

## Correlating Disease Genes and Phenotypes

An NCBI Mini-Course

This mini-course focuses on the correlation of a disease gene to the phenotype. It demonstrates how the NCBI resources such as the literature, expression and structure information can help provide potential functional information for disease genes.

Mutations in the HFE gene are associated with the hemochromatosis disease. A laboratory working on the hemochromatosis disease wants to elucidate the biochemical and structural basis for the function of the mutant protein.

### Outline:

In this exercise, we have the following goals:

1. Determine what is known about the HFE gene and protein (using Entrez Gene).
2. Determine identified SNPs and their locations in the HFE gene (using dbSNP).
3. Learn more about hemochromatosis and its genetic testing (using OMIM and Gene Tests)
4. Elucidate the biochemical and structural basis for the function of the wild type and mutant proteins, if possible.

During the first hour, an overview will be given using one disease gene, followed by an hour of hands-on session to practice using another disease gene. The following handout contains the screenshots of the overview.

URL: <http://www.ncbi.nlm.nih.gov/Class/minicourses/pheno.html>

Course Developed by Medha Bhagwat ([bhagwat@ncbi.nlm.nih.gov](mailto:bhagwat@ncbi.nlm.nih.gov))

## Problem 1

Mutations in the HFE gene are associated with the hemochromatosis disease. A laboratory working on the hemochromatosis disease wants to elucidate the biochemical and structural basis for the function of the mutant protein.

### **Outline:**

In this exercise, we have the following goals:

1. Determining what is known about the HFE gene and protein (using Entrez Gene).
2. Determining identified SNPs and their locations in the HFE gene (using dbSNP).
3. Learning more about the hemochromatosis disease and its genetic testing (using OMIM and Gene Tests)
4. Elucidating the biochemical and structural basis for the function of the wild type and the mutant protein, if possible (using CDD).

### **Step 1. Determining what is known about the HFE gene and protein (using Entrez Gene):**

Search for 'HFE' in [Entrez Gene](#). One entry is for the human HFE gene. Retrieve the entry by clicking on the HFE link.

What is the location and orientation of the HFE gene on the human genome? List the genes adjacent to it. How many alternatively spliced products have been annotated for the HFE gene when the RefSeq mRNA entries were reviewed? What are the differences in the spliced products? List some of the HFE gene aliases. What are the phenotypes associated with the mutations in the HFE gene? What is the name and function of the protein encoded by the HFE gene? What is the conserved domain in the protein? To which cellular component(s) is the protein localized? Obtain the locations of exons and introns for each transcript by choosing "Gene Table" from the Display pull down menu.

### **Step 2. Determining identified SNPs and their locations in the HFE gene:**

From the Links menu on the top right hand side of the page, click on the "SNP: GeneView" to access a list of the known SNPs (reported in dbSNP). By default, the SNPs in the coding region of a gene are reported. Additional SNPs such as in the upstream region or the introns can be viewed by clicking on the "in gene region" button. Currently, how many non-synonymous SNPs are placed on the longest hemochromatosis transcript variant, NM\_000410? How many of these have links to OMIM? We will concentrate on the cys282tyr mutant in the following analysis.

### ***Step 3. Learning more about the hemochromatosis disease and its genetic testing:***

Click on the OMIM link next to the one of the SNPs in the SNP report. What are the clinical features of hemochromatosis? List the 5 types of iron-overload disorders labeled hemochromatosis. Which of these is associated with mutations in the HFE gene? How many allelic variants of the HFE gene have been reported? What is the phenotype associated with the Cys282Tyr mutant?

Click on the Gene Tests link at top of the page. Identify some of the laboratories performing the clinical testing for hemochromatosis. Now refer to the Reviews section. Mutation analysis is available for which of the HFE alleles? List one explanation for the hemochromatosis phenotype caused by the Cys282Tyr mutant.

### ***Step 4. Elucidating the biochemical and structural basis for the function of the wild type and mutant proteins, if possible:***

Go back to the Entrez Gene report. Click on the first protein, NP\_000401. Select the Blink link. Click on the 3D structures button. The output contains a list of similar proteins with known 3D structures. The first entry, 1DE4G, represents the G chain of the hemochromatosis protein (complexed with transferrin receptor). Click on the blue dot next to 1DE4G to get the sequence alignment of the query protein to the G chain of 1DE4. Click on the "View 3D Structure" button. This downloads the structure of G chain of 1DE4 and its sequence alignment with the query protein. Zoom in the area of the disulphide bridge (colored in tan) by pressing "z" on the keyboard. Select the cysteine residues forming the disulphide bridge by double clicking on them. Mouse over the corresponding cysteine residues on the third query line in the alignment and view the amino acid number at the bottom left of the window. One of them is the cysteine at position 282. It is the same cysteine which is mutated to tyrosine causing the hemochromatosis phenotype.

### ***You can now easily explain why the C282Y mutant has an altered function.***

#### **Summary:**

This mini-course describes how to obtain information about the HFE gene, known SNPs in it, and elucidate the biochemical and structural basis for the function of the wild type and Cys282Tyr mutant protein.

- Summary:
1. The HFE gene is located on chromosome 6 and has at least 11 alternatively spliced products.
  2. Currently, there are 8 non-synonymous SNPs annotated on the protein NP\_000401.
  3. The Cys282Tyr mutant is associated with the hemochromatosis disease and the site of mutation is used in hemochromatosis genetic testing.

4. The HFE protein functions to regulate iron absorption by regulating the interaction of the transferrin receptor with transferrin where as the Cys282Tyr mutant fails to regulate this interaction leading to iron overload. The conserved cysteine 282 in the immunoglobulin constant region domain in the HFE protein is involved in formation of a disulphide bridge. Its mutation to tyrosine will alter the folding of the protein.

**NCBI** National Center for Biotechnology Information  
National Library of Medicine National Institutes of Health

PubMed All Databases BLAST OMIM Books TaxBrowser Structure

Search All Databases for  Go

**SITE MAP**  
Alphabetical List  
Resource Guide

▶ **What does NCBI do?** Hot Spots

Established in 1988 as a national resource for molecular biology information. NCBI creates ▶ Assembly Archive

**NCBI** Entrez, The Life Sciences Search Engine

HOME SEARCH SITE MAP PubMed Entrez Human Genome GenBank Map Viewer BLAST

Search across databases  GO CLEAR Help

Welcome to the new Entrez cross-database search page

<b>PubMed:</b> biomedical literature citations and abstracts	<b>Books:</b> online books
<b>PubMed Central:</b> free, full text journal articles	<b>OMIM:</b> online Mendelian Inheritance in Man
<b>Nucleotide:</b> sequence database (GenBank)	<b>Site Search:</b> NCBI web and FTP sites
<b>Protein:</b> sequence database	<b>UniGene:</b> gene-oriented clusters of transcript sequences
<b>Genome:</b> whole genome sequences	<b>CDD:</b> conserved protein domain database
<b>Structure:</b> three-dimensional macromolecular structures	<b>3D Domains:</b> domains from Entrez Structure
<b>Taxonomy:</b> organisms in GenBank	<b>UniSTS:</b> markers and mapping data
<b>SNP:</b> single nucleotide polymorphism	<b>PopSet:</b> population study data sets
<b>Gene:</b> gene-centered information	<b>GEO Profiles:</b> expression and molecular abundance profiles
<b>HomoloGene:</b> eukaryotic homology groups	<b>GEO DataSets:</b> experimental sets of GEO data
<b>PubChem Compound:</b> small molecule chemical structures	<b>Cancer Chromosomes:</b> cytogenetic databases
<b>PubChem Substance:</b> chemical substances screened for bioactivity	<b>PubChem BioAssay:</b> bioactivity screens of chemical substances
<b>Genome Project:</b> genome project information	<b>GENSAT:</b> gene expression atlas of mouse central nervous system
<b>Journals:</b> detailed information about the journals indexed in PubMed and other Entrez databases	<b>MeSH:</b> detailed information about NLM's controlled vocabulary
<b>NLM Catalog:</b> catalog of books, journals, and audiovisuals in the NLM collections	

**Enter** terms and **click 'GO'** to run the search against ALL the databases, **OR**  
**Click** Database Name or Icon to go directly to the Search Page for that database, **OR**  
**Click** Question Mark for a short explanation of that database.

NCBI Entrez Gene

Search Gene for hfe Go Clear

Entrez Gene is a searchable database of genes, from RefSeq genomes, and defined by sequence and/or located in the NCBI Map Viewer

News: New "has ccds" property added. [News archives...](#)

Sample Searches

Find genes by... free text, partial name and multiple species, chromosome and symbol

Search text: human muscular dystrophy, transporter[hfe] AND ("Drosophila melanogaster"[orgn] OR "Mus musculus"[orgn]), [H[chr] OR 2[chr]] AND adh\*[sym]

NCBI Entrez Gene

Search Gene for hfe Go Clear Save Search

Display Summary Show 20 Send to

All: 26 Current Only: 26 Genes Genomes: 26 SNP GeneView: 24

Items 1 - 20 of 26 Page 1 of 2 Next

1: [HFE](#) MGC cDNA clone, Links  
**Official Symbol:** HFE and **Name:** hemochromatosis [*Homo sapiens*]  
**Other Aliases:** HFE1, HH, HLA-H, MGC103790, dJ221C16.10.1  
**Other Designations:** MHC class I-like protein HFE; hemochromatosis protein; hereditary hemochromatosis protein HLA-H  
**Chromosome:** 6; **Location:** 6p21.3  
**GeneID:** 3077

2: [Hfe](#) Links  
**Official Symbol:** Hfe and **Name:** hemochromatosis [*Mus musculus*]

NCBI Entrez Gene

Search Gene for hfe Go Clear

Display Full Report Show 5 Send to

All: 1 Current Only: 1 Genes Genomes: 1 SNP GeneView: 1

1: [HFE hemochromatosis](#) [*Homo sapiens*]  
 GeneID: 3077 updated 14-Jan-2007

Summary

**Official Symbol** HFE provided by HGNC

**Official Full Name** hemochromatosis provided by HGNC

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

**See related** [HPRD:01993](#); [MIM:235200](#)

**Gene type** protein coding

**RefSeq status** Reviewed

**Organism** [Homo sapiens](#)

**Lineage** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo

**Also known as** HH; HFE1; HLA-H; MGC103790; dJ221C16.10.1

**Summary** The protein encoded by this gene is a membrane protein that is similar to MHC class I-type proteins and associates with beta2-microglobulin (beta2M). It is thought that this protein functions to regulate iron absorption by regulating the interaction of the transferrin receptor with transferrin. The iron storage disorder, hereditary haemochromatosis, is a recessive genetic disorder that results from defects in this gene. At least eleven alternatively spliced variants have been described for this gene. Additional variants have been found but their full-length nature has not been determined.

Genomic regions, transcripts, and products  
 Go to [reference sequence details](#)

NC\_000006.10

[26195487] 5' [26205938] 3'

NL\_129805.2  
 NL\_129806.2  
 NL\_024433.2  
 NL\_129804.2

NP\_028574.1 isoform 5 precursor  
 NP\_028571.1 isoform 2 precursor  
 NP\_014451.1 isoform 1 precursor  
 NP\_028572.1 isoform 4 precursor  
 CC04628.1  
 CC04679.1

**Genomic regions, transcripts, and products**

Go to [reference sequence details](#)

**NC\_000006.10**

■ - coding region ■ - untranslated region

**Genomic context**

chromosome: 6; Location: 6p21.3

[See HFE in MapViewer](#)

HIST1H3C HIST1H3C HFE HIST1H3C HIST1H3T

[Full text in PMC](#)  
[Probe](#)  
[Protein](#)  
[PubMed](#)  
[PubMed \(GeneRIF\)](#)  
[SNP](#)  
[SNP: Genotype](#)  
[SNP: GeneView](#)  
[Taxonomy](#)  
[UniSTS](#)  
[AppView](#)  
[CCDS](#)  
[GDB](#)  
[GeneTests for MIM: 235200](#)  
[HGMD](#)  
[HGNC](#)  
[HPRD](#)  
[KEGG](#)  
[MGC](#)  
[UniGene](#)  
[LinkOut](#)

**▼ Entrez Gene Info**  
**▶ Feedback**  
**▼ Subscriptions**

**Interactions**

**Description .....**

Product	Interactant	Other Gene	Complex	Source	Pubs
NP_000401.1	Beta 2 microglobulin	<a href="#">B2M</a>		<a href="#">HPRD</a>	<a href="#">PubMed</a>
NP_000401.1	Transferrin receptor 2	<a href="#">TFR2</a>		<a href="#">HPRD</a>	<a href="#">PubMed</a>
NP_000401.1	<a href="#">NP_003225.1</a>	<a href="#">TFRC</a>		<a href="#">HPRD</a>	<a href="#">PubMed</a>

**General gene information**

**Markers**

RH46796(e-PCR)  
 Links: [UniSTS:18176](#)  
 Alternate name: stSG24898

WI-17546(e-PCR)  
 Links: [UniSTS:30510](#)  
 Alternate names: EST261382; RH61086

RH46637(e-PCR)  
 Links: [UniSTS:36001](#)  
 Alternate name: stSG24673

A004R25(e-PCR)  
 Links: [UniSTS:41641](#)  
 Alternate name: RH25814

STS-U60319(e-PCR)  
 Links: [UniSTS:47384](#)  
 Alternate names: RH75899; sts-U60319

D6S2377(e-PCR)  
 Links: [UniSTS:57170](#)  
 Alternate names: GDB:558419S; sy899g1-19

**Phenotypes**

**General protein information**

**Names**  
hemochromatosis protein  
MHC class I-like protein HFE  
hereditary hemochromatosis protein HLA-H

**NCBI Reference Sequences (RefSeq)**

**Genomic**

- NG\_001335.1 Reference**  
Range: 71162..80773  
Download: [GenBank](#), [FASTA](#)

**mRNA and Protein(s)**

- NM\_000410.3–NP\_000401.1 hemochromatosis protein isoform 1 precursor**  
Description: Transcript Variant: This variant (1) encodes the longest isoform.  
Source sequence(s): [AF115265](#), [AJ249337](#), [U91328](#)  
Consensus CDS: [CCDS4578.1](#)  
Conserved Domains (2): [summary](#)

<b>cd00098</b> Location:223-298 Blast Score:169	IGC; Immunoglobulin domain constant region subfamily; members of the IGC subfamily are components of immunoglobulins, T-cell receptors, CD1 cell surface glycoproteins, secretory glycoproteins A/C, and Major Histocompatibility Complex (MHC) class I/II molecules
<b>pfam00129</b> Location:27-202 Blast Score:314	MHC_I; Class I Histocompatibility antigen, domains alpha 1 and 2
- NM\_139002.2–NP\_620571.1 hemochromatosis protein isoform 2 precursor**  
Description: Transcript Variant: This variant (2) lacks a large 3' region including the 3' CDS and UTR but has an alternate 3' exon, as compared to variant 1. The resulting protein (isoform 2) has a unique carboxy terminus.

**Related Sequences**

Nucleotide	Protein
Genomic <a href="#">AF184234.1</a>	<a href="#">AAF01222.1</a>
Genomic <a href="#">AF204869.1</a>	None
Genomic <a href="#">AF331065.1</a>	<a href="#">AAK16502.1</a>
Genomic <a href="#">AF525359.1</a>	<a href="#">AAM82608.1</a>
Genomic <a href="#">AF525499.1</a>	<a href="#">AAM91950.1</a>
Genomic <a href="#">CS187189.1</a>	<a href="#">CAJ42862.1</a>
Genomic <a href="#">U80914.1</a>	<a href="#">AAD00449.1</a>
Genomic <a href="#">U91328.1</a>	<a href="#">AAB82083.1</a>
Genomic <a href="#">Y09801.1</a>	<a href="#">CAA70934.1</a>
Genomic <a href="#">Z92910.1</a>	<a href="#">CAB07442.1</a>
mRNA <a href="#">AF079407.1</a>	<a href="#">AAC62646.1</a>
mRNA <a href="#">AF079408.1</a>	<a href="#">AAC62647.1</a>
mRNA <a href="#">AF079409.1</a>	<a href="#">AAC62648.1</a>
mRNA <a href="#">AF109385.1</a>	<a href="#">AAD52104.1</a>
mRNA <a href="#">AF115264.1</a>	<a href="#">AAG29571.1</a>

Search Gene for [ ] Go Clear

Limits Preview/Index History Clipboard Details

Display: Full Report Show 5 Send to

All: Summary Genomes: 1 SNP GeneView: 1

Gene: [ ]

updated 14-Jan-2007 Entrez Gene Home

▼ Table Of Contents

- Summary
- Genomic regions, transcripts...
- Genomic context
- Bibliography
- HIV-1 protein interactions
- Interactions
- General gene information
- General protein information
- Reference Sequences
- Related Sequences

Primary source: [HGNC:4886](#)

See related: [HPRD:01993](#); [MIM:235200](#)

Gene type: protein coding

RefSeq status: Reviewed

mRNA	bp	exons	Protein	aa	exons
<a href="#">NM_139005.2</a>	1417	5	<a href="#">NP_620574.1</a>	277	5
<a href="#">NM_139002.2</a>	878	4	<a href="#">NP_620571.1</a>	162	4
<a href="#">NM_000410.3</a>	2222	6	<a href="#">NP_000401.1</a>	349	6
<a href="#">NM_139004.2</a>	1946	5	<a href="#">NP_620573.1</a>	257	5
<a href="#">NM_139003.2</a>	1904	5	<a href="#">NP_620572.1</a>	243	5
<a href="#">NM_139009.2</a>	2153	6	<a href="#">NP_620578.1</a>	326	6
<a href="#">NM_139007.2</a>	1958	5	<a href="#">NP_620576.1</a>	261	5
<a href="#">NM_139008.2</a>	1916	5	<a href="#">NP_620577.1</a>	247	5
<a href="#">NM_139010.2</a>	1682	4	<a href="#">NP_620579.1</a>	169	4
<a href="#">NM_139011.2</a>	1406	3	<a href="#">NP_620580.1</a>	77	3
<a href="#">NM_139006.2</a>	1180	6	<a href="#">NP_620575.1</a>	335	6

**Exon information:**  
[NM\\_139005.2](#) length: 1417 bp, number of exons: 5  
[NP\\_620574.1](#) length: 277 aa, number of exons: 5

EXON		Coding EXON		INTRON	
coords	length	coords	length	coords	length
<a href="#">62 - 297</a>	236 bp	<a href="#">222 - 297</a>	76 bp	<a href="#">298 - 3621</a>	3324 bp
<a href="#">3622 - 3885</a>	264 bp	<a href="#">3622 - 3885</a>	264 bp	<a href="#">3886 - 4094</a>	209 bp
<a href="#">4095 - 4370</a>	276 bp	<a href="#">4095 - 4370</a>	276 bp	<a href="#">4371 - 5465</a>	1095 bp
<a href="#">5466 - 5667</a>	202 bp	<a href="#">5466 - 5667</a>	202 bp	<a href="#">5668 - 9171</a>	3504 bp
<a href="#">9172 - 9610</a>	439 bp	<a href="#">9172 - 9184</a>	13 bp		

[NM\\_139002.2](#) length: 878 bp, number of exons: 4  
[NP\\_620571.1](#) length: 162 aa, number of exons: 4

EXON		Coding EXON		INTRON	
coords	length	coords	length	coords	length
<a href="#">62 - 297</a>	236 bp	<a href="#">222 - 297</a>	76 bp	<a href="#">298 - 3621</a>	3324 bp
<a href="#">3622 - 3885</a>	264 bp	<a href="#">3622 - 3885</a>	264 bp	<a href="#">3886 - 4094</a>	209 bp

[AceView](#)  
[CCDS](#)  
[Evidence Viewer](#)  
[GDB](#)  
[GeneTests for MIM: 235200](#)  
[HGMD](#)  
[HGNC](#)  
[HPRD](#)  
[KEGG](#)  
[MGC](#)  
[ModelMaker](#)  
[UniGene](#)  
[LinkOut](#)  


---

[Entrez Gene Info](#)  


---

[Feedback](#)  


---

[Subscriptions](#)

History
Clipboard
Details

Show 5
Send to

Genes Genomes: 1
SNP GeneView: 1

**hemochromatosis [Homo sapiens]** updated 14-Jan-2007 [Entrez Gene Home](#)

**HFE** provided by [HGNC](#)

hemochromatosis provided by [HGNC](#)

[HGNC:4886](#)

[HPRD:01993](#); [MIM:235200](#)

protein coding

Reviewed

[Homo sapiens](#)

*Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo*

HH; HFE1; HLA-H; MGC103790; dJ221C16.10.1

The protein encoded by this gene is a membrane protein that is similar to MHC class I-type proteins and associates with beta2-microglobulin (beta2M). It is thought that this protein functions to regulate iron absorption by regulating the interaction of the transferrin receptor with transferrin. The iron storage disorder, hereditary haemochromatosis, is a recessive genetic disorder that results from defects in this gene. At least eleven alternatively spliced variants have been described for this gene. Additional variants have been found but their full-length nature has not been determined.

**Transcripts, and products**

[Transcript details](#)

**NC\_000006.10**

**Table of Contents**

- Summary
- Genomic regions, transcripts...
- Genomic context
- Bibliography
- HIV-1 protein interactions
- Interactions
- General gene information
- General protein information
- Reference Sequence
- Related Sequences
- Additional Links

**Links** Explain

- Order cDNA clone
- Books
- Conserved Domains
- Genome
- GEO Profiles
- HomoloGene
- Map Viewer
- Nucleotide
- OMIA
- OMIM
- Full text in PMC
- Probe
- Protein
- PubMed
- PubMed (GeneRef)
- SNP
- SNP: GeneView
- SNP: GeneView
- Taxonomy
- UniSTS
- AceView
- CCDS
- Evidence Viewer
- GDB

NCBI Single Nucleotide Polymorphism

PubMed Nucleotide Protein Genome Structure PopSet Taxonomy OMIM Books SNP

Search Entrez SNP for  Go

BUILD 126 SNP linked to Gene (geneID:3077)

SNP are linked from gene [HFE](#) via the following methods:  
[Contig Annotation](#) [GenBank\(mrna\) Mapping](#)

Send all rs# to Batch Query Download all rs# to file [GENE GENOTYPE REPORT](#)

Gene Model (mRNA alignment) information from genome sequence

Total gene model (contig mRNA transcript): 22						
mRNA	transcript	protein	mRNA orientation	Contig	Contig Label	snp list
NM_000410	plus strand	NP_000401	forward	NT_007592	reference	currently shown
NM_000410	plus strand	NP_000401	forward	NW_922984	Celera	<a href="#">view</a>
NM_139002	plus strand	NP_620571	forward	NT_007592	reference	<a href="#">view</a>
NM_139002	plus strand	NP_620571	forward	NW_922984	Celera	<a href="#">view</a>
NM_139003	plus strand	NP_620572	forward	NT_007592	reference	<a href="#">view</a>
NM_139003	plus strand	NP_620572	forward	NW_922984	Celera	<a href="#">view</a>
NM_139004	plus strand	NP_620573	forward	NT_007592	reference	<a href="#">view</a>
NM_139004	plus strand	NP_620573	forward	NW_922984	Celera	<a href="#">view</a>
NM_139005	plus strand	NP_620574	forward	NT_007592	reference	<a href="#">view</a>
NM_139005	plus strand	NP_620574	forward	NW_922984	Celera	<a href="#">view</a>
NM_139006	plus strand	NP_620575	forward	NT_007592	reference	<a href="#">view</a>
NM_139006	plus strand	NP_620575	forward	NW_922984	Celera	<a href="#">view</a>
NM_139007	plus strand	NP_620576	forward	NT_007592	reference	<a href="#">view</a>
NM_139007	plus strand	NP_620576	forward	NW_922984	Celera	<a href="#">view</a>
NM_139008	plus strand	NP_620577	forward	NT_007592	reference	<a href="#">view</a>
NM_139008	plus strand	NP_620577	forward	NW_922984	Celera	<a href="#">view</a>

Population Class Publication Locus Information STS Markers Mouse Strains

HAPLOTYPE Submission Specifications Sample HapSet Sample Individual RELATED SITES Genome Variation Working Group Whole Genome Association dbMHC

in gene region  cSNP  has frequency  double hit  haplotype tagged refresh

gene model Contig Label Contig mRNA protein mRNA orientation transcript snp  
 (contig mRNA transcript): reference NT\_007592 NM\_000410 NP\_000401 forward plus strand 8,0

Region	Contig position	mRNA pos	dbSNP rs# cluster id	Heterozygosity	Validation	3D	OMIM	Function	dbSNP allele	Protein residue	Codon pos
exon_2	16949347	325	<a href="#">rs2242956</a>	N.D.		H	Yes	nonsynonymous	C	Thr [T]	2
								contig reference	T	Met [M]	2
	16949430	408	<a href="#">rs1799945</a>	0.139		H	Yes	nonsynonymous	G	Asp [D]	1
								contig reference	C	His [H]	1
	16949436	414	<a href="#">rs1800730</a>	N.D.			Yes	nonsynonymous	T	Cys [C]	1
								contig reference	A	Ser [S]	1
	16949520	498	<a href="#">rs28934597</a>	N.D.			Yes	nonsynonymous	C	Arg [R]	1
								contig reference	G	Gly [G]	1
	16949557	535	<a href="#">rs28934596</a>	N.D.			Yes	nonsynonymous	C	Thr [T]	2
								contig reference	T	Ile [I]	2
exon_3	16949833	602	<a href="#">rs28934595</a>	N.D.			Yes	nonsynonymous	C	His [H]	3
								contig reference	A	Gln [Q]	3
exon_4	16951197	871	<a href="#">rs4986950</a>	0.005		H	Yes	nonsynonymous	T	Ile [I]	2
								contig reference	C	Thr [T]	2
	16951392	1066	<a href="#">rs1800562</a>	0.043		H	Yes	nonsynonymous	A	Tyr [Y]	2
								contig reference	G	Cys [C]	2

Search for  in  Highlight: NCBI PubMed Gene Nucleotide My NCBI Clear Uninstall Links

**.0001 HEMOCHROMATOSIS [HFE, CYS282TYR] dbSNP**

**NCBI**

MIM +235200

Description  
Clinical Features  
Other Features  
Inheritance  
Mapping  
Heterogeneity  
Molecular Genetics  
Genotype/Phenotype  
Correlations  
Diagnosis  
Clinical Management  
Population Genetics  
Pathogenesis  
Cloning  
Biochemical Features  
Gene Structure  
Gene Function  
Nomenclature  
Normal Model  
Allelic Variants  
View List

See Also  
References  
Contributors  
Creation Date  
Edit History

• Clinical Synopsis  
• Gene map

Entrez: Gene  
Nomenclature

**PORPHYRIA VARIEGATA, INCLUDED**  
**HEMOCHROMATOSIS, JUVENILE, DIGENIC, INCLUDED**  
**ALZHEIMER DISEASE, SUSCEPTIBILITY TO, INCLUDED**

In patients with hemochromatosis, [Feder et al. \(1996\)](#) identified an 845G-A transition in the HFE gene (which they referred to as HLA-H or cDNA 24), resulting in a cys282-to-tyr (C282Y) substitution. This missense mutation occurs in a highly conserved residue involved in the intramolecular disulfide bridging of MHC class I proteins, and could therefore disrupt the structure and function of this protein. Using an allele-specific oligonucleotide-ligation assay on their group of 178 patients, they detected the C282Y mutation in 85% of all HFE chromosomes. In contrast, only 10 of the 310 control chromosomes (3.2%) carried the mutation, a carrier frequency of 10/155 = 6.4%. One hundred forty-eight of 178 HH patients were homozygous for this mutation, 9 were heterozygous, and 21 carried only the normal allele. These numbers were extremely discrepant from Hardy-Weinberg equilibrium. The findings corroborated heterogeneity among the hemochromatosis patients, with 83% of cases related to C282Y homozygosity. 🗨

[Jazwinska et al. \(1996\)](#) provided convincing evidence that the C282Y mutation in homozygous form in the HFE gene is the cause of hemochromatosis. In studies in Australia, patients properly characterized at the genotypic and phenotypic level all showed homozygosity for the C282Y amino acid substitution. Irrespective of haplotype, all HH heterozygotes were cys/tyr heterozygotes, and all homozygous normal controls were cys/cys homozygotes. The presence of a single mutation in all patients contrasted with the data of [Feder et al. \(1996\)](#), who reported a lower frequency of the mutation. [Jazwinska et al. \(1996\)](#) suggested that different clinical criteria for the diagnosis of HH may account for the difference, or that HH may not be as homogeneous as previously believed. They noted that a key question is why there is a variation in severity of iron loading in HH that is haplotype-related when the mutation is identical in all haplotypes tested. [Jazwinska et al. \(1996\)](#) hypothesized that the HFE locus is the primary HH locus, but that there are likely to be other 6p-linked modifying genes that would explain both the HLA-linked haplotype variation in expression of the disorder and the large region of linkage disequilibrium present in all populations and spanning at least 4.5 Mb distal of D6S265. 🗨

[Jouanolle et al. \(1996\)](#) commented on the significance of the C282Y mutation on the basis of a group of 65 unrelated affected individuals who had been under study in France for more than 10 years and identified by stringent criteria. Homozygosity for the C282Y mutation was found in 59 of 65 patients (90.8%), 3 of the patients were compound heterozygotes for the C282Y mutation and the H63D mutation ([235200.0002](#)); 1 was homozygous for the H63D mutation; and 2 were heterozygous for H63D. These results corresponded to an allelic frequency of 93.1% for the C282Y and 5.4% for the H63D mutations, respectively. Of note, the C282Y mutation was never observed in the family-based controls, while it was present in 5.8% of the general Breton population. In contrast, the H63D allelic frequency was nearly the same in both control groups (15% and 16.5% in the family-based and general population controls, respectively). The C282Y mutation was never observed, even in heterozygous form, in the family-based controls in whom all signs of iron overload had been excluded, whereas the general population displayed 5.8% of heterozygotes. This corresponds to a theoretical frequency of about 1 per 1,000 for the disease, which is slightly lower than generally estimated. While the experience of [Jouanolle et al. \(1996\)](#) appeared to indicate a close relationship of C282Y to hemochromatosis, the implication of the H63D variant was not clear. 🗨

[Beutler et al. \(1996\)](#) reported mutation analysis of 147 patients with hereditary hemochromatosis and 193 controls; 121 (82.3%) HH patients were homozygous for the C282Y mutation, while 10 (6.8%) were heterozygous. All of the C282Y homozygous patients were also homozygous for the wildtype nucleotide 187C (see H63D; [235200.0002](#)), and all C282Y heterozygotes had at least 1 copy of 187C. Thus, the 2 nucleotides, 845 and 187, were in complete linkage disequilibrium; nucleotide 187 was a

Done Local intranet

NCBI **OMIM** Online Mendelian Inheritance in Man Johns Hopkins University

All Databases PubMed Nucleotide Protein Genome Structure PMC Taxonomy OMIM

Search OMIM for Go Clear

Limits Preview/Index History Clipboard Details

Display Allelic Variants Show 20 Send to

All: 1 OMIM dbSNP: 1 OMIM UniSTS: 1

**+235200** **HEMOCHROMATOSIS; HFE** [GeneTests, Links](#)

**ALLELIC VARIANTS**  
(selected examples)

- [0001 HEMOCHROMATOSIS](#) [HFE, CYS282TYR ] **dbSNP**
- [0002 HEMOCHROMATOSIS](#) [HFE, HIS63ASP ] **dbSNP**
- [0003 HEMOCHROMATOSIS](#) [HFE, SER65CYS ] **dbSNP**
- [0004 HFE INTRONIC POLYMORPHISM](#) [HFE, 5569G-A]
- [0005 HFE POLYMORPHISM](#) [HFE, VAL53MET ] **dbSNP**
- [0006 HFE POLYMORPHISM](#) [HFE, VAL59MET ] **dbSNP**
- [0007 PORPHYRIA VARIEGATA](#) [HFE, GLN127HIS ] **dbSNP**
- [0008 HEMOCHROMATOSIS](#) [HFE, ARG330MET]
- [0009 HEMOCHROMATOSIS](#) [HFE, ILE105THR ] **dbSNP**
- [0010 HEMOCHROMATOSIS](#) [HFE, GLY93ARG] **dbSNP**
- [0011 HEMOCHROMATOSIS](#) [HFE, GLN283PRO ]

Copyright © 1966-2006 Johns Hopkins University

Display Allelic Variants Show 20 Send to

<a href="#">Home Page</a>	<a href="#">About GeneTests</a>	<a href="#">GENEReviews</a>	<a href="#">Laboratory Directory</a>	<a href="#">Clinic Directory</a>	<a href="#">Educational Materials</a>
---------------------------	---------------------------------	-----------------------------	--------------------------------------	----------------------------------	---------------------------------------

Funded by the National Institutes of Health 

The result of your search (below) includes a group of related disorders with your search term in **bold** or an alphabetical listing of the individual entries that match your search term. For more information about search results, see [Interpreting Your Search Results](#).

Search Result for OMIM# 235200

HFE- Associated Hereditary Hemochromatosis [Testing](#) [Research](#) [Reviews](#) [Resources](#)

Home Page	About GeneTests	GENEReviews	Laboratory Directory	Clinic Directory	Educational Materials			
Funded by the National Institutes of Health								
<h3>HFE- Associated Hereditary Hemochromatosis</h3> <p align="center">Select all clinical laboratories</p>								
Laboratories offering clinical testing:								
		Sequencing of entire coding region	Sequencing of select exons	Mutation scanning	Targeted mutation analysis	Prenatal diagnosis	Preimplantation diagnosis	Clinical confirmation of mutations identified in a research lab
ARUP Laboratories, Inc. <a href="#">ARUP Laboratories</a> Salt Lake City, UT					●			
Elaine Lyon, PhD; Rong Mao, MD; Edward R Ashwood, MD; Marzia Pasquali, PhD								
Acibadem Healthcare Group <a href="#">Acibadem Genetic Diagnostic Center</a> Istanbul, Turkey					●			
Ender Altioak, MD, PhD								
Alberta Children's Hospital <a href="#">Molecular Diagnostic Laboratory</a> Calgary, Alberta, Canada					●			
Peter Bridge, PhD, FCCMG, FACMG; Jillian Parboosingh, PhD, FCCMG								

Home Page	About GeneTests	GENEReviews	Laboratory Directory	Clinic Directory	Educational Materials	
Funded by the National Institutes of Health						
<p>The result of your search (below) includes a group of related disorders with your search term in <b>bold</b> or an alphabetical listing of the individual entries that match your search term. For more information about search results, see <a href="#">Interpreting Your Search Results</a>.</p>						
<p><b>Search Result for OMIM# 235200</b></p> <p>HFE- Associated Hereditary Hemochromatosis</p>						
			Testing	Research	Reviews	Resources

<a href="#">Home Page</a>	<a href="#">About GeneTests</a>		<a href="#">Laboratory Directory</a>	<a href="#">Clinic Directory</a>	<a href="#">Educational Materials</a>
---------------------------	---------------------------------	--	--------------------------------------	----------------------------------	---------------------------------------

Funded by the NIH • Developed at the University of Washington, Seattle

**HFE-Associated Hereditary Hemochromatosis**

- [Summary](#)
- [Diagnosis](#)
- [Clinical Description](#)
- [Prevalence](#)
- [Differential Diagnosis](#)
- [Management](#)
- [Genetic Counseling](#)
- [Molecular Genetics](#)
- [Resources](#)
- [References](#)
- [Author Information](#)
- [Top of Page](#)

**Printable Copy**  
(Disable glossary before printing)

**Disable Glossary**  
(Returns to top)

**Title Index**

## HFE-Associated Hereditary Hemochromatosis

---

**Authors:** Kris V Kowdley, MD  
Jonathan F Tait, MD, PhD  
Robin L Bennett, MS  
Arno G Motulsky, MD

**About the Authors**

**Initial Posting:** 3 April 2000      **Last Update:** 4 December 2006

---

### Summary

**Disease characteristics.** *HFE*-associated hereditary hemochromatosis (*HFE*-HHC) is characterized by inappropriately high absorption of iron by the gastrointestinal mucosa, resulting in excessive storage of iron particularly in the liver, skin, pancreas, heart, joints, and testes. Abdominal pain, weakness, lethargy, and weight loss are early symptoms. Without therapy, males may develop symptoms between age 40 and 60 years and females after menopause. Hepatic fibrosis or cirrhosis may occur in untreated individuals after age 40 years. Other findings in untreated individuals may include progressive increase in skin pigmentation, diabetes mellitus, congestive heart failure and/or arrhythmias, arthritis, and hypogonadism.

This description applies to individuals with clinical expression of *HFE*-HHC. A large, but yet as undefined, fraction of **homozygotes** for *HFE*-HHC do not develop clinical symptoms (i.e., **penetrance** is low).

**Diagnosis/testing.** The diagnosis of *HFE*-HHC in individuals with clinical symptoms consistent with *HFE*-HHC and/or biochemical

Click on **defined terms**; definition displays here.

**HFE-Associated Hereditary Hemochromatosis**

- [Summary](#)
- [Diagnosis](#)
- [Clinical Description](#)
- [Prevalence](#)
- [Differential Diagnosis](#)
- [Management](#)
- [Genetic Counseling](#)
- [Molecular Genetics](#)
- [Resources](#)
- [References](#)
- [Author Information](#)
- [Top of Page](#)

**Printable Copy**  
(Disable glossary before printing)

**Disable Glossary**  
(Returns to top)

**Title Index**

region, complementation group from OMIM, protein name from Swiss-Prot.

**OMIM Entries for HFE-Associated Hereditary Hemochromatosis**

235200	HEMOCHROMATOSIS; HFE
--------	----------------------

**Genomic Databases for HFE-Associated Hereditary Hemochromatosis**

Gene Symbol	Entrez Gene	HGMD	GeneCards	GDB	GenAtlas
<i>HFE</i>	235200	HFE	<i>HFE</i>	119309	<i>HFE</i>

For a description of the genomic databases listed, click [here](#).

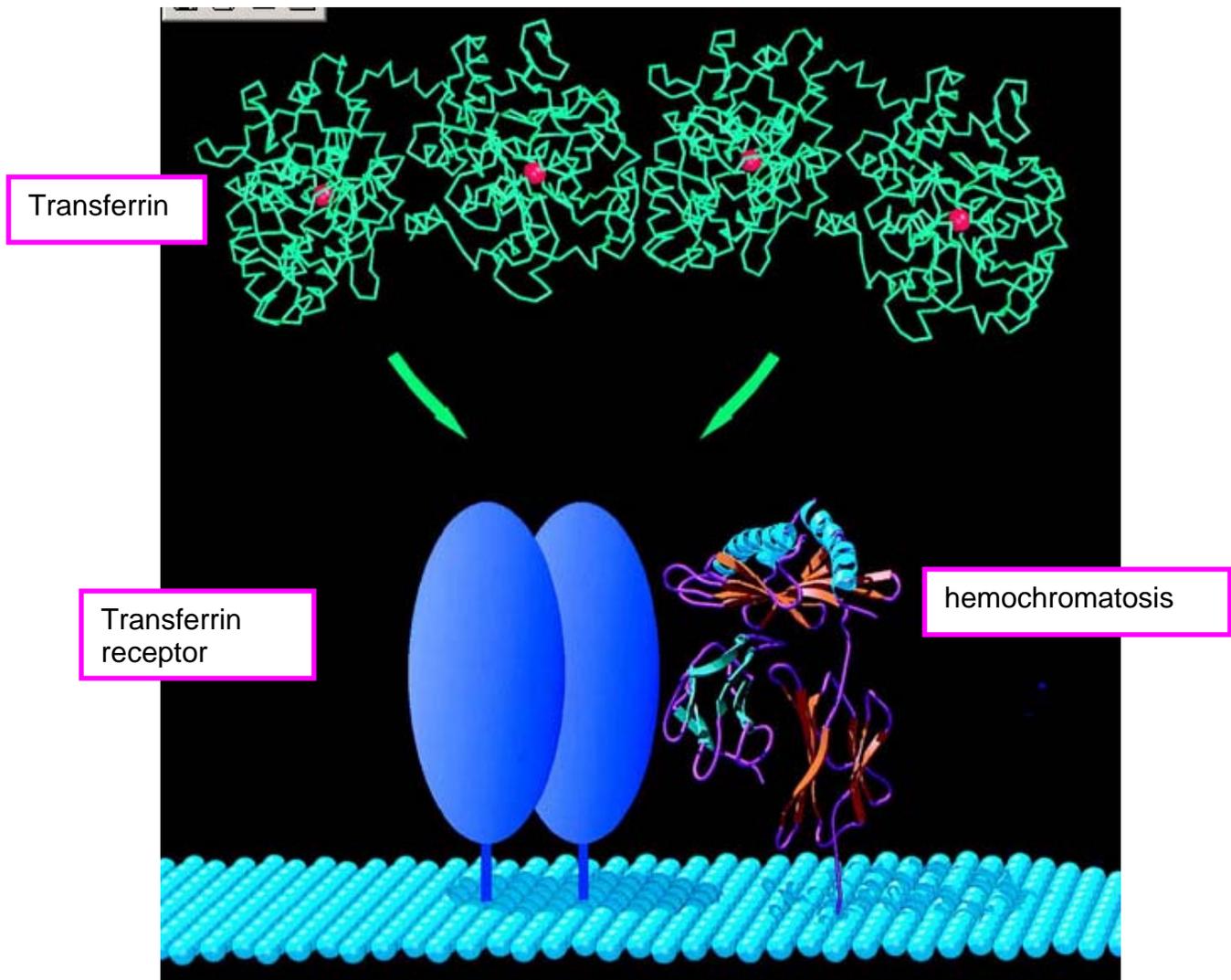
**Normal allelic variants:** The *HFE* gene is about 13 kb in size and contains seven **exons** [Feder et al 1996 , Albig 1998]; *HFE* gives rise to at least eleven alternative transcripts encoding four to seven **exons**.

**Pathologic allelic variants:** At least 28 distinct **mutations** have been reported, most being missense or **nonsense mutations**. Two **missense mutations** account for the vast majority of disease-causing **alleles** in the population:

- Cys282Tyr (p.C282Y; **nucleotide** 845G>A). This **missense mutation** removes a highly conserved cysteine residue that normally forms an intermolecular disulfide bond with beta-2-microglobulin, and thereby prevents the **protein** from being expressed on the cell surface.
- His63Asp (p.H63D; **nucleotide** 187C>G). This **missense mutation** may alter a pH-dependent intramolecular salt bridge, possibly affecting interaction of the HFE **protein** with the transferrin receptor.

**Normal gene product:** The largest predicted primary **translation product** is 348 amino acids, which gives rise to a mature **protein** of about 321 amino acids after cleavage of the signal sequence. The HFE **protein** is similar to HLA Class I molecules at the primary [Feder et al 1996] and tertiary structure [Lebron et al 1998] levels. The mature **protein** is expressed on the cell surface as a heterodimer with beta-2-microglobulin, and this interaction is necessary for normal presentation on the cell surface. The normal HFE **protein** binds to transferrin receptor 1 on the cell surface and may reduce cellular iron uptake; however, the exact means by which the HFE **protein** regulates iron uptake is as yet unclear [Fleming et al 2004].

**Abnormal gene product:** The p.C282Y **mutation** destroys a key cysteine residue that is required for disulfide bonding with beta-2-microglobulin. As a result, the HFE **protein** does not mature properly and becomes trapped in the endoplasmic reticulum and Golgi apparatus, leading to decreased cell-surface expression. The mechanistic basis for the **phenotypic** effect of other *HFE* **mutations** is not clear at present.



Bacon et al. Gastroenterology, 116:193-207, Figure 4

**The hemochromatosis protein functions to regulate iron absorption by regulating the interaction of the transferrin receptor with transferrin.**

NCBI OMIM Online Mendelian Inheritance in Man Johns Hopkins University

All Databases PubMed Nucleotide Protein Genome Structure PMC Taxonomy OMIM

Search OMIM for Go Clear

Limits Preview/Index History Clipboard Details

Display Detailed Show 20 Send to

All: 1 OMIM dbSNP: 1 OMIM UniSTS: 1

**+235200**  
**HEMOCHROMATOSIS; HFE**

**Alternative titles; symbols**

HLAH  
**HEMOCHROMATOSIS, HEREDITARY; HH**  
**HFE GENE, INCLUDED; HFE, INCLUDED**

Gene map locus [6p21.3](#)

TEXT

DESCRIPTION

The clinical features of hemochromatosis include cirrhosis of the liver, diabetes, hypermelanotic pigmentation of the skin, and heart failure. Prr (HCC; [114550](#)), complicating cirrhosis, is responsible for about one-third of deaths in affected homozygotes. Since hemochromatosis is a relati diagnosed, this is a form of preventable cancer. ☺

Links

- Bo
- Gene
- GEO Profiles
- HomoloGene
- OMIA
- Free in PMC
- PubMed (calculated)
- PubMed (cited)
- Gene Genotype
- GeneView in dbSNP
- UniGene
- Related Entries
- Nucleotide
- Protein
- SNP
- Structure

Genomic regions, transcripts, and products

Go to [reference sequence details](#)

**NC\_000006.10**

[ 26195427 ] 5' [ 26205038 ] 3'

NM\_139005.2  
 NM\_139002.2  
 NM\_00410.3  
 NM\_139004.2  
 NM\_139003.2  
 NM\_139009.2  
 NM\_139007.2  
 NM\_139010.2  
 NM\_139011.2  
 NM\_139006.2

NP\_620574 isoform 5 precursor  
 NP\_620571 isoform 2 prccu  
 NP\_004401 isoform 1 prccu  
 NP\_620573 isoform 4 prccu  
 NP\_620572 isoform 3 prccu  
 NP\_620578 isoform 9 prccu  
 NP\_620576 isoform 7 prccu  
 NP\_620577 isoform 8 prccu  
 NP\_620579 isoform 10 prccu  
 NP\_620580 isoform 11 prccu  
 NP\_620575 isoform 6 prccu

- coding region  
 - untranslated region

Links

- FASTA
- GENE
- Blink
- Conserved Domains

Genomic context

chromosome: 6; Location: 6p21.3

See HFE in MapViews:

[ 26153618 ] HIST1H3C [ 26216343 ] HIST1H4C

HIST1H3C HFE HIST1H4C

Map Viewer

- Nucleotide
- OMIA
- OMIM
- Full text in PMC
- Probe
- Protein
- PubMed
- PubMed (GeneRIF)
- SNP
- SNP: Genotype
- SNP: GeneView
- Taxonomy
- UniSTS
- AcView
- CCDS
- Evidence Viewer
- GDB
- GeneTests for MIM: 235200
- HGMD
- HSNC
- HPRD
- KEGG
- MGC
- ModelMaker
- UniGene
- LinkOut

Entrez Gene Info

Feedback

NCBI

BLAST Protein Structure PubMed Taxonomy  
Genome Nucleotide 3D-Domains Books Help

Query: gi|4504377 hemochromatosis protein isoform 1 precursor [Homo sapiens]  
Matching gi: 109658670, 109658506, 83323630, 80748852, 57114069, 38502807, 29709343, 22854810, 20250786, 15115850, 14100030, 11094315, 2497915, 2370111, 2088551, 1890180, 1469790

Hide identical Best hits Common Tree Taxonomy Report 3D structures CDD-Search Gl list Run BLAST

200 BLAST hits to 25 unique species [Sort by taxonomy proximity](#)

0 Archaea 0 Bacteria 199 Metazoa 0 Fungi 0 Plants 0 Viruses 0 Other Eukaryotae

Keep only  Cut-Off 100 Select Reset New search by GI: 4504377 Go

348 aa

SCORE	P	ACCESSION	GI	PROTEIN DESCRIPTION
<a href="#">Conserved Domain Database hits</a>				
1870	31	AAI17202	109658670	Hemochromatosis [Homo sapiens]
1870	31	AAI17204	109658506	Hemochromatosis [Homo sapiens]
1870	29	NP_001...	57114069	hemochromatosis protein [Pan troglodytes]
1870	29	P60018	38502807	Hereditary hemochromatosis protein homolog precursor (HLA-H)
1870	29	AA009793	22854810	hereditary hemochromatosis [Pan troglodytes]
1870	31	AAG29572	11094315	hemochromatosis termination variant terE6; HFE [Homo sapiens]
1870	31	Q30201	2497915	Hereditary hemochromatosis protein precursor (HLA-H)
1870	31	CAA70934	2370111	HFE [Homo sapiens]
1870	31	AAB82083	2088551	hereditary hemochromatosis [Homo sapiens]
1870	31	CAB07442	1890180	HFE [Homo sapiens]
1870	31	AAC51823	1469790	HLA-H
1776	31	AHH74721	50960016	HFE protein [Homo sapiens]
1772	31	NP_620575	21040347	hemochromatosis protein isoform 6 precursor [Homo sapiens]
1772	31	AAC62646	3695107	hemochromatosis splice variant del14E4 [Homo sapiens]
1713	31	NP_620578	21040353	hemochromatosis protein isoform 9 precursor [Homo sapiens]
1713	31	CAC67792	15485419	hemochromatosis protein [Homo sapiens]
1625	26	XP_001...	109069870	PREDICTED: hemochromatosis isoform 6 [Macaca mulatta]
1517	31	1DE4G	6980500	Chain G, Hemochromatosis Protein Hfe Complexed With Transferrin
1517	31	1DE4D	6980497	Chain D, Hemochromatosis Protein Hfe Complexed With Transferrin
1517	31	1DE4A	6980494	Chain A, Hemochromatosis Protein Hfe Complexed With Transferrin

NCBI

BLAST Protein Structure PubMed Taxonomy  
Genome Nucleotide 3D-Domains Books Help

Query: gi|4504377 hemochromatosis protein isoform 1 precursor [Homo sapiens]  
Matching gi: 109658670, 109658506, 83323630, 80748852, 57114069, 38502807, 29709343, 22854810, 20250786, 15115850, 14100030, 11094315, 2497915, 2370111, 2088551, 1890180, 1469790

[Get Cn3D Now!](#)

Hide identical Best hits Common Tree Taxonomy Report 3D structures CDD-Search Gl list Run BLAST

200 BLAST hits to 3 unique species [Sort by taxonomy proximity](#)

0 Archaea 0 Bacteria 200 Metazoa 0 Fungi 0 Plants 0 Viruses 0 Other Eukaryotae

Keep only  Cut-Off 100 Select Reset New search by GI: 4504377 Go

348 aa

SCORE	P	ACCESSION	GI	PROTEIN DESCRIPTION
<a href="#">Conserved Domain Database hits</a>				
1517	31	1DE4G	6980500	Chain G, Hemochromatosis Protein Hfe Complexed With Transferrin F
1517	31	1DE4D	6980497	Chain D, Hemochromatosis Protein Hfe Complexed With Transferrin F
1517	31	1DE4A	6980494	Chain A, Hemochromatosis Protein Hfe Complexed With Transferrin F
1517	31	1A62C	4699712	Chain C, Hfe (Human) Hemochromatosis Protein
1517	31	1A62A	4699710	Chain A, Hfe (Human) Hemochromatosis Protein
525	31	1B1IA	3891929	Chain A, The Crystal Structure Of H-2dd Mhc Class I In Complex With
507	31	1S7TD	48425604	Chain D, Crystal Structures Of The Murine Class I Major Histocomp
507	31	1S7TA	48425601	Chain A, Crystal Structures Of The Murine Class I Major Histocomp
507	31	1S7SA	48425598	Chain A, Crystal Structures Of The Murine Class I Major Histocomp
507	31	1S7RD	48425595	Chain D, Crystal Structures Of The Murine Class I Major Histocomp
507	31	1S7RA	48425592	Chain A, Crystal Structures Of The Murine Class I Major Histocomp
507	31	1S7QA	48425589	Chain A, Crystal Structures Of The Murine Class I Major Histocomp
502	31	2BCKD	88192434	Chain D, Crystal Structure Of Hla-A2402 Complexed With A Telomere
502	31	2BCKA	88192431	Chain A, Crystal Structure Of Hla-A2402 Complexed With A Telomere
502	31	1X7QA	73535522	Chain A, Crystal Structure Of Hla-A1101 With Sars Nucleocapsid Pe
502	31	1QVOD	49258587	Chain D, Structures Of Hla-A1101 In Complex With Immunodominant N
502	31	1QVOA	49258584	Chain A, Structures Of Hla-A1101 In Complex With Immunodominant N

NCBI **Related Structures**

HOME SEARCH SITE MAP PubMed Blast Entrez Structure Help

**Query:** hemochromatosis protein isoform 1 precursor [Homo sapiens]  
[gi: 4504377]

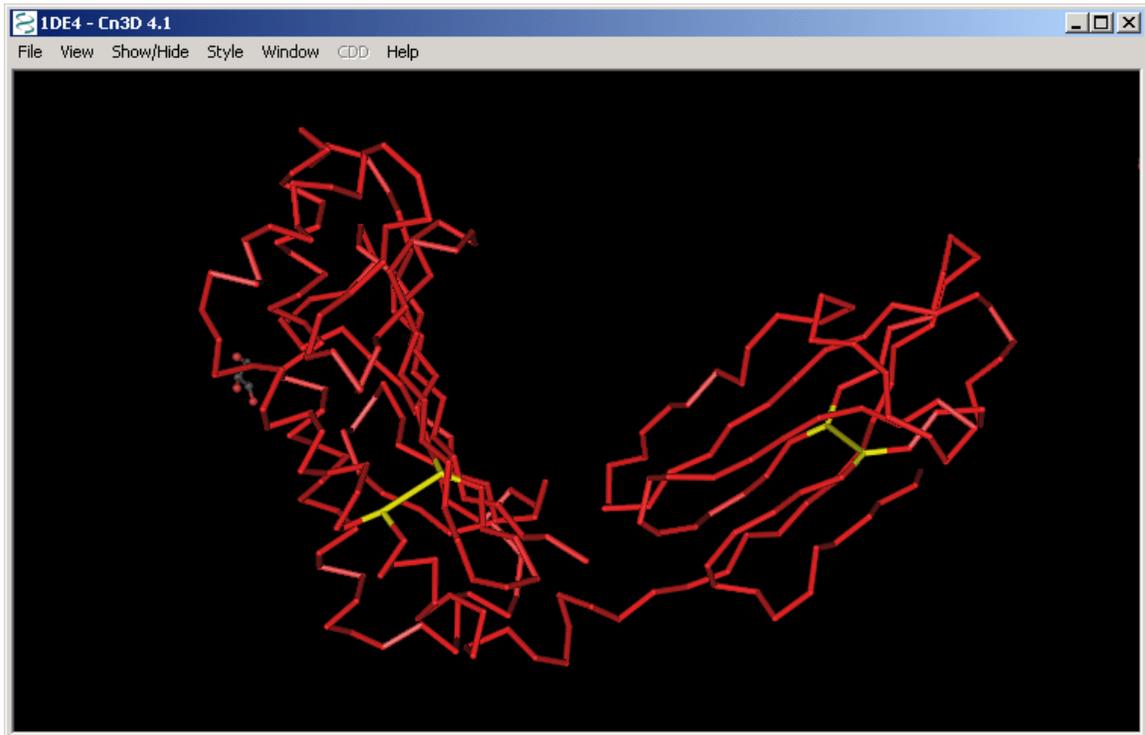
**Structure:** 1DE4 Chain G, Hemochromatosis Protein Hfe Complexed With Transferrin Receptor

**Reference:** [MMDB] [PubMed]

Get 3D Structure data to: View in Cn3D (To display structure, download Cn3D)

E-value = 7e-168, Bit score = 588, Aligned length = 275, Sequence Identity = 100%

		10	20	30	40	50	60	70	80					
gi 4504377	23	RLLRSHSLHYL	FMGASEQDLG	LGLS	LFEALGYVDD	QLFV	FDHESRRVE	PRT	PWSSR	ISSQMWLQ	LSQSLK	GWDHMF	TVDF	102
1DE4 G	1	RLLRSHSLHYL	FMGASEQDLG	LGLS	LFEALGYVDD	QLFV	FDHESRRVE	PRT	PWSSR	ISSQMWLQ	LSQSLK	GWDHMF	TVDF	80
		90	100	110	120	130	140	150	160					
gi 4504377	103	WTIMENHNH	SKESTLQV	ILGCENQED	NSTEGYWKY	GYDGDH	LEFCP	DTLDWRAA	EPR	AWPTKLE	WERH	KIRARQ	NRAY	182



1DE4 - Sequence/Alignment Viewer

View Edit Mouse Mode Unaligned Justification Imports

1DE4 G	TSSVTT	LR	C	R	A	L	N	Y	P	Q	N	I	T	M	K	W	L	K	D	K	Q	P	M	D	A	K	E	F	E	P	K	D	V	L	P	N	G	D	G	T	Y	Q	G	W	I	T	L	A	V	P	P	G	E	E	Q	R	Y	T	C	Q	V	E	H	P	G	L	D	Q	P	L	I	V	W	~	~
gi 4504377	TSSVTT	LR	C	R	A	L	N	Y	P	Q	N	I	T	M	K	W	L	K	D	K	Q	P	M	D	A	K	E	F	E	P	K	D	V	L	P	N	G	D	G	T	Y	Q	G	W	I	T	L	A	V	P	P	G	E	E	Q	R	Y	T	C	Q	V	E	H	P	G	L	D	Q	P	L	I	V	W	e	p

gi 4504377, loc 282 Block 1, Row 2

**Problem 2:**

<http://www.ncbi.nlm.nih.gov/Class/minicourses/pheno2.html>

Mutations in the HBB gene are associated with sickle cell anemia. A laboratory working on sickle cell anemia wants to elucidate the biochemical and structural basis for the function of the mutant HBB protein.

***Step 1. Determining what is known about the HBB gene and protein (using Entrez Gene):***

Search for 'HBB' in [Entrez Gene](#). One entry is for the human HBB gene. Retrieve the entry by clicking on the HBB link.

What is the location and orientation of the HBB gene on the human genome? List the genes adjacent to it. How many alternatively spliced products have been annotated for the HBB gene when the RefSeq mRNA entries were reviewed? List some of the HBB gene aliases. What are the phenotypes associated with the mutations in the HBB gene? Where are the mouse and rat HBB genes located?

What is the name and function of the protein encoded by the HBB gene? What is the conserved domain in the protein? To which cellular component(s) is the protein localized? Beta hemoglobin is a subunit of which protein? Name other subunit(s) in that protein.

Obtain the locations of exons and introns for each transcript by choosing "Gene Table" from the Display pull down menu. Go back to the description page.

***Step 2. Determining other identified SNPs and their locations in the HBB gene:***

From the Links menu on the top right hand side of the page, click on the "SNP: GeneView" to access a list of the known SNPs (reported in dbSNP). By default, the SNPs in the coding region of a gene are reported. Additional SNPs such as in the upstream region or the introns can be viewed by clicking on the "in gene region" button. Currently, how many non-synonymous SNPs are placed on the beta hemoglobin transcript NM\_000518? How many of these have links to OMIM? We will concentrate on the Glu7Val mutant in the following analysis.

***Step 3. Learning more about sickle cell anemia disease and its genetic testing:***

Go back to the Entrez Gene report. Click on the OMIM link and then HBB link. What are the phenotypes caused by mutations in HBB, the absence of HBB and reduced amounts of HBB? What is the clinical synopsis of sickle cell anemia? What is its prominent feature? What is its mode of inheritance? How many allelic variants of the HBB gene have been reported? As mentioned in the OMIM report, the allelic variants are listed for the mature beta hemoglobin protein which lacks

an initiator methionine. Hence, the allelic variants in the OMIM report are off by one amino acid compared to the precursor protein in NP\_000509. Click on the Allelic Variant “View list” to get information about the mutant proteins from patients. Is the Glu6Val variant mentioned in the list? (It is the variant number 0243). Which phenotype does it cause? What is the name of the mutant hemoglobin (hemoglobin S).

Click on the Gene Tests link at top of the page. Identify some of the laboratories performing the clinical testing for sickle cell anemia. Now refer to the Reviews section for Sickle Cell Disease, Mutation analysis is available for which of the HBB alleles? List one explanation for the sickle cell anemia phenotype caused by the Glu7Val mutant beta hemoglobin.

#### ***Step 4. Elucidating the biochemical and structural basis for the function of the wild type and mutant proteins, if possible:***

##### **A. Information about the wild type protein**

Go back to the OMIM report by clicking the back button on the web browser. Go to the Gene report through the Links menu. Based on the RefSeq summary and the PubMed articles, describe the biochemical functions of beta hemoglobin and hemoglobin S. PubMed articles in the Entrez Gene report indicate that the 3-D structure of hemoglobin S is available.

Let us first take a look at the structure of the wild type protein. Click on the NP\_000509 protein link and select Blink. Click on the “Show identical” button and then on the “3D structures” button. The output contains a list of similar proteins with 3D structures known. The entry, 1DXTD, represents the structure of deoxyhemoglobin chain D. Click on the blue dot next to 1DXTD to get the sequence alignment of the query protein to the D chain of 1DXTD. To view the 3D structure of dexoxyhemoglobin (all chains, 2 alpha and 2 beta), click on the MMDB link. That takes us to the MMDB structure summary page for 1DXT. Access the PDB entry, by clicking on 1DXT. Note that the chains A and C in the structure represent alpha chains, and B and D represent beta chains. Go back to the MMDB summary page. View the deoxyhemoglobin tetramer by clicking on the "View 3D Structure button".

Search for the structure of the mutant (deoxyhemoglobin S) in the structure database. Two entries, 1HBS and 2HBS, are retrieved. Click on the 2HBS link. Then click on the PubMed link from the MMDB and PDB entries (under Reference). The abstracts indicate that the mutated valine residue of the beta chain contacts with another hemoglobin tetramer molecule to form hemoglobin polymers which are building blocks for the sickle cell fiber.

**B. To show the side chains of the mutant residue and view its interaction with another hemoglobin molecule:** Download the structure 2HBS by clicking on View 3D Structure. For easier viewing, remove the helix and strand objects using Style--Edit global style, and unclick the boxes next to the Helix objects and Strand objects. Highlight valine 6 from the H chain (one of the beta chains). To show the side chains of the residue, use the Structure window--Style--Annotate--new. Give a name to this annotation such as "valine" and then click on Edit Style. Change the protein backbone "Rendering" to "Space Fill", Color Scheme to "charge" or "hydrophobicity". Repeat these steps for the Protein Sidechains row and click the Protein Sidechains on. To show the amino acid number, choose the Labels panel, and change the Protein Backbone spacing to 1. Click on the "Done", "OK" then "Done" buttons. The valine interacts with a pocket between the two helices on another tetramer. Identify the residues from other molecules within 4 angstroms of the valine, use Show/Hide--Select by distance--other molecules. To unselect the highlighted residues, click on the white portion of the sequence window.

***You can now easily explain why the Glu7Val mutant has an altered function.***

**Summary:**

This mini-course describes how to obtain information about the HBB gene, known SNPs in it, and elucidate the biochemical and structural basis for the function of the wild type and Glu7Val mutant protein.

- Summary:
1. The HBB gene is located on chromosome 11 and has no alternatively spliced products annotated.
  2. Currently, there are 7 non-synonymous SNPs annotated on the protein NP\_000509.
  3. The Glu7Val mutant is associated with the sickle cell anemia disease and the site of mutation is used in sickle cell anemia genetic testing.
  4. The HBB gene encodes beta hemoglobin which is a part of hemoglobin along with alpha hemoglobin. Hemoglobin is a tetramer consisting of 2 beta and 2 alpha chains. Mutation of the 7th negatively charged amino acid, glutamic acid, to hydrophobic valine leads to polymerization of hemoglobin forming a sickle fiber that changes the shape of red blood cells leading to sickle cell anemia.