American Society of Human Genetics
Boston, MA

October 22, 2013

Presented by the
Bioinformatics Research Laboratory

Baylor College of Medicine

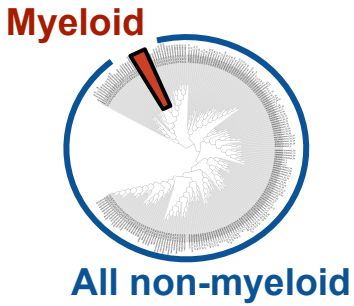
Purpose of the supplemental slides are to find differentially modified histone regions for myeloid lineages

Use Case 18

Clustering/Heatmap: select experimental tracks from the Human Epigenome Atlas to find myeloid cell lineage consisting of CD14 and CD15 cell types

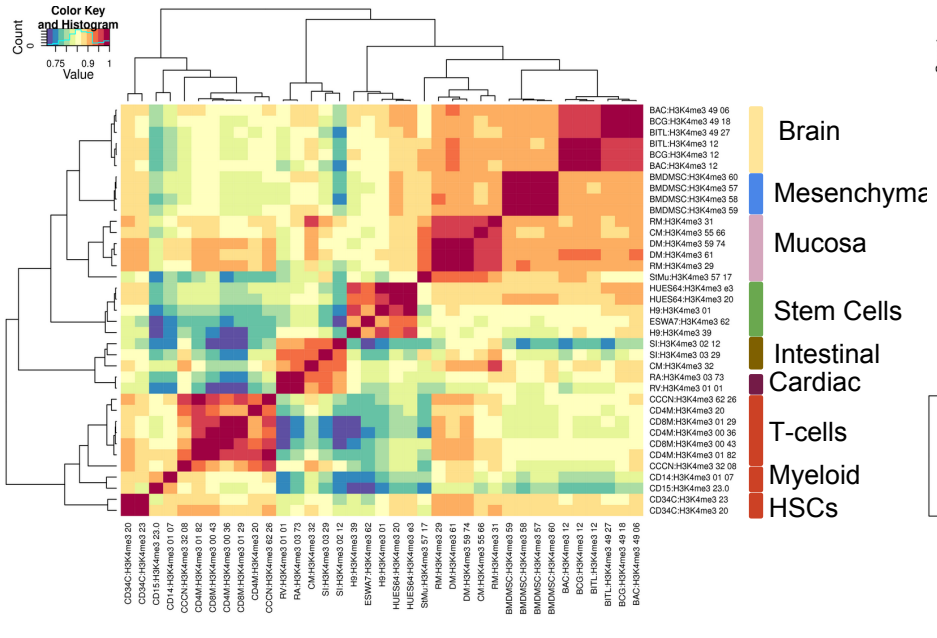
Use Case 19 Supplemental Slides

LIMMA: to find enhancer regions with differentially modified histone signals between two groups of data tracks - Myeloid vs non-myeloid

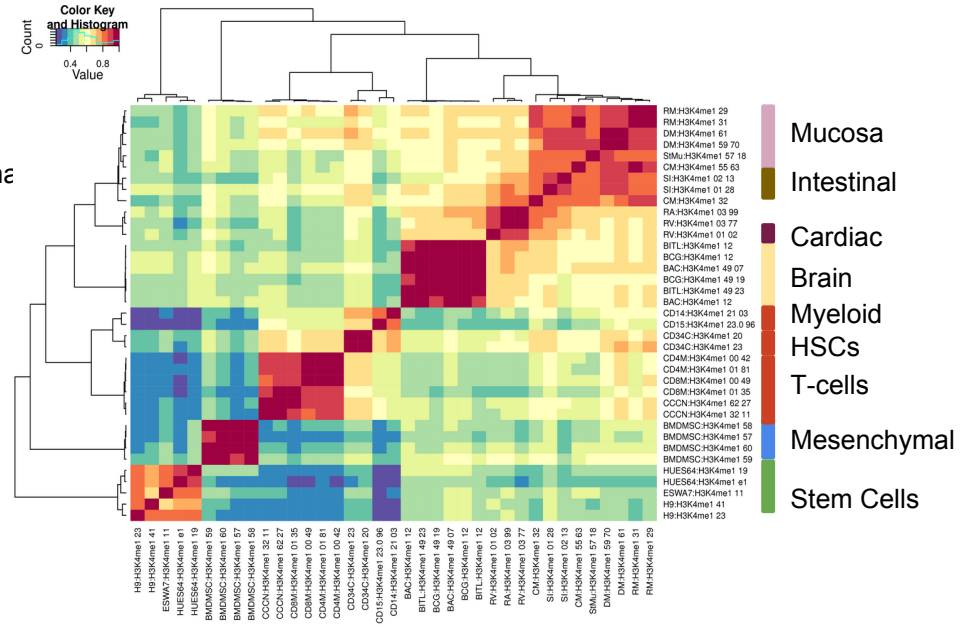


In Use Case 18 we show clustering of samples in myeloid lineage

H3K4me3 signal over protein coding gene promoters on the NIH Roadmap Epigenome data



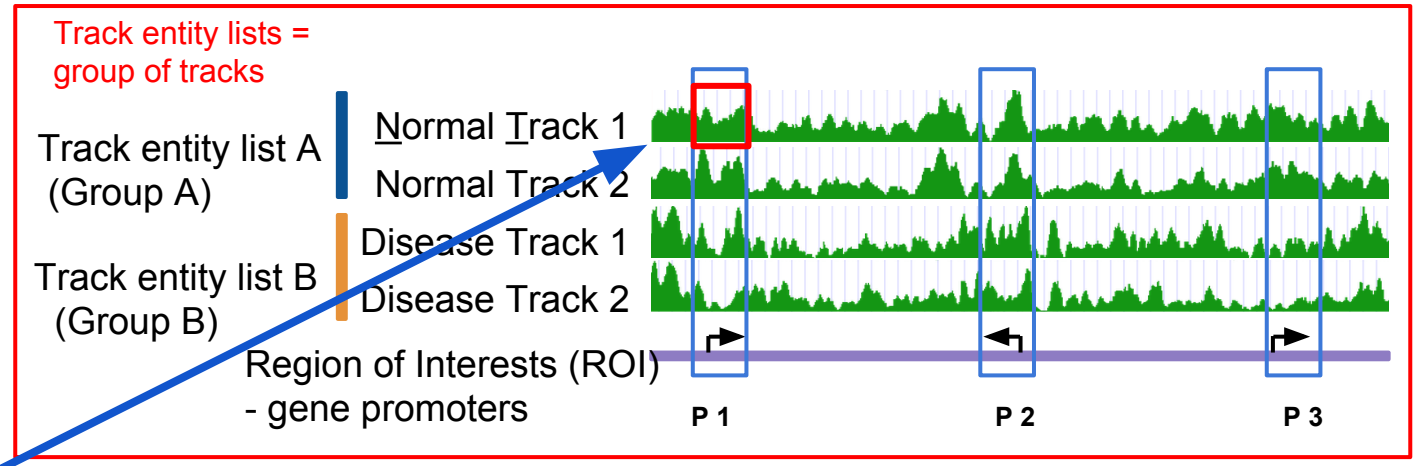
H3K4me1 signal over lincRNA gene promoters on the NIH Roadmap Epigenome data



To find differentially modified histone regions in myeloid vs non-myeloid samples, we can perform LIMMA analysis between these two groups

Methodology: LIMMA (Linear Model for Microarray Analysis)

I. Data Selection



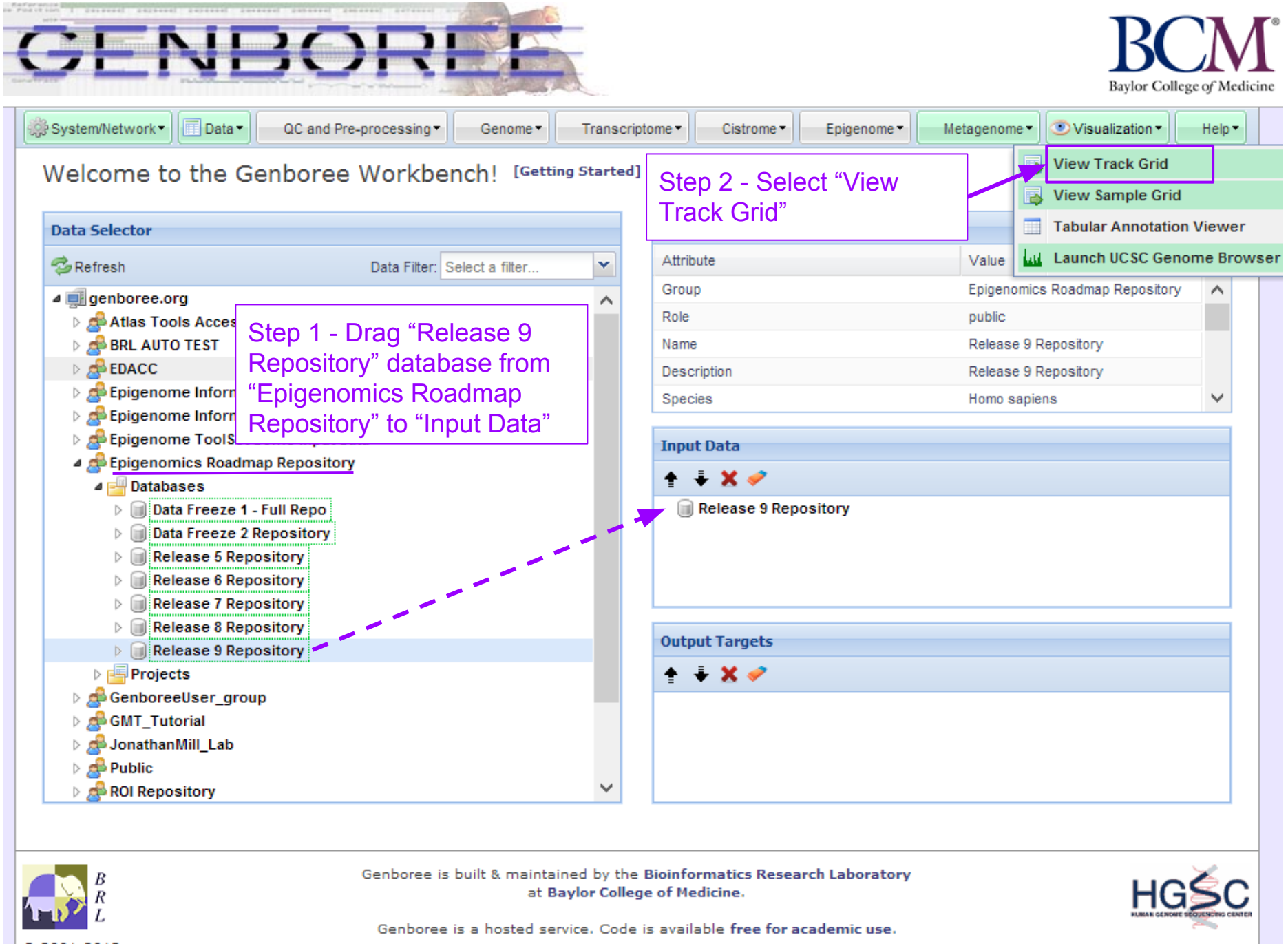
II. Signal processing- Tool calculates average signal for each ROI and each track

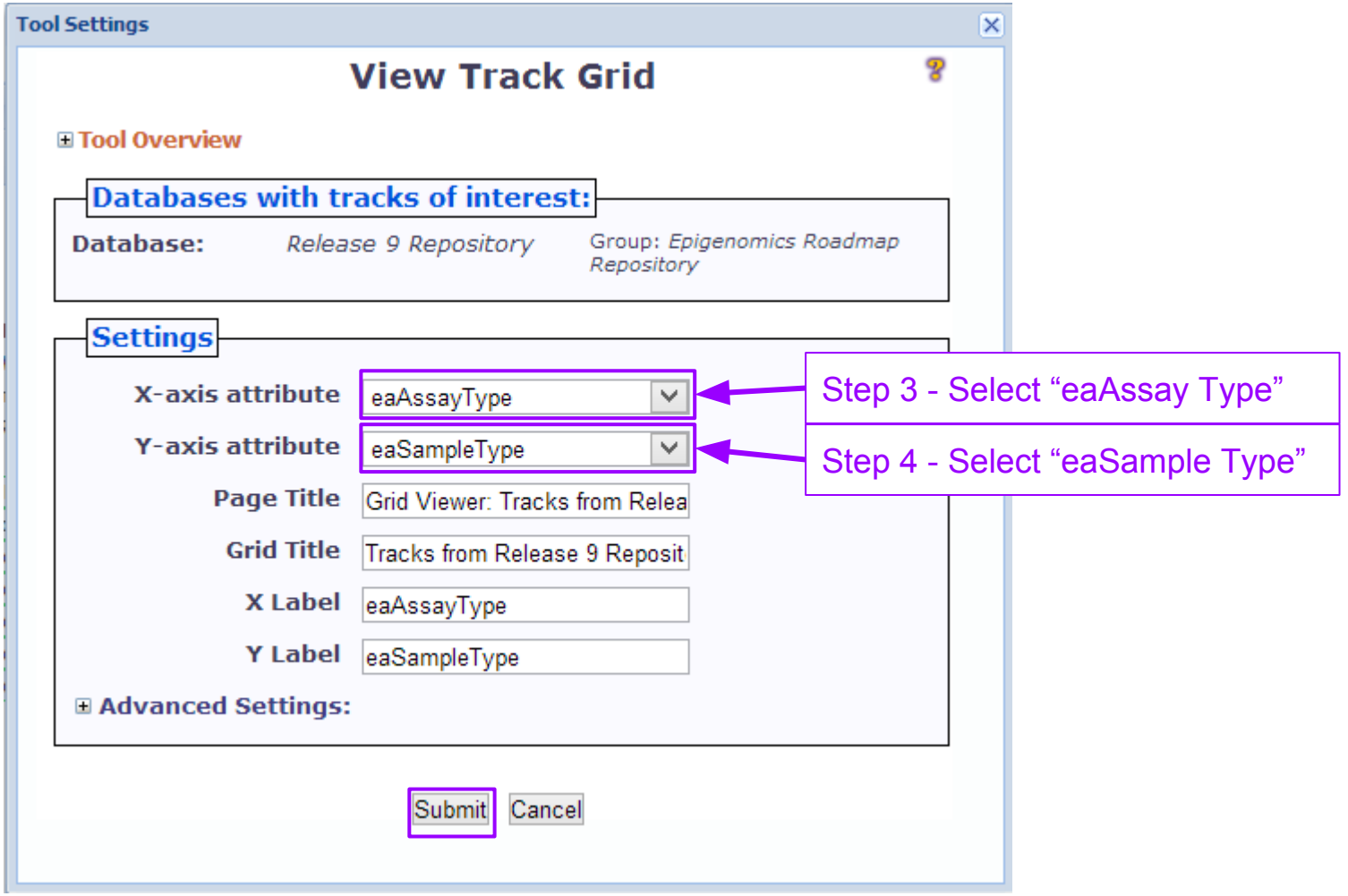
	Group A		Group B	
	NT1	NT 2	DT1	DT2
P 1	0.8	0.7	0.3	0.2
P 2	0.65	0.6	0.7	0.5
P 3	0.8	0.8	0.2	0.15

Data are normalized and LIMMA tool compares average signal for each ROI (row) between two groups. Note: for limma to work, need at least two tracks in each group.

III. Results output

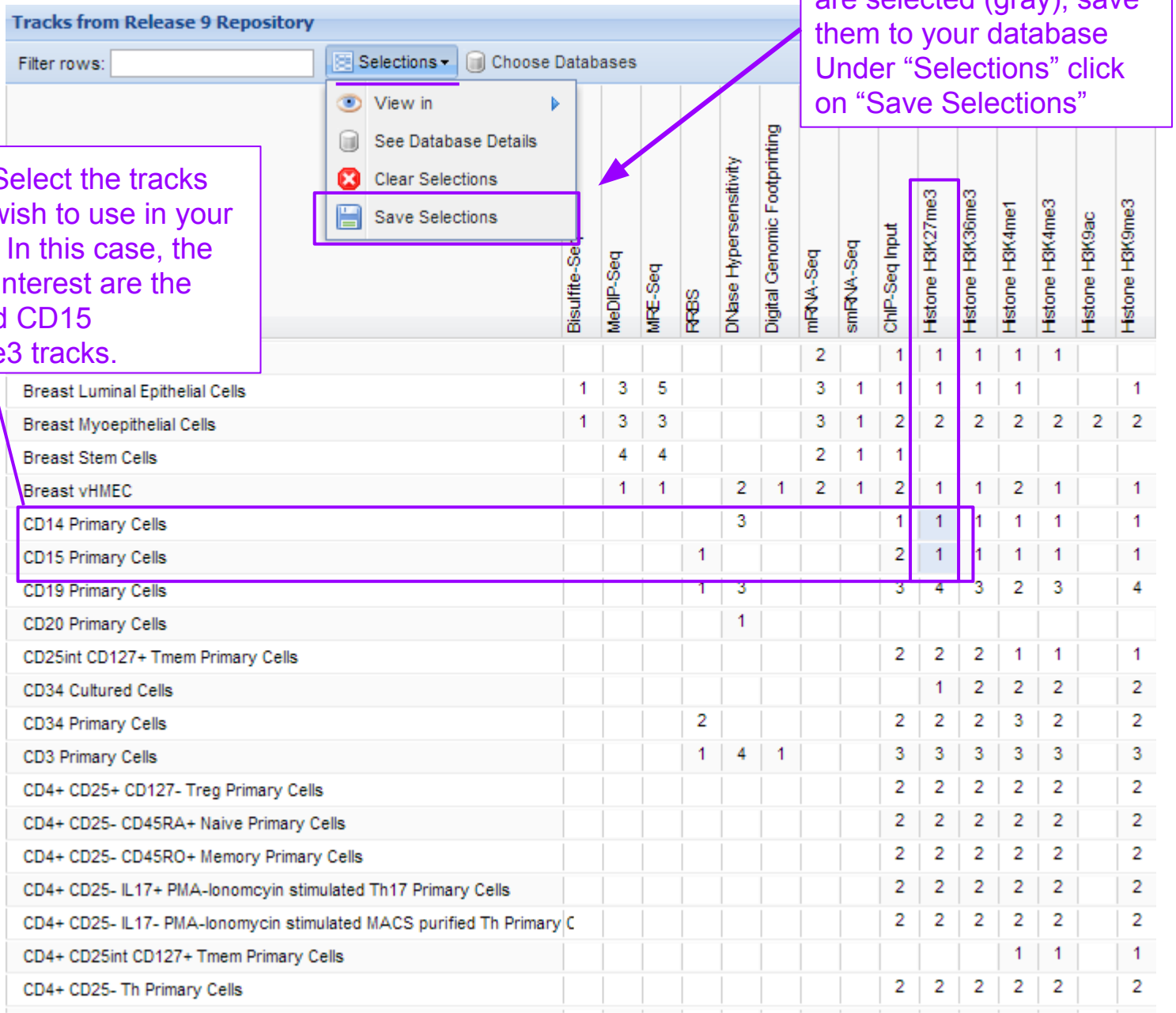
ROIs (i.e. promoter, enhancers, etc) that exhibit significantly different signal between the groups compared, are provided to users as a region track which can be downloaded or used for downstream analysis


GENBOREE
BCM
Baylor College of Medicine
System/Network
Data
QC and Pre-processing
Genome
Transcriptome
Cistrome
Epigenome
Metagenome
Visualization
Help
Welcome to the Genboree Workbench!
[Getting Started]
Step 2 - Select "View Track Grid"
View Track Grid
View Sample Grid
Tabular Annotation Viewer
Launch UCSC Genome Browser
Data Selector
Refresh
Data Filter:
Select a filter...
genboree.org
Atlas Tools Acces
BRL AUTO TEST
EDACC
Epigenome Inforn
Epigenome Inforn
Epigenome ToolS
Epigenomics Roadmap Repository
Databases
Data Freeze 1 - Full Repo
Data Freeze 2 Repository
Release 5 Repository
Release 6 Repository
Release 7 Repository
Release 8 Repository
Release 9 Repository
Projects
GenboreeUser_group
GMT_Tutorial
JonathanMill_Lab
Public
ROI Repository
Step 1 - Drag "Release 9 Repository" database from "Epigenomics Roadmap Repository" to "Input Data"
Attribute
Value
Group
Epigenomics Roadmap Repository
Role
public
Name
Release 9 Repository
Description
Release 9 Repository
Species
Homo sapiens
Input Data
Release 9 Repository
Output Targets
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

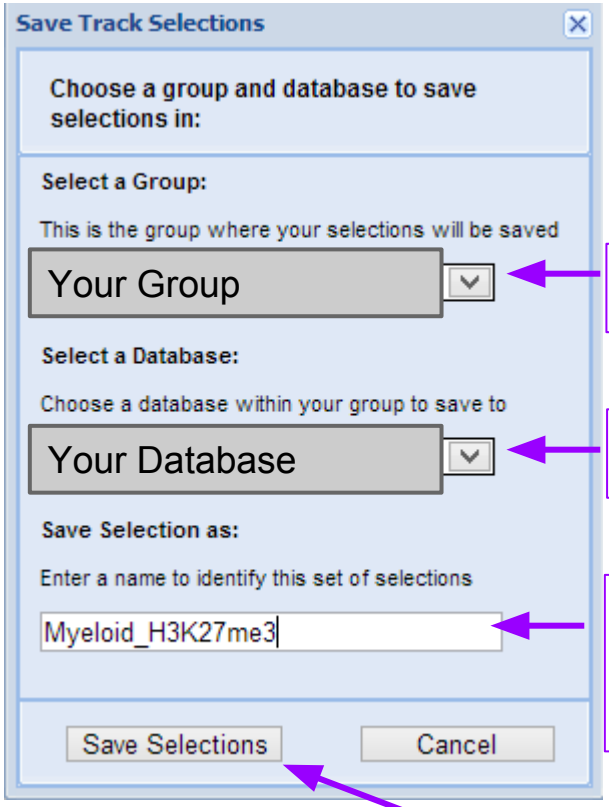

Tool Settings
View Track Grid
Tool Overview
Databases with tracks of interest:
Database:
Release 9 Repository
Group: Epigenomics Roadmap Repository
Settings
X-axis attribute
eaAssayType
Y-axis attribute
eaSampleType
Page Title
Grid Viewer: Tracks from Relea
Grid Title
Tracks from Release 9 Reposit
X Label
eaAssayType
Y Label
eaSampleType
Advanced Settings:
Submit
Cancel
Step 3 - Select “eaAssay Type”
Step 4 - Select “eaSample Type”

Step 6 - Once the tracks are selected (gray), save them to your database Under “Selections” click on “Save Selections”

Step 5 - Select the tracks that you wish to use in your analysis. In this case, the tracks of interest are the CD14 and CD15 H3K27me3 tracks.



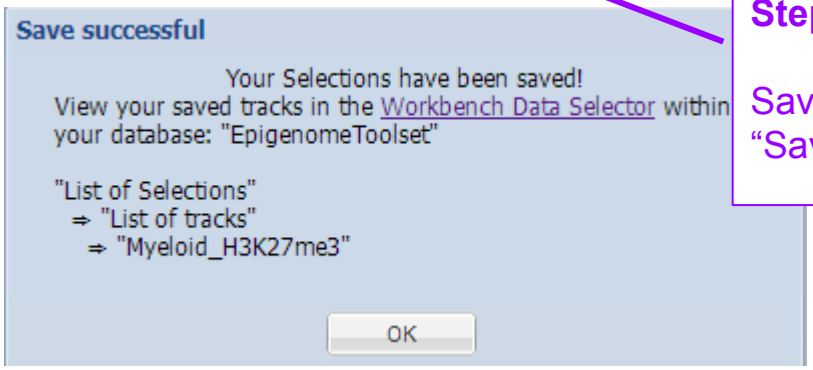
	Bisulfite-Seq	MeDIP-Seq	MRE-Seq	RRBS	DNase Hypersensitivity	Digital Genomic Footprinting	mRNA-Seq	smRNA-Seq	ChIP-Seq Input	Histone H3K27me3	Histone H3K36me3	Histone H3K4me1	Histone H3K4me3	Histone H3K9ac	Histone H3K9me3
							2		1	1	1	1	1		
Breast Luminal Epithelial Cells	1	3	5				3	1	1	1	1	1			1
Breast Myoepithelial Cells	1	3	3				3	1	2	2	2	2	2	2	2
Breast Stem Cells		4	4				2	1	1						
Breast vHMEC		1	1		2	1	2	1	2	1	1	2	1		1
CD14 Primary Cells					3				1	1	1	1	1		1
CD15 Primary Cells				1					2	1	1	1	1		1
CD19 Primary Cells				1	3				3	4	3	2	3		4
CD20 Primary Cells					1										
CD25int CD127+ Tmem Primary Cells									2	2	2	1	1		1
CD34 Cultured Cells										1	2	2	2		2
CD34 Primary Cells				2					2	2	2	3	2		2
CD3 Primary Cells				1	4	1			3	3	3	3	3		3
CD4+ CD25+ CD127- Treg Primary Cells									2	2	2	2	2		2
CD4+ CD25- CD45RA+ Naive Primary Cells									2	2	2	2	2		2
CD4+ CD25- CD45RO+ Memory Primary Cells									2	2	2	2	2		2
CD4+ CD25- IL17+ PMA-Ionomcyin stimulated Th17 Primary Cells									2	2	2	2	2		2
CD4+ CD25- IL17- PMA-Ionomycin stimulated MACS purified Th Primary C									2	2	2	2	2		2
CD4+ CD25int CD127+ Tmem Primary Cells												1	1		1
CD4+ CD25- Th Primary Cells									2	2	2	2	2		2



Step 7 - Select your Group

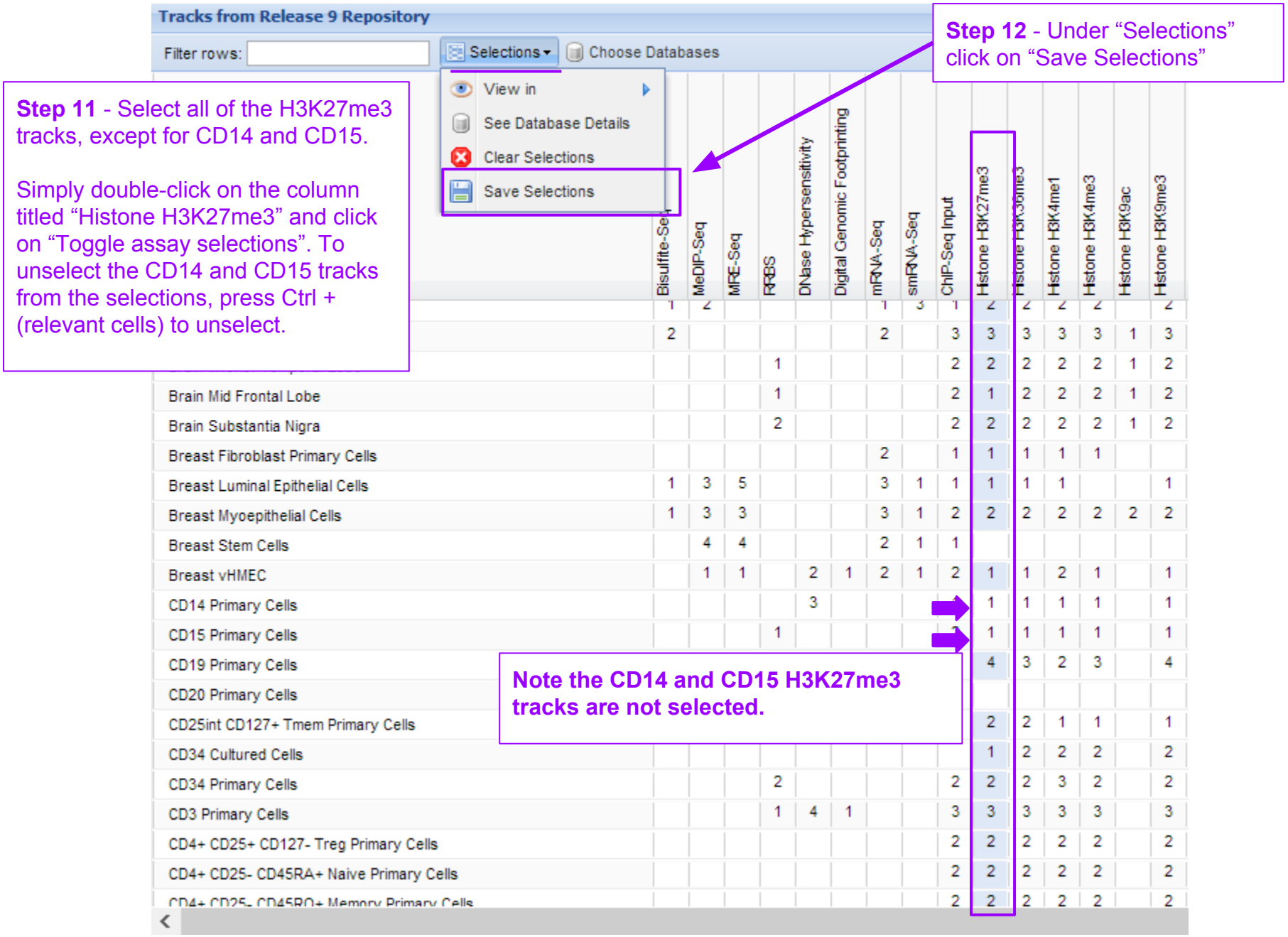
Step 8 - Select your Database

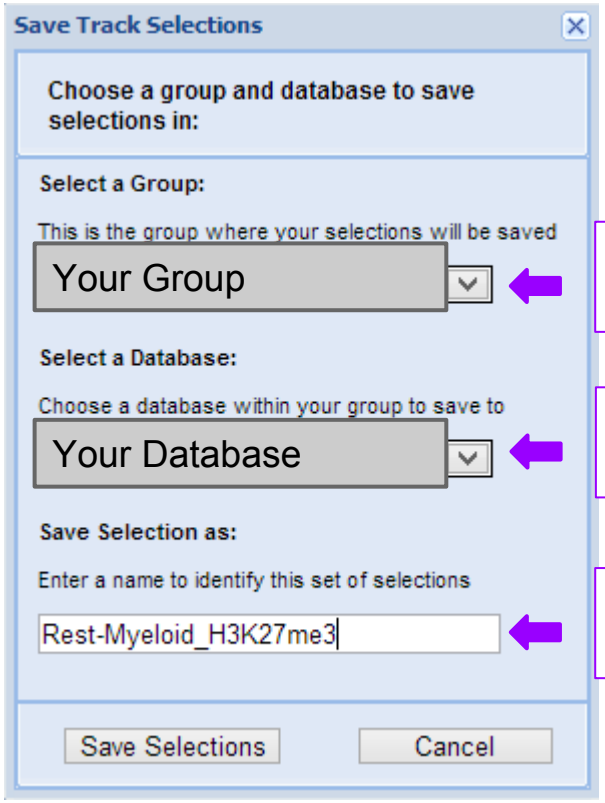
Step 9 - Name your selection something that is meaningful to you. We use “Myeloid_H3K27me3”



Step 10 - Save selection.

Saving the selection will generate a “Save successful” message.

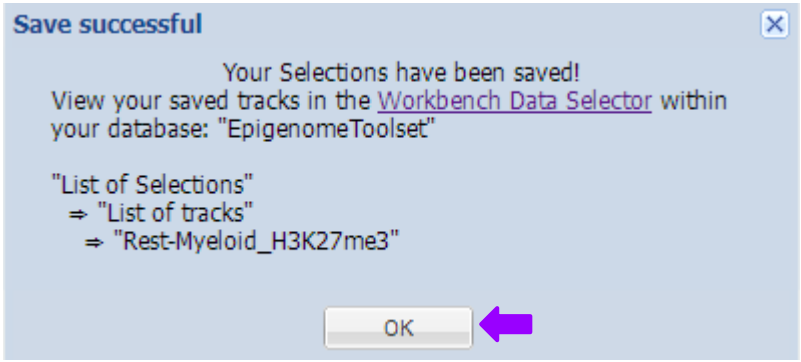

Tracks from Release 9 Repository
Filter rows:
Selections
Choose Databases
View in
See Database Details
Clear Selections
Save Selections
Bisulfite-Seq
MeDIP-Seq
MRE-Seq
RRBS
DNase Hypersensitivity
Digital Genomic Footprinting
mRNA-Seq
smRNA-Seq
ChIP-Seq Input
Histone H3K27me3
Histone H3K36me3
Histone H3K4me1
Histone H3K4me3
Histone H3K9ac
Histone H3K9me3
Brain Mid Frontal Lobe
Brain Substantia Nigra
Breast Fibroblast Primary Cells
Breast Luminal Epithelial Cells
Breast Myoepithelial Cells
Breast Stem Cells
Breast vHMEC
CD14 Primary Cells
CD15 Primary Cells
CD19 Primary Cells
CD20 Primary Cells
CD25int CD127+ Tmem Primary Cells
CD34 Cultured Cells
CD34 Primary Cells
CD3 Primary Cells
CD4+ CD25+ CD127- Treg Primary Cells
CD4+ CD25- CD45RA+ Naive Primary Cells
CD4+ CD25- CD45RO+ Memory Primary Cells
Step 11 - Select all of the H3K27me3 tracks, except for CD14 and CD15.
Simply double-click on the column titled “Histone H3K27me3” and click on “Toggle assay selections”. To unselect the CD14 and CD15 tracks from the selections, press Ctrl + (relevant cells) to unselect.
Step 12 - Under “Selections” click on “Save Selections”
Note the CD14 and CD15 H3K27me3 tracks are not selected.

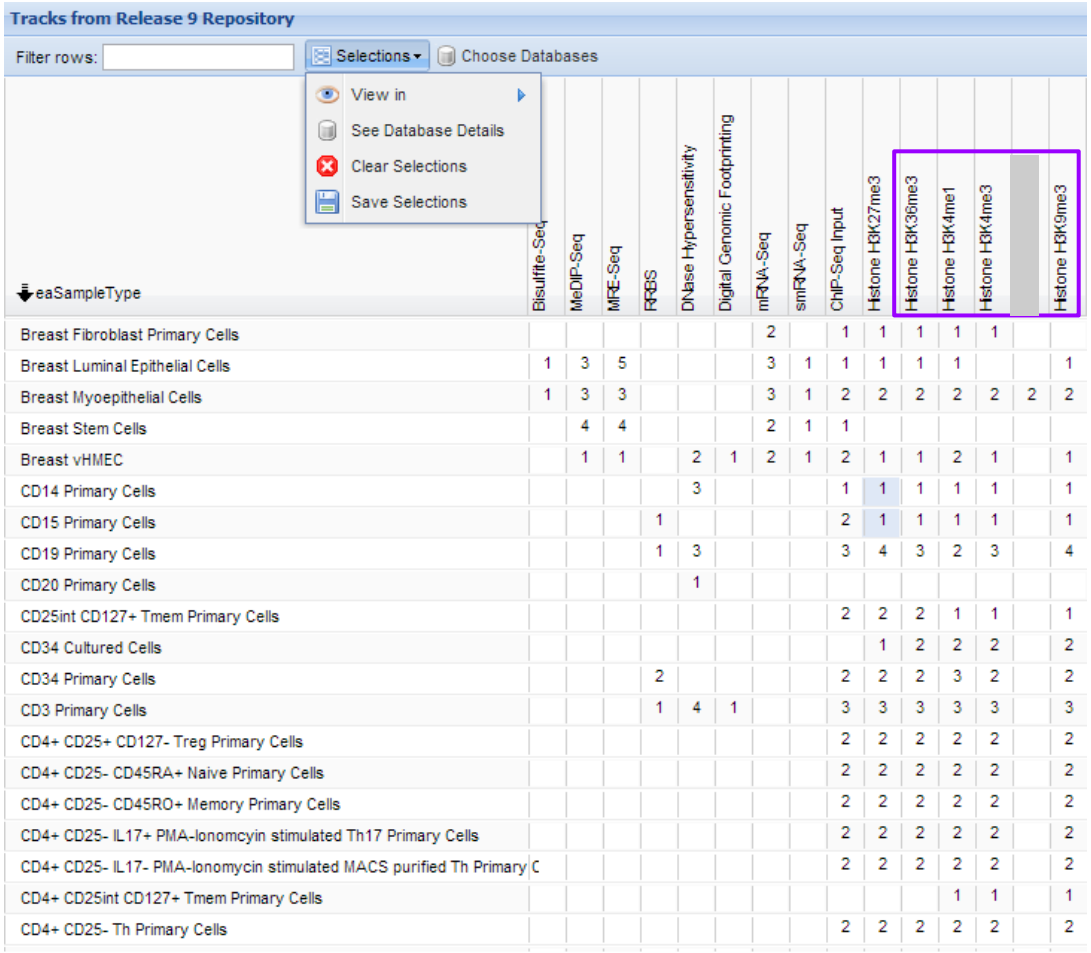


Step 13 - Select your Group

Step 14 - Select your Database

Step 15 - Save Selection as "Rest-Myeloid_H3K27me3"





Similarly, you must now generate a track entity list for each of the other four histone marks, since they will be part of separate analyses:

- Myeloid_H3K4me1 & Rest-Myeloid_H3K4me1
- Myeloid_H3K4me3 & Rest-Myeloid _H3K4me3
- Myeloid_H3K9me3 & Rest-Myeloid_H3K9me3
- Myeloid_H3K36me3 & Rest-Myeloid_H3K36me3

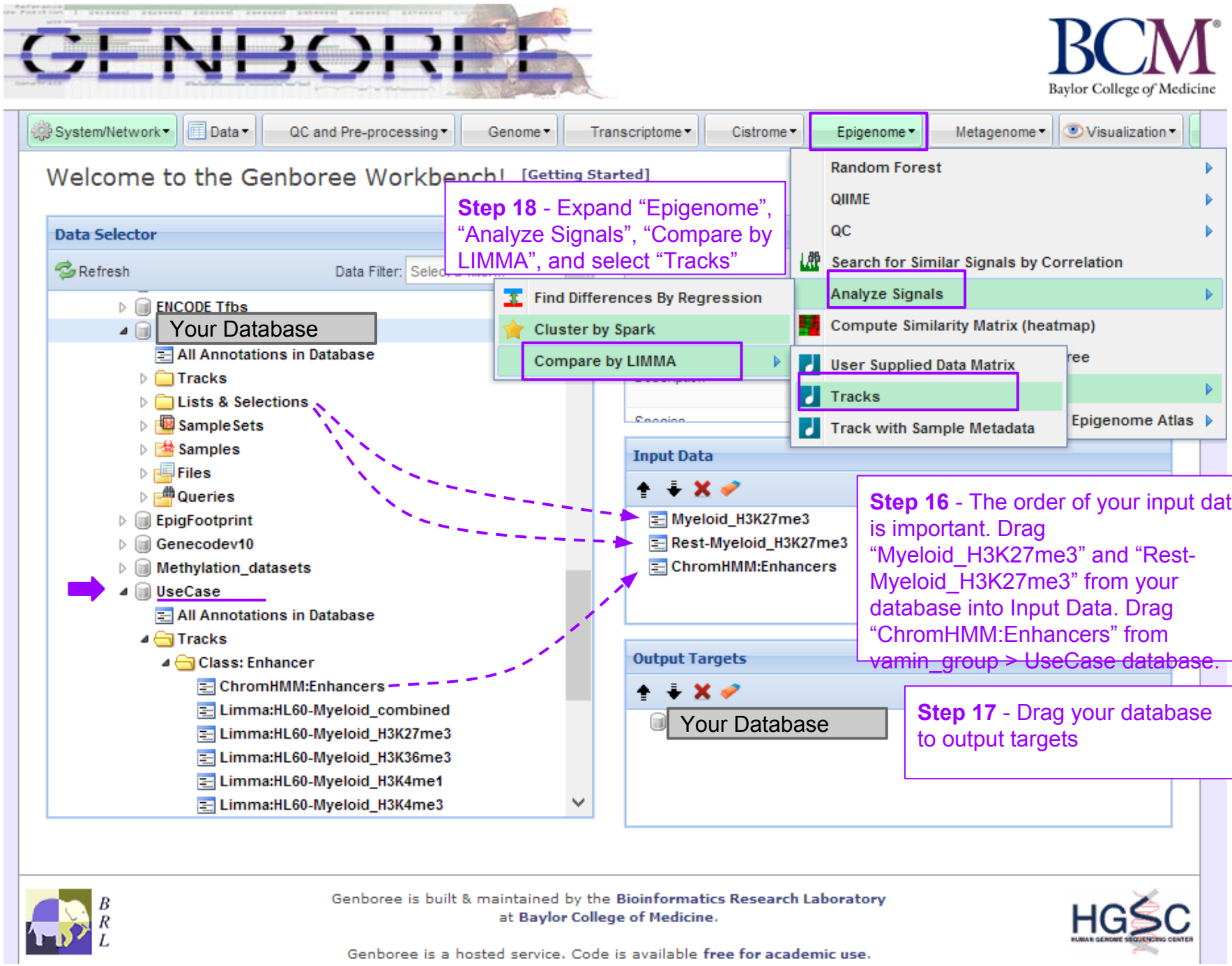
Checkpoint

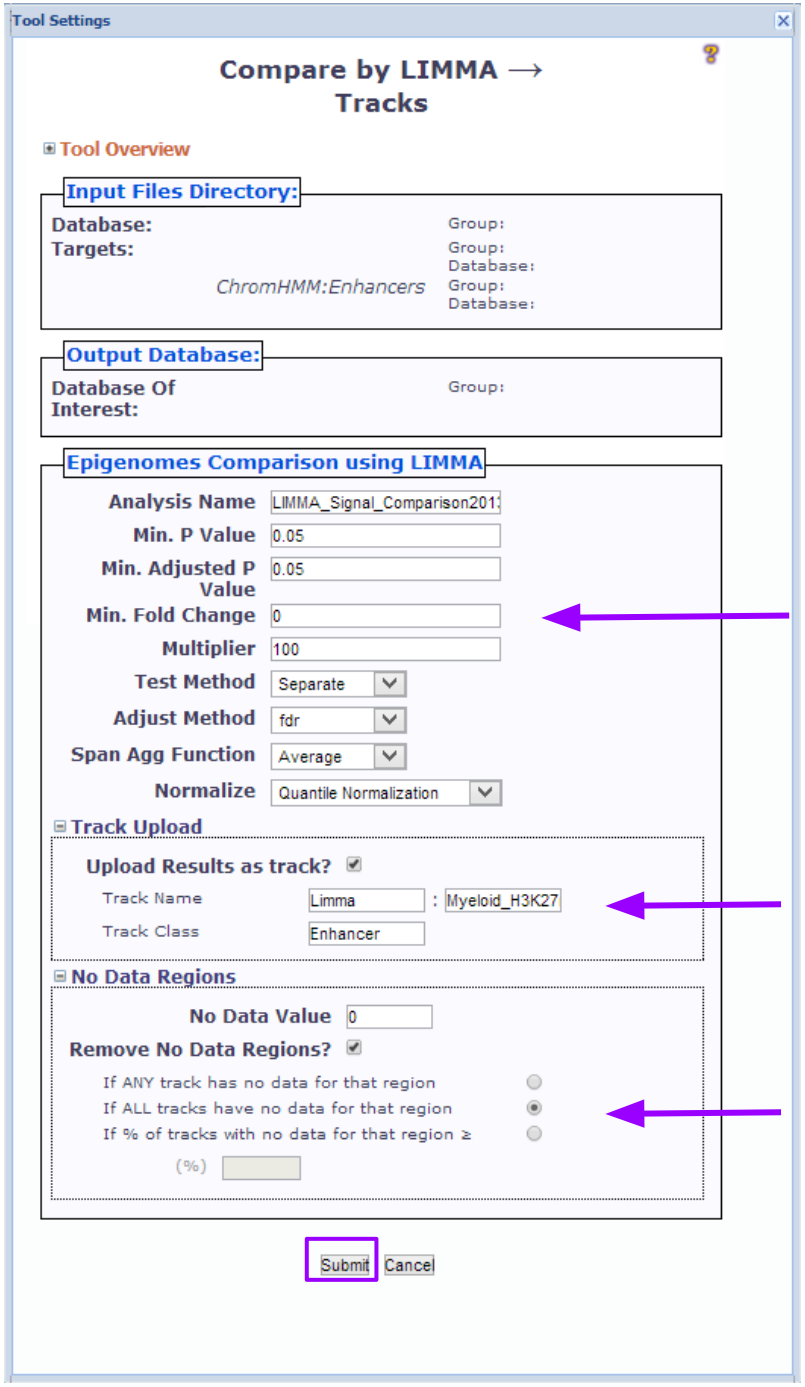
- Myeloid_H3K27me3
- Myeloid_H3K4me1
- Myeloid_H3K4me3
- Myeloid_H3K9me3
- Myeloid_H3K36me3

- Rest-Myeloid_H3K27me3
- Rest-Myeloid_H3K4me1
- Rest-Myeloid _H3K4me3
- Rest-Myeloid_H3K9me3
- Rest-Myeloid_H3K36me3

Once you have completed all the steps you should see following 10 track entity lists in your database under “Lists & Selections” > “List of Tracks”

Again, the purpose of making these track entity lists = group of tracks, so that we can do LIMMA analysis between the groups to find differentially histone modified enhancer regions.


GENBOREE
BCM
Baylor College of Medicine
System/Network
Data
QC and Pre-processing
Genome
Transcriptome
Cistrome
Epigenome
Metagenome
Visualization
Welcome to the Genboree Workbench! [Getting Started]
Step 18 - Expand "Epigenome", "Analyze Signals", "Compare by LIMMA", and select "Tracks"
Random Forest
QIIME
QC
Search for Similar Signals by Correlation
Analyze Signals
Compute Similarity Matrix (heatmap)
Find Differences By Regression
Cluster by Spark
Compare by LIMMA
User Supplied Data Matrix
Tracks
Track with Sample Metadata
Epigenome Atlas
Data Selector
Refresh
Data Filter:
ENCODE Tfbs
Your Database
All Annotations in Database
Tracks
Lists & Selections
SampleSets
Samples
Files
Queries
EpigFootprint
Genecodev10
Methylation_datasets
UseCase
All Annotations in Database
Tracks
Class: Enhancer
ChromHMM:Enhancers
Limma:HL60-Myeloid_combined
Limma:HL60-Myeloid_H3K27me3
Limma:HL60-Myeloid_H3K36me3
Limma:HL60-Myeloid_H3K4me1
Limma:HL60-Myeloid_H3K4me3
Input Data
Myeloid_H3K27me3
Rest-Myeloid_H3K27me3
ChromHMM:Enhancers
Step 16 - The order of your input data is important. Drag "Myeloid_H3K27me3" and "Rest-Myeloid_H3K27me3" from your database into Input Data. Drag "ChromHMM:Enhancers" from vamin_group > UseCase database.
Output Targets
Your Database
Step 17 - Drag your database to output targets
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

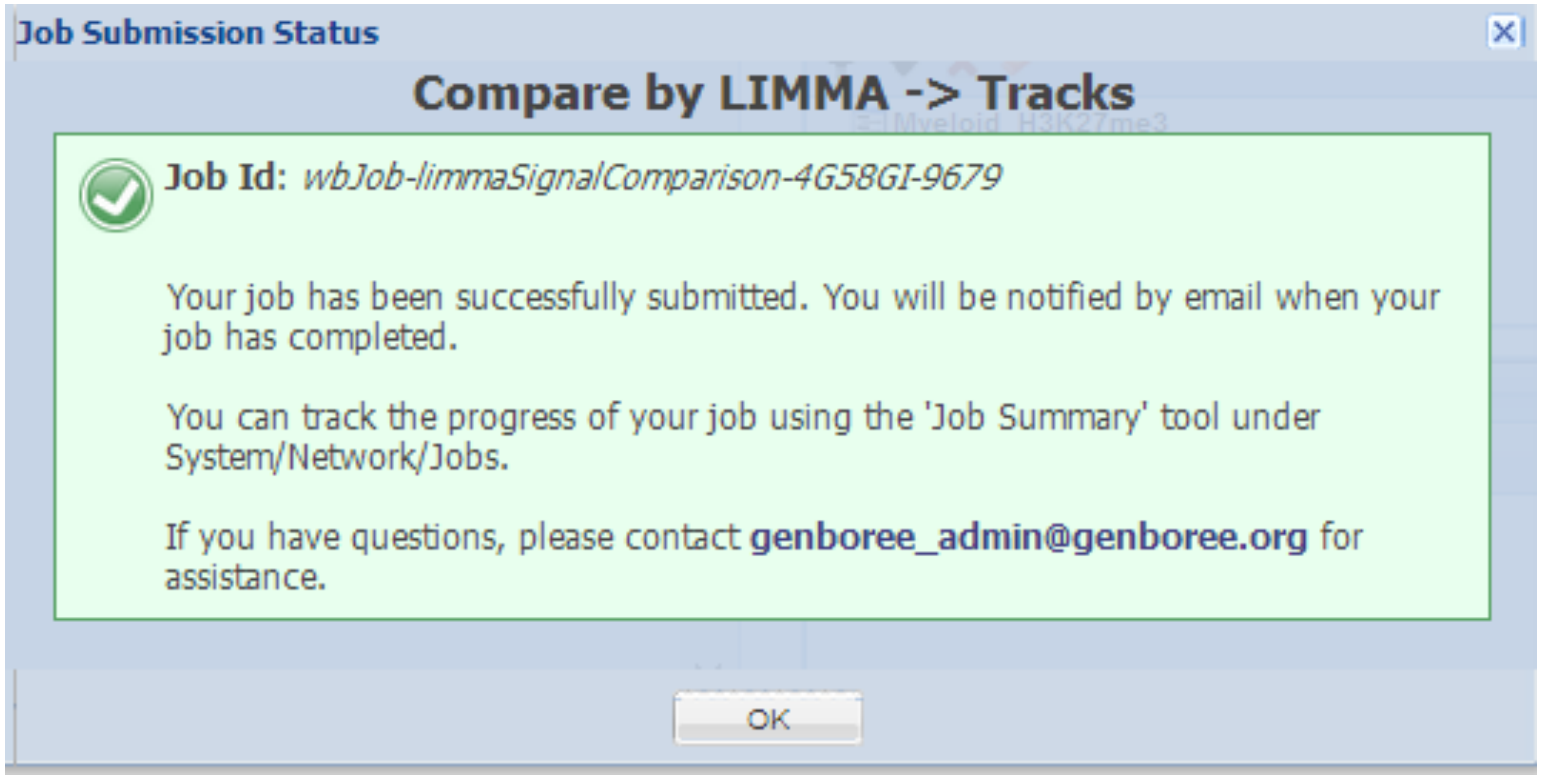


Step 19 - Set Min. Fold Change to 0

Step 20 - Select “Upload Results as track” and give track name “Limma:Myeloid_H3K27me3” and Class “Enhancer”

Step 21 - Select “If ALL tracks have no data for that region”

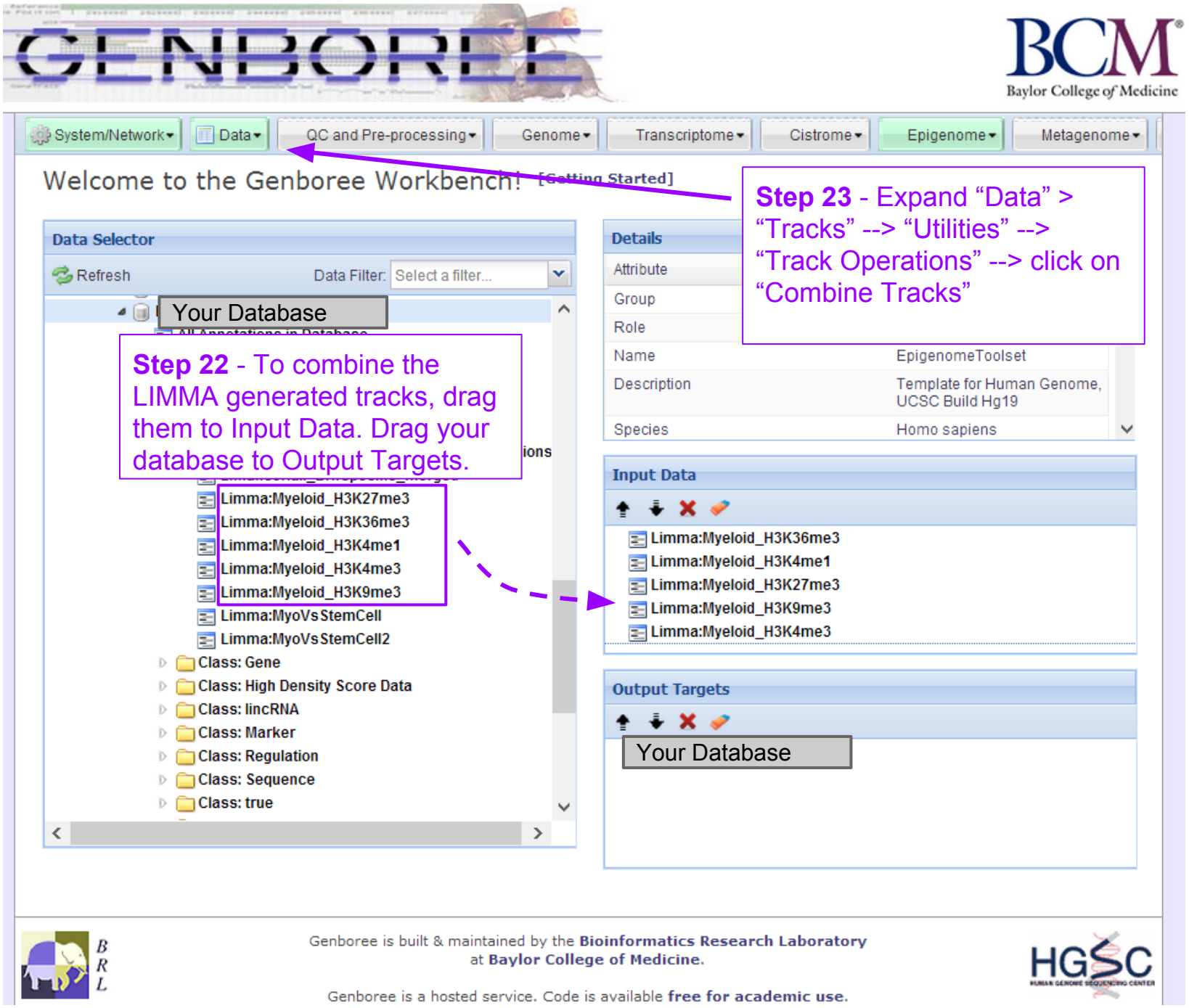
You will see the message below upon successful submission of the LIMMA job

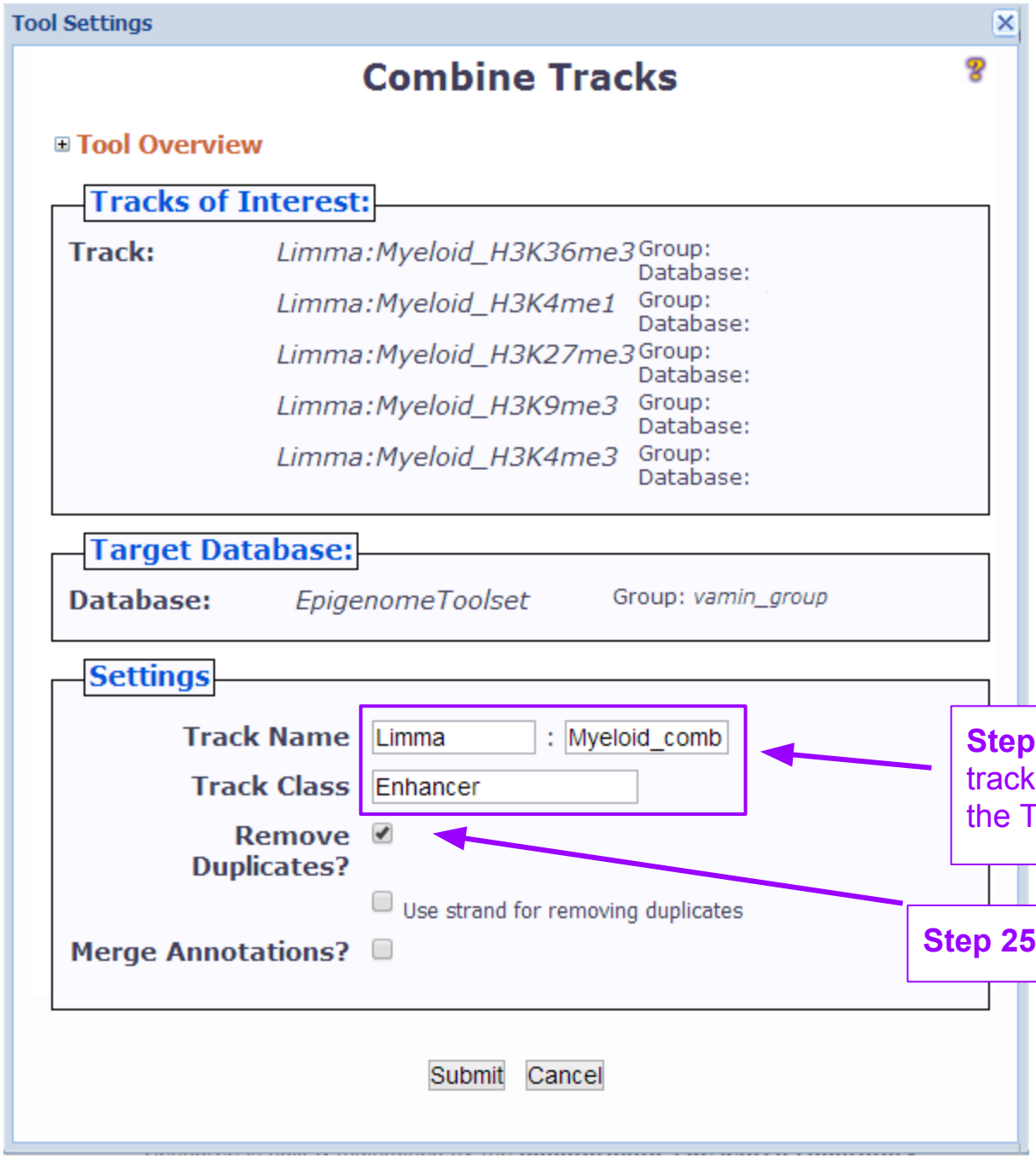


Submit Limma jobs for the other comparisons

- Submitted limma job for Myeloid_H3K27me3 vs Rest-Myeloid_H3K27me3

- Now submit limma jobs for other comparisons:
  - Myeloid_H3K4me1 vs Rest-Myeloid_H3K4me1
  - Myeloid_H3K4me3 vs Rest-Myeloid_H3K4me3
  - Myeloid_H3K9me3 vs Rest-Myeloid_H3K9me3
  - Myeloid_H3K36me3 vs Rest-Myeloid_H3K36me3

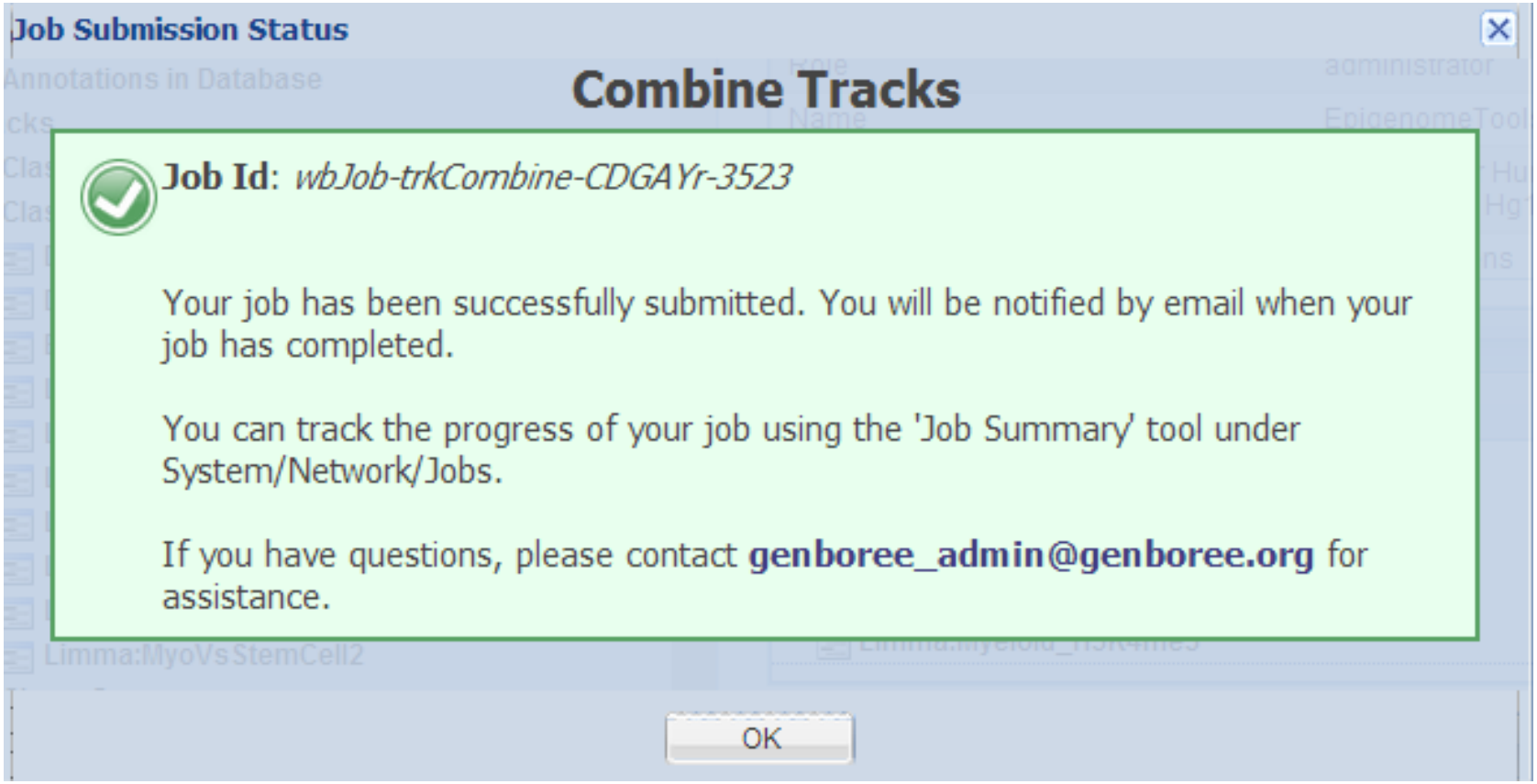

GENBOREE
BCM
Baylor College of Medicine
System/Network
Data
QC and Pre-processing
Genome
Transcriptome
Cistrome
Epigenome
Metagenome
Welcome to the Genboree Workbench!
[Getting Started]
Step 23 - Expand "Data" > "Tracks" --> "Utilities" --> "Track Operations" --> click on "Combine Tracks"
Data Selector
Refresh
Data Filter: Select a filter...
Your Database
Step 22 - To combine the LIMMA generated tracks, drag them to Input Data. Drag your database to Output Targets.
Limma:Myeloid_H3K27me3
Limma:Myeloid_H3K36me3
Limma:Myeloid_H3K4me1
Limma:Myeloid_H3K4me3
Limma:Myeloid_H3K9me3
Limma:MyoVsStemCell
Limma:MyoVsStemCell2
Class: Gene
Class: High Density Score Data
Class: lincRNA
Class: Marker
Class: Regulation
Class: Sequence
Class: true
Details
Attribute
Group
Role
Name
EpigenomeToolset
Description
Template for Human Genome, UCSC Build Hg19
Species
Homo sapiens
Input Data
Limma:Myeloid_H3K36me3
Limma:Myeloid_H3K4me1
Limma:Myeloid_H3K27me3
Limma:Myeloid_H3K9me3
Limma:Myeloid_H3K4me3
Output Targets
Your Database
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



Step 24 - Change the name of the track - “Limma:Myeloid_comb”. Name the Track Class as “Enhancer”

Step 25 - Select “Remove Duplicates”

You will see the message below upon successful submission of the Combine Tracks job



Now you have successfully created track of regions that differentially modified histone regions for Myeloid lineage. Continue to Use Case 19 to find regions that are undergoing epigenomic changes during myeloid differentiation.