**Title: Structure and function of the human genome**

**Exercise 1: Collate technical information about the ACE2 gene in human genome**

1. Go to <https://www.ncbi.nlm.nih.gov/>
2. Enter ***“Homo sapiens” AND “ACE2”*** as the search term to search all databases in the NCBI
3. What is the basic description provided for the ACE2 gene in humans?
4. What are the alternate names of the ACE2 gene?
5. What are the transcript and protein sequence IDs for the ACE2 gene?
6. What are respective lengths of transcript and protein sequence for the ACE2 gene?
7. Click on the Genome Data Viewer to visualize the gene body.
8. What are the genes immediately upstream and downstream of the ACE2 gene?
9. What is the location of the ACE2 gene in the human genome assembly version GRCh38.p13?
   1. Chromosome name
   2. Start coordinate
   3. End coordinate
   4. Strand
10. Perform BLAST search using the mRNA sequence against the mouse genome and transcriptome database.
11. What are the alignment properties of the best alignment against the mouse genome?
    1. Accession ID of the top hit
    2. E-value
    3. Percent Identity
    4. Percent of the query sequence covered by the alignment
    5. Start and end position of the first aligned block for the query and the subject sequence.

Think of other useful information that you could mine from the NCBI database by exploring the results page and clicking through to various databases and tools.

1. GenBank records
2. OMIM records
3. Conserved domains
4. GEO Profiles

**Exercise 2: Ensembl database to understand the evolution of the human X chromosome and paralogous genes of chromosome 22**

1. Go to [http://www.ensembl.org](http://www.ensembl.org/)
2. Check out the number of species for which information is available in Ensembl.
3. Click on the human as the species from the dropdown list for “All genomes”.
4. Spend a minute or two to understand the types of information available for the human genome from Ensembl database.
5. Click on the “View Karyotype”
6. Select chromosome X and click through to the chromosome summary.
7. Spend a minute or so to understand various graphs displayed for the chromosome.
8. Click on Synteny from the menu on the left.
9. What chromosome the human X is homologous to in the following genomes? You can change species name from dropdown menu on the right of the image.
   1. Mouse:
   2. Dog:
   3. Opossum:
   4. Platypus:
   5. Chicken:

*What does this homology tell you about the human X chromosome evolution?*

1. Explore information about the ACE2 gene by searching for ACE2 in the search bar.
2. Select the entry for human ACE2 gene.
3. How many alternate transcripts does ACE2 gene have in the human genome?
4. What are different types of transcripts produced by the ACE2 gene?
5. Select the Gene gain/loss tree from the menu on the left.
   1. When did the ACE2 gene evolve?
   2. What group of animals is it common to?
   3. How many copies of ACE2 gene is found in most species?
   4. Are there any species with more than one copy of ACE2 gene?
6. Click on Biomart at the top of the webpage.
7. Select “Ensembl Genes 104” and “Human Genes”
8. Click on the Filters tab:
   1. Select Chromosome 22 in Regions
   2. Select protein coding in Gene type filter
9. Click on Attributes in the left panel, select Homologues radio button
10. Select Gene stable ID and Chromosome name from the Gene attribute list
11. Select Paralogous ID, Paralog chromosome and Paralog last common ancestor from the Paralog attribute. Export results as XLS.
12. Explore properties of paralogous genes for chromosome 22:
    1. How many paralogous genes per chromosome? Is it meaningful in anyway based on the size or number of genes on each chromosome?
    2. Properties of duplication based on when did the paralogous gene evolve?
       1. Bilateria, amniota or homo sapiens

NOTES: Ensembl serves following datasets as well:

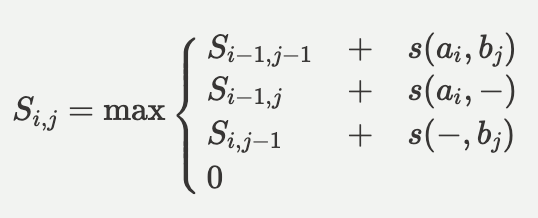
1. Metazoa: <http://metazoa.ensembl.org/>
2. Plants: <http://plants.ensembl.org/>
3. Fungi: <http://fungi.ensembl.org/>
4. Protists: <http://protists.ensembl.org/>
5. Bacteria: <http://bacteria.ensembl.org/>

**Exercise 3: UCSC genome browser to explore conservation property and regulation of the BRCA2 gene**

1. Go to <https://genome.ucsc.edu/>
2. Search for ACE2 gene in the human genome.
3. Explore at the bottom of the image various types of tracks available for display.
4. Select Conservation track and “full” display mode. Drag the Conservation track to the top in the image by clicking and dragging the grey bar on the left of the image.
5. Compare the conservation of ACE2 gene body across vertebrates.
   1. Is there a trend in conservation for exon vs introns?
6. Select Regulation track for:
   1. *H3K4Me1* histone modification: activation of transcription
   2. *H3K27Ac* histone modification: distinguishes active enhancers from poised enhancers.
   3. *H3K4Me3* histone modification: enriched in transcriptionally active promoters
7. Select GTEx track:
   1. Click on the bar plot to see the expression patterns of the ACE2 gene in various tissues in humans.
      1. What tissue has the most abundant expression of ACE2?
8. Compare the expression bar plots of ACE2 with other genes from the pathway in GTEx data.

**Sequence Similarity:**

Based on the equation below, we will fill the table with appropriate values in each cell.

****

Match = 1

Mismatch = -1

Gap = -2

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | A | C | G | T | C | A |
|  | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| T | 0 |  |  |  |  |  |  |
| C | 0 |  |  |  |  |  |  |
| G | 0 |  |  |  |  |  |  |
| T | 0 |  |  |  |  |  |  |
| A | 0 |  |  |  |  |  |  |
| A | 0 |  |  |  |  |  |  |