

ENCODE and Epigenomics Roadmap Workshop



Workshop Overview

1. Introduction

- Overview of the projects
- What data are available
- New features

2. How the scientific community has used ENCODE and Roadmap Epigenomics data

3. Interactive tutorial

- Commonly asked questions about data access
- Questions about analyzing the data
- We want to hear your questions



Project goals

NIH Roadmap Epigenomics Program: Research to transform our understanding of how epigenetics contributes to disease
(<http://commonfund.nih.gov/epigenomics>)

ENCODE (Encyclopedia of DNA Elements): Creating comprehensive, high quality catalogs of functional elements in the human and mouse genomes using high-throughput technologies. Completed analysis of the genomes of the fruit fly and roundworm.
(<http://www.genome.gov/encode/>)



Common purposes

1. Create a public resource of data
 - Experimental data available pre-publication
 - Encourage community use of datasets
2. Promote technology development
 - For epigenomics research
 - Identification and validation of functional elements
3. Development of standards
 - Definition of metadata
 - Standards for performing the assay
 - Uniform analysis of data
 - Standards for data quality



What data types are available?

	ENCODE	Roadmap Epigenomics
ChIP-seq (transcription factors, RNA-binding proteins)	•	
ChIP-seq (histone marks)	•	•
RNA-seq	•	•
RNA-seq after shRNA/siRNA knockdown	•	
DNA methylation	•	•
DNase	•	•
RNA-binding assays (CLIP-seq, RIP-seq)	•	
Enhancer assays	•	
5C	•	
Total data	~3400 datasets	~120 complete epigenomes



How can the data be used?

Hypothesis generation and refinement of gene regulation:

- What are the enhancers and promoter regions for a gene?
- What is the target gene for the regulatory regions?
- Is the regulation specific for certain cell types?

Many genetic findings for human disease map to non-protein coding regions of the human genome. ENCODE identified regions can help develop hypotheses:

- What is the causal variant?
- What is the function of the variant?



ENCODE data access

Consortium homepage <http://encodeproject.org>

- View metadata, protocols, and standards
- Data visualization at <http://genome.ucsc.edu>, <http://ensembl.org/>

NCBI <http://www.ncbi.nlm.nih.gov/bioproject/30707> (human)
<http://www.ncbi.nlm.nih.gov/bioproject/50617> (mouse)

- View and download data
- Compare samples

modENCODE <http://modencode.org>

- View and download data



Roadmap Epigenomics data access

Consortium homepage <http://roadmapepigenomics.org>

- Genome browser
- View protocols and standards

NCBI <http://ncbi.nlm.nih.gov/epigenomics>
<http://ncbi.nlm.nih.gov/geo/roadmap/epigenomics>

- View and download data
- Compare samples

Human Epigenome Atlas <http://epigenomeatlas.org>

- View data on genome or with Atlas gene browser
- Download data
- Use Genboree Workbench tools

WashU VizHub <http://vizhub.wustl.edu>

- Next-gen browser <http://epigenomegateway.wustl.edu>
- UCSC visualization hub at <http://genome.ucsc.edu>



Roadmap Epigenomics Mapping Consortium

- Reference Epigenome Mapping Centers
 - Brad Bernstein & Alex Meissner* (Broad)
 - Joe Costello* (UCSF)
 - Bing Ren* (UCSD)
 - John Stamatoyannopoulos* (Washington)
- Epigenomics Data Analysis and Coordination Center
 - Aleks Milosavljevic* (Baylor)



ENCODE Consortium

ENCODE Production Centers

Bradley Bernstein (John Rinn, Manolis Kellis)

Thomas Gingeras (Carrie Davis, Roderic Guigo)

Brenton Graveley (Christopher Burge, Xiang-Dong Fu, Eugene Yeo)

Richard Myers (Devin Absher, Gregory Cooper, Shawn Levy, Florencia Pauli Behn, Ross Hardison, Ali Mortazavi, Timothy Reddy, Barbara Wold)

Bing Ren (Joseph Ecker, Len Pennacchio, Axel Visel, Wei Wang)

Michael Snyder (Kevin White, Sherman Weissman, Peggy Farnham)

John Stamatoyannopoulos (Ralph Hansen, Rajinder Kaul, Patrick Navas, George Stamatoyannopoulos, Piper Treuting, Michael Bender, Job Dekker, Mark Groudine)

ENCODE Data Coordination Center

Mike Cherry (Jim Kent)

ENCODE Data Analysis Center

Zhiping Weng (Mark Gerstein, Manolis Kellis, Roderic Guigo, Rafael Irizarry, Xiaole Shirley Liu, William Stafford Noble)

Additional ENCODE Participants

Timothy Hubbard (Mark Gerstein, Roderic Guigo, Jen Harrow, Rachel Harte, David Haussler, Manolis Kellis, Alexandre Reymond, Stephen Searle, Alfonso Valencia)

David Gilbert (Tamer Kahveci)

ENCODE Computational Analysis Groups

Peter Bickel (Haiyan Huang, Leonard Lipovich, Bin Yu)

David Gifford (Tommi Jaakkola)

Sunduz Keles (Emery Bresnick, Colin Dewey)

Robert Klein (Christina Leslie, Souma Raychaudhuri, Ross Levine, Kenneth Offit)

Jonathan Pritchard (Yoav Gilad)

Xinshu Xiao

ENCODE Technology Development Groups

Christopher Burge (Wendy Gilbert, Brenton Graveley, Robert Horvitz)

Barak Cohen and Joseph Corbo

Peggy Farnham (Victor Jin, David Jay Segal)

R. David Hawkins

Christina Leslie (Christopher Mason)

Jason Lieb (Karen Mohlke, Eran Segal)

Mats Ljungman (Thomas Wilson)

Tarjei Mikkelsen

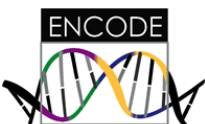
Jay Shendure and Nadav Ahituv (Michael McManus)

Alexey Wolfson

Guo-Cheng Yuan (Stuart Orkin)

... and many senior scientists, postdocs, students, technicians, computer scientists, statisticians and administrators in these groups

Current ENCODE participants: <http://www.genome.gov/26525220>



Future directions

1. Shared metadata standard among multiple projects: ENCODE, modENCODE, Roadmap Epigenomics
2. Annotation of metadata using ontologies
3. Biologically intuitive displays of metadata
4. Easier identification of data sets using metadata



Biologically relevant view of metadata

Eurie Hong (#1462T)

ENCODE Antibodies Biosamples Experiments Targets Esther Chan

Biosample > Tissue > Human


ENCBS086RNA / Tissue

Term name: camera-type eye
Term ID: UBERON:0000019
Description: Human fetal whole eye
Source: [BioChain Institute Inc](#)
Product ID: [R1244108-10-D02](#)
Lot ID: B608426
Project: ENCODE3
Submitted By: Carrie Davis
Lab: Thomas Gingeras, CSHL
Grant: U54HG007004
Note: Purchased as total RNA

Donor Information

Donor ID: ENCD0064AAA
Age: 20 week fetal
Sex: Female
Strain: N/A
Health status: Normal

Protocols and supporting documents

 **Data Sheet**
BioChain Human fetal whole eye B608426 data sheet
Submitted By: Carrie Davis
Lab: Thomas Gingeras, CSHL
Grant: U54HG007004
Download: [B608426.pdf](#)

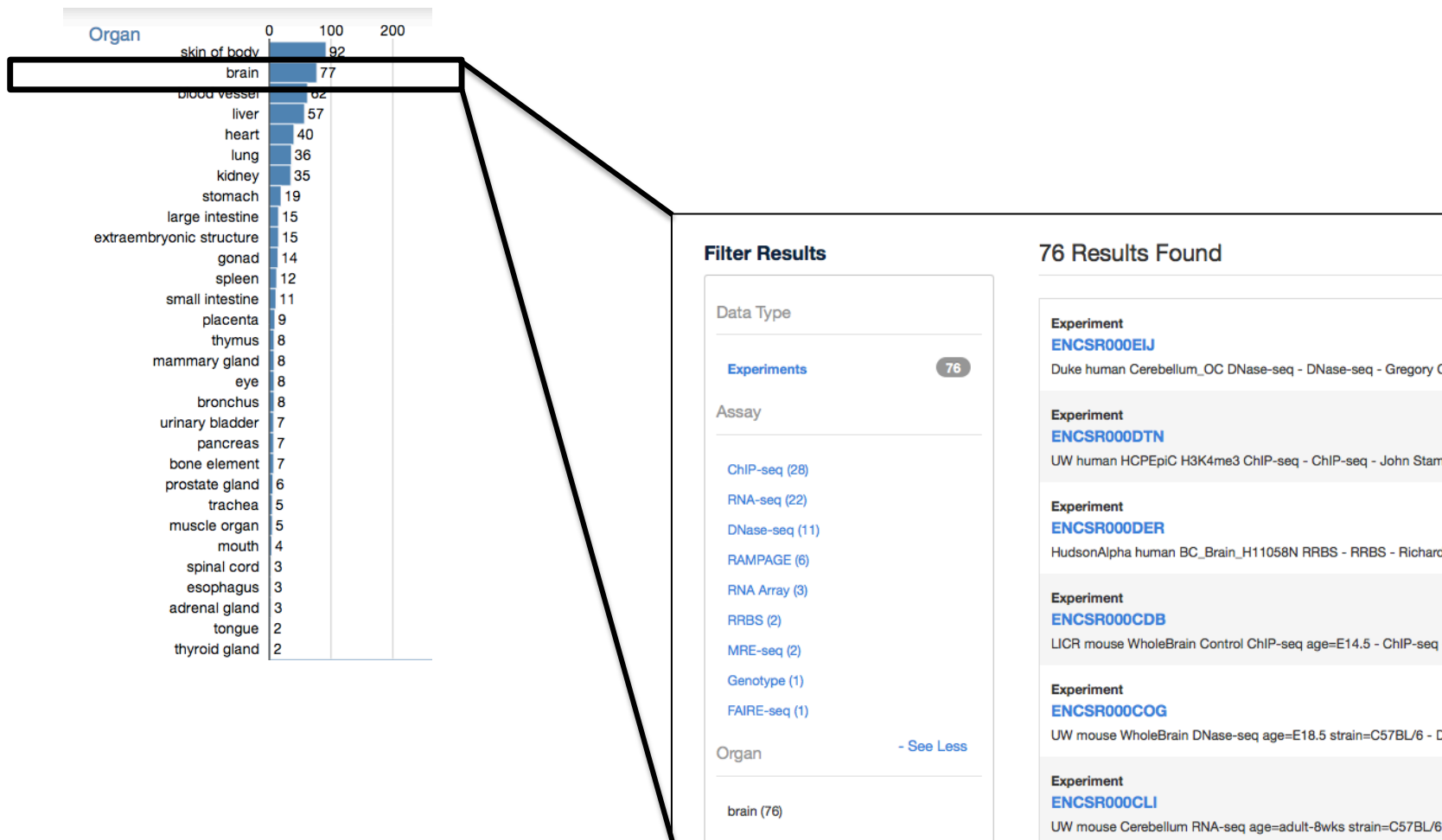
Details describing the biosample

De-identified information about the donor

Protocols and additional documentation

Using ontologies for searching

Venkat Malladi (#1613F)



Contact us



encode-help@lists.stanford.edu



[@encodedcc](https://twitter.com/encodedcc)

Speakers

Genetics and Genomics of Maternal Glycemia During Pregnancy: New Insights from Genome Wide Mapping and ENCODE

William L. Lowe, Jr., M.D.

Professor of Medicine and Vice Dean for Academic Affairs

Northwestern University

Expression and Regulation of Lung Function Genome-Wide Associated Locus

Maen Obeidat, Ph.D.

Post-doctoral fellow (PI: Peter Pare, Ph.D.)

University of British Columbia

Using the Epigenome Roadmap data to analyze genetic studies of type 2 diabetes

Vineeta Agarwala (Ph.D. student; PI: David Altshuler, M.D., Ph.D.; Broad Institute)

Abhishek Sarkar (Ph.D. student; PI: Manolis Kellis, Ph.D.; Broad Institute)

Kyle Gaulton (Post-doctoral fellow; PI: Mark McCarthy, Peter Donnelly; Wellcome Trust Centre for Human Genetics)



Tutorial

- How can ENCODE data help with analysis of SNP data from GWAS studies? Jill Moore, Luke Ward, Robert Klein
- How can ENCODE data help with annotation of non-coding variants? Jill Moore, Ekta Khurana, Robert Klein
- How can ENCODE data help identify how my gene of interest is regulated? Alon Goren, Brian Lee, Adam Frankish (single gene level analysis), Ross Hardison
- What can ENCODE data tell me about my gene's expression levels? Brian Lee, Eurie Hong, Adam Frankish (single gene level analysis)
- How can I use mouse ENCODE data in my analyses? Feng Yue, Ross Hardison, Venkat Malladi
- Analyzing ChIP-seq data Alon Goren, Anshul Kundaje
- Basic introduction to accessing ENCODE/Roadmap Epigenomics data Anshul Kundaje, Jiang Zhu, Eurie Hong

