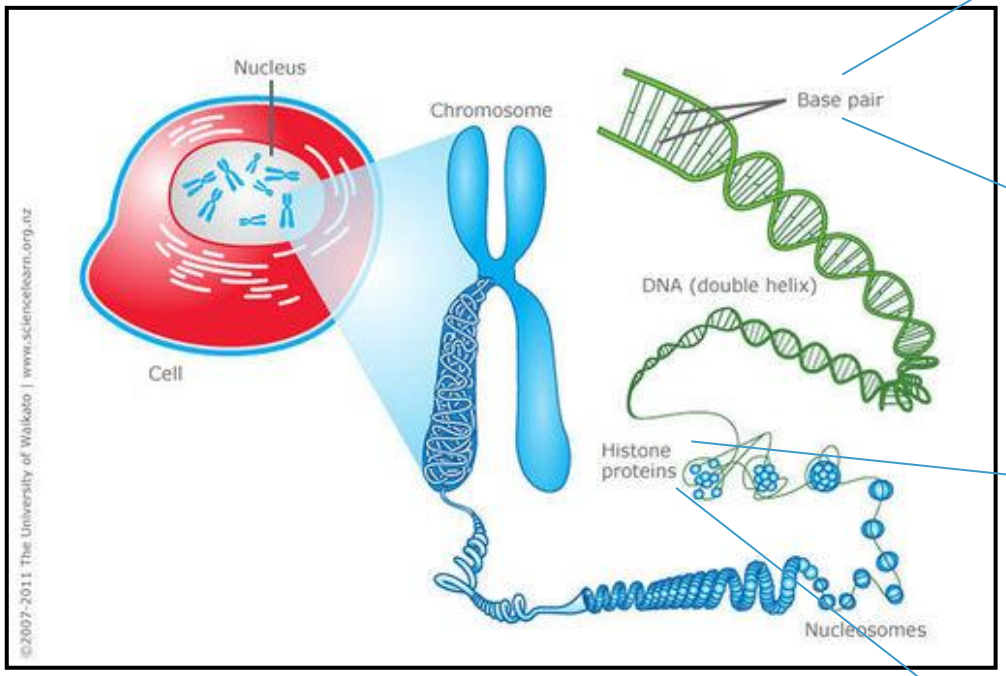
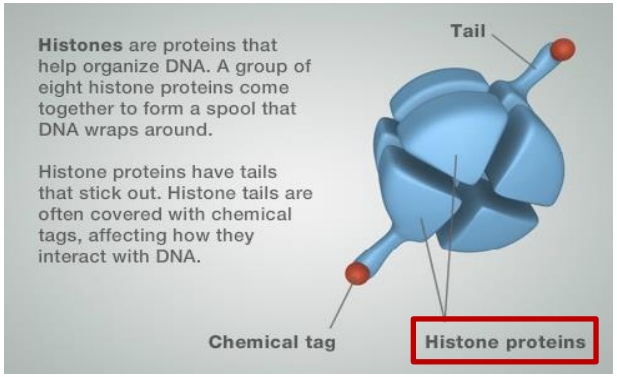
Why DNA Isn't Your Destiny

Courtesy of: Matthew Forsythe

GENOME: *a genome is an organism's complete set of DNA, including all of its genes. Each genome contains all of the information needed to build and maintain that organism*

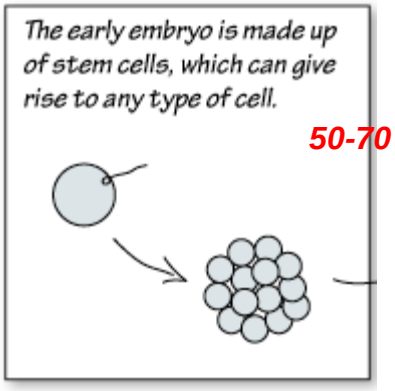


modified by mashable.com



modified by http://learn.genetics.utah.edu/content/epigenetics/epi_learns/

From GENOME to EPIGENOME to CELLS to ORGANISMS

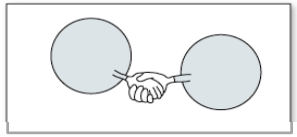


50-70 trillion cells!!!!

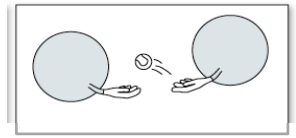
Cells listen for signals

Types of signals

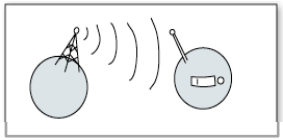
1. Direct Contact



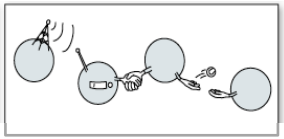
2. Transmission (factor release)



3. Hormones



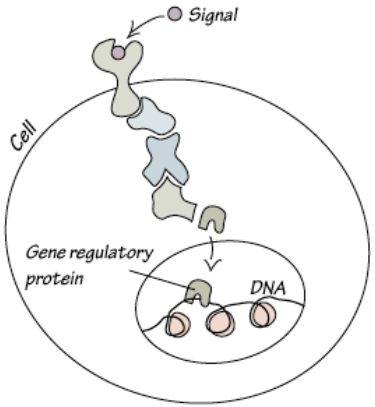
4. Combo



modified by http://learn.genetics.utah.edu/content/epigenetics/epi_learns/

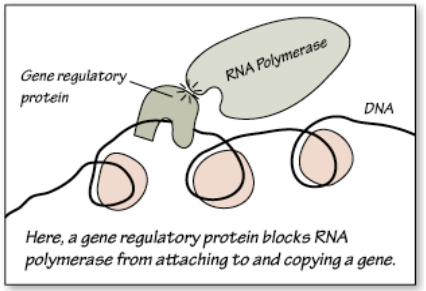
From GENOME to EPIGENOME to CELLS to ORGANISMS

Proteins Carry Signals to the DNA

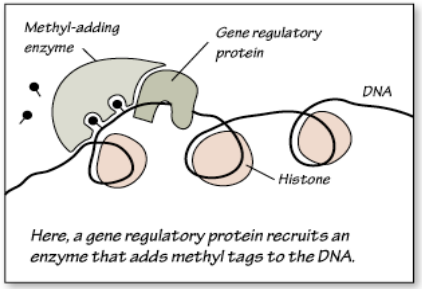


Gene Regulatory Proteins Have Two Functions

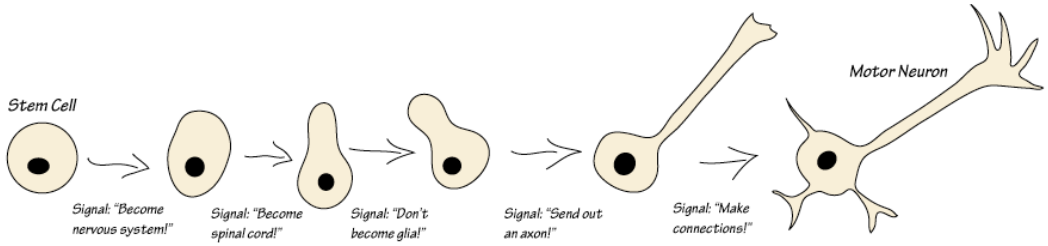
Switch genes ON/OFF



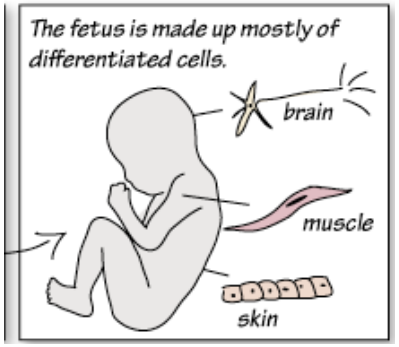
Recruit epi tag-ENZYMES



Paradigm



ORGANISM

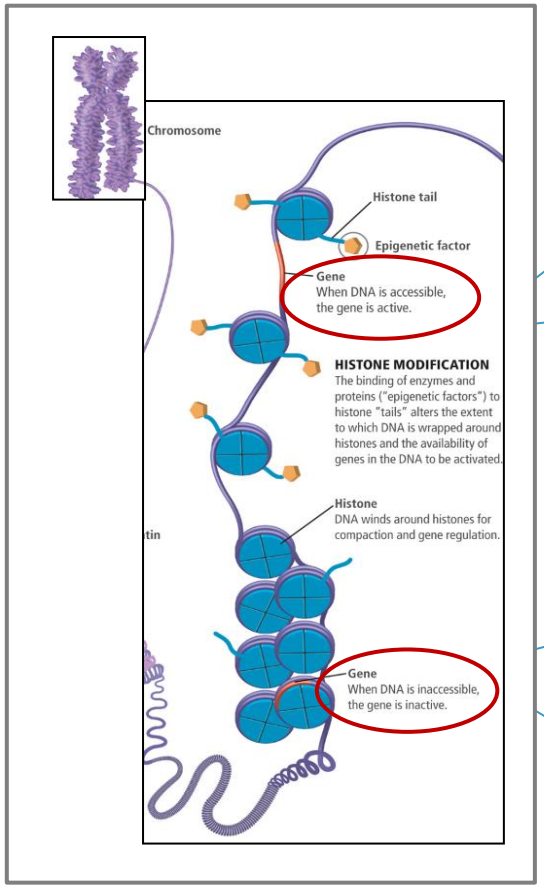


modified by http://learn.genetics.utah.edu/content/epigenetics/epi_learns/

“Force a rethink of the definition of a gene and of the minimum unit of heredity.”

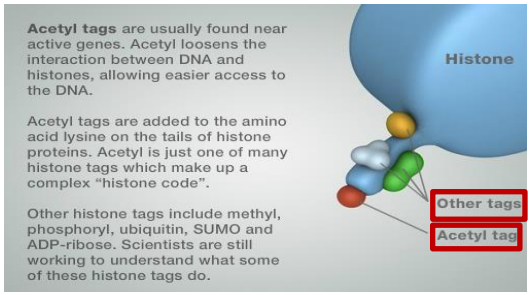
EPIGENOME: *the epigenome comprises all of the chemical modifications (tags) added to the genome as a way to regulate the activity (expression) of all the genes.*

MODIFICATIONS

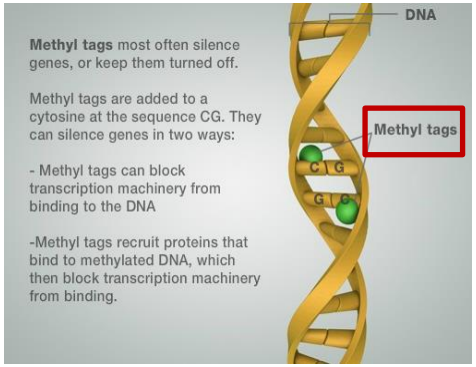


modified by http://ehp.niehs.nih.gov/120-a396/

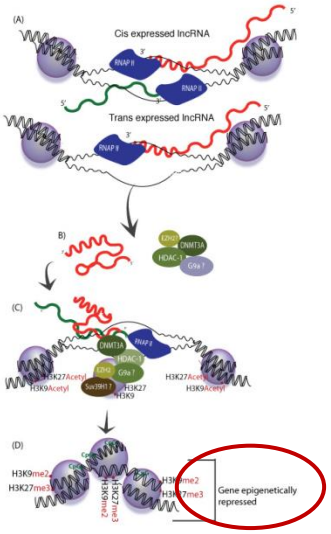
II. Histone Modifications



I. DNA methylation



III. ncRNA associated gene silencing



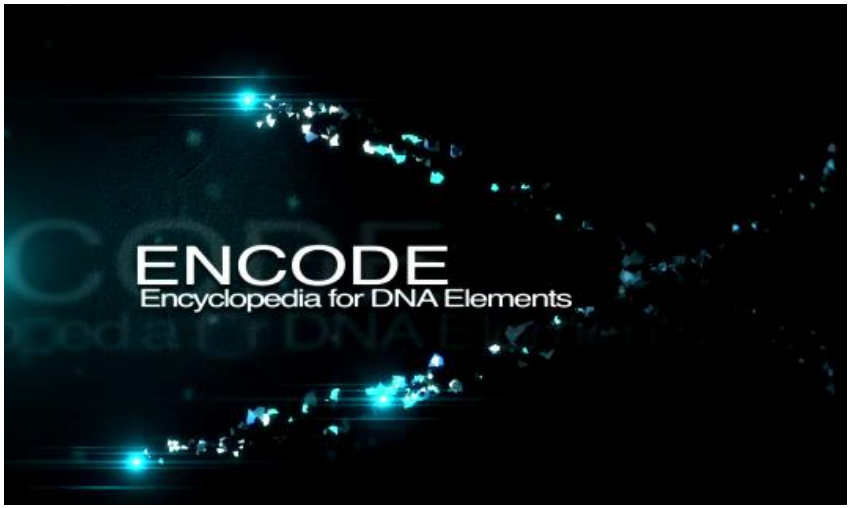
www.morrislab.unsw.edu.au

modified by http://learn.genetics.utah.edu/content/epigenetics/epi_learns/

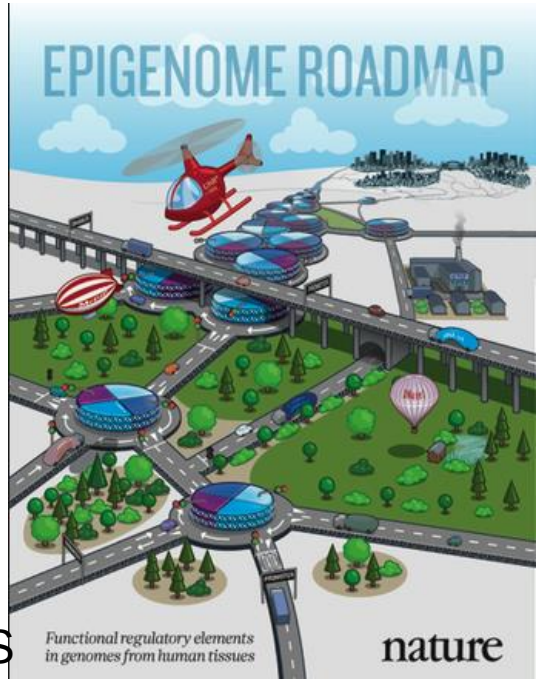
RESEARCH CONSORTIA

BLUEPRINT epigenome

BLOOD CELLS



CELL LINES



111 CELLS & TISSUES

AIMS

1. How is genetic information interpreted by single cells?

2. Annotate cis-regulatory elements.

3. Create *maps* of epigenomic modifications.

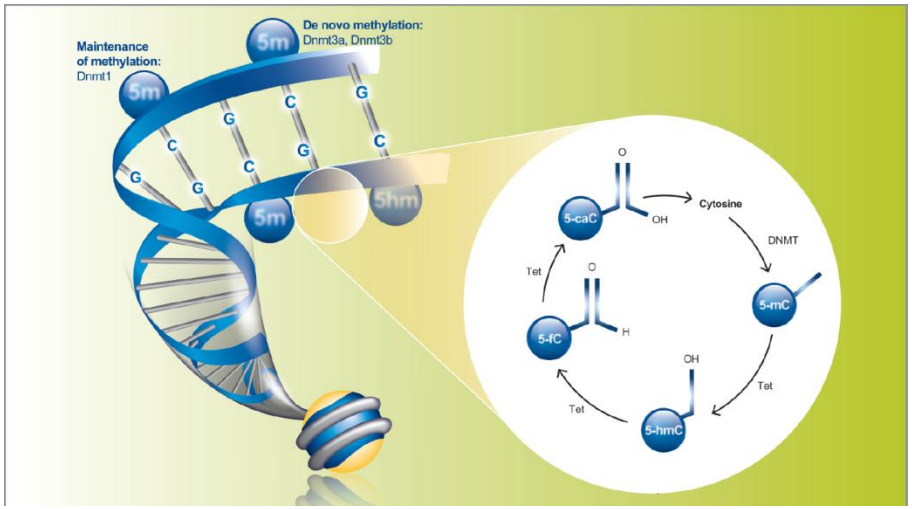
4. Produce clinically usefull epigenetic information.

5. Dissect gene regulatory programs in development and disease.

METHODS

METHODS

I. DNA methylation



a) Sodium bisulfite modification

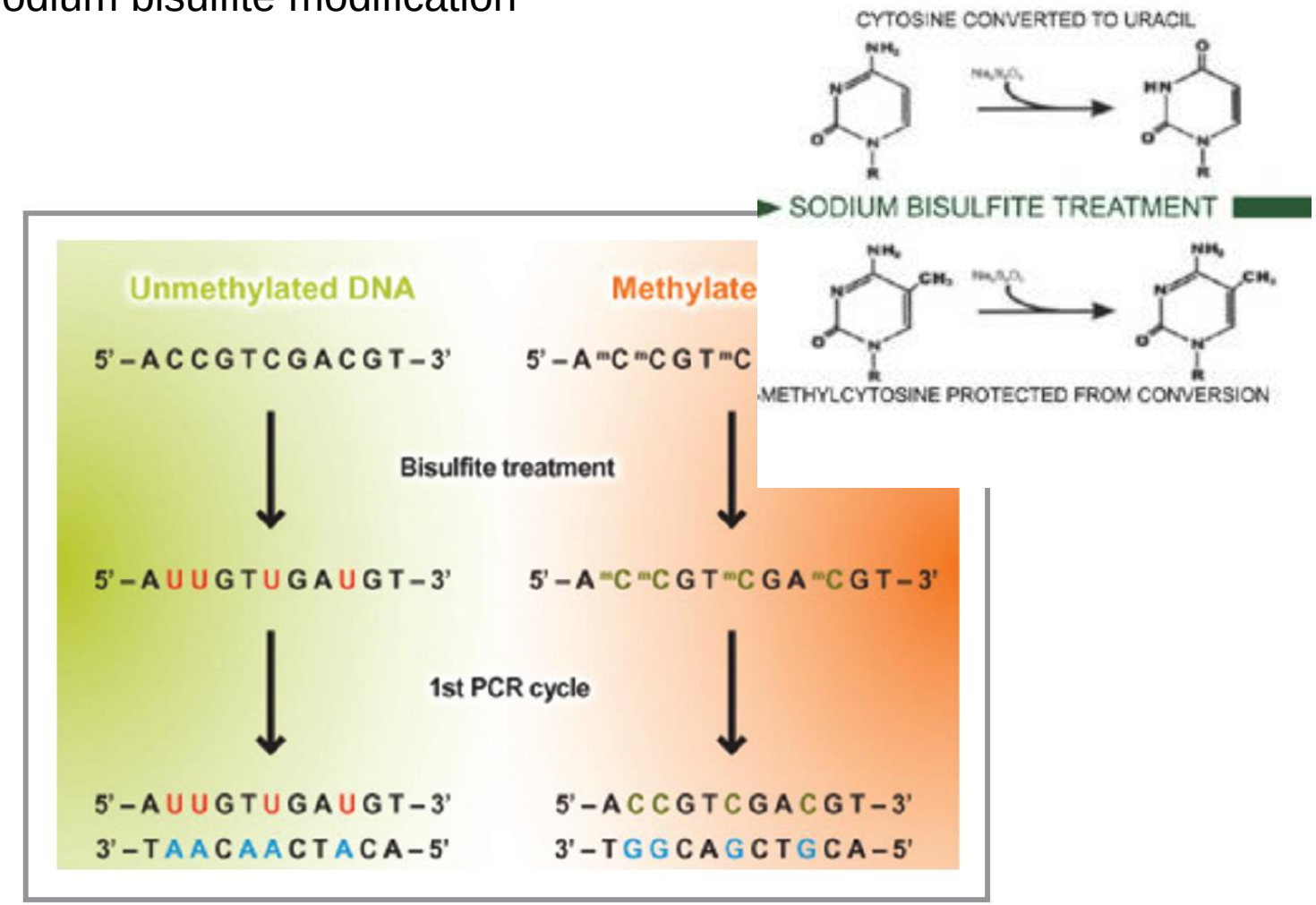
b) Sequence-specific enzyme digestion

c) Capture/quantification of methylated DNA

http://docs.abcam.com/pdf/chromatin/a-guide-to-epigenetics

METHODS

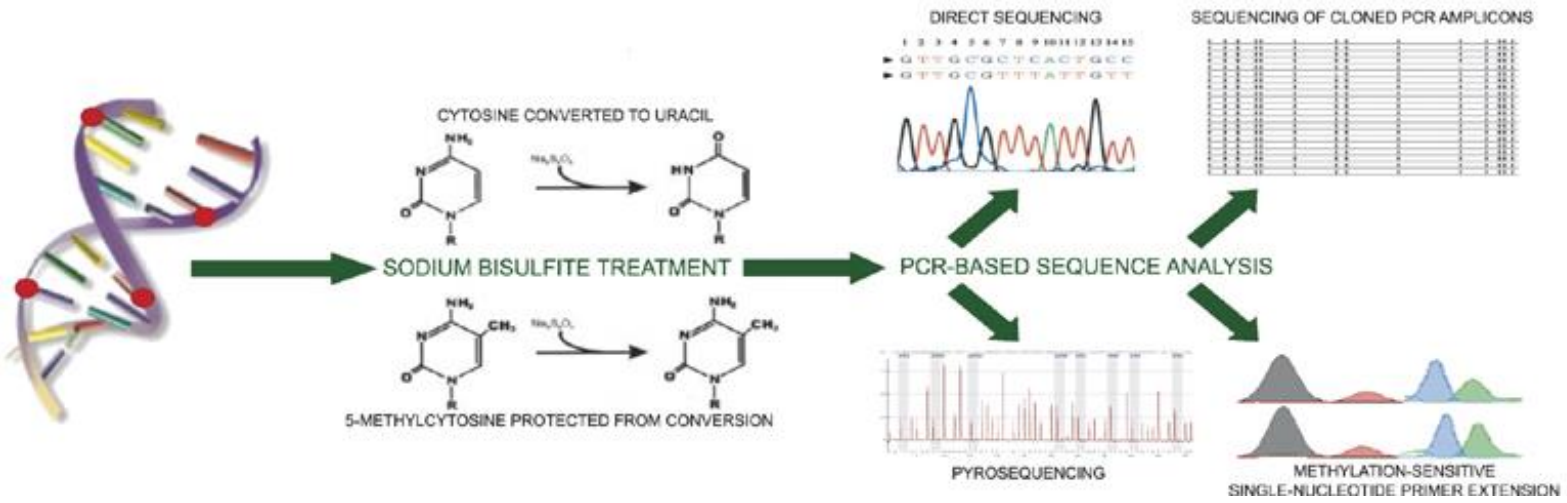
a) Sodium bisulfite modification



http://docs.abcam.com/pdf/chromatin/a-guide-to-epigenetics

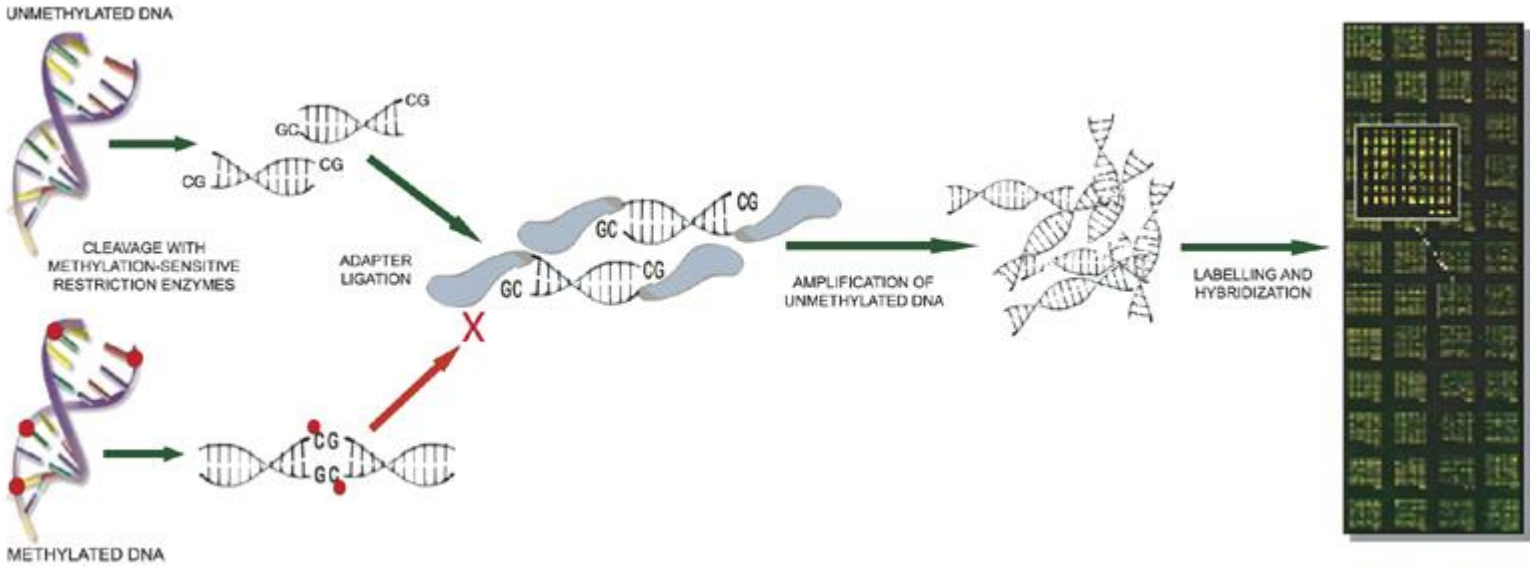
METHODS

a) Sodium bisulfite modification



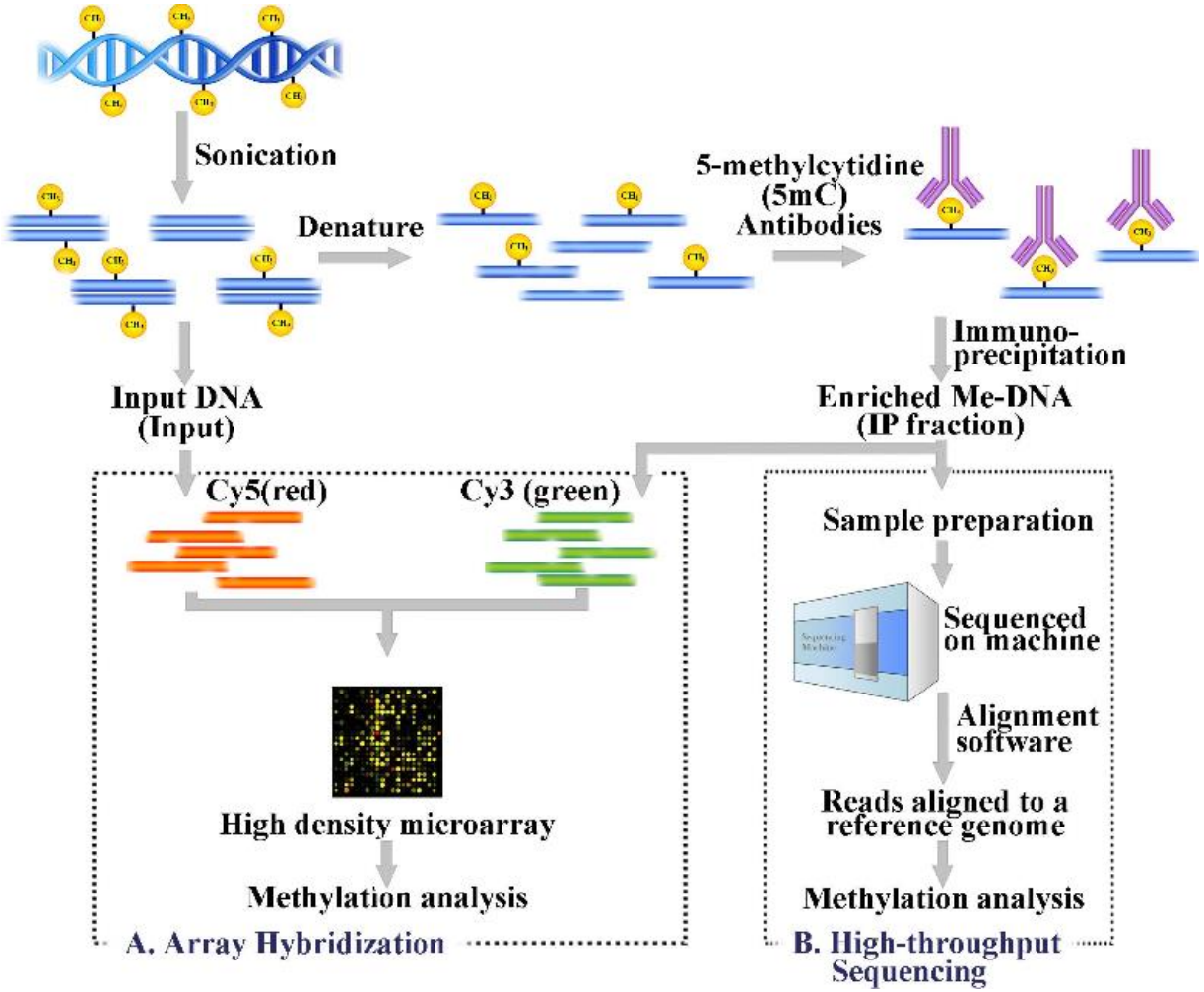
Molecular Psychiatry (2007) **12,** 799–814; doi:10.1038/sj.mp.4001992

b) Sequence-specific enzyme digestion



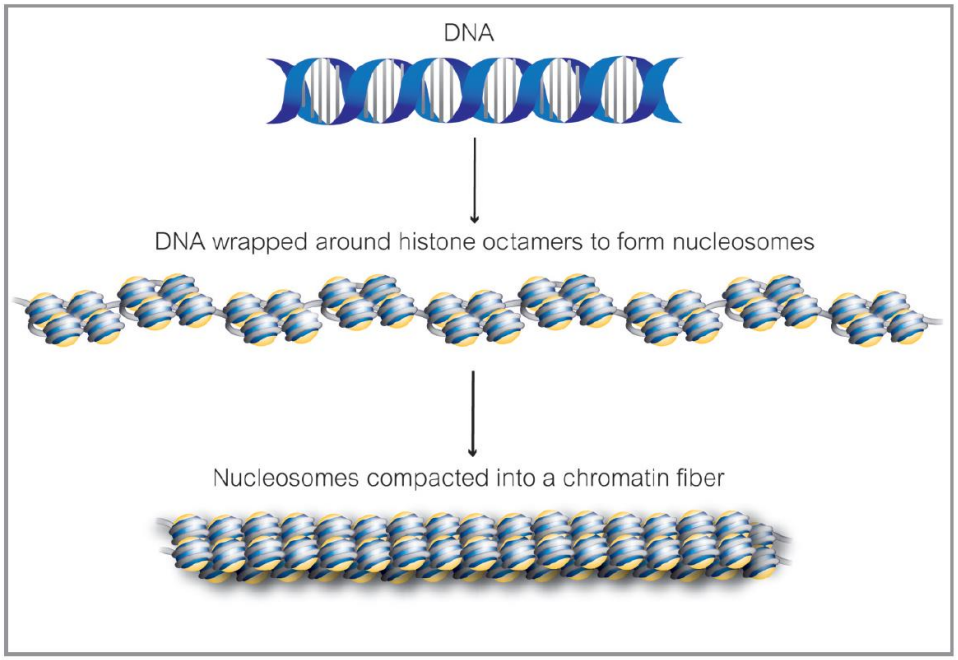
Molecular Psychiatry (2007) **12,** 799–814; doi:10.1038/sj.mp.4001992

c) Capture/quantification of methylated DNA (MeDIP)



http://docs.abcam.com/pdf/chromatin/a-guide-to-epigenetics

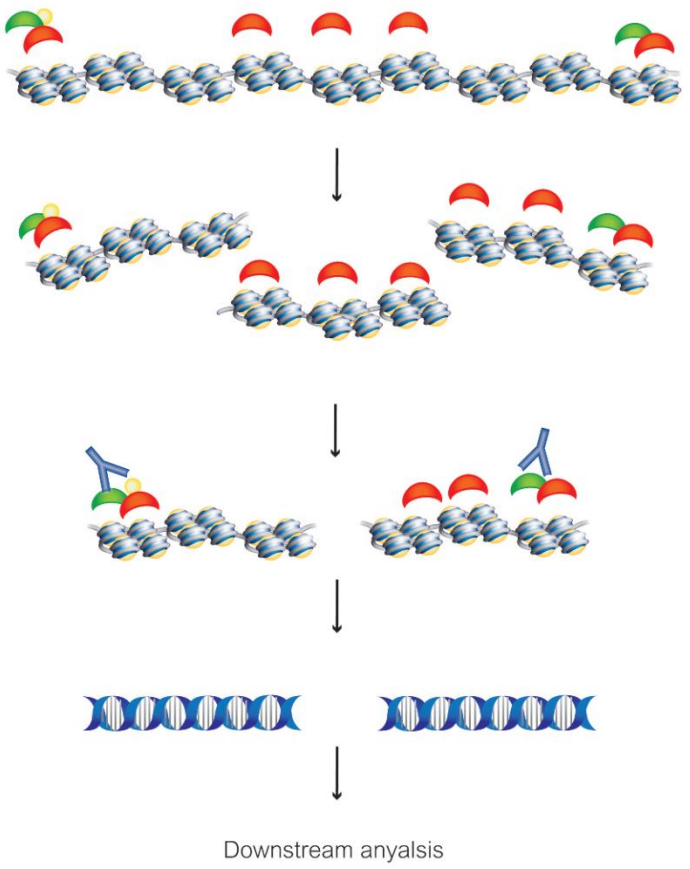
II. Chromatin Modifications



a) Chromatin Immunoprecipitation

http://docs.abcam.com/pdf/chromatin/a-guide-to-epigenetics

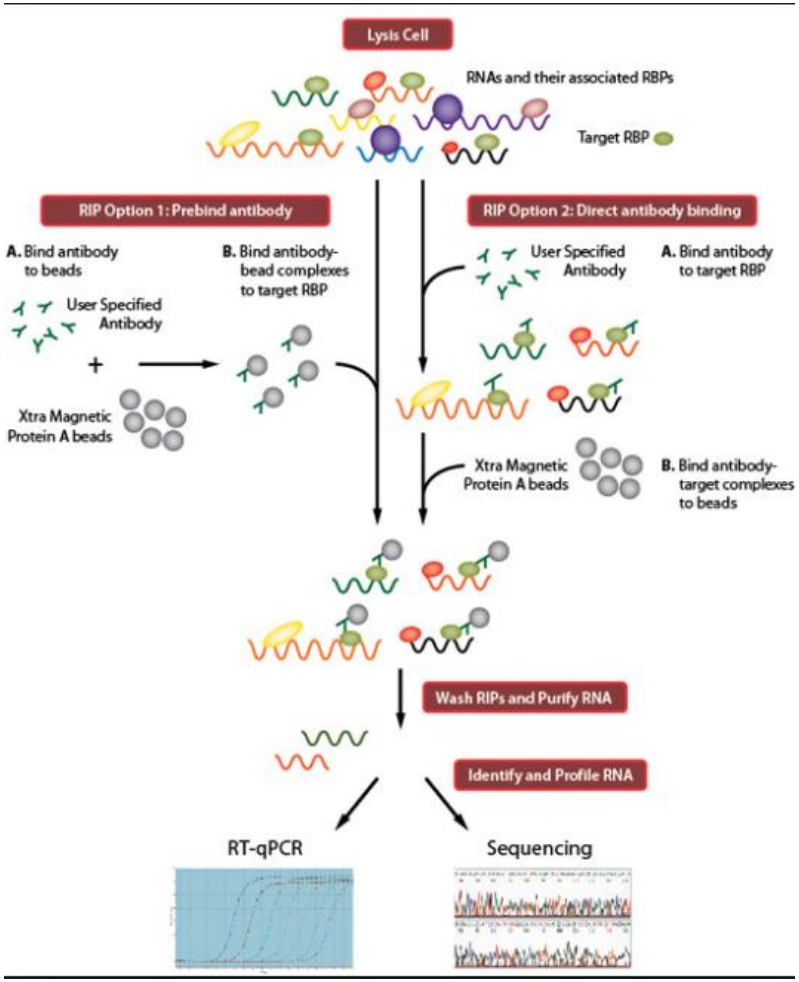
a) Chromatin Immunoprecipitation (ChIP)



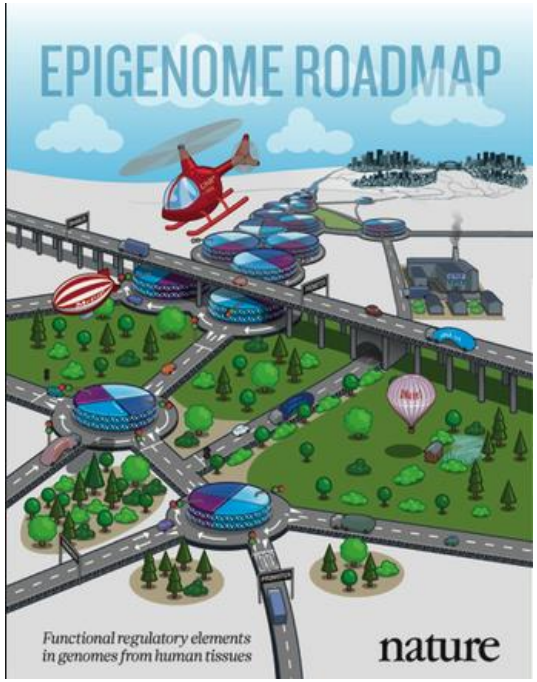
1. Cross-link DNA and proteins (optional)
2. Chromatin fragmentation by sonication or by enzymatic methods (e.g. Micrococcal nuclease)
3. Immunoprecipitation of the chromatin fragments interacting with the target protein/modification.
4. Reverse cross-link (when necessary) and DNA purification
5. Analysis of the immunoprecipitated fraction to determine abudance of target sequence(s) relative to input. Common methods include qPCR, ChIPseq and ChIP-Chip.

http://docs.abcam.com/pdf/chromatin/a-guide-to-epigenetics

a) RNA-protein interactions (RIP)



http://www.sigmaaldrich.com/life-science/epigenetics/imprint-rna.html



ARTICLE

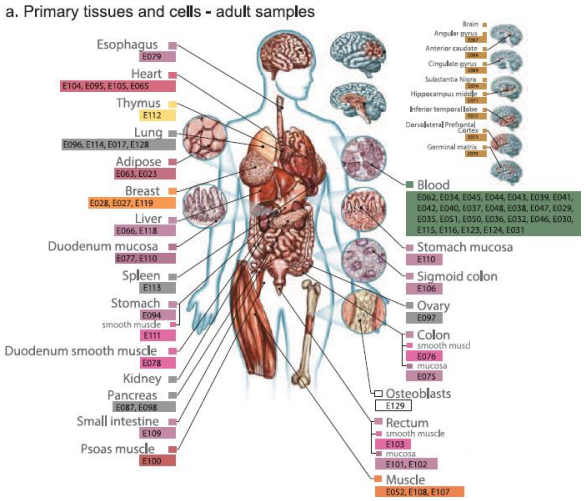
OPEN

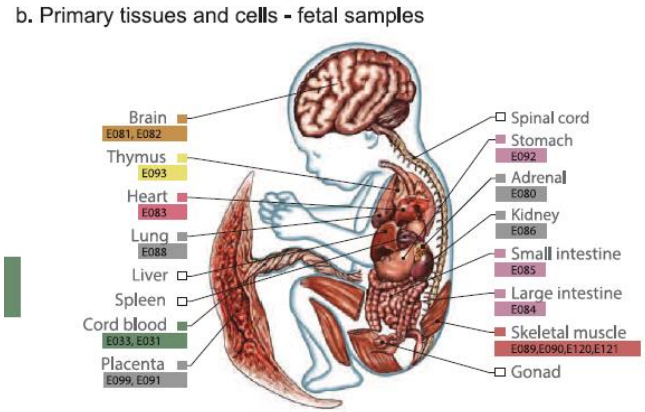
doi:10.1038/nature14248

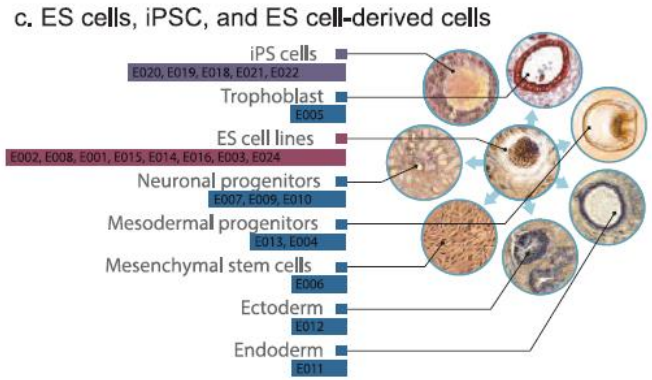
Integrative analysis of 111 reference human epigenomes

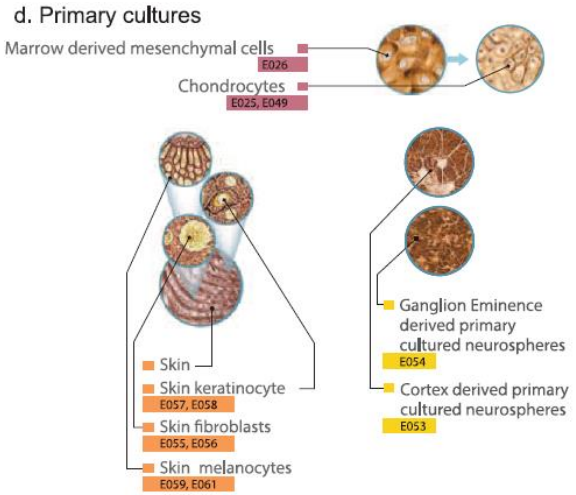
Roadmap Epigenomics Consortium†, Anshul Kundaje[1,2,3]*, Wouter Meuleman[1,2]*, Jason Ernst[1,2,4]*, Misha Bilenky[5]*, Angela Yen[1,2], Alireza Heravi-Moussavi[5], Pouya Kheradpour[1,2], Zhizhuo Zhang[1,2], Jianrong Wang[1,2], Michael J. Ziller[2,6], Viren Amin[7], John W. Whitaker[8], Matthew D. Schultz[9], Lucas D. Ward[1,2], Abhishek Sarkar[1,2], Gerald Quon[1,2], Richard S. Sandstrom[10], Matthew L. Eaton[1,2], Yi-Chieh Wu[1,2], Andreas R. Pfenning[1,2], Xinchen Wang[1,2,11], Melina Claussnitzer[1,2], Yaping Liu[1,2], Cristian Coarfa[7], R. Alan Harris[7], Noam Shoresh[2], Charles B. Epstein[2], Elizabeta Gjoneska[2,12], Danny Leung[8,13], Wei Xie[8,13], R. David Hawkins[8,13], Ryan Lister[9], Chibo Hong[14], Philippe Gascard[15], Andrew J. Mungall[5], Richard Moore[5], Eric Chuah[5], Angela Tam[5], Theresa K. Canfield[10], R. Scott Hansen[16], Rajinder Kaul[16], Peter J. Sabo[10], Mukul S. Bansal[1,2,17], Annaick Carles[18], Jesse R. Dixon[8,13], Kai-How Farh[2], Soheil Feizi[1,2], Rosa Karlic[19], Ah-Ram Kim[1,2], Ashwinikumar Kulkarni[20], Daofeng Li[21], Rebecca Lowdon[21], GiNell Elliott[21], Tim R. Mercer[22], Shane J. Neph[10], Vitor Onuchic[7], Paz Polak[2,23], Nisha Rajagopal[8,13], Pradipta Ray[20], Richard C. Sallari[1,2], Kyle T. Siebenthall[10], Nicholas A. Sinnott-Armstrong[1,2], Michael Stevens[21,42], Robert E. Thurman[10], Jie Wu[24,25], Bo Zhang[21], Xin Zhou[21], Arthur E. Beaudet[26], Laurie A. Boyer[11], Philip L. De Jager[2,23,27], Peggy J. Farnham[28], Susan J. Fisher[29], David Haussler[30], Steven J. M. Jones[5,31,32], Wei Li[33], Marco A. Marra[5,32], Michael T. McManus[34], Shamil Sunyaev[2,23,27], James A. Thomson[35,41], Thea D. Tlsty[15], Li-Huei Tsai[2,12], Wei Wang[8], Robert A. Waterland[36], Michael Q. Zhang[20,37], Lisa H. Chadwick[38], Bradley E. Bernstein[2,39,40]§, Joseph F. Costello[14]§, Joseph R. Ecker[9]§, Martin Hirst[5,18]§, Alexander Meissner[2,6]§, Aleksandar Milosavljevic[7]§, Bing Ren[8,13]§, John A. Stamatoyannopoulos[10]§, Ting Wang[21]§ & Manolis Kellis[1,2]§

MATERIAL

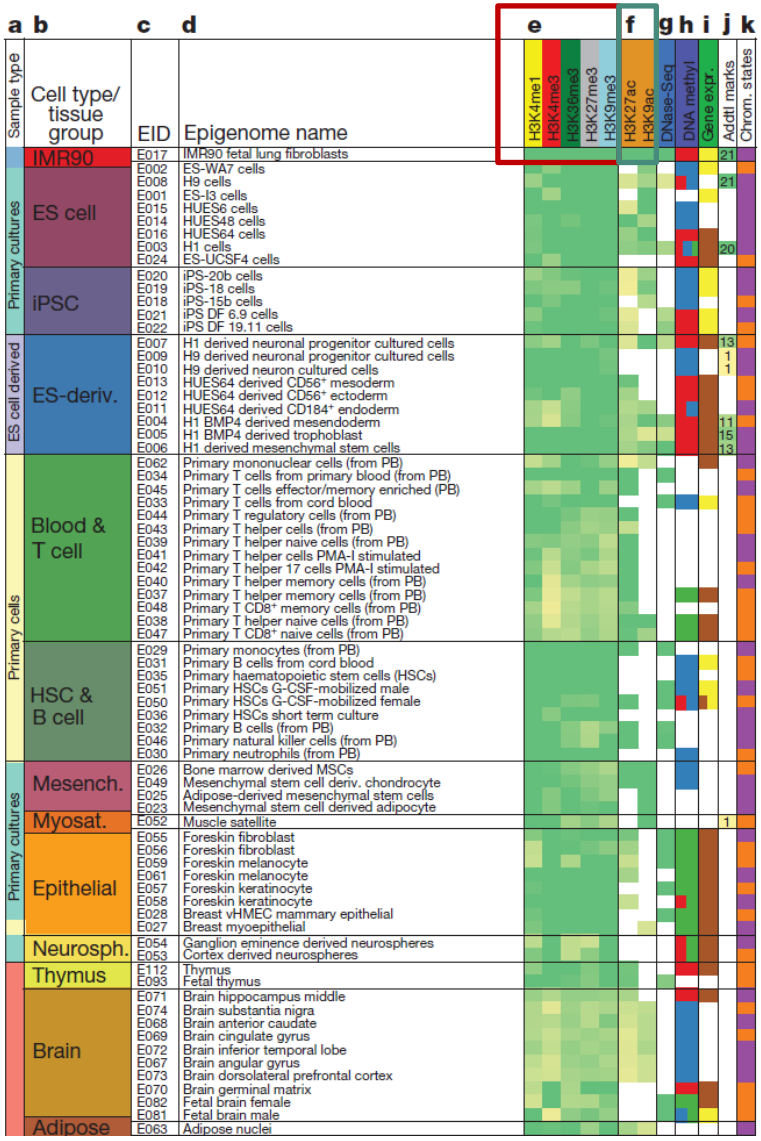

a. Primary tissues and cells - adult samples
Esophagus
E079
Heart
E104, E095, E105, E065
Thymus
E112
Lung
E096, E114, E017, E128
Adipose
E063, E023
Breast
E028, E027, E119
Liver
E066, E118
Duodenum mucosa
E077, E110
Spleen
E113
Stomach
E094
smooth muscle
E111
Duodenum smooth muscle
E078
Kidney
Pancreas
E087, E098
Small intestine
E109
Psoas muscle
E100
Brain
Angular gyrus
Anterior caudate
Cingulate gyrus
Substantia Nigra
Hippocampus middle
Inferior temporal lobe
Dorsolateral Prefrontal Cortex
Germinal matrix
Blood
E062, E034, E045, E044, E043, E039, E041, E042, E040, E037, E048, E038, E047, E029, E035, E051, E050, E036, E032, E046, E030, E115, E116, E123, E124, E031
Stomach mucosa
E110
Sigmoid colon
E106
Ovary
E097
Colon
smooth muscl
E076
mucosa
E075
Osteoblasts
E129
Rectum
smooth muscle
E103
mucosa
E101, E102
Muscle
E052, E108, E107

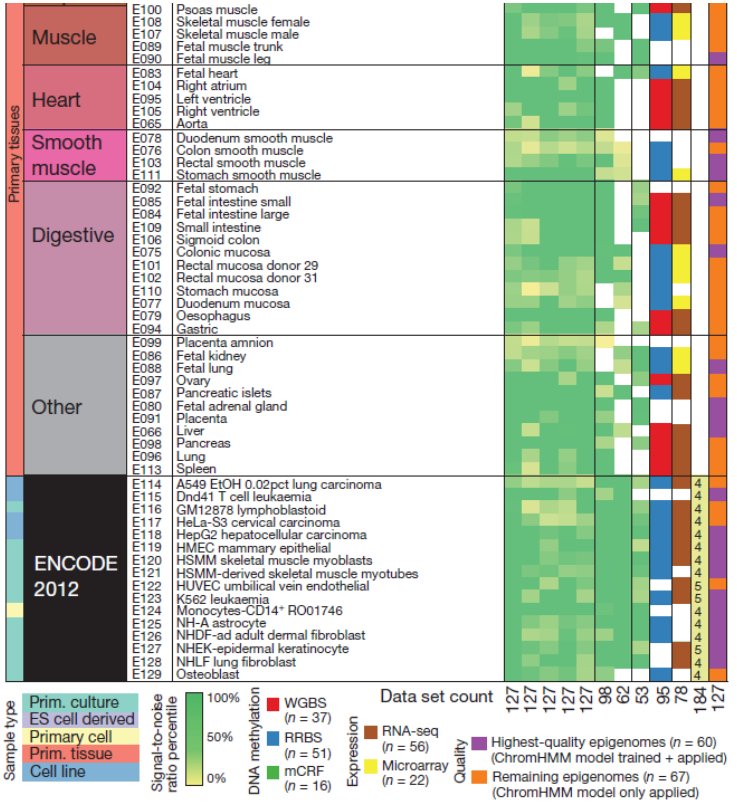

b. Primary tissues and cells - fetal samples
Brain
E081, E082
Thymus
E093
Heart
E083
Lung
E088
Liver
Spleen
Cord blood
E033, E031
Placenta
E099, E091
Spinal cord
Stomach
E092
Adrenal
E080
Kidney
E086
Small intestine
E085
Large intestine
E084
Skeletal muscle
E089,E090,E120,E121
Gonad


c. ES cells, iPSC, and ES cell-derived cells
iPS cells
E020, E019, E018, E021, E022
Trophoblast
E005
ES cell lines
E002, E008, E001, E015, E014, E016, E003, E024
Neuronal progenitors
E007, E009, E010
Mesodermal progenitors
E013, E004
Mesenchymal stem cells
E006
Ectoderm
E012
Endoderm
E011

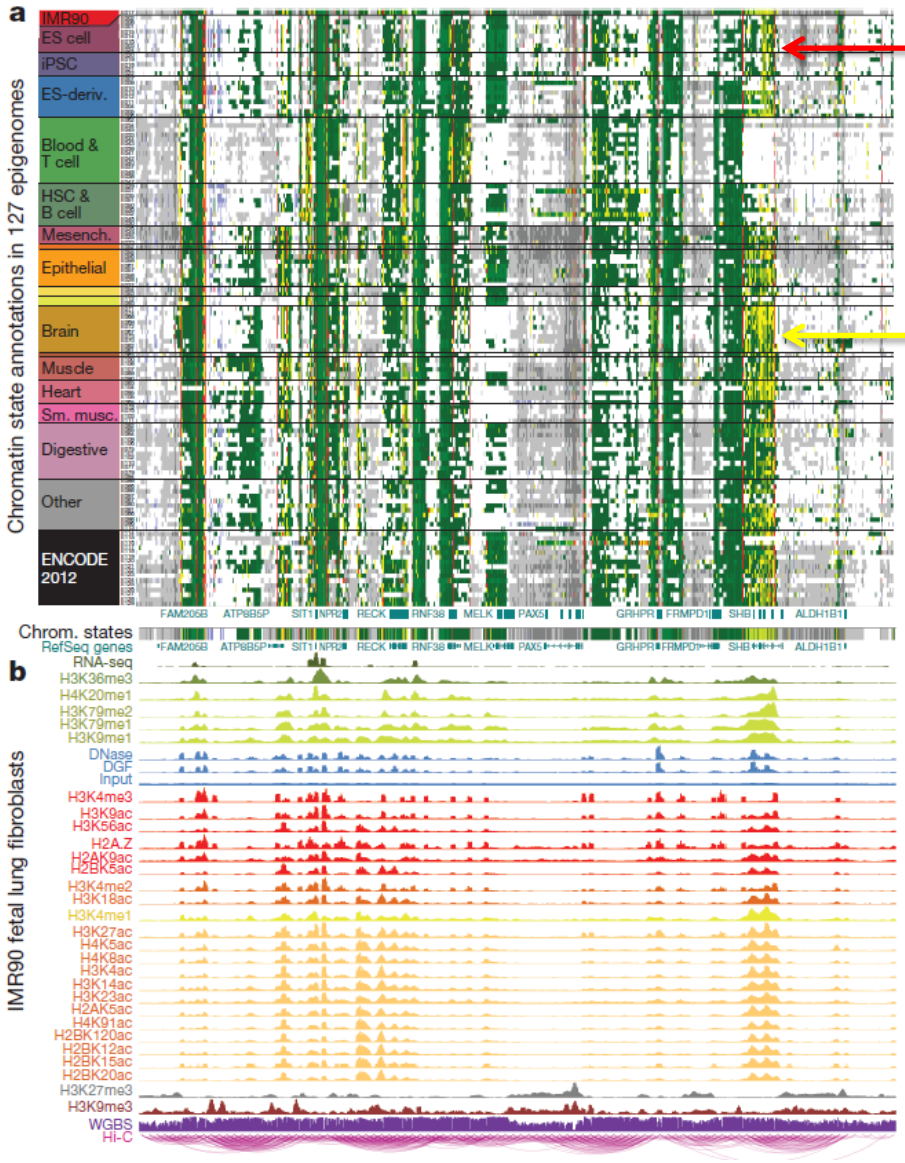

d. Primary cultures
Marrow derived mesenchymal cells
E026
Chondrocytes
E025, E049
Skin
Skin keratinocyte
E057, E058
Skin fibroblasts
E055, E056
Skin melanocytes
E059, E061
Ganglion Eminence derived primary cultured neurospheres
E054
Cortex derived primary cultured neurospheres
E053

DATA SETS for each epigenome





Chromatin State Annotation across tissues



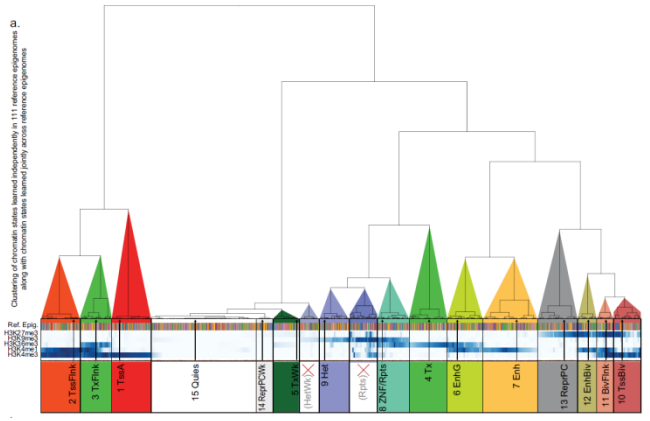
Vertical Red Lines = PROMOTERS

Dispersed Yellow Regions = ENHANCERS

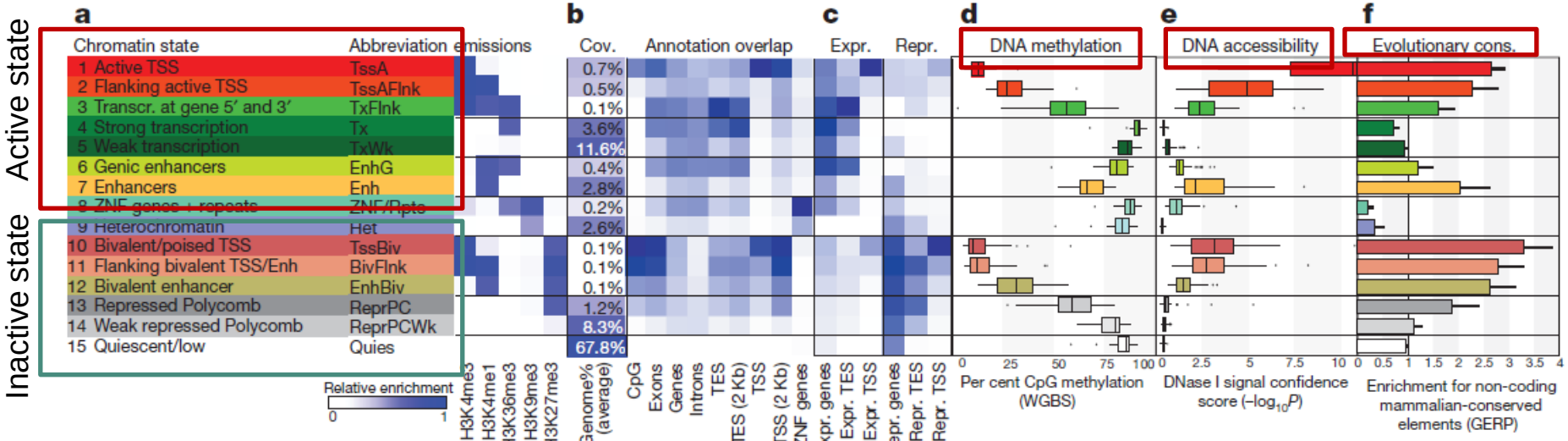
Promoters are primarily constitutive, while Enhancers are highly dynamic

Chromatin State and DNA methylation dynamics

Chromatine States

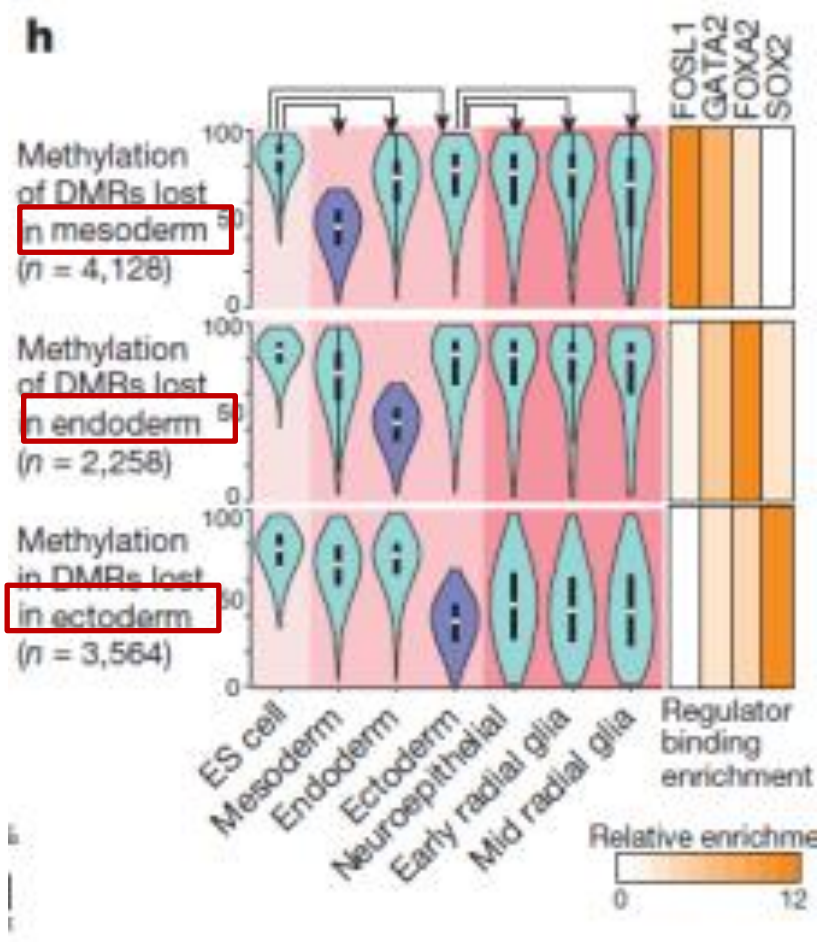


Human ES cells



Chromatin State and DNA methylation dynamics During Lineage Specification

ES cell differentiation



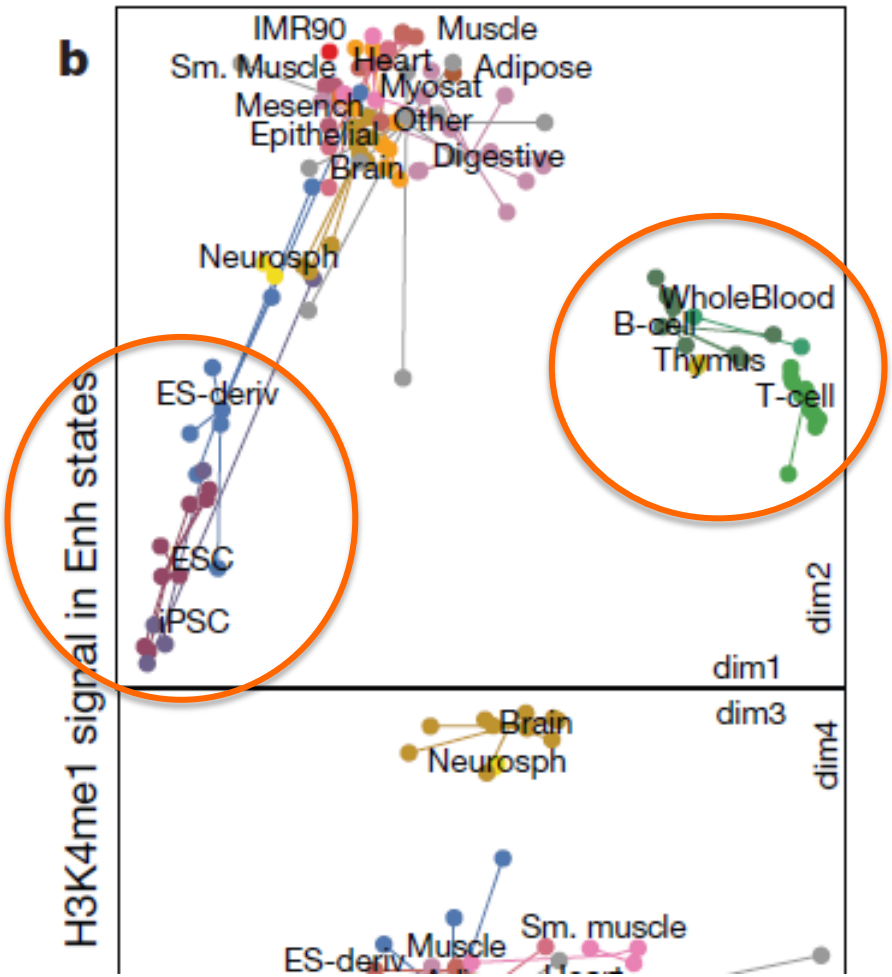
Chromatin State and DNA methylation dynamics During Lineage Specification

Skin Cells

Keratinocytes	Surface Ectoderm
Melanocytes	Neural Crest
Fibroblasts	Mesoderm

“Low overlap in DNA methylation & histone modification signatures.”

Epigenome Relationships

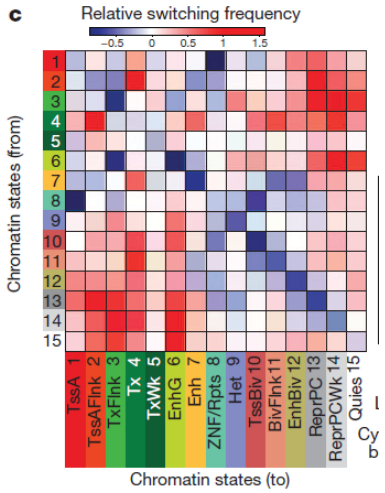


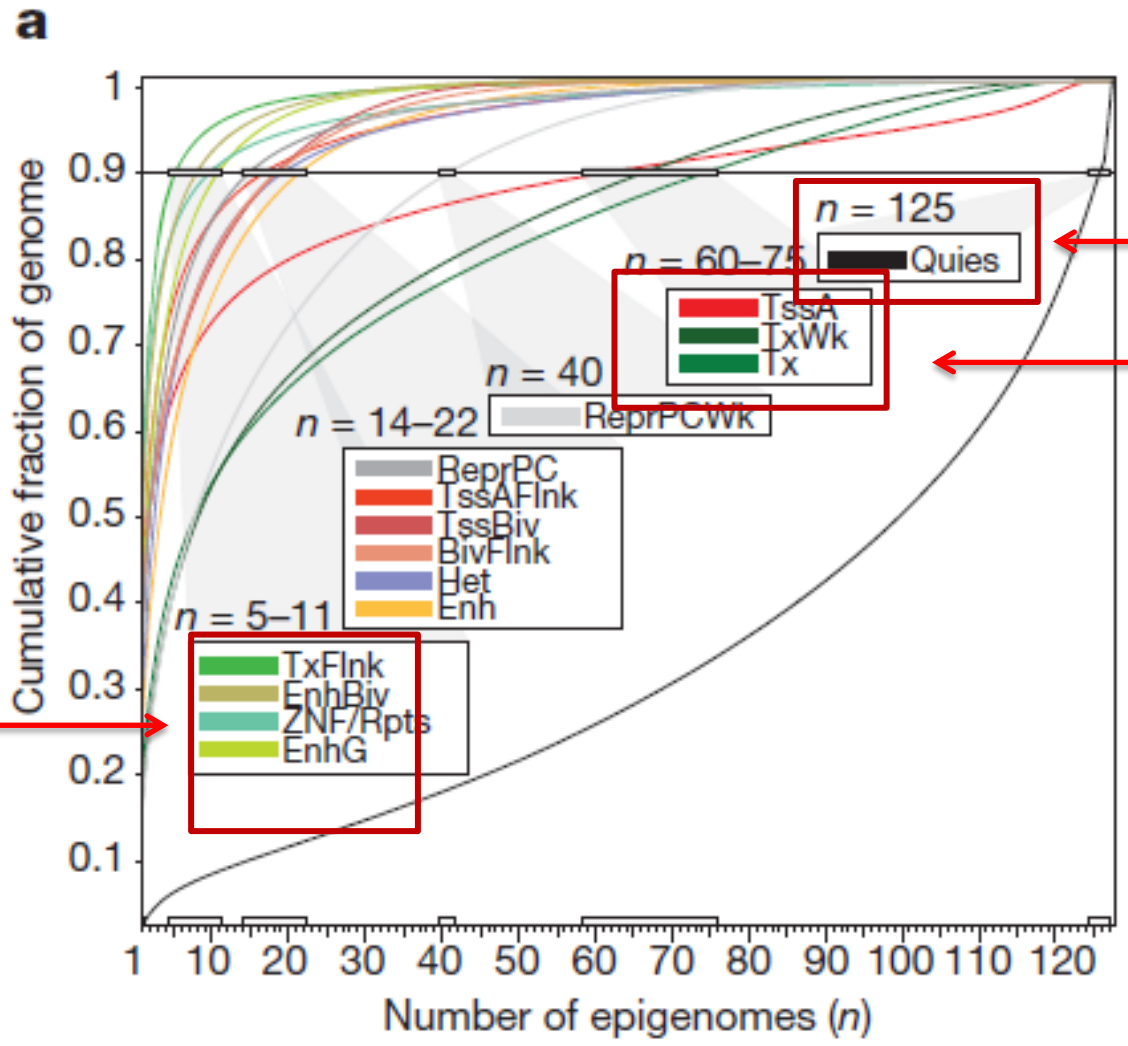
"Lines with common developmental origins show similar epigenetic modification patterns."

CONCLUSION

“Common developmental origins can be a primary determinant of global DNA methylation patterns.”

Cell Type differences in Chromatin states





Most Constitutive Quiescent Regions

Constitutive – Active states

H3K4me1 Tissue Specific – Inactive states

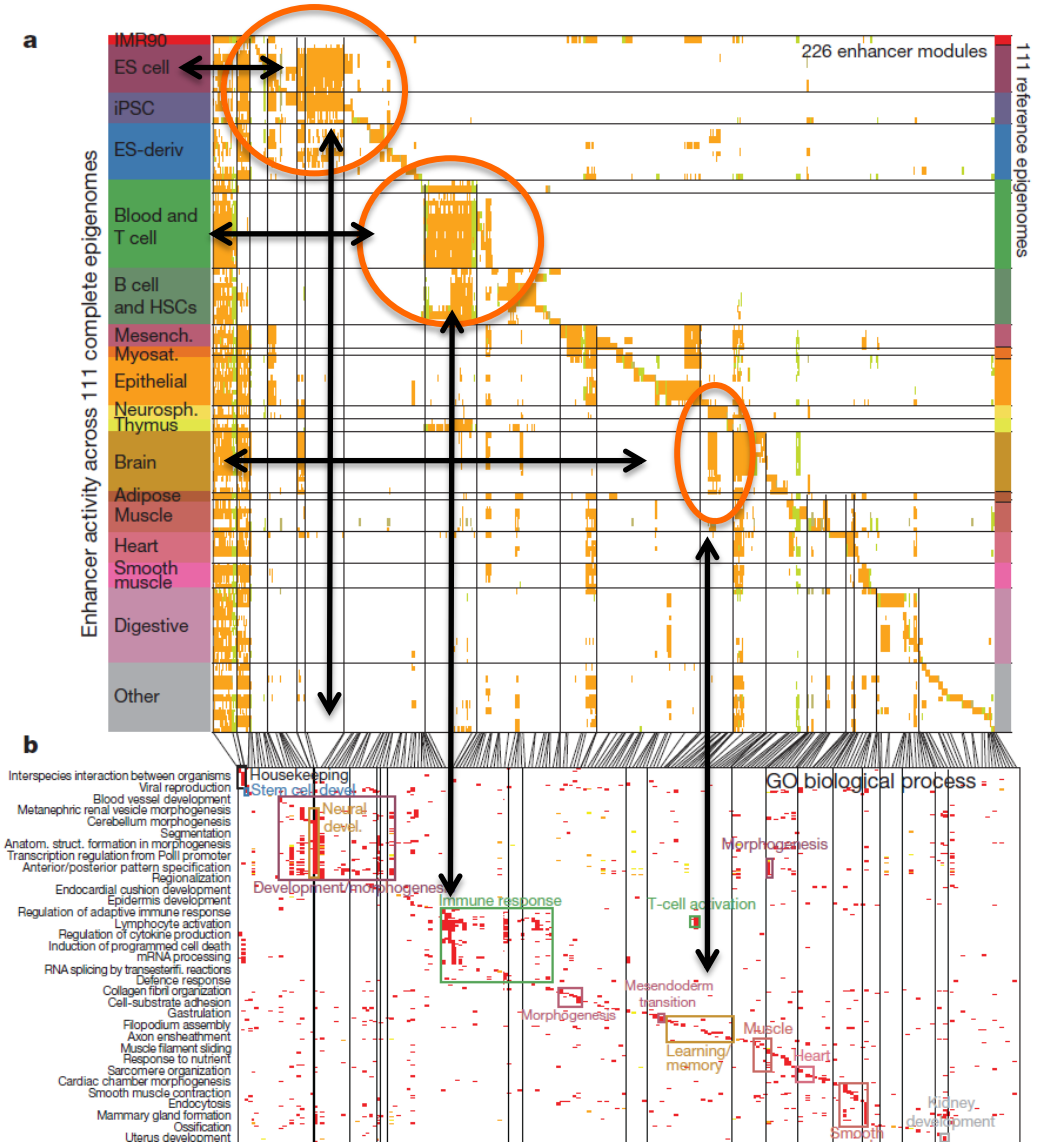
CELL TYPE SPECIFICITY

Chromatin states

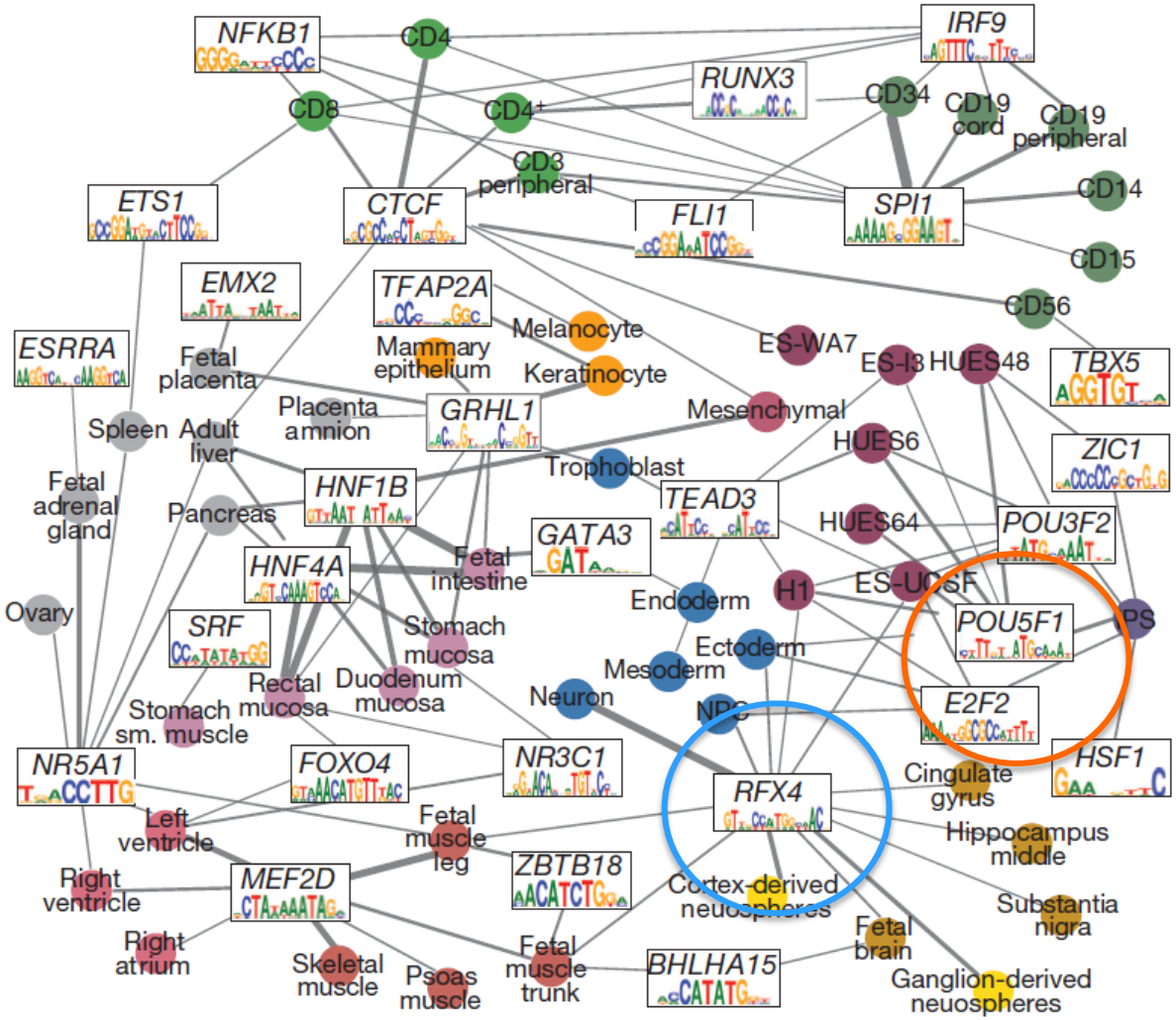
1. ***Hematopoietic stem cells and Immune cells*** *show a consistent and previously unrecognized depletion of active and bivalent promoters (TssA, TssBiv) and weakly transcribed states (TxWk).*

2. ***ES cells and iPS cells*** *show enrichment of TssBiv, consistent with previous studies. They also show a depletion of ReprPCWk, possibly due to restriction of H3K27m3-establishing Polycomb proteins to promoter regions.*

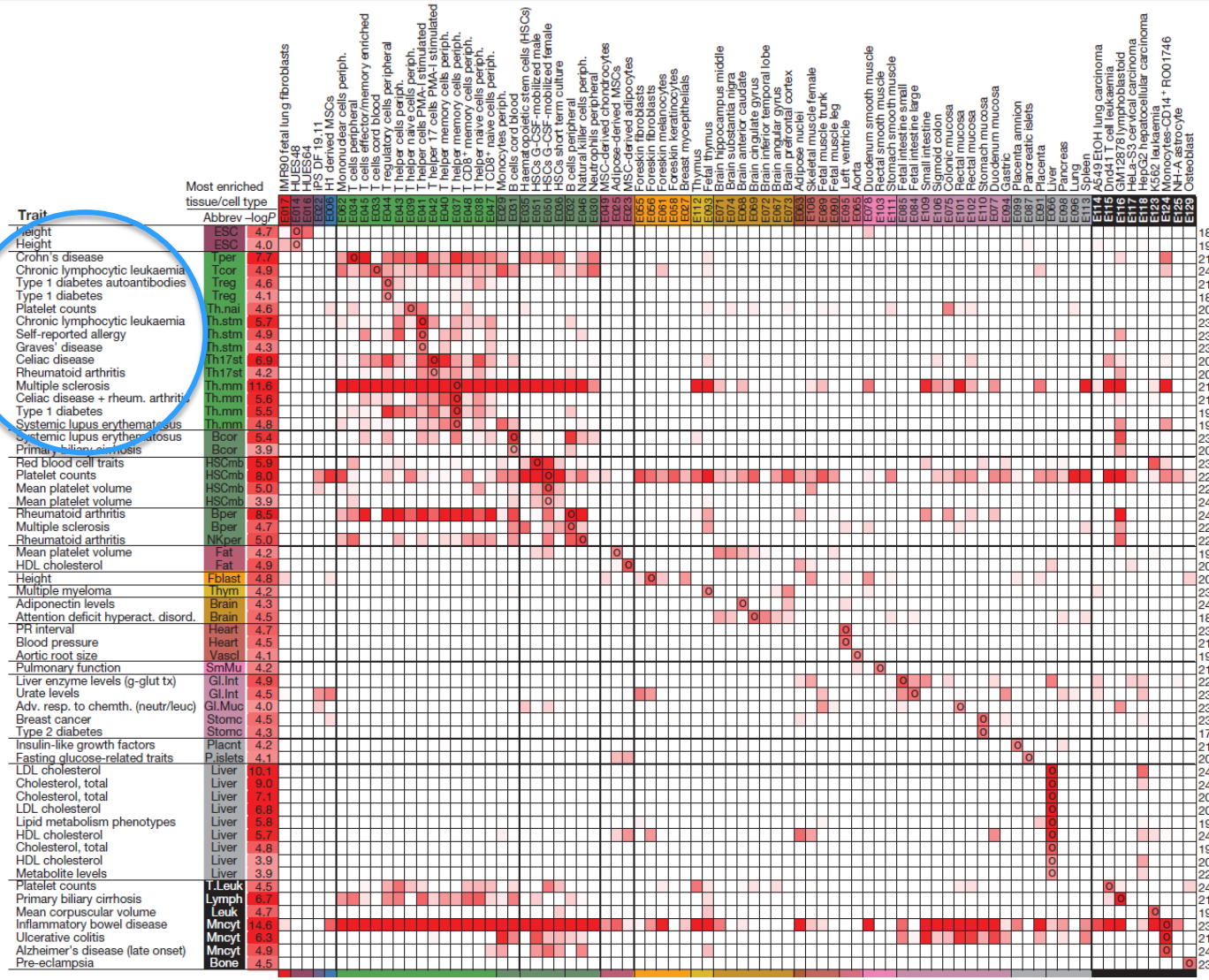
Epigenetic Dynamics - Regulators

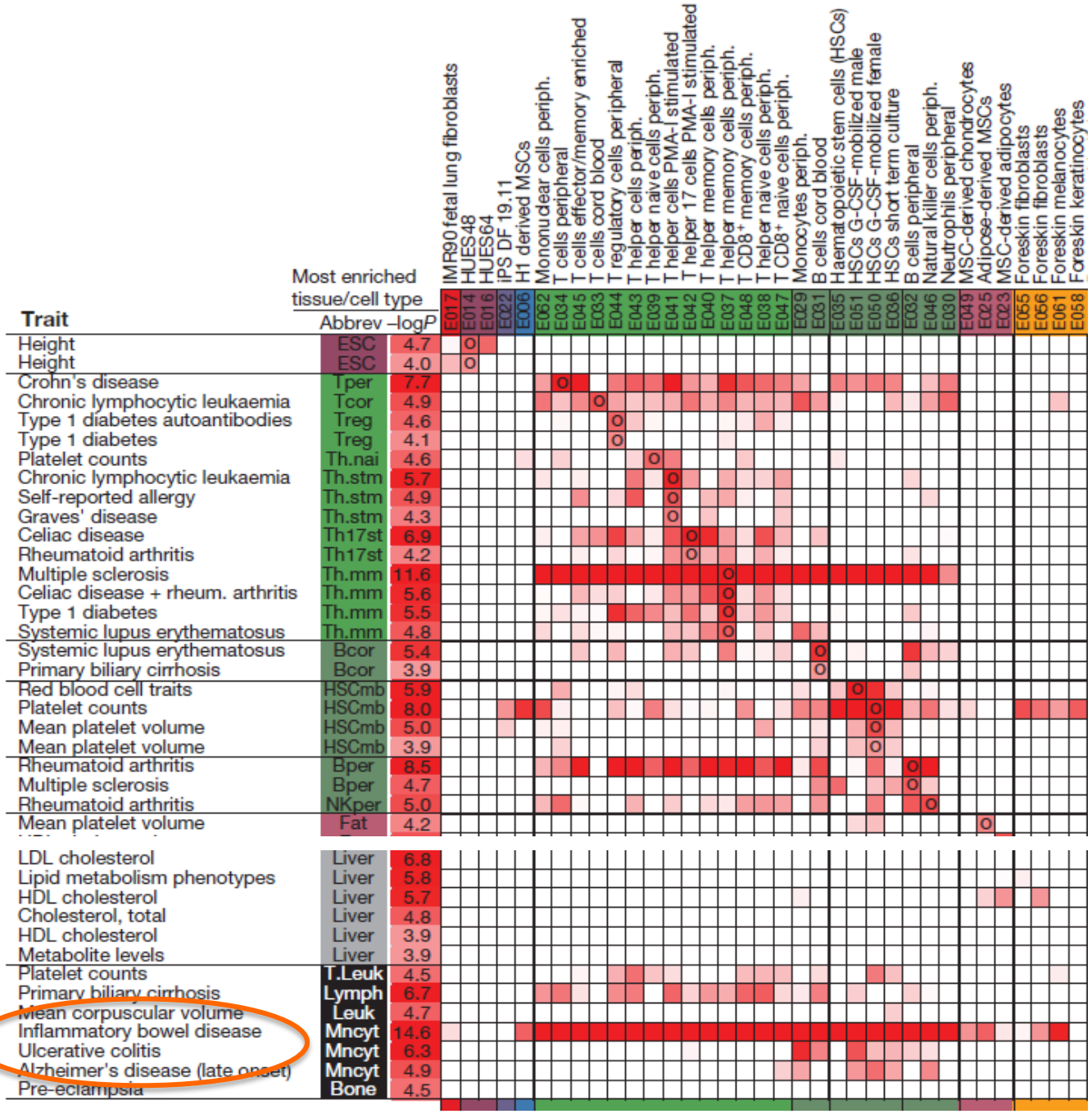


Linking Regulators to tissues & cell types



Linking Epigenomic Enrichments to Disease traits

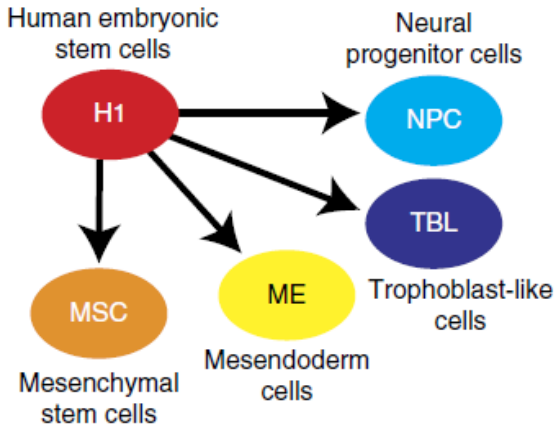



Trait
Most enriched tissue/cell type
Abbrev
−logP
IMR90 fetal lung fibroblasts
HUES48
HUES64
iPS DF 19.11
H1 derived MSCs
Mononuclear cells periph.
T cells peripheral
T cells effector/memory enriched
T cells cord blood
T regulatory cells peripheral
T helper cells periph.
T helper naive cells periph.
T helper cells PMA-I stimulated
T helper 17 cells PMA-I stimulated
T helper memory cells periph.
T helper memory cells periph.
T helper memory cells periph.
T CD8+ memory cells periph.
T helper naive cells periph.
T CD8+ naive cells periph.
Monocytes periph.
B cells cord blood
Haematopoietic stem cells (HSCs)
HSCs G-CSF-mobilized male
HSCs G-CSF-mobilized female
HSCs short term culture
B cells peripheral
Natural killer cells periph.
Neutrophils peripheral
MSC-derived chondrocytes
Adipose-derived MSCs
MSC-derived adipocytes
Foreskin fibroblasts
Foreskin fibroblasts
Foreskin melanocytes
Foreskin keratinocytes
E017
E014
E016
E002
E006
E062
E034
E045
E033
E044
E043
E039
E041
E042
E040
E037
E048
E038
E047
E029
E031
E035
E051
E050
E036
E032
E046
E030
E049
E025
E023
E055
E056
E061
E058
Height ESC 4.7
Height ESC 4.0
Crohn's disease Tper 7.7
Chronic lymphocytic leukaemia Tcor 4.9
Type 1 diabetes autoantibodies Treg 4.6
Type 1 diabetes Treg 4.1
Platelet counts Th.nai 4.6
Chronic lymphocytic leukaemia Th.stm 5.7
Self-reported allergy Th.stm 4.9
Graves' disease Th.stm 4.3
Celiac disease Th17st 6.9
Rheumatoid arthritis Th17st 4.2
Multiple sclerosis Th.mm 11.6
Celiac disease + rheum. arthritis Th.mm 5.6
Type 1 diabetes Th.mm 5.5
Systemic lupus erythematosus Th.mm 4.8
Systemic lupus erythematosus Bcor 5.4
Primary biliary cirrhosis Bcor 3.9
Red blood cell traits HSCmb 5.9
Platelet counts HSCmb 8.0
Mean platelet volume HSCmb 5.0
Mean platelet volume HSCmb 3.9
Rheumatoid arthritis Bper 8.5
Multiple sclerosis Bper 4.7
Rheumatoid arthritis NKper 5.0
Mean platelet volume Fat 4.2
LDL cholesterol Liver 6.8
Lipid metabolism phenotypes Liver 5.8
HDL cholesterol Liver 5.7
Cholesterol, total Liver 4.8
HDL cholesterol Liver 3.9
Metabolite levels Liver 3.9
Platelet counts T.Leuk 4.5
Primary biliary cirrhosis Lymph 6.7
Mean corpuscular volume Leuk 4.7
Inflammatory bowel disease Mncyt 14.6
Ulcerative colitis Mncyt 6.3
Alzheimer's disease (late onset) Mncyt 4.9
Pre-eclampsia Bone 4.5

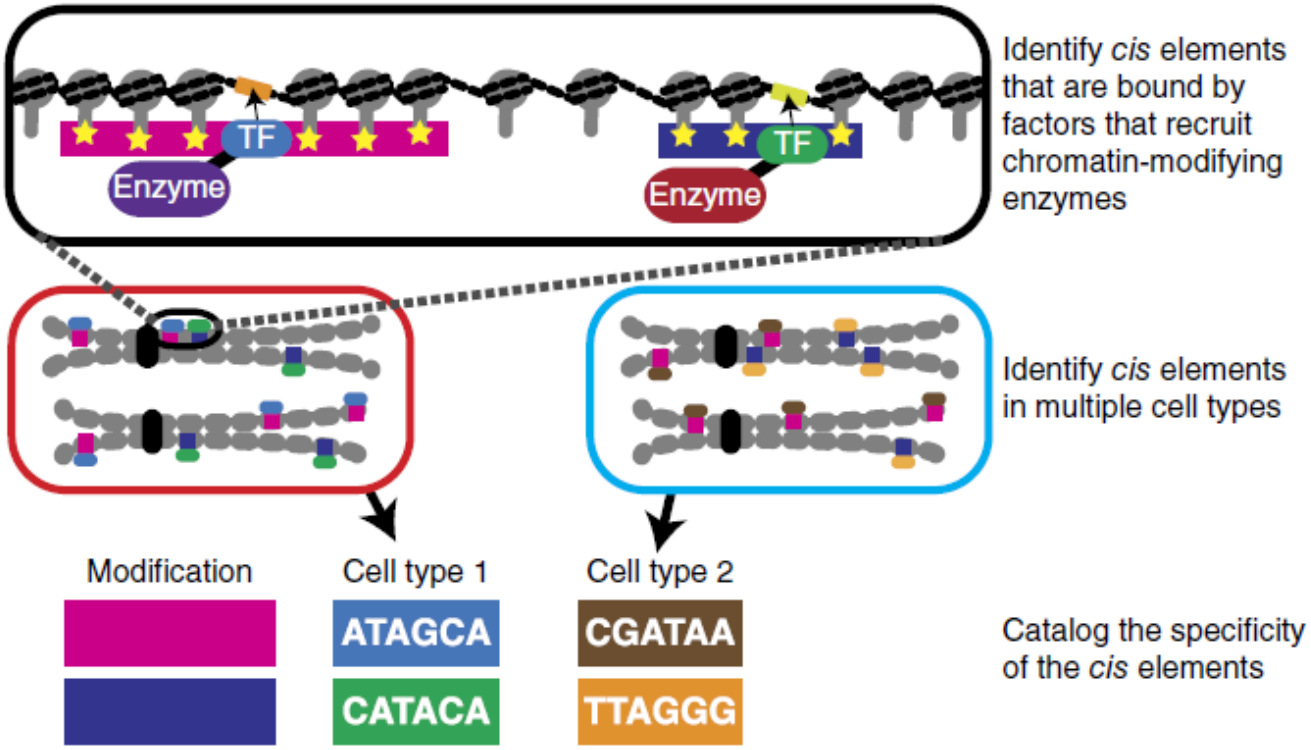
NATURE METHODS | VOL.12 NO.3 | MARCH 2015

Predicting the human epigenome from DNA motifs

John W Whitaker[1,3], Zhao Chen[1] & Wei Wang[1,2]



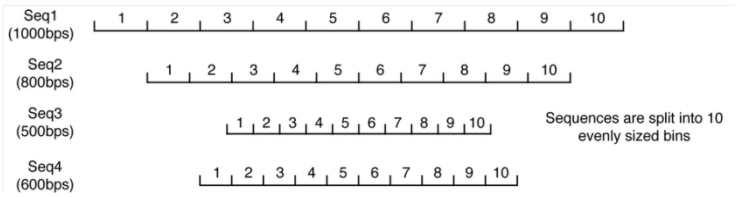
Analysis Pipeline



Analysis Pipeline

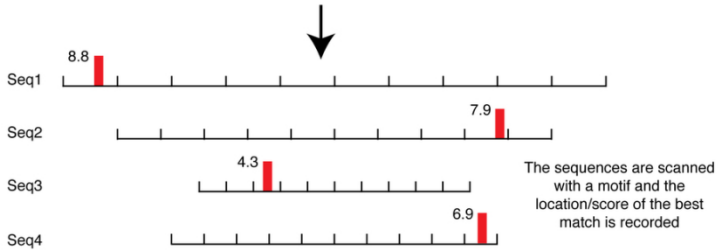
H3 modification	Function
K4me1	Enhancers
K4me3	Promoters
K9me3	Repressive
K27ac	Active
K27me3	Repressive
K36me3	Transcription
DNA methylation	Repressive

Large number of sequences

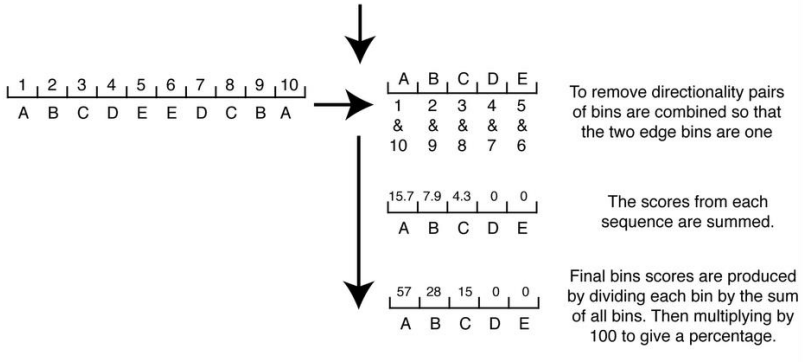


Variable size regions

Sequences are greatly unbalanced for G-C content



Motif Combinations



Analysis Pipeline -EPIGRAM

Reads normalized for G+C biases

Peak calling

Sequence-set balancing (SSB)

De novo motif discovery

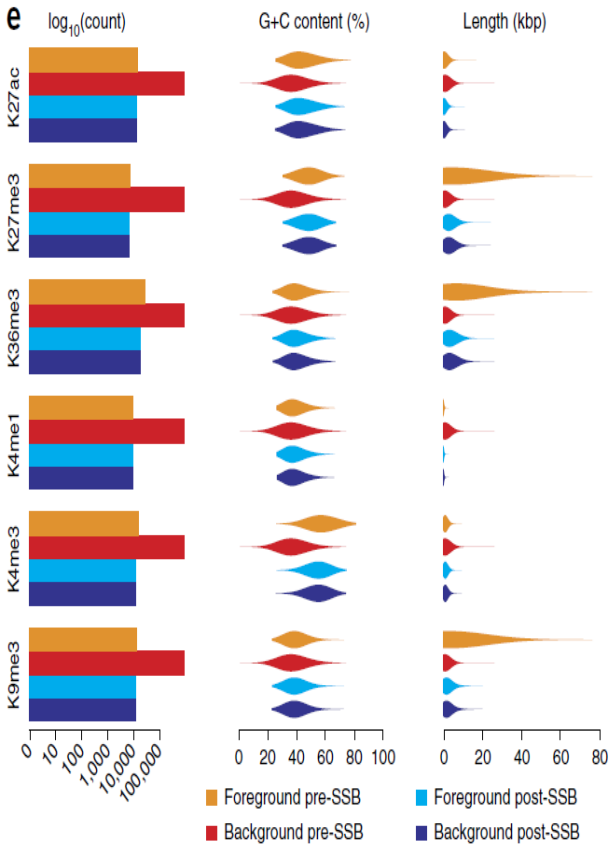
HOMER & EPIGRAM algorithms

Feature selection

Full set of motifs

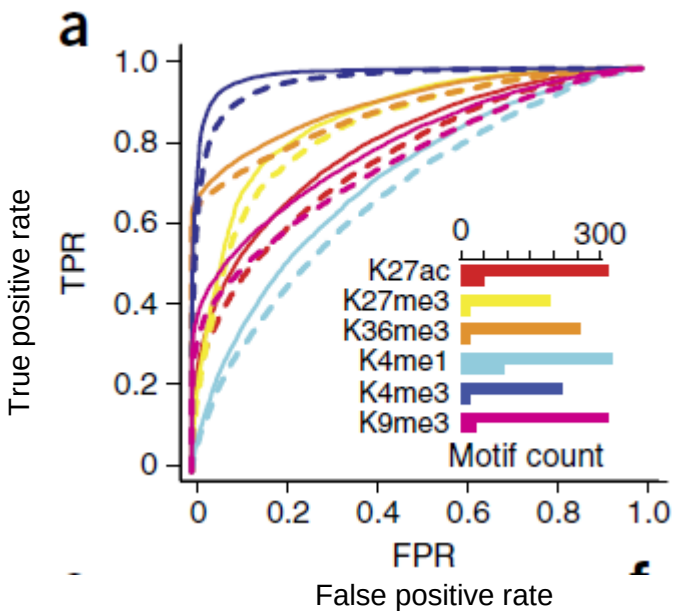
Random Forest classifier

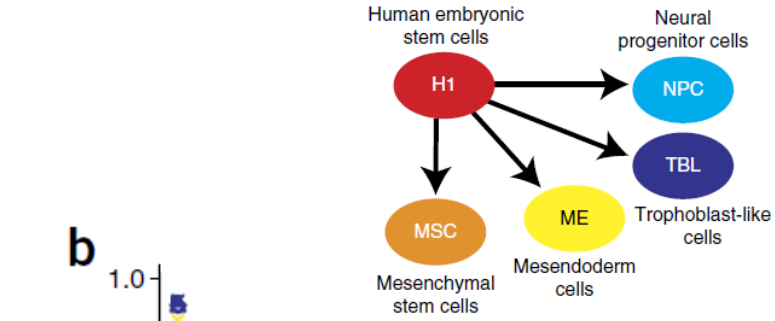
Effect of sequence-set balancing (SSB) on sequence sets

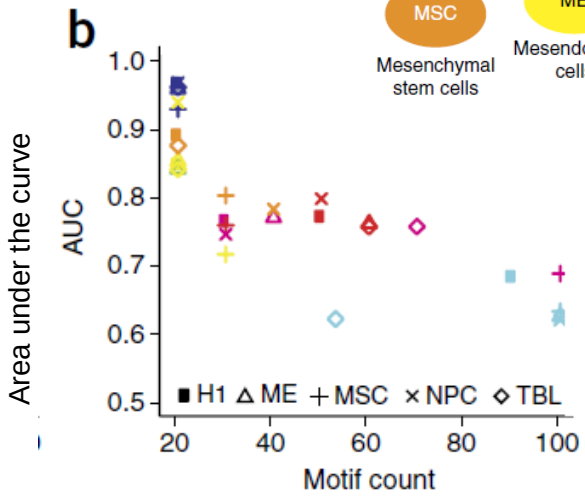


Prediction Performance

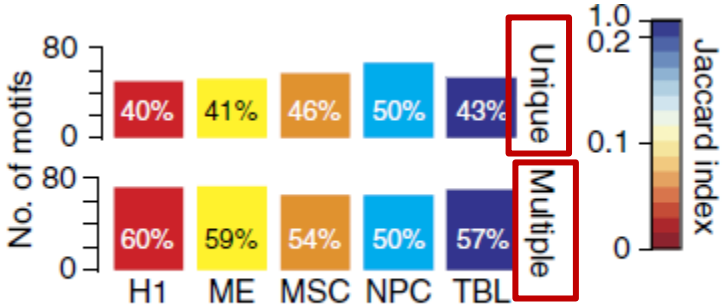
Receiver Operating Characteristic (ROC) curve



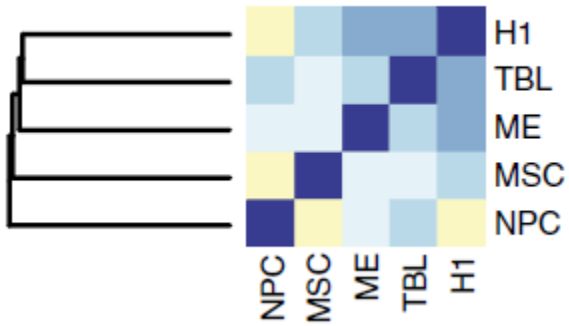




Motifs predictive of epigenetic modification

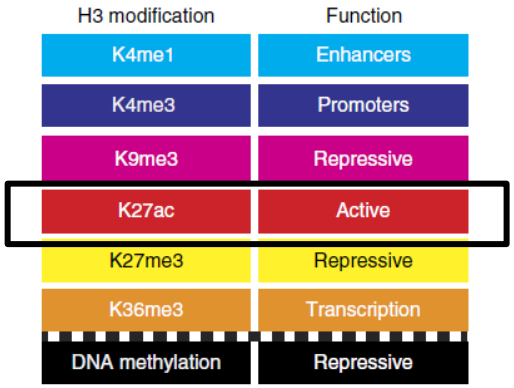


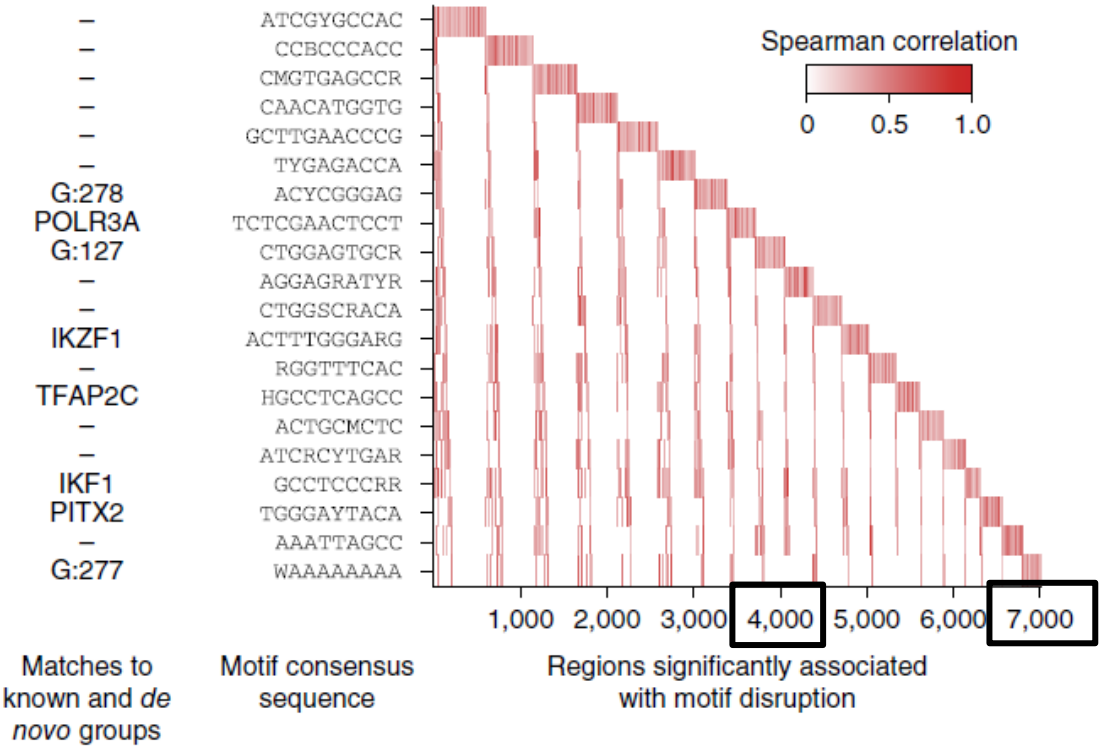
Heat Map



Motif distribution Paradigm

Motif distribution is correlated with H3K27ac variation





SUMMARY

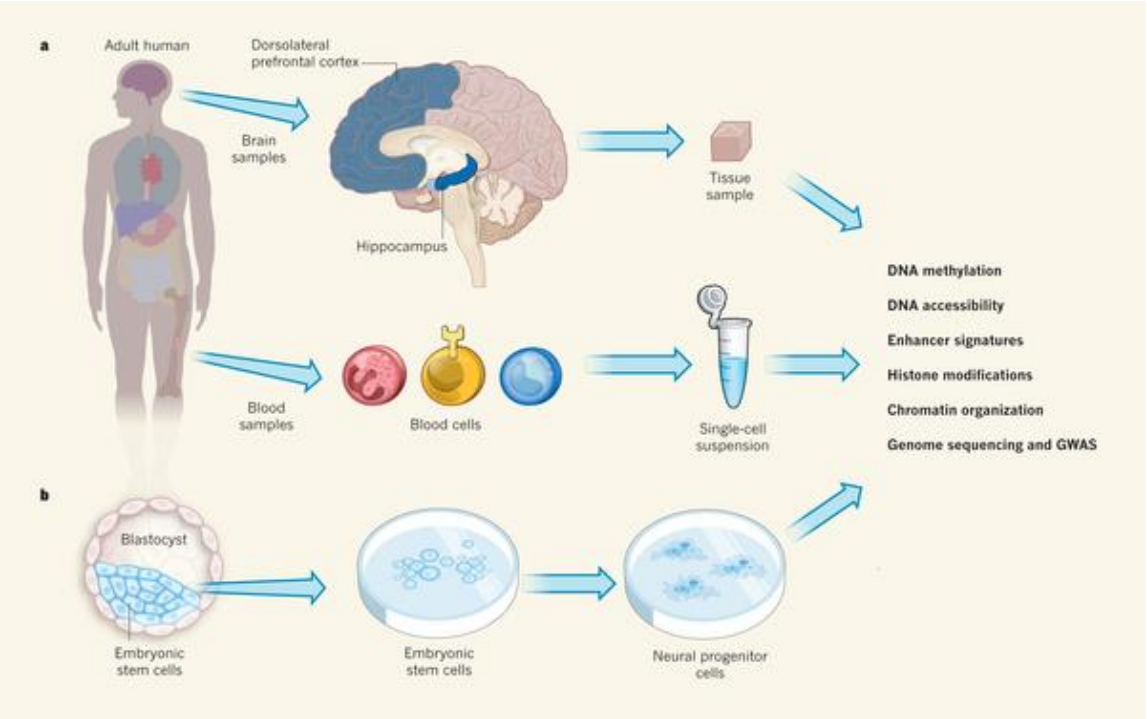
The first comprehensive catalog of DNA motifs

Define the mechanisms by which DNA motifs orchestrate the epigenome

In light of the genome editing technologies,
these approaches can be used to guide ***locus-specific epigenome editing*** *through alteration of the regulatory cis-elements.*

OVERVIEW

1. *How the epigenome affects gene expression?*
2. *How the epigenome changes during stem-cell differentiation (normal development)?*
3. *How the epigenome changes during disease?*



CAVEATS

1. *Studies are based on analysis of cell populations*

→ *Cellular Variability within populations*

2. *Tissue Sample: Enhancer landscapes represent the composite of cell types that make up the tissue*

→ *Use purified cell populations from in vivo sources*

3. *The DNA sequences found in cell specific enhancers provide clues for TF that regulate the enhancers*

→ *these clues must be validated experimentally*

Follow the Threads

EPIGENOMIC ANNOTATIONS

EPI TAGS

DEVELOPMENT

REGULATORY MODELS

Thread articles

THREAD 1

1. Annotation of the non-coding genome

Highlight referenced papers ▸

THREAD 2

2. Relationship between different epigenomic marks: DNA accessibility and methylation, histone marks, and RNA

Highlight referenced papers ▸

THREAD 3

3. Epigenomic changes during differentiation and development

Highlight referenced papers ▸

THREAD 4

4. Regulatory models: networks, motifs, modules, sequence drivers and predictive models

Highlight referenced papers ▸

THREAD 5

5. Interpreting variation: GWAS, cancer, genotype, evolution and allelic

Highlight referenced papers ▸

THREAD 6

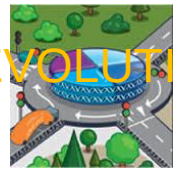
6. Epigenomic changes in human disease and during cancer progression

Highlight referenced papers ▸

THREAD 7

7. Brain epigenomes

Highlight referenced papers ▸



EVOLUTION

DISEASE

BRAIN

Thank you!