

Roadmap Epigenomics Workshop

WASHU EPIGENOME BROWSER

Epigenomics 2016

January 31st- February 5th, 2016

Wyndham Grand Resort, Puerto Rico

Presenters: Nicole Rockweiler, Renee Sears, & Ting Wang

Tutorial Overview:

WashU EpiGenome Browser Tutorial

0.	Tutorial key	1
1.	Getting started with the EpiGenome Browser	2
2.	Loading data on the EpiGenome Browser	3
3.	Navigating the EpiGenome Browser	4
4.	Saving, sharing, and retrieving EpiGenome Browser sessions	12
5.	Using Apps on the EpiGenome Browser	14
6.	Creating Custom Tracks on the EpiGenome Browser	21
7.	Training and support	24

Roadmap EpiGenome Browser Tutorial

8.	Loading data on the Roadmap EpiGenome Browser	25
9.	Navigating the Roadmap EpiGenome Browser	26
10.	Configuring the Roadmap EpiGenome Browser	32
11.	Saving, sharing, and retrieving Roadmap EpiGenome Browser sessions	36
12.	Creating Custom Tracks on the Roadmap EpiGenome Browser	37
13.	Annotating non-coding variants with epigenomic data	38

WashU EpiGenome Browser:

<http://epigenomegateway.wustl.edu/browser>

Roadmap EpiGenome Browser:

<http://epigenomegateway.wustl.edu/browser/roadmap/>

0. Tutorial key

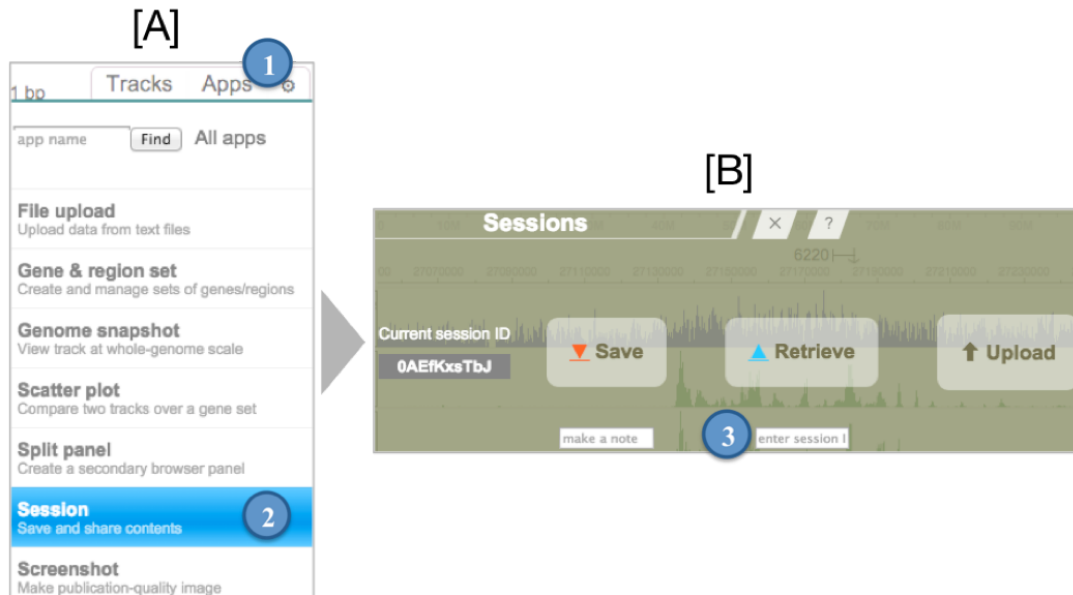
- To follow along with this tutorial, use instructions marked by ➤. We have provided screenshots for guidance (ordered by bracketed alphabets: [A], [B], ...)
- Click in the order of the numbered circles on the screenshots to follow this tutorial.



0.1 How to catch up

0.1.1 WashU EpiGenome Browser tutorial

- To catch up with any section of the *WashU EpiGenome Browser tutorial*, use the provided **session ID** and the **session status**, listed at the end of that section. For this, follow these steps:
 - [A] Click the **Apps** button to display the Apps menu. Then click **Session**.
 - [B] Under the **Retrieve** button, enter the **session ID** and then click the **Retrieve** button.
 - This will generate a list of **session statuses**. Click the session of interest.



0.1.2 Roadmap EpiGenome Browser tutorial

- To catch up with any section of the *Roadmap EpiGenome Browser tutorial*, click the provided **session link**, listed at the end of that section.

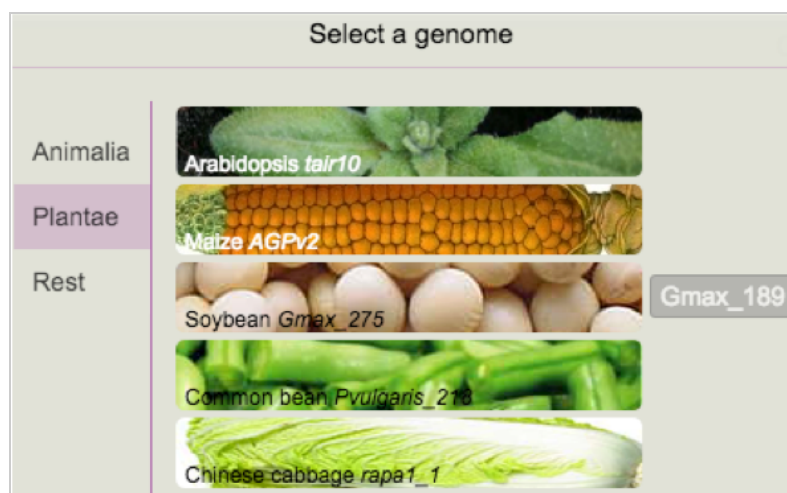
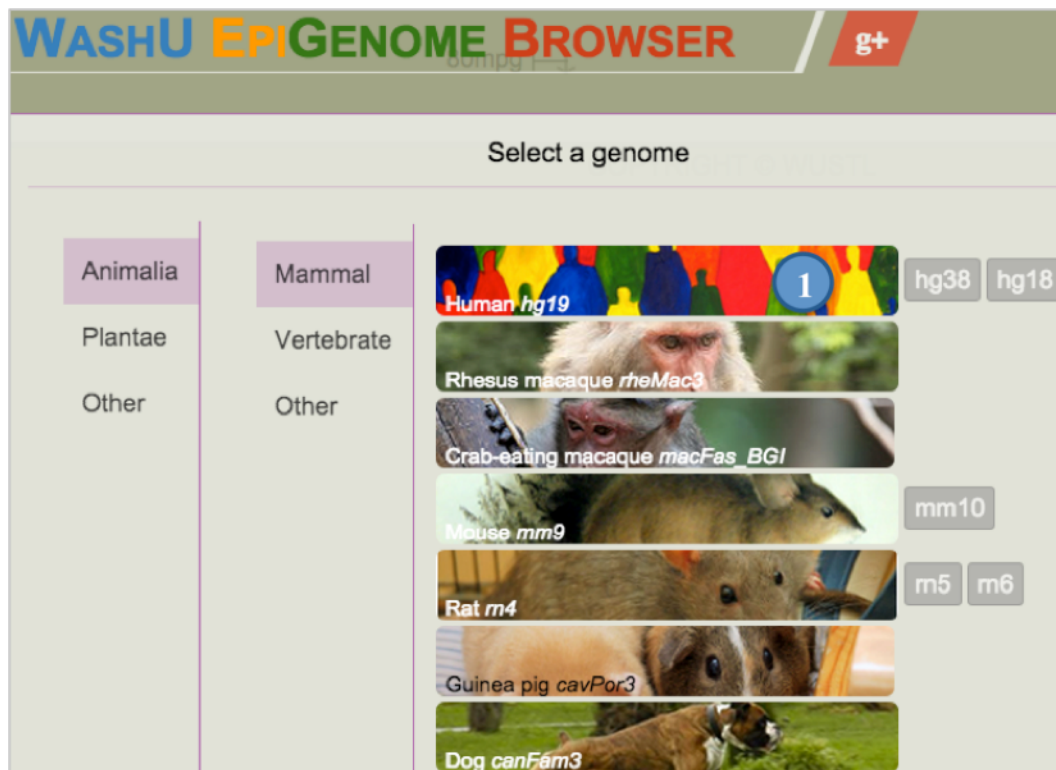
1. Getting started with the EpiGenome Browser

1.1. Loading the browser

- Access the browser at <http://epigenomegateway.wustl.edu/browser>

1.2. Selecting the genome assembly of interest

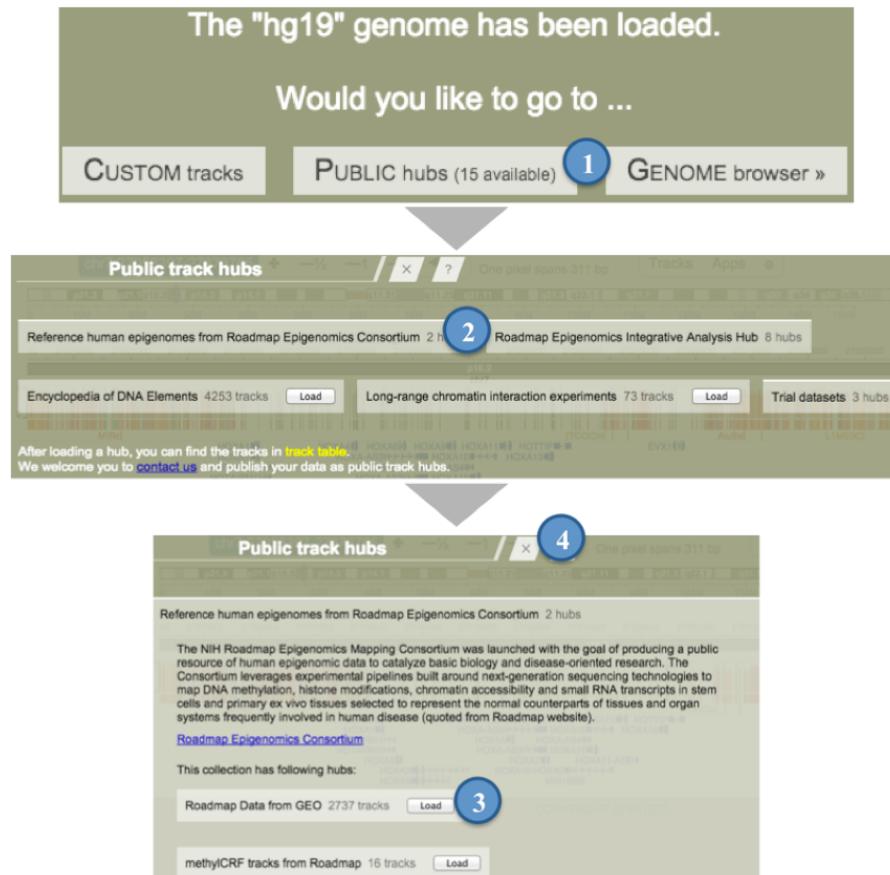
- For the purpose of this tutorial, please select **Human hg19**.
- However, you can see that many other genomes are also available.



2. Loading data on the EpiGenome Browser

2.1. **Data hubs:** A data hub is a collection of datasets (also called **tracks**) that can be viewed on the browser.

- [A] Click the **PUBLIC hubs (15 available)** button to view all the available public datasets on the EpiGenome Browser.
- [B] Click the **Reference human epigenomes from Roadmap Epigenomics Consortium (2 hubs)** button.
- [C] This will generate a list of available data hubs. Click the **Load** button on Roadmap Data from GEO (2737 tracks) box. Once the datasets are loaded, exit the data hub menu by clicking the **X** at the top-right of the panel or pressing **Esc**.



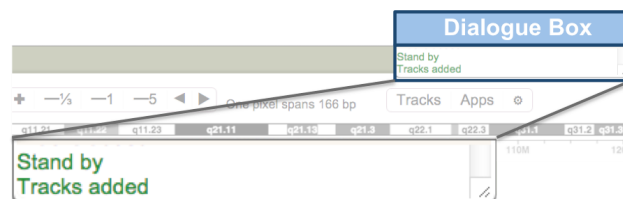
3. Navigating the EpiGenome Browser

3.1. Layout of the EpiGenome Browser



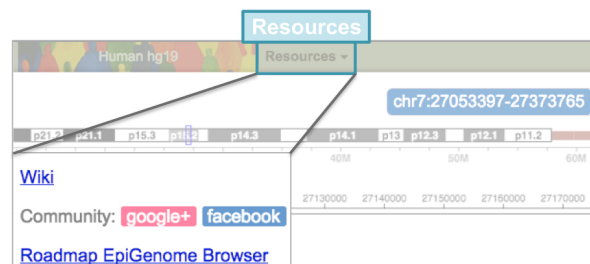
3.1.1. Dialogue box

- This box logs all the actions performed by the EpiGenome browser. Any errors the browser encounters are displayed.



3.1.2. Resources

- This window contains links to the documentation on the browser ([Wiki](#)); social-media discussion forums (on [Google+](#) and [Facebook](#)) for the browser; and the [Roadmap EpiGenome Browser](#), a companion browser for visualizing and analyzing data from the Roadmap Epigenomics Consortium.



3.1.3. WashU EpiGenome Browser

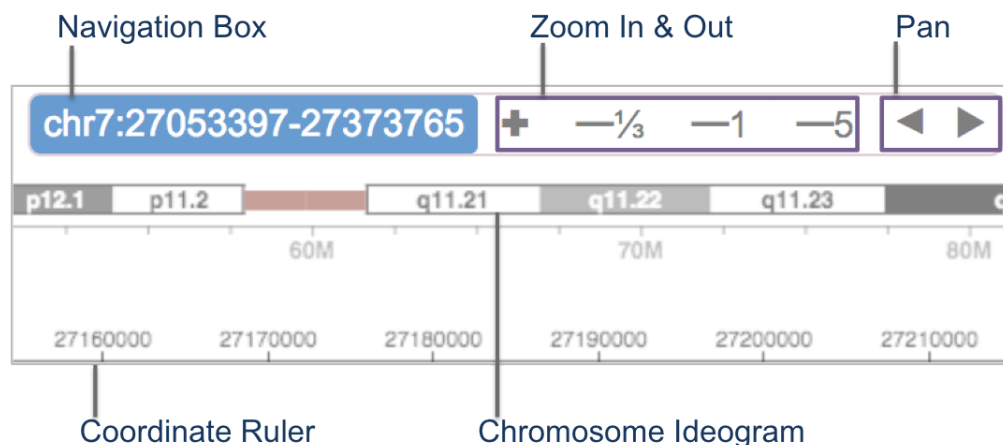
- This region contains the data tracks aligned to the genome along with reference genes and a chromosome ideogram. Additional public track hubs, annotation tracks, and custom tracks can be added by the user.



3.2. Navigating the genome on the EpiGenome Browser

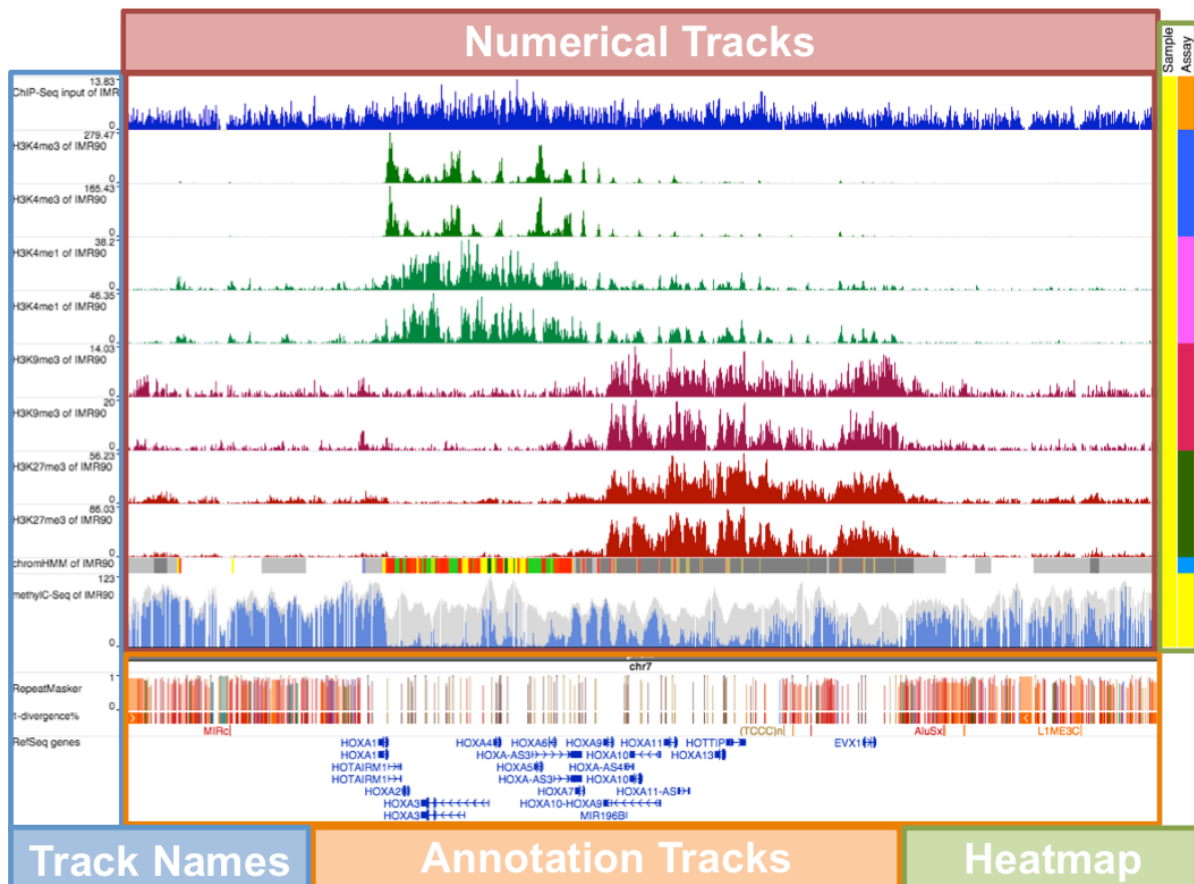
3.2.1. Genome navigation controls

- The controls enable exploration across the genome using the **zoom in**, **zoom out**, and **pan** buttons.
- Alternatively, click the **navigation box** to enter the genomic coordinates of a region of interest or a gene name. Click **Go** to move to the new region.



3.2.2. Data visualization

- **Numerical tracks** represent the density of reads from sequencing experiments aligned to the genome.
- **Track names** for each track are listed on the left-hand side of the track.
- On the right-hand side of the tracks is a **heatmap**, which represents the **metadata** for the tracks. Metadata are terms used for annotating tracks with experimental and sample information.
- Below the numerical tracks are **annotation tracks**, including annotations of genes and transposable elements.



3.3. Adding datasets on the EpiGenome Browser

3.3.1. Searching for datasets using the facet table: The facet table organizes all the loaded datasets into a table to allow the user to search for datasets using metadata terms. **Rows** represent samples and **columns** represent assay types. **Cells** with numbers represent the number of datasets for a particular sample and assay combination. For example, in the cell with numbers **11/130**, **11** represents the number of datasets currently loaded on the browser, while **130** represents the total number of datasets available.

- [A] To display the facet table, click the **Tracks** button near the top of the browser to display the Tracks menu. Click the teal box at the top of the menu to display the facet table.

- [B] To search for datasets, click on **Fetal Cells/Tissues** and then select **Fetal Brain**. Click the cell corresponding to the **Epigenetic Mark** column and the **Fetal Brain** row to list the datasets in this category.
- [C] To add the datasets, first select the datasets **H3K9me3 of Fetal Brain** and **H3K4me3 of Fetal Brain**. Then, click the **Add 2 tracks** button.
- To exit this window, click the **X** at the top-right of the panel or press **Esc**.

[A]

1 Tracks

2737 TOTAL / 11 SHOWN

CLICK FOR TRACK TABLE

track name Find ?

PUBLIC track hubs

ANNOTATION tracks

CUSTOM tracks

[B]

Experimental assay tracks

Row Sample Column Assay

Epigenetic Mark	Expression	Long Range Interaction	Other Assays	Transcription Regulator	
0/1225	0/80	n/a	n/a	n/a	Adult Cells/Tissues
0/4	n/a	n/a	n/a	n/a	Cancer Cells
0/703	0/30	n/a	n/a	n/a	ES/iPS Cells
					Fetal Cells/Tissues
0/14	n/a	n/a	n/a	n/a	Fetal Adrenal Gland
					Fetal Brain
0/32	0/4	n/a	n/a	n/a	Fetal Brain
0/4	n/a	n/a	n/a	n/a	brain, fetal day122 U
0/1	n/a	n/a	n/a	n/a	brain, fetal day112 U
0/1	n/a	n/a	n/a	n/a	brain, fetal day117 F
0/3	n/a	n/a	n/a	n/a	brain, fetal day122 M
0/2	n/a	n/a	n/a	n/a	brain, fetal day85 F
0/3	n/a	n/a	n/a	n/a	brain, fetal day96 F
0/1	n/a	n/a	n/a	n/a	brain, fetal day142 F
0/3	n/a	n/a	n/a	n/a	brain, dorsal neocortex, fetal
0/30	n/a	n/a	n/a	n/a	Fetal Heart
0/27	n/a	n/a	n/a	n/a	Fetal Intestine, Large
0/23	n/a	n/a	n/a	n/a	Fetal Intestine, Small
0/73	n/a	n/a	n/a	n/a	Fetal Kidney
					Fetal Liver
11/130	0/12	n/a	n/a	n/a	Fetal Lung
					Fetal Membrane
0/72	0/31	n/a	n/a	n/a	Fetal Muscle
0/1	0/1	n/a	n/a	n/a	Fetal Ovary
0/14	n/a	n/a	n/a	n/a	Fetal Placenta
0/120	0/18	n/a	n/a	n/a	Fetal Skin
0/5	0/2	n/a	n/a	n/a	Fetal Spinal Cord
0/1	n/a	n/a	n/a	n/a	Fetal Spleen
0/22	n/a	n/a	n/a	n/a	Fetal Stomach
0/2	n/a	n/a	n/a	n/a	Fetal Testes
0/21	n/a	n/a	n/a	n/a	Fetal Thymus
0/14	0/8	n/a	n/a	n/a	Placenta

Remove all

To get additional tracks, load public track hubs.

[C]

6 H3K9me3 of Fetal Brain bigwig i

7 DNase hypersensitivity of Feta... bigwig i

7 H3K4me3 of Fetal Brain bigwig i

H3K27me3 of Fetal Brain bigwig i

ChIP-Seq input of Fetal Brain bigwig i

DNase hypersensitivity of Feta... bigwig i

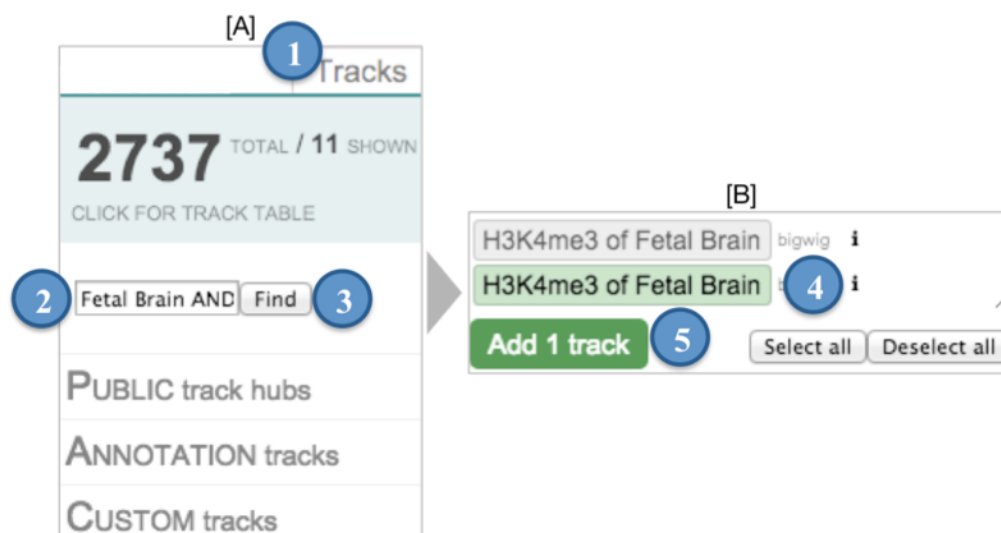
DNase hypersensitivity of Feta... bigwig i

DNase hypersensitivity of Feta... bigwig i

8 Add 2 tracks Select all Deselect all

3.3.2. Searching for datasets using the track search box

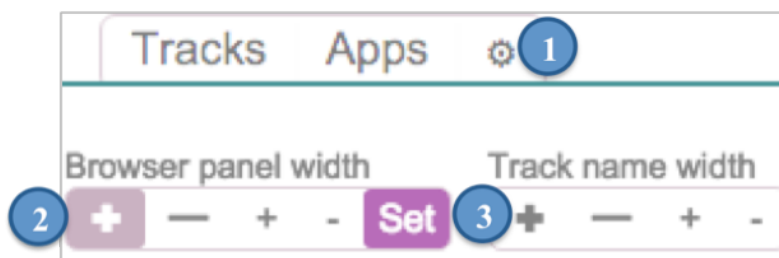
- [A] Click the **Tracks** button. In the search box, type **Fetal Brain AND H3K4me3**, and then click **Find**. The **AND** is a logical operator so must be uppercased, but search terms are otherwise case-insensitive.
- [B] This will list all replicates for this sample and assay type. Only one of the two boxes listed can be selected since the other replicate has already been loaded as indicated by the light gray box. Click the second **H3K4me3 and Fetal Brain** and then click **Add 1 track**.
- Repeat this process to add one more replicate track for **Fetal Brain AND H3K9me3** and click outside the floating window to get back to the browser.



To catch up to this point of the workshop, click on the Apps menu and then select Session. Enter the session ID **PRworkshop** in the Retrieve text box and click the status named **Add datasets**. For more details about sessions, refer to section 4.3 on page 13.

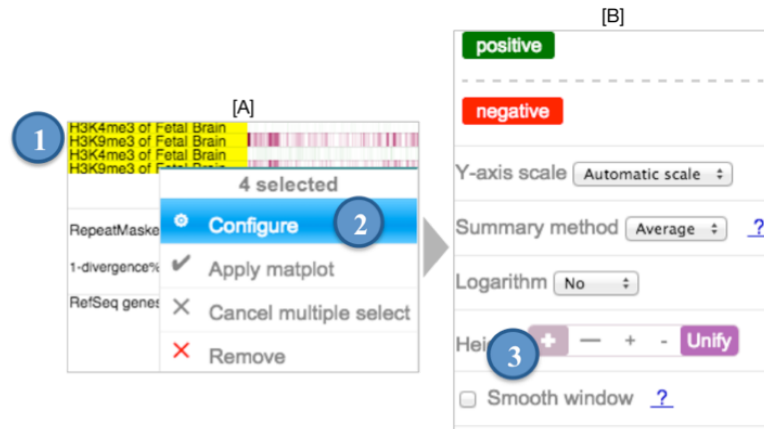
3.3.3. Formatting the view range

- To adjust the **browser width**, click the button near the top of the browser. The + and - buttons increase or decrease the browser width, respectively. Click + then **Set** to change the browser width.
- Similarly, the **width of the track names** can be changed using the + and - buttons.

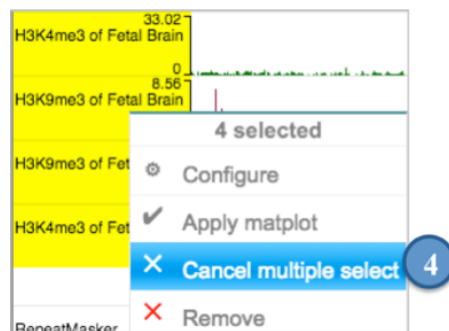


3.3.4. Changing the height of tracks

- [A] To select multiple tracks, hold down the **shift key** and click the names of the **four Fetal Brain tracks** that were just added. The track name will be highlighted in yellow when selected. Right-click the selected tracks then click **Configure**.
- [B] Click the **+** sign repeatedly to increase the height of the tracks. This does not change the y-axis scale.
- Click outside the configuration menu to close the menu.



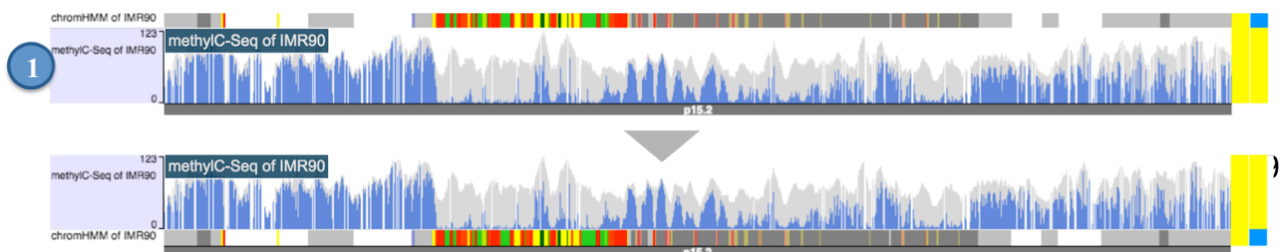
- To deselect the four tracks, right-click the highlighted track names and select **Cancel multiple select**.



3.3.5. Reordering tracks

3.3.5.1 Reordering tracks by clicking-and-dragging

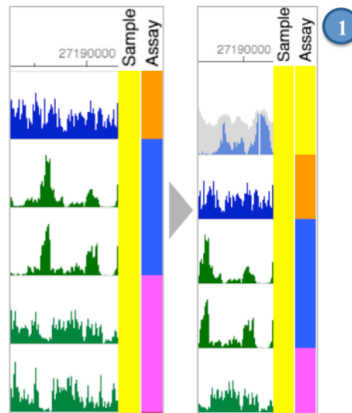
- Tracks can be reordered by clicking on a track name and moving that track to a new position.
- Select the **methyIC-Seq of IMR90** track and move the track above the **chromHMM** track.



To catch up to this point of the workshop, click on the Apps menu and then select Session. Enter the session ID **PRworkshop** in the Retrieve text box and click the status named **Configure browser 1**. For more details about sessions, refer to section 4.3 on page 13.

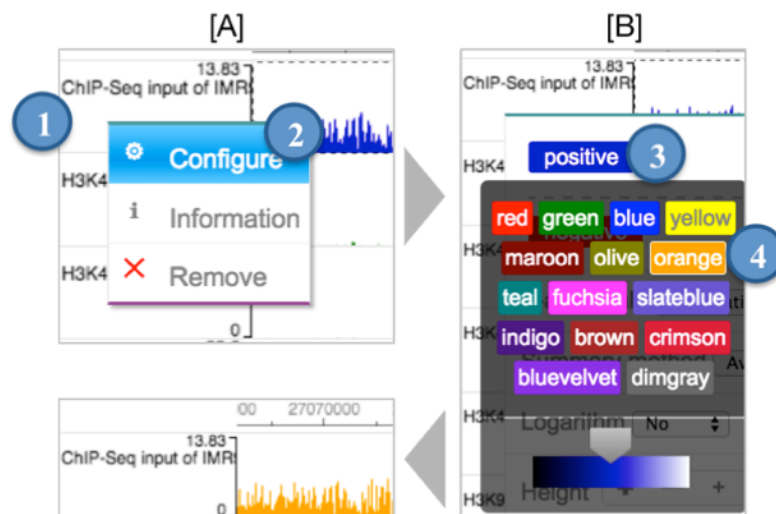
3.3.5.2 Reordering by metadata terms

- Alternatively, tracks can be reordered based on assay type, click the **Assay** term above the heatmap.



3.3.6. Changing the color of tracks

- [A] Select the **ChIP-Seq input of IMR90** track by right-clicking on the track name and then click **Configure**.
- [B] In the configuration menu, click the **positive** button and then select the **color** of your choice to change the color of the track.

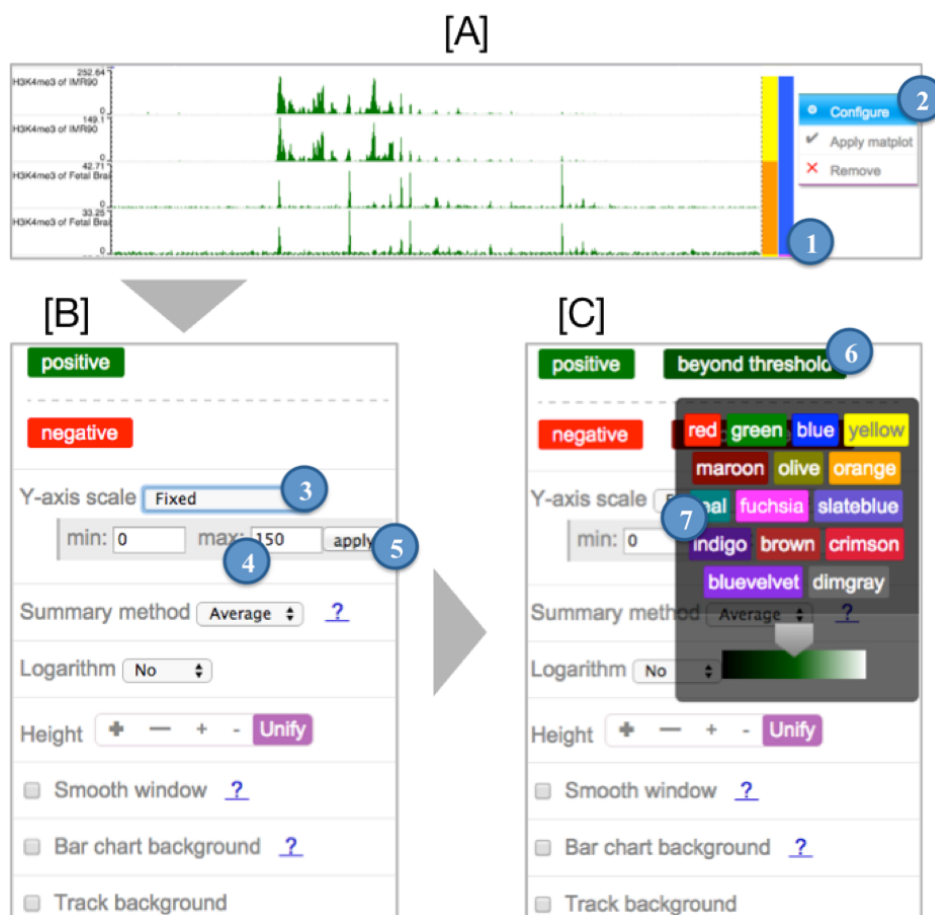


- Click outside the configuration menu to close the menu.

To catch up to this point of the workshop, click on the Apps menu and then select Session. Enter the session ID **PRworkshop** in the Retrieve text box and click the status named **Configure browser 2**. For more details about sessions, refer to section 4.3 on page 13.

3.3.7. Changing the y-axis scale for tracks

- [A] To select the four H3K4me3 tracks, right-click the **blue rectangle** in the heatmap (under the Assay metadata term column) and click **Configure**.
- [B] To change the height of the selected tracks, click the **y-axis scale** drop-down menu and then select **Fixed**. This will generate textboxes in which a y-axis range can be entered. Type **150** in the max textbox. Click **apply** to change the y-axis scale.
- [C] To change the color used to mark regions where the data is beyond the range of the y-axis, click the **beyond threshold** button. This will generate a window with color options; select the **color** of your choice.



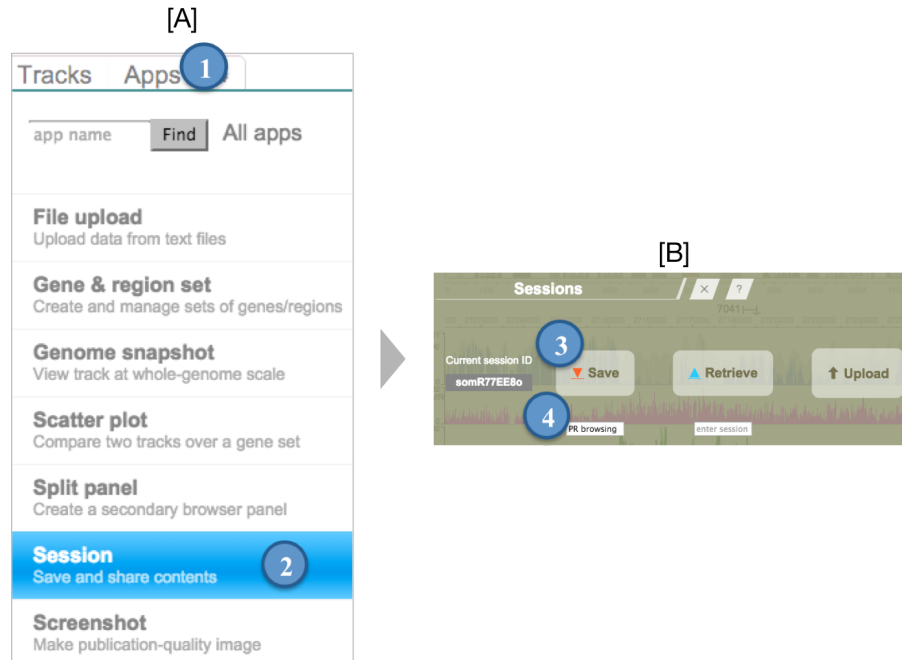
- Click anywhere outside the **Configure** box to close the menu.

To catch up to this point of the workshop, click on the Apps menu and then select Session. Enter the session ID **PRworkshop** in the Retrieve text box and click the status named **Configure browser 3**. For more details about sessions, refer to section 4.3 on page 13.

4. Saving, sharing, and retrieving EpiGenome Browser sessions

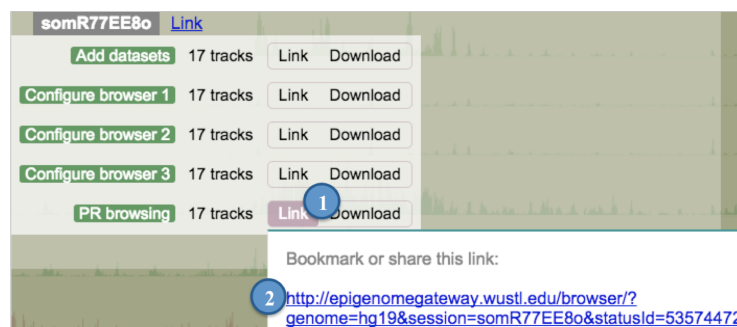
4.1. Saving sessions

- [A] Click the **Apps** button and then select **Session**.
- [B] Type **PR browsing** in the dialogue box under the Save button and then click **Save**. Click the **X** at the top-right of the panel to close the menu



4.2. Sharing Links for Collaboration

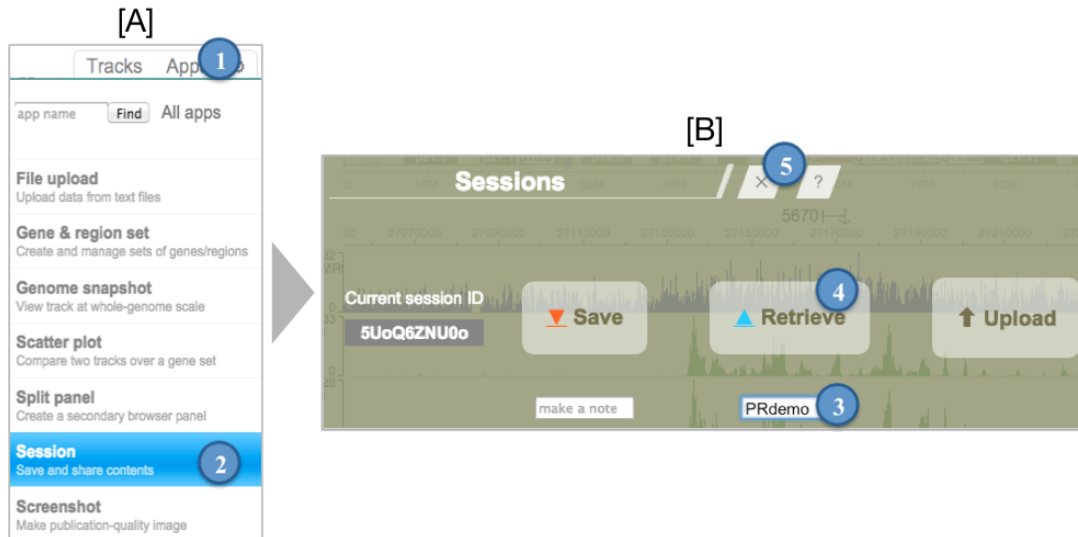
- To generate the link for this session, click the **Link** button next to the session name.



This will create a floating window with a **URL** that can be shared.

4.3. Retrieving Sessions

- Launch a new instance of the EpiGenome browser in a new window by going to <http://epigenomegateway.wustl.edu/browser>. Select the **Human hg19** genome assembly. Click the **GENOME Browser »** button to proceed to the browser.
- [A] Click the **Apps** button, and then select **Sessions**.
- [B] Enter the session ID **PRdemo** in the search box under the Retrieve button, and then click **Retrieve**. This will list all the available statuses under this session ID. Click on the **Start** status. This will retrieve a new session that will be used for the last section of this tutorial.

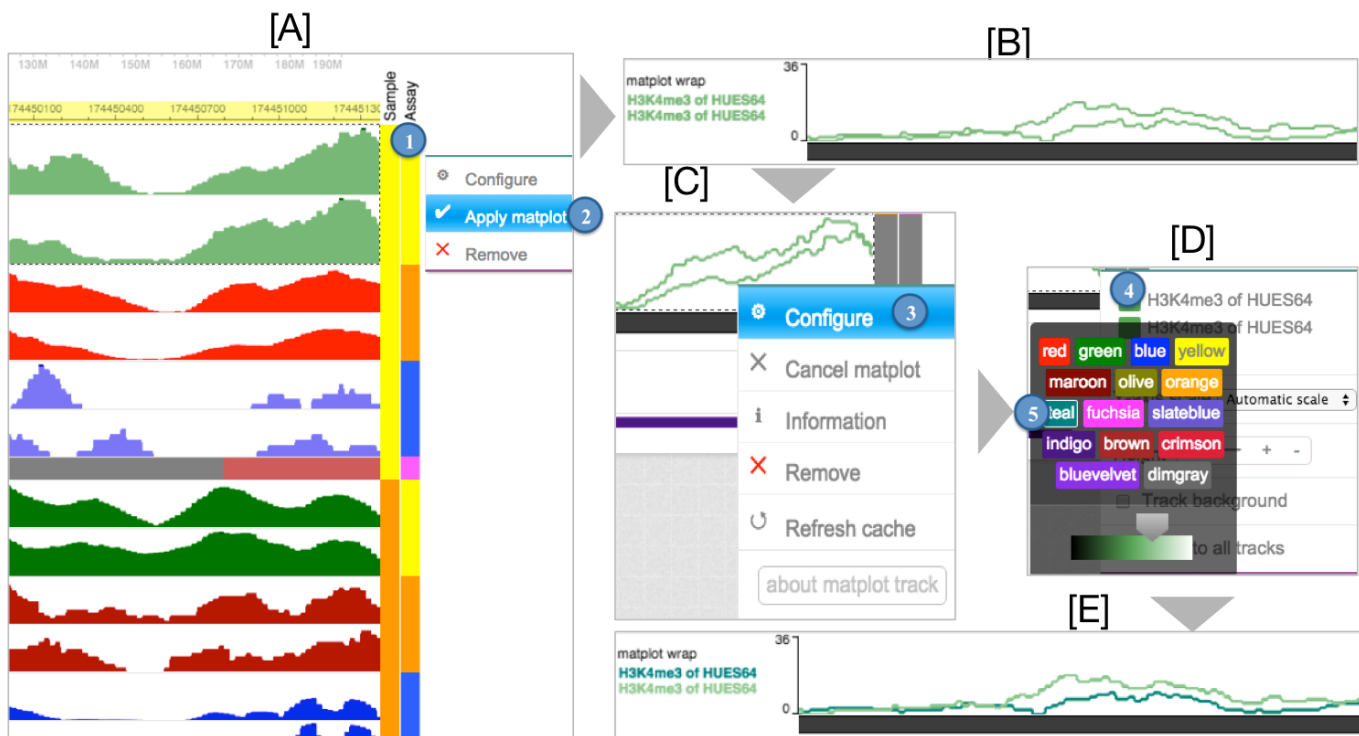


- Click the **X** at the top-right of the panel or press **Esc** to close the menu.
- In this session, we will be using Roadmap human HUES64 cells and left ventricle tissue data to analyze bivalent promoters of genes associated with heart development.
- The **Start** session contains pre-loaded and formatted H3K4me3, H3K27me3, RNA-Seq, and chromHMM tracks for the left ventricle (heart) and ESC line HUES64.

5. Using Apps on the EpiGenome Browse

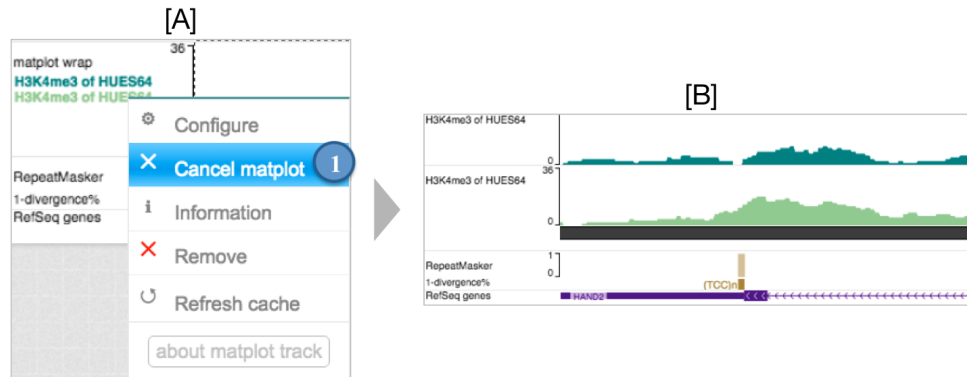
5.1. Matplot: A matplot (line plot) is used to **compare** two or more **numerical** tracks by plotting datasets on the same **X** and **Y axes**. Move to the gene ***HAND2*** in the browser.

- [A] **Right-click** the first yellow box in the Assay column of the metadata heatmap (H3K4me3 of HUES64) and click **Apply matplot**.
- [B] This will generate a matplot track.
- [C] To change the color of the numerical tracks, **right-click** the name of the matplot track and click **Configure**.
- [D] Click the **green-box** beside the first track. Select **teal** to change the color of light green to teal.
- [E] This results in a track where the H3K4me3 data from the two HUES64 replicates share the same y-axis and can be easily compared. It is evident from this view that these tracks both show the same pattern of H3K4me3 across the *HAND2* gene.



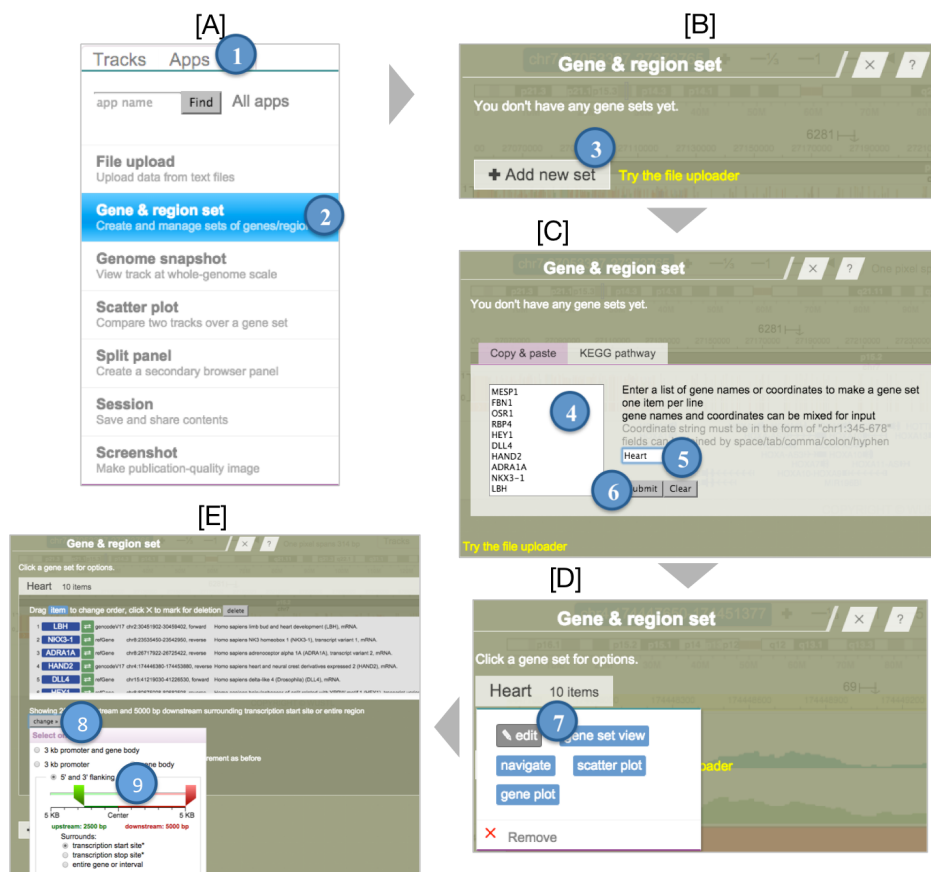
To catch up to this point of the workshop, click on the Apps menu and then select Session. Enter the session ID **PRdemo** in the Retrieve text box and click the status named **Matplot**. For more details about sessions, refer to section 4.3 on page 13.

- [A] To exit from the matplot view, **right-click** the matplot track name and click **Cancel matplot**.
- [B] This ungroups the tracks, but keeps their new colors. These tracks can then be moved back to the original order.



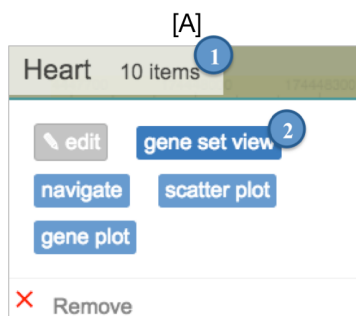
5.2. Gene & region set: The Gene & region set app is used to visualize multiple genomic regions in parallel. The app enables track data to be displayed over nonadjacent regions on a chromosome or even regions on different chromosomes. We are going to use this app to look at a couple of genes with bivalent promoters in the ESC line HUES64, but active transcription in the left ventricle sample.

- Go to the Epigenomics2016 browser workshop page <http://epigenomegateway.wustl.edu/support/workshop2016Epigenomics.html>. Scroll down to the Heart Gene Set and copy the 10 genes in the table.
- [A] Click the **Apps** button and then select **Gene & region set**.
- [B] To add a set of genes click **+ Add new set**.
- [C] Paste the 10 genes into the large textbox, name the set “Heart”, and click **submit**.
- [D] Select the **Heart** button and click **edit**.
- [E] Under the list of genes, click the **change »** button to modify the view range. First select the radio button for **5' and 3' flanking** and then select the radio button for **entire gene or interval**. Adjust the **green** slider to 2.5 KB and the **red** to 5 KB. Click outside the window to exit this window.



To catch up to this point of the workshop, click on the Apps menu and then select Session. Enter the session ID **PRdemo** in the Retrieve text box and click the status named **Gene Set**. For more details about sessions, refer to section 4.3 on page 13.

- [A] To view the “Heart” gene set go to apps Gene & region set, select the **Heart** button, and then click **gene set view**.



- Change the y-axis scale on all RNA-Seq tracks to a fixed scale with a max of 150.
- Reconfigure the track name width to see the full name of each track.
- This results in a tiled view of 10 genes arranged beside each other. By setting the same y-axis scale for all RNA-seq data, the expression of each gene in the set can be better appreciated. Additionally, looking closely at each gene reveals that the HUES64 tracks have strong H3K4me3 peaks over the TSS and promoter region along with H3K27me3, but the H3K27me3 signal is not seen in the left ventricle tracks.



To catch up to this point of the workshop, click on the Apps menu and then select Session. Enter the session ID **PRdemo** in the Retrieve text box and click the status named **Gene Set View**. For more details about sessions, refer to section 4.3 on page 13.

- To exit the gene set view, click the **Showing entire set | X** button near the top left side of the browser.

Showing entire set | X +

5.3. Navigate regions: The navigate regions app allows the user to navigate across the regions specified in a gene & region set.

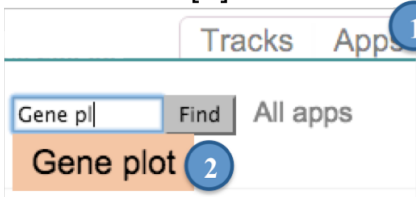
- [A] Click the **Apps** button and then select **Gene & region set**.
- [B] Select the **Heart** button and click **navigate**.
- [C] A panel will now be displayed with the 10 genes in the Heart gene set. Click **HAND2** to navigate to the *HAND2* gene.

To catch up to this point of the workshop, click on the Apps menu and then select Session. Enter the session ID **PRdemo** in the Retrieve text box and click the status named **Navigate**. For more details about sessions, refer to section 4.3 on page 13.

5.4. Gene plot: The Gene plot app summarizes the data distribution across multiple regions by splitting the regions into bins.

- Change the gene view area in each Gene & region set to **gene body only**. This app can also be accessed by clicking on a gene set.
- [A] Click the **Apps** button. In the search box, type **Gene plot** and select the **Gene plot** app.
- [B] To make a gene plot, follow these steps:
 - In the **0 – Gene set** section, click the **Choose a gene set** button. Select the **Heart** gene set.
 - In the **1 – Data track** section, click the **Select numerical track »** button and then click **H3K4me3 of HUES64**.
 - In the **2 – Graph type** section, use the default graph type, a box plot. Click the **plot average values** check-box. To change the box color, click the **box color** button and then select the color of your choice.
 - Click the **Make gene plot** button to create the gene plot.

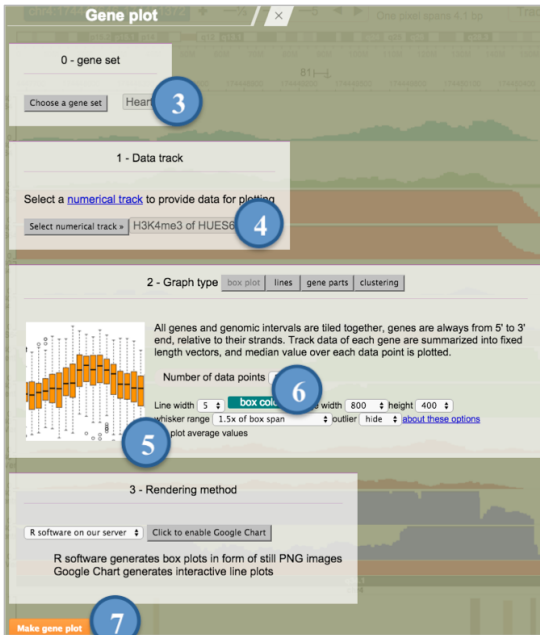
[A]



1: Click the **Apps** button.
2: Click the **Gene plot** app.

➔

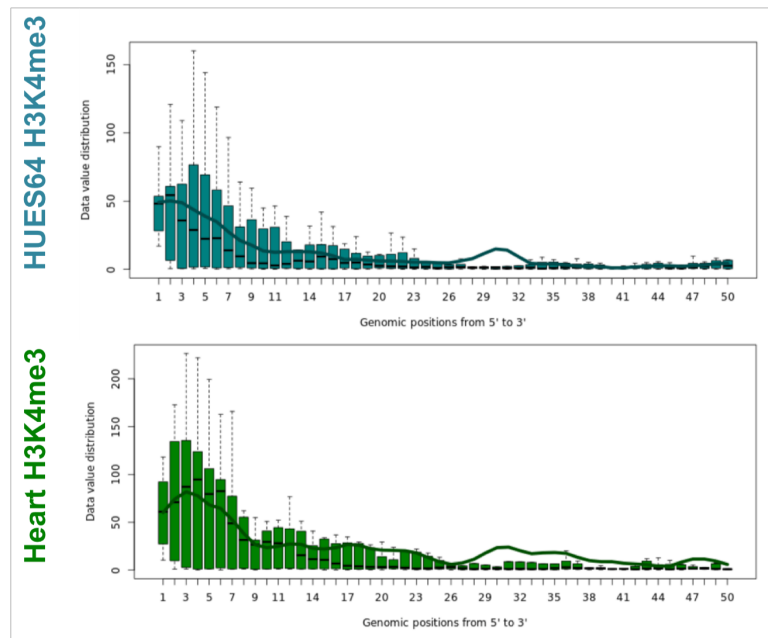
[B]



3: Click **Choose a gene set** and select **Heart**.
4: Click **Select numerical track »** and select **H3K4me3 of HUES64**.
5: In the **2 - Graph type** section, click **plot average values**.
6: Click **box color** to choose a color.
7: Click **Make gene plot**.

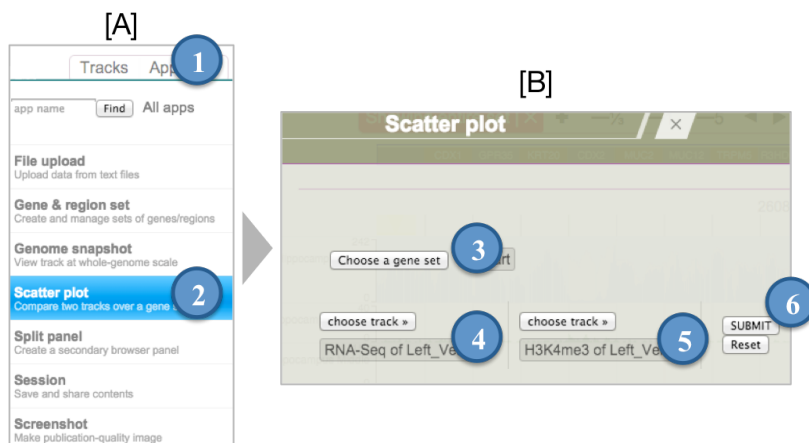
- Repeat these steps with the Heart gene set and the **H3K4me3 of Left Ventricle** dataset. Use a different color for the boxplot.

- Comparison of the two gene plots reveals that the H3K4me3 signal is the strongest in the 5' upstream region of the gene with an appreciable depletion when moving towards the 3' end of the gene.

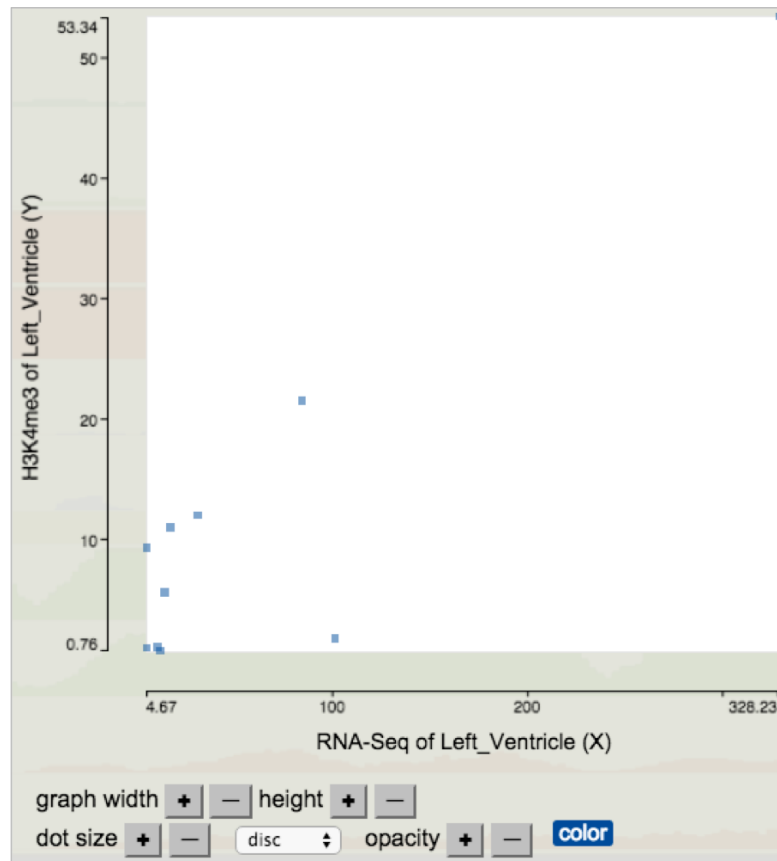


5.5. Scatter plot: The Scatter plot app is used to assess the relationship between two different datasets across multiple genomic regions.

- [A] Click the **Apps** button and then select **Scatter plot**.
- [B] To select the genomic regions to plot, click the **Choose a gene set** button and then click the gene set **Heart**. To select the numerical track to plot on the x-axis, first click the **Choose track »** button above the **for X axis** button then select the second **RNA-Seq of Left Ventricle** track. To select the numerical track to plot on the y-axis, first click the **Choose track »** button above the **for Y axis** button then select the second **H3K4me3 of Left_Ventricle** track. Click **SUBMIT** to create the plot.



- The resulting scatter plot shows the correlation between gene expression and H3K4me3 for the left ventricle across the gene set. H3K4me3 is associated with active promoters, so the correlation is expected.



- To exit the app, click the **X** at the top-right of the panel or press **Esc**.

6. Creating Custom Tracks on the EpiGenome Browser

Now that you are more familiar with the WashU EpiGenome Browser it is time to explore your data.

6.1. Creating custom tracks

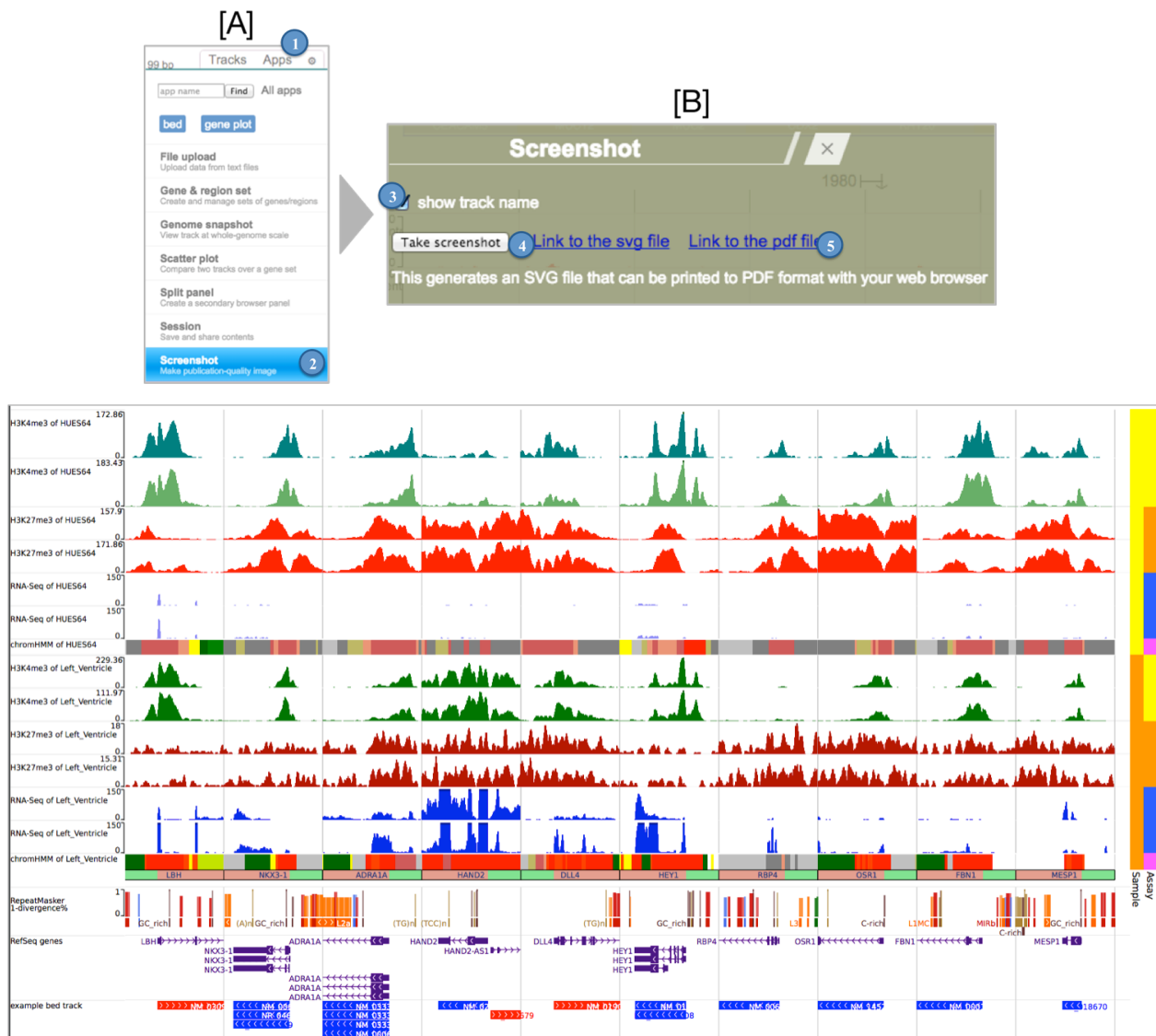
- [A] Click the **Tracks** button then click **Custom Tracks**.
- [B] Click the **+ Add new tracks** button.
- [C] The browser supports many file formats for custom tracks. Data must be hosted on a web server accessible to the browser. Click **BED** to create your first custom track.
- [D] Click **Use example** and then **SUBMIT**. The track now appears in the browser window, click **X** to return to the browser window.



To catch up to this point of the workshop, click on the Apps menu and then select Session. Enter the session ID **PRdemo** in the Retrieve text box and click the status named **Custom Track**. For more details about sessions, refer to section 4.3 on page 13.

6.2. Saving EpiGenome Browser screenshots

- [A] Finally, you may want to take a screenshot of your final tracks, a gene of interest, or a gene set. Reapply the Heart gene set view with 2.5 KB upstream and 5 KB downstream of the TSS. Click the **Apps** button then click **Screenshot**.
- [B] Click the **show track name** checkbox then click **Take screenshot**. Click either the **Link to the svg file** button or the **Link to the pdf file** button to save a high quality screenshot.



To catch up to this point of the workshop, click on the Apps menu and then select Session. Enter the session ID **PRdemo** in the Retrieve text box and click the status named **End**. For more details about sessions, refer to section 4.3 on page 13.

7. Training and support

- For more documentation and tutorials, visit:
<http://epigenomegateway.wustl.edu/support/index.html>
- Follow the WashU EpiGenome Browser on



<http://epigenomegateway.wustl.edu/+>



<http://epigenomegateway.wustl.edu/fb>



<https://twitter.com/wuepgg>



<https://github.com/epgg/eg>

8. Loading data on the Roadmap EpiGenome Browser

The **Roadmap EpiGenome Browser** is a companion browser to the WashU EpiGenome Browser to serve as a point-of-access to explore and analyze comprehensive epigenomics data generated by the Roadmap and ENCODE projects (>20k datasets). The Roadmap EpiGenome Browser performs real-time data clustering analyses on one or more data types to reveal cell and tissue type-specificity of epigenetic marks.

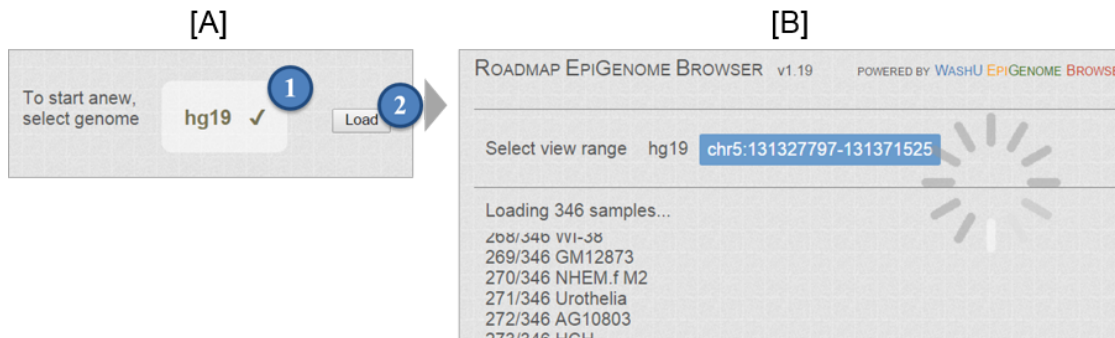
In this tutorial, we will demonstrate how to annotate non-coding genetic variants with epigenetic data using the Roadmap EpiGenome Browser.

8.1. Launching the browser

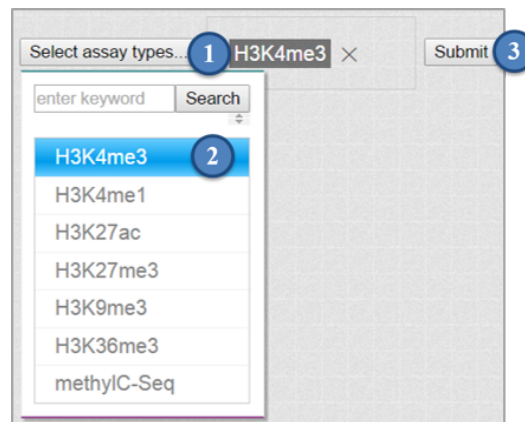
- Access the browser at <http://epigenomegateway.wustl.edu/browser/roadmap/>.

8.2. Loading data

- [A] Click the **hg19** button then click **Load**.
- [B] The browser will load all of the Roadmap Epigenomics and ENCODE reference epigenomes.



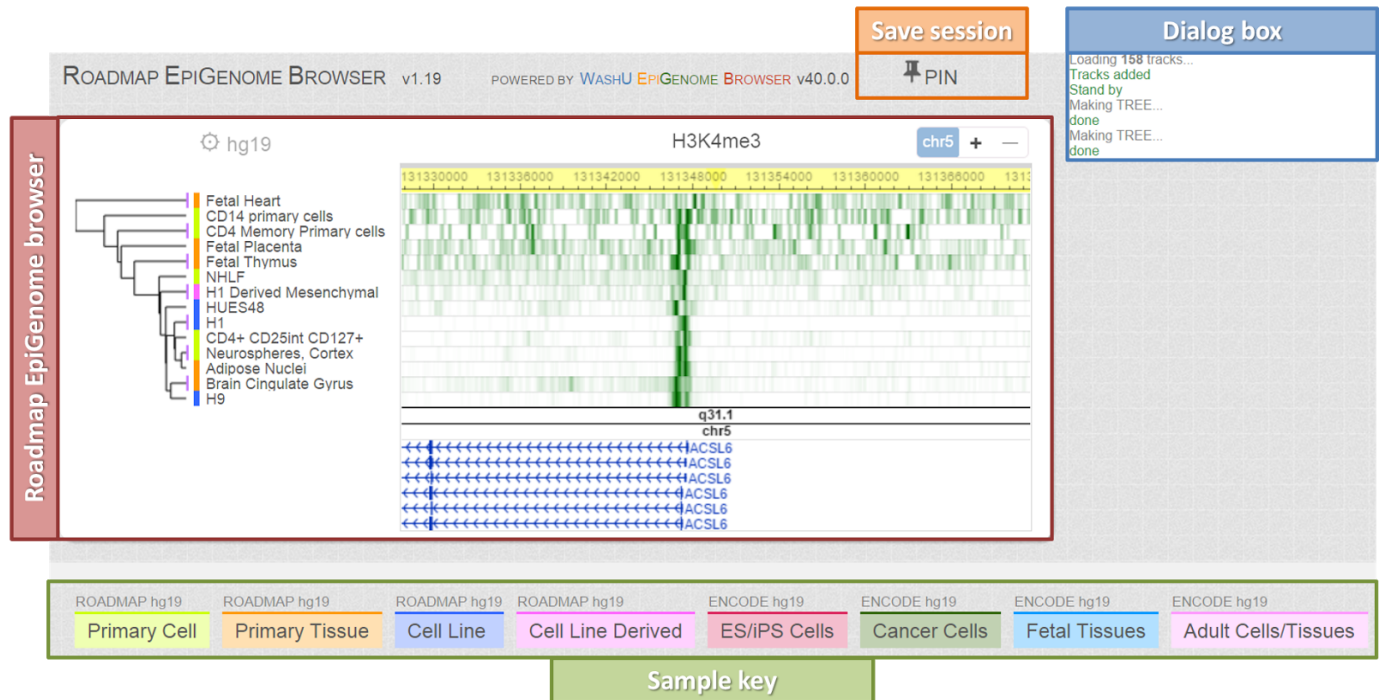
- To display tracks for a specific assay, click the **Select assay types...** button. Select **H3K4me3**. H3K4me3 is associated with actively transcribed promoters. Click **Submit**. The browser will display the H3K4me3 tracks for all samples.



To catch up to this point in the workshop, go to <http://epigenomegateway.wustl.edu/browser/roadmap/?pin=http://epigenomegateway.wustl.edu/browser/roadmap/t/903722733.json>.

9. Navigating the Roadmap EpiGenome Browser

9.1. Layout of the Roadmap EpiGenome Browser



9.1.1. Dialog box

- This box logs all actions performed by the Roadmap EpiGenome Browser. Any errors the browser encounters are displayed.

9.1.2. Save session

- Users can save their current browsing status for future reference or sharing.

9.1.3. Roadmap EpiGenome Browser

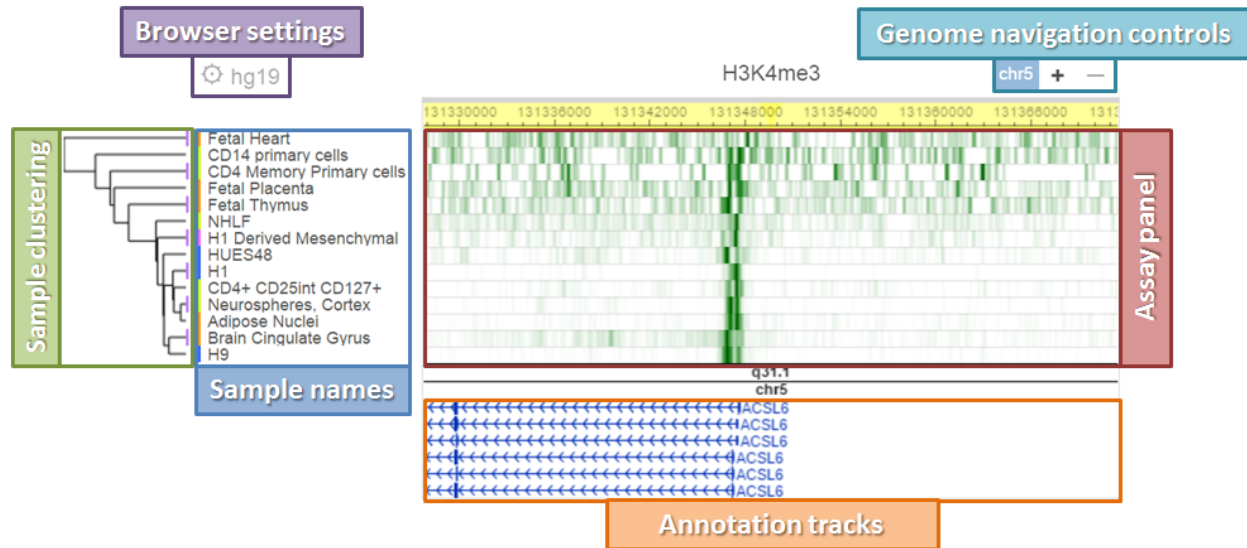
- This region contains the data tracks aligned to the genome along with reference genes and a chromosome ideogram. Additional public track hubs, annotation tracks, and custom tracks can be added by the user.

9.1.4. Sample key

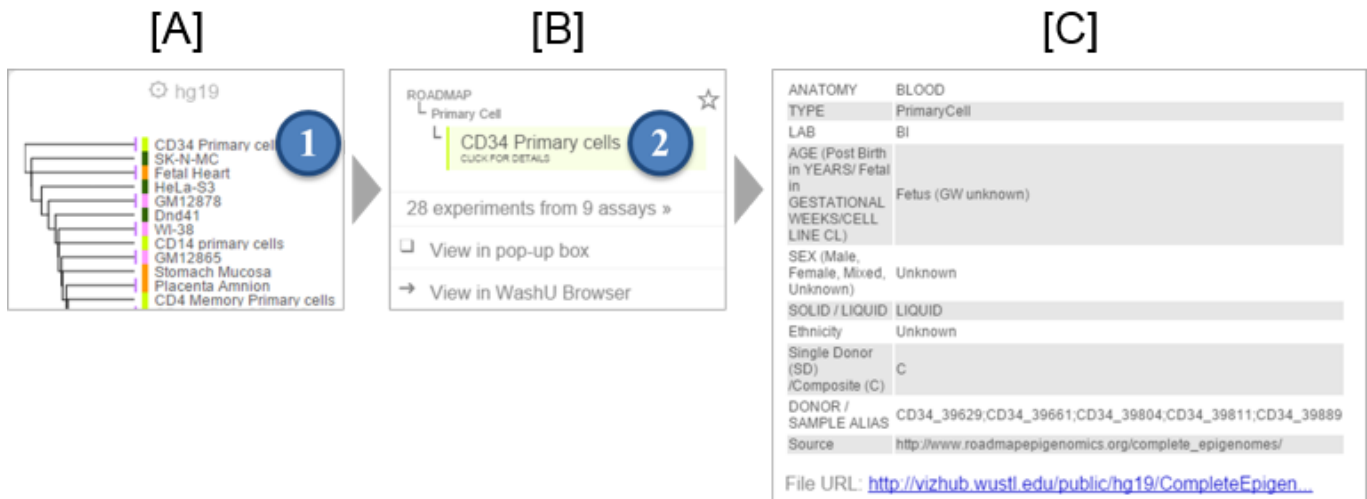
- This region contains the key for mapping the sample type and origin of each sample.

9.2. Navigating the genome on the Roadmap EpiGenome Browser

9.2.1. Data visualization



- **Browser settings** contain options to change which genomic regions, assay types, and samples are displayed as well as the **view range**.
- **Assay panel** contains the numerical tracks for each sample. All tracks are of the same assay type, for example, H3K4me3 ChIP-seq.
- **Sample names** for each track are listed on the left-hand side of the track. The colored bars indicate the sample type and origin of each sample.
- [A] To view the metadata for a sample, click on a **sample name**.
- [B] In the popup menu, click on the **sample name**.
- [C] The sample's metadata is displayed.
- Click outside the configuration menu to close the menu.




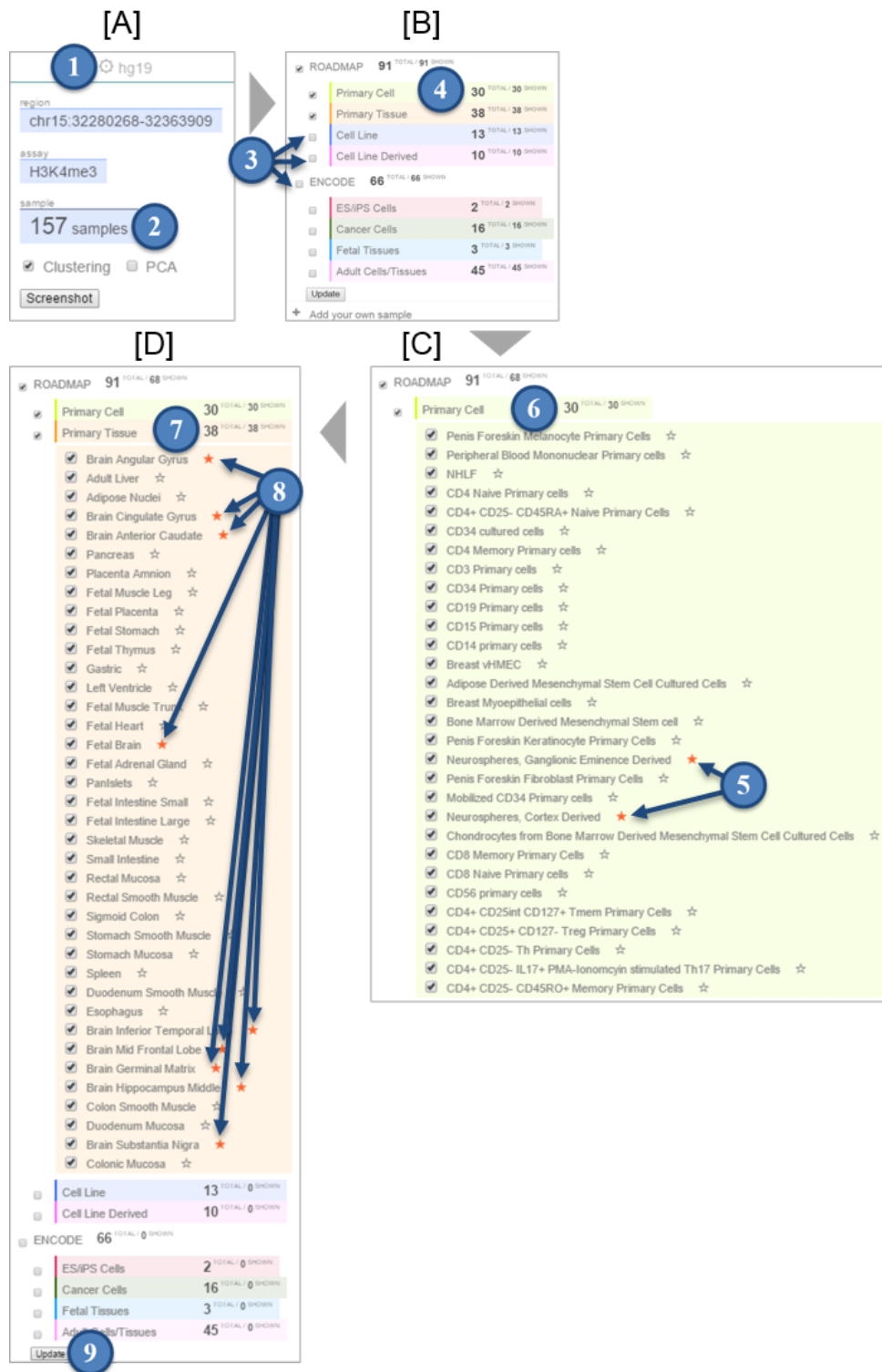
- The dendrogram shows **hierarchical clustering** of the samples based on the track data in the current view range. Whenever the assay panels are

updated, for example by scrolling, the hierarchical clustering analysis is automatically rerun to reflect the new sample relationships within the updated view range.

- Below the numerical tracks are **annotation tracks**, including annotations of genes.
- The **genome navigation controls** enable exploration across the genome.

9.2.2. Adding and removing samples using the browser settings menu

- For this workshop, we are interested in exploring the epigenomes of primary cells and tissue samples from Roadmap. In particular, we want to compare how brain samples relate to the other samples in this dataset.
- [A] By default, the browser displays all available samples. To change which samples are displayed, click the  button in the top-left corner and click [sample](#).
- [B] The number of available samples in each category is listed. [Uncheck](#) the Cell Line and Cell Line Derived samples under the ROADMAP header and uncheck all of the ENCODE samples. To list primary cell samples, click on [Primary Cell](#).
- [C] “Favorite” the brain samples by clicking on the star icon next to each brain sample (2 total: “Neurospheres, Ganglionic Eminence Derived” and “Neurospheres, Cortex Derived”). This will add a star icon to the selected names in the main browser as we will see shortly. Click on [Primary Cell](#) to collapse the sample list.
- [C] To list primary tissue samples, click on [Primary Tissue](#). “Favorite” the brain samples by clicking on the star icon next to each brain sample (9 total: “Brain Angular Gyrus,” “Brain Cingulate Gyrus,” “Brain Anterior Caudate,” “Brain Fetal Brain,” “Brain Inferior Temporal Lobe,” “Brain Mid Frontal Lobe,” “Brain Germinal Matrix,” “Brain Hippocampus Middle,” and “Brain Substantia Nigra”).
- To apply the changes, click the [Update](#) button.



9.2.3. Genome navigation controls

- To zoom in or out, click on the **+** or **-** buttons. Alternatively, **select a region in the coordinate ruler** to zoom in on a region of interest.
- To pan left or right, click on any track and drag left or right, respectively.
- [A] Click the blue **navigation box** to move to a region of interest.

- [B] The region can be specified by a genomic coordinate, a gene name, or a SNP rsID. In the gene name textbox, type **CHRNA7** and click **Go**. Multiple gene models are shown for the gene. Select the **first gene model** to move to its location.

[A]

1 chr5 + -

[B]

The screenshot shows a genomic browser interface. At the top, a chromosome scale from 0 to 180M is displayed. Below it, a search bar contains 'CHRNA7' and a 'Go' button. A second search bar contains 'rs75669958' and a 'Find SNP' button. Below the search bars, a list of gene models is shown, including 'refGene' and 'gencodeV17' tracks. The first gene model is highlighted in yellow. A blue circle with the number 1 is placed over the 'chr5' button in the top left. A blue circle with the number 2 is placed over the 'CHRNA7' text box. A blue circle with the number 3 is placed over the 'Go' button. A blue circle with the number 4 is placed over the first gene model in the list.

- To move the view range to the promoter of **CHRNA7**, click on the Assay Panel and drag the mouse. To zoom in on the promoter, click upstream of the transcription start site in the annotation track and drag the mouse.
- [A] To better resolve the variation in H3K4me3 signal across samples, right-click on any **H3K4me3 track** and select **Configure**.
- [B] Check the **Apply to all tracks checkbox**. Change the Y-axis scale drop-down menu to **Fixed**, set the max to **40**, and click **apply**.

[A]

1

2

Configure

Information

Remove

[B]

positive beyond threshold

negative beyond threshold

Y-axis scale Fixed

min: 0 max: 40 apply

Summary method Average

Logarithm No

Height + - + - Unify

Smooth window

Track background

Curve only

3

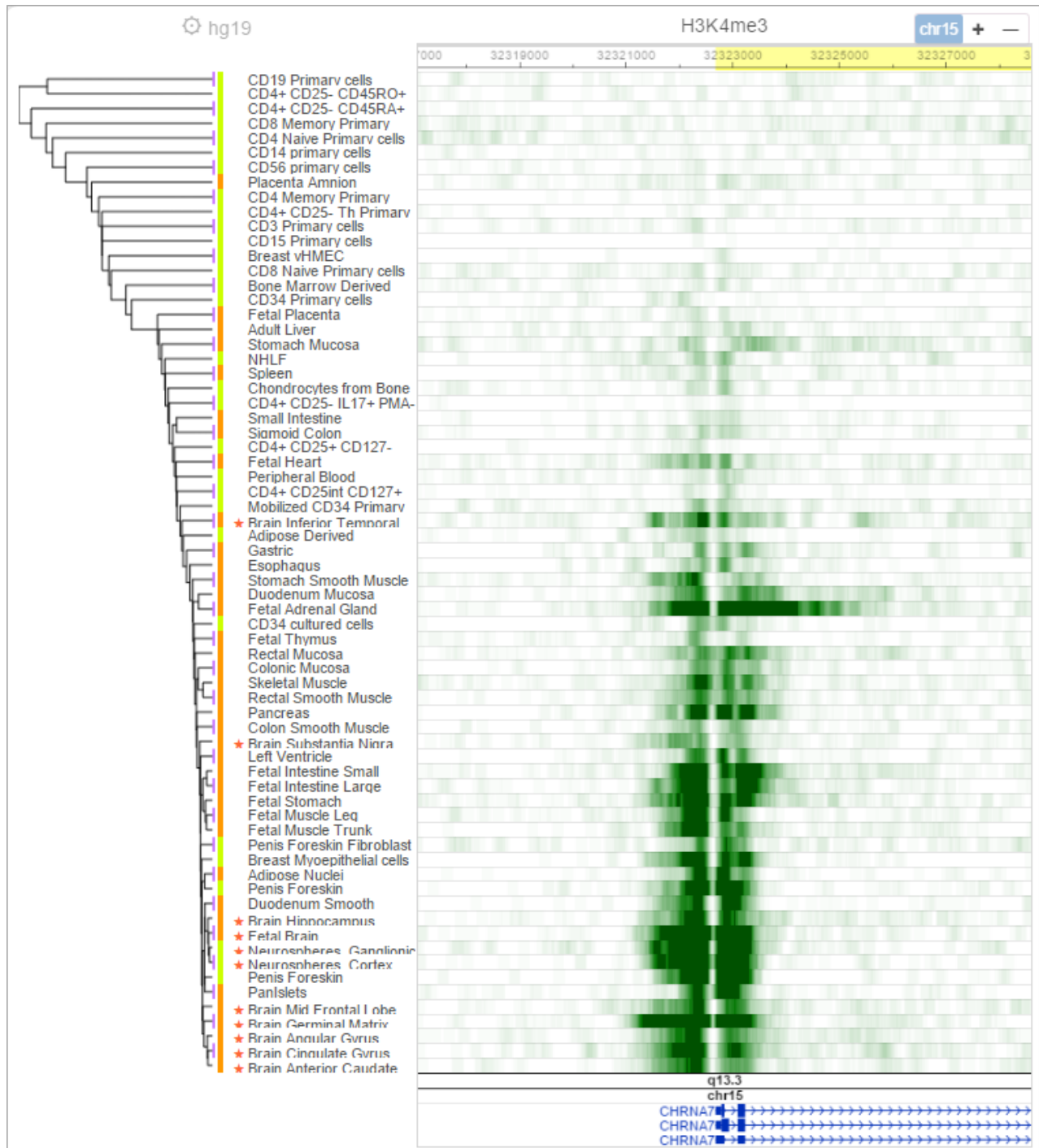
4

5

6

The screenshot shows the configuration panel for an H3K4me3 track. On the left, a track is shown with a right-click context menu open, displaying options: 'Configure' (selected), 'Information', and 'Remove'. A blue circle with the number 1 is placed over the track, and a blue circle with the number 2 is placed over the 'Configure' button. On the right, the configuration panel is shown. It has a 'positive' button (green) and a 'beyond threshold' button (green). Below these are 'negative' (red) and 'beyond threshold' (red) buttons. The 'Y-axis scale' is set to 'Fixed'. The 'min' value is 0 and the 'max' value is 40. There is an 'apply' button. The 'Summary method' is set to 'Average'. The 'Logarithm' is set to 'No'. The 'Height' section has a '+' button, a '-' button, and a 'Unify' button. There are checkboxes for 'Smooth window', 'Track background', and 'Curve only'. At the bottom, there is a checkbox labeled 'Apply to all tracks' which is checked. A blue circle with the number 3 is placed over this checkbox. A blue circle with the number 4 is placed over the 'Fixed' dropdown menu. A blue circle with the number 5 is placed over the 'max' input field. A blue circle with the number 6 is placed over the 'apply' button.

- Many of the brain samples have a strong H3K4me3 signal over the *CHRNA7* promoter while the blood samples do not.



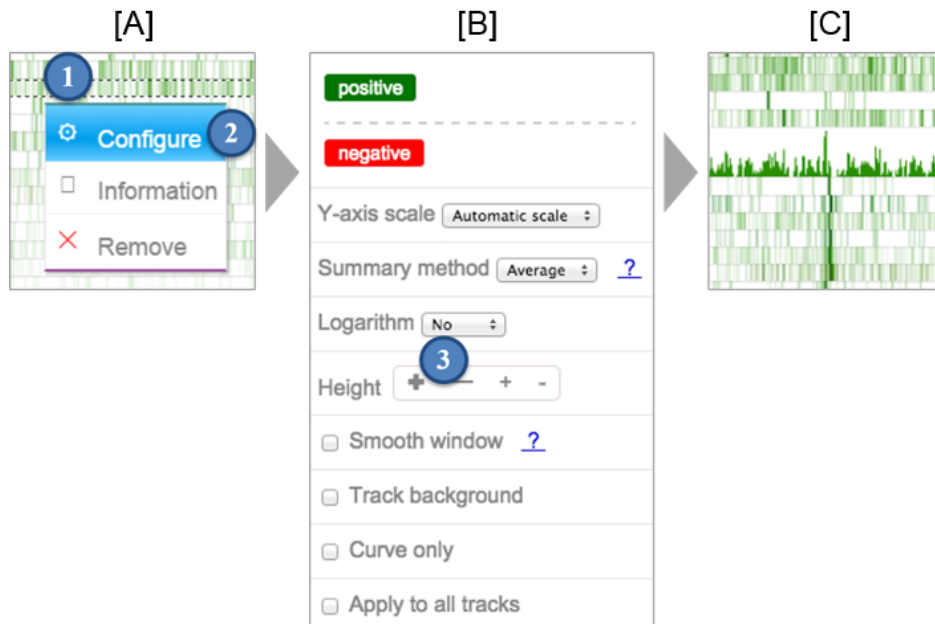
To catch up to this point in the workshop, go to

<http://epigenomegateway.wustl.edu/browser/roadmap/?pin=http://epigenomegateway.wustl.edu/browser/roadmap/t/937613542.json>.

10. Configuring the Roadmap EpiGenome Browser

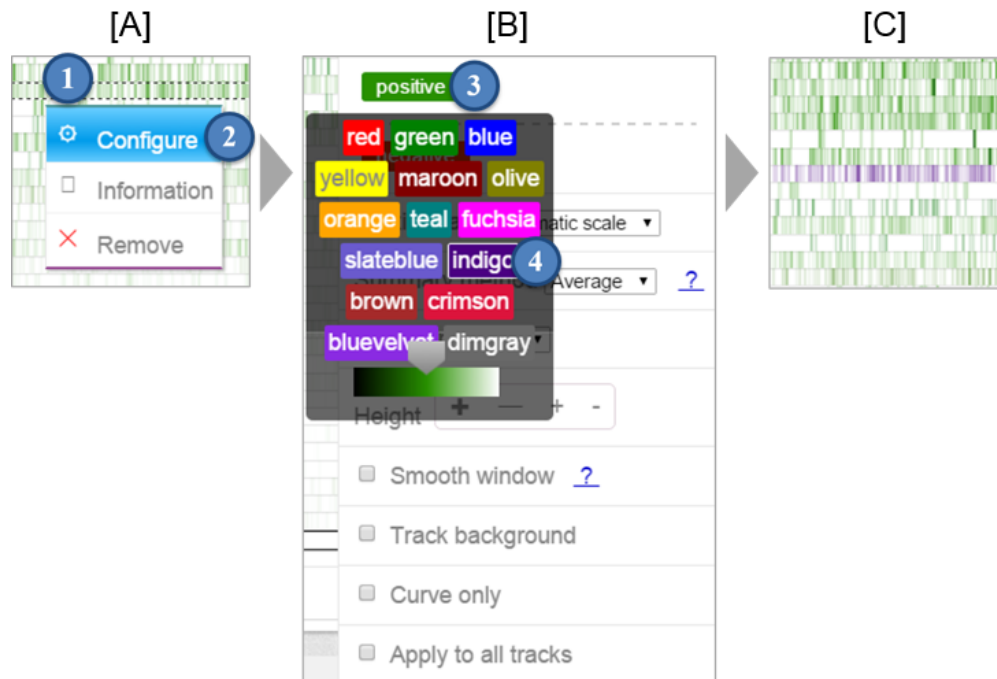
10.1. Changing the height of tracks

- [A] Right-click on a track and select **Configure**.
- [B] Click the large **+** sign repeatedly to increase the height of the track to your preference.
- [C] The track is now displayed as a bar plot instead of a heatmap.
- Click outside the configuration menu to close the menu.



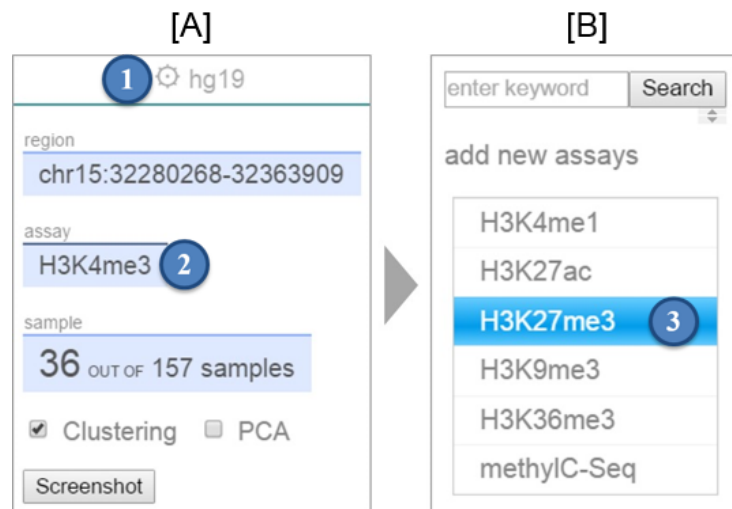
10.2. Changing the color of tracks

- [A] Right-click a track and select **Configure**.
- [B] In the configuration menu, click the **positive** button and then select the **color** of your choice to change the color of the track.
- [C] The track is now colored purple instead of green.
- Click outside the configuration menu to get back to the browser.

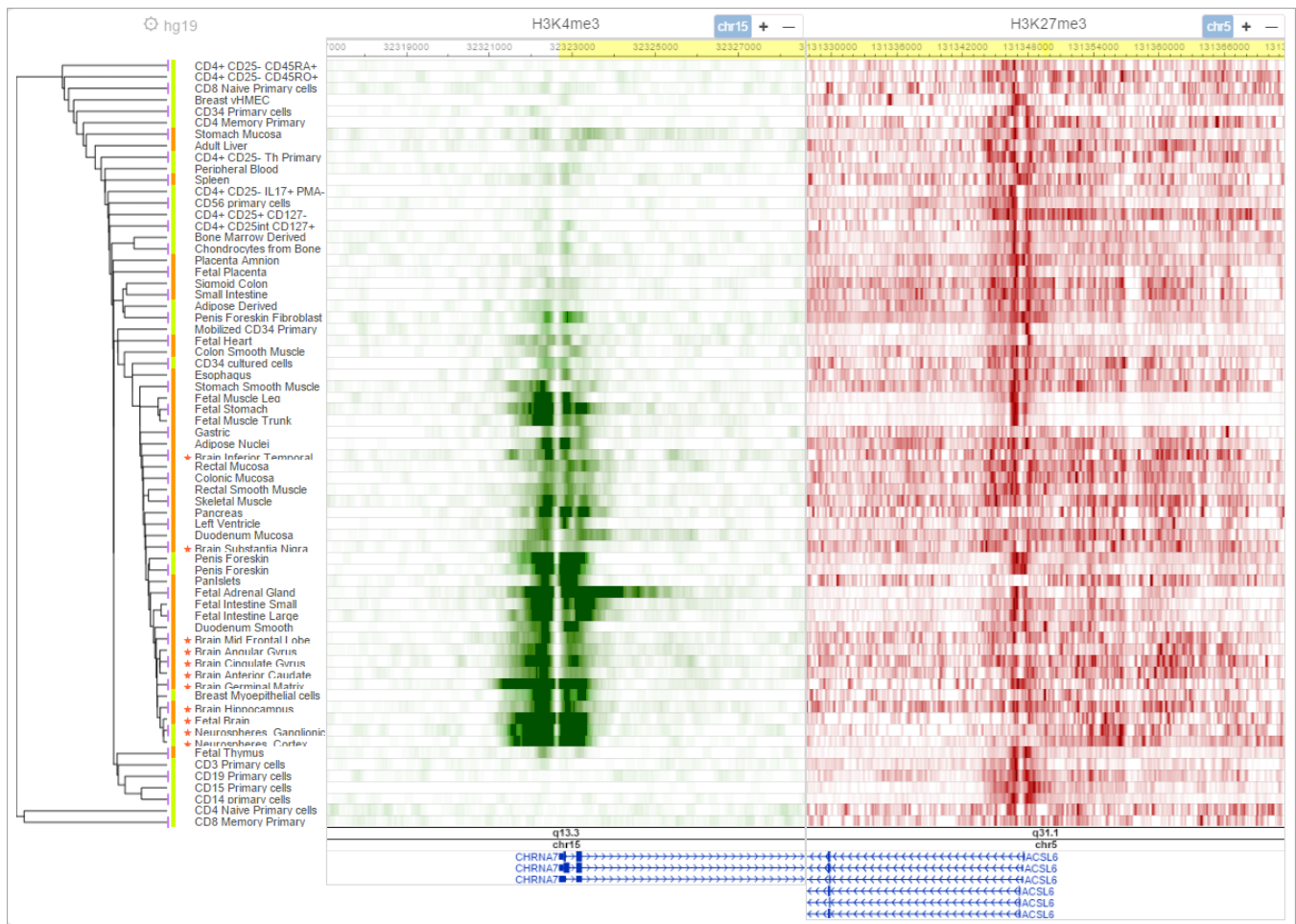


10.3. Adding assays on the Roadmap EpiGenome Browser

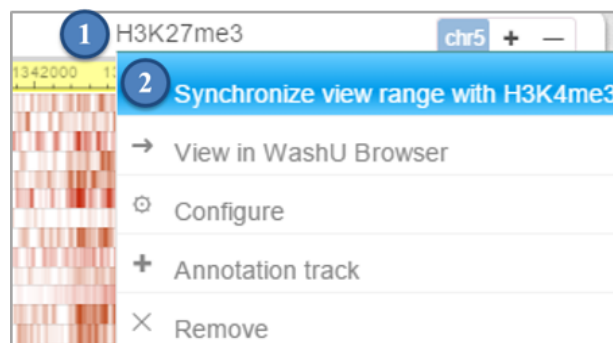
- [A] To add H3K27me3 tracks, click the button near the top-left corner of the browser and then click **assay**.
- [B] Select **H3K27me3**. H3K27me3 marks polycomb repressive chromatin.



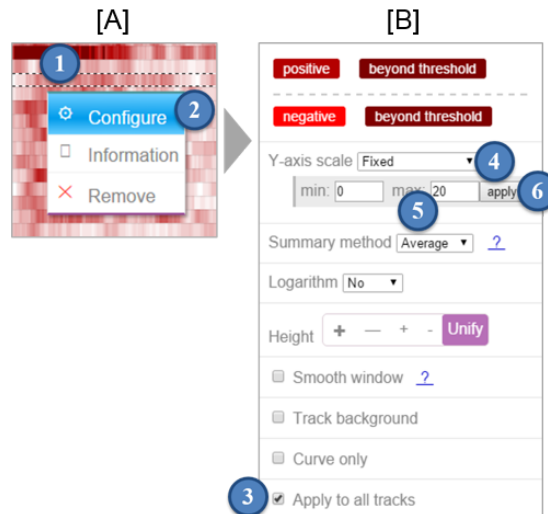
- The H3K27me3 data is shown in a new panel. Each row now contains H3K4me3 and H3K27me3 for a particular sample. The samples are clustered using the data in *both* panels. By default, the browser loads all samples with both H3K4me3 and H3K27me3 data. Repeat the steps in section 9.2.2 to filter for the Roadmap primary cell and tissue samples.



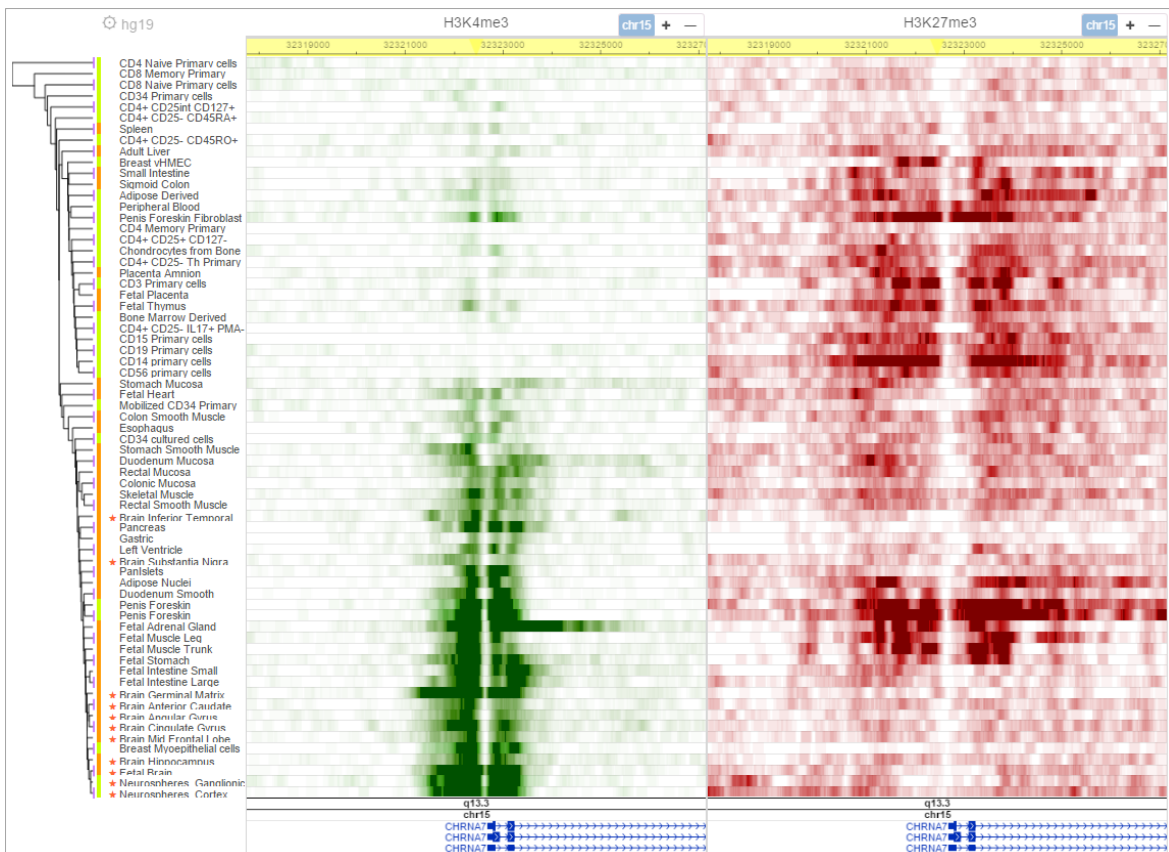
- The new assay panel is over the default genomic region. By default, the view range of each assay panel is independent of the others. To synchronize the view range of the panels, click the **banner of the H3K27me3** assay panel and select **Synchronize view range with H3K4me3**.



- [A] To better resolve the variation in H3K27me3 signal across samples, right-click on any **H3K27me3 track** and select **Configure**.
- [B] Check the **Apply to all tracks checkbox**. Change the Y-axis scale drop-down menu to **Fixed**, set the max to **20**, and click **apply**.



- In the brain samples, the H3K4me3 signal is high and H3K27me3 is low over the transcription start site while in the blood samples, H3K4me3 is low and H3K27me3 is high. Several of the fetal and penis foreskin samples show patterns of a poised/bivalent promoter with both active and repressive chromatin marks present. The browser makes it easy to see the covariation of epigenetic marks across samples for a region.

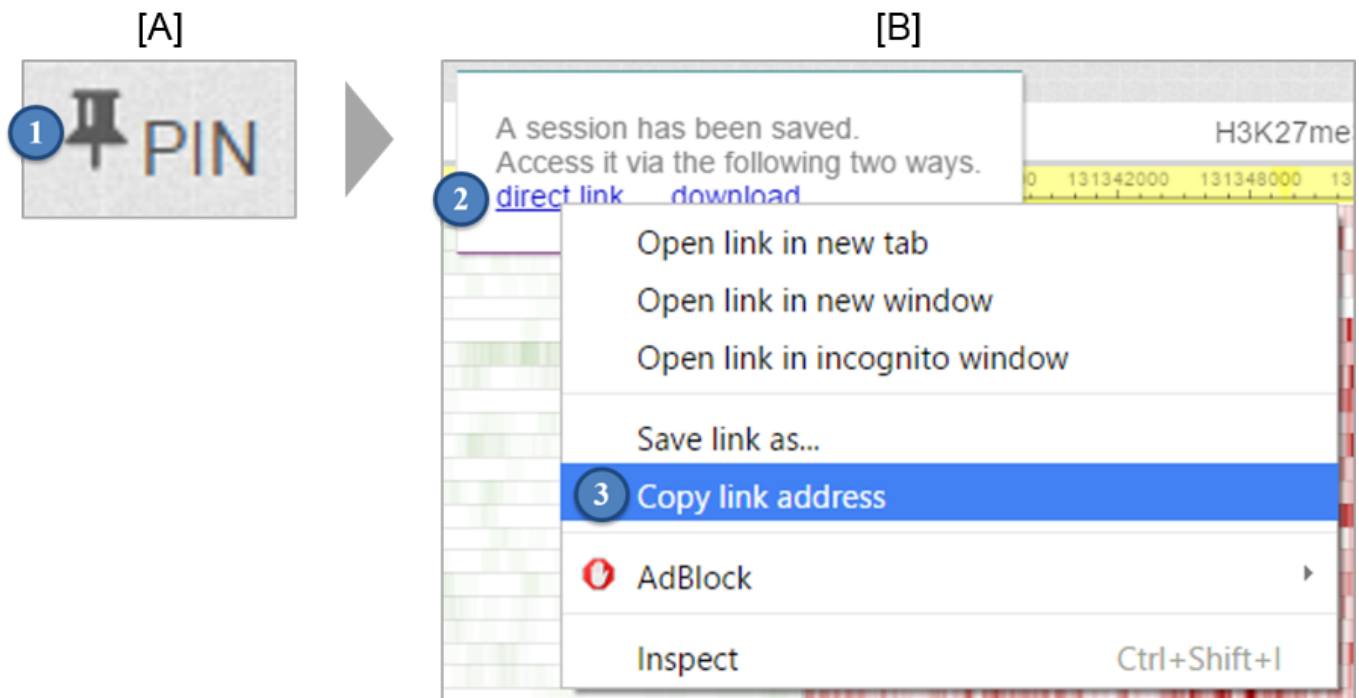


To catch up to this point in the workshop, go to

<http://epigenomegateway.wustl.edu/browser/roadmap/?pin=http://epigenomegateway.wustl.edu/browser/roadmap/t/1138692289.json>.

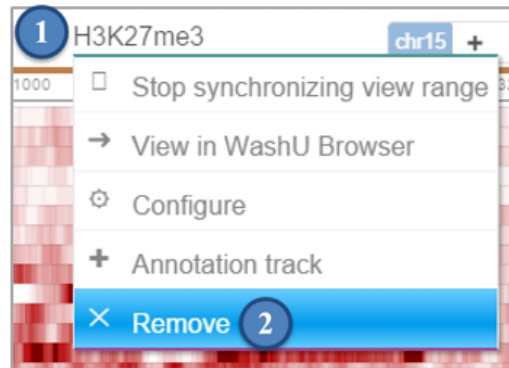
11. Saving, sharing, and retrieving Roadmap EpiGenome Browser sessions


- [A] To save a session, click the **pin icon** at the top of the browser.
- [B] Right-click **direct link** and select **Copy link address** to copy the URL to your clipboard. Paste the URL somewhere where you'll be able to find it again.
- To retrieve a session, paste the session URL in a web browser.

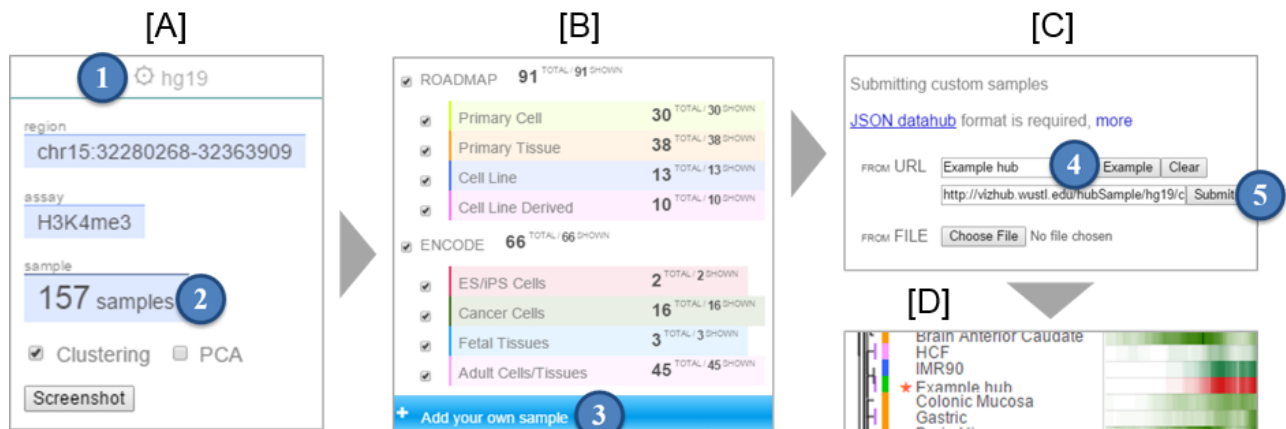


12. Creating Custom Tracks on the Roadmap EpiGenome Browser

- Users can submit their own datasets to analyze them in the context of public samples from the Browser. For this tutorial, our custom sample is IMR90 H3K4me3 ChIP-seq data. First, remove the H3K27me3 assay panel by clicking the **H3K27me3 banner** and select **Remove**.



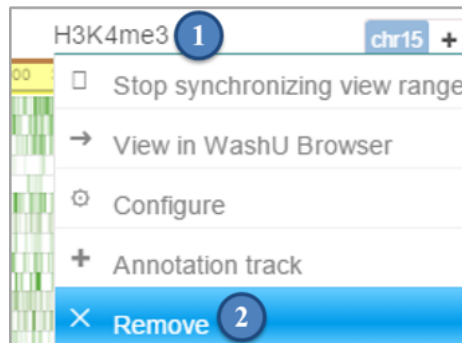
- [A] To add the custom sample, click the  button and click **sample**.
- [B] Select **Add your own sample**.
- [C] In the pop-up window, the user can enter the **URL** of the datahub file or upload the datahub file and enter a sample name. In this tutorial, we'll use the example custom sample. Click **Example** and click **Submit**.
- [D] The custom sample is shown as a **red** heatmap and is marked by a star.




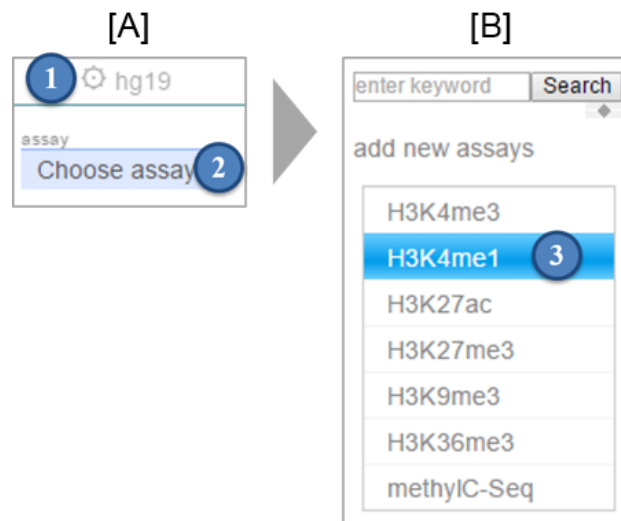
13. Annotating non-coding variants with epigenomic data

13.1. Viewing a non-coding SNP in the context of H3K4me1 data

- A genome-wide association study by Sawcer et al. found SNP rs756699 to be associated with multiple sclerosis (MS). In this section, we will investigate if this SNP is associated with any regulatory elements. First, remove the H3K4me3 assay panel by clicking the **H3K4me3 banner** and selecting **Remove**.

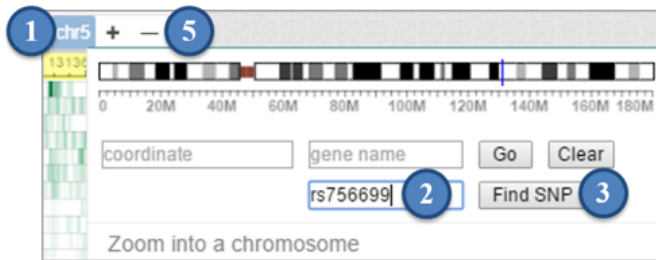


- To add H3K4me1 tracks, click the  button and then click **Choose assay**. Select **H3K4me1**. H3K4me1 marks active enhancers.



- Repeat the steps in section 9.2.2 to filter for the Roadmap primary cell and tissue samples.
- [A] To go to the SNP location, click the blue **navigation box**. Enter **rs756699** in the SNP textbox. Then click **Find SNP**.
- [B] Select **rs756699 chr5:133446574-133446575**. The SNP is marked by a yellow rectangle. Zoom out by clicking the **-** button multiple times.

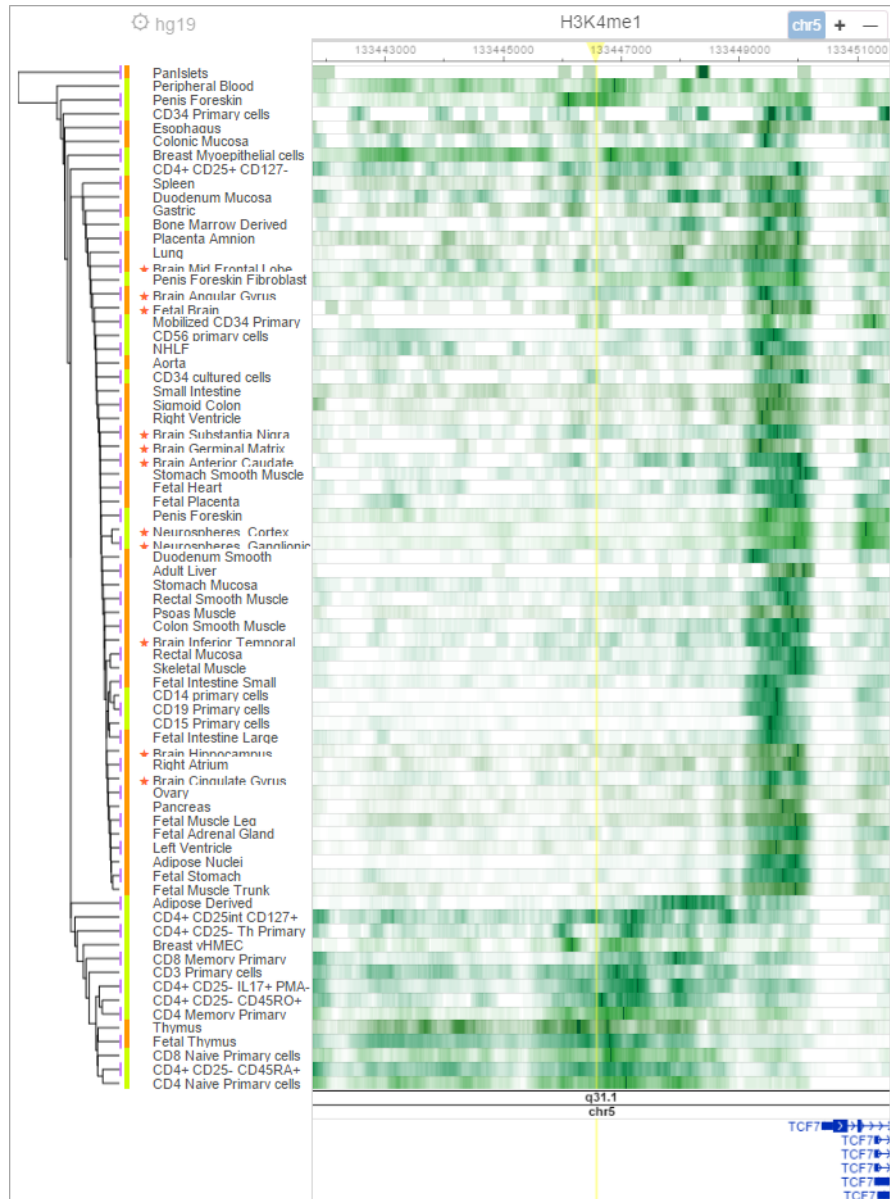
[A]



[B]

rs756699 chr5:133446574-133446574 4


- The SNP lies in a putative blood-specific enhancer upstream of the *TCF7* transcription start site.

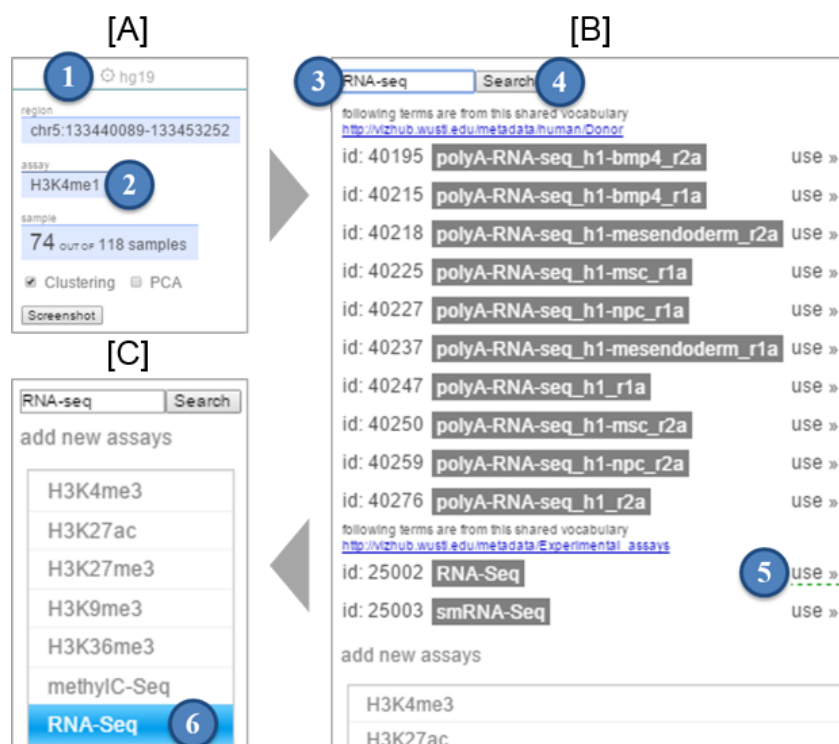


To catch up to this point in the workshop, go to

<http://epigenomegateway.wustl.edu/browser/roadmap/?pin=http://epigenomegateway.wustl.edu/browser/roadmap/t/893995130.json>.

13.2. Viewing a non-coding SNP in the context of H3K4me1 and expression data

- [A] Next, we want to determine if the putative blood-specific enhancer regulates *TCF7* expression in a blood-specific fashion. To add the RNA-seq assay, click the  button and then click **assay**.
- [B] In the search textbox, type **RNA-seq** and click **Search**. In the search results, click the **use »** link next to the RNA-seq assay. The RNA-seq list has now been added to the list of assays.
- [C] Select **RNA-seq**.

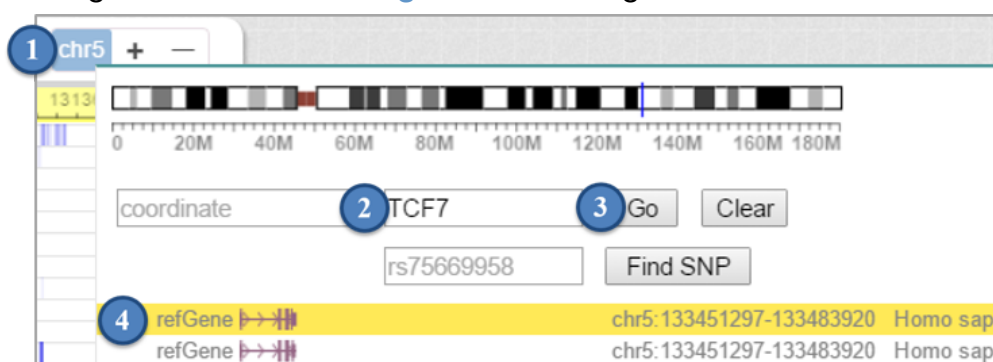


[A] Screenshot of the 'assay' dropdown menu. The 'region' is set to 'chr5:133440089-133453252'. The 'assay' is 'H3K4me1'. The 'sample' is '74 out of 118 samples'. The 'Clustering' checkbox is checked. The 'PCA' checkbox is unchecked. The 'Screenshot' button is visible.

[B] Screenshot of the search results for 'RNA-seq'. The search box contains 'RNA-seq' and the 'Search' button is clicked. The results list includes 'polyA-RNA-seq_h1-bmp4_r2a', 'polyA-RNA-seq_h1-bmp4_r1a', 'polyA-RNA-seq_h1-mesendoderm_r2a', 'polyA-RNA-seq_h1-mesendoderm_r1a', 'polyA-RNA-seq_h1_r1a', 'polyA-RNA-seq_h1-misc_r2a', 'polyA-RNA-seq_h1-npc_r2a', 'polyA-RNA-seq_h1_r2a', 'RNA-Seq', and 'smRNA-Seq'. The 'use »' link is visible next to 'RNA-Seq'.

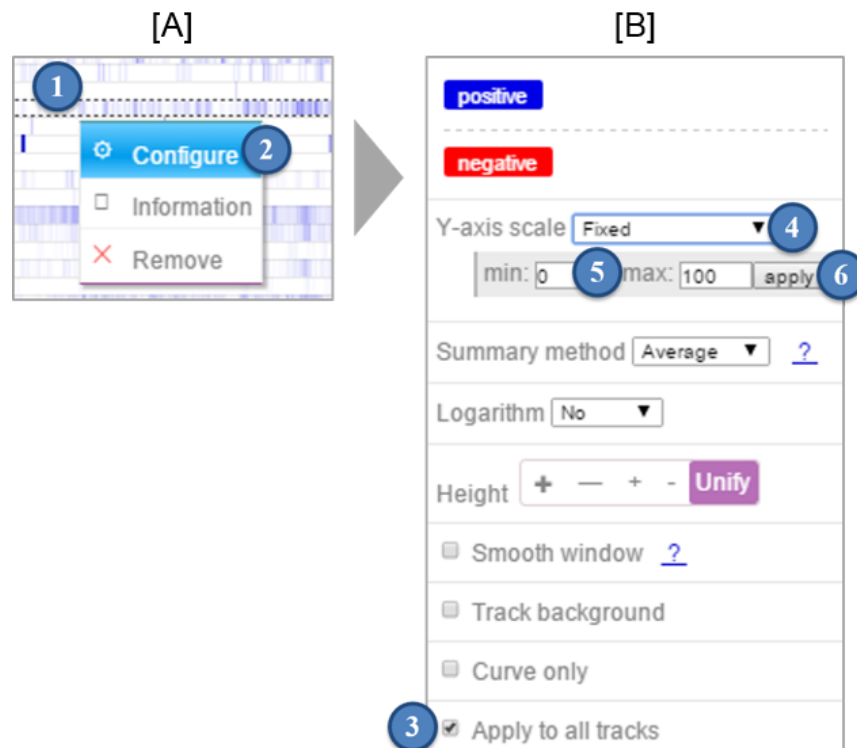
[C] Screenshot of the 'add new assays' list. The search box contains 'RNA-seq' and the 'Search' button is clicked. The list includes 'H3K4me3', 'H3K27ac', 'H3K27me3', 'H3K9me3', 'H3K36me3', 'methylC-Seq', and 'RNA-Seq'. The 'RNA-Seq' option is selected.

- Repeat the steps in section 9.2.2 to filter for the Roadmap primary cell and tissue samples.
- In the RNA-seq panel (not the H3K4me3 panel), click the blue **navigation box**. In the gene name textbox, type **TCF7** and click **Go**. Multiple gene models are shown for the gene. Select the **first gene model** to go to its location.



Screenshot of the gene model view. The 'coordinate' field contains 'TCF7' and the 'Go' button is clicked. The 'Find SNP' button is also visible. The gene model shows the 'refGene' track with the 'TCF7' gene model highlighted. The 'refGene' track shows the gene structure with exons and introns. The 'TCF7' gene model is highlighted in yellow.

- [A] To set all of the RNA-seq tracks to the same scale, right-click on any **RNA-seq track** and select **Configure**.
- [B] Check the **Apply to all tracks checkbox**. Change the Y-axis scale drop-down menu to **Fixed**, set the max to **100**, and click **apply**.



- Click on the **RNA-seq panel** and drag the mouse to center the view range on the transcription start site of *TCF7*.
- The blood samples with high H3K4me1 signal around the SNP also have high *TCF7* expression. This pattern is not observed in the brain samples. This suggests a potential regulatory relationship between the putative enhancer and *TCF7* and thus the SNP may contribute to the disruption of *TCF7* gene expression.



To catch up to this point in the workshop, go to

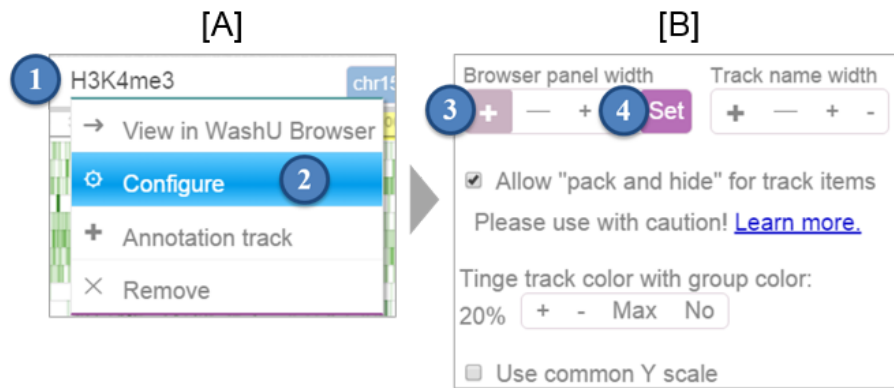
<http://epigenomegateway.wustl.edu/browser/roadmap/?pin=http://epigenomegateway.wustl.edu/browser/roadmap/t/1703831292.json>.

13.3. Viewing multiple non-coding SNPs simultaneously


- The **Gene Set View** allows users to look at data from multiple locations across the genome within the same browser panel. First, remove the RNA-seq assay panel by clicking the **RNA-seq banner** and select **Remove**. Repeat the steps in section 9.2.2 to filter for the Roadmap primary cell and tissue samples.

13.3.1. Formatting the view range

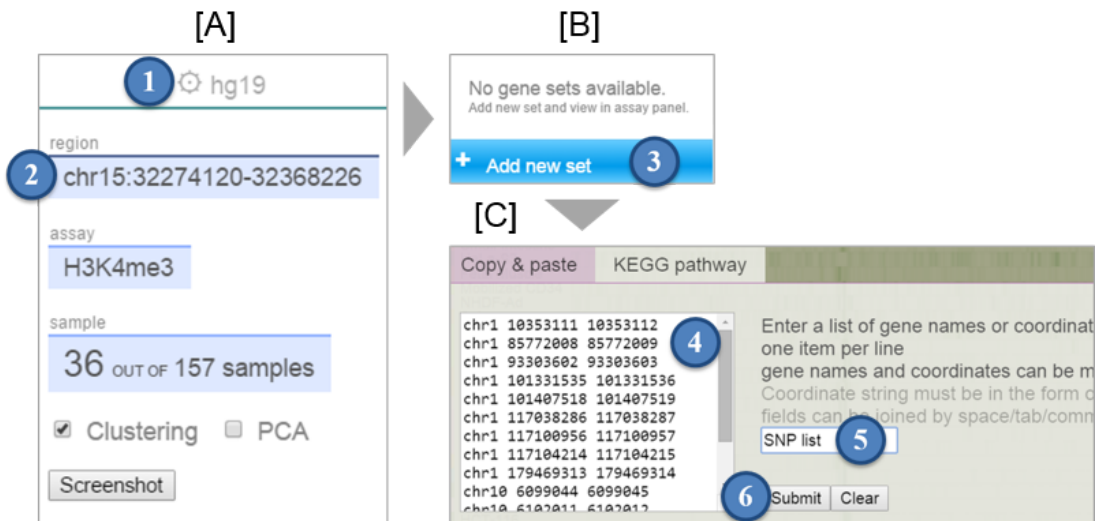
- [A] To view multiple locations in the genome in one browser panel more easily, increase the browser width. To adjust the **browser panel width**, click the **banner** containing the assay name near the top of the browser. Select **Configure**.
- [B] The + and - buttons increase or decrease the browser width, respectively. (The smaller + and - buttons increase or decrease the browser width by a smaller amount.) Click + then **Set** to change the browser width.
- Similarly, the **width of the track names** can be changed using the + and - buttons.



13.3.2. Creating a gene/region set

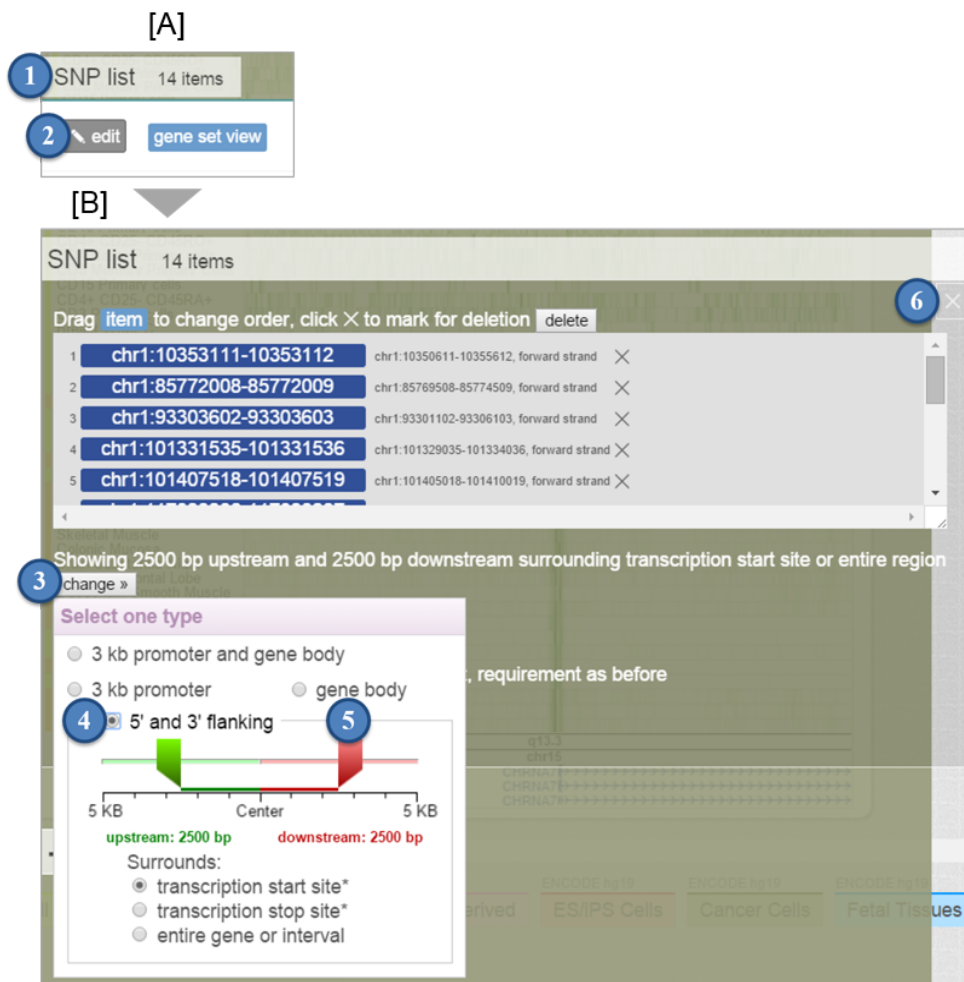
- [A] To visualize a set of SNPs associated with multiple sclerosis, click the  button and then click **region**.
- [B] Click **Add new set**.
- [C] Paste the list of SNP coordinates (listed below) in the textbox, enter a **name**, and click **Submit**.

```
chr1 10353111 10353112
chr1 85772008 85772009
chr1 93303602 93303603
chr1 101331535 101331536
chr1 101407518 101407519
chr1 117038286 117038287
chr1 117100956 117100957
chr1 117104214 117104215
chr1 179469313 179469314
chr10 6099044 6099045
chr10 6102011 6102012
chr10 6110828 6110829
chr10 43814048 43814049
chr10 81036006 81036007
```



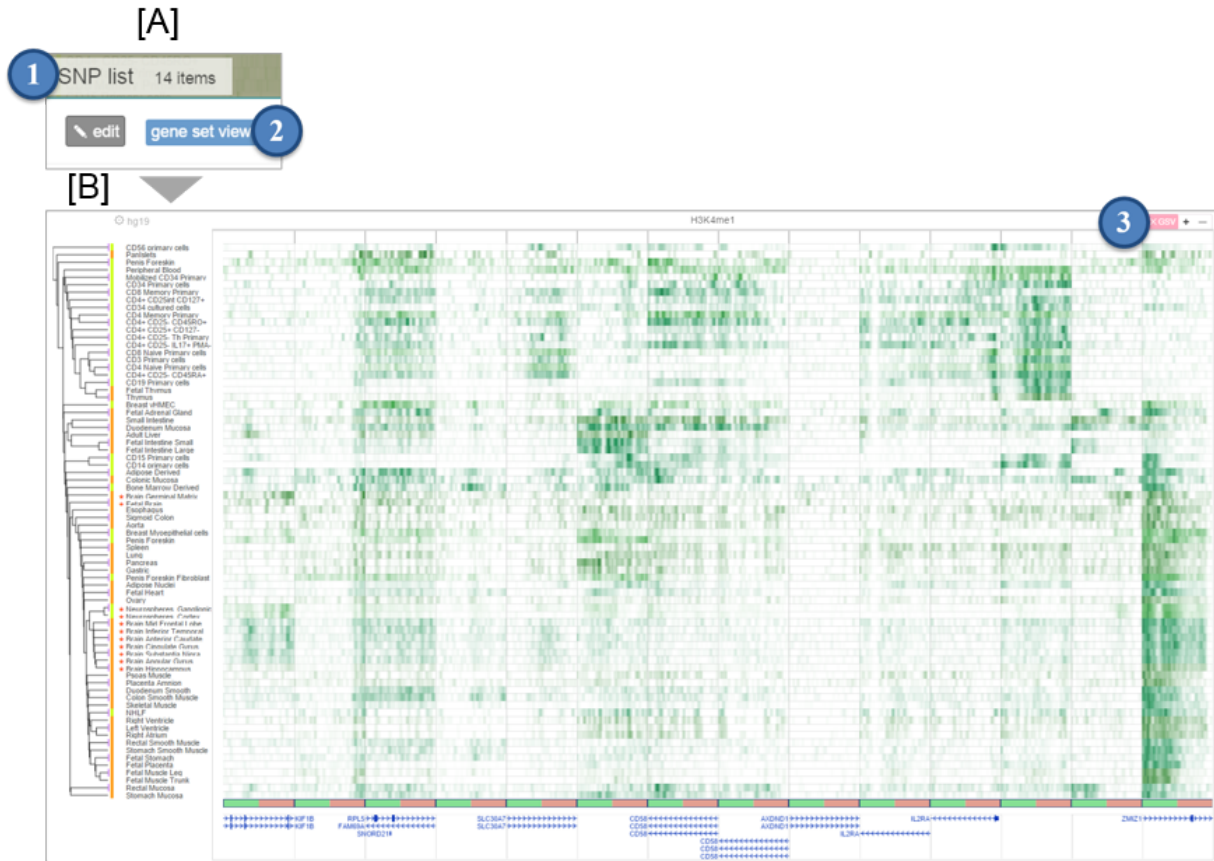
13.3.3. Editing a gene/region set

- [A] To focus on a 5kb window surrounding each of the SNPs, click on the **gene set banner** and then click **edit**.
- [B] Click **change »**, and then select **5' and 3' flanking**. Move the **red slider** to 2.5 KB. Click the **X** button *in the the edit window* to exit.



13.3.4. Viewing a gene/region set

- [A] Click on the **gene set banner** and then click **gene set view**.
- [B] The browser shows the H3K4me1 profile over a 5 KB window centered on each of the SNPs. Several of the SNPs have tissue-specific H3K4me1. To exit the gene set view, click the **X GSV** button in top-right corner.



To catch up to this point in the workshop, go to

<http://epigenomegateway.wustl.edu/browser/roadmap/?pin=http://epigenomegateway.wustl.edu/browser/roadmap/t/530450316.json>.

13.3.5. Viewing complete epigenomic profiles

- [A] To view all of the datasets for a particular sample in the WashU EpiGenome Browser, click on the sample name **PanIslets** and then select **View in WashU Browser**.
- [B] A WashU EpiGenome Browser window will be opened. All of the available datasets for pancreatic islets will loaded over the default view range.

