

# Intro to the ENCODE portal

Jason Hilton, PhD

Data Wrangler

ENCODE Data Coordination Center

# ENCODE Data Coordination Center

encode-help@lists.stanford.edu  
@EncodeDCC  
<https://github.com/ENCODE-DCC>



Mike



Idan



Ben



Seth

## Data Wrangling

## Pipelines Development



Meenakshi



Jennifer



Yunhai



Jason



Casey



Forrest



Jin



Bek



Vanessa



Jennifer



Paul



Zack



Khine



Emma



Phil



Otto

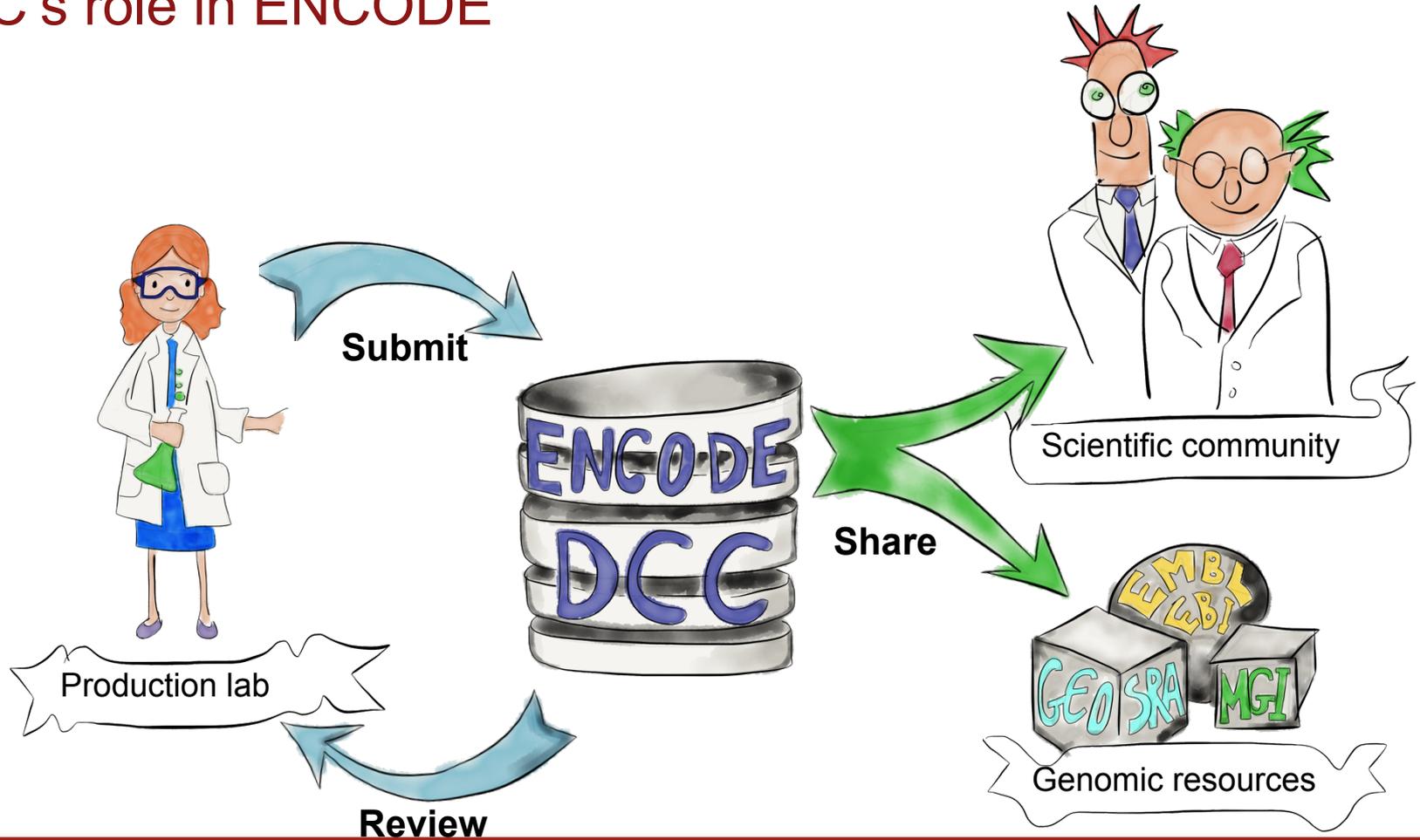


Keenan

## Software

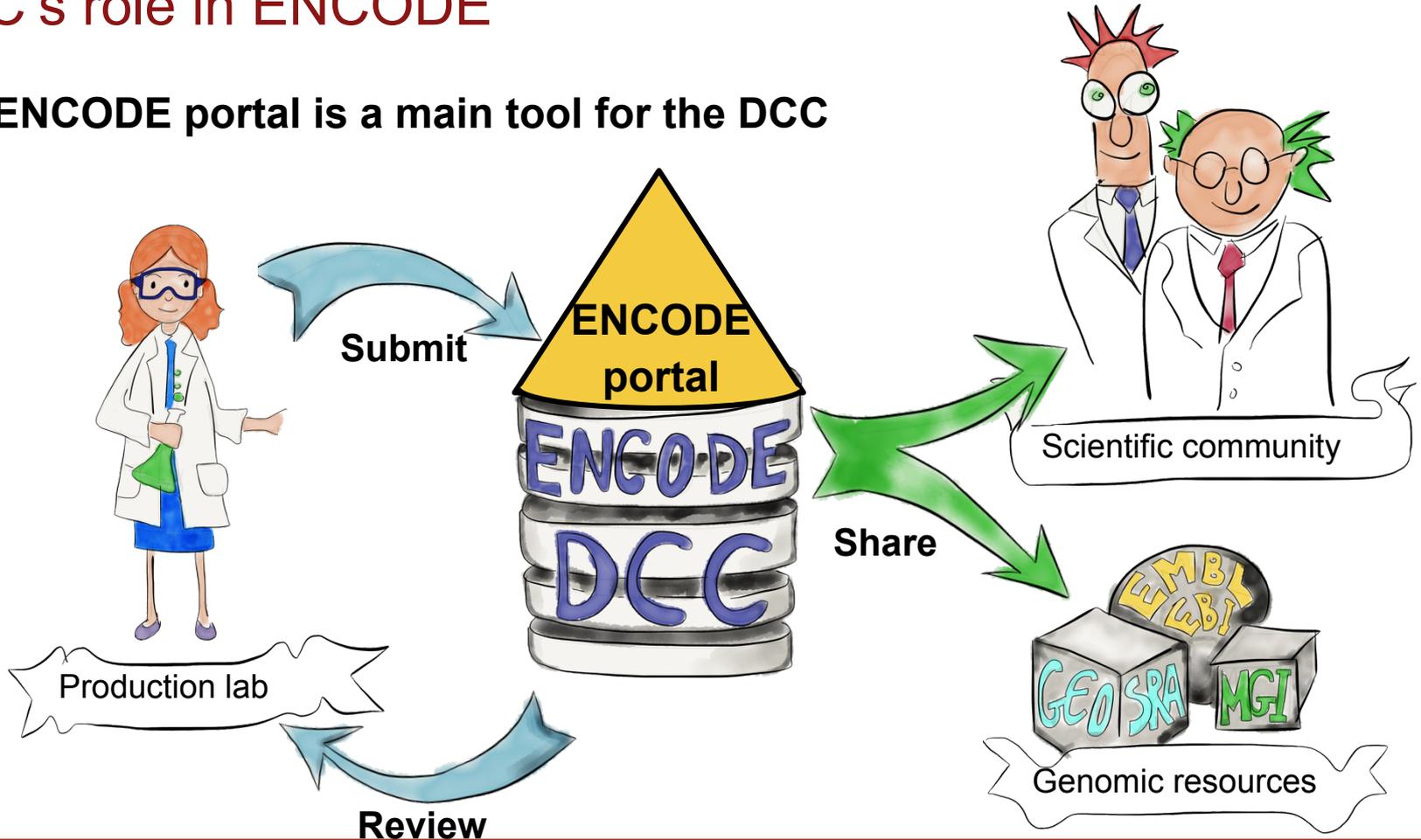
NIH U24 HG009397

# DCC's role in ENCODE



# DCC's role in ENCODE

The ENCODE portal is a main tool for the DCC

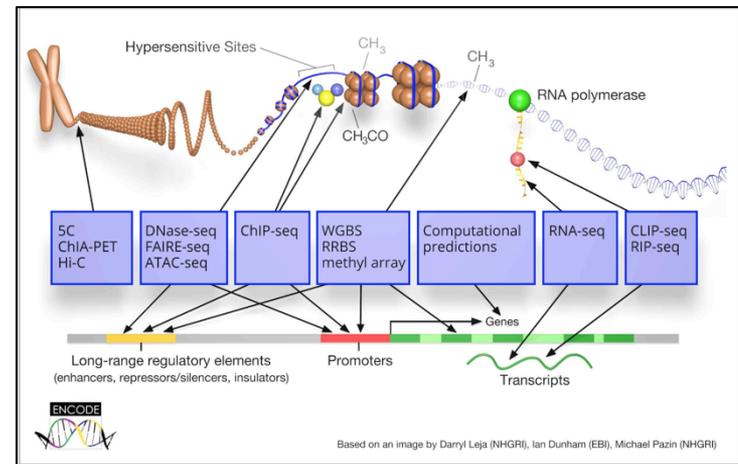


# encodeproject.org

15K experiments

46 assay “flavors”

600TB data files



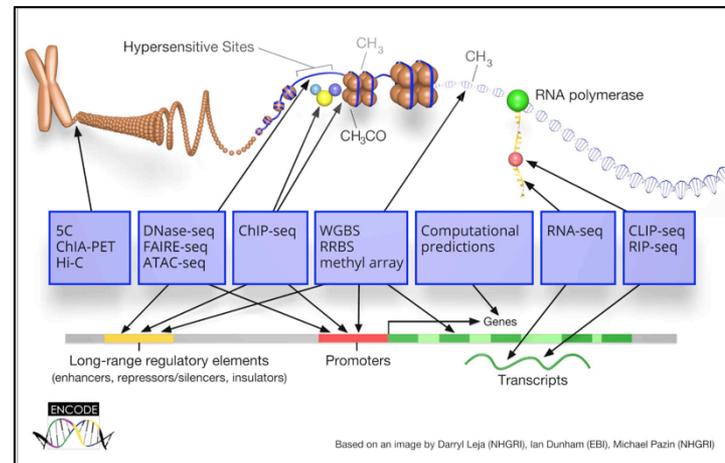
# encodeproject.org

15K experiments

46 assay “flavors”

600TB data files

Primary organizational unit  
on ENCODE portal



# Tutorial Outline

**An ENCODE experiment**

**Browse & Search experiments**

**Download & Visualize files from many experiments**

**Access the ENCODE portal programmatically**

**Data Collections**



ENCODE Data Encyclopedias Materials & Methods Help

Experiments / CHIP-seq / Homo sapiens / ascending aorta

### Experiment summary for ENCSR982QIF

#### Summary

**Status:** ● released

**Assay:** [ChIP-seq](#)

**Target:** [H3K27ac](#)

**Biosample summary:** [Homo sapiens ascending aorta female adult \(51 year\)](#)

**Biosample Type:** [tissue](#)

**Replication type:** [unreplicated](#)

**Nucleic acid type:** [DNA](#)

**Size range:** [200-600](#)

**Fragmentation method:** [sonication \(generic\)](#)

**Strand specificity:** [Non-strand-specific](#)

**Platform:** [Illumina HiSeq 2500](#)

**Controls:** [ENCSR494YJW](#)

Attribution

Lab:

Award:

Project:

External resource:

Aliases:

Date submitted:

Date released:

Related datasets:

Tags:

#### Isogenic replicates

Isogenic replicate	Technical replicate	Summary
1	1	<a href="#">female adult (51 year) ascending aorta tissue</a>

#### Files

[GRCh38](#)

Association graph File details

### Summary

<b>Status:</b>	<span style="background-color: #2e7d32; color: white; padding: 2px;">● released</span>
<b>Assay:</b>	<a href="#">ChIP-seq</a>
<b>Target:</b>	<a href="#">H3K27ac</a>
<b>Biosample summary:</b>	<a href="#">Homo sapiens ascending aorta female adult (51 year)</a>
<b>Biosample Type:</b>	<a href="#">tissue</a>
<b>Replication type:</b>	<a href="#">unreplicated</a>
<b>Nucleic acid type:</b>	<a href="#">DNA</a>
<b>Size range:</b>	<a href="#">200-600</a>
<b>Fragmentation method:</b>	<a href="#">sonication (generic)</a>
<b>Strand specificity:</b>	<a href="#">Non-strand-specific</a>
<b>Platform:</b>	<a href="#">Illumina HiSeq 2500</a>
<b>Controls:</b>	<a href="#">ENCSR494YJW</a>

## Attribution

**Lab:** Bradley Bernstein, Broad

**Award:** [U54HG006991](#) (Bradley Bernstein, Broad)

**Project:** ENCODE

**External resources:** [GEO:GSE101384](#) ↗

**Aliases:** bradley-bernstein:Project Element 2547

**Date submitted:** November 13, 2016

**Date released:** June 8, 2017

**Related datasets:** [ENCSR818FUR](#) ⓘ

## Tags:



## Attribution

**Lab:** Bradley Bernstein, Broad

**Award:** [U54HG006991](#) (Bradley Bernstein, Broad)

**Project:** ENCODE

**External resources:** [GEO:GSE101384](#) ↗

**Aliases:** bradley-bernstein:Project Element 2547

**Date submitted:** November 13, 2016

**Date released:** June 8, 2017

**Related datasets:** [ENCSR818FUR](#) ⓘ

**Tags:**  
 GTEx

ENCODE Data Encyclopedias Materials & Methods Help

Search

### Isogenic replicates

Isogenic replicate ^	Technical replicate v	Summary	Biosample v	Antibody v	Library v
1	1	female adult (51 year) ascending aorta tissue	<a href="#">ENCBS273PSC</a>	<a href="#">ENCAB000AQN</a>	<a href="#">ENCLB626HJU</a>

Replication type: unreplicated  
Nucleic acid type: DNA  
Size range: 200-600  
Fragmentation method: sonication (generic)  
Strand specificity: Non-strand-specific  
Platform: Illumina HiSeq 2500  
Controls: [ENCSR484YJW](#)

Date released: June 8, 2017  
Related datasets: [ENCSR818FJR](#)  
Tags:

### Isogenic replicates

Isogenic replicate ^	Technical replicate v	Summary	Biosample v	Antibody v	Library v
1	1	female adult (51 year) ascending aorta tissue	<a href="#">ENCBS273PSC</a>	<a href="#">ENCAB000AQN</a>	<a href="#">ENCLB626HJU</a>

Files

GRCh38 UCSC Visualize

Association graph File details Include deprecated files

Replicate 1

ENCBS273PSC (unreplicated)

ENCAB000AQN (non-strand-specific)

ENCLB626HJU (unreplicated)

# An ENCODE experiment

<https://www.encodeproject.org/ENCSR982QIF>

# Experimental entities

Isogenic replicate ^	Technical replicate v	Summary	Biosample v	Antibody v	Library v
1	1	female adult (51 year) ascending aorta tissue	<a href="#">ENCBS273PSC</a>	<a href="#">ENCAB000AQN</a>	ENCLB626HJU

## Biosample

product info, treatments,  
genetic modifications,  
donor age & sex, ...  
all Experiments on that donor

## Antibody

source, product & lot IDs,  
characterizations

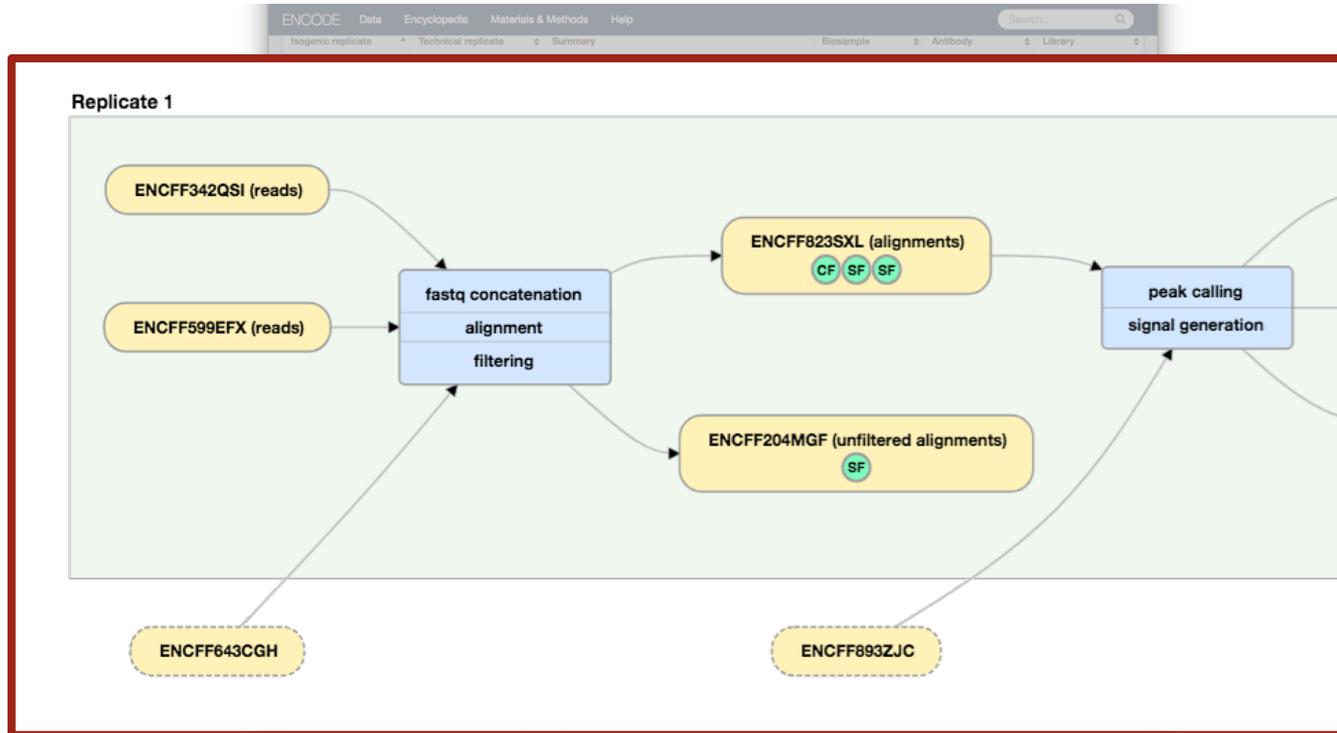
The screenshot displays the ENCODE project website interface. At the top, there is a navigation bar with 'ENCODE' and links for 'Data', 'Encyclopedia', 'Materials & Methods', and 'Help'. A search bar is located on the right. Below the navigation bar, a table provides experiment details: '1' for isogenic replicate, '1' for technical replicate, 'female adult (51 year) ascending aorta tissue' for the sample, 'ENCBS273PSC' for the biosample, 'ENCAB000AQN' for the antibody, and 'ENCLB628HJU' for the library.

The 'Files' section is active, showing 'GRCh38' as the reference genome. Below this, there are tabs for 'Association graph' and 'File details'. The 'Association graph' is selected, displaying a flowchart of the experimental pipeline for 'Replicate 1'. The graph starts with 'ENCFFA089 (reads)' and 'ENCFFA090 (reads)', leading to 'Map/normalization/alignment filtering'. This step branches into 'ENCFFA089 (reads)' and 'ENCFFA090 (reads)'. The next step is 'ENCFFA089 (reads)' and 'ENCFFA090 (reads)', which leads to 'ENCFFA089 (reads)' and 'ENCFFA090 (reads)'. The final step is 'ENCFFA089 (reads)' and 'ENCFFA090 (reads)'. A 'Download Graph' button is located below the graph.

The 'Documents' section contains a grid of document cards:

- High Resolution Pathology Slide Image**: Description excerpt: High-resolution whole slide digital images of pathological specimens (SVS...). File: 84258.svs.
- Pipeline Protocol**: Description: CHIP mapping pipeline: Includes overview and references for the pipeline. File: CHIP-seq\_Mapping\_Pipeline\_Overview.pdf.
- General Protocol**: Description: PRC Case Summary Report For Case 4. File: ENC\_Case-4\_DEJ\_PRCcsr\_Redacted.pdf.
- General Protocol**: Description: GTEx ENCODE Tissue Recovery Form II, Case 4. File: ENC\_Case-4\_DEJ\_TRF\_revised.pdf.
- Data Sheet**: Description: Mapping of high resolution images (SVS format) ids to tissues. File: Encode Public IDs.pdf.
- General Protocol**: Description: Epigenomics Alternative Mag Bead ChIP Protocol v1.1 exp. File: Epigenomics\_Alternative\_Mag\_Bead\_CHIP\_Protocol.pdf.
- Extraction Protocol**: Description: GTEx Tissue Harvesting Work Instruction. File: GTEx Tissue Harvesting Work Instruction.pdf.

The footer contains the ENCODE logo, 'Stanford University', and links for 'Citing ENCODE', 'Privacy', 'Contact', and 'Submitter sign-in'.



The screenshot displays the ENCODE Data Portal interface. At the top, navigation tabs include 'Data', 'Encyclopedia', 'Materials & Methods', and 'Help'. Below this, a breadcrumb trail shows 'Isogenic replicate' > 'Technical replicate' > 'Summary'. The main content area is titled 'Replicate 1' and contains a workflow diagram. The diagram shows a sequence of steps: 'ENCF342QSI (reads)' leads to 'ENCF823SXL (alignments)', which then leads to 'peak calling signal generation'. A second step, 'ENCF204MGF (unfiltered alignments)', also feeds into the 'peak calling signal generation' step. A third step, 'ENCF893ZJC', is shown in a dashed box and also feeds into the 'peak calling signal generation' step. The 'ENCF823SXL (alignments)' step is highlighted with a red border and contains three circular icons labeled 'CF', 'SF', and 'SF'. The 'ENCF204MGF (unfiltered alignments)' step contains one circular icon labeled 'SF'. The 'ENCF893ZJC' step is enclosed in a dashed-line box. A modal window titled 'bam ENCF823SXL' is overlaid on the left side of the screen, providing detailed information about the file.

**Replicate 1**

ENCF342QSI (reads)

ENCF823SXL (alignments)  
CF SF SF

peak calling  
signal generation

ENCF204MGF (unfiltered alignments)  
SF

ENCF893ZJC

**bam ENCF823SXL**

Status: released

Output: alignments

Biological replicate(s): [1]

Technical replicate(s): [1\_1]

Mapped read length: 76

Mapping assembly: GRCh38

Lab: ENCODE Processing Pipeline

Date added: 2016-11-16

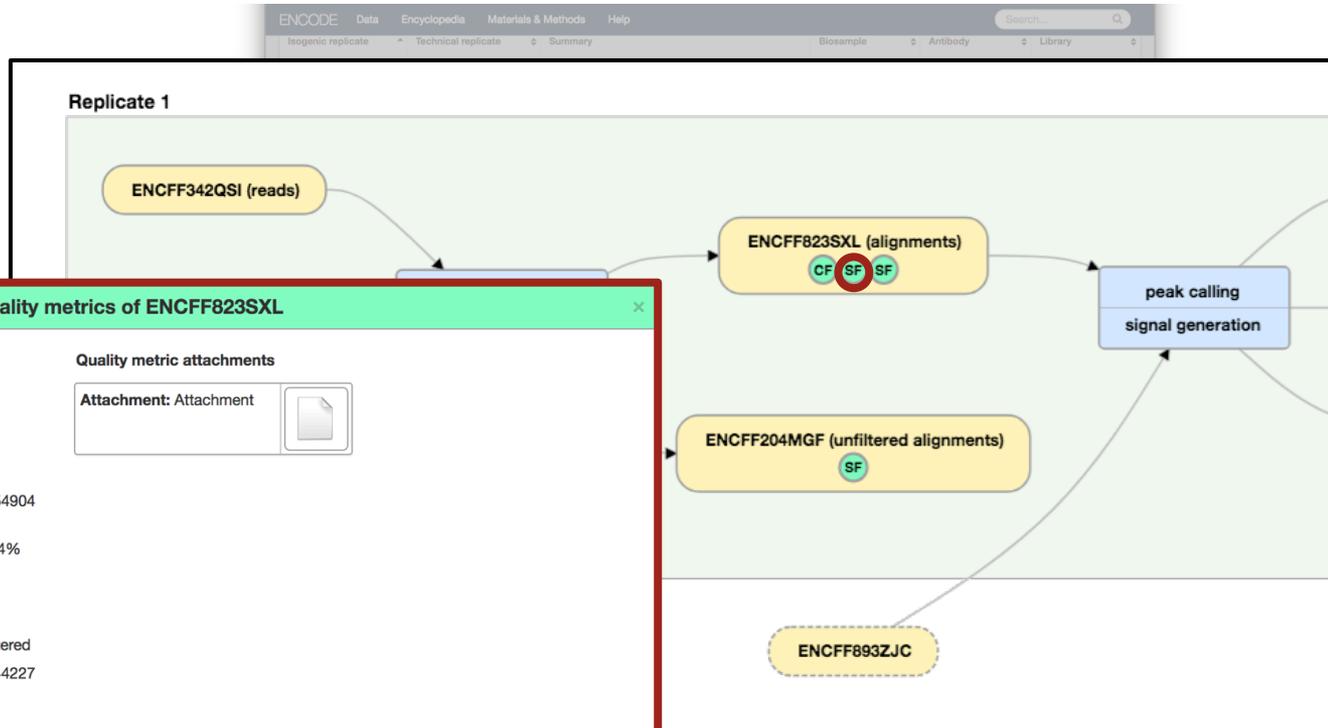
Software: bwa 0.7.10 samtools 1.0 picard 1.92 bedtools 2.17.0 phantompeakqualtools 1.1

File size: 1.96 GB

File download: [ENCF823SXL](#)

File quality metrics: Chipseq filter quality metrics Samtools flagstats quality metrics Samtools flagstats quality metrics

Close



**Samtools flagstats quality metrics of ENCF823SXL**

# reads with duplicates passing QC:	0
# reads with duplicates failing QC:	0
# reads mapped passing QC:	48354904
% reads mapped passing QC:	96.24%
# reads mapped failing QC:	0
Processing stage:	unfiltered
# of total reads passing QC:	50244227
# of total reads failing QC:	0

Quality metric attachments

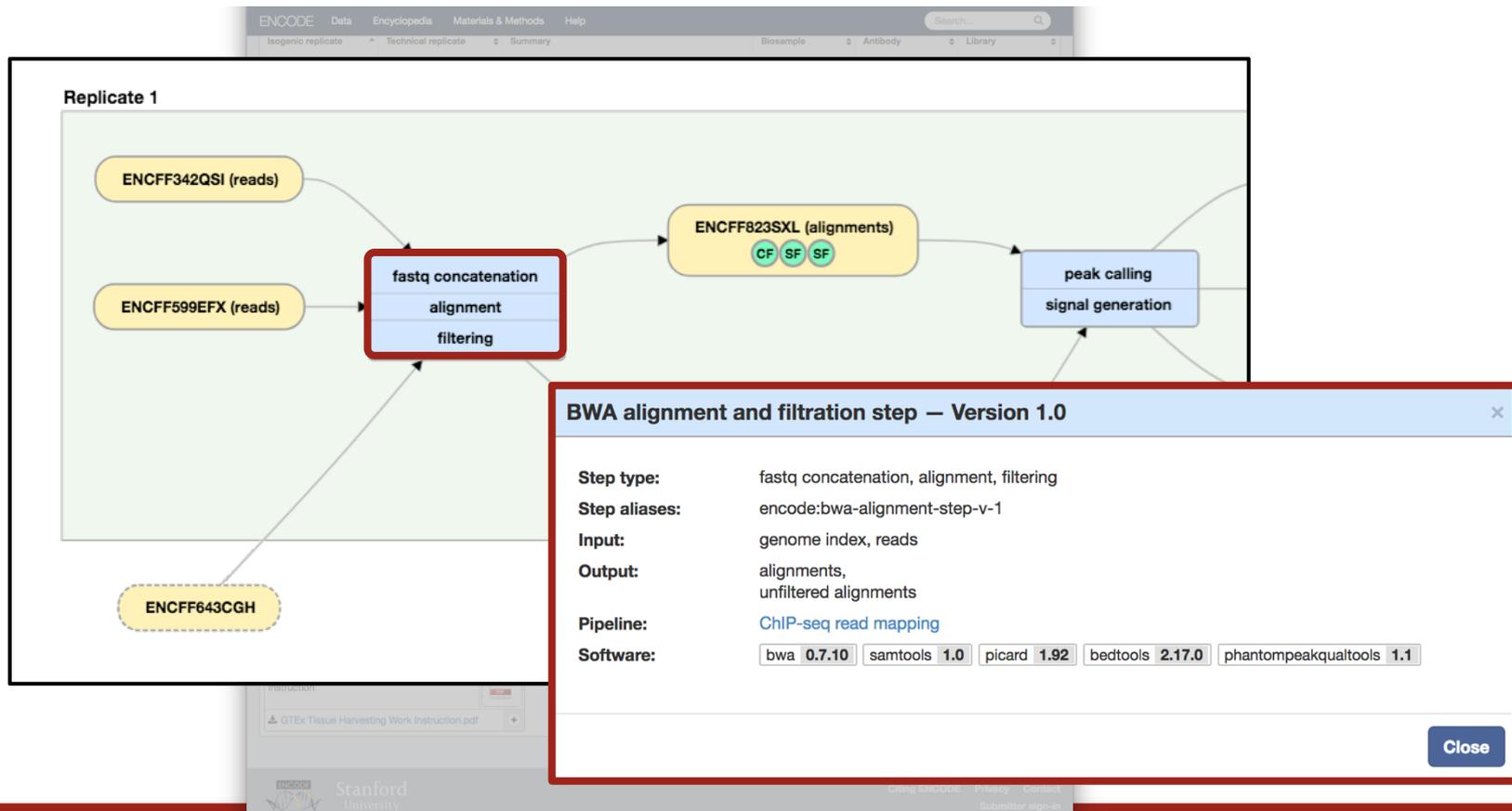
Attachment: Attachment

Close

# An ENCODE experiment

<https://www.encodeproject.org/ENCSR982QIF>

# Pipeline step metadata



The screenshot shows the ENCODE Data Portal interface. At the top, there are navigation tabs: ENCODE, Data, Encyclopedia, Materials & Methods, and Help. A search bar is on the right. Below the navigation, the 'Files' section is active, showing 'GRCh38' as the genome build and 'UCSC' as the track. The 'File details' tab is selected, indicated by a red arrow. The page displays 'Raw sequencing data' and 'Processed data' tables.

**Raw sequencing data**

Isogenic replicate	Library	Accession	File type	Run type	Read	Lab	Date added	File size	Audit status	File status
1	ENCLB626HJU	ENCF599EFX	fastq	SE76nt		Bradley Bernstein, Broad	2016-11-13	1.34 GB		released
1	ENCLB626HJU	ENCF342QSI	fastq	SE76nt		Bradley Bernstein, Broad	2016-11-13	1.33 GB		released

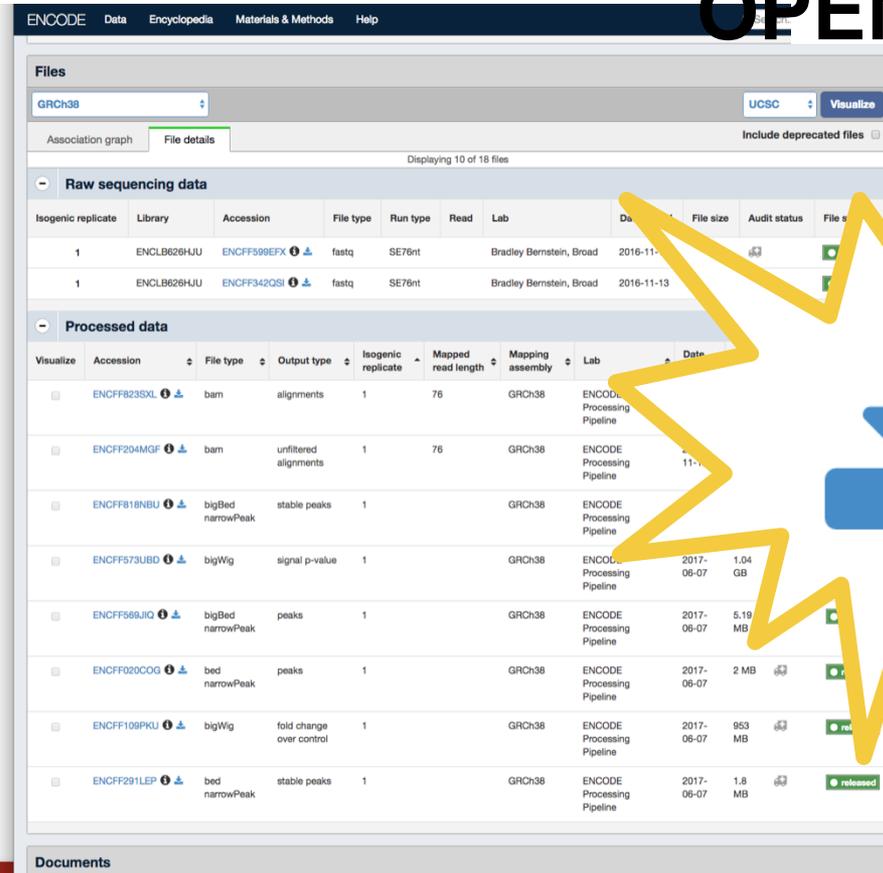
**Processed data**

Visualize	Accession	File type	Output type	Isogenic replicate	Mapped read length	Mapping assembly	Lab	Date added	File size	Audit status	File status
<input type="checkbox"/>	ENCF823SXL	bam	alignments	1	76	GRCh38	ENCODE Processing Pipeline	2016-11-16	1.96 GB		released
<input type="checkbox"/>	ENCF204MGF	bam	unfiltered alignments	1	76	GRCh38	ENCODE Processing Pipeline	2016-11-16	2.33 GB		released
<input type="checkbox"/>	ENCF818NBU	bigBed narrowPeak	stable peaks	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	3.39 MB		released
<input type="checkbox"/>	ENCF573UBD	bigWig	signal p-value	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	1.04 GB		released
<input type="checkbox"/>	ENCF569JIG	bigBed narrowPeak	peaks	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	5.19 MB		released
<input type="checkbox"/>	ENCF020COG	bed narrowPeak	peaks	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	2 MB		released
<input type="checkbox"/>	ENCF109PKU	bigWig	fold change over control	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	953 MB		released
<input type="checkbox"/>	ENCF291LEP	bed narrowPeak	stable peaks	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	1.8 MB		released

# An ENCODE experiment

<https://www.encodeproject.org/ENCSR982QIF>

# ENCODE IS ALL OPEN ACCESS!!!!

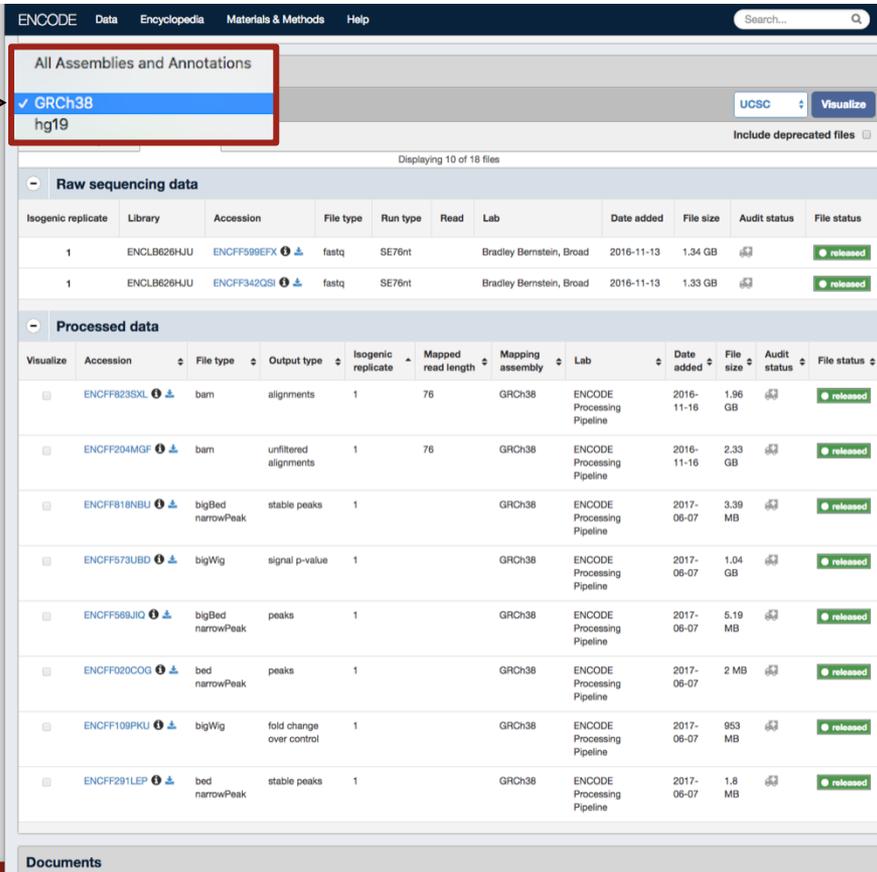


The screenshot displays the ENCODE Data Portal interface for experiment ENCSR982QIF. The 'Raw sequencing data' table shows two entries:

Isogenic replicate	Library	Accession	File type	Run type	Read	Lab	Date	File size	Audit status	File status
1	ENCLB626HJU	ENCFF599EFX	fastq	SE76nt		Bradley Bernstein, Broad	2016-11-13			
1	ENCLB626HJU	ENCFF342QSI	fastq	SE76nt		Bradley Bernstein, Broad	2016-11-13			

The 'Processed data' table shows various data products:

Visualize	Accession	File type	Output type	Isogenic replicate	Mapped read length	Mapping assembly	Lab	Date	File size	Audit status	File status
<input type="checkbox"/>	ENCFF823SXL	bam	alignments	1	76	GRCh38	ENCODE Processing Pipeline				
<input type="checkbox"/>	ENCFF204MGF	bam	unfiltered alignments	1	76	GRCh38	ENCODE Processing Pipeline	2017-06-07			
<input type="checkbox"/>	ENCFF818NBU	bigBed narrowPeak	stable peaks	1		GRCh38	ENCODE Processing Pipeline				
<input type="checkbox"/>	ENCFF573UBD	bigWig	signal p-value	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	1.04 GB		
<input type="checkbox"/>	ENCFF569JIG	bigBed narrowPeak	peaks	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	5.19 MB		
<input type="checkbox"/>	ENCFF020COG	bed narrowPeak	peaks	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	2 MB		
<input type="checkbox"/>	ENCFF109PKU	bigWig	fold change over control	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	953 MB		
<input type="checkbox"/>	ENCFF291LEP	bed narrowPeak	stable peaks	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	1.8 MB		released



ENCODE Data Encyclopedia Materials & Methods Help

All Assemblies and Annotations

- ✓ GRCh38
- hg19

UCSC Visualize

Include deprecated files

Displaying 10 of 18 files

### Raw sequencing data

Isogenic replicate	Library	Accession	File type	Run type	Read	Lab	Date added	File size	Audit status	File status
1	ENCLB626HJU	ENCF599EFX	fastq	SE76nt		Bradley Bernstein, Broad	2016-11-13	1.34 GB	released	released
1	ENCLB626HJU	ENCF342QSI	fastq	SE76nt		Bradley Bernstein, Broad	2016-11-13	1.33 GB	released	released

### Processed data

Visualize	Accession	File type	Output type	Isogenic replicate	Mapped read length	Mapping assembly	Lab	Date added	File size	Audit status	File status
<input type="checkbox"/>	ENCF823SXL	bam	alignments	1	76	GRCh38	ENCODE Processing Pipeline	2016-11-16	1.96 GB	released	released
<input type="checkbox"/>	ENCF204MGF	bam	unfiltered alignments	1	76	GRCh38	ENCODE Processing Pipeline	2016-11-16	2.33 GB	released	released
<input type="checkbox"/>	ENCF818NBU	bigBed narrowPeak	stable peaks	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	3.39 MB	released	released
<input type="checkbox"/>	ENCF573UBD	bigWig	signal p-value	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	1.04 GB	released	released
<input type="checkbox"/>	ENCF569JIG	bigBed narrowPeak	peaks	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	5.19 MB	released	released
<input type="checkbox"/>	ENCF020COG	bed narrowPeak	peaks	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	2 MB	released	released
<input type="checkbox"/>	ENCF109PKU	bigWig	fold change over control	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	953 MB	released	released
<input type="checkbox"/>	ENCF291LEP	bed narrowPeak	stable peaks	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	1.8 MB	released	released

Documents

The screenshot shows the ENCODE Data Portal interface. At the top, there are navigation tabs: ENCODE, Data, Encyclopedia, Materials & Methods, and Help. A search bar is located on the right. Below the navigation, the 'Files' section is active, showing a dropdown menu for 'GRCh38'. A red box highlights the 'Visualize' button, which has opened a menu with options: 'UCSC' (checked), 'Quick View', and 'Ensembl'. Below this, there are two main data sections: 'Raw sequencing data' and 'Processed data'. The 'Raw sequencing data' table has columns for Isogenic replicate, Library, Accession, File type, Run type, Read, Lab, Date added, File size, Audit status, and File status. The 'Processed data' table has columns for Visualize, Accession, File type, Output type, Isogenic replicate, Mapped read length, Mapping assembly, Lab, Date added, File size, Audit status, and File status. The 'Documents' section is partially visible at the bottom.

Isogenic replicate	Library	Accession	File type	Run type	Read	Lab	Date added	File size	Audit status	File status
1	ENCLB626HJU	ENCF599EFX	fastq	SE76nt		Bradley Bernstein, Broad	2016-11-13	1.34 GB	released	released
1	ENCLB626HJU	ENCF342QSI	fastq	SE76nt		Bradley Bernstein, Broad	2016-11-13	1.33 GB	released	released

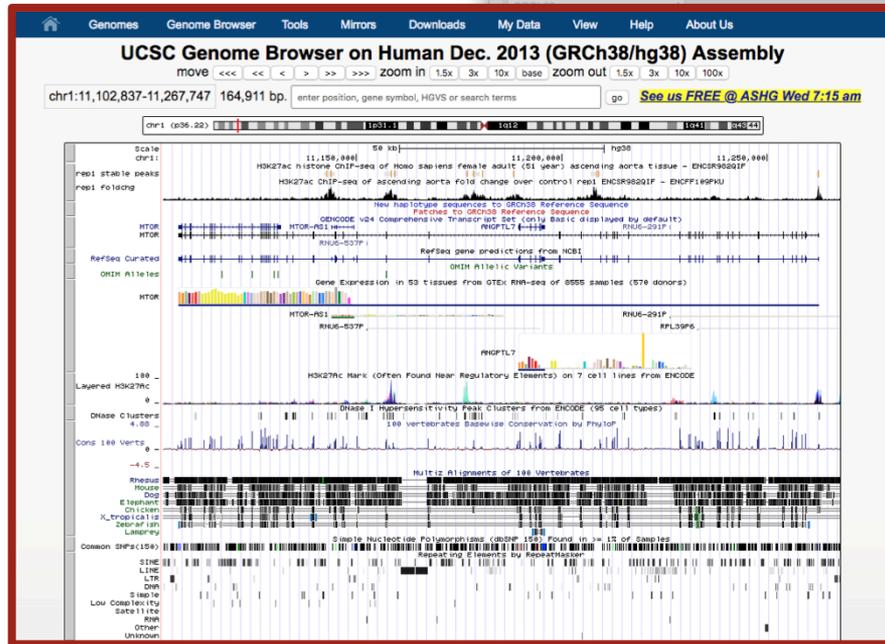
  

Visualize	Accession	File type	Output type	Isogenic replicate	Mapped read length	Mapping assembly	Lab	Date added	File size	Audit status	File status
<input type="checkbox"/>	ENCF823SXL	bam	alignments	1	76	GRCh38	ENCODE Processing Pipeline	2016-11-16	1.96 GB	released	released
<input type="checkbox"/>	ENCF204MGF	bam	unfiltered alignments	1	76	GRCh38	ENCODE Processing Pipeline	2016-11-16	2.33 GB	released	released
<input type="checkbox"/>	ENCF818NBU	bigBed narrowPeak	stable peaks	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	3.39 MB	released	released
<input type="checkbox"/>	ENCF573UBD	bigWig	signal p-value	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	1.04 GB	released	released
<input type="checkbox"/>	ENCF569JIG	bigBed narrowPeak	peaks	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	5.19 MB	released	released
<input type="checkbox"/>	ENCF020COG	bed narrowPeak	peaks	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	2 MB	released	released
<input type="checkbox"/>	ENCF109PKU	bigWig	fold change over control	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	953 MB	released	released
<input type="checkbox"/>	ENCF291LEP	bed narrowPeak	stable peaks	1		GRCh38	ENCODE Processing Pipeline	2017-06-07	1.8 MB	released	released

# An ENCODE experiment

<https://www.encodeproject.org/ENCSR982QIF>

# Visualize at UCSC/Ensembl



ENCODE Data Encyclopedias Materials & Methods Help

Files

UCSC Quick View Ensembl

Read	Lab	Date added	File size	Audit status	File status
76	GRCh38	2016-11-16	1.96 GB	released	released
76	GRCh38	2016-11-16	2.33 GB	released	released
GRCh38	ENCODE Processing Pipeline	2017-06-07	3.39 MB	released	released
GRCh38	ENCODE Processing Pipeline	2017-06-07	1.04 GB	released	released
GRCh38	ENCODE Processing Pipeline	2017-06-07	5.19 MB	released	released
GRCh38	ENCODE Processing Pipeline	2017-06-07	2 MB	released	released
GRCh38	ENCODE Processing Pipeline	2017-06-07	853 MB	released	released
GRCh38	ENCODE Processing Pipeline	2017-06-07	1.8 MB	released	released

# An ENCODE experiment

<https://www.encodeproject.org/ENCSR982QIF>

# Quick View on ENCODE portal

The screenshot displays the ENCODE portal interface. On the left, a sidebar contains search filters for 'Data file search', 'Content category', 'Content type', 'File format', 'Specific file format type', 'File type', 'Project', and 'Mapping assembly'. The main area shows 'Showing 4 of 4 results' for a genomic track at coordinates 22:29,890,000..30,050,000. The track includes annotations for RINUG-331P, AC003681.1, RFP3-304A18.1, MIR6818, and HDORMAD2-AS1. A 'Quick View' button is highlighted in red. On the right, a table lists search results with columns for 'Date added', 'File size', 'Audit status', and 'File status'. A red box highlights the 'UCSC' and 'Quick View' options in the top right corner of the interface.

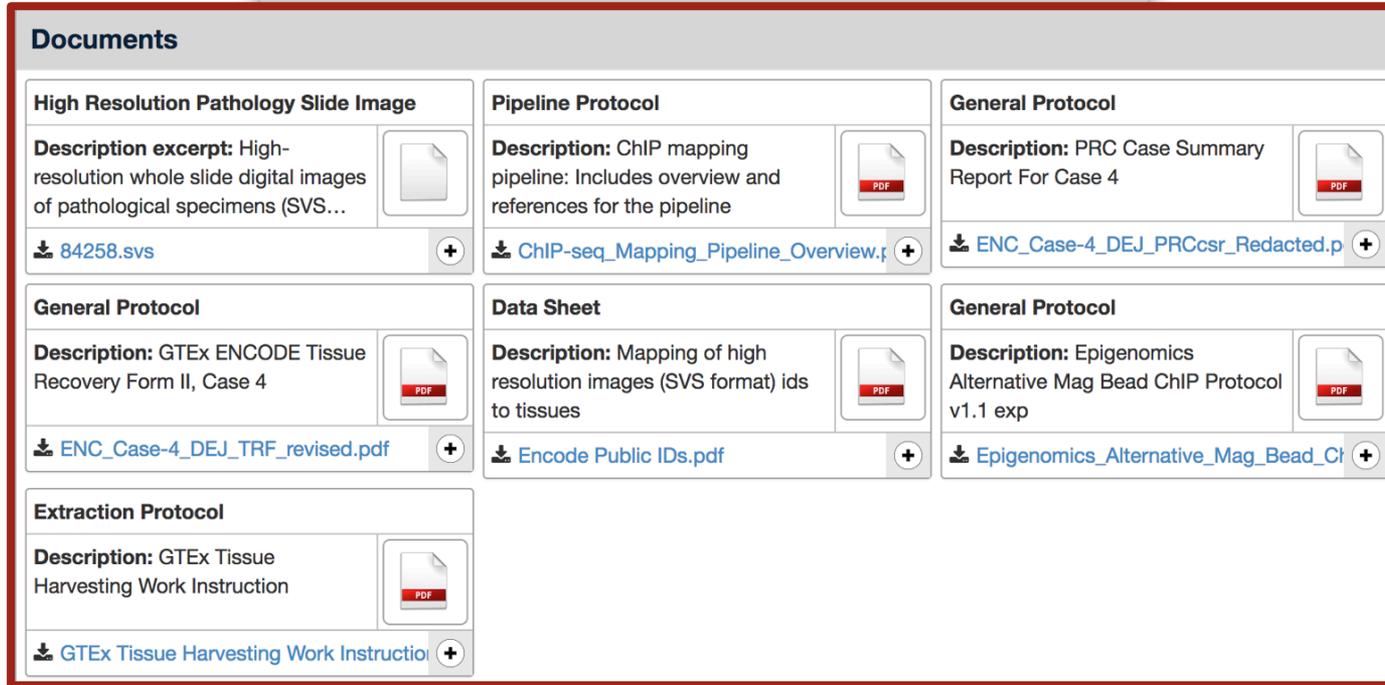
The screenshot displays the ENCODE data portal interface for experiment ENCSR982QIF. At the top, navigation tabs include 'Data', 'Encyclopedia', 'Materials & Methods', and 'Help'. Below this, a table lists experiment details: '1' for isogenic replicate, '1' for technical replicate, 'female adult (51 year) ascending aorta tissue' for the biosample, 'ENCBS273PSC' for the antibody, and 'ENCLB628HJU' for the library.

The 'Files' section shows 'GRCh38' as the reference genome, with 'UCSC' and 'Visualize' options. Below this, an 'Association graph' is displayed, showing a flowchart of the experimental pipeline. The graph starts with 'ENCFF889I (reads)' and 'ENCFF889J (reads)', leading to 'Map/normalization alignment filtering' and 'ENCFF889F (filtered alignment)'. These lead to 'peak calling signal generation', which then branches into 'ENCFF889D (signal values)', 'ENCFF889C (reads)', and 'ENCFF889E (peak calling)'. The 'ENCFF889D' path leads to 'PIL format conversion' and 'ENCFF889K (reads)'. The 'ENCFF889C' path leads to 'peaklist conversion' and 'ENCFF889L (reads peaklist)'. The 'ENCFF889E' path leads to 'PIL format conversion' and 'ENCFF889M (reads peaklist)'. A 'Download Graph' button is located below the graph.

The 'Documents' section contains a grid of document cards:

- High Resolution Pathology Slide Image**: Description excerpt: High-resolution whole slide digital images of pathological specimens (SVS...). File: 84258.svs.
- Pipeline Protocol**: Description: CHIP mapping pipeline. Includes overview and references for the pipeline. File: CHIP-seq\_Mapping\_Pipeline\_Overview.pdf.
- General Protocol**: Description: PRC Case Summary Report For Case 4. File: ENC\_Case-4\_DEJ\_PRCcsr\_Redacted.pdf.
- General Protocol**: Description: GTEx ENCODE Tissue Recovery Form II, Case 4. File: ENC\_Case-4\_DEJ\_TRF\_revised.pdf.
- Data Sheet**: Description: Mapping of high resolution images (SVS format) kds to tissues. File: Encode Public IDs.pdf.
- General Protocol**: Description: Epigenomics Alternative Mag Bead CHIP Protocol v1.1 exp. File: Epigenomics\_Alternative\_Mag\_Bead\_CHIP\_Protoc...
- Extraction Protocol**: Description: GTEx Tissue Harvesting Work Instruction. File: GTEx Tissue Harvesting Work Instruction.pdf.

The footer of the page includes the ENCODE logo, Stanford University logo, and links for 'Citing ENCODE', 'Privacy', 'Contact', and 'Submitter sign-in'.



The screenshot displays the ENCODE Data Portal interface for a specific experiment. At the top, a navigation bar includes 'ENCODE', 'Data', 'Encyclopedia', 'Materials & Methods', and 'Help'. Below this, a table lists experimental details such as 'Isogenic replicate', 'Technical replicate', and 'Summary' for a 'female adult (51 year) ascending aorta tissue' sample. The main content area is titled 'Documents' and contains a grid of document cards. Each card includes a title, a description excerpt, a PDF icon, and a download link with a plus sign for more options.

Document Title	Description Excerpt	Download Link
High Resolution Pathology Slide Image	Description excerpt: High-resolution whole slide digital images of pathological specimens (SVS...	<a href="#">84258.svs</a>
Pipeline Protocol	Description: ChIP mapping pipeline: Includes overview and references for the pipeline	<a href="#">ChIP-seq_Mapping_Pipeline_Overview.p...</a>
General Protocol	Description: PRC Case Summary Report For Case 4	<a href="#">ENC_Case-4_DEJ_PRCcsr_Redacted.p...</a>
General Protocol	Description: GTEx ENCODE Tissue Recovery Form II, Case 4	<a href="#">ENC_Case-4_DEJ_TRF_revised.pdf</a>
Data Sheet	Description: Mapping of high resolution images (SVS format) ids to tissues	<a href="#">Encode Public IDs.pdf</a>
General Protocol	Description: Epigenomics Alternative Mag Bead ChIP Protocol v1.1 exp	<a href="#">Epigenomics_Alternative_Mag_Bead_Cl...</a>
Extraction Protocol	Description: GTEx Tissue Harvesting Work Instruction	<a href="#">GTEx Tissue Harvesting Work Instructio...</a>

# An ENCODE experiment

<https://www.encodeproject.org/ENCSR982QIF>

ENCODE Data Encyclopedia Materials & Methods Help Search...

Experiments / [ChIP-seq](#) / [Homo sapiens](#) / [ascending aorta](#)

## Experiment summary for ENCSR982QIF

**Summary**

**Status:** ● released

**Assay:** ChIP-seq

**Target:** [H3K27ac](#)

**Biosample summary:** *Homo sapiens* ascending aorta female adult (51 year)

**Biosample Type:** tissue

**Replication type:** unreplicated

**Nucleic acid type:** DNA

**Size range:** 200-600

**Fragmentation method:** sonication (generic)

**Strand specificity:** Non-strand-specific

**Platform:** [Illumina HiSeq 2500](#)

**Controls:** [ENCSR494YJW](#)

**Attribution**

**Lab:** Bradley Bernstein, Broad

**Award:** [U54HG006991](#) (Bradley Bernstein, Broad)

**Project:** ENCODE

**External resources:** [GEO:GSE101384](#)

**Aliases:** bradley-bernstein:Project Element 2547

**Date submitted:** November 13, 2016

**Date released:** June 8, 2017

**Related datasets:** [ENCSR818FJR](#)

**Tags:**

**Isogenic replicates**

Isogenic replicate	Technical replicate	Summary	Biosample	Antibody	Library
1	1	female adult (51 year) ascending aorta tissue	<a href="#">ENCBS273PSC</a>	<a href="#">ENCA8000AQN</a>	<a href="#">ENCLB626HJU</a>

**Files** GRCh38 UCSC Visualize

Association graph File details Include deprecated files

**Documents**

**High Resolution Pathology Slide Image**

**Description excerpt:** High-resolution whole slide digital images of pathological specimens (SVS...)

[84258.svs](#)

**Pipeline Protocol**

**Description:** ChIP mapping pipeline. Includes overview and references for the pipeline

[ChIP-seq\\_Mapping\\_Pipeline\\_Overview.pdf](#)

**General Protocol**

**Description:** PRC Case Summary Report For Case 4

[ENC\\_Case-4\\_DEJ\\_PRCcr\\_Redacted.pdf](#)

**General Protocol**

**Description:** GTEX ENCODE Tissue Recovery Form II, Case 4

[ENC\\_Case-4\\_DEJ\\_TRF\\_revised.pdf](#)

**Data Sheet**

**Description:** Mapping of high resolution images (SVS format) Ids to tissues

[Encode Public Ids.pdf](#)

**General Protocol**

**Description:** Epigenomics Alternative Mag Bead ChIP Protocol v1.1 exp

[Epigenomics\\_Alternative\\_Mag\\_Bead\\_ChIP\\_Protocol](#)

**Extraction Protocol**

**Description:** GTEX Tissue Harvesting Work Instruction

[GTEX Tissue Harvesting Work Instruction.pdf](#)

# Browsing the ENCODE portal

ENCODE Data Encyclopedia Materials & Methods Help

Search...

## ENCODE: Encyclopedia of DNA Elements

[About ENCODE Project](#) [Getting Started](#) [Experiments](#)

Search ENCODE portal

ENCODE Q

[About ENCODE Encyclopedia](#) [Candidate Regulatory Elements](#)

Search for Candidate Regulatory Elements

Hosted by SCREEN

Human hg19 Q Mouse mm10 Q

Based on an image by Darryl Leja (NIHGR), Ian Dunham (EBI), Michael Pazin (NIHGR)

HUMAN MOUSE WORM FLY

Data Matrix

### Project

14659

- ENCODE
- Roadmap
- modERN
- modENCODE
- GGR
- community

### Biosample Type

14659

- cell line
- tissue
- whole organisms
- primary cell
- in vitro differentiated cells
- cell-free sample
- single cell

### Assay Categories

Bar chart showing the number of assays for various categories:

Assay Category	Approximate Count
DNA binding	8500
Transcription	3000
DNA accessibility	1500
RNA binding	1000
DNA methylation	800
Replication timing	600
Chromatin looping	400
3D chromatin structure	300
Proteomics	200
DNA sequencing	150
RNA structure	100

# Browsing the ENCODE portal

# Jumping in: matrix

The screenshot shows the ENCODE portal interface. At the top, there is a navigation bar with 'Data', 'Encyclopedia', 'Materials & Methods', and 'Help'. A red arrow points to the 'Data' menu, which is open, showing options: 'Matrix', 'Search', 'Summary', 'Search by region', 'Reference epigenomes', and 'Publications'. Below the navigation is a search bar and a 'Search ENCODE portal' button. The main content area features a diagram of a DNA segment with various annotations: 'Long-range regulatory elements (enhancers, repressors/silencers, insulators)', 'Promoters', and 'Transcripts'. Above the DNA, there are icons for 'SC ChIA-PET Hi-C', 'DNase-seq FAIRE-seq ATAC-seq', 'CHIP-seq', 'WGBS RRBS methyl array', 'Computational predictions', 'RNA-seq', and 'CLIP-seq RIP-seq'. A red arrow points to the 'Matrix' option in the 'Data' menu. Below the diagram, there are buttons for 'About ENCODE Project', 'Getting Started', and 'Experiments'. Further down, there are buttons for 'About ENCODE Encyclopedia' and 'Candidate Regulatory Elements'. At the bottom, there are buttons for 'Human hg19 Q.' and 'Mouse mm10 Q.'. The 'Data Matrix' section shows three charts: 'Project' (donut chart with 14659 total), 'Biosample Type' (donut chart with 14659 total), and 'Assay Categories' (bar chart showing the number of assays for each category).

**Project**

Category	Count
ENCODE	~10000
Roadmap	~2000
modERN	~1000
modENCODE	~1000
GGR	~500
community	~150

**Biosample Type**

Category	Count
cell line	~10000
tissue	~2000
whole organisms	~1000
primary cell	~1000
in vitro differentiated cells	~500
cell-free sample	~500
single cell	~150

**Assay Categories**

Assay Category	Count
DNA binding	~8500
Transcription	~3000
DNA accessibility	~2000
RNA binding	~1500
DNA methylation	~1000
Replication timing	~500
Chromatin looping	~500
3D chromatin structure	~500
Proteomics	~500
DNA sequencing	~500
RNA structure	~500

Filter by  
assay category  
& organism

Jump to  
filtered  
matrix

The screenshot shows the ENCODE portal interface. At the top, there's a navigation bar with 'ENCODE', 'Data', 'Encyclopedia', 'Materials & Methods', and 'Help'. A search bar is on the right. Below the navigation bar is a diagram titled 'ENCODE: Encyclopedia of DNA Elements' illustrating various genomic features like Hypersensitive Sites, CH<sub>3</sub>, CH<sub>3</sub>CO, RNA polymerase, and various assays: SC, ChIA-PET, Hi-C, DNase-seq, FAIRE-seq, ATAC-seq, CHIP-seq, WGBS, RRBS, methyl array, Computational predictions, RNA-seq, and CLIP-seq, RIP-seq. Below the diagram are buttons for 'About ENCODE Project', 'Getting Started', and 'Experiments'. A search bar for the ENCODE portal is present. Further down, there are buttons for 'About ENCODE Encyclopedia' and 'Candidate Regulatory Elements', along with a search bar for candidate regulatory elements. At the bottom of the main content area, there are buttons for 'Human hg19 Q' and 'Mouse mm10 Q'. Below this is a 'Data Matrix' section with three donut charts: 'Project' (14659), 'Biosample Type' (14659), and 'Assay Categories'. The 'Project' and 'Biosample Type' charts use a color-coded legend. The 'Assay Categories' chart is a bar graph showing the number of assays for each category.

ENCODE: Encyclopedia of DNA Elements

Filter by assay category & organism

Jump to filtered matrix

Project: 14659

Biosample Type: 14659

Assay Categories

Assay Category	Count (approx.)
DNA binding	8500
Transcription	3000
DNA accessibility	1500
RNA binding	1000
DNA methylation	800
Replication timing	600
Chromatin structure	400
3D chromatin structure	300
Proteomics	200
DNA sequencing	150
RNA structure	100

The screenshot shows the ENCODE portal homepage. At the top, there is a navigation bar with links for "ENCODE", "Data", "Encyclopedia", "Materials & Methods", and "Help", along with a search bar. Below the navigation bar is a large diagram titled "ENCODE: Encyclopedia of DNA Elements". The diagram illustrates various genomic features and the assays used to study them. On the left, "Hypersensitive Sites" are shown as loops of DNA. In the center, "CH<sub>3</sub>" and "CH<sub>3</sub>CO" methyl groups are shown on DNA. On the right, "RNA polymerase" is shown transcribing DNA into "Transcripts". Below the diagram, several assay categories are listed in boxes: "3C ChIA-PET Hi-C", "DNase-seq FAIRE-seq ATAC-seq", "CHIP-seq", "WGBS RRBS methyl array", "Computational predictions", "RNA-seq", and "CLIP-seq RIP-seq". Below these boxes, a genomic track shows "Long-range regulatory elements (enhancers, repressors/silencers, insulators)", "Promoters", and "Genes".

Below the diagram, there are several search and navigation options:

- Buttons for "About ENCODE Project", "Getting Started", and "Experiments".
- A search bar labeled "Search ENCODE portal" with a magnifying glass icon.
- A button labeled "ENCODE Q".
- Buttons for "About ENCODE Encyclopedia" and "Candidate Regulatory Elements".
- A search bar labeled "Search for Candidate Regulatory Elements" with a magnifying glass icon.
- A note "Hosted by SCREEN".
- Buttons for "Human hg19 Q" and "Mouse mm10 Q".

Below the search options, there are tabs for "HUMAN", "MOUSE", "WORM", and "FLY". A red arrow points to the "HUMAN" tab. Below the tabs is a "Filtered Data Matrix" button, with another red arrow pointing to it. Below the "Filtered Data Matrix" button are three charts:

- Project:** A donut chart showing the distribution of projects. The total count is 10217. The legend includes ENCODE (blue), Roadmap (orange), GGI (red), and community (grey).
- Biosample Type:** A donut chart showing the distribution of biosample types. The total count is 10217. The legend includes cell line (blue), tissue (orange), primary cell (red), and in vitro differentiated cells (purple).
- Assay Categories:** A bar chart showing the number of assays for various categories. The y-axis ranges from 0 to 6000. The x-axis categories include DNA binding, Transcription, DNA accessibility, RNA binding, DNA methylation, Replication timing, Genotyping, Polymorphisms, DNA sequencing, and RNA structure.

ENCODE Data Encyclopedia Materials & Methods Help

## Experiment matrix

Enter search terms to filter the experiments included in the matrix.

**ASSAY**

**10217 results**

Clear Filters

**BIOSAMPLE**

**Assay category**

- DNA binding 5733
- Transcription 2160
- DNA accessibility 902
- RNA binding 511
- DNA methylation 481

**Assay**

- ChIP-seq 5733
- DNase-seq 727
- shRNA RNA-seq 523
- eCLIP 459
- lola RNA-seq 392

**Experiment status**

Selected filters: [released](#)

- released** 10217
- archived 728
- revoked 241

**Project**

- ENCODE 7637
- Roadmap 2156
- GGR 418
- community 6

**Genome assembly**

- hg19 9036
- GRCv38 8853
- GRCv38-minimal 1

**Target category**

- narrow histone mark 895
- recombinant protein 393
- chromatin remodeler 182
- other post-translational modification 66
- other context 51

**Target of assay**

**cell line**

cell line	K562	569	10	268	245	19	3	11	7	10	1	2	1	2	6	1	9	77	9	50
HepG2	357	3	255	210	11	3	5	3	6	2	2	2	6	2	6	2	6	1		
A549	374	14				27	2		9	2	2	1	2	1	5	3				
GM12878	226	2				13	3	6	7	1	2	1	1	6	2	6	2			
HEK293	257						2		1	2				2						

+ See 199 more...

**tissue**

tissue	stomach	65	20	10	3	5	4	5	6	1	2	
adrenal gland	52	11	4	8	5	4	2	4	5	2	1	3
transverse colon	48	4		4	4	4	4	4	2		4	
sigmoid colon	59	2		4	4	4	4	2			4	
liver	73	2		3	1	2	1	2		1		

+ See 127 more...

**primary cell**

primary cell	foreskin keratinocyte	37	2	5	3	13	3	13	
endothelial cell of umbilical vein	35	2	5	1	1	2	1	6	5
common myeloid progenitor, CD34-positive	42	12	1				1		
keratinocyte	23	2	5	3	6		6	5	
B cell	35	5	2	1	1		1	1	

+ See 123 more...

**in vitro differentiated cells**

in vitro differentiated cells	dendritic cell	11		20		24
neural stem progenitor cell	32	1	2	1	2	
mesenchymal stem cell	29	1	2	1	1	
neural cell	24		1	3	1	1
trophoblast cell	27	1	1	1	1	

+ See 31 more...

See all biosamples

[Download](#) [Filter to 500 to visualize](#)

ENCODE Data Encyclopedia Materials & Methods Help

Experiment matrix

Enter search terms to filter the experiments included in the matrix.

Q Enter search term(s)

Assay category

- DNA binding 5733
- Transcription 2160
- DNA accessibility 902
- RNA binding 511
- DNA methylation 441

Assay

Q Search

- ChIP-seq 5733
- DNase-seq 727
- shRNA RNA-seq 523
- eCLIP 459
- polyA RNA-seq 392

Experiment status

Selected filters:  released

- released 10217
- archived 728
- revoked 241

Project

- ENCODE 7637
- Roadmap 2156
- GGR 418
- community 6

Genome assembly

- hg19 9036
- GRCv38 8853
- GRCv38-minimal 1

Target category

- enhancer 895
- narrow histone mark 895
- recombinant protein 353
- chromatin remodeller 182
- other post-translational modification 66
- other context 51

Target of assay

Q Search

10217 results

Clear Filters

cell line

cell line	K562	HepG2	A549	GM12878	HEK293
K562	669	3	14	2	2
HepG2	357	3	14	2	2
A549	374	14	27	2	2
GM12878	226	2	13	2	2
HEK293	257	2	2	2	2

+ See 199 more...

tissue

tissue	stomach	adrenal gland	transverse colon	sigmoid colon	liver
stomach	65 20	10 3	5 4	5 6	5 6
adrenal gland	52 11	4 8	5 4	2 4	5 2
transverse colon	48 4	4 4	4 4	4 2	4 2
sigmoid colon	59 2	4 4	4 4	4 2	4 2
liver	73 2	3 1	2 1	2 1	2 1

+ See 127 more...

primary cell

primary cell	foreskin keratinocyte	endothelial cell of umbilical vein	common myeloid progenitor, CD34-positive keratinocyte	B cell
foreskin keratinocyte	37 2	5 3	13 3	37 2
endothelial cell of umbilical vein	35 2	5 1	1 2	35 2
common myeloid progenitor, CD34-positive keratinocyte	42 12	1	1	42 12
B cell	23 2	5	3 6	23 2

+ See 31 more...

See all biosamples

Download Filter to 800 to visualize

BIOSAMPLES

## ASSAY

10217 results



Clear Filters

ChIP-seq  
 DNase-seq  
 shRNA RNA-seq  
 eCLIP  
 polyA RNA-seq  
 DNase RNA-seq  
 total RNA-seq  
 small RNA-seq  
 RNA microarray  
 RAMPAGE  
 genotyping array  
 WGBS

cell line	K562	HepG2	A549	GM12878	HEK293
K562	669	3	14	2	2
HepG2	357	3	14	2	2
A549	374	14	27	2	2
GM12878	226	2	13	2	2
HEK293	257	2	2	2	2

+ See 199 more...

tissue	stomach	adrenal gland	transverse colon	sigmoid colon	liver
stomach	65 20	10 3	5 4	5 6	5 6
adrenal gland	52 11	4 8	5 4	2 4	5 2
transverse colon	48 4	4 4	4 4	4 2	4 2
sigmoid colon	59 2	4 4	4 4	4 2	4 2
liver	73 2	3 1	2 1	2 1	2 1

+ See 127 more...

primary cell	foreskin keratinocyte	endothelial cell of umbilical vein	common myeloid progenitor, CD34-positive keratinocyte	B cell
foreskin keratinocyte	37 2	5 3	13 3	37 2
endothelial cell of umbilical vein	35 2	5 1	1 2	35 2
common myeloid progenitor, CD34-positive keratinocyte	42 12	1	1	42 12
B cell	23 2	5	3 6	23 2

+ See 31 more...

# Browsing the ENCODE portal

# Filter results using facets

The screenshot shows the ENCODE portal interface. At the top, there are navigation links for ENCODE, Data, Encyclopedias, Materials & Methods, and Help. A search bar is located in the top right corner. The main content area is titled "Experiment matrix" and displays a table of 10217 results. The table has columns for "cell line" and "ASSAY". The "cell line" column lists various cell lines such as K562, HepG2, A549, GM12878, and HEK293. The "ASSAY" column lists various assays such as ChIP-seq, DNase-seq, eCLIP, RNA-seq, and others. The table is filtered to show only "released" experiments. On the left side, there are several filter panels: "Assay category" (DNA binding: 5733, Transcription: 2160, DNA accessibility: 902, RNA binding: 511, DNA methylation: 481), "Assay" (ChIP-seq: 5733, DNase-seq: 727, shRNA RNA-seq: 523, eCLIP: 459, lola RNA-seq: 392), "Experiment status" (released: 10217, archived: 728, revoked: 241), "Project" (ENCODE: 7637, Roadmap: 2156, GGR: 418, community: 6), "Genome assembly" (hg19: 9036, GRCh38: 8853, GRCh38-minimal: 1), and "Target category" (narrow histone mark: 895, recombinant protein: 393, chromatin remodeler: 182, other post-translational modification: 66, other context: 51). At the bottom of the filter panels, there is a search bar for "Target of assay".

# Browsing the ENCODE portal

# Filter results using facets

DNA binding	8699
Transcription	3132
DNA accessibility	1100
RNA binding	699
DNA methylation	560

ChIP-seq	8699
DNase-seq	835
polyA RNA-seq	770
shRNA RNA-seq	523
total RNA-seq	496

transcription factor	3817
histone	3034
control	2599
broad histone mark	1690
DNA binding protein	1450

ENCODE Data Encyclopedias Materials & Methods Help

Experiment matrix

Enter search terms to filter the experiments included in the matrix.

10217 results

Clear Filters

ASSAY

cell line

K562	560	10	268	245	19	3	11	7	10	1	2	1	2	6	1	9	77	9	50
HepG2	37	3	258	210	11	3	5	3	6	2	2	2	6	2	6	1			
A549	27	14																	
GM12878	226	2																	
HEK293	297																		

+ See 199 more...

tissue

stomach	65	20																	
adrenal gland	52	11																	
transverse colon	48	4																	
sigmoid colon	59	2																	
liver	73	2																	

+ See 127 more...

primary cell

foreskin keratinocyte	37	2																	
endothelial cell of umbilical vein	35	2																	
common myeloid progenitor, CD34-positive	42	12																	
keratinocyte	23	2																	
B cell	35	5																	

+ See 123 more...

in vitro differentiated cells

dendritic cell	11																		
neural stem progenitor cell	32	1																	
mesenchymal stem cell	29	1																	
neural cell	24																		
trophoblast cell	27	1																	

+ See 31 more...

See all biosamples

Download Filter to 500 to visualize

# Browsing the ENCODE portal

# Filter results using facets

DNA binding	8699
Transcription	3132
DNA accessibility	1100
RNA binding	699
DNA methylation	560

ChIP-seq	8699
DNase-seq	835
polyA RNA-seq	770
shRNA RNA-seq	523
total RNA-seq	496

transcription factor	3817
histone	3034
control	2599
broad histone mark	1690
DNA binding protein	1450

The screenshot shows the ENCODE portal interface with a search for 'pol'. The main results table shows a heatmap of data points across various biological categories. Several facets are visible on the left side of the interface:

- Assay category:** DNA binding (5733), Transcription (2160), DNA accessibility (902), RNA binding (511), DNA methylation (481).
- Assay:** ChIP-seq (5733), DNase-seq (727), shRNA RNA-seq (523), eCLIP (459), total RNA-seq (392).
- Experiment status:** released (10217), archived (728), revoked (241).
- Project:** ENCODE (7637), Roadmap (2156), GGR (418), community (6).
- Genome assembly:** hg19 (9036), GRCh38 (8853), GRCh38-minimal (1).
- Target category:** narrow histone mark (895), recombinant protein (393), chromatin remodeler (182), other post-translational modification (66), other context (51).
- Target of assay:** POLR2A (122), POLR2AphosphoS5 (45), POLR2AphosphoS2 (6), POLR2G (5), POLR2G eCLIP mock input (2).

# Browsing the ENCODE portal

# Filter results using facets

Assay category	
DNA binding	8699
Transcription	3132
DNA accessibility	1100
RNA binding	699
DNA methylation	560

Assay	
<input type="text" value="Search"/>	
ChIP-seq	8699
DNase-seq	835
polyA RNA-seq	770
shRNA RNA-seq	523
total RNA-seq	496

Target category	
transcription factor	3817
histone	3034
control	2599
broad histone mark	1690
DNA binding protein	1450

The screenshot shows the ENCODE portal interface with several facets highlighted in red boxes. The main search bar contains the text 'pol'. The 'Assay category' facet shows DNA binding (8699), Transcription (3132), DNA accessibility (1100), RNA binding (699), and DNA methylation (560). The 'Assay' facet shows ChIP-seq (8699), DNase-seq (835), polyA RNA-seq (770), shRNA RNA-seq (523), and total RNA-seq (496). The 'Target category' facet shows transcription factor (3817), histone (3034), control (2599), broad histone mark (1690), and DNA binding protein (1450). The 'Target of assay' facet shows POLR2A (122), POLR2AphosphoS5 (45), POLR2AphosphoS2 (6), POLR2G (5), and POLR2G eCLIP mock input (2). The 'Organism' facet shows Homo sapiens (10217), Mus musculus (1896), Drosophila melanogaster (1422), Caenorhabditis elegans (957), and Drosophila pseudoobscura (4). The 'Biosample type' facet shows cell line (5637), tissue (2775), primary cell (1303), and in vitro differentiated cells (502). The 'Cell' facet shows cancer cell (3985), epithelial cell (2105), hematopoietic cell (656), stem cell (645), and leukocyte (510). The 'Organ' facet shows blood (2492), bodily fluid (2492), liver (1204), lung (902), and embryo (950). The main content area shows the 'Experiment matrix' with search terms and filters for 'released' (10217) and 'archived' (728).

Target of assay	
<input type="text" value="pol"/>	
POLR2A	122
POLR2AphosphoS5	45
POLR2AphosphoS2	6
POLR2G	5
POLR2G eCLIP mock input	2

Organism	
Selected filters: <span>✖</span> Homo sapiens	
<b>Homo sapiens</b>	<b>10217</b>
Mus musculus	1896
Drosophila melanogaster	1422
Caenorhabditis elegans	957
Drosophila pseudoobscura	4

Biosample type	
cell line	5637
tissue	2775
primary cell	1303
in vitro differentiated cells	502

Cell	
<input type="text" value="Search"/>	
cancer cell	3985
epithelial cell	2105
hematopoietic cell	656
stem cell	645
leukocyte	510

Organ	
<input type="text" value="Search"/>	
blood	2492
bodily fluid	2492
liver	1204
lung	902
embryo	950

# Browsing the ENCODE portal

# Filter results using facets

Assay category	
DNA binding	8699
Transcription	3132
DNA accessibility	1100
RNA binding	699
DNA methylation	560

Assay	
<input type="text" value="Search"/>	
ChIP-seq	8699
DNase-seq	835
polyA RNA-seq	770
shRNA RNA-seq	523
total RNA-seq	496

Target category	
transcription factor	3817
histone	3034
control	2599
broad histone mark	1690
DNA binding protein	1450

The screenshot shows the ENCODE portal interface with a search bar at the top containing 'pol'. The main content area displays a list of results for 'POLR2A' and other related terms. On the left side, there are several facets for filtering results, including Assay category, Assay, Target category, Assay category, Assay, Experiment status, Project, Genome assembly, Target category, and Target of assay. The Assay category facet shows DNA binding (8699), Transcription (3132), DNA accessibility (1100), RNA binding (699), and DNA methylation (560). The Assay facet shows ChIP-seq (8699), DNase-seq (835), polyA RNA-seq (770), shRNA RNA-seq (523), and total RNA-seq (496). The Target category facet shows transcription factor (3817), histone (3034), control (2599), broad histone mark (1690), and DNA binding protein (1450). The Assay category facet shows DNA binding (5733), Transcription (2160), DNA accessibility (902), RNA binding (511), and DNA methylation (481). The Assay facet shows ChIP-seq (5733), DNase-seq (727), shRNA RNA-seq (523), eCLIP (459), and total RNA-seq (392). The Experiment status facet shows released (10217), archived (728), and revoked (241). The Project facet shows ENCODE (7637), Roadmap (2156), GGR (418), and community (6). The Genome assembly facet shows hg19 (9036), GRCh38 (8853), and GRCh38-minimal (1). The Target category facet shows transcription factor (3817), histone (3034), control (2599), broad histone mark (1690), DNA binding protein (1450), narrow histone mark (985), recombinant protein (393), chromatin remodeler (182), other post-translational modification (66), and other context (51). The Target of assay facet shows POLR2A (122), POLR2AphosphoS5 (45), POLR2AphosphoS2 (6), POLR2G (5), and POLR2G eCLIP mock input (2).

Target of assay	
<input type="text" value="pol"/>	
POLR2A	122
POLR2AphosphoS5	45
POLR2AphosphoS2	6
POLR2G	5
POLR2G eCLIP mock input	2

Organism	
Selected filters: <span>✖</span> Homo sapiens	
<input type="text" value="Search"/>	
<i>Homo sapiens</i>	10217
<i>Mus musculus</i>	1896
<i>Drosophila melanogaster</i>	1422
<i>Caenorhabditis elegans</i>	957
<i>Drosophila pseudoobscura</i>	4

Biosample type	
cell line	5637
tissue	2775
primary cell	1303
in vitro differentiated cells	502

Cell	
<input type="text" value="Search"/>	
cancer cell	3985
epithelial cell	2105
hematopoietic cell	656
stem cell	645
leukocyte	510

Organ	
<input type="text" value="Search"/>	
blood	2492
bodily fluid	2492
liver	1204
lung	902
embryo	850

and many more...

# Browsing the ENCODE portal

# Filter results using facets

**Organism**

Selected filters: ✖ **Homo sapiens**

- Mus musculus 93
- Homo sapiens 79**

**Assay category**

Selected filters: ✖ **Transcription**

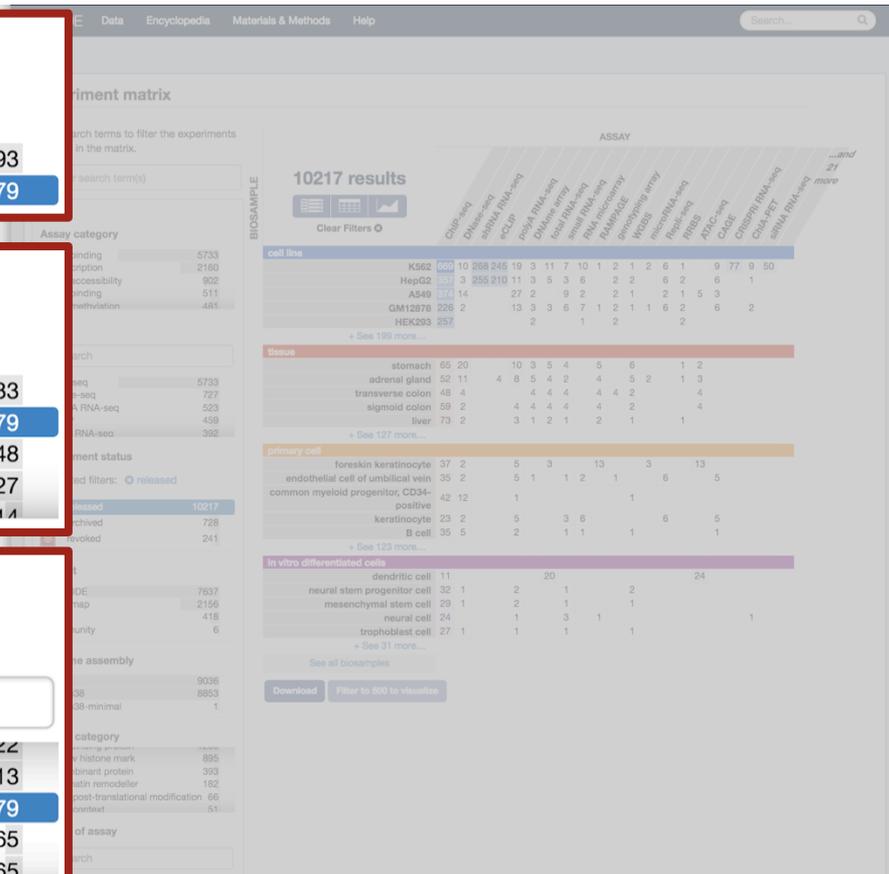
- DNA binding 233
- Transcription 79**
- DNA accessibility 48
- DNA methylation 27
- Coexisting 14

**Organ**

Selected filters: ✖ **brain**

Search

- connective tissue 122
- musculature of body 113
- brain 79**
- embryo 65
- vasculature 65



# Browsing the ENCODE portal

# Experiment matrix

## Organism

Selected filters: ✖ Homo sapiens

Mus musculus 93

Homo sapiens 79

## Assay category

Selected filters: ✖ Transcription

DNA binding 233

Transcription 79

DNA accessibility 48

DNA methylation 27

Coexisting 14

## Organ

Selected filters: ✖ brain

Q Search

connective tissue 122

musculature of body 113

brain 79

embryo 65

vasculature 65

The screenshot shows the ENCODE portal interface with the following elements:

- Navigation tabs: Data, Encyclopedia, Materials & Methods, Help
- Search bar: Search term(s)
- Assay category: Transcription
- Assay type: RNA-seq
- Assay status: released
- Selected filters: ✖ Homo sapiens
- 79 results summary with icons for list, table, and chart views.
- Download and Visualize buttons.
- Table of results with columns for cell line, tissue, and primary cell.

**ASSAY**

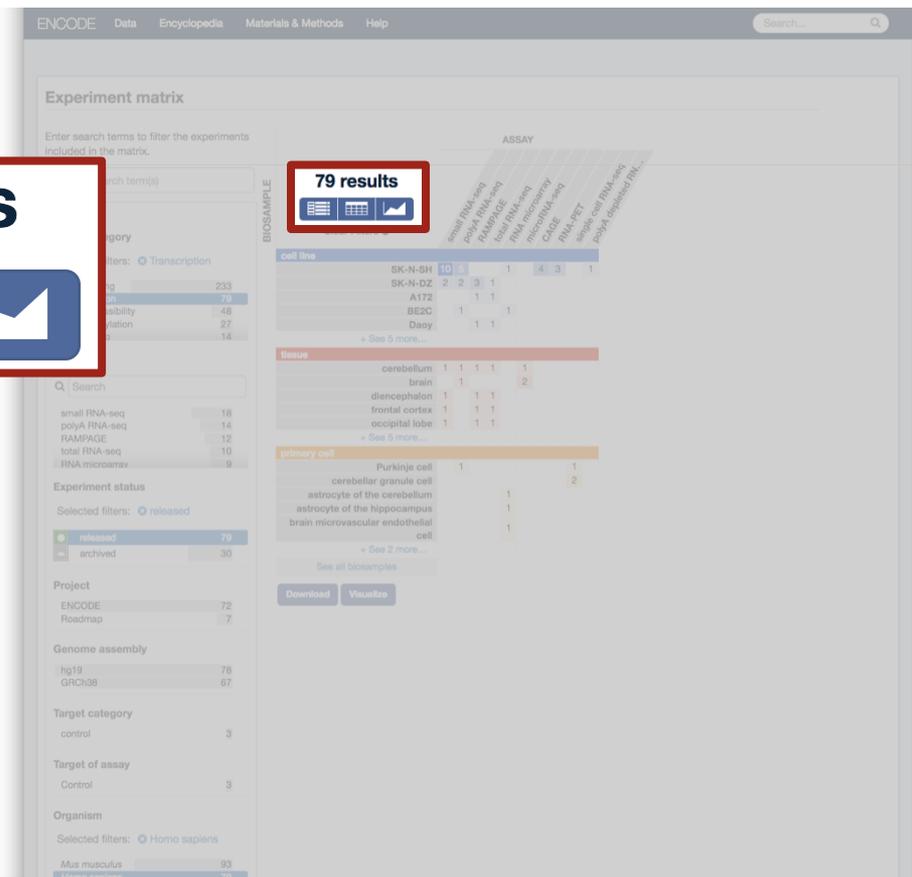
79 results

Clear Filters ✖

**BIOSAMPLE**

BIOSAMPLE	small RNA-seq	polyA RNA-seq	RAMPAGE	total RNA-seq	RNA microarray	microRNA-seq	CAGE	RNA-PET	single cell RNA-seq	polyA depleted RN...
<b>cell line</b>										
SK-N-SH	10	5		1			4	3		1
SK-N-DZ	2	2	3	1						
A172			1	1						
BE2C		1								
Daoy			1	1						
+ See 5 more...										
<b>tissue</b>										
cerebellum	1	1	1	1				1		
brain		1						2		
diencephalon	1		1	1						
frontal cortex	1		1	1						
occipital lobe	1		1	1						
+ See 5 more...										
<b>primary cell</b>										
Purkinje cell		1								1
cerebellar granule cell										2
astrocyte of the cerebellum						1				
astrocyte of the hippocampus						1				
brain microvascular endothelial cell						1				
+ See 2 more...										
See all biosamples										

**79 results**



ENCODE Data Encyclopedias Materials & Methods Help

Experiment matrix

Enter search terms to filter the experiments included in the matrix.

Search term(s)

Category

Filters:  Transcription

233

79

48

27

14

Search

small RNA-seq 16

polyA RNA-seq 14

RAMPAGE 12

total RNA-seq 10

RNA microarray 9

Experiment status

Selected filters:  released

released 79

archived 30

Project

ENCODE 72

Roadmap 7

Genome assembly

hg19 76

GRCh38 67

Target category

control 3

Target of assay

Control 3

Organism

Selected filters:  Homo sapiens

Mus musculus 93

79 results

BIOSAMPLE

ASSAY

cell line

SK-N-SH	10	5	1	4	3	1
SK-N-DZ	2	2	3	1		
A172			1	1		
BE2C			1		1	
Dazy			1	1		

+ See 5 more...

tissue

cerebellum	1	1	1	1	1
brain		1			2
diencephalon		1	1	1	
frontal cortex		1	1	1	
occipital lobe		1	1	1	

+ See 5 more...

primary cell

Purkinje cell		1			1
cerebellar granule cell					2
astrocytes of the cerebellum			1		
astrocytes of the hippocampus				1	
brain microvascular endothelial cell				1	

+ See 2 more...

See all biosamples

Download Visualize

# Browsing the ENCODE portal

# Search view

ENCODE Data Encyclopedia Materials & Methods Help Search...

### Experiment search

Clear Filters

**Assay category**

Selected filters: **Transcription**

DNA binding	233
<b>Transcription</b>	<b>79</b>
DNA accessibility	48
DNA methylation	27
Genetics	14

**Assay**

Q Search

small RNA-seq	18
polyA RNA-seq	14
RAMPAGE	12
total RNA-seq	10
RNA microarray	9

**Experiment status**

Selected filters: **released**

<b>released</b>	<b>79</b>
archived	30

**Project**

ENCODE	72
Roadmap	7

**RFA**

ENCODE3	36
ENCODE2	35
Roadmap	7
ENCODE2-Mouse	1

**Genome assembly**

hg19	78
GRCh38	67

**Target category**

control	3
---------	---

**Target of assay**

Control	3
---------	---

**Organism**

Selected filters: **Homo sapiens**

<i>Mus musculus</i>	93
<b><i>Homo sapiens</i></b>	<b>79</b>

### Showing 25 of 79 results

View All Download Visualize

Add all items to cart

- microRNA-seq of area 11 of Brodmann**  
*Homo sapiens* area 11 of Brodmann male adult (53 years)  
Lab: Joseph Costello, UCSF  
Project: Roadmap  
Experiment ENCSR396DOV **released**
- microRNA-seq of cerebellum**  
*Homo sapiens* cerebellum male embryo (20 weeks)  
Lab: Joseph Costello, UCSF  
Project: Roadmap  
Experiment ENCSR5840X **released**
- polyA RNA-seq of cerebellum**  
*Homo sapiens* cerebellum male embryo (20 weeks)  
Lab: Joseph Costello, UCSF  
Project: Roadmap  
Experiment ENCSR470JHE **released**
- microRNA-seq of insula**  
*Homo sapiens* insula male adult (53 years)  
Lab: Joseph Costello, UCSF  
Project: Roadmap  
Experiment ENCSR559FQR **released**
- microRNA-seq of brain**  
*Homo sapiens* brain female embryo (17 weeks)  
Lab: Joseph Costello, UCSF  
Project: Roadmap  
Experiment ENCSR778DXH **released**
- microRNA-seq of brain**  
*Homo sapiens* brain female embryo (17 weeks)  
Lab: Joseph Costello, UCSF  
Project: Roadmap  
Experiment ENCSR325SAK **released**
- polyA RNA-seq of layer of hippocampus**  
*Homo sapiens* layer of hippocampus male adult (81 year)  
Lab: Bradley Bernstein, Broad  
Project: Roadmap  
Experiment ENCSR239GFM **released**
- polyA RNA-seq of brain**  
*Homo sapiens* brain female adult (66 years)  
Lab: Michael Snyder, Stanford  
Project: ENCODE  
Experiment ENCSR274JRR **released**  
1 1
- polyA RNA-seq of SK-N-DZ**  
*Homo sapiens* SK-N-DZ nuclear fraction  
Lab: Thomas Gingeras, CSHL  
Experiment ENCSR265NYO **released**



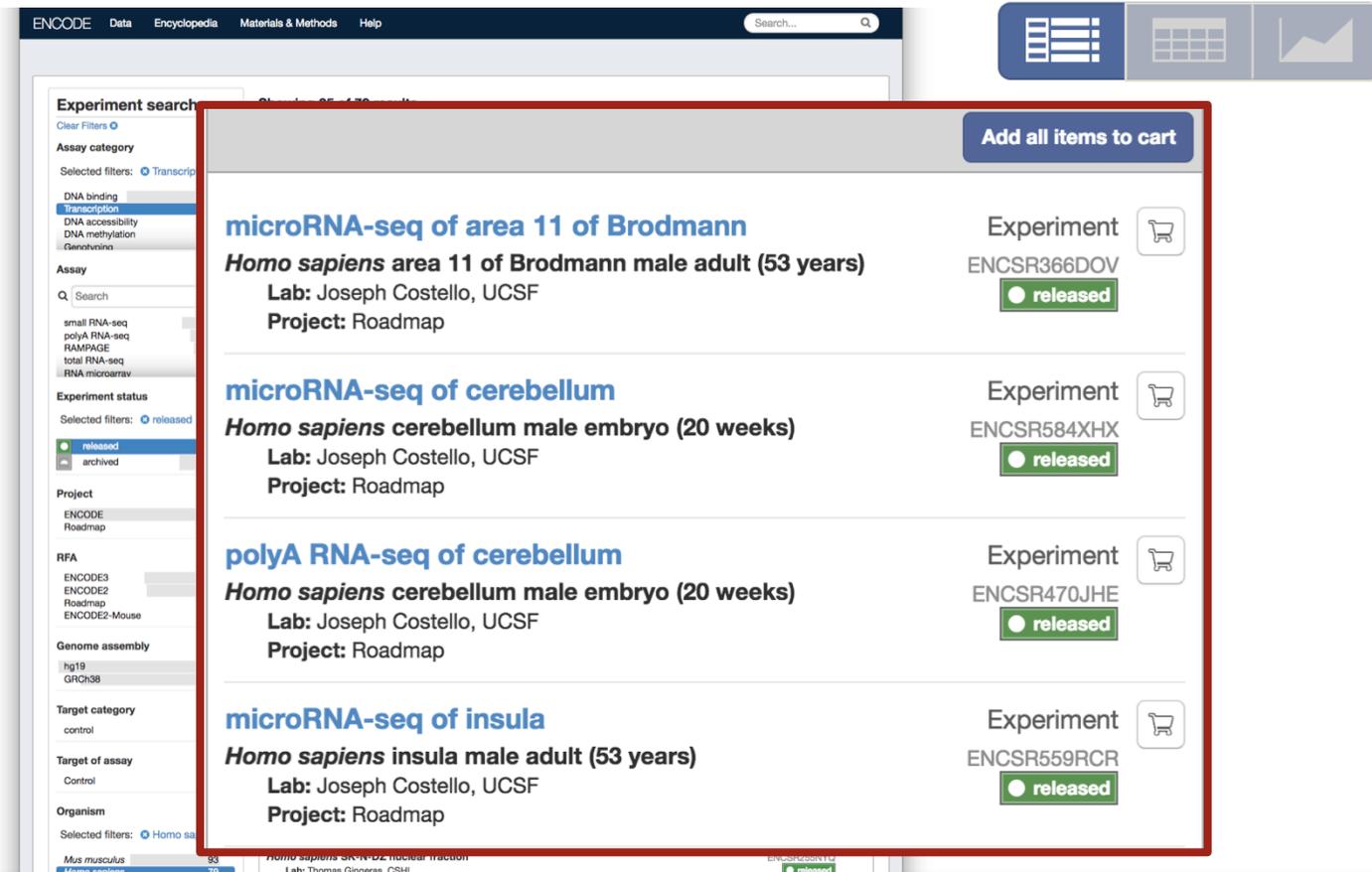
The screenshot shows the ENCODE portal search results page. The top navigation bar includes 'ENCODE', 'Data', 'Encyclopedia', 'Materials & Methods', and 'Help'. A search bar is located in the top right. On the left, there is a sidebar with filters for 'Assay category', 'Experiment status', 'Project', 'RFA', 'Genome assembly', 'Target category', 'Target of assay', and 'Organism'. The main content area displays a list of experiments. A red box highlights a section of the results, which includes the following entries:

- microRNA-seq of area 11 of Brodmann**  
*Homo sapiens* area 11 of Brodmann male adult (53 years)  
Lab: Joseph Costello, UCSF  
Project: Roadmap  
Experiment: ENCSR366DOV  
Status: released
- microRNA-seq of cerebellum**  
*Homo sapiens* cerebellum male embryo (20 weeks)  
Lab: Joseph Costello, UCSF  
Project: Roadmap  
Experiment: ENCSR584XHX  
Status: released
- polyA RNA-seq of cerebellum**  
*Homo sapiens* cerebellum male embryo (20 weeks)  
Lab: Joseph Costello, UCSF  
Project: Roadmap  
Experiment: ENCSR470JHE  
Status: released
- microRNA-seq of insula**  
*Homo sapiens* insula male adult (53 years)  
Lab: Joseph Costello, UCSF  
Project: Roadmap  
Experiment: ENCSR559RCR  
Status: released

Add experiments  
to Cart



(revisit later for file  
downloads)



The screenshot shows the ENCODE portal search results page. The top navigation bar includes 'ENCODE', 'Data', 'Encyclopedia', 'Materials & Methods', and 'Help'. A search bar is located in the top right. Below the navigation, there are three icons: a list view, a grid view, and a line graph. The main content area is titled 'Experiment search' and includes a 'Clear Filters' button and an 'Assay category' section with filters for 'Transcription', 'DNA binding', 'DNA accessibility', 'DNA methylation', and 'Gene expression'. The 'Assay' section has a search bar and filters for 'small RNA-seq', 'polyA RNA-seq', 'RAMPAGE', 'total RNA-seq', and 'RNA microarray'. The 'Experiment status' section has filters for 'released' and 'archived'. The 'Project' section has filters for 'ENCODE' and 'Roadmap'. The 'RFA' section has filters for 'ENCODE3', 'ENCODE2', 'Roadmap', and 'ENCODE2-Mouse'. The 'Genome assembly' section has filters for 'hg19' and 'GRCh38'. The 'Target category' section has a filter for 'control'. The 'Target of assay' section has a filter for 'Control'. The 'Organism' section has a filter for 'Homo sa'. The main results area shows a list of experiments, each with a title, description, lab, project, and a 'released' status indicator. A red box highlights the following experiments:

Assay	Experiment	Status
microRNA-seq of area 11 of Brodmann	ENCSR366DOV	released
microRNA-seq of cerebellum	ENCSR584XHX	released
polyA RNA-seq of cerebellum	ENCSR470JHE	released
microRNA-seq of insula	ENCSR559RCR	released

ENCODE Data Encyclopedia Materials & Methods Help Search...

### Experiment report

Clear Filters

#### Assay category

Selected filters: [Transcription](#)

DNA binding	233
<b>Transcription</b>	<b>79</b>
DNA accessibility	48
DNA methylation	27
Genes/proteins	14

#### Assay

Q Search

small RNA-seq	18
polyA RNA-seq	14
RAMPAGE	12
total RNA-seq	10
RNA microarray	9

#### Experiment status

Selected filters: [released](#)

<b>released</b>	<b>79</b>
archived	30

#### Project

ENCODE	72
Roadmap	7

#### RFA

ENCODE3	36
ENCODE2	35
Roadmap	7
ENCODE-Mouse	1

#### Genome assembly

hg19	78
GRCCh38	67

#### Target category

control	3
---------	---

#### Target of assay

Control	3
---------	---

#### Organism

Selected filters: [Homo sapiens](#)

<i>Mus musculus</i>	93
<b><i>Homo sapiens</i></b>	<b>79</b>

### Showing results 1 to 25 of 79

Columns Download TSV

ID	Accession	Assay Type	Assay Nickname	Target label	Target gene	Biosample summary	Biosample	Description
ENCSP366DOV	ENCSR366DOV	microRNA-seq	microRNA-seq			area 11 of Brodmann male adult (53 years)	area 11 of Brodmann	
ENCSR5840HX	ENCSR5840HX	microRNA-seq	microRNA-seq			cerebellum male embryo (20 weeks)	cerebellum	
ENCSR470JHE	ENCSR470JHE	RNA-seq	polyA RNA-seq			cerebellum male embryo (20 weeks)	cerebellum	
ENCSR559RCR	ENCSR559RCR	microRNA-seq	microRNA-seq			insula male	insula	



# Browsing the ENCODE portal

# Report view

The screenshot shows the ENCODE portal interface. At the top, there is a navigation bar with 'ENCODE', 'Data', 'Encyclopedia', 'Materials & Methods', and 'Help'. A search bar is on the right. Below the navigation bar, the main content area displays an 'Experiment report' for 'Assay category: Transcription'. It shows 'Showing results 1 to 25 of 79' and a 'Columns' button. A table of results is visible, with columns for Accession, Assay Type, Assay Nickname, Target label, Target gene, Biosample summary, Biosample, and Description. A dialog box titled 'Select columns to view' is open in the foreground, listing various columns with checkboxes. The dialog has 'Select all', 'Select ID only', and 'Default sort' options. At the bottom of the dialog are 'Cancel' and 'View selected columns' buttons.

**Select columns to view**

Select all | Select ID only | Default sort

- ID
- Assay Nickname
- Biosample summary
- Lab
- Biosample accession
- Linked Antibody
- Age
- Term ID
- Duration
- Post-synchronization time
- Submitter comment
- Documents
- External identifiers
- Schema Version
- Submitted by
- Lab aliases
- Experiment classification
- Additional data files
- Assay type
- Superseded by
- Developmental slims
- Revoked files
- Biosample synonyms
- Replication type
- System slims
- Accession
- Target label
- Biosample
- Project
- Biological replicate
- Species
- Age Units
- Concentration
- Duration units
- Post-synchronization time units
- Ontology ID
- References
- Date released
- Notes
- Lab
- Date submitted
- Controls
- Internal status
- Assay synonyms
- Assay category
- Assay ID
- Assay type
- Organ slims
- Assay objective
- Assembly
- Assay Type
- Target gene
- Description
- Status
- Technical replicate
- Life stage
- Treatment
- Concentration units
- Synchronization
- Replicates
- Biosample type
- Alternate accessions
- Internal tags
- Date created
- Grant
- Target
- Supersedes
- Pipeline error message
- Hub
- Month released
- Contributing files
- Original files
- Files
- Related series

Cancel | View selected columns



# Browsing the ENCODE portal

# Report view

The screenshot shows the ENCODE portal interface. At the top, there is a navigation bar with 'ENCODE', 'Data', 'Encyclopedia', 'Materials & Methods', and 'Help'. A search bar is on the right. Below the navigation bar, the main content area is titled 'Experiment report' and 'Showing results 1 to 25 of 79'. A 'Download TSV' button is highlighted with a red box. Below this, a table of results is visible with columns for 'Accession', 'Assay Type', 'Assay Nickname', 'Target label', 'Target gene', and 'Biosample summary'. A modal window titled 'Select columns to view' is open in the foreground, showing a list of columns with checkboxes. The 'Download TSV' button is also highlighted with a red box in the foreground.

Accession	Assay Type	Assay Nickname	Target label	Target gene	Biosample summary
ENCSR366DOV	microRNA-seq	microRNA-seq			area 11 of Brodmann male adult (53 years)
ENCSR584XHX	microRNA-seq	microRNA-seq		cerebellum male embryo (20 weeks)	cerebellum
ENCSR470JHE	RNA-seq	polyA RNA-seq		cerebellum male embryo (20 weeks)	cerebellum

Download TSV

## Select columns to view

Select all Select ID only

Default sort

- ID
- Assay Nickname
- Biosample summary
- Lab
- Biosample accession
- Linked Antibody
- Age
- Term ID
- Duration
- Post-synchronization time
- Submitter comment
- Documents
- External identifiers
- Schema Version
- Submitted by
- Lab aliases
- Experiment classification
- Additional data files
- Assay type
- Superseded by
- Developmental slims
- Revoked files
- Biosample synonyms
- Replication type
- System slims
- Accession
- Target label
- Biosample
- Project
- Biological replicate
- Species
- Age Units
- Concentration
- Duration units
- Post-synchronization time units
- Ontology ID
- References
- Date released
- Notes
- Lab
- Date submitted
- Controls
- Internal status
- Assay synonyms
- Assay category
- Assay ID
- Assay type
- Organ slims
- Assay objective
- Assembly
- Assay Type
- Target gene
- Description
- Status
- Technical replicate
- Life stage
- Treatment
- Concentration units
- Synchronization
- Replicates
- Biosample type
- Alternate accessions
- Internal tags
- Date created
- Grant
- Target
- Supersedes
- Pipeline error message
- Hub
- Month released
- Contributing files
- Original files
- Files
- Related series

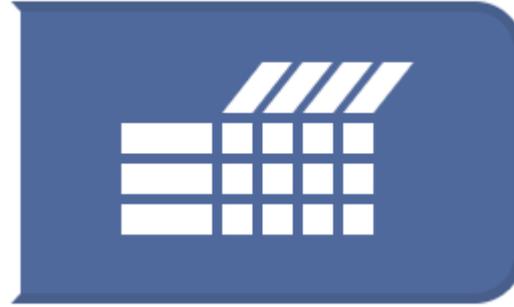
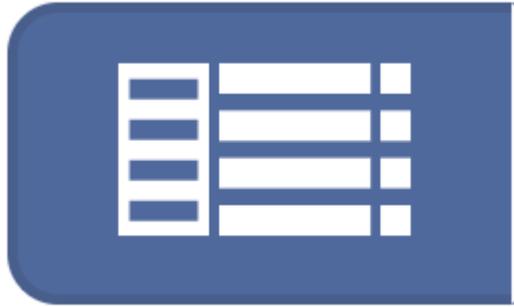
Cancel

View selected columns

Mus musculus 93  
Homo sapiens 79  
ENCSR559RGR ENCSR559RGR microRNA- microRNA- insula male insula

# Batch file download & visualization

**From search or matrix view**



# Batch file download

The screenshot shows the ENCODE Experiment matrix interface. The top navigation bar includes 'ENCODE', 'Data', 'Encyclopedia', 'Materials & Methods', and 'Help'. A search bar is located in the top right corner.

**Experiment matrix**

Enter search terms to filter the experiments included in the matrix.

Q

**Assay category**

Selected filters: **Transcription**

- DNA binding: 233
- Transcription: 70**
- DNA accessibility: 48
- DNA methylation: 27
- Genotyping: 14

**Assay**

Q

- small RNA-seq: 16
- polyA RNA-seq: 14
- RAMPAGE: 12
- total RNA-seq: 10
- RNA microarray: 9

**Experiment status**

Selected filters: **released**

- released: 79**
- archived: 30

**Project**

- ENCODE: 72
- Roadmap: 7

**Genome assembly**

- hg19: 78
- GRCh38: 67

**Target category**

- control: 3

**Target of assay**

- Control: 3

**Organism**

Selected filters: **Homo sapiens**

- Mus musculus*: 93
- Homo sapiens*: 79**

**79 results**

Clear Filters

**BIOSAMPLE**

**ASSAY**

- small RNA-seq
- polyA RNA-seq
- RAMPAGE
- total RNA-seq
- RNA microarray
- microRNA-seq
- CG
- RNA-FRET
- single cell RNA-seq
- polyA depletion RNA-seq

**cell line**

cell line	SK-N-SH	10	5	1	4	3	1
SK-N-DZ	2	2	3	1			
A172			1	1			
BEZC	1			1			
Dazy			1	1			

+ See 5 more...

**tissue**

tissue	cerebellum	1	1	1	1	1	
brain		1				2	
diencephalon	1		1	1			
frontal cortex	1		1	1			
occipital lobe	1		1	1			

+ See 5 more...

**primary cell**

primary cell	Purkinje cell	1				1	
cerebellar granule cell							2
astrocyte of the cerebellum				1			
astrocyte of the hippocampus					1		
brain microvascular endothelial cell						1	

+ See 2 more...

See all biosamples

**Download** **Visualize**

Download

# Batch file download

### 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 has the URL or command line 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 -L 1 curl -O -L < files.txt
```

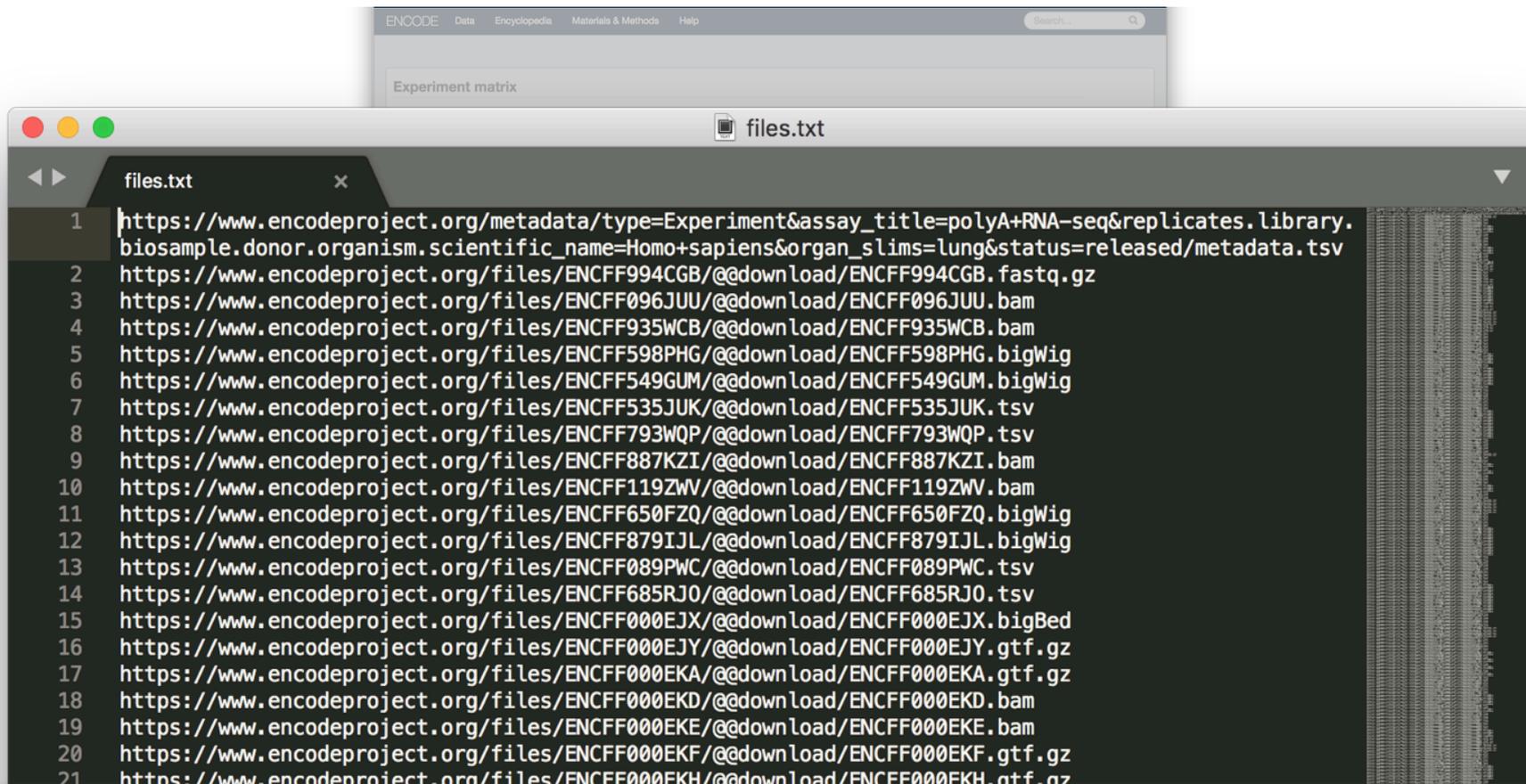
[Close](#) [Download](#)

See all 12000 genes

<b>Project</b>	
ENCODE	72
Roadmap	7
<b>Genome assembly</b>	
hg19	78
GRCh38	67
<b>Target category</b>	
control	3
<b>Target of assay</b>	
Control	3
<b>Organism</b>	
Selected filters: <span>Homo sapiens</span>	
Mus musculus	93
Homo sapiens	78

[Download](#) [Visualize](#)

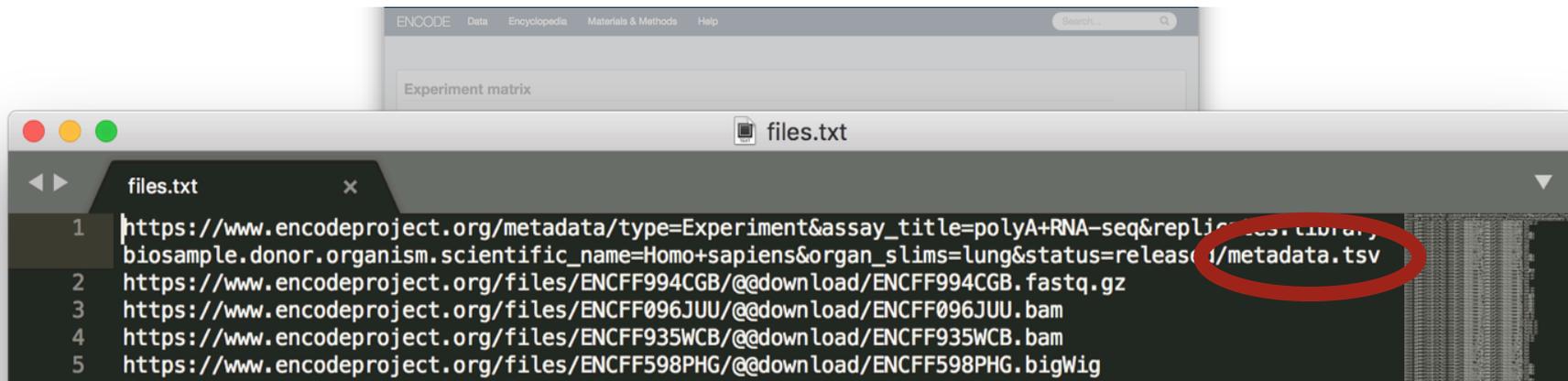
# Batch file download



The image shows a web browser window displaying an ENCODE experiment matrix. The browser's address bar shows the ENCODE website. Below the browser, a text editor window titled 'files.txt' is open, displaying a list of 21 URLs for downloading files from the ENCODE project. The URLs are numbered 1 through 21 and include file names such as metadata.tsv, fastq.gz, bam, bigWig, and tsv files.

```
1 https://www.encodeproject.org/metadata/type=Experiment&assay_title=polyA+RNA-seq&replicates.library.biosample.donor.organism.scientific_name=Homo+sapiens&organ_slims=lung&status=released/metadata.tsv
2 https://www.encodeproject.org/files/ENCFF994CGB/@download/ENCFF994CGB.fastq.gz
3 https://www.encodeproject.org/files/ENCFF096JUU/@download/ENCFF096JUU.bam
4 https://www.encodeproject.org/files/ENCFF935WCB/@download/ENCFF935WCB.bam
5 https://www.encodeproject.org/files/ENCFF598PHG/@download/ENCFF598PHG.bigWig
6 https://www.encodeproject.org/files/ENCFF549GUM/@download/ENCFF549GUM.bigWig
7 https://www.encodeproject.org/files/ENCFF535JUK/@download/ENCFF535JUK.tsv
8 https://www.encodeproject.org/files/ENCFF793WQP/@download/ENCFF793WQP.tsv
9 https://www.encodeproject.org/files/ENCFF887KZI/@download/ENCFF887KZI.bam
10 https://www.encodeproject.org/files/ENCFF119ZVW/@download/ENCFF119ZVW.bam
11 https://www.encodeproject.org/files/ENCFF650FZQ/@download/ENCFF650FZQ.bigWig
12 https://www.encodeproject.org/files/ENCFF879IJL/@download/ENCFF879IJL.bigWig
13 https://www.encodeproject.org/files/ENCFF089PWC/@download/ENCFF089PWC.tsv
14 https://www.encodeproject.org/files/ENCFF685RJ0/@download/ENCFF685RJ0.tsv
15 https://www.encodeproject.org/files/ENCFF000EJX/@download/ENCFF000EJX.bigBed
16 https://www.encodeproject.org/files/ENCFF000EJY/@download/ENCFF000EJY.gtf.gz
17 https://www.encodeproject.org/files/ENCFF000EKA/@download/ENCFF000EKA.gtf.gz
18 https://www.encodeproject.org/files/ENCFF000EKD/@download/ENCFF000EKD.bam
19 https://www.encodeproject.org/files/ENCFF000EKE/@download/ENCFF000EKE.bam
20 https://www.encodeproject.org/files/ENCFF000EKF/@download/ENCFF000EKF.gtf.gz
21 https://www.encodeproject.org/files/ENCFF000EKH/@download/ENCFF000EKH.gtf.gz
```

# Batch file download



The image shows a web browser window displaying an ENCODE experiment matrix. Below the browser, a text editor window titled 'files.txt' is open, showing a list of URLs for data files. The first line of the list is circled in red:

```
1 https://www.encodeproject.org/metadata/type=Experiment&assay_title=polyA+RNA-seq&replicates=10&library_biosample.donor.organism.scientific_name=Homo+sapiens&organ_slims=lung&status=released/1/metadata.tsv
2 https://www.encodeproject.org/files/ENCFF994CGB/@download/ENCFF994CGB.fastq.gz
3 https://www.encodeproject.org/files/ENCFF096JUU/@download/ENCFF096JUU.bam
4 https://www.encodeproject.org/files/ENCFF935WCB/@download/ENCFF935WCB.bam
5 https://www.encodeproject.org/files/ENCFF598PHG/@download/ENCFF598PHG.bigWig
```

## metadata.tsv

one line for each data file in files.txt

file format, output type, read info, assembly

also assay, biosample, library specifics

## Batch file download

**To download all files, run command...**

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

## Batch file download

**To download all files, run command...**

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

Download a subset of files (e.g. only hg19 alignments)

**Or copy/paste metadata.tsv link into browser,  
filter metadata.tsv (assembly, file format, ...),  
collect download links (column AK in metadata.tsv),  
put in new file to run command**

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

# Batch file download

# Using your Cart

The screenshot shows the ENCODE website interface. At the top, there is a navigation bar with 'ENCODE', 'Data', 'Encyclopedia', 'Materials & Methods', and 'Help'. A search bar is located on the right. Below the navigation bar, there is a sidebar on the left for 'Experiment search' and a main content area on the right for 'Showing 25 of 79 results'. The sidebar includes filters for 'Assay category', 'Assay', 'Experiment status', 'Project', 'RFA', 'Genome assembly', 'Target category', 'Target of assay', and 'Organism'. The main content area displays a list of experiments, each with a title, description, lab, and project. A 'Cart' icon is visible next to each experiment entry.

Assay category	Count
DNA binding	233
Transcription	79
DNA accessibility	48
DNA methylation	27
Genetics	14

Assay	Count
small RNA-seq	18
polyA RNA-seq	14
RAMPAGE	12
total RNA-seq	10
RNA microarray	9

Experiment status	Count
released	79
archived	30

Project	Count
ENCODE	72
Roadmap	7

RFA	Count
ENCODE3	36
ENCODE2	35
Roadmap	7
ENCODE2-Mouse	1

Genome assembly	Count
hg19	78
GRCh38	67

Target category	Count
control	3

Target of assay	Count
Control	3

Organism	Count
Mus musculus	93
Homo sapiens	79

Experiment	Count
microRNA-seq of area 11 of Brodmann	1
microRNA-seq of cerebellum	1
polyA RNA-seq of cerebellum	1
microRNA-seq of insula	1
microRNA-seq of brain	1
microRNA-seq of brain	1
polyA RNA-seq of layer of hippocampus	1
polyA RNA-seq of brain	1
polyA RNA-seq of SK-N-DZ	1

Add experiments to Cart



# Batch file download

# Using your Cart

The screenshot shows the ENCODE website interface. At the top, the navigation bar includes 'ENCODE', 'Data', 'Encyclopedia', 'Materials & Methods', and 'Help'. A search bar is located on the right. The main content area is divided into two columns. The left column contains filters for 'Experiment search', 'Assay category', 'Experiment status', 'Project', 'RFA', 'Genome assembly', 'Target category', 'Target of assay', and 'Organism'. The right column displays a list of 25 experiments, each with a title, description, lab, and project. A red box highlights the '19' icon in the top navigation bar, which represents the number of items in the cart. A red-bordered box on the right side of the screenshot contains a shopping cart icon with the number '19' and the text 'View cart'.

ENCODE Data Encyclopedia Materials & Methods Help 19 Search...

Experiment search  
Clear Filters  
Assay category  
Selected filters: Transcription  
DNA binding 233  
Transcription 79  
DNA accessibility 48  
DNA methylation 27  
Genome assembly 14  
Assay  
Search  
small RNA-seq 18  
polyA RNA-seq 14  
RAMPAGE 12  
total RNA-seq 10  
RNA microarray 9  
Experiment status  
Selected filters: released  
released 79  
archived 30  
Project  
ENCODE 72  
Roadmap 7  
RFA  
ENCODE3 36  
ENCODE2 35  
Roadmap 7  
ENCODE2-Mouse 1  
Genome assembly  
hg19 78  
GRCh38 67  
Target category  
control 3  
Target of assay  
Control 3  
Organism  
Selected filters: Homo sapiens  
Mus musculus 93  
Homo sapiens 79

Showing 25 of 79 results  
View All Download Visualize

microRNA-seq of area 11 of Brodmann  
Homo sapiens area 11 of Brodmann male adult (53 years)  
Lab: Joseph Costello, UCSF  
Project: Roadmap

microRNA-seq of cerebellum  
Homo sapiens cerebellum male embryo (20 weeks)  
Lab: Joseph Costello, UCSF  
Project: Roadmap

polyA RNA-seq of cerebellum  
Homo sapiens cerebellum male embryo (20 weeks)  
Lab: Joseph Costello, UCSF  
Project: Roadmap

microRNA-seq of insula  
Homo sapiens insula male adult (53 years)  
Lab: Joseph Costello, UCSF  
Project: Roadmap

microRNA-seq of brain  
Homo sapiens brain female embryo (17 weeks)  
Lab: Joseph Costello, UCSF  
Project: Roadmap

microRNA-seq of brain  
Homo sapiens brain female embryo (17 weeks)  
Lab: Joseph Costello, UCSF  
Project: Roadmap

polyA RNA-seq of layer of hippocampus  
Homo sapiens layer of hippocampus male adult (81 year)  
Lab: Bradley Bernstein, Broad  
Project: Roadmap

polyA RNA-seq of brain  
Homo sapiens brain female adult (66 years)  
Lab: Michael Snyder, Stanford  
Project: ENCODE

polyA RNA-seq of SK-N-DZ  
Homo sapiens SK-N-DZ nuclear fraction  
Lab: Thomas Gingeras, CSHL

Experiment ENCSR470JHE released  
Experiment ENCSR559RQB released  
Experiment ENCSR778DXH released  
Experiment ENCSR325SAK released  
Experiment ENCSR239GFM released  
Experiment ENCSR274JRR released  
Experiment ENCSR265NYO released

19  
View cart

# Batch file download

# Using your Cart

Download

756 files selected for download

File type	
fastq	630
bigWig	42
bam	32
tsv	16
bed narrowPeak	15
bigBed narrowPeak	15
bed bed3+	2
bed broadPeak	2
bigBed broadPeak	2

ENCODE Data Encyclopedia Materials & Methods Help 19 Search... Jason Hilton

### Cart

Download Clear cart

756 files selected for download

19 datasets in cart

<b>total RNA-seq of IMR-90</b> <i>Homo sapiens</i> IMR-90 Lab: Joe Ecker, Salk Project: Roadmap	Experiment ENCSR424FAZ released 2 2
<b>single cell RNA-seq of forelimb</b> <i>Mus musculus</i> C57BL/6 forelimb embryo (15.5 days) Lab: Barbara Wold, Caltech Project: ENCODE	Experiment ENCSR4300IC released 3
<b>single cell RNA-seq of forelimb</b> <i>Mus musculus</i> C57BL/6 forelimb embryo (13 days) Lab: Barbara Wold, Caltech Project: ENCODE	Experiment ENCSR182WHH released 3
<b>single cell RNA-seq of forelimb</b> <i>Mus musculus</i> C57BL/6 forelimb embryo (14.5 days) Lab: Barbara Wold, Caltech Project: ENCODE	Experiment ENCSR548KBB released 3
<b>WGBS of heart left ventricle</b> <i>Homo sapiens</i> heart left ventricle female adult (51 year) Lab: Richard Myers, HAIB Project: ENCODE	Experiment ENCSR307KDA released 2
<b>DNase-seq of liver</b> <i>Mus musculus</i> C57BL/6 liver adult (8-10 weeks) treated with darkness for 70 days Lab: John Stamatoyannopoulos, UW Project: ENCODE	Experiment ENCSR802MKV released 1 2
<b>Hi-C of mature B cell</b> <i>Mus musculus</i> C57BL/6 mature B cell treated with 50 µg/mL lipopolysaccharide for 24 hours, 5 ng/mL Interleukin-4 for 24 hours, 500 ng/mL anti-CD180 for 24 hours, 99.7 ng/mL oligomycin for 2 hours, 820 µg/mL 2-Deoxy-D-glucose for 2 hours Lab: Erez Lieberman-Aiden, Baylor Project: ENCODE	Experiment ENCSR847UWT released 1 1
<b>Hi-C of mature B cell</b> <i>Mus musculus</i> C57BL/6 mature B cell treated with 50 µg/mL lipopolysaccharide for 72 hours, 5 ng/mL Interleukin-4 for 72 hours, 500 ng/mL anti-CD180 for 72 hours Lab: Erez Lieberman-Aiden, Baylor Project: ENCODE	Experiment ENCSR471ZWB released 1

The screenshot shows the ENCODE Cart interface. At the top, there is a navigation bar with 'ENCODE', 'Data', 'Encyclopedia', 'Materials & Methods', 'Help', and a search bar. Below the navigation bar, the page title is 'Cart'. On the right side of the cart area, there are 'Download' and 'Clear cart' buttons. The main content area shows '19 datasets in cart'. On the left, a table lists '756 files selected for download' with columns for 'File type' and a count. The table includes file types like 'fastq', 'bigWig', 'bam', 'tsv', 'bed narrowPeak', 'bigBed narrowPeak', 'bed bed3+', 'bed broadPeak', and 'bigBed broadPeak'. The main list of datasets includes:

- total RNA-seq of IMR-90**  
*Homo sapiens* IMR-90  
Lab: Joe Ecker, Salk  
Project: Roadmap
- single cell RNA-seq of forelimb**  
*Mus musculus* C57BL/6 forelimb embryo (15.5 days)  
Lab: Barbara Wold, Caltech  
Project: ENCODE
- single cell RNA-seq of forelimb**  
*Mus musculus* C57BL/6 forelimb embryo (13 days)  
Lab: Barbara Wold, Caltech  
Project: ENCODE
- single cell RNA-seq of forelimb**  
*Mus musculus* C57BL/6 forelimb embryo (14.5 days)  
Lab: Barbara Wold, Caltech  
Project: ENCODE
- WGBS of heart left ventricle**  
*Homo sapiens* heart left ventricle female adult (51 year)  
Lab: Richard Myers, HAIB  
Project: ENCODE
- DNase-seq of liver**  
*Mus musculus* C57BL/6 liver adult (8-10 weeks) treated with darkness for 70 days  
Lab: John Stamatoyannopoulos, UW  
Project: ENCODE

Each dataset entry includes an 'Experiment' label, an experiment ID, a 'released' status, and a download icon with a count.

**Future Cart developments:  
multiple Carts, naming Carts, sharing Carts...**

# Visualizing data

The screenshot shows the ENCODE Experiment matrix interface. The top navigation bar includes 'ENCODE', 'Data', 'Encyclopedia', 'Materials & Methods', and 'Help'. A search bar is located in the top right. The main content area is titled 'Experiment matrix' and contains a search input field with the placeholder 'Enter search term(s)'. Below this are several filter sections: 'Assay category' with 'Transcription' selected, showing counts for DNA binding (233), Transcription (70), DNA accessibility (48), DNA methylation (27), and Genotyping (14); 'Assay' with a search field and a bar chart showing counts for small RNA-seq (16), polyA RNA-seq (14), RAMPAGE (12), total RNA-seq (10), and RNA microarray (9); 'Experiment status' with 'released' selected, showing 79 released and 30 archived; 'Project' with ENCODE (72) and Roadmap (7); 'Genome assembly' with hg19 (78) and GRCh38 (67); 'Target category' with 'control' selected, showing 3; 'Target of assay' with 'Control' selected, showing 3; and 'Organism' with 'Homo sapiens' selected, showing 93 Mus musculus and 79 Homo sapiens. The main results area shows '79 results' and a table with columns for 'BIOSAMPLE' and 'ASSAY'. The table lists cell lines (SK-N-SH, SK-N-DZ, A172, BEZC, Dazl) and tissues (cerebellum, brain, diencephalon, frontal cortex, occipital lobe) and primary cells (Purkinje cell, cerebellar granule cell, astrocyte of the cerebellum, astrocyte of the hippocampus, brain microvascular endothelial cell). A 'Download' button and a 'Visualize' button are at the bottom of the table.

## Open visualization browser

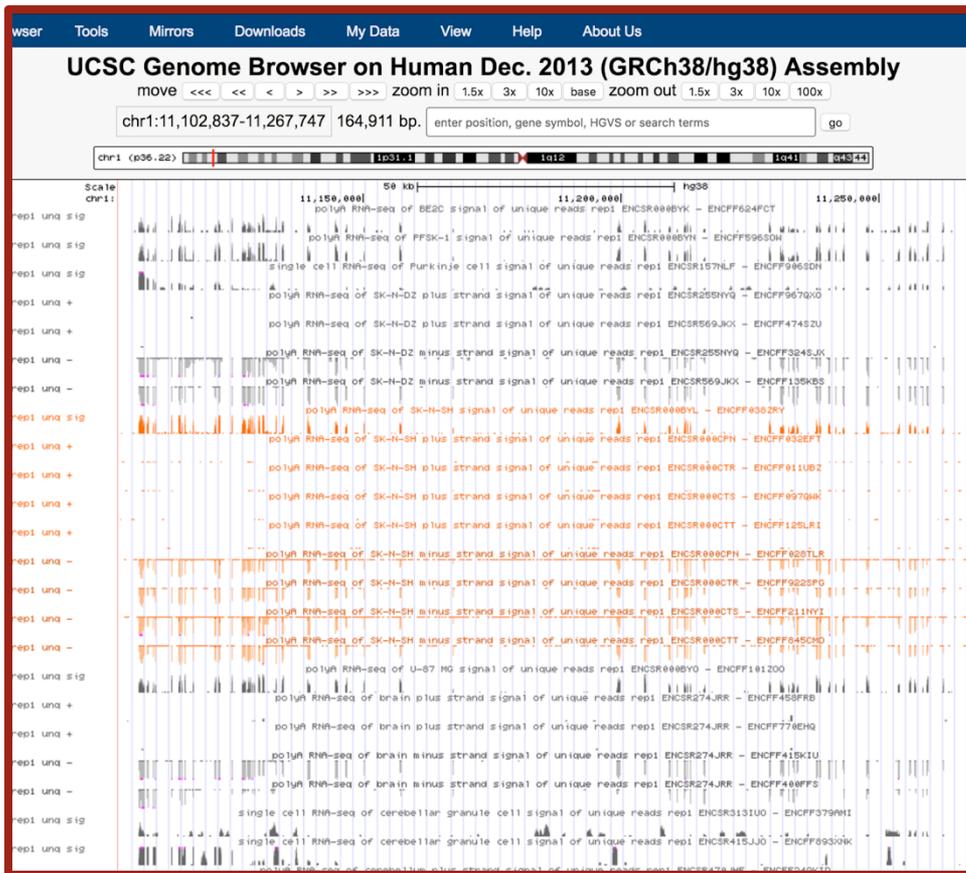
Assembly Visualize with browser...

GRCh38: [UCSC](#) [Ensembl](#)

hg19: [UCSC](#)

[Close](#)

# Visualizing data



## Open visualization browser

Assemble Visualize with browser...

GRCh38: **UCSC** Ensembl

hg19: **UCSC**

Close

The image shows a screenshot of the ENCODE portal website. At the top, there is a navigation bar with the following items: ENCODE, Data, Encyclopedia, Materials & Methods, and Help. A search bar is located on the right side of the navigation bar. The 'Help' menu is open, showing a list of options: Getting started, Tutorials, REST API (highlighted with a blue bar), Cart, Project overview, Acknowledgements, Contact, News, and Citing ENCODE. A red arrow points from the 'REST API' option to the main content area. The main content area features a diagram of a DNA molecule with various regulatory elements and associated data tracks. The diagram includes labels for 'Hypersensitive Sites', 'CH<sub>3</sub>', and 'CH<sub>3</sub>CO'. Below the diagram are several data tracks: 5C ChIA-PET Hi-C, DNase-seq FAIRE-seq ATAC-seq, CHIP-seq, WGBS RRBS methyl array, Computational predictions, RNA-seq, and CLIP-seq RIP-seq. These tracks are connected to a DNA sequence with labels for 'Long-range regulatory elements (enhancers, repressors/silencers, insulators)', 'Promoters', 'Genes', and 'Transcripts'. On the right side of the page, there are several buttons and search boxes: 'About ENCODE Project', 'Getting Started', 'Experiments', 'Search ENCODE portal', 'ENCOD Q', 'About ENCODE Encyclopedia', 'Candidate Regulatory Elements', 'Search for Candidate Regulatory Elements', 'Hosted by SCREEN', 'Human hg19 Q', and 'Mouse mm10 Q'.

# Programmatic access

# JSON objects



to view json (w/ a json plug-in)

## Object pages

## Search pages

ENCODE Data Encyclopedia Materials & Methods Help Search...

EXPERIMENTS / CHIP-SEQ / HOMO SAPIENS / ASCENDING AORTA

### Experiment summary for ENCSR982QIF

Summary Attribution

Summary	Attribution
Status: <span>released</span>	Lab: Bradley Bernstein, Broad
Assay: ChIP-seq	Award: U54HG006991 (Bradley Bernstein, Broad)
Target: H3K27ac	Project: ENCODE

ENCODE Data Encyclopedia Materials & Methods Help Search...

Showing 25 of 49 results

Assay category: Transcription (49)

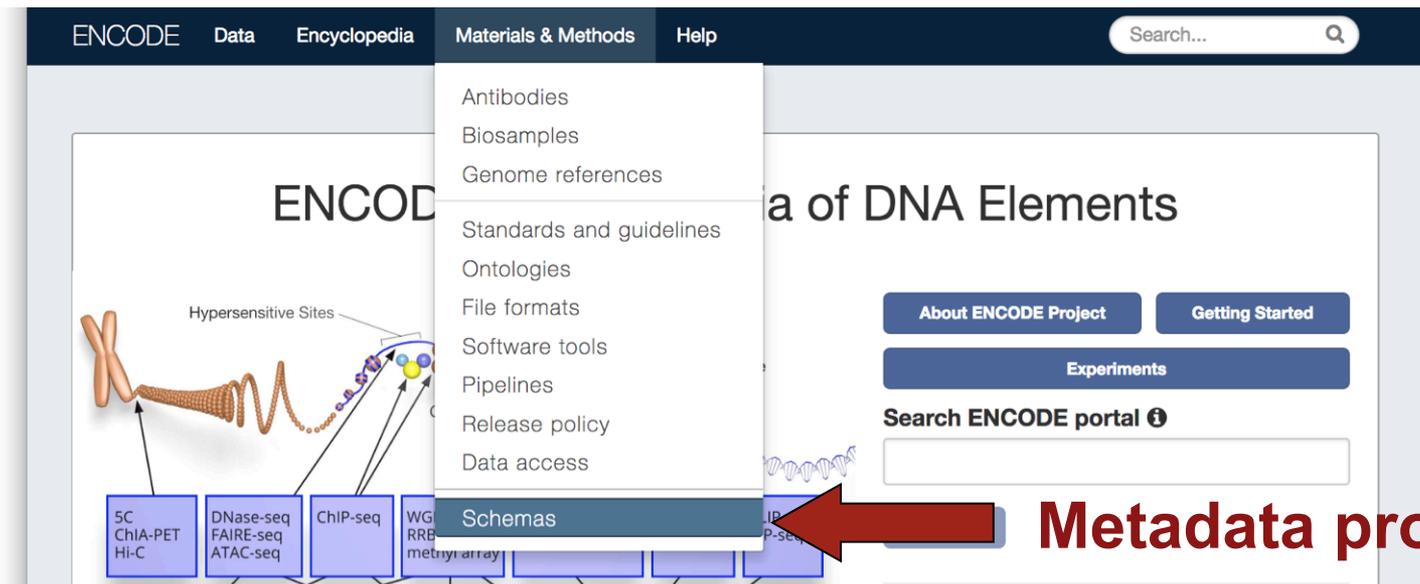
Assay: ChIP-seq (545), DNase-seq (69), **polyA RNA-seq (49)**, small RNA-seq (26), DNase array (18)

Experiment status: **released (49)**

**polyA RNA-seq of A549**  
Homo sapiens A549 treated with 100 nM dexamethasone for 25 minutes  
Lab: Tim Reddy, Duke  
Project: GGR  
Experiment: ENCSR525SHS released

**polyA RNA-seq of A549**  
Homo sapiens A549 treated with 100 nM dexamethasone for 20 minutes  
Lab: Tim Reddy, Duke  
Project: GGR  
Experiment: ENCSR525SHS released





**Metadata properties**

## ENTEx (ENCODE + Gtex)

31 tissue types from 4 donors

## Mouse development series

13 tissues across 9 developmental time points

## ENCORE: Encyclopedia of RNA Elements

>250 RNA-Binding Proteins in HepG2 & K562

## Reference Epigenomes

Methylation, transcription, & 'core' histone marks on the same cell/tissue type

# Encyclopedia resources

The screenshot displays the ENCODE website's 'Encyclopedia of DNA Elements' page. The navigation bar includes 'ENCODE', 'Data', 'Encyclopedia', 'Materials & Methods', and 'Help'. A search bar is located in the top right. The 'Encyclopedia' dropdown menu is highlighted with a red circle and contains the following items: 'About', 'Visualize (SCREEN)', 'Matrix', 'Search', and 'Methods'. The main content area features a diagram of a DNA segment with various annotations. Above the DNA, there are labels for 'Hypersensitive Sites', 'CH<sub>3</sub>', and 'CH<sub>3</sub>CO'. A 'RNA polymerase' is shown transcribing the DNA into a 'Transcript'. Below the DNA, there are boxes for different assays: '5C', 'ChIA-PET', 'Hi-C', 'DNase-seq', 'FAIRE-seq', 'ATAC-seq', 'ChIP-seq', 'WGBS', 'RRBS', 'methyl array', 'Computational predictions', 'RNA-seq', and 'CLIP-seq', 'RIP-seq'. Arrows indicate the relationship between these assays and the genomic features. The 'Search for Candidate Regulatory Elements' section is also circled in red and includes buttons for 'About ENCODE Project', 'Getting Started', 'Experiments', and 'Candidate Regulatory Elements'. It also features a search bar and buttons for 'hg19' and 'Mouse mm10'.

# THANK YOU!!!

encode-help@lists.stanford.edu  
@EncodeDCC  
<https://github.com/ENCODE-DCC>



Mike



Idan



Ben



Seth



Meenakshi



Jennifer



Yunhai



Jason



Casey



Forrest



Jin



Bek



Vanessa



Jennifer



Paul



Zack



Khine



Emma



Phil



Otto



Keenan

NIH U24 HG009397