# Post-Lab Activity

## Analysis and Interpretation of Results

**Detailed Gel Analysis**

**Does molecular evidence support or refute the theory of evolution?**

**Does your molecular evidence support or refute your predictions?**

Create a cladogram using your results to find out.

From your gel you can create a cladogram based on proteins that the fish have in common. You can then determine whether your cladogram supports your predictions and/or matches the evolutionary relatedness of the fish species determined by morphological analysis in the evolutionary tree provided. Each protein band that a fish has in common with another fish is a shared characteristic. Cladistic analysis assumes that when two organisms share a characteristic, they had a common ancestor that had that characteristic, and this can be represented as a node on a cladogram with two branches coming from that node representing the descendent organisms.

In this exercise you will define the shared characteristics (i.e., make a list of all the different proteins in fish muscle), find which proteins (characteristics) are shared between fish, and construct a cladogram based on your data.

## Procedures

### Generate a standard curve to calculate protein sizes

The different protein bands in your gel can be defined by their different molecular masses. Indeed many proteins are named for their molecular weights. For example p53, a protein implicated in tumor progression is 53 kD in size. To determine the molecular masses of the proteins, a standard curve is created plotting the known molecular masses of the proteins in the Precision Plus Protein Kaleidoscope prestained standards against the distance they have migrated down the gel from the base of the well.

A 15% polyacrylamide gel is designed to separate small proteins- proteins less than 40 kD. Your gel analysis will concentrate on this size range. Note: If a different percentage acrylamide gel or an agarose gel has been run, analyze the section of the gel that has the best separation.

1. As shown in the figure below draw a line between the 37 and 25 kD bands of the prestained standards. Your gel analysis will be restricted to the proteins below this line.

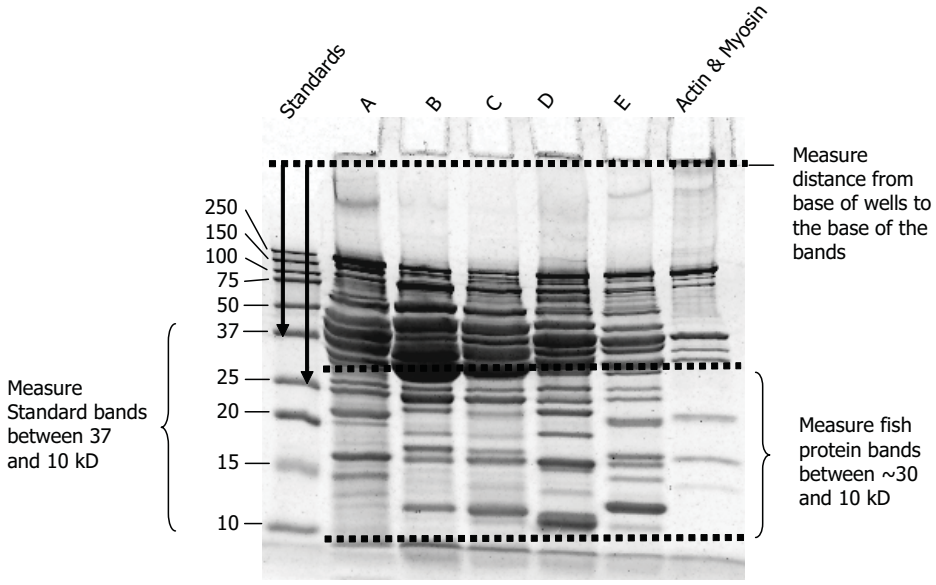



**Fig. 18. Image of fish muscle proteins separated by SDS-PAGE and stained with Bio-Safe Coomassie stain.** Lines illustrate measurement of bands for constructing the standard curve.

2. To create the standard curve measure and record in the table below the distances the five sub 40 kD protein bands of the prestained standard have migrated from the base of the well i.e. measure the 37, 25, 20, 15 and 10 kD bands. Accuracy to 0.5 mm is required.

| Precision Plus Protein Kaleidoscope Prestained Standards Molecular Weight (kD) | Distance Migrated (mm) |
|---|---|
| 37 | |
| 25 | |
| 20 | |
| 15 | |
| 10 | |

3. On the graph paper provided, plot the distances migrated in mm on the x-axis against the molecular masses of the prestained protein bands in kD on the y-axis as a scatter plot. Draw a line of best fit through the points. On semi-logarithmic graph paper with the molecular mass of the proteins on a logarithmic scale the data should result in a linear (straight line) curve.

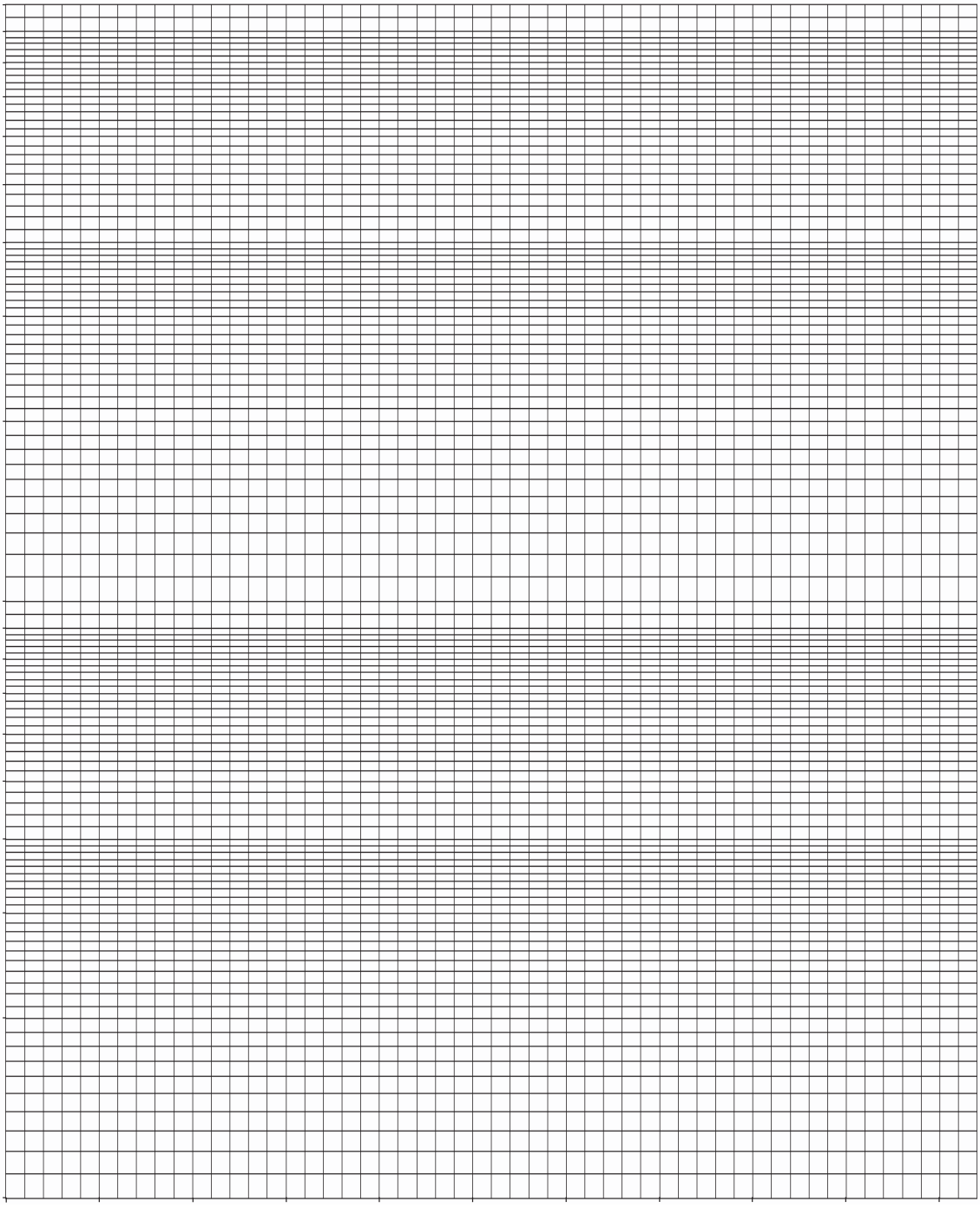


**Fig. 19. Two-cycle semi-logarithmic graph paper to construct curve of the protein molecular mass against the distance migrated.**

**Define the characteristics (proteins) of the different fish**

4. For each fish sample that has been analyzed, determine the molecular masses of the proteins below the 25-37 kD line. Measure the distance each band has migrated from the base of its well. Find that distance on the x-axis of the standard curve. Draw a line up from the x-axis to the curve. Read across to the y-axis to determine the molecular mass.

Alternatively, use graphing software to generate the standard curve. Make a line of best fit (or trend line) through the points and formulate an equation to calculate the mass of the unknown proteins on the gel.

5. Enter this data into a table with the molecular masses of the proteins for each fish (see example below).

| Fish Species A | |
|---|---|
| Distance Migrated (mm) | Molecular Mass (kD) |
| 25 | 32.5 |
| 26.5 | 31 |
| 29 | 28.6 |
| 36 | 21.7 |
| 36.5 | 21.2 |
| 39 | 18.8 |
| 44 | 13.9 |
| 52 | 6 |

**Determine which fish have each characteristic (protein)**

6. Make a table with a row for every band size you have recorded for all your fish samples and a column for each type of fish on your gel. Then make a mark in each cell of the table where the fish has that size band (see example below).

**Characteristic**

| Distance Migrated (mm) | Protein Molecular Mass (kDa) | Species A | Species B | Species C | Species D | Species E |
|---|---|---|---|---|---|---|
| 25 | 32.5 | X | | | | |
| 26 | 31.5 | | X | X | X | X |
| 26.5 | 31.0 | X | | | | |
| 27.5 | 30.0 | | X | X | X | X |
| 28.5 | 29.1 | | | | | |
| 29 | 28.6 | X | X | X | X | |
| 30 | 27.6 | | | X | | X |
| 30.5 | 27.1 | | | | | X |
| 32 | 25.6 | | X | X | X | |
| 33 | 24.7 | | | | | X |
| 34.5 | 23.2 | | X | X | | |
| 35.5 | 22.2 | | | | | X |
| 36 | 21.7 | X | | | | |
| 36.5 | 21.2 | X | X | X | X | |
| 37 | 20.7 | | | | | X |
| 37.5 | 20.2 | | X | X | | |
| 38 | 19.7 | | | | X | |
| 38.5 | 19.3 | | | | X | |
| 39 | 18.8 | X | | | | X |
| 39.5 | 18.3 | | | | | X |
| 40.5 | 17.3 | | X | X | | |
| 41 | 16.8 | | | | X | |
| 41.5 | 16.3 | | | | | |
| 42 | 15.8 | | X | X | | X |
| 43 | 14.8 | | | | | |
| 44 | 13.9 | X | | | | X |
| 45 | 12.9 | | X | X | | |
| 46 | 11.9 | | | | X | |
| 46.5 | 11.4 | | | X | | |
| 47 | 10.9 | | | | | X |
| 47.5 | 10.4 | | | | X | |
| 51.5 | 6.5 | | | X | | |
| 52 | 6.0 | X | | | | |
| | **COUNT** | **8** | **10** | **13** | **10** | **12** |

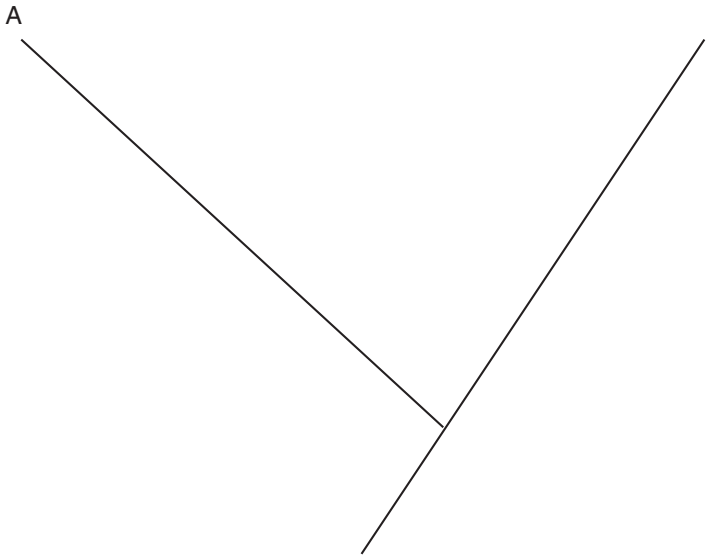
**Find the number of characteristics shared by each of the fish**

7. In the table below both the row and column headings are the types of fish. From the table above, separately compare the number of bands (X's) in common with every other fish sample from your gel and put those numbers into the table below, such that each fish is individually compared with every other fish. In this example, species A and B have just 2 bands in common while species B and C have 10 bands in common. Your table will be the basis for drawing your cladogram.

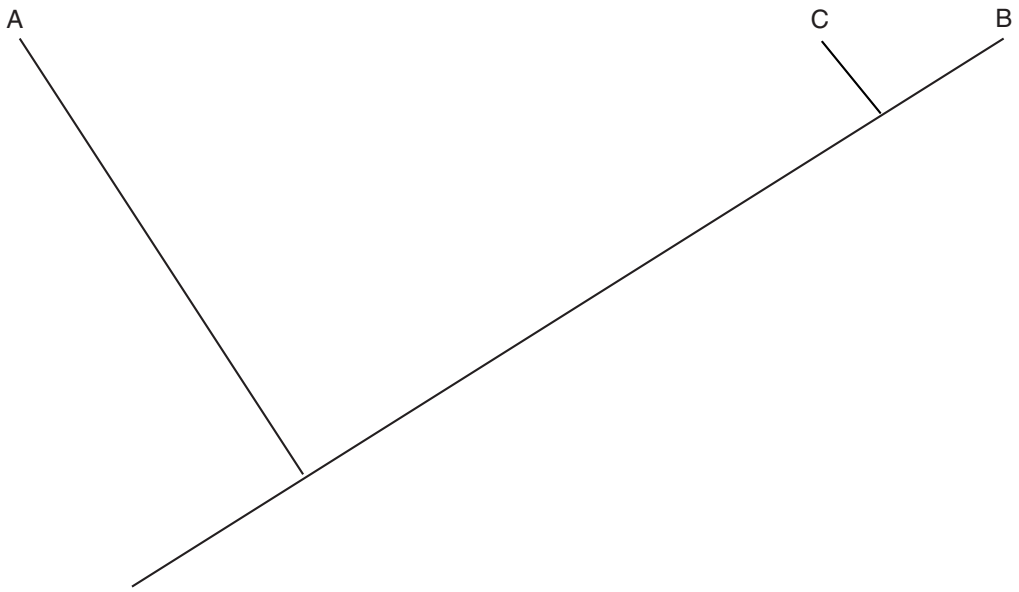
| | **Species A** | **Species B** | **Species C** | **Species D** | **Species E** |
|---|---|---|---|---|---|
| **Species A** | 8 | 2 | 2 | 2 | 2 |
| **Species B** | | 10 | 10 | 5 | 3 |
| **Species C** | | | 13 | 5 | 4 |
| **Species D** | | | | | 2 |
| **Species E** | | | | | 12 |

**Construct your cladogram**

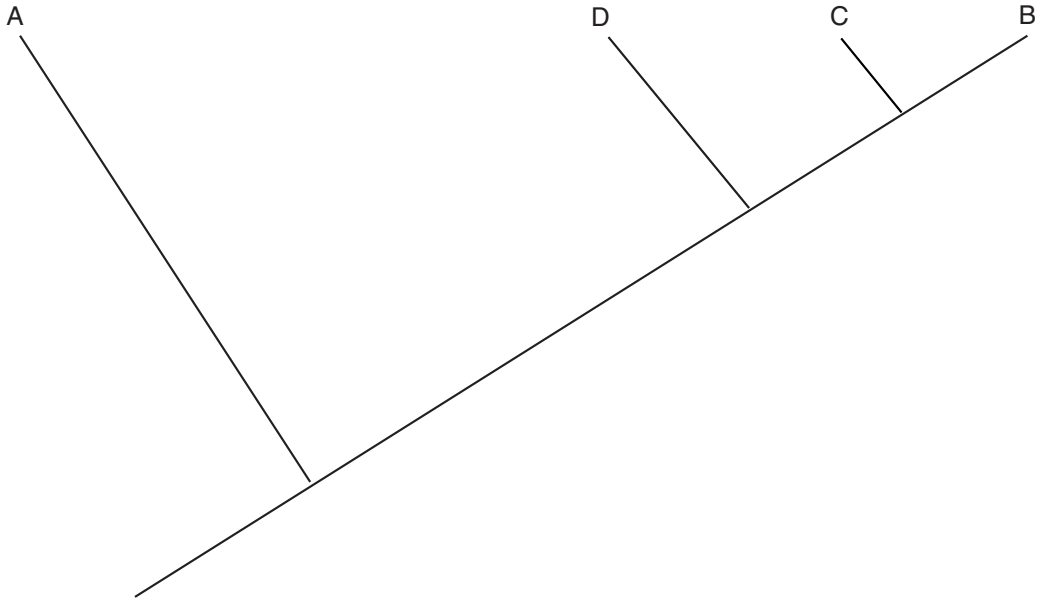
8. Now you are ready to construct your cladogram. First draw a line to form the trunk of your cladogram. Find the fish with the least bands in common. In the example above it is species A, which has only 2 bands in common with any of the other fish. Then draw a side branch off the line near the bottom of the trunk and label that branch with the fish's name, in this case, species A. This fish is the outlier, i.e., it is the least similar to any of the others. The node (where the side branch meets the trunk) represents an ancestor that is common to all the fish in this analysis.




Now, find the two fish with the most bands in common (in this example it is species B and C, which have 10 bands in common). Draw a side branch off the trunk near the top and label the two ends with the fishes' names, in this case, species B and species C (it doesn't matter which branch gets which label). The node represents a common ancestor of species B and species C that had all the same characteristics (proteins).




Now, identify those fish species with the next most bands in common. In this example, species D has five bands in common with species B and species C, which indicates species D is the same cladistic distance from B and C (i.e. species D is not more closely related to either B or C). Draw a branch further down the trunk. This node represents an ancestor that is common to species B, C, and D that had these 5 characteristic proteins.

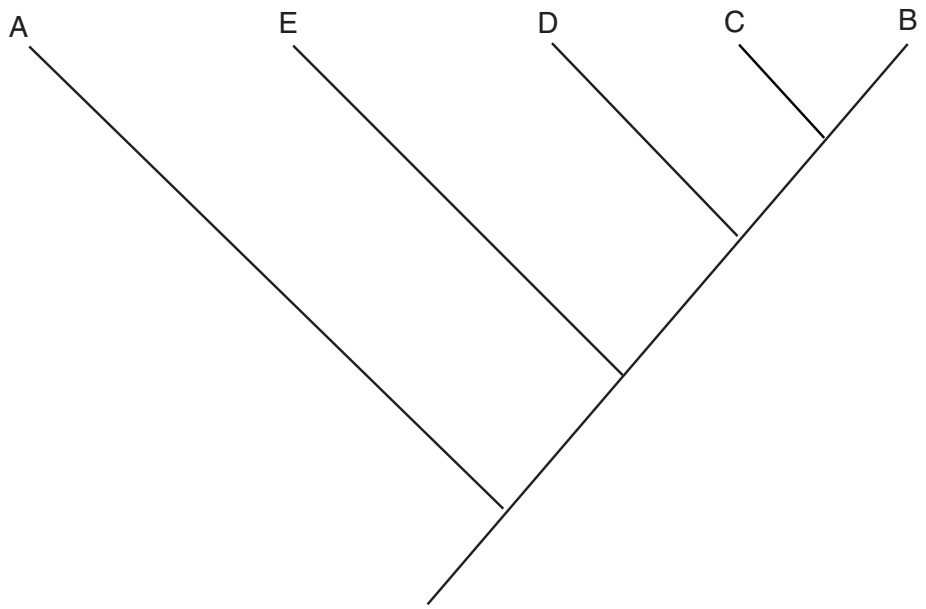


The last fish to add to the cladogram in this example is species E, which shares four bands with species C, three bands with species B, and only two bands with species A and D. This fish may seem trickier to place than the others because it shares more characteristics with species B and C than it does with D, but D shares more characteristics with B and C than E does. So, to place this fish you might ask:

Does species E share the five proteins that the common ancestor of species B, C, and D had? Answer (no).

Does species E share more proteins with B, C, and D than A? Answer (yes).

Therefore, species E gets its own branch in between the D and A branches to indicate that it has more shared characteristics with B, C, and D than A, but fewer shared characteristics with B and C than D.




Now compare your cladogram with your original predictions. Write your deductions below.

# Appendix B: Using databases to obtain real amino acid sequence data to create cladograms

In order to determine how closely related species are, scientists often will study amino acid sequences of essential proteins. Any difference in the amino acid sequence is noted and a phylogenetic tree is constructed based on the number of differences. More closely related species have fewer differences (i.e., they have more amino acid sequence in common) than more distantly related species.

There are many tools scientists can use to compare amino acid sequences of muscle protein. One such tool is the National Center for Biotechnology Information protein databases (http://www.ncbi.nlm.nih.gov/). By entering the amino acid sequence of a protein you are interested in, the BLAST search tool compares that sequence to all others in its database. The data generated provides enough information to construct cladograms.

The purpose of this activity is to use data obtained from NCBI to construct an evolutionary tree based on the amino acid sequences of the myosin heavy chain. In this example we have input a 60 amino acid sequence from myosin heavy chain of rainbow trout and then pulled out matching sequences using BLAST, which include chum salmon, zebra fish, common carp, and bluefin tuna, and then compared each of these sequences with each other.

You may either use the data provided below or go online and obtain data directly by performing BLAST searches. A quick guide to performing BLAST searches is given at the end of this activity.

The data below was obtained by entering a 60 amino acid sequence from the heavy myosin chain of rainbow trout (*Oncorhynchus mykiss*). The database search tool returned all sequences that were a close match. The results are formatted as such:

```
gi|755771|emb|CAA88724.1|  myosin heavy chain [Oncorhynchus mykiss]
        Length=698

Score =  119 bits (299),  Expect = 2e-26
Identities = 60/60 (100%), Positives = 60/60 (100%), Gaps = 0/60 (0%)

Query  1
VAKAKGNLEKMCRTLEDQLSELKTKNDENVRQVNDISGQRARLLTENGEFGRQLEEKEAL  60

VAKAKGNLEKMCRTLEDQLSELKTKNDENVRQVNDISGQRARLLTENGEFGRQLEEKEAL

Sbjct  1
VAKAKGNLEKMCRTLEDQLSELKTKNDENVRQVNDISGQRARLLTENGEFGRQLEEKEAL  60
```

The value for 'identities' is the number of amino acids exactly in common, the value for 'positives' is the number of amino acids that are similar to each other (such as serine and threonine), and the value for 'gaps' is the number of amino acid positions that are absent one of the sequences. 'Query' is the original trout sequence, 'Sbjct' is the aligned sequence, and the middle sequence shows the mismatches: a '+' indicates a positive and a space indicates a mismatch that is not a positive. There are resources on the NCBI web site to help you understand more about the information a BLAST search generates.

The data below compares rainbow trout to salmon, zebra fish, carp, and tuna, and then compares salmon to zebra fish, carp, and tuna, then zebra fish to carp and tuna, and finally carp to tuna.

Use the data provided to determine how many amino acid differences exist between the organisms. Organize your data in charts.

### Rainbow trout compared to Chum Salmon (*Oncorhynchus keta)*

```
gi|21623523|dbj|BAC00871.1|  myosin heavy chain [ Oncorhynchus keta]

        Length=1937

Score =  119 bits (299),  Expect = 2e-26
Identities = 60/60 (100%), Positives = 60/60 (100%), Gaps = 0/60 (0%)

Query  1
AKAKGNLEKMCRTLEDQLSELKTKNDENVRQVNDISGQRARLLTENGEFGRQLEEKEAL  60

VAKAKGNLEKMCRTLEDQLSELKTKNDENVRQVNDISGQRARLLTENGEFGRQLEEKEAL

Sbjct  1240
VAKAKGNLEKMCRTLEDQLSELKTKNDENVRQVNDISGQRARLLTENGEFGRQLEEKEAL  1299
```

### Rainbow Trout compared to Zebra Fish (*Danio rerio*)

```
gi|68360600|ref|XP_708916.1|  PREDICTED: myosin, heavy polypeptide 1, skeletal muscle
[ Danio rerio]

         Length=2505

Score =  108 bits (269),  Expect = 6e-23
Identities = 52/60 (86%), Positives = 57/60 (95%), Gaps = 0/60 (0%)

Query  1
VAKAKGNLEKMCRTLEDQLSELKTKNDENVRQVNDISGQRARLLTENGEFGRQLEEKEAL  60

VAKAK NLEKMCRTLEDQLSE+K+KNDEN+RQ+ND+S QRARL TENGEFGRQLEEKEAL

Sbjct  1240
VAKAKANLEKMCRTLEDQLSEIKSKNDENLRQINDLSAQRARLQTENGEFGRQLEEKEAL  1299
```

### Rainbow Trout compared to Common Carp (*Cyprinus carpio*)

```
gi|806515|dbj|BAA09069.1|  myosin heavy chain [ Cyprinus carpio]

        Length=955

Score =  104 bits (259),  Expect = 8e-22
Identities = 51/60 (85%), Positives = 56/60 (93%), Gaps = 0/60 (0%)

Query  1
VAKAKGNLEKMCRTLEDQLSELKTKNDENVRQVNDISGQRARLLTENGEFGRQLEEKEAL  60

VAKAK NLEKMCRTLEDQLSE+KTK+DENVRQ+ND++ QRARL TENGEF RQLEEKEAL

Sbjct  259
VAKAKANLEKMCRTLEDQLSEIKTKSDENVRQLNDMNAQRARLQTENGEFSRQLEEKEAL  318
```

### Rainbow Trout compared to Bluefin Tuna (*Thunnus thynnus*)

```
gi|1339977|dbj|BAA12730.1|  skeletal myosin heavy chain [ Thunnus thynnus]

         Length=786

Score =  104 bits (259),  Expect = 8e-22
Identities = 49/60 (81%), Positives = 57/60 (95%), Gaps = 0/60 (0%)

Query  1
VAKAKGNLEKMCRTLEDQLSELKTKNDENVRQVNDISGQRARLLTENGEFGRQLEEKEAL  60

VAK+KGNLEKMCRT+EDQLSELK KNDE+VRQ+ND++GQRARL TENGEF RQ+EEK+AL

Sbjct  88
VAKSKGNLEKMCRTIEDQLSELKAKNDEHVRQLNDLNGQRARLQTENGEFSRQIEEKDAL  147
```

### Chum Salmon compared to Zebra Fish

```
gi|68360600|ref|XP_708916.1|  PREDICTED: myosin, heavy polypeptide 1, skeletal muscle
[ Danio rerio]

        Length=2505

Score =  108 bits (269),  Expect = 6e-23
Identities = 52/60 (86%), Positives = 57/60 (95%), Gaps = 0/60 (0%)

Query  1
VAKAKGNLEKMCRTLEDQLSELKTKNDENVRQVNDISGQRARLLTENGEFGRQLEEKEAL  60

VAKAK NLEKMCRTLEDQLSE+K+KNDEN+RQ+ND+S QRARL TENGEFGRQLEEKEAL

Sbjct  1240
VAKAKANLEKMCRTLEDQLSEIKSKNDENLRQINDLSAQRARLQTENGEFGRQLEEKEAL  1299
```

### Chum Salmon compared to Common Carp

```
gi|806515|dbj|BAA09069.1|  myosin heavy chain [ Cyprinus carpio]

         Length=955

Score =  104 bits (259),  Expect = 8e-22
Identities = 51/60 (85%), Positives = 56/60 (93%), Gaps = 0/60 (0%)

Query  1
VAKAKGNLEKMCRTLEDQLSELKTKNDENVRQVNDISGQRARLLTENGEFGRQLEEKEAL  60

VAKAK NLEKMCRTLEDQLSE+KTK+DENVRQ+ND++ QRARL TENGEF RQLEEKEAL

Sbjct  259
VAKAKANLEKMCRTLEDQLSEIKTKSDENVRQLNDMNAQRARLQTENGEFSRQLEEKEAL  318
```

### Chum Salmon compared to Bluefin Tuna

```
gi|1339977|dbj|BAA12730.1|  skeletal myosin heavy chain [ Thunnus thynnus]

        Length=786

Score =  104 bits (259),  Expect = 8e-22
Identities = 49/60 (81%), Positives = 57/60 (95%), Gaps = 0/60 (0%)

Query  1
VAKAKGNLEKMCRTLEDQLSELKTKNDENVRQVNDISGQRARLLTENGEFGRQLEEKEAL  60

VAK+KGNLEKMCRT+EDQLSELK KNDE+VRQ+ND++GQRARL TENGEF RQ+EEK+AL

Sbjct  88
VAKSKGNLEKMCRTIEDQLSELKAKNDEHVRQLNDLNGQRARLQTENGEFSRQIEEKDAL  147
```

### Zebra Fish compared to Common Carp

```
gi|806515|dbj|BAA09069.1|  myosin heavy chain [ Cyprinus carpio]

        Length=955

Score =  108 bits (271),  Expect = 4e-23
Identities = 53/60 (88%), Positives = 59/60 (98%), Gaps = 0/60 (0%)

Query  1
VAKAKANLEKMCRTLEDQLSEIKSKNDENLRQINDLSAQRARLQTENGEFGRQLEEKEAL  60

VAKAKANLEKMCRTLEDQLSEIK+K+DEN+RQ+ND++AQRARLQTENGEF RQLEEKEAL

Sbjct  259
VAKAKANLEKMCRTLEDQLSEIKTKSDENVRQLNDMNAQRARLQTENGEFSRQLEEKEAL  318
```

### Zebra Fish compared to Bluefin Tuna

```
gi|1339977|dbj|BAA12730.1|  skeletal myosin heavy chain [ Thunnus thynnus]

        Length=786

Score =  102 bits (253),  Expect = 4e-21
Identities = 47/60 (78%), Positives = 57/60 (95%), Gaps = 0/60 (0%)

Query  1
VAKAKANLEKMCRTLEDQLSEIKSKNDENLRQINDLSAQRARLQTENGEFGRQLEEKEAL  60

VAK+K NLEKMCRT+EDQLSE+K+KNDE++RQ+NDL+ QRARLQTENGEF RQ+EEK+AL

Sbjct  88
VAKSKGNLEKMCRTIEDQLSELKAKNDEHVRQLNDLNGQRARLQTENGEFSRQIEEKDAL  147
```

**Common Carp compared to Bluefin Tuna**

```
gi|1339977|dbj|BAA12730.1|  skeletal myosin heavy chain [ Thunnus thynnus]

Length=786

Score =  104 bits (259),  Expect = 9e-22
Identities = 49/60 (81%), Positives = 57/60 (95%), Gaps = 0/60 (0%)

Query  1
VAKAKANLEKMCRTLEDQLSEIKTKSDENVRQLNDMNAQRARLQTENGEFSRQLEEKEAL  60

VAK+K NLEKMCRT+EDQLSE+K K+DE+VRQLND+N QRARLQTENGEFSRQ+EEK+AL

Sbjct  88
VAKSKGNLEKMCRTIEDQLSELKAKNDEHVRQLNDLNGQRARLQTENGEFSRQIEEKDAL  147
```

1. Construct a table of your data containing the number of amino acid differences between each of the different fish.

| | Rainbow Trout | Chum Salmon | Zebra Fish | Common Carp | Bluefin Tuna |
|---|---|---|---|---|---|
| **Rainbow Trout** | **0** | | | | |
| **Chum Salmon** | **X** | **0** | | | |
| **Zebra Fish** | **X** | **X** | **0** | | |
| **Common Carp** | **X** | **X** | **X** | **0** | |
| **Bluefin Tuna** | **X** | **X** | **X** | **X** | **0** |

2. Which two fish share the most amino acids in their myosin heavy chains based on your data?

3. Which two fish share the fewest amino acids?

4. Are there any fish that share more amino acids with each other than each does with the two fish in question one? If yes, which fish?

5. Construct a cladogram based on this data:

6. The myosin heavy chain of white croaker (*Pennahia argentata*) (BAB12571) has the following amino acid differences with the five fish above.

| | **Rainbow Trout** | **Chum Salmon** | **Zebra Fish** | **Common Carp** | **Bluefin Tuna** |
|---|---|---|---|---|---|
| **White Croaker** | 4 | 4 | 11 | 9 | 11 |

Add this fish to your cladogram and explain why you placed it where you did.

Taxonomic data can be derived from many sources: DNA sequences, protein sequences, morphology, and paleontology. Classification of organisms derives from these sources. Inconsistencies in the phylogenetic trees generated between molecular and taxonomic data emphasize why data from different sources is required to generate phylogenetic trees and why there is still much dispute in the field of phylogenetics on the correct placement of organisms within phylogenetic trees. Bear in mind that myosin heavy chain is around 1900 amino acids in length and our molecular data is based on a 60 amino acid region – just 3% of the entire protein. The amount of work required to process the small amount of data provided here also emphasizes the need for skilled bioinformaticists to process and analyze the vast amount of data generated by genomic and proteomic research.

Examine the taxonomic classification of the fishes below. The large phylogenetic tree figure will be useful for this exercise.

**Rainbow Trout** (*Oncorhynchus mykiss*)
Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

**Chum Salmon** (*Oncorhynchus ket*a)
Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

**Zebra Fish** (*Danio rerio*)
Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

**Common Carp** (*Cyprinus carpio*)
Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Cyprinus.

**Bluefin Tuna** (*Thunnus thynnus*)
Vertebrata; Euteleostomi; ctinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei; Scombridae; Thunnus.

**White Croaker** (*Pennahia argentata*)
Vertebrata; Euteleostomi; ctinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; canthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei; Sciaenidae; Pennahia.

7. Construct a phylogenetic tree based on the taxonomic classification of the fishes above.

8. Does the taxonomic classification support the molecular data? Please explain your answer.

9. What reasons might there be for a discrepancy between the molecular data and the taxonomic classification.

10. Why do scientists need to examine multiple data sets before determining evolutionary relatedness?

**Quick Guide to BLAST searching**

Please note, this is a quick guide to obtain a list of fish myosin sequences, there are many refinements you can make to your search and many different ways to use BLAST searches. Further information can be found on the NCBI web site.

1) Go to http://www.ncbi.nlm.nih.gov/ and choose BLAST
2) Choose Protein BLAST.
3) Enter your myosin sequence into the search box.

Rainbow Trout Myosin Heavy Chain Protein Sequence (CAA88724):

VAKAKGNLEKMCRTLEDQLSELKTKNDENVRQVNDISGQRARLLTENGEFGRQLEEKEAL

4) Leave the other fields as found and hit the BLAST button.
5) A new window should pop up. Hit the Format button.
6) After a short wait the BLAST results window will come up and may well be hundreds of pages long — don't worry. There should be a long list of sequences that produced significant alignments. Although the search may pick up hundreds of sequences, they are in order of homology, so the ones you are interested in should be in the first 25 or so.
7) Further down the BLAST results page, after the list of sequences, each sequence will be aligned with the original trout sequence (as shown in the example) so that you can see how the two compare.
8) To compare your second fish, say bluefin tuna, with the other fish, you must perform a second BLAST search with the tuna sequence to obtain the protein alignments of tuna with the other fish. Alternatively, you can align 5 protein sequences yourself from your original search in a word processing document (use Courier font, this aligns sequences because all the letters are the same width) and have your students manually compare them.

# Appendix C: Glossary

**Actin** – major muscle protein organized into thin filaments

**Amino acids** – basic building blocks of proteins

**Anode** – positive electrode

**Bioinformatics** – use of data storage and analysis technologies to extract meaningful information from large quantities of biological data

**BME (β-mercaptoethanol)** – a reducing agent that breaks disulfide bonds

**Cathode** – negative electrode

**Charge density** – ratio of charge to mass of a protein

**Cladogram** – tree-like relationship-diagrams that demonstrate the evolutionary relatedness between organisms

**Codon** – a set of three nucleotides (DNA bases) that code for an amino acid

**Dalton (Da)** – unit of molecular weight equal to the mass of a hydrogen atom, 1.66 x 10–24 g

**Denature** – to disrupt a protein's 3-dimensional structure

**Disulfide bond** – S-S bond between amino acids in a polypeptide chain; contributes to tertiary and quaternary structure of proteins

**DTT (dithiothreitol)** – a reducing agent that breaks disulfide bonds

**Exon** – region of a gene that is translated into amino acids (compare to intron)

**Fingerprint** – distinct pattern of bands on a protein gel, useful as an identifying characteristic of a sample or species

**Gel electrophoresis** – technique used to separate molecules that carry electric charges. The molecules separate from each other according to the different rates at which they migrate through an electric field set up in a gel soaked in a chemical solution.

**Gene** – a defined region of DNA that encodes information for the synthesis of a single polypeptide

**Genome** – the entire complement of genes in an organism

**Genomics** – the study of all the nucleotide sequences in the chromosomes of an organism

**Homology** – similarity between genes of different species due to common ancestry

**Intron** – region of a gene that is not translated into amino acids (compare to exon)

**Kilodalton (kD)** – 1,000 daltons

**mRNA** – message derived from a gene, with information to make one polypeptide

**Myosin** – major muscle protein organized into thick filaments

**Native** – the natural structure of a protein or protein complex, as found within the organism

**PAGE** – polyacrylamide gel electrophoresis

**Phylogeny** – the evolutionary relationship of species based on lineage and history of descent

**Polypeptide** – a chain of amino acids

**Posttranscriptional modification** – alterations to mRNA that allow one gene to code for many proteins, such as alternate splicing

**Posttranslational modification** – alterations of proteins after they are synthesized by the cell, such as phosphorylation or cleavage

**Protein** – a functional assembly of one or more polypeptides, made of sequences of amino acids

**Protein folding** – the process by which a protein bends and twists to achieve its normal three-dimensional shape

**Proteome (protein complement expressed by a genome)** – the complete protein profile found under given conditions in a biological sample

**Proteomics** – the study of the proteome in specific cells, tissues, organs, organ systems, or organisms during a specific time period (e.g., during development)

**SDS** – sodium dodecyl sulfate

**SDS-PAGE** – sodium dodecyl sulfate-polyacrylamide gel electrophoresis; a form of electrophoresis that treats samples with SDS to denature proteins

**Transcription** – production of mRNA from DNA genetic information

**Translation** – production of a protein from messenger RNA (mRNA)

**tRNA** – transfer RNA which acts as adaptor molecule between mRNA and an amino acid



APPENDIX C