III (1) Test case output:  
s1 to s2

score: 20

s1: 1 D-EADLY

D EA LY

s2: 2 DGEARLY

[0 0 0 0 0 0 0 0 0 0]

[0 6 6 2 2 0 0 0 0 0]

[0 2 8 4 7 3 0 0 0 1]

[0 0 4 8 4 11 7 3 0 0]

[0 6 6 4 10 7 9 5 1 0]

[0 2 2 2 6 9 5 13 9 5]

[0 0 0 0 2 5 7 9 20 16]

p-value: 2.000000e-02

0.000000000000000000e+00 1.430000000000000000e+02 1.500000000000000000e+03 1.020000000000000000e+03 9.780000000000000000e+02 8.930000000000000000e+02 4.280000000000000000e+02 3.680000000000000000e+02 1.180000000000000000e+02 5.600000000000000000e+01

0.000000000000000000e+00 0.000000000000000000e+00 1.280000000000000000e+02 1.290000000000000000e+02 1.280000000000000000e+02 1.120000000000000000e+02 1.440000000000000000e+02 1.230000000000000000e+02 1.560000000000000000e+02 6.600000000000000000e+01

0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 1.043000000000000000e+03 1.002000000000000000e+03 9.250000000000000000e+02 4.400000000000000000e+02 3.670000000000000000e+02 1.180000000000000000e+02 5.200000000000000000e+01

0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 1.147000000000000000e+03 1.093000000000000000e+03 4.480000000000000000e+02 4.140000000000000000e+02 1.200000000000000000e+02 6.100000000000000000e+01

0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 1.104000000000000000e+03 4.500000000000000000e+02 4.100000000000000000e+02 1.200000000000000000e+02 7.200000000000000000e+01

0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 4.490000000000000000e+02 4.100000000000000000e+02 1.170000000000000000e+02 6.200000000000000000e+01

0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 4.460000000000000000e+02 1.250000000000000000e+02 4.500000000000000000e+01

0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 1.240000000000000000e+02 7.400000000000000000e+01

0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 6.700000000000000000e+01

0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00 0.000000000000000000e+00III(3)

p-value for P15172 and Q10574: 1.000000e-03

p-value for P15172 and O95363: 5.200000e-01

I did 999 trials for each.

II (b) One sentence summary of the protein's (known) function. Based on that, write whether you expect each protein to be similar to the others or not.

1. MYOD\_HUMAN: An activator gene that promotes transcription of muscle-specific genes and helps with muscle differentiation. Since some other genes are from humans and muscle or myoblast-related as well, I expect this protein to be similar to the others.
2. TAL1\_HUMAN: Associated with hemopoietic—or blood cellular—malignancies, including leukemia. It is expected to be related to blood cell differentiation as well. Other genes seem to be more muscle cell-related, which leads me to expect this protein to be less similar to the others.
3. MYOD1\_MOUSE: An activator gene that promotes transcription of muscle-specific genes and helps with muscle differentiation. This gene’s description is near-identical to that of MYOD1\_HUMAN, which makes me believe that this gene will end up being similar to others.
4. MYOD1\_CHICK: In the chicken, acts as a activator gene that promotes transcription of muscle-specific genes and plays a role in muscle differentiation. Again, this is similar to other genes on the list who are both MyoD and muscle development-related genes, which makes me expect this gene to be similar to the others.
5. MYODA\_XENLA: In the African clawed frog, may act as a transcriptional activator that promotes transcription of muscle-specific target genes and plays a role in muscle differentiation. This is very similar to other organisms with a MyoD- or muscle-related functions, so I do expect this gene to be similar to the others.
6. MYOD1\_DANRE: In zebrafish, may act as a transcriptional activator that promotes transcription of muscle-specific target genes and plays a role in muscle differentiation. In the same way as MYODA\_XENLA, I would expect this gene to be similar to the others since there seems to be a common function, and zebrafish are also a terrific model organism.
7. Q8IU24\_BRABE: No known function, but is expected to be associated with MyoD and muscle organ development. Since this organism may be key to the evolution of vertebrates, I expect this gene to be somewhat similar to the others, but maybe not as much.
8. MYOD\_DROME: Is expected to be important to muscle development in the fruit fly. This seems pretty similar to other genes’ association with muscle development and MyoD, so I expect this gene to be pretty similar to the others.
9. LIN32\_CAEEL: This gene in C. elegans is expected to be involved in the development and differentiation of neurons. Since this gene deals with nerve-related functions rather than muscles, I don’t expect this gene to be similar to the others.
10. SYFM\_HUMAN: Linked to dealing with tRNA in mitochondrial translation and other tRNA functions, such as delivering tRNA to the ribosome and incorporating damaged amino acid into proteins. This doesn’t seem as explicitly MyoD- or muscle-related as the other genes, although it is in humans, so I don’t expect this gene to be as similar to the others.

Reflection

After seeing the scores, my predicted similarities have changed very slightly, but generally have stayed the same, with MyoD-related genes being quite similar to each other, while other genes were not as similar to the others.