Michael Deising

April 22, 2024

Bioinformatics

Assignment 6

1.) How many splice variants does the human BRCA2 gene have? **15**

2.) How many exons are in the longest splice variant(s) of BRCA2? **27**

3.) How many introns are in the longest splice variant(s) of BRCA2? **26**

4.) What chromosome is this gene on? **Chromosome 13**

5.) What are the gene synonyms of BRCA2?

Gene synonyms: BRCC2, FACD, FAD, FAD1, FANCD, FANCD1, XRCC11

6.) How many orthologs of human BRCA2 are found in the African ostrich? **23 orthologues**

7.) In the human hemoglobin transcript ENST00000330597.5 HBG1-201, what is the sequence of the second exon? GCTCCTGGTTGTCTACCCATGGACCCAGAGGTTCTTTGACAGCTTTGGCAACCTGTCCTCTGCCTCTGCCATCATGGGCAACCCCAAAGTCAAGGCACATGGCAAGAAGGTGCTGACTTCCTTGGGAGATGCCACAAAGCACCTGGATGATCTCAAGGGCACCTTTGCCCAGCTGAGTGAACTGCACTGTGACAAGCTGCATGTGGATCCTGAGAACTTCAAG

8.) What is the length of the first intron of this hemoglobin transcript? **The length is 122bp**

9.) How many pseudogenes are in the narwal genome annotation (NGI\_Narwhal\_1, INSDC Assembly GCA\_005190385.2)? **723**

Using Genomicus web-code version 2021-08-15 and database version 109.01, answer the following questions:

10.) What two genes flank BRCA2 in the human genome? **N4BP2L1 and ZARIL**

11.) What genes flank the sonic hedgehog gene(s) in the betta fish? **Rnf32 and rbm33a**

12.) How many copies of shh are in the betta fish? **4**

13.) How many copies of shh are in Atlantic salmon? **4**

Using NCBI answer the following questions:

14.) How many genome assemblies are available for Bos taurus on NCBI? **26 genome assemblies**

15.) How many of the above assemblies are assembled into chromosomes? **11 chromosome**

16.) How many chromosomes does Bos taurus have excluding sex chromosomes? **29**

17.) Which cow chromosome is the longest? **Chromosome 1**

18.) How many microRNAs are annotated on chromosome the assembly: Bos taurus ARS-UCD2.0? **1115 microRNA**

19.) If I am interested in learning more about the gene associate with the transcript: NM\_001104822, what is a good publication to read on PubMed?

Gene evolution and gene expression after whole genome duplication in fish: the PhyloFish database

20.) What is the translated coding sequence of NM\_001104822?

**atggacc tgctggccaa agcatctgtg**

**61 ctgctgttgc tgctcctgag cctcagcaac gctcaaactg acaatatgga ggaagcagaa**

**121 aacggttcat caaaggagga aatagacgag tctgaactgg aggacgtgtc ctccatcatc**

**181 ttcagaatga acaacaactc tatggaggaa ctgttggaag gagatcttgt tcttcccaaa**

**241 accaggaatg ccatgaagtg ctttggcgct ccagatagct gccgctggcc aaagtcttcc**

**301 aatggcatcg tgaaggttcc ttatgtggtt agcgacaact atgaaagtga cgagaaggaa**

**361 accattcgga acgccatgaa ggagtttgca gaaaaaacct gcattcactt tgttcctcgc**

**421 aacaatgaga gggcctacct gagccttgaa cccagatttg gctgcaagtc tatgatgggc**

**481 tatgttggtg acaaacaagt ggtggtgctt cagcggtttg gctgcataaa gcacgccgtc**

**541 atccagcatg agctcctgca tgctctgggt ttctaccacg agcacactcg gagcgaccgc**

**601 gaccagcatg taaaaatcaa ctgggaaaac atcattaaag atttcacaca caactttgat**

**661 aagaatgata ccgacaatct gggcaccccg tatgactatg gctccatcat gcactatgga**

**721 agaactgcct ttggaaaaga cagaaaggag accataaccc ccatccctaa ccccaaagct**

**781 gccattggcc aaacagagag gatgtcggac atcgatatac ttcgagtcaa taagctttac**

841 aaatgttga