

How to download the ENCODE data

You Zhou

2024-12-20

Table of contents

1	Introduction	1
2	Data download	1
2.1	Getting metadata	4

1 Introduction

This report shows how to down load the ENCODE data from (<https://www.encodeproject.org/>).

2 Data download

To correctly download the selected files. We need to use the function cart in ENCODE. We can first We can first go to experiment search and select the wanted data.

Note, to get the knockdown and knockout data, we need to select the option `Hide control experiments == no`. Because the default in ENCODE select only the `no perturbed` data, by default there is not KO or KD experiment exist in `Assay title`.

Then add them to cart and go to cart page.

Afterwards, we need to remove the show default data only and select the correct filtering standard for the wanted data. Note, the default data only contains the result from replicate 1. If we want to download both, we need to use the right filtering.

In the end, when we download the data, we need to use the download button and select for the right files. It always shows the default file automatically so we need to select the right files.

ENCODE

Data

Encyclopedia

Materials & Methods

Help

Experiment

Clear all selected

Assay

Assay type

Assay title

Search

TF ChIP-seq

Histone ChIP-seq

DNase-seq

total RNA-seq

Mint-ChIP-seq

polyA plus fraction

ATAC-seq

microRNA-seq

snRNA-seq

snATAC-seq

Target category

Target of analysis

Search

H3K4me3

H3K27ac

H3K27me3

H3K4me1

Functional Genomics data ?

Experiment search

Experiment matrix

ChIP-seq matrix

Human and mouse body maps

Functional genomics series

Single-cell experiments

Functional Characterization data ?

Experiment search

Experiment matrix

Cloud Resources

AWS Open Data

Azure Open Datasets

Collections

RNA-protein interactions (ENCORE)

Epigenomes from four individuals (ENTEx)

Rush Alzheimer's disease study

Stem cell differentiation

Deeply profiled cell lines

Human donor matrix

Immune cells

Human reference epigenomes

Mouse reference epigenomes

Mouse development matrix

Protein knockdown (Degron)

Search by region

Publications

RNA-Get (gene expression)

Figure 1: Experiment search

Assay title

Search

TF ChIP-seq	4690
Histone ChIP-seq	3694
DNase-seq	1723
total RNA-seq	946
Mint-ChIP-seq	895
polyA plus RNA-seq	753
ATAC-seq	459
microRNA-seq	412
snRNA-seq	395
snATAC-seq	364

Target category

Target of assay

Hide control experiments

☒ yes

17252

☐ no

20909

Biosample

Figure 2: Select to not hide the control

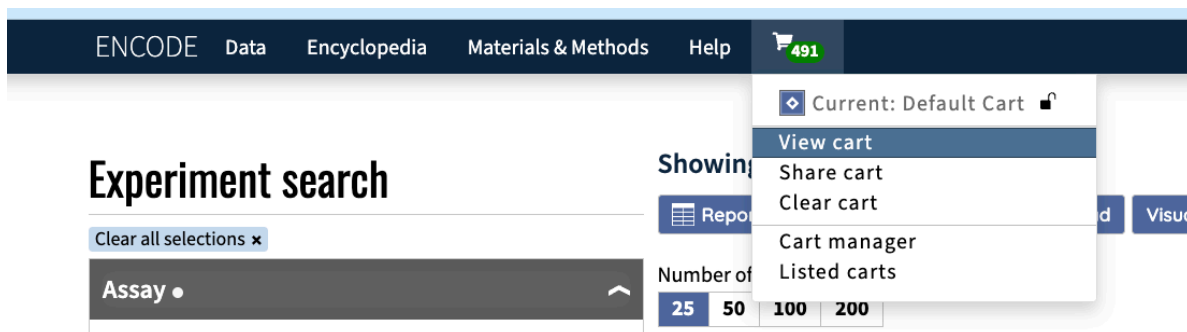


Figure 3: Cart page

2.1 Getting metadata

To get the metadata, we can use the dataset report function and get a tsv table back. In this page, we can also decide which information we would like to include in the meta data by selecting them in the `column` option. By default, the control information is not included in this table. Therefore, do not forget to select the `controls` in the `column` options.

Default Cart

◇ unlisted

No description

🔒 Lock

Clear cart

List cart

Select single dataset type to download

Dataset



28,161 files selected

- ☐ Show default data only
- ☐ Show visualizable data only

DATASET TYPE

☐ Experiment 491

ANALYSIS/ASSEMBLY

☐ GRCh38 15376 ▲

☐ hg19 10803 ▲

☐ ENCODE4 v1.2.3
GRCh38 V29 2088 ▲

☐ ENCODE4 v1.2.1
GRCh38 V29 6336 ▲

☐ ENCODE4 GRCh38 V29 414 ▲

More filters

Figure 4: Filtering

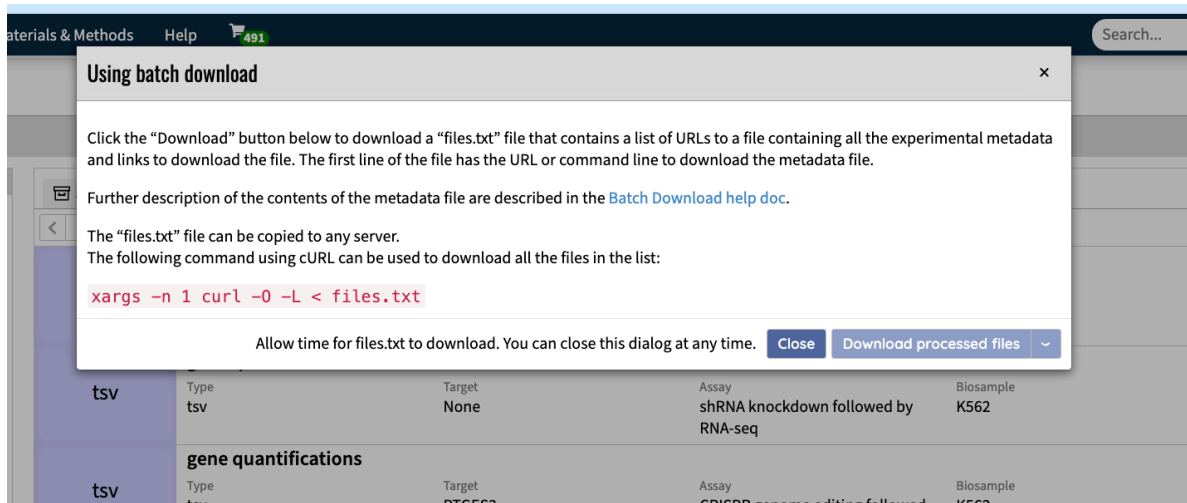


Figure 5: Select right files

Default Cart

Default Cart

unlisted

No description

Lock Clear cart List cart

Download Dataset report

704 files selected

☐ Show default data only
☐ Show visualizable data only

DATASET TYPE

☒ Experiment 491

ANALYSIS/ASSEMBLY

☒ GRCh38 704

☐ ENCODE4 v1.2.3 GRCh38 V29 232

Series 0 **Selected experiments** 491 **Genome browser** 0 **Processed** 704

1 OF 29

Type	Target	Assay
tsv	SYNCRIP	CRISPR ger by RNA-sec
gene quantifications		
tsv	SYNCRIP	CRISPR ger by RNA-sec
gene quantifications		
tsv	None	CRISPR ger by RNA-sec

Figure 6: Dataset report

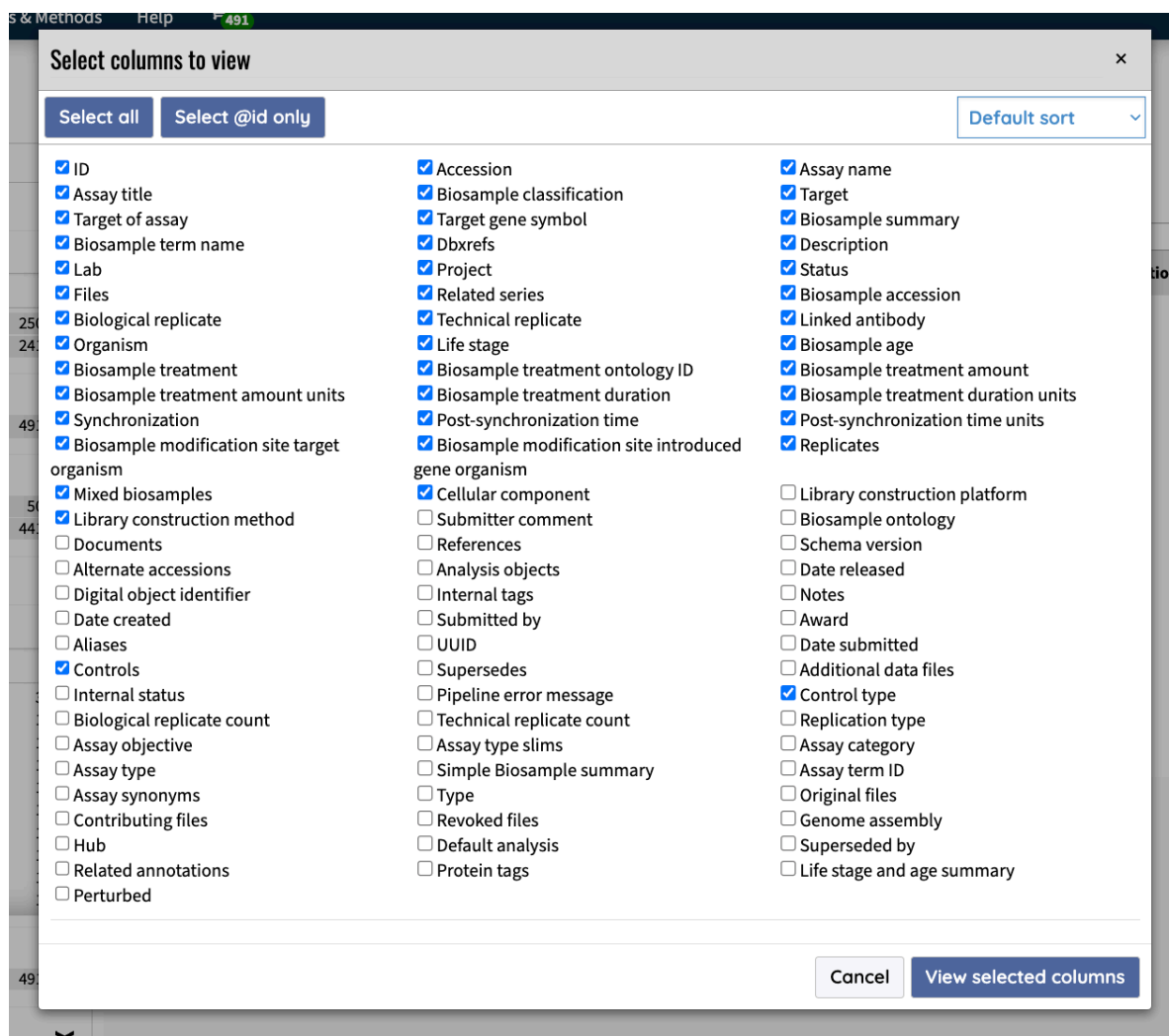


Figure 7: Download tsv

Experiment report

ColumnsDownload TSV

Items per page: 2550100200

Showing results 1 to 25 of 491

ID	Accession	Assay name	Assay title	Biosample classification	Target	Target of assay	Target gene symbol	Biosample summary	Biosample term name
ENCSR727WPW	ENCSR727WPW	CRISPR genome editing followed by RNA-seq	CRISPR RNA-seq	cell line	/targets/SEPT11-human/	SEPT11	SEPT11	Homo sapiens K562 genetically modified (deletion) using CRISPR targeting H. sapiens SF3B2	K562
ENCSR033MIV	ENCSR033MIV	CRISPR genome editing followed by RNA-seq	CRISPR RNA-seq	cell line	/targets/SRA1-human/	SRA1	SRA1	Homo sapiens K562 genetically modified (deletion) using CRISPR targeting H. sapiens SRA1	K562

Clear all selections x

Assay type

Assay title

Status

Perturbation

Target category

Target of assay

Figure 8: Controls