tab.

Chart settings

X axis:

Axis label

X-axis

Provide a label for the axis.

Axis value type

Auto

Select the value type of the axis.

Y axis:

Axis label

Y-axis

Provide a label for the axis.

Axis value type

Auto

Select the value type of the axis.

Others:

Show legend

Yes No

Would you like to add a legend?

Color scheme

Jet

Select a color scheme for your heatmap

Url template

http://someurl.com?id=__LABEL__

Enter a url to link the labels with external sources. Use __LABEL__ as placeholder.



Heatmap specific options are **highlighted**. Feel free to set **axis labels** or other options.

Define a URL template

Yes

No

Would you like to add a legend?

Jet

Select a color scheme for your heatmap

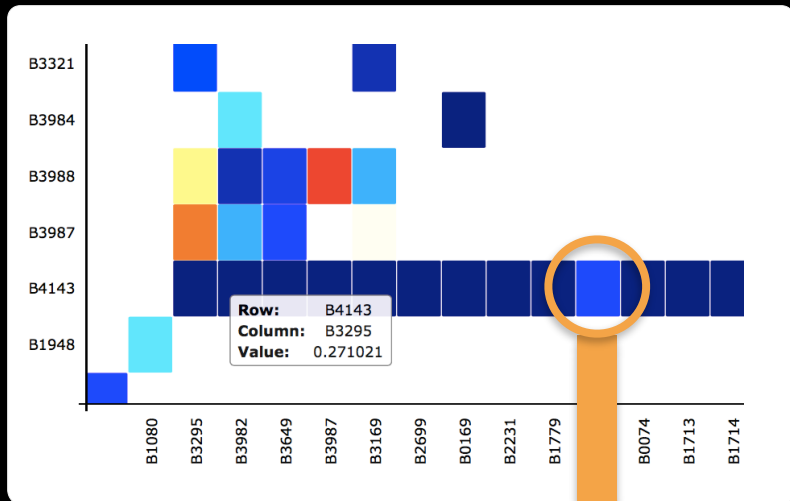


http://www.ncbi.nlm.nih.gov/geoprofiles/?term=__LABEL__

Enter a url to link the labels with external sources. Use __LABEL__ as placeholder

Paste a **database URL** into the template URL field and add the **__LABEL__** tag. You may use **<http://www.ncbi.nlm.nih.gov>** or any other database. Click on **Draw** to redraw the chart.

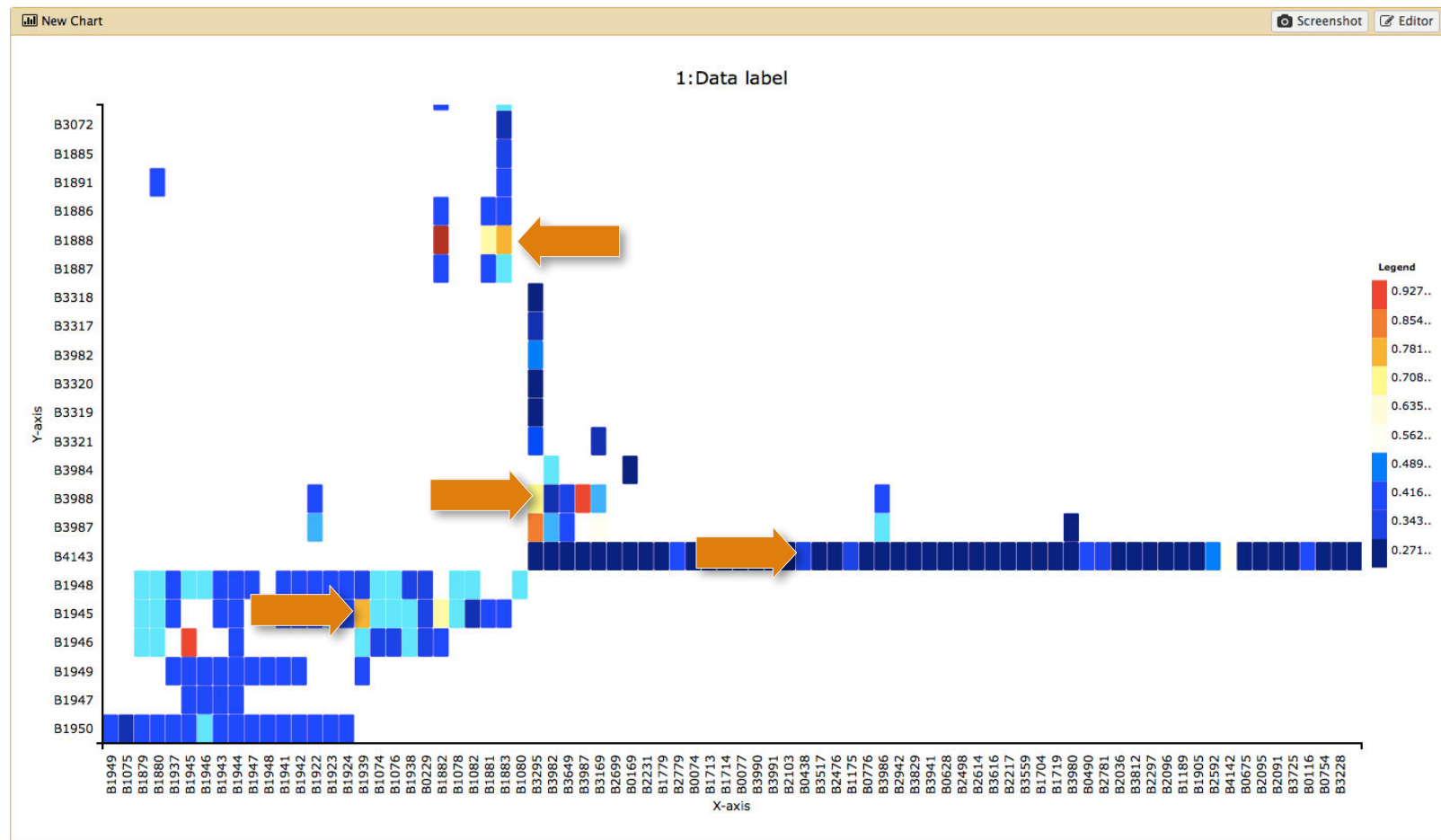
Data points linked to web sources



Double click on a **box** and the browser will open two new tabs using the previously defined **URL template**.

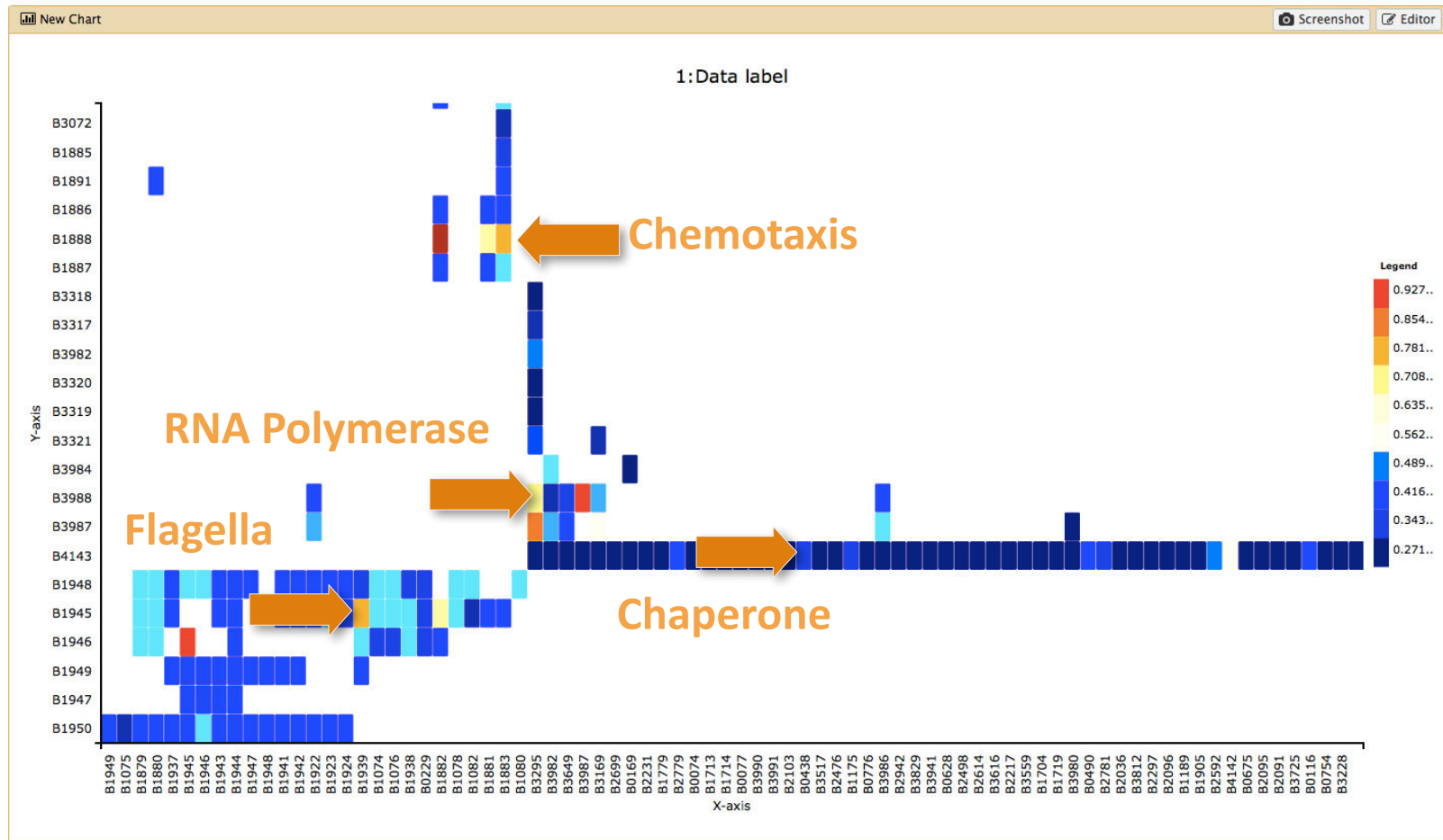
Two overlapping browser screenshots showing the NCBI GEO Profiles page for gene B4143. The top screenshot shows the search results for B4143, including the gene name "groEL - Stress factor RpoS regulon in exponential-phase bacteria" and its expression profile. The bottom screenshot shows the search results for B4143, including the gene name "groL - Indole-3-acetic acid effect on Escherichia coli" and its expression profile.

Cluster selection and analysis



Select one element from each **highlighted row**. What are the corresponding **protein functions**?

Identified protein categories



Please return to the **Editor**.

Make a new chart (3 of 4)

1 → S1: http://www.compsysbio.org/bacteriome/dataset/functional_interactions.txt [eye] [pencil] [x]

3,989 lines
format: **tabular**, database: ?

uploaded tabular file

[save] [info] [refresh] [chart] 2 →

3 → Charts

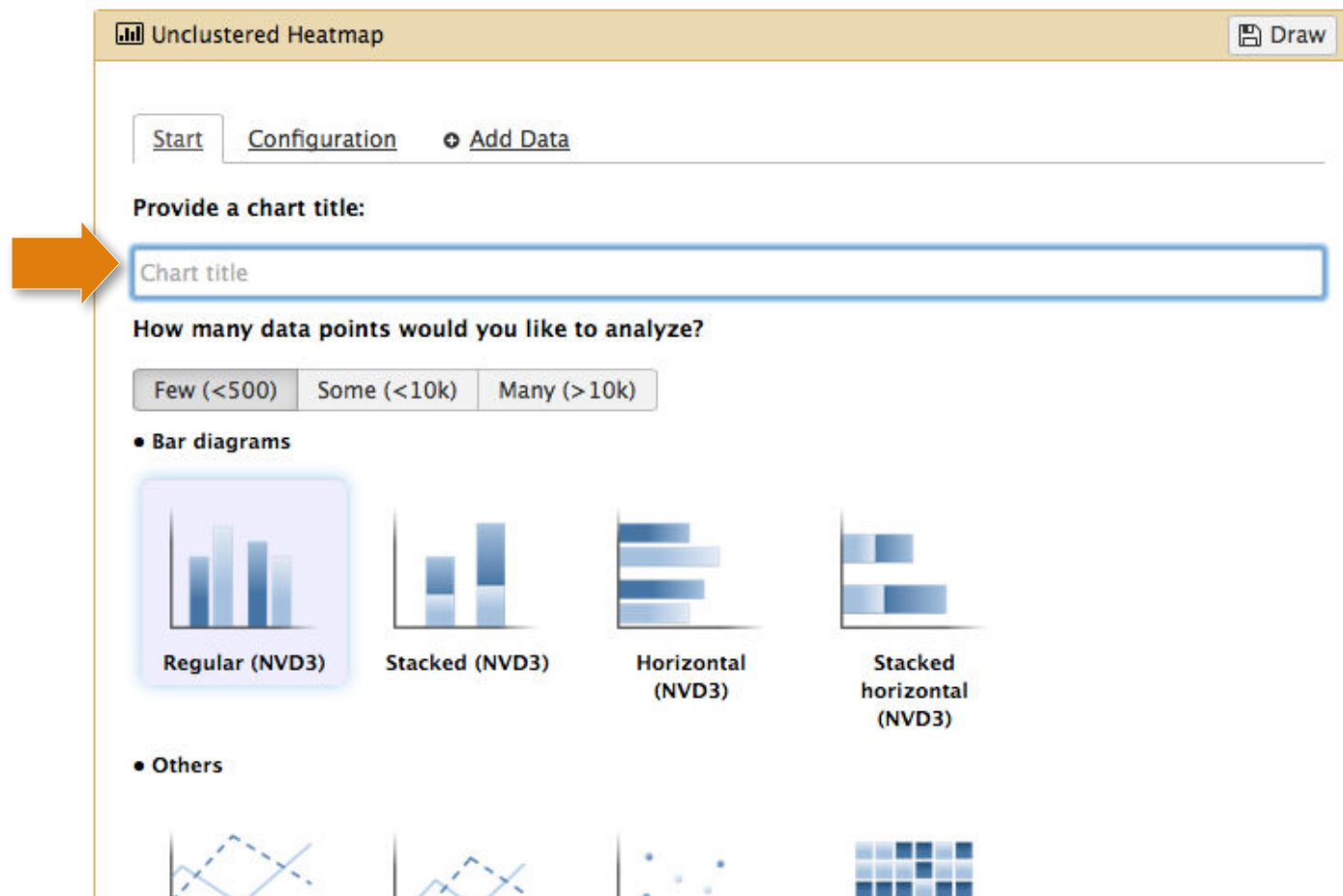
- Scatterplot
- Trackster

B4200 B4202 0.933934
B0779 B4058 0.933934
B0032 B0033 0.933183

The screenshot shows a web interface for uploading and visualizing a dataset. An orange arrow labeled '1' points to the URL input field. The URL is 'http://www.compsysbio.org/bacteriome/dataset/functional_interactions.txt'. Below the URL, it shows '3,989 lines' and 'format: tabular, database: ?'. A green box highlights the 'uploaded tabular file' section. Below this, there is a toolbar with icons for save, info, refresh, and chart. An orange arrow labeled '2' points to the chart icon. A dropdown menu is open, showing 'Charts' as the selected option, with 'Scatterplot' and 'Trackster' as sub-options. An orange arrow labeled '3' points to the 'Charts' option in the dropdown menu. Below the dropdown, a table of data is visible, showing columns for gene IDs and a numerical value.

Select your **Dataset** and click on the **Visualization Icon** then select **Charts**.

Give your chart a name



Unclustered Heatmap Draw

[Start](#) [Configuration](#) [Add Data](#)

Provide a chart title:

Chart title

How many data points would you like to analyze?

Few (<500) Some (<10k) Many (>10k)

• Bar diagrams

Regular (NVD3) Stacked (NVD3) Horizontal (NVD3) Stacked horizontal (NVD3)

• Others

Name your chart **Score Histogram**.

Analyze the score distribution

- Area charts



QRegular (NVD3)



QExpanded (NVD3)



QStream (NVD3)



Pie chart (NVD3)

- Data processing (requires 'charts' tool from Toolshed)



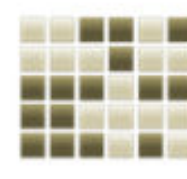
Histogram (NVD3)



QDiscrete Histogram (jqPlot)



QBox plot (jqPlot)



QClustered Heatmap (Custom)

Double click on the **Histogram** icon and click on **Draw**.

Give your chart a name

 Unclustered Heatmap   Draw

Please select data columns before drawing the chart.

Start Configuration 1: Data label   Add Data

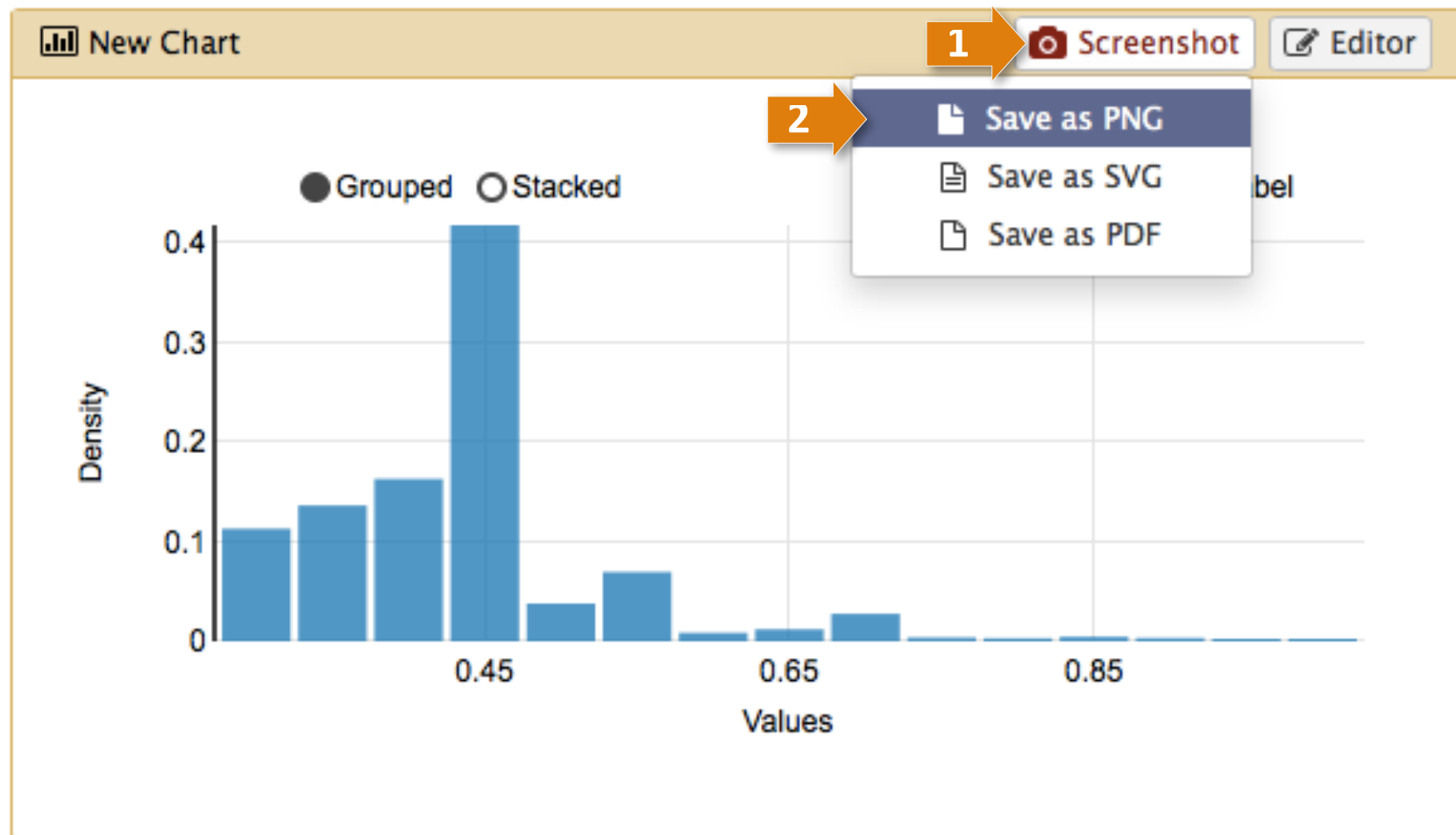
Provide a label:

Select columns:

Observations 

Click on **Draw**.

Export as PNG



Click on **Screenshot** and select **Save as PNG**. Finally, return to the **Editor** again.

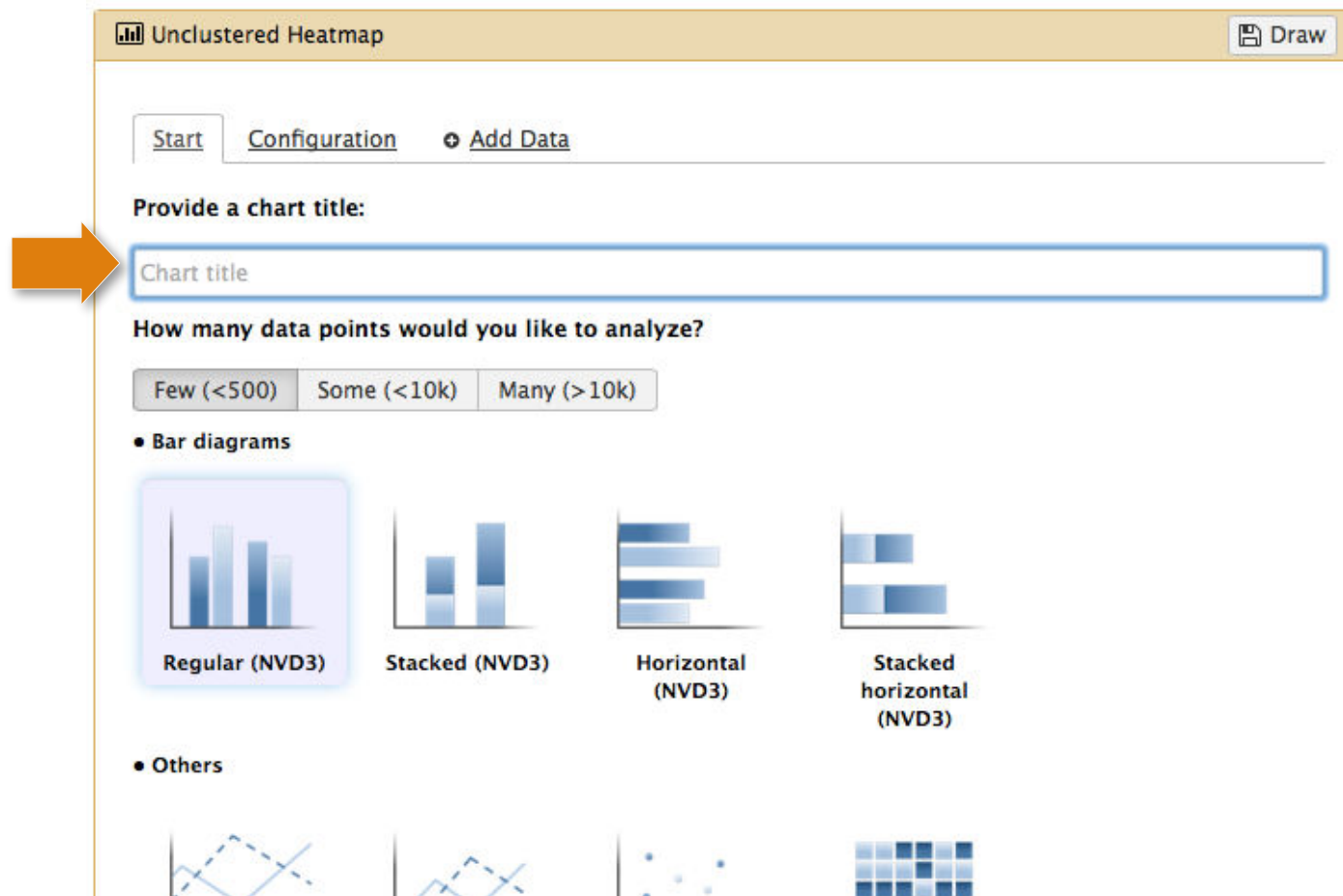
Make a new chart (4 of 4)

The screenshot shows a web interface with a green background. At the top, an orange arrow labeled '1' points to the URL: http://www.compsysbio.org/bacteriome/dataset/functional_interactions.txt. Below the URL, it says '3,989 lines' and 'format: tabular, database: ?'. A text box contains 'uploaded tabular file'. Below this is a toolbar with icons for save, info, refresh, and visualization. An orange arrow labeled '2' points to the visualization icon. A dropdown menu is open, showing 'Charts' (highlighted), 'Scatterplot', and 'Trackster'. An orange arrow labeled '3' points to the 'Charts' option. Below the menu, a table of data is visible:

B4200	B4202	0.933934
B0779	B4058	0.933934
B0032	B0033	0.933183

Select your **Dataset** and click on the **Visualization Icon** then select **Charts**.

Give your chart a name



Unclustered Heatmap Draw

[Start](#) [Configuration](#) [Add Data](#)

Provide a chart title:

Chart title

How many data points would you like to analyze?

Few (<500) Some (<10k) Many (>10k)

• Bar diagrams

Regular (NVD3) Stacked (NVD3) Horizontal (NVD3) Stacked horizontal (NVD3)

• Others

Name your chart **Discrete Histogram**.

Analyze the protein distribution

- Area charts



ⓈRegular (NVD3)



ⓈExpanded (NVD3)



ⓈStream (NVD3)



Pie chart (NVD3)

- Data processing (requires 'charts' tool from Toolshed)



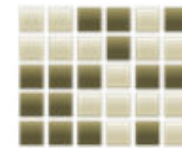
Histogram (NVD3)



ⓈDiscrete Histogram (jqPlot)



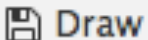
ⓈBox plot (jqPlot)



ⓈClustered Heatmap (Custom)

Double click on the **Discrete Histogram** icon.

Add more data

 Unclustered Heatmap 

[Start](#) [Configuration](#) [1: Data label](#)  [+ Add Data](#) 

Provide a label:

Select columns:

Observations



Click on **Add Data**.

Select a second data group

The screenshot shows a 'New Chart' dialog box with a yellow header bar. The header bar contains a bar chart icon, the text 'New Chart', and two buttons: 'Cancel' and 'Draw'. An orange arrow with the number '2' points to the 'Draw' button. Below the header bar, there are four tabs: 'Start', 'Configuration', '1: Data label', and '2: Data label'. The '2: Data label' tab is selected. To the right of the tabs is a button labeled '+ Add Data'. Below the tabs, there is a section titled 'Provide a label:' with a text input field containing the text 'Data label'. Below this, there is a section titled 'Select columns:' with a list box. The list box has a blue background and contains the text 'Observations' and 'Column: 2 [str]'. An orange arrow with the number '1' points to the 'Observations' text. To the right of the list box is a dropdown arrow.

New Chart [Cancel] [Draw]

Start Configuration 1: Data label 2: Data label + Add Data

Provide a label:

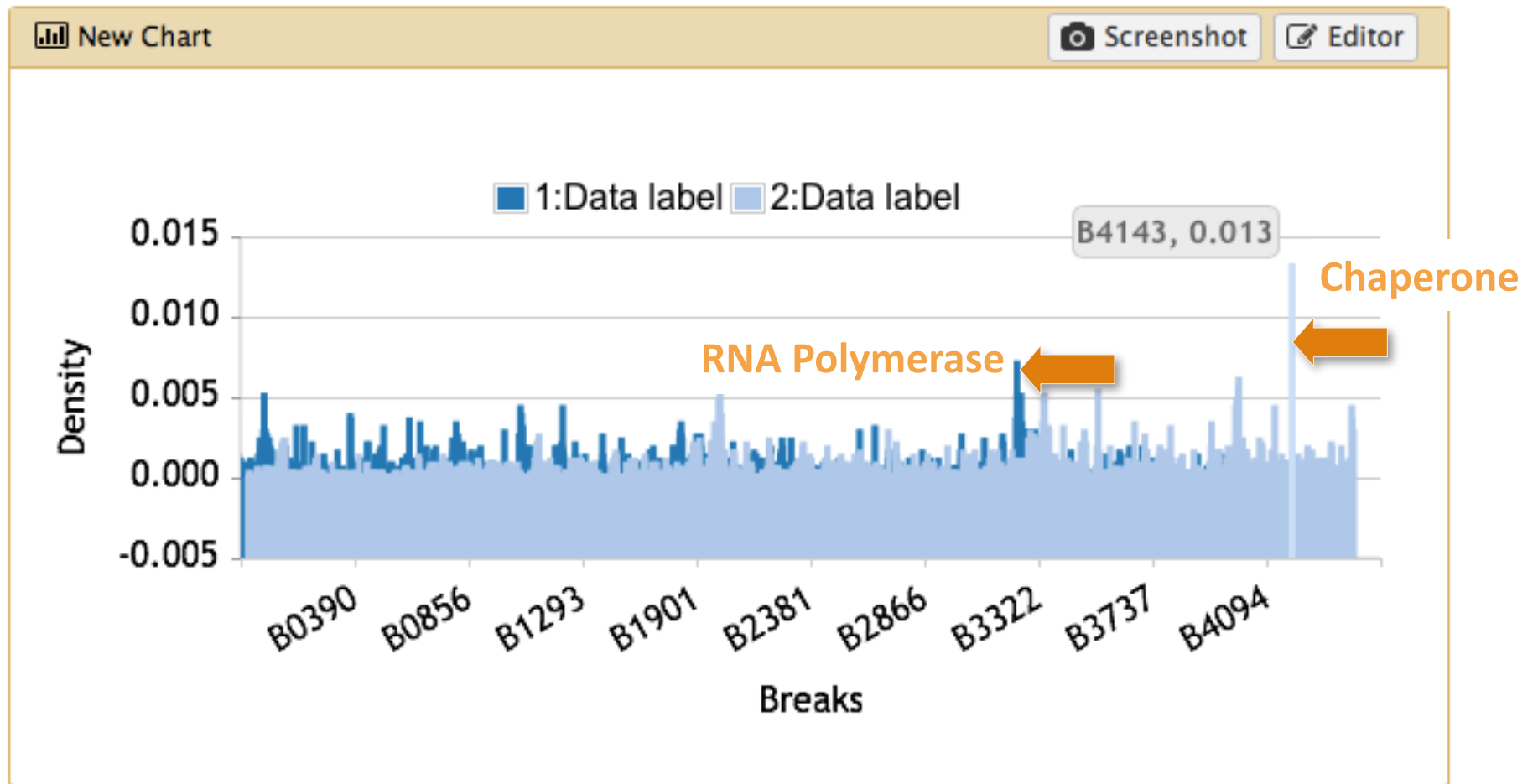
Data label

Select columns:

Observations Column: 2 [str]

At first click on **Observations** and select **Column 2**. Then, click on **Draw**.

Which proteins have most interactions?



Done with Part I.

Scratchbook

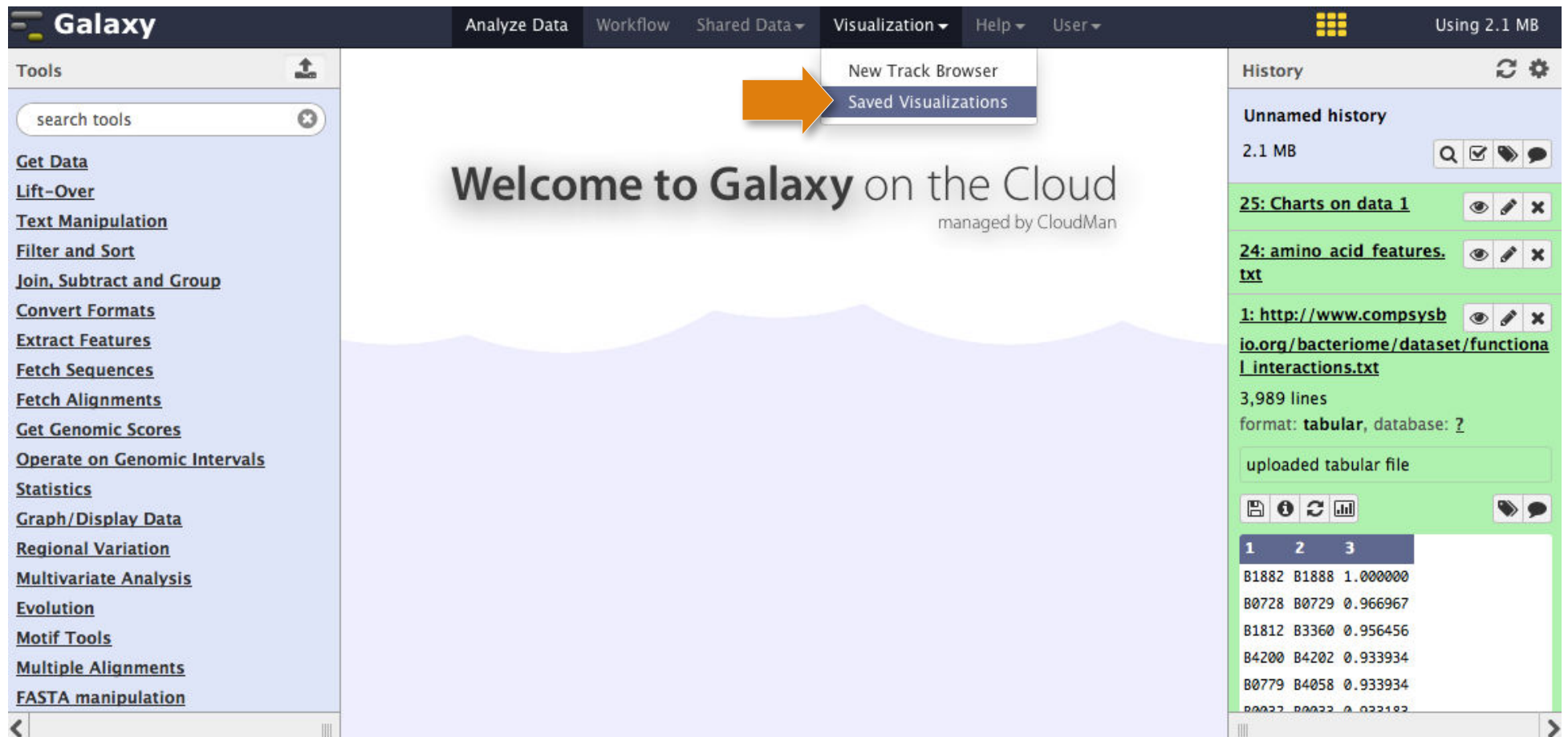
Activate the Scratchbook

The screenshot shows the Galaxy web interface. The top navigation bar includes links for Analyze Data, Workflow, Shared Data, Visualization, Help, and User. On the right side of the navigation bar, there is a grid icon representing the Scratchbook, which is highlighted by an orange arrow. The main content area displays a welcome message: "Welcome to Galaxy on the Cloud" managed by CloudMan. The left sidebar contains a list of tools categorized by function, such as Get Data, Text Manipulation, and Statistics. The right sidebar shows the History panel, which lists recent jobs. The most recent job is titled "25: Charts on data 1" and shows a preview of a tabular file with 3,989 lines. Below the preview, a table of data is visible.

1	2	3
B1882	B1888	1.000000
B0728	B0729	0.966967
B1812	B3360	0.956456
B4200	B4202	0.933934
B0779	B4058	0.933934
B0023	B0023	0.023182

Activate the **Scratchbook** by clicking on the above icon.

Activate the Scratchbook



The screenshot shows the Galaxy web interface. The top navigation bar includes 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The 'Visualization' menu is open, showing 'New Track Browser' and 'Saved Visualizations'. An orange arrow points to 'Saved Visualizations'. The left sidebar lists various tools under the 'Tools' section. The main content area displays 'Welcome to Galaxy on the Cloud' managed by CloudMan. The right sidebar shows the 'History' section with a list of recent data items, including '25: Charts on data 1', '24: amino acid features.txt', and '1: http://www.compsysbio.org/bacteriome/dataset/functional_interactions.txt'.

Galaxy | Analyze Data | Workflow | Shared Data | Visualization | Help | User | Using 2.1 MB

Tools

- search tools
- Get Data
- Lift-Over
- Text Manipulation
- Filter and Sort
- Join, Subtract and Group
- Convert Formats
- Extract Features
- Fetch Sequences
- Fetch Alignments
- Get Genomic Scores
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Regional Variation
- Multivariate Analysis
- Evolution
- Motif Tools
- Multiple Alignments
- FASTA manipulation

Welcome to Galaxy on the Cloud
managed by CloudMan

Visualization

- New Track Browser
- Saved Visualizations

History

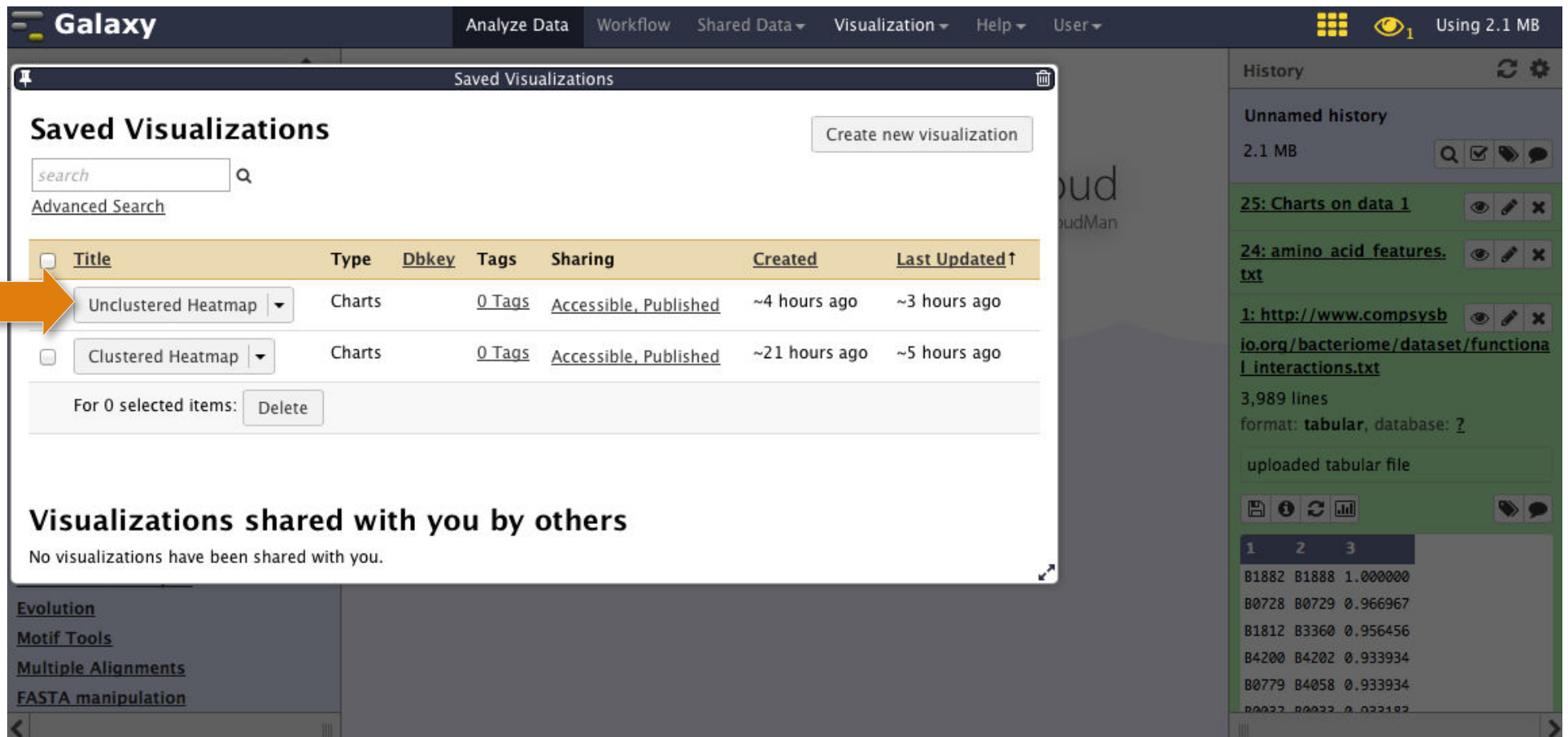
Unnamed history
2.1 MB

- 25: Charts on data 1
- 24: amino acid features.txt
- 1: http://www.compsysbio.org/bacteriome/dataset/functional_interactions.txt
3,989 lines
format: tabular, database: ?
uploaded tabular file

1	2	3
B1882	B1888	1.000000
B0728	B0729	0.966967
B1812	B3360	0.956456
B4200	B4202	0.933934
B0779	B4058	0.933934
B0022	B0022	0.022182

Click on **Saved Visualizations**.

Activate the Scratchbook



The screenshot shows the Galaxy web interface. The top navigation bar includes 'Galaxy', 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'User'. The 'Saved Visualizations' panel is open, displaying a table of saved visualizations. An orange arrow points to the 'Unclustered Heatmap' visualization. The table has columns for Title, Type, Dbkey, Tags, Sharing, Created, and Last Updated. Below the table, there is a section for 'Visualizations shared with you by others' which is currently empty. On the right side, the 'History' panel shows a list of recent workflows, including '25: Charts on data 1' and '24: amino acid features.txt'.

Title	Type	Dbkey	Tags	Sharing	Created	Last Updated
Unclustered Heatmap	Charts		0 Tags	Accessible, Published	~4 hours ago	~3 hours ago
Clustered Heatmap	Charts		0 Tags	Accessible, Published	~21 hours ago	~5 hours ago

For 0 selected items: [Delete](#)

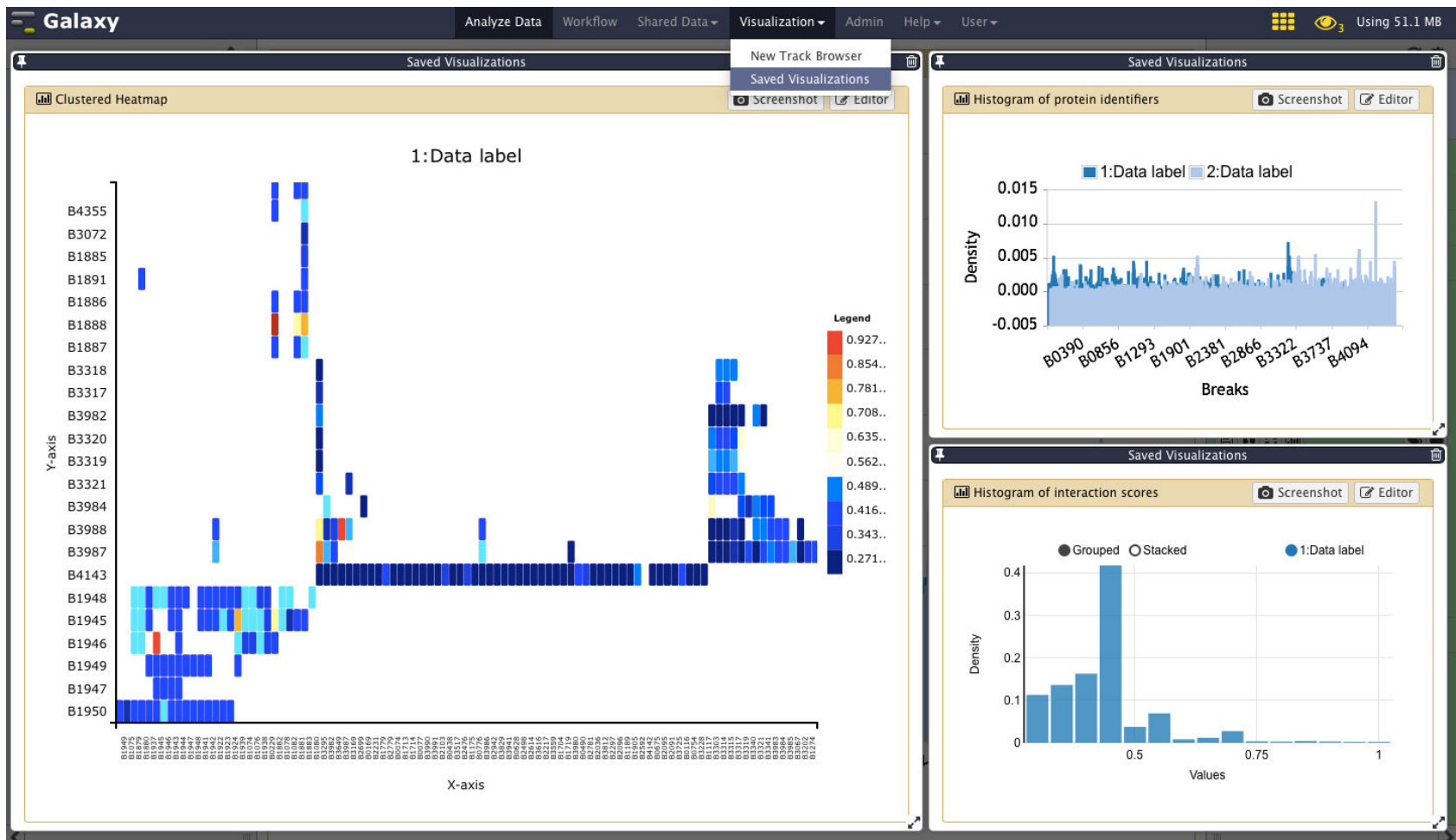
Visualizations shared with you by others
No visualizations have been shared with you.

History
Unnamed history
2.1 MB
25: Charts on data 1
24: amino acid features.txt
1: http://www.compsysbio.org/bacteriome/dataset/functionality_interactions.txt
3,989 lines
format: tabular, database: ?
uploaded tabular file

1	2	3
B1882	B1888	1.000000
B0728	B0729	0.966967
B1812	B3360	0.956456
B4200	B4202	0.933934
B0779	B4058	0.933934
B0022	B0022	0.022182

Select a Visualization and repeat the process by selecting **Saved Visualizations** again.

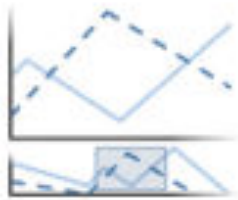
Scratchbook for multiple charts



Resize all visualizations so they fit into the screen.

More Examples

Create a pie chart



🔍 Line with focus
(NVD3)



🔍 Line chart
(NVD3)



🔍 Scatter plot
(NVD3)



🔍 Heatmap
(Custom)

• Area charts



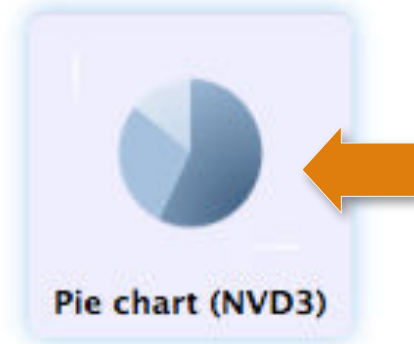
🔍 Regular (NVD3)



🔍 Expanded
(NVD3)



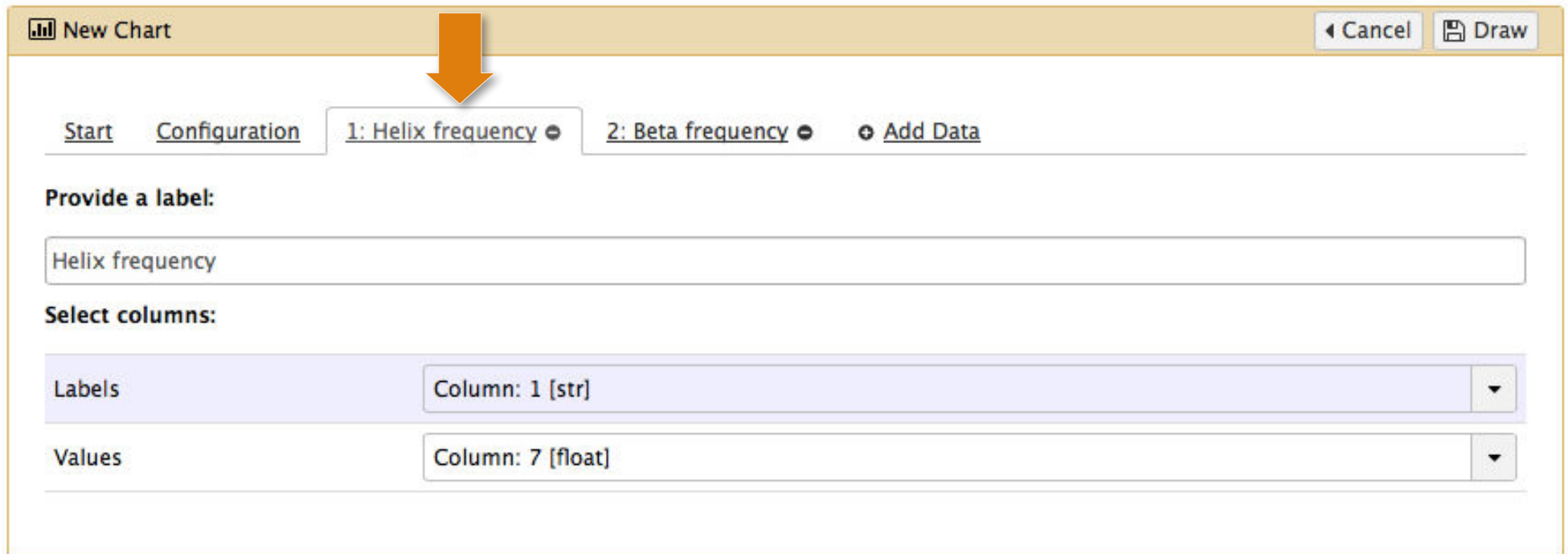
🔍 Stream (NVD3)



Pie chart (NVD3)

Select the imported datasets, create a new chart and select **Pie chart**. Then, click on **Add data**.

Add first data group



The screenshot shows a 'New Chart' dialog box with a yellow title bar. In the center, an orange arrow points down to the '1: Helix frequency' tab. The dialog has three tabs: 'Start', 'Configuration', and '1: Helix frequency'. To the right of the tabs are 'Cancel' and 'Draw' buttons. Below the tabs, there is a section 'Provide a label:' with a text input field containing 'Helix frequency'. Below that is a section 'Select columns:' with two rows. The first row is labeled 'Labels' and has a dropdown menu showing 'Column: 1 [str]'. The second row is labeled 'Values' and has a dropdown menu showing 'Column: 7 [float]'.

New Chart

Cancel Draw

Start Configuration 1: Helix frequency 2: Beta frequency Add Data

Provide a label:

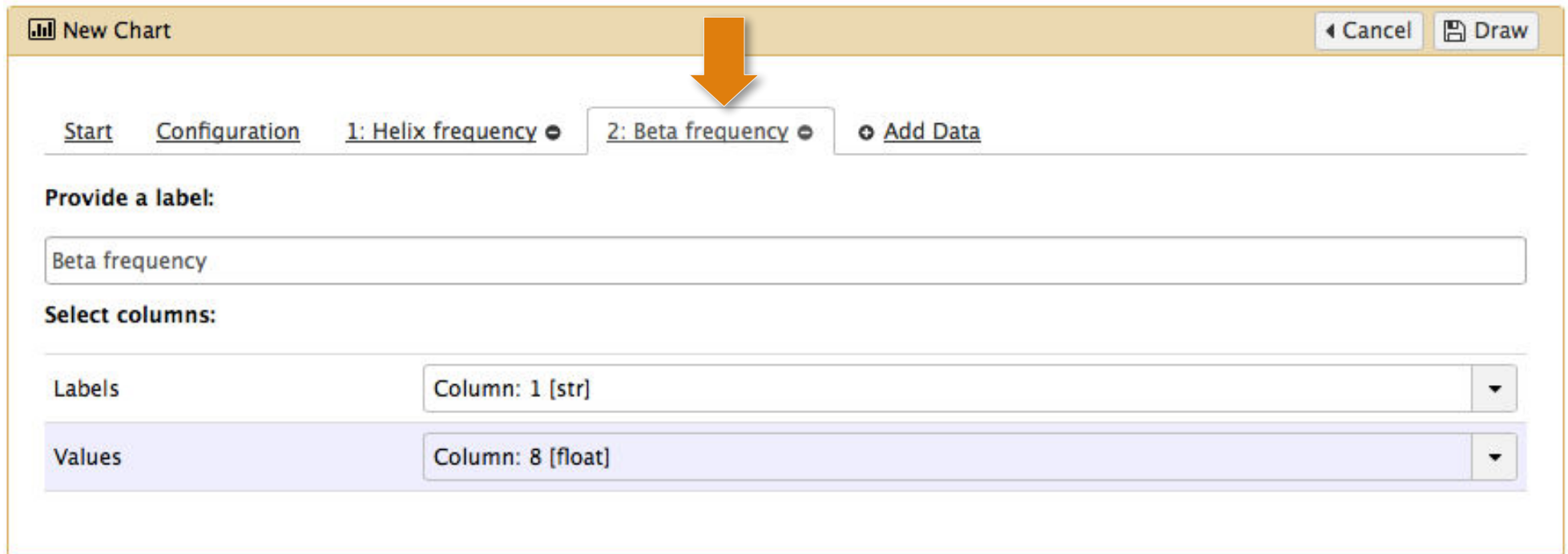
Helix frequency

Select columns:

Labels	Column: 1 [str]
Values	Column: 7 [float]

Configure the **Helix frequency** column.

Add second data group



The screenshot shows a 'New Chart' dialog box with a yellow title bar. In the top right corner are 'Cancel' and 'Draw' buttons. The main area has tabs: 'Start', 'Configuration', '1: Helix frequency', and '2: Beta frequency'. An orange arrow points to the '2: Beta frequency' tab. To the right of the tabs is an 'Add Data' button. Below the tabs, the 'Provide a label:' section has a text input field containing 'Beta frequency'. The 'Select columns:' section has two rows: 'Labels' with a dropdown menu showing 'Column: 1 [str]', and 'Values' with a dropdown menu showing 'Column: 8 [float]'. The 'Values' row is highlighted with a light blue background.

New Chart Cancel Draw

Start Configuration 1: Helix frequency **2: Beta frequency** + Add Data

Provide a label:

Beta frequency

Select columns:

Labels	Column: 1 [str]
Values	Column: 8 [float]

Configure the **Beta frequency** column.

Configure the pie chart

New Chart

Cancel Draw

Start Configuration 1: Helix frequency 2: Beta frequency Add Data

Pie chart settings:

Donut ratio 50% Determine how large the donut hole will be.

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

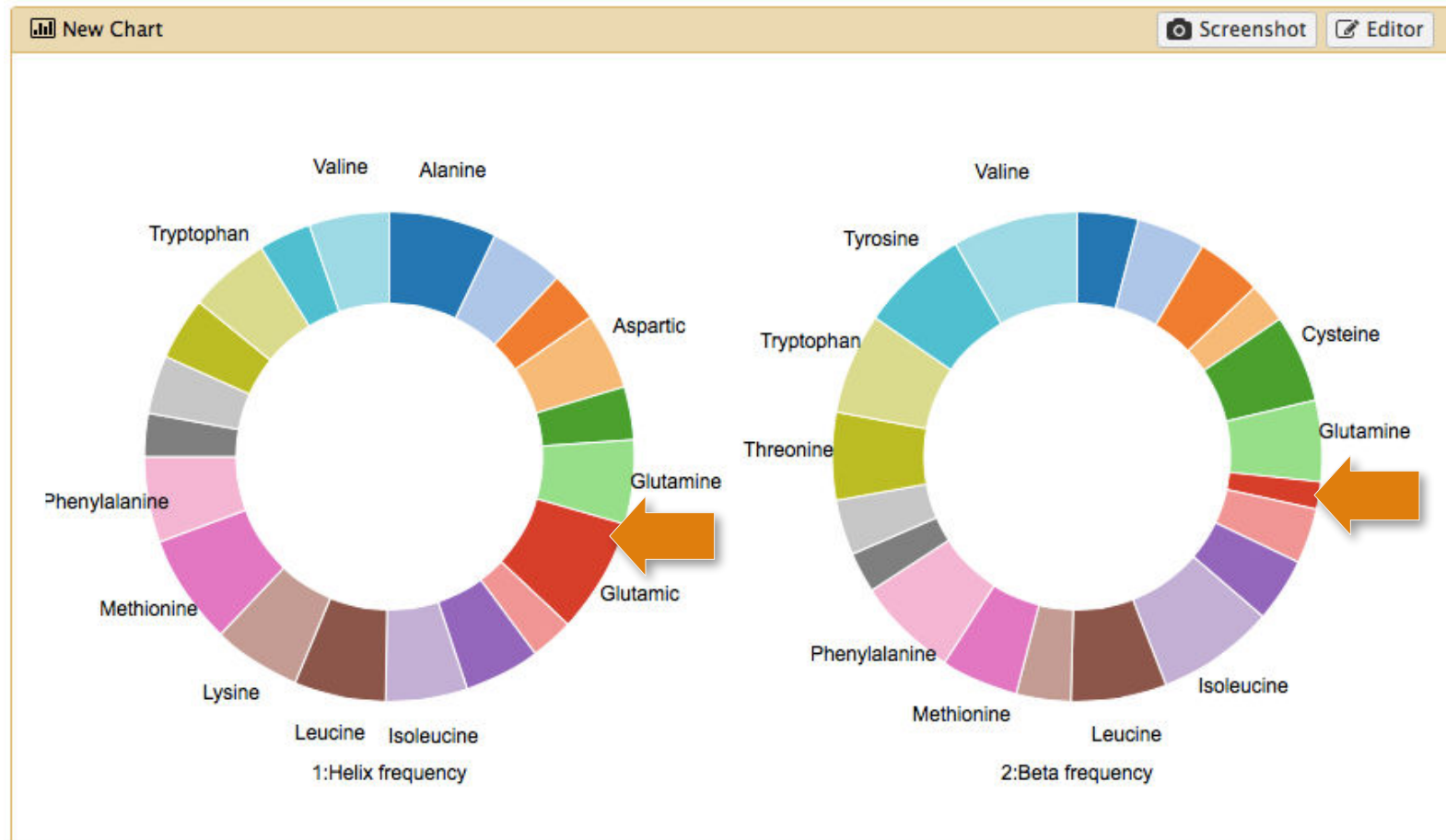
Label settings:

Donut label Label column What would you like to show for each slice?

Show outside Yes No Would you like to show labels outside the donut?

Configure the **Pie chart** as shown above. Then, click on **Draw**.

Configure the pie chart



Glutamic acids seem to fit much better into **helices** than **beta sheets**. In other words, “Aspartic and Glutamic Acids are Important for Alpha-helix Folding”, JBSD 2007.

Create a bar diagram

Amino Acid Features (Bar diagram) Cancel Draw

[Start](#) [Configuration](#) [1: Hydrophobicity](#) [2: Membrane](#) [3: Flexibility](#) [4: Helix](#) [5: Beta](#) [Add Data](#)

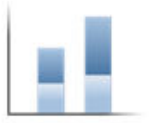
Provide a chart title:

Amino Acid Features (Bar diagram)

How many data points would you like to analyze?

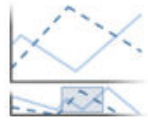
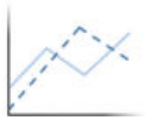
☒ Few (<500) ☐ Some (<10k) ☐ Many (>10k)

• Bar diagrams

Regular (NVD3) Stacked (NVD3) Horizontal (NVD3) Stacked horizontal (NVD3)

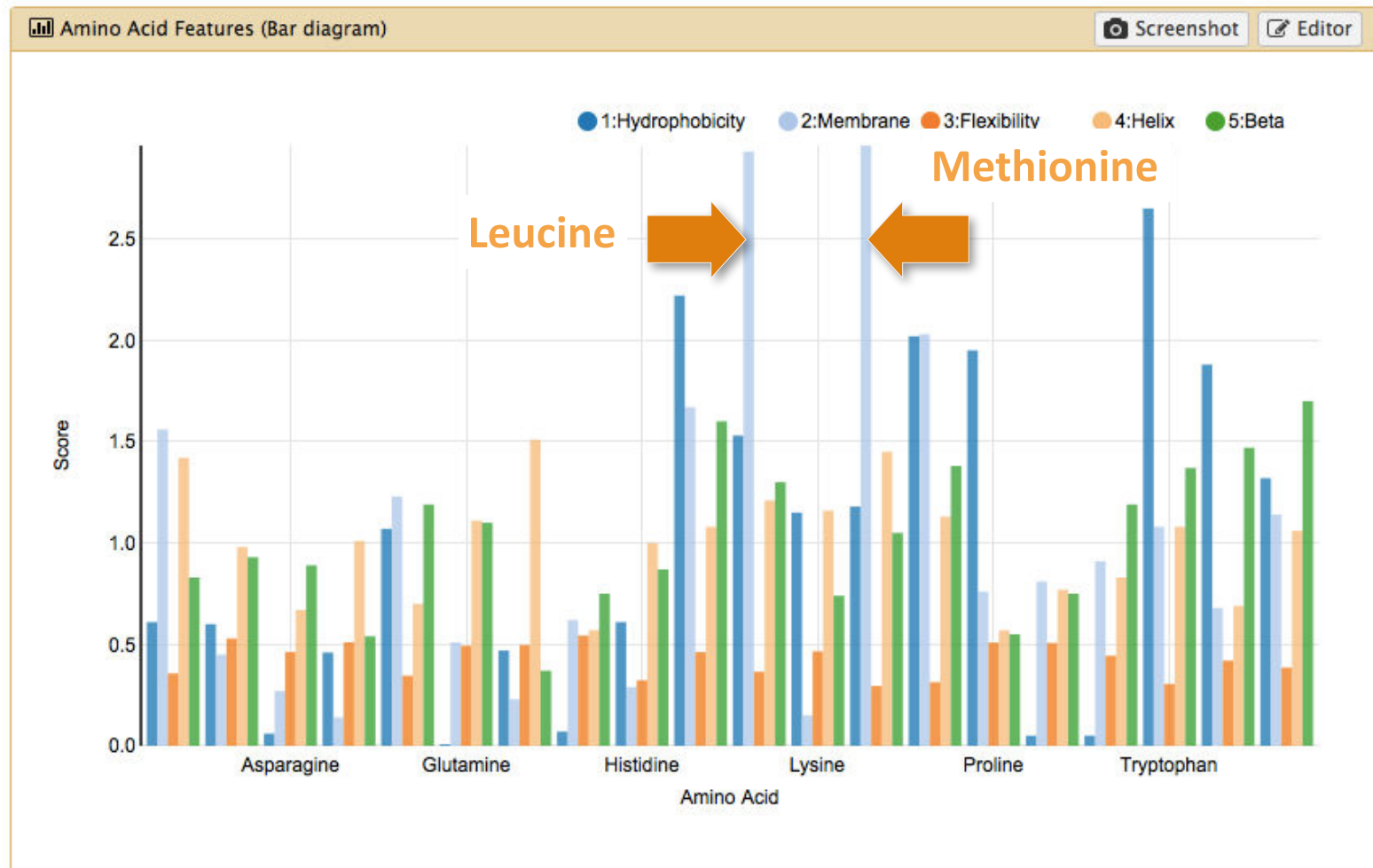
• Others

Line with focus Line chart Scatter plot Heatmap

Create data groups for the following features: **Hydrophobicity**, **Membrane frequency**, **Flexibility**, **Helix frequency** and **Beta frequency**.

Bar diagram of amino acid features



Use the **tooltips** to identify the amino acids which are likely to be found within membrane proteins.

Topics

Visualization history and introduction

Biological Visualizations

Numerical Visualizations

Adding your own visualizations

Adding your own Visualizations

Go to **config/plugins/visualizations/charts**

charts/others/YOURVIZNAME

Add three files to this directory:

Logo (logo.png)

Configuration (config.js)

Wrapper (wrapper.js)

charts/types.js

Rebuild by typing 'npm install' and 'grunt'