Finding ChromHMM States

I. ENCODE ChromHMM states: 3

II. BigBed to Bed 5

III. Roadmap ChromHMM states: 6

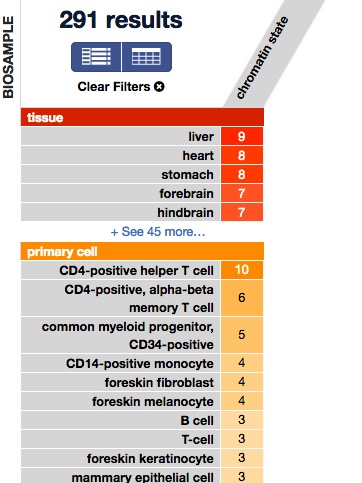
IV. UCSC ENCODE ChromHMM states: 8

# ENCODE ChromHMM states:

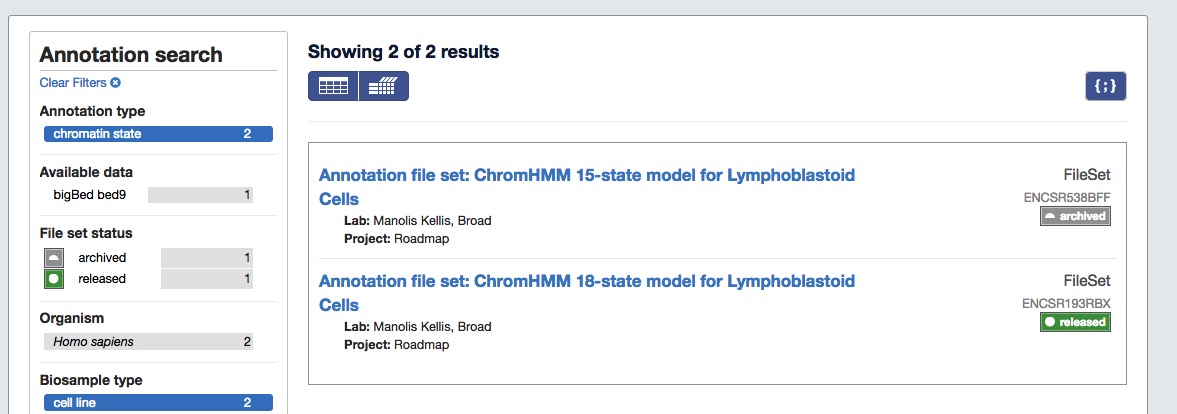
**The List of all ENCODE/Roadmap ChromHMM states can be found here:**

[**https://www.encodeproject.org/matrix/?type=Annotation&annotation\_type=chromatin+state&limit=all&searchTerm=chromHMM**](https://www.encodeproject.org/matrix/?type=Annotation&annotation_type=chromatin+state&limit=all&searchTerm=chromHMM)

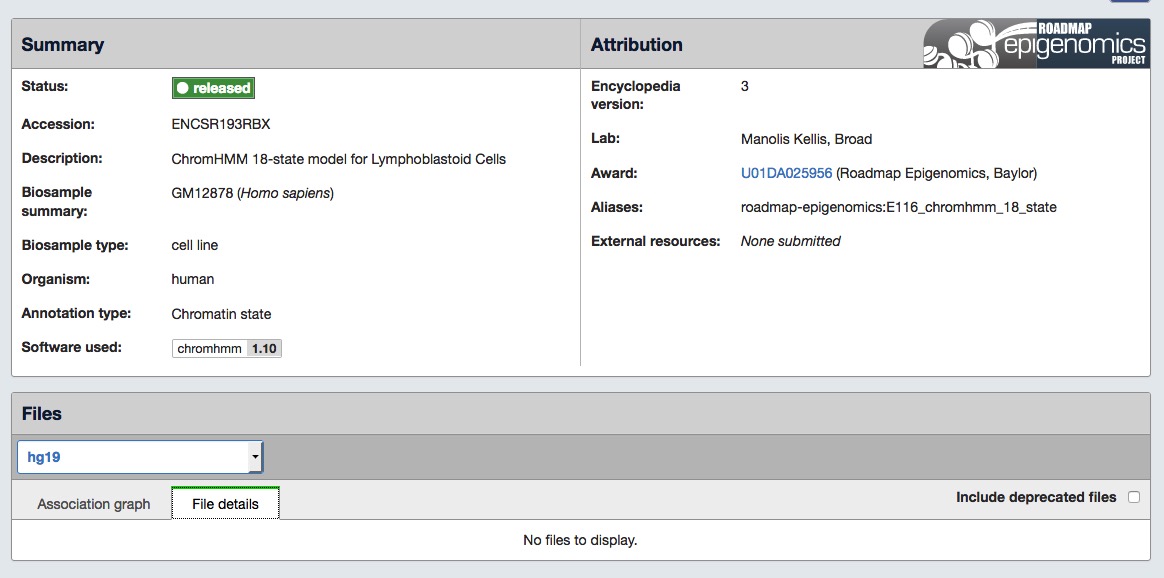
**Search the matrices to identify the cell type of interest:**



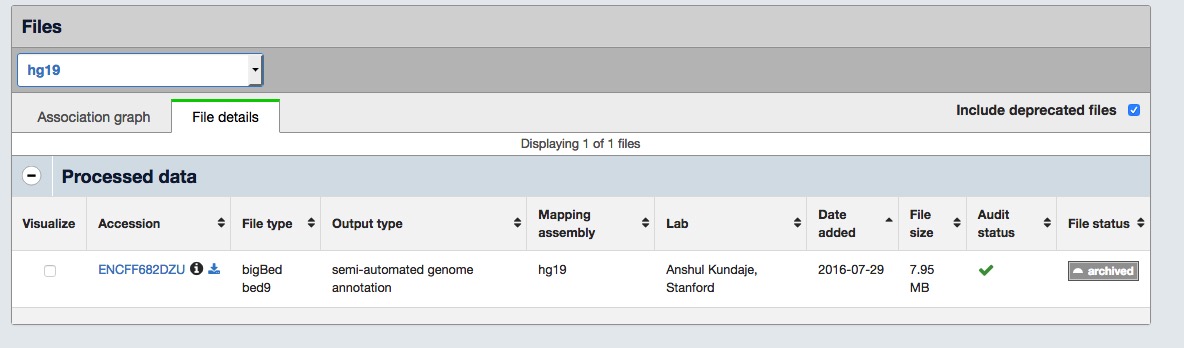
**Select the chromHMM states and navigating through the list of states associated with that cell type:**



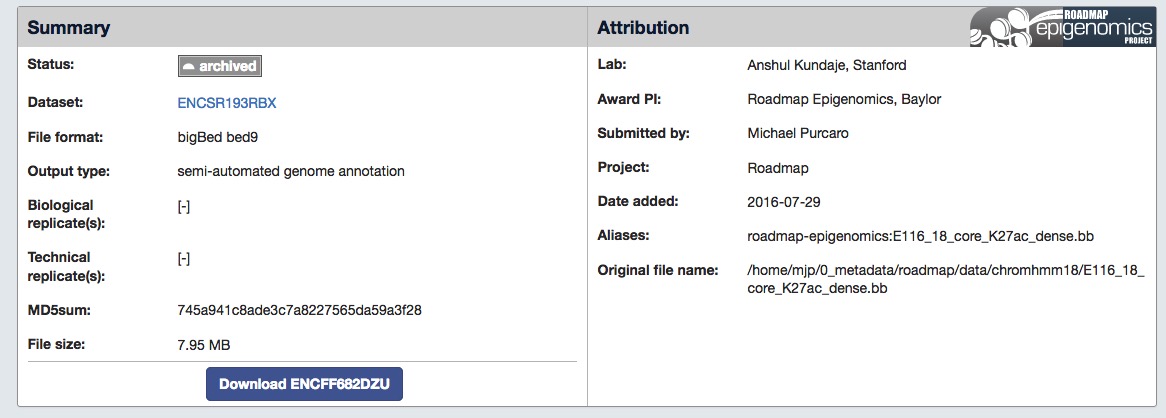
**After selecting the fileset to obtain ChromHMM states from, you will proceed to a summary of the files.**



**Select the file details to identify the bigbed/bed files with the chromHMM states. Some of the files may be hidden because they are listed as “deprecated”. Select “include deprecated”.**



**Select the file, and download the bigbed/bed file on the next page:**

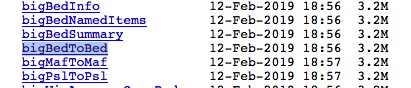


# BigBed to Bed

**UCSC has a tool for extracting bed files from bigbed files. This tool can be obtained from the following URL:**

[**http://hgdownload.soe.ucsc.edu/admin/exe/**](http://hgdownload.soe.ucsc.edu/admin/exe/)

**Identify the operating system that you will need to run the tool and download the bigBedToBed file:**

****

**Add this file to the directory that holds the bigbed files from the ENCODE portal.**

**Set the permissions to run the tool using:**

**chmod +x bigBedToBed**

**Then run the tool by providing it two inputs: the bigbed input file and the desired output bed file.**

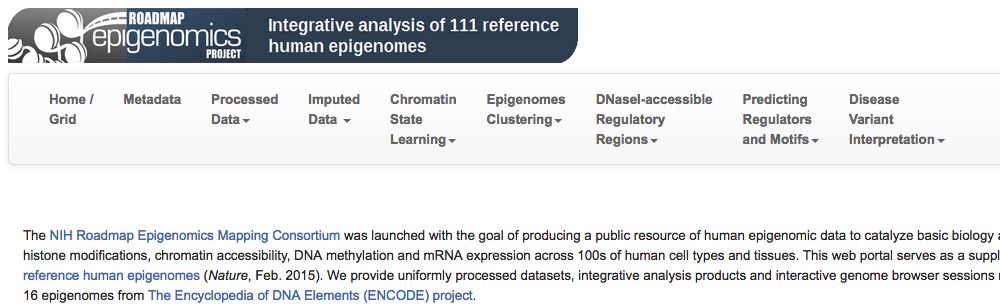
**bigBedToBed input.bigbed output.bed**

# Roadmap ChromHMM states:

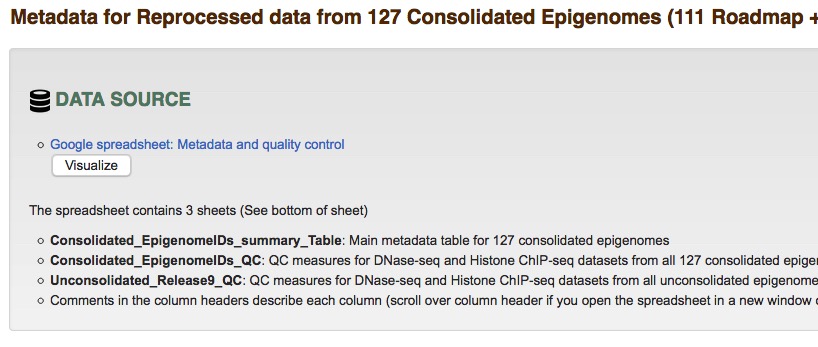
**Roadmap states can be found using the following web portal:**

[**https://egg2.wustl.edu/roadmap/web\_portal/index.html**](https://egg2.wustl.edu/roadmap/web_portal/index.html)

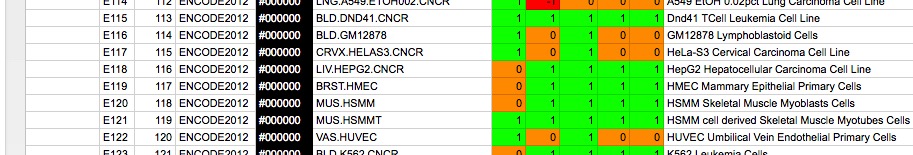
**To get the chromHMM states, we first need to identify the code by looking at the Metadata. Click the “Metadata” link at the top.**



**Next, initialize the visualize the metadata spreadsheet:**

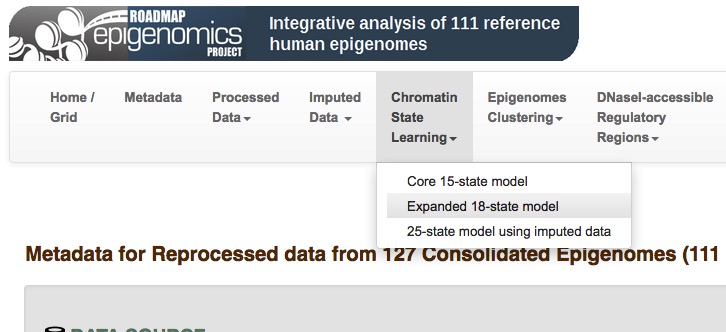


**The full grid of cell types available will then be displayed. Locate the cell type of interest and write down the code associated with it. For example, if we are interested in GM12878 states we would scroll through this list to identify this cell type and write down the code E116.**



**There multiple columns with cell type identifiers, it may be useful to check multiple identifiers to ensure the reference ChromHMM states better match the cell type you are looking for.**

**Navigate to the chromHMM states by traversing to “Chromin State Learning” and “Expanded 18-state model”.**



**Select the download URL to traverse to the list of ChromHMM states:**



**You will then see a list of many files representing all the ChromHMM states that Roadmap has. The file you will be looking for will begin with the code you identified previously and end with “\_mnemonics.bed.gz”.**



**The chromHMM states will be for the hg19 reference genome unless specified as using hg38.**

# UCSC ENCODE ChromHMM states:

**Different ENCODE chromHMM states from UCSC can be found here:**

[**http://hgdownload.soe.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeBroadHmm/**](http://hgdownload.soe.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeBroadHmm/)

**These states may be useful as they include the insulator state (CTCF).**