Marine Bioinformatics Lab 1: BLAST

100 Points

Do as much as you can in class, ask lots of questions, finish at home if you need to.

Answer the red questions and upload edited doc to Canvas.

Due Tuesday Feb 7 at 11:55 PM.

1. The 18S ribosomal RNA (rRNA) gene is a common marker gene used to construct

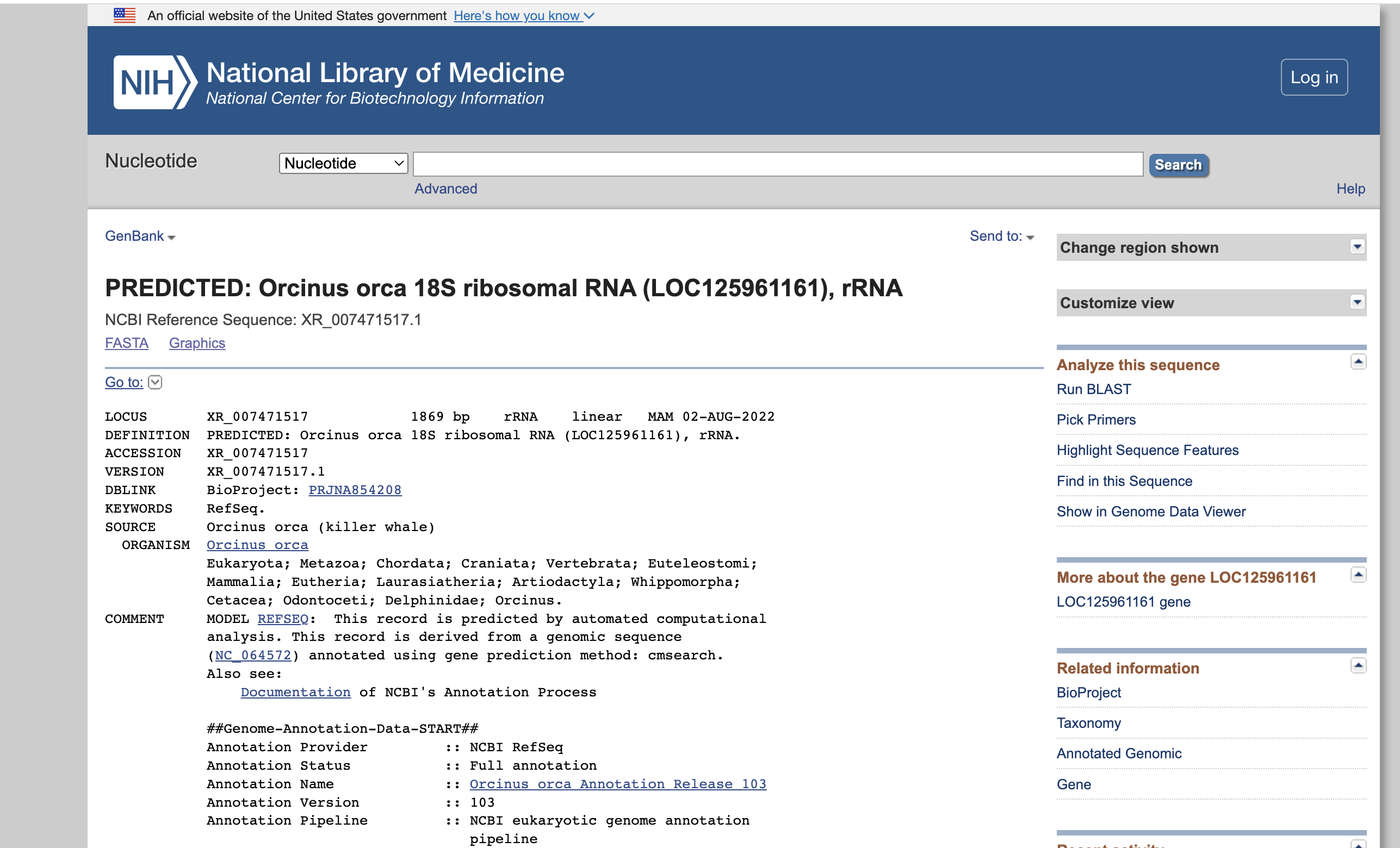
evolutionary relationships, especially among vertebrates (animals with backbones). It’s computationally overwhelming to compare the entire genomes of two species, but it’s easy to pick and analyze just one gene that many organisms have. Download the nucleotide sequence for the 18S rRNA gene of *Orcinus orca* (killer whale, I know that sounds harsh but it’s deserved). To do this, browse to <https://www.ncbi.nlm.nih.gov/> and type “Orcinus orca 18S rRNA” into the search field. To the left of the search field is a pulldown menu that says “All Databases”. Change it to “Nucleotide”. You’ll *probably* see this. (“Probably” because the database can change at any moment … this screenshot was taken on Jan 20, 2023.) Click on the first entry with length roughly = 1869 bp (“bp” = base pairs).



What is the accession number for the record you selected?

Accession Number: XR\_007471517

1. You are now looking at the “GP Page” for your record. It provides the sequence and lots of other useful information. The top looks like this:



Where on the page do you see the sequence? Other than the sequence, what are 3 other pieces of information that you see on this page?

The sequence is at the bottom of the page under “ORIGIN”

-6: missing other pieces of information

1. Now browse to <https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&BLAST_SPEC=GeoBlast&PAGE_TYPE=BlastSearch> (you can just click on the link back there). This is where you can request a nucleotide blast, executed on NCBI’s computers, against the enormous GenBank database. Under “Enter Query Sequence”, type the accession number. Don’t change any of the default configurations, just click the “BLAST” button. When the blast completes, you’ll get a results screen with 4 tabs: “Descriptions”, “Graphic Summary”, “Alignments”, and “Taxonomy”. You might have to scroll down a bit to see these tabs. Make sure the “Descriptions” tab is selected and the “Show” pulldown specifies 50 or 100. Paste a screenshot of the top 10 hits.

Graphical user interface, application

Description automatically generated

1. What is the range of E-values that you see?

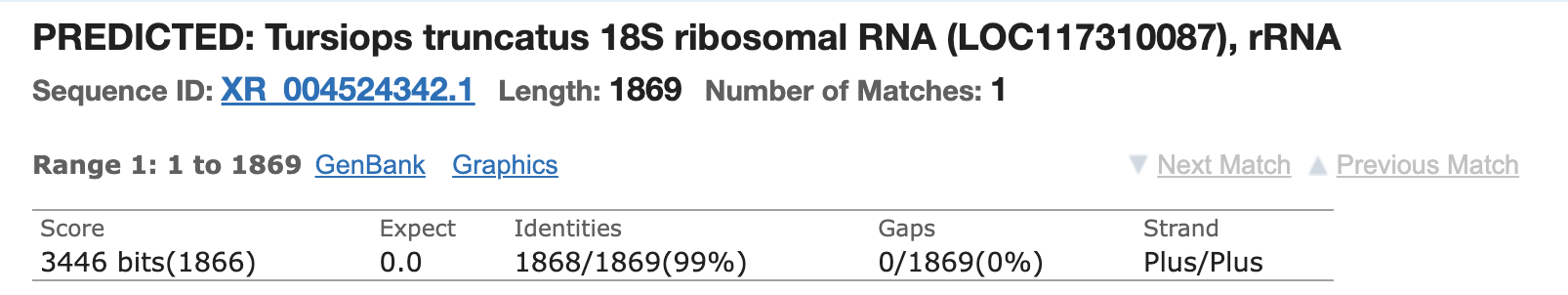
The range of E-values that I see is only 0.0 to 0.0 .

1. What are the scientific and common names of the first hit target (“subject”) with the highest % identity (“Per. Ident”) that is not *Orcinus orca*? (A scientific name is almost always 2 words: the genus and then the species. Both should be *in italics*. The genus should start with a capital letter, the species should not.) If the name is too long to fit in its column and ends with “… “ then rest the mouse on the name. The full name will appear after 1 second. In a new browser window, google the scientific name of the subject. What is the common name? Paste a picture of the organism.

Tursiops truncates



1. In the “Descriptions” column, click in the row for the closest-relative hit that you just identified. You’ll see an alignment. The top of the alignment section looks like this:



What is the length of the subject sequence? How many match (identity) columns are in the alignment? What % identity does the alignment report? BLAST reports % identity as an integer, rounding the fractional part if necessary. Any fraction >.99 and <1 is rounded to 99%, so the only way to get 100% identity is to be completely identical.

The length of the subject sequence is 1869. There is 1 match. There are 99% identity for the alignment report.

-5: Match identity columns refers to the amount of identities. Example case: 1868

1. When you navigated to the alignment, your display switched to the “Alignments” tab. Scroll to the top of the page and select the “Descriptions” tab. What organisms, other than *Orcinus orca* but including the subject you reported in Question 4, are reported? Enter them into the table on the next page, remembering to format the scientific name correctly. (If you see a 3-word scientific name, the 3rd word is the subspecies. Enter it in lower case italics, e.g. *Elephas maximus indicus*.) Don’t enter any subject more than once. Don’t enter more than 5 organisms. Also paste a picture of each organism in the 3rd column.

.

|  |  |  |
| --- | --- | --- |
| Scientific  Name | Common name | Picture |
| Tursiops truncates | Common bottlenose dolphin |  |
| Cavia porcellus | Guinea pig |  |
| globicephala melas  -2: Globicephala melas | Long-finned pilot whale |  |
| Lagenorhynchus obliquidens | Pacific white-sided dolphin | Pacific White-Sided Dolphin:Lagenorhynchus obliquidens - ScienceDirect |
| *Ovis aries* | Sheep |  |

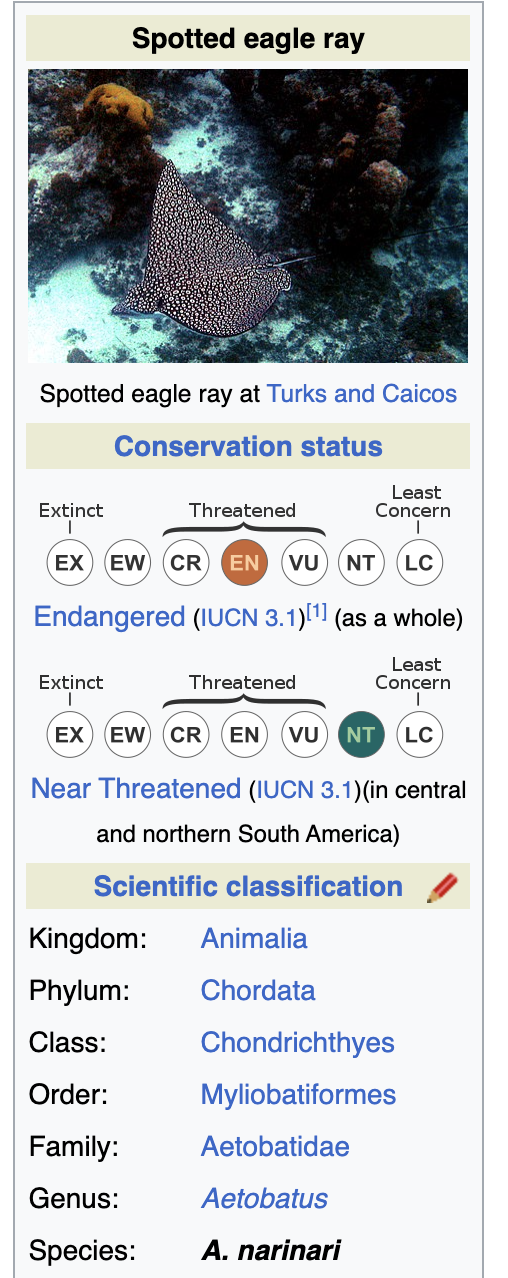
1. Did you find Step 7 to be tedious? (It’s ok if you did, it’s *supposed* to be!)

Not that bad.

1. Learning about the world scientifically often starts with tedious data collecting. The really interesting part happens when you think about the data that you tediously collected. Do that now. Look at your pix from Step 8. They are all closely related to killer whales. Which ones surprise you? Which ones are not surprising at all?

The sheep and guinea pig surprises me. The pacific white-sided dolphin, Long-finned pilot whale, and common bottlenose dolphin do not surprise me at all.

1. Look up the Wikipedia page for “orca”. Usually when you look up any organism, Wikipedia gives you an excellent box at the right edge of the page, telling you the complete taxonomy of the organism. It looks something like this (from the page for the lovely spotted eagle ray):



The taxonomy appears under “Scientific classification”.

1. It’s often informative to learn about the higher taxonomic levels of an organism that you’re studying (its genus, family, order, etc.) What is the taxonomy of killer whales?

|  |  |
| --- | --- |
| Taxonomic level | Name |
| Kingdom | Animalia |
| Phylum | Chordata |
| Class | Mammalia |
| Order | Artiodactyla |
| Family | Delphinidae |
| Genus | *Orcinus* |
| Species | *orca* |

1. Are killer whales really whales? To find out, read about their genus. If that isn't helpful, read about their family, and so on until you find the answer. What taxonomic level gave you the answer? **Killer whales are not really whales.** **Delphinidae was the taxonomic level that gave me the answer**