Use Case 13: Analysis of epigenomic variation in breast tumors

Epigenome Informatics Workshop
Bioinformatics Research Laboratory

BCM Baylor College of Medicine

The Data Used in Case Study 13

1. Dedeurwaerder, S.et al. (2011) "Evaluation of the Infinium Methylation 450K technology", Epigenomics 3(6):771-84.

2. **16 breast tissue samples were profiled** (8 normal, 8 primary tumor samples)

3. The paper evaluates 450K technology and does not report any analysis concerning cancer biology.

4. We analyze the data to explore epigenomic states and cell type composition of normal and tumor samples.

Overview of Case Study 13

Part 1: Cluster all 16 breast tissue samples using the Epigenomic Heatmap tool

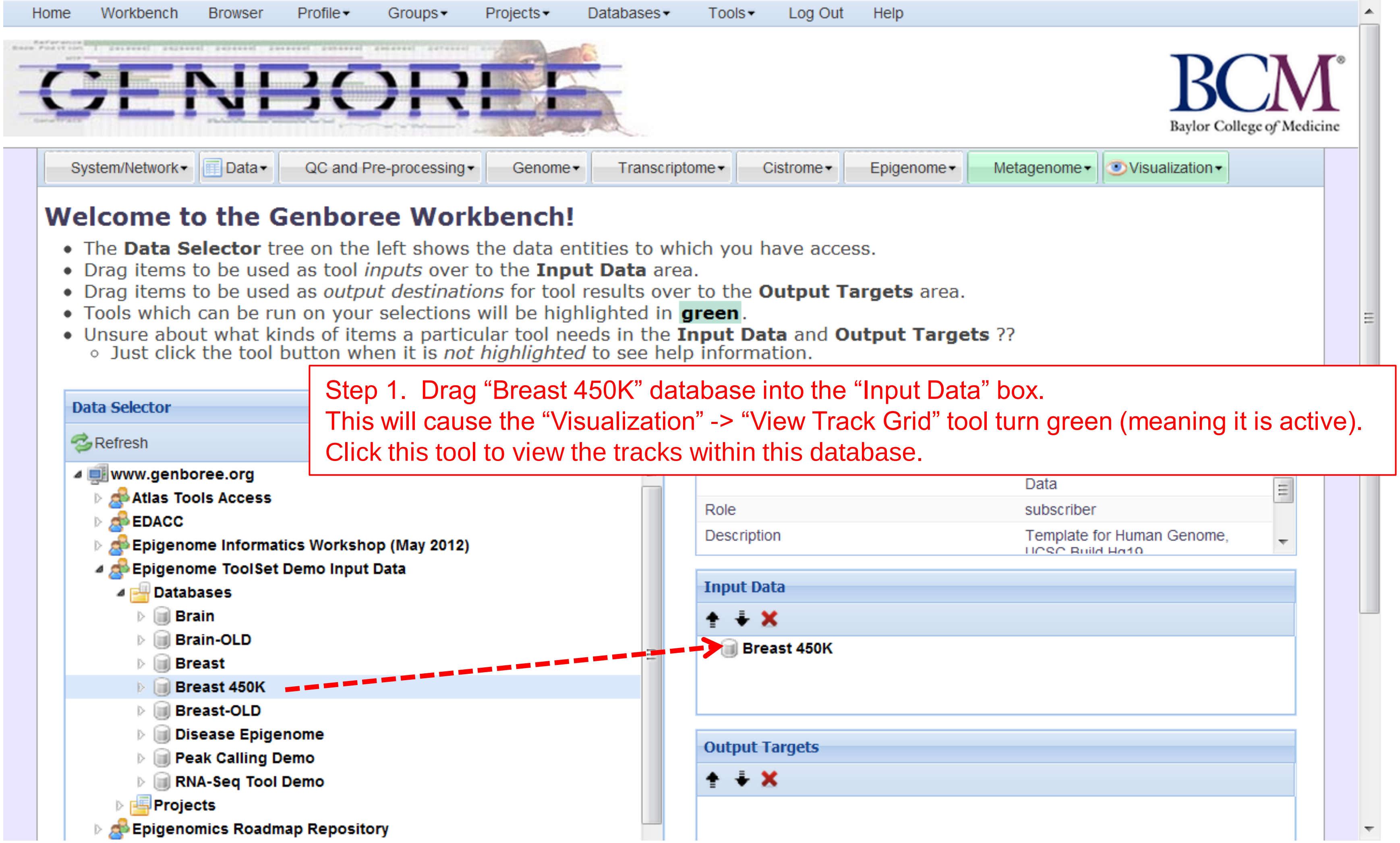
Part 2: Compare the 16 samples against the Human Epigenome Atlas to determine tissue compositions using the Epigenomic Heatmap tool

Part 3: Identify differentially methylated probes, genes, and pathways between 8 normal and 8 tumor samples using the LIMMA comparison tool

Part 1: Cluster all 16 breast tissue samples

Part 1: Introduction

- We will apply **Epigenomic Heatmap tool** to cluster all 16 breast tissue samples.
- We anticipate a pattern discriminating 8 tumor samples from 8 normal samples.
- The tumor samples may be more heterogeneous than normal samples


Home
Workbench
Browser
Profile
Groups
Projects
Databases
Tools
Log Out
Help
GENBOREE
BCM
Baylor College of Medicine
System/Network
Data
QC and Pre-processing
Genome
Transcriptome
Cistrome
Epigenome
Metagenome
Visualization
Data Selector
Refresh
www.genboree.org
Atlas Tools Access
EDACC
Epigenome Informatics Workshop (May 2012)
Epigenome ToolSet Demo Input Data
Databases
Brain
Brain-OLD
Breast
Breast 450K
Breast-OLD
Disease Epigenome
Peak Calling Demo
RNA-Seq Tool Demo
Projects
Epigenomics Roadmap Repository
Input Data
Breast 450K
Output Targets

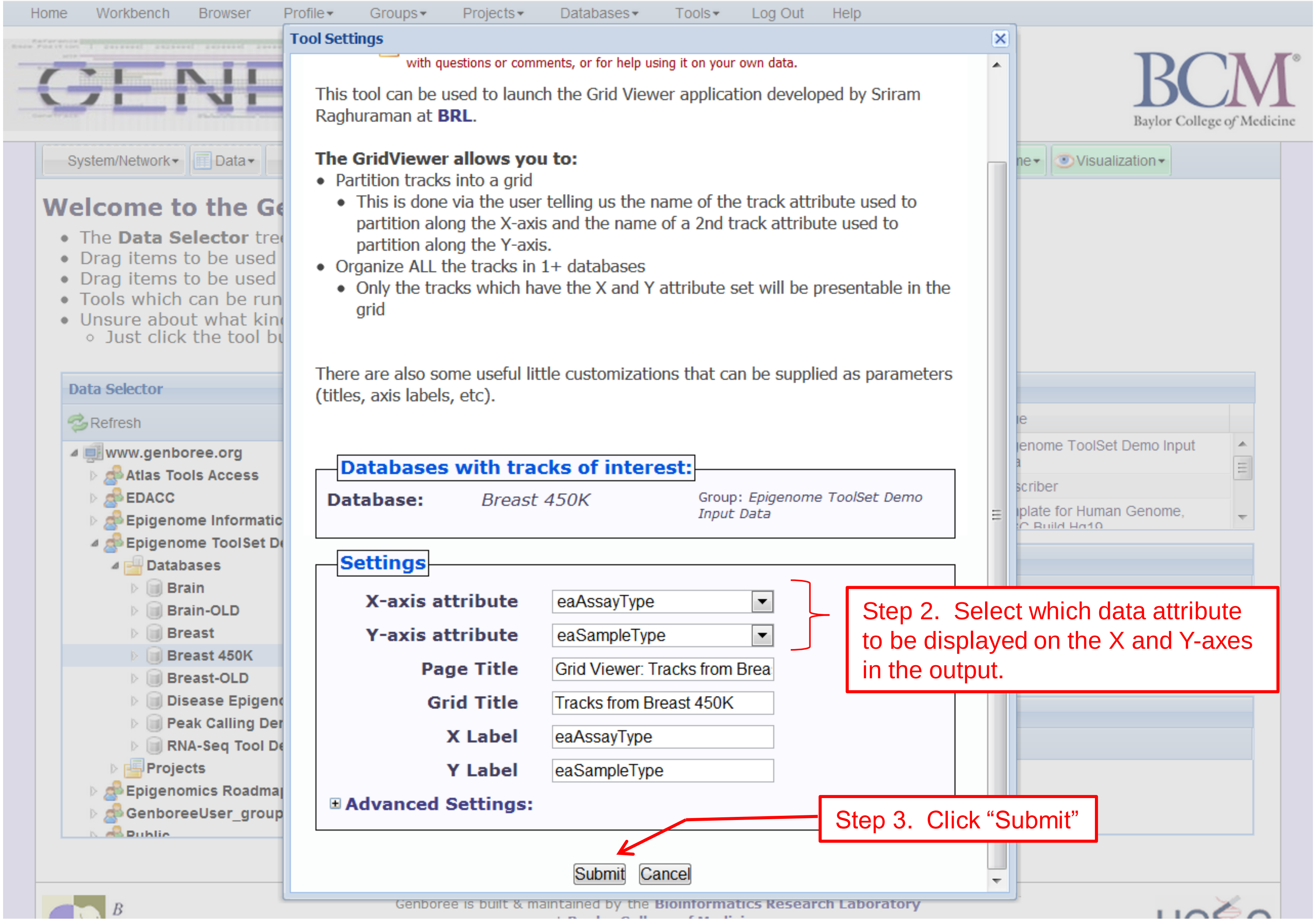
Welcome to the Genboree Workbench!

- The **Data Selector** tree on the left shows the data entities to which you have access.
- Drag items to be used as tool *inputs* over to the **Input Data** area.
- Drag items to be used as *output destinations* for tool results over to the **Output Targets** area.
- Tools which can be run on your selections will be highlighted in **green**.
- Unsure about what kinds of items a particular tool needs in the **Input Data** and **Output Targets** ??
  - Just click the tool button when it is *not highlighted* to see help information.

	Data
Role	subscriber
Description	Template for Human Genome, UCSC Build Hg19

Step 1. Drag “Breast 450K” database into the “Input Data” box.
This will cause the “Visualization” -> “View Track Grid” tool turn green (meaning it is active).
Click this tool to view the tracks within this database.

Part 1: Create Your Data Grid-Con’d


Home
Workbench
Browser
Profile
Groups
Projects
Databases
Tools
Log Out
Help
System/Network
Data
Visualization
BCM
Baylor College of Medicine
Data Selector
Refresh
www.genboree.org
Atlas Tools Access
EDACC
Databases
Brain
Brain-OLD
Breast
Breast 450K
Breast-OLD
Projects
Tool Settings

with questions or comments, or for help using it on your own data.

This tool can be used to launch the Grid Viewer application developed by Sriram Raghuraman at **BRL**.

The GridViewer allows you to:

- Partition tracks into a grid
  - This is done via the user telling us the name of the track attribute used to partition along the X-axis and the name of a 2nd track attribute used to partition along the Y-axis.
- Organize ALL the tracks in 1+ databases
  - Only the tracks which have the X and Y attribute set will be presentable in the grid

There are also some useful little customizations that can be supplied as parameters (titles, axis labels, etc).

Databases with tracks of interest:

Database: *Breast 450K* Group: *Epigenome ToolSet Demo Input Data*

Settings

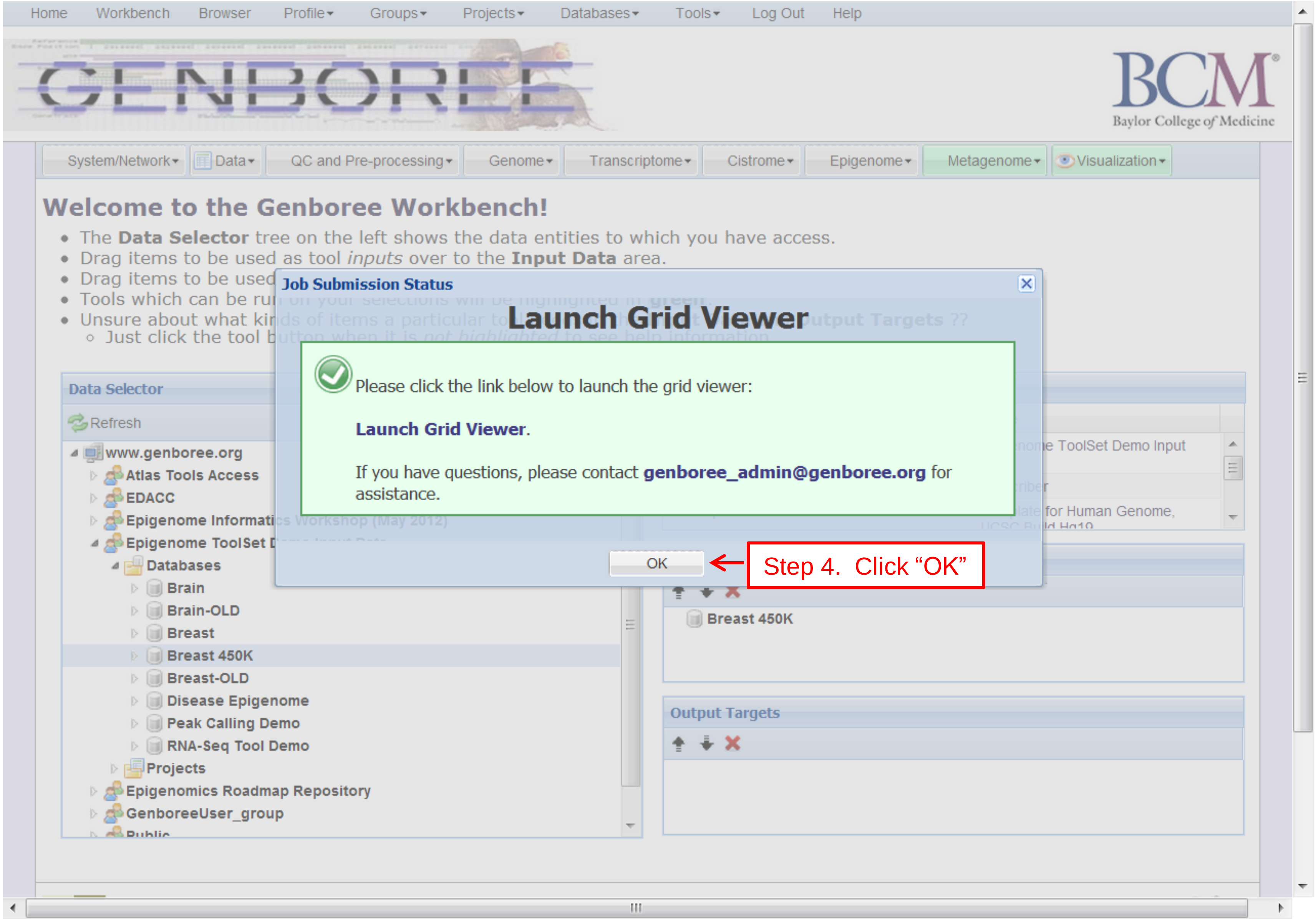
X-axis attribute	eaAssayType
Y-axis attribute	eaSampleType
Page Title	Grid Viewer: Tracks from Brea
Grid Title	Tracks from Breast 450K
X Label	eaAssayType
Y Label	eaSampleType

⊞ **Advanced Settings:**

Submit Cancel

Step 2. Select which data attribute to be displayed on the X and Y-axes in the output.

Step 3. Click “Submit”


Home
Workbench
Browser
Profile
Groups
Projects
Databases
Tools
Log Out
Help
GENBOREE
BCM
Baylor College of Medicine
System/Network
Data
QC and Pre-processing
Genome
Transcriptome
Cistrome
Epigenome
Metagenome
Visualization
Welcome to the Genboree Workbench!
Data Selector
Refresh
www.genboree.org
Atlas Tools Access
EDACC
Epigenome Informatics Workshop (May 2012)
Epigenome ToolSet D
Databases
Brain
Brain-OLD
Breast
Breast 450K
Breast-OLD
Disease Epigenome
Peak Calling Demo
RNA-Seq Tool Demo
Projects
Epigenomics Roadmap Repository
GenboreeUser_group
Public
Breast 450K
Output Targets
OK

Job Submission Status

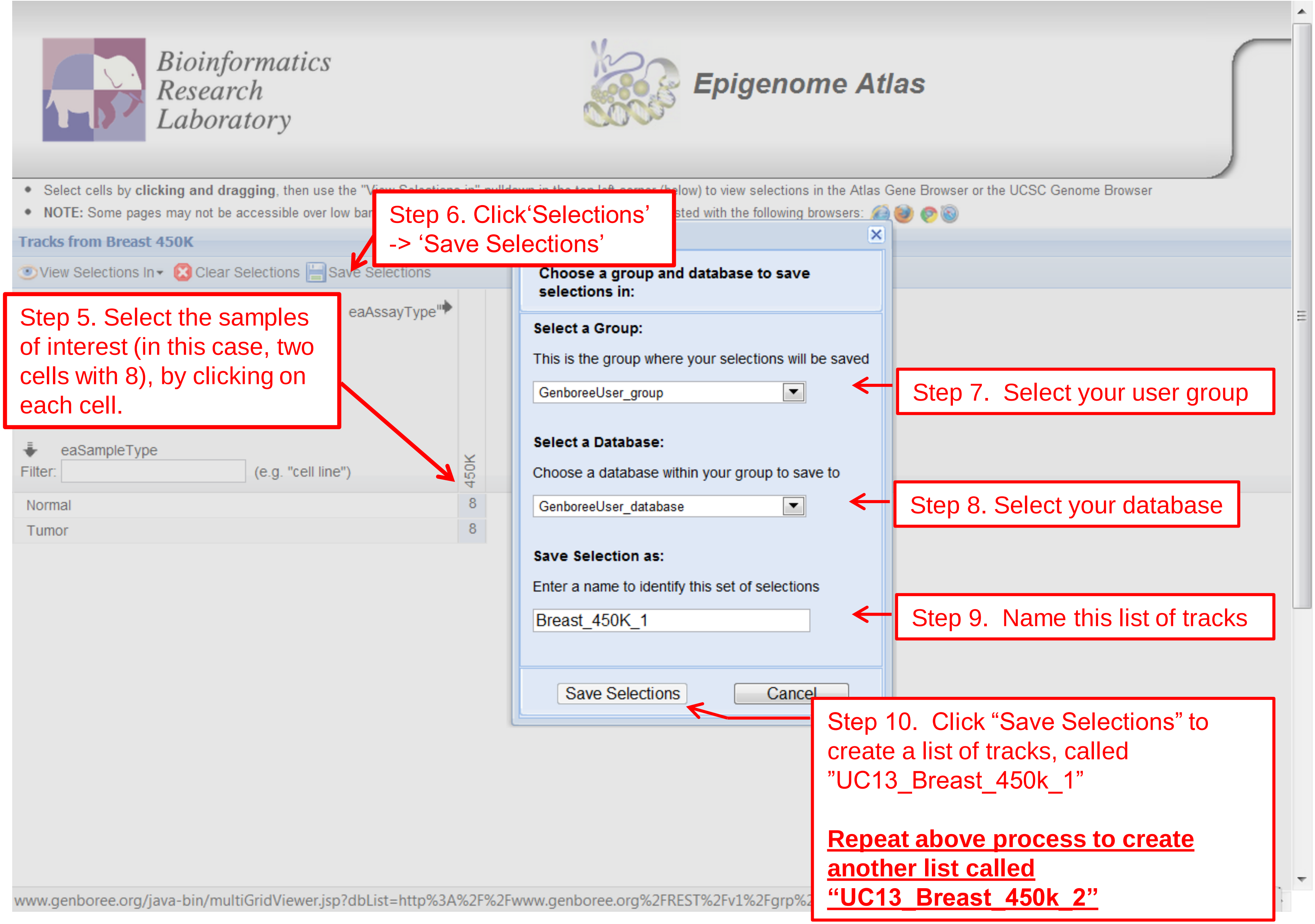
Launch Grid Viewer

Please click the link below to launch the grid viewer:

Launch Grid Viewer.

If you have questions, please contact **genboree_admin@genboree.org** for assistance.

Step 4. Click “OK”


Bioinformatics Research Laboratory
Epigenome Atlas

- Select cells by **clicking and dragging**, then use the "V… (below) to view selections in the Atlas Gene Browser or the UCSC Genome Browser
- **NOTE:** Some pages may not be accessible over low ban… sted with the following browsers:

Tracks from Breast 450K

View Selections In
Clear Selections
Save Selections
eaAssayType
eaSampleType
Filter:
(e.g. "cell line")

	450K
Normal	8
Tumor	8

Choose a group and database to save selections in:

Select a Group:

This is the group where your selections will be saved

GenboreeUser_group

Select a Database:

Choose a database within your group to save to

GenboreeUser_database

Save Selection as:

Enter a name to identify this set of selections

Breast_450K_1

Save Selections
Cancel

www.genboree.org/java-bin/multiGridViewer.jsp?dbList=http%3A%2F%2Fwww.genboree.org%2FREST%2Fv1%2Fgrp%2

Step 5. Select the samples of interest (in this case, two cells with 8), by clicking on each cell.

Step 6. Click'Selections' -> 'Save Selections'

Step 7. Select your user group

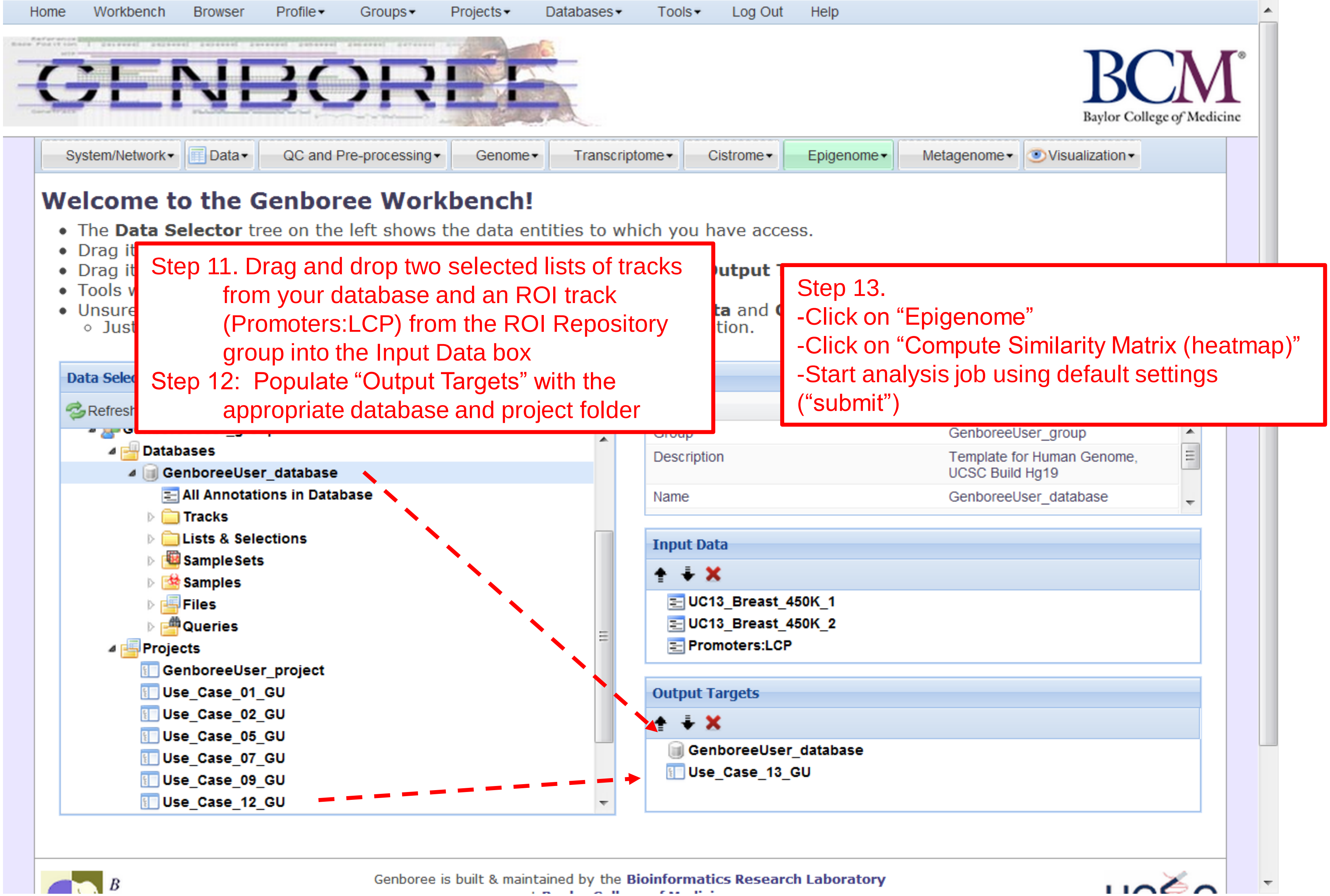
Step 8. Select your database

Step 9. Name this list of tracks

Step 10. Click "Save Selections" to create a list of tracks, called "UC13_Breast_450k_1"

Repeat above process to create another list called "UC13_Breast_450k_2"

Part 1: Tool Launch


Home
Workbench
Browser
Profile
Groups
Projects
Databases
Tools
Log Out
Help
GENBOREE
BCM
Baylor College of Medicine
System/Network
Data
QC and Pre-processing
Genome
Transcriptome
Cistrome
Epigenome
Metagenome
Visualization
Data Selec
Refresh
Databases
GenboreeUser_database
All Annotations in Database
Tracks
Lists & Selections
SampleSets
Samples
Files
Queries
Projects
GenboreeUser_project
Use_Case_01_GU
Use_Case_02_GU
Use_Case_05_GU
Use_Case_07_GU
Use_Case_09_GU
Use_Case_12_GU
Input Data
Output Targets

Welcome to the Genboree Workbench!

- The **Data Selector** tree on the left shows the data entities to which you have access.
- Drag it
- Drag it … utput …
- Tools w
- Unsure … ta and …
  - Just … tion.

Group	GenboreeUser_group
Description	Template for Human Genome, UCSC Build Hg19
Name	GenboreeUser_database

Input Data:
- UC13_Breast_450K_1
- UC13_Breast_450K_2
- Promoters:LCP

Output Targets:
- GenboreeUser_database
- Use_Case_13_GU

Genboree is built & maintained by the **Bioinformatics Research Laboratory**

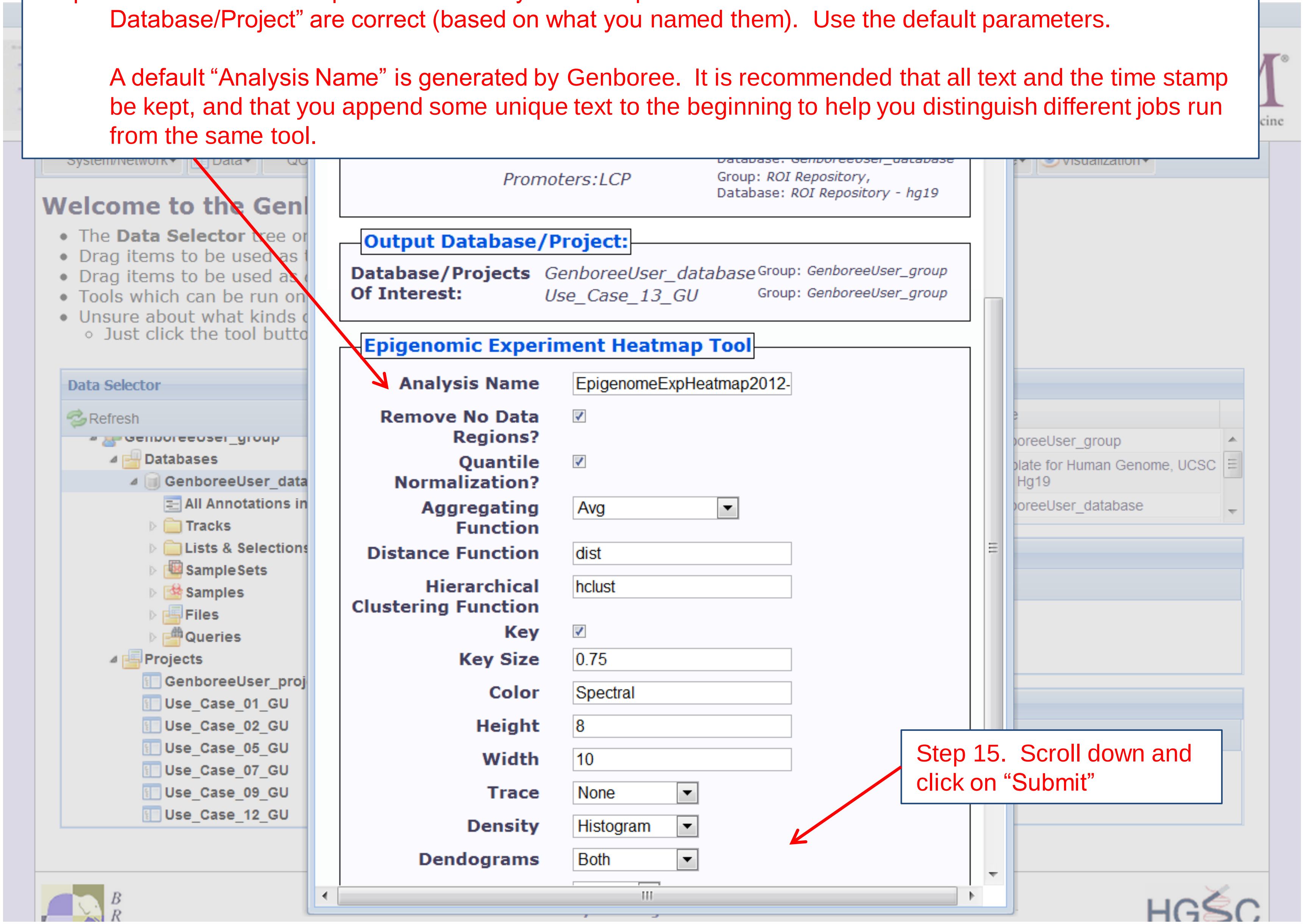
Step 11. Drag and drop two selected lists of tracks from your database and an ROI track (Promoters:LCP) from the ROI Repository group into the Input Data box

Step 12: Populate "Output Targets" with the appropriate database and project folder

Step 13.
-Click on "Epigenome"
-Click on "Compute Similarity Matrix (heatmap)"
-Start analysis job using default settings ("submit")

Step 14. Check that the “Input Files Directory” and “Output Database/Project” are correct (based on what you named them). Use the default parameters.

A default “Analysis Name” is generated by Genboree. It is recommended that all text and the time stamp be kept, and that you append some unique text to the beginning to help you distinguish different jobs run from the same tool.



Promoters:LCP
Group: *ROI Repository*,
Database: *ROI Repository - hg19*

Output Database/Project:

Database/Projects Of Interest:	*GenboreeUser_database*	Group: *GenboreeUser_group*
	Use_Case_13_GU	Group: *GenboreeUser_group*

Epigenomic Experiment Heatmap Tool

Analysis Name	EpigenomeExpHeatmap2012-
Remove No Data Regions?	☑
Quantile Normalization?	☑
Aggregating Function	Avg
Distance Function	dist
Hierarchical Clustering Function	hclust
Key	☑
Key Size	0.75
Color	Spectral
Height	8
Width	10
Trace	None
Density	Histogram
Dendograms	Both

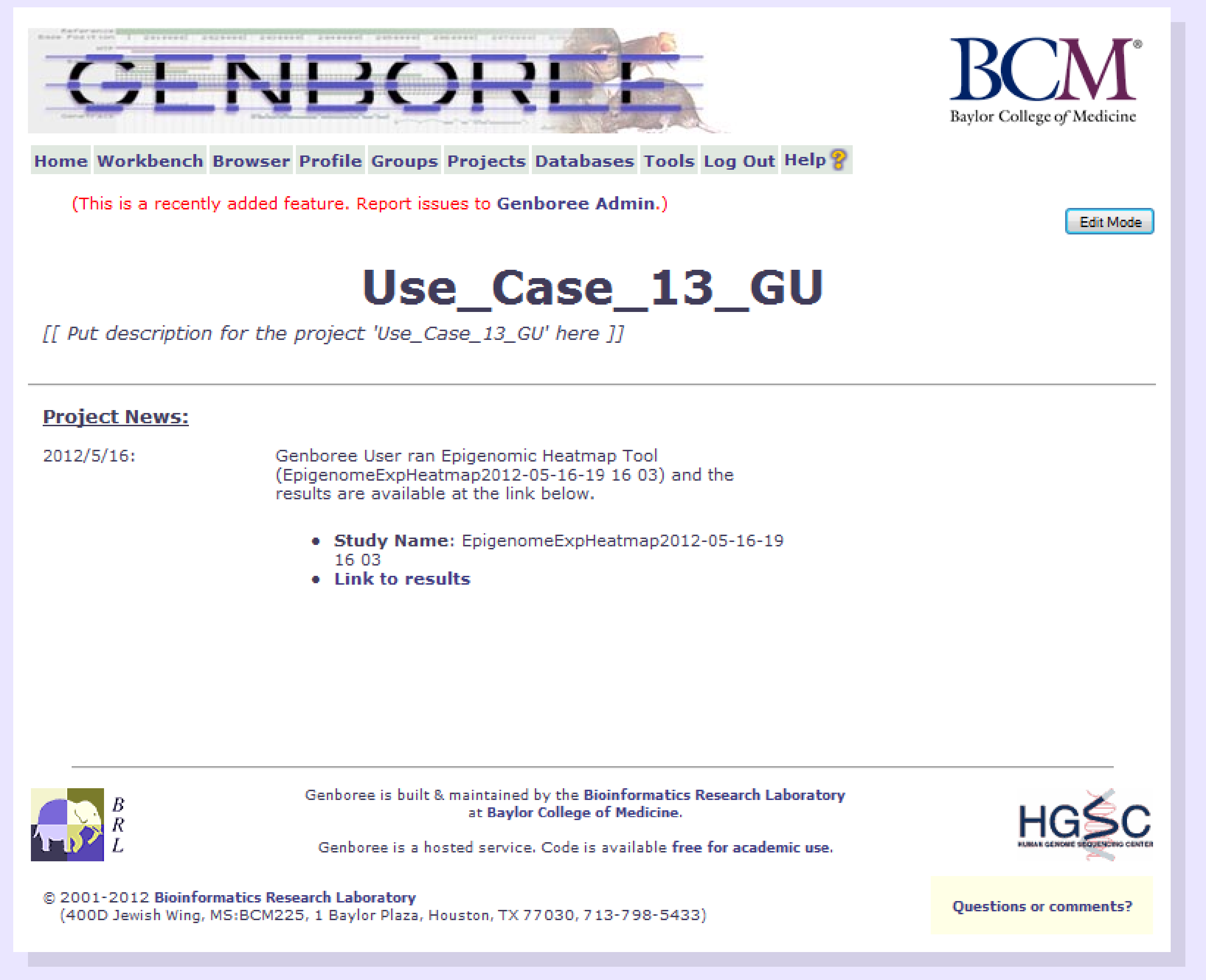
Welcome to the Genl
The Data Selector tree or
Drag items to be used as i
Drag items to be used as
Tools which can be run on
Unsure about what kinds
Just click the tool butto
Data Selector
Refresh
GenboreeUser_group
Databases
GenboreeUser_data
All Annotations in
Tracks
Lists & Selections
SampleSets
Samples
Files
Queries
Projects
GenboreeUser_proj
Use_Case_01_GU
Use_Case_02_GU
Use_Case_05_GU
Use_Case_07_GU
Use_Case_09_GU
Use_Case_12_GU
System/Network
Data
Visualization
boreeUser_group
plate for Human Genome, UCSC Hg19
boreeUser_database

Step 15. Scroll down and click on “Submit”

Part 1: Access You Result

You will receive an email with the following message:
“Genboree: Your Epigenomic Experiment Heatmap Tool job is complete”

The body of the email will contain the following note and a link to the results file: “Result File Location in the Genboree Workbench:
http://www.genboree.org/java-bin/project.jsp?projectName=Use_Case_13


GENBOREE
BCM
Baylor College of Medicine
Home
Workbench
Browser
Profile
Groups
Projects
Databases
Tools
Log Out
Help
Edit Mode

(This is a recently added feature. Report issues to **Genboree Admin**.)

Use_Case_13_GU

[[Put description for the project 'Use_Case_13_GU' here]]

Project News:

2012/5/16: Genboree User ran Epigenomic Heatmap Tool (EpigenomeExpHeatmap2012-05-16-19 16 03) and the results are available at the link below.

- **Study Name:** EpigenomeExpHeatmap2012-05-16-19 16 03
- **Link to results**

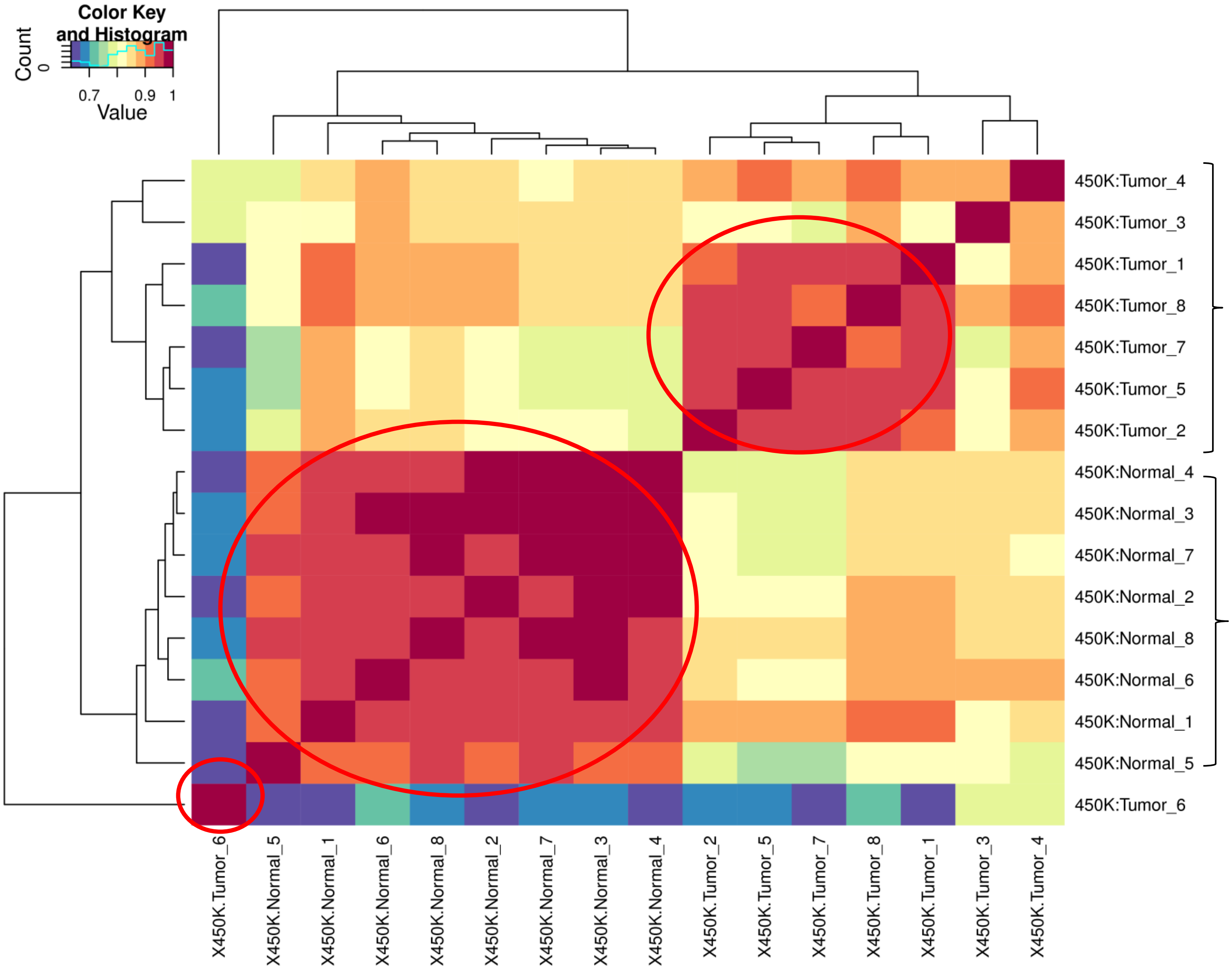
Genboree is built & maintained by the **Bioinformatics Research Laboratory** at **Baylor College of Medicine.**

Genboree is a hosted service. Code is available **free for academic use.**

B
R
L
HGSC
HUMAN GENOME SEQUENCING CENTER
Questions or comments?

© 2001-2012 **Bioinformatics Research Laboratory**
(400D Jewish Wing, MS:BCM225, 1 Baylor Plaza, Houston, TX 77030, 713-798-5433)

Part 1: Heatmap Output (Promoters:LCP)

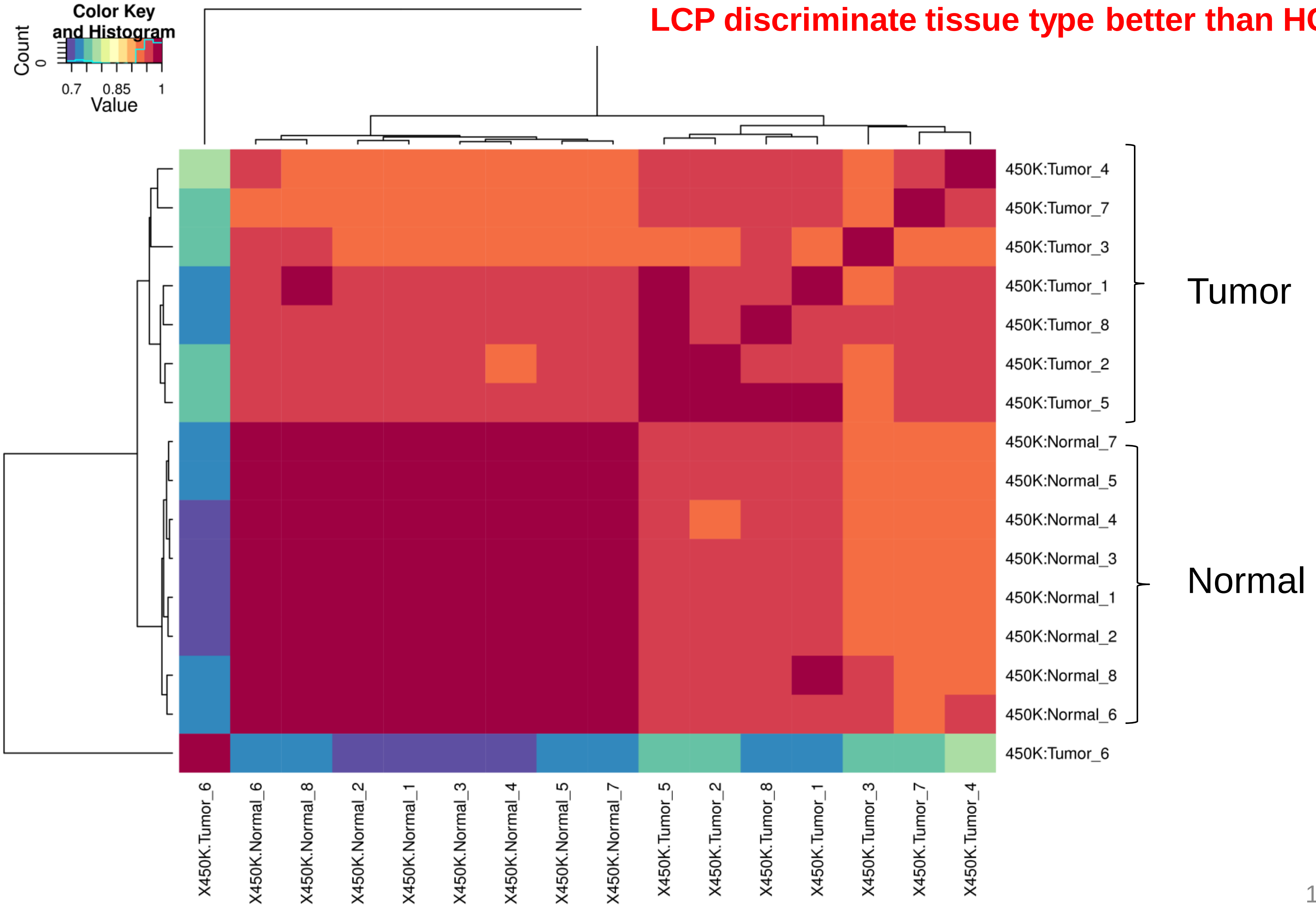


Tumor

Normal

Part 1: Heatmap Output (Promoters:HCP)

LCP discriminate tissue type better than HCP



Part 2: Compare 450K profiles (8 tumor, 8 normal) against reference epigenomes from the Epigenome Atlas

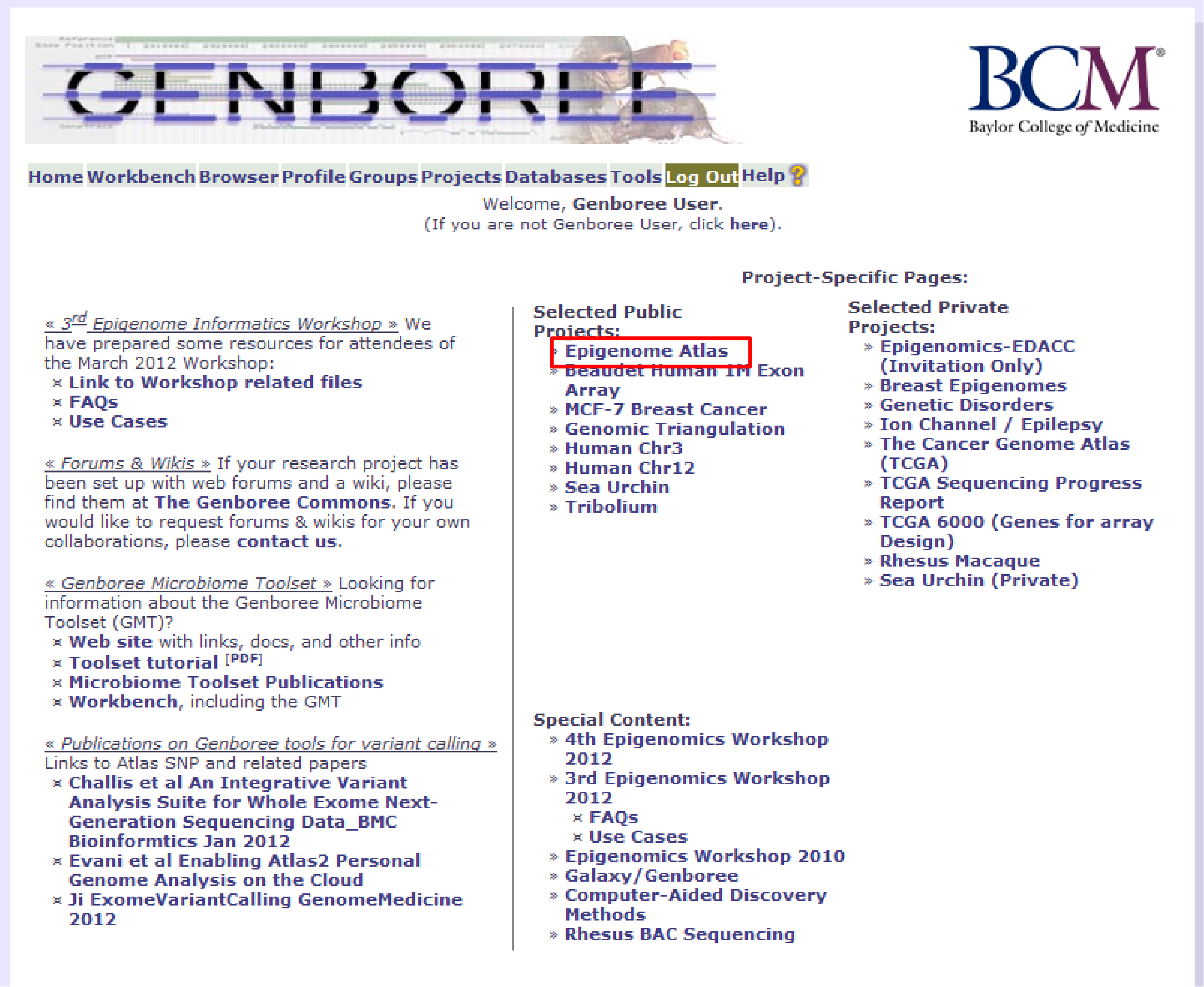
Part 2: Introduction

- In Part 1, we found that tumor and normal samples are different. Now we ask if the difference is due to tissue composition by comparing the samples against the Human Epigenome Atlas.
- Plan: Compare 8 tumor and 8 normal 450K profiles against normal tissue profiles from the Human Epigenome Atlas using the Epigenomic Heatmap tool
- Note: We compare epigenomic profiles obtained using different technologies (Illumina 450K vs MeDIP-seq).

Reference MeDIP profiles of Normal Tissues and Cell Lines from the Atlas

1. Breast Luminal Epithelial Cells
2. Breast Myoepithelial Cells
3. Breast Stem Cells
4. Peripheral blood mononuclear primary cells
5. Fetal Brain
6. H1 Cell Line
7. CD4 Memory Primary Cells
8. CD4 Naive Primary Cells

Access the “Epigenome Atlas” from the Genboree homepage


GENBOREE
BCM Baylor College of Medicine
Home Workbench Browser Profile Groups Projects Databases Tools Log Out Help

Welcome, **Genboree User**.
(If you are not Genboree User, click **here**).

Project-Specific Pages:

« 3rd Epigenome Informatics Workshop » We have prepared some resources for attendees of the March 2012 Workshop:

- **Link to Workshop related files**
- **FAQs**
- **Use Cases**

« Forums & Wikis » If your research project has been set up with web forums and a wiki, please find them at **The Genboree Commons**. If you would like to request forums & wikis for your own collaborations, please **contact us**.

« Genboree Microbiome Toolset » Looking for information about the Genboree Microbiome Toolset (GMT)?

- **Web site** with links, docs, and other info
- **Toolset tutorial** [PDF]
- **Microbiome Toolset Publications**
- **Workbench**, including the GMT

« Publications on Genboree tools for variant calling » Links to Atlas SNP and related papers

- **Challis et al An Integrative Variant Analysis Suite for Whole Exome Next-Generation Sequencing Data_BMC Bioinformtics Jan 2012**
- **Evani et al Enabling Atlas2 Personal Genome Analysis on the Cloud**
- **Ji ExomeVariantCalling GenomeMedicine 2012**

Selected Public Projects:

- **Epigenome Atlas**
- **Beaudet Human 1M Exon Array**
- **MCF-7 Breast Cancer**
- **Genomic Triangulation**
- **Human Chr3**
- **Human Chr12**
- **Sea Urchin**
- **Tribolium**

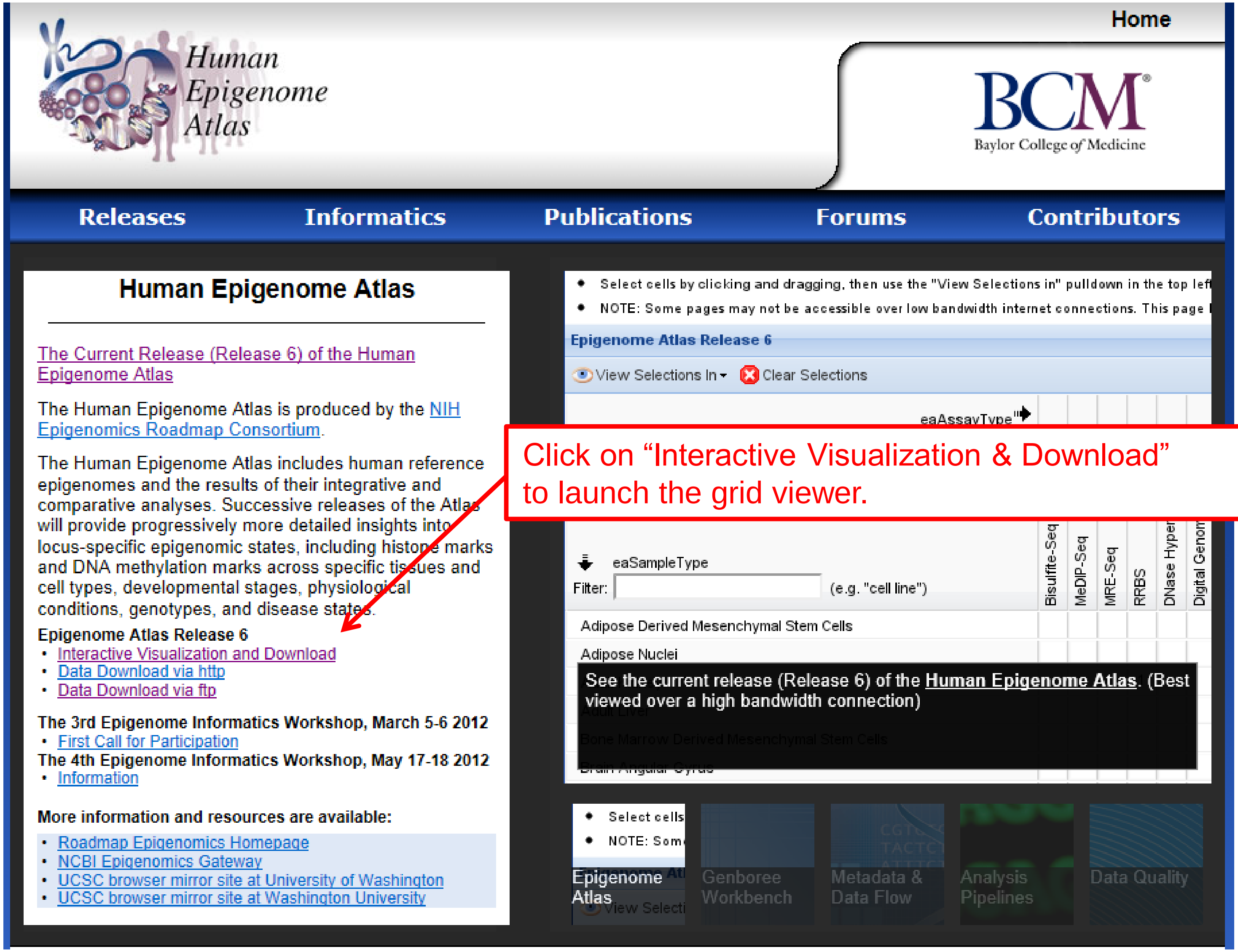
Special Content:

- **4th Epigenomics Workshop 2012**
- **3rd Epigenomics Workshop 2012**
  - **FAQs**
  - **Use Cases**
- **Epigenomics Workshop 2010**
- **Galaxy/Genboree**
- **Computer-Aided Discovery Methods**
- **Rhesus BAC Sequencing**

Selected Private Projects:

- **Epigenomics-EDACC (Invitation Only)**
- **Breast Epigenomes**
- **Genetic Disorders**
- **Ion Channel / Epilepsy**
- **The Cancer Genome Atlas (TCGA)**
- **TCGA Sequencing Progress Report**
- **TCGA 6000 (Genes for array Design)**
- **Rhesus Macaque**
- **Sea Urchin (Private)**

Access the Human Epigenome Atlas Data


Home
Human Epigenome Atlas
BCM
Baylor College of Medicine
Releases
Informatics
Publications
Forums
Contributors

Human Epigenome Atlas

The Current Release (Release 6) of the Human Epigenome Atlas

The Human Epigenome Atlas is produced by the NIH Epigenomics Roadmap Consortium.

The Human Epigenome Atlas includes human reference epigenomes and the results of their integrative and comparative analyses. Successive releases of the Atlas will provide progressively more detailed insights into locus-specific epigenomic states, including histone marks and DNA methylation marks across specific tissues and cell types, developmental stages, physiological conditions, genotypes, and disease states.

Epigenome Atlas Release 6
- Interactive Visualization and Download
- Data Download via http
- Data Download via ftp

The 3rd Epigenome Informatics Workshop, March 5-6 2012
- First Call for Participation

The 4th Epigenome Informatics Workshop, May 17-18 2012
- Information

More information and resources are available:
- Roadmap Epigenomics Homepage
- NCBI Epigenomics Gateway
- UCSC browser mirror site at University of Washington
- UCSC browser mirror site at Washington University

- Select cells by clicking and dragging, then use the "View Selections in" pulldown in the top lef
- NOTE: Some pages may not be accessible over low bandwidth internet connections. This page

Epigenome Atlas Release 6

View Selections In ▾
Clear Selections
eaAssayType
eaSampleType
Bisulfite-Seq
MeDIP-Seq
MRE-Seq
RRBS
DNase Hyper
Digital Genom

Filter: (e.g. "cell line")

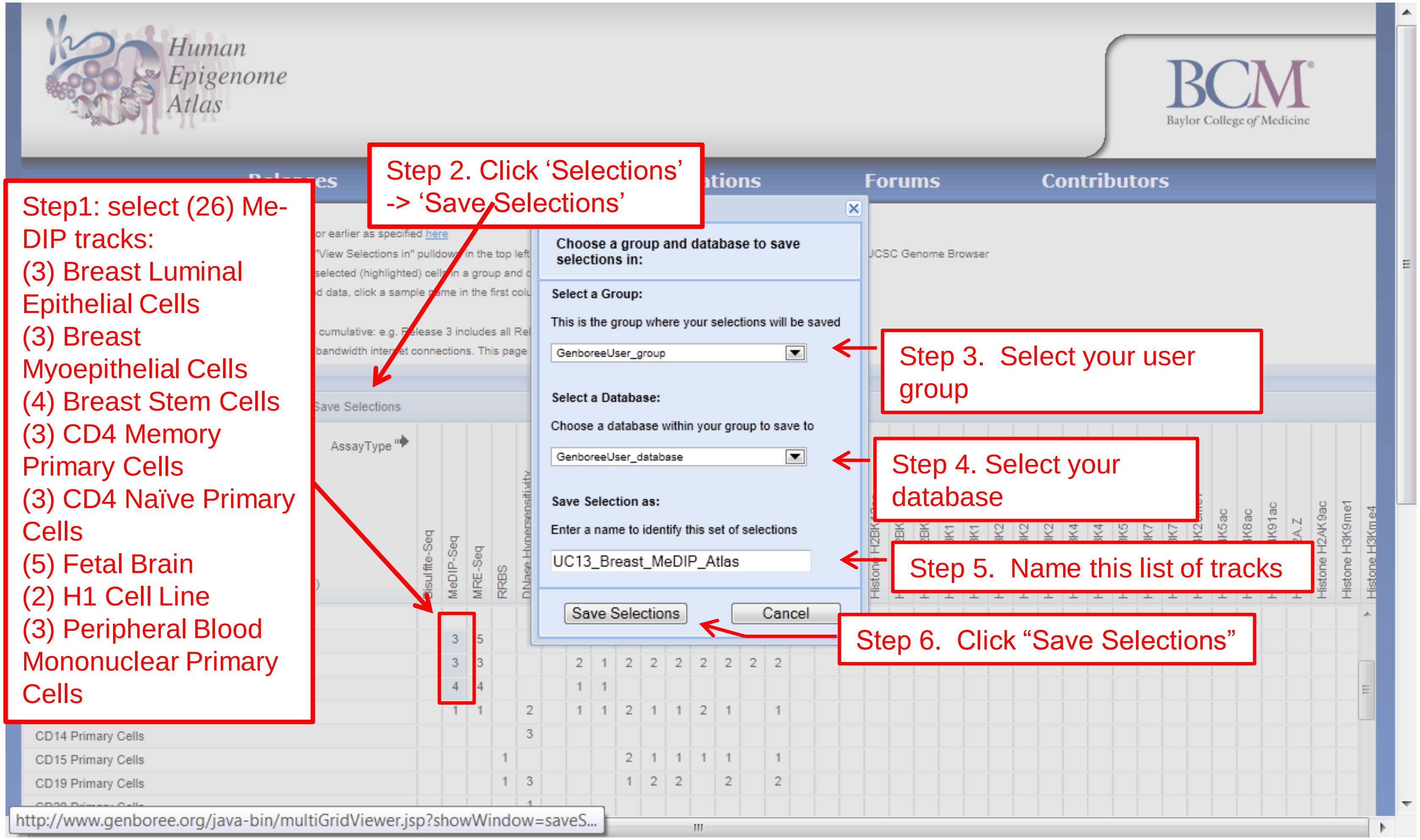
Adipose Derived Mesenchymal Stem Cells

Adipose Nuclei

See the current release (Release 6) of the Human Epigenome Atlas. (Best viewed over a high bandwidth connection)

Epigenome Atlas
Genboree Workbench
Metadata & Data Flow
Analysis Pipelines
Data Quality

Click on “Interactive Visualization & Download” to launch the grid viewer.


Human Epigenome Atlas
BCM
Baylor College of Medicine
Forums
Contributors
Save Selections
Cancel
AssayType
MeDIP-Seq
MRE-Seq
RRBS
CD14 Primary Cells
CD15 Primary Cells
CD19 Primary Cells

Step1: select (26) Me-DIP tracks:
(3) Breast Luminal Epithelial Cells
(3) Breast Myoepithelial Cells
(4) Breast Stem Cells
(3) CD4 Memory Primary Cells
(3) CD4 Naïve Primary Cells
(5) Fetal Brain
(2) H1 Cell Line
(3) Peripheral Blood Mononuclear Primary Cells

Step 2. Click ‘Selections’ -> ‘Save Selections’

Step 3. Select your user group

Step 4. Select your database

Step 5. Name this list of tracks

Step 6. Click “Save Selections”

Choose a group and database to save selections in:

Select a Group:

This is the group where your selections will be saved

GenboreeUser_group

Select a Database:

Choose a database within your group to save to

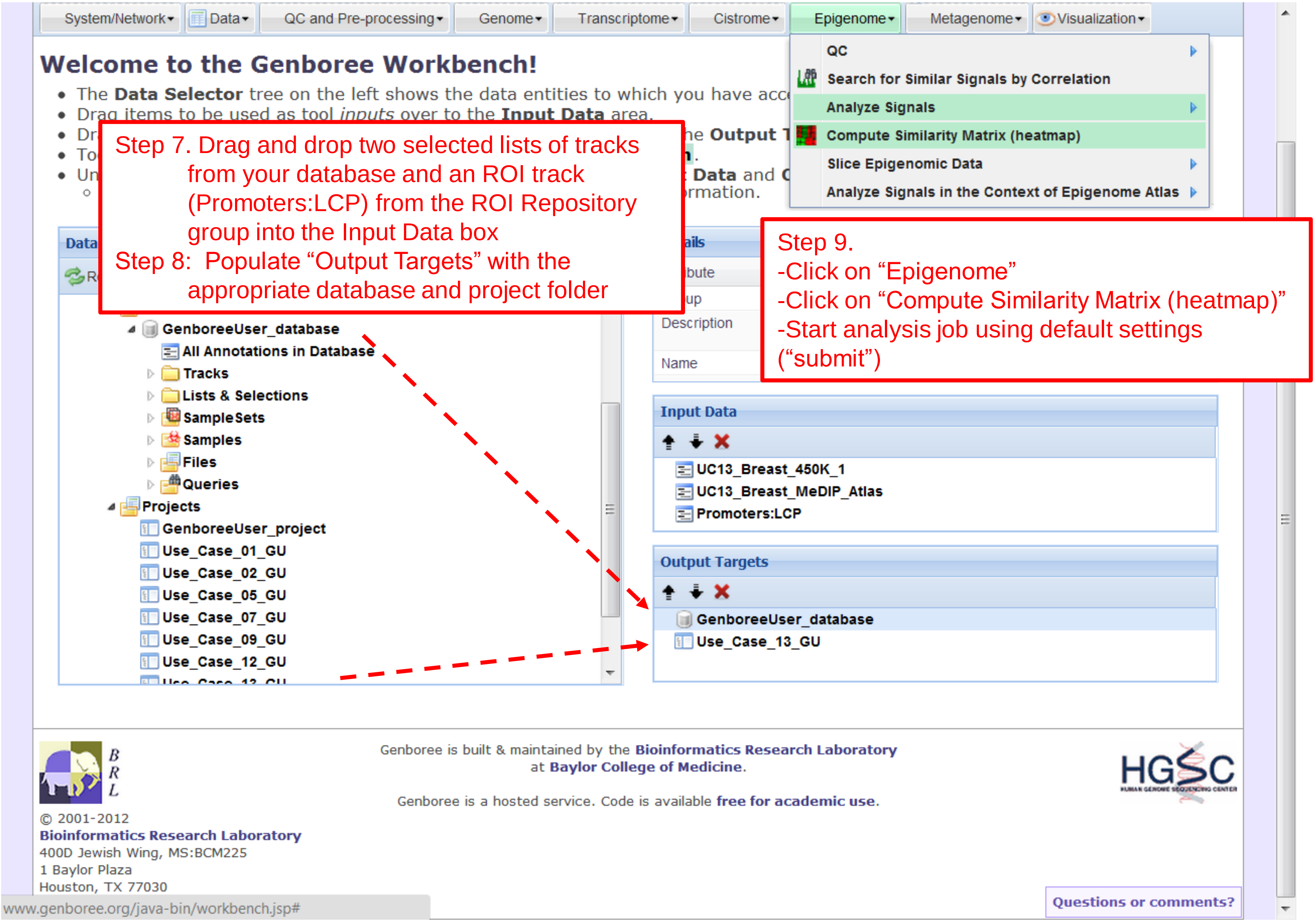
GenboreeUser_database

Save Selection as:

Enter a name to identify this set of selections

UC13_Breast_MeDIP_Atlas

http://www.genboree.org/java-bin/multiGridViewer.jsp?showWindow=saveS...



Step 7. Drag and drop two selected lists of tracks from your database and an ROI track (Promoters:LCP) from the ROI Repository group into the Input Data box

Step 8: Populate “Output Targets” with the appropriate database and project folder

Step 9.
-Click on “Epigenome”
-Click on “Compute Similarity Matrix (heatmap)”
-Start analysis job using default settings (“submit”)

Welcome to the Genboree Workbench!

- The **Data Selector** tree on the left shows the data entities to which you have acc
- Drag items to be used as tool *inputs* over to the **Input Data** area.
- Dr
- To
- Un

Genboree is built & maintained by the **Bioinformatics Research Laboratory** at **Baylor College of Medicine**.

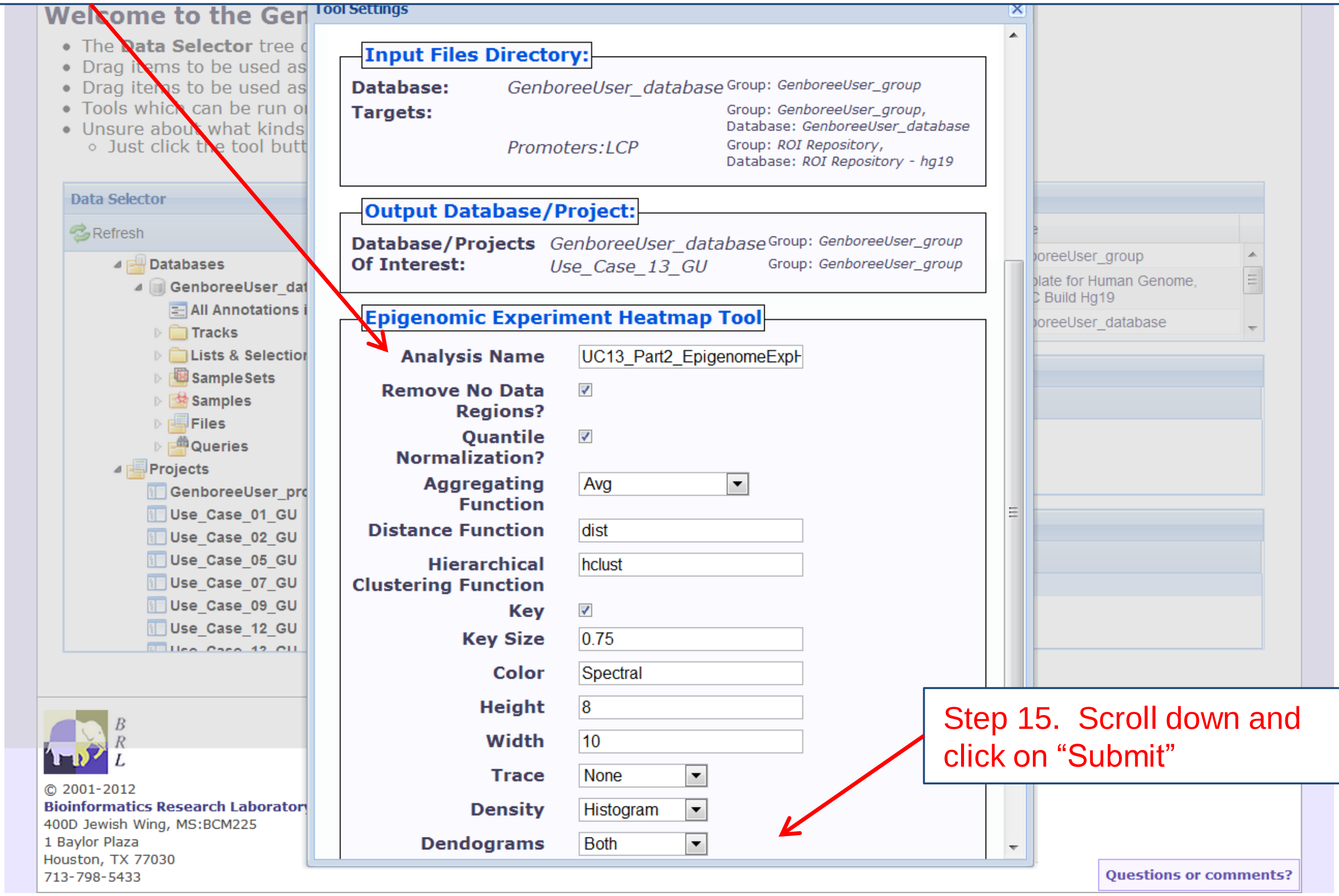
Genboree is a hosted service. Code is available **free for academic use**.

© 2001-2012
Bioinformatics Research Laboratory
400D Jewish Wing, MS:BCM225
1 Baylor Plaza
Houston, TX 77030

System/Network
Data
QC and Pre-processing
Genome
Transcriptome
Cistrome
Epigenome
Metagenome
Visualization
QC
Search for Similar Signals by Correlation
Analyze Signals
Compute Similarity Matrix (heatmap)
Slice Epigenomic Data
Analyze Signals in the Context of Epigenome Atlas
he Output T
Data and C
rmation.
Data
R
ails
ibute
up
Description
Name
GenboreeUser_database
All Annotations in Database
Tracks
Lists & Selections
SampleSets
Samples
Files
Queries
Projects
GenboreeUser_project
Use_Case_01_GU
Use_Case_02_GU
Use_Case_05_GU
Use_Case_07_GU
Use_Case_09_GU
Use_Case_12_GU
Input Data
UC13_Breast_450K_1
UC13_Breast_MeDIP_Atlas
Promoters:LCP
Output Targets
GenboreeUser_database
Use_Case_13_GU
BRL
HGSC
HUMAN GENOME SEQUENCING CENTER
Questions or comments?
www.genboree.org/java-bin/workbench.jsp#

Step 14. Check that the “Input Files Directory” and “Output Database/Project” are correct (based on what you named them). Use the default parameters.

A default “Analysis Name” is generated by Genboree. It is recommended that all text and the time stamp be kept, and that you append some unique text to the beginning to help you distinguish different jobs run from the same tool.



Input Files Directory:

Database: *GenboreeUser_database* Group: *GenboreeUser_group*

Targets: Group: *GenboreeUser_group*, Database: *GenboreeUser_database*

Promoters:LCP Group: *ROI Repository*, Database: *ROI Repository - hg19*

Output Database/Project:

Database/Projects Of Interest: *GenboreeUser_database* Group: *GenboreeUser_group*

Use_Case_13_GU Group: *GenboreeUser_group*

Epigenomic Experiment Heatmap Tool

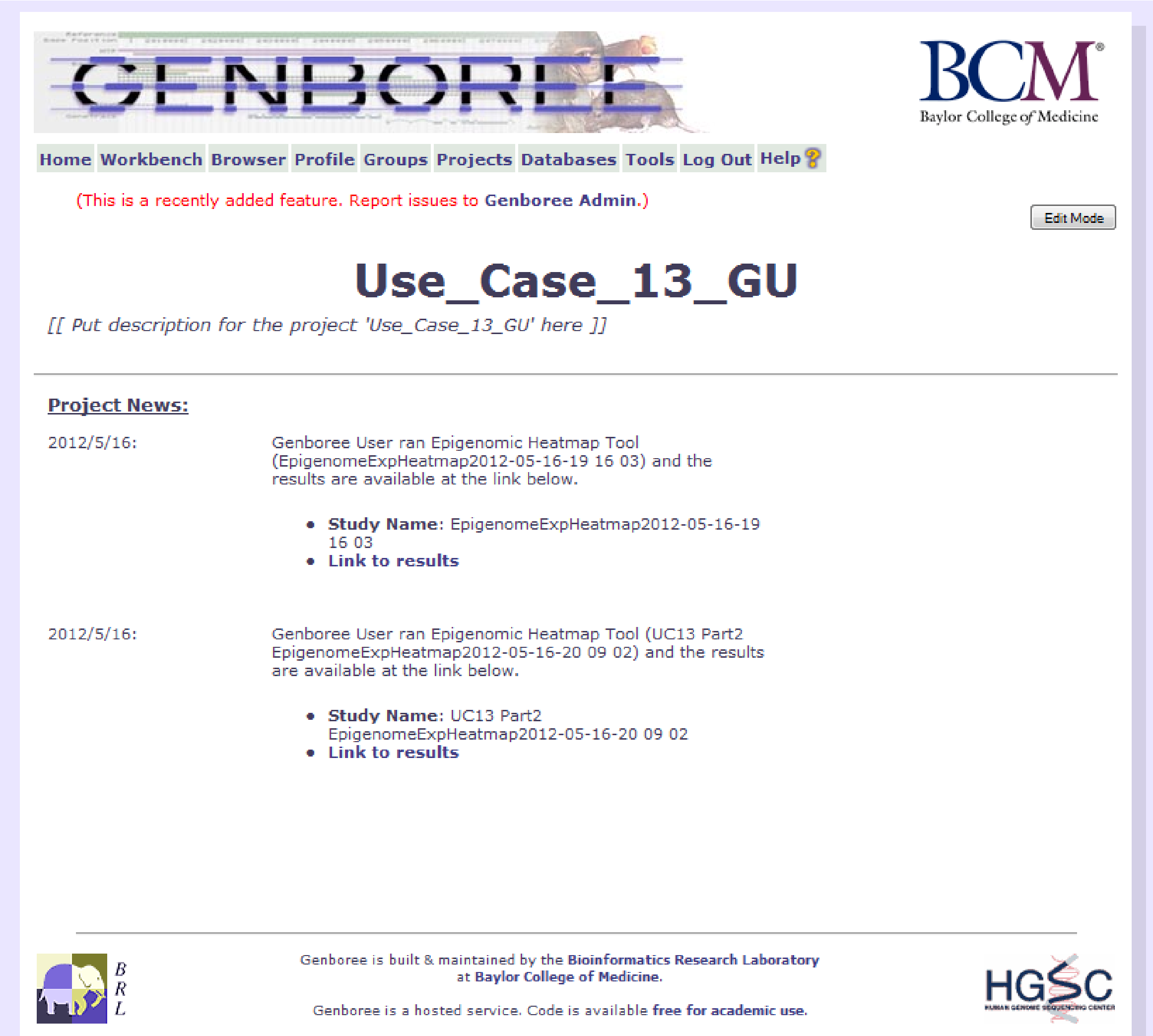
Analysis Name	UC13_Part2_EpigenomeExpH
Remove No Data Regions?	☑
Quantile Normalization?	☑
Aggregating Function	Avg
Distance Function	dist
Hierarchical Clustering Function	hclust
Key	☑
Key Size	0.75
Color	Spectral
Height	8
Width	10
Trace	None
Density	Histogram
Dendograms	Both

© 2001-2012
Bioinformatics Research Laboratory
400D Jewish Wing, MS:BCM225
1 Baylor Plaza
Houston, TX 77030
713-798-5433

Step 15. Scroll down and click on “Submit”

You will receive an email with the following message:
“Genboree: Your Epigenomic Experiment Heatmap Tool job is complete”

The body of the email will contain the following note and a link to the results file: “Result File Location in the Genboree Workbench:
http://www.genboree.org/java-bin/project.jsp?projectName=Use_Case_13


GENBOREE
BCM
Baylor College of Medicine
Home Workbench Browser Profile Groups Projects Databases Tools Log Out Help
Edit Mode

(This is a recently added feature. Report issues to **Genboree Admin**.)

Use_Case_13_GU

[[Put description for the project 'Use_Case_13_GU' here]]

Project News:

2012/5/16:

Genboree User ran Epigenomic Heatmap Tool (EpigenomeExpHeatmap2012-05-16-19 16 03) and the results are available at the link below.

- **Study Name:** EpigenomeExpHeatmap2012-05-16-19 16 03
- **Link to results**

2012/5/16:

Genboree User ran Epigenomic Heatmap Tool (UC13 Part2 EpigenomeExpHeatmap2012-05-16-20 09 02) and the results are available at the link below.

- **Study Name:** UC13 Part2 EpigenomeExpHeatmap2012-05-16-20 09 02
- **Link to results**

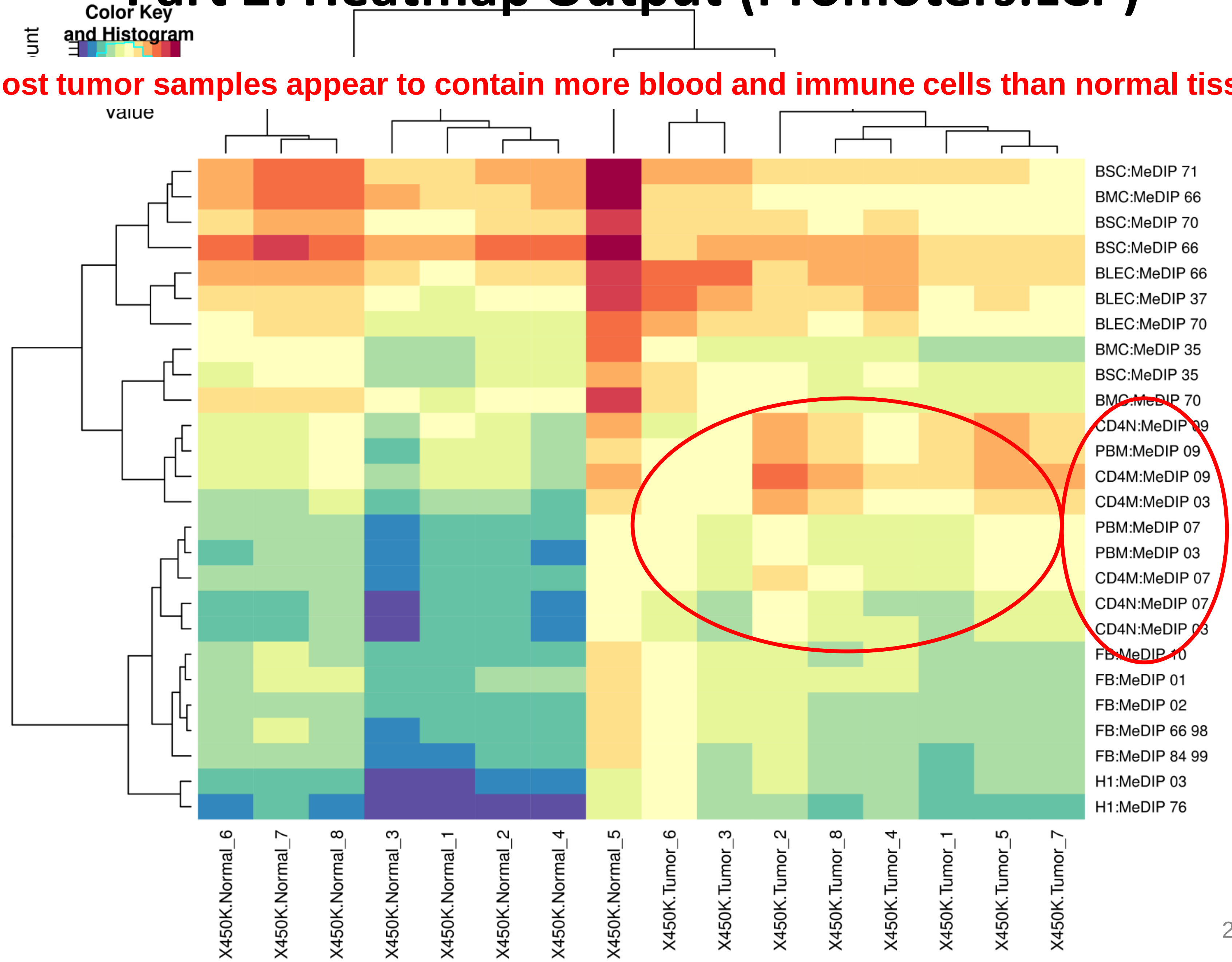
Genboree is built & maintained by the **Bioinformatics Research Laboratory** at **Baylor College of Medicine.**

Genboree is a hosted service. Code is available **free for academic use.**

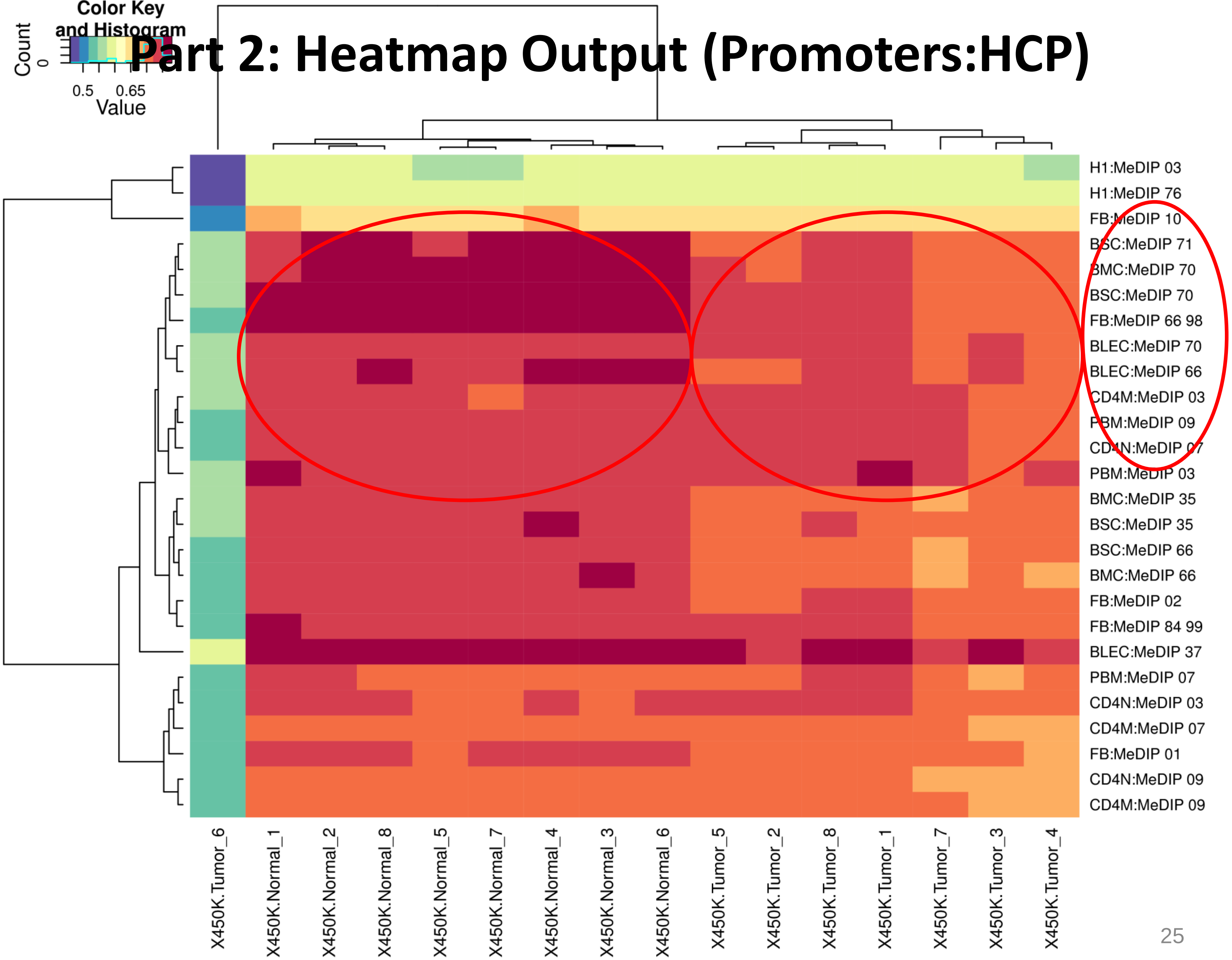
BRL
HGSC
HUMAN GENOME SEQUENCING CENTER

Part 2: Heatmap Output (Promoters:LCP)

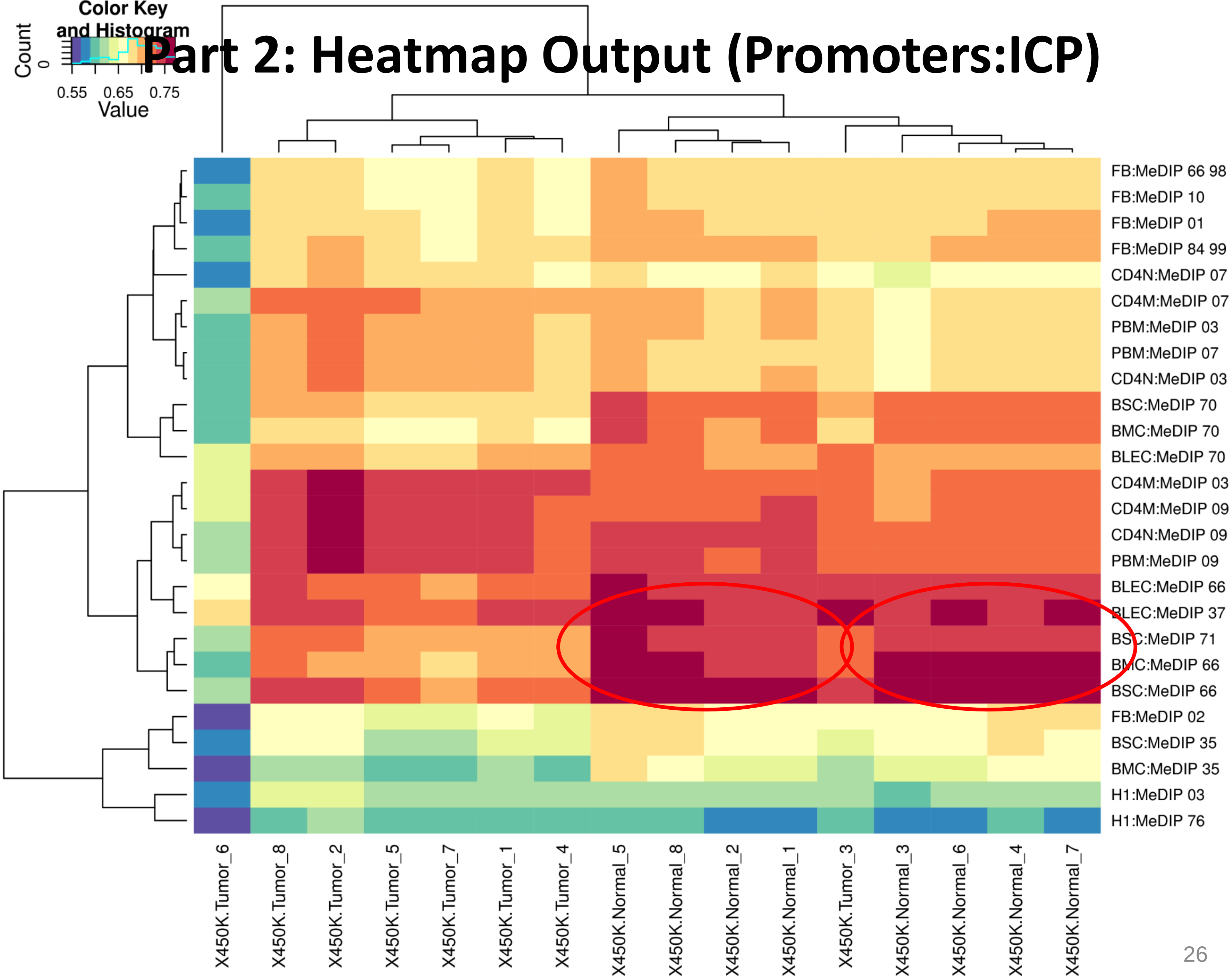
Most tumor samples appear to contain more blood and immune cells than normal tissue.



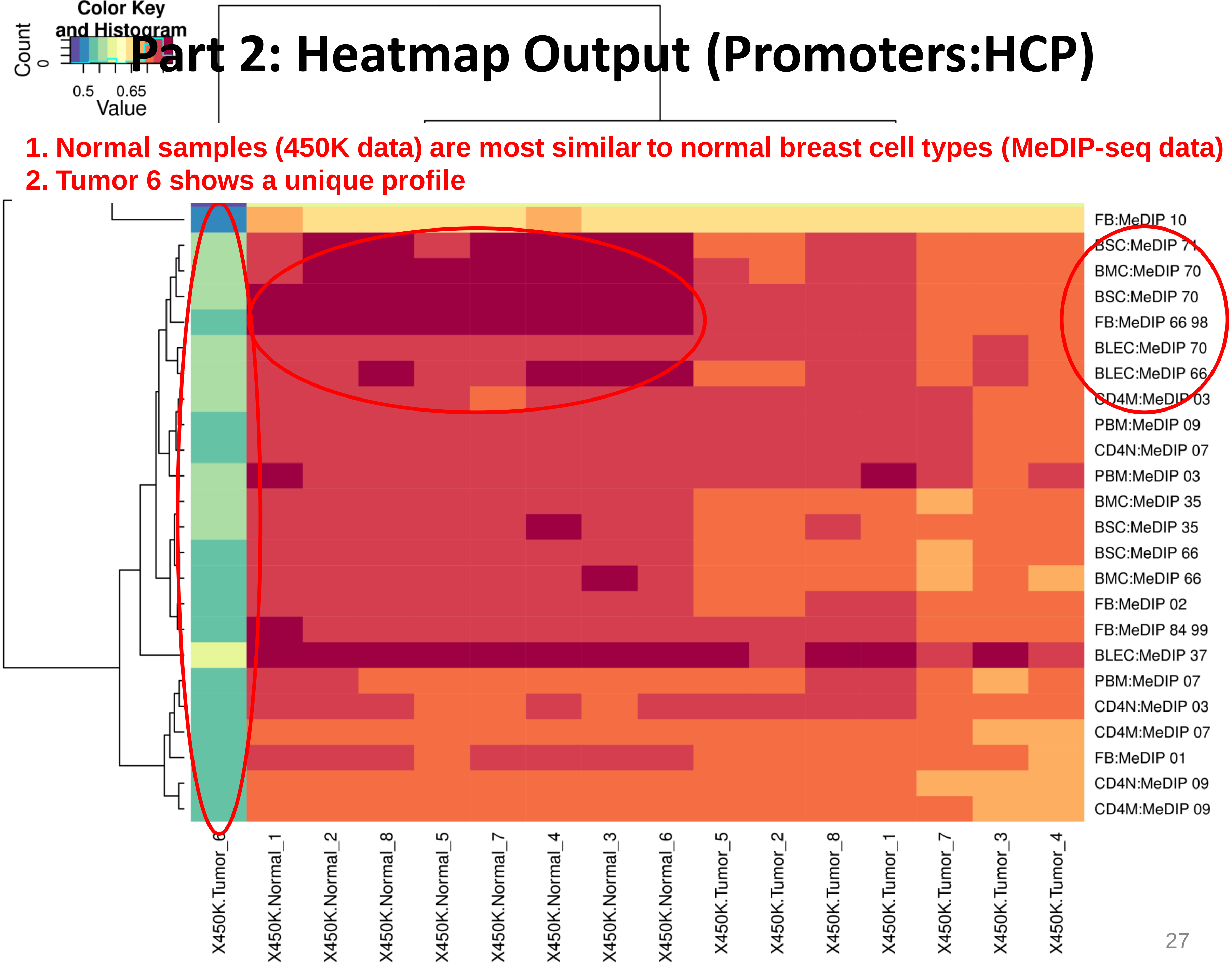
Part 2: Heatmap Output (Promoters:HCP)



Part 2: Heatmap Output (Promoters:ICP)



Part 2: Heatmap Output (Promoters:HCP)

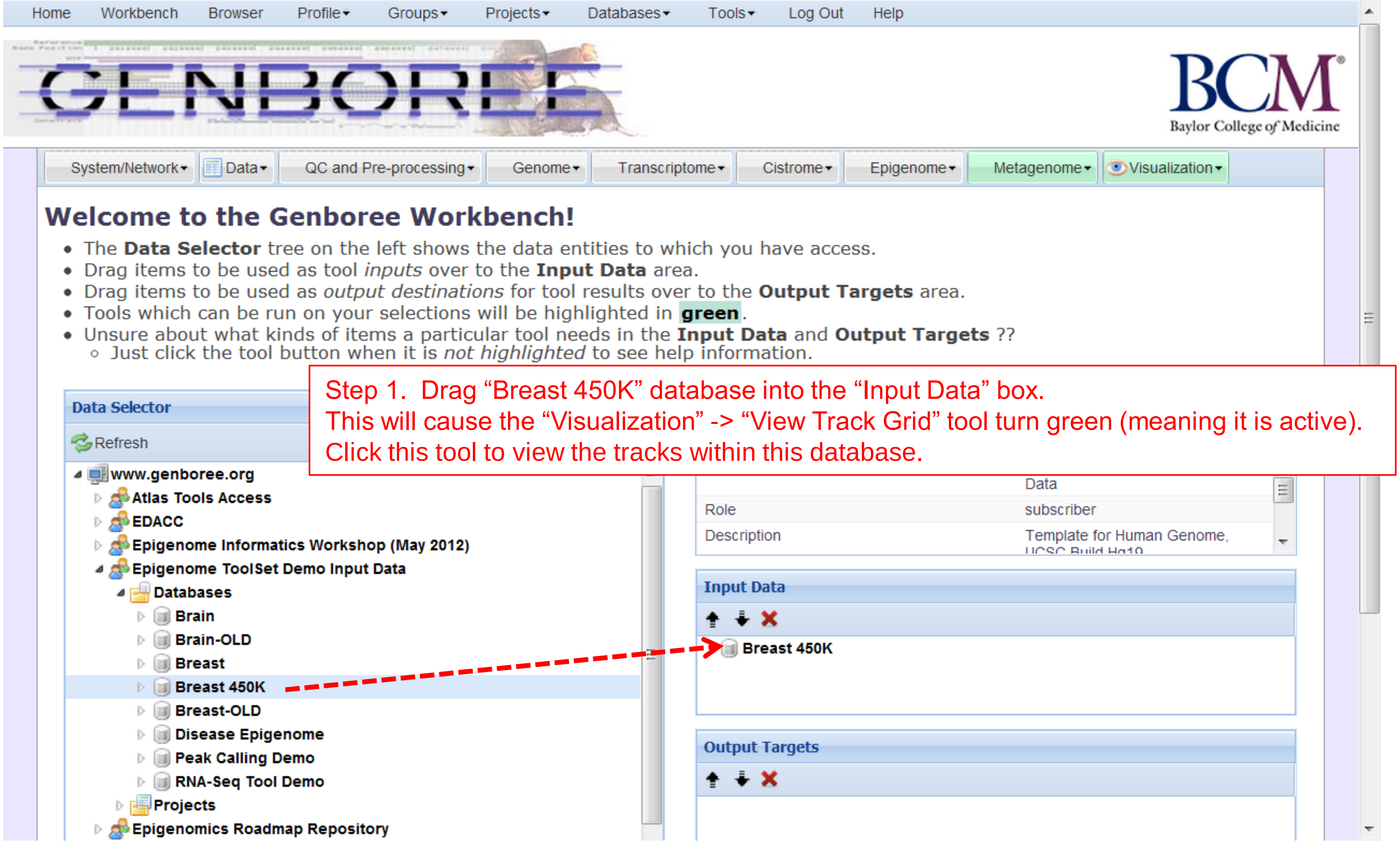


1. Normal samples (450K data) are most similar to normal breast cell types (MeDIP-seq data)
2. Tumor 6 shows a unique profile

Part 3: LIMMA_Comparison: Breast Normal vs Breast Tumor

Part 3: Introduction

- In Part 2, we found that **most breast tumor samples appear to contain an excess of blood and immune cells.**
- Hypothesis: Comparison of normal and tumor tissue should reveal **differentially methylated genes that are involved in immune related pathways or biological processes.**
- Plan: Identify differentially methylated probes, genes, and pathways by comparing normal and tumor tissues using the LIMMA comparison tool.


Home
Workbench
Browser
Profile
Groups
Projects
Databases
Tools
Log Out
Help
GENBOREE
BCM
Baylor College of Medicine
System/Network
Data
QC and Pre-processing
Genome
Transcriptome
Cistrome
Epigenome
Metagenome
Visualization

Step 1. Drag “Breast 450K” database into the “Input Data” box.
This will cause the “Visualization” -> “View Track Grid” tool turn green (meaning it is active).
Click this tool to view the tracks within this database.

Welcome to the Genboree Workbench!

- The **Data Selector** tree on the left shows the data entities to which you have access.
- Drag items to be used as tool *inputs* over to the **Input Data** area.
- Drag items to be used as *output destinations* for tool results over to the **Output Targets** area.
- Tools which can be run on your selections will be highlighted in **green**.
- Unsure about what kinds of items a particular tool needs in the **Input Data** and **Output Targets** ??
    - Just click the tool button when it is *not highlighted* to see help information.

Data Selector

Refresh

- **www.genboree.org**
    - **Atlas Tools Access**
    - **EDACC**
    - **Epigenome Informatics Workshop (May 2012)**
    - **Epigenome ToolSet Demo Input Data**
        - **Databases**
            - **Brain**
            - **Brain-OLD**
            - **Breast**
            - **Breast 450K**
            - **Breast-OLD**
            - **Disease Epigenome**
            - **Peak Calling Demo**
            - **RNA-Seq Tool Demo**
        - **Projects**
    - **Epigenomics Roadmap Repository**

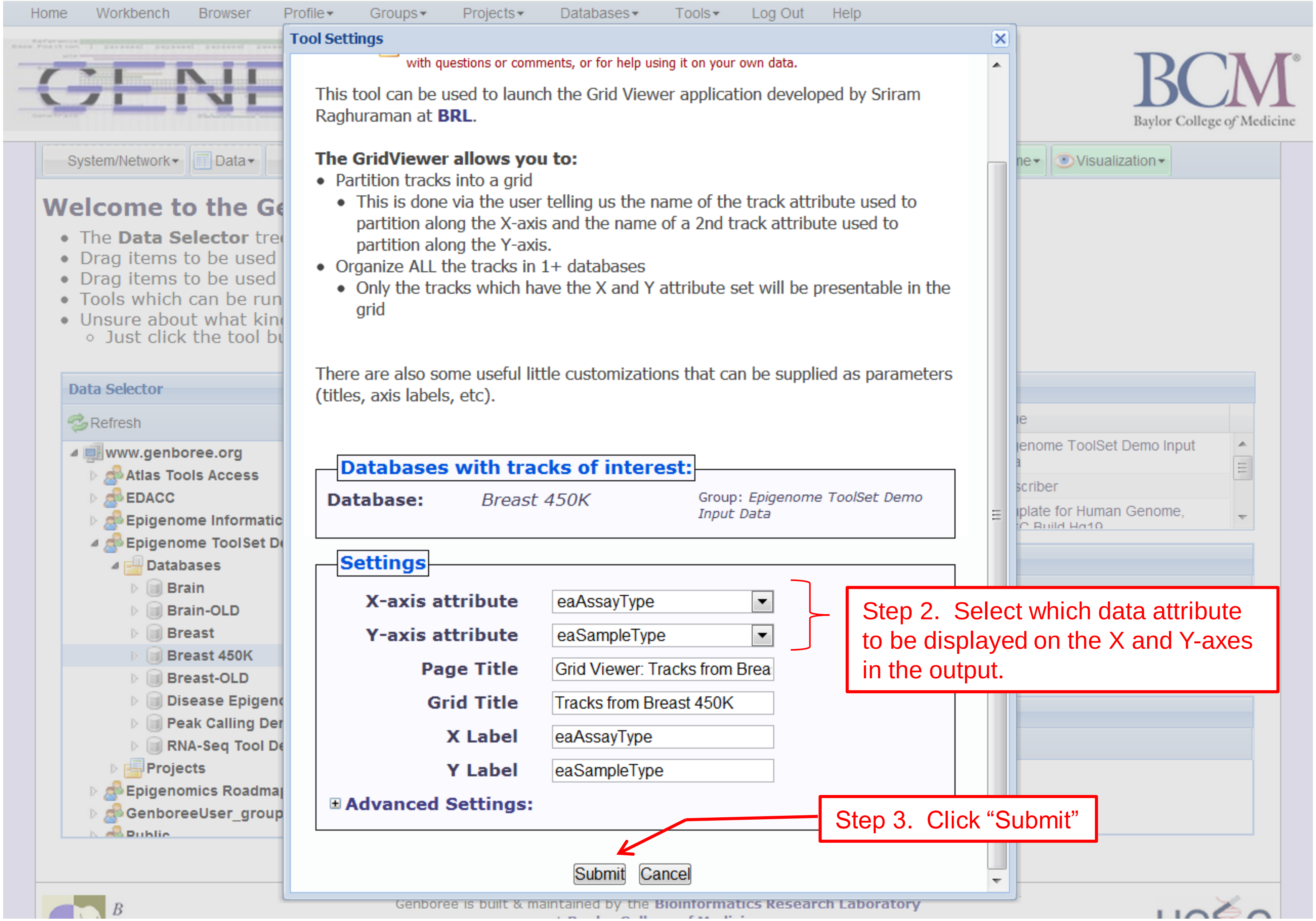
	Data
Role	subscriber
Description	Template for Human Genome, UCSC Build Hg19

Input Data

Breast 450K

Output Targets

Part 1: Create Your Data Grid-Con’d


Home
Workbench
Browser
Profile
Groups
Projects
Databases
Tools
Log Out
Help
System/Network
Data
Visualization
BCM
Baylor College of Medicine
Data Selector
Refresh
www.genboree.org
Atlas Tools Access
EDACC
Breast 450K
Submit
Cancel

Tool Settings

with questions or comments, or for help using it on your own data.

This tool can be used to launch the Grid Viewer application developed by Sriram Raghuraman at **BRL**.

The GridViewer allows you to:

- Partition tracks into a grid
  - This is done via the user telling us the name of the track attribute used to partition along the X-axis and the name of a 2nd track attribute used to partition along the Y-axis.
- Organize ALL the tracks in 1+ databases
  - Only the tracks which have the X and Y attribute set will be presentable in the grid

There are also some useful little customizations that can be supplied as parameters (titles, axis labels, etc).

Databases with tracks of interest:

Database: *Breast 450K* Group: *Epigenome ToolSet Demo Input Data*

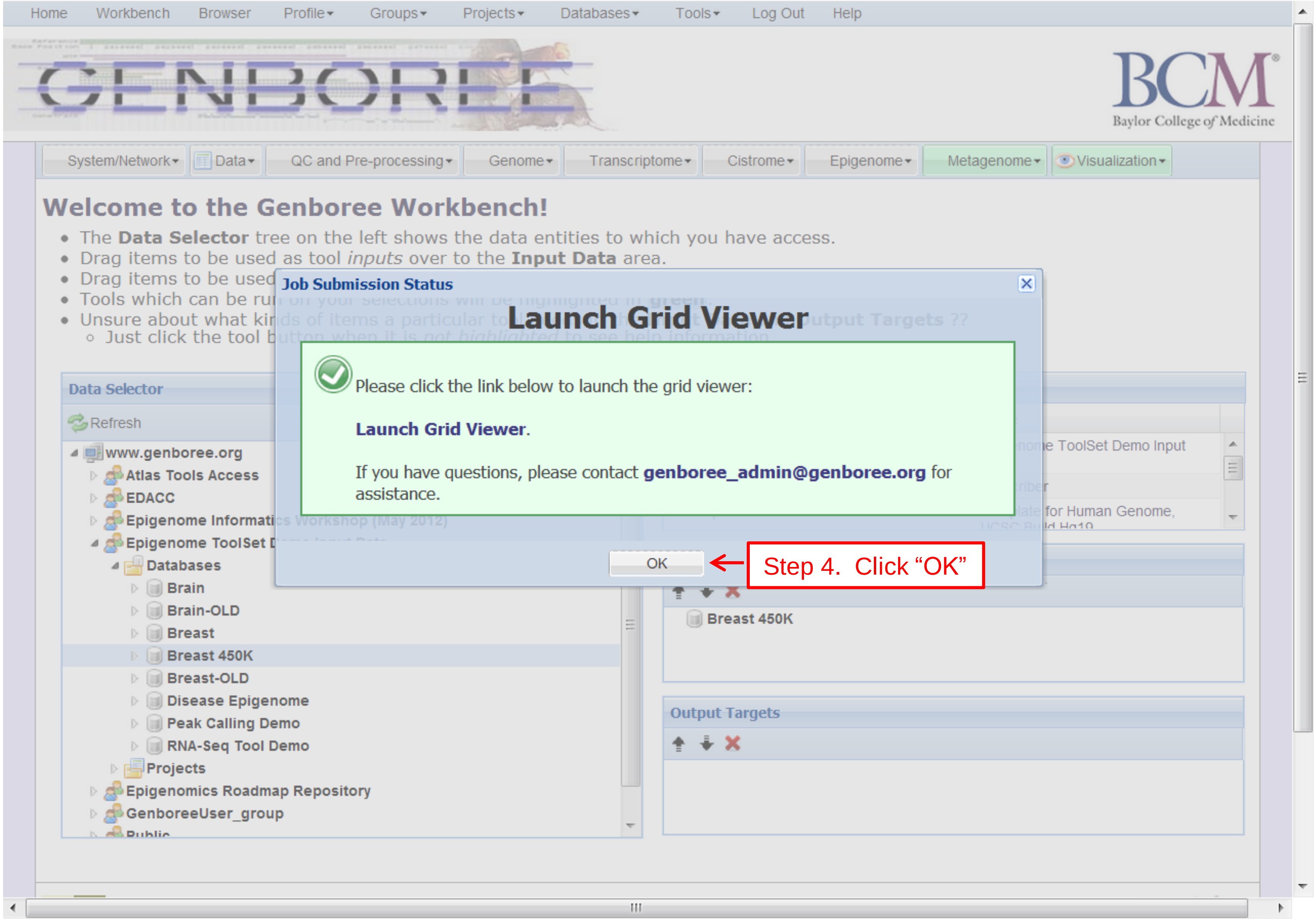
Settings

X-axis attribute	eaAssayType
Y-axis attribute	eaSampleType
Page Title	Grid Viewer: Tracks from Brea
Grid Title	Tracks from Breast 450K
X Label	eaAssayType
Y Label	eaSampleType

⊞ **Advanced Settings:**

Step 2. Select which data attribute to be displayed on the X and Y-axes in the output.

Step 3. Click “Submit”


Home
Workbench
Browser
Profile
Groups
Projects
Databases
Tools
Log Out
Help
GENBOREE
BCM
Baylor College of Medicine
System/Network
Data
QC and Pre-processing
Genome
Transcriptome
Cistrome
Epigenome
Metagenome
Visualization
Data Selector
Refresh
www.genboree.org
Atlas Tools Access
EDACC
Epigenome Informatics Workshop (May 2012)
Epigenome ToolSet D
Databases
Brain
Brain-OLD
Breast
Breast 450K
Breast-OLD
Disease Epigenome
Peak Calling Demo
RNA-Seq Tool Demo
Projects
Epigenomics Roadmap Repository
GenbooreeUser_group
Public
Breast 450K
Output Targets
OK

Welcome to the Genboree Workbench!

- The **Data Selector** tree on the left shows the data entities to which you have access.
- Drag items to be used as tool *inputs* over to the **Input Data** area.
- Drag items to be used
- Tools which can be run on your selections will be highlighted in green.
- Unsure about what kinds of items a particular tool Output Targets ??
  - Just click the tool button when it is *not highlighted* to see help information.

Job Submission Status

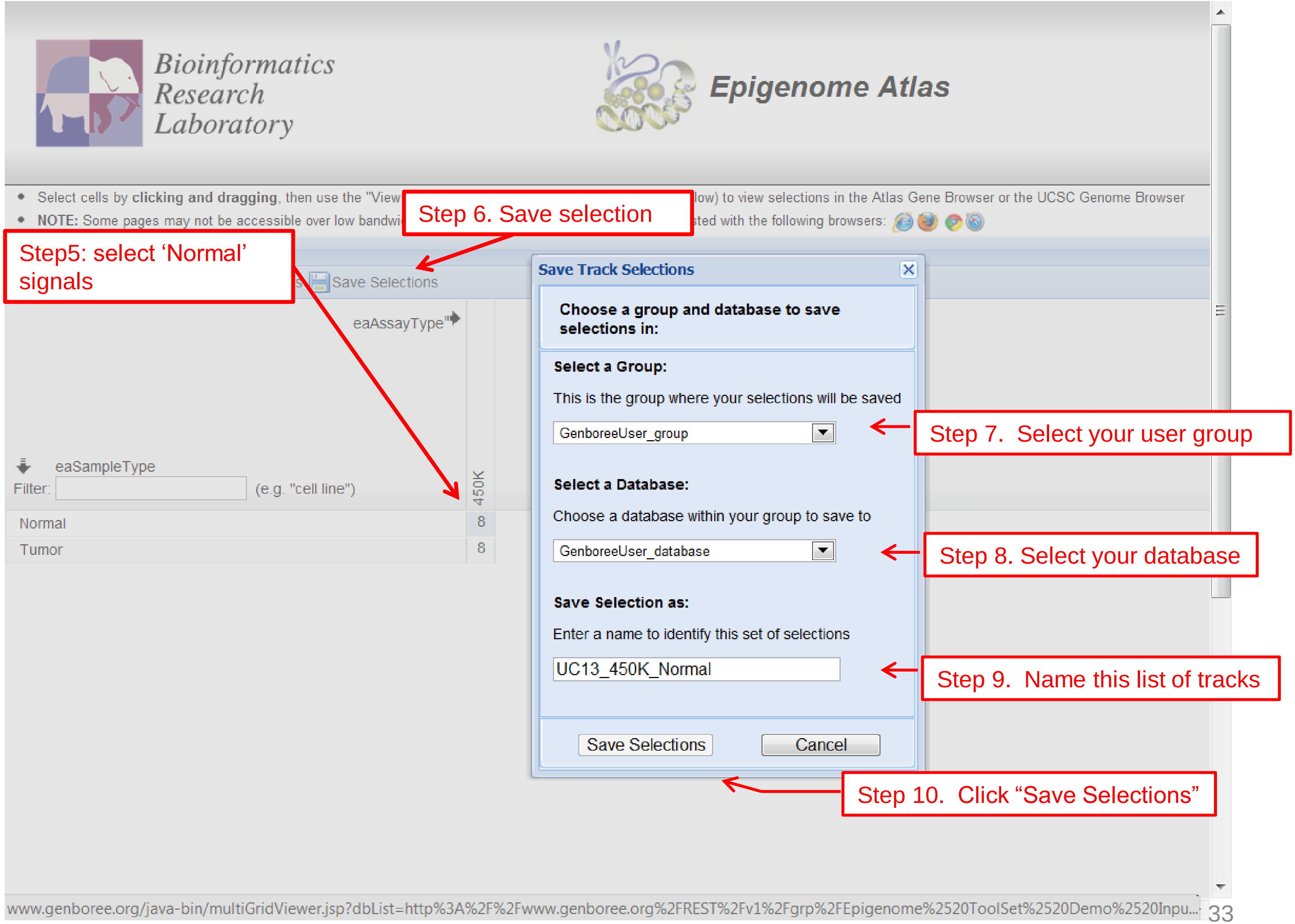
Launch Grid Viewer

Please click the link below to launch the grid viewer:

Launch Grid Viewer.

If you have questions, please contact **genboree_admin@genboree.org** for assistance.

Step 4. Click "OK"



Bioinformatics Research Laboratory
Epigenome Atlas

- Select cells by **clicking and dragging**, then use the "View … low) to view selections in the Atlas Gene Browser or the UCSC Genome Browser
- **NOTE:** Some pages may not be accessible over low bandwi… ted with the following browsers:

Save Selections
eaAssayType
eaSampleType
Filter:
(e.g. "cell line")
450K
Normal 8
Tumor 8

Save Track Selections

Choose a group and database to save selections in:

Select a Group:

This is the group where your selections will be saved

GenboreeUser_group

Select a Database:

Choose a database within your group to save to

GenboreeUser_database

Save Selection as:

Enter a name to identify this set of selections

UC13_450K_Normal

Save Selections
Cancel

Step5: select 'Normal' signals

Step 6. Save selection

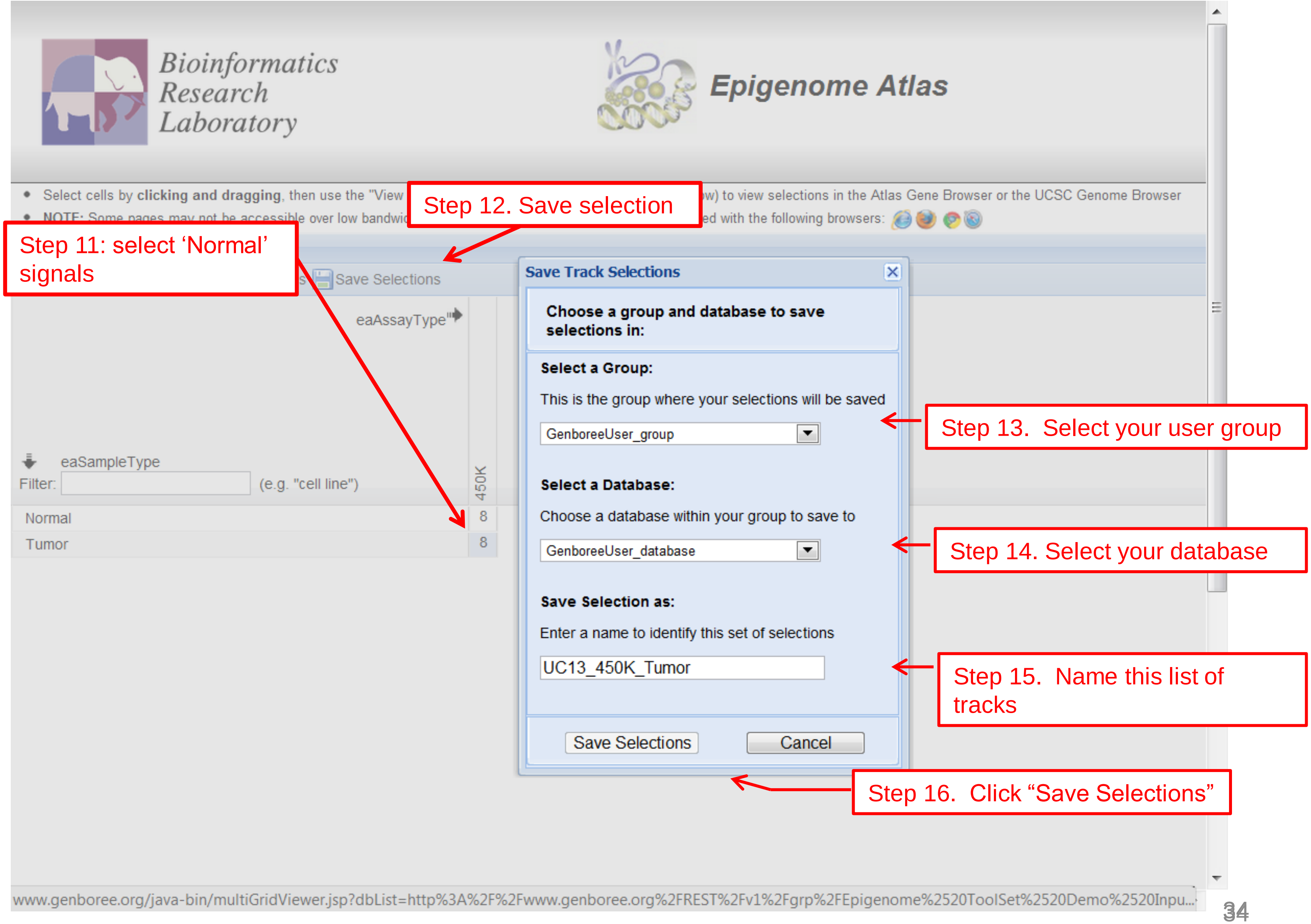
Step 7. Select your user group

Step 8. Select your database

Step 9. Name this list of tracks

Step 10. Click "Save Selections"

www.genboree.org/java-bin/multiGridViewer.jsp?dbList=http%3A%2F%2Fwww.genboree.org%2FREST%2Fv1%2Fgrp%2FEpigenome%2520ToolSet%2520Demo%2520Inpu...


Bioinformatics
Research
Laboratory
Epigenome Atlas
Save Selections
eaAssayType
eaSampleType
450K
Save Selections
Cancel

- Select cells by **clicking and dragging**, then use the "View … w) to view selections in the Atlas Gene Browser or the UCSC Genome Browser
- NOTE: Some pages may not be accessible over low bandwid… ed with the following browsers:

Filter: (e.g. "cell line")

	450K
Normal	8
Tumor	8

Save Track Selections

Choose a group and database to save selections in:

Select a Group:

This is the group where your selections will be saved

GenboreeUser_group

Select a Database:

Choose a database within your group to save to

GenboreeUser_database

Save Selection as:

Enter a name to identify this set of selections

UC13_450K_Tumor

www.genboree.org/java-bin/multiGridViewer.jsp?dbList=http%3A%2F%2Fwww.genboree.org%2FREST%2Fv1%2Fgrp%2FEpigenome%2520ToolSet%2520Demo%2520Inpu...

Step 11: select 'Normal' signals

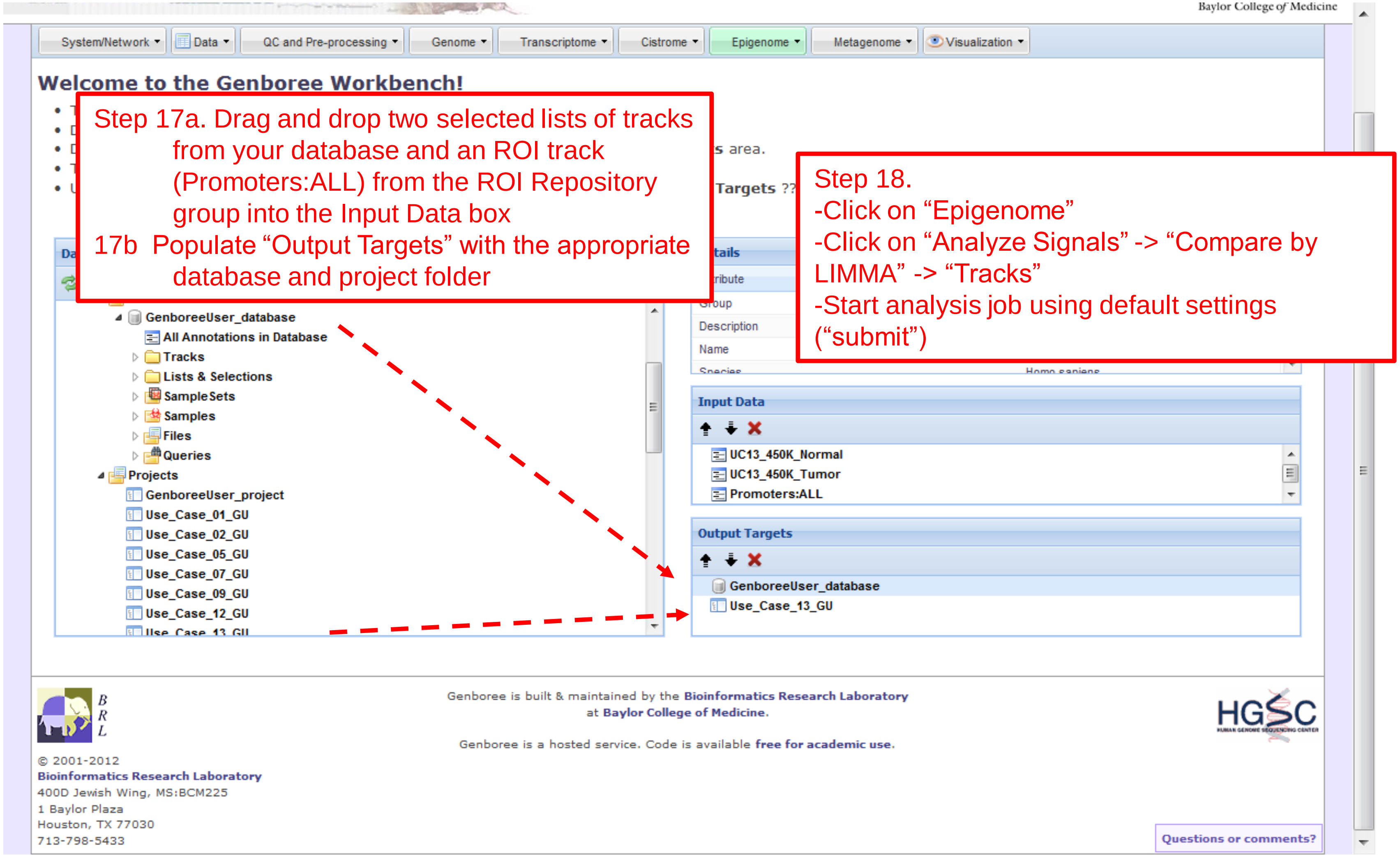
Step 12. Save selection

Step 13. Select your user group

Step 14. Select your database

Step 15. Name this list of tracks

Step 16. Click "Save Selections"


Baylor College of Medicine
System/Network
Data
QC and Pre-processing
Genome
Transcriptome
Cistrome
Epigenome
Metagenome
Visualization
Welcome to the Genboree Workbench!
s area.
Targets ??
Da
tails
ribute
Group
Description
Name
Species
Homo sapiens
GenboreeUser_database
All Annotations in Database
Tracks
Lists & Selections
SampleSets
Samples
Files
Queries
Projects
GenboreeUser_project
Use_Case_01_GU
Use_Case_02_GU
Use_Case_05_GU
Use_Case_07_GU
Use_Case_09_GU
Use_Case_12_GU
Use_Case_13_GU
HGSC
HUMAN GENOME SEQUENCING CENTER
B
R
L
Questions or comments?

Step 17a. Drag and drop two selected lists of tracks from your database and an ROI track (Promoters:ALL) from the ROI Repository group into the Input Data box

17b Populate “Output Targets” with the appropriate database and project folder

Step 18.

-Click on “Epigenome”

-Click on “Analyze Signals” -> “Compare by LIMMA” -> “Tracks”

-Start analysis job using default settings (“submit”)

Input Data

UC13_450K_Normal
UC13_450K_Tumor
Promoters:ALL

Output Targets

GenboreeUser_database
Use_Case_13_GU

Genboree is built & maintained by the **Bioinformatics Research Laboratory** at **Baylor College of Medicine**.

Genboree is a hosted service. Code is available **free for academic use**.

© 2001-2012
Bioinformatics Research Laboratory
400D Jewish Wing, MS:BCM225
1 Baylor Plaza
Houston, TX 77030
713-798-5433

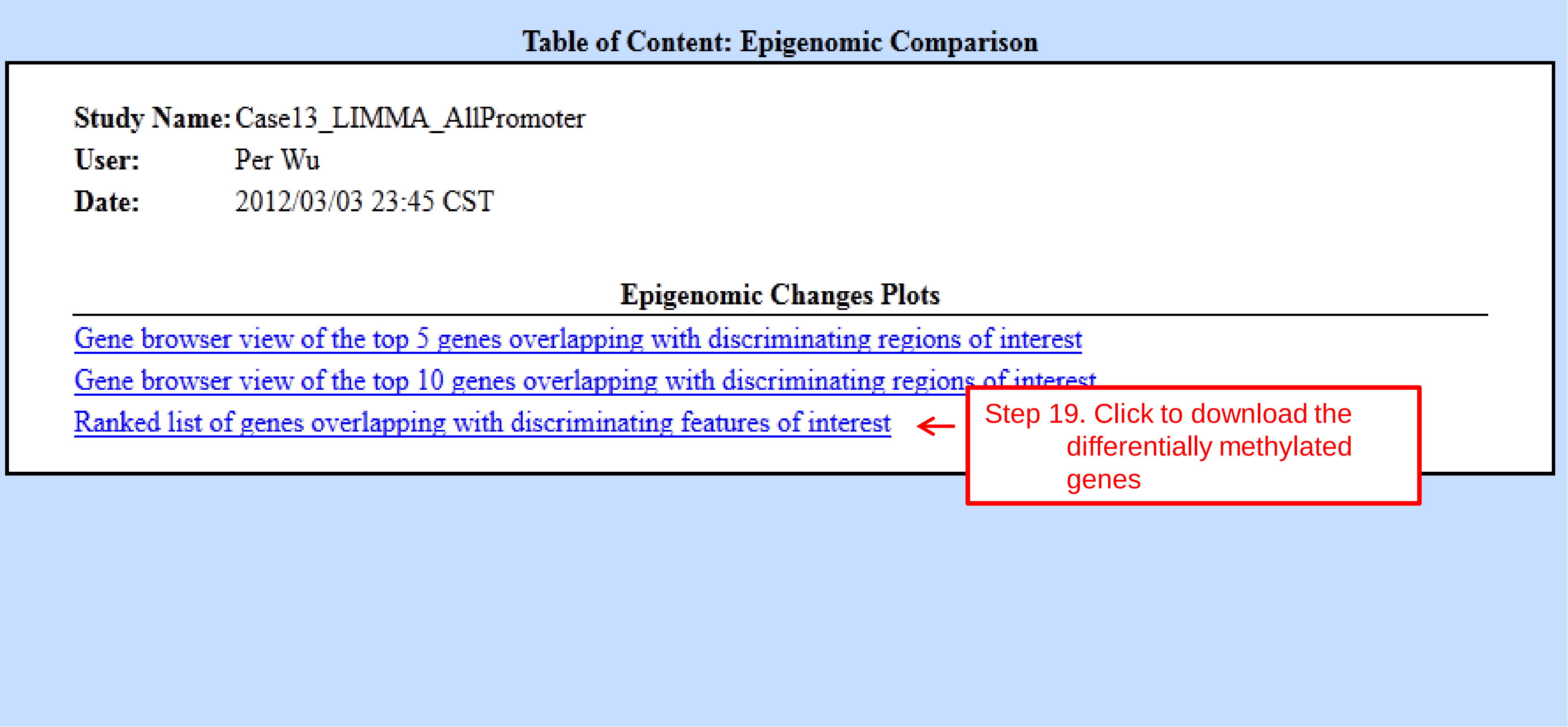


Table of Content: Epigenomic Comparison

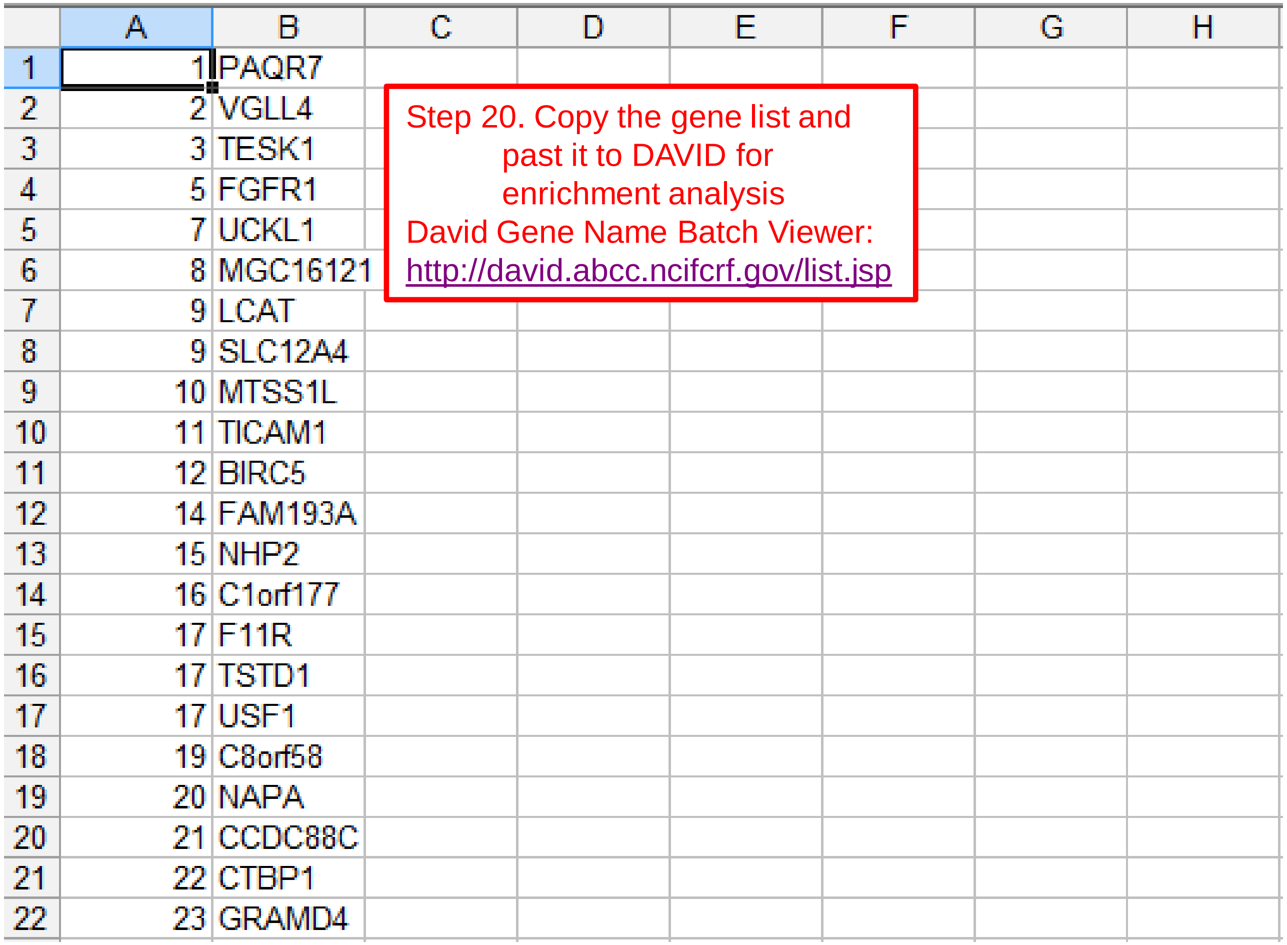
Study Name: Case13_LIMMA_AllPromoter
User: Per Wu
Date: 2012/03/03 23:45 CST

Epigenomic Changes Plots

Gene browser view of the top 5 genes overlapping with discriminating regions of interest
Gene browser view of the top 10 genes overlapping with discriminating regions of interest
Ranked list of genes overlapping with discriminating features of interest

Step 19. Click to download the differentially methylated genes

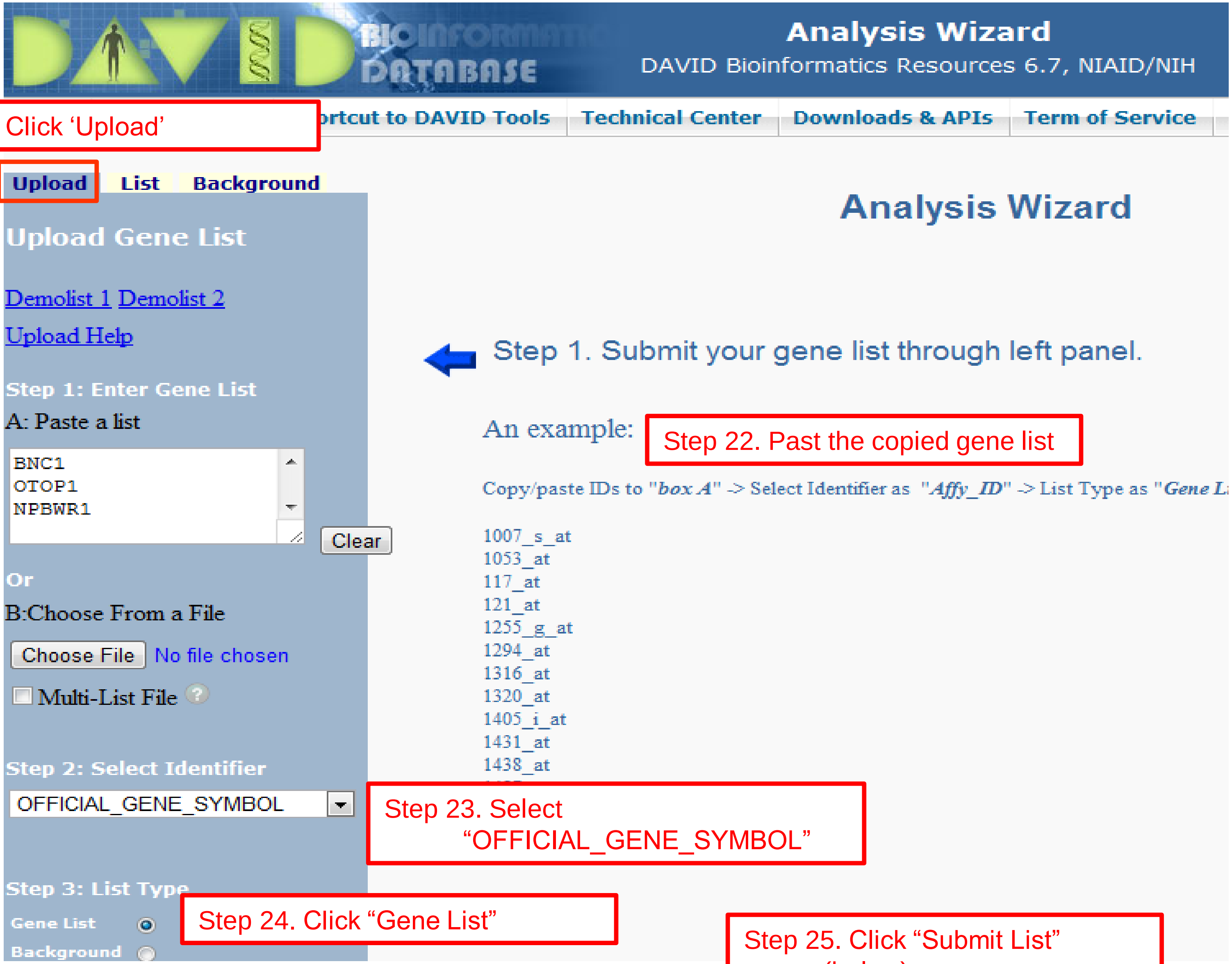
Part 3: The Gene List



	A	B	C	D	E	F	G	H
1	1	PAQR7						
2	2	VGLL4						
3	3	TESK1						
4	5	FGFR1						
5	7	UCKL1						
6	8	MGC16121						
7	9	LCAT						
8	9	SLC12A4						
9	10	MTSS1L						
10	11	TICAM1						
11	12	BIRC5						
12	14	FAM193A						
13	15	NHP2						
14	16	C1orf177						
15	17	F11R						
16	17	TSTD1						
17	17	USF1						
18	19	C8orf58						
19	20	NAPA						
20	21	CCDC88C						
21	22	CTBP1						
22	23	GRAMD4						

Step 20. Copy the gene list and
past it to DAVID for
enrichment analysis
David Gene Name Batch Viewer:
http://david.abcc.ncifcrf.gov/list.jsp

Part 3: The Gene List



Analysis Wizard
DAVID Bioinformatics Resources 6.7, NIAID/NIH

rtcut to DAVID Tools
Technical Center
Downloads & APIs
Term of Service
Upload
List
Background

Upload Gene List

Demolist 1 Demolist 2

Upload Help

Step 1: Enter Gene List

A: Paste a list

```
BNC1
OTOP1
NPBWR1
```

Clear

Or

B:Choose From a File

Choose File

 No file chosen

☐ Multi-List File

Step 2: Select Identifier

OFFICIAL_GENE_SYMBOL

Step 3: List Type

Gene List ●
Background ○

Analysis Wizard

Step 1. Submit your gene list through left panel.

An example:

Copy/paste IDs to "*box A*" -> Select Identifier as "*Affy_ID*" -> List Type as "*Gene L*

1007_s_at
1053_at
117_at
121_at
1255_g_at
1294_at
1316_at
1320_at
1405_i_at
1431_at
1438_at

Step 21. Click 'Upload'

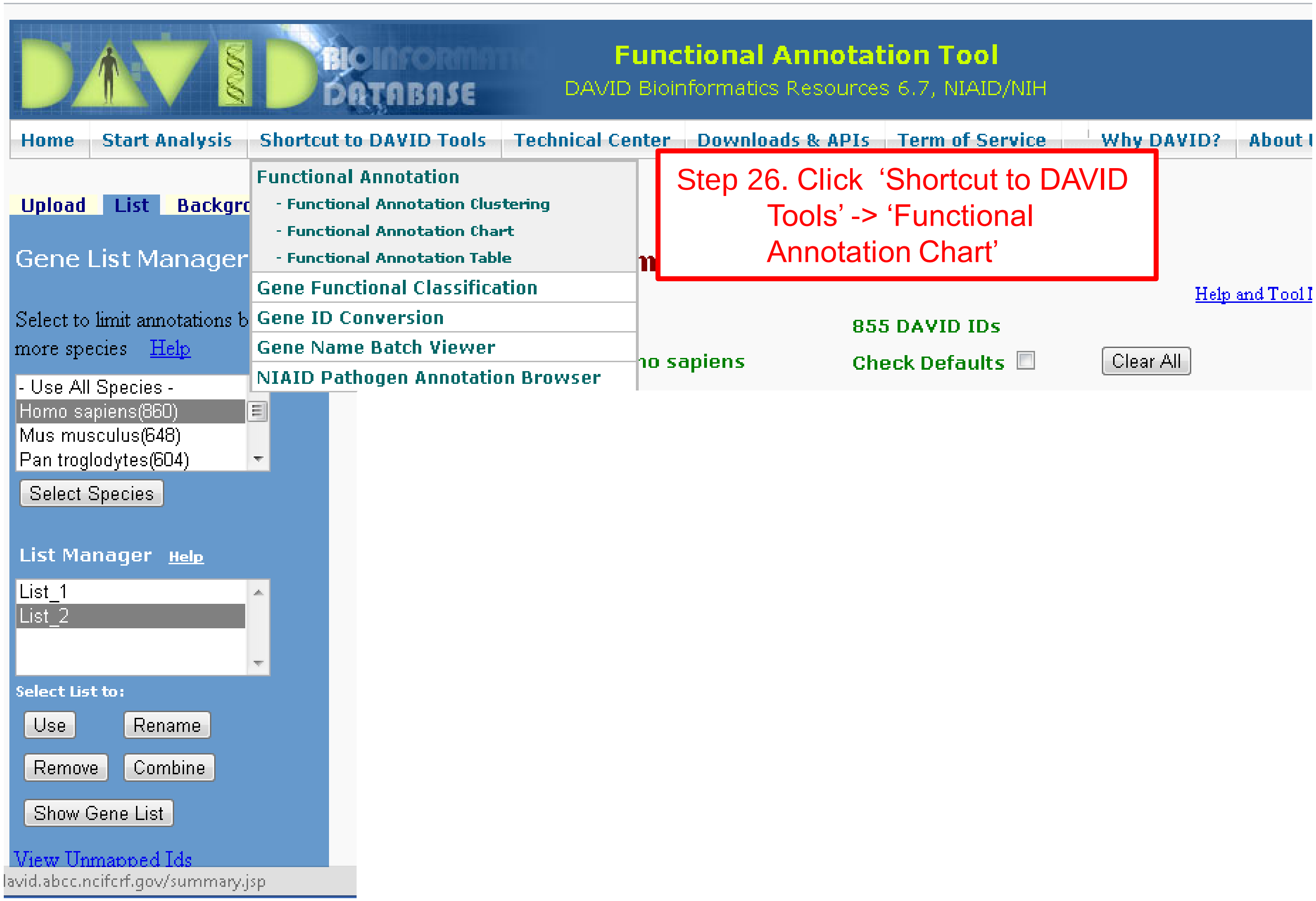
Step 22. Past the copied gene list

Step 23. Select
"OFFICIAL_GENE_SYMBOL"

Step 24. Click "Gene List"

Step 25. Click "Submit List"
(below)

DAVID – Functional Annotation Chart


Functional Annotation Tool
DAVID Bioinformatics Resources 6.7, NIAID/NIH
Home
Start Analysis
Shortcut to DAVID Tools
Technical Center
Downloads & APIs
Term of Service
Why DAVID?
About
Functional Annotation
- Functional Annotation Clustering
- Functional Annotation Chart
- Functional Annotation Table
Gene Functional Classification
Gene ID Conversion
Gene Name Batch Viewer
NIAID Pathogen Annotation Browser
Upload
List
Backgro
Gene List Manager
Select to limit annotations b
more species Help
- Use All Species -
Homo sapiens(860)
Mus musculus(648)
Pan troglodytes(604)
Select Species
List Manager Help
List_1
List_2
Select List to:
Use
Rename
Remove
Combine
Show Gene List
View Unmapped Ids
Help and Tool
855 DAVID IDs
no sapiens
Check Defaults ☐
Clear All
lavid.abcc.ncifcrf.gov/summary.jsp

Step 26. Click ‘Shortcut to DAVID Tools’ -> ‘Functional Annotation Chart’

DAVID – Gene Ontology

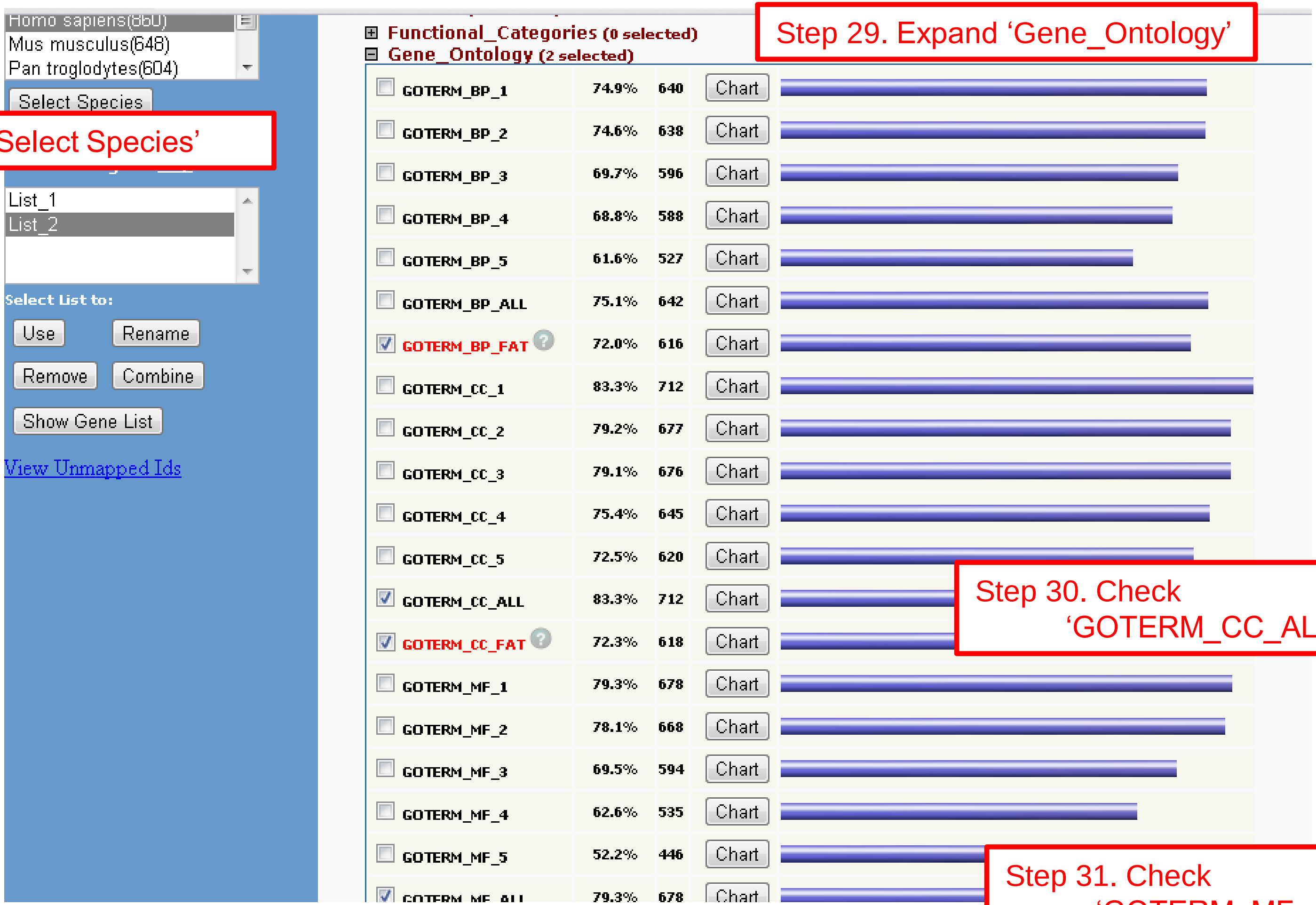
Step 27. Select Homo sapiens

Step 28. Click ‘Select Species’

Step 29. Expand ‘Gene_Ontology’

Step 30. Check ‘GOTERM_CC_ALL’

Step 31. Check ‘GOTERM_MF_ALL’



Homo sapiens(860)
Mus musculus(648)
Pan troglodytes(604)

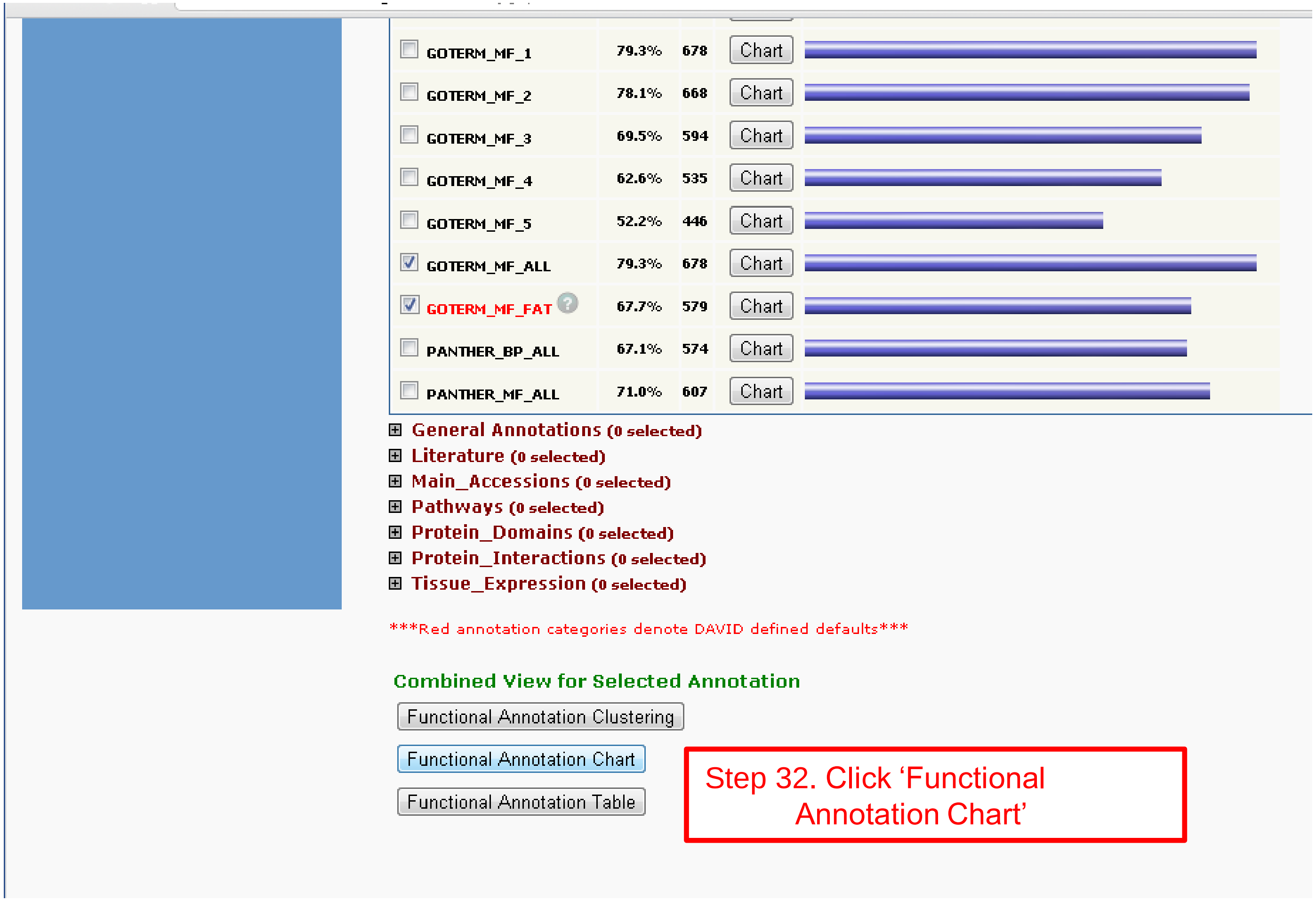
List_1
List_2

Select List to:

⊞ Functional_Categories (0 selected)
⊟ Gene_Ontology (2 selected)

☐	GOTERM_BP_1	74.9%	640	Chart
☐	GOTERM_BP_2	74.6%	638	Chart
☐	GOTERM_BP_3	69.7%	596	Chart
☐	GOTERM_BP_4	68.8%	588	Chart
☐	GOTERM_BP_5	61.6%	527	Chart
☐	GOTERM_BP_ALL	75.1%	642	Chart
☑	GOTERM_BP_FAT	72.0%	616	Chart
☐	GOTERM_CC_1	83.3%	712	Chart
☐	GOTERM_CC_2	79.2%	677	Chart
☐	GOTERM_CC_3	79.1%	676	Chart
☐	GOTERM_CC_4	75.4%	645	Chart
☐	GOTERM_CC_5	72.5%	620	Chart
☑	GOTERM_CC_ALL	83.3%	712	Chart
☑	GOTERM_CC_FAT	72.3%	618	Chart
☐	GOTERM_MF_1	79.3%	678	Chart
☐	GOTERM_MF_2	78.1%	668	Chart
☐	GOTERM_MF_3	69.5%	594	Chart
☐	GOTERM_MF_4	62.6%	535	Chart
☐	GOTERM_MF_5	52.2%	446	Chart
☑	GOTERM_MF_ALL	79.3%	678	Chart

DAVID – View Functional Annotation Chart



☐ GOTERM_MF_1	79.3%	678	Chart	
☐ GOTERM_MF_2	78.1%	668	Chart	
☐ GOTERM_MF_3	69.5%	594	Chart	
☐ GOTERM_MF_4	62.6%	535	Chart	
☐ GOTERM_MF_5	52.2%	446	Chart	
☑ GOTERM_MF_ALL	79.3%	678	Chart	
☑ GOTERM_MF_FAT	67.7%	579	Chart	
☐ PANTHER_BP_ALL	67.1%	574	Chart	
☐ PANTHER_MF_ALL	71.0%	607	Chart	

- ⊞ General Annotations (0 selected)
- ⊞ Literature (0 selected)
- ⊞ Main_Accessions (0 selected)
- ⊞ Pathways (0 selected)
- ⊞ Protein_Domains (0 selected)
- ⊞ Protein_Interactions (0 selected)
- ⊞ Tissue_Expression (0 selected)

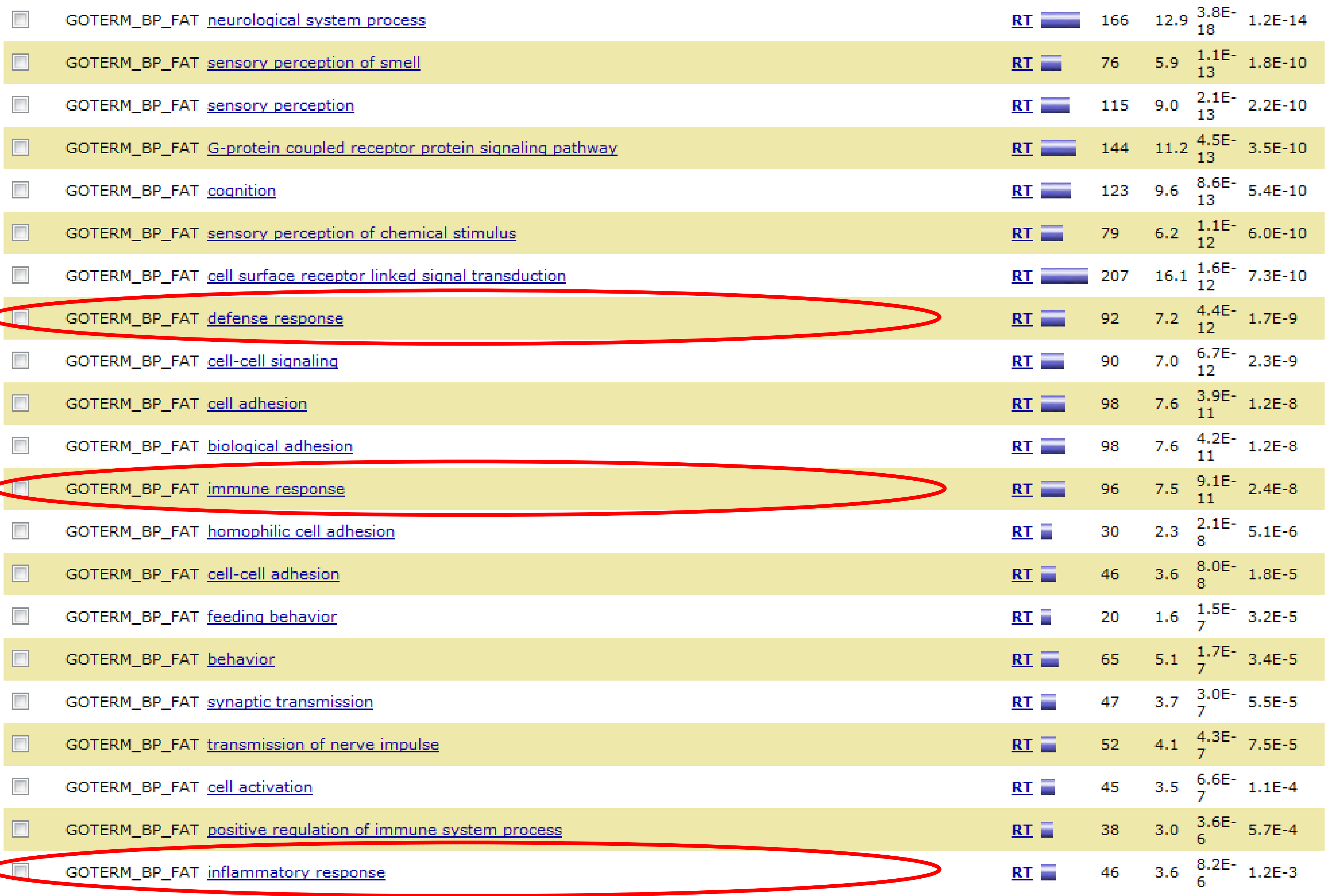
Red annotation categories denote DAVID defined defaults

Combined View for Selected Annotation

Functional Annotation Clustering
Functional Annotation Chart
Functional Annotation Table

Step 32. Click ‘Functional Annotation Chart’

Part 3: Enriched Biological Process



☐	GOTERM_BP_FAT	neurological system process	RT	166	12.9	3.8E-18	1.2E-14
☐	GOTERM_BP_FAT	sensory perception of smell	RT	76	5.9	1.1E-13	1.8E-10
☐	GOTERM_BP_FAT	sensory perception	RT	115	9.0	2.1E-13	2.2E-10
☐	GOTERM_BP_FAT	G-protein coupled receptor protein signaling pathway	RT	144	11.2	4.5E-13	3.5E-10
☐	GOTERM_BP_FAT	cognition	RT	123	9.6	8.6E-13	5.4E-10
☐	GOTERM_BP_FAT	sensory perception of chemical stimulus	RT	79	6.2	1.1E-12	6.0E-10
☐	GOTERM_BP_FAT	cell surface receptor linked signal transduction	RT	207	16.1	1.6E-12	7.3E-10
☐	GOTERM_BP_FAT	defense response	RT	92	7.2	4.4E-12	1.7E-9
☐	GOTERM_BP_FAT	cell-cell signaling	RT	90	7.0	6.7E-12	2.3E-9
☐	GOTERM_BP_FAT	cell adhesion	RT	98	7.6	3.9E-11	1.2E-8
☐	GOTERM_BP_FAT	biological adhesion	RT	98	7.6	4.2E-11	1.2E-8
☐	GOTERM_BP_FAT	immune response	RT	96	7.5	9.1E-11	2.4E-8
☐	GOTERM_BP_FAT	homophilic cell adhesion	RT	30	2.3	2.1E-8	5.1E-6
☐	GOTERM_BP_FAT	cell-cell adhesion	RT	46	3.6	8.0E-8	1.8E-5
☐	GOTERM_BP_FAT	feeding behavior	RT	20	1.6	1.5E-7	3.2E-5
☐	GOTERM_BP_FAT	behavior	RT	65	5.1	1.7E-7	3.4E-5
☐	GOTERM_BP_FAT	synaptic transmission	RT	47	3.7	3.0E-7	5.5E-5
☐	GOTERM_BP_FAT	transmission of nerve impulse	RT	52	4.1	4.3E-7	7.5E-5
☐	GOTERM_BP_FAT	cell activation	RT	45	3.5	6.6E-7	1.1E-4
☐	GOTERM_BP_FAT	positive regulation of immune system process	RT	38	3.0	3.6E-6	5.7E-4
☐	GOTERM_BP_FAT	inflammatory response	RT	46	3.6	8.2E-6	1.2E-3