疾患遺伝子と表現型の相関関係

このミニコースでは、ある疾患と、その関連遺伝子について知られていることをどのように見つけ出すかその方法 について学びます。またタンパク質の変異体が引き起こす表現型生化学的、構造的な基礎を説明していきます。

課題

このミニコースでは疾患遺伝子とその表現型の相関関係について焦点を当ててみていきます。以下の各ステップでは、文献、発現、構造情報といったNCBIのリソースがどのように疾患遺伝子に関して予想される機能的な情報を得るのにどう役立つのか、紹介していくことにします。

HFE遺伝子の突然変異は、血色素症と関連しています。血色素症の病気について研究しているある研究室では、変異タンパク質の働きに関する生化学的、構造的な基礎を説明していきたいと思っています。

概要

この例題では、以下のような目標が設定されています。

- 1. HFE遺伝子とタンパク質についてわかっていることを見つけ出します。 (Entrez Geneを使います)
- 2. HFE遺伝子に存在する既知のSNPsとその位置を見つけ出します。 (dbSNPを使います)
- 3. 血色素症とその遺伝子検査について様々なことを学びます。 (OMIMやGene Testsを使います)
- 4. できたら野生型タンパク質、変異型タンパク質の機能に関して、生化学的、構造的な基礎を 説明していきます。

Step.1 HFE遺伝子とタンパク質についてわかっていることを見つけ出す(Entrez Geneの利用)

■ Step. 1-1_

Entrez Geneを用いて "HFE"を検索して下さい。ヒトのHFE遺伝子について一つのエントリが見つかるはずです。 HFEと書かれたリンクをクリックして、その情報を取得してください。

■ Step. 1-2_

ヒトゲノム上で、HFE遺伝子の位置や方向はどのようになっていますか?隣接した遺伝子をリストアップしてみて下さい。 RefSeqのmRNAに注目した場合、HFE遺伝子にはいくつの選択的スプライス転写産物がアノテーションされているでしょうか? 選択的スプライス転写産物の間にはどのような違いがありますか? HFEの遺伝子の別名をリストアップして下さい。 HFEの遺伝子の変異と関連付けられる表現型にはどのようなものがあるでしょうか?

■ Step. 1-3_

HFE遺伝子がコードしているタンパク質の名前と機能は何でしょうか? HFEタンパク質に保存されたドメイン領域にはどのようなものがあるでしょうか?どの細胞分画にHFEタンパク質は局在化されているのでしょうか?

■ Step. 1-4

Displayのプルダウンメニューから"Gene Table"を選んで、それぞれのトランスクリプトに関して、エクソンとイントロンの位置を求めて下さい。

Step. 2 HFE遺伝子に存在する既知のSNPsとその位置を見つける

■ Step. 2-1_

ページ右側の上部にあるLinkメニューから、既知のSNPsのリスト (dbSNPに含まれている) にアクセスするため に"SNP:GeneView"をクリックして下さい。初期設定状態では遺伝子のコード領域にあるSNPが表示されます。 "in gene region"ボタンをクリックすると、遺伝子上流領域あるいはイントロンに存在するような SNPを追加表示する事が出来ます。

■ Step. 2-2

今、一番長い血色素症の一番長い遺伝子転写産物であるNM_000410上にいくつの非同義SNPが確認されているでしょうか?それらのうちいくつがOMIMへとリンクされているでしょうか?以下の解析ではcys282tyr変異に注目していきましょう。

Step.3 血色素症とその遺伝子検査について様々なことを学ぶ

■ Step. 3-1

SNPレポートの中で、どれかひとつのSNP情報にあるOMIMリンクをクリックして下さい。血色素症の臨床兆候は何でしょうか?血色素症と呼ばれる鉄過剰疾患の5つの種類をリストアップして下さい。それらの疾患のうち、HFE遺伝子の変異と関連しているのはどれでしょうか。 HFE遺伝子の対立遺伝子多型はいくつ報告されているでしょうか? Cys282Tyr変異体と関連している表現型は何でしょうか?

■ Step. 3-2_

ページの一番上にある"GeneTests"へのリンクをクリックして下さい。血色素症に関して臨床試験を行っている研究室をいくつか確認して下さい。では、次にReviewsを参照して下さい。どのHFE遺伝子の対立遺伝子に対して変異体解析が利用出来るでしょうか? Cys282Tyr変異体によって生じた血色素症の表現型についての説明を一つリストアップして下さい。

Step. 4 できたら野生型タンパク質、変異型タンパク質の機能に関して、生化学的、構造的な基礎を説明していく

■ Step. 4-1

Entrez Geneのページに戻って下さい。タンパク質で一番上にあるNP_000401をクリックして下さい。そして、BLinkをクリックして下さい。このタンパク質配列上で保存されたドメインおよび、それらが含まれているデータベースを挙げて下さい。

3D structuresと書かれているボタンをクリックして下さい。表示されている結果は、NP_000401と配列が似ていて、立体構造が既知であるを持つタンパク質のリストになっています。一番初めにエントリーされている1DE46は、血色素症タンパク質のG鎖に相当します。(トランスフェリン受容体と複合体を形成します) 1DE4のG鎖と問い合わせのアミノ酸配列とのアライメントを得るため、 1DE4Gの付近にある青いドットをクリックして下さい。

Step. 4-2_

"Get 3D Structure data"をクリックして下さい。これによって、1DE4のG鎖の構造情報と問い合わせアミノ酸配列との配列アライメントデータがダウンロードされ、Cn3Dによる可視化が可能となります。キーボードの"z"を押して、ジスルフィド結合(黄褐色の色)の部分を拡大して下さい。それらをダブルクリックして、ジスルフィド結合を形作るシステイン残基を選択して下さい。別ウィンドウで表示されているアライメントの画面上で、今選択したシステイン残基に該当する文字上にマウスを重ねることでアミノ酸番号を表示して下さい。それらの残基のうちの一つが282番目にあるシステインとなっています。血色素症を引き起こす原因となるチロシンへと変異したのと同じシステインです。

これで、あなたも何故C282Y変異体が変化した機能を持っているか簡単に説明できますね

要旨

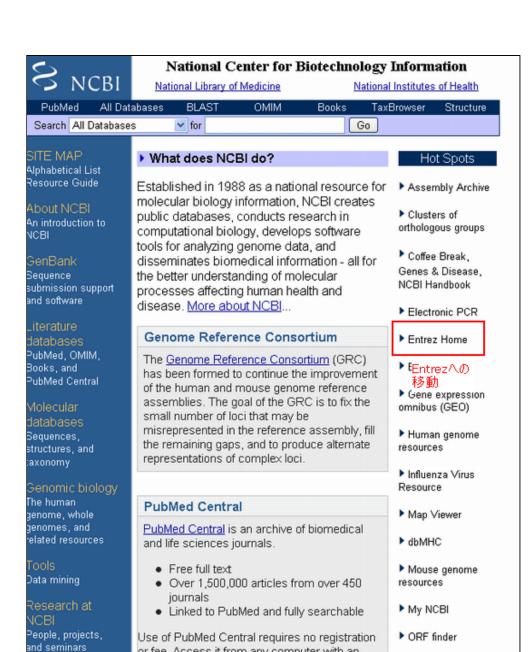
このミニコースでは、既知のSNPsを含むHFE遺伝子についての情報取得方法や、野生型とCys282Tyr変異型タンパク質の機能に関する生化学的、構造的な基礎の説明方法を示しています。

- 1. HFE遺伝子は、6番染色体上にあり、少なくとも11選択的スプライスによる生成物から成り立っています。
- 2. 今のところ、NP_000401タンパク質配列上には8つの非同義アミノ酸変異を引き起こすSNPが報告されています。
- 3. Cys282Tyr変異体は血色素症の疾病と関連し、変異している部位は、血色素症の遺伝子検査に用いられています。
- 4. HFEタンパク質は、トランスフェリン受容体とトランスフェリンの相互作用を制御することで鉄吸収をコントロールする働きがあります。 しかし、Cys282Tyrの変異体では、この相互作用が制御できず鉄過剰を 引き起こします。HFEタンパク質において、免疫グロブリン定常領域に保存されているシステイン282は、ジスルフィド結合 の形成に関わっており、チロシンへの突然変異は、タンパク質のフォールディングを変えてしまいます。

解答•解説

■ 解答1-1_

問題に示されたリンクあるいは、NCBIのトップページ右側にあるメニューからEntrez Home (図中赤四角)を選択し、その次のページからGene (図中赤四角)を選択することで、Entrez Geneのトップ画面に移動します。



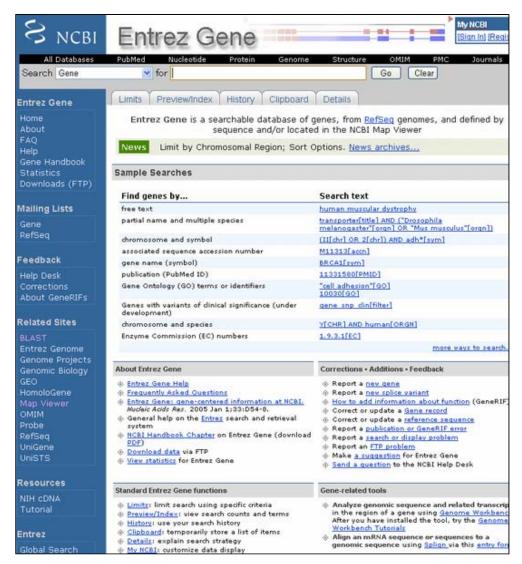
or fee. Access it from any computer with an

Internet connection.

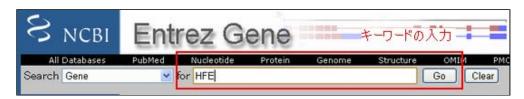
Rat genome

resources

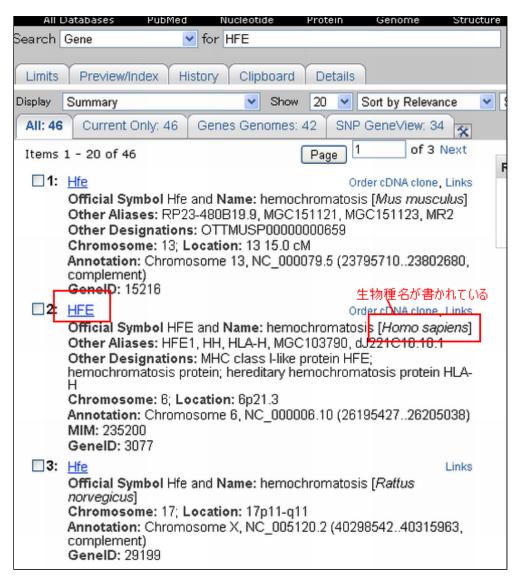




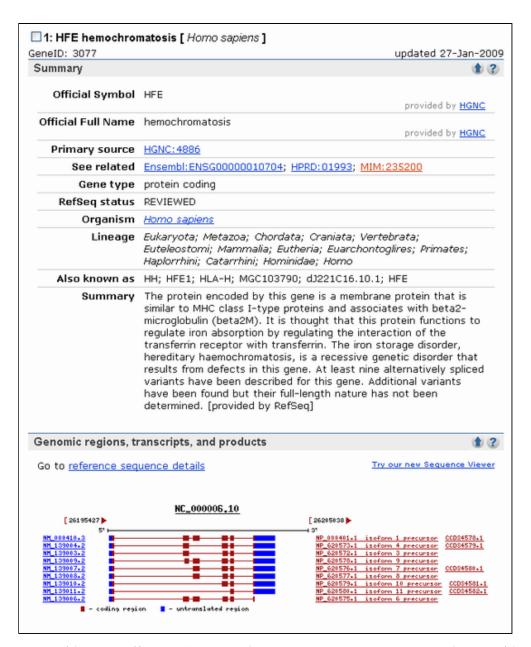
続いて、上部のテキストボックスに"HFE"と入力し、その右横の"GO"ボタンを押すことで HFEをキーワードにした 検索を実行します。



すると下図のように46件もの検索結果が得られます。しかし、Human(Homo Sapiens)に関する結果は2件目のみです。



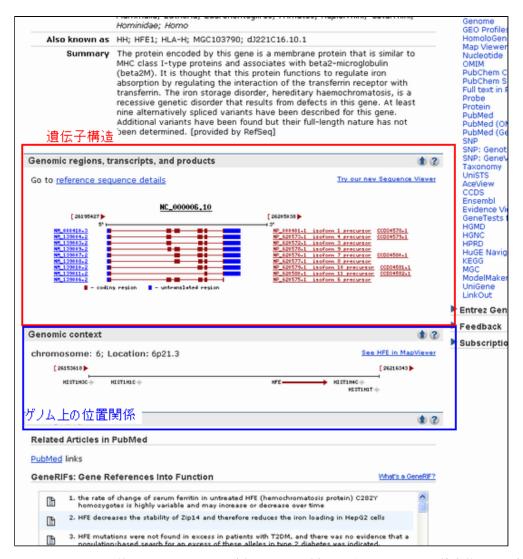
赤四角で囲まれたHFEへのリンクをクリックして詳細情報を取得してください。



ヒトHFE遺伝子関する情報が示されたページへと移動したと思います。このページではHFE遺伝子に関し、ゲノム上での位置、遺伝子構造、発現情報、文献情報などさまざまな情報が得られるようになっています。

■ 解答1-2_

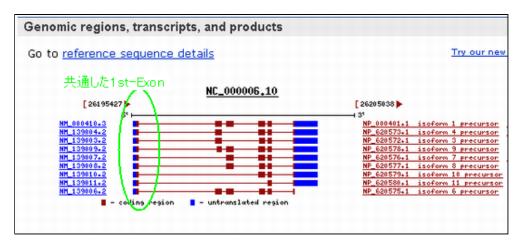
画面を下に少しスクロールすると、遺伝子構造やゲノム上での位置関係についての情報が得られます。



Genomic contextの箇所に注目するとHFE遺伝子がエンジ色で示されており、6番染色体上の順方向の鎖に位置していることが確認できます。また、その前後には青四角で示したHIST1H1C、HIST1H4C という遺伝子が存在することも確認できます。



その少し上の"Genomic regions, transcipts, and products"の欄では遺伝子構造が示されています。



図からわかるようにHFE遺伝子には9個の選択的スプライス転写産物が今のところ知られています。 1番目のエクソンは皆共通ですが、2番目以降のエクソンに関しては図で示されたように転写産物によってエクソンの使われ方や、またエクソンの長さ自体も異なっていることが確認できます。

HFEの別名 (エイリアス) に関しては、画面上部の"summary"内に"also known as"として示されています。

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

HH, HFE1, HLA-A, MGC103790, dJ221C16.10.1といった名前で呼ばれていることが確認できます。

HFE遺伝子の変異に関連した表現型は画面中央付近の"Phenotypes"の欄に示されています。

Phenotypes

Hemochromatosis

MIM: 235200

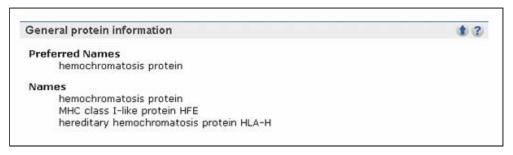
Porphyria variegata

MIM: 176200

これらからHFE遺伝子の変異は血色素症と異型ポルフィリン症に関連していることがわかります。より詳細な疾患に関する情報はMIM番号で示されたOMIMへのリンクをクリックすることで得られます。

■ 解答1-3

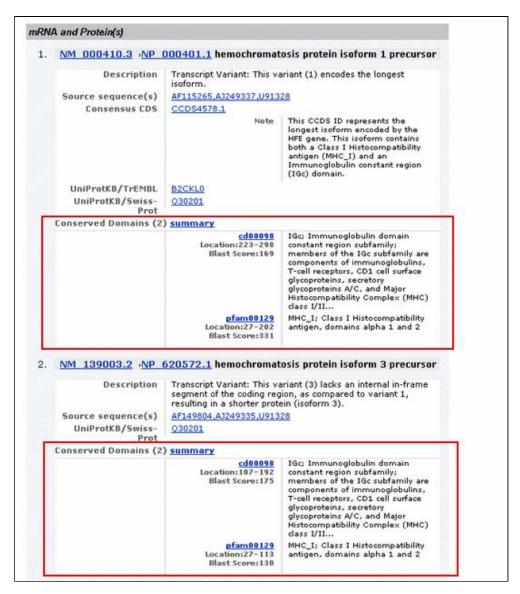
HFE遺伝子がコードするタンパク質に関する情報は画面中央部分のGeneral protein informationや Gene Ontology の部分にその概要が示されています。まず名前ですが名前はGeneral protein informationの箇所から hemochromatosis protein (血色素症タンパク)、MHC class I-like protein HFE (MHCクラスI様タンパクHFE)、hereditary hemochromatosis protein HLA-H(遺伝子性血色素症タンパクHLA-H)などと呼ばれていることが確認できます。



また機能としては、その上のGene Ontologyのfunctionから「鉄イオンの結合」があることがわかります。その下のProcessから抗原処理、抗原提示や鉄イオン輸送に関連することも確認できます。

eOntology		Provided by			
Function	Evidence				
iron ion binding		IEA			
Process	Evidence				
antigen processing and presentation	IEA				
antigen processing and presentation of peptide antigen via Mi class I	HC IEA				
cellular iron ion homeostasis	TAS	PubMed			
cellular response to iron ion starvation	IEA				
female pregnancy	IEA				
hormone biosynthetic process	IEA				
immune response	IEA				
ion transport	IEA				
iron ion homeostasis	IEA				
iron ion transport	TAS	PubMed			
protein complex assembly	TAS	PubMed			
receptor-mediated endocytosis	TAS	PubMed			
Component	Ev	Evidence			
MHC class I protein complex	IEA				
apical part of cell	IEA				
cytoplasm	TAS	PubMed			
integral to plasma membrane	TAS	PubMed			
membrane	IEA				
plasma membrane	TAS	PubMed			

HFE遺伝子がコードするタンパク質に含まれるドメインは、NCBI Reference Sequence (RefSeq)の mRNA and Protein(s)に示されています。

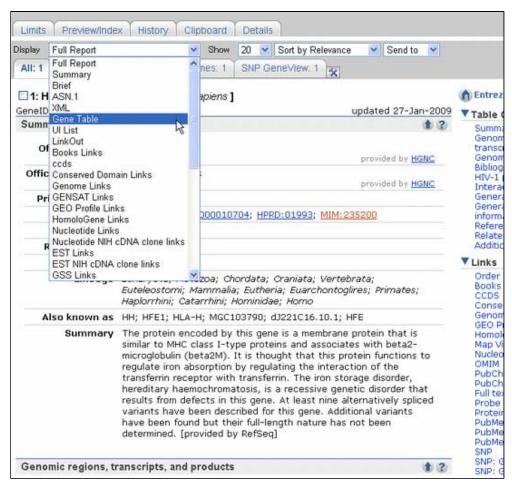


図中赤で示されたconserved domainの欄に各転写産物ごとに含まれるドメインが示されています。この遺伝子では、免疫グロブリン関連のサブファミリやMHCクラスI組織適合抗原ドメインなどが含まれていることが確認できます。

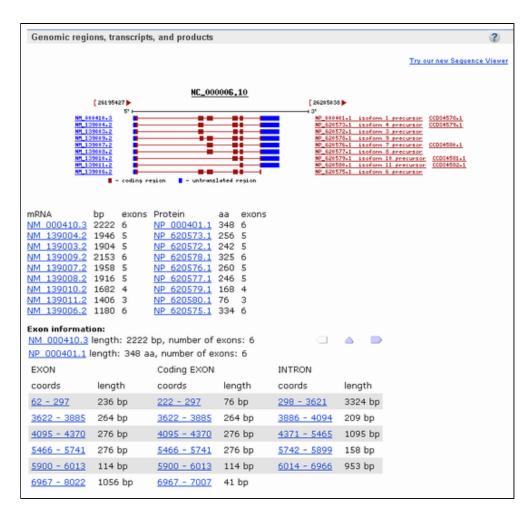
またタンパク質の細胞内での局在情報は、Gene OntologyのComponentの箇所から原形質膜などに存在することが読み取れます。

■ 解答1-4

画面上部の"Display"横のプルダウンメニューから"Gene Table"を選んでください。



すると、下図のような遺伝子構造を詳細に表示した画面へと移動します。

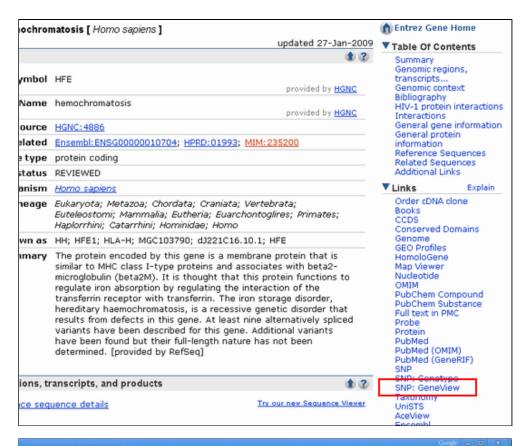


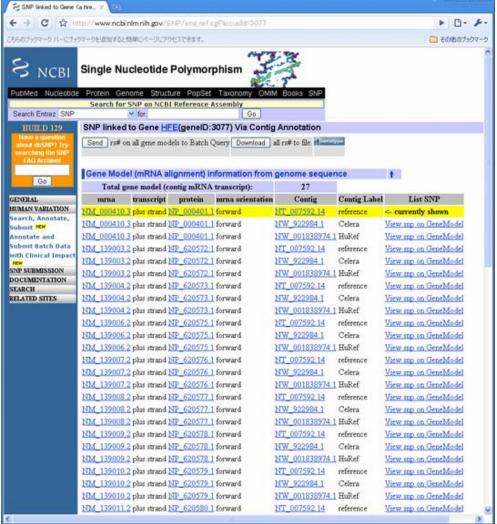
このページでは赤四角で囲んだ領域で、エクソン、イントロンの位置情報が表形式で示されていて、それを読むことで位置がbp単位で得られます。

解答・解説2

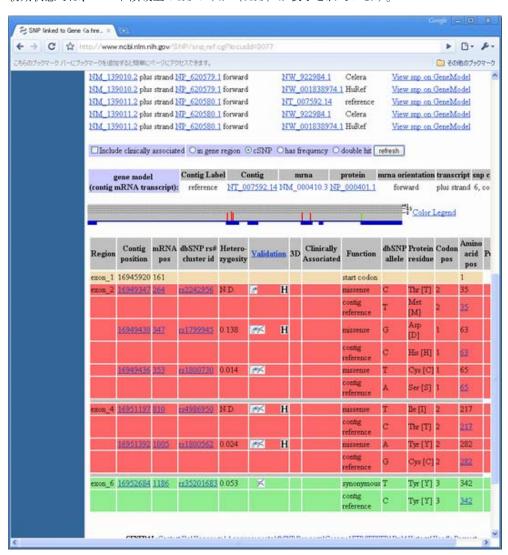
■ 解答2-1_

画面右側のリンクメニューからSNP:GeneViewをクリックして、既知SNP情報へと移動してください。





初期状態では、コード領域上のSNPのみが(cSNP)が表示されています。



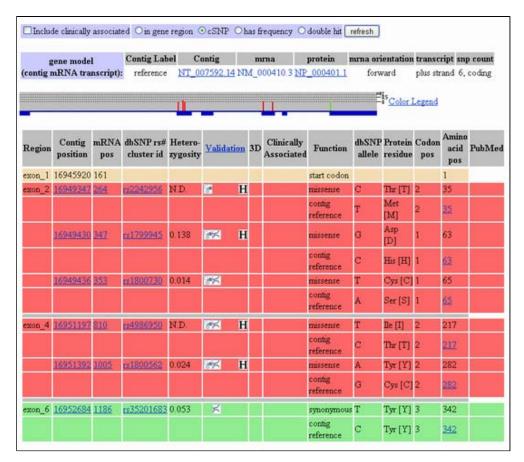
"in gene region"ボタンをクリックし、refreshボタンをクリックすることで、上流やイントロン上に存在するSNPも表示されるようになります。(画面上黄色の領域など)





■ 解答2-2_

ひとつ画面を戻り、NM_000410 (デフォルト) についてのSNP情報を詳細に見てみましょう。



画面よりデフォルトの設定では6個のコード領域に含まれているSNPが表示されていることがわかります。 (1SNPにつき2行で表示されている) その中で、赤く表示された非同義置換のSNPが5個あることも確認できます。

次にOMIMへのリンクがあるデータを表示させて見ましょう。デフォルトの設定では表示されていないのでまず、OMIMを含む診療関連データと対応の取れているデータを表示させることが可能なオプションを用いてこれらのSNPsを表示させて見ましょう。



SNPs一覧の上にある箇所から"Include clinically assocciated"横のチェックボックスをクリックして"refresh"ボタンを押し、診療関連データと対応の取れているデータを表示させてください。

Region	Contig position	mRNA pos	dbSNP rs# cluster id		Validation	3D	Clinically Associated	Function	100000000000000000000000000000000000000	Protein residue		Amino acid pos	PubMed
exon_1	16945920	161						start codon				1	
	16949347	264	rs2242956	N.D.	a H			missense	C	Thr [T]	2	35	
								contig reference	Т	Met [M]	2	35	
	16949400	317	rs28934889	0.011			X	missense	A	Met [M]	1	53	
							丟	contig reference	G	Val [V]	1	53	
	16949418	335	rs28934890	N.D.			X	missense	A	Met [M]	1	59	
							**	contig reference	G	Val [V]	1	59	
	16949430	347	rs1799945	0.138	#≤ H			missense	G	Asp [D]	1	63	
		i de				Š		contig reference	С	His [H]	1	63	
	16949436	353	rs1800730	0.014	8°K			missense	T	Cys [C]	1	65	
								contig reference	A	Ser [S]	1	65	
	16949520	437	t#28934597	N.D.			*	missense	C	Arg [R]	1	93	
							X	contig reference	G	Gly [G]	1	93	
	16949557	474	ts28934596	N.D			*	missense	C	The [T]	2	105	
							Ħ	contig reference	Т	lle [I]	2	105	
exon_3	16949833	541	ts28934595	N.D.			*	missense	C	His [H]	3	127	
							莱	contig reference	A	Gin [Q]	3	127	
	16951197	310	rs4986950	N.D.	₩ H		9	missense	T	lle [I]	2	217	
								contig reference	c	Thr [T]	2	217	
	16951392	1005	rs1800562	0.024	<i>≫</i> H			missense	A	Tyr [Y]	2	282	
								contig reference	G	Cys [C]	2	282	
exon_6	16952684	1186	rs35201683	0.053	×			synonymous	Т	Tyr [Y]	3	342	
								contig reference	c	Тут [Ү]	3	342	

[&]quot;Clinically Assocciated"欄にヒトの形をしたOMIMへのリンクがあるものが5つあることも確認できます。

解答・解説3

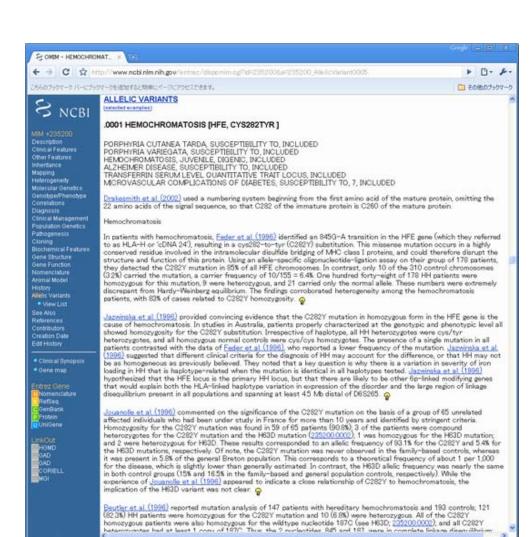
■ 解答3-1_

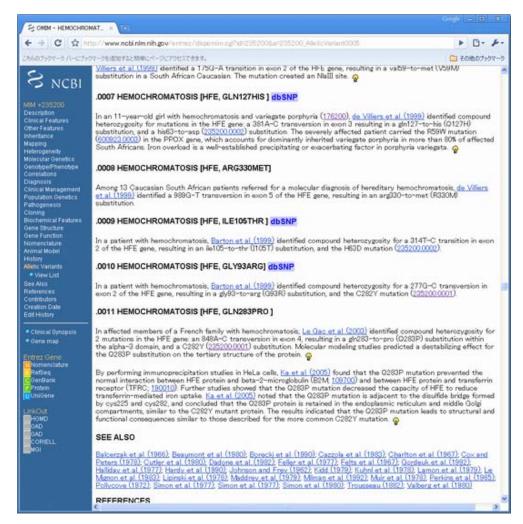
SNP情報の表中に"Clinically Assocciated"欄のヒトの形をしたOMIMへのリンクのどれかをクリックしてください。血色素症に関するOMIMへと移動します。一番上までスクロールしてみてください。はじめのdescriptionのところにこの病気に関する概要が示されており、下に詳細な情報が記述されています。



Descriptionの記述によると、血色素症は肝硬変、糖尿病、皮膚へのメラニン色素沈着、心不全などを引き起こします。またこの疾患は、6p21.3上のHFE遺伝子変異により引き起こされる古典的な血色素症(HFE)、1q21上の hemojuvelin(HJV)遺伝子変異によるもの、若年型血色素症と呼ばれる19q13上の遺伝子(HAMP)変異によるもの (HFE2)、7q22上の遺伝子(TFR2)変異によるもの (HFE3)、2q32上の遺伝子(SLC40A1)変異によるもの (HFE4)の5種類が知られています。それらの疾患のうち、HFEと呼ばれる古典的な血色素症が6p21.3上のHFE遺伝子における変異と関連しています。

HFE遺伝子の対立遺伝子多型についてはこのページの中央よりやや上部の Allelic variantsに記載されています。

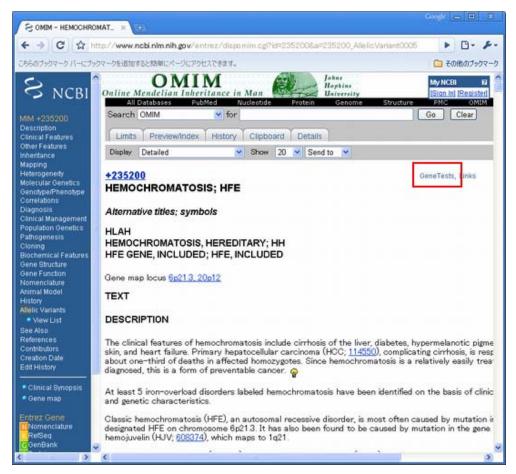




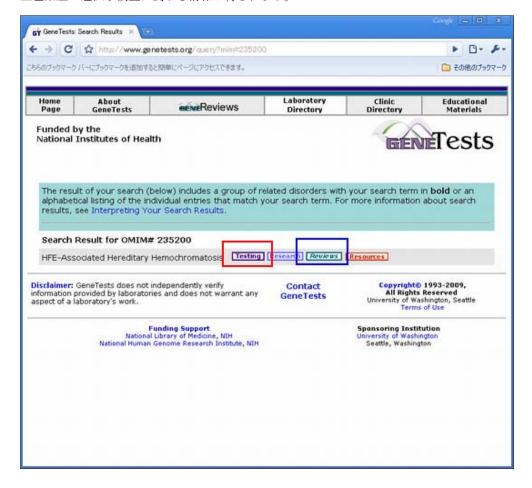
現在OMIMでは11種類のHFE遺伝子に関連する多型が記述されており、 Cys282Tyr変異は、その一番目に記載されています。表現型としては血色素症となっています。

■ 解答3-2

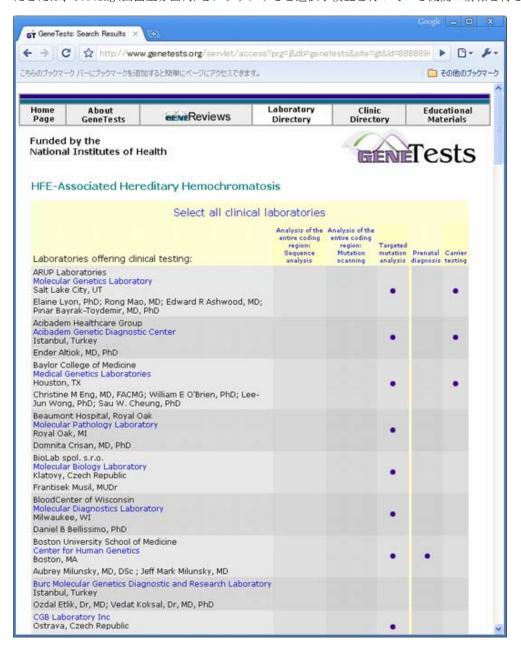
OMIMページの一番上にある"GeneTests"へのリンクをクリックしてください。



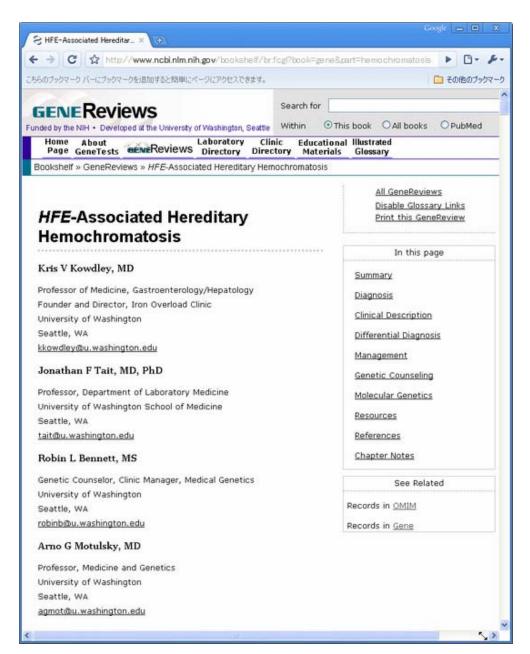
血色素症の遺伝子検査に関する情報が得られます。



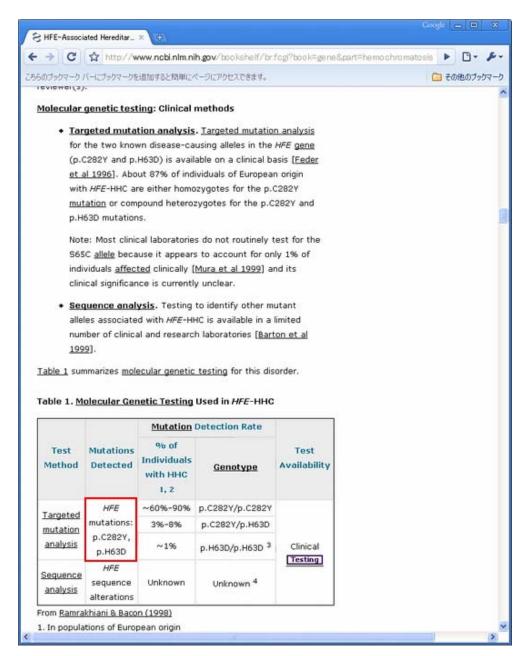
たとえば、Testing(画面上赤四角)をクリックすると遺伝子検査を行っている機関の情報を得ることができます。



一つ画面を戻ってReview(画面上青四角)sをクリックすることでHFE遺伝子に関する遺伝子検査についての概要を知ることができます。



画面1/4程の箇所付近にある"Molecular Genetic Testing"の項目では、HFE遺伝子のどの多型を検査できるかが示されています。



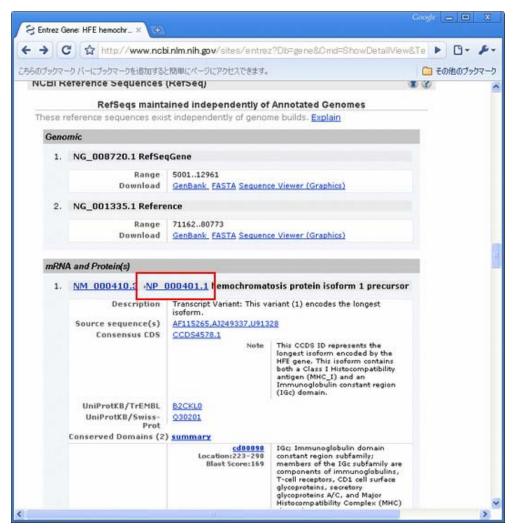
赤四角で囲んだ表の記述から282残基目のCys/Tyr変異、および63残基目のHis/Asp変異に対する検査が可能なことがわかります。

282残基目のCys/Tyr変異に関する表現型は随所に記載が見られます。この変異により鉄の代謝異常が起こり貯蔵鉄増加による血色素症を引き起こします。

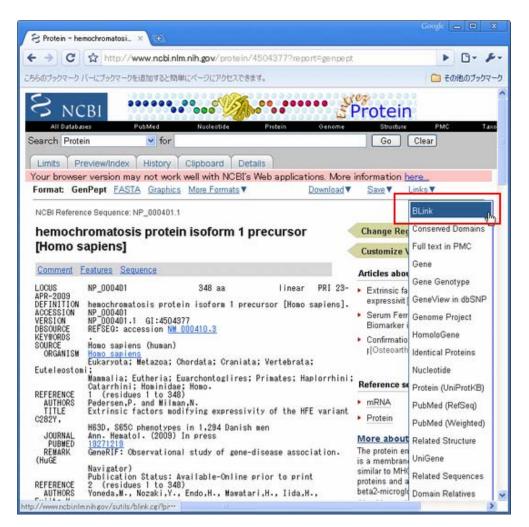
解答・解説4

■ 解答4-

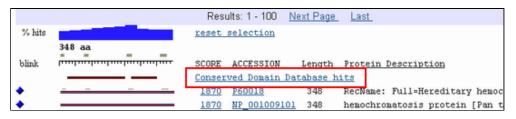
ブラウザのバックボタンを用いてEntrez Geneの画面に戻り、タンパク質のアミノ酸配列で一番上に書かれている NP_000401をクリックしてください。

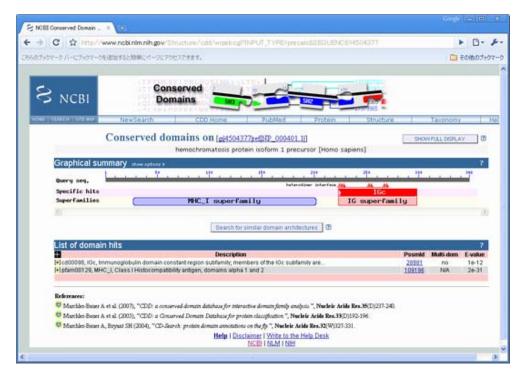


続いて表示された画面上部のLinksプルダウンメニューからBLinkをクリックしてください。

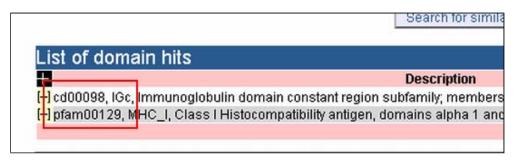


さらに、Conserved Domain Database hitsをクリックしてドメイン情報を取得してください。



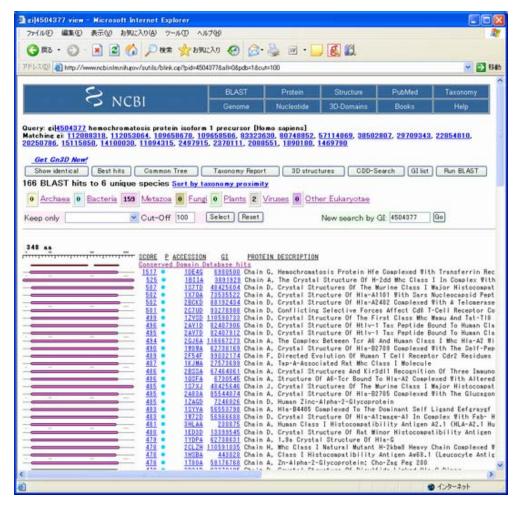


CDDに含まれるcd00098とpfamに含まれるpfam00129の二つのドメインがこの配列上に確認できます。



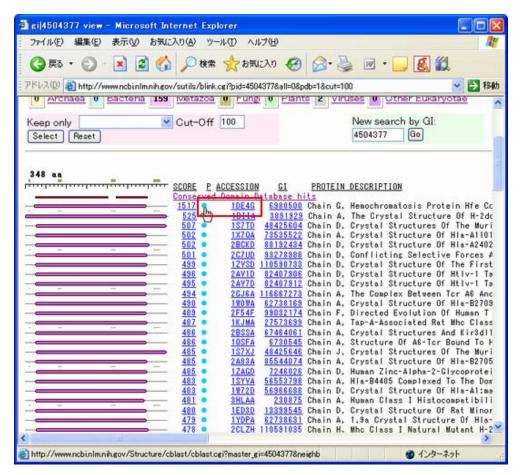
NP_000401と配列が似ていて、立体構造が既知であるを持つタンパク質を調べるために、ブラウザのバックボタンを用いて戻り、画面上部の3D structuresボタンをクリックしてください。(この機能はなくなってしまったため、昔の画面を用いて流れを解説します。)



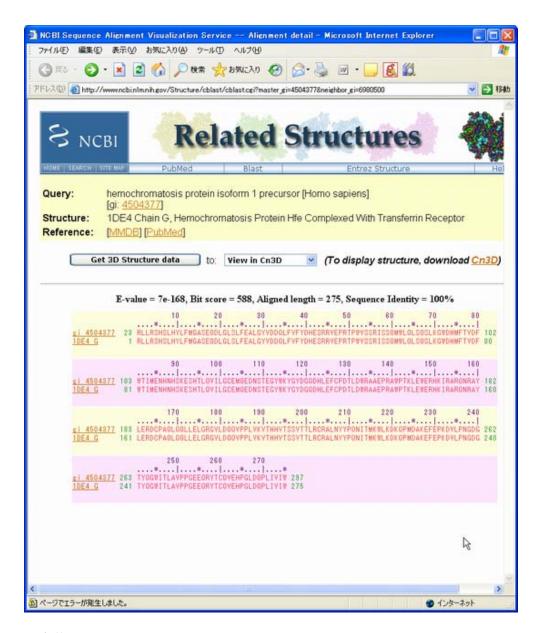


表示されている結果は、既知の立体構造を持つタンパク質の中で問い合わせ配列と類似のもののリストになっています。一番初めにエントリーされている1DE4Gは、血色素症タンパク質のG鎖に相当します。(トランスフェリン受容体と複合体を形成します)

1DE4のG鎖と問い合わせのアミノ酸配列とのアライメントを得るため、 1DE4Gの付近にある青いドットをクリックして下さい。

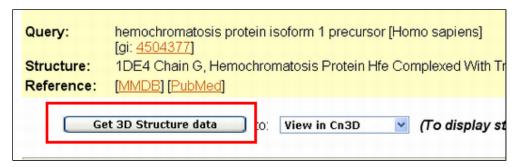


アライメント結果が別画面に表示されます。

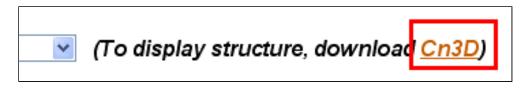


■ 解答4-2_

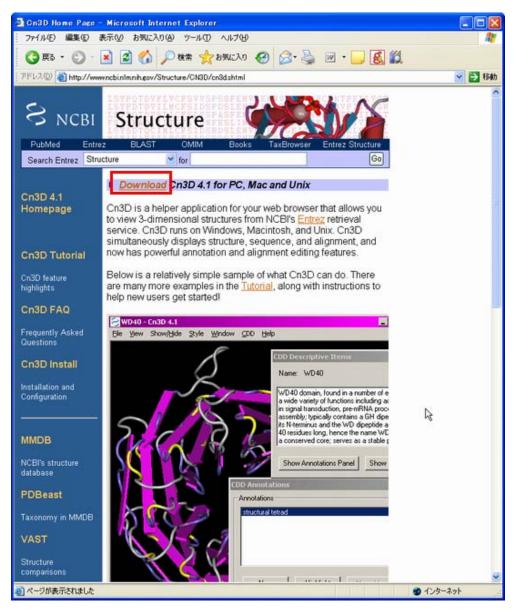
アライメント画面内でGet 3D Structure dataボタンをクリックしてください。



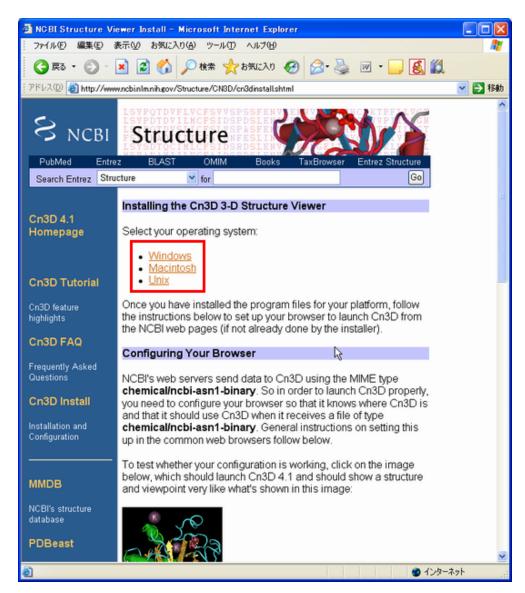
この際、Cn3Dがインストールされていない場合には正しく表示することができないのでまずCn3Dをインストールする必要があります。インストールされている人はこの部分は読み飛ばしてください。まず、Todisplay structure, download Cn3Dと書かれた箇所をクリックします。



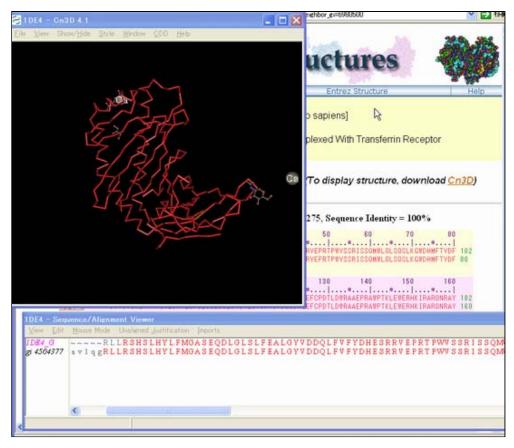
続いて表示されたNCBI StructureのページからCn3Dをダウンロードします。画面上部のDownloadをクリックします。



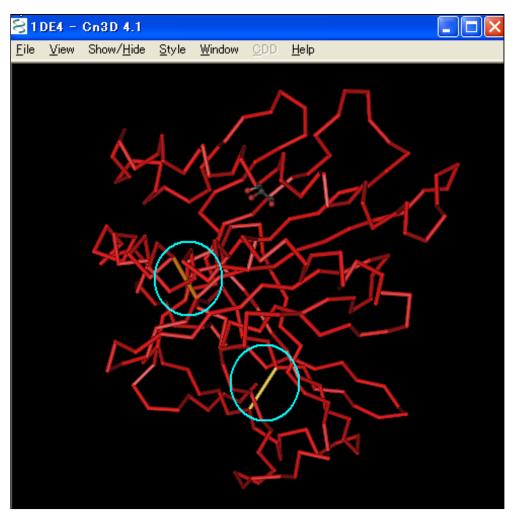
するとインストールするコンピュータのOSを聞かれますので、該当するOSの部分をクリックし、その先はOSごとに書かれた指示に従ってCn3Dをインストールしてください。



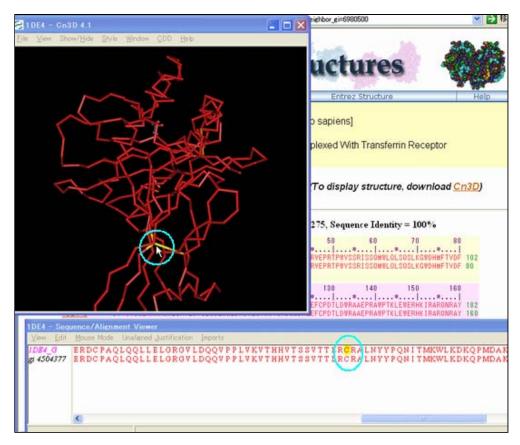
インストールが正常にできているとGet 3D Structure dataボタンをクリックすると、下図のようなワイヤーモデルで立体構造が示された画面とSequence/Alignment Viewerが立ち上がります。



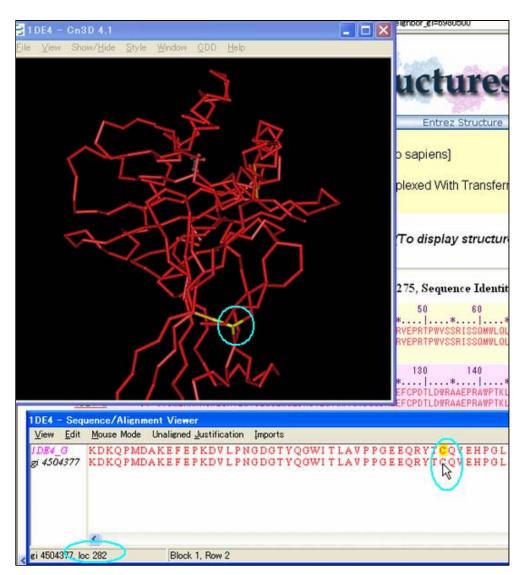
ワイヤーモデルの中を"z"ボタンを使ってズームインしたり、あるいはマウスを使って回転させていくと、画像中水色で丸く囲んだ箇所に黄褐色で示されたジスルフィド結合(S-S結合)が認められます。



そのジスルフィド結合をマウスでダブルクリックすると残基が画面上黄色くなり、 Sequence/Alignment Viewerの 方でも該当する残基が黄色くハイライトされます。



Sequence/Alignment Viewer上で問い合わせ配列である2行目の配列上(gi 4504377)で黄色くなっている残基(システイン"C")にマウスを重ねると左下にそのポジションが表示されます。これをいくつかのジスルフィド結合を作っているシステイン上で繰り返すと、その中のひとつに 282番と表示されるものがあります。



血色素症を引き起こす原因となるチロシンへと変異したのと同じシステインです。これにより、正常なNP_000401配列上では282番目のアミノ酸はジスルフィド結合を形成しており、これがチロシンへと変わることで立体構造の形成に重要な役割をするジスルフィド結合が形成できなくなり、構造が大きく変化することがわかります。

原文更新日: 2005年6月7日 日本語版更新日: 2006年1月31日

All Rights Reserved, Copyright(C) 1997-2006 Japan Science and Technology Agency(JST)

