

# How to view CSREP output for Roadmap and Epimap data on the genome browser

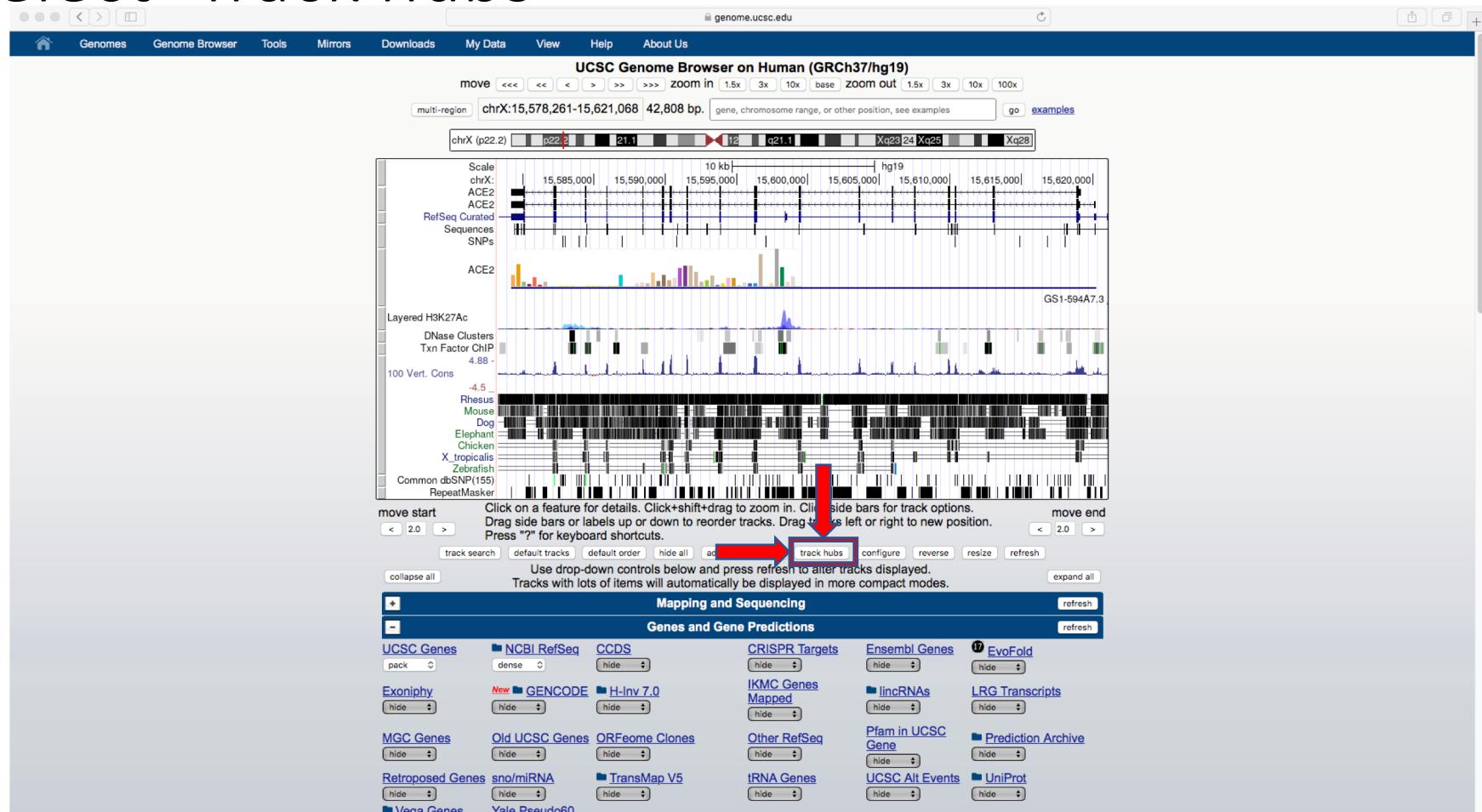
Ha Vu

09/21/2022

## 1. Visit the [UCSC genome browser website](#)

- Note, for us, using Google Chrome as the browser hid away multiple panels of the browser, so we used Safari

# Select 'Track hubs'



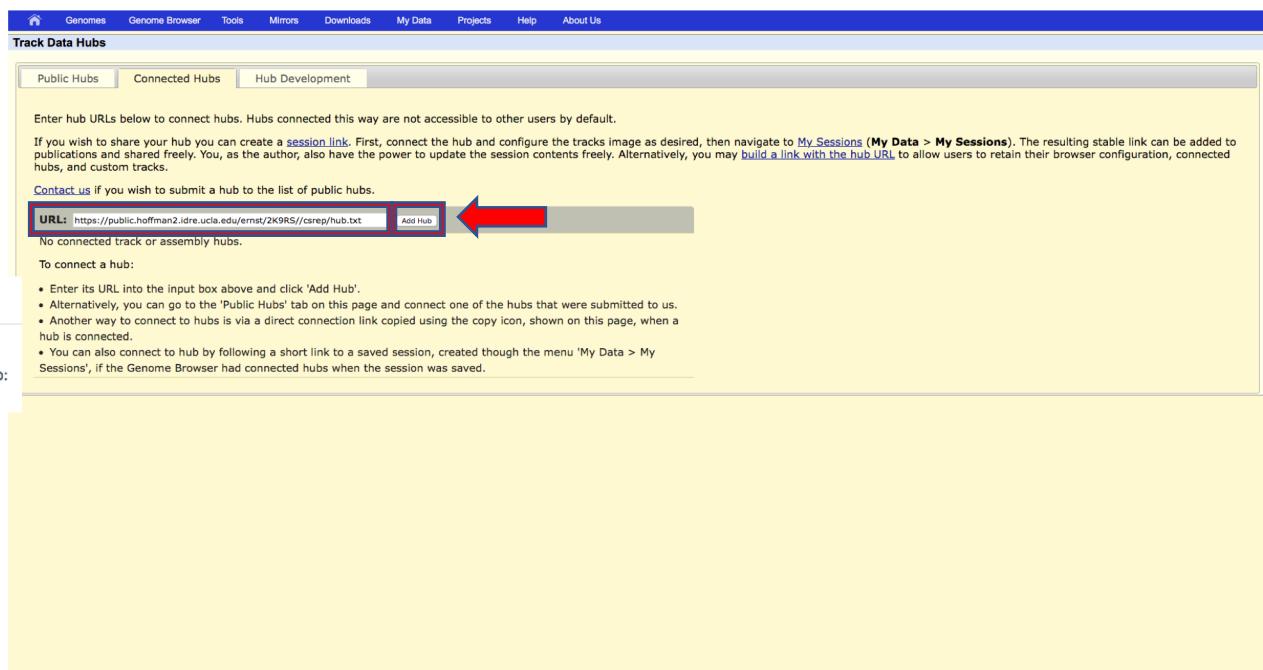
# On the ‘Connected Hubs’ tab, paste the link to CSREP track hub

- The link to CSREP track hub is most updated in our [Github](#), then ‘Add Hub’

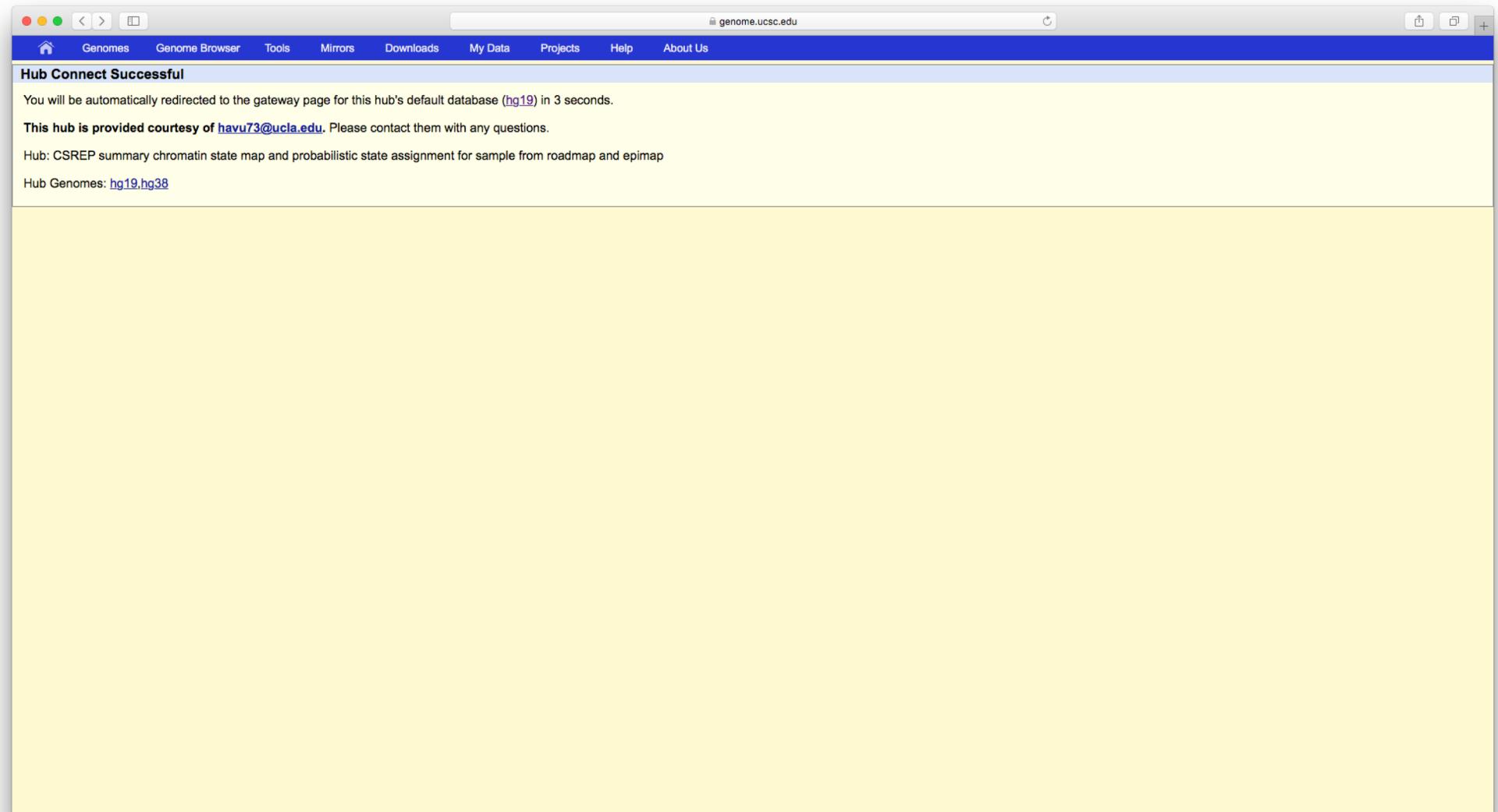
On Github, find the trackhub link:

## Viewing summary chromatin state maps on USCS Genome Browser

Users can easily view the provided summary chromatin state maps (state assignment probabilities and state annotations) for cell groups in Roadmap and Epimap on USCS Genome Browser, using the link to the track hub:  
<https://public.hoffman2.idre.ucla.edu/ernst/2K9RS//csrep/hub.txt>



# Wait a few seconds



# Select a genomic region and ‘Go’

The screenshot shows the UCSC Genome Browser Gateway homepage. At the top, there's a navigation bar with links for Genomes, Genome Browser, Tools, Mirrors, Downloads, My Data, Projects, Help, and About Us. Below the navigation bar is a search bar labeled "Find Position" with a dropdown menu showing "Human Assembly" and "Feb. 2009 (GRCh37/hg19)". There's also a "GO" button and a "Position/Search Term" input field containing "chrX:15,578,261-15,621,068".

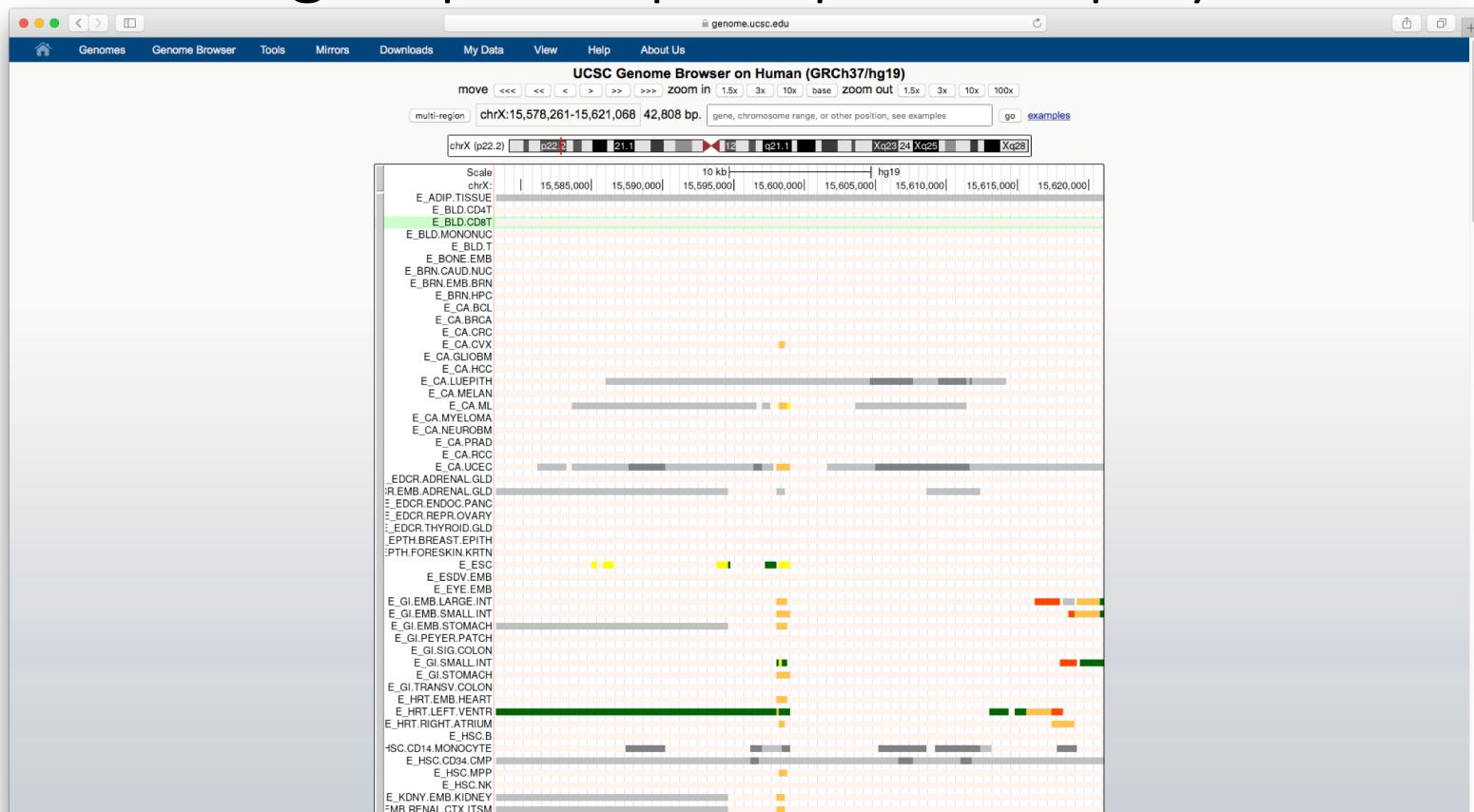
On the left, there's a "Browse/Select Species" section with a "POPULAR SPECIES" grid (Human, Mouse, Rat, Zebrafish, Fruitfly, Worm, Yeast) and a search bar for "species, common name or assembly ID". Below this is a link to "Unable to find a genome? Send us a request." and a "UCSC SPECIES TREE AND CONNECTED ASSEMBLY HUBS" tree diagram.

The main content area is titled "Human Genome Browser - hg19 assembly". It includes a "view sequences" button and a graphic of a human figure with DNA helixes. A note states: "The February 2009 human reference sequence (GRCh37) was produced by the Genome Reference Consortium. For more information about this assembly, see GRCh37 in the NCBI Assembly database." Below this is a "Sample position queries" section with a table of requests and their responses:

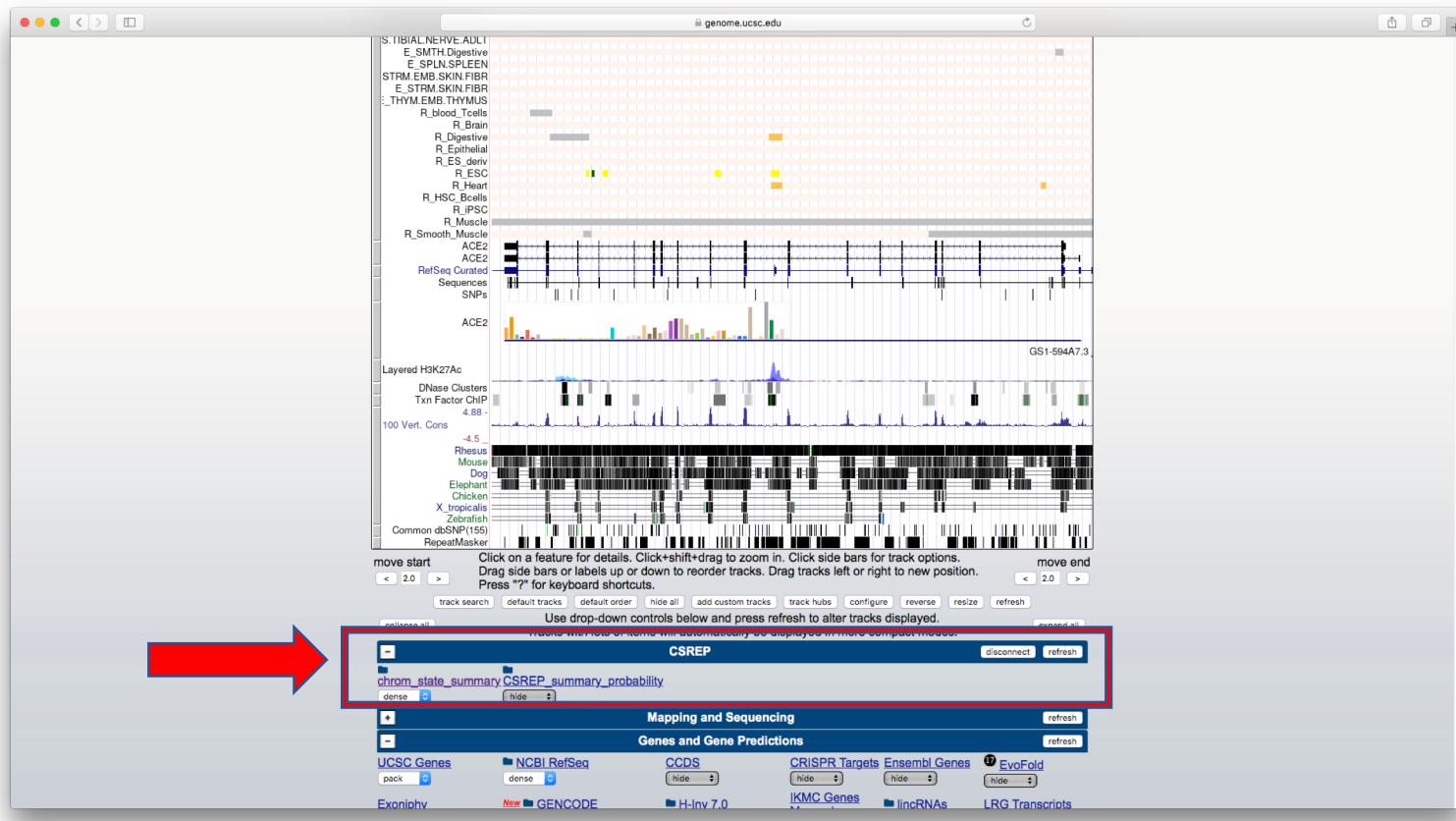
Request:	Genome Browser Response:
chr7	Displays all of chromosome 7
chrUn_gi000212	Displays all of the unplaced contig gi000212
20p13	Displays region for band p13 on chr 20
chr3:1000000	Displays first million bases of chr 3, counting from p-arm telomere
chr3:1000000+2000	Displays a region of chr3 that spans 2000 bases, starting with position 1000000
RH18061;RH80175 15q11;15q13 rs1042522;rs1800370	Displays region between genome landmarks, such as the STS markers RH18061 and RH80175, or chromosome bands 15q11 to 15q13, or SNPs rs1042522 and rs1800370. This syntax may also be used for other range queries, such as between uniquely determined ESTs, mRNAs, refSeqs, etc.
D16S3046 AA205474 AC008101 AF083811 PRNP NM_017414 NP_059110	Displays region around STS marker D16S3046 from the Genethon/Marshfield maps. Includes 100,000 bases on each side as well. Displays region of EST with GenBank accession AA205474 in BRCA1 cancer gene on chr 17 Displays region of clone with GenBank accession AC008101 Displays region of mRNA with GenBank accession number AF083811 Displays region of genome with HUGO Gene Nomenclature Committee identifier PRNP Displays the region of genome with RefSeq identifier NM_017414 Displays the region of genome with protein accession number NP_059110
pseudogene mRNA homeobox caudal zinc finger kruppel zinc finger huntington zahler	Lists transcribed pseudogenes, but not cDNAs Lists mRNAs for caudal homeobox genes Lists many zinc finger mRNAs Lists only kruppel-like zinc fingers Lists candidate genes associated with Huntington's disease Lists mRNAs deposited by scientist named Zahler

# Summary chromatin state maps for 11 groups in Roadmap and 75 groups in Epimap are displayed

- Group name follows the format <project>\_<cell group>, where project can be E (Epimap) or R (Roadmap)
- Add. File 2 in manuscript shows all the metadata of samples for each group.



The CSREP tab (scroll down) will show all the available data that users can choose to view



# If you click 'chrom\_state\_summary\_assignment\_by\_CSREP'

- You can (un)select the cell groups that you would like to view.

chrom\_state\_summary Track Settings

chrom\_state\_summary\_assignment\_by\_CSREP

Display mode: dense   Submit  Cancel Reset to defaults

All subtracks:

List subtracks:  only selected/visible  all (86 of 86 selected)

Subtrack	Description	Schema
E.ADIP.TISSUE	CSREP summary track for ADIP.TISSUE from epimap	Schema
E.BLD.CD4T	CSREP summary track for BLD.CD4T from epimap	Schema
E.BLD.CD8T	CSREP summary track for BLD.CD8T from epimap	Schema
E.BLD.MONONUC	CSREP summary track for BLD.MONONUC from epimap	Schema
E.BLD.T	CSREP summary track for BLD.T from epimap	Schema
E.BONE.EMB	CSREP summary track for BONE.EMB from epimap	Schema
E.BRN.CAUD.NUC	CSREP summary track for BRN.CAUD.NUC from epimap	Schema
E.BRN.EMB.BRN	CSREP summary track for BRN.EMB.BRN from epimap	Schema
E.BRN.HPC	CSREP summary track for BRN.HPC from epimap	Schema
E.CA.BCL	CSREP summary track for CA.BCL from epimap	Schema
E.CA.BRCA	CSREP summary track for CA.BRCA from epimap	Schema
E.CA.CRC	CSREP summary track for CA.CRC from epimap	Schema
E.CA.GVX	CSREP summary track for CA.GVX from epimap	Schema
E.CA.GLIOBM	CSREP summary track for CA.GLIOBM from epimap	Schema
E.CA.HCC	CSREP summary track for CA.HCC from epimap	Schema
E.CA.LUETPITH	CSREP summary track for CA.LUETPITH from epimap	Schema
E.CA.MELAN	CSREP summary track for CA.MELAN from epimap	Schema
E.CA.ML	CSREP summary track for CA.ML from epimap	Schema
E.CA.MYELOMA	CSREP summary track for CA.MYELOMA from epimap	Schema
E.CA.NEUROBM	CSREP summary track for CA.NEUROBM from epimap	Schema
E.CA.PRAD	CSREP summary track for CA.PRAD from epimap	Schema
E.CA.RCC	CSREP summary track for CA.RCC from epimap	Schema
E.CA.UCEC	CSREP summary track for CA.UCEC from epimap	Schema
E.EDCR.ADRENAL.GLD	CSREP summary track for EDCR.ADRENAL.GLD from epimap	Schema
E.EDCR.EMB.ADRENAL.GLD	CSREP summary track for EDCR.EMB.ADRENAL.GLD from epimap	Schema
E.EDCR.ENDOC.PANC	CSREP summary track for EDCR.ENDOC.PANC from epimap	Schema
E.EDCR.REPR.OVARY	CSREP summary track for EDCR.REPR.OVARY from epimap	Schema
E.EDCR.THYROID.GLD	CSREP summary track for EDCR.THYROID.GLD from epimap	Schema
E.EPHT.BREAST.EPITH	CSREP summary track for EPHT.BREAST.EPITH from epimap	Schema
E.EPTH.I.FORESKIN.KRTN	CSREP summary track for EPTH.I.FORESKIN.KRTN from epimap	Schema
E.ESC	CSREP summary track for ESC from epimap	Schema
E.ESDV.EMB	CSREP summary track for ESDV.EMB from epimap	Schema
E.EYE.EMB	CSREP summary track for EYE.EMB from epimap	Schema
E.GI.EMB.LARGE.INT	CSREP summary track for GI.EMB.LARGE.INT from epimap	Schema
E.GI.EMB.SMALL.INT	CSREP summary track for GI.EMB.SMALL.INT from epimap	Schema
E.GI.EMB.STOMACH	CSREP summary track for GI.EMB.STOMACH from epimap	Schema
E.GI.PEYER.PATCH	CSREP summary track for GI.PEYER.PATCH from epimap	Schema
E.GI.SIG.COLON	CSREP summary track for GI.SIG.COLON from epimap	Schema

# If you click 'state\_summary\_probability\_by\_CSREP'

You can select to view the tracks showing probabilities of state assignment for different chromatin states, in different cell groups.

The screenshot shows the 'CSREP\_summary\_probability Track Settings' page. At the top, there are settings for 'Display mode' (set to 'bar'), 'Track height' (20 pixels), 'Data view scaling' (using vertical viewing range setting), 'Vertical viewing range' (min: 0, max: 1), 'Transform function' (NONE), 'Windowing function' (mean), 'Smoothing window' (0 pixels), and 'Negate values' (unchecked). Below these are options for 'Draw y indicator lines' at y = 0.0 and y = 1.0. A 'Graph configuration help' link is also present. The main area is titled 'List subtracks' with a radio button for 'only selected/visible' (selected) and 'all' (1548 of 1548 selected). It displays a table with columns: tissue<sup>1</sup>, state<sup>2</sup>, and Track Name<sup>3</sup>. The table lists numerous tracks, each with a 'Configure' link and a 'hide' checkbox. The tracks include various chromatin states across different tissues and cell types, such as EnhA2, EnhWk, ZNFRpts, Het, TssBiv, EnhBiv, ReprPC, ReprPCWk, Quies, TssA, TssFlnk, TssFlnkU, TssFlnkD, Tx, TxWk, EnhC1, EnhG2, EnhA1, BLD\_CD4T, EnhWk, ZNFRpts, Het, TssBiv, EnhBiv, ReprPC, and BLD\_CD4T\_16\_ReprPC. Most entries have a 'Schema' link next to them.

tissue <sup>1</sup>	state <sup>2</sup>	Track Name <sup>3</sup>	Schema
hide	Configure epimap ADIP.TISSUE	10 EnhA2	Schema
hide	Configure epimap ADIP.TISSUE	11 EnhWk	Schema
hide	Configure epimap ADIP.TISSUE	12 ZNFRpts	Schema
hide	Configure epimap ADIP.TISSUE	13 Het	Schema
hide	Configure epimap ADIP.TISSUE	14 TssBiv	Schema
hide	Configure epimap ADIP.TISSUE	15 EnhBiv	Schema
hide	Configure epimap ADIP.TISSUE	16 ReprPC	Schema
hide	Configure epimap ADIP.TISSUE	17 ReprPCWk	Schema
hide	Configure epimap ADIP.TISSUE	18 Quies	Schema
hide	Configure epimap ADIP.TISSUE	1 TssA	Schema
hide	Configure epimap ADIP.TISSUE	2 TssFlnk	Schema
hide	Configure epimap ADIP.TISSUE	3 TssFlnkU	Schema
hide	Configure epimap ADIP.TISSUE	4 TssFlnkD	Schema
hide	Configure epimap ADIP.TISSUE	5 Tx	Schema
hide	Configure epimap ADIP.TISSUE	6 TxWk	Schema
hide	Configure epimap ADIP.TISSUE	7 EnhC1	Schema
hide	Configure epimap ADIP.TISSUE	8 EnhG2	Schema
hide	Configure epimap ADIP.TISSUE	9 EnhA1	Schema
hide	Configure epimap BLD.CD4T	10 EnhA2	Schema
hide	Configure epimap BLD.CD4T	11 EnhWk	Schema
hide	Configure epimap BLD.CD4T	12 ZNFRpts	Schema
hide	Configure epimap BLD.CD4T	13 Het	Schema
hide	Configure epimap BLD.CD4T	14 TssBiv	Schema
hide	Configure epimap BLD.CD4T	15 EnhBiv	Schema
hide	Configure epimap BLD.CD4T	16 ReprPC	Schema
hide	Configure epimap BLD.CD4T	17 ReprPCWk	Schema
hide	Configure colmap BLD.CD4T		Schema

# Feedbacks and questions

- If you have any feedbacks or questions, please contact Prof. Jason Ernst ([jason.ernst@ucla.edu](mailto:jason.ernst@ucla.edu)) or graduate student Ha Vu ([havu73@ucla.edu](mailto:havu73@ucla.edu)). We are happy to help you.