

Introduction to the Third Epigenome Informatics Workshop

NIH Roadmap
Epigenomics Data Analysis and Coordination Center



March 5-6 2012
Houston, Texas

Workshop Objective:

Catalyze Conversion of Epigenomic Profiling Data into Biological Insights through Integrative Analysis

- Methods (Assays, Data Processing)
- Standards (Metadata, Interoperability)
- Data Resources (Human Epigenome Atlas)
- Tools (Epigenomic Toolset, Genboree Workbench)
- Use Cases / Case Studies
- Collaborative Opportunities / Networking / Exchange of Experience

Workshop Participants

Monday, March 5th 2012

Session 1

9:00 – 9:45 am Introduction to the Workshop – Aleks Milosavljevic

9:45 – 10:30 am *Quantitative profiling of histone modifications, peak calling and segmentation of epigenomic signals, and small RNA analysis* – Cristi Coarfa (EDACC)

10:30 – 11:15 am *Methylome mapping using MeDIP-seq, WGBS and RRBS and analysis of allelic imbalances* - Alan Harris (EDACC)

11:15 – 11:30 am *Break*

11:30 – 12:15 pm *Whole-genome bisulfite sequencing: comparative analysis of programs for mapping bisulfite reads* – Govind K Ramamoorthy (EDACC, Rob Waterland Laboratory)

12:15 – 1:00 pm *Methylome mapping using Illumina Human Methylome 450K arrays*- Bekim Sadikovic (Art Beaudet Laboratory)

1:00 – 2:00 pm Boxed lunch (outside auditorium)

Monday, March 5th 2012

Session 1

9:00 – 9:45 am Introduction to the Workshop – Aleks Milosavljevic

9:45 – 10:30 am *Quantitative profiling of histone modifications, peak calling and segmentation of epigenomic signals, and small RNA analysis* – Cristi Coarfa (EDACC)

10:30 – 11:15 am *Methylome mapping using MeDIP-seq, WGBS and RRBS and analysis of allelic imbalances* - Alan Harris (EDACC)

11:15 – 11:30 am *Break*

11:30 – 12:15 pm *Whole-genome bisulfite sequencing: comparative analysis of programs for mapping bisulfite reads* – Govind K Ramamoorthy (EDACC, Rob Waterland Laboratory)

12:15 – 1:00 pm *Methylome mapping using Illumina Human Methylome 450K arrays*- Bekim Sadikovic (Art Beaudet Laboratory)

1:00 – 2:00 pm Boxed lunch (outside auditorium)

Monday, March 5th 2012

Session 1

9:00 – 9:45 am Introduction to the Workshop – Aleks Milosavljevic

9:45 – 10:30 am *Quantitative profiling of histone modifications, peak calling and segmentation of epigenomic signals, and small RNA analysis* – Cristi Coarfa (EDACC)

10:30 – 11:15 am *Methylome mapping using MeDIP-seq, WGBS and RRBS and analysis of allelic imbalances* - Alan Harris (EDACC)

11:15 – 11:30 am *Break*

11:30 – 12:15 pm *Whole-genome bisulfite sequencing: comparative analysis of programs for mapping bisulfite reads* – Govind K Ramamoorthy (EDACC, Rob Waterland Laboratory)

12:15 – 1:00 pm *Methylome mapping using Illumina Human Methylome 450K arrays*- Bekim Sadikovic (Art Beaudet Laboratory)

1:00 – 2:00 pm Boxed lunch (outside auditorium)

Monday, March 5th 2012

Session 1

9:00 – 9:45 am Introduction to the Workshop – Aleks Milosavljevic

9:45 – 10:30 am *Quantitative profiling of histone modifications, peak calling and segmentation of epigenomic signals, and small RNA analysis* – Cristi Coarfa (EDACC)

10:30 – 11:15 am *Methylome mapping using MeDIP-seq, WGBS and RRBS and analysis of allelic imbalances* - Alan Harris (EDACC)

11:15 – 11:30 am *Break*

11:30 – 12:15 pm *Whole-genome bisulfite sequencing: comparative analysis of programs for mapping bisulfite reads* – Govind K Ramamoorthy (EDACC, Rob Waterland Laboratory)

12:15 – 1:00 pm *Methylome mapping using Illumina Human Methylome 450K arrays*- Bekim Sadikovic (Art Beaudet Laboratory)

1:00 – 2:00 pm Boxed lunch (outside auditorium)

Monday, March 5th 2012

Session 2

2:00 – 2:45 pm *NIH Epigenomics Roadmap: reference epigenomes, metadata standards, interoperability, and integrative data analysis* – Aleks Milosavljevic (EDACC)

2:45 – 3:15 pm *Introduction to Genboree Workbench features that will be used in Case Studies* - Aleks Milosavljevic (EDACC)

3:15 – 3:30 pm *Break*

3:30 – 5:00 pm *Preparation for Case Studies on Day 2 – Setting up projects, databases, accessing files, navigating Genboree* - Matt Roth, Kevin Riehle, Chia-Chin Wu, Yuan Yuan, Cristi Coarfa, Alan Harris, Aleks Milosavljevic

6:00 pm Depart for Houston Livestock Show & Rodeo or dinner on your own

Monday, March 5th 2012

Session 2

2:00 – 2:45 pm *NIH Epigenomics Roadmap: reference epigenomes, metadata standards, interoperability, and integrative data analysis* – Aleks Milosavljevic (EDACC)

2:45 – 3:15 pm *Introduction to Genboree Workbench features that will be used in Case Studies* - Aleks Milosavljevic (EDACC)

3:15 – 3:30 pm *Break*

3:30 – 5:00 pm *Preparation for Case Studies on Day 2 – Setting up projects, databases, accessing files, navigating Genboree* - Matt Roth, Kevin Riehle, Chia-Chin Wu, Yuan Yuan, Cristi Coarfa, Alan Harris, Aleks Milosavljevic

6:00 pm Depart for Houston Livestock Show & Rodeo or dinner on your own

Tuesday, March 6th 2012

Session 3

9:00 – 9:30 am *Review of Day 1 and of Case Studies* – Matt Roth (EDACC)

9:30 – 11:30 pm **Case Studies 1, 2, 5:** *Epigenomic Variation between Tissues (Part 1)* – Matt Roth, Kevin Riehle, Chia-Chin Wu, Yuan Yuan, Cristi Coarfa, Alan Harris, Aleks Milosavljevic

11:30 - 11:45 pm Break

11:45 – 1:00 pm **Case Studies 10, 9:** *Epigenomic Variation Between Tissues (Part 2)* – Matt Roth, Kevin Riehle, Chia-Chin Wu, Yuan Yuan, Cristi Coarfa, Alan Harris, Aleks Milosavljevic

1:00 – 2:00 pm Boxed lunch (outside auditorium)

The data for several of the workshop Use Cases was kindly provided by Dr. Jonathan Mill (King's College London, UK), and is under review for this publication:

"Tissue-specific epigenetic variation across brain and blood: functional annotation of the human brain methylome". Matthew Davies¹, Manuela Volta¹, Abhishek Dixit¹, Simon Lovestone¹, Cristian Coarfa², R. Alan Harris², Aleksandar Milosavljevic², Claire Troakes¹, Safa Al-Sarraj¹, Richard Dobson¹, Leonard C. Schalkwyk¹, Jonathan Mill^{1*}

¹Institute of Psychiatry, King's College London. UK. ²Baylor College of Medicine, Houston, Texas. USA. *Corresponding Author: Dr. Jonathan Mill, Address: Institute of Psychiatry, SGDP Centre, De Crespigny Park, Denmark Hill, London.

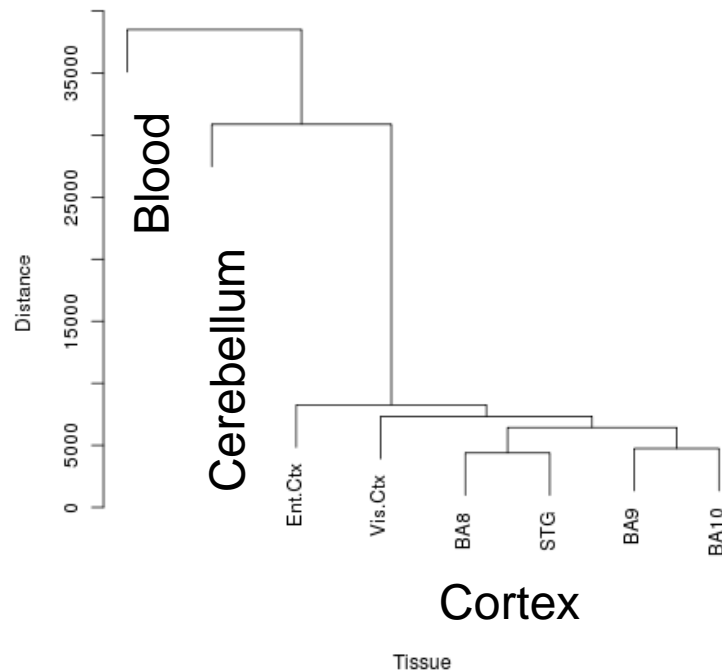
Since the paper is under review (Davies et al), it can not be shared with anyone outside of the workshop at this time, and we ask that you consider the data confidential. We will notify you when the data can be shared. Thank you for your understanding.

MEDIPS-processed signal averaged over 500bp windows genome-wide

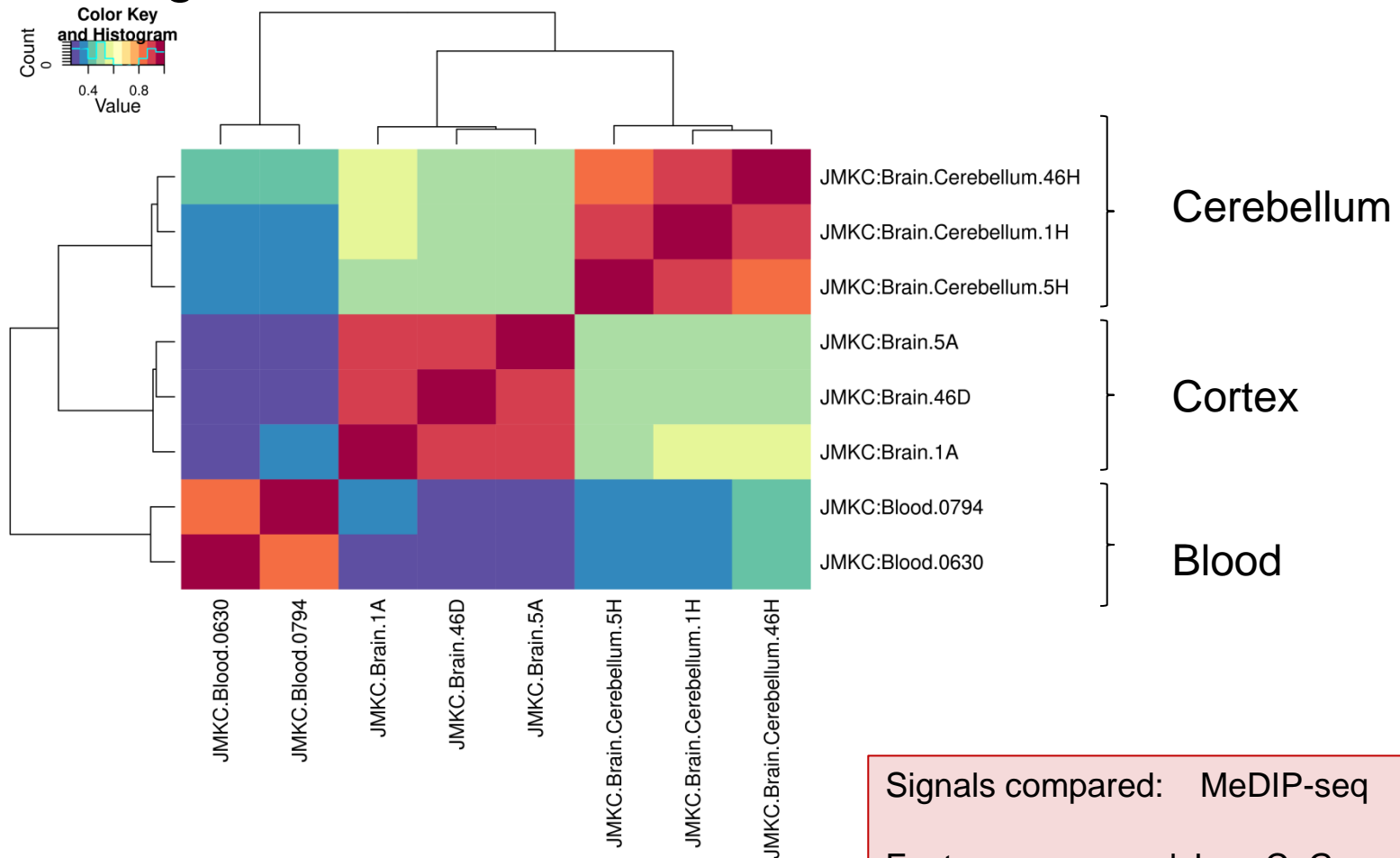
Pairwise correlation between signals

	BA9	BA10	BA8	STG	Ent. Ctx	Vis. Ctx	Cerebellum	Blood
BA9		0.97	0.95	0.94	0.92	0.92	0.57	0.64
BA10	0.97		0.96	0.95	0.94	0.93	0.59	0.66
BA8	0.95	0.96		0.96	0.92	0.92	0.57	0.58
STG	0.94	0.95	0.96		0.93	0.92	0.58	0.59
Ent. Ctx	0.92	0.94	0.92	0.93		0.90	0.63	0.75
Vis. Ctx	0.92	0.93	0.92	0.92	0.90		0.57	0.60
Cerebellum	0.57	0.59	0.57	0.58	0.63	0.57		0.49
Blood	0.64	0.66	0.58	0.59	0.75	0.60	0.49	

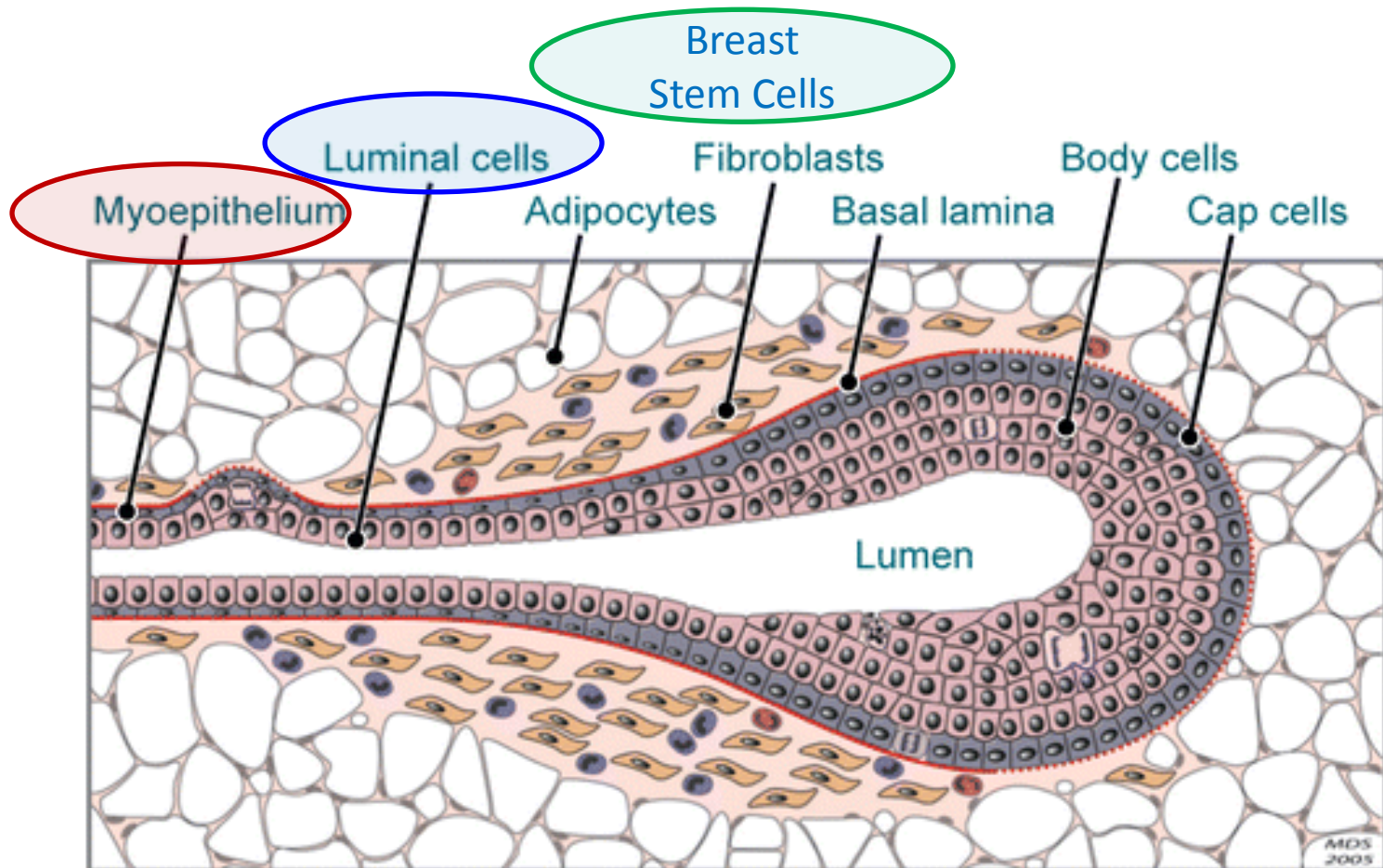
Hierarchical clustering discriminates tissues



Use Case 1: Genomewide Patterns of Methylation can Distinguish Between Blood, Cerebellum, and Cortex

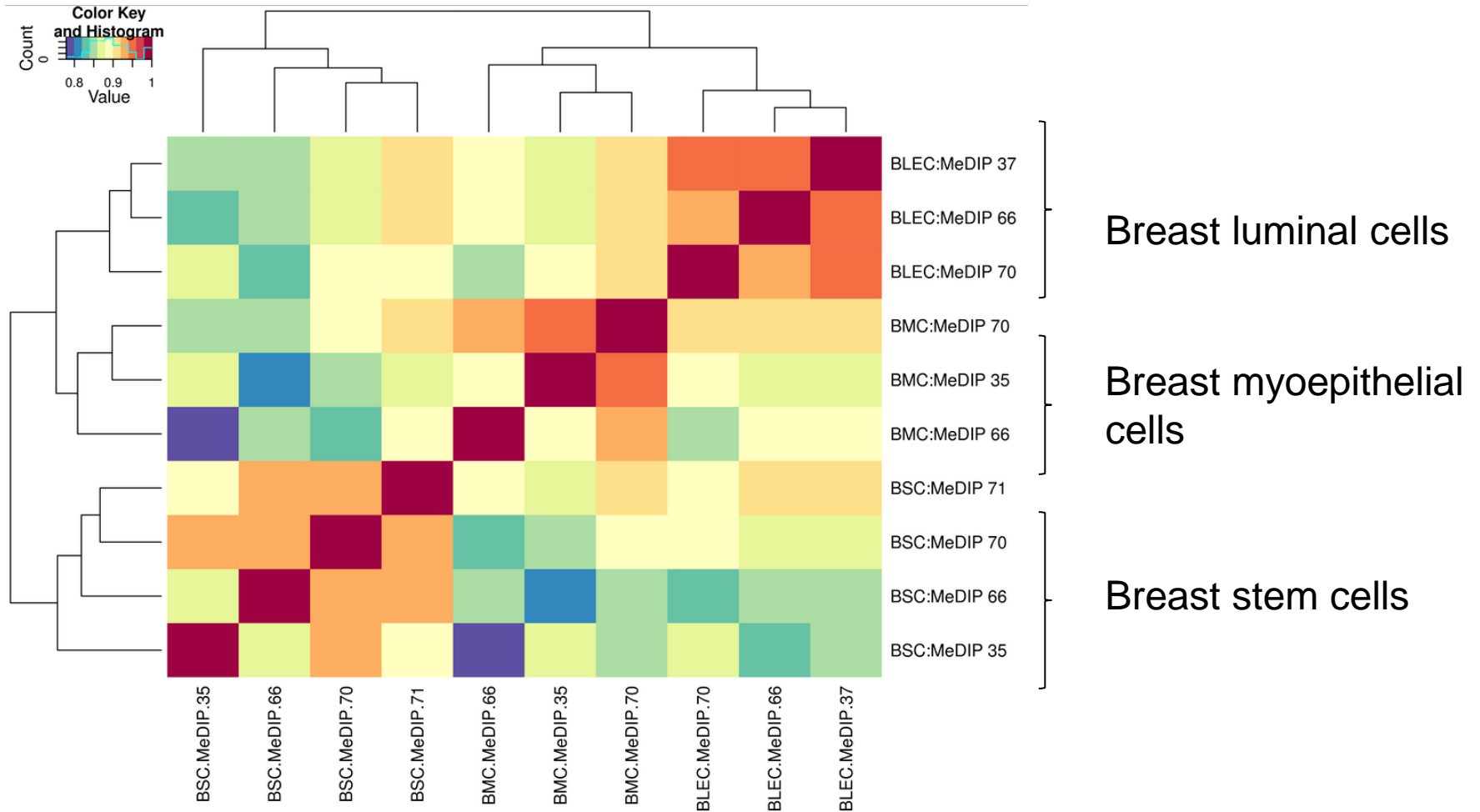


Use Case 2: Genomewide Patterns of Methylation can Distinguish Between Breast Cell Types



AR Hebner C, et al. 2008.
Annu. Rev. Pathol. Mech. Dis. 3:313–39

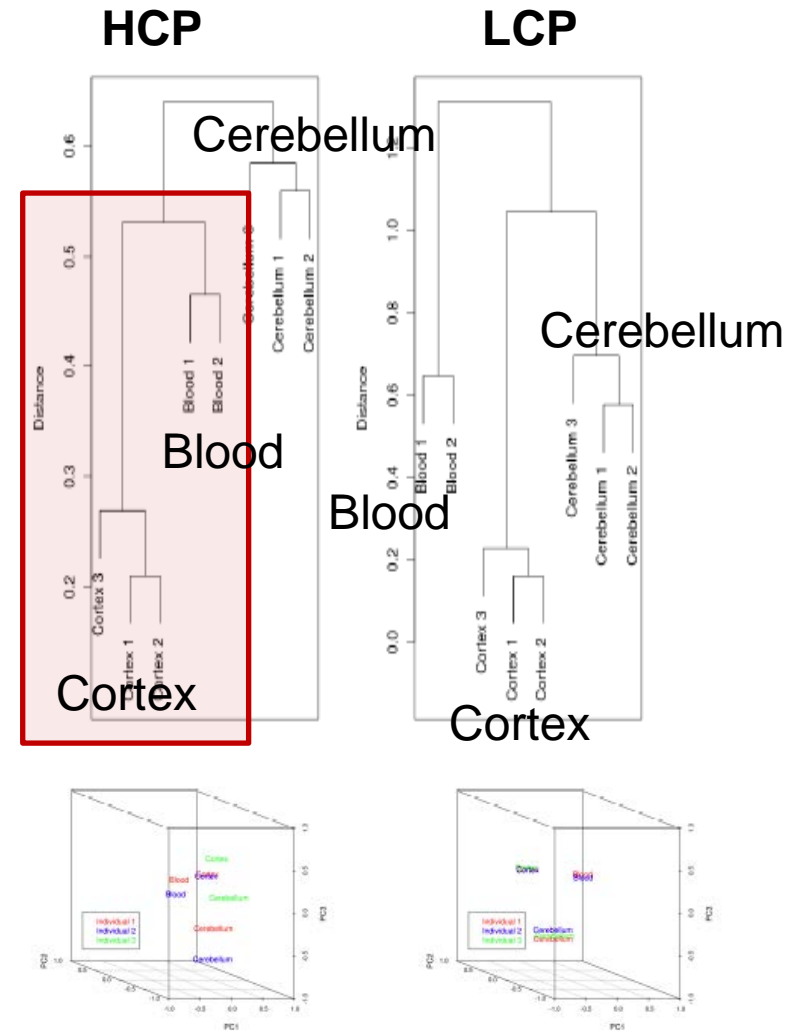
Use Case 2: Breast Cell Types Cluster Correctly Based on Their MeDIP-seq Profiles (Epigenome Atlas and UCSF REMC data)



Data from: Epigenome Atlas, Release 5

Use Case 5: Methylation Profiles of Some Features More Informative About Cell Type Than Profiles of Others

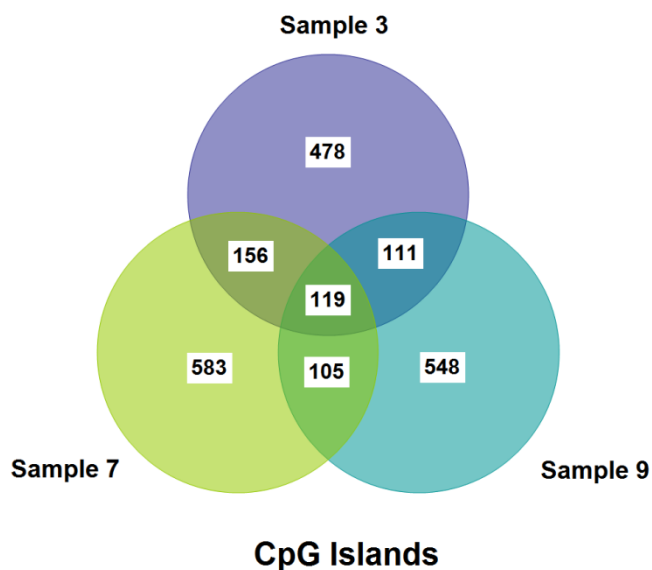
Methylation of LCPs conveys more information about tissue type than methylation of HCPs.



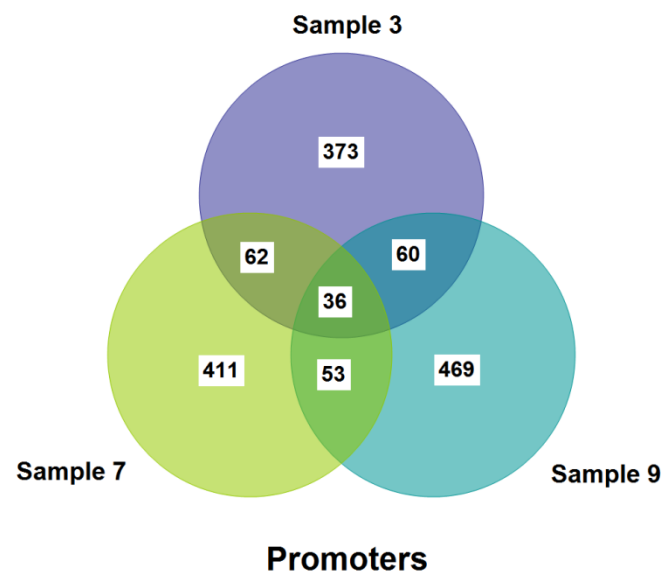
Use Case 10: Methylation changes during CD4+ T-cell maturation (Epigenome Atlas and UCSF REMC data)

CD4+ Memory Primary T cells / CD4+ Naïve Primary T cells methylation differences across 3 individuals

Genes with changes in associated CpG islands







Genes with changes in promoters



Use Case 10: Methylation changes during CD4+ T-cell maturation (Epigenome Atlas and UCSF REMC data)

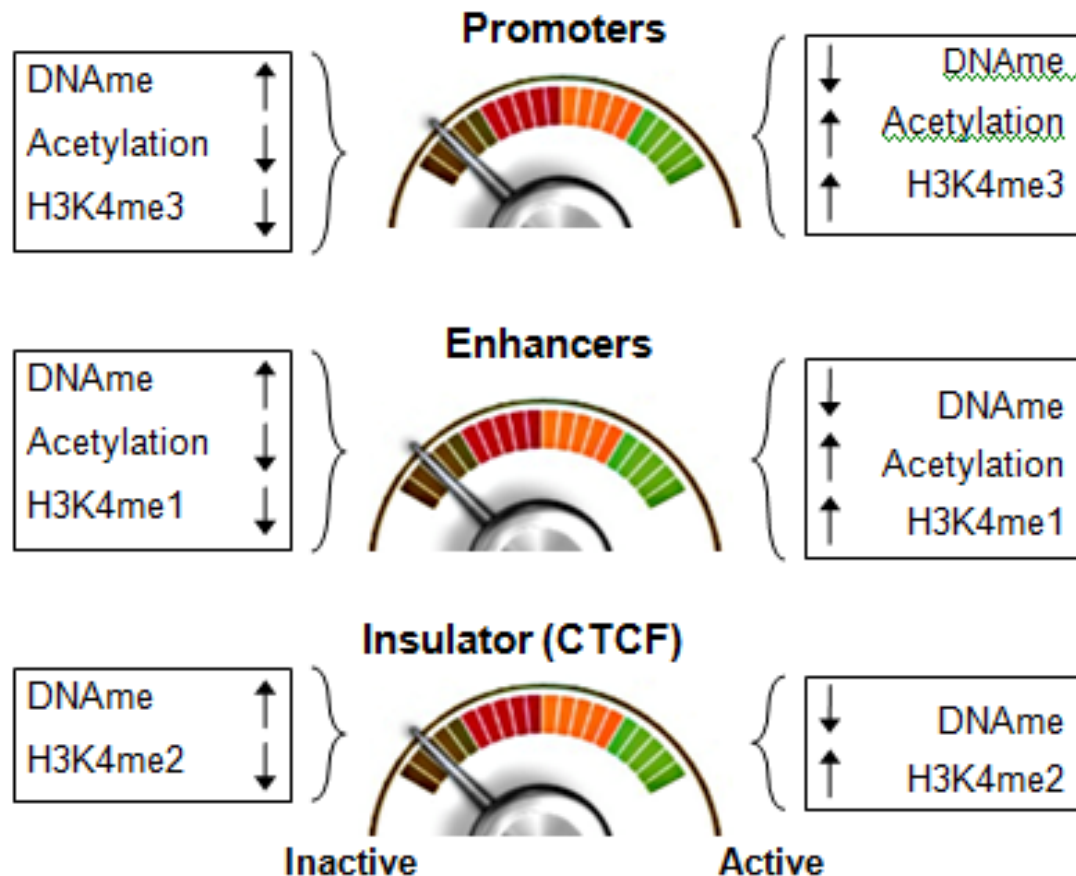
Methylation changes in **CpG islands**

KEGG pathways (via DAVID web site):

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
<input type="checkbox"/>	KEGG_PATHWAY →	T cell receptor signaling pathway	RT		5	5.9	2.3E-3	8.9E-2
<input type="checkbox"/>	KEGG_PATHWAY →	Natural killer cell mediated cytotoxicity	RT		5	5.9	4.9E-3	9.4E-2
<input type="checkbox"/>	KEGG_PATHWAY →	Wnt signaling pathway	RT		4	4.7	4.5E-2	4.6E-1
<input type="checkbox"/>	KEGG_PATHWAY	Pathways in cancer	RT		5	5.9	9.2E-2	6.2E-1

Use Case 9: Coordinated Changes of Epigenomic Marks Across Tissue Types

“Active” or “inactive” states of TF binding elements (promoters, enhancers) may be inferred with certain probability based on the state of a handful of correlated epigenomic marks.



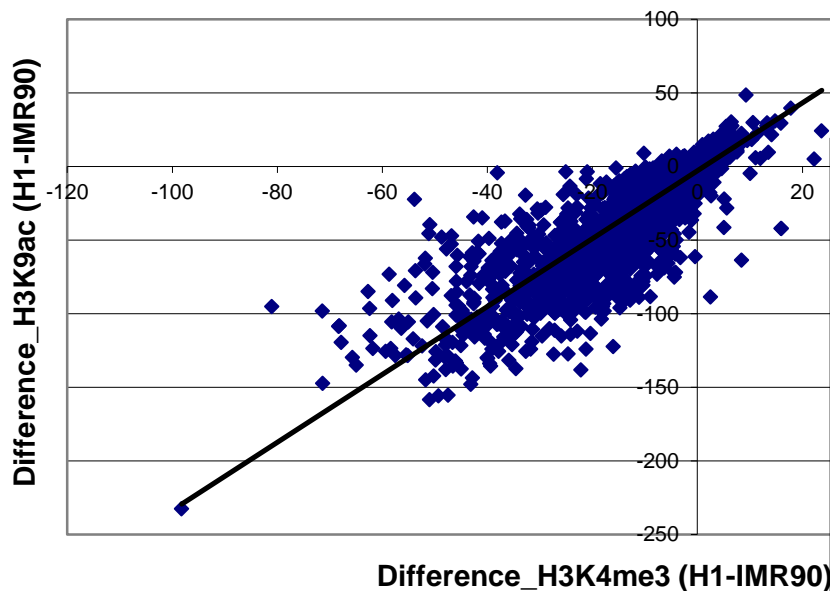
Use Case 9: Coordinated Changes of Epigenomic Marks Across Tissue Types

- High- and Intermediate CpG (HCP and ICP) promoters are regulated by different classes of TFs (including Polycomb complex) than Low-CpG (LCP) promoters.
- It is therefore possible that HCP, ICP, and LCP promoters experience different patterns of epigenomic changes during cellular differentiation.
- Approach: **Examine changes in H3K9ac and H3K4me3 over HCP, ICP, and LCP promoters between an ES cell line (H1) and a fibroblast cell line (IMR90).**

Use Case 9: Coordinated Changes of Epigenomic Marks Across Tissue Types

H3K9ac vs H3K4me3 (ICP)

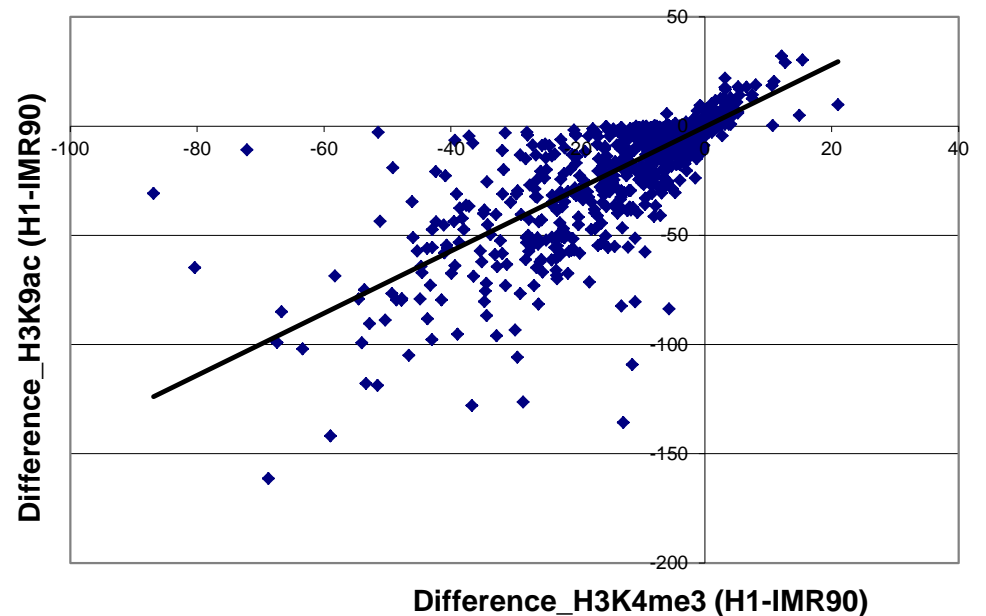
$$y = 2.306x - 2.8494$$
$$R^2 = 0.7975$$



Scatter Plots of H3K9ac and H3K4me3 Differences Between H1 and IMR90 Over ICP and LCP Promoters

H3K9ac vs H3K4me3 (LCP)

$$y = 1.4214x - 0.3166$$
$$R^2 = 0.6976$$



Tuesday, March 6th 2012

Session 4

2:00 – 3:30 pm **Case Study 7:** *Epigenomic Variation Between Individuals* – Matt Roth, Kevin Riehle, Chia-Chin Wu, Yuan Yuan, Cristi Coarfa, Alan Harris, Aleks Milosavljevic

3:30 – 4:30 pm *Case Studies: Epigenomic Variation in Cancer* – Matt Roth, Kevin Riehle, Chia-Chin Wu, Yuan Yuan, Cristi Coarfa, Alan Harris, Aleks Milosavljevic

4:30 – 5:00 pm Open discussion and wrap-up

5:00 pm Adjourn

Use Case 7: Detecting Coordinated Changes in Cortex, Cerebellum and Blood Between Individuals

Data Source: Matthew Davies, M. et al. (submitted) *"Tissue-specific epigenetic variation across brain and blood: functional annotation of the human brain methylome"*.

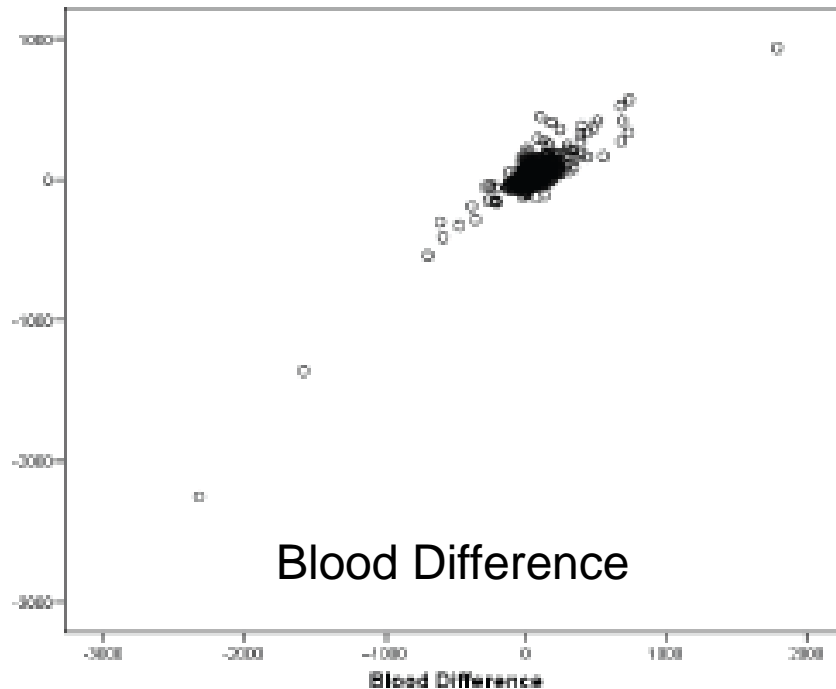
Use Case 7: Detecting Coordinated Changes in Cortex, Cerebellum and Blood Between Individuals

- To which extent the easily accessible peripheral tissue (e.g. whole blood) can be used to ask questions about **inter-individual phenotypic variation manifest in a phenotypically relevant but inaccessible tissue such as brain.**
- It is also of interest to delineate inter-individual differences that are genetic (say, due to copy-number polymorphisms) from those that may arise due to environmental, developmental, and physiological differences and disease processes.

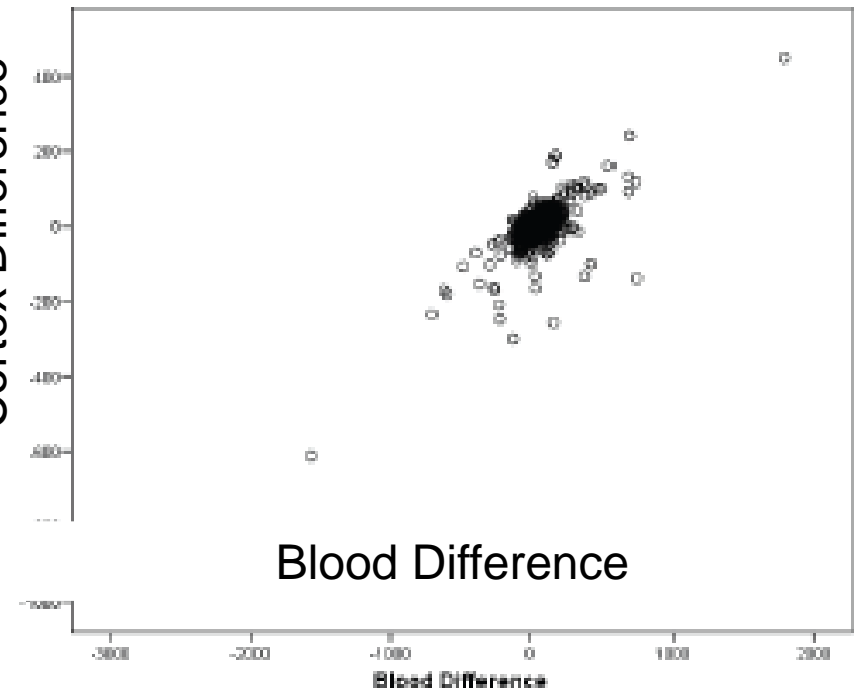
Use Case 7: Detecting Coordinated Changes in Cortex, Cerebellum and Blood Between Individuals

Between-individual variation in DNA methylation is correlated between blood and brain

Cerebellum Difference



Cortex Difference



Tuesday, March 6th 2012

Session 4

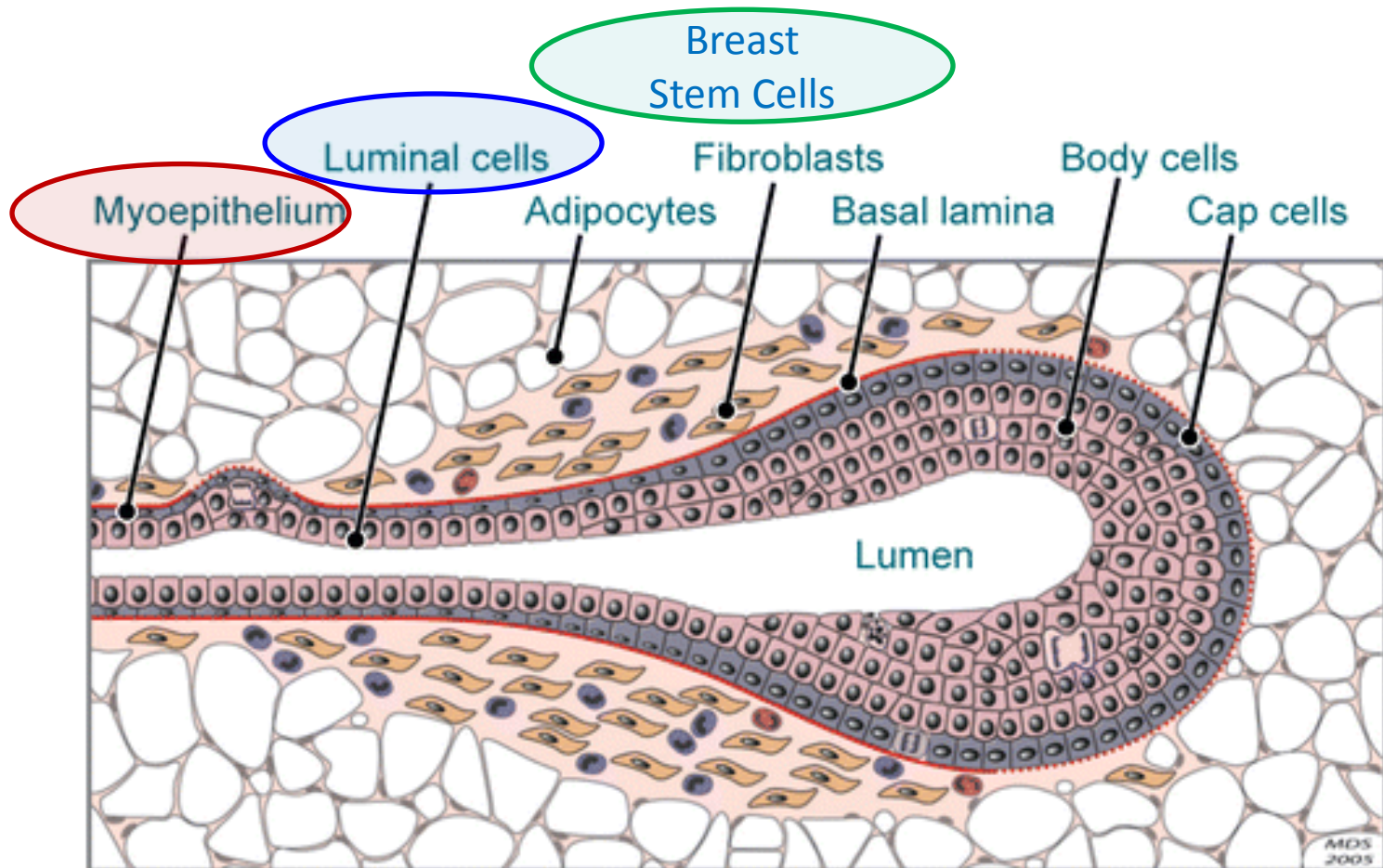
2:00 – 3:30 pm *Case Studies: Epigenomic Variation Between Individuals* – Matt Roth, Kevin Riehle, Chia-Chin Wu, Yuan Yuan, Cristi Coarfa, Alan Harris, Aleks Milosavljevic

3:30 – 4:30 pm ***Case Studies 12, 13:*** *Epigenomic Variation in Cancer* – Matt Roth, Kevin Riehle, Chia-Chin Wu, Yuan Yuan, Cristi Coarfa, Alan Harris, Aleks Milosavljevic

4:30 – 5:00 pm Open discussion and wrap-up

5:00 pm **Adjourn**

Use Case 12: Determining Cell Type of Origin of Breast Cancer



AR Hebner C, et al. 2008.
Annu. Rev. Pathol. Mech. Dis. 3:313–39

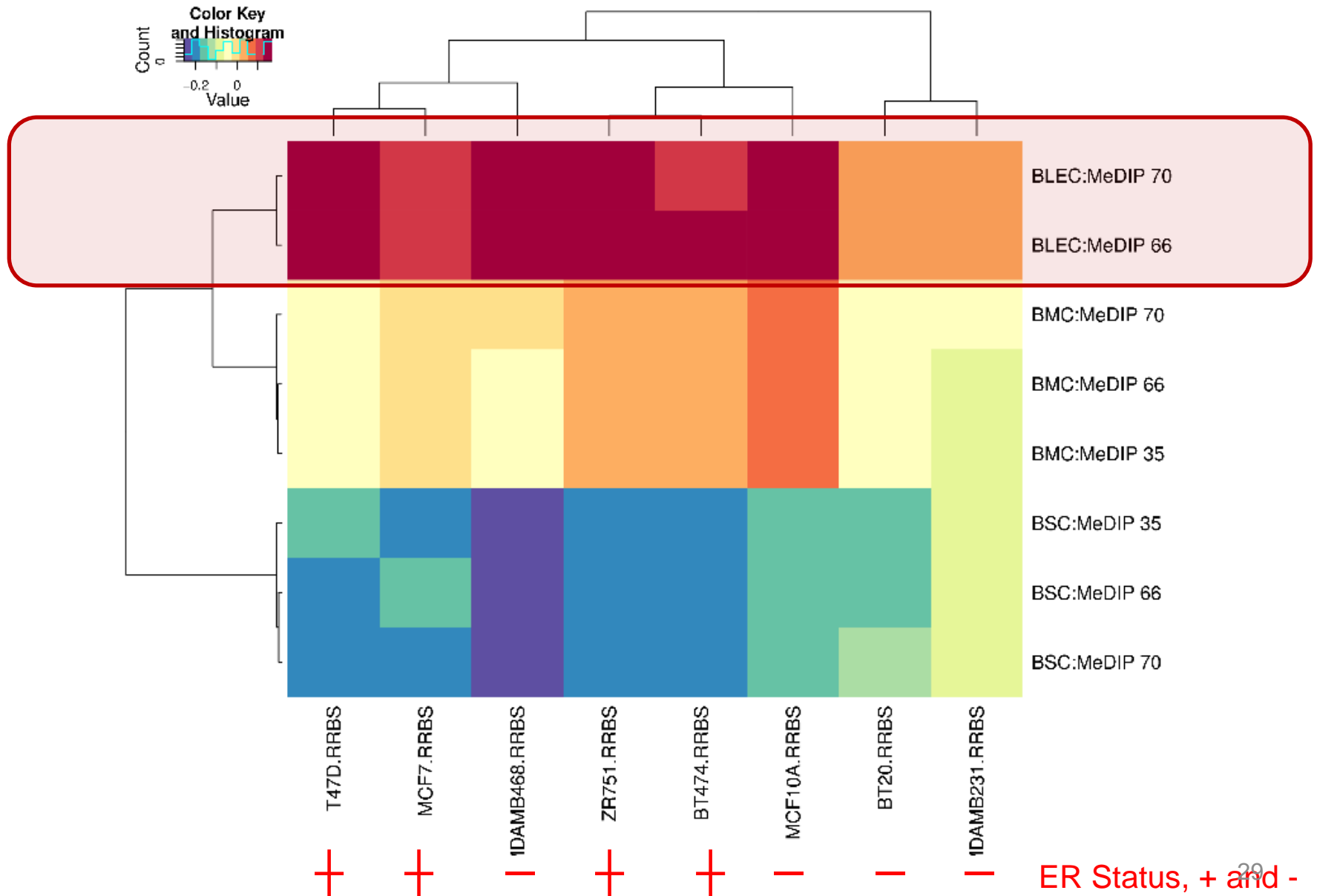
Use Case 12: Determining Cell Type of Origin of Breast Cancer

Compare profiles of 4 ER+ and 4 ER- breast cancer cell lines (RRBS profiles)¹ against the reference epigenomes of normal breast cell types (MeDIP-seq profiles) from the Human Epigenome Atlas.

Data Source:

¹ Sun, Z., et al. (2011) Integrated Analysis of Gene Expression, CpG Island Methylation, and Gene Copy Number in Breast Cancer Cells by Deep Sequencing. PLoS ONE 6(2): e17490.

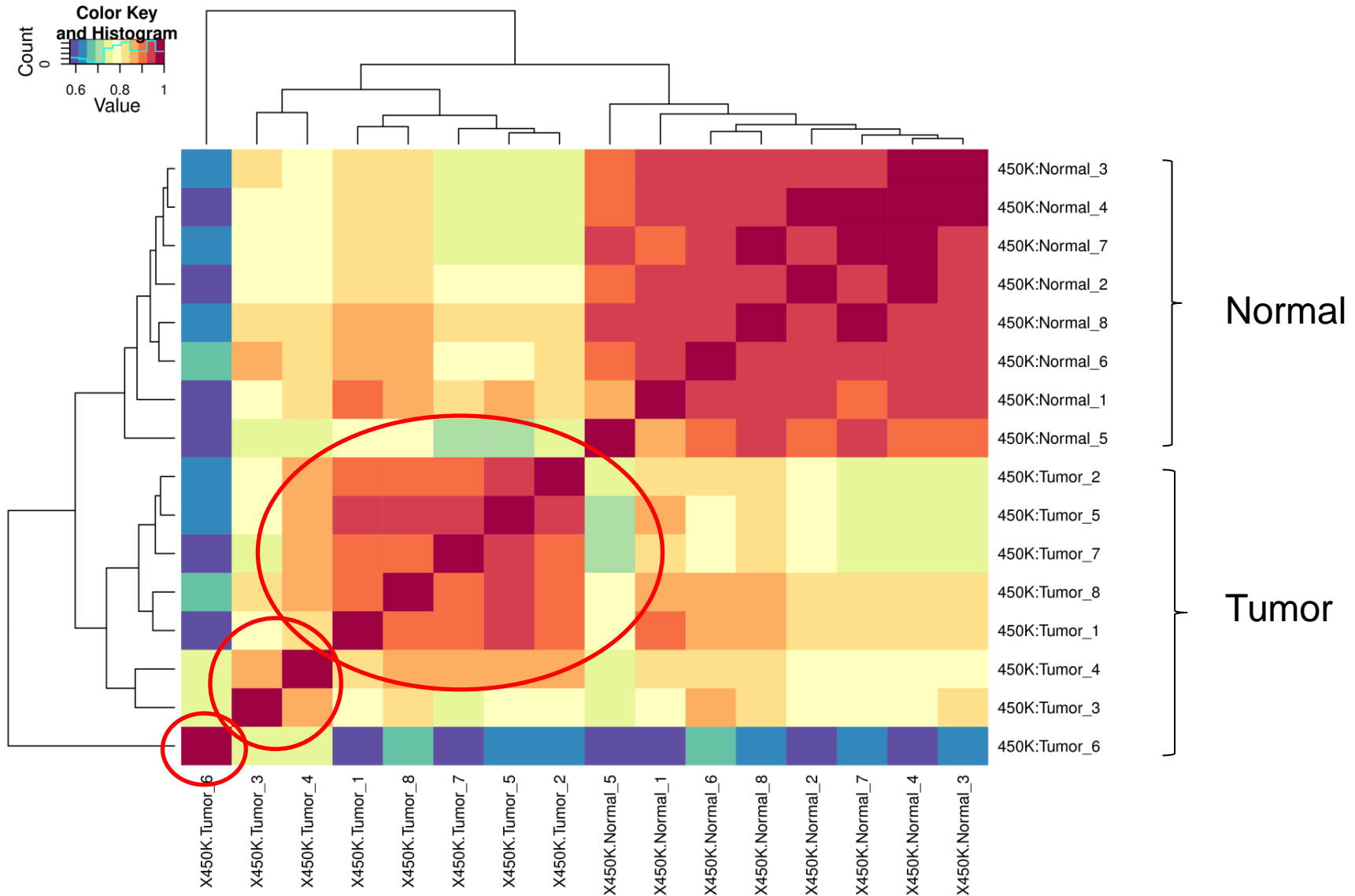
Use Case 12: Determining Cell Type of Origin of Breast Cancer



Use Case 13: Analysis of Epigenomic Variation in Breast Tumors

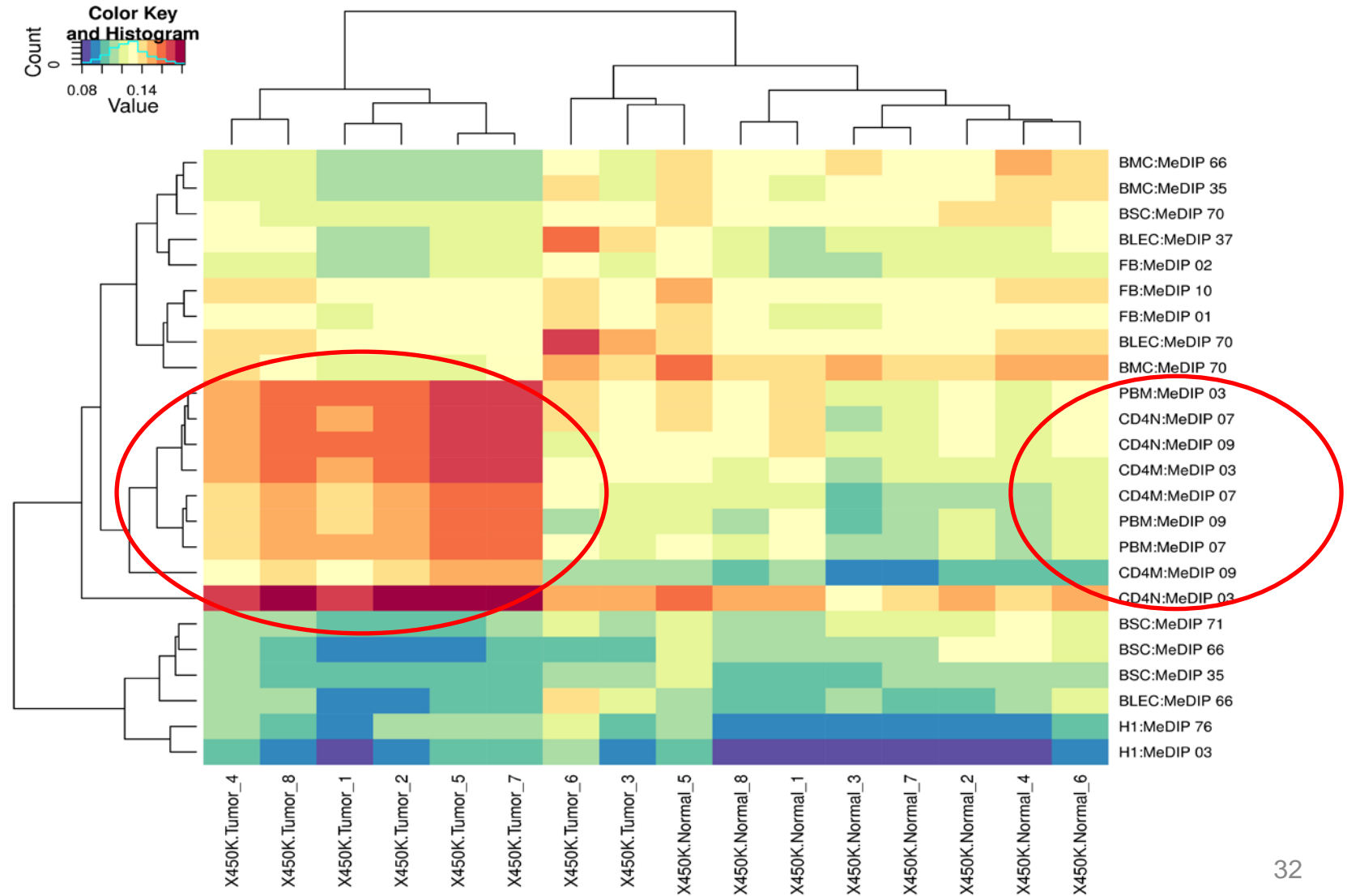
- Data Source: Dedeurwaerder, S. et al. (2011) "Evaluation of the Infinium Methylation 450K technology", Epigenomics 3(6):771-84.
- **16 breast tissue samples were profiled** (8 normal, 8 primary tumor samples)
- The paper evaluates 450K technology and does not report any analysis concerning cancer biology.
- We analyze the data to explore epigenomic states and cell type composition of normal and tumor samples.

Use Case 13: Analysis of Epigenetic Variation in Breast Tumors



Use Case 13: Analysis of Epigenomic Variation in Breast Tumors

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



Use Case 13: Analysis of Epigenomic Variation in Breast Tumors

- Most breast tumor samples appear to contain an excess of blood and immune cells.
- Comparison of normal and tumor tissue should therefore reveal **differentially methylated genes that are involved in immunity-related pathways or biological processes.**

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

Tuesday, March 6th 2012

Session 4

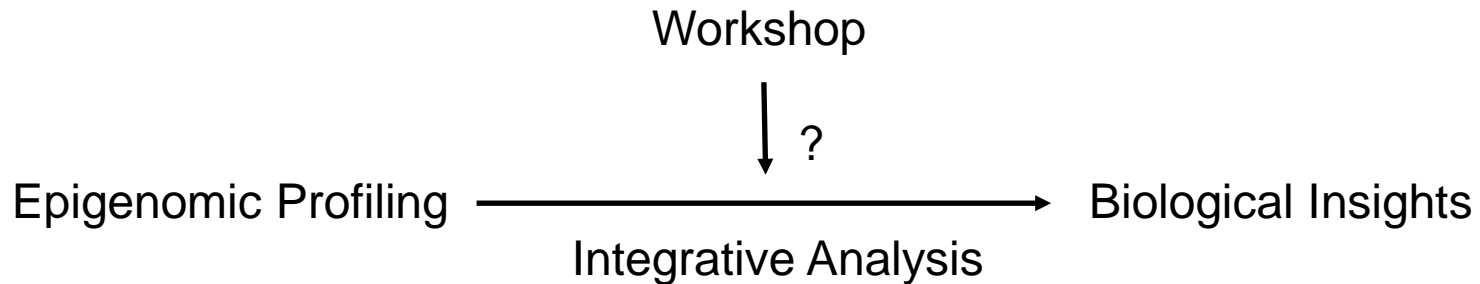
2:00 – 3:30 pm *Case Studies: Epigenomic Variation Between Individuals* – Matt Roth, Kevin Riehle, Chia-Chin Wu, Yuan Yuan, Cristi Coarfa, Alan Harris, Aleks Milosavljevic

3:30 – 4:30 pm *Case Studies: Epigenomic Variation in Cancer* – Matt Roth, Kevin Riehle, Chia-Chin Wu, Yuan Yuan, Cristi Coarfa, Alan Harris, Aleks Milosavljevic

4:30 – 5:00 pm Open discussion and wrap-up

5:00 pm Adjourn

Workshop Evaluation



- Methods (Assays, Data Processing)
- Standards (Metadata, Interoperability)
- Data Resources (Human Epigenome Atlas)
- Tools (Epigenomic Toolset, Genboree Workbench)
- Use Cases / Case Studies
- Collaborative Opportunities / Networking / Exchange of Experience

NIH Roadmap Epigenomics:
Reference Epigenomes
and the Human Epigenome Atlas



NIH Roadmap
Epigenomics Data Analysis and Coordination Center

March 5-6 2012
Houston, Texas

NIH Roadmap Epigenomics Project

Hypothesis:

Origins of health and susceptibility to disease are, in part, result of epigenetic regulation

Goal:

Transform biomedical research by

- **Developing comprehensive reference epigenome maps**
- **Developing new technologies for epigenomic analyses**

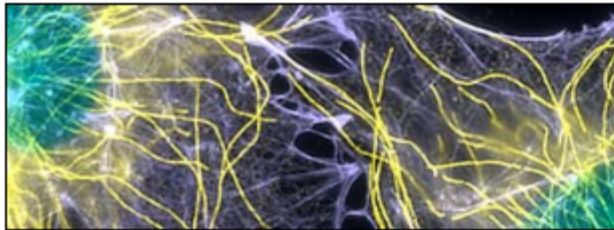
Including cyberinfrastructure for epigenomic research

NIH Roadmap Epigenomics Project: Data Flow



Quarterly cumulative releases of the Human Epigenome Atlas
(Human Epigenome Atlas Release 6 being completed)

Epigenomics Portal at NCBI



Epigenomics

Genomics maps of stable, yet reprogrammable nuclear changes that control gene expression and influence our health.

How to...

[How to Use the Sample Browser](#)

[How to Manage Collections of Samples](#)

[How to View Genome Tracks](#)

[How to Download Genome Tracks](#)

Epigenomics Tools

[Browse Experiments](#)

[Browse Samples](#)

[Compare Samples Beta](#)

[Advanced Search](#)

Scientific Background

[About Epigenetics](#)

[About DNA Methylation](#)

[About Histone Modifications](#)

[About Chromatin Structure](#)

Latest Studies

CTCF Binding Sites by ChIP-seq from ENCODE/University of Washington [ESS000141]

Genome-wide remodeling of the epigenetic landscape during myogenic differentiation [ESS000133]

Histone Modifications by ChIP-seq from ENCODE/Broad Institute [ESS000130]

DNaseI hypersensitivity by digital DNaseI from ENCODE/University of Washington [ESS000129]

The Cohesin Complex Cooperates with Pluripotency Transcription Factors in the Maintenance of Embryonic Stem Cells [ESS000137]

[See more...](#)

NIH Roadmap Epigenomics



About the Project
Reference epigenome maps and their applications to human health.

[Roadmap Epigenomics Web Resources](#)

[Roadmap Data in GEO](#)

[Data Access Policies](#)

Recent Review Articles

Epigenetics in disease and cancer. [Malays J Pathol. 2011]

Role of TGF- β and the tumor microenvironment during mammary tumorigenesis. [Gene Expr. 2011]

[Epigenetic regulation of transcription in heart development]. [Seikagaku. 2011]

[Neurobiology of suicidal behaviour]. [Psychiatr Pol. 2011]

Histone methylation makes its mark on longevity. [Trends Cell Biol. 2012]

[See more...](#)

Human Epigenome Browser at Wash U



Epigenomic Data

	“chip” data	“seq” data
Level 0	image	reads
Level 1	extracted features	mapped reads
Level 2	normalized intensities	read density maps
Level 3	epigenomic state (per feature such as enhancer or genomic segments outside of features)	
Level 4	comparative analysis results (e.g., cell type – specific state)	

Epigenomic Data Standards

www.ihec-epigenomes.org



The screenshot shows the IHEC website with a blue border. The main header features the IHEC logo (a globe with a blue arc) and the text "IHEC International Human Epigenome Consortium". Below the header is a navigation bar with orange buttons: "Home", "Areas of focus", "Standard Operating Procedures", "Tools / Useful Information", "Policies and Guidelines", "IHEC Structure", "Outreach and Training", "Feedback", and "Intranet". The "Standard Operating Procedures" button is circled in red. On the left, a "Links" section includes logos for "The EPIGENOME Network of Excellence" and "Centre for Epigenetics". The main content area is titled "IHEC Recommendations for Epigenomic Analysis (DRAFT)". It contains a "Download" section with text about data and metadata models developed by the NIH Roadmap Epigenomics Project, and a list of standards: "RNA-seq Standards", "ChIP-seq Standards", and "MethylC-seq Standards". The entire content area is circled in red.

IHEC
International Human Epigenome Consortium

Home Areas of focus **Standard Operating Procedures** Tools / Useful Information Policies and Guidelines IHEC Structure Outreach and Training Feedback Intranet

Links

 The EPIGENOME Network of Excellence
The focal point for the European epigenetics research community

 Centre for Epigenetics
Centre of Excellence funded by The Danish National Research Foundation.

IHEC Recommendations for Epigenomic Analysis (DRAFT)

Download

Data and Metadata Models
Developed by the NIH Roadmap Epigenomics Project

(Under Consideration by the International Human Epigenome Consortium (IHEC) as a Basis for IHEC Recommendation)

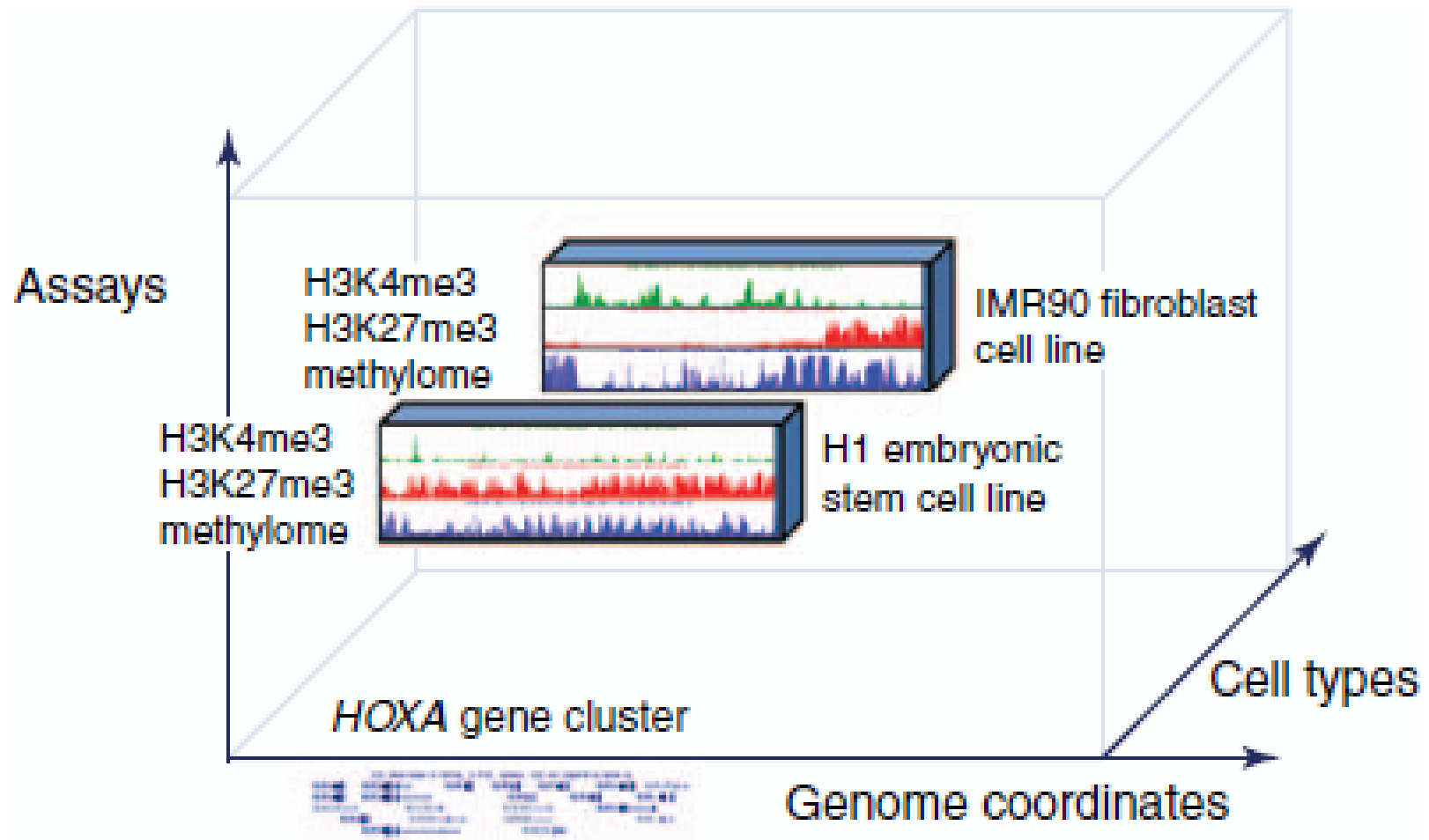
These assay standards developed/in use by the NIH Roadmap Reference Epigenome Mapping Centers are being considered for potential use by IHEC.

RNA-seq Standards

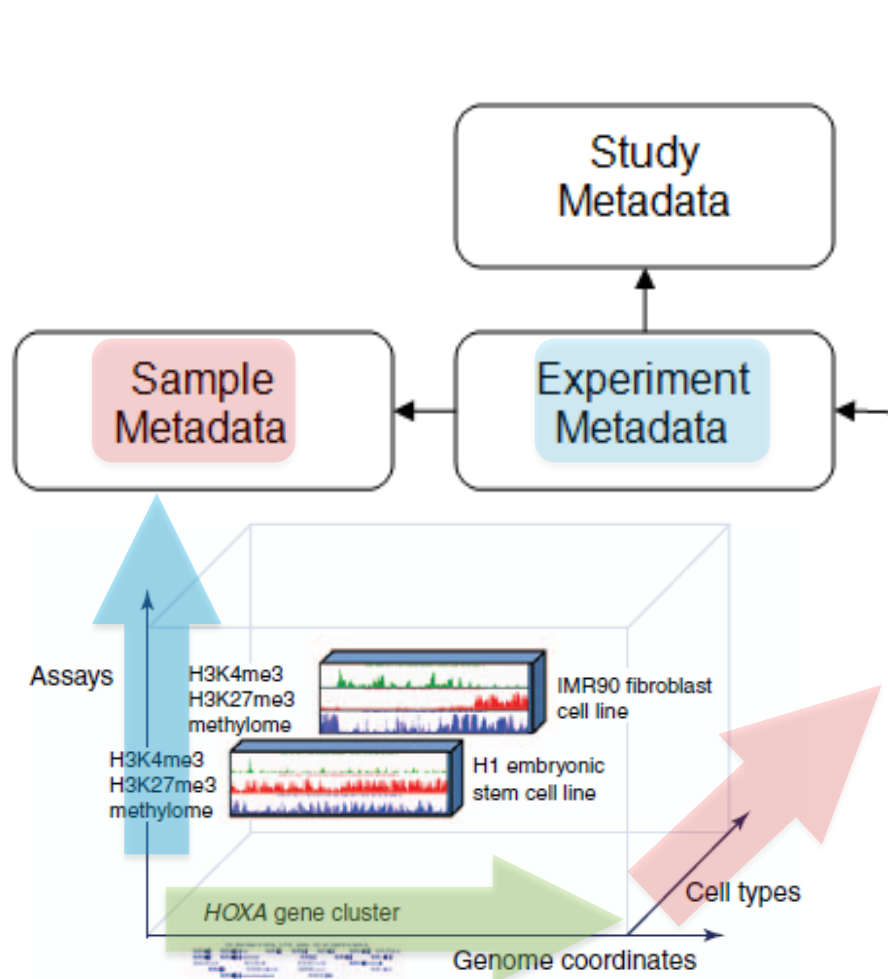
ChIP-seq Standards

MethylC-seq Standards

www.epigenomeatlas.org



Epigenomic Metadata



Level	Data	Metadata
0	Fastq / SRF	Run

Level	Analysis	Metadata
1	BED / BAM	Software Parameters QC
2	Wig / Counts	Software Parameters QC
3	Peaks / GFF3 / LFF	Software Parameters QC
4	Comparative Analysis / GFF3 / LFF	Software Parameters QC

Epigenomic MetaData Standards

www.ihec-epigenomes.org

The screenshot shows the IHEC website with a blue and white patterned border. The IHEC logo, featuring a globe and the acronym 'IHEC', is at the top center, with the full name 'International Human Epigenome Consortium' below it. A horizontal navigation bar contains several orange buttons: 'Home', 'Areas of focus', 'Standard Operating Procedures' (circled in red), 'Tools / Useful Information', 'Policies and Guidelines', 'IHEC Structure', 'Outreach and Training', 'Feedback', and 'Intranet'. On the left, a 'Links' section lists 'The EPIGENOME Network of Excellence' and 'Centre for Epigenetics'. The main content area is titled 'IHEC Recommendations for Epigenomic Analysis (DRAFT)'. It includes a 'Download' section with the text 'Data and Metadata Models Developed by the NIH Roadmap Epigenomics Project' and '(Under Consideration by the International Human Epigenome Consortium (IHEC) as a Basis for IHEC Recommendation)'. To the right, there is a box titled 'These assay standards developed/in use by the NIH Roadmap Reference Epigenome Mapping Centers are being considered for potential use by IHEC.' containing links for 'ChIP-seq Standards', 'ChIP-seq Standards', and 'MethylC-seq'. A red circle highlights the 'Download' section, and a red arrow points from it to a separate box on the right.

IHEC
International Human Epigenome Consortium

Home Areas of focus **Standard Operating Procedures** Tools / Useful Information Policies and Guidelines IHEC Structure Outreach and Training Feedback Intranet

Links

- The EPIGENOME Network of Excellence
The focal point for the European epigenetics research community
- Centre for Epigenetics
Centre of Excellence funded by The Danish National Research Foundation

IHEC Recommendations for Epigenomic Analysis (DRAFT)

Download

Data and Metadata Models Developed by the NIH Roadmap Epigenomics Project

(Under Consideration by the International Human Epigenome Consortium (IHEC) as a Basis for IHEC Recommendation)

These assay standards developed/in use by the NIH Roadmap Reference Epigenome Mapping Centers are being considered for potential use by IHEC.

ChIP-seq Standards

ChIP-seq Standards

MethylC-seq

Data and Metadata Models Developed by the NIH Roadmap Epigenomics Project

(Under Consideration by the International Human Epigenome Consortium (IHEC) as a Basis for IHEC Recommendation)

R. Alan Harris¹, Robert A. Waterland^{1,2}, Ryan Lister³, Martin Hirst⁴, Aleksandar Milosavljevic¹

¹Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas

²USDA/ARS Children's Nutrition Research Center, Department of Pediatrics, Baylor College of Medicine, Houston, Texas, USA.

³Genomic Analysis Laboratory, The Salk Institute for Biological Studies, La Jolla, California, USA.

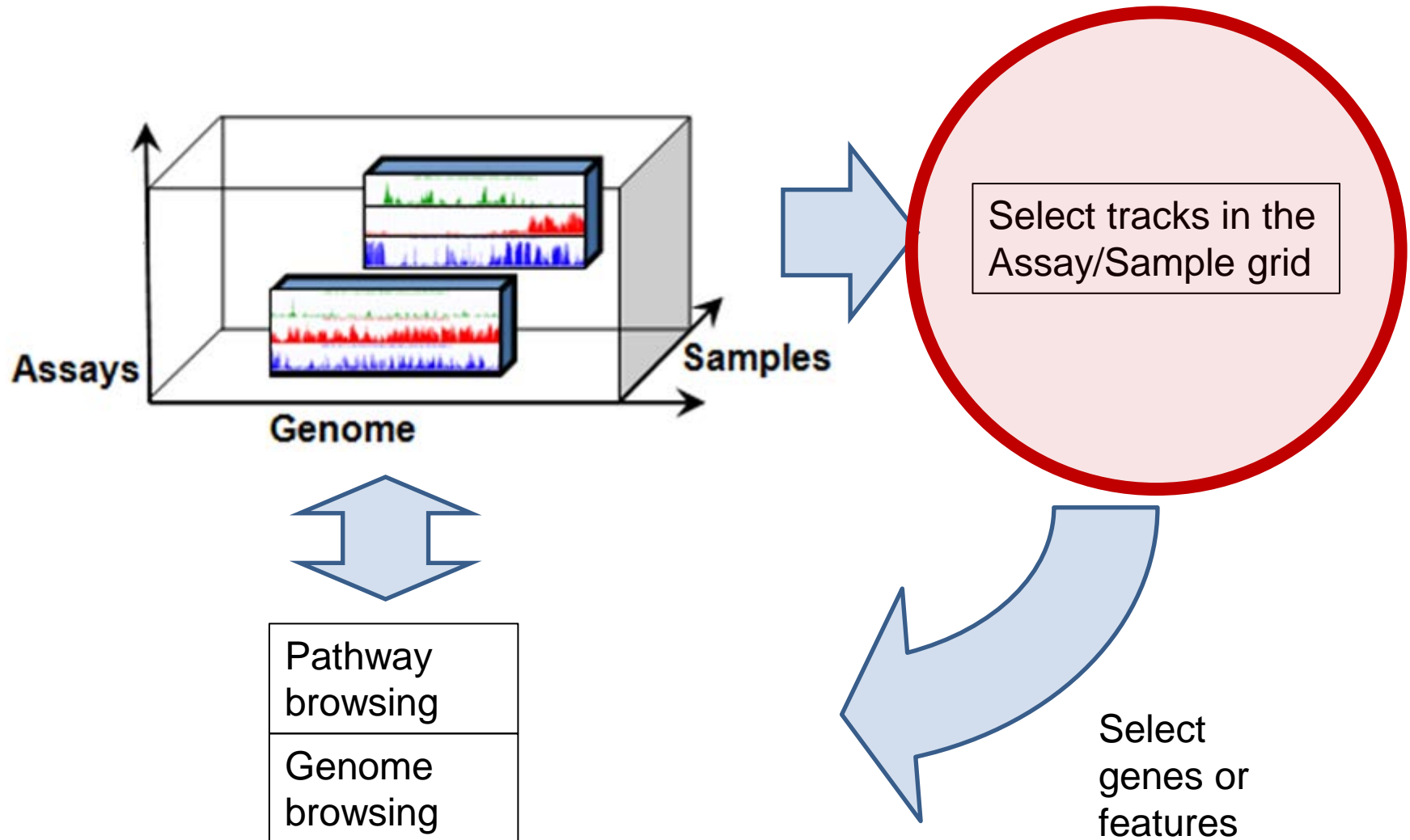
⁴Genome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia, Canada.

1. Introduction

The data produced by the NIH Epigenomics Roadmap Initiative is illustrated in Figure 1

Epigenome Atlas Grid

www.epigenomeatlas.org



www.epigenomeatlas.org

[Home](#)

Releases





Informatics

Publications

Forums

Contributors

Human Epigenome Atlas [Release 4](#) (hg19)

- [Data Access Policy](#)
- Data embargo period: from 04/14/2011 - 01/14/2012 or earlier as specified [here](#)
- Select cells by **clicking and dragging**, then use the "View Selections in" pulldown in the top left corner (below) to view selections in the Atlas Gene Browser or the UCSC Genome Browser
- To see data authors, other metadata, and to download data, click a sample name in the first column or an assay type in the header row
- Human Epigenome Atlas releases are intended to be cumulative: e.g. Release 3 includes all Release 2 data and additional submissions
- **NOTE:** Some pages may not be accessible over low bandwidth internet connections. This page has been tested with the following browsers:    

Human Epigenome Atlas Release 4 (hg19)

 View Selections In ▾ Clear Selections

Assay¹

Sample

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

[illegible]

Epigenome Atlas Release 5

www.epigenomeatlas.org



Home



Releases





Informatics

Publications



Forums


Contributors

Human Epigenome Atlas Release 4 (hg19)

- [Data Access Policy](#)
- Data embargo period: from 04/14/2011 - 01/14/2012 or earlier as specified [here](#)
- Select cells by clicking and dragging, then use the "View Selections in" pulldown in the top left corner (below) to view selections in the Atlas Gene Browser or the UCSC Genome Browser
- To see data authors, other metadata, and to download data, click a sample name in the first column or an assay type in the header row
- Human Epigenome Atlas releases are intended to be cumulative: e.g. Release 3 includes all Release 2 data and additional submissions
- NOTE: Some pages may not be accessible over low bandwidth internet connections. This page has been tested with the following browsers:    

Human Epigenome Atlas Release 4 (hg19)

 View Selections in  Clear Selections



Assay 																																									
Sample	Filter	Bisulfite-Seq	MeDIP-Seq	MRE-Seq	RRBS	DNAse Hypersensitivity	Digital Genomic Footprinting	mRNA-Seq	smRNA-Seq	Expression Array	ChIP-Seq Input	Histone H3K27me3	Histone H3K36me3	Histone H3K4me1	Histone H3K4me3	Histone H3K9ac	Histone H3K9me3	Histone H2AK5ac	Histone H2BK5ac	Histone H2BK120ac	Histone H2BK12ac	Histone H2BK15ac	Histone H2BK20ac	Histone H3K14ac	Histone H3K18ac	Histone H3K23ac	Histone H3K23me2	Histone H3K27ac	Histone H3K4ac	Histone H3K4me2	Histone H3K56ac	Histone H3K79me1	Histone H3K79me2	Histone H4K20me1	Histone H4K5ac	Histone H4K8ac	Histone H4K91ac				
Brain Substantia Nigra											1	1	1	1	1	1	1																								
Breast Luminal Epithelial Cells		4	5					2				1	1	1				1																							
Breast Myoepithelial Cells		3	3					2			2	2	2	2	2	2	2	2																							
Breast Stem Cells		4	4					1																																	
Breast vHMEC		1	1		2			1	1		2	1	1	2	1			1																							
CD14 Primary Cells						2																																			
CD15 Primary Cells					1							1			1																										
CD19 Primary Cells					1	3					1	2	2		2		2																								

48

Epigenome Atlas Release 5

www.epigenomeatlas.org

Human Epigenome Atlas Release 4 (hg19)

- [Data Access Policy](#)
- Data embargo period: from 04/14/2011 - 01/14/2012 or earlier as specified [here](#)
- Select cells by clicking and dragging, then use the "View Selections In" pulldown in the top left corner (below) to view selections in the Atlas Gene B
- To see data authors, other metadata, and to download data, click a sample name in the first column or an assay type in the header row
- Human Epigenome Atlas releases are intended to be cumulative: e.g. Release 3 includes all Release 2 data and additional submissions
- NOTE: Some pages may not be accessible over low bandwidth Internet connections. This page has been tested with the following browsers:  

Human Epigenome Atlas Release 4 (hg19)

 View Selections In  Clear Selections

Assay															
Sample	Filter: <input type="text"/> (e.g. "cell line")	Bisulfite-Seq	MeDIP-Seq	MRE-Seq	RRBS	DNase Hypersensitivity	Digital Genomic Footprinting	mRNA Seq	smRNA Seq	Expression Array	ChIP-Seq Input	Histone H3K27me3	Histone H3K36me3	Histone H3K4me1	
Brain Substantia Nigra											1	1	1	1	
Breast Luminal Epithelial Cells													1	1	
Breast Myoepithelial Cells													2	2	
Breast Stem Cells															
Breast vHMEC													1	2	
CD14 Primary Cells						2									
CD15 Primary Cells					1							1			
CD19 Primary Cells					1	3					1	2	2		

Breast Luminal Epithelial Cells 

 Toggle sample selections

 [Download data](#)

 [See Metadata](#)

Epigenome Atlas Release 5

www.epigenomeatlas.org

Human Epigenome Atlas Release 4 (hg19)

- [Data Access Policy](#)
- Data embargo period: from 04/14/2011 - 01/14/2012 or earlier as specified [here](#)
- Select of Submissions for Breast Luminal Epithelial Cells
- To see data
- Human E
- NOTE: S

Human E

View S

Sample

Filter

UCSF-UBC

Histone H3K27me3

BLEC-H3K27me3 80 Accession coming soon!

UCSF-UBC

MeDIP-Seq

BLEC:MeDIP 35 [GSM493615](#)

BLEC:MeDIP 70 [GSM613843](#)

BLEC:MeDIP 37 [GSM613864](#)

BLEC:MeDIP 66 [GSM613856](#)

MRE-Seq

BLEC:MRE 66 [GSM613833](#)

BLEC:MRE 35 [GSM493620](#)

BLEC:MRE 70 [GSM613818](#)

BLEC:MRE 37 [GSM493617](#)

BLEC:MRE 71 [GSM613826](#)

mRNA-Seq

BLEC:mRNA 35 [GSM543029](#)

BLEC:mRNA 80 [GSM669620](#)

Histone H3K36me3

BLEC-H3K36me3 80 [GSM669597](#)

Histone H3K4me1

BLEC-H3K4me1 80 [GSM669595](#)

Histone H3K9me3

BLEC-H3K9me3 80 [GSM669596](#)

down in the top left corner (below) to view selections in the Atlas Gene B
e in the first column or an assay type in the header row
e 3 includes all Release 2 data and additional submissions
ctions. This page has been tested with the following browsers:

Bisulfite-Seq	MeDIP-Seq	MRE-Seq	RRBS	DNase Hypersensitivity	Digital Genomic Footprinting	mRNA-Seq	smRNA-Seq	Expression Array	ChIP-Seq Input	Histone H3K27me3	Histone H3K36me3	Histone H3K4me1
---------------	-----------	---------	------	------------------------	------------------------------	----------	-----------	------------------	----------------	------------------	------------------	-----------------

Brain Substantia Nigra

Breast Luminal Epithelial Cells

Breast Myoepithelial Cells

Breast Stem Cells

Breast vHMEC

CD14 Primary Cells

CD15 Primary Cells

CD19 Primary Cells

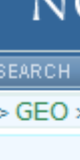

Breast Luminal Epithelial Cells

Toggle sample selections

[Download data](#)

[See Metadata](#)

www.epic.org

[HOME](#)
[SEARCH](#)
[SITE MAP](#)

[GEO Publications](#)

NCBI > [GEO](#) > **Accession Display** [?](#)

GEO help: Mouse over screen elements for information.

Scope: Format: Amount: GEO accession:

Sample GSM669597 [Query DataSets](#)

Status	Public on Mar 31, 2011
Title	H3K36me3 Luminal Epithelial Cell; HS2795-1
Sample type	SRA
Source name	Breast, Luminal Epithelial Cells, RM080; HS2795-1
Organism	Homo sapiens
Characteristics	sample alias: Breast Luminal Epi RM080 sample common name: Breast, Luminal Epithelial Cells disease: None biomaterial_provider: Thea Tlsty Lab UCSF biomaterial_type: Primary Cell cell_type: Luminal Epithelial Cell markers: CD227+ donor_id: RM080 donor_age: 33 years donor_health_status: Disease Free donor_sex: Female donor_ethnicity: African-American passage_if_expanded: NA karyotype: 46, XX

Viewing selections

Human Epigenome Atlas Release 4 (hg19)

- [Data Access Policy](#)
- Data embargo period: from 04/14/2011 - 01/14/2012 or earlier as specified [here](#)
- Select cells by clicking and dragging, then use the "View Selections In" pulldown in the top left corner (below) to view
- To see data authors, other metadata, and to download data, click a sample name in the first column or an assay type in
- Human Epigenome Atlas releases are intended to be cumulative: e.g. Release 3 includes all Release 2 data and add
- NOTE: Some pages may not be accessible over low bandwidth Internet connections. This page has been tested with

Human Epigenome Atlas Release 4 (hg19)

View Selections In ✖ Clear Selections

Atlas Gene Browser

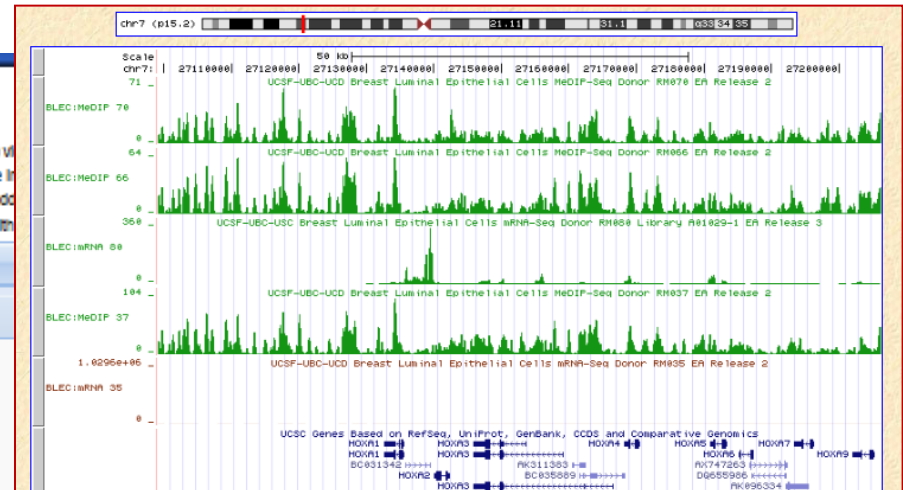
Genome Browser ▶ Choose a Mirror

- Local UCSC browser mirror (Fast)
- UCSC genome browser (Slow)

Assay

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

Sample	Bisulfite-Seq	MeDIP-Seq	MRE-Seq	RRBS	DNAse Hypersensitivity	Digital Genomic Footprinting	mRNA-Seq	smRNA-Seq	Expression
Brain Substantia Nigra									
Breast Luminal Epithelial Cells	4	5					2		
Breast Myoepithelial Cells	3	3					2		
Breast Stem Cells	4	4					1		
Breast vHMEC		1	1		2		1	1	
CD14 Primary Cells					2				
CD15 Primary Cells				1					
CD19 Primary Cells				1	3				
CD20 Primary Cells					1				



Viewing selections

Human Epigenome Atlas Release 4 (hg19)

- [Data Access Policy](#)
- Data embargo period: from 04/14/2011 - 01/14/2012 or earlier as specified [here](#)
- Select cells by clicking and dragging, then use the "View Selections In" pull-down in the top left corner (below) to
- To see data authors, other metadata, and to download data, click a sample name in the first column or an assay type
- Human Epigenome Atlas releases are intended to be cumulative: e.g. Release 3 includes all Release 2 data and a
- NOTE: Some pages may not be accessible over low bandwidth Internet connections. This page has been tested w

Human Epigenome Atlas Release 4 (hg19)

View Selections In ✖ Star Selections

Atlas Gene Browser

Genome Browser ▶ Choose a Mirror

- Local UCSC browser mirror (Fast)
- UCSC genome browser (Slow)

Sample

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

Sample	Brain Substantia Nigra	Breast Luminal Epithelial Cells	Breast Myoepithelial Cells	Breast Stem Cells	Breast vHMEC	CD14 Primary Cells	CD15 Primary Cells	CD19 Primary Cells	CD20 Primary Cells	Assay
Brain Substantia Nigra										
Breast Luminal Epithelial Cells		4	5						2	mDNA-Seq
Breast Myoepithelial Cells		3	3						2	smRNA-Seq
Breast Stem Cells		4	4						1	Expression Ar
Breast vHMEC		1	1		2				1	1
CD14 Primary Cells					2					
CD15 Primary Cells					1					
CD19 Primary Cells					1	3				
CD20 Primary Cells					1					

Human Epigenome Atlas Release 4 (hg19)

Gene: SCGB2A2

Gene: BRBB2

Assay: mDNA-Seq

Assay: MeDIP-Seq

Assay: DNase Hypersensitivity

Assay: Digital Genomic Footprinting

Assay: mDNA-Seq

Assay: smRNA-Seq

Assay: Expression Ar

Human Epigenome Atlas Release 4 (hg19)

Gene: BRBB2

Assay: mDNA-Seq

Assay: MeDIP-Seq

Assay: DNase Hypersensitivity

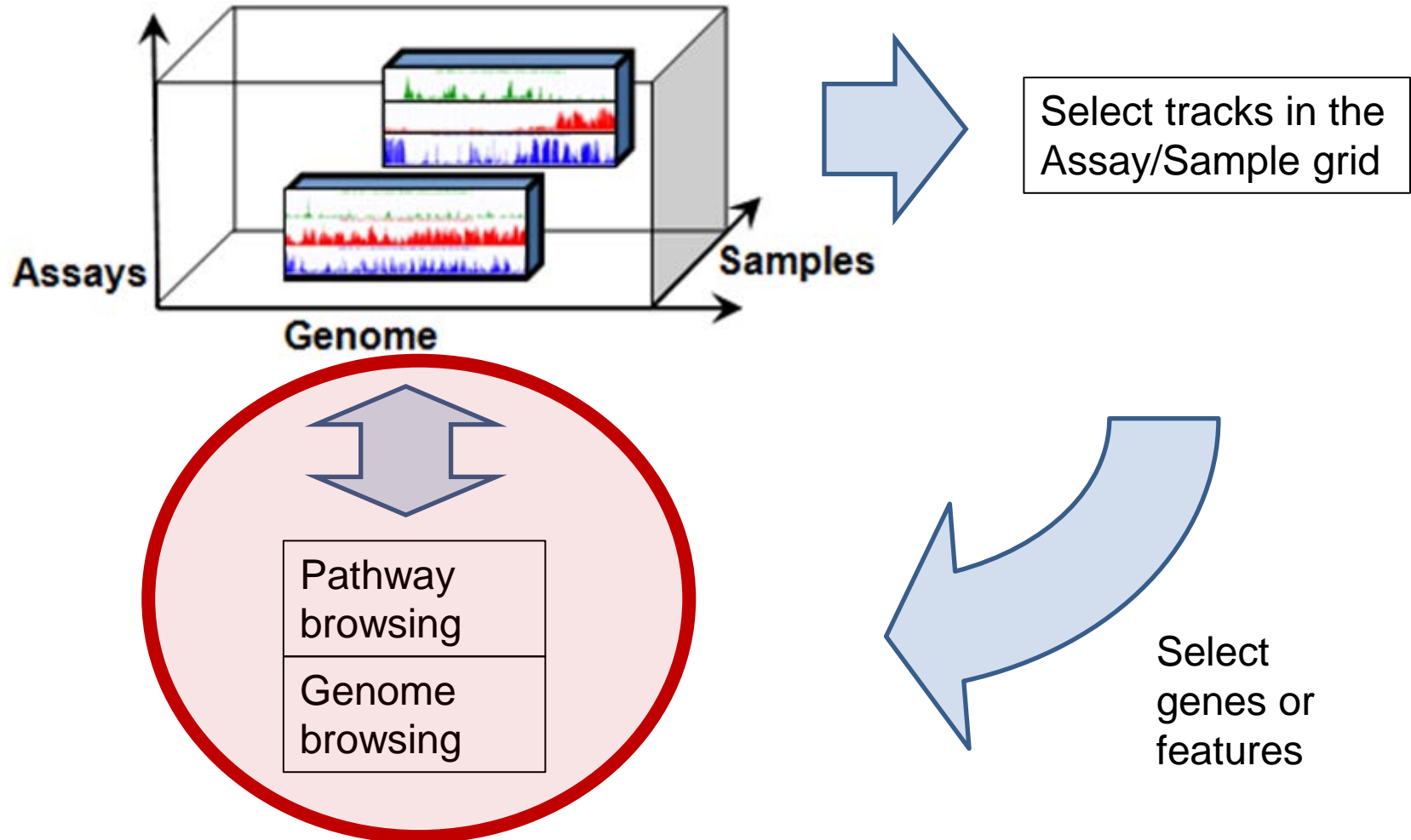
Assay: Digital Genomic Footprinting

Assay: mDNA-Seq

Assay: smRNA-Seq

Assay: Expression Ar

Epigenome Atlas Grid and the Atlas Gene Browser



Atlas Gene Browser

Download scores
into spreadsheet

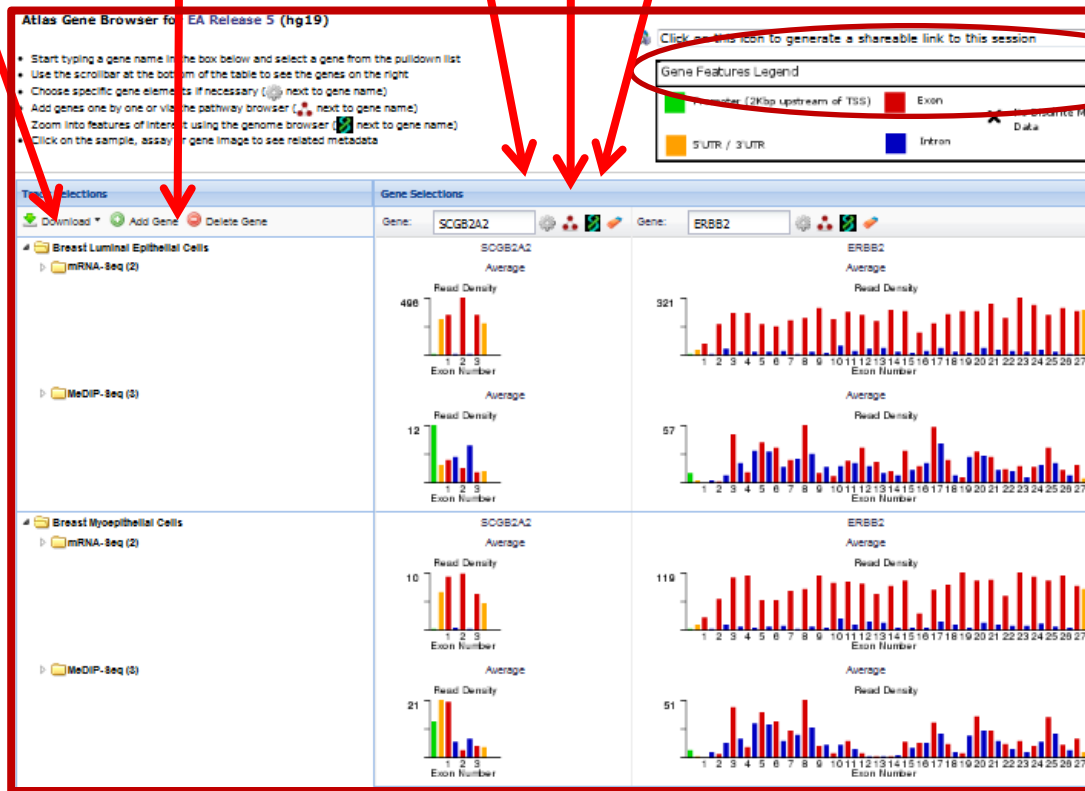
Configure gene elements

Add/delete genes

Select genes from shared
KEGG pathways

View in
UCSC genome browser

Share view (URL)



Including More Genes in the Same Pathway

Atlas Gene Browser for EA Release 5 (hg19)

- Start typing a gene name in the box below and select a gene from the pulldown list
- Use the scrollbar at the bottom of the table to see the genes on the right
- Choose specific gene elements if necessary (🔍 next to gene name)
- Add genes one by one or via the pathway browser (🔗 next to gene name)
- Zoom into features of interest using the genome browser (🔍 next to gene name)
- Click on the sample, assay or gene image to see related metadata

Click on this icon to generate a shareable link to this session

Gene Features Legend

- Promoter (2Kbp upstream of TSS)
- Exon
- 5'UTR / 3'UTR
- Intron
- No Biofite Med Data

Track Selections

Download Add Gene Delete Gene

Breast Luminal Epithelial Cells

mRNA-Seq (2)

Gene Selections

Gene: SCGB2A2

Gene: ERBB2

Pathways associated with the gene "ERBB2":

- Cellular Processes :
 - Cell Communication :
 - "Adherens junction" (Kegg Pathway Id: hsa04520)
 - "Focal adhesion" (Kegg Pathway Id: hsa04510)
 - Environmental Information Processing :
 - Signal Transduction :
 - "Calcium signaling pathway" (Kegg Pathway Id: hsa04020)
 - "ErbB signaling pathway" (Kegg Pathway Id: hsa04012)
 - Human Diseases :
 - Cancers :
 - "Bladder cancer" (Kegg Pathway Id: hsa05219)
 - "Endometrial cancer" (Kegg Pathway Id: hsa05213)
 - "Non-small cell lung cancer" (Kegg Pathway Id: hsa05223)
 - "Pancreatic cancer" (Kegg Pathway Id: hsa05212)
 - "Pathways in cancer" (Kegg Pathway Id: hsa05200)
 - "Prostate cancer" (Kegg Pathway Id: hsa05215)

Genes Selected:
of Genes: (0)

Choose from genes in a pathway

Note: This Is Wrapped Pathway Data
The pathway from Kegg has been wrapped by Genboree. The original Kegg pathway page contains useful links to their own database information about pathways, genes, etc.

Click a pathway node to select genes. Click [🔍] to save selections or click [X] to discard changes. When a gene is saved, all nodes containing that gene will be highlighted. When you are finished choosing genes click 'Done' to return to the previous screen.

ERBB SIGNALING PATHWAY

Calcium signaling pathway

ER

C²⁺

CAMK

- Cellular targets
- PKC
- Cellular targets
- Cbl
- Receptor ubiquitylation
- Degradation
- STAT3
- Src
- FAK
- Adhesion Migration
- Crk
- Abl
- BTC
- (HER2/Neu)
- ErbB-1
- ErbB-2
- EGF
- TOFe
- AR
- EGFR

Genes Selected:
of Genes: (0)

Done