

# Option 1: UCSC Genome Browser...genome.ucsc.edu

UNIVERSITY OF CALIFORNIA  
SANTA CRUZ  **Genome Browser**

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

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

**tools**

- Genome Browser**  
interactively visualize genomic data
- LiftOver**  
rapidly align sequences to the genome
- Table Browser**  
load data from the Genome Browser database
- Variant Annotation Integrator**  
functional effect predictions for variant calls
- Data Integrator**  
combine data sources from the Genome Browser database
- Gene Sorter**  
find genes that are similar by expression and other metrics

- Genome Browser in a Box (GBiB)**  
run the Genome Browser on your laptop or server
- In-Silico PCR**  
rapidly align PCR primer pairs to the genome
- LiftOver**  
convert genome coordinates between assemblies
- VisiGene**  
interactively view *in situ* images of mouse and frog

[More tools...](#)

**Our story**

On June 22, 2000, UCSC and the other members of the International Human Genome Project consortium completed the first working draft of the human genome assembly, forever [marking the beginning of the information age](https://genome.ucsc.edu/cgi-bin/hgTables) on it.

<https://genome.ucsc.edu/cgi-bin/hgTables>

**What's new**

Nov. 07, 2016 - **New CRISPR track for many assemblies**

Nov. 03, 2016 - **New chromosome aliases search support**

Oct. 17, 2016 - **UCSC Genome Browser 2017**

# Select your species

Genomes   Genome Browser   Tools   Mirrors   Downloads   My Data   Help   About Us

## Table Browser

Use this program to retrieve the data associated with a track in text format, to calculate intersections between tracks, and to retrieve DNA sequence covered by a track. For help in using this application see [Using the Table Browser](#) for a description of the controls in this form, the [User's Guide](#) for general information and sample queries, and the [OpenHelix Table Browser tutorial](#) for a narrated presentation of the software features and usage. For more complex queries, you may want to use [Galaxy](#) or our [public MySQL server](#). To examine the biological function of your set through annotation enrichments, send the data to [GREAT](#). Send data to [GenomeSpace](#) for use with diverse computational tools. Refer to the [Credits](#) page for the list of contributors and usage restrictions associated with these data. All tables can be downloaded in their entirety from the [Sequence and Annotation Downloads](#) page.

clade: Mammal   genome: Human  
group: Repeats   assembly: Dec. 2013 (GRCh38/hg38)  
table: chainSel   describe table  
region:  genome  position or identifiers (names/acceSSIONS)  
filter: [create](#)  
intersection: [create](#)  
output format: all fields from selected table  
output file: [create](#)  
file type returned:  plain text  
  
[get output](#)   [summary/statistics](#)

To reset **all** user cart settings (including filters)

**Using the Table Browser**

This section provides brief line-by-line descriptions of the controls in the Table Browser.

- **clade:** Specifies which clade to search.
- **genome:** Specifies which genome sequence to use.
- **assembly:** Specifies which assembly to use.
- **group:** Selects the type of tracks to use. This is an alphabetical list of all available groups. Select 'All Tracks' to see all tracks including those not associated with a group. Select 'All Tables' to see all tables including those not associated with a track.
- **database:** (with "All Tables") Selects which database should be used for options in table menu.
- **track:** Selects the annotation tracks. This list displays all tracks belonging to the group specified in the group list. Some tracks are not available when

Human  
Mouse  
Alpaca  
Armadillo  
Baboon  
Bonobo  
Bushbaby  
Cat  
Chimp  
Chinese hamster  
Cow  
Crab-eating macaque  
Dog  
Dolphin  
Elephant  
Ferret  
Gibbon  
Gorilla  
Green monkey  
Guinea pig  
Hedgehog  
Horse  
Kangaroo rat  
Malayan flying lemur  
Manatee  
Marmoset  
Megabat  
Microbat  
Minke whale  
Mouse lemur  
Naked mole-rat  
Opossum  
Orangutan  
Panda  
Pig  
Pika

Send output to  Galaxy  GREAT  GenomeSpace  
keep output in browser

[click here.](#)

The Table Browser controls. For more information on using this program, see the [Table Browser User's Guide](#).

# Select your assembly (which "build" of the reference genome)

Genomes   Genome Browser   Tools   Mirrors   Downloads   My Data   Help   About Us

## Table Browser

Use this program to retrieve the data associated with a track in text format, to calculate intersections between tracks, and to retrieve DNA sequence covered by a track. For help in using this application see [Using the Table Browser](#) for a description of the controls in this form, the [User's Guide](#) for general information and sample queries, and the OpenHelix Table Browser [tutorial](#) for a narrated presentation of the software features and usage. For more complex queries, you may want to use [Galaxy](#) or our [public MySQL server](#). To examine the biological function of your set through annotation enrichments, send the data to [GREAT](#). Send data to [GenomeSpace](#) for use with diverse computational tools. Refer to the [Credits](#) page for the list of contributors and usage restrictions associated with these data. All tables can be downloaded in their entirety from the [Sequence and Annotation Downloads](#) page.

clade: Mammal   genome: Human   assembly: ✓ Dec. 2013 (GRCh38/hg38)  
group: Repeats   track: Self Chain   add  
table: chainSelf   describe table schema  
region:  genome  position chr6:31995680-31996859   lookup   define  
identifiers (names/accessions): [paste list](#) [upload list](#)  
filter: [create](#)  
intersection: [create](#)  
output format: all fields from selected table    Send output to  Galaxy  GREAT  GenomeSpace  
output file:  (leave blank to keep output in browser)  
file type returned:  plain text  gzip compressed  
  
[get output](#) [summary/statistics](#)

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

## Using the Table Browser

This section provides brief line-by-line descriptions of the Table Browser controls. For more information on using this program, see the [Table Browser User's Guide](#).

- **clade:** Specifies which clade the organism is in.
- **genome:** Specifies which organism data to use.
- **assembly:** Specifies which version of the organism's genome sequence to use.
- **group:** Selects the type of tracks to be displayed in the *track* list. The options correspond to the track groupings shown in the Genome Browser. Select 'All Tracks' for an alphabetical list of all available tracks in all groups. Select 'All Tables' to see all tables including those not associated with a track.
- **database:** (with "All Tables" group option) Determines which database should be used for options in table menu.
- **track:** Selects the annotation track data to work with. This list displays all tracks belonging to the group specified in the *group* list. Some tracks are not available when

# Let's get RefSeq genes: choose "Genes and Gene Predictions"

The screenshot shows the UCSC Genome Browser Table Browser interface. The top navigation bar includes links for Genomes, Genome Browser, Tools, Mirrors, Downloads, My Data, Help, and About Us. A sidebar on the left lists various databases and tools, with 'Genes and Gene Predictions' currently selected. The main content area provides a detailed description of the Table Browser's features, including its use for calculating intersections between tracks and retrieving DNA sequence. It also describes how to use the Table Browser for general information and sample queries, and for more complex queries through annotation enrichments. Below this, there are several input fields: 'clade' (set to 'man'), 'assembly' (set to 'Feb. 2009 (GRCh37/hg19)'), 'track' (set to 'Self Chain'), and buttons for 'add custom tracks' and 'track hubs'. There are also sections for 'region' (radio buttons for 'genome', 'ENCODE Pilot regions', and 'position', with 'chr21:33031597-33041570' selected), 'identifiers (names/accessions)' (with 'paste list' and 'upload list' buttons), 'filter' (with 'create' button), 'intersection' (with 'create' button), 'output format' (dropdown set to 'all fields from selected table'), 'Send output to' (checkboxes for Galaxy, GREAT, and GenomeSpace), 'output file' (text input field with placeholder '(leave blank to keep output in browser)'), and 'file type returned' (radio buttons for 'plain text' and 'gzip compressed'). At the bottom, there are buttons for 'get output' and 'summary/statistics', and a link to reset user cart settings.

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

## Using the Table Browser

This section provides brief line-by-line descriptions of the Table Browser controls. For more information on using this program, see the [Table Browser User's Guide](#).

- **clade:** Specifies which clade the organism is in.
- **genome:** Specifies which organism data to use.
- **assembly:** Specifies which version of the organism's genome sequence to use.
- **group:** Selects the type of tracks to be displayed in the *track* list. The options correspond to the track groupings shown in the Genome Browser. Select 'All Tracks' for an alphabetical list of all available tracks in all groups. Select 'All Tables' to see all tables including those not associated with a track.
- **database:** (with "All Tables" group option) Determines which database should be used for options in table menu.
- **track:** Selects the annotation track data to work with. This list displays all tracks belonging to the group specified in the *group* list. Some tracks are not available when

# Now select RefSeq genes

Genomes    Genome Browser    Tools    Mirrors    Downloads    My Data    Help    About Us

## Table Browser

Use this program to retrieve the data associated with a track in text format, to calculate intersections between tracks, and to retrieve DNA sequence covered by a track. For help in using this application see [Using the Table Browser](#) for a description of the controls in this form, the [User's Guide](#) for general information and sample queries, and the OpenHelix Table Browser [tutorial](#) for a narrated presentation of the software features and usage. For more complex queries, you may want to use [Galaxy](#) or our [public MySQL server](#). To examine the biological function of your set through annotation enrichments, send the data to [GREAT](#). Send data to [GenomeSpace](#) for use with diverse computational tools. Refer to the [Credits](#) page for the list of contributors and usage restrictions associated with these data. All tables can be downloaded in their entirety from the [Sequence and Annotation Downloads](#) page.

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

group: Genes and Gene Predictions    track: UCSC Genes  
RefSeq Genes

table: knownGene    describes:

region:  genome  ENCODE Pilot regions

identifiers (names/accessions):

filter:

intersection:

correlation:

output format: all fields from selected table

output file:  (leave empty to return to browser)

file type returned:  plain text  gzip compressed

To reset all user cart settings (including custom tracks)

**Using the Table Browser**

This section provides brief line-by-line descriptions of the controls. For more information on using this program, see the [Table Browser User's Guide](#).

- **clade:** Species which clade the organism
- **genome:** Species which organism data is derived from
- **assembly:** Specifies which version of the genome to use
- **group:** Selects the type of tracks to be displayed. This is an alphabetical list of all available tracks
- **database:** (with "All Tables" group option)

add custom tracks    track hubs

-33041570    lookup    define regions

Galaxy     GREAT     GenomeSpace

UCSC Genes  
RefSeq Genes  
GENCODE Gene V24lift37  
GENCODE Genes V19  
GENCODE Genes V17  
GENCODE Genes V14  
GENCODE Genes V7  
TransMap UCSC  
TransMap RefGene  
TransMap mRNA  
TransMap ESTs  
AveView Genes  
Augustus  
CCDS  
CRISPR Regions  
CRISPR Targets  
Ensembl Genes  
EvoFold  
Exoniphy  
Geneid Genes  
Genscan Genes  
H-Inv 7.0  
IKMC Genes Mapped  
lncRNA RNA-Seq Reads  
lncRNA Transcripts  
LRG Transcripts  
MGC Genes  
N-SCAN  
Old UCSC Genes  
ORFeome Clones  
Other RefSeq  
Pfam in UCSC Gene  
Retroposed Genes  
SGP Genes  
SST Genes

# Let's ask for the annotations in "BED" format. More on this later.

Genomes    Genome Browser    Tools    Mirrors    Downloads    My Data    Help    About Us

## Table Browser

Use this program to retrieve the data associated with a track in text format, to calculate intersections between tracks, and to retrieve DNA sequence covered by a track. For help in using this application see [Using the Table Browser](#) for a description of the controls in this form, the [User's Guide](#) for general information and sample queries, and the OpenHelix Table Browser [tutorial](#) for a narrated presentation of the software features and usage. For more complex queries, you may want to use [Galaxy](#) or our [public MySQL server](#). To examine the biological function of your set through annotation enrichments, send the data to [GREAT](#). Send data to [GenomeSpace](#) for use with diverse computational tools. Refer to the [Credits](#) page for the list of contributors and usage restrictions associated with these data. All tables can be downloaded in their entirety from the [Sequence and Annotation Downloads](#) page.

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

group: Genes and Gene Predictions    track: RefSeq Genes    add custom tracks    track hubs

table: refGene    describe table schema

region:  genome  ENCODE Pilot regions  position chr21:33031597-33041570    lookup    define regions

identifiers (names/accessions): paste list    upload list

filter: [create](#)

intersection: [create](#)

correlation: [create](#)

output format  all fields from selected table  
 selected fields from primary and related tables  
 sequence  
 GTF - gene transfer format  
 CDS FASTA alignment from multiple alignment  
 BED - browser extensible data

output file:

file type return:  get output     summary  
 BED - browser extensible data  
 custom track  
 hyperlinks to Genome Browser

Send output to  Galaxy     GREAT     GenomeSpace  
keep output in browser)

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

# Let's save the annotations as a file called "refseq.bed"

Genomes    Genome Browser    Tools    Mirrors    Downloads    My Data    Help    About Us

## Table Browser

Use this program to retrieve the data associated with a track in text format, to calculate intersections between tracks, and to retrieve DNA sequence covered by a track. For help in using this application see [Using the Table Browser](#) for a description of the controls in this form, the [User's Guide](#) for general information and sample queries, and the OpenHelix Table Browser [tutorial](#) for a narrated presentation of the software features and usage. For more complex queries, you may want to use [Galaxy](#) or our [public MySQL server](#). To examine the biological function of your set through annotation enrichments, send the data to [GREAT](#). Send data to [GenomeSpace](#) for use with diverse computational tools. Refer to the [Credits](#) page for the list of contributors and usage restrictions associated with these data. All tables can be downloaded in their entirety from the [Sequence and Annotation Downloads](#) page.

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

group: Genes and Gene Predictions    track: RefSeq Genes    add custom tracks    track hubs

table: refGene    describe table schema

region:  genome  ENCODE Pilot regions  position chr21:33031597-33041570    lookup    define regions

identifiers (names/accessions): paste list    upload list

filter: create

intersection: create

correlation: create

output format: BED - browser extensible data    Send output to  Galaxy  GREAT  GenomeSpace

output file: **refseq.bed** (leave blank to keep output in browser)

file type returned:  plain text  gzip compressed

[get output](#)    [summary/statistics](#)

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



Then click "get output"

# We want the “Whole Gene” in this case. Other options...

Genomes    Genome Browser    Tools    Mirrors    Downloads    My Data    Help    About Us

## Output refGene as BED

**Include custom track header:**

name= tb\_refGene  
description= table browser query on refGene  
visibility= pack  
url=

**Create one BED record per:**

Whole Gene  
 Upstream by 200 bases  
 Exons plus 0 bases at each end  
 Introns plus 0 bases at each end  
 5' UTR Exons  
 Coding Exons  
 3' UTR Exons  
 Downstream by 200 bases

Note: if a feature is close to the beginning or end of a chromosome and upstream/downstream bases are added, they may be truncated in order to avoid extending past the edge of the chromosome.

# What is in the file?

1. head -30 ~/Desktop/refseq.bed   column -t   less -S (less)																			
chr1	66999251	67216822	NM_001308293	0	+	67000041	67208778	0	22	104,123,64,25,57,55,176,25,52,86,93,75,128,127,66,112,156,133,203,65,165,8067,	0,677,92278,99501,106208,109241,109975,137426,138375,139712,143435,146109,155579,156621,160879,18								
chr1	66999638	67216822	NM_032291	0	+	67000041	67208778	0	25	413,64,25,72,57,55,176,12,12,25,52,86,93,75,501,128,127,60,112,156,133,203,65,165,8067,	0,91891,99114,101988,105821,108854,109588,126557,133574,137039,137988,139325,143048,145722,147913								
chr1	16767166	16786584	NM_001145277	0	+	16767256	16785491	0	7	182,101,105,82,109,178,1248,	0,2960,7198,7388,8421,11166,18170,								
chr1	16767166	16786584	NM_001145278	0	+	16767256	16785385	0	8	104,101,105,82,109,178,76,1248,	0,2960,7198,7388,8421,11166,15146,18170,								
chr1	16767166	16786584	NM_018090	0	+	16767256	16785385	0	8	182,101,105,82,109,178,76,1248,	0,2960,7198,7388,8421,11166,15146,18170,								
chr1	33547778	33567493	NR_126031	0	+	33567493	33567493	0	8	177,174,173,172,166,163,113,60,	0,1776,9872,11067,12370,14529,15889,19655,								
chr1	48998526	50489626	NM_001323575	0	-	48999844	50489468	0	13	1439,97,163,153,112,115,90,40,217,95,125,123,192,	0,6787,54149,57978,101638,120482,130297,334336,512729,712915,1164458,1318541,1490908,								
chr1	48998526	50489626	NM_032291	0	-	48999844	50489468	0	14	1439,27,97,163,153,112,115,90,40,217,95,125,123,192,	0,2035,6787,54149,57978,101638,120482,130297,334336,512729,712915,1164422,1318541,1490908,								
chr1	48998526	50489626	NR_136623	0	-	48999844	50489626	0	13	1439,97,163,153,112,115,90,124,40,217,95,123,192,	0,6787,54149,57978,101638,120482,130297,156925,334336,512729,712915,1318541,1490908,								
chr1	25071759	25170815	NM_013943	0	+	25072044	25167428	0	6	357,110,126,107,182,3552,	0,52473,68825,81741,94591,95504,								
chr1	48998526	50489626	NM_032785	0	-	48999844	50489468	0	14	1439,27,97,163,153,112,115,90,40,217,95,125,123,192,	0,2035,6787,54149,57978,101638,120482,130297,334336,512729,712915,1164458,1318541,1490908,								
chr1	33546713	33586132	NM_001293562	0	+	33547850	33585783	0	11	182,118,177,174,173,135,166,163,113,215,488,	0,278,1065,2841,10937,12169,13435,15594,16954,36789,38931,								
chr1	48998526	50489626	NM_001323573	0	-	48999844	50489468	0	13	1439,97,163,153,112,115,90,40,217,95,161,123,192,	0,6787,54149,57978,101638,120482,130297,334336,512729,712915,1164422,1318541,1490908,								
chr1	33546713	33586132	NM_052998	0	+	33547850	33585783	0	12	182,121,212,177,174,173,135,166,163,113,215,488,	0,275,498,1065,2841,10937,12169,13435,15594,16954,36789,38931,								
chr1	8378144	8404227	NM_001080897	0	+	8378168	8404073	0	9	102,421,93,225,728,154,177,206,421,	0,6221,7213,7733,12124,17352,19731,21408,25662,								
chr1	33547778	33567493	NM_001301826	0	+	33547850	33567493	0	8	177,174,173,135,166,163,113,60,	0,1776,9872,11104,12370,14529,15889,19655,								
chr1	33547778	33586132	NM_001301825	0	+	33547850	33585783	0	9	177,174,173,135,166,163,173,215,488,	0,1776,9872,11104,12370,14529,15829,35724,37866,								
chr1	33546729	33586132	NM_001301824	0	+	33557656	33585783	0	8	380,173,135,166,163,113,215,488,	0,10921,12153,13419,15578,16938,36773,38915,								
chr1	33546729	33586132	NM_001301823	0	+	33557656	33585783	0	9	380,174,173,135,166,163,113,215,488,	0,2825,10921,12153,13419,15578,16938,36773,38915,								
chr1	92145899	92371559	NM_001195684	0	-	92149295	92370788	0	18	3515,108,42,121,300,159,141,153,335,190,148,169,184,138,185,174,61,177,	0,15329,17746,28320,31900,35893,36225,38969,39550,41612,47316,49462,54433,78270,116944,181128,219								
chr1	92145899	92351836	NM_001195683	0	-	92149295	92370788	0	17	3515,108,42,121,300,159,141,153,335,190,148,169,184,138,185,174,402,	0,15329,17746,28320,31900,35893,36225,38969,39550,41612,47316,49462,54433,78270,116944,181128,205								
chr1	100652477	100715409	NM_0019198	0	-	100661810	100715376	0	11	9501,72,192,78,167,217,122,182,76,124,84,	0,19308,19523,23772,27895,29061,31704,43811,48514,53839,62848,								
chr1	92145899	92351836	NM_003243	0	-	92149295	92327088	0	17	3515,108,42,121,300,159,141,153,338,190,148,169,184,138,185,174,402,	0,15329,17746,28320,31900,35893,36225,38969,39550,41612,47316,49462,54433,78270,116944,181128,205								
chr1	92145899	92351836	NR_036634	0	-	92351836	92351836	0	18	3515,108,42,121,300,159,141,153,338,190,148,169,184,138,185,97,174,402,	0,15329,17746,28320,31900,35893,36225,38969,39550,41612,47316,49462,54433,78270,116944,120616,181								



# Option 2: Ensembl biomart

www.ensembl.org/index.html?redirect=no

**e!Ensembl** Login/Register

Search: All species for  Go

e.g. [BRCA2](#) or [rat 5:62797383-63627669](#) or [rs699](#) or [coronary heart disease](#)

**Browse a Genome**

Ensembl is a genome browser for vertebrate genomes that supports research in comparative genomics, evolution, sequence variation and transcriptional regulation. Ensembl annotate genes, computes multiple alignments, predicts regulatory function and collects disease data. Ensembl tools include BLAST, BLAT, BioMart and the Variant Effect Predictor (VEP) for all supported species.

**Favourite genomes**

 <b>Human</b> GRCh38.p7	 <b>Human</b> GRCh37
 <b>Mouse</b> GRCm38.p4	 <b>Zebrafish</b> GRCz10

[Edit favourites](#)

**All genomes**

-- Select a species --

[View full list of all Ensembl species](#)

Other species are available in [Ensembl Pre!](#) and [EnsemblGenomes](#)

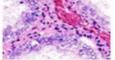
**Still using Human GRCh37?**

[Go to e!GRCh37](#)

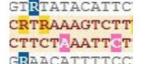
**Variant Effect Predictor**



**Gene expression in different tissues**



**Find SNPs and other variants for my gene**



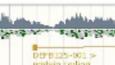
**Retrieve gene sequence**

GCCTGACTTCGGGGGG;  
GGGCTTGTGGCGGAGC;  
QCCTCTCTCTCCGCCCT;  
AGGGGACAGAATTGGTGA;  
CACCTCTGGAGCGGGTT;  
CCCAATCCAGCGGGCG

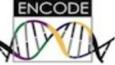
**Compare genes across species**



**Use my own data in Ensembl**



**ENCODE data in Ensembl**



**What's New in Ensembl Release 86 (October 2016)**

- Mouse Strains
- Chicken new assembly and gene set
- Macaque new assembly and genebuild
- Mouse lemur new assembly and genebuild
- Zebrafish: update to Ensembl-Havana merged gene set

[Full details](#) | [All web updates, by release](#) | [More news on our blog](#)

- 28 Oct 2016: [Ensembl genomes 33 is out!](#)
- 17 Oct 2016: [Ensembl helpdesk maintenance](#)
- 05 Oct 2016: [Ensembl 86 has been released!](#)

[Go to Ensembl blog](#)

# Select a species: [www.ensembl.org](http://www.ensembl.org)

**e!Ensembl** BLAST/BLAT | BioMart | Tools | Downloads | Help & Documentation | Blog | Mirrors

Login/Register

Search: All species for  Go

e.g. BRCA2 or rat 5:62797383-63627669 or rs699 or coronary heart disease

**Browse a Genome**

Ensembl is a genome browser for vertebrate genomes that supports research in comparative genomics, evolution, sequence variation and transcriptional regulation. Ensembl annotates genes, computes multiple alignments, predicts regulatory function and collects disease data. Ensembl tools include BLAST, BLAT, BioMart and the Variant Effect Predictor (VEP) for all supported species.

**Favourite genomes**

Human GRCh38.p7 Human GRCh37 Zebrafish GRCz10

-- Select a species --

- Favourites
  - Human GRCh38
  - Human GRCh37
  - Mouse
  - Zebrafish
- Primates
  - Bushbaby
  - Chimpanzee
  - Gibbon
  - Gorilla
- Human
- Macaque
- Marmoset
- Mouse Lemur
- Olive baboon
- Orangutan
- Tarsier

Still using Human GRCh37? Go to eGRCh37

Variant Effect Predictor Ve!P

Gene expression in different tissues

Find SNPs and other variants for my gene

GT<sup>R</sup>TATAACATTC  
CT<sup>R</sup>TAAGTCTT  
CTTC<sup>A</sup>ATT<sup>T</sup>  
GRAACATTTCC

Retrieve gene sequence

GCCTGACTTCGGGTGG;  
GGGCTTGTGGCGAAC;  
GGGGCTCTGCTGGGCT;  
AAGGGACAGATTGTGAA;  
CACCTCTGGACGGGTT;  
CCCCATCCAGCGTGGC

Compare genes across species

Use my own data in Ensembl

ENCODE data in Ensembl

Ensembl supports data from external projects through

**What's New in Ensembl Release 86 (October 2016)**

- Mouse Strains
- Chicken new assembly and gene set
- Macaque new assembly and genebuild
- Mouse lemur new assembly and genebuild
- Zebrafish: update to Ensembl-Havana merged gene set

[Full details](#) | [All web updates, by release](#) | [More news on our blog](#)

- 28 Oct 2016: [Ensembl genomes 33 is out!](#)
- 17 Oct 2016: [Ensembl helpdesk maintenance](#)
- 05 Oct 2016: [Ensembl 86 has been released!](#)

[Go to Ensembl blog](#)

# If necessary, select a genome build

**e!Ensembl** BLAST/BLAT | BioMart | Tools | Downloads | Help & Documentation | Blog | Mirrors

Human (GRCh38.p7) ▾

 **Human**  
*Homo sapiens*

Search all categories ▾ Search Human... Go

e.g. [BRCA2](#) or [17:63973115-64437414](#) or [rs1333049](#) or [osteoarthritis](#)

**Genome assembly: GRCh38.p7 (GCA\_000001405.22)**

-  More information and statistics
-  Download DNA sequence (FASTA)
-  Convert your data to GRCh38 coordinates
-  Display your data in Ensembl

Other assemblies

GRCh37 Full Feb 2014 archive with BLAST, VEP and BioMart ▾ Go

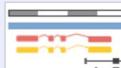
**Comparative genomics**

What can I find? Homologues, gene trees, and whole genome alignments across multiple species.

-  More about comparative analysis
-  Download alignments (EMF)

**Regulation**

What can I find? DNA methylation, transcription factor binding sites, histone modifications, and regulatory features such as enhancers and repressors, and

 View karyotype  
 Example region

**Gene annotation**

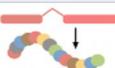
What can I find? Protein-coding and non-coding genes, splice variants, cDNA and protein sequences, non-coding RNAs.

-  More about this genebuild, including RNASeq gene expression models
-  Download genes, cDNAs, ncRNA, proteins (FASTA)
-  Update your old Ensembl IDs

 Additional manual annotation can be found in Vega

 Pax6 INS  
FOXP2  
BRCA2  
DMD ssh

Example gene

 Example transcript

**Variation**

What can I find? Short sequence variants and longer structural variants; disease and other phenotypes

-  More about variation in Ensembl
-  Download all variants (GVF)
-  Variant Effect Predictor

 ATCGAGCT  
ATCCAGCT  
ATCGAGAT

Example variant

 Example phenotype





# Now choose "BioMart"

The screenshot shows the GRCh37 Human genome browser interface. At the top, there's a navigation bar with links for BLAST/BLAT, BioMart, Tools, Downloads, Help & Documentation, and Blog. A green arrow points from the left towards the 'BioMart' link. To the right of the navigation bar is a search bar labeled 'Search Human...' with a magnifying glass icon. In the top right corner, there are 'Login/Register' links.

**Human (GRCh37.p13) ▾**

**Human**  
*Homo sapiens*

Search all categories ▾ Search Human... Go

e.g. [BRCA2](#) or [17:63973115-64437414](#) or [rs1333049](#) or [osteoarthritis](#)

**Genome assembly: GRCh37.p13 (GCA\_000001405.14)**

- [More information and statistics](#)
- [Download DNA sequence \(FASTA\)](#)
- [Convert your data to GRCh37.p13 coordinates](#)
- [Display your data in Ensembl](#)

**Other assemblies**

GRCh38 (Ensembl release 86) Go

**Comparative genomics**

What can I find? Homologues, gene trees, and whole genome alignments across multiple species.

- [More about comparative analysis](#)
- [Download alignments \(EMF\)](#)

**Regulation**

What can I find? DNA methylation, transcription factor binding sites, histone modifications, and regulatory features such as enhancers and repressors, and grch37.ensembl.org/biomart/martview

**Gene annotation**

What can I find? Protein-coding and non-coding genes, splice variants, cDNA and protein sequences, non-coding RNAs.

- [More about this genebuild, including RNASeq gene expression models](#)
- [Download genes, cDNAs, ncRNA, proteins \(FASTA\)](#)
- [Update your old Ensembl IDs](#)

Vega\* Additional manual annotation can be found in Vega

**What's New in Human GRCh37**

- This GRCh37 archive will be updated every release with new web features, and periodically with new human data.

**Variation**

What can I find? Short sequence variants and longer structural variants; disease and other phenotypes

- [More about variation in Ensembl](#)
- [Download all variants \(GVF\)](#)
- [Variant Effect Predictor](#)

**Pax6 INS FOXP2 BRCA2 DMD ssh**

Example gene

**Example transcript**

**ATCGAGCT ATCCAGCT ATCGAGAT**

Example variant

Example phenotype

# Let's choose Ensembl gene

The screenshot shows the e!GRCh37 search interface. At the top, there is a navigation bar with links for BLAST/BLAT, BioMart, Tools, Downloads, Help & Documentation, and Blog. On the right side of the navigation bar is a "Login/Register" link and a search bar labeled "Search all species..." with a magnifying glass icon. Below the navigation bar is a toolbar with buttons for New, Count, Results, URL, XML, Perl, and Help. On the left, there is a sidebar titled "Dataset" with the sub-label "[None selected]". A dropdown menu titled "✓ - CHOOSE DATABASE -" is open, listing "Ensembl Gene" (which is highlighted in blue), "Ensembl Variation", "Ensembl Regulation", and "Vega". The main content area below the sidebar is currently empty.

Datasets -> Filters (filtering and inputs) -> Attributes (desired output) -> Results



# Pick the appropriate assembly

**e!GRCh37** BLAST/BLAT | BioMart | Tools | Downloads | Help & Documentation | Blog

Login/Register

Search all species... 

New Count Results

**Dataset**  
Homo sapiens genes (GRCh37.p13)  
**Filters**  
[None selected]  
**Attributes**  
Ensembl Gene ID  
Ensembl Transcript ID

- CHOOSE DATASET -  
Danio rerio genes (Zv9)  
Gallus gallus genes (Galgal4)  
✓ Homo sapiens genes (GRCh37.p13)  
Mus musculus genes (GRCm38.p2)  
Rattus norvegicus genes (Rnor\_5.0)

Ailuropoda melanoleuca genes (ailMe1)  
Anas platyrhynchos genes (BGI\_duck\_1.0)  
Anolis carolinensis genes (AnoCar2.0)  
Astyanax mexicanus genes (AstMex102)  
Bos taurus genes (UMD3.1)  
Caenorhabditis elegans genes (WBcel235)  
Callithrix jacchus genes (C\_jacchus3.2.1)  
Canis familiaris genes (CanFam3.1)  
Cavia porcellus genes (cavPor3)  
Chloepus hoffmanni genes (choHof1)  
Ciona intestinalis genes (KH)  
Cliona savignyi genes (CSAV2.0)  
Dasypus novemcinctus genes (Dasnov3.0)  
Dipodomys ordii genes (dipOrd1)  
Drosophila melanogaster genes (BDGP5)  
Echinops telfairi genes (TENREC)  
Equus caballus genes (EquCab2)  
Erinaceus europaeus genes (eriEur1)  
Felis catus genes (Felis\_catus\_6.2)  
Ficedula albicollis genes (FicAlb\_1.4)  
Gadus morhua genes (gadMor1)  
Gasterosteus aculeatus genes (BROADS1)  
Gorilla gorilla genes (gorGor3.1)  
Ictidomys tridecemlineatus genes (spetri2)  
Latimeria chalumnae genes (LatCha1)  
Lepisosteus oculatus genes (LepOcu1)

URL XML Perl Help

Outputs) -> Attributes (desired output) -> Results

# Select the attributes of interest

The screenshot shows the GRCh37 BioMart interface. On the left, there's a sidebar with sections for 'Dataset' (Homo sapiens genes (GRCh37.p13)), 'Filters' ([None selected]), and 'Attributes' (Ensembl Gene ID, Ensembl Transcript ID). The main area has a 'Dataset' section with dropdown menus set to 'Ensembl Gene' and 'Homo sapiens genes (GRCh37.p13)'. At the top, there are navigation links for New, Count, Results, URL, XML, Perl, Help, and a search bar for species.

Datasets -> Filters (filtering and inputs) -> Attributes (desired output) -> Results



# Select "structures"

The screenshot shows the GRCh37 BioMart interface. On the left, there's a sidebar with sections for Dataset (Homo sapiens genes (GRCh37.p13)), Filters ([None selected]), and Attributes (Ensembl Gene ID, Ensembl Transcript ID). The main area has tabs for New, Count, and Results, with Results selected. At the top right are links for Login/Register, a search bar, and download options (URL, XML, Perl). A message in the center says "Please select columns to be included in the output and hit 'Results' when ready". Below it, another message says "Missing non coding genes in your mart query output, please check the following [FAQ](#)". There are two groups of radio buttons: one for Features/Sequences and Variant types, and another for Structures/Homologues. Under "Attributes", there are checkboxes for GENE and EXON.

Datasets -> Filters (filtering and inputs) -> Attributes (desired output) -> Results



# Let's grab starts and ends for transcripts, and chrom

The screenshot shows the e!GRCh37 BioMart interface. On the left, there are three tabs: 'New' (selected), 'Count', and 'Results'. At the top right are links for 'Login/Register' and a search bar with a magnifying glass icon. Below the search bar is a row of icons: URL (star icon), XML, Perl, and Help.

**Dataset:** Homo sapiens genes (GRCh37.p13)

**Filters:** [None selected]

**Attributes:**

- Ensembl Gene ID
- Ensembl Transcript ID
- Chromosome Name
- Transcript Start (bp)
- Transcript End (bp)

---

**Dataset:** [None Selected]

**Output Selection:**

Please select columns to be included in the output and hit 'Results' when ready

Missing non coding genes in your mart query output, please check the following [FAQ](#)

Features     Sequences  
 Structures     Variant (Germline)  
 Homologues     Variant (Somatic)

GENE:  
**Ensembl**  
 Ensembl Gene ID  
 Ensembl Transcript ID  
 Ensembl Protein ID  
 Chromosome Name  
 Gene Start (bp)  
 Gene End (bp)  
 Transcript Start (bp)  
 Transcript End (bp)  
 Transcription Start Site (TSS)  
 Transcript length (including UTRs and CDS)  
 Strand

Associated Gene Name  
 Associated Gene DB  
 5' UTR Start  
 5' UTR End  
 3' UTR Start  
 3' UTR End  
 CDS Length  
 Transcript count  
 Description  
 Gene type

EXON:

Datasets -> Filters (filtering and inputs) -> Attributes (desired output) -> Results

# ...and starts and ends for exons

The screenshot shows the e!GRCh37 web application interface. At the top, there is a navigation bar with links for BLAST/BLAT, BioMart, Tools, Downloads, Help & Documentation, and Blog. On the right side of the top bar are Login/Register and a search bar labeled "Search all species..." with a magnifying glass icon.

The main content area has tabs for New, Count, and Results at the top left, and URL, XML, Perl, and Help buttons at the top right.

**Dataset:** Homo sapiens genes (GRCh37.p13)

**Filters:** [None selected]

**Attributes:**

- Ensembl:**
  - Ensembl Gene ID
  - Ensembl Transcript ID
  - Ensembl Protein ID
  - Chromosome Name
  - Gene Start (bp)
  - Gene End (bp)
  - Transcript Start (bp)
  - Transcript End (bp)
  - Transcription Start Site (TSS)
  - Transcript length (including UTRs and CDS)
  - Strand
- Associated Gene Name
- Associated Gene DB
- 5' UTR Start
- 5' UTR End
- 3' UTR Start
- 3' UTR End
- CDS Length
- Transcript count
- Description
- Gene type

- Exon:**
- Exon Information:**
  - Exon Chr Start (bp)
  - Exon Chr End (bp)
  - Constitutive Exon
  - Exon Rank in Transcript
  - start phase
  - cDNA coding start
- cDNA coding end
- Genomic coding start
- Genomic coding end
- Ensembl Exon ID
- CDS Start
- CDS End

Datasets -> Filters (filtering and inputs) -> Attributes (desired output) -> Results

# Let's display the results



The screenshot shows the e!GRCh37 BioMart interface. At the top, there is a navigation bar with links for BLAST/BLAT, BioMart, Tools, Downloads, Help & Documentation, and Blog. On the right of the navigation bar is a search bar labeled "Search all species..." with a magnifying glass icon. Below the navigation bar is a toolbar with buttons for New, Count, Result (which is highlighted with a red box and has a green arrow pointing to it), URL, XML, Perl, and Help.

**Dataset**  
Homo sapiens genes (GRCh37.p13)

**Filters**  
[None selected]

**Attributes**

- Ensembl Gene ID
- Ensembl Transcript ID
- Chromosome Name
- Transcript Start (bp)
- Transcript End (bp)
- Exon Chr Start (bp)
- Exon Chr End (bp)

---

**Dataset**  
[None Selected]

**Result**

**GENE:**

**Ensembl**

- Ensembl Gene ID
- Ensembl Transcript ID
- Ensembl Protein ID
- Chromosome Name
- Gene Start (bp)
- Gene End (bp)
- Transcript Start (bp)
- Transcript End (bp)
- Transcription Start Site (TSS)
- Transcript length (including UTRs and CDS)
- Strand

- Associated Gene Name
- Associated Gene DB
- 5' UTR Start
- 5' UTR End
- 3' UTR Start
- 3' UTR End
- CDS Length
- Transcript count
- Description
- Gene type

**EXON:**

**Exon Information**

- Exon Chr Start (bp)
- Exon Chr End (bp)
- Constitutive Exon
- Exon Rank in Transcript
- start phase
- cDNA coding start

- cDNA coding end
- Genomic coding start
- Genomic coding end
- Ensembl Exon ID
- CDS Start
- CDS End

Datasets -> Filters (filtering and inputs) -> Attributes (desired output) -> Results

# Download the first 200 results

The screenshot shows the GRCh37 bioinformatics tool interface. The top navigation bar includes links for BLAST/BLAT, BioMart, Tools, Downloads, Help & Documentation, and Blog. A search bar at the top right allows searching across all species. The main interface has tabs for New, Count, and Results. The Results tab is active, displaying a table of search results. The table has columns for Ensembl Gene ID, Ensembl Transcript ID, Chromosome Name, Start (bp), End (bp), and Transcript Start (bp). The table shows the first 200 rows of results for Homo sapiens genes (GRCh37.p13). The table includes a dropdown menu for selecting the number of rows (10, 20, 50, 100, 150, 200) and a dropdown for file format (URL, XML, Perl, TSV). A checkbox for "Unique results only" is also present.

Ensembl Gene ID	Ensembl Transcript ID	Chromosome Name	Start (bp)	End (bp)	Transcript Start (bp)
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66119285
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66298434
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66314236
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66320895
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66339743
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66341024
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66424056
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66440552
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66447170
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66448221
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66455382
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66320895
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66339743
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66341024
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66424056
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66440552
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66447170
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66448221
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66448221
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66449101

Datasets -> Filters (filtering and inputs) -> Attributes (desired output) -> Results

# Download the first 200 results

The screenshot shows the GRCh37 bioinformatics tool interface. The top navigation bar includes links for BLAST/BLAT, BioMart, Tools, Downloads, Help & Documentation, and Blog. A search bar at the top right allows searching across all species. The main interface has tabs for New, Count, and Results. The Results tab is active, displaying a table of search results. The table has columns for Ensembl Gene ID, Ensembl Transcript ID, Chromosome Name, Start (bp), End (bp), and Transcript Start (bp). The table shows the first 200 rows of results for Homo sapiens genes (GRCh37.p13). The table includes a dropdown menu for selecting the number of rows (10, 20, 50, 100, 150, 200) and a dropdown for file format (URL, XML, Perl, TSV). A checkbox for "Unique results only" is also present.

Ensembl Gene ID	Ensembl Transcript ID	Chromosome Name	Start (bp)	End (bp)	Transcript Start (bp)
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66119285
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66298434
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66314236
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66320895
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66339743
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66341024
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66424056
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66440552
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66447170
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66448221
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66455382
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66320895
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66339743
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66341024
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66424056
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66440552
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66447170
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66448221
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66448221
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66449101

Datasets -> Filters (filtering and inputs) -> Attributes (desired output) -> Results

Or choose "All" and download as a file in TSV (tab separated) format

**e!GRCh37**

BLAST/BLAT | BioMart | Tools | Downloads | Help & Documentation | Blog

Search all species...

New Count Results

Dataset

Homo sapiens genes (GRCh37.p13)

Filters

[None selected]

Attributes

Ensembl Gene ID  
Ensembl Transcript ID  
Chromosome Name  
Transcript Start (bp)  
Transcript End (bp)  
Exon Chr Start (bp)  
Exon Chr End (bp)

Dataset

[None Selected]

Export all results to

Email notification to

View 200 rows as TSV Unique results only Go

All

Ensembl Gene ID	Ensembl Transcript ID	Chromosome	Name	Transcript Start (bp)	Transcript End (bp)
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66119285
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66298434
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66314236
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66320895
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66339743
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66341024
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66424056
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66440552
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66447170
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66448221
ENSG00000261657	ENST00000566782	HG991_PATCH	66119285	66456619	66455382
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66320895
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66339743
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66341024
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66424056
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66440552
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66447170
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66448221
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66456619
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66447170
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66448221
ENSG00000261657	ENST00000562780	HG991_PATCH	66320895	66455748	66448221

Datasets -> Filters (filtering and inputs) -> Attributes (desired output) -> Results