04_Ensembl

Genom böngészők: pl. UCSC(University of California Santa Cruz), NCBI (National Center for Biotechnology Information)

A *-al jelölt részek, csak definíciók, kiegészítések

Genomböngészés ensembl-el www.ensemblgenomes.org

Az Ensembl egy közös project az EBI (European Bioinformatics Institute) és a Wellcome Trust Sanger Institute részvételével 1999 –ban indult, itt történt az első humán draft genom bemutatása, és azóta is céluként szerepel a humán gének annotációja gén modellek alapján. További céluk a gerinchúros és gerinces élőlények (ma ár több mint 100 faj) genomjának annotálása, olyan élőlényeké amelyek fontosaka mezőgazdaságnak, orvostudománynak, veszélyeztettek, vagy evolúciós jeltőséggel bírnak illetve model állatok (Drosophila, C.elegans, S. cerevisiae)Nagy hangsúlyt fektettenek a humán, egér, és zebradánió fajokra Ezenfelül élesztő és féreg gén szettek is bekerültek ebbe böngészőbe összehasonlító analízis céljából az Ensembl compara csoportnak köszönhetően. A legtöbb annotációt kéthavonta frissítik, ez egyre frisseb Ensembl verzió megjelenést eredményez. További céljuk az automatikus ámde mégis pontos génannotációt. Elérhető genomok, Összefűzni az annotációkat az elérhető biológiai adatokkal. web és program alapú interface. Az Ensembl-ön elérhetőek gének, és egyéb annotációk, mint pl. regulátor régiók, konzervált bázispárok a fajok között és különböző szekvencia variációk. Az ensembl-nek van egy testvér browsere: at www.ensemblgenomes.org, ahová nem chordata élőlények gerinchúr nélküliek kerülnek. (baktériumok, növények, metazoa-k, egysejtűek).

* A genom annotáció nem más, mint biológiai információ hozzárendelése egy DNS szekvenciához (általában a génekre értik). A biológiai információ lehet strukturális (pl. kromoszómán való pontos elhelyezkedésének megállapítása), ill. funkcionális(például részvétel a szabályozásban, kölcsönhatásokban). (http://web.med.u-szeged.hu/mdbio/hun/anyagok/2013-2014/I.felev/smge/6/1.A-Genom-Szerkezete-word.pdf)

A core (alap) adatok:

- Genom szekvenciák
- Gén / transzkript / protein modelek
- Külső referenciák
- Feltérképezett cDNS-ek, proteinek, microarray probák, BAC klónok, kromószóma sáv térképek (cytogenetic band) repeat-ek, markerek etc.
- Komparatív adatok:
- orthologok and paralogok, fehérje családok, teljes genom illesztések, szinténikus régiók, Variant adatok:
- szekvenvencia variációk, strukturális variációk, linkage disequilibrium Regulációs adatok
- "best guess" (legjobbnak becsült)regulációs elemek sorozata

linkek más adatbázisokhoz(directly or via Distributed Annotated System DAS)

Az Ensembl gén model-ek alapjául az ENA/ GenBank /DDBJ (International Nucleotide Sequence Database Collaboration) INSDC , cDNS, EST,RNAseq az UniProtKB : Swiss-Prot(manuálisan javított),

TrEMBL (fehérje adatbázisok)és NCBI RefSeq adatbázisokban felelhető protein és mRNS szekvenciák, szolgáltak melyek transzkriptjeit (és még EST RNA-seq, cDNS, protein etc.,)illesztették aztán klaszterezték úgy, hogy az átfedő kódoló szekvenciák segítségével pontosan meg lehessen adni az Ensembl gének szekvenciáit és helyzetét. Automatikus annotáció: Genome-wide (Teljes genom)meghatározás Ensembl génalapú pipeline (parancssorozatok) felhasználásával történt. *Ab intio* gén predikció zajlott le. Ezenfelül manuálisan annotált géneket vettek át a VEGA/Havana csoporttól, majd az automatikusan és a manuálisan annotált géneket összemergelték, összvetteék egymással és így jött létre a Genecode geneset. Adatokat importáltak még a FlyBase, WormBase, és yeast (SGD) adatbázisokból is.

*Az intrinsic (ab initio, de novo) gén predikciós megközelítések a géneket a fehérjekódoló génekre jellemző tulajdonságok felismerése alapján azonosítják a genom szekvenciában. Például az annotációs pipeline megkeresi a promoter régióra, transzkripció start helyre, exon/intron határra, poliadenilációs helyre stb. utaló speciális szekvencia jeleket vagy felismeri a fehérjekódoló gének jellegzetes statisztikus a nem-kódoló részektől különböző tulajdonságait. (http://math.bme.hu/~gnagy/mmsz/eloadasok/PatthyLaszlo2012.pdf)

* <u>Vertebrate Genome Annotation</u> (VEGA), - <u>Wellcome Trust Sanger Institute</u>, A Havana csoport a consensus-coding sequence (CCDS) eggyüttműködés része, illetve a teljes emberi genom kézi annotációjáért felelős ENCODE project kiterjesztése, ami így elérhető referenciaka genomként illetve komparatív analízisekhez.

Gén annotácios folyamat:

• Targeted build: (cél alapú):

Fajspecifikus proteineket és cDNS-ket illesztenek a genomra, hogy iniciális transzkript modelleket hozzanak létre.

• Similarity build (Azonosság alapú):

Nagyon közelrokon fehérjéit és cDNS-t illesztik fel, hogy annotálhassák a géneket. Az ismeretlen transzkriptumokat nem térképezik a fajspecifikus proteinekhez. Egybevetik a redundáns transzkriptumokat és klaszterekbe rendezve az alternatív transzkriptumokat hozzáteszik a megfelelő génekhez.

A CCDS-ek (core set of protein coding regions) –magas minőségű meghatározás, embernél és egérnél

*Genome assembly: read-> contig-> scaffold-> chromosome

* A referencia genom a haploid genom (embernél a 24-féle kromoszóma + a mitokondrium DNS) pszeudoszekvenciáit tartalmazó konszenzus szekvenciák összessége, ami önmagában ilyen formában nem létezik és nem is létezett. A célja az, hogy a többi szekvenciát ehhez viszonyítsuk. A referencia genomszekvenciákat folyamatosan újraépítik, ami azt jelentheti, hogy a szekvencia koordináták elcsúszhatnak, ezért vigyázni kell, hogy ha nem szekvencia, hanem koordináta alapon keresünk például két szekvencia között átfedést, akkor ugyanazt az összerakott genomot használjuk (pl. hg18 vs hg18 vagy mm8 vs mm8). A referencia genomszekvenciákat a nagy genomikai portálokról lehet (érdemes) letölteni. Javaslom az ensembl, az NCBI vagy az UCSC webhelyét. (http://genomika.blog.hu/2010/04/23/mi_az_a_referencia_genom)

Elérhetőek variációs adatok, Összehasonlítóakegymással genomok, és a regulatórikus régiók tulajdonságait is felderíthetjük. **Genom szekvencia assembly-k is hozzáférhetőek pl** -Mouse NCBI m37 (mm9)-Zebrafish Zv9 –Human GRCh37 (hg19). folyamatosan frissül, de a korábbi verziók is megtalálhatóak a preensembl-benaz alternatív haplotípusok is megtekinthetőek.

Minden adat ingyen hozzáférhető online <u>www.ensembl.org</u>. A PERL programnyelvet használók pedig közvetlenül hozzátudnak férni az ENSEMBL adatbázisokhoz az Application Programming Interfaces (**Perl APIs**) (felhasználói felület))keresztül. A gén szekvenciák letölthetőek az ENSEMBL browserről közvetelenül vagy a BioMart web felhasználói felület-en keresztül, ami lehetővé teszi, hogy programozói ismeretek nélkül , egyszerű felhasználóként információkat szerezzünk az Ensembl adatbázisairól.

Mire jó az Ensembl?

- Megnézheted a gének különböző annotációi a kromoszómán
- Megtekintheted az adott gén alternatív transzkriptumaii (azaz splice variánsai).
- Megkeresheted a több mint 100 faj (120-25) bármely génjének homológjai, amelyek alapján filogenetikus fákat lehet készíteni
- Teljes genom alignmenteket és konzervált régiókat hasonlíthatsz össze a fajok között.
- Láthatod az Ensembl génekhez illeszkedő mikroarray-szekvenciákat
- Megnézheted meg EST-eket, klónokat, mRNS-eket és fehérjéket bármilyen kromoszomális régión.
- *klón genetikilag megegyezik az anyaszervezettel, EST: Expressed sequence tag: A CDNS darabjai.
- Megtalálhatod a nukleotid polimorfizmusokat (SNP) a géneken vagy kromoszómális régiókon.
- Láthatod az SNP-ket a fajok (patkány, egér), a populációk (emberi) vagy fajták (kutyák) között.
- Megnézheted az Ensembl génekhez illeszkedő mRNS-ek és fehérjék pozícióját és szekvenciáját.
- Feltöltheted a saját adatait.
- Használhatod a BLAST-ot vagy a BLAT-ot bármelyik Ensembl genommal szemben.

• Exportálhatod a szekvenciákat, vagy létrehozhatsz egy táblázatokat a génadatokról a BioMart segítségével.

• Láthatod az a saját génvariánsaid milyen hatással vannak a génekra és a transzkriptumokra a Variant Effect Predictor segítségével.

Ensem	blast/blat BioMart VEP	Tools Downloads Help & Docs	Blog	Login/Ro
,				
Tools All tools	BioMart > Export custom datasets from Ensembl with this data-mining tool	BLAST/BLAT > Search our genomes for your DNA or protein sequence	Variant Effect Predictor > Analyse your own variants and predict the functional consequences of known and unknown variants	Ensembl is a genome browser for vertebrate genomes that support research in comparative genomics, evolution, sequence variation in transcriptional regulation. Ensembl annotate genes, computes multip alignments, predicts regulatory function and colets: disease data. Ensembl tools include BLAST, BLAT, BioMart and the Variant Effect Predictor (VEP) for all supported species. Ensembl Release 92 (April 2018)
				New goat annotation on the ARS1 assembly
	Search			 Update of Marmoset assembly and genebuild
	All species	▼ for		Mouse: update to Ensembl-Havana GENCODE gene set
			Go	Update to Ensembl-Havana human GENCODE gene set (release 28)
	e.g. BRCA2 or rat 5:62	797383-63627669 or rs699 or coronary	heart disease	New command line tool for LD
				Full details All web updates, by release More news on our
ll genomes		Eavourite genomes		• 10 Apr 2018: Do you use transcripts for your work?
il genomes		Tavounte genomes		 05 Apr 2018: Ensembl 92 has been released! Ø
 Select a specie: 	·s · · ·	GRCb38 p12		 23 Mar 2018: <u>2018 – a year of conferences</u>
 View full list of 	all Ensemblisheries	Greise.prz		Go to Ensembleb
- Editoria factor	wites	Still using GRCh37?		2

A felső kék csíkban linkeket mutat számos különbző ensembl alkalmazáshoz. felsorolás mik vannak

A kereső rész search bar lehetőséget ad hogy az ensembl összes fajában keressünk genom koordináták, gén nevek, próba szettek, variant ID-k, fenotípusok, és regulatórikus jellemzők ID –je (azonosítója)alapján.

Fent találhatóak az újdonságok, 2-3 hónapos frissülés miatt, a változások az annotációkban.

Összes genom elérhető:

<u>View full list of all Ensembl species</u>

és a genomok régi verziója is a preensembl-n, a keresőben a gén nevek a HGNC azonosító alapján a HUGO gene nomenclatura committee által jóváhagyott formában kereshetőek pl BRCA2

Nézzük meg a humán genom honlapját klikk a humán -ra GRCh38 -as assembly-re

	loce Blog	A source all species	Login/Register
Human (GRCh38.p12) ▼	loco blog		
Search Human (Homo sapiens) Search all categories Search Human Go e.g. BRCA2 or 17:63992802-64038237 or rs1333049 or osteoarthritis		What's New in Human release 92 Update to Ensembl-Havana human GENCODE gene set (release 28) Updated human otherfeatures db: New CCDS import HGMD-Public dataset	
Genome assembly: GRCh38.p12 (GCA_000001405.27) More information and statistics Download DNA sequence (FASTA) Convert your data to GRCh38 coordinates Display your data in Ensembl Other assemblies GRCh37 Full Feb 2014 archive with BLAST, VEP and BioMart Go	karyotype	Gene annotation What can I find? Protein-coding and non-coding genes, splice variants, cDNA and protein sequences, non-coding RNAs. Image: Comparison of the sequences of the seque	More news Pax6 INS DBRCA2 D
Comparative genomics	-C 766	Variation	ATCGAGCT

Info a genom assembly-ről, a komparatív genomról, a gén annotációról, és a varációkról.

BRCA2 keresőbe GO

Rew Search Jobs	nbl blas	T/BLAT BioMart VEP Tools Downloads Help & Docs Blog	Log 🛃 🕶 Search all species
Current selection: < all Species Only searching Hu	uman	Only searching Human BRCA2	
Restrict category to:		<u>Did you mean ▼</u>	
Gene	9	BRCA2 (Human Gene)	
Transcript	47	ENSG00000139618 13:32315474-32400266-1 BRCA2 DNA repair associated [Source: HGNC: Symbol: Acc: HGNC: 1101]	
Phenotype	6	LRG_293 (LRG display in Ensembl gene record; description: Locus Reference Genomic record for	
Somatic Mutation	1837	BRCA2,) is an external reference matched to Gene ENSG00000139618	
GeneTree	1	Variant table • Prenotypes • Location • External Refs. • Regulation • Orthologues • Gene free	
ProbeFeature	75	BRCA2-203 (Human Transcript) ENST00000528762 13:32370971-32379495:1	
Clones & Regions	1	BRCA2, DNA repair associated [Source:HGNC Symbol;Acc:HGNC:1101]. Location • External Refs. • cDNA seq. • Exons • Variant table • Protein seq. • Population • Protein summary	
Protein Domain	6	BRCA2-204 (Human Transcript)	
Protein Family	2	ENST00000530893 13:32315505-32333291:1 BRCA2_DNA renair associated (Source: HGNC Symbol: Acc: HGNC: 1101)	
Variant	17490	Location * External Refs. * cDNA seq. * Exons * Variant table * Population	
Per page:		BRCA2-207 (Human Transcript) ENST00000614259 13:32316461-32362659:1 BRCA2, DNA repair associated [Source:HGNC Symbol;Acc:HGNC:1101].	

Kattintsatok az első linkre



Gén azonosító Gene:BRCA2 ENSG00000139618 ez minden ensembl release, újabb genom assemblyben meg fog egyezni!

Gén nevek az Ensembl-ben:

ENS (ENSEMBL)MUS (fajra jellemző, humánnál nincsen!)G/T/P/E /R(gén,transzkript,protein,exon, regulációs sajátság)11 számjegy

Fenn látható 3 fül amelyekből mi most a gene-ben vagyunk, baloldalon e génre vonatkozó speciális információk találhatóak. pl. gén szekvencia, szekvencia alignmentek, philogenetikei törzsfák, homológiák, gén ontológia, variációs és gén experessziós adatok, itt hét annotált transzkript látható. Show transcript table

Location: 13:32,315,474-32,400,266	Gene: BRCA	2 Jobs 🔻									
Gene-based displays Burnmary Gene: BRCA2 ENSG00000139618											
Spilce variants Transcript comparison Gene alleles Sequence Sequence Comparative Genomics Comparative Genomics Genomic alignments Gene tree Gene gainloss tree Orthologues	Description Synonyms Location About this ge Transcripts	ane	BF BF GF Th	CA2, DNA r CC2, FACD romosome 1 RCh38:CM00 is gene has Hide transcri	epair associated [Source:HGM , FAD, FAD1, FANCD, FANCD <u>3: 32,315,474-32,400,266</u> for 0675.2 7 transcripts (<u>splice variants</u>), pt table	IC Symbol;Acc: 11, XRCC11 ward strand. 89 orthologues	HGNC:1101	외 r of <u>1 Ensembl pr</u>	<u>otein fami</u>	ly and is associate	ed with <u>100 phe</u>
Ensembl protein families Ontologies GO: Biological process	Show/hide	columns (1 hidden)	he i	Proto la c	Plature	CODE	Un IDent	Deffer		Filter	
 GO: Molecular function GO: Cellular component 	BRCA2-201	ENST00000380152.7	11986	3418aa	Protein coding	CCDS9344@	P51587	-	TSL:5	GENCODE basic	APPRIS P1
Phenotypes Genetic Variation	BRCA2-206	ENST00000544455.5	10984	<u>3418aa</u>	Protein coding	<u>CCDS9344</u> 교	<u>P51587</u> @	<u>NM_000059</u> 관 NP_000050 <i>@</i>	TSL:1	GENCODE basic	APPRIS P1
 Variant table Variant image 	BRCA2-202	ENST00000470094.1	842	<u>186aa</u>	Nonsense mediated decay	-	<u>H0YE37</u> ₽	-	C	DS 5' incomplete	TSL:5
Structural variants	BRCA2-203	ENST0000528762.1	495	<u>64aa</u>	Nonsense mediated decay	-	<u>H0YD86</u> ¢P	-	C	DS 5' incomplete	TSL:4
 Gene expression Pathway 	BRCA2-207	ENST0000614259.1	7950	No protein	Processed transcript	-	-	-		TSL:2	
- Regulation	BRCA2-204	ENST00000530893.6	2011	No protein	Processed transcript	-	-	-		TSL:1	
 External references Supporting evidence 	BRCA2-205	ENST00000533776.1	523	No protein	Retained intron	-	-	-		TSL:3	
ID History Gene history	Summar	Summary @									

Az első két találat a CCDS consensus-coding sequence szett tagja vagyis ezek olyan protein kódoló transzkriptumok melyek más genom browserekkel (NCBI, UCSC, HAVANA) való együttműködés révén határoztak meg. Established as collaborative effort A sárga a HAVANA génekkel összevetett verziót mutattja. A kék vonal jelöli a kontigot, ami egy genome szekvencia. A felette lévő trk a vezető forward szálon vannak, az alatta lévőek a maradó szálon lévő reverz transzkriptek. A boxok téglalapok exonokat, az ezeket összekötö vonalak intronokat jelölnek. Ha a boxok színezettek akkor kódoló szekveniát jelölnek ha nem akkor untranszlated régiókat. A transzkriptumok színe a biológiai típusokra utal a sárga az ensembl automatikus annotációs pipeline-al annotált de a havana project kézzel annotált génjeivel összvetett, összefésült transzkriptumot jelöli. A piros csak automatikusan annotált ensembl-ös. A kékek nem kódolóak. Kétféle módon nézhetjük meg a transzkriptet: vagy a Transcript ID alapján vagy

Go to Reg	jion in Detail f	for more tracks and nav	igation options (e.g. zoomi	ng)		
🌣 🖪 < 🖽 🖬 🎭						Drag/Select: 😝 📋
	L			104.79 kb		Forward strand 📂
		32.32Mb	32.34Mb	32.36Mb	32.38Mb	32.40Mb
pap				No features from pa	ap on this strand	
Genes (Comprehensive set		BRCA2-201 >				
Comprenensive acc.		protein coding				
		BRCA2-206				
		protein coding				
		processed transcript			nonsense mediated decay	retained intron
		processed transcript			BRCA2-202 > nonsense me	diated decay
					IFIII	.P1-201 > essed pseudogene
Contigs				AL445212.9 >		AL137247.14 >
Genes		-M				
(Comprehensive set	< ZAR1L-202 protein codina					< N4BP2L1-205 nonsense media
	< ZAR1L-201					< N4BP2L1-211
	protein county					
						< N4BP2L1-20
						honsense med
						< N4BP
						retained
						< N4BP2L1-204
8						protein coding

HGNC Syn	nbol: BRCA2
Gene	BRCA2, DNA repair associated
	ENSG00000139618
Location	Chromosome 13: 32,315,474-32,400,266
Exon	11 of 27
Transcript	BRCA2-201
	ENST00000380152.7
	Exons
	cDNA Sequence
Protein	EN SP00000369497
	Protein Variations
Gene type	Protein coding
Transcript type	Protein coding
Strand	Forward
Base pairs	11,986
Amino acids	3,418
Source	Ensembl/Havana merge

vagy úgy hogy rákkantintunk a képre és megjelennek róla az információk. Itt azazonosítót választjuk. kinyílik egy újabb fül és láthatjuk a trk jellemzőeit

Summary 🛛		
BRCA2-201,>	B4.79 kb	
protein coding		
Statistics	Exons: 27, Coding exons: 26, Transcript length: 11,986 bps, Translation length: 3,418 residues	
CCDS	This transcript is a member of the Human CCDS set: CCDS9344@	
Uniprot	This transcript corresponds to the following Uniprot identifiers: P51587 @	
Transcript Support Level (TSL)	TSL:5	
Version	ENST0000380152.7	
Туре	Protein coding	
Annotation Method	Transcript where the Ensembl genebuild transcript and the Vega @ manual annotation have the same sequence, for every base pair. See article.	
GENCODE basic gene	This transcript is a member of the <u>Gencode basic</u> gene set.	
Ensembl release 92 - April 2018 © EMBL-EBI	Permanent link - View in archive	site

ha külső general identifiers re kattintunk akkor láthatjuk e transzkriptnek a találatait más külső adatbázisokban pl UCSC, de megnézhetjük az oligipróbáit, vagy a protein summeryben a proteinjeire jellemző tulajdonságokat,

Most nézzük meg e gén melletti szekvenciákat úgy hogy rámegyünk a Location fülre.

CEnsembl 🖬	AST/BLAT BioMart Too	ls Downloads Help & Docu	mentation Biog Mirrors		🚮 • Search of species	Login/Register
Human (GRCh38.p10	») ▼					
Location: 13:32,315,434-32,400,2	66 Gene BRCA2 Tran	script: BRCA2-201 Jobs +				
ocation-based displays Whole genome Chromotome summary	Chromosome 1	3: 32,315,474-32,400,	266			
Region overview Region in detail	♦ < <					
Comparative Genomics - Syntany - Alignments (image) - Alignments (text)	Assembly exceptions Ohr. 13 Assembly exceptions		C 11151(214.11	6143 C415	C33.4.	
Genetic Variation Markets Markets	Region in detai	0				
Other genome browsers	< 3 2 3 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	e		Scrol: <	Track height: I C C	Drag/Select: ++ []
NCBI# Ensembl GRCh37#	Chromosome bands	52.00 Ma	52.20 MB	12.40 Mg	32.90 No	DUBU NO
Configure this page	Contigs Genes	A1371.04>	AL138002.26 >	AL445152.8 >	Z6440711 5 Z754061	> AL13HIDG 11
Custom tracks	(Comprehensive set from GENCODE 26)	FIDPS> LINCO1073 > <fryasi PRV ></fryasi 	8		7247.1 < N469P2L2-IT2 ATE PC0006 > ATP8A2P2.>	< 79/1/18/4 AL 13/08/20.3 >
🛃 Export data		< ACCORSES 1	_	< 140	PER	
< Share this cape	Gene Legend	32.00 Mb	32.20.946	12.40 Mb	32.60 Mb	32.80 Mb
P Bookmark this page		pseudogene		Processing the INVA game	1021	
	Location: 13:323154	4-32400266	00		contraction and a second	
	Gene:		Ga			
	0 A < 1 2 0	5				Drag/Select ++ 11

Felül a kromoszóma látható a piros négyzetek a gén körüli régiókat mutatják, amit kinagyítva látunk a következő képen, ahol 1 Mb nagyságban látszik a gén környzete, a szomszédos gének , a kontigok világosak és sötét kéket, hogy el lehessen különíteni őket. a 3. nagy kép részlet zoom-olható, kifejezetten a génre fókuszál, itt olyan plusz információk is megjelennek mint pl e régióra jellemző fenotípus ezek ael is tüntethetőek a Configure this page opcióval, pl dbSNP variant normal, uniprot keresőbe, nézzünk információt és írjuk be az emlőst

C 1 www.ensembl.org/Homo_sapie	ns/Location/View?db=core;g=EN	VSG00000139618;r=13:3	2315474-32400266;1	t=ENST00000380152				☆
谷 副 < 回	■ % ³						Drag/Selec	:t ↔
				84.79 kb			Forward	l stran
Chromosome t	ands 32.32Mb 32	2.33Mb 32.34Mb	32.35Mb	32.36Mb q13.1	32.37Mb	32.38Mb	32.39Mb	
pap				No features from	pap on this strand			
- CR	Constrained elements for 68	8 eutherian mammals EPO_	LOW_COVERAGE	1.1.1.1.11	11	1.1.1		1
elements								
Proteins (marr	mal)							<u> </u>
from UniProt								
28 A.	Showing 8 of 14 features, d	lue to track being limited to	6 rows by default - clic	k to show more				
Human cDNAs								-
(RefSeg/ENA)					M			_
CCDS set	CCDS9344.1 > protein coding							1
Genes (Comprehensie	e set BRCA2-201 >				n	1 11		-
51	h h		-+					-
	protein coding							
	BRCA2-204 >	+++0			BRCA2-2		TT1P1-201 >	
	processed transcript				nonsens	e mediated decay pr	ocessed pseudogen	e
	BRCA2-207 >					BRCA2-202	>	
2 ²	processed transcript					nonsense m	ediated decay	P
								BF
Contigs				AL445212.9 >				
dbSNP SNPs/ir	dels dbSNP - short variants (SN	IPs and indels)						
						· · · · · · · · · · · · · · · · · · ·	·····	.
All phenotype-	assoc All phenotype-associated -	short variants (SNPs and in	dels)	an a constant and			ann an	1.00
2	SV - 1000 Genomes 3 - All							
SV - 1KG 3 - A	11							
Regulatory Bui	ld.							



Információk a variant-ról:

*Frameshift kereteltolódás: blokk- értelmezése deléció, hol van

Nem okoznak változást a kódolt polipeptid aminosavsorrendjében. Ezek az úgynevezett szinoním (synonymous) vagy más néven csendes (silent) mutációk.

Ha egy pontmutáció eredményeképpen a kódolt aminosav minősége megváltozik, akkor misszensz mutációról beszélünk. aminosav csere

Előfordul, hogy egy aminosavat kódoló kodon a transzlációt termináló, úgynevezett STOP kodonra változik, ezáltal a fehérje transzlációja idő előtt befejeződik. Az ilyen mutációkat nevezzük nonszensz (nonsense , értelmetlen) mutációknak.

Regulációs feature információk:

Regula	tory I	Feature					
Stable	ID	EN SR0	0000060	894			
Туре		Promoter					
Core b	р	13:3231	5000-32	316601			
		2 Q					
Round	r hn	42-2224	4402.23	247400			
Bound	s nh	20	4402-52	317133			
		- 4					
Motif Ir	form	ation					
Name	PWM	I ID	Score				
Gabp	PB0	<u>020.1</u> දුව	8.314				
Gabp	MAG	<u>062.2</u> ද්ට	9.552				
Gabp	MAG	<u>062.1</u> ද්ට	9.503				
USF1	MAG	<u>093.2</u> ද්ව	9.831				
USF1	MAG	<u>093.2</u> ද්ව	9.235				
USF1	MAG	<u>281.1</u> 값	8.184				
E2F4	MAG	<u>)541.1</u> द्व	5.94				
E2F4	MAG	<u>541.1</u> 값	8.994				
E2F6	MAG	471.1 🖓	9.773				
PU1	PB0	<u>058.1</u> 군	9.023				
Pax5	MAG	014.2 🔊	8 9 1 4				

A region detail segít abban hogy mi is van az oldalon.

Exon intron:

Transcriptra kattintani majd az exonra, flanking zöld intron kicsi kék, lila UTR, egyben az intron configure this page show full introns.

Exons/ Introns	Translated sequen	ce Flanking	sequence Intr	on sequen	ce UTR			
Variants	Coding sequence	Frameshift	Inframe deletio	n Inframe	e insertion	Missense	Protein	n altering variant
	Splice acceptor	Splice donor	Splice region	Start lost	Start retain	ed Stop g	gained	Synonymous
Markup	loaded							

Show	All v entries			Show/hide co	olumns		Filter
No.	Exon / Intron	Start	End	Start Phase	End Phase	Length	Sequence
	5' upstream sequence						cgggtggtgcgtgtgctgcgtgtcgcgtcacggcgtcacgtggccagcgc
1	ENSE00001184784	32,315,474	32,315,667	-	-	194	GGCTTTGGCCCEACTICTAAACTACECGCACEGGCGCGCCCCCCCACAACTCT GCCCTTCCTECGCCCCCGGTCTTTTTGCCCCCGCGCGCGC
	Intron 1-2	32,315,668	32,316,421			754	gttagtggtggtggtagtgggttggtgcattttggtcttctgttttgcag
2	ENSE00001484009	32,316,422	32,316,527	-	1	106	<mark>a</mark> ctta <mark>ii</mark> tac <mark>oarceat</mark> tgegegata <u>icet</u> acgt <mark>aaaate</mark> cc <mark>ta</mark> tt <mark>ec</mark> atc <mark>oaar</mark> eg Acc <mark>caace</mark> tti <mark>tuisaartuiina</mark> s <mark>cacescegegaacaar</mark> eca <mark>c</mark>
	Intron 2-3	32,316,528	32,319,076			2,549	gtattgacaaattttatataactttaaggtgggatttttttttaaatag
3	ENSE00003666217	32,319,077	32,319,325	1	1	249	ATT RECEACED THE ASTOT TARTIES TO SAME ACTITICS TO AGAINST CACCE TARA ATTOTICAD STOCE SMACANT STEAD & TABANCH OF A TARACH O ATTACCH O AN ACCT THIN A AG TO SEC A SACCANCE TO TABAN SACCACE TO A ACCE O ATTACT ATT TO AGA ACCENCES TRACT TO SECOND STATUS AT SACCACE OF TO A ACCE ATTACT AND AND

Kattinsd végig exon cDNS

Adj hozzá saját adatot

Most egy nagyon egyszerű megoldás

transzkript re menj rá, aztán custom tracks –ra add data sokféle formátumú kézellbeírt, vagy fájlként feltöltött adatot(bam, gff, bed) betölthet pl. BRCA2 –t megkeres a UCSC-ben a table browser bed – ként és egyszerűen bemásolod az ensembl- configure page -add custom track-jébe (Vigyázat assembly egyezésre!)

Custom tracks	Add a custom track							
Track Hub Registry Search								
Manage Configurations	Please note that track hubs and indexed files (BAM, BigBed, etc) do not work with certain cloud services, including Google Drive and Dropbox Please see our support page for more information.							
	Name for this data (optional):	UCSC						
	Species:	Human (Homo sapiens) Assembly: GRCh38						
	Data:	0 + <u>chr13</u> 32394688 32394933 <u>uc658wfx.1 cds</u> 2_0_ <u>chr13</u> 32394689_f 0 + <u>chr13</u> 32395990 32396092 <u>uc658wfx.1 cds</u> 3_0_ <u>chr13</u> 32395991_f 0 +						
		Or upload file (max 20MB) Fájl kiválasztása Nincs fájl kiválasztva						
	Data format:	BED •						
		Help on supported formats, display types, etc						
		Add data						

Blast:/blat:

PDF to DOCX - Convert ×	Tutorials × e! BLAST/BLAT search - H	Ho 🗙 🕼 Genomic Sequence Nea 🗙 🌾 https://genome-euro.uc 🗴 🔽 Mónika show-Szilvi felp 🗴 💭 🗐 Nóra 🗕	0	×
← → C ☆ ③ www.ensembl	lorg/Homo_sapiens/Tools/Blast?db=core;r=13:18	705047-41742466	7 💦	0
Location: 13:18,705,047-41,742,466	BLAST/BLAT 🔻			-
Web Tools Web Tools BLAST/BLAT	BLAST/BLAT search @			
 Variant Effect Predictor Linkage Diseguilibrium Calculato 	BLAST/BLAT for Human GRCh37			
 File Chameleon Assembly Converter ID History Converter 	If you are looking for BLAST/BLAT for Human GRO	Ch37, please go to <u>GRCh37 website</u> tP.		J
VCF to PED Converter Data Slicer	Sequence data:	Maximum of 30 sequences (type in plain text, FASTA or sequence ID)		1
Configure this page				1
		Or upload sequence file Fájl kiválasztása Nincs fájl kiválasztva		
< Share this page				
Bookmark this page		• DNA		
		O Protein		
	Search against:	Homo_saplens X		ł
		Add/remove species		
		DNA database Genomic sequence		
		Protein database Proteins (GENCODE/Ensembl)		
	Search tool:	BLAT		
	Description (optional):			
🛋 🎒 💾 🚞		🕜 스 4 🕅 📶 👘 HUN	3:56 2018.04.1	16.

Például UCSC kivesszük BRCA2-t protein szekvenciát! (de ugyanígy blasztolhatsz saját szekvenciát vagy például az NCBI-ból, ENA-ból sstb.-ből szekvenciákat.)

Állítsuk át proteinre, állítsuk be a chimpanze-t referencia szekvenciaként, mindent kibontunk de alapbeállításon hagyunk tblastn –el futtatunk. megtudjuk nézni a szekvenciát lokációt, de igazán informatív az ha az alignmenteket nézzük meg. A result table-ben a query vagyis a kereső, blasztolandó szekvencia a humán fasta. A reference chimpanze. Töltsük le download result file. HSP distribution on query sequence.

*A BLAST futás eredményeként olyan találatokat kapunk, amelyek az adatbázisban tárolt szekvenciák közül szignifikáns hasonlóságot mutatnak a célszekvenciával. A program sorba állítja ezeket a szekvenciapárokat, kezdve a legnagyobb hasonlóságot mutatóval. A szignifikanciát egy E-vel jelölt, a véletlen hasonlóság mértékéhez viszonyított várható érték (*expectation*) jelzi, valamint egy "score" érték, ami az azonos, hasonló és "rés" (gap) pozíciókat számolja egy nukleotid vagy aminosav hasonlósági mátrix alapján. Ezek közül leggyakrabban az ún. BLOSUM (BLOcks SUbstitution Matrix) mátrixokat használják. Ha E < 0,01, akkor a két szekvencia minden bizonnyal homológ (azaz evolúciós rokonságot mutat, közös ősből származik). **BLAST (Basic Local Alignment Search Tool)** : HSP-k (high scoring pair) keres kibővít.

Program	Kereső szekvencia	Célszekvencia
BLASTN	Nukleotid	Nukleotid
BLASTP	Protein	Protein
BLASTX	Nukleotid, 6 leolvasási keretben	Protein
TBLASTN	Protein	Nukleotid, 6 leolvasási keretben
TBLASTX	Nukleotid, 6 leolvasási keretben	Nukleotid, 6 leolvasási keretben

(http://elte.prompt.hu/sites/default/files/tananyagok/BevBiokemGyak/ch11s03.html)

Adatok exportálása:

BRCA2 transcript, export data CSV formátumban, és FASTA formátumban, bekattintani amit lehet

Az BRCA2 gén összehasonlítása emberben és más fajokban.

A gén

-Genomic alignment- select on alignment kiválasztjuk az állatokat! block-ot 1 configure, az állatok passzoló régiói tkiírva chr start end, orientáció, alignments text ,szekvencia :piros exont jelent, vonal nincs semmi, konzervált régiók aláhúzása: configurate dispage display opció, Show conservation regions kékkel kiemelve

LEGEND	Gorilla Gorilla Gorilla Gorilla Gow Gow Gow	
Branch Length — x1 branch length x100 branch length x100 branch length	Species SpeciesName high quality assembly SpeciesName low quality assembly	Expanded Alignments gap aligned seq
Human > chromosome:GRCh38:13:32263179:32400866:1 Bonobo > chromosome:panpan1.11:3:32056171:32093416:1 Gorilla > chromosome:gorGor4.13:1403093114068684:1 Mouse > chromosome:GRCm38:5:150551455:150570786:1 Cow > chromosome:UMD3.112:2863559:28653312:-1 Horse > chromosome:EquCab2.17:11549286:11576061:1 Cat > chromosome:Falis_catus_8.0:A1:11569562:1159458	<u>9.1</u>	
() Alignment		
Currently showing the alignment for first 120 columns only. T Display full alignment	o display the full alignment, please click the button b	elow.
Human NINTGATACGGRAATTGATAGAAGCAGARGATCGGCTA' Bonobo ATATGATATGGAAATTGATAGAAGCAGAAGATCGGCTA' Gorilla ATATGATATGGAAATTGATAGAAGCAGAAGATCGGCTA' Mouse ATACGATATGGAAATTGATAGAAGCAGAAGATCGGCTA' Cow ATATGATGGAAATTGATAGAAGCAGAAGATCGGCGGA'	TAAAAAGATAATGGAAGGGATGRCACAGCTGCAAAA TAAAAAGATAATGGAAGGGATGACACAGCTGCAAAA TAAAAAGATAATGGAAAGGGATGACACAGCTGCAAAA TAAAGAAGATAATGGAAAGGGATGACAGCTGCAAAA TAAAAAGATAATGGAAAGGGATGACACAGCTGCAAAA TAAAAAGATAATGGAAAGGGATGACACAGCTGCAAAA	CACTGTTCTCTGTGTTTCTGACATAATTCATTGAGGGGAAA CACTGTTCTCTGTGTTTCTGACATAATTTCATTGAGGGGAA C.CCTTGTTCTGTGTTTCTGACATAATTCATTGAGGGGAA CCCTGTTCTCTGTATTCAGACTGATTGACGACAAA CACTGTTCTCTGTGTTTCTGAAATAATTCATTAAGGACAGA

Lássuk a variant-okat Configure this page show variációk yes and show links

• Filter:	applied					
Filters have been applied to this sequence. If you no longer wish to use these filters, use "Configure this page" to remove them.						
• Hide	variants from sources: Archive dbSNP					
Human	WARGANAGR <mark>UNPATURRUMS</mark> A <mark>RC<mark>HIC</mark>HACHTCEKCTRYPARAAAGBUTEYGGAAAGGKRUGRSARAGUTCOCAABAAGAGYUKYYCUYUUCU</mark> GAC <mark>RH</mark> AA <mark>Y</mark> III <mark>B</mark> AM <mark>I</mark> GA <mark>RYPCAAA 1</mark>	rs766885				
orilla	HIA IGHIAI GGARAFIGATAGARGARGARGAFGGOTATAGARARGATARIGGARAGGGATGACACROTIGGARARGACHIGITOTOTOTOTOTOTOTOTOTOTOTOTOTAGAAAAATTCATIGARGGGARA ATATGATATGGARATTGATAGARGCRGARGATCGGOTATAAAAARGATARIGGARAGGGATGACACRGOTGCAAAAACACTTGTTCTCTGTGTTTCTGACATAATTCATTGACGCARA					
louse Cow	niacsaterogaaatteacaacascasgasatcogetchaagaasatactegasagegargacacasetocaaaaacocttettetettetetteacatcattteacaasa Itate <mark>xn</mark> et Isa <mark>i</mark> attgatagaagcaaaagatcagegataaaaagataatggaaagggatgacacagetgcaaaaacaettgttetetgttetcgaaataattteattaascaecaga <u>1</u>	rs45136				
Horse Cat	ATATGATATAGAAATTGATAGAAGCAGAAGATCAGCTATAAAAAAGATAATGGAAAGGGATGACACAGCTGCAAAAAGACTTGTTCTGTGTGTTTCTGAAAAAATTTCATCCAGTACAGA ATACGATATGGAAATTGATAGAAGCAAAAGATCAGCTATTAAAAAGATCATGGAAAGGGATGATACCGCAGCAAAAACACTTGTTCTCGCATTTCTGAAACAATTTCATCAAGCACAGA					

Ensembl release 92 - April 2018 © EMBL-EBI

Permanent link - View in archive site

Variation: <u>rs451362933</u>						
Class	SNP					
Location	12:28653312					
Source	db SNP					
Alleles	T/C					
Consequences	splice region variant					
synonymous variant						
Explore this variant						
Gene/Transcript Locations						

Nézzük meg a cow Y-t -SNP -tovább az rs-es nevére



Információk róla pl. phylogenetic context



90 faj BRCA2 génje alapján gene tree a zöld szín intenzitása jelzi az illesztés erősségét, a piros a mi génünk, a tölcsér összegyűjtött mód, fehér gap-eket jelöl, a node-kra kattintva megkapjuk azok jellemzőit. A piros node duplikációt jelöl, világos kék kétértelmű kék speciális. az evolúciós távolságok is láthatóakView alignment image: kékek a kontigok pirosak a breaing pointok,

🌣 < 🖬 🕁 🇞					
Human		100kb	427.24 kb 200kb	300kb	400kb
Genes (Comprehensive set	BRCA2 >				IFIT1P1 >
Contigs			AL445212.9 >		
Bonobo		100kb	200kb	300kb	400kb
Genes (Ensembl)	BRCA2 >				+++1
Contigs	AJFE02094368.1 >	>A	JFE02094368.1 >	AJFE02094370.:	1 >
Chimpanzee		100kb	200kb	300kb	400kb
Genes (Ensembl)	H BRCA2 >		1 1 1 1		+++4
Contigs	AAC204071714.1	· · · · · · · · · · · · · · · · · · ·	AAC2040/1/14.1 >	AAC2040/1/14.1 >	
Gorilla		100kb	200kb	300kb	400kb
Genes (Ensembl)	BRCA2 >		1 1 1		+++1
Contigs	CABD030088451.1 >		CABD030088452.1 >	CABD030088452.1	>
Orangutan		100kb	200kb	300kb	400kb
Genes (Ensembl)	H H H H				+++
Contigs	< Contig230.54	4	< Contig230.53	< Contig230.50	

Syntenies: A színténia boxok színes különböző kromoszómáról származó blokk különböző színű, fajokat lehet változtani



Nézzük az ortholog-okat: válasszuk a primates-eket-> selected orthologs, nézzük meg a bonobo alignmentjét

Show All v entries		Show/hide	e columns					Filter	
Species	Туре	Orthologue		dN/dS	Target %id	Query %id	GOC Score	WGA Coverage	High Confidence
Angola colobus	1-to-1	ENSCANG0000028513		0.53567	91.38 %	91.22 %	100	58.69	Yes
(Colobus angolensis palliatus)	View Gene Tree	Compare Regions (KN981 1,632,495:-1)	104.1:1,541,886-						
		View Sequence Alignments	2						
Black snub-nosed monkey	<u>1-to-1</u>	ENSRBIG0000001621		0.42170	87.56 %	9.89 %	50	55.36	No
(Rhinopithecus bieti)	View Gene Tree	Compare Regions (MCGX) 8,856:-1)	01034939.1:7,696-						
		View Sequence Alignments	2						
Bonobo	1-to-1	ENSPPAG0000002166		0.31967	99.12 %	99.09 %	100	100.00	Yes
(Pan paniscus)	View Gene Tree	Compare Regions (13:32,0 32,091,332:1)	10,066-						
		View Sequence Alignments	2						
Bushbaby	<u>1-to-1</u>	BRCA2 (ENSOGAG0000	1010588)	0.49387	72.61 %	72.35 %	100	56.13	Yes
(Otolemur gamettii)	View Gene Tree	Compare Regions (GL873) 415,772:-1)	639.1:315,573-						
		View Sequence Alignment	2						
Capuchin (Cabua capucipus	<u>1-to-1</u>	ENSCCAG00000033858		0.55600	88.17 %	87.62 %	100	100.00	Yes
Species		Gene ID	Peptide ID		Peptide le	ngth %i (cD	dentity NA)	% coveraç	ge Genomic lo
Human (Homo sap	iens)	ENSG00000139618	ENSP00000369	9497	3418 aa	99	%	99 %	13:3231547
Bonobo (Pan panis	scus)	ENSPPAG0000002166	ENSPPAP0000	0000836	3417 aa	99	%	100 %	<u>13:3201006</u>
CLUSTAL W (1.81) multiple sequence alignment ENSP00000369497/1-10254 ATGCTATTGGATCCAAAGAGAGGCCAACATTTTTGAAATTTTTAAGACACGCTGCAAC ENSPPAP00000000836/1-10251 ATGCTTATTGGATCCAAAGAGGCCAACATTTTTGAAATTTTTAAGACACGCTGCAAC									
ENSP0000036949 ENSPPAP0000000	7/1-10254 0836/1-10251	AAAGCAGATTTAGGACCAATA AAAGCAGATTTAGGACCAATA ******	AGTCTTAATTGGI AGTCTTAATTGGI ******	TTTGAAGAP TTTGAAGAP	ACTITICITICA ACTITICITICA	GAAGCTCC GAAGCTCC *******	A A *		
ENSP00000369497/1-10254 CCCTATAATTCTGAACCTGCAGAAGAAGTCTGAACAATAAAAACAACAATTACGAACCAAAC ENSPPAP00000000836/1-10251 CCCTATAATTCTGAACCTGCAGAAGAAGTATCGAACAATAAAAACAACAATTATGAACCAAAC									
ENSP0000036949 ENSPPAP00000000	7/1-10254 0836/1-10251	CTATTTAAAACTCCACAAAGG CTATTTAAAACTCCGCAAAGG ******	AAACCATCTTATA AAACCATCTTATA	ATCAGCTO	GGCTTCAACT GGCTTCAACT	CCAATAAT CCAATAAT *******	A A *		
ENSP0000036949 ENSPPAP0000000	7/1-10254 0836/1-10251	TTCAAAGAGCAAGGGCTGACI TTCAAAGAGCAAGGGCTGACI	CTGCCGCTGTACC	CAATCTCCI	IGTAAAAGAA IGTAAAAGAA	TTAGATAA TTAGATAA *******	A A *		

Nézzük meg a protein familiest, az összes izoforma Markov clusterezése adja meg ezeket a protein családokat. 1 találat van! Nézzük meg a tagjait:

Ensembl protein families @

Show/hide columns			Filter
Family ID	 Consensus annotation 	Other BRCA2 proteins in this family	Multiple alignments
PTHR11289_SF0 (2 genes)	BREAST CANCER TYPE 2 SUSCEPTIBILITY HOMOLOG FANCONI ANEMIA GROUP D1 HOMOLOG	ENSP00000369497 (BRCA2-201) ENSP00000434898 (BRCA2-202) ENSP00000439902 (BRCA2-206)	126 Ensembl members of this family <u>Wasabi viewer</u> 1170 members of this family <u>Wasabi viewer</u>

Menjünk a lokáció fülre nézzük meg a kromoszóma sávtérképét itt a gén piros boxxal jelölt.

5

HUMAN genes in this family

		gen	.5 m	una	, iuii	iiiy										
Ensen	ıbl gen	ies col	ntainin	g prot	eins ir	n family	y <u>PTH</u>	R11289	<u></u>							
								9	10		12					
14	15	16	17	18	19				() ×	ļ	МĪ					
Gene ID and Location Gene Name Descript						Desc	riptio	a (if known)	Protein ID(s)							
ENSO	00000	00000139618 BRCA2 @							BRCA2, DNA repair associated [Source:HGNC Symbol;Acc:HGNC:1101]							
LRG_2	293 93: 5.00k	o. az.azr	n		B	RCA2 d	₽				BRCA2, DNA repair associated [Source:HGNC Symbol;Acc:HGNC:1101]					

Config this page: Comparative genomics menü, conservation regiók beállítása



*Az izozimek vagy izoenzimek kifejezés olyan enzimekre utal, melyek aminosav-szekvenciájukban eltérnek, de ugyanazt a kémiai reakciót katalizálják különböző alléljei által kódolt enzimekre, és az izozimek jelentik a különböző génekből származó, de ugyanazt a reakciót katalizáló enzimeket, a két szót általában felcserélhető módon használják. (wikipédia)

* Homológia a közös evolúciós eredetet jelenti (szekvenciák esetében). Nincs mértéke (nem fejezhető ki %–ban!!!), két szekvencia vagy homológ, vagy nem. A homológia két típusa: ortológia és paralógia

*Ortológia:Két gén ortológ, ha két különbözőfajban találhatóak, és egy közös ős–génből származnak, mely a két faj közös ősében volt jelen. Ugyanazt a funkciót szolgálják, a két fajban. Példa: laktát dehidrogenáz (ill. génje) az emberben és az egérben.

*Paralógia:Két gén paralóg, ha ugyanabban az organizmusban találhatóak, és egy közös ős–génből génduplikáció és azt követő-divergens evolúció útján alakultak ki. Többnyire különböző, de egymással összefüggésben lévőfunkciójuk van. Példa: a hisztidin–bioszintézis enzimei (ill. ezek génjei) emberben (nagyon hasonló szerkezetűek, de más–más reakciót katalizálnak). (wikipédia)

Biomart:

- Ez egy alkalmazás a böngészőn belül
- Nem kell programozni az adatok kiexportálásához
- Néhány klikkel kereső szekvenciát hozhatsz létre
- Az ereményket megnézheted online vagy letöltheted fasta, xls, csv, tsv html formátumokban

choose database ensembl-choose dataset human genes-Filters gene Input external refernces HGNC symbolt írjuk be a BRCA2, ATM

count: gyors összegzés

Attributes: External HGNC symbol, Refseq peptid predicted ID,

unque results only , hogy a multiple transzkriptek redundanciáját csökkentsük

export : letölthető zippelt formátumban, vagy el lehet küldeni saját emai-címre is.

💿 Browsing RNA-Seq Data x 🐨 Izoenzim – Wikipédia x 🖉 www.ensembl.org/Dioma X M bioinfo terem - bana.a.n: x 📃 🖤 Niko 📮 🗗 🗙								
\leftrightarrow \rightarrow C \triangle (i) www.ensembl.c	rg/biomart/martview/d573d149e49e4d320d06eb34855dcde4	☆ 💦 :						
CENSEMBI BLAST/	SLAT BioMart BLAST VEP Tools Downloads Help & Docs Blog	Login/Register						
> New Count Results	URL XML 2 Perl O Help							
Dataset [None selected]	CHOOSE DATABASE ▼							
	In order to maintain service for all users. BioMart browser sessions running for more than 5 minutes are terminated. If you have							
🛋 🖨 💾 🧮	Q V 2 A 0 x	HUN 12:30 HU 2018.04.16.						

RNA-seq adatok vizualizációja



A human USP25 a zsírszövetben expresszálódik –e? Configure this page RNA-seq models az összeset klikkeld be

Active tracks		Gones and transcripts	Find a track
Favourite tracks		Genes and transcripts	
Track order		Enable/disable all RNA-Seq models	
Search results		Adipose intron-spanning reads	* 0
Sequence and assembly	(4/24)	Adrenal intron-spanning reads	* 0
- Sequence	(2/4)	Blood intron-spanning reads	* 0
 Markers Ditan features 	(0/1)	Brain intron-spanning reads	* 0
- Simple features	(1/4)	Breast intron-spanning reads	* 0
 Misc. regions & clones 	(1/13)	Colon intron-spanning reads	* 0
Genes and transcripts	(37/43)	Heart intron-spanning reads	* 0
 Genes Prediction transcripts 	(5/9)	Kidney intron-spanning reads	* 0
RNA-Seq models	(32/32)	Liver intron-spanning reads	* 0
mRNA and protein alignments	(2/15)	Lung intron-spanning reads	* 0
- mRNA alignments	(1/3)	Lymph intron-spanning reads	* 0
 Protein alignments 	(1/5)	Ovary intron-spanning reads	* 0
 Protein features 	(0/6)	Prostate intron-spanning reads	* 0
ncRNA	(0/3)	Skeletal muscle intron-spanning reads	* 0
Other DNA alignments	(1/1)	Testes intron-spanning reads	* 0
Contraction Contraction			

Az RNA-seq-ben is látható intron exon rész, Az adipose gén model lehelyezem a sárgákhoz ezek cDNS-kből származnak. A consensus coding sequnce-st bezárom CCDS, így látszik hogy a az adipose RNA-seq-em nagyon hasonlít (ugyanaz) mint a USP25-001 transzkript, vagyis ez a splice variant az zsírszövetben expresszálódik-> hipotézis igaz,

Intron modellek:



Ontology és phenotypes cancer gene

Több segítség:

Ensembl <u>documentaciók, videó segédlet (tutorials) a YouTube-on, FAQs, gyakorlati példák,</u> publikációk, online kurzusok.email a helpdesk-nek: <u>helpdesk@ensembl.org</u>, kövesd az Ensembl <u>blogot</u>, <u>íratkozz fel a levelező listára</u>

Flicek, P. *et al* **Ensembl 2013** Nucleic Acids Res. Advanced Access (Databaselssue) <u>http://www.ncbi.nlm.nih.gov/pubmed/23203987</u>

Ensembl Methods Series http://www.biomedcentral.com/series/ENSEMBL2010

Xosé M. Fernández---Suárez and Michael K. Schuster **Using the Ensembl Genome Server to Browse Genomic Sequence Data.** UNIT 1.15 in Current Protocols in Bioinformatics, Jun 2010.

Giulietta M Spudich and Xosé M Fernández---Suárez **Touring Ensembl: A practical guide to genome browsing** BMC Genomics 2010, 11:295 (11 May 2010)