



Introductory bioinformatics for human genomics

Introduction to the UCSC genome browser

Never Stand Still

Medicine

Prince of Wales Clinical School

Dominik Beck

NHMRC Peter Doherty Fellow

Stem Cell Group, Adult Cancer Program, Lowy Cancer Research Centre

Prince of Wales Clinical School, Faculty of Medicine, UNIVERSITY OF NEW SOUTH WALES, SYDNEY NSW 2052



ADULT CANCER PROGRAM



CONTENT

- Background
- Genome Assemblies
- Annotation Tracks: data set that can be linked to the genome given some coordinates
- Associated Tools
- Practical Exercise



Genome Browser

<http://genome.ucsc.edu/>

UCSC Genome Bioinformatics

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

- Genome Browser
- ENCODE
- Neandertal
- Blat
- Table Browser
- Gene Sorter
- In Silico PCR
- Genome Graphs
- Galaxy
- VisiGene
- Utilities
- Downloads
- Release Log
- Custom Tracks
- Cancer Browser
- Microbial Genomes
- Mirrors
- Training
- Credits
- Publications
- Cite Us
- Licenses
- Jobs
- Staff
- Contact Us

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. We encourage you to explore these sequences with our tools. The [Genome Browser](#) zooms and scrolls over chromosomes, showing the work 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 data 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 Center for Genome Sciences and Policy. If you have feedback or questions concerning the tools or data on this website, feel free to contact us on our [public mailing list](#).

News

To receive announcements of new genome assembly releases, new software features, updates and training seminars by email, subscribe to the [Genome Browser News](#) mailing list.

24 October 2013 - Job Opening: UCSC Genome Browser Trainer

The [Center for Biomolecular Science and Engineering](#) (CBSE) at University of California Santa Cruz seeks an articulate, self-motivated educational professional for in-person training on the UCSC Genome Browser at universities, hospitals, institutes, and professional meetings in the United States and international locations. The position requires a Master's degree in a biological science, depth in molecular biology, experience in a research environment, working knowledge of video production, and a strong interest in teaching or training in a scientific environment. Preferred qualifications include a PhD in a relevant field, experience with video production, and a strong interest in teaching or training in a scientific environment.

For more information and to apply for this position, see [Job #1304619](#) on the UCSC Staff Employment website.

23 October 2013 - dbSNP Build 138 Available for hg19

We are pleased to announce the release of four tracks derived from NCBI [dbSNP](#) Build 138 data, available on the human assembly (GRCh37/hg19). The tracks include corresponding coloring and filtering options in the Genome Browser.

As was the case for the annotations based on the previous dbSNP build 137, there are four tracks in this release. One is a track containing all the SNPs in the build and the other three are subsets of this track and show interesting and easily defined subsets of dbSNP:

- Common SNPs (138): uniquely mapped variants that appear in at least 1% of the population or are 100% non-reference
- Flagged SNPs (138): uniquely mapped variants, excluding Common SNPs, that have been flagged by dbSNP as "clinically associated"
- Mult. SNPs (138): variants that have been mapped to more than one genomic location

By default, only the Common SNPs (138) are visible; other tracks must be made visible using the track controls.

You will find the four SNPs (138) tracks on the Human Feb. 2009 (GRCh37/hg19) browser in the "Variation and Repeats" group.

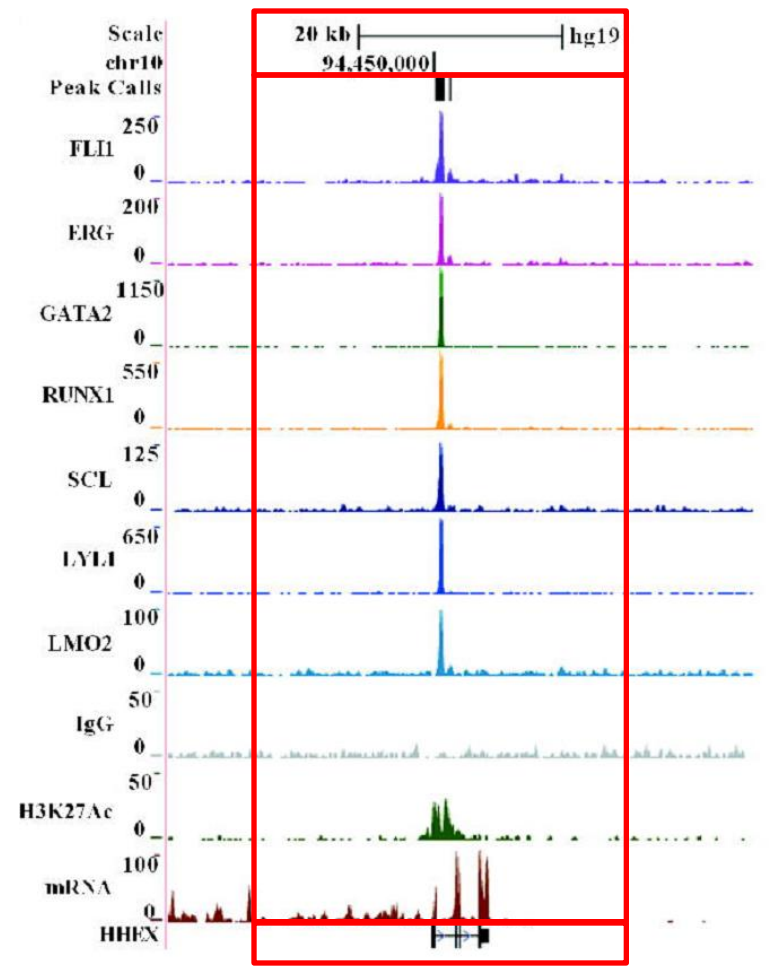
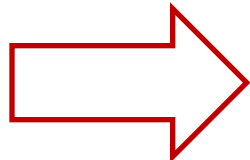
The tracks were produced at UCSC by Angie Hinrichs and Luvina Guruvadoo. We'd like to thank the dbSNP group at NCBI for providing access to the data.



Background

← → ↻ 🔍 Web genome.ucsc.edu/cgi-bin/hgc

```
TTTAAACATTTTTTTTGTDRDRHSRTAATGTGTGTCTTCTCAGTAAATGRTTTTAAARTTA  
CAAGTTAGAAAAGTTTATCTAGTTGTTCTATATAGCATTAATCTGGGCTGCTTTGTCAA  
AAATTTATTTATTTTGAATCAGGGTGTCACTCTGTACCACAGGCTAGTGGCGTATCA  
CAGCTCACTGCAGCCTTGAACCTCTGGGACGTGCAAGTGTGCACACCACTCCGGCCTTT  
TGTATTTATTTATTTTTTGTAGAGACTGGGTACTCTCATGTGCCACAGGCTGGTCCCG  
GCTCAAGCGATCCACCAGCCTCTGTCTTCAAAGTGTGGGATACAGGTTGAGTCAAC  
AAGTTCCTTTAAACCAAGGTTAAAGATAAAAGAAAGGACAGAAATTTGGTGTATTTTT  
GAGTGTCCAAAGTAAATTTTATCTCTAATTTGTCACTCACAGCTTATTTTGAATGAA  
TTAATGAGAAGCTAAATAAACTAGCTTTTAAAAAATATTTCTCTCTCTCCAGGCA  
ATGGGTCACTTTGGGGGCTGCAGAACATCTGTGCTTTCCCTGTGCCCATCACTCTCC  
AGCTCTGAGCTGTAGCCATCTCTGCTGCACCTGTCTGAGAGGCGAGCTCTCCGGCTC  
CTGCTCTGGGTTGCAGTGCACATCTGTGCAGAAAGCTTAGGAGAAGCAGTGAAGCTT  
CTAACGACAGCTCTGCACAAAGCCTTTGTGAACCTCCGGGACATAAGGGCATCAA  
ACTAGGAACAAAACACAGCTCTTGAAGTGTGTTGGTAAATTTACACATACAAGGCGCG  
AAAGCCAAATGAAGCAAGAAAAGAACAAATCTCTTGCAGCAGGATGAAGTTGTAC  
AGGAGCTGTAGTATACATCAGGGCAAGCAGAGACAGGGCAACAGAGAGAAAAGTT  
CATGCAATTTAATATGACCTGGAGGGGCTTTGGAAATGAGTTACACAGTAAACAGG  
AGTTAATGCAATTTGAACAATTTAAGAATATGATTTGAAGGAACATCAGTATATCA  
TGTATCCAAAATTTCCAGTTGCTCATAGACTGTCTAATAACCTGGACTGGGGTGG  
AGGGTTGAGGAGCTAATGGAAAACATCTGTTAAAATGAATATGATAATCATGCTTGT  
TTTTATGATCTGTCTCATGAGACAGCTTGAAGAGAGACTGAGTAAATTCAGATAAA  
TGTGATTTTCTAATGAAGAATGAGGTGTCTTGAAGTTCGAAATTTAAAATAAAATGAC  
TGAATTTTTCCATTAGCACAGTTTGGGTGAGCTAACTGCAGCTTTCATGATGTATTT  
ATTTATTTCCCTAAGGAGATACGGGCACTTTGTGGGCTTCCAGGCTGTGCAAAAATG  
ATGGATGGTGGGCTCGTCTAATCAAAATGACCGGGGTTCAGGAGCATTTGGTAATCG  
TGCTCGCGACTTCAAAAGCCACTGCCAAAACATCCACAGCAGAGGTTTGGCAACT  
CACATAATCATAGCAAAAGGACTGCAAGCTGAAAACAGGCTGAAAACAGCTATGAA  
AAGGGCCAGTTCAGAAACCTATTCAAGTAAAGTGTTCGGTAGGCTTCCGACGGTCA  
CCTTGTATTTTCACTTTTGAAGTGTCTTCACTGGACCAAGGTTGGCAGCATTTGTAT  
TCAAAGAAAAGATGCCAGGGAACTGACTAGATATGGAATAAATGATTTGCTTCAA  
GGATAGCCAGCAACATCAGGCTCAGGGCTAGTGAATCCCAAGCACAGTCCCAAGTAA  
CTCTGATGTAGCAGGACTAAAGCTGTCTACTAGTGAAGCTCTCAGAAAAGAAAACCCG  
CAGCAAACTAGACGTTATCCCTTGTCTTGTAAAGTGAAGAAATGACGCTATCCATGAC  
GCTTTATTTGCAAGTAAATAGAGTTGCTTACAGGTTGCTTACAGTGTGCTTCACTGT  
CAAAATGATTTAATCATAGCAGGAAATTAAGGTTGATTTGACAGGTTGCTGACATGA  
ATTTCTCAATGACAAACCCAGCTTCAAGGCTCTTCCATCCAGGCAATAATGGAG  
GCTCAAATGTAAACCCGGGCTCTTGTCTCAAAGGGGCTAGAATAAAAAGCAGGA  
GGGAGGCAAGAAAGCCTGGAGAGGCTGACGCAATTTGGGTGCAAAACATCTATTTC  
TTGGCTCTCCCTTGCACAAAGTTCTGGACAAAAGTAAATATAAACAAAATCCACAACT  
TCAGCACATGTTTCAATTAAGCAACTTTAGTCACTAAAAGAAAGTGAAGTGAAGCTCT  
GTATAAATCTGATTTGCTGATGCTAGGCAAGCTTATTTTATACATACATTTGCTTCT  
AAGAACATAAATCTCATATGTAAACATTAAGCATACAGGTTAAATTTCAAGGCCACA  
TTATATCGATTTGCTCTTTGCTGTTGCTTTGGCTGGCCGAGATCAACTGTAAGTGA  
TAAATGCATAAGTTATATAATATATAAAAAGGGGAAAACATTTGACTGTATACT  
TCATCTGACAAACAGCAGGCTTGCAGCTCAAGGAAAACCTGGAGCCGCTACCCAA  
AATGCTGCGTGATTTCTGATTTGGCAGCCAAGAGGCTCATCTTACCTGACCTGTG  
GAGAACAAGGCCCAACATAATGAGGCTCTGAATGTTTCTCTAAATACAGCAAT  
TCCAGTTAAAATTTTCAATTTGACAAAACAAAAGAAAGATGCGCATTTTGTCTGAAAT  
CTACTACTTCCCTTTCTCATTAGCCTGTGCTTCTCTTCAACTGGGTTGGTATA  
ACACTGACTGATGAACCTCGAGTCTCATAACTCAITGTACACAGTCCCAAAATGCTC  
TGAGTCCAGCTTCTATTGCTTGAAGGTTCTTCCAGGCTTGGCTCTCCCTCTCTG  
CCTCTTCTCTCTCTCAAGCCTCTGTCTCAATCACACGCTCACTCTATACACACT  
CTCAGTCCCACTTTTGTGTTTATGTTCCGGTAACTGTAAAGGAGTTGAAAATTTGG  
GTCACTTCAAGTGTGAAGAGCTGACAGCTGTCAATCAACGGAATGATTTGATTTCC  
AACAAAACAGCACATGCCATGAGTTGCATATCAAGCTGTGTTGATGGGCCACAGCTCT  
TCTGCTGCTTTTCTCTGTTTGTATGATGTTCTATTTTAAAATCAGGTTGTTTCT  
TAAAATGGCATATAAAGCTGATGTTAGGTTCAAACCTGTTGTTTCTATTTGCTTGT  
TCACTCCAGTTGCAATGAGTGGGATCCAAACTCACTCAAAAGTTATACATTTCTTAA  
GACCACTTTCTTGGCACTTTGCTTAAAGCTTCAAGCTTCACTACATACAGTCTCTTTA  
CAGGTCAGCTCAAGGATGATGTTCCATTTGTCATTTGTCATTTTAACTCTCTCATTT  
CTGTAGTAAGACTTCAGTAACTCCCTCAAGAGTCTTTGGATCTTCCCGGCTCTCT  
TCCACAGCCAGTAAGACTTTTTCATTCTCTTGAAGGCTTTTGTTCATGTTCTCC  
GATAGATTTGTGGCAGTGGCTGGTGAATGACAGCTGATGGGAAAAGCTCCCGCAGT  
CTTTAGTAAAGTGCACAGATGAGAAGGCATAGGCTGGTGGGAGGCTAGTGTGGGG  
TATATACCCCAAGTTGGTGAATCCAGTATGGGTTGGGCGCAAAAACATGGAAGAT  
GTACAGGGGAGGCTGGAGGGTGGGGGCTCAAAAAGTTCATCTCTGGGTTGGGCTGA  
TAGGAGCCATGTACGGGAGGCTGAGGGGACTTGTACAGAGATGACTCCGGGGGGTGG  
GGCTGGAGGGCTGGGCGATCCGTGGAAAGTCAAGCTTGTAGGCGTAGGCTTCCATGG  
ACCTTGGCTATGATGTTCTTGTATAGTAAAGGAGGGGCGCGGTGAGCTTATGTAG
```



Background

Visualization of genomic data

- Graphical viewpoint on the very large amount of genomic sequence produced by the Human Genome Project.

Human Genome: 3,156,105,057 bp

- Focus turned from accumulating and assembling sequences to identifying and mapping functional landmarks

Genetic markers

Genes

SNPs

Points of regulation

- Visualization of Next-generation-sequencing data



Background

Client-side

Integrative Genomics Viewer*

- ❑ Application (Java) on the user's machine
- ❑ Often difficult to install
- ❑ Does not have the extensive third-party data of the other browsers
- ❑ Much faster than web-based browsers



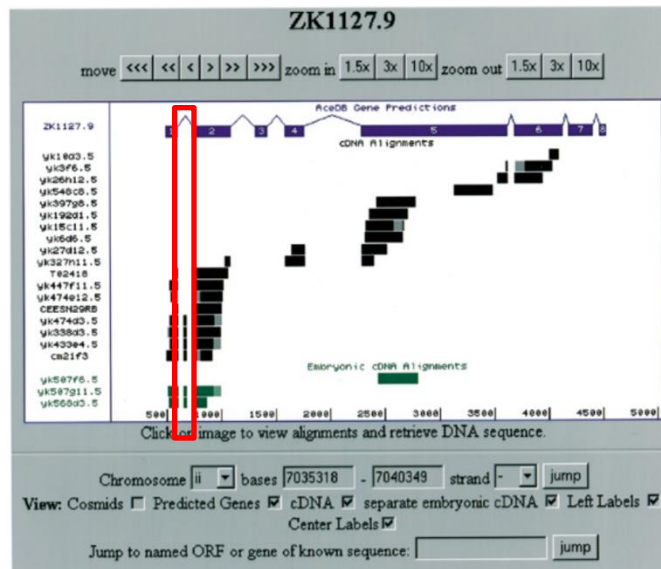
The screenshot shows the homepage of the Integrative Genomics Viewer (IGV) website. The browser address bar displays "www.broadinstitute.org/igv/". The page features a navigation menu on the left with links for Home, Downloads, Documents, Hosted Genomes, FAQ, IGV User Guide, File Formats, Release Notes, Credits, and Contact. Below the menu is a search box and the Broad Institute logo. The main content area includes a large banner with the IGV logo and a visualization of genomic data. Below the banner are sections for "What's New" (announcing a workshop in October 2013 and a new release in April 2013), "Citing IGV" (providing citation information for publications), "Overview" (describing IGV as a high-performance visualization tool), "Downloads" (with a registration link), and "Funding" (listing funding sources like the National Cancer Institute and the National Human Genome Research Institute). The footer contains logos for the National Cancer Institute, National Human Genome Research Institute, and GENOME SPACE.

Background

- ❑ Intronerator was developed by J. Kent to map the exon–intron structure of *C. elegans* RNAs mapped against genomic coordinates

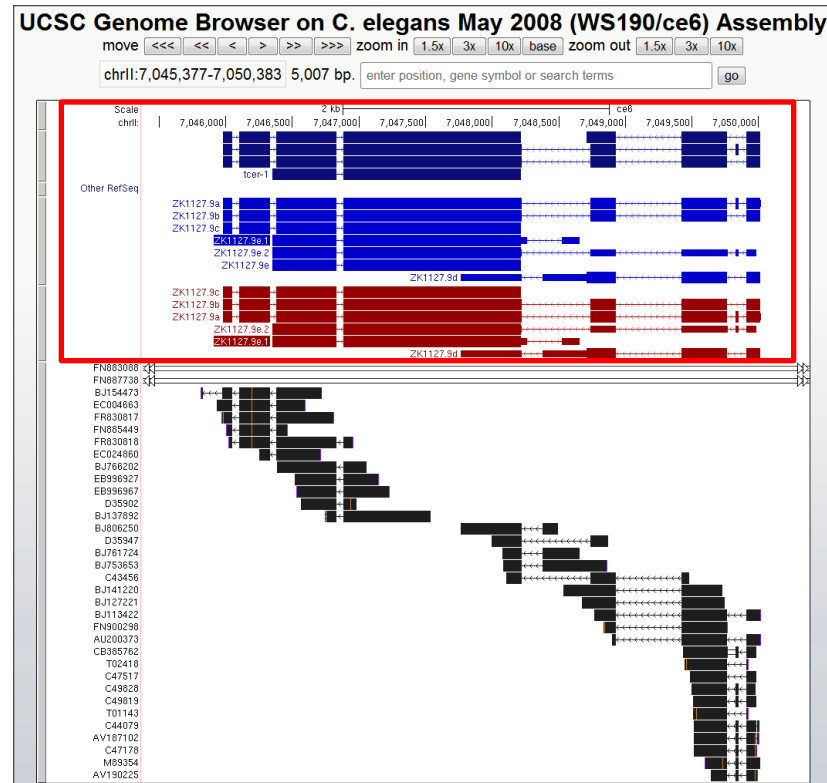


Jim Kent



Background

- ❑ Draft human genome sequence became available at the UCSC in 2000
- ❑ Intronerator was used as the graphics engine



CONTENT

- Background
- Genome Assemblies
- Annotation Tracks
- Associated Tools
- Practical Exercise



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	genome	assembly	position	search term	
Mammal	Human	Feb. 2009 (GRCh37/hg19)	chr10:123,227,429-123,343,066	FGFR2	<input type="button" value="submit"/>

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

FGFR2 (Homo sapiens fibroblast growth factor receptor 2 (FGFR2), transcript variant 2, mRNA.)

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

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.

Sample position queries

A genome position can be specified by the accession number of a sequenced genomic clone, an mRNA or EST or STS marker, a chromosomal coordinate range, or keywords from the GenBank description of an mRNA. The following list shows examples of valid position queries for the human genome. See the [User's Guide](#) for more information.

Request:

Genome Browser Response:

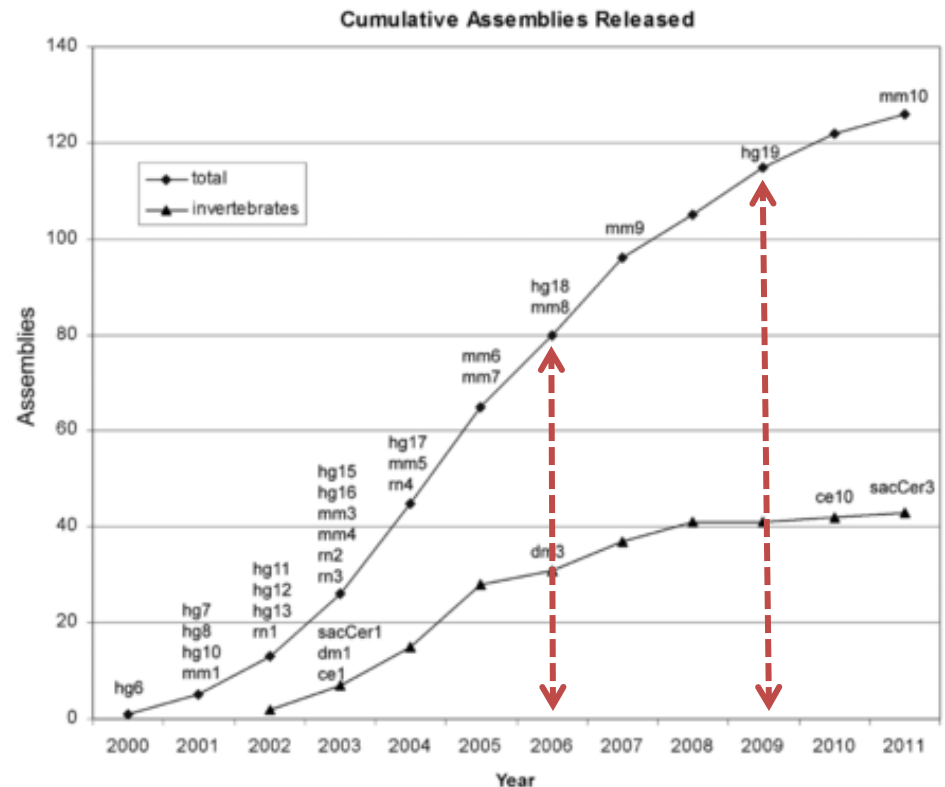
chr7	Displays all of chromosome 7
chrUn_gi000212	Displays all of the unplaced contig gi000212
20p13	Displays region for band p13 on chr 20
chr3:1-1000000	Displays first million bases of chr 3, counting from p-arm telomere
chr3:1000000+2000	Displays a region of chr3 that spans 2000 bases, starting with 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	Displays region around STS marker D16S3046 from the Genethon/Marshfield maps. Includes 100,000 bases on each side as well.
AA205474	Displays region of EST with GenBank accession AA205474 in BRCA1 cancer gene on chr 17
AC008101	Displays region of clone with GenBank accession AC008101
AF083811	Displays region of mRNA with GenBank accession number AF083811
PRNP	Displays region of genome with HUGO Gene Nomenclature Committee identifier PRNP
NM_017414	Displays the region of genome with RefSeq identifier NM_017414
NP_059110	Displays the region of genome with protein accession number NP_059110
pseudogene mRNA	Lists transcribed pseudogenes, but not cDNAs
homeobox caudal	Lists mRNAs for caudal homeobox genes
zinc finger	Lists many zinc finger mRNAs
kruppel zinc finger	Lists only kruppel-like zinc fingers
huntington	Lists candidate genes associated with Huntington's disease
zahler	Lists mRNAs deposited by scientist named Zahler
Evans,J.E.	Lists mRNAs deposited by co-author J.E. Evans



Homo sapiens
(Graphic courtesy of [CBSE](#))

Genome Assemblies

- ❑ Regular updates to genome assemblies to close gaps in genomic sequence, troubleshoot assembly problems and otherwise improve the genome assemblies
- ❑ Shifting coordinates for known sequences and a potential for confusion and error among researchers, particularly when reading literature based on older versions.
- ❑ Frequently used assemblies hg18/hg19
- ❑ New assemblies increase genomic coverage 6-fold and have been deposited in GenBank.
- ❑ 127 genome assemblies have been released on 58 organisms (April 2012)



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	genome	assembly	position	search term
Mammal	Human	Feb. 2009 (GRCh37/hg19)	chr10:123,227,429-123,343,066	FGFR2 <input type="submit" value="submit"/>

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

FGFR2 (Homo sapiens fibroblast growth factor receptor 2 (FGFR2), transcript variant 2, mRNA)

Human Genome Browser – hg19 assembly (sequences)

The February 2009 human reference sequence (GRCh37) was produced by the [GRCh37](#) in the NCBI Assembly database.

Sample position queries

A genome position can be specified by the accession number of a sequenced or keywords from the GenBank description of an mRNA. The following list shows for more information.

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
chr3:1-1000000	Displays first million bases of chr 3, counting from p-arm
chr3:1000000+2000	Displays a region of chr3 that spans 2000 bases, starting at 1000000
RH18061;RH80175	Displays region between genome landmarks, such as BAC clones or SNPs rs1042522 and rs1800370. This syntax may also be used for mRNAs, refSeqs, etc.
15q11;15q13	
rs1042522;rs1800370	
D16S3046	Displays region around STS marker D16S3046 from the genome
AA205474	Displays region of EST with GenBank accession AA205474
AC008101	Displays region of clone with GenBank accession AC008101
AF083811	Displays region of mRNA with GenBank accession number AF083811
PRNP	Displays region of genome with HUGO Gene Nomenclature Committee symbol PRNP
NM_017414	Displays the region of genome with RefSeq identifier NM_017414
NP_059110	Displays the region of genome with protein accession number NP_059110
pseudogene mRNA	Lists transcribed pseudogenes, but not cDNAs
homeobox caudal	Lists mRNAs for caudal homeobox genes
zinc finger	Lists many zinc finger mRNAs
kruppel zinc finger	Lists only kruppel-like zinc fingers
huntington	Lists candidate genes associated with Huntington's disease
zahler	Lists mRNAs deposited by scientist named Zahler
Evans,J.E.	Lists mRNAs deposited by co-author J.E. Evans

RefSeq Genes

[FGFR2 at chr10:123237844-123353481](#) - (NM_001144914) fibroblast growth factor receptor 2 isoform 4 precursor
[FGFR2 at chr10:123237844-123356159](#) - (NM_001144915) fibroblast growth factor receptor 2 isoform 5 precursor
[FGFR2 at chr10:123241367-123357972](#) - (NM_001144919) fibroblast growth factor receptor 2 isoform 9 precursor
[FGFR2 at chr10:123237844-123357972](#) - (NM_001144917) fibroblast growth factor receptor 2 isoform 7 precursor
[FGFR2 at chr10:123237844-123357972](#) - (NM_022970) fibroblast growth factor receptor 2 isoform 2 precursor
[FGFR2 at chr10:123237844-123353772](#) - (NM_001144916) fibroblast growth factor receptor 2 isoform 6 precursor
[FGFR2 at chr10:123237844-123357972](#) - (NM_001144918) fibroblast growth factor receptor 2 isoform 8 precursor
[FGFR2 at chr10:123237844-123357972](#) - (NM_000141) fibroblast growth factor receptor 2 isoform 1 precursor
[FGFR2 at chr10:123241367-123353481](#) - (NM_001144913) fibroblast growth factor receptor 2 isoform 3 precursor
[FGFR2 at chr10:123237844-123353481](#) - (NM_023029) fibroblast growth factor receptor 2 isoform 11 precursor
[FGFR2 at chr10:123237844-123357972](#) - (NR_073009)

Non-Human RefSeq Genes

[fgfr2 at chr10:123239371-123324098](#) - (NM_178303) fibroblast growth factor receptor 2 isoform 3 precursor
[fgfr2 at chr10:123239371-123325219](#) - (NM_001090663) fibroblast growth factor receptor 2 precursor
[fgfr2 at chr10:123239023-123325219](#) - (NM_001102856) fibroblast growth factor receptor 2 precursor
[FGFR2 at chr10:123239371-123353399](#) - (NM_205319) fibroblast growth factor receptor 2 precursor
[FGFR2 at chr10:123237856-123269807](#) - (NM_001131221) fibroblast growth factor receptor 2
[FGFR2 at chr10:123237856-123353434](#) - (NM_001003336) fibroblast growth factor receptor 2 precursor
[FGFR2 at chr10:123238077-123357741](#) - (NM_001163863) fibroblast growth factor receptor 2 precursor
[FGFR2 at chr10:123239230-123353378](#) - (NM_001099924) fibroblast growth factor receptor 2 precursor
[FGFR2 at chr10:123239559-123353331](#) - (NM_001082688) fibroblast growth factor receptor 2 precursor
[Fgfr2 at chr10:123237846-123358315](#) - (NM_201601) fibroblast growth factor receptor 2 isoform IIIb
[Fgfr2 at chr10:123237846-123358315](#) - (NM_010207) fibroblast growth factor receptor 2 isoform IIIc
[Fgfr2 at chr10:123237873-123357855](#) - (NM_001109893) fibroblast growth factor receptor 2 isoform c
[Fgfr2 at chr10:123237873-123357855](#) - (NM_001109896) fibroblast growth factor receptor 2 isoform f
[Fgfr2 at chr10:123237873-123357855](#) - (NM_012712) fibroblast growth factor receptor 2 isoform a
[Fgfr2 at chr10:123237873-123357855](#) - (NM_001109894) fibroblast growth factor receptor 2 isoform d
[Fgfr2 at chr10:123237873-123357855](#) - (NM_001109892) fibroblast growth factor receptor 2 isoform b
[Fgfr2 at chr10:123237873-123357855](#) - (NM_001109895) fibroblast growth factor receptor 2 isoform e
[FGFR2 at chr10:123237846-123357550](#) - (NM_001205310) fibroblast growth factor receptor 2
[fgfr2 at chr10:123239371-123324098](#) - (NM_001243004) fibroblast growth factor receptor 2 isoform 1 precursor
[fgfr2 at chr10:123239371-123324098](#) - (NM_001243005) fibroblast growth factor receptor 2 isoform 2 precursor
[fgfr2 at chr10:123239371-123324098](#) - (NM_001243006) fibroblast growth factor receptor 2 isoform 4 precursor

Basic Gene Annotation Set from ENCODE/GENCODE Version 17

[FGFR2 at chr10:123237848-123353481](#)
[FGFR2 at chr10:123237848-123356159](#)
[FGFR2 at chr10:123237855-123357598](#)
[FGFR2 at chr10:123237878-123290828](#)
[FGFR2 at chr10:123238586-123357972](#)
[FGFR2 at chr10:123238732-123357812](#)
[FGFR2 at chr10:123239133-123357966](#)



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

move



zoom in

1.5x

3x

10x

base

zoom out

1.5x

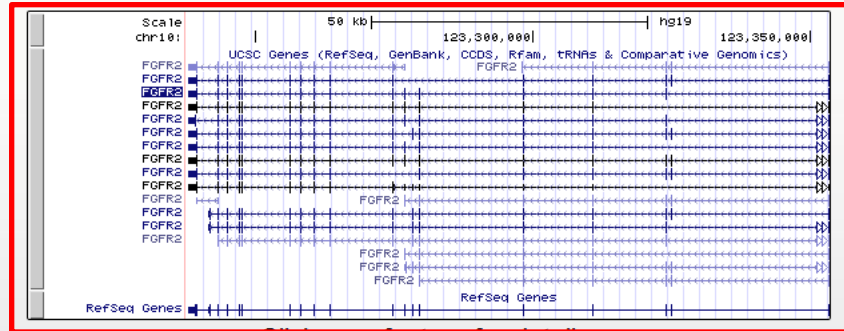
3x

10x

chr10:123,237,844-123,353,481 115,638 bp.

enter position, gene symbol or search terms

go

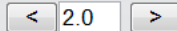


Click on a feature for details.

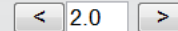
Click or drag in the base position track to zoom in.

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



move end



track search

default tracks

default order

hide all

add custom tracks

track hubs

configure

reverse

resize

refresh

collapse all

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.

expand all

- Mapping and Sequencing Tracks refresh
- Phenotype and Disease Associations refresh
- Genes and Gene Prediction Tracks refresh
- Literature refresh
- mRNA and EST Tracks refresh
- Expression refresh

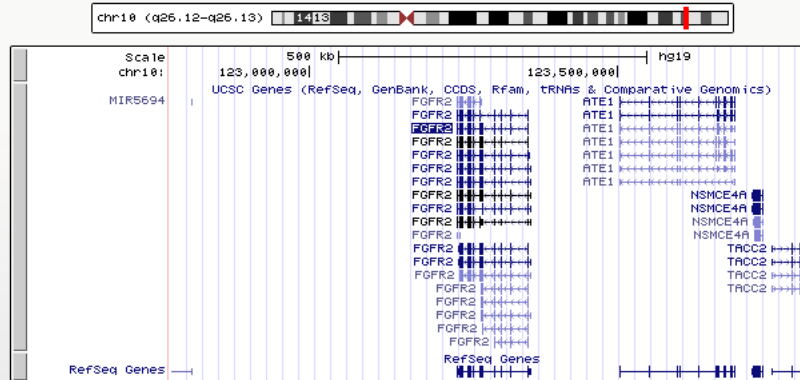




UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly

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

chr10:122,775,292-123,816,033 1,040,742 bp.



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

collapse all

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.

expand all

- Mapping and Sequencing Tracks
- Phenotype and Disease Associations
- Genes and Gene Prediction Tracks
- Literature
- mRNA and EST Tracks
- Expression





Configure Image

image width: pixelslabel area width: characterstext size:

<input checked="" type="checkbox"/>	Display chromosome ideogram above main graphic
<input checked="" type="checkbox"/>	Show light blue vertical guidelines
<input checked="" type="checkbox"/>	Display labels to the left of items in tracks
<input checked="" type="checkbox"/>	Display description above each track
<input checked="" type="checkbox"/>	Show track controls under main graphic
<input type="checkbox"/>	Next/previous item navigation
<input checked="" type="checkbox"/>	Next/previous exon navigation

Configure Tracks on UCSC Genome Browser: Human Feb. 2009 (GRCh37/hg19)

Tracks: Groups:

Control track and group visibility more selectively below.

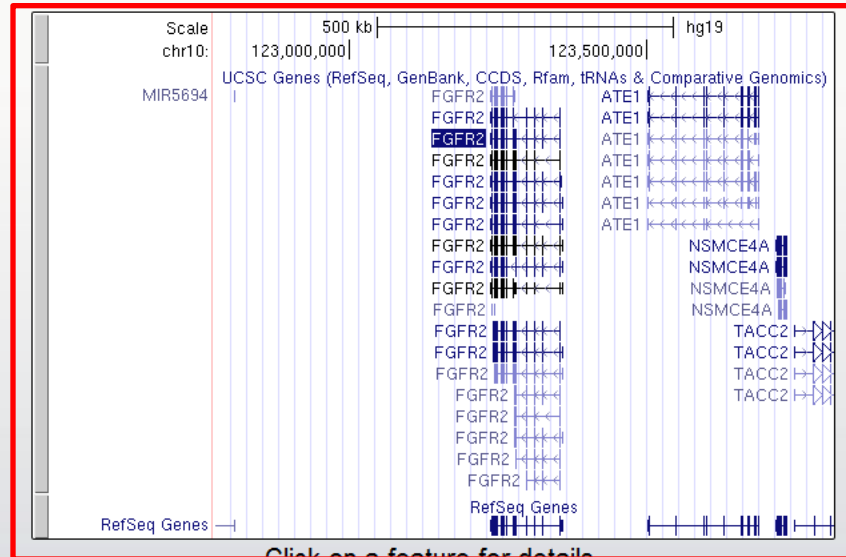
+ Mapping and Sequencing Tracks **+ Phenotype and Disease Associations** **+ Genes and Gene Prediction Tracks** **+ Literature** **+ mRNA and EST Tracks** **+ Expression** **+ Regulation**



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

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

chr10:122,775,292-123,816,033 1,040,742 bp.



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 >

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



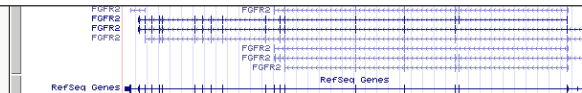
Mapping and Sequencing Tracks



Phenotype and Disease Associations

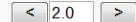


Annotation tracks

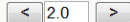


Click on a feature for details.
Click or drag in the base position track to zoom in.
Click side bars for track options. Drag side bars or labels up or down to reorder tracks. Drag tracks left or right to new position.

move start



move end



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

collapse all

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.

expand all

+	Mapping and Sequencing Tracks	refresh
+	Phenotype and Disease Associations	refresh
+	Genes and Gene Prediction Tracks	refresh
+	Literature	refresh
+	mRNA and EST Tracks	refresh
+	Expression	refresh
+	Regulation	refresh
-	Comparative Genomics	refresh

Conservation Cons Indels MmCf GERP Evo Cpg Primate Chain/Net Placental Chain/Net

Vertebrate Chain/Net

+	Neandertal Assembly and Analysis	refresh
+	Denisova Assembly and Analysis	refresh
+	Variation and Repeats	refresh

Genomes Genome Browser Tools Mirrors Downloads My Data About Us Help

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

Search Advanced

conservation

search clear cancel

return to browser (2 of 17 selected)

Visibility	Track Name	Sort: by Relevance Alphabetically
<input checked="" type="checkbox"/> full	Conservation	Vertebrate Multiz Alignment & Conservation (46 Species)
<input type="checkbox"/> hide	Primate Cons	Primate Conservation by PhastCons
<input type="checkbox"/> hide	Vertebrate Cons	Vertebrate Conservation by PhastCons
<input type="checkbox"/> hide	Primate Cons	Primate Basewise Conservation by PhyloP
<input type="checkbox"/> hide	Vertebrate Cons	Vertebrate Basewise Conservation by PhyloP
<input type="checkbox"/> hide	Mammal Cons	Placental Mammal Conservation by PhastCons
<input type="checkbox"/> hide	Cons Indels MmCf	Indel-based Conservation for human hg19, mouse mm8 and dog canFam3
<input checked="" type="checkbox"/> full	Mammal Cons	Placental Mammal Basewise Conservation by PhyloP
<input type="checkbox"/> hide	Mod Hum Variants	Variant Calls from 11 Modern Human Genome Sequences
<input type="checkbox"/> hide	Denisova Variants	Variant Calls from High-Coverage Genome Sequence of an Archaic Denisovan
<input type="checkbox"/> hide	CCDS	Consensus CDS
<input type="checkbox"/> hide	TransMap	TransMap Alignments
<input type="checkbox"/> hide	TransMap UCSC	TransMap UCSC Gene Mappings
<input type="checkbox"/> hide	TransMap RefGene	TransMap RefSeq Gene Mappings
<input type="checkbox"/> hide	TransMap mRNA	TransMap GenBank mRNA Mappings
<input type="checkbox"/> hide	TransMap ESTs	TransMap Spliced EST Mappings
<input type="checkbox"/> hide	GERP	GERP scores for mammalian alignments

Return to Browser (2 of 17 selected)

Tracks so marked are containers which group related data tracks. Containers may need additional configuration (icon) before they can be viewed in the browser.



Annotation tracks

- The database may contain any data that can be mapped to genomic coordinates and therefore can be displayed in the Genome Browser
- Overview of tracks: <http://genome.ucsc.edu/cgi-bin/hgTracks>
- Three different categories:
 - computed at UCSC
 - computed elsewhere and displayed at UCSC
 - computed and hosted entirely elsewhere



Annotation tracks computed at UCSC

- ❑ Comparative genomic annotations as well as Convert and liftOver capabilities
- ❑ mRNAs and ESTs in GenBank are aligned to the reference assembly in separate tracks (75 million GenBank RNAs and ESTs, ~3 billion bases of the human reference assembly → 2 CPU-years of computing time)
- ❑ The Conservation composite track displays the results of the multiz algorithm that aligns the results from up to 46 pairwise Blastz alignments to the reference assembly (e.g. hg19 human assembly consumed 10 CPU-years)



mRNA and EST Tracks refresh

Human mRNAs pack	Spliced ESTs hide	Human ESTs hide	Other mRNAs hide	Other ESTs hide	<input checked="" type="checkbox"/> H-Inv hide
UniGene hide	Gene Bounds hide	SIB Alt-Splicing hide	<input checked="" type="checkbox"/> Poly(A) hide	PolyA-Seq hide	<input checked="" type="checkbox"/> CGAP SAGE hide

[Human RNA Editing](#)
hide

+ **Expression** refresh

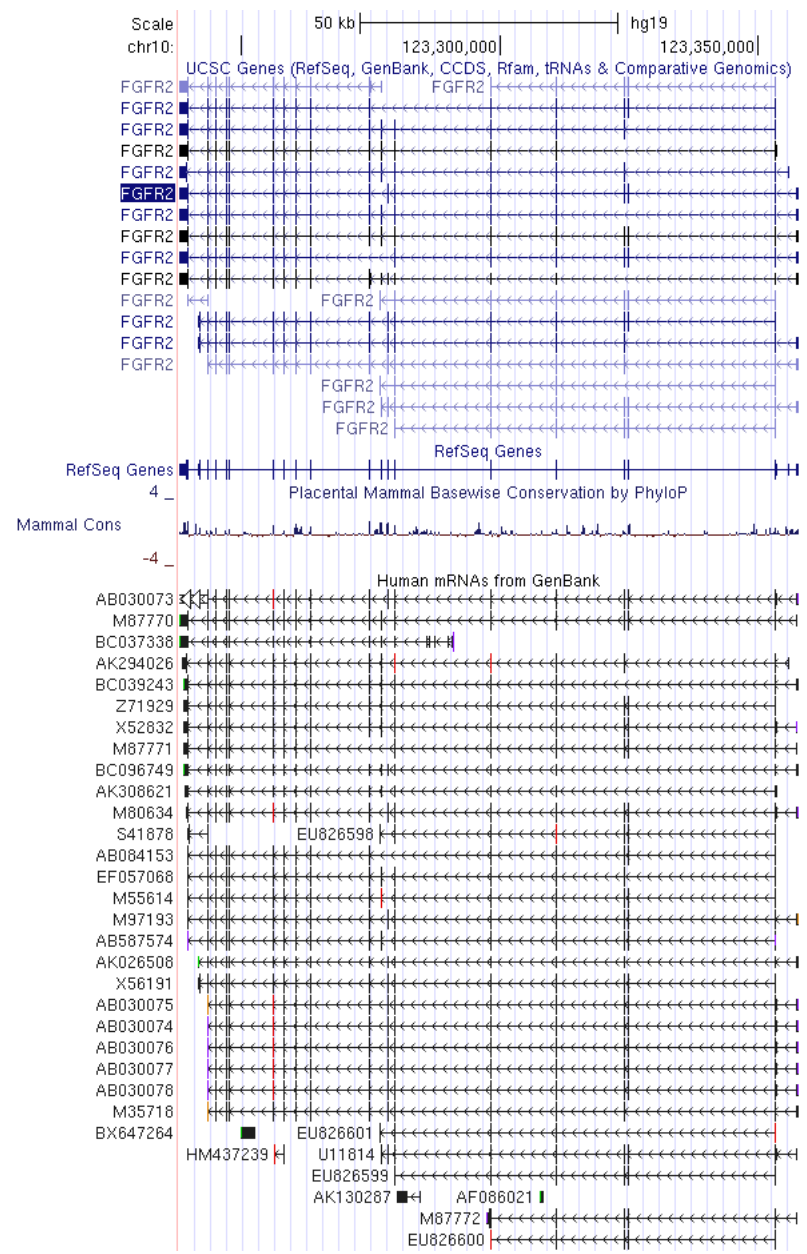
+ **Regulation** refresh

- **Comparative Genomics** refresh

Conservation full	<input checked="" type="checkbox"/> Cons Indels MmCf hide	GERP hide	<input checked="" type="checkbox"/> Evo Cpg hide	Primate Chain/Net hide	Placental Chain/Net hide
--------------------------------------	--	------------------------------	---	---	---

[Vertebrate Chain/Net](#)
hide

+ **Neandertal Assembly and Analysis** refresh



Annotation tracks computed elsewhere and displayed at UCSC

Annotations that are not post-processed by the UCSC

- Probe sets for commercially available microarrays, copy-number variation from the Database of Genomic Variants and expression data from the GNF Expression Atlas
- Data Coordination Center for the ENCODE project allowing access to a large number of functional annotations in relation to gene regulation

Annotations that are post-processed by the UCSC

- dbSNP (Common SNPs, Flagged SNPs, Mult. SNPs)
- OMIM (OMIM Allelic Variant SNPs, OMIM Genes, OMIM Phenotypes)



Phenotype and Disease Associations refresh

[GAD View](#) [DECIPHER](#) **OMIM AV SNPs** [OMIM Genes](#) [OMIM Pheno Loci](#) [COSMIC](#)

[LOVD Variants](#) [HGMD Variants](#) [UniProt Variants](#) [ClinVar Variants](#) [GWAS Catalog](#) [ISCA](#)

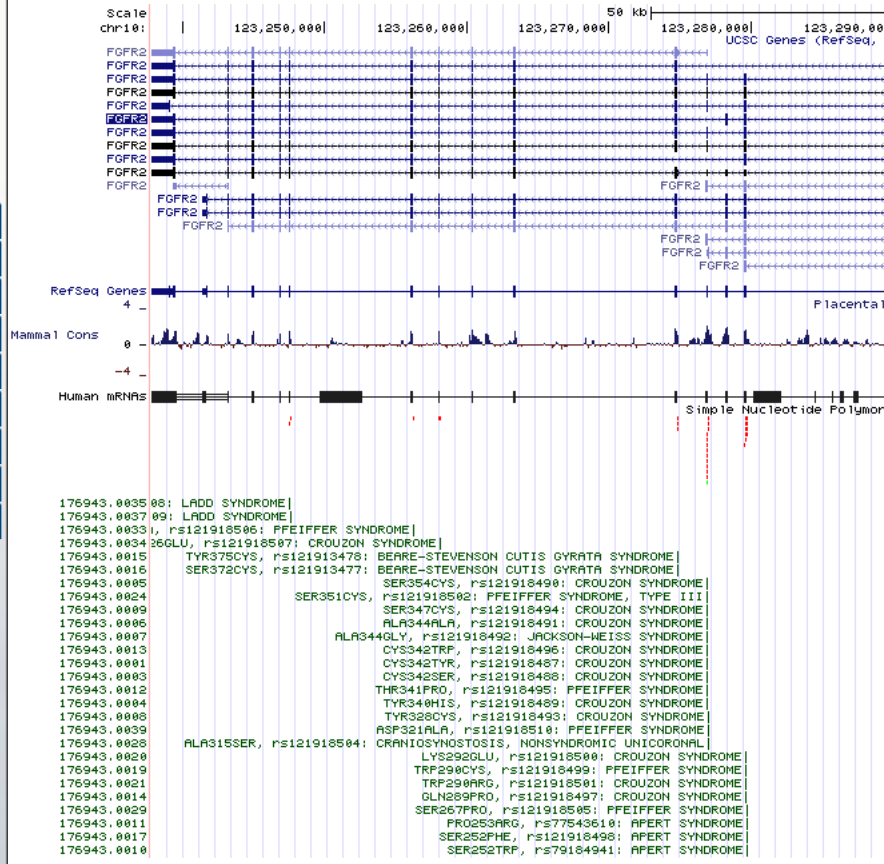
[Coriell CNVs](#) [RGD Human QTL](#) [RGD Rat QTL](#) [MGI Mouse QTL](#) [GeneReviews](#)

Genes and Gene Prediction Tracks refresh

[Literature](#) refresh
[mRNA and EST Tracks](#) refresh
[Expression](#) refresh
[Regulation](#) refresh
[Comparative Genomics](#) refresh
[Neandertal Assembly and Analysis](#) refresh
[Denisova Assembly and Analysis](#) refresh

Variation and Repeats refresh

Common SNPs (138) <input type="button" value="hide"/>	Flagged SNPs (138) <input type="button" value="squish"/>	Mult. SNPs (138) <input type="button" value="hide"/>	All SNPs (138) <input type="button" value="hide"/>	Common SNPs (137) <input type="button" value="hide"/>	Flagged SNPs (137) <input type="button" value="hide"/>
Mult. SNPs (137) <input type="button" value="hide"/>	All SNPs (137) <input type="button" value="hide"/>	Common SNPs (135) <input type="button" value="hide"/>	Flagged SNPs (135) <input type="button" value="hide"/>	Mult. SNPs (135) <input type="button" value="hide"/>	All SNPs (135) <input type="button" value="hide"/>
1000G Ph1 Vars <input type="button" value="hide"/>	1000G Ph1 Accsbl <input type="button" value="hide"/>	GIS DNA PET <input type="button" value="hide"/>	HAIB Genotype <input type="button" value="hide"/>	SNP/CNV Arrays <input type="button" value="hide"/>	HGDP Allele Freq <input type="button" value="hide"/>
HapMap SNPs <input type="button" value="hide"/>	DGV Struct Var <input type="button" value="hide"/>	Segmental Dups <input type="button" value="hide"/>	RepeatMasker <input type="button" value="hide"/>	Interrupted Rpts <input type="button" value="hide"/>	Simple Repeats <input type="button" value="hide"/>



Tracks from the Epigenome project

CpG and MRE sites refresh

Base Position
[CpG Islands](#) [GC Percent](#) [CpG_MRE sites](#)
 dense ▾ full ▾ hide ▾ hide ▾

methylMnM refresh

methylCRF refresh

Epigenome Atlas Data Complete Collection Composite Tracks refresh

Broad Histone [UCSD Histone](#) [UCSF-UBC-USC Histone](#) [DNase](#) [Footprinting](#) [RNA](#)
 hide ▾ hide ▾ hide ▾ hide ▾ dense ▾ hide ▾

DNA Methylation [By Assay...](#) [By Sample...](#) [Roadmap ChromHMM](#) [Roadmap ChromHMM 15 state](#) [Roadmap Uniformly Signal](#)
 hide ▾ hide ▾ hide ▾ hide ▾ hide ▾

Epigenome Atlas Data Complete Collection Integrative Tracks refresh

[Assay Summary...](#) [Sample Summary...](#) [Methylation Summary...](#)
 hide ▾ hide ▾ show ▾

Mapping and Sequencing Tracks refresh

Phenotype and Disease Associations refresh

[GWAS Catalog](#)
 pack ▾

Genes and Gene Prediction Tracks refresh

[UCSC Genes](#) [Alt Events](#) [GENCODE Genes V7](#) [CCDS](#) [RefSeq Genes](#) [Other RefSeq](#)
 pack ▾ hide ▾ hide ▾ hide ▾ dense ▾ hide ▾

[MGC Genes](#) [ORFeome Clones](#) [TransMap...](#) [Ensembl Genes](#) [N-SCAN](#) [Exoniphy](#)
 hide ▾ hide ▾ hide ▾ hide ▾ hide ▾

mRNA and EST Tracks refresh

Expression refresh

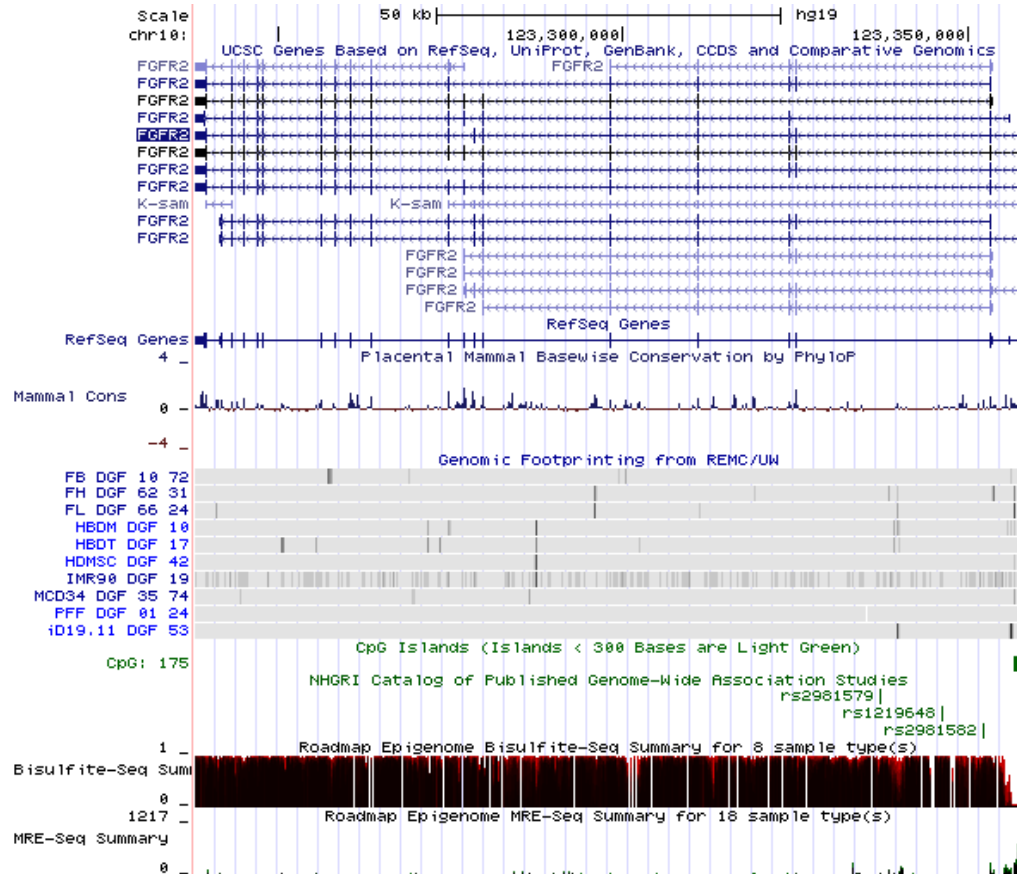
Regulation refresh

Comparative Genomics refresh

[Conservation Chimp Chain/Net](#) [caJac1 Chain/Net](#) [felCat3 Chain/Net](#) [bosTau4 Chain/Net](#) [Primate Chain/Net](#)
 full ▾ hide ▾ hide ▾ hide ▾ hide ▾ hide ▾

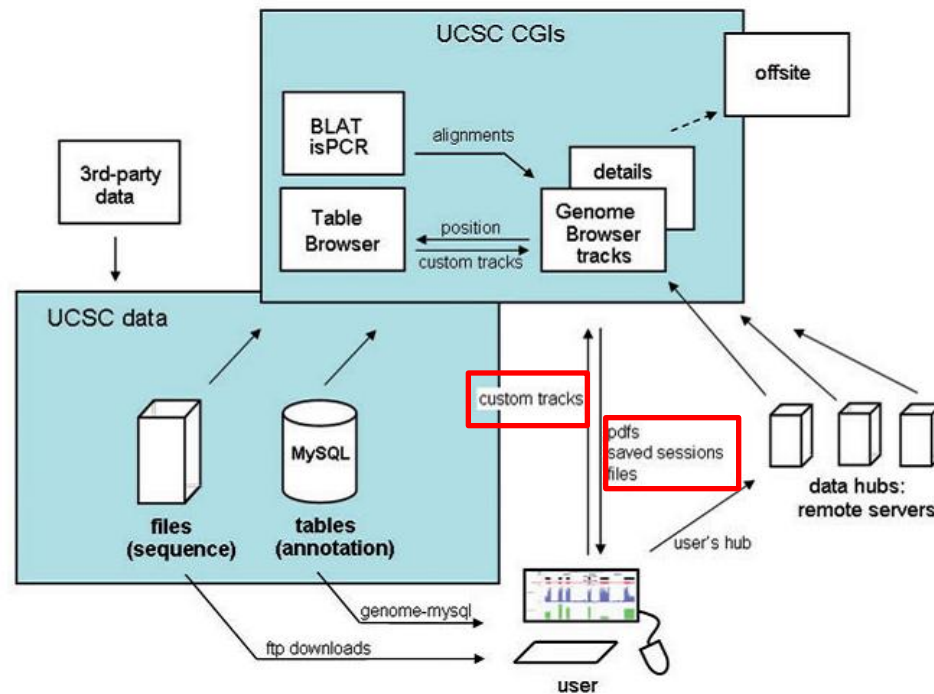
[Placental Chain/Net](#) [Vertebrate Chain/Net](#) [Lizard Chain/Net](#) [xenTro2 Chain/Net](#) [Zebrafish Chain/Net](#) [Sea hare Chain/Net](#)
 hide ▾ hide ▾ hide ▾ hide ▾ hide ▾ hide ▾

Variation and Repeats refresh



Associated Tools

- ❑ Tools other than the main graphic image account for 42% of traffic on the UCSC server



Sessions

Tools Mirrors Downloads My Data About Us View Help

Genome Browser on Human GRCh37/hg19 Assembly

844-123,357,972 120,129 bp. enter position, gene symbol or search terms go

Scale chr10 (q26.13) 50 kb hg19 123,300,000 123,350,000

UCSC Genes (RefSeq, GenBank, CDS, Refam, tRNAs & Comparative Genomics)

RefSeq Genes

Mammal Cons

Human mRNAs from GenBank

OMIM Allelic Variant SNPs

176943.0035 NDROME |
 176943.0037 NDROME |
 176943.0033 FER SYNDROME |
 176943.0034 UZON SYNDROME |
 176943.0015 ION CUTIS GYRATA SYNDROME |
 176943.0016 ION CUTIS GYRATA SYNDROME |
 176943.0005 1918490: CROUZON SYNDROME |
 176943.0024 PEIFFER SYNDROME, TYPE II |
 176943.0009 1918494: CROUZON SYNDROME |
 176943.0005 1918491: CROUZON SYNDROME |
 176943.0007 32: JACKSON-WEISS SYNDROME |
 176943.0013 1918496: CROUZON SYNDROME |
 176943.0011 1918487: CROUZON SYNDROME |
 176943.0003 1918488: CROUZON SYNDROME |
 176943.0012 1918495: PFEIFFER SYNDROME |
 176943.0004 1918489: CROUZON SYNDROME |
 176943.0002 1918492: CROUZON SYNDROME |
 176943.0039 1918510: PFEIFFER SYNDROME |
 176943.0028 121918508: NONSYNDROMIC UNICORNIUM |
 176943.0020 121918509: CROUZON SYNDROME |
 176943.0019 1918499: PFEIFFER SYNDROME |
 176943.0021 121918501: CROUZON SYNDROME |
 176943.0014 121918497: CROUZON SYNDROME |
 176943.0029 21918505: PFEIFFER SYNDROME |
 176943.0011 rs77543610: APERT SYNDROME |
 176943.0017 rs121918498: APERT SYNDROME |
 176943.0010 rs79184941: APERT SYNDROME |

Save Settings

Save current settings as named session:
 name: allow this session to be loaded by others

Save current settings to a local file:
 file: file type returned:
 (leave file blank to get output in browser window)

Restore Settings

Use settings from another user's saved session:
 user: session name:

Use settings from a local file:

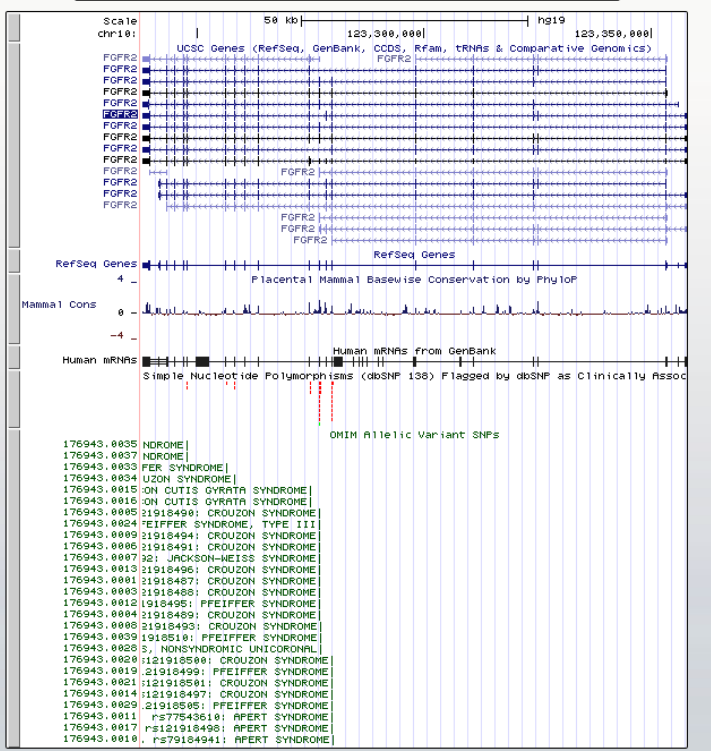
Use settings from a URL (http://..., ftp://...):

Sharing Sessions

- There are several ways to share saved sessions with others.
- Each previously saved named session appears with Browser and Email links. That session loaded. The resulting Genome Browser page can be bookmarked. Email link invokes your email tool with a message containing the Genome Browser URL.
 - If you have saved your settings to a local file, you can send email to others with genome.ucsc.edu/cgi-bin/hgSession.
 - If a saved settings file is available from a web server, you can send email to [hgSession?hgS_doLoadUrl=submit&hgS_loadUrlName=U](mailto:genome.ucsc.edu/cgi-bin/hgSession?hgS_doLoadUrl=submit&hgS_loadUrlName=U) where **U** is the URL of the settings file, e.g. <http://mySession.txt>. In this type of link, you can replace "hgSession" with "hgTrack".

Custom track

- Sessions
- Track Hubs
- Custom Tracks



Add Custom Tracks

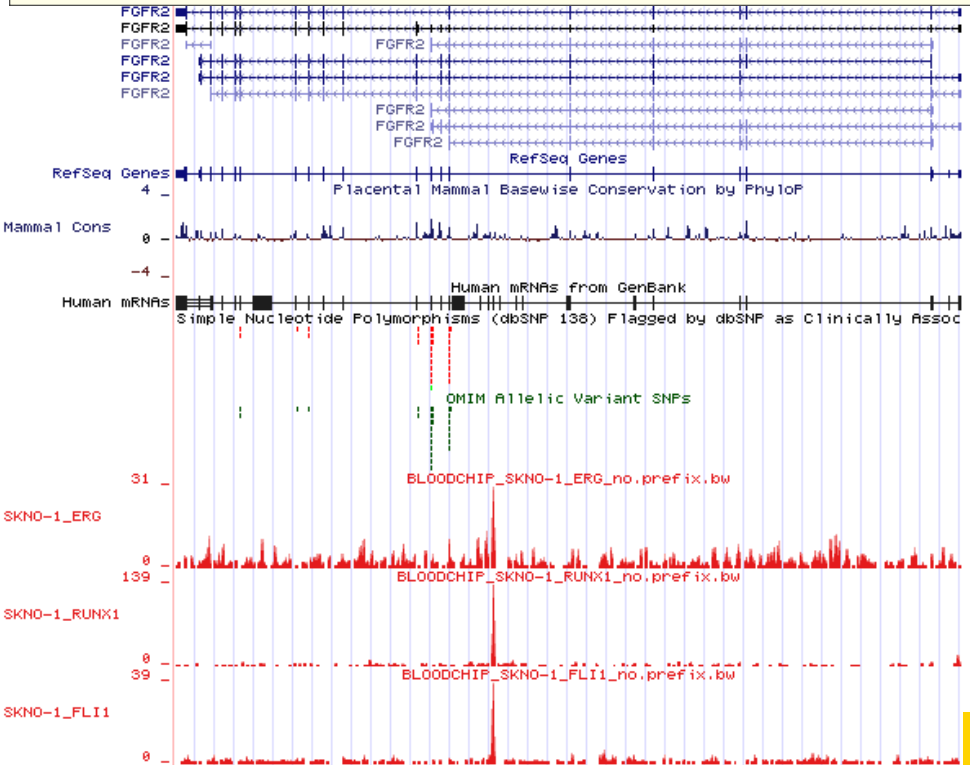
clade Mammal genome Human assembly Feb. 2009 (GRCh37/hg19)

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 must be provided via a URL 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:

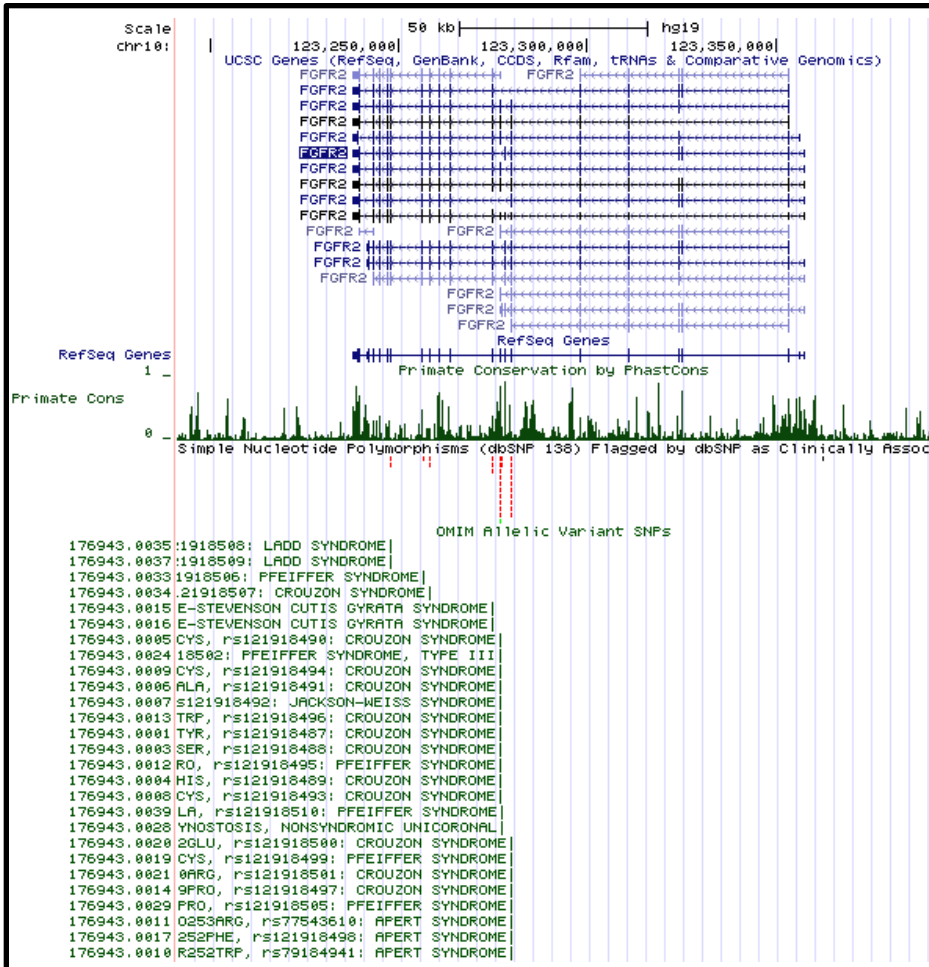
Optional track documentation: Or upload:

Click [here](#) for an HTML document template that may be used for Genome Browser track descriptions.

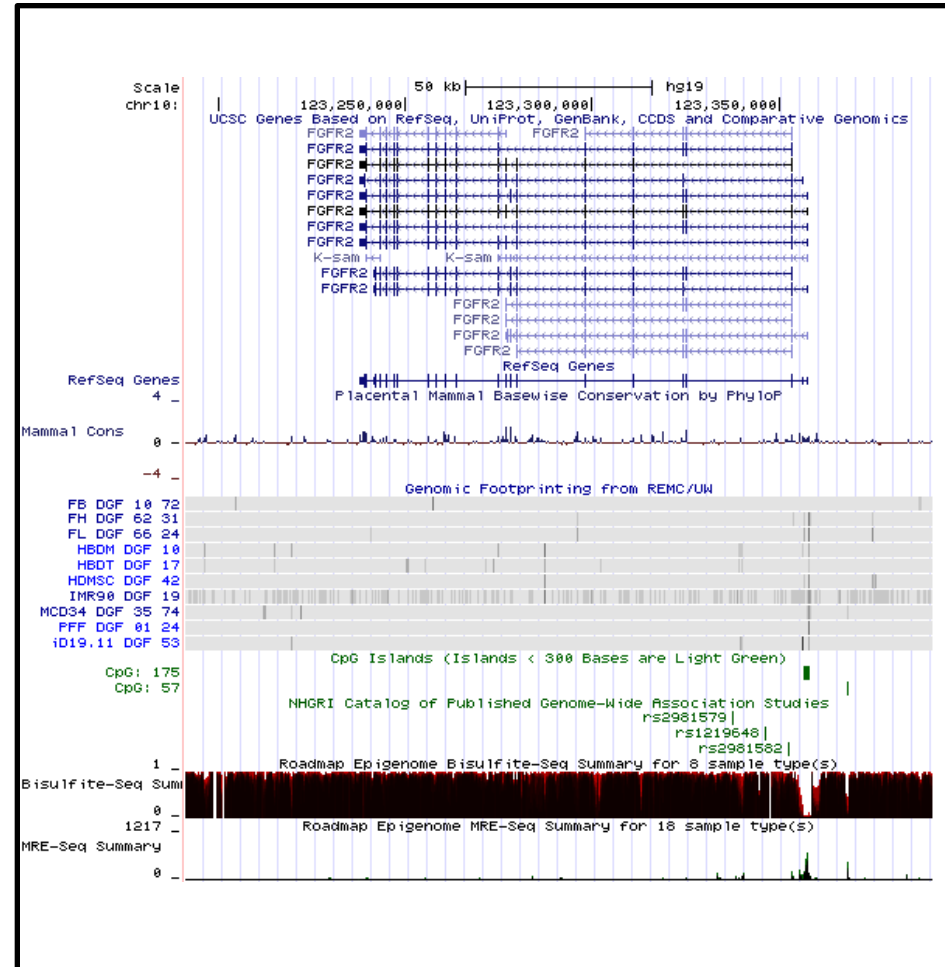


Exercise

UCSC Genome Browser



Epigenome Browser



Exercise

UCSC Genome Browser

- Search for gene of interest

- Clean up the UCSC Browser tracks,
 - show UCSC genes
 - show RefSeq genes

- Find the conservation tracks
 - use “Track Search” and Drop-down tables
 - Visualize Conservation tracks
 - Visualize PhastCons tracks

- Find the dbSNP tracks,
 - Visualize the Flagged SNPs track
 - Modify the “Flagged SNP track” and apply the “squish” visualization

- Find the OMIM tracks
 - Visualize the OMIM AV SNP track.
 - Modify the OMIM AV SNP track and apply the “full” visualization

- Modify text sizes, browser resolution and track colors using the configure button

- Save session as a txt file



Exercise

Epigenome Browser : <http://vizhub.wustl.edu/>

- Apply genome coordinates from gene of interest

- Clean up the Epigenome browser
 - show UCSC genes
 - show RefSeq genes
 - conservation tracks

- Find the Epigenome Atlas Data Complete Collection Composite Tracks
 - Visualize Footprinting tracks and adjust visualization

- Find the CpG and MRE sites
 - Visualize the CpG Island track and adjust visualization

- Find the Phenotype and Disease Associations tracks
 - Visualize GWAS Catalog and adjust visualization

- Find the Epigenome Atlas Data Complete Collection Integrative Tracks
 - Visualize Methylation Summary and hide MeDIP0seq and RRBS Summary

- Modify text sizes, browser resolution and track colors using the configure button

- Save session as a txt file

