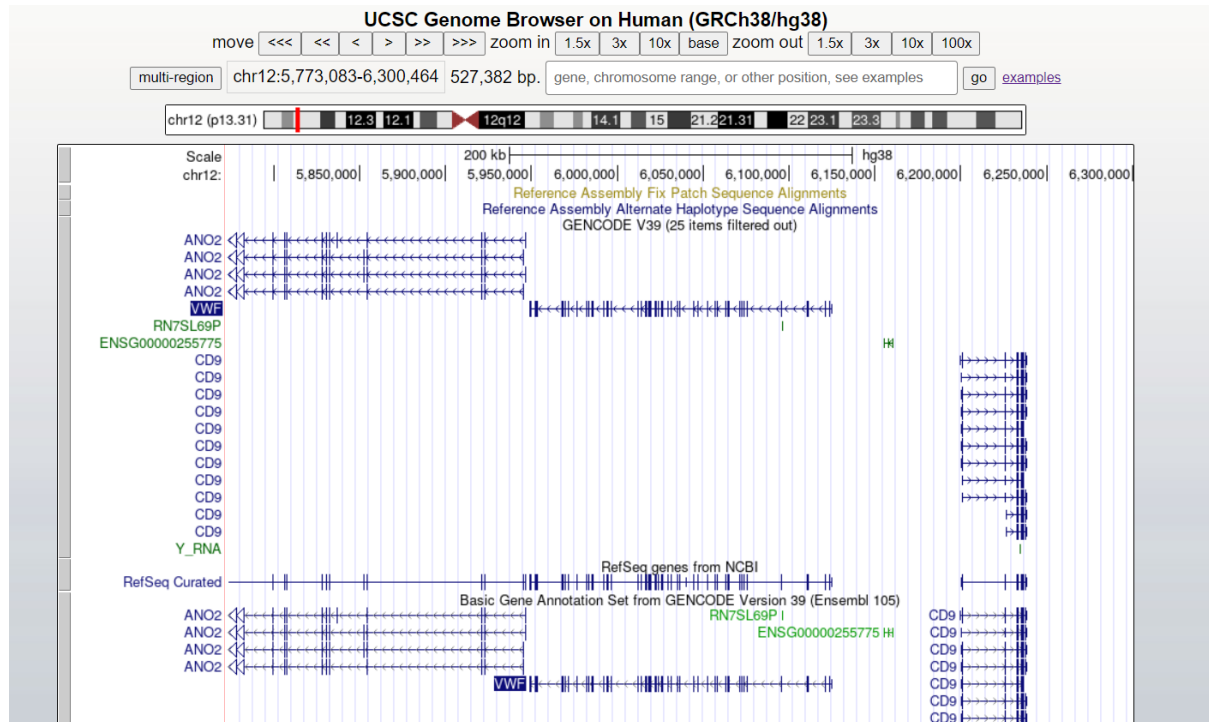


Отчет по ДЗ №2

Предмет: прикладная биоинформатика

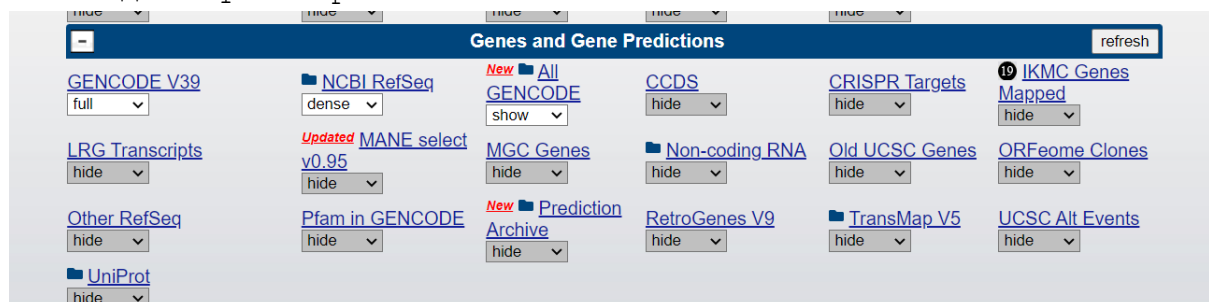
Исполнительница: Смолкина Ю.А

1.1. Выберите какой-нибудь белок человека и найдите информацию о гене этого белка в геномном браузере UCSC.



Укажите:

- имя гена - Human (GRCh38/hg38)
- идентификатор гена в Gencode - V39



- на какой цепи он закодирован - standard -

Human Gene ANO2 (ENST00000327087.12) from GENCODE V39

Description: Acts as a calcium-activated chloride channel (CaCC), mostly in photoreceptors. May mediate olfactory amplification in olfactory sensory neurons (OSNs) and light perception amplification in retina (By similarity). (from UniProt Q9NQ90)

RefSeq Summary (NM_001278597): ANO2 belongs to a family of calcium-activated chloride channels (CaCCs) (reviewed by Hartzell et al., 2009 [PubMed 19015192]).[supplied by OMIM, Jan 2011]. Sequence Note: The RefSeq transcript and protein were derived from genomic sequence to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on alignments.

Gencode Transcript: ENST00000327087.12

Gencode Gene: ENSG00000047617.17

Transcript (Including UTRs)

Position: hg38 chr12:5,562,651-5,946,232 **Size:** 383,582 **Total Exon Count:** 27 **Strand:** -

Coding Region

- в какой хромосоме находится - chr 12 - (см скрин сверху)
- к какому плечу и полосе принадлежит участок -chr12: p13.31
- сколько альтернативных продуктов (транскриптов) закодировано в гене;

и для каждого транскрипта укажите идентификатор Gencode, координаты в хромосоме (включая UTR'ы), число экзонов (общее) и длину последовательности белка



1ый

Description: Homo sapiens von Willebrand factor (VWF), mRNA. (from RefSeq NM_000552)

RefSeq Summary (NM_000552): This gene encodes a glycoprotein involved in hemostasis. The encoded preproprotein is proteolytically processed following assembly into large multimeric complexes. These complexes function in the adhesion of platelets to sites of vascular injury and the transport of various proteins in the blood. Mutations in this gene result in von Willebrand disease, an inherited bleeding disorder. An unprocessed pseudogene has been found on chromosome 22. [provided by RefSeq, Oct 2015].

Gencode Transcript: ENST00000261405.10

Gencode Gene: ENSG00000110799.14

Transcript (Including UTRs)

Position: hg38 chr12:5,948,877-6,124,670 **Size:** 175,794 **Total Exon Count:** 52 **Strand:** -

Coding Region

Position: hg38 chr12:5,949,015-6,123,196 **Size:** 174,182 **Coding Exon Count:** 51

Page Index	Sequence and Links	UniProtKB Comments	MalaCards	CTD	RNA-Seq Expression
Microarray Expression	RNA Structure	Protein Structure	Other Species	GO Annotations	mRNA Descriptions
Pathways	Other Names	GeneReviews	Methods		

Data last updated at UCSC: 2022-01-17 08:30:34

GENCODE V39 Transcript Annotation ENST00000356134.9 (ANO2)

GENCODE V39 Transcript Annotation ENST00000356134.9 (ANO2)

	Transcript	Gene
GENCODE id	ENST00000356134.9	ENSG00000047617.17
Protein id	ENSP00000348453.5	
HAVANA manual id	OTTHUMT00000399019.4	OTTHUMG00000168264.7
Position	chr12:5562654-5946232	chr12:5562651-5946232
Strand	-	
Biotype	protein_coding	protein_coding
Annotation Level	manual (2)	
Annotation Method	manual	manual & automatic
Transcription Support Level	tsl5	
HGNC gene symbol	ANO2	
CCDS	CCDS44807.2	
GeneCards	ANO2	
APPRIS		ANO2

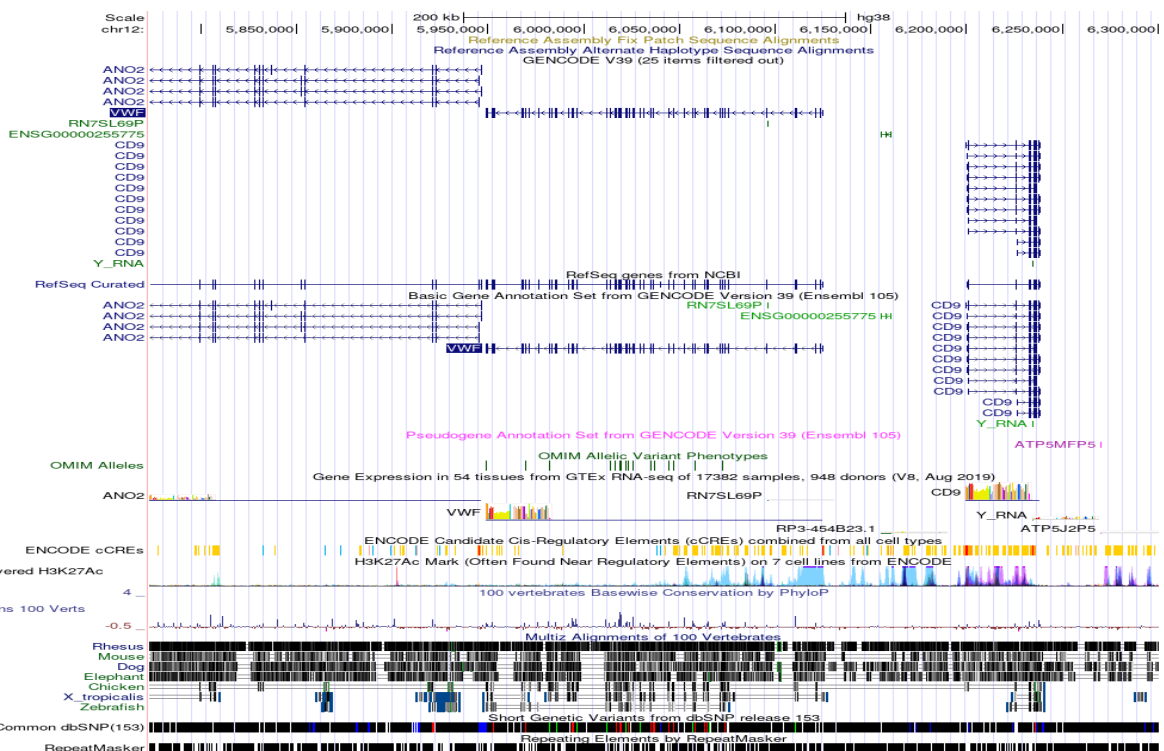
2ой

Tags

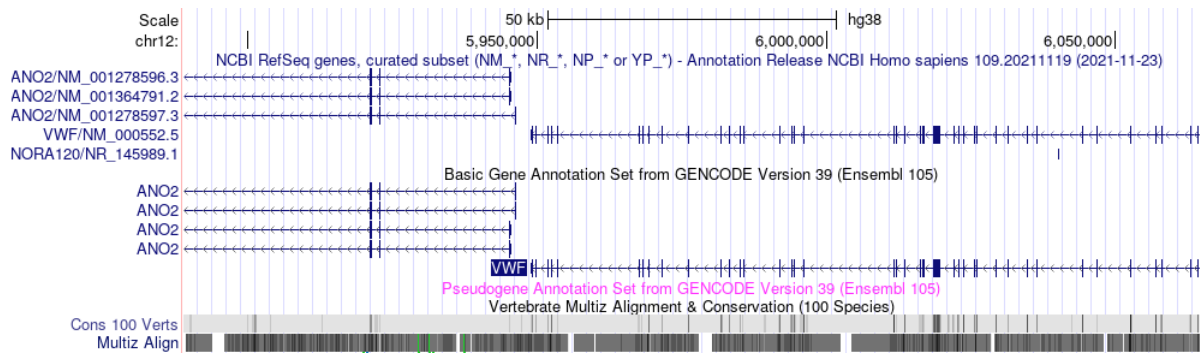
1.2. Сохраните картинку окрестности гена из Genome Browser с треками:

- транскрипты GENCODE и RefSeq,
- консервативность последовательности среди позвоночных (Conservation),
- частые полиморфизмы (Common SNPs) последней версии (153).

Все остальные треки скройте.



Через *view image*
после скрывтия



2. Ensembl

- 2.1 Постройте выравнивание выбранного гена человека с гомологичным геном шимпанзе (прикрепить к письму)
- 2.2 Оцените процент различий на 100 нуклеотидов.
- 2.3. Сравните с полногеномной оценкой. (источник указать)

Выбран Gene: GEM ENSG00000164949

Ensembl BLAST/BLAT | VEP | Tools | BioMart | Downloads | Help & Docs | Blog

Human (GRCh38.p13) Location: 8:94,249,253-94,262,350 Gene: GEM

Gene: GEM ENSG00000164949

Description: GTP binding protein overexpressed in skeletal muscle [Source: HGNC Symbol; Acc: HGNC:42346]

Gene Synonyms: KIR

Location: Chromosome 8: 94,249,253-94,262,350 reverse strand. GRCh38: CM000570.2

About this gene: This gene has 5 transcripts (splice variants), 202 orthologues and 35 paralogues.

Transcripts: [Show transcript table](#)

Genomic alignments

Alignment: [Select an alignment](#)

No alignment specified

Please select the alignment(s) to display from the table.

Select an alignment

Alignments Selector

Start typing the name of a species...

All Alignments

Multiple Pairwise

Multiple

- 9 primates EPO
 - ☐ --Ancestral_sequences--
 - ☐ Bonobo
 - ☒ Chimpanzee
 - ☐ Crab-eating macaque
 - ☐ Gibbon
 - ☐ Gorilla
 - Human
 - ☐ Macaque
 - ☐ Mouse Lemur
 - ☐ Vervet-AGM
- 24 primates EPO-Extended
- 43 eutherian mammals EPO
- 58 amniota vertebrates Mercator-Pecan
- 91 eutherian mammals EPO-Extended

Reset All Cancel Apply

Selected species

Chimpanzee X

https://dec2021.archive.ensembl.org/Homo_sapiens/Share/a791eefalbba8ff8cf5a918cf78a6555?redirect=no;mobileredirect=no

Скачен файл в формате фасты

С помощью команды infoalign пакета EMBOSS

Name	SeqLen	AlignLen	Gaps	GapLen	Ident	Similar	Differ	% Change	Weight	Description
homo_sapiens_1-17022	13098	17022	182	3924	13098	0	0	23.052521	1.000000	
pan_troglodytes_1-17022	13089	17022	179	3933	12948	0	141	23.933733	1.000000	

Gaps	GapLen	Ident	Similar	Differ	% Change
182	3924	13098	0	0	23.052521
179	3933	12948	0	141	23.933733

Как видно из чисел , различия небольшие
разница в Gaps ~2% Gaplen <1% Ident ~2%

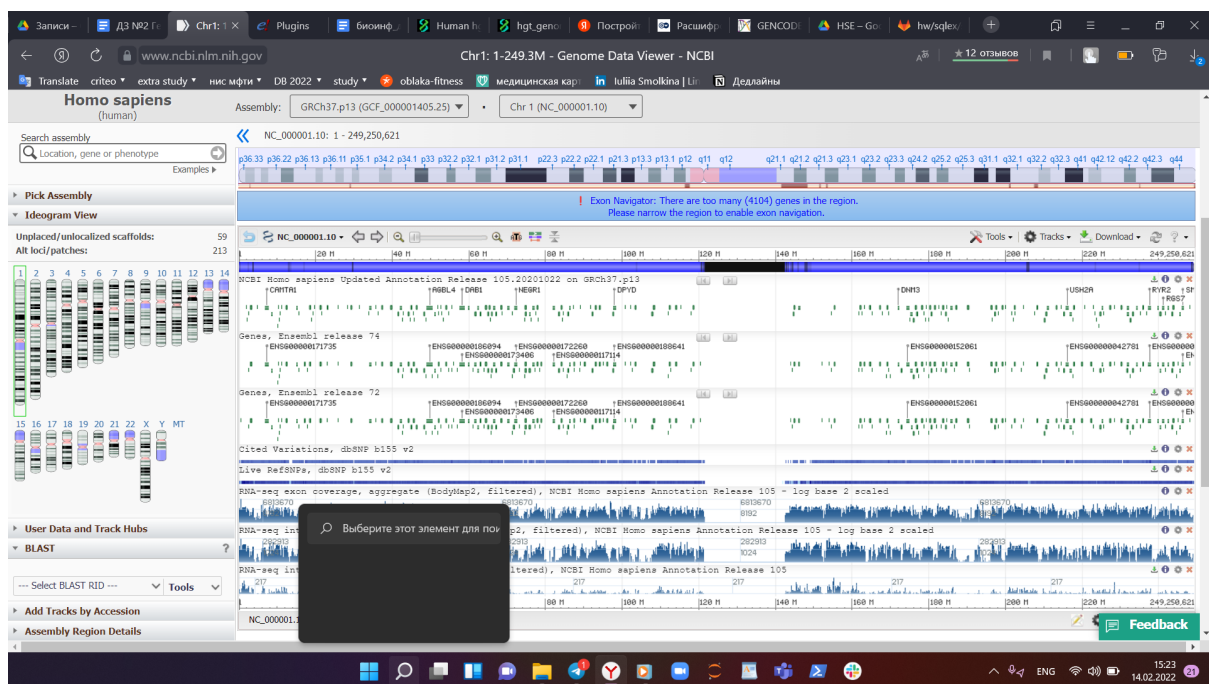
The difference between the two genomes ~4%—comprising ~35 million single nucleotide differences

<http://m.genome.cshlp.org/content/15/12/1746>

3. Genome Data Viewer

Рассмотрите тот же фрагмент генома в Genome Data Viewer.

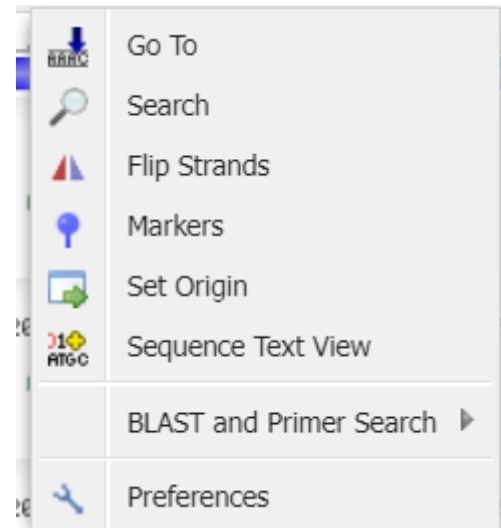
Опишите, в чем отличие от других браузеров, в чем он удобнее/неудобнее, какую информацию можно извлечь только из этого браузера.



Удобно : можно нажать share page и скопировав ссылку поделиться именно той частью последовательности, с которой работаешь.

Также можно сразу скачать в формате ФАСТА

Неудобно : Выбор того, что показывать удобнее в UCSC, тк не уходя со страницы генома tools довольно скучный , также я не нашла



аналогии кнопки Зум (см пример ниже)

