

Mutant X- A Bioinformatics Exercise

- 1.) Choose 3 genes and identify

Gene 3: Homo sapiens huntingtin (HTT), transcript variant 2, mRNA

```
ATTGCCCGGTGCTGAGCGGCCGCGAGTCGGCCGAGGCCTCCGGGACTGCCG  
TGCCGGCGGGAGA  
CCGCCATGGCGACCCCTGGAAAAGCTGATGAAGGCCTCGAGTCCCTCAAGTCCTC  
CAGCAGCAGCAGCA  
GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA  
GCAGCAGCAGCAGCAGCAGCAGCAG  
CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG  
CAGCAGCAGCAGCAACAGCCGCCAC CGCCGCCG
```

Gene 4:

Homo sapiens BRCA1 DNA repair associated (BRCA1), transcript variant 1, mRNA

```
CTTAGCGGTAGCCCCTGGTTCCGTGGCAACGGAAAAGCGCGGAAATTACAG  
ATAAATTAAAATCGCA  
CTGCGCGCGTGGACTCGCTGAGACTCCTGGACGGGGACAGGCTGTGGGGT  
TTCTCAGATAACTGGGC  
CCCTGCCTCAGGAGGCCTCACCCCTGCTCTGGTAAAGTTCATGGAACAG  
AAAGAAATGGATTAT  
CTGCTCTCGCGTTGAAGAAGTACAAAAGGTCTTAATGCTATGCAGAAAATCT  
TAGAGTGTCCCCTG
```

Protein 5: M-phase inducer phosphatase 2 isoform 1 (Homo Sapiens)

```
dgfvfkmpwk pthpssthal aewasrreaf agrpssapdl mclspdrkme veelsplalg  
rfsltpaegd teeddgfvdi lesdlkddd vppgmeslis aplvktleke eekdlvmyesk  
cqrlfrspsm pcsvirpilk rlerpqdrdt pvqnkrrrsv tppeeqqeae epkarvlqsk  
slchdeienl lsdhrelig dyskaflqt vdgkhqdlky ispetmvall tgkfsnivdk  
fvivdcrypy eyegghikta vnlplerdae sflkspiap csldkrvili fhcefsserg
```

Identity of genes:

Gene 3: Homo sapiens huntingtin (HTT), transcript variant 2, mRNA
Scientific Name: Homo sapiens

Gene 4: Homo sapiens BRCA1 DNA repair associated (BRCA1), transcript variant 1, mRNA
Scientific Name: Homosapiens

Protein 5: M-phase inducer phosphatase 2 isoform 1 (Homo Sapiens)

Part 1 Question

1. Where is the mutation located and what is the nature of the mutation? (example substitution, nonsense mutation, deletion, insertion).
 - Gene 3: 187. Substitution mutation
 - Gene 4: 238. Substitution mutation
 - Protein 5: 178. Insertion mutation

Accession numbers:

- Gene 3: NM_002111.8
- Gene 4: NM_007294.4
- Protein 5: NM_021873

Part 2:

1. In regards to genomic DNA, would you expect to see introns or exons or both?
 - We would expect to see both introns and exons, it will have more exons than introns though.
2. How many introns and exons are in the first transcript listed?
 - a. Gene 3: 67 exons introns 35
 - b. Gene 4: 22 exons and 40 introns
 - c. Protein 5: 16 exons
3. What is the length of the mRNA transcript?
 - a. Gene 3: 13498 bp
 - b. Gene 4: 7088 bp
 - c. Protein 5: 3119 bp
4. How many amino acids are there?
 - a. Gene 3: 3144 aa
 - b. Gene 4: 1863 aa
 - c. Protein 5: 580 aa
5. What is the molecular weight of the protein?
 - a. Gene 3: 347730
 - b. Gene 4: 207591
 - c. Protein 5: 64587

Part 3:

1. How were the gene mutations initially identified?
 - a. The gene mutation was initially identified in NCBI BLAST by finding the different nucleotide pairing.

2. State which diseases you can find this mutated gene
 - a. Gene 3: Huntington
 - b. Gene 4: breast-ovarian cancer syndrome
 - c. Protein 5: Retinoblastoma and Breast Cancer.
3. What chromosome is the gene located on?
 - a. Gene 3: Chromosome 4
 - b. Gene 4: Chromosome 17
 - c. Protein 5: Chromosome 20
4. What is the function of the normal gene?
 - a. Gene 3: The HTT gene encodes huntingtin, a ubiquitously expressed nuclear protein that binds to a number of transcription factors to regulate transcription.
 - b. Gene 4: BRCA1 plays critical roles in DNA repair, cell cycle checkpoint control, and maintenance of genomic stability. BRCA1 forms several distinct complexes through association with different adaptor proteins, and each complex forms in a mutually exclusive manner
 - c. Protein 5: The cell division cycle 25 (CDC25) family of proteins are highly conserved dual specificity phosphatases that activate cyclin-dependent kinase (CDK) complexes, which in turn regulate progression through the cell division cycle. CDC25 phosphatases are also key components of the checkpoint pathways that become activated in the event of DNA damage. In mammalian cells, 3 isoforms have been identified: CDC25A (116947), CDC25B, and CDC25C (157680). CDC25B and CDC25C are required primarily for entry into mitosis. CDC25B is proposed to be responsible for the initial activation of CDK1 (116940)-cyclin B (123836) at the centrosome during the G2-M transition, which is then followed by a complete activation of CDK1-cyclin B complexes by CDC25C in the nucleus at the onset of mitosis (summary by Boutros et al., 2007).

Multiple Sequence Alignment and Analysis

Part 1: BLAST and PSI-BLAST in the MIP family

When BLAST is finished, have a look at the names of the domain hits and families BLAST returns; are they all from the AQPs and GLPs sub-family? If not, what do you think is happening here?

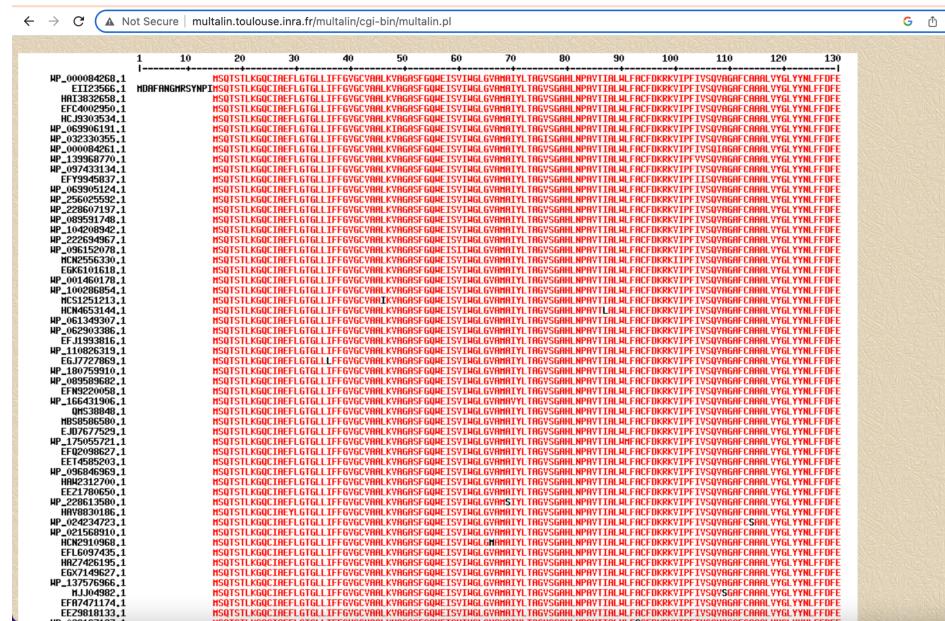
- No they are not all from AQPs and GLPs sub-family, what could be happening here is variants due to evolution or mutations.

From the Blast output, select all AQPs 1(Aquaporin-1), AQPs 2(Aquaporin-2) and GLPs (glycerol facilitators) proteins and download them as FASTA.

Part 2: Alignment and Quality Scoring

1. Multiple Aligning

- Used CLUSTAL, MUSCLE, COBALT, Multalin (Praline was not working)



Results for job clustalo-l20230612-131245-0857-96863420-p1m

[Alignments](#) [Result Summary](#) [Guide Tree](#) [Phylogenetic Tree](#) [Results Viewers](#) [Submission Details](#)

[Download Alignment File](#) [Show Colors](#)

CLUSTAL O (1.2.4) multiple sequence alignment

```

HAN3763063.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
HAW2312700.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
EET4585203.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
HNC2190968.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
WP_14871274.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
E121217478.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
EPR472685.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGRWEISV 46
WP_000084633.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
EFL6097435.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
HA2742095.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
MCS1251213.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
WP_148711804.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
WP_14871274.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
HAW1880889.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
WP_13756966.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
WP_137518125.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
EPR4097407.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
HON4209060.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
EBB7655455.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
EBBS072902.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
WP_13756966.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
WP_137518125.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
HAW1091877.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
EFT00185315.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
WP_024213623.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
EFTK514385.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
WP_021293120.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
WP_074491225.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46
WP_175127226.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWKISV 46
WP_176065270.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWKISV 46
EJD7677529.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV 46

```

Results for job muscle-l20230612-131251-0088-84383656-p1m

[Alignments](#) [Result Summary](#) [Phylogenetic Tree](#) [Results Viewers](#) [Submission Details](#)

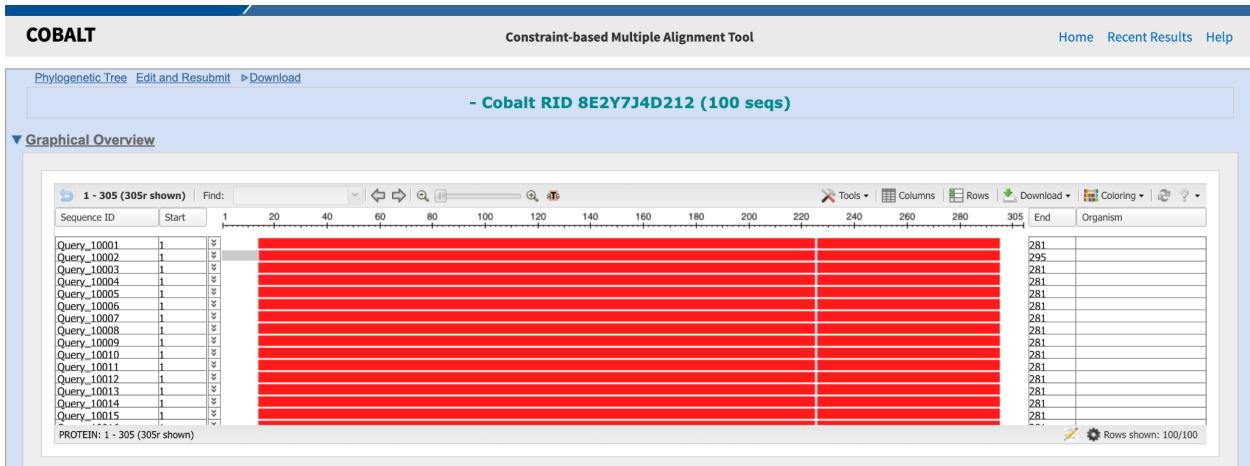
[Download Alignment File](#) [Show Colors](#)

CLUSTAL multiple sequence alignment by MUSCLE (3.8)

```

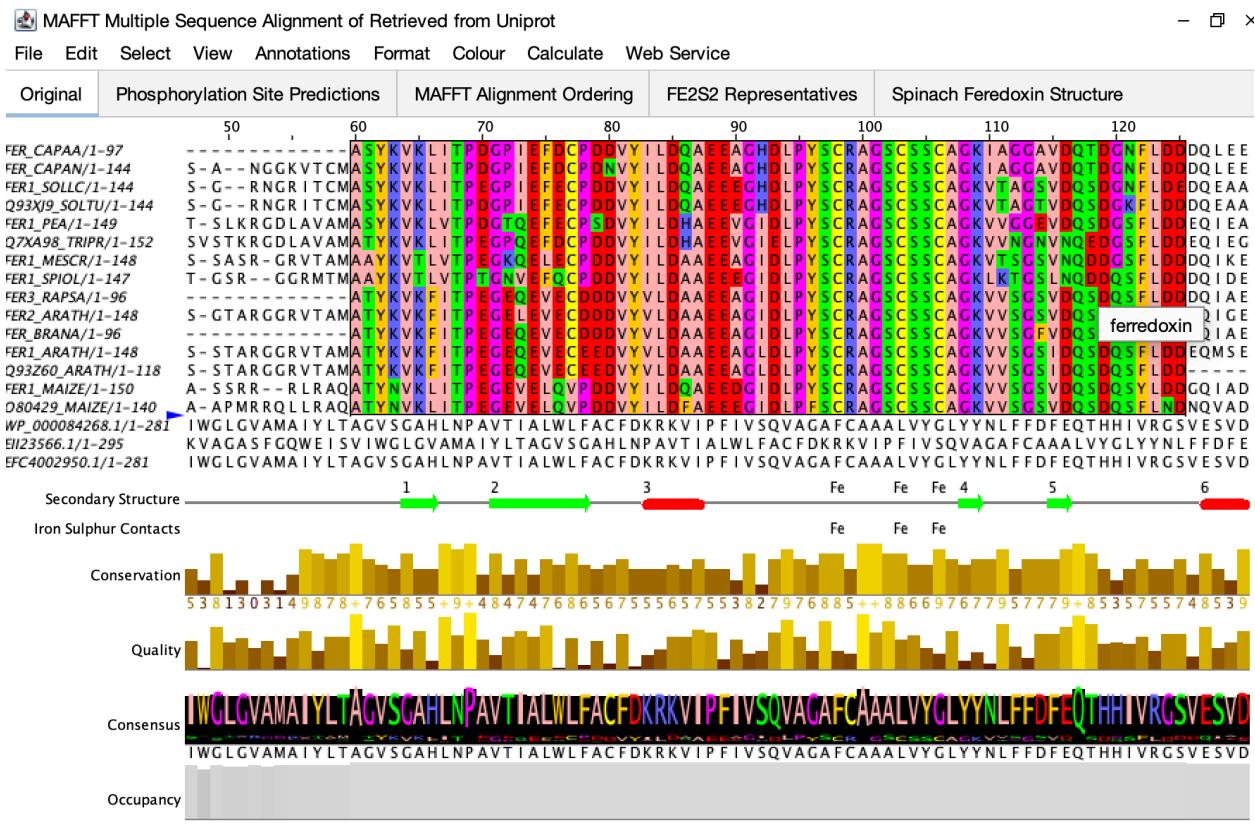
HBC1484665.1 -----MSKTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
HAW1880889.1 -----MNQQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
E1123566.1 -----MDAFANGHRSYNNMSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
HAW2312700.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
EET4585203.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
HNC2190968.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
WP_000084633.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
EFL6097435.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
HA27426195.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
MCS1251213.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
WP_14871274.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
WP_137518125.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
EPR4097407.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
HON4209060.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
MCN4029060.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
EES072902.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
WP_13756966.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
WP_052896116.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
WP_094171713.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
HAN4091877.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
WP_021293120.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
EFA7471174.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
MBS8586580.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
WP_096846969.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
WP_053886403.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
EFG2098627.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
EE21780650.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
WP_228613580.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
HAV8830186.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
WP_024234723.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
WP_021568910.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV
WP_175055721.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWKISV
EJD7677529.1 -----MSQTSTLKGQCIAEFLGTGLLIPFGVGCVAALKVAGASFGQWEISV

```



2. Alignment Visualization

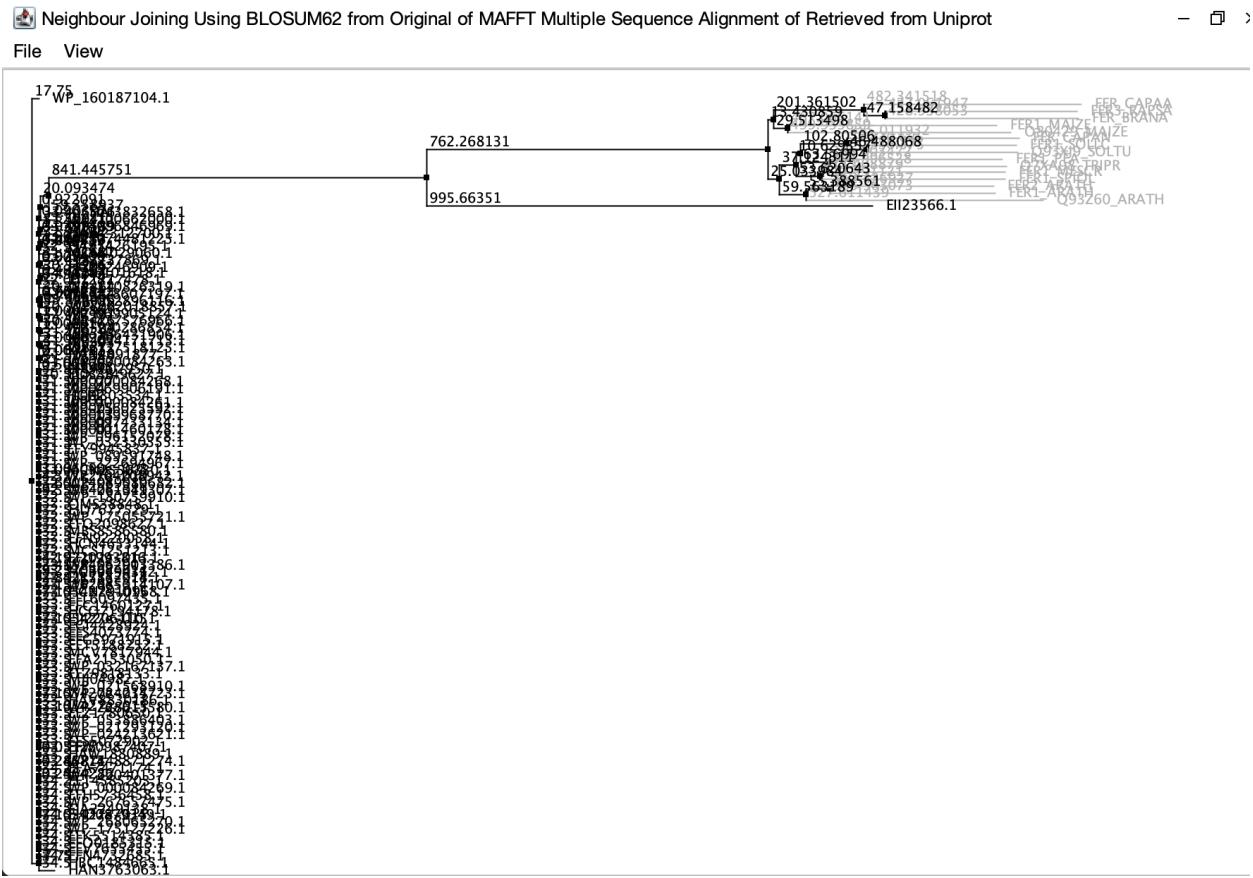
- Used Jalview



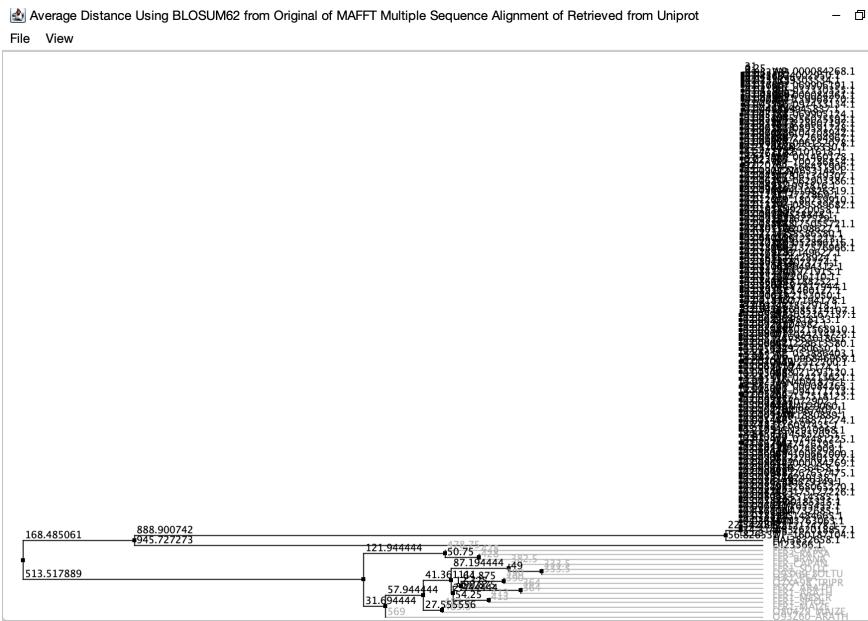
3. Optional- Assessing Alignment Quality

Part 3: Phylogeny and Specificity Scoring

1. Phylogeny vs Functional Annotations
 - Neighbor Joining Tree



- Average Distance Tree



Calculate NJ and UPGMA trees for the alignments. Compare the trees you get. Pay attention to how sequences of the 3 different MIP subfamilies are distributed along the phylogenetic tree.

Are all members of each subfamily grouped together in a single clade?

- Yes they are grouped together

Are subfamily clades side-by-side in the tree, or are subfamilies in a sub-clade of another subfamily?

- They are mostly side-by-side in the tree

2. Calculating Specificity

- Used Multiharmony



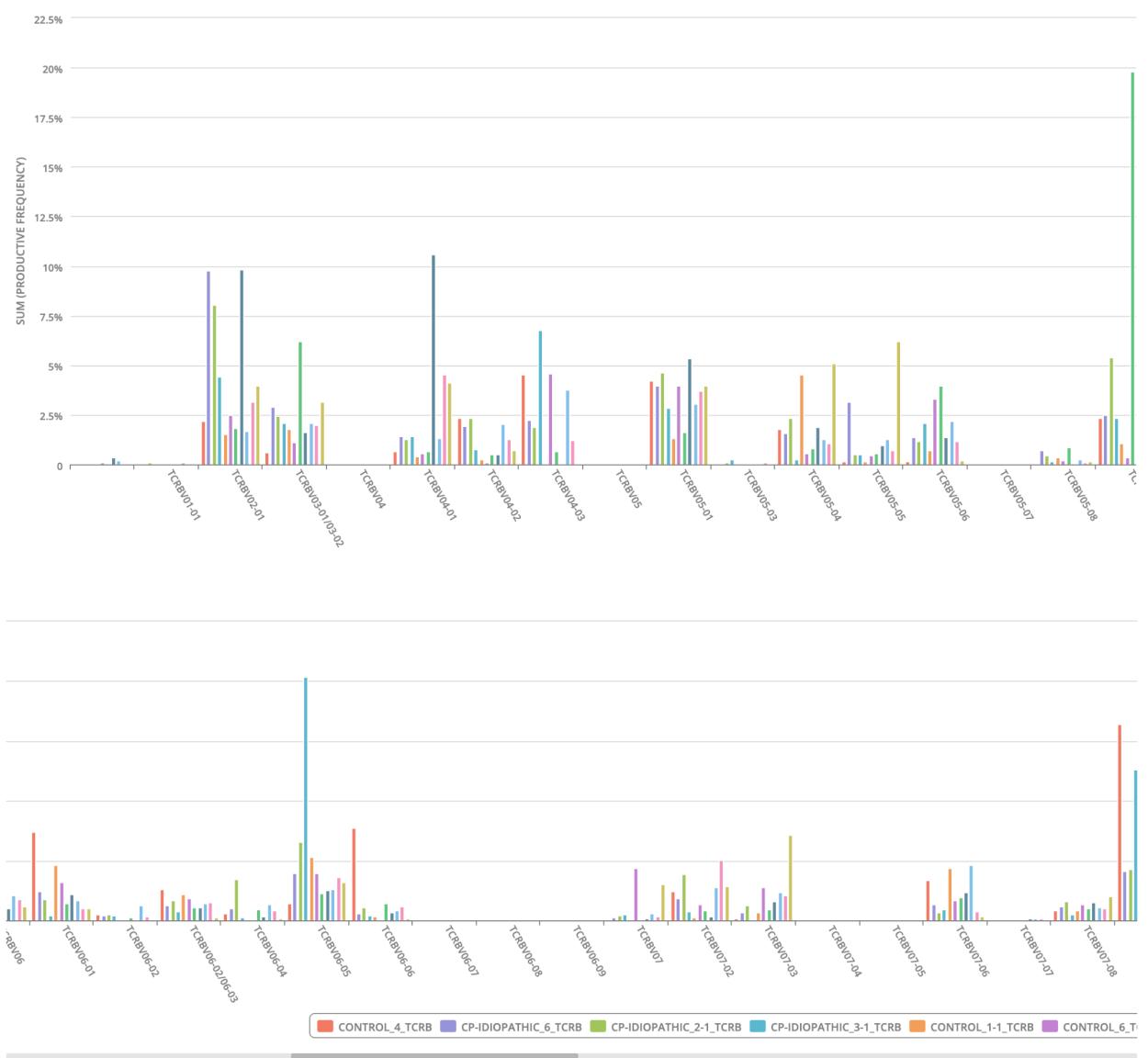
3. Evaluating Specificity Detection

- Used SH and ROC plots

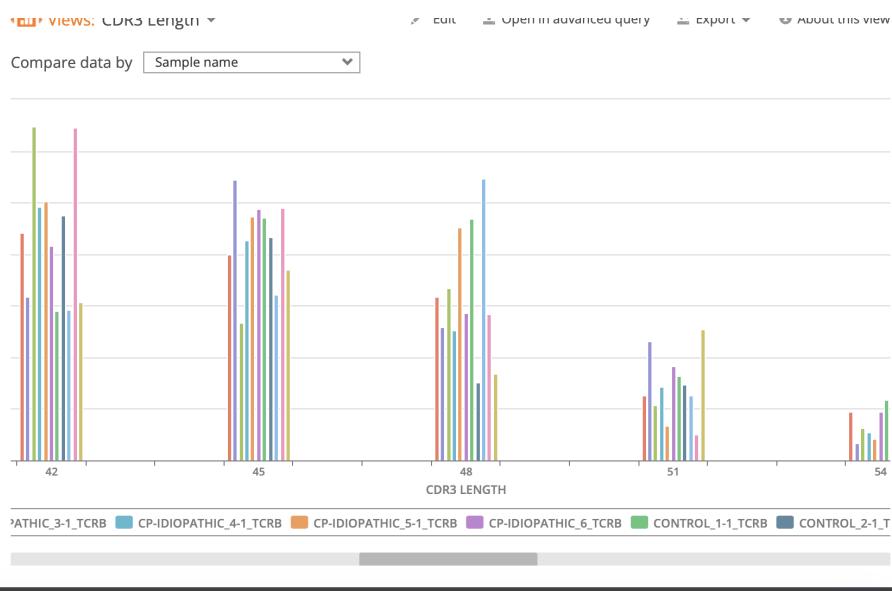
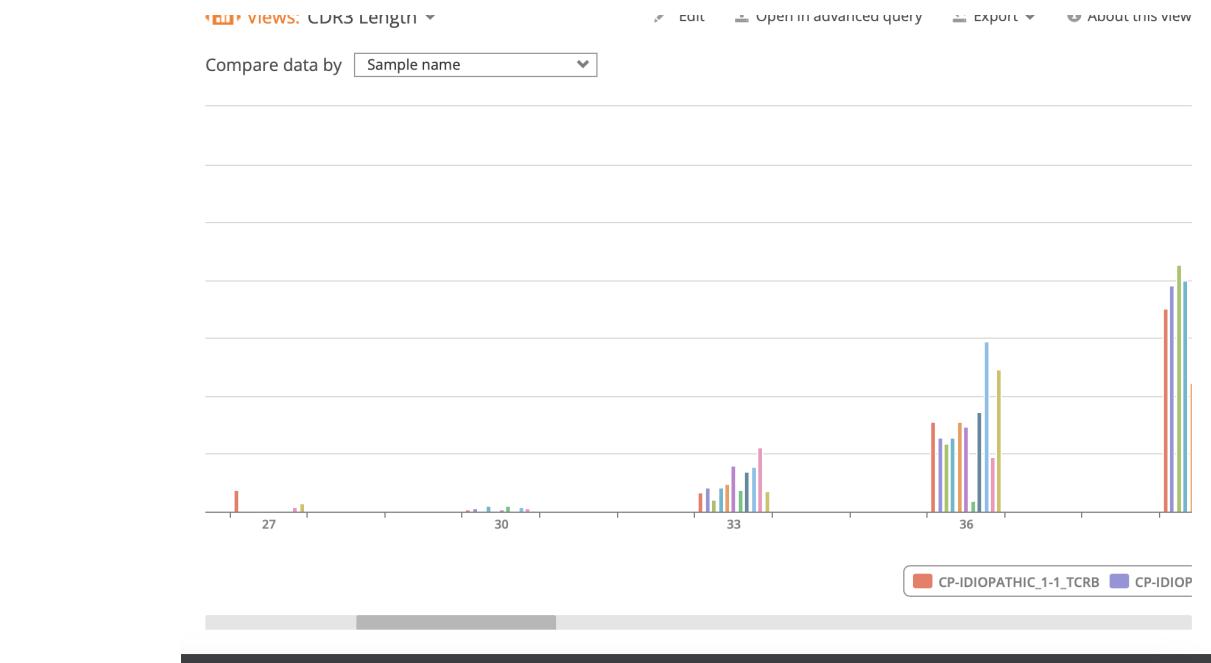
Immune Sequencing

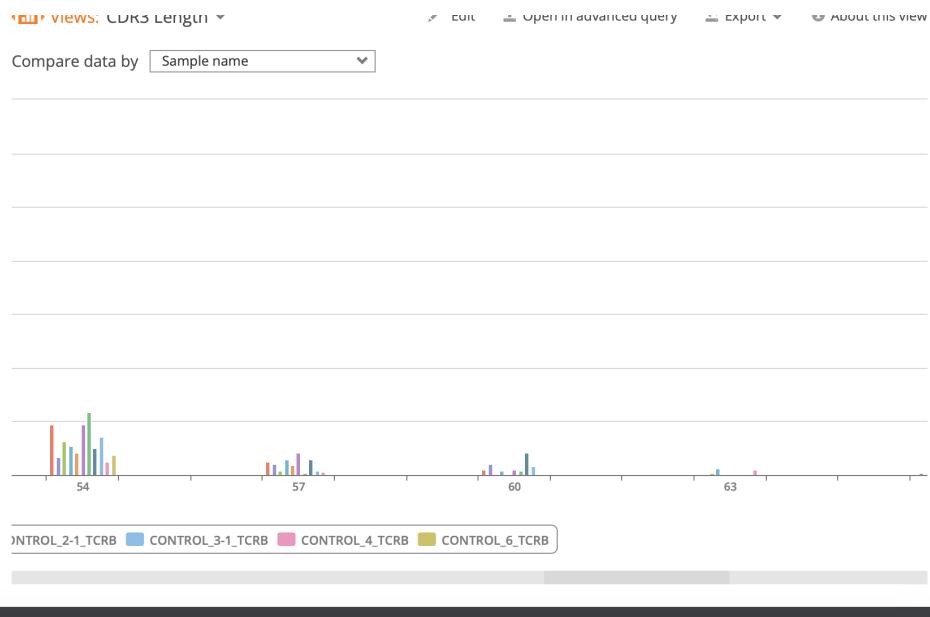
Dataset:

1. Gene Usage

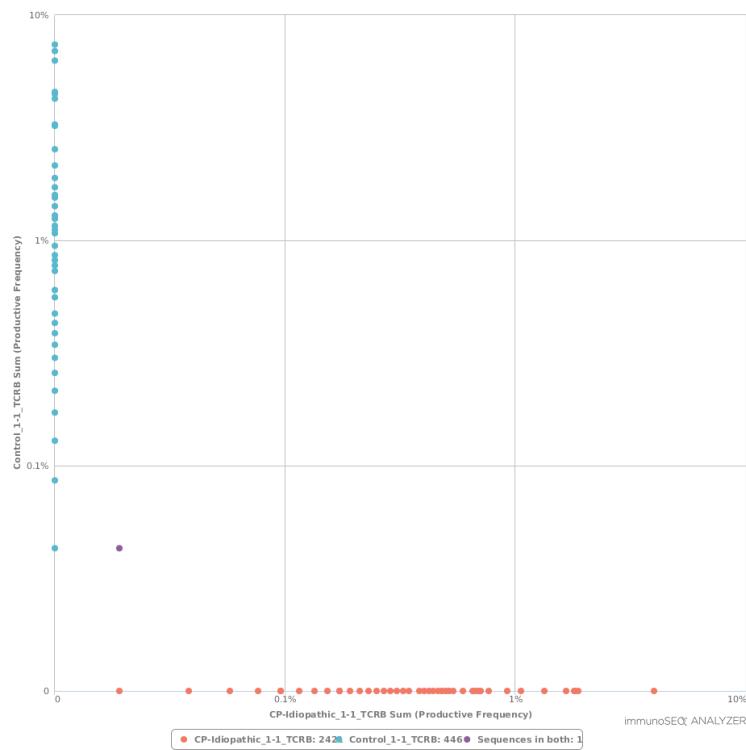


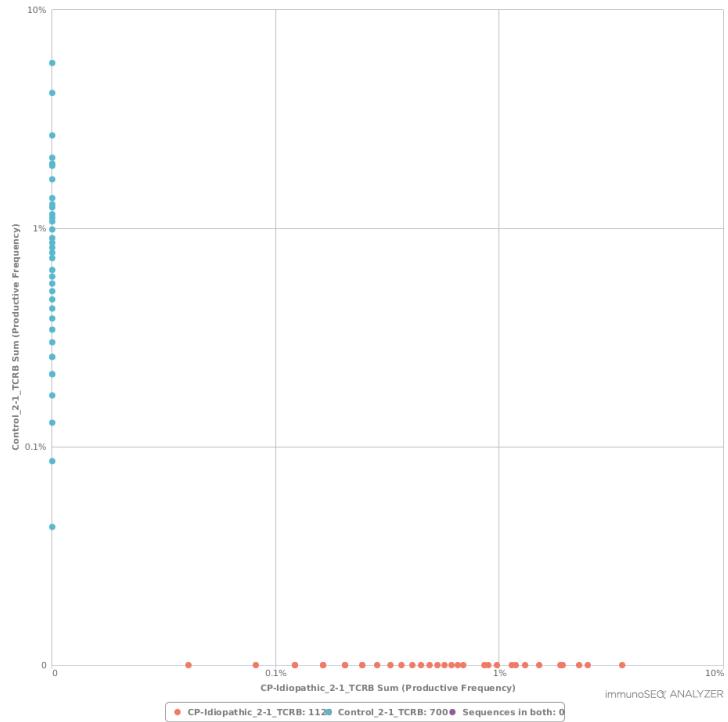
2. CDR3 length





3. Pairwise scatterplot



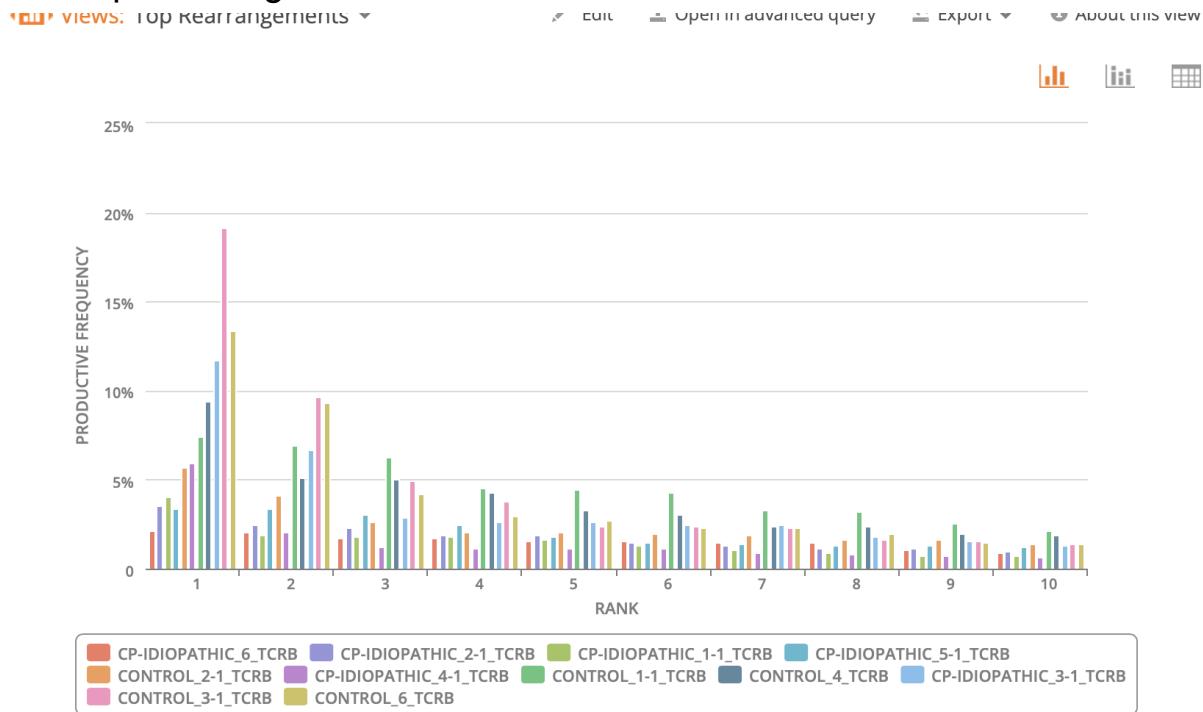


4. Combined rearrangements

Select to:	START NEW ANALYSIS	INCLUDE	EXCLUDE	21,899 entries					
<input type="checkbox"/> Amino Acid	Sum (Productive Frequency)	Present In	CP-Idiopathic_1-1_TCRB	CP-Idiopathic_2-1_TCRB	CP-Idiopathic_3-1_TCRB	CP-Idiopathic_4-1_TCRB			
<input type="checkbox"/> CASSGGGLAAGELFF	19.1649551%	1	na	na	na	na			
<input type="checkbox"/> CASSLGGPYEQYF	13.35616438%	1	na	na	na	na			
<input type="checkbox"/> CASSPGQGYEQYF	11.73402868%	1	na	na	na	11.73402868%			
<input type="checkbox"/> CSVPEGNTIYF	9.61670089%	1	na	na	na	na			
<input type="checkbox"/> CASSSQGANPYEQYF	9.42028986%	1	na	na	na	na			
<input type="checkbox"/> CASSGPSAGRSTDTQYF	9.36073059%	1	na	na	na	na			
<input type="checkbox"/> CAISGALDRDGNTIYF	7.39784946%	1	na	na	na	na			
<input type="checkbox"/> CASSYSKGTGELFF	7.13196034%	1	na	na	na	na			
<input type="checkbox"/> CATSEKRAGNTEAFF	6.92473118%	1	na	na	na	na			
<input type="checkbox"/> CASSPVWGTTEAFF	6.64928292%	1	na	na	na	6.64928292%			
<input type="checkbox"/> CASSLDAGTSVYEQYF	6.27956289%	1	na	na	na	na			
<input type="checkbox"/> CASSLNQGGNTIYF	5.94947025%	1	na	na	na	5.94947025%			
<input type="checkbox"/> CASSLEWFGQPQHF	5.71060541%	1	na	na	na	na			
<input type="checkbox"/> CASSNSYEQYF	5.03432494%	1	na	na	na	na			

Select to:	START NEW ANALYSIS	INCLUDE	EXCLUDE	21,899 entries					
<input type="checkbox"/> Amino Acid	Sum (Productive Frequency)	Present In	CP-Idiopathic_1-1_TCRB	CP-Idiopathic_2-1_TCRB	CP-Idiopathic_3-1_TCRB	CP-Idiopathic_4-1_TCRB			
<input type="checkbox"/> CASIRSTGLSVGQNEKI ...	0.04078303%	1	na	0.04078303%	na				
<input type="checkbox"/> CASKYGYEGLVAFF	0.04078303%	1	na	0.04078303%	na				
<input type="checkbox"/> CASSPTSGGGRTQYF	0.04078303%	1	na	0.04078303%	na				
<input type="checkbox"/> CASSLELKDTQYF	0.04078303%	1	na	0.04078303%	na				
<input type="checkbox"/> CASSLGTTSREQYF	0.04078303%	1	na	0.04078303%	na				
<input type="checkbox"/> CASSPPGQQIRGRTQYF	0.04078303%	1	na	0.04078303%	na				
<input type="checkbox"/> CASSLGGGLAFETQYF	0.04078303%	1	na	0.04078303%	na				
<input type="checkbox"/> CASATGSTDTQYF	0.04078303%	1	na	0.04078303%	na				
<input type="checkbox"/> CATSGRGMAEAFF	0.04078303%	1	na	0.04078303%	na				
<input type="checkbox"/> CSVRGPRGRDEEQYF	0.04078303%	1	na	0.04078303%	na				
<input type="checkbox"/> CASSQDAGEHQPOHF	0.04078303%	1	na	0.04078303%	na				
<input type="checkbox"/> CASSGEQGGEQFF	0.04078303%	1	na	0.04078303%	na				
<input type="checkbox"/> CASSERGSSTDTQYF	0.04078303%	1	na	0.04078303%	na				
<input type="checkbox"/> CASSQYVENTEAFF	0.04078303%	1	na	0.04078303%	na				

5. Top rearrangements



6. Clonality

	Sample Name	Total Templates	Total Productive	Fraction Productive	Rearrangement	Productive Rearrangement	Productive Simpson Clonality	Max Productive Frequency	Locus	Sample Tags
<input checked="" type="checkbox"/>	Control_3-1_TCRB	3,426	2,922	0.8529	878	681	0.232	19.164956%	TCRB	35 Years, 35-39 Years, Control, Male
<input type="checkbox"/>	CP-Hereditary_3-1_TCRB	10,069	8,806	0.8746	2,862	2,202	0.21	17.067908%	TCRB	55-59 Years, 57 Years, Chronic Pancreatitis, Male
<input checked="" type="checkbox"/>	Control_6_TCRB	1,068	876	0.8202	437	329	0.184	13.356164%	TCRB	45-49 Years, 48 Years, Control, Female
<input checked="" type="checkbox"/>	Control_1-1_TCRB	3,130	2,325	0.7428	609	451	0.1627	7.397849%	TCRB	40-44 Years, 43 Years, Control, Male
<input checked="" type="checkbox"/>	CP-Idiopathic_3-1_TCRB	930	767	0.8247	409	307	0.1579	11.734029%	TCRB	45-49 Years, 49 Years, Chronic Pancreatitis, Male
<input checked="" type="checkbox"/>	Control_4_TCRB	3,407	2,622	0.7696	756	597	0.1502	9.42029%	TCRB	30 Years, 30-34 Years, Control, Female, HLA MHC Class I
<input checked="" type="checkbox"/>	Control_2-1_TCRB	2,881	2,329	0.8084	923	725	0.1036	5.710606%	TCRB	60-64 Years, 63 Years, Control, Female
<input type="checkbox"/>	CP-Hereditary_2-1_TCRB	56,605	45,569	0.805	21,949	17,455	0.099	9.506463%	TCRB	20-24 Years, 21 Years, Chronic Pancreatitis, Female
<input checked="" type="checkbox"/>	CP-Idiopathic_5-1_TCRB	3,306	2,653	0.8025	1,343	1,087	0.0856	3.354693%	TCRB	55-59 Years, 58 Years, Chronic Pancreatitis, Female
<input checked="" type="checkbox"/>	CP-Idiopathic_2-1_TCRB	3,329	2,452	0.7366	1,442	1,136	0.073	3.548124%	TCRB	60-64 Years, 62 Years, Chronic Pancreatitis, Female
<input checked="" type="checkbox"/>	CP-Idiopathic_4-1_TCRB	7,764	6,135	0.7902	4,304	3,374	0.0729	5.94947%	TCRB	15-19 Years, 17 Years, Chronic Pancreatitis, Male
<input type="checkbox"/>	CP-Hereditary_5-1_TCRB	6,281	4,972	0.7916	2,166	1,748	0.0726	3.05712%	TCRB	15-19 Years, 18 Years, Chronic Pancreatitis, Female
<input type="checkbox"/>	CP-Hereditary_6_TCRB	5,332	4,149	0.7781	2,616	2,048	0.0706	4.989154%	TCRB	30-34 Years, 32 Years, Chronic Pancreatitis, Female
<input checked="" type="checkbox"/>	CP-Idiopathic_1-1_TCRB	6,713	5,210	0.7761	3,144	2,477	0.0663	4.011516%	TCRB	40-44 Years, 41 Years, Chronic Pancreatitis, HLA MHC Class I, Male
<input type="checkbox"/>	CP-Hereditary_7-1_TCRB	28,533	21,900	0.7675	12,443	9,579	0.0606	4.990868%	TCRB	40-44 Years, 41 Years, Chronic Pancreatitis, Male
<input checked="" type="checkbox"/>	CP-Idiopathic_6_TCRB	29,921	24,252	0.8105	14,179	11,386	0.0573	2.131783%	TCRB	40-44 Years, 42 Years, Chronic Pancreatitis, Female
<input type="checkbox"/>	CP-Hereditary_4-1_TCRB	4,435	3,475	0.7835	2,706	2,074	0.0417	1.294964%	TCRB	10-14 Years, 13 Years, Chronic Pancreatitis, Male
<input type="checkbox"/>	CP-Hereditary_1-1_TCRB	19,831	15,604	0.7868	9,392	7,286	0.0363	1.281723%	TCRB	20-24 Years, 21 Years, Chronic Pancreatitis, Male