

## Часть 1

[Human Gene TTN \(ENST00000589042.5\) from GENCODE V44 \(ucsc.edu\)](#)

1.1 Имя гена: TTN

1.2 GeneCode ID: ENST00000589042.5

1.3 Цепь: -

1.4 Номер хромосомы: 2

1.5 Длинное плечо, 31.2

1.6 Кол-во альтернативных продуктов: 9

ENST00000604692.1

- Координаты hg38 chr2:178,541,125-178,541,799
- Кол-во экзонов 1
- Длина последовательности белка 675

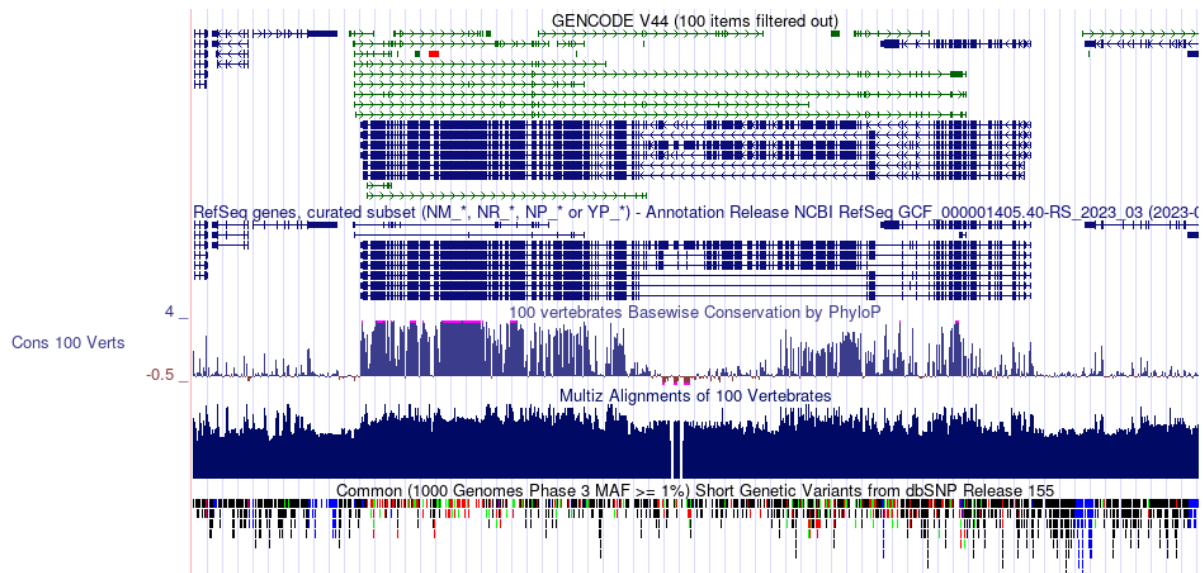
ENST00000603415.1

- Координаты hg38 chr2:178,548,884-178,550,681
- Кол-во экзонов 1
- Длина последовательности белка 1798

ENST00000624360.1

- Координаты hg38 chr2:178,554,561-178,558,966
- Кол-во экзонов 1
- Длина последовательности белка 4406

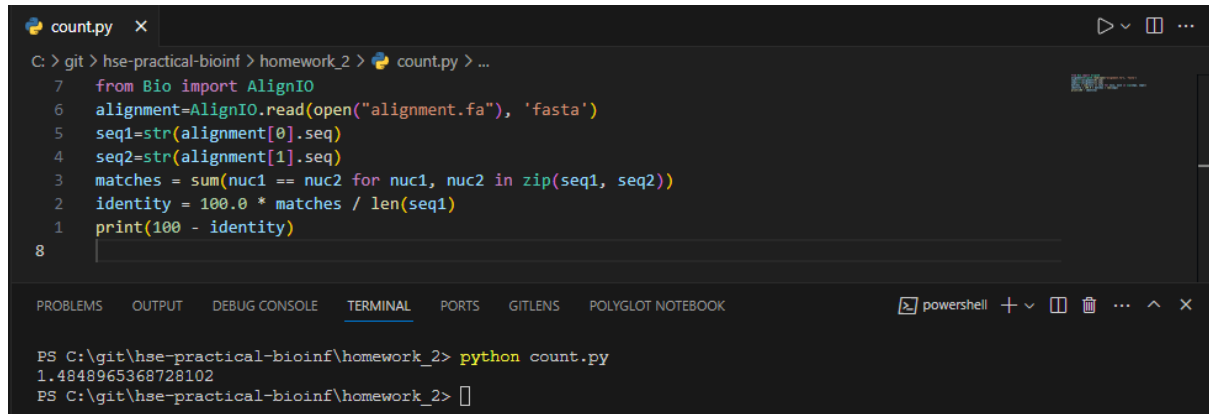
1.7



Самым информативным, на мой взгляд треком, является трек SNP155, на котором красным, согласно описанию в документации, указаны полиморфизмы в белок-кодирующих участках. Я бы использовал это для подтверждения гипотезы о том, что эластичность мышечной ткани (за которую отвечает этот белок) может сильно адаптироваться в зависимости от расположения этой ткани.

## Часть 2

### 2.2 Процент различий на 100 нуклеотидов: 1.49



```
count.py x
C: > git > hse-practical-bioinf > homework_2 > count.py > ...
7 from Bio import AlignIO
6 alignment=AlignIO.read(open("alignment.fa"), 'fasta')
5 seq1=str(alignment[0].seq)
4 seq2=str(alignment[1].seq)
3 matches = sum(nuc1 == nuc2 for nuc1, nuc2 in zip(seq1, seq2))
2 identity = 100.0 * matches / len(seq1)
1 print(100 - identity)
8

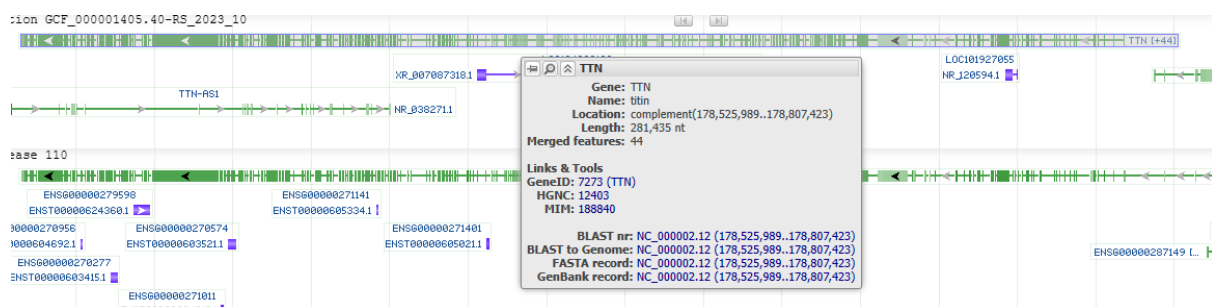
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS GITLENS POLYGLOT NOTEBOOK powershell + v [] [] ... ^ x

PS C:\git\hse-practical-bioinf\homework_2> python count.py
1.4848965368728102
PS C:\git\hse-practical-bioinf\homework_2> []
```

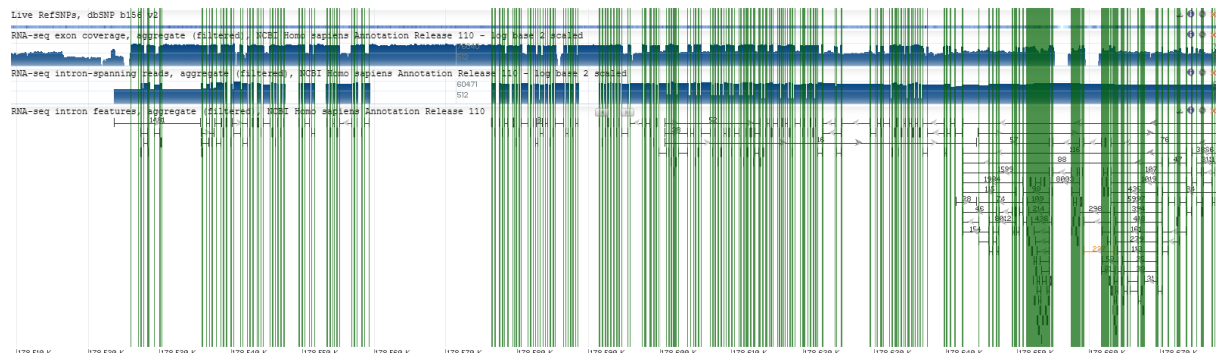
2.3 Различия геномов человека и шимпанзе оценивается в диапазоне от 1.23% до 3%. Я думаю, что близкое совпадение полученных оценок неслучайно: все-таки шимпанзе считается самым близким генетическим родственником человека. Также, рассматриваемый мной ген не является специфичным для человека, поэтому странно было бы ожидать каких-то больших различий.

## Часть 3

Мне понравилось, что в данном геномном браузере удобно сделан tooltip, с быстрой информацией об интересующем объекте.



Также понравились линии, отображающие начало и конец экзонов

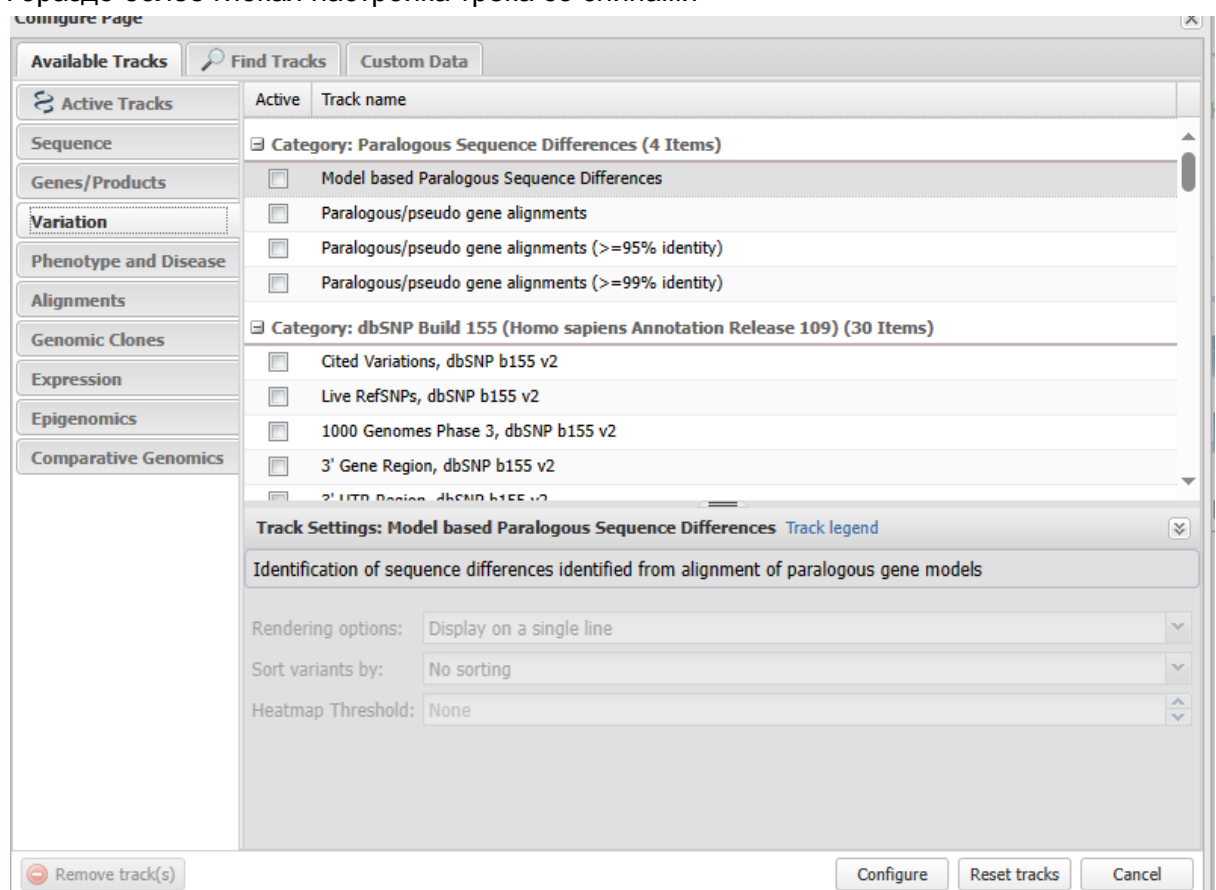


Также то, что вся информация уже загружена и треки недолго обновляются до новых координат.

Из минусов можно отметить то, что дорожки настраиваются не совсем очевидным образом.

Из полезной информации, которую я не нашел в других браузерах:

1. Гораздо более гибкая настройка трека со снипами



2. CpG островки и содержание GC

Category: Features (7 Items)

- ☐ Segmental Duplications, Eichler Lab
- ☒ G+C content (50 bp)
- ☒ CpG Islands
- ☐ Six-frame translations
- ☐ Repeats identified by WindowMasker
- ☐ Repeats identified by RepeatMasker
- ☐ Symmetric unique K-mers for GRCh38 human assembly