Epigenomic Toolset

Mapping the Ontogenetic Pathways of Cellular Differentiation from the Human Epigenome Atlas Data Using the **Epigenomic Toolset** Within the Genboree Workbench

Aleks Milosavljevic
Baylor College of Medicine

Epigenomics: A Roadmap to the Living Genome
Boston, October 22nd, 2013

Acknowledgements

Epigenomics Data Analysis and Coordination Center (EDACC)

BCM Baylor College of Medicine

Cristian Coarfa
Alan Harris
Viren Amin
Vitor Onuchic
Andrew Jackson
Sriram Raghuraman
Sameer Paithankar
Matt Roth
Kevin Riehle
Tim Charnecki

ROADMAP epigenomics PROJECT

Govind Ramamoorthy
Eleonora Laritsky
Robert Waterland

UCSF University of California San Francisco

Joe Costello

Cydney Nielsen
Martin Hirst

CANADA'S MICHAEL SMITH GENOME SCIENCES CENTRE

Steven Jones

IHEC International Human Epigenome Consortium

AIST CBRC

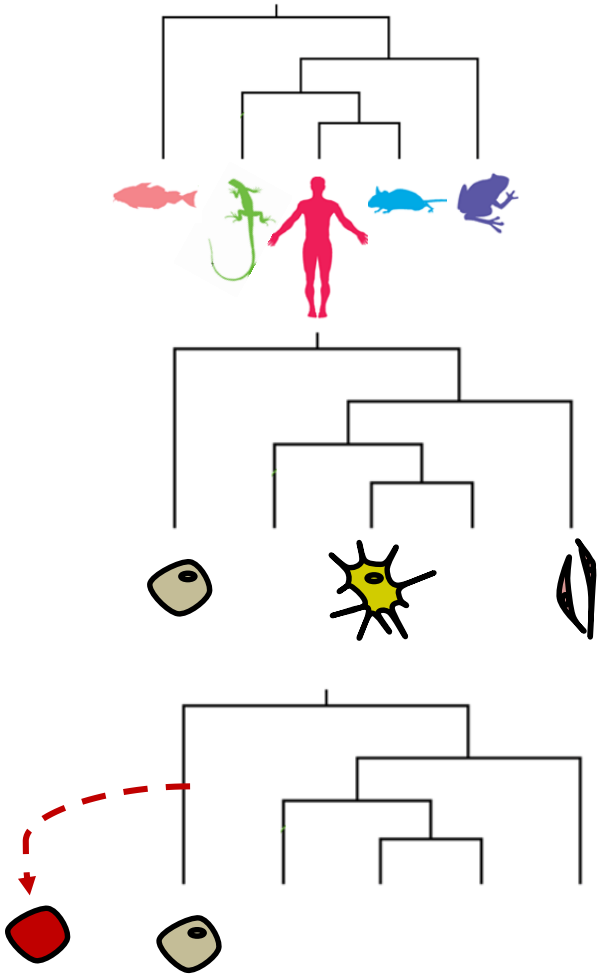
Toutai Mituyama

The scope of the Epigenomic Toolset

Phylogeny

Reconstruct ontogeny
Ontogenetic pathways of cellular differentiation

Interpret epigenomic perturbations
Disease-associated



Applications of the Epigenomic Toolset

1. Identify cell-type specific epigenomic states

2. Inform experiment design by identifying informative marks

3. Interpret epigenomic changes during cellular differentiation

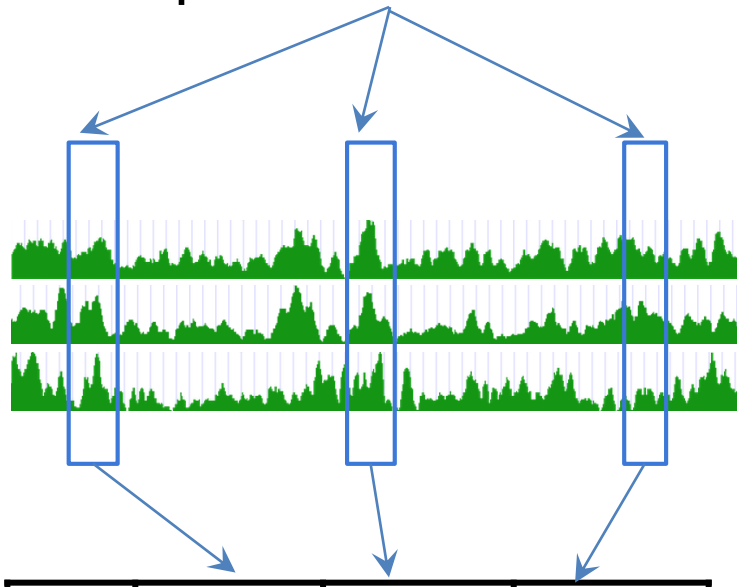
4. Detect and interpret epigenomic perturbations in disease (cancer cells) relative to epigenomes of normal cells

5. Detect shifts in cell type composition in complex tissues (tumor vs. normal) using epigenomes of constituent cells

6. Enable world-wide collaboration and integrative analysis by virtual integration of data, tools, and computing resources

Epigenomic Data Slice

Regions Of Interest (ROIs)
(enhancers, promoters, ChIP-seq peaks, etc.)



Epigenomic marks
(e.g., histone marks)

	ROI 1	ROI 2	ROI 3
Cell type 1	0.8	0.7	0.3
Cell type 2	0.7	0.6	0.7
Cell type 3	0.6	0.8	0.2

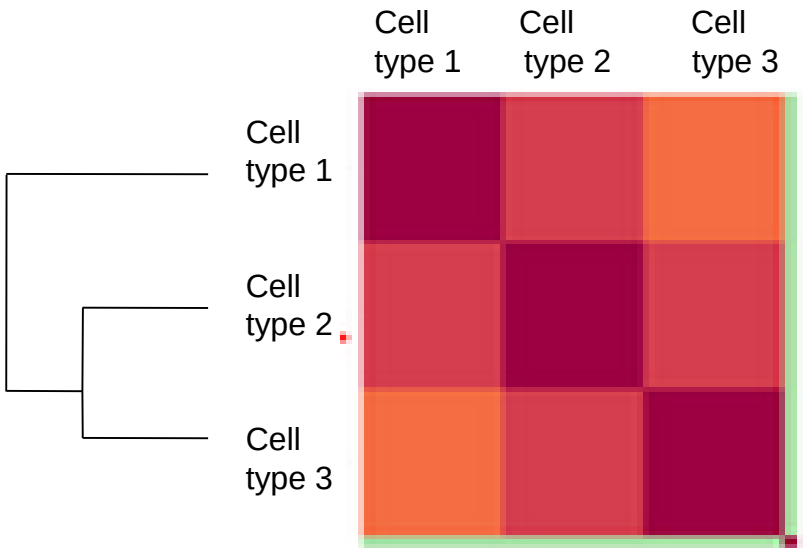
Epigenomic
Data Slice

Heatmaps and Dendograms

	ROI 1	ROI 2	ROI 3
Cell type 1	0.8	0.7	0.3
Cell type 2	0.7	0.6	0.7
Cell type 3	0.6	0.8	0.2

Epigenomic Data Slice

Self-correlation



Dendogram

Heatmap

BCM
Baylor College of Medicine

System/Network | Data | QC and Pre-processing | Genome | Transcriptome | Cistrome | Epigenome | Metagenome | Visualization | Help

Random Forest
QIIME
QC
Search for Similar Signals by Correlation
Analyze Signals
Compute Similarity Matrix (heatmap)
Create Track Lists from Newick Tree
Slice Epigenomic Data
Analyze Signals in the Context of Epigenome Atlas

Welcome to the Genboree Workbench! [Getting Started]

Data Selector

Refresh | Data Filter: Select a filter...

- Peak Calling Demo
- RNA-Seq Tool Demo
- Projects
- Epigenomics Roadmap Repository
- GenboreeUser_group
  - Databases
    - GenboreeUser_database
  - 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
    - Use_Case_14_GU
    - Use_Case_18_GU
    - Use_Case_19_GU
    - Use_Case_20_GU
    - Use_Case_21_GU
- GMT_Tutorial

Drag Input

Drag Output

Details

Attribute	
View Link	
Group	
Name	

Input Data

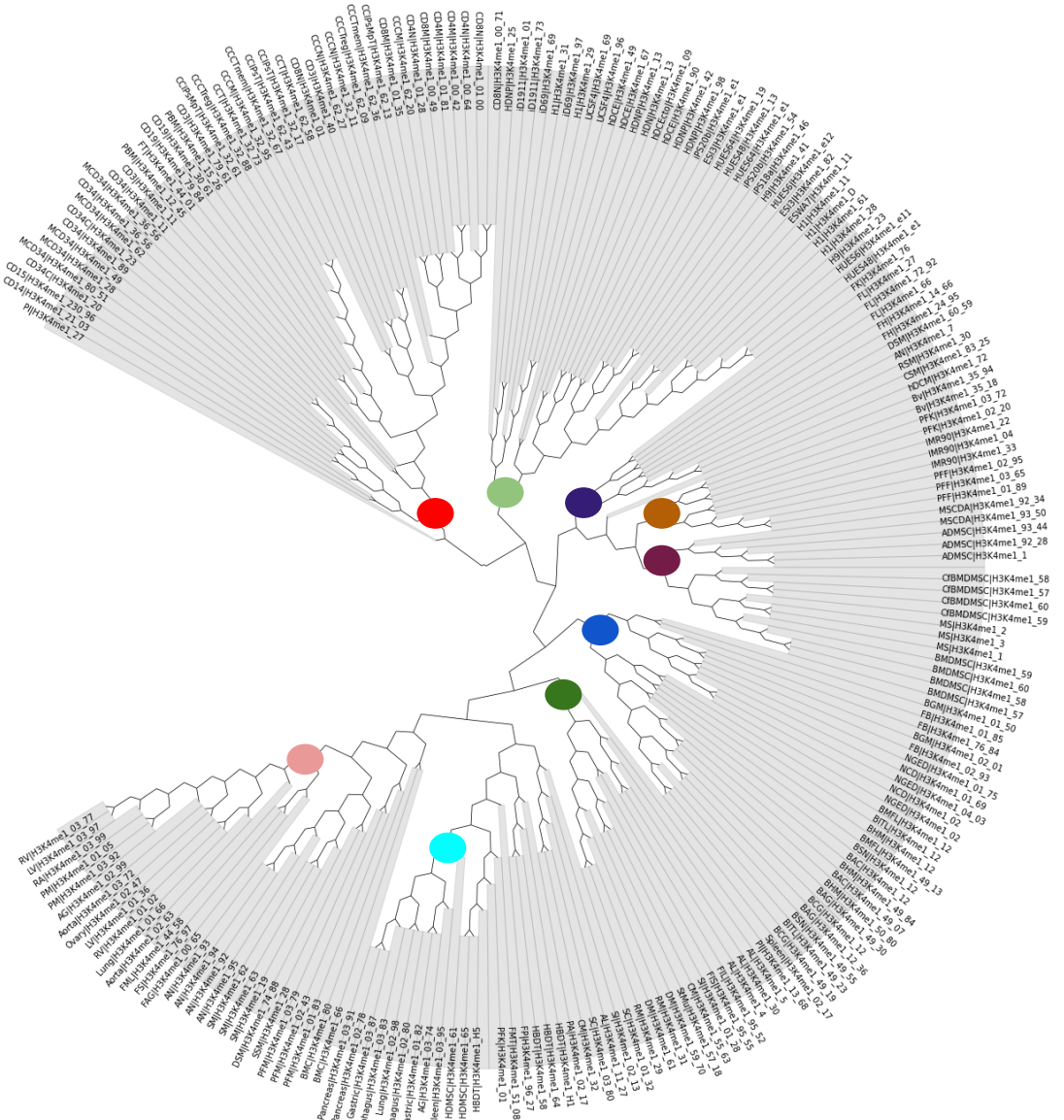
- Gene:ProteinCoding_promoter ← ROIs
- release9_H3K4me3_subset ← Epigenomic marks

Output Targets

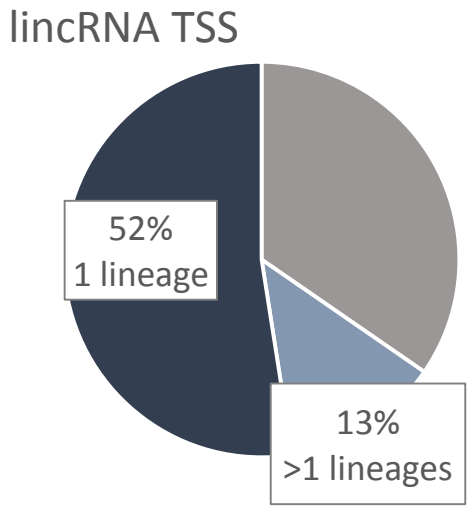
- GenboreeUser_database ← Database
- Use_Case_18_GU ← Project page

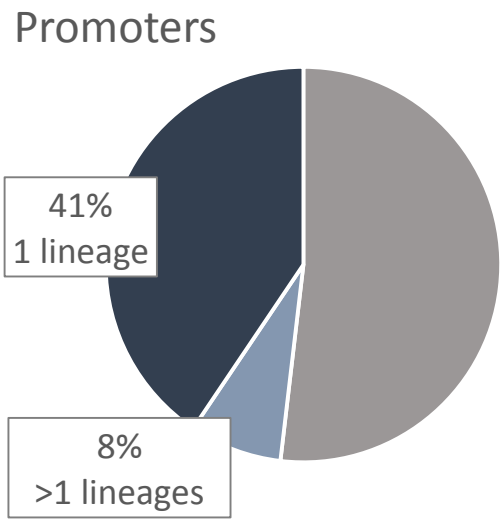
Dendogram of 100 epigenomes constructed from ~0.5M enhancers and H3K4me1

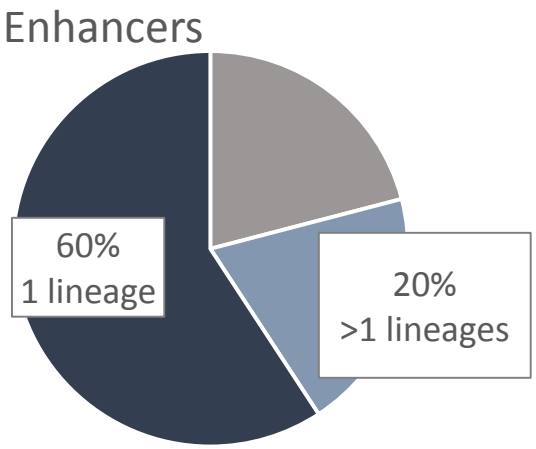
- Immune
- Endoderm
- Cardiac
- StemCells
- Ectodermal Epithelial
- Gastrointestinal
- Brain
- Fetal Organs
- Mesenchymal Stem Cells



ROIs with lineage-specific active epigenomic marks







Adult Brain	41441
Cardiac	18675
Epidermal	20097
Endoderm	12309
Fetal Muscle	6467
Fetal Organ	995
GastroInt	13531
HSCs	21931
Mesenchymal	51164
Myeloid	13993
NeuroProg	19849
Placenta	7757
Smooth Muscle	16427
Stem Cell	87597
T-cells	59808
Total	**392041**

Applications of the Epigenomic Toolset

1. Identify cell-type specific epigenomic states

2. Inform experiment design by identifying informative marks

3. Interpret epigenomic changes during cellular differentiation

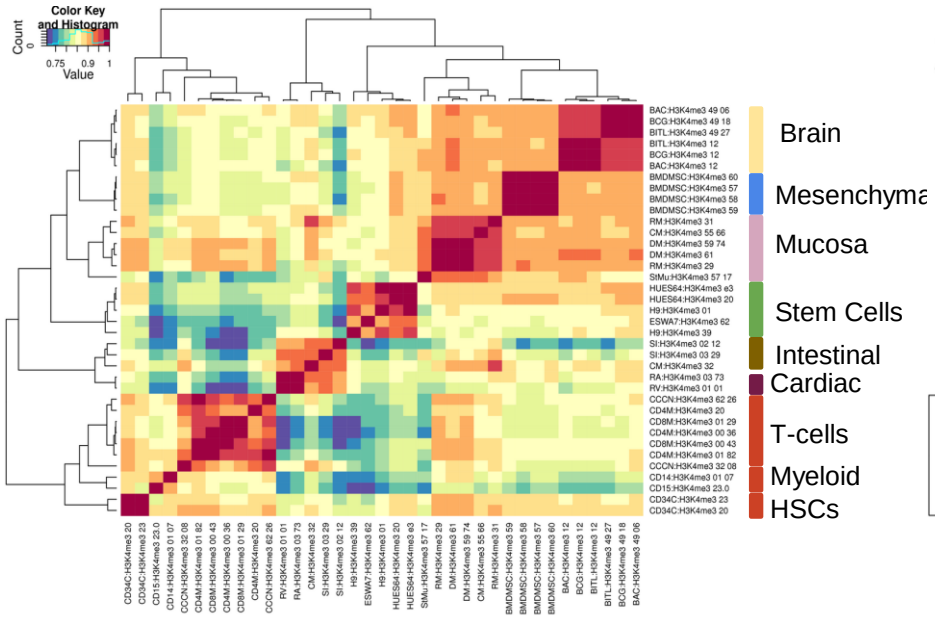
4. Detect and interpret epigenomic perturbations in disease (cancer cells) relative to epigenomes of normal cells

5. Detect shifts in cell type composition in complex tissues (tumor vs. normal) using epigenomes of constituent cells

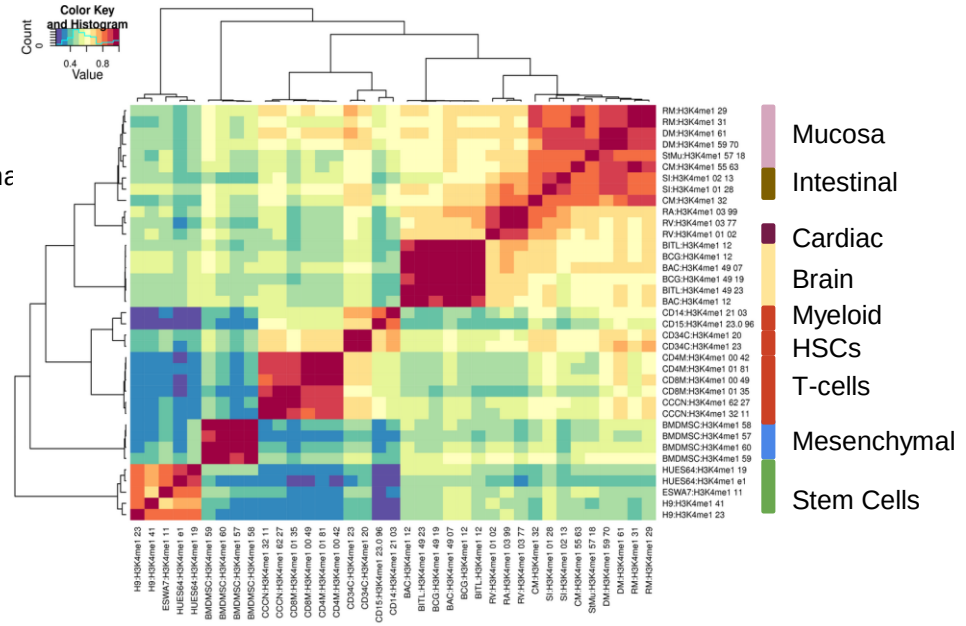
6. Enable world-wide collaboration and integrative analysis by virtual integration of data, tools, and computing resources

Recurrent tree branches

H3K4me3 signal over promoters

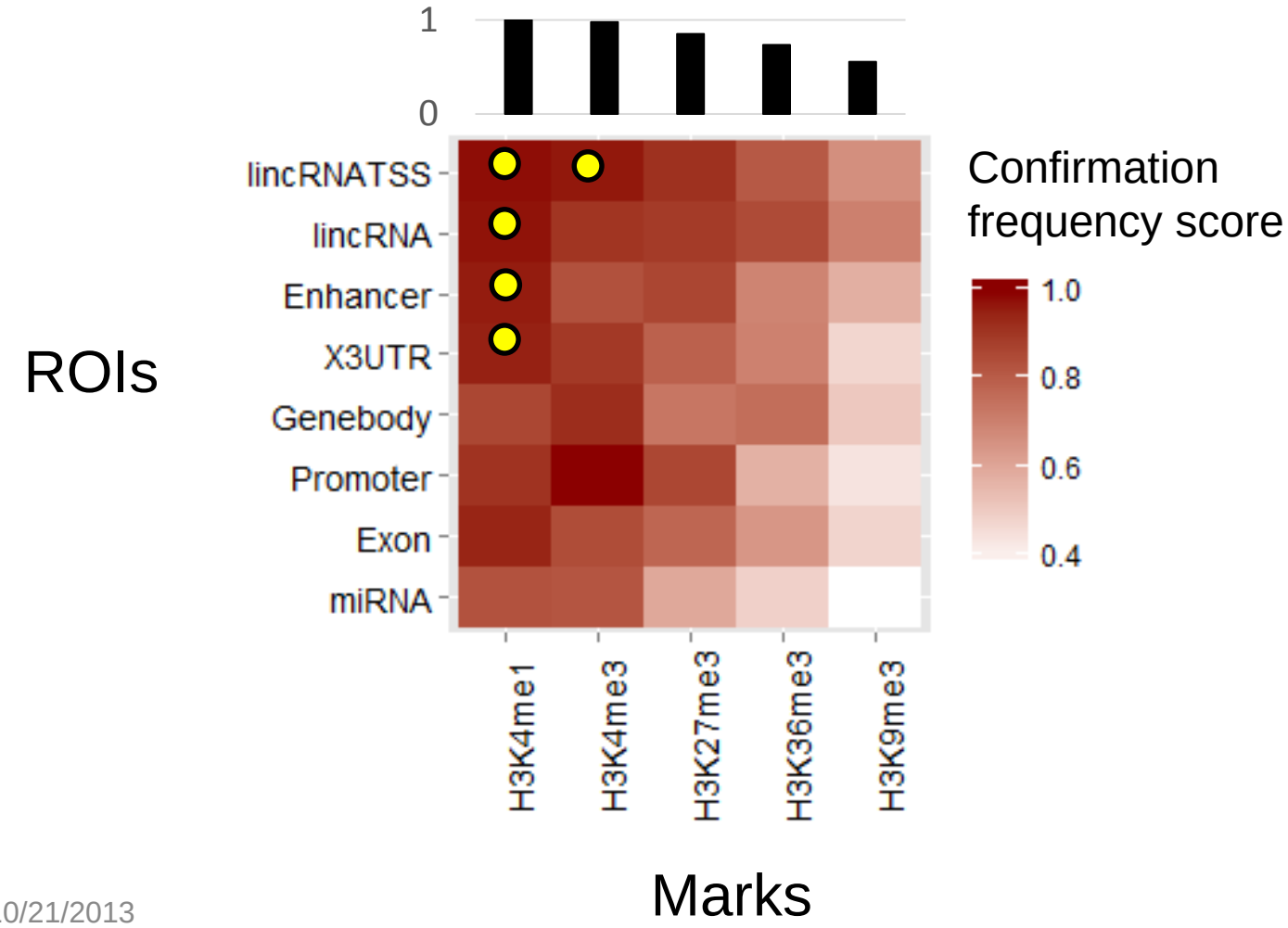


H3K4me1 signal over lincRNA Transcription Start Site (TSS)



Hypothesis: a branch shared by many trees (*confirmed by many mark-ROI combinations*) corresponds to an ontogenetic pathway

Informative epigenomic marks and mark-ROI combinations



Applications of the Epigenomic Toolset

1. Identify cell-type specific epigenomic states

2. Inform experiment design by identifying informative marks

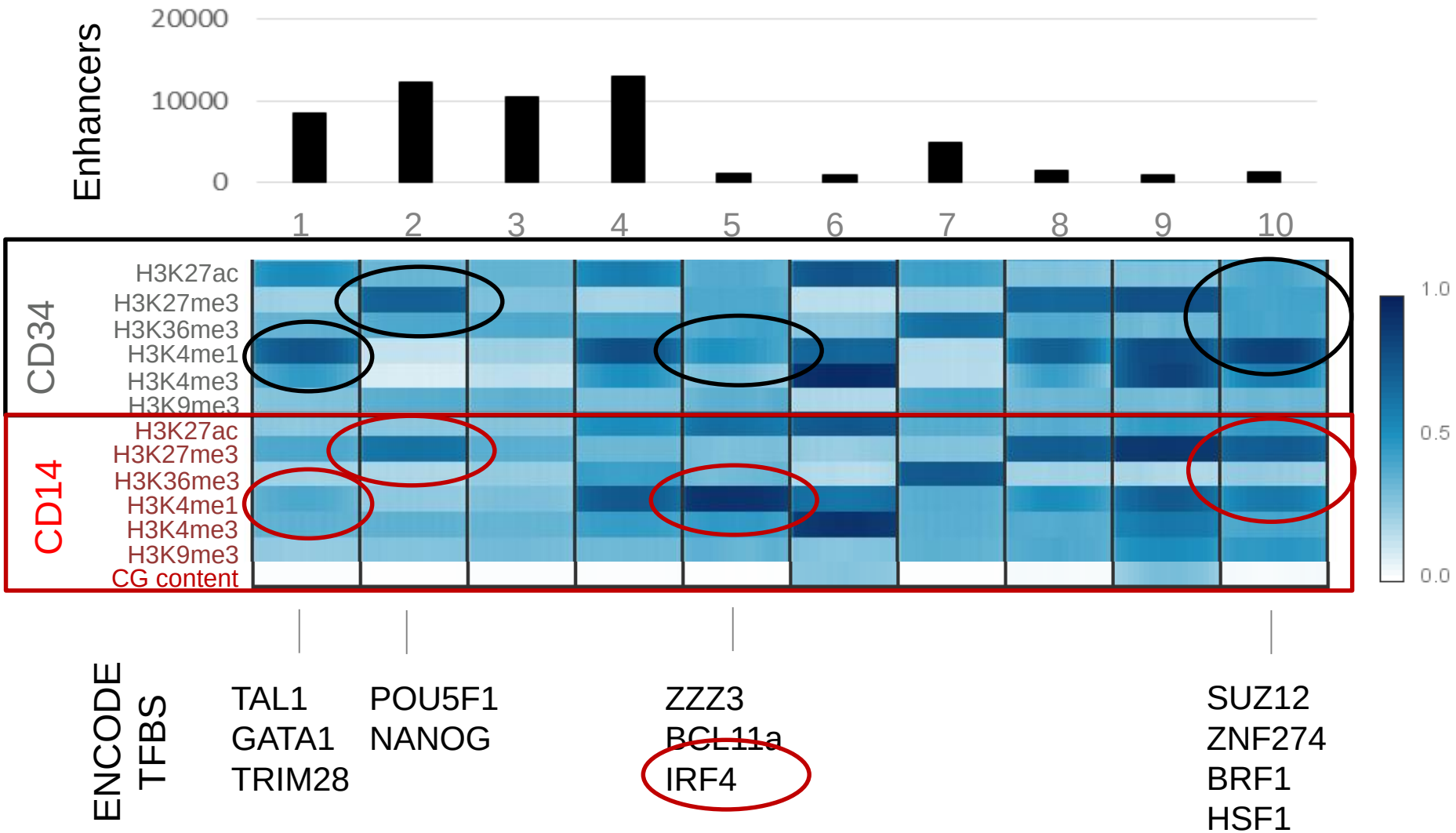
3. Interpret epigenomic changes during cellular differentiation

4. Detect and interpret epigenomic perturbations in disease (cancer cells) relative to epigenomes of normal cells

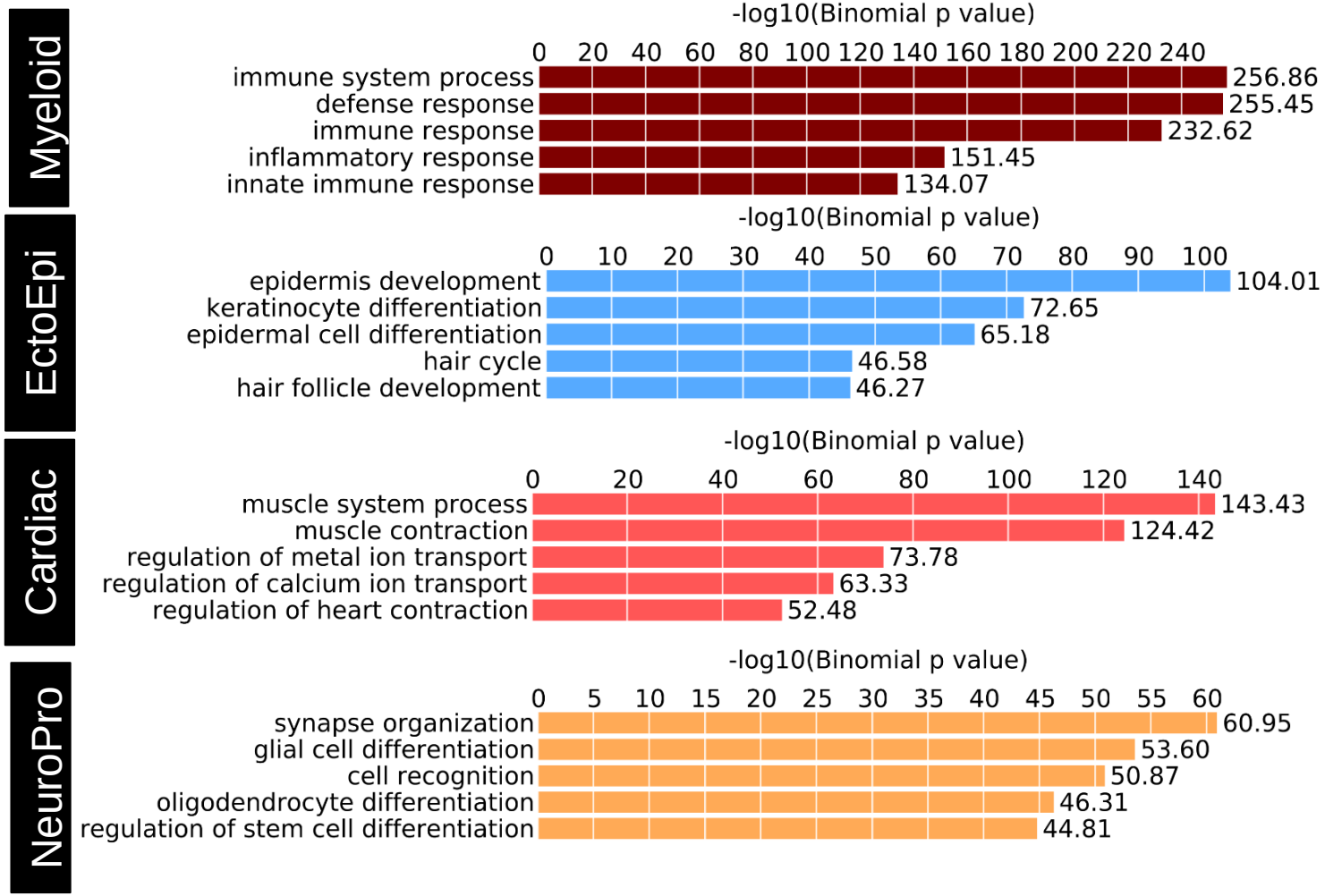
5. Detect shifts in cell type composition in complex tissues (tumor vs. normal) using epigenomes of constituent cells

6. Enable world-wide collaboration and integrative analysis by virtual integration of data, tools, and computing resources

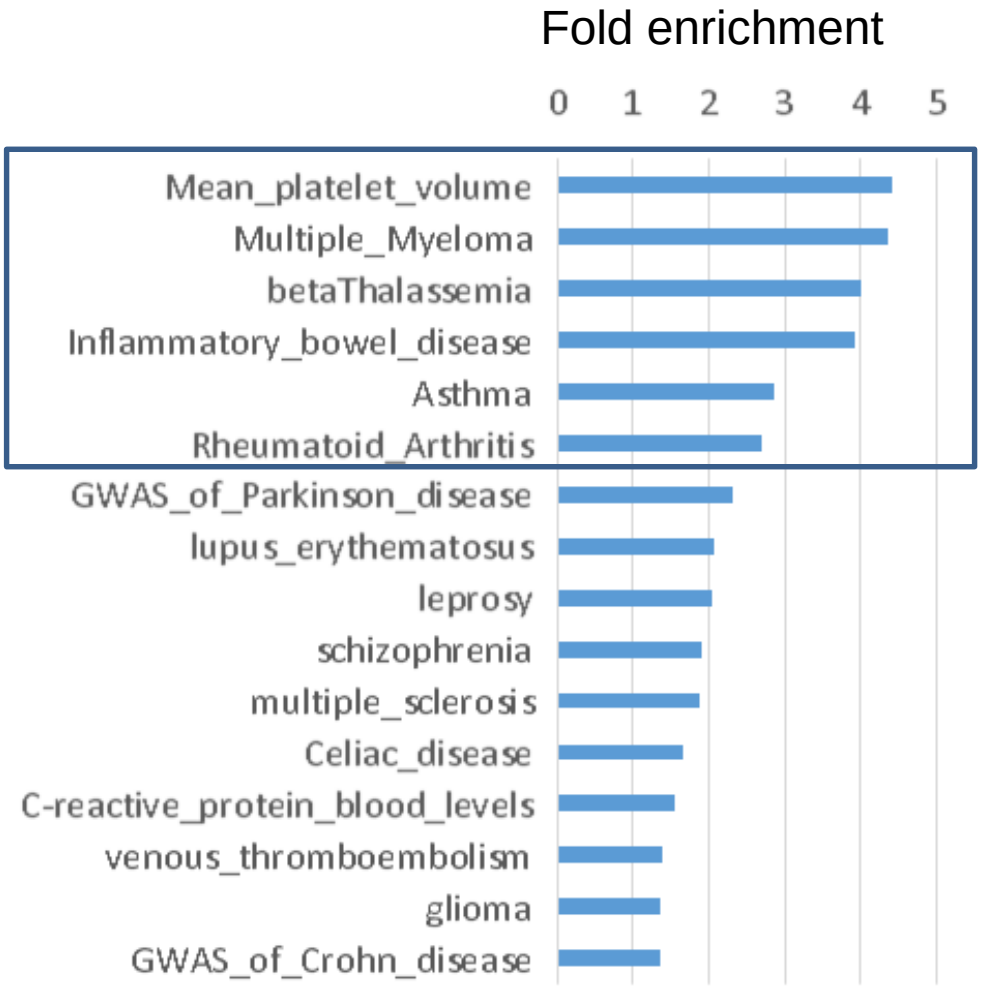
Mapping epigenomic changes using Spark tool



Pathway enrichment using H3K4me1 over enhancers and GREAT tool



GWAS regions enriched for enhancers with **immune-lineage specific** H3K4me1 mark



Applications of the Epigenomic Toolset

1. Identify cell-type specific epigenomic states

2. Inform experiment design by identifying informative marks

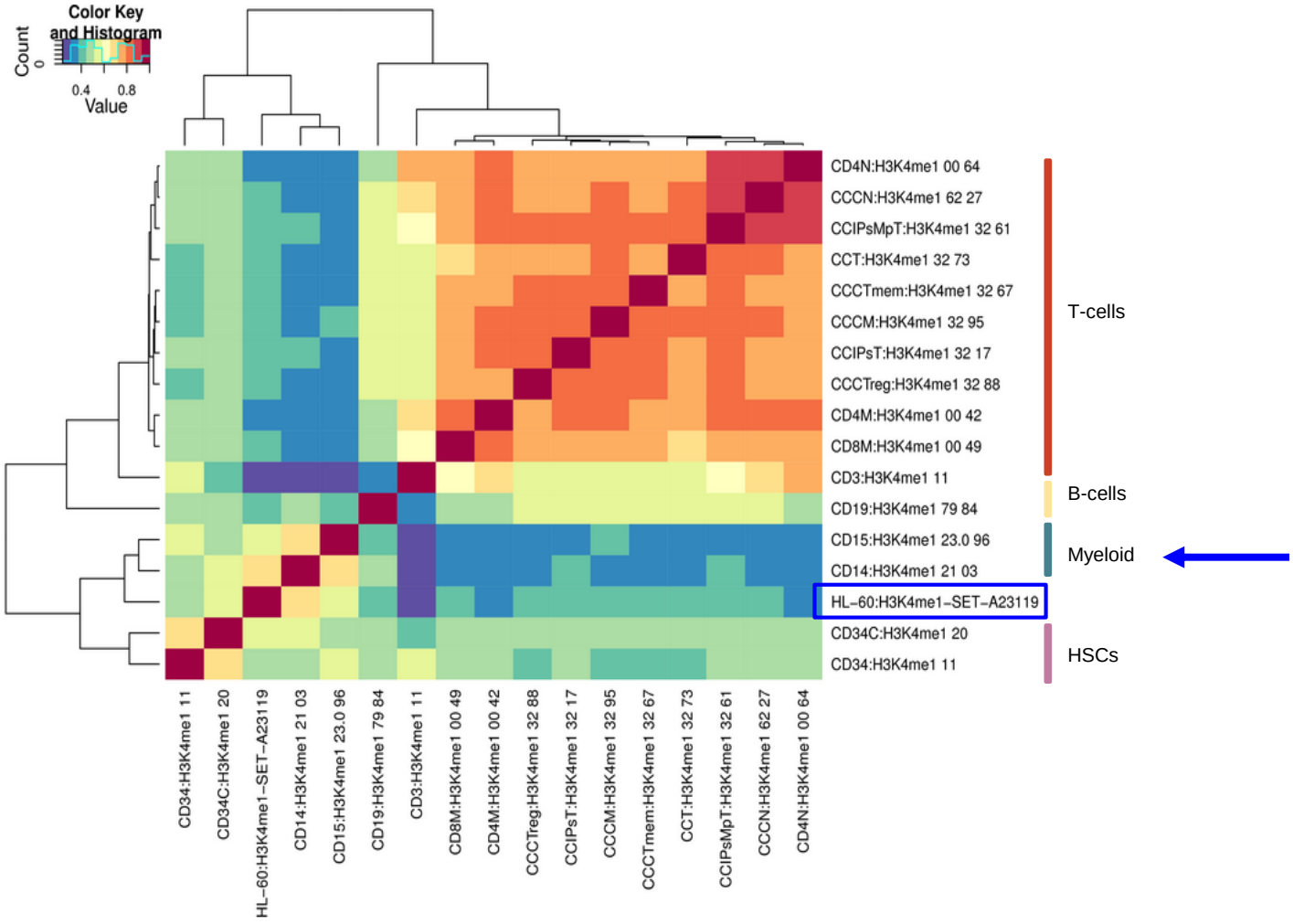
3. Interpret epigenomic changes during cellular differentiation

4. Detect and interpret epigenomic perturbations in disease (cancer cells) relative to epigenomes of normal cells

5. Detect shifts in cell type composition in complex tissues (tumor vs. normal) using epigenomes of constituent cells

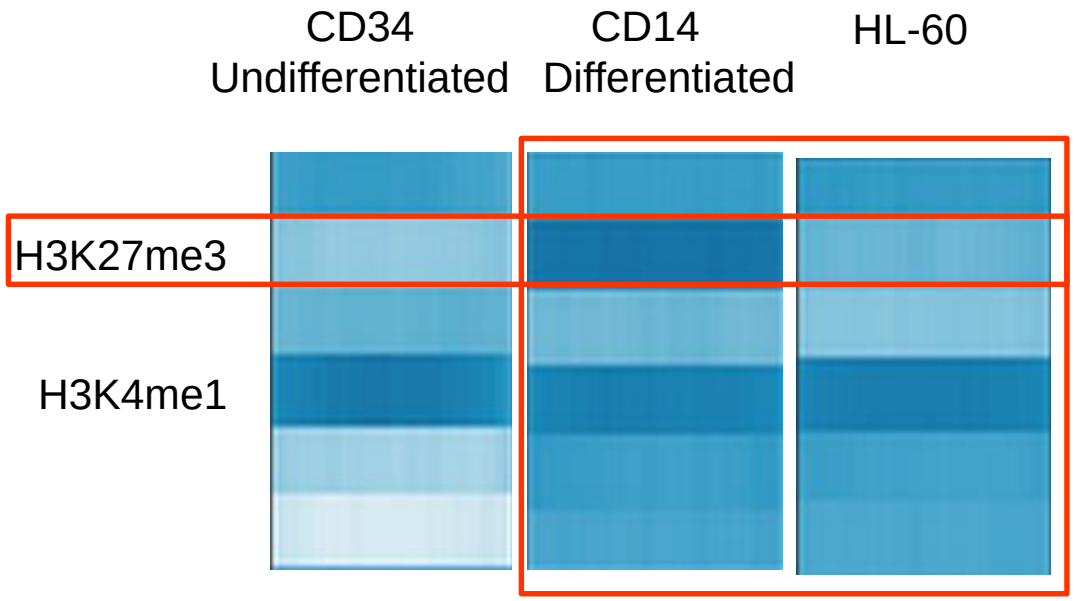
6. Enable world-wide collaboration and integrative analysis by virtual integration of data, tools, and computing resources

Dendogram including HL-60, a ***promyelocytic leukemia*** cell line confirms ***myeloid lineage***



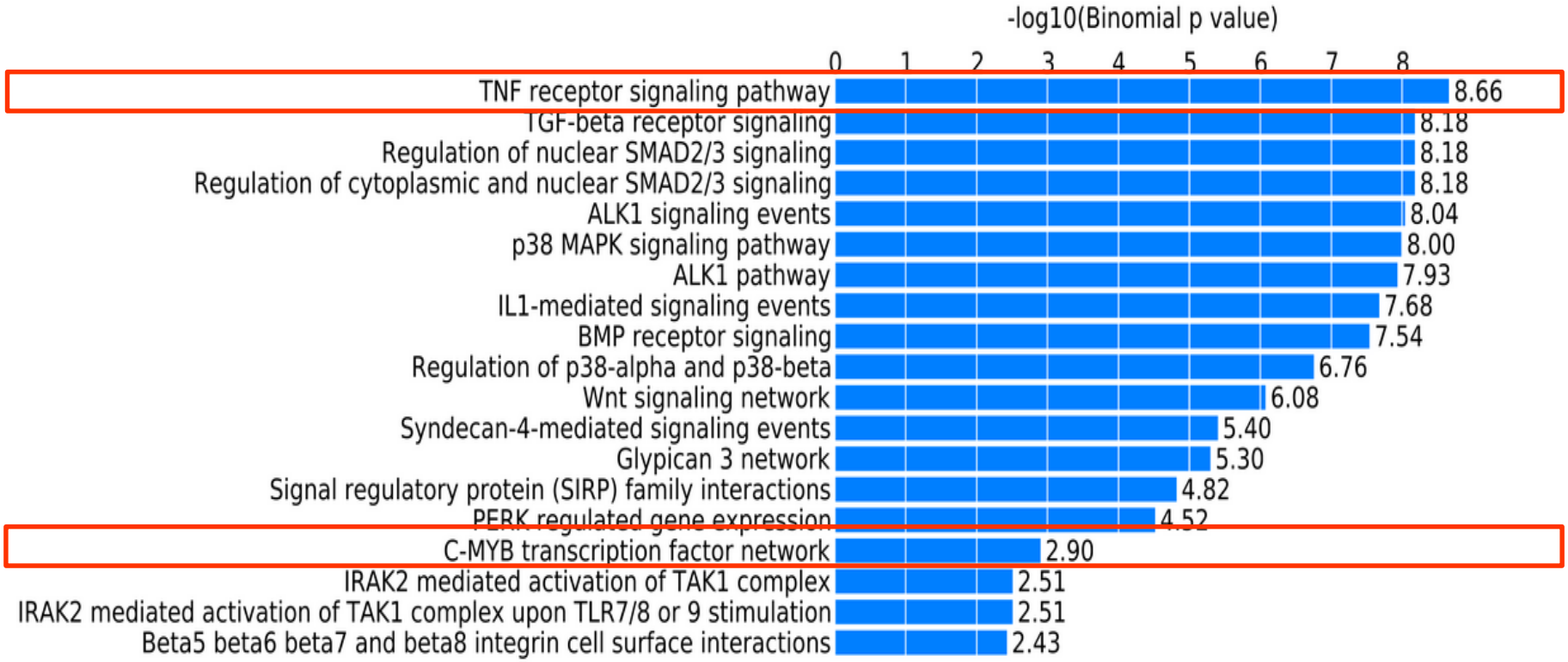
Spark analysis involving closest reference epigenomes reveals lack of differentiation

Epigenomic patterns over 1025 enhancers that change epigenomic state during myeloid differentiation



Pathway analysis reveals up-regulation of oncogenetic pathways

Analysis of enhancers with up-regulated H3K4me1 using GREAT tool



Applications of the Epigenomic Toolset

1. Identify cell-type specific epigenomic states

2. Inform experiment design by identifying informative marks

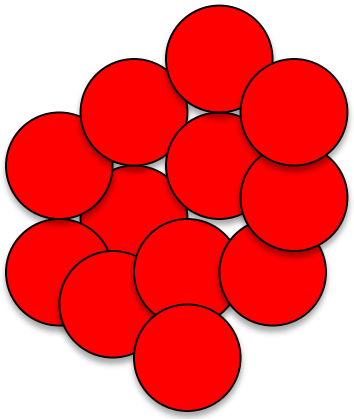
3. Interpret epigenomic changes during cellular differentiation

4. Detect and interpret epigenomic perturbations in disease (cancer cells) relative to epigenomes of normal cells

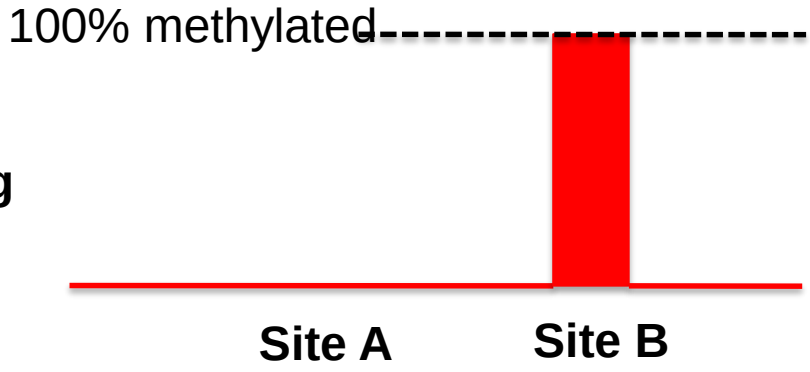
5. Detect shifts in cell type composition in complex tissues (tumor vs. normal) using epigenomes of constituent cells

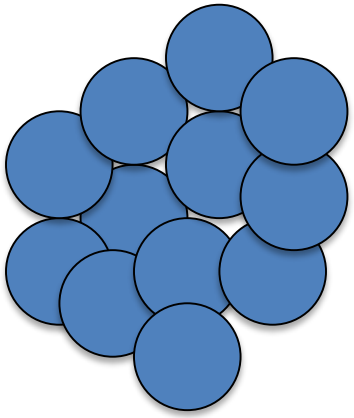
6. Enable world-wide collaboration and integrative analysis by virtual integration of data, tools, and computing resources

Samples from pure cell types

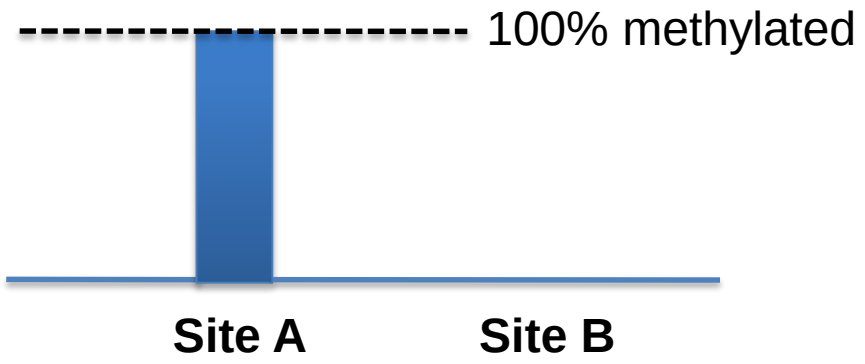


Methylation Profiling

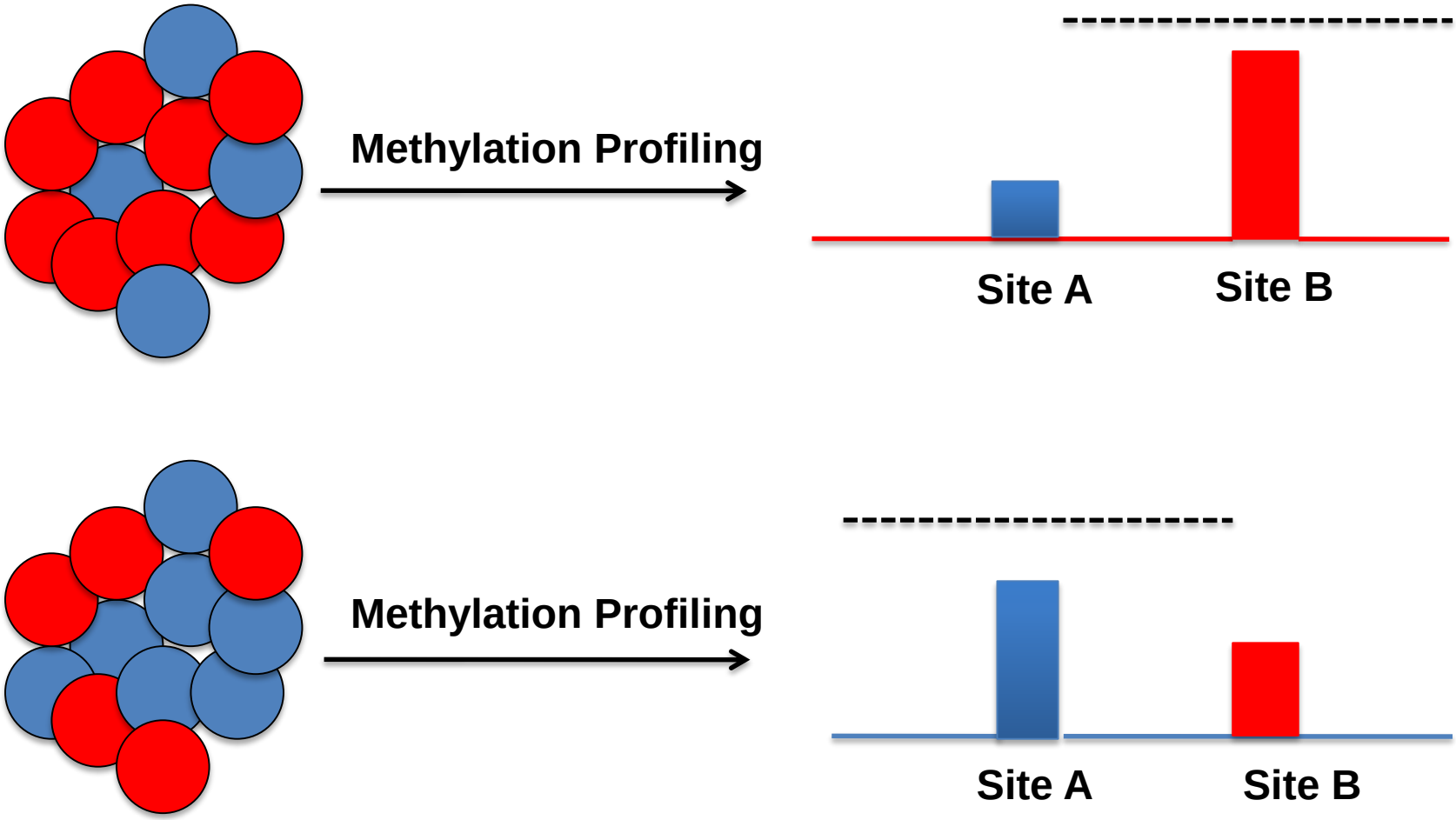




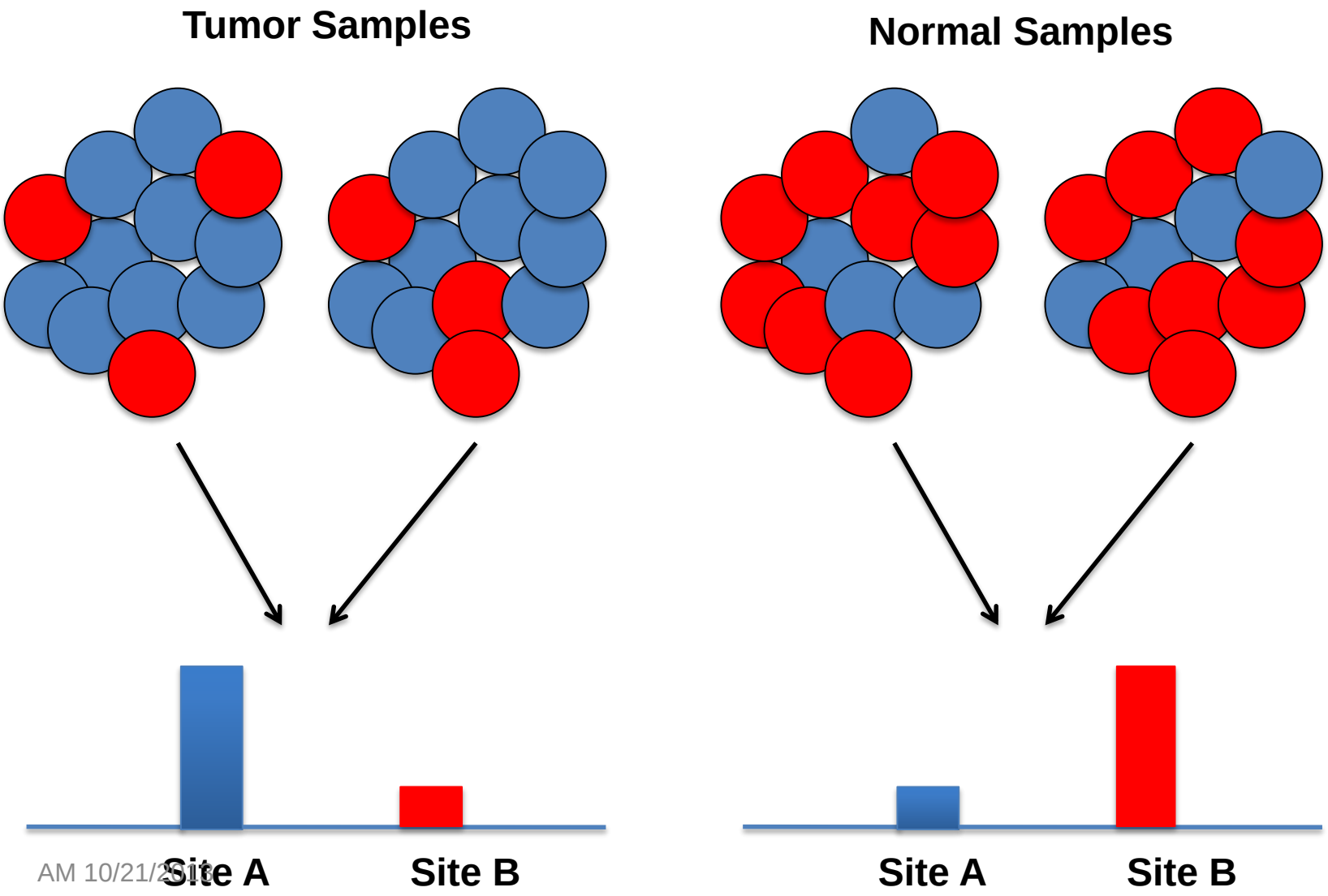
Methylation Profiling



Samples from a complex tissue



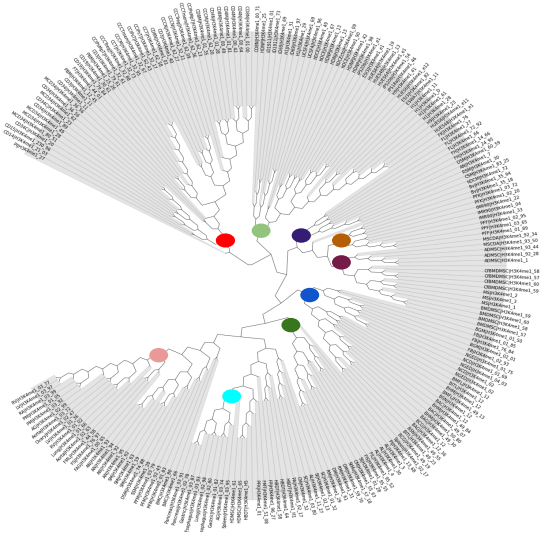
Tumor vs. normal tissue



Marker CpGs for Illumina arrays

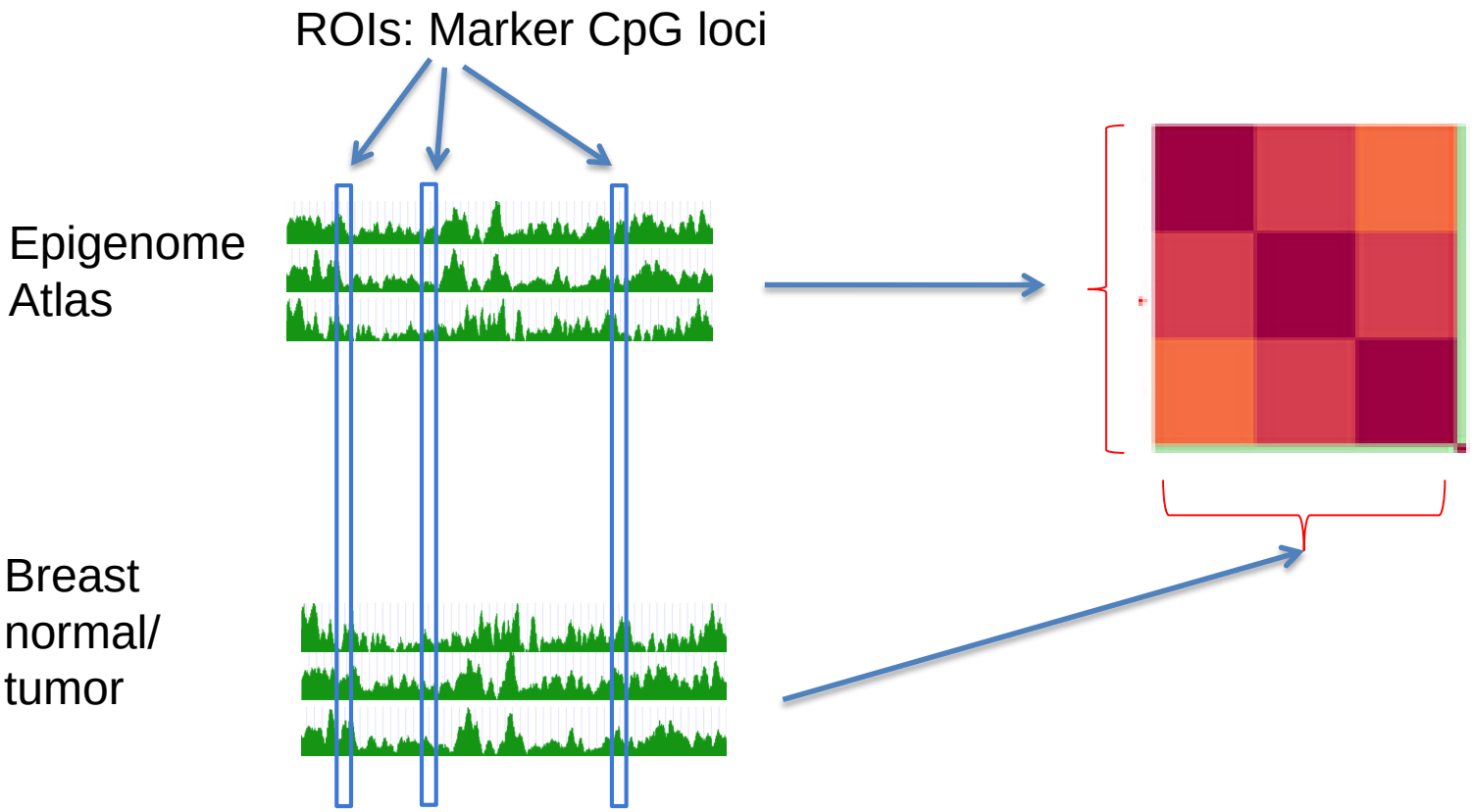
LIMMA tool within the Genboree Workbench used to select **marker CpGs** that show lineage-specific methylation

- Illumina 450k array CpGs
- Illumina 27k array CpGs
- Shared 450K and 27K array CpGs

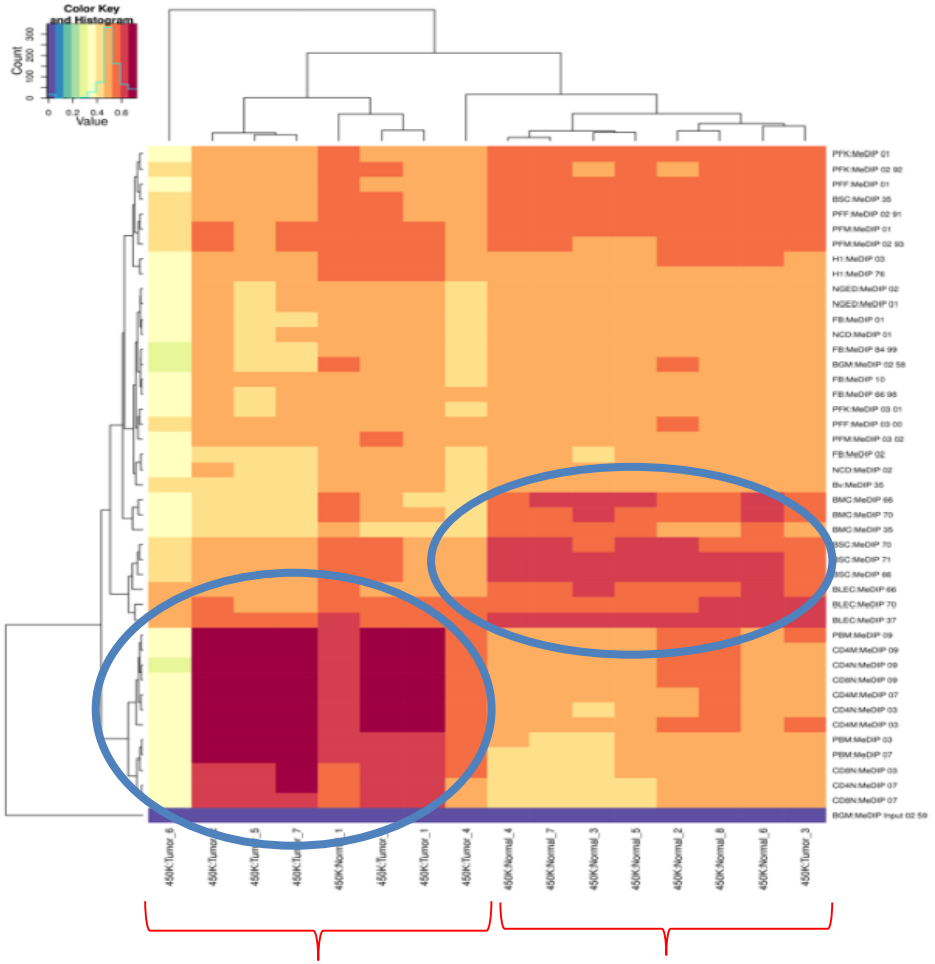


Tracks with Illumina marker CpGs available at www.genboree.org (Use case 21)

Comparison against reference epigenomes



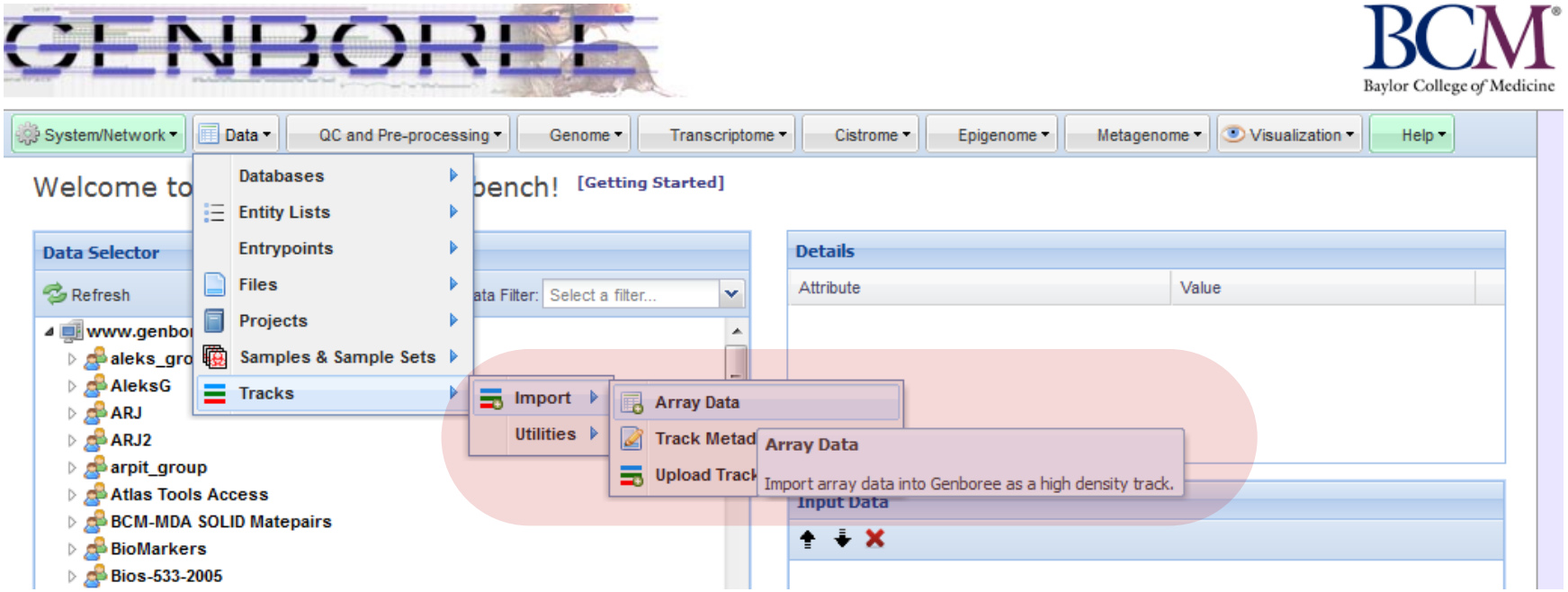
Comparison against reference epigenomes



Illumina HM 450K **8 tumor and 8 normal samples** from Dedeurwaerder, S.et al. (2011) "Evaluation of the Infinium Methylation 450K technology", Epigenomics 3(6):771-84.

Comparing methylation arrays against the methylomes in the Epigenome Atlas

Step 1:
Array import via the Genboree Workbench:



Applications of the Epigenomic Toolset

1. Identify cell-type specific epigenomic states

2. Inform experiment design by identifying informative marks

3. Interpret epigenomic changes during cellular differentiation

4. Detect and interpret epigenomic perturbations in disease (cancer cells) relative to epigenomes of normal cells

5. Detect shifts in cell type composition in complex tissues (tumor vs. normal) using epigenomes of constituent cells

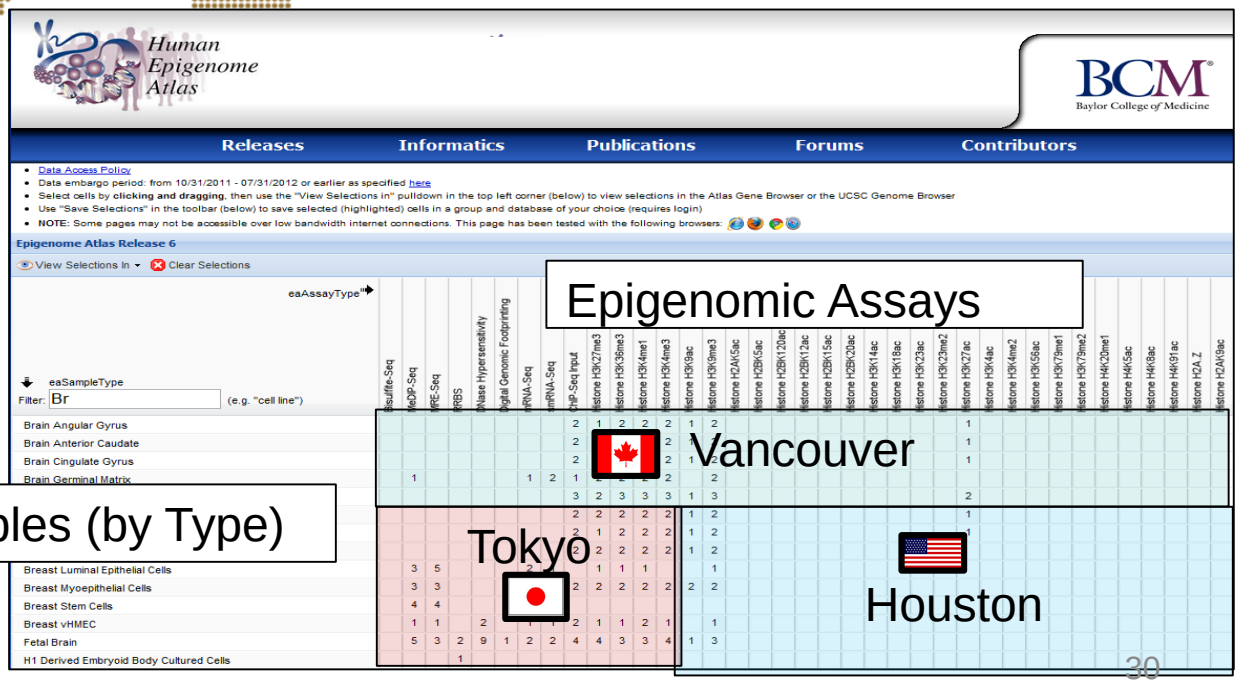
6. Enable world-wide collaboration and integrative analysis by virtual integration of data, tools, and computing resources

Virtual integration across the International Human Epigenome Consortium (IHEC)

Vancouver
Steven Jones
Martin Hirst

Houston
NIH Roadmap
Epigenomics

Tokyo
Toutai Mituyama

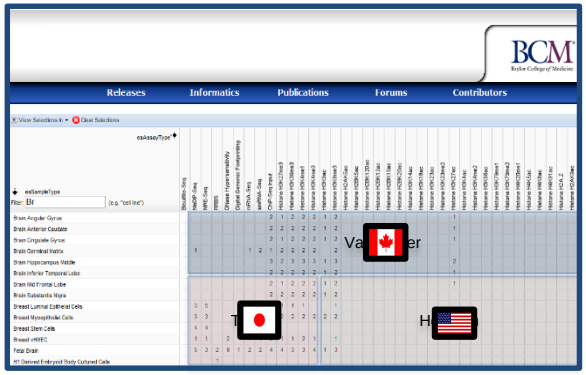


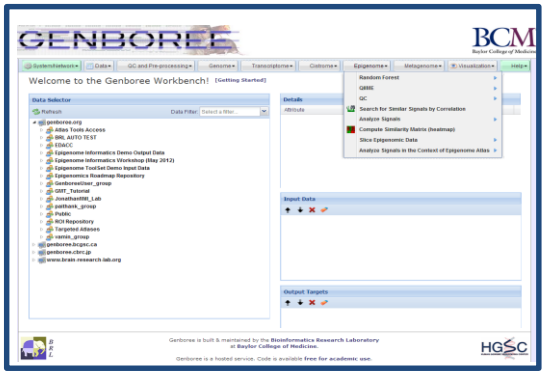
Virtual integration

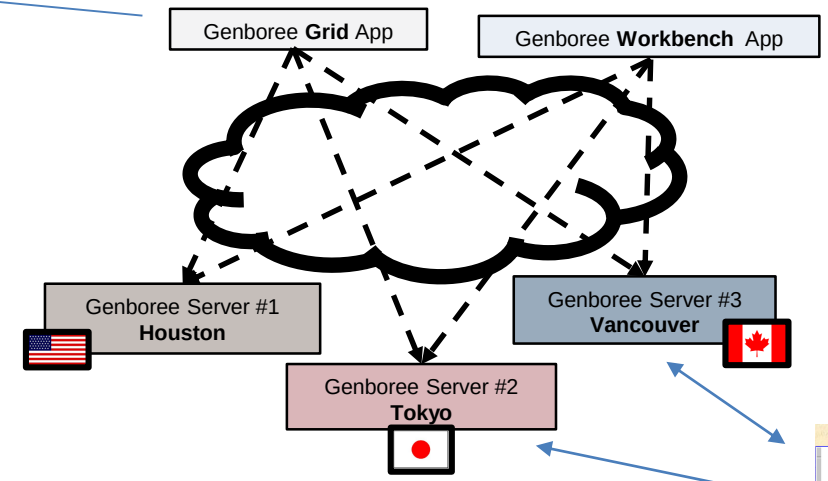
Browsing

Analysis

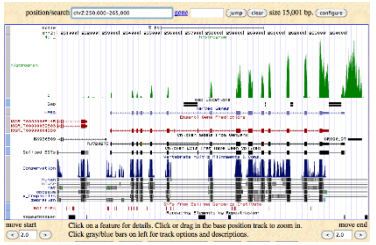
Only the “data slices” transferred across sites (10MB vs 10GB)







Visualization in UCSC Genome Browser



Applications of the Epigenomic Toolset

1. Identify cell-type specific epigenomic states

2. Inform experiment design by identifying informative marks

3. Interpret epigenomic changes during cellular differentiation

4. Detect and interpret epigenomic perturbations in disease (cancer cells) relative to epigenomes of normal cells

5. Detect shifts in cell type composition in complex tissues (tumor vs. normal) using epigenomes of constituent cells

6. Enable world-wide collaboration and integrative analysis by virtual integration of data, tools, and computing resources

American Society of Human Genetics October 22-26 2013

BOSTON

Human Epigenome Atlas and Epigenomic Profile Analysis Using the Genboree Workbench

Separate advance registration required. Please check the appropriate box on the meeting registration form.

Tuesday, October 22

2:00 pm – 3:30 pm

Room 104AB, Level 1, Convention Center

SORRY! SOLD OUT

www.genboree.org

GENBOREE

BCM Baylor College of Medicine

You are currently not logged in.

ASHG workshop tomorrow

Use cases available here

« *American Society of Human Genetics 2013 Workshop (Boston)* »

- **Link to announcement**
- **Use Cases**

« *6th Epigenome Informatics Workshop* »

- **Use Cases**
- **FAQs**
- **Link to Workshop Presentations**

Login Name:

Guest/Public View

New to Genboree? **Register here!**