

## PEC1. GENÒMICA COMPUTACIONAL

Amèlia Martínez Sequera.

### Exercici 1

(S'annexa PDF article PubMed: Rhinn H., Abeliovich A. *Differential Aging Analysis in Human Cerebral Cortex Identifies Variants in TMEM106B and GRN that Regulate Aging Phenotypes. Cell Syst. 2017*)

Rhinn i Abeliovich van estudiar l'expressió del gen TMEM106B sobretot al córtex frontal (humà) i, en particular, del SNPs rs1990622 i quantes còpies contenia de l'al·lel de risc A (0,1,2), i del rs708384 al gen de la progranulina, però en aquest cas no s'obté significació estadística. Fan un metanàlisi i conclouen que l'expressió màxima de l'al·lel de risc al rs190622 a les cèl·lules del córtex frontal està relacionat amb un augment de la inflamació neuronal i un dèficit cognitiu, tal i com es va veure en estudis anteriors, quan es va seqüenciar el RNA del gen tmem106b a ratolins.

**Localització genòmica:** TMEM106B chr7:12,211,294-12,243,367 *Homo sapiens*, hg 38. RefSeq Genes (UCSC RefSeq)

*Homo sapiens*, hg19: chr7:12250920-12282993

*Mus musculus*, mm10: chr6:13,069,759-13,089,269 No existeixen dades en mm9

chr6:13069841-13092696 (amb LiftOver)

chr6:13,071,744-13,084,326 (amb BLAT)

### **Gens propers:**

**VWDE** (RefSeq NM\_001346973), hg38 chr7:12,330,885-12,403,865

**THSD7A**, precursor del domini de la tromboespondina humana, mRNA

RefSeq(NM\_015204.3), hg38 chr7:11370365-11832198

### OMIM

TMEM106B es troba al chr. 7p21.3, controla el nombre, la mida, mobilitat, tràfic i acidificació dels lisosomes. Es va associar a la degeneració del lòbul frontotemporal, amb inclusions de TDP43, i es va detectar la expressió en el córtex frontal i en línies cel·lulars linfoblastoides (Van Deerlin et al. 2010). Conté 274 aa i una massa molecular de 43kD en línies cel·lulars de humà i ratolí (Brady et al. 2013), i, al menys, 5 possibles llocs de N-glicosilació C-terminal al domini

transmembrana predictiu(Chen-Plotkin et al. 2012). Ortòlegs de TMEM106B han sigut detectats en diferents llinatges de vertebrats, amb casi absoluta conservació de la regió trans-membrana i el domini membrana-proximal luminal (Brady et al. 2014). S'expressa en el SNC, inclosos el oligodendròcits (Simons et al. 2017).

MicroRNAs, MIR132 i MIR212, regulen la transcripció i expressió de proteïnes, disminuint-la quan s'uneixen al 3'UTR del TMEM106B. En FTLD (degeneració del lòbul frontotemporal) i demència es va veure que augmentava significativament la expressió de TMEM106B i disminuïa la de MIR132/MIR212, comparat amb un cervell normal. La sobreexpressió de TMEM106B en humans causa anomalies en la morfologia endolisosomal i la seva acidificació, i també augmenta la progranulina intracel·lular (Chen-Plotkin et al. 2012). També s'ha vist que provoquen l'engrossiment de lisosomes i inhabiliten la capacitat de fusió amb el endosomes i l'apoptosi (Brady et al. 2013)

Hi ha una correlació entre l'al·lel de risc A del SNP rs1990622 i l'increment de l'expressió del TMEM106B en molts tipus cel·lulars. En cèl·lules HeLa i neurones resulta en un engrossiment de lisosomes i un augment de la mort cel·lular (Gallagher et al.2017).

En 4 malalts de Leucodistrofia hipomielinitzant-16 es va identificar una mutació *missense* heterozigòtica (Simons et al. 2017), i posteriorment en una nena xinesa (Yan et al. 2018). La mutació tenia lloc al dinucleòtid CpG, el que assenyalava que és una mutació recurrent perquè és un punt d'accés dintre del gen. Variants al·leliques en HLD-16: c.754G-A transició (NM\_001134232) a l'exó 8, resultant en la substitució asp252-to-asn (D252N) de la part intraluminal de la proteïna.

Tmem106b <sup>-/-</sup> de ratolí eren aparentment normals, però tenien regulats a la baixa molts enzims de lisosomes en maduració (Klein et al. 2017).

### **Entrez Gene**

TMEM106B

Prot.transmembrana 106B. Localització 7p21.3, 9 exons.

Homòlegs: el gen TMEM106B està conservat en chimpanzè, Rhesus, gos, vaca, ratolí, pollastre, zebrafish i granota. <https://www.ncbi.nlm.nih.gov/sites/homologene/56806>

Ortòlegs: 286 organismes tenen ortòlegs amb el TMEM106B humà.

<https://www.orthodb.org/?ncbi=1&query=54664>. Enllaç Ortho DB

[https://www.ncbi.nlm.nih.gov/gene/?Term=ortholog\\_gene\\_54664\[group\]](https://www.ncbi.nlm.nih.gov/gene/?Term=ortholog_gene_54664[group]). Enllaç NCBI

Fenotips: [https://www.ncbi.nlm.nih.gov/gtr/tests/?term=TMEM106B\[gene\]](https://www.ncbi.nlm.nih.gov/gtr/tests/?term=TMEM106B[gene]) NCBI

Malalties associades: <https://www.ncbi.nlm.nih.gov/gtr/conditions/C4693779>

mRNA i proteïnes:

NM\_001134232.2 → NP\_001127704.1 transmembrane protein 106B. Variant 2 que difereix de

la variant 1 en el 5'UTR (la proteïna és la mateixa).

NM\_018374.4 → NP\_060844.2 transmembrane protein 106B. Variant 1.

[https://www.ncbi.nlm.nih.gov/CCDS/CcdsBrowse.cgi?](https://www.ncbi.nlm.nih.gov/CCDS/CcdsBrowse.cgi?REQUEST=CCDS&GO=MainBrowse&DATA=CCDS5358.1)

REQUEST=CCDS&GO=MainBrowse&DATA=CCDS5358.1. Seqüència CDS NCBI. També hi ha un enllaç amb UniProt.

**Gene Ontology** <https://www.ebi.ac.uk/GOA/>

TMEM106B humà:

- 32098 bps chr7 des de 12211270 fins 12243367.

- té 8 transcrits que contenen un total de 29 exons (+).

- Enllaç a ENSEMBL. [http://www.ensembl.org/Homo\\_sapiens/Gene/Sequence?](http://www.ensembl.org/Homo_sapiens/Gene/Sequence?db=core;g=ENSG00000106460;r=7:12211270-12243367)  
db=core;g=ENSG00000106460;r=7:12211270-12243367

- Variacions: 8290 SNPs. [http://www.ensembl.org/Homo\\_sapiens/Gene/Variation\\_Gene/Image?](http://www.ensembl.org/Homo_sapiens/Gene/Variation_Gene/Image?db=core;g=ENSG00000106460;r=7:12211270-12243367)  
db=core;g=ENSG00000106460;r=7:12211270-12243367

- Ortòlegs: 239 en ENSEMBL.

[http://www.ensembl.org/Homo\\_sapiens/Gene/ compara\\_Ortholog?](http://www.ensembl.org/Homo_sapiens/Gene/ compara_Ortholog?db=core;g=ENSG00000106460;r=7:12211270-12243367)  
db=core;g=ENSG00000106460;r=7:12211270-12243367

- Regulació: 5 elements en la regió de TMEM106B.

[http://www.ensembl.org/Homo\\_sapiens/Gene/Regulation?](http://www.ensembl.org/Homo_sapiens/Gene/Regulation?db=core;g=ENSG00000106460;r=7:12211270-12243367)  
db=core;g=ENSG00000106460;r=7:12211270-12243367

Tmem106B *Mus musculus*: (inclou també enllaços corresponents a ENSEMBL)

- 19511 bps chr6 des de 13069759 fins 13089269.

- té 6 transcrits que contenen 21 exons (+).
- 738 SNPs i té també 5 elements de regulació.

### ***CDS TMEM106 humà:***

>hg38\_refGene\_NM\_018374 range=chr7:12214811-12231975 5'pad=0 3'pad=0 strand=+ repeatMasking=none

```
ATGGGAAAGTCTCTTTCTCATTTGCCTTTGCATTCAAGCAAAGAAGATGC
TTATGATGGAGTCACATCTGAAAACATGAGGAATGGACTGGTTAATAGTG
AAGTCCATAATGAAGATGGAAGAAATGGAGATGTCTCTCAGTTTCCATAT
GTGGAATTTACAGGAAGAGATAGTGTACCTGCCCTACTTGTGAGGGAAC
AGGAAGAATTCTAGGGGGCAAGAAAACCAACTGGTGGCATTGATTCCAT
ATAGTGATCAGAGATTAAGGCCAAGAAGAACAAAGCTGTATGTGATGGCT
TCTGTGTTTGTCTGTCTACTCCTTTCTGGATTGGCTGTGTTTTTCCTTT
CCCTCGCTCTATCGACGTGAAATACATTGGTGTAATCAGCCTATGTCA
GTTATGATGTTGAGAAGCGTACAATTTATTTAAATACACAAACACTA
AATATAACAAACAATAACTATTACTCTGTGGAAGTTGAAAACATCACTGC
CCAAGTTCAATTTCAAAAACAGTTATTGAAAGGCACGCTTAAACAACA
TAACCATTATTGGTCCACTTGATATGAAACAAATTGATTACACAGTACCT
ACCGTTATAGCAGAGGAAATGAGTTATATGTATGATTTCTGTACTCTGAT
ATCCATCAAAGTGCATAACATAGTACTCATGATGCAAGTTACTGTGACAA
CAACATACTTTGGCCACTCTGAACAGATATCCAGGAGAGGTATCAGTAT
GTCGACTGTGGAAGAAACACAACCTTATCAGTTGGGGCAGTCTGAATATT
AAATGTACTTCAGCCACAACAGTAA
```

### ***Proteïna:***

>NP\_001127704 length=274

```
MGKSLSHLPLHSSKEDAYDGVTSENMRNGLVNSEVHNEDGRNGDVSQFPY
VEFTGRDSVTCPTCQGTGRIPRGQENQLVALIPYSDQRLRPRRTKLYVMA
SVFVCLLSGLAVFFLPRSIDVKYIGVKSAYVSYDVQKRTIYLNITNTL
NITNNNYSVEVENITAQVQFSKTVIGKARLNNITIIGPLDMKQIDYTVP
TVIAEEMSYMYDFCTLSIKVHNIVLMMQVTVTTTYFGHSEQISQERYQY
VDCGRNTTYQLGQSEYLNVLQPQQ
```

### ***CDS tmem106b Mus musculus***

>mm10\_refGene\_NM\_027992 range=chr6:13071744-13084326 5'pad=0 3'pad=0 strand=+ repeatMasking=none

ATGGGAAAGTCTCTTTCTCACTTACCTTTGCATTCAAATAAAGAAGATGG  
CTATGATGGCGTTACATCGACAGACAATATGAGAAATGGATTGGTTAGCA  
GTGAAGTGCACAACGAAGACGGAAGAAATGGAGATGTCTCTCAGTTCCCA  
TATGTGGAATTTACTGGAAGAGATAGTGTCACTTGTCCCACTTGCCAAGG  
AACAGGAAGAATTCCTAGGGGACAAGAAAACCAACTGGTGGCATTGATTC  
CATATAGTGATCAGCGGTTACGGCCAAGAAGAACAAGCTGTATGTGATG  
GCGTCTGTGTTTGTCTGCCTGCTCCTGTCTGGATTGGCTGTGTTTTTCT  
TTTCCTCGATCTATTGAGGTGAAGTACATTGGAGTAAATCAGCCTATG  
TCAGCTACGACGCTGAAAAGCGAACCATATATTTAAATATCACGAACACA  
CTAAATATAACAATAATAACTATTATTCTGTTGAAGTTGAAAACATCAC  
TGCTCAAGTCCAGTTTTCAAACCGTGATTGGAAAGGCTCGTTAAACA  
ACATAACTAACATTGGCCCACTTGATATGAAGCAGATTGATTATACGGTA  
CCCACAGTTATTGCAGAGGAAATGAGTTACATGTATGATTCTGTACACT  
GCTCTCCATCAAAGTGCACAACATAGTACTCATGATGCAAGTTACTGTAA  
CAACAGCATACTTTGGACACTCTGAGCAGATATCTCAGGAAAGGTACCAG  
TATGTCGACTGTGGAAGGAACGACTTACCAGTTGGCCAGTCTGAGTA  
TCTAAATGTCCTTCAGCCACAACAATAA

### ***Proteïna:***

>NP\_082268 length=275

MGKSLSHLPLHSNKEDGYDGVSTDNMRNGLVSSEVHNEDGRNGDVSQFP  
YVEFTGRDSVTCPTCQGTGRIPRGQENQLVALIPYSDQRLRPRRTKLYVM  
ASV FVCLLSGLAVFFLPR SIEVKYIGVKSAYVSYDAEKRTIYLNITNT  
LNITNNNY SVEVENITAQVQFSKTVIGKARLNNITNIGPLDMKQIDYTV  
PTVIAEEMSYMYDFCTLLSIKVHNIVLMMQVTVTTAYFGHSEQISQERYQ  
YVDCGRNTTYQLAQSEYLNVLQPQQ

Comparativa de seqüències ENSEMBL



## CLUSTAL Omega

### Seqüència DNA

# Aligned\_sequences: 2

# 1: TMEM106B

# 2: tmem106b

# Matrix: EDNAFULL

# Gap\_penalty: 16

# Extend\_penalty: 4#

# Length: 828

# Identity: 733/828 (88.5%)

# Similarity: 733/828 (88.5%)

# Gaps: 3/828 ( 0.4%)

# Score: 3273

```

      10      20      30      40      50
EMBOSS ATGGGAAAGTCTCTTTCTCATTTCCTTTGCCTTTGCATTCAAGCAAAGAAGATGC
EMBOSS .....C..A.....AT.....G

      60      70      80      90
EMBOSS TTATGATGGAGTCACATCT---GAAAACATGAGGAATGGACTGGTTAATA
EMBOSS C.....C..T....GACA..C..T....A.....T.....GC..

     100     110     120     130     140
EMBOSS GTGAAGTCCATAATGAAGATGGAAGAAATGGAGATGTCTCTCAGTTTCCA
EMBOSS .....G..C..C.....C.....C.....C...

     150     160     170     180     190
EMBOSS TATGTGGAATTTACAGGAAGAGATAGTGTACCTGCCCTACTTGTTCAGGG
EMBOSS .....T.....T...T..C....C..A..

     200     210     220     230     240
EMBOSS AACAGGAAGAATTCCTAGGGGGCAAGAAAACCAACTGGTGGCATTGATTC
EMBOSS .....A.....

     250     260     270     280     290
EMBOSS CATATAGTGATCAGAGATTAAGGCCAAGAAGAACAAAGCTGTATGTGATG
EMBOSS .....C.G...C.....

     300     310     320     330     340
EMBOSS GCTTCTGTGTTTGTCTGTCTACTCCTTTCTGGATTGGCTGTGTTTTCTT
EMBOSS ..G.....C..G....G.....T..

     350     360     370     380     390
EMBOSS TTTCCCTCGCTCTATCGACGTGAAATACATTGGTGTAATCAGCCTATG
EMBOSS .....A....T..G....G.....A.....

     400     410     420     430     440
EMBOSS TCAGTTATGATGTTCAGAAGCGTACAATTTATTTAAATATCACAAACACA
EMBOSS ...C..C..C.C.G.A....A..C..A.....G.....

```

## Proteīna

Aligned\_sequences: 2

# 1: Proteīna transmembrana 106B humana

# 2: Proteīna transmembrana 106b ratolī

# Matrix: EBLOSUM62

# Gap\_penalty: 14

# Extend\_penalty: 4

# Length: 275

# Identity: 263/275 (95.6%)

# Similarity: 269/275 (97.8%)

# Gaps: 1/275 ( 0.4%)

# Score: 1360

[illegible]

Identitat: 95.636%

### **Conclusió:**

La gran similitud i identitat que hi ha entre la seqüència humana i la de ratolí (gairebé del 90%) pot ser indicativa de que provenen d'una seqüència ancestral comú i que s'ha conservat al llarg de l'evolució. També ens indica que existeix un rol biològic compartit, ja que les regions coincidents són aquelles implicades en la realització de les funcions conservades.

Amb la comparativa de PhyloP i PhatCons també podem veure, de manera més gràfica, com és una seqüència molt conservada dins dels vertebrats.

El fet que les proteïnes tinguin més d'un 95% de coincidència pot indicar que aquestes molècules tenen una funció biològica semblant a les dues espècies. Molt del canvis nucleotídics



no afecten a la proteïna, són substitucions amb més grau d'acceptació a la selecció natural.

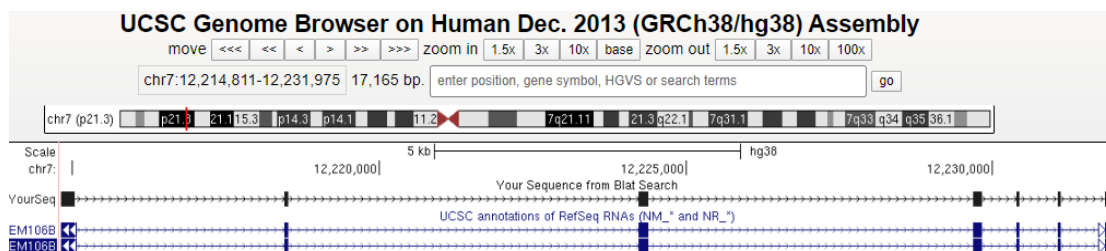


## Exercici 2

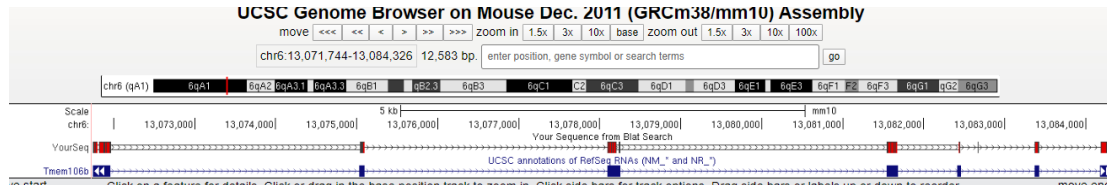
BLAT està dissenyat per trobar seqüències amb un 95% o més de similitud, amb una mida de 25 pb o més. Troba coincidències perfectes de seqüències de 20 bases. En proteïnes troba seqüències del 80% o més similitud, de la mida de 20 aa o més. DNA BLAT treballa bé en primats, i protein BLAT en vertebrats.

DNA BLAT treballa mantenint un índex del genoma total en memòria. L'índex consta de tots els 11-mers que es solapen, excepte que participin en les repeticions, per trobar àrees de probable homologia, i després es carreguen a la memòria per l'alineament. Protein BLAT, treballa d'igual manera, però són 4-mers en comptes de 11-mers.

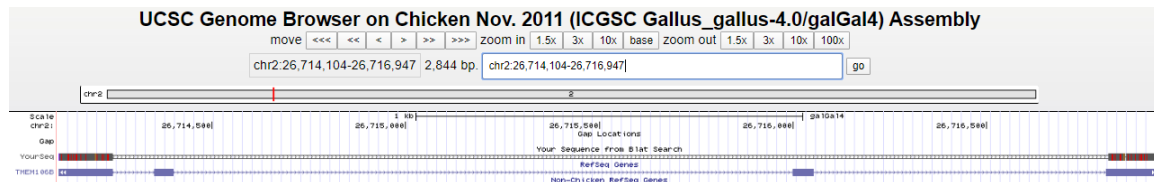
chr7:12,214,811-12,231,975



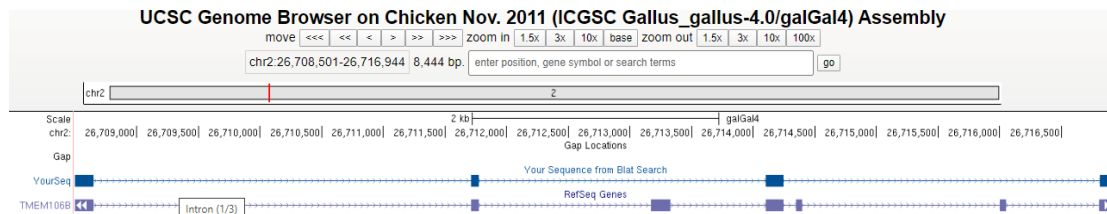
en genoma de ratolí mm10: chr6:13,071,744-13,084,326



en genoma de pollastre: chr2:26,714,104-26,716,947



Proteïna humana amb pollastre (BLAT): chr2:26,708,501-26,716,944



### Conclusió:

Degut a la distància evolutiva (major entre aus/homes que entre mamífers), les seqüències de pollastre i humà no són tan similars, però molts són canvis que no afecten a la proteïna resultant, probablement de nou pel rol biològic que comparteixen.

### Exercici 3

El "table browser" s'utilitza per interrogar sistemàticament al navegador UCSC sobre un conjunt específic de registres d'una pista dins una part d'un genoma seleccionat (que pot arribar a ser el genoma sencer), i realitzar operacions de selecció, intersecció i correl·lació entre aquests. Podem recuperar la informació a través de pistes pròpies o de fitxers de text.

El genoma humà hg38, segons refGene, conté 85561 transcrits, i 3257347282 parells de bases

item count	85,561
item bases	1,443,021,625 (46.61%)
item total	5,441,489,486 (175.76%)
smallest item	21
average item	63,598
biggest item	2,298,757
block count	847,215
block bases	96,308,767 (3.11%)
block total	274,547,488 (8.87%)
smallest block	1
average block	324
biggest block	91,671

[Genomes](#)
[Genome Browser](#)
[Tools](#)
[Mirrors](#)
[Downloads](#)
[My Data](#)
[Projects](#)

## Table Browser

Use this program to retrieve the data associated with a track in text format, to calculate intersections between tracks, to calculate the overlap between tracks, to calculate the overlap between tracks, to calculate the overlap between tracks.
 [Using the Table Browser](#) for a description of the controls in this form, and the [User's Guide](#) for general information.
 To examine the biological function of your set through annotation enrichments, send the data to the [MySQL server](#).
 All tables can be downloaded in their entirety from the [Sequence and Annotation Downloads](#) page.

**clade:** 
**genome:** 
**assembly:**

**group:** 
**track:**

**table:**

**region:** ☒ genome ☐ position

**identifiers (names/accessions):**

**filter:**

**subtrack merge:**

**intersection:**

**correlation:**

**output format:** 
☐ Send output to [Galaxy](#) ☐ [GREAT](#)

**output file:**  (leave blank to keep output in browser)

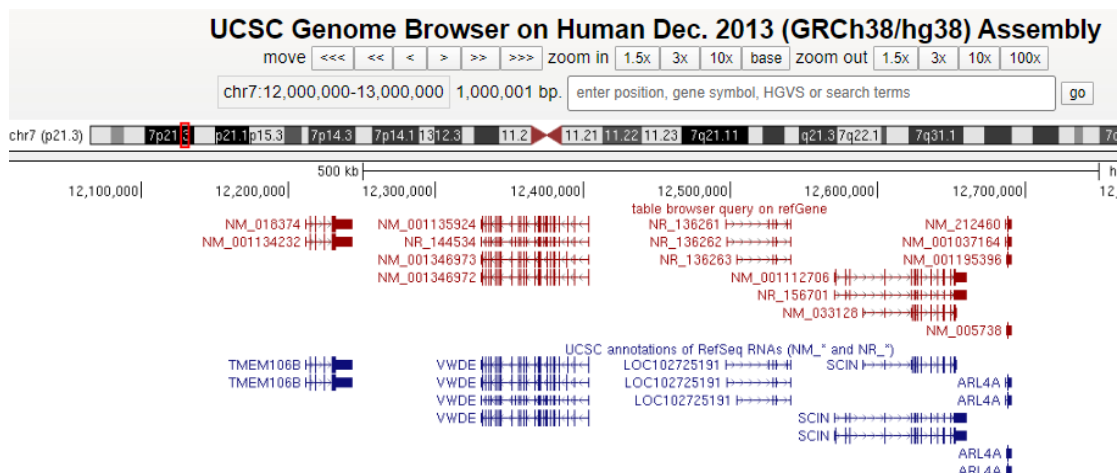
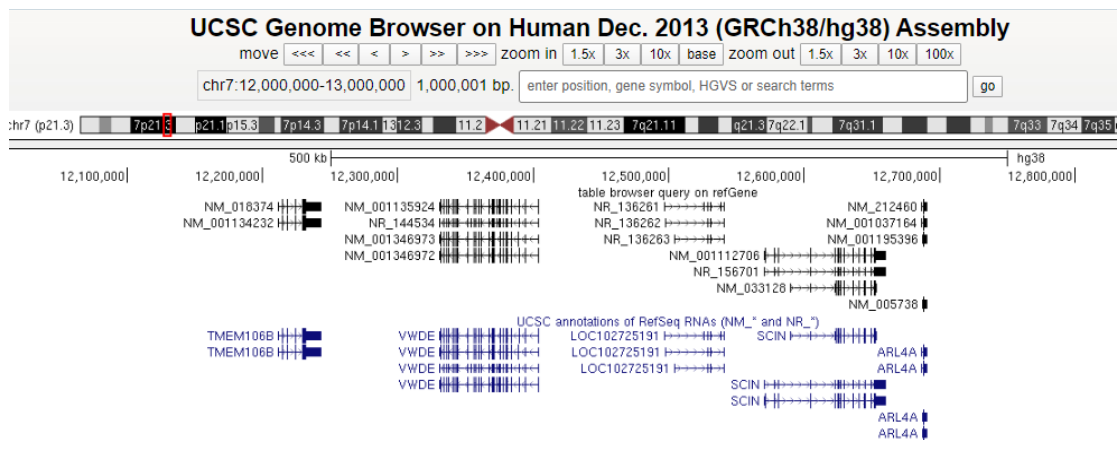
**file type returned:** ☒ plain text ☐ gzip compressed

#name	chrom	strand	txStart	txEnd	exonCount	name2	
NM_000299	chr1	+	201283451	201332993	15	PKP1	
NM_001276351	chr1	-	67092165	67134970	8	C1orf141	
NM_001276352	chr1	-	67092165	67134970	9	C1orf141	
NM_001005337	chr1	+	201283505	201332989	14	PKP1	
NR_075077	chr1	-	67092165	67134970	10	C1orf141	
NM_001042681	chr1	-	8352403	8817640	23	RERE	
NR_038261	chr1	-	41847188	42035934	4	HIVEP3	
NM_001042682	chr1	-	8352403	8423832	13	RERE	
NM_052896	chr1	-	33513998	34165842	70	CSMD2	
NM_012102	chr1	-	8352403	8817640	24	RERE	
NM_001281956	chr1	-	33513997	34165230	71	CSMD2	
NM_001319683	chr1	-	92246397	92299009	18	GLMN	
NM_001320285	chr1	-	75202128	75611114	26	SLC44A5	
NM_001130058	chr1	-	75202128	75611114	24	SLC44A5	
NM_001320287	chr1	-	75202128	75611114	25	SLC44A5	
NM_053274	chr1	-	92246401	92298987	19	GLMN	
NR_135089	chr1	-	92246397	92299009	18	GLMN	
NM_152697	chr1	-	75204474	75611114	24	SLC44A5	
NM_001320283	chr1	-	75204474	75616163	22	SLC44A5	
NM_001350214	chr1	-	83865023	83999132	22	TTLL7	
NM_006699	chr1	+	117367448	117528870	13	MAN1A2	
NM_001350215	chr1	-	83865023	83999132	20	TTLL7	
NM_024686	chr1	-	83865023	83999132	21	TTLL7	
NM_001146191	chr1	+	167721981	167791919	3	MPZL1	
NM_024569	chr1	+	167721981	167791919	5	MPZL1	
NM_030806	chr1	+	184387028	184629019	6	C1orf21	
NM_003568	chr1	+	150982291	150995633	14	ANXA9	
NR_126003	chr1	-	192935743	192948257	3	RGS2-AS1	
NM_001001740	chr1	-	175944825	176207244	19	COP1	

Regió chr7:12000000-13000000

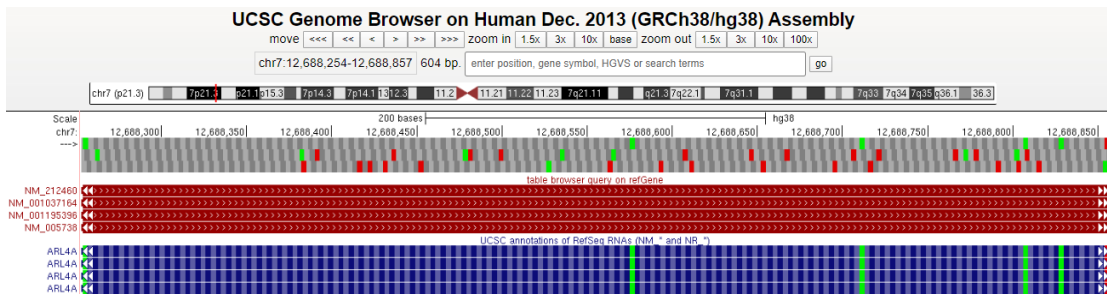
16 transcrits:

#name	chrom	strand	txStart	txEnd	exonCount	name2	
NM_001134232	chr7	+	12211293	12243367	8	TMEM106B	
NM_018374	chr7	+	12211293	12243367	9	TMEM106B	
NM_001346972	chr7	-	12330884	12403941	27	VWDE	
NM_001346973	chr7	-	12330884	12403941	27	VWDE	
NR_144534	chr7	-	12330884	12403941	30	VWDE	
NM_001135924	chr7	-	12330884	12403865	29	VWDE	
NR_136261	chr7	+	12496428	12541135	6	LOC102725191	
NR_136262	chr7	+	12497245	12541135	4	LOC102725191	
NR_136263	chr7	+	12504396	12541135	4	LOC102725191	
NR_156701	chr7	+	12570719	12660181	15	SCIN	
NM_001112706	chr7	+	12570719	12660181	16	SCIN	
NM_033128	chr7	+	12589522	12653603	14	SCIN	
NM_001037164	chr7	+	12686826	12690933	2	ARL4A	
NM_212460	chr7	+	12686826	12690933	2	ARL4A	
NM_001195396	chr7	+	12687285	12690933	2	ARL4A	
NM_005738	chr7	+	12687632	12690958	2	ARL4A	

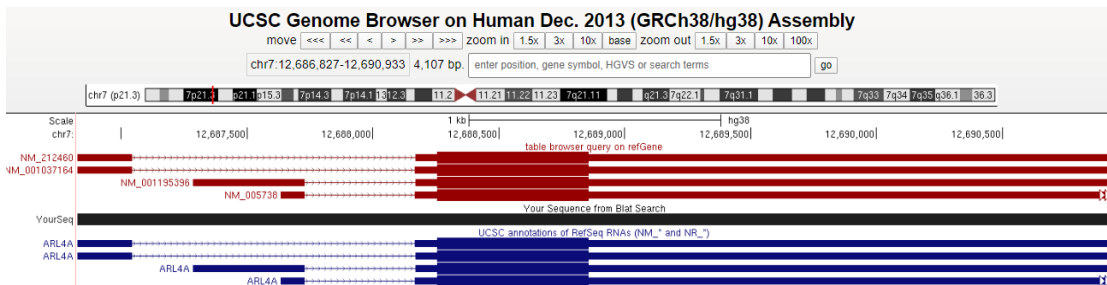


#chrom	name	strand	thickStart	thickEnd
chr7	NM_018374	+	12214810	12231975
chr7	NM_001134232	+	12214810	12231975
chr7	NM_001135924	-	12331182	12403716
chr7	NR_144534	-	12403941	12403941
chr7	NM_001346973	-	12331182	12388817
chr7	NM_001346972	-	12331182	12403716
chr7	NR_136261	+	12541135	12541135
chr7	NR_136262	+	12541135	12541135
chr7	NR_136263	+	12541135	12541135
chr7	NM_001112706	+	12570786	12652715
chr7	NR_156701	+	12660181	12660181
chr7	NM_033128	+	12622875	12652715
chr7	NM_212460	+	12688254	12688857
chr7	NM_001037164	+	12688254	12688857
chr7	NM_001195396	+	12688254	12688857
chr7	NM_005738	+	12688254	12688857

agafem, per exemple, chr7:12,688,254-12,688,857



Si fem BLAT:



Comprobem que hi ha coincidència.

## Exercici 4

Export all results to  TSV ☒ Unique results only

Email notification to

View 20 rows as HTML ☒ Unique results only

Gene description	Gene start (bp)	Gene end (bp)	Strand	Gene name	GO term name	Superfamily ID	Ensembl Protein Family ID(s)	GO term accession	GO term definition
replication protein A3 [Source: HGNC; Symbol: Acc: HGNC:10291]	7636518	7718607	-1	RPA3	DNA binding	SSF50249	PTHR15:14_SF1	GO:0003677	Any molecular function by which a interacts selectively and non-covalently with deoxyribonucleic acid.
replication protein A3 [Source: HGNC; Symbol: Acc: HGNC:10291]	7636518	7718607	-1	RPA3	nucleus	SSF50249	PTHR15:14_SF1	GO:0005634	A membrane-bounded organelle in which chromosomes are housed. In most cells, the nucleus contains a chromatin body, the nucleolus, and is the site of RNA synthesis and ribosome assembly. In some species, or in specialized cells, it is involved in metabolism or DNA replication.
replication protein A3 [Source: HGNC; Symbol: Acc: HGNC:10291]	7636518	7718607	-1	RPA3	DNA repair	SSF50249	PTHR15:14_SF1	GO:0006281	The process of restoring DNA after damage. Genomes are subject to damage by physical agents in the environment (ionizing radiation, chemical mutagens, bacterial toxins, etc.) and by free radicals and alkylating agents endogenously or metabolically. DNA is also damaged during its replication. A variety of

New

Count

Results

URL

XML

Perl

Help

Dataset

Human genes (GRCh38.p13)

Filters

Chromosome/scaffold: 7

Start: 7680240

End: 7905217

Attributes

Variant name

Karyotype band

Variant alleles

Variant consequence

Version (protein)

Version (transcript)

Gene name

Export all results to

File

TSV

Unique results only

Email notification to

View

20

rows as

HTML

Unique results only

Variant name	Karyotype band	Variant alleles	Variant consequence	Version (protein)	Version (transcript)	Gene name
rs1480289293	p21.3	T/C	3_prime_UTR_variant	4	8	RPA3
rs1382051635	p21.3	G/C	3_prime_UTR_variant	4	8	RPA3
rs1363066569	p21.3	T/C	3_prime_UTR_variant	4	8	RPA3
rs1174929906	p21.3	G/C	3_prime_UTR_variant	4	8	RPA3
rs10280185	p21.3	T/C	3_prime_UTR_variant	4	8	RPA3
rs1174356030	p21.3	C/A	3_prime_UTR_variant	4	8	RPA3
rs1375007161	p21.3	CA/CACA	3_prime_UTR_variant	4	8	RPA3
rs1436648954	p21.3	A/T	3_prime_UTR_variant	4	8	RPA3
rs1019056785	p21.3	T/C	3_prime_UTR_variant	4	8	RPA3
rs1401428341	p21.3	C/G	3_prime_UTR_variant	4	8	RPA3
rs532718808	p21.3	A/C	3_prime_UTR_variant	4	8	RPA3
rs1343704491	p21.3	T/C	3_prime_UTR_variant	4	8	RPA3

mart\_export (3): Bloc de notas

Archivo Edición Formato Ver Ayuda

Gene end (bp)	Strand	Gene name	GO term name	Superfamily ID	Ensembl Protein Family ID(s)	GO term acce
-1	RPA3	DNA binding	SSF50249	PTHR15114_SF1	GO:0003677	
-1	RPA3	nucleus	SSF50249	PTHR15114_SF1	GO:0005634	
-1	RPA3	DNA repair	SSF50249	PTHR15114_SF1	GO:0006281	
-1	RPA3	DNA replication	SSF50249	PTHR15114_SF1	GO:0006260	
-1	RPA3	DNA recombination	SSF50249	PTHR15114_SF1	GO:0006310	
-1	RPA3	protein binding	SSF50249	PTHR15114_SF1	GO:0005515	
-1	RPA3	cellular response to DNA damage stimulus	SSF50249	PTHR15114_SF1	GO:0006974	
-1	RPA3	single-stranded DNA binding	SSF50249	PTHR15114_SF1	GO:0003697	
-1	RPA3	mismatch repair	SSF50249	PTHR15114_SF1	GO:0006298	
-1	RPA3	nucleoplasm	SSF50249	PTHR15114_SF1	GO:0005654	
-1	RPA3	regulation of mitotic cell cycle	SSF50249	PTHR15114_SF1	GO:0007346	
-1	RPA3	regulation of cell population proliferation	SSF50249	PTHR15114_SF1	GO:0042127	
-1	RPA3	regulation of cellular response to heat	SSF50249	PTHR15114_SF1	GO:1900034	
-1	RPA3	interstrand cross-link repair	SSF50249	PTHR15114_SF1	GO:0036297	
-1	RPA3	base-excision repair	SSF50249	PTHR15114_SF1	GO:0006284	
-1	RPA3	regulation of signal transduction by p53 class mediator	SSF50249	PTHR15114_SF1	GO:1901796	
-1	RPA3	G1/S transition of mitotic cell cycle	SSF50249	PTHR15114_SF1	GO:0000082	
-1	RPA3	double-strand break repair via homologous recombination	SSF50249	PTHR15114_SF1	GO:0000724	
-1	RPA3	translesion synthesis	SSF50249	PTHR15114_SF1	GO:0019985	
-1	RPA3	transcription-coupled nucleotide-excision repair	SSF50249	PTHR15114_SF1	GO:0006283	
-1	RPA3	damaged DNA binding	SSF50249	PTHR15114_SF1	GO:0003684	
-1	RPA3	DNA replication factor A complex	SSF50249	PTHR15114_SF1	GO:0005663	