# Map Viewer クイックスタート

Map Viewer(ヒトゲノム版)の使い方を遺伝子の探索を通して学びます。 なお、ゲノムデータのバージョンによって図の結果と必ずしも一致しないことがあります。

## 課題1

例として、グリア細胞由来の神経栄養因子遺伝子GDNFを使う事にしましょう。 GDNFは変異を起こすと、ヒルシュ スプルング病の原因となると推測されている遺伝子です。

## ■ 課題1-1\_

ヒトゲノムに関するMap Viewerを立ち上げて"GDNF"遺伝子を検索してみて下さい。 Map Viewerをクリック、Human Map Viewerにアクセスしてください。続いて検索語としてGDNFを入力して下さい。

#### ■ 課題1-2\_

この遺伝子のある染色体を確認して下さい。複数のヒット候補の中から、 "Assembly"の列にreference、 "Map element"の列にGDNFと書かれたヒットの、 "Maps"の列にあるGenes\_seqと書かれたリンクをクリックして下さ い。 表示されているマップはどのようなものでしょうか? Maps & Optionsをクリックして表示される画面を用い て、座標の表示を選択して下さい。 この遺伝子は染色体のどこに位置していますか?その位置を塩基数で答えて ください。 染色体上で、遺伝子はどちらを向いていますか? "dl"と書かれている箇所をクリックして、遺伝子配 列をダウンロードしてください。 (必要であれば、ストランドの向きを変更してください。) プロモータ因子を 探索するために、ダウンロードする範囲を上流5000塩基に調節して、 遺伝子配列の上流部分にある5000塩基も合 わせてダウンロードして下さい。

#### ■ 課題1-3\_

Maps & Optionsをクリックして表示される画面を用いて、この領域に対する、クローン情報、 コンポーネント情報、コンティグ地図を追加して下さい。 この領域を含むゲノム配列に対応するコンティグとGenBankのアクセッションナンバーを確認してください。 解読が完了した(訳注:ドラフトでない)配列はありますか?この領域にマップされたクローンはありますか? もしあるのであれば、どうやってそのクローンを注文する事が出来るでしょうか?

#### ■ 課題1-4

Genes\_seqを除いた全てのマップを削除し、Ab initio(モデル)とTranscript(RNA)マップを追加して下さい。 遺伝子の予測結果が現在表示されている遺伝子のアノテーションと一致していますか? この遺伝子に対して、一 体いくつの選択的スプライスによる転写産物がアノテーションされているでしょうか? 現在表示されているデー タについて"Table View"をクリックして表形式にしてみて下さい。

### ■ 課題1-5

Model Maker (mm)を使って、選択的スプライス転写産物候補やその転写産物から翻訳されたアミノ酸配列を取得して下さい。 それらに対してBLASTを使って類似したタンパク質を検索して下さい。

### ■ 課題1-6\_

Map Viewerの表示画面に戻って下さい。Genes\_seq以外の全てのマップを削除し、マウス、チンパンジー、ラットのGene maps を追加して下さい。 この3つの生物種の遺伝子構造は似ていますか?

## ■ 課題1-7\_

ヒトのGeneMapを除いた全てのマップを削除し、phenotype mapを追加して下さい。 GDNF遺伝子が関連する病気を 確認して下さい。 その病気に関してリンクされているOMIMデータから、その病気についてさらなる情報を取得し て下さい。

## 課題2

今度は、ヒトのPRNP遺伝子を使って上述した手順を繰り返して下さい。

Map Viewerで"PRNP"を検索してみて下さい。 <u>Map Viewer</u> をクリックしてHuman Map Viewerにアクセスして、検索 語としてPRNPを入力して下さい。

この遺伝子がある染色体を確認して下さい。そして、referenceアセンブリ内の要素であるPRNPからリンクが貼ら

れているGenes seq mapと書かれた箇所をクリックして下さい。

表示されているマップはどのようなものですか? Maps & Optionsをクリックして表示される画面を用いて、座標の表示を選択して下さい。この遺伝子は染色体のどこに位置していますか?塩基数で答えてください。染色体上で、遺伝子はどちらを向いていますか?

"dl"と書かれている箇所をクリックして、遺伝子配列をダウンロードしてください。(必要であれば、ストランドの向きを変更してください。)プロモータ因子を探索するために、ダウンロードする範囲を上流5000塩基に調節して、遺伝子配列の上流部分にある5000塩基も合わせてダウンロードして下さい。

Maps & Optionsをクリックして表示される画面を用いて、この領域に対するクローン情報、コンポーネント情報、 コンティグ地図を追加して下さい。この領域を含むゲノム配列に対応するコンティグとGenBankのアクセッション ナンバーを確認してください。解読が完了した(訳注:ドラフトでない)配列はありますか?この領域にマップさ れたクローンはありますか?もしあるのであれば、どうやってそのクローンを注文する事が出来るでしょうか?

Genes\_seqを除いた全てのマップを削除し、Ab initio(モデル)とTranscript(RNA)マップを追加して下さい。遺伝子の予測結果が現在表示されている遺伝子のアノテーションと一致していますか?この遺伝子に対して、一体いくつの選択的スプライスによる転写産物がアノテーションされているでしょうか?現在表示されているデータを"Table View"をクリックして表形式にしてみて下さい。

ModelMaker (mm)を使って、選択的スプライス転写産物候補やその転写産物から翻訳されたアミノ酸配列を取得して下さい。それらに対してBLASTを使って類似したタンパク質を検索して下さい。

MapViewerの表示画面に戻って下さい。Genes\_seq以外の全てのマップを削除し、マウス、チンパンジー、ラットの Gene maps を追加して下さい。この3つの生物種の遺伝子構造は似ていますか?

ヒトのGeneMapを除いた全てのマップを削除し、phenotype mapを追加して下さい。PRNP遺伝子が関連する病気を確認して下さい。その病気に関してリンクされているOMIMデータから、その病気についてさらなる情報を取得して下さい。

# 解答・解説

■ 解答1-1\_

<u>ヒトゲノムに関するMap Viewerの立ち上げ</u>

ヒトゲノムに関するMap Viewerのトップページを開きましょう。

下に示した<u>Map Viewerのトップページ</u>からHomo Sapiensと書かれた横にあるゲノムのバージョンをクリックして (現在はBuild36.3とBuild35.1の三種類のゲノムのバージョンが公開されているので好きな方をクリックしてくだ さい) ヒトゲノムに関するMap Viewerのトップページに移動するか、

← → C ☆ http://www.ncbi.nlm.nih	www.mapview/ NORIフッゴピュ	പതനം പ്രസംസം 🖉			• 0	· \$
SNCBI Home GenBank	BLAST	-701924-2				
Map Viewer Home					Hel	p
17	e Map Viewer provides a wide variety of	genome mapping and sequencing data.	More		(W)A	a
Casesh	v Vertebrates				(15	6
Search *	▼ Mammals				(13	0
Search: Select Group or Organism 💌	* Primates				(3	0
for.	Scientific name	Common name	Build	Tools	1-	
ヒトゲ	L Homo sapiens	human 18 States	Build 36.3	989	0	
	-7	N=292<	Build 35.1	9.8	-	
Tools Legend v	Macaca mulatta	rhesus macaque	Build 1.1	9.8	0	
(b) Search or Browse the Genome	Pan troglodytes	chimpanzee	Build 2.1	9.8	0	
B BLAST	* Rodents				(7	5
G Clone Finder	Scientific name	Common name	Build	Tools		
@ Genome Resources page	Mus musculus	laboratory mouse	Build 37.1	989	0	
Name	1		Build 36.1	9.8	-	
News +	Rettus norvegicus	rat	RGSC v3.4	9.8	0	
Annotation update released for Mar24	Monotremes				(1	0
human genome build 36	Marsupials				(T	5
MCBI Build 36.30 more 王初程	Other Mammals				(6	6
Show all	Other Vertebrates				(2	9
Related Resources v	Invertebrates				(10	0
	▶ Protozoa (8)				(18	0
NCBI Home	► Plants (3)				(45	0
NCBI Site map	▼ Fungi 🔍 🖲				(17	9
Genome Biology	Scientific name	Common name	Build	Tools		
Taxonomy Entret (Global Quero	Aspergillus clavatus		Build 1.1	9.8		
BLAST	Aspengillus fumigatus		Build 2.1	9.8	0	
Map Viewer FTP	Aspergillus niger		Build 1.1	88	0	
Small Genomes V	Candida glabrata		Build 1.1	9.8		
Since Generatives	Cryptococcus neoformans		Build 2.1	9.8		
Bacteria 🛞	Debaryomyces hansenii		Build 1.1	9.8		
Organelles (B)	Encephalitozoon cuniculi		Build 1.1	9.8		
Viruses (8)	Eremothecium gossypii		Build 3.1	9.8		
	Gibberella zeae		Build 1.2	9.8		
	Klugveromyces Jacbs		Build 1.1	0.0		
	Maghaporthe grisea	rice blast fungus	Build 3.1	0.0		
	Pickie alle Na		Build 2.1	0.0		
	Pacing oppose	holosie usoel	Duild 2.1			
	Sal Charomyces cerevisiae Schiznegor honory og provins	bakers yeast fection yeard	Build 1 4	0.0		
	Unblace martia	instanti yeast	Duild 1.1	0.0		
	Vannuis linnkhira		Build 1 1	0.0		
	<ul> <li>Lawrana shrubara</li> </ul>		00101.1			

あるいは直接http://www.ncbi.nlm.nih.gov/projects/mapview/map\_search.cgi?taxid=9606とブラウザに入力する ことで、ヒトゲノムMap Viewerのトップページに移動できます。



# <u>Map Viewer内を遺伝子名で検索する</u>

上図で青く示したテキストボックスに検索したい遺伝子である「GDNF」と入力し、Findボタンをクリックします。



# 206個ものヒットがあったことが確認できます。

結果は、左から染色体番号、アセンブル名(現在のヒトゲノム情報は公的機関が出しているヒトゲノム情報 (reference)とセレラ社が公開している情報(Celera)、JCVIが公開している情報(HuRef)、それ以外にも6番、7番 染色体など部分的にハプロタイプの違うゲノム配列などが提供されています)、検索キーワードのマッチ情報、マ ップ内の要素名、タイプ(表現型、マーカー、遺伝子などを示します)、マップ名の順に示されています。上記結 果では、5番・6番・8番・10番・20番染色体および位置関係のわからない配列断片上に「GDNF」というキーワード を含むヒットが認められています。ヒットのおよその位置は上図中の青丸で囲まれた部分にヒットの数と共に示さ れています。

提供元が違うとゲノム上の位置関係が多少違うものの本質的には同じものなので、情報を整理するためここでは公的研究機関から出されたreferenceと書かれている情報のみに注目してください。このフィルタは先ほどの検索時にFindボタンの左にある"assembly"からreferenceを選択して検索することでも実現できます。

PubMed	Nucleotide	Protein	Genome	Gene	Structure	PopSet	Texonom	y Help
Search for	GDNF	on chro	mosome(s)		assembly	All	Find	Advanced Search
Inconsection and a consection	Destroyed in					All		
			77477	1110 253	7 7	reference		
Homo sap	nens (numan) g	enome view	アビノ	ノクの事業が	<u>~</u> 07	CRA_TCAGchr7v248	search t	ne numan genome
Build 36.3 st	tatistics Switch to	previous build				Celera		
	2					HuRef		

<u>Hon</u> Build	no sa 1 36.3 :	<u>pie</u> stat	istics	huma Swit	<u>n)</u> ge ch to	enom previo	e vie us bui	w ild							В	LAST search	the human genome
Hitsi			2	2		- - - 33	1 1 3		- <b> -</b>   7	2	10 16	ļ	1	<b>Ì</b>			
Hitss	Ì		İ	15	1	<b>j</b> 19	12	T 20 9	<b>1</b> 21	<b>i</b> 22	8	Î Y	ů	=] not placed			
	Sear	ch 1	result	s for	query	"GD	NF	69 hi	its (as:	sembly	refer	ence)	$\geq$	ビット 動 絞込み	マの ができてい	5	0.1171
Chr		Ma	tch									Map	eleme	ent	Туре	Maps	Gene
5		all r Sirr GE tk0 RE Rat	natch ilar to NF f 5a07. CEPT tus no	GDN amily r x1 NC COR A	NF fan recept CI_CC LPHL rus glis	nily rec or alp? 3AP_I A al cell.	ceptor na 3 .u24 E .Gdnf	alpha Iomo ), mRl	3 prep .GDN NA	roprot IF FA	ein MILY	Bt.9 Hs.5 AI88 NM	5288 8042 7844 0191	1 39.1	Bt_EST_Cl Hs_EST_Cl TRANSCRIP1 TRANSCRIP1	<u>Bt UniG</u> Hs UniG Ms RNA	Transcript : all RefSeq STS Unigene
		Rat	tus no tial	avegio	us net	urotrop	bic	GDNF	(Gdn	f) mRl	VA,	AF4	97634	1	TRANSCRIPT	RnRNA	Filter

このフィルタにより、結果は絞り込まれますがそれでも69個にヒットします。しかし、得られた結果をよく見ると GDNFファミリ受容体といった関連遺伝子であったり、マーカーであったりして、今探したいGDNFという遺伝子は、 赤四角で囲んだ5番染色体に存在することがわかります。

glial cell line-derived neurotrophic factor receptor alpha-3	GFRA3	GENE	<u>Genes_cyto</u> <u>Genes_seq</u>
$\mathbf{GDNF}$ : glial derived neurotrophic factor	GDNF	GENE	Genes_cyto   Genes_seq
GDNF : glial cell derived neurotrophic factor	GDNF	PHENOTYPE	Pheno

このように遺伝子を探したいのであれば、下図に示されたQuick FilterからGeneを選択してFilterボタンを押すことで絞り込むこともできます。

Match	Map element	ту <mark>クイックフィル ps</mark>	Quick Filter
all matches		ターによる	Transcript :
Similar to GDNF family receptor alpha 3 preproprotein	Bt.95288	Bt 税込み Unic	🗌 all
GDNF family receptor alpha 3	Hs.58042	Hs_EST_Cl Hs UniG	RefSeq
tk05a07.x1 NCI_CGAP_Lu24 HomoGDNF FAMILY RECEPTOR ALPHA	AI887844.1	TRANSCRIPT HSRNA	STS Unigene
Rattus norvegicus glial cellGdnf), mRNA	NM_019139.1	TRANSCRIPT Rn RNA	
Rattus norvegicus neurotrophicGDNF (Gdnf) mRNA, partial	AF497634.1	TRANSCRIPT R. RNA	Filter
D			

下の例でわかるように20候補に絞り込まれています。しかし、assemblyによるフィルタは解除されてしまいます。

	Search res	ults for query "GDNF AND gene[obj_type]": 20 hit	<ul> <li>20hitに該</li> </ul>	り込ま	れている
Chr	Assembly	Match	Map element	Type	Maps
5	reference	all matches			
		glial cell line-derived neurotrophic factor receptor alpha-3	GFRA3	GENE	Genes_cyto Genes_seq
		GDNF : glial derived neurotrophic factor	GDNF	GENE	Genes cyto Genes seq
		GDNF : ENSG00000168621	GDNE	GENE	ensGenes
5	Celera	all matches			
		glial cell line-derived neurotrophic factor receptor alpha-3	GFRA3	GENE	craGenes Genes_seq
		GDNF : ghal derived neurotrophic factor	GDNF	GENE	craGenes Genes_seq

■ 解答1-2\_\_\_\_

Map Viewerを用いて遺伝子の位置情報を調べる

公的機関が出したreference Map上でのGDNFに関する詳細を見てみましょう。

	Search res	suits for query "GDINF AND genelon_type]": 20 hits			
Chr	Assembly	Match	Map element	Туре	Maps
5	reference	all matches			
		glial cell line-derived neurotrophic factor receptor alpha-3	GFRA3	GENE	Genes_cyto   Genes_seq
		GDNF : ghal derived neurotrophic factor	GDNF	GENE	Genes cyto Genes seq
		GDNF : ENSG00000168621	GDNE	GENE	ensGenes
5	Celera	all matches			
		ghal cell line-derived neurotrophic factor receptor alpha-3	GFRA3	GENE	craGenes   Genes_seq
		GDNF : shal derived neurotrophic factor	GDNF	GENE	craGenes   Genes seq

上記赤く囲った箇所をクリックします。

S NCBI	🛞 - NCBI М	ap Viewer			
PubMed	Entrez B	LAST	OMM To	ixonomy S	Structure
Search		Find	Find in This View	Advanced Se	arch
Human genome overview page (Build	Homo supiens (human	) Build 36.3 (C)	urrent)	BLAST The Hur	nan Genome
36.3)	Chromosome: 1234[	5 6 7 8 9 10 11 12	13 14 15 16 17 18 19	2 20 21 22 X Y MT	
overview page (Build	Query: GDNF [clear]	人 5 悉达/	<b>邑休</b>	Checkler Strength Strength	-
35.1)	Master Map: Genes On	Sequence	Summary of M	lans 🚺	aps & Options
Map Viewer Home	Region Displayed: 37,833K-	37,894K bp	D	ownload/View Sequ	encerevidence
Map Viewer Help	Genes_seq Symbol Q	Link		E Cyto	Descriptio
Human Maps Help	and the second second second				2010 CO. 1
F IP Data As Table View					
Maps & Options					
Compress Map	→ 表示範囲	動拡大・縮小	V.		
Region Shown					
37,833K	/	-			
37,894K Go	/ 位置の概要	2			
200m					
in in	1				
You are here:					
211					
5410 -					
総計 二 5422	GDNE + OMIN	HGNC sypr dlevn	am hm sts CCDS SNP	best RefSeq 5p13.1-	p12 glial cell d
5423 - 5431 -					
5432 - 5433 - 5434 -					
5405 -					
<ul> <li>default</li> </ul>	4				
<ul> <li>master</li> </ul>					

上図のような結果が表示されます。この図は、選択した5番染色体のGDNF遺伝子周辺を拡大表示したもので、ゲノムは上から下へと並んでいます。左中央部のイデオグラム内に、染色体内で今見ている場所がどこかが模式図的に示されており、表示範囲の移動、拡大、縮小もイデオグラムおよびその上のZoomボタンの利用で行えます。 Map Viewerでは、このように染色体のある位置にアノテーションされた様々な情報をユーザが選択して表示させることが可能です。複数の情報(Mapと呼ぶ)を同時に表示させることも可能ですし、表示方法をユーザが選ぶことも可能です。現在は、ユーザが選択したGenes\_seq (Genes On Sequence) というマップが表示されています。

図中赤丸で示したMaps & Optionsをクリックすると下に示した画面がポップアップされ、表示させるマップの選択 や表示方法をユーザが選ぶことが可能です。今は、赤丸で示したGeneのマップだけが右のリストにあるため、この Geneマップだけが表示されています。他のマップを合わせて表示したい場合には左のAvailable Mapsから表示した いマップを選択し、真ん中のADDボタンをクリックすることで可能になります。同様に、表示をやめたい場合には 右のMap Displayedから削除したいマップを選択し、真ん中のREMOVEボタンをクリックします。



座標を表示させます。座標をどのマップに表示させたいかを右のリストから選択し(今はGeneしかない)、その後 画面青四角で示したToggle Rulerをクリックしてください。次に示したように座標を表示させたいマップの横に 「R」マークがつきます。

📸 – Map Viewer – Google Chrome	00	8
http://www.ncbinlm.nih.gov/projects/mapview/maps_options.cg?YTAXID=9606&CHR=5&MAPS=genes{3783348	8.25:3789356	50.75]8
Organism: Homo sapiens Chromosome: 5 Region Shown: 37833488 37893560	E	<u>Help</u>
Available Maps: Maps Displayed (left to Org: human V Assembly: reference V Change Assembly	right):	
Ab initio Assembly Celera Genes Celera Transcripts Clone Component Contig CpG Island	<u>ve to Botto</u> ans 'ruler s	əm set')
More Options: Show Connections ♥Verbose Mode Compress Map: off ♥ Auto Compress if > 350 px Page Length: 30 Thumbnail View: ④ default (ideogram) ○ master OK Apply Close		

一通り設定後、OKボタンを押してください。 Maps & Option画面は自動的に閉じられ、Map Viewerの画面 Genes\_Seqの左横に座標情報が増えたのが確認できると思います。



座標を読み取りましょう。青い四角で示したようにGenes\_Seqマップ上で遺伝子は水色で示されています。この図 はエクソンーイントロン構造を表しており、太くなっている部分がエクソンです。黒線で表されたゲノムの左右で 遺伝子の向きを示しています。黒線の右側に表示されている遺伝子は、図中上から下向きに5'→3'、左側の遺伝子 は図中下から上向きに5'→3'となっており、それぞれ+鎖、一鎖を表しています。また、ユーザが検索あるいは選 択した遺伝子はハイライトされ、赤のバーが横についていてわかりやすくなっています。遺伝子名(GDNF)の右横 にも遺伝子の向きが矢印で示されています。



GDNF遺伝子は5番染色体37.85~37.87Mb付近に、一鎖方向に位置していることが確認できます。

Map ViewerでGenes\_Seqマップを表示させている場合、画面には個々の遺伝子について、左から遺伝子シンボル、 向き、リンク、エビデンス、バンド情報、説明が順に表示されています(ページの上に簡単な説明があります)。 リンクにはOMIM、HGNC、sv、pr、dl、ev、mm、hm、sts、CCDS、SNPがあり、それぞれOMIM、HUGOの遺伝子に関する 説明のページ、 NCBI sequence viewer、アミノ酸配列、配列のダウンロード、evidence viewer、model maker、 homologene、 UniSTS、CCDS、dbSNPsへのリンクとなっています。この中でdlと書かれたリンクをクリックしてく ださい。

Map Viewerを用いて遺伝子周辺のゲノム配列を取得する

< → C ☆	http://www	.ncbi.nlm.	nih.go	v/projec	ots/mapview/se	a q_reg.cgi?taxid=9	606&chr=5&fro
Homo sapiens (hur Region to retrieve Chromosome: 5	<u>nan)</u> (Build (in chromo	36.3) some coo Strand:	ordina plus	ates):			
from: 37851510	adjust by	-0K					
to: 37875539	adjust by	+0K		Change	Region/Strand	)	
This chromoson	ne region o	corresp	onds	s to the	contig region	n(s):	
Contig s NT_006576.15 377	<b>tart s</b> 88510 3781	top st 12539	rand +	<u>Display</u>	<u>Save to Disk</u>	View Evidence	ModelMaker
						Disclaimer   Write to NCBI   NLM	the Help Desk 4   NIH

上のようなダウンロード領域の設定画面が開きます。「ダウンロード」では、ゲノム上の指定した領域を切り出し て表示、保存することができます。但し、デフォルトでは向きが+鎖となっています。今回の課題のように遺伝子 配列(遺伝子がコードされているゲノム配列、イントロンを含む)をダウンロードしたい場合には遺伝子の向きに 合わせてStrandを変更する必要がある場合があります。 GDNFは-鎖にアノテートされているので、上図赤丸で示 したstrandをminusに変更して、"Change Region/Strand"をクリックしてください。すると、下図のようにダウン ロードする領域が相補鎖になりますので、その後Save to Diskをクリックして配列を保存してください。

← → C ☆ http://www	v.ncbi.nlm.nih.gov/projects/mapview/seq.reg.cgl?taxid=9606&.chr=5&.from
<u>Homo sapiens (human)</u> (Build Region to retrieve (in chromo	36.3) como scordinatos)
Chromosome: 5	Strand: minus 💌
from: 37851510 adjust by	-OK plus
to: 37875539 adjust by	T+UK Change Region/Strand
Sequence Format: FASTA	corresponds to the contig region(s):
Contig start s NT_006576.15 37788510 378	t <b>op strand</b> 12539 + <u>Display Save to Disk View Evidence ModelMaker</u>
	Disclaimer   Write to the Help Desk NCBI   NLM   NIH

Home Region	n to retrieve (i nosome: 5	<u>an)</u> (Build n chromo	36.3) some coo Strand	ordinates):
from:	37851510	adjust by	-0K	
to:	37875539	adjust by	+OK	Change Region/Strand
Reve	rse Complen	nent of th	nis chro	mosome region corresponds to the contig region(s):
Reve Ca NT_00	ontig sta 06576.1537788	nent of th art s 3510 3781	top 2539 st	trand - Visplay Save to Disk View Evidence ModelMaker

ダウンロード領域設定画面で、切り出す範囲を遺伝子の5'側に5kb広げてください。この遺伝子は-鎖にアノテートされているので、ゲノム上では切り出す領域はq-arm側(toに表示されている方)に5kb延ばす必要があることになります。 Toの横に表示された数字の右のadjust by:に+5Kと入力し、Change Region/Strandをクリックしてください。下のダウンロード範囲も広がります。その後、Save to Diskをクリックして配列をダウンロードしてください。

<ul> <li></li></ul>	
Homo sapiens (human) (Build 36.3)         Region to retrieve (in chromosome coordinates):         Chromosome: 5       Strand: minus ♥         from: 37851510       adjust by:-0K         to: 37875539       adjust by:+5K         Change Region/Strand         Sequence Format:       FASTA         FASTA       上流5Kb         Reverse Complement of this chromosome region corresponds to the contig region(s):         Contig       start         stop       strand         NT_006576.15       37788510         adjust       - Display         Save to Disk       View Evidence         Model/Maker         Disclaimer   Write to the Help Desk NCBI   NLM   NHH	
Region to retrieve (in chromosome coordinates): Chromosome: 5 Strand: minus ♥ from: 37851510 adjust by: 0K to: 37875539 adjust py: 45k Change Region/Strand Sequence Format: FASTA ♥ 上流5Kb Reverse Complement of this chromosome region corresponds to the contig region(s): Contig start stop strand NT_006576.15 37788510 37812539 - Display Save to Disk View Evidence ModelMaker Disclaimer   Write to the Help Desk NCBI   NLM   NIH Change Region/Strand Disclaimer   Write to the Help Desk NCBI   NLM   NIH	
Chromosome:       5       Strand: minus ▼         from:       37851510       adjust by:       0K         to:       37875539       adjust by:       +5K         Change Region/Strand       Sequence Format:       FASTA       上流5Kb         Reverse Complement of this chromosome region corresponds to the contig region(s):         Contig       start       stop       strand         NT_006576.15       37788510       37812539       – Display       Save to Disk       View Evidence       ModelMaker         Disclaimer   Write to the Help Desk       NCBI   NLM   NIH       NCBI   NLM   NIH       NCBI   NLM   NIH	
from: 37851510 adjust by:+0K to: 37875539 adjust by:+5K Change Region/Strand Sequence Format: FASTA ▲ 上流5Kb Reverse Complement of this chromosome region corresponds to the contig region(s): Contig start stop strand NT_006576.15 37788510 37812539 - Display Save to Disk View Evidence ModelMaker Disclaimer   Write to the Help Desk NCB    NLM   NIH Contig http://www.ncbi.nlm.nih.gov/projects/mapview/seq.reg.cgl?TAXID=96068AQSTR=GD	
to: 37875539 adjust by +5k Change Region/Strand Sequence Format: FASTA ・ 上流5Kb Reverse Complement of this chromosome region corresponds to the contig region(s): Contig start stop strand NT_006576.15 37788510 37812539 - Display Save to Disk View Evidence ModelMaker Disclaimer   Write to the Help Desk NCBI   NLM   NIH Contig start stop strand Disclaimer   Write to the Help Desk NCBI   NLM   NIH	
Sequence Format: FASTA ▲ 上流5Kb Reverse Complement of this chromosome region corresponds to the contig region(s): Contig start stop strand NT_006576.15 37788510 37812539 - Display Save to Disk View Evidence ModelMaker Disclaimer   Write to the Help Desk NCBI   NLM   NIH C A http://www.ncbi.nlm.nih.gov/projects/mapview/seq.reg.cgl?TAXID=96068&QSTR=GD	
Reverse Complement of this chromosome region corresponds to the contig region(s):         Contig       start       stop       strand         NT_006576.15       37788510       37812539       - Display       Save to Disk       View Evidence       ModelMaker         Disclaimer   Write to the Help Desk       NCBI   NLM   NIH         C       Image: NCBI   NLM   NIH	
Contig       start       stop       strand         NT_006576.15       37788510       37812539       - Display       Save to Disk       View Evidence       ModelMaker         Display         Display       Save to Disk       View Evidence       ModelMaker         Display         Display <td cols<="" td=""></td>	
NT_006576.15       37788510       37812539       – Display       Save to Disk       View Evidence       Model/Maker         Disclaimer   Write to the Help Desk       NCBI   NLM   NIH       Disclaimer   Write to the Help Desk         NCBI   NLM   NIH       NCBI   NLM   NIH       Disclaimer   Write to the Help Desk         C       ☆       http://www.ncbi.nlm.nih.gov/projects/mapview/seq.reg.cgl?TAXID=9606&QSTR=GD	
Disclaimer   Write to the Help Desk NCBI   NLM   NIH ← → C ☆ http://www.ncbi.nlm.nih.gov/projects/mapview/seq.reg.cgi?TAXID=9606&QSTR=GD	
NCBINLMINH         (← →)       ()         ()       ()      <	
C A http://www.ncbi.nlm.nih.gov/projects/mapview/seq_reg.cgi?TAXID=9606&QSTR=GD	
C ☆ http://www.ncbi.nlm.nih.gov/projects/mapview/seq.reg.cgi?TAXID=9606&QSTR=GD	
C ☆ http://www.ncbi.nlm.nih.gov/projects/mapview/seq_reg.cgi?TAXID=9606&QSTR=GD	
Homo saplens (human) (Build 36.3)	
Region to retrieve (in chromosome coordinates):	
Chromosome: 5 Strand: minus S	
from: 37851510 adjust by: -0K	
to: 37880539 adjust by: +0K Change Region/Strand	
Sequence Format: FASTA 🕑5kb広がった	
Reverse Complement of this chromosome region corresponds to the contig region(s):	
Contig         start         stop           NT_006576.15         37788510         37817539         - Display         Save to Disk         View Evidence         Model/Maker	
Disclaimer   Write to the Help Desk	
NCBI   NLM   NIH	

## Map Viewerにクローン関連情報を表示する

Map Viewerの画面に戻って、Maps & Options を再度クリックしてください。

表示された上記画面の左に表示されたAvailable MapsからClone, Component, Contigを選択して(複数選択はCtrl キーを押しながら行ってください)画面中央のADDボタンを押してください。

Organism: Homo sap Chromosome: 5	iens Region Shown:	37833488 37893560	Help
Available Maps: Org: human 💌	Assembly:	Maps Displayed (left to Change Assembly	right): ]
Sequence Maps Ab initio Assembly Celera Genes Celera Transcripto Clone Component Contis Contis Contis		[] Gene Move UP Move DOWN Make Master/Move Tossle Ruler ([R] before map mean	e to Bottom ns 'ruler set')
More Options: Show Connections Compress Map: off Page Length: 30 Thumbnail View: Ode OK Apply Close	♥ Verbose Mode Auto Compress if > fault (ideogram) ○ ma	350 px aster	

すると、選択された情報は右側のMaps Displayedのカラムに追加されます。



その後、左下のOKボタンを押すと、Maps & Options の画面は自動的に閉じられ、 Map Viewerのページが次の図の ようになります。画面にClone, Comp, Contigのマップが追加されたことが確認できます。 Cloneのマップでは FISHおよびend配列によってこの領域にマッピングされたクローン情報が、 Contigのマップでは、この領域に該当 するNCBIコンティグ情報が、そしてそのアセンブルされたコンティグを構成している各クローンの情報が



Componentのマップに表示されています。ですから、Componentには配列解読されたクローンが並びます。

図からコンティグは、NT\_006576、GenBankのアクセッション番号はAC008869.5であることが確認できます。また、Compのマップに表示されているクローンはオレンジ色のものがドラフト、水色のものが完成配列を示しています。AC008869.5は水色なので完成配列です。

Map Viewerに表示されたクローンの詳細情報を知る

Cloneマップを見てみましょう。この画面上GDNF遺伝子の付近には数多くのクローンがマップされており、それぞれをクリックすると詳細情報がポップアップされます。



その中でRP11やCTで始まる緑色で示されているクローンをクリックすると(例えばCTD-2653K15)



さらに別ページにリンクされており、クリックすることで次のようなクローンに関する詳細なページへとジャンプ

します。

< → C ☆ ht	ttp://www.ncbi.	nlm.nih.gov/genom	e/clone/clnam	e.cgi?styp	e=Name&list=C`
S NCBI		Registry st OMIM	тероатессора ос ататаса сасасаса геосатаертеттерс Тактакосаат геос Тахопо	ATATATTATA GOCGCGOAT, ATOC ATOGOC ACACOGCGO My S	TAGE TEGATEGATE AGEATGAET GATET. I TAGAGAG AG OGET. I COMATEGRE GE GAG Structure
Search CibneRegistr	y v 101		<u> </u>		
Clone Registry Home Page	Clone-ID: Comment:	CTD-2653K15 EndSeqInfo prod	CurrState: cessing	Free	
	FingerPrint	Information: No	>		
Clone	End-Sequer	nce Information	エンド配列情	報	
- Hornendare	Accession	Seqlen(bp)	Repeat	Hit	End
BAC End Report	AQ625576	567	No	Yes	M13F
	AQ625580	519	No	Yes	R
Distributor			品作者中国		
nionnauon	Library Info	College burger	DAO Ilhumus D		
	Organiem:	Homo canions	BAC library D	Type	BAC
	Annroved:	Yes		Sex:	Male
	Source:	sperm		Rsite:	ivere .
	Distributor	Information	币元情報		
	Available from	n clone library			
	Name:	Research Gene	tics		
	Email:	libraries@resgen	.com	0.1	1.0
	Web:	http://www.resg	gen.com/resour	rces/inde	x.php3
	Transaction	History			
	G. Center	Trans-State	Chrom	Result	Submit-date
	SERVER	Free		Accept	06/22/1999
				24	
	0.5				

このページではクローンCTD-2653K15に関して、エンド配列の情報やゲノムにアライメントできたかなどの詳細情 報が書かれています。さらに赤く囲んだところにこのクローンの配布元が記載されています。このクローンでは、 Research Genetics社から手に入れることが可能なことがわかります。

## ■ 解答1-4\_\_\_\_

Map Viewerに遺伝子予測結果やRefseq情報を表示する

Maps & Optionsをクリックして、ポップアップした画面からGene以外を選択してその後REMOVEを押してください。



続いて、左のリストからAb initioとRefseq Transcriptsを選択し(ctrlキーを押しながら)、 ADDボタンを押し てください。

🛅 - Map Viewer - Google Chrome	_ 0 X
http://www.ncbinlm.nih.gov/projects/mapview/maps_options.cgi?TAXID=960	068CHR=58MAPS=clone,comp.cntg.genes[37833488:25:378!
Organism: Hom <u>o</u> sapiens Chromosome: 5 Region Shown: 37833488 3	87893560]
Available Maps: N	Maps Displayed (left to right):
Org: human 🗹 Assembly: reference 💟	Change Assembly
GenBank DNA	RI Gene Move UP
Phenetype	Move DOWN
RefSeq Transcripts ADD>>	Make Master/Move to Bottom
maBt < <remove< td=""><td></td></remove<>	
maGga	Tossle Ruler
rnaMm M	[[R] before map means 'ruler set')
More Options:	
Show Connections Verbose Mode	
Compress Map: off 💌 Auto Compress if > 350 px	
Page Length: 30	
Thumbnail View: 💿 default (ideogram) 🔿 master	
OK Apply Close	

最後にOKボタンを押してください。



下記のような画面が表示されると思います。



一番左にはAb intio法で予測された遺伝子構造が、真ん中のRefseqにはRefseqの配列をマップした情報が表示されています。一番右にはGDNFの代表的な遺伝子構造がアノテーションされていますが、一番左にある予測結果と比べて下さい。どうでしょうか?完全一致はしないものの塗りつぶされた四角のエクソン領域の多くは一致することが確認できます。

上記図の真ん中のRefseqを見てみると3種類のスプライスバリアントがアノテーションされているのが確認できます。

Map Viewerに表示されたデータを表形式で見る



図中左側に示されたData As Table Viewをクリックしてください。

Homo s	aniens (h	uman) Buil	d 36 3 (Current)		BLAST T	he Human Genome
Data As T	able View		d dot o (ourreing	Down	load All !	ne ridinari Genome
-	<u>^</u>	Genes –				
Genes On Region Dis ポジ	Sequence played: 37, ション	833K-37,894 nosome: 144	All Sequence Mar IK bp 8 [4 not localized]	Download/V	iew Sequence/Evidence	Download Data
スタート	、エンド	Symbol Q	Links		E Cyto	Description
37851510	37875539	GDNE - ON	AMHGNC sv prdlev mm	hm sts CCDS SN	P best RefSeq 5p13.1-p	o12 glial cell derived ne
Ab initio m Region Dis Total mod models in	played: 37, els On Chro Region: 3	o initio 833K-37,894 omosome: 351	II Sequence Maps IK bp I9 [7 not localized]	Download/V	iew Sequence/Evidence	<u>prev next</u> Download Data
start	stop	Symbol	Q Best hit Best	hit description		
37841490	37848120	hmm1886934	-			
37851510	37875533	hmm2052293	3 - NP 0005051 glial o	ell derived neuro	trophic factor isoform 1	l preproprotein [Homo
			and the second se			
37870721 RefSeq Tr Region Dis	3/881574 anscripts C played: 37,	n Sequence 833K-37,894	RefSeq All	Sequence Maps Download/V	iew Sequence/Evidence	prev Download Data
37870721 RefSeq Tr Region Dis Total RefS RefSeq Tr	anscripts C played: 37, Seq Transcr anscripts in	In Sequence 833K-37,894 ripts On Chron h Region: 3	RefSeq All Ik bp mosome: 1455 <u>(3 not k</u>	Sequence Maps Download/Vi calized]	iew Sequence/Evidence	prev Download Data
37870721 RefSeq Tr Region Dis Total RefS RefSeq Tr start	3/8815/4 anscripts C played: 37, Seq Transcr anscripts in stop	hmm1288934 In Sequence 833K-37,894 ripts On Chron Region: 3 Accession	HefSeq         All           1k bp         mosome: 1455 [3 not log           Locus Q         Links	Sequence Maps Download/V calized] Align quality	iew Sequence/Evidence Description	prev Download Data
37870721 RefSeq Tr Region Dis Total RefS RefSeq Tr start 37851510	37881574 anscripts C played: 37, Seq Transcr anscripts in stop ) 37870655	Imm 1888934 In Sequence 833K-37,894 ripts On Chron n Region: 3 Accession NM 199234.1	RefSeq All No bp mosome: 1455 (3 not lo Locus Q Links GDNE - uzsv prev BL	Sequence Maps Download/V calized Alizn quality ink identical	iew Sequence/Evidence Description glial cell derived neurot	Download Data
37870721 RefSeq Tr Region Dis Total RefS RefSeq Tr start 37851510 37851510	37881574 anscripts C played: 37, Seq Transcr anscripts in stop 37870655 37851857	Imm1888934 In Sequence 833K-37,894 ripts On Chron h Region: 3 Accession NM 199234.1 exon	RefSeq All Ik bp mosome: 1455 <u>(3 not lo</u> Locus Q Links <u>GDNE - ugsvprev BL</u> -	Sequence Maps Download/Vi calized Align quality ink identical CDS	iew Sequence/Evidence Description glial cell derived neurot 348 bp	prev Download Data rophic factor
37870721 RefSeq Tr Region Dis Total RefS RefSeq Tr start 37851510 37851510 37851858	37881574 anscripts ( played: <b>37</b> , Seq Transcr anscripts in stop 37870655 37851857 37851857	hmm1888934 In Sequence 833K-37,894 ripts On Chron Region: 3 Accession NM 199234.1 exon intron	RefSeq All Ik bp mosome: 1455 <u>(3 not le</u> Locus <u>Q</u> Links <u>GDNF</u> – ugsvprev BL – –	Sequence Maps Download/V calized Align quality ink identical CDS	iew Sequence/Evidence Description glial cell derived neurot 348 bp 18736 bp	prev Download Data rophic factor
37870721 RefSeq Tr Region Dis Total RefS RefSeq Tr start 37851510 37851510 37851858 37870594	37881574 anscripts ( played: 37, seq Transcr anscripts in stop 37870655 37851857 37870593 37870593	https://www.initescondinguesian https://wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww	RefSeq All Ik bp mosome: 1455 <u>[3 not k</u> Locus <u>Q</u> Links GDNE - uzsverev BL - - -	Sequence Mans Download/V calized Align quality ink identical CDS CDS	iew Sequence/Evidence Description glial cell derived neurot 348 bp 18736 bp 54 bp	prev Download Data
37870721 RefSeq Tr Region Dis Total RefS RefSeq Tr start 37851510 37851510 37851510 37851858 37870594 37870648	37881574 anacripts ( played: 37, Seq Transcr anscripts in stop 37870655 37851857 37870593 37870593 37870647 37870655	https://www.commission	F RefSeq All In bp mosome: 1455 <u>[3 not le</u> Locus Q Links GDNE - uzsverev BL - - - -	Sequence Maps Download/V calized Align quality ink identical CDS CDS UTR	iew Sequence/Evidence Description glial cell derived neurot 348 bp 18736 bp 54 bp 5 bp	prev Download Data
37870721 RefSeq Tr Region Dis Total RefS RefSeq Tr start 37851510 37851858 37851858 37870594 37870648 37851510	37881574 anscripts C played: <b>37</b> , Seq Transcripts in stop 37870655 37870655 37870593 37870647 37870647 37870655 37870655	hmm13838344 In Sequence 833K-37,894 ripts On Chron Region: 3 Accession NM 199234.1 exon intron exon exon NM 199231.1	RefSeq All No bp mosome: 1455 (3 not lo Locus Q Links GDNE - uzsvprev BL - - - - GDNE - uzsvprev BL	Sequence Maps Download/V calized Alian quality ink identical CDS UTR ink identical	iew Sequence/Evidence Description glial cell derived neurot 348 bp 18736 bp 54 bp 54 bp glial cell derived neurot	prev Download Data trophic factor
37870721 RefSeq Tr Region Dis Total RefS RefSeq Tr start 37851510 37851510 37851510 37870594 37870648 37851510 37851510	37881574 anacripts C played: <b>37</b> , Seq Transci anscripts in stop 37870655 37870655 37870593 37870647 37870647 37870655 37870655 37871350 37851994	hmm13838344 In Sequence 833K-37,894 ripts On Chron a Region: 3 Accession NM 199234.1 exon exon exon NM 199231.1 exon	RefSeq All No bp moscome: 1455 (3 not lo Locus Q Links GDNE - uzsv prev BL - - - - - - - - - - - - -	Sequence Maps Download/V calized Alian quality ink identical CDS UTR UTR ink identical CDS	iew Sequence/Evidence Description glial cell derived neurot 348 bp 18736 bp 54 bp 54 bp glial cell derived neurot 485 bp	prev Download Data trophic factor
37870721 RefSeq Tr Region Dis Total RefS RefSeq Tr start 37851510 37851510 37851510 37851510 37870648 37870648 37851510 37851510 37851510	37881574 anacripts C played: <b>37</b> , seq Transci anscripts in stop 37870655 37870655 37870655 37870647 37870655 37870655 37871350 37851994 37870582	hmm1388844 In Sequence 833K-37,894 ripts On Chron a Region: 3 Accession NM 199234.1 exon exon NM 199231.1 exon intron	RefSeq All No bp moscome: 1455 (3 not is Locus Q Links GDNE - uzsyprey BL - - - - - - - - - - - - -	Sequence Maps Download/V calized Mian quality ink identical CDS UTR UTR ink identical CDS	iew Sequence/Evidence Description glial cell derived neurot 348 bp 18736 bp 54 bp 8 bp glial cell derived neurot 485 bp 18588 bp	prev Download Data
37870721 RefSeq Tr Region Dis Total RefS RefSeq Tr start 37851510 37851510 37851510 37851510 37870594 37870594 37870594 37851510 37851510 37851510 37851510 37851510 37851510 37851510 37851510 37870583 37870585 37870585 37870585 37870585 37870585 37870585 37870585 37870585 37870585 37870585 37870585 37870585 37870585 37870585 37870585 37850585 378705858 3787058585 3787058585 3787058585 378705	37881574 anacripts C played: <b>37</b> , seq Transci anscripts in stop 37870655 37870655 37870655 37870647 37870655 37871350 37851994 37870552 37870655	hmm1388844 In Sequence 833K-37,894 intro On Chron n Region: 3 Accession NM 199234.1 exon exon NM 199231.1 exon intron exon intron exon intron exon	RefSeq All RefSeq All No bp moscome: 1455 (3 not is Locus Q Links GDNE - ussyprey BL - - - - - - - - - - - - -	Sequence Mana Download/V calized ink identical CDS UTR ink identical CDS CDS CDS	ew Sequence/Evidence Description glial cell derived neurot 348 bp 18736 bp 54 bp 8 bp glial cell derived neurot 485 bp 18588 bp 73 bp	prev <u>Download Data</u> crophic factor
37870721 RefSeq Tr Region Dis Total RefS RefSeq Tr start 37851510 37851510 37870648 37870648 37851510 37851510 37851995 37870583 378706583	37881574 anscripts C played: 37, Seq Transcripts in stop 37870655 37870655 37870655 37870655 37870655 37871350 37851994 37870582 37870582 37870655 37870655	hmm1388844 In Sequence 833K-37,894 intro A Chron h Region: 3 Accession NM 199234.1 exon exon NM 199231.1 exon intron exon intron exon exon intron exon	RefSeq All No bp mosome: 1455 <u>[3 not le</u> Locus <u>Q</u> Links <u>GDNF - ugsvprev BL</u> - - - - - - - - - - - - - - - - - - -	Sequence Mana Download/V calized ink identical CDS UTR ink identical CDS UTR CDS UTR CDS UTR	iew Sequence/Evidence Description glial cell derived neurot 348 bp 18736 bp 54 bp glial cell derived neurot 485 bp 18588 bp 73 bp 26 bp	prev Download Data
37870721 RefSeq Tr Region Dia Total RefS RefSeq Tr start 37851510 37851510 37850594 37870648 37851510 37851510 37851510 37851510 37850583 37870682 37870682	37881574 anacripts C played: 37, Seq Transci anscripts in stop 37870655 37870655 37870655 37870647 37870647 37870655 37871350 37851994 37870582 37870582 37870655 37870655 37870681 37870681	hmm13838344 In Sequence 833K-37,894 ripts On Chron Region: 3 Accession NM 1992341 exon intron exon NM 1992311 exon intron exon intron exon intron exon intron exon intron exon intron	RefSeq All No bp mosome: 1455 <u>[3 not le</u> Locus <u>Q</u> Links <u>GDNF</u> - ugsvprev <u>BL</u> - - - - - - - - - - - - - - - - - - -	Sequence Man Download/Vi calized ink identical CDS UTR ink identical CDS CDS UTR CDS UTR	iew Sequence/Evidence Description glial cell derived neurot 348 bp 18736 bp 54 bp glial cell derived neurot 485 bp 18588 bp 73 bp 26 bp 572 bp	prev Download Data
37870721 RefSeq Tr Region Dis Total RefS RefSeq Tr start 37851510 37851510 37851510 37870594 37870548 37851510 37851510 37851510 37851510 37851510 37850583 37870583 37870682 37870682 37871254	37881574 anacripts C played: 37, Seq Transci anscripts in stop 37870655 37870655 37870655 37870647 37870647 37870655 37871050 37870655 3787065	hmm13838344 In Sequence 833K-37,894 ripts On Chron Region: 3 Accession NM 199234.1 exon intron exon NM 199231.1 exon intron exon intron exon intron exon intron exon intron exon intron exon exon intron exon exon intron exon exon intron exon exon intron exon exon intron exon exon intron exon exon intron exon	RefSeq All No bp mosome: 1455 <u>[3 not ls</u> Locus <u>Q</u> Links <u>GDNF</u> - ugsvprev <u>BL</u> - - - <u>GDNF</u> - ugsvprev <u>BL</u> - - - - - - - - - - - - -	Sequence Mans Download/V calized ink identical CDS UTR ink identical CDS UTR CDS UTR UTR UTR	ew Sequence/Evidence Description glial cell derived neurot 348 bp 18736 bp 54 bp 8 bp glial cell derived neurot 485 bp 18588 bp 73 bp 26 bp 572 bp 97 bp	prev Download Data

上記のように図で表示されていたデータが表形式で示されます。図では確認しづらかったゲノム上の位置が数値と して表現されているなど、表形式ならではのメリットがあることがわかります。

# Model Makerを使って転写産物候補を新たに作成する

Backボタンを使って画面を一つ戻ってください。

GDNFに関するLinkの中からmmをクリックしてください。



以下のようなGDNF遺伝子に関するModel Makerのページが表示されます。



このModel Makerでは、ある遺伝子に関してmRNAやESTの配列に由来するエクソン候補をユーザが自由に組み合わせ ることで転写産物候補を自由に組み立ててその配列を得ることができるようになっています。まず、図の上部に Evidenceと書かれた領域があり、その部分に当該遺伝子に関する転写産物が水色で描かれています。その下に Putative Exonsと書かれた領域があり、上部の転写産物から予測されるエクソン候補が緑色で示されています。こ のGDNFでは11個のエクソン候補があることがわかります。さらにその下のYour Modelの部分では、この11個のエク ソン候補を組み合わせてユーザが転写産物を組み立てることができます。一番下のPutative Exonsの中から使いた いエクソンをユーザが選択すると、そのエクソンをつないだ配列が上のText Boxに表示されます。しかし、任意の エクソン同士を自由に組み立てることはできません。フレームの連続性を満足させる必要やエクソンが重なっては いけないという制約があるからです。例えば、1番目のエクソンの次には5番目のエクソンしか並べられないことが 示されています。

Put	ative e	exons (table view):	custom exons intron bases 2
		37812539 CGICCGCCTICCA	37812533-37812366 GAGIGT = 5
	2	CT   CCA	37808686-37808493 ATG GT => 4 or 5
	3	CC CAT	37808350-37808254 TGG GT => 4
	4	2 or 3 <= AG GTG	37807681-37807583 CCG GT => 10
	5	<u>1</u> or <u>2</u> <= AG GTG	37807681-37807505 ACT GT => 10
	6	AG ATG	37807655-37807594 CCC GC => 11
	7	AG ATG	37807655-37807583 CCG GT => 9 or 10
	8	AG ATG	37807655-37807505 ACT GT => 10
	9	$\overline{2} \le AG   GGC$	37796948-37796869 GAG GT => 10
	10 <u>4</u> «	or $5$ or $7$ or $8$ or $9 <=$ AG CAA	37788994-37788513 ATC TGA CT 37788510
	11	<u>6</u> <= CA   GOT	37788857-37788510 TGA CT

下の図は、これらの制約を満足するように1-5-10番目のExonを選んだ例です。

Your model:	clear
AF053748.1, NM_0005142 (algn), hmm20522933. NM 000514 (annot) ユーザモデルと同じ 既存情報の表示	42 模式図
COGCCTCCAGOGCCCCTGCTGCCCCCGCGGACCCCAGGATTGCGAACTCTTGCC ACCTGTTGGCGGGGCTCCGCCTCCAGCCATCAGCCCGATGGCTCTCTGCTG TTGGGCGACTGGATTAATGTCCAACCTAGGCTCGCGGAGCCCGATCGCT CCGCCGGACGGGACTTTAAGATGAACTTATGGGATGTCGGCGGTGTCCCCGGG	CCTG GGAC GCCG CTGC
Frame ORF=42 CDD      Frame2 ORF=22 DD	Frame? ORF=211 ODD
ppparpccparpqdcellpl	assapllpratperriapd A llggaprsshapgwvswigi gapgynvaprvcgdpirgaa agrafkMKLWDVVAVCLVLL
Putative exons (table view): gustom exons	intron bases:2
<ul> <li>図 1 37812539 CGICCGCCTICCA 37812533-37812366 G</li> <li>CTICCA 378006888-378006884 A</li> <li>CCICAT アミノルマンノルマックの設長</li> <li>CCICAT アミノルマンノルマンノルマンノルマンノルマンノルマンノルマンノルマンノルマンノルマン</li></ul>	AG GT =>5 TG GT => 4 or 5 GG GT => 4 ∰∥%Ω75
4 1,5,10を 2 or 3 <= AG GTG 37007001-37007003 J 1 or 2 <= AG GTG 37807681-37807505 A 違択 AG ATG 37807681-37807555 A	oelet ⇒ 10 cTlet ⇒ 00 minute 満足
AGIATG 37807655-37807583 C	CG GT => 9 or 10
AG ATG 37807655-37807505 A	CT GT => <u>10</u>
9 <u>7</u> <= AG GGC 37796948-37796869 G	AG GT => <u>10</u>
ID 4 or 5 or 7 or 8 or 9 <= AG CAA 37788994-37788513 A	TC TGA CT 37788510
11 6 <= CA GCT 37788857-37788510 T	GACT

赤く、Exon-Intron構造が模式図として表示され、この組み合わせはデータベース内のAF053748, NM\_000514などと 同じことが示されています。また、その下のText Boxには3つのExonをつなげた「転写産物」配列が、またその下 には3つのフレームそれぞれで予測される一番長いアミノ酸配列が示されています。この例ではFrame3の211AAから なるアミノ酸が一番長い候補となります。

Model Maker予測結果と既知アミノ酸配列との相同性検索

上の説明で示された、アミノ酸配列を用いてホモロジー検索を実施してもいいのですが、ここでは、もうひとつのやり方を紹介します。予測された「転写産物」の右横のORF Finderをクリックしてみてください。



上図がポップアップします。これは、cDNA配列などを入力とし、その中からORF候補を見つけてくれるORF Finder の結果画面になります。 +3のフレームに636bp (211AA:ストップコドンを除く)のORF候補があるのが確認でき ます。これをクリックしてみてください。



選択したORF候補がピンク色になるとともに、下にその配列が表示されます。また、上部でこのORFを入力としたホ モロジー検索が実施できます。例えば、Databaseをswissprotにしてblaspを用いたblast検索を実行してみましょう。

Program blastp 💌 Database	swissprot 💌	BLAST with parameters Cognitor
	nr month	
View 1 GenBank V Rec	swissprot pdb yeast ecoli	SixFrames Frame from to Length +3 ■ 201835 636 -2 ■ 599745 147

Databaseの横のプルダウンメニューからswissprotを選択し、 BLASTボタンを押せば次の画面のようにblast検索が 実行されます。

, BLAST					
Home Recen	t Results Saved S	itrategies Help			
CBI/BLAST/Format R	equest				
Query	Icl 5262 (211 letters)				
Database	nr				
Job title	Icl 5262 (211 letters)				
Request ID	Request ID BWRJJUH016 View report Show results in a new window				
Format	Show	Alignment S as HTML S Advanced View Use old BLAST report format	Reset form to defaul		
	Alignment View	Pairwise			
	Display	Graphical Overview Linkout Sequence Retrieval INCBK-gl			
		Masking Character: Lower Case 💌 Masking Color: Grey 💌			
	Limit results	Descriptions: 100 💌 Orsphical overview: 100 💌 Alignments: 100 💌			
		Organism Type common name, binomial, taxid, or group name. Only 20 top taxa will be shown.			
		Enter organism name or id-completions will be suggested			
		Entrez query			
		Expect Mix: Expect Mix:			
	5 200 0000298				

しばらく待ってからView Reportボタンを押すと、結果が表示されます。当然、GDNFに関する遺伝子が上位に並ん でいることが確認できます。



■ 解答1-6\_

オーソログ遺伝子情報の構造情報を見る

Model Makerの画面からBackボタンを用いてMap Viewerの画面に戻ってください。 Maps & Optionsをクリックし

て、Ab initio, Refseq TranscriptsをREMOVEしてください。続いてOrg:をchimpにしてGeneを加えてください。



同様にmouseのGene、ratのGeneを加えて、OKを押してください。



次の図のようにChimp, Mouse, Rat, HumanのGDNF周辺が並んで表示されているのが確認できると思います。



3'側の2つのコーディングエクソンの間に長いイントロンがあるなどの構造はよく似ているのが確認できます。

■ 解答1-7\_

## 疾患情報へのリンクをたどる

Maps & Optionsをクリックしてポップアップされる画面から、 Chimp, Rat, MouseのGenesを削除して、Phenotypeを 追加し、OKを押してください。



下のような画面が見られると思います。Phenotypeのところには、ゲノム上の位置が特定された表現型がマップされています。これらの中にはGDNFの例のように遺伝子まで特定されているものや、マーカーを利用したQTL解析でLOCUSとしてある領域に落とし込まれた結果など(下図中MHS6など。この疾患は5pに原因遺伝子があることのみがわかっている)が表示されています。



Phenotypeのところにある、GDNFをクリックしてください。簡単な説明画面がポップアップするので、その中のリ ンクをクリックしてください。



GDNFに関するOMIMのページが開きます。



OMIMはヒトのメンデル遺伝する疾患に関する情報を論文から抽出したもので、人手で作られているデータベースです。この文章を読むことで詳細なGDNFが関係する病気の情報が得られます。

# 解答・解説2

課題1の手順でGDNFのところをPRNPに変えて同じ事を実行してみてください。

原文更新日: 2005年2月15日 日本語版更新日: 2008年8月23日

All Rights Reserved, Copyright(C) 1997-2008 Japan Science and Technology Agency(JST)

