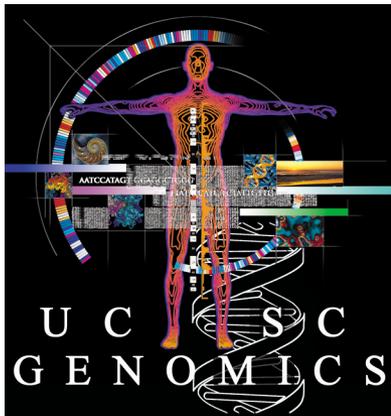


Visualizing ENCODE Data in the UCSC Genome Browser



Pauline Fujita, Ph.D.
UCSC Genome Bioinformatics Group

Training Resources

genome@soe.ucsc.edu

- Genomewiki: genomewiki.ucsc.edu
- Mailing list archives:
genome.ucsc.edu/FAQ/
- Training page:
genome.ucsc.edu/training.html
- Twitter  @GenomeBrowser
- Tutorial videos: YouTube channel
- Open Helix: openhelix.com/ucsc

Outline

- Basics: search, display, more info
- Tools for finding ENCODE data
- Annotating a BED file: RNAseq example
- Annotating a VCF file
- Track Hubs: What are they? How do I make one?
- Exercises

Basic Navigation: Main Display

genome.ucsc.edu/cgi-bin/hgTracks?db=hg19

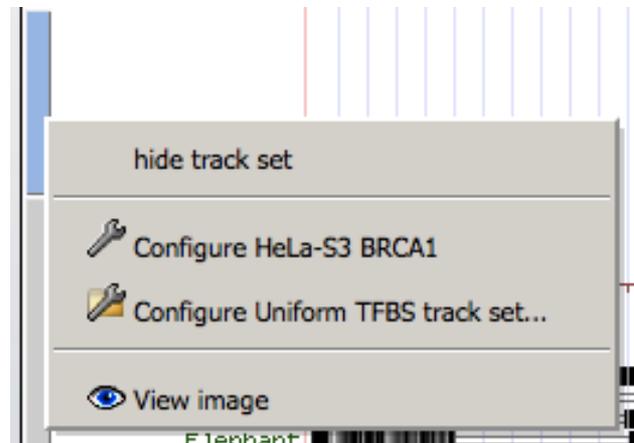
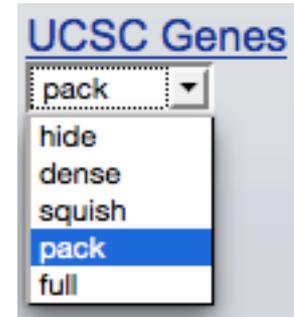
The screenshot displays the UCSC Genome Browser interface for the human genome (hg19). The main display area is titled "UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly". The current view is centered on chromosome 21, specifically the region chr21:33,031,597-33,041,570 (9,974 bp). The interface includes a navigation bar at the top with links for Genomes, Genome Browser, Tools, Mirrors, Downloads, My Data, View, Help, and About Us. Below the navigation bar, there are controls for moving and zooming, including a search box for coordinates or gene symbols and a "go" button. The main display area is divided into several tracks: a scale bar (2 kb), a track for UCSC Genes (RefSeq, GenBank, CCDS, Rfam, tRNAs & Comparative Genomics), a track for Sequences (Publications: Sequences in Scientific Articles), a track for SNPs, a track for Layered H3K27Ac, a track for DNase Clusters (dnase1 hypersensitivity clusters in 126 cell types from ENCODE v3), a track for Txn Factor CHIP (Transcription Factor CHIP-seq (161 factors) from ENCODE with Factorbook Motifs), a track for 100 Vert. Cons (100 vertebrates Basewise Conservation by PhyloP), and a track for Multiz Alignments of 100 Vertebrates (Rhesus, Mouse, Dog, Elephant, Chicken, X_tropicalis, Zebrafish, Lamprey). The interface also includes a "More on-site workshops available!" link and a "move start" / "move end" control at the bottom left and right. At the bottom, there are buttons for track search, default tracks, default order, hide all, add custom tracks, track hubs, configure, reverse, resize, refresh, collapse all, and expand all. A footer bar at the very bottom contains the text "Mapping and Sequencing" and a "refresh" button.

Click on a feature for details. Click or drag in the base position track to zoom in. Click side bars for track options. Drag side bars or labels up or down to reorder tracks. Drag tracks left or right to new position.

Use drop-down controls below and press refresh to alter tracks displayed. Tracks with lots of items will automatically be displayed in more compact modes.

Display Configuration

- Visibility:
hide, dense, squish, pack, full
- Track ordering: drag and drop
- Drag and zoom/highlighting
- Configuration page
- Right click menu



How to find more info

The image shows a genomic browser interface with several tracks. A red box highlights a vertical bar in the RepeatMasker track. Another red box highlights a blue bar in the UCSC Genes track. A white box labeled 'Item Description' points to the UCSC Genes track. Another white box labeled 'Track Description' points to the UCSC Genes track in the track selection area. The interface includes a 'move start' control with a value of 2.0, a 'track search' input, and various track management buttons like 'default tracks', 'hide all', and 'add custom tracks'. The track selection area is titled 'Mapping and Sequencing' and 'Genes and Gene Predictions', listing tracks like UCSC Genes, RefSeq Genes, AceView Genes, CCDS, Ensembl Genes, Exoniphy, GENCODE..., Geneid Genes, Genscan Genes, and H-Inv 7.0.

RepeatMasker

UCSC Genes (RefSeq, GenBank, CCDS, Rfam, tRNAs & Comparative Genomics)

Item Description

Track Description

move start
< 2.0 >

track search default tracks default order hide all add custom tracks track hubs configure reverse

Use drop-down controls below and press refresh to alter tracks displayed.
Tracks with lots of items will automatically be displayed in more compact modes

Mapping and Sequencing

Genes and Gene Predictions

UCSC Genes RefSeq Genes AceView Genes CCDS Ensembl Genes

Exoniphy GENCODE... Geneid Genes Genscan Genes H-Inv 7.0

More info: Track Description

[Genomes](#)[Genome Browser](#)[Tools](#)[Mirrors](#)[Downloads](#)[My Data](#)[Help](#)[About Us](#)

UCSC Genes Track Settings

UCSC Genes (RefSeq, GenBank, CCDS, Rfam, tRNAs & Comparative Genomics) ([▲All Genes and Gene Predictions tracks](#))

Display mode:

Label: gene symbol UCSC Known Gene ID UniProt Display ID OMIM ID

Show: non-coding genes splice variants

Color track by codons: [Help on codon coloring](#)

Show codon numbering:

[View table schema](#)

Data last updated: 2013-06-14

Description

The UCSC Genes track is a set of gene predictions based on data from RefSeq, GenBank, CCDS, Rfam, and the [tRNA Genes](#) track. The track includes both protein-coding genes and non-coding RNA genes. Both types of genes can produce non-coding transcripts, but non-coding RNA genes do not produce protein-coding transcripts. This is a moderately conservative set of predictions. Transcripts of protein-coding genes require the support of one RefSeq RNA, or one GenBank RNA sequence plus at least one additional line of evidence. Transcripts of non-coding RNA genes require the support of one Rfam or tRNA prediction. Compared to RefSeq, this gene set has generally about 10% more protein-coding genes, approximately four times as many putative non-coding genes, and about twice as many splice variants.

More info: Item Description

[Genomes](#)[Genome Browser](#)[Tools](#)[Mirrors](#)[Downloads](#)[My Data](#)[Help](#)[About Us](#)

Human Gene SCAF4 (uc002ypd.2) Description and Page Index

Description: Homo sapiens SR-related CTD-associated factor 4 (SCAF4), transcript variant 1, mRNA.

RefSeq Summary (NM_020706): This gene likely encodes a member of the arginine/serine-rich splicing factor family. A similar protein in Rat appears to bind the large subunit of RNA polymerase II and provide a link between transcription and pre-mRNA splicing. Alternatively spliced transcript variants have been described. [provided by RefSeq, Feb 2009].

Transcript (Including UTRs)

Position: chr21:33,043,313-33,104,431 **Size:** 61,119 **Total Exon Count:** 20 **Strand:** -

Coding Region

Position: chr21:33,043,712-33,104,005 **Size:** 60,294 **Coding Exon Count:** 20

Page Index	Sequence and Links	UniProtKB Comments	CTD	Gene Alleles	RNA Structure
Protein Structure	Other Species	GO Annotations	mRNA Descriptions	Other Names	Model Information
Methods					

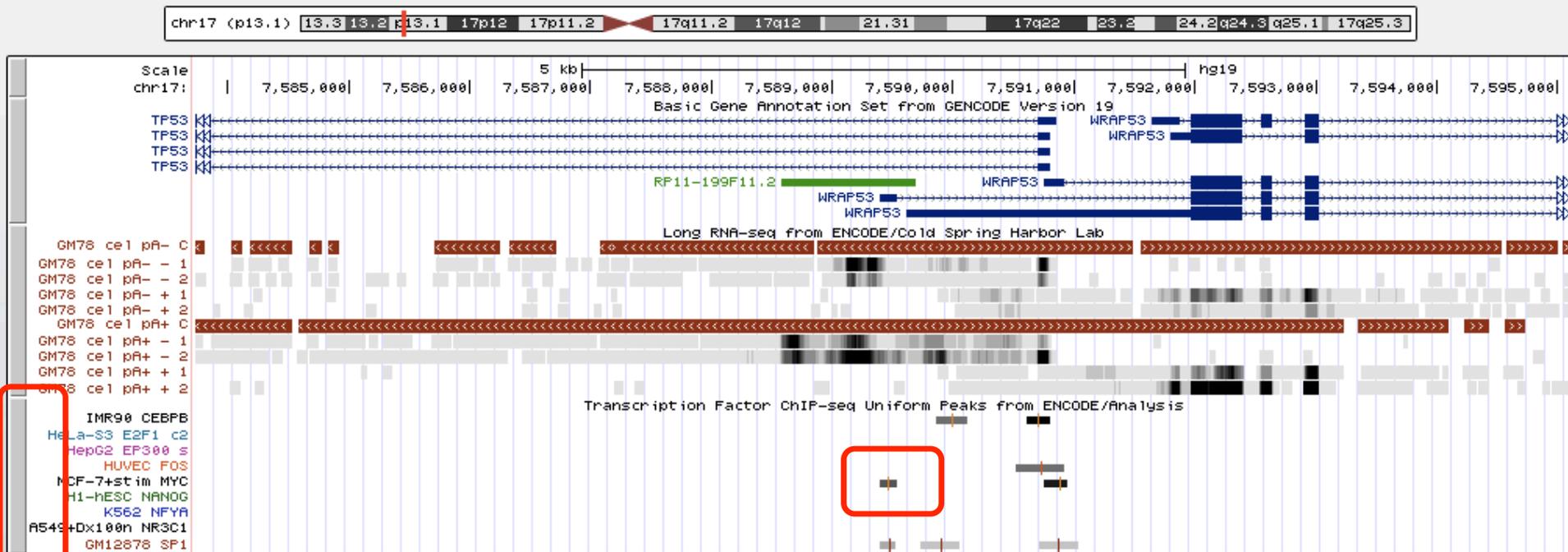
Data last updated: 2013-06-14

Sequence and Links to Tools and Databases

Genomic Sequence (chr21:33,043,313-33,104,431)		mRNA (may differ from genome)		Protein (1147 aa)	
Gene Sorter	Genome Browser	Protein FASTA	VisiGene	Table Schema	BioGPS
CGAP	Ensembl	Entrez Gene	ExonPrimer	GeneCards	GeneNetwork
Gepis Tissue	H-INV	HGNC	HPRD	Jackson Lab	Lynx
MOPED	neXtProt	OMIM	PubMed	Stanford SOURCE	UniProtKB

Comments and Description Text from UniProtKB

ENCODE



move start Click on a feature for details. Click or drag in the base position track to zoom in. Click side bars for track options. Drag side bars or labels up or down to reorder tracks. Drag tracks left or right to new position. move end

Regulation refresh

<input checked="" type="checkbox"/> ENCODE Regulation... hide	<input checked="" type="checkbox"/> CD34 Dnase1 hide	<input checked="" type="checkbox"/> CpG Islands... hide	<input checked="" type="checkbox"/> ENC Chromatin... hide	<input checked="" type="checkbox"/> ENC DNA Methyl... hide	<input checked="" type="checkbox"/> ENC DNase/FAIRE... hide
<input checked="" type="checkbox"/> ENC Histone... hide	<input checked="" type="checkbox"/> ENC RNA Binding... hide	<input checked="" type="checkbox"/> ENC TF Binding... show	<input checked="" type="checkbox"/> FSU Repli-chip hide	<input checked="" type="checkbox"/> Genome Segments hide	<input checked="" type="checkbox"/> NKI Nuc Lamina... hide
<input checked="" type="checkbox"/> Stanf		<input checked="" type="checkbox"/> SUNY	<input checked="" type="checkbox"/> SwitchGear		

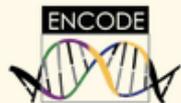
ENCODE: Super-track Settings

Regulation refresh

<input checked="" type="checkbox"/> ENCODE Regulation... hide	<input checked="" type="checkbox"/> 18 CD34 DnaseI hide	<input checked="" type="checkbox"/> CpG Islands... hide	<input checked="" type="checkbox"/> ENC Chromatin... hide	<input checked="" type="checkbox"/> ENC DNA Methyl... hide	<input checked="" type="checkbox"/> ENC DNase/FAIRE... hide
<input checked="" type="checkbox"/> ENC Histone... hide	<input checked="" type="checkbox"/> ENC RNA Binding... hide	<input checked="" type="checkbox"/> ENC TF Binding... show	<input checked="" type="checkbox"/> FSU Repli-chip hide	<input checked="" type="checkbox"/> Genome Segments hide	<input checked="" type="checkbox"/> 18 NKI Nuc Lamina... hide

[Home](#) [Genomes](#) [Genome Browser](#) [Tools](#) [Mirrors](#) [Downloads](#) [My Data](#) [Help](#) [About Us](#)

ENC TF Binding Super-track Settings



ENCODE Transcription Factor Binding Tracks ([^All Regulation tracks](#))

Display mode:

- All**
- | | |
|---|--|
| <input checked="" type="checkbox"/> <input type="text" value="dense"/> Uniform TFBS | Transcription Factor ChIP-seq Uniform Peaks from ENCODE/Analysis <small>ENCODE March 2012 Freeze</small> |
| <input type="checkbox"/> <input type="text" value="hide"/> HAIB TFBS | Transcription Factor Binding Sites by ChIP-seq from ENCODE/HAIB |
| <input type="checkbox"/> <input type="text" value="hide"/> SYDH TFBS | Transcription Factor Binding Sites by ChIP-seq from ENCODE/Stanford/Yale/USC/Harvard |
| <input type="checkbox"/> <input type="text" value="hide"/> UChicago TFBS | Transcription Factor Binding Sites by Epitope-Tag from ENCODE/UChicago |
| <input type="checkbox"/> <input type="text" value="hide"/> UTA TFBS | Open Chromatin TFBS by ChIP-seq from ENCODE/Open Chrom(UT Austin) <small>ENCODE July 2011 Freeze</small> |
| <input type="checkbox"/> <input type="text" value="hide"/> UW CTCF Binding | CTCF Binding Sites by ChIP-seq from ENCODE/University of Washington |

Description

ENCODE: Item Details



move start move end

Click on a feature for details. Click or drag in the base position track to zoom in. Click side bars for track options. Drag side bars or labels up or down to reorder tracks. Drag tracks left or right to new position.

[Genomes](#) [Genome Browser](#) [Tools](#) [Mirrors](#) [Downloads](#) [My Data](#) [Help](#) [About Us](#)

MCF-7 (serum_stimulated) TFBS Uniform Peaks of c-Myc from ENCODE/UT-A/Analysis

Position: [chr17:7589391-7589530](#)

Peak point: 7589461

Score: 688

Signal value: 107.768

Q-value (FDR): 4.212

View table: [schema](#), [downloads](#), [metadata](#) ▾

[Go to Uniform TFBS track controls](#)

Data version: ENCODE Mar 2012 Freeze

Data last updated: 2013-04-12

Description

This track represents a comprehensive set of human transcription factor binding sites based on ChIP-seq experiments generated by production in the ENCODE Consortium from the inception of the project in September 2007, through the March 2012 internal data freeze. The track represents

ENCODE Tools

UCSC Genome Bioinformatics

[Genomes](#) - [Blat](#) - [Tables](#) - [Gene Sorter](#) - [PCR](#) - [VisiGene](#) - [Session](#) - [FAQ](#) - [Help](#)

- [Genome Browser](#)
- [Ebola](#)
- [Blat](#)
- [Table Browser](#)
- [Gene Sorter](#)
- [In Silico PCR](#)
- [Genome Graphs](#)
- [Galaxy](#)
- [VisiGene](#)
- [Utilities](#)
- [Downloads](#)
- [Release Log](#)
- [Custom Tracks](#)
- [Cancer Browser](#)
- [Microbial Genomes](#)
- [ENCODE](#)

About the UCSC Genome Bioinformatics Site

Welcome to the UCSC Genome Browser website. This site contains the reference sequence and working draft assemblies for a large collection of genomes. It also provides portals to [ENCODE](#) data at UCSC (2003 to 2012) and to the [Neandertal](#) project. Download or purchase the Genome Browser source code, or the Genome Browser in a Box ([GBIB](#)) at our [online store](#).

We encourage you to explore these sequences with our tools. The [Genome Browser](#) zooms and scrolls over chromosomes, showing the work of annotators worldwide. The [Gene Sorter](#) shows expression, homology and other information on groups of genes that can be related in many ways. [Blat](#) quickly maps your sequence to the genome. The [Table Browser](#) provides convenient access to the underlying database. [VisiGene](#) lets you browse through a large collection of *in situ* mouse and frog images to examine expression patterns. [Genome Graphs](#) allows you to upload and display genome-wide data sets.

The UCSC Genome Browser is developed and maintained by the Genome Bioinformatics Group, a cross-departmental team within the [UC Santa Cruz Genomics Institute](#) and the Center for Biomolecular Science and Engineering ([CBSE](#)) at the University of California Santa Cruz ([UCSC](#)). If you have feedback or questions concerning the tools or data on this website, feel free to contact us on our [public mailing list](#).

The Genome Browser project team relies on public funding to support our work. Donations are welcome -- we have many more ideas than our funding supports! If you have ideas, drop a comment in our [suggestion box](#).

[DONATE NOW](#)

News

[News Archives](#) ►

To receive announcements of new genome assembly releases, new software features, updates and training seminars by email, subscribe to the [genome-announce](#) mailing list. Please see our [blog](#) for posts about Genome Browser tools, features, projects and more.

03 March 2015 - Use UDR for Quick Sizable Downloads

The UCSC Genome Browser is pleased to share a download protocol to use when downloading large sets of files from our download servers: **UDR (UDT Enabled Rsync)**. UDR utilizes rsync as the transport mechanism, but sends the data over the UDT protocol, which enables huge amounts of data to be downloaded efficiently over long distances.

Protocols like http, ftp and rsync can have a problem in that the further away the download source is from the user, the slower the speed becomes. Protocols like UDT/UDR allow for many UDP packets to be sent in batch, thus allowing for much higher transmit speeds over long distances. UDR will be especially useful for users who are downloading from places that are far away from California. The US East Coast and the international community will likely see much higher download speeds by using UDR rather than rsync, http or ftp. [Read more](#).

ENCODE: Experiment Matrix

Assays

search for: tracks files

[view matrix](#)

Cell Types

	DNA Methylation	Methyl Array	Methyl RRBS	Open Chromatin	DNase-DGF	DNase-seq	FAIRE-seq	RNA Binding Proteins	RIP Gene ST	RIP Tiling Array	RIP Validation	RIP-seq	RNA Profiling	CAGE	Exon Array	RNA-chip	RNA-PET	RNA-seq	Small RNA-seq	TFBS & Histones	ChIP-seq	Other
Tier 1																						
GM12878		1	1			2	1		7	4		4		6	2	6	2	12	5		133	
H1-hESC		1	1			2	1		3					4	1		1	10	3		91	
K562		1	1		3	16	3		6	4		4		9	7	9	6	17	7		224	
Tier 2																						
A549		1	1		1	2	1							3	2		3	10	9		87	
CD20+														1				2	1		4	
CD20+_RO01778					1	1															2	
CD20+_RO01794						1															5	
H1-neurons																			3		4	
HeLa-S3		1	1			3	3		4					6	4		3	8	3		93	
HeLa-S3		1	1		1	2	1		4					6	2	5	2	2	2		114	

ENCODE: Experiment Summary



ENCODE Experiment Summary *hg19* 2007 - 2012

search for: tracks files

Genome Annotation	
GENCODE	4
Mappability	10
ORChID	1
Total: 3	15

All Experiments		view matrix by cell line
5C	13	
CAGE	78	
ChIA-PET	8	
ChIP-seq	1479	
Combined	26	
DNA-PET	6	
DNase-DGF	56	
DNase-seq	219	
Exon Array	158	
FAIRE-seq	37	
Genotype	64	
Methyl Array	62	

ChIP-seq Experiments		view matrix by cell line
ARID3A	2	
ATF1	1	
ATF2	2	
ATF3	7	
BACH1	2	
BATF	1	
BCL11A	3	
BCL3	3	
BCLAF1	2	
BDP1	2	
BHLHE40	5	
BRCA1	4	

ENCODE: Track Search

Search for Tracks in the Human Feb. 2009 (GRCh37/hg19) Assembly

Search

Advanced

Track Name: contains

and Description: contains

and Group: is Any

and Data Format: is Any

ENCODE terms

+ and Antibody or target protein is among [Antibody or target protein](#)

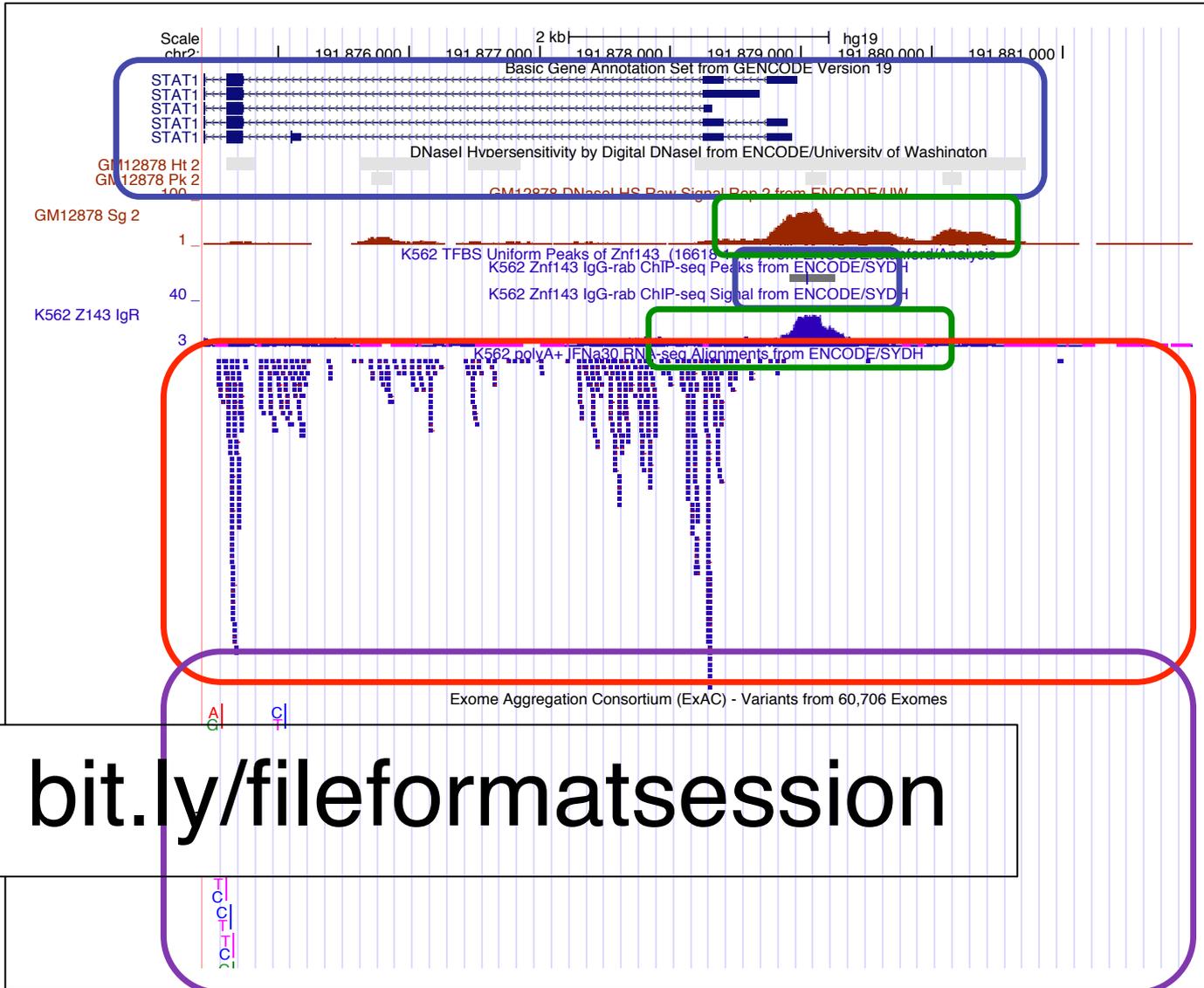
+ and View - Peaks or Signals is among

[return to browser](#)

(12 of 12 selected)

+ -	Visibility	Track Name	Sort: <input checked="" type="radio"/> by Relevance <input type="radio"/> Alphabetically <input type="radio"/> by Hierarchy
<input checked="" type="checkbox"/>	pack <input type="button" value="v"/>	GM12878 BRCA1	GM12878 TFBS Uniform Peaks of BRCA1_(A300-000A) from ENCODE/Stanford/Analysis ▾
<input checked="" type="checkbox"/>	pack <input type="button" value="v"/>	H1-hESC BRCA1	H1-hESC TFBS Uniform Peaks of BRCA1_(A300-000A) from ENCODE/Stanford/Analysis ▾
<input checked="" type="checkbox"/>	pack <input type="button" value="v"/>	HeLa-S3 BRCA1	HeLa-S3 TFBS Uniform Peaks of BRCA1_(A300-000A) from ENCODE/Stanford/Analysis ▾
<input checked="" type="checkbox"/>	pack <input type="button" value="v"/>	HepG2 BRCA1 <small>Display track details</small>	HepG2 TFBS Uniform Peaks of BRCA1_(A300-000A) from ENCODE/Stanford/Analysis ▾
<input checked="" type="checkbox"/>	pack <input type="button" value="v"/>	GM78 BRC1 IgM	GM12878 BRCA1 IgG-mus ChIP-seq Peaks from ENCODE/SYDH ▾
<input checked="" type="checkbox"/>	full <input type="button" value="v"/>	GM78 BRC1 IgM	GM12878 BRCA1 IgG-mus ChIP-seq Signal from ENCODE/SYDH ▾
<input checked="" type="checkbox"/>	pack <input type="button" value="v"/>	H1ES BRC1 IgR	H1-hESC BRCA1 IgG-rab ChIP-seq Peaks from ENCODE/SYDH ▾
<input checked="" type="checkbox"/>	full <input type="button" value="v"/>	H1ES BRC1 IgR	H1-hESC BRCA1 IgG-rab ChIP-seq Signal from ENCODE/SYDH ▾
<input checked="" type="checkbox"/>	pack <input type="button" value="v"/>	HeLa BRC1 IgR	HeLa-S3 BRCA1 IgG-rab ChIP-seq Peaks from ENCODE/SYDH ▾
<input checked="" type="checkbox"/>	full <input type="button" value="v"/>	HeLa BRC1 IgR	HeLa-S3 BRCA1 IgG-rab ChIP-seq Signal from ENCODE/SYDH ▾
<input checked="" type="checkbox"/>	pack <input type="button" value="v"/>	HEPG BRC1 IgR	HEPG2 BRCA1 IgG-rab ChIP-seq Peaks from ENCODE/SYDH ▾
<input checked="" type="checkbox"/>	full <input type="button" value="v"/>	HEPG BRC1 IgR	HEPG2 BRCA1 IgG-rab ChIP-seq Signal from ENCODE/SYDH ▾

File Formats



BED

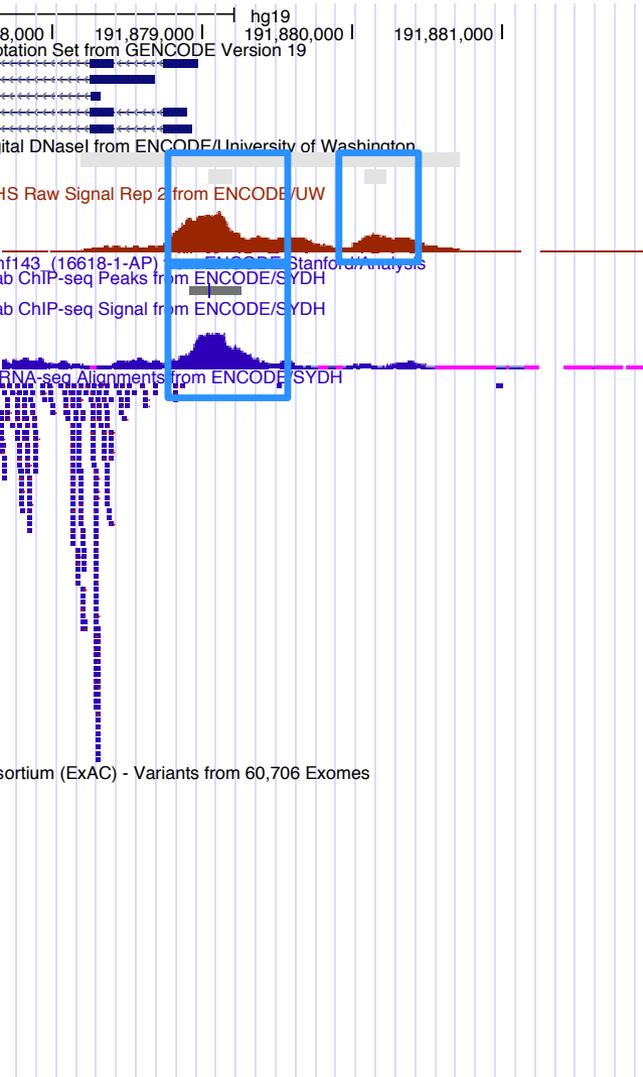
wig(gle)

BAM

VCF

bit.ly/fileformatsession

File Formats



BED

Positional annotations. (ex. Regions w/: enriched ChIP-seq signal for TF binding, Δ I methylation, splice jxns from RNA-seq)

wig(gle)

Continuous signal data. # of reads (ex. DNase I HS and ChIP-seq signals)

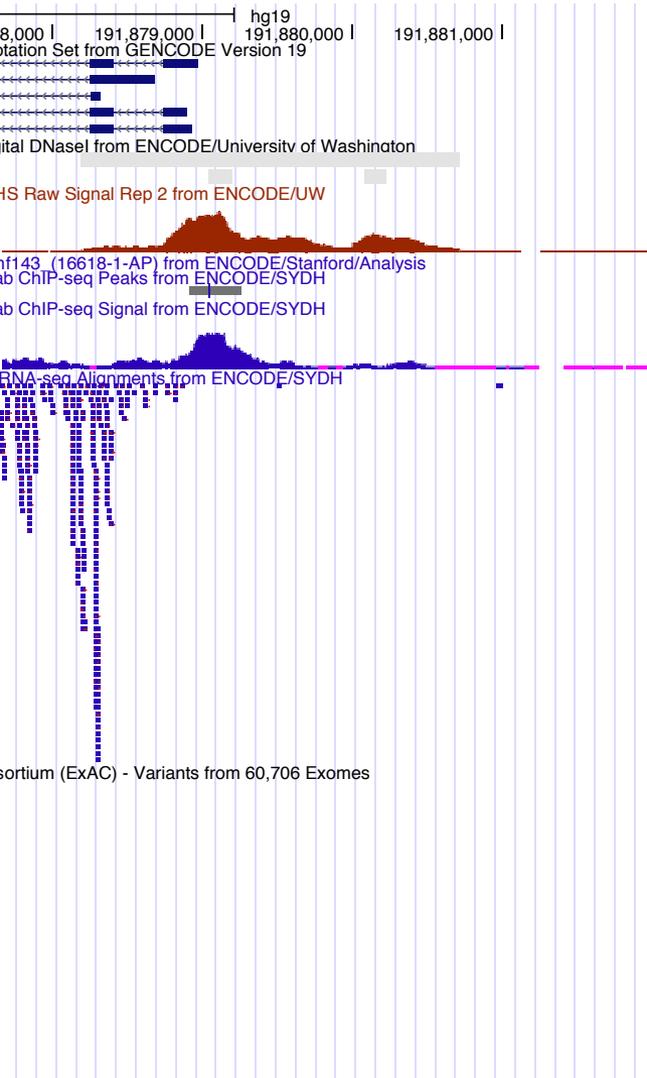
BAM

Alignments of seq. reads, mapped to genome (ex. RNA-seq alignments)

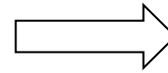
VCF

Variation data: SNPs, indels, Copy Number Variants, Structural Variants (ex. ExAC data)

Indexed File Formats

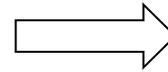


BED



bigBed

wig(gle)



bigWig

BAM

VCF

Indexed File Formats

- Only displayed portions of files transferred to UCSC
- Display large files (would time out)
- File + index on your web-accessible server (http, https, or ftp)
- Faster display
- More user control

File Formats

[Genomes](#)[Genome Browser](#)[Tools](#)[Mirrors](#)[Downloads](#)[My Data](#)[Help](#)[About Us](#)

Add Custom Tracks

clade genome assembly

Display your own data as custom annotation tracks in the browser. Data must be formatted in [BED](#), [bigBed](#), [bedGraph](#), [GFF](#), [GTF](#), [WIG](#), [bigWig](#), [MAF](#), [BAM](#), [BED detail](#), [Personal Genome SNF](#), [VCF](#), [broadPeak](#), [narrowPeak](#), or [PSL](#) formats. To configure the display, set [track](#) and [browser](#) line attributes as described in the [User's Guide](#). Data in the bigBed, bigWig, BAM and VCF formats can be provided via only a URL or embedded in a track line in the box below. Publicly available custom tracks are listed [here](#). Examples are [here](#).

Paste URLs or data:

Or upload:

No file selected.

File Formats

Genomes Genome Browser Tools Mirrors Downloads My Data View Help About Us

UCSC Genome Browser on Human Feb. 2009 (GRCh37)

move <<< << < > >> >>> zoom in 1.5x 3x 10x base ZOOM

chr21:33,031,597-33,041,570 9,974 bp. enter position, gene symbol or search terms

chr21 (q22.11) p13 p12 p11.2 21q21.1 21.2 21q21.3 q22.1

Scale chr21: 33,035,000 2 kb 33,040,000

UCSC Genes (RefSeq, GenBank, CCDS, Rfam, tRNAs & Comparative Genomics)

BC041449 SOD1

UCSC Genome Bioinformatics

Home - Genomes - Blat - Tables - Gene Sorter - PCR - VisiGene - Session - Help

FAQ Table of Contents

This page contains responses to questions frequently asked by our user community and subscribers to the [Genome Browser mailing list](#).

- [Display Problems](#)
- [Assembly Releases and Versions](#)
- [Data and Downloads](#)
- [Genome Browser Tracks](#)
- [Custom Annotation Tracks](#)

- [Data File Formats](#)

- [Blat](#)

- [ENCODE Resources and FAQ](#)

File Formats

UCSC Genome Bioinformatics

[Home](#) - [Genomes](#) - [Blat](#) - [Tables](#) - [Gene Sorter](#) - [PCR](#) - [VisiGene](#) - [Session](#) - [Help](#)

Frequently Asked Questions: Data File Formats

General formats:

- [Axt format](#)
- [BAM format](#)
- [BED format](#)
- [BED detail format](#)
- [bedGraph format](#)
- [bigBed format](#)
- [bigGenePred table format](#)
- [bigWig format](#)

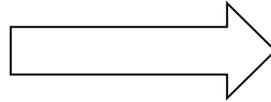
ENCODE-specific formats:

- [ENCODE broadPeak format](#)
- [ENCODE gappedPeak format](#)
- [ENCODE narrowPeak format](#)
- [ENCODE pairedTagAlign format](#)
- [ENCODE peptideMapping format](#)
- [ENCODE RNA elements format](#)
- [ENCODE tagAlign format](#)

File Formats

www.encodeproject.org/help/file-formats/

Help



File formats

ENCODE Data ▾ Methods ▾ About ENCODE ▾ Help ▾ Search ENCODE 🔍 Sign In

Getting started
REST API
File formats
Tutorials
Contact

Common File Formats the ENCODE Consortium

Overview

The ENCODE consortium uses several file formats to store, display, and disseminate data:

- **FASTQ**: a text-based format for storing nucleotide sequences (reads) and their quality scores. [1]
- **BAM**: The Sequence Alignment/Mapping (SAM) format is a text-based format for storing read alignments against reference sequences and it is interconvertible with the binary BAM format. [2]
- **bigWig**: The bigWig format is an indexed binary format for rapid display of continuous and dense data in the UCSC Genome Browser.
- **bigBed**: The bigBed format is also an indexed binary format for rapid display of annotation items such as a linked collection of exons or the binding peaks of a transcription factor.

These file formats were originally designed to be generic and flexible. As the ENCODE consortium is a collaborative effort, the consortium has made several specifications on the file formats to facilitate data archival, presentation, and distribution, as well as integrative analysis on the data. The consortium considers FASTQ as the basic file format for archival purpose and thus the FASTQ format's specifications aim to preserve the raw

Custom Tracks

Software Copyright (c) The Regents of the University of California. All rights reserved.

group genome assembly position search term

Mammal Human Feb. 2009 (GRCh37/hg19) chr17:7,583,734-7,595,119 enter position, gene symbol or search terms submit

[Click here to reset](#) the browser user interface settings to their defaults. **More on-site workshops available!**

track search **add custom tracks** track hubs configure tracks and display

Genomes Genome Browser Tools Mirrors Downloads My Data View Help About Us

UCSC Genome Browser on Human Feb 2009 (GRCh37/hg19) Assembly

move <<< << < > >> >>> zoom in 1.5x 3x

chr17:7,583,734-7,595,119 11,386 bp. enter position, gene symbol or search term **on-site workshops available!**

chr17 (p13.1) 13.3 13.2 13.1 17p12 17p11.2 17q11.2 17q12 21.31 17q22 23.2 24.2q24.3 q25.1 17q25.3

Scale 5 kb hg19

chr17: 7,585,000 7,586,000 7,587,000 7,588,000 7,589,000 7,590,000 7,591,000 7,592,000 7,593,000 7,594,000 7,595,000

Basic Gene Annotation Set from GENCODE Version 19

TP53 WRAP53

RP11-199F11.2 WRAP53

move start < 2.0 > move end < 2.0 >

track search default tracks default order hide all **add custom tracks** track hubs configure reverse resize refresh

Use drop-down controls below and press refresh to alter tracks displayed

Custom Tracks

genome.ucsc.edu/cgi-bin/hgCustom



Genomes

Genome Browser

Tools

Mirrors

Downloads

My Data

Help

About Us

Add Custom Tracks

clade genome assembly

Display your own data as custom annotation tracks in the browser. Data must be formatted in [BED](#), [bigBed](#), [bedGraph](#), [GFF](#), [GTF](#), [WIG](#), [bigWig](#), [MAF](#), [BAM](#), [BED detail](#), [Personal Genome SNP](#), [VCF](#), [broadPeak](#), [narrowPeak](#), or [PSL](#) formats. To configure the display, set [track](#) and [browser](#) line attributes as described in the [User's Guide](#). Data in the bigBed, bigWig, BAM and VCF formats can be provided via only a URL or embedded in a track line in the box below. Publicly available custom tracks are listed [here](#). Examples are [here](#).

Paste URLs or data:

Or upload:

No file selected.

BED format

[Index](#) ▶

BED format provides a flexible way to define the data lines that are displayed in an annotation track. BED lines have three required fields and nine additional optional fields. The number of fields per line must be consistent throughout any single set of data in an annotation track. The order of the optional fields is binding: lower-numbered fields must always be populated if higher-numbered fields are used.

If your data set is BED-like, but it is very large and you would like to keep it on your own server, you should use the [bigBed](#) data format.

The first three required BED fields are:

1. **chrom** - The name of the chromosome (e.g. chr3, chrY, chr2_random) or scaffold (e.g. scaffold10671).
2. **chromStart** - The starting position of the feature in the chromosome or scaffold. The first base in a chromosome is numbered 0.
3. **chromEnd** - The ending position of the feature in the chromosome or scaffold. The *chromEnd* base is not included in the display of the feature. For example, the first 100 bases of a chromosome are defined as *chromStart=0*, *chromEnd=100*, and span the bases numbered 0-99.

The 9 additional optional BED fields are:

4. **name** - Defines the name of the BED line. This label is displayed to the left of the BED line in the Genome Browser window when the track is

Custom Tracks

genome.ucsc.edu/cgi-bin/hgCustom



Genomes

Genome Browser

Tools

Mirrors

Downloads

My Data

Help

About Us

Add Custom Tracks

clade genome assembly

Display your own data as custom annotation tracks in the browser. Data must be formatted in [BED](#), [bigBed](#), [bedGraph](#), [GFF](#), [GTF](#), [WIG](#), [bigWig](#), [MAF](#), [BAM](#), [BED detail](#), [Personal Genome SNP](#), [VCF](#), [broadPeak](#), [narrowPeak](#), or [PSL](#) formats. To configure the display, set [track](#) and [browser](#) line attributes as described in the [User's Guide](#). Data in the bigBed, bigWig, BAM and VCF formats can be provided via only a URL or embedded in a track line in the box below. Publicly available custom tracks are listed [here](#). Examples are [here](#).

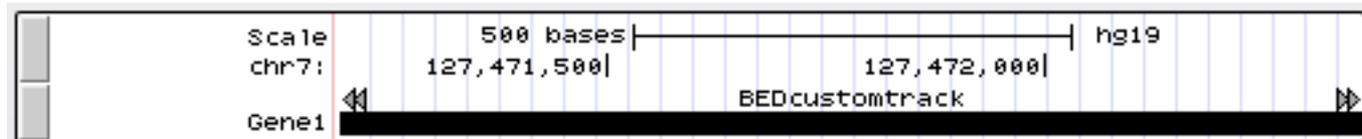
Paste URLs or data:

Or upload: No file selected.

```
track name="BED_custom_track"  
chr7 127471196 127472363 Gene1
```

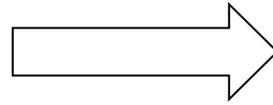
Optional track documentation:

Or upload: No file selected.



Annotating your data: BED

Tools



Data Integrator

UCSC Genome Browser

Genomes Genome Browser Tools Mirrors Downloads My Data View Help About Us

chr21:33,031,597-33,041,5...
chr21 (q22.11) 21p13 21p12

Human Feb. 2009 (GRCh37/hg19) Assembly

Scale chr21: 33,033,000

BC041449 SOD1

RefSeq Genes

Sequences SNPs

Human mRNAs

Spliced ESTs

Layered H3K27Ac

DNase Clusters

Txn Factor ChIP

100 Vert. Cons

Blat
Table Browser
Variant Annotation Integrator
Data Integrator
Gene Sorter
Genome Graphs
In-Silico PCR
LiftOver
VisiGene
Other Utilities

H3K27Ac Mark (Often Found Near Active Regulatory Elements) on 7 cell lines from ENCODE

DNaseI Hypersensitivity Clusters in 125 cell types from ENCODE (V3)

Transcription Factor ChIP-seq (161 factors) from ENCODE with Factorbook Motifs

100 vertebrates Basewise Conservation by PhyloP

Data Integrator

genome.ucsc.edu/cgi-bin/hgIntegrator

Data Integrator

Select Genome Assembly and Region

group: Mammal genome: Human assembly: Feb. 2009 (GRCh37/hg19)

region to annotate: position or search term: chr21:33031597-33041570

Configure Data Sources

- ↑ SYDH_TFBS [View table schema](#) ×
- ↑ GENCODE Genes V19 - Basic [View table schema](#) ×

Add Data Source

track group: Genes and Gene Predictions track: GENCODE Genes V19... view: Genes... subtrack: Basic (wgEncodeGencodeBasicV19)

[View table schema](#)

get more data:

Output Options

Send output to file

Data Integrator

Data Integrator Undo Redo

Select **Choose Fields** [X]

group
Mamm

region
positio

Confi

↑ SY
↑ GE

Add **Basic**
[X] [X]

track
Gene
View

get m
track

Output

Se
Choos

Get

Using

SYDH_TFBS
Set all Clear all

<input checked="" type="checkbox"/>	chrom	Reference sequence chromosome or scaffold
<input checked="" type="checkbox"/>	chromStart	Start position in chromosome
<input checked="" type="checkbox"/>	chromEnd	End position in chromosome
<input checked="" type="checkbox"/>	name	Name of item.
<input checked="" type="checkbox"/>	score	Score (0-1000)
<input type="checkbox"/>	strand	+ or - for strand

Basic
Set all Clear all

<input checked="" type="checkbox"/>	name	Name of gene (usually transcript_id from GTF)
<input type="checkbox"/>	chrom	Reference sequence chromosome or scaffold
<input type="checkbox"/>	strand	+ or - for strand
<input type="checkbox"/>	txStart	Transcription start position
<input type="checkbox"/>	txEnd	Transcription end position
<input type="checkbox"/>	cdsStart	Coding region start
<input type="checkbox"/>	cdsEnd	Coding region end
<input type="checkbox"/>	exonCount	Number of exons
<input type="checkbox"/>	exonStarts	Exon start positions
<input type="checkbox"/>	exonEnds	Exon end positions
<input type="checkbox"/>	score	score
<input checked="" type="checkbox"/>	name2	Alternate name (e.g. gene_id from GTF)
<input type="checkbox"/>	cdsStartStat	enum('none','unk','incmpl','cmpl')

Data Integrator

Data Integrator

Select Genome Assembly and Region

group genome assembly

region to annotate

position or search term

Configure Data Sources

↓ **SYDH_TFBS** [View table schema](#)

↓ **GENCODE Genes V19 - Basic** [View table schema](#)

Add Data Source

track group track view subtrack

[View table schema](#)

get more data:

Output Options

Send output to file

Data Integrator

```
#ct_SYDHTFBS_4733.chrom ct_SYDHTFBS_4733.chromStart
ct_SYDHTFBS_4733.chromEnd ct_SYDHTFBS_4733.name ct_SYDHTFBS_4733.score
wgEncodeGencodeBasicV19.name wgEncodeGencodeBasicV19.name2
```

chr21	33031473	33032186	.	608	ENST00000449339.1	
AP000253.1						
chr21	33031473	33032186	.	608	ENST00000270142.6	SOD1
chr21	33031473	33032186	.	608	ENST00000389995.4	SOD1
chr21	33031473	33032186	.	608	ENST00000470944.1	SOD1

Annotating your VCF file

1. Make a VCF custom track
2. Go to the Variant Annotation Integrator
3. Choose your track
4. Add annotations

Remotely Hosted Custom Tracks

- Put data file (bigBed/bigWig/BAM/VCF, etc) in internet accessible location
- Must have: 1. track info, 2. bigDataUrl
- VCF example:

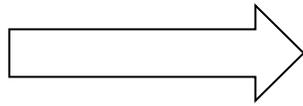
```
track type=vcfTabix
name="VCF_Example"
description="VCF Ex. 1: 1000 Genomes phase 1
interim SNVs"
bigDataUrl=
http://hgwdev.cse.ucsc.edu/~pauline/presentations/
vcfExample.vcf.gz
```

Variant Annotation Integrator

- Upload pgSnp or VCF custom track
- Associate UCSC annotations with your uploaded variant calls
- Add dbSNP info if dbSNP identifier found
- Select custom track and VAI options

Variant Annotation Integrator

Tools



Variant Annotation Integrator

The screenshot displays the UCSC Genome Browser interface. The top navigation bar includes links for Genomes, Genome Browser, Tools, Mirrors, Downloads, My Data, View, Help, and About Us. The main content area shows the Human Feb. 2009 (GRCh37/hg19) Assembly for chromosome 21, with a zoomed-in view of a region around 33,031,596-33,041,000. A red circle highlights the 'Tools' menu, which is open and shows the 'Variant Annotation Integrator' option selected. Other tools listed include Blat, Table Browser, Gene Sorter, Genome Graphs, In-Silico PCR, LiftOver, VisiGene, and Other Utilities. The interface also features a search bar for gene symbols or search terms, a zoom control (1.5x, 3x, 10x, base, zoom out, 1.5x, 3x, 10x, 100x), and various tracks for genomic data, including H2A.2 Histone Mods by ChIP-seq Peaks and Signal from ENCODE/Broad.

Variant Annotation Integrator

genome.ucsc.edu/cgi-bin/hgVai



Genomes

Genome Browser

Tools

Mirrors

Downloads

My Data

Help

About Us

Variant Annotation Integrator

Select Genome Assembly and Region

clade genome assembly

region to annotate

Select Variants

If you have more than one custom track or hub track in [pgSnp](#) or [VCF](#) format, please select the one you wish to annotate:

variants:

maximum number of variants to be processed:

[manage custom tracks](#)

[track hubs](#)

To reset **all** user cart settings (including custom tracks), [click here](#).

Select Genes

The gene predictions selected here will be used to determine the effect of each variant on genes, for example intronic, missense, splice site, intergenic, etc.

Select Regulatory Annotations

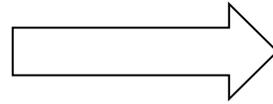
The annotations in this section provide predicted regulatory regions based on various experimental data. When a variant overlaps an annotation, a consequence term [regulatory region variant](#) will be assigned. Follow the links to description pages that explain how each dataset was constructed. A significant portion of the genome and it may be desirable to filter these annotations by cell type and/or score in order to avoid an overabundance of annotations.

Track Data Hubs

- Remotely hosted
- Data persistence
- File formats:
 - bigBED, bigWig, BAM, VCF
- Track organization:
 - groups, supertracks
- multiWigs
- Assembly hubs

Track Hubs

My Data



Track Hubs

Human (*Homo sapiens*) Genome Browser Gateway

The UCSC Genome Browser was created by the [Genome Bioinformatics Group of UC Santa Cruz](#).
Software Copyright (c) The Regents of the University of California. All rights reserved.

group: Mammal | genome: Human | assembly: Feb. 2009 (GRCh37/hg19) | position: chr17:7,583,734-7,595,119 | search term: enter position, gene symbol or search terms | submit

[Click here to reset](#) the browser user interface settings to their defaults. [More on-site workshops available!](#)

track search | add custom tracks | track hubs | configure tracks and display

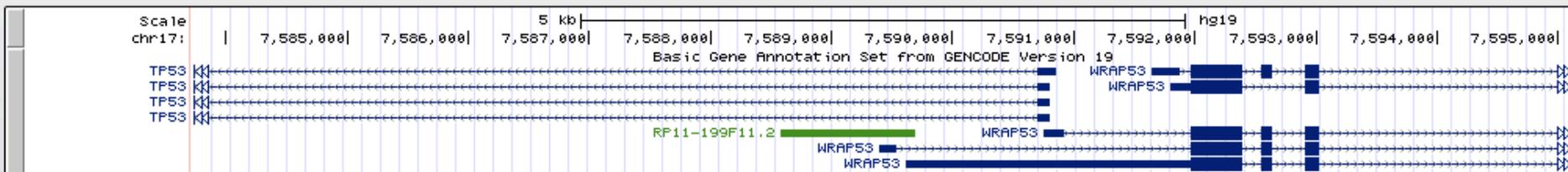
Genomes | Genome Browser | Tools | Mirrors | Downloads | My Data | View | Help | About Us

UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly

- Sessions
- Track Hubs
- Custom Tracks

chr17:7,583,734-7,595,119 11,386 bp. enter position, gene symbol or search terms

chr17 (p13.1) | 13.3 | 13.2 | 13.1 | 17p12 | 17p11.2 | 17q11.2 | 17q12 | 21.31 | 17q22 | 23.2 | 24.2 | 24.3 | 25.1 | 17q25.3



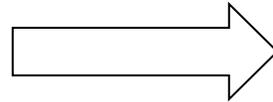
move start | < 2.0 > | Click on a feature for details. Click or drag in the base position track to zoom in. Click side bars for track options. Drag side bars or labels up or down to reorder tracks. Drag tracks left or right to new position. | move end | < 2.0 >

track search | default tracks | default order | hide all | add custom tracks | track hubs | configure | reverse | resize | refresh

Track Hubs

genome.ucsc.edu/cgi-bin/hgHubConnect

My Data



Track Hubs



Genomes

Genome Browser

Tools

Mirrors

Downloads

My Data

Help

About Us

Track Data Hubs

Track data hubs are collections of external tracks that can be imported into the UCSC Genome Browser. Hub tracks show up under the label bar on the main browser page, as well as on the configure page. For more information, see the [User's Guide](#). To import a public hub "Connect" button below.

NOTE: Because Track Hubs are created and maintained by external sources, UCSC is not responsible for their content.

Public Hubs

My Hubs

Enter search terms to find in public track hub description pages:

Search Public Hubs

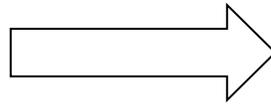
Clicking Connect redirects to the gateway page of the selected hub's default assembly.

Display	Hub Name	Description	Assemblies
<input type="button" value="Connect"/>	Roadmap Epigenomics Data Complete Collection at Wash U VizHub	Roadmap Epigenomics Human Epigenome Atlas Data Complete Collection, VizHub at Washington University in St. Louis	hg19
		UMassMed H3K4me3 ChIP-seq data for	

My Hubs

genome.ucsc.edu/cgi-bin/hgHubConnect

My Data



Track Hubs



Genomes

Genome Browser

Tools

Mirrors

Downloads

My Data

Help

About Us

Track Data Hubs

Track data hubs are collections of external tracks that can be imported into the UCSC Genome Browser. Hub tracks show up under the hub's own blue label bar on the main browser page, as well as on the configure page. For more information, see the [User's Guide](#). To import a public hub click its "Connect" button below.

NOTE: Because Track Hubs are created and maintained by external sources, UCSC is not responsible for their content.

Public Hubs

My Hubs

URL:

Add Hub

No Unlisted Track Hubs

Contact genome@soe.ucsc.edu to add a public hub.

Make Your Own Track Hub

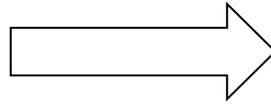
You will need:

- Data (compressed binary index formats: bigBed, bigWig, BAM, VCF)
- Text files to define properties of the track hub
- Internet-enabled web/ftp server
- *Assembly Hubs:*
a twoBit sequence file

Track Hubs

genome.ucsc.edu/cgi-bin/hgHubConnect

My Data



Track Hubs



myHub/ - directory containing track hub files



hub.txt - a short description of hub properties



genomes.txt - list of genome assemblies included



hg19/ - directory of data for the hg19 human assembly



Data files! BAM, bigBed, bigWig, VCF

An Example Assembly Hub

An Arabidopsis hub:

*[http://genome-test.cse.ucsc.edu/
~pauline/hubs/Plants/hub.txt](http://genome-test.cse.ucsc.edu/~pauline/hubs/Plants/hub.txt)*



Acknowledgements

UCSC Genome Browser team

- David Haussler – co-PI
- Jim Kent – Browser Concept, BLAT, Team Leader, PI
- Bob Kuhn – Associate Director, Outreach – co-PI
- Donna Karolchik, Ann Zweig – Project Management

Engineering

Angie Hinrichs

Kate Rosenbloom

Hiram Clawson

Galt Barber

Brian Raney

Max Haeussler

QA, Docs, Support

Katrina Learned

Pauline Fujita

Luvina Guruvadoo

Steve Heitner

Brian Lee

Jonathan Caspar

Matt Speir

Sys-admins

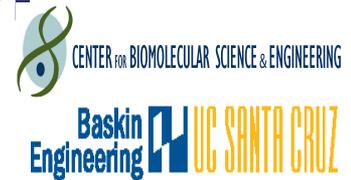
Jorge Garcia

Erich Weiler

Gary Moro



THE GB TEAM



UNIVERSITY OF CALIFORNIA
SANTA CRUZ

THE ORIGINAL
AUTHORITY ON
QUESTIONING
AUTHORITY

Funding Sources

National Human Genome Research Institute (NHGRI)

National Cancer Institute (NCI)

National Institute for Dental and Cranio-Facial Research (NIDCR)

National Institute for Child Health and Human Development (NICHD)

QB3 (UCBerkeley, UCSF, UCSC)

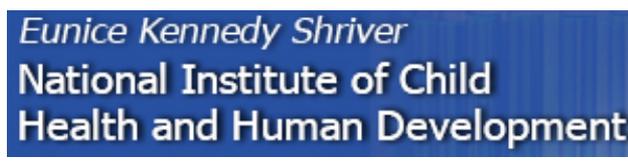
American Recovery and Reinvestment Act (ARRA) stimulus funds



UNIVERSITY OF CALIFORNIA

SANTA CRUZ

Genomics
Institute



genome.ucsc.edu

THANK YOU!

Exercises

1. Load example BED and VCF tracks via url
2. Look at custom track data by pasting url into a web browser.
3. Annotate the TFBS custom track using the Data Integrator.
4. Annotate the VCF custom track using the Variant Annotation Integrator.

Exercise 1

Load example BED and VCF tracks via url

1. Go to the Custom tracks menu
 - My Data -> Custom Tracks
2. Input this url: <http://bit.ly/customtracks> (note that you must include the "http" part of this url or you will get an error) and click [submit].
3. Click the [Go to genome browser] button.
4. Once in the main Browser, jump to this position:
 - chr21:33,034,804-33,037,719
5. See if you can drag your 2 custom tracks to the top of the display

Exercise 2

Exploring your BED and VCF tracks

1. Now that you have 2 custom tracks loaded, take a look at the data by pasting that same url into a web browser:
2. These custom tracks are actually data copied from some existing tracks, see if you can find them, turn them on, and observe that the original tracks and custom tracks look the same in the browser:
 - Track 1 (BED format): Group (Regulation), Super Track (ENC TF Binding), Track (SYDH TFBS)
 - Track 2 (VCF format): Group (Variation), Track (1000G Ph1 Vars)
3. Navigate to this position for best comparison (esp. for the VCF track): chr21:33,034,804-33,037,719

Exercise 3

Annotate your BED with the Data Integrator

1. Go to the Data Integrator
2. Once there select:
 1. Region to annotate: chr21:33031597-33041570
 2. Add data source: group (custom tracks), track (SYDH...) [click add]
3. Now choose which annotations you want to add by [add]ing more tracks to the list – ex:
 1. *Find the genes that overlap with your regions:* group (Genes and Gene Prediction), track (GENCODE V19), view (Genes), subtrack (Basic) [add]
 2. Find the SNPs that overlap with your regions: group (Variation), track (Common SNPs) [add]

Choose which fields to include in your output:

Output options -> Choose fields [Done] -> [get output]

Exercise 4

Annotate your VCF with the Variant Annotation Integrator

1. Go to the Variant Annotation Integrator
 - Tools -> V.A.I.
2. Select Variants:
 - Variants: “VCF Ex. 1...”
3. Now choose which annotations you want to add:
 - To determine which gene regions your variants fall into, select a gene track (Select Genes = “Basic Gene Annotation Set... GENCODE”)
 - Add regulatory annotations: Under “Select Regulatory Annotations” click the “+” button to choose which TFs to include (or select none to include all binding sites)

Bonus Material!

Where to search genome.ucsc.edu/cgi-bin/hgGateway

Genomes Genome Browser Tools Mirrors Downloads My Data Help About Us

(Homo sapiens) Genome Browser Gateway

The UCSC Genome Browser was created by the [Genome Bioinformatics Group of UC Santa Cruz](#).
Software Copyright (c) The Regents of the University of California. All rights reserved.

group	genome	assembly	position	search term	
normal	Human	Feb. 2009 (GRCh37/hg19)	chr21:33,031,597-33,041,570	brc	submit

[Click here to reset](#) the browser user interface settings to their defaults. **[More on-site workshops available!](#)**

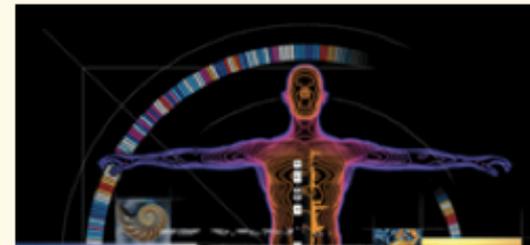
[track search](#) [add custom tracks](#) [track hubs](#) [configure tracks and display](#)

Genome Browser – hg19 assembly ([sequences](#))

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.

Position queries

A position can be specified by the accession number of a sequenced genomic clone or



Where to search genome.ucsc.edu/cgi-bin/hgGateway

Genomes Genome Browser Tools Mirrors Downloads My Data Help About Us

(Homo sapiens) Genome Browser Gateway

The UCSC Genome Browser was created by the [Genome Bioinformatics Group of UC Santa Cruz](#).
Software Copyright (c) The Regents of the University of California. All rights reserved.

group	genome	assembly	position	search term
mammal	Human	Feb. 2009 (GRCh37/hg19)	chr21:33,031,597-33,041,570	brca

[Click here to reset](#) the browser user interface settings to their defaults. [More](#)

BRCA1 (Homo sapiens breast cancer 1, early onset (BRCA1), transcript variant 2, mRNA.)
BRCA2 (Homo sapiens breast cancer 2, early onset (BRCA2), mRNA.)
BRCC3 (Homo sapiens BRCA1/BRCA2-containing complex, subunit 3 (BRCC3), transcript variant 1, mRNA.)

Genome Browser – hg19 assembly ([sequences](#))

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.

Genome position queries



Where to search: Main Browser genome.ucsc.edu/cgi-bin/hgTracks

The screenshot displays the UCSC Genome Browser interface. At the top, a navigation bar includes links for Genomes, Genome Browser, Tools, Mirrors, Downloads, My Data, View, Help, and About Us. Below this, the main title reads "UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly". A search bar contains the text "smad" and is highlighted with a red box. To the right of the search bar is a "go" button and a link to "More on-site workshop". Below the search bar, a list of gene transcripts is displayed, including SMAD1 through SMAD9, with their full names and transcript variants. The interface also shows a genomic track for chromosome 17 (q21.31) with a scale of 41,210,000 bp. The track displays various data tracks, including BRCA1 tracks and ChIP-seq data for GM12878 cells targeting BRCA1.

Genomes Genome Browser Tools Mirrors Downloads My Data View Help About Us

UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly

move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x 100x

1,196,312-41,277,500 81,189 bp. go [More on-site workshop](#)

chr17 (q21.31) 13

Scale chr17: 41,210,000

BRCA1 BRCA1 BRCA1 BRCA1 BRCA1 BRCA1 BRCA1 BRCA1 BRCA1 BRCA1

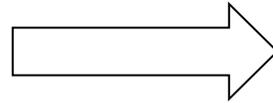
SMAD1 (Homo sapiens SMAD family member 1 (SMAD1), transcript variant 1, mRNA.)
SMAD2 (Homo sapiens SMAD family member 2 (SMAD2), transcript variant 2, mRNA.)
SMAD3 (Homo sapiens SMAD family member 3 (SMAD3), transcript variant 1, mRNA.)
SMAD4 (Homo sapiens SMAD family member 4 (SMAD4), mRNA.)
SMAD5 (Homo sapiens SMAD family member 5 (SMAD5), transcript variant 3, mRNA.)
SMAD5-AS1 (Homo sapiens SMAD5 antisense RNA 1 (SMAD5-AS1), non-coding RNA.)
SMAD6 (Homo sapiens SMAD family member 6 (SMAD6), transcript variant 1, mRNA.)
SMAD7 (Homo sapiens SMAD family member 7 (SMAD7), transcript variant 1, mRNA.)
SMAD9 (Homo sapiens SMAD family member 9 (SMAD9), transcript variant a, mRNA.)

BRCA1 BRCA1 BRCA1 BRCA1 BRCA1 BRCA1

ChIP-seq of GM12878 - ENCSR000DZ3 (Target - BRCA1)

Public Hubs

My Data



Track Hubs

Human (*Homo sapiens*) Genome Browser Gateway

The UCSC Genome Browser was created by the [Genome Bioinformatics Group of UC Santa Cruz](#).
Software Copyright (c) The Regents of the University of California. All rights reserved.

group: Mammal | genome: Human | assembly: Feb. 2009 (GRCh37/hg19) | position: chr17:7,583,734-7,595,119 | search term: enter position, gene symbol or search terms | submit

[Click here to reset](#) the browser user interface settings to their defaults. [More on-site workshops available!](#)

track search | add custom tracks | track hubs | configure tracks and display

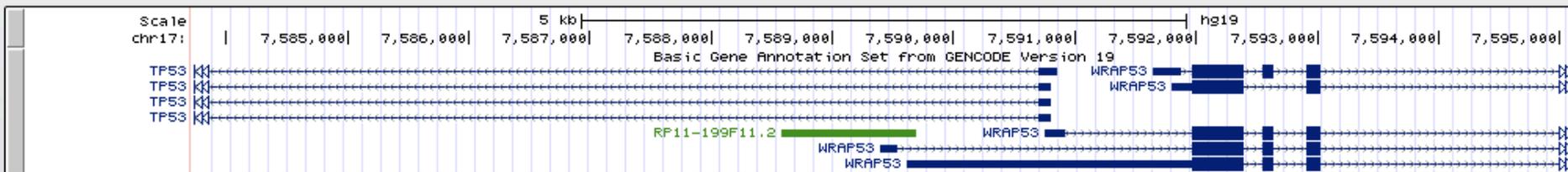
Genomes | Genome Browser | Tools | Mirrors | Downloads | My Data | View | Help | About Us

UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly

- Sessions
- Track Hubs
- Custom Tracks

chr17:7,583,734-7,595,119 11,386 bp. enter position, gene symbol or search terms

chr17 (p13.1) | 13.3 | 13.2 | 13.1 | 17p12 | 17p11.2 | 17q11.2 | 17q12 | 21.31 | 17q22 | 23.2 | 24.2 | 24.3 | 25.1 | 17q25.3



move start | < 2.0 > | Click on a feature for details. Click or drag in the base position track to zoom in. Click side bars for track options. Drag side bars or labels up or down to reorder tracks. Drag tracks left or right to new position. | move end | < 2.0 >

track search | default tracks | default order | hide all | add custom tracks | track hubs | configure | reverse | resize | refresh

Where to search

genome.ucsc.edu/cgi-bin/hgHubConnect

[Genomes](#)[Genome Browser](#)[Tools](#)[Mirrors](#)[Downloads](#)[My Data](#)[Help](#)[About Us](#)

Track Data Hubs

Track data hubs are collections of external tracks that can be imported into the UCSC Genome Browser. Hub tracks show up under their own blue label bar on the main browser page, as well as on the configure page. For more information, see the [User's Guide](#). To import a hub click its "Connect" button below.

NOTE: Because Track Hubs are created and maintained by external sources, UCSC is not responsible for their content.

[Public Hubs](#)[My Hubs](#)

Enter search terms to find in public track hub description pages:

Clicking Connect redirects to the gateway page of the selected hub's default assembly.

Display	Hub Name	Description	Assemblies
<input type="button" value="Connect"/>	Roadmap Epigenomics Data Complete Collection at Wash U VizHub	Roadmap Epigenomics Human Epigenome Atlas Data Complete Collection, VizHub at Washington University in St. Louis	hg19
<input type="button" value="Connect"/>	Cancer genome polyA site &	An in-depth map of polyadenylation sites in cancer fresh and archived	hg19

Track search

(Homo sapiens) Genome Browser Gateway

The UCSC Genome Browser was created by the [Genome Bioinformatics Group of UC Santa Cruz](#).
Software Copyright (c) The Regents of the University of California. All rights reserved.

group genome assembly position search term
normal Human Feb. 2009 (GRCh37/hg19) chr21:33,031,597-33,041,570 brc submit

[Click here to reset](#) the browser user interface settings to their defaults. **More on-site workshops available!**

track search

add custom tracks

track hubs

configure tracks and display

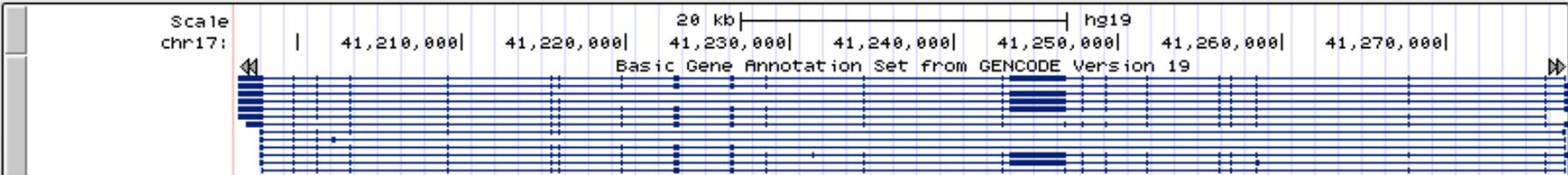
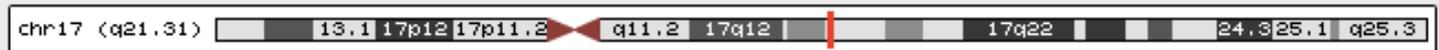
UCSC Genome Browser on Human Feb. 2009

Assembly

move <<< << < > >> >>> zoom in 1.5x 3x 10x 10x 100x

chr17:41,196,312-41,277,500 81,189 bp. enter position, gene symbol or search terms

- My Data
- View
- Help
- Sessions
- Track Hubs
- Custom Tracks



Click on a feature for details. Click or drag in the base position track to zoom in. Click side bars for track options. Drag side bars or labels up or down to reorder tracks. Drag tracks left or right to new position.

move start < 2.0 >

move end < 2.0 >

track search

default tracks

default order

hide all

add custom tracks

track hubs

configure

reverse

resize

refresh

Track search

[Genomes](#)[Genome Browser](#)[Tools](#)[Mirrors](#)[Downloads](#)[My Data](#)[Help](#)[About Us](#)

Search for Tracks in the Human Feb. 2009 (GRCh37/hg19) Assembly

Search

Advanced

search

clear

cancel

About Track Search

Search for terms in track names, descriptions, groups, and ENCODE metadata. If multiple terms are entered, only tracks with all terms will be part of the results.

[more help](#)

Track search

[Genomes](#)[Genome Browser](#)[Tools](#)[Mirrors](#)[Downloads](#)[My Data](#)[Help](#)[About Us](#)

Search for Tracks in the Human Feb. 2009 (GRCh37/hg19) Assembly

[Search](#)[Advanced](#)

Track Name:

contains

and Description:

contains

and Group:

is

and Data Format:

is [ENCODE terms](#)

+

and

is among

[Cell, tissue or DNA sample](#)

+

and

is among

[Antibody or target protein](#)

+ -

Visibility

Track Name

hide

[GM12878](#)

GM12878 Methylation 450K Bead Array from ENCODE/HAIB

▼

hide

[H1-hESC](#)

H1-hESC Methylation 450K Bead Array from ENCODE/HAIB

▼

(0 of 2 selected)

About Track Search