

FIND AND VISUALIZE ENCODE DATA

WWW.ENCODEPROJECT.ORG

J. Seth Strattan, PhD

ENCODE Data Coordinating Center (DCC)

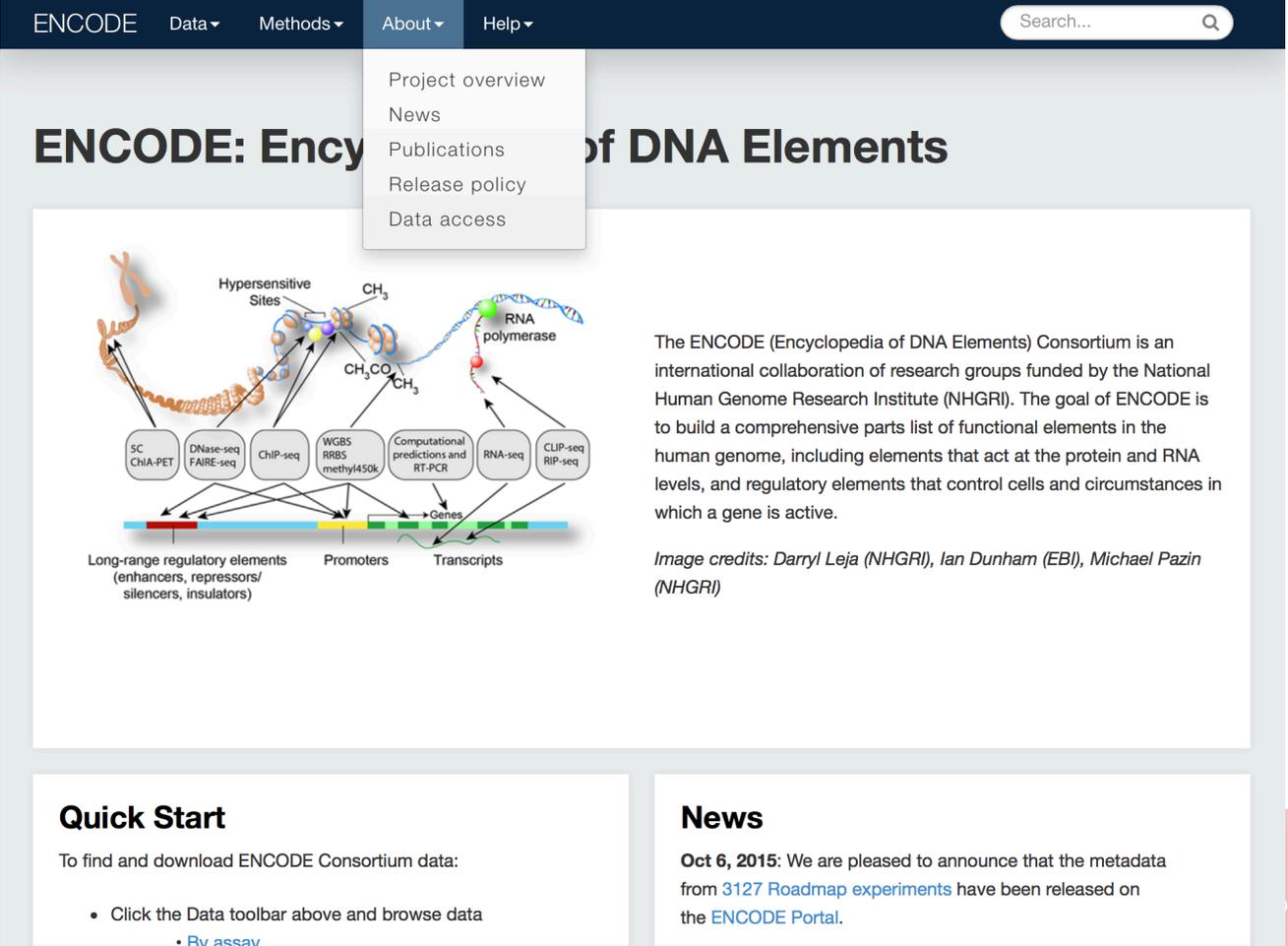
Keystone Chromatin and Epigenetics

March, 2016



ENCODE Portal www.encodeproject.org

- The ENCODE Portal is the **canonical source** for ENCODE metadata and data.
- The Portal also documents **ENCODE standards** like antibody standards, data release.
- The Portal links to **documentation and tutorials**.
- Use the Portal to **browse, search, and visualize** what ENCODE has done.
- Understand ENCODE **data analyses**, and how you can replicate them.
- REST API for programmatic **search and download** of ENCODE metadata and data.

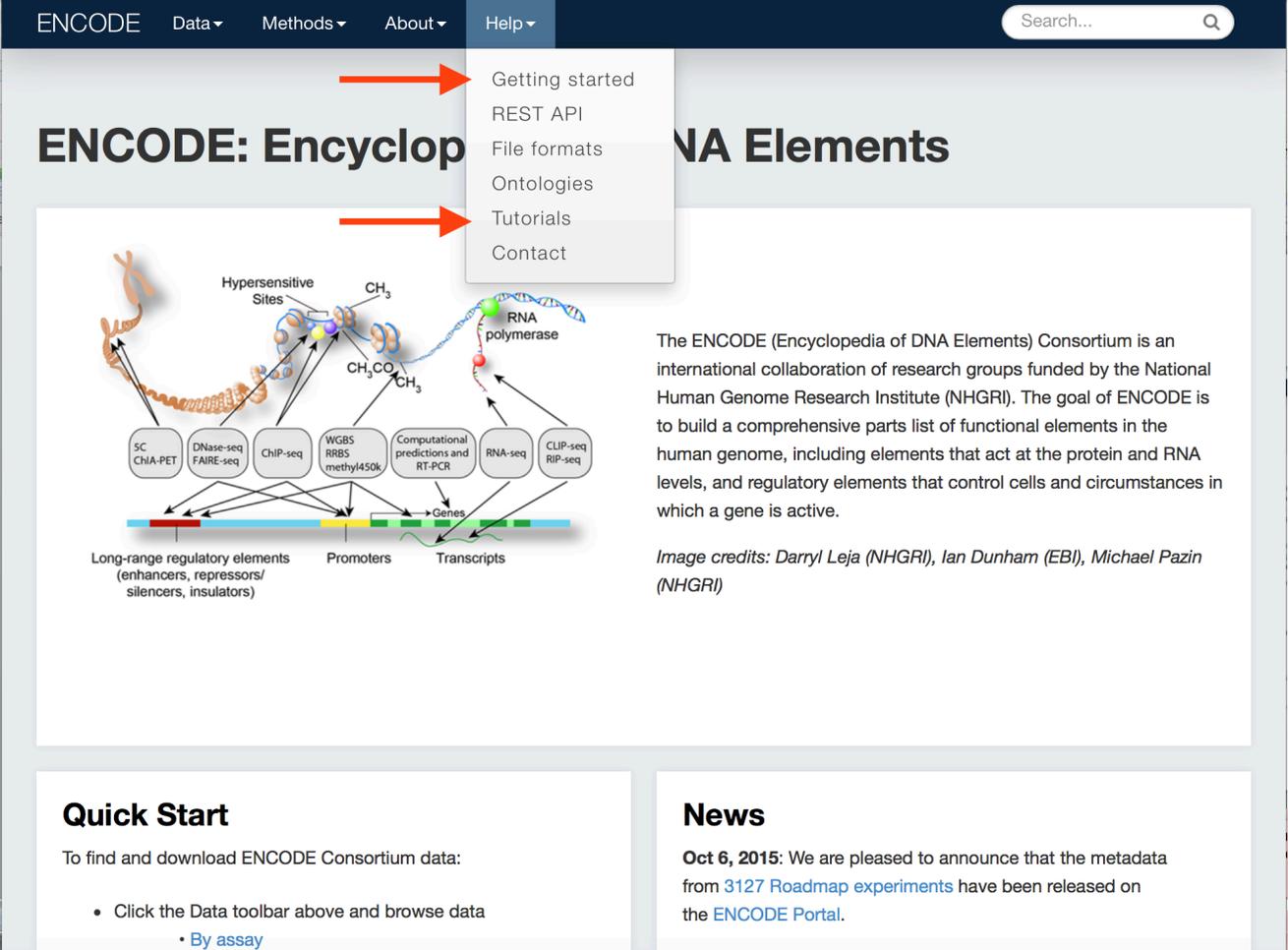


The screenshot shows the ENCODE Portal website. At the top, there is a navigation bar with 'ENCODE' and dropdown menus for 'Data', 'Methods', 'About', and 'Help'. A search bar is located on the right. Below the navigation bar, the main heading reads 'ENCODE: Encyclopedia of DNA Elements'. A diagram illustrates the structure of DNA with various elements and assays. The diagram shows a DNA strand with 'Long-range regulatory elements (enhancers, repressors/silencers, insulators)', 'Promoters', and 'Transcripts'. Above the DNA, 'Hypersensitive Sites' are marked with blue circles, and 'CH₃' and 'CH₃CO' groups are shown. 'RNA polymerase' is depicted as a green structure. Below the DNA, several assay types are listed in boxes: 'SC ChIA-PET', 'DNase-seq FAIRE-seq', 'ChIP-seq', 'WGBS RRBS methyl450k', 'Computational predictions and RT-PCR', 'RNA-seq', and 'CLIP-seq RIP-seq'. Arrows indicate the relationship between these assays and the DNA elements. To the right of the diagram, a text block explains the ENCODE Consortium's goal: '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.' Below this text, it says 'Image credits: Darryl Leja (NHGRI), Ian Dunham (EBI), Michael Pazin (NHGRI)'. At the bottom of the page, there are two sections: 'Quick Start' and 'News'. The 'Quick Start' section provides instructions on how to find and download ENCODE Consortium data, including a link to 'By assay'. The 'News' section features a date 'Oct 6, 2015' and a short announcement about the release of metadata from 3127 Roadmap experiments.

www.encodeproject.org ... Getting Started

Help Menu Documents the Portal

- *Getting Started*: Portal organization
 - <https://www.encodeproject.org/help/getting-started>
- *Tutorials*:
 - <https://www.encodeproject.org/tutorials>
 - Links to slides from today's workshop
- *Contact*:
 - encode-help@lists.stanford.edu
 - [@encodedcc](#)



The screenshot shows the ENCODE website interface. At the top, there is a navigation bar with 'ENCODE' and dropdown menus for 'Data', 'Methods', 'About', and 'Help'. A search bar is located on the right. The 'Help' menu is open, showing options: 'Getting started', 'REST API', 'File formats', 'Ontologies', 'Tutorials', and 'Contact'. Two red arrows point to 'Getting started' and 'Tutorials'. Below the navigation bar, the main heading reads 'ENCODE: Encyclopedia of DNA Elements'. A central diagram illustrates the genome with various elements: 'Long-range regulatory elements (enhancers, repressors/silencers, insulators)', 'Promoters', and 'Transcripts'. Above the diagram, several assays are listed: 'SC ChIA-PET', 'DNase-seq FAIRE-seq', 'ChIP-seq', 'WGBS RRBS methyl450k', 'Computational predictions and RT-PCR', 'RNA-seq', and 'CLIP-seq RIP-seq'. The diagram also shows 'Hypersensitive Sites', 'CH₃', 'CH₃CO', and 'RNA polymerase'. To the right of the diagram, a text block describes the ENCODE Consortium's goal. Below the diagram, there are sections for 'Quick Start' and 'News'.

ENCODE Data Methods About Help Search...

ENCODE: Encyclopedia of DNA Elements

Getting started
REST API
File formats
Ontologies
Tutorials
Contact

Hypersensitive Sites
CH₃
CH₃CO
RNA polymerase

SC ChIA-PET
DNase-seq FAIRE-seq
ChIP-seq
WGBS RRBS methyl450k
Computational predictions and RT-PCR
RNA-seq
CLIP-seq RIP-seq

Long-range regulatory elements (enhancers, repressors/silencers, insulators)
Promoters
Transcripts
Genes

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)

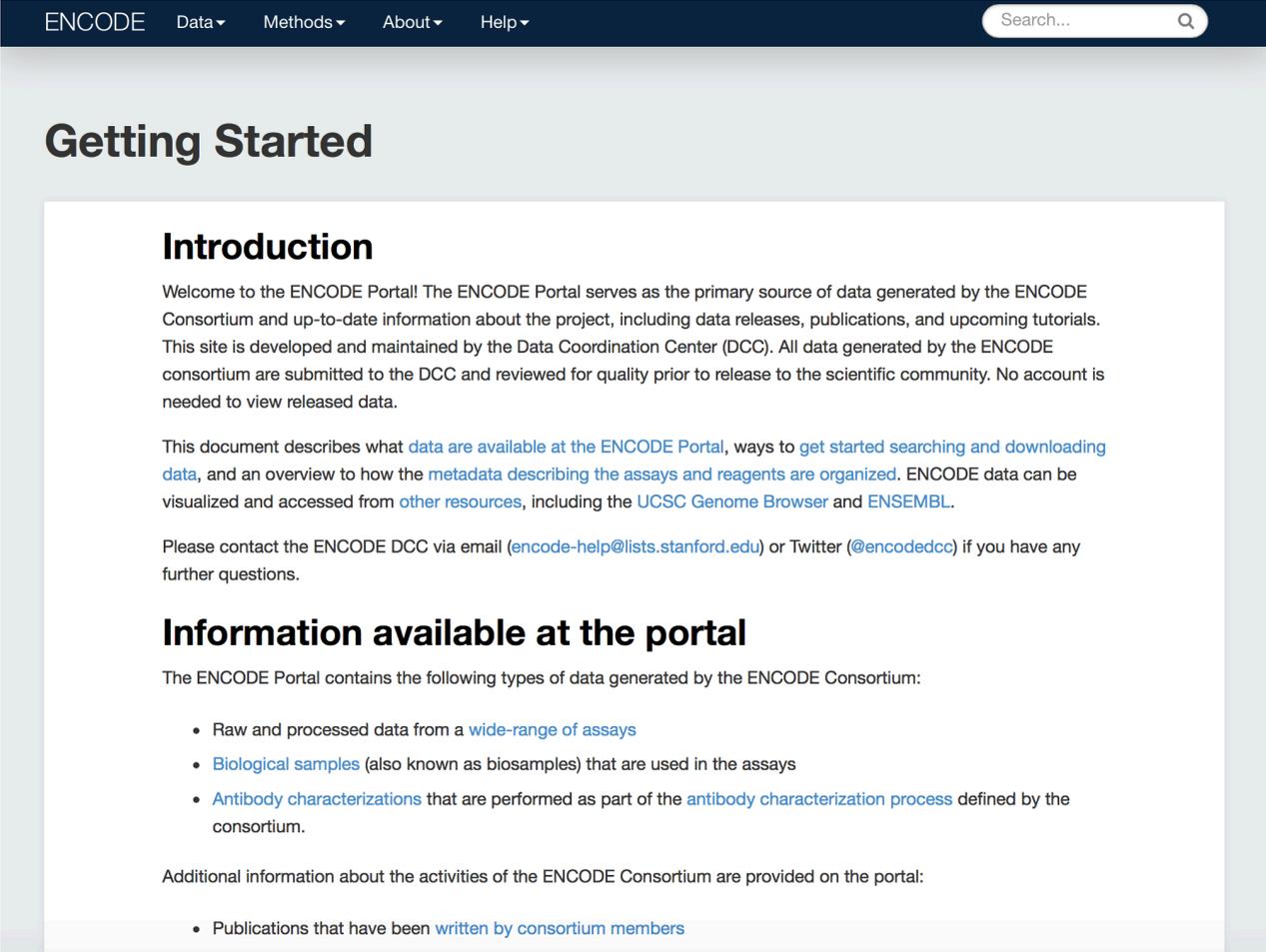
Quick Start
To find and download ENCODE Consortium data:

- Click the Data toolbar above and browse data
 - By assay

News
Oct 6, 2015: We are pleased to announce that the metadata from 3127 Roadmap experiments have been released on the ENCODE Portal.

www.encodeproject.org ... Getting Started

- <https://www.encodeproject.org/help/getting-started>
 - How to browse and facet.
 - How to access one file at a time.
 - Bulk download of files.
 - ENCODE data model and data organization.



The screenshot shows the ENCODE Portal's 'Getting Started' page. The navigation bar includes 'ENCODE', 'Data', 'Methods', 'About', and 'Help', along with a search bar. The main content area is titled 'Getting Started' and contains an 'Introduction' section. The introduction welcomes users to the ENCODE Portal, stating it is the primary source of data generated by the ENCODE Consortium. It also provides information on how to access data, including links to 'data are available at the ENCODE Portal', 'ways to get started searching and downloading data', and 'metadata describing the assays and reagents are organized'. A contact information section follows, providing an email address (encode-help@lists.stanford.edu) and a Twitter handle (@encodedcc). Below this is a section titled 'Information available at the portal' which lists the types of data generated by the ENCODE Consortium: raw and processed data from a wide-range of assays, biological samples, and antibody characterizations. Finally, it mentions that additional information about the activities of the ENCODE Consortium is provided on the portal, including publications written by consortium members.

www.encodeproject.org ... Faceted Browsing

Faceted Browsing of Assays

- Click on Data ... Assays
- <https://www.encodeproject.org/search/?type=Experiment>
- In the center is a list of all ENCODE experiments.
- To the left are facets you can choose to filter the list.

The screenshot displays the ENCODE project website's search results page. The top navigation bar includes 'ENCODE', 'Data', 'Methods', 'About', and 'Help', along with a search bar. The main content area is divided into several sections:

- Assay category:** A horizontal bar chart showing the number of experiments for each category: DNA binding (2814), Transcription (1725), DNA accessibility (394), DNA methylation (307), and RNA binding (304). A '+ See more...' link is provided.
- Assay:** A horizontal bar chart showing the number of experiments for each assay type: ChIP-seq (2814), RNA-seq (685), shRNA RNA-seq (445), DNase-seq (355), and RNA microarray (180). A '+ See more...' link is provided.
- Project:** A horizontal bar chart showing the number of experiments for each project: ENCODE (5822), Roadmap (3081), and modENCODE (874).
- RFA:** A horizontal bar chart showing the number of experiments for each RFA: ENCODE2 (2756), ENCODE3 (2474), ENCODE2-Mouse (557), and ENCODE (35).
- Experiment status:** A horizontal bar chart showing the number of experiments for each status: released (5787) and revoked (35).
- Genome assembly (visualization):** A horizontal bar chart showing the number of experiments for each genome assembly: hg19 (3451), mm10 (770), and mm9 (569).

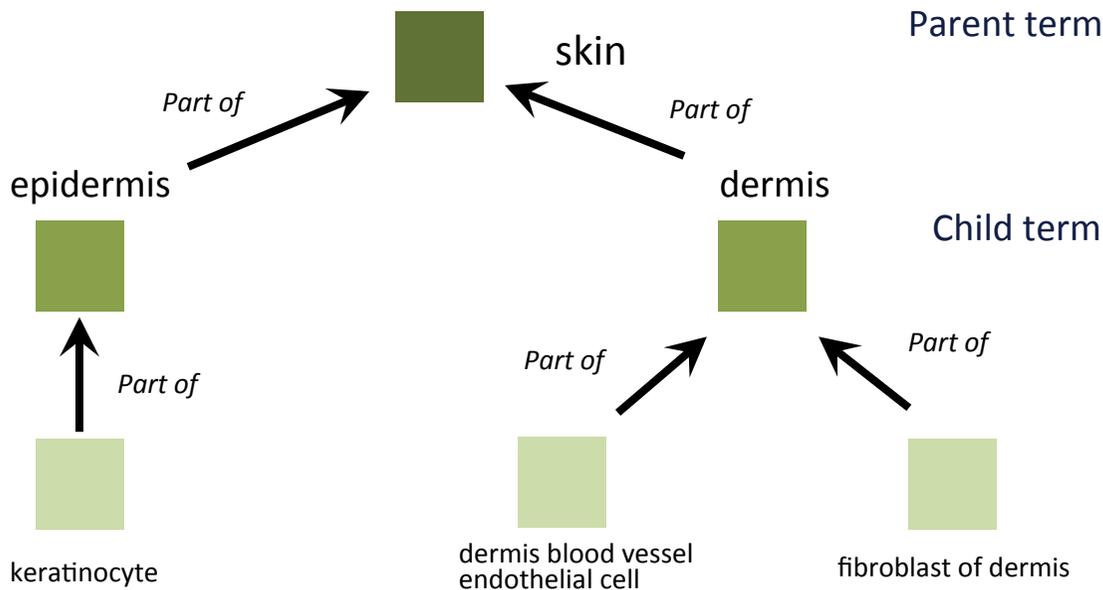
The main results area shows 'Showing 25 of 5822 results' with buttons for 'View All', 'Download', and 'Filter to 100 to visualize'. The results list includes:

- whole-genome shotgun bisulfite sequencing of forebrain (*Mus musculus*, embryonic 12.5 day)** Experiment ENCSR007JSP released. Lab: Joe Ecker, Salk; Project: ENCODE.
- CRISPR genome editing followed by RNA-seq of HepG2 (*Homo sapiens*, child 15 year)** Experiment ENCSR845MNM released. Target: SRSF7; Lab: Brenton Graveley, UConn; Project: ENCODE.
- CRISPR genome editing followed by RNA-seq of HepG2 (*Homo sapiens*, child 15 year)** Experiment ENCSR000SKS released. Target: RPLP0; Lab: Brenton Graveley, UConn; Project: ENCODE.
- CRISPR genome editing followed by RNA-seq of HepG2 (*Homo sapiens*, child 15 year)** Experiment ENCSR105NML released. Target: Non-specific target control; Lab: Brenton Graveley, UConn; Project: ENCODE.
- CRISPR genome editing followed by RNA-seq of HepG2 (*Homo sapiens*, child 15 year)** Experiment ENCSR268XHM released. Target: IGF2BP1; Lab: Brenton Graveley, UConn.

www.encodeproject.org ... The Search Box

Keyword search powered by ontologies

- Search for “skin” ... select Experiments.
- Finds “skin of body” and “keratinocyte”



The screenshot shows the ENCODE search results for the keyword "skin". The search bar at the top right contains the text "skin". The results are displayed in a table format with various filters and categories.

Assay category

| | |
|--------------------|----|
| Transcription | 69 |
| DNA binding | 54 |
| DNA accessibility | 26 |
| DNA methylation | 16 |
| Replication timing | 8 |

+ See more...

Assay

| | |
|----------------|----|
| ChIP-seq | 54 |
| RNA-seq | 29 |
| DNase-seq | 25 |
| RNA microarray | 24 |
| CAGE | 9 |

+ See more...

Project

| | |
|-----------|-----|
| ENCODE | 178 |
| Roadmap | 151 |
| modENCODE | 1 |

RFA

| | |
|---------|-----|
| ENCODE2 | 158 |
| ENCODE3 | 18 |
| ENCODE | 2 |

Experiment status

| | |
|----------|-----|
| released | 176 |
| revoked | 2 |

Genome assembly (visualization)

| | |
|--------|-----|
| hg19 | 154 |
| GRCh38 | 9 |
| mm9 | 2 |

Showing 25 of 178 results

View All Download Filter to 100 to visualize

RNA-seq of skin of body (*Homo sapiens*, fetal) Experiment
Lab: Thomas Gingeras, CSHL
Project: ENCODE
ENCSR000AGA released

RNA-seq of skin of body (*Homo sapiens*, fetal) Experiment
Lab: Thomas Gingeras, CSHL
Project: ENCODE
ENCSR000AFG released

RAMPAGE of skin of body (*Homo sapiens*, fetal) Experiment
Lab: Thomas Gingeras, CSHL
Project: ENCODE
ENCSR000AGU released

ChIP-seq of keratinocyte (*Homo sapiens*) Experiment
Target: H2AFZ
Lab: Bradley Bernstein, Broad
Project: ENCODE
ENCSR000ARL released

ChIP-seq of fibroblast of dermis (*Homo sapiens*, adult) Experiment
Target: H3K4me1
Lab: Bradley Bernstein, Broad
Project: ENCODE
ENCSR000ARV released

ChIP-seq of fibroblast of dermis (*Homo sapiens*, adult) Experiment
Target: H3K27ac
Lab: Bradley Bernstein, Broad
Project: ENCODE
ENCSR000APN released

ChIP-seq of fibroblast of dermis (*Homo sapiens*, adult) Experiment

www.encodeproject.org ... The Matrix

Click the matrix icon near “results”

The screenshot shows the ENCODE project website interface. At the top, there is a navigation bar with 'ENCODE' and dropdown menus for 'Data', 'Methods', 'About', and 'Help'. A search bar on the right contains the text 'skin'. Below the navigation bar, the main content area is divided into several sections. On the left, there are four filter panels: 'Assay category', 'Assay', 'Project', and 'RFA'. Each panel contains a list of categories with corresponding horizontal bar charts and counts. The 'Project' panel shows 'ENCODE' with 178 results. Below these panels, there are sections for 'Experiment status' and 'Genome assembly (visualization)'. The main right-hand section displays a list of search results. At the top of this section, it says 'Showing 25 of 178 results' followed by a grid icon and a matrix icon. A red arrow points from the top right towards the matrix icon. Below this, there are three buttons: 'View All', 'Download', and 'Filter to 100 to visualize'. The results list includes several entries, each with a title, lab name, and project name. The titles are: 'RNA-seq of skin of body (Homo sapiens, fetal)', 'RNA-seq of skin of body (Homo sapiens, fetal)', 'RAMPAGE of skin of body (Homo sapiens, fetal)', 'ChIP-seq of keratinocyte (Homo sapiens)', 'ChIP-seq of fibroblast of dermis (Homo sapiens, adult)', 'ChIP-seq of fibroblast of dermis (Homo sapiens, adult)', and 'ChIP-seq of fibroblast of dermis (Homo sapiens, adult)'. Each entry also includes the lab name (e.g., Thomas Gingeras, CSHL or Bradley Bernstein, Broad) and the project name (ENCODE).

| Assay category | Count |
|--------------------|-------|
| Transcription | 69 |
| DNA binding | 54 |
| DNA accessibility | 26 |
| DNA methylation | 16 |
| Replication timing | 8 |

| Assay | Count |
|----------------|-------|
| ChIP-seq | 54 |
| RNA-seq | 29 |
| DNase-seq | 25 |
| RNA microarray | 24 |
| CAGE | 9 |

| Project | Count |
|-----------|-------|
| ENCODE | 178 |
| Roadmap | 151 |
| modENCODE | 1 |

| RFA | Count |
|---------|-------|
| ENCODE2 | 158 |
| ENCODE3 | 18 |
| ENCODE | 2 |

| Experiment status | Count |
|-------------------|-------|
| released | 176 |
| revoked | 2 |

| Genome assembly (visualization) | Count |
|---------------------------------|-------|
| hg19 | 154 |
| GRCh38 | 9 |
| mm9 | 2 |

| Result Title | Lab | Project |
|--|--------------------------|---------|
| RNA-seq of skin of body (Homo sapiens, fetal) | Thomas Gingeras, CSHL | ENCODE |
| RNA-seq of skin of body (Homo sapiens, fetal) | Thomas Gingeras, CSHL | ENCODE |
| RAMPAGE of skin of body (Homo sapiens, fetal) | Thomas Gingeras, CSHL | ENCODE |
| ChIP-seq of keratinocyte (Homo sapiens) | Bradley Bernstein, Broad | ENCODE |
| ChIP-seq of fibroblast of dermis (Homo sapiens, adult) | Bradley Bernstein, Broad | ENCODE |
| ChIP-seq of fibroblast of dermis (Homo sapiens, adult) | Bradley Bernstein, Broad | ENCODE |
| ChIP-seq of fibroblast of dermis (Homo sapiens, adult) | Bradley Bernstein, Broad | ENCODE |

www.encodeproject.org ... The Matrix

The same list in matrix form

- See the search results as a matrix.
- Facts work here too.
- Select **ChIP-seq ... histone**.

The screenshot shows the ENCODE Data Matrix interface. At the top, there are navigation tabs for ENCODE, Data, Methods, About, and Help. A search bar on the right contains the text 'skin'. Below the navigation is the 'Experiment Matrix' section, which includes a search filter box with 'skin' entered. To the right of the search box are four summary tables: Assay, Assay category, Target of assay, and Date released. Below these are filter panels for Organism, Biosample type, Organ, Project, and Genome assembly (visualization). The main content area displays a matrix of 178 results, with columns representing assays and rows representing biosamples. The matrix is color-coded by biosample type: primary cell (orange), tissue (blue), and stem cell (green). Buttons for 'Download' and 'Visualize' are visible at the bottom of the matrix.

Assay

| | |
|----------------|----|
| ChIP-seq | 54 |
| RNA-seq | 29 |
| DNase-seq | 25 |
| RNA microarray | 24 |
| CAGE | 9 |

Assay category

| | |
|--------------------|----|
| Transcription | 69 |
| DNA binding | 54 |
| DNA accessibility | 26 |
| DNA methylation | 16 |
| Replication timing | 8 |

Target of assay

| | |
|----------------------|----|
| histone | 31 |
| histone modification | 31 |
| narrow histone mark | 16 |
| broad histone mark | 15 |
| transcription factor | 13 |

Date released

| | |
|----------------|----|
| February, 2011 | 29 |
| April, 2011 | 14 |
| March, 2011 | 14 |
| March, 2012 | 14 |
| October, 2011 | 13 |

Available data

| | |
|------------------|-----|
| fastq | 134 |
| bigWig | 128 |
| bam | 127 |
| bed broadPeak | 73 |
| bigBed broadPeak | 73 |

Organism

| | |
|--------------|-----|
| Homo sapiens | 175 |
| Mus musculus | 2 |

Biosample type

| | |
|--------------|-----|
| primary cell | 169 |
| tissue | 7 |
| stem cell | 2 |

Organ

| | |
|------------------|-----|
| skin of body | 178 |
| blood vessel | 9 |
| lymphatic vessel | 6 |

Project

| | |
|-----------|-----|
| ENCODE | 178 |
| Roadmap | 151 |
| modENCODE | 1 |

Genome assembly (visualization)

| | |
|--------|-----|
| hg19 | 154 |
| GRCh38 | 9 |
| mm9 | 2 |

ASSAY

178 results

Clear all filters

| BIOSAMPLE | ChIP-seq | RNA-seq | DNase-seq | RNA microarray | CAGE | DNAme array | RIBS | Repli-seq | genotyping array | RAMPAGE | RNA-PET | Repli-chip | FAIRE-seq | WGAS | microRNA counts | microRNA-seq |
|--------------------------------------|----------|---------|-----------|----------------|------|-------------|------|-----------|------------------|---------|---------|------------|-----------|------|-----------------|--------------|
| primary cell | | | | | | | | | | | | | | | | |
| keratinocyte | 22 | 11 | 2 | 6 | 5 | | | 6 | | 2 | 1 | | | | | |
| fibroblast of dermis | 14 | 2 | 1 | 1 | 1 | | | | | | | | | | | |
| skin fibroblast | 2 | 9 | 5 | | | | 1 | | | | | | | | | |
| foreskin fibroblast | 6 | 2 | 2 | 1 | 1 | | 1 | | | | | | | | | |
| dermis blood vessel endothelial cell | | 1 | 4 | 4 | | | | | | | | | | | | |
| ...and 10 more | | | | | | | | | | | | | | | | |
| tissue | | | | | | | | | | | | | | | | |
| skin of body | 2 | | | | | | | | 1 | | | | | | | |
| lower leg skin | 2 | | | | | | | | | | | | | | | |
| zone of skin | | | | 1 | 1 | | | | | | | | | | | |
| stem cell | | | | | | | | | | | | | | | | |
| stem cell of epidermis | | | | | | | | | | | 2 | | | | | |

Download Visualize

www.encodeproject.org ... The Matrix

Skin ChIP-seq ... histone

- Cells in the matrix are clickable.
- Produces a list of experiments that match the selected criteria.

The screenshot shows the ENCODE project website interface. At the top, there is a navigation bar with 'ENCODE', 'Data', 'Methods', 'About', and 'Help'. A search bar on the right contains the text 'skin'. Below the navigation bar, the 'Experiment Matrix' section is visible. It includes a search input field with 'skin' entered. To the right of the search field, there are several filter categories: 'Assay' (ChIP-seq, 31), 'Assay category' (DNA binding, 31), 'Target of assay' (histone, 31), 'Date released' (February, 2011: 22; March, 2012: 7; August, 2012: 1; November, 2011: 1), and 'Available data' (bam: 31; bed broadPeak: 31; bigBed broadPeak: 31; bigWig: 31; fastq: 31). Below these filters, there are sections for 'Organism' (Homo sapiens, 31), 'Biosample type' (primary cell, 31), 'Organ' (skin of body, 31), 'Project' (Roadmap: 49; ENCODE: 31), and 'Genome assembly (visualization)' (hg19, 31). The main results section, titled 'ASSAY', shows '31 results' and a 'Clear all filters' button. A 'BIOSAMPLE' column is visible on the left of the results table. The results table lists various cell types and their counts: primary cell (15), keratinocyte (15), fibroblast of dermis (11), foreskin fibroblast (2), fibroblast of pedal digit skin (1), fibroblast of skin of abdomen (1), and fibroblast of upper leg skin (1). At the bottom of the results section, there are 'Download' and 'Visualize' buttons.

| Assay | Assay category | Target of assay | Date released | Available data |
|-------------|----------------|-------------------------|-------------------|---------------------|
| ChIP-seq 31 | DNA binding 31 | histone 31 | February, 2011 22 | bam 31 |
| | | histone modification 31 | March, 2012 7 | bed broadPeak 31 |
| | | narrow histone mark 16 | August, 2012 1 | bigBed broadPeak 31 |
| | | broad histone mark 15 | November, 2011 1 | bigWig 31 |
| | | transcription factor 13 | | fastq 31 |

| Organism | Biosample type | Organ | Project | Genome assembly (visualization) |
|-----------------|-----------------|-----------------|-------------------------|---------------------------------|
| Homo sapiens 31 | primary cell 31 | skin of body 31 | Roadmap 49 ENCODE 31 | hg19 31 |

| BIOSAMPLE | Count |
|--------------------------------|-------|
| primary cell | 15 |
| keratinocyte | 15 |
| fibroblast of dermis | 11 |
| foreskin fibroblast | 2 |
| fibroblast of pedal digit skin | 1 |
| fibroblast of skin of abdomen | 1 |
| fibroblast of upper leg skin | 1 |



www.encodeproject.org ... Metadata to Data

Use metadata to find data

- Filter: RNAseq, ENCODE, mm10 assembly

- **Select an experiment:**

<https://www.encodeproject.org/>

[experiments/ENCSR466KZY/](https://www.encodeproject.org/experiments/ENCSR466KZY/)

- **Note metadata on protocols, replicates, etc.**

- Graph: files are related by processing steps

- Download files from the graph or a list

ENCODE Data Methods About Help Search...

Showing 25 of 102 results

View All Download Filter to 100 to visualize

RNA-seq of stomach (*Mus musculus*, embryonic 16.5 day) Experiment
Lab: Barbara Wold, Caltech
Project: ENCODE
ENCSR466KZY released

RNA-seq of kidney (*Mus musculus*, embryonic 16.5 day) Experiment
Lab: Barbara Wold, Caltech
Project: ENCODE
ENCSR537GNQ released

RNA-seq of intestine (*Mus musculus*, embryonic 16.5 day) Experiment
Lab: Barbara Wold, Caltech
Project: ENCODE
ENCSR848GST released

RNA-seq of lung (*Mus musculus*, embryonic 16.5 day) Experiment
Lab: Barbara Wold, Caltech
Project: ENCODE
ENCSR992WBR released

RNA-seq of liver (*Mus musculus*, embryonic 16.5 day) Experiment
Lab: Barbara Wold, Caltech
Project: ENCODE
ENCSR826HIQ released

RNA-seq of heart (*Mus musculus*, embryonic 16.5 day) Experiment
Lab: Barbara Wold, Caltech
Project: ENCODE
ENCSR020DGG released

RNA-seq of hindbrain (*Mus musculus*, embryonic 16.5 day) Experiment
Lab: Barbara Wold, Caltech
Project: ENCODE
ENCSR285WZV released

Assay category
Transcription 102

Assay
ChIP-seq 468
RNA-seq 102
microRNA counts 61
microRNA-seq 58
single cell RNA-seq 56
+ See more...

Project
ENCODE 102

RFA
ENCODE3 95
ENCODE2 6
ENCODE2-Mouse 1
ENCODE 0

Experiment status
released 102

Genome assembly (visualization)
hg19 397
GRCh38 104
mm10 102
mm9 99

Organism
Mus musculus 102

Biosample type
tissue 86
primary cell 10

https://www.encodeproject.org

www.encodeproject.org ... Metadata to Data

Use metadata to find data

- Filter: RNAseq, ENCODE, mm10 assembly
- **Select an experiment:**
<https://www.encodeproject.org/experiments/ENCSR466KZY/>
- **Note metadata on protocols, replicates, etc.**
- Graph: files are related by processing steps
- Download files from the graph or a list

The screenshot displays the ENCODE website interface. At the top, there is a navigation bar with 'ENCODE' and dropdown menus for 'Data', 'Methods', 'About', and 'Help'. A search bar is located on the right. Below the navigation bar, the breadcrumb trail reads 'EXPERIMENTS / RNA-SEQ / MUS MUSCULUS / STOMACH'. The main heading is 'Experiment summary for ENCSR466KZY'. Below the heading, there are two status indicators: 'Status: released' (in green) and 'Validation: pending' (in orange). The main content area contains a table of metadata:

| | |
|---------------------------|---|
| Assay: | RNA-seq |
| Replication type: | isogenic |
| Biosample summary: | stomach (<i>Mus musculus</i> , embryonic 16.5 day mixed) |
| Type: | tissue |
| Description: | RNA-Seq on embryonic 16.5 day mouse stomach |
| Lab: | Barbara Wold, Caltech |
| Award PI: | Richard Myers, HAIB |
| Project: | ENCODE |
| Date released: | 2015-10-13 |

Below the main summary, there is a section for 'Assay details' with the following information:

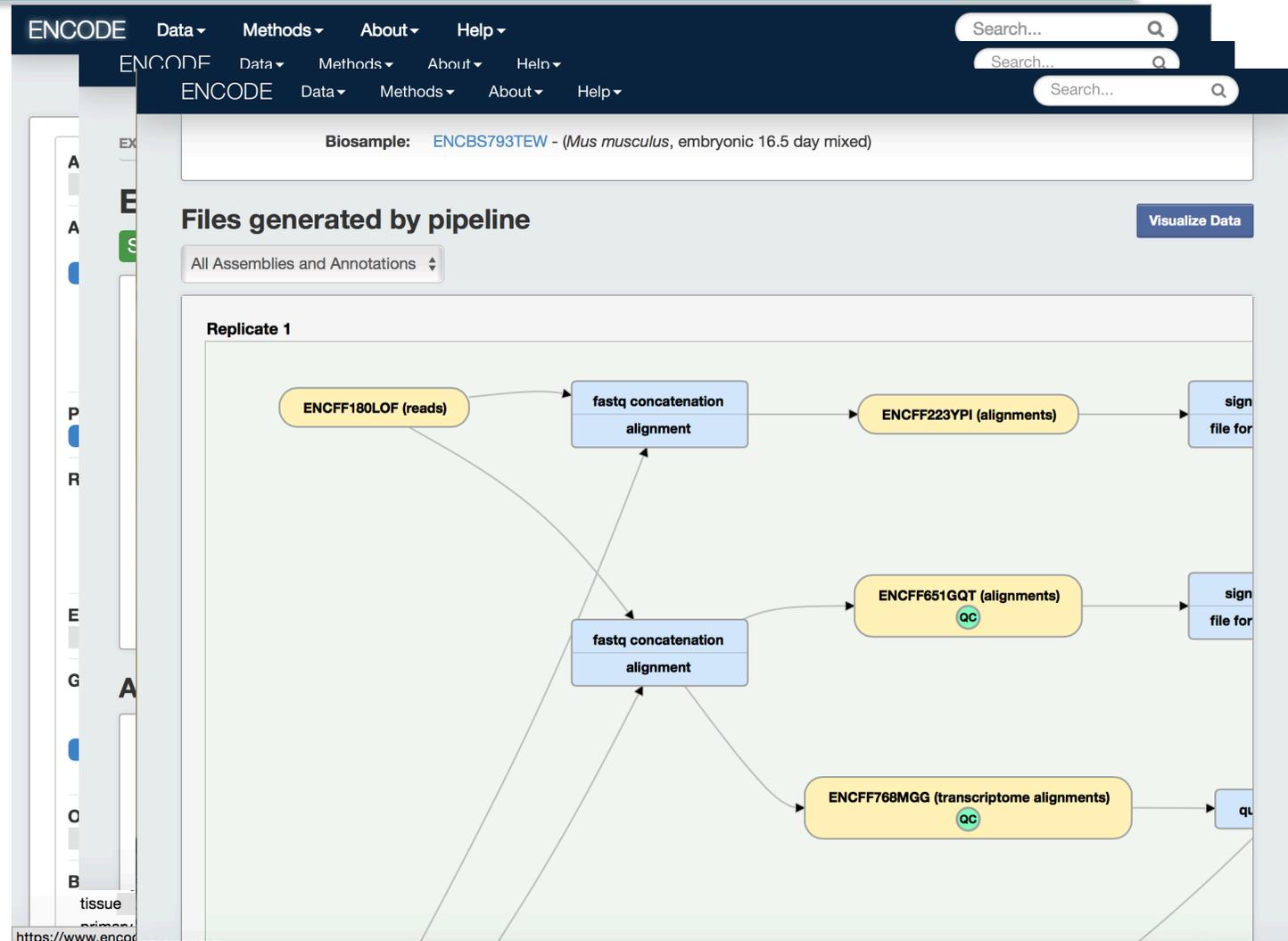
| | |
|------------------------------|---------------------------------|
| Nucleic acid type: | RNA |
| Lysis method: | Ambion mirVana |
| Extraction method: | Ambion mirVana |
| Fragmentation method: | chemical (Nextera tagmentation) |

At the bottom of the page, there is a footer with the URL 'https://www.encodeproject.org' and the text 'Project: ENCODE released'.

www.encodeproject.org ... Metadata to Data

Use metadata to find data

- Filter: RNAseq, ENCODE, mm10 assembly
- **Select an experiment:**
<https://www.encodeproject.org/experiments/ENCSR466KZY/>
- **Note metadata on protocols, replicates, etc.**
- Graph: files are related by processing steps.
- Download files from the graph or a list.



www.encodeproject.org ... Metadata to Data

Use metadata to find data

- Filter: RNAseq, ENCODE, mm10 assembly
- **Select an experiment:**
<https://www.encodeproject.org/experiments/ENCSR466KZY/>
- **Note metadata on protocols, replicates, etc.**
- Graph: files are related by processing steps.
- Download files from the graph or a list.

The screenshot displays the ENCODE project website interface. At the top, there is a navigation bar with 'ENCODE' and dropdown menus for 'Data', 'Methods', 'About', and 'Help'. A search bar is located on the right. Below the navigation bar, a workflow graph is shown with nodes: 'ENCFF494GQK (reads)', 'fastq concatenation alignment', and 'ENCFF084TIX (transcriptome alignments)'. A 'Download Graph' button is visible. Below the graph, a metadata table is displayed:

| | |
|--------------------------|--|
| Format: | bam |
| Output: | transcriptome alignments |
| Biological Replicate(s): | [2] |
| Technical Replicate: | 1 |
| Mapping assembly: | mm10 |
| Genome annotation: | M4 |
| Lab: | ENCODE Processing Pipeline |
| Date added: | 2015-10-02 |
| Software: | concat-fastqs 1.0.2 star 2.4.0k samtools 0.1.19-96b5f2294a |
| File download: | Download |

At the bottom of the screenshot, the text 'Files linked to ENCSR466KZY' is visible.

www.encodeproject.org ... Metadata to Data

Use metadata to find data

- Filter: RNAseq, ENCODE, mm10 assembly
- **Select an experiment:**
<https://www.encodeproject.org/experiments/ENCSR466KZY/>
- **Note metadata on protocols, replicates, etc.**
- Graph: files are related by processing steps.
- Download files from the graph or a list.

ENCODE Data Methods About Help Search... Download Graph

Files linked to ENCSR466KZY

Raw data

| Accession | File type | Biological replicate | Technical replicate | Read length | Run type | Paired end | Mapping assembly | Lab | Date added | Validation status |
|--|-----------|----------------------|---------------------|-------------|--------------|------------|------------------|-----------------------|------------|-------------------|
| ENCFF180LOF Download 4.65 GB | fastq | 1 | 1 | 116 nt | single-ended | | | Barbara Wold, Caltech | 2015-09-24 | pending |
| ENCFF494GQK Download 2.87 GB | fastq | 2 | 1 | 116 nt | single-ended | | | Barbara Wold, Caltech | 2015-09-24 | pending |

Processed data

| Accession | File type | Output type | Biological replicate(s) | Technical replicate | Mapping assembly | Genome annotation | Lab | Date added | Validation status |
|--|-----------|------------------------|-------------------------|---------------------|------------------|-------------------|----------------------------|------------|-------------------|
| ENCFF493DUV Download 3.48 GB | bam | alignments | 2 | 1 | mm10 | M4 | ENCODE Processing Pipeline | 2015-10-02 | pending |
| ENCFF223YPI Download 4.38 GB | bam | alignments | 1 | 1 | mm10 | M4 | ENCODE Processing Pipeline | 2015-10-02 | pending |
| ENCFF910QNE Download 175 MB | bigWig | signal of unique reads | 1 | 1 | mm10 | M4 | ENCODE Processing Pipeline | 2015-10-02 | pending |
| ENCFF618CCD Download | bigWig | signal of all reads | 1 | 1 | mm10 | M4 | ENCODE Processing Pipeline | 2015-10-02 | pending |

Files linked to ENCSR466KZY

www.encodeproject.org ... Visualization

Use metadata to find data:

- Search for “H3K9ac neural tube”.
- Select an experiment, for example <https://www.encodeproject.org/experiments/ENCSR087PLZ/>
- Note metadata on protocols, replicates.
- Graph: files are related by processing steps.
- Download from the graph or a list.
- Click on “Visualize Data” to visualize the results of this experiment.

The screenshot shows the ENCODE project website interface. At the top, there is a navigation bar with 'ENCODE' and dropdown menus for 'Data', 'Methods', 'About', and 'Help'. A search bar on the right contains the text 'H3K9ac neural tube'. Below the navigation bar, there is a sidebar on the left with various metadata filters, each with a count of 4. The main content area displays 'Showing 4 of 4 results' with a grid icon and two buttons: 'Download' and 'Visualize'. Below this, there are four experiment entries, each with a title, target, lab, and project information. A red arrow points to the second entry, 'ChIP-seq of neural tube (Mus musculus, embryonic 13.5 day)'. The sidebar filters include: Assay category (DNA binding), Assay (ChIP-seq), Project (ENCODE), RFA (ENCODE3), Experiment status (released), Genome assembly (visualization) (mm10), Organism (Mus musculus), Target of assay (histone, histone modification, narrow histone mark), Biosample type (tissue), Life stage (embryonic), and Pipeline (Histone ChIP-seq, Transcription factor ChIP-seq).

| Assay category | Count |
|----------------|-------|
| DNA binding | 4 |

| Assay | Count |
|----------|-------|
| ChIP-seq | 4 |

| Project | Count |
|---------|-------|
| ENCODE | 4 |

| RFA | Count |
|---------|-------|
| ENCODE3 | 4 |

| Experiment status | Count |
|-------------------|-------|
| released | 4 |

| Genome assembly (visualization) | Count |
|---------------------------------|-------|
| mm10 | 4 |

| Organism | Count |
|--------------|-------|
| Mus musculus | 4 |

| Target of assay | Count |
|----------------------|-------|
| histone | 4 |
| histone modification | 4 |
| narrow histone mark | 4 |

| Biosample type | Count |
|----------------|-------|
| tissue | 4 |

| Life stage | Count |
|------------|-------|
| embryonic | 4 |

| Pipeline | Count |
|-------------------------------|-------|
| Histone ChIP-seq | 4 |
| Transcription factor ChIP-seq | 4 |

Showing 4 of 4 results

Download Visualize

ChIP-seq of neural tube (*Mus musculus*, embryonic 15.5 day) Experiment ENCSR571HOT released
Target: H3K9ac
Lab: Bing Ren, UCSD
Project: ENCODE

ChIP-seq of neural tube (*Mus musculus*, embryonic 13.5 day) Experiment ENCSR087PLZ released
Target: H3K9ac
Lab: Bing Ren, UCSD
Project: ENCODE

ChIP-seq of neural tube (*Mus musculus*, embryonic 14.5 day) Experiment ENCSR511LWL released
Target: H3K9ac
Lab: Bing Ren, UCSD
Project: ENCODE

ChIP-seq of neural tube (*Mus musculus*, embryonic 11.5 day) Experiment ENCSR547PLI released
Target: H3K9ac
Lab: Bing Ren, UCSD
Project: ENCODE

www.encodeproject.org ... Visualization

Use metadata to find data:

- Search for “H3K9ac neural tube”
- Facet on ChIP-seq; mouse; mm10 assembly
- Select an experiment, for example <https://www.encodeproject.org/experiments/ENCSR087PLZ/>
- Note metadata on protocols, replicates
- Graph: files are related by processing steps
- Download from the graph or a list
- Click on “Visualize Data” ... mm10 to visualize the results of this experiment.

The screenshot displays the ENCODE project interface. At the top, there are navigation menus for 'Data', 'Methods', 'About', and 'Help', along with a search bar. The main content area shows a workflow graph for 'Replicate 2'. The graph starts with 'ENCF531KNS (reads)' leading to an 'alignment' step (fastq concatenation, filtering). This leads to 'ENCF891NNX (alignments)' (QC QC QC) and 'ENCF101KOM (alignments)'. Both alignment sets lead to a 'peak calling' step (signal generation). Below the graph is a 'Download Graph' button. Underneath is a 'File summary' section with a 'Visualize Data' button. The 'Raw data files' table is as follows:

| Accession | File type | Biological replicate | Library | Run type | Read | Lab | Date added | File size | Validation status |
|------------|-----------|----------------------|-------------|----------|------|----------------|------------|-----------|-------------------|
| ENCF833RGB | fastq | 1 | ENCLB960XUI | SE 50nt | | Bing Ren, UCSD | 2015-06-19 | 3.03 GB | pending |
| ENCF531KNS | fastq | 2 | ENCLB045XUT | SE 50nt | | Bing Ren, UCSD | 2015-06-19 | 2.71 GB | pending |

The 'Processed data files' section is currently empty.

Visualize the Experiment

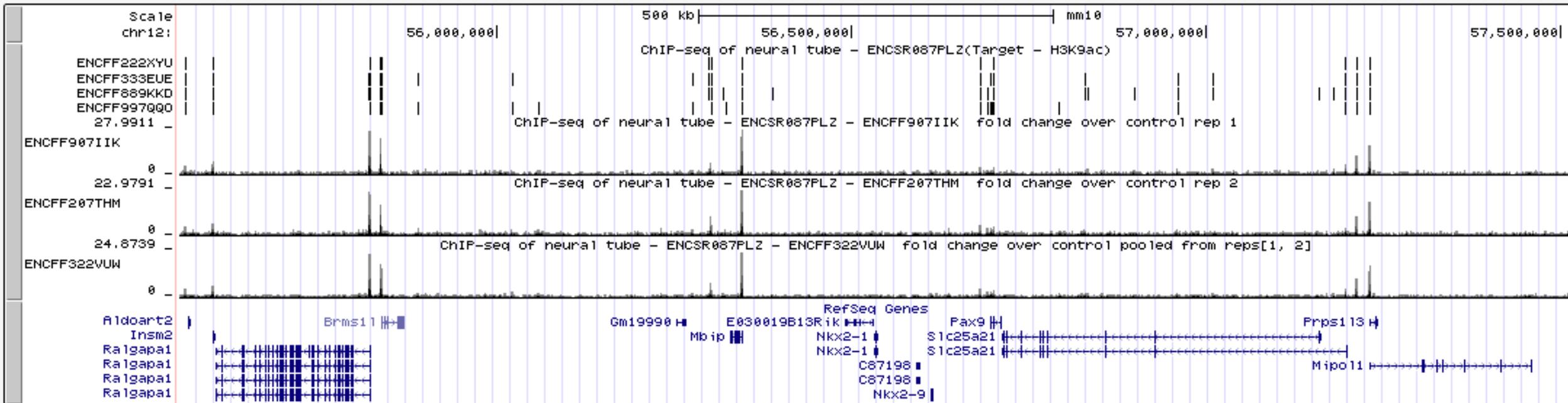
Adjust the browser settings to display fold-over-signal in "full"

UCSC Genome Browser on Mouse Dec. 2011 (GRCm38/mm10) Assembly

move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x 100x

chr12:55,552,346-57,515,345 1,963,000 bp.

chr12 (qC1) 12qA1.1 qA1.3 12qA2 12qA3 12qB1 12qB3 12qC1 12qC2 12qC3 qD1 qD2 12qD3 12qE 12qF1 12qF2



Find Several Experiments

Use metadata to find data:

- Search for “H3K9ac neural tube”
- Get a list of several experiments
- Click on “Visualize Data” to visualize all the experiments matching this search.

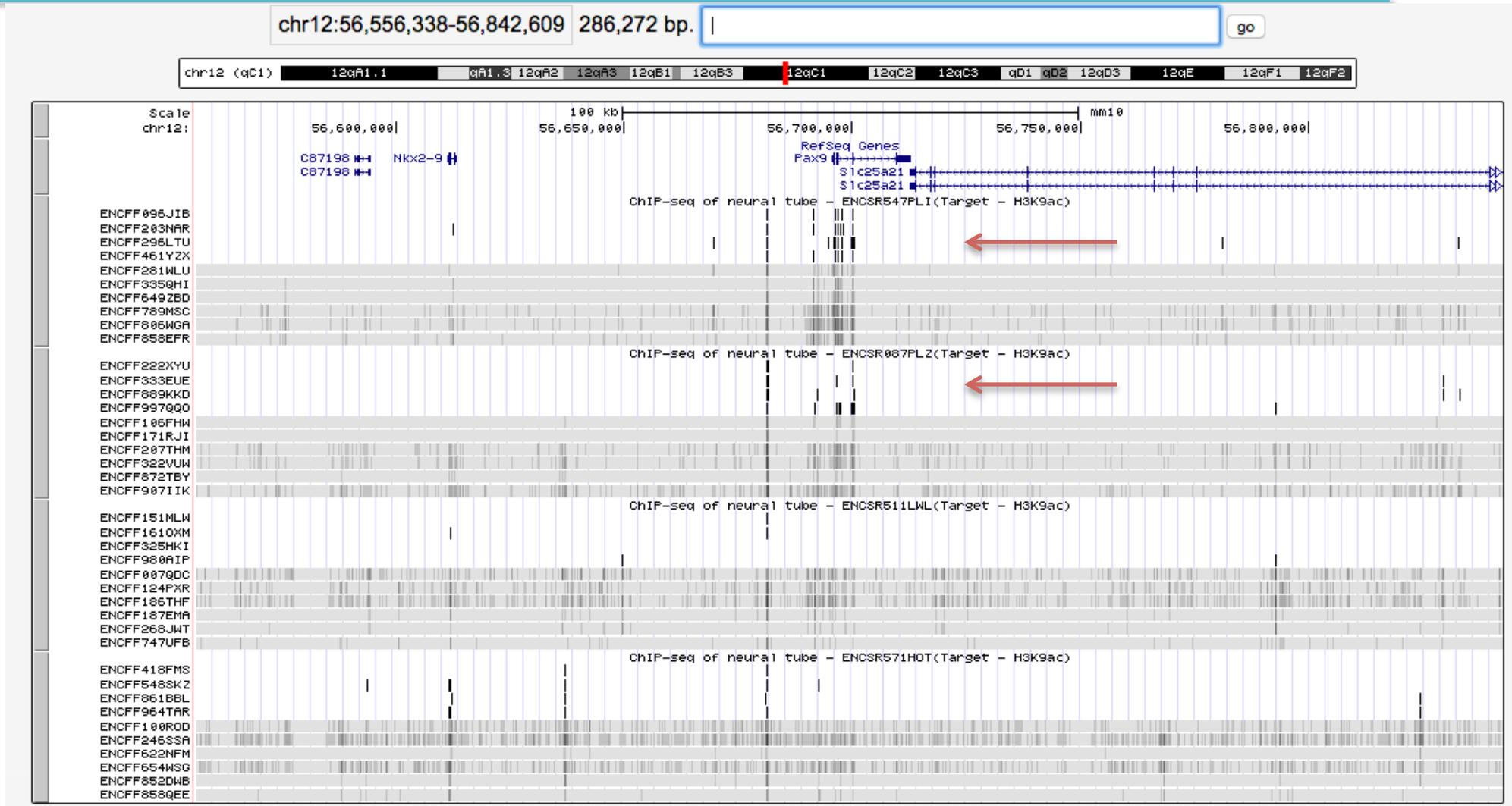
The screenshot shows the ENCODE website interface. At the top, there is a navigation bar with 'ENCODE' and dropdown menus for 'Data', 'Methods', 'About', and 'Help'. A search bar on the right contains the text 'H3K9ac neural tube'. Below the navigation bar, there is a sidebar on the left with filters for 'Assay category', 'Assay', 'Project', 'RFA', 'Experiment status', 'Genome assembly (visualization)', 'Organism', 'Target of assay', and 'Biosample type'. The main content area shows 'Showing 4 of 4 results' with a grid icon. Below this, there are two buttons: 'Download' and 'Visualize'. A red arrow points from the 'Visualize' button to the first result. The results are listed as follows:

| Experiment Title | Target | Lab | Project | Status |
|---|--------|----------------|---------|----------|
| ChIP-seq of neural tube (<i>Mus musculus</i> , embryonic 15.5 day) | H3K9ac | Bing Ren, UCSD | ENCODE | released |
| ChIP-seq of neural tube (<i>Mus musculus</i> , embryonic 13.5 day) | H3K9ac | Bing Ren, UCSD | ENCODE | released |
| ChIP-seq of neural tube (<i>Mus musculus</i> , embryonic 14.5 day) | H3K9ac | Bing Ren, UCSD | ENCODE | released |
| ChIP-seq of neural tube (<i>Mus musculus</i> , embryonic 11.5 day) | H3K9ac | Bing Ren, UCSD | ENCODE | released |



Visualize Several Experiments

Stage-dependent
H3K9ac signal
present at Pax9 in
neural tube at
e11.5, e13.5.



Find & Download Several Experiments

Use metadata to find data:

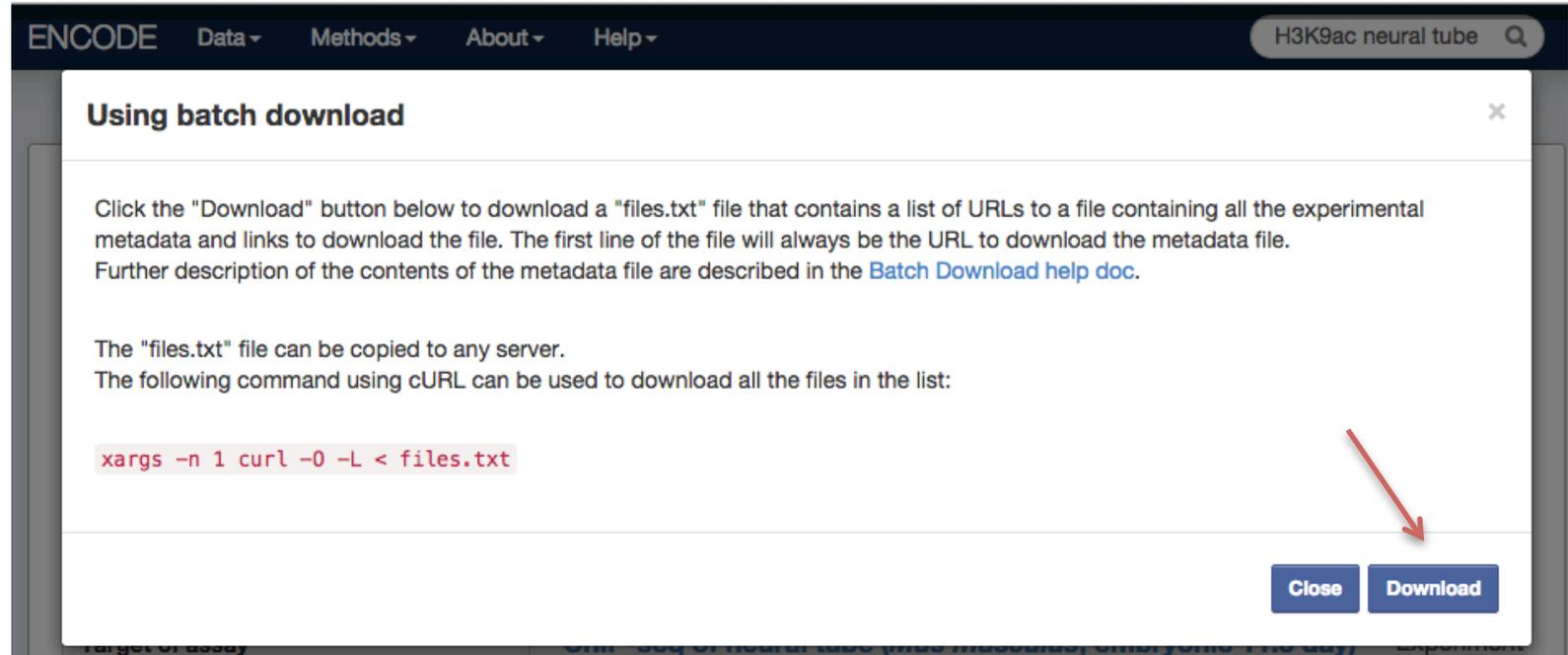
- Search for “H3K9ac neural tube”
- Get a list of several experiments
- Click on “Download” to download selected metadata and complete links to data.

The screenshot shows the ENCODE website interface. At the top, there is a navigation bar with 'ENCODE' and menu items 'Data', 'Methods', 'About', and 'Help'. A search bar on the right contains the text 'H3K9ac neural tube'. Below the navigation bar, a red arrow points to the search results area. On the left side, there is a filter sidebar with the following categories and counts: Assay category (DNA binding: 4), Assay (ChIP-seq: 4), Project (ENCODE: 4), RFA (ENCODE3: 4), Experiment status (released: 4), Genome assembly (visualization) (mm10: 4), Organism (Mus musculus: 4), Target of assay (histone: 4, histone modification: 4, narrow histone mark: 4), and Biosample type (tissue: 4). The main content area displays 'Showing 4 of 4 results' with a grid icon. Below this, there are two buttons: 'Download' and 'Visualize'. The results list four experiments, each with a title, target, lab, and project information. The experiments are: 1. ChIP-seq of neural tube (Mus musculus, embryonic 15.5 day) - Experiment ENCSR571HOT released. 2. ChIP-seq of neural tube (Mus musculus, embryonic 13.5 day) - Experiment ENCSR087PLZ released. 3. ChIP-seq of neural tube (Mus musculus, embryonic 14.5 day) - Experiment ENCSR511LWL released. 4. ChIP-seq of neural tube (Mus musculus, embryonic 11.5 day) - Experiment ENCSR547PLI released.

Download Several Experiments

Use metadata to find data:

- Search for “H3K9ac neural tube”
- Get a list of several experiments
- Click on “Download” to download selected metadata and complete links to data.



ENCODE Data Methods About Help H3K9ac neural tube

Using batch download

Click the "Download" button below to download a "files.txt" file that contains a list of URLs to a file containing all the experimental metadata and links to download the file. The first line of the file will always be the URL to download the metadata file. Further description of the contents of the metadata file are described in the [Batch Download help doc](#).

The "files.txt" file can be copied to any server.
The following command using cURL can be used to download all the files in the list:

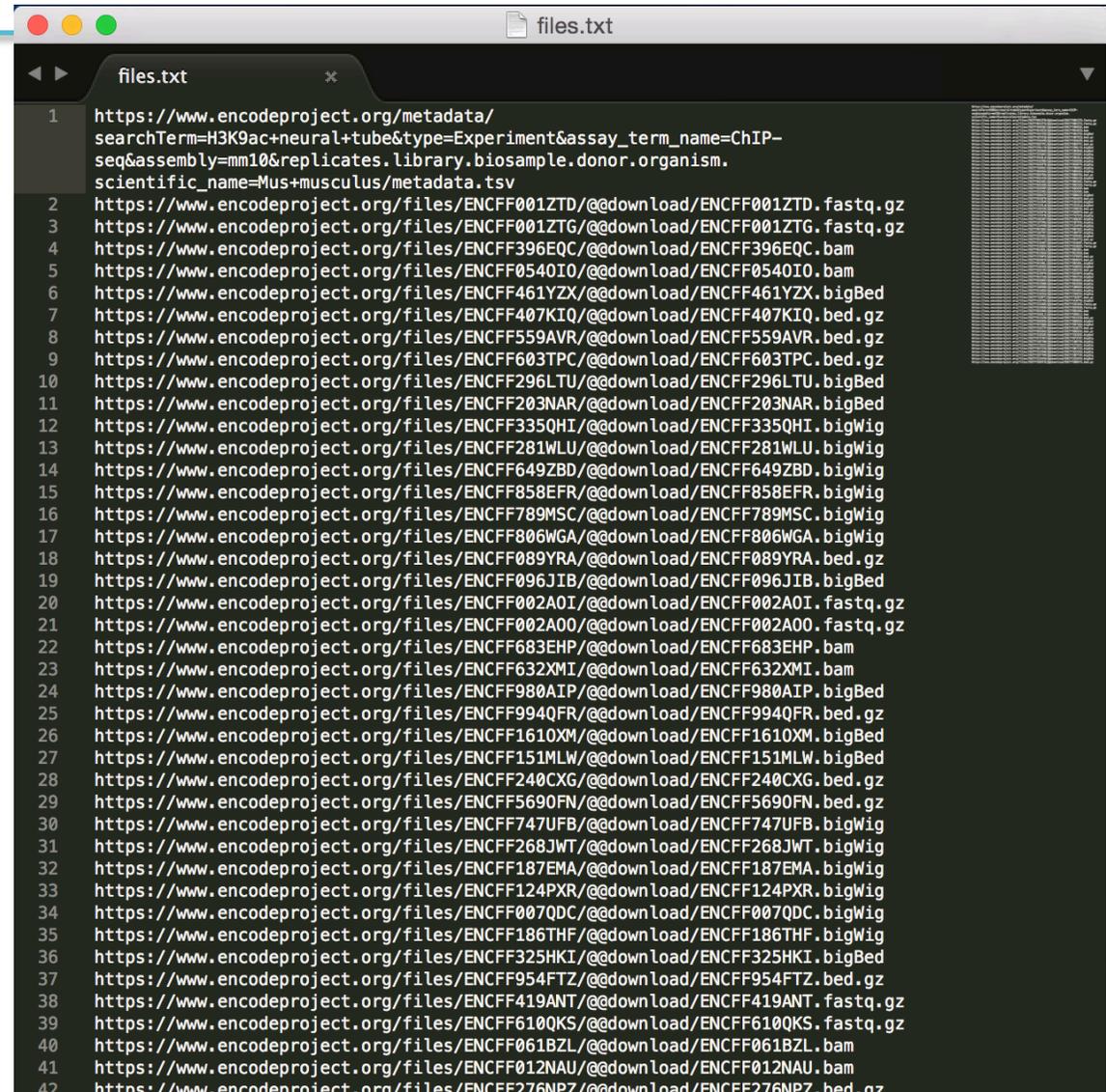
```
xargs -n 1 curl -O -L < files.txt
```

Close Download



Download Several Experiments

- “Download” produces a file with a list of links to all the files for all the experiments in your search.
- You can iterate through the list in your own script.
- Or:
`xargs -n 1 curl -O -L < files.txt`
- The first link is to a file called metadata.tsv that contains metadata you need to interpret what each file is.



```
files.txt
1 https://www.encodeproject.org/metadata/
  searchTerm=H3K9ac+neural+tube&type=Experiment&assay_term_name=ChIP-
  seq&assembly=mm10&replicates.library.biosample.donor.organism.
  scientific_name=Mus+musculus/metadata.tsv
2 https://www.encodeproject.org/files/ENCF001ZTD/@download/ENCF001ZTD.fastq.gz
3 https://www.encodeproject.org/files/ENCF001ZTG/@download/ENCF001ZTG.fastq.gz
4 https://www.encodeproject.org/files/ENCF396EQC/@download/ENCF396EQC.bam
5 https://www.encodeproject.org/files/ENCF0540I0/@download/ENCF0540I0.bam
6 https://www.encodeproject.org/files/ENCF461YZX/@download/ENCF461YZX.bigBed
7 https://www.encodeproject.org/files/ENCF407KI0/@download/ENCF407KI0.bed.gz
8 https://www.encodeproject.org/files/ENCF559AVR/@download/ENCF559AVR.bed.gz
9 https://www.encodeproject.org/files/ENCF603TPC/@download/ENCF603TPC.bed.gz
10 https://www.encodeproject.org/files/ENCF296LTU/@download/ENCF296LTU.bigBed
11 https://www.encodeproject.org/files/ENCF203NAR/@download/ENCF203NAR.bigBed
12 https://www.encodeproject.org/files/ENCF335QHI/@download/ENCF335QHI.bigWig
13 https://www.encodeproject.org/files/ENCF281WLU/@download/ENCF281WLU.bigWig
14 https://www.encodeproject.org/files/ENCF649ZBD/@download/ENCF649ZBD.bigWig
15 https://www.encodeproject.org/files/ENCF858EFR/@download/ENCF858EFR.bigWig
16 https://www.encodeproject.org/files/ENCF789MSC/@download/ENCF789MSC.bigWig
17 https://www.encodeproject.org/files/ENCF806WGA/@download/ENCF806WGA.bigWig
18 https://www.encodeproject.org/files/ENCF089YRA/@download/ENCF089YRA.bed.gz
19 https://www.encodeproject.org/files/ENCF096JIB/@download/ENCF096JIB.bigBed
20 https://www.encodeproject.org/files/ENCF002A0I/@download/ENCF002A0I.fastq.gz
21 https://www.encodeproject.org/files/ENCF002A00/@download/ENCF002A00.fastq.gz
22 https://www.encodeproject.org/files/ENCF683EHP/@download/ENCF683EHP.bam
23 https://www.encodeproject.org/files/ENCF632XMI/@download/ENCF632XMI.bam
24 https://www.encodeproject.org/files/ENCF980AIP/@download/ENCF980AIP.bigBed
25 https://www.encodeproject.org/files/ENCF994QFR/@download/ENCF994QFR.bed.gz
26 https://www.encodeproject.org/files/ENCF1610XM/@download/ENCF1610XM.bigBed
27 https://www.encodeproject.org/files/ENCF151MLW/@download/ENCF151MLW.bigBed
28 https://www.encodeproject.org/files/ENCF240CXG/@download/ENCF240CXG.bed.gz
29 https://www.encodeproject.org/files/ENCF5690FN/@download/ENCF5690FN.bed.gz
30 https://www.encodeproject.org/files/ENCF747UFB/@download/ENCF747UFB.bigWig
31 https://www.encodeproject.org/files/ENCF268JWT/@download/ENCF268JWT.bigWig
32 https://www.encodeproject.org/files/ENCF187EMA/@download/ENCF187EMA.bigWig
33 https://www.encodeproject.org/files/ENCF124PXR/@download/ENCF124PXR.bigWig
34 https://www.encodeproject.org/files/ENCF007QDC/@download/ENCF007QDC.bigWig
35 https://www.encodeproject.org/files/ENCF186THF/@download/ENCF186THF.bigWig
36 https://www.encodeproject.org/files/ENCF325HKI/@download/ENCF325HKI.bigBed
37 https://www.encodeproject.org/files/ENCF954FTZ/@download/ENCF954FTZ.bed.gz
38 https://www.encodeproject.org/files/ENCF419ANT/@download/ENCF419ANT.fastq.gz
39 https://www.encodeproject.org/files/ENCF610QKS/@download/ENCF610QKS.fastq.gz
40 https://www.encodeproject.org/files/ENCF061BZL/@download/ENCF061BZL.bam
41 https://www.encodeproject.org/files/ENCF012NAU/@download/ENCF012NAU.bam
42 https://www.encodeproject.org/files/ENCF276NPZ/@download/ENCF276NPZ.bed.gz
```



Download Several Experiments

- metadata.tsv: Each line contains metadata on a file from the download package.

| | A | B | C | D | E | F | G | H | I | J | K |
|----|----------------|-------------------|--------------------------|----------------------|----------|-------------------|---------------------|----------------|----------------------|---------------|--------------------|
| 1 | File accession | File format | Output type | Experiment accession | Assay | Biosample term id | Biosample term name | Biosample type | Biosample life stage | Biosample sex | Biosample organism |
| 2 | ENCF001ZTD | fastq | reads | ENCSR547PLI | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 3 | ENCF001ZTG | fastq | reads | ENCSR547PLI | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 4 | ENCF396EQC | bam | alignments | ENCSR547PLI | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 5 | ENCF054OIO | bam | alignments | ENCSR547PLI | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 6 | ENCF461YZX | bigBed narrowPeak | peaks | ENCSR547PLI | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 7 | ENCF407KIQ | bed narrowPeak | peaks | ENCSR547PLI | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 8 | ENCF559AVR | bed narrowPeak | peaks | ENCSR547PLI | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 9 | ENCF603TPC | bed narrowPeak | peaks | ENCSR547PLI | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 10 | ENCF296LTU | bigBed narrowPeak | peaks | ENCSR547PLI | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 11 | ENCF203NAR | bigBed narrowPeak | peaks | ENCSR547PLI | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 12 | ENCF335QHI | bigWig | signal p-value | ENCSR547PLI | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 13 | ENCF281WLU | bigWig | signal p-value | ENCSR547PLI | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 14 | ENCF649ZBD | bigWig | signal p-value | ENCSR547PLI | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 15 | ENCF858EFR | bigWig | fold change over control | ENCSR547PLI | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 16 | ENCF789MSC | bigWig | fold change over control | ENCSR547PLI | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 17 | ENCF806WGA | bigWig | fold change over control | ENCSR547PLI | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 18 | ENCF089YRA | bed narrowPeak | replicated peaks | ENCSR547PLI | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 19 | ENCF096JIB | bigBed narrowPeak | replicated peaks | ENCSR547PLI | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 20 | ENCF002AOI | fastq | reads | ENCSR511LWL | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 21 | ENCF002AOO | fastq | reads | ENCSR511LWL | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 22 | ENCF683EHP | bam | alignments | ENCSR511LWL | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 23 | ENCF632XMI | bam | alignments | ENCSR511LWL | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 24 | ENCF980AIP | bigBed narrowPeak | peaks | ENCSR511LWL | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 25 | ENCF994QFR | bed narrowPeak | peaks | ENCSR511LWL | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 26 | ENCF161OXM | bigBed narrowPeak | peaks | ENCSR511LWL | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 27 | ENCF151MLW | bigBed narrowPeak | peaks | ENCSR511LWL | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |
| 28 | ENCF240CYC | bed narrowPeak | peaks | ENCSR511LWL | ChIP-seq | UBERON:0001049 | neural tube | tissue | embryonic | mixed | Mus musculus |



Programmatic access via the ENCODE REST API

- All Portal content is accessible via URL's; just add `?format=json`
- The database record is returned in JSON format
- JSON can be parsed in your language of choice

```
GET_object.py *
1  #!/usr/bin/env python
2
3  import requests
4
5  URL = 'https://www.encodeproject.org/experiments/ENCSR236EGS/?format=json'
6
7  response = requests.get(URL)
8
9  experiment = response.json()
10
11 print experiment['accession']
12 print experiment['description']
13
```



Programmatic access via the ENCODE REST API

```
GET_object.py *
1  #!/usr/bin/env python
2
3  import requests
4
5  URL = 'https://www.encodeproject.org/experiments/ENCSR236EGS/?format=json'
6
7  response = requests.get(URL)
8
9  experiment = response.json()
10
11 print experiment['accession']
12 print experiment['description']
13
```

```
jseth:Keystone Epigenomics 2015 jseth$ ./GET_object.py
ENCSR236EGS
RNA-seq on a dissected area of layer V from an 8 month old male wild type C57B16 mouse
jseth:Keystone Epigenomics 2015 jseth$ █
```



Programmatic access via the ENCODE REST API

```
GET_search.py *
1  #!/usr/bin/env python
2
3  import requests
4
5  URL = ('https://www.encodeproject.org/search/?'
6        'type=experiment&'
7        'assay_term_name=ChIP-seq&'
8        'replicates.library.biosample.donor.organism.scientific_name=Homo sapiens&'
9        'target.investigated_as=transcription factor&'
10       'replicates.library.biosample.biosample_type=in vitro differentiated cells&'
11       'format=json')
12
13  response = requests.get(URL)
14
15  search_result = response.json()['@graph']
16
17  #extract and print the target for each experiment
18  print '\n'.join([experiment['target']['label'] for experiment in search_result])
19
```



The ENCODE Portal: Recap

- Interactive access to ENCODE metadata via faceted browsing and search
- Interactive retrieval of ENCODE data one file at a time
- Batch download of ENCODE metadata and data files
- Programmatic access using the ENCODE REST API

Next: ENCODE's Integrative Results and More Visualization



Contributors

ENCODE Data Coordinating Center

Mike Cherry, PI, Stanford

Jim Kent, co-PI, UCSC

Eurie Hong, Project Manager

Pipeline Developers

Ben Hitz, WGBS, Software Lead

Tim Dreszer, RNA-seq, DNase-seq

J. Seth Strattan, CHIP-seq

Portal Developers

Laurence Rowe

Nikhil Poddaturi

Forrest Tanaka

Data Wranglers

Esther Chan

Jean Davidson

Venkat Malladi

Cricket Sloan

J. Seth Strattan

QA & Biocuration Assistance

Brian Lee

Marcus Ho

Aditi Narayanan

Support Staff

Stuart Miyasato

Matt Simison

Zhenhua Wang

ENCODE Data Analysis Center

Zhiping Weng, PI, University of Massachusetts

Mark Gerstein, co-PI, Yale

Methylation

Junko Tsuji, U Mass

Eric Mendenhall, U Alabama, HAIB

RNA-seq

Alex Dobin, CSHL

Carrie Davis, CSHL

Rafael Irizarry, Harvard

Xintao Wei, UConn

Brent Gravely, UConn

Colin Dewey, U Wisconsin

Roderic Guigó, CRG

Sarah Djebali, CRG

ChIP-seq

Anshul Kundaje, Stanford

Nathan Boley, Stanford

Jin Lee, Stanford



@encodedcc



encode-help@lists.stanford.edu

<https://github.com/ENCODE-DCC/>

