

Entrez Geneクイックスタート

NCBIのEntrez Geneでは、染色体上の位置、配列、発現、構造、機能、ホモロジーデータのような遺伝子をベースとした情報を提供しています。それぞれのレコードは、ある生物種由来のある遺伝子を表しています。Entrez Geneには、RefSeqゲノムデータが存在する生物種由来の遺伝子も含まれています。

このコースでは、mRNAやゲノム配列情報、遺伝子構造情報（エクソン、イントロンの位置）、機能、変異と関連する表現型といったヒト遺伝子に関連する情報をどのように手に入れるかを学びます。また、遺伝子のコード領域に存在するSNPsがその産物であるタンパク質の機能を変更することが知られているかどうかについて、どのように調べるかについても学びます。

Entrez Geneは、LocusLinkの後継となるものです。このミニコースでは、LocusLinkで見つけられたのと同じ情報を、Entrez Geneから見つける方法についてを扱っています。さらには、効率的な検索オプションや、バクテリアやウィルスを含むゲノムが完全解読された生物種に関する遺伝子特異的な情報の取り扱いといったEntrez Geneならではの長所も併せて紹介しています。

課題. 1

■ Step. 1-1

[Entrez Gene](#)を用いて、“プリオンタンパク質”に関係するヒトのエントリを検索して下さい。プリオンタンパク質(PRPN)の遺伝子情報を確認して下さい。ヒトゲノム上でこの遺伝子が位置する場所をマップ上で確認して下さい。このタンパク質の機能は何でしょうか？この遺伝子の別名は何でしょうか？この遺伝子の変異に関連した表現型を確認して下さい。

■ Step. 1-2

RefSeqのmRNAのデータに目を通しましたか？この遺伝子には、どれくらいの選択的スプライス転写産物がアノテートされているのでしょうか？

他の真核生物についてのホモログについて知るため、Homologeneを選択して下さい。画面を“Alignment Scores”というオプションの表示に変えて下さい。ヒトとマウスのタンパク質を比べてみて、どの程度の相同性(percent identity)があるのでしょうか？“Blast”と表示されている箇所をクリックし、アライメントを表示して下さい

■ Step. 1-3

Entrez Geneに戻って下さい。SNP:GeneViewをクリックして、この遺伝子上にアノテートされた塩基置換を確認して下さい。これらの変異のうち、どれくらい非同義置換があるのでしょうか？また、遺伝子のコード領域に存在する既知のSNPsが、表現型に関連しているかどうかを調べるため、SNPレポート中のOMIMの列に“Yes”と書かれている箇所をクリックして OMIMデータにアクセスして下さい。（今のバージョンでは“Clinically Associated”と記述され、LSDBなどへの情報も含まれます。）

SNPレポート中の非同義置換とOMIMデータ中の“ALLELIC VARIANTS”とを比較してください。プリオンタンパク質機能変異の原因だとされているSNPsはありますか？

■ Step. 1-4

Entrez Geneに戻って下さい。タンパク質NP_000302と書かれた箇所をクリックし、ポップアップするリストからBlinkリンクを選択して、NP_000302と類似したタンパク質のリストを表示して下さい。

（これ以降の部分は現在では、提供されていません。）

立体構造上で変異のある部位を確認するために、ヒトのプリオンタンパク質の立体構造上に NP_000302タンパク質配列を重ねて表示して下さい。（3D structuresと書かれたボタンを押し、最初にある青色のドットをクリックし、さらに表示されたページから立体構造のデータを取得して下さい。）立体構造上で変異した残基を確認し、強調して表示して下さい。

課題. 2

■ Step. 2-1

[Entrez Gene](#)を用いて、“結腸ガン”に関係するヒトのエントリを検索して下さい。MLH1の遺伝子情報を確認して下さい。ヒトゲノム上でこの遺伝子が位置する場所をマップ上で確認して下さい。このタンパク質の機能は何でしょうか？この遺伝子の別名は何でしょうか？この遺伝子の変異に関連した表現型を確認して下さい。

■ Step. 2-2

RefSeqのmRNAのデータに目を通しましたか？この遺伝子には、どれくらいの選択的スプライス転写産物がアノテートされているのでしょうか？

他の真核生物についてのホモログについて知るため、Homologeneを選択して下さい。画面を“Alignment Scores”というオプションの表示に変えて下さい。ヒトとマウスのタンパク質を比べてみて、どの程度の相同性(percent identity)があるのでしょうか？“Blast”と表示されている箇所をクリックし、アライメントを表示して下さい

■ Step. 2-3

Entrez Geneに戻って下さい。SNP:GeneViewをクリックして、この遺伝子上にアノテートされた塩基置換を確認して下さい。これらの変異のうち、どれくらい非同義置換があるのでしょうか？また、遺伝子のコード領域に存在する既知のSNPsが、表現型に関連しているかどうかを調べるため、SNPレポート中のOMIMの列に“Yes”と書かれている箇所をクリックして OMIMデータにアクセスして下さい。（今のバージョンでは“Clinically Associated”と記述され、LSDBなどへの情報も含まれます。）

SNPレポート中の非同義置換とOMIMデータ中の“ALLELIC VARIANTS”とを比較してください。MLH1タンパク質機能変異の原因だとされているSNPsはありますか？

■ Step. 2-4

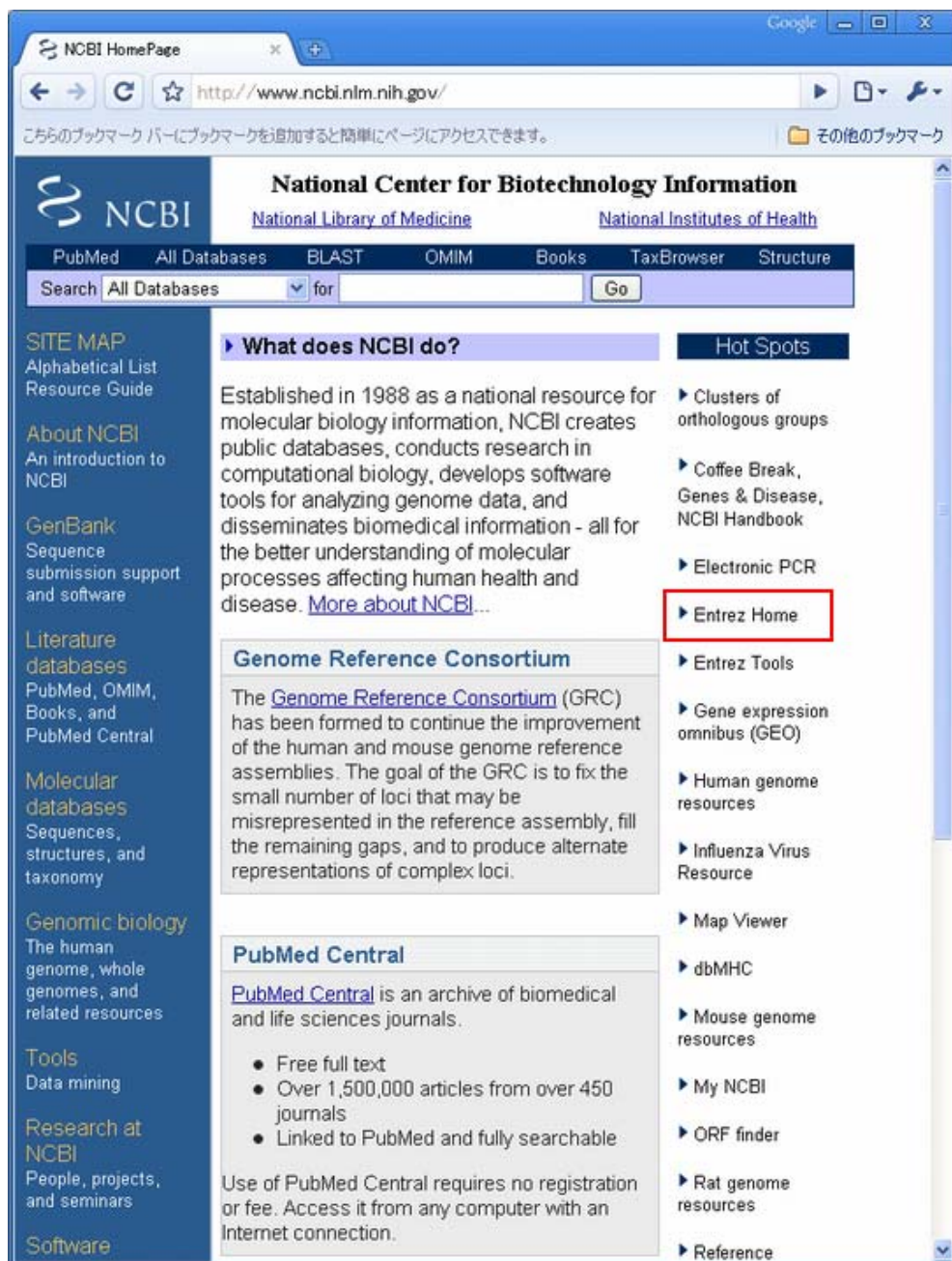
Entrez Geneに戻って下さい。タンパク質NP_000240と書かれた箇所をクリックし、ポップアップするリストからBlinkリンクを選択して、NP_000240と類似したタンパク質のリストを表示して下さい。立体構造上で変異のある部位を確認するために、大腸菌1BKNBタンパク質の立体構造上に NP_000240タンパク質配列を重ねて表示して下さい。（3D structuresと書かれたボタンを押し、最初にある青色のドットをクリックし、さらに表示されたページから立体構造のデータを取得して下さい。）立体構造上でヒトMLH1の32残基目のイソロイシンに該当する残基を確認し、強調して表示して下さい。大腸菌のタンパク質ではヒトMLH1の32残基目に対応するアミノ酸は何ですか？この情報に基づいて、ヒトMLH1のI32Vの変異は機能を変えてしまうと思いますか？あなたが得た知見をMLH1に関するOMIMの記述の中で確認してください。

解答・解説

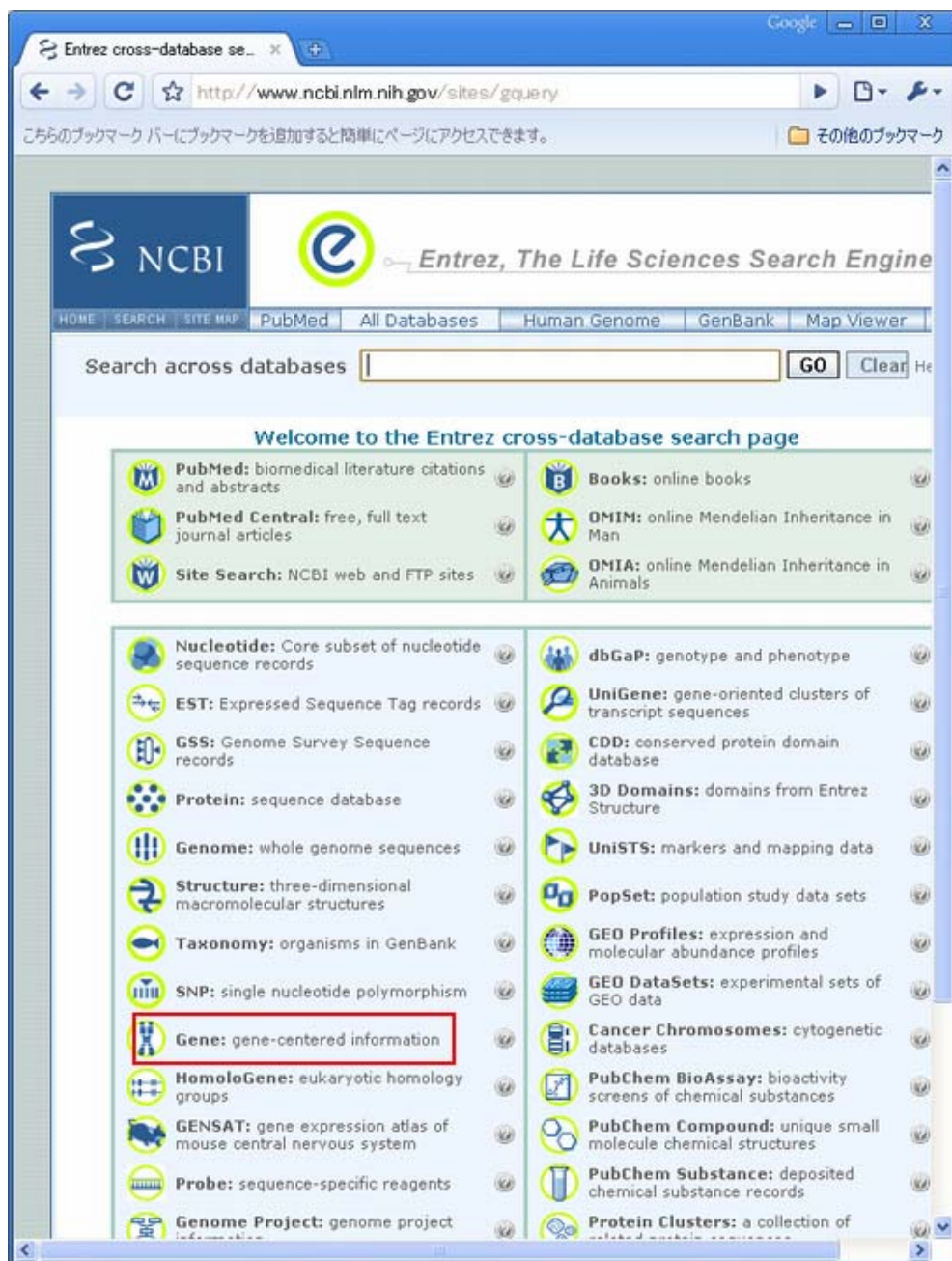
■ 解答1-1

[Entrez Geneへの移動](#)

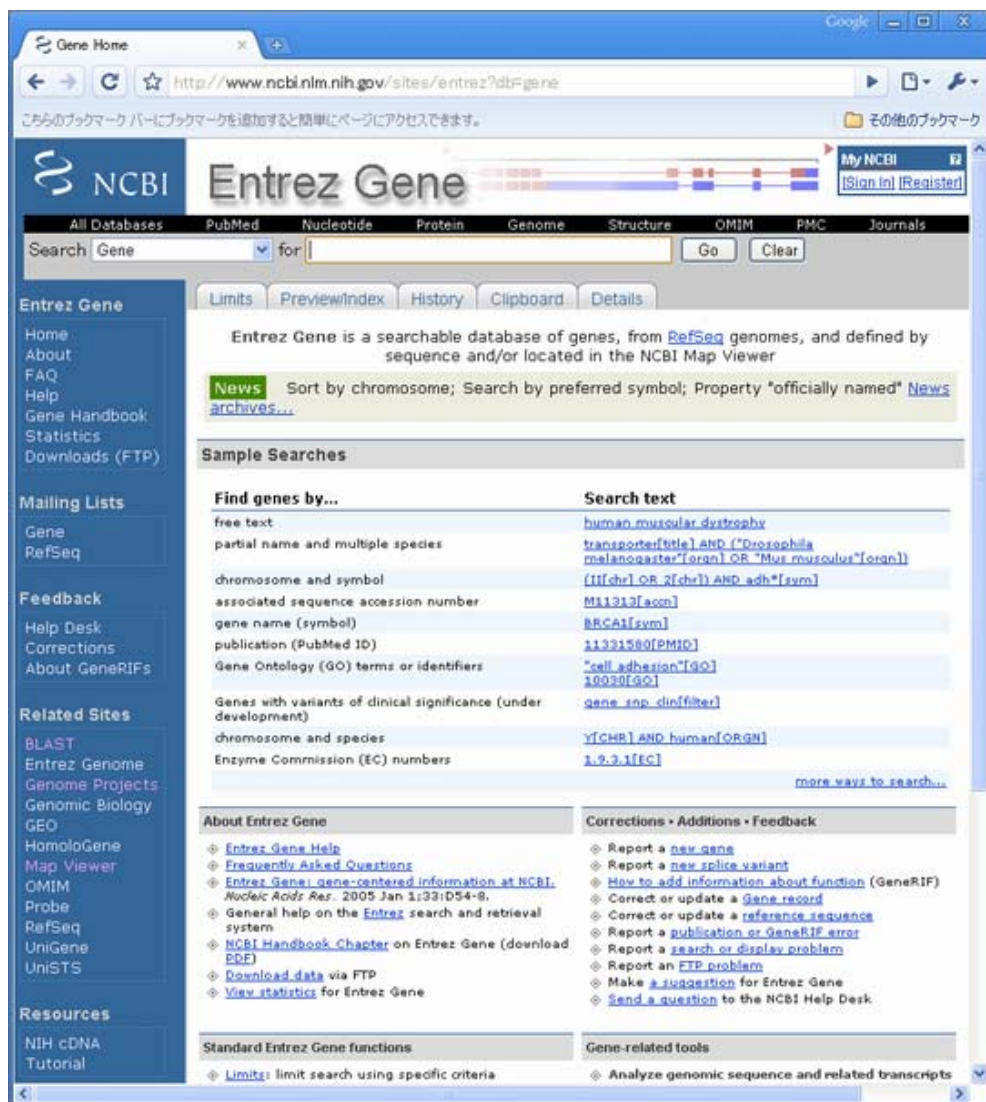
課題に示されたリンクあるいは、NCBIのトップページ右側にあるメニューから Entrez Home (図中赤四角) を選択し、



その次のページからGene(図中赤四角)を選択することで、

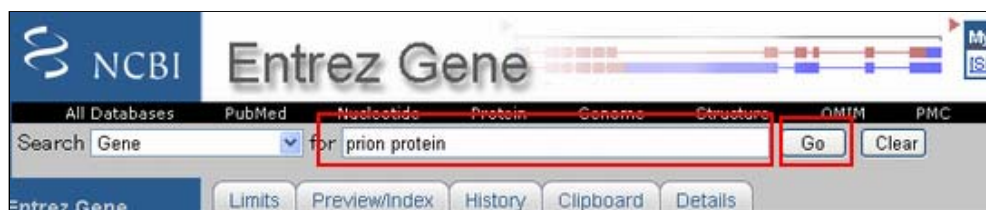


Entrez Geneのトップ画面に移動します。



Entrez Geneでのキーワード検索

続いて、上部のテキストボックスに“prion protein”と入力し、その右横の“Go”ボタンを押すことで“prion protein”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複数の検索結果が得られます。その中にHuman (Homo Sapiens)に関する PRNP(プリオンタンパク質)が1番目に表示されていることが確認できます。

prion protein - Gene Result

http://www.ncbi.nlm.nih.gov/sites/entrez

その他のブックマーク

NCBI Entrez Gene

All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Journals

Search: Gene for prion protein Go Clear Save Search

Limits Preview/Index History Clipboard Details

Display Summary Show 20 Sort by Relevance Send to

All: 372 Current Only: 322 Genes Genomes: 285 SNP GeneView: 117

Items 1 - 20 of 372 Page 1 of 19 Next

☐ 1: [PRNP](#) [Order cDNA clone, Links](#)

Official Symbol PRNP and **Name:** prion protein [*Homo sapiens*]
Other Aliases: ASCR, CD230, CJD, GSS, MGC26679, PrP^{Sc}, PrP^{Pr}, PrP27-30, PrP33-35C, PrPc, prion
Other Designations: CD230 antigen; OTTHUMP0000000004, OTTHUMP00000000163, major prion protein; p27-30, prion protein PrP; prion-related protein
Chromosome: 20; **Location:** 20p13
Annotation: Chromosome 20, NC_000020.9 (4614797..4630234)
MIM: 176640
GeneID: 5621

☐ 2: [Pitrm4](#) [Links](#)

Official Symbol Pitrm4 and **Name:** prion incubation time 4 [*Mus musculus*]
Chromosome: 15; **Location:** 15 49.6 cM
GeneID: 117009

☐ 3: [Pitrm3](#) [Links](#)

Official Symbol Pitrm3 and **Name:** prion incubation time 3 [*Mus musculus*]
Chromosome: 8; **Location:** 8 43.0 cM
GeneID: 117008

☐ 4: [Pitrm2](#) [Links](#)

Official Symbol Pitrm2 and **Name:** prion incubation time 2 [*Mus musculus*]
Chromosome: 4; **Location:** 4 42.5 cM
GeneID: 117007

☐ 5: [Pitrm1](#) [Links](#)

Official Symbol Pitrm1 and **Name:** prion incubation time 1 [*Mus musculus*]
Chromosome: 2; **Location:** 2 34.0 cM
GeneID: 117006

Recent Activity

Turn Off

Your browsing activity is empty.

Entrez Gene: PRNP prion p... x

http://www.ncbi.nlm.nih.gov/sites/entrez?Db=gene&Cmd=ShowDetailView&Term=1

こちらのブックマーク バーにブックマークを1追加すると簡単にページにアクセスできます。

NCBI Entrez Gene

All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Journals

Search Gene for [] Go Clear

Limits Preview/Index History Clipboard Details

Display Full Report Show 20 Sort by Relevance Send to

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

1: PRNP prion protein [*Homo sapiens*]

GeneID: 5621 updated 12-Jan-2009

Summary

Official Symbol PRNP provided by HGNC

Official Full Name prion protein provided by HGNC

Primary source HGNC:9449

See related Ensembl: ENSG00000171867; HPRD:01453; MIM:176640

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 CJD; GSS; PrP; ASCR; PRIP; PrPc; CD230; prion; MGC26679; PrP27-30; PrP33-35C; PRNP

Summary The protein encoded by this gene is a membrane glycosylphosphatidylinositol-anchored glycoprotein that tends to aggregate into rod-like structures. The encoded protein contains a highly unstable region of five tandem octapeptide repeats. This gene is found on chromosome 20, approximately 20 kbp upstream of a gene which encodes a biochemically and structurally similar protein to the one encoded by this gene. Mutations in the repeat region as well as elsewhere in this gene have been associated with Creutzfeldt-Jakob disease, fatal familial insomnia, Gerstmann-Straussler disease, Huntington disease-like 1, and kuru. Alternative splicing results in multiple transcript variants encoding the same protein. [provided by RefSeq]

Entrez Gene

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
Additional Links

Links

Order cDNA
CCDS
Conserved
Genome
GEO Profile
HomoloGen
Map Viewer
Nucleotide
EST
OMIM
PubChem C
PubChem S
Full text in
Probe
Protein
PubMed
PubMed (O)
PubMed (G)
SNP

少し下にスクロールすると、ゲノム上での位置などに関する情報がGenomic context, Genomic regions などの欄に示されています。赤四角で囲まれた情報からPRNPはヒトゲノム20p13の4614797bpから4630234bpに位置することがわかります。

Official Symbol	PRNP	provided by HGNC
Official Full Name	prion protein	provided by HGNC
Primary source	HGNC:9449	
See related	Ensembl:ENSG00000171867 ; HPRD:01453 ; MIM:176640	
Gene type	protein coding	
RefSeq status	REVIEWED	
Organism	Homo sapiens	
Lineage	<i>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo</i>	
Also known as	CJD; GSS; PrP; ASCR; PRIP; PrPc; CD230; prion; MGC26679; PrP27-30; PrP33-35C; PRNP	
Summary	<p>The protein encoded by this gene is a membrane glycosylphosphatidylinositol-anchored glycoprotein that tends to aggregate into rod-like structures. The encoded protein contains a highly unstable region of five tandem octapeptide repeats. This gene is found on chromosome 20, approximately 20 kbp upstream of a gene which encodes a biochemically and structurally similar protein to the one encoded by this gene. Mutations in the repeat region as well as elsewhere in this gene have been associated with Creutzfeldt-Jakob disease, fatal familial insomnia, Gerstmann-Straussler disease, Huntington disease-like 1, and kuru. Alternative splicing results in multiple transcript variants encoding the same protein. [provided by RefSeq]</p>	

機能に関する記載も

あるいはGene Ontologyの欄にオントロジーを利用して記述されています。

Function	Evidence
copper ion binding	IEA
copper ion binding	TAS PubMed
identical protein binding	IPI PubMed
microtubule binding	IDA PubMed
Process	Evidence
anti-apoptosis	IEA
cellular copper ion homeostasis	NAS PubMed
metabolic process	TAS PubMed
protein homooligomerization	IEA
response to cadmium ion	IEA
response to copper ion	IEA
response to oxidative stress	IEA
response to oxidative stress	ISS
Component	Evidence
Golgi apparatus	IEA
Golgi apparatus	ISS
anchored to membrane	IEA
cytoplasm	IEA
cytoplasm	TAS PubMed
endoplasmic reticulum	IEA
endoplasmic reticulum	ISS
extrinsic to membrane	TAS PubMed
membrane	IEA
membrane raft	IEA
membrane raft	ISS
plasma membrane	IEA
plasma membrane	ISS

これらの情報によるとPRNPは膜タンパク質の一種であるグリコシルホスファチジルイノシトールアンカー型タンパク質で、銅イオンと結合するといったことがわかっています。

また、SummaryのAlso known asの記述から、CJD, GSS, PrP, ASCR, PRIPなど多くの別名で呼ばれていることも確認できます。

Official Symbol	PRNP	provided by HGNC
Official Full Name	prion protein	provided by HGNC
Primary source	HGNC:9449	
See related	Ensembl:ENSG00000171867 ; HPRD:01453 ; MIM:176640	
Gene type	protein coding	
RefSeq status	REVIEWED	
Organism	Homo sapiens	
Lineage	<i>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo</i>	
Also known as	CJD; GSS; PrP; ASCR; PRIP; PrPc; CD230; prion; MGC26679; PrP27-30; PrP33-35C; PRNP	
Summary	The protein encoded by this gene is a membrane glycosylphosphatidylinositol-anchored glycoprotein that tends to aggregate into rod-like structures. The encoded protein contains a highly unstable region of five tandem octapeptide repeats. This gene is found on chromosome 20, approximately 20 kbp upstream of a gene which encodes a biochemically and structurally similar protein to the one encoded by this gene. Mutations in the repeat region as well as elsewhere in this gene have been associated with Creutzfeldt-Jakob disease, fatal familial insomnia, Gerstmann-Straussler disease, Huntington disease-like 1, and kuru. Alternative splicing results in multiple transcript variants encoding the same protein. [provided by RefSeq]	

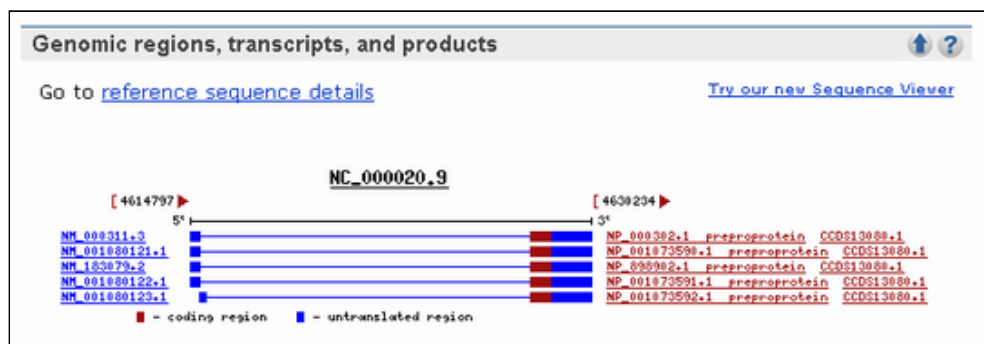
この遺伝子の変異に関する表現型は画面中央付近Phenotypesの欄に示されていて、クロイツフェルト・ヤコブ病を初めとして5つの病気が知られていることが確認できます。

Phenotypes
Creutzfeldt-Jakob disease MIM: 123400
Gerstmann-Straussler disease MIM: 137440
Huntington disease-like 1 MIM: 603218
Insomnia, fatal familial MIM: 600072
Prion disease with protracted course MIM: 606688

■ 解答1-2

[Entrez Geneからの選択的スプライス転写産物情報の取得](#)

Refseqに登録されているmRNAの情報は“Genomic regions, transcripts, and products”の欄に詳細に示されています。この遺伝子では5種類の選択的スプライス転写産物があることが確認できます。



[Entrez Geneからのホモログ情報の取得](#)

他の真核生物についてのホモログについて知るため、右側のメニューにあるLinksからHomologeneをクリックして下さい。

etazoa; Chordata; Craniata; Vertebrata;
Mammalia; Eutheria; Euarchontoglires; Primates;
Cetartarrhini; Hominidae; Homo

P; ASCR; PRIP; PrPc; CD230; prion; MGC26679;
P33-35C; PRNP

ncoded by this gene is a membrane
phatidylinositol-anchored glycoprotein that tends to
o rod-like structures. The encoded protein contains a
le region of five tandem octapeptide repeats. This
on chromosome 20, approximately 20 kbp upstream
ich encodes a biochemically and structurally similar
e one encoded by this gene. Mutations in the repeat
as elsewhere in this gene have been associated with
akob disease, fatal familial insomnia, Gerstmann-
ease, Huntington disease-like 1, and kuru. Alternative
s in multiple transcript variants encoding the same
vided by RefSeq]

Additional Links

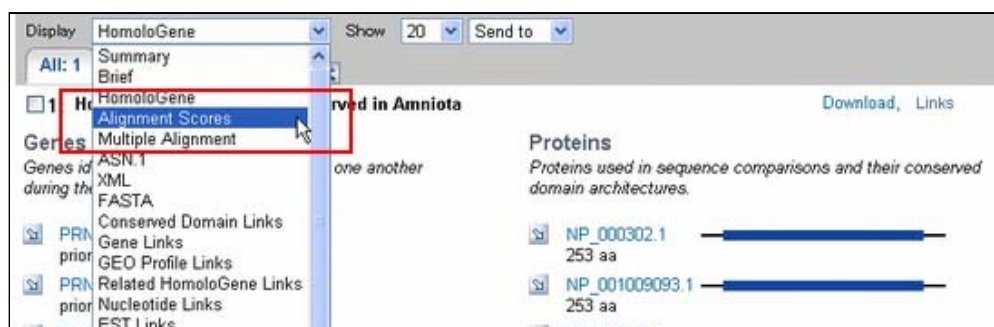
▼ **Links** **Explain**

- [Order cDNA clone](#)
- [CCDS](#)
- [Conserved Domains](#)
- [Genome](#)
- [GEO Profiles](#)
- [HomoloGene](#)
- ✓ [Map Viewer](#)
- [Nucleotide](#)
- [EST](#)
- [OMIM](#)
- [PubChem Compound](#)
- [PubChem Substance](#)
- [Full text in PMC](#)
- [Probe](#)
- [Protein](#)
- [PubMed](#)
- [PubMed \(OMIM\)](#)
- [PubMed \(GeneRIF\)](#)

以下のような画面が開きます。Homologeneではゲノム解読が進んでいる真核生物を対象にある遺伝子とホモログ関係にあると考えられる遺伝子をひとつにまとめた情報を提供しているデータベースです。これにより、どの系統まで同様の遺伝子が存在するのか、どの程度それらの配列は似ているかなどの情報を得ることができます。

The screenshot shows the NCBI HomoloGene web interface. The search bar contains 'HomoloGene'. The results section displays a list of genes and their corresponding proteins. The 'Genes' section lists several prion protein entries, including PRNP from Homo sapiens, Pan troglodytes, Canis lupus familiaris, Bos taurus, Mus musculus, Rattus norvegicus, and Gallus gallus. The 'Proteins' section lists the corresponding protein sequences, such as NP_000302.1 (253 aa) for Homo sapiens PRNP. The interface also includes a 'Protein Alignments' section at the bottom, which shows a multiple sequence alignment of the prion protein sequences.

配列間の相同性を知るために、Displayオプションから“Alignment Scores”を選択してください。



すると下図のように各生物種の配列間の類似度やSubstitution Rateが表示されます。

Gene		Identity (%)		Substitution Rates ¹			
Species	Symbol	Protein	DNA	d	d _p /d _s	d _{NR} /d _{NC}	
Homo sapiens							
vs. Pan troglodytes	PRNP	99.2	99.2	0.008	0.138	0.548	Blast
vs. Canis lupus familiaris	PRNP	87.7	87.3	0.139	0.126	0.282	Blast
vs. Bos taurus	PRNP	92.0	88.0	0.130	0.064	0.348	Blast
vs. Mus musculus	Pmp	90.1	85.3	0.163	0.077	0.342	Blast
vs. Rattus norvegicus	Pmp	89.7	86.8	0.145	0.090	0.260	Blast
vs. Gallus gallus	PRNP	47.1	57.4	0.631	0.399	0.878	Blast
Pan troglodytes							
vs. Homo sapiens	PRNP	99.2	99.2	0.008	0.138	0.548	Blast
vs. Canis lupus familiaris	PRNP	87.7	87.3	0.139	0.126	0.253	Blast
vs. Bos taurus	PRNP	92.0	87.9	0.132	0.063	0.300	Blast
vs. Mus musculus	Pmp	90.1	85.2	0.165	0.075	0.307	Blast
vs. Rattus norvegicus	Pmp	89.7	86.7	0.146	0.088	0.229	Blast
vs. Gallus gallus	PRNP	47.1	56.9	0.642	0.351	0.882	Blast
Canis lupus familiaris							
vs. Homo sapiens	PRNP	87.7	87.3	0.139	0.126	0.282	Blast
vs. Pan troglodytes	PRNP	87.7	87.3	0.139	0.126	0.253	Blast
vs. Bos taurus	PRNP	91.4	87.3	0.139	0.068	0.259	Blast
vs. Mus musculus	Pmp	86.9	82.5	0.200	0.092	0.335	Blast
vs. Rattus norvegicus	Pmp	86.9	83.6	0.185	0.099	0.300	Blast

表よりヒト・マウス間の配列類似度はアミノ酸レベルで90.1%、核酸レベルで85.3%であることがわかります。

また、右端のBlastリンクをクリックすると、BLAST2seqを利用したアミノ酸配列同士のアライメントを得ることもできます。

Blast Result

http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?program=blastp&\$log

こちらのブックマーク バーにブックマークを追加すると簡単にページにアクセスできます。

その他のブックマーク

NCBI Blast 2 Sequences results

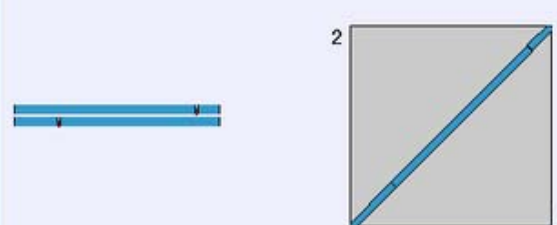
PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.18 [Mar-02-2008]

Matrix: BLOSUM62 gap open: 11 gap extension: 1
 x_dropoff: 0 expect: 10.00 wordsize: 3 Filter View option: Standard
 Masking character option: X for protein, n for nucleotide Masking color option: Black
☐ Show CDS translation Align

Sequence 1: [gi4506113|prion protein preproprotein \[Homo sapiens\]](#) > [gi34335270|refNP_898902.1|prion protein preproprotein \[Homo sapiens\]](#) > [gi122056623|refNP_001073590.1|prion protein preproprotein \[Homo sapiens\]](#) > [gi122056625|refNP_001073591.1|prion protein preproprotein \[Homo sapiens\]](#) > [gi122056628|refNP_001073592.1|prion protein preproprotein \[Homo sapiens\]](#) > [gi130912|sp|P04156.1|PRIO HUMAN RecName: Full=Major prion protein; Short=PrP; AltName: Full=PrP27-30; AltName: Full=PrP33-35C; AltName: Full=ASCR; AltName: CD a](#)
 Length = 253 (1 .. 253)

Sequence 2: [gi13173473|prion protein \[Mus musculus\]](#) > [gi130914|sp|P04925.2|PRIO MOUSE RecName: Full=Major prion protein; Short=PrP; AltName: Full=PrP27-30; AltName: Full=PrP33-35C; AltName: CD antigen=CD230; Flags: Precursor](#) > [gi200529|gb|AAA39997.1|prion protein](#) > [gi2865215|gb|AAC02804.1|short incubation prion protein Prnpa \[Mus musculus\]](#) > [gi13879449|gb|AAH06703.1|Prion protein \[Mus musculus\]](#) > [gi71060019|emb|CAJ18553.1|Prnp \[Mus musculus\]](#) > [gi74182795|db|BAE34724.1|unnamed protein product \[Mus musculus\]](#) > [gi74186646|d](#)
 Length = 254 (1 .. 254)



NOTE: Bitscore and expect value are calculated based on the size of the nr database.

Score = 502 bits (1293), Expect = 1e-140
 Identities = 227/255 (89%), Positives = 243/255 (95%), Gaps = 3/255 (1%)

Query	1	MANLGQMLVLFVATWSDGLCKKRPKPGGWTGGSRYPGGSPGGNRYPPGGGGWGGP	60
Sbjct	1	MANLGQMLVLFVATWSDGLCKKRPKPGGWTGGSRYPGGSPGGNRYPPGGGGWGGP	59
Query	61	HGGGPGGPHGGGPGGPHGGGPGGPHGGGPGGPGGTHSQNKPSKPKTNMKHMAAGAAAGA	120

■ 解答1-3

Entrez GeneからのSNPs情報の取得

Entrez Geneのページに戻り、右側メニューからSNP:GeneViewをクリックしてこの遺伝子に関するSNPs情報へと移動してください。

one encoded by this gene. Mutations in the repeat as elsewhere in this gene have been associated with job disease, fatal familial insomnia, Gerstmann-ase, Huntington disease-like 1, and kuru. Alternative in multiple transcript variants encoding the same ed by RefSeq]

Full text in PMC
 Probe
 Protein
 PubMed
 PubMed (OMIM)
 PubMed (GeneRIF)
 SNP
 SNP: Genotype
 SNP: GeneView
 Taxonomy
 UniSTS

products

SNP linked to Gene Ca bre...

http://www.ncbi.nlm.nih.gov/SNP/snp.cgi?to_cusid=5621

このページのブックマークバーにブックマークを追加すると簡単にページにアクセスできます。

その他のブックマーク

GENERAL

HUMAN VARIATION

Search, Annotate, Submit

Annotate and Submit Batch Data with Clinical Impact

SNP SUBMISSION

DOCUMENTATION

SEARCH

RELATED SITES

Total gene model (contig mRNA transcript):

Contig	mRNA	transcript	protein	mRNA orientation	Contig	Contig Label	List SNP
NM_000311.3	plus strand	NP_000302.1	forward	NT_011387.8	reference	< currently shown	
NM_000311.3	plus strand	NP_000302.1	forward	NW_927317.1	Celera	View snp on GeneModel	
NM_000311.3	plus strand	NP_000302.1	forward	NW_001838652.1	HuRef	View snp on GeneModel	
NM_001080121.1	plus strand	NP_001073590.1	forward	NT_011387.8	reference	View snp on GeneModel	
NM_001080121.1	plus strand	NP_001073590.1	forward	NW_927317.1	Celera	View snp on GeneModel	
NM_001080121.1	plus strand	NP_001073590.1	forward	NW_001838652.1	HuRef	View snp on GeneModel	
NM_001080122.1	plus strand	NP_001073591.1	forward	NT_011387.8	reference	View snp on GeneModel	
NM_001080122.1	plus strand	NP_001073591.1	forward	NW_927317.1	Celera	View snp on GeneModel	
NM_001080122.1	plus strand	NP_001073591.1	forward	NW_001838652.1	HuRef	View snp on GeneModel	
NM_001080123.1	plus strand	NP_001073592.1	forward	NT_011387.8	reference	View snp on GeneModel	
NM_001080123.1	plus strand	NP_001073592.1	forward	NW_927317.1	Celera	View snp on GeneModel	
NM_001080123.1	plus strand	NP_001073592.1	forward	NW_001838652.1	HuRef	View snp on GeneModel	
NM_183079.2	plus strand	NP_898202.1	forward	NT_011387.8	reference	View snp on GeneModel	
NM_183079.2	plus strand	NP_898202.1	forward	NW_927317.1	Celera	View snp on GeneModel	
NM_183079.2	plus strand	NP_898202.1	forward	NW_001838652.1	HuRef	View snp on GeneModel	

☐ Include clinically associated ☐ in gene region ☒ cSNP ☐ has frequency ☐ double hit

gene model (contig mRNA transcript):

Contig Label	Contig	mRNA	protein	mRNA orientation	transcript snp count
reference	NT_011387.8	NM_000311.3	NP_000302.1	forward	plus strand 24, coding

Color Legend

Region	Contig position	mRNA pos	dbSNP rs#	Heterozygosity	Validation	3D	Clinically Associated	Function	dbSNP allele	Protein residue	Codon pos	Amino acid pos	PubMed
exon_1	4619867	373						start codon				1	
exon_2	4619942	448	rs11538755	N.D.				missense	A	Thr [T]	1	26	
								contig reference	C	Pro [P]	1	26	
	4619949	455	rs11538762	N.D.				missense	A	His [H]	2	28	
								contig reference	C	Pro [P]	2	28	
	4619975	481	rs11538763	N.D.				synonymous	A	Arg [R]	1	37	

この表を読み取ると非同義置換(表中赤く示されている)は17個存在することがわかります。

続いて、OMIMへのリンクを見てみましょう。デフォルトの設定ではOMIMへのリンクは表示されていません。

画面中の"Include clinically associated"と書かれたチェックボックスにチェックを入れて右端のrefreshボタンを押すことで表中にOMIMへのリンクが表示されます。

☒ Include clinically associated ☐ in gene region ☒ cSNP ☐ has frequency ☐ double hit

gene model (contig mRNA transcript):



Contig Label	Contig	mRNA	protein	mRNA orientation	transcript snp count
reference	NT_011387.8	NM_000311.3	NP_000302.1	forward	plus strand 24, coding

Color Legend

Region	Contig position	mRNA pos	dbSNP rs#	Heterozygosity	Validation	3D	Clinically Associated	Function	dbSNP allele	Protein residue	Codon pos	Amino acid pos	PubMed
exon_1	4619867	373						start codon				1	
exon_2	4619942	448	rs11538755	N.D.				missense	A	Thr [T]	1	26	
								contig reference	C	Pro [P]	1	26	
	4619949	455	rs11538762	N.D.				missense	A	His [H]	2	28	
								contig reference	C	Pro [P]	2	28	

OMIMへのリンクは、表中の"Clinically Associated"のカラムにヒトの形をしたマークで表示されます。



Region	Contig position	mRNA pos	dbSNP rs#	Heterozygosity	Validation	3D	Clinically Associated	Function	dbSNP allele	Protein residue	Codon pos	Amino acid pos	PubMed
exon_1	4619867	373						start codon				1	
exon_2	4619942	448	rs11538755	N.D.				missense	A	Thr [T]	1	26	
								contig reference	C	Pro [P]	1	26	
	4619949	455	rs11538762	N.D.				missense	A	His [H]	2	28	
								contig reference	C	Pro [P]	2	28	

4620464	970	rs28933385	N.D.			missense	A	Lys [K]	1	200	
						config reference	G	Glu [E]	1	200	
4620488	994	rs5826226	N.D.			missense	T	Cys [C]	1	208	
						config reference	C	Arg [R]	1	208	

表よりOMIMデータベースへのリンクが含まれるSNPs は200残基目の1箇所であることがわかります（上下二カラムに書いてあるが同じポジションのデータ）。

これをクリックし、OMIMの記載とSNPsテーブルとを比較してください。同じアミノ酸位置の、同じ非同義置換情報がSNPsのページとOMIMの両者に記載されていることが確認できます。

200残基：Glu/Lys

4620464	970	rs28933385	N.D.			missense	A	Lys [K]	1	200	
						config reference	G	Glu [E]	1	200	

.0006 CREUTZFELDT-JAKOB DISEASE [PRNP, **GLU200LYS] dbSNP**

FATAL FAMILIAL INSOMNIA, INCLUDED

In 2 patients with Creutzfeldt-Jakob disease ([123400](#)) from the same family, [Goldgaber](#) transition in the PRNP gene, resulting in a glu200-to-lys (E200K) substitution.

Studying an unusual cluster of cases of CJD in rural Slovakia, [Goldfarb et al. \(1990\)](#) found 4 tested cases of 'focal CJD,' in 12 of 40 healthy first-degree relatives, and in 6 of 23 other extrafocal cases or their relatives had the mutation; nor did any unrelated individuals with the mutation. One of the healthy individuals with the E200K mutation was the 75-year-old mother of a patient. This was interpreted as indicating that the mutation is a necessary, but not sufficient, factor for the disease as scrapie-infected sheep was proposed. 🤖

PRNPの変異に関連した病気の原因遺伝子であることがOMIMのエントリより読み取れます。

■ 解答1-4

Entrez Geneから類似配列情報の取得

Entrez Geneのページに戻り、“Genomic regions, transcripts, and products”に示された NP_000302をクリックし、さらに表示されるプルダウンメニューからBlinkを選択してください。

Genomic regions, transcripts, and products

Go to [reference sequence details](#) [Try our new Sequence](#)

NC_000020.9

[4614797] 5' [4630234] 3'

■ = coding region ■ = untranslated region

Links

PROTEIN LINKS

FASTA

GENPEPT

Blink

Conserved Domains

それによりNP_000302に類似した配列が一覧表示されます。Blinkはあらかじめblastを用いて計算された類似配列情報です。

BLINK precomputed BLAST

Home Taxonomy Report Multiple Alignment Blast Help

Pre-computed BLAST results for: [gil4506113refNP_000302.1 prion protein preproprotein \[Homo sapiens\]](#)

Matching gis: [3996152.3999540.4000210.5993964.40075008.10053355.11079226.34335270.15108368.217228859.217228861.217256691.1](#)

Total (score > 100) : 56 hits in 53 proteins in 15 species

Selected: 56 hits in 53 proteins in 15 species Filter: Min Score: 100 | Keep Only: 3D structures (PDB) |

Other views (Reports): [Taxonomy report](#) [Multiple Alignment](#) [Blast](#)

[Reset all filters](#)

[Choose Display Options](#)

0 Archaea 0 Bacteria 54 Metazoa 0 Fungi 0 Plants 0 Viruses 2 The Others [reset selection](#)

Results: 1 - 53

% hits 253 aa

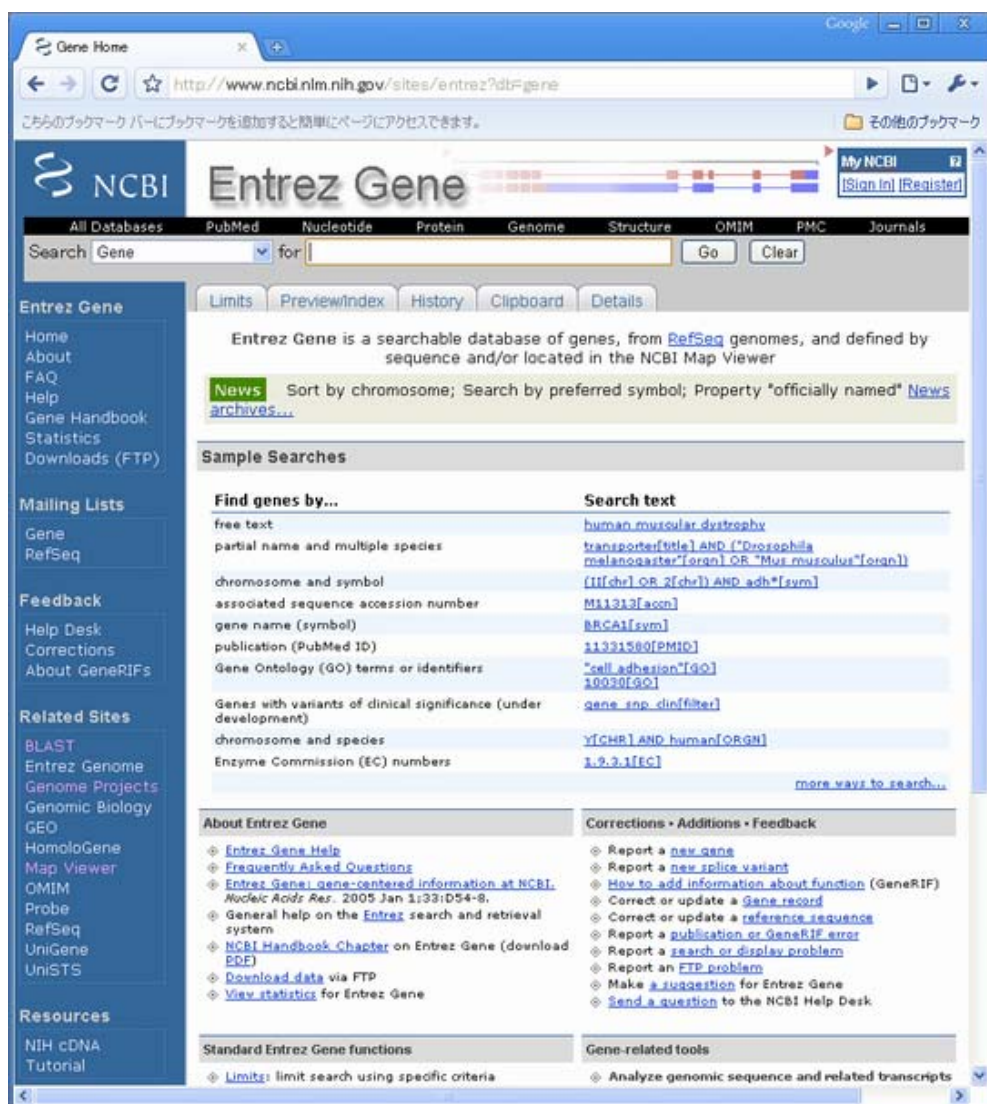
blink

SCORE	ACCESSION	Length	Protein Description
Conserved Domain Database hits			
1203	1Q1X-A	210	Chain A, Human Prion Protein
1203	1Q1X-A	210	Chain A, Human Prion Protein
1136	1QX0-A	219	Chain A, Bovine Prion Protein Residues 23-230
771	1QM1-A	143	Chain A, Human Prion Protein Fragment 90-230
771	1QM0-A	143	Chain A, Human Prion Protein Fragment 90-230
768	1F07-A	142	Chain A, Human Prion Protein Mutant E200k Fragment 90-231
768	1F07-A	142	Chain A, Human Prion Protein (Mutant E200k) Fragment 90-231
682	1B10-A	142	Chain A, Solution NMR Structure Of Recombinant Syrian Hamster Prion Protein
683	2Z0H-A	148	Chain A, NMR Structure Of Rabbit Prion Protein Mutation S146G
678	2F13-A	138	Chain A, NMR Solution Of Rabbit Prion Protein (91-228)
677	2Z0H-A	148	Chain A, NMR Structure Of Rabbit Prion Protein Mutation I146V
616	1TFX-A	121	Chain A, Ovine Recombinant Prp(114-234), Arg Variant In C
604	1QM3-A	112	Chain A, Human Prion Protein Fragment 121-230
604	1QM2-A	112	Chain A, Human Prion Protein Fragment 121-230
587	1I4H-A	108	Chain A, Crystal Structure Of The Human Prion Protein Rev
584	1QW7-A	112	Chain A, Bovine Prion Protein Fragment 121-230
584	1QW7-A	112	Chain A, Bovine Prion Protein Fragment 121-230
583	1H01-A	112	Chain A, Human Prion Protein 121-230 M166C221C
576	1Y2S-A	113	Chain A, Ovine Prion Protein Variant R168
569	1XCV-A	111	Chain A, Solution Structure Of The Sheep Prion Protein Wit
563	1XCV-A	111	Chain A, Elk Prion Protein

(これ以降のペアワイズアライメントから立体構造を表示させる機能は、現在では提供されていません。)

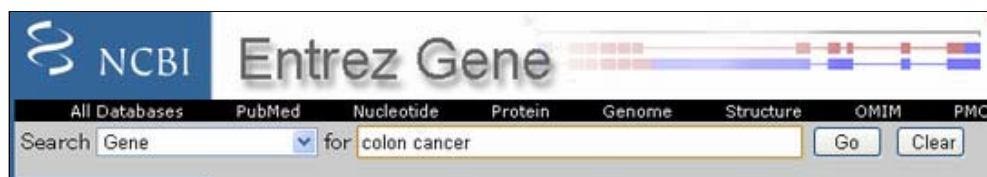
■ 解答2-1

課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を選択し、その次のページからGeneを選択することで、Entrez Geneのトップ画面に移動します。



Entrez Geneでのキーワード検索2

続いて、上部のテキストボックスに“colon cancer”と入力し、その右横の“Go”ボタンを押すことで“colon cancer”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複数の検索結果が得られます。その中にHuman (Homo Sapiens)に関する MLH1が3番目に表示されていることが確認できます。

colon cancer - Gene Resu... x

http://www.ncbi.nlm.nih.gov/sites/entrez

こちらのブックマーク バーにブックマークを追加すると簡単にページにアクセスできます。 その他のブックマーク

NCBI Entrez Gene

All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Jo

Search Gene for colon cancer Go Clear Save Se

Limits Preview/Index History Clipboard Details

Display Summary Show 20 Sort by Relevance Send to

All: 673 Current Only: 602 Genes Genomes: 586 SNP GeneView: 452

Items 1 - 20 of 673 Page 1 of 34 Next

☐ 1: [PTPRJ](#) Order cDNA clone, Links
Official Symbol PTPRJ and Name: protein tyrosine phosphatase, receptor type, J [*Homo sapiens*]
Other Aliases: CD148, DEP1, HPTPeta, R-PTP-ETA, SCC1
Other Designations: CD148 antigen; human density enhanced phosphatase-1; protein tyrosine phosphatase, receptor type, J polypeptide; receptor-type tyrosine-protein phosphatase eta; susceptibility to colon cancer 1, mouse, homolog of
Chromosome: 11; Location: 11p11.2
Annotation: Chromosome 11, NC_000011.8 (47958686..48148970)
MIM: 600925
GeneID: 5795

☐ 2: [MSH2](#) Order cDNA clone, Links
Official Symbol MSH2 and Name: mutS homolog 2, colon cancer, nonpolyposis type 1 (E. coli) [*Homo sapiens*]
Other Aliases: COCA1, FCC1, HNPCC, HNPCC1, LCFS2
Other Designations: mutS homolog 2
Chromosome: 2; Location: 2p22-p21
Annotation: Chromosome 2, NC_000002.10 (47483767..47563864)
MIM: 609309
GeneID: 4436

☐ 3: [MLH1](#) Order cDNA clone, Links
Official Symbol MLH1 and Name: mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli) [*Homo sapiens*]
Other Aliases: COCA2, FCC2, HNPCC, HNPCC2, MGC5172, hMLH1
Other Designations: DNA mismatch repair protein Mlh1; MutL protein homolog 1
Chromosome: 3; Location: 3p21.3
Annotation: Chromosome 3, NC_000003.10 (37009983..37067341)
MIM: 120436
GeneID: 4292

☐ 4: [Bmi_50400](#) Links
colon cancer-associated protein Mic1 [*Brugia malayi*]

Recent Activity
Turn
Your browsing activit
empty.

[Entrez Gene Reportの読み方2](#)

赤四角で囲まれたMLH1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。

Entrez Gene: MLH1 mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli) [Homo sapiens]

GeneID: 4292 updated 15-Jan-2009

Summary

Official Symbol MLH1 provided by HGNC

Official Full Name mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli) provided by HGNC

Primary source HGNC:7127

See related Ensembl: ENSG00000076242; HPRD:00390; MIM:120436

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 FCC2; COCA2; HNPCC; hMLH1; HNPCC2; MGC5172; MLH1

Summary This gene was identified as a locus frequently mutated in hereditary nonpolyposis colon cancer (HNPCC). It is a human homolog of the E. coli DNA mismatch repair gene mutL, consistent with the characteristic alterations in microsatellite sequences (RER+ phenotype) found in HNPCC. Alternatively spliced transcript variants encoding different isoforms have been described, but their full-length natures have not been determined. [provided by RefSeq]

Genomic regions, transcripts, and products

Go to [reference sequence details](#) Try our new [Sequence Viewer](#)

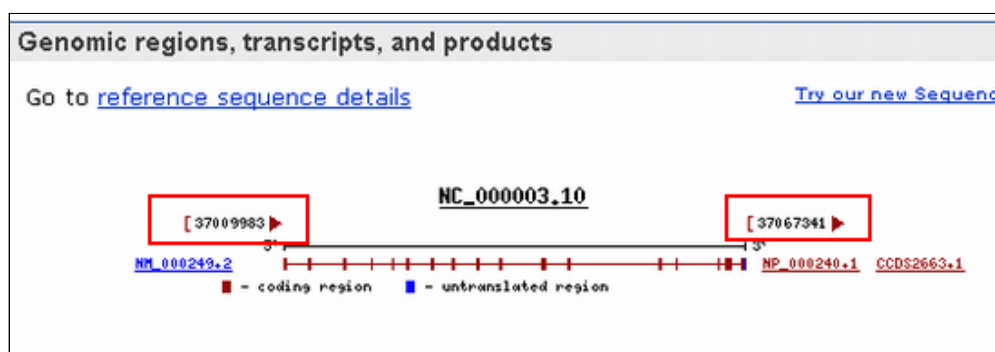
Table Of Contents

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

Links

- Order cDNA clone
- Books
- CCDS
- Conserved Domains
- Genome
- GEO Profiles
- HomoloGene
- Map Viewer
- Nucleotide
- EST
- OMIM
- PubChem Compound
- PubChem Substance
- Full text in PMC
- Probe
- Protein
- PubMed
- PubMed (OMIM)
- PubMed (GeneRIF)
- SNP
- SNP VarView
- SNP: Genotype
- SNP: GeneView

少し下にスクロールすると、ゲノム上での位置などに関する情報がGenomic context, Genomic regions などの欄に示されています。これからMLH1はヒトゲノム3p21.3の37,009,983bpから 37,067,341bpに位置することがわかります。



機能に関する記述はSummaryの部分に文章で、あるいはGene Ontologyの欄にオントロジーを利用して記述されています。これらの情報によるとMLH1はミスマッチ修復遺伝子としての機能を持っており、大腸菌ミスマッチ修復遺伝子mutLのヒト相同体であり、HNPCCに見られるマイクロサテライト配列(RER+表現型)の特徴的な変更が一致しているといったことがわかっています。

Summary	
Official Symbol	MLH1 <small>provided by HGNC</small>
Official Full Name	mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli) <small>provided by HGNC</small>
Primary source	HGNC:7127
See related	Ensembl:ENSG00000076242 ; HPRD:00390 ; MIM:120436
Gene type	protein coding
RefSeq status	REVIEWED
Organism	Homo sapiens
Lineage	<i>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo</i>
Also known as	FCC2; COCA2; HNPCC; hMLH1; HNPCC2; MGC5172; MLH1
Summary	This gene was identified as a locus frequently mutated in hereditary nonpolyposis colon cancer (HNPCC). It is a human homolog of the E. coli DNA mismatch repair gene mutL, consistent with the characteristic alterations in microsatellite sequences (RER+ phenotype) found in HNPCC. Alternatively spliced transcript variants encoding different isoforms have been described, but their full-length natures have not been determined. [provided by RefSeq]

また、SummaryのAlso known asの記述からFCC2, COCA2, HNPCCなど多くの別名で呼ばれていることも確認できます。

Summary	
Official Symbol	MLH1 <small>provided by HGNC</small>
Official Full Name	mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli) <small>provided by HGNC</small>
Primary source	HGNC:7127
See related	Ensembl:ENSG00000076242 ; HPRD:00390 ; MIM:120436
Gene type	protein coding
RefSeq status	REVIEWED
Organism	Homo sapiens
Lineage	<i>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo</i>
Also known as	FCC2; COCA2; HNPCC; hMLH1; HNPCC2; MGC5172; MLH1
Summary	This gene was identified as a locus frequently mutated in hereditary nonpolyposis colon cancer (HNPCC). It is a human homolog of the E. coli DNA mismatch repair gene mutL, consistent with the characteristic alterations in microsatellite sequences (RER+ phenotype) found in HNPCC. Alternatively spliced transcript variants encoding different isoforms have been described, but their full-length natures have not been determined. [provided by RefSeq]

この遺伝子の変異に関する表現型は画面中央付近Phenotypesの欄に示されていて、5つの病気が知られていることが確認できます。

Phenotypes

Cafe-au-lait spots with glioma or leukemia

[MIM: 114030](#)

Colorectal cancer, hereditary nonpolyposis, type 2

[MIM: 609310](#)

Mismatch repair cancer syndrome

[MIM: 276300](#)

Muir-Torre syndrome

[MIM: 158320](#)

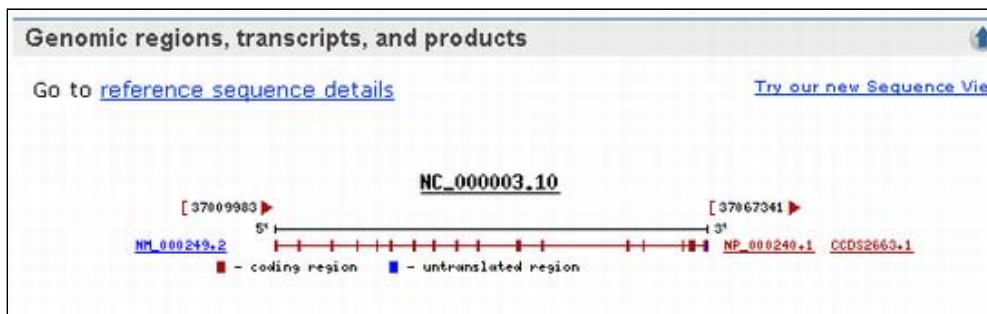
Turcot syndrome with glioblastoma

[MIM: 276300](#)

■ 解答2-2

[Entrez Geneからの選択的スプライス転写産物情報の取得2](#)

Refseqに登録されているmRNAの情報は“Genomic regions, transcripts, and products”の欄に詳細に示されています。この遺伝子では選択的スプライス転写産物が知られていないことが確認できます。



[Entrez Geneからのホモログ情報の取得2](#)

他の真核生物についてのホモログについて知るため、右側のメニューにあるLinksからHomologeneをクリックして下さい。

以下のような画面が開きます。

HomoloGene Result

http://www.ncbi.nlm.nih.gov/sites/entrez?itool=genefullreport&DbFrom=genefullreport&Cmd=Link&

こちらのブックマーク バーにブックマークを追加すると簡単にページにアクセスできます。

NCBI HomoloGene Discover Homologs

All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Journals

Search HomoloGene for Go Clear

Limits Preview/Index History Clipboard Details

Display HomoloGene Show 20 Send to

All: 1 Fungi: 0 Mammals: 0

1: HomoloGene:208. Gene conserved in Fungi/Metazoa group

Download, Links

Genes
Genes identified as putative homologs of one another during the construction of HomoloGene.

Proteins
Proteins used in sequence comparisons and their conserved domain architectures.

MLH1, <i>Homo sapiens</i> mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	NP_000240.1 756 aa
MLH1, <i>Pan troglodytes</i> mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	XP_001170433.1 756 aa
MLH1, <i>Canis lupus familiaris</i> mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	XP_534219.2 757 aa
MLH1, <i>Bos taurus</i> mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	NP_001069462.1 233 aa
Mlh1, <i>Mus musculus</i> mutL homolog 1 (E. coli)	NP_081086.1 760 aa
Mlh1, <i>Rattus norvegicus</i> mutL homolog 1 (E. coli)	NP_112315.1 757 aa
MLH1, <i>Gallus gallus</i> mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	XP_418828.1 757 aa
mlh1, <i>Danio rerio</i> mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	NP_956953.1 724 aa
Mlh1, <i>Drosophila melanogaster</i> Mlh1	NP_477022.1 664 aa
AgaP_AGAP012192, <i>Anopheles gambiae</i> AGAP012192-PA	XP_320342.2 671 aa
ENSANGG00000010995, <i>Anopheles gambiae</i> ENSANGP00000013484	XP_307435.2 395 aa

配列間の相同性を知るために、Displayオプションから“Alignment Scores”を選択してください。すると下図のように各生物種の配列間の類似度やSubstitution Rateが表形式で表示されます。

HomoloGene Result

http://www.ncbi.nlm.nih.gov/sites/entrez

NCBI HomoloGene Discover Homologs

Search HomoloGene for [] Go Clear

Limits Preview Index History Clipboard Details

Display Alignment Scores Show 20 Send to

All: 1 Fungi: 0 Mammals: 0

1: HomoloGene:208. Gene conserved in Fungi/Metazoa group Download Links

Pairwise Alignment Scores

Species	Gene	Symbol	Identity (%)		Substitution Rates ¹			
			Protein	DNA	d	d _N /d _S	d _{NR} /d _{NC}	
Homo sapiens	MLH1							
vs. Pan troglodytes	MLH1		99.6	99.7	0.003	0.287	1.226	Blast
vs. Canis lupus familiaris	MLH1		93.4	90.9	0.097	0.091	0.411	Blast
vs. Bos taurus	MLH1		94.8	91.3	0.093	0.059	0.190	Blast
vs. Mus musculus	Mlh1		88.6	87.3	0.140	0.125	0.680	Blast
vs. Rattus norvegicus	Mlh1		87.3	85.4	0.162	0.126	0.580	Blast
vs. Gallus gallus	MLH1		80.4	73.7	0.325	0.050	0.743	Blast
vs. Danio rerio	mlh1		71.4	66.6	0.442	undef	0.865	Blast
vs. Drosophila melanogaster	Mlh1		52.5	55.6	0.672	undef	0.888	Blast
vs. Anopheles gambiae	AgaP_AGAP012192		52.9	55.5	0.676	undef	0.752	Blast
vs. Anopheles gambiae	ENSANGG00000010995		46.6	51.9	0.770	undef	0.799	Blast
vs. Schizosaccharomyces pombe	mlh1		42.1	48.7	0.865	undef	0.944	Blast
vs. Saccharomyces cerevisiae	MLH1		40.6	48.0	0.886	undef	0.811	Blast
vs. Kluyveromyces lactis	KLLA000995g		41.4	49.4	0.843	undef	0.831	Blast
vs. Eremothecium gossypii	AGOS_AFL199C		40.5	47.4	0.908	undef	0.893	Blast
vs. Magnaporthe oryzae	MGG_07121		46.0	49.4	0.842	undef	0.787	Blast
vs. Neurospora crassa	NCU08309.1		48.9	51.6	0.778	undef	0.850	Blast
Pan troglodytes	MLH1							
vs. Homo sapiens	MLH1		99.6	99.7	0.003	0.287	1.226	Blast
vs. Canis lupus familiaris	MLH1		93.5	90.9	0.097	0.090	0.422	Blast
vs. Bos taurus	MLH1		94.8	91.4	0.091	0.061	0.190	Blast

表赤く囲んだ領域よりヒト・マウス間の配列類似度はアミノ酸レベルで88.6%、核酸レベルで87.3%であることがわかります。

また、右端のBlastリンクをクリックすると、BLAST2seqを利用したアミノ酸配列同士のアライメントを得ることができます。

Blast Result

http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?program=blastp&\$log=

こちらのブックマーク バーにブックマークを追加すると簡単にページにアクセスできます。

その他のブックマーク

NCBI Blast 2 Sequences results

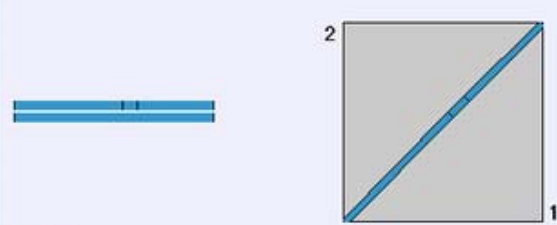
PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.18 [Mar-02-2008]

Matrix: BLOSUM62 gap open: 11 gap extension: 1
 x_dropoff: 0 expect: 10.00 wordsize: 3 Filter ☐ View option: Standard
 Masking character option: X for protein, n for nucleotide Masking color option: Black
☐ Show CDS translation Align

Sequence 1: [gi|4557757|MutL protein homolog 1 \[Homo sapiens\]](#)
[>gi|730028|sp|P40692.1|MLH1 HUMAN RecName: Full=DNA mismatch repair protein Mh1; AltName: Full=MutL protein homolog 1](#) [>gi|463989|gb|AAC50285.1| DNA mismatch repair protein homolog \[Homo sapiens\]](#) [>gi|1079787|gb|AA082079.1| DNA mismatch repair protein homolog](#)
[>gi|13905126|gb|AAH08850.1| MutL homolog 1, colon cancer, nonpolyposis type 2 \(E. coli\) \[Homo sapiens\]](#) [>gi|27805155|gb|AAO22994.1| mutL homolog 1, colon cancer, nonpolyposis type 2 \(E. coli\) \[Homo sapiens\]](#) [>gi|11...](#)
 Length = 756 (1 ... 756)

Sequence 2: [gi|19387852|MutL protein homolog 1 \[Mus musculus\]](#) [>gi|18255308|gb|AAH21815.1| MutL homolog 1 \(E. coli\) \[Mus musculus\]](#)
 Length = 760 (1 ... 760)



NOTE: Bitscore and expect value are calculated based on the size of the nr database.

Score = 1328 bits (3438), Expect = 0.0
 Identities = 670/760 (88%), Positives = 713/760 (93%), Gaps = 4/760 (0%)

Query	1	MSFVAGVIRRLDET VVNRIAAGEVIORPANA IKEMIENCLDAKSTSIQVIVKEGGLKLIQ	60
Sbjct	1	M+ FVAGVIRRLDET VVNRIAAGEVIORPANA IKEMIENCLDAKST+IQV+VKEGGLKLIQ	60
Query	61	IQDNGTGIRKEDLDIVCERFTTSKLQSFEDLASISTYGFGEALASISHVAHYTITTKTA	120

■ 解答2-3

Entrez GeneからのSNPs情報の取得2

Entrez Geneのページに戻り、右側メニューからSNP:GeneViewをクリックしてこの遺伝子に関するSNPs情報へと移動してください。

SNP linked to Gene Ca free...

http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?locus=4292

あなたのブックマーク バーにブックマークを追加すると簡単にページにアクセスできます。

NCBI Single Nucleotide Polymorphism

PubMed Nucleotide Protein Genome Structure PopSet Taxonomy QMM Books SNP

Search Entrez: SNP for Go

Search for SNP on NCBI Reference Assembly

BUILD 129

Need a question about dbSNP? Try searching the SNP FAQ Archive

Go

GENERAL

HUMAN VARIATION

Search, Annotate, Submit

Submit

Annotate and Submit Batch Data with Clinical Impact

SNP SUBMISSION

DOCUMENTATION

SEARCH

RELATED SITES

SNP linked to Gene **MLH1** (geneID:4292) Via Contig Annotation

Send rs# on all gene models to Batch Query Download all rs# to file

Gene Model (mRNA alignment) information from genome sequence

Total gene model (contig mRNA transcript):				3	Contig Label	List SNP
mRNA	transcript	protein	mRNA orientation	Contig		
NM_000249.2	plus strand	NP_000240.1	forward	NT_022517.17	reference	<- currently shown
NM_000249.2	plus strand	NP_000240.1	forward	NW_921651.1	Celera	View snp on GeneModel
NM_000249.2	minus strand	NP_000240.1	reverse	NW_001838877.2	HuRef	View snp on GeneModel

☐ Include clinically associated ☐ in gene region ☒ cSNP ☐ has frequency ☐ double hit [refresh](#)

gene model	Contig Label	Contig	mRNA	protein	mRNA orientation	transcript	snp count
(contig mRNA transcript):	reference	NT_022517.17	NM_000249.2	NP_000240.1	forward	plus strand 15, coding	

Color Legend

Region	Contig position	mRNA pos	dbSNP rs#	Heterozygosity	Validation	3D	Clinically Associated	Function	dbSNP allele	Protein residue	Codon pos	Amino acid pos	PubMed
exon_1	36975043	61						start codon				1	
exon_1	36975081	92	rs1890143	N.D.		Yes		synonymous	A	Glu [E]	3	13	
								coding reference	G	Glu [E]	3	13	
exon_4	36985967	438	rs34665507	N.D.		Yes		synonymous	T	Tyr [Y]	3	126	
								coding reference	C	Tyr [Y]	3	126	
exon_6	36990329	534	rs4647256	0.036	WH	Yes		synonymous	T	Asn [N]	3	158	
								coding reference	C	Asn [N]	3	158	
exon_7	36993321	612	rs35225190	N.D.		Yes		synonymous	T	Ser [S]	3	184	
								coding reference	A	Ser [S]	3	184	

SNP linked to Gene Ca free...

http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?pos=4282

こちらのブックマークバーにブックマークを追加すると簡単にページにアクセスできます。

その他のブックマーク

exon_7	36993321	612	rs35225190	N.D.	Yes	synonymous	T	Ser [S]	3	184
						contig	A	Ser [S]	3	184
	36993324	616	rs11541862	N.D.		missense	T	Tyr [Y]	1	186
						contig	C	His [H]	1	186
	36993352	643	rs35847123	N.D.		frame shift	-		1	195
						contig	A	Lys [K]	1	195
exon_9	36996039	850	rs35386320	N.D.		missense	G	Gly [G]	1	264
						contig	C	Arg [R]	1	264
exon_12	37007284	1251	rs35164771	N.D.		synonymous	T	Leu [L]	3	397
						contig	G	Leu [L]	3	397
	37007447	1414	rs34385582	N.D.		missense	T	Ser [S]	1	452
						contig	A	Thr [T]	1	452
exon_15	37033823	1790	rs36185292	N.D.		missense	T	Leu [L]	2	577
						contig	C	Ser [S]	2	577
exon_17	37030023	1268	rs1800145	N.D.		synonymous	A	Leu [L]	3	636
						contig	G	Leu [L]	3	636
exon_18	37030474	1225	rs11542214	N.D.		missense	A	Lys [K]	1	689
						contig	C	Gln [Q]	1	689
exon_19	37031925	2178	rs1800147	N.D.		synonymous	T	Gly [G]	3	706
						contig	C	Gly [G]	3	706
	37032031	2214	rs1800148	N.D.		synonymous	T	His [H]	3	718
						contig	C	His [H]	3	718
	37032042	2245	rs1800149	N.D.		missense	G	Val [V]	1	729
						contig	C	Leu [L]	1	729

GENERAL: Contact Us | Homepage | Announcements | SNP Summary | Genome | FTP SERVER | Build History | Handle Request
DOCUMENTATION: FAQ | Searchable FAQ Archive | Consensus | How to Submit | RefSNP Summary Info | Database Schema
SEARCH: Enter a SNP | Blast SNP | Batch Query | By Submitter | By Batch | By Method | Population | Publication | Batch | Locus Info | Between Markers
HAFLPOTYPE: Submission | Specifications | Sample HapSet | Sample Indexes
NCBI: PubMed | Entrez | BLAST | OMIM | Taxonomy | Structure

この表を読み取ると非同義置換(表中赤く示されている)は6個存在することがわかります。

続いて、OMIMへのリンクを見てみましょう。デフォルトの設定ではOMIMへのリンクは表示されていません。

画面中の“Include clinically associated”と書かれたチェックボックスにチェックを入れて右端のrefreshボタンを押すことで表中にOMIMへのリンクが表示されます。

OMIMへのリンクは、表中の“Clinically Associated”のカラムにヒトの形をしたマークで表示されます。

exon_5	36988499	454	rs28930073	N.D.		missense	C	His [H]	1	132
						contig	G	Asp [D]	1	132

表よりOMIMデータベースへのリンクが含まれるSNPsは132残基目の1箇所であることがわかります(上下二カラムに書いてあるが同じポジションのデータ)。

これをクリックし、OMIMの記載とSNPsテーブルとを比較してください。同じアミノ酸位置の、同じ非同義置換情報がSNPsのページとOMIMの両者に記載されていることが確認できます。

132残基 : His/Asp

exon_5	36988499	454	rs28930073	N.D.		missense	C	His [H]	1	132
						contig	G	Asp [D]	1	132

.0019 COLORECTAL CANCER, SPORADIC, SUSCEPTIBILITY TO [MLH1, ASP132HIS] dbSNP

Using a novel high density oligonucleotide array (HNPCC Chip) to look for variants in the MLH1, MSH2 (609309), and M (600678) genes in Israeli probands with familial colorectal cancer (CRC; 114500) unstratified with respect to the microsatellite instability phenotype, Lipkin et al. (2004) identified a 415G-C translation in the MLH1 gene, resulting in an asp132-to-h amino acid substitution. MLH1 415C conferred clinically significant susceptibility to CRC. In contrast to classic HNPCC associated with MLH1 415C usually did not have the microsatellite instability (MSI) defect, which is important for clinic mutation screening. Structural and functional analyses showed that the normal ATPase function of MLH1 was attenuated but not eliminated, by the MLH1 415G-C mutation.

MLH1の変異に関連した病気の原因遺伝子であることがOMIMのエントリより読み取れます。

■ 解答2-4

Entrez Geneから類似配列情報の取得

Entrez Geneのページに戻り、“Genomic regions, transcripts, and products”に示された NP_000240をクリックし、さらに表示されるプルダウンメニューからBlinkを選択してください。それによりNP_000240に類似した配列が一覧表示されます。

BLINK precomputed BLAST

Pre-computed BLAST results for: gi4557757refNP_000240.1 MutL protein homolog 1 [Homo sapiens]

Matching gis: 33738032.53932122.155685496.157928134.157928839.730028.741682.1079787.91132684.155119205.463989.119584889.27805155.211637

Total (score > 100) : 4517 hits in 4457 proteins in 1313 species

Selected: 4517 hits in 4457 proteins in 1313 species Filter: Min Score: 100

Other views (Reports): Taxonomy report Multiple Alignment Blast

Reset all filters

Choose Display Options

Archaea 2470 Bacteria 441 Metazoa 326 Fungi 63 Plants 0 Viruses 1177 The Others reset selection

Results: 1 - 100 Next Page Last

SCORE	ACCESSION	Length	Protein Description
3862	ABV03362	756	mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli) [synth
3862	ABH02701	756	mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli) [synth
3862	F40492	756	RecName: Full=DNA mismatch repair protein Mlh1; AltName: Full=MutL
3862	U1741685	756	DNA mismatch repair protein
3862	AAA02079	756	DNA mismatch repair protein homolog
3862	AA020285	756	DNA mismatch repair protein homolog [Homo sapiens]
3862	KAF64485	756	mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli), isofor
3862	AA023229	756	mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli) [Homo s
3862	AAH06350	756	MutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli) [Homo s
3862	AA002400	757	mutL-like 1, colon cancer, nonpolyposis type 2 [synthetic construct
3868	AAA17379	756	human homolog of E. coli mutL gene product, Swiss-Prot Accession Nu
3865	BAG25827	756	unnamed protein product [Homo sapiens]
3864	AAT44531	755	hypothetical rhabdomyosarcoma antigen MU-RMS-40.5 [Homo sapiens]
3860	XP_001170432	756	PREDICTED: MutL protein homolog 1 isoform 8 [Pan troglodytes]
3860	XP_001170274	756	PREDICTED: MutL protein homolog 1 isoform 3 [Pan troglodytes]
3850	BAD26530	756	MutL protein homolog 1 variant [Homo sapiens]
3833	AAA05887	752	MMLN1 gene product
3660	XP_001170390	725	PREDICTED: MutL protein homolog 1 isoform 6 [Pan troglodytes]
3648	XP_001170129	723	PREDICTED: MutL protein homolog 1 isoform 1 [Pan troglodytes]
3620	XP_001482268	758	PREDICTED: mutL homolog 1, colon cancer, nonpolyposis type 2 (E. co
3615	XP_534212	757	PREDICTED: similar to MutL protein homolog 1 isoform 1 [Canis famil
3573	XP_001720404	750	PREDICTED: MutL protein homolog 1 [Bos taurus]
3450	BAG28292	758	mismatch repair protein [Mesocricetus auratus]
3444	BAG40671	760	unnamed protein product [Mus musculus]
3443	ED177036	765	cCG26088 [Rattus norvegicus]
3440	AAE45154	760	MutL homolog 1 protein [Mus musculus]
3440	Q9JF21	760	RecName: Full=DNA mismatch repair protein Mlh1; AltName: Full=MutL
3438	NP_081026	760	MutL protein homolog 1 [Mus musculus]
3438	AAH21815	760	MutL homolog 1 (E. coli) [Mus musculus]
3428	ED148277	760	mutL homolog 1 (E. coli), isoform CPA_c [Mus musculus]

NP_000240と類似の配列は、画面から1313生物種の4457アミノ酸配列と相同性を持つことが分かります。

その中で立体構造既知な配列に絞り込みます。そのために、画面上部の“Choose Display Options”をクリックして、表示オプションメニューをまず表示させてください。

続いてkeep onlyのメニューから“3D Structures”を選択し、左下の“BLINK”ボタンをクリックして、立体構造が既知の配列に絞込みを実施してください。

すると立体構造が既知なタンパク質に対するヒットのみが表示されます。

BLINK precomputed BLAST

Home Taxonomy Report Multiple Alignment Blast Help

Pre-computed BLAST results for: [gi|4557757|refNP_000240.1 Mutl. protein homolog 1 \[Homo sapiens\]](#)

Matching gis: [33738032.53932122.155685496.157928134.157928839.730028.741682.1079787.91132884.155119205.453989.119584889.2780](#)

Total (score > 100) : 13 hits in 13 proteins in 2 species

Selected: 13 hits in 13 proteins in 2 species Filter: Min Score: 100 | Keep Only: 3D structures (PDB) |

Other views (Reports): Taxonomy report Multiple Alignment Blast

[Reset all filters](#)

[Choose Display Options](#)

0 Archaea 7 Bacteria 6 Metazoa 0 Fungi 0 Plants 0 Viruses 0 The Others [reset selection](#)

Results: 1 - 13

% hits 756 aa [reset selection](#)

blat	SCORE	ACCESSION	length	Protein Description
◆	524	1B62-R	352	Chain B, Crystal Structure Of An N-Terminal 40kd Fragment Of
◆	524	1B62-A	352	Chain A, Crystal Structure Of An N-Terminal 40kd Fragment Of
◆	524	1B62-A	355	Chain A, Mutl Complexed With Adp
◆	520	1M04-A	333	Chain A, Crystal Structure Of N-Terminal 40kd Mutl Protein
◆	520	1M04-A	333	Chain A, Crystal Structure Of N-Terminal 40kd Mutl (Ln40) Co
◆	520	1B63-A	333	Chain A, Mutl Complexed With Adp
◆	515	1M04-A	333	Chain A, Crystal Structure Of N-Terminal 40kd MutlA100P MUT
◆	420	1EA6-R	364	Chain B, N-Terminal 40kda Fragment Of Hpas2 Complexed With
◆	420	1EA6-A	364	Chain A, N-Terminal 40kda Fragment Of Hpas2 Complexed With
◆	420	1H7U-R	365	Chain B, Hpas2-Atpgs
◆	420	1H7U-A	365	Chain A, Hpas2-Atpgs
◆	418	1H7S-R	365	Chain B, N-Terminal 40kda Fragment Of Human Pas2
◆	418	1H7S-A	365	Chain A, N-Terminal 40kda Fragment Of Human Pas2

Results: 1 - 13

(これ以降のペアワイズアライメントから立体構造を表示させる機能は、現在では提供されていません。)

原文更新日：2005年6月7日 日本語版更新日：2009年1月19日

All Rights Reserved, Copyright(C) 1997-2009 Japan Science and Technology Agency(JST)