

WEBLEM 9

Introduction to Genomics & its various browser (UCSC, ENSEMBL, GDV)

Genomics is the **study of genomes**. Genomic studies are characterized by **simultaneous analysis of a large number of genes** using **automated data gathering tools**. The topics of **genomics range** from **genome mapping sequencing**, and **functional genomic analysis** to **comparative genomic analysis**. The **advent of genomics** and the ensuing explosion of **sequence information** are the main driving force behind the **rapid development** of bioinformatics today.

Genomic study can be **tentatively divided** into **structural genomics** and **functional genomics**. Structural genomics refers to the **initial phase of genome analysis**, which includes construction of **genetic and physical maps** of a genome, identification of genes, **annotation of gene features**, and **comparison of genome structures**. Functional genomics refers to the analysis of **global gene expression** and gene functions in a genome.

Genome browsers are resources that **integrate data at the genomic level**, thereby allowing visualization of related **genomic information** in one space. These data can **include genes, noncoding elements** that regulate **gene expression, genetic variation** and the results of **comparative genomics analyses**, among other forms of annotation. Commonly used genome **browsers include** Ensembl, the UCSC Genome Browser and GDV.

1. UCSC Genome Browser

The **University of California Santa Cruz (UCSC) Genome Browser** (genome.ucsc.edu) is a popular **Web-based tool** for quickly displaying a requested **portion of a genome** at any scale, accompanied by a series of **aligned annotation “tracks”**. The annotations generated by the **UCSC Genome Bioinformatics Group** and external collaborators display **gene predictions, mRNA and expressed sequence tag alignments, simple nucleotide polymorphisms, expression and regulatory data, phenotype and variation data, and pairwise and multiple-species comparative genomics data**. All **information relevant** to a region is presented in **one window**, facilitating **biological analysis and interpretation**. The database tables **underlying the Genome Browser tracks** can be **viewed, downloaded, and manipulated** using another **Web-based application**, the UCSC Table Browser. Users can upload data as **custom annotation tracks** in both browsers for research or educational use.

The **vast size** of vertebrate genome data sets **presents challenges** in **efficient data storage and retrieval**. In addition, the **burgeoning number** of versions of a **particular genome** demands a process that can **rapidly integrate new data** and annotations into the database while **implementing creative solutions** for maintaining and **enhancing views of the data**. Through **software algorithmic refinements** and optimizations to **both the database and hardware**, the UCSC Genome Browser viewer maintains the **same interactive response time** on the **large Homo sapiens and Mus musculus genomes** that its predecessor had on the **much smaller Caenorhabditis elegans genome**.

Sequence and annotation data for each genome assembly are **stored in a MySQL relational database**, which is **quite efficient at retrieving data** from indexed files. The **database is loaded** in large batches and is used **primarily as a read-only database**. To improve performance, each of the Genome Browser web servers has a **copy of the database on its local disk**. UCSC **generates several annotations** based on **mRNA alignments**. The **mRNA and EST sequences** are extracted from GenBank, and are **aligned** against the genome **using the BLAST-like Alignment Tool (BLAT)**, a fast sequence alignment tool developed by Jim Kent. The **data is filtered based on percentage identity** and near best in genome to select only those alignments that **best match the sequence**. The **spliced EST annotation** is computed from the **filtered data by analyzing** the EST alignments for **evidence of splicing**.

2. ENSEMBL gene browser

The Ensembl project was initially launched in 1999 with the aim of **developing methodologies** for **automatic annotation** of (human) **genomic sequence** with genes and their **constituent transcripts**. Since that time, the project has broadened substantially in scope; the **Ensembl Genome Browser**, which came online in 2000, now includes **reference genomic sequence** and **annotation** for nearly **100 chordate organisms**. Ensembl is **rapidly incorporating** new data, including **whole clades** of new species' genomes and reference sequence for **multiple strains of existing species**, such as mouse. In addition, existing annotation is **regularly augmented** by the **inclusion of new data sets**. Ensembl's sister site, **Ensembl Genomes**, provides access to **nonvertebrate genomes** through dedicated portals for **Bacteria, Fungi, Plants, Metazoa, and Protists**.

Ensembl data, annotations, and analyses are updated **every 2–3 months**, alongside **software updates** to both the **public-facing website** and the **underlying databases**. A dedicated site is also maintained for the **GRCh37 reference human genome assembly**, which is annotated with **new data on a limited basis**; partial data from **ongoing genome annotation** can be accessed via the preview Pre! site.

Data from Ensembl can be **accessed at multiple scales**. Data can be accessed through the browser web pages and **via BioMart**, a web-based tool that allows **customized retrieval** of data from the **Ensembl databases**. However, data can also be **accessed programmatically** via our **Perl and REST APIs**. Files containing **genome-wide data** are available for all species **represented in Ensembl** via an **FTP site**; data from all releases of **Ensembl** can be **retrieved** from the **FTP site**, or from our databases via the **Perl APIs, in perpetuity**.

3. Genome Data Viewer

GDV is composed of an **embedded instance of SV** that displays **sequence and track data**, along with **additional page elements** that allow a user to **search within an entire genome** assembly and efficiently **narrow in on their chromosome, sequence, region, or gene of interest**. GDV replaced the NCBI Map Viewer, NCBI's **previous** tool for whole-genome display. Researchers using GDV **can go directly** to the NCBI BLAST service from the browser and **load BLAST results** as alignment tracks that can be viewed **side by side** with **gene annotation** and other data. Variation Viewer, a related browser associated with NCBI's variation resources, is **functionally similar to GDV** and also **incorporates an instance of SV** but is **configured with features specifically intended for analyzing human variation data**. GDV and Variation Viewer can both **display the same types of NCBI variation track data**.

The GDV **can be accessed** from its own **home page** and can also be **found via links** from other **NCBI resources, including gene, assembly, GEO, and dbGaP record pages**. GDV provides users a graphical gateway to data at the NCBI, especially **RefSeq** and **refSNP** annotation. Below, we highlight **some of the functions** of GDV and other instances of the NCBI SV and provide context for **GDV's features** with respect to the **broader collection** of publicly available **genome browsers**, including the **UCSC** and **Ensembl genome browsers, JBrowse, and IGV**. **GDV** was designed specifically to **support visualization and analysis** of the wide range of genomes and assemblies annotated at the NCBI. **RefSeq gene annotation** data tracks are shown by default in the **graphical view** for these assemblies. **NCBI refSNP data tracks** are also shown by default for **human assemblies**. **Gene and SNP tracks** are **automatically updated** in **GDV** and **SV** embedded instances upon **new releases** of the NCBI databases, so that **users of the NCBI graphical viewers** always have immediate access to the **latest versions of RefSeq and SNP annotation**.

GDV offers users the ability to **customize the displays of individual tracks**. Users can **hide or configure** tracks from the **track configuration panel** or by using the icons at the right end of each track. Different **public genome browsers provide conceptually similar**, but somewhat **distinct options**, for **visualizing gene, graphical, and alignment data**. In this section, we highlight **track data visualizations** in the GDV browser and other instances of the **SV graphical view** component that support various **analysis scenarios**.

REFERENCES:

1. Xiong, J. (2008). Genome Mapping, Assembly, and Comparison. Essential bioinformatics. Cambridge: Cambridge University Press. 243.
2. Baxevanis, Andreas D.; Petsko, Gregory A.; Stein, Lincoln D.; Stormo, Gary D. (2002). Current Protocols in Bioinformatics || The UCSC Genome Browser. , (), -. doi:10.1002/0471250953.bi0104s28
3. Karolchik, D. (2003). The UCSC Genome Browser Database. , 31(1), 51–54. doi:10.1093/nar/gkg129
4. Birney, E. (2004). An Overview of Ensembl. Genome Research, 14(5), 925–928. doi:10.1101/gr.1860604
5. Kollmar, Martin (2018). [Methods in Molecular Biology] Eukaryotic Genomic Databases Volume 1757 || The Ensembl Genome Browser: Strategies for Accessing Eukaryotic Genome Data. , 10.1007/978-1-4939-7737-6(Chapter 6), 115–139. doi:10.1007/978-1-4939-7737-6_6
6. Rangwala, S. H., Kuznetsov, A., Ananiev, V., Asztalos, A., Borodin, E., Evgeniev, V., Joukov, V., Lotov, V., Pannu, R., Rudnev, D., Shkeda, A., Weitz, E. M., & Schneider, V. A. (2020). Accessing NCBI data using the NCBI Sequence Viewer and Genome Data Viewer (GDV). Genome Research, gr.266932.120. <https://doi.org/10.1101/gr.266932.120>

WEBLEM 9a**UCSC Genome Browser**

(URL: <https://genome.ucsc.edu/>)

AIM:

To explore UCSC genome browser in order to understand the gene, its related studies & protein level information.

INTRODUCTION:

The **University of California Santa Cruz (UCSC) Genome Browser** (genome.ucsc.edu) is a popular **Web-based tool** for quickly displaying a requested **portion of a genome** at any scale, accompanied by a series of **aligned annotation “tracks”**. The annotations generated by the **UCSC Genome Bioinformatics Group** and external collaborators display **gene predictions, mRNA and expressed sequence tag alignments, simple nucleotide polymorphisms, expression and regulatory data, phenotype and variation data, and pairwise and multiple-species comparative genomics data**. All **information relevant** to a region is presented in **one window**, facilitating **biological analysis and interpretation**. The database tables **underlying the Genome Browser tracks** can be **viewed, downloaded, and manipulated** using another **Web-based application**, the UCSC Table Browser. Users can upload data as **custom annotation tracks** in both browsers for research or educational use.

METHODOLOGY:

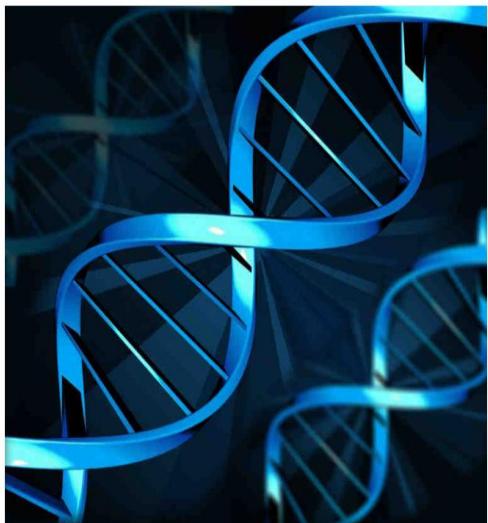
1. Open homepage for UCSC browser (URL: <https://genome.ucsc.edu/>)
2. Select genome browser.
3. Select human assembly (GRCh3/hg38)
4. Navigate results through gene name, SNP id, Ref_Seq, OMIM id, coordinates and cytological band.
5. Use tools for zooming tracks in and out, configuration by right click, drag and select and various option available at bottom of the page.
6. Observer and interpret the results.

OBSERVATION:

UNIVERSITY OF CALIFORNIA
SANTA CRUZ Genomics Institute 

Genome Browser

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](#)

Fig1. Homepage of UCSC browser

UNIVERSITY OF CALIFORNIA
SANTA CRUZ Genomics Institute 

Genome Browser Gateway

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

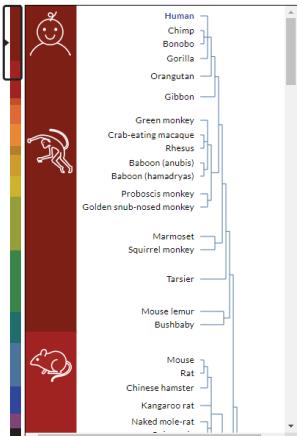
Browse/Select Species

POPULAR SPECIES



Enter species, common name or assembly ID
[Can't find a genome assembly?](#)

REPRESENTED SPECIES



Find Position

Human Assembly
Dec. 2013 (GRCh38/hg38)

Position/Search Term
Enter position, gene symbol or search terms
Current position: chrX:15,560,138-15,602,945

Human Genome Browser - hg38 assembly

UCSC Genome Browser assembly ID: hg38
Sequencing/Assembly provider ID: Genome Reference Consortium Human GRCh38.p13 (GCA_000001405.28)
Assembly date: Dec. 2013 initial release; Dec. 2017 patch release 13
Assembly accession: GCA_000001405.28
NCBI Genome ID: 51 (Homo sapiens (human))
NCBI Assembly ID: GCF_000001405.39 (GRCh38.p13, GCA_000001405.28)
BioProject ID: PRJNA31257

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 HTTP**
- **Using FTP**
- **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 GRCh37 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 numbers for 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 a glossary of assembly-related terms, see the [GRC Assembly Terminology page](#).


Homo sapiens
(Graphic courtesy of CBSE)

Fig2. Genome browser gateway

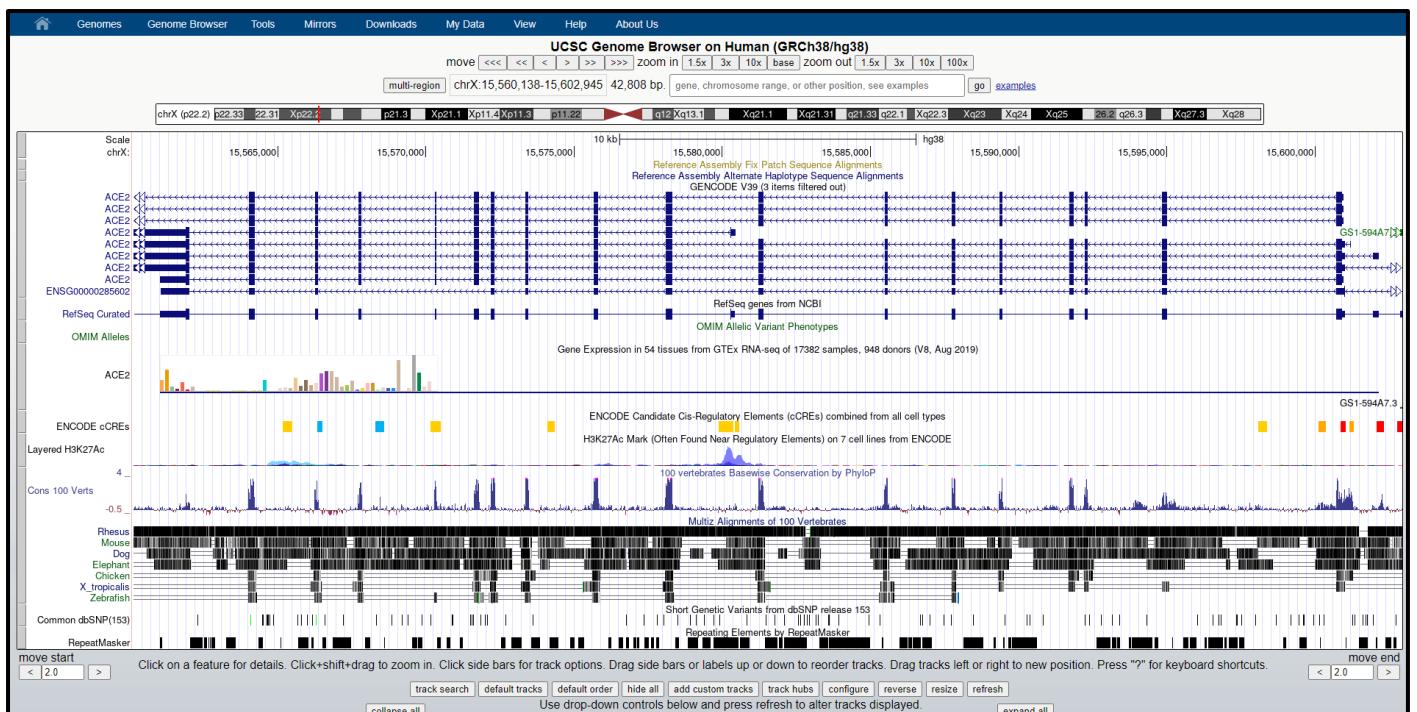


Fig3. UCSC genome browser on human (GRCh38/hg38)

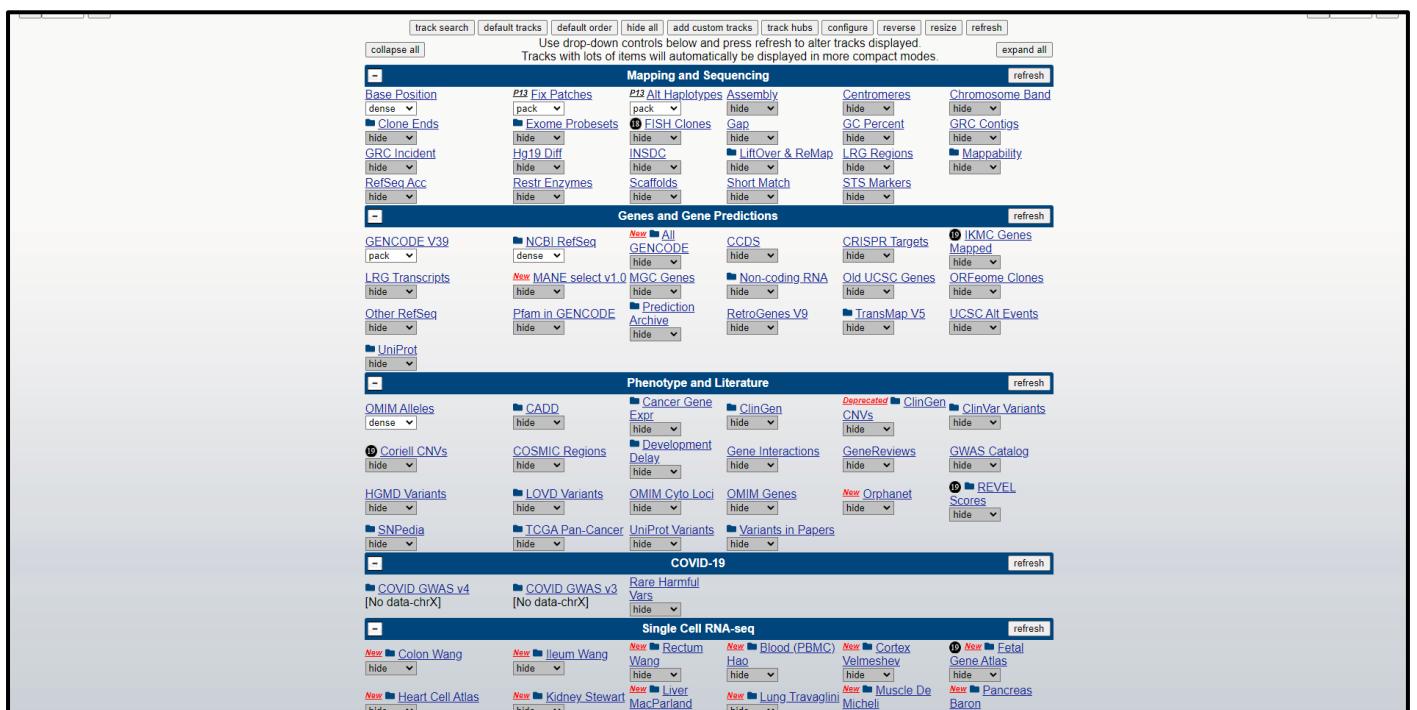


Fig4. Options for customization



Fig5. Results after customization



Fig6. Navigation by OMIM id: 115500

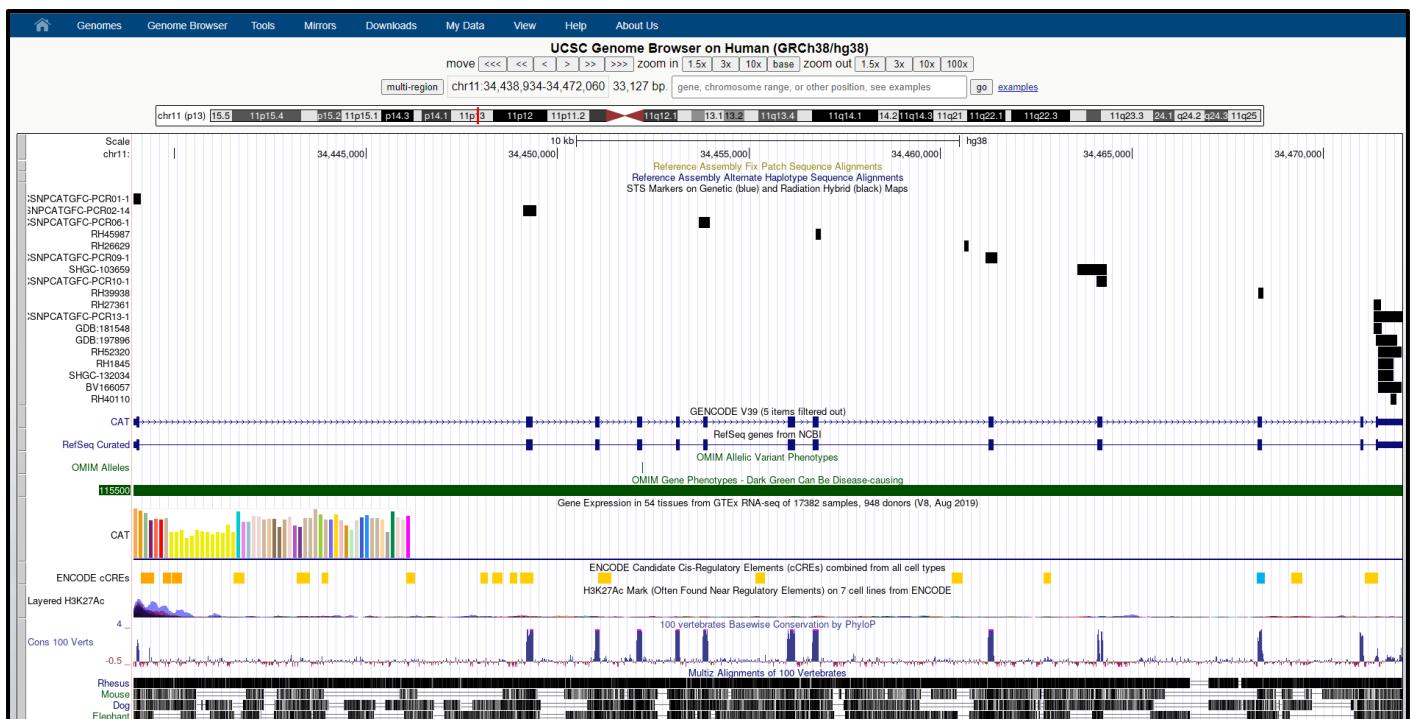


Fig7. Results for OMIM id: 115500

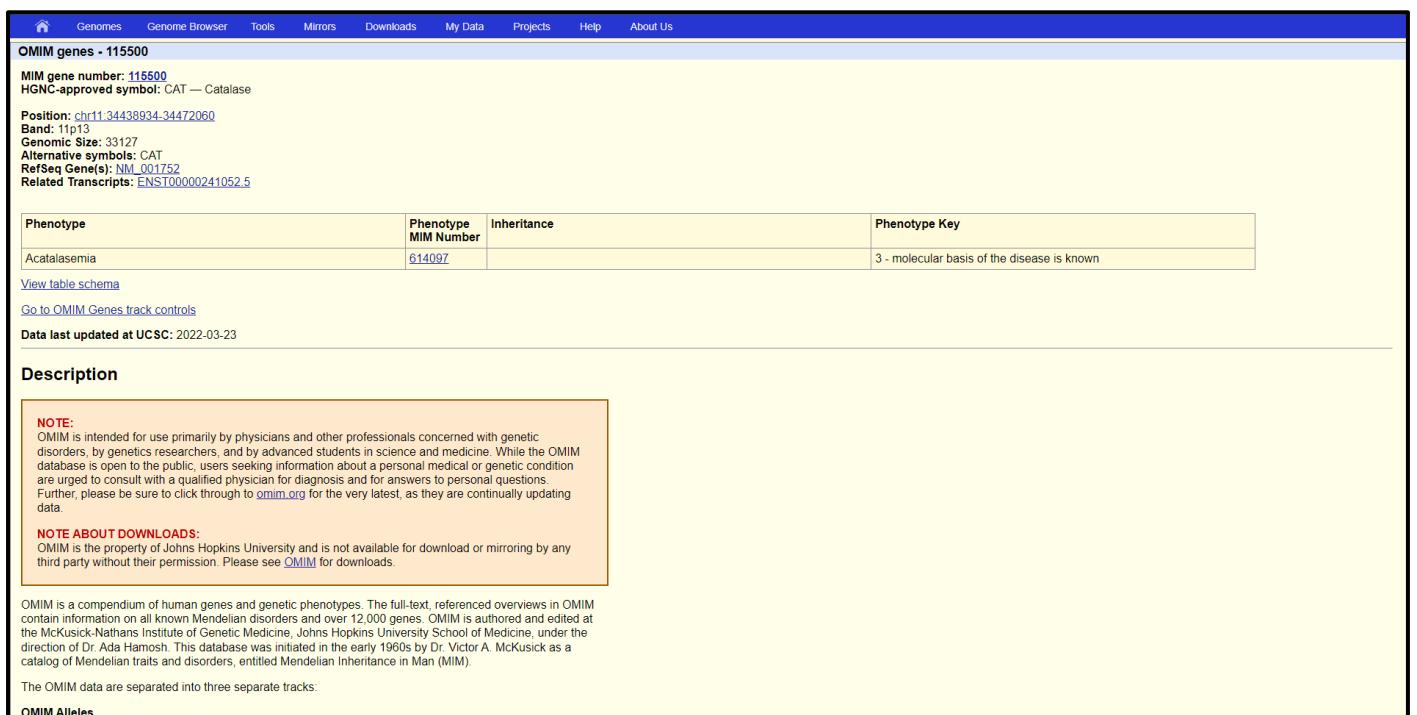


Fig8. Description for OMIM gene

[Home](#) [Genomes](#) [Genome Browser](#) [Tools](#) [Mirrors](#) [Downloads](#) [My Data](#) [Projects](#) [Help](#) [About Us](#)

Human Gene CAT (ENST00000241052.5) from GENCODE V39

Description: Homo sapiens catalase (CAT), mRNA. (from RefSeq NM_001752)

RefSeq Summary (NM_001752): This gene encodes catalase, a key antioxidant enzyme in the bodies defense against oxidative stress. Catalase is a heme enzyme that is present in the peroxisome of nearly all aerobic cells. Catalase converts the reactive oxygen species hydrogen peroxide to water and oxygen and thereby mitigates the toxic effects of hydrogen peroxide. Oxidative stress is hypothesized to play a role in the development of many chronic or late-onset diseases such as diabetes, asthma, Alzheimer's disease, systemic lupus erythematosus, rheumatoid arthritis, and cancers. Polymorphisms in this gene have been associated with decreases in catalase activity but, to date, acatalasemia is the only disease known to be caused by this gene. [provided by RefSeq, Oct 2009].

Gene ID: ENST00000241052.5
Gene Name: CAT
Gene Symbol: ENST00000241052.5
Gene Type: Transcript (including UTRs)
Position: hg38 chr11:34,438,934-34,472,060 **Size:** 33,127 **Total Exon Count:** 13 **Strand:** +
Coding Region:
Position: hg38 chr11:34,439,014-34,471,433 **Size:** 32,420 **Coding Exon Count:** 13

Page Index [Sequence and Links](#) [UniProtKB](#) [Comments](#) [MalaCards](#) [CTD](#) [RNA-Seq Expression](#)
[Microarray Expression](#) [RNA Structure](#) [Protein Structure](#) [Other Species](#) [GO Annotations](#) [mRNA Descriptions](#)
[Pathways](#) [Other Names](#) [Methods](#)

Data last updated at UCSC: 2022-01-18 01:30:34

Sequence and Links to Tools and Databases

Genomic Sequence (chr11:34,438,934-34,472,060)	mRNA (may differ from genome)	Protein (527 aa)
Gene Sorter	Genome Browser	Other Species FASTA
Ensembl	Entrez Gene	ExonPrimer
GeneCards	HPRD	Lynx
OMIM	Reactome	UniProtKB
PubMed		Wikipedia

Comments and Description Text from UniProtKB

ID: CATA_HUMAN
DESCRIPTION: ReChain: Full-Catalase; EC=1.11.1.6.
FUNCTION: Occurs in almost all aerobically respiring organisms and serves to protect cells from the toxic effects of hydrogen peroxide. Promotes growth of cells including T-cells, B-cells, myeloid leukemia cells, melanoma cells, mastocytoma cells and normal and transformed fibroblast cells.
CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
COFACTOR: Heme group
COFACTOR: NADH
SUBUNIT: Homotetramer
SUBCELLULAR LOCATION: Peroxisome
PTM: The N-terminus is blocked.
DISEASE: Defects in CAT are the cause of acatalasemia (ACATLAS) [MIM:614097]. A metabolic disorder characterized by absence of catalase activity in red cells and is often associated with ulcerating oral lesions.
SIMILARITY: Belongs to the catalase family.
WEB RESOURCE: Name=Wikipedia, Note=Catalase entry, URL="http://en.wikipedia.org/wiki/Catalase".
WEB RESOURCE: Name=SeattleSNPs, URL="http://pga.gs.washington.edu/data/cat/".

MalaCards Disease Associations

MalaCards Gene Search: [CAT](#)

Fig9. Related transcripts information

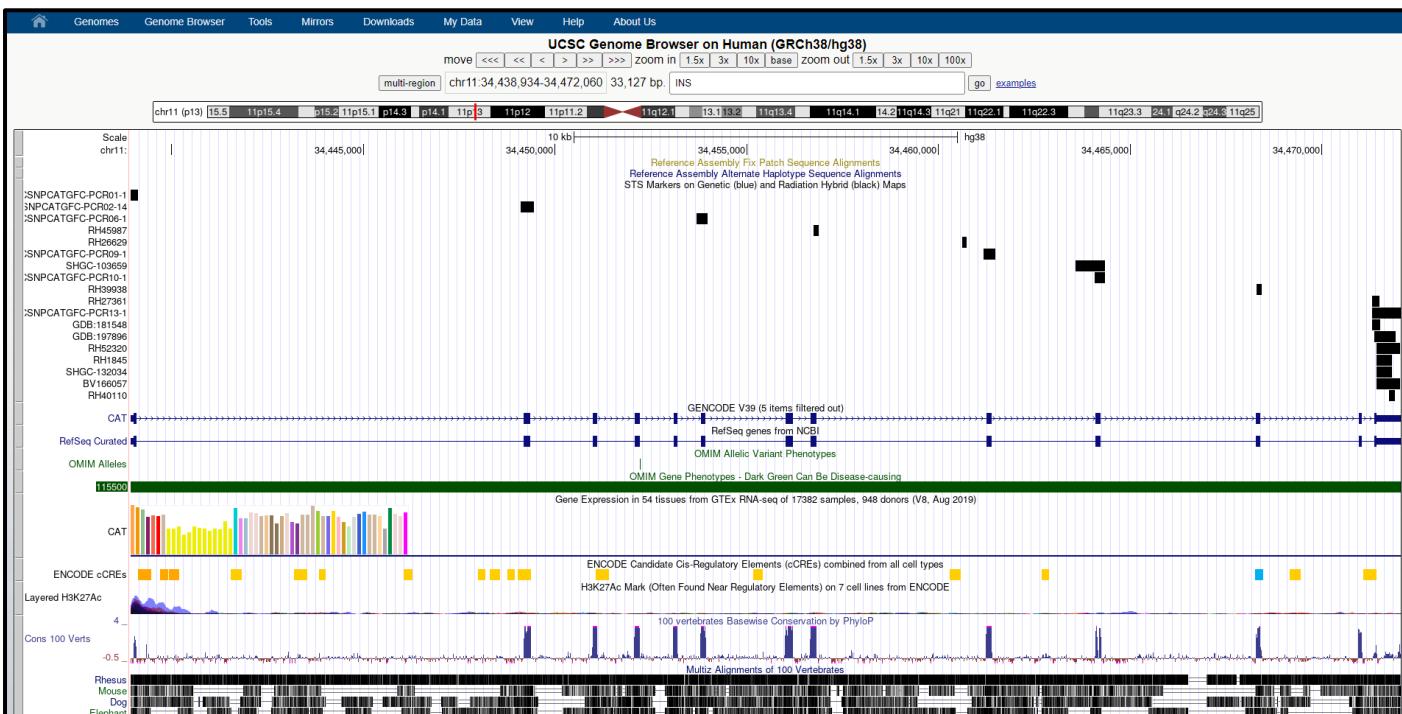


Fig10. Navigation by gene name: INS

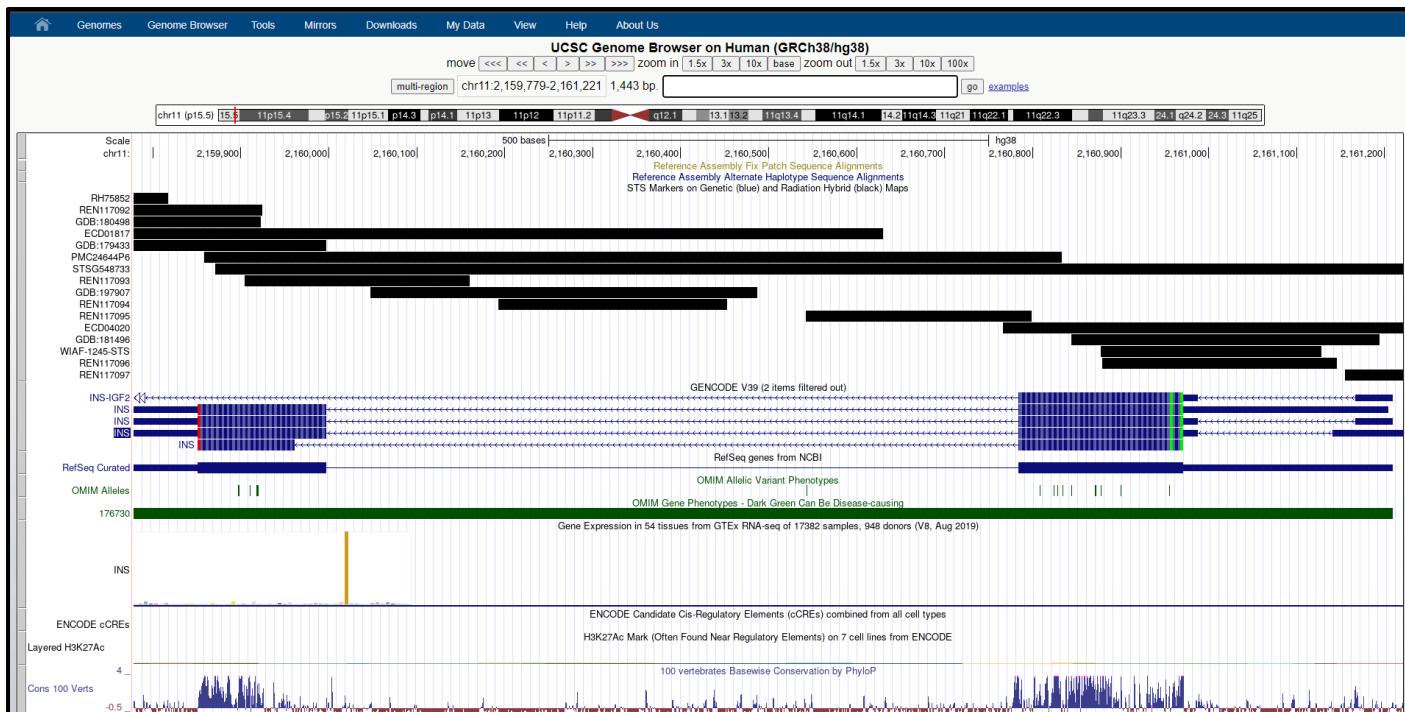


Fig11. Result for INS gene



Fig12. Option to configure tracks

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

Configure Image

submit
 image width: 1883 pixels
 label area width: 20 characters
 text size: 12
 font: Helvetica
 style: Normal

Display chromosome ideogram above main graphic
 Show light blue vertical guidelines, or light red vertical window separators in multi-region view
 Display labels to the left of items in tracks
 Display description above each track
 Show track controls under main graphic
 Next/previous item navigation
 Next/previous exon navigation
 Show exon numbers
 Enable highlight with drag-and-select (if unchecked, drag-and-select always zooms to selection)

Configure Tracks on UCSC Genome Browser: Human Dec. 2013 (GRCh38/hg38)

Tracks [track search] [hide all] [show all] [default] Groups [collapse all] [expand all]

Control track and group visibility more selectively below.

Mapping and Sequencing

Base Position	dense	Chromosome position in bases. (Clicks here zoom in 3x)
PtP Fix Patches	pack	Reference Assembly Fix Patch Sequence Alignments
PtP All Haplotypes	pack	Reference Assembly Alternate Haplotype Sequence Alignments
Assembly	hide	Assembly from Fragments
Centromeres	hide	Centromere Locations
Chromosome Band	hide	Chromosome Bands Localized by FISH Mapping Clones
Clone Ends	hide	Mapping of clone libraries end placements
Exome Probesets	hide	Exome Capture Probesets and Targeted Region
FISH Clones	hide	Clones Placed on Cytogenetic Map Using FISH
Gap	hide	Gap Locations
GC Percent	hide	GC Percent in 5-Base Windows
GRC Contigs	hide	Genome Reference Consortium Contigs
GRC Incident	hide	GRC Incident Database
Hg19 Diff	hide	Contigs New to GRCh38/hg38, Not Carried Forward from GRCh37/hg19
INSDC	hide	Accession at INSDC - International Nucleotide Sequence Database Collaboration
LiftOver & ReMap	hide	UCSC LiftOver and NCBI ReMap: Genome alignments to convert annotations to hg19

Fig13. Configuration settings

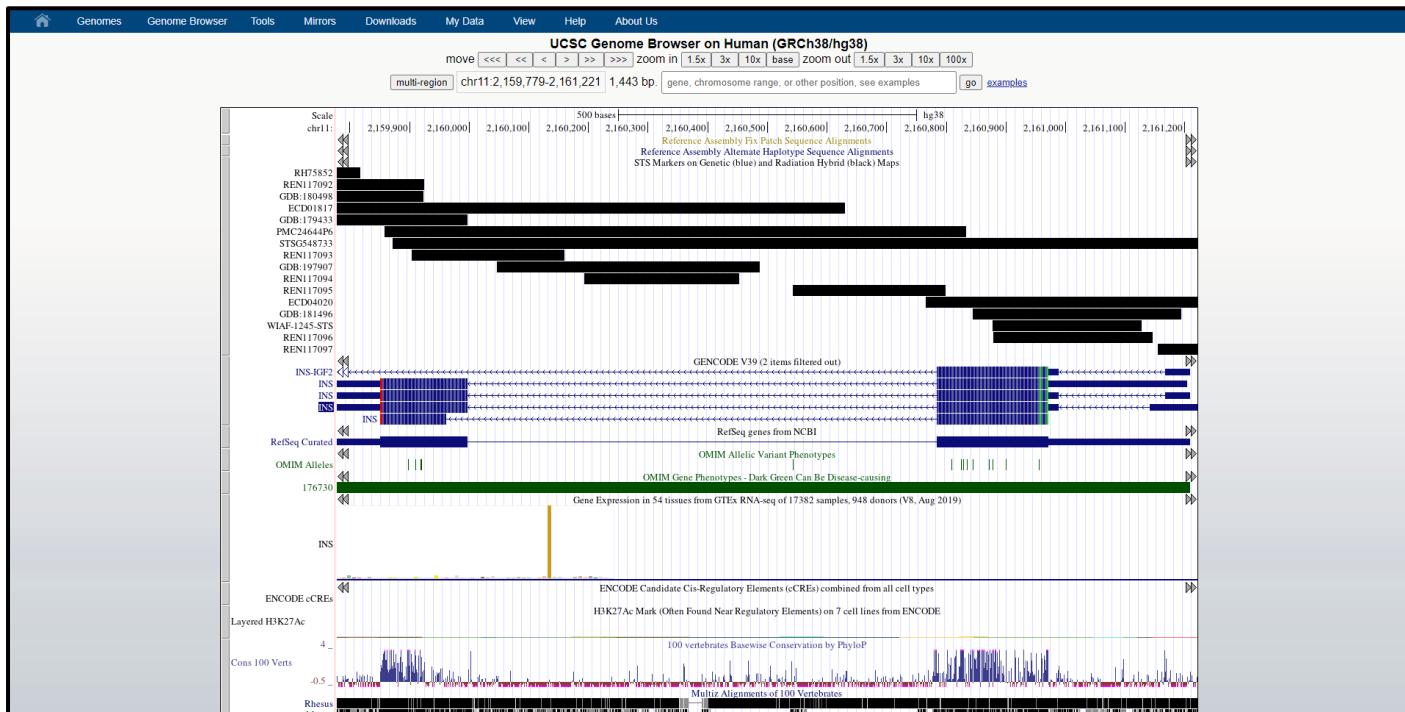


Fig14. Configured tracks



Fig15. Navigation by SNP id: rs10000



Fig16. Result for SNP id: rs10000

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

All Short Genetic Variants from dbSNP Release 153 (rs10000)

dbSNP: rs10000
Position: chr7:5973522-5973522
Band: 7p22.1
Genomic Size: 1
View DNA for this feature (hg38/Human)

Reference allele: A
Alternate allele: G
Allele frequency counts:

Allele	1000Genomes	GnomAD_exomes	ExAC	GnomAD
A	4694/5008 (0.937300)	121595/136684 (0.889607)	19069/24180 (0.788627)	11310/12746 (0.887337)
G	314/5008 (0.062700)	15089/136684 (0.110393)	5111/24180 (0.211373)	1436/12746 (0.112663)

Functional effects: synonymous_variant coding_sequence_variant nc_transcript_variant
ClinVar: RCV000030369.5 (benign), RCV000162401.1 (benign), RCV000174851.6 (benign), RCV000627740.1 (benign)
Submitted by: 1000GENOMES CGM KYOTO CNGFU CORRELAGEN CSHL EVA EVA_EXAC EVA_GENOME_DK EVA_SAMSUNG_MC GMI GNOMAD JILAB KHV_HUMAN_GENOMES OMUKHERJEE_ABDS SSMP SWEGEN TOPMED URBANLAB WEILL CORNELL_DGM
Publications in PubMed: PMID16619 , PMID10479499 , PMID15256438 , PMID16472587 , PMID20186688 , PMID20205264 , PMID25741868
Variation class/type: snv

Interesting or anomalous conditions noted by UCSC:

- Variant is in ClinVar
- Variant is in ClinVar with clinical significance of benign and/or likely benign.
- Variant is "common", i.e. has a Minor Allele Frequency of at least 1% in all projects reporting frequencies.
- This variant overlaps another variant with a different type/class.

[View table schema](#)

[Go to dbSNP 153 track controls](#)

Data last updated at UCSC: 2019-11-25 15:00:45

This track shows short genetic variants (up to approximately 50 base pairs) from dbSNP build 153: single-nucleotide variants (SNVs), small insertions, deletions, and complex deletion/insertions (indels), relative to the reference genome assembly. Most variants in dbSNP are rare, not true polymorphisms, and some variants are known to be pathogenic.

For hg38 (GRCh38), approximately 667 million distinct variants (RefSNP clusters with rs# ids) have been mapped to more than 702 million genomic locations including alternate haplotypes and fix patch sequences.

Fig17. Description for SNP id: rs10000



Fig18. Configuration option by right click

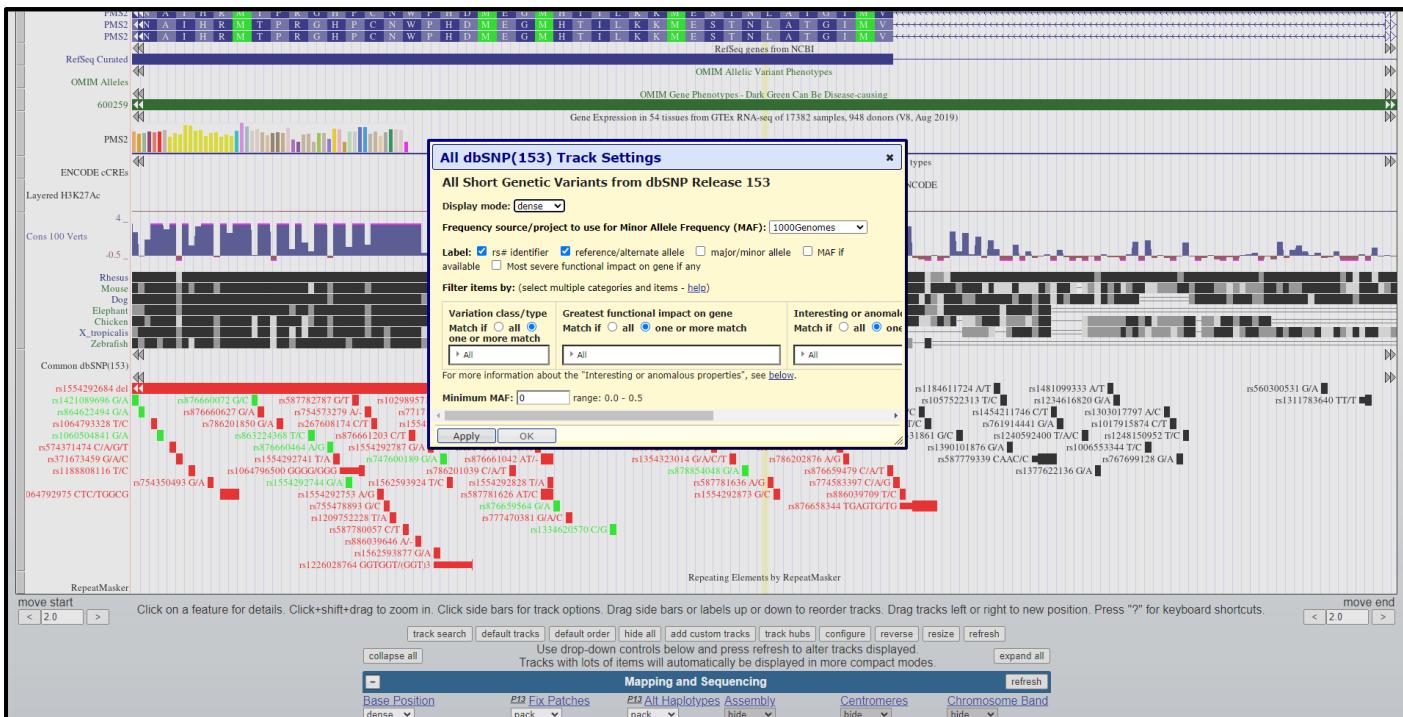


Fig19. Configuration applied

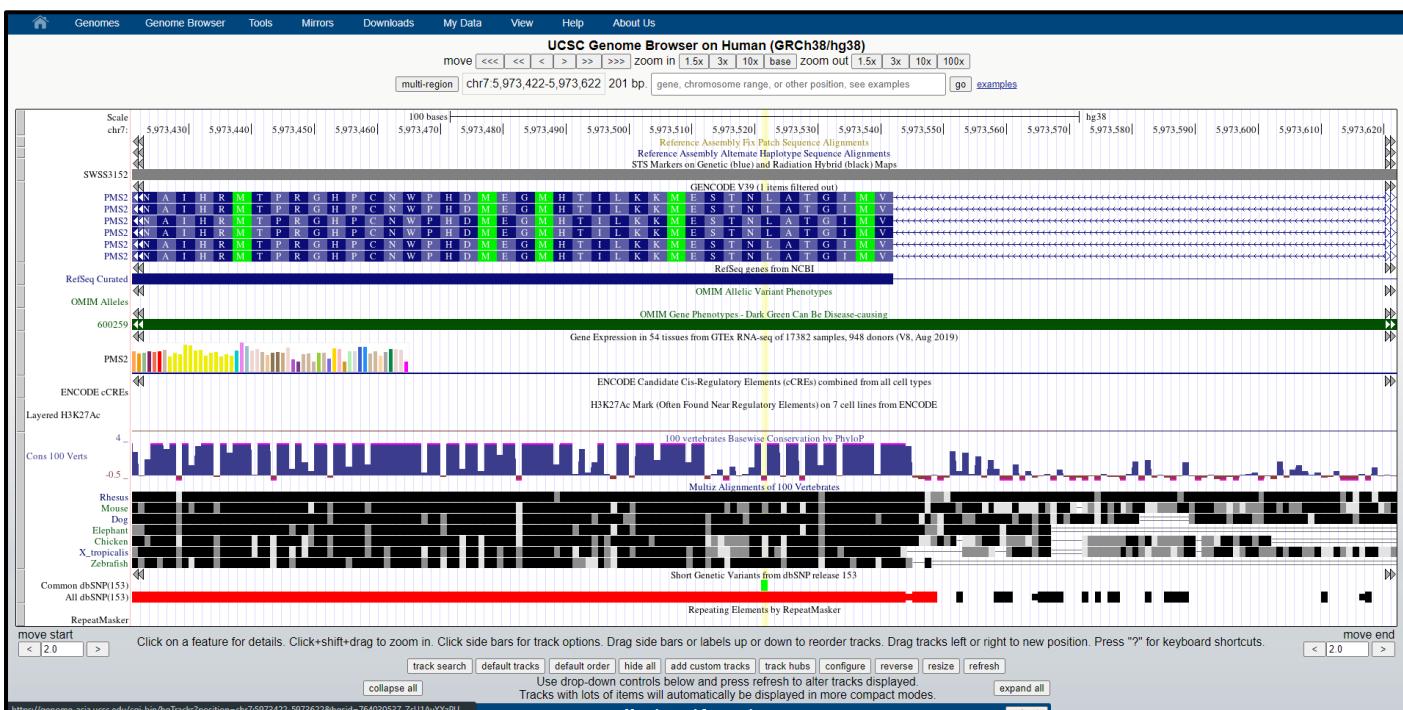


Fig20. Result after configuration



Fig21. Navigation by Ref_Seq: NM_014877.1 (HEL gene)

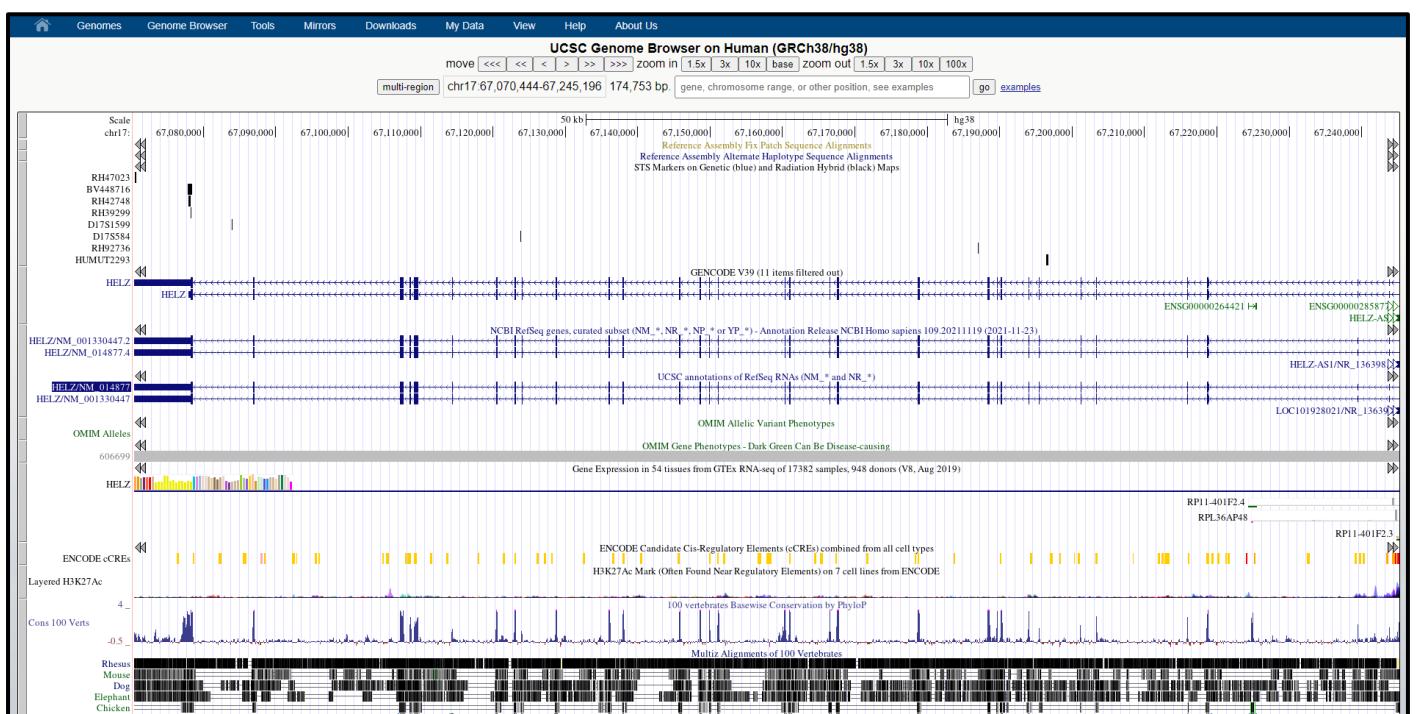


Fig22. Result for Ref Seq: NM_014977.1 (HELZ gene)

[Home](#) [Genomes](#) [Genome Browser](#) [Tools](#) [Mirrors](#) [Downloads](#) [My Data](#) [Projects](#) [Help](#) [About Us](#)

RefSeq Gene

RefSeq Gene HELZ

RefSeq: NM_014977.4 **Status:** Validated
Description: Homo sapiens helicase with zinc finger (HELZ), transcript variant 1, mRNA.
CCDS: CCDS42374.1
CDS: full length
OMIM: 606699
Entrez Gene: 9931
PubMed on Gene: HELZ
PubMed on Product: probable helicase with zinc finger domain isoform 1
GeneCards: HELZ
GeneView: HELZ

Summary of HELZ

HELZ is a member of the superfamily I class of RNA helicases. RNA helicases alter the conformation of RNA by unwinding double-stranded regions, thereby altering the biologic activity of the RNA molecule and regulating access to other proteins (Wagner et al., 1999 [PubMed 10471385]) [supplied by OMIM, Mar 2008]. Sequence Note: This RefSeq record was created from transcript and genomic sequence data because no single transcript was available for the full length of the gene. The extent of this transcript is supported by transcript alignments.

mRNA/Genomic Alignments

BRASER	SIZE	IDENTITY	CHROMOSOME	STRAND	START	END	QUERY	START	END	TOTAL
browser	13817	100.0%	17	-	67076444	67245196	NM_014977	1	13817	13817

[View details of parts of alignment within browser window.](#)

Position: chr17:67070444-67245196
Band: 17q24.2
Genomic Size: 174753
Strand:
Gene Symbol: HELZ
CDS Start: complete
CDS End: complete

Links to sequence:

- [Predicted Protein](#)
- [mRNA Sequence](#) (may be different from the genomic sequence)
- [Genomic Sequence](#) from assembly
- [CDS FASTA alignment](#) from multiple alignment

Fig23. Description for Gene HELZ

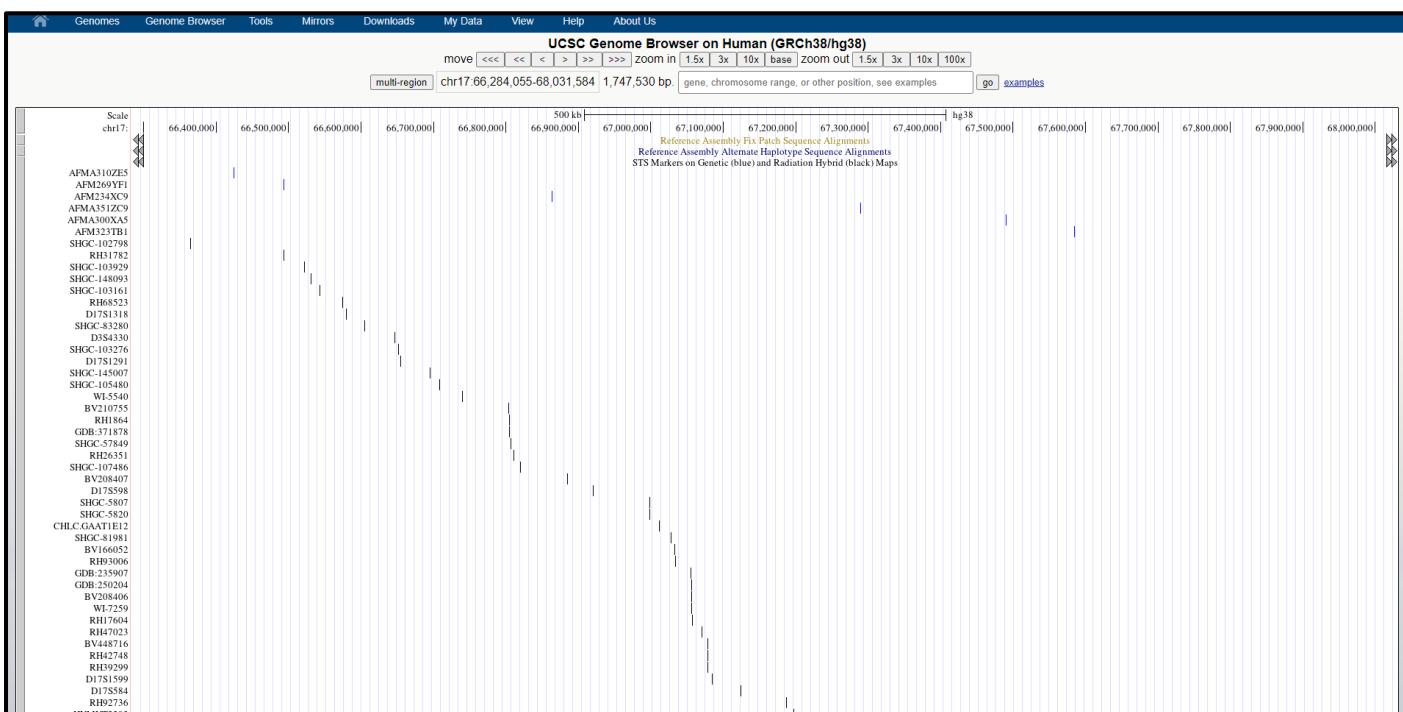


Fig24. Result after zooming 10x



Fig25. Navigation by cytological band: 11p15.1

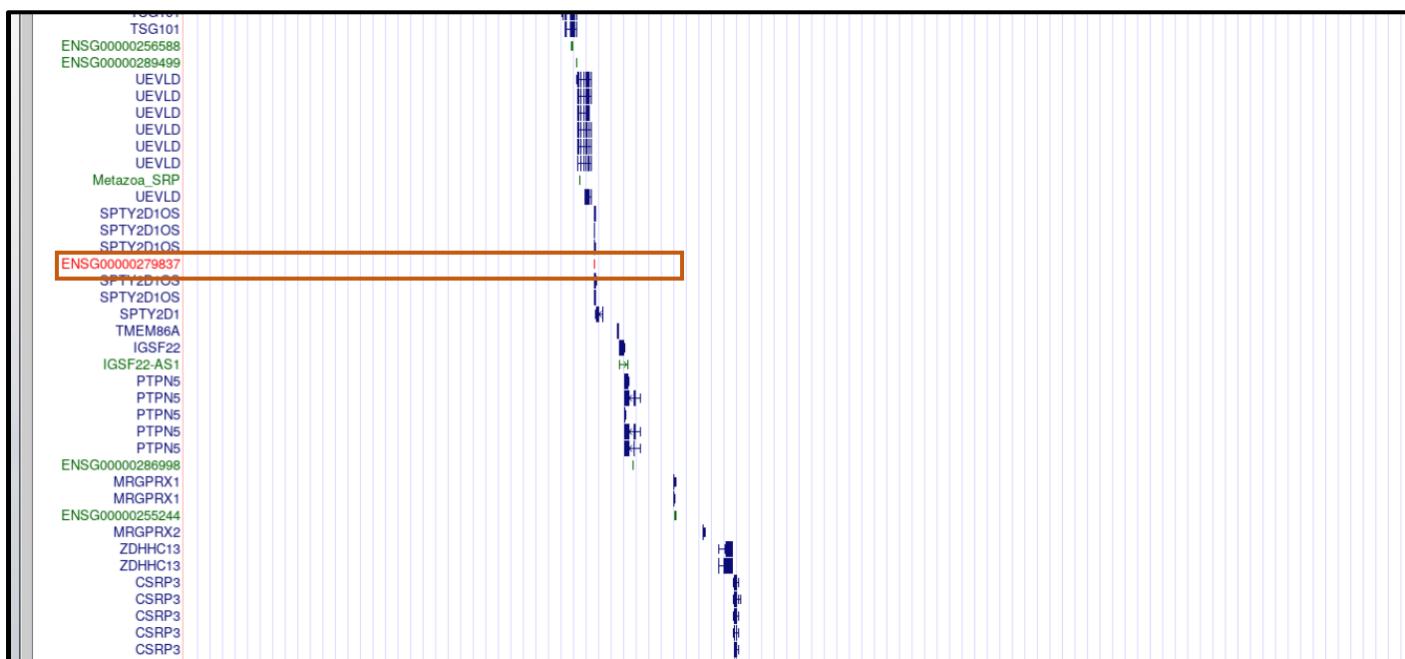


Fig26. Result for cytological band

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

Human Gene ENSG00000279837 (ENST00000623697.1) from GENCODE V39

Description: ENSG00000279837 (from geneSymbol)
 Gencode Transcript: ENST00000623697.1
 Gencode Gene: ENSG00000279837.1
 Transcript (Including UTRs)
 Position: hg38 chr11:18,601,882-18,602,649 Size: 768 Total Exon Count: 1 Strand: +

[Page Index](#) [Sequence and Links](#) [Other Species](#) [mRNA Descriptions](#) [Other Names](#) [Methods](#)

Data last updated at UCSC: 2022-01-17 08:30:34

Sequence and Links to Tools and Databases

Genomic Sequence (chr11:18,601,882-18,602,649)	mRNA (may differ from genome)	No protein			
Gene Sorter	Genome Browser	Other Species FASTA	Table Schema	Ensembl	ExonPrimer
Gencode	PubMed				

Orthologous Genes in Other Species

Orthologies between human, mouse, and rat are computed by taking the best BLASTP hit, and filtering out non-syntenic hits. For more distant species reciprocal-best BLASTP hits are used. Note that the absence of an ortholog in the table below may reflect incomplete annotations in the other species rather than a true absence of the orthologous gene.

Mouse	Rat	Zebrafish	D. melanogaster	C. elegans	S. cerevisiae
No ortholog	No ortholog	No ortholog	No ortholog	No ortholog	No ortholog

Fig27. Description for cytological band: 11p15.1

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

UCSC Genome Browser on Human (GRCh38/hg38)

move <<< << < > >> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x 100x

multi-region chr11:16,900,001-22,000,000 5,100,000 bp. chr11:16,900,001-22,000,000 go examples

chr11 (p15.1) 11p15.4 p15.1 14.3 14.1 11p13 11p12 p11.2 12.1 11q14.1 q13.4 q14.3 q21 q22.1 11q22.3 11q23.3 24.2 q25

Scale chr11: 17,500,000 18,000,000 18,500,000 19,000,000 19,500,000 20,000,000 20,500,000 hg38 21,000,000 21,500,000 Reference Assembly Fix Patch Sequence Alignments

chr11_ML143360v1_fix

Reference Assembly Alternate Haplotype Sequence Alignments Chromosome Bands Localized by FISH Mapping Clones 11p15.1

GENCODE V39 (337 items filtered out)

PLEKHA7 RPS13 SNORD14A PIK3C2A RNU6-593P NUCB2 ENSG00000285545 NORQLG1 KCNJ11 ENSG00000260196 KCNJ11 KCNJ11 KCNJ11 ABCC8 ABCC8

Fig28. Navigation by coordinates

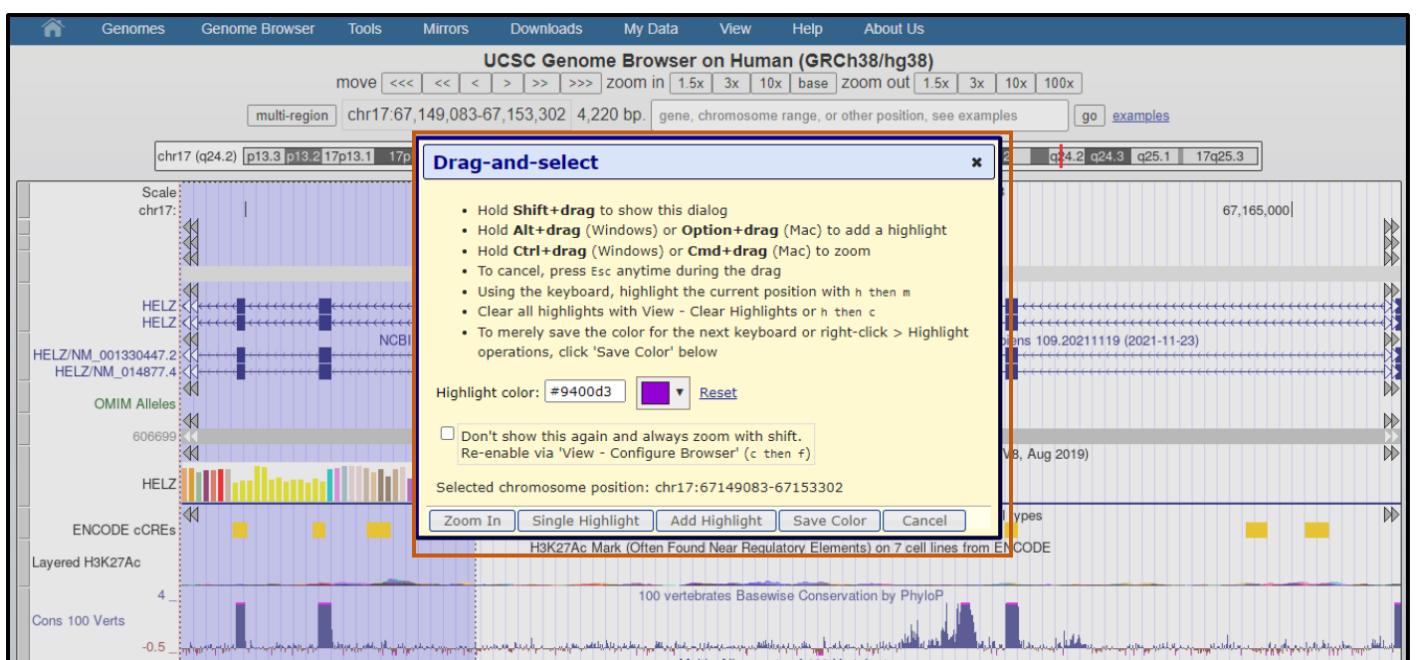


Fig29. Drag and select option for configuration

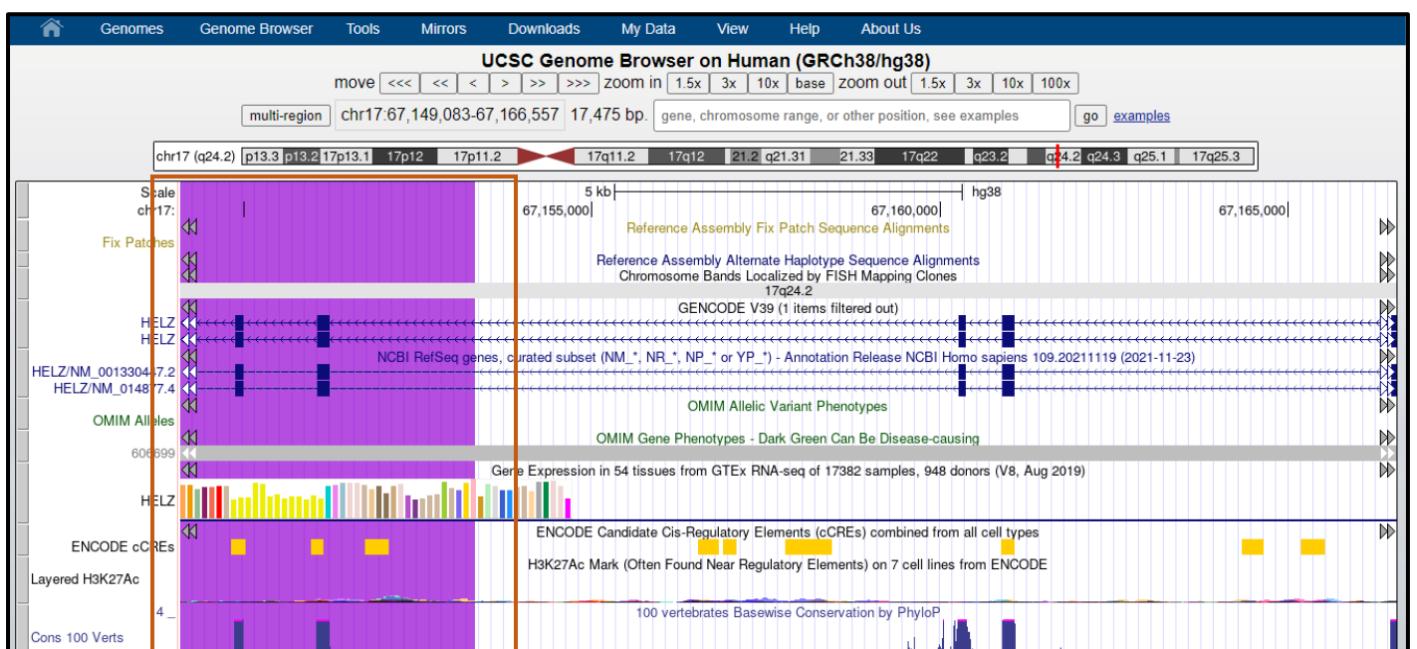


Fig30. Result after configuration

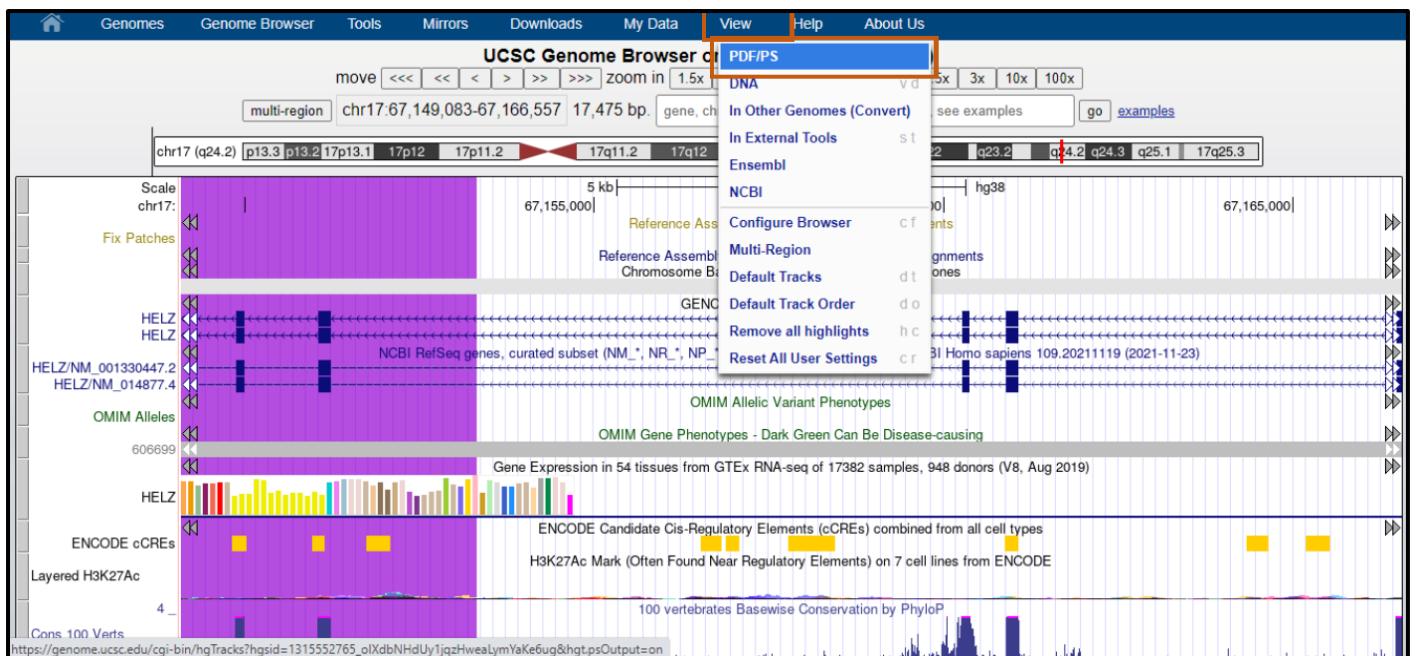


Fig32. Steps to export result as PDF

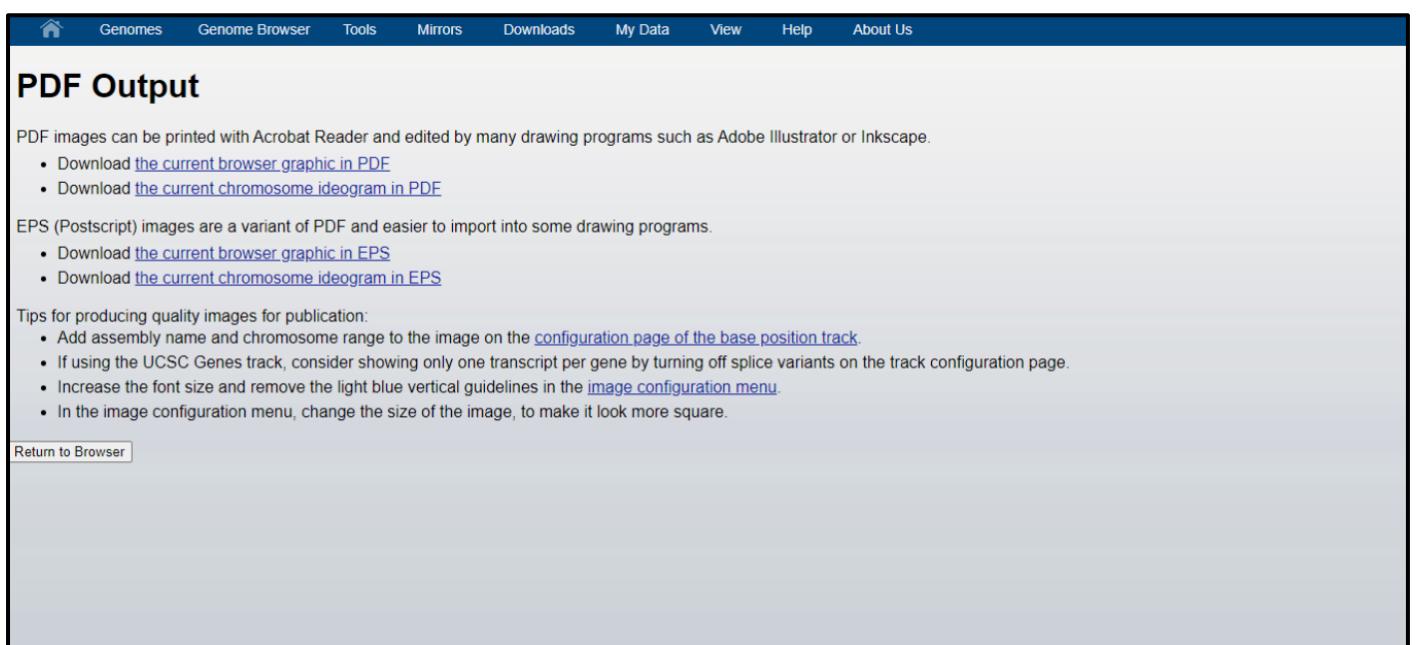


Fig33. Option to download PDF

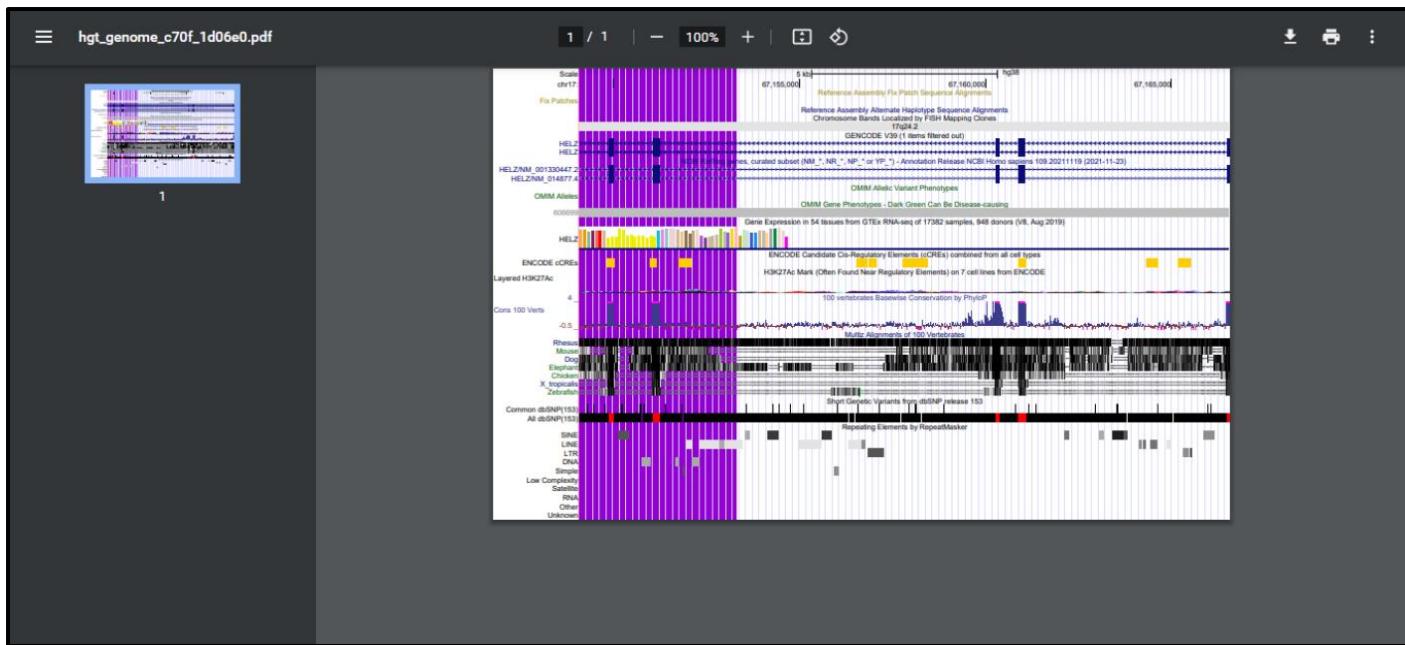


Fig34. Result in PDF format

RESULT:

UCSC genome browser was used for setting for GRCH/hg38 browser and search option used were:

- Navigation by gene name
- Navigation by SNP id
- Navigation by Ref_Seq: NM_014877.4
- Navigation by OMIM Id: 115500
- Navigation by cytological band: 11p15.1

Various options for configuration of tracks, zooming in and out the results, saving results in PDF format, etc were also used.

CONCLUSION:

UCSC genome browser can be used for gene predictions, mRNA and expressed sequence tag alignments, simple nucleotide polymorphisms, expression and regulatory data, phenotype and variation data, and pairwise and multiple-species comparative genomics data. All information relevant to a region is presented in one window, facilitating biological analysis and interpretation. It also provides various tools for configuration of tracks and refining the results. Options are available for zooming in and out the results and downloading the results in PDF format.

REFERENCES:

1. Baxevanis, Andreas D.; Petsko, Gregory A.; Stein, Lincoln D.; Stormo, Gary D. (2002). Current Protocols in Bioinformatics || The UCSC Genome Browser. , (), -. doi:10.1002/0471250953.bi0104s28
2. UCSC Genome Browser Home. (2019). Ucsc.edu. Retrieved March 28, 2022, from <https://genome.ucsc.edu/>
3. UCSC Genome Browser Gateway. (2018). Ucsc.edu. Retrieved March 28, 2022, from <https://genome.ucsc.edu/cgi-bin/hgGateway>
4. Human hg38 chr11%3A34438934%2D34472060 UCSC Genome Browser v428. (n.d.).Genome.ucsc.edu. Retrieved March 28, 2022, from https://genome.ucsc.edu/cgi-bin/hgTracks?db=hg38&lastVirtModeType=default&lastVirtModeExtraState=&virtModeType=default&virtMode=0&nonVirtPosition=&position=chr11%3A34438934%2D34472060&hgsid=1315552765_oIXdbNHdUy1jqzHweaLymYaKe6ug

5. **Navigation by OMIM id:** OMIM genes - 115500. (n.d.). Genome.ucsc.edu. Retrieved March 28, 2022, from https://genome.ucsc.edu/cgi-bin/hgc?hgsid=1315629613_43o3KD070AuOsLgYd19FgktiKt2H&db=hg38&c=chr11&l=34438933&r=34472060&o=34438933&t=34472060&g=omimGene2&i=115500
6. **Navigation by SNP id:** All Short Genetic Variants from dbSNP Release 153 (rs10000). (n.d.). Genome.ucsc.edu. Retrieved March 28, 2022, from https://genome.ucsc.edu/cgi-bin/hgc?hgsid=1315552765_oIXdbNHdUy1jqzHweaLymYaKe6ug&db=hg38&c=chr7&l=5973421&r=5973622&o=5973521&t=5973522&g=dbSnp153&i=rs10000
7. **Navigation by Ref_Seq:** NCBI RefSeq genes, curated subset (NM_*, NR_*, NP_* or YP_*) - NM_014877.4. (n.d.). Genome.ucsc.edu. Retrieved March 28, 2022, from https://genome.ucsc.edu/cgibin/hgc?hgsid=1315552765_oIXdbNHdUy1jqzHweaLymYaKe6ug&db=hg38&c=chr17&l=67070443&r=67245196&o=67070443&t=67245196&g=ncbiRefSeqCurated&i=NM_014877.4
8. **Navigation by Cytological band:** Human Gene ENSG00000279837 (ENST00000623697.1) from GENCODE V39. (n.d.). Genome.ucsc.edu. Retrieved March 28, 2022, from https://genome.ucsc.edu/cgibin/hgGene?hgg_gene=ENST00000623697.1&hgg_chrom=chr11&hgg_start=18601881&hgg_end=18602649&hgg_type=knownGene&db=hg38

WEBLEM 9b

Ensemble Genome Browser

(URL: <https://asia.ensemble.org/index.html>)

AIM:

To explore Ensembl genome browser in order to gather information for annotated genes/genome/protein/transcript etc.

INTRODUCTION:

The Ensembl project was initially launched in 1999 with the aim of developing methodologies for automatic annotation of (human) genomic sequence with genes and their constituent transcripts. Since that time, the project has broadened substantially in scope; the Ensembl Genome Browser, which came online in 2000, now includes reference genomic sequence and annotation for nearly 100 chordate organisms. Ensembl is rapidly incorporating new data, including whole clades of new species' genomes and reference sequence for multiple strains of existing species, such as mouse. In addition, existing annotation is regularly augmented by the inclusion of new data sets. Ensembl's sister site, Ensembl Genomes, provides access to nonvertebrate genomes through dedicated portals for Bacteria, Fungi, Plants, Metazoa, and Protists.

METHODOLOGY:

1. Open homepage for Ensembl genome browser. (URL: <https://asia.ensembl.org/index.html>)
 2. Select human (GRCH38.p13) genome assembly
 3. Search for *helad1* gene.
 4. Observer the results.
 5. Use the configuration tools for the tracks.
 6. Interpret the results.

OBSERVATION:

Ensembl MSB

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

Search all species... Login/Register

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

Pig breeds
Pig reference genome and 12 additional breeds

[View full list of all species](#)

Favourite genomes

 Human	GRCh38.p13
<small>Still using GRCh37??</small>	
 Mouse	GRCh39
 Zebrafish	GRCh31

Ensembl Rapid Release

New assemblies with gene and protein annotation every two weeks.

Note: species that already exist on this site will continue to be updated with the full range of annotations.

Rapid Release news

The Ensembl Rapid Release website provides annotation for recently produced, publicly available vertebrate and non-vertebrate genomes from biodiversity initiatives such as Darwin Tree of Life, the Vertebrate Genomes Project and the Earth BioGenome Project.

Rapid Release news

Other news from our blog

- 18 Mar 2022: [Disruption to Ensembl services](#)
- 22 Feb 2022: [Refining our public databases for Ensembl releases 24 to 47](#)
- 11 Feb 2022: [Free Ensembl Browser and REST API virtual workshops in March](#)

Compare genes across species 	Find SNPs and other variants for my gene 	Gene expression in different tissues 	Retrieve gene sequence 	Find a Data Display 	Use my own data in Ensembl 
--	--	--	--	---	--

EMBL-EBI  Ensembl creates, integrates and distributes reference datasets and analysis tools that enable genomic. We are based at [EMBL-EBI](#) and our software and data are freely available. Our [acknowledgements](#) page includes a list of current and previous funding bodies. How to [cite Ensembl](#) in your own publications.

This website requires cookies, and the limited processing of your personal data in order to function. By using the site you are agreeing to this as outlined in our [Privacy Policy](#) and [Terms of Use](#)

I Agree

elixir Core Data Resources

Fig1. Homepage for Ensembl

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

Human (GRCh38.p13) ▾

Search Human (Homo sapiens)

Search all categories ▾ Search... Go

e.g. BRCA2 or 17:63992802-64038237 or rs699 or osteoarthritis

Genome assembly: GRCh38.p13 (GCA_000001405.28)

- 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 microarray annotations.

- More about the Ensembl regulatory build and microarray annotation
- Experimental data sources
- Download all regulatory features (GFF)

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
- Download FASTA files for genes, cDNAs, ncRNA, proteins
- Download GTF or GFF3 files for genes, cDNAs, ncRNA, proteins
- Update your old Ensembl IDs

Pax6 IINS
Pax6
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

VelP

Example phenotype

Example structural variant

Ensembl release 105 - Dec 2021 © EMBL-EBI

Permanent link - View in archive site

Fig2. Homepage for Human (GRCH28.p13) genome assembly

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

Human (GRCh38.p13) ▾

Search Human (Homo sapiens)

Search all categories ▾ Search... Go

e.g. BRCA2 or 17:63992802-64038237 or rs699 or osteoarthritis

Genome assembly: GRCh38.p13 (GCA_000001405.28)

- 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 microarray annotations.

- More about the Ensembl regulatory build and microarray annotation
- Experimental data sources
- Download all regulatory features (GFF)

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
- Download FASTA files for genes, cDNAs, ncRNA, proteins
- Download GTF or GFF3 files for genes, cDNAs, ncRNA, proteins
- Update your old Ensembl IDs

Pax6 IINS
Pax6
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

VelP

Example phenotype

Example structural variant

Ensembl release 105 - Dec 2021 © EMBL-EBI

Permanent link - View in archive site

Fig3. Search for helad1 gene

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

New Search

Current selection: Only searching Human ▾ helad1

Only searching Human 1 results match helad1 when restricted to species: Human

Restrict category to: Gene 1

NAV2 (Human Gene) ENSG00000166833 11:19350724-20121601.1

Neuron navigator 2 [Source:HGNC Symbol;Acc:HGNC_15997]. Variant table • Phenotypes • Location • External Refs • Regulation • Orthologues • Gene tree

Per page: 10 25 50 100

Layout: Standard Table

Tip: Help and Documentation can be searched from the homepage! Just type in a term you want to know more about, like non-synonymous SNP

About Us Get help Our sister sites Follow us

About us Using this website Ensembl Bacteria Blog

Contact us Adding custom tracks Ensembl Fungi Twitter

Citing Ensembl Downloading data Ensembl Plants Facebook

Privacy policy Video tutorials Ensembl Protists

Disclaimer Variant Effect Predictor Ensembl Metazoa (VEP)

Best gene match Human Gene Human NAV2 11:19350724-20121601.1

Protein coding gene HGNC Symbol: Acc:HGNC_15997 neuron navigator 2

Fig4. Hit page for helad1 gene

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

Human (GRCh38.p13) ▾ Gene: NAV2

Location: 11:19,350,724-20,121,601

Gene-based displays

- Summary
- Splice variants
- Transcript comparison
- Gene alleles
- Sequences
- Secondary Structure
- Comparative Genomics
- Genomic alignments
- Gene tree
- Gene gain/loss tree
- Orthologues
- Paralogues
- Ensembl protein families

Ontologies

- GO: Biological process
- GO: Molecular function
- GO: Cellular component

Phenotypes

- Genetic Variation
- Variant table
- Variant image
- Structural variants
- Gene expression
- Pathways
- Regulation
- External references
- Supporting evidence
- ID History
- Gene history

Configure this page

Custom tracks

Export data

Share this page

Bookmark this page

Go to Region in Detail for more tracks and navigation options (e.g. zooming)

NAV2 (HGNC Symbol)

This gene is a member of the Human CCDS set: CCDS44552.1, CCDS53612.1, CCDS58126.1, CCDS7850.1, CCDS7851.2

This gene has proteins that correspond to the following UniProtKB identifiers: Q8V1L1

This Ensembl/Gencode gene contains transcript(s) for which we have selected identical RefSeq transcript(s). If there are other RefSeq transcripts available they will be in the External references table

ENSG00000166833.23

There is no ungapped mapping of this gene onto the GRCh37 assembly. View this locus in the GRCh37 archive: ENSG00000166833

Protein coding

Annotation for this gene includes both automatic annotation from Ensembl and Havana manual curation, see article.

overlapping locus [Definitions]

Loading component

Ensembl release 105 - Dec 2021 © EMBL-EBI

Permanent link - View in archive site

About Us Get help Our sister sites Follow us

About us Using this website Ensembl Bacteria Blog

Fig5. Result for NAV2 gene

Location: 11:19,350,724-20,121,601 Gene: NAV2 Transcript: NAV2-201

- Transcript-based displays
- Summary
- Sequence
 - Exons
 - cDNA
 - Proteins
- Protein Information
 - Protein summary
 - Domains & features
 - Variants
 - PDB 3D protein model
 - AlphaFold predicted model
- Genetic Variation
 - Variant table
 - Allele coverage
 - Haplotypes
 - Population comparison
 - Comparison image
- External References
 - General Identifiers
 - Gene IDs
 - Protein IDs
 - Supporting evidence
 - ID History
 - Transcript history
 - Protein history

Configure this page

Custom tracks

Export data

Share this page

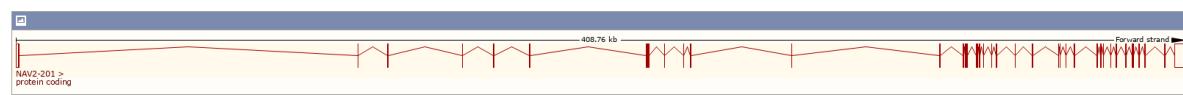
Bookmark this page

Transcript: ENST00000349880.9 NAV2-201

Description: neuron navigator 2 [Source HGNC Symbol Acc: HGNC_15997.6]
 Gene: NAV2
 Location: Chromosome 11: 19,12,837-20,121,601 forward strand.
 About this transcript: This transcript has 38 exons, is annotated with 68 domains and features, is associated with 110693 variant alleles and maps to 1154 oligo probes.
 Gene: This transcript is a product of gene ENSG00000166833.23 Hide transcript table

Transcript ID	Name	bp	Protein	Biotype	CCDS	UniProt Match	RefSeq Match	Flags
ENST00000349880.9	NAV2-201	11492	2429aa	Protein coding	CCDS7850.6	Q8VVL1-3.0	NM_145117.5	MANE Select v0.95 Ensembl Canonical GENCODE basic APPRIS P3 TSL1
ENST00000349885.6	NAV2-203	11501	2432aa	Protein coding	CCDS7851.0	Q8VVL1-2.0	-	GENCODE basic APPRIS ALT1 TSL5
ENST00000360655.8	NAV2-202	10667	2395aa	Protein coding	CCDS53612.0	Q8VVL1-4.0	-	GENCODE basic APPRIS ALT2 TSL1
ENST00000396087.7	NAV2-204	7882	2488aa	Protein coding	CCDS58126.0	Q8VVL1-1.0	-	GENCODE basic TSL5
ENST00000333917.5	NAV2-212	5084	1493aa	Protein coding	CCDS44552.0	Q8VVL1-5.0	-	GENCODE basic TSL2
ENST00000525322.5	NAV2-206	2716	815aa	Protein coding	-	E9PNV5.0	-	TSL2 CDS 3' incomplete
ENST00000530408.1	NAV2-210	556	160aa	Protein coding	-	E9PLU3.0	-	TSL5 CDS 3' incomplete
ENST000005650578.1	NAV2-215	256	63aa	Protein coding	-	A0A3B3ISY2.0	-	CDS 3' incomplete
ENST00000534229.1	NAV2-213	684	No protein	Processed transcript	-	-	-	TSL3
ENST00000534299.5	NAV2-214	668	No protein	Processed transcript	-	-	-	TSL5
ENST00000526575.1	NAV2-207	607	No protein	Processed transcript	-	-	-	TSL5
ENST00000526008.1	NAV2-208	570	No protein	Processed transcript	-	-	-	TSL4
ENST00000533746.1	NAV2-211	2289	No protein	Retained intron	-	-	-	TSL2
ENST00000525205.1	NAV2-205	627	No protein	Retained intron	-	-	-	TSL3
ENST00000526923.1	NAV2-209	563	No protein	Retained intron	-	-	-	TSL2

Summary



Statistics

Exons: 38, Coding exons: 38, Transcript length: 11,492 bps, Translation length: 2,429 residues

CCDS

This transcript is a member of the Human CCDS set: CCDS7850.6

Fig6. Result for NAV2-201

Summary

Name: NAV2 (HGNC Symbol)
 CCDS: This gene is a member of the Human CCDS set: CCDS44552.0, CCDS53612.1, CCDS58126.1, CCDS7850.1, CCDS7851.2

UniProtKB: This gene has proteins that correspond to the following UniProtKB identifiers: Q8VVL1

RefSeq: This Ensembl/Gencode gene contains transcript(s) for which we have selected identical RefSeq transcript(s). If there are other RefSeq transcripts available they will be in the External references table

Ensembl version: ENSG00000166833.23

Other assemblies: There is no unaligned mapping of this gene onto the GRCh37 assembly.

View this locus in the GRCh37 archive: ENSG00000166833.0

Gene type: Protein coding

Annotation method: Annotation for this gene includes both automatic annotation from Ensembl and Havana manual curation, see article

overlapping locus [Definitions]

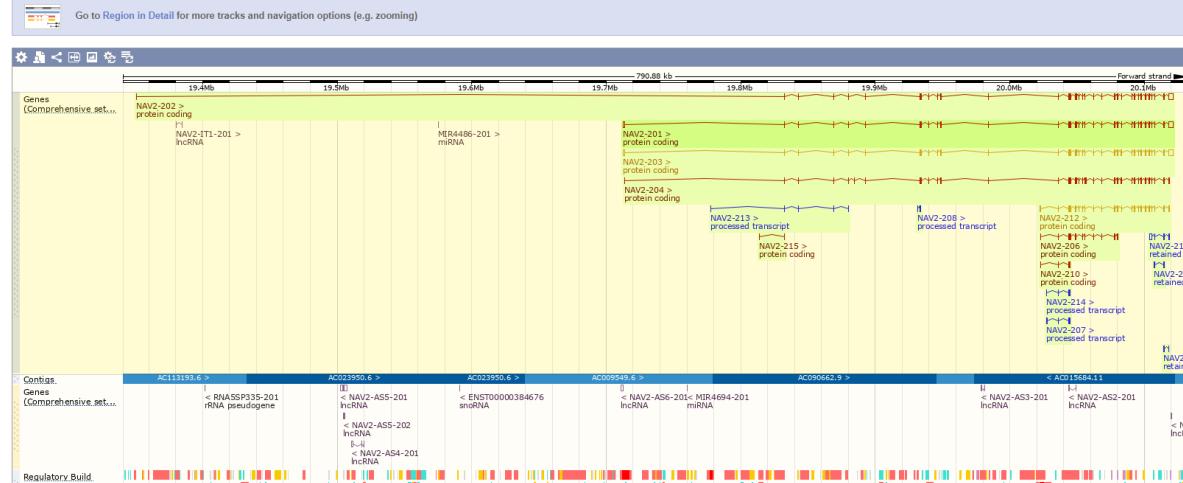


Fig7. Result for NAV2

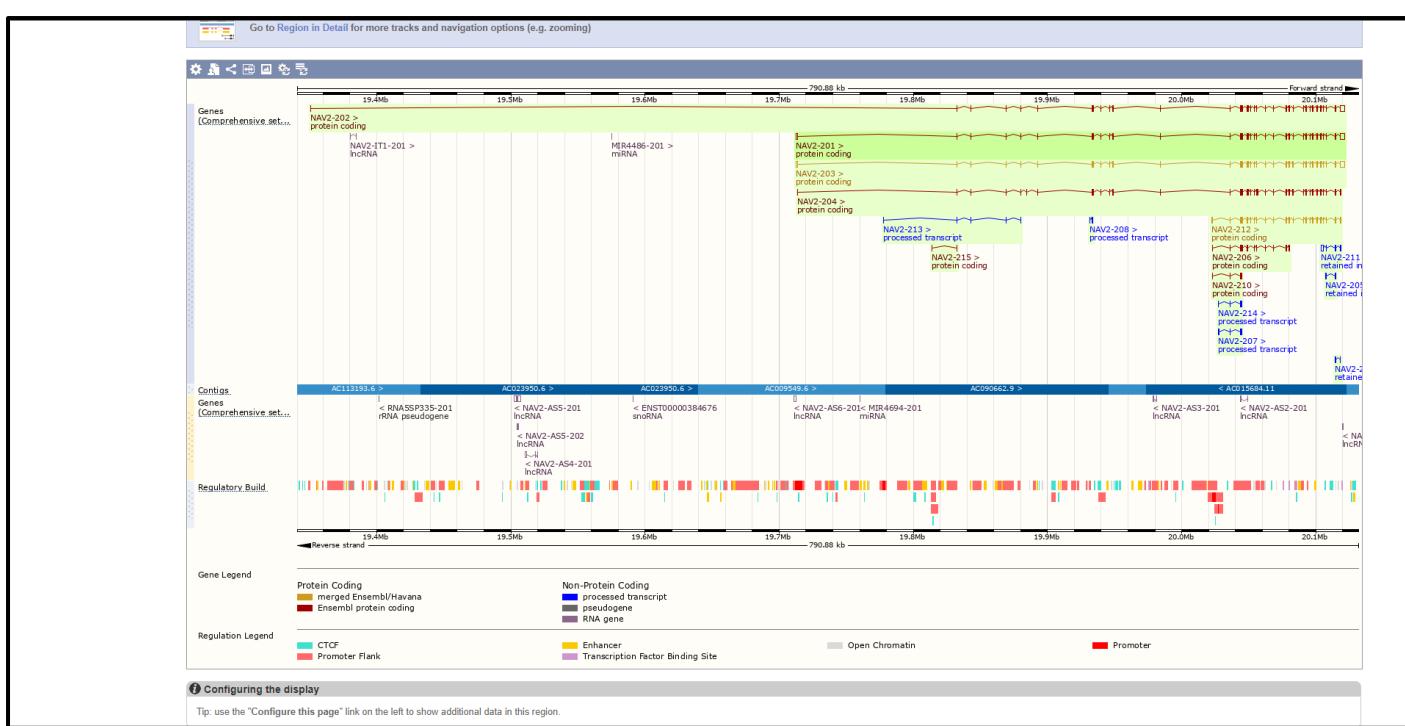


Fig6. Tracks information

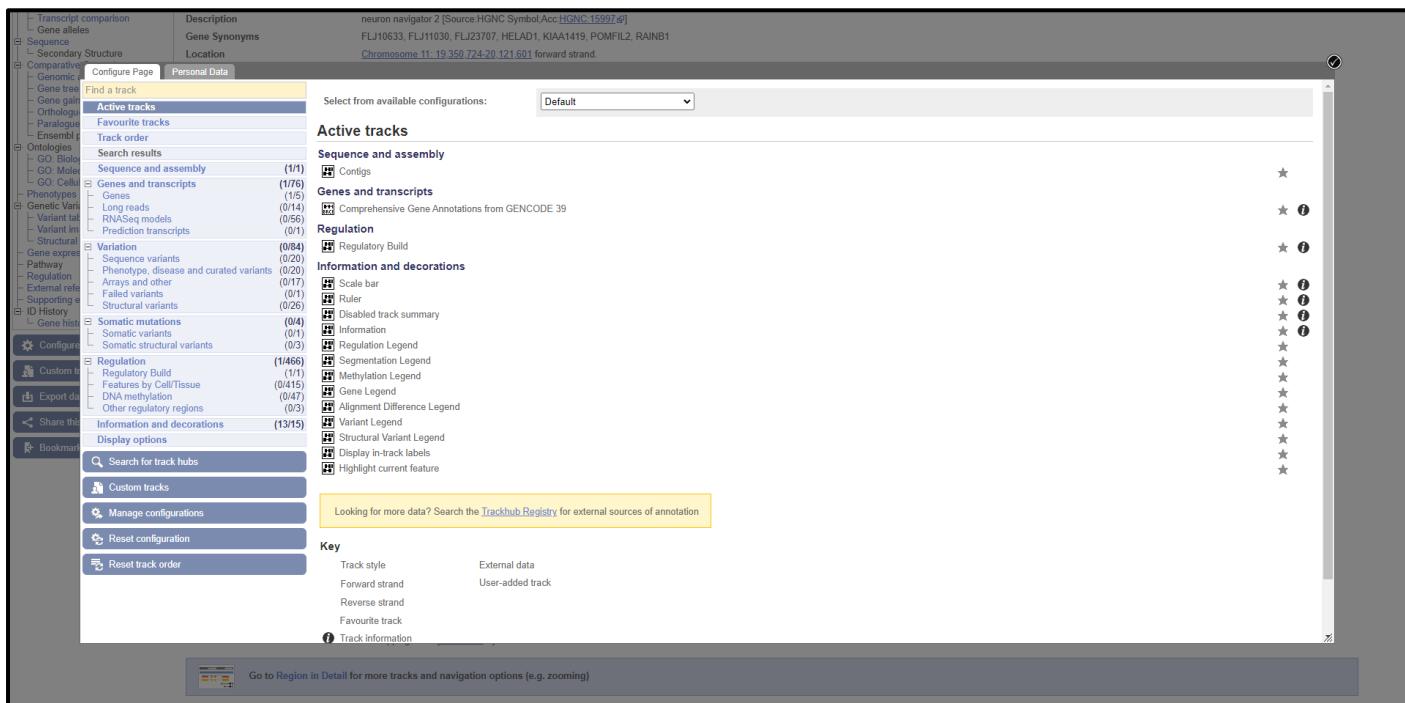


Fig7. Option for tracks configuration

Transcript comparison
Gene alleles
Sequence
Secondary Structure
Comparative
Genomic
- Gene tree
- Gene gall
- Gene gall
- Orthologues
- Paralogues
- Ensembl
- Ontologics
- GO Biolo
- GO Mole
- GO Cellu
- Phenotypes
- Genetic Vari
- Variant ta
- Variant im
- Structural
- Gene express
- Pathway
- Regulation
- External refe
- Supporting d
- ID History
- Gene hist
- Configure
- Custom t
- Export da
- Share thi
- Bookmar

Configure Page Personal Data

Find a track Active tracks

Description Gene Synonyms Location

neuron navigator 2 [Source: HGNC Symbol Acc: HGNC_15997 #]
FLJ10633, FLJ11030, FLJ23707, HELAD1, KIAA1419, POMF1L2, RAINB1
Chromosome 11:19 350 724-20 121 601 forward strand.

Select from available configurations: Current unsaved Save current configuration

Active tracks

Sequence and assembly

Genes and transcripts

Regulation

Information and decorations

Key

Track style External data

Forward strand User-added track

Reverse strand

Favourite track

Track information

Go to Region in Detail for more tracks and navigation options (e.g. zooming)

Fig8. Tracks configuration

Transcript comparison
Gene alleles
Sequence
Secondary Structure
Comparative
Genomic
- Gene tree
- Gene gall
- Gene gall
- Orthologues
- Paralogues
- Ensembl
- Ontologics
- GO Biolo
- GO Mole
- GO Cellu
- Phenotypes
- Genetic Vari
- Variant ta
- Variant im
- Structural
- Gene express
- Pathway
- Regulation
- External refe
- Supporting d
- ID History
- Gene hist
- Configure
- Custom t
- Export da
- Share thi
- Bookmar

Configure Page Personal Data

Find a track Active tracks

Description Gene Synonyms Location

neuron navigator 2 [Source: HGNC Symbol Acc: HGNC_15997 #]
FLJ10633, FLJ11030, FLJ23707, HELAD1, KIAA1419, POMF1L2, RAINB1
Chromosome 11:19 350 724-20 121 601 forward strand.

Select from available configurations: Current unsaved Save current configuration

Genes and transcripts

Enable/disable all Genes

Change track style CODE 39

Genes and transcripts

Regulation

Information and decorations

Configure RNASeq models

Prediction transcripts

Key

Track style External data

Forward strand User-added track

Reverse strand

Favourite track

Track information

Go to Region in Detail for more tracks and navigation options (e.g. zooming)

Fig9. Tracks configuration

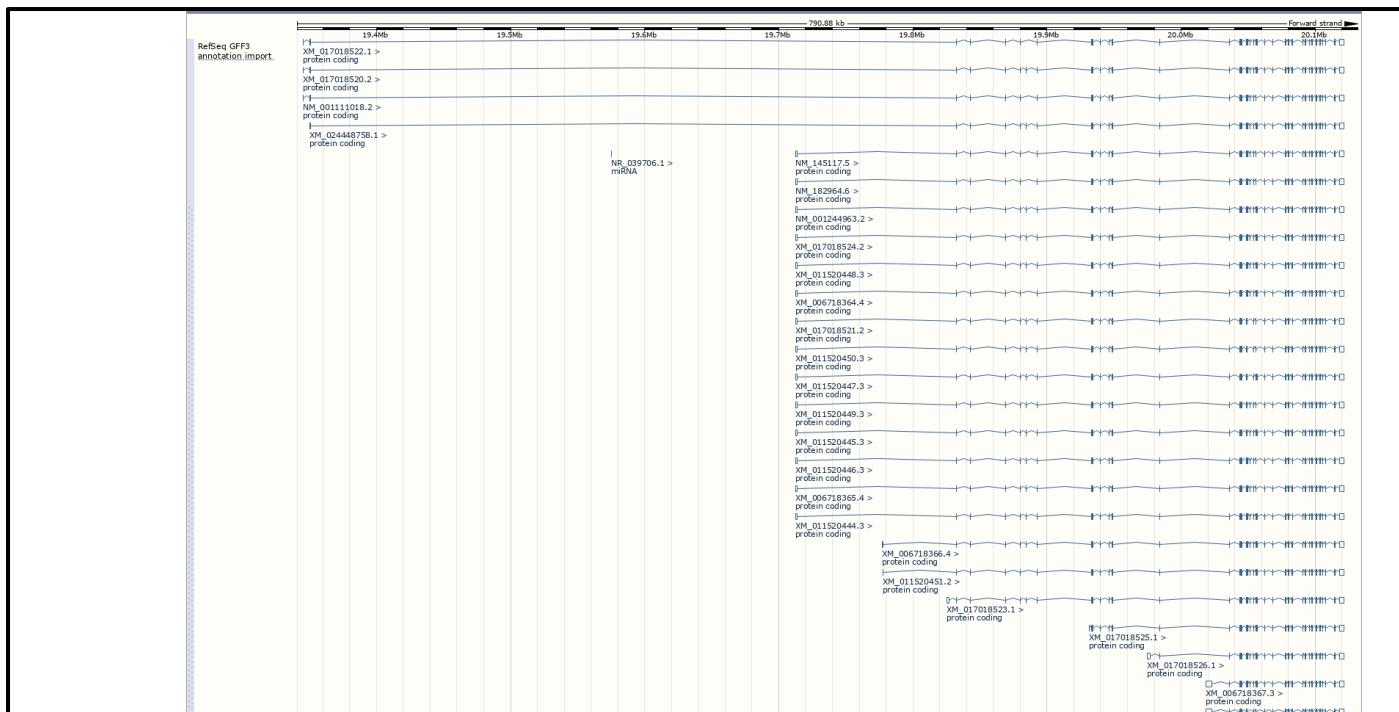


Fig10. Updated results after track configuration

RESULT:

Ensembl genome browser was used to search for helad1 gene under human genome assembly and was explored for various tracks configuration options.

CONCLUSTION:

Ensembl genome browser provides annotation of (human) genomic sequence with genes and their constituent transcripts. Beyond providing access to data related to publicly available genome annotation, Ensembl integrates a number of tools designed to process or analyze your own data. Sequence alignment using BLAST and BLAT against Ensembl genes, genomes and proteins is also available, along with a suite of tools developed as part of the 1000 Genomes Project that can be accessed on the dedicated GRCh37 browser site.

REFERENCES:

1. Karolchik, D. (2003). The UCSC Genome Browser Database. , 31(1), 51–54. doi:10.1093/nar/gkg129
2. Ensembl genome browser 100. (n.d.-b). Uswest.ensembl.org. Retrieved March 28, 2022, from <https://asia.ensembl.org/index.html>
3. Homo_sapiens - Ensembl genome browser 104. (2014). Ensembl.org. Retrieved March 28, 2022, from https://asia.ensembl.org/Homo_sapiens/Info/Index
4. helad1 - Search - Homo_sapiens - Ensembl genome browser 105. (2021b). Ensembl.org. Retrieved March 28, 2022, from <https://asia.ensembl.org/Human/Search/Results?q=helad1>
5. Summary - Homo sapiens - Ensembl genome browser 100. (n.d.). Uswest.ensembl.org. Retrieved March 28, 2022, from https://asia.ensembl.org/Homo_sapiens/Gene/Summary?db=core
6. Summary - Homo sapiens - Ensembl genome browser 100. (n.d.). Uswest.ensembl.org. Retrieved March 28, 2022, from https://asia.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000166833;r=11:19350724-20121601;t=ENST00000349880
7. Summary - Homo sapiens - Ensembl genome browser 100. (n.d.). Uswest.ensembl.org. Retrieved March 28, 2022, from https://asia.ensembl.org/Homo_sapiens/Transcript/Summary?db=core;g=ENSG00000166833;r=11:19350724-20121601;t=ENST00000349880

WEBLEM 9c

Genome Data Viewer

(URL: <https://www.ncbi.nlm.nih.gov/genome/gdv/>)

AIM:

To explore graphical displays of features on NCBI's assembly of human genomic sequence data as well as cytogenetic, genetic, physical, and radiation hybrid maps using Genome Data Viewer (GDV).

INTRODUCTION:

GDV is composed of an **embedded instance of SV** that displays **sequence and track data**, along with **additional page elements** that allow a user to **search within an entire genome** assembly and efficiently **narrow in on their chromosome, sequence, region, or gene of interest**. GDV replaced the NCBI Map Viewer, NCBI's **previous** tool for whole-genome display. Researchers using GDV **can go directly** to the NCBI BLAST service from the browser and **load BLAST results** as alignment tracks that can be viewed **side by side** with **gene annotation** and other data. Variation Viewer, a related browser **associated with NCBI's** variation resources, is **functionally similar to GDV** and also **incorporates an instance of SV** but is **configured with features specifically intended for analyzing human variation data**. GDV and Variation Viewer can both **display the same types of NCBI** variation track data.

METHODOLOGY:

1. Open homepage for GDV genome browser (URL: <https://www.ncbi.nlm.nih.gov/genome/gdv/>)
2. Select human genome assembly
3. Search for DNA repair in genome
4. Select BRCA1 gene
5. Observe the results
6. Use various configuration options
7. Interpret the results.

OBSERVATION:

The screenshot shows the NCBI Genome Data Viewer homepage. At the top, there are links for U.S. National Library of Medicine and NCBI National Center for Biotechnology Information, along with a 'Log in' button. The main title is 'Genome Data Viewer'. Below the title, there is a 'Switch view' button (grid icon) and a 'Search organisms' input field containing 'Homo sapiens (human)'. A note says: 'To view more organisms in the tree, click on nodes that have '+' signs. Press and hold the '+' to expand and reveal all the subgroups. Or, search for an organism using the search box above.' Another note says: 'New! Click on Switch view at the top to see another way of navigating genomes.' The central feature is a phylogenetic tree diagram showing the evolutionary relationships between various organisms. The tree includes nodes for yeast, fruit fly, nematode, Aedes albopictus, human, chimpanzee, rat, mouse, cattle, sheep, pig, horse, dog, zebrafish, chicken, maize, rice, Plasmodium falciparum 3D7, Arabidopsis, grape, and soybean. The 'Homo sapiens (human)' node is highlighted with a green circle. To the right of the tree, there is a detailed assembly summary for 'Homo sapiens (human)'. It includes the following information:

Assembly details	
Name	GRCh38.p13
RefSeq accession	GCF_000001405.39
GenBank accession	GCA_000001405.28
Submitter	Genome Reference Consortium
Level	Chromosome
Category	Reference genome
Replaced by	GCF_000001405.25

Annotation details:

Annotation Release	109
Release date	Nov 22, 2021

Feedback

Fig1. Homepage for Genome Data Viewer

Genome Data Viewer

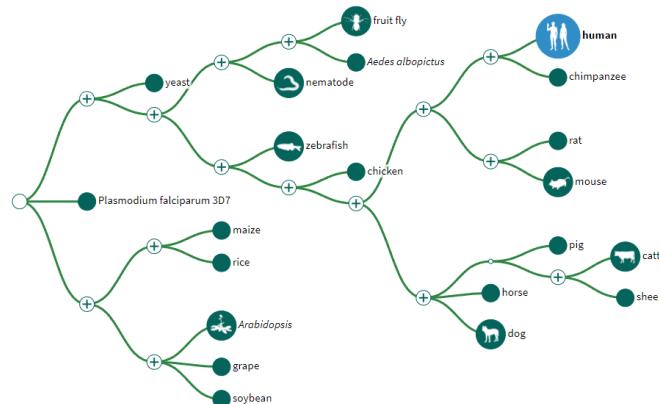


Search organisms

Homo sapiens (human)

To view more organisms in the tree, click on nodes that have '+' signs. Press and hold the '+' to expand and reveal all the subgroups. Or, search for an organism using the search box above.

New! Click on Switch view at the top to see another way of navigating genomes.



GDV supports the exploration and analysis of *NCBI-annotated* and selected non-NCBI annotated eukaryotic genome assemblies. Currently, assemblies from over 1500 organisms are available.

Homo sapiens (human)

Search in genome

DNA repair



Assembly

GRCh38.p13

Browse genome

BLAST genome

Download via NCBI Datasets

Assembly details

Name GRCh38.p13

RefSeq accession GCF_000001405.39

GenBank accession GCA_000001405.28

Submitter Genome Reference Consortium

Level Chromosome

Category Reference genome

Replaced by GCF_000001405.25

Annotation details

Annotation Release 109

Release date Nov 22, 2021

Feedback

Fig2. Search for DNA repair in GRCh38.p13 assembly

Genome Data Viewer

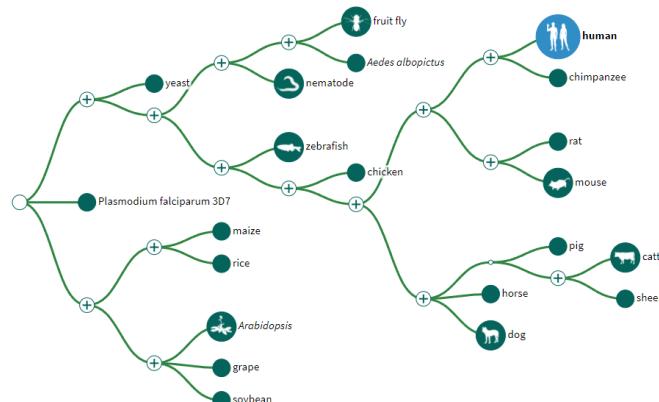


Search organisms

Homo sapiens (human)

To view more organisms in the tree, click on nodes that have '+' signs. Press and hold the '+' to expand and reveal all the subgroups. Or, search for an organism using the search box above.

New! Click on Switch view at the top to see another way of navigating genomes.



GDV supports the exploration and analysis of *NCBI-annotated* and selected non-NCBI annotated eukaryotic genome assemblies. Currently, assemblies from over 1500 organisms are available.

Homo sapiens (human)

Search in genome

DNA repair



Genes

Other features

	Name	Location
BRCA1	Chr17: 43,044,295 - 43,125,364	
ERCC2	Chr19: 45,340,837 - 45,370,647	
BRCA2	Chr13: 32,315,508 - 32,400,268	
APEX1	Chr14: 20,455,226 - 20,457,767	
XRCC5	Chr2: 216,109,348 - 216,205,293	
ERCC4	Chr16: 13,920,137 - 13,952,348	
XRCC6	Chr22: 41,621,295 - 41,664,041	
ERCC3	Chr2: 127,257,290 - 127,294,166	

Examples: TP53, chr17:7667000-7689000, DNA repair

Assembly

GRCh38.p13

Browse genome

BLAST genome

Download via NCBI Datasets

Assembly details

Name GRCh38.p13

Feedback

Fig3. Results for DNA repair genes

The GDV team is hard at work building a new Comparative Genome Viewer. Please send us a comment using the [Feedback survey](#) if you'd like to help with beta testing!

Genome Data Viewer

Homo sapiens
(human)

Assembly: GRCh38.p13 (GCF_000001405.39) • Chr 17 (NC_000017.11)

Home Share this page Reset All More Info

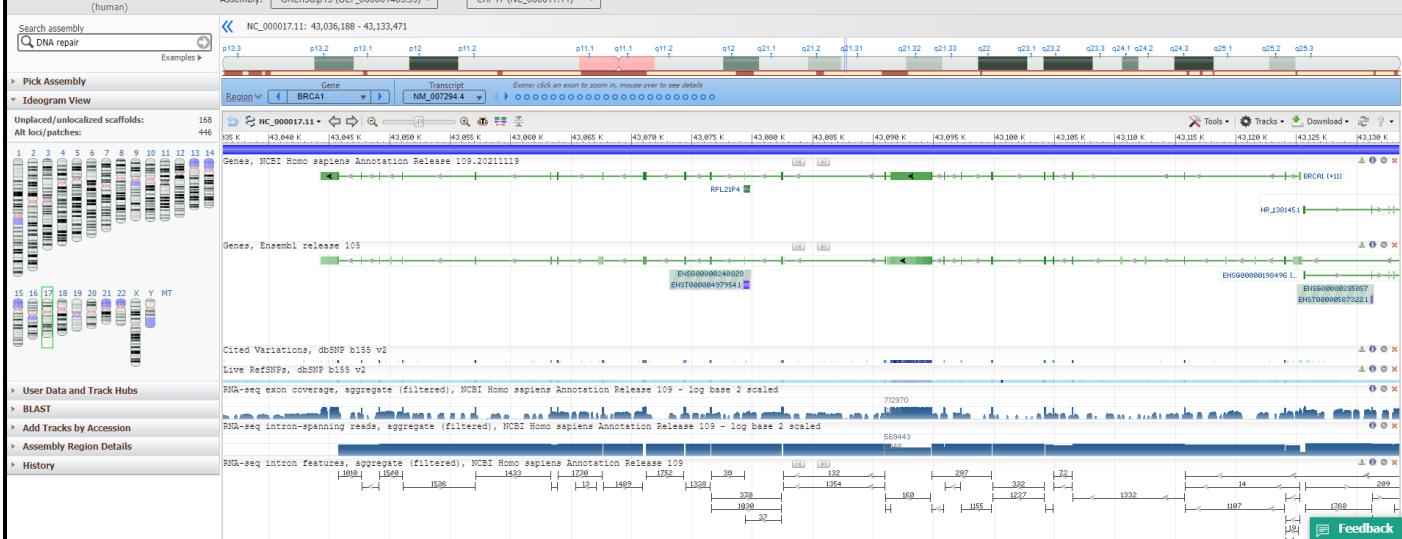


Fig4. Result for BRCA1 gene

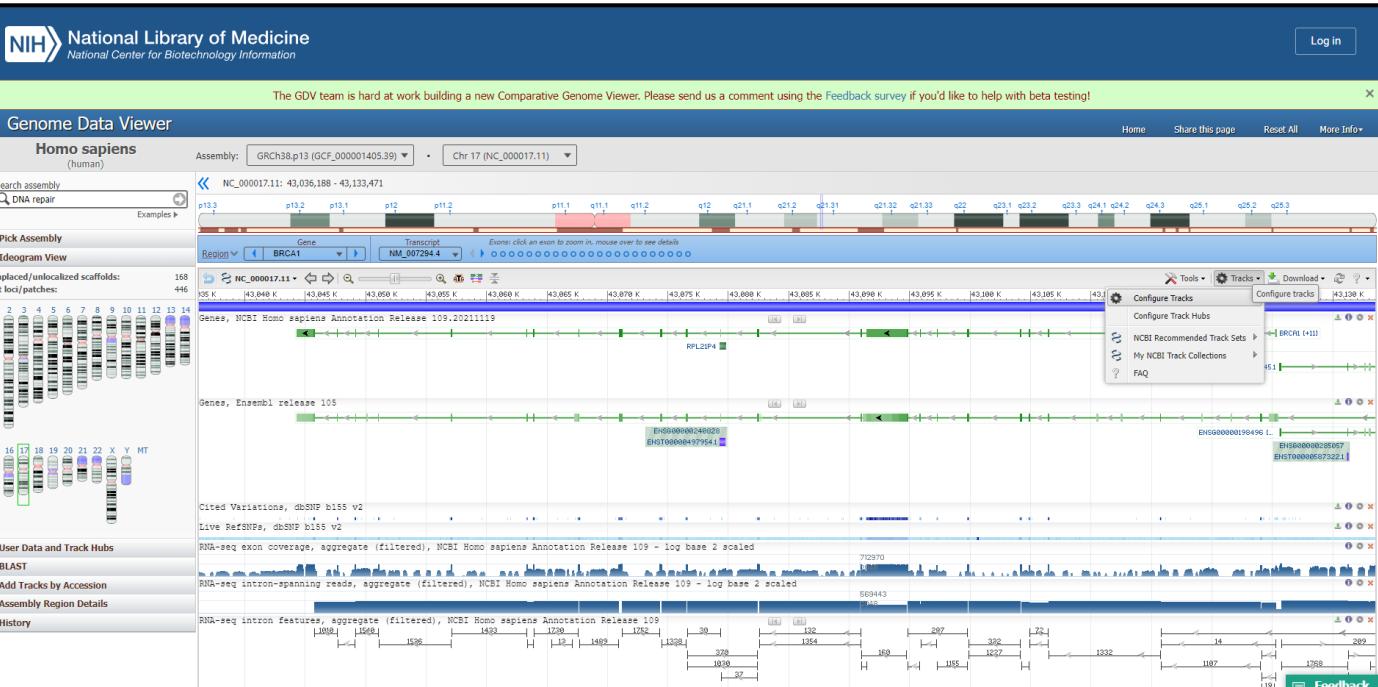


Fig5. Steps to configure tracks

The GDV team is hard at work building a new Comparative Genome Viewer. Please send us a comment using the Feedback survey if you'd like to help with beta testing!

Genome Data Viewer

Homo sapiens
(human)

Assembly: GRCh38.p13 (GCF_000001405.3)

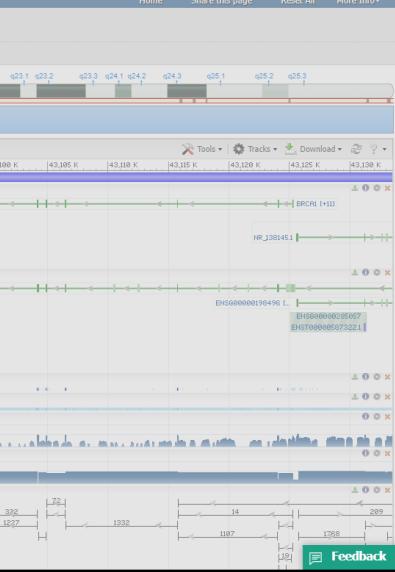
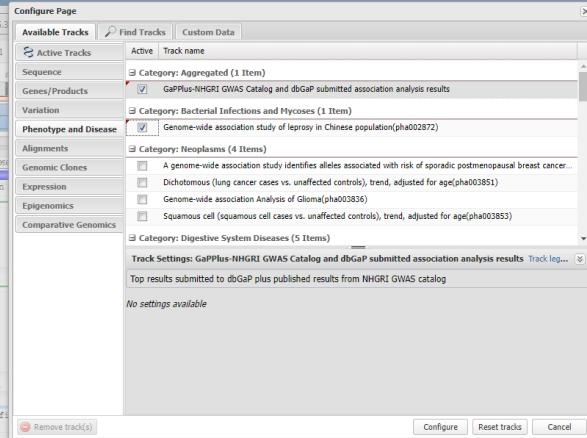
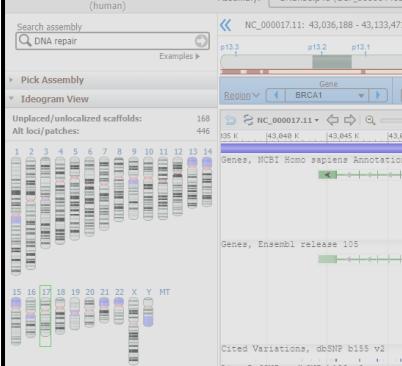


Fig6. Configuration page for phenotype and disease

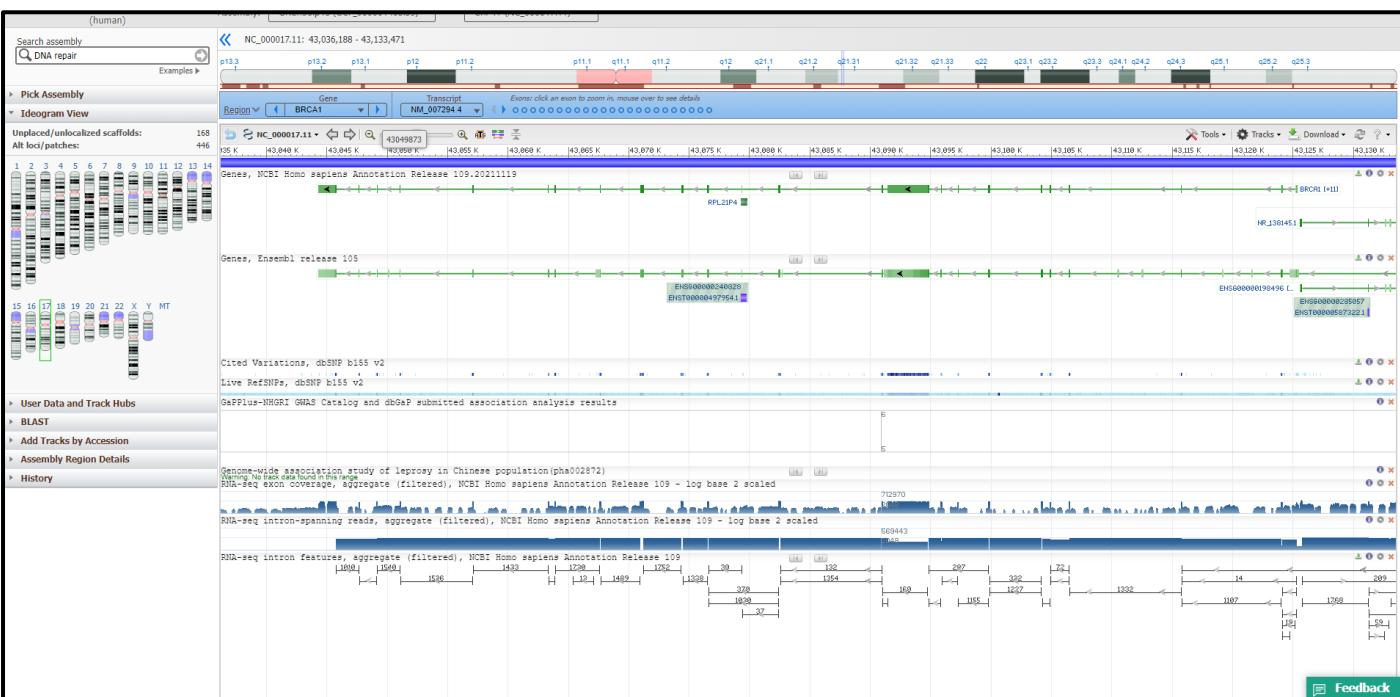


Fig7. Updated result after configuration

The GDV team is hard at work building a new Comparative Genome Viewer. Please send us a comment using the [Feedback survey](#) if you'd like to help with beta testing!

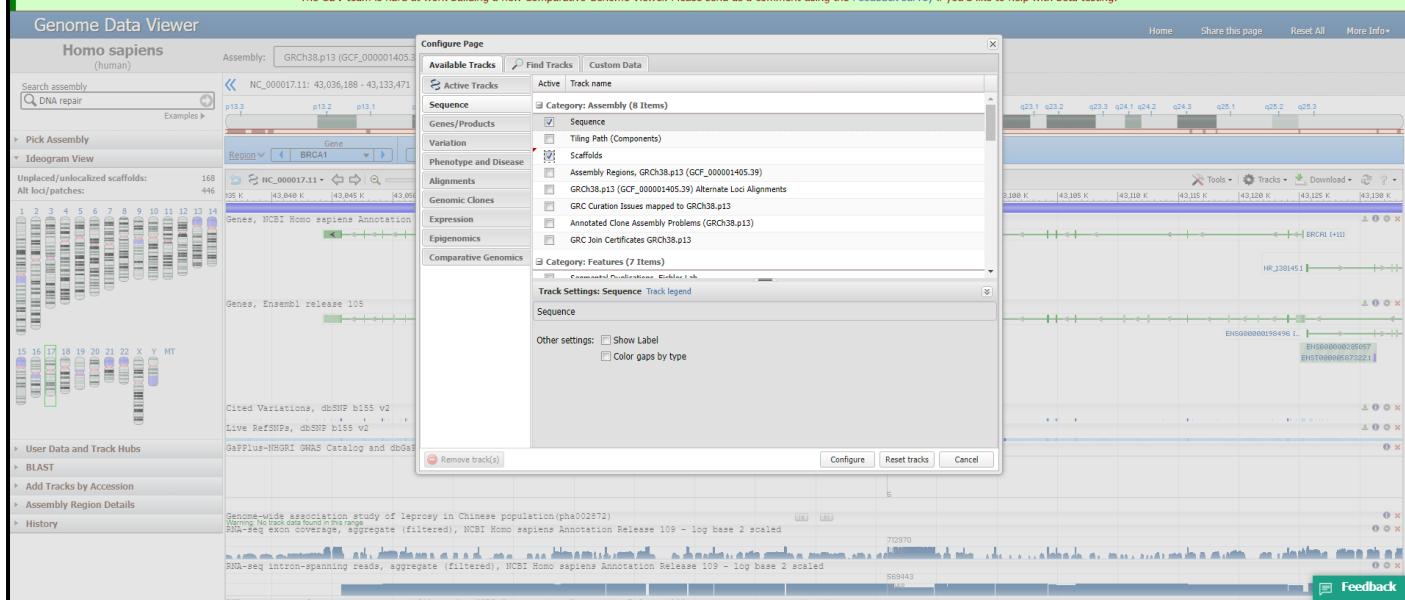


Fig8. Configuration page for sequence

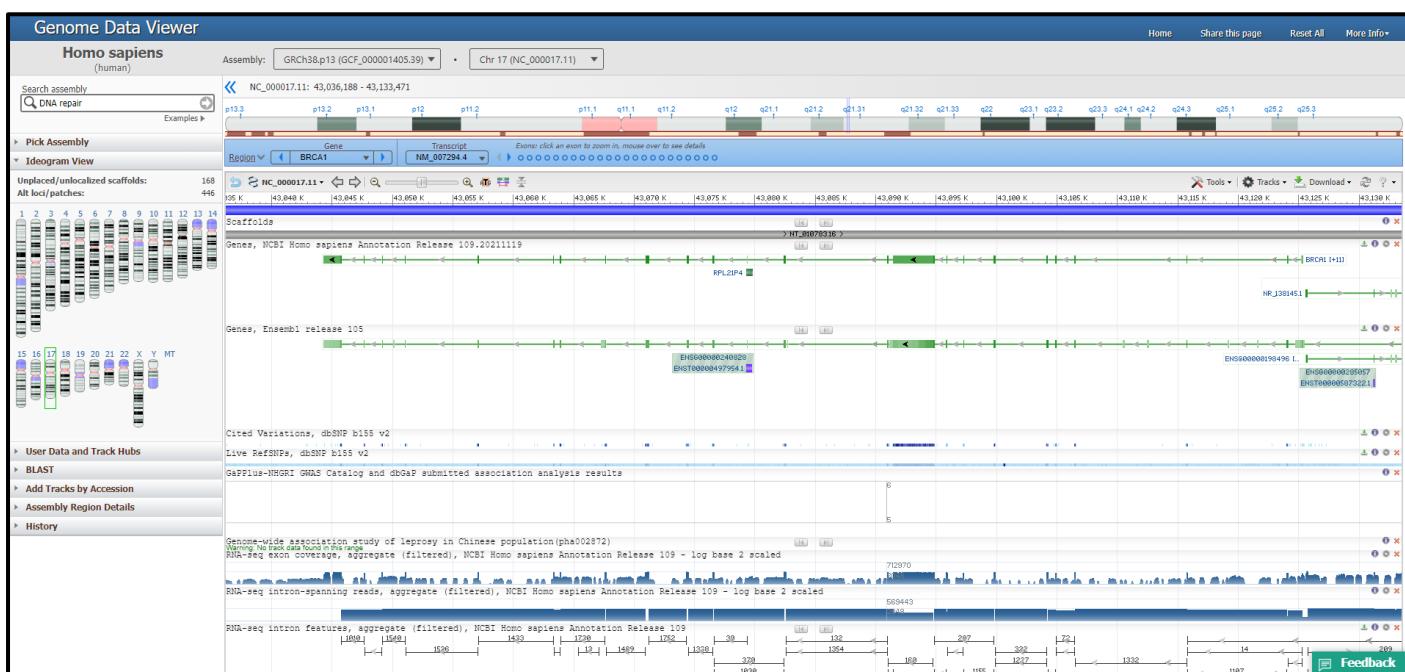


Fig9. Updated result after configuration

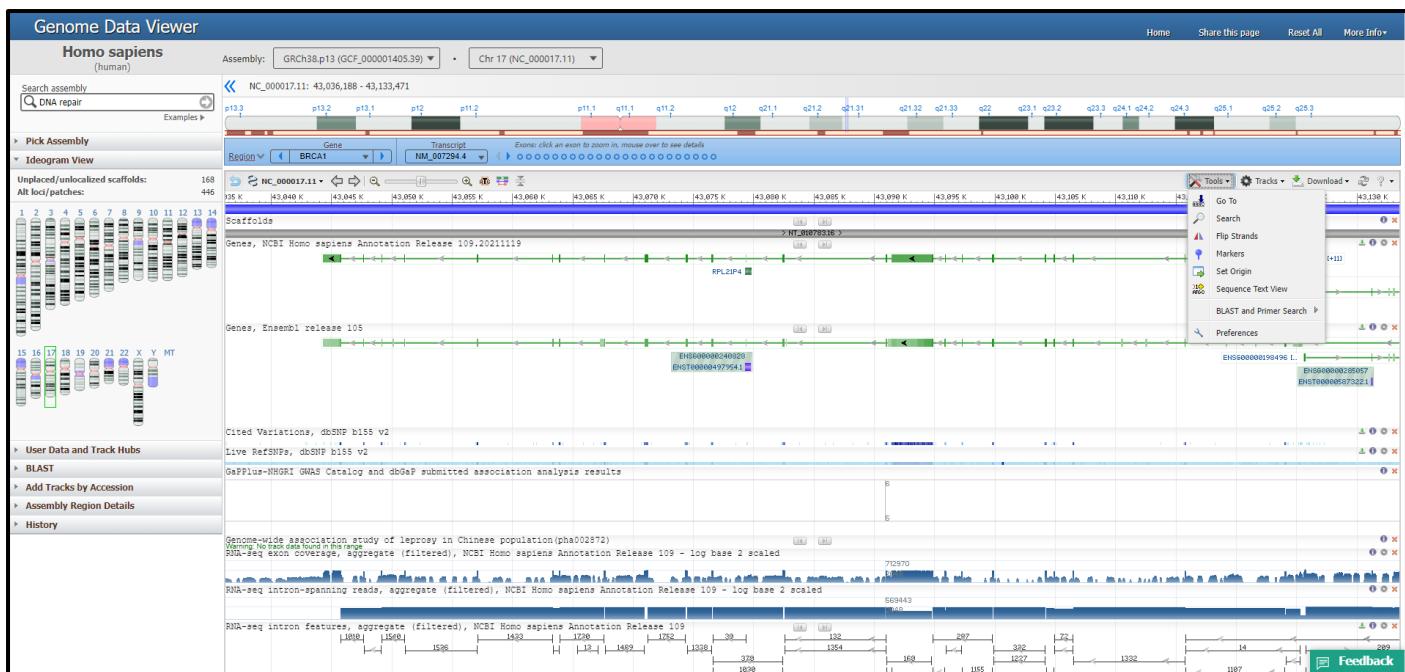


Fig10. Steps for sequence text view

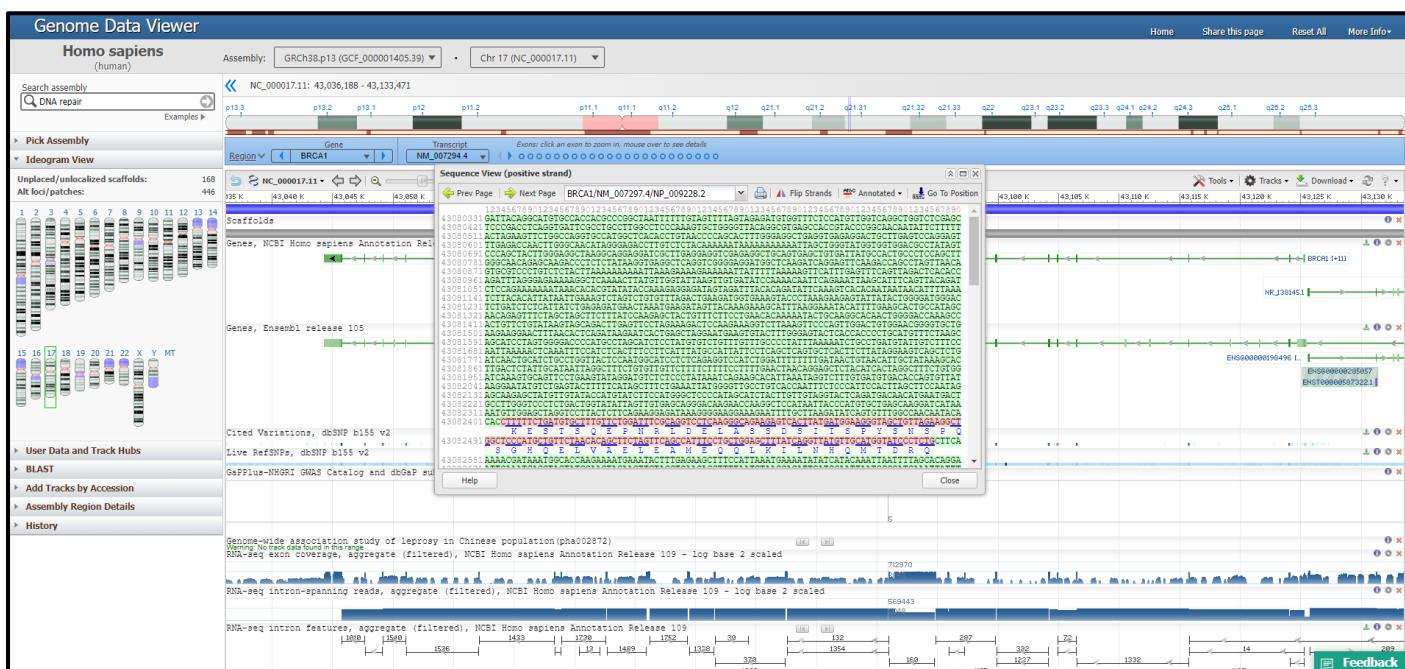


Fig11. Result for sequence text view

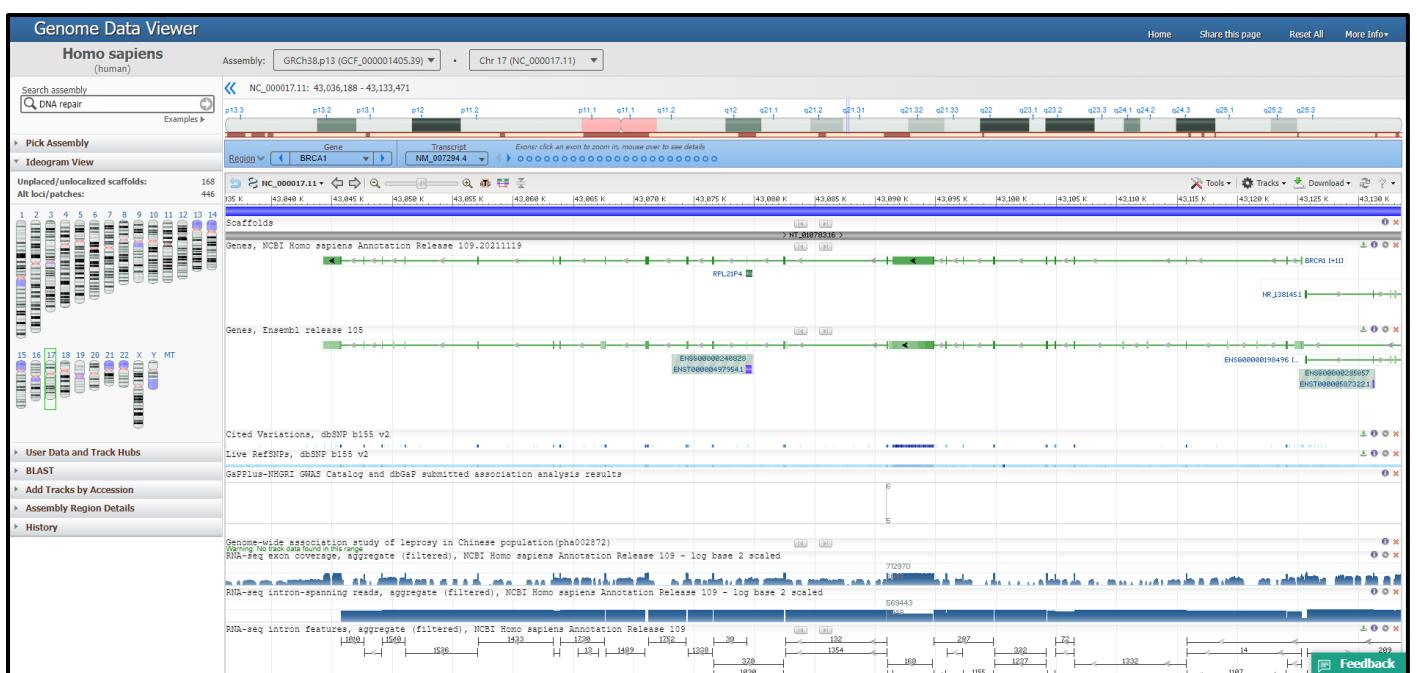


Fig12. Options to view exon information

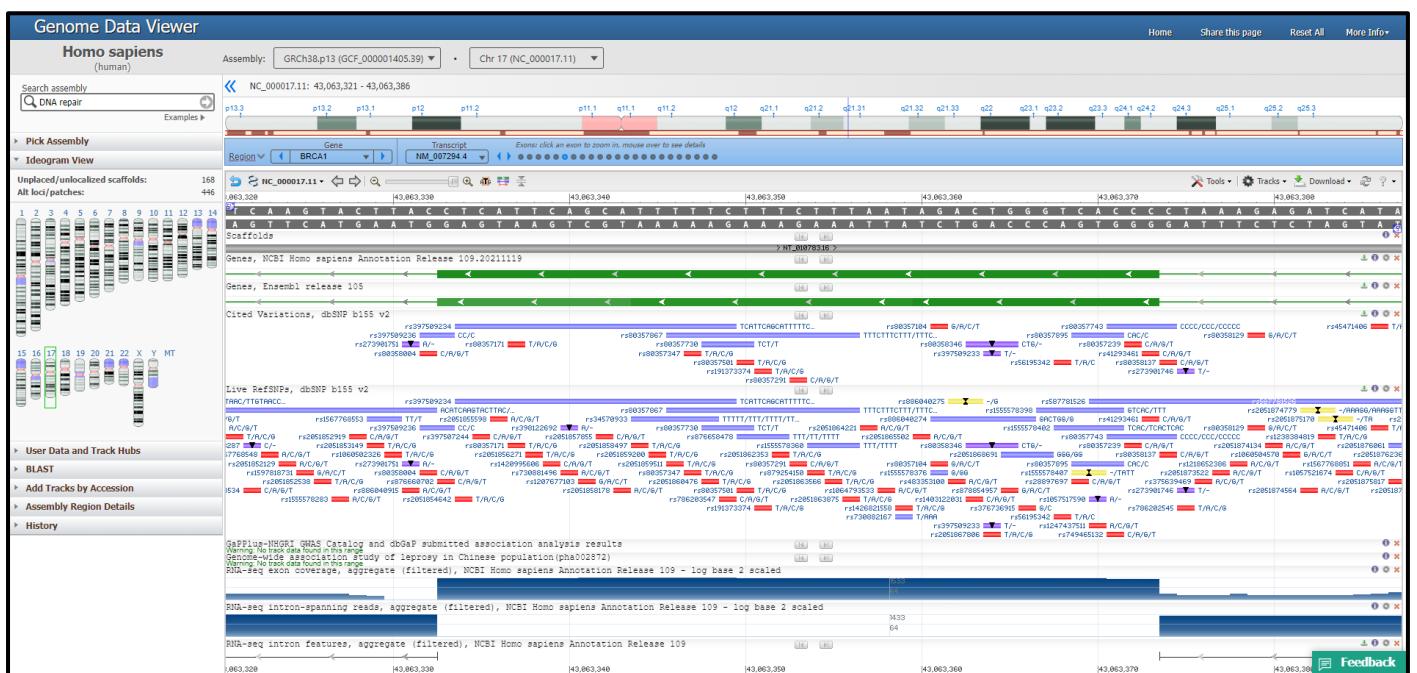


Fig13. Result for exon 18

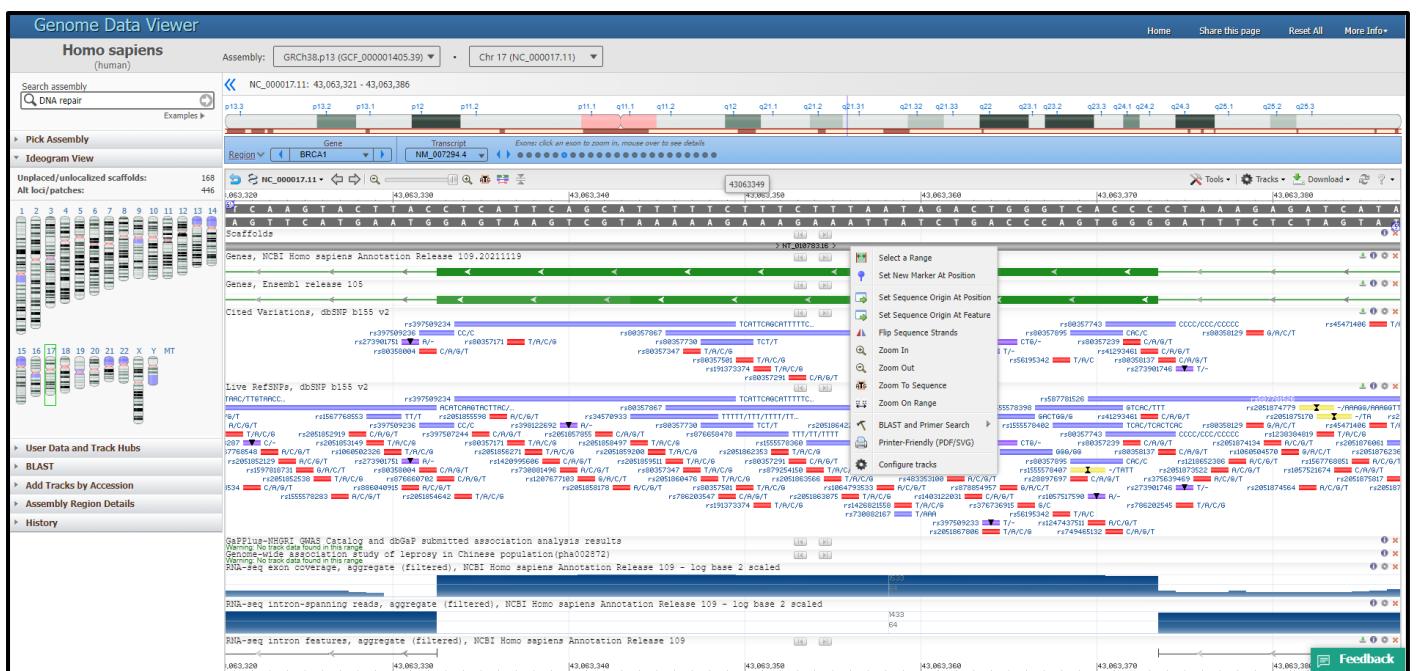


Fig14. Drag and select option for configuring tracks

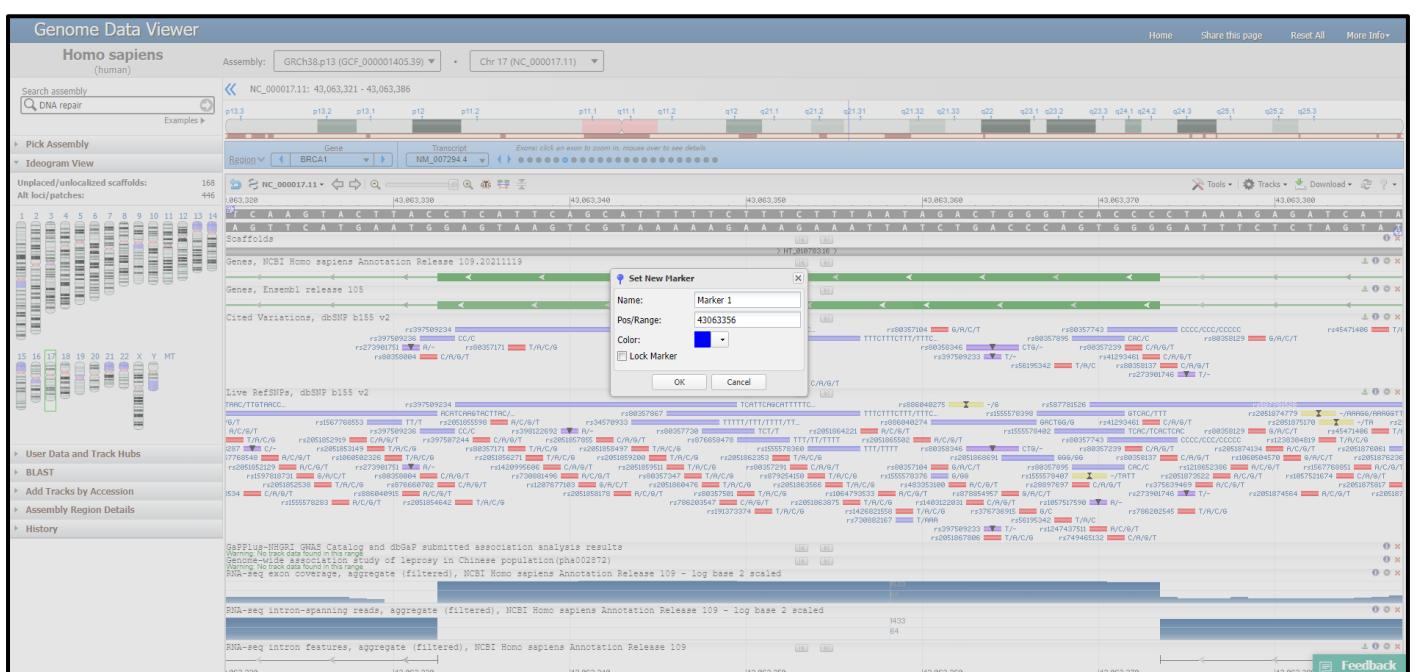


Fig15. Option to set new marker

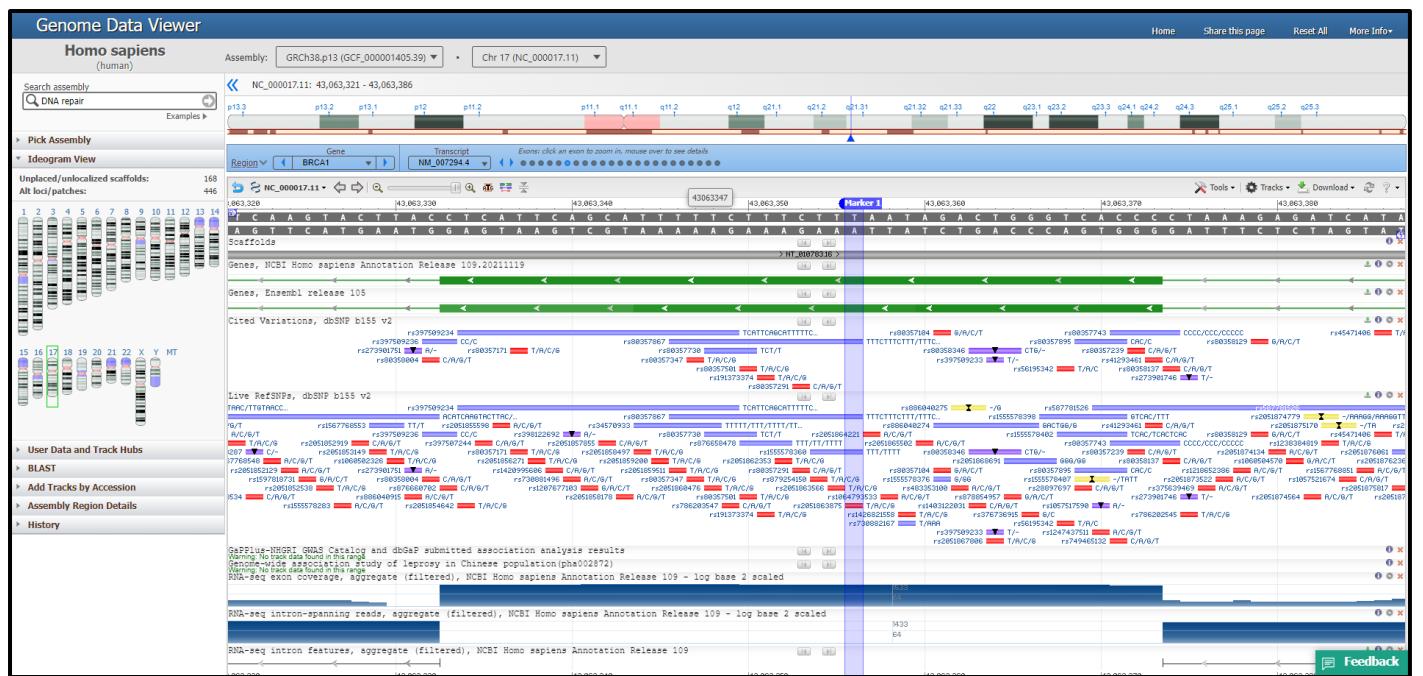


Fig16. Result for new marker

RESULT:

GDV genome browser was used to search for DNA repair under human genome assembly and results were observed for BRCA1 gene. Various options for tracks configuration were explored and information regarding sequence and exons were also viewed.

CONCLUSION:

GDV can be used for visualization and analysis of the wide range of genomes and assemblies annotated at the NCBI. RefSeq gene annotation data tracks are shown by default in the graphical view for these assemblies. NCBI ref SNP data tracks are also shown by default for human assemblies. GDV offers users the ability to customize the displays of individual tracks. Users can hide or configure tracks from the track configuration panel or by using the icons at the right end of each track.

REFERENCES:

1. Rangwala, S. H., Kuznetsov, A., Ananiev, V., Asztalos, A., Borodin, E., Evgeniev, V., Joukov, V., Lotov, V., Pannu, R., Rudnev, D., Shkeda, A., Weitz, E. M., & Schneider, V. A. (2020). Accessing NCBI data using the NCBI Sequence Viewer and Genome Data Viewer (GDV). *Genome Research*, gr.266932.120. <https://doi.org/10.1101/gr.266932.120>
2. NCBI Genome Data Viewer. (n.d.). [Www.ncbi.nlm.nih.gov](https://www.ncbi.nlm.nih.gov/genome/gdv/). Retrieved March 28, 2022, from <https://www.ncbi.nlm.nih.gov/genome/gdv/>
3. Genome Data Viewer. (n.d.). [Www.ncbi.nlm.nih.gov](https://www.ncbi.nlm.nih.gov/genome/gdv/browser/genome/?id=GCF_000001405.39). Retrieved March 28, 2022, from https://www.ncbi.nlm.nih.gov/genome/gdv/browser/genome/?id=GCF_000001405.39