**Ashlee Gardiner**

**Bioinformatics**

**Assignment 6 - UPDATED**

1. **Using Ensembl Ensembl Release 111 answer the following questions:**

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? **BRCC2, FACD, FAD, FAD1, FANCD, FANCD1, XRCC11**

6.) How many orthologs of human BRCA2 are found in the African ostrich? **172 orthologues**

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

**G**[**C**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1429943325;vf=735355253)**T**[**C**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1847923528;vf=741376696)**CTG**[**G**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1343156232;vf=731218509)[**T**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1589906467;vf=739070804)**TGTCT**[**A**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1335958398;vf=730873977)**CC**[**C**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs41404150;vf=706976586)**A**[**T**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs35700518;vf=706959283)**GGACCC**[**A**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs35977759;vf=706969491)**GA**[**G**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs33974602;vf=706896498)**G**[**T**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1589906450;vf=739070801)**T**[**C**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1847923183;vf=741376580)**T**[**T**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1394056016;vf=733644608)**T**[**G**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs41475844;vf=706976788)**A**[**C**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1395960485;vf=733732899)**AGCTT**[**T**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1589906444;vf=739070800)**G**[**G**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1847923033;vf=741376523)**CAACCTGT**[**CCTC**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1409662507;vf=734391537)

[**T**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1409662507;v=rs1589906439;vf=734391537;vf=739070798)[**GCCTCTG**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1409662507;vf=734391537)[**C**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1409662507;v=rs35746147;vf=734391537;vf=706960945)[**C**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1409662507;vf=734391537)[**A**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1173059624;vf=723098646)[**T**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs778172072;vf=713544984)**CA**[**T**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1434319312;vf=735562609)**G**[**G**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1422261857;vf=734988779)**G**[**C**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1187559666;vf=723792915)**AACCC**[**C**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1485842420;vf=738016136)[**A**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs539011424;vf=709947693)**AAGT**[**C**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1847922732;vf=741376424)[**A**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs34747494;vf=706924389)**A**[**G**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1847922666;vf=741376399)[**G**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1847922641;vf=741376390)**CAC**[**A**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1213522116;vf=725034610)**T**[**G**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs747508521;v=rs756489020;vf=712059015;vf=712495096)[**G**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs747508521;vf=712059015)**CAAG**[**A**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1847922517;vf=741376345)[**A**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1847922517;v=rs1589906410;vf=741376345;vf=739070788)[**G**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1208095608;vf=724775097)**GT**[**G**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs752995575;vf=712325850)**CTGACTTC**

[**C**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1589906400;vf=739070785)**T**[**T**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1225872760;vf=725625116)**G**[**G**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs34049890;vf=706898654)**GA**[**G**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs33965337;vf=706896416)**A**[**T**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1847922133;vf=741376219)**G**[**C**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1256614625;vf=727088779)**CA**[**C**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1061234;vf=706477848)**AAAG**[**C**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs755174349;vf=712431430)[**A**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1589906380;vf=739070780)**CCTG**[**G**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs34435255;vf=706912864)**A**[**T**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1847921855;vf=741376110)[**G**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs33947112;vf=706896259)**ATCTCAAGGG**[**C**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1359816903;vf=732014326)**AC**[**C**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs751636715;vf=712259925)**TTTGCC**[**C**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1847921684;vf=741376065)[**A**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1415219099;vf=734657005)**GCTG**[**A**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1376275271;vf=732796568)[**G**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs766448778;vf=712977746)**TGA**

**AC**[**T**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1464732955;vf=737010311)**GCACT**[**G**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1847921541;vf=741376013)**T**[**G**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs527266118;vf=709349841)**ACA**[**A**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1847921469;vf=741375992)**GCT**[**G**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs750370911;vf=712198199)**C**[**A**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs34127117;vf=706901542)[**T**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1418054559;vf=734788224)**GTGGAT**[**C**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1190498683;vf=723934545)**CT**[**G**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1267554412;vf=727611202)**AGAACTT**[**C**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs765208645;vf=712917744)**A**[**A**](https://useast.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000213934;r=11:5248269-5249857;t=ENST00000330597;v=rs1257266010;vf=727120011)**G**

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

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

1. **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 ZAR1L**

11.) What genes flank the sonic hedgehog gene(s) in the betta fish? **rnf32 and rbm33a**

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

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

1. **Using NCBI answer the following questions:**

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

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

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? **1,115**

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?

**A good publication to read is Gene evolution and gene expression after whole genome duplication in fish: The PhyloFish database by Pasquier et al.**

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