



NCBI Gene

A portal to gene-centered information from different sources

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

National Center for Biotechnology Information • National Library of Medicine • National Institutes of Health • Department of Health and Human Services

Scope and Access

The Gene database [1, 2] provides detailed information for known and predicted genes defined by nucleotide sequence or map position. Currently, Gene contains over 17 million entries and includes data from all major taxonomic groups. Each record in the database corresponds to a single gene and is derived from processing by the NCBI Reference Sequence [3, 4] and genome annotation groups [5]. You can access data from the Gene database on the web through the Gene homepage, programmatically through the Entrez Programming Utilities, or by file transfer through its FTP site.

- Gene Homepage: www.ncbi.nlm.nih.gov/gene/
 - Entrez Programming Utilities: www.ncbi.nlm.nih.gov/books/NBK25501
 - Gene Data Files: ftp.ncbi.nlm.nih.gov/gene/DATA/
 - Gene Records in ASN.1 format: ftp.ncbi.nlm.nih.gov/gene/DATA/ASN_BINARY/
- You can obtain technical details on the Gene database from the NCBI Bookshelf:
- Gene Help: www.ncbi.nlm.nih.gov/books/NBK3841/
 - Gene FAQ: www.ncbi.nlm.nih.gov/books/NBK3840/

The screenshot shows the NCBI Gene database search results for the query 'hemochromatosis'. The interface includes a search bar (A), a sidebar with filters (B), a search results table (E), and a 'Send to' dropdown menu (F) with a 'File' option selected. The table lists genes such as HFE, HJV, Hfe, and Tf, with columns for Name/Gene ID, Description, Location, Aliases, and MIM.

Name/Gene ID	Description	Location	Aliases	MIM
HFE ID: 3077	homeostatic iron regulator [<i>Homo sapiens</i> (human)]	Chromosome 6, NC_000006.12 (26087281..26096216)	HFE1, HH, HLA-H, MVCD7, TFQTL2	613609
HJV ID: 148738	hemojuvelin BMP co-receptor [<i>Homo sapiens</i> (human)]	Chromosome 1, NC_000001.11 (146017468..146021822, complement)	HFE2, HFE2A, JH, RGMC	608374
Hfe ID: 15216	homeostatic iron regulator [<i>Mus musculus</i> (house mouse)]	Chromosome 13, NC_000079.6 (23702730..23710986, complement)	MR2	
Hfe ID: 29199	homeostatic iron regulator [<i>Rattus norvegicus</i> (Norway rat)]	Chromosome 17, NC_005116.4 (43661276..43669327)		
Tf ID: 22041	transferrin [<i>Mus musculus</i> (house mouse)]	Chromosome 9, NC_000075.6 (103208876..103230286, complement)	AI266983, Cd176, HP, Tf, Tfn, hpx	

Searching Gene

Enter a set of query terms and click the **Search** button (A) to find gene records of interest. Select filters in the left column (B) to narrow down the retrieved list according to chosen criteria. Click the “SIDEBAR” (C) to expand and access functions listed in the right-hand column. Use the pull-down menus (D) at the top to change the display format or sorting order. From left to right, the default tabular display (E) provides information on the official symbol and NCBI Gene ID, gene name and source organism, chromosome location, alias symbols known to be associated with the gene, and MIM id’s if available. You can save retrieved Gene records in various formats using the “File” option available under the “Send to” link (F). Click the official symbol to open the full report of that Gene record.

Advanced Search Builder

This tool (right) makes indexed fields accessible through a pull-down menu under the arrow (H), and terms indexed through the **Show index list** link (I). Select a term to enter it in the query box above (J) with the selected Boolean operator. Click **Add to history** link (K) to preview and save a search to the History. Click the number under the **Items found** column (L) to retrieve these records. Unlock the query box by clicking the **Edit** link to customize the query terms, such as combining searches (M). With a My NCBI account, you can save your gene search results permanently and customize the display format using options under the “NCBI Site Preferences” [9].

The screenshot shows the NCBI Gene Advanced Search Builder interface. It includes a search box (M) with the query '#5 AND rodentia[organisms]', a pull-down menu (H) for 'genotype protein coding', a list of indexed terms (J) including 'genotype ncrna', 'genotype other', 'genotype protein coding', 'genotype pseudo', and 'genotype rna', a 'Show index list' link (I), a 'Search' button (K), and a 'History' table (L) showing previous searches.

Search	Add to builder	Query	Items found	Time
#5	Add	Search hemochromatosis Schema: base Filters: Current	386	09:00:25
#3	Add	Search alive[prop] Schema: base Filters: Current	779342	08:58:37

Contents of a Gene record

Information found in a Gene record depends on publicly available data. All records, like the well-studied human HFE gene (right), contain a **Summary** section (A), which provides an overview of the gene. The HFE gene record also contains links to corresponding records from HUGO Gene Nomenclature Committee (B), other external links such as OMIM [7] (C). It also provides a list of unofficial symbols associated with this gene (D) and a brief summary of the gene and its product (E). For ease of navigation, the **Table of Contents** (F) provides direct links to individual sections of a Gene report. For example, click **Bibliography** (G) to scroll the display to that section

where a collection of publications relevant to the Gene record is listed. Similarly, click the **Reference sequences** (H) to bring the NCBI Reference Sequences section to focus and access individual sequences relevant to this gene. Click the "all" link next to **Orthologs** (I) to see orthologs of the gene derived from NCBI Eukaryotic Genome Annotation pipeline.

Full Report Send to: ▾

HFE homeostatic iron regulator [*Homo sapiens* (human)]

Gene ID: 3077, updated on 12-Mar-2019

Summary

Official Symbol HFE provided by [HGNC](#)

Official Full Name homeostatic iron regulator provided by [HGNC](#)

Primary source [HGNC:HGNC:4886](#)

See related [Ensembl:ENSG0000010704](#) [MIM:613609](#)

Gene type protein coding

RefSeq status REVIEWED

Organism [Homo sapiens](#)

Lineage Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eumarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo

Also known as HH; HFE1; HLA-H; MVCD7; TFQTL2

Summary The protein encoded by this gene is a membrane protein that is similar to and associates with beta2-microglobulin (beta2M). It is thought that this protein is involved in iron absorption by regulating the interaction of the transferrin receptor with beta2-microglobulin. Defects in this gene, hereditary haemochromatosis, is a recessive genetic disorder. Additional variants have been found but their full-length nature has not been determined. [MIM:235200] (OMIM)

Expression Ubiquitous expression in thyroid (RPKM 5.0), gall bladder (RPKM 4.6) and

Orthologs [mouse](#) [all](#)

Table of contents

- Summary
- Genomic context
- Genomic regions, transcripts, and products
- Expression
- Bibliography
- Phenotypes
- Variation
- HIV-1 interactions
- Pathways from BioSystems
- Interactions
- General gene information
 - Markers, Clone Names, Homology, Gene Ontology
- General protein information
- NCBI Reference Sequences (RefSeq)
- Related sequences
- Additional links
 - Locus-specific Databases
- Genome Browsers**
 - Genome Data Viewer
 - Variation Viewer (GRCh37.p13)
 - Variation Viewer (GRCh38)
 - 1000 Genomes Browser (GRCh37.p13)
 - Ensembl
 - UCSC

Graphical presentation of the gene on the genome

The Sequence Viewer (SV) panel presents the RefSeqs for a Gene record in an interactive and customizable graphical display. This panel (below) depicts the genomic structure of the human HFE gene, its size, and single nucleotide polymorphisms (SNPs) mapped to it. The ruler provides the chromosome mapping information (J). The display lists alternatively spliced transcripts and their protein products and depicts them as bars (exons) linked by thin lines (introns) (K). Mouseover the accession of a transcript to get the summary information in a popup window (L). Click the arrow (M) in the upper left-hand to pop the SV display into a full-screen mode for optimal results and more functionality. Features available in this mode allow you to customize the display, import data in various formats, view and download of subsequences, as well as select a region to BLAST, and bring the alignment results back into the display. See SV document linked off the full-screen mode for more details.

Genomic Sequence: [NC_000006.12](#) Chromosome 6 Reference GRCh38.p12 Primary Assembly

Go to nucleotide: [Graphics](#) [FASTA](#) [GenBank](#)

26,886 K 26,888 K 26,890 K 26,892 K 26,894 K 26,896 K

Genes, NCBI Homo sapiens Annotation Release 109

HFE

NH_000410.3
NP_000410.1

NH_001390074.1
NP_001390074.1

NH_001390082.2
NP_001390082.2

NH_001390083.2
NP_001390083.2

NH_001390089.2
NP_001390089.2

NH_001390087.2
NP_001390087.2

NH_001390088.2
NP_001390088.2

NH_001390102.2
NP_001390102.2

NH_001390112.2
NP_001390112.2

XI_001151454.3
XR_241893.4

LOC108783645

NR_144383.1

Cited Variations, dbSNP b152 v2

rs2794719 T/C/G	rs28934889 G/A	rs907289 C/G/T	rs1572982 G/R/T	rs707889 G/R/T
rs9366637 C/T	rs11833557 G/A	rs140080192 G/R/C	rs1805521 G/A	rs18455
	rs1799945 C/G	rs1800582 G/A		
	rs1800736 A/T	rs11833553 A/C		
	rs28934537 G/C	rs1800758 G/A		
	rs28934596 T/C	rs1800708 T/C		
	rs2071303 T/C	rs11833558 G/C/T		
	rs28934995 A/C	rs2858996 G/R/C/T		
	rs148519482 G/C/T			
	rs797045145 G/A			
	rs2032451 G/T			

RNA-seq exon coverage, aggregate (filtered), NCBI Homo sapiens Annotation Release 109 - log

RNA-seq intron-spanning reads, aggregate (filtered), NCBI Homo sapiens Annotation Release 109 - log

26,886 K 26,888 K 26,890 K 26,892 K 26,894 K 26,896 K

Summary

mRNA: [NM_000410.3](#)

Title: mRNA-homeostatic iron regulator, transcript variant 1

Location: [26,087,281..26,095,241](#)

[Length]

Span on [NC_000006.12](#): 7,961 nt

Aligned length: 2,222 nt

Sequence length: 2,222 nt

[Positional Info]

[NC_000006.12](#) position: 26,094,564

mRNA position: 1,545

mRNA sequence: TTTTGGAGGAGGACT[C]CTTAAATTTGGGGG

Download: [NM_000410.3](#)

Links & Tools

View GeneID: [3077](#) (HFE)

View HGNC: [4886](#)

View MIM: [613609](#)

BLAST Genome-specific: [NC_000006.12](#) (26,087,281..26,095,241), [NM_000410.3](#)

BLAST Genomic: [NC_000006.12](#) (26,087,281..26,095,241)

BLAST mRNA: [NM_000410.3](#)

FASTA View: [NC_000006.12](#) (26,087,281..26,095,241), [NM_000410.3](#)

GenBank View: [NC_000006.12](#) (26,087,281..26,095,241), [NM_000410.3](#)

Graphical View: [NM_000410.3](#)

Expression

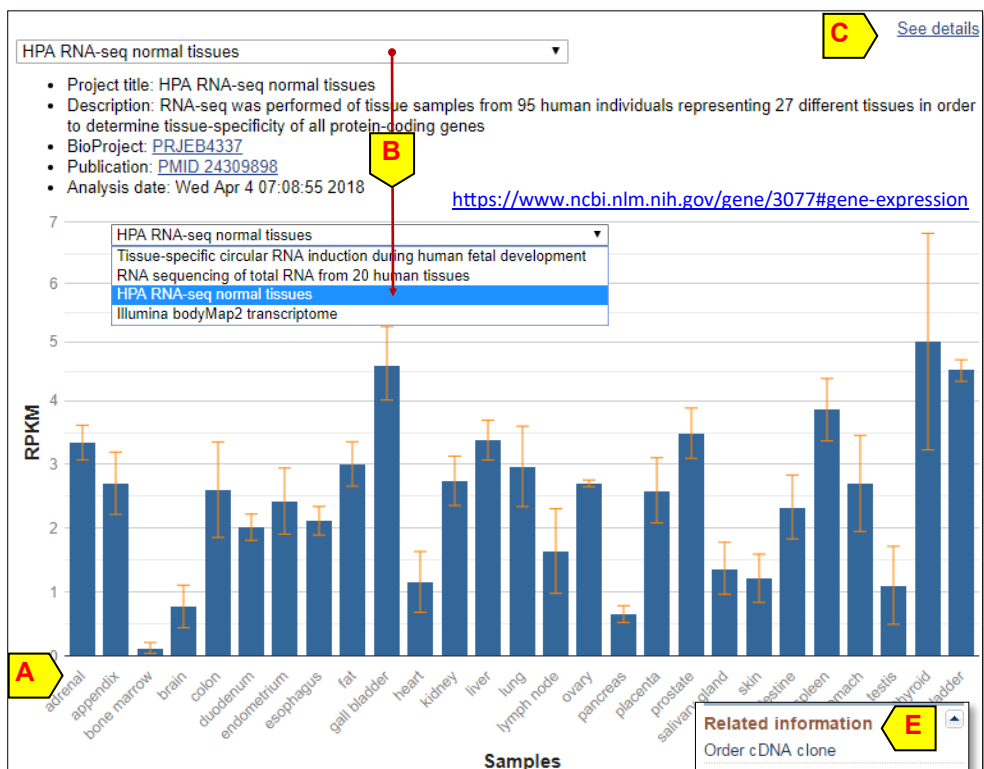
This newly added section sums up level of gene expression across various tissues (A), with information derived from alignment of Body Map 2 RNA-seq data (<https://www.ncbi.nlm.nih.gov/bioproject/144517>). You can use the pull-down menu to select source data from different studies (B). Click "See details" link (C) to see this report in a new window, along with the underlying data table and detailed counts from individual samples.

Phenotype Information

The "Phenotypes" section summarizes available information from OMIM, NHGRI, and PubMed as a list of links. The Phenotype and Genotype Integrator (PheGenI, D) tool, linked at the top, provides a user-friendly interface to collect genotype/phenotype data from available GWAS.

Related Records

A powerful feature of the Gene database is its integration with other gene-related resources at NCBI and elsewhere, via links in the "Related information" section (E). For the human HFE gene, this section contains genetic testing information from GTR (F), protein, RNA and genomic Reference Sequence records for the gene (G), various subsets of SNPs for the gene in dbSNP [6] (H), and Information on genetic disorders from/through OMIM [7] (I). For genes involved in known biologic pathways, links to BioSystems and outside resource like Kyoto Encyclopedia of Genes and Genomes (KEGG) are also available. Certain records also have links to databases outside NCBI to provide disease information or other details. The NCBI genome annotation process also analyzes and collects homologs for available genes and makes them available as the Ortholog link in the Summary section (p. 2).



Phenotypes

[Find tests for this gene in the NIH Genetic Testing Registry \(GTR\)](#)
[Review eQTL and phenotype association data in this region using PheGenI](#)

Associated conditions

Description	Tests
<p>Alzheimer's disease</p> <p>MedGen: C0002395, OMIM: 104300</p> <p>GeneReviews: Alzheimer Disease Overview, Early-Onset Familial Alzheimer Disease</p>	<p>Compare labs</p>
<p>Familial porphyria cutanea tarda</p> <p>MedGen: C0268323, OMIM: 176100</p> <p>GeneReviews: Familial Porphyria Cutanea Tarda, Hepatoerythropoietic Porphyria</p>	<p>Compare labs</p>

NHGRI GWAS Catalog

Items 1 - 10 of 24 << First < Prev Page 1 of 3 Next > Last >>

Description

A genome-wide association study of red blood cell traits using the electronic medical record.

NHGRI GWA Catalog
[NHGRI GWA Catalog](#), [PubMed](#)

A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium.

NHGRI GWA Catalog
[NHGRI GWA Catalog](#), [PubMed](#)

Related information

- Order cDNA clone
- 3D structures
- BioAssay
- BioAssay by Target (List)
- BioAssay by Target (Summary)
- BioAssay, by Gene target
- BioAssays, RNAi Target, Tested
- BioProjects
- Books
- CCDS
- ClinVar
- Conserved Domains
- dbVar
- Full text in PMC
- Full text in PMC_nucleotide
- Gene neighbors
- Genome
- GEO Profiles
- GTR
- HomoloGene
- Map Viewer
- MedGen
- Nucleotide
- OMIM
- Probe
- Protein
- PubChem Compound
- PubChem Substance
- PubMed
- PubMed (GeneRIF)
- PubMed (OMIM)
- PubMed(nucleotide/PMC)
- RefSeq Proteins
- RefSeq RNAs
- RefSeqGene
- SNP
- SNP: GeneView
- Taxonomy
- UniGene
- Variation Viewer

Displaying and downloading gene records

Using the pull-down menu (A), you can change the display of a Gene record to other formats. Selecting a format from the list browser will automatically switch the display. The “Tabular_(text)” option provides more values (such as exon counts) than visible in the “Tabular” display itself, and the Gene Table format provides transcript-specific exon details.

You can save a gene record to a local file using the “Send to” link (B) using the following steps: Check the “File” radio button (C), select file format from the pull-down list (D), and click the “Create File” button. Doing so from a summary display of multiple gene records saves all the records in the list.

The screenshot shows the top of the NCBI Gene record for HFE. A pull-down menu (A) is open, showing options: Full Report, Full Report (text), Expression, Gene Table, Gene Table (text), GeneRIF, Summary, Summary (text), Tabular, Tabular (text), ASN.1, XML, and UI List. To the right, the 'Send to' dialog box (B) is open, with the 'File' radio button (C) selected. A pull-down menu (D) is open from the 'Format' field in the dialog, showing options: Full Report (text), Gene Table (text), Summary (text), ASN.1, XML, and UI List.

Special Gene Table display

The screenshot shows the Gene Table display for HFE. The table has columns for Exon, Coding, Exon, Coding, and Intron under the heading 'Interval (exons 5' to 3')'. The table contains six rows of data. A pull-down menu (G) is open for the 'Coordinate System' field, showing 'Genomic-based'. A pull-down menu (H) is open for the 'Exon table for RefSeq mRNA NM_000410.3 and protein NP_000401.1 Interval (exons 5' to 3')' field, showing 'Top of page' and 'Gene Table help'. A pull-down menu (I) is open for the 'Interval (exons 5' to 3')' field, showing '26087441-26087516'. A pull-down menu (J) is open for the 'Length (bp)' field, showing '236'.

Interval (exons 5' to 3')					Length (bp)		
Exon	Coding	Exon	Coding	Intron			
26087281-26087516	26087441-26087516	236	76	3324			
26090841-26091104	26090841-26091104	264	264	209			
26091314-26091589	26091314-26091589	276	276	1095			
26092685-26092960	26092685-26092960	276	276	158			
26093119-26093232	26093119-26093232	114	114	953			
26094186-26095241	26094186-26094226	1056	41				

When genome annotation is available, the “Gene Table” format will be available to display genomic coordinates of exons and introns for all annotated transcript variants of the gene record (E). Clicking a set of numerical coordinates (F) retrieves the subsequence for that element in FASTA format. Default coordinates are based on the highest level reference assembly for that genome, such as the reference assembly for the chromosome. If needed, you can also select to use coordinates on other RefSeq records through options in the pull-down menus (G). One of the

splicing variants for human HFE displayed under “Gene Table” format is NM_000410.3 encoding a protein product NP_000401.1 (H). This variant has six exons, all of which are coding (I). Columns under the Length (bp) heading (J) contain the exon lengths and the length of the intervening introns.

Homo sapiens chromosome 6, GRCh38.p7 Primary Assembly

NCBI Reference Sequence: NC_000006.12
[GenBank](#) [Graphics](#)
 >gl|568815592:26087281-26087516 Homo sapiens chromosome 6, GRCh38.p7 Primary Assembly
 CTAAGTTCTGAAAGACCTGTGCTTTTACACAGGAAGTTTACTGGGCATCTCCTGAGCCTAGGCAATA
 GCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCGCCCCCAAAGAAGCGGAGATTTAACGGGGAC
 GTGCGCCAGAGCTGGGAAATGGGCCCGGAGCCAGGCCGCGCTTCTCCTCTGATGCTTTTGACAGAC
 CGCGTCTGACGAGGGCGCTTGCTGC

References

- Maglott D, et al. Gene: a gene-centered information resource at NCBI. [Nucleic Acids Res. 2014. pii: gku1055. \[Epub ahead of print\]](#)
- Maglott D, et al. Gene, The NCBI Handbook, 2nd Ed. NCBI, 2013. [www.ncbi.nlm.nih.gov/books/NBK169435/](#)
- Pruitt K, et al. RefSeq: an update on mammalian reference sequences. [Nucleic Acids Res. 2014. 42\(Database issue\):D756-63.](#)
- Pruitt K, et al. Chapter 18, The NCBI Handbook. NCBI, 2003. [www.ncbi.nlm.nih.gov/books/NBK21091/](#)
- Thibaud-Nissen, et al. Eukaryotic Genome Annotation Pipeline, The NCBI Handbook, 2nd Ed. [www.ncbi.nlm.nih.gov/books/NBK169439/](#)
- Kitts A, Sherry S. Chapter 5, The NCBI Handbook. NCBI, 2003. [www.ncbi.nlm.nih.gov/books/NBK21088/](#)
- Amberger JS, et al. OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. [Nucleic Acids Res. 2014 Nov 26. pii: gku1205. \[Epub ahead of print\]](#)
- Gene Help Manual [www.ncbi.nlm.nih.gov/books/NBK3841/](#)
- MyNCBI help manual. [https://www.ncbi.nlm.nih.gov/books/NBK3842/](#)