Marine Bioinformatics

Homework 1

Due by Canvas upload 11:55 PM Feb 17

Edit this doc and submit the edited version

100 points

1. **(10 points)** What is the score of this alignment? Show all your work (zero points if you don’t). Use the standard scoring scheme: +1 for a match, -1 for a mismatch, -2 for an indel.

ATCGGC---GCA

TC---GTTCGCC

A|T = -1

T|C = -1

C|- = -2

G|- = -2

G|- = -2

C|G = -1

-|T = -2

-|T = -2

-|C = -2

G|G = +1

C|C = +1

A|C = -1

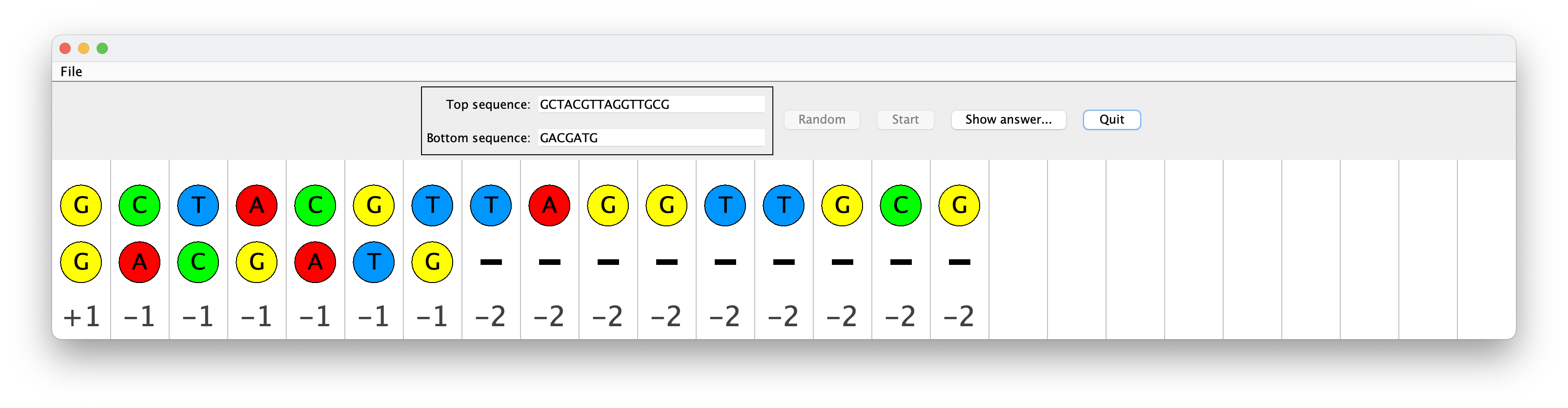
-8-1-6+2-1=-15+2-1=-13-1=**-14**

1. **(10 points)** An indel lowers an alignment score by twice as much as a mismatch. Explain why in terms of evolution.

**A mismatch caused by a polymorphism only affects that particular codon(3 nucleotide sequence). An insertion or deletion occurring between an ancestor and a descendant can completely change the function of proteins and or lead to premature stoppage of a protein’s transcription due to the shifting of amino acids changing all future codons. This means that indels can create a much higher chance of causing a negative outcome (negatively selected for trait such as a non-functioning protein).**

1. **(20 points)** Start the “ManualPWA.jar” app that you downloaded with this assignment. Or, if it’s easier, use the copy that you downloaded with the previous assignment. (They are identical). To run the app, open a Terminal window (Mac) or a Command Prompt (windows). cd into the folder that contains the files for this assignment. (It was created when you unzipped the .zip file that you downloaded.) Then type or paste “java -jar ManualPWA.jar”. ***If you still can't get it to run, see Anson or Dr. Heller asap! Or if you’re running out of time, it’s ok to use a classmate’s computer. But, seriously, get ahead of this!***

Click on the “Random” button, and then on the “Start” button. You’ll see something like this:



Use your mouse to drag the dots in either row into different columns. The app won't let you change the order of the dots. Try to improve on the initial score, which will probably be pretty bad. To challenge yourself, try to make an optimal alignment (one with the highest possible score … there might be many ways to do this). When you’re done trying, click “Show answer…” to see the initial alignment, your solution, and an optimal solution. Paste a screenshot of that display into your homework doc. ***You’ll get full marks for submitting any reasonable answer.***

***Diagram

Description automatically generated***

1. **(20 points)** Run the other app that you downloaded with this assignment: PrematureStop.jar. Start it like you started ManualPWA.jar. Click “Start” and run to completion. Paste a screenshot of the final screen into your homework doc. Explain the horizontal and vertical axes of the diagram. If you can't, this is your opportunity to come to office hours. One purpose of office hours is to get help when you’re stuck. Please don’t be shy!

Graphical user interface, text

Description automatically generated with medium confidence

**The horizontal axes of the diagram is the number of amino acids.**

**The vertical axes of the diagram is the number of experiments.**

**The graph represents the number of experiments that stop with respect to a certain number of amino acids.**

1. **(10 points)** What is the score of this global alignment? Use the standard nucleotide scoring scheme: +1 for a match, -1 for a mismatch, -2 for an indel. (Show your work, zero points if you don’t.) Why would local alignment be a much better option for this situation?



**There are 1000 indels at the front and 1000 indels at the back which when matched with any nucleotide results in a score of -2 for each. The nucleotides in between the indels 1000 beginning and ending indels are of length 10. The nucleotides in the middle all match each other so they get a score of +10. The 1000 indels at the beginning get a score of 1000\*-2=-2000. The same goes for the indels at the end =-2000. This gives a combined score of -2000+-2000+10 = -3990. Local alignment is built for aligning sequences whose lengths vary greatly. Finding where in the longer sequence the shorter sequence best aligns is much more useful. A score for global alignment for sequences whose lengths vary greatly requires many indels be added and therefore skews the alignment score highly negatively.**

1. **(30 points)** In Lab 1, you used BLAST at the NCBI website to determine closest relatives to *Orca orsinus*. Do the same (following the Lab1 instructions) for 2 additional marine animals. The first should be the animal in the table below next to your name. The second can be anything you choose. Start by looking up the *Genus* *species* scientific names of your animals. If you were assigned a broad category (like “cephalopods”), choose any species; if that species isn't in GenBank, choose another that is. For each of your animals, submit a table like the one in Step 7 of Lab 1, showing scientific name, common name, and picture.

Whale shark ACC #: XR\_007261513

|  |  |  |
| --- | --- | --- |
| Scientific Name | Common Name | Picture |
| Rhincodon typus | Whale shark |  |
| Megachasma pelagios | Megamouth Shark | The Massive Filter Feeding Shark You Ought to Know | Smithsonian Ocean |
| Alopias vulpinus | Common thresher |  |
| Carcharodon carcharias | Great white shark |  |
| Squalus acanthias | Spurdogs | Squalus acanthias | CMS |

Moon Jelly ACC #: KC767902

|  |  |  |
| --- | --- | --- |
| Scientific  Name | Common name | Picture |
| Aurelia aurita | Moon Jelly |  |
| Hyperia galba | big-eye amphipod |  |
| Aurelia solida | Moon Jelly(different species than aurita). |  |
| Aurelia pseudosolida | Moon Jelly(different species than aurita). | Water | Free Full-Text | Description of Aurelia pseudosolida sp. nov.  (Scyphozoa, Ulmaridae) from the Adriatic Sea |
| Aurelia labiata | Moon Jelly(different species than aurita). |  |

|  |  |
| --- | --- |
| **NAME** | **ANIMAL** |
| Alcain, Abigail | Seahorse |
| Butelet, Julia | Dolphin |
| Calderon, Kenia | Jellyfish |
| Casio, Julia | Sea turtle |
| Girivasan, Anjali | Humpback whale |
| Hallisey, Ryan | Purple sea urchin |
| Huang, Kevin | Cuttlefish |
| Irawan, Ashley | Sea star |
| Klofas, Andrew | Great white shark |
| Lazaro, Christina | Jellyfish |
| Lee, Shin Hyung | Mantis shrimp |
| Luu, Sophie | Cephalopod |
| Martinez, Shalsy | Dolphin |
| Oliver Brown, David | Sea urchin |
| Reimann, Jessica | Sea otter |
| Rothstein, Anna | Staghorn coral |
| Salter, Daniel | Sea otter |
| Schroeder, Victoria | Humpback whale |
| Sena, Jasmine | Sea lion |
| Tajima, Keigo | Spotted eagle ray |
| Vrooman, Ethan | Whale shark |
| Warshawsky, David | Penguin |
| Anybody else | Manatee |