UCSCゲノムブラウザ チュートリアル

UCSCゲノムブラウザはゲノム解読がなされている真核生物を対象として自動アノテーションを行い、その結果をデ ータベースとして公開している UCSCが進めているプロジェクトです。NCBI MapViewerのようにゲノムベースでそ の上にアノテーションされている遺伝子などの情報を閲覧すると共に、ホモロジー検索や必要なデータのダウンロ ードなどの機能を提供しています。

UCSCゲノムブラウザでは、データの品質を高めるために人手を介する部分を極力減らす代わりに、非常に多様な計算結果を提供しており、ユーザ側で複数のトラックを並べて表示したりしながら情報の絞込みを行っていくような 使い方に向いています。

用いているゲノム情報はNCBI, Ensemblと同じものですが、アノテーションされている情報は独自に計算したものやNCBI, Ensemblの情報など多岐に渡っています。高速に自動アノテーションするため、表示されている情報自身は新しいものが多くなっているのも一つの特徴です。

このチュートリアルでは、UCSCゲノムブラウザ(ヒトゲノム版)の使い方をGDNF遺伝子の探索を通して学びます。 用いている例はMapViewerミニコースに合わせてありますので、照らし合わせながら見ていくとNCBI MapViewerや Ensemblとの違いなどがよくわかると思います。

1. ヒト遺伝子GDNFをUCSCゲノムブラウザで探してみる

では、グリア細胞由来の神経栄養因子遺伝子GDNFを例にとってUCSCゲノムブラウザの使い方を見て行くことにしましょう。 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。

■1-1. ヒト遺伝子の情報をUCSCゲノムブラウザで探してみる

<u>UCSCゲノムブラウザのトップページを開いてみる</u>

まず、UCSCゲノムブラウザのトップページを開いてみましょう。 をクリックしてみてください。

UCSC Genome Bioinformatics				
Genomes -	Blat - Tables - Gene Sorter - PCR - VisiGene - Proteome - Session - FAQ - Help			
Genome	About the UCSC Genome Bioinformatics Site			
Browser ENCODE	Welcome to the UCSC Genome Browser website. This site contains the reference sequence and working draft assemblies for a large collection of genomes. It also provides a postal to the ENCODE project.			
Blat Table Browser	We encourage you to explore these sequences with our tools. The <u>Genome Rrowser</u> zooms and scrolls over chromosomes, showing the work of annotators worldwide. The <u>Gene Sorter</u> shows expression, homology and other information on groups of genes that can be related in many ways. Biat quickly maps your sequence to the genome. The <u>Table Browser</u> provides convenient access to the underlying database. <u>VisiGene</u> lets you browse through a large collection of <i>in situ</i> mouse and frog images to examine expression patterns. <u>Genome Graphs</u> allows you to upload and display genome-wide data sets.			
Gene Sorter In Silico PCR	The UCSC Genome Browser is developed and maintained by the Genome Bioinformatics Group, a cross-departmental team within the Center for Biomolecular Science and Engineering (<u>CBSE</u>) at the University of California Santa Cruz (<u>UCSC</u>). If you have feedback or questions concerning the tools or data on this website, feel free to contact us on our <u>public mailing list</u> . To view the results of the Genome Browser users' survey we conducted in May 2007, click <u>here</u> .			
Genome Graphs				
Galaxy	News Archives >			
VisiGene	To receive announcements of new genome assembly releases, new software features, updates and training seminars by email, subscribe to the genome-announce mailing list.			
Proteome	17 October 2008 - Quality Assurance Engineer Job Opening			
Utilities	The UCSC Genome browser group has re-opened a search for a Quality Assurance engineer. Biological Database Testing and User Support Technician. Applicants should have a strong background in biology and computation, and excellent skills in scientific writing/editing. For details see the job listing			
Downloads	17 September 2008 - Seven Assemblies Archived			
Release Log	We have archived seven older assemblies of the human (hg15), mouse (mm5, mm6), rat (m2), C. elegans (ce1), zebrafish (danRer2), and cow (bosTau1) genomes. You can continue to browse these assemblies on the Genome Browser archive sever. At least two later assemblies for each			
Custom Tracks	ot these organisms remain available on the main Genome Browser ate. We do not provide blat servers or updated GenBank data for archived assemblies. However, it is still possible to lift coordinates between these			
Archaeal Genomes	assemblies and those on the main browser site.			
Mirrors	15 September 2008 - Request for Input on ENCODE Data Release Plan: The National Human Genome Research Institute (NHGRI) has updated the data release plan for the <u>ENCODE</u> and model organism ENCODE (<u>modENCODE</u>) projects. <u>Read</u>			
Archives	more.			
Training	3 September 2008 - Updated UCSC Genes Set Released for hg18: We have released an updated UCSC Genes data set on the latest human assembly (hg18, March 2006). Read more.			
Credits				

上図のようなトップページが立ち上がります。この中で、ゲノム情報のブラウズは画面左上の "Genome Browser"と書かれた箇所をクリックした下図の画面を通して行います。

clade	genome	assembly	position or search term	image width
Mammal 💌	Human 💽	Mar. 2006 💌	chrX:151,073,054-151,383,976	620 submit
	Human Chimp Orangutan Rhesus Marmoset Mouse Rat Guinea Pig	eset the browse racks config	er user interface settings to their da ure tracks and display Celear po	efaults. sition
Mar. 2006 (hg18)	Cat	equences)		
reference sequence (P	Dog Horse Cow	1) was produce	ed by the International Human Ge	nome Sequencing Consortium
ieries	Opossum Platypus			

左から分類(哺乳類、脊椎動物など)、生物種名、ゲノムのバージョンを選択するようになっています。例えばヒ トゲノムの最新版を見る場合には、左から"Mammal", "Human", "Mar.2006" を選択します。

clade	genome	assembly	position or search term	image width
018/00	Poulome		Popular of peptien term	and o most
Mammal 🛛 🎽	Human 💌	Mar. 2006 ⊻ 🛛	chrX:151,073,054-151,383,976	620 submit
	Click here to re	Mar. 2006 May 2004	user interface settings to their de	faults.
	add custom tr	July 2003 Ju	re tracks and display clear pos	ition

<u>UCSCゲノムブラウザでの遺伝子検索</u>

そして、その右横の検索窓に表示したい領域、あるいは検索キーを入力して"submit"ボタンをクリックします。

例えば、"chr3:1-100000"というように[染色体番号]:[start position]-[end position]と指定するとその範囲の ゲノムブラウザが立ち上がります。

また、検索キーワードを入力すると検索結果が一覧表示され、その中から希望するものをクリックすることで該当 するゲノム領域へと移動します。

GDNFと検索窓に入力し、検索を実行してみましょう。

clade	genome	assembly	position or search term	image width
Mammal 🛛 🖌	Human 💌	Mar. 2006 💌 G	DNF	620 submit
	Click here to re	set the browser	user interface settings to their de	faults.
	add custom tr	acks configur	e tracks and display clear po	sition

下図のようにUCSCゲノムブラウザで文字列検索を実施すると、UCSC遺伝子、RefSeq遺伝子、ヒト以外のRefSeq遺伝子(をヒトゲノム上にマッピングしたもの)などにヒットすることが確認できます。

■1-2. ヒト遺伝子をUCSCゲノムブラウザ上で見る

遺伝子名に検索キーワードが含まれている遺伝子を探す

検索結果を順に見ていくと、UCSC遺伝子名にGDNFが含まれているものは上からの7件で、いずれも5番染色体の 37.86Mb付近にヒットしていることがわかります。(Alternativeであることが示唆されます。)

UCSC Genes					
DNF (uc010ivd.1) at chr5:37851510-37870767 - glial cell derived neurotrophic factor isoform 2 DNF (uc010ivc.1) at chr5:37851510-37870681 - glial cell derived neurotrophic factor isoform DNF (uc00311a.1) at chr5:37851510-37870655 - glial cell derived neurotrophic factor isoform 1 DNF (uc0031k2.1) at chr5:37851510-37871686 - glial cell derived neurotrophic factor isoform 1 DNF (uc0031k2.1) at chr5:37851510-37871686 - glial cell derived neurotrophic factor isoform 2 DNF (uc0031k2.1) at chr5:37851510-37871686 - glial cell derived neurotrophic factor isoform 1 DNF (uc0031k2.1) at chr5:37851510-37871686 - glial cell derived neurotrophic factor isoform 2 DNF (uc0031k2.1) at chr5:37851510-37871686 - glial cell derived neurotrophic factor isoform 1 DNF (uc0031k2.1) at chr5:37851510-37871686 - glial cell derived neurotrophic factor isoform 1 DNF (uc0031k2.1) at chr5:37851510-37871850 - glial cell derived neurotrophic factor isoform 1 DNF (uc0031k2.1) at chr5:37851510-37871850 - glial cell derived neurotrophic factor isoform 1 DNF (uc0031k2.1) at chr5:37851510-37871850 - glial cell derived neurotrophic factor isoform 1 DNF (uc0031k2.1) at chr5:37851510-37871850 - glial cell derived neurotrophic factor isoform 2 DNF (uc0031k2.1) at chr5:37851510-37871850 - glial cell derived neurotrophic factor isoform 1 DNF (uc0031k2.1) at chr5:37851510-37871850 - glial cell derived neurotrophic factor isoform 2 DNF (uc0031k2.1) at chr5:37851510-37871850 - glial cell derived neurotrophic factor isoform 2 DNF (uc0031k2.1) at chr5:37851510-37871850 - glial cell derived neurotrophic factor isoform 2 DNF (uc0031k2.1) at chr5:37851510-37871850 - glial cell derived neurotrophic factor isoform 2 DNF (uc0031k2.1) at chr5:37851510-37871850 - glial cell derived neurotrophic factor isoform 2 DNF (uc0031k2.1) at chr5:37851510-37871850 - glial cell derived neurotrophic factor isoform 2 DNF (uc0031k2.1) at chr5:37851510-37871850 - glial cell derived neurotrophic factor isoform 2 DNF (uc0031k2.1) at chr5:37851510 - glial cell derive					
ZF1 (uc010gdb.1) at chr20:23290819-23301683 - GDNF-inducible zinc finger protein 1					

目的の遺伝子をゲノムブラウザ上で見る

ここでは、一番目のGDNF遺伝子と書かれたデータを見ていくことにしましょう。クリックしてください。



UCSCゲノムブラウザでは基本的に、画面下部で表示するように設定したトラックが上のブラウザに表示されるよう になっています。



画面上部には表示範囲を設定するコントローラーが配置されていて、上流/下流に移動したり拡大縮小したりする ことができます。

ビューアにはデフォルトの設定では、上から順にUCSCがアノテートした遺伝子(UCSC gene)、Refseq、GenBank内のmRNA, EST、multiz/phastconsを用いた比較ゲノム解析により保存されている領域、SNPs、Repeatの分布が表示されています。

ゲノムブラウザ内へのトラックの表示/非表示方法

UCSCゲノムブラウザ内で、ユーザが選択したトラックを表示/非表示させる方法を紹介します。

UCSCゲノムブラウザの下部には、表示するトラック選択のためのプルダウンメニューが並んでいます。

-	Mapping and	Sequencing Tra	icks	refresh
Base Position dense 💙	Chromosome Band hide	STS Markers	FISH Clones	Recomb Rate
Map Contigs hide	Assembly hide	Gap hide	Coverage hide	hide
hide	hide	hide	hide 💙	
*	Phenotype and	Disease Associ	ations	refresh
+	Genes and Ge	ne Prediction Tr	acks	refresh
•	mRNA a	nd EST Tracks		refresh
•	Ex	pression		refresh
•	Re	gulation		refresh
•	Compara	tive Genomics		refresh
•	Variatio	n and Repeats		refresh
•	ENCODE R	egions and Ger	ies	refresh
•	ENCODI	3 Transcription		refresh
•	ENCODE Chroma	tin Immunopreo	ripitation	refresh
•	ENCODE CI	hromatin Struct	ure	refresh
•	ENCODE Comparati	ve Genomics ar	d Variation	refresh

メニューは大きく13カテゴリーに分けられており、それぞれの左前にある+アイコンをクリックするとそのカテゴ リに含まれるトラックの一覧が表示されます。

例として、上から3つ目のカテゴリ"Genes and Gene Prediction tracks"(遺伝子および予測遺伝子)に含まれる トラックを見ていくことにしましょう。"Genes and Gene Prediction tracks"の左にある+のアイコンをクリッ クして、トラックを表示させてください。

-	Genes and G	Jene Prediction Tra	cks	refresh
UCSC Genes	Old UCSC Genes	Alt Events	CCDS	RefSeq Genes
pack		hide	hide	dense 💌
Other RefSeq	MGC Genes	ORFeome Clones	<u>TransMap</u>	Vega Genes
hide	pack		hide <mark>▼</mark>	hide
Ensembl Genes	AceView Genes	SIB Genes	N-SCAN	CONTRAST
hide		hide	hide	hide
SGP Genes	Geneid Genes	Genscan Genes	Exoniphy	Augustus
hide	hide		hide	hide
RNA Genes	ACEScan	<u>EvoFold</u>	sno/miRNA	Pos Sel Genes
	hide	hide <mark>❤</mark>	hide	hide

左上から横に順に、"UCSC Genes (UCSCが予測した遺伝子)","Old UCSC Genes (前のversionのUCSC予測遺伝 子)","Alt Events (UCSC遺伝子でのAlternativeスプライス情報)","CCDS (NCBI, Ensembl, UCSC遺伝子で共通 する遺伝子), "Refseq Genes (NCBI Refseqプロジェクトによるヒト遺伝子)"Other RefSeq (他生物種のRefSeq (をヒトゲノム上にマッピングしたもの))","MGC Genes (Mammalian Gene Collectionによる遺伝子情報)", "ORFeome (The ORFeome Collaborationからサンプルが提供されている遺伝子)","Trans Map (他脊椎動物由来の cDNAをマッピングしたもの)","Vega Genes (Sanger CentreのVegaプロジェクトでアノテートされた遺伝子)", "Ensembl Genes (Ensemblによりアノテートされた遺伝子)"などが選択できるようになっています。 各トラックを表示させたり、非表示にさせたりするには、そのトラックのプルダウンメニューから適切なモードを 選択し、最後に"Refresh"ボタンを押す必要があります。

Genes and Gene Prediction Tracks				
Old UCSC Genes hide MGC Genes pack AceView Genes hide	Alt Events hide ORFeome Clones hide SIB Genes hide	CCDS hide TransMap hide N-SCAN hide	RefSeq Genes full ✓ hide s dense squish pack ST full	

モードには5種類存在し、それぞれをRefSeqを使った例で説明します。

非表示

dense 一行に圧縮表示

RefSeq Genes

squish

低い高さに圧縮して全件表示(名前の記述なし)

RefSeq Genes

pack

全件を名前入りで表示 但し重ならないも	のは一行にまとめて表示
₩DR78 <mark>→++++++++++++++++++++++++++++++++++++</mark>	RefSeq Genes GDNF + GDNF +

full

全件を名前入り)で表示 各エントリにつき一行で表示
LIDE 74	RefSeq Genes
GDNF	here b
GDNF	
GDNF	<u>▶+++</u> ≱

これらの5種類のモードから表示方法を選択します。

■1-3. 遺伝子周辺のゲノム配列をUCSCゲノムブラウザからダウンロードする

遺伝子周辺のゲノム配列をダウンロードする

次に、GDNF周辺のゲノム配列をダウンロードしてみましょう。そのためには、まずダウンロードしたい遺伝子(転 写産物)上をマウスでクリックして、以下のようなページへと移動してください。この例はUCSC遺伝子の一番上に 表示されている遺伝子をクリックしたものです。

Human Gene GDNF (uc010ivb.1) Description and Page Index									
Description: glial cell derived neurotrophic factor isoform RefSeq Summary (NM_199234): This gene encodes a highly conserved neurotrophic factor. The recombinant form of this protein was shown to promote the survival and differentiation of dopaminergic neurons in culture, and was able to prevent apoptosis of motor neurons induced by axotomy. The encoded protein is processed to a mature secreted form that exists as a homodimer. The mature form of the protein is a ligand for the product of the RET (rearranged during transfection) protooncogene. In addition to the transcript encoding GDNF, two additional alternative transcripts encoding distinct proteins, referred to as astrocyte-derived trophic factors, have also been described. Mutations in this gene may be associated with Hirschsprung disease. Strand: - Genomic Size: 19146 Exon Count: 2 Coding Exon Count: 2									
Page Index	Sequence a	nd Links Genetic	Associations (TD	Microarray	Other Species			
mRNA Desc	riptions Other Nam	es Model I	information 1	dethods					
						·			
- Seguer	ce and Links to	Tools and De	tahases						
Seque	ice and Links to	TOOLS and Da	ttabases						
Genomic Se	quence (chr5:37,851	,510-37,870,655) mRNA (may	differ fro	om genome)	Protein (133 aa)			
Gene Sorter	Genome Browser	Protein FASTA	VisiGene	Table	Schema	CGAP			
Ensembl	Entrez Gene	ExonPrimer	GeneCards	Gene	lests .	Gepis Tissue			
HGNC	HuGE	Jackson Lab	OMIM	PubM	ed	Stanford SOUR	CE		
Treefam									
- Geneti	Genetic Association Studies of Complex Diseases and Disorders								
 □ Genetic Association Studies of Complex Diseases and Disorders Genetic Association Database: GDNF CDC HuGE Published Literature: GDNF Positive Disease Associations: schizophrenia Related Studies: 1. schizophrenia Michelato, A. et al. 2004, 3' UTR (AGG)n repeat of glial cell line-derived neurotrophic factor (GDNF) gene polymorphism in schizophrenia, Neuroscience letters. 2004 Mar,357(3):235-7. [PubMed <u>15003293</u>] These results support that the (AGG)(n) >/=15 alleles could be protective factors against schizophrenia and thus they suggest a possible involvement of GDNF gene in the genetic hability to illness. 2. schizophrenia 									

このページは、クリックした遺伝子(転写産物)に関しての詳細を記述したもので、配列情報や他のデータベース へのリンク、疾患との関連解析の情報、発現解析情報、他生物種でのオーソログ遺伝子情報、由来となった配列情 報など様々な情報から成り立っています。

その中で"Sequence and Links to Tools and Databases"と書かれた箇所から"Genomic Sequence"と書かれたリンク(下図の赤四角の箇所)をクリックしてください。

📃 Sequen					
Genomic Seq	Protein (133 aa)				
Gene Sorter	Genome Browser	Protein FASTA	VisiGene	Table Schema	CGAP
Ensembl	Entrez Gene	ExonPrimer	GeneCards	GeneTests	Gepis Tissue
HGNC	HuGE	Jackson Lab	OMIM	PubMed	Stanford SOURCE
Treefam					

下図に示したような遺伝子周辺のゲノム配列取得ページへと移動します。

Out Owner is Summer New Owner							
Get Genomic Sequence Near Gene							
Note: if you would prefer to get DNA for more than one feature of this track at a time, try the Table Browser using the output format sequence.							
Sequence Retrieval Region Options:							
Promoter/Upstream by 1000 bases 5' UTR Exons CDS Exons 3' UTR Exons I Introns							
Downstream by 1000 bases							
One FASTA record per gene. One FASTA record per region (exon, intron, etc.) with extra bases upstream (5') and extra downstream (3') Control of the second per region (exon, intron, etc.) with EASTA record per region (control of the second per reg							
Note: if a feature is close to the beginning or end of a chromosome and upstream/downstream bases are added, they may be							
truncated in order to avoid extending past the edge of the chromosome.							
Sequence Formatting Options:							
Exons in upper case, everything else in lower case.							
O CDS in upper case, UTR in lower case.							
O All upper case.							
O All lower case.							
Mask repeats: • to lower case • to N							
submit							

このページでは、遺伝子周辺のゲノム配列を取得するに当たっての条件を設定することが可能です。

まず赤く囲んだ領域で、配列取得の範囲を設定します。6つのチェックボックスから成っていて、チェックされた 領域をダウンロードすることができます。上から順に遺伝子上流配列(何塩基上流までかをテキスト領域に記述し ます)、5'UTRエクソン、CDS領域、3'UTRエクソン、イントロン、遺伝子下流配列(何塩基下流までかをテキスト 領域に記述します)です。

次に青く囲んだ領域で、データを一つのFASTA形式にまとめて出力するか、あるいは領域(CDSとかイントロンとか)ごとにマルチFASTAで出力するか(その際には各領域の上流下流に糊代を塩基数で設定できます)を選択します。また、領域ごとに出力する際にUTRエクソンとCDSエクソンとも分けるかをチェックボックスで指定します。

最後に緑四角で囲んだ領域で出力フォーマットを一つ選択します。設定し終わったらsubmitをクリックしてみましょう。

>hg18_knownGene_uc010ivb.1 range=chr5:37851510-37870655 ATGAAGTTATGGGATGTCGTGGCTGTTGCCTGGTGCTGCTCCACACCGC GTCCGCCTTCCCGct gccgcgcgcgt aasaggcctcccgaggcgcgcg aagaccgdtccctcgccgcgcgcggcggcggcggcggcggcggcggcg	5'pad=0	3'pad=0	strand=-	repeatMasking=none
ccaaaataaagcaggaatgagctgtgtgtaactaggataccaccctggct				
cgcttgaaacttactatggtcataaccacagtagtagtcctgcagtagtc				
aget gt geget gagt eecceaaacceaat ggeet t gget aggaaacaace				
cagtaaatgcaatgtgcttattttatgcttaattatagcaacaaaagcac				
t agait and of the again agait of the stage of the second at the second				
taaaggettttattteeteacaaatttaatgtatataaaggttteteea				

遺伝子周辺のゲノム配列が取得できました。

■1-4. クローン関連情報を見る

クローン情報を表示する

続いてUCSCゲノムブラウザ上でこのゲノム領域に対応したクローン情報を表示してみましょう。まずは、Genome Browserの画面まで戻ってください。



もう少し広範囲を見るために表示範囲を3倍ほど広げてみてください。それには画面上部のZoom Out から"x3"をクリックすることで実現できます。



では、この領域に位置するクローンを表示させて見ましょう。そのためには、下部の設定領域"Mapping and Sequencing Tracks"から適切なトラックを選択追加します。



"FISH Clones", "Assembly", "Bac END pairs"を"dense"に設定し、"refresh"ボタンを押して画面に反映させましょう。



上から順にFISHによりマッピングされたクローン情報、この領域のゲノムアセンブルに用いられた配列データのア クセッション番号、Bac-end情報からマッピングされたクローンが表示されています。(NCBI MapViewer にあった クローンの注文などの機能はありません。)

ocation of RP11-1022O21 using BAC end sequences RP11-1022O21								
Chromoso	me: chr5							
Start:	37849013							
End:	38046574							
Length:	197562							
Strand:								
Score:	1000							
Band:	5p13.2							
This is the Senomic	only location found for alignments of BAC	RP11-102202 ends:	21					
This is the Genomic AQ5924(81ZE IDENI	only location found for alignments of BAC 19	RP11-102202 ends:	END	QUERY	START	END	TOTAL	
This is the Genomic AQ5924(SIZE IDENI 372 94.	alignments of BAC	RP11-102202 ends: D START 37849013 3784	END 49403	QUERY 40592409	START 3	END 392	TOTAL 456	
Chis is the Genomic AQ5924(81ZE IDEN) 372 94. Jiew details	only location found for alignments of BAC 19 1TY CHROMOSOME STRAN 4% 5 + of parts of alignment wit	RP11-1022O2 ends: D START 37849013 3784	END 49403 1dow	QUERY AG592409	START 3	END 392	TOTAL 456	
'his is the 'enomic \Q59241 812E IDENI 372 94. 'lew details \Q70274 812E IDENI	only location found for alignments of BAC 19 11Y CHROMOSOME STRAN 4% 5 + of parts of alignment wit 18 11Y CHROMOSOME STRAN	RP11-1022O2 ends: D START <u>37849013 3784</u> hin browser win D START	END 49403 10000	OUERY AG592409 QUERY	START 3 START	END 392 END	TOTAL 456 TOTAL	

また、それぞれのクローンなどをクリックするとより詳細な情報が表示されます。

■1-5. 様々な転写産物データを表示する

様々な転写産物データを表示する

Genome Browserの画面まで戻ってください。

- Genes and Gene Prediction Tracks refresh							
UCSC Genes pack Other RefSeq hide Ensembl Genes hide SGP Genes hide RNA Genes	Old UCSC Genes hide MGC Genes pack AceView Genes hide Geneid Genes hide ACEScan	Alt Events hide ORFeome Clones hide SIB Genes hide Genscan Genes hide EvoFold	CCDS hide TransMap hide N-SCAN hide Exoniphy hide sno/miRNA	RefSeq Genes dense ✓ Vega Genes hide ✓ CONTRAST hide ✓ Augustus hide ✓ Pos Sel Genes			
hide 👻	hide 💙 mRN	hide A and EST Tracks	hide 🔽	hide refresh			
Human mRNAs dense V H-Inv hide V CGAP SAGE hide V	Spliced ESTs dense V UniGene hide V	Human ESTs hide	Other mRNAs hide SIB Alt-Splicing hide	Other ESTs hide V Poly(A) hide V			

転写産物の表示は下部の"Genes and Gene Prediction Tracks"および"mRNA and EST Tracks"からの設定で行います。

"Genes and Gene Prediction Tracks"については先ほど紹介しましたが、"mRNA and EST Tracks"からも、 mRNAや EST, UniGene, SAGEデータなど多様なデータが選択可能となっています。

例としてNscanとUnigene, Ensembl Geneを新たに表示させて見ましょう。



上から順に、UCSC遺伝子、RefSeq、Ensembl Gene、N-scan予測結果が表示されており、GDNF遺伝子の周辺を見るといずれも似ているものの、UTRなどが微妙に異なった遺伝子構造を持っていることがわかります。

このようにUCSCゲノムブラウザでは、NCBIから提供されているUnigeneやEnsembl遺伝子などの情報を表示すること も可能ですが、モデルメーカーのようなユーザがインタラクティブな操作で遺伝子候補を作成したりすることはで きません。

■1-6. 他生物種の遺伝子データを表示する

他生物種の遺伝子データを表示する(その1)

さらにこの画面上に他生物種の遺伝子データを表示してみましょう。

そのためにはヒトの場合と同様に、"Genes and Gene Prediction Tracks"および"mRNA and EST Tracks"からの設 定で表示させたいトラックを選択します。

Genes and Gene Prediction Tracks						
UCSC Genes	Old UCSC Genes	Alt Events	CCDS	RefSeq Genes		
pack 💌	hide 💙	hide 💙	hide 💙	dense 💌		
Other RefSeq	MGC Genes	ORFeome Clones	TransMap	Vega Genes		
hide 💙	pack 💌	hide 💙	hide 🔽	hide 💙		
Ensembl Genes	AceView Genes	SIB Genes	N-SCAN	CONTRAST		
hide 💙	hide 💌	hide 💌	hide 💌	hide 💌		
SGP Genes	Geneid Genes	Genscan Genes	Exoniphy	Augustus		
hide 🔽	hide 💌	hide 💌	hide 💌	hide 💌		
RNA Genes	ACEScan	EvoFold	sno/miRNA	Pos Sel Genes		
hide 💙	hide 💙	hide 💙	hide 💙	hide 💙		
•	mRN	A and EST Tracks		refresh		
Human mRNAs	Spliced ESTs	Human ESTs	Other mRNAs	Other ESTs		
dense 💌	dense 💌	hide 💌	hide 🔽	hide 🔽		
H-Inv	UniGene	Gene Bounds	SIB Alt-Splicing	Poly(A)		
hide 💙	hide 💙	hide 💙	hide 💙	hide 💙		
CGAP SAGE	San Sing	Sale Sale	Sala Sala	Sun Jan		

例として、"Other RefSeq"を表示させて見ましょう。



RefSeqの下にNon-Human RefSeq Genesが表示され、マウス、ラット、ゼブラフィッシュのRefSeqがマッピングされ ていることが確認できます。ヒトのGDNFと比較すると、この3種のRefSeqの中ではラットのものが一番遺伝子構造 としては似ていそうなことがわかります。

作成日: 2008年10月30日

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

