

Cancer Data Analysis with Caleydo StratomeX and enRoute

Marc Streit

Johannes Kepler University Linz, Austria

Alexander Lex

Harvard University, Cambridge, USA



JKU
JOHANNES KEPLER
UNIVERSITÄT LINZ



HARVARD

School of Engineering
and Applied Sciences



CALEYDO

What is Caleydo?

Software for visualizing biomolecular data

tabular data

numerical & categorical

e.g., mRNA, microRNA, copy number variation, methylation, mutation status, etc.

clinical data

pathways

KEGG, WikiPathways



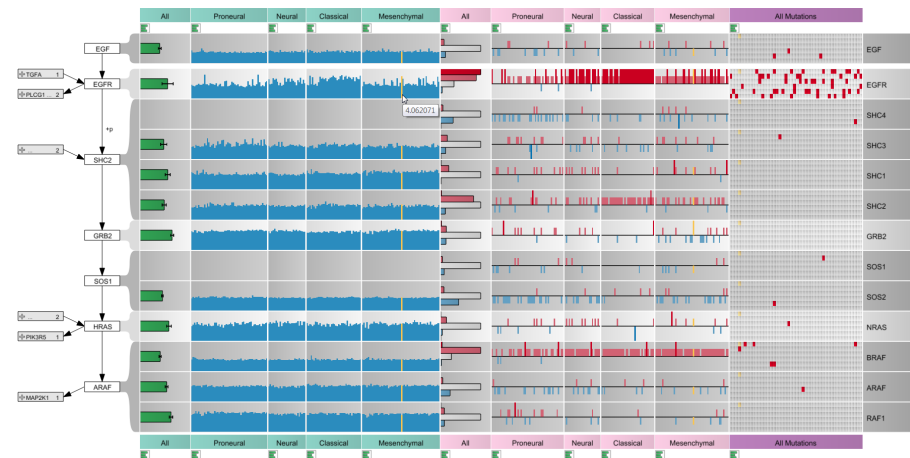
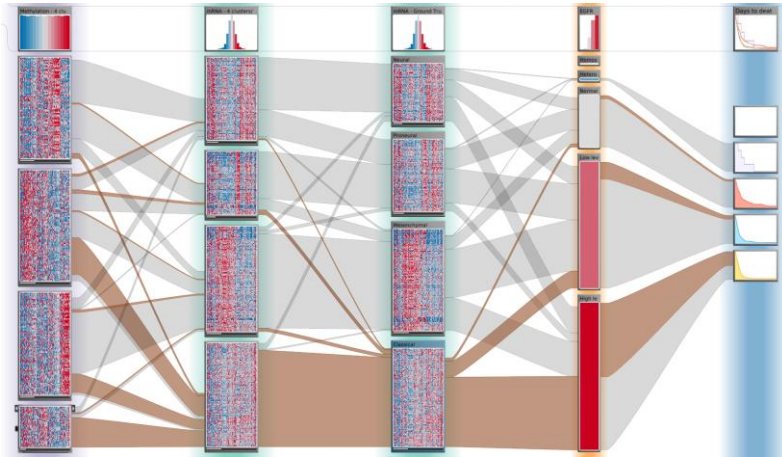
CALEYDO

Caleydo Core Features

Multi-Dataset Analysis. Want to see....

...relationships between multiple datasets?

...relationships between tabular and graph data?



What is Caleydo?

Software for doing **research in visualization**

developed in academic setting

platform for trying out radically new visualization ideas

Quest for compromise between **academic prototyping** and **ready-to-use software**



What is Caleydo?

Open source platform for developing
visualization and data analysis techniques
easily extendible due to plug-in architecture
you can create your own views
you can plug-in your own algorithms

Who is Caleydo?

Marc Streit

Johannes Kepler University Linz, AT

Alexander Lex

Harvard University, Cambridge, USA

Christian Partl

Graz University of Technology, AT

Samuel Gratzl

Johannes Kepler University Linz, AT

Nils Gehlenborg

Harvard Medical School, Boston, USA

Dieter Schmalstieg

Graz University of Technology, AT

Outline

14:00 - 15:15	Intro Data Format & Loading Data Assignment, Clustering Tabular Data & Subtype Visualization I
15:15 - 15:30	Break
15:30 - 16:30	Tabular Data & Subtype Visualization II Pathway Visualization

DATA FORMAT & LOADING

Data Example

Comma-separated text files

various types of separators

Column Identifiers (Samples)

Header

Row Identifiers (Genes)

Irrelevant Column

Data

Marc Streit & Alexander Lex

	A	B	C	D
1	#1.2			
2	1500	490		
3	Name	Description	TCGA.02.0001.01C.01R.0177.01	TCGA.02.0003.01A.01R.0177.01
4	LTF	LTF		
5	POSTN	POSTN	0.982	0.9933333333
6	TMSL8	TMSL8	0.0686666667	0.1046666667
7	RP11-35N6.1	RP11-35N6.1	0.062	0.0593333333
8	HLA-DQA1	HLA-DQA1	0.234	0.9986666667
9	STMN2	STMN2	0.1246666667	0.0526666667
10	DCX	DCX	0.166	0.0886666667
11	IL13RA2	IL13RA2	0.098	0.1593333333
12	SLN	SLN	0.1306666667	0.114
13	AGXT2L1	AGXT2L1	0.122	0.0513333333
14	MEOX2	MEOX2	0.1386666667	0.3253333333
15	COL11A1	COL11A1	0.8793333333	0.4953333333
16	CXCL14	CXCL14	0.0753333333	0.2553333333
17	NNMT	NNMT	0.7946666667	0.744
18	F13A1	F13A1	0.6186666667	0.948
19	MBP	MBP	0.2173333333	0.1393333333
20	GABRB1	GABRB1	0.1686666667	0.0546666667
21	ITF	ITF	0.022	0.0093333333
22	SLPI	SLPI	0.8266666667	0.9893333333
23	PLA2G5	PLA2G5	0.2013333333	0.1506666667
24	KCNJ16	KCNJ16	0.09	0.1006666667
25	KCND2	KCND2	0.2373333333	0.136
26	COL6A3	COL6A3	0.996	0.9973333333
27	MOXD1	MOXD1	0.5986666667	0.678
28	AQP1	AQP1	0.072	0.614
29	JARID1D	JARID1D	0.682	0.1966666667
30	PROM1	PROM1	0.1213333333	0.9233333333
31	PTX3	PTX3	0.0646666667	0.0453333333
32	CHI3L2	CHI3L2	0.1226666667	0.1226666667

Data Format

Data Import Wizard

Load Dataset
Specify the dataset you want to load.

Input File
Open Data File: /home/alexsb/Desktop/mrna.gct

Dataset Name
mrna

Row Configuration
Row ID Class: GENE
Number of Header Rows: 3
Column with Row IDs: 1

Column Configuration
Column ID Class: SAMPLE
Column ID Type: SAMPLE
Row with Column IDs: 1
☒ Columns use same Scale

Separated by (delimiter)
☒ TAB ☐ ; ☐ , ☐ . ☐ SPACE ☐ Other

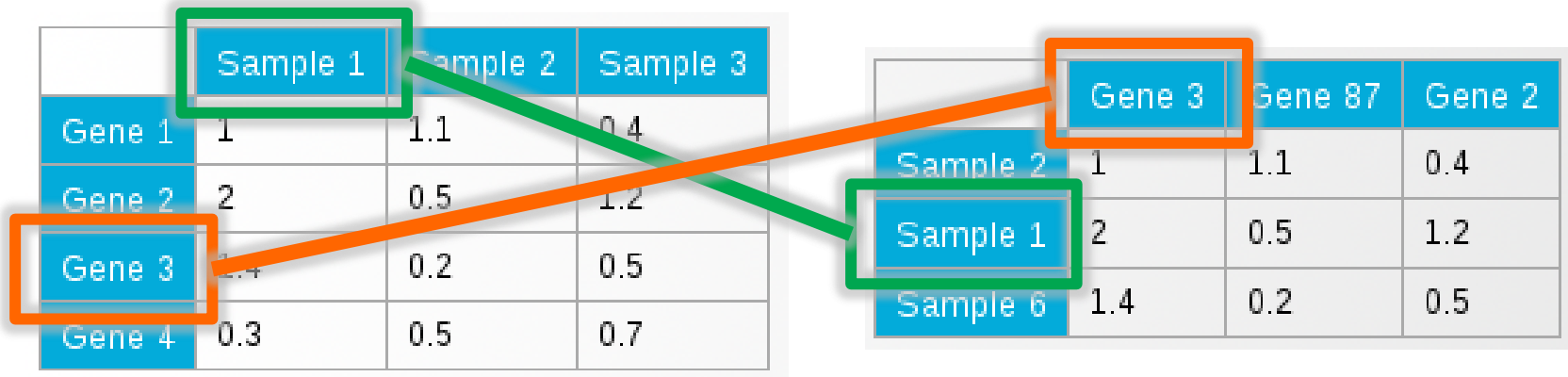
Column Selection
Select All Select None

Use column		☒	2	☒	3	☒	4	☒	5	☒	6	☒	7	☒	8	☒	9	☒	10
1	#1.2																		
2	1500		490																
3	Name		Description		TCGA.02.0001.01		TCGA.02.0003.01		TCGA.02.0004.01		TCGA.02.0007.01		TCGA.02.0009.01		TCGA.02.0010.01		TCGA.02.0011.01		TCGA.02.0014.01A.011
4	LTF		LTF		0.332		0.957333333333		0.985333333333		0.089333333333		0.075333333333		0.5		0.079333333333		0.0033333333333333
5	POSTN		POSTN		0.982		0.993333333333		0.994		0.061333333333		0.992666666666		0.112666666666		0.068		0.002
6	TMSL8		TMSL8		0.068666666666		0.104666666666		0.477333333333		0.11		0.047333333333		0.523333333333		0.138		0.996
7	RP11-35N6.1		RP11-35N6.1		0.062		0.059333333333		0.002666666666		0.138		0.072666666666		0.064		0.110666666666		0.9733333333333333
8	HLA-DQA1		HLA-DQA1		0.234		0.998666666666		0.966		0.266		0.988666666666		0.381333333333		0.332		0.526
9	STMN2		STMN2		0.124666666666		0.052666666666		0.061333333333		0.231333333333		0.149333333333		0.104		0.176666666666		0.714
10	DCX		DCX		0.166		0.088666666666		0.009333333333		0.256666666666		0.091333333333		0.219333333333		0.249333333333		1
11	IL13RA2		IL13RA2		0.098		0.159333333333		0.96		0.140666666666		0.052666666666		0.82		0.146666666666		0.7826666666666667
12	FLN		FLN		0.430666666666		0.444		0.76		0.333666666666		0.083		0.444		0.318		0.0533333333333333

Data Format

Multiple heterogeneous tables

Matching based on IDs



Data Format

Caleydo has many **genetic identifiers built-in**

ID Classes group ID types

ID types in a class are different but can be mapped

e.g., RefSeq ID maps to gene short name

Data Format

Data Import Wizard

Load Dataset

Specify the dataset you want to load.

Input File

Open Data File

/home/alexs/b/Desktop/mrna.gct

Dataset Name

mrna

Row Configuration

Row ID Class

GENE

New

Row ID Type

GENE_SYMBOL

New

Define Parsing

Column with Row IDs

1

Column Configuration

Column ID Class

SAMPLE

New

Column ID Type

SAMPLE

New

Define Parsing

Row with Column IDs

1

☒ Columns use same Scale

Separated by (delimiter)

TAB

,

.

SPACE

Other

Column Selection

Select All

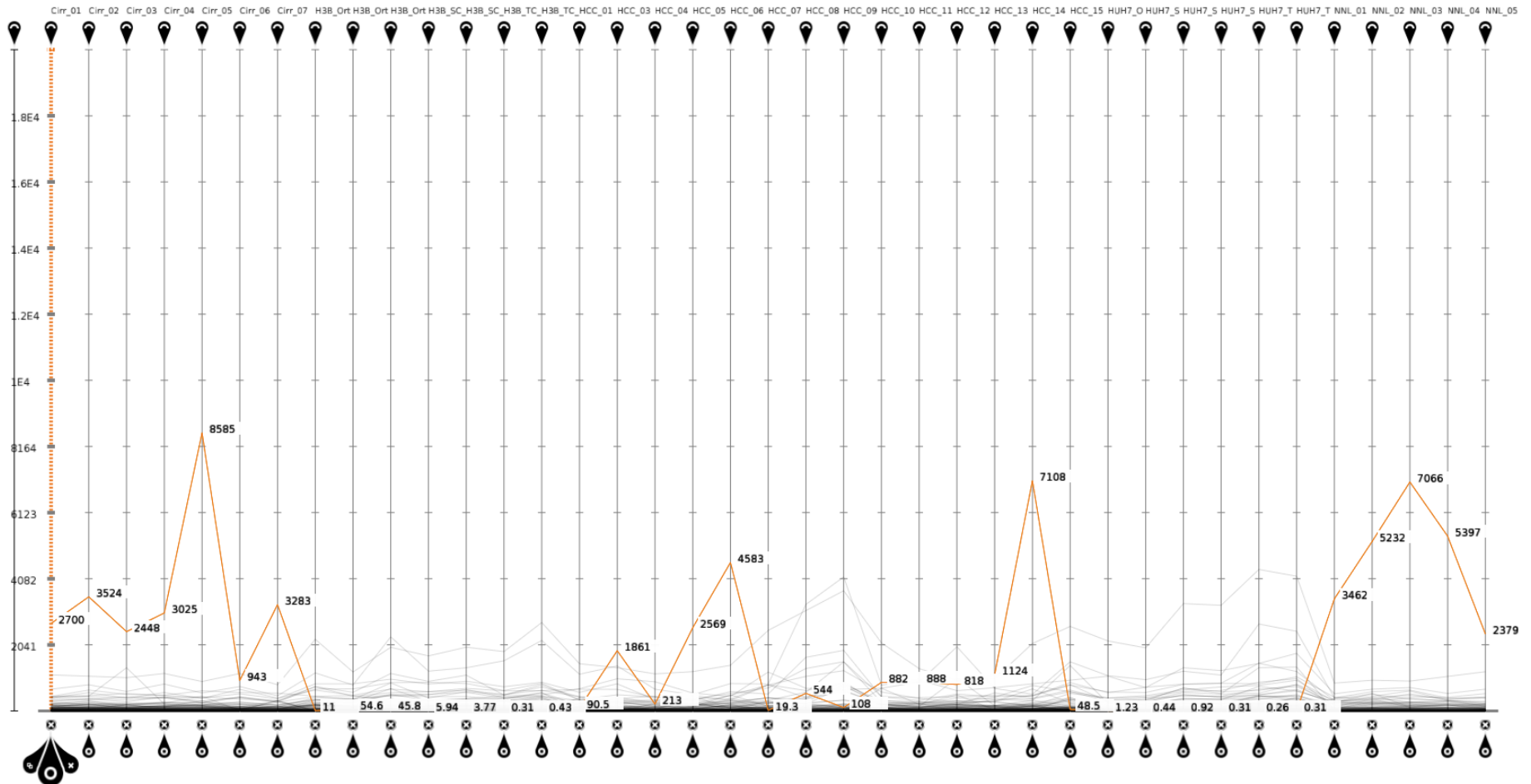
Select None

Use column		☒	2	☒	3	☒	4	☒	5	☒	6	☒	7	☒	8	☒	9	☒	10
1	#1.2	☒		☒		☒		☒		☒		☒		☒		☒		☒	
2	1500	☒	490	☒		☒		☒		☒		☒		☒		☒		☒	
3	Name	☒	Description	☒	TCGA.02.0001.01	☒	TCGA.02.0003.01	☒	TCGA.02.0004.01	☒	TCGA.02.0007.01	☒	TCGA.02.0009.01	☒	TCGA.02.0010.01	☒	TCGA.02.0011.01	☒	TCGA.02.0014.01A.0
4	LTF	☒	LTF	☒	0.332	☒	0.957333333333	☒	0.985333333333	☒	0.089333333333	☒	0.075333333333	☒	0.5	☒	0.079333333333	☒	0.003333333333
5	POSTN	☒	POSTN	☒	0.982	☒	0.993333333333	☒	0.994	☒	0.061333333333	☒	0.992666666666	☒	0.112666666666	☒	0.068	☒	0.002
6	TMSL8	☒	TMSL8	☒	0.068666666666	☒	0.104666666666	☒	0.477333333333	☒	0.11	☒	0.047333333333	☒	0.523333333333	☒	0.138	☒	0.996
7	RP11-35N6.1	☒	RP11-35N6.1	☒	0.062	☒	0.059333333333	☒	0.002666666666	☒	0.138	☒	0.072666666666	☒	0.064	☒	0.110666666666	☒	0.973333333333
8	HLA-DQA1	☒	HLA-DQA1	☒	0.234	☒	0.998666666666	☒	0.966	☒	0.266	☒	0.988666666666	☒	0.381333333333	☒	0.332	☒	0.526
9	STMN2	☒	STMN2	☒	0.124666666666	☒	0.052666666666	☒	0.061333333333	☒	0.231333333333	☒	0.149333333333	☒	0.104	☒	0.176666666666	☒	0.714
10	DCX	☒	DCX	☒	0.166	☒	0.088666666666	☒	0.009333333333	☒	0.256666666666	☒	0.091333333333	☒	0.219333333333	☒	0.249333333333	☒	1

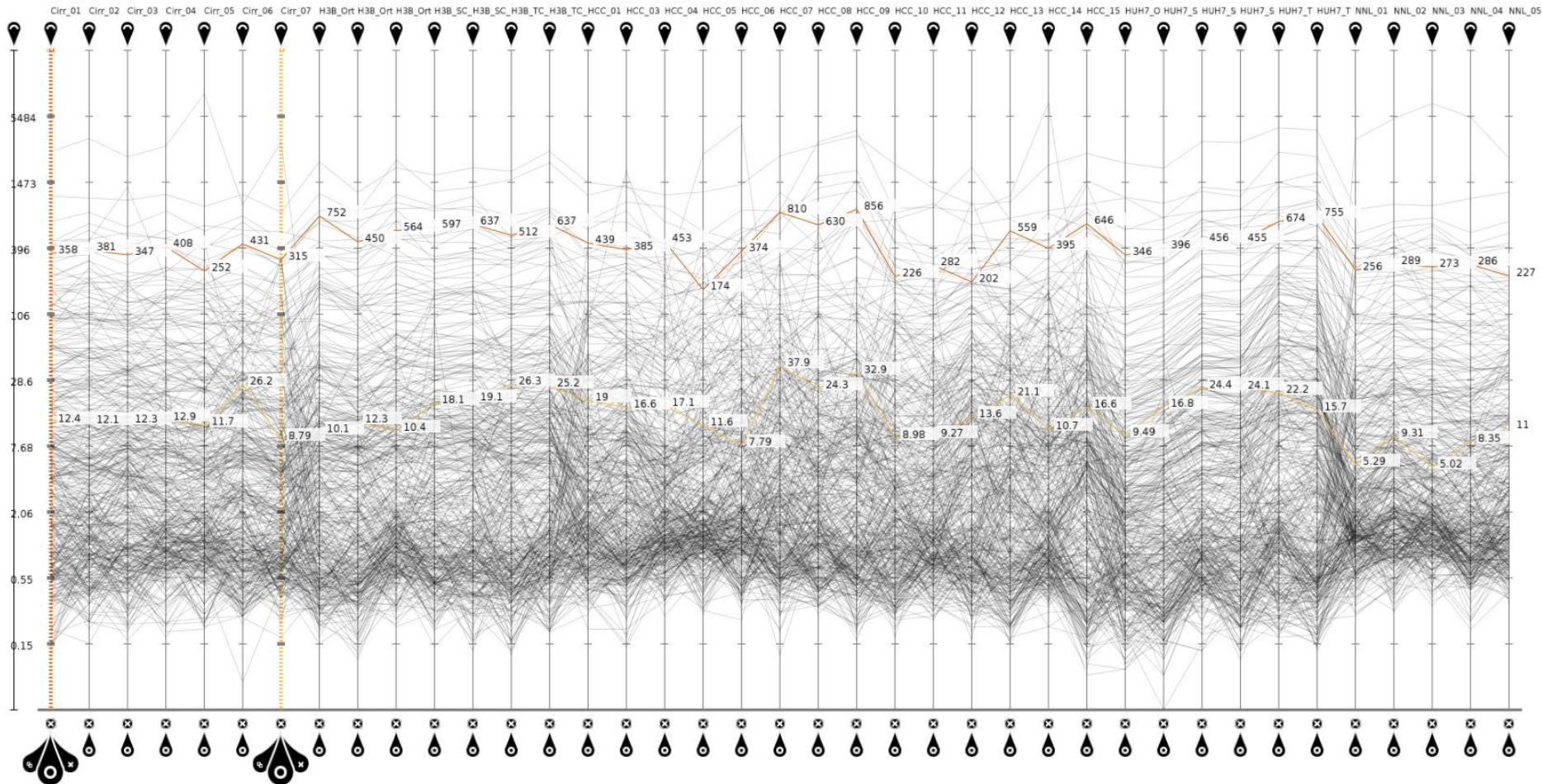
Transformations

*How to transform your data to
make sense for visualization?*

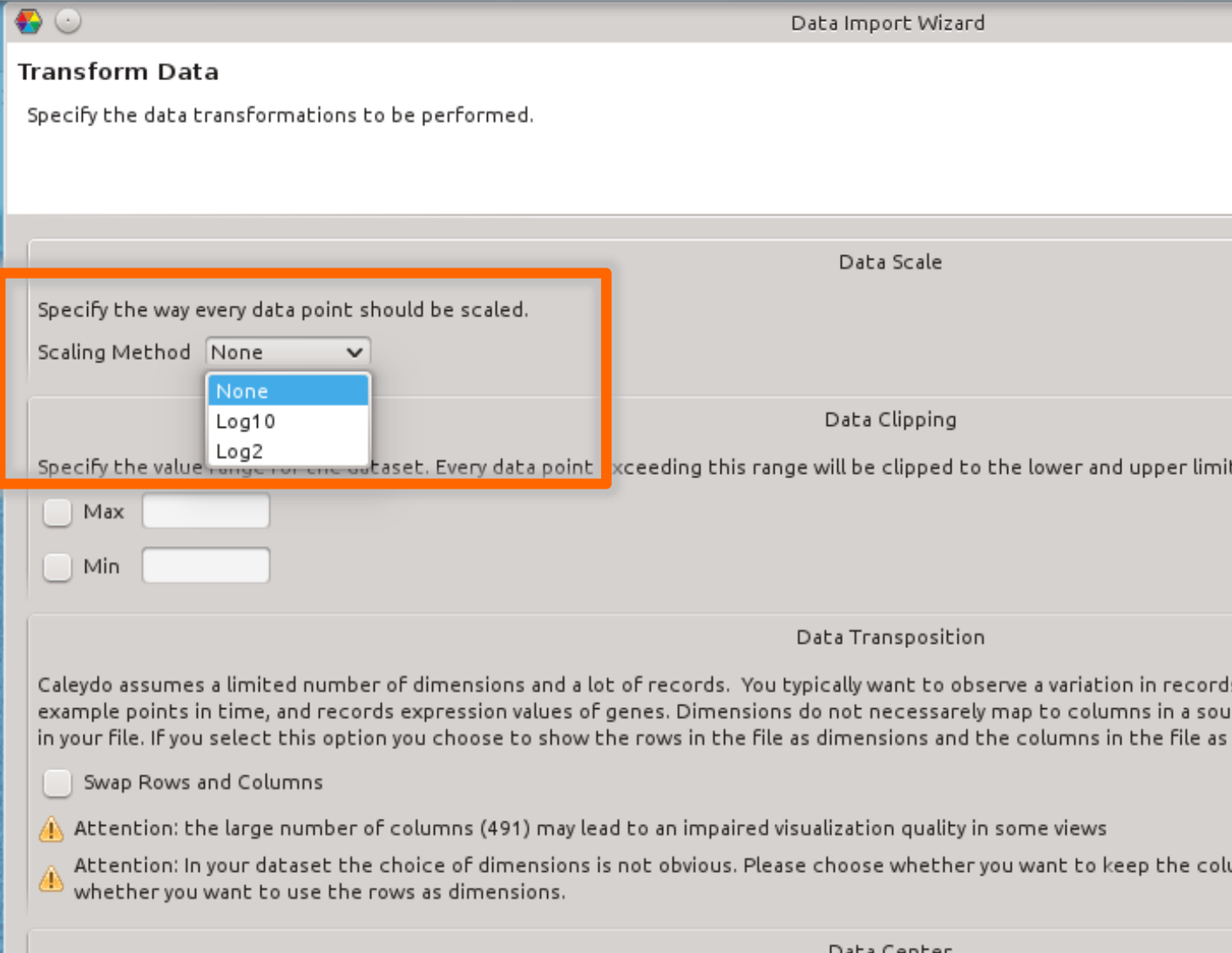
Scale Transformation



Scale Transformation - LOG



Scale Transformation



Data Import Wizard

Transform Data

Specify the data transformations to be performed.

Data Scale

Specify the way every data point should be scaled.

Scaling Method: **None** (dropdown menu open showing: None, Log10, Log2)

Data Clipping

Specify the value range for the dataset. Every data point exceeding this range will be clipped to the lower and upper limit.

☐ Max

☐ Min

Data Transposition

Caleydo assumes a limited number of dimensions and a lot of records. You typically want to observe a variation in record, for example points in time, and records expression values of genes. Dimensions do not necessarily map to columns in a source file in your file. If you select this option you choose to show the rows in the file as dimensions and the columns in the file as records.

☐ Swap Rows and Columns

Attention: the large number of columns (491) may lead to an impaired visualization quality in some views

Attention: In your dataset the choice of dimensions is not obvious. Please choose whether you want to keep the columns as dimensions or whether you want to use the rows as dimensions.

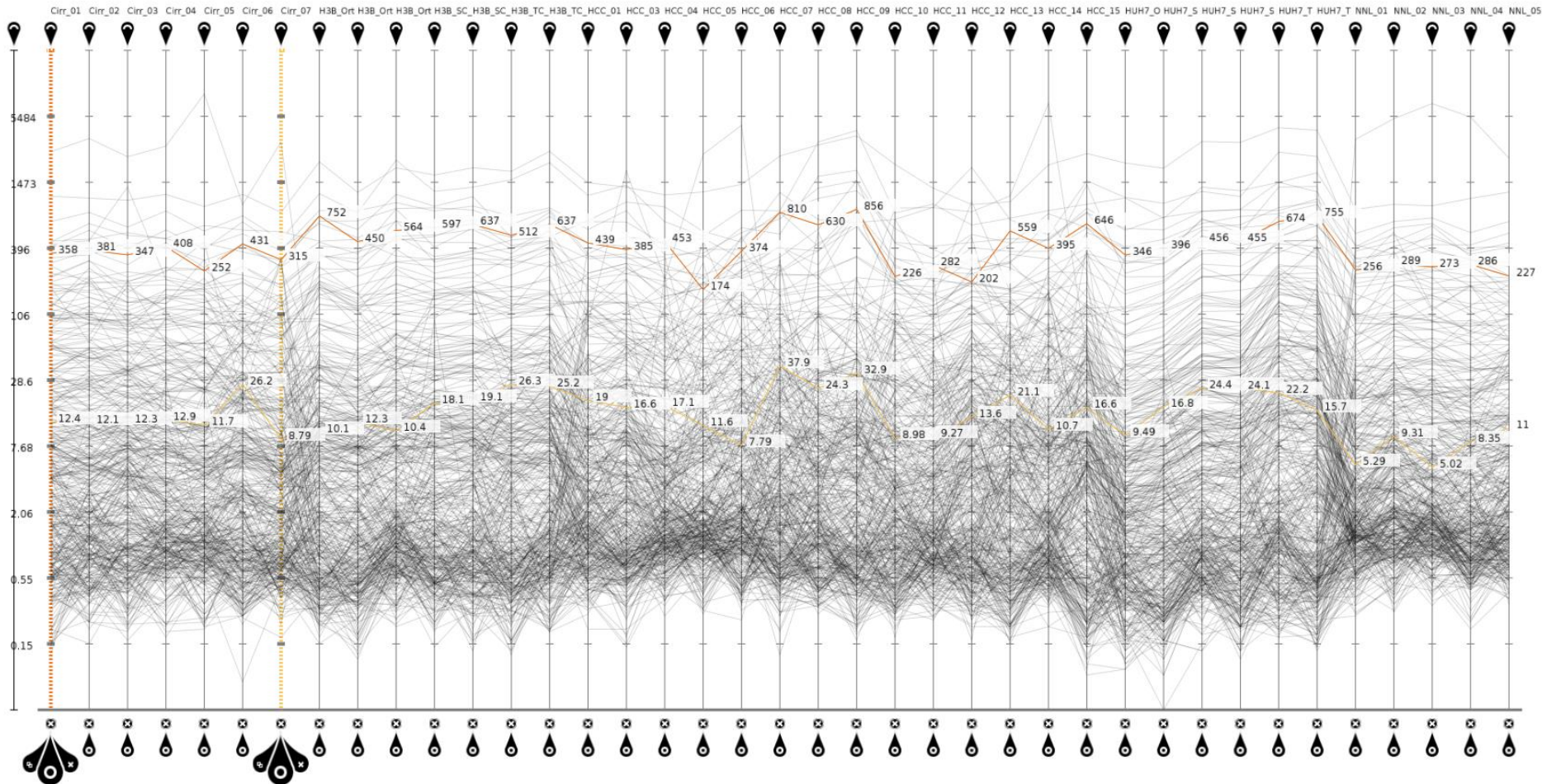
Data Center

Switch Columns with Rows

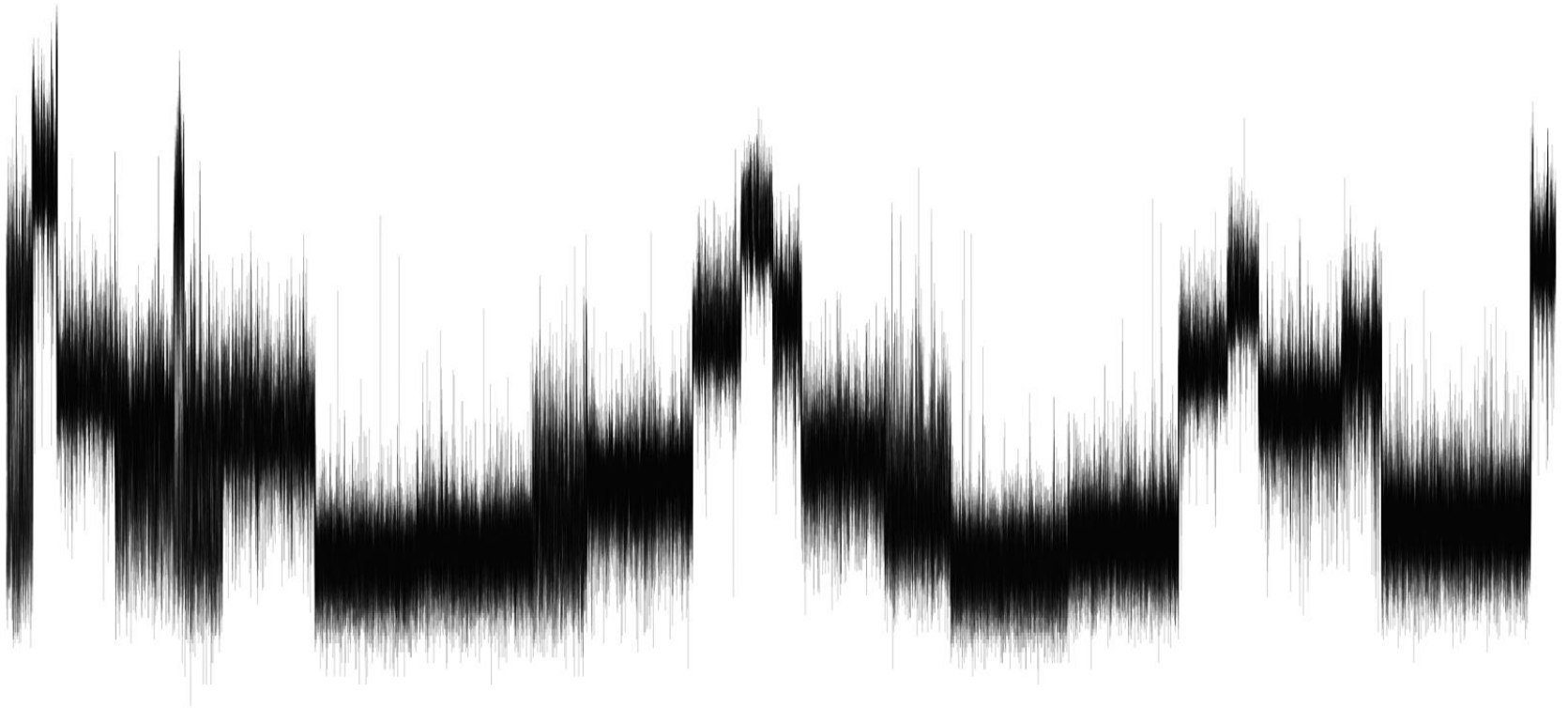
Transpose data

	Sample 1	Sample 2	Sample 3
Gene 1	1	1.1	0.4
Gene 2	2	0.5	1.2
Gene 3	1.4	0.2	0.5
Gene 4	0.3	0.5	0.7

Columns & Rows as in Source

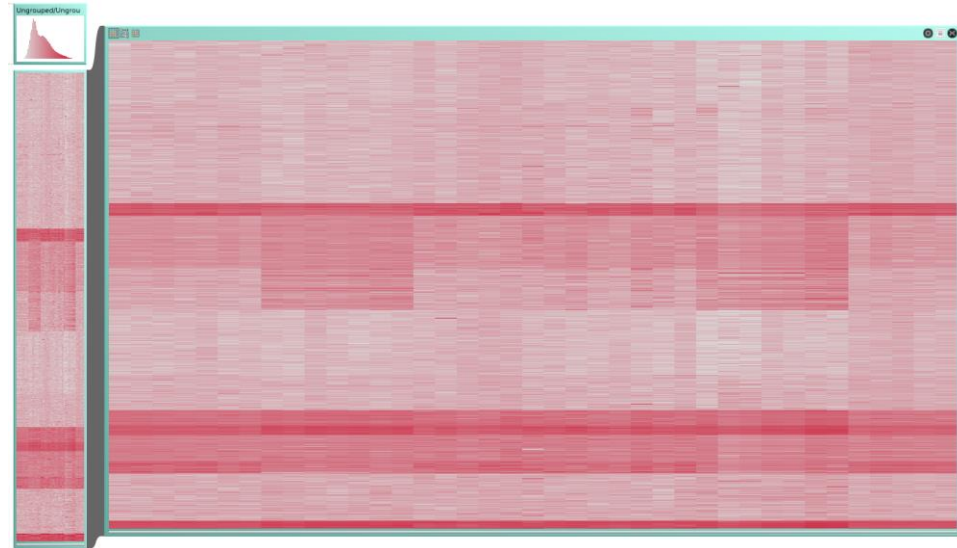


Columns & Rows Switched

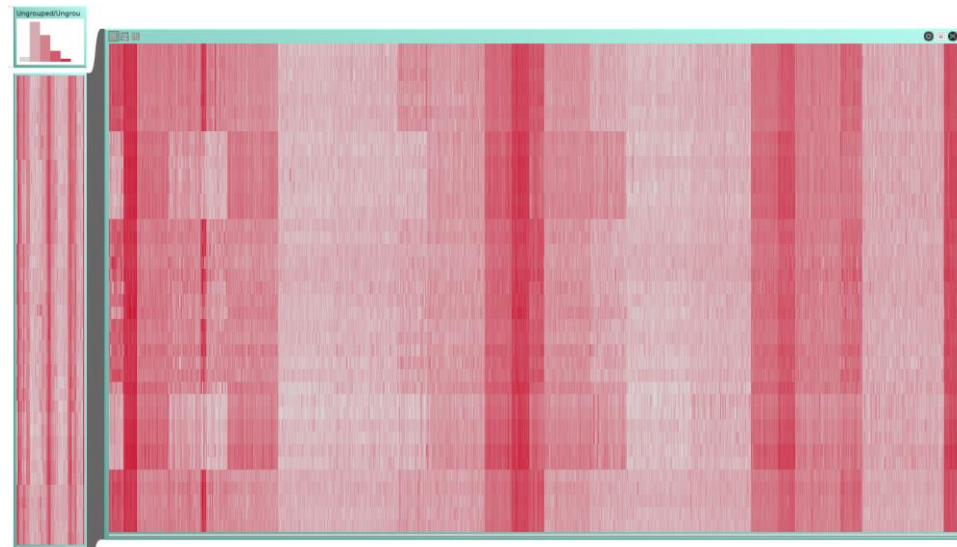


Switch Columns with Rows

Genes = Rows
Samples = Columns



Genes = Columns
Samples = Rows



Switch Columns with Rows



Data Import Wizard

Transform Data
Specify the data transformations to be performed.

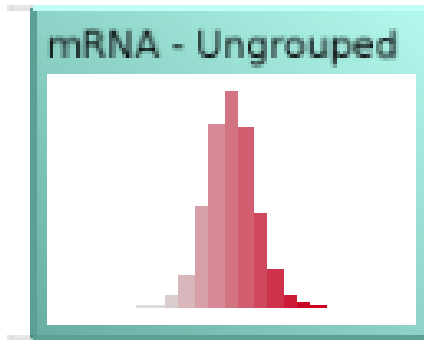
Data Scale
Specify the way every data point should be scaled.
Scaling Method:

Data Clipping
Specify the value range for the dataset. Every data point exceeding this range will be clipped to the lower and upper limits respectively.
☐ Max
☐ Min

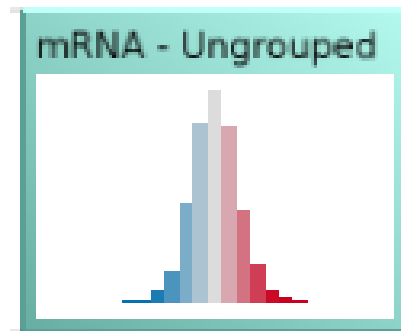
Data Transposition
Caleydo assumes a limited number of dimensions and a lot of records. You typically want to observe a variation in records over dimensions, where dimensions would be, for example points in time, and records expression values of genes. Dimensions do not necessarily map to columns in a source file and equally records must not be the rows in your file. If you want to use the columns in the file as records, you can swap rows and columns.
☒ Swap Rows and Columns
 Attention: the large number of columns (1503) may lead to an impaired visualization quality in some views

Data Center
The data center is a value that, if set, determines a neutral center point of the data. A common example is that 0 is the neutral value, lower values are in the negative and larger values are in the positive range. If the data center is set it is assumed that the extend into both, positive and negative direction is the same. For example, for a dataset [-0.5, 0.7] with a center set at 0, the value range will be set to -0.7 to 0.7.
☒ Use data center

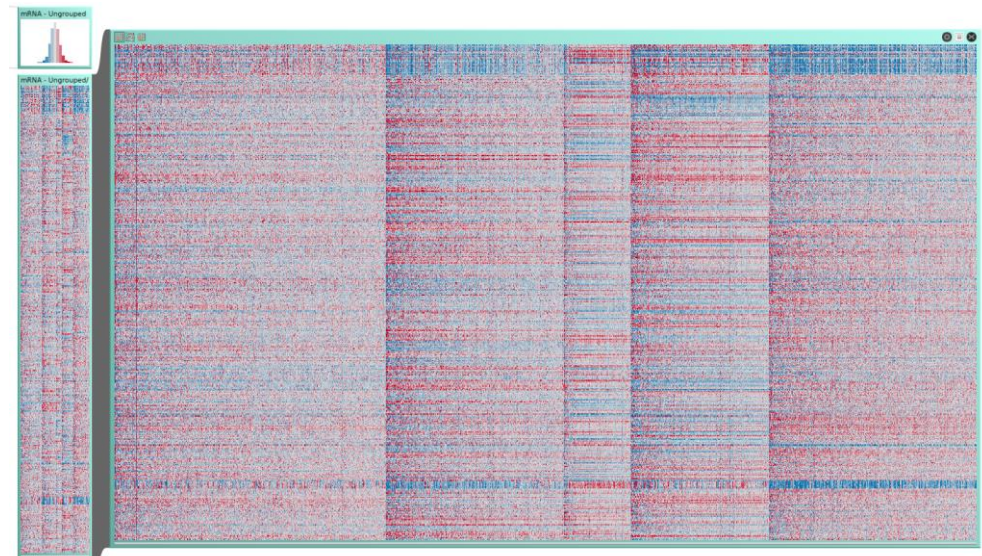
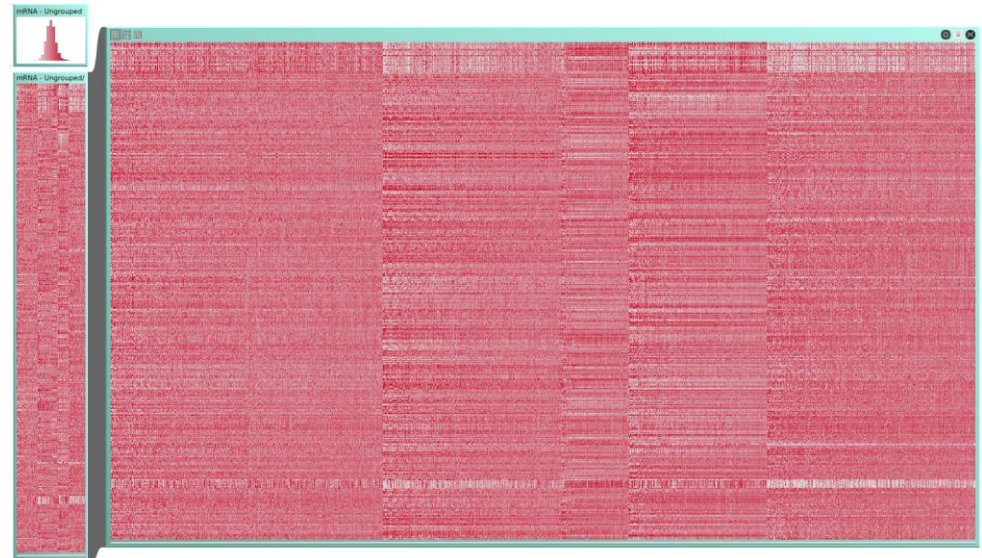
Data Center



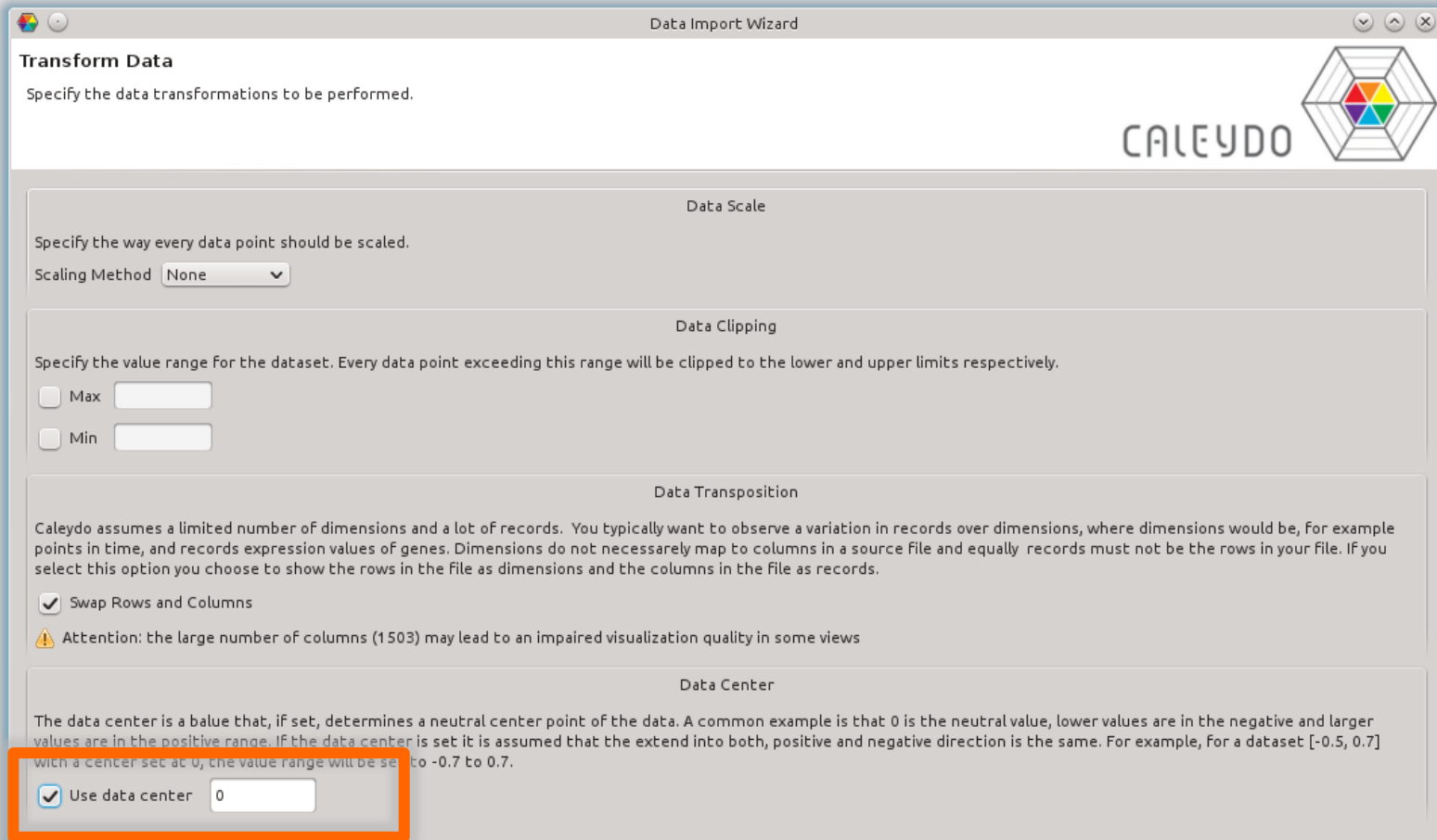
Data not centered
White-red color map



Data centered (at 0)
Blue-white-red color map



Switch Columns with Rows



Data Import Wizard

Transform Data
Specify the data transformations to be performed.

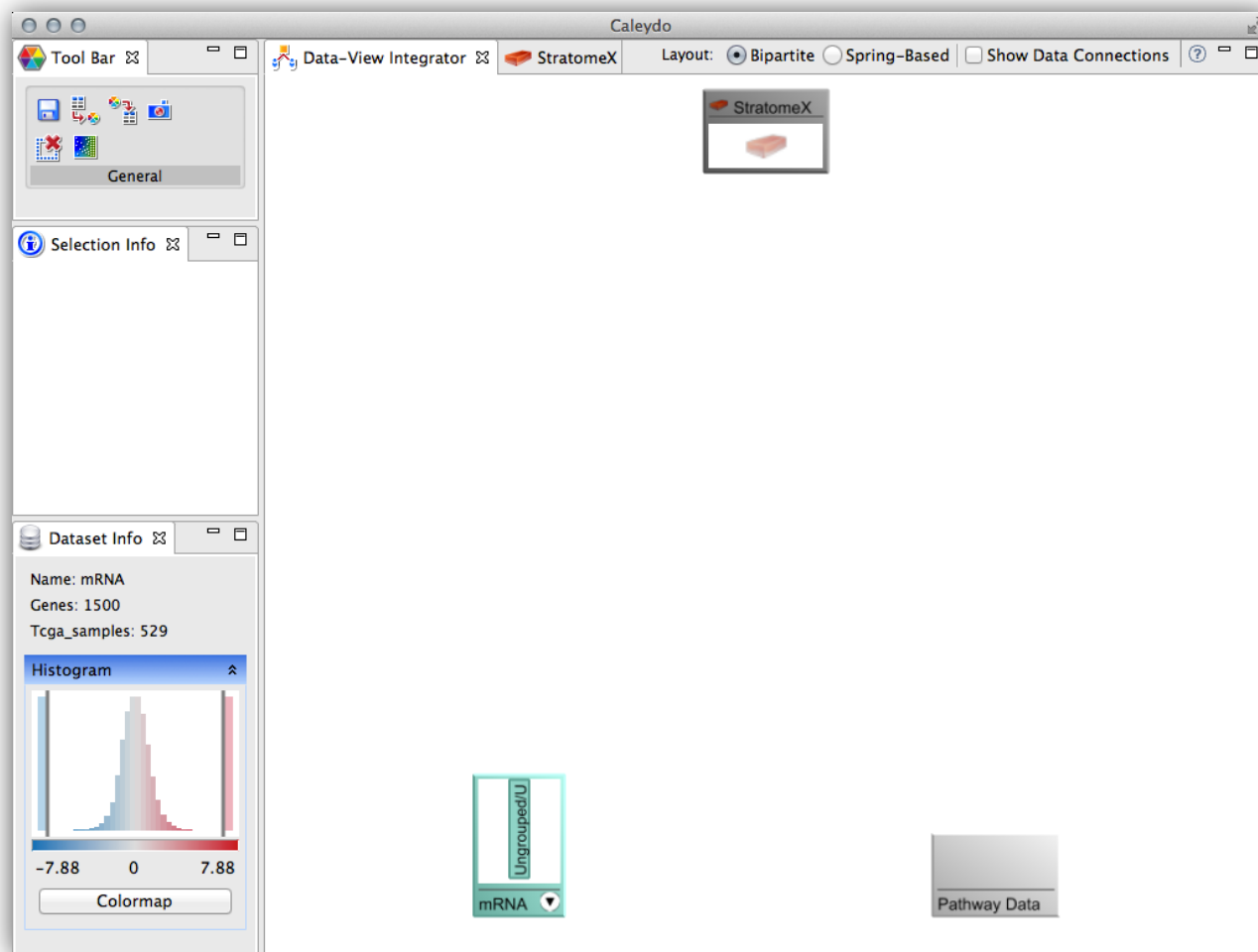
Data Scale
Specify the way every data point should be scaled.
Scaling Method:

Data Clipping
Specify the value range for the dataset. Every data point exceeding this range will be clipped to the lower and upper limits respectively.
☐ Max
☐ Min

Data Transposition
Caleydo assumes a limited number of dimensions and a lot of records. You typically want to observe a variation in records over dimensions, where dimensions would be, for example points in time, and records expression values of genes. Dimensions do not necessarily map to columns in a source file and equally records must not be the rows in your file. If you select this option you choose to show the rows in the file as dimensions and the columns in the file as records.
☒ Swap Rows and Columns
⚠ Attention: the large number of columns (1503) may lead to an impaired visualization quality in some views

Data Center
The data center is a value that, if set, determines a neutral center point of the data. A common example is that 0 is the neutral value, lower values are in the negative and larger values are in the positive range. If the data center is set it is assumed that the extend into both, positive and negative direction is the same. For example, for a dataset [-0.5, 0.7] with a center set at 0, the value range will be set to -0.7 to 0.7.
☒ Use data center

Data Successfully loaded



What's next?

Start **exploring** your data

Load **another** dataset

Save as **project**

Scripted Data Loading

```
private DataSetDescription setUpMRNAData() {
    DataSetDescription mrnaData = new DataSetDescription(ECreateDefaultProperties.NUMERICAL);
    mrnaData.setDataSetName("mRNA");

    mrnaData.setDataSourcePath(MRNA);
    mrnaData.setNumberOfHeaderLines(3);

    ParsingRule parsingRule = new ParsingRule();
    parsingRule.setFromColumn(2);
    parsingRule.setParseUntilEnd(true);
    parsingRule.setColumnDescription(new ColumnDescription());
    mrnaData.addParsingRule(parsingRule);
    mrnaData.setTransposeMatrix(true);

    IDSpecification geneIDSpecification = new IDSpecification();
    geneIDSpecification.setIDTypeGene(true);
    geneIDSpecification.setIdType("GENE_SYMBOL");
    mrnaData.setRowIDSpecification(geneIDSpecification);
    mrnaData.setColumnIDSpecification(sampleIDSpecification);

    GroupingParseSpecification firehoseClustering = new GroupingParseSpecification(
        MRNA_GROUPING);
    firehoseClustering.setContainsColumnIDs(false);
    firehoseClustering.setRowIDSpecification(sampleIDSpecification);
    mrnaData.addColumnGroupingSpecification(firehoseClustering);

    DataProcessingDescription dataProcessingDescription = new DataProcessingDescription();
    ClusterConfiguration clusterConfiguration = new ClusterConfiguration();
    clusterConfiguration.setDistanceMeasure(EDistanceMeasure.EUCLIDEAN_DISTANCE);
    KMeansClusterConfiguration kMeansAlgo = new KMeansClusterConfiguration();
    kMeansAlgo.setNumberOfClusters(5);
    clusterConfiguration.setClusterAlgorithmConfiguration(kMeansAlgo);
    dataProcessingDescription.addRowClusterConfiguration(clusterConfiguration);
    mrnaData.setDataProcessingDescription(dataProcessingDescription);

    return mrnaData;
}
```

Projects contain

the **data** you loaded
the **views** you opened
your **clustering** results and filters
your **color** codes
etc.,

Stored in .cal files

2 sample datasets i
complex glioblastom
datasets and clusteri
simple mRNA datase
and other conditions



HANDING OVER TO MARC

DATA ASSIGNMENT, CLUSTERING

Challenge 1

Challenge 2

The diagram illustrates a complex flow of information, likely representing a biological or chemical process. It is structured into three main columns: Input, Processing, and Output.

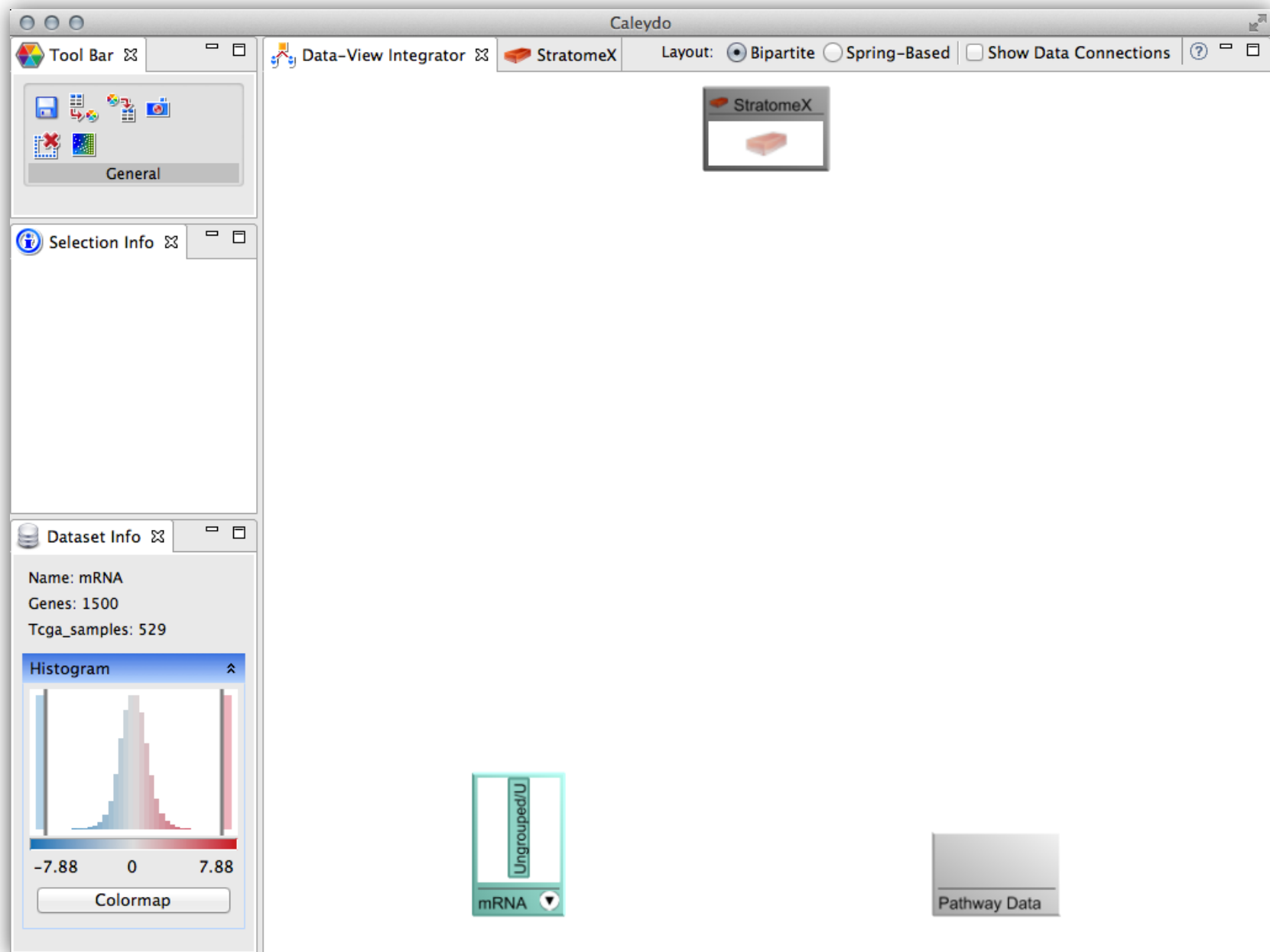
- Input Column:** Contains five stacked rectangular blocks. From top to bottom, they are labeled: "Memory", "History", "Normal", "Low level", and "High level". The "Low level" and "High level" blocks are significantly larger and colored red, while the others are smaller and colored grey.
- Processing Column:** Contains four stacked rectangular blocks. From top to bottom, they are labeled: "Neural", "Metasternal", "Postnatal", and "Chemical". Each block contains a dense, colorful pattern of red and blue dots, suggesting a complex internal structure or data representation.
- Output Column:** Contains five stacked rectangular blocks. From top to bottom, they are labeled: "Output", "Output", "Output", "Output", and "Output". Each block contains a different type of visualization: the top block is a line graph, the second is a bar chart, the third is a line graph, the fourth is a bar chart, and the bottom is a line graph.

Flow lines (Sankey diagram) connect the blocks between columns, showing the distribution of information. The "Low level" and "High level" input blocks have the most prominent flows, leading into the "Neural" and "Chemical" processing blocks, which then lead into the "Output" blocks.

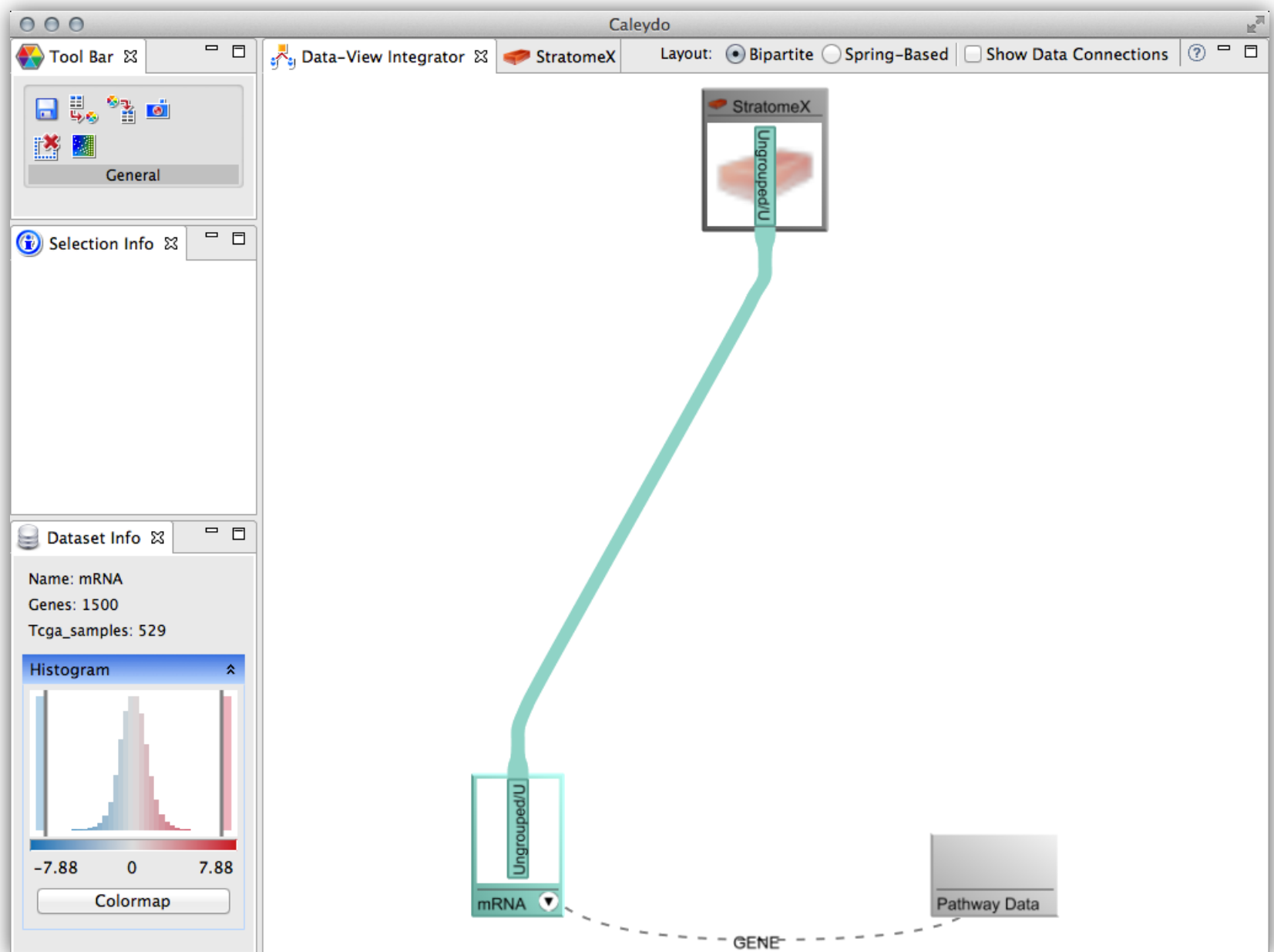
mRNA Dataset:

Data assignment + heatmap + color
mapping

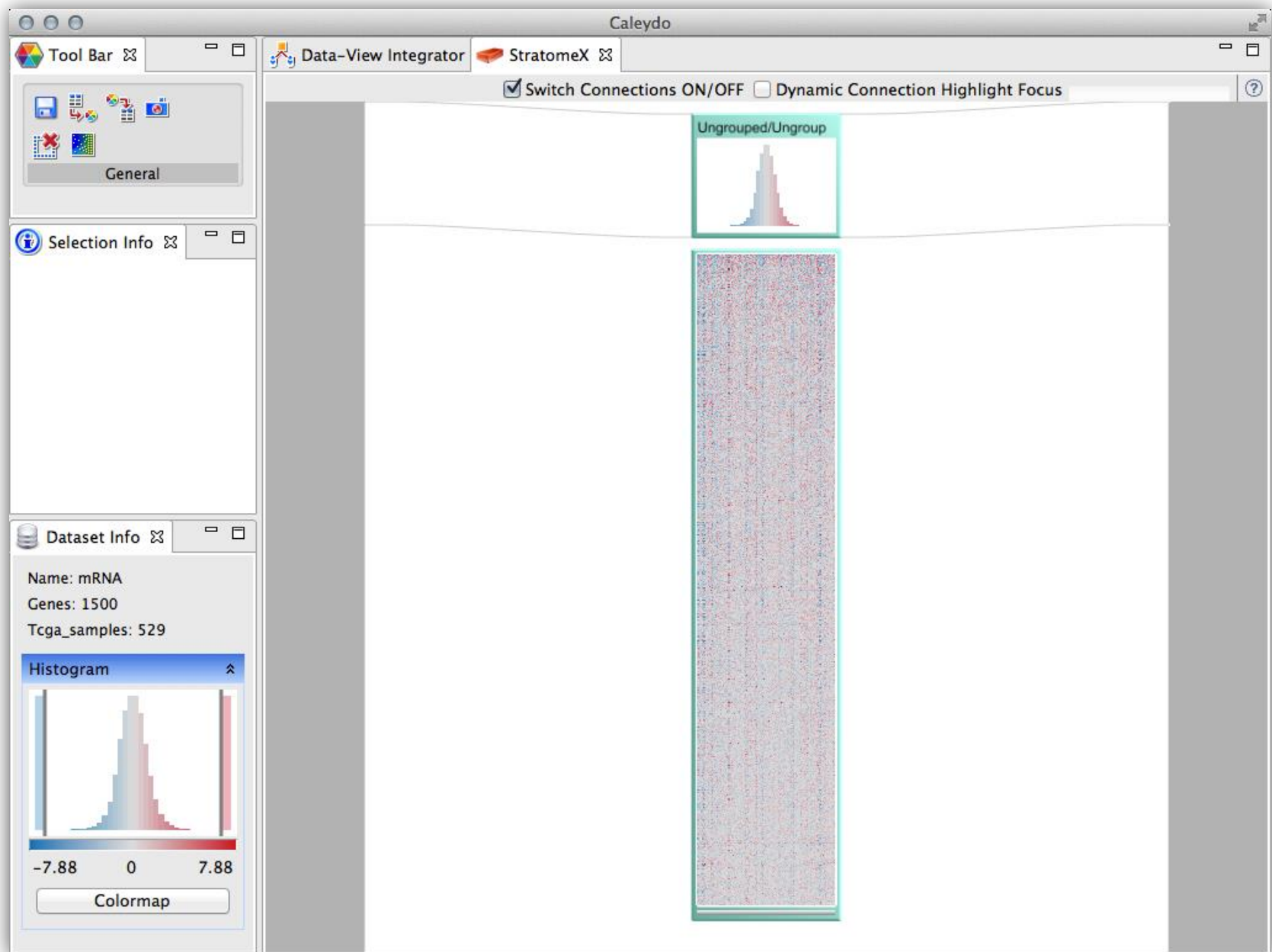
GETTING STARTED



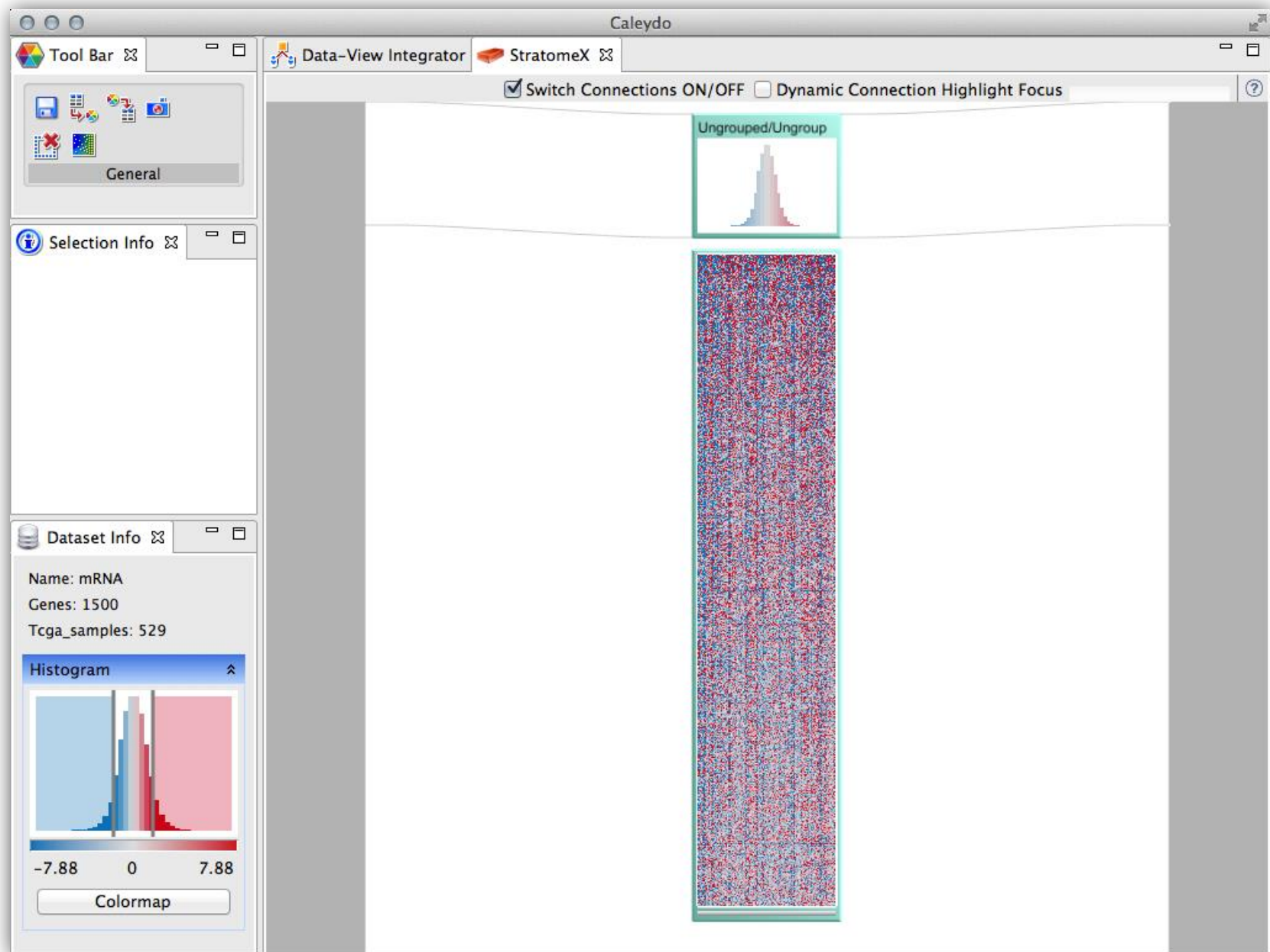
Loaded data at the bottom – available views on the top



Use drag-and-drop to add data to StratomeX view



Switch to StratomeX view: heatmap rows = samples, columns = genes



Use interactive histogram to update color mapping

Cancer

Cancer types are not homogeneous

They are divided into **Subtypes**

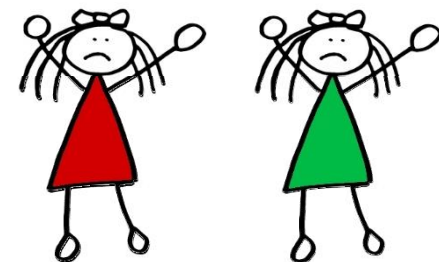
- different histology

- different **molecular alterations**

Subtypes have serious implications

- different **treatment** for subtypes

- prognosis** varies between subtypes



Cancer Subtype Analysis

Done using *many different types of data*,
for *large numbers of patients*.

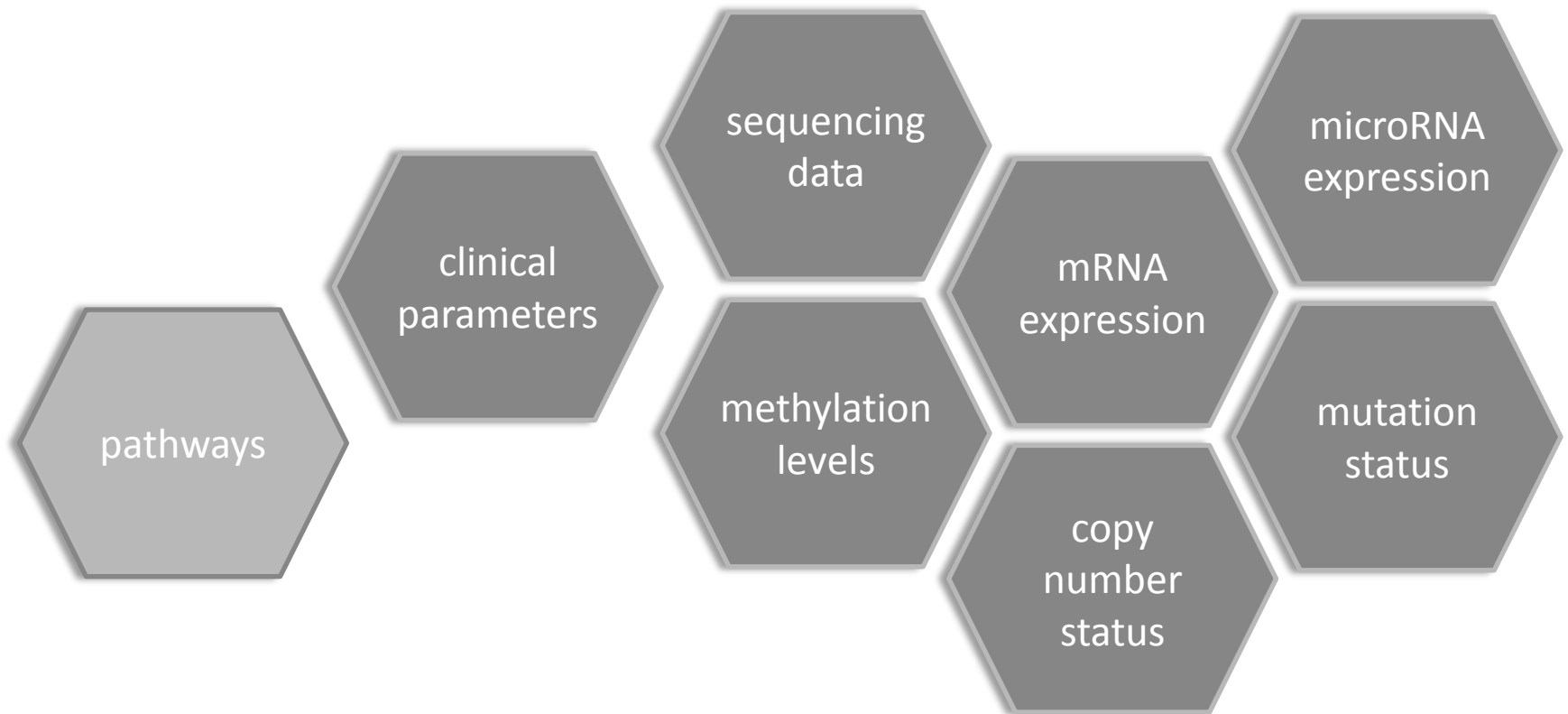
Large-scale project to catalogue genetic mutations responsible for cancer

20 tumor types

500 patient samples each

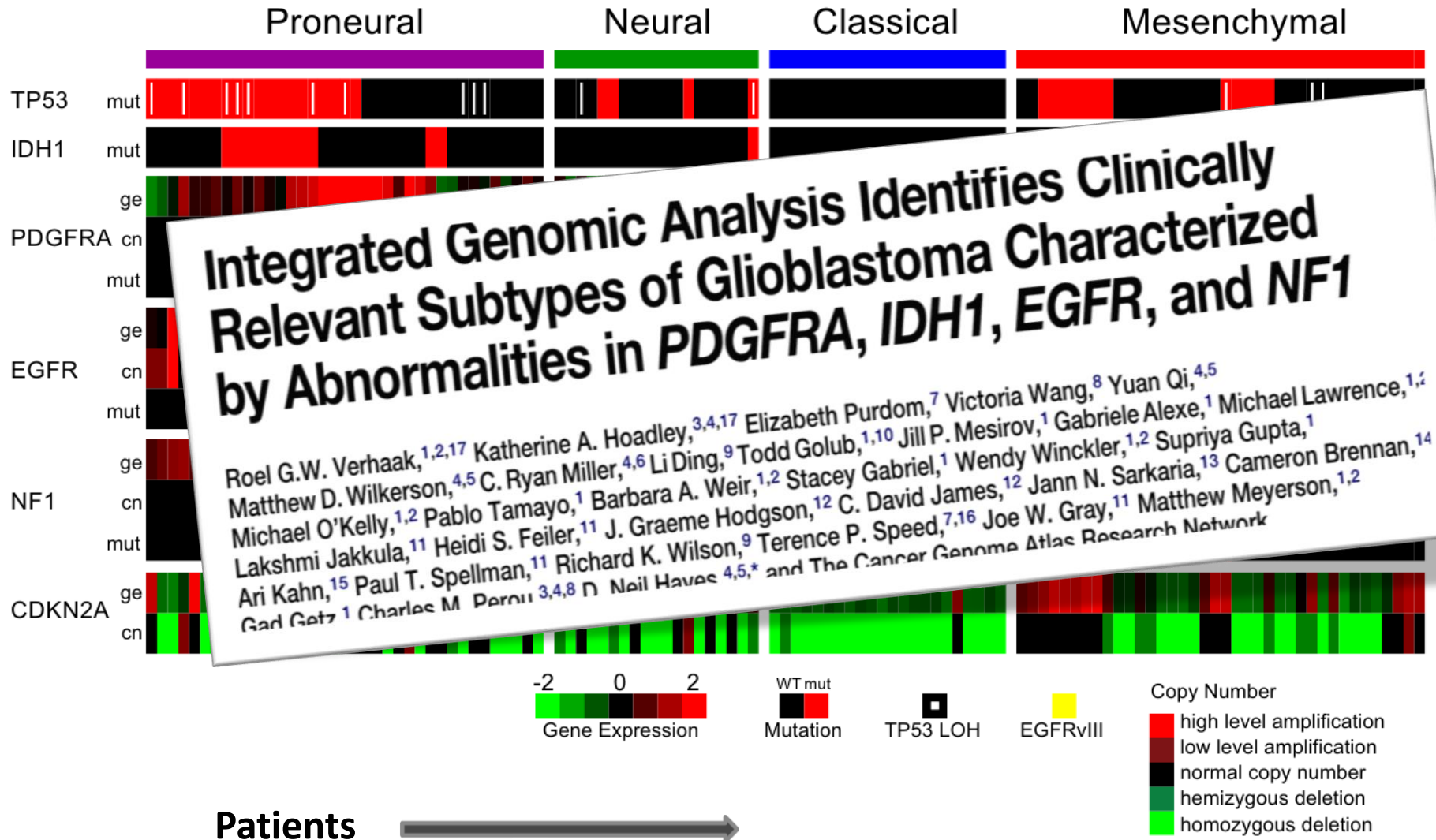
Extensive molecular profiling for each patient

TCGA Data



Our goal is to support
tumor subtype characterization
through integrative
visual analysis of cancer genomics data sets.

Subtype Identification

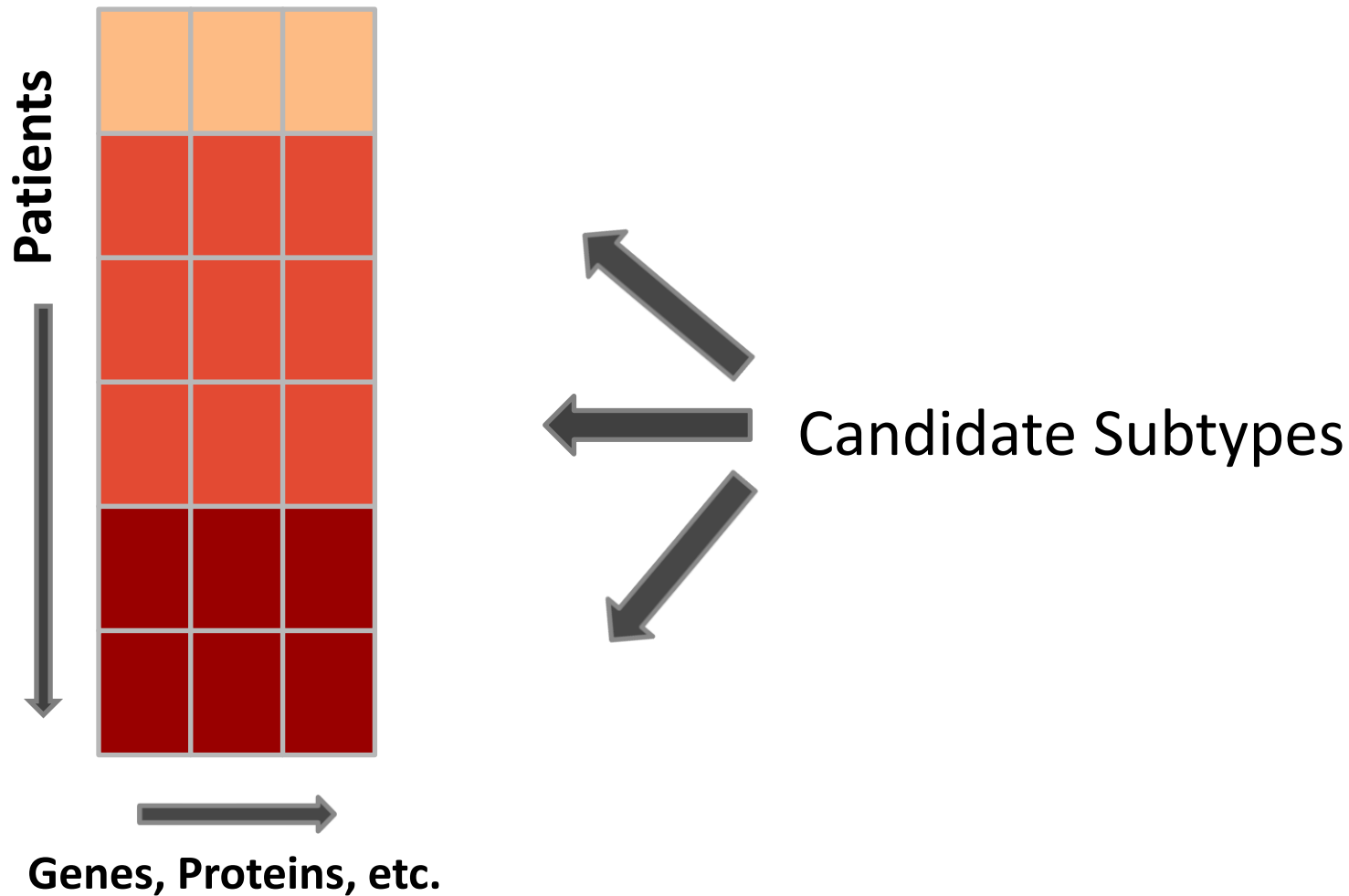


Subtype Identification Process

Step 1: Determine candidate subtypes

Step 2: Find supporting evidence

Tabular Data

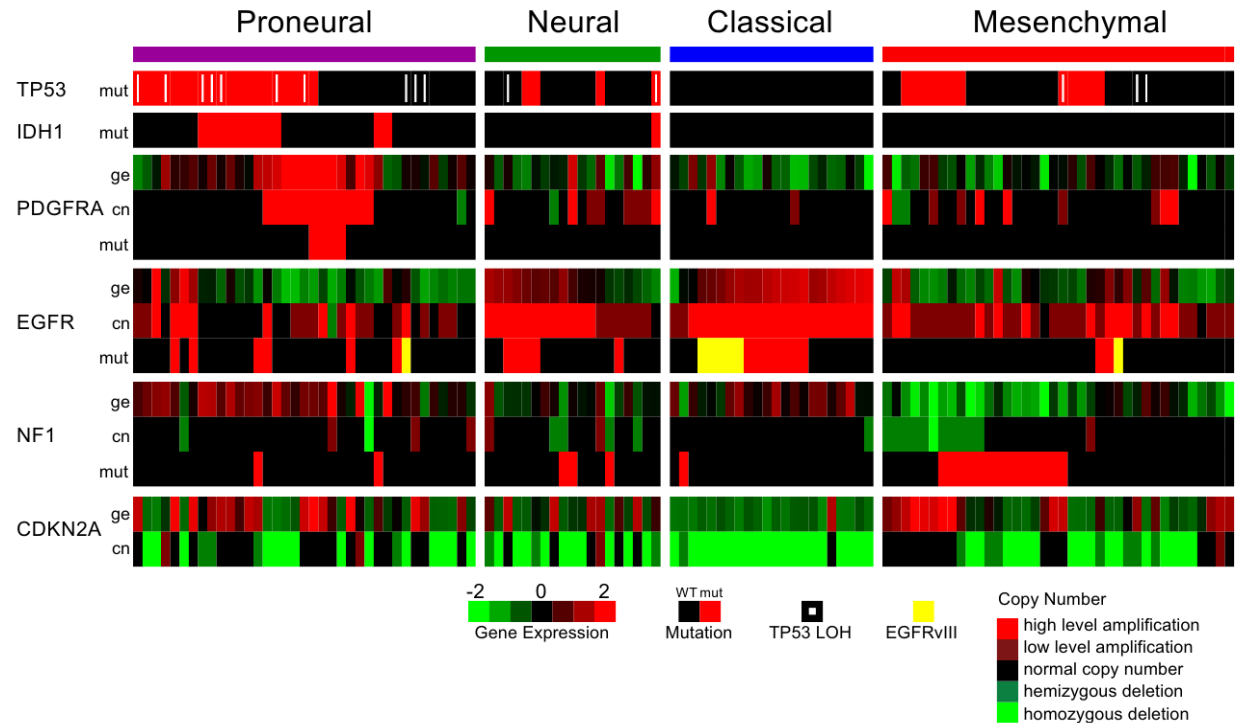


Stratification of a Single Dataset

Cluster A1

Cluster A2

Cluster A3

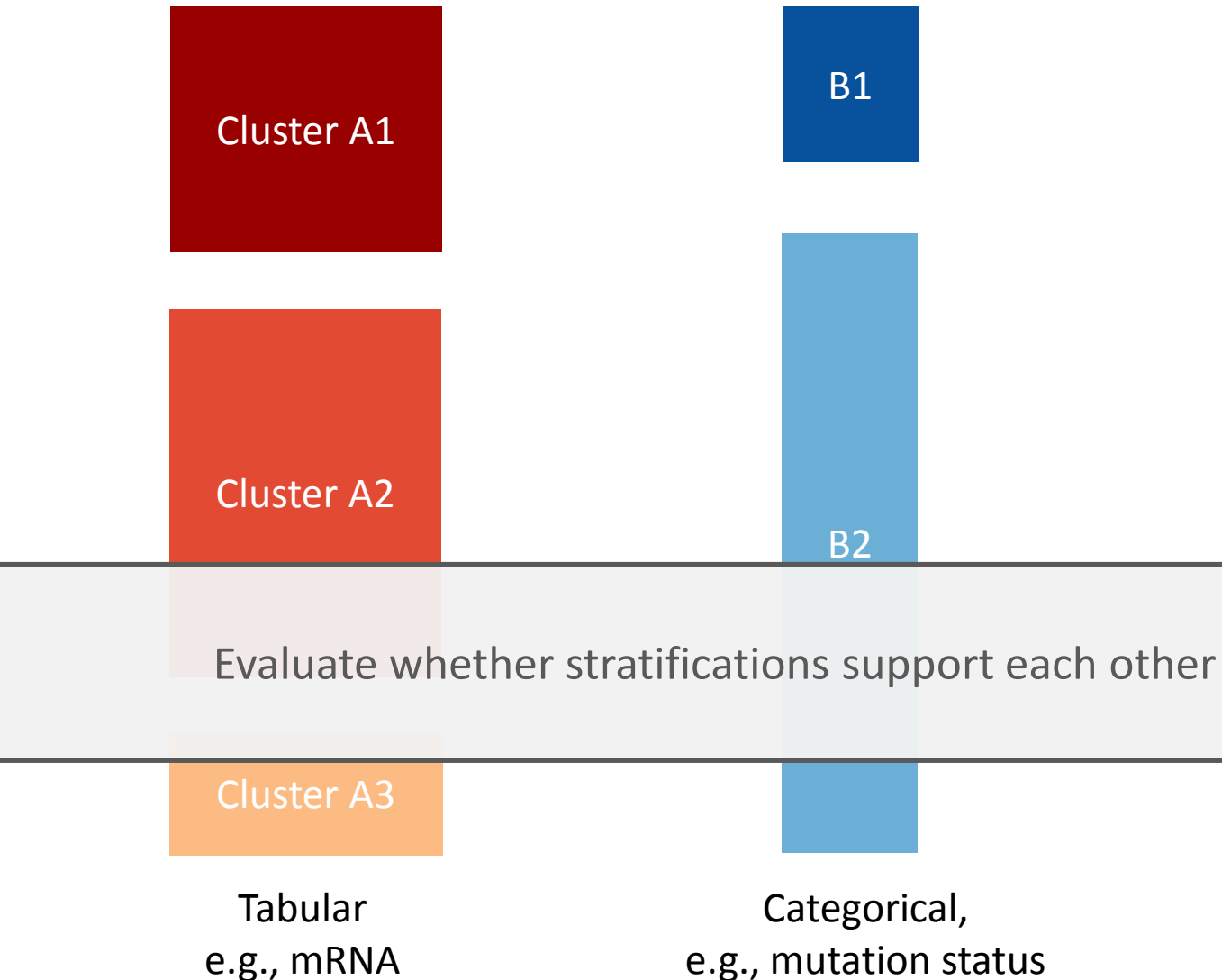


Subtype Identification Process

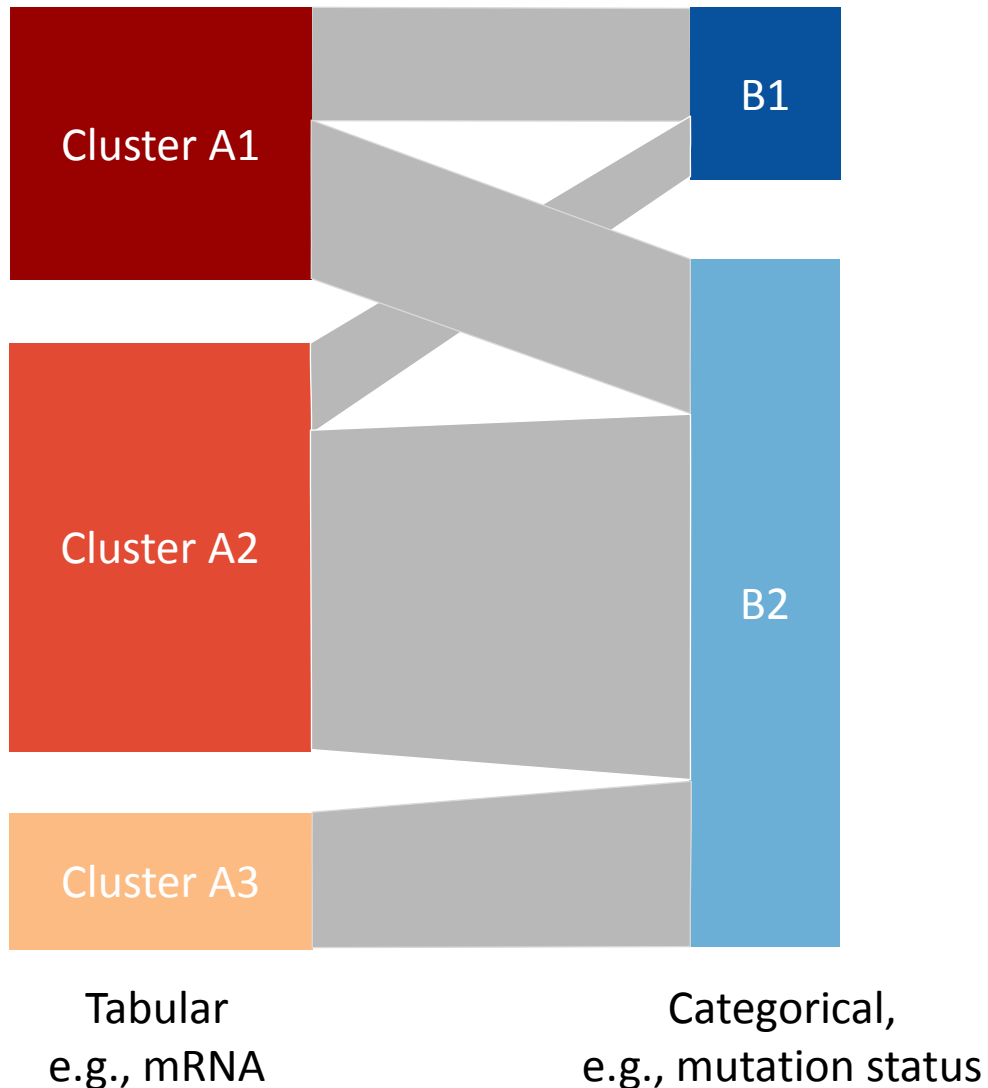
Step 1: Determine candidate subtypes

Step 2: Find supporting evidence

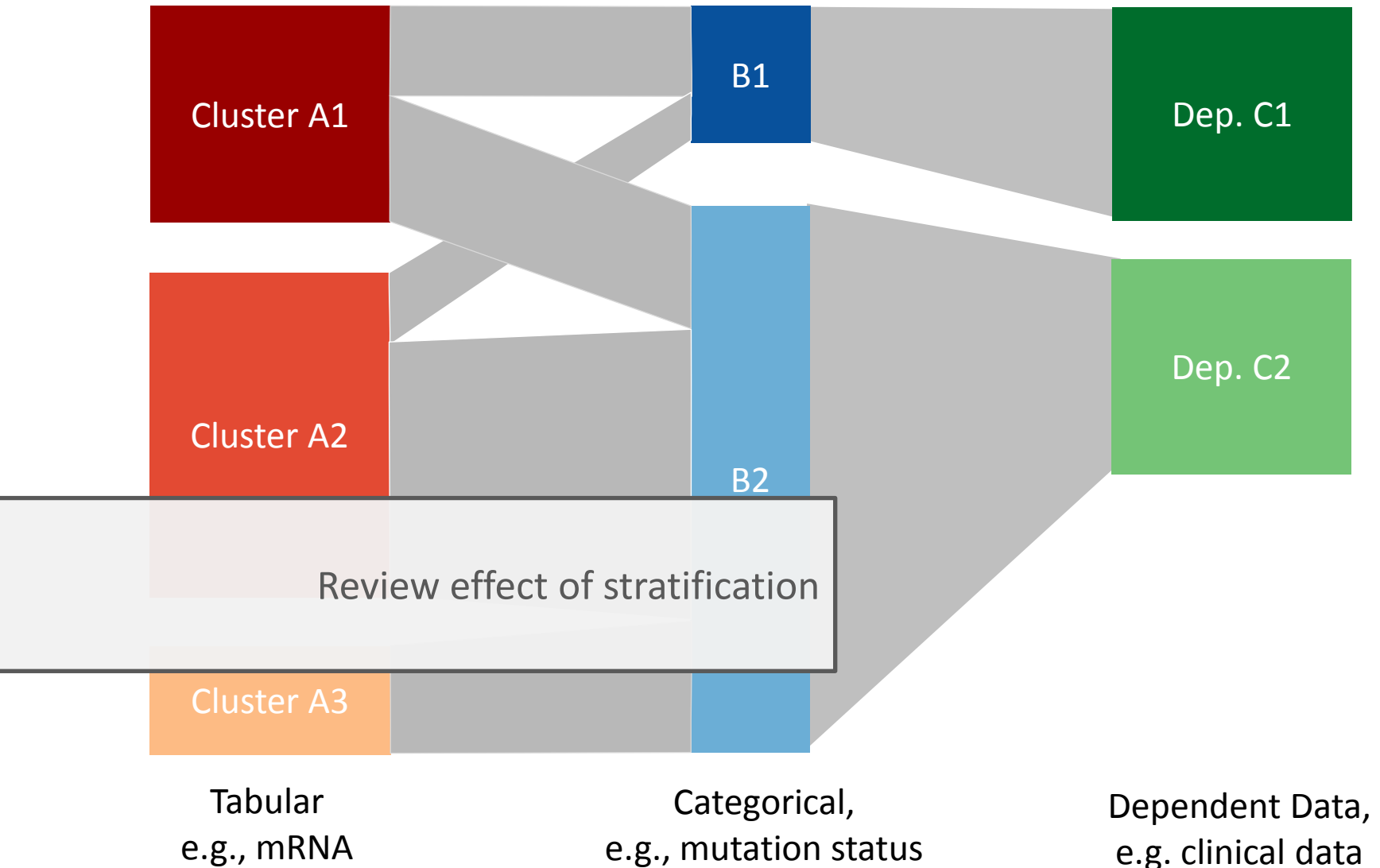
Stratification of Multiple Datasets



Stratification of Multiple Datasets



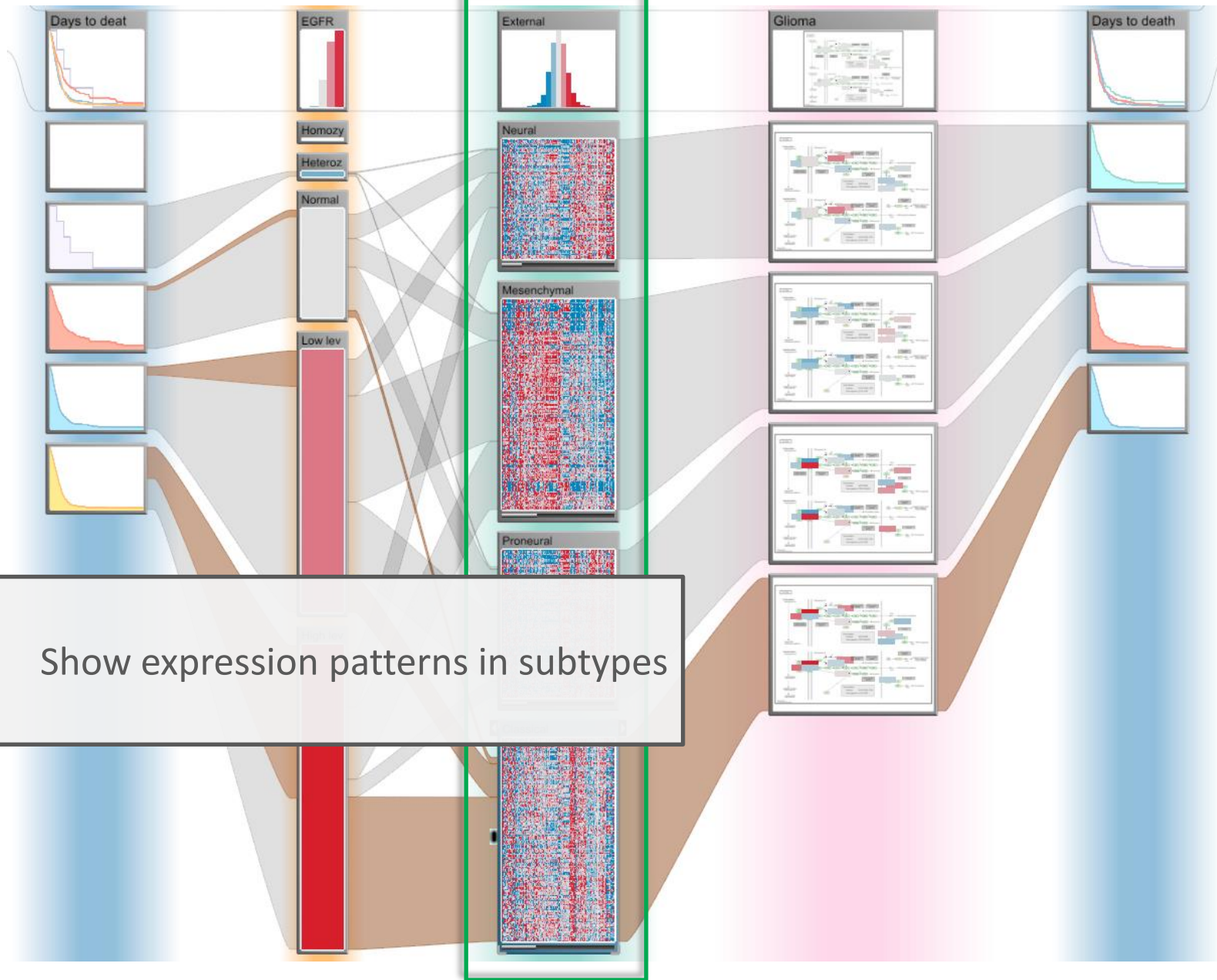
Stratification of Multiple Datasets

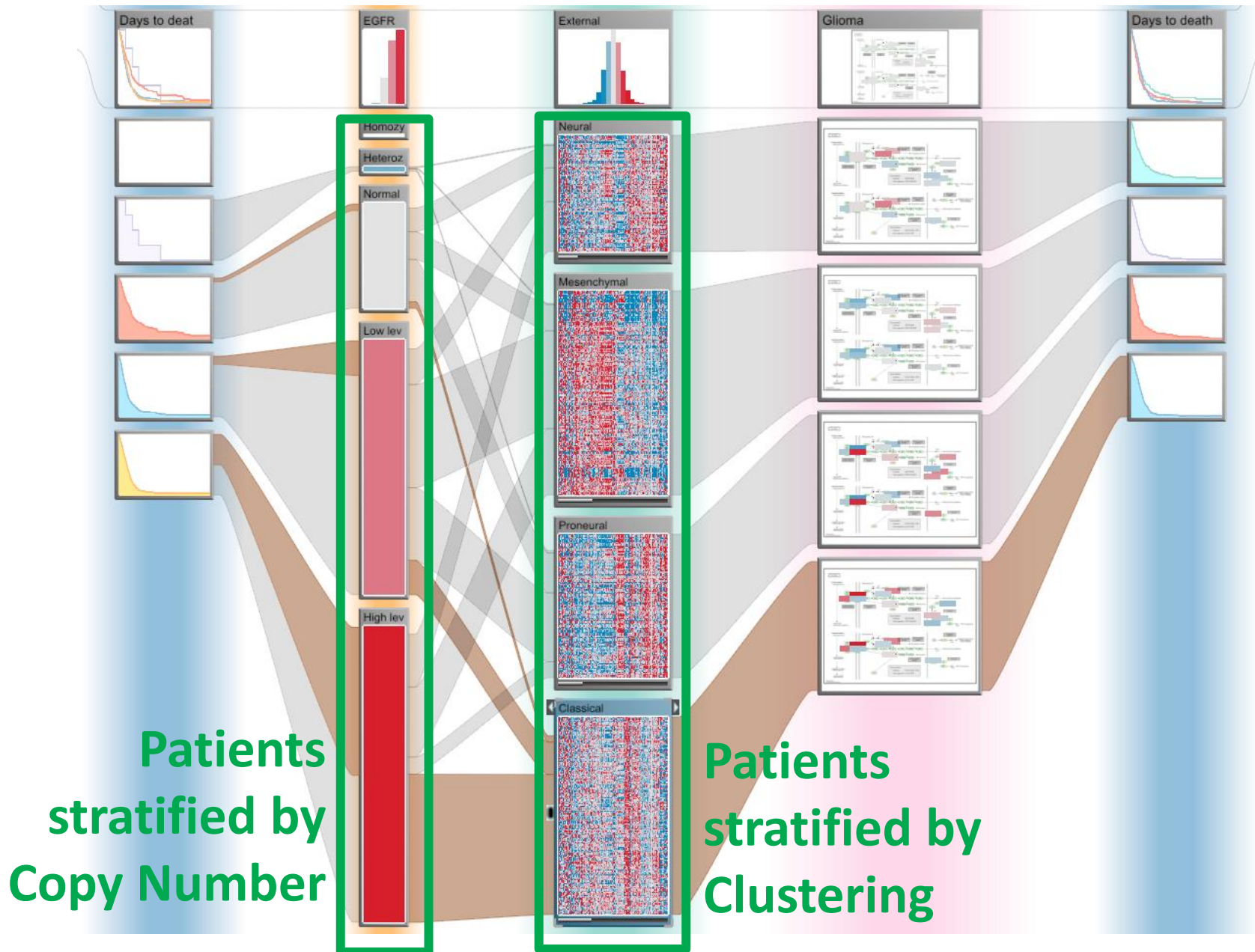


Cate-
gorical

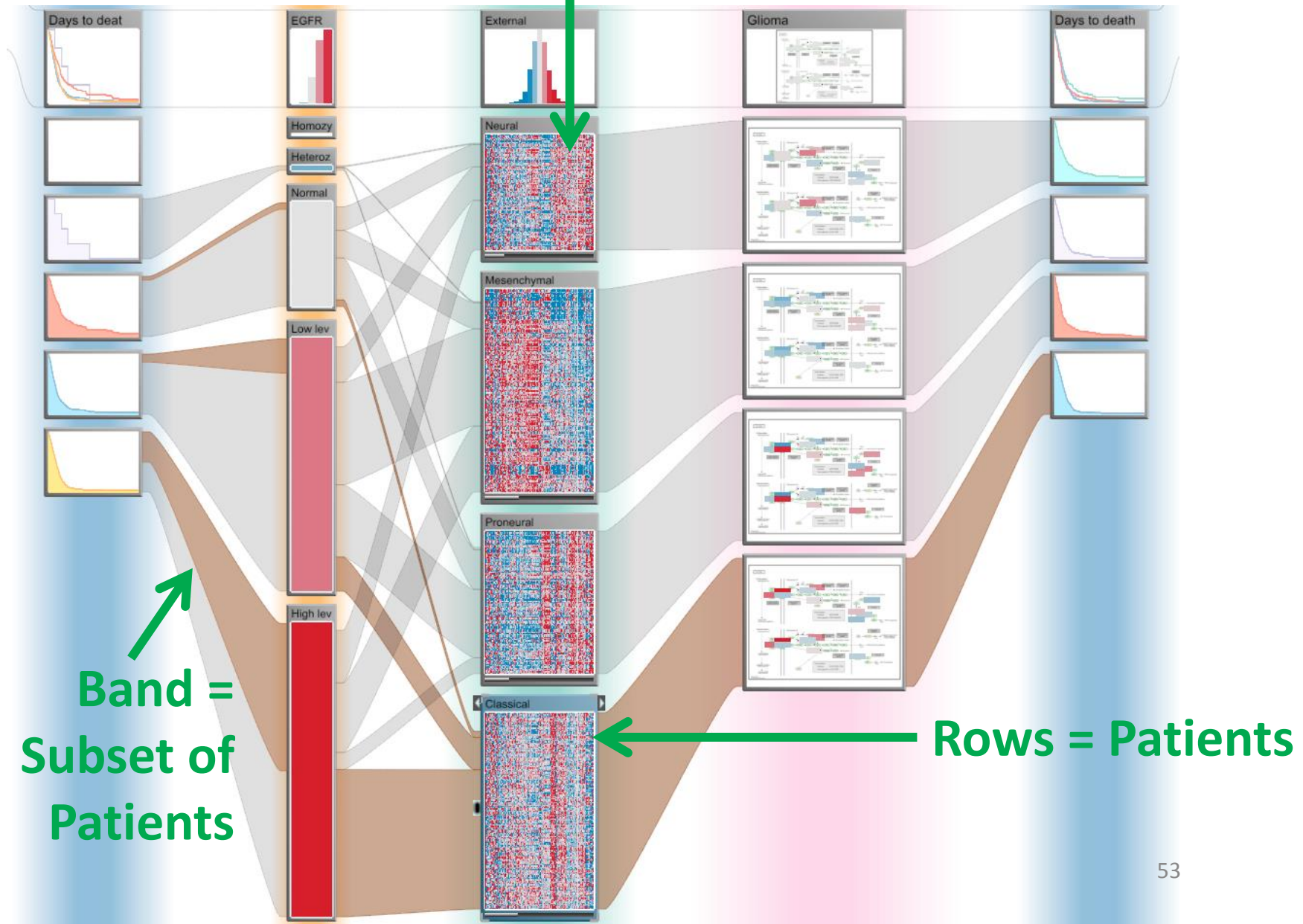
Table

Depen-
dent





Columns = Genes



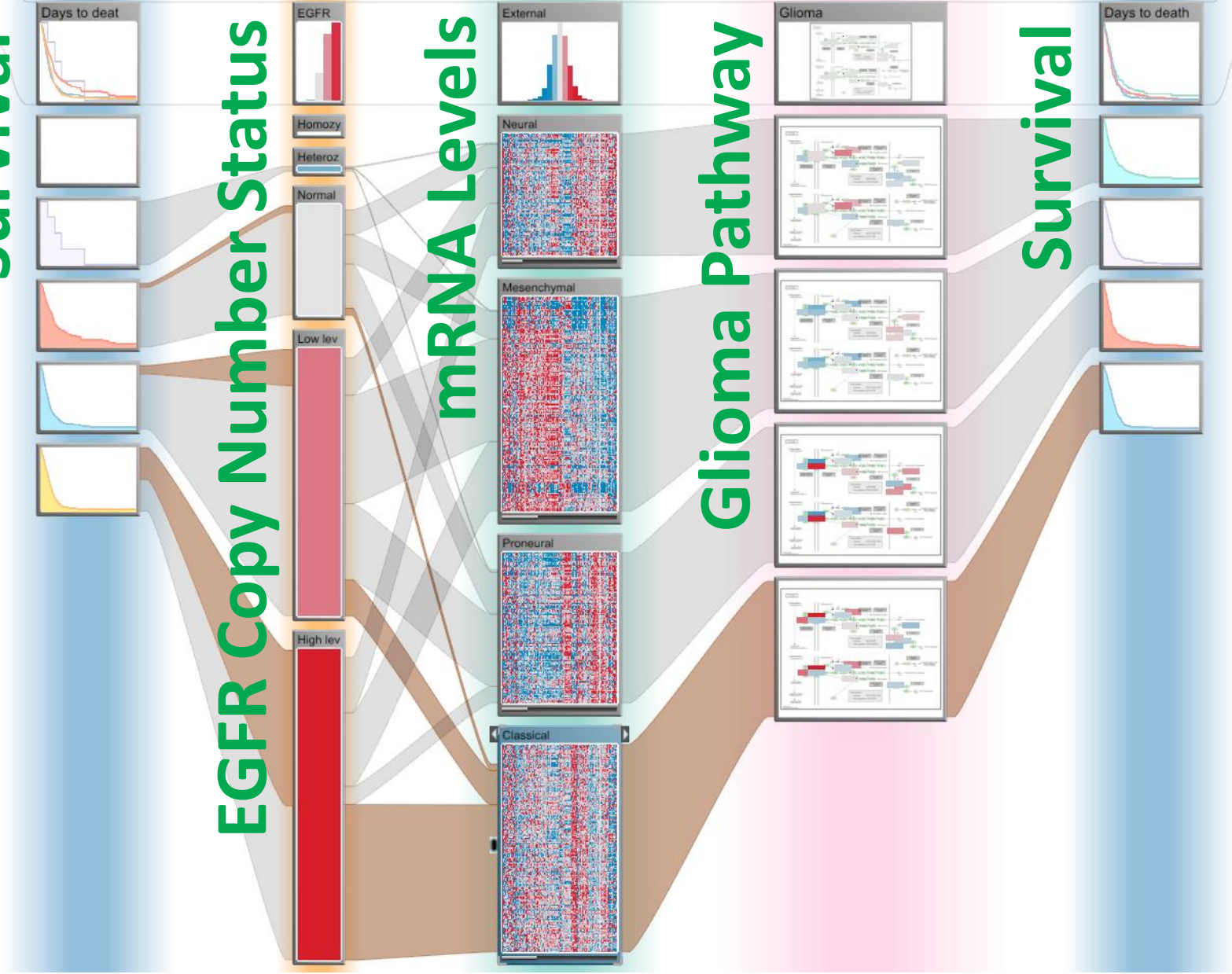
Survival

EGFR Copy Number Status

mRNA Levels

Glioma Pathway

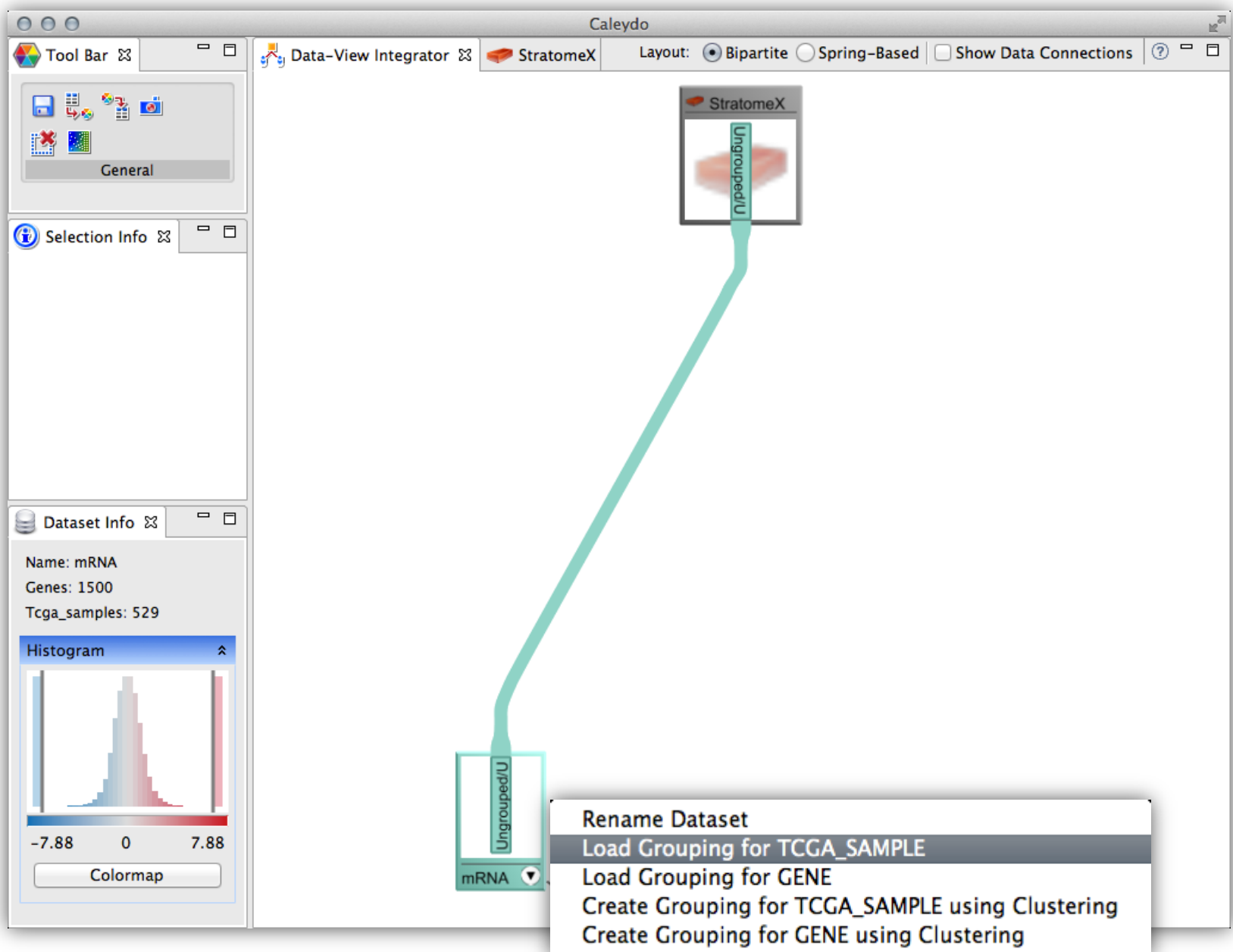
Survival



mRNA Dataset:

Stratification loading / clustering / heatmap
comparison

LIVE DEMO



Right-click on data set box to load a stratification

Import Grouping

Input File

Open Grouping File

Grouping Name

Row Configuration

Row ID Class TCGA_SAMPLE

Row ID Type

Number of Header Rows

Column with Row IDs

Separated by (delimiter)

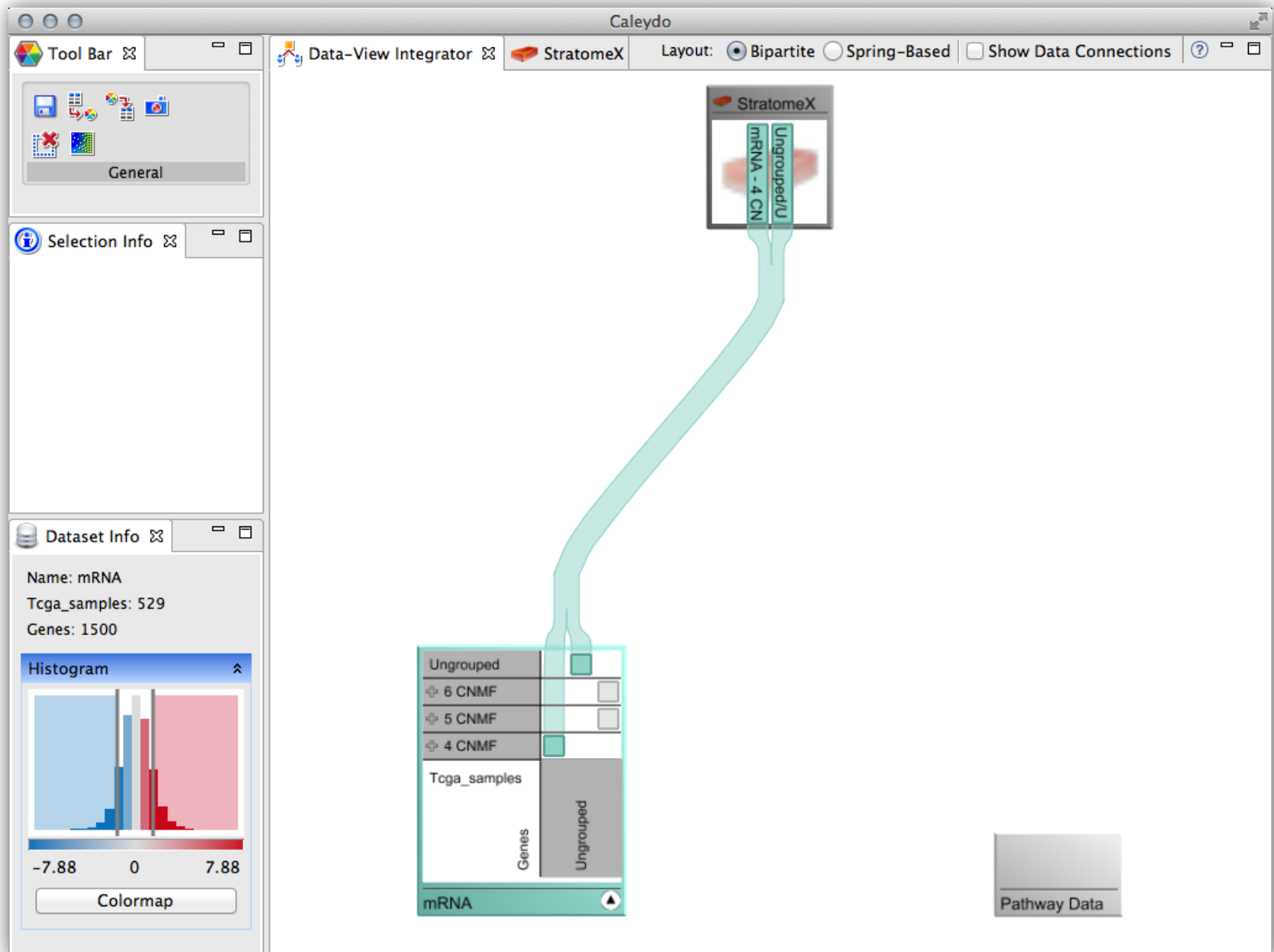
☒ TAB ☐ ; ☐ , ☐ . ☐ SPACE ☐ Other

Column Selection

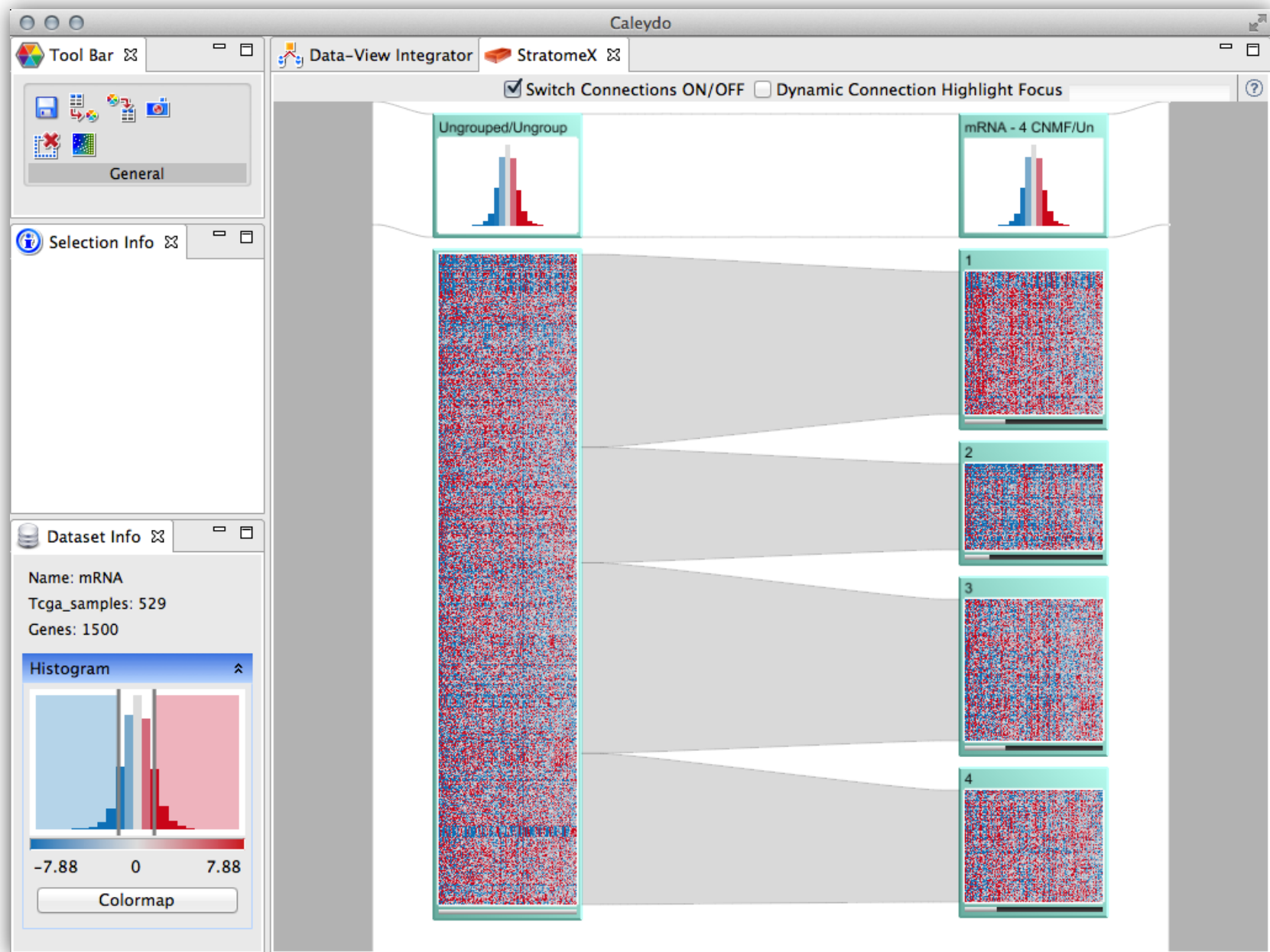
Use column		☐ 2	☒ 3	☒ 4	☒ 5	☐ 6	☐ 7
1	ID	membership	membership.1	membership.2	membership.3	membership.4	membership.5
2	TCGA-02-0...	1	1	1	1	1	1
3	TCGA-02-0...	1	1	1	1	1	1
4	TCGA-02-0...	1	1	2	1	1	1
5	TCGA-02-0...	1	2	1	1	1	1
6	TCGA-02-0...	1	1	1	1	1	1
7	TCGA-02-0...	1	2	1	1	1	1
8	TCGA-02-0...	1	2	1	1	1	1
9	TCGA-02-0...	2	2	3	2	2	2
10	TCGA-02-0...	1	2	3	3	3	3
11	TCGA-02-0...	3	3	4	4	4	4
12	TCGA-02-0...	1	2	1	1	1	1
13	TCGA-02-0...	3	3	4	4	4	5
14	TCGA-02-0...	1	2	1	1	1	1
15	TCGA-02-0...	1	1	2	1	1	1
16	TCGA-02-0...	2	2	3	2	2	2
17	TCGA-02-0...	1	1	2	1	1	1

7 of 7 Columns shown | ☒ Show all Columns

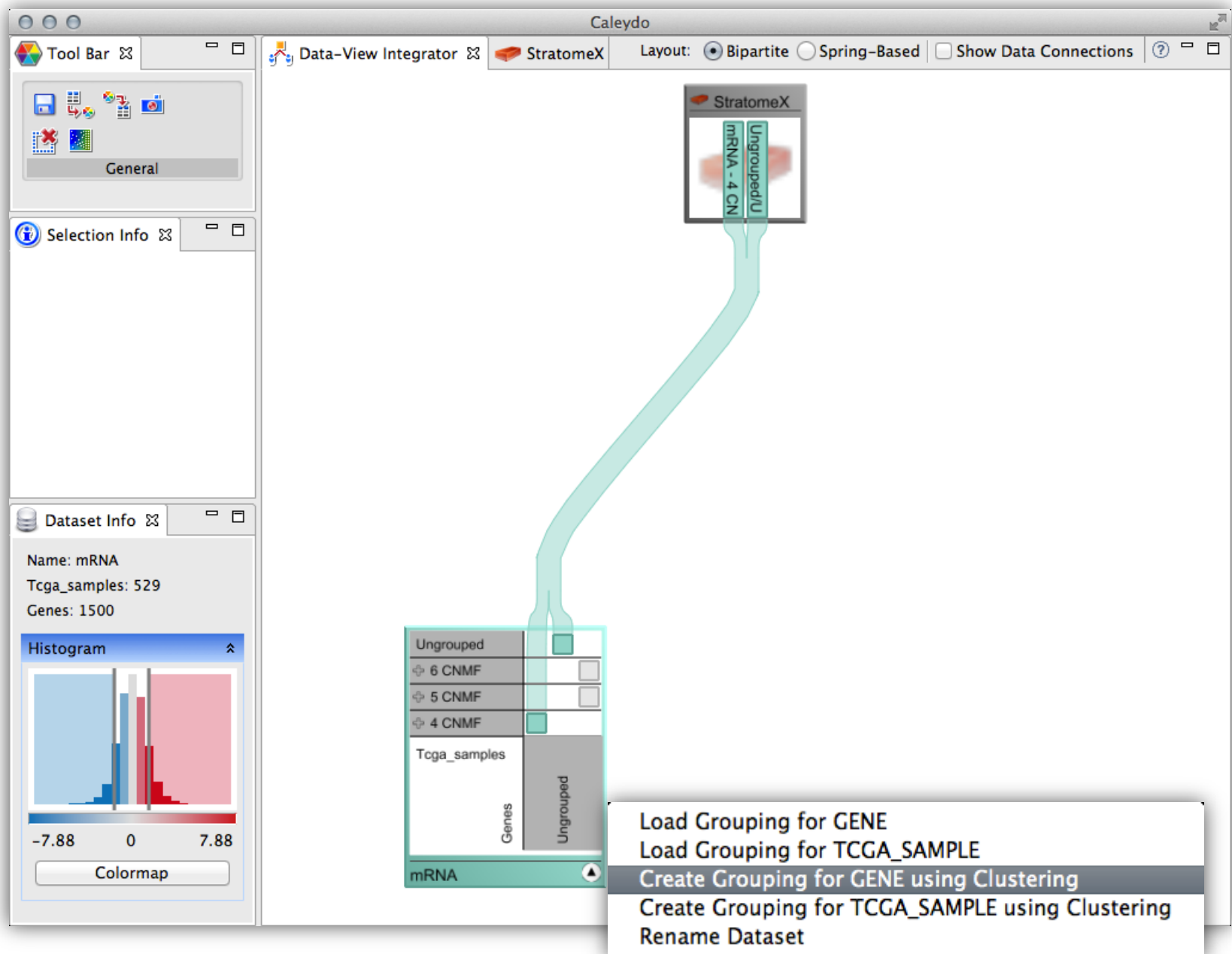
Choose one or more stratifications



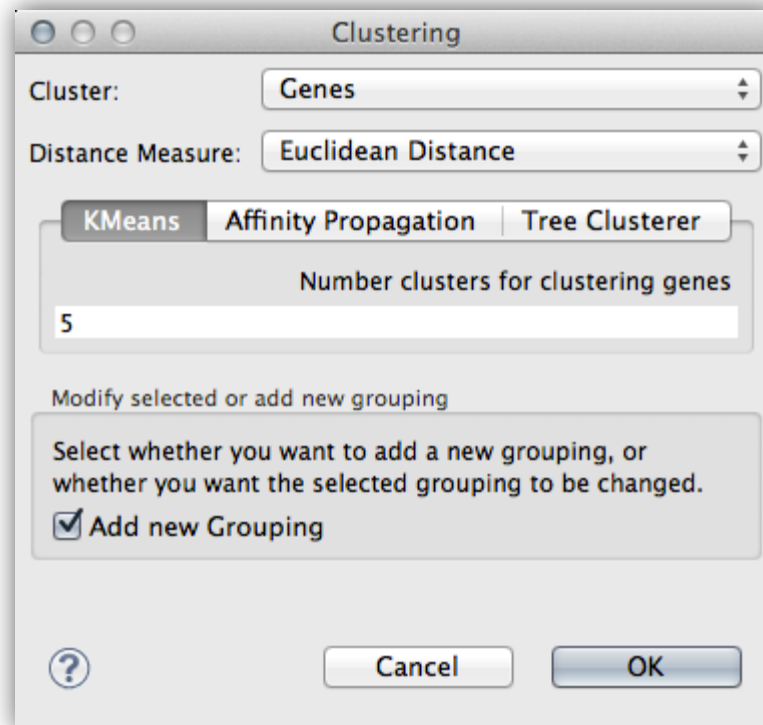
Click arrow on data set box to see new stratifications and drag onto StratomeX view



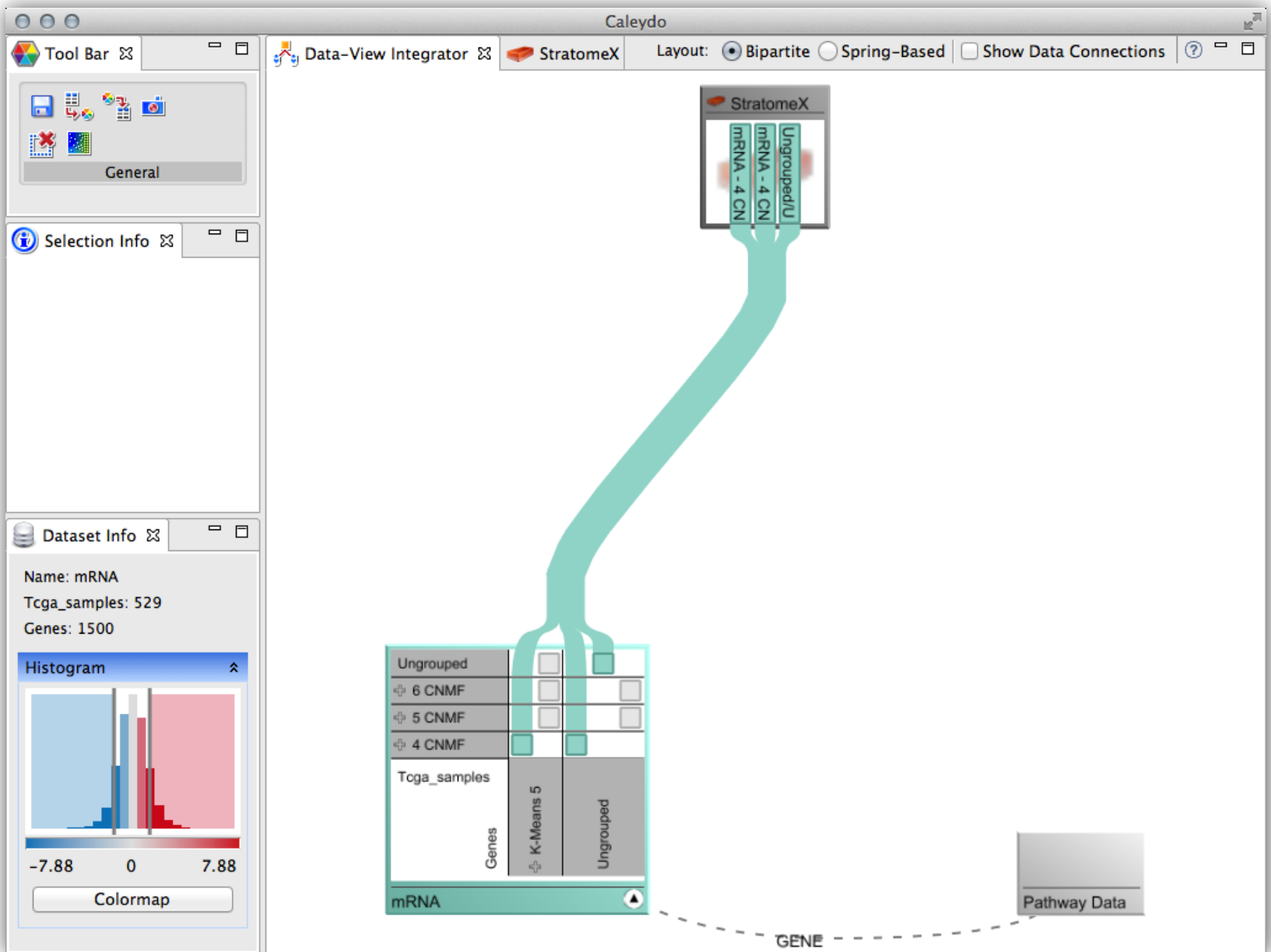
Switch to StratomeX view



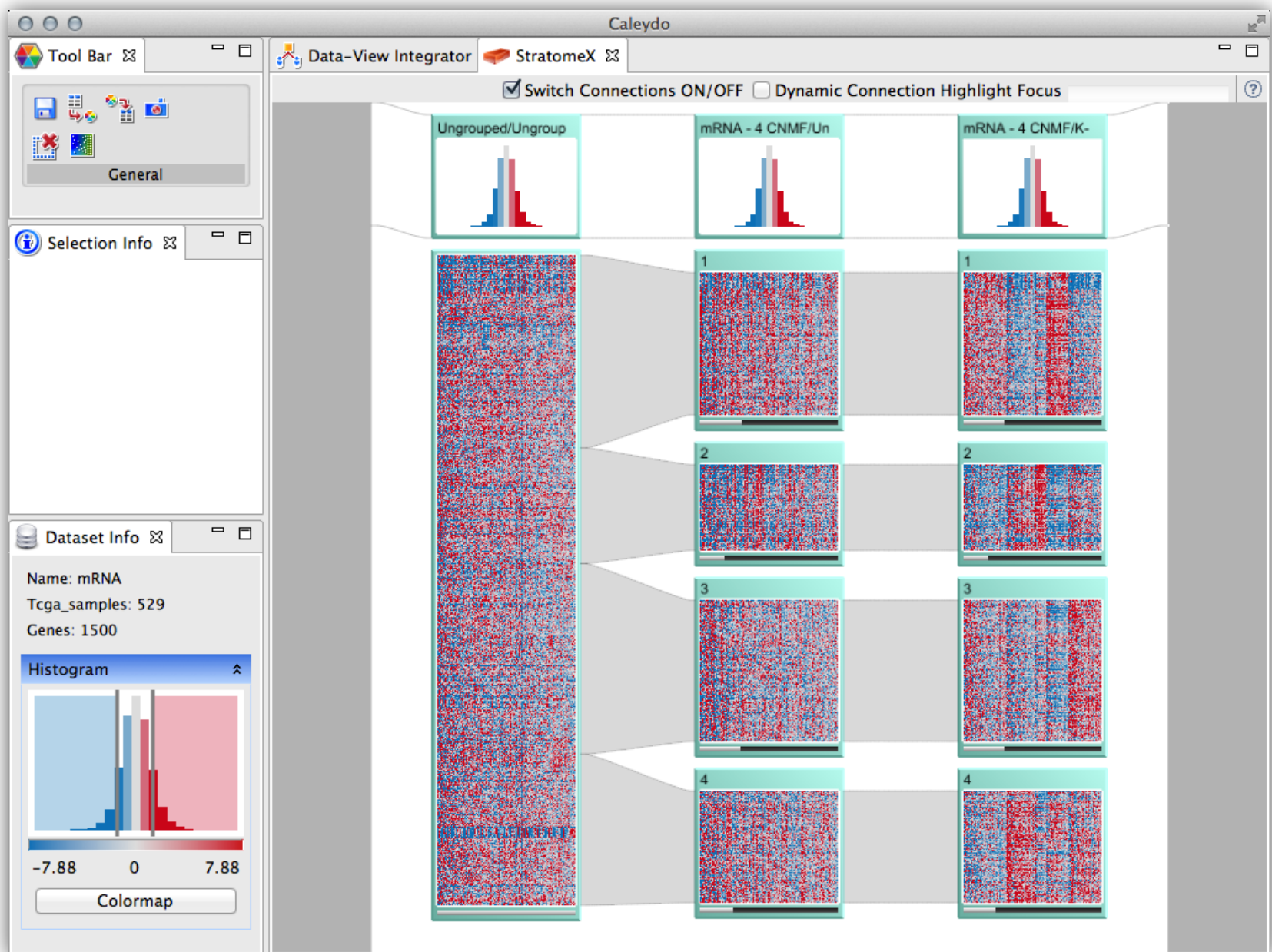
Right-click on data set box to cluster genes



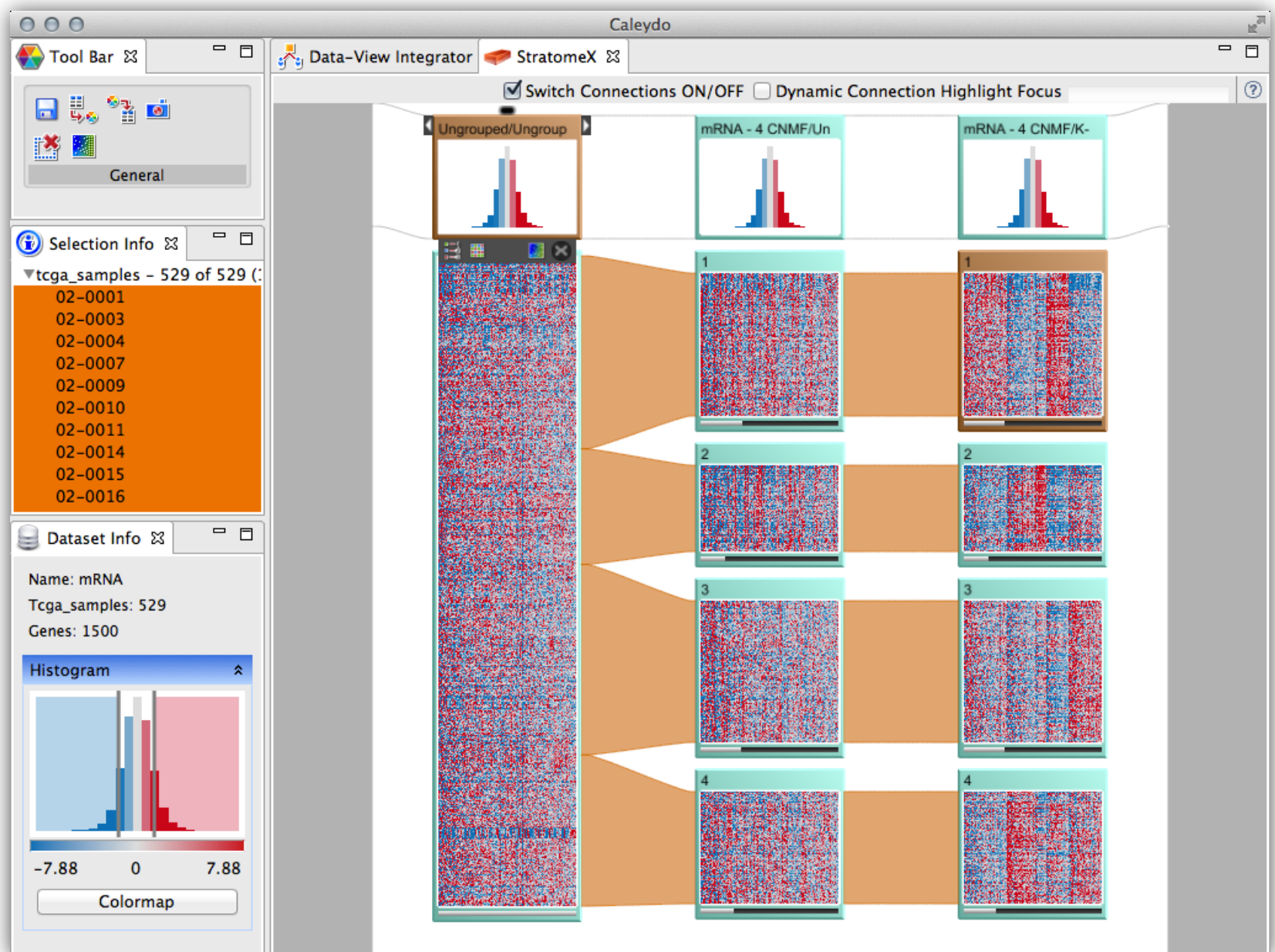
Run with defaults



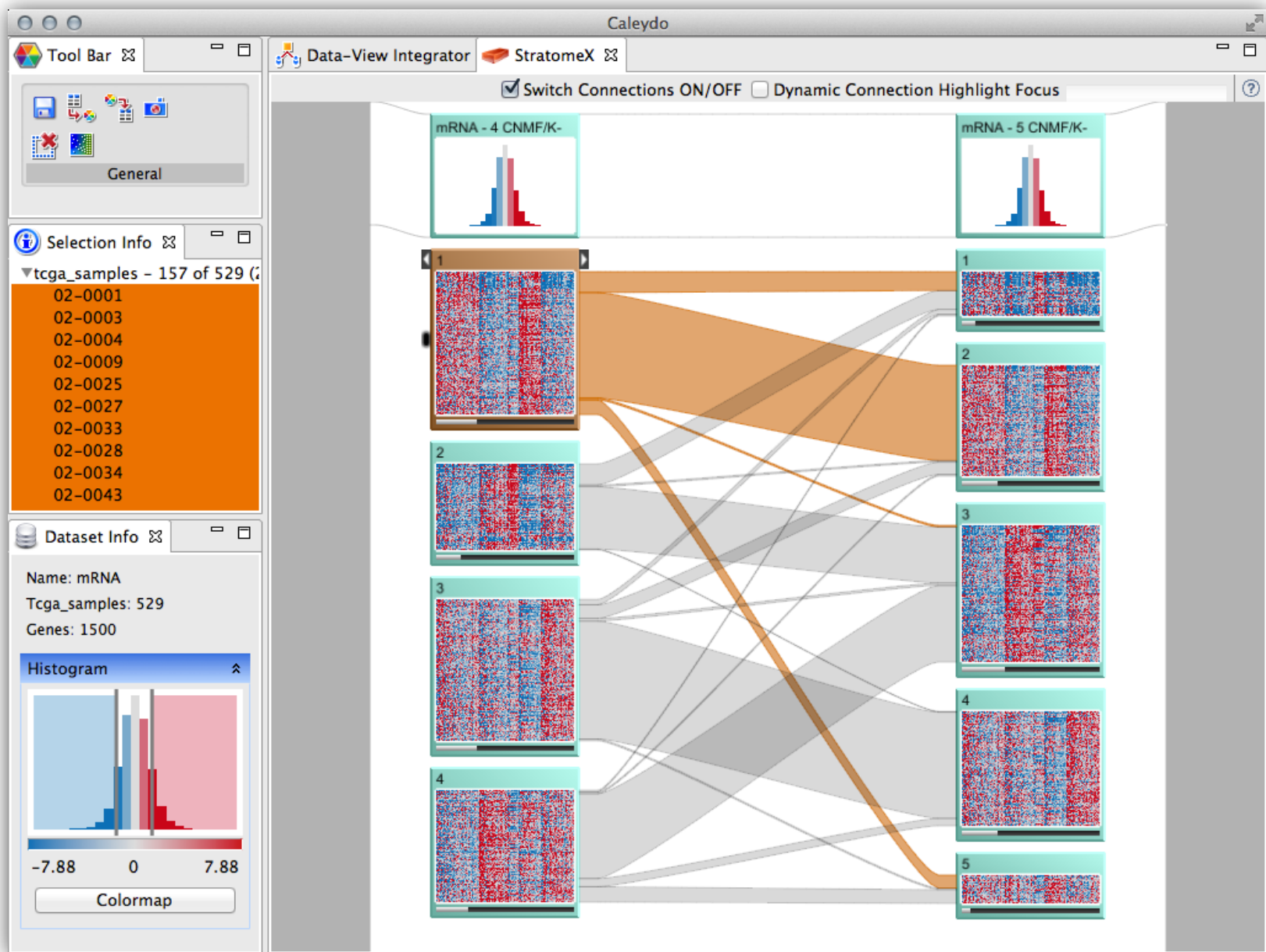
**A new column “K-Means 5” appears in the gene dimension
Drag stratification with 4 clusters to view**



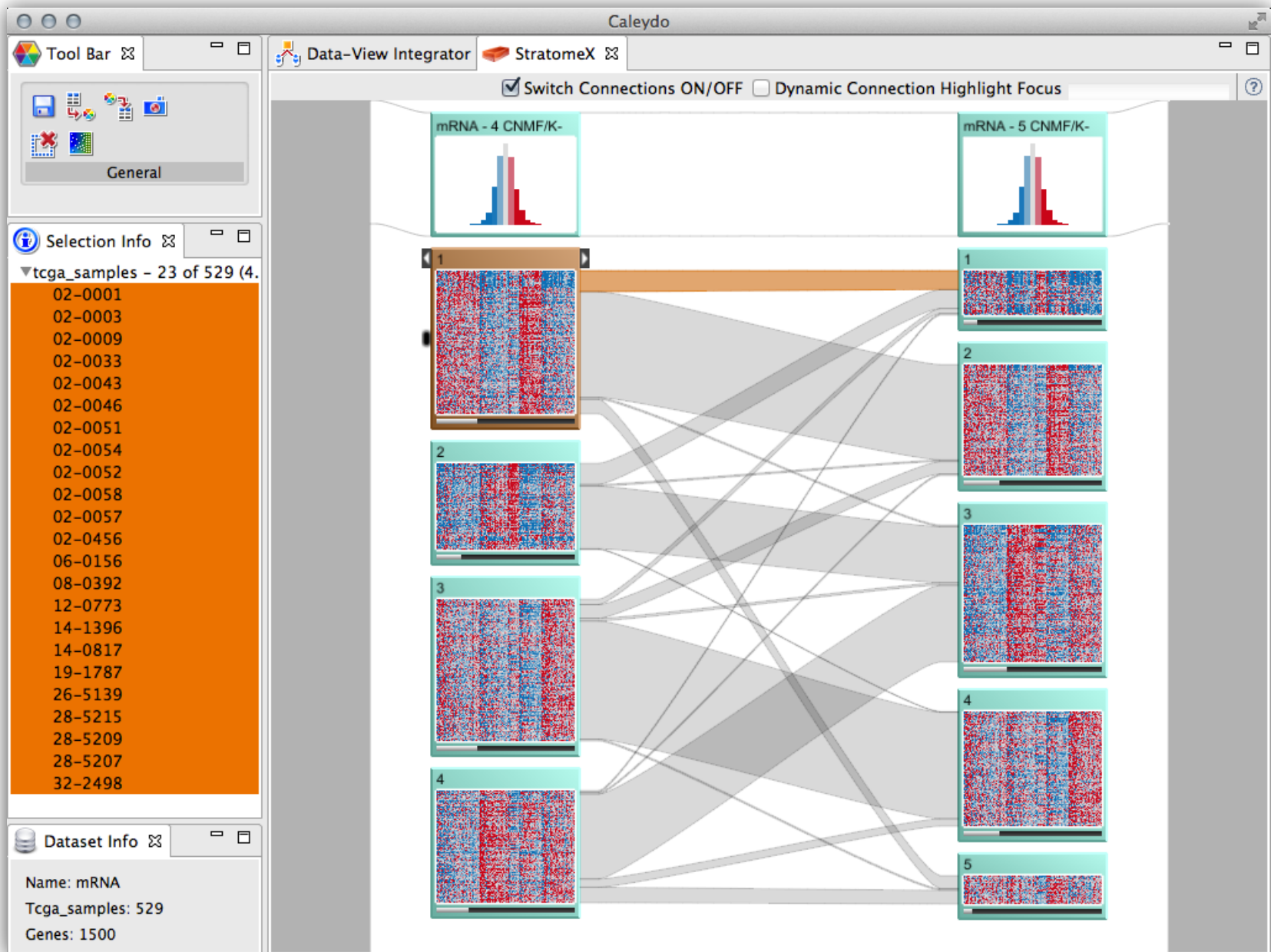
Compare heatmaps



Hover over header brick and click on “X” in menu bar below to remove the column

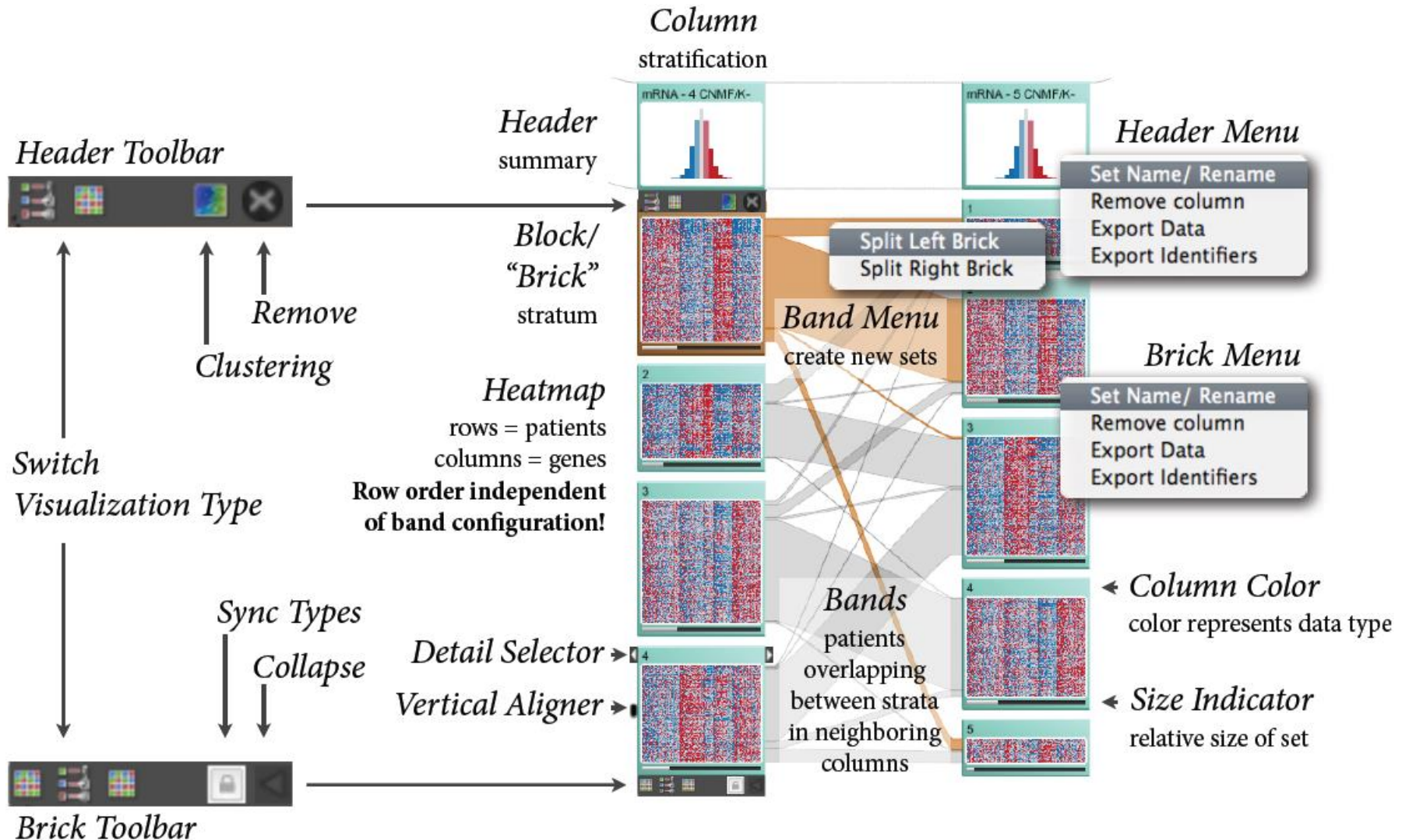


Compare clusterings with 4 and 5 clusters
Click on brick to highlight all patients in this set



Click on band to highlight all patients in the intersection of two sets
Use “Selection Info” view to see the list of patient ids

Components of the StratomeX view



GETTING MORE DATA

Import your own data

Data Import Wizard

Load Dataset
Specify the dataset you want to load.

Input File
Choose Data File... /home/alexs/b/Dropbox/TCGA GDAC/Omics

Dataset Name
GBM.medianexp

Row Configuration
Row ID Class: GENE
Row ID Type: GENE_SYMBOL
Number of Header Rows: 2
Column with Row IDs: 1

Column Configuration
Column ID Class: TCGA_SAMPLE
Column ID Type: TCGA_SAMPLE
Row with Column IDs: 1
☒ Columns use same Scale

Separated by (delimiter)
☒ TAB ☐ ; ☐ , ☐ . ☐ SPACE ☐ Other

Use column		2	3	4	5	6	7
1	Hybridization R	TCGA-02-0001	TCGA-02-0003	TCGA-02-0004	TCGA-02-0007	TCGA-02-0009	TCGA-02-0010
2	Composite Ele	Signal	Signal	Signal	Signal	Signal	Signal
3	AACS	0.5141497264	0.4083453003	1.0383928087	1.0836972281	1.4622804132	1.8951899
4	FSTL1	2.7432607737	3.6635003116	5.7200949805	-1.1581416311	4.6273377963	2.8307503
5	ELMO2	-0.4750396240	0.0830813038	0.7122830511	-0.8727501044	0.4079185628	1.4515957
6	CREB3L1	-1.1034492518	-1.2946236287	-0.2270108466	-0.2845884812	-0.8795445794	-0.0134788
7	RPS11	4.9983821656	4.6803448505	4.0969186092	4.3741099540	4.4245101714	4.9007129
8	PNMA1	1.5487908925	0.8670330481	2.7926044547	2.2529230516	0.7301441750	1.9435549
9	MMP2	2.6876081283	3.2176903345	4.9090056741	-1.6736736037	3.2394738023	-1.5990954
10	SAMD4A	-0.9538500010	-1.1039392598	-0.7803396410	-0.9272558396	-1.0478434536	-1.1306851
11	SMARCD3	-1.2754315263	-0.8031662482	1.9287820058	-1.6627241164	-1.2605498216	2.5381451
12	A4GNT	-0.8779238902	-1.7822941328	-2.2870781166	-1.2790108426	-2.0081527046	-1.3711382

< Back Next > Finish Cancel

Data Import Wizard

Input File
Open Grouping File /home/alexs/b/Dropbox/TCGA GDAC/Omics Integration/testd2

Grouping Name

Row Configuration
Row ID Class: TCGA_SAMPLE
Row ID Type: TCGA_SAMPLE
Number of Header Rows: 1
Column with Row IDs: 1

Separated by (delimiter)
☒ TAB ☐ ; ☐ , ☐ . ☐ SPACE ☐ Other

Use column		2	3	4	5	6	7
1	membership	membership.1	membership.2	membership.3	membership.4	membership.5	membership.6
2	TCGA-02-0001	1	1	1	1	1	1
3	TCGA-02-0002	1	1	1	2	2	2
4	TCGA-02-0003	2	2	2	3	3	3
5	TCGA-02-0006	1	3	3	4	4	3
6	TCGA-02-0007	2	2	2	3	5	4
7	TCGA-02-0009	2	2	2	3	1	5
8	TCGA-02-0010	1	3	3	4	4	6
9	TCGA-02-0011	1	3	3	4	4	6
10	TCGA-02-0014	1	3	3	4	4	6
11	TCGA-02-0021	2	2	2	1	1	1
12	TCGA-02-0024	1	3	3	4	4	6

8 of 8 Columns ☐ Show all Columns

OK Cancel

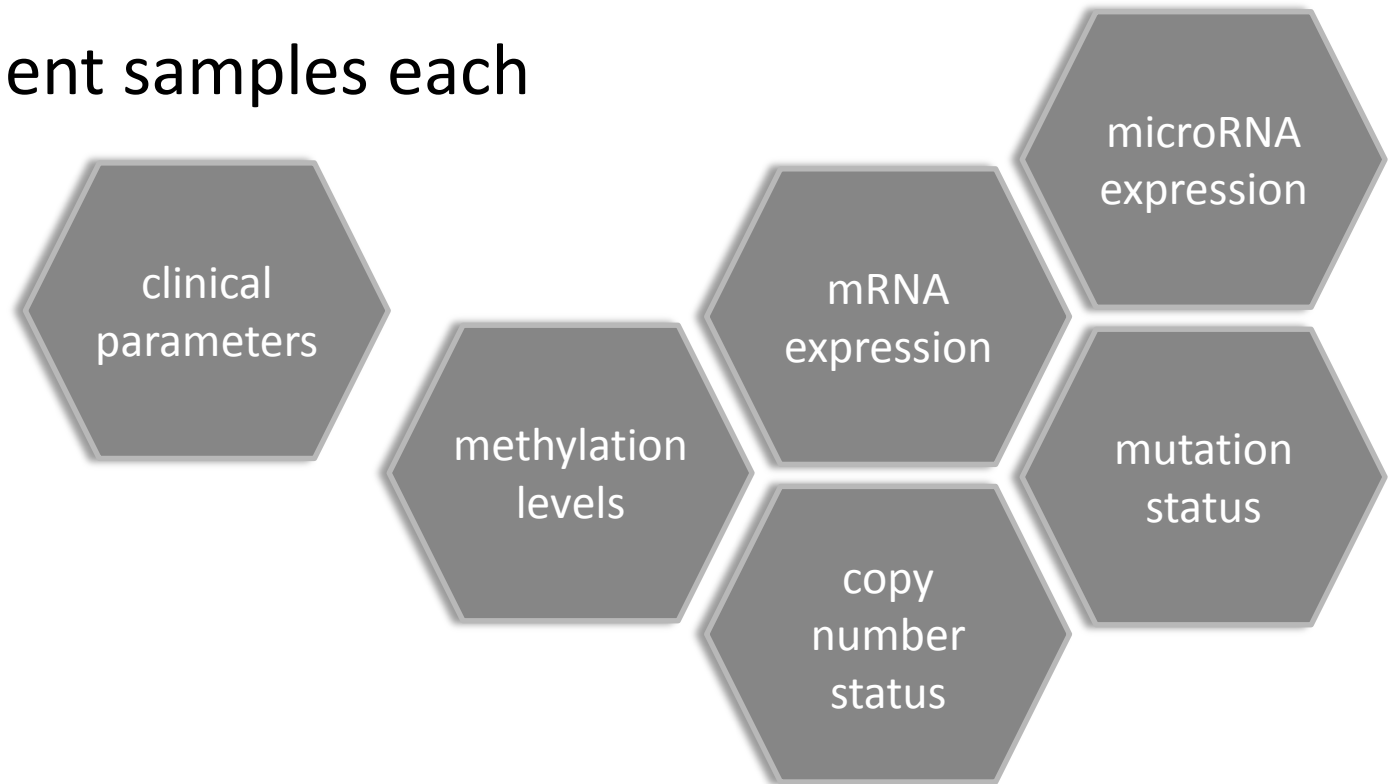
Import data matrices

Import stratifications

TCGA Data

20 tumor types

500 patient samples each





Analysis Results

[illegible]



FIREHOSE
Broad GDAC



Analysis
Results



StratomeX
Data Packaging
(1x per month)



<http://tcga.caleydo.org>

<http://tcga.caleydo.org>



FIREHOSE
Broad GDAC



Analysis
Results



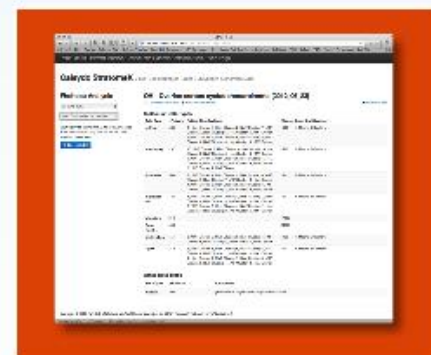
Instead of downloading and dealing with over a dozen archives, a single click is required to load and visualize the data for a tumor type.

StratomeX
Data Packaging
(1x per month)



CALEYDO
StratomeX

Java Web
Start



<http://tcga.caleydo.org>

StratomeX

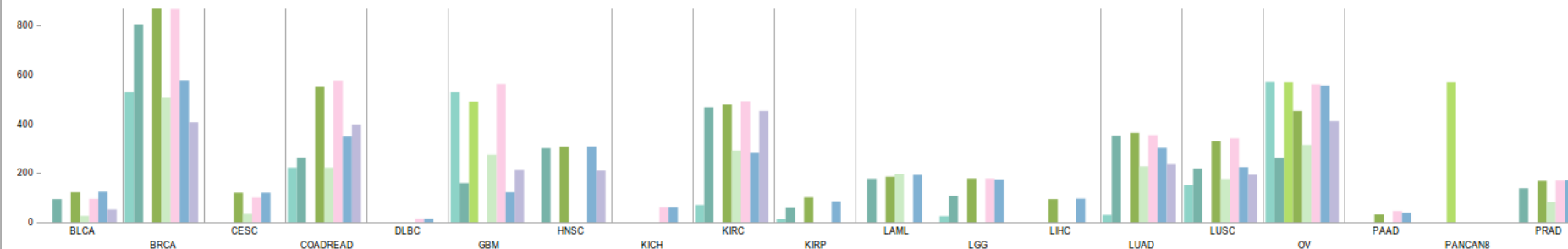
Caleydo StratomeX is a visual exploration tool for identification and characterization of tumor subtypes in cancer genomics data. This website provides StratomeX users with direct access to data for over twenty tumor types studied by the [The Cancer Genome Atlas](#).

The data packages are automatically generated based on the output of monthly analyses performed by the [Firehose pipeline](#) of the TCGA Genome Data Analysis Center at the Broad Institute.

[View Data](#) [Run StratomeX](#) [Read FAQ](#)

Firehose Analysis Run

24 October 2012



Data Type	Clear	All
☒ mRNA	☒	☐
☒ mRNA-seq	☒	☐
☒ microRNA	☒	☐
☒ microRNA-seq	☒	☐
☒ Mutations	☒	☐
☒ Copy Number	☒	☐
☒ Methylation	☒	☐
☒ RPPA	☒	☐

Tumor Type	Clear	All
☒ BLCA	☒	☐
☒ BRCA	☒	☐
☒ CESC	☒	☐
☒ COADREAD	☒	☐
☒ DLBC	☒	☐
☒ GBM	☒	☐
☒ HNSC	☒	☐
☒ KICH	☒	☐
☒ KIRC	☒	☐
☒ KIRP	☒	☐
☒ LAML	☒	☐
☒ LGG	☒	☐
☒ LIHC	☒	☐
☒ LUAD	☒	☐
☒ LUSC	☒	☐
☒ OV	☒	☐
☒ PAAD	☒	☐
☒ PANCAN8	☒	☐
☒ PRAD	☒	☐
☒ SKCM	☒	☐
☒ STAD	☒	☐
☒ SARC	☒	☐

<http://tcga.caleydo.org>

StratomeX

Caleydo StratomeX is a visual exploration tool for identification and characterization of tumor subtypes in cancer genomics data. This website provides StratomeX users with direct access to data for over twenty tumor types studied by the [The Cancer Genome Atlas](#).

The data packages are automatically generated based on the output of monthly analyses performed by the [Firehose pipeline](#) of the TCGA Genome Data Analysis Center at the Broad Institute.

[View Data](#) [Run StratomeX](#) [Read FAQ](#)

Firehose Analysis Run

24 October 2012

Tumor Type

GBM - Glioblastoma multiforme

[▶ Run](#) [⏸](#) [📄](#) [🔍](#) [⚠ Mac Users](#)

Important Notice Please be advised that downloading data from this site constitutes agreement to the *The Cancer Genome Atlas* policies and guidelines on [data usage](#) and [publications](#).

Molecular Data Types

Data Type		#Patients	#Patient Stratifications	#Genes	#Gene Stratifications
mRNA		529	14	1500	1
mRNA-seq		161	16	1500	1
microRNA		491	14	150	1
Mutations		276	0	17848	0
Copy Number		563	0	24174	0
Methylation		124	7	7783	1
RPPA		214	14	150	1

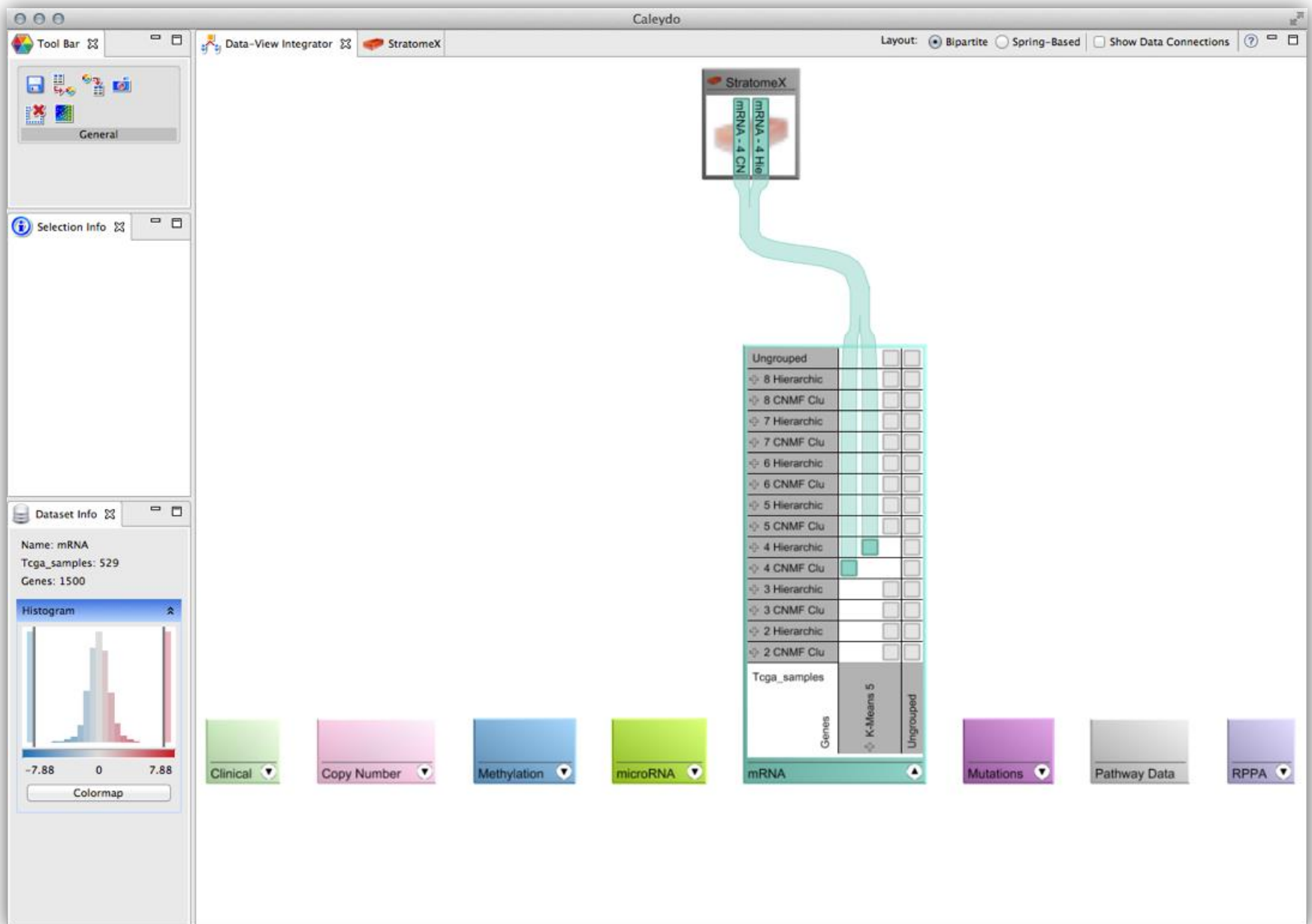
Other Data Types

Data Type	#Patients	Parameters
Clinical	565	yearstobirth; daystodeath; daystolastfollowup

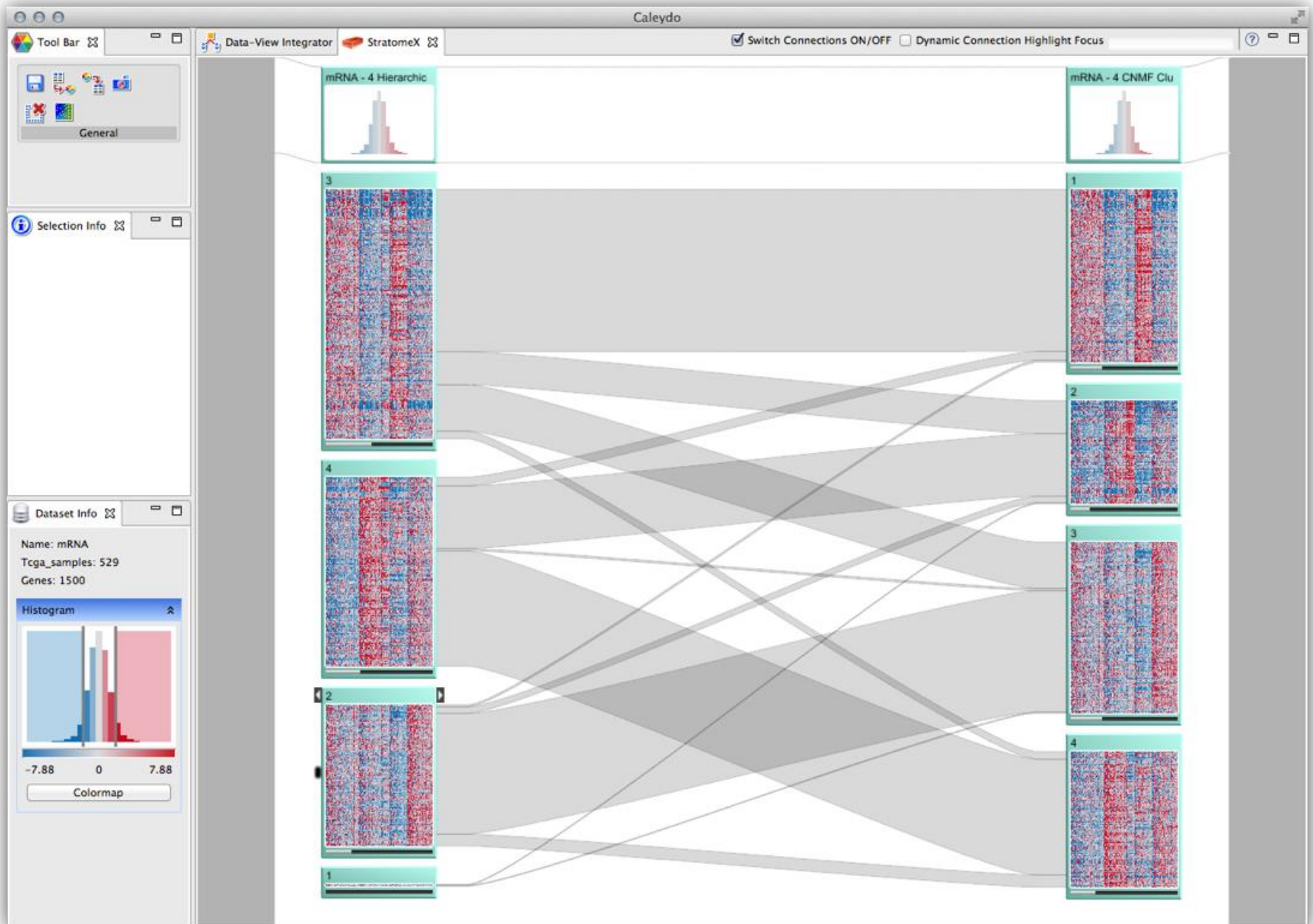
USE CASE / LIVE DEMO

Using TCGA Glioblastoma Multiforme

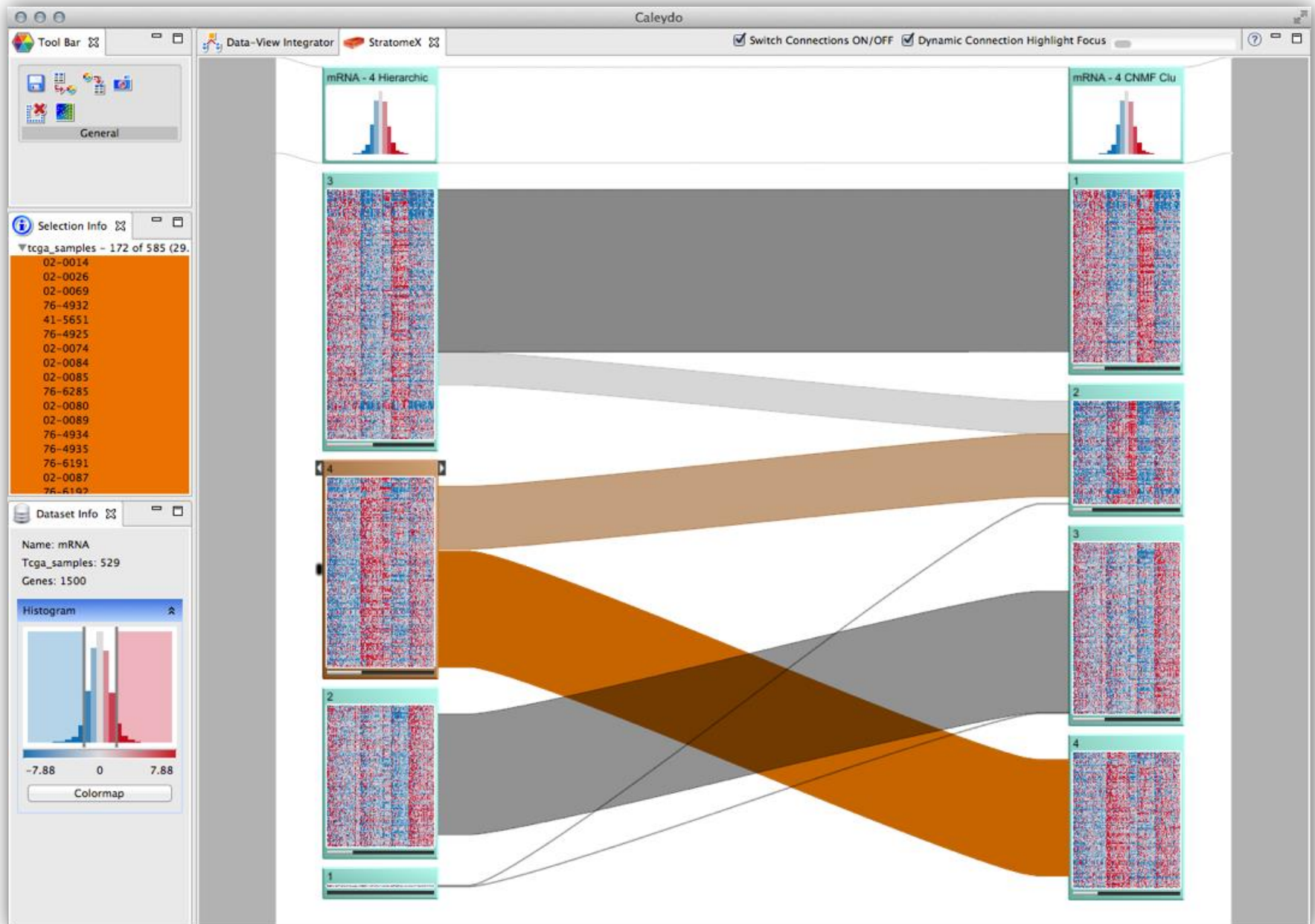




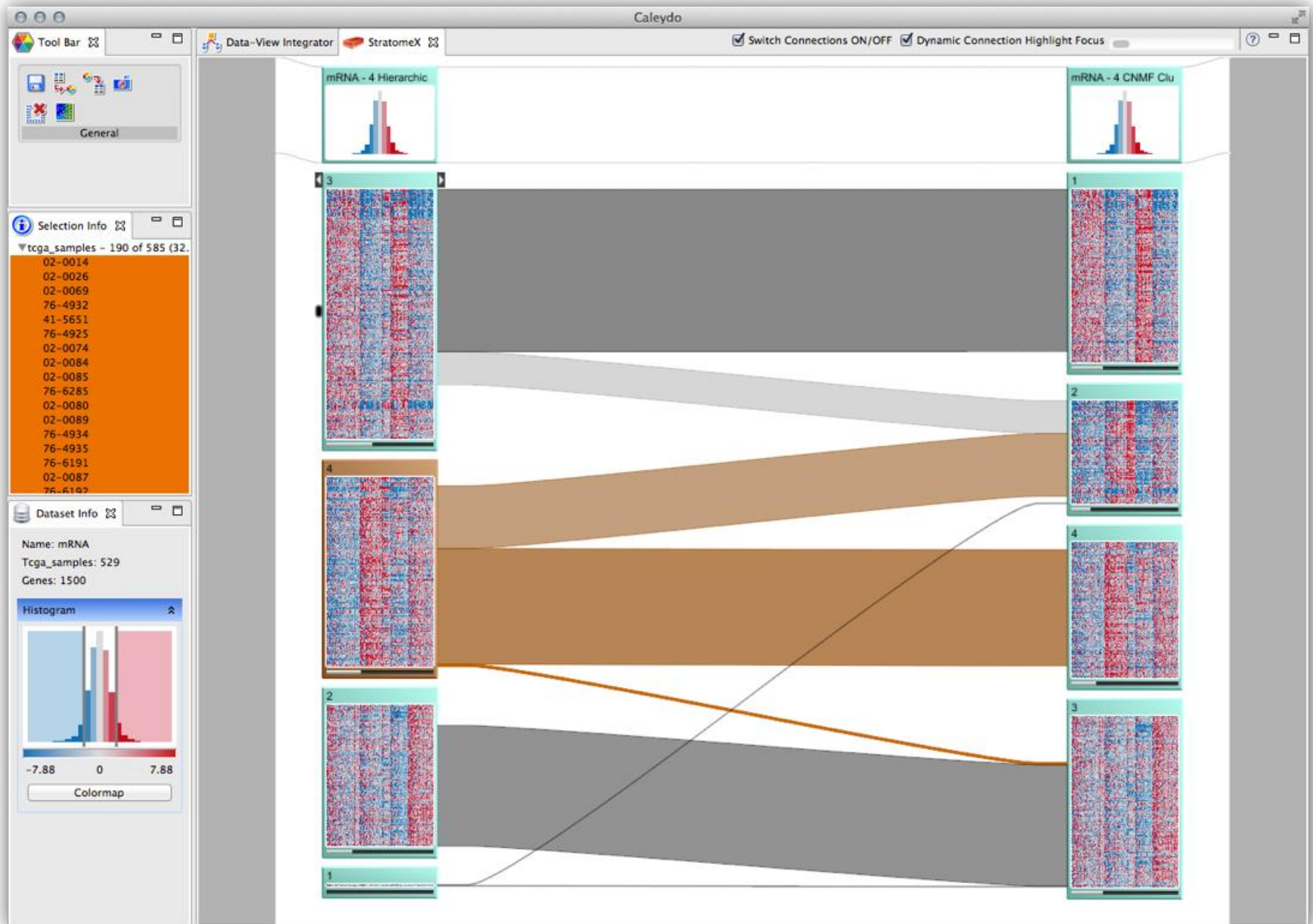
Glioblastoma Multiforme
mRNA clustering CNMF and hierarchical with 4 clusters



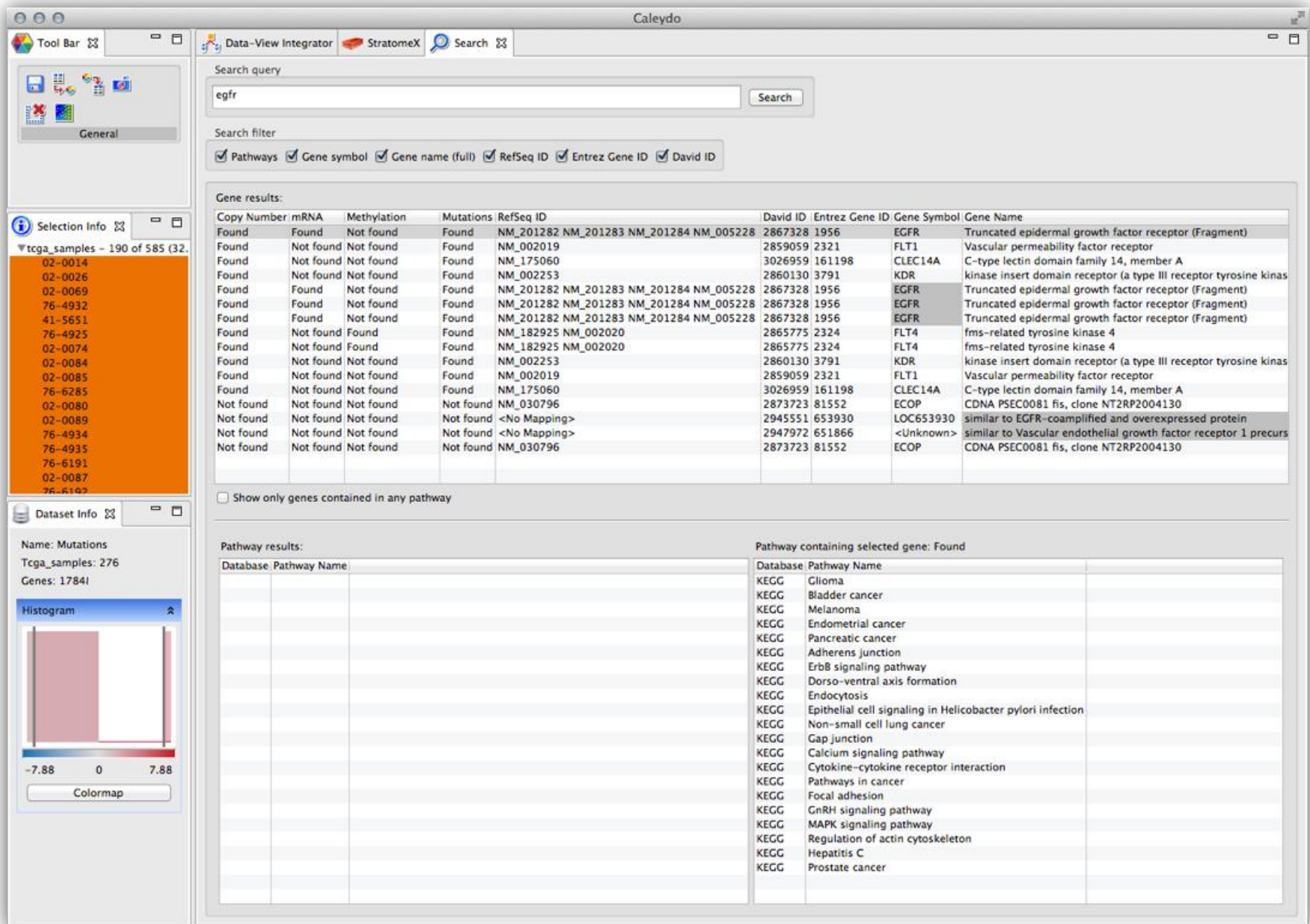
Glioblastoma Multiforme
mRNA clustering CNMF and hierarchical with 4 clusters



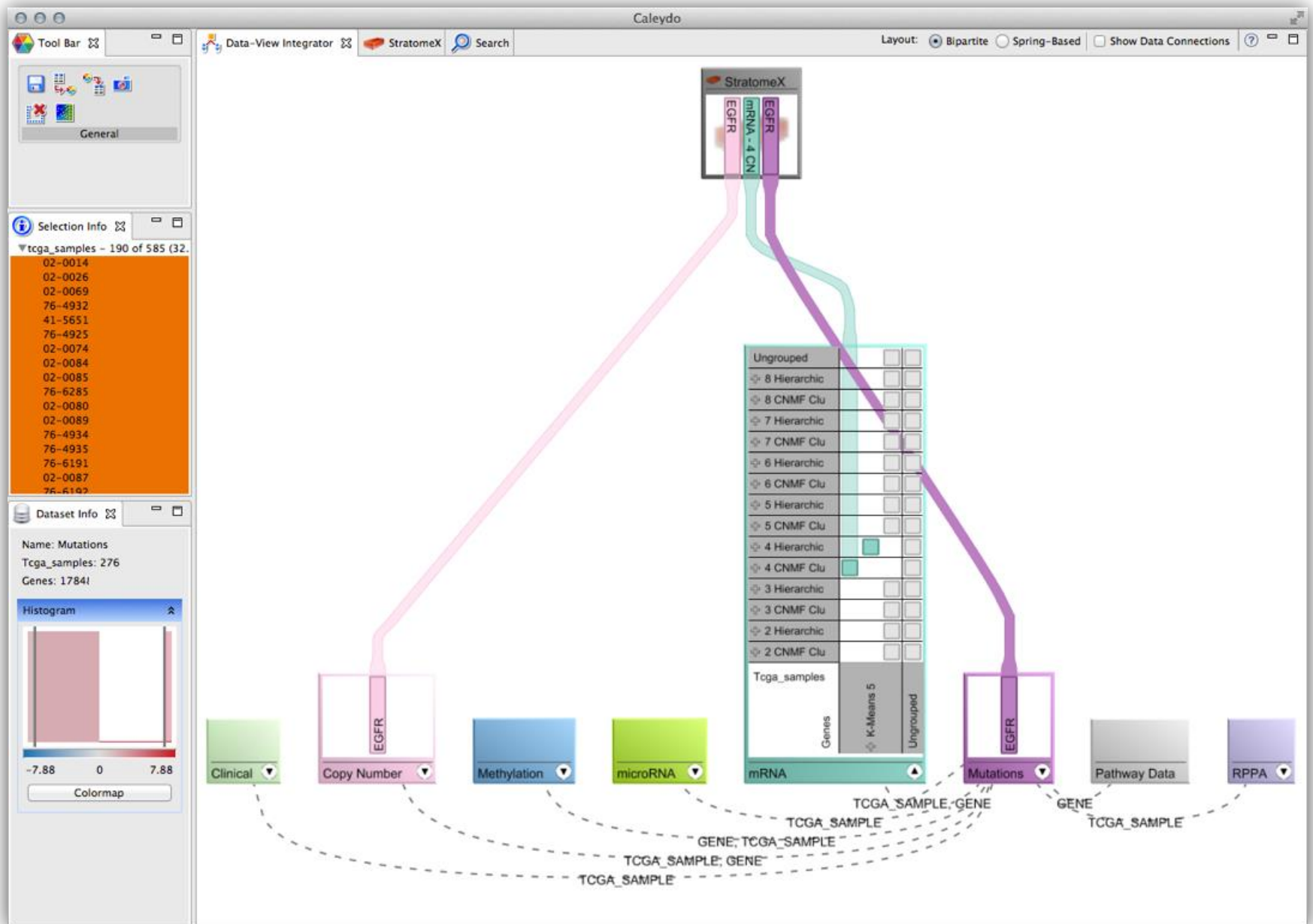
Glioblastoma Multiforme
Use “Dynamic Connection Highlight Focus”



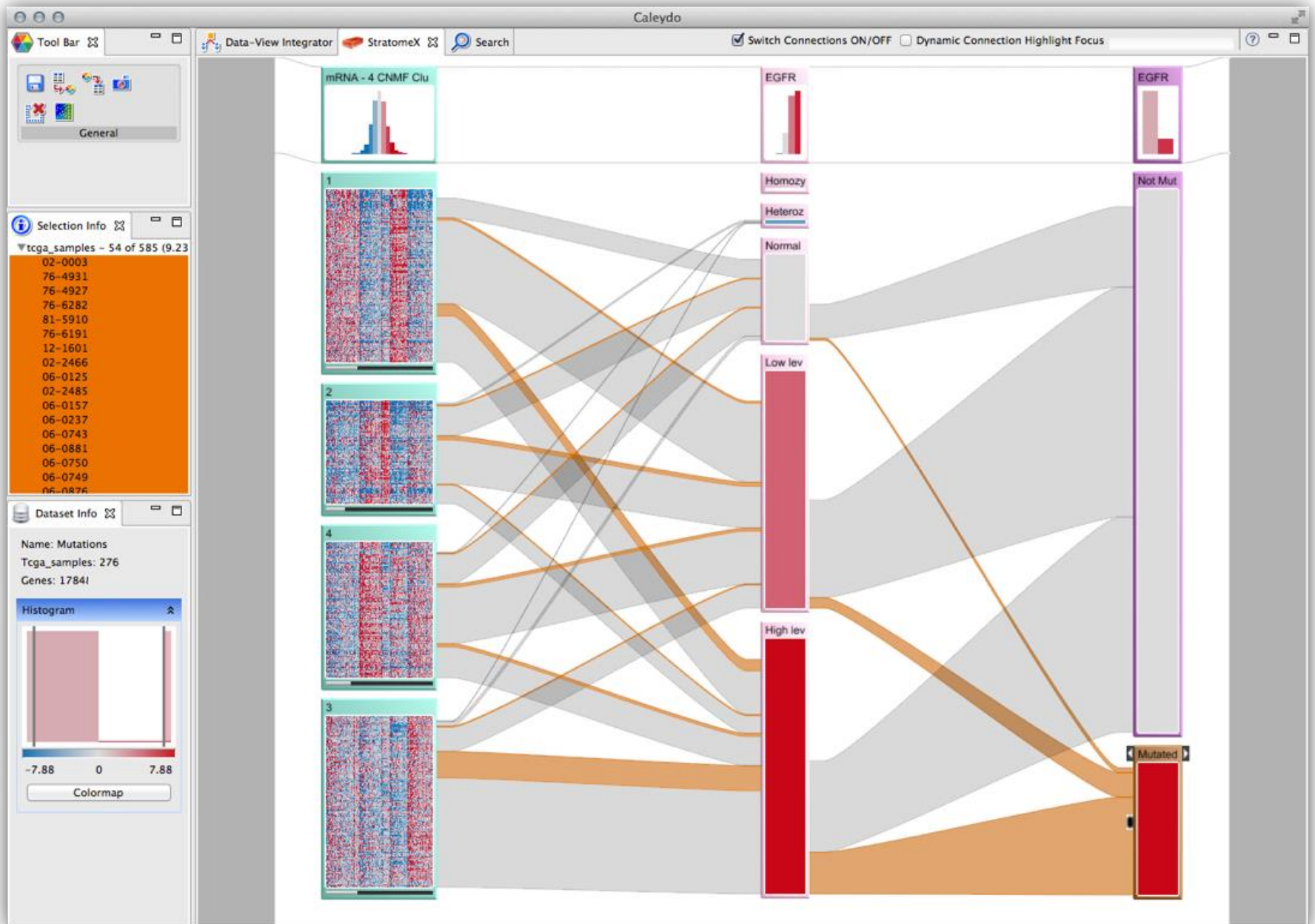
Glioblastoma Multiforme
Reorder bricks within column to detangle bands



Glioblastoma Multiforme
Add EGFR copy number and mutation status



Glioblastoma Multiforme
Add EGFR copy number and mutation status



Glioblastoma Multiforme
Mutated EGFR almost always means amplified EGFR

Create Small Pathway Multiples Group

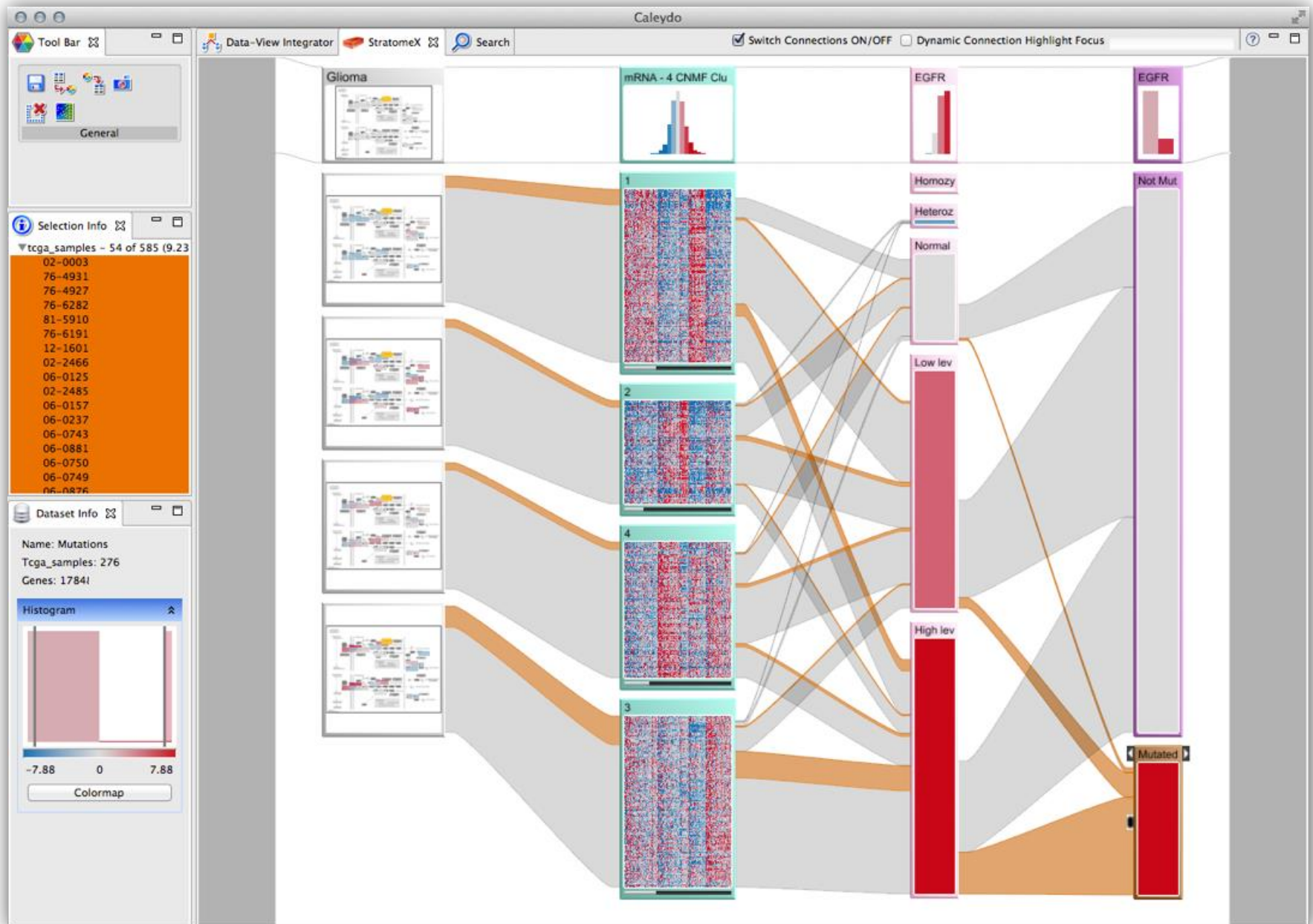
Select the pathways for the group.

Pathway	Database	Gene Occurrences
☐ Fanconi anemia pathway	KEGG	3
☐ Fat digestion and absorption	KEGG	4
☐ Fatty acid elongation	KEGG	2
☐ Fatty acid metabolism	KEGG	5
☐ Fc epsilon RI signaling pathway	KEGG	10
☐ Fc gamma R-mediated phagocytosis	KEGG	20
☐ Focal adhesion	KEGG	46
☐ Folate biosynthesis	KEGG	2
☐ GABAergic synapse	KEGG	18
☐ Gap junction	KEGG	14
☐ Gastric acid secretion	KEGG	12
☒ Glioma	KEGG	10
☐ Glutamatergic synapse	KEGG	18
☐ Glutathione metabolism	KEGG	6
☐ Glycerolipid metabolism	KEGG	5
☐ Glycerophospholipid metabolism	KEGG	7
☐ Glycine, serine and threonine metabolism	KEGG	9
☐ Glycolysis / Gluconeogenesis	KEGG	3
☐ Glycosaminoglycan biosynthesis - chondroitin sulfate	KEGG	5
☐ Glycosaminoglycan biosynthesis - heparan sulfate	KEGG	4
☐ Glycosaminoglycan biosynthesis - keratan sulfate	KEGG	2
☐ Glycosphingolipid biosynthesis - ganglio series	KEGG	3
☐ Glycosphingolipid biosynthesis - globo series	KEGG	2
☐ Glycosphingolipid biosynthesis - lacto and neolacto series	KEGG	3
☐ Glyoxylate and dicarboxylate metabolism	KEGG	2
☐ GnRH signaling pathway	KEGG	8

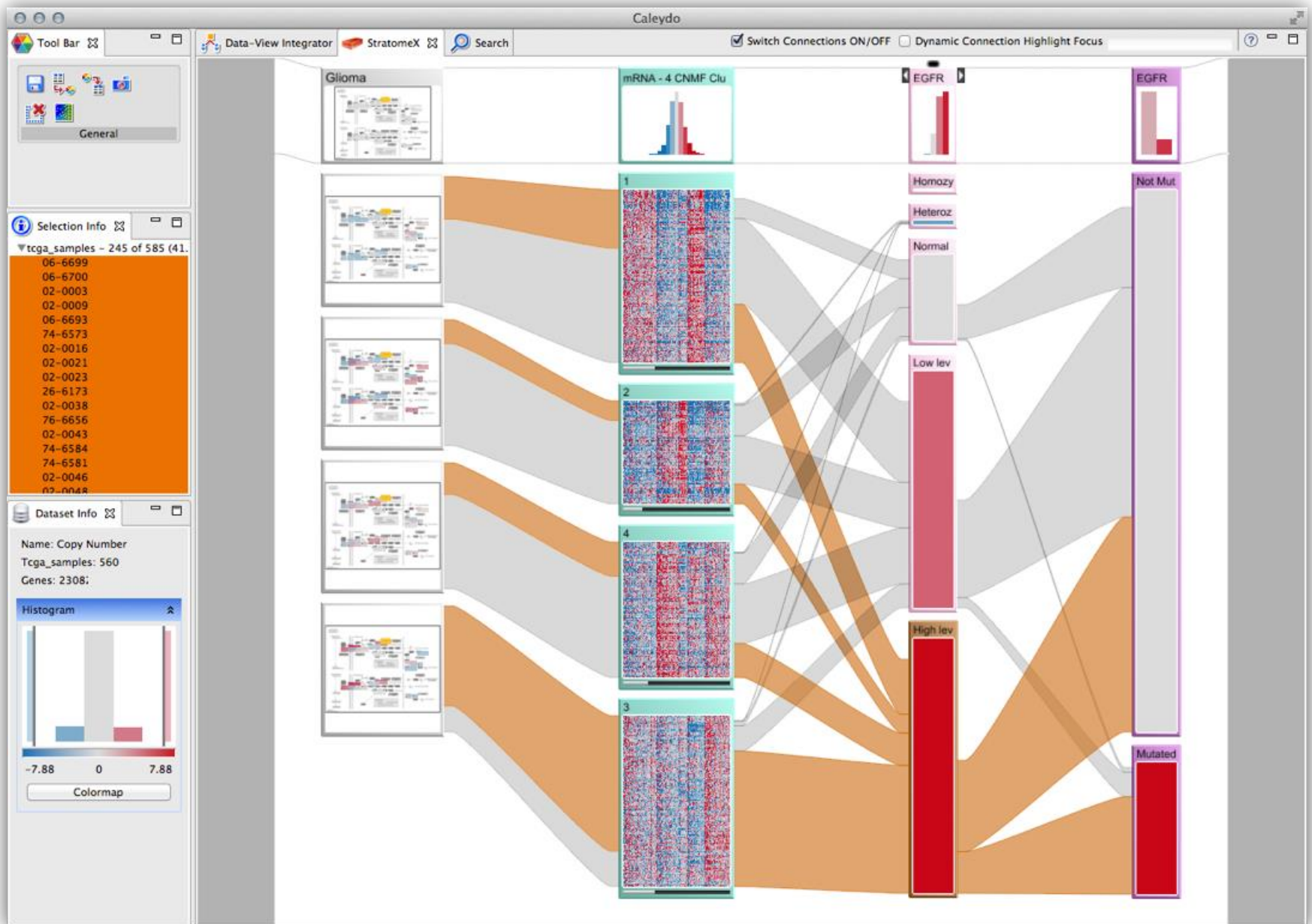
Cancel OK

Glioblastoma Multiforme

Right-click on mRNA header brick and add glioma pathway

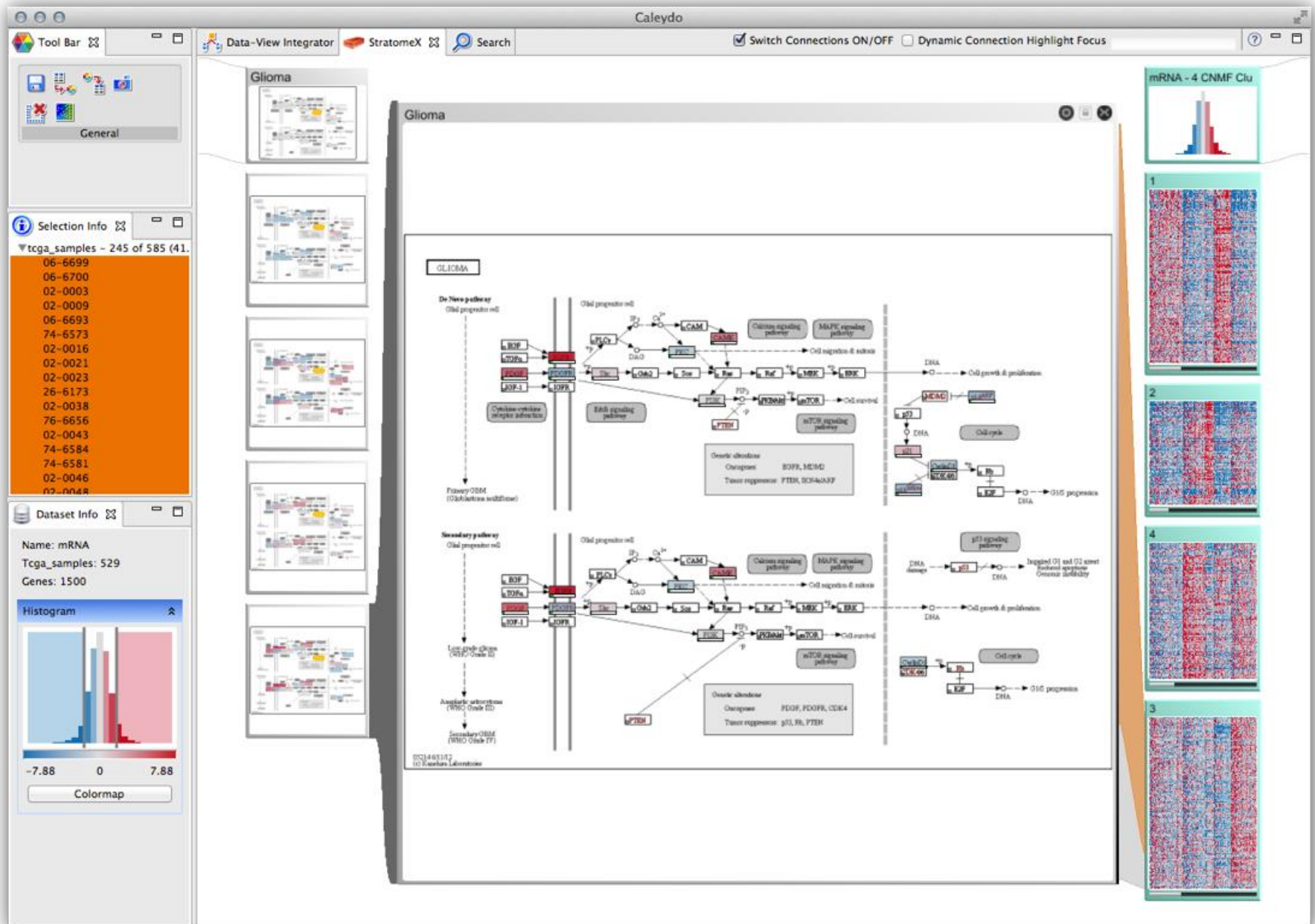


Glioblastoma Multiforme
mRNA mapped to glioma pathway - even distribution of EGFR mutations



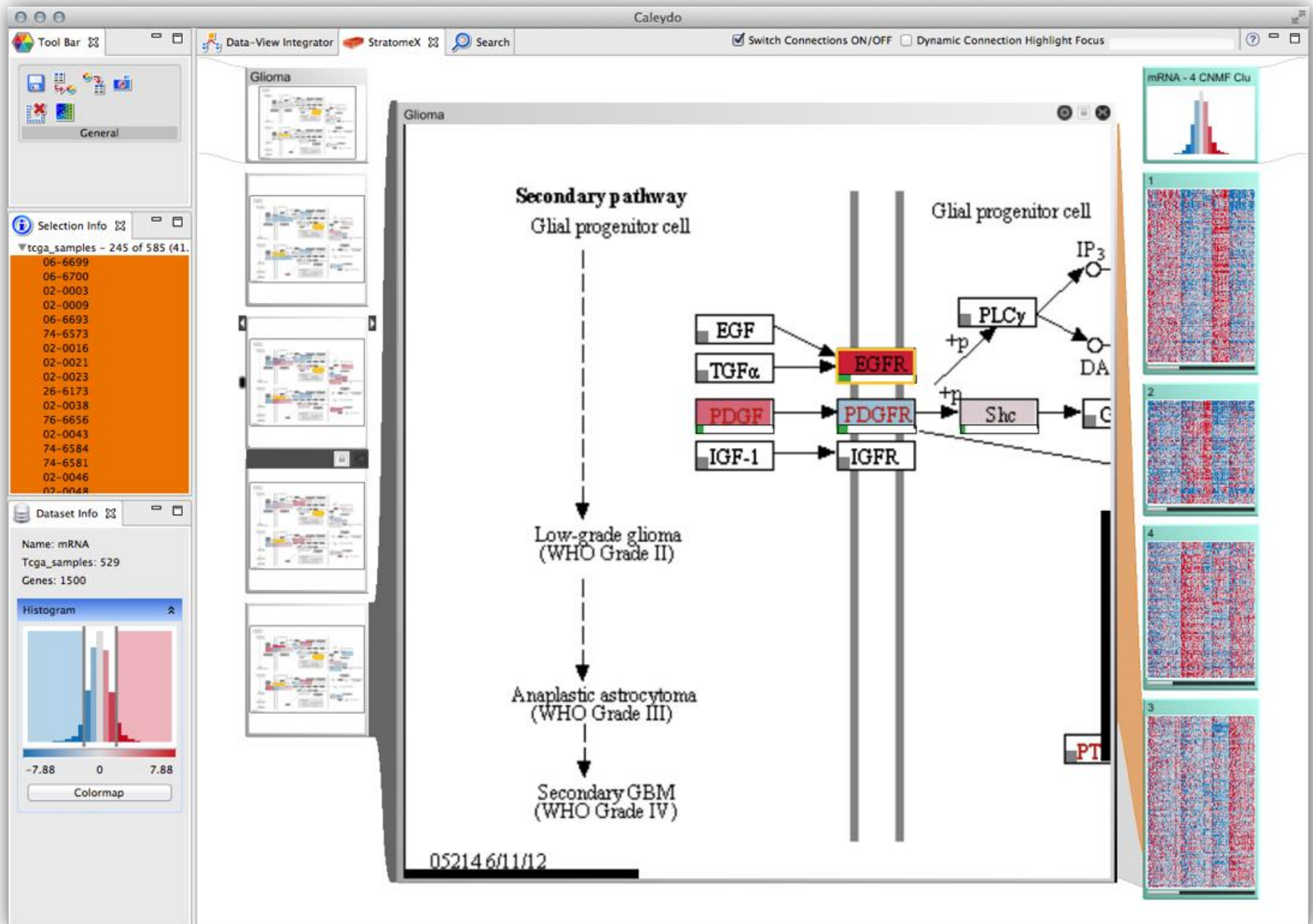
Glioblastoma Multiforme

mRNA mapped to glioma pathway - high amplification frequent in 3rd cluster



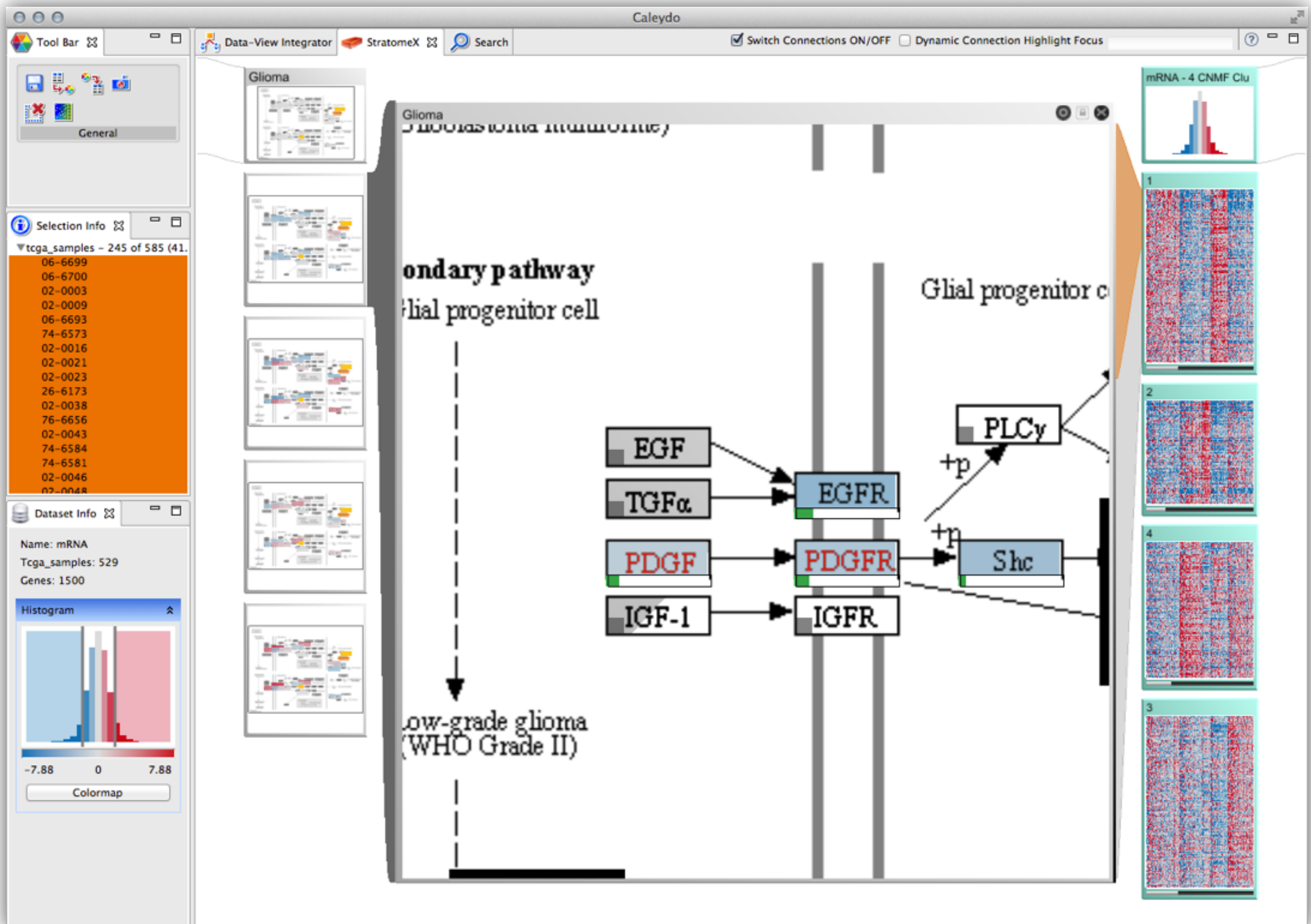
Glioblastoma Multiforme

View pathway map for glioma and 3rd cluster in detail



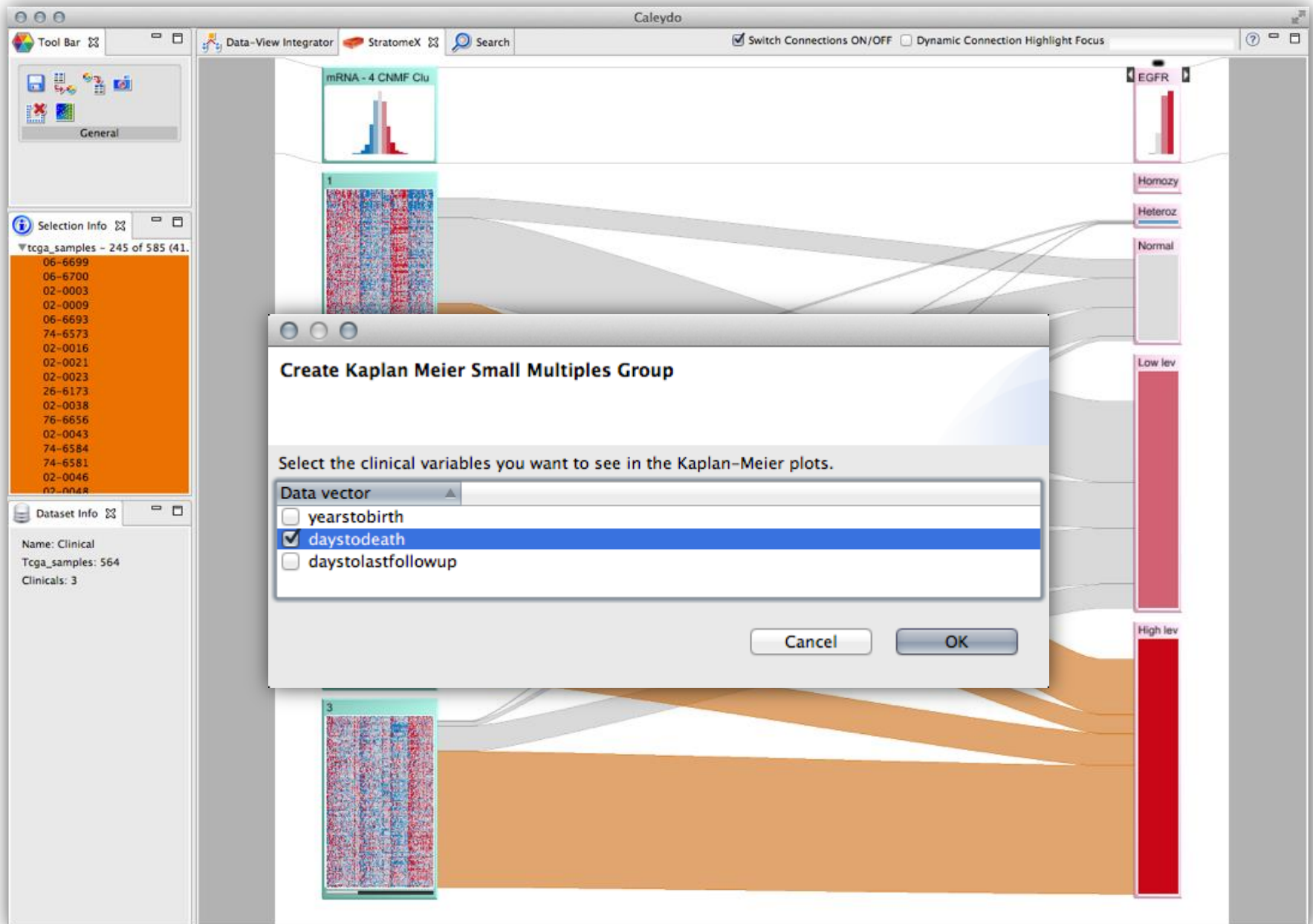
Glioblastoma Multiforme

View pathway map for glioma and 3rd cluster in detail - zoomed into map

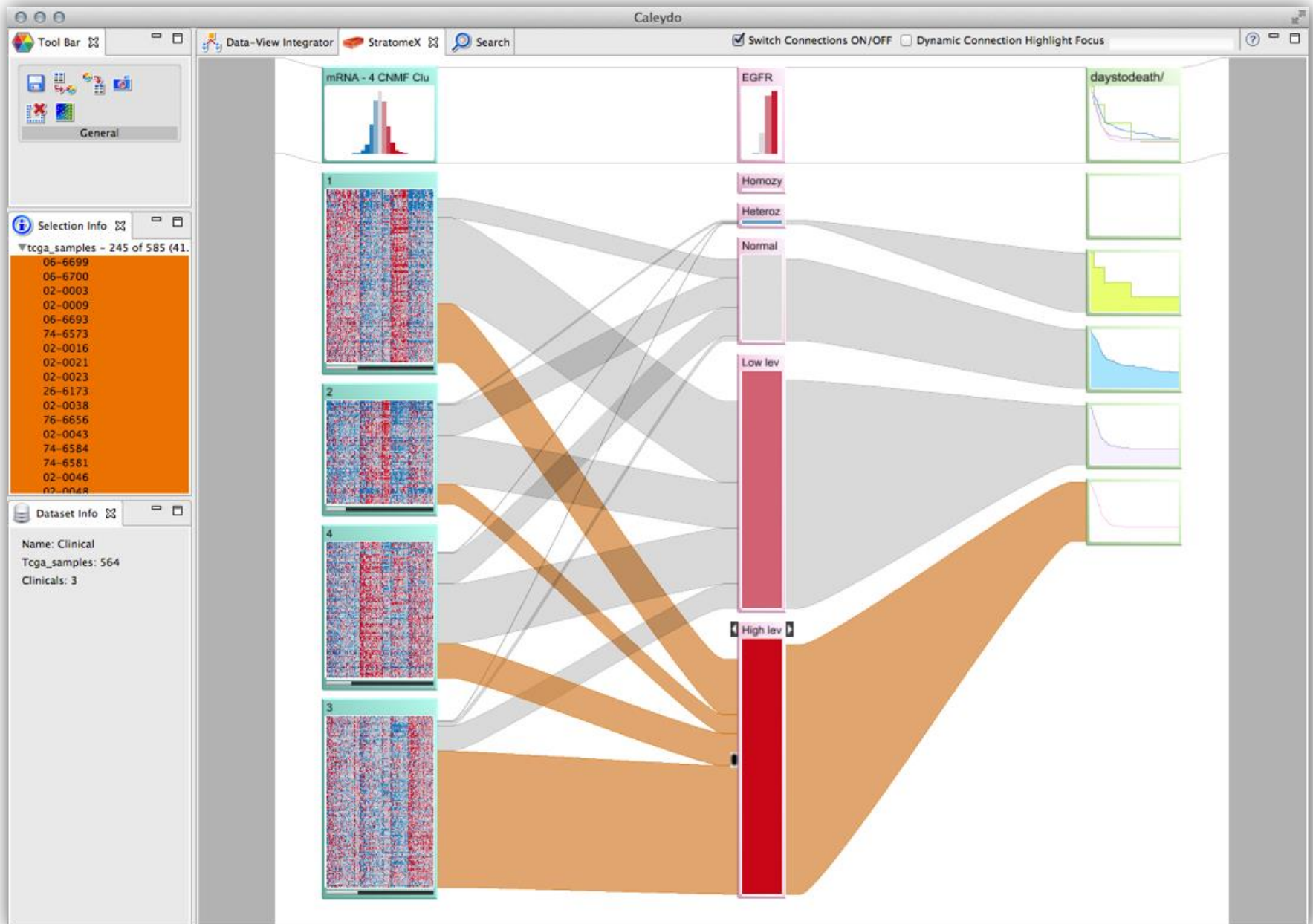


Glioblastoma Multiforme

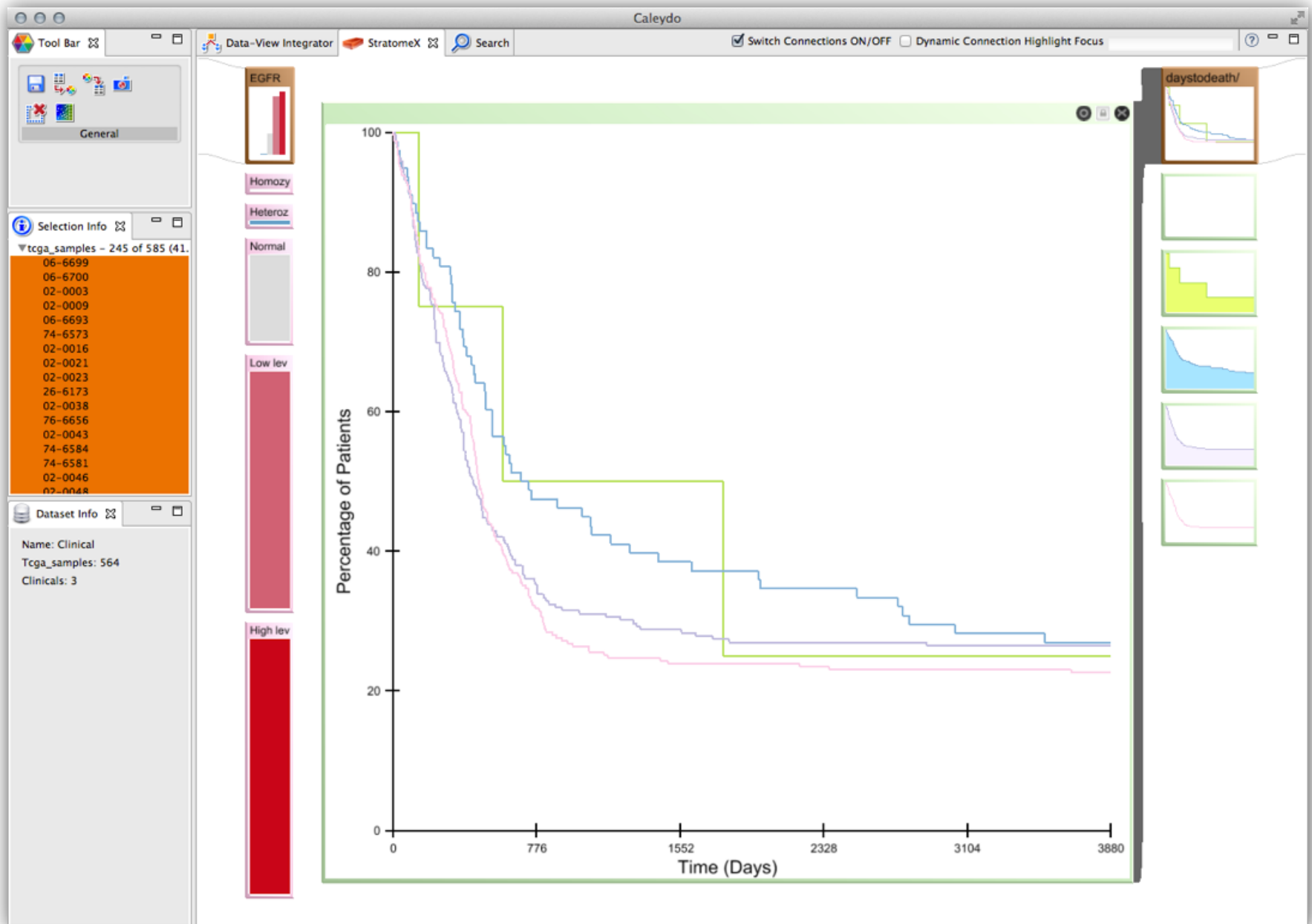
View pathway map for glioma and 1st cluster in detail - EGFR down regulated



Glioblastoma Multiforme
Add clinical variable: survival

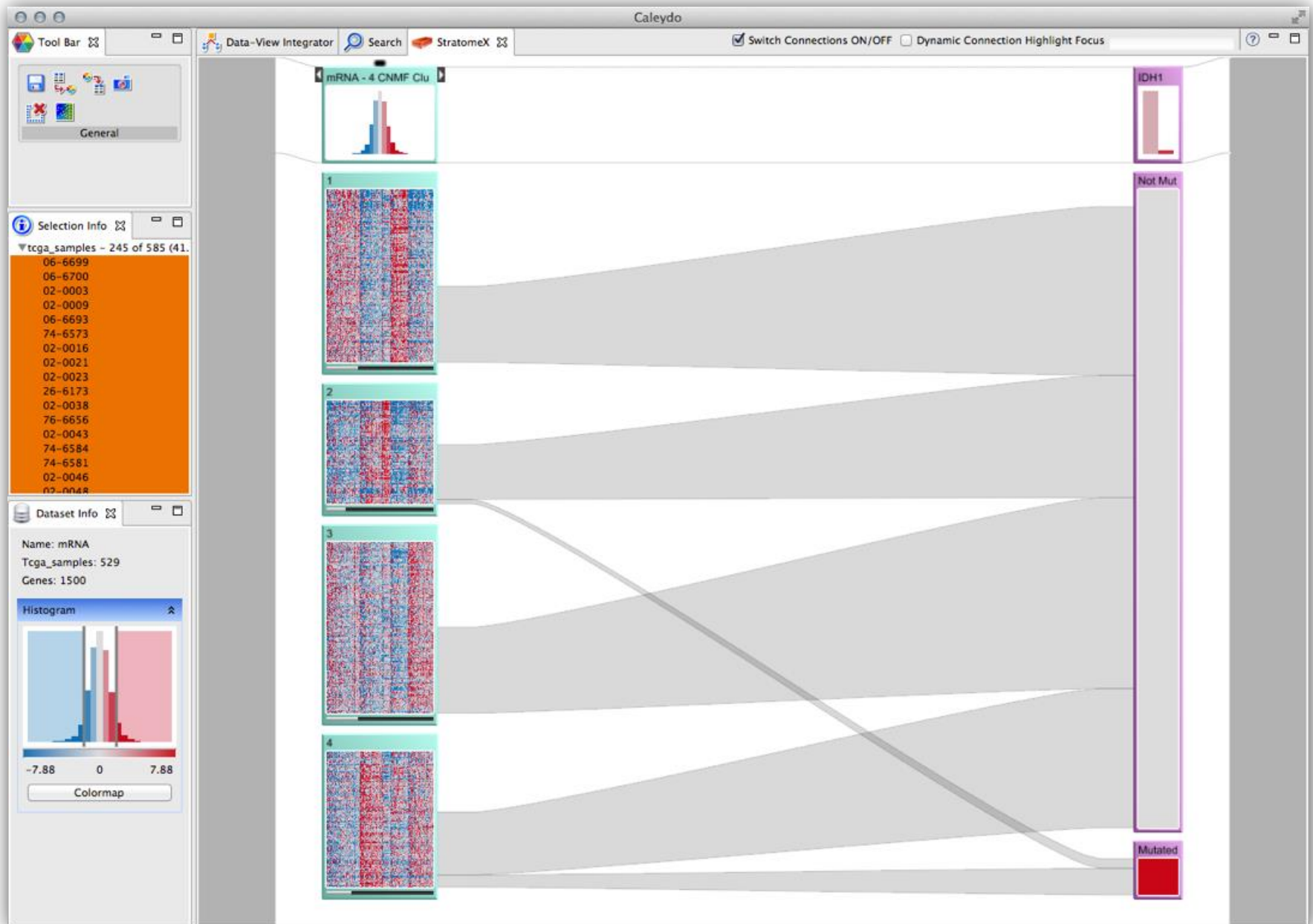


Glioblastoma Multiforme
Add clinical variable: survival



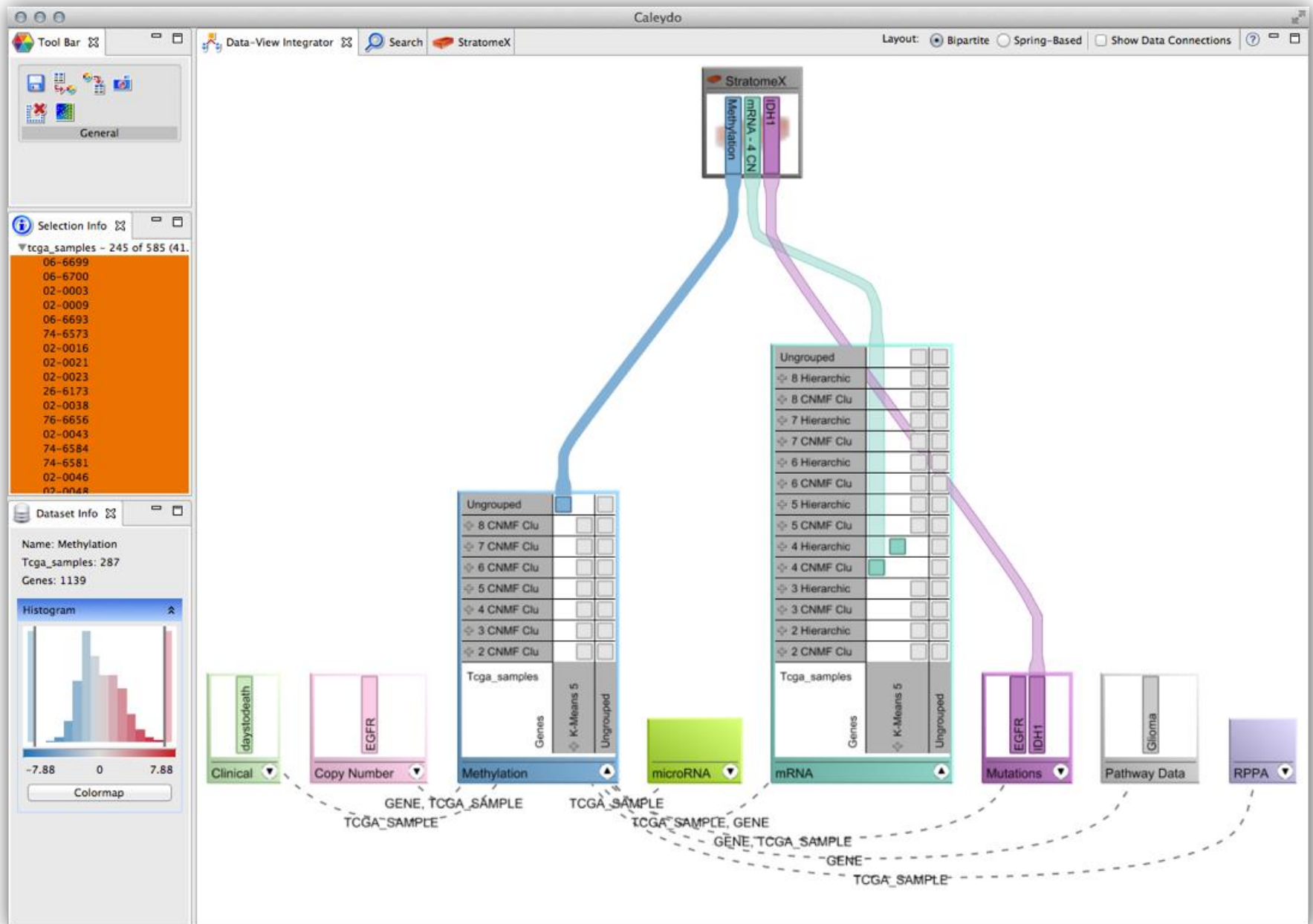
Glioblastoma Multiforme

Add clinical variable: survival - EGFR amplification is worse

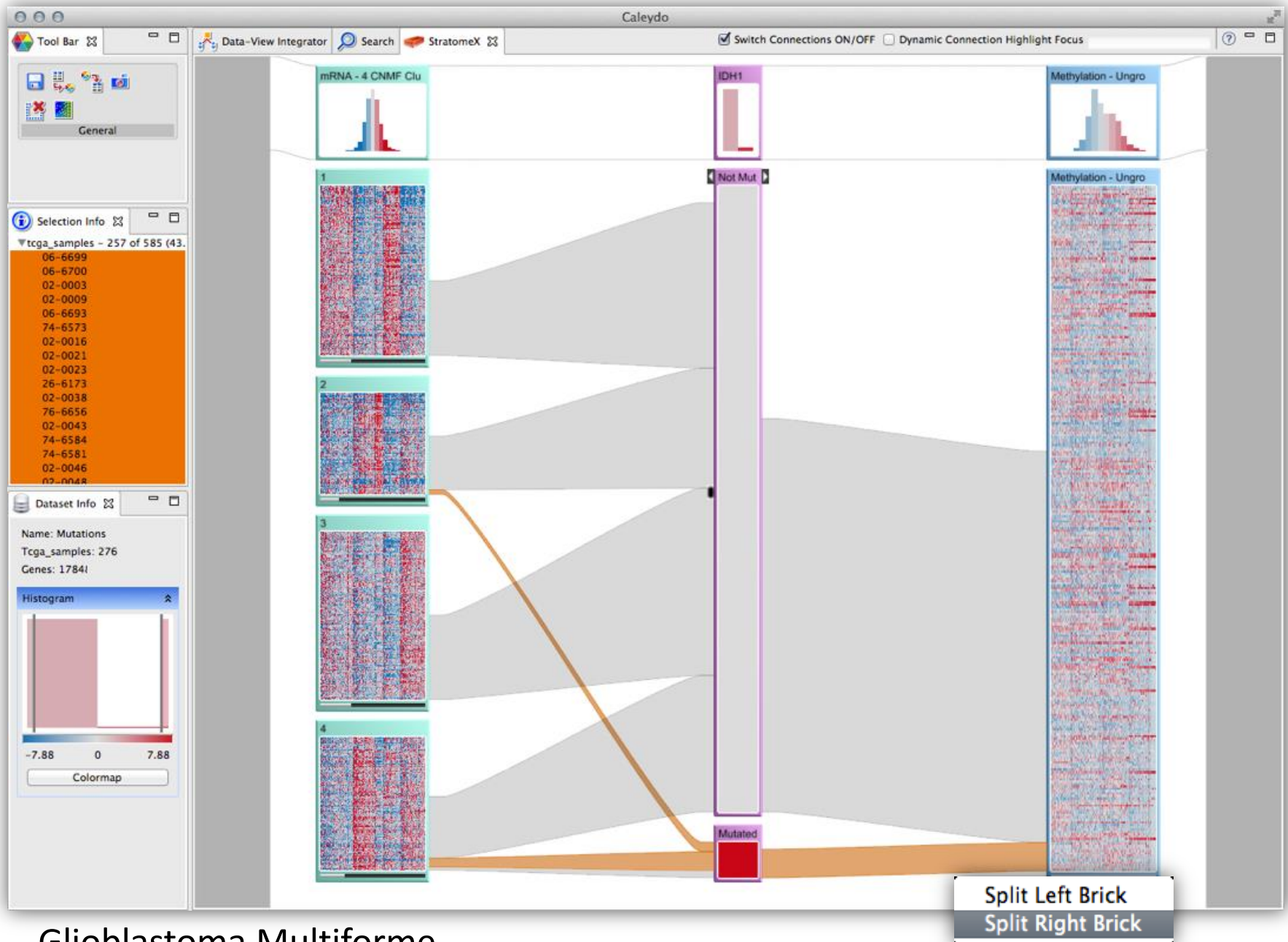


Glioblastoma Multiforme

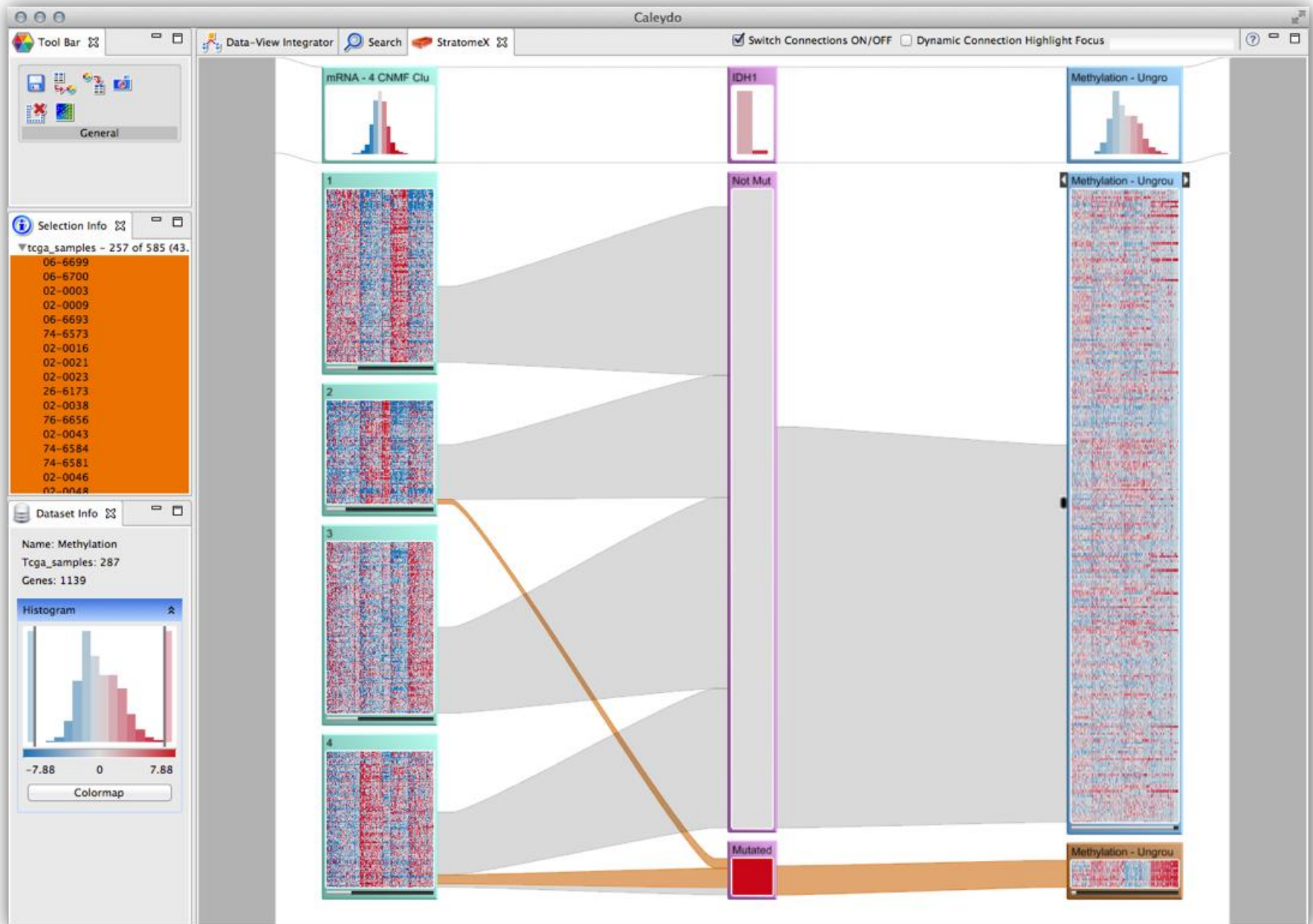
Add IDH1 mutation status - IDH1 mutation is strongly associated with Cluster 4



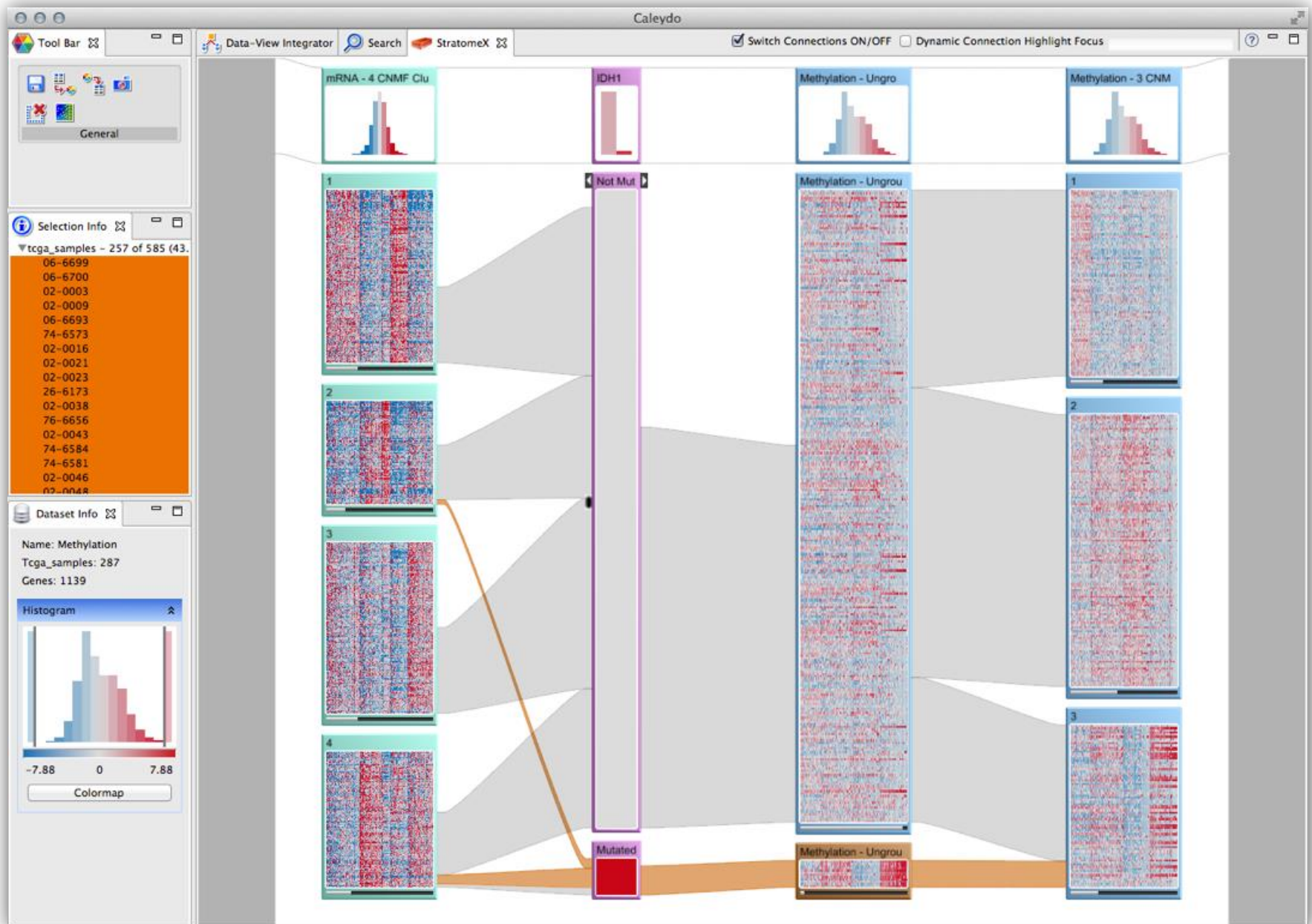
Glioblastoma Multiforme
Add unclustered methylation data



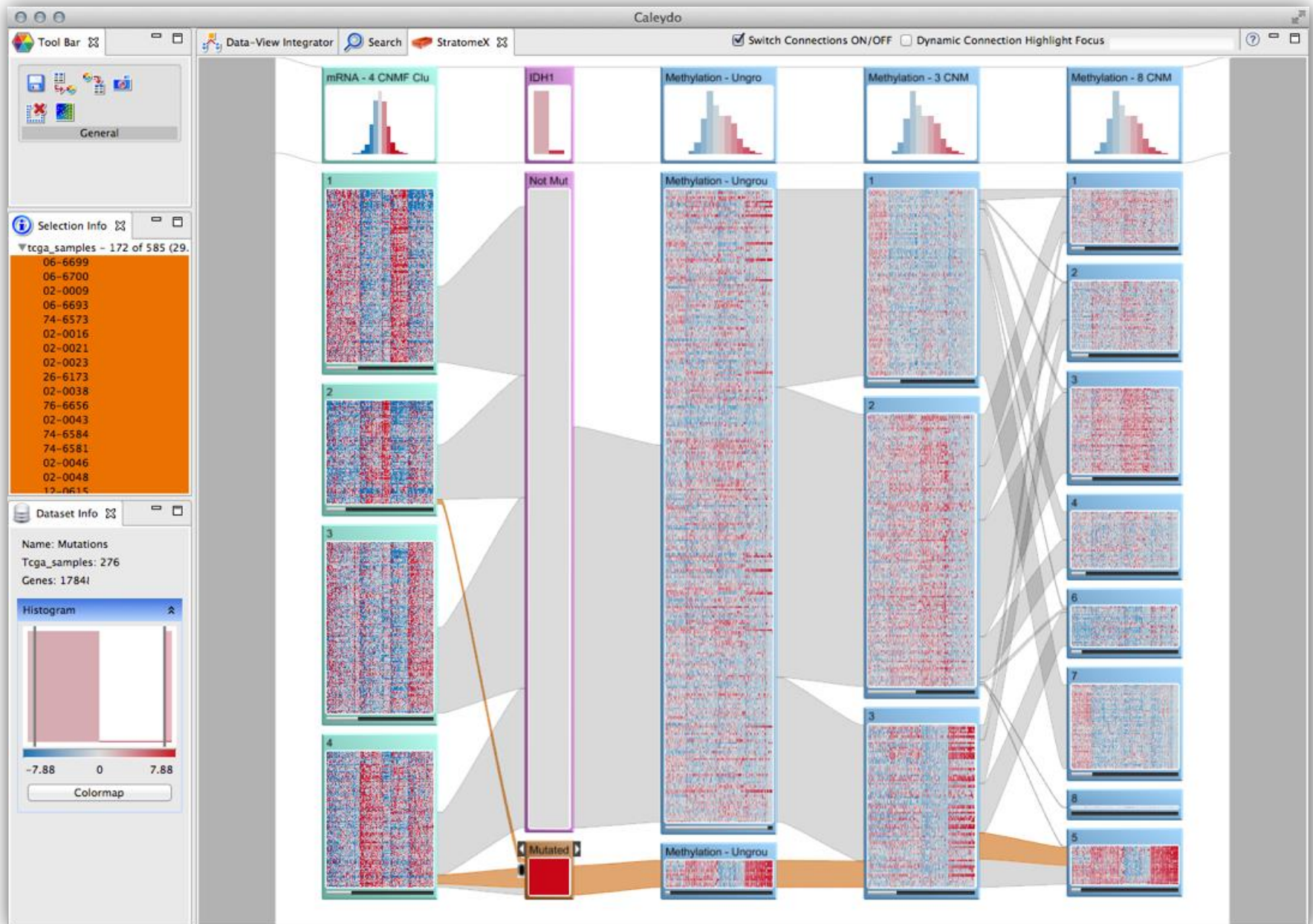
Glioblastoma Multiforme
Create methylation cluster based on IDH1 status



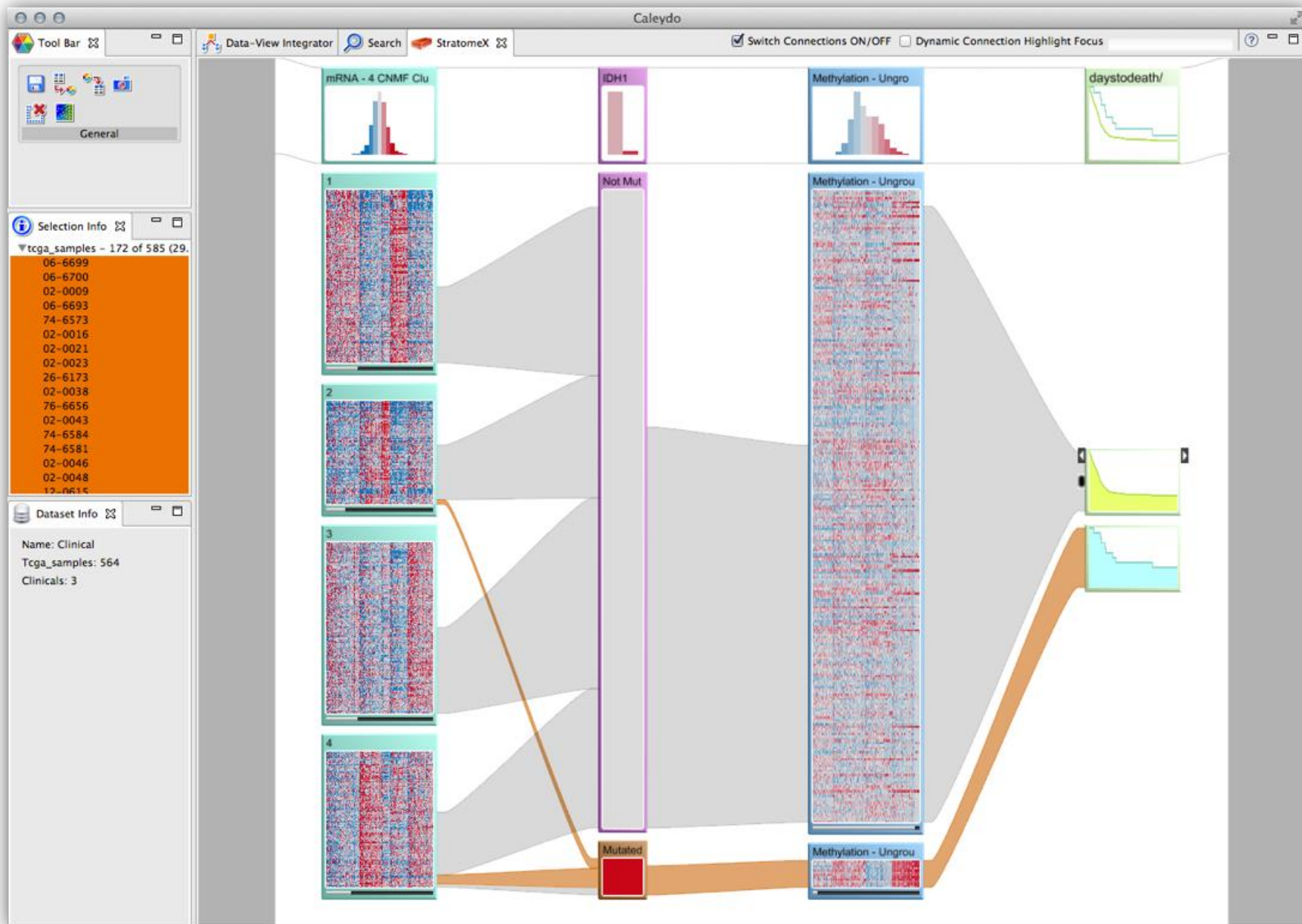
Glioblastoma Multiforme
Create methylation cluster based on IDH1 status: striking pattern emerges



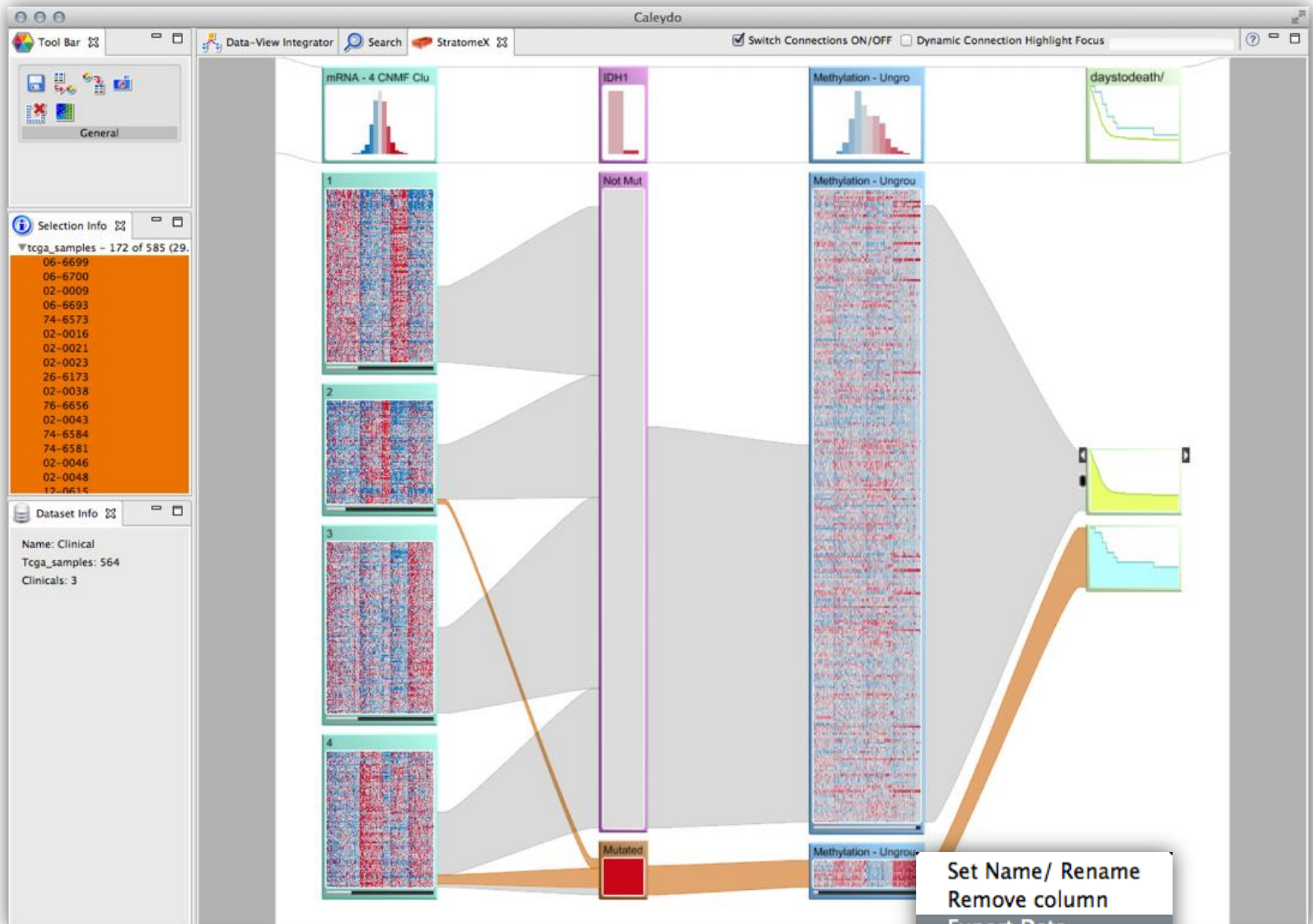
Glioblastoma Multiforme
Add CNMF methylation clustering with 3 clusters



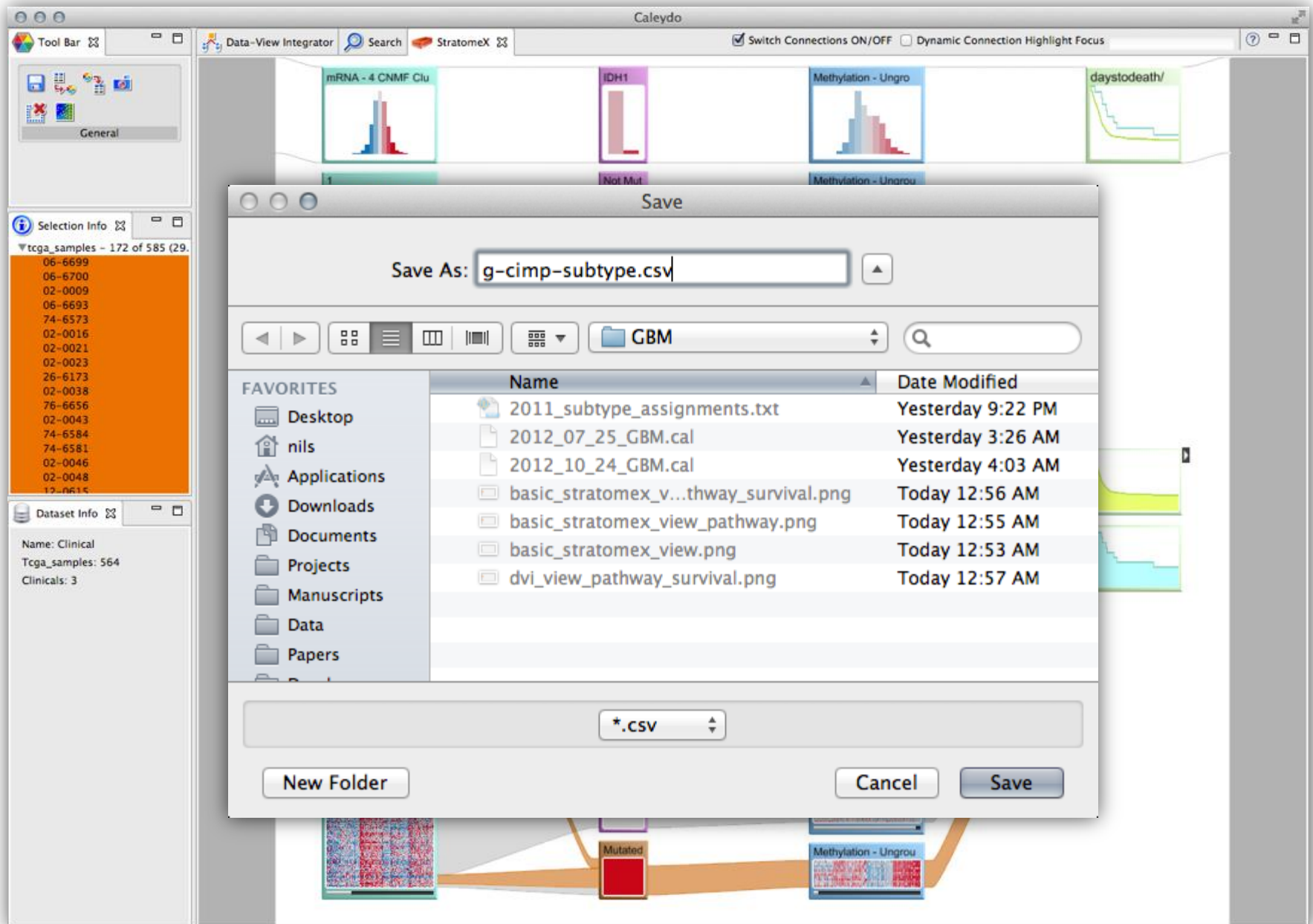
Glioblastoma Multiforme
Add CNMF methylation clustering with 8 clusters



Glioblastoma Multiforme
Add survival data



Glioblastoma Multiforme
Export data



Glioblastoma Multiforme
Export data

g-cimp-subtype.csv

Search in Sheet

Home Layout Tables Charts SmartArt Formulas Data Review

B1 2/1/2483

	A	B	C	D	E	F	G	H	I	J	K
1	Identifier	Feb-83	06-0128	06-0129	Jun-70	12-0818	12-0827	14-1458	14-1821	16-0850	19-2629
2	PAX3	-0.1417896	0.11429937	-0.3375288	-0.2687174	0.41211167	0.025824534	0.28384444	-0.09339912	0.18420377	-0.191
3	FSD1	-0.08075349	-0.2783441	-0.40741664	-0.44710276	0.33394808	-0.013799776	-0.21399914	-0.4125057	-0.47387525	0.214
4	HOXB1	-0.17096336	-0.256667	-0.42307696	-0.45638907	-0.32295564	-0.05438675	-0.10711608	-0.07609512	-0.4616507	-0.28
5	LRTM1	-0.26256612	-0.087932535	-0.3923521	-0.4137606	-0.35225746	0.20410466	0.114409216	-0.49363917	-0.24358872	-0.293
6	MRI1	0.52148366	-0.29675996	-0.2756883	0.088104345	0.53619486	-0.34041375	-0.3065707	0.010070764	0.5537861	0.432
7	ARPP-21	0.020761115	-0.31807184	-0.42144486	-0.4305186	-0.42465714	-0.1244007	-0.30161503	-0.37284482	-0.46526018	-0.261
8	DEFB123	-0.1923641	-0.25150293	-0.4641215	-0.26961613	-0.1000148	-0.28770006	0.28340194	0.2653647	0.005994942	0.0613
9	SLC17A4	-0.31467038	-0.27672243	-0.431508	-0.41168833	-0.36524487	-0.061333694	0.26783326	-0.09931746	-0.43470803	-0.150
10	LIPE	-0.16355085	-0.2057257	-0.40440673	-0.38160163	0.29892504	0.122576095	0.1004104	0.20973328	-0.39580178	-0.205
11	MYL4	-0.102354944	-0.19977987	-0.45640662	-0.30327943	0.32728568	0.24660082	0.37093848	-0.48242703	-0.106939346	0.080
12	PCOLCE	0.16873473	-0.060603883	0.14316115	0.36924142	0.16061684	-0.3404151	-0.24301422	0.3602322	-0.37927496	0.127
13	C19ORF21	0.3087737	0.16148421	0.27055407	0.3561273	0.35662952	0.20585556	0.31406486	0.35922503	0.25179705	0.2
14	SERPINB7	-0.12861061	0.037382588	-0.3850914	0.1519851	0.28978908	-0.23642868	0.31525913	0.22680324	0.13706173	0.137
15	LRRC4	-0.33419588	-0.40359664	-0.4558562	-0.4762344	-0.45512238	-0.036328483	-0.48569387	-0.5082524	-0.2696056	0.107
16	RIMS3	-0.0551132	-0.06975174	-0.30769977	0.26118246	0.45369178	0.38229224	0.2668557	0.26791584	-0.30687025	-0.020
17	GNAS	0.30027902	-0.027930172	-0.17651857	0.002783225	0.4111265	0.36731106	0.40524366	0.38852462	-0.10159901	0.314
18	JPH4	-0.22371887	-0.11420198	-0.31419784	-0.32766497	0.25921133	-0.10835368	-0.28695214	-0.3721775	-0.4428355	-0.123
19	FAM12B	-0.23671702	-0.16780111	-0.32869387	-0.36372742	-0.3418818	0.22051738	0.19346273	-0.048124313	-0.2654656	0.210
20	ANK3	-0.11880269	-0.2569762	-0.2837809	-0.41483262	0.065270714	0.18258646	-0.37288272	-0.2372996	-0.36835268	-0.182
21	SLC34A1	-0.2081686	-0.30003256	-0.41136366	-0.34625614	-0.39861625	-0.10222819	0.35125035	-0.3248961	-0.22747482	0.0846
22	TM4SF1	-0.1327866	0.06710029	0.15667441	-0.17877707	0.23100285	0.25099844	0.28077552	-0.1562967	0.34177822	-0.032
23	NUPR1	0.23173398	-0.35026714	-0.25457463	-0.033180553	0.38761973	0.14017034	0.3422673	0.14159	0.20451283	-0.09
24	ZP4	-0.23404218	-0.05524534	-0.39515772	-0.29391825	-0.36461675	-0.24992093	0.35941556	0.28970224	0.07766051	0.124
25	KALRN	-0.025574567	0.16688342	0.22365151	0.263183	0.2791145	0.011530072	0.21209803	-0.17796098	0.26628864	0.211
26	MPP7	0.12534429	0.026998634	-0.121099435	0.27216002	0.29677257	0.11272653	0.32750252	0.26103935	0.34128648	0.282
27	C10ORF11	0.21831873	-0.05703577	-0.34773836	0.23048413	0.26526666	0.07562882	0.35189787	0.3401315	-0.10411584	0.0473
28	CCL16	-0.1704086	-0.21435413	-0.48918238	-0.24238631	-0.2367581	0.07165758	0.30762985	0.20658514	0.21251778	0.098
29	OR1G1	-0.26554856	-0.19975024	-0.3487445	-0.31132045	-0.2778506	-0.15856531	0.07690581	-0.14481743	-0.39109546	0.174
30	APOD	-0.31046936	-0.19764917	-0.31000727	-0.32856318	0.23192082	-0.008822635	0.07476304	-0.25195557	-0.43233263	-0.254
31	SRBM2	0.003401006	0.13816880	0.42616758	0.23054852	0.08812471	0.13025866	0.11574127	0.2462023	0.4602042	0.264

Sheet1

Normal View Ready

Sum= Feb-83

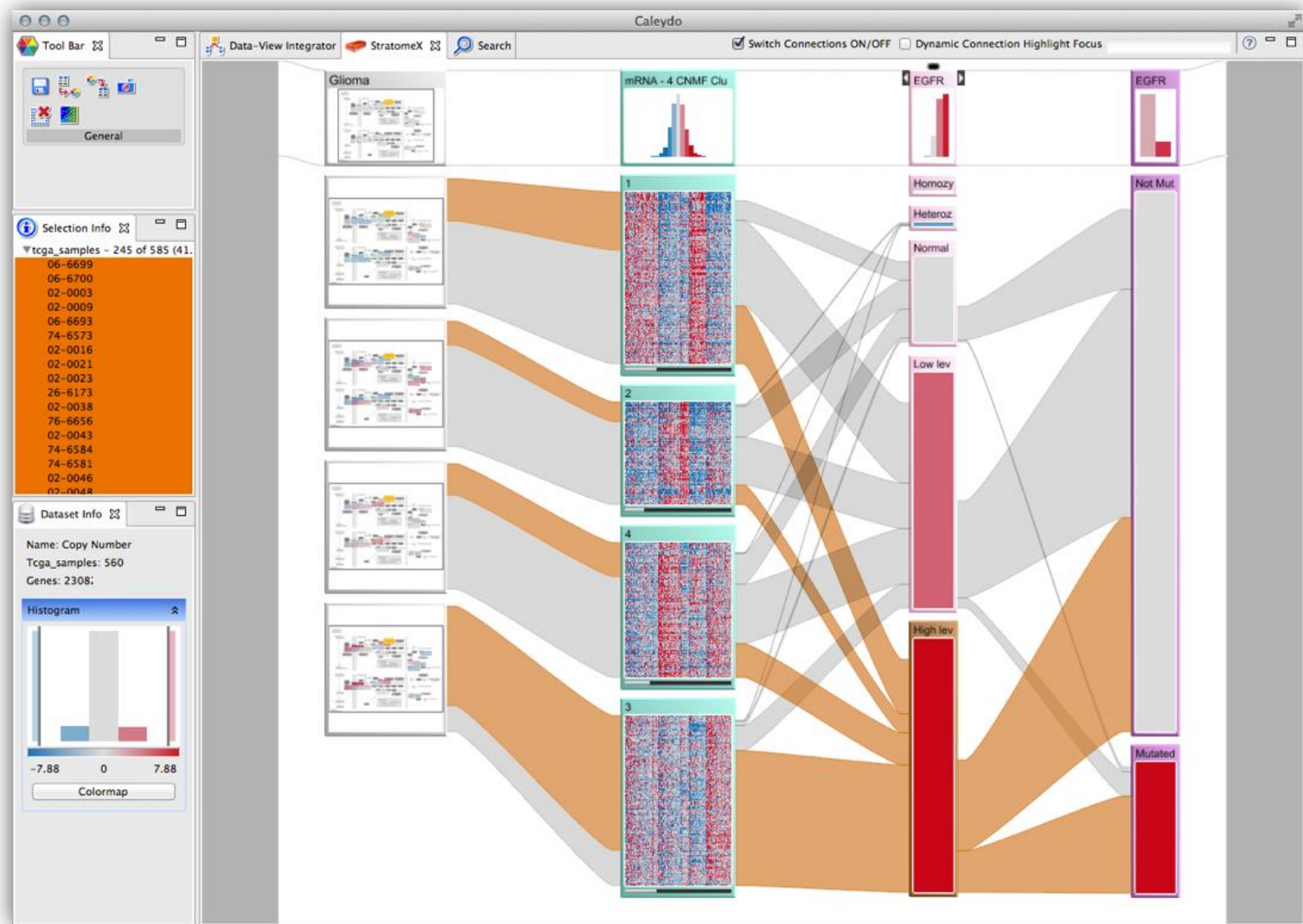
Glioblastoma Multiforme

Downstream analysis on exported matrix

Please cite the StratomeX paper

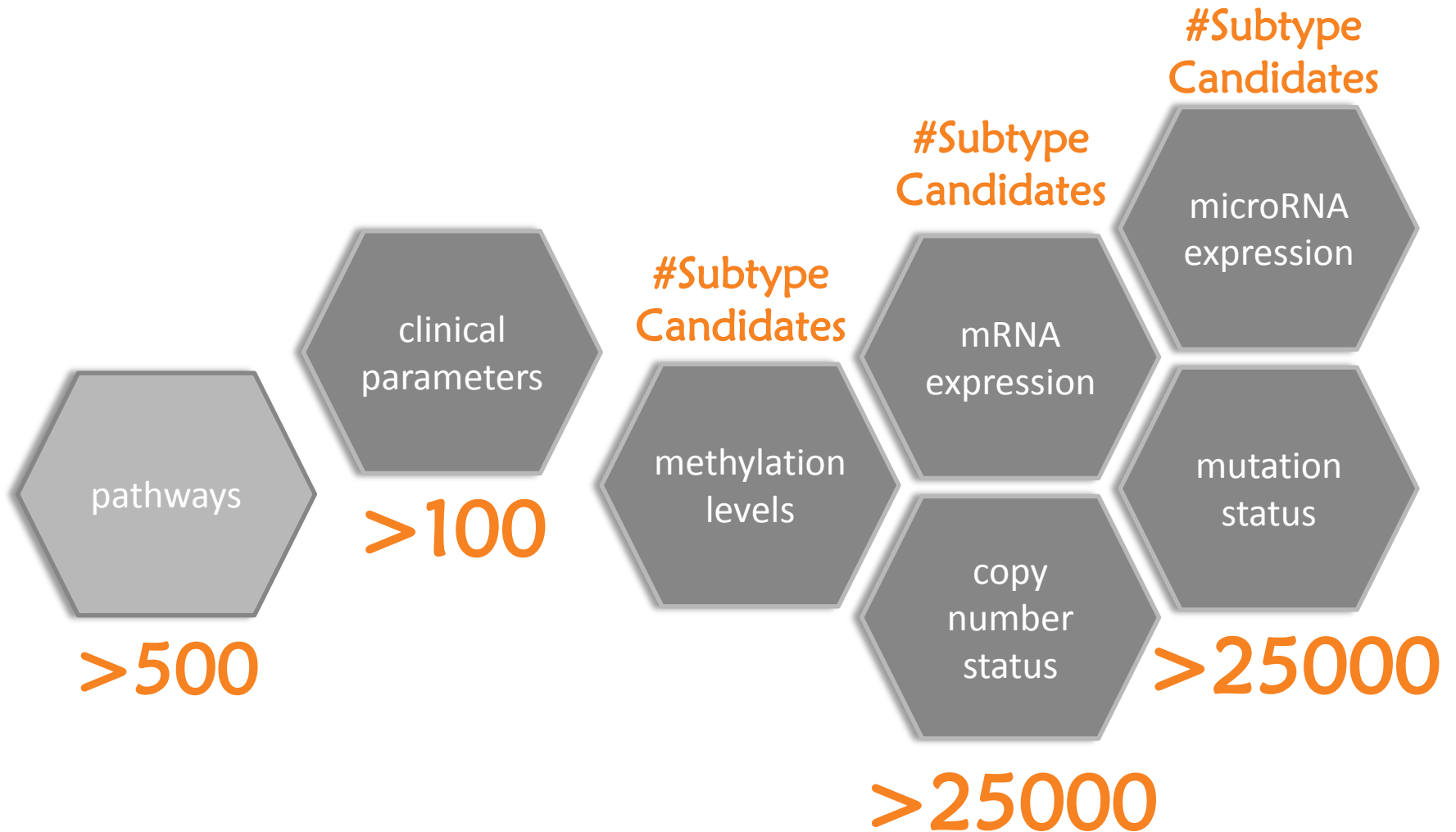
A Lex, M Streit, H-J Schulz, C Partl, D Schmalstieg, PJ Park and N Gehlenborg,
*"StratomeX: Visual Analysis of Large-Scale Heterogeneous Genomics Data for
Cancer Subtype Characterization"*, Computer Graphics Forum (EuroVis '12),
31:3, 1175-1184 (2012)

CURRENT DEVELOPMENTS



StratomeX strong for hypothesis validation

TCGA Data

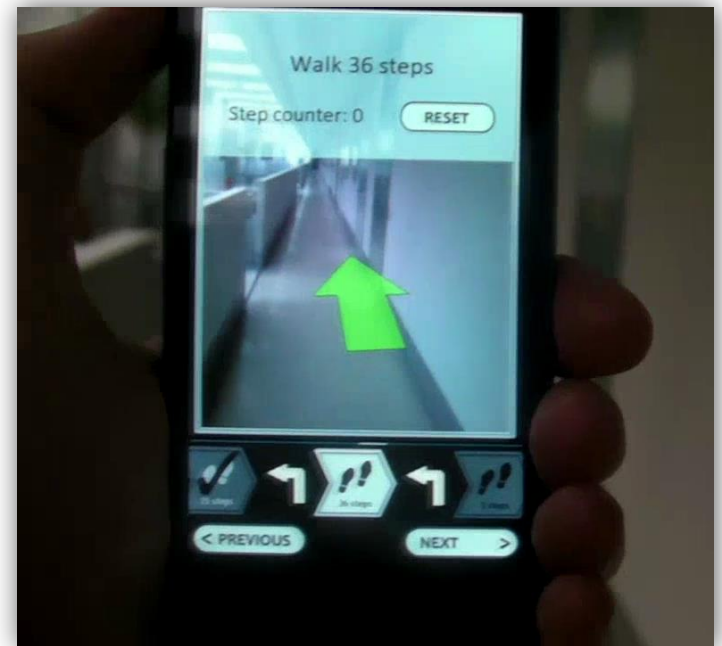


Guided Exploration

Visually guide analyst through analysis

Pinpoint to possibly interesting patterns and correlations

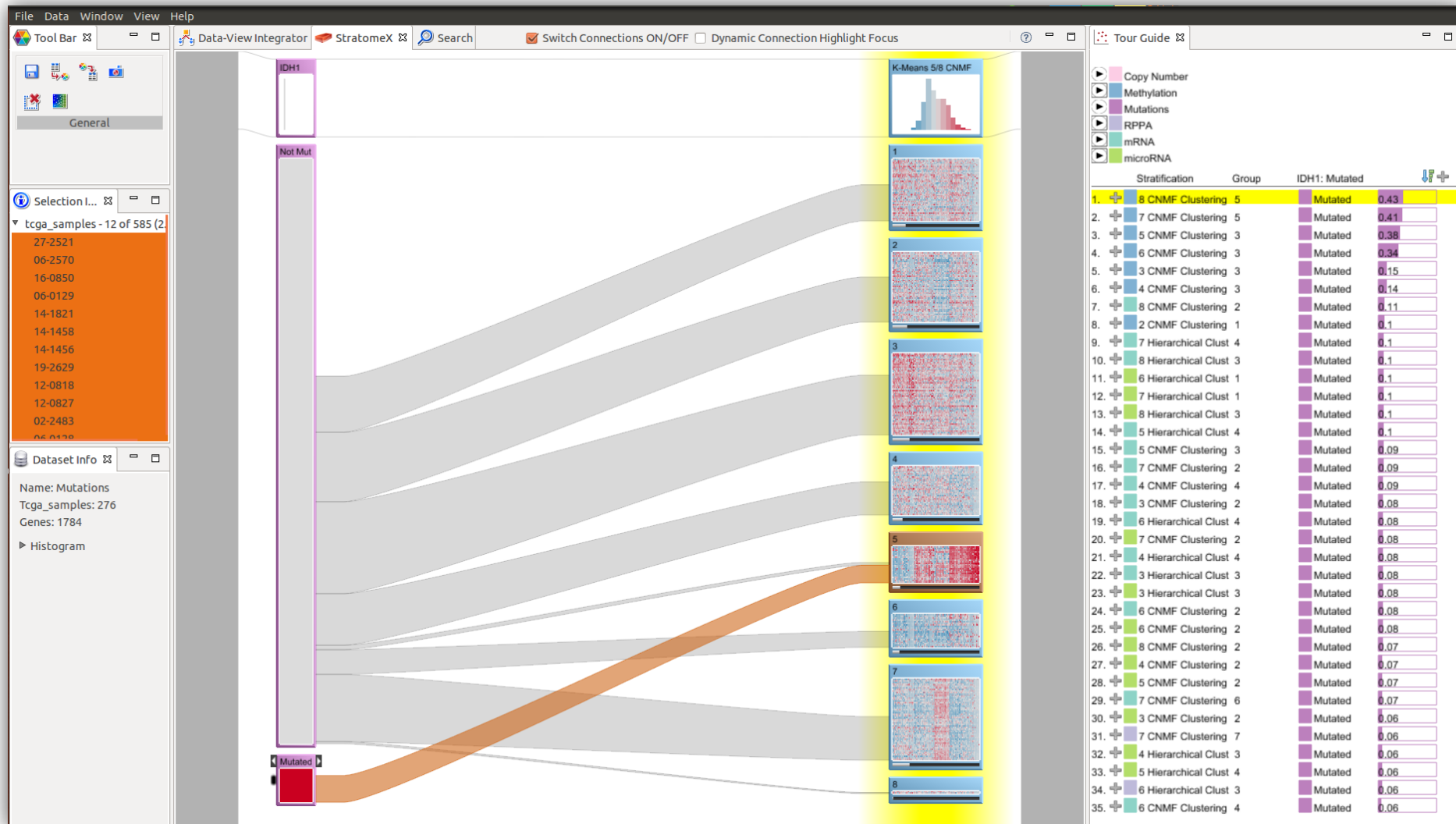
Tour Guide



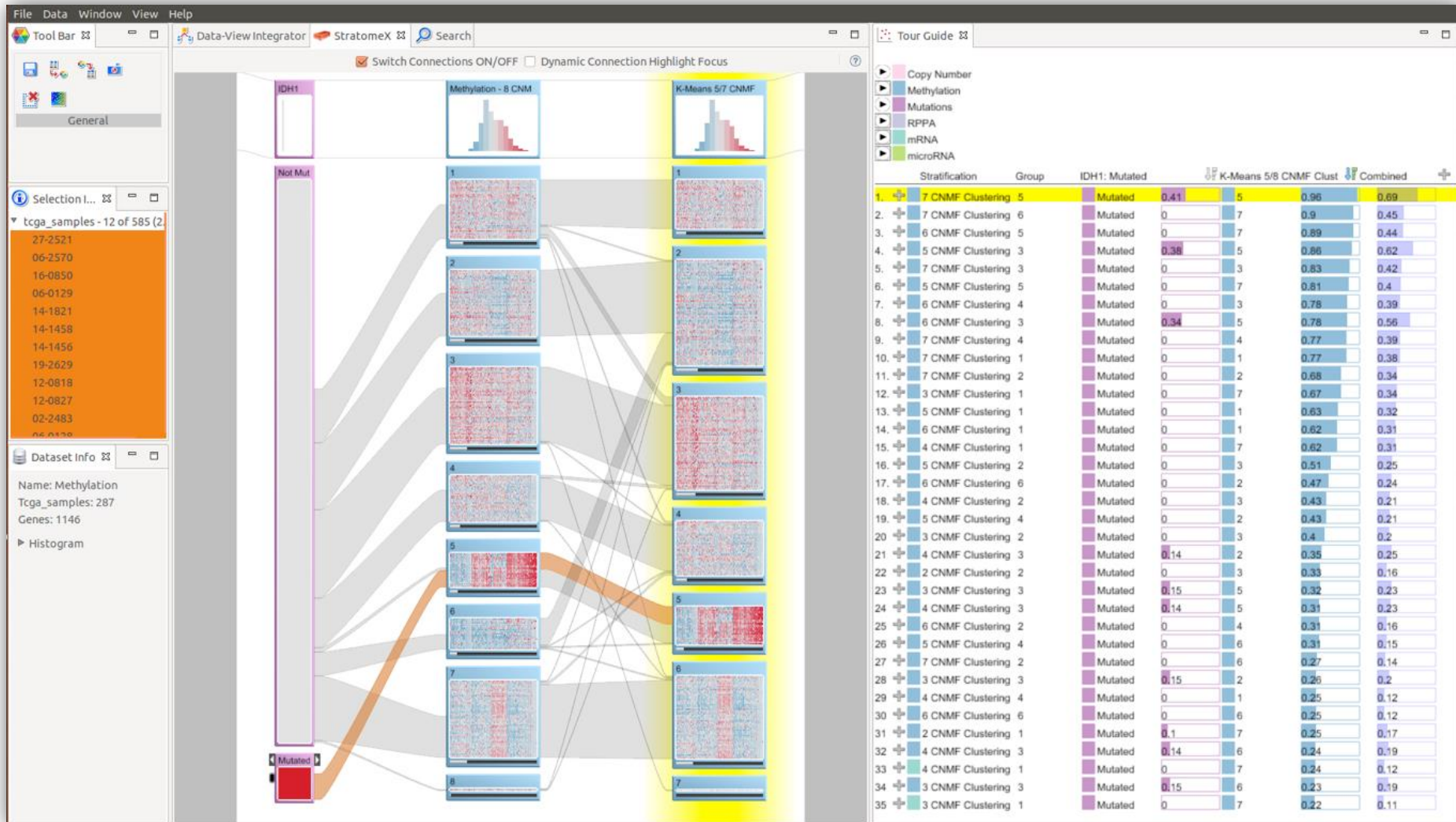
Tour Guide Approach

- Q: Which stratifications are most similar to this mRNA clustering?*
- Q: Which stratifications contain a subset with the same patients who also have a homozygous deletion of this gene?*
- Q: What are the top pathways containing genes with copy number alteration or mutations?*

StratomeX Tour Guide



Combine Complex Queries



BACK TO ALEX...

PATHWAY VISUALIZATION

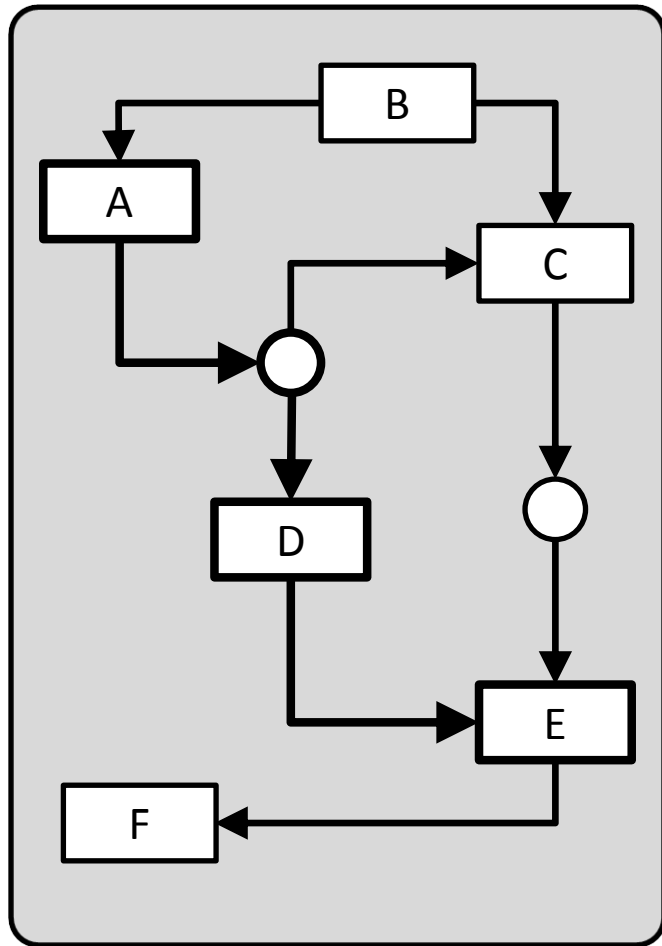
Experimental Data and Pathways

Pathways represent **consensus knowledge** for a healthy organism or specific disease

Cannot account for **variation** found in real-world data

Branches can be **(in)activated** due to
mutation,
changed gene expression,
modulation due to drug treatment,
etc.

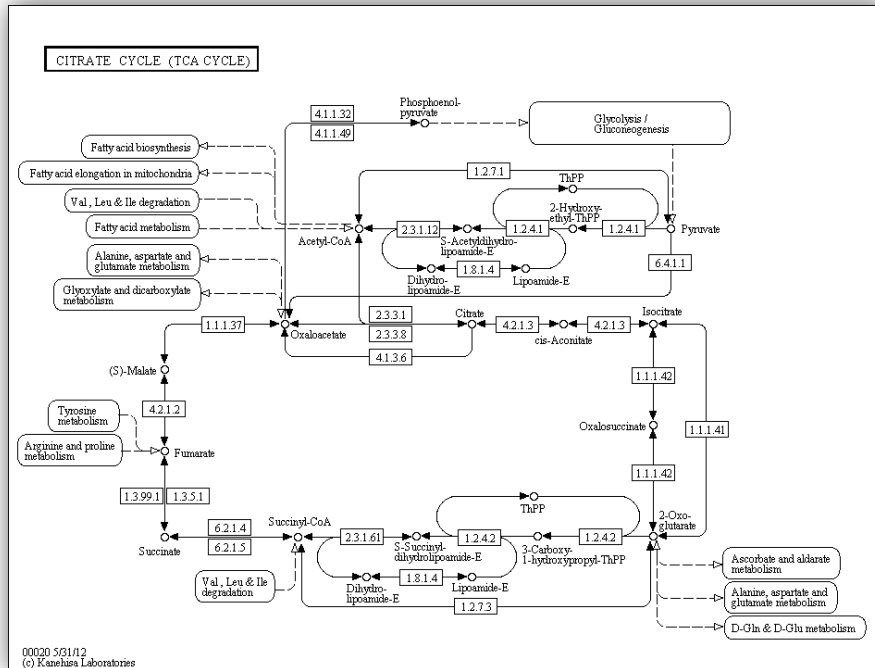
Why use Visualization?



Efficient communication of information

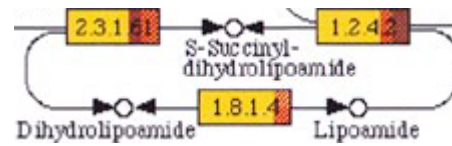
A	-3.4
B	2.8
C	3.1
D	-3
E	0.5
F	0.3

Experimental Data and Pathways

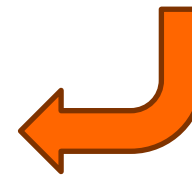


	A	B	C	D	E	F	G
1	Hybridization	0001-01C-01R-0177-01	0003-01A-01R-0177-01	0004-01A-01R-0298-01	0007-01A-01R-0177-01	0009-01A-01R-0177-01	0010-01A-01R-0177-01
2	Composite Ek	Signal	Signal	Signal	Signal	Signal	Signal
3	AACS	0.51414972642765	0.408345300354386	103.839.280.872.892	108.369.722.813.568	146.228.041.324.369	189.518.994.867.335
4	FSTL1	274.326.077.373.904	366.350.031.165.199	572.009.498.050.765	-115.814.163.115.937	462.733.779.632.644	283.075.033.965.805
5	ELMO2	-0.47503962404922	0.083081303877976	0.71228305118261	-0.872750104404229	0.407918562894304	145.159.574.522.917
6	CREB3L1	-110.344.925.188.626	-129.462.362.872.386	-0.227010846693110	-0.284588481211029	-0.879544579484095	-0.0134788669146966
7	RPS11	499.838.216.564.566	468.034.485.053.486	409.691.860.929.175	437.410.995.409.265	442.451.017.149.714	490.071.296.665.235
8	PNMA1	154.879.089.253.498	0.867033048134666	27.926.044.547.714	225.292.305.165.001	0.730144175027465	194.355.497.056.736
9	MMP2	268.760.812.835.858	321.769.033.459.614	490.900.567.411.305	-167.367.360.374.371	323.947.380.237.154	-159.909.546.919.508
10	SAMD4A	-0.95385000100184	-110.393.925.981.711	-0.780339641097821	-0.927255839038269	-104.784.345.382.755	-113.068.517.857.268
11	SMARCD3	-12.754.315.263.248	-0.803166248235193	192.878.200.580.422	-166.272.411.648.149	-126.054.982.166.486	253.814.516.975.109
12	A4GNT	-0.8779238907219	-178.229.413.280.745	-22.870.781.160.139	-127.901.084.268.365	-200.815.270.464.235	-137.113.826.496.947
13	C9orf39	-119.755.123.311.627	-199.689.453.294.211	-170.879.710.400.263	-197.903.696.598.943	-184.196.451.349.235	-123.143.505.917.242
14	PKNOX2	-0.95454191801256	-129.980.559.337.036	-129.725.303.429.725	-117.863.280.467.411	-169.952.892.104.581	-0.914002701548987
15	RALYL	-163.237.829.190.046	-223.568.546.117.929	-202.107.025.240.506	-167.211.920.052.532	-226.223.976.628.069	-180.973.590.502.948
16	ZHX3	-0.92856964251013	-111.244.404.270.456	0.0565663069364497	-0.794043792350179	-101.938.096.745.295	-0.413263822075167
17	ERCC5	0.95243371908686	0.957680305439156	0.9955322244338	0.238642819120941	153.006.779.612.285	-0.01226896339307
18	RXFP3	-206.267.350.950.005	-247.628.615.578.476	-258.794.340.437.298	-243.076.988.610.523	-261.197.965.543.047	-216.635.887.704.531
19	APBB2	-101.794.466.906.032	-139.226.919.342.780	0.7396843984963	-160.525.972.599.955	-129.292.861.167.009	-153.166.942.410.299
20	BBOX1	-219.765.356.842.015	-216.619.193.833.608	189.217.944.622.057	-223.116.920.533.355	-242.172.086.385.896	-172.405.872.547.507
21	PRO0478	-208.735.486.814.708	-201.266.491.273.900	-264.514.012.419.739	-183.469.408.961.363	-228.483.271.853.159	-222.688.943.536.626
22	GCSH	214.667.175.620.209	214.425.000.089.483	282.468.317.897.215	379.400.938.009.765	234.157.818.929.887	305.821.981.214.471
23	XDH	-140.335.573.729.949	-154.352.054.860.075	-195.598.156.263.592	-164.927.258.900.381	-153.786.759.540.717	-150.754.123.452.277
24	EDN1	-0.27265835556173	0.0601082045242967	-0.686221828128451	-0.0703902946165886	0.241048141573394	-0.681241448310897
25	MTERF	0.33456496873088	0.602787120981156	116.205.672.714.072	0.290137408663012	0.339834697015724	0.376698230530083
26	PDCL3	0.22161805747679	192.980.867.583.586	171.282.568.732.192	129.218.742.469.691	183.026.664.772.145	211.448.961.587.135
27	CLK4	0.11771672245876	0.837990640842366	0.48153932013633	0.101190616384591	0.461222798318555	0.547200380309174
28	KCNQ1	-145.735.076.947.667	-154.967.690.506.818	-173.914.329.989.365	-168.541.602.431.989	-196.875.877.389.228	-190.608.188.831.993
29	CXCR4	-187.536.347.744.716	263.335.862.666.286	228.888.923.555.987	-145.426.654.150.337	277.993.551.197.589	32.964.742.520.002
30	DEC1	267.353.839.615.507	243.484.300.334.252	340.595.900.405.511	332.766.227.083.077	300.755.122.393.034	195.692.044.856.851
31	SALL1	-131.231.146.088.135	-206.830.763.066.117	-167.501.082.187.998	-196.742.986.889.255	-226.876.961.736.839	-0.851620028342527
32	PTPRR	-175.072.128.842.408	-215.050.278.107.699	-201.907.311.040.685	-0.618976507439188	-0.453120032708616	166.423.598.869.425
33	CAD4	-159.091.940.638.954	-196.549.713.788.350	-224.919.487.589.083	-163.188.842.517.916	-202.074.958.245.880	-179.879.615.356.428
34	IRAK1	315.954.017.145.736	299.254.943.418.361	356.221.429.575.433	368.207.420.675.573	332.140.514.206.449	197.043.270.833.661

[KEGG]

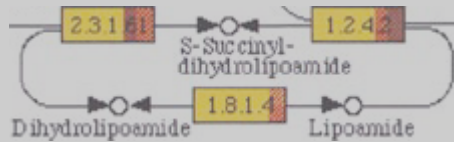


[Lindroos2002]

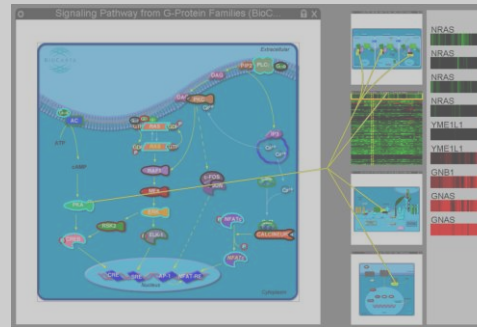


Visualization Approaches

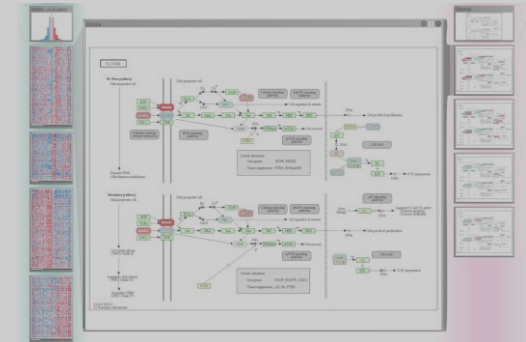
[Lindroos 2002]



On-Node Mapping

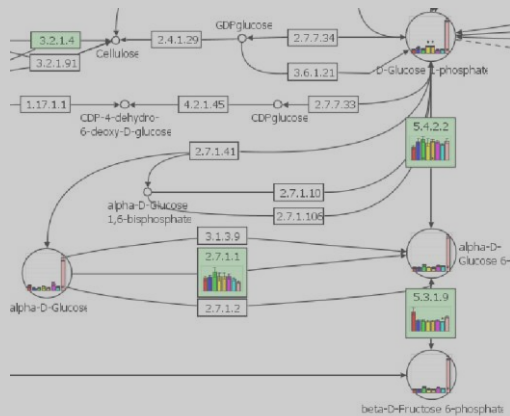


Separate Linked Views



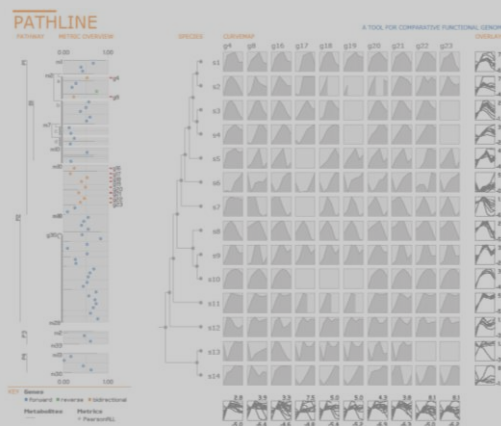
Small Multiples

[Junker 2006]

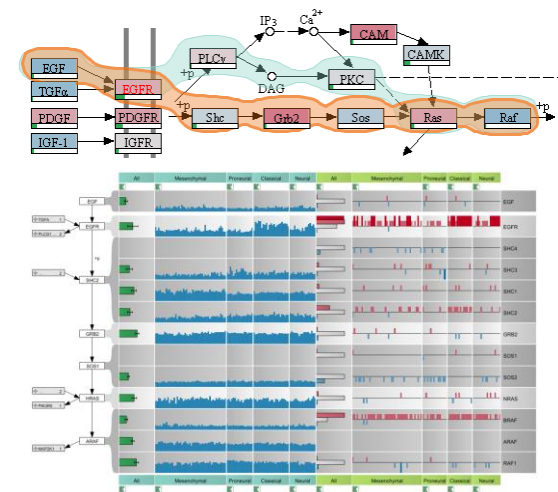


Layout Adaption

[Meyer 2010]



Linearization



Path-Extraction

REQUIREMENTS FOR VISUALIZING PATHWAYS AND EXPERIMENTAL DATA

What to Consider when Visualizing Experimental Data and Pathways

Conflicting Goals

Preserving **topology** of pathways

Showing **lots of experimental data**

Five Requirements

Ideal visualization technique addresses all

R I: Data Scale

Large **number of experiments**

Large datasets have more than 500 experiments

Multiple **groups/conditions**

R II: Data Heterogeneity

Different **types** of data, e.g.,

mRNA expression

numerical

mutation status

categorical

copy number variation

ordered categorical

metabolite concentration

numerical

Require **different visualization** techniques

R III: Multi-Mapping

Pathways nodes are biomolecules

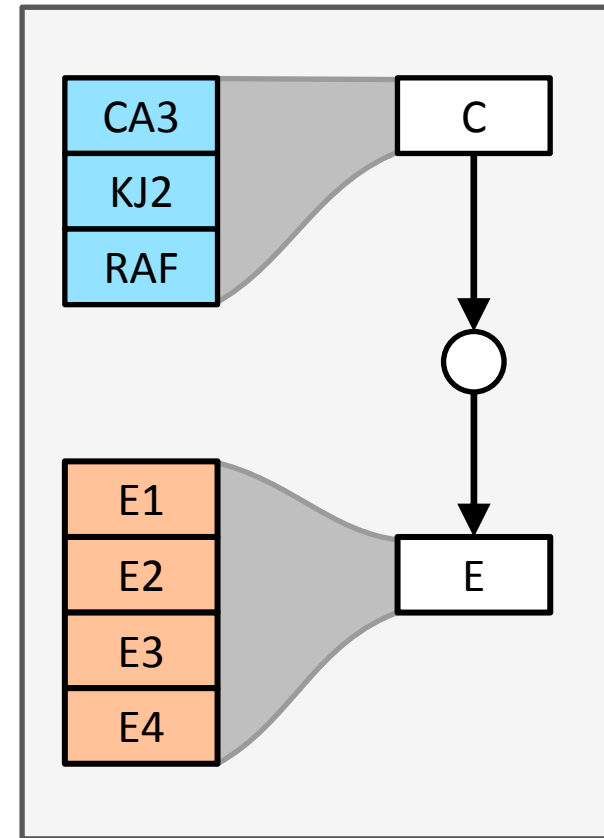
Proteins, nucleic acids, lipids,
metabolites

Experimental data often
on a „gene“ level

Multiple genes can produce protein

Multiple genes encode one protein

Result: many „gene“ values map
to one pathway node



R IV: Preserving the Layout

Pathways are available in
carefully designed layouts

e.g., KEGG, WikiPathways, Biocarta

Users are familiar with layouts

Goal: preserve layouts as much as possible

Two approaches:

Emulate drawing conventions

Use original layouts

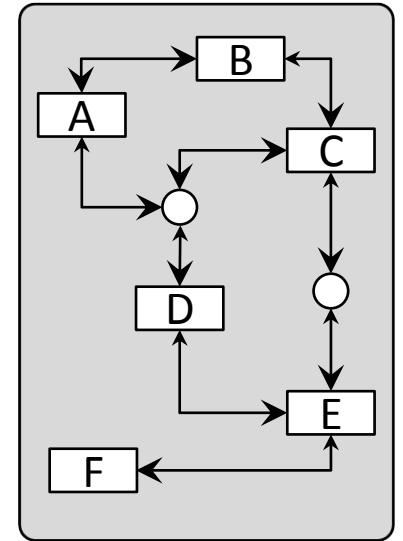
R V: Supporting Multiple Tasks

Two central tasks:

Explore **topology** of pathway

Explore the **attributes** of the nodes
(experimental data)

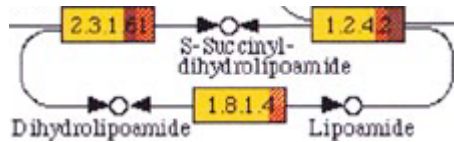
Need to support both!



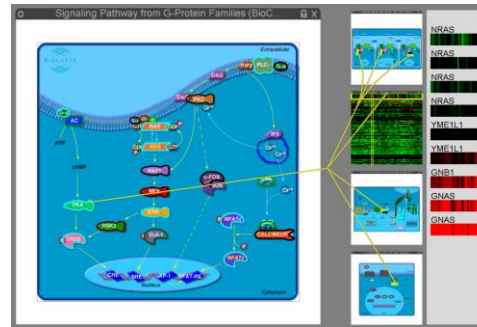
	Sample 1	Sample 2	Sample 3
Gene 1	1	1.1	0.4
Gene 2	2	0.5	1.2
Gene 3	1.4	0.2	0.5
Gene 4	0.3	0.5	0.7

Visualization Approaches

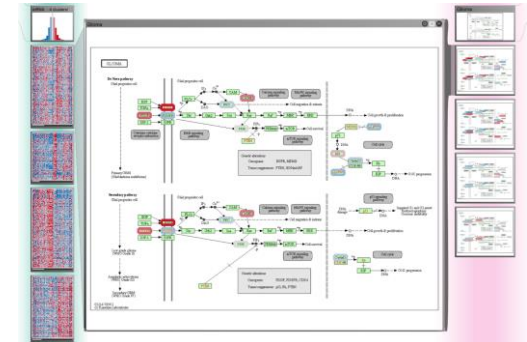
[Lindroos 2002]



On-Node Mapping

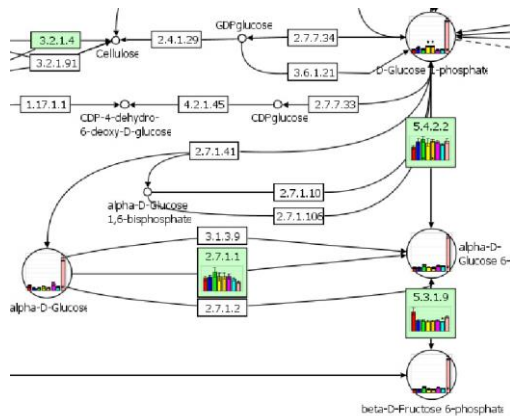


Separate Linked Views



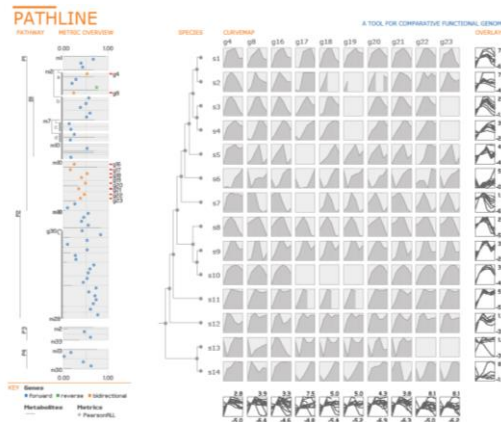
Small Multiples

[Junker 2006]

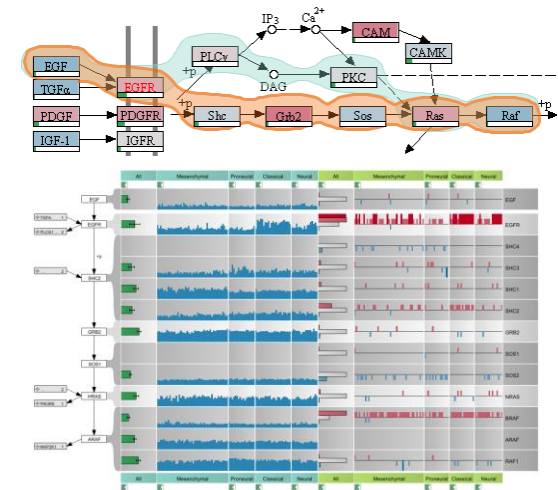


Layout Adaption

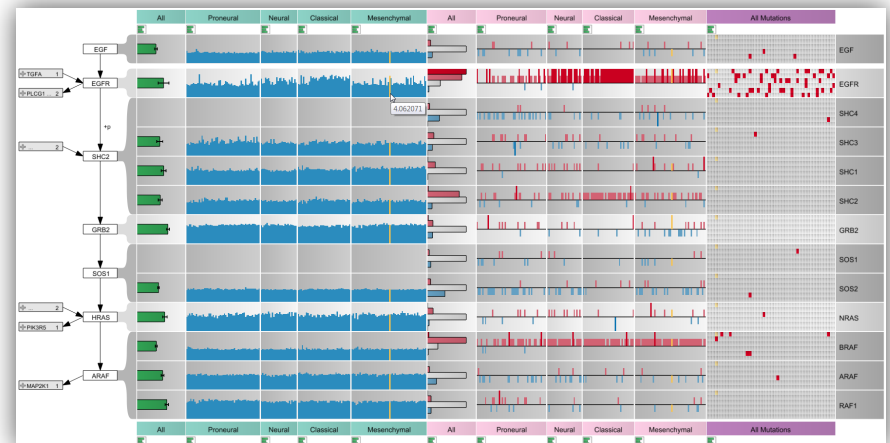
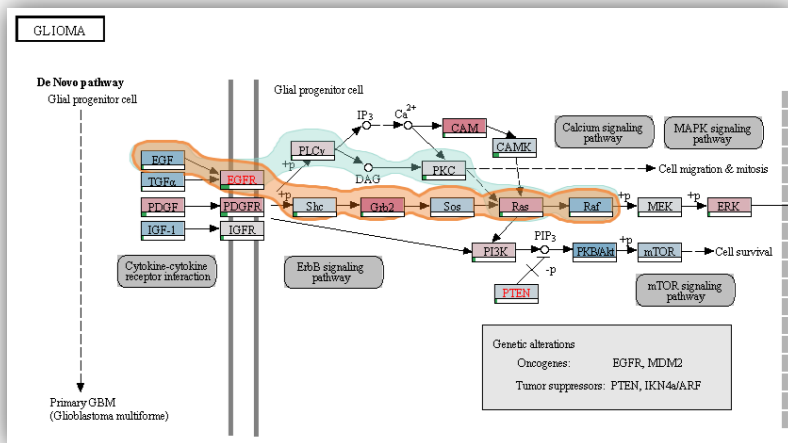
[Meyer 2010]



Linearization

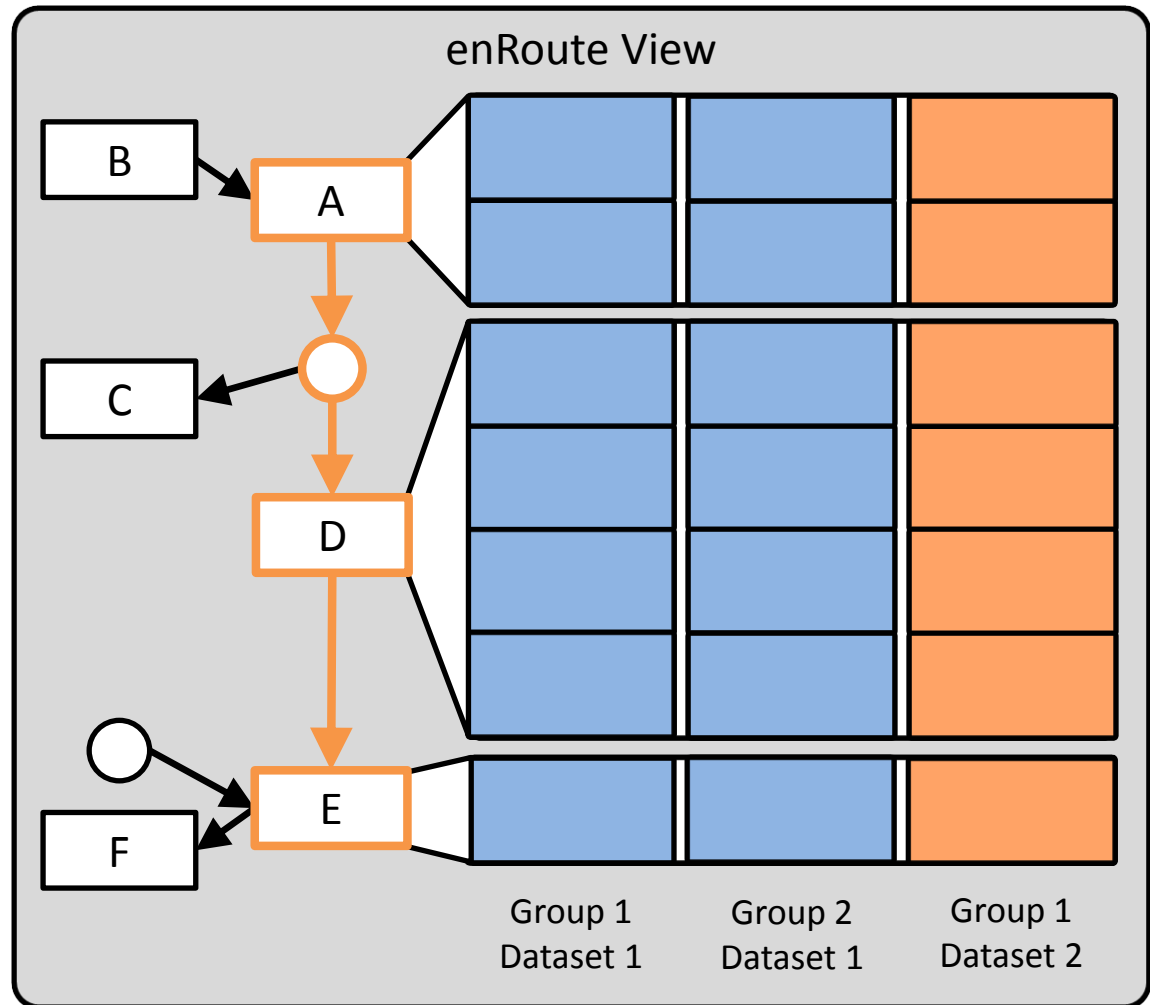
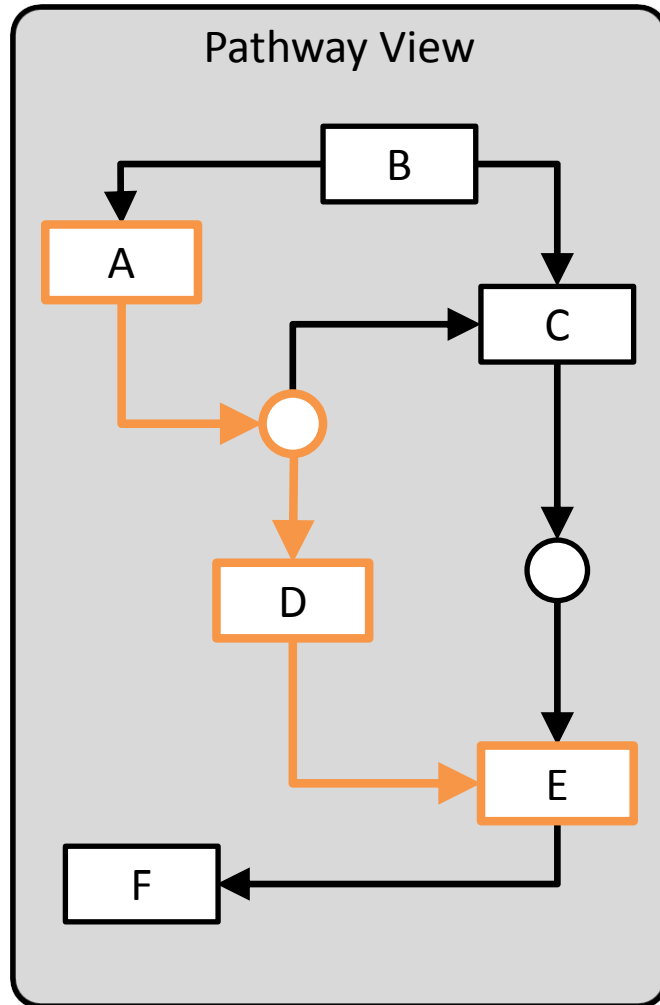


Path-Extraction

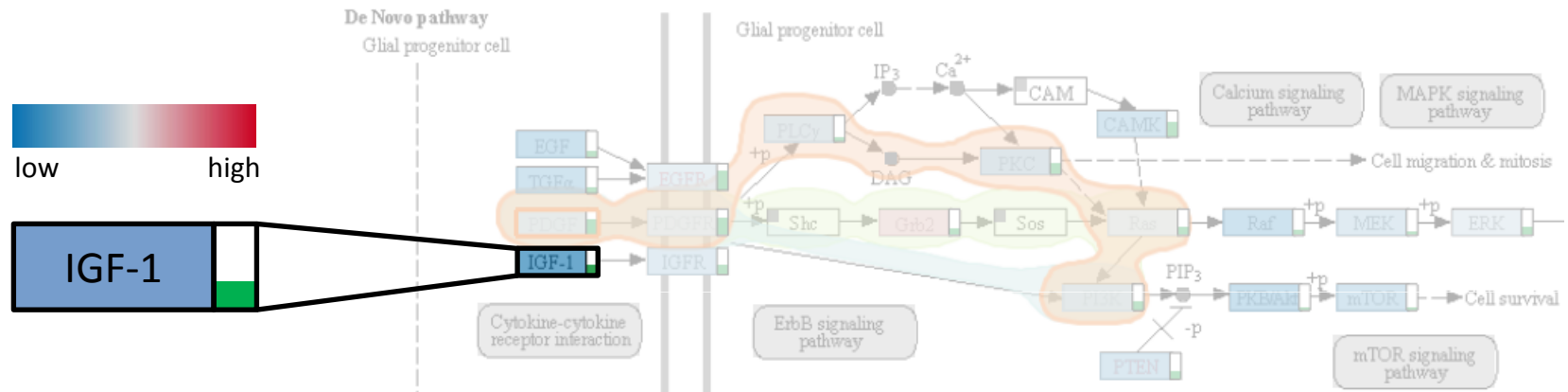


CALEYDO ENROUTE

Concept



Pathway View



On-Node Mapping

Path highlighting with Bubble Sets [Collins2009]

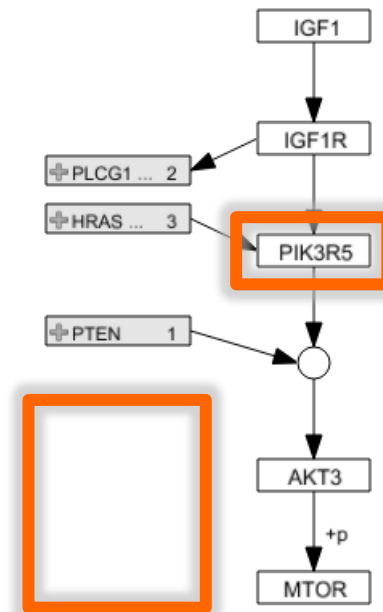
Selection

Start- and end node

Iterative adding of nodes

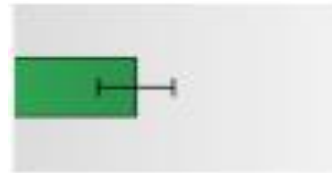
enRoute View – Path Representation

- Design of KEGG [Kanehisa2008]
- Abstract branch nodes
 - Additional topological information
 - Incoming vs. outgoing branches
 - Expandable
- Branch switching

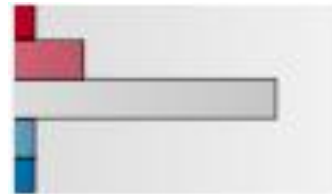
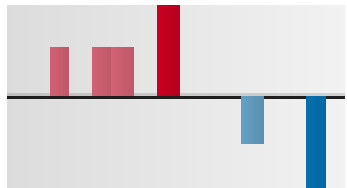


Experimental Data Representation

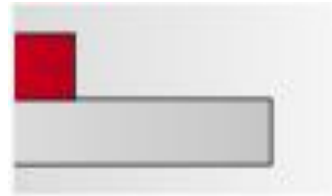
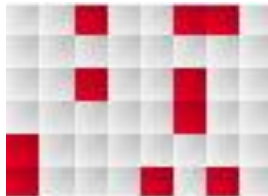
Gene Expression Data (Numerical)



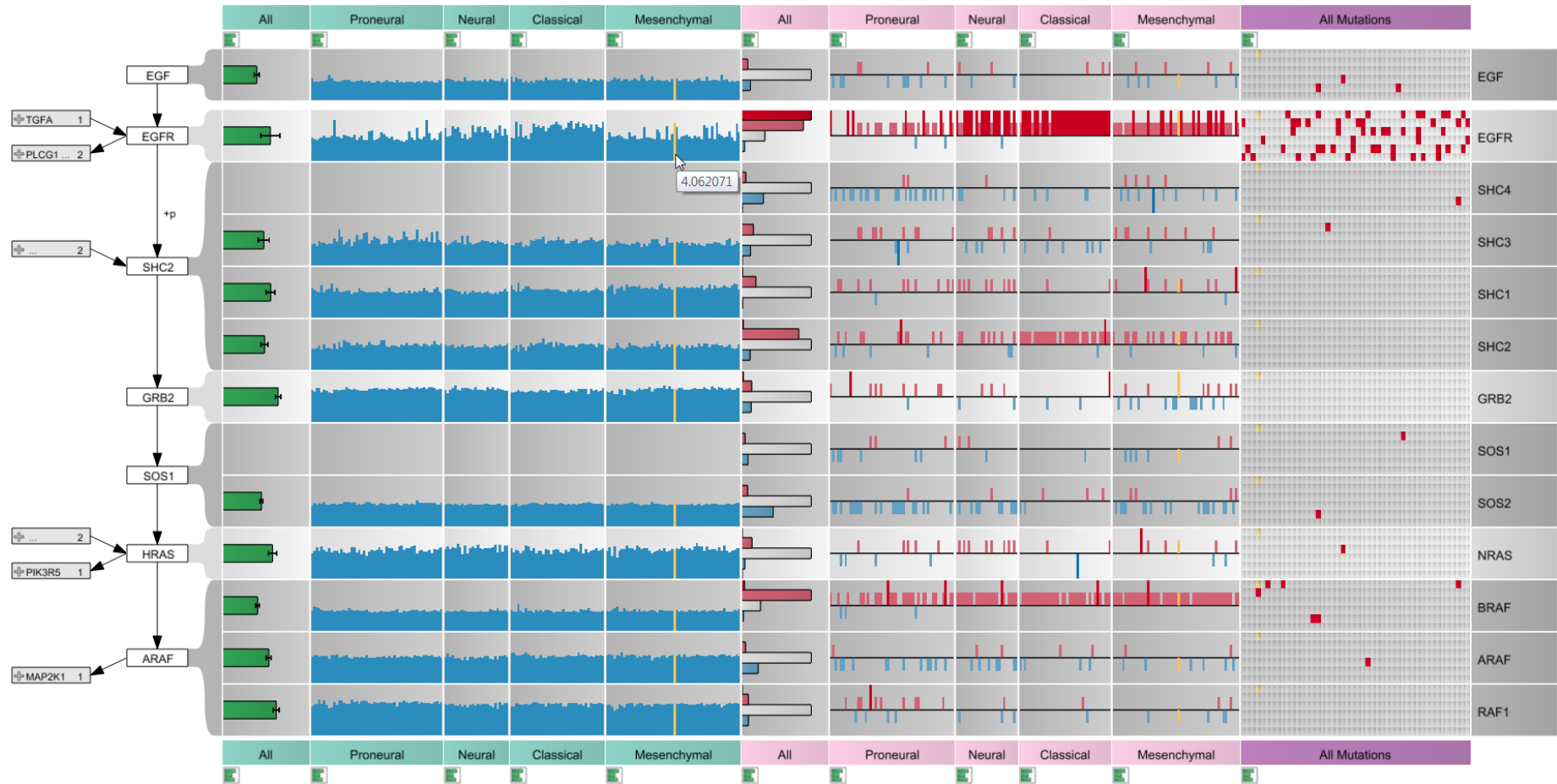
Copy Number Data (Ordered Categorical)



Mutation Data



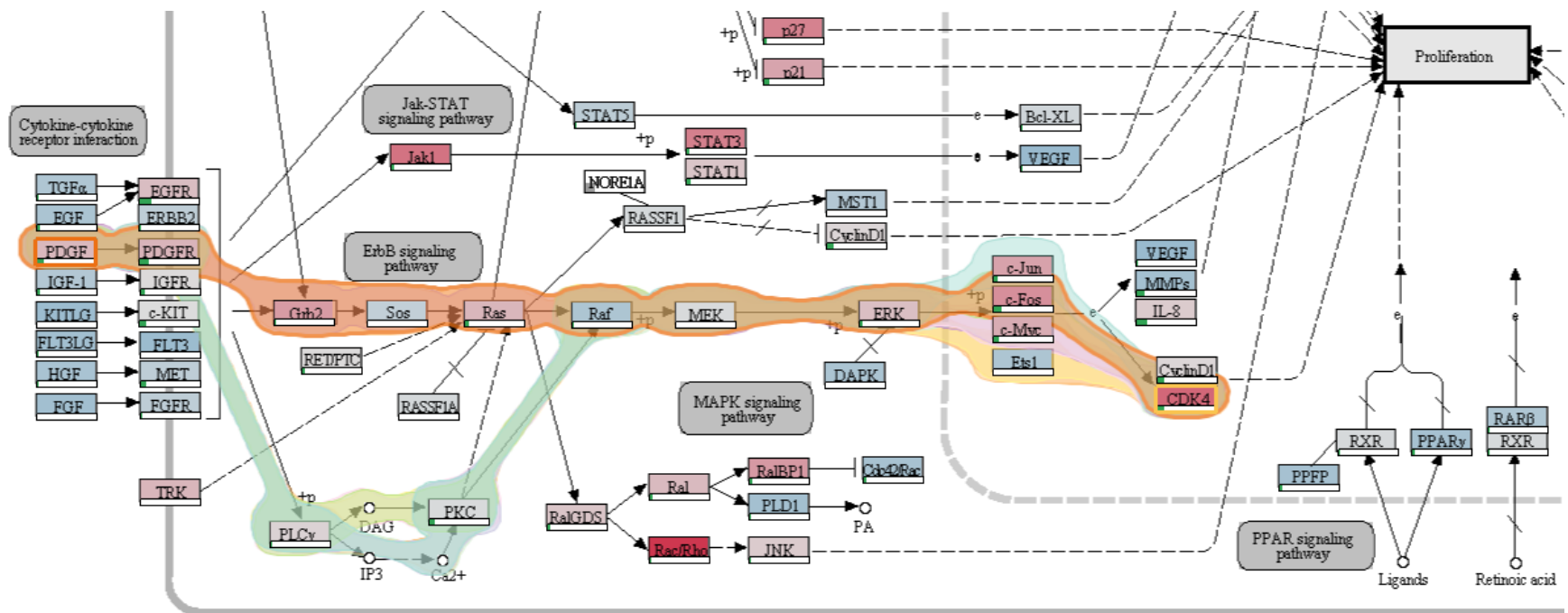
enRoute View – Putting All Together



Demo Video: <http://enroute.caleydo.org>

Live Demonstration!

Glioblastoma Multiforme Example



Glioblastoma Multiforme Example



Please cite the enRoute paper

C. Partl, A. Lex, M. Streit, D. Kalkofen, K. Kashofer, and D. Schmalstieg, *enRoute: Dynamic Path Extraction from Biological Pathway Maps for In-Depth Experimental Data Analysis*, in Proceedings of the IEEE Symposium on Biological Data Visualization (BioVis '12), 2012, pp. 107–114.

CONCLUSION

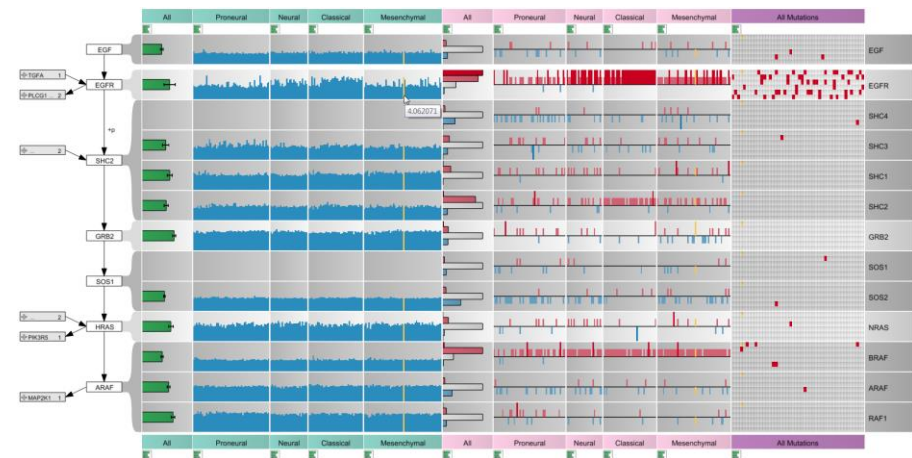
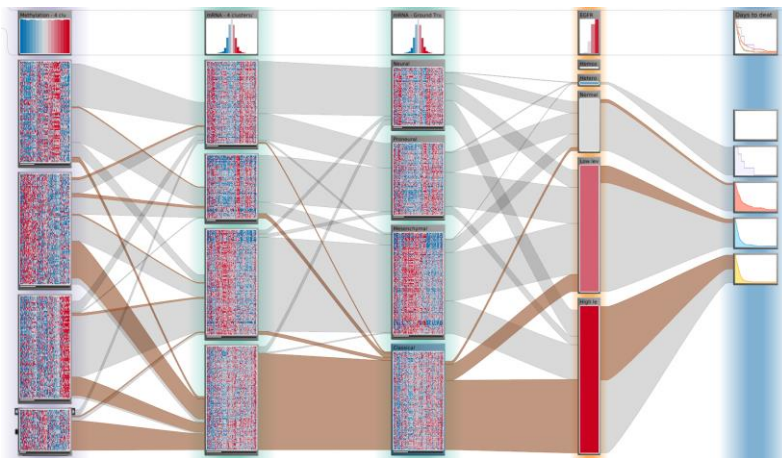
Caleydo – Multi-Dataset Analysis

For relationships of

multiple datasets and tabular and graph data

Free to use for all

Feedback and help welcome anytime!



Resources

Project website

<http://caleydo.org>

Caleydo help

<http://help.caleydo.org>

Pre-packaged TCGA data

<http://tcga.caleydo.org>

Contact

caleydo@icg.tugraz.at



Cancer Data Analysis with Caleydo StratomeX and enRoute

Christian Partl
Samuel Gratzl
Nils Gehlenborg
Dieter Schmalstieg
Peter J. Park
Hans-Jörg Schulz

Alexander Lex, Harvard University
alex@seas.harvard.edu
<http://alexander-lex.com>



HARVARD

**School of Engineering
and Applied Sciences**

Marc Streit, JKU Linz
marc.streit@jku.at
<http://marc.streit.com>



JKU
JOHANNES KEPLER
UNIVERSITÄT LINZ



CALEYDO