

Browsing Genomes with Ensembl

The screenshot displays the Ensembl genome browser interface. At the top, there is a navigation bar with the Ensembl logo and 'Archive! Ensembl' branding. Below this is a search bar with the text 'Search all Ensembl: [Anything] Go'. The main content area is divided into several sections:

- Search Ensembl:** A search box with the text 'Search [All species] for [] Go'. Below it, a note says '+ e.g. mouse chromosome 7 or X10000_70000 or human gene BRCA2'.
- Ensembl tools:** A list of tools including 'Start a sequence search', 'Mine Ensembl with BioMart', 'Customise Your Ensembl', and 'Fetch data with the Ensembl API'.
- Ensembl headlines:** A section titled 'Release 42 (December 2006)' listing new species and assemblies, such as 'New species - Duck-billed Platypus', 'New Dog assembly and genebuild', 'New Chicken assembly and genebuild', and 'New Human Ensembl Vega'.
- Popular genomes:** A list of popular genomes including 'Homo sapiens', 'Mus musculus', and 'Danio rerio'. Below this is a list of 'More genomes' with various species names and their Ensembl IDs.
- Other Ensembl websites:** A list of other Ensembl websites including 'archive - past releases of Ensembl', 'VEGA - Vertebrate Genome Annotation', 'EBI Genome Reviews database', and 'Trace server'.

On the left side, there is a sidebar with navigation links such as 'Your Ensembl', 'Help & Documentation', 'Select a species', and 'Ensembl Archive'. At the bottom of the page, there is a footer with the text '© 2007 HTI / EBI. Ensembl is available to download for public use - please see the code licence for details.'

e!
E
N
S
E
M
B
L

Ensembl - Project

- **Joint project**
 - EMBL – European Bioinformatics Institute (EBI)
 - Wellcome Trust Sanger Institute
- **Produce accurate, automatic genome annotation**
- **Focused on selected eukaryotic genomes**
- **Integrate external (distributed) biological data**
- **Presentation of the analysis to all via the Web at**
<http://www.ensembl.org>
- **Open distribution of the analysis the community**
- **Development of open, collaborative software (databases and APIs)**

e!
Ensembl

Beyond classical *ab initio* gene prediction

- Ensembl automatic gene prediction relies on homology 'supporting evidence' to avoid overprediction.
- Classical *ab initio* gene prediction (eg *GENSCAN*) relies partly on global statistics of protein coding potentials, not used in the cell
- Genes are just a series of short signals
 - Transcription start site
 - Translation start site
 - 5' & 3' Intron splicing signals
 - Termination signals
- Short signal sequences difficult to recognise over background noise in large genomes

Ensembl

Ensembl 42 Prel species










Popular genomes

 **Rattus norvegicus**
R6SC 3.4

More genomes

- ▶ **Aedes aegypti** AaegL1
- ▶ **Anopheles gambiae** AgamP3
- ▶ **Bos taurus** Btau 2.0
- ▶ **Caenorhabditis elegans** WS160
- ▶ **Canis familiaris** CanFam 2.0 **UPDATED!**
- ▶ **Ciona intestinalis** JGI 2
- ▶ **Ciona savignyi** CSAV 2.0
- ▶ **Danio rerio** Zv6
- ▶ **Dasypus novemcinctus** ARMA
- ▶ **Drosophila melanogaster** BDGP 4.3
- ▶ **Echinops telfairi** TENREC
- ▶ **Gallus gallus** WASHUC2 **UPDATED!**
- ▶ **Gasterosteus aculeatus** BROAD S1
- ▶ **Homo sapiens** NCBI 36
- ▶ **Loxodonta africana** BROAD E1
- ▶ **Macaca mulatta** MMUL 1.0
- ▶ **Monodelphis domestica** MonDom 4.0
- ▶ **Mus musculus** NCBI m36
- ▶ **Ornithorhynchus anatinus** Dana-5.0 **NEW!**
- ▶ **Oryctolagus cuniculus** RABBIT
- ▶ **Oryzias latipes** HdR
- ▶ **Pan troglodytes** PanTro 2.1
- ▶ **Saccharomyces cerevisiae** SGD1.01 **UPDATED!**
- ▶ **Takifugu rubripes** FUGU 4.0
- ▶ **Tetraodon nigroviridis** TETRAODON 7
- ▶ **Xenopus tropicalis** JGI 4.1

Prel species

-  **Cavia porcellus**
cavPor1 **NEW!**
-  **Erinaceus europaeus**
eriEur1
-  **Felis catus**
CAT
-  **Myotis lucifugus**
MICROBAT 1
-  **Otolemur garnettii**
BUSHBABY 1
-  **Sorex araneus**
soriAra1
-  **Spermophilus tridecemlineatus**
speTri1 **NEW!**
-  **Sus scrofa**
[clone map]
-  **Tupaia belangeri**
tupBel1 **NEW!**

DAS Registry

<http://www.dasregistry.org>

[[home](#) | [list services](#) | [validate](#) | [register new](#) | [statistics](#) | [help](#)]

AVAILABLE DAS SERVICES

available DAS services

organism:	authority:	type:	capability:	label:
any	any	any	any	ENSEMBL

id	clients nickname	status	capabilities	coordinateSystem	description
DS_109	uniprot		sequence features	UniProt,Protein Sequence	The UniProt DAS Reference Server serves both sequence and feature data from UniProt/Swiss-Prot and UniProt/TrEMBL. In addition to this, feature links to InterPro for all the proteins in UniProt are provided, including basic information (position, id and a brief description) and internet addresses to the relevant data source. The following kinds of protein ID / Accession number can be used: UniProt Accession numbers e.g. O35502 Swiss-Prot IDs e.g. A4_Human TrEMBL IDs e.g. Q12368 IPI IDs e.g. IPI00015171 UniParc IDs e.g. UPI0000125656 For a full description of the service, visit http://www.ebi.ac.uk/uniprot-das/ go to site



UniProt DAS Services at the EBI

The distributed annotation system (DAS) is a client-server system in which a single client integrates information from multiple servers. The UniProt DAS server allows researchers to show their research results, for example identified peptides or signal sequences, on the UniProt reference sequence server, in the context of UniProt/Swiss-Prot annotation.

The UniProt Protein DAS Reference Server

- ◆ [Access the UniProt DAS Server](#)

The UniProt DAS Reference Server serves both sequence and feature data from UniProt/Swiss-Prot and UniProt/TrEMBL. In addition to this, feature links to InterPro for all the proteins in UniProt are provided, including basic information (position, id and a brief description) and internet addresses to the relevant data source.

The server can be queried using any of the following:

- ◆ *UniProt Accession numbers e.g. O35502 [sequence](#), [features](#)
- ◆ *Swiss-Prot ID's e.g. A4_Human [sequence](#), [features](#)
- ◆ *TrEMBL ID's e.g. Q12368 [sequence](#), [features](#)
- ◆ IPI ID's e.g. IPI00015171 [sequence](#)
- ◆ UniParc ID's e.g. UPI0000125656 [sequence](#)

All of these kinds of identifier will return sequence information. Identifier types marked with an * will also return UniProt and InterPro feature hits.

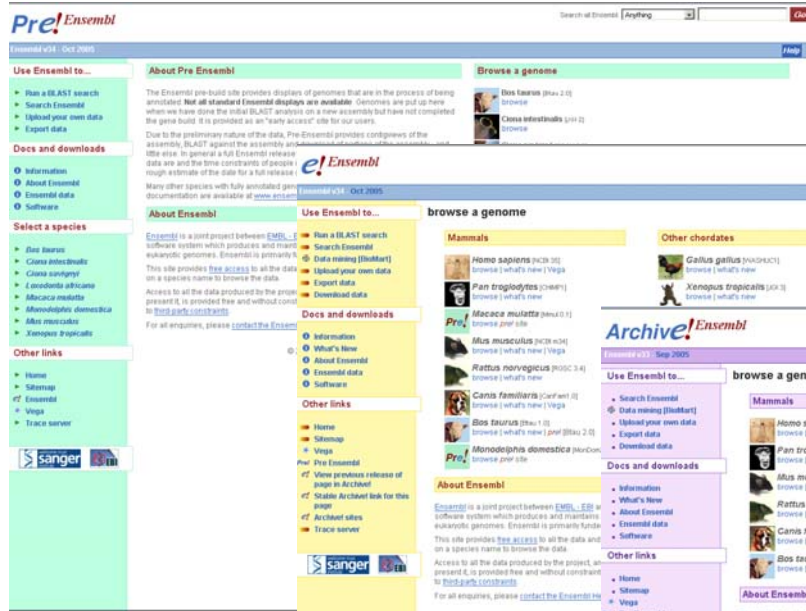
For further details of the UniProt DAS server click on the link above that will take you to the UniProt DAS home page. This includes example requests using the DAS protocol for real proteins that can be found in UniProt.

omo sapiens	ensembl genebuild NCBI36 CDS go to site
sequence	Ensembl exon boundary and snp data projected on UniProt sequence. Based on the COMPARA ENSP-UniProt mapping.
	Associated directly or via orthologues or protein families

e!
Ensembl

Pre! and Archive! sites

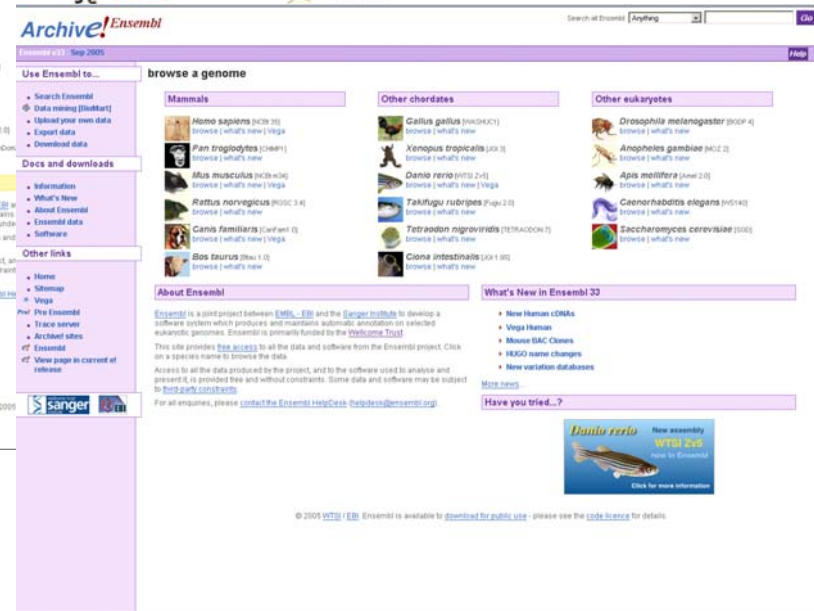
<http://pre.ensembl.org>



The screenshot shows the Pre! Ensembl website. At the top, there is a search bar with the text "Search of Ensembl [Anything] Go". Below the search bar, the page is divided into several sections. On the left, there is a navigation menu with "Use Ensembl to..." (containing links for BLAST search, upload data, export data), "Docs and downloads" (information, Ensembl data, software), "Select a species" (listing various species like Bos taurus, Canis familiaris, etc.), and "Other links" (Home, Sitemap, Vega, Pre Ensembl, Archive link, Trace server). The main content area features "About Pre-Ensembl" (explaining the pre-build site's purpose), "Browse a genome" (with a search bar and a list of species like Bos taurus, Canis familiaris, etc.), and "Use Ensembl to..." (with links for BLAST search, data mining, upload data, export data, download data). At the bottom, there are "Other links" and "What's New" sections.

<http://www.ensembl.org>

<http://archive.ensembl.org>



The screenshot shows the Archive! Ensembl website. It features a search bar at the top with "Search of Ensembl [Anything] Go". The main content area is titled "Archive! Ensembl" and includes "Use Ensembl to..." (with links for BLAST search, data mining, upload data, export data, download data), "Docs and downloads" (information, what's new, Ensembl data, software), "Other links" (Home, Sitemap, Vega, Pre Ensembl, Archive link, Trace server), and "What's New in Ensembl 33" (listing new human cDNAs, Vega names, Mouse IAC Clones, HSKO name changes, and new variation databases). The "Browse a genome" section is prominent, with a search bar and a list of species including Homo sapiens, Pan troglodytes, Mus musculus, Rattus norvegicus, Canis familiaris, Bos taurus, Gallus gallus, Xenopus tropicalis, Drosophila melanogaster, Anopheles gambiae, Apis mellifera, Caenorhabditis elegans, and Saccharomyces cerevisiae. A "Have you tried...?" section at the bottom highlights the "Danio rerio" assembly.

Open source open standards



- Object model
 - **standard interface makes it easy for others to build custom applications on top of Ensembl data**
- Open discussion of design (ensembl-dev@ebi.ac.uk)
- Most major pharma and many academics represented on mailing list and code is being actively developed externally
- Ensembl locally
 - **Both industry & academia**

Ensembl – Open source



Central *Aspergillus*
Data REpository



Cattle Public EST Contig Browser

About Cattle ContigBrowser

Sigenae Species

The Cattle EST contig browser is an **INRA** project aimed to produce and maintain an automatic annotation of Cattle EST libraries.

Cattle Bos taurus
Chicken Gallus gallus

Current annotation of Cattle **GRAMENE** *Genome_browser*

Ageneae Project's We Agena ESTs Project.

Find in Genomes

Browser Links

- Genomes Home
- Genomes Help
- Switch Species

Find anything

Search

Search of Ensembl: Anything

NASC e! AtEnsembl

Use Ensembl to...

- Run a BLAST search
- Search Ensembl
- Hybridize your data
- Export data

Docs and downloads

- Information
- What's New
- About Ensembl
- Ensembl data
- Software

Other links

- Home
- SiteMap

Arabidopsis genomes

Arabidopsis thaliana [TAIR3] (update) what's new

Parallel display of TAIR3 (pepares TGFS) and MIPS annotations, and all other features such as inserts stocked at NASC, Affymetrix probes and other alignment data

Arabidopsis thaliana [MIPS] (update) what's new

MIPS assembly and annotation with its other features

What's New in AtEnsembl 35

- New TAIR3 annotations replaces TGFS (Arabidopsis thaliana)
- CAZMs and RFLPs to genes re-annotated by TAIR3 genome (Arabidopsis thaliana)
- Protein, EST, and cDNA alignments updated (Arabidopsis thaliana)

More news...

Have you tried...?

InsertWatch

Registers a sequence or gene name with InsertWatch and receive automatic updates of any new inserts as they are released.

Aspergillus nidulans (FGSC A4)

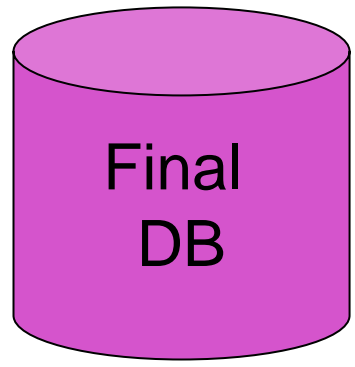
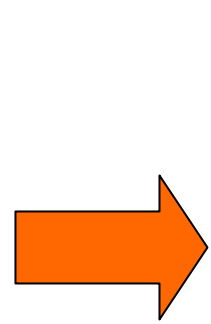
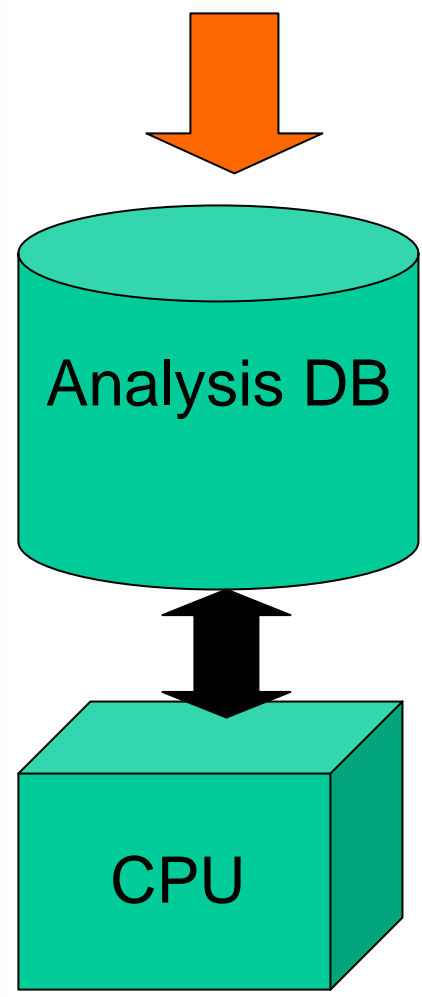
UPDATED! August 2006

- CADF
- Asper
- Asper
- genom
- sequ
- CADF
- ▶ E
- ▶ F
- ▶ Ar
- Info
- ▶ At
- ▶ M:
- ▶ Si
- Refer
- Plea
- J.E
- For al

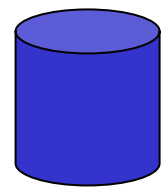
Making genomes useful

- **Interpretation**
 - Where are the interesting parts of the genome?
 - What do they do?
 - How are they related to elements in other genomes?
- **Access**
 - for bench biologists
 - for non-programming mid-scale groups
 - for good programming groups

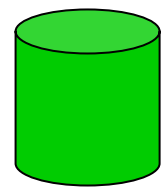
Ensembl



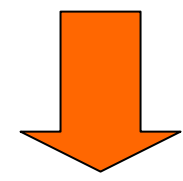
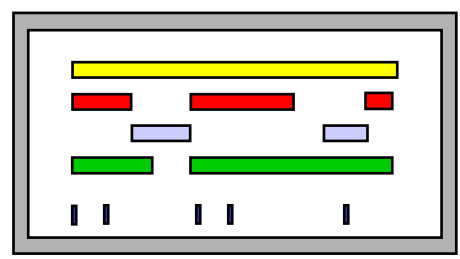
Supporting Databases



SNP



Manual Annotation



MartView

Genome browsing ***why present the whole genome?***

- Explore what is in a chromosome region
- See features in and around a specific gene
- Search & retrieve across the whole genome
- Investigate genome organization
- Compare to other genomes

Basic Genome Annotation

- **Genes**
 - **Genomic location**
 - **Gene model structures**
 - Exons
 - Introns
 - UTRs
 - **Transcript(s)**
 - Pseudogenes
 - Non-coding RNA
 - **Protein(s)**
 - **Links to other sources of information**

Advanced Genome Annotation

- **Cytogenetic bands**
- **Polymorphic markers**
 - **Sequence Tagged Sites (STS)**
- **Genetic variation**
 - **Single Nucleotide Polymorphisms (SNPs)**
 - **Deletion-Insertion Polymorphisms (DIPs)**
 - **Short Tandem Repeats (STRs)**
- **Repetitive sequences**
- **Expressed Sequence Tags (ESTs)**
- **cDNAs or mRNAs from related species**
- **Regions of sequence homology**

How to get started



- **Species homepage**
- **Map View**
- **Text search**
- **BLAST**
- **SSAHA**

Homepage

e!Ensembl

Ensembl release 42 - Dec 2006 (ecs3:3307 - homo_sapiens_core_42_36d) HOME · BLAST · BIOMART · SITEMAP **HELP**

Your Ensembl

-  [Show account](#) - [Log out](#)
-  [Save bookmark](#)




Healthchecks

-  [Health checks](#)
-  [Old Health checks](#)

Help & Documentation

-  [Table of Contents](#)
-  [Helpdesk](#)
-  [About Ensembl](#)
-  [Downloading data](#)
-  [Displaying your own data](#)
-  [Ensembl software](#)

Select a species

-  [Mammals](#)
-  [Other chordates](#)
-  [Other eukaryotes](#)

Ensembl Archive

-  [View previous release of page in Archive!](#)
-  [Stable Archive! link for this page](#)







Search Ensembl

Search: for

e.g. mouse chromosome 2 or X:10000..20000 or human gene BRCA2

Ensembl tools

-  **Start a sequence search** →
Search Ensembl for nucleotide and peptide sequences with BLAST and SSAHA.
-  **Mine Ensembl with BioMart** →
Cross-reference Ensembl datasets with BioMart, a powerful data-mining tool.
-  **Customise Your Ensembl** →
Register with Ensembl to bookmark your favourite pages, customise your home page and much more!
-  **Fetch data with the Ensembl API** →
Learn how to extract data from the public Ensembl database with this tutorial.

You are logged in as **Giulietta**. Go to your account home page or [log out](#).

Ensembl headlines: Release 42 (December 2006)

-  **New - User accounts** (all species)
-  **New species - Duck-billed Platypus** (*Ornithorhynchus anatinus*)
-  **New Dog assembly and genebuild** (*Canis familiaris*)
-  **New Chicken assembly and genebuild** (*Gallus gallus*)
-  **New Human Ensembl-Vega** (*Homo sapiens*)

[More news...](#)

Go to your account to customise this news panel

About Ensembl

Ensembl is a joint project between EMBL - EBI and the Sanger Institute to develop a software system which produces and maintains automatic annotation on selected eukaryotic genomes. Ensembl is primarily funded by the Wellcome Trust.

This site provides free access to all the data and software from the Ensembl project. Click on a species name to browse the data.

Access to all the data produced by the project, and to the software used to analyse and present it, is provided free and without constraints. Some data and software may be subject to third-party constraints.

For all enquiries, please contact the [Ensembl HelpDesk \(helpdesk@ensembl.org\)](mailto:helpdesk@ensembl.org).

Other Ensembl websites

- ▶ [archive!](#) - past releases of Ensembl
- ▶ [VEGA](#) - Vertebrate Genome Annotation
- ▶ [EBI Genome Reviews](#) database - mainly archaea and bacteria.
- ▶ [Trace server](#)

Other sites using Ensembl software...

Ensembl 42 [Prel species](#)

Popular genomes - [Reorder](#)

-  **Homo sapiens**
NCBI 36 | Vega
-  **Ciona intestinalis**
JGI 2
-  **Mus musculus**
NCBI m38 | Vega
-  **Danio rerio**
Zv6 | Vega

More genomes - [Reorder](#)

- ▶ [Aedes aegypti](#) AaegL1
- ▶ [Anopheles gambiae](#) AgamP3
- ▶ [Bos taurus](#) Btau 2.0
- ▶ [Caenorhabditis elegans](#) WS160
- ▶ [Canis familiaris](#) CanFam 2.0 **UPDATED!**
- ▶ [Ciona savignyi](#) CSAV 2.0
- ▶ [Dasyypus novemcinctus](#) ARMA
- ▶ [Drosophila melanogaster](#) BDGP 4.3
- ▶ [Echinops telfairi](#) TENREC
- ▶ [Gallus gallus](#) WASHUC2 **UPDATED!**
- ▶ [Gasterosteus aculeatus](#) BROAD S1
- ▶ [Loxodonta africana](#) BROAD E1
- ▶ [Macaca mulatta](#) MMUL 1.0
- ▶ [Monodelphis domestica](#) MonDom 4.0
- ▶ [Ornithorhynchus anatinus](#) Oana-5.0 **NEW!**
- ▶ [Oryctolagus cuniculus](#) RABBIT
- ▶ [Oryzias latipes](#) HfR
- ▶ [Pan troglodytes](#) PanTro 2.1
- ▶ [Rattus norvegicus](#) RGSC 3.4
- ▶ [Saccharomyces cerevisiae](#) SGD1.01 **UPDATED!**
- ▶ [Takifugu rubripes](#) FUGU 4.0
- ▶ [Tetraodon nigroviridis](#) TETRAODON 7
- ▶ [Xenopus tropicalis](#) JGI 4.1



Ensembl

MapView

e! **Ensembl** Human *MapView*

Search e!Human:

e.g. [5_22](#)

Ensembl v32 - May 2005 [Help](#)

Chromosome X

- View Chromosome X
- View Chr X Synteny
- Map your data onto this chromosome
- Browse OMIM diseases

Use Ensembl to...

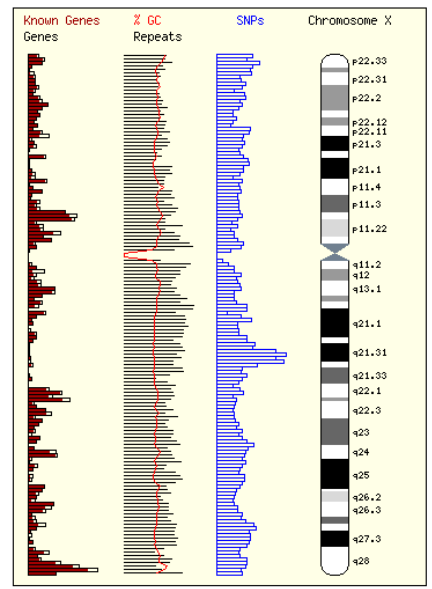
- Run a BLAST search
- Search Ensembl
- Data mining [BioMart]
- Upload your own data
- Download data
- Export data

Docs and downloads

- Information
- About Ensembl
- Using Ensembl data
- Software

Other links

- Home
- Sitemap
- What's New
- Stable (archive) link for this page



Click on the image above to zoom into that point

Chromosome X

Length: 154,824,264 bps

Gene Count: 931

Known Gene Count: 766

PseudoGene Count: 380

SNP Count: 320,997

Change Chromosome

Chromosome

Fields marked with * are required

Jump to ContigView

Choose two features from this chromosome as anchor points and display the region between them. Both features must be mapped to the current Ensembl golden tiling path. If you select "None" for the second feature, the display will be based around the first feature.

Please note that if you select widely spaced features there may be an significant delay while Ensembl builds the DNA display.

Region

From (type):

To (type):

Context

Bp downstream

Bp upstream

Fields marked with * are required

BLAST and SSAHA

new **SETUP** <> CONFIG <> RESULTS <> DISPLAY

refresh **Online Help**

Summary

- ▶ **setup** Not yet initialised
- ▶ **configure** Not yet initialised
- ▶ **results** Not yet initialised
- ▶ **display** Not yet initialised

Enter the Query Sequence

Either Paste sequences (max 30) in FASTA or plain text:

```
ACCACCCCTACCCAGGCTACAGAACTACATGGCAGCCAGAAGCATGCGGAAGC
CTGCCTCTCCTTGAACAACACATCCGGATACGTTCTAAAGAATCTGAGAAAA
ACGACTCGATTGGCTGACGCCAAACCGTAATAAATTCACGAGTTTTATCGTATC
ATTACACCATAT
```

Or Upload a file containing one or more FASTA sequences
Choose File no file selected

Or Enter a sequence ID or accession (EMBL, UniProt, RefSeq)
 Retrieve

Or Enter an existing ticket ID:
 Retrieve

dna queries
 peptide queries

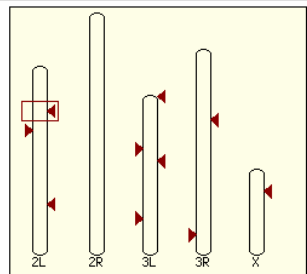
Select the databases to search against

Select species:
Use 'ctrl' key to select multiple species

- Apis_mellifera
- Caenorhabditis_elegans
- Canis_familiaris

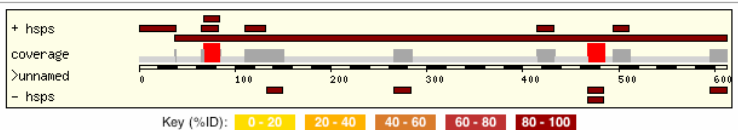
dna database
 peptide database

Alignment Locations vs. Karyotype (click arrow to hide)



Key (%ID): 0-20 20-40 40-60 60-80 80-100

Alignment Locations vs. Query (click arrow to hide)



Key (%ID): 0-20 20-40 40-60 60-80 80-100

Alignment Summary (click arrow to hide)

Select rows to include in table, and type of sort (Use the 'ctrl' key to select multiples)

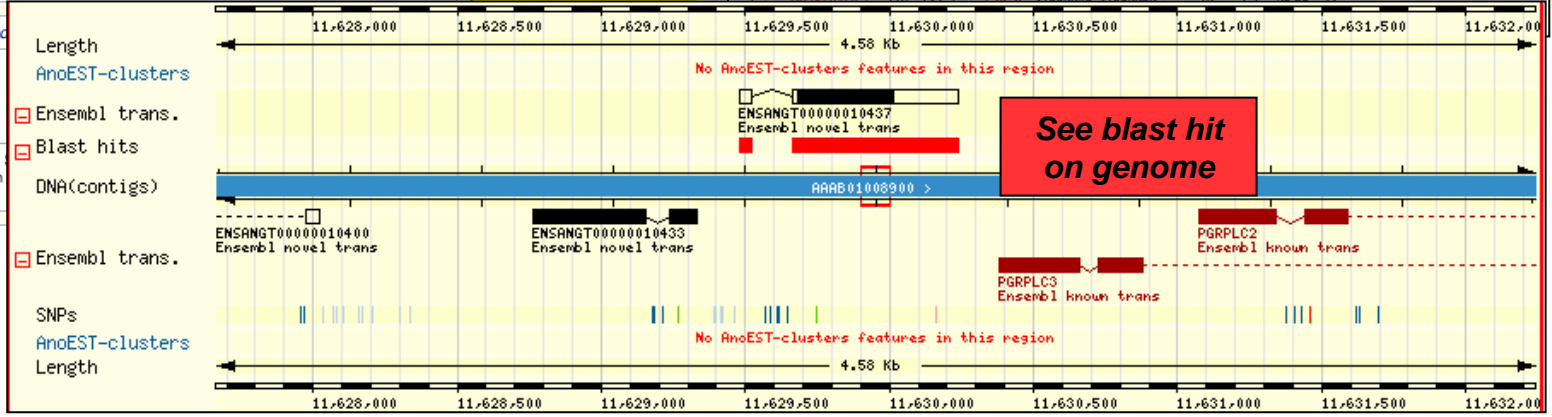
refresh

Query	Subject	Chromosome	Scaffold	Chunk	Stats	Sort By
Name	Start	Name	_off_ Name	_off_ Name	E-val	>Score
Start	Start	Start	Start	Start	P-val	<E-val
End	End	End	End	End	%ID	>E-val
Ori	Ori	Ori	Ori	Ori	Length	<P-val
Links	Query	Chromosome	Stats			
	Start End Ori	Name Start End	Ori Score E-val %ID Length			
	[A] [S] [G] [C]					
	37 612 +	Chr:2L 11629664 11630239 +	576 0. 100.00 576			
	[A] [S] [G] [C]					
	1 38 +	Chr:2L 11629483 11629520 +	38 0. 100.00 38			
	[A] [S] [G] [C]					
	65 83 +	Chr:3R 48244909 48244927 +	19 1.2 100.00 19			
	[A] [S] [G] [C]					
	110 120	Chr:3R 48244909 48244927 +	19 1.2 100.00 19			

Select the Search Tool

BLASTN
 SSAHA
 TBLASTX

Optimise search



Regions, maps and markers

ContigView

CytoView

SyntenView

MultiContigView

MarkerView

SNPView

GeneSNPView

Chromosome X
138,229,875 - 138,686,223

- View of Chromosome X
- Graphical view
- Graphical overview
- Export sequence as FASTA
- Export EMBL file
- Export Gene info in region
- Export SNP info in region
- Export Vega info in region
- View alongside ...
- View Syntenic regions ...
- View region in NCBI browser
- View region in UCSC browser

Use Ensembl to...

- Run a BLAST search
- Search Ensembl
- Data mining [BioMart]
- Upload your own data
- Download data
- Export data

Docs and downloads

- Information
- About Ensembl
- Using Ensembl data
- Software

Other links

- Home
- Sitemap
- What's New
- Stable (archive) link for this page

Ensembl ContigView

The screenshot displays the Ensembl Human ContigView interface for Chromosome X, specifically the region 138,229,875 - 138,686,223. The interface is divided into several sections:

- Overview:** Shows a high-level view of the genomic region with various tracks including Gene, Transcription Start Site (TSS), and Repeat.
- Detailed view:** Provides a more granular view of the genomic region, showing individual contigs and their associated features. The contigs are color-coded (yellow, green, red) and labeled with their IDs (e.g., chrX:138,229,875-138,686,223). The detailed view includes tracks for Gene, Transcription Start Site (TSS), Repeat, and other genomic features.
- Basepair view:** Shows the raw sequence data in a color-coded format, with the sequence aligned to the reference genome. The basepair view includes tracks for Gene, Transcription Start Site (TSS), Repeat, and other genomic features.

The interface also includes a navigation pane on the left with options for viewing the region in different ways (e.g., Graphical view, Graphical overview, Export sequence as FASTA, Export EMBL file, Export Gene info in region, Export SNP info in region, Export Vega info in region, View alongside ..., View Syntenic regions ..., View region in NCBI browser, View region in UCSC browser).

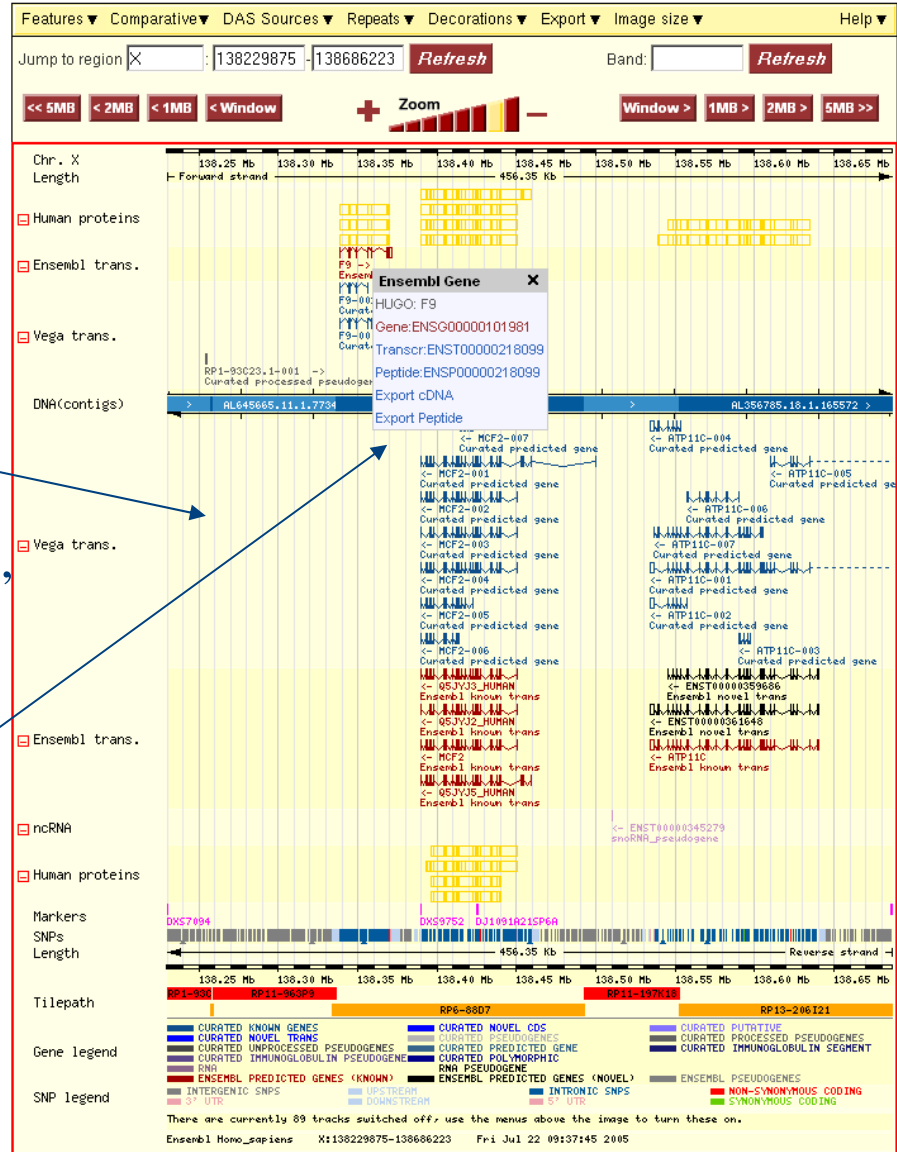
ContigView

close-up

Detailed view

Transcripts
red & black
(Ensembl predictions)
Blue (Vega) & gold (HAVANA,
only in human)

Pop-up
menu

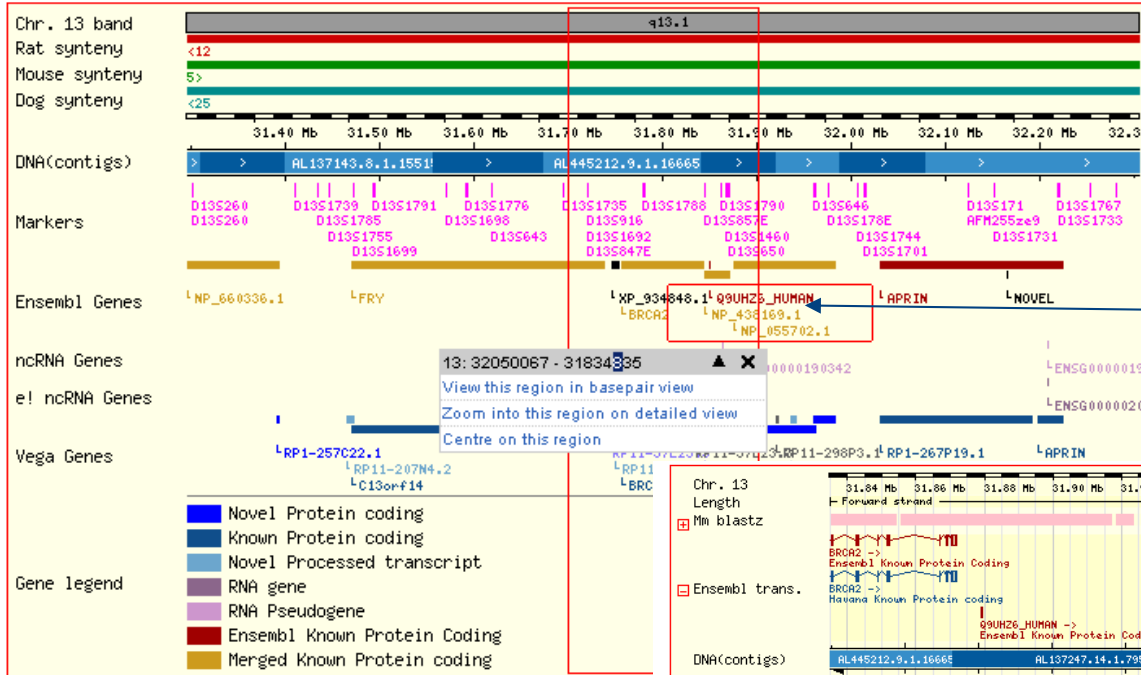


Basepair view

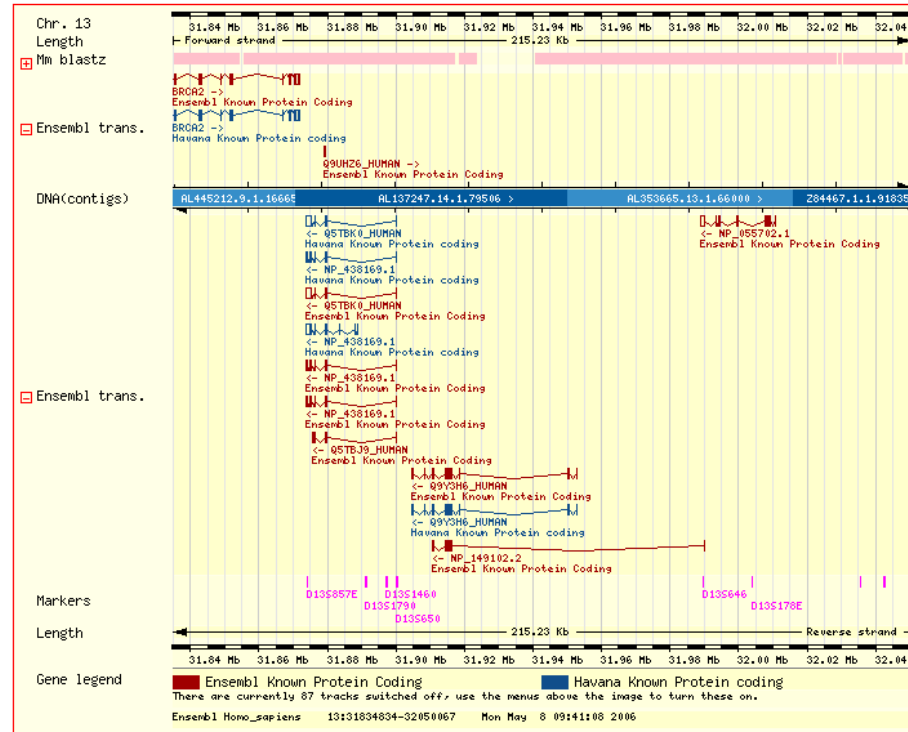


Contig View - Navigation

Overview

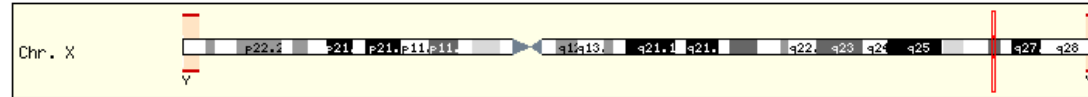


Click and drag mouse to select region

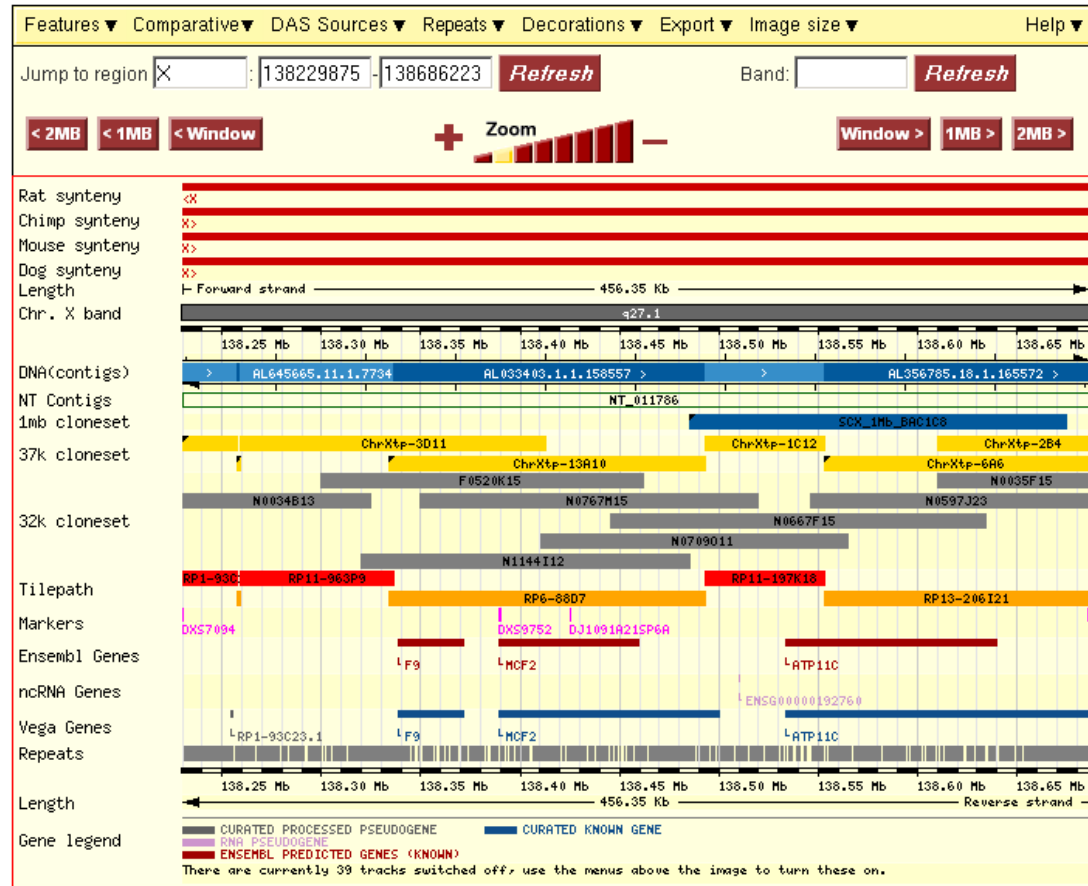


CytoView

☐ Chromosome X



☐ Detailed view



☐ Export data

Select Set of features to render

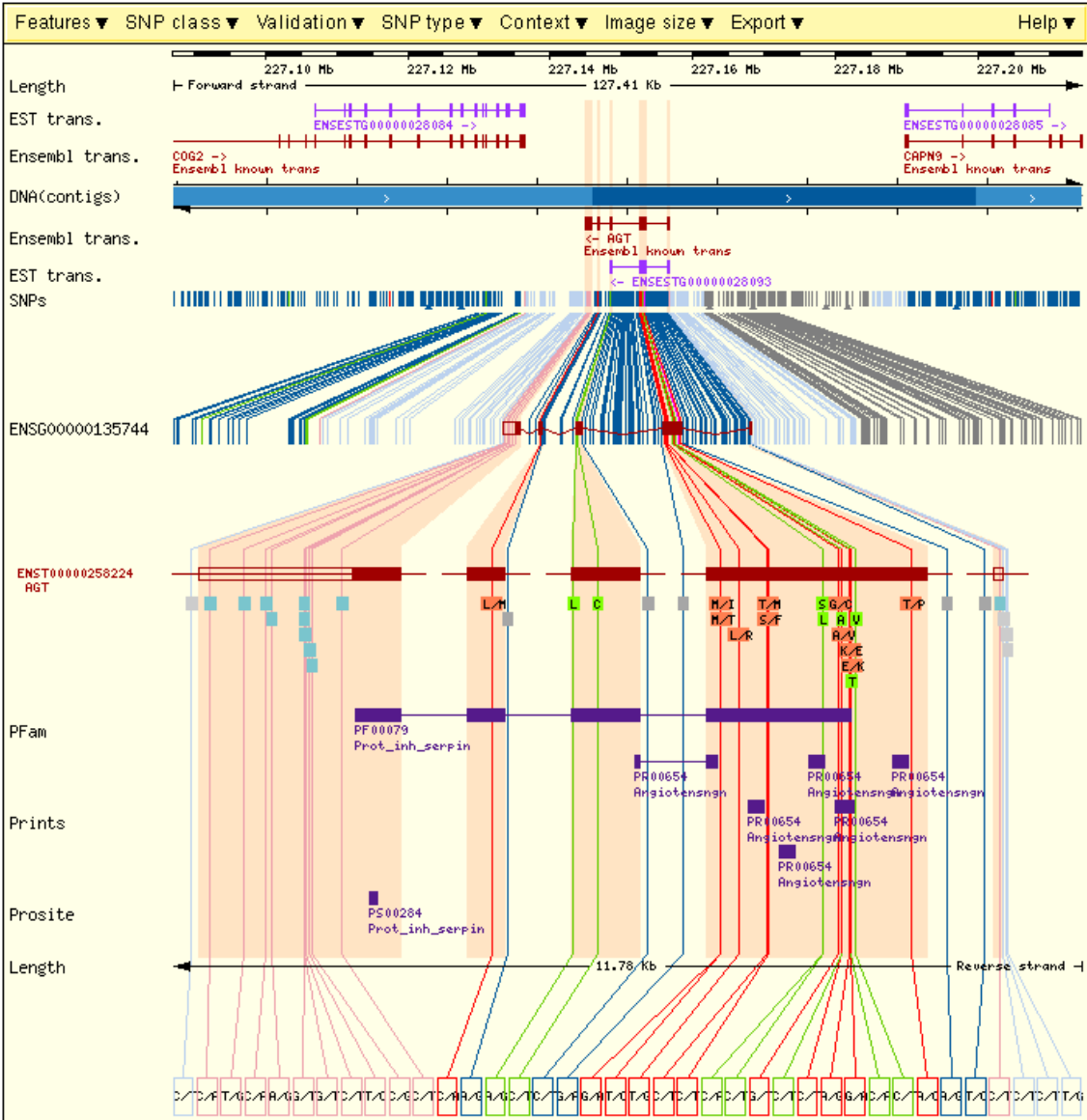
Output format

Select type to export

Export

GeneSNP View

Variations in region of gene ENSG00000135744



SNPView

SNP Report

SNP	rs699 (dbSNP124)			
Synonyms	HGvbase SNP00000510			
Alleles	T/C (ambiguity code: Y)			
Validation status	Proven by cluster, frequency, doublehit (SNP tested and validated by a non-computational method).			
Linkage disequilibrium data	Links to LDview per population: PERLEGEN_AFD_CHN_PANEL PERLEGEN_AFD_AFR_PANEL	CSHL-HAPMAP_HapMap-YRI CSHL-HAPMAP_HapMap-CEU	CSHL-HAPMAP_HapMap-JPT PERLEGEN_AFD_EUR_PANEL	CSHL-HAPMAP_HapMap-HCB
Sequence region	<pre> CCCTCCAGGCTGTACAGGCGCTGCTAGTGGCCAGGGCAGGGCGTATAGCCAGGGCCAGC TGTCTGCTCCACGGTGTGGGGCTGTTCACAGCCCAAGGCTGACCTGAAGCAGCCGCT TTGTGACGGGCTGGCTCTATACCCCTGTGGCTCTCCACGCTCTCTGGAGTTCACAG AACTGGATTCTCTGTGAGAGATTGACAGGTTGATCCGGCTGTGACAGATGGAAAG CTGGCTGCTCCCTGATGGGAGCAGCTGTGGACAGCAGCCTGGCTTCAACACTACTCC ACTTCCAAGATTAAGGCAAACTCTCTGCTGGCTCTGGCCCTAGGACTAGTATCCAAATGT GTAGCTGAGTACAGCCAGTCAAGCCTTGGAGATGGGAGGGGGCAGCCTGGGAGACATCC TGGTGACACCTTGGAGATGGGAGATGGCTGGCTCCCTGGCTCCCTGGATGGGGCT CCACTCTGGACCTGCTCCCTACTATGTG (SNP highlight) </pre>			

Genotype frequencies per population

Population	Genotypes T/T	Genotypes C/T	Genotypes C/C	Description
PARC-PARC-EUROPEAN-PANEL	0.217	0.652	0.130	This population of 23 individuals (12 male/11 female) is composed of CEPH Parent DNA available from the Coriell Cell Repository (CCR). CCR ID CEPH PEDIGREE ID PARC ID SEX:NA12580 French 12 ED01 M,NA12547 French 66 E002 M,NA10945 Utah 1424 ED03 M,NA10893 Uta
WIPGA POP-WIPGA-12-05-2003	0.179	0.436	0.385	Each assay with this population used a subset of the individuals listed below. To determine which specific individuals were used for a given polymorphism, please consult the genotypes associated with the polymorphism. CL420 female,Caucasian,Coriell,CL388
PARC-PARC-AFRICAN-PANEL	0.042	0.250	0.708	This population of 24 individuals (12 male/12 female) is composed of DNA available from the Coriell Cell Repository. These individuals were selected from the human variation panel of 50 African Americans (H050AA). CCR ID SEX:NA17101 D001 M,NA17102

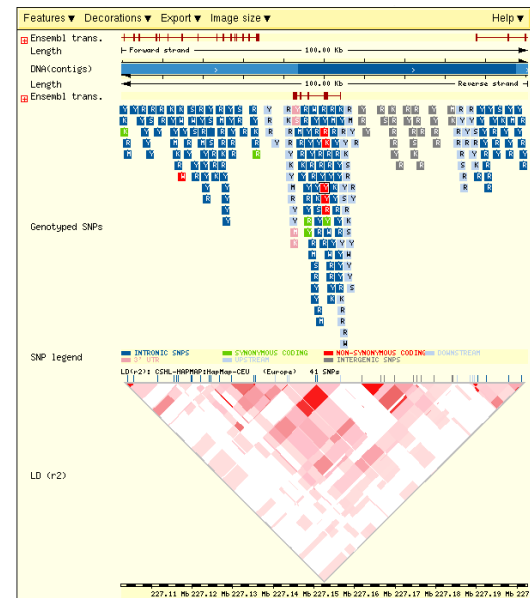
Allele frequencies per population

Population	Alleles T	Alleles C	Description
ARAVINDA.HYP1	0.100	0.900	80 chromosomes are from Harare, Zimbabwe; Samples represent individuals from both extremes (2.5th percentiles) of a community blood pressure distribution.
PARC-PARC-EUROPEAN-PANEL	0.543	0.457	This population of 23 individuals (12 male/11 female) is composed of CEPH Parent DNA available from the Coriell Cell Repository (CCR). CCR ID CEPH PEDIGREE ID PARC ID SEX:NA12580 French 12 ED01 M,NA12547 French 66 E002 M,NA10845 Utah 1424 ED03 M,NA10893 Uta
YUSUKE.JBIC-allele	0.188	0.812	752 anonymous unrelated Japanese volunteers; Nation:Japan
PHARMGB.PARC-PA128275074	0.351	0.649	-
CGAP-GAI POOLED CEPH	0.572	0.428	Pooled genomic DNA from 94 unrelated CEPH individuals (caucasian)
WIPGA POP-WIPGA-12-05-2003	0.397	0.603	Each assay with this population used a subset of the individuals listed below. To determine which specific individuals were used for a given polymorphism, please consult the genotypes associated with the polymorphism. CL420 female,Caucasian,Coriell,CL388
PARC-PARC-AFRICAN-PANEL	0.167	0.833	This population of 24 individuals (12 male/12 female) is composed of DNA available from the Coriell Cell Repository. These individuals were selected from the human variation panel of 50 African Americans (H050AA). CCR ID SEX:NA17101 D001 M,NA17102

SNP rs699 is located in the following transcripts

Genomic location (strand)	Transcript: start-end	Gene/SNP view link	Translation: start-end	Peptide allele	Consequence
1:227152529-227152529 (-1)	ENST00000258224 : 842-842	SNP in gene context	ENSP00000258224 : 268-268	MT	NON_SYNONYMOUS_CODING

SNP Context - chromosome 1 227152529



MarkerView

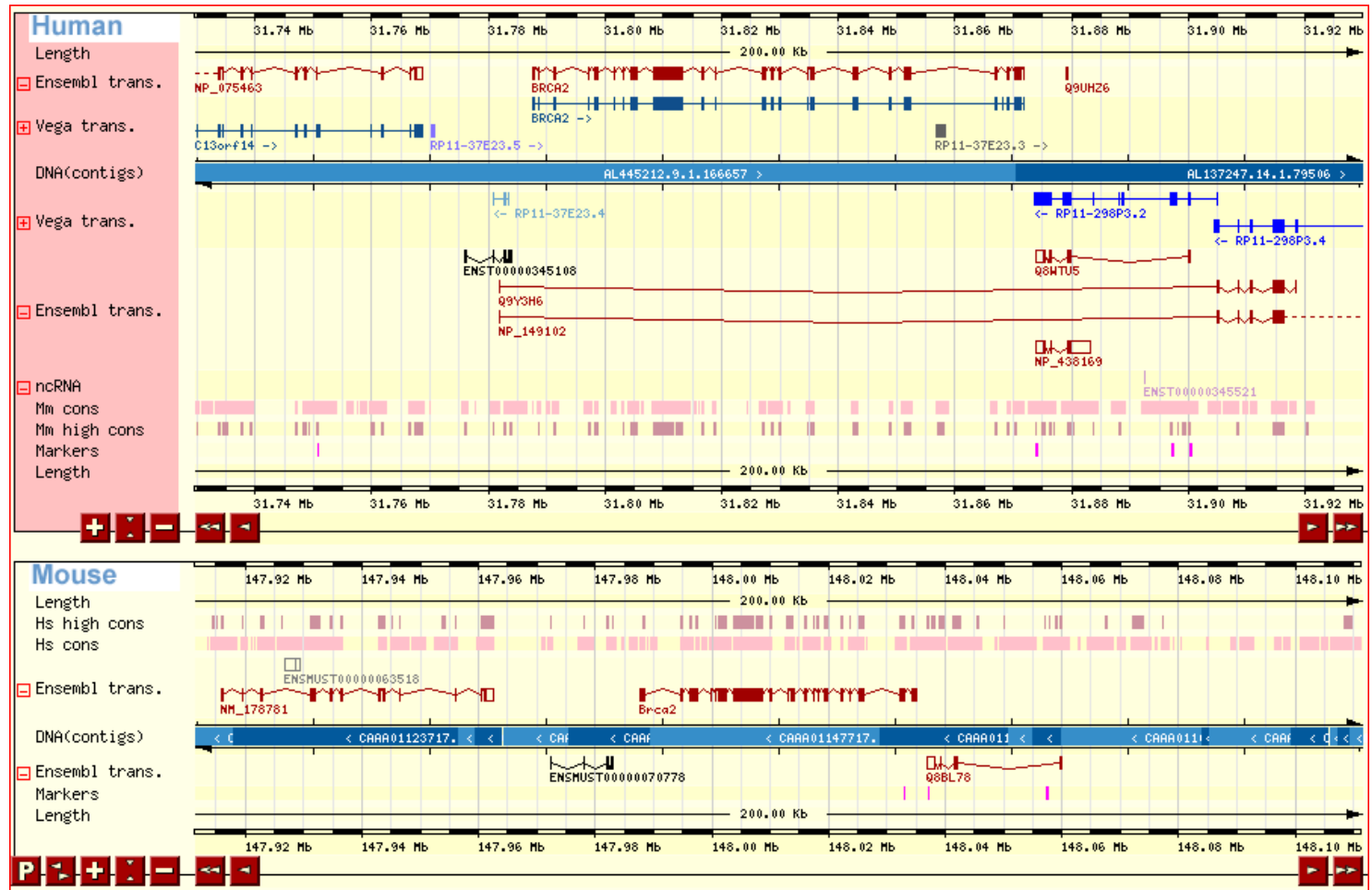
Chromosome Map Marker DXS9752

Marker Source	82913 (database: unists)		
Marker Location	Basepairs 138389786 - 138390044 on chromosome X [Export data]		
Marker Synonyms	Gdb : GDB:737728 GDB:738733 Genbank : G13636 Other : SHGC-11927 DXS9752 RH8108		
Marker Primers	Expected Product Size 259	Left Primer TTTTCAGGTTAATGGACACGC	Right Primer CCATTTTGCAGCCGTAATTT

Marker DXS9752 map locations

Map Name	Synonym	Chromosome	Position	LOD Score
gm99g3	RH8108	X	4259	3.5

MultiContig View



Genes & gene products

GeneView

TransView

ExonView

ProteinView

FamilyView

GOView

Ensembl Gene View

ENSG00000101981

- Gene information
- Gene splice site image
- Gene variation info.
- Genomic sequence
- Export data
- Transcript information
- Exon information
- Peptide information

Chromosome X 138,338,415 - 138,371,137

- View of Chromosome X
- Graphical view
- Graphical overview
- Export information about region
- Export sequence as FASTA
- Export EMBL file
- Export Gene info in region
- Export SNP info in region
- Export Vega info in region

Use Ensembl to...

- Run a BLAST search
- Search Ensembl
- Data mining [BioMart]
- Upload your own data
- Download data
- Export data

Docs and downloads

- Information
- About Ensembl
- Using Ensembl data
- Software

Other links

- Home
- Sitemap
- What's New
- Stable (archive) link for this page



Ensembl Gene Report for ENSG00000101981

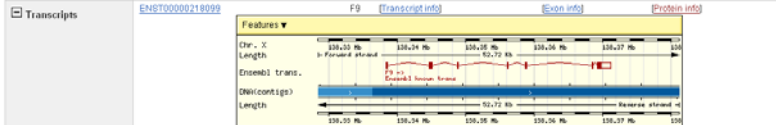
Gene [F9 \(HUGO ID\)](#) (to view all Ensembl genes linked to the name [@f9.html](#))
 This gene is a member of the human CCDS set: [CCDS14666](#)

Ensembl Gene ID **ENSG00000101981**

Genomic Location This gene can be found on Chromosome X at location: [138,338,415-138,371,137](#)
 This start of this gene is located in [Contig AL033403.1](#) 159897

Description Coagulation factor IX precursor ([EC:3.4.21.27](#)) (Christmas factor) (Plasma thromboplastin component) (PTC) [Source: UniProt/SwissProt:P00730](#)

Prediction Method Genes were annotated by the Ensembl automatic analysis pipeline using either a GeneWise model from a human/vertebrate protein, a set of aligned human cDNAs followed by Genomewise for CDS prediction or from GenScan exons supported by protein, cDNA and EST evidence. GeneWise models are further combined with available aligned cDNAs to annotate UTRs.



- Orthologue Prediction** The following gene(s) have been identified as putative orthologues by reciprocal BLAST analysis:
- | Species | Type | dNDS | Gene Identifier |
|---------------------------|------|---------|---|
| <i>Canis familiaris</i> | UBRH | 0.22373 | ENSCAF00000001898 (F9A_CANFA) MultiCastView blast |
| <i>Bos taurus</i> | UBRH | 0.28398 | ENSBTCAC0000004093 (CF9P96E8_BOVIN) MultiCastView blast |
| <i>Danio rerio</i> | UBRH | | ENSDAR000000010027 (F9) MultiCastView blast |
| <i>Mus musculus</i> | UBRH | 0.18856 | ENSMUSG000000031138 (F9) MultiCastView blast |
| <i>Fugu rubripes</i> | UBRH | | SRRRFL02000001472924 (Novel Ensembl prediction) MultiCastView blast |
| <i>Gallus gallus</i> | UBRH | | ENGGAL00000008513 (NM_204347) MultiCastView blast |
| <i>Xenopus tropicalis</i> | UBRH | | ENSGXET000000003472 (Novel Ensembl prediction) MultiCastView blast |
| <i>Rattus norvegicus</i> | UBRH | 0.25220 | ENSRNOP00000003430 (Novel Ensembl prediction) MultiCastView blast |
| <i>Tetraodon lineolus</i> | UBRH | | GSTENOG0001200001 (GOSTEN0001200001) MultiCastView blast |
| | RHG | | GSTENOG0002125101 (GOSTEN0002125101) MultiCastView blast |
- UBRH* = (Ubiquitin) (Best Orthologous ODP)
MBLE = one of (Multiple) (Best Orthologous ODPs)
RHG = Reciprocal Hit based on Synteny around BRH
DNDS = Derived from Whole Genome Alignment

[View alignments of homologies.](#)
 UBRH = (Ubiquitin) (Best Orthologous ODP)
 MBLE = one of (Multiple) (Best Orthologous ODPs)
 RHG = Reciprocal Hit based on Synteny around BRH
 DNDS = Derived from Whole Genome Alignment

Disease Matches This Ensembl entry corresponds to the following OMIM disease identifiers:
Hemophilia B (3)
 [OMIM ID: 306900] - [View disease information](#)

Gene DAS Report

- DAS Sources**
- [AltSplice](#) (Alternative splice database)
 - [GAD](#) (Genetic Association Database)
 - [IRIGO_text](#) (PubMed text mining via IRIGO symbol)
 - Phenotypes** (Associated directly or via orthologues or protein families)
 - [Protonef](#) (Global classification of proteins into hierarchical clusters)
 - [Reactions](#) (Knowledgebase of biological processes)
- [Manage Sources](#)

Transcript ENST00000218099

Transcript [F9 \(HUGO ID\)](#) (to view all Ensembl genes linked to the name [@f9.html](#))
 This transcript is a member of the human CCDS set: [CCDS14666](#)

Transcript Information Exons: 8 **Transcript length:** 2,802 bps **Protein length:** 461 residues
[Further Transcript info](#) [Exon information](#) [Protein information](#)

- Similarity Matches** This Ensembl entry corresponds to the following database identifiers:
- AFVHG Focus:** [207218_at](#)
 - AFVHG U133 PLUS 2:** [207218_at](#)
 - AFVHG U133A:** [207218_at](#)
 - AFVHG U133A 2:** [207218_at](#)
 - AFVHG U95Av2:** [35556_at](#)
 - AFV U133 X3P:** [g10518507_sp_at](#)
 - CCDS:** [CCDS14666.1](#)
 - EMBL:** [AF536327](#) [\[tag|g|AL033403\]](#) [\[tag|AY226143\]](#) [\[tag|AY269424\]](#) [\[tag|AY269425\]](#) [\[tag|Z00128\]](#) [\[tag|Z00137\]](#) [\[tag|Z02040\]](#) [\[tag|Z02049\]](#) [\[tag|Z02051\]](#) [\[tag|Z02052\]](#) [\[tag|Z02053\]](#) [\[tag|Z02040\]](#) [\[tag|Z02051\]](#) [\[tag|Z02052\]](#) [\[tag|Z02053\]](#) [\[tag|Z02040\]](#) [\[tag|Z02051\]](#) [\[tag|Z02052\]](#) [\[tag|Z02053\]](#) [\[tag|Z02040\]](#) [\[tag|Z02051\]](#) [\[tag|Z02052\]](#) [\[tag|Z02053\]](#) [\[tag|Z02040\]](#) [\[tag|Z02051\]](#) [\[tag|Z02052\]](#) [\[tag|Z02053\]](#) [\[tag|Z02040\]](#) [\[tag|Z02051\]](#) [\[tag|Z02052\]](#) [\[tag|Z02053\]](#) [\[tag|Z02040\]](#) [\[tag|Z02051\]](#) [\[tag|Z02052\]](#) [\[tag|Z02053\]](#)
 - EntrezGene:** [2158](#)
 - IRIGO:** [Search GeneCards for F9](#)
 - RPK:** [IPI00266178.2](#) (Target Nid: 100, Query Nid: 100)
 - MMB:** [1345101](#) [134540](#) [388980](#)
 - PGC:** [1CFH1](#) [1CF7](#) [1EDM](#) [1B95](#) [1MOX](#) [18DZ](#)
 - Protein ID:** [AA51881.1](#) [\[tag|AA52023.1\]](#) [\[tag|AA52456.1\]](#) [\[tag|AA52783.1\]](#) [\[tag|AA56822.1\]](#) [\[tag|AA60726.1\]](#) [\[tag|AAD20588.1\]](#) [\[tag|AAD29758.1\]](#) [\[tag|AA51881.1\]](#) [\[tag|AA52023.1\]](#) [\[tag|AA52456.1\]](#) [\[tag|AA52783.1\]](#) [\[tag|AA56822.1\]](#) [\[tag|AA60726.1\]](#) [\[tag|AAD20588.1\]](#) [\[tag|AAD29758.1\]](#)

ENSG00000101981

- Gene information
- Gene splice site image
- Gene variation info.
- Genomic sequence
- Export data
- Transcript information
- Exon information
- Peptide information

Chromosome X 138,338,415 - 138,371,137

- View of Chromosome X
- Graphical view
- Graphical overview
- Export information about region
- Export sequence as FASTA
- Export EMBL file
- Export Gene info in region
- Export SNP info in region
- Export Vega info in region

Use Ensembl to...

- Run a BLAST search
- Search Ensembl
- Data mining [BioMart]
- Upload your own data
- Download data
- Export data

Docs and downloads

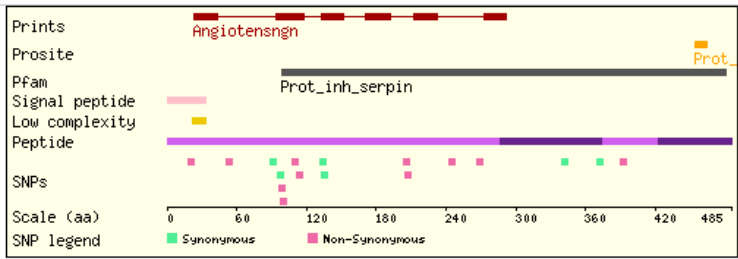
- Information
- What's New
- About Ensembl
- Using Ensembl data
- Software

Other links

- Home
- Sitemap
- Archive! sites
- Vega
- Trace server
- Stable (archive) link for this page

Protein View

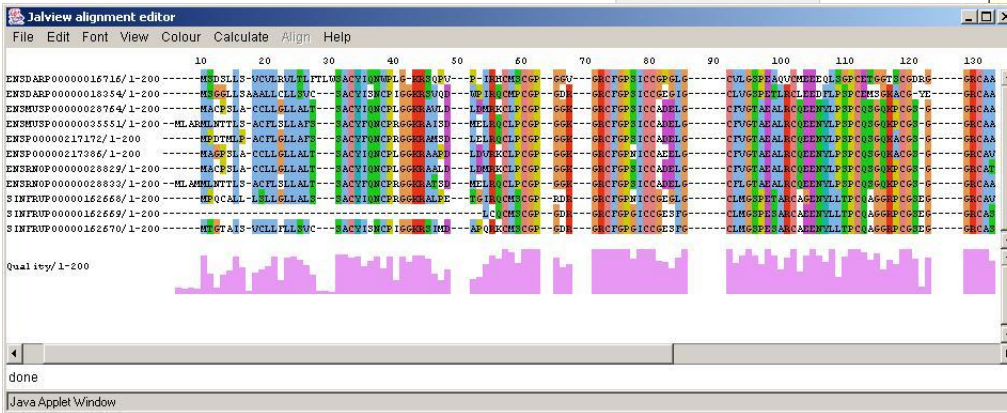
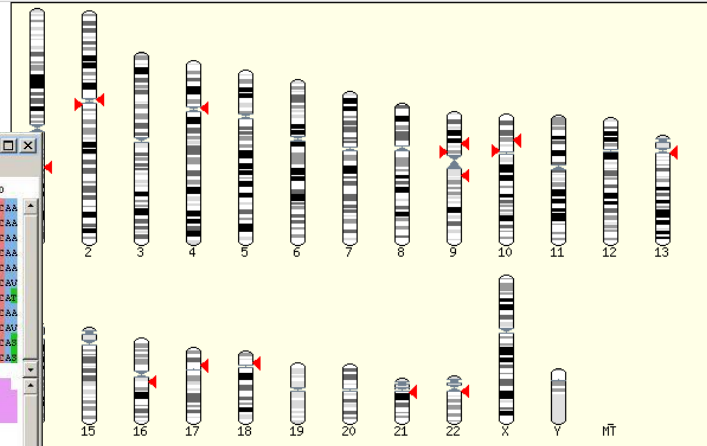
Ensembl Protein Report

Peptide	<p>AGT (HUGO ID) (to view all Ensembl genes linked to the name click here)</p> <p>This peptide is a member of the human CCDS set: CCDS1585</p>
Ensembl Peptide ID	ENSP00000258224
Translation information	This peptide is a translation of transcript: ENST00000258224 , which is a product of gene: ENSG00000135744 .
Genomic Location	<p>This peptide can be found on Chromosome 1 at location: 227,145,622-227,153,331</p> <p>This start of this peptide is located in Contig AL158214.33.1.181175.</p>
Description	Angiotensinogen precursor [Contains: Angiotensin I (Ang I); Angiotensin II (Ang II); Angiotensin III (Ang III) (Des-Asp[1]- angiotensin II)]. Source: Uniprot/SWISSPROT P01019
Prediction Method	Genes were annotated by the Ensembl automatic analysis pipeline using either a GeneWise model from a human/vertebrate protein, a set of aligned human cDNAs followed by GenomeWise for ORF prediction or from Genscan exons supported by protein, cDNA and EST evidence. GeneWise models are further combined with available aligned cDNAs to annotate UTRs.
InterPro	<p>IPR000227 Angiotensinogen - [View other genes with this domain]</p> <p>IPR000215 Proteinase inhibitor I4, serpin - [View other genes with this domain]</p>
Protein Family	<p>ENSF00000005553 : ANGIOTENSINOGEN PRECURSOR [CONTAINS: ANGIOTENSIN I ANG I ; ANGIOTENSIN II ANG II ; ANGIOTENSIN III ANG III DES ASP[1] ANGIOTENSIN II]</p> <p>This cluster contains 1 Ensembl gene member(s)</p>
Protein Features	
Peptide Sequence	<p>MRKRAPQSEMAPAGVSLRATILCLLANAGLAAGDRVYIHPFHLVIHNESTCEQLAKANAG KPKDPTFIPAPIQARTSPVDEKALQDQLVLVAAKLDTEDKLRANVGHLANFLGFRFYCH HSELWCVVHGATVLSPTAVFCTLASLYLGCALDHTADRLQAILCVPWKDKNCTSRLDAAHKV LSALQAVQGLLVAQGRADSQAQLLLSTVVGCVFTAPGLHLKQPFVQGLALYTPVWLPRLSD FTELDVAAEKIDRFHQAVTCWKTGCSLMCASVDSTLAFNTYVHFQCKHKQFSLLAEPQEF WVDNSTSVSVPMLSGCHTFQHWSDIQDNFSVTQVPFTESACLLIQPHYASDLDRVEGLT FQQNSLNVWKKLSPRITIIHLTMPQLVLQCSYDLQDLLAQAEIPALLHTELNMLQKLSNDRIR VCEVLNSIFFLEADEREPTTESTQQLNKPVLVLTNRPFLFAVYDQSATALHFLGRVAN PLSTA</p>

Family View

Ensembl Family ENSF0000000497

Family ID	ENSF0000000497
Consensus annotation	AMBIGUOUS
Prediction method	Protein families were generated using the MCL (Markov CLustering) package available at http://micans.org/mcl/ . The application of MCL to biological graphs was initially proposed by Enright A.J., Van Dongen S. and Ouzounis C.A. (2002) "An efficient algorithm for large-scale detection of protein families." Nucl. Acids. Res. 30, 1575-1584.
Multiple alignments	<p>Click to view multiple alignments of the 203 Ensembl members of this family. JalView</p> <p>Click to view multiple alignments of the 249 members of this family. JalView</p>
Ensembl genes containing peptides in family ENSF0000000497	



g family ENSF0000000497

Gene ID	Gene Name	Chromosome	Description (if known)
ENSG00000177195	-novel-	Chromosome 2: 88.94m	MGC12538 protein (Novel protein). [Source:Uniprot/SPTREMBL;Acc:Q96IX9]
ENSG00000172014	-novel-	Chromosome 2: 94.85m	No description
ENSG00000135976	-novel-	Chromosome 2: 95.94m	No description
ENSG00000174501	KIAA1641	Chromosome 2: 97.30m	No description
ENSG00000187998	NP_066021.1	Chromosome 2: 97.58m	No description
ENSG00000178305	Q8NF67_HUMAN	Chromosome 4: 49.33m	No description
ENSG00000159712	-novel-	Chromosome 9: 33.51m	No description
ENSG00000180071	ANKRD18A	Chromosome 9: 30.53m	Ankyrin repeat domain protein 18A. [Source:Uniprot/SWISSPROT;Acc:Q8IVF6]

GOView

Ensembl GO Search

GO Accession	GO:0004867 [serine-type endopeptidase inhibitor activity]																						
GO Database	GO data is provided by the Gene Ontology Consortium																						
Search GO	<p>Search GO database for: <input type="text" value="GO:0004867"/> * [e.g. GO:0004867, *vesicle, *calcium binding*]</p> <p>Search</p> <p style="text-align: right;">Fields marked with * are required</p>																						
Go Graph	<table border="1"> <thead> <tr> <th>tree</th> <th>Ensembl Gene Matches</th> </tr> </thead> <tbody> <tr> <td>all [all]</td> <td></td> </tr> <tr> <td>molecular_function [GO:0003674]</td> <td></td> </tr> <tr> <td>enzyme regulator activity [GO:0030234]</td> <td>5 gene(s)</td> </tr> <tr> <td>enzyme inhibitor activity [GO:0004857]</td> <td>22 gene(s)</td> </tr> <tr> <td>protease inhibitor activity [GO:0030414]</td> <td>1 gene(s)</td> </tr> <tr> <td>endopeptidase inhibitor activity [GO:0004866]</td> <td>17 gene(s)</td> </tr> <tr> <td>serine-type endopeptidase inhibitor activity [GO:0004867]</td> <td>80 gene(s)</td> </tr> <tr> <td>chymotrypsin inhibitor activity [GO:0030569]</td> <td>1 gene(s)</td> </tr> <tr> <td>plasmin inhibitor activity [GO:0030568]</td> <td>1 gene(s)</td> </tr> <tr> <td>trypsin inhibitor activity [GO:0030304]</td> <td>3 gene(s)</td> </tr> </tbody> </table>	tree	Ensembl Gene Matches	all [all]		molecular_function [GO:0003674]		enzyme regulator activity [GO:0030234]	5 gene(s)	enzyme inhibitor activity [GO:0004857]	22 gene(s)	protease inhibitor activity [GO:0030414]	1 gene(s)	endopeptidase inhibitor activity [GO:0004866]	17 gene(s)	serine-type endopeptidase inhibitor activity [GO:0004867]	80 gene(s)	chymotrypsin inhibitor activity [GO:0030569]	1 gene(s)	plasmin inhibitor activity [GO:0030568]	1 gene(s)	trypsin inhibitor activity [GO:0030304]	3 gene(s)
tree	Ensembl Gene Matches																						
all [all]																							
molecular_function [GO:0003674]																							
enzyme regulator activity [GO:0030234]	5 gene(s)																						
enzyme inhibitor activity [GO:0004857]	22 gene(s)																						
protease inhibitor activity [GO:0030414]	1 gene(s)																						
endopeptidase inhibitor activity [GO:0004866]	17 gene(s)																						
serine-type endopeptidase inhibitor activity [GO:0004867]	80 gene(s)																						
chymotrypsin inhibitor activity [GO:0030569]	1 gene(s)																						
plasmin inhibitor activity [GO:0030568]	1 gene(s)																						
trypsin inhibitor activity [GO:0030304]	3 gene(s)																						

Data retrieval

BioMart

Export View

Data sets on ftp site

MySQL queries of databases

Perl API access to databases

ExportView

Chromosome X
138,333,415 - 138,376,137

- View of Chromosome X
- Graphical view
- Graphical overview
- Export information about region
- Export sequence as FASTA
- Export EMBL file
- Export Gene info in region
- Export SNP info in region
- Export Vega info in region
- View alongside ...
- View Syntenic regions ...
- View region in NCBI browser
- View region in UCSC browser

Use Ensembl to...

- Run a BLAST search
- Search Ensembl
- Data mining [BioMart]
- Upload your own data
- Download data
- Export data

Docs and downloads

- Information
- About Ensembl
- Using Ensembl data
- Software

Other links

- Home
- Sitemap
- What's New
- Stable (archive) link for this page



Results

```

X chr: chromosome: chromosome: NCB135.X: 138333415: 138376137: 1
ACTGCCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT CAGGAAACTCATCT
CACATGCAGAGACACACATAGGCTCABAATAAAGGGATGGAGGAAAGATCTACCAAGCAAA
TGGAAAACAAAAGGGCAGGGTTTGCATCTCTACTCTGTATAAAACAGACTTTAAAACAA
CAAAGTCAAAGAGACAAAGAGGGCATTACATTAATGGTAAAGGGATCAATTCAACAAAG
AGGAGCTAATCTCTTAATATATATGCACCCAAACAGGAGCACCAGATTCAATAAAGC
AATCTTATAGAGACTACAAGAGACTTATAGACTCCCACATTAATTAATGGGAGACTTTA
ACACCATCTGTCAACATTAAGACAGATCAACGAGACAGAAAGTTAACAAGGATATCCAGG
AATTGAATCACTTCTGCACAGCAGGACTAATAGACATCTACAGAACTCTCCACCCCA
AATCAACAGAAATATACATTCTTCTCAGCACTACACCACTACTTCCAAAATTGATCAGC
TAGTTGAAAGTAAAGCCTCTCTCAGCAATGTAAAAGAACAGAAATTAATACAAACTATC
TCTCAGACCACAGTGCATCAAACTAGAATCAGGATTAAGAAACTCACTCAAACCACT
CGACTACATGAAAACGAAACAACTGCTCCTGAATGACTACTGGGTACATACGAAATGA
AGGCAGAAATAAAGTGTCTTTGAAACCACAAAGAACAAAGAGCAGCACTACCGAAATC
TCTGGGACACATTAAGGACGTGTAGAGGGAAATTTATAGCAATAAATGCCCAAGA
GAAAGCAGGACGGATCTAAAATGGATACCTAACATCACAATAAAAGACTAGAGAAGC
AGAGCAACACATTAAGAGCTAGCAGAGGCAAGAAATTAATAGATCAGAGCAGAGA
TGAAGGAAATAGAGACACAAAACCTTTCAAAGAAATCAATGAACCCAGGAGCTGGTTTT
TTGAAAAGATCAACAAAACCGATAGACCTAGCAAGACTAATAAGAAAGAAAAGAGAA
AGAAATCAATAGATGCATAAAATGATAAAGGGATATCACCCAGCTCCCGCAGAAAT
TACAAACTACCATCAGAGAACTATAAACACTCTATACAAATAACTAGAAAATCTAG
AAGAAATGGATAAATCTCTGACACATACACCTCCCAAGACTAAACCAGAAAGAGTTG
AATCTCTGAATAGACCAATACAGGAGCTGAAATGAGGCATTAATTAATAGCTTACTGA
CCAAAAGAGTCCAGGACAGATGGATTCACAGCCAAATCTACCAGAGTACAAGGAGG
AGCTGGTACCATTCTTCTGAAACTATCCAAATCAATAGAAAAGAGGGAAATCTCCCTA
ACTCATTTTATGAGGCGAGCATCTCTGTATACCAAGCCTGGCAGAGACACAAACAAA
AAGAGAAATTTAGACCAATATCCCTGATGAACATCAATGCAAAAATCTCAATAAATAC
TGGCAACCAAAATCCAGCAGCACATAAAGAGCTTGTCCACCATTAATCAAGTGGGCTC
TCCCTGGGATGCAAGGCTGGTTCAACATATGCAAAATCAATTAATGATTAATCCAGCAT
ATACAGAACCAAGACAAAACCAATGATTAATCTCAATAGATGCGAAGAAAAGGCTTTGAC
AAATTCACAACTTCATGCTAAAATCTCTCAATAAATAGCTTTGATGGGATCTATCT
CAAAATCATAGAGCTATCTATGACAAACCCACAGCCATATCATACTGAATGGGCAAG
ACTGGAAGCTTCCCTTTGAAACTGGCACAAGCAGGGATGCCCTCTCTCAGCACTCCT
ATTCAACATAGTGTGGAAGCTTCTGTCAGGGCAATCGGCAAGGAAAAGGAAATAAAGGG
TATTCAATAGGAAAAGAGGAAATCAATTTGTCCTGTTGCAAGATGACATGATTTGATAT
TCTAGAAAACCCCTATTGCTCATCCAAATCTCCTTAAGCTATAGGCACTTCAGCA
AGTCTCAGGATACAAATCAATGTGCAAAAATCACAGGCATTTCTATACACCAATAGCAG
ACAAACAGCAGCCAAATCAATGAGTGAATCCCTTCAAAATGCTTCAAAGAGAAATAAA
ATCTAGGAAATCTACTTACAAGGGATGTGAAGGACCTTTCAGGAGAACTACAAAAC
ACTGCTCAATGAAATAAAAGAGGATACAAACAAATGGAAGCACTTACTGCTCATGGGT
AGGAAAGAAATCAATATCATGAAAATGGCCATACTGCCCAAGGTAAATTAATAGATTCA
MTCCCTCATAGCTACCATGACTTTCTTCAAGAAATGGAAAAATCTACTTTAAATG
CATMTGGAACTAAAAGAGGCGGATGGCCAACTCAATCTAGCCAAAAGAACAAAGC
TGGAGGCATCATGCTACCTGACTTCAAATCTACTCAAGGCTACAGTAACCAAAACAGC
ATGGTACTGGTACCAAAACAGAGATACAGACCAATGGAACAGAACAGAGCCCTCAGA
AATGGCCACATATCTACAATCTCTGATCTTTGAAAACCTGACAAAACAAAGAAATGGG
AAAGGATCCCTATTAATAAATGGTGTGGGAAAACCTGGCTAGCCATATGTAGAAAGCT
GAACTGGATCCCTTCTTATACCTTATACAAAATTAATTAAGATGGATTAAGACTT
CAATGTTAGACTAAAACCTAAAACCTTAAAGAAAACCTAGGCAATACCTTTCAGGA
CATAGGCTGGGCAAGGACTTCAATGTCTAAAACACCAAAAGCAATGGCAACAAAAGC
AATGACAAATGGGATCAATGAAACTAAAGAGCTTCTGCACAGCAAGAAAACCTACCT
CAGACTGAACAGGCACCTAAAGAAATGGGAGAAAATTTGCAATCTACTCATCTGCTCC
AGGGCTAATACCTAGAACTAAAATGAACTCAAAACAAATTTACAGAAAACCAAAACAC
CCATCAACAACTGGGTGAAAGATATGAAAGACACTTCTCAAAAGAAACATTTATGCA
GCCAACAGACACATGAAAATGCTCAGCATCAGCCGCTCAGAGAAATGCAAAATCAA
ACCACATGAGATACCATCTCACAAAGTTAAGATGGCGATCATCAAAAACCTCAGGAAGC
AAGAGGTGCTGGAGAGGATGTGGGAAATAGGAACACTTTGACTACTGTTGGGACTGT
AAACTAGTCAACATTTGGGAGTCAAGTGTGGGACTTCTCAGGGATCTAGACTAGAA
ATACCATCTGACCCAGCCATCCCTATTGGGTATATACCCAAAGTATTAAATCATGC
TGCTATAAGACACATGCACAGCTATGTTTATTGGGCACTTTCACAATAGCAATGACT
TGGAAACACCCAAATGTCCAAATGATAGACTGGATTAAGAAAATGTGGCACATATAC
ACCATGGAATACTAGGACCCATAAAAAGAAATGAGTTCATGTCCTTTGTAGGGACATG
GATGAGCTAGAAACCTCTTCTCAGCAAACTATCGCAAGGACAAAACCAACACCG
CTMTCTCACTCATAGCTGGAAATGAACAAATGAGCACTTTGGACATAGGAAGGGGAA
CATCACACACCGGGCTGTTGTGGGTGGGGGCTCAGGGGAGGGATAGCAATAGGGAT
ATACCTAATGCTAATGACGAGTTAATGGGTACAGCACCAACATGGCACATGTATACA
TATGTACAAACCTGCTGTTGTGCACATGTACCCTAAAATTAAGATTAATTAATAAA
AAAAGATCATTTCAAATTTATACAGCCCTTGAAGACTTAAAATATCTTACAAAAG
AAGAAATAAAGTTGGGAAATCACTCTACCTAATATAAGTCTTACTACATAGCTACAGT
ATTATGACAGTGTATATTGGCAGAGGGATAAATCAATCAATGGGCAAGAAATAGATAG
AGAACTGGAAATAGACCCAAAACAAATGTTTACTGACTACGAAAATTTCAAGAA

```

Select region/feature to Export

Choose one or two features from the same chromosome as anchor points and display select "None" for the second feature, the display will be based around the first feature.

Please note that there is an upper limit of 5Mb that we will export.

Region

Chromosome name/fragment

From (type):

To (type):

Context

Bp downstream

Bp upstream

Output format

Output Format

FASTA File

- FASTA format text file

Feature List

- CSV (Comma Separated)
- GFF format
- Tab separated values

Flat File

- EMBL
- GenBank

PIP (%age identity)

- Pipmaker / zPicture
- Vista Format

© 2005 WTSI / EMBL

Help!

- context sensitive help pages - click

[Help](#)

- access other documentation via generic home page

- email the helpdesk

[Contact helpdesk](#)

[Go](#)

[Help](#)

Ensembl Team

Leaders	<u>Ewan Birney</u> (EBI), <u>Tim Hubbard</u> (Sanger Institute)
Database Schema and Core API	<u>Glenn Proctor</u> , <u>Andreas Kähäri</u> , <u>Ian Longden</u> , <u>Patrick Meidl</u>
BioMart	<u>Arek Kasprzyk</u> , <u>Syed Haider</u> , <u>Damian Smedley</u> , <u>Richard Hollandr</u>
Distributed Annotation System (DAS)	<u>Eugene Kulesha</u>
Outreach	<u>Xosé M Fernández</u> , <u>Bert Overduin</u> , <u>Michael Schuster</u> , <u>Giulietta Spudich</u>
Web Team	<u>James Smith</u> , <u>Fiona Cunningham</u> , <u>Anne Parker</u> , <u>Steve Trevanion</u> (VEGA), <u>Matt Wood</u>
Comparative Genomics	<u>Abel Ureta-Vidal</u> , <u>Kathryn Beal</u> , <u>Benoît Ballester</u> , <u>Stephen Fitzgerald</u> , <u>Javier Herrero Sánchez</u> , <u>Albert Vilella</u>
Analysis and Annotation Pipeline	<u>Val Curwen</u> , <u>Steve Searle</u> , <u>Browen Aken</u> , <u>Juilo Banet</u> , <u>Laura Clarke</u> , <u>Sarah Dyer</u> , <u>Jan-Hinnerck Vogel</u> , <u>Kevin Howe</u> , <u>Felix Kokocinski</u> , <u>Stephen Rice</u> , <u>Simon White</u>
Functional Genomics	<u>Paul Flicek</u> , <u>Yuan Chen</u> , <u>Stefan Gräf</u> , <u>Nathan Johnson</u> , <u>Daniel Rios</u>
Zebrafish Annotation	<u>Kerstin Howe</u> , <u>Mario Caccamo</u> , <u>Ian Sealy</u>
VectorBase Annotation	<u>Martin Hammond</u> , <u>Dan Lawson</u> , <u>Karyn Megy</u>
Systems & Support	<u>Guy Coates</u> , <u>Tim Cutts</u> , <u>Shelley Goddard</u>
Research	<u>Damian Keefe</u> , <u>Guy Slater</u> , <u>Michael Hoffman</u> , <u>Alison Meynert</u> , <u>Benedict Paten</u> , <u>Daniel Zerbino</u> , <u>Dace Ruklisa</u>