

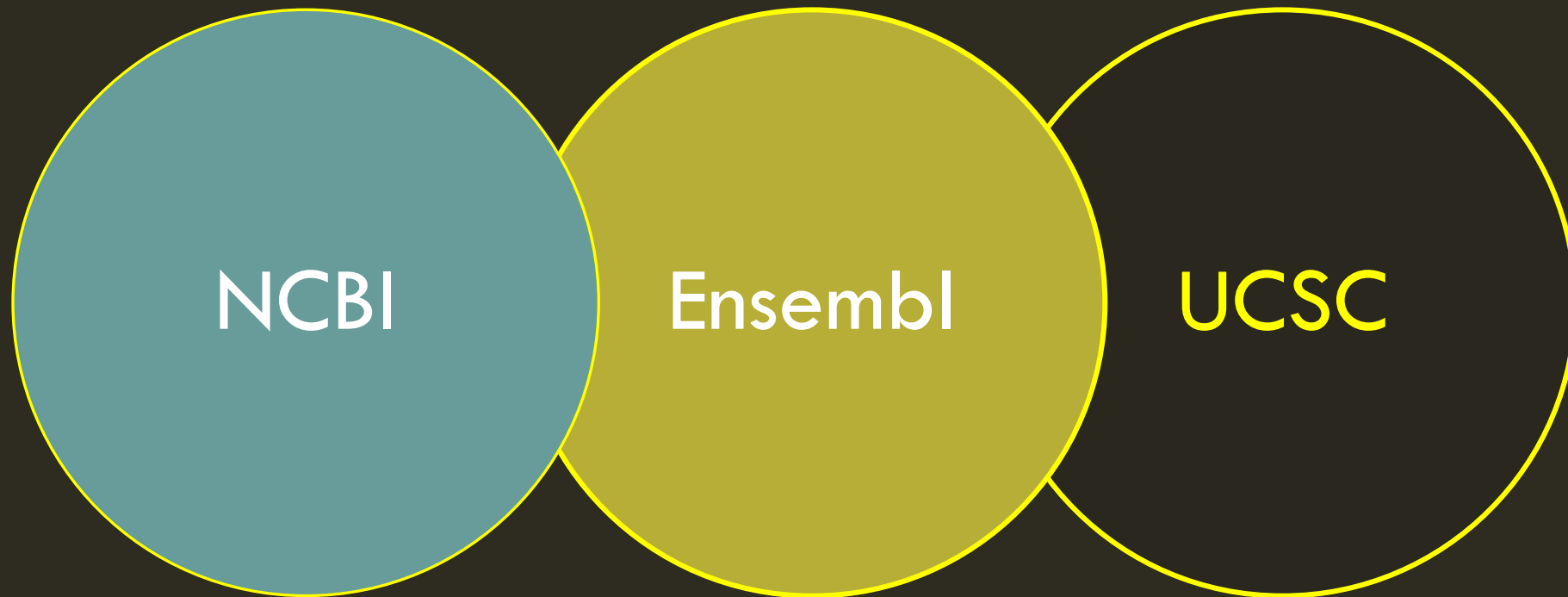
Niniejsze opracowanie zostało stworzone przez dr Magdę Mielczarek, pracownika Uniwersytetu Przyrodniczego we Wrocławiu w ramach wykonywania obowiązków związanych z kształceniem studentów i jest przeznaczone dla studentów Bioinformatyki (Wydział Biologii i Hodowli Zwierząt) na potrzeby dydaktyczne bez prawa do dalszego rozpowszechniania.



POPULARNE PRZEGLĄDARKI GENOMOWE

Pracownia Informatyczna

WYBRANE BIOLOGICZNE BAZY DANYCH POPULARNE PRZEGLĄDARKI GENOMOWE



NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION



- Utworzone w 1988
- Dziedziny nauki: biologia molekularna, biochemia, genetyka
- Zadania
 - Przechowywanie → bazy danych
 - Analiza danych → oprogramowanie
 - Udostępnianie baz danych i oprogramowania do badań naukowych
 - Koordynacja gromadzenia danych na poziomie międzynarodowym

Search NCBI databases

[Help](#)

Search

Literature

Books	books and reports
MeSH	ontology used for PubMed indexing
NLM Catalog	books, journals and more in the NLM Collections
PubMed	scientific & medical abstracts/citations
PubMed Central	full-text journal articles

Health

ClinVar	human variations of clinical significance
dbGaP	genotype/phenotype interaction studies
GTR	genetic testing registry
MedGen	medical genetics literature and links
OMIM	online mendelian inheritance in man
PubMed Health	clinical effectiveness, disease and drug reports

Genomes

Assembly	genome assembly information
BioCollections	museum, herbaria, and other biorepository collections
BioProject	biological projects providing data to NCBI
BioSample	descriptions of biological source materials

Genes

EST	expressed sequence tag sequences
Gene	collected information about gene loci
GEO DataSets	functional genomics studies
GEO Profiles	gene expression and molecular abundance profiles
HomoloGene	homologous gene sets for selected organisms
PopSet	sequence sets from phylogenetic and population studies
UniGene	clusters of expressed transcripts

Proteins

Conserved Domains	conserved protein domains
Protein	protein sequences
Protein Clusters	sequence similarity-based protein clusters
Structure	experimentally-determined biomolecular structures

Chemicals

BioSystems	molecular pathways with links to genes, proteins and chemicals
PubChem BioAssay	bioactivity screening studies
PubChem Compound	chemical information with structures, information and

NCBI ASSEMBLY & GENOME DATABASES

Assembly: A database providing information on the structure of assembled genomes, assembly names and other meta-data, statistical reports, and links to genomic sequence data.

Genome: This resource organizes information on genomes including sequences, maps, chromosomes, assemblies, and annotations.

NCBI ASSEMBLY & GENOME DATABASES

Search results

Items: 12

i Filters activated: Latest, Exclude anomalous. [Clear all](#) to show

[ARS-UCD1.2](#)

1. Organism: **Bos taurus** (cattle)
Intraspecific name: Breed: Hereford
Sex: female
Submitter: USDA ARS
Date: 2018/04/11
Assembly level: Chromosome
Genome representation: full
RefSeq category: representative genome
GenBank assembly accession: GCA_002263795.2 (latest)
RefSeq assembly accession: GCF_002263795.1 (latest)
IDs: 1677391 [UID] 6369068 [GenBank] 6386598 [RefSeq]

[Bos taurus UMD 3.1.1](#)

2. Organism: **Bos taurus** (cattle)
Intraspecific name: Breed: Hereford
Sex: pooled male and female
Submitter: Center for Bioinformatics and Computational Biology, Un
Date: 2014/11/25
Assembly level: Chromosome
Genome representation: full
GenBank assembly accession: GCA_000003055.5 (latest)
RefSeq assembly accession: GCF_000003055.6 (suppressed)
Release type: Minor
IDs: 228231 [UID] 1400768 [GenBank] 1425528 [RefSeq]

Bos taurus (cattle)

Reference genome: [Bos taurus \(assembly Bos_taurus_UMD_3.1.1\)](#)

Download sequences in FASTA format for [genome](#), [transcript](#), [protein](#)

Download genome annotation in [GFF](#), [GenBank](#) or [tabular](#) format

BLAST against Bos taurus [genome](#)

All 3 genomes for species:

Browse the [list](#)

Download sequence and annotation from [RefSeq](#) or [GenBank](#)

Display Settings: **v** Overview

Send to: **v**

Organism Overview ; [Genome Assembly and Annotation report \[3\]](#) ; [Organelle Annotation Report \[2\]](#)

ID: 82



Bos taurus (cattle)

cow

Lineage: [Eukaryota\[3281\]](#); [Metazoa\[1076\]](#); [Chordata\[526\]](#); [Craniata\[516\]](#); [Vertebrata\[516\]](#); [Euteleostomi\[508\]](#); [Mammalia\[178\]](#); [Eutheria\[172\]](#); [Laurasiatheria\[89\]](#); [Cetartiodactyla\[39\]](#); [Ruminantia\[21\]](#); [Pecora\[21\]](#); [Bovidae\[15\]](#); [Bovinae\[7\]](#); [Bos\[4\]](#); [Bos taurus\[1\]](#)

Bos taurus (cow) is an agriculturally important animal; beef and milk production are the largest manufacturing industries in the United States. The cow is an important model organism for health research in obesity, female health, and infectious diseases. Cow is also used in studies of endocrinology, physiology and reproductive techniques. The [More...](#)

Summary

Sequence data:	genome assemblies: 3; sequence reads: 32 (See Genome Assembly and Annotation report)
Statistics:	median total length (Mb): 2715.85
	median protein count: 42497
	median GC%: 41.8685

NCBI Annotation Release: 106

NCBI GENOME DATABASE

NCBI Resources How To Sign in to NCBI

Genome Genome mus musculus[orgn] Search

Create alert Limits Advanced Help

Mus musculus (house mouse)
Reference genome: **Mus musculus**
Download sequences in FASTA form
Download genome annotation in GFF3
BLAST against Mus musculus genome
All 22 genomes for species:
Browse the list
Download sequence and annotation

NCBI Resources


Representative (genome information for reference and representative genomes)

Reference genome:

- Mus musculus GRCm38.p6**
Submitter: Genome Reference Consortium

Display Settings: Overview

Organism Overview : Genome /



Mus musculus
The laboratory mouse genome has been sequenced and annotated.

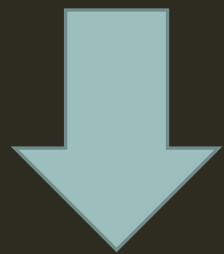
Lineage: Eukaryota, Mammalia, Euarchontoglires
The mouse is one of the major model organisms, providing a wealth of different genetic backgrounds, chemically-induced, and transgenic strains.

Loc	Type	Name	RefSeq	INSDC	Size (Mb)	GC%	Protein	rRNA	tRNA	Other RNA	Gene	Pseudogene
	Chr	1	NC_000067.6	CM000994.2	195.47	41.3	4,731	-	37	2,031	2,687	579
	Chr	2	NC_000068.7	CM000995.2	182.11	42.2	6,282	-	8	2,783	3,491	609
	Chr	3	NC_000069.6	CM000996.2	160.04	40.7	3,507	-	40	1,663	2,225	480
	Chr	4	NC_000070.6	CM000997.2	156.51	42.5	4,710	-	8	1,991	2,622	497
	Chr	5	NC_000071.6	CM000998.2	151.84	42.7	4,634	-	10	1,952	2,507	413
	Chr	6	NC_000072.6	CM000999.2	149.74	41.6	3,844	1	53	1,572	2,597	555
	Chr	7	NC_000073.6	CM001000.2	145.44	43.2	6,336	-	10	2,087	3,798	935
	Chr	8	NC_000074.6	CM001001.2	129.4	42.6	3,653	32	7	1,673	2,177	376
	Chr	9	NC_000075.6	CM001002.2	124.6	42.9	4,406	-	7	1,698	2,276	374

FASTA

Prosty i popularny → czytelny dla wielu programów do analizy bioinformatycznej

Zapis sekwencji kwasów nukleinowych oraz białek



Identyfikator
sekwencji

Sekwencja

>gi|52693750|dbj|AB175071.1

ATGACCAACTTTCGAAAAACCCATCCATTAATAAAAATTCTTAACAACCTCATTTCATCGAC
CATCAAACATTCATCATGATGAAATTCGGGTCCCTTCTAGGATTGTGCCTAGTAATC

NCBI REFERENCE SEQUENCE DATABASE

NCBI RefSeq:

- kompleksowy, zintegrowany, niepowtarzalny dobrze opisany zestaw sekwencji referencyjnych
- genomy, transkrypty, białka
- baza niepowtarzalna (non-redundant)
- informacje sprawdzone (często weryfikowane manualnie)

NCBI REFERENCE SEQUENCE DATABASE

RefSeq

RefSeq

mus musculus

Search



RefSeq: NCBI Reference Sequence Database

A comprehensive, integrated, non-redundant, well-annotated set of reference sequences including genomic, transcript, and protein.

Using RefSeq

[About RefSeq](#)

[Human Reference Genome](#)

[Prokaryotic RefSeq Genomes](#)

[FAQ](#)

[NCBI Handbook](#)

[Factsheet](#)

RefSeq Access

[Human Genome Resources and Download](#)

[RefSeq FTP](#)

[RefSeq genomes FTP](#)

[New RefSeq genomic \(last 30 days\)](#)

[New RefSeq transcripts \(last 30 days\)](#)

[New RefSeq proteins \(last 30 days\)](#)

[Searching for RefSeq records \(Queries\)](#)

Announcements

September 13, 2021

RefSeq Release 208 is available for FTP

This release includes:

Proteins: 210,703,648

Transcripts: 40,213,945

Organisms: 113,002

Available at: <ftp://ftp.ncbi.nlm.nih.gov/refseq/release/>

Documentation: [Release Notes](#)

See [previous announcements](#), follow [NCBI on Twitter](#), or subscribe to [NCBI's refseq-announce mail list](#) to receive announcements.

NCBI REFERENCE SEQUENCE DATABASE

NCBI Resources How To Sign in to NCBI

Nucleotide Search

Create alert Advanced Help

Primary 20 per page Sort by Default order Send to: Filters: [Manage Filters](#)

Items: 1 to 20 of 129096

<< First < Prev Page 1 of 6455 Next > Last >>

- [Mus musculus esterase D/formylglutathione hydrolase \(Esd\), transcript variant 2, mRNA](#)
 - 1,988 bp linear mRNA
Accession: NM_016903.5 GI: 550544231
[Protein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)
- [Mus musculus esterase D/formylglutathione hydrolase \(Esd\), transcript variant 1, mRNA](#)
 - 2,121 bp linear mRNA
Accession: NM_001285423.1 GI: 550544229
[Protein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)

Results by taxon

Top Organisms [\[Tree\]](#)

- Mus musculus (116395)
- Escherichia coli (2339)
- Lactobacillus reuteri (1328)
- Yersinia pestis PY-36 (592)
- Yersinia pestis PY-48 (585)
- All other taxa (7857)

More...

Find related data

Database:

Find items

NCBI REFERENCE SEQUENCE DATABASE

NCBI Resources How To Sign in to NCBI

Nucleotide Nucleotide mus musculus AND srcdb_refseq[PROP] Search

Create alert Advanced Help

Species Summary 20 per page Sort by Default order Send to: Filters: Manage Filters

Animals (116,577) Plants (13) Fungi (175) Protists (256) Bacteria (12,056) Viruses (20) Customize ...

Molecule types genomic DNA/RNA (20,205) mRNA (77,471) rRNA (61) Customize ...

Source databases RefSeq (129,096) Customize ...

Genetic

Items: 1 to 20 of 129096

<< First < Prev Page 1 of 6455 Next > Last >>

[Mus musculus esterase D/formylglutathione hydrolase \(Esd\), transcript variant 2, mRNA](#)

1. 1,988 bp linear mRNA
Accession: NM_016903.5 GI: 550544231
[Protein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)

[Mus musculus esterase D/formylglutathione hydrolase \(Esd\), transcript variant 1, mRNA](#)

2. 2,121 bp linear mRNA
Accession: NM_001285423.1 GI: 550544229
[Protein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)

Results by taxon

Top Organisms [Tree]


- Mus musculus (116395)
- Escherichia coli (2339)
- Lactobacillus reuteri (1328)
- Yersinia pestis PY-36 (592)
- Yersinia pestis PY-48 (585)
- All other taxa (7857)

More...

Find related data

Database: Select

Find items



NCBI REFERENCE SEQUENCE DATABASE

Species Summary ▾ 20 per page ▾ Sort by Default order ▾ Send to: ▾ Filters: [Manage Filters](#)

Animals (116,577)
Plants (13)
Fungi (175)
Protists (256)
Bacteria (12,056)
Viruses (20)
Customize ...

Molecule types
genomic DNA/RNA (20,205)
mRNA (77,471)
rRNA (61)
Customize ...

Source database
RefSeq (129,056)
Customize ...

Genetic compartment
Mitochondrion
Plasmid (55)

Sequence length
Custom range

Release date
Custom range

Items: 1 to 20 of 129096

<< First < Prev Page 1 of 6455 Next > Last >>

[Mus musculus esterase D/formylglutathione hydrolase \(Esd\), transcript variant 2, mRNA](#)
1. 1,988 bp linear mRNA
Accession: NM_016903.5 GI: 550544231
[Protein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)

[formylglutathione hydrolase \(Esd\), transcript variant 1, mRNA](#)
550544229

[ate isomerase 1 \(Tpi1\), mRNA](#)
958348

[ultra-variant 4 \(Ulv4\), transcript variant 1, mRNA](#)

Results by taxon

Top Organisms [\[Tree\]](#)
Mus musculus (116395)
Escherichia coli (2339)
Lactobacillus reuteri (1328)
Yersinia pestis PY-36 (592)
Yersinia pestis PY-48 (585)
All other taxa (7857)
More...

Find related data

Database:

Search details

```
("Mus musculus"[Organism] OR mus musculus[All Fields]) AND srcdb_refseq[PROP]
```

NCBI DOWNLOAD FTP (FILE TRANSFER PROTOCOL)

NCBI Resources How To

NCBI
National Center for
Biotechnology Information

All Databases

NCBI Home

Resource List (A-Z)

All Resources

Chemicals & Bioassays

Data & Software

DNA & RNA

Domains & Structures

Genes & Expression

Genetics & Medicine

Genomes & Maps

Homology

Literature

Welcome to NCBI

The National Center for Biotechnology Information advances scientific biomedical and genomic information.

[About the NCBI](#) | [Mission](#) | [Organization](#) | [NCBI News & Blog](#)

Submit
Deposit data or manuscripts into NCBI databases

Download
Transfer NCBI data to your computer

Indeks – /genomes/M_musculus/

[\[katalog główny\]](#)

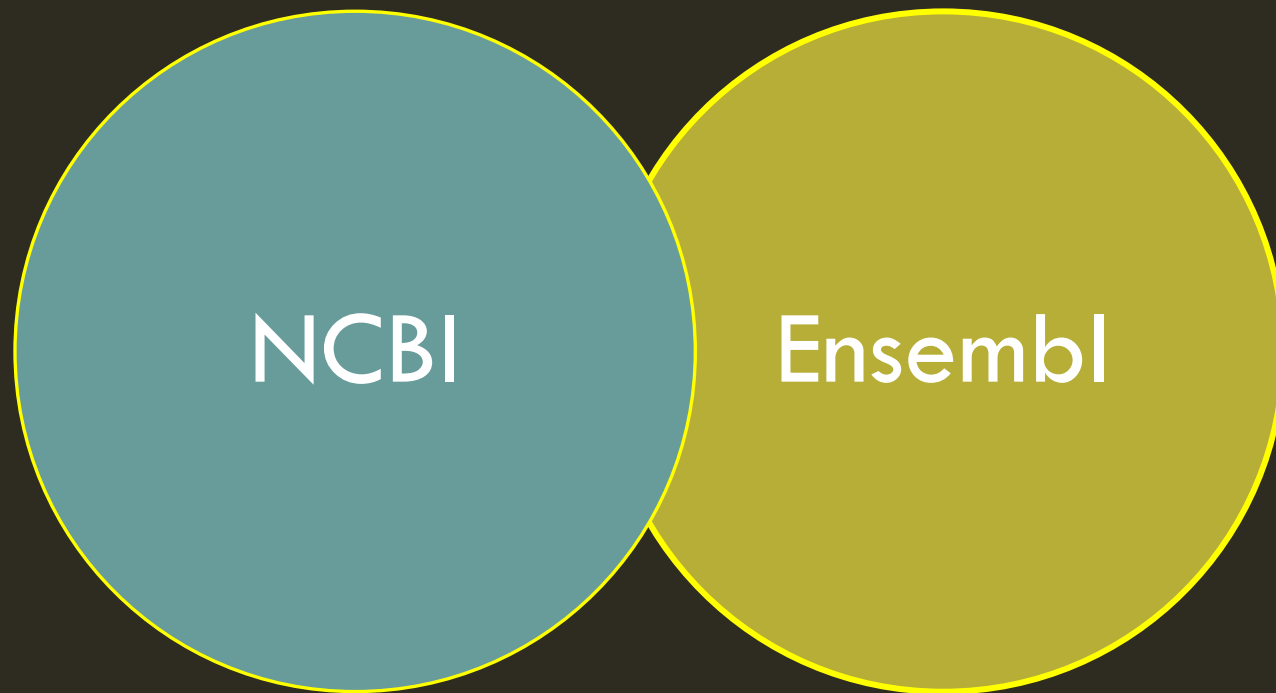
Nazwa	Rozmiar	Data modyfikacji
ARCHIVE/		24.06.2016, 02:00:00
Assembled_chromosomes/		24.06.2016, 02:00:00
CHR_01/		27.07.2016, 02:00:00
CHR_02/		27.07.2016, 02:00:00
CHR_03/		27.07.2016, 02:00:00
CHR_04/		27.07.2016, 02:00:00
CHR_05/		27.07.2016, 02:00:00
CHR_06/		27.07.2016, 02:00:00
CHR_07/		27.07.2016, 02:00:00
CHR_08/		27.07.2016, 02:00:00
CHR_09/		27.07.2016, 02:00:00
CHR_10/		27.07.2016, 02:00:00
CHR_11/		27.07.2016, 02:00:00
CHR_12/		27.07.2016, 02:00:00
CHR_13/		27.07.2016, 02:00:00
CHR_14/		27.07.2016, 02:00:00
CHR_15/		27.07.2016, 02:00:00
CHR_16/		27.07.2016, 02:00:00
CHR_17/		27.07.2016, 02:00:00

NCBI REFERENCE SEQUENCE DATABASE

Identyfikatory:

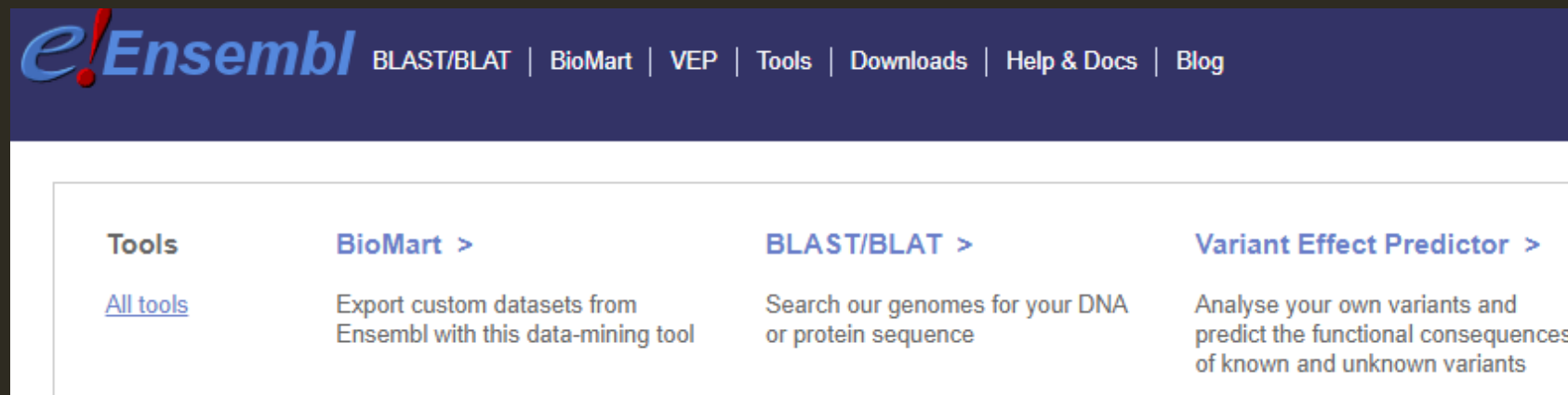
- **N**M_numer / **X**M_numer mRNA
 - **N**P_numer / **X**P_numer białka
 - **N**R_numer / **X**R_numer niekodujące RNA
 - **N**C_numer, **N**G_numer kontigi, sekwencje genomowe
-
- **N** – sekwencje uzyskane z wyników eksperymentów
 - **X** – sekwencje z adnotacji (sekwencja została przewidziana np. przez zmapowanie białka do genomu spokrewnionego organizmu)

GENOMY I ICH ADNOTACJE



ENSEMBL

- „Przeglądarka” genomów **kręgowców**
- Wspiera badania z zakresu genomiki porównawczej, ewolucji i regulacji ekspresji genów
- Dostępne narzędzia: **BLAST**, **BLAT**, **BioMart**, Variant Effect Predictor (**VEP**)



The screenshot shows the top navigation bar of the Ensembl website. The Ensembl logo is on the left, followed by a horizontal menu with links: BLAST/BLAT, BioMart, VEP, Tools, Downloads, Help & Docs, and Blog. Below the navigation bar is a grid of four tool cards. Each card has a title, a link with a chevron, and a brief description of the tool's function.

Tools	BioMart >	BLAST/BLAT >	Variant Effect Predictor >
All tools	Export custom datasets from Ensembl with this data-mining tool	Search our genomes for your DNA or protein sequence	Analyse your own variants and predict the functional consequences of known and unknown variants

ENSEMBL

- Ensembl
- EnsemblGenomes
- EnsemblFungi
- EnsemblMetazoa
- EnsemblProtists
- EnsemblBacteria
- EnsemblPlants
- PreEnsembl

The screenshot shows the top navigation bar of the Ensembl website. The logo 'e!Ensembl' is on the left, followed by links for 'BLAST/BLAT', 'BioMart', 'VEP', 'Tools', 'Downloads', 'Help & Docs', and 'Blog'. Below the navigation bar, there are four main tool categories: 'Tools' (with a link to 'All tools'), 'BioMart >' (described as 'Export custom datasets from Ensembl with this data-mining tool'), 'BLAST/BLAT >' (described as 'Search our genomes for your DNA or protein sequence'), and 'Variant Effect Predictor >' (described as 'Analyse your own variants and predict the functional consequences of known and unknown variants'). At the bottom of the screenshot, there is a search bar with a dropdown menu set to 'All species', a text input field, and a 'Go' button. Below the search bar, there is an example text: 'e.g. BRCA2 or rat 5:62797383-63627669 or rs699 or coronary heart disease'.

ENSEMBL

Identyfikatory (*Homo sapiens*):

- ENS**G** → Gene
- EN**ST** → Transcript
- EN**SP** → Protein
- EN**SE** → Exon

Identyfikatory (*inne gatunki*):

- ENS**FCAT**00000032635 *Felis catus*
- EN**SRNOG**00000050313 *Rattus norvegicus*
- EN**SCAFG**00000022708 *Canis lupus familiaris*
- EN**BTAT**00000064726 *Bos taurus*

Zagadka:

ENST00000471181.7

?

?

ENSRNOT00000075759.1

?

?

ENSSSCG00000018060

?

?

ENSEMBL

The screenshot shows the Ensembl website interface. At the top is a dark blue navigation bar with the Ensembl logo and links for BLAST/BLAT, BioMart, VEP, Tools, Downloads, Help & Docs, and Blog. Below this is a white navigation area with four main sections: Tools (with a link to 'All tools'), BioMart (described as an export tool), BLAST/BLAT (described as a search tool for DNA/protein), and Variant Effect Predictor (described as a variant analysis tool). The central part of the page features a search bar with a dropdown menu set to 'All species', a text input field, and a 'Go' button. Below the search bar, there are two columns: 'All genomes' with a species selection dropdown and links to view all species and edit favourites; and 'Favourite genomes' which lists Human (GRCh38.p12), Mouse (GRCm38.p6), and Zebrafish (GRCz11) with small representative images.

e!Ensembl BLAST/BLAT | BioMart | VEP | Tools | Downloads | Help & Docs | Blog

Tools
[All tools](#)

BioMart >
Export custom datasets from Ensembl with this data-mining tool

BLAST/BLAT >
Search our genomes for your DNA or protein sequence

Variant Effect Predictor >
Analyse your own variants and predict the functional consequences of known and unknown variants

Search

All species ▼ for


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


All genomes


-- Select a species -- ▼

- [View full list of all Ensembl species](#)
- [Edit your favourites](#)

Favourite genomes

 **Human**
GRCh38.p12
[Still using GRCh37?](#)

 **Mouse**
GRCm38.p6

 **Zebrafish**
GRCz11

ENSEMBL

e!Ensembl [BLAST/BLAT](#) | [BioMart](#) | [VEP](#) | [Tools](#) | [Downloads](#) | [Help & Docs](#) | [Blog](#)

Tools [All tools](#)

BioMart > Export custom datasets from Ensembl with this data mining tool

BLAST/BLAT > Search our genomes for your DNA or protein sequence

Variant Effect Predictor > Analyse your own variants and predict the functional consequences of known and unknown variants

Search

Cow for
19:12,571,690-12,571,788 **Go**


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


All genomes


-- Select a species --

- [View full list of all Ensembl species](#)
- [Edit your favourites](#)

Favourite genomes

 **Human**
GRCh38.p12
[Still using GRCh37?](#)

 **Mouse**
GRCm38.p6

 **Zebrafish**
GRCz11

ENSEMBL

Chromosome 19: 12,571,690-12,571,788

Chr. 19

Region in detail

Scroll: Track height: Drag/Select:

Contigs
Genes (Ensembl)

Gene Legend
■ Ensembl protein coding
■ RNA gene
■ pseudogene

Location: 19:12571690-12571788 Go Gene: Go

Genes (Ensembl)
Sequence
Contigs
Sequence
Genes (Ensembl)

Gene Legend
Protein Coding
■ Ensembl protein coding

There are currently 57 tracks turned off.
Ensembl Bostaurus version 87.31 (UMD3.1) Chromosome 19: 12,571,690 - 12,571,788

ENSEMBL

The screenshot shows the Ensembl website interface. At the top, the Ensembl logo is followed by navigation links: BLAST/BLAT, BioMart, VEP, Tools, Downloads, Help & Docs, and Blog. Below this is a horizontal menu with four items: Tools, BioMart, BLAST/BLAT, and Variant Effect Predictor. A large green arrow points to the BLAST/BLAT link. The BLAST/BLAT section contains a search form with a dropdown menu set to 'Human' and a text input field containing 'BRCA1'. A 'Go' button is to the right of the input field. Below the search form, there are examples of search terms: 'e.g. BRCA2 or rat 5:62797383-63627669 or rs699 or coronary heart disease'. At the bottom of the page, there are two sections: 'All genomes' with a species selection dropdown and links to 'View full list of all Ensembl species' and 'Edit your favourites'; and 'Favourite genomes' which lists 'Human' (GRCh38.p12), 'Mouse' (GRCm38.p6), and 'Zebrafish' (GRCz11), each with a small representative image.

Na którym chromosomie leży gen BRCA1?
Ile ma form splicingowych?

ENSEMBL

Gene: BRCA1 ENSG0000012048

Description BRCA1 DNA repair associated [Source:HGNC Symbol;Acc:HGNC:1100]

Gene Synonyms BRCC1, FANCS, PPP1R53, RNF53

Location [Chromosome 17: 43,044,295-43,170,245](#) reverse strand.
GRCh38:CM000679.2

About this gene This gene has 34 transcripts ([splice variants](#)), [195 orthologues](#) and is associated with [134 phenotypes](#).

Transcripts [Hide transcript table](#)

Show/hide columns (1 hidden)								Filter				
Name	Transcript ID	bp	Protein	Biotype	CCDS	UniProt Match	RefSeq Match	Flags				
BRCA1-203	ENST00000357654.9	7088	1863aa	Protein coding	CCDS11453	P38398-1	NM_007294.4	MANE Select v0.93	Ensembl Canonical	GENCODE basic	APPRIS P2	TSL:1
BRCA1-210	ENST00000471181.7	7270	1884aa	Protein coding	CCDS11456	P38398-7	-	GENCODE basic	APPRIS ALT2	TSL:1		
BRCA1-221	ENST00000493795.5	5732	1816aa	Protein coding	CCDS11459	P38398-8	-	GENCODE basic	TSL:5			
BRCA1-208	ENST00000468300.5	3273	699aa	Protein coding	CCDS11455	P38398-6	-	GENCODE basic	TSL:1			
BRCA1-219	ENST00000491747.6	2379	759aa	Protein coding	CCDS11454	P38398-3	-	GENCODE basic	TSL:5			
BRCA1-202	ENST00000354071.7	4497	1399aa	Protein coding	-	Q5YLB2	-	GENCODE basic	TSL:1			
BRCA1-201	ENST00000352993.7	3668	721aa	Protein coding	-	P38398-5	-	GENCODE basic	TSL:5			
BRCA1-232	ENST00000644379.1	2571	659aa	Protein coding	-	A0A2R8Y7V5	-	CDS 5' incomplete				
BRCA1-230	ENST00000634433.1	2534	798aa	Protein coding	-	A0A0U1RRA9	-	TSL:5	CDS 3' incomplete			
BRCA1-234	ENST00000652672.1	2291	601aa	Protein coding	-	A0A494C182	-	CDS 3' incomplete				
BRCA1-209	ENST00000470026.5	2108	649aa	Protein coding	-	E7EWN5	-	TSL:1	CDS 3' incomplete			
BRCA1-214	ENST00000477152.5	1980	622aa	Protein coding	-	E9PH68	-	TSL:1	CDS 3' incomplete			
BRCA1-215	ENST00000478531.5	1972	623aa	Protein coding	-	E7EUM2	-	TSL:1	CDS 3' incomplete			

Na którym chromosomie leży gen BRCA1?
Ile ma form splicingowych?

ENSEMBL

Gene: BRCA1 ENSG0000012048

Description BRCA1 DNA repair associated [Source:HGNC Symbol;Acc:HGNC:1100]

Gene Synonyms BRCC1, FANCS, FANCD1, RNF53

Location Chromosome 17q21.31:43,170,245 reverse strand.
GRCh38:CM000679

About this gene This gene has 34 transcripts (splice variants), 195 orthologues and is associated with 134 phenotypes.

Transcripts

Hide transcript table

Name	Transcript ID	bp	Protein	Biotype	CCDS	UniProt Match	RefSeq Match	Flags
BRCA1-203	ENST00000357654.9	7088	1863aa	Protein coding	CCDS11453	P38398-1	NM_007294.4	MANE Select v0.93 Ensembl Canonical GENCODE basic APPRIS P2 TSL:1
BRCA1-210	ENST00000471181.7	7270	1884aa	Protein coding	CCDS11456	P38398-7	-	GENCODE basic APPRIS ALT2 TSL:1
BRCA1-221	ENST00000493795.5	5732	1816aa	Protein coding	CCDS11459	P38398-8	-	GENCODE basic TSL:5
BRCA1-208	ENST00000468300.5	3273	699aa	Protein coding	CCDS11455	P38398-6	-	GENCODE basic TSL:1
BRCA1-219	ENST00000491747.6	2379	759aa	Protein coding	CCDS11454	P38398-3	-	GENCODE basic TSL:5
BRCA1-202	ENST00000354071.7	4497	1399aa	Protein coding	-	Q5YLB2	-	GENCODE basic TSL:1
BRCA1-201	ENST00000352993.7	3668	721aa	Protein coding	-	P38398-5	-	GENCODE basic TSL:5
BRCA1-232	ENST00000644379.1	2571	659aa	Protein coding	-	A0A2R8Y7V5	-	CDS 5' incomplete
BRCA1-230	ENST00000634433.1	2534	798aa	Protein coding	-	A0A0U1RRA9	-	TSL:5 CDS 3' incomplete
BRCA1-234	ENST00000652672.1	2291	601aa	Protein coding	-	A0A494C182	-	CDS 3' incomplete
BRCA1-209	ENST00000470026.5	2108	649aa	Protein coding	-	E7EWN5	-	TSL:1 CDS 3' incomplete
BRCA1-214	ENST00000477152.5	1980	622aa	Protein coding	-	E9PH68	-	TSL:1 CDS 3' incomplete
BRCA1-215	ENST00000478531.5	1972	623aa	Protein coding	-	E7EUM2	-	TSL:1 CDS 3' incomplete

Który transkrypt jest najdłuższy?

ENSEMBL

Gene: **BRCA1** ENSG0000012048

Description BRCA1 DNA repair associated [Source:HGNC Symbol;Acc:HGNC:1100]

Gene Synonyms BRCC1, FANCS, PPP1R53, RNF53

Location [Chromosome 17: 43,044,295-43,170,245](#) reverse strand.
GRCh38:CM000679.2

About this gene This gene has 34 transcripts ([splice variants](#)), [195 orthologues](#) and is associated with [134 phenotypes](#).

Transcripts [hide transcript table](#)

Show/hide columns (1 hidden) Filter

Name	Transcript ID	bp	Protein	Biotype	CCDS	UniProt Match	RefSeq Match	Flags
BRCA1-203	ENST00000357654.9	7088	1863aa	Protein coding	CCDS11453	P38398-1	NM_007294.4	MANE Select v0.93 Ensembl Canonical GENCODE basic APPRIS P2 TSL:1
BRCA1-210	ENST00000471181.7	7270	1884aa	Protein coding	CCDS11456	P38398-7	-	GENCODE basic APPRIS ALT2 TSL:1
BRCA1-221	ENST00000493795.5	5732	1816aa	Protein coding	CCDS11459	P38398-8	-	GENCODE basic TSL:5
BRCA1-208	ENST00000468300.5	3273	699aa	Protein coding	CCDS11455	P38398-6	-	GENCODE basic TSL:1
BRCA1-219	ENST00000491747.6	2379	759aa	Protein coding	CCDS11454	P38398-3	-	GENCODE basic TSL:5
BRCA1-202	ENST00000354071.7	4497	1399aa	Protein coding	-	Q5YLB2	-	GENCODE basic TSL:1
BRCA1-201	ENST00000352993.7	3668	721aa	Protein coding	-	P38398-5	-	GENCODE basic TSL:5
BRCA1-232	ENST00000644379.1	2571	659aa	Protein coding	-	A0A2R8Y7V5	-	CDS 5' incomplete
BRCA1-230	ENST00000634433.1	2534	798aa	Protein coding	-	A0A0U1RRA9	-	TSL:5 CDS 3' incomplete
BRCA1-234	ENST00000652672.1	2291	601aa	Protein coding	-	A0A494C182	-	CDS 3' incomplete
BRCA1-209	ENST00000470026.5	2108	649aa	Protein coding	-	E7EWN5	-	TSL:1 CDS 3' incomplete
BRCA1-214	ENST00000477152.5	1980	622aa	Protein coding	-	E9PH68	-	TSL:1 CDS 3' incomplete
BRCA1-215	ENST00000478531.5	1972	623aa	Protein coding	-	E7EUM2	-	TSL:1 CDS 3' incomplete

Downloads → FTP Download

ENSEMBL

Show 10 entries		Show/hide columns												Filter		
★	Species	DNA (FASTA)	cDNA (FASTA)	CDS (FASTA)	ncRNA (FASTA)	Protein sequence (FASTA)	Annotated sequence (EMBL)	Annotated sequence (GenBank)	Gene sets	Whole databases	Variation (GVF)	Variation (VCF)	Variation (VEP)	Regulation (GFF)	Data files	BAM/BigWig
Y	Human <i>Homo sapiens</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	GVF	VCF	VEP	Regulation (GFF)	Regulation data files	BAM/BigWig
Y	Mouse <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	GVF	VCF	VEP	Regulation (GFF)	Regulation data files	BAM/BigWig
Y	Zebrafish <i>Danio rerio</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	GVF	VCF	VEP	-	-	BAM/BigWig
	Algerian mouse <i>Mus spretus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-
	Alpaca <i>Vicugna pacos</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-
	Amazon molly <i>Poecilia formosa</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	BAM/BigWig
	Angola colobus <i>Colobus angolensis palliatus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-
	Anole lizard <i>Anolis carolinensis</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	BAM/BigWig

Downloads → Download data via FTP

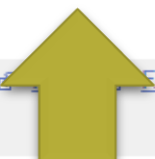
ENSEMBL



Show	10	entries	Show/hide columns											mus		
★	Species	DNA (FASTA)	cDNA (FASTA)	CDS (FASTA)	ncRNA (FASTA)	Protein sequence (FASTA)	Annotated sequence (EMBL)	Annotated sequence (GenBank)	Gene sets	Whole databases	Variation (GVF)	Variation (VCF)	Variation (VEP)	Regulation (GFF)	Data files	BAM/BigWig
	Mouse <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	GVF	VCF	VEP	Regulation (GFF)	Regulation data files	BAM/BigWig
	Algerian mouse <i>Mus spretus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-
	Ferret <i>Mustela putorius furo</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	BAM/BigWig
	Mouse 129S1/SvImJ <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-
	Mouse A/J <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-
	Mouse AKR/J <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-
	Mouse BALB/cJ <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-
	Mouse C3H/HeJ	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-

ENSEMBL

Show 10 entries		Show/hide columns												mus		
★	Species	DNA (FASTA)	cDNA (FASTA)	CDS (FASTA)	ncRNA (FASTA)	Protein sequence (FASTA)	Annotated sequence (EMBL)	Annotated sequence (GenBank)	Gene sets	Whole databases	Variation (GVF)	Variation (VCF)	Variation (VEP)	Regulation (GFF)	Data files	BAM/BigWig
Y	Mouse <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	GVF	VCF	VEP	Regulation (GFF)	Regulation data files	BAM/BigWig
	Algerian mouse <i>Mus spretus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-
	Ferret <i>Mustela putorius furo</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	BAM/BigWig
	Mouse	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF	MySQL	-	-	VEP	-	-	-



```
156.17.191.216 - PuTTY
>ENSMUST00000196221.1 cdna chromosome:GRCm38:14:54113468:54113476:1 gene:ENSMUSG00000096749.2 gene_biotype:TR_D_gene transcript_biotype:TR_D_gene gene_symbol:Trddl description:T cell receptor delta diversity 1 [Source:MGISymbol;Acc:MGISymbol:4439547]
ATGGCATAT
>ENSMUST00000179664.1 cdna chromosome:GRCm38:14:54113468:54113478:1 gene:ENSMUSG00000096749.2 gene_biotype:TR_D_gene transcript_biotype:processed_transcript gene_symbol:Trddl description:T cell receptor delta diversity 1 [Source:MGISymbol;Acc:MGISymbol:4439547]
ATGGCATATCA
>ENSMUST00000177564.1 cdna chromosome:GRCm38:14:54122226:54122241:1 gene:ENSMUSG00000096176.1 gene_biotype:TR_D_gene transcript_biotype:TR_D_gene gene_symbol:Trdd2 description:T cell receptor delta diversity 2 [Source:MGISymbol;Acc:MGISymbol:4439546]
ATCGGAGGGATACGAG
```

ENSEMBL

Show 10 entries		Show/hide columns												mus		
★	Species	DNA (FASTA)	cDNA (FASTA)	CDS (FASTA)	ncRNA (FASTA)	Protein sequence (FASTA)	Annotated sequence (EMBL)	Annotated sequence (GenBank)	Gene sets	Whole databases	Variation (GVF)	Variation (VCF)	Variation (VEP)	Regulation (GFF)	Data files	BAM/BigWig
Y	Mouse <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	GVF	VCF	VEP	Regulation (GFF)	Regulation data files	BAM/BigWig
	Algerian mouse <i>Mus spretus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-
	Ferret <i>Mustela putorius furo</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	BAM/BigWig
	Mouse	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF	MySQL	-	-	VEP	-	-	-



```

156.17.191.216 - PuTTY
>ENSMUST00000196221.1 cds chromosome:GRCm38:14:54113468:54113476:1 gene:ENSMUSG00000096749.2 gene_biotype:TR_D_gene transcript_biotype:TR_D_gene gene_symbol:Trdd1 description:T cell receptor delta diversity 1 [Source:MGI Symbol;Acc:MGI:4439547] ATGGCATAT
>ENSMUST00000177564.1 cds chromosome:GRCm38:14:54122226:54122241:1 gene:ENSMUSG00000096176.1 gene_biotype:TR_D_gene transcript_biotype:TR_D_gene gene_symbol:Trdd2 description:T cell receptor delta diversity 2 [Source:MGI Symbol;Acc:MGI:4439546] ATCGGAGGGGATACGAG
>ENSMUST00000178537.1 cds chromosome:GRCm38:6:41533201:41533212:1 gene:ENSMUSG00000095668.1 gene_biotype:TR_D_gene transcript_biotype:TR_D_gene gene_symbol:Trbd1 description:T cell receptor beta, D region 1 [Source:MGI Symbol;Acc:MGI:4439571] GGGACAGGGGGC
>ENSMUST00000178862.1 cds chromosome:GRCm38:6:41542163:41542176:1 gene:ENSMUSG00000094569.1 gene_biotype:TR_D_gene transcript_biotype:TR_D_gene gene_symbol:Trbd2 description:T cell receptor beta, D region 2 [Source:MGI Symbol;Acc:MGI:4439727] GGGACTGGGGGGGC
    
```

Non-coding RNA gene predictions.

ENSEMBL

Show 10 entries Show/hide columns mus

★	Species	DNA (FASTA)	cDNA (FASTA)	CDS (FASTA)	ncRNA (FASTA)	Protein sequence (FASTA)	Annotated sequence (EMBL)	Annotated sequence (GenBank)	Gene sets	Whole databases	Variation (GVF)	Variation (VCF)	Variation (VEP)	Regulation (GFF)	Data files	BAM/BigWig
Y	Mouse <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	GVF	VCF	VEP	Regulation (GFF)	Regulation data files	BAM/BigWig
	Algerian mouse <i>Mus spretus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-
	Ferret <i>Mustela putorius furo</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	BAM/BigWig
	Mouse	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-



```
156.17.191.216 - PuTTY
>ENSMUST00000083969.1 ncRNA chromosome:GRCm38:X:64800511:64800616:1 gene:ENSMUSG00000065903.1 gene_biotype:snRNA transcript_biotype:snRNA gene_symbol:Gm26111 description:predicted gene, 26111 [Source:MGI Symbol;Acc:MGI:5455888]
GTGCTCACTTCCGCAGCACATACTAAAATTGGAACAAAACAGAGAAGGTTAGCATGGC
CCTTGTACAAGTATGACATGCACGTTTCATGAAGCTTCCATATTTTT
>ENSMUST00000157441.1 ncRNA chromosome:GRCm38:2:71495271:71495374:1 gene:ENSMUSG00000088066.1 gene_biotype:snRNA transcript_biotype:snRNA gene_symbol:Gm23253 description:predicted gene, 23253 [Source:MGI Symbol;Acc:MGI:5453030]
GCTGTGCTCACTTCAGCAGCAAACAGTAAAATCGGAATGAGCAGAGGTTAGTATGACCTT
TGTACAAGGGTGACACAGAAAATTGTGAGGTGCTCCATTAATAAAA
>ENSMUST00000178877.1 ncRNA chromosome:GRCm38:16:14807163:14807269:-1 gene:ENSMUSG00000095165.1 gene_biotype:snRNA transcript_biotype:snRNA gene_symbol:Gm25547 description:predicted gene, 25547 [Source:MGI Symbol;Acc:MGI:5455324]
GTGCTCGCTTCGGCAGCACATATACTAAAATTGGAACGATACAGAGAAGATTAGCATGGC
CCCTGCGCAAGGATGACACGCAAATTCGTGAAGCGTTCATATTTTT
```


ENSEMBL

Show 10 entries		Show/hide columns												mus		
★	Species	DNA (FASTA)	cDNA (FASTA)	CDS (FASTA)	ncRNA (FASTA)	Protein sequence (FASTA)	Annotated sequence (EMBL)	Annotated sequence (GenBank)	Gene sets	Whole databases	Variation (GVF)	Variation (VCF)	Variation (VEP)	Regulation (GFF)	Data files	BAM/BigWig
Y	Mouse <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	GVF	VCF	VEP	Regulation (GFF)	Regulation data files	BAM/BigWig
	Algerian mouse <i>Mus spretus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-
	Ferret <i>Mustela putorius furo</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	BAM/BigWig
	Mouse 129S1/SvJm.1	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-



```

156.17.191.216 - PuTTY
>ENSMUSP00000142546.1 pep chromosome:GRCm38:14:54113468:54113476:1 gene:ENSMUSG00000096749.2 transcript:ENSMUST00000196221
.1 gene_biotype:TR_D_gene transcript_biotype:TR_D_gene gene_symbol:Trdd1 description:T cell receptor delta diversity 1 [Source:MGI Symbol;Acc:MGI:4439547]
MAY
>ENSMUSP00000142955.1 pep chromosome:GRCm38:14:54122226:54122241:1 gene:ENSMUSG00000096176.1 transcript:ENSMUST00000177564
.1 gene_biotype:TR_D_gene transcript_biotype:TR_D_gene gene_symbol:Trdd2 description:T cell receptor delta diversity 2 [Source:MGI Symbol;Acc:MGI:4439546]
IGGIR
>ENSMUSP00000141764.1 pep chromosome:GRCm38:6:41533201:41533212:1 gene:ENSMUSG00000095668.1 transcript:ENSMUST00000178537.
1 gene_biotype:TR_D_gene transcript_biotype:TR_D_gene gene_symbol:Trbd1 description:T cell receptor beta, D region 1 [Source:MGI Symbol;Acc:MGI:4439571]
GTGG
    
```

ENSEMBL

More extensive sequence annotation by means of feature tables and contain thus the genome sequence as annotated by the automated Ensembl genome annotation pipeline. Each nucleotide sequence record in a flat file represents a 1 Mb slice of the genome sequence.

Show	10	entries	Show/hide columns													mus
★	Species	DNA (FASTA)	cDNA (FASTA)	CDS (FASTA)	ncRNA (FASTA)	Protein sequence (FASTA)	Annotated sequence (EMBL)	Annotated sequence (GenBank)	Gene sets	Whole databases	Variation (GVF)	Variation (VCF)	Variation (VEP)	Regulation (GFF)	Data files	BAM/BigWig
Y	Mouse <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	GVF	VCF	VEP	Regulation (GFF)	Regulation data files	BAM/BigWig
	Algerian mouse <i>Mus spretus</i>	FASTA	FASTA	FASTA	FASTA	FASTA		GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-
	Ferret <i>Mustela putorius furo</i>	FASTA	FASTA	FASTA	FASTA	FASTA										
	Mouse 129S1/SvImJ <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA										
	Mouse A/J <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA										
	Mouse AKR/J <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA										
	Mouse BALB/cJ <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA										
	Mouse C3H/HeJ <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA										

156.17.191.216 - PuTTY

```
LOCUS      1 195471971 bp DNA HTG 9-MAR-2018
DEFINITION Mus musculus chromosome 1 GRCm38 full sequence 1..195471971 reannotated via Ensembl
ACCESSION  chromosome:GRCm38:1:1:195471971:1
VERSION    1GRCm38
KEYWORDS   .
SOURCE     house mouse
ORGANISM   Mus musculus
            Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Bilateria;
            Deuterostomia; Chordata; Craniata; Vertebrata; Gnathostomata;
            Teleostomi; Euteleostomi; Sarcopterygii; Dipnotetrapodomorpha;
            Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Boreoeutheria;
            Euarchontoglires; Glires; Rodentia; Myomorpha; Muroidea; Muridae.
COMMENT    This sequence was annotated by Ensembl(www.ensembl.org). Please visit the Ensembl
            or EnsemblGenomes web site, http://www.ensembl.org/ or
            http://www.ensemblgenomes.org/ for more information.
```

ENSEMBL

GTF → Gene sets for each species. These files include annotations of both coding and non-coding genes.

GFF3 → provides access to all annotated transcripts which make up an Ensembl gene set.

Show	10	entries	Show/hide columns													mus
★	Species	DNA (FASTA)	cDNA (FASTA)	CDS (FASTA)	ncRNA (FASTA)	Protein sequence (FASTA)	Annotated sequence (EMBL)	Annotated sequence (GenBank)	Gene sets	Whole databases	Variation (GVF)	Variation (VCF)	Variation (VEP)	Regulation (GFF)	Data files	BAM/BigWig
Y	Mouse <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	GVF	VCF	VEP	Regulation (GFF)	Regulation data files	BAM/BigWig
	Algerian mouse <i>Mus spretus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank		MySQL	-	-	VEP	-	-	-
	Ferret <i>Mustela putorius furo</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	BAM/BigWig
	Mouse 129S1/SvImJ <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-
	Mouse A/J <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-
	Mouse AKR/J <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-
	Mouse BALB/cJ <i>Mus musculus</i>	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-
	Mouse C3H/HeJ	FASTA	FASTA	FASTA	FASTA	FASTA	EMBL	GenBank	GTF GFF3	MySQL	-	-	VEP	-	-	-

GTF → Gene sets for each species. These files include annotations of both coding and non-coding genes

ENSEMBL

```
156.17.191.216 - PuTTY
#!genome-build GRCh38.p6
#!genome-version GRCh38
#!genome-date 2012-01
#!genome-build-accession NCBI:GCA_000001635.8
#!genebuild-last-updated 2018-01
1   havana  gene      3073253 3074322 .      +      .      gene_id "ENSMUSG00000102693"; gene_version "1"; gene_name
"4933401J01Rik"; gene_source "havana"; gene_biotype "TEC";
1   havana  transcript  3073253 3074322 .      +      .      gene_id "ENSMUSG00000102693"; gene_version "1"; tr
anscript_id "ENSMUST00000193812"; transcript_version "1"; gene_name "4933401J01Rik"; gene_source "havana"; gene_biotype "T
EC"; transcript_name "4933401J01Rik-201"; transcript_source "havana"; transcript_biotype "TEC"; tag "basic"; transcript_su
pport_level "NA";
1   havana  exon       3073253 3074322 .      +      .      gene_id "ENSMUSG00000102693"; gene_version "1"; transcript
_id "ENSMUST00000193812"; transcript_version "1"; exon_number "1"; gene_name "4933401J01Rik"; gene_source "havana"; gene_b
iotype "TEC"; transcript_name "4933401J01Rik-201"; transcript_source "havana"; transcript_biotype "TEC"; exon_id "ENSMUSE0
0001343744"; exon_version "1"; tag "basic"; transcript_support_level "NA";
1   ensembl gene      3102016 3102125 .      +      .      gene_id "ENSMUSG00000064842"; gene_version "1"; gene_name
"Gm26206"; gene_source "ensembl"; gene_biotype "snRNA";
1   ensembl transcript  3102016 3102125 .      +      .      gene_id "ENSMUSG00000064842"; gene_version "1"; tr
anscript_id "ENSMUST00000082908"; transcript_version "1"; gene_name "Gm26206"; gene_source "ensembl"; gene_biotype "snRNA"
; transcript_name "Gm26206-201"; transcript_source "ensembl"; transcript_biotype "snRNA"; tag "basic"; transcript_support_
level "NA";
1   ensembl exon      3102016 3102125 .      +      .      gene_id "ENSMUSG00000064842"; gene_version "1"; transcript
_id "ENSMUST00000082908"; transcript_version "1"; exon_number "1"; gene_name "Gm26206"; gene_source "ensembl"; gene_biotyp
e "snRNA"; transcript_name "Gm26206-201"; transcript_source "ensembl"; transcript_biotype "snRNA"; exon_id "ENSMUSE0000052
2066"; exon_version "1"; tag "basic"; transcript_support_level "NA";
```

GFF3 → provides access to all annotated transcripts which make up an Ensembl gene set

ENSEMBL

```
156.17.191.216 - PuTTY
##gff-version 3
##sequence-region 1 1 195471971
#!genome-build Ensembl GRCm38.p6
#!genome-version GRCm38
#!genome-date 2012-01
#!genome-build-accession NCBI:GCA_000001635.8
#!genebuild-last-updated 2018-01
1 ensembl_havana gene 3205901 3671498 . - . ID=gene:ENSMUSG00000051951;Name=Xkr4;biotype=protein;description=X-linked Kx blood group related 4 [Source:MGI Symbol%3BAcc:MGI:3528744];gene_id=ENSMUSG00000051951;logic_name=ensembl_havana_gene;version=5
1 havana lnc_RNA 3205901 3216344 . - . ID=transcript:ENSMUST00000162897;Parent=gene:ENSMUSG00000051951;Name=Xkr4-203;biotype=processed_transcript;transcript_id=ENSMUST00000162897;transcript_support_level=1;version=1
1 havana exon 3205901 3207317 . - . Parent=transcript:ENSMUST00000162897;Name=ENSMUSE00000866652;constitutive=0;ensembl_end_phase=-1;ensembl_phase=-1;exon_id=ENSMUSE00000866652;rank=2;version=1
1 havana exon 3213609 3216344 . - . Parent=transcript:ENSMUST00000162897;Name=ENSMUSE00000858910;constitutive=0;ensembl_end_phase=-1;ensembl_phase=-1;exon_id=ENSMUSE00000858910;rank=1;version=1
1 havana lnc_RNA 3206523 3215632 . - . ID=transcript:ENSMUST00000159265;Parent=gene:ENSMUSG00000051951;Name=Xkr4-202;biotype=processed_transcript;transcript_id=ENSMUST00000159265;transcript_support_level=1;version=1
1 havana exon 3206523 3207317 . - . Parent=transcript:ENSMUST00000159265;Name=ENSMUSE00000867897;constitutive=0;ensembl_end_phase=-1;ensembl_phase=-1;exon_id=ENSMUSE00000867897;rank=2;version=1
1 havana exon 3213439 3215632 . - . Parent=transcript:ENSMUST00000159265;Name=ENSMUSE00000863980;constitutive=0;ensembl_end_phase=-1;ensembl_phase=-1;exon_id=ENSMUSE00000863980;rank=1;version=1
1 ensembl_havana mRNA 3214482 3671498 . - . ID=transcript:ENSMUST00000070533;Parent=gene:ENSMU
```

VEP → Analyse your own variants and predict the functional consequences of known and unknown variants

ENSEMBL TOOLS

The screenshot shows the Ensembl website interface. At the top, there is a navigation bar with the Ensembl logo and links for BLAST/BLAT, BioMart, VEP, Tools, Downloads, Help & Docs, and Blog. Below this, there are four main tool categories: Tools, BioMart, BLAST/BLAT, and Variant Effect Predictor. The Variant Effect Predictor is highlighted with a green arrow. Below the navigation bar, there is a search bar with a dropdown menu for species (currently set to 'All species') and a 'Go' button. Below the search bar, there are two columns: 'All genomes' and 'Favourite genomes'. The 'All genomes' column has a dropdown menu for species and links for 'View full list of all Ensembl species' and 'Edit your favourites'. The 'Favourite genomes' column lists three species: Human (GRCh38.p12), Mouse (GRCm38.p6), and Zebrafish (GRCz11).

Tools [All tools](#)

BioMart > Export custom datasets from Ensembl with this data-mining tool

BLAST/BLAT > Search our genomes for your DNA or protein sequence

Variant Effect Predictor > Analyse your own variants and predict the functional consequences of known and unknown variants

Search

All species for


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


All genomes


-- Select a species --

- [View full list of all Ensembl species](#)
- [Edit your favourites](#)

Favourite genomes

 **Human**
GRCh38.p12
[Still using GRCh37?](#)

 **Mouse**
GRCm38.p6

 **Zebrafish**
GRCz11

VEP → Analyse your own variants and predict the functional consequences of known and unknown variants

VEP

The screenshot shows the Ensembl website's Variant Effect Predictor (VEP) page. At the top, the Ensembl logo is on the left, and navigation links for BLAST/BLAT, BioMart, VEP, Tools, Downloads, Help & Docs, and Blog are in the center. On the right, there is a search bar for species and a Login/Register link. Below the navigation bar, there are tabs for 'Using this website', 'Annotation and prediction', 'Data access', 'API & software', and 'About us'. The 'API & software' tab is active, and the breadcrumb trail shows the path: Home > Help & Documentation > API & Software > Ensembl Tools > Variant Effect Predictor. On the left side, there is a sidebar menu titled 'In this section' with categories like 'Web interface' (Input form, Results) and 'VEP script' (Tutorial, Download and install, Running VEP, Annotation sources, Filtering results, Custom annotations, Plugins, Examples and use cases, Other information). Below the sidebar is a search box for documentation. The main content area features the 'Variant Effect Predictor' title, the 'Ve!P' logo, and a description: 'VEP determines the effect of your variants (SNPs, insertions, deletions, CNVs or structural variants) on genes, transcripts, and protein sequence, as well as regulatory regions.' It then lists the information provided by VEP: Genes and Transcripts affected, Location of variants, Consequence of variants on protein sequence, Known variants from the 1000 Genomes Project, SIFT and PolyPhen scores, and more. A large blue button with a play icon and the text 'Launch Ve!P' is positioned in the bottom right corner of the screenshot.

VEP

VEP → Analyse your own variants and predict the functional consequences of known and unknown variants

e!Ensembl BLAST/BLAT | BioMart | Tools | Downloads

Species: **VEP**

Web Tools

- Web Tools
 - BLAST/BLAT
 - Variant Effect Predictor**
 - File Chameleon
 - Assembly Converter
 - ID History Converter
- Configure this page
- Custom tracks
- Export data
- Share this page
- Bookmark this page

Variant Effect Predictor

VEP for Human GRCh37

If you are looking for VEP for Human

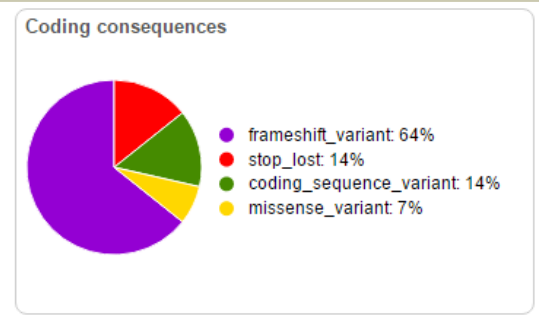
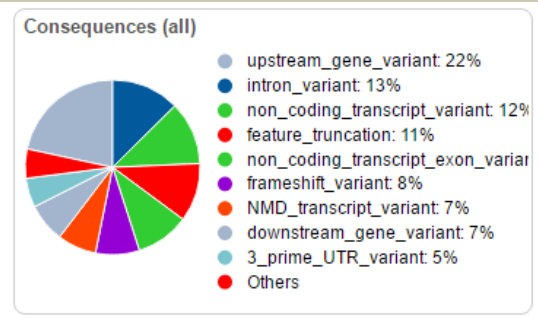
Species:

Name for this job (optional):

Either paste data:

Or upload file:

Category	Count
Variants processed	6
Variants filtered out	0
Novel / existing variants	5 (83.3) / 1 (16.7)
Overlapped genes	10
Overlapped transcripts	70
Overlapped regulatory features	1



Results preview

Navigation: Page: 1 of 1 | Show: All variants

Filters: is

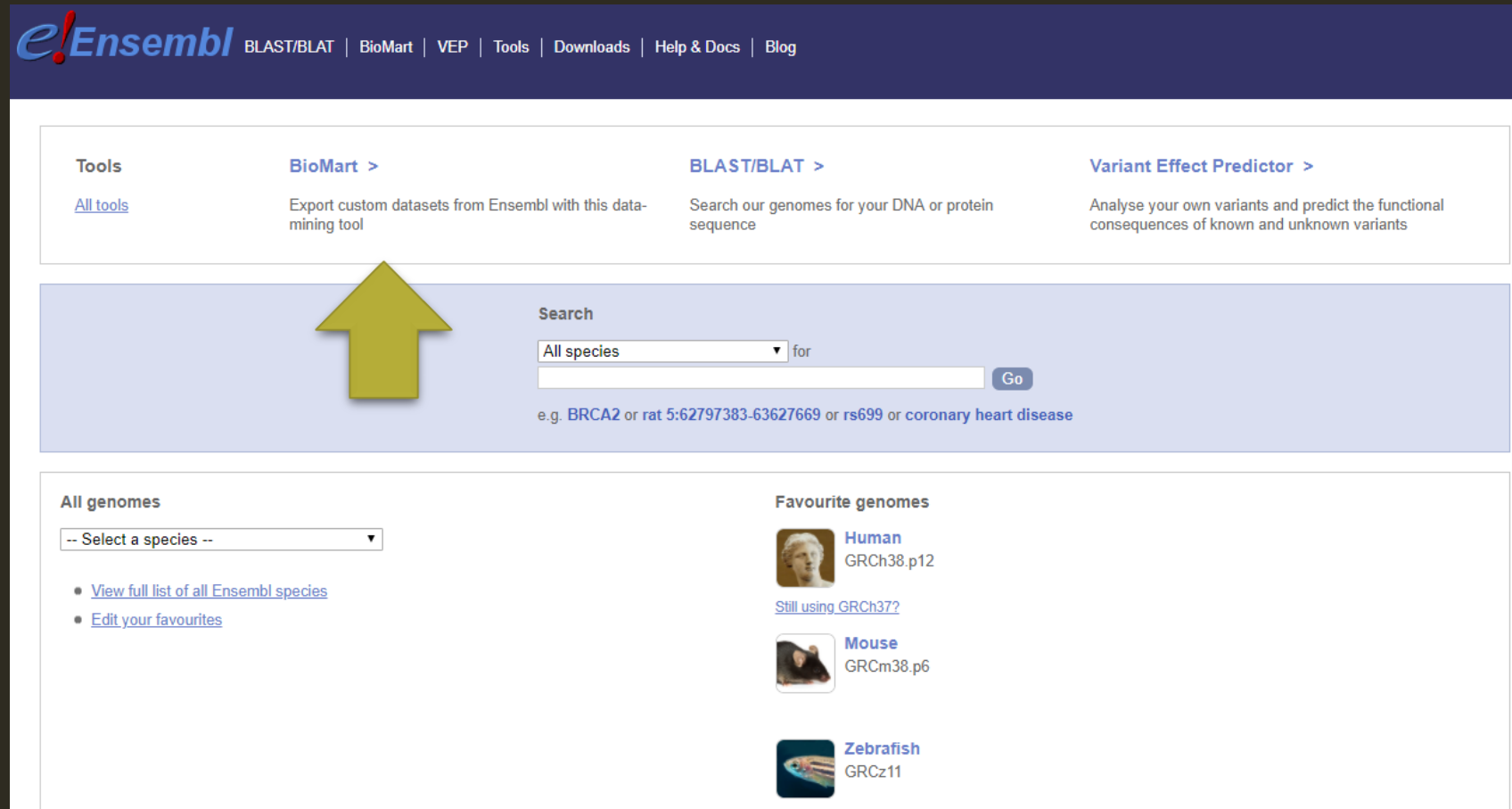
Download: All: [VCF VEP TXT](#) | BioMart: [Variants](#) [Genes](#)

Show/hide columns

Uploaded variant	Location	Allele	Consequence	Impact	Symbol	Gene	Feature type	Feature	Biotype
1_160283_duplication	1:160282-160282	duplication	upstream gene variant	MODIFIER	RNU6-1100P	ENSG00000222623	Transcript	ENST00000410691	snRNA
1_160283_duplication	1:160282-160282	duplication	non coding transcript exon variant, intron variant, non coding transcript variant	MODIFIER	RP11-34P13.13	ENSG00000241860	Transcript	ENST00000466557	lincRNA
1_160283_duplication	1:160282-160282	duplication	non coding transcript exon variant, intron variant, non coding transcript variant	MODIFIER	RP11-34P13.13	ENSG00000241860	Transcript	ENST00000491962	lincRNA
1_160283_duplication	1:160282-160282	duplication	non coding transcript exon variant, intron variant, non coding transcript variant	MODIFIER	RP11-34P13.9	ENSG00000241599	Transcript	ENST00000496488	lincRNA
1_182712_A/C	1:182712-182712	C	downstream gene variant	MODIFIER	FO538757.1	ENSG00000279457	Transcript	ENST00000623083	protein_coding
1_182712_A/C	1:182712-182712	C	downstream gene variant	MODIFIER	FO538757.1	ENSG00000279457	Transcript	ENST00000623834	protein_coding
1_182712_A/C	1:182712-182712	C	missense variant	MODERATE	FO538757.2	ENSG00000279928	Transcript	ENST00000624431	protein_coding
1_182712_A/C	1:182712-182712	C	downstream gene variant	MODIFIER	FO538757.1	ENSG00000279457	Transcript	ENST00000624735	protein_coding
1_1385015_deletion	1:1385014-1385014	deletion	stop lost, coding sequence variant, 3 prime UTR variant, feature truncation	HIGH	CCNL2	ENSG00000221978	Transcript	ENST00000400809	protein_coding

BioMart → Export custom datasets from Ensembl with this data-mining tool

BIOMART



The screenshot shows the Ensembl website's navigation bar and main content area. The navigation bar includes links for BLAST/BLAT, BioMart, VEP, Tools, Downloads, Help & Docs, and Blog. Below the navigation bar, there are four main tool categories: Tools, BioMart, BLAST/BLAT, and Variant Effect Predictor. The BioMart category is highlighted with a yellow arrow and includes the description: "Export custom datasets from Ensembl with this data-mining tool". Below the navigation bar, there is a search bar with a dropdown menu set to "All species" and a "Go" button. Below the search bar, there are two sections: "All genomes" and "Favourite genomes". The "All genomes" section has a dropdown menu set to "-- Select a species --" and links for "View full list of all Ensembl species" and "Edit your favourites". The "Favourite genomes" section lists three genomes: Human (GRCh38.p12), Mouse (GRCm38.p6), and Zebrafish (GRCz11).

Ensembl BLAST/BLAT | BioMart | VEP | Tools | Downloads | Help & Docs | Blog

Tools [All tools](#)

BioMart > Export custom datasets from Ensembl with this data-mining tool

BLAST/BLAT > Search our genomes for your DNA or protein sequence

Variant Effect Predictor > Analyse your own variants and predict the functional consequences of known and unknown variants

Search

All species for


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


All genomes


-- Select a species --

- [View full list of all Ensembl species](#)
- [Edit your favourites](#)

Favourite genomes

 **Human**
GRCh38.p12
[Still using GRCh37?](#)

 **Mouse**
GRCm38.p6

 **Zebrafish**
GRCz11

BioMart → Export custom datasets from Ensembl with this data-mining tool

BIOMART

Dataset
[None selected]

- CHOOSE DATABASE - ▾

↑

Dataset:
Ensembl genes
Mouse strains
Ensembl Variation
Ensembl Regulation

BioMart → Export custom datasets from Ensembl with this data-mining tool

BIOMART

Ensembl BLAST/BLAT | BioMart | BLAST | VEP | Tools | Downloads | Help & Docs | Blog Login/Register

Search all species...

New Count Results URL XML Perl Help

Dataset: Ensembl Genes 92
Cow genes (UMD3.1)
Filters: [None selected]
Attributes: Gene stable ID, Transcript stable ID

Dataset	[None Selected]
---------	-----------------

Filters → specyfikacja przeszukiwań

Attributes → określenie formatowania danych wyjściowe

Count → dostępne rekordy o określonych parametrach

Results → wyniki, eksport plików

BioMart → Export custom datasets from Ensembl with this data-mining tool

BIOMART

The screenshot displays the Ensembl BioMart interface. At the top, the Ensembl logo is followed by navigation links: BLAST/BLAT, BioMart, BLAST, VEP, Tools, Downloads, Help & Docs, and Blog. A search bar on the right contains the text "Search all species...". Below the navigation bar, there are buttons for "New", "Count", and "Results". To the right of these are buttons for "URL", "XML", "Perl", and "Help". The main content area features a dropdown menu for "Dataset" with "Ensembl Variation 92" selected. Below this is another dropdown menu for "Cow Structural Variants (UMD3.1)". On the left side, there is a sidebar with sections for "Filters" (Limit to variants from this source: DGVa) and "Attributes" (Study accession, Structural variant name, Chromosome/scaffold name, Chromosome/scaffold position start (bp), Chromosome/scaffold position end (bp)). At the bottom of the sidebar, there is a "Dataset" section with "[None Selected]". Three large green arrows are overlaid on the image: one pointing down to the "Ensembl Variation 92" dropdown, one pointing down to the "Cow Structural Variants (UMD3.1)" dropdown, and one pointing left towards the "Attributes" section.

BioMart → Export custom datasets from Ensembl with this data-mining tool

BIOMART

Ensembl BLAST/BLAT | BioMart | BLAST | VEP | Tools | Downloads | Help & Docs | Blog Login Register

Search all species...

New Count Results URL XML Perl Help

Dataset
Cow Structural Variants (UMD3.1)

Filters
Limit to variants from this source: DGVa
Chromosome/scaffold: 29

Attributes
Study accession
Structural variant name
Chromosome/scaffold name
Chromosome/scaffold position start (bp)
Chromosome/scaffold position end (bp)

Dataset
[None Selected]

Please restrict your query using criteria below
(If filter values are truncated in any lists, hover over the list item to see the full text)

REGION

Chromosome/scaffold 25 ▲
26
27
28
29
X ▼

Coordinates
Start
End

Marker
Marker Start
Marker End

BioMart → Export custom datasets from Ensembl with this data-mining tool

BIOMART

e!Ensembl BLAST/BLAT | BioMart | BLAST | VEP | Tools | Downloads | Help & Docs | Blog Login/Register

New **Count** **Results** [★ URL](#) [XML](#) [Perl](#) [Help](#)

Dataset 365 / 10462 Entries
Cow Structural Variants (UMD3.1)

Filters
Limit to variants from this source: DGVa
Chromosome/scaffold: 29

Attributes
Study accession
Structural variant name
Chromosome/scaffold name
Chromosome/scaffold position start (bp)
Chromosome/scaffold position end (bp)

Dataset
[None Selected]

Please restrict your query using criteria below
(If filter values are truncated in any lists, hover over the list item to see the full text)

REGION

Chromosome/scaffold

Coordinates

Start

End

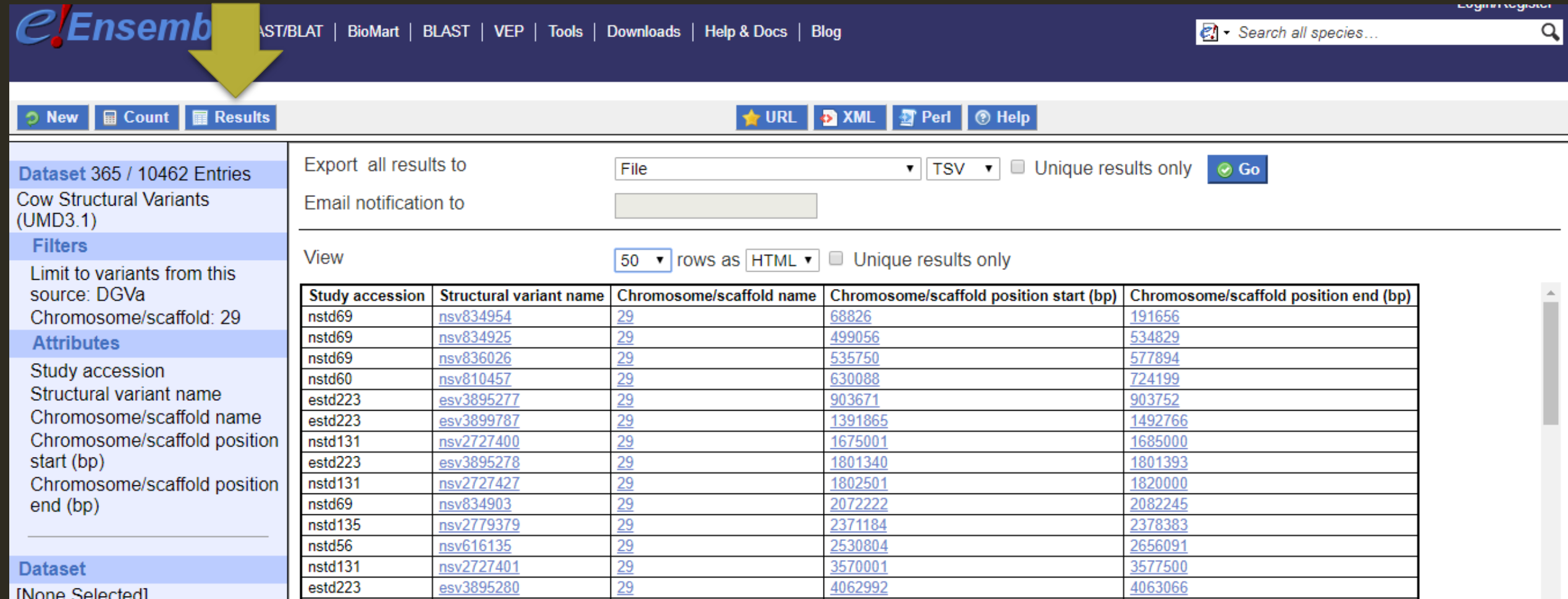
Marker

Marker Start

Marker End

BioMart → Export custom datasets from Ensembl with this data-mining tool

BIOMART



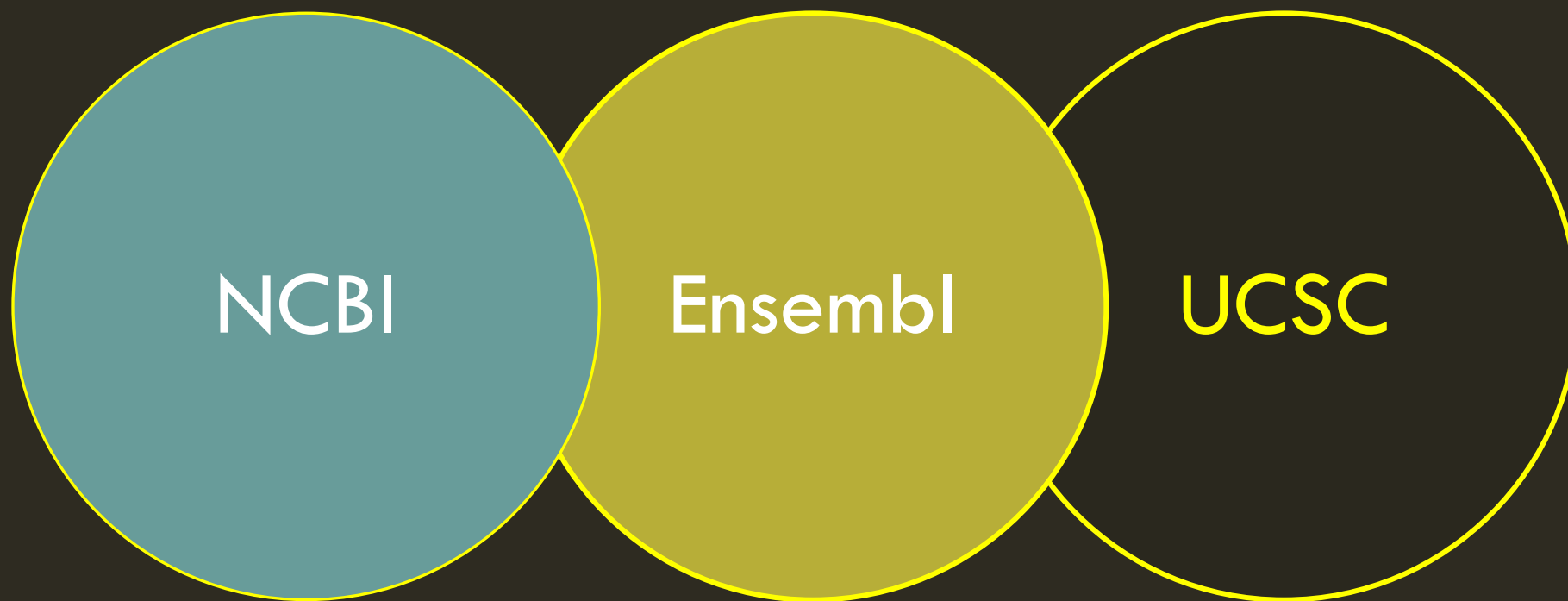
The screenshot displays the Ensembl BioMart interface. At the top, the Ensembl logo is on the left, and navigation links for BLAST/BLAT, BioMart, BLAST, VEP, Tools, Downloads, Help & Docs, and Blog are in the center. A search bar on the right contains the text "Search all species...". Below the navigation bar, there are buttons for "New", "Count", and "Results". A yellow arrow points from the "BIOMART" title to the BioMart link in the navigation bar.

The main content area shows the export options for a dataset. The dataset is identified as "Dataset 365 / 10462 Entries" and "Cow Structural Variants (UMD3.1)". The "Export all results to" section has a dropdown menu set to "File", a format dropdown set to "TSV", and a checkbox for "Unique results only" which is unchecked. A "Go" button is present. Below this, the "View" section shows a dropdown set to "50" rows as "HTML", with a "Unique results only" checkbox also unchecked.

The table below displays the first 15 rows of the dataset. The columns are: Study accession, Structural variant name, Chromosome/scaffold name, Chromosome/scaffold position start (bp), and Chromosome/scaffold position end (bp). The table contains 15 rows of data, each with a unique study accession and structural variant name, all located on chromosome/scaffold 29.

Study accession	Structural variant name	Chromosome/scaffold name	Chromosome/scaffold position start (bp)	Chromosome/scaffold position end (bp)
nstd69	nsv834954	29	68826	191656
nstd69	nsv834925	29	499056	534829
nstd69	nsv836026	29	535750	577894
nstd60	nsv810457	29	630088	724199
estd223	esv3895277	29	903671	903752
estd223	esv3899787	29	1391865	1492766
nstd131	nsv2727400	29	1675001	1685000
estd223	esv3895278	29	1801340	1801393
nstd131	nsv2727427	29	1802501	1820000
nstd69	nsv834903	29	2072222	2082245
nstd135	nsv2779379	29	2371184	2378383
nstd56	nsv616135	29	2530804	2656091
nstd131	nsv2727401	29	3570001	3577500
estd223	esv3895280	29	4062992	4063066

GENOMY I ICH ADNOTACJE





UNIVERSITY OF CALIFORNIA Genomics Institute UCSC Genome Browser

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

Our tools

- **Genome Browser**
interactively visualize genomic data
- **COVID-19 Research**
use the SARS-CoV-2 genome browser and explore coronavirus datasets
- **BLAT**
rapidly align sequences to the genome
- **Table Browser**
download data from the Genome Browser database
- **Variant Annotation Integrator**
get functional effect predictions for variant calls
- **Data Integrator**
combine data sources from the Genome Browser database
- **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
- **Track Hubs**
import and view external data tracks
- **REST API**
returns data in JSON format

[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 ensuring free public access to the genome and the information it contains. A few weeks later, on July 7, 2000, the newly assembled genome was released on the web at <http://genome.ucsc.edu>, along with the initial prototype of a graphical viewing tool, the UCSC Genome Browser. In the ensuing years, the website has grown to include a broad collection of vertebrate and model organism assemblies and annotations, along with a large suite of tools for viewing, analyzing and downloading data. Learn more about our history on the [UCSC Genome Browser Project History](#) page and by watching [this video](#).

What's new

- Sep. 27, 2021 - [JASPAR tracks for human \(hg19/hg38\)](#)
- Sep. 13, 2021 - [Cactus 241-way comparative genomics \(hg38\)](#)
- Aug. 31, 2021 - [GENCODE Genes VM27 for mouse \(mm39\)](#)






[More news...](#)

[Subscribe](#)

The UCSC Genome Browser is developed and maintained by the [Genome Bioinformatics Group](#), a cross-departmental team within the [UCSC Genomics Institute](#).

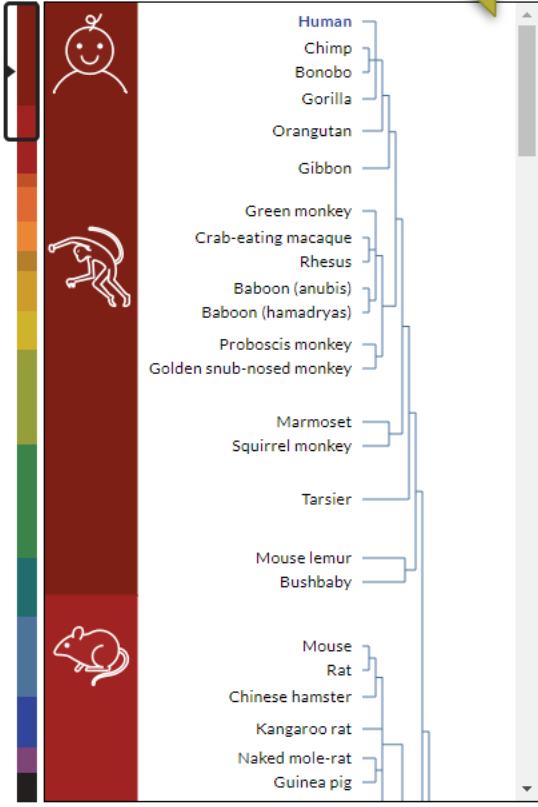
Browse/Select Species

POPULAR SPECIES

 Human
  Mouse
  Rat
  Fruitfly
  Worm


Enter species or common name

REPRESENTED SPECIES




Find Position

Human Assembly

Dec. 2013 (GRCh38/hg38) 

Position/Search Term

Enter position, gene symbol or search terms

Current position: chr1:11,102,837-11,267,747 

Human Genome Browser - hg38 assembly

[view sequences](#)

UCSC Genome Browser assembly ID: hg38
 Sequencing/Assembly provider ID: Genome Reference Consortium Human GRCh38 (GCA_000001405.15)
 Assembly date: Dec. 2013
 Accession ID: [GCA_000001405.15](#)
 NCBI Genome ID: 51 (Homo sapiens (human))
 NCBI Assembly ID: 883148 (GRCh38, GCA_000001405.15)
 BioProject ID: 31257



Homo sapiens
(Graphic courtesy of CBSE)

Search the assembly:

- **By position or search term:** Use the "position or search term" box to find areas of the genome associated with many different attributes, such as a specific chromosomal coordinate range; mRNA, EST, or STS marker names; or keywords from the GenBank description of an mRNA. [More information](#), including sample queries.
- **By gene name:** Type a gene name into the "search term" box, choose your gene from the drop-down list, then press "submit" to go directly to the assembly location associated with that gene. [More information](#).
- **By track type:** Click the "track search" button to find Genome Browser tracks that match specific selection criteria. [More information](#).

Download sequence and annotation data:

- [Using rsync \(recommended\)](#)
- [Using FTP](#)
- [Using HTTP](#)
- [Data use conditions and restrictions](#)
- [Acknowledgments](#)

Assembly Details

The GRCh38 assembly is the first major revision of the human genome released in more than four years. As with the previous assembly, the [Genome Reference Consortium \(GRC\)](#) is now the primary source for human genome assembly data submitted to GenBank. Beginning with this release, the UCSC Genome Browser version of the human assemblies now match those of the GRC to minimize version confusion. Hence, the GRCh38 assembly is referred to as "hg38" in the Genome Browser datasets and documentation. For more assembly-related terms, see the [GRC Assembly Terminology](#) page.

UCSC



Genome Browser

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



Our tools

- **Genome Browser**
interactively visualize genomic data
- **COVID-19 Research**
use the SARS-CoV-2 genome browser and explore coronavirus datasets
- **BLAT**
rapidly align sequences to the genome
- **Table Browser**
download data from the Genome Browser database
- **Variant Annotation Integrator**
get functional effect predictions for variant calls
- **Data Integrator**
combine data sources from the Genome Browser database
- **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
- **Track Hubs**
import and view external data tracks
- **REST API**
returns data in JSON format

[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 ensuring free public access to the genome and the information it contains. A few weeks later, on July 7, 2000, the newly assembled genome was released on the web at <http://genome.ucsc.edu>, along with the initial prototype of a graphical viewing tool, the UCSC Genome Browser. In the ensuing years, the website has grown to include a broad collection of vertebrate and model organism assemblies and annotations, along with a large suite of tools for viewing, analyzing and downloading data. Learn more about our history on the [UCSC Genome Browser Project History](#) page and by watching [this video](#).

What's new

Sep. 27, 2021 - [JASPAR tracks for human \(hg19/hg38\)](#)

Sep. 13, 2021 - [Cactus 241-way comparative genomics \(hg38\)](#)

Aug. 31, 2021 - [GENCODE Genes VM27 for mouse \(mm39\)](#)

[More news...](#)

[Subscribe](#)

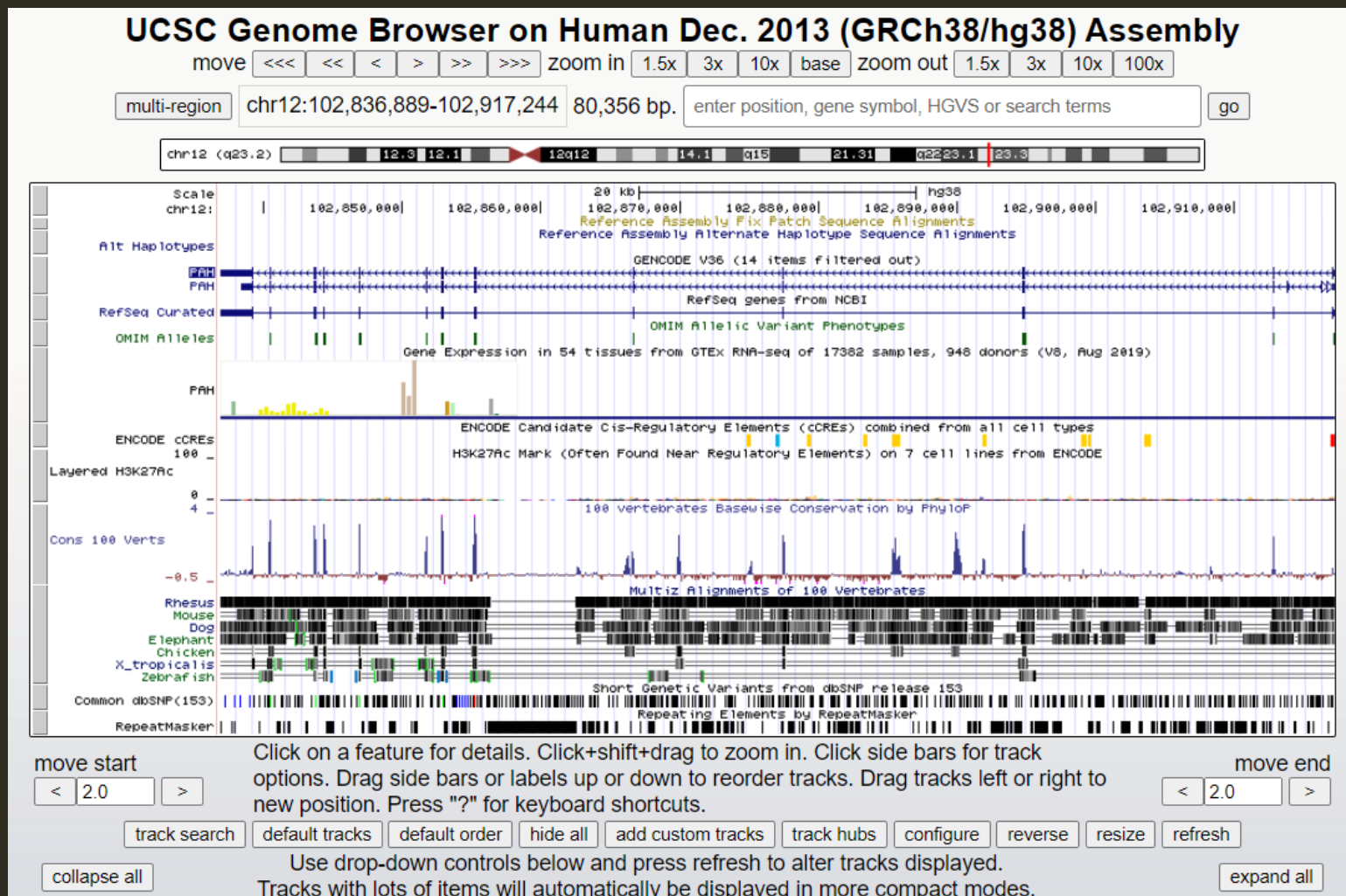
The UCSC Genome Browser is developed and maintained by the [Genome Bioinformatics Group](#), a cross-departmental team within the [UCSC Genomics Institute](#).

UCSC – GEN PAH

Gen PAH (enzym hydroksylazy fenyloalaninowej)

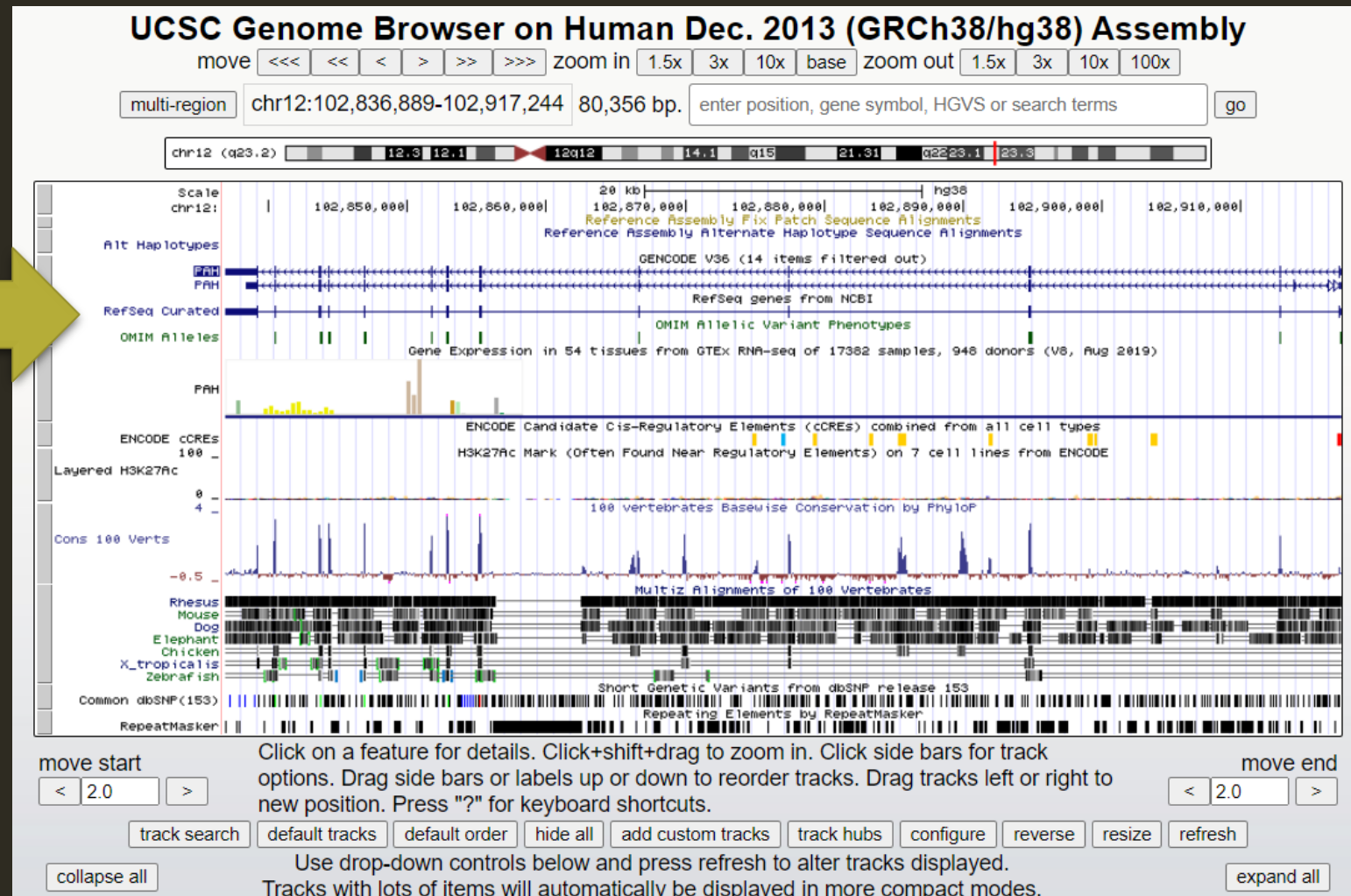
Na którym chromosomie jest zlokalizowany?

Jakiej długości jest sekwencja?



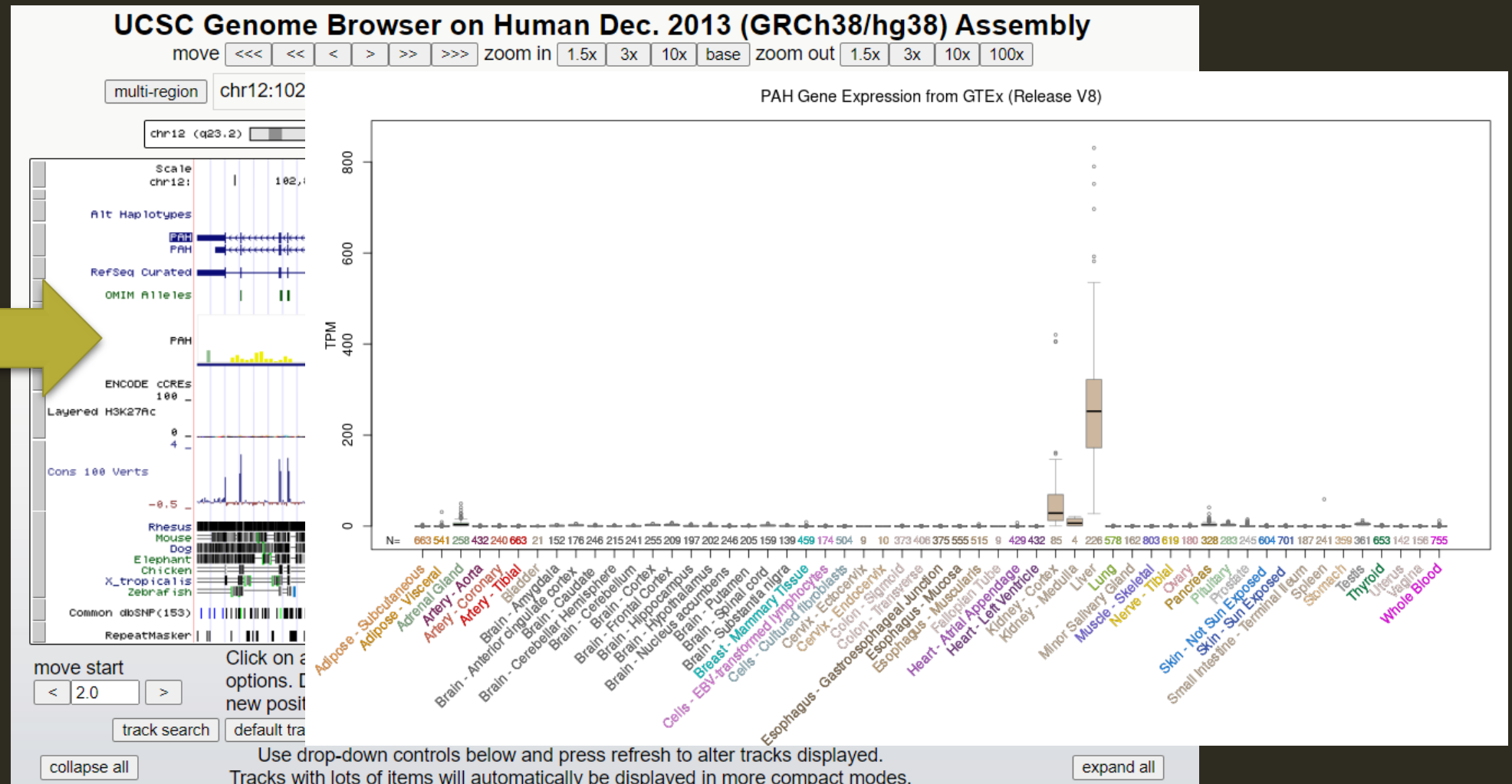
Gen PAH (enzym hydroksylazy fenyloalaninowej) Ile ma egzonów?

UCSC – GEN PAH



Gen PAH (enzym hydroksylazy fenyloalaninowej)
Gdzie ulega najwyższej ekspresji?

UCSC – GEN PAH



UCSC



UNIVERSITY OF CALIFORNIA
SANTA CRUZ UCSC Genome Browser

Genomes Genome Browser Tools Mirrors Downloads My Data Help About Us

Our tools

- **Genome Browser**
interactively visualize genomic data
- **BLAT**
rapidly align sequences to the genome
- **Table Browser**
download data from the Genome Browser database

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: Genes and Gene Predictions **track:** GENCODE v24

table: knownGene

region: genome position chr12:102836885-102917603

identifiers (names/accessions):

filter:

intersection:

correlation:

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

the first working draft of the human genome assembly, forever ensuring free public access to the genome and the information it contains. A few weeks later, on July 7, 2000, the newly assembled genome was released on the web at <http://genome.ucsc.edu>, along with the initial prototype of a graphical viewing tool, the UCSC Genome Browser. In the ensuing years, the website has grown to include a broad collection of vertebrate and model organism assemblies and annotations, along with a large suite of tools for viewing, analyzing and downloading data.

Apr. 03, 2018 - [New tool: Track Collection Builder](#)

Feb. 20, 2018 - [New video: Visibility control in the Browser](#)

[More news...](#)

[Subscribe](#)

The UCSC Genome Browser is developed and maintained by the Genome Bioinformatics Group, a cross-departmental team within the UCSC Genomics Institute.

Wszystkie znane geny z bazy UCSC dla myszy
Chromosome 1

UCSC — TABLE BROWSER

Format BED

The screenshot shows the UCSC Table Browser interface. At the top is a navigation bar with links: Genomes, Genome Browser, Tools, Mirrors, Downloads, My Data, Help, and About Us. Below the navigation bar is the 'Table Browser' title and a brief description of the tool's purpose. The main area contains several form fields and buttons for configuring the query:

- clade:** Mammal (dropdown)
- genome:** Mouse (dropdown)
- assembly:** Dec. 2011 (GRCm38/mm10) (dropdown)
- group:** Genes and Gene Predictions (dropdown)
- track:** UCSC Genes (dropdown)
- Buttons: add custom tracks, track hubs
- table:** knownGene (dropdown)
- Button: describe table schema
- region:** genome (radio), position (radio) chr1 (text input)
- Buttons: lookup, define regions
- identifiers (names/accessions):** paste list (button), upload list (button)
- filter:** create (button)
- intersection:** create (button)
- correlation:** create (button)
- output format:** BED - browser extensible data (dropdown)
- Send output to: Galaxy (checkbox), GREAT (checkbox), GenomeSpace (checkbox)
- output file:** (text input) (leave blank to keep output in browser)
- file type returned:** plain text (radio), gzip compressed (radio)
- Buttons: get output, summary/statistics

UCSC – TABLE BROWSER

Wszystkie znane geny z bazy UCSC dla myszy
Chromosome 1
Format BED

The screenshot shows the UCSC Table Browser interface. The top navigation bar includes links for Genomes, Genome Browser, Tools, Mirrors, Downloads, My Data, Help, and About Us. The main content area is titled 'Table Browser' and contains a form for configuring the output. The 'Output knownGene as BED' section is active, showing options to include a custom track header with fields for name (tb_knownGene), description (table browser query on knownGene), visibility (pack), and url. Below the form is a table of gene coordinates for Chromosome 1. The table has 11 columns: chromosome, start, end, gene name, score, strand, start, end, score, strand, and score. The data rows are:

chr1	3205903	3215632	uc007aet.1	0	-	3205903	3205903	0	2	1414,2194, 0,7535,
chr1	3214481	3671498	uc007aeu.1	0	-	3216021	3671348	0	3	2487,200,947, 0,207220,456070,
chr1	3648310	3658904	uc007aev.1	0	-	3648310	3648310	0	2	2199,58, 0,10536,
chr1	4290845	4409241	uc007aew.1	0	-	4292980	4409187	0	4	2167,172,636,72, 0,61064,61356,118324,

Below the table, there are radio buttons for selecting the output file type: 5' UTR Exons, Coding Exons, 3' UTR Exons, and Downstream by 200 bases. A note states: '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.' Buttons for 'get BED' and 'cancel' are visible at the bottom of the configuration panel. Two green arrows point upwards from the bottom left towards the 'get BED' button.