Biological Data Resources

Benjamin Hitz, PhD
4/1/2015 CBPSS Panel
http://cherylab.stanford.edu

SGD

- The original model organism database for the first eukaryotic genome (*Saccharomyces cerevisiae*) to be sequenced
- Stores genes and biological annotations extracted from scientific literature *via* manual curation

http://www.yeastgenome.org/ … since 1993

SGD Annotations

GO: FUS3 encodes a MAP protein kinase as shown by direct enzyme assay in Bao et al. (2004).

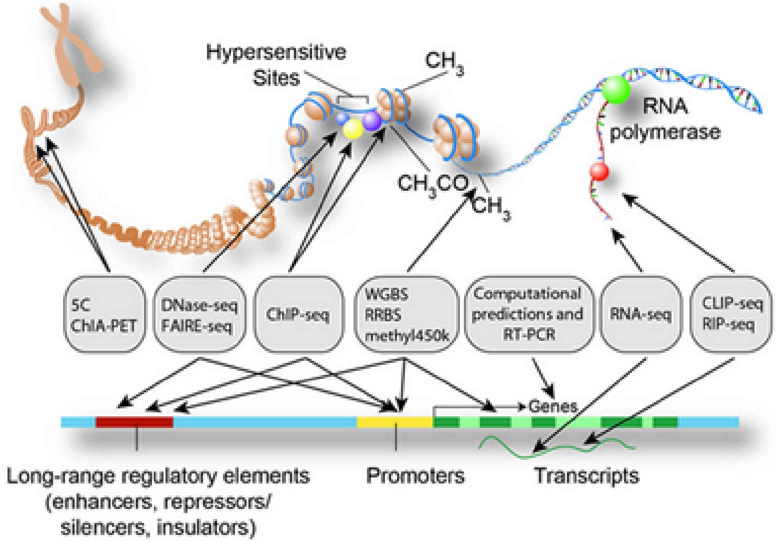
Phenotype: A null mutant of VAC14 has abnormal vacuolar morphology in S288C strain background as shown in Alghamdi et al. (2013).

Bioentity (what is it)	**Bioconcept** (what it does)	**Reference** (who said it)	**Experiment** (how is it known)	**Strain** (details)
FUS3	GO: MAP kinase activity	PMID:15620357	Direct enzyme assay	
VAC14	Phenotype: Vacuolar morphology	PMID: 23389034	Classical genetics	S288C

Additional properties (e.g., allele, conditions, qualifier) can be attached to any annotation.

ENCODE Data▾ Methods▾ About ENCODE▾ Help▾ Search ENCODE Sign in

ENCODE: Encyclopedia of DNA Elements



The ENCODE (Encyclopedia of DNA Elements) Consortium is an international collaboration of research groups funded by the National Human Genome Research Institute (NHGRI). The goal of ENCODE is to build a comprehensive parts list of functional elements in the human genome, including elements that act at the protein and RNA levels, and regulatory elements that control cells and circumstances in which a gene is active.

Image credits: Darryl Leja (NHGRI), Ian Dunham (EBI), Michael Pazin (NHGRI)

Data

To find and download ENCODE Consortium data:

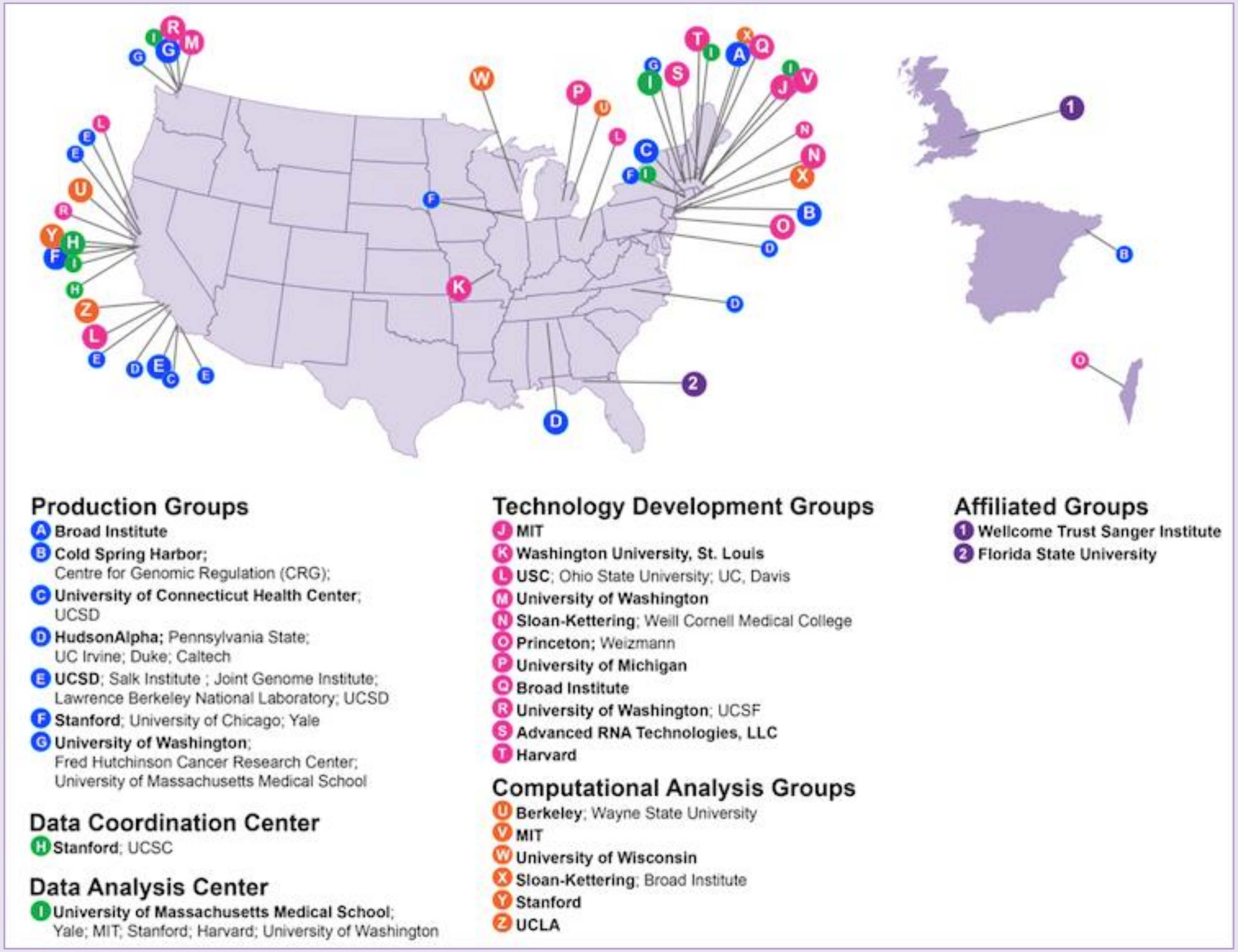
- Click the Data toolbar above and browse data

News

Sept 12, 2014: Data release: 23 human and 5 mouse datasets. [read more]

August 28, 2014: modENCODE and ENCODE comparison papers

What is the ENCODE Consortium?



Production Groups

- A **Broad Institute**
- B **Cold Spring Harbor;** Centre for Genomic Regulation (CRG);
- C **University of Connecticut Health Center**; UCSD
- D **HudsonAlpha;** Pennsylvania State; UC Irvine; Duke; Caltech
- E **UCSD**; Salk Institute ; Joint Genome Institute; Lawrence Berkeley National Laboratory; UCSD
- F **Stanford**; University of Chicago; Yale
- G **University of Washington**; Fred Hutchinson Cancer Research Center; University of Massachusetts Medical School

Data Coordination Center

- H **Stanford**; UCSC

Data Analysis Center

- I **University of Massachusetts Medical School**; Yale; MIT; Stanford; Harvard; University of Washington

Technology Development Groups

- J **MIT**
- K **Washington University, St. Louis**
- L **USC**; Ohio State University; UC, Davis
- M **University of Washington**
- N **Sloan-Kettering**; Weill Cornell Medical College
- O **Princeton;** Weizmann
- P **University of Michigan**
- Q **Broad Institute**
- R **University of Washington**; UCSF
- S **Advanced RNA Technologies, LLC**
- T **Harvard**

Computational Analysis Groups

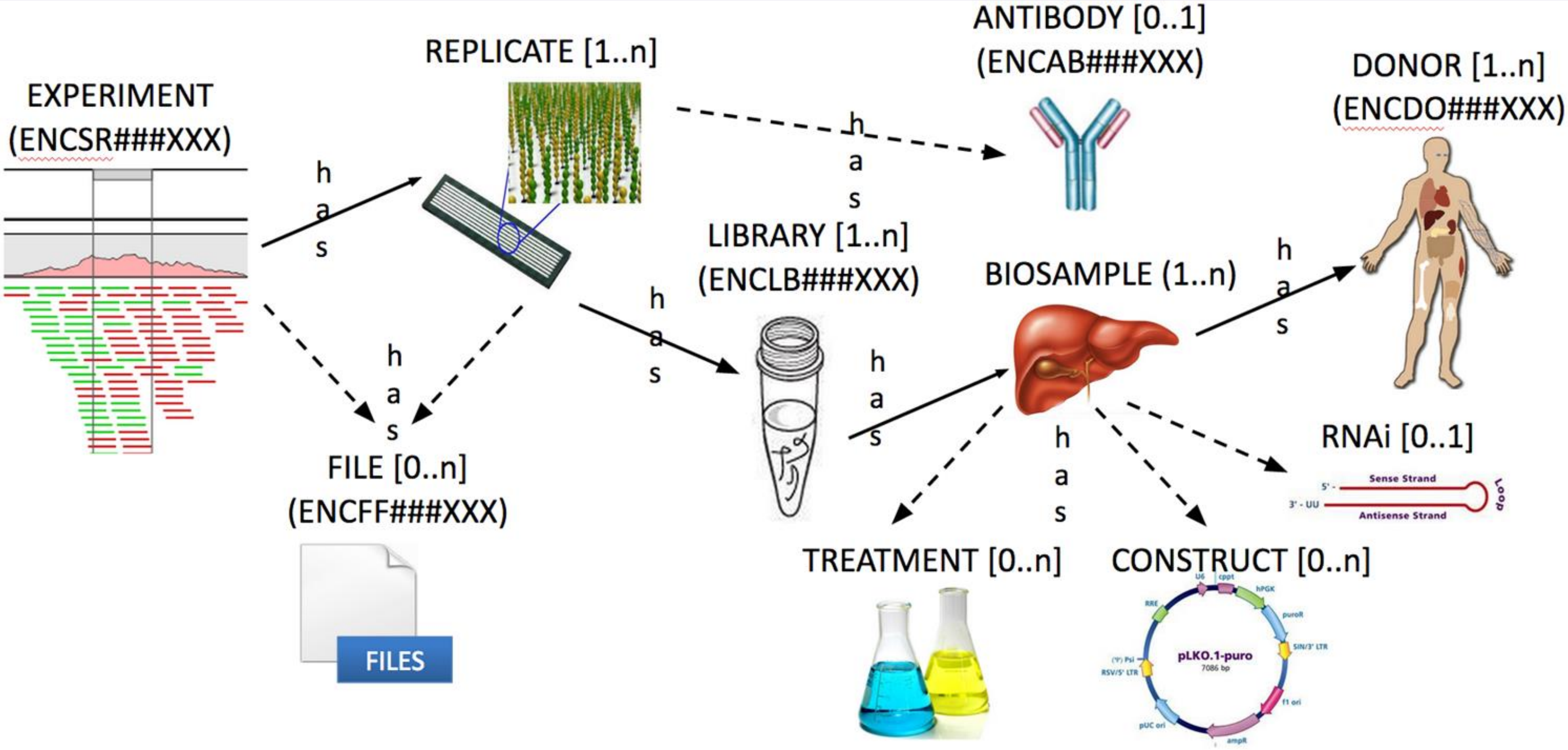
- U **Berkeley**; Wayne State University
- V **MIT**
- W **University of Wisconsin**
- X **Sloan-Kettering**; Broad Institute
- Y **Stanford**
- Z **UCLA**

Affiliated Groups

- 1 **Wellcome Trust Sanger Institute**
- 2 **Florida State University**

Image credit: NHGRI

Metadata model



Metadata submission

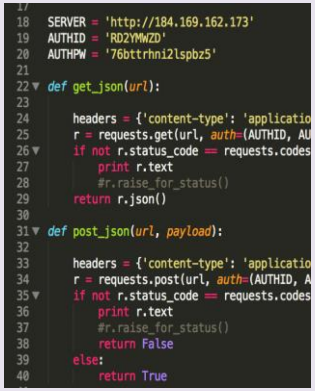
Metadata db

REST API

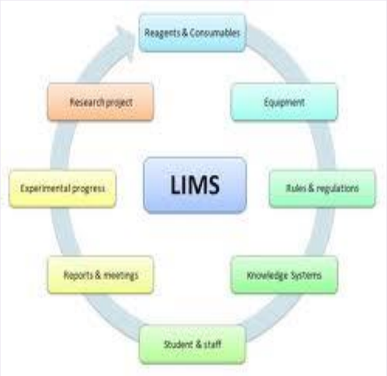
JSON Objects (GET, PATCH, POST)

Scripts

Wrangler Spreadsheets

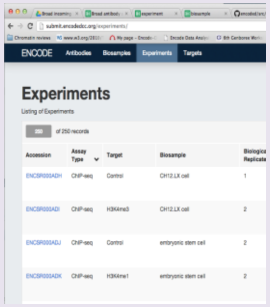


```
SERVER = 'http://184.169.162.173'
AUTHID = 'RD2YMWZD'
AUTHPW = '76bttrhni2lspbz5'

def get_json(url):

    headers = {'content-type': 'applicatio
    r = requests.get(url, auth=(AUTHID, AU
    if not r.status_code == requests.codes
        print r.text
        #r.raise_for_status()
    return r.json()

def post_json(url, payload):

    headers = {'content-type': 'applicatio
    r = requests.post(url, auth=(AUTHID, A
    if not r.status_code == requests.codes
        print r.text
        #r.raise_for_status()
        return False
    else:
        return True
```



Code

Web App

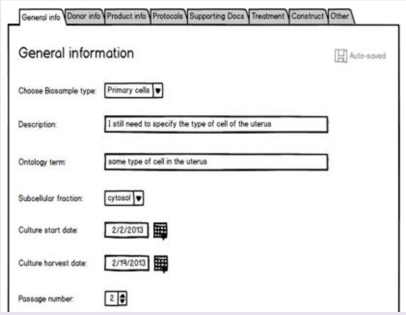


Experiments

Listing of Experiments

250 of 250 records

Accession	Assay Type	Target	Biosample	Biologica Replicate
ENCSR000ADH	ChIP-seq	Control	CH12.LX cell	1
ENCSR000ADI	ChIP-seq	H3K4me3	CH12.LX cell	2
ENCSR000ADJ	ChIP-seq	Control	embryonic stem cell	2
ENCSR000ADK	ChIP-seq	H3K4me1	embryonic stem cell	2



General information

Choose Biosample type: Primary cells

Description: I still need to specify the type of cell of the uterus

Ontology term: some type of cell in the uterus

Subcellular fraction: cytosol

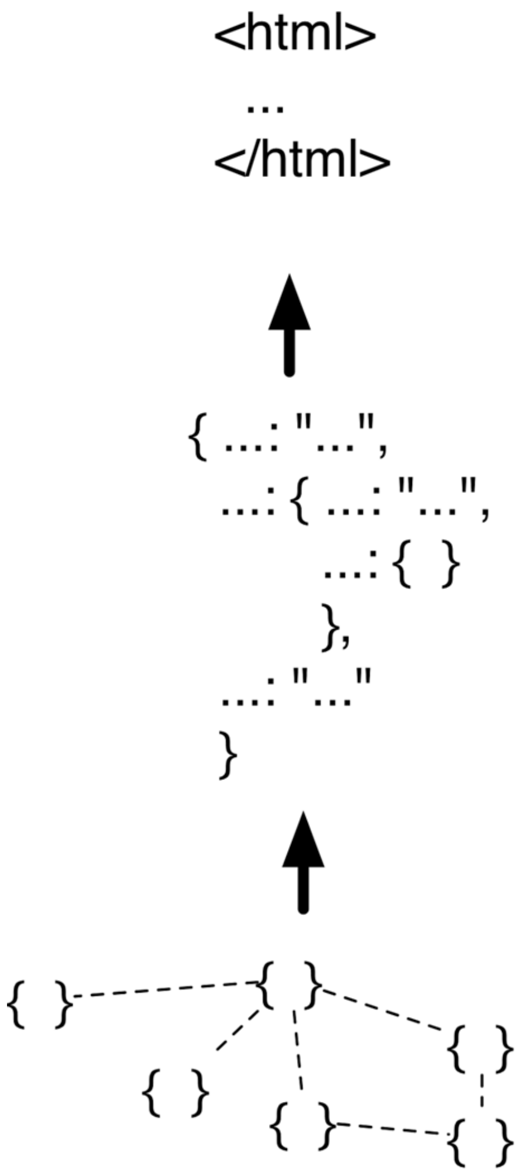
Culture start date: 2/2/2013

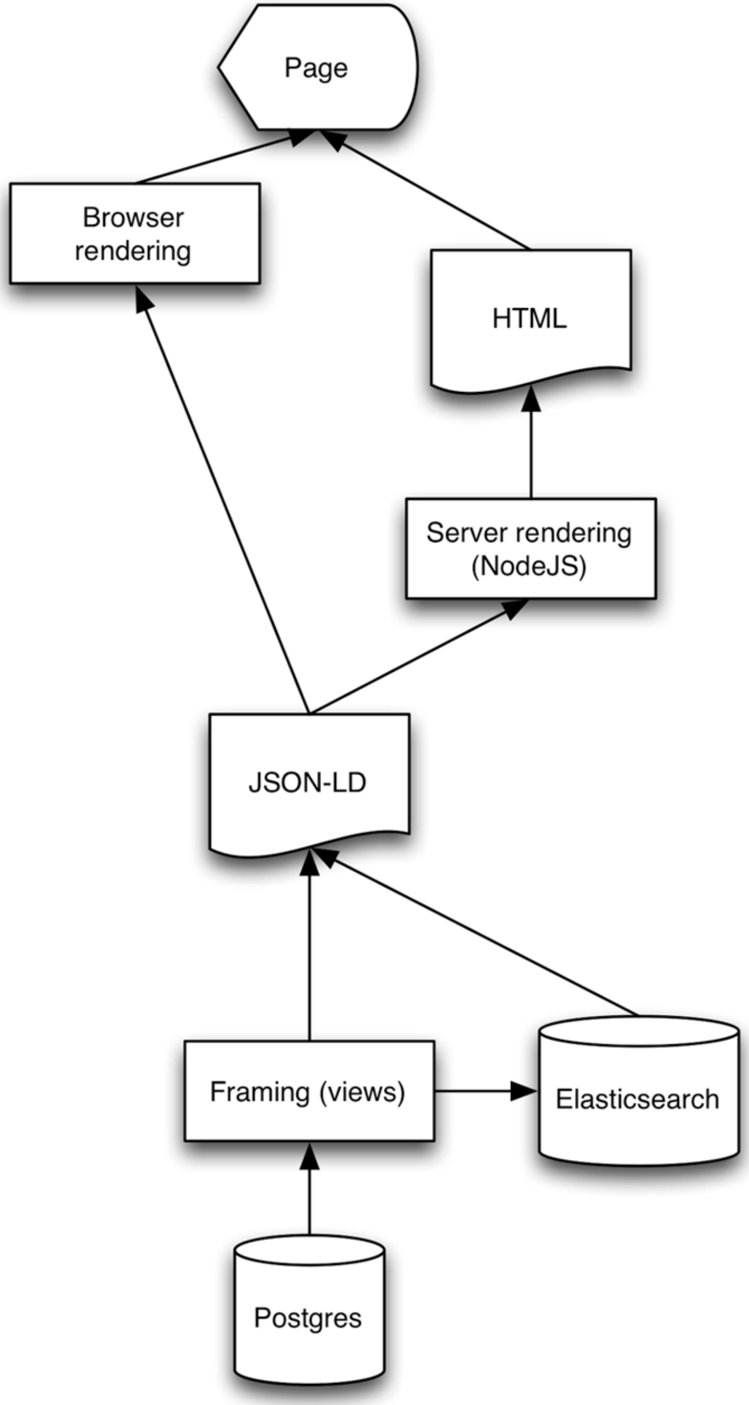
Culture harvest date: 2/19/2013

Passage number: 2

Web Forms

Software Stack




Page

Browser rendering

HTML

Server rendering (NodeJS)

JSON-LD

Framing (views)

Elasticsearch

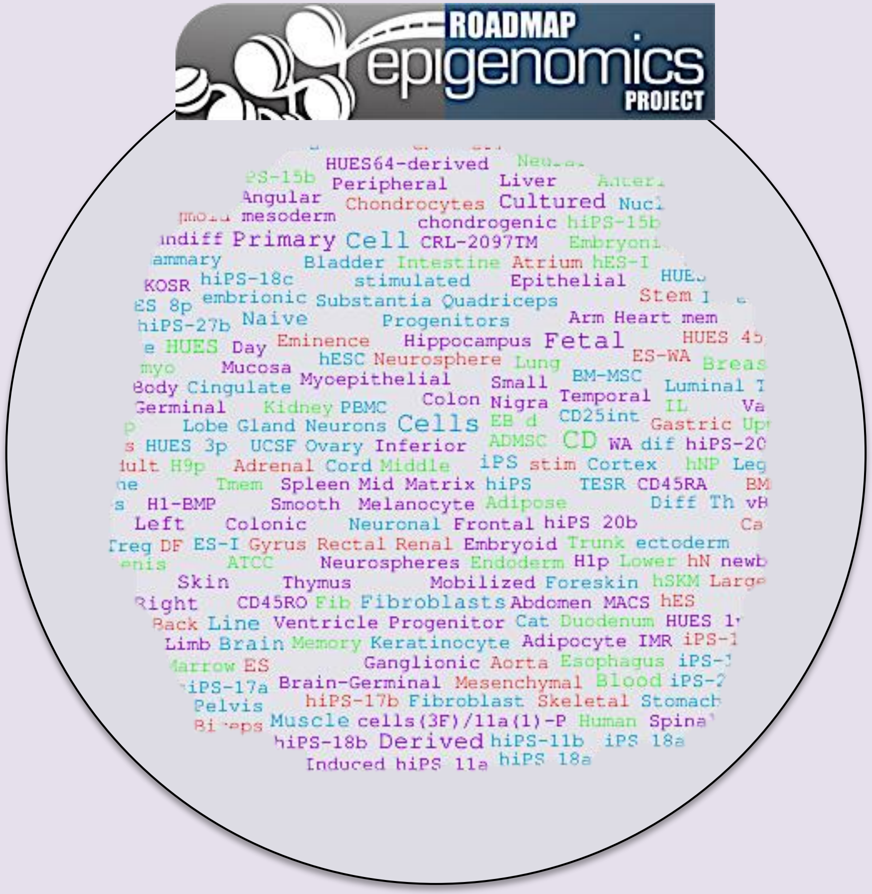
Postgres

Find Common Biosamples Between Two Consortia

ENCODE
HMVEC-dNeo BC Brain H11058N
Adipose UHN HSMMtube FSHD Panls
BC Prostate Gland H12817N Monocytes-CD14+ RO Medull
Myometr Cord CD4 naive HMVEC-dBl-Neo HWP
HPDE6-E6E BC Pericardium H12529N HSMM FSHD BC Colon T-47D
tate HSMMtube HMVEC-dBl-Ad BC Skeletal Muscle SK-N-
HPC-PL HFL24W H9ES-EBD hMNC-PB Cerebellum OC PBDE
HMVEC-dLy-Ad hMSC-BM BC Esophagus H9ES-CM HIPEpiC M059
HSMM emb BC Left Ventricle N H1-neurons HMVEC-dAd Kidney
MRT A hMSC-UC HEK293-T-REx GM12878-XiMat SkMC Treg Wb
HPIEpC H9ES-EB BC Lung HCH 0011308.2P HSaVEC Dnd NHDF
hMSC-BM hMSC-AT BC Pancreas H12817N HFF HSaVEC MRT T
MRT G HPAEC Caco H7-hESC Adult CD4 Th Chorion SKMC HPAEpi
HMF hMSC-UC bone marrow MSC BC Uterus BN HFL11W HMVEC-
CWRU HMEC HL CD4+ Naive Wb Fibrobl GM HConF HCF BC Skin Decidua
CaP HRE HPIEpC HEEpiC BG02ES GM FibroP H9ES Liver STL
NHBE HOB HEK293(b Colo BC Leukocyte UHN HAoEC HA-sp hMNC-PB
IMR Th1 Wb HBVP BC Kidney H12817N Th BC Kidney HVMF HTR8svn
iPS NIHi HOB Cord CD4 Th Hepatocytes Fibrobl HFDPC CMK hMSC-AT
Melano ECC Heart OC FibroP AG HBVSMC HCM Glioblа HPF Kidney
HRCEpiC HeLa-S HFF-Myc Esophagus BC HCH HAL NHDF
HLF HSMM CD20+ RO CD AG BC Lung H12817N HepG HMEpC
Mel HVMF HGF HA-h HCFaa HepG2b HCH HEK H9ES-E Astrocy RP
HUVEC BJ HPC-PL GC B cell BE2 C BC Liver CLL HWP
K562b HCPEpiC HAoAF Frontal cortex OC HCT HT HRGEC
HRPEpiC HAc H9ES-AFP Cerebrum frontal OC HEK293T PBM
HPAF HPdLF Huh BC Esophagus H12817N H1-hESC HAoAF BC Testis N
HMEpC HAEpiC Breast OC BC Penis H12817N HHSEC BC Breast
cytes-CD Th2 Wb bone marrow HS27a HFDPC HAoEC AoAF NB
Jurkat NHEM M2 hMNC-CB CD34+ Mobilized Colon OC iPS
NHBE RA hMNC-CB BC Jejunum H12817N HBMEC BC Stomach
BC BC Rectum N BC Small Intestine Colon BC HNPCEpiC
AoSMC BC Skeletal Muscle H12817N NHEM.f M2 Ishikawa
HEM M BC Bladder BC Spleen H12817N iPS hFib2 iPS Ste
steobl HSMMtube emb BC Placenta UHN BG02ES EBD
OC HMVEC-LLy BC Stomach H12817N
Colon H12817N HMVEC-dLy

356 terms

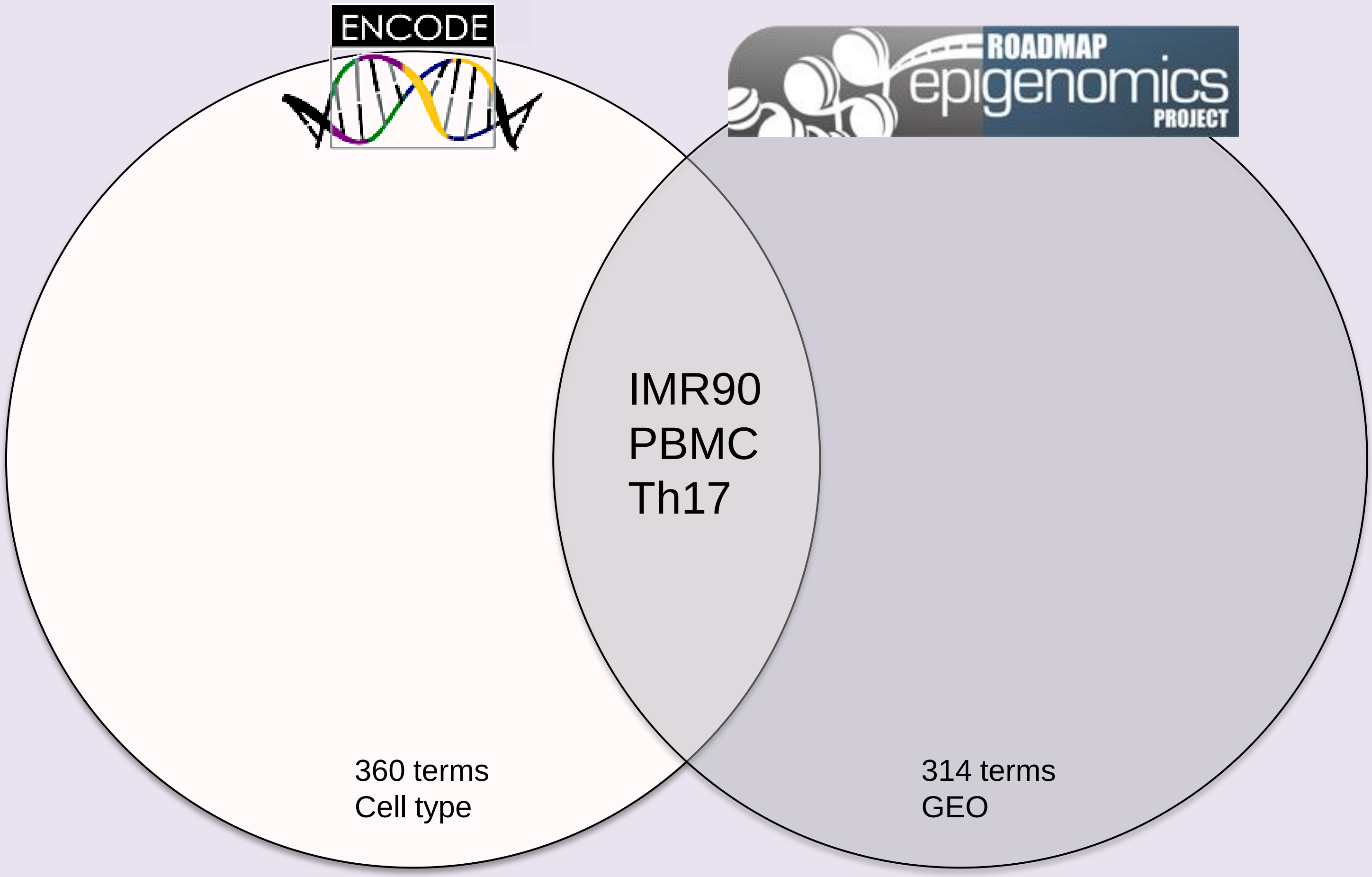
http://genome.ucsc.edu/ENCODE/cellTypes.html


ROADMAP epigenomics PROJECT
HUES64-derived Neural
iPS-15b Peripheral Liver Anterior
Angular Chondrocytes Cultured Nuclei
mesoderm chondrogenic hiPS-15b
undiff Primary Cell CRL-2097TM Embryonic
mammary Bladder Intestine Atrium hES-I
KOSR hiPS-18c stimulated Epithelial HUES
ES 8p embrionic Substantia Quadriceps Stem
hiPS-27b Naive Progenitors Arm Heart mem
HUES Day Eminence Hippocampus Fetal HUES 45
hESC Neurosphere Lung ES-WA Breast
Mucosa Myoepithelial Small BM-MSC Luminal
Body Cingulate Colon Nigra Temporal
Germinal Kidney PBMC CD25int IL
Lobe Gland Neurons Cells EB d Gastric
HUES 3p UCSF Ovary Inferior ADMSC CD WA dif hiPS-20
Adult H9p Adrenal Cord Middle iPS stim Cortex hNP Leg
Tmem Spleen Mid Matrix hiPS TESR CD45RA
H1-BMP Smooth Melanocyte Adipose Diff Th vB
Left Colonic Neuronal Frontal hiPS 20b
Treg DF ES-I Gyrus Rectal Renal Embryoid Trunk ectoderm
ATCC Neurospheres Endoderm H1p Lower hN newb
Skin Thymus Mobilized Foreskin hSKM Large
Right CD45RO Fib Fibroblasts Abdomen MACS hES
Back Line Ventricle Progenitor Cat Duodenum HUES
Limb Brain Memory Keratinocyte Adipocyte IMR iPS-
Marrow ES Ganglionic Aorta Esophagus iPS-
hiPS-17a Brain-Germinal Mesenchymal Blood iPS-
Pelvis hiPS-17b Fibroblast Skeletal Stomach
Biceps Muscle cells(3F)/11a(1)-P Human Spinal
hiPS-18b Derived hiPS-11b iPS 18a
Induced hiPS 11a hiPS 18a

314 terms

GEO characteristics: common_name, tissue_type, cell_type, lines

Labs were internally consistent



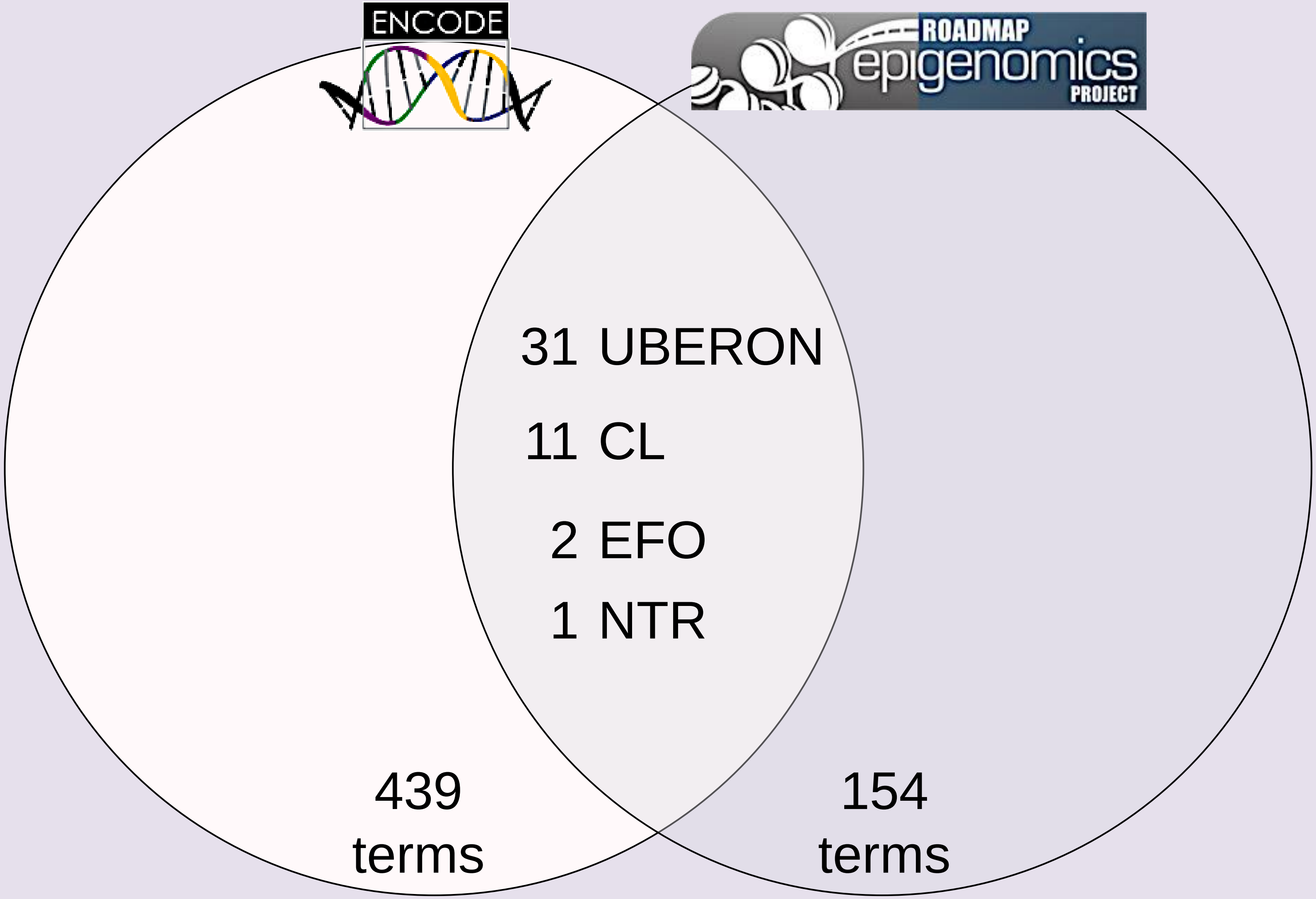
IMR90
PBMC
Th17

360 terms
Cell type

314 terms
GEO

… but only 3 biosample names match exactly between projects

45 Biosamples in Common Between Current ENCODE & REMC



31 UBERON

11 CL

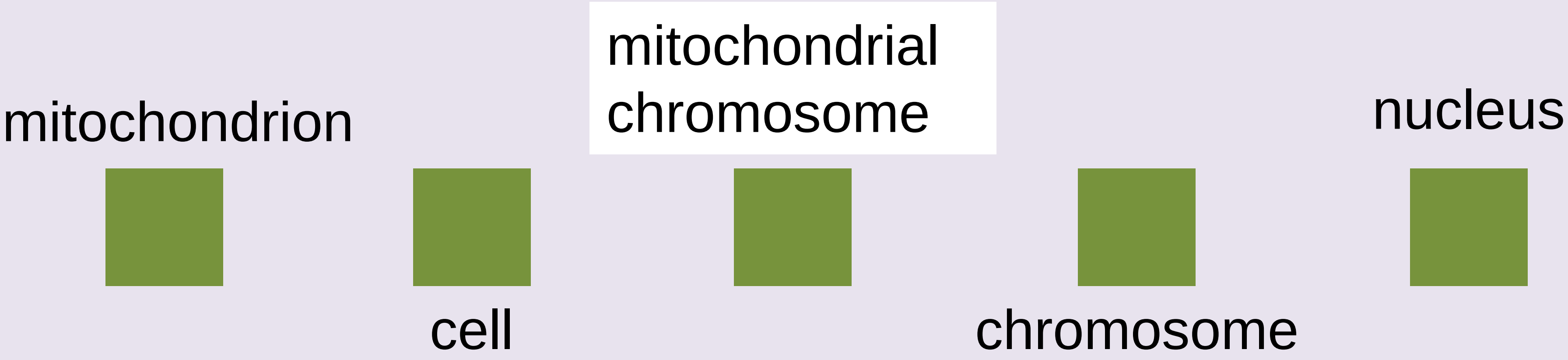
2 EFO

1 NTR

439 terms

154 terms

An ontology is a set of words...



mitochondrion

cell

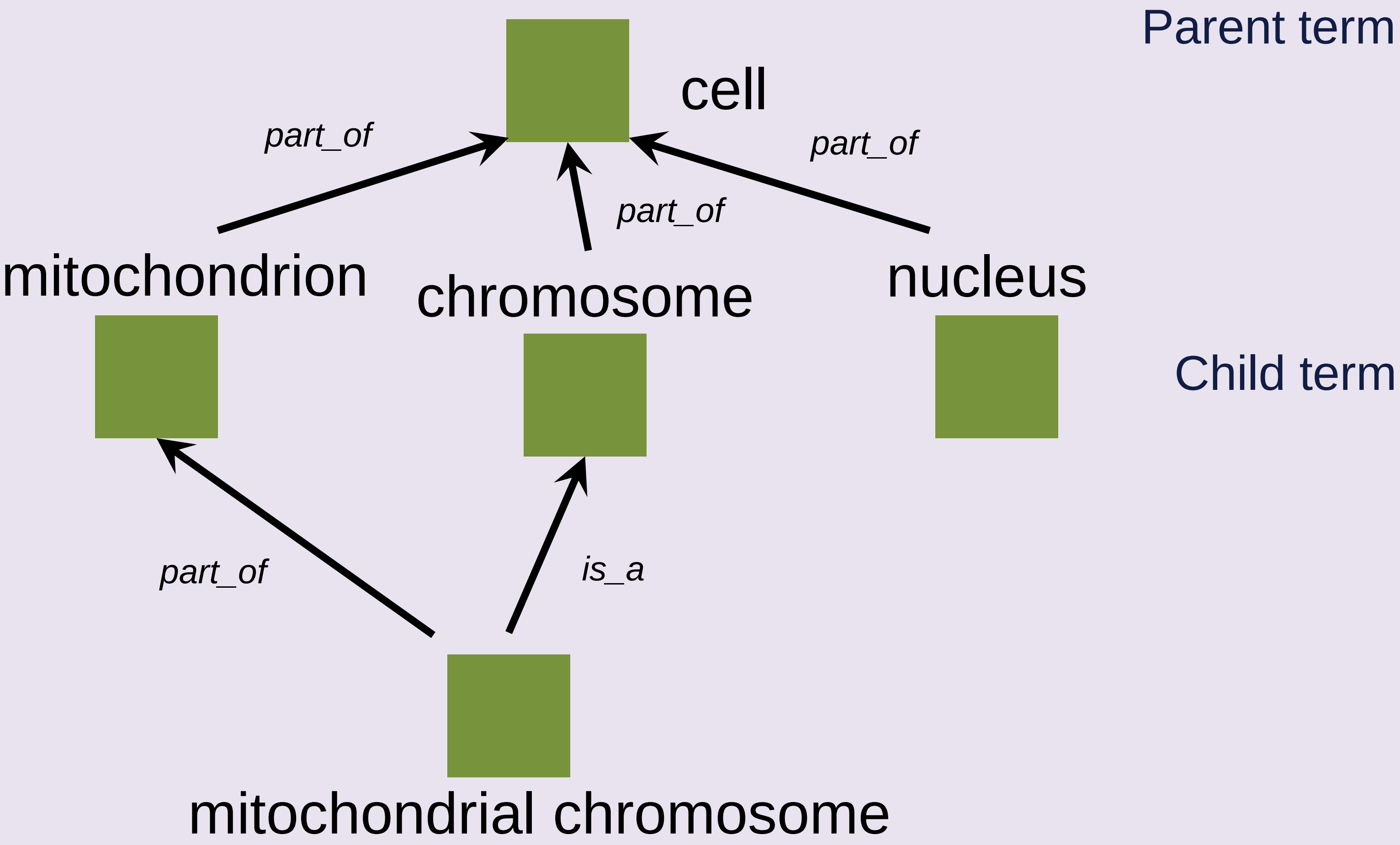
mitochondrial chromosome

chromosome

nucleus

An ontology is a set of words...

.. with different types of relationships to each other.



Why use ontologies?

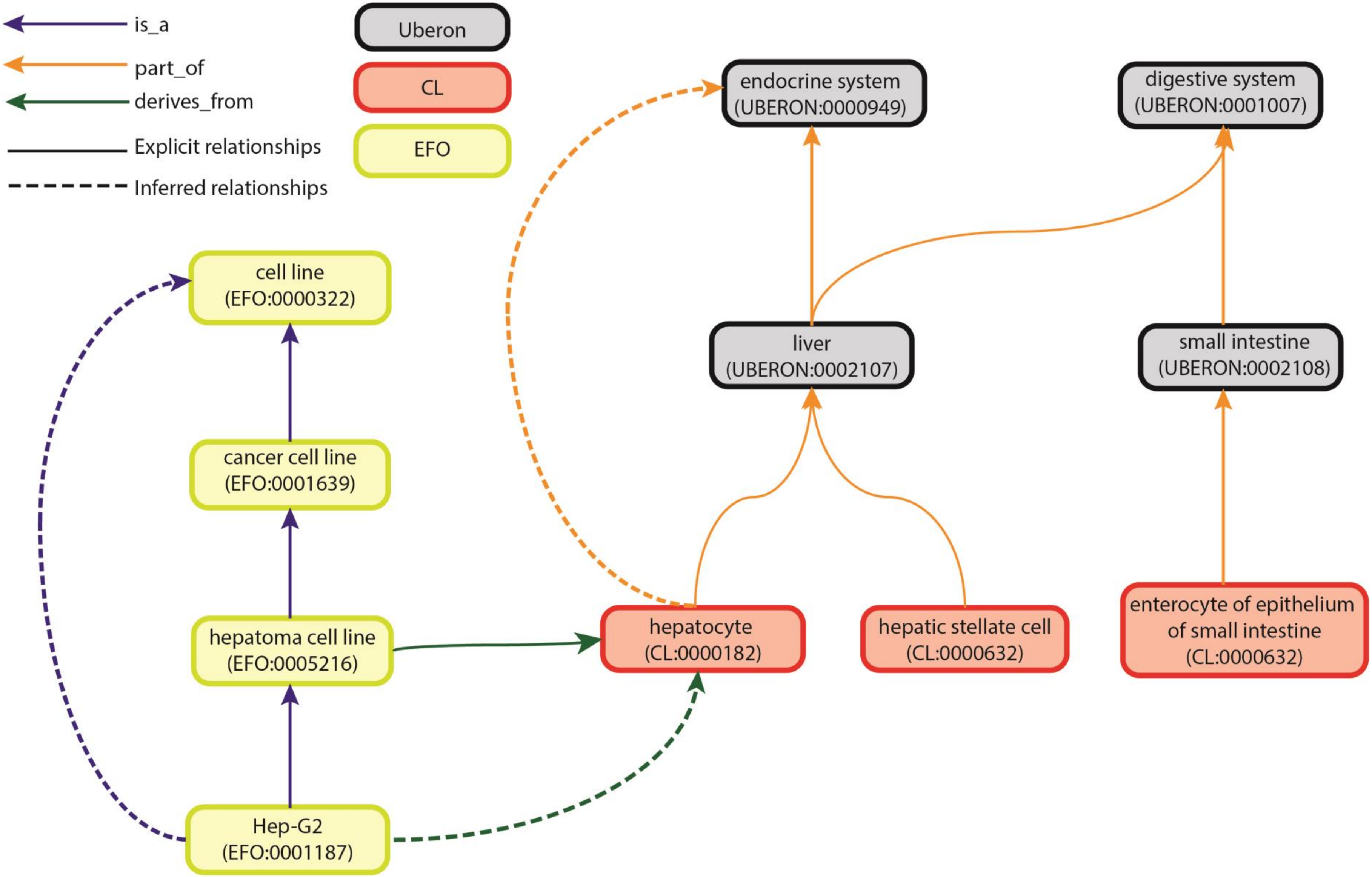
- Consistency of language and identifiers facilitates identification of data programmatically. Alternative spellings & phrases are synonyms. Independent of a particular data model.

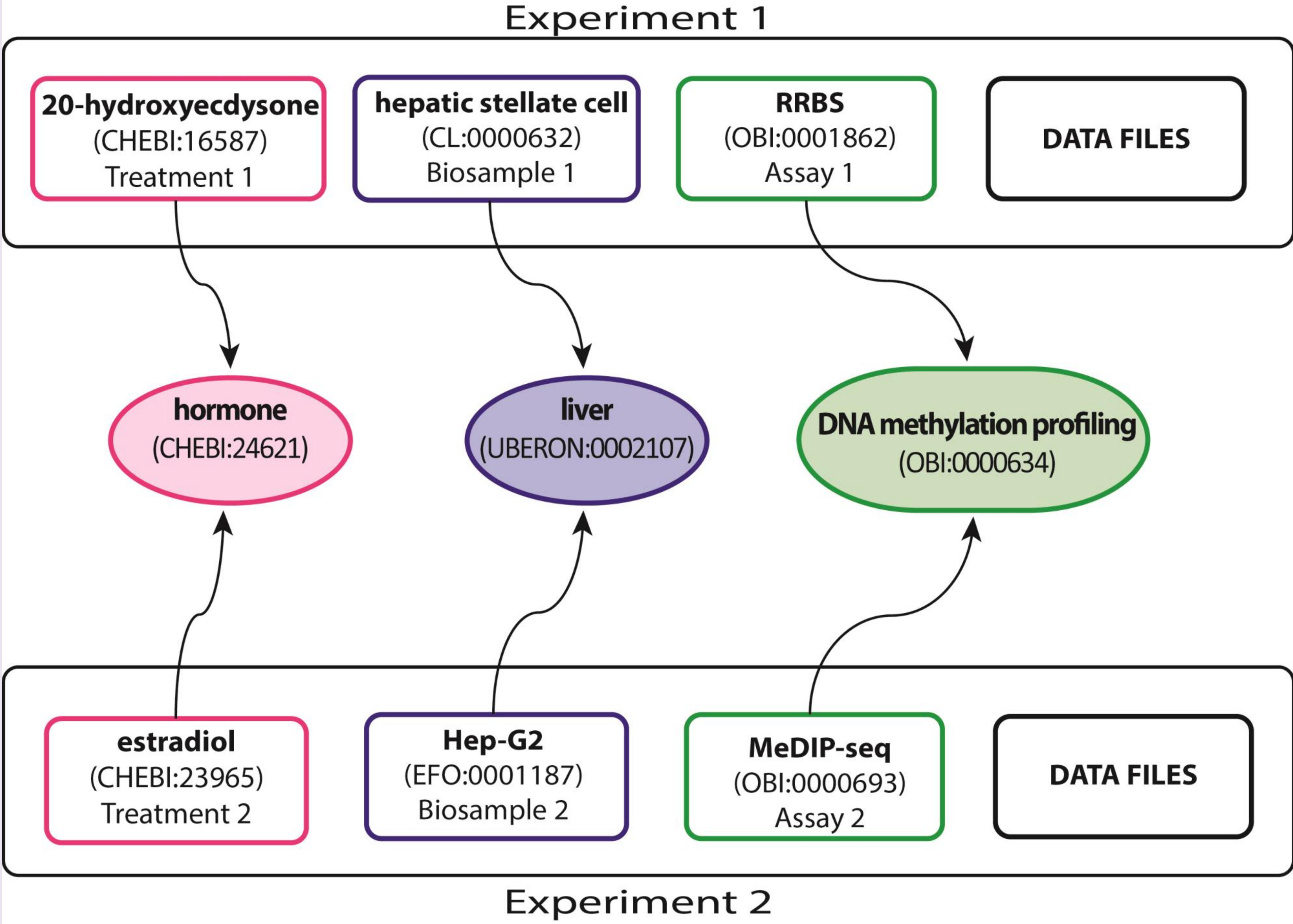
F ≠ f ≠ Female ≠ female

- Biological concepts are defined to provide scope

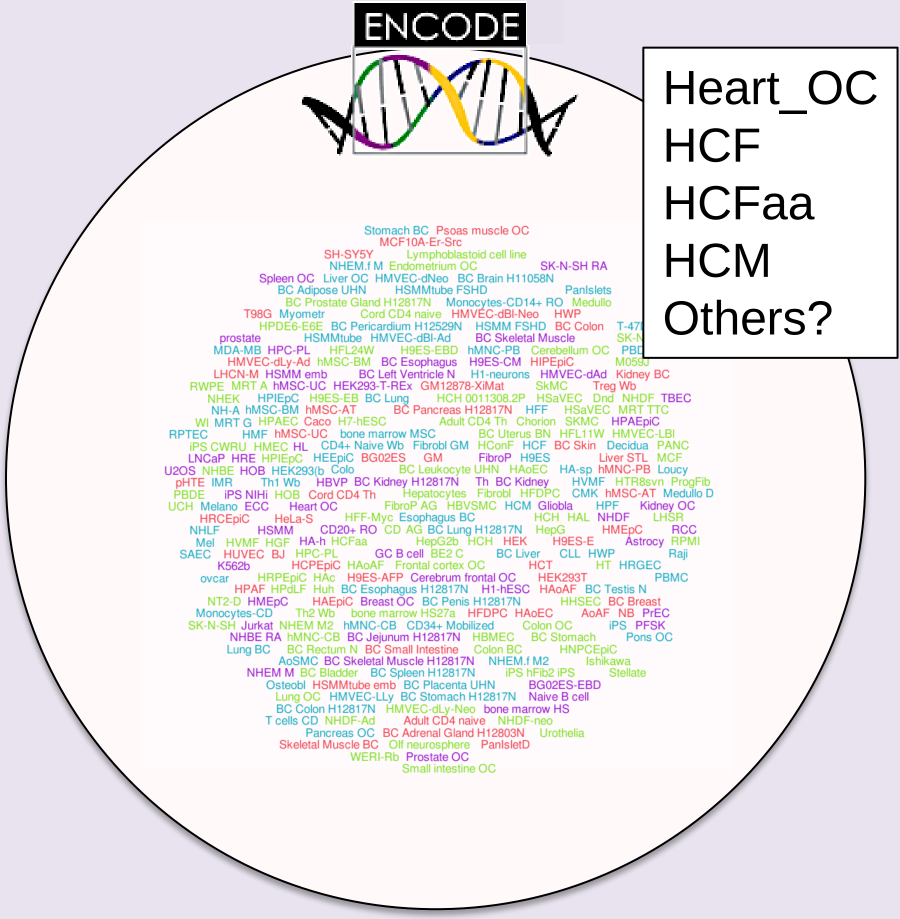
Mitochondria: A semiautonomous, self replicating organelle that occurs in varying numbers, shapes, and sizes in the cytoplasm of virtually all eukaryotic cells. It is notably the site of tissue respiration.

- Relationships between terms can be computed to provide additional annotation details for grouping, searching, or analysis

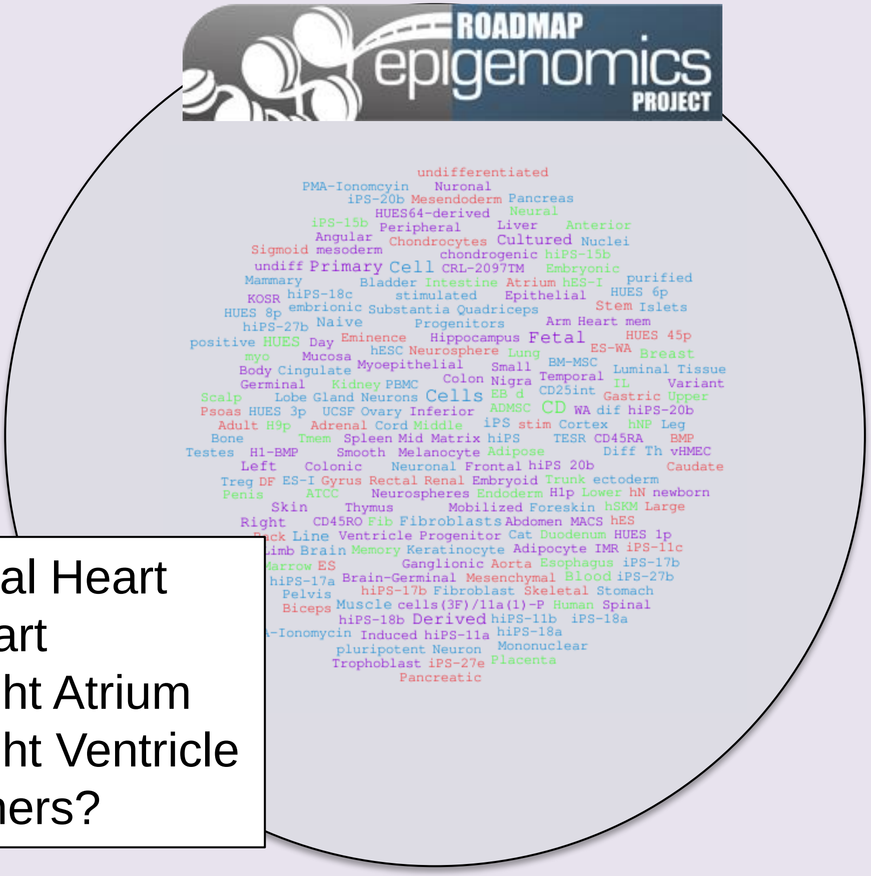

is_a
part_of
derives_from
Explicit relationships
Inferred relationships
Uberon
CL
EFO
endocrine system (UBERON:0000949)
digestive system (UBERON:0001007)
cell line (EFO:0000322)
liver (UBERON:0002107)
small intestine (UBERON:0002108)
cancer cell line (EFO:0001639)
hepatoma cell line (EFO:0005216)
hepatocyte (CL:0000182)
hepatic stellate cell (CL:0000632)
enterocyte of epithelium of small intestine (CL:0000632)
Hep-G2 (EFO:0001187)


Experiment 1
20-hydroxyecdysone
(CHEBI:16587)
Treatment 1
hepatic stellate cell
(CL:0000632)
Biosample 1
RRBS
(OBI:0001862)
Assay 1
DATA FILES
hormone
(CHEBI:24621)
liver
(UBERON:0002107)
DNA methylation profiling
(OBI:0000634)
estradiol
(CHEBI:23965)
Treatment 2
Hep-G2
(EFO:0001187)
Biosample 2
MeDIP-seq
(OBI:0000693)
Assay 2
DATA FILES
Experiment 2

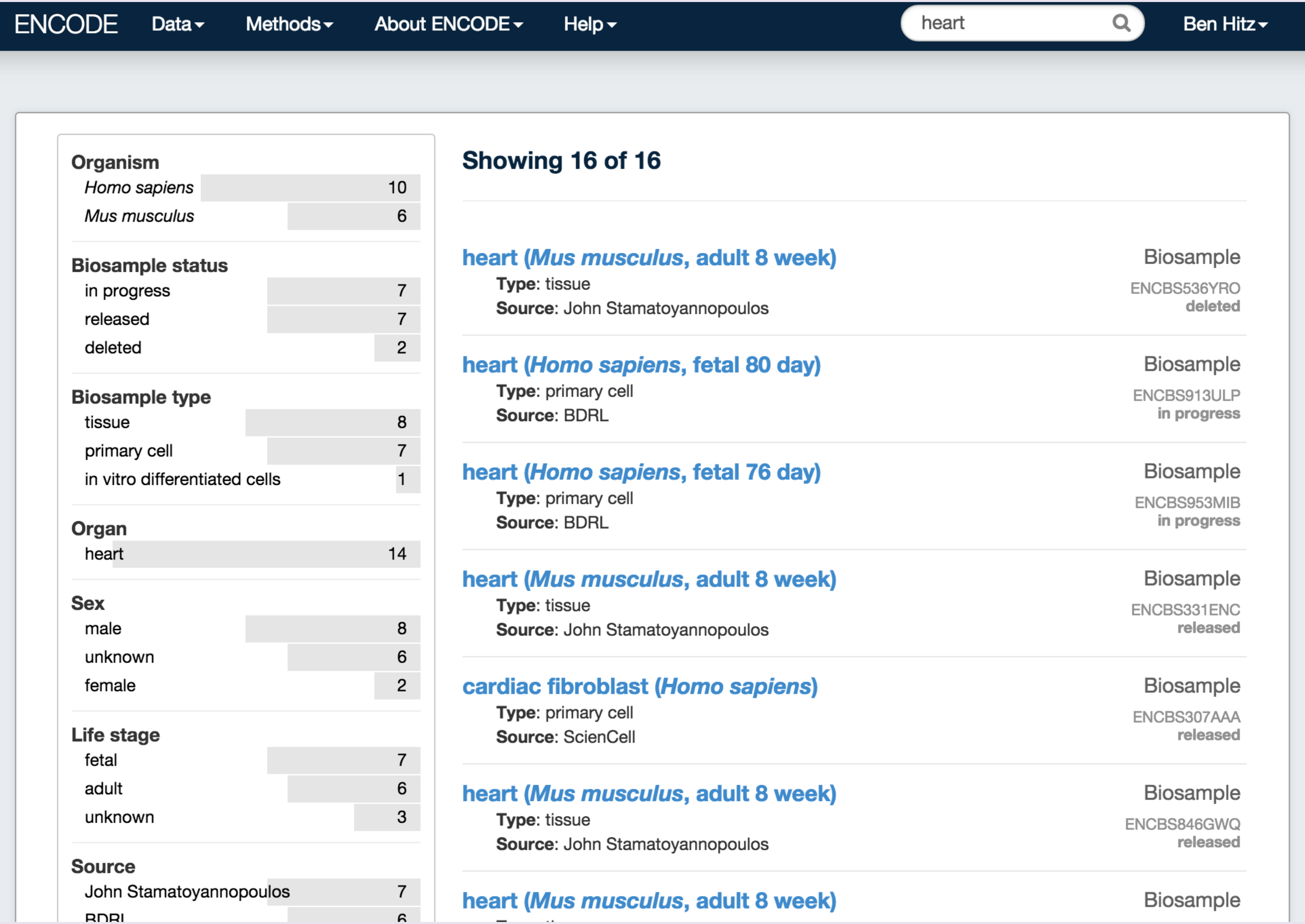
Challenge: Find all heart-related tissues?


ENCODE
Stomach BC Psoas muscle OC
MCF10A-Er-Src
SH-SY5Y Lymphoblastoid cell line
NHEM.f M Endometrium OC SK-N-SH RA
Spleen OC Liver OC HMVEC-dNeo BC Brain H11058N
BC Adipose UHN HSMMtube FSHD PanIslets
BC Prostate Gland H12817N Monocytes-CD14+ RO Medullo
T98G Myometr Cord CD4 naive HMVEC-dBl-Neo HWP
HPDE6-E6E BC Pericardium H12529N HSMM FSHD BC Colon T-47I
prostate HSMMtube HMVEC-dBl-Ad BC Skeletal Muscle SK-N
MDA-MB HPC-PL HFL24W H9ES-EBD hMNC-PB Cerebellum OC PBD
HMVEC-dLy-Ad hMSC-BM BC Esophagus H9ES-CM HIPEpiC M059J
LHCN-M HSMM emb BC Left Ventricle N H1-neurons HMVEC-dAd Kidney BC
RWPE MRT A hMSC-UC HEK293-T-REx GM12878-XiMat SkMC Treg Wb
NHEK HPIEpC H9ES-EB BC Lung HCH 0011308.2P HSaVEC Dnd NHDF TBEC
NH-A hMSC-BM hMSC-AT BC Pancreas H12817N HFF HSaVEC MRT TTC
WI MRT G HPAEC Caco H7-hESC Adult CD4 Th Chorion SKMC HPAEpiC
RPTEC HMF hMSC-UC bone marrow MSC BC Uterus BN HFL11W HMVEC-LBl
iPS CWRU HMEC HL CD4+ Naive Wb Fibrobl GM HConF HCF BC Skin Decidua PANC
LNCaP HRE HPIEpC HEEpiC BG02ES GM FibroP H9ES Liver STL MCF
U2OS NHBE HOB HEK293(b Colo BC Leukocyte UHN HAoEC HA-sp hMNC-PB Loucy
pHTE IMR Th1 Wb HBVP BC Kidney H12817N Th BC Kidney HVMF HTR8svn ProgFib
PBDE iPS NIHi HOB Cord CD4 Th Hepatocytes Fibrobl HFDPC CMK hMSC-AT Medullo D
UCH Melano ECC Heart OC FibroP AG HBVSMC HCM Gliobla HPF Kidney OC
HRCEpiC HeLa-S HFF-Myc Esophagus BC HCH HAL NHDF LHSR
NHLF HSMM CD20+ RO CD AG BC Lung H12817N HepG HMEpC RCC
Mel HVMF HGF HA-h HCFaa HepG2b HCH HEK H9ES-E Astrocy RPMI
SAEC HUVEC BJ HPC-PL GC B cell BE2 C BC Liver CLL HWP Raji
K562b HCPEpiC HAoAF Frontal cortex OC HCT HT HRGEC
ovcar HRPEpiC HAc H9ES-AFP Cerebrum frontal OC HEK293T PBMC
HPAF HPdLF Huh BC Esophagus H12817N H1-hESC HAoAF BC Testis N
NT2-D HMEpC HAEpiC Breast OC BC Penis H12817N HHSEC BC Breast
Monocytes-CD Th2 Wb bone marrow HS27a HFDPC HAoEC AoAF NB PrEC
SK-N-SH Jurkat NHEM M2 hMNC-CB CD34+ Mobilized Colon OC iPS PFSK
NHBE RA hMNC-CB BC Jejunum H12817N HBMEC BC Stomach Pons OC
Lung BC BC Rectum N BC Small Intestine Colon BC HNPCEpiC
AoSMC BC Skeletal Muscle H12817N NHEM.f M2 Ishikawa
NHEM M BC Bladder BC Spleen H12817N iPS hFib2 iPS Stellate
Osteobl HSMMtube emb BC Placenta UHN BG02ES-EBD
Lung OC HMVEC-LLy BC Stomach H12817N Naive B cell
BC Colon H12817N HMVEC-dLy-Neo bone marrow HS
T cells CD NHDF-Ad Adult CD4 naive NHDF-neo
Pancreas OC BC Adrenal Gland H12803N Urothelia
Skeletal Muscle BC Olf neurosphere PanIsletD
WERI-Rb Prostate OC
Small intestine OC

Heart_OC
HCF
HCFaa
HCM
Others?


ROADMAP
epigenomics
PROJECT
undifferentiated
PMA-Ionomcyin Nuronal
iPS-20b Mesendoderm Pancreas
HUES64-derived Neural
iPS-15b Peripheral Liver Anterior
Angular Chondrocytes Cultured Nuclei
Sigmoid mesoderm chondrogenic hiPS-15b
undiff Primary Cell CRL-2097TM Embryonic
Mammary Bladder Intestine Atrium hES-I purified
KOSR hiPS-18c stimulated Epithelial HUES 6p
HUES 8p embrionic Substantia Quadriceps Stem Islets
hiPS-27b Naive Progenitors Arm Heart mem
positive HUES Day Eminence Hippocampus Fetal HUES 45p
myo Mucosa hESC Neurosphere Lung ES-WA Breast
Body Cingulate Myoepithelial Small BM-MSC Luminal Tissue
Germinal Kidney PBMC Colon Nigra Temporal IL Variant
Scalp Lobe Gland Neurons Cells EB d CD25int Gastric Upper
Psoas HUES 3p UCSF Ovary Inferior ADMSC CD WA dif hiPS-20b
Adult H9p Adrenal Cord Middle iPS stim Cortex hNP Leg
Bone Tmem Spleen Mid Matrix hiPS TESR CD45RA BMP
Testes H1-BMP Smooth Melanocyte Adipose Diff Th vHMEC
Left Colonic Neuronal Frontal hiPS 20b Caudate
Treg DF ES-I Gyrus Rectal Renal Embryoid Trunk ectoderm
Penis ATCC Neurospheres Endoderm H1p Lower hN newborn
Skin Thymus Mobilized Foreskin hSKM Large
Right CD45RO Fib Fibroblasts Abdomen MACS hES
ack Line Ventricle Progenitor Cat Duodenum HUES 1p
Limb Brain Memory Keratinocyte Adipocyte IMR iPS-11c
Marrow ES Ganglionic Aorta Esophagus iPS-17b
hiPS-17a Brain-Germinal Mesenchymal Blood iPS-27b
Pelvis hiPS-17b Fibroblast Skeletal Stomach
Biceps Muscle cells(3F)/11a(1)-P Human Spinal
hiPS-18b Derived hiPS-11b iPS-18a
A-Ionomycin Induced hiPS-11a hiPS-18a
pluripotent Neuron Mononuclear
Trophoblast iPS-27e Placenta
Pancreatic

Fetal Heart
Heart
Right Atrium
Right Ventricle
Others?



Organism

Homo sapiens	10
Mus musculus	6

Biosample status

in progress	7
released	7
deleted	2

Biosample type

tissue	8
primary cell	7
in vitro differentiated cells	1

Organ

heart	14

Sex

male	8
unknown	6
female	2

Life stage

fetal	7
adult	6
unknown	3

Source

John Stamatoyannopoulos	7
BDRL	6

Showing 16 of 16

heart (*Mus musculus*, adult 8 week) — Biosample
Type: tissue
Source: John Stamatoyannopoulos
ENCBS536YRO
deleted

heart (*Homo sapiens*, fetal 80 day) — Biosample
Type: primary cell
Source: BDRL
ENCBS913ULP
in progress

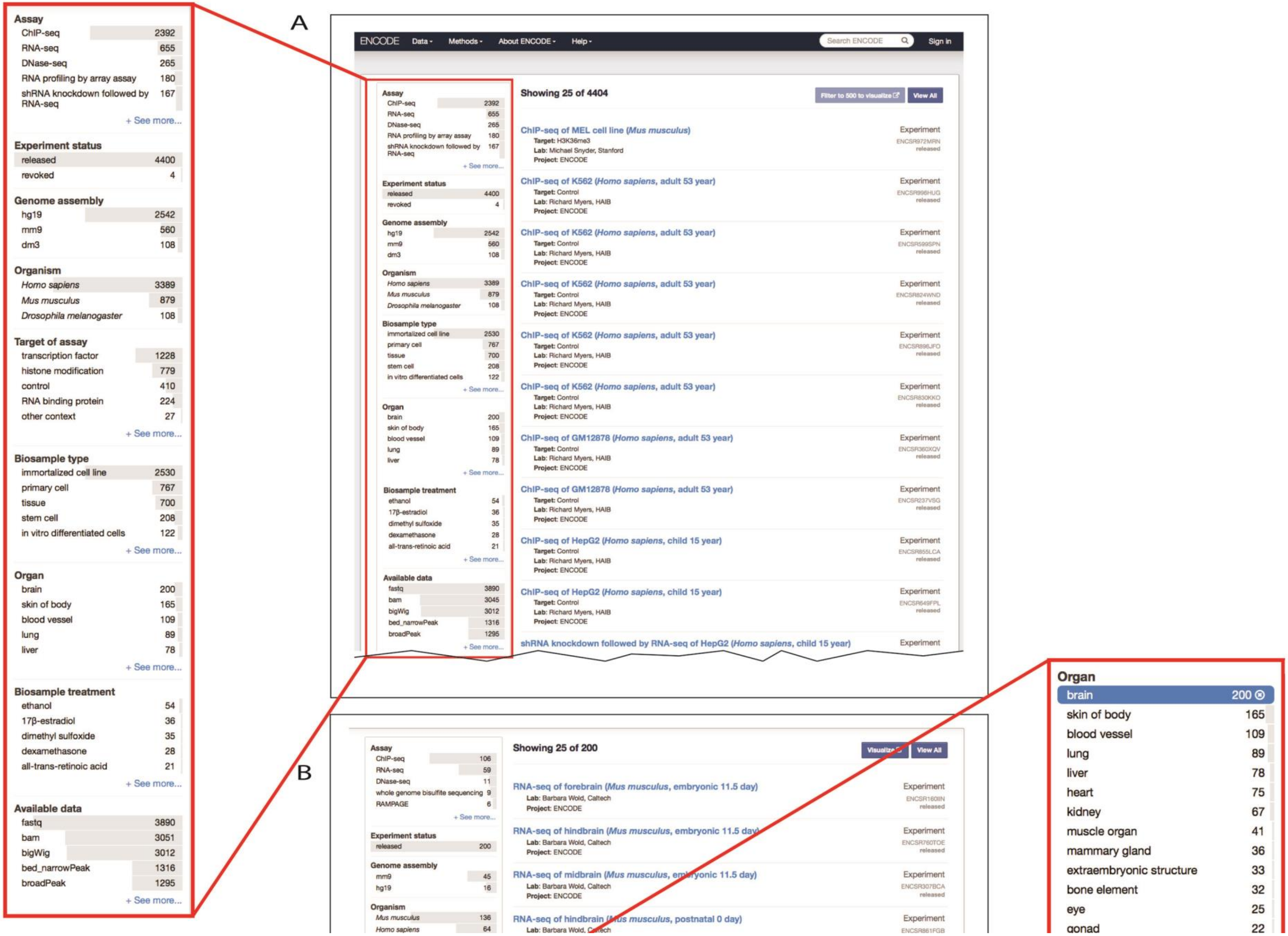
heart (*Homo sapiens*, fetal 76 day) — Biosample
Type: primary cell
Source: BDRL
ENCBS953MIB
in progress

heart (*Mus musculus*, adult 8 week) — Biosample
Type: tissue
Source: John Stamatoyannopoulos
ENCBS331ENC
released

cardiac fibroblast (*Homo sapiens*) — Biosample
Type: primary cell
Source: ScienCell
ENCBS307AAA
released

heart (*Mus musculus*, adult 8 week) — Biosample
Type: tissue
Source: John Stamatoyannopoulos
ENCBS846GWQ
released

heart (*Mus musculus*, adult 8 week) — Biosample



Assay	
ChIP-seq	2392
RNA-seq	655
DNase-seq	265
RNA profiling by array assay	180
shRNA knockdown followed by RNA-seq	167
+ See more...	

Experiment status	
released	4400
revoked	4

Genome assembly	
hg19	2542
mm9	560
dm3	108

Organism	
Homo sapiens	3389
Mus musculus	879
Drosophila melanogaster	108

Target of assay	
transcription factor	1228
histone modification	779
control	410
RNA binding protein	224
other context	27
+ See more...	

Biosample type	
immortalized cell line	2530
primary cell	767
tissue	700
stem cell	208
in vitro differentiated cells	122
+ See more...	

Organ	
brain	200
skin of body	165
blood vessel	109
lung	89
liver	78
+ See more...	

Biosample treatment	
ethanol	54
17β-estradiol	36
dimethyl sulfoxide	35
dexamethasone	28
all-trans-retinoic acid	21
+ See more...	

Available data	
fastq	3890
bam	3051
bigWig	3012
bed_narrowPeak	1316
broadPeak	1295
+ See more...	

ENCODE Data ▾ Methods ▾ About ENCODE ▾ Help ▾ Search ENCODE Sign in

Assay	
ChIP-seq	2392
RNA-seq	655
DNase-seq	265
RNA profiling by array assay	180
shRNA knockdown followed by RNA-seq	167
+ See more...	

Experiment status	
released	4400
revoked	4

Genome assembly	
hg19	2542
mm9	560
dm3	108

Organism	
Homo sapiens	3389
Mus musculus	879
Drosophila melanogaster	108

Biosample type	
immortalized cell line	2530
primary cell	767
tissue	700
stem cell	208
in vitro differentiated cells	122
+ See more...	

Organ	
brain	200
skin of body	165
blood vessel	109
lung	89
liver	78
+ See more...	

Biosample treatment	
ethanol	54
17β-estradiol	36
dimethyl sulfoxide	35
dexamethasone	28
all-trans-retinoic acid	21
+ See more...	

Available data	
fastq	3890
bam	3045
bigWig	3012
bed_narrowPeak	1316
broadPeak	1295
+ See more...	

Showing 25 of 4404

Filter to 500 to visualize View All

ChIP-seq of MEL cell line (*Mus musculus*)
Target: H3K36me3
Lab: Michael Snyder, Stanford
Project: ENCODE
Experiment
ENCSR972MRN
released

ChIP-seq of K562 (*Homo sapiens*, adult 53 year)
Target: Control
Lab: Richard Myers, HAIB
Project: ENCODE
Experiment
ENCSR996HUG
released

ChIP-seq of K562 (*Homo sapiens*, adult 53 year)
Target: Control
Lab: Richard Myers, HAIB
Project: ENCODE
Experiment
ENCSR599SPN
released

ChIP-seq of K562 (*Homo sapiens*, adult 53 year)
Target: Control
Lab: Richard Myers, HAIB
Project: ENCODE
Experiment
ENCSR824WND
released

ChIP-seq of K562 (*Homo sapiens*, adult 53 year)
Target: Control
Lab: Richard Myers, HAIB
Project: ENCODE
Experiment
ENCSR896JFO
released

ChIP-seq of K562 (*Homo sapiens*, adult 53 year)
Target: Control
Lab: Richard Myers, HAIB
Project: ENCODE
Experiment
ENCSR830KKO
released

ChIP-seq of GM12878 (*Homo sapiens*, adult 53 year)
Target: Control
Lab: Richard Myers, HAIB
Project: ENCODE
Experiment
ENCSR360XQV
released

ChIP-seq of GM12878 (*Homo sapiens*, adult 53 year)
Target: Control
Lab: Richard Myers, HAIB
Project: ENCODE
Experiment
ENCSR237VSG
released

ChIP-seq of HepG2 (*Homo sapiens*, child 15 year)
Target: Control
Lab: Richard Myers, HAIB
Project: ENCODE
Experiment
ENCSR855LCA
released

ChIP-seq of HepG2 (*Homo sapiens*, child 15 year)
Target: Control
Lab: Richard Myers, HAIB
Project: ENCODE
Experiment
ENCSR649FPL
released

shRNA knockdown followed by RNA-seq of HepG2 (*Homo sapiens*, child 15 year)
Experiment

Assay	
ChIP-seq	106
RNA-seq	59
DNase-seq	11
whole genome bisulfite sequencing	9
RAMPAGE	6
+ See more...	

Experiment status	
released	200

Genome assembly	
mm9	45
hg19	16

Organism	
Mus musculus	136
Homo sapiens	64

Showing 25 of 200

Visualize View All

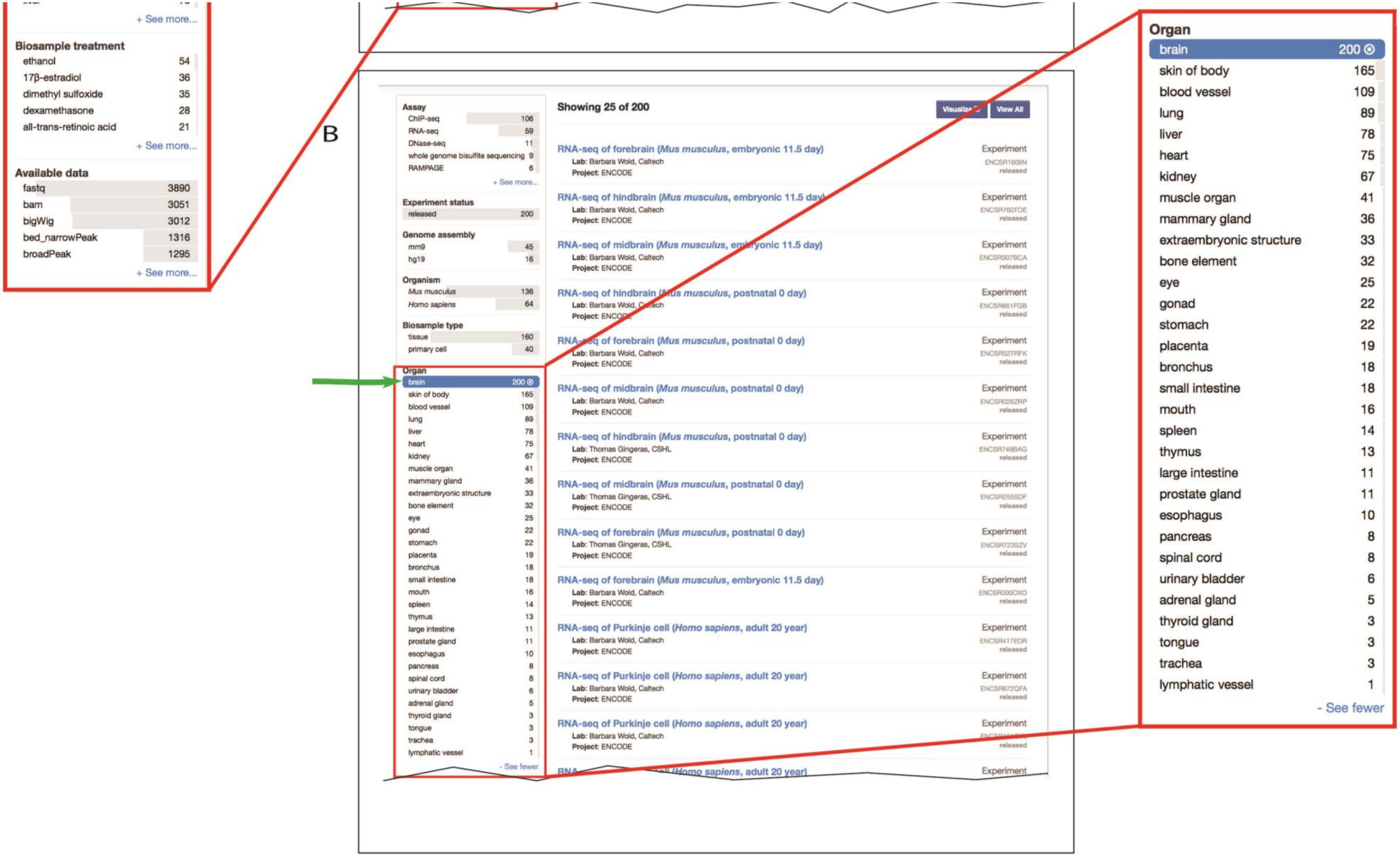
RNA-seq of forebrain (*Mus musculus*, embryonic 11.5 day)
Lab: Barbara Wold, Caltech
Project: ENCODE
Experiment
ENCSR160IIN
released

RNA-seq of hindbrain (*Mus musculus*, embryonic 11.5 day)
Lab: Barbara Wold, Caltech
Project: ENCODE
Experiment
ENCSR760TOE
released

RNA-seq of midbrain (*Mus musculus*, embryonic 11.5 day)
Lab: Barbara Wold, Caltech
Project: ENCODE
Experiment
ENCSR307BCA
released

RNA-seq of hindbrain (*Mus musculus*, postnatal 0 day)
Lab: Barbara Wold, Caltech
Experiment
ENCSR861FGB

Organ	
brain	200 ⊗
skin of body	165
blood vessel	109
lung	89
liver	78
heart	75
kidney	67
muscle organ	41
mammary gland	36
extraembryonic structure	33
bone element	32
eye	25
gonad	22



Biosample treatment

ethanol	54
17β-estradiol	36
dimethyl sulfoxide	35
dexamethasone	28
all-trans-retinoic acid	21

+ See more...

Available data

fastq	3890
bam	3051
bigWig	3012
bed_narrowPeak	1316
broadPeak	1295

+ See more...

Assay

ChIP-seq	106
RNA-seq	59
DNase-seq	11
whole genome bisulfite sequencing	9
RAMPAGE	6

+ See more...

Experiment status

released	200

Genome assembly

mm9	45
hg19	16

Organism

Mus musculus	136
Homo sapiens	64

Biosample type

tissue	160
primary cell	40

Organ

brain	200 ⊗
skin of body	165
blood vessel	109
lung	89
liver	78
heart	75
kidney	67
muscle organ	41
mammary gland	36
extraembryonic structure	33
bone element	32
eye	25
gonad	22
stomach	22
placenta	19
bronchus	18
small intestine	18
mouth	16
spleen	14
thymus	13
large intestine	11
prostate gland	11
esophagus	10
pancreas	8
spinal cord	8
urinary bladder	6
adrenal gland	5
thyroid gland	3
tongue	3
trachea	3
lymphatic vessel	1

- See fewer

Showing 25 of 200

RNA-seq of forebrain (*Mus musculus*, embryonic 11.5 day) — Experiment ENCSR160IIN released
Lab: Barbara Wold, Caltech
Project: ENCODE

RNA-seq of hindbrain (*Mus musculus*, embryonic 11.5 day) — Experiment ENCSR760TOE released
Lab: Barbara Wold, Caltech
Project: ENCODE

RNA-seq of midbrain (*Mus musculus*, embryonic 11.5 day) — Experiment ENCSR307BCA released
Lab: Barbara Wold, Caltech
Project: ENCODE

RNA-seq of hindbrain (*Mus musculus*, postnatal 0 day) — Experiment ENCSR861FGB released
Lab: Barbara Wold, Caltech
Project: ENCODE

RNA-seq of forebrain (*Mus musculus*, postnatal 0 day) — Experiment ENCSR527RFK released
Lab: Barbara Wold, Caltech
Project: ENCODE

RNA-seq of midbrain (*Mus musculus*, postnatal 0 day) — Experiment ENCSR026ZRP released
Lab: Barbara Wold, Caltech
Project: ENCODE

RNA-seq of hindbrain (*Mus musculus*, postnatal 0 day) — Experiment ENCSR749BAG released
Lab: Thomas Gingeras, CSHL
Project: ENCODE

RNA-seq of midbrain (*Mus musculus*, postnatal 0 day) — Experiment ENCSR255SDF released
Lab: Thomas Gingeras, CSHL
Project: ENCODE

RNA-seq of forebrain (*Mus musculus*, postnatal 0 day) — Experiment ENCSR723SZV released
Lab: Thomas Gingeras, CSHL
Project: ENCODE

RNA-seq of forebrain (*Mus musculus*, embryonic 11.5 day) — Experiment ENCSR000OXO released
Lab: Barbara Wold, Caltech
Project: ENCODE

RNA-seq of Purkinje cell (*Homo sapiens*, adult 20 year) — Experiment ENCSR417EDR released
Lab: Barbara Wold, Caltech
Project: ENCODE

RNA-seq of Purkinje cell (*Homo sapiens*, adult 20 year) — Experiment ENCSR672QFA released
Lab: Barbara Wold, Caltech
Project: ENCODE

RNA-seq of Purkinje cell (*Homo sapiens*, adult 20 year) — Experiment released
Lab: Barbara Wold, Caltech
Project: ENCODE

RNA (*Homo sapiens*, adult 20 year) — Experiment

Organ

brain	200 ⊗
skin of body	165
blood vessel	109
lung	89
liver	78
heart	75
kidney	67
muscle organ	41
mammary gland	36
extraembryonic structure	33
bone element	32
eye	25
gonad	22
stomach	22
placenta	19
bronchus	18
small intestine	18
mouth	16
spleen	14
thymus	13
large intestine	11
prostate gland	11
esophagus	10
pancreas	8
spinal cord	8
urinary bladder	6
adrenal gland	5
thyroid gland	3
tongue	3
trachea	3
lymphatic vessel	1

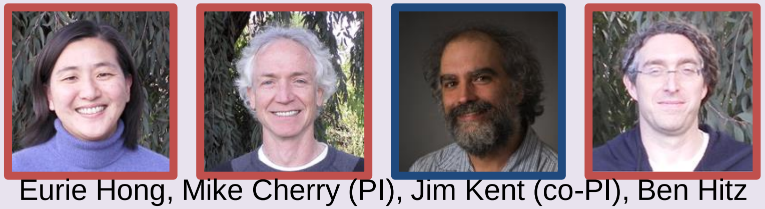
- See fewer

ENCODE DCC

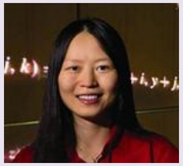
https://www.encodeproject.org/

STANFORD SCHOOL OF MEDICINE

Baskin Engineering UC SANTA CRUZ



Eurie Hong, Mike Cherry (PI), Jim Kent (co-PI), Ben Hitz



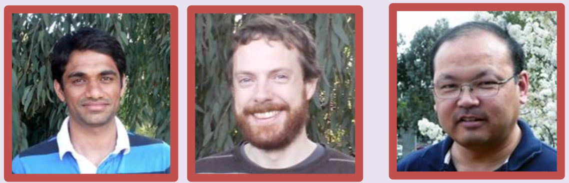
Zhiping Weng, ENCODE DAC

University of Massachusetts UMASS. Medical School

Data Wranglers

Esther Chan, Jean Davidson, Venkat Malladi, Cricket Sloan, J. Seth Strattan

Software Engineers



Nikhil Podduturi, Laurence Rowe, Forrest Tanaka

Funding Source: NHGRI

QA, administration, biocuration

Brian Lee, Stuart Miyasato, Matt Simison, Zhenhua Wang, Marcus Ho

Malladi et al. ; Database, 2015, 1–11 doi: 10.1093/database/bav010

@encodedcc encode-help@lists.stanford.edu https://github.com/ENCODE-DCC/

Using ontologies for metadata annotation

1. **Uber Anatomy ontology** (UBERON; http://uberon.org/)
    - tissues: heart, blood, brain

2. **Cell Ontology** (CL; http://cellontology.org/)
    - primary cell types: hepatocyte, cardiomyocyte

3. **Experimental Factor Ontology** (EFO; http://www.ebi.ac.uk/efo/)
    - immortalized cell lines: K562, HepG2, MCF-7

4. **Ontology for Biomedical Investigations** (OBI; http://obi-ontology.org/page/Main_Page)
    - experimental assays: RNA-seq, CLIP-seq, ChIP-seq, etc

5. **Chemical Entities of Biological Interest** (ChEBI; http://www.ebi.ac.uk/chebi/)
    - chemical treatments: estradiol, ethanol, etc

6. **Sequence Ontology** (SO; http://www.sequenceontology.org/)
    - nucleic acid being sequenced: microRNA, poly-A+ mRNA, etc

7. **Gene Ontology** (GO; http://www.geneontology.org/)
    - group gene products that are targets of ChIP-seq or RNAi experiments