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.


Instructor: Medha Bhagwat (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.
Gene: HFE

**Official Symbol:** HFE

**Official Full Name:** hemochromatosis

**Other Aliases:** HFE1, HHC, HLA-H, MGC:13790, 6p21.31

**Other Designations:** MHC class I-like protein HFE, hemochromatosis protein, hereditary hemochromatosis protein, HLA-H

**Chromosome:** 6, **Location:** 6p21.3

**Gene ID:** 3077

**Summary:**

The protein encoded by this gene is a membrane protein that is similar to MHC class I-like proteins and associates with beta2-microglobulin (B2M). 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 hemochromatosis, 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 context

chromosome: 6, Location: 6p21.3

Bibliography

Related Articles in PubMed

GeneRIFs: Gene References Into Function

1. determined race-specific frequencies of the HFE mutations, C282Y and H63D
2. Glucose intolerance may be important risk factor for the development of hepatic fibrosis in subjects with the C282Y/H63D HFE genotype.
3. Potential interaction between HFE genotypes and hema iron intake in relation to the risk of type 2 diabetes

Interactions

Product       Interactant              Other Gene  Complex  Source  Pubs
NP_000401.1   Beta 2 microglobulin      BPM        HFEQ      PubMed
NP_000401.1   Transferrin receptor 2   TFR2       HFEQ      PubMed
NP_000401.1   NP_000255.1             TFRC       HFEQ      PubMed

General gene information

Markers

RH66790(e-POR)
Links: Unigene:18176
Alternate name: s1505246998

W1-175446(e-POR)
Links: Unigene:30516
Alternate name: EST251388; RH51056

RH66637(e-POR)
Links: Unigene:30031
Alternate name: s1650244673

A0B8H22K(e-POR)
Links: Unigene:4134
Alternate name: RHE5914

STS-U660419(e-POR)
Links: Unigene:47384
Alternate name: RHE5999; s15-U02229

D6622577(e-POR)
Links: Unigene:57170
Alternate name: GEO:5584195; sy095g1-19

Phenotypes

Hemochromatosis
OMIM: 228600
Porphyria variegata
OMIM: 127600
<table>
<thead>
<tr>
<th>mRNA</th>
<th>bp</th>
<th>exons</th>
<th>Protein</th>
<th>aa</th>
<th>exons</th>
</tr>
</thead>
<tbody>
<tr>
<td>NM_139005.2</td>
<td>1417</td>
<td>5</td>
<td>NP_625574.1</td>
<td>277</td>
<td>5</td>
</tr>
<tr>
<td>NM_139006.2</td>
<td>678</td>
<td>4</td>
<td>NP_625571.1</td>
<td>163</td>
<td>4</td>
</tr>
<tr>
<td>NM_139007.2</td>
<td>822</td>
<td>6</td>
<td>NP_625561.1</td>
<td>340</td>
<td>6</td>
</tr>
<tr>
<td>NM_139008.2</td>
<td>1946</td>
<td>5</td>
<td>NP_625573.1</td>
<td>257</td>
<td>5</td>
</tr>
<tr>
<td>NM_139009.2</td>
<td>1904</td>
<td>5</td>
<td>NP_625572.1</td>
<td>243</td>
<td>5</td>
</tr>
<tr>
<td>NM_139010.2</td>
<td>8135</td>
<td>6</td>
<td>NP_625578.1</td>
<td>359</td>
<td>6</td>
</tr>
<tr>
<td>NM_139011.2</td>
<td>1988</td>
<td>5</td>
<td>NP_625580.1</td>
<td>261</td>
<td>5</td>
</tr>
<tr>
<td>NM_139012.2</td>
<td>1910</td>
<td>5</td>
<td>NP_625577.1</td>
<td>247</td>
<td>5</td>
</tr>
<tr>
<td>NM_139013.2</td>
<td>1652</td>
<td>4</td>
<td>NP_625579.1</td>
<td>160</td>
<td>4</td>
</tr>
<tr>
<td>NM_139014.2</td>
<td>1406</td>
<td>3</td>
<td>NP_625581.1</td>
<td>77</td>
<td>3</td>
</tr>
<tr>
<td>NM_139015.2</td>
<td>1180</td>
<td>6</td>
<td>NP_625575.1</td>
<td>335</td>
<td>6</td>
</tr>
</tbody>
</table>

Exon Information:

- **NM_139005.2**: length: 1417 bp, number of exons: 5
- **NP_625574.1**: length: 277 aa, number of exons: 5

### EXON
- **Coding EXON**:
- **INTRON**:

<table>
<thead>
<tr>
<th>coords</th>
<th>length</th>
<th>coords</th>
<th>length</th>
<th>coords</th>
<th>length</th>
</tr>
</thead>
<tbody>
<tr>
<td>62 - 297</td>
<td>235 bp</td>
<td>252 - 297</td>
<td>26 bp</td>
<td>293 - 292</td>
<td>324 bp</td>
</tr>
<tr>
<td>3652 - 3888</td>
<td>234 bp</td>
<td>3722 - 3888</td>
<td>264 bp</td>
<td>3886 - 4004</td>
<td>238 bp</td>
</tr>
<tr>
<td>4005 - 4087</td>
<td>276 bp</td>
<td>4095 - 4087</td>
<td>276 bp</td>
<td>4087 - 4165</td>
<td>278 bp</td>
</tr>
<tr>
<td>5466 - 5667</td>
<td>262 bp</td>
<td>5466 - 5667</td>
<td>262 bp</td>
<td>5667 - 5871</td>
<td>205 bp</td>
</tr>
<tr>
<td>2072 - 2080</td>
<td>13 bp</td>
<td>2072 - 2080</td>
<td>13 bp</td>
<td>2080 - 2093</td>
<td>13 bp</td>
</tr>
</tbody>
</table>

- **NM_139006.2**: length: 678 bp, number of exons: 4
- **NP_625571.1**: length: 162 aa, number of exons: 4

### EXON
- **Coding EXON**:
- **INTRON**:

<table>
<thead>
<tr>
<th>coords</th>
<th>length</th>
<th>coords</th>
<th>length</th>
<th>coords</th>
<th>length</th>
</tr>
</thead>
<tbody>
<tr>
<td>62 - 297</td>
<td>236 bp</td>
<td>262 - 297</td>
<td>76 bp</td>
<td>299 - 301</td>
<td>302 bp</td>
</tr>
<tr>
<td>2652 - 2685</td>
<td>24 bp</td>
<td>2662 - 2685</td>
<td>24 bp</td>
<td>2685 - 4004</td>
<td>203 bp</td>
</tr>
</tbody>
</table>
### gene model (contig mRNA transcript)

<table>
<thead>
<tr>
<th>Region</th>
<th>Contig</th>
<th>mRNUA pos</th>
<th>dbSNP rs no.</th>
<th>chr position in</th>
<th>Heterozygosity</th>
<th>Validation</th>
<th>3D OMMIM</th>
<th>Function</th>
<th>dbSNP allele</th>
<th>Protein residue</th>
<th>Codon pos</th>
</tr>
</thead>
<tbody>
<tr>
<td>exon_2</td>
<td>16949347</td>
<td>325</td>
<td>NA</td>
<td>1</td>
<td>N.D.</td>
<td>H: Yes</td>
<td>nonsynonymous</td>
<td>C</td>
<td>Thr [T]</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>17292954</td>
<td>2</td>
<td>N.D.</td>
<td>H: Yes</td>
<td>nonsynonymous</td>
<td>G</td>
<td>Asp [D]</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>18007078</td>
<td>3</td>
<td>N.D.</td>
<td>H: Yes</td>
<td>nonsynonymous</td>
<td>T</td>
<td>Ser [S]</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>18282458</td>
<td>4</td>
<td>N.D.</td>
<td>Yes</td>
<td>nonsynonymous</td>
<td>A</td>
<td>Thr [T]</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>18333459</td>
<td>5</td>
<td>N.D.</td>
<td>Yes</td>
<td>nonsynonymous</td>
<td>F</td>
<td>Ile [I]</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>exon_3</td>
<td>16949333</td>
<td>602</td>
<td>NA</td>
<td>1</td>
<td>N.D.</td>
<td>Yes</td>
<td>nonsynonymous</td>
<td>C</td>
<td>His [H]</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>23445225</td>
<td>2</td>
<td>N.D.</td>
<td>Yes</td>
<td>nonsynonymous</td>
<td>A</td>
<td>Glu [Q]</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>exon_4</td>
<td>16951321</td>
<td>8371</td>
<td>14289955</td>
<td>1</td>
<td>0.005</td>
<td>H: Yes</td>
<td>nonsynonymous</td>
<td>T</td>
<td>Ile [I]</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>16951332</td>
<td>1066</td>
<td>13020562</td>
<td>3</td>
<td>0.043</td>
<td>H: Yes</td>
<td>nonsynonymous</td>
<td>A</td>
<td>Tyr [Y]</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>
In patients with hemochromatosis, Faller et al. (1996) identified an 840-A transition in the HFE gene (which they referred to as A840G or C282Y), resulting in a cysteine to tyrosine (C282Y) substitution. This missense mutation occurs as a highly conserved residue involved in the indispensable dimerization of HFE-C282Y protein, and could thereby disrupt the structure and function of this protein. Using allele-specific oligonucleotide linkage analysis on 140 patients, they detected the C282Y mutation in 65% of the HFE chromosomes. In 100 of the 140 patients (71%), the C282Y mutation was identified, which is a frequency of 57.1%. By contrast, the normal allele was found in only 28.8% of the 100 alleles. These results were consistent with findings in previous studies. The allele frequencies were significantly different from the Hardy-Weinberg equilibrium. The findings confirmed heterogeneity among the hemochromatosis patients, with 83% of cases linked to C282Y heterozygosity.

Jacquemont et al. (1996) provided convincing evidence that the C282Y mutation is homogenous in the HFE gene in the group of hemochromatosis. In studies in Australia, patients were more frequently characterized by the genotype and phenotype level all showed homozygosity for the C282Y allele and heterozygosity for the HFE allele. The presence of a single mutation in all patients contrasted with the data of Faller et al. (1996), who reported a lower frequency of the mutation. Jacquemont et al. (1996) suggested that a different racial origin for the diagnosis of HFE may account for these differences. Furthermore, the data showed that the HFE locus is the primary HFE locus, but that there is not likely to be other linked modifying genes that may explain both the HLA-linked holoerythroderma and the inactivation of the disease and the large region of linkage disequilibrium present in all populations and spanning at least 4.5 Mb of the HFE locus.

Jacquemont et al. (1996) concentrated 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 by stringent criteria. Heterozygosity for the C282Y mutation was found in 59 of 65 patients (90.7%). Only 2 of the patients were homozygous for the C282Y mutation and the HFE mutation (95.8% and 9.8%) and 1 was homozygous for the HFE mutation, and 2 were heterozygous for HFE. These results corresponded to an allele frequency of 95.8% for the C282Y and 9.8% for the HFE mutation, respectively. Of these, the C282Y mutation was more frequently observed in the null mutation context, while it was present in 5.7% of the general French population. To further analyze the HFE allele frequency in the white control groups, the null mutation context was divided into control groups, 9.7% of the general control groups were found to have C282Y mutations. These results were consistent with the studies of Faller et al. who showed that the C282Y mutation is homogenous in the HFE gene. The authors concluded that the C282Y mutation is the primary HFE locus, and that there is not likely to be other linked modifying genes that may explain both the HLA-linked holoerythroderma and the inactivation of the disease and the large region of linkage disequilibrium present in all populations and spanning at least 4.5 Mb of the HFE locus.

Benoist et al. (1996) reported mutation analysis of 147 patients with hereditary hemochromatosis and 30 controls. Of the 147 patients, 62 (42.4%) were homozygous for the C282Y mutation, while 100 (68.4%) were heterozygous for the C282Y mutation. Of the 147 patients, 62 (42.4%) were homozygous for the C282Y mutation. The frequency of the C282Y mutation was significantly higher in patients with hereditary hemochromatosis compared to controls. The authors concluded that the C282Y mutation is the primary HFE locus, and that there is not likely to be other linked modifying genes that may explain both the HLA-linked holoerythroderma and the inactivation of the disease and the large region of linkage disequilibrium present in all populations and spanning at least 4.5 Mb of the HFE locus.
**HEMOCHROMATOSIS; HFE**

**ALLELIC VARIANTS**

*selected examples*

- **0001 HEMOCROMATOSIS** [HFE, CYS282TYR] dbSNP
- **0002 HEMOCROMATOSIS** [HFE, H663ASAP] dbSNP
- **0003 HEMOCROMATOSIS** [HFE, SER65CYS] dbSNP
- **0004 HFE INTRONIC POLYMORPHISM** [HFE, 2559G-A]
- **0005 HFE POLYMORPHISM** [HFE, VAL59MET] dbSNP
- **0006 HFE POLYMORPHISM** [HFE, VAL59MET] dbSNP
- **0007 PORPHYRIA VARIEGATA** [HFE, GLN127HIS] dbSNP
- **0008 HEMOCROMATOSIS** [HFE, ARG380ARG]
- **0009 HEMOCROMATOSIS** [HFE, ILE105THR] dbSNP
- **0010 HEMOCROMATOSIS** [HFE, GLN93ARG] dbSNP
- **0011 HEMOCROMATOSIS** [HFE, GLN283PRO]

---

**GeneTests, Links**

---

**GeneTests**

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
**HFE- Associated Hereditary Hemochromatosis**

Labs offering clinical testing:
- ARUP Laboratories, Inc.
- ARUP Laboratories
- Salt Lake City, UT
- Elaine Lyer, PhD; Kong Mao, MD; Edward R Ashwood, MD; Marzia Pandolui, PhD
- Ashademi Healthcare Group
- Ashademi Genetic Diagnostic Center
- Istanbul, Turkey
- Bander Alhok, MD, PhD
- Alberta Children's Hospital
- Molecular Diagnostic Laboratory
- Calgary, Alberta, Canada
- Peter Bridge, PhD, FCOMG, FACMG; Jilian Parboosingh, PhD, FCOMG

**Select all clinical laboratories**

- Sequencing of entire coding region
- Sequencing of select exons
- Microdeletion analysis
- Prenatal diagnosis
- Postnatal diagnosis
- Clinical confirmation of genotype identified in a research lab

---

**Search Result for OMIM # 235200**

HFE- Associated Hereditary Hemochromatosis

[Testing] [Research] [Reviews] [Resources]
HFE-Associated Hereditary Hemochromatosis

**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 40 and 60 years of age and

Normal allelic variants: A serine at position 65 to cysteine (S65C) has been identified. The effect of this mutation is unclear.

Pathologic allelic variants: Two missense mutations have been identified, a cysteine at position 282 to tyrosine (C282Y) and histidine at position 63 to aspartate (H63D).

- Cys282 Tyr (synonym: C282Y; nucleotide 845G>A) This missense mutation removes a highly conserved cysteine residue that normally forms an intramolecular disulfide bond, and thereby prevents the protein from being expressed on the cell surface.
- His63 Asp (synonym: H63D; nucleotide 187C>G) This missense mutation may impair interaction of the HFE-encoded protein with the transferrin receptor on the cell surface.

Normal gene product: A cell-surface protein of 321 amino acids with sequence similarity to HLA class I molecules. The normal protein forms a heterodimer with beta-2-microglobulin, and this interaction is necessary for normal presentation on the cell surface. The normal protein binds to the transferrin receptor, and may act by modulating its affinity for transferrin.

Abnormal gene product: An impaired cell-surface protein is apparently formed. This protein does not migrate to the cell surface and does not bind transferrin (bound to ferric iron). Therefore, lack of internalization of transferrin into the small bowel absorptive cell may lead to compensatory increase in iron absorption [Biscovec et al 1999].
The hemochromatosis protein functions to regulate iron absorption by regulating the interaction of the transferrin receptor with transferrin.
Problem 2:

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):**


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 deoxyhemoglobin (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.