# 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\_

### <u>Entrez Geneへの移動</u>

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、

S Entrez cross-datal	base se_ ×			Google 👝 🖻	X
+ → C ☆	http://www.ncbi.nlm.nih.gov/sites	/gque	ery	<ul> <li>D-</li> </ul>	s-
こちらのブックマーク バーにつ	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	t.	🗀 その他のブッ	ウマーク
					-
S NC	BI 🕐 Entre	z, T	he L	ife Sciences Search Engl	ine
HOME SEARCH S	TTE MAP PubMed All Databases	H	luman	Genome GenBank Map Viewe	r
Search ac	ross databases			GO Clea	n He
	Welcome to the Entre	z cro	ss-d	atabase search page	
💓 Pu	ubMed: biomedical literature citations	0	1	Books: online books	
Pt jo	ubMed Central: free, full text urnal articles		t	OMIM: online Mendelian Inheritance in Man	
💓 si	te Search: NCBI web and FTP sites		0	OMIA: online Mendelian Inheritance in Animals	۲
Se Se	ucleotide: Core subset of nucleotide equence records	0		dbGaP: genotype and phenotype	۲
🕂 E5	ST: Expressed Sequence Tag records	0	B	UniGene: gene-oriented clusters of transcript sequences	
E Ge	SS: Genome Survey Sequence cords	9		CDD: conserved protein domain database	۲
😯 Pr	rotein: sequence database		1	3D Domains: domains from Entrez Structure	۲
()) G	enome: whole genome sequences	0	D	UniSTS: markers and mapping data	۲
🔁 st	tructure: three-dimensional acromolecular structures	9	00	PopSet: population study data sets	۲
🔿 Ta	axonomy: organisms in GenBank	0		GEO Profiles: expression and molecular abundance profiles	
42 🔟	NP: single nucleotide polymorphism	0		GEO DataSets: experimental sets of GEO data	0
🖹 G	ene: gene-centered information			Cancer Chromosomes: cytogenetic databases	
te He	omoloGene: eukaryotic homology oups	0		PubChem BioAssay: bioactivity screens of chemical substances	۲
	ENSAT: gene expression atlas of ouse central nervous system		00	PubChem Compound: unique small molecule chemical structures	1
Pr	obe: sequence-specific reagents		0	PubChem Substance: deposited chemical substance records	
89	enome Project: genome project	0	0	Protein Clusters: a collection of	0
199 (B)		_			1

Entrez Geneのトップ画面に移動します。

S Gene Home	×	Coople 👝 🖬 🗙
< → C ☆ h	ttp://www.ncbi.nlm.nih.gov/sites/entrez?db=gene	► B- ₽
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
S NCBI	Entrez Gene	My NCBI R
All Databases Search Gene	PubMed Nucleotide Protein Genome	Structure OMIM PMC Journals Go Clear
ntrez Gene	Limits PreviewIndex History Clipboard	Details
iome Ibout	Entrez Gene is a searchable database of g sequence and/or locate	genes, from <u>RefSeq</u> genomes, and defined by d in the NCBI Map Viewer
AQ Help Sene Handbook	News Sort by chromosome; Search by pre archives	ferred symbol; Property "officially named" News
Statistics Downloads (FTP)	Sample Searches	
lailing Lists	Find genes by	Search text
ene efSeg	free text partial name and multiple species	human muzsular dystrophy transporter[title] AND ("Drozophila
adback	chromosome and symbol	(III(chr] OR 2[chr]) AND adh*[sym]
elo Desk	gene name (symbol)	BRCAI[sym]
orrections	publication (PubMed ID)	11331580[PMID]
bout GeneRIFs	Gene Ontology (GO) terms or identifiers	[cell_adhesion[190] 10030[60]
lated Sites	Genes with variants of clinical significance (under development)	gene and din[filter]
	chromosome and species	Y[CHR] AND human[ORGN]
ntrez Genome enome Projects	Enzyme Commission (EC) numbers	1.9.3.1[EC] more ways to search
enomic Biology	About Entrez Gene	Corrections - Additions - Feedback
omoloGene ap Viewer MIM	Entres Gene Help     Frequently Asked Questions     Entres Gene gene-centered information at NCBL.     Nocket Acids Res. 2005 Jan 1;33:054-8.	Report a <u>new gane</u> Report a <u>new splice variant</u> Hov to add information about function (GeneRIF)     Correct or update a Gene record
robe efSeq niGene	<ul> <li>General help on the <u>Entres</u> search and retrieval system</li> <li><u>NCBL Handbook, Chapter</u> on Entres Gene (download <u>POF</u>)</li> </ul>	Correct or update a reference requence     Report a <u>publication or GeneRIF error</u> Report a <u>search or display problem</u> Benort an FTP problem
esources	<u>Download data</u> via FTP <u>View statistics</u> for Entrez Gene	Make <u>a suggestion</u> for Entrez Gene <u>Send a question</u> to the NCBI Help Desk
IH cDNA	Standard Entrez Gene functions	Gene-related tools
utorial	Limits: limit search using specific oriteria	Analyze genomic sequence and related transcripts

<u>Entrez Geneでのキーワード検索</u>

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。

S NCBI	Entrez Gene	
All Databases	PubMed Nucleatide Protein Conome Structur	OMIM PMC
Search Gene	Ter prion protein	GoClear
Entrez Gene	Limits Preview/Index History Clipboard Details	

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。



# Entrez Gene Reportの読み方

PRNP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。

S Entrez Gene: PRNP prion	p× €	ogk 🗕 🗆 🗶
← → C ☆ http/	//www.ncbi.nlm.nih.gov/sites/entrez?Db=gene&Cmd=ShowDetailView&Term	► D- F-
こちらのブックマーク パーにブックマー	ウ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
	Entrez Gene	My NCBI
Search Gene	v for Go Clea	r
Limits Preview/Inde	x 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 pro	tein [ Homo sapiens ]	ntrez Gen
GeneID: 5621	updated 12-Jan-2009	Table Of Cc
Official Symbol	PRNP provided by HGNC	Summary Genomic rec transcripts. Genomic col Bibliography
Official Full Name	provided by HGNC	HIV-1 prote interactions
Primary source	HGNC:9449	Interactions General cor
See related	Ensembl:ENSG00000171867; HPRD:01453; MIM:176640	information
Gene type	protein coding	General pro information
RefSeq status	REVIEWED	Reference
Organism	Homo sapiens	Related Sec
Lineage	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Additional L Links Order cDNA
Also known as	CJD; GSS; PrP; ASCR; PRIP; PrPc; CD230; prion; MGC26679; PrP27-30; PrP33-35C; PRNP	CCDS Conserved
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- Strauseler disease, Huntington disease-like 1, and kuru. Alternative splicing results in multiple transcript variants encoding the same protein. [provided by RefSeq]	GEO Profiles HomoloGen Map Viewer Nucleotide EST OMIM PubChem C PubChem S Full text in F Probe Protein PubMed PubMed (Of PubMed (Ge 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。



より詳細なゲノム上での位置情報などは緑で囲んだ"See PRNP in Mapviewer"から見ることができます。



機能に関する記述はSummaryの部分に文章で、

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			
<b>RefSeq status</b>	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 機能に関す る記載も	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]			

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。

eOntology		Provided by		
Function	Ev	Evidence		
copper ion binding	IEA			
copper ion binding	TAS	PubMed		
identical protein binding	IPI	PubMed		
microtubule binding	IDA	PubMed		
Process	Ev	idence		
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	Ev	idence		
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	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]

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、クロイトフェルツ・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, transripts, 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。

Additional Links etazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Euarchontoglires; Primates; Explain Links atarrhini; Hominidae; Homo Order cDNA clone CCDS P; ASCR; PRIP; PrPc; CD230; prion; MGC26679; **Conserved Domains** 933-35C; PRNP Genome ncoded by this gene is a membrane **GEO Profiles** HomoloGene hatidylinositol-anchored glycoprotein that tends to Map Viewei o rod-like structures. The encoded protein contains a Nucleotide le region of five tandem octapeptide repeats. This EST on chromosome 20, approximately 20 kbp upstream OMIM ich encodes a biochemically and structurally similar PubChem Compound one encoded by this gene. Mutations in the repeat PubChem Substance as elsewhere in this gene have been associated with Full text in PMC Probe akob disease, fatal familial insomnia, Gerstmann-Protein ease, Huntington disease-like 1, and kuru. Alternative PubMed s in multiple transcript variants encoding the same PubMed (OMIM) ided by RefSeq] 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。

S HomoloGene Result ×	Couje — 🖬 🗙
← → C ☆ http://www.ncbi.nlm.nih.gov/sites/en	entrez?itool=gene_full_report&DbFrom=gene&Omd=Link2 🕨 🕒 🖌
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	きます。 🛅 その他のブックマーク
All Databases Publicad Nucleotide Protein	ene MyNG Bion ein Genome Structure OMIM PMC Journals Bo Go Clear
Climits         Previewindex         History         Clipboard         Det           Your browser version may not work well with NCBI's         Display         HomoloGene         Show         20 S         S           All: 1         Fungl: 0         Mammals: 0         S         S         S	s Web applications. More information <u>here</u> Send to 💌
1: HomoloGene:7904. Gene conserved in Amniota	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.
PRNP, Homo sapiens prion protein	253 aa
PRNP, Pan troglodytes prion protein	3 NP_001009093.1
PRNP, Canis Appus familians prion protein (p27-30) (Creutzfeldt-Jakob disease, Gerstmann-Strausler-Scheinker syndrome, fatal familial insomnia)	5 XP_542906.2
PRNP, Bos taurus prion protein (p27-30) (Creutzfeldt-Jakob disease, Gerstmann-Strausler-Scheinker syndrome, fatal familial insomnia)	9 NP_851358.1
Pmp, Mus musculus prion protein	SI NP_035300.1
Prop. Rattus norvegicus	3 NP_036763.1
PRNP, Gallus gallus prion protein (p27-30) (Creutzfeldt-Jakob disease, Gerstmann-Strausler-Scheinker syndrome, fatal familial insomnia)	Si NP_990796.1 267 aa
Protein Alignments Protein multiple alignment, pairwise similarity scores and evolutionary distances.	Conserved Domains Conserved Domains from CDD found in protein sequences by rpsblast searching.
Show Multiple Alignment	(cl02493)

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。

Display	HomoloGene	Show 20 💌 Ser	nd to 💌
All:	1 Summary Brief		
	He HomoloGene Alignment Scores	rved in Amniota	Download, Links
Gere	s Multiple Alignment		Proteins
Genes during	id ASN.1 the XML FASTA	one another	Proteins used in sequence comparisons and their conserved domain architectures.
D PI	RN Gene Links ior GEO Profile Links	1	MP_000302.1 — 253 aa
D PI	RN Related HomoloGene Links ior Nucleotide Links 		NP_001009093.1

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 🛛 🛪 🕞							AND
→ C ☆ http://www.ncb	inlmnih gov/sites/entr	0.2					Þ B- J
5のブックマーク バーにブックマーク転越加すると	:簡単にページにアクセスできます	la l					🗀 その他のブックマ
S Mary B	HomoloGer	10					My NCBI K
> NCBI	Discover Homo	logs		Help			[Sign In] [Register
All Databases PubMed	Nucleotide Protein	Genome:	Structure	MINO	PMC	Journals	Books
sarch HomoloGene	1000				lear		
Limits Preview/index History	y Clipboard Detail	5					
isplay Alignment Scores	🖌 Show 20 🖌 Se	nd to 💌					
All: 1 Fungi: 0 Mammals: 0	2						
D1. HomoloGana 7904 Gene of	ancested in Ampirta				Downl	and Links	
LI I: Homologene./ Sola. Gene ci	onserved in Adminuta				Domin	ond, canks	
Pairwise Alignment Scores							
Gene	0.2102020	Identi	CY (%)	Subst	itution R	ates '	
Species	Symbol	Protein	DNA.	d	$d_{\rm H}/d_{\rm S}$	d <sub>NR</sub> /d <sub>NC</sub>	
Homo sapiens	PRNP						
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	PRNP						
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	PRNP						
vs. Homo sapiens	PRNP	87.7	87.3	0.139	0.126	0.282	Blast
		87.7	87.3	0.139	0.126	0.253	Blast
vs. Pan troglodytes	PRNP						
vs. Pan troglodytes vs. Bos taurus	PRNP	91.4	87.3	0.139	0.068	0.259	Blast
vs. Pan troglodytes vs. Bos taurus vs. Mus musculus	PRNP PRNP Pmp	91.4 86.9	87.3 82.5	0.139	0.068	0.259	Blast

表よりヒトー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。



### ■ 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。

as elsewhere in this gene. Matadons in the repeat ob disease, fatal familial insomnia, Gerstmann- ase, Huntington disease-like 1, and kuru. Alternative in multiple transcript variants encoding the same led by RefSeq]	Probe Protein PubMed PubMed (OMIM) PubMed (GeneRIF) SNP SNP: Genotype
products 🔹 🌒	Taxonomy

- C H	tiger a woww	nco.nm.	nin gov		a carriece									k
ちのブックマークパーにブッ	クマークを通知	すると効果に	-shop	クセスできます。								0	2 2018/00	195
		Total ge	ene mod	el (contig n	RNA tra	uscript):	12	5	102214				_	
ENERAL	nu	ma	transc	ript pr	otein	mma orientation	a Con	ntig C	ontig Lal	el	List	SNP		
MAN VARIATION	NIM_000	311.3	phus str	and NP_00	03021	forward	NT_0113	87.8 n	ference	<- 0	arrently	shown		
bmit HCH	NM_000	311.3	plus str	and NP_00	0302.1	forward	NW_9273	317.1 C	elera	View	mp on	Greek	(odel	
notate and	NM_000	311.3	plus str	and NP_00	03021	forward	NW_0018	838652 1 H	iuRef	View	mp on	GeneN	fodel	
omit Batch Data	NM 001	080121	plus str	and MP_00	1073590	1 forward	NT_0113	87.8 m	ference	View	r imp on	GeneN	lodel	
h Clinical Impact	NM 001	080121	phus str	and MP 00	1073590	forward	NW 9273	317.1 C	elera	View	mp on	GeneN	fodel	
SUBMISSION	NM 001	080121	plus str	and MP 00	1073590	1 forward	NW 0018	838652 1 H	uRef	View	snp on	GeneN	fodel	
CUMENTATION	ND4 001	080122	ehos ate	and MP 00	1073591	1 Forward	NT 0113	87.8 **	forence	View		Greek	Condet	
ARCH	MM 001	020122	phar etc	and MP 00	1073501	1 Formard	MMI 0273	317.1 0	lera	Viett	ting on	Genald	foodal.	
LATED SITES	304 001	000100	print and	- 4 MD 00	1072501	1 formed	1707 0010	02024511	-D-F	15	inter su	(Card)	for dist	
	200 001	000126	pere an	and the out	1072201	I forward	19.W_0013	0200261	ILLINES .	VIEW	tub cu	CHERRY	1290	
	NM_001	080123	plus str	and MP_00	1073592	I forward	N1_0115	<u>5/8</u> ft	terence	Y153	mp on	Genely	10.05	
	NM_001	080123	plas str	and NP_00	1075592	1 forward	<u>NW_927</u>	517.1 C	elera	View	r mp on	GeneN	lodel	
	NM_001	080123	bjm an	and NP_00	1073592	1 forward	NW_0018	838652 1 H	huRef	Vitw	no qui	GeneN	fodel	
	NM_183	079.2	phus str	and <u>NP_89</u>	8902.1	forward	NT_0113	87.8 r	ference	View	r stop of	GeneN	lodel	
	NM_183	079.2	plus str	and NP_89	8902.1	forward	NW_927	317.1 0	elera	View	r inp-on	GeneN	fodel	
	and the second second second													
	NM_183	e cinically	plus str	and <u>NP_89</u> ted Oin ge	ne region	forward OcSNP Ohas fr	NW_001s	838652 1 B	refresh	View	r snp. on	GeneN	(ode)	
	UM_183	079_2 le cinical) ene mod aRNA tra	phus str e associa el anscript	and <u>NP_89</u> ted Oin ge Contig L treferen	abel C	forward ©cSNP Ohas fr contig mars 011387.8 14M_00	14W_0013 equency O m J 0311.3 NP	838652 1 H double hit protein _000302 1	refresh mrna orie forwa	<u>View</u> ntation rd	transer plus str	igenely ript sup and 24,	Codel count coding	
	NM_183	079.2 le cinical) ene mode aRNA tra	plus str e associa el anscript	and <u>MP_89</u> ted Oin ge Contig L referen	abel C	forward ©cSNP Ohas fr contig enry 011387.8 NM_00	equency O na J 0311 3 <u>NP</u>	838652 1 H double hit protein _000302 1	iuRef refresh mrna orie forwa	View ntation rd	transer plus str <sup>3</sup> Color	igenel ript sup and 24, Legend	count coding	
	NM_183	079.2 le cluical? nRNA tra Contig position	plus str e associa anscript mRNA pos	and <u>NP_89</u> ted Oin ge Contig L fr referen dbSNP rs# cluster id	8902 1 ne region abel C ce <u>NT</u> Hetero- zygosity	forward ©cSNP Ohas fr contig mar 011387.8 NM_000 Yalidation 3D	17W_0013 equency O in J 0311.3 MP_ Clinically Associated	BIRGED 1 H double hit protein 000302 1  Function	taRef refresh forwa dbSNP allele	View ntation rd	transer plus str <sup>5</sup> Color.) Codon pos	ript sup and 24, Legend Amino acid pos	fedel count coding PubMed	
	NM_183	079.2 le clinical ene mode aRNA tra Contig position 4619867	phur str r asrocia el anseript pos 373	and <u>NP_89</u> ted Oin ge Contig L br referen dbSNP rs# cluster id	8902 1 ne region abel C ce <u>NT</u> Hetero- zygosity	forward ©cSNP Ohas fr Contig mar 011387.8 NM_000 Validation 3D	17W_0013 equency O na J 0311.3 NP_ Clinically Associated	S38652   H double hit protein 000302   Function start codes	iuRef refresh mma orie forw dbSNP allele	View ntation rd Fl Protein residue	transer plus str SColorJ Codon pos	ript sup and 24, Legend Amino acid pos	fedel count coding PubMed	
	NM_183	ene mode aRNA tra Contig position 4619867	phur str r associa el anscript pos 373	and <u>NP_89</u> ted Oin ge Contig L b: referen dbSNP rs# cluster id	8902 1 ne region abel C ce NT_ Hetero- zygosity	forward ©cSNP Ohas fr Contig nurr 011387.8 NM_000 Validation 3D	equency O ia J 0311.3 NP Clinically Associated	Protein 000302 1 Function ntart coder	iuRef refresh forwa dbSNP allele	View ntation rd Fi Protein residue	transer plus str Color pos	ript sup and 24, Legend Amino acid pos 1	fedel count coding PubMed	
	NM_183	079.2 le cinical ene mod aRNA tri Contig position 4619867 4619942	phus str e associa el anscript pos 373	and <u>NP_89</u> ted Oin ge Contig L is referen dbSNP rs# cluster id	8902 1 ne region abel C ce NT Ygosity N.D	forward ©cSNP Ohas fr contig nerr 011387.8 NM_000 Validation 3D	NW_0013 equency O na p 0311 3 NP_ Clinically Associated	A contract of the second secon	iuRef refresh mma orie forwa dbSNP allele A C	View ntation rd Fi Protein residue Thr (T) Pro (P)	transer plus str <sup>5</sup> Color) Codon pos	ript sup and 24, Legend Amino acid pos 1 26	fodel count coding PabMed	
	NM_183	contig position 4619967	phus str e associa mRNA pos 373 441	and <u>NP_89</u> ted Oin ge Contig L fr reference dbSNP rs# cluster id	8902 1 ne region abel C ce NT Hetero- rygosity ND	forward ©cSNP Ohas fr contig mar 011387.8 NM_000 Validation 3D	equency O an J 0311.3 MP Clinically Associated	Vertice of the second s	inRef refresh forwa dbSNP allele	View ntation rd Fi Protein residue Dir (T) Pro (P) Has (H)	transco plus str SColor) Codon pos	ript sup and 24, Legend Amino acid pos 1 26	fodel count coding	
	NM_183	0079.2 ie clinical ane modu annA tra Contig position 4619942 4612942	phus str e associa anseript pas 373 441 455	and <u>NP_89</u> ted Oin ge Contig L referen dbSNP rat cluster id m11538752	8902.1 ne region abel C ce NT_ Hetero- xygosity N.D	forward ©cSNP Ohas fr Contig nerr 011387.8 NM_000 Validation 3D	NW_0013 equency O na p 0311.3 NP Clinically Associated	Function reference missense contig reference missense contig	inRef refresh forwa dbSNP allele A C A C	View ntation rd Protein residue The [T] Pro [P] His [H] Fro [P]	transer plus str Codor 1 1 2 2	ript sapa and 24, Legend Amino acid pos 1 26 28 28 28	Count count coding	

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 Dingene region ⊙cSNP Ohas frequency Odouble ht													
\$	zene mod	el	Contig La	abel C	ontig	m	na p	rotein n	uma orio	entation	transcr	ipt snp	count
(contig	mRNA tr	anscript	: referenc	e NT (	11387.8 N	M_00	0311.3 NP	000302.1	forw	ard	plus stra	and 24,	coding
Region Contig mRNA dbSNP rs# Hetero- position pos cluster id zygosity Validation 3D Clinically Associated Function dbSNP Frotein Codon acid pos													
Region	position	pos	dbSNP rs# cluster id	Hetero- zygosity	Validation	3D	Associated	Function	allele	Protein residue	Codon pos	acid pos	PubMed
exon_1	<b>position</b> 4619867	pos 373	dbSNP rs# cluster id	Hetero- zygosity	Validation	3D	Associated	Function start codon	allele	Protein residue	Codon pos	acid pos 1	PubMed
exon_1	position 4619867 4619942	pos 373 448	dbSINP rs# cluster id s11538755	Hetero- zygosity N.D.	Validation	3D	Associated	Function start codon missense	allele A	Protein residue Thr (T)	Codon pos	acid pos 1 26	PubMed

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# cluster id	Hetero- zygosity	<u>Validation</u>	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	с	Pro [P]	1	26	
	4619949	455	rs11538762	N.D.				missense	A	His [H]	2	28	
								contig reference	с	Pro [P]	2	28	

4620464 970	rs28933385 N.D.	missense	A	Lys [K]	1	200
		contig reference	G	Glu [E]	1	200
4620488 994	rs55826236 N.D.	missense	Т	Cys [C]	1	208
		contig reference	с	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.	<b>X</b>	missense	A	Lys [K] 1	200
				88	contig reference	G	Głu [E] 1	200

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

FATAL FAMILIAL INSOMNIA, INCLUDED

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

Studying an unusual cluster of cases of CJD in rural Slovakia, <u>Goldfarb et al. (1990)</u> fou tested cases of 'focal CJD,' in 12 of 40 healthy first-degree relatives, and in 6 of 23 oth extrafocal cases or their relatives had the mutation; nor did any unrelated individuals wi One of the healthy individuals with the E200K mutation was the 75-year-old mother of high incidence of CJD in the Orava and Lucenec regions of Slovakia appeared to be of interpreted this as indicating that the mutation is a necessary, but not sufficient, factor 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, transcrips, 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 <u>reference sequence details</u>	Try our new Sequence
NC_000020.9	[4639234]- (3' 19:000201 - representation 000010886.1 19:00107 Links 2010 19:00107 PROTEIN LINKS 2010.1 10:0011 10

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。

S BLD	NK × (O)				Costr = 0 X
- >	C A http://www.ncb	i.nlm.nih.gov	/sutils/blink	.ogi?al=	18.pid=4506113
55075	ックマーク バーにブックマークを通加すると	と簡単にページに	アクセスできます。		(二) その他のブックマー
BI	LINK		ore	comou	ted BLAST
-	fome Taxonomy Report	Multiple A	lignment	Blast	Help (Size h) (Besister)
re-cor	mputed BLAST results for: gi	14506113(ref)	NP_000.902.1	prion p	protein preproprotein [Homo sapiens]
atchir	ng gis: 3996152,399540,4000	1210;5993964	40075008.10	1053300	110/9226;34335270;15108388;217228859;217228861;217258691;122056
stal (s	score > 100) : 15261 hits in 80	127 proteins	in 798 specie	es	
electe	ed: 15261 hits in 8027 protein	ns in 798 spe	cies Filter:	Min Sci	ore: 100
there	views (Deports): Taxonomy	manad Mu	Ninla Alianma		last
mer v	news (reports). Taxonomy i	report inu	inpie Augnme		10151
eset a	<u>sil filters</u>				
N ON	and Director Outlines				
- Cent	obse Display Options				
			[man] c	Iner	
<ul> <li>Au</li> </ul>	chaea 2436 Bactena 742	Metazoa	850 Fung	2950	Plants 131 Viruses 1466 The Others reset selection
_			Desuiter 1	100 M	Press Last
			Hesuits: 1 -	100 10	ent Page. Last.
hits		reset sel	ection		
	253 aa				
ink.	turniturniturniturnit	SCORE AC	CESSION	Length	Protein Description
	the second s	Conserved	Domain Dat	abase t	uite .
		1435 AA	621693	253	prion protein [Homo sapiens]
		1435 NP	898902	253	prion protein preproprotein [Homo sapiens]
		1435 NP	001073590	253	prion protein preproprotein [Nomo sapiens]
		1435 NP	001073591	253	prion protein preproprotein [Homo sapiens]
		1435 NP	001073592	253	prion protein preproprotein [Homo sapiens]
		1435 BA	673579	253	prion protein [synthetic construct]
		1435 20	4156	253	RecName: Full-Major prion protein: Short-PrP: AltName: Full
		1435 AA	A60182	253	prion protein
		1435 EA	632276	253	prion [Homo sepiens]
		1435 EA	X10449	253	prion protein (p27-30) (Creutzfeld-Jakob disease, Gerstmann
		1435 EA	X10450	253	prion protein (p27-30) (Creutzfeld-Jakob disease, Gerstnam
		1435 CA	875503	253	prion protein (p27-30) (Creutzfeldt-Jakob disease, Gerstman
		1435 AA	580162	253	prion protein [Homo sapiens]
	5 <u></u>	1432 94	0252	253	RecName: Full=Major prion protein; Short=PrF: AltName: Full
		1432 AA	C50089	253	prion protein
		1432 AA	X42953	254	prion protein [synthetic construct]
		1432 41	V38282	254	prion protein (p27-30) (Creutzfeld-Jakob disease, Gerstnaw
		ALUA DO.	¥32089	254	prion protein (synthetic construct)
		1432 44			The second
		1432 AA	M82244	253	prion protein (p27-30) (Creutzfeldt-Jakob disease Gerstman
		1432 AA 1431 AB	102244 102244	253	prion protein (p27-30) (Creutzfeldt-Jakob disease, Gerstman prion protein (p27-30) (Creutzfeldt-Jakob disease, Gerstman
		1432 AA 1431 AB 1431 AB	<u>M82244</u> <u>M85428</u> 632277	253 253 253	prion protein (p27-30) (Creutzfeldt-Jakob disease, Gerstwar prion protein (p27-30) (Creutzfeldt-Jakob disease, Gerstwar prion (Nama samiens)

NP\_000302と類似の配列は、画面から798生物種の8027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。

Choose Display Optic	ons		
Archaea 2436 B	acteria <b>7424</b> Metazoa	850 Fungi 2950 Plants 131	Viruses 1466 The Others reset selection
		Reculter 1 100 Next Page Las	1
Choose Display Opt	ions		
inter nits	O best hits O all hits	s O hide identical 🦦 keep only	Select the database
Minimum Hit Score	100	Maximum Hit Score	3D structures (PDB) SWISSPROT
New Seach By GI		GO Items per page	REFSEQ Complete Genomes
BUNK			

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。

S BLINK × 🕁				Cogle 👝 🖬 🕱
← → C ☆ http://www.nch	oi nim nih go	v/sutils/blir	nk.og/?mode=	result&pid=4506113&page=0&taxon_mode= 🕨 🖸 - 🌾
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	₫+:	🧰 その他のブックマーク
BLINK Home Taxonomy Report	Multiple	pi Alignment	Blast H	BLAST My NCBI 12 A
Pre-computed BLAST results for: g Matching gis: 2396152: 3999540.400 Total (score > 100) : 56 hits in 53 pr Selected: 56 hits in 53 proteins in 1 Other views (Reports): Taxonomy Reset all filters	il4506113(re 0210.599396 oteins in 15 15 species report M	ANP_000302 4:40075008 5 species Filter: Min S lultiple Alignn	2.1 prion prot 10053355.110 Score: 100   K nent Blast	ein preproprotein [Homo sapiens] 79226.34335270,15108368.217228859.217228861:217256691.1; Geep Only: 3D structures (PDB)   
O Archaea O Bacteria 54 M	etazoa 🚺	Fungi 0	Plants 0	Viruses 2 The Others reset selection
% hits	teset se	lection	esuns. 1 - 55	
blink	SCORE A	CCESSION L	ength Prote	in Description
<u> </u>	1203 1	QLX-A	210 Chair 219 Chair	A, Human Prion Protein A Bourine Prion Protein
	771 1	OMD-A	143 Chair	A, Human Prion Protein Fragment 90-230
:	768 1 768 1	F07-A	142 Chair 142 Chair	A, Human Prion Protein Mutant E200k Fragment 90-231 A, Human Prion Protein (Mutant E200k) Fragment 90-23
	<u>698</u> 1 <u>683</u> 2 678 2	810-A 30H-A	142 Chair 148 Chair 138 Chair	A, Solution Nur Structure Of Recombinant Syrian Hamilt, A, Nur Structure Of Rabbit Prion Protein Mutation S. A. Nur Solution Of Rabbit Prion Protein (91-228)
:	677 2 616 1	TPX-A	148 Chair 121 Chair	A, Mar Structure Of Rabbit Prion Protein Mutation I: A, Ovine Recombinant Prp(114-234), Arg Variant In C
:	604 1 604 1	<u>A-680</u>	112 Chair 112 Chair	A, Human Prion Protein Fragment 121-230 A, Human Prion Protein Fragment 121-230
	587 1 584 1 584 1	DWZ-A	108 Chair 112 Chair 112 Chair	A, Crystal Structure Of The Human Prion Protein Reve A, Bowine Prion Protein Fragment 121-230 A, Bowine Prion Protein Fragment 121-230
:	583 1 576 1	HOL-A	112 Chair 113 Chair	A, Muman Prion Protein 121-230 M166cE221C A, Ovine Prion Protein Variant R168
;	<u>569</u> 1 563 1	A-UTX	111 Chair 111 Chair	A, Solution Structure Of The Sheep Prion Protein Wit 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。

S Gene Home	×	Coople 👝 🖬 🗙
< → C ☆ h	ttp://www.ncbi.nlm.nih.gov/sites/entrez?db=gene	► B- ₽
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
S NCBI	Entrez Gene	My NCBI R
All Databases Search Gene	PubMed Nucleotide Protein Genome	Structure OMIM PMC Journals Go Clear
ntrez Gene	Limits PreviewIndex History Clipboard	Details
iome Ibout	Entrez Gene is a searchable database of g sequence and/or locate	genes, from <u>RefSeq</u> genomes, and defined by d in the NCBI Map Viewer
AQ Help Sene Handbook	News Sort by chromosome; Search by pre archives	ferred symbol; Property "officially named" News
Statistics Downloads (FTP)	Sample Searches	
lailing Lists	Find genes by	Search text
ene efSeg	free text partial name and multiple species	human muzsular dystrophy transporter[title] AND ("Drozophila
adback	chromosome and symbol	(III(chr] OR 2[chr]) AND adh*[sym]
elo Desk	gene name (symbol)	BRCAI[sym]
orrections	publication (PubMed ID)	11331580[PMID]
bout GeneRIFs	Gene Ontology (GO) terms or identifiers	[cell_adhesion[190] 10030[60]
lated Sites	Genes with variants of clinical significance (under development)	gene and din[filter]
	chromosome and species	Y[CHR] AND human[ORGN]
ntrez Genome enome Projects	Enzyme Commission (EC) numbers	1.9.3.1[EC] more ways to search
enomic Biology	About Entrez Gene	Corrections - Additions - Feedback
omoloGene ap Viewer MIM	Entres Gene Help     Frequently Asked Questions     Entres Gene gene-centered information at NCBL.     Nocket Acids Res. 2005 Jan 1;33:054-8.	Report a <u>new gane</u> Report a <u>new splice variant</u> Hov to add information about function (GeneRIF)     Correct or update a Gene record
robe efSeq niGene	<ul> <li>General help on the <u>Entres</u> search and retrieval system</li> <li><u>NCBL Handbook, Chapter</u> on Entres Gene (download <u>POF</u>)</li> </ul>	Correct or update a reference requence     Report a <u>publication or GeneRIF error</u> Report a <u>search or display problem</u> Benort an FTP problem
esources	<u>Download data</u> via FTP <u>View statistics</u> for Entrez Gene	Make <u>a suggestion</u> for Entrez Gene <u>Send a question</u> to the NCBI Help Desk
IH cDNA	Standard Entrez Gene functions	Gene-related tools
utorial	Limits: limit search using specific oriteria	Analyze genomic sequence and related transcripts

<u>Entrez Geneでのキーワード検索2</u>

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。

S	NCBI	Ent	rez G	ene				=
All	Databases	PubMed	Nucleotide	Protein	Genome	Structure	OMIM	PMC
Search	Gene	<b>~</b>	for colon cance	r			Go C	lear

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。



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。



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。これらから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。

Genomic regions, transcripts, and products	
Go to reference sequence details	Try our new Sequence
NC_000003.10 S NM_000249.2 - coding region - untranslated region	[37067341] 3* ** ** **

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に見られるミ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 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]

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	¢ 2
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
<b>RefSeq status</b>	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]

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	
<u>MIM: 114030</u>	
Colorectal cancer, hereditary nonpolyposis, type 2	
<u>MIM: 609310</u>	
Mismatch repair cancer syndrome	
<u>MIM: 276300</u>	
Muir-Torre syndrome	
<u>MIM: 158320</u>	
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, transripts, 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。

Genomic regions, transcripts, and products	
Go to <u>reference sequence details</u>	Try our new Sequence Vie
NC_000003.10 [37009983] 5" MM_000249.2 = coding region = untranslated region	[37967361] 3 <sup>4</sup> 11 11 <u>NP_000240+1</u> <u>CCD52663+1</u>

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。

8	HamoloGene Result ×				Google 🖃 🗷
+	→ C A http://www.ncbi.nlm.nih.gov/sites/ent	niz?ltool	=gene_full,report	&DbFrom=gene&	Omd=Link& 🕨 🖸 - 🖋 -
2550	カブックマーク 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	₫ <sub>4</sub>			🎦 その他のブックマーク
3		ne			
	All Databases PubMed Nucleotide Protein	Ger	nome Structu	re OMIM	PMC Journals Bo
Sea	rch HomoloGene 🖌 for			Go Cle	ar .
L	imits Preview/Index History Clipboard Deta	as			
Disp	HomoloGene Show 20 💌 St	end to 📑	8		
A	II: 1 Fungt 0 Mammals: 0 🔀				
	1: HomoloGene:208. Gene conserved in Fungi/Metazo	oa group			Download, Links
Ge	nes	Pro	oteins		
dun	es identified as putative normologs of one another ng the construction of HomoloGene.	don	teins used in sequ nain architectures.	ence companisons	and their conserved
12	MLH1, Homo septens mult homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	2	NP_000240.1 756 aa		
2	MLH1, Pan troglodytea mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	2	XP_001170433.1 756 aa		
(J)	MLH1, Canie kypus familianis mutt. homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	2	XP_534219.2 757 aa		
2	MLH1, Bos taurus mutt. homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	2	NP_001069452.1 233 aa		
2	Mih1, Mus musculus mutL homolog 1 (E. coli)	3	NP_081086.1 760 aa		
2	Mih1, Rattus nonvegicus mutL homolog 1 (E. coli)	2	NP_112315.1 757 aa		3
2	MLH1, Gallus gallus mutt. homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	2	XP_416828.1 757 aa		
2	mih1, Danio renio mutt. homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	2	NP_956953.1 724 aa		
17	Mlh1, Drosophila melanogaster Mlh1	2	NP_477022.1 664 aa		
10	AgaP_AGAP012192, Anopheles gambiae AGAP012192-PA	54	XP_320342.2 671 aa		
12	ENSANGG0000010995, Anopheles gambiae ENSANGP00000013484	21	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 🛛 🛪 💽							Coogle = 0
→ C ☆ http://www.ncbi.nlm	nih.gov/sites/entrez						Þ 🗗 .
らのブックマーク バーにブックマークを追加すると問題	にページにアクセスできます。						こ その他のブック・
S NCBI 💽 🗄	lomoloGene Discover Homologs			Help			My NCEI IX (Sion In) (Rooiston
All Databases PubMed Nuc	Seatide Protein Ge	nome	Structure	OMIM	PMC	Journals	Books
sarch HomoloGene Mor [					ar j		
Limits PreviewIndex History	Clipboard Details						
Alignment Scores Alignment Scores	Show 20 💌 Send to	2					
1. HamploGene:208. Gene conserv	red in Fungi/Metazoa group					Download,	Links
	and a second second						
Pairwise Alignment Scores							
Gene		Identit	¥ (%)	Substit	ution Ra	tes 1	
Species	Symbol	Protein	DNA	d	d <sub>W</sub> /d <sub>S</sub>	d <sub>NR</sub> /d <sub>NC</sub>	
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	Mih1	88.6	87.3	0.140	0.125	0.680	Blast
vs. Rattus norvegicus	Mih1	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	mih1	71.4	66.6	0.442	undef	0.865	Blast
vs. Drosophila melanogaster	Mih1	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	ENSANGG0000010996	46.6	51.9	0.770	undef	0.799	Blast
vs. Schizosaccharomyces pombe	mih1	42.1	48.7	0.865	undef	0.944	Blast
vs. Saccharomyces cerevisiae	MLH1	40.6	48.0	0.686	undef	0.811	Blast
vs. Kluyveromyces lactis	KLLA0009955g	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 grisea	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	MUH1	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 taunus	MIHI	94.8	91.4	0.091	0.061	0 190	Blast

表赤く囲んだ領域よりヒトー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もできます。





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。

Serie Calter. ×	9												
tttp://www	ncbi.nim.ni	h.gov/S	NP/stip.ref.c	gPhoul	d±4202								▶ [
につかった意思	ロオると知識こと	-9070	せんできます。										C 2018
Single	e Nucleo	otide	Polymor	phism	T	P.							
olide Protein Searc	Genome h for SNP (	Structu on NCB1	re PopSet Reference /	Taxono	my OMM	Book	s SNP						
SNP	v for		Liff and all	400011	Go		abalian						
SNPI	nked to G	ene mu	HI (geneiu	-42921	via Contig	Ann	otation						
Send	rs# on all g	ene mod	els to Batch (	Query Do	ownload all	rs# to	file Historic						
Gene	Model (mi	RNA ali	ignment) in	formatio	on from ge	enon	ne sequen	ce	+				
1	otal gene n	nodel (c	ontig mRNA	transcri	ipt):		3	10 T					
IN I	sa tra	nscript	protein	mma o	orientation		Contig	Contig La	bel	List S	SNP		
NM 00	0249-2 plus	strand	NP_000240	1 forwar	d 1	NT_0	22517.17	reference	<- B	urrently	shown		
NM_00	0249.2 plus	strand	NP_000240	1 forwar	d 1	NW.	921651.1	Celera	View	e inp en	GeneM.	odel	
nta pact Dinthe N	de clinically	associate	ed Oin gene Contig Lab	region C	cSNP Ob	as fre	quency Od	ouble hit re	fresh ma orie	ntation t	transcri	ipt sup	count
nta pact Incha N (contig	de clinically gene model mRNA trar	associate ascript):	d Oin gene Contig Lab reference	region C el Co NT_02	0cSNP Oh mtig 2517.17 NJ	nas fre mrr M_00	quency Od na pr 0249.2 <u>NP</u>	ouble hit [ref rotein nu 2002:40.1	fresh ma orie forwa	ntation t rd p	transcri olus stra	ipt sup nd 15, (	count.
nta poct N (contig	de clinically : gene model mRNA trar	associate ascript):	od Oin gene Contig Lab reference	region C el Co NT_02	0.cSNP Oh mtig 22517.12 ND	mr mr M_00	quency Od na pr 0249 2 <u>NP_</u> 1	ouble hit ref rotein mi 200240_1	fresh ma orie forwa	ntation t ed p	transeri plus stra for Legr	ipt smp nd 15, o end	count
n (contig	de clinically pene model mRNA trar Contig position	nseript): mRNA pos	d Oin gene Contig Lab reference dbSNP rs# cluster id	region C el Co NT_02 Hetero- rygosity	DeSNP Oh mtig 2517.12 NJ	mr Mr M_00	quency Od na pi 0249.2 <u>NP_</u> Clinically Associated	ouble hit ref rotein mi 200240_1 ,	fresh forwa dbSNP allele	ntation f ed p F <sup>1</sup> s Co Protein residue	transeri plus stra lor Legr Codon pos	ipt sup nd 15, o md Amino acid pos	count coding PubMed
ta pact Incha N (contig Region exon_1	de clinically gene model mRNA trur Contig position 36975043	mRNA 61	d Oin gene Contig Lab reference dbSNP rs# cluster id	region C el Co <u>NT_02</u> Hetero- zygosity	Validation	as fre mm d_00	quency Od na pr 0249.2 NP_1 Clinically Associated	ouble hit ref notein mu 200240.1 Function	fresh forwa dbSNP allele	ntation f ed p Protein residue	transeri olur stra for Lege Codon pos	ipt sup nd 15, o end Amino acid pos 1	rount coding PubMed
tha piact Incha Region exon_1 exon_1	de clinically : prime model mRNA tran Contig position 36975043 36975043	mRNA pos 61 22	d Oin gene Contig Lab reference dbSNP rs# cluster id rr1800143	region C el Co NT_02 Hetero sygosity N.D.	OcSNP Oh mtig 22517.12 NJ Validation	mr mr d_00	quency O d na pr D2249.2 NP_ Clinically Associated	votein mi votein mi vooged 1 Function start coden synceymoud	fresh forwa forwa dbSNP allele	ntation f ed p Protein residue Glu (E)	transeri olur stra lor Legr Codon pos	and 15, or and Amino acid pos 1 13	count coding PubMed
ta pact (anno (cantig Region exon_1 exon_1	de clinically : pene model mRNA trur Contig position 36975043 36975043	mRNA 61 22	d Oin gene Contig Lab- reference dbSNP rs# cluster id rr1800143	region C el Ce NT_02 Hetero- rygosity N.D.	Validation	mr mr M_00 3D	quency 0 d	Votein nu votein nu votein nu votein nu votein nu Function start codon symceymous config reference	fesh forwa dbSNP allele A G	ntation t ed p Protein Frosidue Glu (E) Glu (E)	transeri olur stra lor Legr pos 3 3	ipt sup nd 15, o end Amino acid pos 1 13 13	rount coding PubMed
ta pact (Inclu (contig Region exon_1 exon_1 exon_4	de clinically : pene model mRNA trur Contig position 36975083 36975083	mRNA pos 61 92	d Oin gene Contig Lab- reference dbSNP rs# cluster id rr1800143	region C el Ce NT_02 Hetero- rygosity N.D.	Validation	mar mar M_00 Xes Xes	quency Od	Function synceymour reference synceymour cong	fesh forwa forwa dbSNP allele A G T	ntation f ed f Protein residue Glu (E) Glu (E) Tyr (Y)	transcri olur stra lor Lege Codon pos 3 3 3	Amino acid pos 1 13 126	rount coding PubMed
N (contig Region exon_1 exon_4	de clinically : pene model mRNA trar Contig position 36975043 36925967	mRNA pos 61 92 438	d Om gene Contig Labr reference dbSNP rs# cluster id rr1820143	region C el Co NT_02 Hetero- rygosity N.D.	Validation	aas fre mar M_000	quency. O d	Function reference reference reference reference	fesh forwa dbSNP allele A G T C	rtation ( ed ) Protein residue Glu (E) Glu (E) Tyr (Y) Tyr (Y)	transeri olus stra lor Lege Codon pos 3 3 3 3 3	ipt sup and 15, 4 and acid pos 1 13 13 126 126	rount coding PubMed
N Region exon_1 exon_4 exon_6	de clinically : prine model mRNA tran Contig position 36975043 36975043 36925967 36920329	mRNA 61 22 534	d Oin gene Contig Labr reference dbSNP rs# cluster id rc1800143 rc34665507	region C el Ce NT_02 Heterro xygosity N.D. N.D.	Validation	aas free mari 4_000 1_3D Yes Yes	quency O d	Puble hit (ref otein mi otoco240.1  Function start codon synceymous config reference synceymous synceymous synceymous synceymous synceymous synceymous	tesh ] forwa dbSNP allele A G T C C	ntation f rd g Protein residue Glu (Ε) Glu (Ε) Tyr (Υ) Tyr (Υ) Am (N)	transeri lor Lege Codon pos 3 3 3 3 3 3 3 3 3	ipt sup and 15, c and acid pos 1 13 12 126 126 126	count coding
An inclusion in the second sec	de clinically i pene model mRNA tran Contig position 36975043 36925967 36920329	mRNA pos 61 22 534	d Oin gene Contig Labo reference dbSNP rs# cluster id rs1800143 rs34665507 rs4547256	region C el Co NT_02 Hetero- xygosity N.D. N.D.	Validation	as fre mr (d_00 ) 3D Yes Yes	quency O d	Function synceptional synceptional synceptional contag reference synceptional contag reference	tesh forwa dbSNPP allele A G T C	ntation 1 erd 1 Protein Gu (E) Gu (E) Tyr (Y) Tyr (Y) Tyr (Y) Am [N]	transeri lor Lege Codon pos 3 3 3 3 3 3 3 3 3 3 3	ipt sup md 15, i md acid pos 1 13 13 126 126 158	PobMed
N (contig Ragion exon_1 exon_4 exon_6 exon_7	de clinically : pene model mRNA trur Contig position 36975061 36925967 36920322	mRNA pos 61 22 534 534	d Om gene Contig Labo reference dbSNP rs# chaster id rs1820143 rs34665597 rs4647256	region C el Ce NT_02 Hetero- sygosity ND 0.036	Validation	as fre mr 4_00 Yes Yes I Yes	quency O d	Puble hat ref rotein na rotein na rotein na rotein na Function rfart codon synceymous conng reference	tesh forwa dbSNPP allele A G T C C T	ntation 1 erd 1 Protein Gu (E) Gu (E) Tyr (Y) Tyr (Y) Tyr (Y) Am [N] Ser (S1	transeri lor Lege Codon pos 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	ipt smp md 15, o end Aminu acid pos 1 13 126 126 126 158 158 158	count coding

error 7	36993321	612	++35225190	ND	10	Ver	-	Ŧ	Sec (S)	2	184	1
sava	SACCESSION A	N,A.M	TAILORNICALLY			-	config		0101	2	104	
	22002202	240		AT IS			reference	~	aer (a)		192	
	200922223	0.10	HILLSRINGS	24.10		0.0	contra		Theft		180	
							reference	с	His [H]	1	1106	
	36993352	643	rr35847123	N.D.			frame shift			1	195	
							reference	A	Lys [K]	1	195	
exon_9	36096039	850	1535338630	ND			missense	G	Gly [G]	1	264	
							comg	с	Arg (R)	1	264	
eron 12	37007224	1251	er35164771	ND			etenorementer	T	Leu(L)	3	397	
cave_re		Abach	LESS CONTINUE				config	a	Levilla	2	307	
							reference	-	Ten (r)	2	221	
	37007447	1914	1154280587	ND.	-		config		Ser [S]	1	452	
							reference	Å	194 (T)		452	
exon_15	37703-798443	1790	1150185999	ND.			missense	T	Leu [L]	2	577	
							reference	c	Set (S)	2	512	
exon 17	37030023	1968	rs1800145	N.D.		TT T	SADORADORS	A	LeufLl	3	636	
							config	G	LeufLi	3	636	
_							reference	ñ.,		1	-	
exon_18	37030474	2122	1041542214	N.D.	-		mattenae	A	Lyr [K]	1	689	
						- J. J.	reference	с	Gh [Q]	1	689	
exon_19	37031995	2178	m1500147	ND.			synonymous	т	Gty [G]	3	706	
							config	с	Gly [G]	3	706	
	37032031	2214	rs1800148	N.D.			synonymous	т	Har [H]	3	718	
							config	с	Has [H]	3	718	
	37032062	2245	er1800149	ND			massense	G	VallVI	1	729	
			and the second states		-		contig		Terr (T.)		100	
							reference	ň.,	new (m)		Card I	

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。



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.	3	t I⊉	missense	С	His [H]	1	132
		9	t Ie	contig reference	G	Asp [D]	1	132



Using a novel high density oligonucleotide array (HNPCC Chip) to look for variants in the MLH1, MSH2 (<u>609309</u>), and M (<u>600678</u>) genes in Israeli probands with familial colorectal cancer (CRC; <u>114500</u>) unstratified with respect to the micros instability phenotype, <u>Lipkin et al. (2004</u>) 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 attenua not eliminated, by the MLH1 415G-C mutation. Q

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, transcrips, 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。



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。

	×					
· > C f	http://www.ncb	i.nlm.nih.gov/	sutils/blink.cg	Pmode=result&pid=4557757&page=0&taxon,mode=al 🕨 🗋 • 🧳		
うのブックマーク 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。	この他のブックマ		
BLINK			preco	mouted BLAST		
Home T	axonomy Report	Multiple Ali	gnment Bl	ast Help Bion M Register		
e-computed B	LAST results for: a	145577571cefN	P 000240.1 M	utl. protein homolog 1 [Homo sapiens]		
atching gis: 33	738032 53932122 15	55685496 1579	28134:1579288	39.730028.741682.1079787.91132884.155119205.463989.119584889.278		
atal (score > 10	10) : 13 hits in 13 pr	oteins in 2 spe	cies			
alacted: 13 hits	in 13 proteins in 3	enacios Eilt	ar Min Score	100   Keen Only 3D structures (PDB)		
necied: 13 mis	s in 15 proteins in 2	species rite	er: min Score	: too ( Neep Only: 50 structures (PDB) (		
ther views (Rep	ports): Taxonomy	report Multi	ple Alignment	Blast		
eset all filters						
Change Dien	lau Ontione					
FUIDORE FUIDE	MATPHINUS					
Archaea 7	Bacteria E Ma	tazon 🗖 Eur	noi 🔲 Plant	Vinces I The Others reset selection		
0 Archaea 7	Bacteria 6 Me	tazoa 🚺 Eur	ngi 🚺 Plant	I Viruses The Others reset selection		
0 Archaea 7	Bacteria 6 Me	tazoa 🚺 Eur	ngi 🚺 Plant Resu	a Viruses The Others reset selection		
0 Archaea 7	Bacteria 6 Me	tazoa 间 Eur	ngi 🚺 Elantı Resu	a Viruses The Others reset selection		
Archaea 7	Bacteria 6 Me	tazoa 间 Eu reset sele	ngi 🚺 Elantı Resu ction	a Vinuses The Others reset selection ts: 1 - 13		
0 Archaea 7	Bacteria 6 Me	tazoa D Eur	Resu Resu	ts: 1 - 13		
Archaea 7	Bacteria 6 Me	tazoa 💽 Eu reset sele: SCORE ACCI Conserved I	ngi D Elanti Resu CLION ESSION Lengt	te: 1 - 13		
Archaea 7	Bacteria 6 Me	tazoa D Eur reset selev SCORE ACCI Conserved J 524 1863	ngi Plant Resu ction ESSION Lengt Domain Datab N-B 352	ts: 1 - 13		
Archaea 7	Bacteria 6 Me	tazoa  Eur reset selev SCORE ACCI Conserved 524 IBHO 524 IBHO	ngi O Plant Resu stion ESSION Lengt Domain Datah N-B 352 N-A 352	ts 1 - 13 <b>Exotein Description</b> has his Chain B, Crystal Structure Of An N-Terminal 40kd Fragment Chain B, Crystal Structure Of An N-Terminal 40kd Fragment		
Archaea 7	Bacteria 6 Me	tazon 0 Eur reset sele: SCORE ACCI Conserved 1 524 150 524 1565 524 1565	Resu Resu ESSION Lengt Denain Patah N=B 352 V=A 352 Z=A 355	te 1 - 13 The Others reset selection the 1 - 13 A Protein Description are hits Chain B, Crystal Structure Of An N-Terminal 40kd Fragment Chain A, Muti Complexed With Adm		
0 Archaea 7	Bacteria 6 Me	tazoa D Fut reset sele: SCORE ACCI Conserved J 524 JB60 524 JB60 524 JB60 524 JB60 524 JB60 524 JB60 524 JB60	Resu Resu Ction ESSION Lengt Denote Datab U-B 352 U-A 352 Z-A 353	te 1 • 13 The Others reset selection the 1 • 13 Protein Rescription are hits Chain B, Crystal Structure Of An N-Terminal 40kd Fragment Chain A, Crystal Structure Of An N-Terminal 40kd Fragment Chain A, Crystal Structure Of N-Terminal 40kd Fragment Chain A, Crystal Structure Of N-Terminal 40kd Fragment Chain A, Crystal Structure Of N-Terminal 40kd Mutl Protein		
) Archaea 7 hits 756 aa iak	Bacteria 6 Me	tazoa D Fut reset selet SCORE ACCI Conserved J 524 1885 524 1885 524 1885 520 1885 520 1885	Resu Ction Ction Ction Conain Patah N=B 352 N=A 352 C=A 355 H=A 333	te: 1 - 13		
Archaea 7	Bacteria 6 Me	tazoa le Eur reset sele: SCORE ACCI Conserved J S24 JB55 S24 JB55 S24 JB55 S20 JM6 S20 JM6 S20 JB6	ngi O Plant Resu ESSION Lengt Conain Patah W-B 352 2-A 355 1-A 333 I-A 333	Vinuses      The Others reset selection ts: 1 - 13      Exotein Percription     Arc Arc Arc Arc Arc Arc Arc Arc Arc		
0 Archaea 7 htts 756 ea	Bacteria 6 Me	tazoa  Europeano	Image         Plant:           Result           CELION           ESSION Lengt           Demain Detain	Vinuses The Others reset selection tts: 1 - 13      Protein Description     Anno. Crystal Structure Of An N-Terminal 40kd Fragment     Chain A, Crystal Structure Of An N-Terminal 40kd Fragment     Chain A, Crystal Structure Of N-Terminal 40kd Mutl Protein     Chain A, Crystal Structure Of N-Terminal 40kd Mutl Protein     Chain A, Crystal Structure Of N-Terminal 40kd Mutl (Ln40)     Chain A, Mutl Complexed With Adpnp     Chain A, Crystal Structure Of N-Terminal 40kd Mutl (Ln40)     Chain A, Crystal Structure Of N-Terminal 40kd Mutl (Ln40)     Chain A, Crystal Structure Of N-Terminal 40kd MutlA100P MU		
Archaea     Archaea     Archaea     Totae     Totae	Bacteria 6 Me	tazoa rezet solo SCORE ACCI Conserved S24 IBS S24 IBS S24 IBS S20 IMB S20 IMB S20 IBS S15 IMB S15 IMB	Image         Plant:           Result           CESSION Lengt:           CESSION LENGT: <td c<="" colspan="2" td=""><td><ul> <li>Yinuses The Others reset selection</li> <li>The Others reset selection</li> <li>The Others reset selection</li> <li>The Exotein Description</li> <li>Chain B, Crystal Structure Of An N-Terminal 40kd Fragment</li> <li>Chain A, Crystal Structure Of An N-Terminal 40kd Fragment</li> <li>Chain A, Crystal Structure of N-Terminal 40kd Mutl Protein</li> <li>Chain A, Crystal Structure of N-Terminal 40kd Mutl Protein</li> <li>Chain A, Crystal Structure of N-Terminal 40kd Mutl (In40)</li> <li>Chain A, Crystal Structure of N-Terminal 40kd Mutl (In40)</li> <li>Chain A, Crystal Structure of N-Terminal 40kd MutlA100P MU</li> <li>Chain B, M-Terminal 40kd Fragment of Mnpms2 Complexed Wat</li> </ul></td></td>	<td><ul> <li>Yinuses The Others reset selection</li> <li>The Others reset selection</li> <li>The Others reset selection</li> <li>The Exotein Description</li> <li>Chain B, Crystal Structure Of An N-Terminal 40kd Fragment</li> <li>Chain A, Crystal Structure Of An N-Terminal 40kd Fragment</li> <li>Chain A, Crystal Structure of N-Terminal 40kd Mutl Protein</li> <li>Chain A, Crystal Structure of N-Terminal 40kd Mutl Protein</li> <li>Chain A, Crystal Structure of N-Terminal 40kd Mutl (In40)</li> <li>Chain A, Crystal Structure of N-Terminal 40kd Mutl (In40)</li> <li>Chain A, Crystal Structure of N-Terminal 40kd MutlA100P MU</li> <li>Chain B, M-Terminal 40kd Fragment of Mnpms2 Complexed Wat</li> </ul></td>		<ul> <li>Yinuses The Others reset selection</li> <li>The Others reset selection</li> <li>The Others reset selection</li> <li>The Exotein Description</li> <li>Chain B, Crystal Structure Of An N-Terminal 40kd Fragment</li> <li>Chain A, Crystal Structure Of An N-Terminal 40kd Fragment</li> <li>Chain A, Crystal Structure of N-Terminal 40kd Mutl Protein</li> <li>Chain A, Crystal Structure of N-Terminal 40kd Mutl Protein</li> <li>Chain A, Crystal Structure of N-Terminal 40kd Mutl (In40)</li> <li>Chain A, Crystal Structure of N-Terminal 40kd Mutl (In40)</li> <li>Chain A, Crystal Structure of N-Terminal 40kd MutlA100P MU</li> <li>Chain B, M-Terminal 40kd Fragment of Mnpms2 Complexed Wat</li> </ul>
) Archaea 7 hits 756 aa ak	Bacteria 6 Me	tazoa le Fut reset sele: SCORE ACCI Conserved J 524 JBS3 524 JBS3 524 JBS3 524 JBS3 520 JBS 520 JBS 515 JMR 420 JEA	Image         Plant:           Result           CESSION Lengt:           CEA 352           CEA 353           CEA 333           CEA 333           CEA 333           CEA 364	<ul> <li>Yinuses The Others reset selection</li> <li>Frotein Rescription</li> <li>Chain B, Crystal Structure Of An N-Terminal 40kd Fragment Chain A, Crystal Structure Of An N-Terminal 40kd Fragment Chain A, Crystal Structure Of N-Terminal 40kd Mutl Protein Chain A, Crystal Structure Of N-Terminal 40kd Mutl Protein Chain A, Crystal Structure Of N-Terminal 40kd Mutl (Ln40) Chain A, Crystal Structure Of N-Terminal 40kd Mutl (Ln40) Chain A, Mutl Complexed With Adpup Chain A, Crystal Structure Of N-Terminal 40kd MutlA100P MU Chain B, N-Terminal 40kda Fragment Of Mnpms2 Complexed Wit Chain A, M-Terminal 40kda Fragment Of Mnpms2 Complexed Wit</li> </ul>		
Archaea     Archaea     Archaea     Total     Archaea     Total     Tot	Bacteria 6 Me	tazoa le Eur reset sele: SCORE ACCI Conserved J S24 1893 S24 1893 S24 1893 S24 1893 S24 1893 S24 1893 S20 1894 S20 1844 420 1844 420 1847	Image         Plant           Result         Result           CELION         Second Patch           Consain Patch         352           VA         352           Z=A         355           H=A         333           I=A         333           I=A         333           I=A         364           S=A         364           S=A         364           I=B         365	<ul> <li>Vinuses The Others reset selection</li> <li>Exotein Description</li> <li>Arystal Structure Of An N-Terminal 40kd Fragment Chain A, Crystal Structure Of An N-Terminal 40kd Fragment Chain A, Crystal Structure Of N-Terminal 40kd Fragment Chain A, Mutl Complexed With Adp Chain A, Crystal Structure Of N-Terminal 40kd Mutl Protein Chain A, Crystal Structure Of N-Terminal 40kd Mutl (Ln40) Chain A, Mutl Complexed With Adpm Chain A, Crystal Structure Of N-Terminal 40kd MutlA100P MU Chain A, Crystal Structure Of N-Terminal 40kd MutlA100P MU Chain B, M-Terminal 40kda Fragment Of Nhpms2 Complexed Wit Chain A, Bpms2-Atpgs</li> </ul>		
0 Archaea 7 756 ea ink	Bacteria 6 Me	tazoa	Image         Plant:           Result           CELION           CELION <td><ul> <li>Vinuses The Others reset selection</li> <li>International selection</li> <li>Protein Description</li> <li>An Protein Description</li> <li>Chain B, Crystal Structure Of An N-Terminal 40kd Fragment</li> <li>Chain A, Crystal Structure Of An N-Terminal 40kd Fragment</li> <li>Chain A, Ruti Complexed With Adp</li> <li>Chain A, Crystal Structure Of N-Terminal 40kd Mutl Protein</li> <li>Chain A, Crystal Structure Of N-Terminal 40kd Mutl (Ln40)</li> <li>Chain A, Crystal Structure Of N-Terminal 40kd Mutl (Ln40)</li> <li>Chain A, Mutl Complexed With Adpnp</li> <li>Chain A, Crystal Structure Of N-Terminal 40kd MutlAl00P MU</li> <li>Chain B, W-Terminal 40kda Fragment Of Mnpms2 Complexed Witt</li> <li>Chain B, Hpms2-Atpps</li> </ul></td>	<ul> <li>Vinuses The Others reset selection</li> <li>International selection</li> <li>Protein Description</li> <li>An Protein Description</li> <li>Chain B, Crystal Structure Of An N-Terminal 40kd Fragment</li> <li>Chain A, Crystal Structure Of An N-Terminal 40kd Fragment</li> <li>Chain A, Ruti Complexed With Adp</li> <li>Chain A, Crystal Structure Of N-Terminal 40kd Mutl Protein</li> <li>Chain A, Crystal Structure Of N-Terminal 40kd Mutl (Ln40)</li> <li>Chain A, Crystal Structure Of N-Terminal 40kd Mutl (Ln40)</li> <li>Chain A, Mutl Complexed With Adpnp</li> <li>Chain A, Crystal Structure Of N-Terminal 40kd MutlAl00P MU</li> <li>Chain B, W-Terminal 40kda Fragment Of Mnpms2 Complexed Witt</li> <li>Chain B, Hpms2-Atpps</li> </ul>		
Archaea     7	Bacteria 6 Me	tazoa	Image         Plant:           Result           Result           CESSION Lengt:           CESSION<	<ul> <li>Ymuses The Others reset selection</li> <li>The Others reset selection</li> <li>the 1 - 13</li> <li>Protein Rescription</li> <li>Chain B, Crystal Structure Of An N-Terminal 40kd Fragment</li> <li>Chain A, Crystal Structure Of An N-Terminal 40kd Fragment</li> <li>Chain A, Crystal Structure Of N-Terminal 40kd Mutl Protein</li> <li>Chain A, Crystal Structure Of N-Terminal 40kd Mutl (In40)</li> <li>Chain A, Crystal Structure Of N-Terminal 40kd Mutl (In40)</li> <li>Chain A, Crystal Structure Of N-Terminal 40kd Mutl (In40)</li> <li>Chain A, Crystal Structure Of N-Terminal 40kd MutlA100P MU</li> <li>Chain A, N-Terminal 40kda Fragment Of Mnpms2 Complexed Wit</li> <li>Chain B, Mpms2-Atpgs</li> <li>Chain B, M-Terminal 40kda Fragment Of Human Pms2</li> </ul>		
Archaea     7     Archaea     7     Archaea     7     7     Archaea     7     A	Bacteria 6 Me	tazca	Image         Plant:           Result           CESSION Lengt:	<ul> <li>Yinuses The Others reset selection</li> <li>Frotein Description</li> <li>Protein Description</li> <li>Chain B, Crystal Structure Of An N-Terminal 40kd Fragment Chain A, Crystal Structure Of An N-Terminal 40kd Fragment Chain A, Crystal Structure Of N-Terminal 40kd Mutl Protein Chain A, Crystal Structure Of N-Terminal 40kd Mutl Protein Chain A, Crystal Structure Of N-Terminal 40kd Mutl (Ln40) Chain A, Crystal Structure Of N-Terminal 40kd Mutl (Ln40) Chain A, Crystal Structure Of N-Terminal 40kd MutlAl00P MU Chain A, N-Terminal 40kda Fragment Of Mnpms2 Complexed Witt Chain B, M-Terminal 40kda Fragment Of Mnpms2 Complexed Witt Chain B, Hpms2-Atpgs Chain B, N-Terminal 40kda Fragment Of Human Pms2 Chain B, N-Terminal 40kda Fragment Of Human Pms2</li> </ul>		

(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)

