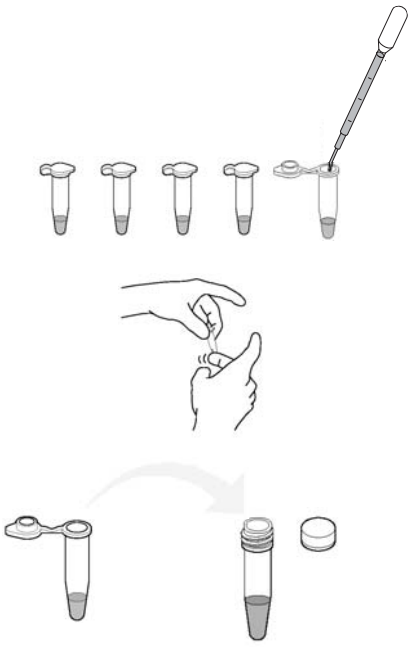
Comparative Proteomics: Protein Profiler Kit – Quick Guide

Lesson 1 Quick Guide



1 Label one 1.5 ml **fliptop** micro tube for each of five fish samples. Also label one **screwcap** micro tube for each fish sample.

2. Add 250 µl of Bio-Rad Laemmli sample buffer to each labeled **fliptop** microtube.

3. Cut a piece of each fish muscle about 0.25 x 0.25 x 0.25 cm^3 ( ) and transfer each piece into a labeled **fliptop** micro test tube. Close the lids.

4. Flick the microtubes 15 times to agitate the tissue in the sample buffer.

5. Incubate for 5 minutes at room temperature.

6. Carefully transfer the buffer by pouring from **each fliptop** microtube into a labeled **screwcap** microtube. Do not transfer the fish!

7. Heat the fish samples in screwcap microtubes for 5 minutes at 95°C.

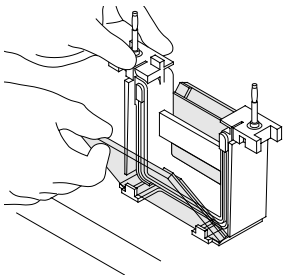
Lesson 2 Quick Guide

1. Set up Mini-PROTEAN 3 gel box and add 1x TGS electrophoresis buffer to the chamber.

2. Prepare a Ready Gel cassette by cutting along the black line on the bottom of the cassette with a razor blade and pulling off the plastic strip, as indicated on gel cassette.

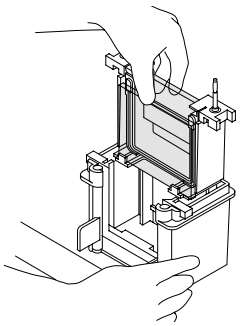
3. Remove the comb from the Ready Gel cassette.

4. Place Ready Gel cassette into the electrode assembly with the short plate facing inward. Place a buffer dam or another Ready Gel cassette on the opposite side of the electrode assembly, with notch on buffer dam facing inward.

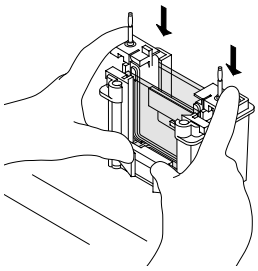


QUICK GUIDE

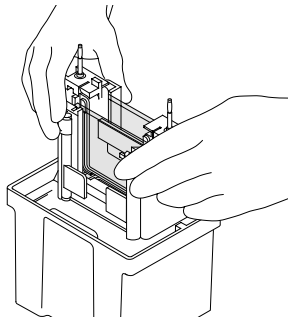
5. Slide gel cassette, buffer dam, and electrode assembly into the clamping frame.



6. Press down the electrode assembly while closing the two cam levers of the clamping frame.

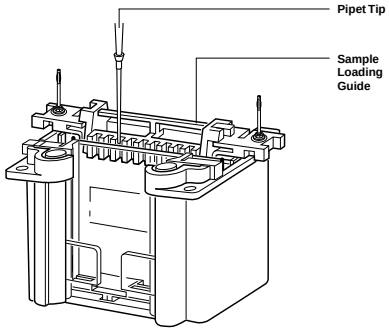


7. Lower the inner chamber into the mini tank.



8. Completely fill the inner chamber with 1x TGS electrophoresis buffer, making sure the buffer covers the short plate (~150 ml).
9. Fill mini tank approximately 200 ml of 1x TGS electrophoresis buffer.

10. Plate sample loading guide on top of the electrode assembly.

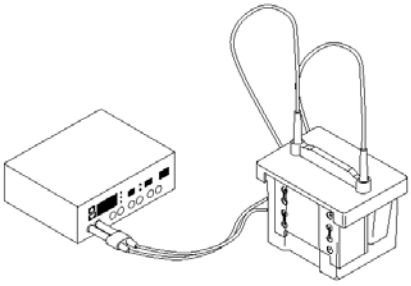


11. Heat fish samples and actin and myosin standard to 95°C for 2–5 min.

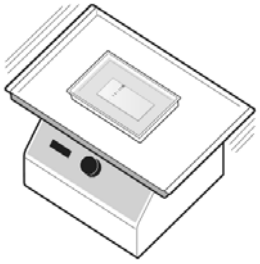
12. Load your gel:

Lane	Volume	Sample
1 & 2	empty	empty
3	10 µl	Precision Plus Protein Kaleidoscope prestained standards (Stds)
4	10 µl	fish sample 1
5	10 µl	fish sample 2
6	10 µl	fish sample 3
7	10 µl	fish sample 4
8	10 µl	fish sample 5
9	10 µl	actin and mysin standard (AM)
10	empty	empty

13. Electrophorese for 30 minutes at 200 V in 1x TGS electrophoresis buffer.

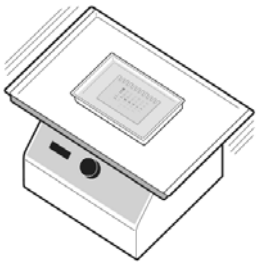


14. After electrophoresis, remove gel from cassette and transfer gel to a container with 40 ml Bio-Safe Coomassie blue stain and stain gel for 1 hour, with gentle shaking for best results.



Lesson 3 Quick Guide

1. Discard stain and destain gels in a large volume of water for at least 30 minutes overnight, changing the water at least once. Blue-stained bands will be visible on a clear gel after destaining.
2. Dry gels using GelAir cellophane.



QUICK GUIDE

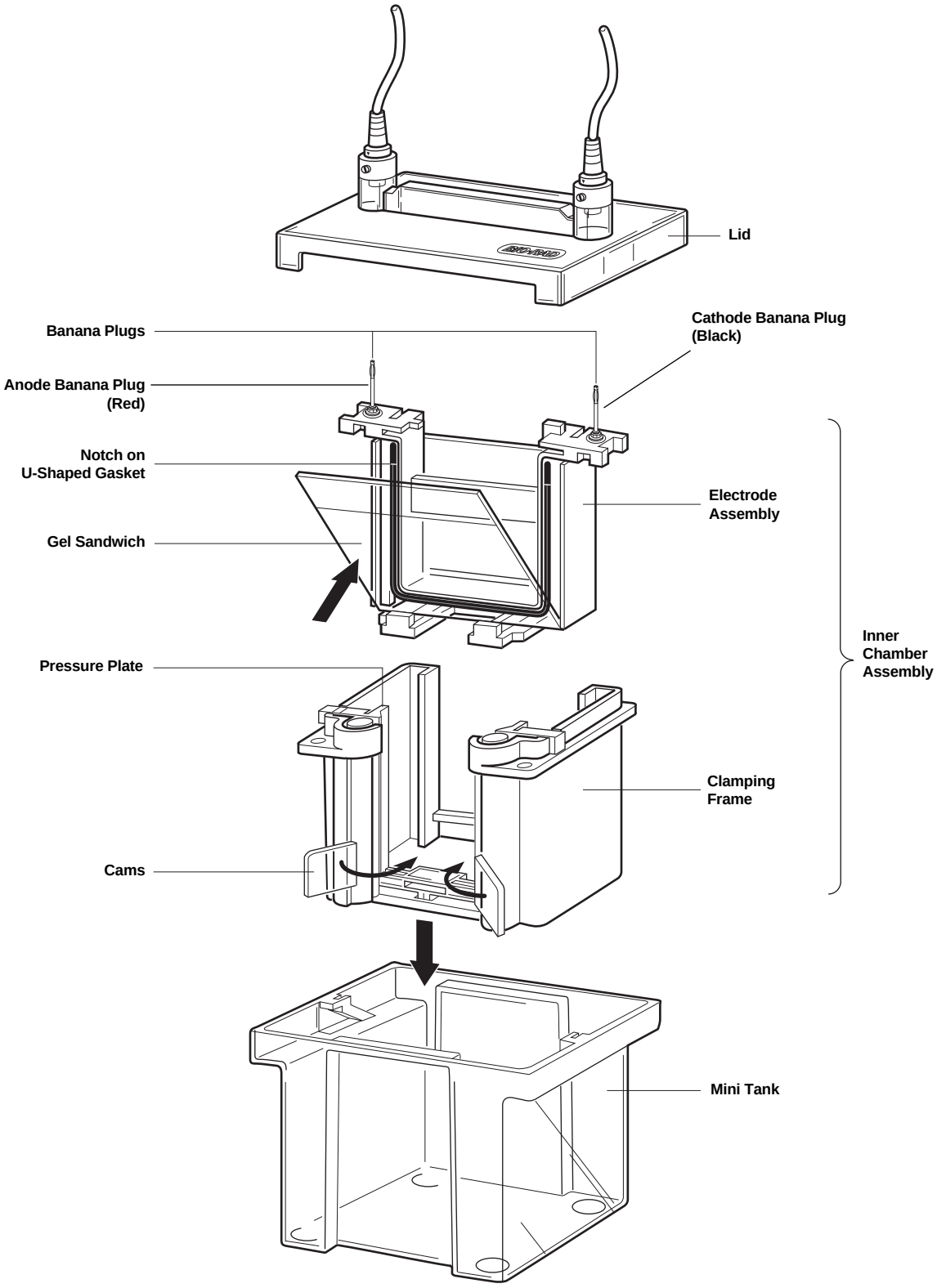


Fig. 11. Assembling the Mini-PROTEAN 3 cell.

Student Manual

Does molecular evidence support or refute the theory of evolution? DNA gets a lot of attention, but proteins do all the work. Proteins determine an organism's form, function, and phenotype. As such, proteins determine the traits that are the raw material of natural selection and evolution.

In this lab you will use protein gel electrophoresis, the technique most widely used in biotechnology research, to examine muscle proteins from closely and distantly related fish species, and to identify similarities and differences in these organisms' protein profiles, or fingerprints.

Analogous in principle to DNA fingerprinting, protein profiles can also reveal genetic similarities or differences, and from such molecular data it is possible to infer relatedness. Each protein band that a fish has in common with another fish is considered a shared characteristic. A fish family tree, or cladogram, can be constructed based on protein bands that the fish have in common. Cladistic analysis assumes that when two organisms share a common characteristic, they also share a common ancestor with that same characteristic.

Muscle protein consists mainly of actin and myosin, but numerous other proteins also make up muscle tissue. While actin and myosin are highly conserved across all animal species, the other proteins are more diverse, varying even among closely related species.

During this laboratory-based scientific investigation you are asked: Can molecular data show similarities and differences among species? You will compare the similarities and differences in the protein profiles of various fish species, create a cladogram (family tree) from your own gel results, and compare your data to published evolutionary data. Then you will be asked: Do the data agree? Why or why not? What explanations can you suggest?

Molecular biology has unlocked secrets of mystifying new diseases, given us the premier tools for defining biological identity, and created a pillar of data to support Darwin's theory of common descent. In short, molecular biology and its elegant techniques have revolutionized our understanding of life's origins and mechanisms.

Is it just genes that determine what proteins will be made? Current research in the field of proteomics suggests not. The following section is designed as a review of important background information for this laboratory investigation.

Background

Proteomics, Evolution, and Classification of Fish

Proteomics

Proteomics is the study of proteins, particularly their structures and functions. This term was coined to make an analogy with genomics, and while it is often viewed as the "next step", proteomics is much more complicated than genomics. Most importantly, while the genome is a rather constant entity, the proteome differs from cell to cell and is constantly changing through its biochemical interactions with the genome and the environment. One organism will have radically different protein expression in different parts of its body, in different stages of its life cycle, and in different environmental conditions.

The entirety of proteins in existence in an organism throughout its life cycle, or on a smaller scale the entirety of proteins found in a particular cell type under a particular type of stimulation, are referred to as the *proteome* of the organism or cell type, respectively.

With completion of a rough draft of the human genome, many researchers are now looking at how genes and proteins interact to form other proteins. The large increase in protein diversity is thought to be due to alternative splicing and posttranslational modification of proteins. This discrepancy implies that protein diversity cannot be fully characterized by gene expression analysis alone, making proteomics a useful tool for characterizing cells, tissues, and organisms of interest.

To catalog all human proteins and ascertain their functions and interactions presents a daunting challenge for scientists. An international collaboration to achieve these goals is being coordinated by the Human Proteome Organization (HUPO).

Evolution

The term evolution probably brings to mind Charles Darwin and the Theory of Natural Selection. In short, this theory states that there are more organisms brought into the environment than can be supported by the environment. Each of these individuals are different – even among the same species. The environment selects organisms best suited to survive and reproduce based on those differences. Adaptations are the differences that make one organism more suited to the environment than another individual. These adaptations are phenotypic (physical) characteristics such as finch beaks that are determined by a genetic component. The genetic component is inherited from the parent in the form of genes.

Variations in an organism's proteins may reflect physiological adaptations to an ecological niche and environment, but they originate as chance DNA mutations. Such random mutation events, if favorable, persist through the natural selection process and contribute to the evolution of new species – with new specialized functions.

The discovery of the chemical structure of DNA by Watson, Crick, Wilkins, and Franklin and our understanding of how the triplet code of nitrogen bases leads to the synthesis of proteins (which is the phenotypic expression) convinced us that adaptations are the result of changes in the DNA code (mutations). However, current research in the field of proteomics is leading some scientists to question whether or not DNA is the final determining factor in the synthesis of proteins and thus the determining factor in evolution.

The central dogma of molecular biology of DNA→RNA→protein has given us a comfortable explanation of how the information encoded by our DNA is translated and used to make an organism. It describes how a gene made of DNA is transcribed by messenger RNA and

then translated into a protein by transfer RNA in a complex series of events utilizing ribosomal RNA and amino acids. New discoveries about alternative roles for RNA, multiple forms of proteins being encoded by single genes in our cells, and changes to proteins after translation are changing this comfortable scenario and we are finding that things (as ever in biology) are not so simple. Although in essence the central dogma remains true, investigations into genomics and proteomics are revealing a complexity that we had never imagined.

In 1990, a massive research effort took place to sequence what was estimated to be the 100,000 genes that coded for each protein synthesized by humans (the human genome). This study, the Human Genome Project, took 13 years to complete. When the study began, scientists estimated that there were over 100,000 human genes. Now, years after the genome has been sequenced, there is still no consensus on the actual number of human genes, but the current estimate is down to around 22,000 human genes, this only a few thousand more genes than encodes the genome of a much simpler organism, *C. elegans*, a nematode worm that has around 19,000 genes.

So why are a similar number of genes required to make a worm and a person? Importantly, a human has a much larger total genome (3 billion base pairs) than a worm (100 million base pairs) suggesting that the total amount of DNA rather than the actual number of genes may be what gives rise to complexity. In addition, recent developments have shown it is quite common in complex organisms for a single gene to encode multiple proteins. Moreover, changing when, to what level and where a protein is expressed, or changing a protein after it has been translated (posttranslational modification) can result in proteins with very different functions. This realization of the importance and diversity of proteins started a whole new field termed **proteomics**.

Proteomics was initially defined as the effort to catalog all the proteins expressed in all cells at all stages of development. That definition has now been expanded to include the study of protein functions, protein-protein interactions, cellular locations, expression levels, and posttranslational modifications of all proteins within all cells and tissues at all stages of development. Thus, it is hypothesized that a large amount of the noncoding DNA in the human genome functions to highly regulate protein production, expression levels, posttranslational modification etc., and it is this regulation of our complex proteomes, rather than our genes, that makes us different from worms.

Research in the proteomic field has discovered a number of modification systems that allow one gene to code for many proteins and mechanisms that finely regulate the sub- and extracellular locations and expression levels of proteins. These include alternative splicing of exons, use of different promoters, posttranscriptional modification, translational frameshifting, posttranslational modification, and RNA editing.

Muscle Proteins

Our most familiar daily movements, from walking to simply breathing, are driven by the interactions between specialized proteins in our muscle fibers. The basic contractile elements of the muscle cells are the myofibrils that are bundled into muscle fibers. Each myofibril consists of a linear series of contractile units called sarcomeres.

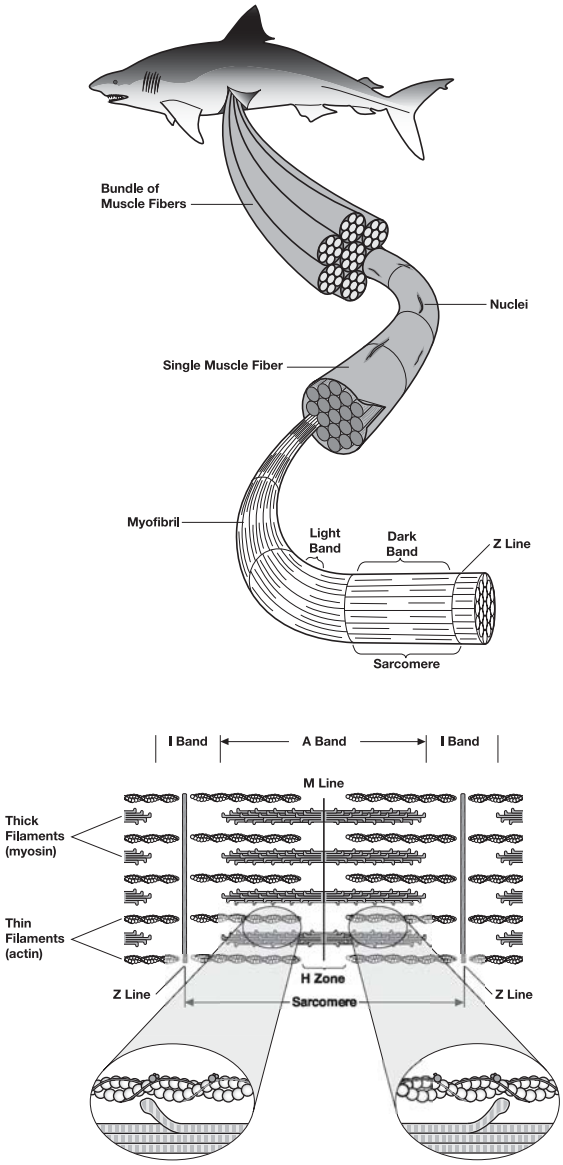


Fig. 12. Telescopic view of muscle structure: Thick myosin filaments and thin actin filaments form myofibrils, which are bundled together to make muscle fibers. (Figure modified from Campbell 1996 with permission.)

Sarcomeres are precisely arranged assemblies of actin and myosin protein filaments. Thin filaments of actin are aligned with thick filaments of myosin in a parallel and partly overlapping manner. The sarcomere shortens when myosin hydrolyzes ATP to slide along the actin filament, pulling the ends of the sarcomere towards each other. The combined contraction of many sarcomeres along a muscle fiber causes contraction of the entire muscle. It is important to note that, although actin and myosin are the major components, other proteins are also found in muscle tissue.

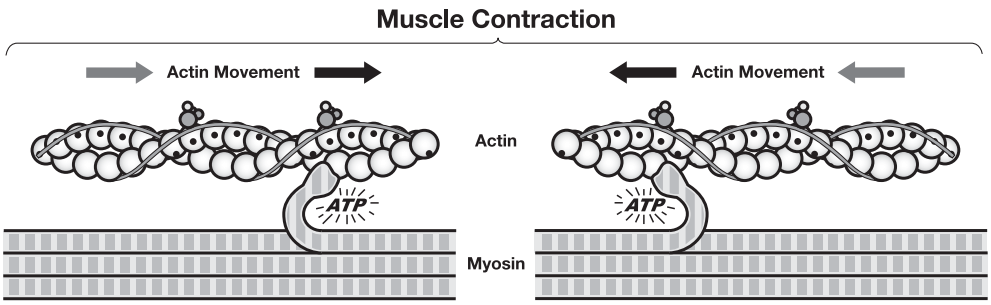


Fig. 13. Hydrolysis of ATP causes myosin and actin filaments to slide past one another, shortening the sarcomere and contracting the muscle. (Figure modified from Campbell 1996 with permission.)

Other Muscle Proteins

Numerous proteins besides actin and myosin are also required for muscle contraction (please refer to the table below). While actin and myosin are highly conserved across all animal species, other muscle proteins show more variability. These variations in an organism's muscle proteins may reflect refinements of muscle function and performance that are adaptive to particular niches, environments, or physiological stresses.

Table 2. Characterized muscle proteins, in order of decreasing size, adapted from Alberts et al. (1994).

Protein	MW (in kD)	Function
titin	3,000	centers myosin in sarcomere
dystrophin	400	anchoring to plasma membrane
filamin	270	crosslinks actin filaments into a gel
myosin heavy chain	210	slides actin filaments
spectrin	265	attaches filaments to plasma membrane
M1/M2	190	myosin breakdown product
M3	150	myosin breakdown product
C protein	140	myosin breakdown product
nebulin	107	regulates actin assembly
α-actinin	100	bundles actin filaments
gelsolin	90	fragments actin filaments
fimbrin	68	bundles actin filaments
actin	42	forms filaments
tropomyosin	35	strengthens actin filaments
troponin (T)	30	regulates contraction
myosin light chains	15–25	slide actin filaments
troponin (I)	19	regulates contraction
troponin (C)	17	regulates contraction
thymosin	5	sequesters actin monomers

Pre-Lab Activity: Using Computer Databases to Predict What Factors Affect Fish Muscle Profiles

Your teacher has assigned you 5 different species of fish. You will fill in a Fish Data Sheet for each fish. You will then use the information you have gathered to make predictions on how the muscle proteins of these fish will be similar or different.

Go to http://www.fishbase.net

You should see a screen very similar to this:

First follow the example below to become familiar with how the database works.

Example – Rainbow Trout

Type "rainbow trout" into the Common Name field on the FishBase home page.

This search term returns at least 22 different types of fish called rainbow trout. All species are the same. Click on the species *Oncorhynchus mykiss* from the USA. Use the web page this link takes you to in order to fill in the Fish Data Sheet. Note: some information, such as swim type, may need to be reached by following an additional link on the page.

Common Name: Rainbow trout

Scientific Name: *Oncorhynchus mykiss*

Taxonomic Classification:

Family: Salmonidae (Salmonids)

Order: Salmoniformes (Salmons)

Class: Actinopterygii (Ray-finned fishes)

Size: Max weight 25.4 kg

Environment: benthopelagic; anadromous freshwater; brackish; marine; depth range 0–200 m

Biology: Survive better in lakes than streams. Needs fast flowing well oxygenated waters for spawning. Can adapt to sea water if necessary.

Swim Type: Moves body and caudal fin

Definitions of unfamiliar terms: Benthopelagic – feeds on bottom, midwaters, and near surface. Hovers near bottom.

Anadromous – ascend rivers to spawn.

Now it is time to investigate your own fish species. Enter the common name of one of the species of fish you will be investigating into the FishBase search field. Then click on the scientific name link and use the information it brings up to fill in the fields of your Fish Data Sheet. If you think an additional factor may be important in predicting what proteins are expressed in fish muscles, feel free to add it to your data sheet in the additional factors field, making sure you specify that factor and fill it in for each fish species. Fill in a Fish Data Sheet for each fish species you will be investigating. If you find terms you are unfamiliar with, find out what they mean and write definitions in the field for definitions of unfamiliar terms. **Hint**: performing an Internet search with the unfamiliar term and then writing "definition" after it (for example. "benthopelagic definition") usually brings up dictionary definitions of terms.

Once your Fish Data Sheets are filled in, use the information you have gathered to answer the questions below.

Focus Questions

Which of the fields you have filled in on your Fish Data Sheets (taxonomic classification, environment, biology, swim type and additional factors) do you think will help you most in predicting which fish have the most similar muscle protein profiles?

For each field you have chosen, state why you think these fields will help you to make these predictions, and for those fields you haven't chosen, state why you don't think these will help.

Using the fields you have stated above to help you predict which fish will have similar muscle protein profiles, predict which two fish will have the most similar protein profiles. Describe how you have come to this prediction.

Similarly, predict which two fish will have the most contrasting protein profiles, i.e., the fewest proteins in common. Describe how you have come to this prediction.

After the lab is complete, remember to come back and see whether your predictions were correct.

Fish Data Sheet

Common Name:

Scientific Name:

Taxonomic Classification:

- Family
- Order
- Class

Size:

Environment:

Biology:

Swim Type:

Additional Factors:

Definitions of unfamiliar terms:

Alternative Pre-Lab Activity

Evolution and Classification of Fish

An evolutionary tree shows the evolutionary lineages of different species over relative time. Evolutionary trees, (also called cladograms), can be based on many different types of data. Some trees are constructed using a single type of data and some trees use multiple types of data. The traditional way of constructing evolutionary trees was to look at the physical morphology of organisms, including sizes, shapes and developmental structures of both living organisms and fossils. Today, similarities and differences in protein and DNA sequences are being used. Both methods are valuable and often complement each other but they may not always agree. Can you propose why this may be so?

Use the evolutionary tree below to make predictions about the relatedness of the fish species you will examine in this lab. Following the analysis and interpretation of your electrophoresis results, you will create a cladogram from your own results and compare your cladogram with your predictions. Will your lab data support or refute your own predictions? Why, or why not?

For this activity, and in order to generate meaningful data, it is useful to compare both closely related and distantly related fish. In addition to using the tree below to make your predictions, we recommend that you research additional information on the evolutionary histories of fishes, using the Internet, and biology and zoology books.

The data used to construct the evolutionary tree below was obtained from the cladograms on the tree of life web page from the University of Arizona (www.tolweb.org). (Please note that the field of phylogenetics is ever changing and different methods used to construct a phylogenetic tree often result in differences between trees, hence the data on the tree of life web page may not concur exactly with "textbook" evolutionary trees.)

Most fish are contained within the superclass Gnathostoma (jawed vertebrates), which also includes all tetrapods. Only hagfish and lampreys are outside this group. These two fish types are sometimes classed together as Agnatha, but can also be separated into Hyperotreti and Hyperoartia. Hyperotreti (hagfish) are craniates (animals with skulls), but not vertebrates because they have no backbone, while Hyperoartia (lamprey) are very primitive vertebrates, but do not have a jaw. The Gnathostoma fishes are divided into the classes Chondrichthyes (cartilaginous fishes) and Osteichthyes (bony fishes). The Chondrichthyes include the sharks and rays, and the Osteichthyes, include all other modern fishes and all tetrapods (amphibians, birds, and mammals). Below are brief descriptions of some of the major fish groups, in order from most ancient to most recently diverged.

Hyperotreti (e.g., hagfish) are eel-like, jawless fishes that have a skull, but no backbone with parasitic and scavenging lifestyles. They are very primitive and may approach the condition of the common ancestor to all craniates.

Hyperoartia (e.g., lamprey) are eel-like, jawless fishes that are primitive vertebrates. They are identified by a single nostril and a sucker-like mouth with which they attach to fishes and rocks.

Chondrichthyes (e.g., shark, ray, skate, & sawfish) have a cartilaginous rather than bony skeletons that reflects a more evolutionarily ancestral state. Their skin is thick and without true scales, and they do not have swim bladders or lungs.

Osteichthyes (e.g., coelancanth, tuna & haddock). The bony fishes are the most diverse class of fish. The class is characterized by having bony skeletons, true scales, paired fins, and movable rays in their fins and tail. Osteichthyes are divided into two subclasses:

- The lobe-finned fish, Sarcopterygians which contains the living fossil, the coelacanth, and the tetrapods (amphibians, reptiles, mammals, birds and dinosaurs)
- The ray-finned fish, Actinopterygians, which contains most other fish

Sarcopterygians (e.g., lungfish and coelacanth) also include modern amphibians, reptiles, birds, and mammals. Coelacanth were thought to have become extinct at about the same time as the dinosaurs, until a live specimen was found in 1938. They form an important evolutionary link between fish and four-legged land animals.

Actinopterygian (e.g., gar, sturgeon, mackerel & anglerfish) is the subclass encompassing most modern ray-finned fish including the chondrostei, semionotiformes, and teleosts.

- **Chondrostei (e.g., sturgeon)** are considered relic bony fishes. They lack scales on most of the body, have a cartilaginous skeleton, and have developed a shark-like, heterocercal tail and a rostrum extending past the mouth.
- **Semionotiformes (e.g., gar)** are also ancient fish, they have bony scales and a mainly cartilaginous skeleton.

Teleosts (e.g., herring, carp & pufferfish) comprise the remainder of the bony fishes. These include Clupeomorpha (e.g., herring, sardine & anchovy), Ostariophysi (e.g., carp, catfish, minnow, piranha & electric eel), Salmoniformes (e.g., salmon, trout & smelt), Esociformes (e.g., pike), and the diverse group, Acanthomorphia (e.g., tuna, cod & pufferfish).

Acanthomorphia (e.g. pollock, bass & sole) comprises two main superorders, Paracanthopterygians (e.g., cod, pollock & anglerfish) and Acanthopterygians. The Acanthopterygians include the Perciformes (e.g., the scombridae (e.g., swordfish, mackerel & tuna) and the serranidae (e.g., bass, snapper & grouper)), the Pleuronectiformes (e.g., flat fish, flounders & sole) and the Tetradontiformes (e.g., pufferfish).

Pre-Lab Focus Questions

1. Name three proteins found in muscle

   i. _____

   ii. _____

   iii. _____

2. Make a prediction of which two of your fish species will have the most similar protein profiles and why. You may want to use a database on fish (http://www.fishbase.net) to find more information on the different fish you are analyzing.

3. Give a prediction of which two of your fish species will have the least similar protein profiles and why.

4. Predict how the other fish in the study will compare by drawing an evolutionary tree and describing your reasoning.

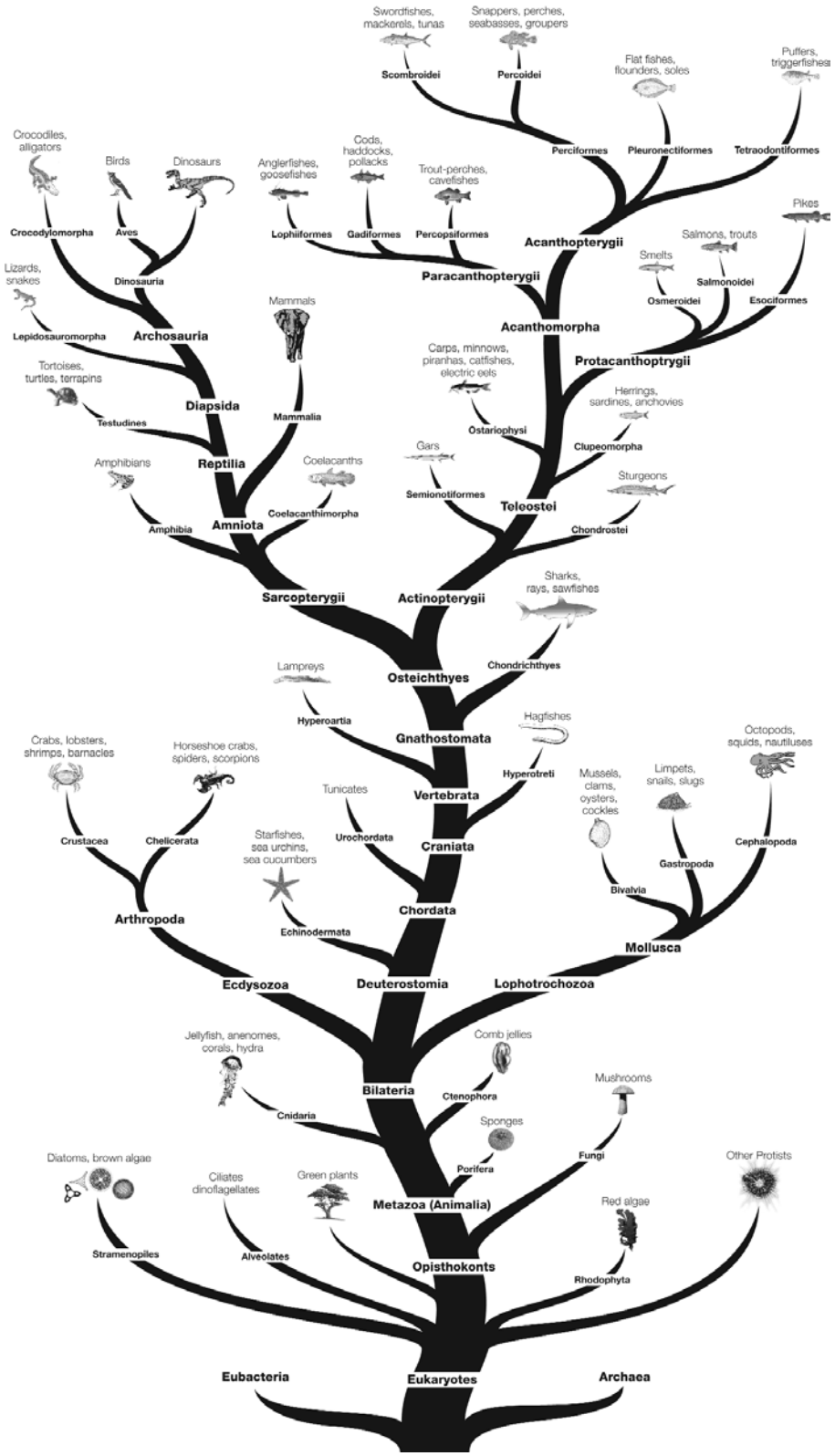


Fig. 14. Evolutionary tree showing the relationships of eukaryotes. (Figure adapted from the tree of life web page from the University of Arizona (www.tolweb.org).)

Lesson 1: Introduction to Protein Electrophoresis and SDS-PAGE

How Can We Study Proteins Found in Muscle?

Polyacrylamide gel electrophoresis (PAGE) can be used to separate small molecules such as proteins. Understanding protein structure is important to understanding how we can use PAGE for protein analysis.

Proteins are made of smaller units (monomers) called amino acids. There are 20 common amino acids. The sequence and interaction between these different amino acids determine the function of the protein they form. Amino acids are joined together by peptide bonds to form polypeptide chains. Chains of amino acids constitute a protein. In turn these chains may interact with other polypeptides to form multi-subunit proteins.

Amino acids can be combined in many different sequences. The sequence of the amino acids in the chain is referred to as the primary protein structure. All amino acids have the same basic structural component (Figure 15).

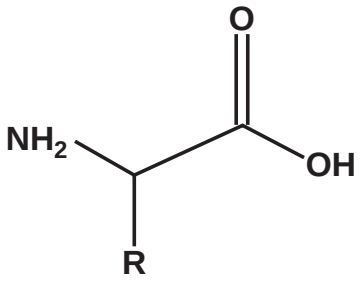


Fig. 15. Chemical structure of an amino acid.

The "R" group may be charged or uncharged, or may be a long side chain. Thus, each amino acid has different properties and can interact with other amino acids in the chain. Hydrogen bonding between these side chains, primarily between the C=O and the N-H groups, causes the protein to bend and fold to form helices, pleated sheets, reverse turns, and non-ordered arrangements. Disulfide bonds between methionines can also bend and loop the amino acid chain. This is considered the secondary structure of the protein.

The tertiary structure of the protein is determined by the interaction of the hydrophilic and hydrophobic side chains with the aqueous environment. The hydrophobic regions aggregate to the center of the molecule. The hydrophilic regions orient themselves toward the exterior. These ordered bends and folds make the protein compact. Examples of tertiary protein structure are structural and globular proteins.

The quaternary structure of proteins is achieved from the interaction of polypeptide chains with others. Multiple polypeptides can combine to form complex structures such as the muscle protein, myosin, or the blood protein, hemoglobin, which are both composed of four polypeptide chains. These complex proteins are often held together by disulfide bonds between methionines. In fact, PAGE analysis was first carried out in 1956 to show the genetic disease sickle cell anemia is caused by a change to a single amino acid of the hemoglobin protein (Ingram 1956).

Prior to electrophoresis, the proteins are treated with the detergent sodium dodecyl sulfate (SDS) and heated. SDS and heat denatures (destroys) the protein tertiary and quaternary structures, so that the proteins become less three dimensional and more linear. SDS also gives the protein an overall negative charge with a strength that is relative to the length of its polypeptide chain, allowing the mixture of proteins to be separated according to size.

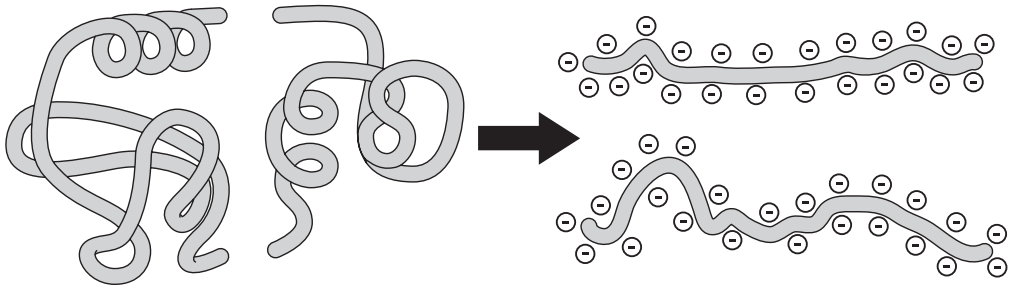


Fig. 16. The combination of heat and the detergent SDS denatures proteins for SDS-PAGE analysis.

The proteins, in their SDS-containing Laemmli sample buffer, are separated on a gel with a matrix that acts to sieve the proteins by size upon addition of an electric current. A polyacrylamide gel is positioned in a buffer-filled chamber between two electrodes, protein samples are placed in wells at the top of the gel, and the electrodes are connected to a power supply that generates a voltage gradient across the gel. The SDS-coated, negatively charged proteins migrate through the gel away from the negatively charged anode toward the cathode, with the larger proteins moving more slowly than the smaller proteins. This technique was developed by U.K. Laemmli, whose 1970 Nature paper has received the highest number of citations of any scientific paper. SDS-PAGE is still the predominant method used in vertical gel electrophoresis of proteins.

As soon as the electric current is applied, the SDS-coated proteins begin their race toward the positive electrode. The smaller proteins can move through the gel more quickly than the larger ones, so over time, the proteins will be separated according to their sizes.

Protein size is measured in **daltons**, a measure of molecular mass. One dalton is defined as the mass of a hydrogen atom, which is 1.66×10^{-24} gram. Most proteins have masses on the order of thousands of daltons, so the term **kilodalton** (kD) is often used to describe protein molecular weight. Given that the average weight of an amino acid is 110 daltons, the number of amino acids in a protein can be approximated from its molecular weight.

- Average amino acid = 110 daltons
- Approximate molecular weight of protein = number of amino acids x 110 daltons

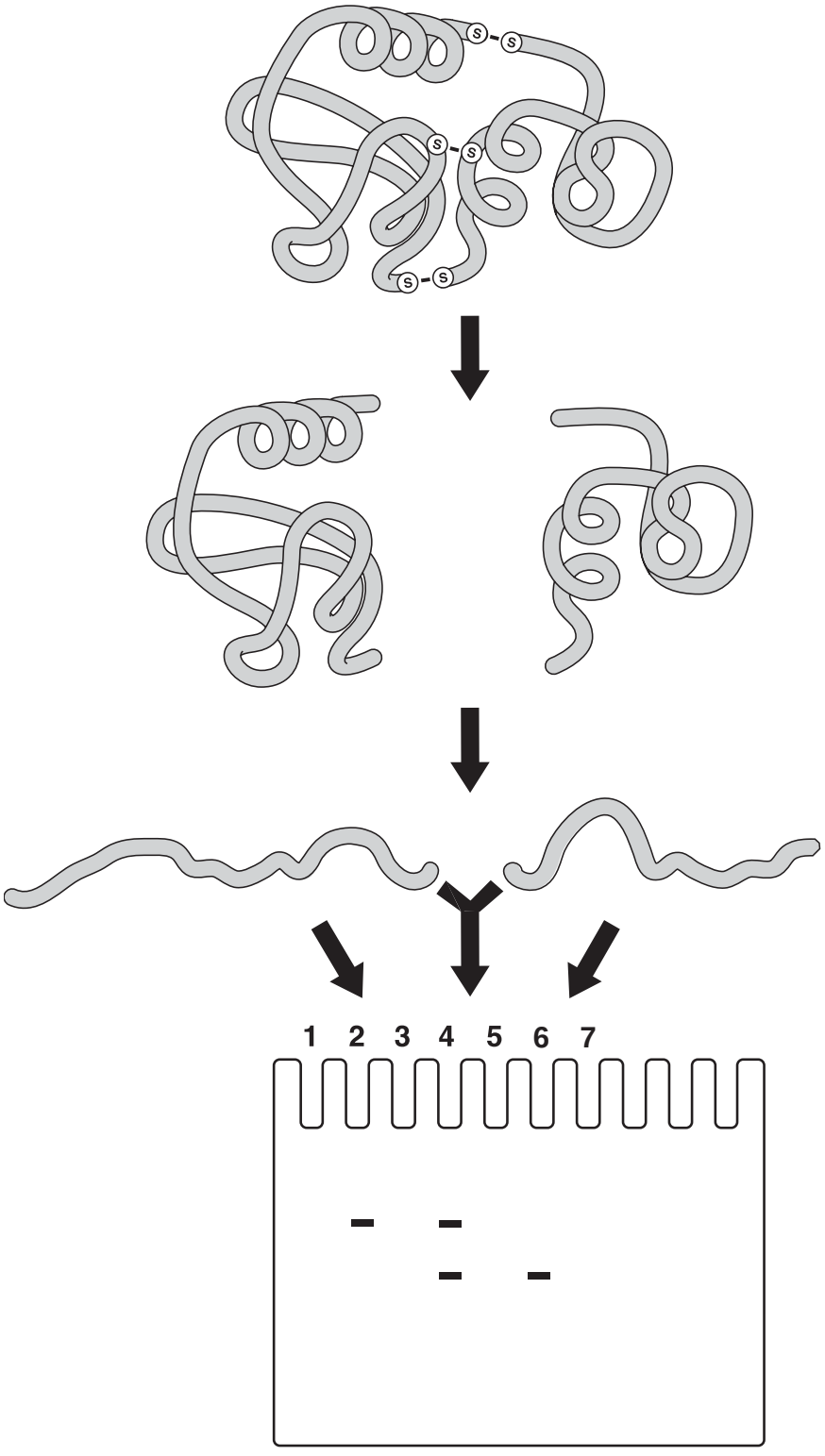


Fig. 17. A quaternary protein complex denatured with reducing agents, heat, and SDS, can be separated into individual proteins and resolved by size using SDS-PAGE.

In this investigation, you will use SDS-PAGE to separate and analyze the protein profiles of the muscle tissue of different fish. By comparing the protein profiles of different fish species you can test the hypothesis that protein profiles are indicators of genetic and evolutionary relatedness.

Visualizing your proteins

Proteins in your samples are not visible while the gel is running. The only visible proteins will be those in the Precision Plus Protein Kaleidoscope standard that have been prestained with covalently attached dyes. You should see these proteins resolve into multicolored bands that move down the gel as power is run through the gel. If the electric current is left on for too long, the proteins will run off the bottom of the gel. To guard against this and to show you the progress of your gel if you did not have the standards, a blue tracking dye is mixed with the Laemmli sample buffer used to prepare your protein samples. This blue dye is negatively charged and is also drawn toward the positive electrode. Since the dye molecules are smaller than the proteins expected in most samples, they move ahead of the proteins in the gel.

After turning off the electric current and removing the gel from the glass plates that hold it in place, the gel is placed in a stain. The stain used in this technique was originally developed to dye wool, which, like your own hair, is composed of protein. This stain binds specifically to proteins and not to other macromolecules such as DNA or lipids. After destaining, distinct blue bands appear on the gel, each band representing on the order of 10^{12} molecules of a particular protein that have migrated to that position: the larger the amount of protein, the more intense the blue staining.

Lesson 1: Protein Extraction From Muscle

Your first task is to extract proteins from muscle tissue, unfold and denature them, and give each protein an overall negative charge using Laemmli sample buffer, mechanical forces, and heat. In this lab you will add tiny pieces of muscle to Laemmli sample buffer and manually disrupt the tissue by flicking the tubes. This will release muscle specific proteins from the cells, unfold them, and add an overall negative charge to each protein. You will then pour off the extract and heat the extracted proteins to 95°C, which will complete their denaturation.

Student Workstations

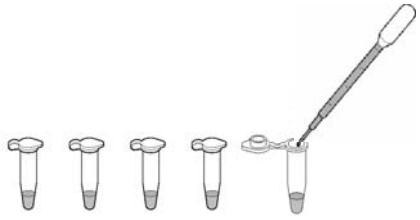
Material	Quantity
1.5 ml fliptop microtubes	5
1.5 ml screwcap microtubes	5
1 ml transfer pipet	1
Fish samples, labeled 1–5	5 species
Marking pen	1
Laemmli sample buffer	1.5 ml
Knife or scissors to cut fish samples	1

Common Workstation

Material	Quantity
Water bath set to 95°C	1

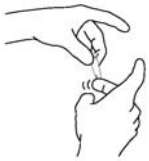
Procedure

1. To make this a blind study, assign a letter (e.g., A–E) to each fish sample to be investigated. Keep a record of which fish got which number and hide their true identities until after the analysis is complete.
2. Label 1.5 ml fliptop microtubes with the number of the fish species to be analyzed. There should be one labeled tube for each fish sample being prepared for electrophoresis.
3. Add 250 µl of Laemmli sample buffer to each labeled tube.

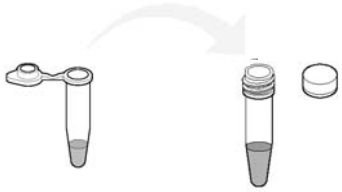


4. For each sample, obtain a piece of fish muscle (avoid skin, fat, and bones) approximately 0.25 x 0.25 x 0.25 cm^3 (), and transfer it to the appropriately labeled microtube. Close the lid.

5. Gently flick the microtube 15 times with your finger to agitate the tissue in the sample buffer.



6. Incubate the samples for 5 minutes at room temperature to extract and solubilize the proteins.
7. Pour the buffer containing the extracted proteins, but not the solid fish piece, to a labeled 1.5 ml **screwcap** tube. Note: It's not necessary to transfer all of the fluid to the screwcap tube, since only a small volume (<20 µl) is actually needed for gel loading.



8. Heat your fish samples **and** the actin & myosin (AM) standard in their screwcap tubes for 5 minutes at 95°C to denature the proteins in preparation for electrophoresis.
9. Store the samples at room temperature if they are to be loaded onto gels within 3–4 hours, or store them at –20°C for up to several weeks.

Lesson 1 Focus Questions

1. Why did you add Laemmli sample buffer to your fish samples?

2. What was the purpose of heating the samples?

3. How are the proteins extracted from the fish samples?

4. Have all the proteins been extracted from the fish slice or are some still left after the extraction? How could you test your hypothesis?

Lesson 2: Electrophoresis: Gel Loading, Running, and Staining

So far you have extracted, denatured, and given the proteins from fish muscle tissue a negative charge. Now they can be separated according to their molecular weights using gel electrophoresis, which will generate profiles for various fish species

Student Workstations

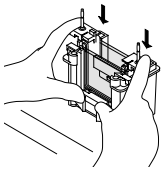
Material	Quantity
Fish protein extracts from lab one	5 species
Actin & myosin standard, 12.5 ul	1 vial
Precision Plus Protein Kaleidoscope prestained standards, 6 µl	1 vial
15%, 10-well, Ready Gel precast gel	1
1–20 µl adjustable-volume micropipet	1
Prot/Elec pipet tips for gel loading	7 tips
Mini-PROTEAN 3 electrophoresis module (gel box)	1 per 2 gels
1x Tris-glycine-SDS (TGS) running buffer	500 ml per gel box
Power supply (200 V constant) to be shared between workstations	1
Sample Loading Guide – for 10 well comb	1 per gel box
Buffer dam (only required if running 1 gel/box)	1
Staining trays	1 per 2 gels
Bio-Safe Coomassie stain for proteins	50 ml per 2 gels

Common Workstation

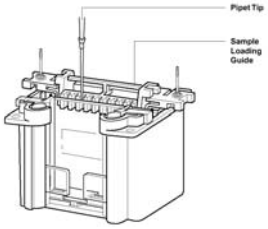
Material	Quantity
Water bath set at 80–95°C	1
Water for gel destaining (tap water is fine)	

Procedure

1. Reheat frozen fish samples and actin and myosin standard at 95°C for 2–5 minutes to redissolve any precipitated detergent. Note: If you have prepared your fish samples in this lesson, there is no need to reheat them.
2. Assemble gel boxes. Use the pictorial guide found in the Quick Guide to insert your Ready Gel polyacrylamide gels into the vertical electrophoresis module if your instructor has not preassembled them.
3. Double-check that the buffer in the inner buffer chamber is well above the top of the smaller plate. If it is not, you may have a leak; consult with your instructor.
   Note: If you do have a leak, the outer chamber of the gel box can be filled to above the inner small plates, to equalize the buffer levels in both reservoirs.



4. If available, place a yellow sample loading guide on the top of the electrode assembly. The guide will direct your pipet tip to the correct position for loading each sample in a well.

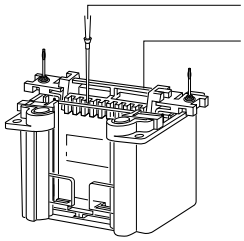


5. Record in which well of your gel you will load which of your samples in the table below:

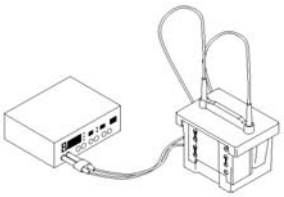
Well	Volume	Sample Name
1	empty	none
2	empty	none
3	5 µl Stds	Precision Plus Protein Kaleidoscope prestained standard (Stds)
4	*10 µl sample A	_____
5	*10 µl sample B	_____
6	*10 µl sample C	_____
7	*10 µl sample D	_____
8	*10 µl sample E	_____
9	*10 µl AM	actin & myosin standard (AM)
10	empty	none

*Note: If you are going on to perform the western blot module load 5 µl of each fish sample and the actin & myosin standard. This will prevent overloading the lanes.

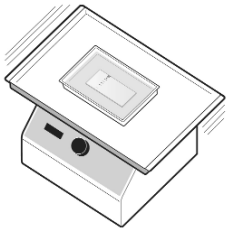
6. Load **5 µl** of Precision Plus Protein Kaleidoscope Prestained standard gently into well # 3 using a thin gel loading tip. **Note**: The fine barrel of the gel loading tips means liquid is slower to go into the tip than normal tips. You must therefore release the plunger of the micropipet very slowly, otherwise you will not pipet the correct volume.



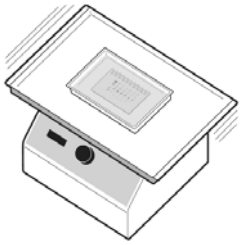
7. Using a fresh tip each time, load **10 µl** of each of your protein samples gently into the wells designated in your table above. Note: If you are going on to perform the western blot module, load 5 µl of each protein sample.
8. Using a fresh tip, load **10 µl** of the actin & myosin standard gently into well # 9. Note: If you are going to perform the western blot module load 5 µl of the actin & myosin standard.
9. After loading all samples, remove the yellow sample loading guide (if used), place the lid on the tank, and insert the leads into the power supply, matching red to red and black to black. Set the voltage to 200 V and run the gels for 30 minutes. Watch for the separation of the standard.



10. When gels are finished running, turn off the power supply and disconnect the leads. Remove the lid and lift out the electrode assembly and clamping frame.
11. Pour out the running buffer from the electrode assembly. Open the cams and remove the gel cassettes.
12. Now it's time to stain the proteins in your gel. Lay your gel cassette flat on the bench with the short plate facing up. Cut the tape along the sides of the gel cassette. Carefully pry apart the gel plates, using your fingertips. The gel will usually adhere to one of the plates. Transfer the plate with the gel adhering to it to a tray containing tap water allowing the liquid to detach the gel from the plate. The gel may also be lifted directly (and gently!) from the plate and placed into the water. If there is sufficient time, rinse the gel 3 times with tap water for 5 minutes by carefully pouring out the water and replacing it.
13. Carefully pour out the water and replace with 50 ml of Bio-Safe Coomassie stain per 2 gels.



14. Allow the gels to stain for 1 hour, with shaking if available.
15. After 1 hour discard the stain and replace it with a large volume of water to destain the gel overnight with rocking action if available.



Lesson 2 Focus Questions

1. Why do SDS-coated proteins move when placed in an electric field?

2. What is the purpose of the actin & myosin standards and the Precision Plus Protein Kaleidoscope prestained standard?

3. Which proteins will migrate farthest? Why?

4. What is the purpose of the stain?

Lesson 3: Destain and Dry Gels

You will now destain your gel and dry it to make a permanent record.

Student Workstations

Material	Quantity
Container of tap water to wet cellophane	1
GelAir cellophane support sheets	2
Square plastic container and 2 rubber bands (if not using GelAir drying frame)	1
Ruler & graph paper if performing band analysis	

Common Workstation

Material	Quantity
GelAir drying frames	1
GelAir assembly table (optional)	1
GelAir dryer (optional)	1

- Examine your gel. Blue protein bands should be visible on the clear, destained gel. You may want to make a photocopy of your gel so that you can perform the detailed analysis sooner. Your instructor will tell you if you will be drying your gel with GelAir drying frames, or by the plastic container method. Follow the instructions below accordingly.
- Answer the focus questions on qualitative comparisons of protein profiles before drying your gels.

GelAir drying frame method:

1. Precast polyacrylamide gels must have the ridge at the bottom of the gel removed by chopping them off (not slicing) using a plastic card, e.g., an I.D. card or ruler.
2. Prewet 2 sheets of cellophane in a container of water for 15–20 seconds.
3. Place a plastic drying frame on the GelAir assembly table. Center one sheet of cellophane on the assembly table.
4. Carefully lay your gel on the cellophane, positioning it to accommodate other gels (up to 6 total). If there are bubbles between the gel and the cellophane, gently push them out with your gloved finger.
5. Flood the gels with water and lay the second sheet of cellophane on top of them, trying not to trap any bubbles in the sandwich. If there are bubbles, gently push them out with a gloved finger. Bubbles will cause cracks in the gel during drying!
6. Place the square metal frame on top of the cellophane sandwich. Secure the eight clamps onto the frame, two on each side. If you are not using a GelAir Dryer oven, place the frames upright in a well-ventilated area for 12–36 hours. If you have a GelAir dryer, place up to four drying frames into the oven, turn the heater switch on, and set the dial to 3 hours. The dryer will shut off automatically.

7. When the gels are completely dry, they will be flat. Remove the clamps and take the gel/cellophane sandwich from the frame. Trim the excess cellophane surrounding your dried gel with scissors.
8. In contrast to Fast Blast™ Stain for DNA staining, Bio-Safe Coomassie protein stain is not light sensitive.

Alternatively, you may use the GelAir cellophane sandwich and plastic container method:

1. Wet two pieces of cellophane in a large volume of water, around 500 ml.
2. Place one sheet of cellophane over a plastic container. Pull the cellophane taut so that it makes a flat surface over the top of the container, and use a rubber band to secure the sheet in place.
3. Place a gel onto the cellophane. Remove any air bubbles that are under or around the gel. Flooding the surface of the cellophane around the gel with water will aid in the removal of bubbles.
4. Place the second sheet of wetted cellophane over the gel, being careful not to trap any bubbles.
5. Secure the second sheet of cellophane to the box with a second rubber band. Allow gel to dry for several days in a well-ventilated area. Cutting the bottom out of the plastic container can speed up drying.
6. In contrast to Fast Blast™ Stain for DNA staining, Bio-Safe Coomassie protein stain is not light sensitive.

Lesson 3 Focus Questions

1. Which two fish have the most similar protein profiles?

2. Which two fish have the least similar protein profiles?

3. Give an explanation for why you think the protein profiles of some fish species share more bands than other fish species.

4. Did your predictions from your Pre-Lab Activity turn out to be true or not? If not, why do you think that was?

Post-Lab Activity

Analysis and Interpretation of Results

Detailed Gel Analysis

Does your molecular evidence support or refute your predictions?

Each protein band that a fish has in common with another fish is considered a shared characteristic. A fish family tree, or cladogram, can be constructed based on proteins bands that the fish have in common. Cladistic analysis assumes that when two organisms share a common characteristic that they also share a common ancestor with that same characteristic.

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 molecular weights

(Optional) Although it is not strictly necessary for this exercise, you may want to create a standard curve from your gel and determine the actual size of each protein band. Alternatively, the cladogram can be generated just using the distance in millimeters the different protein bands have migrated from the wells of the gel.

To create the standard curve measure and record the distances the five visible protein bands contained in the Precision Plus Protein Kaleidoscope prestained standards. Start from the green 37 kD band down to the yellow 10 kD band that has migrated from the wells. Accuracy to 0.5 mm is required.

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

On the graph paper provided below, plot the distances migrated in mm on the x-axis against the molecular weight of the bands in kD on the y-axis as a scatter plot.

Draw a line through the points. On a logarithmic scale, plotting the molecular weights against the distances migrated for each protein in the standard should result in a linear (straight line) curve. Alternatively, you can use a graphing computer program to generate the chart and make a line of best fit (or a trend-line) through these points and to formulate an equation to calculate the MW of the unknown proteins on the gel.

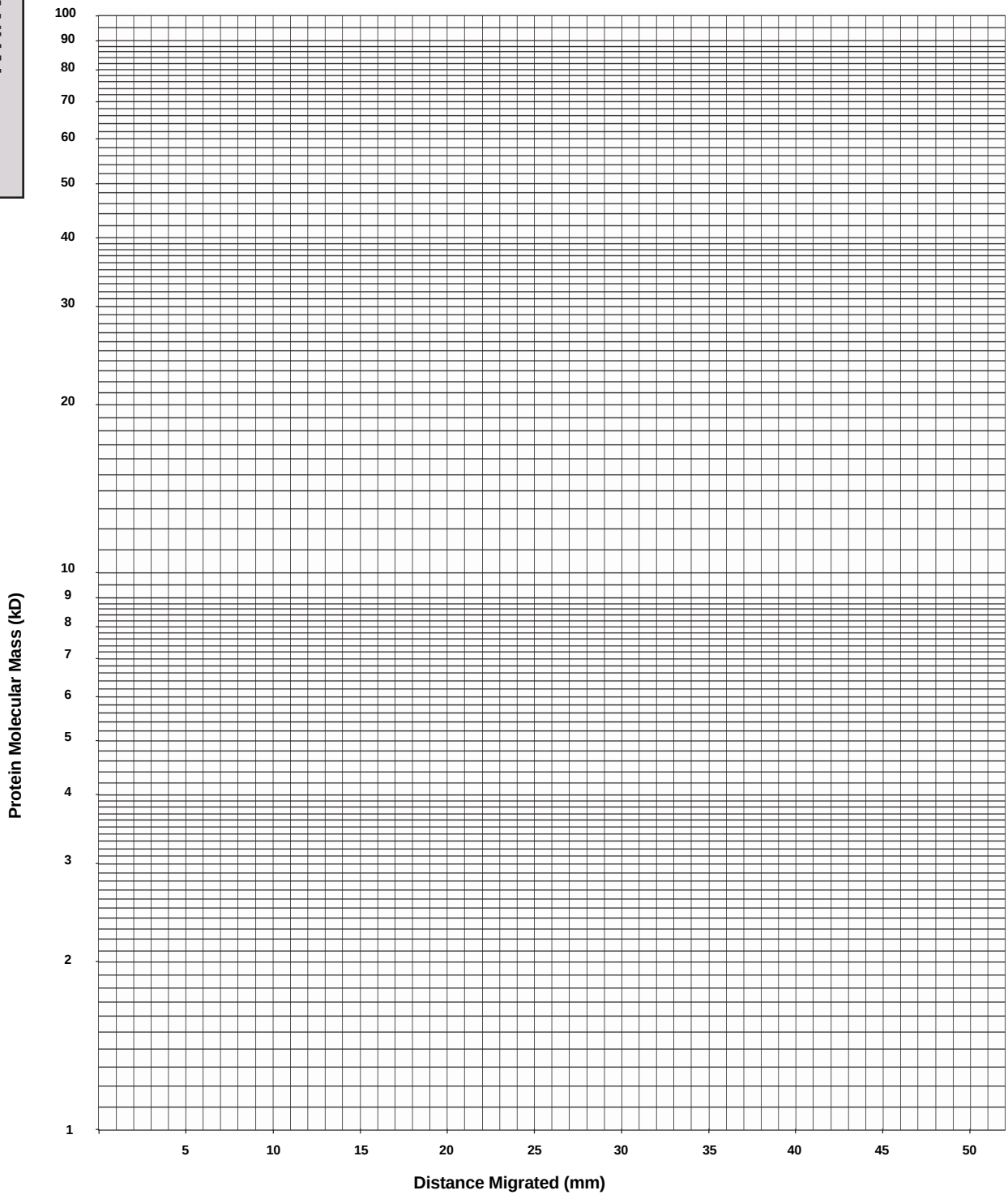


Fig. 16. Two-cycle semi-logarithmic graph paper to construct curve of the protein molecular mass against the distance migrated.

Define the characteristics (proteins) of the different fish

Make a horizontal line on the dried gel (or gel image) between the 37 kD (green) and 25 kD (pink) markers below the fat bands that occur at around 30 kD (see gel above). Then, for each band below the line for each fish sample, measure the distance the protein band has migrated from the wells (and, if required, determine its size in kD using the standard curve or the formula generated from the standard curve) and record this data (see example below):

	Species A	**Species B**	**Species C**	**Species D**	**Species E**
Distance protein bands migrated in mm	25, 26.5, 29, 36, 36.5, 39, 44, 52	26, 27.5, 29, 32, 34.5, 36.5, 37.5, 40.5, 42, 45	26, 27.5, 29, 29.5, 32, 34.5, 36.5, 37.5, 40.5, 42, 45, 46.5, 51.5	26, 27.5, 29, 32, 36.5, 38, 38.5, 41, 46, 47.5 44, 47	26, 27.5, 30, 30.5, 33, 35.5, 37, 39, 39.5, 42,

Determine which fish have each characteristic (protein)

Make a table with a row for every band size you have recorded 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**

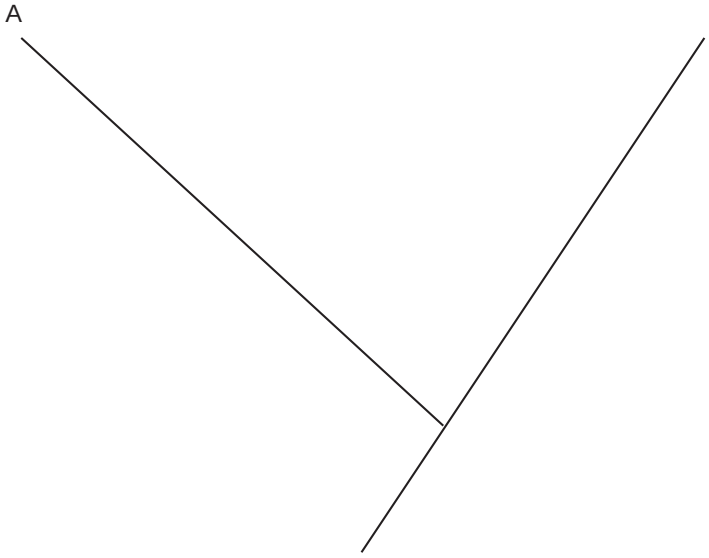
Find the number of characteristics shared by each of the fish

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. The table below will be the basis for drawing your cladogram.

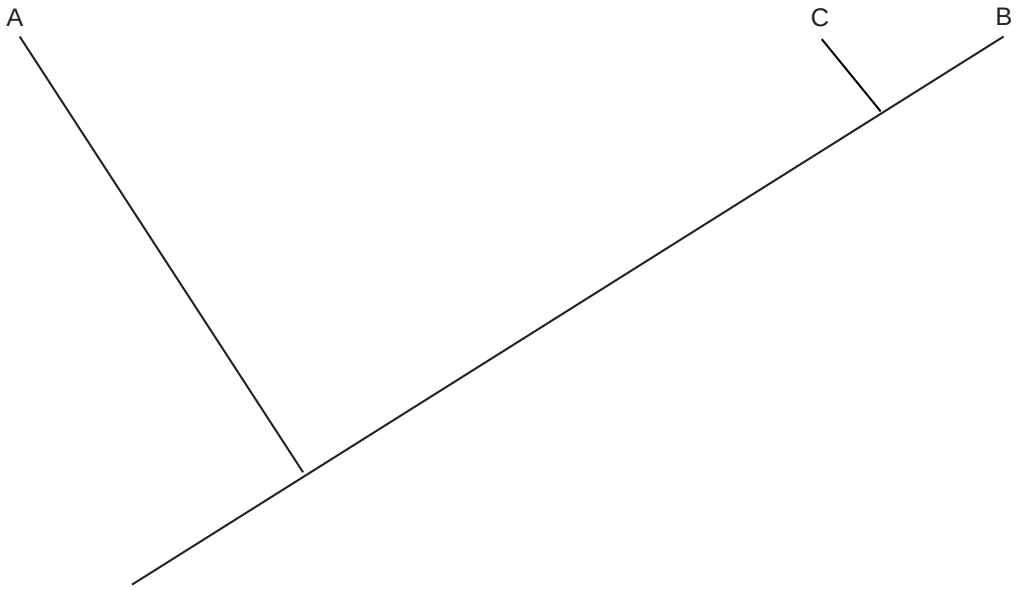
	Species A	**Species B**	**Species C**	**Species D**	**Species E**
Species A	8	2	2	2	2
Species B		10	10	5	3
Species C			13	5	4
Species D				10	2
Species E					12

Construct your cladogram

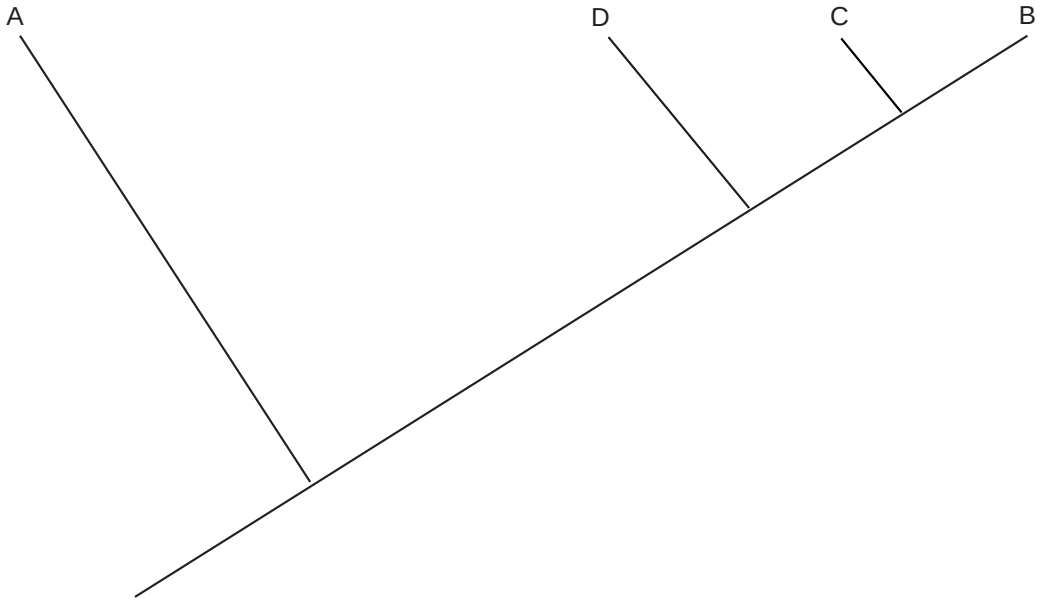
Now you are ready to construct your cladogram. First draw a line to form the trunk of your cladogram. Find the fish with the least bands in common. In the example above it is species A, which has only 2 bands in common with any of the other fish. Then draw a side branch off the line near the bottom of 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.

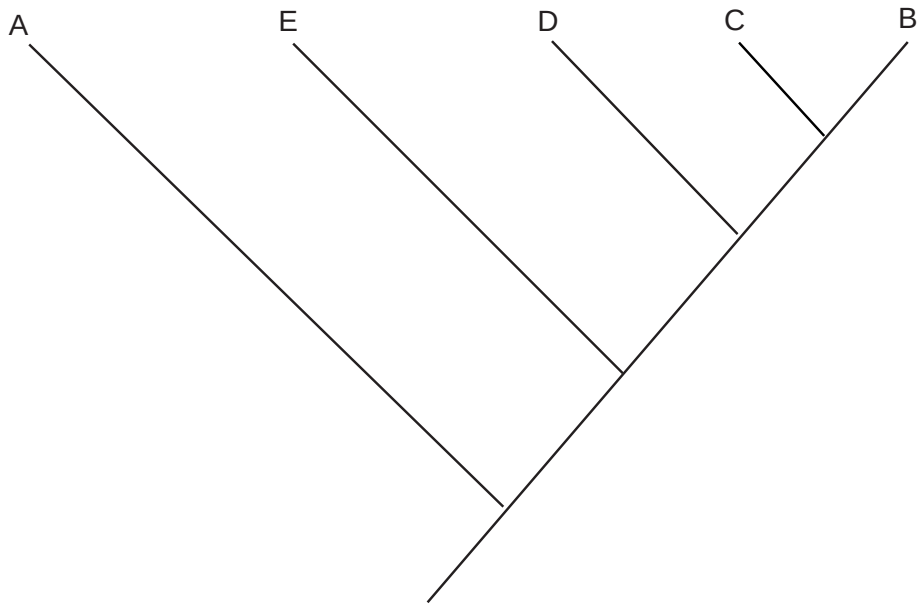


The last fish to add to the cladogram in this example is species E, which shares four bands with species C, three bands with species B, and only two bands with species A and D. This fish may seem trickier to place than the others because it shares more characteristics with species B and C than it does 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 A: Protein Electrophoresis Using Agarose Gels and Horizontal Gel Electrophoresis Apparatus

The principles of protein separation by electrophoresis can be taught using agarose gels and horizontal electrophoresis apparatus that is typically used for separating DNA. Proteins are rarely separated on agarose in the real world because proteins are much smaller than DNA (compare the number of carbons in a base pair to an amino acid). Polyacrylamide has a tighter matrix than agarose which, in addition to the discontinuous system used in SDS-PAGE (see Lesson 1: Introduction to Protein Electrophoresis SDS-PAGE), is much better at resolving most proteins into distinct bands. However, agarose gels are used in scientific laboratories to separate large proteins, for example serum proteins from blood. The following protocol describes a method to separate fish muscle proteins using agarose gels and horizontal gel electrophoresis apparatus. The same experimental principles apply to this method as SDS-PAGE, such as the SDS conveying a net negative charge and the denaturation of proteins using heat. The main difference is that the gel is not a discontinuous system like the polyacrylamide gels.

An example of the results is shown in Figure 18 below. You can see that the bands are not as distinct as with polyacrylamide gels, and that the larger proteins separate better with more distinct bands than the smaller ones, which diffuse in the more porous agarose.

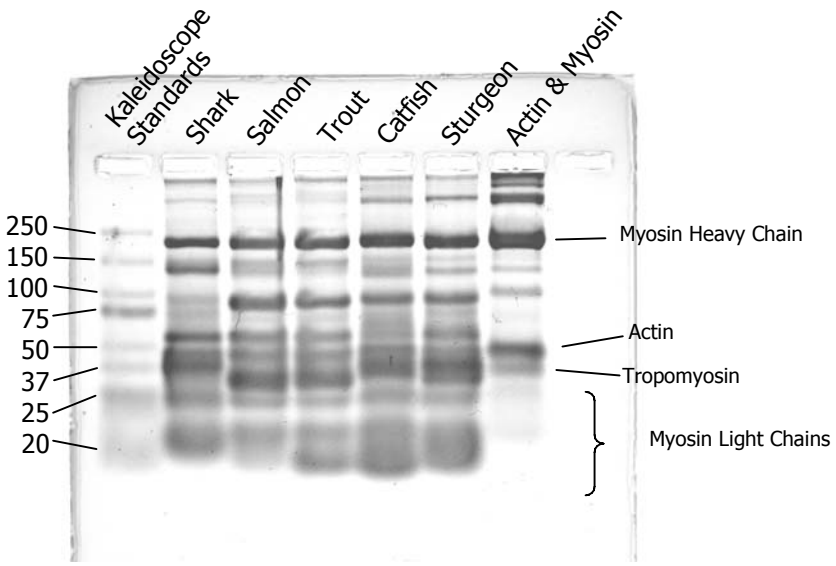


Fig. 18. 4% agarose gel electrophoresed at 100 V for 45 minutes, fixed, stained with Bio-Safe Coomassie stain, and destained in water.

It is important to note that both the type of agarose and the running buffer are different from what is usually used to separate DNA. **Regular agarose will not work with this protocol**. The agarose used to make these gels is high-percentage PCR low-melt agarose and must be purchased separately from the kit (Certified PCR low-melt agarose, 25 g, catalog #161-3113). This special type of agarose is superior to regular molecular biology agarose at separating very small molecules. Low-melt agarose is different from regular agarose because once dissolved and formed into gels, it melts at around 65°C and will remain molten to around 34°C. Therefore, if the gels are prepared in a hot climate, it may be necessary to put the gels in the refrigerator to set. **Tip**: Thinner agarose gels (e.g., 5 mm) will resolve proteins better than thick gels, so pour your gels carefully with minimal agarose. The running buffer and the buffer

used to dissolve the agarose is 1x Tris-glycine-SDS (TGS) buffer. This is the same running buffer used for polyacrylamide gels.

It is important to prepare the agarose gels by the method described below using a hot water bath, and not in the microwave or on a hot plate as is often done when preparing DNA agarose gels. SDS is a detergent, and boiling the TGS-agarose in a microwave or hot plate will result in a boiling foam that will not form good gels and is a safety hazard.

The protocol for performing protein electrophoresis using agarose gels differs from that used for polyacrylamide gels described in lesson 2 by a few points: extra electrophoresis time; a fixation step, a longer staining step and extra destaining. Note: an extra day may be required to complete the lab when using this protocol.

- The gels are electrophoresed for 45 minutes at 100 V. The extra electrophoresis time is necessary to see good separation of protein bands. Note: Increasing the voltage to reduce the running time is not recommended since this may increase the temperature of your gel running buffer, which could melt the low-melt agarose gels.
- After the gels have been run, the proteins need to be fixed into the agarose, otherwise they will diffuse and the protein bands will become very fuzzy. The fixative is 10% acetic acid and 40% ethanol. The fixation causes the proteins to aggregate and partially precipitate. This makes them larger and so they diffuse less in the porous agarose matrix. Vinegar can replace acetic acid, but will not work quite as well. Gels should be fixed for 1 hour to overnight.
- The gels are stained for 2 hours to overnight, this extra time is required because the gels are much thicker than polyacrylamide gels.
- The water used to destain the gels should be changed at least 3 times to remove the stain from the agarose.

Post-lab activities: The activities to determine protein sizes and the relatedness of the fish species described in the Post-Lab Activity can be performed from agarose gel data. However, the analysis should concentrate on the larger proteins rather than the smaller proteins. The 250 kD to 25 kD molecular mass markers should be used to generate a standard curve, rather than the 37 to 10 kD. Your students may have difficulty deciding which point on the protein band to measure to, because the protein bands can be quite fat in the agarose gels. So long as they are consistent, they can either measure to the base of each band or to the middle of each band.

Student Protocol
Lesson 2: Agarose Electrophoresis – Agarose Gel Loading, Running, Fixing, and Staining

So far you have extracted, denatured, and given the proteins from fish muscle tissue a negative charge. Now they can be separated according to their molecular weights using gel electrophoresis, which will generate profiles for various fish species

Student Workstations

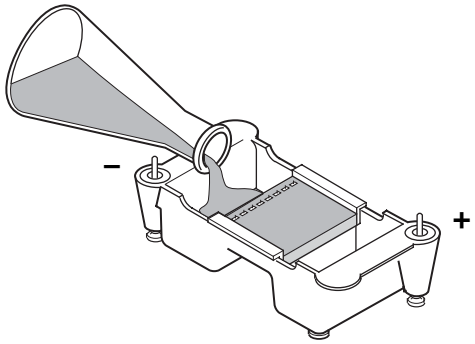
Materials Required	Quantity
Fish protein extracts from lesson one	5 species
Actin & myosin standard, 25 ul	1 vial
Precision Plus Protein Kaleidoscope prestained standards, 6 µl	1 vial
4% TGS PCR low-melt agarose gel	1
1–20 µl adjustable-volume micropipet	1
Pipet tips for gel loading	7 tips
Horizontal electrophoresis module (gel box)	1
1x Tris-glycine-SDS (TGS) running buffer	275 ml per gel box
Power supply (100 V constant) to be shared between workstations	1
Staining trays	1 per 2 gels
Gel fixative solution	100 ml
Bio-Safe Coomassie stain for proteins	100 ml

Common Workstation

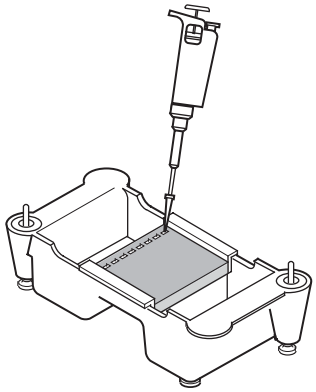
Materials Required	Quantity
Water bath set at 95°C	1
Rocker	1
Water for gel destaining (tap water is fine)	

Protocol

1. Obtain a TGS low-melt agarose gel from your teacher, or, if your teacher instructs you to do so, pour your own gel.
2. Place the casting tray, with the solidified gel in it, onto the central platform in the gel box. The wells should be at the negative (cathode) end of the box where the black electrical lead is connected. Very carefully remove the comb from the gel by pulling it straight up.
3. Pour about 275 ml of TGS gel running buffer into the electrophoresis chamber. Pour enough buffer into the box until it just covers the wells of the gel by 1–2 mm.



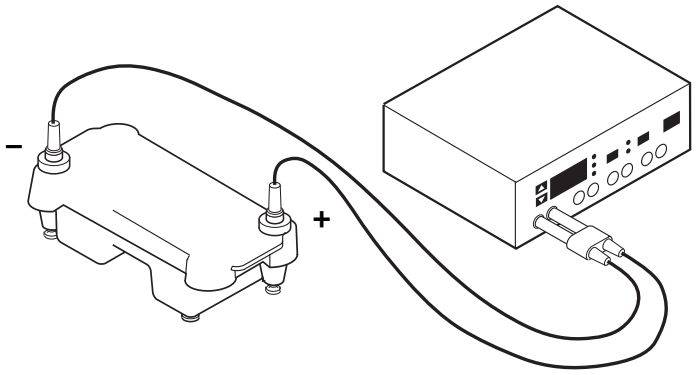
4. Obtain your 5 muscle extracts from the previous lesson. Heat to 95°C for 2–5 minutes to redissolve any precipitated detergent. Note: If you have prepared your muscle extracts today, there is no need to reheat the samples.
5. Using a fresh pipet tip and either an adjustable micropipet set to 5 µl, or a fixed-volume 5 µl micropipet, load 5 µl of Precision Plus Protein Kaleidoscope prestained standards to lane 1 of your TGS agarose gel. **Important: use a fresh tip each time**.



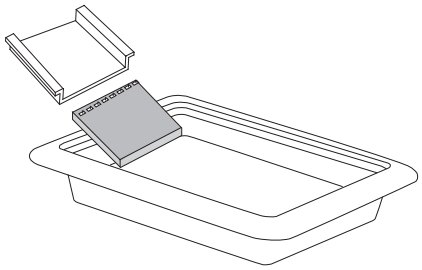
6. Using the table below as a guide, load 20 µl of each of your protein samples and the Actin & Myosin standard into your gel in the order indicated below. Record your gel loading in the table below.

Lane	Volume	Sample Name
1	5 µl Stds	Precision Plus Protein Kaleidoscope prestained standard (Stds)
2	20 µl sample A	_____
3	20 µl sample B	_____
4	20 µl sample C	_____
5	20 µl sample D	_____
6	20 µl sample E	_____
7	20 µl AM	Actin & myosin standard (AM)
8	empty	none

5. Slide the cover of the chamber into place, and connect electrical leads to the power supply, anode to anode (red to red) and cathode to cathode (black to black). Make sure both electrical leads are attached to the same channel of the power supply.
6. Electrophorese at 100 V for 45 minutes. Shortly after the current is applied, the loading dye can be seen moving through the gel toward the positive side of the gel chamber.



7. When electrophoresis is complete, turn off the power supply, disconnect the leads from the inputs, and remove the lid from the gel box.
8. Carefully remove the gel tray and the gel from the gel box. Be careful, the gel is very slippery. Nudge the gel off the gel tray with your thumb and carefully slide it into the gel staining tray for fixing. Add 100 ml of gel fixative and place gel on rocking platform for 1 hour to overnight. Taking the extra time to fix the gels overnight results in much better resolution. Multiple gels can be fixed in the same container.



9. After fixing, pour off the gel fixative and add sufficient Bio-Safe Coomassie stain to cover the gel and leave to stain on a rocking platform for 2 hours to overnight. Multiple gels can be stained in the same container.
10. After staining, pour off the stain and rinse the gel with distilled water and then destain the gel on a rocking platform overnight with at least 3 changes of water. Note: you probably will not see any bands on the gel when you begin to destain. The bands will develop slowly as the gel destains. Destain until you see a nice contrast between the protein bands and the gel background.
11. If you wish to keep a permanent record of your gel, the gel can be dried between cellophane sheets in the manner described in lesson 3. Please note, however, that you must not use a gel drying oven for this process since the gels will melt. You can also peel away the cellophane once the agarose gel has dried; this is not possible with polyacrylamide gels.

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 have your class go online and obtain their 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. 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

```
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

```
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

```
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

```
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
```

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**

APPENDIX B

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

Which two fish share the fewest amino acids?

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?

Construct a cladogram based on this data:

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. 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 and construct a phylogenetic tree based on that data. 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.

Phylogenetic Tree of Fish

Does the taxonomic classification support the molecular data?

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.

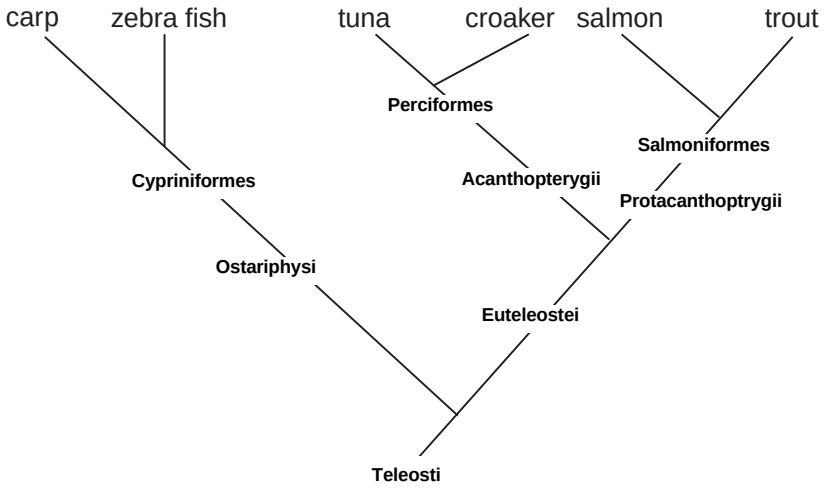
1) Go to http://www.ncbi.nlm.nih.gov/ and choose BLAST
2) Choose Protein-Protein BLAST.
3) Enter your myosin sequence into the search box.

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.

Construct a simple phylogenetic tree based on the taxonomic data (the large phylogenetic tree figure will be useful here).



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

The trees do not entirely match. Both trees show a close relationship between salmon and trout and zebra fish and carp. However, tuna is in the same sub-phylum (Euteleostei) as salmon and trout, yet this does not concur with the molecular data and croaker is in the same order as tuna (Perciformes) and yet the amino acid sequence of croaker's myosin is much closer to salmon than tuna.

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

The statistical relevance of data grows as the size of the data set increases. The 60 amino acid segment of myosin heavy chain constitutes just 3% of the myosin heavy chain molecule, which is around 1,900 amino acids long. Performing a BLAST search with a larger portion of the molecule generates a cladogram with different relationships, demonstrating that the 60 amino acid piece is not large enough to provide a full picture of relatedness. However, even if the full-length myosin were compared, that is just a single protein out of the thousands generated by the organism. The data would be much stronger if the sequences of multiple proteins were compared and stronger still if molecular data were used with other types of classification data such as morphological data.

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

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

Posttranscriptional modification – alterations that allow one gene to code for many proteins

Posttranslational modification – alterations of proteins after they are synthesized by the cell

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