



Лекция № 6

Геномные базы данных Геномные браузеры

Разработал профессор П. М. Васильев
Кафедра фармакологии и биоинформатики

Для студентов, обучающихся по направлению 06.03.01 «Биология»
профили Биохимия, Генетика
при изучении дисциплины «Биоинформатика»

П л а н л е к ц и и

- **Геномные базы данных**
- **Геномные браузеры**
- **Аннотация генома**
- **Распознавание структурно-функциональных мотивов в генетических текстах**

Геномные базы данных

GenBank

The screenshot shows the NCBI GenBank website. At the top, there is a navigation bar with "NCBI Resources" and "How To" menus, and a "Sign in to NCBI" link. Below this is a search bar with "GenBank" selected and a "Nucleotide" dropdown menu. A "Search" button is to the right of the search bar. Below the search bar is a horizontal menu with various database categories: GenBank, Submit, Genomes, WGS, Metagenomes, TPA, TSA, INSDC, and Other. The main content area is divided into two columns. The left column contains a "GenBank Overview" section with a sub-heading "What is GenBank?". This section includes a paragraph explaining that GenBank is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences. It also mentions that GenBank is part of the International Nucleotide Sequence Database Collaboration, which includes the DNA DataBank of Japan (DDBJ), the European Nucleotide Archive (ENA), and GenBank at NCBI. A second paragraph states that a GenBank release occurs every two months and is available from the ftp site, with links to release notes and previous releases. A third paragraph mentions GenBank growth statistics for both traditional divisions and the WGS division. A fourth paragraph provides an example of an annotated sample GenBank record for a *Saccharomyces cerevisiae* gene. The right column contains a "GenBank Resources" section with links to GenBank Home, Submission Types, Submission Tools, Search GenBank, and Update GenBank Records. Below the "What is GenBank?" section is an "Access to GenBank" section with a paragraph stating there are several ways to search and retrieve data from GenBank. This is followed by a bulleted list of search methods: searching for sequence identifiers and annotations with Entrez Nucleotide (divided into CoreNucleotide, dbEST, and dbGSS), searching and aligning GenBank sequences to a query sequence using BLAST, searching, linking, and downloading sequences using NCBI e-utilities, and accessing ASN.1 and flatfile formats at NCBI's anonymous FTP server.

<https://www.ncbi.nlm.nih.gov/genbank/>

copying, or distribution of the information contained in GenBank.

Confidentiality

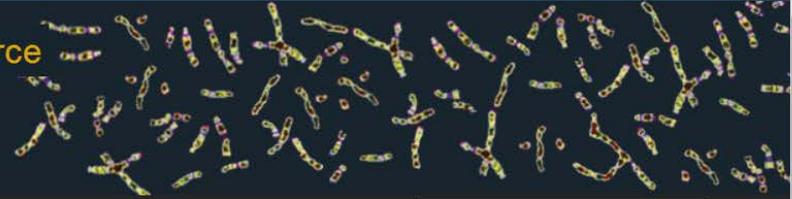
Some authors are concerned that the appearance of their data in GenBank prior to publication will compromise their work. GenBank will, upon request, withhold release of new submissions for a specified period of time. A date must be specified; we can not hold a sequence indefinitely pending publication. However, if a paper citing the sequence or accession number is published prior to the specified date, the sequence will be released upon publication. In order to prevent the delay in the appearance of published sequence data, we urge authors to inform us of the appearance of the published data. As soon as it is available, please send the full publication data--all authors, title, journal, volume, pages and date--to the following address: update@ncbi.nlm.nih.gov

Геномные базы данных

1000 Genomes

IGSR: The International Genome Sample Resource

Providing ongoing support for the 1000 Genomes Project data



[Home](#) [About](#) [Data](#) [Portal](#) [Analysis](#) [Contact](#) [Browser](#) [FAQ](#)

Search 1000genomes



IGSR and the 1000 Genomes Project



Links

- [Announcements](#)
- [IGSR Sample Collection Principles](#)
- [1000 Genomes Project Publications](#)
- [File formats](#)
- [Software tools](#)
- [Download data](#)
- [Twitter](#)

<http://www.internationalgenome.org/>

Populations: ● - African; ● - American; ● - East Asian; ● - European; ● - South Asian;

The International Genome Sample Resource (IGSR) was established to ensure the ongoing usability of data generated by the 1000 Genomes Project and to extend the data set. More information is available [about the IGSR](#).

Геномные базы данных

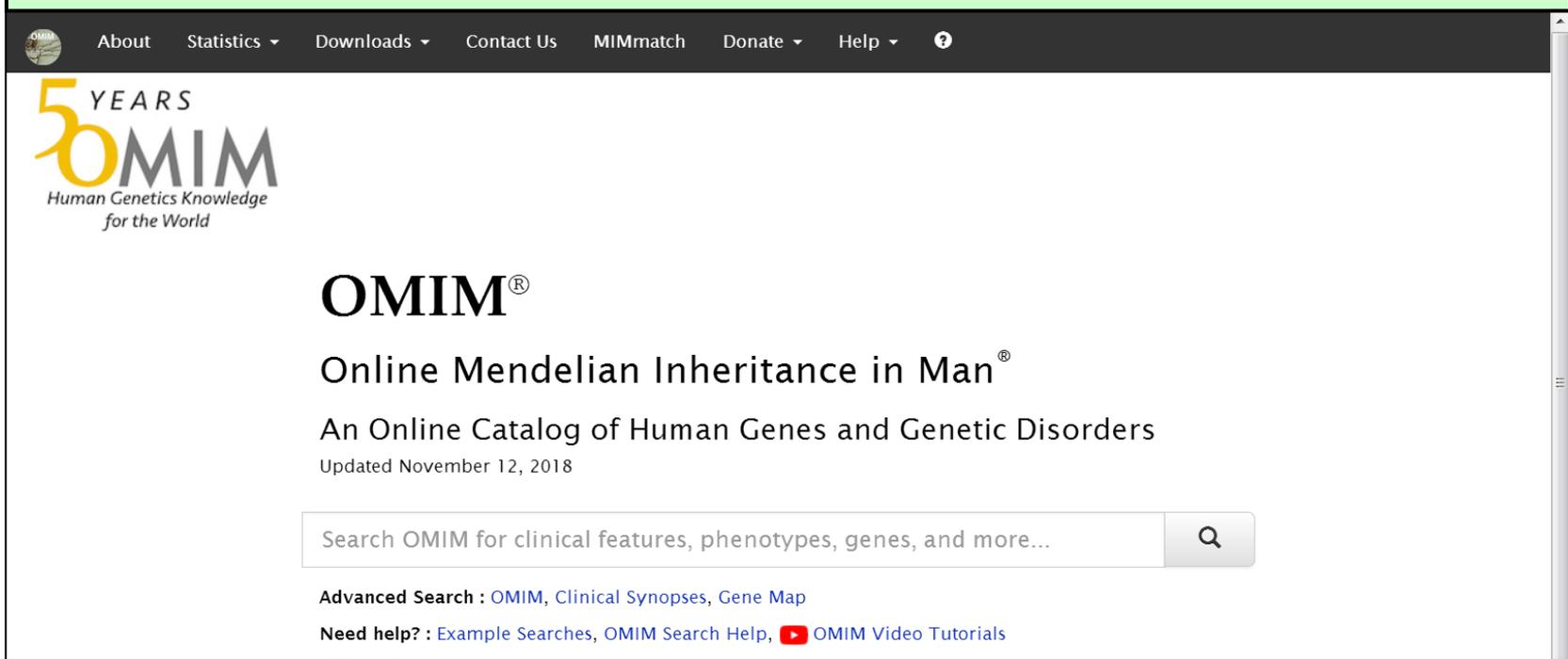
The GDB Human Genome Database

The screenshot shows the homepage of the GDB Human Genome Database. At the top center, there are three blue circles containing the letters 'G', 'D', and 'B'. Below this is the title 'The GDB Human Genome Database' and a subtitle 'The Official World-Wide Database for the Annotation of the Human Genome'. It mentions it is hosted by RTI International, North Carolina USA, and available at mirror sites worldwide. There are links for 'New Database Information' and 'Log in'. A navigation menu on the left includes: Help, Site Map, Reports, Resources, GDB Operations, CF Mutations, HUGO Chromosomes, and GDB Redevelopment Survey. The main content area features a 'Simple Search' section with two columns of radio buttons: 'Search:' (Genomic Segments, All Biological Data, People, Citations) and 'By:' (Name/GDB ID, DNA Sequence ID). There is a search input field and 'Submit' and 'Reset' buttons. A note below the search section states: 'Note: For Name/ID searches, adding * to the end of the search text may improve results.' Below the search section are two columns of links: 'Advanced Search' (Preferences, Other Search Options) and 'Submit and Edit Data' (How to submit data, get D seg #s NEW!, Insert Object).

<http://morissardjerome.free.fr/infobiogen/>
www.gdb.org/gdb/

Геномные базы данных

Online Mendelian Inheritance in Man



The screenshot shows the OMIM website homepage. At the top, there is a navigation bar with links for About, Statistics, Downloads, Contact Us, MIMmatch, Donate, and Help. Below the navigation bar is a logo celebrating 50 years of OMIM, with the tagline "Human Genetics Knowledge for the World". The main heading is "OMIM® Online Mendelian Inheritance in Man®", followed by the subtitle "An Online Catalog of Human Genes and Genetic Disorders" and the date "Updated November 12, 2018". A search bar is present with the placeholder text "Search OMIM for clinical features, phenotypes, genes, and more...". Below the search bar, there are links for "Advanced Search : OMIM, Clinical Synopses, Gene Map" and "Need help? : Example Searches, OMIM Search Help, OMIM Video Tutorials".

<https://www.omim.org/>

Описание генетических заболеваний и генов, их вызывающих
Фенотипические проявления генетических синдромов
Генетические карты

Геномные базы данных

Human Gene Mutation Database



The Human Gene Mutation Database
at the Institute of Medical Genetics in Cardiff

Home Search help Statistics New genes What is new Background Publications Contact Register Login LSDBs Other links

Gene symbol Go! Symbol: Missense/nonsense Go!

The Human Gene Mutation Database (HGMD®) represents an attempt to collate all known (published) gene lesions responsible for human inherited disease and is maintained in Cardiff by D.N. Cooper, E.V. Ball, P.D. Stenson, A.D. Phillips, K. Evans, S. Heywood, M.J. Hayden, M.M. Chapman, M.E. Mort, L. Azevedo and M. Mort

Get HGMD Professional *Please note that this less up-to-date public version of our database is freely available only to [registered](#) users from academic institutions/non-profit organisations. All commercial users are required to purchase a license from QIAGEN®, our commercial partner. A license to [HGMD Professional](#) is available to both commercial and academic non-profit users wishing to access the most up-to-date version of the database (visit QIAGEN® to request a [free trial](#) of HGMD Professional). Read more about how HGMD is [funded](#). You may not copy, store or re-distribute HGMD data without express written permission (i) from the curators or (ii) via your license agreement. Copyright © Cardiff University 2017. All rights reserved. [Register for Public Version](#)

| Table: | Description: | Public entries: <small>This site. Academic non-profit users only</small> | Total entries: <small>HGMD Professional 2018.3</small> |
|---|--|---|---|
| Mutation totals (as of 2018-11-13) | | 157114 | 240269 |
| Gene symbol | The gene description, gene symbol (as recommended by the HUGO Nomenclature Committee) and chromosomal location is recorded for each gene. In cases where a gene symbol has not yet been made official, a provisional symbol has been adopted which is denoted by lower-case letters. | 6531 | 9976 |
| cDNA sequence | cDNA reference sequences are provided, numbered by codon. | 6531 | 10339 |
| Genomic coordinates | Genomic (chromosomal) coordinates have been calculated for missense/nonsense, splicing, regulatory, small deletions, small insertions and small indels. | 0 | 214308 |
| HGVS nomenclature | Standard HGVS nomenclature has been obtained for missense/nonsense, splicing, regulatory, small deletions, small insertions and small indels. | 0 | 214691 |
| Missense/nonsense | Single base-pair substitutions in coding regions are presented in terms of a triplet change with an additional flanking base included if the mutated base lies in either the first or third position in the triplet. | 87397 | 137354 |
| Splicing | Mutations with consequences for mRNA splicing are presented in brief with information specifying the relative position of the lesion with respect to a numbered intron donor or acceptor splice site. Positions given as positive integers refer to a 3' (downstream) location, negative integers refer to a 5' (upstream) location. | 14317 | 21222 |
| Small insertions | Missense insertions (20 bp or less) are presented in terms of the inserted bases in lower case plus, in upper case, 10 bp DNA sequence flanking both sides of the lesion. The numbered codon is preceded in the given sequence by the caret character (^). | 9844 | 14746 |
| Missense indels | Missense indels (20 bp or less) are presented in terms of the deleted/inserted bases in lower case plus, in upper case, 10 bp DNA sequence flanking both | | |
| Complex rearrangements | Information regarding the nature and location of each lesion is logged in narrative form because of the extremely variable quality of the original data reported. | 1575 | 2031 |

<http://www.hgmd.cf.ac.uk/ac/index.php>

Описание мутаций ядерных генов,
приводящих к наследственным заболеваниям

GenBank

Файл Правка Вид Журнал Закладки Инструменты Справка

genbank - Google Search x FDA Whole Genome Sequencin... x GenBank Overview x +

https://www.ncbi.nlm.nih.gov/genbank/ 110% Поиск

Часто посещаемые

NCBI Resources How To Sign in to NCBI

GenBank Nucleotide Search

GenBank Submit Genomes WGS Metagenomes TPA TSA INSDC Other

GenBank Overview

What is GenBank?

GenBank® is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences (*Nucleic Acids Research*, 2013 Jan;41(D1):D9-D20). It is part of the *International Nucleotide Sequence Database Collaboration*, which comprises the DNA DataBank of Japan (DDB), the European Nucleotide Archive (ENA), and GenBank at NCBI. These three organizations exchange data on a daily basis.

A GenBank release occurs every two months and is available from the [ftp site](#). The [release notes](#) for the current version of GenBank provide detailed information about the release and notifications of upcoming changes to GenBank. Release notes for [previous GenBank releases](#) are also available. GenBank growth statistics for both the traditional GenBank divisions and the WGS division are available from each release. GenBank growth [statistics](#) for both the traditional GenBank divisions and the WGS division are available from each release.

An [annotated sample GenBank record](#) for a *Saccharomyces cerevisiae* gene demonstrates many of the features of the GenBank flat file format.

Access to GenBank

There are several ways to search and retrieve data from GenBank.

- Search GenBank for sequence identifiers and annotations with [Entrez Nucleotide](#), which is divided into three divisions: [CoreNucleotide](#) (the main collection), [dbEST](#) (Expressed Sequence Tags), and [dbGSS](#) (Genome Survey Sequences).
- Search and align GenBank sequences to a query sequence using [BLAST](#) (Basic Local Alignment Search Tool). BLAST searches CoreNucleotide, dbEST, and dbGSS independently; see [BLAST info](#) for more information about the numerous BLAST databases.
- Search, link, and download sequences programatically using [NCBI e-utilities](#). The ASN.1 and flatfile formats are available at NCBI's anonymous FTP server (<ftp://ftp.ncbi.nlm.nih.gov/ncbiasn1/>) and (<ftp://ftp.ncbi.nlm.nih.gov/genbank/>).

GenBank Resources

- [GenBank Home](#)
- [Submission Types](#)
- [Submission Tools](#)
- [Search GenBank](#)
- [Update GenBank Records](#)

GenBank Data Usage

The GenBank database is designed to provide and encourage access within the scientific community to the most up-to-date and comprehensive DNA sequence information. Therefore, NCBI places no restrictions on the use or distribution of the GenBank data. However, some submitters may claim patent, copyright, or other intellectual property rights in all or a portion of the data they have submitted. NCBI is not in a position to assess the validity of such claims, and therefore cannot provide comment or unrestricted

Genomes → Complete Genome

GenBank

Файл Правка Вид Журнал Закладки Инструменты Справка

genebank - Google Search x FDA Whole Genome Sequencin... x Prokaryotic and Eukaryotic ... x +

https://www.ncbi.nlm.nih.gov/genbank/genomesubmit/ 110% Поиск

Часто посещаемые

NCBI Resources How To Sign in to NCBI

GenBank Nucleotide Search

GenBank Submit Genomes WGS Metagenomes TPA TSA INSDC Other

Prokaryotic and Eukaryotic Genomes Submission Guide

Both WGS and non-WGS genomes, including gapless complete bacterial chromosomes, can be submitted via the Submission Portal. You will be asked to choose whether the genome being submitted is considered WGS or not. The differences for GenBank purposes are:

non-WGS genome

WGS genome

Chr1 p1 p2 Chr2

Chr1 p1 Chr2

= Gap (Ns)

non-WGS

- Each chromosome is in a single sequence and there are no extra sequences
- Each sequence in the genome must be assigned to a chromosome, plasmid or organelle
- Plasmids and organelles can still be in multiple pieces.

WGS

- One or more chromosomes are in multiple pieces and/or some sequences are not assembled into chromosomes

In both cases

- There can still be gaps within the sequences; you will supply that information in the submission
- Plasmids and organelles can still be in multiple pieces.
- Internal sequences must be arranged in the correct order and orientation.
- Sequences concatenated in unknown order are not allowed.

Genome Resources

- [About WGS](#)
- [WGS Browser](#)
- [Genome Submission Guide](#)
- [Genome Submission Portal](#)
- [Update Genome Records](#)
- [FAQ](#)
- [tbl2asn](#)
- [Create Submission Template](#)
- [Eukaryotic Annotation Guide](#)
- [Prokaryotic Annotation Guide](#)
- [Annotation Example Files](#)
- [Discrepancy Report](#)
- [NCBI Prokaryotic Genome Annotation Pipeline](#)
- [AGP Format](#)
- [Complex Assembly Submission Guide](#)
- [Metagenome Submission Guide](#)
- [BioProject](#)
- [Structured Comment](#)

WGS Browser (Whole Genome Shotgun)

GenBank — WGS Browser

The screenshot shows the NCBI Sequence Set Browser interface. On the left, the 'Available Facets' panel has 'Organism' selected. In the 'Project type' section, the 'All' button is highlighted. The search results table shows 1,344 projects, with the first row (AACC02) circled in red. The text 'Organism — Homo sapiens' is overlaid in large purple font at the bottom.

Available Facets:

- Type
- Source database
- Targeted Locus Name
- Project type
- Organism**
- Bioproject
- Biosample
- Strain
- Breed
- Cultivar
- Isolate
- Tissue Type
- Host
- Isolation Source
- Dev Stage

Project type: All WGS TLS TSA

Search: Term: Search Reset All

Found 1,344 projects Download Columns List Page 1 2 3 4 5 6 7 8 ... 27 (50) per page

| # | Prefix | Type | DIV | Organism | Bioproject | Biosample | Intraspecific Name | Other Source | Contigs | | | | | |
|---|--------|------|-----|--------------|------------|--------------|--------------------|--------------|-----------------------|---------|--------|-----------|--------|--------|
| | | | | | | | | | Total Length (Mbases) | # | # Prot | Has Annot | # | |
| 1 | AACC02 | WGS | PRI | Homo sapiens | PRJNA10793 | SAMN02981216 | | | 137.7 | 171 | 109 | Y | 20 | |
| 2 | AADB02 | WGS | PRI | Homo sapiens | PRJNA1431 | SAMN02981219 | | | 3,087.3 | 401,931 | | | 14,095 | 64,... |
| 3 | AADD01 | WGS | PRI | Homo sapiens | PRJNA1431 | SAMN02981219 | | | 2,756.6 | 61,715 | | | 2,639 | |
| 4 | AADD01 | WGS | PRI | Homo sapiens | PRJNA1431 | SAMN02981219 | | | 2,695.6 | 211,493 | | | 1,372 | |

Organism — Homo sapiens

GenBank — Project: AACCC02

Файл Правка Вид Журнал Закладки Инструменты Справка

GenBank - Поиск в Google x AACCC00000000.2 Homo sa... x +

https://www.ncbi.nlm.nih.gov/Traces/wgs/AACCC02?display=contigs&page=1 97% Поиск

NCBI Sequence Set Browser Sign In to NCBI

Project: AACCC02 Search List of all Projects

AACCC0000000.2 Homo sapiens

Master Contigs Proteins Download

| | | |
|----------------------|---|--|
| # of Contigs: | 171 | On Jul 19, 2004 this sequence version replaced gi:29727032. |
| # of Proteins: | 818 | The Homo sapiens whole genome shotgun (WGS) project has the project accession AACCC00000000. This version of the project (02) has the accession number AACCC02000000, and consists of sequences AACCC02000001-AACCC02000171. |
| # of Scaffolds/Chrs: | 20 | |
| Total length: | 137,712,494 bp | |
| BioProject: | PRJNA10793 | |
| BioSample: | SAMN02981216 | |
| Keywords: | WGS | |
| Annotation: | Contigs, Scaffolds | |
| Organism: | Homo sapiens show lineage | |
| Biosource: | /chromosome = 7 | |
| | /mol_type = genomic genomic | |
| WGS: | AACCC02000001:AACCC02000171 | |
| Scaffolds: | CH236947:CH236966 | |
| | 20 scaffolds, 709 proteins, total length is 123,763,301 bases | |
| Reference: | Human chromosome 7: DNA sequence and biology : Science 300 (5620), 767-772 (2003) – show 90 authors | |
| Submission: | Submitted (27-MAR-2003) Department of Genetics and Genomic Biology, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada – show 88 authors | |
| Submission: | Submitted (29-JUN-2004) Department of Genetics and Genomic Biology, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada – show 88 authors | |

National Center for Biotechnology Information, U.S. National Library of Medicine
8600 Rockville Pike, Bethesda MD, 20894 USA
[Policies and Guidelines](#) | [Contact](#)

Chromosome 7

NATIONAL LIBRARY OF MEDICINE NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION USA.gov

GenBank — Project: AACCC02

GenBank - Поиск в Google x AACCC00000000.2 Homo sa... x +

https://www.ncbi.nlm.nih.gov/Traces/wgs/AACCC02?display=proteins&page=1

NCBI Sequence Set Browser Sign In to NCBI

Project: AACCC02 Search List of all Projects

AACCC00000000.2 Homo sapiens

Master **Contigs** **Proteins** Download

Accession Product name Search Clear

Found 109 proteins Download Page 1 2 3 (50) per page

| # | Accession | Length | Product name | View |
|----|------------|--------|---|---------------------|
| 1 | EAL24413.1 | 519 | Rho guanine nucleotide exchange factor (GEF) 5 | FASTA GenPept ASN.1 |
| 2 | EAL24414.1 | 490 | similar to OG-2 homeodomain protein-like; similar to U65067 (PID:g1575526) | FASTA GenPept ASN.1 |
| 3 | EAL24415.1 | 243 | thiamin pyrophosphokinase 1 | FASTA GenPept ASN.1 |
| 4 | EAL24416.1 | 366 | similar to Importin alpha-2 subunit (Karyopherin alpha-2 subunit) (SRP1-alpha) (RAG cohort protein 1) | FASTA GenPept ASN.1 |
| 5 | EAL24417.1 | 1,331 | contactin associated protein-like 2 | FASTA GenPept ASN.1 |
| 6 | EAL24418.1 | 174 | similar to Mtr3 (mRNA transport regulator 3)-homolog; Mtr3 (mRNA transport regulator 3)-homolog (yeast) | FASTA GenPept ASN.1 |
| 7 | EAL24419.1 | 132 | similar to 60S ribosomal protein L32 | FASTA GenPept ASN.1 |
| 8 | EAL24420.1 | 177 | chromosome 7 open reading frame 33 | FASTA GenPept ASN.1 |
| 9 | EAL24421.1 | 101 | similar to Huntingtin interacting protein K | FASTA GenPept ASN.1 |
| 10 | EAL24422.1 | 776 | cullin 1 | FASTA GenPept ASN.1 |
| 11 | EAL24423.1 | 707 | enhancer of zeste homolog 2 (Drosophila) | FASTA GenPept ASN.1 |
| 12 | EAL24424.1 | 751 | enhancer of zeste homolog 2 (Drosophila) | FASTA GenPept ASN.1 |
| 13 | EAL24425.1 | 645 | protein disulfide isomerase related protein (calcium-binding protein, intestinal-related) | FASTA GenPept ASN.1 |
| 14 | EAL24426.1 | 782 | hypothetical protein DKFZp7621137 | FASTA GenPept ASN.1 |
| 15 | EAL24427.1 | 752 | hypothetical protein LOC155054 | FASTA GenPept ASN.1 |

Proteins

GenBank — Project: AACCC02

The screenshot shows the NCBI Sequence Set Browser interface. The browser window title is "GenBank - Поиск в Google" and the address bar shows "https://www.ncbi.nlm.nih.gov/Traces/wgs/AACCC02". The page title is "NCBI Sequence Set Browser" and it includes a "Sign In to NCBI" link. The project name "AACCC02" is entered in the search field. Below the search bar, the project details "AACCC0000000.2 Homo sapiens" are displayed. A red circle highlights the "Download" button in the navigation tabs. The "Contigs" section lists three files: "AACC02.1.gbff.gz" (54.8 Mb), "AACC02.1.fsa_nt.gz" (39.6 Mb), and "AACC02.1.bbs.gz" (31.4 Mb). The "Proteins" section lists two files: "AACC02.1.gnp.gz" (57.7 kb) and "AACC02.1.fsa_aa.gz" (28 kb). The footer contains the NCBI address and logos for the National Library of Medicine and USA.gov.

Project: AACCC02

AACCC0000000.2 Homo sapiens

| Contigs: | | Proteins: | |
|----------|--|-----------|--|
| GenBank: | AACC02.1.gbff.gz 54.8 Mb | GenPept: | AACC02.1.gnp.gz 57.7 kb |
| FASTA: | AACC02.1.fsa_nt.gz 39.6 Mb | FASTA: | AACC02.1.fsa_aa.gz 28 kb |
| ASN.1: | AACC02.1.bbs.gz 31.4 Mb | | |

National Center for Biotechnology Information, U.S. National Library of Medicine
8600 Rockville Pike, Bethesda MD, 20894 USA
[Policies and Guidelines](#) | [Contact](#)

Download — Chromosome 7

GenBank — Reference Genom

The screenshot shows the NCBI Assembly website interface. The browser address bar displays <https://www.ncbi.nlm.nih.gov/assembly>. The search bar contains the text "Homo sapiens genome", which is circled in red. The word "Assembly" in the navigation menu is circled in blue. A pink banner at the top provides information about COVID-19, including links to CDC and NIH resources. Below the banner, a "Filters removed" message is visible. The main content area features a blue header with the word "Assembly" and a description: "A database providing information on the structure of assembled genomes, assembly names and other meta-data, statistical reports, and links to genomic sequence data." The page is organized into three columns: "Using Assembly" (with links like Assembly Help, Browse by Organism, and NCBI Assembly Data Model), "Submitting an Assembly" (with links like Submission Information and Genome Size Check), and "Related Resources" (with links like NCBI Datasets, Genome, and Genome Reference Consortium). At the bottom, a yellow banner displays the URL <https://www.ncbi.nlm.nih.gov/assembly>.

GenBank — Reference Genom

GenBank - Поиск в Google x Homo sapiens genome - A... x +

https://www.ncbi.nlm.nih.gov/assembly/?term=Homo+sapiens+genome

NCBI Resources How To Sign in to NCBI

Assembly Assembly Homo sapiens genome Search

Create alert Advanced Browse by organism Help

COVID-19 is an emerging, rapidly evolving situation.
Get the latest public health information from CDC: <https://www.coronavirus.gov>.
Get the latest research from NIH: <https://www.nih.gov/coronavirus>.
Find NCBI SARS-CoV-2 literature, sequence, and clinical content: <https://www.ncbi.nlm.nih.gov/sars-cov-2/>.

Organism group Summary 20 per page Sort by Significance Send to Filters: Manage Filters

Animals (903)
Customize ...

Status clear Download Assemblies

Latest (903)
Latest GenBank (903)
Latest RefSeq (1)

Assembly level

Complete genome (0)
Chromosome (18)
Scaffold (806)
Contig (79)

RefSeq category

Reference (1)
Representative (0)

Exclude clear

Exclude partial (1)
Exclude anomalous (0)
Customize ...

Annotation status

Has annotation (3)
GenBank has annotation (1)
RefSeq has annotation (3)

Relation to type material

Assembly from any type (0)
Assembly from type (0)
Assembly from synonym type (0)
Assembly from pathotype (0)
Assembly designated as neotype (0)

GENOME ASSEMBLY Was this helpful?

GRCh38.p13

[Homo sapiens \(human\)](#)

Genome Reference Consortium (March 2019)
RefSeq GCF_000001405.39
[PubMed \(27\)](#) [Genome Resources](#)

Genome Browser BLAST Get data

Assembly statistics +

Search results << First < Prev Page 1 of 46 Next > Last >>

Items: 1 to 20 of 903

Filters activated: Latest, Latest GenBank, Exclude anomalous. [Clear all](#) to show 1068 items.

GRCh38.p13

1. Genome Reference Consortium Human Build 38 patch release 13 (GRCh38.p13)

NCBI Datasets Download a genome dataset including genome, transcript and protein sequence, annotation and a data report. [Learn more](#) Download Datasets

Find related data Database: Select Find items

Search details "Homo sapiens"[Organism] AND ((latest[filter] OR "latest genbank"[filter]) AND all[filter] NOT anomalous[filter]) Search See more...

Recent activity Turn Off Clear

Homo sapiens genome AND ((latest[filter] OR "latest genbank"[filt... (903) Assembly

GRCh38.p13 - Genome - Assembly - NCBI Assembly

Референсный геном человека

Файл Правка Вид Журнал Закладки Инструменты Справка

GenBank - Поиск в Google x GRCh38.p13 - Genome - A... x +

https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.39

NCBI Resources How To Sign in to NCBI

Assembly Assembly Search Help

Advanced Browse by organism

Full Report v Send to: v

GRCh38.p13

Description: Genome Reference Consortium Human Build 38 patch release 13 (GRCh38.p13)

Organism name: [Homo sapiens \(human\)](#)

BioProject: [PRJNA31257](#)

Submitter: Genome Reference Consortium

Date: 2019/02/28

Assembly type: haploid-with-alt-loci

Release type: patch

Assembly level: Chromosome

Genome representation: full

RefSeq category: reference genome

GenBank assembly accession: GCA_000001405.28 (latest)

RefSeq assembly accession: GCF_000001405.39 (latest)

RefSeq assembly and GenBank assembly identical: no ([hide details](#))

- Only in GenBank: 1 unplaced scaffold (in primary assembly-unit)
- Data displayed for RefSeq version

IDs: 2334371 [UID] 8687898 [GenBank] 8765528 [RefSeq]

History ([Show revision history](#))

Comment

See [Genome](#) information for **Homo sapiens**

There are 917 assemblies for this organism
[See more](#)

Download Assembly

Access the data

- Genome Data Viewer
- RefSeq Annotation Report
- BLAST the assembly
- Full sequence report
- Statistics report
- Regions report
- FTP directory for RefSeq assembly

Assembly information

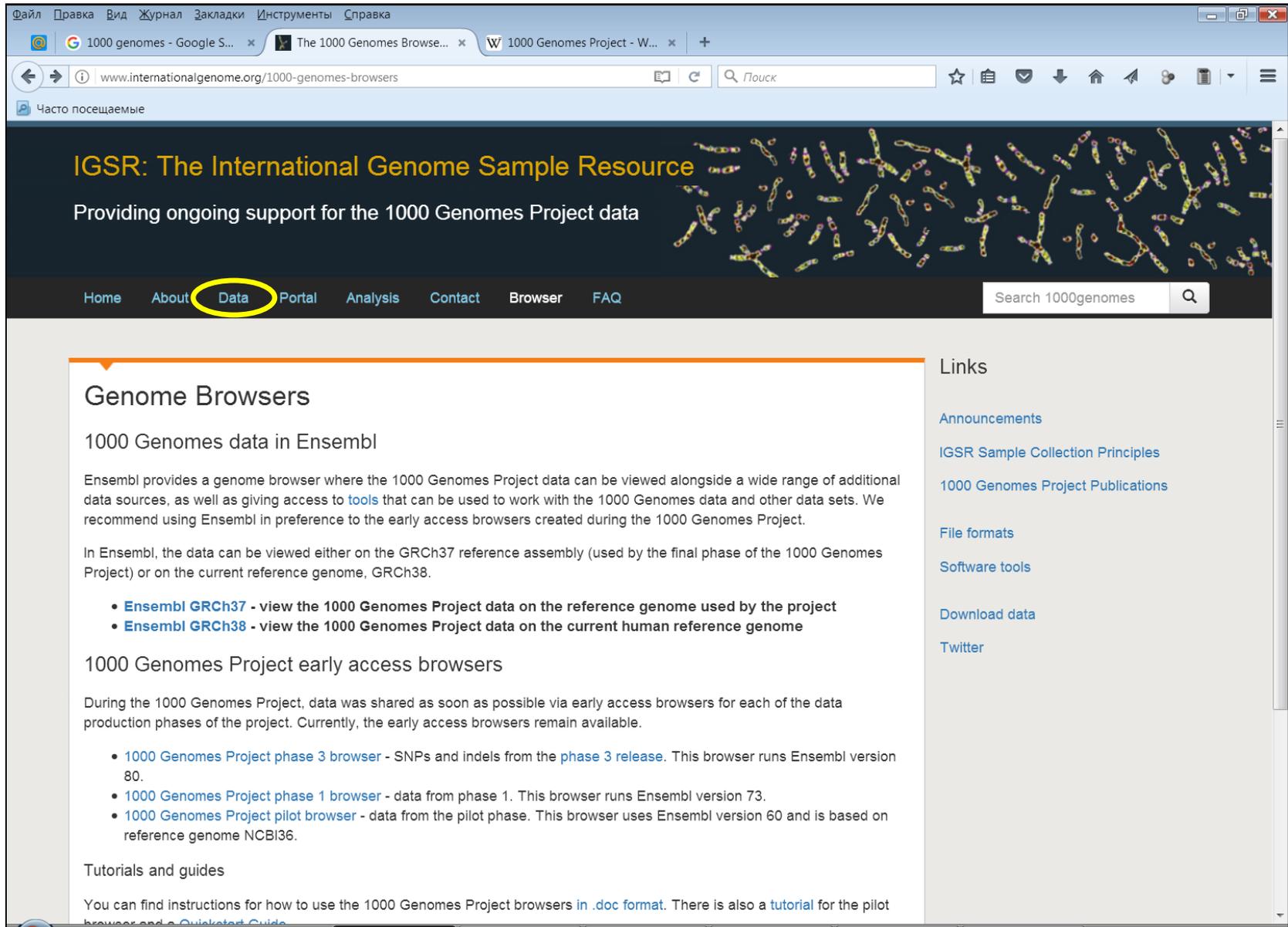
- Assembly Help
- Assembly Basics
- NCBI Assembly Data Model

https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.39

Всего доступно
917 различных сборок

Сборка GRCh38.p13 — февраль 2019
Genome Reference Consortium Human Build 38 patch release 13

1000 Genomes



Файл Правка Вид Журнал Закладки Инструменты Справка

1000 genomes - Google S... The 1000 Genomes Browse... 1000 Genomes Project - W... +

www.internationalgenome.org/1000-genomes-browsers Поиск

Часто посещаемые

IGSR: The International Genome Sample Resource

Providing ongoing support for the 1000 Genomes Project data

Home About **Data** Portal Analysis Contact Browser FAQ

Search 1000genomes

Genome Browsers

1000 Genomes data in Ensembl

Ensembl provides a genome browser where the 1000 Genomes Project data can be viewed alongside a wide range of additional data sources, as well as giving access to [tools](#) that can be used to work with the 1000 Genomes data and other data sets. We recommend using Ensembl in preference to the early access browsers created during the 1000 Genomes Project.

In Ensembl, the data can be viewed either on the GRCh37 reference assembly (used by the final phase of the 1000 Genomes Project) or on the current reference genome, GRCh38.

- [Ensembl GRCh37](#) - view the 1000 Genomes Project data on the reference genome used by the project
- [Ensembl GRCh38](#) - view the 1000 Genomes Project data on the current human reference genome

1000 Genomes Project early access browsers

During the 1000 Genomes Project, data was shared as soon as possible via early access browsers for each of the data production phases of the project. Currently, the early access browsers remain available.

- [1000 Genomes Project phase 3 browser](#) - SNPs and indels from the [phase 3 release](#). This browser runs Ensembl version 80.
- [1000 Genomes Project phase 1 browser](#) - data from phase 1. This browser runs Ensembl version 73.
- [1000 Genomes Project pilot browser](#) - data from the pilot phase. This browser uses Ensembl version 60 and is based on reference genome NCBI36.

Tutorials and guides

You can find instructions for how to use the 1000 Genomes Project browsers [in .doc format](#). There is also a [tutorial](#) for the pilot browser and a [Quickstart Guide](#).

Links

- [Announcements](#)
- [IGSR Sample Collection Principles](#)
- [1000 Genomes Project Publications](#)
- [File formats](#)
- [Software tools](#)
- [Download data](#)
- [Twitter](#)

1000 Genomes — Data

The screenshot shows the IGSR website interface. At the top, there's a navigation bar with links: Home, About, Data, Portal, Analysis, Contact, Browser, and FAQ. A search bar is on the right. The main content area features a section titled "Using data from IGSR" with a sub-section "Data portal beta". Below this is a table of genomic data. The "Population" column in the table is circled in red. To the right of the table is a "Links" sidebar with various resource links.

IGSR: The International Genome Sample Resource
Providing ongoing support for the 1000 Genomes Project data

Home About Data Portal Analysis Contact Browser FAQ

Search 1000genomes

Using data from IGSR

IGSR provides open data to support the community's research efforts. You can see our terms of use in our [data disclaimer](#).

Data portal *beta*

We are developing a [new data portal](#) to make it easier to find and browse data in IGSR. You can use the development version to [explore the data set](#). Let us know what you think at info@1000genomes.org.

| Sample | Sex | Population | Exome | Low cov WLL | High cov WLL | HD genotyp | Complete |
|---------|--------|------------|-------|-------------|--------------|------------|----------|
| HG00513 | Female | CHS | • | • | • | • | |
| HG01112 | Male | CLM | • | • | • | • | |
| HG00759 | Female | CDX | • | • | • | • | |
| HG01500 | Male | IBS | • | • | • | • | |
| HG03006 | Male | BEB | • | • | • | | |
| NA18525 | Female | CHB | • | • | • | • | |
| NA19648 | Female | MXL | • | • | • | • | |
| HG00096 | Male | GBR | • | • | • | • | |
| HG00419 | Female | CHS | • | • | • | • | |

Links

- [Announcements](#)
- [IGSR Sample Collection Principles](#)
- [1000 Genomes Project Publications](#)
- [File formats](#)
- [Software tools](#)
- [Download data](#)
- [Twitter](#)

1000 Genomes — Population

The screenshot shows the IGSR website interface. At the top, there's a navigation bar with links for Home, About, Data, Portal, Analysis, Contact, Browser, and FAQ. A search bar is also present. Below the navigation, there are tabs for Samples, Populations, and Data collections. The 'Samples' tab is active, and a dropdown menu for 'Filter by population' is open, showing a list of populations with 'GBR' selected. The main content area displays a table of samples, with the first two rows highlighted. The word 'British' is overlaid in large purple text on the table.

IGSR: The International Genome Sample Resource
Providing ongoing support for the 1000 Genomes Project data

Home About Data Portal Analysis Contact Browser FAQ Search 1000genomes

Samples Populations Data collections Search

Filter by population Filter by analysis group Filter by data collection Toggle table view Download the list

Filter by population

- ACB
- ASW
- BEB
- CDX
- CEU
- CHB
- CHD
- CHS
- CLM
- ESN
- FIN
- GBR
- GIH
- GWD
- GWF
- GWJ
- GWW
- IBS
- ITU
- JPT
- KHV
- LWK
- MSL
- MXL
- PEL
- PJL
- PUR
- STU
- TSI
- YRI

Filters: GBR

Showing 1 to 50 of 107 samples

« Previous Next »

| Sample | Sex | Population | GRCm3 | Phase 2 | Phase 1 | Platinum pedigree | Structural variation | Gambian variation | 90 Han Chinese | Geuvadis |
|---------|--------|------------|-------|---------|---------|-------------------|----------------------|-------------------|----------------|----------|
| HG00099 | Female | GBR | . | . | . | | | | . | |
| HG00102 | Female | GBR | . | . | . | | | | . | |

British

1000 Genomes — Sample HG00099

The screenshot shows a web browser window with the URL www.internationalgenome.org/data-portal/sample/HG00099. The page header features the IGSR logo and the text "IGSR: The International Genome Sample Resource" and "Providing ongoing support for the 1000 Genomes Project data". A navigation menu includes "Home", "About", "Data", "Portal", "Analysis", "Contact", "Browser", and "FAQ". A search bar is located in the top right corner.

The main content area is titled "Sample HG00099" (circled in red) and includes a "Data portal beta" badge. It is divided into two columns: "HG00099 details" and "GBR population".

HG00099 details:

- Sex: Female
- Biosample ID: [SAME123271](#)
- Search Coriell: [HG00099](#)

GBR population:

- Population: [British](#)
- Code: GBR
- Description: British in England and Scotland
- Superpopulation: European
- Superpopulation code: EUR

Below the details is a section titled "Data collections for HG00099" with tabs for "1000 Genomes on GRCh38", "1000 Genomes phase 3 release", "1000 Genomes phase 1 release", and "Geuvadis". The "1000 Genomes on GRCh38" tab is active, showing "12 matching data files" and a "Download the list" button. A "Data reuse policy for 1000 Genomes on GRCh38" link is also present.

A table of data files is partially visible at the bottom, with columns for "Data types", "File", and "Analysis group". One file is listed: [ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR741/SRR741411/SRR741411_2.fastq.gz](#) with the analysis group "Low coverage".

1000 Genomes — Exome

Sample HG00099

HG00099 details

Sex: Female
Biosample ID: SAME123271
Search Coriell: HG00099

GBR population

Population: British
Code: GBR
Description: British in England and Scotland
Superpopulation: European
Superpopulation code: EUR

Data collections for HG00099

1000 Genomes on GRCh38 | 1000 Genomes phase 3 release | 1000 Genomes phase 1 release | Geuvadis

Data reuse policy for 1000 Genomes on GRCh38 | 2 matching data files | Download the list

Data types

- Alignment
- Sequence

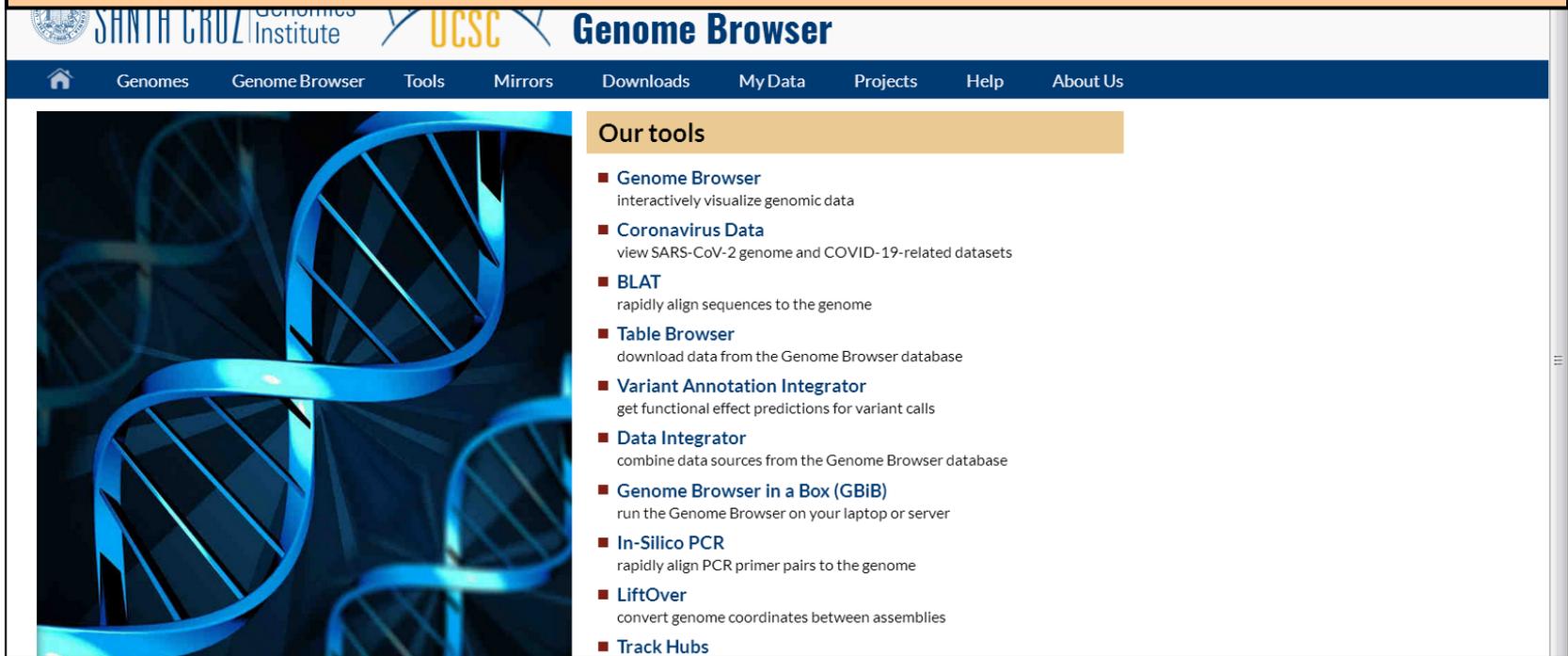
Analysis groups

- Exome
- Low coverage WGS

| File | Analysis group |
|---|----------------|
| ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR765/SRR765993/SRR765993_2.fastq.gz | Exome |
| ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR765/SRR765993/SRR765993_1.fastq.gz | Exome |

Геномные браузеры

UCSC Genome Browser



The screenshot shows the UCSC Genome Browser homepage. At the top, there is a navigation bar with links for Genomes, Genome Browser, Tools, Mirrors, Downloads, MyData, Projects, Help, and About Us. Below the navigation bar, there is a large image of a DNA double helix on the left. To the right of the image, there is a section titled "Our tools" with a list of tools and their descriptions:

- Genome Browser**
interactively visualize genomic data
- Coronavirus Data**
view SARS-CoV-2 genome and COVID-19-related 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**

<http://genome.ucsc.edu/>

Our story

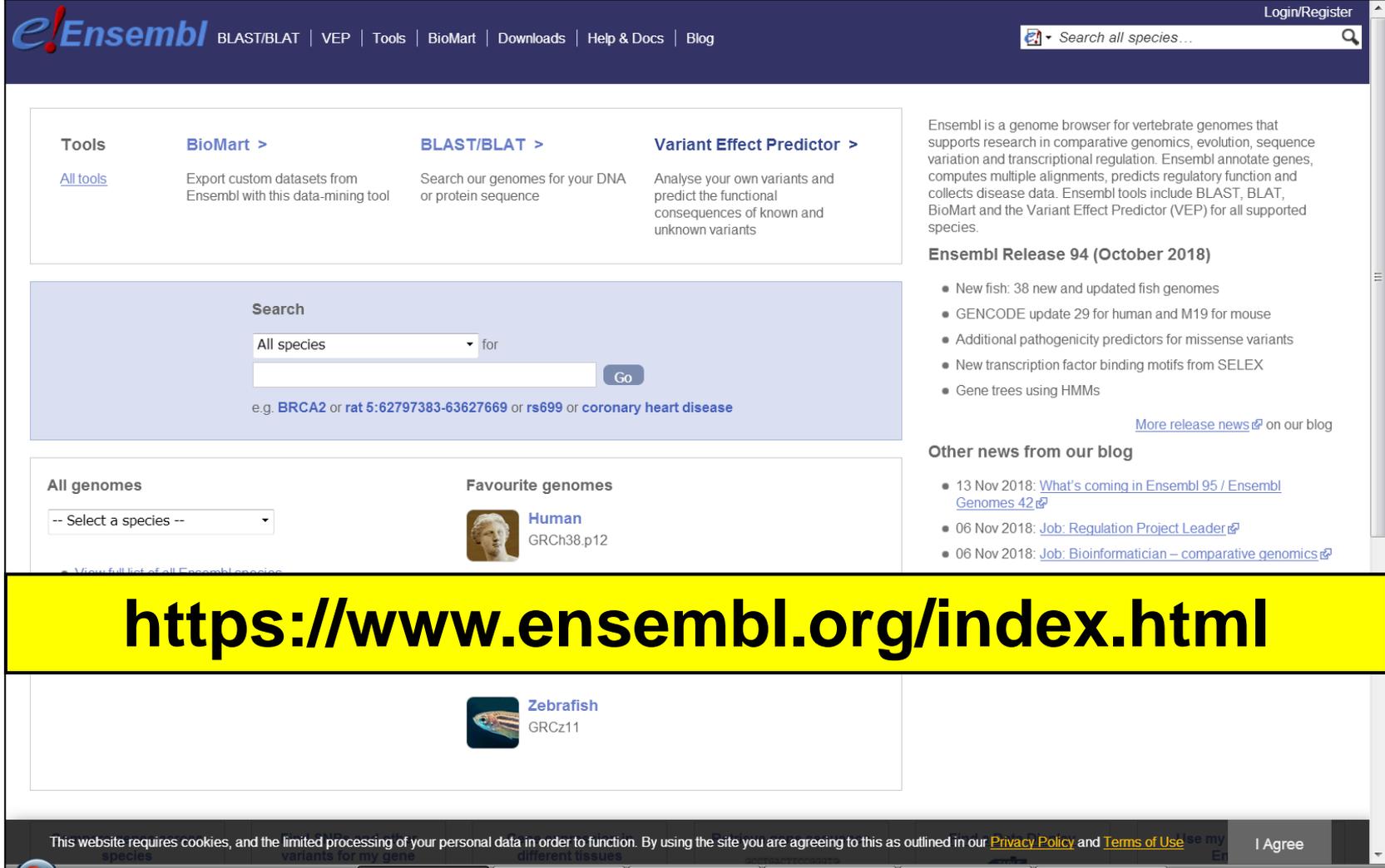
On June 22, 2000, UCSC and the other members of the International Human Genome Project consortium completed the first working draft of the human genome assembly, forever ensuring free public access to the genome and the information it contains. A few weeks later, on July 7, 2000, the newly assembled genome was released on the web at <http://genome.ucsc.edu>, along with the initial prototype of a

What's new

Nov 3, 2020 - [New gnomAD pext track for GRCh37/hg19](#)
Oct 23, 2020 - [New GRCm39 Mouse Genome Browser is here](#)
Oct 21, 2020 - [Oct 21st SARS-CoV-2 release](#)
[More news...](#)

Геномные браузеры

Ensembl Genome Browser



The screenshot shows the Ensembl Genome Browser homepage. At the top, there is a navigation bar with the Ensembl logo and links for BLAST/BLAT, VEP, Tools, BioMart, Downloads, Help & Docs, and Blog. A search bar is located on the right side of the navigation bar. Below the navigation bar, there is a main content area with several sections:

- Tools:** A section with four columns: "All tools", "BioMart >" (Export custom datasets from Ensembl with this data-mining tool), "BLAST/BLAT >" (Search our genomes for your DNA or protein sequence), and "Variant Effect Predictor >" (Analyse your own variants and predict the functional consequences of known and unknown variants).
- Search:** A search box with a dropdown menu set to "All species" and a "Go" button. Below the search box, there is a text input field and a "Go" button. Below the text input field, there is a text input field with the text "e.g. BRCA2 or rat 5:62797383-63627669 or rs699 or coronary heart disease".
- All genomes:** A dropdown menu with the text "-- Select a species --".
- Favourite genomes:** A section with a small image of a classical bust and the text "Human GRCh38.p12".
- Ensembl Release 94 (October 2018):** A section with a list of bullet points: "New fish: 38 new and updated fish genomes", "GENCODE update 29 for human and M19 for mouse", "Additional pathogenicity predictors for missense variants", "New transcription factor binding motifs from SELEX", and "Gene trees using HMMs". Below the list, there is a link "More release news" on our blog.
- Other news from our blog:** A section with a list of bullet points: "13 Nov 2018: What's coming in Ensembl 95 / Ensembl Genomes 42", "06 Nov 2018: Job: Regulation Project Leader", and "06 Nov 2018: Job: Bioinformatician – comparative genomics".
- Zebrafish:** A section with a small image of a zebrafish and the text "Zebrafish GRCz11".

At the bottom of the page, there is a yellow banner with the URL <https://www.ensembl.org/index.html>. Below the banner, there is a footer with a cookie notice: "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](#)." and a link "I Agree".

Геномные браузеры

Vista

The screenshot displays the VISTA web interface. At the top, there is a navigation menu with the following items: VISTA Home, Custom Alignment, **Browser** (highlighted), Enhancer DB, Downloads, Publications, and Help. To the right of the menu are links for About Us, Cite Us, and Contact Us. Below the menu is a search form with the following fields: Clade (Vertebrate), Genome (Human), Release (Feb. 2009), and Position (chr9:107,543,285-107,690,436). There is a Submit button next to the Position field. Below the form are two radio buttons: VISTA-Point (selected) and VISTA Browser (Requires Java2). A Help link is also present. Below the form, the text reads: "Whole Genome Comparative Analysis of the Human Feb. 2009 Genome". It states: "From this page you can access the results of: the multiple alignments" and lists: "Human Feb. 2009, Chimp Mar. 2006, Callithrix jacchus v.2.0.2, Rhesus Jan. 2006 and Pongo pygmaeus abelii v.2.0.2 genomes." It also mentions: "the pairwise alignments of the Human Feb. 2009 genome produced by the Genome Reference Consortium with the following genomes:" and lists: "The Chimp Feb. 2011 genome produced by the Chimpanzee Sequencing and Analysis Consortium", "The Cow Oct. 2011 genome produced by the Baylor College of Medicine -- Cow", and "The Gorilla Dec. 2009 genome produced by The Wellcome Trust Sanger Institute". At the bottom, it says: "Annotation for this genome comes from the UCSC." and provides a link for "Methods and references".

<http://pipeline.lbl.gov/cgi-bin/gateway2>

Геномные браузеры

GBrowse



[Log in / create account](#)

Page [Discussion](#)Read [View form](#) [View source](#) [View history](#)

GBrowse



Contents [\[hide\]](#)

- 1 About Generic Genome Browser (GBrowse)
 - 1.1 Screenshots
- 2 Downloads
- 3 Using GBrowse
 - 3.1 System Requirements
 - 3.2 Installation
 - 3.3 Configuration
- 4 Documentation
 - 4.1 POD documentation
 - 4.2 Installation
 - 4.3 Configuration
 - 4.4 Documentation
 - 4.5 RSS

Status

Mature release
Development: active
Support: active

Licence

 [GPL 2.0](#) [Artistic License 2.0](#)

Resources

[Home page](#)

What links here

- 7 GBrowse in the wild
- 8 See also
- 9 More on GBrowse
 - 9.1 GBrowse Logo

About Generic Genome Browser (GBrowse)

GBrowse is a combination of database and interactive web pages for manipulating and displaying annotations on genomes. Features include:

http://gmod.org/wiki/GBrowse

What links here

- Related changes
- Special pages
- Printable version
- Permanent link
- Page information
- Browse properties
- Print as PDF

What links here

- 7 GBrowse in the wild
- 8 See also
- 9 More on GBrowse
 - 9.1 GBrowse Logo

About Generic Genome Browser (GBrowse)

GBrowse is a combination of database and interactive web pages for manipulating and displaying annotations on genomes. Features include:

UCSC Genome Browser

Уайл Правка Вид Журнал Закладки Инструменты Справка

GenBank - Поиск в Google UCSC Genome Browser Ho... +

genome.ucsc.edu Поиск

UNIVERSITY OF CALIFORNIA SANTA CRUZ Genomics Institute UCSC Genome Browser

Genomes Genome Browser Tools Mirrors Downloads MyData Projects Help About Us

- Human GRCh38/hg38
- Human GRCh37/hg19
- Mouse GRM39/mm39
- Mouse GRM38/mm10
- Mouse: 16 strains
- SARS-CoV-2 (COVID-19)
- Other

Our tools

- Genome Browser**
interactively visualize genomic data
- Coronavirus Data**
view SARS-CoV-2 genome and COVID-19-related 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**

Загружаем референсный геном

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 [first assembled genome](#) was released on the web at [genome.ucsc.edu/cgi-bin/hgGateway](#) long with the initial prototype of a

What's new

- Nov 3, 2020 - [New gnomAD pext track for GRCh37/hg19](#)
- Oct 23, 2020 - [New GRM39 Mouse Genome Browser is here](#)
- Oct 21, 2020 - [Oct 21st SARS-CoV-2 release](#)
- [More news...](#)

UCSC Genome Browser

UCSC Genome Browser on Human Dec. 2013 (GRCh38/hg38) Assembly

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

chr1:11,102,837-11,267,747 164,911 bp. enter position, gene symbol, HGVS or search terms go

chr1 (p36.22) 1p31.1 1q12 1q41 1q43.44

Scale chr1: 11,150,000 50 kb hg38 11,200,000 11,250,000

MTOR MTOR New haplotype sequences to GRCh38 Reference Sequence Patches to GRCh38 Reference Sequence Comprehensive Transcript Set (only Basic displayed by default) RNU6-291P RNU6-537P RNU6-537P

RefSeq Curated RefSeq gene predictions from NCBI OMIM Allelic Variants

OMIM Alleles gene Expression in 53 tissues from GTEx RNA-seq of 8555 samples (570 donors)

MTOR HSK27Ac Mark (Often Found Near Regulatory Elements) on 7 cell lines from ENCODE

Layered H3K27Ac DNase I Hypersensitivity Peak Clusters from ENCODE (95 cell types) 100 vertebrates Basewise Conservation by PhyloP

Cons 100 Verts Multiz Alignments of 100 Vertebrates

Rhesus

Click on a feature for details. Click or drag in the base position track to zoom in. Click side bars for track options. Drag side bars or labels up or down to reorder tracks. Drag tracks left or right to new position. Press "?" for keyboard shortcuts.

move start < 2.0 > move end < 2.0 >

track search default tracks default order hide all add custom tracks track hubs configure multi-region reverse resize refresh

Хромосома 1, locus p36.22

UCSC Genome Browser

UCSC Genome Browser on Human Dec. 2013 (GRCh38/hg38) Assembly

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

chr1:149,806,988-149,971,898 164,911 bp. enter position, gene symbol, HGVS or search terms go

chr1 (q21.2) 1p31.1 1q12 1q41 1q43.44

Scale chr1: 149,850,000 | 50 kb | 149,900,000 | hg38 149,950,000

GENCODE v24 Comprehensive Transcript Set (only Basic displayed by default)
Patches to GRCh38 Reference Sequence

RefSeq Curated OMIM Alleles

Gene Expression in 53 tissues from G-Ex RNA-seq of 8555 samples (570 donors)

100 vertebrates Base-wise Conservation by PhyloP

Multiz Alignments of 100 Vertebrates

Simple Nucleotide Polymorphisms (dbSNP 150) Found in >= 12 of Samples

Repeat Elements by RepeatMasker

Хромосома 1, локус q21.2

UCSC Genome Browser

Файл Правка Вид Журнал Закладки Инструменты Справка

обозначение генов в хро... x Локус — Википедия x Human hg38 chr1:10,965,9... x +

genome.ucsc.edu/cgi-bin/hgTracks?db=hg38&lastVirtModeType=default&lastVirtModeExtraState=

Часто посещаемые

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

UCSC Genome Browser on Human Dec. 2013 (GRCh38/hg38) Assembly

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

chr1:10,965,907-11,130,817 164,911 bp. chr6 go

chr1 (p36.22) 1p31.1 1q12 1q41 1q43.44

Scale chr1: 11,000,000| 11,050,000| 11,100,000| hg38

New haplotype sequences to GRCh38 Reference Sequence
Patches to GRCh38 Reference Sequence
Comprehensive Transcript Set (only Basic displayed by default)

Gene Expression in 53 tissues from GTEx RNA-seq of 8555 samples (570 donors)

Cons 100 Verts

Multiz Alignments of 100 Vertebrates

Simple Nucleotide Polymorphisms (rsSNP 150) Found in >= 1% of Samples

Repeating Elements by RepeatMasker

SINE LINE LTR DNA

Переход на хромосому 6

UCSC Genome Browser

The screenshot displays the UCSC Genome Browser interface for Human hg38. The browser window shows the URL `genome.ucsc.edu/cgi-bin/hgTracks?db=hg38&lastVirtModeType=default&lastVirtModeExtraState=&virt`. The page title is "UCSC Genome Browser on Human Dec. 2013 (GRCh38/hg38) Assembly". The main content area shows a genomic track for chromosome 6, with a scale from 50,000,000 to 150,000,000 bp. The track displays various annotations, including KQ090817.1, KQ031367.1, KN196478.1, and KV765194.1. A yellow banner at the bottom of the browser window reads "Вся хромосома 6".

File Правка Вид Журнал Закладки Инструменты Справка

обозначение генов в хро... x Локус — Википедия x Human hg38 chr6:1-17080... x +

genome.ucsc.edu/cgi-bin/hgTracks?db=hg38&lastVirtModeType=default&lastVirtModeExtraState=&virt Поиск

Часто посещаемые

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

UCSC Genome Browser on Human Dec. 2013 (GRCh38/hg38) Assembly

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

chr6:1-170,805,979 170,805,979 bp. chr6 go

chr6 (p25.3-q27) 6p22.3 6q12 6q13 6q14.1 6q15 6q16.1 6q21 6q22 6q23

Scale chr6: 50,000,000 50 Mb hg38 100,000,000 150,000,000

New haplotype sequences to GRCh38 Reference Sequence

Patches to GRCh38 Reference Sequence

KQ090817.1 KQ031367.1 KN196478.1 KV765194.1

GENCODE v24 Comprehensive Transcript Set (only Basic displayed by default)

RefSeq Curated RefSeq gene predictions from NCBI

OMIM Allelic Variants

UCSC Genome Browser

The screenshot displays the UCSC Genome Browser interface. At the top, the browser window title is "Human hg38 chr6:2-170,80...". The address bar shows the URL "genome.ucsc.edu/cgi-bin/hgTracks?db=hg38&lastVirtModeType=default&lastVirtModeExtraState=". The navigation bar includes "Genomes", "Genome Browser", "Tools", "Mirrors", "Downloads", "My Data", "View", "Help", and "About Us".

The main heading is "UCSC Genome Browser on Human Dec. 2013 (GRCh38/hg38) Assembly". Below this, there are navigation buttons: "move", "zoom in", "zoom out", and "base". The "zoom in" buttons are labeled "1.5x", "3x", "10x", and "100x". The "zoom out" buttons are labeled "1.5x", "3x", "10x", and "100x".

The search bar contains the coordinates "chr6:2-170,805,979 170,805,978 b" and the band "6p21.1". A "go" button is located to the right of the search bar. The "6p21.1" text and the "go" button are circled in red and blue, respectively.

Below the search bar is a genomic track for chromosome 6, labeled "chr6 (p25.3-q27)". The track shows a scale from 50,000,000 to 150,000,000. The track displays various annotations, including "New haplotype sequences to GRCh38 Reference Sequence", "Patches to GRCh38 Reference Sequence", and "GENCODE v24 Comprehensive Transcript Set (only Basic displayed by default)".

A yellow banner at the bottom of the image contains the text "Переход на локус 6p21.1".

UCSC Genome Browser

Файл Правка Вид Журнал Закладки Инструменты Справка

Uniprot - Google Search x dpp4 in UniProtKB x Локус — Википедия x Human hg38 chr2:1619922... x +

genome.ucsc.edu/cgi-bin/hgTracks?db=hg38&lastVirtModeType=default&lastVirtModeExtraState=&virt Поиск

Часто посещаемые

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

UCSC Genome Browser on Human Dec. 2013 (GRCh38/hg38) Assembly

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

chr2:161,992,241-162,074,542 82,302 bp. enter position, gene symbol, HGVS or search terms go

chr2 (q24.2) 21 14 p12 11.2 q34 q35

Scale chr2: 162,000,000 162,010,000 162,020,000 162,030,000 162,040,000 162,050,000 162,060,000 162,070,000 hg38

New haplotype sequences to GRCh38 Reference Sequence
Patches to GRCh38 Reference Sequence
Chromosome Bands Localized by FISH Mapping Clones
2q24.2

GENCODE v24 Comprehensive Transcript Set (only Basic displayed by default)
RefSeq Curated RefSeq gene predictions from NCBI
OMIM Alleles OMIM Allelic Variants
Gene Expression in 53 tissues from GTEx RNA-seq of 8555 samples (570 donors)
DPP4 AC008063.2 (4)

Layered H3K27Ac H3K27Ac Mark (Often Found Near Regulatory Elements) on 7 cell lines from ENCODE
AC008063.2

DNase I Hypersensitivity Peak Clusters from ENCODE (95 cell types)
100 vertebrates Basewise Conservation by PhyloP

Cons 100 Verts
-4.5
Multiz Alignments of 100 Vertebrates
Rhesus
Mouse
Dog
Elephant
Chicken
X_tropicalis
Zebrafish

Ген DPP4 — локус 2q24.2

Satellite
RNA
Other
Unknown

move start < 2.0 > move end < 2.0 >

track search default tracks default order hide all add custom tracks track hubs configure multi-region reverse resize refresh

collapse all

Use drop-down controls below and press refresh to alter tracks displayed.
Tracks with lots of items will automatically be displayed in more compact modes. expand all

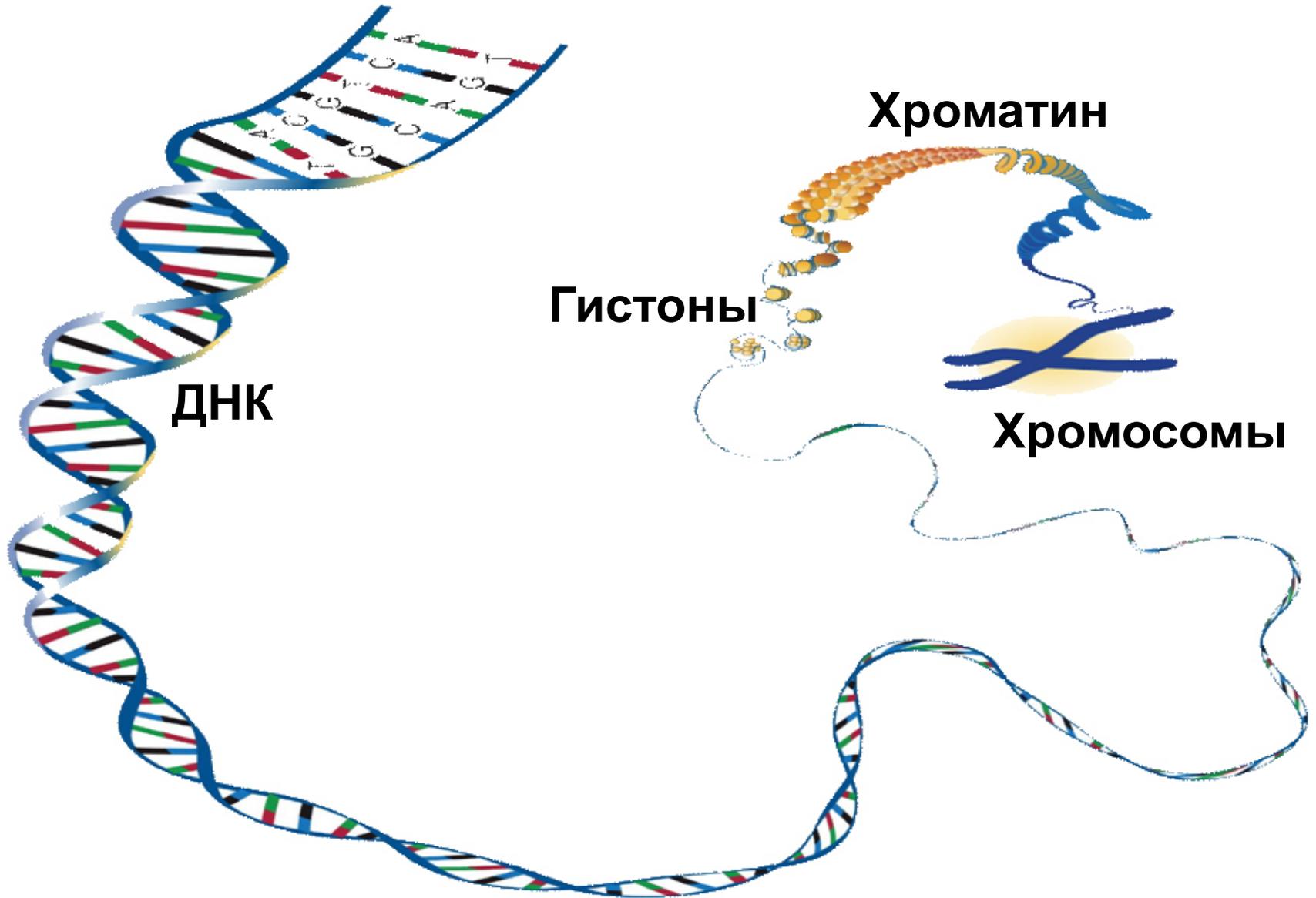
Определения

- Лocus — местоположение определённого гена на генетической или цитологической карте хромосомы
- Аллель — вариант последовательности ДНК в данном локусе
- Генетическая карта — упорядоченный перечень локусов генома
- Генное картирование — определение локуса для специфического биологического признака

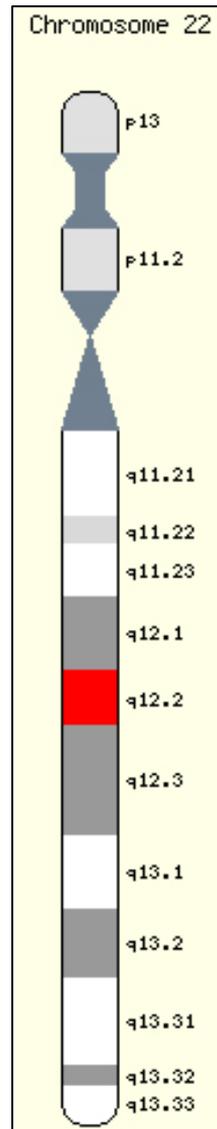
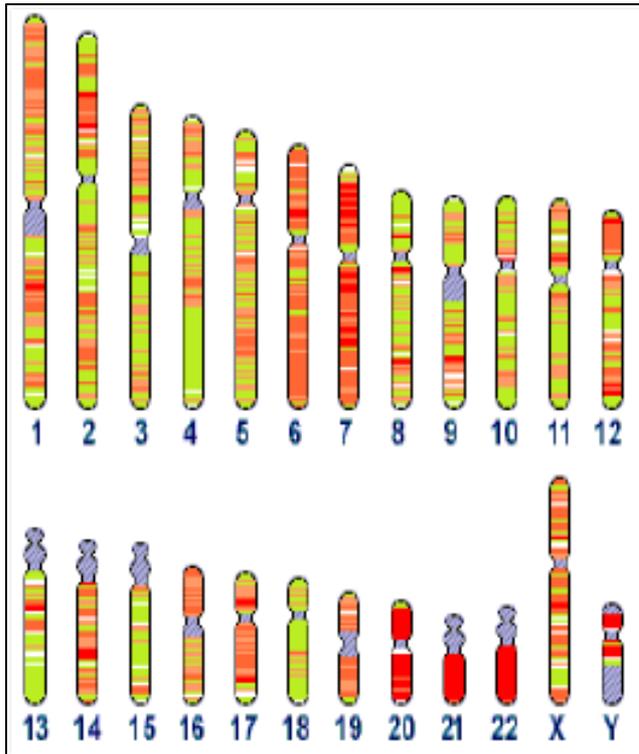
Аннотация генома

- Описание функциональных и структурных характеристик генома
- Местонахождение кодирующих участков генов в геноме, регуляторных элементов, регулирующих транскрипцию, другие функции генома и особенности функционирования генома
- Тканеспецифичность экспрессии, профили экспрессии и коэкспрессию генов

Структура генома



Обозначения локусов



6p21.3

6 – номер хромосомы

p – плечо хромосомы:

p – короткое

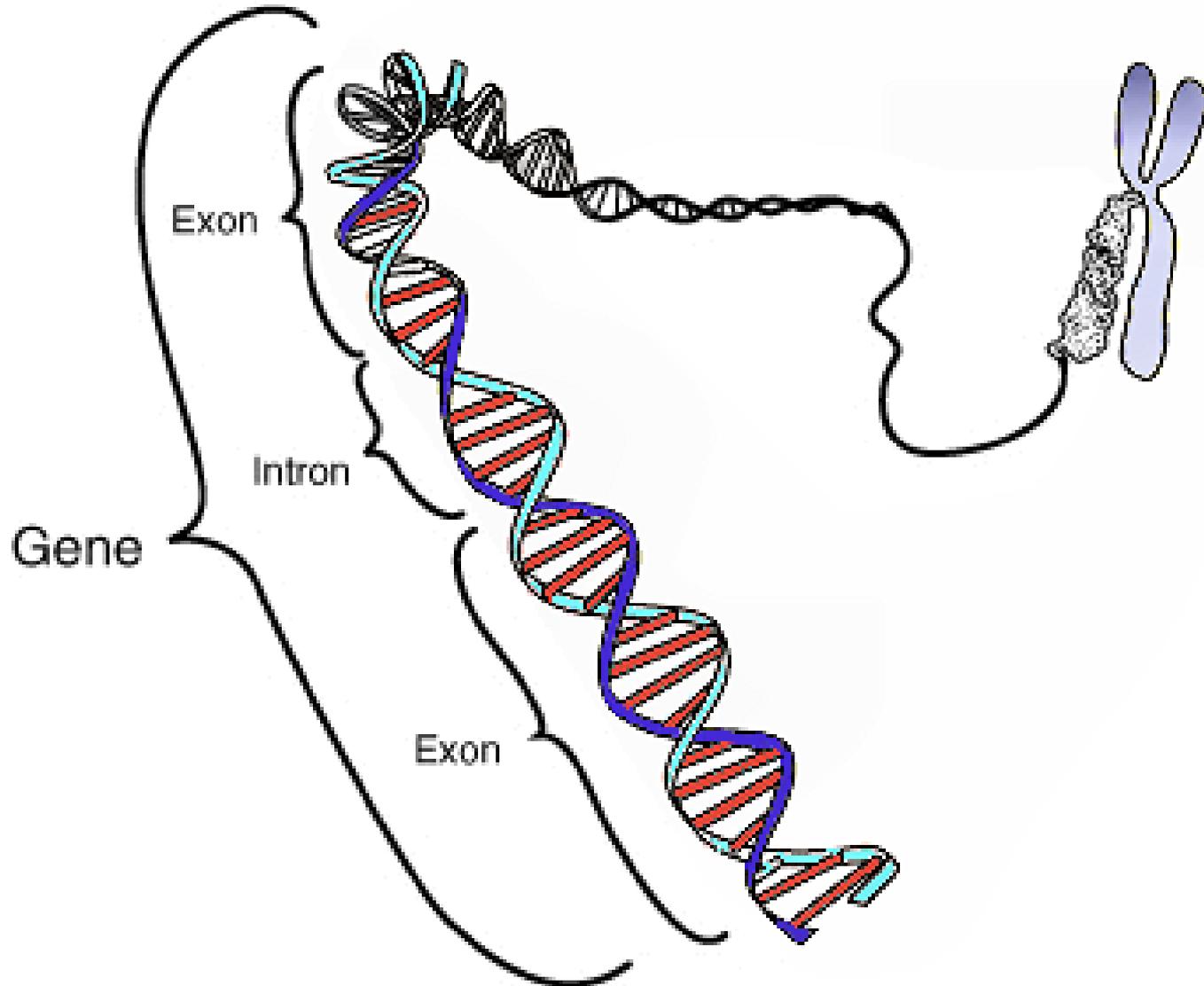
q – длинное

2 – область

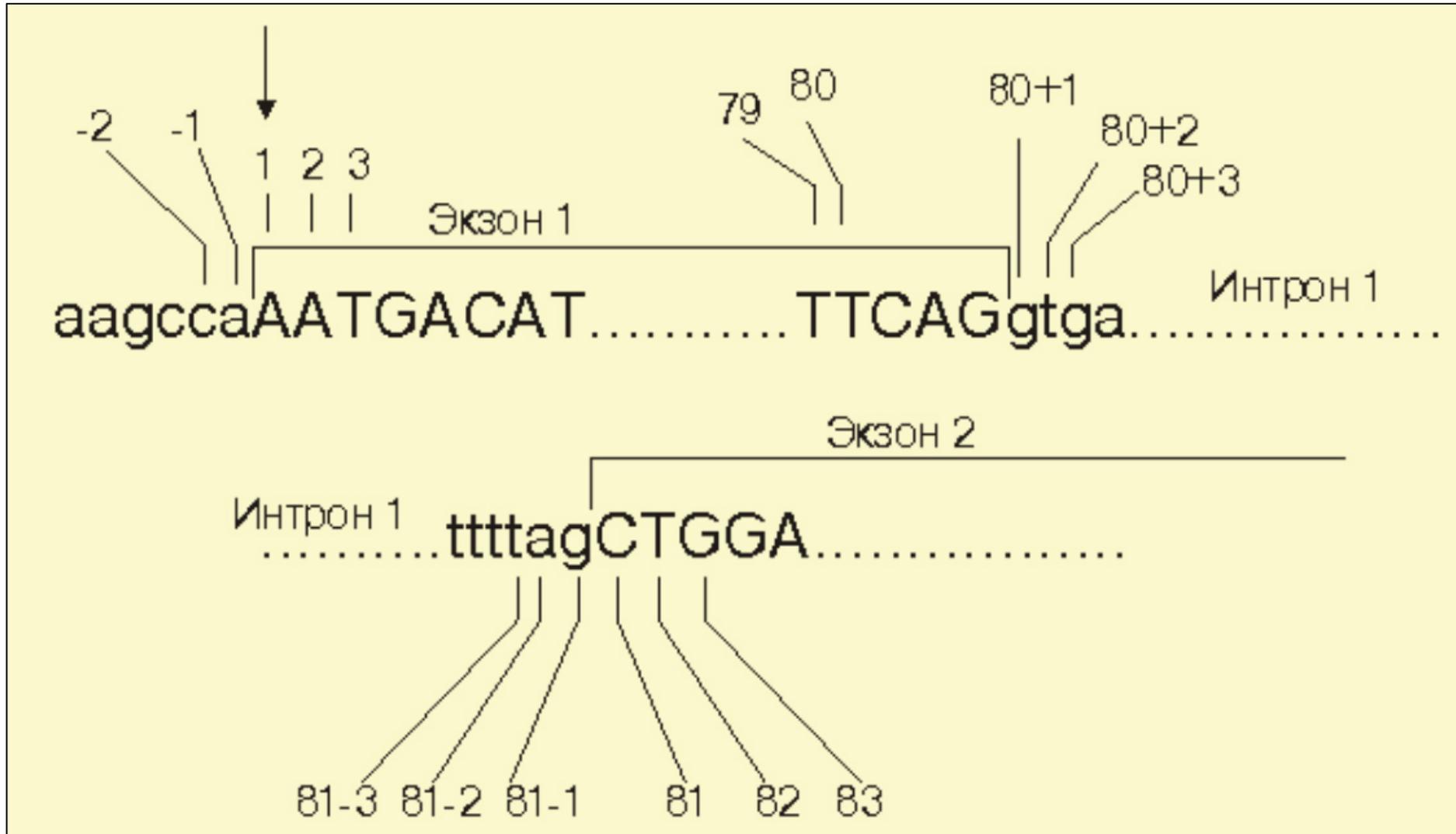
1 – бэнд

.3 – суб-бэнд

Структура генома



Нумерация нуклеотидов в гене



Характеристики генома

- 24 хромосомы
- ~3.2 млрд. пар нуклеотидов
- ~21 тыс. генов
- Один ген — 1000-3000 оснований
- Два генома идентичны на ~99.9 %
- Фенотипические признаки и все заболевания — ~0.1 % генома, т.е. его 1/1000 мутированная часть

To be continued ...

