Attention! Please specify PDF title in FAQ Settings first.

---

1. Getting Started with Genboree

1.1 How Do I Upload Files Onto The Workbench? Generic Example

Go to http://www.genboree.org
Log in if necessary
Click on the 'Workbench' tab or directly visit the Workbench http://genboree.org/java-bin/workbench.jsp
Expand the Group that contains your Database that you would like to upload files to
Expand 'Databases'
Drag your target database from the 'Data Selector' window over into the 'Output Targets' window
Select the 'Data' tab, the 'Files' tab, and then the 'Transfer File' tab
Click the 'Choose File' button and select the file you wish to upload
If you are uploading a multi-file archive (aka. a compressed file that contains multiple files), check the 'Unpack Multi-File Archive' check box
Optionally type in a sub-folder into the 'Create in Sub-Folder' text area if you would like to place the uploaded file in a folder other than the 'File' folder
Optionally type in a file description into the 'File Description' text area
Click 'Submit'

!http://genboree.org/theCommons/attachments/1602/upload_data_1.PNG!

!http://genboree.org/theCommons/attachments/1603/upload_data_2.PNG!

!http://genboree.org/theCommons/attachments/1604/upload_data_3.PNG!

!http://genboree.org/theCommons/attachments/1605/upload_data_4.PNG!

!http://genboree.org/theCommons/attachments/1606/upload_data_5.PNG!

!http://genboree.org/theCommons/attachments/1600/upload_data_6.PNG!

!http://genboree.org/theCommons/attachments/1601/upload_data_7.PNG!

1.2 How do I create a new Genboree Database?

Go to http://www.genboree.org
Log in if necessary
Click on the 'Workbench' tab or directly visit the Workbench http://genboree.org/java-bin/workbench.jsp
Drag your Genboree Group to the Output Targets window. The Genboree Database will be created within this Genboree Group. Also see, "How to Create a Genboree Group":http://genboree.org/theCommons/ezfaq/show/epigenome-workshop?faq_id=466
Click 'Data' -> 'Databases' -> 'Create Database'
Select your Reference Sequence (i.e. 'Template: Human (hg19)'). NOTE: If you are using the Genboree Workbench for Metagenome analysis, plan to upload your reference sequence, or do not know your reference sequence please select '** User Will Upload **' (default).
Enter Database Name and optional Description.
Click 'Submit'
Click Refresh and expand your Genboree Group -> Databases to see your newly created Genboree Database in the Data Selector window

h1. Go to the Genboree Workbench

!http://genboree.org/theCommons/attachments/2701/db_1.png!

h1. Drag the Genboree Group to Output Targets

!http://genboree.org/theCommons/attachments/2702/db_2.png!

h1. Click Data -> Databases -> Create Database

!http://genboree.org/theCommons/attachments/2703/db_3.png!

h1. The default Genboree Database will be empty

!http://genboree.org/theCommons/attachments/2704/db_4.png!

h1. Select the appropriate Reference Sequence (i.e. Human hg19) if applicable

!http://genboree.org/theCommons/attachments/2705/db_5.png!

h1. Enter Database Name and optional Description

!http://genboree.org/theCommons/attachments/2706/db_6.png!

h1. Click OK after you have created your Genboree Database

!http://genboree.org/theCommons/attachments/2707/db_7.png!

h1. Click Refresh and expand your Genboree Group -> Databases to see your newly created Genboree Database

!http://genboree.org/theCommons/attachments/2708/db_8.png!

h1. See details of your Genboree Database in the Details window

!http://genboree.org/theCommons/attachments/2709/db_9.png!

1.3 How do I create a new Genboree Group?
Go to http://www.genboree.org
Log in if necessary
Click on the 'Workbench' tab or directly visit the Workbench http://genboree.org/java-bin/workbench.jsp
Drag the _Host_ to the Output Targets window
** For most users this will be 'genboree.org'
# Click 'System/Network' -> 'Groups' -> 'Create Group'
# Enter a Group Name
# Optionally enter a Description
# Click Submit
# Click _Refresh_ to see your newly created Genboree Group in the Data Selector window

This tool will create/add a new group in Genboree. Note that the group to be created should NOT already exist at the target host (i.e. a Genboree installation at a particular site). The Groups menu also has tools for managing your group, including editing Group information, Deleting a Group, adding new users, and updating roles of Group members (i.e. subscriber, author, administrator). You can also send messages to everyone in your Group through Genboree.

h1. Go to the Genboree Workbench

!http://genboree.org/theCommons/attachments/2685/group_1.png!

h1. Drag the Genboree Host (i.e. 'genboree.org') to Output Targets

!http://genboree.org/theCommons/attachments/2686/group_2.png!

h1. Select System/Network -> Groups -> Create Group

!http://genboree.org/theCommons/attachments/2687/group_3.png!

h1. Enter Genboree Group name and optional description

!http://genboree.org/theCommons/attachments/2688/group_4.png!

h1. Select OK after you have created your Genboree Group

!http://genboree.org/theCommons/attachments/2689/group_5.png!

h1. Click Refresh to see your newly created Genboree Group

!http://genboree.org/theCommons/attachments/2690/group_6.png!

h1. See details of your Genboree Group in Details window

!http://genboree.org/theCommons/attachments/2691/group_7.png!

1.4 How do I create a new Genboree Project?

Go to http://www.genboree.org
Log in if necessary
Click on the 'Workbench' tab or directly visit the Workbench http://genboree.org/java-bin/workbench.jsp
Drag your Genboree Group to the Output Targets window. The Genboree Project will be created within this Genboree Group. Also see, "How to Create a Genboree Group":http://genboree.org/theCommons/ezfaq/show/epigenome-workshop?faq_id=466
Click 'Data' -> 'Projects' -> 'Create Project'
Enter Project Name and optionally Project Title and Project Description. Note: Project Name must be globally unique within Genboree, which means that you cannot duplicate a Project Name that has already been created. If necessary, you can a unique identifier such as your initials or other set of characters that will make the project name unique across Genboree.
Click 'Submit'
Click Refresh and expand your Genboree Group -> Projects to see your newly created Genboree Project in the Data Selector window
Optionally click on 'Link to Project' to see your newly created Genboree Project HTML page

h1. Go to the Genboree Workbench

!http://genboree.org/theCommons/attachments/2692/project_1.png!

h1. Drag the Genboree Group to Output Targets

!http://genboree.org/theCommons/attachments/2693/project_2.png!

h1. Click Data -> Projects -> Create Project

!http://genboree.org/theCommons/attachments/2694/project_3.png!

h1. Enter (unique) Genboree Project Name and optionally Project Title and Project Description

!http://genboree.org/theCommons/attachments/2695/project_4.png!

h1. Click OK after you have created your Genboree Group

!http://genboree.org/theCommons/attachments/2696/project_5.png!

h1. Click Refresh and expand your Genboree Group -> Projects to see your newly created Genboree Project

!http://genboree.org/theCommons/attachments/2697/project_6.png!

h1. See details of your Genboree Project in Details window

!http://genboree.org/theCommons/attachments/2698/project_7.png!

h1. Optionally click on 'Link to Project' in Details window

!http://genboree.org/theCommons/attachments/2699/project_8.png!

h1. New HTML Genboree Project page displays Project Title and Project Description

!http://genboree.org/theCommons/attachments/2700/project_9.png!

1.5 How do I create, edit, delete, and upload annotations / entry points to a Genboree Database?
Direct 'How-to' Link:
* http://valine.brl.bcmd.bcm.edu/java-bin/myrefseq.jsp

Or you can do the following to access the 'how-to' information:
Go to http://www.genboree.org
Log in if necessary
Click on the 'Workbench' tab or directly visit the Workbench http://genboree.org/java-bin/workbench.jsp
Click the 'Databases' tab in the top Genboree Menu Bar
Click 'How-to'

This link addresses the following:
* Create
* Edit, unlock, and publish
* Delete
* Upload annotations
* Upload, delete, and edit chromosomes / entry points

1.6 What are Annotations?
* An Annotation in Genboree is a coordinate range on a defined Entry Point (such as a chromosome).
** Examples include: exons, chromosome bands, mapped BACs, repeat regions, promotor sites, base-conservation scores, etc.

* Genboree also has the concept of Annotation Groups, which are sets of linked annotations that are best displayed via group-aware drawing styles.
** For example, the exons of a gene transcript should always grouped together.
** Other examples include mate-pair reads, the HSPs of a significant BLAST/BLAT hit, etc.

Helpful links:
* http://www.genboree.org/java-bin/showHelp.jsp?topic=uploadAnnoHowto
* http://www.genboree.org/java-bin/showHelp.jsp?topic=definingAGenome

1.7 What are Entry Points?
h2. What are Entry Points?

* Entry points comprise the coordinate system on which annotations are made.
** For example, a chromosome is an entry point. For unassembled genomes, the scaffolds might be the entry points.
* Entry points must be defined within Genboree, with a name and a length, so that annotations can be placed on them.
* Currently, entry points are independent within Genboree; i.e. not directly linked. They are simply the names of chromosomes

Uploading entry points:
* Entry points can be uploaded in two different formats:
** 3-column LFF
*** The entry point name
*** The keyword "Chromosome"
*** The length of the entry point
** Fasta files

Helpful links:
* http://www.genboree.org/java-bin/showHelp.jsp?topic=definingAGenome#whatAreEPs
* http://www.genboree.org/java-bin/showHelp.jsp?topic=uploadEPhowto

h2. Real Example

As an example, you can download the chromosome sizes for Mouse MM10 from UCSC:
* http://hgdownload.cse.ucsc.edu/goldenPath/mm10/database/chromInfo.txt.gz

We can turn this into a 3 column LFF format for uploading onto Genboree with some very light formatting:
<pre>
chr1 Chromosome 195471971
chr2 Chromosome 182113224
chrX Chromosome 171031299
chr3 Chromosome 160039680
chr4 Chromosome 156508116
chr5 Chromosome 151834684
chr6 Chromosome 149736546
chr7 Chromosome 145441459
chr10 Chromosome 130694993
chr8 Chromosome 129401213
chr14 Chromosome 124902244
chr9 Chromosome 124595110
chr11 Chromosome 122082543
chr13 Chromosome 120421639

chr12 Chromosome 120129022
chr15 Chromosome 104043685
chr16 Chromosome 98207768
chr17 Chromosome 94987271
chrY Chromosome 91744698
chr18 Chromosome 90702639
chr19 Chromosome 61431566
chr5_JH584299_random Chromosome 953012
chrX_GL456233_random Chromosome 336933
chrY_JH584301_random Chromosome 259875
chr1_GL456211_random Chromosome 241735
chr4_GL456350_random Chromosome 227966
chr4_JH584293_random Chromosome 207968
chr1_GL456221_random Chromosome 206961
chr5_JH584297_random Chromosome 205776
chr5_JH584296_random Chromosome 199368
chr5_GL456354_random Chromosome 195993
chr4_JH584294_random Chromosome 191905
chr5_JH584298_random Chromosome 184189
chrY_JH584300_random Chromosome 182347
chr7_GL456219_random Chromosome 175968
chr1_GL456210_random Chromosome 169725
chrY_JH584303_random Chromosome 158099
chrY_JH584302_random Chromosome 155838
chr1_GL456212_random Chromosome 153618
chrUn_JH584304 Chromosome 114452
chrUn_GL456379 Chromosome 72385
chr4_GL456216_random Chromosome 66673
chrUn_GL456393 Chromosome 55711
chrUn_GL456366 Chromosome 47073
chrUn_GL456367 Chromosome 42057
chrUn_GL456239 Chromosome 40056
chr1_GL456213_random Chromosome 39340
chrUn_GL456383 Chromosome 38659
chrUn_GL456385 Chromosome 35240
chrUn_GL456360 Chromosome 31704
chrUn_GL456378 Chromosome 31602
chrUn_GL456389 Chromosome 28772
chrUn_GL456372 Chromosome 28664
chrUn_GL456370 Chromosome 26764
chrUn_GL456381 Chromosome 25871
chrUn_GL456387 Chromosome 24685
chrUn_GL456390 Chromosome 24668
chrUn_GL456394 Chromosome 24323
chrUn_GL456392 Chromosome 23629
chrUn_GL456382 Chromosome 23158
chrUn_GL456359 Chromosome 22974
chrUn_GL456396 Chromosome 21240
chrUn_GL456368 Chromosome 20208
chrM Chromosome 16299
chr4_JH584292_random Chromosome 14945
chr4_JH584295_random Chromosome 1976
</pre>

1.8 What data can I upload for analysis onto Genboree?
The answer to this question will require knowing a few things about your data:

What type of analysis do you wish to conduct? (Cistrome (Chip-Seq), Transcriptome (RNA-Seq), Genome (variant detection), Epigenome, Metagenome (16S rRNA))
From what organism does the data originate (i.e. human)? If you list an organism, what reference release of the organism do you have data for (i.e. human hg19, mouse mm10, etc.)?
What is the file format? (BED, BAM/SAM, etc)
How many samples do you have data for? What are the file sizes per sample? What is the file size for the entire experiment? (10GB, 30GB, etc)
From what sequencing platform was the data generated? (i.e. Illumina GAIIx, etc.)
Type of analysis desired (i.e. use RNA-Seq tools to map reads and splice junctions using tophat, use microbiome tools to produce alpha and beta diversity, etc.)

Your response is required for us to initiate the data upload process with you. Please send the answers to the above questions to genboree_admin@genboree.org. We will follow up with you to make sure we have a complete understanding of your data analysis needs before initiating the upload process. Once we have a clear understanding of the data set, we will send you instructions on how to transfer data to our FTP in order to start the data upload process.

If you wish to upload 27K or 450K data, please follow this tutorial:
* "27k / 450k Tutorial":http://genboree.org/theCommons/ezfaq/show/epigenome-workshop?faq_id=474

1.9 What is a Genome in Genboree?
* A "genome" is a set of entry points. It is the complete set of valid chromosomes, or maybe scaffolds, upon which your annotations can be placed.
* In fact, you could have a set of entry points from several species (for example, the X chromosomes of several mammals).

Helpful links:
* http://www.genboree.org/java-bin/showHelp.jsp?topic=definingAGenome

1.10 What is an Entity List?
* An Entity List is a list of sequencing data tracks or files (the â€œentitiesâ€•).
* Some tools in the Genboree workbench require an Entity List as input, on which the operations or computations are performed.

1.11 What is the LFF Annotation Format?
* This is the preferred annotation format, because it is the most expressive and feature-filled for Genboree.
* Particular fields and values in this format enable several features for capturing and displaying your annotation data in the most complete and customizable way.
* The Genboree LFF format is adapted from the LDAS upload format:
** http://www.biodas.org/
* Although it is backwards compatible with the original, Genboree LFF includes some optional fields to support extra annotation information and browser capabilities.
* Due to the importance of the LFF format within Genboree, please read "The LFF File Format" for a description of the format and how the columns relate to Genboree:
** http://www.genboree.org/java-bin/showHelp.jsp?topic=lffFileFormat

2. Using the Grid Viewer

2.1 How do I use the grid to create a List of Tracks? Assay Type vs Sample Type

Go to http://www.genboree.org
Log in if necessary
Click on the 'Workbench' tab or directly visit the Workbench http://genboree.org/java-bin/workbench.jsp
Expand the Group that contains your Database of interest
Expand 'Databases'
Drag your Database into the 'Input Data' window
Click the 'Visualization' tab and then 'View Track Grid'
Select your 'X-axis attribute' - i.e. 'eaAssayType'
Select your 'Y-axis attribute' - i.e. 'eaSampleType'
Click 'Submit'
Click the link titled 'Launch Grid Viewer'
Select your desired tracks by single clicking individual cells or dragging your cursor over the cells
Click 'Selections' -> 'Save Selections'
Select the Group in which your Database exists in which you would like to output this Track List
Select your Database in which you would like to output this Track List
Enter an informative name into the text box to label this group of tracks
Click 'Save Selections'

!http://genboree.org/theCommons/attachments/1575/grid_1.PNG!

!http://genboree.org/theCommons/attachments/1576/grid_2.PNG!

!http://genboree.org/theCommons/attachments/1577/grid_3a.PNG!

!http://genboree.org/theCommons/attachments/1570/grid_4a.PNG!

!http://genboree.org/theCommons/attachments/1572/grid_5a.PNG!

!http://genboree.org/theCommons/attachments/1578/grid_6a.PNG!

!http://genboree.org/theCommons/attachments/1585/grid_7a.PNG!

!http://genboree.org/theCommons/attachments/1587/grid_8a.PNG!

!http://genboree.org/theCommons/attachments/1580/grid_9a.PNG!

!http://genboree.org/theCommons/attachments/1582/grid_10a.PNG!

!http://genboree.org/theCommons/attachments/1588/grid_11a.PNG!

!http://genboree.org/theCommons/attachments/1595/grid_12a.PNG!

!http://genboree.org/theCommons/attachments/1597/grid_13a.PNG!

!http://genboree.org/theCommons/attachments/1590/grid_14a.PNG!

2.2 How do I use the grid to create two Lists of Tracks? Individual vs. Individual

Go to http://www.genboree.org
Log in if necessary
Click on the 'Workbench' tab or directly visit the Workbench http://genboree.org/java-bin/workbench.jsp

Expand the Group that contains your Database of interest
Expand 'Databases'
Drag your Database into the 'Input Data' window
Click the 'Visualization' tab and then 'View Track Grid'
Select your 'X-axis attribute' - i.e. 'eaSampleType'
Select your 'Y-axis attribute' - i.e. 'Individual'
Click 'Submit'
Click the link titled 'Launch Grid Viewer'
Select your desired tracks for group 1 by single clicking individual cells or dragging your cursor over the cells
Click 'Selections' -> 'Save Selections'
Select the Group in which your Database exists in which you would like to output this Track List
Select your Database in which you would like to output this Track List
Enter an informative name into the text box to label this group of tracks
Click 'Save Selections'
Clear the selected tracks by clicking the 'Clear Selections' tab
Select your desired tracks for group 2 by single clicking individual cells or dragging your cursor over the cells
Click the 'Save Selections' tab
Select the Group in which your Database exists in which you would like to output this Track List
Select your Database in which you would like to output this Track List
Enter an informative name into the text box to label this group of tracks
Click 'Save Selections'

!http://genboree.org/theCommons/attachments/1575/grid_1.PNG!

!http://genboree.org/theCommons/attachments/1576/grid_2.PNG!

!http://genboree.org/theCommons/attachments/1579/grid_3b.PNG!

!http://genboree.org/theCommons/attachments/1571/grid_4b.PNG!

!http://genboree.org/theCommons/attachments/1573/grid_5b.PNG!

!http://genboree.org/theCommons/attachments/1584/grid_6b.PNG!

!http://genboree.org/theCommons/attachments/1586/grid_7b.PNG!

!http://genboree.org/theCommons/attachments/1589/grid_8b.PNG!

!http://genboree.org/theCommons/attachments/1581/grid_9b.PNG!

!http://genboree.org/theCommons/attachments/1583/grid_10b.PNG!

!http://genboree.org/theCommons/attachments/1594/grid_11b.PNG!

!http://genboree.org/theCommons/attachments/1596/grid_12b.PNG!

!http://genboree.org/theCommons/attachments/1599/grid_13b.PNG!

!http://genboree.org/theCommons/attachments/1591/grid_14b.PNG!

!http://genboree.org/theCommons/attachments/1592/grid_15b.PNG!

!http://genboree.org/theCommons/attachments/1593/grid_16b.PNG!

!http://genboree.org/theCommons/attachments/1598/grid_17b.PNG!

3. Track Operations

3.1 Converting BED files to Read Density Signal Tracks

Go to http://www.genboree.org
Log in if necessary
Click on the 'Workbench' tab or directly visit the Workbench http://genboree.org/java-bin/workbench.jsp
Click the 'Databases' tab in the top Genboree Menu Bar
Click 'Create Database'
Select the Group in which you would like to create the Database
Select your 'Reference Sequence' (i.e. 'Template: Human (HG19))' or leave this as '** User Will Upload **' if you do not know your Reference Sequence or do not plan on utilizing one
Enter a Database name into the Database' text area
Click 'Create'
Expand the Group that contains your Database of interest
Expand 'Databases'
Drag your Database into the 'Output Targets' window
Drag your target database from the 'Data Selector' window over into the 'Output Targets' window
Select the 'Data' tab, the 'Files' tab, and then the 'Transfer File' tab
Click the 'Choose File' button and select the file you wish to upload
If you are uploading a multi-file archive (aka. a compressed file that contains multiple files), check the 'Unpack Multi-File Archive' check box
Optionally type in a sub-folder into the 'Create in Sub-Folder' text area if you would like to place the uploaded file in a folder other than the 'File' folder
Optionally type in a file description into the 'File Description' text area
Click 'Submit'
Drag your Database into the 'Output Targets' window (if it already doesn't exist within the window)
Drag your uploaded BED file into the 'Input Data' window (if you didn't specify a subfolder when you uploaded it, it will exist within your_group -> Databases -> your_database -> Files -> your_bed_file)
Select the 'Analysis' tab, the 'Track Tools' tab, and 'Coverage Computation'
Customize the coverage tool settings or leave the default
Click 'Submit'
View the output of the read coverage tool by clicking the top menu 'Browser', select your 'Group', select your 'Database', select your 'Entry Point', (optionally select your From and To genomic coordinates) and click 'View'
In order to visualize this track in context of the UCSC browser click on the 'Databases' tab, select your 'Group', select your 'Database', click on '"Big* Files"'
Select 'no file' under the 'BigWig File' column
Click 'Generate Files'
You will receive a message alerting you that your job has been submitted and you will receive an email when this has completed
Click the 'refresh' button to see that your job has been submitted
Click 'refresh' when you receive your response email and you will see that a time stamp has been listed confirming that your job has been completed
Click on the 'Databases' tab
Click on the 'Unlock' tab
Click on 'Generate New Key (Unlock)', which will unlock the database
Click on the 'Databases' tab, select your 'Group', select your 'Database', click on '"Big* Files"', and select your BigWig File (which has the time stamp of your finished job)
Click 'View tracks in UCSC browser'
Finally, select the track you would wish to visualize

!bed1.PNG!

!bed2.PNG!

!bed3.PNG!

!bed4.PNG!

!bed5.PNG!

!bed6.PNG!

!bed7.PNG!

!bed8.PNG!

!bed9.PNG!

!bed10.PNG!

!bed11.PNG!

!bed12.PNG!

!bed13.PNG!

!bed14.PNG!

!bed15.PNG!

!bed16.PNG!

!bed17.PNG!

3.2 How do I copy / move tracks from one database to another?

Go to http://www.genboree.org
Log in if necessary
Click on the 'Workbench' tab or directly visit the Workbench http://genboree.org/java-bin/workbench.jsp
Expand the Group that contains your Database that you would like to copy tracks *TO*
Expand 'Databases'
Drag your target database from the 'Data Selector' window over into the 'Output Targets' window
Expand the Group that contains the Database of tracks that you would like to copy tracks *FROM*
Expand 'Databases'
Expand the Database containing the tracks you wish to copy
Expand 'Tracks'
Expand the 'Class' of tracks you are interested in (i.e. 'Class:High Density Score Data')
Drag the tracks that you wish to copy to the 'Input Data' window (NOTE: you also have the option to drag accros an entire database or class of tracks if you find that more convenient)
Click the 'Data' tab, followed by 'Tracks', and then select 'Copy/Move Tracks'
If you wish to copy tracks leave the 'Copy Tracks?' radio button selected. If you want to move the tracks, click the 'Move Tracks?' radio button.
Click 'Submit'

!http://genboree.org/theCommons/attachments/1613/copy_tracks_1.PNG!

!http://genboree.org/theCommons/attachments/1614/copy_tracks_2.PNG!

!http://genboree.org/theCommons/attachments/1615/copy_tracks_3.PNG!

!http://genboree.org/theCommons/attachments/1616/copy_tracks_4.PNG!

!http://genboree.org/theCommons/attachments/1617/copy_tracks_5.PNG!

!http://genboree.org/theCommons/attachments/1612/copy_tracks_6.PNG!

4. Genboree Workbench Jobs

4.1 How do I check the status of jobs that I have submitted?

Go to http://www.genboree.org
Log in if necessary
Click on the 'Workbench' tab or directly visit the Workbench http://genboree.org/java-bin/workbench.jsp
Click on 'System/Network' -> 'Jobs' -> 'Job Summary'
Select options or keep defaults
** Select specific tools (or keep all tools highlighted)
** Start Date
** End Date
** Sort Order
** Group By
** If you do not enter a Start Date and End Date, your report will contain every job you have ever launched for the selected tool(s).
# Click 'Generate Report'
** The Summary report has five columns:
*** Job Name
*** Tool
*** Submit Date
*** Completed Data
*** Status.
** This tool can be used for generating summary reports for any jobs that you have launched using the Genboree Workbench.
** You can sort each column by clicking on the column headers (i.e. 'Job Name', 'Tool', etc.) or you can choose to sort ascending or descending by clicking the the small down pointing arrow on the right border of each column header.
** You can also choose to omit certain columns by clicking on the small down arrow on the border of each column, click 'Columns', and select or deselect the column you wish to see / omit.

* This window can be left open and the information updated by clicking on the Refresh button in the upper lefthand corner.

5. Working with 27k & 450k Data

5.1 How do I prepare, upload, and evaluate 27k / 450k data? (Multi-block data format)

h1. Introduction

h2. 27k and 450k References

In order to grasp the general procedure of understanding and utilizing the 27k and 450k output we recommend some of the following manuscripts:
* 27k
** DNA methylation profiling reveals a predominant immune component in breast cancers
***
http://onlinelibrary.wiley.com/doi/10.1002/emmm.201100801/abstract;jsessionid=16E66DB216D0684E4EC0D0701F25E0E8.d03t04
* 450k
** Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome
*** http://www.landesbioscience.com/journals/epigenetics/article/16196/?nocache=54549194
** Evaluation of the Infinium Methylation 450K technology
*** http://www.futuremedicine.com/doi/full/10.2217/epi.11.105

h2. Tutorial Data Set

In order to illustrate how to use the Genboree Workbench to evaluate 27k / 450k data, we're going to demonstrate how to utilize a publicly available data set:
* Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome
** http://www.landesbioscience.com/journals/epigenetics/article/16196/?nocache=54549194

The data set that we are going to start with resides in the Supplementary section of the GEO web site:
* http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29290
** Download and decompress this file:
***
http://www.ncbi.nlm.nih.gov/geosuppl/?acc=GSE29290&file=GSE29290%5FMatrix%5FProcessed%2Etxt%2Egz
**** GSE29290_Matrix_Processed.txt.gz
***** GSE29290_Matrix_Processed.txt.zip

Data format:
* The data can be tab or comma delimited format
* One column has to represent the probe ID
** I.e. cg00000029
* One column has to represent the probe score for each sample (i.e. <sample>.AVG_Beta, <sample>.M-value, etc.)

<pre>
ID_REF Sample_1.AVG_Beta Sample_1.Detection Pval Sample_2.AVG_Beta Sample_2.Detection Pval Sample_3.AVG_Beta Sample_3.Detection Pval
cg00000029 0.8296142 0 0.852155 0 0.8956234 0 ...
cg00000108 0.8492596 0 0.8898684 0 0.9276204 0 ...
cg00000109 0.8247395 0 0.8609225 0 0.8725377 0 ...
cg00000165 0.8228635 0 0.8665444 0 0.8800115 0 ...
... ...
</pre>

h1. Preparing Processed Matrix for Genboree Workbench

We need to take the processed matrix and prepare it for import within the 'Array Data Importer' utility. Please read the help on this tool, but we will also post the necessary file format here:

Format:

<pre>
##trackName=myTrackName_1:trackSubName_1
#Probe_Name<tab>Score_Value
<ProbeID_1><tab><ProbeScore_1>
<ProbeID_2><tab><ProbeScore_2>
<ProbeID_3><tab><ProbeScore_3>
...
##trackName=myTrackName_2:trackSubName_2
#Probe_Name<tab>Score_Value
<ProbeID_1><tab><ProbeScore_1>
<ProbeID_2><tab><ProbeScore_2>
<ProbeID_3><tab><ProbeScore_3>
...
</pre>

Actual implementation of above sample (in data format):

##trackName=Sample_1.AVG_Beta:Subset_avg_beta	
#Probe	Score
cg00000029	0.8296142
cg00000108	0.8492596
cg00000109	0.8247395
cg00000165	0.8228635
##trackName=Sample_2.AVG_Beta:Subset_avg_beta	
#Probe	Score
cg00000029	0.852155
cg00000108	0.8898684
cg00000109	0.8609225
cg00000165	0.8665444

You will note the following:
* A header line for each separate sample followed by the column headers
** This will repeat every time you wish to import an additional sample
* A unique track name
* Probe IDs that exist within the ROI (region of interest) annotation track
* Numerical values for the Score data

The output of this process is provided in the following file:
* GSE29290_Matrix_Processed.txt-array_format-450k.tsv.zip

h1. Preparing Metadata for 27k & 450k Data Sets

In order to be able to utilize the Genboree Workbench to analyze your array data, it is most convenient if you produce some metadata for your samples. Providing metadata for your samples will allow you to more easily create sets of tracks (called Track Entity Lists) in order to be able to evaluate your samples in a variety of groups.

Creating Track Metadata
* This example has 2 metadata columns
** cell_type
*** Colorectal_cancer
*** Colorectal_cancer_knock_out
*** Breast_normal
*** Breast_tumor

** experiment_type
*** 450k

#name	cell_type	experiment_type
Sample_1.AVG_Beta:450k_avg_beta	Colorectal_cancer	450k
Sample_2.AVG_Beta:450k_avg_beta	Colorectal_cancer	450k
Sample_3.AVG_Beta:450k_avg_beta	Colorectal_cancer	450k
Sample_4.AVG_Beta:450k_avg_beta	Colorectal_cancer_knock_out	450k
Sample_5.AVG_Beta:450k_avg_beta	Colorectal_cancer_knock_out	450k
Sample_6.AVG_Beta:450k_avg_beta	Colorectal_cancer_knock_out	450k
Sample_7.AVG_Beta:450k_avg_beta	Breast_normal	450k
Sample_8.AVG_Beta:450k_avg_beta	Breast_normal	450k
Sample_9.AVG_Beta:450k_avg_beta	Breast_normal	450k
Sample_10.AVG_Beta:450k_avg_beta	Breast_normal	450k
Sample_11.AVG_Beta:450k_avg_beta	Breast_normal	450k
Sample_12.AVG_Beta:450k_avg_beta	Breast_normal	450k
Sample_13.AVG_Beta:450k_avg_beta	Breast_normal	450k
Sample_14.AVG_Beta:450k_avg_beta	Breast_normal	450k
Sample_15.AVG_Beta:450k_avg_beta	Breast_tumor	450k
Sample_16.AVG_Beta:450k_avg_beta	Breast_tumor	450k
Sample_17.AVG_Beta:450k_avg_beta	Breast_tumor	450k
Sample_18.AVG_Beta:450k_avg_beta	Breast_tumor	450k
Sample_19.AVG_Beta:450k_avg_beta	Breast_tumor	450k
Sample_20.AVG_Beta:450k_avg_beta	Breast_tumor	450k
Sample_21.AVG_Beta:450k_avg_beta	Breast_tumor	450k
Sample_22.AVG_Beta:450k_avg_beta	Breast_tumor	450k

File:
* GSE29290_full_track_metadata.tsv

h1. Using the Genboree Workbench to Evaluate 27k & 450k Data Sets - Step by Step

* Create a new Database
** Drag your group into the 'Output Targets' window
** Click 'Data' -> 'Databases' -> 'Create Database'
** Select 'Template: Human (hg19)
** Enter a Database Name
** Click Submit

* Create a new Project
** Drag your group into the 'Output Targets' window
** Click 'Data' -> 'Projects' -> 'Create Project'
** Enter Project Name
** Click Submit

* Upload your prepared array data ('GSE29290_Matrix_Processed.txt-array_format-450k.tsv.zip')
** (Remove your Group from the 'Output Targets' window
** Drag your Database into the 'Output Targets' window
** Click 'Data' -> 'Files' -> 'Transfer File'
** Choose your file
** Click Submit

* Import your array data

** Drag your Database into the 'Output Targets' window
** Drag your file into the 'Input Data' window
*** This file is located in your_group -> Databases -> your_database -> Files
** Click 'Data' -> 'Tracks' -> 'Import' -> 'Array Data'
** Select 'Hs Methylation:450k'
*** You would select 'Hs Methylation:27k' if you are using 27k data
** Select File Format 'Muti-block'
** Click Submit
** Wait for success email

<pre>
Hello Kevin Riehle,

Your Array Data Importer job has completed successfully.

JOB SUMMARY:
  JobID   : wbJob-arraydataimporter-1347059333_775361

The following array/probe file has been imported:
 GSE29290_Matrix_Processed.txt-array_format-450k.tsv

The following tracks were uploaded in the target database:

Sample_19.AVG_Beta:450k_avg_beta
Sample_2.AVG_Beta:450k_avg_beta
Sample_1.AVG_Beta:450k_avg_beta
Sample_3.AVG_Beta:450k_avg_beta
Sample_15.AVG_Beta:450k_avg_beta
Sample_17.AVG_Beta:450k_avg_beta
Sample_16.AVG_Beta:450k_avg_beta
Sample_14.AVG_Beta:450k_avg_beta
Sample_12.AVG_Beta:450k_avg_beta
Sample_8.AVG_Beta:450k_avg_beta
Sample_9.AVG_Beta:450k_avg_beta
Sample_7.AVG_Beta:450k_avg_beta
Sample_13.AVG_Beta:450k_avg_beta
Sample_21.AVG_Beta:450k_avg_beta
Sample_22.AVG_Beta:450k_avg_beta
Sample_5.AVG_Beta:450k_avg_beta
Sample_11.AVG_Beta:450k_avg_beta
Sample_6.AVG_Beta:450k_avg_beta
Sample_10.AVG_Beta:450k_avg_beta
Sample_4.AVG_Beta:450k_avg_beta
Sample_18.AVG_Beta:450k_avg_beta
Sample_20.AVG_Beta:450k_avg_beta

The Genboree Team
...
</pre>

Add Track Metadata
* Upload track metadata file

** GSE29290_track_metadata.tsv
* Drag your File (i.e. 'GSE29290_full_track_metadata.tsv') to 'Input Data' window
* Drag your Database to 'Output Targets' window
* Click 'Data' -> 'Tracks' -> 'Import' -> 'Track Metadata'
** Uncheck 'Create New Tracks?'
** Click Submit

Quickly Create Track Entity Lists
* Drag your Database into 'Input Data'
* Click 'Visualization' -> 'View Track Grid'
** X-axis attribute
*** cell_type
** Y-axis attribute
*** experiment_type
** Click Submit
** Click the blue hyperlink 'Launch Grid Viewer'

Create Track Entity Lists - All Samples
* Select the (8) Breast_normal cell, the (8) Breast_tumor cell, the (3) Colorectal_cancer cell, and the (3) Colorectal_cancer_knock_out cell.
* Click 'Selections' -> 'Save Selections'
** Select your Group
** Select your Database
** Type in a name (i.e. 'all22samplesTrackEntityList_A')
** Click 'Save Selections'
* Select the (8) Breast_tumor cell
* Click 'Selections' -> 'Save Selections'
** Select your Group
** Select your Database
** Type in a name (i.e. ''all22samplesTrackEntityList_B')
** Click 'Save Selections'

You Can Also Create Track Entity Lists Based on Metadata Labels
* Select the (8) Breast_normal cell
* Click 'Selections' -> 'Save Selections'
** Select your Group
** Select your Database
** Type in a name (i.e. 'Breast_normal_450k')
** Click 'Save Selections'
* Select the (8) Breast_tumor cell
* Click 'Selections' -> 'Save Selections'
** Select your Group
** Select your Database
** Type in a name (i.e. 'Breast_tumor_450k')
** Click 'Save Selections'

h3. Heatmap

* Run Heatmap on All Tracks
** (Clear any entries in the 'Input Data' window if they exist)
** Drag your two track entity lists into the 'Input Data' window (i.e. 22samplesTrackEntityList_A and 22samplesTrackEntityList_B)
*** These track entity lists are located in your_group -> Databases -> your_database -> Lists & Selections -> List of

Tracks
** Drag your desired ROI (regions of interest) track into the 'Input Data' window
*** For example, you can use the Promoters:LCP ROI track
**** This track is located in ROI Repository -> Databases -> ROI Repository - hg19 -> Tracks -> Class: Regulation
** Drag your Database into the 'Output Targets' window
** Drag your Project into the 'Output Targets' window
** Click 'Epigenome' -> 'Compute Similarity Matrix (heatmap)'
** Click Submit
** Wait for a confirmation email

<pre>

Hello Kevin Riehle,

Your Epigenomic Experiment Heatmap Tool job completed successfully.

Job Summary:
JobID - wbJob-epigenomicsheatmap-1347635780_213632
Analysis Name - EpigenomeExpHeatmap2012-09-14-10:15:54
Inputs:
1. Entitylist - 450k_1
2. Entitylist - 450k_2
3. Trk - Promoters%3ALCP
Outputs:
1. Db - 450ktest_full
2. Prj - 450k
Settings:
analysisName - EpigenomeExpHeatmap2012-09-14-10:15:54
clusterQueue - gbMultiCore
color - Spectral
dendograms - both
density - histogram
distfun - dist
fixedResolution - medium
hclustfun - hclust
height - 8
key - TRUE
keySize - 0.75
quantileNormalized - true
removeNoDataRegions - true
spanAggFunction - avg
trace - none
width - 10

- The Genboree Team

Result File Location in the Genboree Workbench:
http://www.genboree.org/java-bin/project.jsp?projectName=450k

</pre>

!matrix.txt.fixed.heatmap-matrix_data-scaled.PNG!

Heatmap results:
* You will see that we witness clustering among:
** 8 Breast_normal
** 7 Breast_tumor (Sample_20 is an outlier)
** 3 Colorectal_cancer
** 3 Colorectal_cancer_knock_out

h3. LIMMA

Run LIMMA
* Drag your first track entity list into the 'Input Data' window (i.e. 'Breast_normal_450k')
* Drag your second track entity list into the 'Input Data' window (i.e. 'Breast_tumor_450k')
* Drag your ROI (regions of interest) track into the 'Input Data' window (i.e. 'Promoters:ALL)
** This track is located in ROI Repository -> Databases -> ROI Repository - hg19 -> Tracks -> Class: Regulation
* Drag your Database into the 'Output Targets' window
* Drag your Project into the 'Output Targets' window
* Click 'Epigenome' -> 'Analyze Signals' -> 'Compare by LIMMA' -> 'Tracks'
** Click Submit
** Wait for confirmation emails:
*** "Genboree: Your Epigenomic Experiment Sets Comparison Using Limma job is complete"
*** "LFF API Upload [SUCCESS]"

h3. SPARK

Run SPARK
* Drag your first track entity list into the 'Input Data' window (i.e. 'Breast_normal_450k')
* Drag your second track entity list into the 'Input Data' window (i.e. 'Breast_tumor_450k')
* Drag your ROI (regions of interest) track into the 'Input Data' window (i.e. 'Promoters:ALL)
** This track is located in ROI Repository -> Databases -> ROI Repository - hg19 -> Tracks -> Class: Regulation
* Drag your Database into the 'Output Targets' window
* Click 'Epigenome' -> 'Analyze Signals' -> 'Cluster by Spark'
** Select your ROI Track
*** Single click on your ROI track (i.e. Promoters:LCP)
** Customize the settings or leave the defaults
** Optionally change track colors
*** I.e. change samples 15-22 to 'green' for Data Track Colors
** Click Submit
** Wait for confirmation email and follow directions

5.2 How do I prepare, upload, and evaluate 27k / 450k data? (Multi-column (i.e. matrix) data format)
h1. Introduction

h2. 27k and 450k References

In order to grasp the general procedure of understanding and utilizing the 27k and 450k output we recommend some of the following manuscripts:
* 27k
** DNA methylation profiling reveals a predominant immune component in breast cancers
***
http://onlinelibrary.wiley.com/doi/10.1002/emmm.201100801/abstract;jsessionid=16E66DB216D0684E4EC0D0701F25E0E8.d03t04
* 450k

** Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome
*** http://www.landesbioscience.com/journals/epigenetics/article/16196/?nocache=54549194
** Evaluation of the Infinium Methylation 450K technology
*** http://www.futuremedicine.com/doi/full/10.2217/epi.11.105

h2. Tutorial Data Set

In order to illustrate how to use the Genboree Workbench to evaluate 27k / 450k data, we're going to demonstrate how to utilize a publicly available data set:
* Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome
** http://www.landesbioscience.com/journals/epigenetics/article/16196/?nocache=54549194

The data set that we are going to start with resides in the Supplementary section of the GEO web site:
* http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29290
** Download and decompress this file:
***
http://www.ncbi.nlm.nih.gov/geosuppl/?acc=GSE29290&file=GSE29290%5FMatrix%5FProcessed%2Etxt%2Egz
**** GSE29290_Matrix_Processed.txt.gz
***** GSE29290_Matrix_Processed.txt.zip

Data format:
* The data is tab delimited
* One column has to represent the probe ID
** I.e. cg00000029
* One column has to represent the probe score for each sample (i.e. <sample>.AVG_Beta, <sample>.M-value, etc.)

<pre>
ID_REF Sample_1.AVG_Beta Sample_1.Detection Pval Sample_2.AVG_Beta Sample_2.Detection Pval Sample_3.AVG_Beta Sample_3.Detection Pval
cg00000029 0.8296142 0 0.852155 0 0.8956234 0 ...
cg00000108 0.8492596 0 0.8898684 0 0.9276204 0 ...
cg00000109 0.8247395 0 0.8609225 0 0.8725377 0 ...
cg00000165 0.8228635 0 0.8665444 0 0.8800115 0 ...
... ...
</pre>

h1. Preparing Processed Matrix for Genboree Workbench

We need to take the processed matrix and prepare it for import within the 'Array Data Importer' utility. Please read the help on this tool, but we will also post the necessary file format here:

Format:
<pre>
#probe<tab><sample1_name><tab><sample2_name><tab><sample3_name>
<ProbeID_1><tab><ProbeScore_1-sample1><tab><ProbeScore_1-sample2><tab><ProbeScore_1-sample3>
<ProbeID_2><tab><ProbeScore_2-sample1><tab><ProbeScore_2-sample2><tab><ProbeScore_2-sample3>
<ProbeID_3><tab><ProbeScore_3-sample1><tab><ProbeScore_3-sample2><tab><ProbeScore_3-sample3>
...
</pre>

Actual implementation of above sample (in multi-column matrix data format):

ID_REF	Sample_1.AVG_Beta	Sample_2.AVG_Beta	Sample_3.AVG_Beta
cg00000029	0.8296142	0.852155	0.8956234
cg00000108	0.8492596	0.8898684	0.9276204
cg00000109	0.8247395	0.8609225	0.8725377
cg00000165	0.8228635	0.8665444	0.8800115

You will note the following:
* A *single* score column for each sample
** All other columns must be removed prior to uploading and importing
* A unique sample name (which will be used to name the track)
* Probe IDs that exist within the ROI (region of interest) annotation track
** Probe IDs that do not exist within the 27K / 450K ROI track will be ignored
* Numerical values for the Score data
* Tab delimited

The output of this process is provided in the following file:
* GSE29290_Matrix_Processed-AVG_Beta.tsv.zip

h1. Preparing Metadata for 27k & 450k Data Sets

In order to be able to utilize the Genboree Workbench to analyze your array data, it is most convenient if you produce some metadata for your samples. Providing metadata for your samples will allow you to more easily create sets of tracks (called Track Entity Lists) in order to be able to evaluate your samples in a variety of groups.

Creating Track Metadata
* This example has 2 metadata columns
** cell_type
*** Colorectal_cancer
*** Colorectal_cancer_knock_out
*** Breast_normal
*** Breast_tumor
** experiment_type
*** 450k

#name	cell_type	experiment_type
Sample_1.AVG_Beta:450K	Colorectal_cancer	450k
Sample_2.AVG_Beta:450K	Colorectal_cancer	450k
Sample_3.AVG_Beta:450K	Colorectal_cancer	450k
Sample_4.AVG_Beta:450K	Colorectal_cancer_knock_out	450k
Sample_5.AVG_Beta:450K	Colorectal_cancer_knock_out	450k
Sample_6.AVG_Beta:450K	Colorectal_cancer_knock_out	450k
Sample_7.AVG_Beta:450K	Breast_normal	450k
Sample_8.AVG_Beta:450K	Breast_normal	450k
Sample_9.AVG_Beta:450K	Breast_normal	450k
Sample_10.AVG_Beta:450K	Breast_normal	450k
Sample_11.AVG_Beta:450K	Breast_normal	450k
Sample_12.AVG_Beta:450K	Breast_normal	450k
Sample_13.AVG_Beta:450K	Breast_normal	450k
Sample_14.AVG_Beta:450K	Breast_normal	450k
Sample_15.AVG_Beta:450K	Breast_tumor	450k
Sample_16.AVG_Beta:450K	Breast_tumor	450k
Sample_17.AVG_Beta:450K	Breast_tumor	450k
Sample_18.AVG_Beta:450K	Breast_tumor	450k

Sample_19.AVG_Beta:450K	Breast_tumor	450k
Sample_20.AVG_Beta:450K	Breast_tumor	450k
Sample_21.AVG_Beta:450K	Breast_tumor	450k
Sample_22.AVG_Beta:450K	Breast_tumor	450k

File:
* GSE29290_full_track_metadata-matrix-format.tsv

h1. Using the Genboree Workbench to Evaluate 27k & 450k Data Sets - Step by Step

* Create a new Database
** Drag your group into the 'Output Targets' window
** Click 'Data' -> 'Databases' -> 'Create Database'
** Select 'Template: Human (hg19)
** Enter a Database Name
** Click Submit

* Create a new Project
** Drag your group into the 'Output Targets' window
** Click 'Data' -> 'Projects' -> 'Create Project'
** Enter Project Name
** Click Submit

* Upload your prepared array data ('GSE29290_Matrix_Processed-AVG_Beta.tsv.zip')
** Remove your Group from the 'Output Targets' window
** Drag your Database into the 'Output Targets' window
** Click 'Data' -> 'Files' -> 'Transfer File'
** Choose your file
** Click Submit

* Import your array data
** Drag your Database into the 'Output Targets' window
** Drag your file ('GSE29290_Matrix_Processed-AVG_Beta.tsv.zip') into the 'Input Data' window
*** This file is located in your_group -> Databases -> your_database -> Files
** Click 'Data' -> 'Tracks' -> 'Import' -> 'Array Data'
** Select 'Hs Methylation:450k'
*** You would select 'Hs Methylation:27k' if you are using 27k data
** Select File Format 'Muti-column' (default)
** Click Submit
** Wait for success email

<pre>
Hello Kevin Riehle,

Your Array Data Importer job has completed successfully.

JOB SUMMARY:
  JobID   : wbJob-arraydataimporter-1347059333_775361

The following array/probe file has been imported:
 GSE29290_Matrix_Processed.txt-array_format-450k.tsv

The following tracks were uploaded in the target database:

Sample_19.AVG_Beta:450k_avg_beta
Sample_2.AVG_Beta:450k_avg_beta
Sample_1.AVG_Beta:450k_avg_beta
Sample_3.AVG_Beta:450k_avg_beta
Sample_15.AVG_Beta:450k_avg_beta
Sample_17.AVG_Beta:450k_avg_beta
Sample_16.AVG_Beta:450k_avg_beta
Sample_14.AVG_Beta:450k_avg_beta
Sample_12.AVG_Beta:450k_avg_beta
Sample_8.AVG_Beta:450k_avg_beta
Sample_9.AVG_Beta:450k_avg_beta
Sample_7.AVG_Beta:450k_avg_beta
Sample_13.AVG_Beta:450k_avg_beta
Sample_21.AVG_Beta:450k_avg_beta
Sample_22.AVG_Beta:450k_avg_beta
Sample_5.AVG_Beta:450k_avg_beta
Sample_11.AVG_Beta:450k_avg_beta
Sample_6.AVG_Beta:450k_avg_beta
Sample_10.AVG_Beta:450k_avg_beta
Sample_4.AVG_Beta:450k_avg_beta
Sample_18.AVG_Beta:450k_avg_beta
Sample_20.AVG_Beta:450k_avg_beta

The Genboree Team
...
</pre>

Add Track Metadata
* Upload track metadata file
** GSE29290_full_track_metadata-matrix-format.tsv
* Drag your File (i.e. 'GSE29290_full_track_metadata-matrix-format.tsv') to 'Input Data' window
* Drag your Database to 'Output Targets' window
* Click 'Data' -> 'Tracks' -> 'Import' -> 'Track Metadata'
** Uncheck 'Create New Tracks?'
** Click Submit

Quickly Create Track Entity Lists
* Drag your Database into 'Input Data'
* Click 'Visualization' -> 'View Track Grid'
** X-axis attribute
*** cell_type
** Y-axis attribute
*** experiment_type
** Click Submit
** Click the blue hyperlink 'Launch Grid Viewer'

Create Track Entity Lists - All Samples
* Select the (8) Breast_normal cell, the (8) Breast_tumor cell, the (3) Colorectal_cancer cell, and the (3) Colorectal_cancer_knock_out cell.
* Click 'Selections' -> 'Save Selections'
** Select your Group

** Select your Database
** Type in a name (i.e. 'all22samplesTrackEntityList')
** Click 'Save Selections'

You Can Also Create Track Entity Lists Based on Metadata Labels
* Select the (8) Breast_normal cell
* Click 'Selections' -> 'Save Selections'
** Select your Group
** Select your Database
** Type in a name (i.e. 'Breast_normal_450k')
** Click 'Save Selections'
* Select the (8) Breast_tumor cell (and deselect the (8) Breast_normal_cell cell if it is still highlighted)
* Click 'Selections' -> 'Save Selections'
** Select your Group
** Select your Database
** Type in a name (i.e. 'Breast_tumor_450k')
** Click 'Save Selections'

h3. Heatmap

* Run Heatmap on All Tracks
** (Clear any entries in the 'Input Data' window if they exist)
** Drag your track entity list(s) into the 'Input Data' window (i.e. 22samplesTrackEntityList)
*** These track entity lists are located in your_group -> Databases -> your_database -> Lists & Selections -> List of Tracks
** Drag your desired ROI (regions of interest) track into the 'Input Data' window
*** For example, you can use the Promoters:LCP ROI track
**** This track is located in ROI Repository -> Databases -> ROI Repository - hg19 -> Tracks -> Class: Regulation
** Drag your Database into the 'Output Targets' window
** Drag your Project into the 'Output Targets' window
** Click 'Epigenome' -> 'Compute Similarity Matrix (heatmap)'
** Click Submit
** Wait for a confirmation email

<pre>
Hello Kevin Riehle,

Your job completed successfully.

Job Summary:
JobID - wbJob-epigenomicsHeatmap-pCvyYR-4758
Analysis Name - all22_self_tutorial-EpigenomeExpHeatmap2013-02-27-11:07:03
Inputs:
1. Entitylist - all22samplesTrackEntityList
2. Trk - Promoters%3ALCP
3. Entitylist - all22samplesTrackEntityList
Outputs:
1. Db - 450k_tutorial_matrix
2. Prj - 450k_tutorial_matrix_project
Settings:
analysisName - all22_self_tutorial-EpigenomeExpHeatmap2013-02-27-11:07:03
color - Spectral
dendograms - both

density - histogram
distfun - dist
hclustfun - hclust
height - 8
key - TRUE
keySize - 0.75
normalization - quant
quantileNormalized - false
removeNoDataRegions - true
spanAggFunction - avg
trace - none
width - 10

- The Genboree Team

Result File Location in the Genboree Workbench:
http://genboree.org/java-bin/project.jsp?projectName=450k_tutorial_matrix_project

</pre>

!matrix.txt.fixed.heatmap-matrix_data-scaled.PNG!

Heatmap results:
* You will see that we witness clustering among:
** 8 Breast_normal
** 7 Breast_tumor (Sample_20 is an outlier)
** 3 Colorectal_cancer
** 3 Colorectal_cancer_knock_out

h3. LIMMA

Run LIMMA
* Drag your first track entity list into the 'Input Data' window (i.e. 'Breast_normal_450k')
* Drag your second track entity list into the 'Input Data' window (i.e. 'Breast_tumor_450k')
* Drag your ROI (regions of interest) track into the 'Input Data' window (i.e. 'Promoters:ALL)
** This track is located in ROI Repository -> Databases -> ROI Repository - hg19 -> Tracks -> Class: Regulation
* Drag your Database into the 'Output Targets' window
* Drag your Project into the 'Output Targets' window
* Click 'Epigenome' -> 'Analyze Signals' -> 'Compare by LIMMA' -> 'Tracks'
** Click Submit
** Wait for confirmation emails:
*** "Genboree: Your Epigenomic Experiment Sets Comparison Using Limma job is complete"
*** "LFF API Upload [SUCCESS]"

h3. SPARK

Run SPARK
* Drag your first track entity list into the 'Input Data' window (i.e. 'Breast_normal_450k')
* Drag your second track entity list into the 'Input Data' window (i.e. 'Breast_tumor_450k')
* Drag your ROI (regions of interest) track into the 'Input Data' window (i.e. 'Promoters:ALL)
** This track is located in ROI Repository -> Databases -> ROI Repository - hg19 -> Tracks -> Class: Regulation
* Drag your Database into the 'Output Targets' window

* Click 'Epigenome' -> 'Analyze Signals' -> 'Cluster by Spark'
** Select your ROI Track
*** Single click on your ROI track (i.e. Promoters:LCP)
** Customize the settings or leave the defaults
** Optionally change track colors
*** I.e. change samples 15-22 to 'green' for Data Track Colors
** Click Submit
** Wait for confirmation email and follow directions

6. Genboree Commons

6.1 6a. What is Genboree Commons?

Genboree Commons is a collaboration site featuring discussion forums, wikis, and document sharing. You will need a Genboree username and password to sign in. You can register with Genboree at http://www.genboree.org/java-bin/login.jsp .

6.2 6b. How do I create a project in Genboree Commons?

Click on *Projects* (upper left of page), then on *New Project* (upper right of page). Fill in all the fields. A description of some of the fields is shown below, under *General Settings*. The Project Identifier is used internally by Redmine (for URLs and other things). Once created, you will not need to use the Project Identifier again.

The following *General Settings* are available:

â€¢ Name: project display name (must be unique).
â€¢ Subproject of: lets you define a parent project to the project being created. Projects can be unlimitedly nested.
â€¢ Description: description that appears on the project overview.
â€¢ Identifier: used by the application for various things (eg. in URLs). It must be unique. Once the project is created, it identifier cannot be modified.
â€¢ Homepage: homepage-link that appears on the project overview.
â€¢ Public: if checked, the project can be viewed by all the users, including those who are not members of the project. If unchecked, only the project members have access to it, according to their role.

6.3 6c. How do I create a forum in Genboree Commons?

You must be an administrator of a project to create a forum, and a project can have more than one administrator. An *administrator* of a project defines a list of forums in the *Project Settings* (authors and subscribers can not create forums). The forums allow users in a project to communicate with each other by creating Conversations (i.e. threads) within the forums.

Each project can have one or more discussion forums. Each forum has the following properties:
* Name: The text you want to be displayed to identify the discussion forum. This field is required.
* Description: A short description to describe the subject of the specific forum. This field is required.

To add a topic in a given forum, click on the forum name, then click on *New Conversation*. You can now enter a subject, a body and attach files to your message.

Two options are available (this is usually left as default. That is, you don't need to select anything here):
* sticky: if checked, the topic will stay displayed at the top of topic list, in bold
* locked: if checked, users can not add replies to the message

6.4 6d. How do I change my email notification settings in Genboree Commons?

You manage your email notifications under *My account* --> *Email notifications*. The default setting is *For any event on all my projects....* With this setting, you will receive email notifications about any event that occurs on the

projects you belong to (posts to forums, new document uploaded, etc).

You can change this setting so that you receive fewer emails, or emails for just specific projects, by selecting the appropriate setting via the pull-down menu.

You may also set your preference so that you do not receive any emails by selecting *Suppress all notifications, even for things I watch or am involved in..*..

Click on *Save* after you have set your desired preference. Your email notification settings can be changed as often as you wish.
