Bioinformatics - lab 1 and 2

Time to submit final report: 6 weeks

File name: `name_surname_1_2_bio.pdf`

Exercise type: two weeks

Goal: The purpose of the exercise is to familiarize the student with the basic concepts operating in biology and bioinformatics. This aim will be achieved, among other things, by analyzing the circulation of information in biology, and multidirectional presentation of the basics of genetics, for example, with the example of human genetics.

Basics of biology for bioinformaticians - exercises

1. Read the "Rules" file first.
2. Before you start following the instructions below, please take 5 minutes to write down what motivated you to choose the course " Basics of Bioinformatics". What do you want to learn in this course? This will help us improve it. Put your thoughts (it can even be a few words) in the file "Bioinformatics.txt" and send it to the teacher via MS Teams. Thank you.

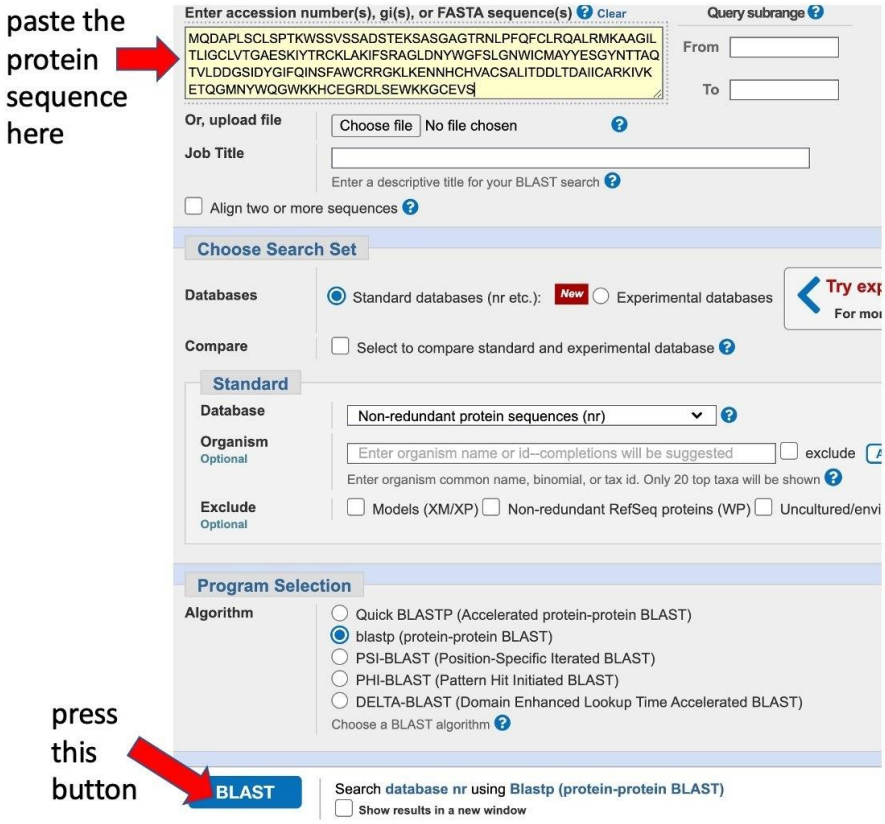
DNA is a molecule that is the main data store in a living cell. It is made up of repetitive elements called nucleotides, and there are known 4 types of them labelled as: A, T, C or G. Data from this data store is used by the organism by reading it in a process called "gene expression." It involves transferring information from DNA to a slightly different molecule called mRNA, and then to a protein. The protein is usually already a fully functional "biorobot" that performs specific tasks. Each of these steps has a name. Transcribing information from DNA to mRNA is transcription, and from mRNA to protein is translation.

3. You will carry out a mini-simulation of the information flow in biology, i.e. the path of genetic information from DNA to protein. To do this, open the tool: BIOMODEL-link
    - Put the following DNA sequence in the "DNA sequence" field:

      ```
      ATGCAGGACGCTCCCCTGAGCTGCCTGTCACCGACTAAGTGGAGCAGTGTTTCTTCCGCAGACTCAACTG
      AGAAGTCAGCCTCTGGGGCAGGCACCAGGAATCTGCCTTTTCAGTTCTGTCTCCGGCAGGCTTTGAGGAT
      GAAGGCTGCGGGCATTCTGACCCTCATTGGCTGCCTGGTCACAGGCGCCGAGTCCAAAATCTACACTCGT
      TGCAAACTGGCAAAAATATTCTCGAGGGCTGGCCTGGACAATTACTGGGGCTTCAGCCTTGGAAACTGGA
      TCTGCATGGCATATTATGAGAGCGGCTACAACACCACAGCCCAGACGGTCCTGGATGACGGCAGCATCGA
      CTATGGCATCTTCCAGATCAACAGCTTCGCGTGGTGCAGACGCGGAAAGCTGAAGGAGAACAACCACTGC
      CATGTCGCCTGCTCAGCCTTGATCACTGATGACCTCACAGATGCAATTATCTGTGCCAGGAAAATTGTTA
      AAGAGACACAAGGAATGAACTATTGGCAAGGCTGGAAGAAACATTGTGAGGGCAGAGACCTGTCCGAGTG
      GAAAAAAGGCTGTGAGGTTTCCTAA
      ```

    - Then click the arrow with the word "transcription" and paste the generated mRNA sequence [**task 1**]. What is the mRNA [**task 2**]? Why is there a "U" instead of a "T" in mRNA. [**task 3**].
    - Then click the arrow with the name "translation" and paste the generated fragment of the protein sequence [**task 4**]. What are proteins generally made of [**task 5**]?

- Check with this tool: BLAST-link what is the name of the protein encoded by this gene fragment [**task 6**], what species it is from [**task 7**] (see picture 1 and 2).



Rysunek 1: BLAST

4. Explain the concept of complementarity in the context of the structure of the DNA molecule. You can search for an explanation here: video-link [**task 8**].

   - With the tool: *BIOMODEL-link* create a sequence complementary to:

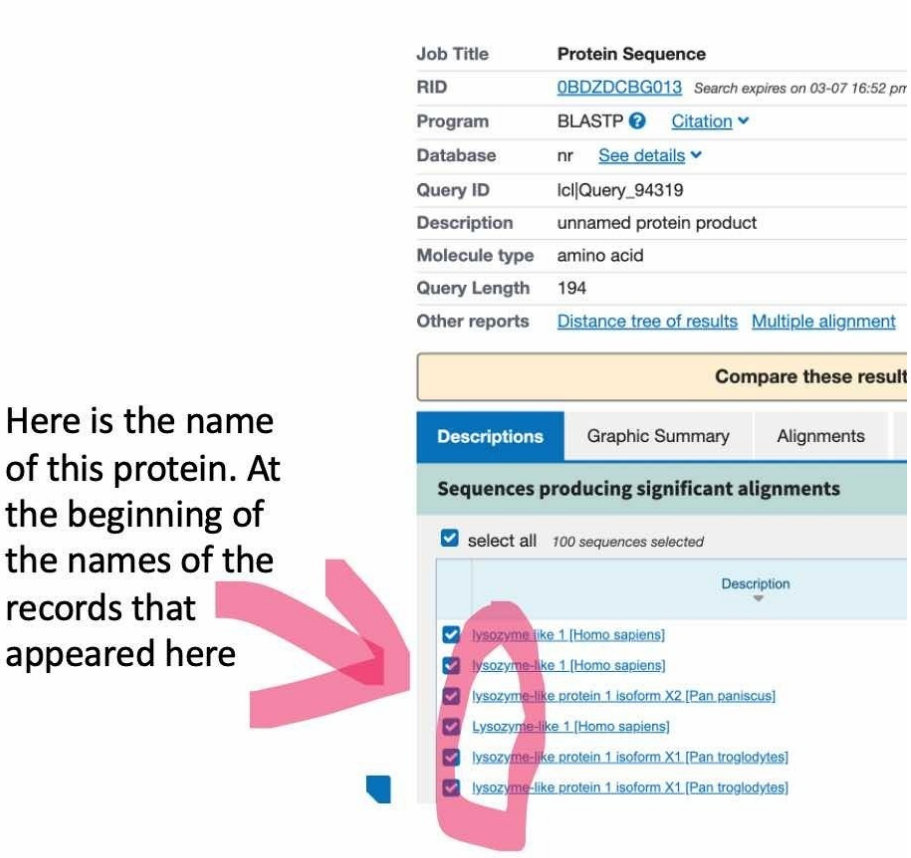
     `ccaagcacatgtggcctggagacataaagggcaattttggacaaactgcata`
     `tctaaacagtaactggttccctgcaggaagtaaacctttcatttaccaggag`

     and then paste both strands as in the example below (remember the font in which the characters take up the same amount of space, e.g. Courier New) [**task 9**]:

     `AACCTTGCGAACCTTGCGAACCTTGCGAACCTTGCG`
     `TTGGAACGCTTGGAACGCTTGGAACGCTTGGAACGC`

5. Let's look at the phenomenon that is called the cell cycle. In general, these are the phases of a cell's life, which take place from one division to the next. Depending on the current phase, such a cell "behaves" slightly differently. Let's try to deepen the topic.

   - Go to the article page: article-link, look at the link and answer the questions: what are cell cycle checkpoints [**task 10**]? What is a checkpoint in the G2/M phase [**task 11**].



Rysunek 2: BLAST2

6. Plant and animal cells are significantly different. Go to cellsalive.com , look at Interactive Cell Models -> Plant/Animal->Start the animation and list (use your own words) 5 differences you notice between a plant cell and an animal cell. Don't do advanced analysis, simply describe the differences between the images [**task 12**].

7. Let's look at the sizes of selected structures in biology. Go to cellsalive.com , watch "How Big is..." and answer the question what is bigger than a rhinovirus and smaller than a *Staphylococcus* bacteria [**task 13**].

8. Time to play. Solve the Memory game found here: cellsalive.com
   Paste a screen of the completed game (with all the cards covered) [**task 14**].

Chromosomes are special structures in our cells that contain genetic material in the form of DNA. Simply put, all the DNA in a cell does not exist as one huge molecule, but is divided into smaller ones, and these are the chromosomes. Each person has 46 chromosomes, which are divided into 23 pairs. Each of us inherits one chromosome from each pair from our parents (one from the mother, one from the father).

mtDNA is a DNA molecule that is found in cellular structures called mitochondria. Mitochondria are cellular energy factories. For this reason, they are numerous in muscle cells, for example. Mitochondria have some DNA of their own, which we call mtDNA. Where did it come from? There is a hypothesis that once upon a time there was an interesting phenomenon in which one cell began to use the other in such a way that it built it into its structures and began to use it as an energy provider. The mitochondrion may be a "descendant" of a bacterium that once existed, which the so-called eukaryotic cells (that's what we call human, animal, plant and other cells more advanced than bacteria) absorbed and began to use for their purposes. What do you make of the fact that there are more mitochondria in your body than your own cells? Who are we?

9. Read the content from the website: igs.org.pl or support.ancestry.com (EN) and in 2 - 4 sentences summarize the study of genetic profiles of Y chromosome and mtDNA - what is it, what is the difference, why is it made? [**task 15**].

10. Read the website: www.eupedia.com. Answer the questions: What is a haplogroup [**task 16**]? What country has the most representatives of haplogroup R1a [**task 17**]? What famous people possessed or have this haplogroup [**task 18**]?

Genetic marker is a fragment of DNA (it can even be one nucleotide) that is associated with a trait, e.g. if there is an "A" at hypothetical position 122 on chromosome 10, it increases the probability of having blonde hair, and if there is a "C" it is black hair. And then this place in the chromosome is a genetic marker, in which there can be two alleles (two versions), A or C.

Predisposition is otherwise known as a tendency to a particular trait. That is, if you have the AA genotype (or two A alleles) in a genetic marker, e.g. rs100001, it means that you have, for example, a greater tendency to be empathetic than people who have CC. This is not determined, it is a tendency. If you don't fasten your seat belt, your risk of dying in an accident increases, but it's not saying that you will 100% die in an accident because you didn't fasten your seat belt. That's how most genetic markers work.

11. Study the content of snpedia.com and write, in a nutshell, what predispositions in humans are associated with the genetic marker labeled rs53576 [**task 19**].

12. Similarly to the task above, briefly describe the predisposition associated with the genetic marker "rs333". Previously find it on snpedia.com [**task 20**].

More and more options for testing your own DNA are appearing on the commercial service market. Today, the cost of knowing the sequence of all own DNA can not exceed $200. From DNA today, you can already read many interesting things thanks to scientific research. It is often the case that once you have your DNA sequence, you can use applications that will allow you to analyze the data from a variety of aspects.

13. Go to sequencing.com and list the 3 reports that most captured your attention [**task 21**].
14. Watch ted.com. List the two issues you found most interesting [**task 22**].
15. What is a dominant trait and what is a recessive trait in biology (you can find the answers here: zpe.gov.pl (EN) or here yourgenome.org) (EN) [**task 23**].

Selected dominant and recessive human traits dependent on gene complexes.

Dominant trait	**Recessive trait**
non-red hair	red hair
dark hair	light hair
protruding ears	ears not protruding
the ability to curl the tongue into a trumpet	no ability to curl the tongue into a trumpet
right hand leading	the left hand leading
the presence of dimples in the cheeks	no dimples in the cheeks
the presence of freckles	no freckles
long eyelashes	short eyelashes
free earlobes	attached earlobes
thumb straight	thumb bent
oval shape eye	round shape eye
oval face shape	quadrangular (square) face shape
placing the left thumb over the right while folding the hands	placing the right thumb over the left while folding the hands

Source: https://zpe.gov.pl/a/dziedziczenie-wybranych-cech-u-czlowieka/D10jpb5vw

16. Using the table above, determine what traits you have that are dominant and which are recessive [**task 24**].

For those who are willing

1. Briefly answer the questions: What is an allele and how does it relate to the gene [**task 25**]? How many inhabitants of Poland are carriers of the phenylketonuria allele [**task 26**]? If both parents are carriers of phenylketonuria, what is the probability that the child will have the disease [**task 27**]?
2. Can humans produce vitamin C [**task 28**]? Why did this happen? What is the mechanism of [**task 29**]?