# Results for test data

Identifiers N1 (deadly) and N2(ddgearlyk)  
Score for optimal alignment : **20**

N1 1 d-eadly

d ea ly

N2 2 dgearly

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | d | d | g | e | a | r | l | y | k |
|  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| d | 0 | 6 | 6 | 2 | 2 | 0 | 0 | 0 | 0 | 0 |
| e | 0 | 2 | 8 | 4 | 7 | 3 | 0 | 0 | 0 | 1 |
| a | 0 | 0 | 4 | 8 | 4 | 11 | 7 | 3 | 0 | 0 |
| d | 0 | 6 | 6 | 4 | 10 | 7 | 9 | 5 | 1 | 0 |
| l | 0 | 2 | 2 | 2 | 6 | 9 | 5 | 13 | 9 | 5 |
| y | 0 | 0 | 0 | 0 | 2 | 5 | 7 | 9 | 20 | 16 |

emperical p-value for k 30, N 999 = 0.031

# Local Sequence Alignment

### Upper Triangle showing score

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | P15172 | P17542 | P10085 | P16075 | P13904 | Q90477 | Q8IU24 | P22816 | Q10574 | O95363 |
| P15172 |  | 143 | 1500 | 1020 | 978 | 893 | 428 | 368 | 118 | 56 |
| P17542 |  |  | 128 | 129 | 128 | 112 | 144 | 123 | 156 | 66 |
| P10085 |  |  |  | 1043 | 1002 | 925 | 440 | 367 | 118 | 52 |
| P16075 |  |  |  |  | 1147 | 1093 | 448 | 414 | 120 | 61 |
| P13904 |  |  |  |  |  | 1104 | 450 | 410 | 120 | 72 |
| Q90477 |  |  |  |  |  |  | 449 | 410 | 117 | 62 |
| Q8IU24 |  |  |  |  |  |  |  | 446 | 125 | 45 |
| P22816 |  |  |  |  |  |  |  |  | 124 | 74 |
| Q10574 |  |  |  |  |  |  |  |  |  | 67 |
| O95363 |  |  |  |  |  |  |  |  |  |  |

# Empirical p-value

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Identifier 1 | Identifier 2 | K | N – Random Trials | p-Value |
| P15172 | Q10574 | 0 | 100000 | 9.99990000099999E-06 |
| P15172 | O95363 | 55940 | 100000 | 0.55940440595594 |

# Expectation on similarity

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Species** | **Accession** | **Summary** | **Similarity** |
| **1** | [Homo sapiens (Human)](https://courses.cs.washington.edu/courses/csep527/18wi/hw/hw2pics/sapiens.jpg) | P15172 | promotes transcription of muscle-specific target genes and plays a major role in regulating [muscle differentiation](https://en.wikipedia.org/wiki/Myogenesis). | Expect it to be more similar to P10085, P16075 and relatively less similar to P13904, P22816, Q90477, Q8IU24 and no similarity to the others |
| **2** | [Homo sapiens (Human)](https://courses.cs.washington.edu/courses/csep527/18wi/hw/hw2pics/sapiens.jpg) | P17542 | Plays a role in the cause of blood cancer and also in hemopoietic differentiation. Serves as a positive regulator in differentiation of oxygen carrying RBC - erythroid | Expect not to be similar to the other proteins |
| **3** | [Mus musculus (Mouse)](https://courses.cs.washington.edu/courses/csep527/18wi/hw/hw2pics/Mouse.jpg) | P10085 | promotes transcription of muscle-specific target genes and plays a major role in regulating [muscle differentiation](https://en.wikipedia.org/wiki/Myogenesis). | Expect it to be more similar to P15172, P16075 and relatively less similar to P13904, P22816, Q90477, Q8IU24 and no similarity to the others |
| **4** | [Gallus gallus (Chicken)](https://courses.cs.washington.edu/courses/csep527/18wi/hw/hw2pics/Chicken.jpg) | P16075 | promotes transcription of muscle-specific target genes and plays a major role in regulating [muscle differentiation](https://en.wikipedia.org/wiki/Myogenesis). | Expect it to be more similar to P15172, P10085 and relatively less similar to P13904, P22816, Q90477, Q8IU24 and no similarity to the others |
| **5** | [Xenopus laevis (African clawed frog)](https://courses.cs.washington.edu/courses/csep527/18wi/hw/hw2pics/Xenopus.jpg) | P13904 | May promote transcription of muscle-specific target genes and plays a major role in regulating [muscle differentiation](https://en.wikipedia.org/wiki/Myogenesis) | Expect it to be more similar to Q90477 and relatively less similar to P22816, P15172, P10085, Q8IU24 and no similarity to the others |
| **6** | [Danio rerio (Zebrafish)](https://courses.cs.washington.edu/courses/csep527/18wi/hw/hw2pics/Zebrafish.jpg) | Q90477 | May promote transcription of muscle-specific target genes and plays a major role in regulating [muscle differentiation](https://en.wikipedia.org/wiki/Myogenesis) | Expect it to be more similar to P13904and relatively less similar to P22816, P15172, P10085, Q8IU24 and no similarity to the others |
| **7** | [Branchiostoma belcheri (Amphioxus)](https://courses.cs.washington.edu/courses/csep527/18wi/hw/hw2pics/Amphioxus.jpg) | Q8IU24 | Related to MYOD genes and plays a role in the formation of the notochord muscle in embryos | Expect it to be similar to P22816 and relatively less similar to Q90477, P13904, P16075, P10085, P15172 and no similarity to others |
| **8** | [Drosophila melanogaster (Fruit fly)](https://courses.cs.washington.edu/courses/csep527/18wi/hw/hw2pics/Drosophila.jpg) | P22816 | May play an important role in the early development of muscle. | Expect it to be similar to Q8IU24 and relatively less similar to Q90477, P13904, P16075, P10085, P15172 and no similarity to others |
| **9** | [Caenorhabditis elegans](https://courses.cs.washington.edu/courses/csep527/18wi/hw/hw2pics/Caenorhabditis.jpg) | Q10574 | Essential for the specification of the neuroblast cell (cell from which nervous tissue is formed) fate in the development of peripheral sense organs. | Expect not to be very similar to others |
| **10** | [Homo sapiens (Human)](https://courses.cs.washington.edu/courses/csep527/18wi/hw/hw2pics/sapiens.jpg) | O95363 | Is responsible for the charging of tRNA(Phe) with phenylalanine in mitochondrial translation | Expect not to be very similar to others |

## Brief Reflection

My expectation on the similarity of the amino acid sequences matched with the score obtained from local sequence alignment. For example, I had expected O95363, Q10574, P17542 to not be similar to the others, the local alignment score also show that they have lower scores (< 156). Another one is I expected P15172 to be more similar to P10085, P16075 - as they play major role in muscle differentiation - and relatively less similar to P13904, P22816, Q90477, Q8IU24 - as these sequences though related to MYOD some were impacting early development of muscle - and had expected no similarity to the others. The local alignment score shows the same.