NATIONAL HUMAN GENOME RESEARCH INSTITUTE Division of Intramural Research

Current Topics in Genome Analysis 2012

Week 3: Genome Browsers

Tyra Wolfsberg, Ph.D.

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES | NATIONAL INSTITUTES OF HEALTH | genome.gov/DIR

JOHNS HOPKINS MEDICINE
CONTINUING MEDICAL EDUCATION

Current Topics in Genome Analysis 2012

Tyra Wolfsberg, Ph.D.

No Relevant Financial Relationships with Commercial Interests

NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research

Accessing the public genome sequence data

- UCSC's Genome Browser ("Golden Path")
  http://genome.ucsc.edu
- Ensembl
  http://www.ensembl.org
- NCBI's Map Viewer
  http://www.ncbi.nlm.nih.gov/mapview/

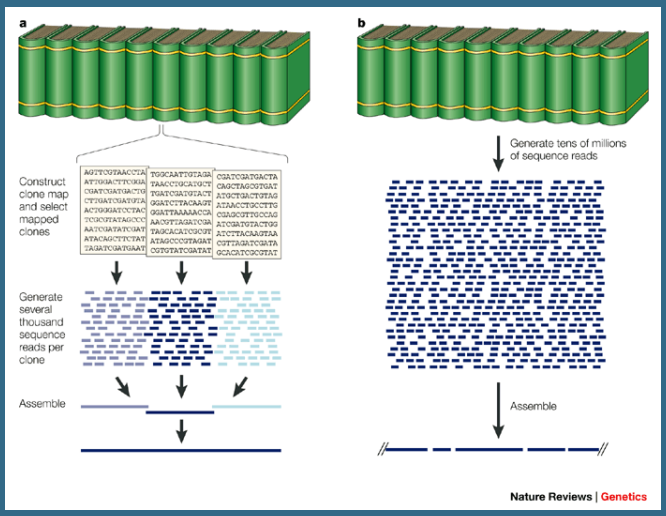
Types of data integrated in genome browsers

- Same starting material for all genome browsers: genomic sequence
- Annotations calculated independently by each genome browser
  - Genes
    - RefSeq mRNAs (non-redundant)
    - GenBank mRNAs (redundant)
    - ESTs
    - Gene predictions
  - SNPs
  - Non-coding functional elements

Overview of genome sequencing strategies

Clone-by-clone shotgun sequencing

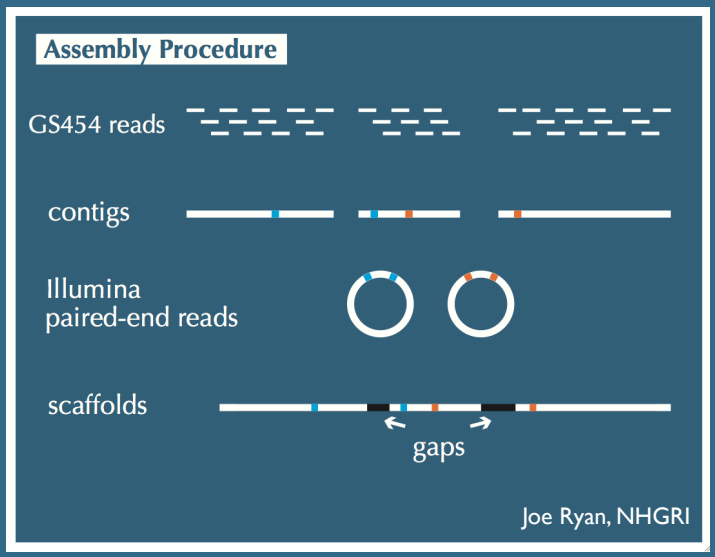
Whole-genome shotgun sequencing



Green ED. Strategies for the systematic sequencing of complex genomes. Nat Rev Genet. 2001. 2:573-83.z

NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research

Next Generation genome assemblies



NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research

Genome Sequence Assemblies

- Complex algorithms needed to incorporate all sequence data
- Assemblies updated periodically as new sequence becomes available
  - Mouse, human, and zebrafish (future) genomes assembled by the Genome Reference Consortium (GRC)
  - Other genomes assembled by sequencing centers or consortia
- Assemblies not updated concurrently by the three Genome Browsers
  - "Pre-release" assemblies and annotations available at
    - UCSC: *http://genome-preview.cse.ucsc.edu/*
    - pre!Ensembl: *http://pre.ensembl.org/*
  - UCSC and Ensembl provide archive of all genome assemblies and annotations; NCBI provides only limited archive
- IF YOU ARE COMPARING DATA FROM DIFFERENT GENOME BROWSERS, MAKE SURE YOU ARE LOOKING AT THE SAME VERSION OF THE ASSEMBLY

NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research

Genome Assembly Versions

	Same assembly?	UCSC	NCBI	Ensembl
Human	Yes	Feb 2009/GRCh37/ hg19	Build 37.3	GRCh37
Mouse	Yes	July 2007 (NCBI37/ mm9)	Build 37.2	NCBIM37
Dog	NO	May 2005 (Broad/ canFam2) canFam3 at *genome-preview.cse.ucsc.edu*	Build 3.1/ CanFam 3.1	BROAD2; CanFam 2.0
Zebrafish	Yes	Jul. 2010 (Zv9/ danRer7)	Zv9	Zv9

NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research

NCBI Reference Sequences (RefSeqs)

- Non-redundant collection of richly annotated DNA, RNA, and protein sequences from diverse taxa
- Each RefSeq represents a single, naturally occurring molecule from one organism

	derived from GenBank submissions	model reference sequences produced by NCBI's Genome Annotation project
mRNA	NM_123456	XM_123456
protein	NP_123456	XP_123456
non-coding transcripts	NR_123456	XR_123456

http://www.ncbi.nlm.nih.gov/RefSeq/key.html

NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research

Beta actin mRNA RefSeq

```
LOCUS       NM_001101               1852 bp    mRNA    linear   PRI 27-DEC-2009
DEFINITION  Homo sapiens actin, beta (ACTB), mRNA.
ACCESSION   NM_001101
VERSION     NM_001101.3  GI:168480144
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1852)
  AUTHORS   Yamaguchi,H., Shiraishi,M., Fukami,K., Tanabe,A., Ikeda-Matsuo,Y.,
            Naito,Y. and Sasaki,Y.
  TITLE     MARCKS regulates lamellipodia formation induced by IGF-I via
            association with PIP2 and beta-actin at membrane microdomains
  JOURNAL   J. Cell. Physiol. 220 (3), 748-755 (2009)
   PUBMED   19475567

COMMENT     REVIEWED REFSEQ: This record has been curated by NCBI staff. The
            reference sequence was derived from AK130157.1 and BC009636.1.
            On Feb 22, 2008 this sequence version replaced gi:5016088.

            Summary: This gene encodes one of six different actin proteins.
            Actins are highly conserved proteins that are involved in cell
            motility, structure, and integrity. This actin is a major
            constituent of the contractile apparatus and one of the two
            nonmuscle cytoskeletal actins. [provided by RefSeq].

            Publication Note:  This RefSeq record includes a subset of the
            publications that are available for this gene. Please see the
            Entrez Gene record to access additional publications.
            COMPLETENESS: complete on the 3' end.

     CDS             85..1212
                     /gene="ACTB"
                     /gene_synonym="PS1TP5BP1"
                     /note="beta cytoskeletal actin; PS1TP5-binding protein 1;
                     actin, cytoplasmic 1"
                     /codon_start=1
                     /product="beta actin"
                     /protein_id="NP_001092.1"
                     /db_xref="GI:4501885"
                     /db_xref="CCDS:CCDS5341.1"
                     /db_xref="GeneID:60"
                     /db_xref="HGNC:132"
                     /db_xref="HPRD:00032"
                     /db_xref="MIM:102630"
                     /translation="MDDDIAALVVDNGSGMCKAGFAGDDAPRAVFPSIVGRPRHQGVM
                     VGMGQKDSYVGDEAQSKRGILTLKYPIEHGIVTNWDDMEKIWHHTFYNELRVAPEEHP
                     VLLTEAPLNPKANREKMTQIMFETFNTPAMYVAIQAVLSLYASGRTTGIVMDSGDGVT
                     HTVPIYEGYALPHAILRLDLAGRDLTDYLMKILTERGYSFTTTAEREIVRDIKEKLCY
                     VALDFEQEMATAASSSSLEKSYELPDGQVITIGNERFRCPEALFQPSFLGMESCGIHE
                     TTFNSIMKCDVDIRKDLYANTVLSGGTTMYPGIADRMQKEITALAPSTMKIKIIAPPE
                     RKYSVWIGGSILASLSTFQQMWISKQEYDESGPSIVHRKCF"

ORIGIN
        1 accgccgaga ccgcgtccgc cccgcgagca cagagcctcg cctttgccga tccgccgccc
       61 gtccacaccc gccgccagct caccatggat gatgatatcg ccgcgctcgt cgtcgacaac
      121 ggctccggca tgtgcaaggc cggcttcgcg ggcgacgatg ccccccgggc cgtcttcccc
```

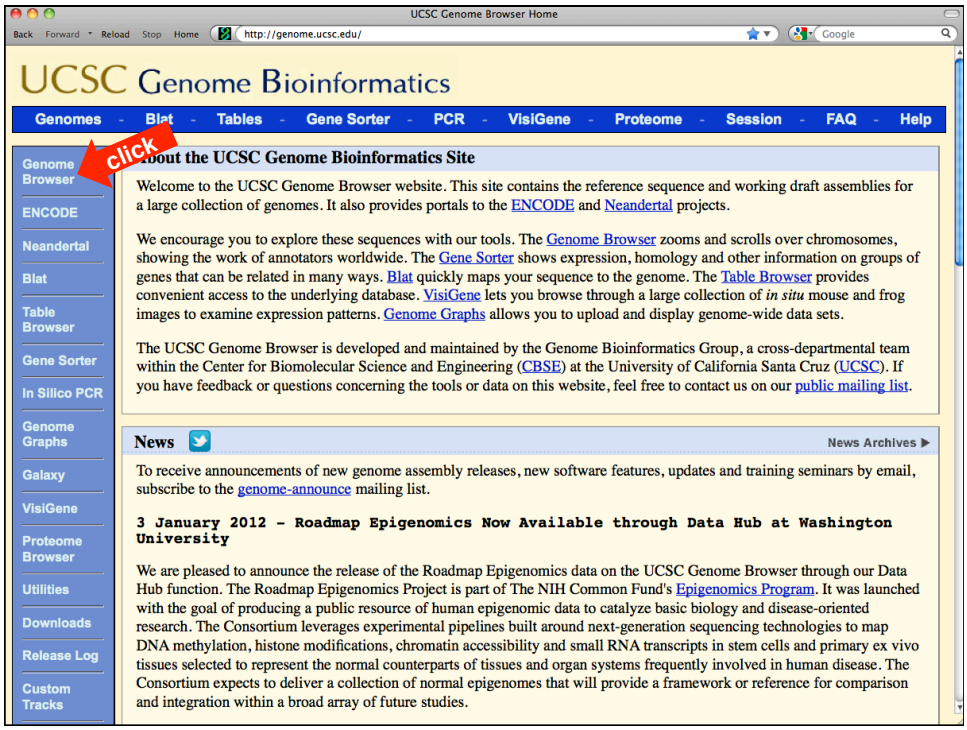
NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research

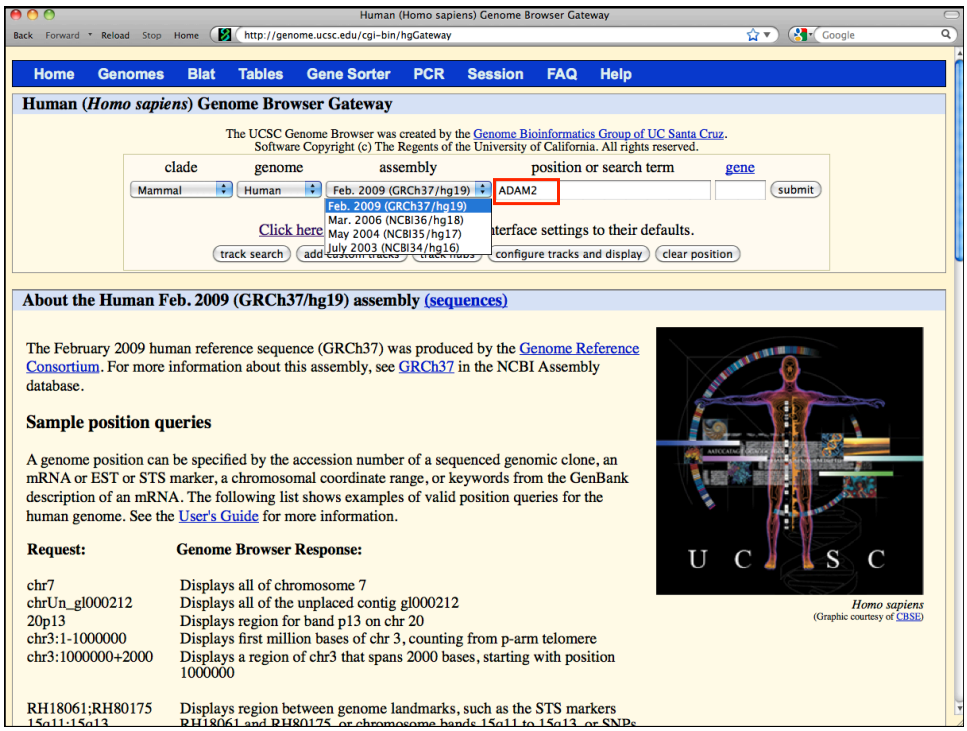
UCSC

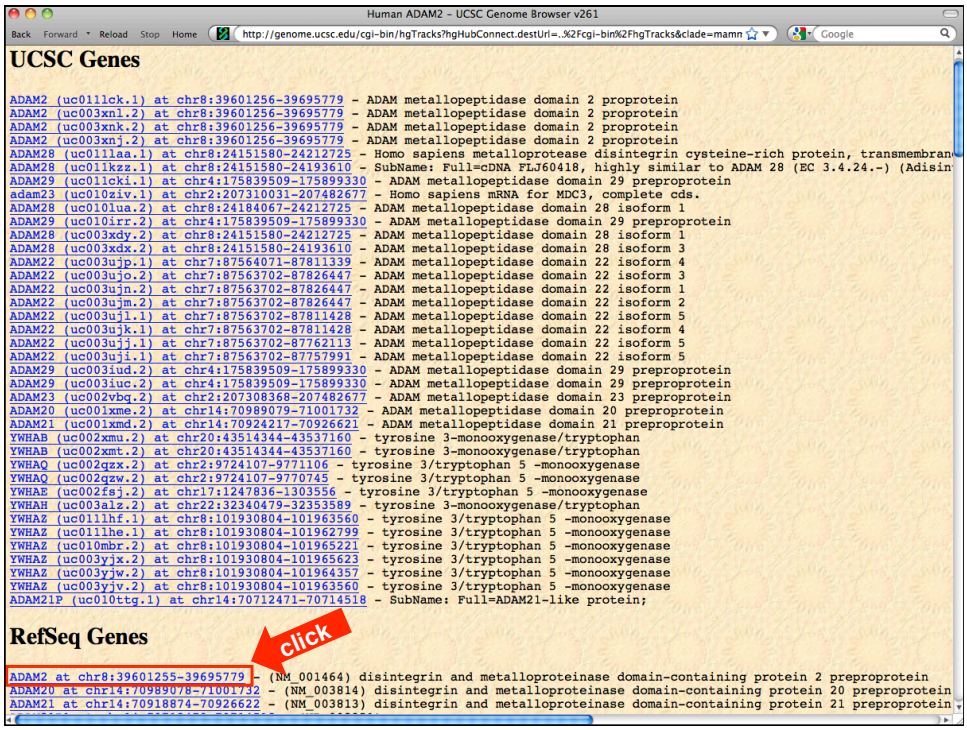
View a region in the genome by querying with a gene symbol

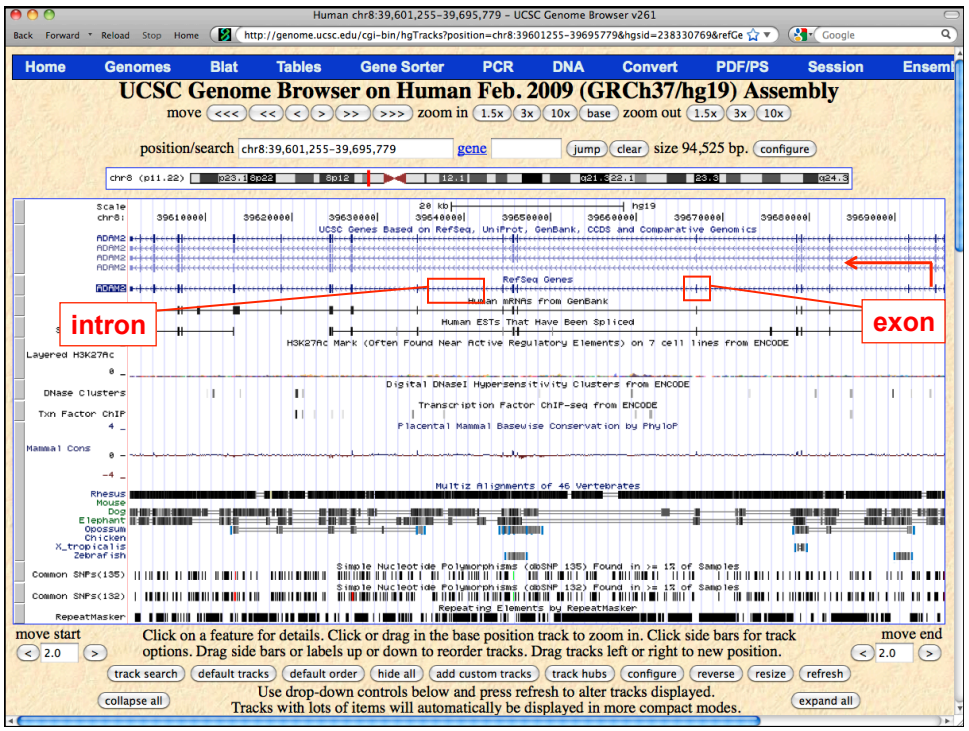
http://genome.ucsc.edu

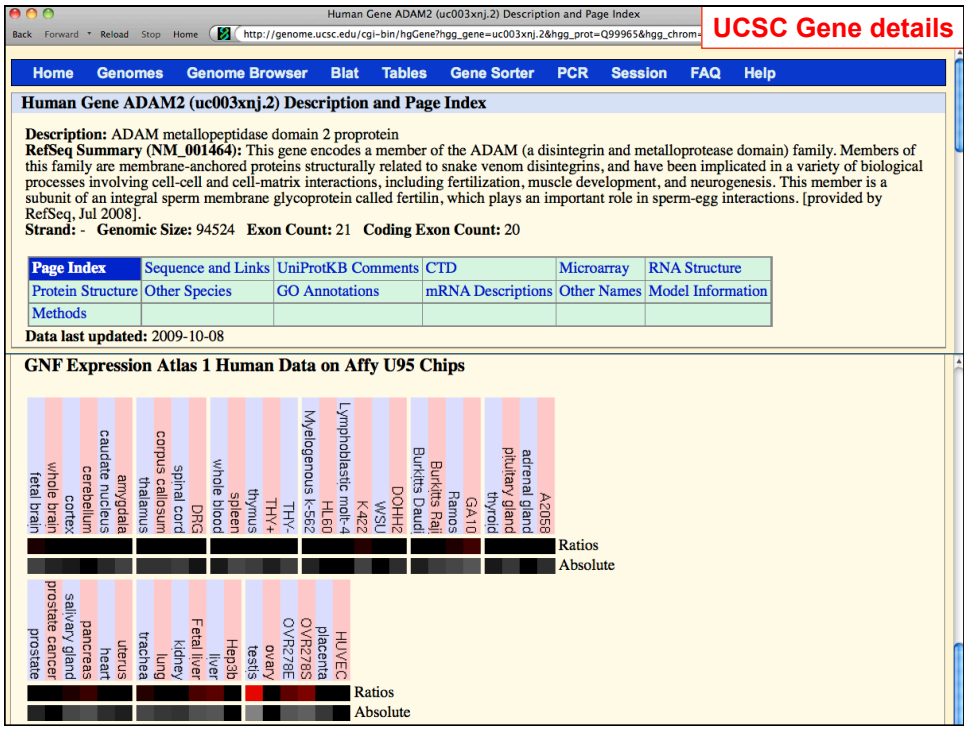
NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research

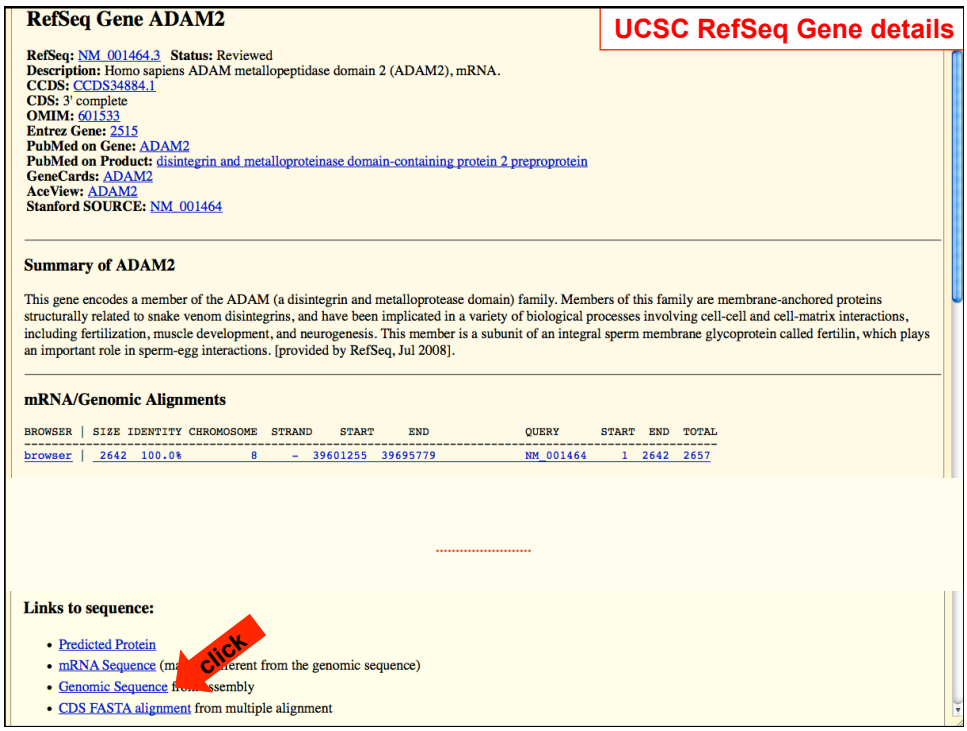


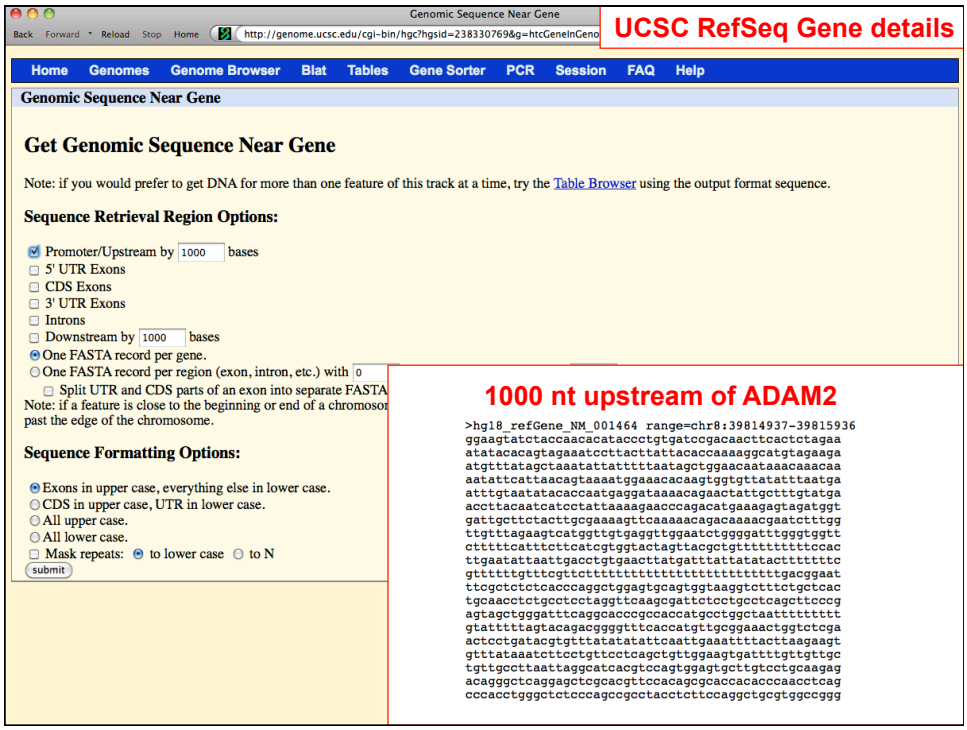

Human (Homo sapiens) Genome Browser Gateway
http://genome.ucsc.edu/cgi-bin/hgGateway
Home Genomes Blat Tables Gene Sorter PCR Session FAQ Help
Human (Homo sapiens) Genome Browser Gateway
The UCSC Genome Browser was created by the Genome Bioinformatics Group of UC Santa Cruz.
Software Copyright (c) The Regents of the University of California. All rights reserved.
clade genome assembly position or search term gene
Mammal Human Feb. 2009 (GRCh37/hg19) ADAM2 submit
Feb. 2009 (GRCh37/hg19)
Mar. 2006 (NCBI36/hg18)
May 2004 (NCBI35/hg17)
July 2003 (NCBI34/hg16)
Click here
track search
configure tracks and display
clear position
About the Human Feb. 2009 (GRCh37/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_gl000212 Displays all of the unplaced contig gl000212
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 Displays region between genome landmarks, such as the STS markers
U C S C
Homo sapiens
(Graphic courtesy of CBSE)


Human ADAM2 - UCSC Genome Browser v261
http://genome.ucsc.edu/cgi-bin/hgTracks?hgHubConnect.destUrl=..%2Fcgi-bin%2FhgTracks&clade=mamm
UCSC Genes
ADAM2 (uc011lck.1) at chr8:39601256-39695779 - ADAM metallopeptidase domain 2 proprotein
ADAM2 (uc003xnl.2) at chr8:39601256-39695779 - ADAM metallopeptidase domain 2 proprotein
ADAM2 (uc003xnk.2) at chr8:39601256-39695779 - ADAM metallopeptidase domain 2 proprotein
ADAM2 (uc003xnj.2) at chr8:39601256-39695779 - ADAM metallopeptidase domain 2 proprotein
ADAM28 (uc011laa.1) at chr8:24151580-24212725 - Homo sapiens metalloprotease disintegrin cysteine-rich protein, transmembran
ADAM28 (uc011kzz.1) at chr8:24151580-24193610 - SubName: Full=cDNA FLJ60418, highly similar to ADAM 28 (EC 3.4.24.-) (Adisin
ADAM29 (uc011cki.1) at chr4:175839509-175899330 - ADAM metallopeptidase domain 29 preproprotein
adam23 (uc010ziv.1) at chr2:207310031-207482677 - Homo sapiens mRNA for MDC3, complete cds.
ADAM28 (uc010lua.2) at chr8:24184067-24212725 - ADAM metallopeptidase domain 28 isoform 1
ADAM29 (uc010irr.2) at chr4:175839509-175899330 - ADAM metallopeptidase domain 29 preproprotein
ADAM28 (uc003xdy.2) at chr8:24151580-24212725 - ADAM metallopeptidase domain 28 isoform 1
ADAM28 (uc003xdx.2) at chr8:24151580-24193610 - ADAM metallopeptidase domain 28 isoform 3
ADAM22 (uc003ujp.1) at chr7:87564071-87811339 - ADAM metallopeptidase domain 22 isoform 4
ADAM22 (uc003ujo.2) at chr7:87563702-87826447 - ADAM metallopeptidase domain 22 isoform 3
ADAM22 (uc003ujn.2) at chr7:87563702-87826447 - ADAM metallopeptidase domain 22 isoform 1
ADAM22 (uc003ujm.2) at chr7:87563702-87826447 - ADAM metallopeptidase domain 22 isoform 2
ADAM22 (uc003ujl.1) at chr7:87563702-87811428 - ADAM metallopeptidase domain 22 isoform 5
ADAM22 (uc003ujk.1) at chr7:87563702-87811428 - ADAM metallopeptidase domain 22 isoform 4
ADAM22 (uc003ujj.1) at chr7:87563702-87762113 - ADAM metallopeptidase domain 22 isoform 5
ADAM22 (uc003uji.1) at chr7:87563702-87757991 - ADAM metallopeptidase domain 22 isoform 5
ADAM29 (uc003iud.2) at chr4:175839509-175899330 - ADAM metallopeptidase domain 29 preproprotein
ADAM29 (uc003iuc.2) at chr4:175839509-175899330 - ADAM metallopeptidase domain 29 preproprotein
ADAM23 (uc002vbq.2) at chr2:207308368-207482677 - ADAM metallopeptidase domain 23 preproprotein
ADAM20 (uc001xme.2) at chr14:70989079-71001732 - ADAM metallopeptidase domain 20 preproprotein
ADAM21 (uc001xmd.2) at chr14:70924217-70926621 - ADAM metallopeptidase domain 21 preproprotein
YWHAB (uc002xmu.2) at chr20:43514344-43537160 - tyrosine 3-monooxygenase/tryptophan
YWHAB (uc002xmt.2) at chr20:43514344-43537160 - tyrosine 3-monooxygenase/tryptophan
YWHAQ (uc002qzx.2) at chr2:9724107-9771106 - tyrosine 3/tryptophan 5 -monooxygenase
YWHAQ (uc002qzw.2) at chr2:9724107-9770745 - tyrosine 3/tryptophan 5 -monooxygenase
YWHAE (uc002fsj.2) at chr17:1247836-1303556 - tyrosine 3/tryptophan 5 -monooxygenase
YWHAH (uc003alz.2) at chr22:32340479-32353589 - tyrosine 3-monooxygenase/tryptophan
YWHAZ (uc011lhf.1) at chr8:101930804-101963560 - tyrosine 3/tryptophan 5 -monooxygenase
YWHAZ (uc011lhe.1) at chr8:101930804-101962799 - tyrosine 3/tryptophan 5 -monooxygenase
YWHAZ (uc010mbr.2) at chr8:101930804-101965221 - tyrosine 3/tryptophan 5 -monooxygenase
YWHAZ (uc003yjx.2) at chr8:101930804-101965623 - tyrosine 3/tryptophan 5 -monooxygenase
YWHAZ (uc003yjw.2) at chr8:101930804-101964357 - tyrosine 3/tryptophan 5 -monooxygenase
YWHAZ (uc003yjv.2) at chr8:101930804-101963560 - tyrosine 3/tryptophan 5 -monooxygenase
ADAM21P (uc010ttg.1) at chr14:70712471-70714518 - SubName: Full=ADAM21-like protein;
RefSeq Genes
click
ADAM2 at chr8:39601255-39695779 - (NM_001464) disintegrin and metalloproteinase domain-containing protein 2 preproprotein
ADAM20 at chr14:70989078-71001732 - (NM_003814) disintegrin and metalloproteinase domain-containing protein 20 preproprotein
ADAM21 at chr14:70918874-70926622 - (NM_003813) disintegrin and metalloproteinase domain-containing protein 21 preproprotein


Human chr8:39,601,255-39,695,779 – UCSC Genome Browser v261
UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly
intron
exon
RefSeq Genes
Human mRNAs from GenBank
Human ESTs That Have Been Spliced
H3K27Ac Mark (Often Found Near Active Regulatory Elements) on 7 cell lines from ENCODE
Digital DNaseI Hypersensitivity Clusters from ENCODE
Transcription Factor ChIP-seq from ENCODE
Placental Mammal Basewise Conservation by PhyloP
Multiz Alignments of 46 Vertebrates
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.


UCSC Gene details
Human Gene ADAM2 (uc003xnj.2) Description and Page Index
Description: ADAM metallopeptidase domain 2 proprotein
RefSeq Summary (NM_001464): This gene encodes a member of the ADAM (a disintegrin and metalloprotease domain) family. Members of this family are membrane-anchored proteins structurally related to snake venom disintegrins, and have been implicated in a variety of biological processes involving cell-cell and cell-matrix interactions, including fertilization, muscle development, and neurogenesis. This member is a subunit of an integral sperm membrane glycoprotein called fertilin, which plays an important role in sperm-egg interactions. [provided by RefSeq, Jul 2008].
Strand: - Genomic Size: 94524 Exon Count: 21 Coding Exon Count: 20
Page Index
Sequence and Links
UniProtKB Comments
CTD
Microarray
RNA Structure
Protein Structure
Other Species
GO Annotations
mRNA Descriptions
Other Names
Model Information
Methods
Data last updated: 2009-10-08
GNF Expression Atlas 1 Human Data on Affy U95 Chips
Ratios
Absolute


RefSeq Gene ADAM2
UCSC RefSeq Gene details
RefSeq: NM_001464.3 Status: Reviewed
Description: Homo sapiens ADAM metallopeptidase domain 2 (ADAM2), mRNA.
CCDS: CCDS34884.1
CDS: 3' complete
OMIM: 601533
Entrez Gene: 2515
PubMed on Gene: ADAM2
PubMed on Product: disintegrin and metalloproteinase domain-containing protein 2 preproprotein
GeneCards: ADAM2
AceView: ADAM2
Stanford SOURCE: NM_001464
Summary of ADAM2
This gene encodes a member of the ADAM (a disintegrin and metalloprotease domain) family. Members of this family are membrane-anchored proteins structurally related to snake venom disintegrins, and have been implicated in a variety of biological processes involving cell-cell and cell-matrix interactions, including fertilization, muscle development, and neurogenesis. This member is a subunit of an integral sperm membrane glycoprotein called fertilin, which plays an important role in sperm-egg interactions. [provided by RefSeq, Jul 2008].
mRNA/Genomic Alignments
BROWSER | SIZE IDENTITY CHROMOSOME STRAND START END QUERY START END TOTAL
browser | 2642 100.0% 8 - 39601255 39695779 NM_001464 1 2642 2657
Links to sequence:
Predicted Protein
mRNA Sequence
Genomic Sequence
CDS FASTA alignment from multiple alignment
click

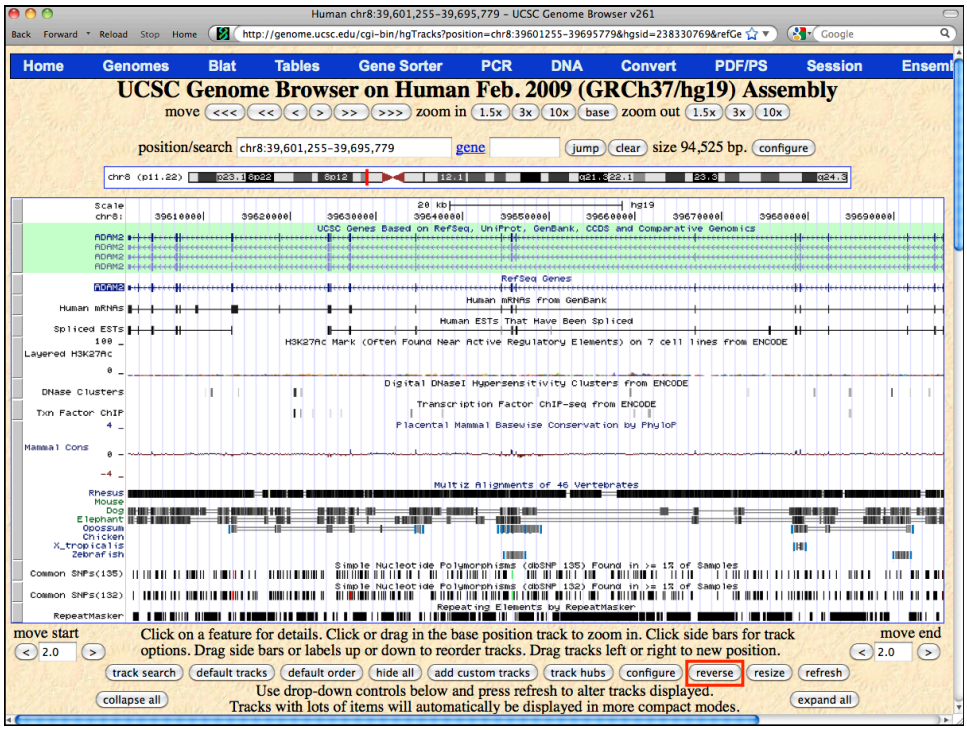

Genomic Sequence Near Gene
UCSC RefSeq Gene details
Home Genomes Genome Browser Blat Tables Gene Sorter PCR Session FAQ Help
Genomic Sequence Near Gene
Get Genomic Sequence Near Gene
Note: if you would prefer to get DNA for more than one feature of this track at a time, try the Table Browser using the output format sequence.
Sequence Retrieval Region Options:
Promoter/Upstream by 1000 bases
5' UTR Exons
CDS Exons
3' UTR Exons
Introns
Downstream by 1000 bases
One FASTA record per gene.
One FASTA record per region (exon, intron, etc.) with 0
Split UTR and CDS parts of an exon into separate FASTA
Sequence Formatting Options:
Exons in upper case, everything else in lower case.
CDS in upper case, UTR in lower case.
All upper case.
All lower case.
Mask repeats: to lower case to N
submit
1000 nt upstream of ADAM2
>hg18_refGene_NM_001464 range=chr8:39814937-39815936

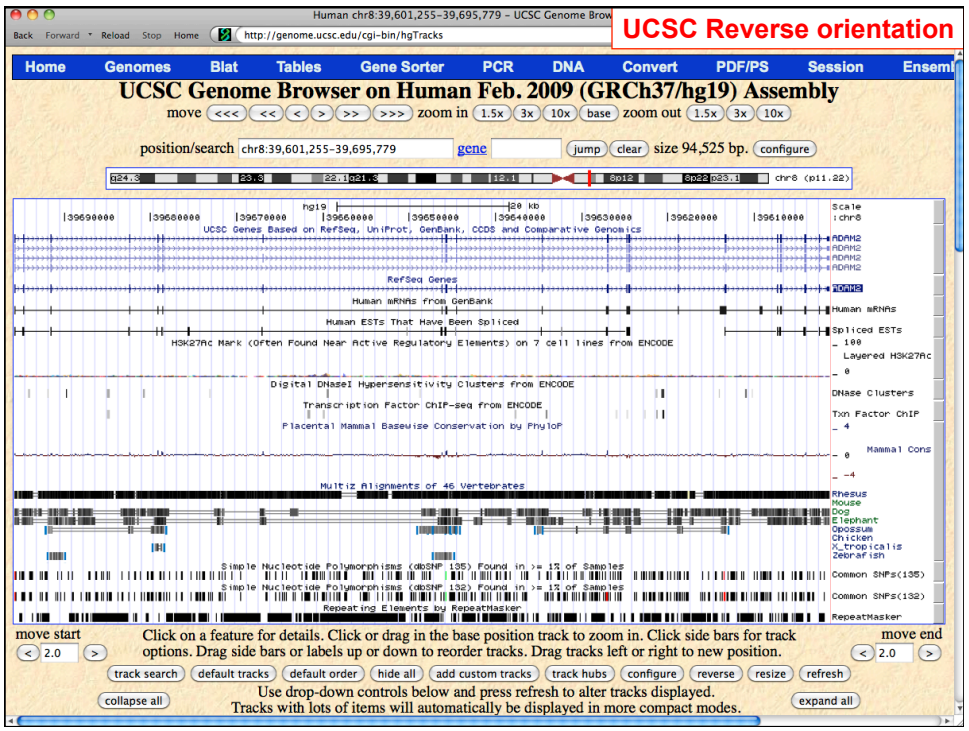
UCSC

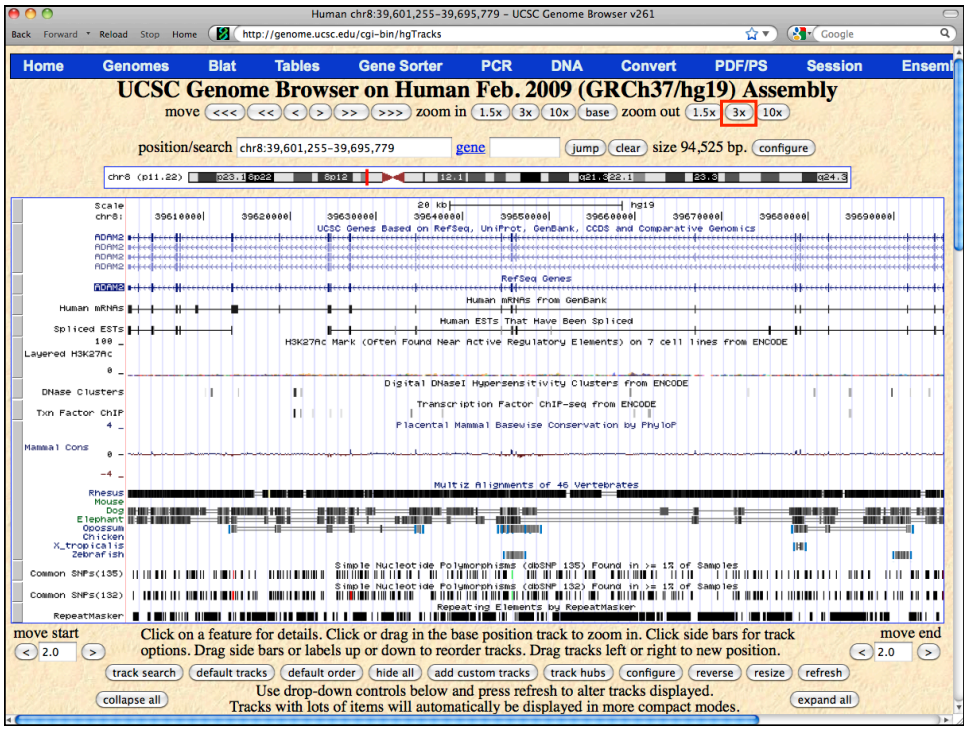
Navigating around the Genome Browser

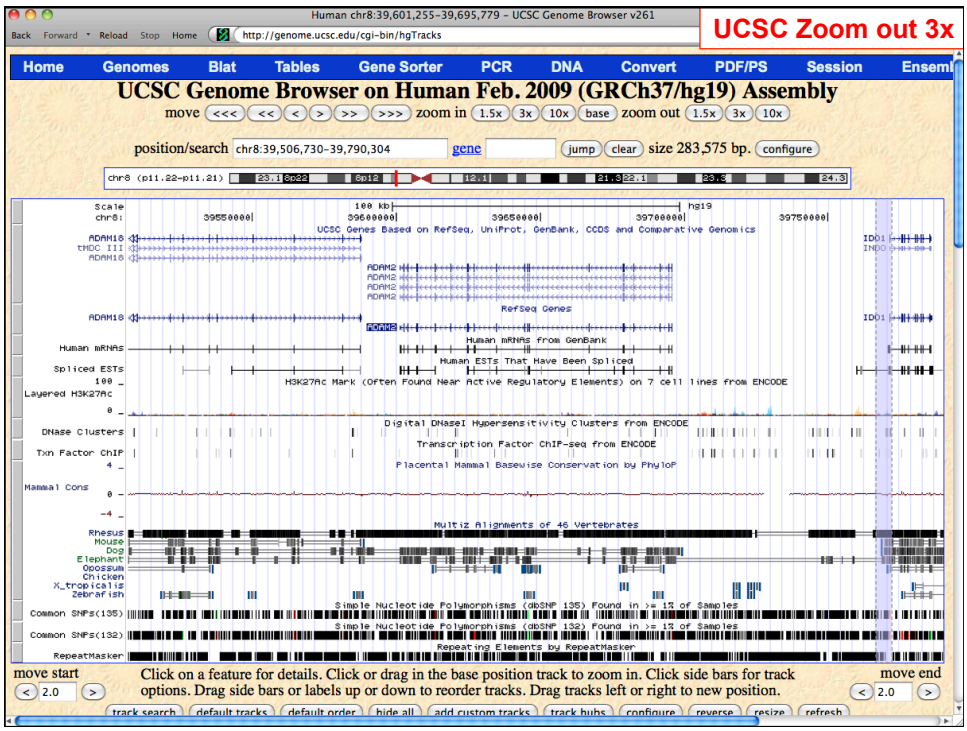
http://genome.ucsc.edu

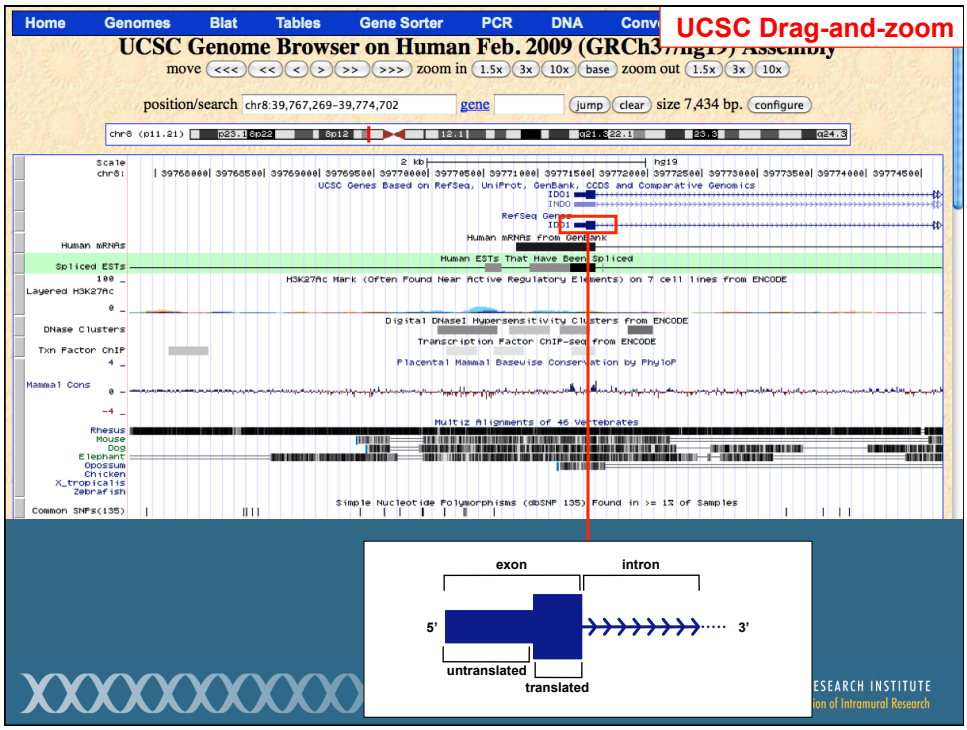
NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research




UCSC Reverse orientation
Human chr8:39,601,255-39,695,779 - UCSC Genome Brow
http://genome.ucsc.edu/cgi-bin/hgTracks
Home Genomes Blat Tables Gene Sorter PCR DNA Convert PDF/PS Session
UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly
move zoom in zoom out
position/search chr8:39,601,255-39,695,779 gene jump clear size 94,525 bp. configure
RefSeq Genes
Human mRNAs from GenBank
Human ESTs That Have Been Spliced
H3K27Ac Mark (Often Found Near Active Regulatory Elements) on 7 cell lines from ENCODE
Digital DNaseI Hypersensitivity Clusters from ENCODE
Transcription Factor ChIP-seq from ENCODE
Placental Mammal Basewise Conservation by PhyloP
Multiz Alignments of 46 Vertebrates
Simple Nucleotide Polymorphisms (dbSNP 135) Found in >= 1% of Samples
Simple Nucleotide Polymorphisms (dbSNP 132) Found in >= 1% of Samples
Repeating Elements by RepeatMasker
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.
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


Human chr8:39,601,255-39,695,779 - UCSC Genome Browser v261
http://genome.ucsc.edu/cgi-bin/hgTracks
Home Genomes Blat Tables Gene Sorter PCR DNA Convert PDF/PS Session
UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly
move zoom in zoom out
position/search chr8:39,601,255-39,695,779 gene jump clear size 94,525 bp. configure
RefSeq Genes
Human mRNAs from GenBank
Human ESTs That Have Been Spliced
H3K27Ac Mark (Often Found Near Active Regulatory Elements) on 7 cell lines from ENCODE
Digital DNaseI Hypersensitivity Clusters from ENCODE
Transcription Factor ChIP-seq from ENCODE
Placental Mammal Basewise Conservation by PhyloP
Multiz Alignments of 46 Vertebrates
Simple Nucleotide Polymorphisms (dbSNP 135) Found in >= 1% of Samples
Simple Nucleotide Polymorphisms (dbSNP 132) Found in >= 1% of Samples
Repeating Elements by RepeatMasker
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.
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


UCSC Zoom out 3x
Human chr8:39,601,255-39,695,779 - UCSC Genome Browser v261
http://genome.ucsc.edu/cgi-bin/hgTracks
Home Genomes Blat Tables Gene Sorter PCR DNA Convert PDF/PS Session
UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly
position/search chr8:39,506,730-39,790,304
size 283,575 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.

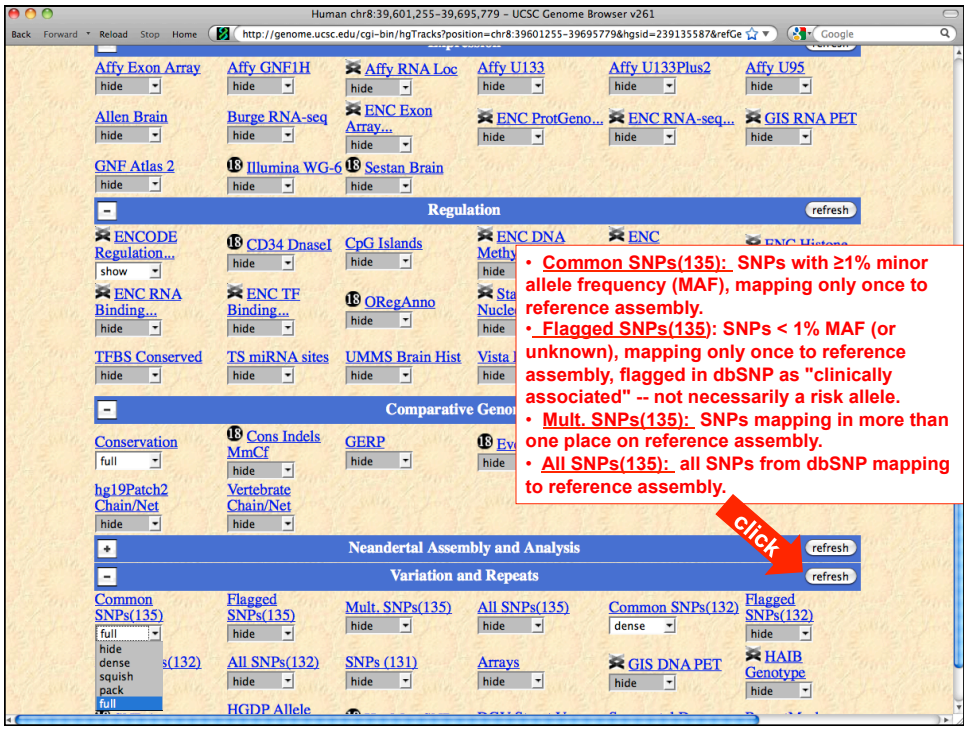

UCSC Drag-and-zoom
Home Genomes Blat Tables Gene Sorter PCR DNA
position/search chr8:39,767,269-39,774,702
size 7,434 bp.
exon
intron
5'
3'
untranslated
translated

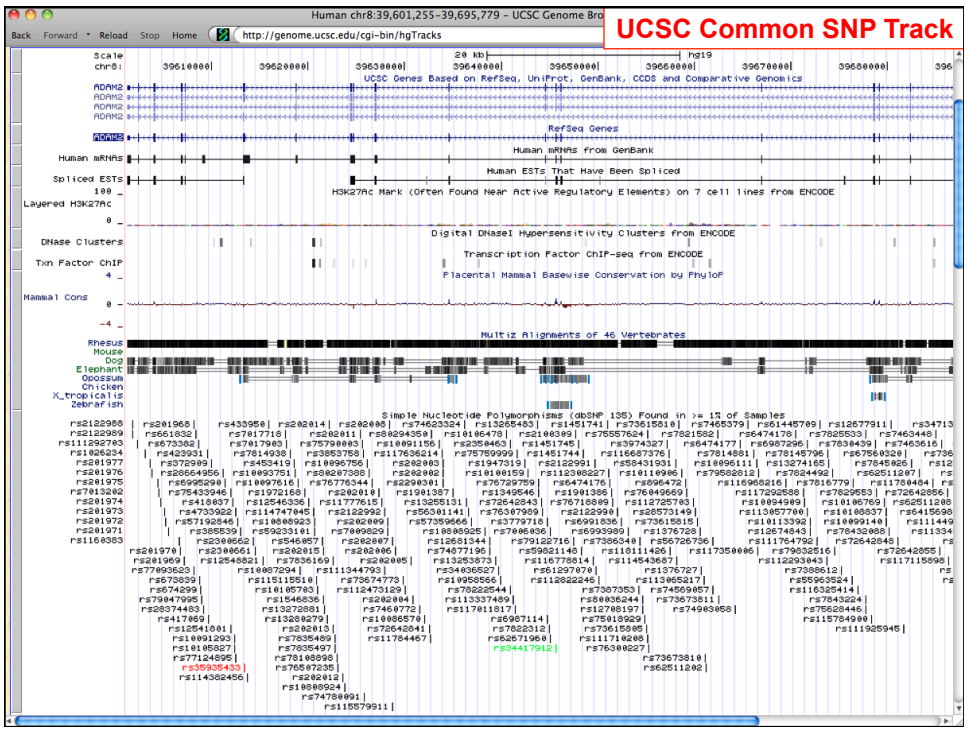
UCSC

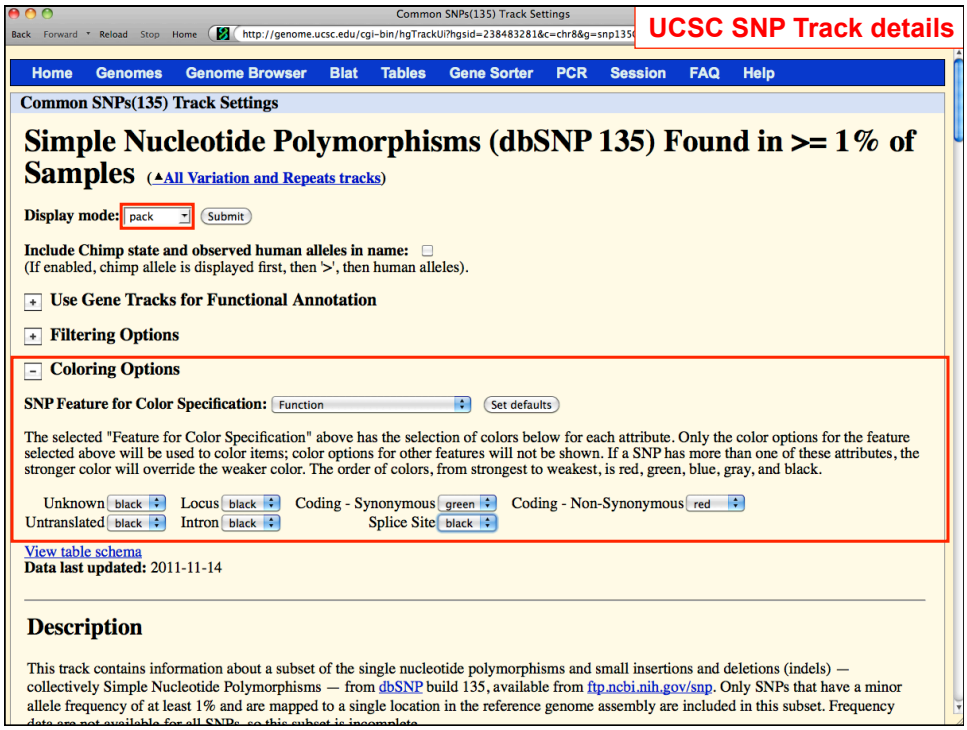
Add a track to the Genome Browser

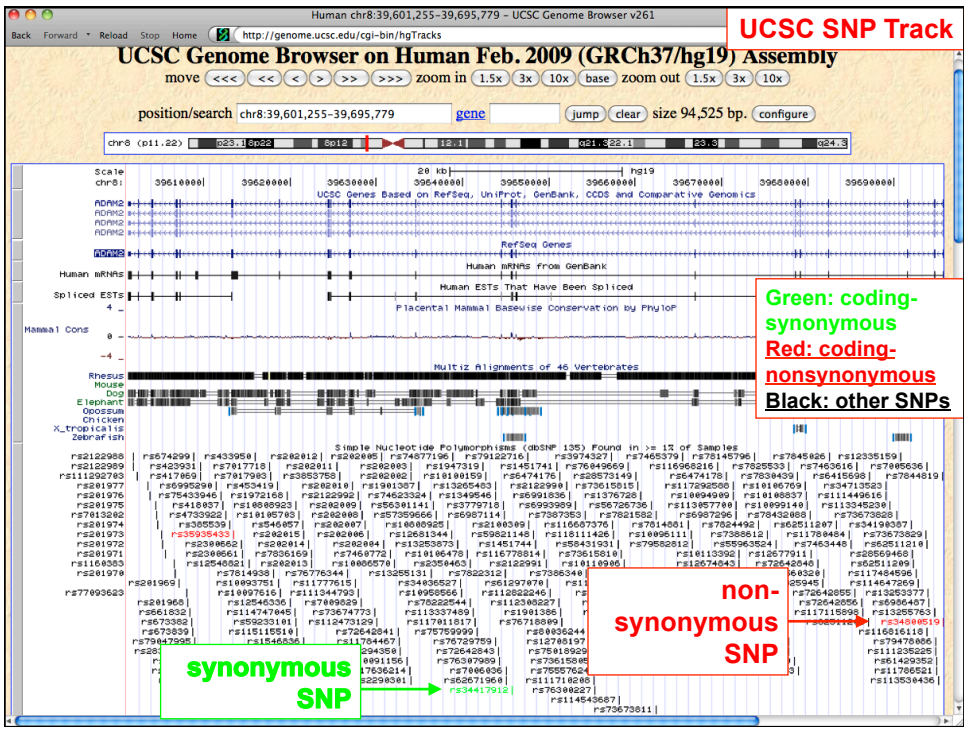
http://genome.ucsc.edu

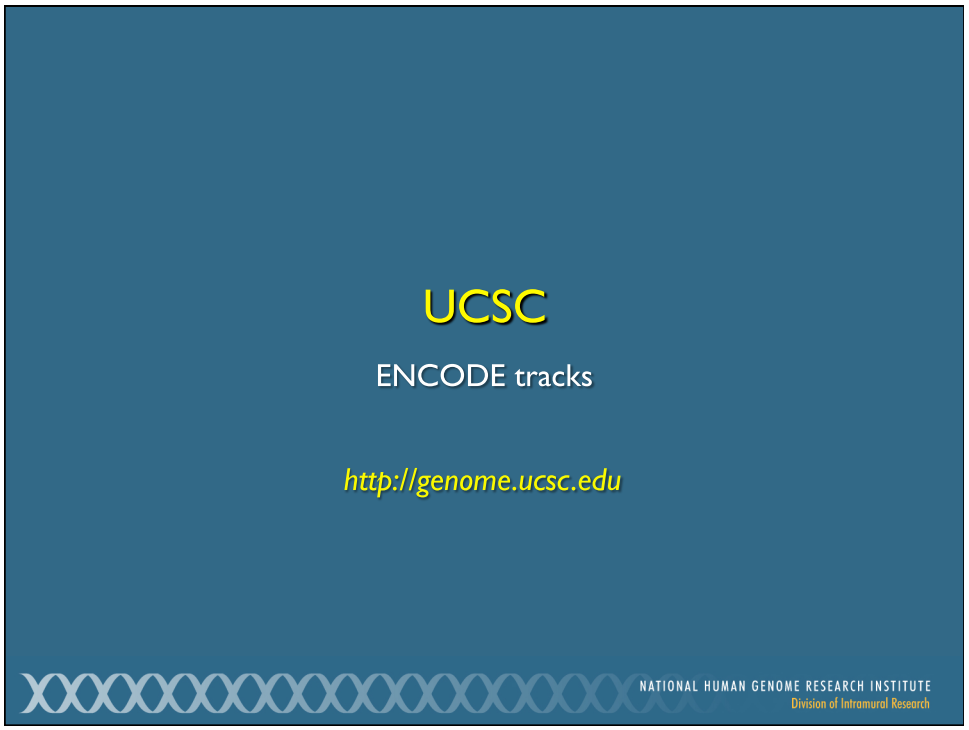
NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research

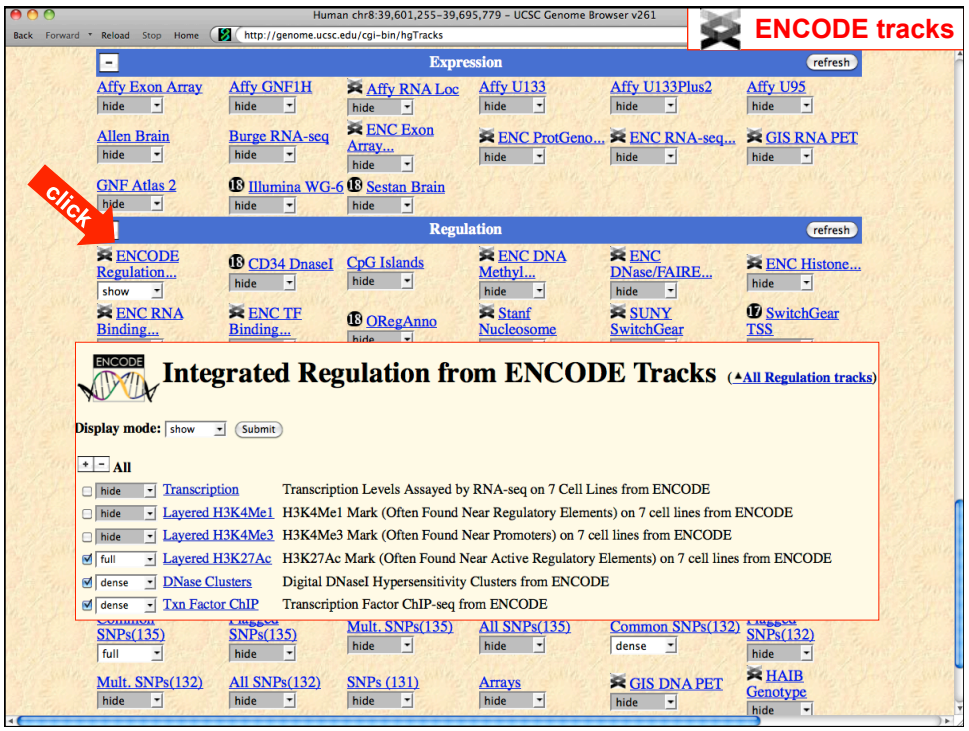


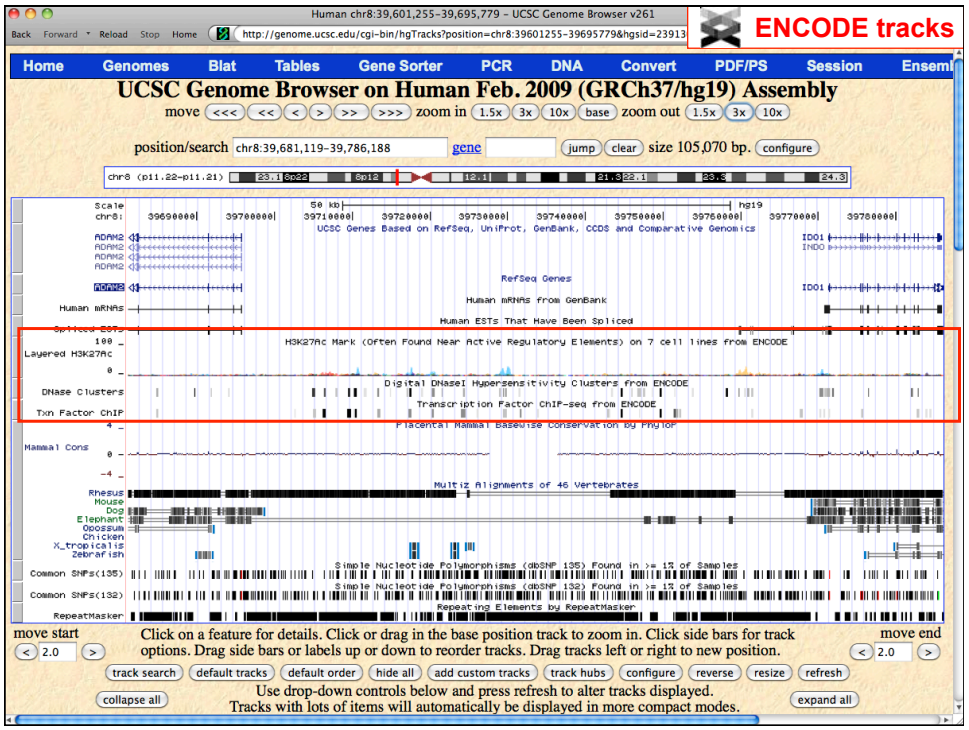

UCSC Common SNP Track
Human chr8:39,601,255-39,695,779 - UCSC Genome Bro
http://genome.ucsc.edu/cgi-bin/hgTracks
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics
RefSeq Genes
Human mRNAs from GenBank
Human ESTs That Have Been Spliced
H3K27Ac Mark (Often Found Near Active Regulatory Elements) on 7 cell lines from ENCODE
Digital DNaseI Hypersensitivity Clusters from ENCODE
Transcription Factor ChIP-seq from ENCODE
Placental Mammal Basewise Conservation by PhyloP
Multiz Alignments of 46 Vertebrates
Simple Nucleotide Polymorphisms (dbSNP 135) Found in >= 1% of Samples

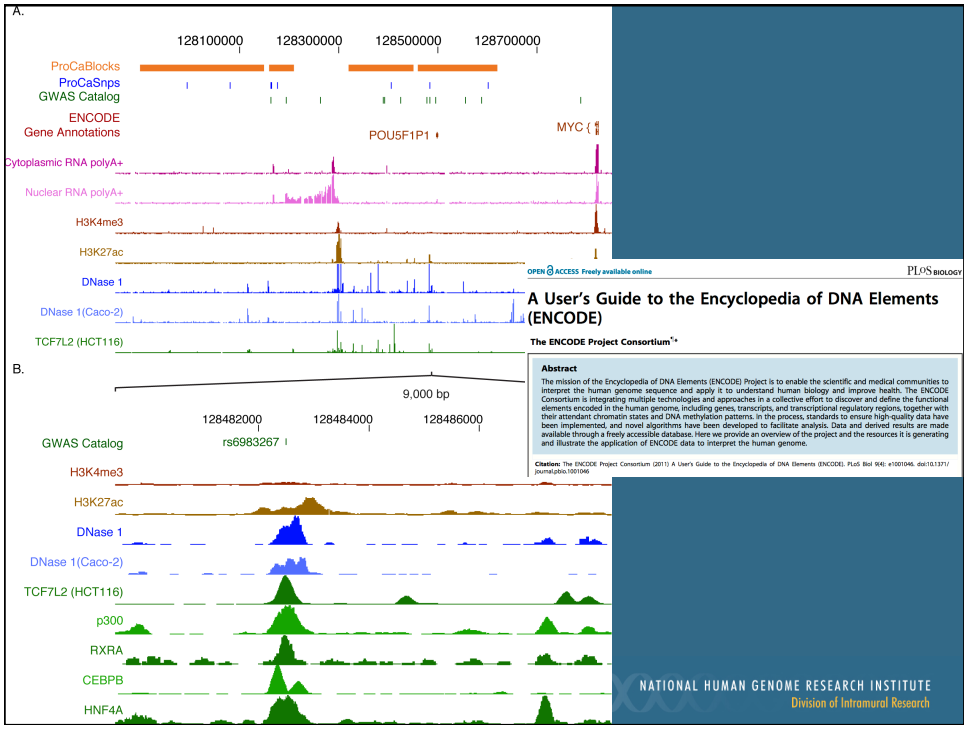

UCSC SNP Track details
Common SNPs(135) Track Settings
Home Genomes Genome Browser Blat Tables Gene Sorter PCR Session FAQ Help
Common SNPs(135) Track Settings
Simple Nucleotide Polymorphisms (dbSNP 135) Found in >= 1% of Samples
All Variation and Repeats tracks
Display mode: pack Submit
Include Chimp state and observed human alleles in name:
(If enabled, chimp allele is displayed first, then '>', then human alleles).
Use Gene Tracks for Functional Annotation
Filtering Options
Coloring Options
SNP Feature for Color Specification: Function Set defaults
The selected "Feature for Color Specification" above has the selection of colors below for each attribute. Only the color options for the feature selected above will be used to color items; color options for other features will not be shown. If a SNP has more than one of these attributes, the stronger color will override the weaker color. The order of colors, from strongest to weakest, is red, green, blue, gray, and black.
Unknown black Locus black Coding - Synonymous green Coding - Non-Synonymous red
Untranslated black Intron black Splice Site black
View table schema
Data last updated: 2011-11-14
Description
This track contains information about a subset of the single nucleotide polymorphisms and small insertions and deletions (indels) — collectively Simple Nucleotide Polymorphisms — from dbSNP build 135, available from ftp.ncbi.nih.gov/snp. Only SNPs that have a minor allele frequency of at least 1% and are mapped to a single location in the reference genome assembly are included in this subset. Frequency


UCSC SNP Track
Human chr8:39,601,255-39,695,779 - UCSC Genome Browser v261
UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly
Green: coding-synonymous
Red: coding-nonsynonymous
Black: other SNPs
synonymous SNP
non-synonymous SNP

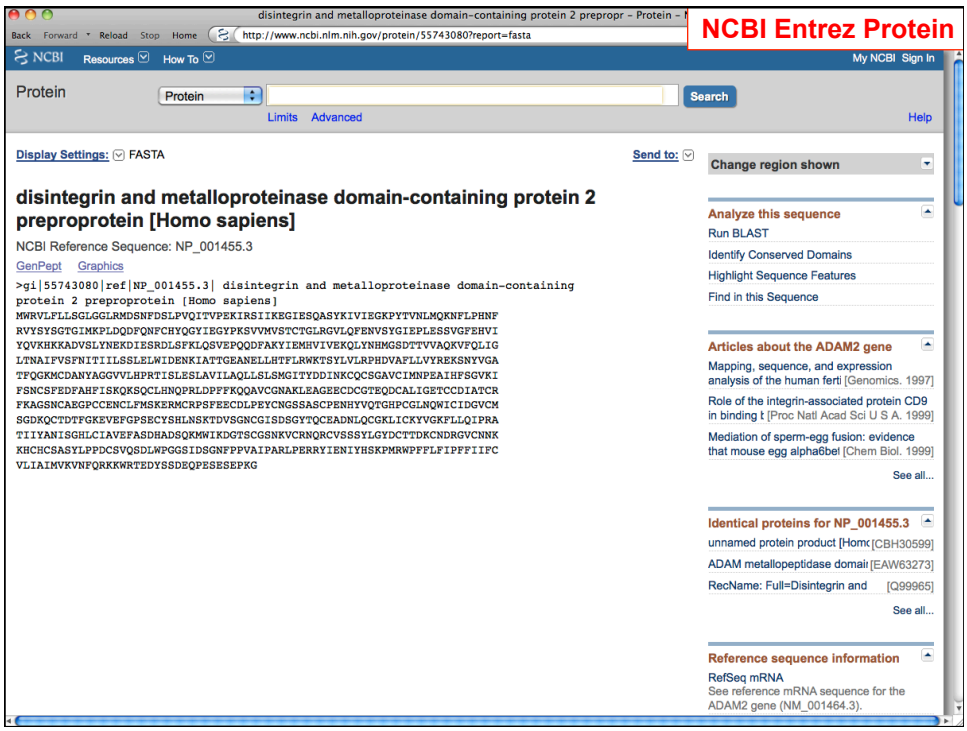

UCSC
ENCODE tracks
http://genome.ucsc.edu
NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research

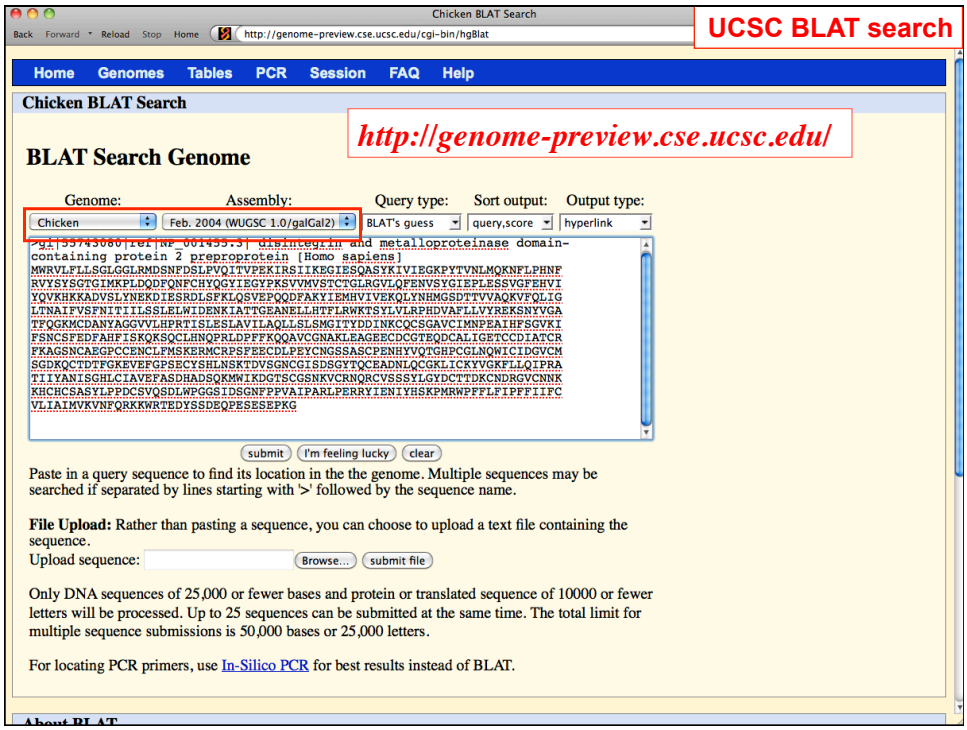

ENCODE tracks
click
Expression
Regulation
Integrated Regulation from ENCODE Tracks
Transcription
Layered H3K4Me1
Layered H3K4Me3
Layered H3K27Ac
DNase Clusters
Txn Factor ChIP

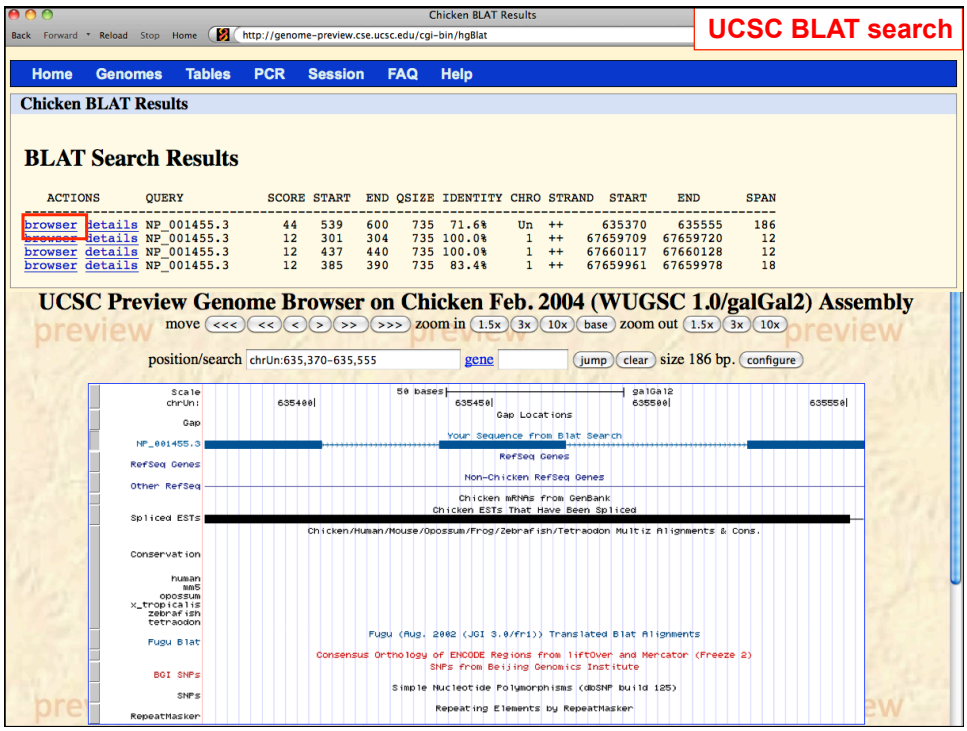

ENCODE tracks
UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly
H3K27Ac Mark (Often Found Near Active Regulatory Elements) on 7 cell lines from ENCODE
Digital DNaseI Hypersensitivity Clusters from ENCODE
Transcription Factor ChIP-seq from ENCODE

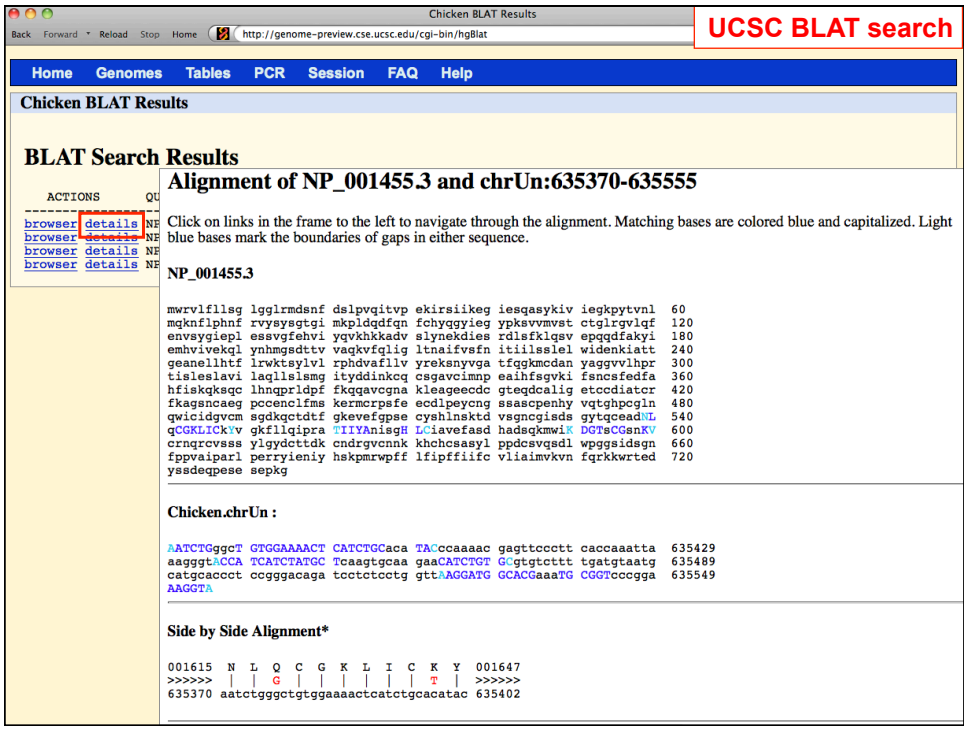

A.
128100000
128300000
128500000
128700000
ProCaBlocks
ProCaSnps
GWAS Catalog
ENCODE
Gene Annotations
POU5F1P1
MYC
Cytoplasmic RNA polyA+
Nuclear RNA polyA+
H3K4me3
H3K27ac
DNase 1
DNase 1(Caco-2)
TCF7L2 (HCT116)
B.
9,000 bp
128482000
128484000
128486000
GWAS Catalog
rs6983267
H3K4me3
H3K27ac
DNase 1
DNase 1(Caco-2)
TCF7L2 (HCT116)
p300
RXRA
CEBPB
HNF4A
OPEN ACCESS Freely available online
PLoS BIOLOGY
A User's Guide to the Encyclopedia of DNA Elements (ENCODE)
The ENCODE Project Consortium*
Abstract
The mission of the Encyclopedia of DNA Elements (ENCODE) Project is to enable the scientific and medical communities to interpret the human genome sequence and apply it to understand human biology and improve health. The ENCODE Consortium is integrating multiple technologies and approaches in a collective effort to discover and define the functional elements encoded in the human genome, including genes, transcripts, and transcriptional regulatory regions, together with their attendant chromatin states and DNA methylation patterns. In the process, standards to ensure high-quality data have been implemented, and novel algorithms have been developed to facilitate analysis. Data and derived results are made available through a freely accessible database. Here we provide an overview of the project and the resources it is generating and illustrate the application of ENCODE data to interpret the human genome.
Citation: The ENCODE Project Consortium (2011) A User's Guide to the Encyclopedia of DNA Elements (ENCODE). PLoS Biol 9(4): e1001046. doi:10.1371/journal.pbio.1001046
NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research


UCSC
Find a chicken homolog of a human protein
http://genome.ucsc.edu
NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research


NCBI Entrez Protein
disintegrin and metalloproteinase domain-containing protein 2 preprotein [Homo sapiens]
NCBI Reference Sequence: NP_001455.3
>gi|55743080|ref|NP_001455.3| disintegrin and metalloproteinase domain-containing protein 2 preproprotein [Homo sapiens]
MWRVLFLLSGLGGLRMDSNFDSLPVQITVPEKIRSIIKEGIESQASYKIVIEGKPYTVNLMQKNFLPHNF
RVYSYSGTGIMKPLDQDFQNFCHYQGYIEGYPKSVVMVSTCTGLRGVLQFENVSYGIEPLESSVGFEHVI
YQVKHKKADVSLYNEKDIESRDLSFKLQSVEPQQDFAKYIEMHVIVEKQLYNHMGSDTTVVAQKVFQLIG
LTNAIFVSFNITIILSSLELWIDENKIATTGEANELLHTFLRWKTSYLVLRPHDVAFLLVYREKSNYVGA
TFQGKMCDANYAGGVVLHPRTISLESLAVILAQLLSLSMGITYDDINKCQCSGAVCIMNPEAIHFSGVKI
FSNCSFEDFAHFISKQKSQCLHNQPRLDPFFKQQAVCGNAKLEAGEECDCGTEQDCALIGETCCDIATCR
FKAGSNCAEGPCCENCLFMSKERMCRPSFEECDLPEYCNGSSASCPENHYVQTGHPCGLNQWICIDGVCM
SGDKQCTDTFGKEVEFGPSECYSHLNSKTDVSGNCGISDSGYTQCEADNLQCGKLICKYVGKFLLQIPRA
TIIYANISGHLCIAVEFASDHADSQKMWIKDGTSCGSNKVCRNQRCVSSSYLGYDCTTDKCNDRGVCNNK
KHCHCSASYLPPDCSVQSDLWPGGSIDSGNFPPVAIPARLPERRYIENIYHSKPMRWPPFLFIPPFIIFC
VLIAIMVKVNFQRKKWRTEDYSSDEQPESESEPKG
Display Settings: FASTA
Send to:
Change region shown
Analyze this sequence
Run BLAST
Identify Conserved Domains
Highlight Sequence Features
Find in this Sequence
Articles about the ADAM2 gene
Mapping, sequence, and expression analysis of the human ferti [Genomics. 1997]
Role of the integrin-associated protein CD9 in binding t [Proc Natl Acad Sci U S A. 1999]
Mediation of sperm-egg fusion: evidence that mouse egg alpha6be! [Chem Biol. 1999]
See all...
Identical proteins for NP_001455.3
unnamed protein product [Homc [CBH30599]
ADAM metallopeptidase domai [EAW63273]
RecName: Full=Disintegrin and [Q99965]
See all...
Reference sequence information
RefSeq mRNA
See reference mRNA sequence for the ADAM2 gene (NM_001464.3).


UCSC BLAT search
Chicken BLAT Search
http://genome-preview.cse.ucsc.edu/cgi-bin/hgBlat
Home Genomes Tables PCR Session FAQ Help
Chicken BLAT Search
http://genome-preview.cse.ucsc.edu/
BLAT Search Genome
Genome: Assembly: Query type: Sort output: Output type:
Chicken Feb. 2004 (WUGSC 1.0/galGal2) BLAT's guess query,score hyperlink
submit I'm feeling lucky clear
Paste in a query sequence to find its location in the the genome. Multiple sequences may be searched if separated by lines starting with '>' followed by the sequence name.
File Upload: Rather than pasting a sequence, you can choose to upload a text file containing the sequence.
Upload sequence: Browse... submit file
Only DNA sequences of 25,000 or fewer bases and protein or translated sequence of 10000 or fewer letters will be processed. Up to 25 sequences can be submitted at the same time. The total limit for multiple sequence submissions is 50,000 bases or 25,000 letters.
For locating PCR primers, use In-Silico PCR for best results instead of BLAT.


Chicken BLAT Results
http://genome-preview.cse.ucsc.edu/cgi-bin/hgBlat
UCSC BLAT search
Home Genomes Tables PCR Session FAQ Help
Chicken BLAT Results
BLAT Search Results
ACTIONS QUERY SCORE START END QSIZE IDENTITY CHRO STRAND START END SPAN
browser details NP_001455.3 44 539 600 735 71.6% Un ++ 635370 635555 186
browser details NP_001455.3 12 301 304 735 100.0% 1 ++ 67659709 67659720 12
browser details NP_001455.3 12 437 440 735 100.0% 1 ++ 67660117 67660128 12
browser details NP_001455.3 12 385 390 735 83.4% 1 ++ 67659961 67659978 18
UCSC Preview Genome Browser on Chicken Feb. 2004 (WUGSC 1.0/galGal2) Assembly
move zoom in zoom out
position/search chrUn:635,370-635,555 gene jump clear size 186 bp. configure
Gap Locations
Your Sequence from Blat Search
RefSeq Genes
Non-Chicken RefSeq Genes
Chicken mRNAs from GenBank
Chicken ESTs That Have Been Spliced
Chicken/Human/Mouse/Opossum/Frog/Zebrafish/Tetraodon Multiz Alignments & Cons.
Fugu (Aug. 2002 (JGI 3.0/fr1)) Translated Blat Alignments
Consensus Orthology of ENCODE Regions from liftOver and Mercator (Freeze 2)
SNPs from Beijing Genomics Institute
Simple Nucleotide Polymorphisms (dbSNP build 125)
Repeating Elements by RepeatMasker


Chicken BLAT Results
http://genome-preview.cse.ucsc.edu/cgi-bin/hgBlat
UCSC BLAT search
Home Genomes Tables PCR Session FAQ Help
Chicken BLAT Results
BLAT Search Results
Alignment of NP_001455.3 and chrUn:635370-635555
Click on links in the frame to the left to navigate through the alignment. Matching bases are colored blue and capitalized. Light blue bases mark the boundaries of gaps in either sequence.
NP_001455.3
mwrvlfllsg lgglrmdsnf dslpvqitvp ekirsiikeg iesqasykiv iegkpytvnl 60
mqknflphnf rvysysgtgi mkpldqdfqn fchyggyieg ypksvvmvst ctglrgvlqf 120
envsygiepl essvgfehvi yqvkhkkadv slynekdies rdlsfklqsv epqqdfakyi 180
emhvivekql ynhmgsdttv vaqkvfqlig ltnaifvsfn itiilssslel widenkiatt 240
geanellhtf lrwktsylvl rphdvafllv yreksnyvga tfqgkmcdan yagqvvlhpr 300
tisleslavi laqllslsmg ityddinkcq csgavcimnp eaihfsgvki fsncsfedfa 360
hfiskqksqc lhnqprldpf fkqqavcgna kleageecdc gteqdcalig etccdiatcr 420
fkagsncaeg pccenclfms kermcrpsfe ecdlpeycng ssascpenhy vqtghpcgln 480
qwicidgvcm sgdkqctdtf gkevefgpse cyshlnsktd vsgncgisds gytqceadNL 540
qCGKLICkYv gkfllqipra TIIYAnisgH LCiavefasd hadsqkmwiK DGTsCGsnKV 600
crnqrcvsss ylgydcttdk cndrgvcnnk khchcsasyl ppdcsvqsdl wpggsidsgn 660
fppvaiparl perryieniy hskpmrwpff lfipffiifc vliaimvkvn fqrkkwrted 720
yssdeqpese sepkg
Chicken.chrUn :
AATCTGggcT GTGGAAAACT CATCTGCaca TACccaaaac gagttccctt caccaaatta 635429
aagggtACCA TCATCTATGC Tcaagtgcaa gaaCATCTGT GCgtgtcttt tgatgtaatg 635489
catgcaccct ccgggacaga tcctctcctg gttAAGGATG GCACGaaaTG CGGTcccgga 635549
AAGGTA
Side by Side Alignment*
001615 N L Q C G K L I C K Y 001647
635370 aatctgggctgtggaaaactcatctgcacatac 635402

UCSC

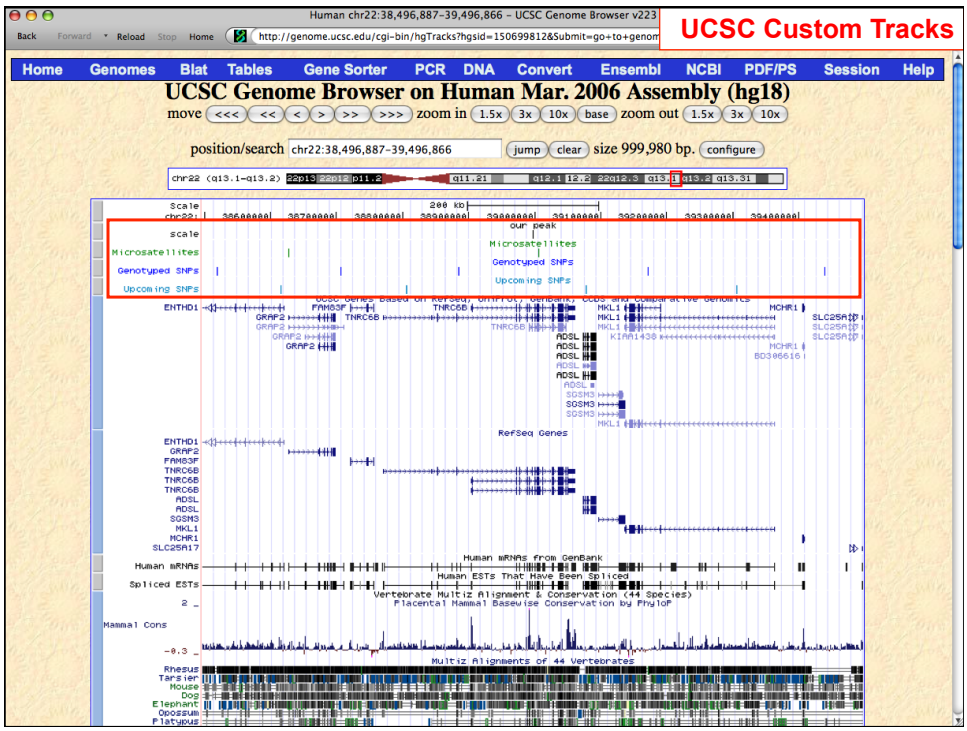
Add your own custom tracks

http://genome.ucsc.edu

UCSC custom track format

```
browser position chr22:38496887-39496866
browser hide cytoBand
browser hide stsMap
browser hide gap
browser hide clonePos
browser full refGene
browser dense mrna
track name="scale" description="our peak"
chr22 38996887 38996888 peak
track name="Microsatellites" description="Microsatellites" color=0,128,0
chr22 38627059 38627060 D22S276
chr22 39005417 39005418 D22S307
track name="Genotyped SNPs" description="Genotyped SNPs" color=0,0,255
chr22 38518342 38518343 ss146131
chr22 38705963 38705964 ss2941443
chr22 38884157 38884158 ss141110
chr22 39171390 39171391 ss22916
chr22 39438769 39438770 ss1479794
track name="Upcoming SNPs" description="Upcoming SNPs" color=0,128,192
chr22 38615712 38615713 ss86855
chr22 38804838 38804839 ss85533
chr22 39077895 39077896 ss141190
chr22 39305065 39305066 ss137027
```

http://research.nhgri.nih.gov/teaching/custom_tracks.shtml



Types of UCSC custom tracks

- Upload annotation data from your computer
  - Only viewable on the machine from which they were uploaded
  - Discarded after 48 hours
- Post annotation data to your Web site
  - URL, with link to Genome Browser, can be shared with anyone
  - Never discarded
- Create a Session to configure your browser with specific track combinations, including custom tracks
  - Can be shared or non-shared
  - Session persists for 4 months; custom tracks for 48 hours
- Contribute your tracks to the UCSC Genome Browser

http://genome.ucsc.edu/goldenPath/help/hgTracksHelp.html#CustomTracks
http://genome.ucsc.edu/goldenPath/help/hgSessionHelp.html

NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research

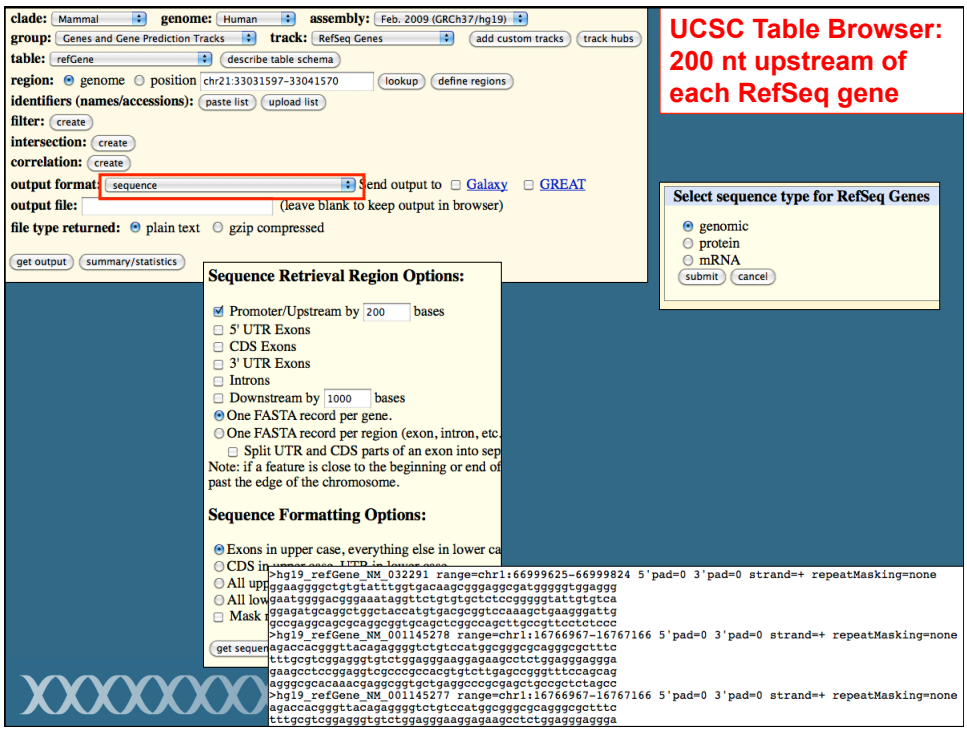
UCSC

Table Browser

http://genome.ucsc.edu

UCSC Table Browser

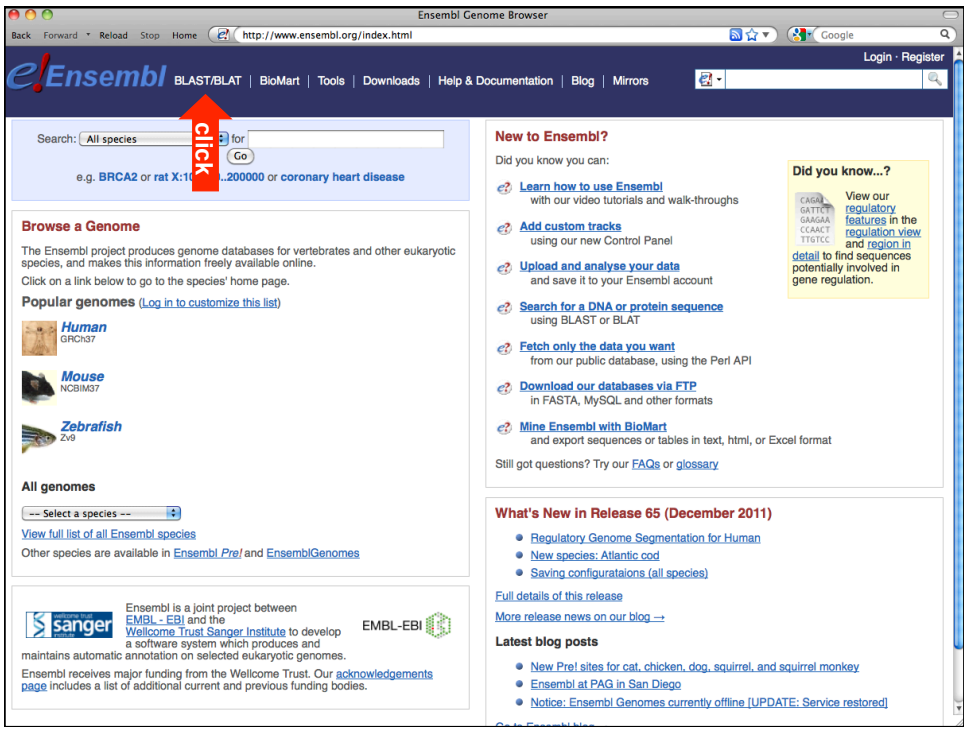
- Download track in text format
- Retrieve DNA sequence covered by a track
  - Get sequence 1 Kb upstream of each RefSeq gene
- Calculate intersections between tracks and view in the Genome Browser.
  - List all SNPs in a gene
- Filter track data based on certain criteria
  - Show all RefSeq genes that contain only one exon

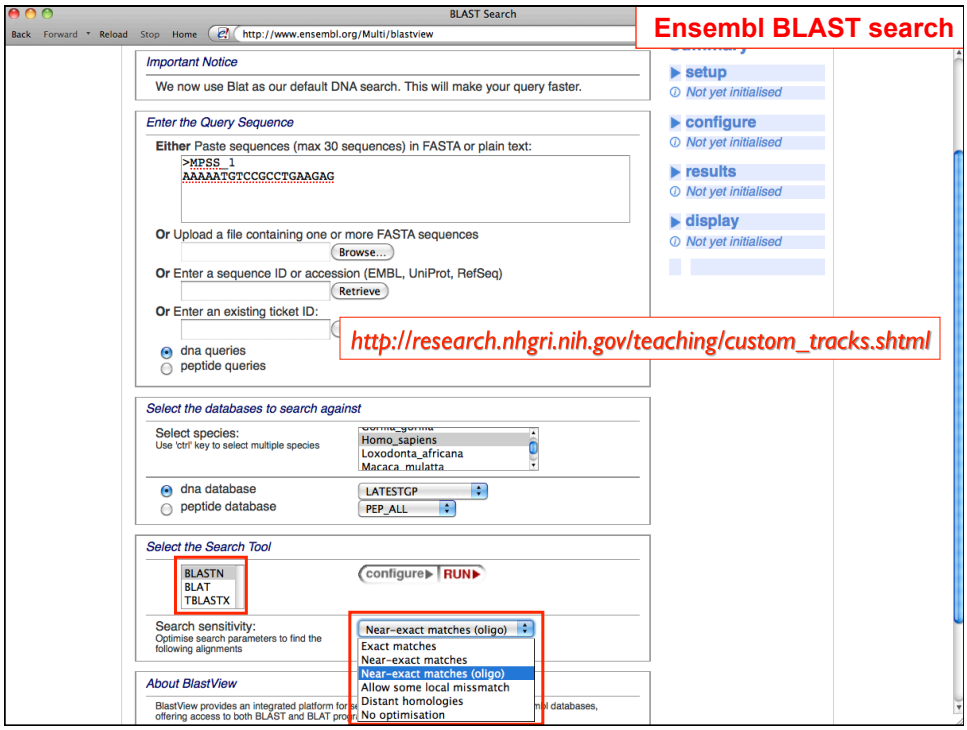


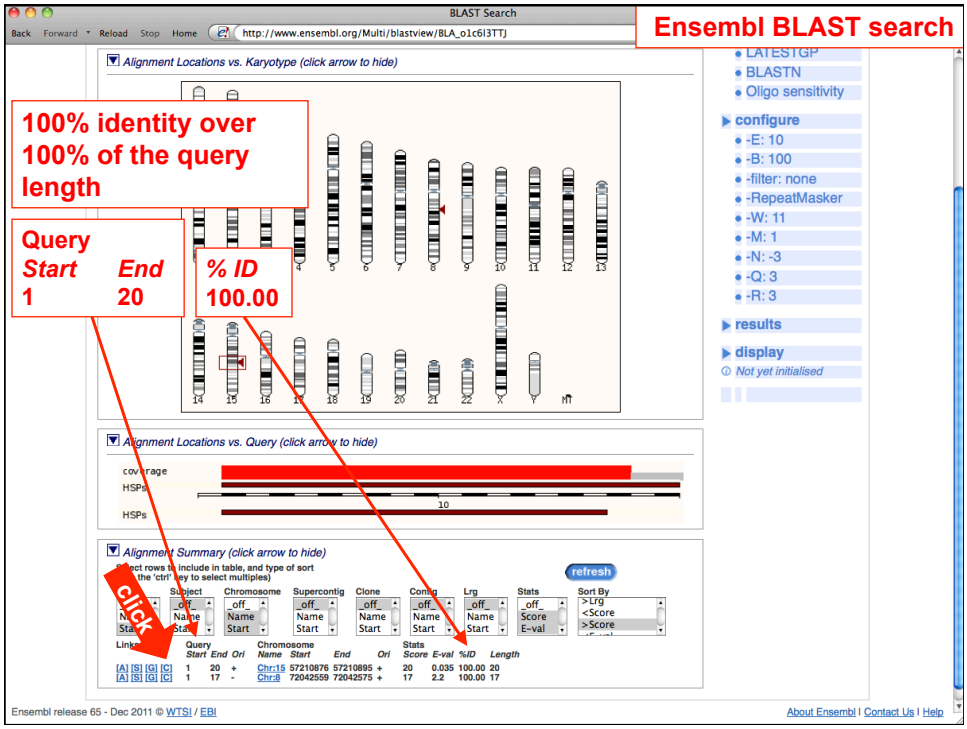
Ensembl

Identify genes that overlap with an oligo tag

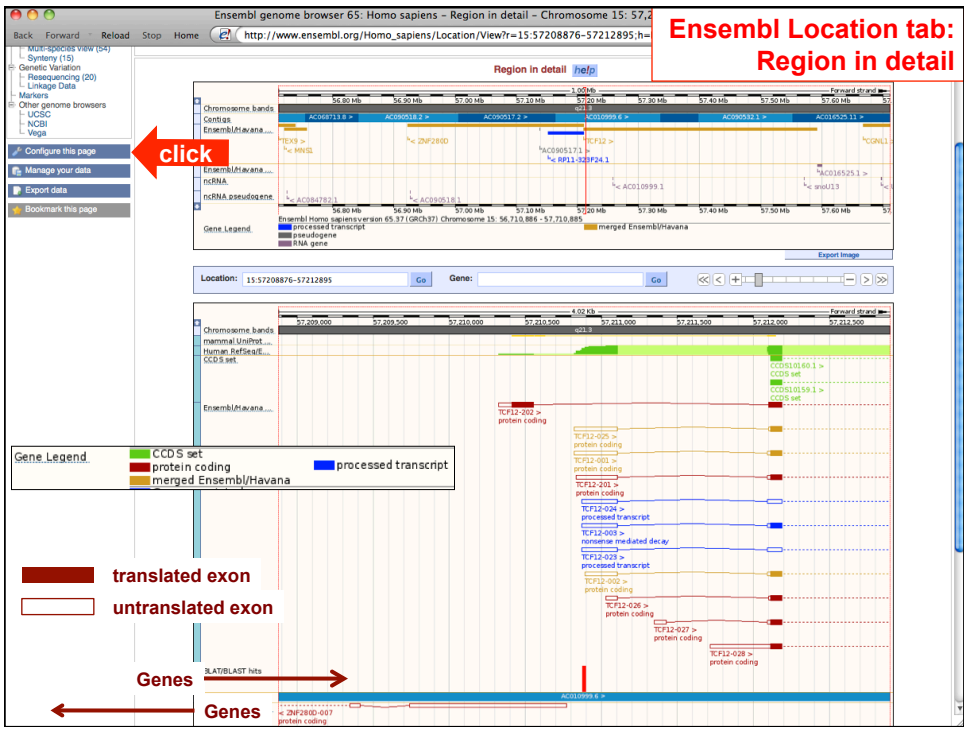
http://www.ensembl.org

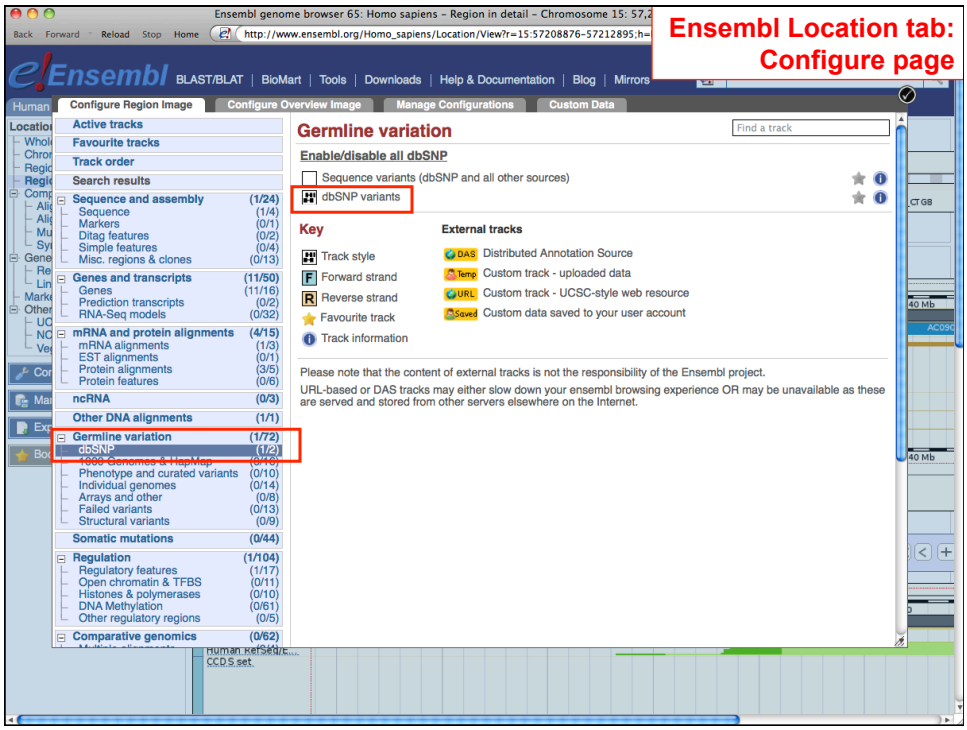

Ensembl Genome Browser
http://www.ensembl.org/index.html
Ensembl
BLAST/BLAT | BioMart | Tools | Downloads | Help & Documentation | Blog | Mirrors
Login · Register
click
Search: All species for Go
e.g. BRCA2 or rat X:10..200000 or coronary heart disease
Browse a Genome
The Ensembl project produces genome databases for vertebrates and other eukaryotic species, and makes this information freely available online.
Click on a link below to go to the species' home page.
Popular genomes (Log in to customize this list)
Human GRCh37
Mouse NCBIM37
Zebrafish Zv9
All genomes
-- Select a species --
View full list of all Ensembl species
Other species are available in Ensembl Pre! and EnsemblGenomes
Ensembl is a joint project between EMBL - EBI and the Wellcome Trust Sanger Institute to develop a software system which produces and maintains automatic annotation on selected eukaryotic genomes.
EMBL-EBI
Ensembl receives major funding from the Wellcome Trust. Our acknowledgements page includes a list of additional current and previous funding bodies.
New to Ensembl?
Did you know you can:
Learn how to use Ensembl with our video tutorials and walk-throughs
Add custom tracks using our new Control Panel
Upload and analyse your data and save it to your Ensembl account
Search for a DNA or protein sequence using BLAST or BLAT
Fetch only the data you want from our public database, using the Perl API
Download our databases via FTP in FASTA, MySQL and other formats
Mine Ensembl with BioMart and export sequences or tables in text, html, or Excel format
Still got questions? Try our FAQs or glossary
Did you know...?
View our regulatory features in the regulation view and region in detail to find sequences potentially involved in gene regulation.
What's New in Release 65 (December 2011)
Regulatory Genome Segmentation for Human
New species: Atlantic cod
Saving configurataions (all species)
Full details of this release
More release news on our blog →
Latest blog posts
New Pre! sites for cat, chicken, dog, squirrel, and squirrel monkey
Ensembl at PAG in San Diego
Notice: Ensembl Genomes currently offline [UPDATE: Service restored]


BLAST Search
http://www.ensembl.org/Multi/blastview
Ensembl BLAST search
Important Notice
We now use Blat as our default DNA search. This will make your query faster.
Enter the Query Sequence
Either Paste sequences (max 30 sequences) in FASTA or plain text:
>MPSS_1
AAAAATGTCCGCCTGAAGAG
Or Upload a file containing one or more FASTA sequences
Browse...
Or Enter a sequence ID or accession (EMBL, UniProt, RefSeq)
Retrieve
Or Enter an existing ticket ID:
dna queries
peptide queries
http://research.nhgri.nih.gov/teaching/custom_tracks.shtml
setup
Not yet initialised
configure
Not yet initialised
results
Not yet initialised
display
Not yet initialised
Select the databases to search against
Select species:
Use 'ctrl' key to select multiple species
Homo_sapiens
Loxodonta_africana
Macaca_mulatta
dna database
peptide database
LATESTGP
PEP_ALL
Select the Search Tool
BLASTN
BLAT
TBLASTX
configure
RUN
Search sensitivity:
Optimise search parameters to find the following alignments
Near-exact matches (oligo)
Exact matches
Near-exact matches
Near-exact matches (oligo)
Allow some local missmatch
Distant homologies
No optimisation
About BlastView
BlastView provides an integrated platform for
databases,
offering access to both BLAST and BLAT

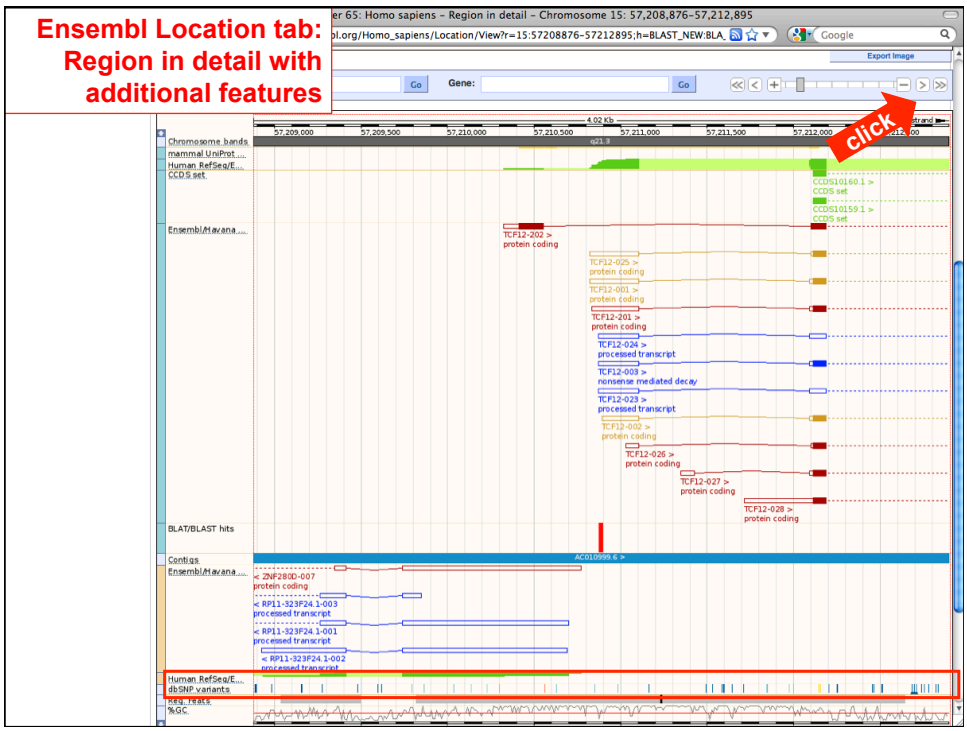

Ensembl BLAST search
100% identity over 100% of the query length
Query Start 1
End 20
% ID 100.00
click
BLAST Search
Alignment Locations vs. Karyotype (click arrow to hide)
Alignment Locations vs. Query (click arrow to hide)
Alignment Summary (click arrow to hide)
refresh
configure
-E: 10
-B: 100
-filter: none
-RepeatMasker
-W: 11
-M: 1
-N: -3
-Q: 3
-R: 3
results
display
Not yet initialised
Ensembl release 65 - Dec 2011 © WTSI / EBI
About Ensembl | Contact Us | Help


Ensembl
Location tab
http://www.ensembl.org
NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research

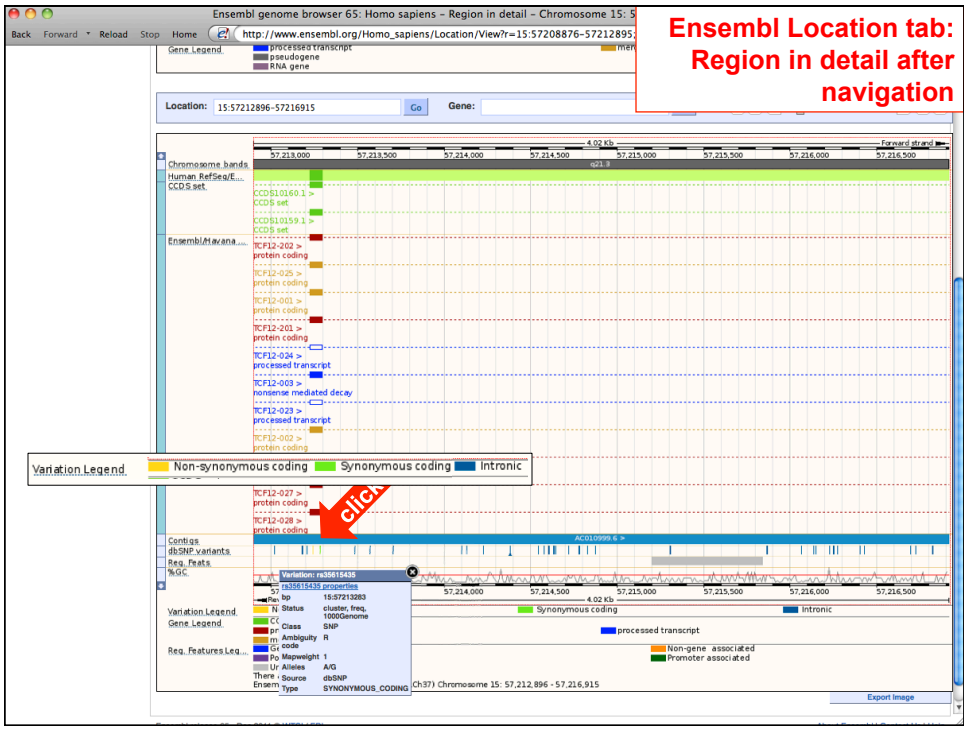

Ensembl Location tab:
Region in detail
click
Region in detail
Gene Legend
CCDS set
protein coding
merged Ensembl/Havana
processed transcript
translated exon
untranslated exon
Genes
Genes

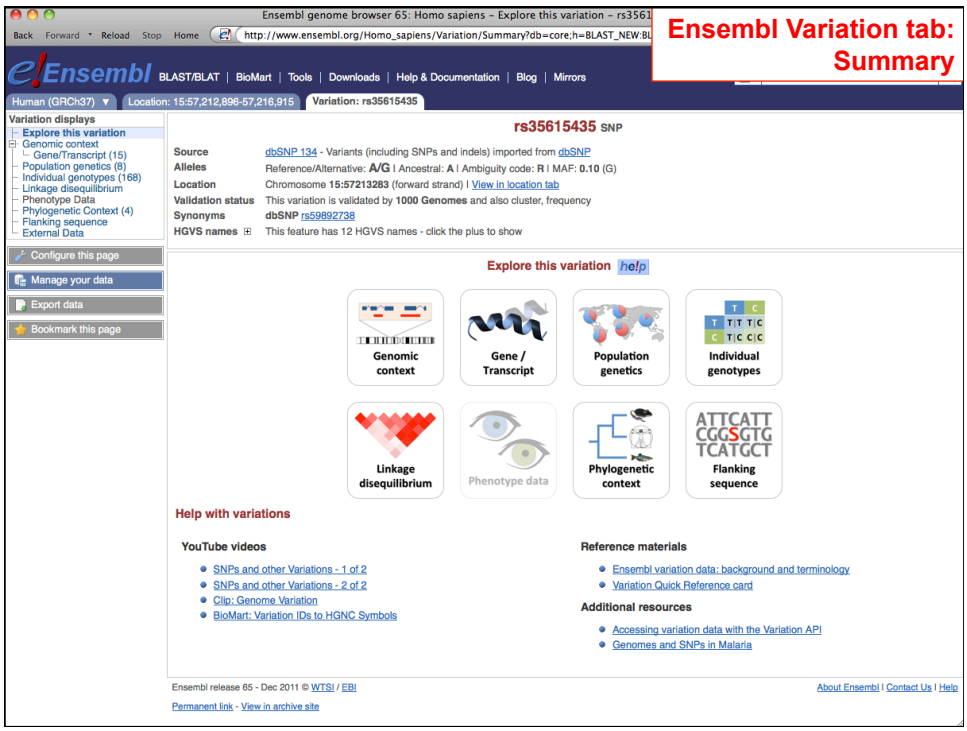

Ensembl Location tab:
Configure page
Configure Region Image
Configure Overview Image
Manage Configurations
Custom Data
Germline variation
Enable/disable all dbSNP
Sequence variants (dbSNP and all other sources)
dbSNP variants
Key
External tracks
Please note that the content of external tracks is not the responsibility of the Ensembl project.

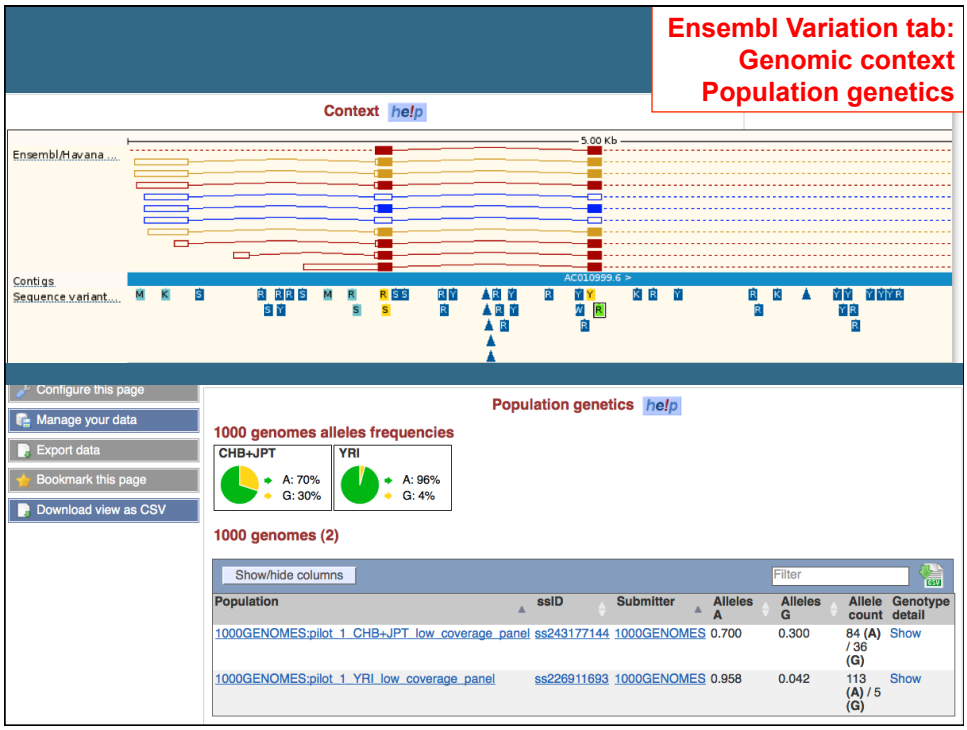
Ensembl Location tab: Region in detail with additional features



Ensembl Location tab: Region in detail after navigation




Ensembl Variation tab:
Summary
Ensembl genome browser 65: Homo sapiens - Explore this variation - rs3561
http://www.ensembl.org/Homo_sapiens/Variation/Summary?db=core;h=BLAST_NEW:BL
Ensembl BLAST/BLAT | BioMart | Tools | Downloads | Help & Documentation | Blog | Mirrors
Human (GRCh37) | Location: 15:57,212,896-57,216,915 | Variation: rs35615435
Variation displays
Explore this variation
Genomic context
Gene/Transcript (15)
Population genetics (8)
Individual genotypes (168)
Linkage disequilibrium
Phenotype Data
Phylogenetic Context (4)
Flanking sequence
External Data
Configure this page
Manage your data
Export data
Bookmark this page
rs35615435 SNP
Source dbSNP 134 - Variants (including SNPs and indels) imported from dbSNP
Alleles Reference/Alternative: A/G | Ancestral: A | Ambiguity code: R | MAF: 0.10 (G)
Location Chromosome 15:57213283 (forward strand) | View in location tab
Validation status This variation is validated by 1000 Genomes and also cluster, frequency
Synonyms dbSNP rs59892738
HGVS names This feature has 12 HGVS names - click the plus to show
Explore this variation help
Genomic context
Gene / Transcript
Population genetics
Individual genotypes
Linkage disequilibrium
Phenotype data
Phylogenetic context
Flanking sequence
Help with variations
YouTube videos
SNPs and other Variations - 1 of 2
SNPs and other Variations - 2 of 2
Clip: Genome Variation
BioMart: Variation IDs to HGNC Symbols
Reference materials
Ensembl variation data: background and terminology
Variation Quick Reference card
Additional resources
Accessing variation data with the Variation API
Genomes and SNPs in Malaria
Ensembl release 65 - Dec 2011 © WTSI / EBI
About Ensembl | Contact Us | Help
Permanent link - View in archive site

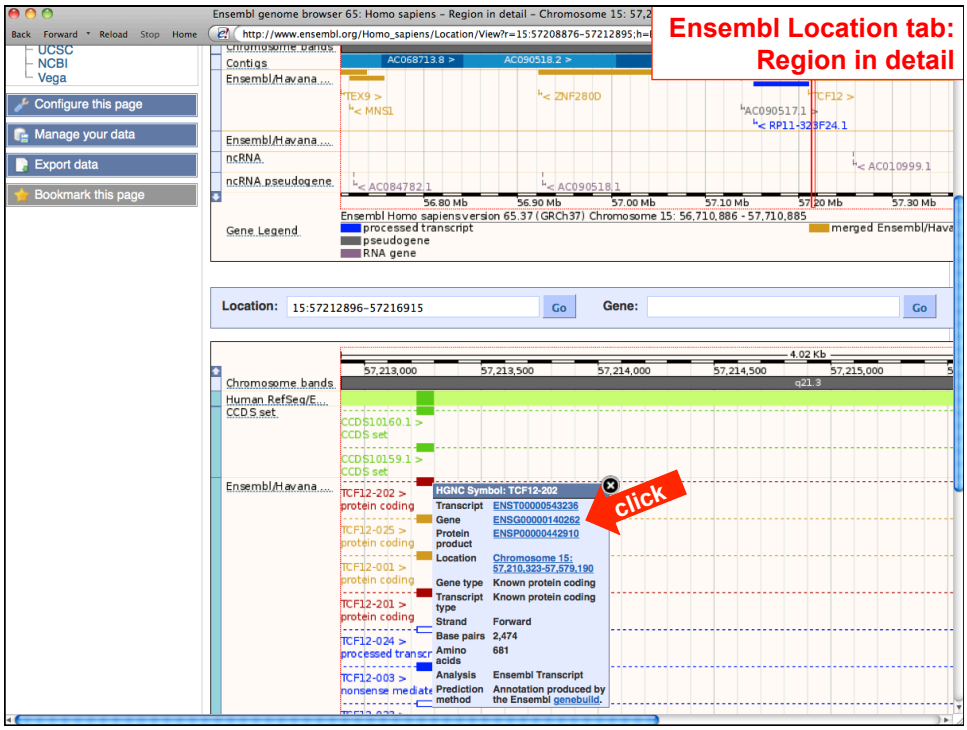

Ensembl Variation tab:
Genomic context
Population genetics
Context help
5.00 Kb
Ensembl/Havana
Contigs
Sequence variant
AC010999.6 >
Configure this page
Manage your data
Export data
Bookmark this page
Download view as CSV
Population genetics help
1000 genomes alleles frequencies
CHB+JPT
A: 70%
G: 30%
YRI
A: 96%
G: 4%
1000 genomes (2)
Show/hide columns
Filter
Population
ssID
Submitter
Alleles A
Alleles G
Allele count
Genotype detail
1000GENOMES:pilot_1_CHB+JPT_low_coverage_panel
ss243177144
1000GENOMES
0.700
0.300
84 (A) / 36 (G)
Show
1000GENOMES:pilot_1_YRI_low_coverage_panel
ss226911693
1000GENOMES
0.958
0.042
113 (A) / 5 (G)
Show

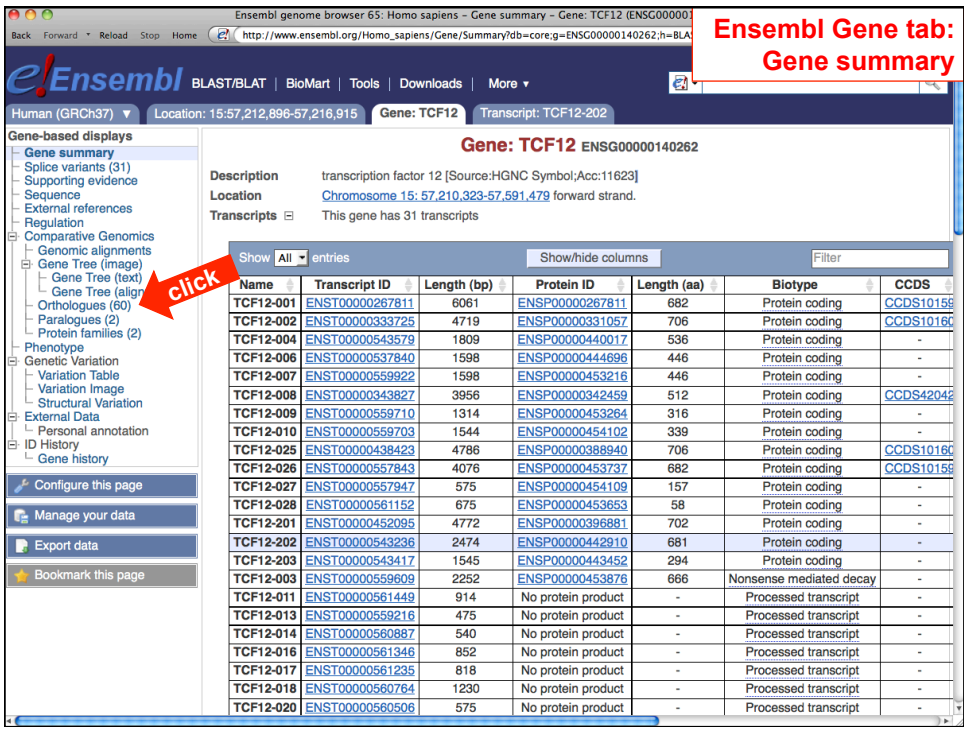
Ensembl

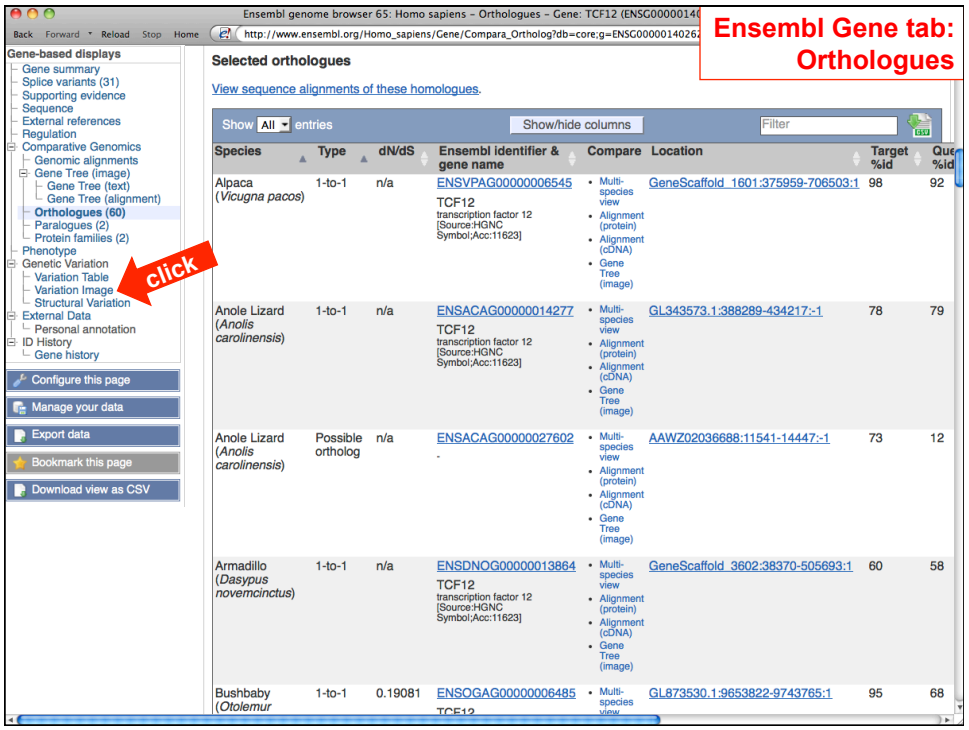
Gene tab

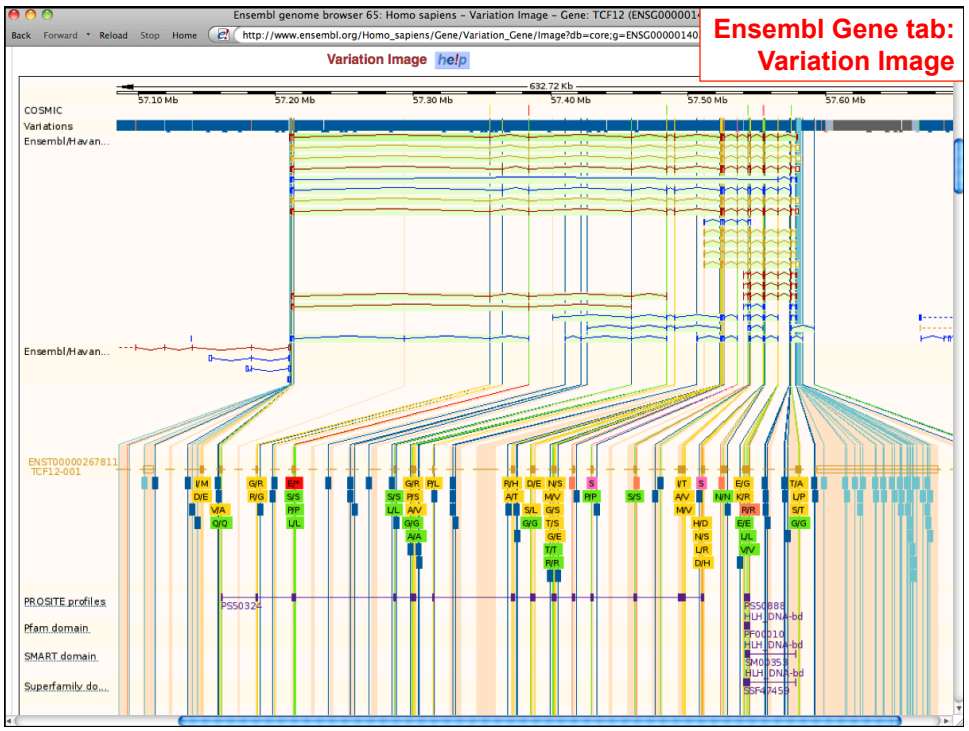
http://www.ensembl.org

NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research




Ensembl Gene tab:
Gene summary
click
Gene: TCF12 ENSG00000140262
Description transcription factor 12 [Source:HGNC Symbol;Acc:11623]
Location Chromosome 15: 57,210,323-57,591,479 forward strand.
Transcripts This gene has 31 transcripts


Ensembl Gene tab:
Orthologues
click
Selected orthologues
View sequence alignments of these homologues.

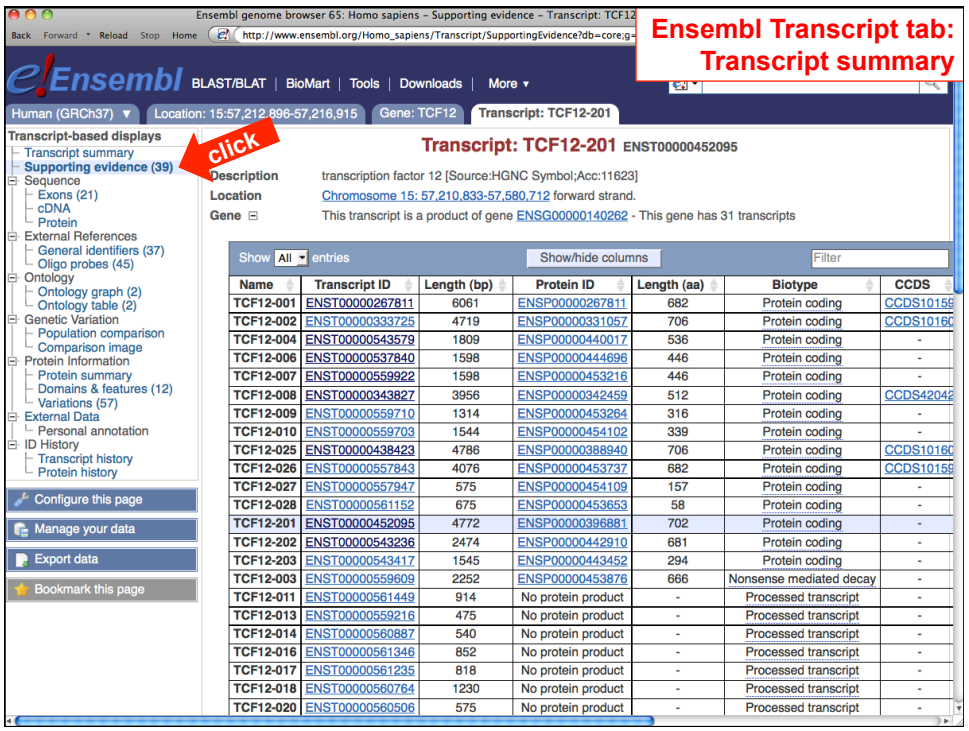


Ensembl

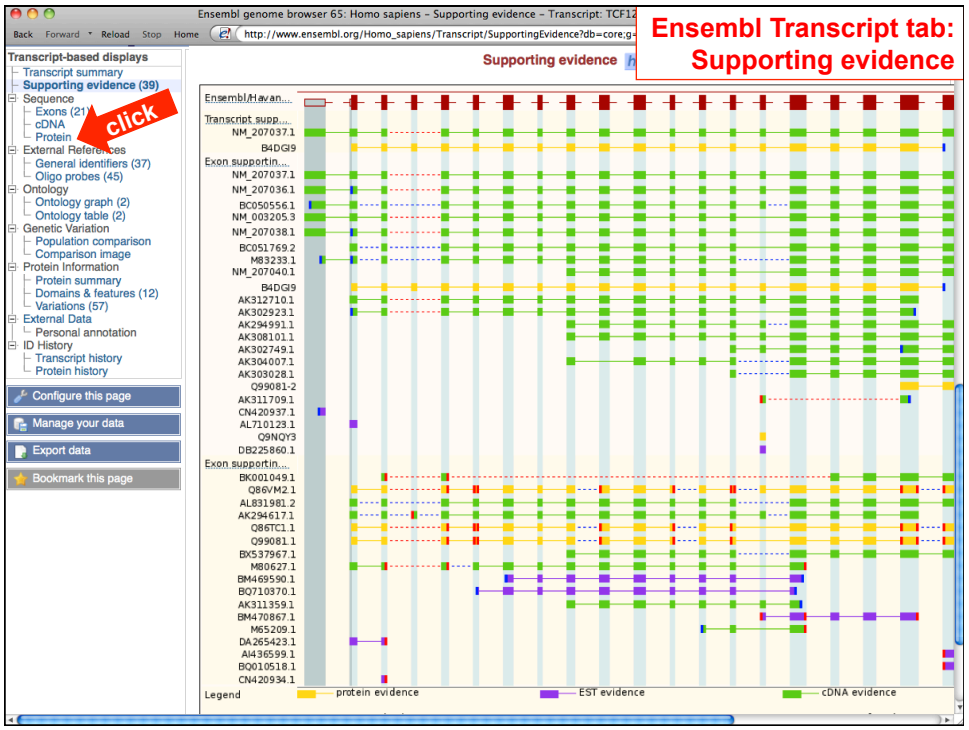
Transcript tab

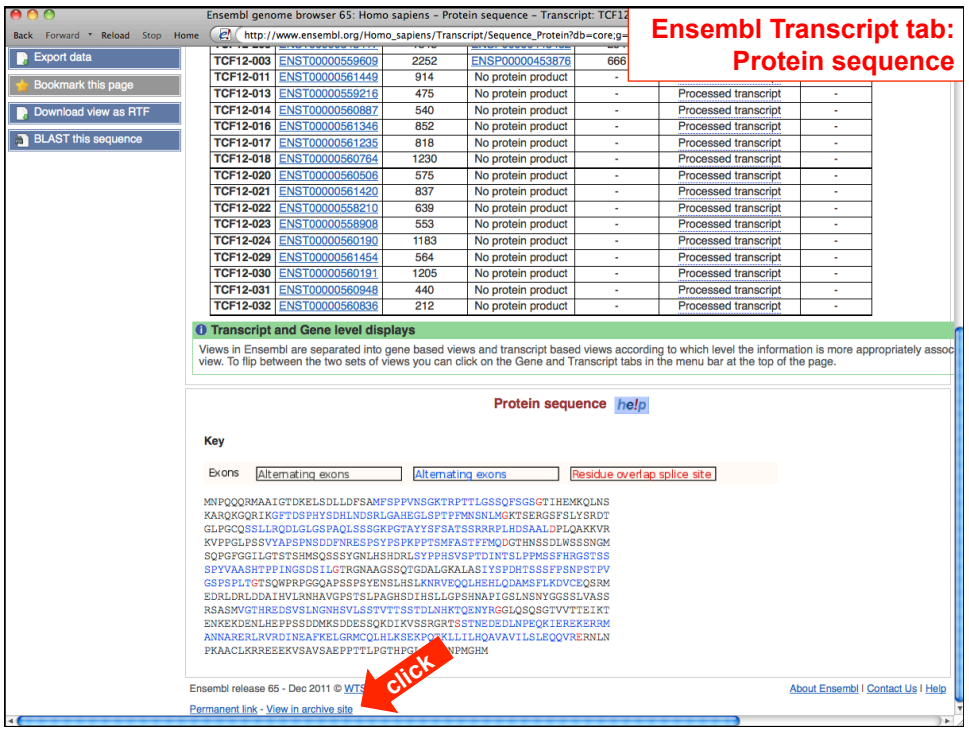
http://www.ensembl.org

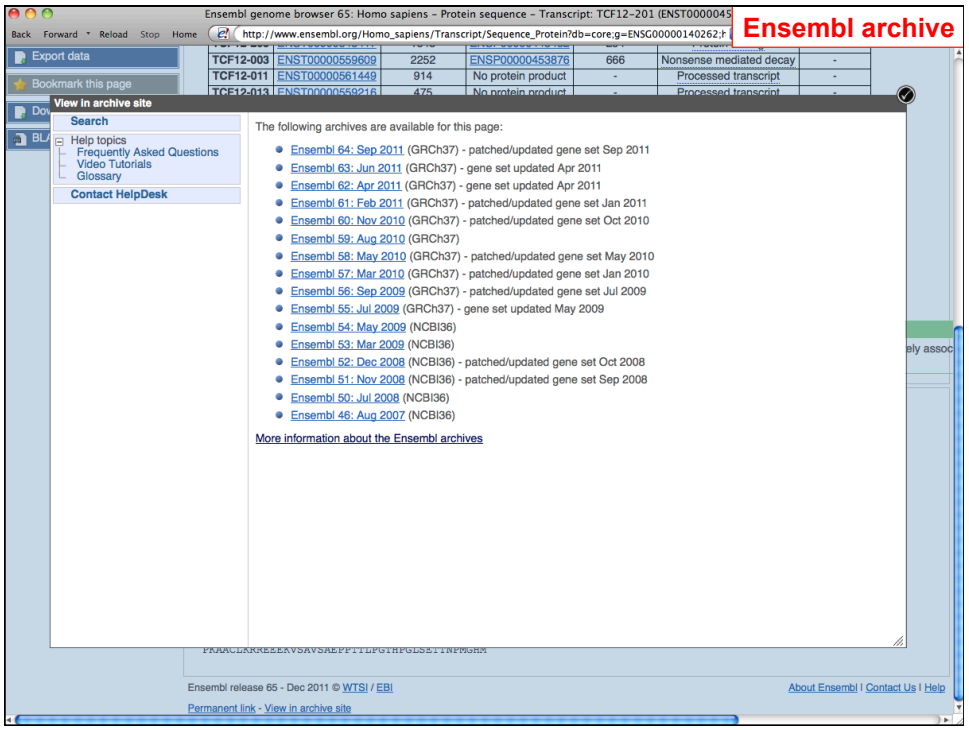
NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research



Name	Transcript ID	Length (bp)	Protein ID	Length (aa)	Biotype	CCDS
TCF12-001	ENST00000267811	6061	ENSP00000267811	682	Protein coding	CCDS10159
TCF12-002	ENST00000333725	4719	ENSP00000331057	706	Protein coding	CCDS10160
TCF12-004	ENST00000543579	1809	ENSP00000440017	536	Protein coding	-
TCF12-006	ENST00000537840	1598	ENSP00000444696	446	Protein coding	-
TCF12-007	ENST00000559922	1598	ENSP00000453216	446	Protein coding	-
TCF12-008	ENST00000343827	3956	ENSP00000342459	512	Protein coding	CCDS42042
TCF12-009	ENST00000559710	1314	ENSP00000453264	316	Protein coding	-
TCF12-010	ENST00000559703	1544	ENSP00000454102	339	Protein coding	-
TCF12-025	ENST00000438423	4786	ENSP00000388940	706	Protein coding	CCDS10160
TCF12-026	ENST00000557843	4076	ENSP00000453737	682	Protein coding	CCDS10159
TCF12-027	ENST00000557947	575	ENSP00000454109	157	Protein coding	-
TCF12-028	ENST00000561152	675	ENSP00000453653	58	Protein coding	-
TCF12-201	ENST00000452095	4772	ENSP00000396881	702	Protein coding	-
TCF12-202	ENST00000543236	2474	ENSP00000442910	681	Protein coding	-
TCF12-203	ENST00000543417	1545	ENSP00000443452	294	Protein coding	-
TCF12-003	ENST00000559609	2252	ENSP00000453876	666	Nonsense mediated decay	-
TCF12-011	ENST00000561449	914	No protein product	-	Processed transcript	-
TCF12-013	ENST00000559216	475	No protein product	-	Processed transcript	-
TCF12-014	ENST00000560887	540	No protein product	-	Processed transcript	-
TCF12-016	ENST00000561346	852	No protein product	-	Processed transcript	-
TCF12-017	ENST00000561235	818	No protein product	-	Processed transcript	-
TCF12-018	ENST00000560764	1230	No protein product	-	Processed transcript	-
TCF12-020	ENST00000560506	575	No protein product	-	Processed transcript	-




Ensembl Transcript tab:
Protein sequence
Export data
Bookmark this page
Download view as RTF
BLAST this sequence
Transcript and Gene level displays
Protein sequence help
Key
Exons
Alternating exons
Alternating exons
Residue overlap splice site
click
Permanent link - View in archive site
About Ensembl | Contact Us | Help

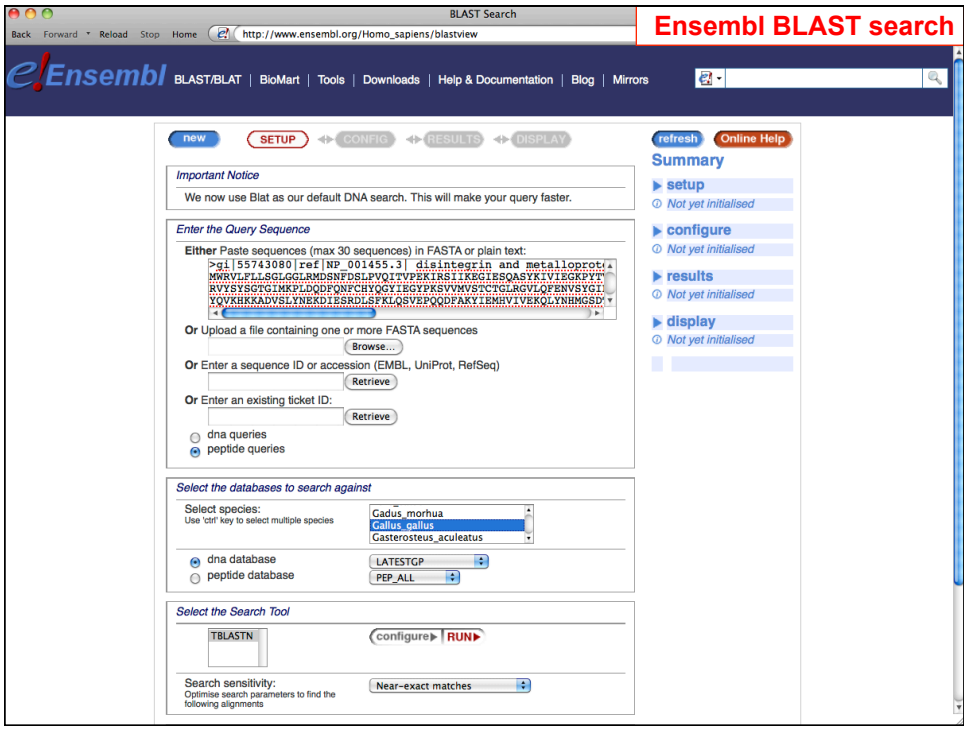

Ensembl archive
View in archive site
Search
Help topics
Frequently Asked Questions
Video Tutorials
Glossary
Contact HelpDesk
The following archives are available for this page:
Ensembl 64: Sep 2011 (GRCh37) - patched/updated gene set Sep 2011
Ensembl 63: Jun 2011 (GRCh37) - gene set updated Apr 2011
Ensembl 62: Apr 2011 (GRCh37) - gene set updated Apr 2011
Ensembl 61: Feb 2011 (GRCh37) - patched/updated gene set Jan 2011
Ensembl 60: Nov 2010 (GRCh37) - patched/updated gene set Oct 2010
Ensembl 59: Aug 2010 (GRCh37)
Ensembl 58: May 2010 (GRCh37) - patched/updated gene set May 2010
Ensembl 57: Mar 2010 (GRCh37) - patched/updated gene set Jan 2010
Ensembl 56: Sep 2009 (GRCh37) - patched/updated gene set Jul 2009
Ensembl 55: Jul 2009 (GRCh37) - gene set updated May 2009
Ensembl 54: May 2009 (NCBI36)
Ensembl 53: Mar 2009 (NCBI36)
Ensembl 52: Dec 2008 (NCBI36) - patched/updated gene set Oct 2008
Ensembl 51: Nov 2008 (NCBI36) - patched/updated gene set Sep 2008
Ensembl 50: Jul 2008 (NCBI36)
Ensembl 46: Aug 2007 (NCBI36)
More information about the Ensembl archives

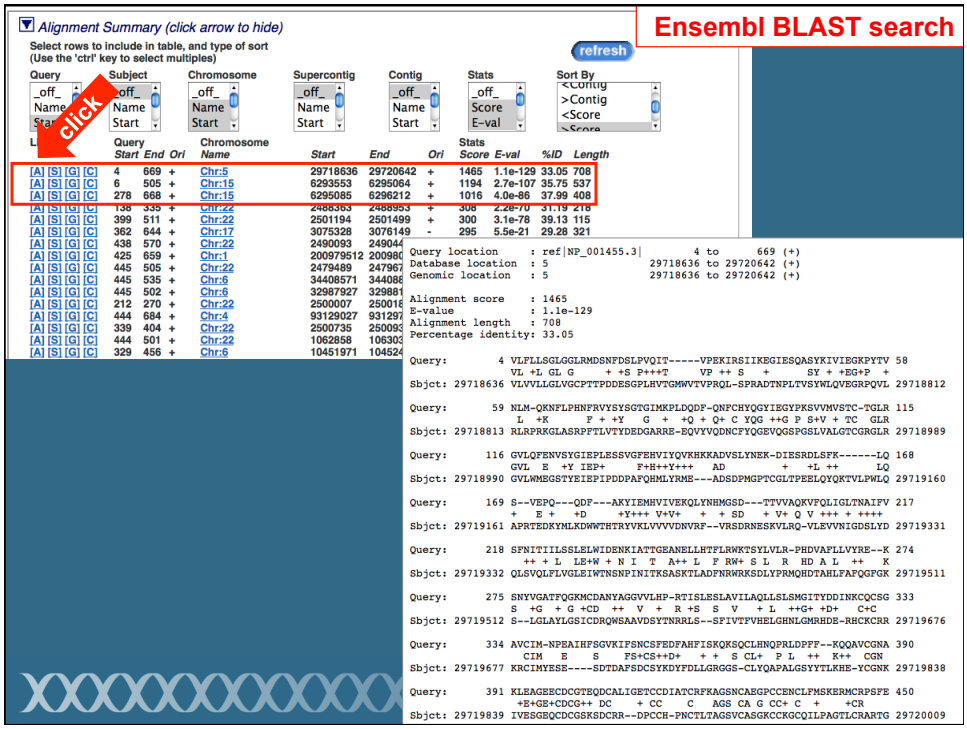
Ensembl

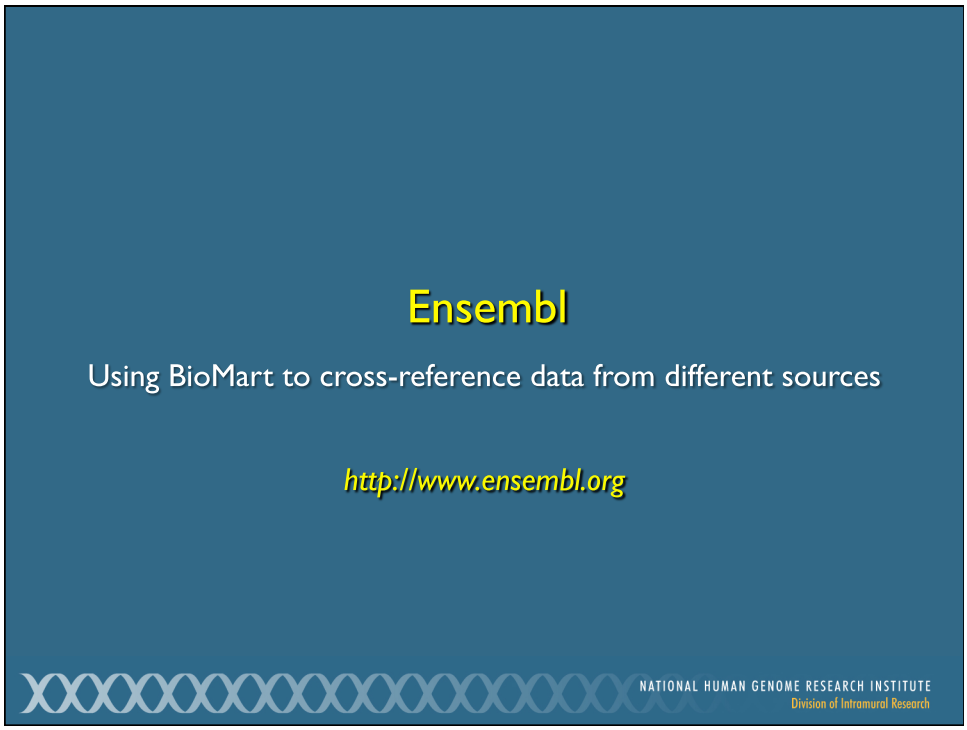
Find a chicken homolog of a human protein

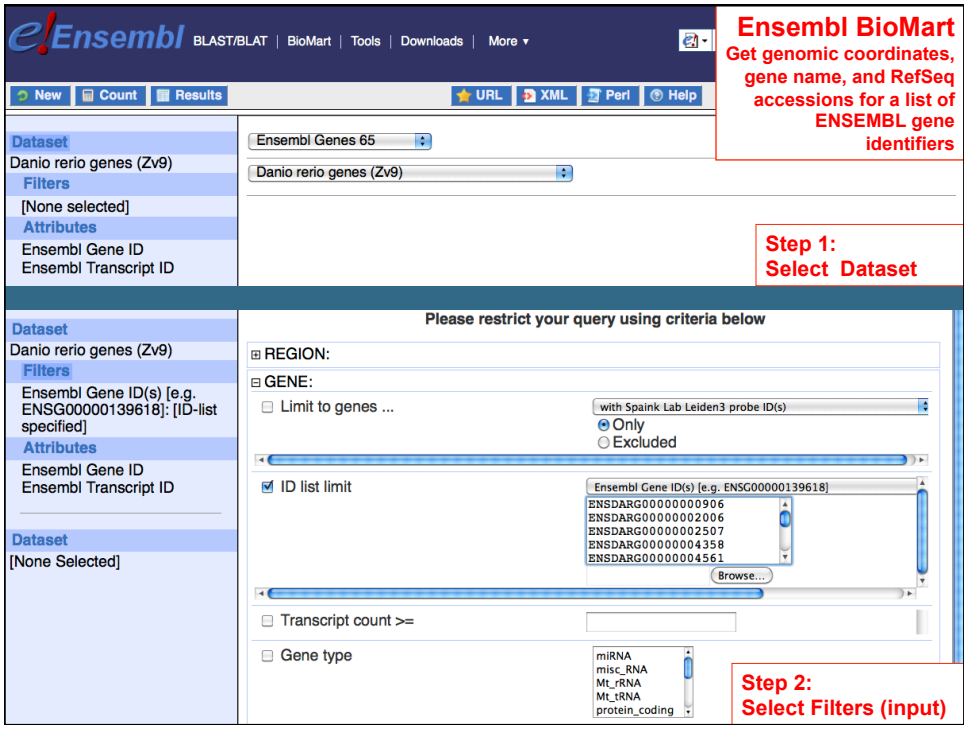
http://www.ensembl.org

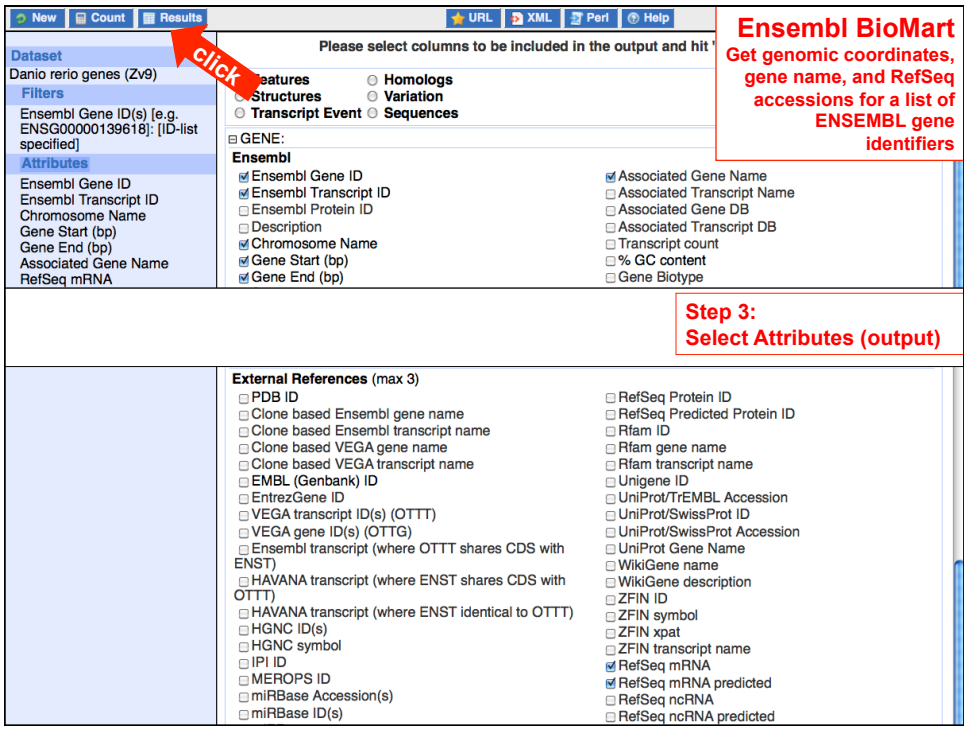
NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research

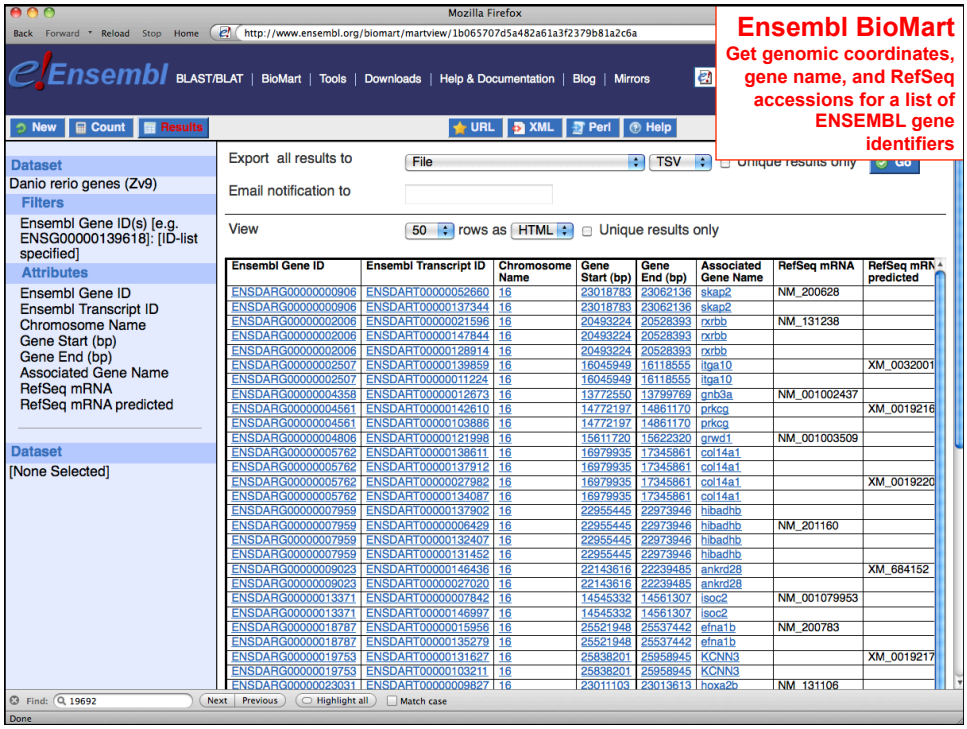


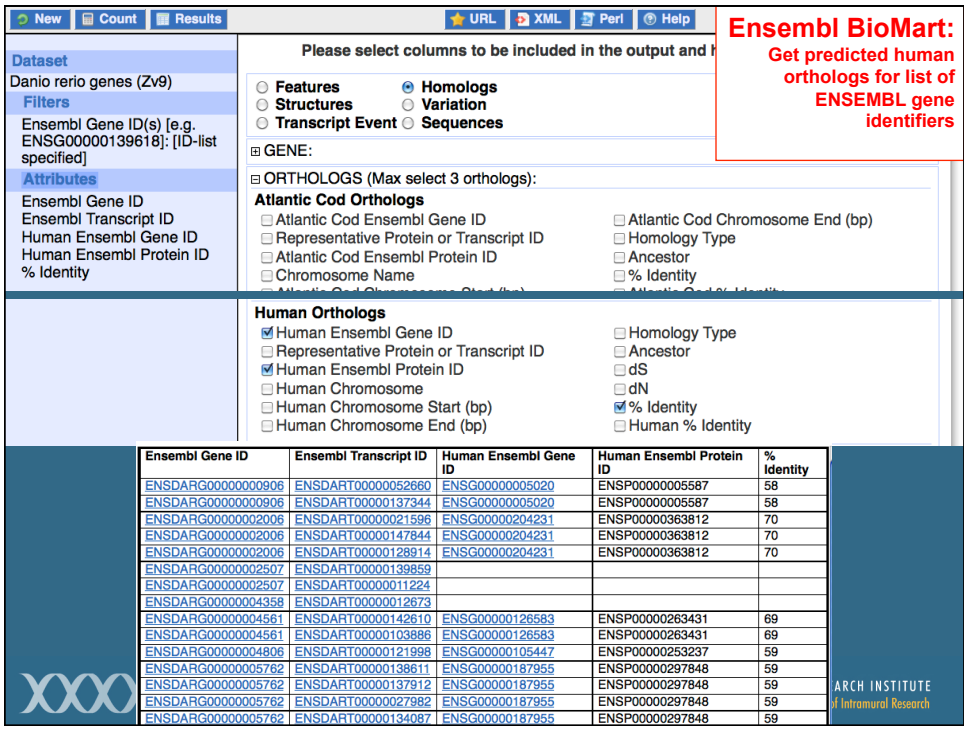

Ensembl BLAST search
Alignment Summary (click arrow to hide)
Select rows to include in table, and type of sort
(Use the 'ctrl' key to select multiples)
refresh
click
Query
Subject
Chromosome
Supercontig
Contig
Stats
Sort By
Score
E-val
Start
Name
Query location : ref|NP_001455.3| 4 to 669 (+)
Database location : 5 29718636 to 29720642 (+)
Genomic location : 5 29718636 to 29720642 (+)
Alignment score : 1465
E-value : 1.1e-129
Alignment length : 708
Percentage identity: 33.05


Ensembl
Using BioMart to cross-reference data from different sources
http://www.ensembl.org
NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research


Ensembl BioMart
Get genomic coordinates, gene name, and RefSeq accessions for a list of ENSEMBL gene identifiers
Ensembl BLAST/BLAT | BioMart | Tools | Downloads | More
New | Count | Results | URL | XML | Perl | Help
Dataset
Danio rerio genes (Zv9)
Filters
[None selected]
Attributes
Ensembl Gene ID
Ensembl Transcript ID
Ensembl Genes 65
Danio rerio genes (Zv9)
Step 1:
Select Dataset
Please restrict your query using criteria below
REGION:
GENE:
Limit to genes ...
with Spaink Lab Leiden3 probe ID(s)
Only
Excluded
ID list limit
Ensembl Gene ID(s) [e.g. ENSG00000139618]
ENSDARG00000000906
ENSDARG00000002006
ENSDARG00000002507
ENSDARG00000004358
ENSDARG00000004561
Browse...
Transcript count >=
Gene type
miRNA
misc_RNA
Mt_rRNA
Mt_tRNA
protein_coding
Step 2:
Select Filters (input)


Ensembl BioMart
Get genomic coordinates, gene name, and RefSeq accessions for a list of ENSEMBL gene identifiers
New | Count | Results | URL | XML | Perl | Help
click
Please select columns to be included in the output and hit
Features
Homologs
Structures
Variation
Transcript Event
Sequences
GENE:
Ensembl
Ensembl Gene ID
Ensembl Transcript ID
Ensembl Protein ID
Description
Chromosome Name
Gene Start (bp)
Gene End (bp)
Associated Gene Name
Associated Transcript Name
Associated Gene DB
Associated Transcript DB
Transcript count
% GC content
Gene Biotype
Step 3:
Select Attributes (output)
External References (max 3)
PDB ID
Clone based Ensembl gene name
Clone based Ensembl transcript name
Clone based VEGA gene name
Clone based VEGA transcript name
EMBL (Genbank) ID
EntrezGene ID
VEGA transcript ID(s) (OTTT)
VEGA gene ID(s) (OTTG)
Ensembl transcript (where OTTT shares CDS with ENST)
HAVANA transcript (where ENST shares CDS with OTTT)
HAVANA transcript (where ENST identical to OTTT)
HGNC ID(s)
HGNC symbol
IPI ID
MEROPS ID
miRBase Accession(s)
miRBase ID(s)
RefSeq Protein ID
RefSeq Predicted Protein ID
Rfam ID
Rfam gene name
Rfam transcript name
Unigene ID
UniProt/TrEMBL Accession
UniProt/SwissProt ID
UniProt/SwissProt Accession
UniProt Gene Name
WikiGene name
WikiGene description
ZFIN ID
ZFIN symbol
ZFIN xpat
ZFIN transcript name
RefSeq mRNA
RefSeq mRNA predicted
RefSeq ncRNA
RefSeq ncRNA predicted


Mozilla Firefox
http://www.ensembl.org/biomart/martview/1b065707d5a482a61a3f2379b81a2c6a
Ensembl BLAST/BLAT | BioMart | Tools | Downloads | Help & Documentation | Blog | Mirrors
New Count Results URL XML Perl Help
Dataset
Danio rerio genes (Zv9)
Filters
Ensembl Gene ID(s) [e.g. ENSG00000139618]: [ID-list specified]
Attributes
Ensembl Gene ID
Ensembl Transcript ID
Chromosome Name
Gene Start (bp)
Gene End (bp)
Associated Gene Name
RefSeq mRNA
RefSeq mRNA predicted
Dataset
[None Selected]
Export all results to
File
TSV
Email notification to
View
50 rows as HTML
Unique results only
Ensembl BioMart
Get genomic coordinates, gene name, and RefSeq accessions for a list of ENSEMBL gene identifiers

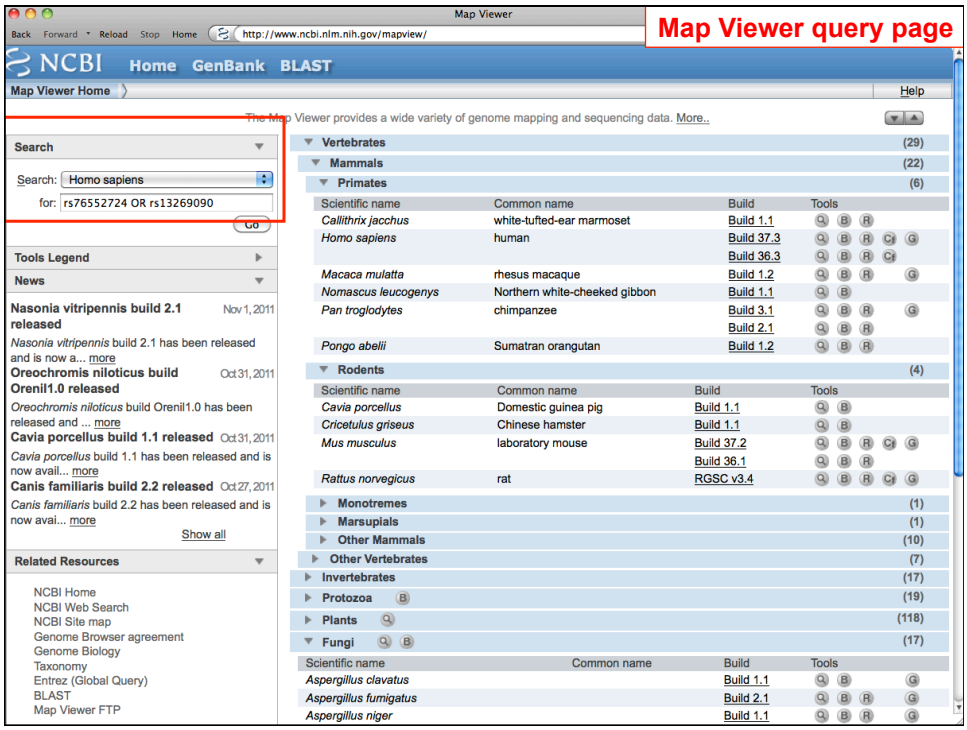

New Count Results URL XML Perl Help
Dataset
Danio rerio genes (Zv9)
Filters
Ensembl Gene ID(s) [e.g. ENSG00000139618]: [ID-list specified]
Attributes
Ensembl Gene ID
Ensembl Transcript ID
Human Ensembl Gene ID
Human Ensembl Protein ID
% Identity
Please select columns to be included in the output and h
Features
Homologs
Structures
Variation
Transcript Event
Sequences
GENE:
ORTHOLOGS (Max select 3 orthologs):
Atlantic Cod Orthologs
Atlantic Cod Ensembl Gene ID
Representative Protein or Transcript ID
Atlantic Cod Ensembl Protein ID
Chromosome Name
Atlantic Cod Chromosome End (bp)
Homology Type
Ancestor
% Identity
Human Orthologs
Human Ensembl Gene ID
Representative Protein or Transcript ID
Human Ensembl Protein ID
Human Chromosome
Human Chromosome Start (bp)
Human Chromosome End (bp)
Homology Type
Ancestor
dS
dN
% Identity
Human % Identity
Ensembl Gene ID | Ensembl Transcript ID | Human Ensembl Gene ID | Human Ensembl Protein ID | % Identity
ENSDARG00000000906 | ENSDART00000052660 | ENSG00000005020 | ENSP00000005587 | 58
ENSDARG00000000906 | ENSDART00000137344 | ENSG00000005020 | ENSP00000005587 | 58
ENSDARG00000002006 | ENSDART00000021596 | ENSG00000204231 | ENSP00000363812 | 70
ENSDARG00000002006 | ENSDART00000147844 | ENSG00000204231 | ENSP00000363812 | 70
ENSDARG00000002006 | ENSDART00000128914 | ENSG00000204231 | ENSP00000363812 | 70
ENSDARG00000002507 | ENSDART00000139859
ENSDARG00000002507 | ENSDART00000011224
ENSDARG00000004358 | ENSDART00000012673
ENSDARG00000004561 | ENSDART00000142610 | ENSG00000126583 | ENSP00000263431 | 69
ENSDARG00000004561 | ENSDART00000103886 | ENSG00000126583 | ENSP00000263431 | 69
ENSDARG00000004806 | ENSDART00000121998 | ENSG00000105447 | ENSP00000253237 | 59
ENSDARG00000005762 | ENSDART00000138611 | ENSG00000187955 | ENSP00000297848 | 59
ENSDARG00000005762 | ENSDART00000137912 | ENSG00000187955 | ENSP00000297848 | 59
ENSDARG00000005762 | ENSDART00000027982 | ENSG00000187955 | ENSP00000297848 | 59
ENSDARG00000005762 | ENSDART00000134087 | ENSG00000187955 | ENSP00000297848 | 59
Ensembl BioMart:
Get predicted human orthologs for list of ENSEMBL gene identifiers

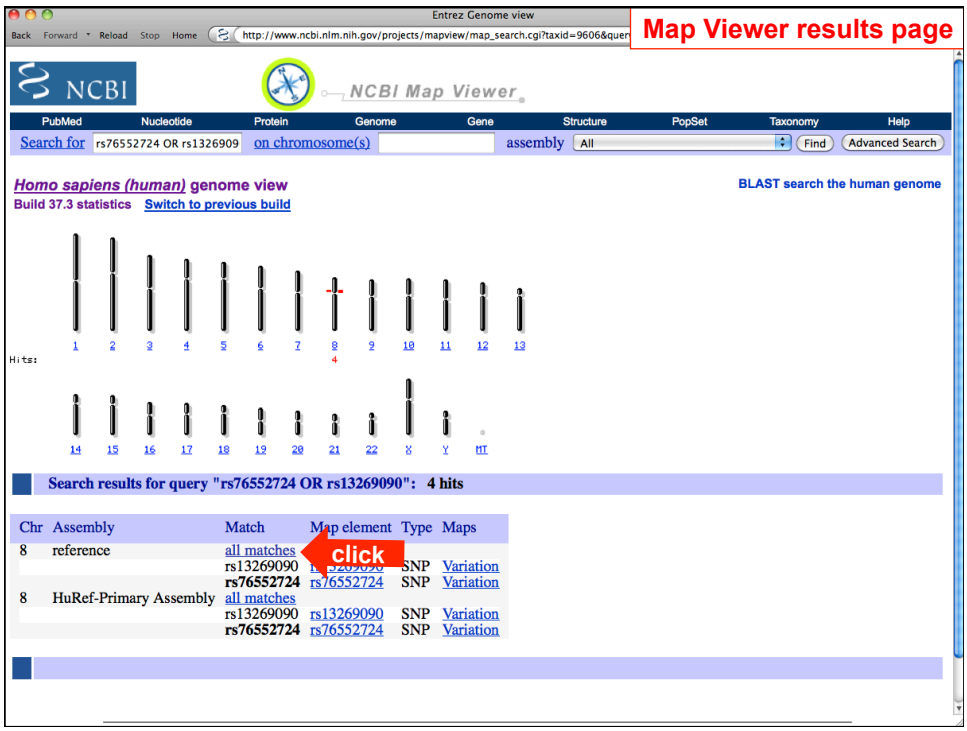
NCBI

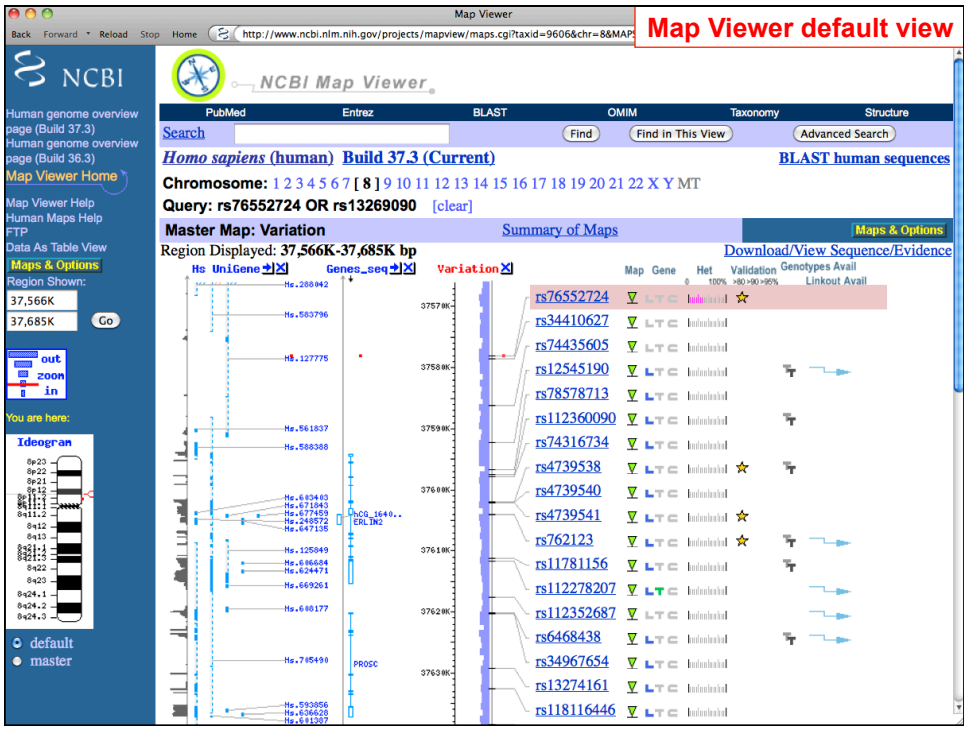
View a genomic region between two SNPs

http://www.ncbi.nlm.nih.gov/mapview

NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research




Map Viewer results page
Entrez Genome view
NCBI
NCBI Map Viewer
PubMed Nucleotide Protein Genome Gene Structure PopSet Taxonomy Help
Search for rs76552724 OR rs1326909 on chromosome(s) assembly All Find Advanced Search
Homo sapiens (human) genome view
Build 37.3 statistics Switch to previous build
BLAST search the human genome
Hits:
Search results for query "rs76552724 OR rs13269090": 4 hits
Chr Assembly Match Map element Type Maps
8 reference all matches
rs13269090 SNP Variation
rs76552724 rs76552724 SNP Variation
8 HuRef-Primary Assembly all matches
rs13269090 rs13269090 SNP Variation
rs76552724 rs76552724 SNP Variation
click

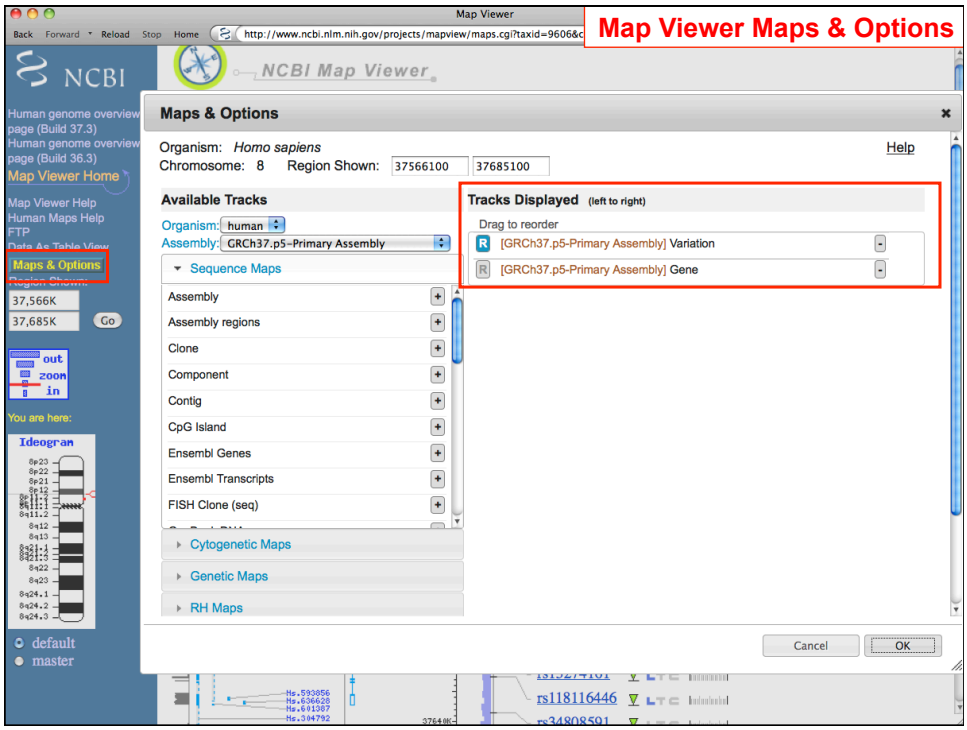

Map Viewer default view
Map Viewer
NCBI
NCBI Map Viewer
PubMed Entrez BLAST OMIM Taxonomy Structure
Search Find Find in This View Advanced Search
Homo sapiens (human) Build 37.3 (Current)
BLAST human sequences
Chromosome: 1 2 3 4 5 6 7 [8] 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y MT
Query: rs76552724 OR rs13269090 [clear]
Master Map: Variation
Summary of Maps
Maps & Options
Region Displayed: 37,566K-37,685K bp
Download/View Sequence/Evidence
Hs UniGene Genes_seq Variation
Map Gene Het Validation Genotypes Avail Linkout Avail
rs76552724
rs34410627
rs74435605
rs12545190
rs78578713
rs112360090
rs74316734
rs4739538
rs4739540
rs4739541
rs762123
rs11781156
rs112278207
rs112352687
rs6468438
rs34967654
rs13274161
rs118116446
Human genome overview page (Build 37.3)
Human genome overview page (Build 36.3)
Map Viewer Home
Map Viewer Help
Human Maps Help
FTP
Data As Table View
Maps & Options
Region Shown:
37,566K
37,685K
Go
You are here:
Ideogram
default
master

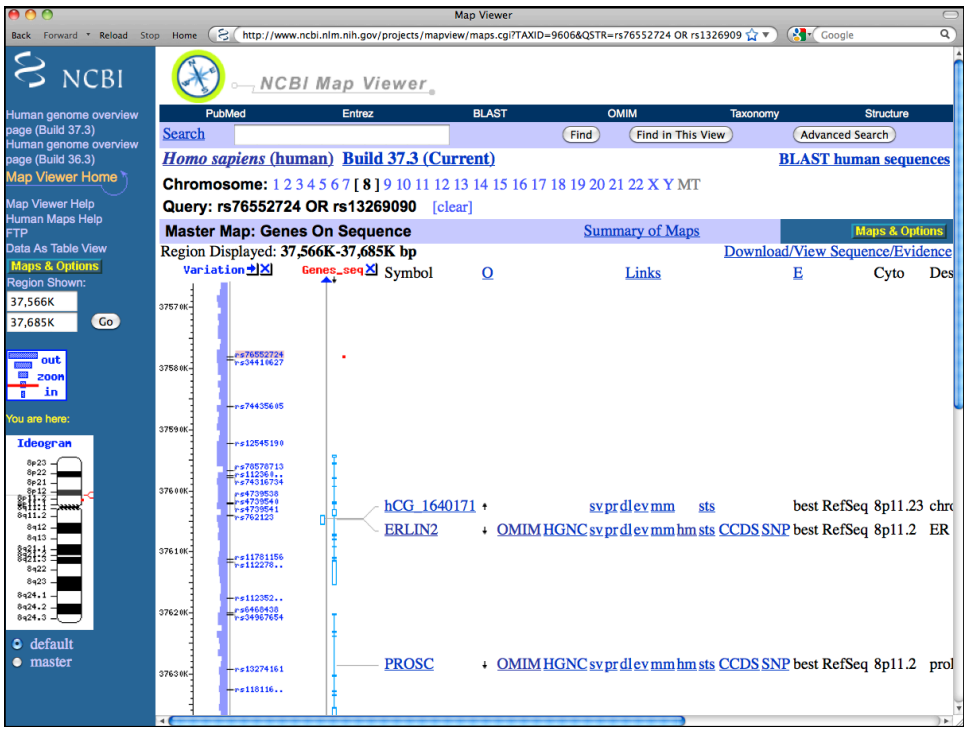
NCBI

Change the maps displayed on the Map Viewer

http://www.ncbi.nlm.nih.gov/mapview

NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research

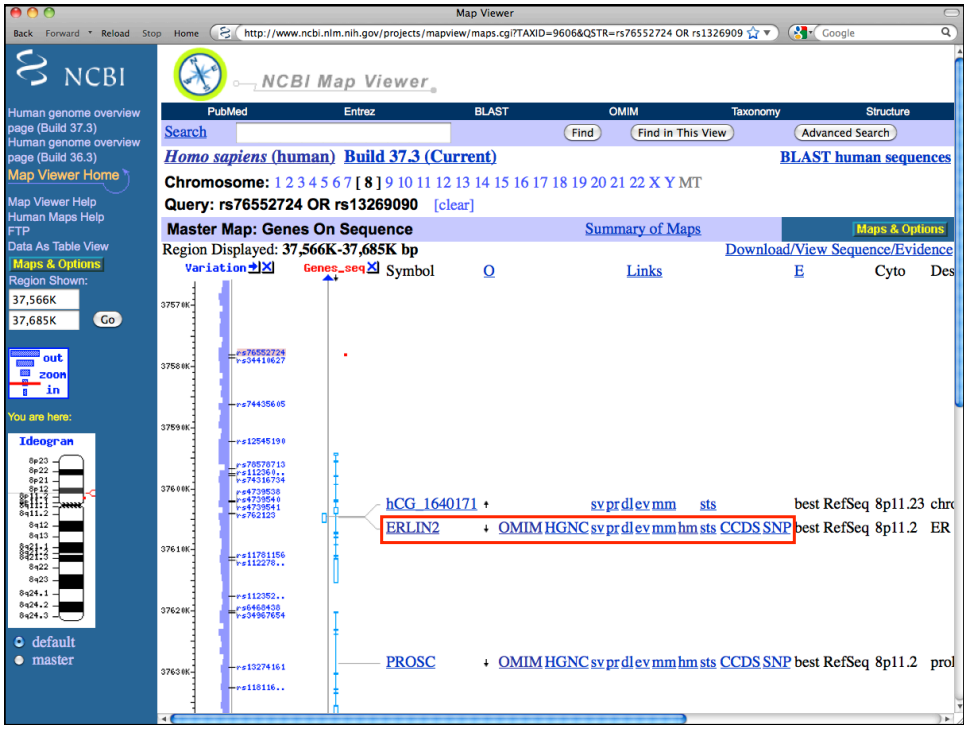


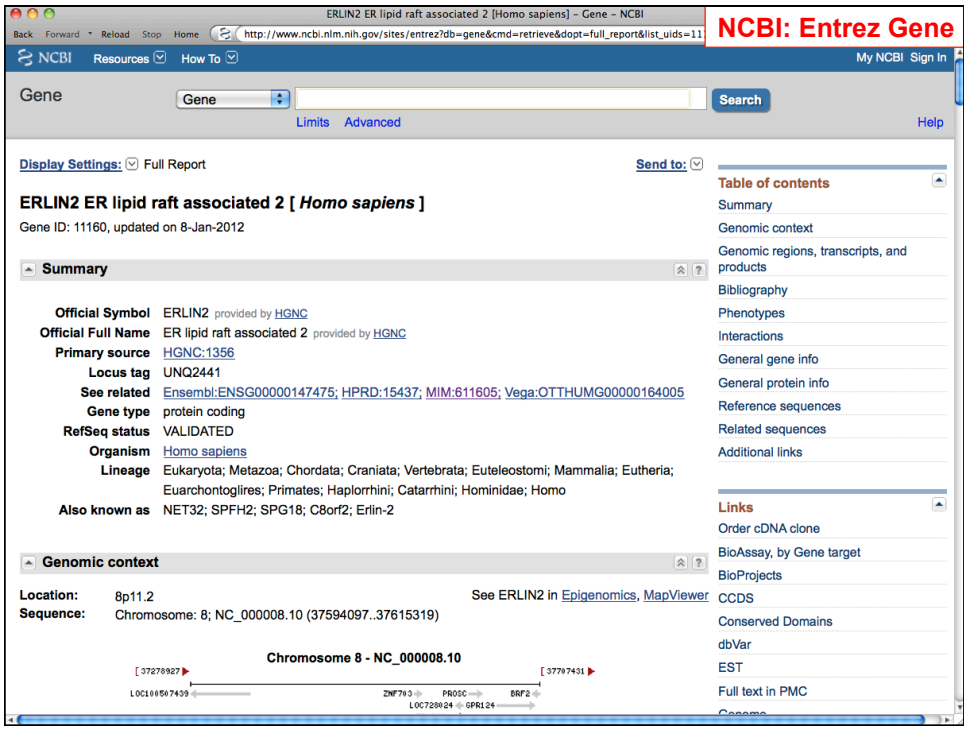


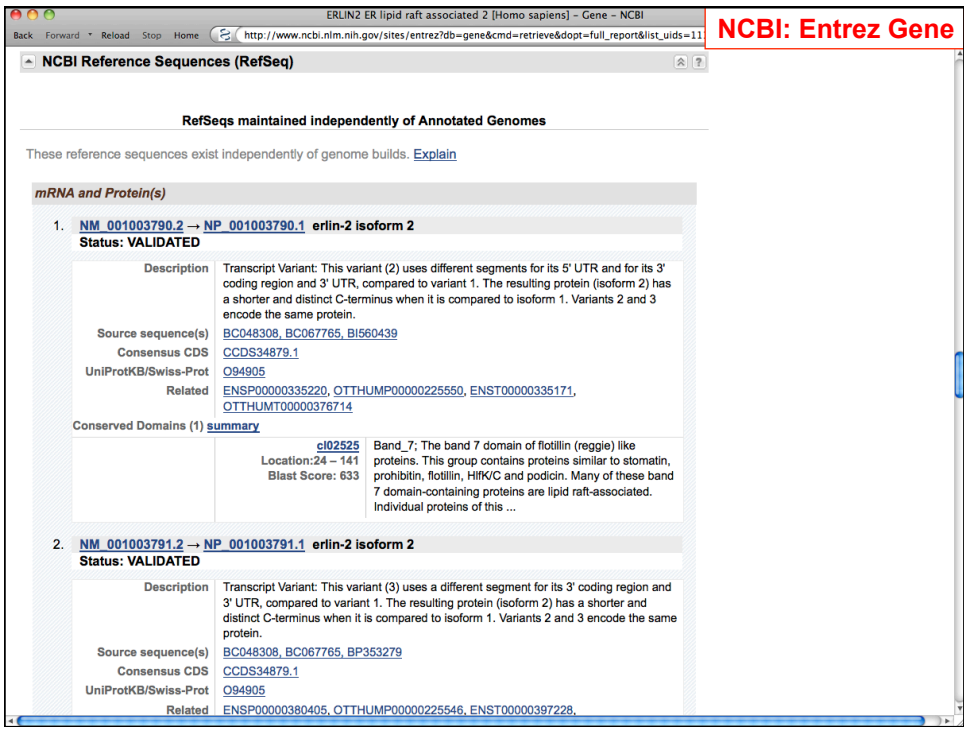
NCBI

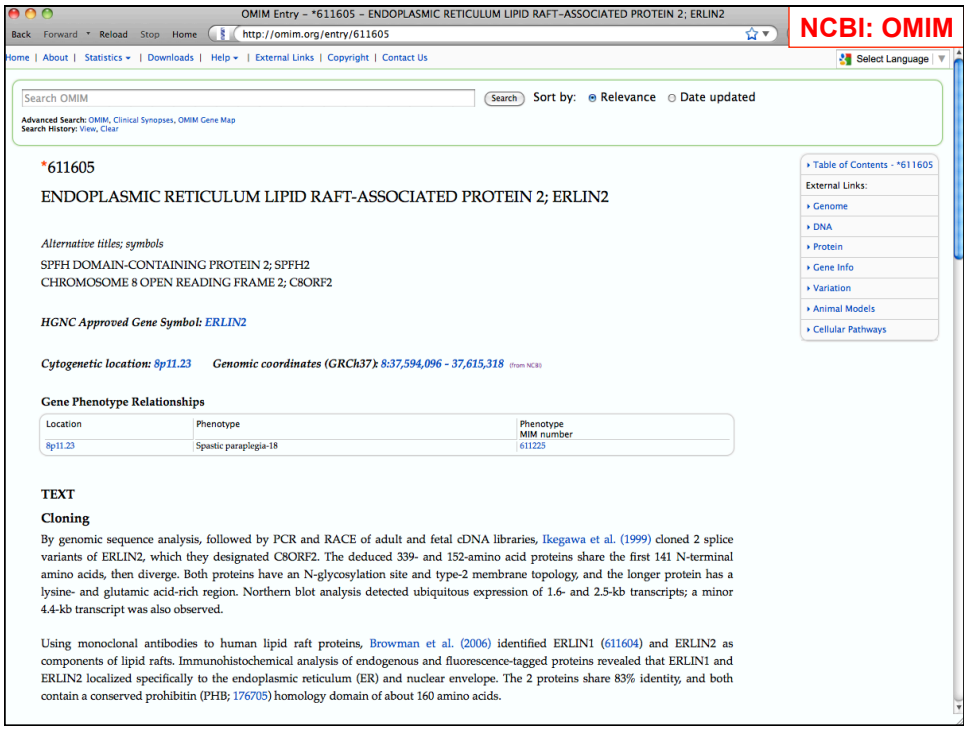
View additional information about a gene

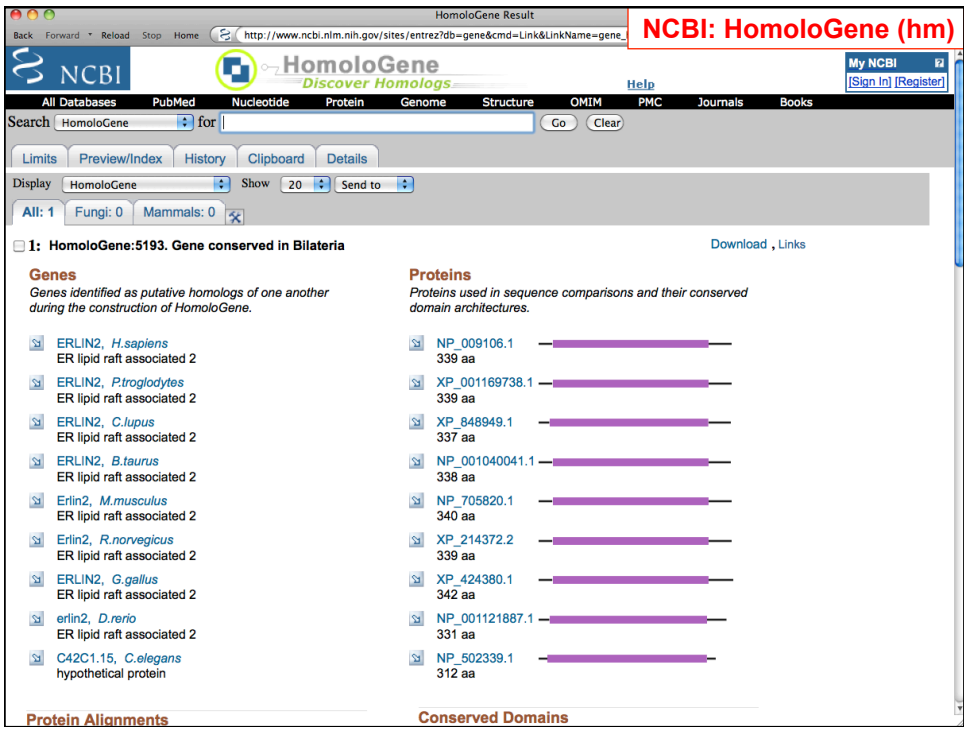
http://www.ncbi.nlm.nih.gov/mapview


Map Viewer
NCBI
NCBI Map Viewer
Human genome overview page (Build 37.3)
Human genome overview page (Build 36.3)
Map Viewer Home
Map Viewer Help
Human Maps Help
FTP
Data As Table View
Maps & Options
Region Shown:
37,566K
37,685K
Go
You are here:
Ideogram
default
master
PubMed
Entrez
BLAST
OMIM
Taxonomy
Structure
Search
Find
Find in This View
Advanced Search
Homo sapiens (human) Build 37.3 (Current)
BLAST human sequences
Chromosome: 1 2 3 4 5 6 7 [8] 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y MT
Query: rs76552724 OR rs13269090 [clear]
Master Map: Genes On Sequence
Summary of Maps
Maps & Options
Region Displayed: 37,566K-37,685K bp
Download/View Sequence/Evidence
Variation
Genes_seq
Symbol
Links
Cyto
hCG_1640171
sv pr dl ev mm sts
best RefSeq 8p11.23
ERLIN2
OMIM HGNC sv pr dl ev mm hm sts CCDS SNP
best RefSeq 8p11.2
PROSC
OMIM HGNC sv pr dl ev mm hm sts CCDS SNP
best RefSeq 8p11.2

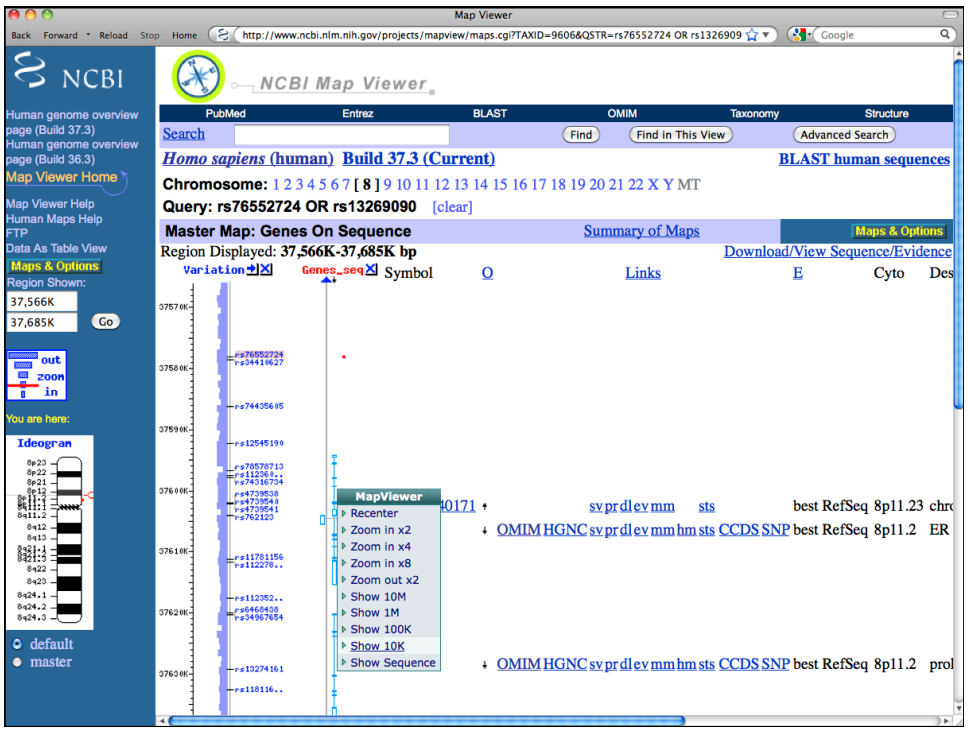

NCBI: Entrez Gene
ERLIN2 ER lipid raft associated 2 [Homo sapiens] - Gene - NCBI
NCBI
Resources
How To
My NCBI Sign In
Gene
Search
Limits Advanced
Help
Display Settings: Full Report
Send to:
ERLIN2 ER lipid raft associated 2 [Homo sapiens]
Gene ID: 11160, updated on 8-Jan-2012
Summary
Official Symbol ERLIN2 provided by HGNC
Official Full Name ER lipid raft associated 2 provided by HGNC
Primary source HGNC:1356
Locus tag UNQ2441
See related Ensembl:ENSG00000147475; HPRD:15437; MIM:611605; Vega:OTTHUMG00000164005
Gene type protein coding
RefSeq status VALIDATED
Organism Homo sapiens
Lineage Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo
Also known as NET32; SPFH2; SPG18; C8orf2; Erlin-2
Genomic context
Location: 8p11.2
Sequence: Chromosome: 8; NC_000008.10 (37594097..37615319)
See ERLIN2 in Epigenomics, MapViewer
Chromosome 8 - NC_000008.10
Table of contents
Summary
Genomic context
Genomic regions, transcripts, and products
Bibliography
Phenotypes
Interactions
General gene info
General protein info
Reference sequences
Related sequences
Additional links
Links
Order cDNA clone
BioAssay, by Gene target
BioProjects
CCDS
Conserved Domains
dbVar
EST
Full text in PMC

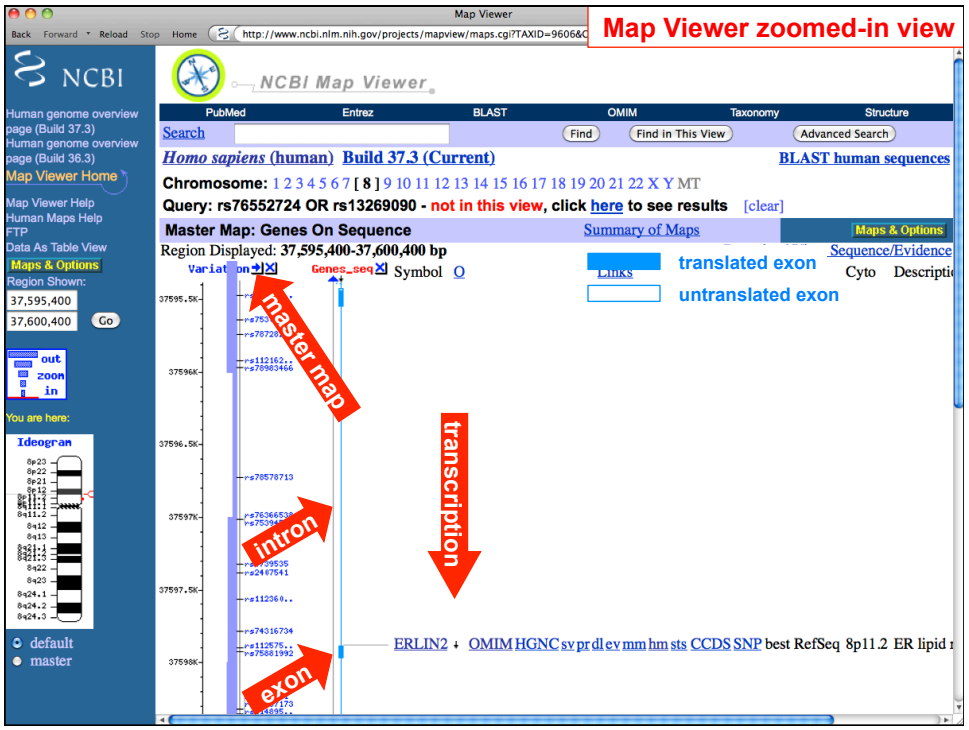

ERLIN2 ER lipid raft associated 2 [Homo sapiens] - Gene - NCBI
Back Forward Reload Stop Home
http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene&cmd=retrieve&dopt=full_report&list_uids=11
NCBI: Entrez Gene
NCBI Reference Sequences (RefSeq)
RefSeqs maintained independently of Annotated Genomes
These reference sequences exist independently of genome builds. Explain
mRNA and Protein(s)
1. NM_001003790.2 → NP_001003790.1 erlin-2 isoform 2
Status: VALIDATED
Description
Transcript Variant: This variant (2) uses different segments for its 5' UTR and for its 3' coding region and 3' UTR, compared to variant 1. The resulting protein (isoform 2) has a shorter and distinct C-terminus when it is compared to isoform 1. Variants 2 and 3 encode the same protein.
Source sequence(s) BC048308, BC067765, BI560439
Consensus CDS CCDS34879.1
UniProtKB/Swiss-Prot O94905
Related ENSP00000335220, OTTHUMP00000225550, ENST00000335171, OTTHUMT00000376714
Conserved Domains (1) summary
cl02525
Location:24 – 141
Blast Score: 633
Band_7; The band 7 domain of flotillin (reggie) like proteins. This group contains proteins similar to stomatin, prohibitin, flotillin, HlfK/C and podicin. Many of these band 7 domain-containing proteins are lipid raft-associated. Individual proteins of this ...
2. NM_001003791.2 → NP_001003791.1 erlin-2 isoform 2
Status: VALIDATED
Description
Transcript Variant: This variant (3) uses a different segment for its 3' coding region and 3' UTR, compared to variant 1. The resulting protein (isoform 2) has a shorter and distinct C-terminus when it is compared to isoform 1. Variants 2 and 3 encode the same protein.
Source sequence(s) BC048308, BC067765, BP353279
Consensus CDS CCDS34879.1
UniProtKB/Swiss-Prot O94905
Related ENSP00000380405, OTTHUMP00000225546, ENST00000397228,

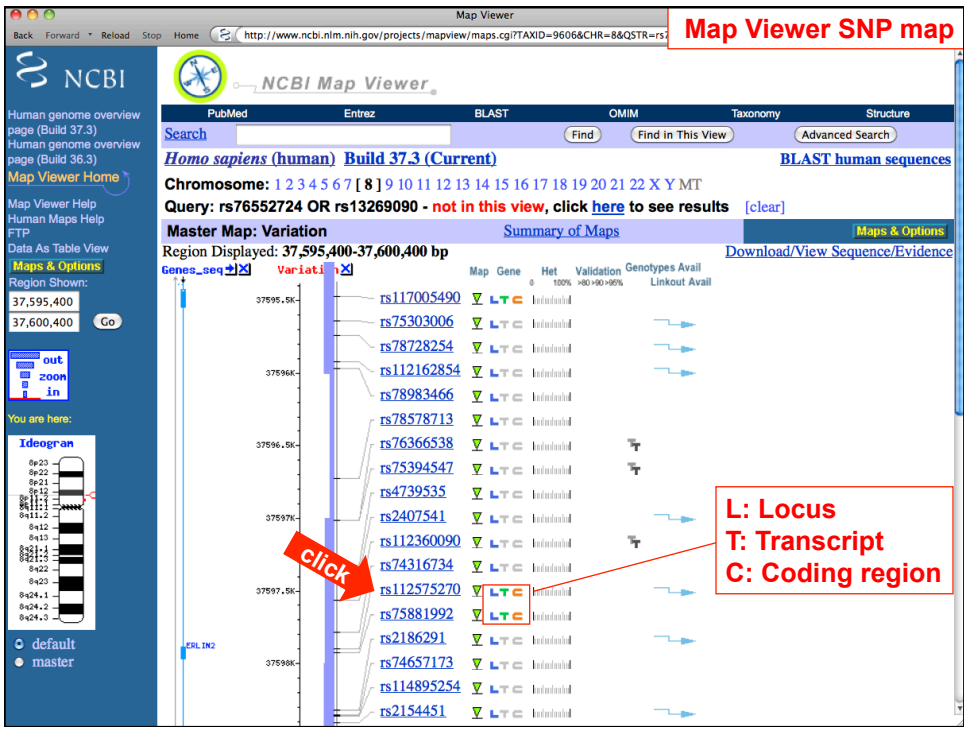

OMIM Entry - *611605 - ENDOPLASMIC RETICULUM LIPID RAFT-ASSOCIATED PROTEIN 2; ERLIN2
Back Forward Reload Stop Home
http://omim.org/entry/611605
NCBI: OMIM
Home | About | Statistics | Downloads | Help | External Links | Copyright | Contact Us
Select Language
Search OMIM
Search
Sort by: Relevance Date updated
Advanced Search: OMIM, Clinical Synopses, OMIM Gene Map
Search History: View, Clear
*611605
ENDOPLASMIC RETICULUM LIPID RAFT-ASSOCIATED PROTEIN 2; ERLIN2
Alternative titles; symbols
SPFH DOMAIN-CONTAINING PROTEIN 2; SPFH2
CHROMOSOME 8 OPEN READING FRAME 2; C8ORF2
HGNC Approved Gene Symbol: ERLIN2
Cytogenetic location: 8p11.23
Genomic coordinates (GRCh37): 8:37,594,096 - 37,615,318
Table of Contents - *611605
External Links:
Genome
DNA
Protein
Gene Info
Variation
Animal Models
Cellular Pathways
Gene Phenotype Relationships
Location
Phenotype
Phenotype MIM number
8p11.23
Spastic paraplegia-18
611225
TEXT
Cloning
By genomic sequence analysis, followed by PCR and RACE of adult and fetal cDNA libraries, Ikegawa et al. (1999) cloned 2 splice variants of ERLIN2, which they designated C8ORF2. The deduced 339- and 152-amino acid proteins share the first 141 N-terminal amino acids, then diverge. Both proteins have an N-glycosylation site and type-2 membrane topology, and the longer protein has a lysine- and glutamic acid-rich region. Northern blot analysis detected ubiquitous expression of 1.6- and 2.5-kb transcripts; a minor 4.4-kb transcript was also observed.
Using monoclonal antibodies to human lipid raft proteins, Browman et al. (2006) identified ERLIN1 (611604) and ERLIN2 as components of lipid rafts. Immunohistochemical analysis of endogenous and fluorescence-tagged proteins revealed that ERLIN1 and ERLIN2 localized specifically to the endoplasmic reticulum (ER) and nuclear envelope. The 2 proteins share 83% identity, and both contain a conserved prohibitin (PHB; 176705) homology domain of about 160 amino acids.

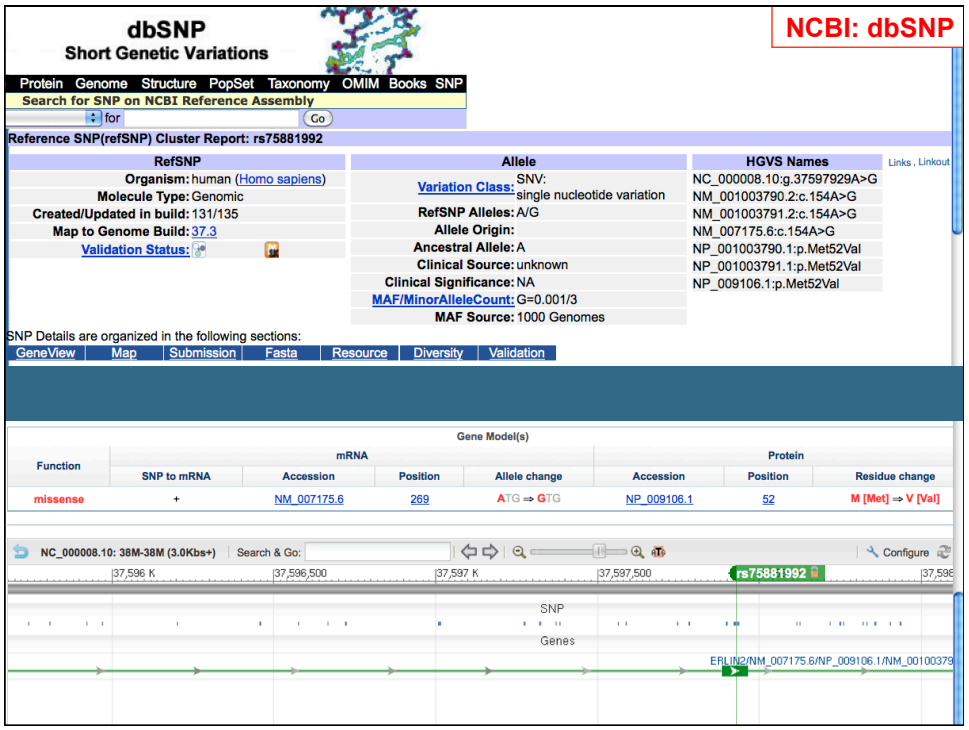

NCBI: HomoloGene (hm)
HomoloGene Result
http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene&cmd=Link&LinkName=gene_
NCBI
HomoloGene
Discover Homologs
Help
My NCBI
[Sign In] [Register]
All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Journals Books
Search HomoloGene for Go Clear
Limits Preview/Index History Clipboard Details
Display HomoloGene Show 20 Send to
All: 1 Fungi: 0 Mammals: 0
1: HomoloGene:5193. Gene conserved in Bilateria
Download , Links
Genes
Genes identified as putative homologs of one another during the construction of HomoloGene.
Proteins
Proteins used in sequence comparisons and their conserved domain architectures.
ERLIN2, H.sapiens ER lipid raft associated 2
NP_009106.1 339 aa
ERLIN2, P.troglodytes ER lipid raft associated 2
XP_001169738.1 339 aa
ERLIN2, C.lupus ER lipid raft associated 2
XP_848949.1 337 aa
ERLIN2, B.taurus ER lipid raft associated 2
NP_001040041.1 338 aa
Erlin2, M.musculus ER lipid raft associated 2
NP_705820.1 340 aa
Erlin2, R.norvegicus ER lipid raft associated 2
XP_214372.2 339 aa
ERLIN2, G.gallus ER lipid raft associated 2
XP_424380.1 342 aa
erlin2, D.rerio ER lipid raft associated 2
NP_001121887.1 331 aa
C42C1.15, C.elegans hypothetical protein
NP_502339.1 312 aa
Protein Alignments
Conserved Domains


NCBI
Zoom in to view greater detail
http://www.ncbi.nlm.nih.gov/mapview
NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research


Map Viewer
NCBI
NCBI Map Viewer
PubMed
Entrez
BLAST
OMIM
Taxonomy
Structure
Search
Find
Find in This View
Advanced Search
Homo sapiens (human) Build 37.3 (Current)
BLAST human sequences
Chromosome: 1 2 3 4 5 6 7 [8] 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y MT
Query: rs76552724 OR rs13269090 [clear]
Master Map: Genes On Sequence
Summary of Maps
Maps & Options
Region Displayed: 37,566K-37,685K bp
Download/View Sequence/Evidence
Variation
Genes_seq
Symbol
Links
Cyto
MapViewer
Recenter
Zoom in x2
Zoom in x4
Zoom in x8
Zoom out x2
Show 10M
Show 1M
Show 100K
Show 10K
Show Sequence
Region Shown:
37,566K
37,685K
Go
You are here:
Ideogram
default
master


Map Viewer zoomed-in view
NCBI
NCBI Map Viewer
Homo sapiens (human) Build 37.3 (Current)
Chromosome: 1 2 3 4 5 6 7 [8] 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y MT
Query: rs76552724 OR rs13269090 - not in this view, click here to see results [clear]
Master Map: Genes On Sequence
Region Displayed: 37,595,400-37,600,400 bp
translated exon
untranslated exon
master map
transcription
intron
exon
ERLIN2
Region Shown:
37,595,400
37,600,400
Go
You are here:
Ideogram
default
master


Map Viewer SNP map
Map Viewer
NCBI
NCBI Map Viewer
Human genome overview page (Build 37.3)
Human genome overview page (Build 36.3)
Map Viewer Home
Map Viewer Help
Human Maps Help
FTP
Data As Table View
Maps & Options
Region Shown:
37,595,400
37,600,400
Go
out
zoom
in
You are here:
Ideogram
default
master
PubMed
Entrez
BLAST
OMIM
Taxonomy
Structure
Search
Find
Find in This View
Advanced Search
Homo sapiens (human) Build 37.3 (Current)
BLAST human sequences
Chromosome: 1 2 3 4 5 6 7 [8] 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y MT
Query: rs76552724 OR rs13269090 - not in this view, click here to see results
[clear]
Master Map: Variation
Summary of Maps
Maps & Options
Region Displayed: 37,595,400-37,600,400 bp
Download/View Sequence/Evidence
Genes_seq
Variation
Map Gene Het Validation Genotypes Avail Linkout Avail
rs117005490
rs75303006
rs78728254
rs112162854
rs78983466
rs78578713
rs76366538
rs75394547
rs4739535
rs2407541
rs112360090
rs74316734
rs112575270
rs75881992
rs2186291
rs74657173
rs114895254
rs2154451
click
L: Locus
T: Transcript
C: Coding region


NCBI: dbSNP
dbSNP
Short Genetic Variations
Protein Genome Structure PopSet Taxonomy OMIM Books SNP
Search for SNP on NCBI Reference Assembly
for
Go
Reference SNP(refSNP) Cluster Report: rs75881992
RefSNP
Organism: human (Homo sapiens)
Molecule Type: Genomic
Created/Updated in build: 131/135
Map to Genome Build: 37.3
Validation Status:
Allele
Variation Class: SNV: single nucleotide variation
RefSNP Alleles: A/G
Allele Origin:
Ancestral Allele: A
Clinical Source: unknown
Clinical Significance: NA
MAF/MinorAlleleCount: G=0.001/3
MAF Source: 1000 Genomes
HGVS Names
Links , Linkout
NC_000008.10:g.37597929A>G
NM_001003790.2:c.154A>G
NM_001003791.2:c.154A>G
NM_007175.6:c.154A>G
NP_001003790.1:p.Met52Val
NP_001003791.1:p.Met52Val
NP_009106.1:p.Met52Val
SNP Details are organized in the following sections:
GeneView Map Submission Fasta Resource Diversity Validation
Gene Model(s)
Function
mRNA
SNP to mRNA
Accession
Position
Allele change
Protein
Accession
Position
Residue change
missense
+
NM_007175.6
269
ATG ⇒ GTG
NP_009106.1
52
M [Met] ⇒ V [Val]
NC_000008.10: 38M-38M (3.0Kbs+)
Search & Go:
Configure
37,596 K
37,596,500
37,597 K
37,597,500
rs75881992
SNP
Genes
ERLIN2/NM_007175.6/NP_009106.1/NM_00100379

Additional resources

- UCSC Human Genome Browser User Guide
  http://genome.ucsc.edu/goldenPath/help/
- Ensembl Tutorials and Worked Examples
  http://www.ensembl.org/info/website/tutorials/
- NCBI MapViewer Help
  http://www.ncbi.nlm.nih.gov/mapview/static/MapViewerHelp.html

NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Division of Intramural Research

Current Protocols in Bioinformatics

The UCSC Genome Browser — *UNIT 1.4*

Donna Karolchik,[1] Angie S. Hinrichs,[1] and W. James Kent[1]

[1]Center for Biomolecular Science and Engineering, University of California Santa Cruz, Santa Cruz, California

ABSTRACT

The University of California Santa Cruz (UCSC) Genome Browse based tool for quickly displaying a requested portion of a genome panied by a series of aligned annotation "tracks." The annotation UCSC Genome Bioinformatics Group and external collaborators— tions, mRNA and expressed sequence tag alignments, simple nucleo expression and regulatory data, phenotype and variation data, and pa species comparative genomics data. All information relevant to a in one window, facilitating biological analysis and interpretation. underlying the Genome Browser tracks can be viewed, downloade using another Web-based application, the UCSC Table Browser. U

Using the NCBI Map Viewer to Browse Genomic Sequence Data — *UNIT 1.5*

Tyra G. Wolfsberg[1]

[1]Bethesda, Maryland

ABSTRACT

This unit includes a Basic Protocol with an introduction to the Map Viewer, describing how to perform a simple text-based search of genome annotations to view the genomic context of a gene, navigate along a chromosome, zoom in and out, and change the displayed maps to hide and show information. It also describes some of NCBI's sequence- inks from the Map Viewer. The Alternate Protocols genome sequence, and also illustrate additional Protocol 1 shows how to perform and interpret the human genome. Alternate Protocol 2 demonstrates ween two STS markers. Finally, Alternate Protocol embers of a gene family. *Curr. Protoc. Bioinform.* cy & Sons, Inc.

e assembly • genomic sequence • gene map

Using the Ensembl Genome Server to Browse Genomic Sequence Data — *UNIT 1.15*

Xosé M. Fernández-Suárez[1] and Michael K. Schuster[1]

[1]EMBL-European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, United Kingdom

ABSTRACT

The Ensembl project provides a comprehensive source of automatic annotation of the human genome sequence, as well as other species of biomedical interest, with confirmed gene predictions that have been integrated with external data sources. This unit describes how to use the Ensembl genome browser (*http://www.ensembl.org/*), the public interface of the project. It describes how to find a gene or protein of interest, how to get additional information and external links, and how to use the comparative genomic

Protoc. Bioinform. 30:1.15.1-1.15.48. © 2010 by John Wiley & Sons, Inc.

Keywords: computer graphics • databases • genetic • genetic variation • ge sequence homology • genome • genome sequence

Access from NIH at
http://onlinelibrary.wiley.com/book/10.1002/0471250953