

FILER tutorial

- Webserver: <https://lisanwanglab.org/FILER>. Three main functions:
 - Browse/download: identify and retrieve relevant FILER tracks for your downstream analyses
 - Search FILER tracks/genomic records using genomic coordinates
 - Analyze/annotate your own experimental data/loci:
 - Upload local file
 - Provide a data URL
- FILER stand-alone version, API (<https://bitbucket.org/wanglab-upenn/FILER>)

FILER webserver: browse/download

A) quickly identify and retrieve relevant experiments and data in specific tissue contexts for downstream analyses

Browse FILER

To select specific FILER datasets, use any of the filters below in any order:

Select Genome build (3):

Select Data Source (13):

Select classification (1):

Select Assay (1):

Select Tissue category (21):

Select cell type (12):

[Reset Form](#)

1. Data selectors

2. Bulk download annotation/data

[Download](#)

3. Individual dataset download

Showing records 1 to 13 out of 13

Identifier	Data Source	Download file	Number of intervals	bp covered	Output type	Genome build	cell type	Biosample type	Biosamples term id	Tissue category	ENCODE Experiment id	Biological replicate(s)	Technical replicate	Antibody	Assay	File format	File size	Release date	Date added to GADB
NGRM005470	ROADMAP_Enhancers	Download	94,007	82,930,758	ChromHMM_enhancer	hg38	Angular gyrus	primary tissue	Not applicable	Brain	Not applicable	Not applicable	Not applicable	consolidated	ChIP-seq	bed bed4	612,819	3/30/18	8/25/18
NGRM005471	ROADMAP_Enhancers	Download	110,876	97,818,524	ChromHMM_enhancer	hg38	Anterior caudate	primary tissue	Not applicable	Brain	Not applicable	Not applicable	Not applicable	consolidated	ChIP-seq	bed bed4	719,252	3/30/18	8/25/18
NGRM005472	ROADMAP_Enhancers	Download	106,044	97,458,143	ChromHMM_enhancer	hg38	Cingulate gyrus	primary tissue	Not applicable	Brain	Not applicable	Not applicable	Not applicable	consolidated	ChIP-seq	bed bed4	691,888	3/30/18	8/25/18

FILER webserver: analyze/annotate your own data

B) quickly search and retrieve all genomic intervals within the genomic regions of interest

Search FILER

The screenshot shows the FILER webserver search interface. At the top, there is a 'Genome:' dropdown menu with 'hg19' selected, circled in orange. A yellow callout box labeled '1. Select genome build' points to this dropdown. Below this, there are three tabs: 'Search by genomic coordinates' (active), 'Analyze your own data', and 'Upload your data'. Under the active tab, there is a text input field labeled 'Enter URL for your data (bed, bed.gz):' containing the text 'BED, BED.GZ URL, e.g., https://tf.lisanwanglab.org/GADB/www/tes'. This input field is also circled in orange. A yellow callout box labeled '2. Provide experimental loci in BED or bed.gz formats' points to this input field. Below the input field, there is a green link labeled 'Example URL' and a dark blue 'submit' button.

Genome: hg19

1. Select genome build

Search by genomic coordinates | Analyze your own data | Upload your data

Enter URL for your data (bed, bed.gz):

BED, BED.GZ URL, e.g., https://tf.lisanwanglab.org/GADB/www/tes

Example URL

submit

2. Provide experimental loci in BED or bed.gz formats

FILER webserver: analysis results page

Search results

Run summary

Number of input intervals: 100

Number of database intervals: 7,113,542,775 (22,123 tracks)

Performed 711,354,277,500 searches in 9.556803 seconds (74,434,334,320.329322 intervals/sec)

Found 292,978 overlapping database intervals.

Giggle search time: .887014313

Combining overlaps time: .702927392

Report time: 7.656749389

Link to the results folder: [FILER_out/bf5d182](#) [763M]

Link to results: [Download results](#) (ZIP)

Annotated overlaps: [FILER_out/bf5d182/giggle_overlaps.with_meta.bed](#) (ZIP)

Input size

Total number of FILER tracks and genomic records searched

of overlaps found

running time stats

Links to zipped results and results folder

Distribution of overlaps across data sources and tissue categories +

Distribution of overlaps across data sources and experimental assays/data types +

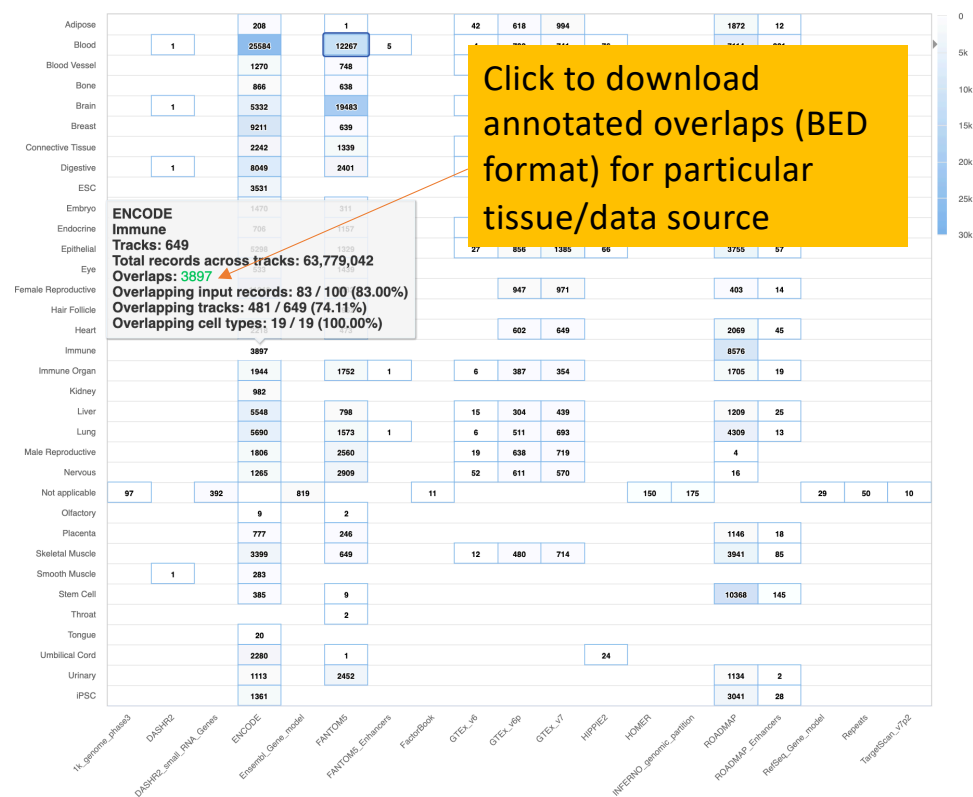
Genomic feature type overlap summary +

Data collection overlap summary +

Download results +

Run log +

Distribution of overlaps across data sources and tissue categories +



FILER: stand-alone usage and API

- FILER code repository: <https://bitbucket.org/wanglab-upenn/FILER>
- Scripts for installing a stand-alone FILER instance (use FILER in your own cluster/cloud computing environment)
- Command line CLI scripts for accessing/querying FILER data: query, retrieve genomic data at scale
- Scripts for genomic data indexing and pre-processing (prepare data for adding to FILER)