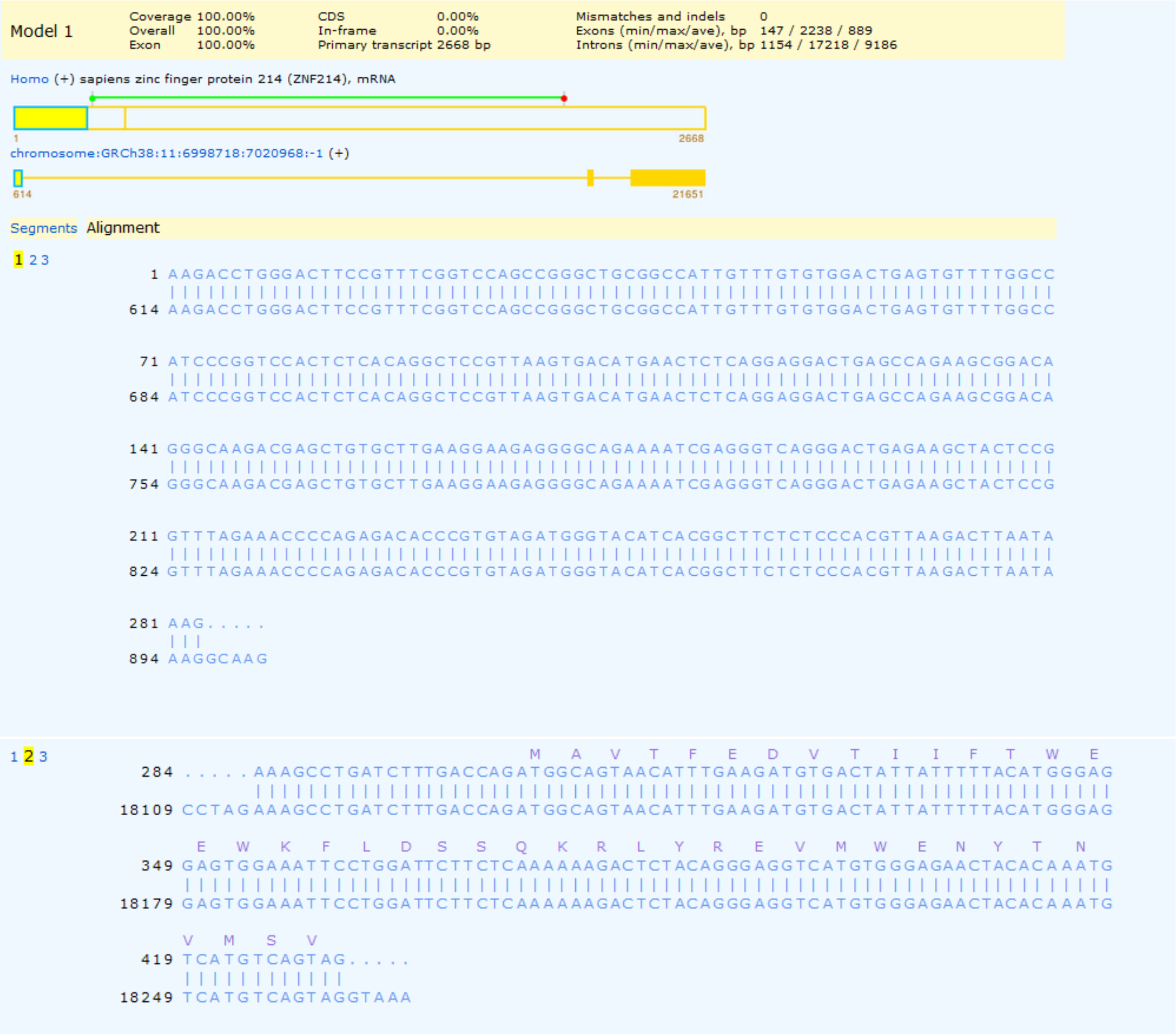
Genomics Spring 2021

Units 3-4 Graded Homework

(For the following problems, submit just one master pdf file and submit it along with other necessary files via Blackboard. Grades will be given mainly based on the pdf file. Screen capture or copy & paste all the results into the pdf file.)

1. 1.25 pts. There are two attached files: znf214_mrna.txt and znf214_genomic.txt. Use Splign to find the mRNA and CDS coordinates in the genomic DNA.

   a. 0.625 pts: Report mRNA locations: **1..283, 284..430, 431..2668**



Model 1

Coverage	100.00%	CDS	0.00%	Mismatches and indels	0
Overall	100.00%	In-frame	0.00%	Exons (min/max/ave), bp	147 / 2238 / 889
Exon	100.00%	Primary transcript	2668 bp	Introns (min/max/ave), bp	1154 / 17218 / 9186

Homo (+) sapiens zinc finger protein 214 (ZNF214), mRNA

chromosome:GRCh38:11:6998718:7020968:-1 (+)

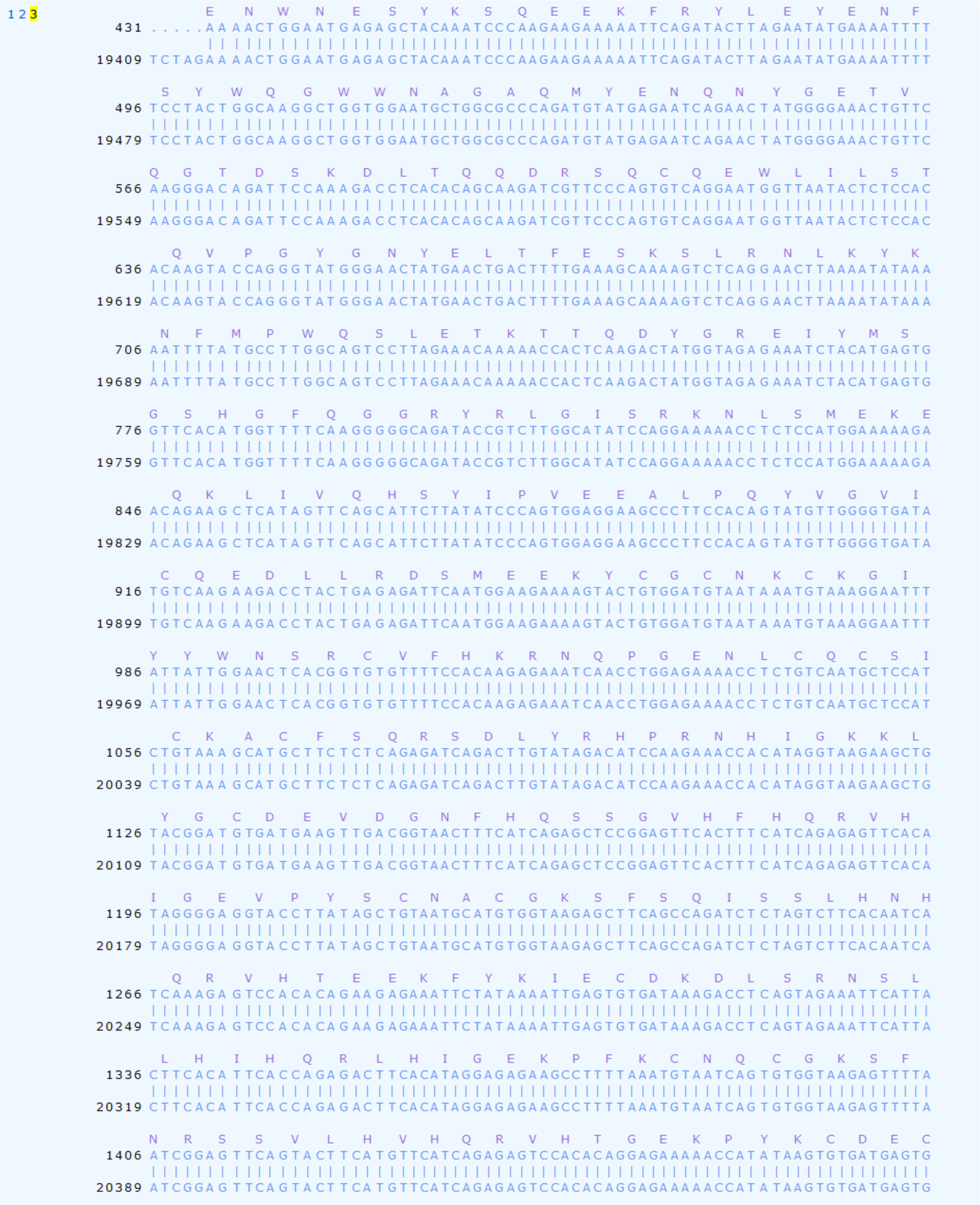
Segments Alignment

1 2 3

```
  1 AAGACCTGGGACTTCCGTTTCGGTCCAGCCGGGCTGCGGCCATTGTTTGTGTGGACTGAGTGTTTTGGCC
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
614 AAGACCTGGGACTTCCGTTTCGGTCCAGCCGGGCTGCGGCCATTGTTTGTGTGGACTGAGTGTTTTGGCC

 71 ATCCCGGTCCACTCTCACAGGCTCCGTTAAGTGACATGAACTCTCAGGAGGACTGAGCCAGAAGCGGACA
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
684 ATCCCGGTCCACTCTCACAGGCTCCGTTAAGTGACATGAACTCTCAGGAGGACTGAGCCAGAAGCGGACA

141 GGGCAAGACGAGCTGTGCTTGAAGGAAGAGGGGCAGAAAATCGAGGGTCAGGGACTGAGAAGCTACTCCG
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
754 GGGCAAGACGAGCTGTGCTTGAAGGAAGAGGGGCAGAAAATCGAGGGTCAGGGACTGAGAAGCTACTCCG

211 GTTTAGAAACCCCAGAGACACCCGTGTAGATGGGTACATCACGGCTTCTCTCCCACGTTAAGACTTAATA
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
824 GTTTAGAAACCCCAGAGACACCCGTGTAGATGGGTACATCACGGCTTCTCTCCCACGTTAAGACTTAATA

281 AAG.....
    |||
894 AAGGCAAG
```

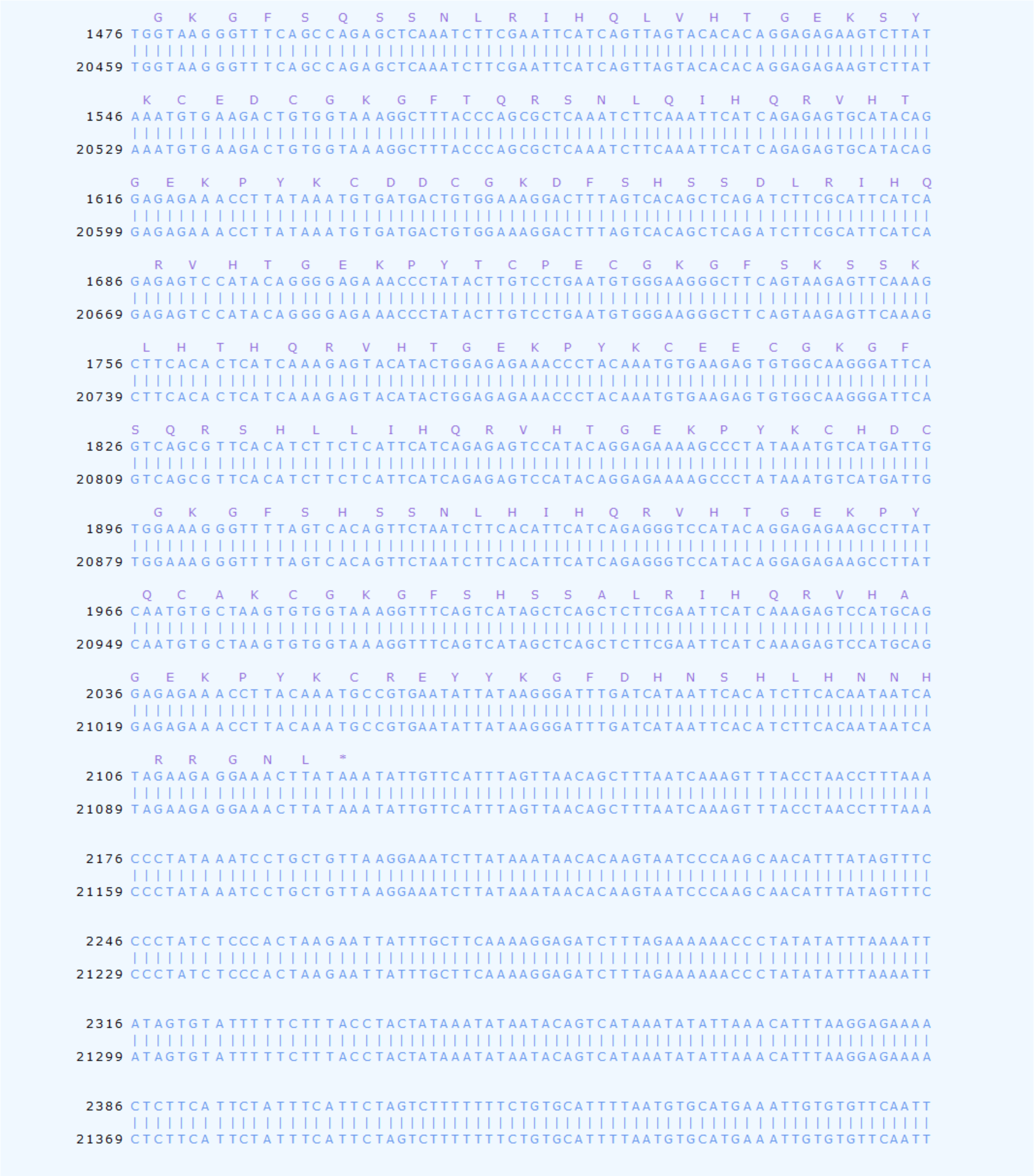
1 2 3

```
                                M  A  V  T  F  E  D  V  T  I  I  F  T  W  E
  284 .....AAAGCCTGATCTTTGACCAGATGGCAGTAACATTTGAAGATGTGACTATTATTTTTACATGGGAG
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
18109 CCTAGAAAGCCTGATCTTTGACCAGATGGCAGTAACATTTGAAGATGTGACTATTATTTTTACATGGGAG

       E  W  K  F  L  D  S  S  Q  K  R  L  Y  R  E  V  M  W  E  N  Y  T  N
  349 GAGTGGAAATTCCTGGATTCTTCTCAAAAAAGACTCTACAGGGAGGTCATGTGGGAGAACTACACAAATG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
18179 GAGTGGAAATTCCTGGATTCTTCTCAAAAAAGACTCTACAGGGAGGTCATGTGGGAGAACTACACAAATG

       V  M  S  V
  419 TCATGTCAGTAG.....
      ||||||||||||
18249 TCATGTCAGTAGGTAAA
```

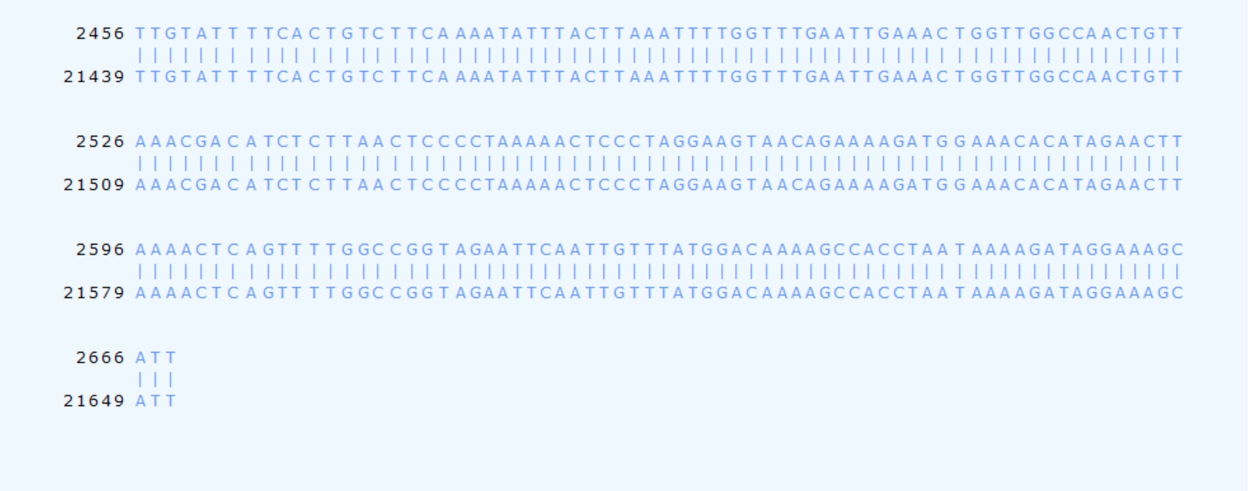


1 2 3

```
           E  N  W  N  E  S  Y  K  S  Q  E  E  K  F  R  Y  L  E  Y  E  N  F
  431 .....AAAACTGGAATGAGAGCTACAAATCCCAAGAAGAAAAATTCAGATACTTAGAATATGAAAATTTT
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
19409 TCTAGAAAACTGGAATGAGAGCTACAAATCCCAAGAAGAAAAATTCAGATACTTAGAATATGAAAATTTT

       S  Y  W  Q  G  W  W  N  A  G  A  Q  M  Y  E  N  Q  N  Y  G  E  T  V
  496 TCCTACTGGCAAGGCTGGTGGAATGCTGGCGCCCAGATGTATGAGAATCAGAACTATGGGGAAACTGTTC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
19479 TCCTACTGGCAAGGCTGGTGGAATGCTGGCGCCCAGATGTATGAGAATCAGAACTATGGGGAAACTGTTC

      Q  G  T  D  S  K  D  L  T  Q  Q  D  R  S  Q  C  Q  E  W  L  I  L  S  T
  566 AAGGGACAGATTCCAAAGACCTCACACAGCAAGATCGTTCCCAGTGTCAGGAATGGTTAATACTCTCCAC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
19549 AAGGGACAGATTCCAAAGACCTCACACAGCAAGATCGTTCCCAGTGTCAGGAATGGTTAATACTCTCCAC

        Q  V  P  G  Y  G  N  Y  E  L  T  F  E  S  K  S  L  R  N  L  K  Y  K
  636 ACAAGTACCAGGGTATGGGAACTATGAACTGACTTTTGAAAGCAAAAGTCTCAGGAACTTAAAATATAAA
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
19619 ACAAGTACCAGGGTATGGGAACTATGAACTGACTTTTGAAAGCAAAAGTCTCAGGAACTTAAAATATAAA

       N  F  M  P  W  Q  S  L  E  T  K  T  T  Q  D  Y  G  R  E  I  Y  M  S
  706 AATTTTATGCCTTGGCAGTCCTTAGAAACAAAAACCACTCAAGACTATGGTAGAGAAATCTACATGAGTG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
19689 AATTTTATGCCTTGGCAGTCCTTAGAAACAAAAACCACTCAAGACTATGGTAGAGAAATCTACATGAGTG

      G  S  H  G  F  Q  G  G  R  Y  R  L  G  I  S  R  K  N  L  S  M  E  K  E
  776 GTTCACATGGTTTTCAAGGGGGCAGATACCGTCTTGGCATATCCAGGAAAAACCTCTCCATGGAAAAAGA
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
19759 GTTCACATGGTTTTCAAGGGGGCAGATACCGTCTTGGCATATCCAGGAAAAACCTCTCCATGGAAAAAGA

        Q  K  L  I  V  Q  H  S  Y  I  P  V  E  E  A  L  P  Q  Y  V  G  V  I
  846 ACAGAAGCTCATAGTTCAGCATTCTTATATCCCAGTGGAGGAAGCCCTTCCACAGTATGTTGGGGTGATA
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
19829 ACAGAAGCTCATAGTTCAGCATTCTTATATCCCAGTGGAGGAAGCCCTTCCACAGTATGTTGGGGTGATA

       C  Q  E  D  L  L  R  D  S  M  E  E  K  Y  C  G  C  N  K  C  K  G  I
  916 TGTCAAGAAGACCTACTGAGAGATTCAATGGAAGAAAAGTACTGTGGATGTAATAAATGTAAAGGAATTT
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
19899 TGTCAAGAAGACCTACTGAGAGATTCAATGGAAGAAAAGTACTGTGGATGTAATAAATGTAAAGGAATTT

      Y  Y  W  N  S  R  C  V  F  H  K  R  N  Q  P  G  E  N  L  C  Q  C  S  I
  986 ATTATTGGAACTCACGGTGTGTTTTCCACAAGAGAAATCAACCTGGAGAAAACCTCTGTCAATGCTCCAT
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
19969 ATTATTGGAACTCACGGTGTGTTTTCCACAAGAGAAATCAACCTGGAGAAAACCTCTGTCAATGCTCCAT

        C  K  A  C  F  S  Q  R  S  D  L  Y  R  H  P  R  N  H  I  G  K  K  L
 1056 CTGTAAAGCATGCTTCTCTCAGAGATCAGACTTGTATAGACATCCAAGAAACCACATAGGTAAGAAGCTG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
20039 CTGTAAAGCATGCTTCTCTCAGAGATCAGACTTGTATAGACATCCAAGAAACCACATAGGTAAGAAGCTG

       Y  G  C  D  E  V  D  G  N  F  H  Q  S  S  G  V  H  F  H  Q  R  V  H
 1126 TACGGATGTGATGAAGTTGACGGTAACTTTCATCAGAGCTCCGGAGTTCACTTTCATCAGAGAGTTCACA
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
20109 TACGGATGTGATGAAGTTGACGGTAACTTTCATCAGAGCTCCGGAGTTCACTTTCATCAGAGAGTTCACA

      I  G  E  V  P  Y  S  C  N  A  C  G  K  S  F  S  Q  I  S  S  L  H  N  H
 1196 TAGGGGAGGTACCTTATAGCTGTAATGCATGTGGTAAGAGCTTCAGCCAGATCTCTAGTCTTCACAATCA
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
20179 TAGGGGAGGTACCTTATAGCTGTAATGCATGTGGTAAGAGCTTCAGCCAGATCTCTAGTCTTCACAATCA

        Q  R  V  H  T  E  E  K  F  Y  K  I  E  C  D  K  D  L  S  R  N  S  L
 1266 TCAAAGAGTCCACACAGAAGAGAAATTCTATAAAATTGAGTGTGATAAAGACCTCAGTAGAAATTCATTA
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
20249 TCAAAGAGTCCACACAGAAGAGAAATTCTATAAAATTGAGTGTGATAAAGACCTCAGTAGAAATTCATTA

       L  H  I  H  Q  R  L  H  I  G  E  K  P  F  K  C  N  Q  C  G  K  S  F
 1336 CTTCACATTCACCAGAGACTTCACATAGGAGAGAAGCCTTTTAAATGTAATCAGTGTGGTAAGAGTTTTA
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
20319 CTTCACATTCACCAGAGACTTCACATAGGAGAGAAGCCTTTTAAATGTAATCAGTGTGGTAAGAGTTTTA

      N  R  S  S  V  L  H  V  H  Q  R  V  H  T  G  E  K  P  Y  K  C  D  E  C
 1406 ATCGGAGTTCAGTACTTCATGTTCATCAGAGAGTCCACACAGGAGAAAAACCATATAAGTGTGATGAGTG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
20389 ATCGGAGTTCAGTACTTCATGTTCATCAGAGAGTCCACACAGGAGAAAAACCATATAAGTGTGATGAGTG
```



```
        G  K  G  F  S  Q  S  S  N  L  R  I  H  Q  L  V  H  T  G  E  K  S  Y
 1476 TGGTAAGGGTTTCAGCCAGAGCTCAAATCTTCGAATTCATCAGTTAGTACACACAGGAGAGAAGTCTTAT
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
20459 TGGTAAGGGTTTCAGCCAGAGCTCAAATCTTCGAATTCATCAGTTAGTACACACAGGAGAGAAGTCTTAT

       K  C  E  D  C  G  K  G  F  T  Q  R  S  N  L  Q  I  H  Q  R  V  H  T
 1546 AAATGTGAAGACTGTGGTAAAGGCTTTACCCAGCGCTCAAATCTTCAAATTCATCAGAGAGTGCATACAG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
20529 AAATGTGAAGACTGTGGTAAAGGCTTTACCCAGCGCTCAAATCTTCAAATTCATCAGAGAGTGCATACAG

      G  E  K  P  Y  K  C  D  D  C  G  K  D  F  S  H  S  S  D  L  R  I  H  Q
 1616 GAGAGAAACCTTATAAATGTGATGACTGTGGAAAGGACTTTAGTCACAGCTCAGATCTTCGCATTCATCA
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
20599 GAGAGAAACCTTATAAATGTGATGACTGTGGAAAGGACTTTAGTCACAGCTCAGATCTTCGCATTCATCA

        R  V  H  T  G  E  K  P  Y  T  C  P  E  C  G  K  G  F  S  K  S  S  K
 1686 GAGAGTCCATACAGGGGAGAAACCCTATACTTGTCCTGAATGTGGGAAGGGCTTCAGTAAGAGTTCAAAG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
20669 GAGAGTCCATACAGGGGAGAAACCCTATACTTGTCCTGAATGTGGGAAGGGCTTCAGTAAGAGTTCAAAG

       L  H  T  H  Q  R  V  H  T  G  E  K  P  Y  K  C  E  E  C  G  K  G  F
 1756 CTTCACACTCATCAAAGAGTACATACTGGAGAGAAACCCTACAAATGTGAAGAGTGTGGCAAGGGATTCA
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
20739 CTTCACACTCATCAAAGAGTACATACTGGAGAGAAACCCTACAAATGTGAAGAGTGTGGCAAGGGATTCA

      S  Q  R  S  H  L  L  I  H  Q  R  V  H  T  G  E  K  P  Y  K  C  H  D  C
 1826 GTCAGCGTTCACATCTTCTCATTCATCAGAGAGTCCATACAGGAGAAAAGCCCTATAAATGTCATGATTG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
20809 GTCAGCGTTCACATCTTCTCATTCATCAGAGAGTCCATACAGGAGAAAAGCCCTATAAATGTCATGATTG

        G  K  G  F  S  H  S  S  N  L  H  I  H  Q  R  V  H  T  G  E  K  P  Y
 1896 TGGAAAGGGTTTTAGTCACAGTTCTAATCTTCACATTCATCAGAGGGTCCATACAGGAGAGAAGCCTTAT
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
20879 TGGAAAGGGTTTTAGTCACAGTTCTAATCTTCACATTCATCAGAGGGTCCATACAGGAGAGAAGCCTTAT

       Q  C  A  K  C  G  K  G  F  S  H  S  S  A  L  R  I  H  Q  R  V  H  A
 1966 CAATGTGCTAAGTGTGGTAAAGGTTTCAGTCATAGCTCAGCTCTTCGAATTCATCAAAGAGTCCATGCAG
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
20949 CAATGTGCTAAGTGTGGTAAAGGTTTCAGTCATAGCTCAGCTCTTCGAATTCATCAAAGAGTCCATGCAG

      G  E  K  P  Y  K  C  R  E  Y  Y  K  G  F  D  H  N  S  H  L  H  N  N  H
 2036 GAGAGAAACCTTACAAATGCCGTGAATATTATAAGGGATTTGATCATAATTCACATCTTCACAATAATCA
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
21019 GAGAGAAACCTTACAAATGCCGTGAATATTATAAGGGATTTGATCATAATTCACATCTTCACAATAATCA

        R  R  G  N  L  *
 2106 TAGAAGAGGAAACTTATAAATATTGTTCATTTAGTTAACAGCTTTAATCAAAGTTTACCTAACCTTTAAA
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
21089 TAGAAGAGGAAACTTATAAATATTGTTCATTTAGTTAACAGCTTTAATCAAAGTTTACCTAACCTTTAAA


 2176 CCCTATAAATCCTGCTGTTAAGGAAATCTTATAAATAACACAAGTAATCCCAAGCAACATTTATAGTTTC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
21159 CCCTATAAATCCTGCTGTTAAGGAAATCTTATAAATAACACAAGTAATCCCAAGCAACATTTATAGTTTC


 2246 CCCTATCTCCCACTAAGAATTATTTGCTTCAAAAGGAGATCTTTAGAAAAAACCCTATATATTTAAAATT
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
21229 CCCTATCTCCCACTAAGAATTATTTGCTTCAAAAGGAGATCTTTAGAAAAAACCCTATATATTTAAAATT


 2316 ATAGTGTATTTTTCTTTACCTACTATAAATATAATACAGTCATAAATATATTAAACATTTAAGGAGAAAA
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
21299 ATAGTGTATTTTTCTTTACCTACTATAAATATAATACAGTCATAAATATATTAAACATTTAAGGAGAAAA


 2386 CTCTTCATTCTATTTCATTCTAGTCTTTTTTTCTGTGCATTTTAATGTGCATGAAATTGTGTGTTCAATT
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
21369 CTCTTCATTCTATTTCATTCTAGTCTTTTTTTCTGTGCATTTTAATGTGCATGAAATTGTGTGTTCAATT
```



```
2456  TTGTATTTTCACTGTCTTCAAAATATTTACTTAAATTTTGGTTTGAATTGAAACTGGTTGGCCAACTGTT
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
21439 TTGTATTTTCACTGTCTTCAAAATATTTACTTAAATTTTGGTTTGAATTGAAACTGGTTGGCCAACTGTT

2526  AAACGACATCTCTTAACTCCCCTAAAAACTCCCTAGGAAGTAACAGAAAAGATGGAAACACATAGAACTT
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
21509 AAACGACATCTCTTAACTCCCCTAAAAACTCCCTAGGAAGTAACAGAAAAGATGGAAACACATAGAACTT

2596  AAAACTCAGTTTTGGCCGGTAGAATTCAATTGTTTATGGACAAAAGCCACCTAATAAAAGATAGGAAAGC
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
21579 AAAACTCAGTTTTGGCCGGTAGAATTCAATTGTTTATGGACAAAAGCCACCTAATAAAAGATAGGAAAGC

2666  ATT
      |||
21649 ATT
```

b. 0.625 pts: Report CDS locations: **614..891, 18109..18265, 19409..21651**

2. 1.25 pts. Create a BED6 file with 2 lines based on the attached paper (Takenaka_et_al-2015-FEBS_Journal.pdf). Figure 3 shows the location of transcription factor DdlR binding to the promoter region of the ddlR-ddl operon in Brevibacillus brevis. The chromosomal location of the ddlR CDS is 2968133..2969623. The zero-based BED6 file should contain the location information of two genomic regions:

a. 0.625 pts: The region bound by the DdlR transcription factor, which we will call the promoter. It is 170 bp in length, begins 140 nucleotides upstream from the start codon, and ends 29 nucleotides downstream from the start codon.

2968133-140 = 2967993(start)   2968133 + 29 = 2968162(end) *need 0base*

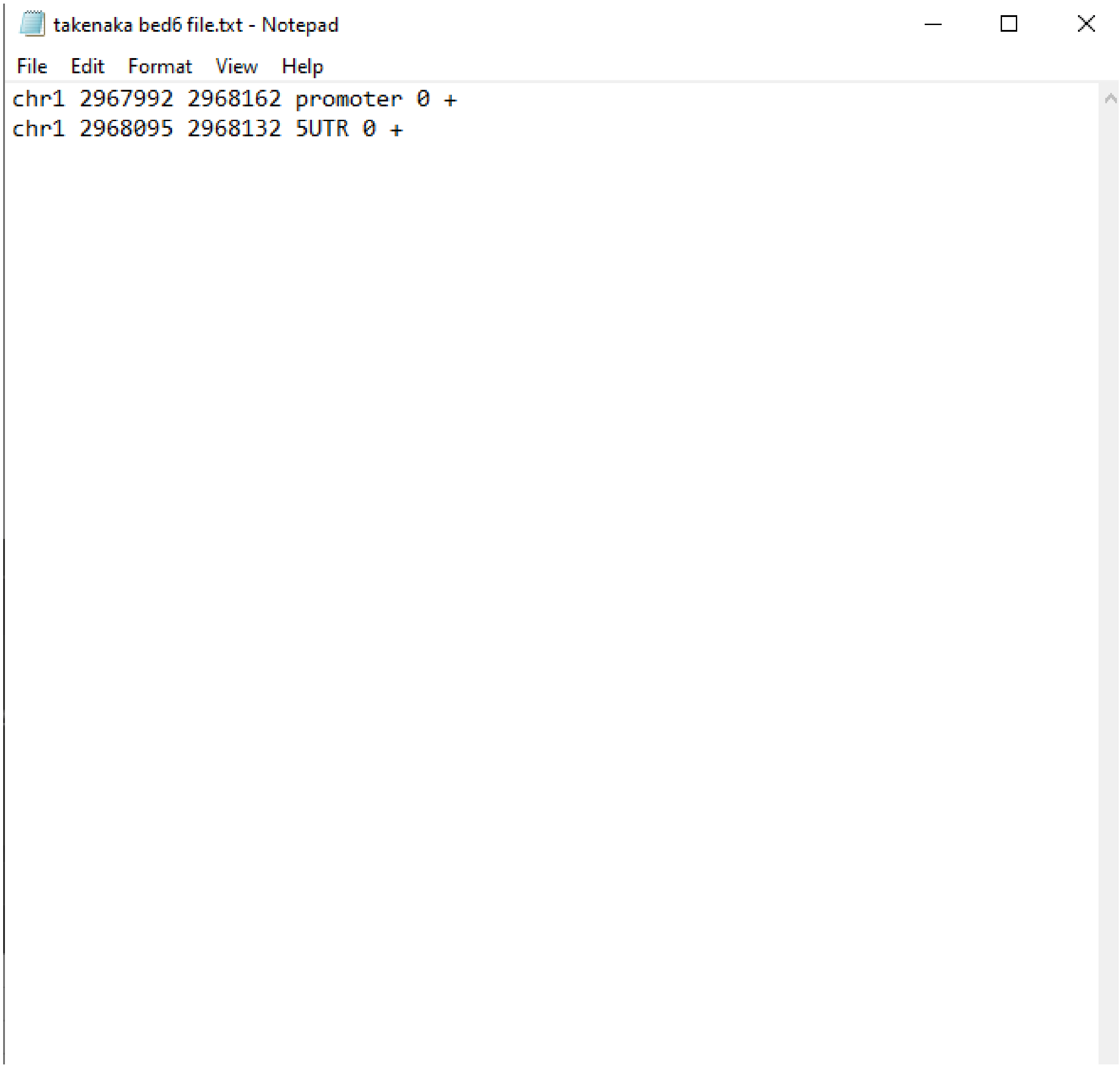
2968162 – 2967993 = 169(…reading graded homework collaboration in slack to attempt to determine why missing one)

chr1 2967992 2968162 promoter 0 +

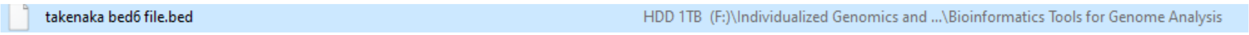
b. 0.625 pts: The 5' UTR, noting that the transcription start site, as predicted by BPROM, begins 38 nucleotides upstream from the start codon. The 5' UTR is defined as the region from the transcription start site through the nucleotide that immediately precedes the start codon.

2968133 – 38 = 2968095(start)   2968132(end)

2968132-2968095= 37+1 = 38bp length of 5’ UTR chr1 2968095 2968132 5UTR 0 +



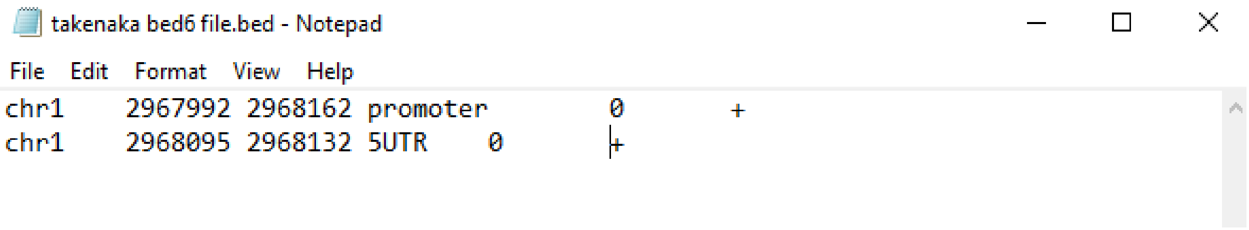
```
chr1 2967992 2968162 promoter 0 +
chr1 2968095 2968132 5UTR 0 +
```


takenaka bed6 file.bed HDD 1TB (F:)\Individualized Genomics and ...\Bioinformatics Tools for Genome Analysis

saved these two lines in this .bed file for question 3

3. 1.25 pts. Submit a screenshot of the BED6 from Problem 2. Using the NCBI Genome Browser for Brevibacillus brevis NBRC 100599, load your BED6 file. Take a screenshot showing the entire promoter, 5' UTR region, and CDS of ddlR. Be sure to zoom in so that these regions take up a majority of the shot.

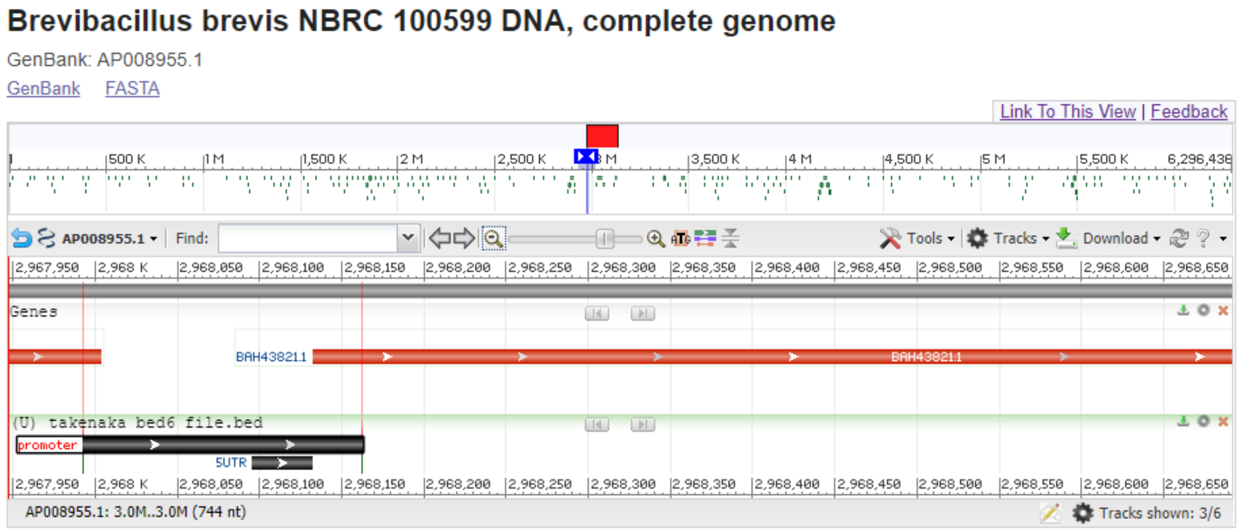
I forgot to set the spaces between each individual item in the .bed file as a ‘tab’, gave me quite the headache until I figured this out!



takenaka bed6 file.bed - Notepad

```
chr1	2967992 2968162 promoter	0	+
chr1	2968095 2968132 5UTR	0	+
```

https://www.ncbi.nlm.nih.gov/nuccore/226092535?report=graph&tracks=[key:sequence_track,name:Sequence,display_name:Sequence,id:STD649220238,annots:Sequence,ShowLabel:false,ColorGaps:false,shown:true,order:1][key:gene_model_track,name:Genes,display_name:Genes,id:STD3194982005,annots:Unnamed,Options:MergeAll,CDSProductFeats:false,NtRuler:true,AaRuler:true,HighlightMode:2,ShowLabel:true,shown:true,order:5][key:feature_track,name:U1CtIwAlPC90B,display_name:(U) takenaka bed6 file.bed,id:U1CtIwAlPC90B,data_key:LZ63RzGc7kVCsqBCYaOWvMYJ7eGykLyVsJOYhYyBnq8Po1Mp6ROZzlp6b0ZN1OOpsrHvpfGBqoTtnvmT_5XzjsGn_KnQlfo,subkey:region,dbname:NetCache,annots:takenaka bed6 file.bed_UUD1614012630DUU_region,Layout:Adaptive,LinkedFeat:Packed,shown:true,order:6]&assm_context=GCA_000010165.1&v=2967880:2968251&c=FF0000&select=gi|226092535-002d49b8-002d4a61-0153-df2ec903-2e039e4c;&slim=0



Brevibacillus brevis NBRC 100599 DNA, complete genome

GenBank: AP008955.1

GenBank FASTA

Link To This View | Feedback

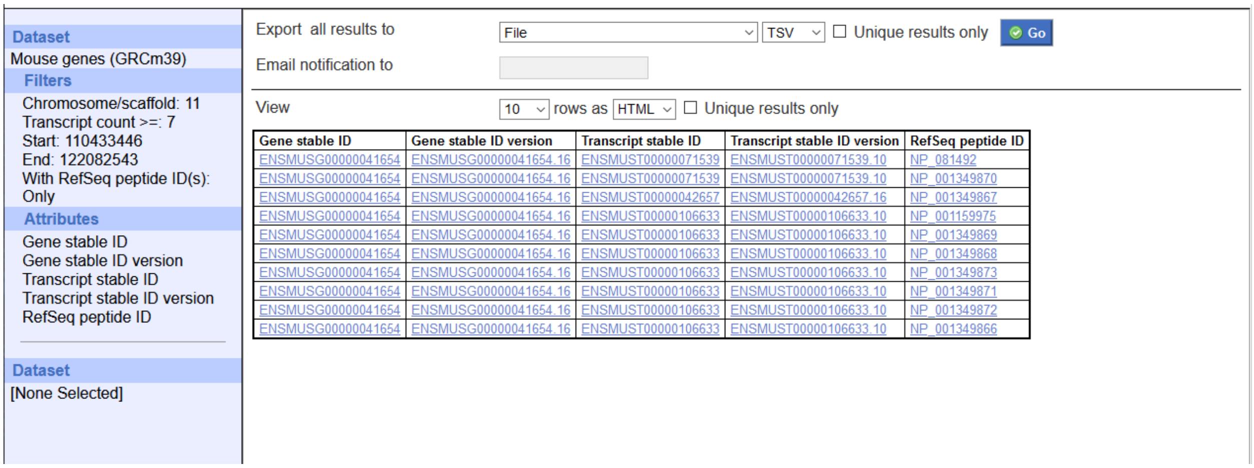
AP008955.1: 3.0M..3.0M (744 nt)

Tracks shown: 3/6

4. 1.25 pts. Use the web-based Biomart in Ensembl to create a dataset and save it as a TSV, CSV, or XLS file. Use the following parameters to make the dataset:

Dataset:
Ensembl Genes 100 (or the latest version)
Mouse genes (GRCm38.p6) (or the latest version)
Filters:
Chromosome 11
Band E2 only (found via ensembl)
Transcript count >=7
Limit to genes with RefSeq protein (peptide) IDs only
Attributes:
Default attributes
Add “RefSeq Protein (peptide) ID”

Get all the results, export the results to a file, and submit the file.



Dataset
Mouse genes (GRCm39)
Filters
Chromosome/scaffold: 11
Transcript count >=: 7
Start: 110433446
End: 122082543
With RefSeq peptide ID(s): Only
Attributes
Gene stable ID
Gene stable ID version
Transcript stable ID
Transcript stable ID version
RefSeq peptide ID

Dataset
[None Selected]

Export all results to File TSV ☐ Unique results only Go

Email notification to

View 10 rows as HTML ☐ Unique results only

Gene stable ID	Gene stable ID version	Transcript stable ID	Transcript stable ID version	RefSeq peptide ID
ENSMUSG00000041654	ENSMUSG00000041654.16	ENSMUST00000071539	ENSMUST00000071539.10	NP_081492
ENSMUSG00000041654	ENSMUSG00000041654.16	ENSMUST00000071539	ENSMUST00000071539.10	NP_001349870
ENSMUSG00000041654	ENSMUSG00000041654.16	ENSMUST00000042657	ENSMUST00000042657.16	NP_001349867
ENSMUSG00000041654	ENSMUSG00000041654.16	ENSMUST00000106633	ENSMUST00000106633.10	NP_001159975
ENSMUSG00000041654	ENSMUSG00000041654.16	ENSMUST00000106633	ENSMUST00000106633.10	NP_001349869
ENSMUSG00000041654	ENSMUSG00000041654.16	ENSMUST00000106633	ENSMUST00000106633.10	NP_001349868
ENSMUSG00000041654	ENSMUSG00000041654.16	ENSMUST00000106633	ENSMUST00000106633.10	NP_001349873
ENSMUSG00000041654	ENSMUSG00000041654.16	ENSMUST00000106633	ENSMUST00000106633.10	NP_001349871
ENSMUSG00000041654	ENSMUSG00000041654.16	ENSMUST00000106633	ENSMUST00000106633.10	NP_001349872
ENSMUSG00000041654	ENSMUSG00000041654.16	ENSMUST00000106633	ENSMUST00000106633.10	NP_001349866

File will be attached to the submission of this assignment!

(Hint) Band E2 is not available directly from Ensembl anymore. But, you may search for other databases, and filter via coordinates. Or you can use Ensembl archive with the specific archive version number.

(Optional) If you are doing this for human genes, you can still filter by karyoband. In this case, try with p12 band.