



dbSNP: Database for Short Genetic Variations

Catalog of nucleotide changes for human and other model organisms

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

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

Scope and Access

The NCBI Short Genetic Variations database (dbSNP) [1], commonly known as dbSNP, catalogs short variations in nucleotide sequences for human. These variations include single nucleotide variations, short nucleotide insertions and deletions, short tandem repeats. Short Genetic Variations may be common, thus representing true polymorphisms, or they may be rare. Some rare human entries have additional information associated with them, including disease associations from ClinVar [2], genotype information and allele origin, as some variations are somatic rather than germline events.

Short nucleotide variation data can be accessed via the SNP homepage and EUtils API:

www.ncbi.nlm.nih.gov/snp and www.ncbi.nlm.nih.gov/projects/SNP/SNPeutils.htm

VCF files and database bcp files are available for download through FTP and Aspera client at:

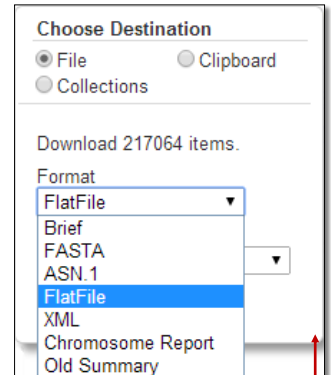
<ftp://ncbi.nlm.nih.gov/snp/> and www.ncbi.nlm.nih.gov/public/?snp/organisms/

The SNP API service, SPDI [3], is available at:

api.ncbi.nlm.nih.gov/variation/v0/

SNP data can also be accessed interactively through Variation Viewer:

www.ncbi.nlm.nih.gov/variation/view/



Searching for and Displaying SNP Records

You can search for variations on the dbSNP homepage by typing a query term in the search box and clicking the **Search** button (A). You can also use the **Advanced** (B) page to create complex queries to produce more precise results. The search below, "hfe[*gene*] AND human[*orgn*]", retrieves variations mapped to the human HFE gene. You can use options in the **Display settings** popup (C) to change the number of records displayed or sort retrieved variations in a different order. You can further narrow down retrieved variations by selecting filters present in the left column (D), or save them to a local file using the **Send to** (E) option. Use links to separate displays to see gene-centric listings (GeneView, F), graphical presentation under the context of genome or mRNA sequences (via HGVS names, G), or gene-centric display in a genomic context (Varview, H). Using the "Find related data" portlet (I), you can retrieve related entries from other NCBI databases for the set of variations in the display.

dbSNP Search: human[orgn] AND HFE[gene]

Display Settings: Summary, 20 per page, Sorted by SNP_ID

Results: 1 to 20 of 841

1. rs1799945 [Homo sapiens]
 TGACCAGCTGTTCTGTTCTATGAT[C/G]ATGAGAGTCGCGCTGTGGAGCCCCG
 Chromosome: 6:26090951
 Gene: HFE (GeneView)
 Functional Consequence: intron variant,missense,nc transcript variant
 Allele Origin: G(germline)/C(germline)
 Clinical significance: other
 Validated: by 1000G,by cluster,by frequency,by hapmap
 Global MAF: G=0.0731/366
 HGVS: NC_000006.11:g.26091179C>G, NC_000006.12:g.26090951C>G, NG_008720.2:g.8671C>G, NM_000410.3:c.187C>G, NM_001300749.1:c.187C>G, NM_139003.2:c.187C>G, NM_139004.2:c.187C>G, NM_139006.2:c.187C>G, NM_139007.2:c.77-363C>G, NM_139008.2:c.77-363C>G, NM_139009.2:c.118C>G, NM_139010.2:c.77-1734C>G, NM_139011.2:c.77-2168C>G, NP_000401.1:p.His63Asp, NP_001287678.1:p.His63Asp, NP_620572.1:p.His63Asp, NP_620573.1:p.His63Asp, NP_620575.1:p.His63Asp, NP_620578.1:p.His40Asp, XM_005249040.1:c.187C>G, XM_011514543.1:c.187C>G, XM_011514544.1:c.187C>G, XP_005249097.1:p.His63Asp, XP_011512845.1:p.His63Asp, XP_011512846.1:p.His63Asp, XR_241893.2:n.309C>G, XR_241894.1:n.434C>G

2. rs1800562 [Homo sapiens]
 CCTGGGGAAAGAGATACGT[A/G]CCAGGTGGAGCACCCAGGCGCTGGAT
 Chromosome: 6:26092913
 Gene: HFE (GeneView)
 Functional Consequence: intron variant,missense,nc transcript variant
 Allele Origin: G(germline)/A(germline)
 Clinical significance: Pathogenic
 Validated: by 1000G,by cluster,by frequency,by hapmap
 Global MAF: A=0.0126/63
 HGVS: NC_000006.11:g.26093141G>A, NC_000006.12:g.26092913G>A, NG_008720.2:g.10633G>A, NM_000410.3:c.845G>A

Find related data: Select Database: BioProject, BioSample, ClinVar, dbGaP, dbVar, Gene, Nucleotide, PMC, Probe, Protein, PubMed, SNP, Sparcle, Structure, Taxonomy

Display Settings: Format: Summary, Items per page: 5, 10, 20, 50, 100, 200, Sort by: Default order, Organism, SNP_ID, Success Rate, Heterozygosity, Chromosome Base Position

The New Reference SNP Report

The new Reference SNP Report linked from rsIDs, such as [rs1800730](#) shown below and on p.3, shows details of a variation record. The summary section at the top (A) provides an overview of the variant in a self-explanatory key/value format. It reports the allele in the forward orientation of the chromosome record. The information in display is also available in JSON format through the API link at the upper right (B). The new report separates details of the variation into various categories (C) and displays them under separate tabs.

dbSNP Short Genetic Variations

Reference SNP (rs) Report ALPHA **A**

rs1800730

Organism: *Homo sapiens*
 Position: chr6:26090957 (GRCh38.p7)
 Alleles: A>T
 Variation Type: SNV Single Nucleotide Variation
 Frequency: T=0.01163 (2506/215486, GnomAD)
 T=0.00957 (1202/125568, TOPMED)
 T=0.01009 (1225/121410, ExAC) (+ 4 more)

Clinical Significance: Reported in ClinVar
 Gene : Consequence: HFE : Missense Variant
 Publications: 16 citations
 Genomic View: **G** [See rs on genome](#)

Variant Details **D**

Genomic Placements **E**

Sequence name	Change
GRCh37.p13 chr 6	NC_000006.11:g.26091185A>T
GRCh38.p7 chr 6	NC_000006.12:g.26090957A>T
HFE RefSeqGene (LRG_748)	NG_008720.2:g.8677A>T

Gene: **HFE, hemochromatosis (plus strand)**

Molecule type	Change	Amino acid(Codon)	SO Term
HFE transcript variant 7	NM_139007.2:c.	N/A	Intron Variant
HFE transcript variant 8	NM_139008.2:c.	N/A	Intron Variant
HFE transcript variant 10	NM_139010.2:c.	N/A	Intron Variant
HFE transcript variant 11	NM_139011.2:c.	N/A	Intron Variant
HFE transcript variant 1	NM_000410.3:c.193A>T	S [AGT]> C [TGT]	Coding Sequence Variant
hereditary hemochromatosis protein isoform 1 precursor	NP_000401.1:p.Ser...	S [Ser]> C [Cys]	Missense Variant

The Variant Details tab (D) is open by default. This tab shows the genomic placement of the variant on the current and previous genome assemblies in HGVS format (E), with the gene/transcript mapping plus the relevant molecular consequences (F) given below.

The “See rs on genome” links to the graphical panel (G) of the display. It puts the variant in the context of genome and other neighboring variants. Different tracks bins mapped variants according to specific attributes, such those with clinical assertions (H). Use the button at upper right (I) to go to Variation Viewer for more detailed examination.

Genomic regions, transcripts, and products

Choose placement: GRCh38.p7 (NC_000006.12)

See rs1800730 in Variation Viewer **I**

Genes, NCBI Homo sapiens Annotation Release 109, 2018-03-27

dbSNP Build 151 (Homo sapiens Annotation Release 108) all data

Clinical, dbSNP b151 v2 Alpha

ClinVar Short Variations based on dbSNP Build 150 (Homo sapiens Annotation Release 108)

Cited Variations, dbSNP b151 v2 Alpha

1000 Genomes Phase 3, dbSNP b151 v2 Alpha

ClinVar Short Variations based on dbSNP Build 150 (Homo sapiens Annotation Release 108)

Cited Variations, dbSNP b151 v2 Alpha

The New Reference SNP Report (cont.)

Other tabs in the new Reference SNP Report provide category-specific information.

The **Clinical Significance** tab (A) lists related clinical assertions for the variant from ClinVar, with IDs linking directly to the records there.

The **Frequency** tab (B, filtered by "1000Genomes") lists allele frequency data from major studies, such as 1000 Genomes, ExAc, and Genome Aggregation Database. It will provide frequencies for major populations (C) if such a frequency breakdown is submitted by these studies. You can use the "Download" link (D) to get the data in a tab-delimited format. This provides a way to evaluate the impact of a variant if there is no information in the **Clinical Significance** and **Publications** sections.

The **Aliases** tab (E) lists HGVS names that can be used to describe the same variant. These names use different reference accessions, but they are equivalent and point to the same genomic variation. The table separates names into genomic, transcript non-coding, and transcript coding groups (F).

The **Submission** tab (G) lists submitter/batches that reported the variant. Submitters include large study projects or individual submitters. Older submission before adoption of asserted location will have ssIDs (H).

The **History** tab (I) tracks the change of the cluster and lists other rsIDs that has merged into this variant. In this case, variants rs115372583 and rs28934888 were determined to be duplication of rs1800730, so they were merged into a single record.

dbSNP established explicit connection between Reference SNP variant and biomedical literature citations through text-mining. The Reference SNP Report displays these connections under the **Publications** tab (J). You can use the "View All in PubMed" button (K) to retrieve the list of citations in PubMed and examine their abstracts for more information. Some of these citations may also have free full-text available from PubMed Central for detailed online reading.

Allele: T (allele ID: 15050) A

ClinVar Accession	Disease Names	Clinical Significance
RCV000000028.7	Hemochromatosis type 1	Conflicting-Interpretations-Of-Pathogenicity
RCV000290779.1	Hereditary hemochromatosis	Uncertain-Significance

Filter: 1000Genomes B D [Download](#)

Study	Population	Group	Sample Size	Ref Allele	Alt Allele
1000Genomes	Global	Study-wide	5008	A=0.996	T=0.004
1000Genomes	African	Sub	1322	A=1.000	T=0.000
1000Genomes	East Asian	Sub	1008	A=1.000	T=0.000
1000Genomes	Europe	Sub	1006	A=0.984	T=0.016
1000Genomes	South Asian	Sub	978	A=1.000	T=0.000
1000Genomes	American	Sub	694	A=1.000	T=0.000

Filter: E

Placement	A=	T	Note
GRCh38.p7 chr 6	NC_000006.12:g.26090957A=	NC_000006.12:g.26090957A>T	
GRCh37.p13 chr 6	NC_000006.11:g.26091185A=	NC_000006.11:g.26091185A>T	
HFE RefSeqGene (LRG_748)	NG_008720.2:g.8677A=	NG_008720.2:g.8677A>T	
HFE transcript variant 1	NM_000410.3:c.193A=	NM_000410.3:c.193A>T	F
HFE transcript variant 6	NM_139006.2:c.193A=	NM_139006.2:c.193A>T	
HFE transcript variant X2	XR_241893.3:n.315A=	XR_241893.3:n.315A>T	
hereditary hemochromatosis protein isoform 1 precursor	NP_000401.1:p.Ser65=	NP_000401.1:p.Ser65Cys	
hereditary hemochromatosis protein isoform 6 precursor	NP_620575.1:p.Ser65=	NP_620575.1:p.Ser65Cys	

7 Frequency, 43 SubSNP, 2 ClinVar submissions G

Filter:

No	Submitter	Submission ID	Date (Build)
7	The Exome Aggregation Consortium	NC_000006.11 - 26091185	May 10, 2018 (151)
8	The Avon Longitudinal Study of Parents and Children	NC_000006.11 - 26091185	May 10, 2018 (151)
9	1000Genomes	NC_000006.11 - 26091185	May 10, 2018 (151)
10	ILLUMINA	ss3644903282	May 10, 2018 (151)
11	ILLUMINA	ss3640757204	May 10, 2018 (151)

Filter: I

Associated ID	History Updated (Build)
rs115372583	Oct 26, 2010 (133)
rs28934888	May 25, 2008 (130)

16 citations for rs1800730 J

Filter:

PMID	Title	Author	Year	Journal
27657935	Association between the HFE C282Y, H63D Polymorphisms and the Risks of Non-Alcoholic Fatty Liver Disease, Liver Cirrhosis and Hepatocellular Carcinoma: An Updated Systematic Review and Meta-Analysis of 5,758 Cases and 14,741 Controls.	Ye Q et al.	2016	PloS one
27317329	Haplotype analysis of the HFE gene among populations of Northern Eurasia, in patients with metabolic disorders or stomach cancer, and in long-lived people.	Mikhailova SV et al.	2016	BMC genetics
27221532	Population-based analysis of the frequency of HFE gene polymorphisms: Correlation with the susceptibility to develop hereditary hemochromatosis.	Katsarou MS et al.	2016	Molecular medicine reports
27173269	Molecular epidemiology of HFE gene polymorphic variants	Alves LN et	2016	Genetics and

[View All in PubMed](#) K

Variation Viewer

The Variation Viewer (p.2, I) provides an interactive display of the variant under the context of annotation of the selected genome assembly. It correlates a variation and its molecular consequences in the data table with its genomic context in the graphical display (A). Filters in the left hand column (not shown) are available to selectively display variants of interest. More information on this tool is available online [4 and 5]

Other Ways to Access dbSNP Data

The SNP database is fully integrated with the Entrez system, enabling the access of variation data through links present in records from other NCBI databases. For example, you can project variations mapped to a RefSeq genomic or mRNA record (with NT_, NG_, NW_ or NM_ accessions) by using the **Customize view** (B) menu in the upper right hand corner of the sequence record, simply check the SNPs checkbox and click **Update View** (C) to activate the selection.

dbSNP also integrates disease-related nucleotide variations that were reported in literature and cited in rsID format, collected by OMIM, or submitted to ClinVar. The table below is the **Allelic Variant** display for OMIM record 613609, which cites the rsIDs in the dbSNP column (D).

References

1. The Database of Short Genetic Variation (dbSNP). Kitts A, Phan L, Ward MH, and Holmes JB. In The NCBI Handbook [Internet], 2nd ed. <https://www.ncbi.nlm.nih.gov/books/NBK174586/>
2. ClinVar: improving access to variant interpretations and supporting evidence. Landrum MJ, et al. Nucleic Acids Res. 2018 Jan 4;46(D1):D1062-D1067. <https://www.ncbi.nlm.nih.gov/pubmed/29165669>
3. New Web Services for Comparing and Grouping Sequence Variants. <https://go.usa.gov/xUeKT>.
4. Variation Viewer factsheet. https://ftp.ncbi.nih.gov/pub/factsheets/Factsheet_Variation_Viewer.pdf
5. Variation Viewer Online video tutorial. <https://www.youtube.com/watch?v=rrWZ9MFBwUM>

The screenshot shows the NCBI Variation Viewer interface. At the top, it displays the genomic region (NC_000006.12) and the selected variant (rs1800730). Below this, there are tracks for ClinVar Short Variations and dbSNP Build 151. A detailed table of alleles associated with rs1800730 is shown below the tracks.

Allele information					ClinVar information				
Variant allele	Transcript change	RefSeq	Protein change	Molecular consequence	Condition	Most severe clinical significance	Submitters	Highest review status	Last evaluated
T	c.193A>T	NM_000410.3	Ser65Cys	missense variant	Hemochromatosis type 1 and 1 more	Pathogenic	5	criteria provided, conflicting interpretations	Jun 21, 2016
T	c.193A>T	NM_001300749.1	Ser65Cys	missense variant	Hemochromatosis type 1 and 1 more	Pathogenic	5	criteria provided, conflicting interpretations	Jun 21, 2016
T	c.193A>T	NM_139003.2	Ser65Cys	missense variant	Hemochromatosis type 1 and 1 more	Pathogenic	5	criteria provided, conflicting interpretations	Jun 21, 2016

The screenshot shows the 'Customize view' menu. It has three main sections: 'Basic Features' with 'Default features' selected; 'Features added by NCBI' with '586 SNPs' checked; and 'Display options' with 'Show sequence' checked. An 'Update View' button is at the bottom.

The screenshot shows the 'FEATURES' panel. It lists genomic features like 'exon' and 'variation' with their location and qualifiers. The 'variation' feature is highlighted, showing its location (353) and qualifiers including the dbSNP ID 'rs1800730'.

613609
HFE GENE; HFE
Allelic Variants (11 Selected Examples) :

Number	Phenotype	Mutation	dbSNP	ExAC	ClinVar
.0001	HEMOCHROMATOSIS, TYPE 1 PORPHYRIA CUTANEA TARDA, SUSCEPTIBILITY TO, INCLUDED PORPHYRIA VARIEGATA, SUSCEPTIBILITY TO, INCLUDED HEMOCHROMATOSIS, JUVENILE, DIGENIC, INCLUDED ALZHEIMER DISEASE, SUSCEPTIBILITY TO, INCLUDED TRANSFERRIN SERUM LEVEL QUANTITATIVE TRAIT LOCUS 2, INCLUDED MICROVASCULAR COMPLICATIONS OF DIABETES, SUSCEPTIBILITY TO, INCLUDED	HFE, CYS282TYR	[rs1800562]	-	[RCV000210820...]
.0002	HEMOCHROMATOSIS, TYPE 1 MICROVASCULAR COMPLICATIONS OF DIABETES, SUSCEPTIBILITY TO, INCLUDED	HFE, HIS63ASP	[rs1799945]	[rs1799945]	[RCV000000027...]
.0003	HEMOCHROMATOSIS, TYPE 1	HFE, SER65CYS	[rs1800730]	-	[RCV000290779...]
.0004	HFE INTRONIC POLYMORPHISM	HFE, 5569G-A	[rs1800758]	[rs1800758]	[RCV000000031]
.0005	HFE POLYMORPHISM	HFE, VAL53MET	[rs28934889]	-	[RCV000000032]
.0006	HFE POLYMORPHISM	HFE, VAL59MET	[rs111033557]	-	[RCV000000033]
.0007	HEMOCHROMATOSIS, TYPE 1	HFE, GLN127HIS	[rs28934595]	-	[RCV000000034]
.0008	HEMOCHROMATOSIS, TYPE 1	HFE, ARG330MET	[rs111033558]	-	[RCV000000035]