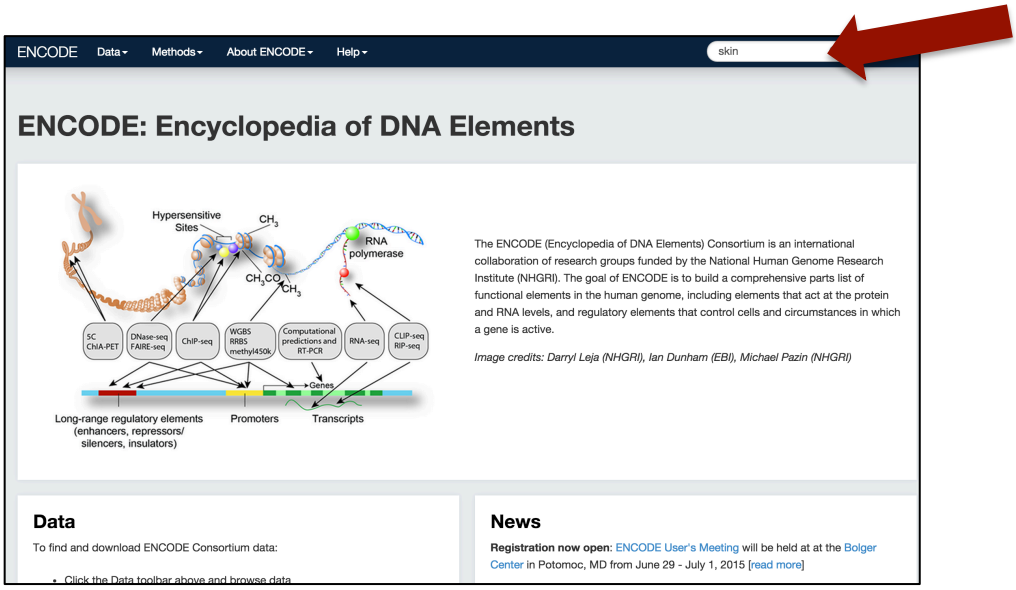
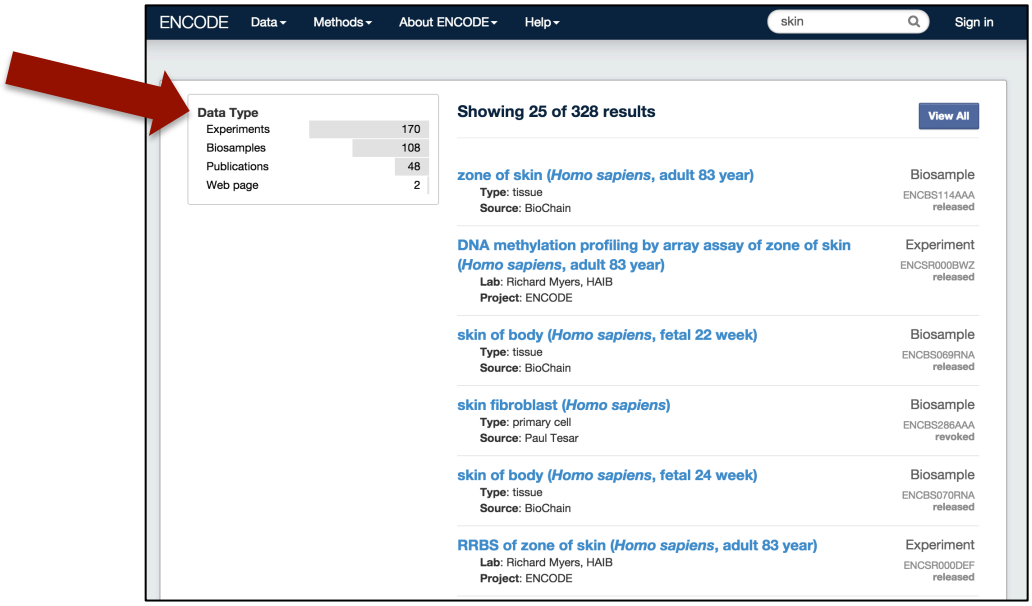
Demo 1: Free text search of ENCODE data

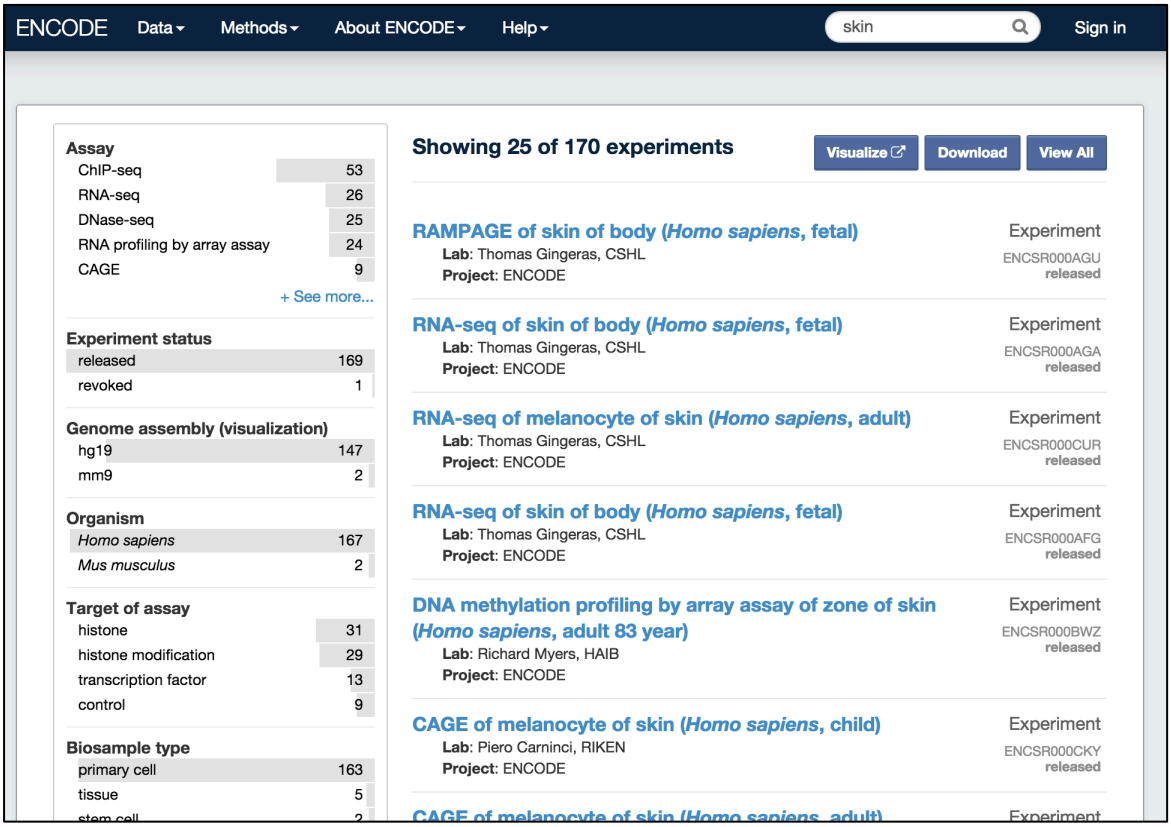
1. Go to https://www.encodeproject.org
2. Enter “skin” into the search box on the upper right hand corner



3. All matches on the website will be shown
4. Select “Experiments”

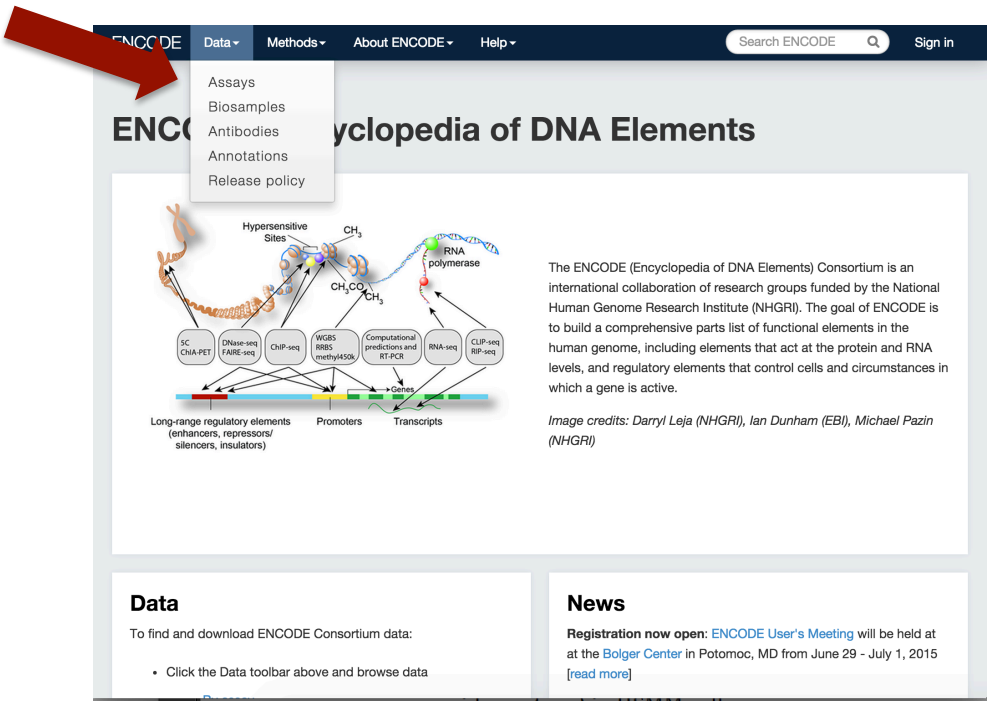


5. View all ENCODE assays that contain a match to “skin”

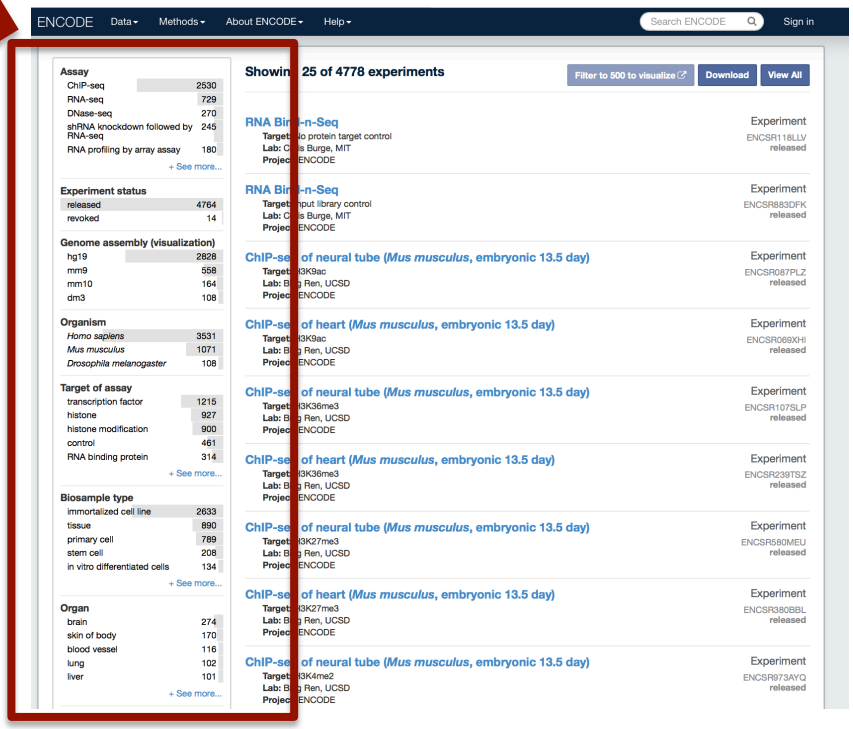


Demo 2: Browsing and filtering of ENCODE data

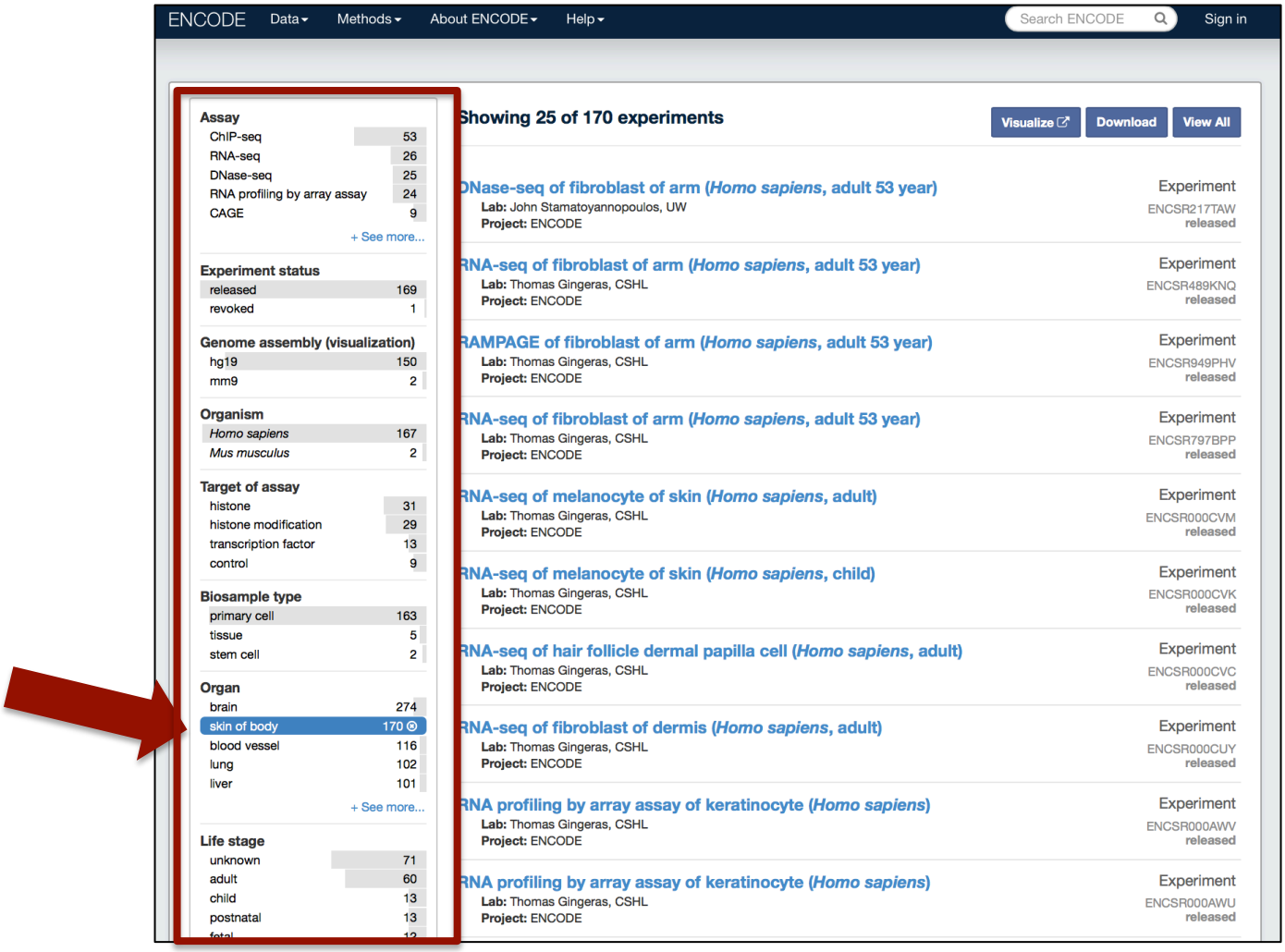
1. Go to https://www.encodeproject.org
2. Under the “Data” menu bar, select “Assays”



3. All publicly available assays are shown. The categories on the left are metadata describing the assays. They can be used to filter your results.

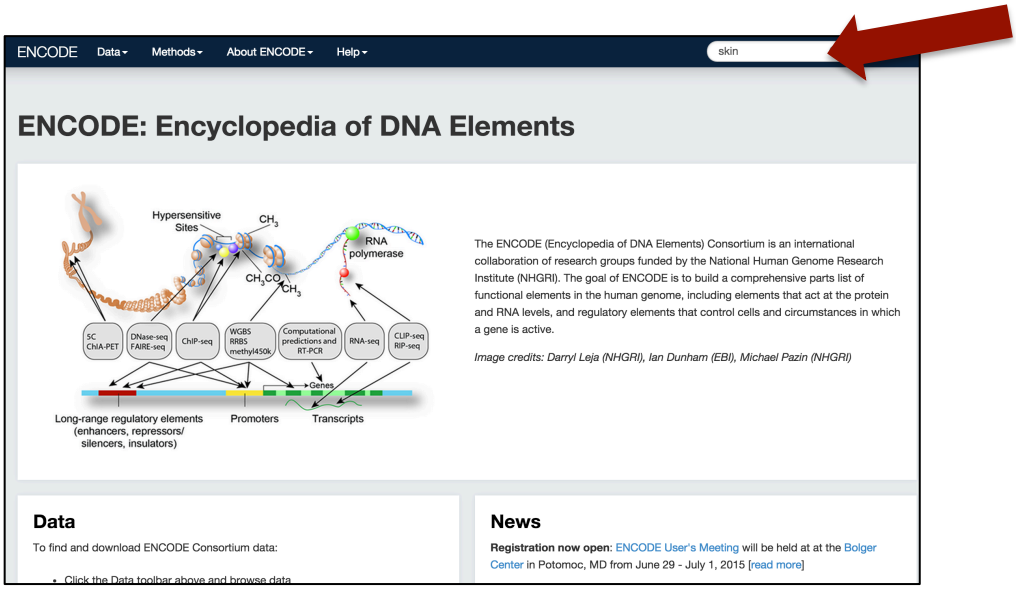


5. View all ENCODE assays that are filtered by selecting “skin of body”. Note that the results are not straight text matches. For example, “fibroblast of arm”, and “keratinocytes” are included in the results list.

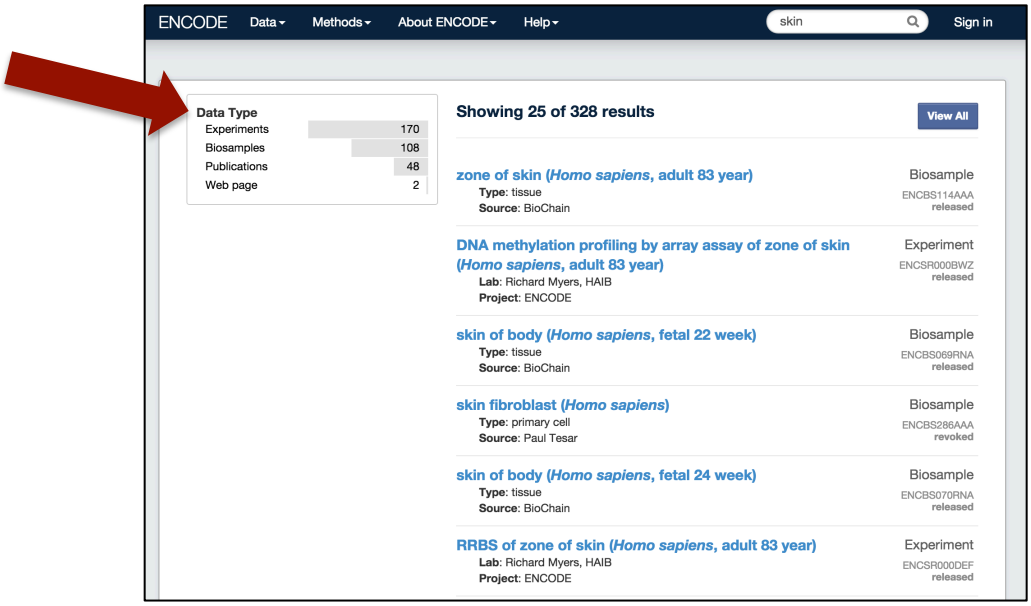


Demo: Combine search & filter of ENCODE data

1. Go to https://www.encodeproject.org
2. Enter “skin” into the search box on the upper right hand corner

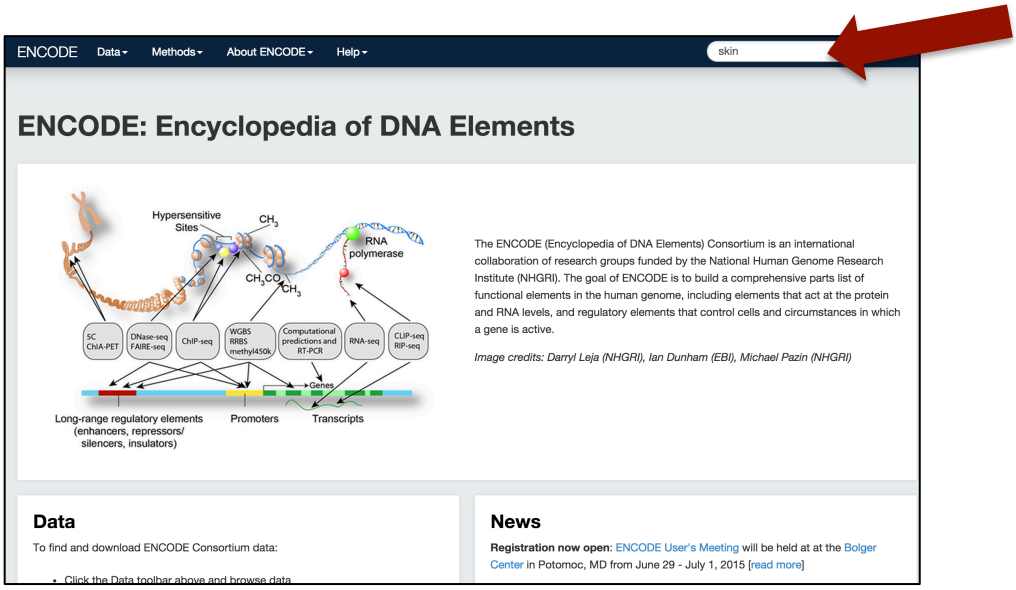


3. All matches on the website will be shown
4. Select “Experiments”

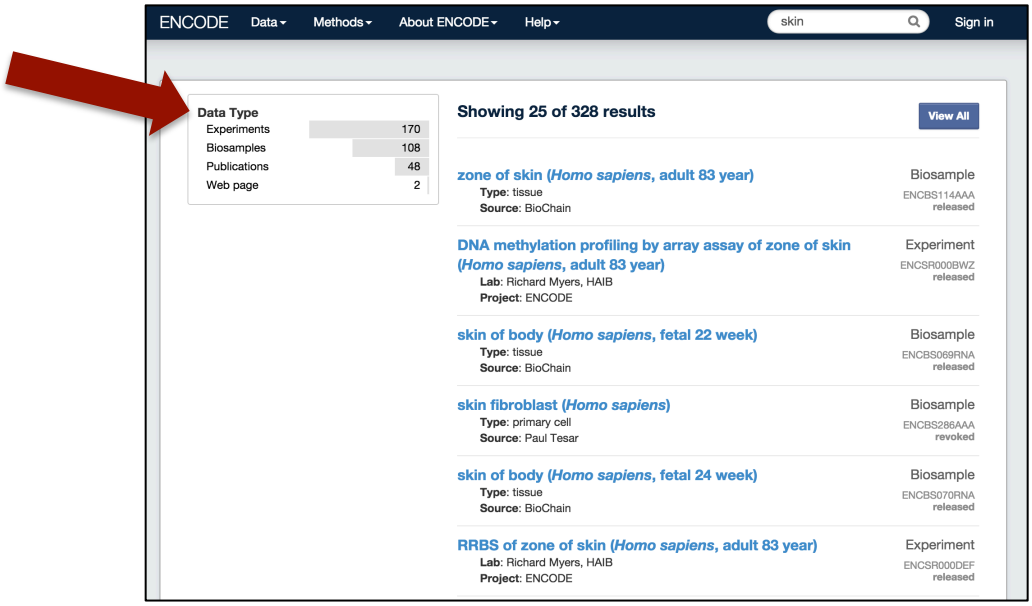


Demo 3: Combine search & filter of ENCODE data

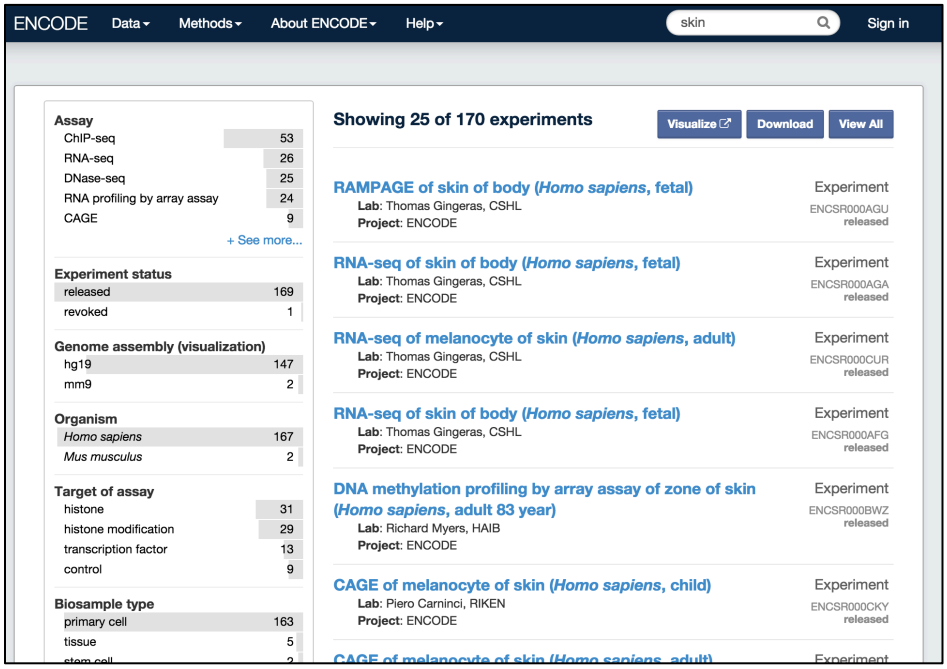
1. Go to https://www.encodeproject.org
2. Enter “skin” into the search box on the upper right hand corner



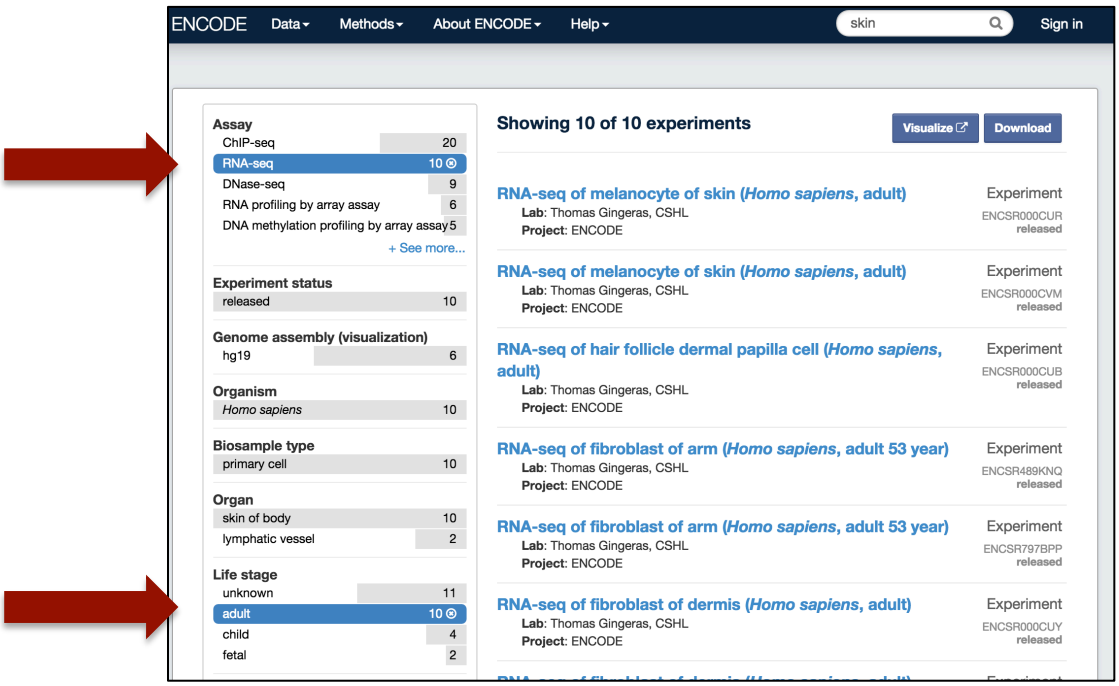
3. All matches on the website will be shown
4. Select “Experiments”



5. View all ENCODE assays that contain a match to “skin”



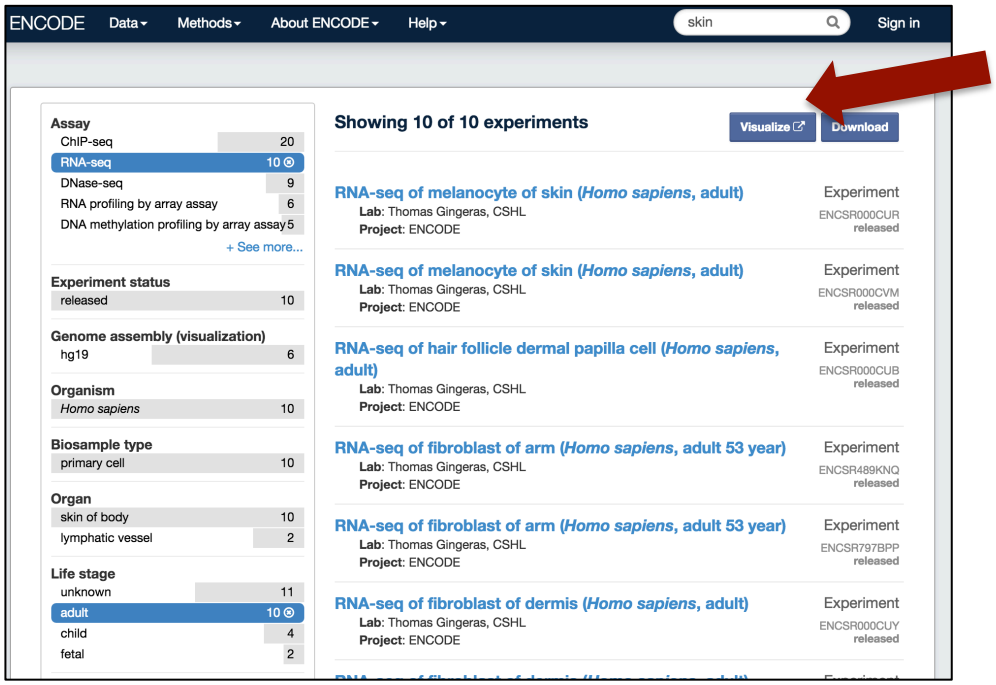
6. If looking for RNA-seq data from adult samples, select “RNA-seq” under Assay and “adult” under Life Stage.



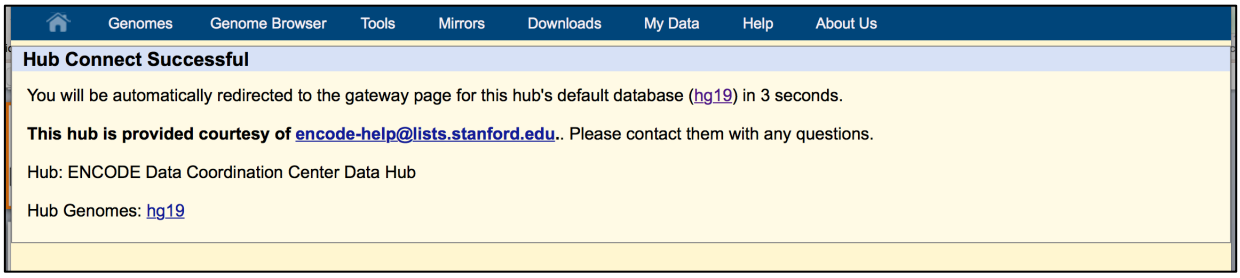
Demo 4: Visualize data

Details of how to configure tracks will be presented during the UCSC Genome Browser workshop. This demo just shows what to expect as you are connected to the browser.

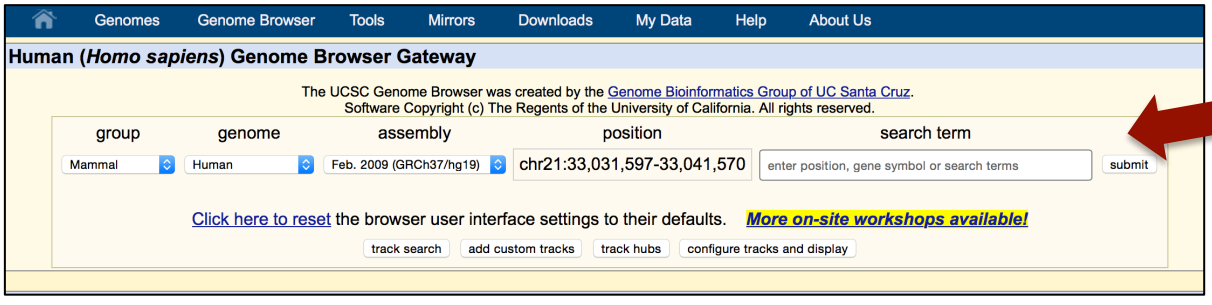
1. Get the list of assays illustrated in Demo 3.
2. Select "Visualize" button on the upper right corner. This will automatically create a trackhub to visualize the results



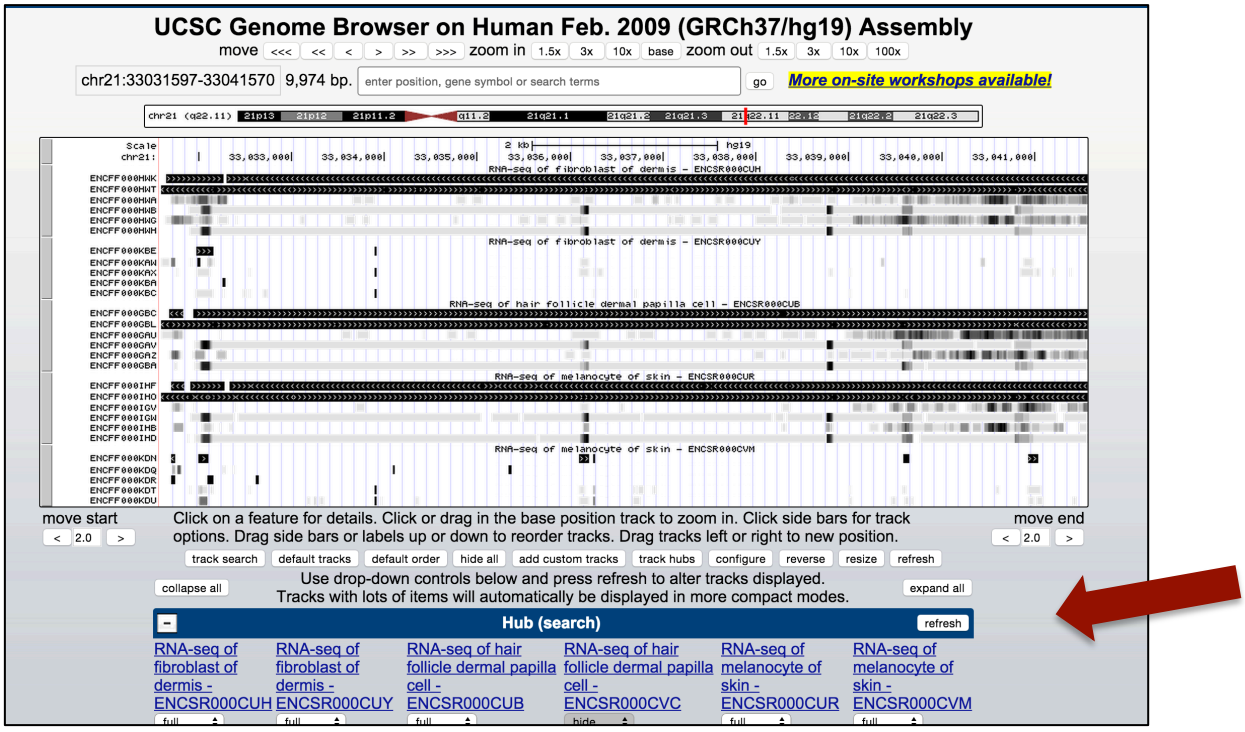
3. An intermediate page listing the reference genome is shown. Trackhubs can be made with different reference



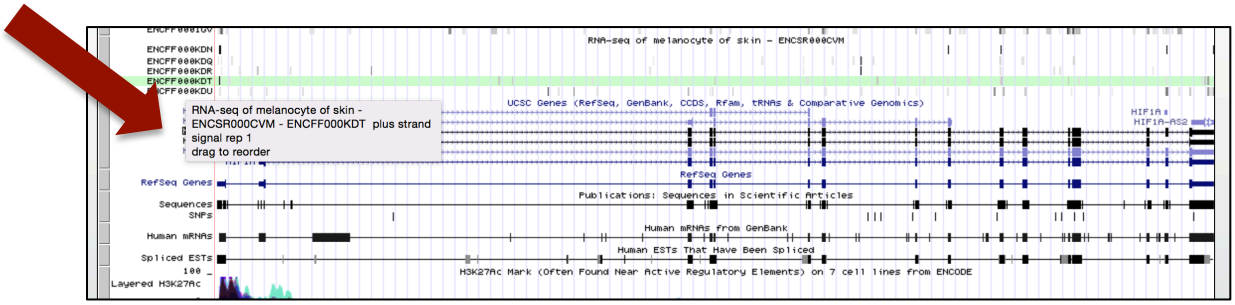
4. The trackhub will be connected. Enter a gene name or enter submit. Details of how to configure tracks will be presented during the UCSC Genome Browser workshop.



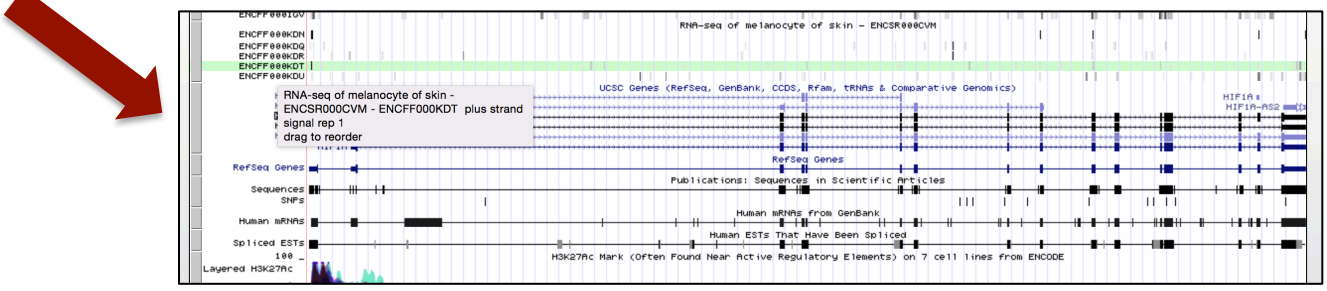
5. The track hub is listed as “Hub (search)”



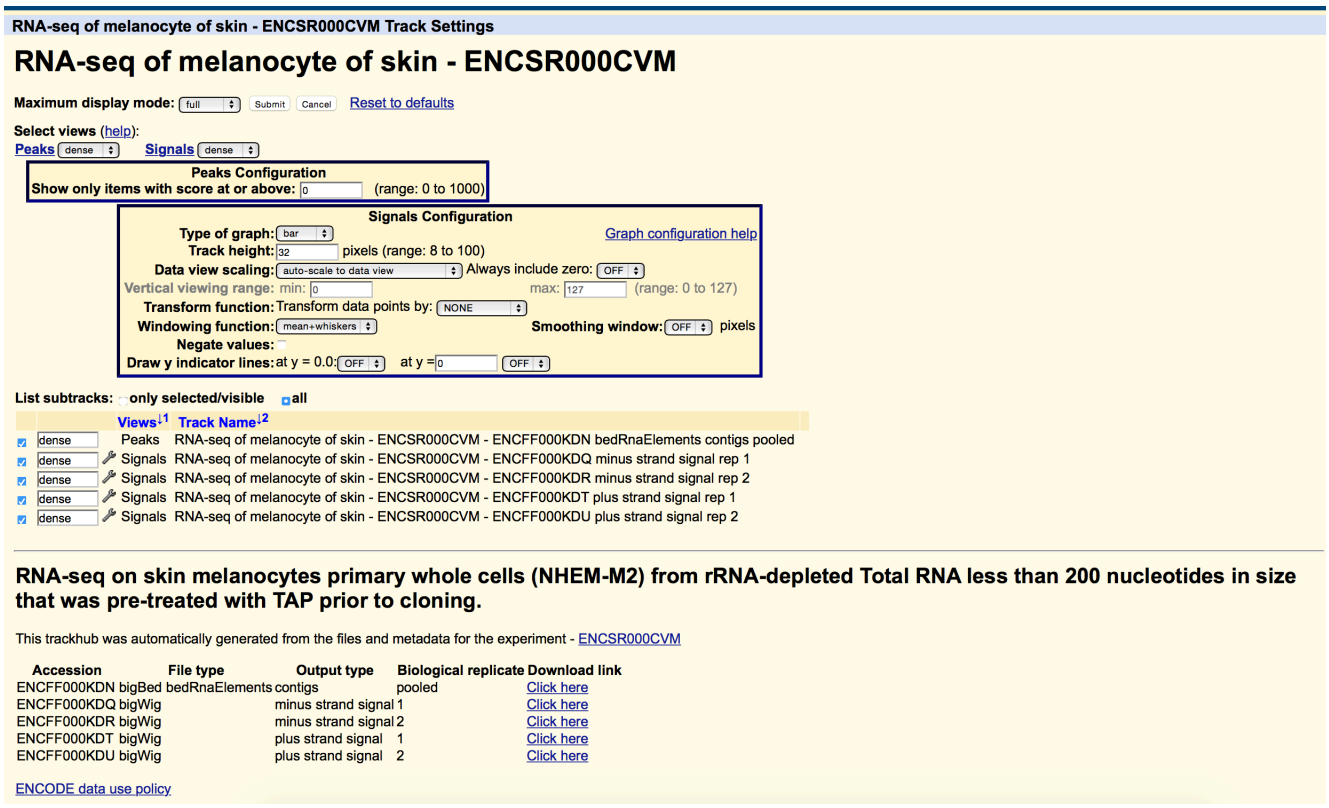
6. Hover over tracks to view file-related metadata.



7. Click on left hand grouping to configure tracks and view more metadata, including file download links and links back to the ENCODE Portal.



8. ENCODE metadata for the track hub



RNA-seq of melanocyte of skin - ENCSR000CVM Track Settings

RNA-seq of melanocyte of skin - ENCSR000CVM

Maximum display mode: full Submit Cancel Reset to defaults

Select views (help):

Peaks dense Signals dense

Peaks Configuration

Show only items with score at or above: 0 (range: 0 to 1000)

Signals Configuration

Type of graph: bar Graph configuration help

Track height: 32 pixels (range: 8 to 100)

Data view scaling: auto-scale to data view Always include zero: OFF

Vertical viewing range: min: 0 max: 127 (range: 0 to 127)

Transform function: Transform data points by: NONE

Windowing function: mean+whiskers Smoothing window: OFF pixels

Negate values:

Draw y indicator lines: at y = 0.0: OFF at y = 0 OFF

List subtracks: only selected/visible all

	Views↓1	Track Name↓2
dense	Peaks	RNA-seq of melanocyte of skin - ENCSR000CVM - ENCFF000KDN bedRnaElements contigs pooled
dense	Signals	RNA-seq of melanocyte of skin - ENCSR000CVM - ENCFF000KDQ minus strand signal rep 1
dense	Signals	RNA-seq of melanocyte of skin - ENCSR000CVM - ENCFF000KDR minus strand signal rep 2
dense	Signals	RNA-seq of melanocyte of skin - ENCSR000CVM - ENCFF000KDT plus strand signal rep 1
dense	Signals	RNA-seq of melanocyte of skin - ENCSR000CVM - ENCFF000KDU plus strand signal rep 2

RNA-seq on skin melanocytes primary whole cells (NHEM-M2) from rRNA-depleted Total RNA less than 200 nucleotides in size that was pre-treated with TAP prior to cloning.

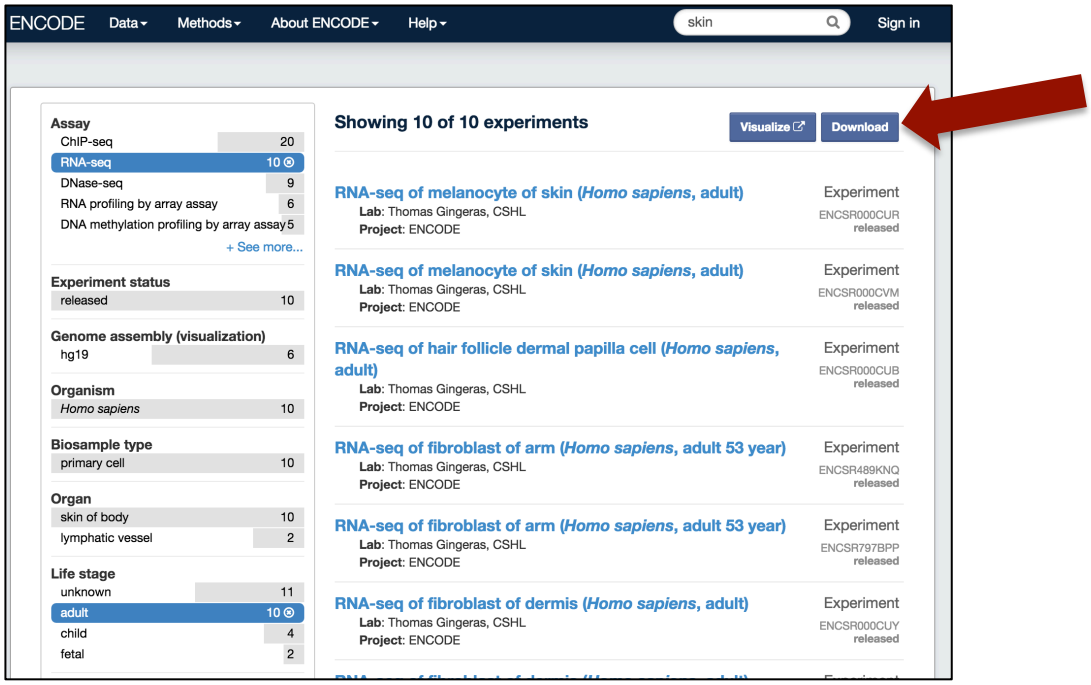
This trackhub was automatically generated from the files and metadata for the experiment - ENCSR000CVM

Accession	File type	Output type	Biological replicate	Download link
ENCFF000KDN	bigBed bedRnaElements	contigs	pooled	Click here
ENCFF000KDQ	bigWig	minus strand signal	1	Click here
ENCFF000KDR	bigWig	minus strand signal	2	Click here
ENCFF000KDT	bigWig	plus strand signal	1	Click here
ENCFF000KDU	bigWig	plus strand signal	2	Click here

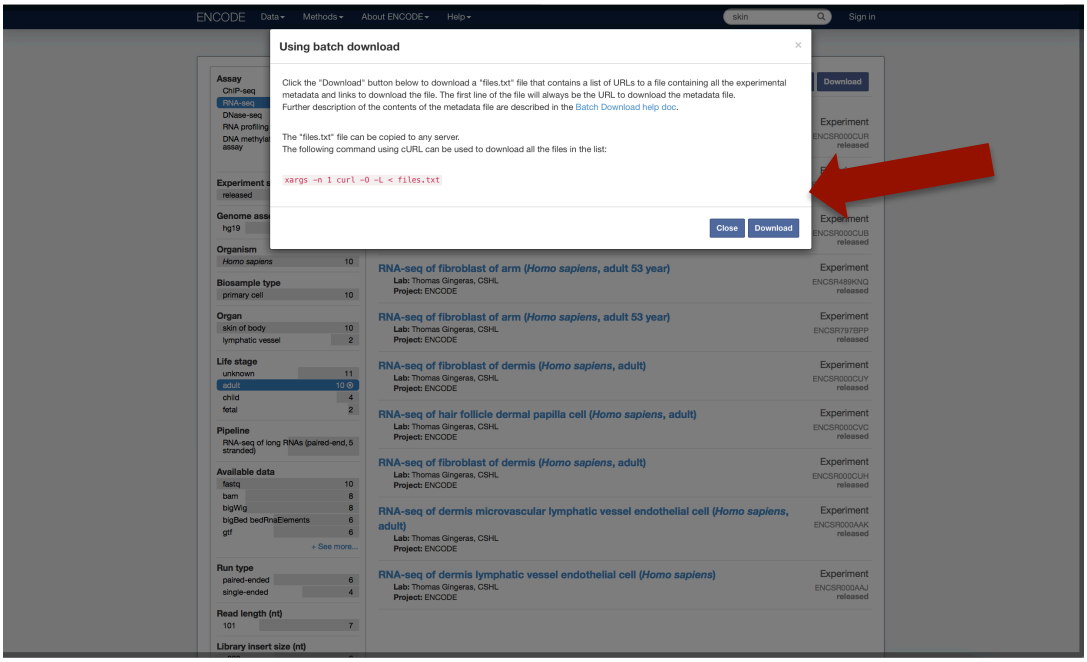
ENCODE data use policy

Demo 5: Batch download data

1. Get the list of assays illustrated in Demo 3.
2. Select “Download” button on the upper right corner.



3. Instructions on the command to batch download files is displayed. Click on ‘Download’ again. This will download a file called ‘files.txt’



4. Open files.txt. This includes a list of links to all the files for those experiments. The first link contains the metadata. You can select the subset of files by using the 'available data' facet.

```
https://www.encodeproject.org/metadata/searchTerm=skin&type=experiment&assay_term_name=RNA-seq&replicates.library.biosample.life_stage=adult/metadata.tsv
https://www.encodeproject.org/files/ENCFF000IGV/@@download/ENCFF000IGV.bigWig
https://www.encodeproject.org/files/ENCFF000IGW/@@download/ENCFF000IGW.bigWig
https://www.encodeproject.org/files/ENCFF000IGY/@@download/ENCFF000IGY.bam
https://www.encodeproject.org/files/ENCFF000IHA/@@download/ENCFF000IHA.bam
https://www.encodeproject.org/files/ENCFF000IHB/@@download/ENCFF000IHB.bigWig
https://www.encodeproject.org/files/ENCFF000IHC/@@download/ENCFF000IHC.fastq.gz
https://www.encodeproject.org/files/ENCFF000IHD/@@download/ENCFF000IHD.bigWig
https://www.encodeproject.org/files/ENCFF000IHF/@@download/ENCFF000IHF.bigBed
https://www.encodeproject.org/files/ENCFF000IHG/@@download/ENCFF000IHG.gtf.gz
https://www.encodeproject.org/files/ENCFF000IHI/@@download/ENCFF000IHI.gtf.gz
https://www.encodeproject.org/files/ENCFF000IHK/@@download/ENCFF000IHK.gtf.gz
https://www.encodeproject.org/files/ENCFF000IHM/@@download/ENCFF000IHM.gtf.gz
https://www.encodeproject.org/files/ENCFF000IHO/@@download/ENCFF000IHO.bigBed
https://www.encodeproject.org/files/ENCFF000IHR/@@download/ENCFF000IHR.gtf.gz
https://www.encodeproject.org/files/ENCFF000IHT/@@download/ENCFF000IHT.gtf.gz
https://www.encodeproject.org/files/ENCFF000IHV/@@download/ENCFF000IHV.fastq.gz
https://www.encodeproject.org/files/ENCFF000IHY/@@download/ENCFF000IHY.fastq.gz
https://www.encodeproject.org/files/ENCFF000IHZ/@@download/ENCFF000IHZ.fastq.gz
https://www.encodeproject.org/files/ENCFF536CFG/@@download/ENCFF536CFG.bigWig
https://www.encodeproject.org/files/ENCFF594YIZ/@@download/ENCFF594YIZ.tsv
https://www.encodeproject.org/files/ENCFF129TCI/@@download/ENCFF129TCI.tsv
```

5. Transfer this file to your server and use
   xargs –n 1 curl –O –L < files.txt

Exercises using the ENCODE Portal

Exercise 1: ChIP-seq assays

- How many ChIP-seq assays are available against H3K27me3 in mouse
- How many unique antibodies are used?
- Of the antibodies used, which one(s) have been fully characterized to current ENCODE standards?

Exercise 2: Controls

- What is the accession for control experiment for ENCSR778SIU?
- How many assays use this control?

Exercise 3: Recently released data

- What was the total number of assays released in June 2015?
- How many of each kind?

Exercise 4: Files

- What are the accession(s) and md5sum(s) to the fastq files for biological replicate 1 in assay ENCSR000AFI?
- What are the accessions of the alignment (bam) files made from the fastq's ?
- Which software tool and version was used to generate them?

Exercise 5: Protein factors

- Which assays have been performed against IGF2BP1?

REST API exercises using the ENCODE Portal

Pre-requisites: download or install these tools to help you get started

- A JSON pretty-printer plugin for your web browser, such as JSONView (for Chrome or Firefox) or JSON Formatter (for Safari)
- A few python modules
  - pip install requests
  - pip install json
  - pip install jsonschema

Help document: https://www.encodeproject.org/help/rest-api/
Sample script: https://github.com/ENCODE-DCC/submission_sample_scripts/blob/master/get.py

Try the exercises listed on the previous page. They can be performed programmatically as well.