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

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 is 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 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. 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. 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.

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.

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).

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.

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.

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.

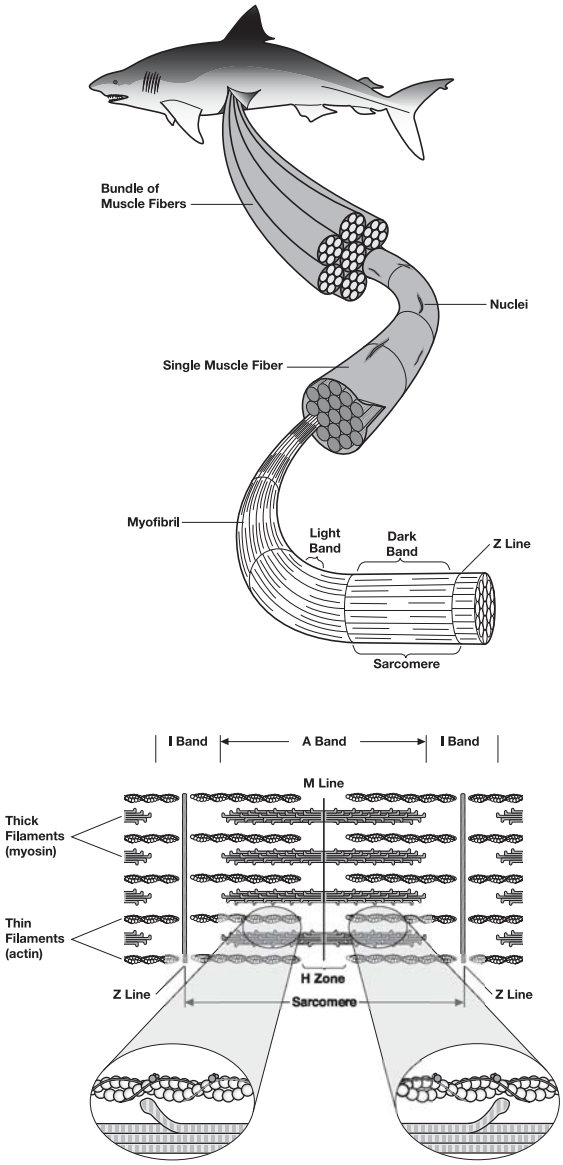


Fig. 12. Telescopic view of muscle structure: Thick myosin filaments and thin actin filaments form myofibrils, which are bundled together to make muscle fibers. (Figure modified from Campbell 1996 with permission.)

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.

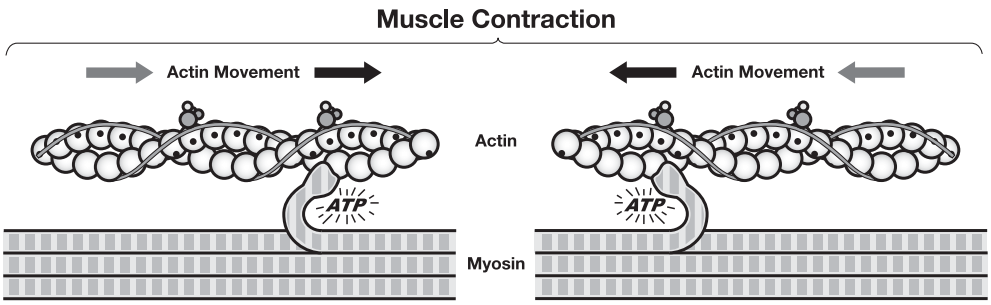


Fig. 13. Hydrolysis of ATP causes myosin and actin filaments to slide past one another, shortening the sarcomere and contracting the muscle. (Figure modified from Campbell 1996 with permission.)

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

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).

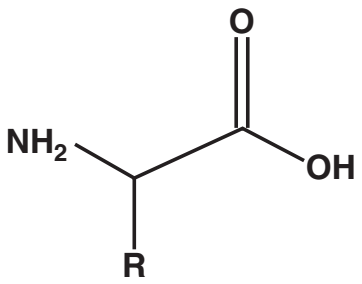


Fig. 15. Chemical structure of an amino acid.

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 cysteines. 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.

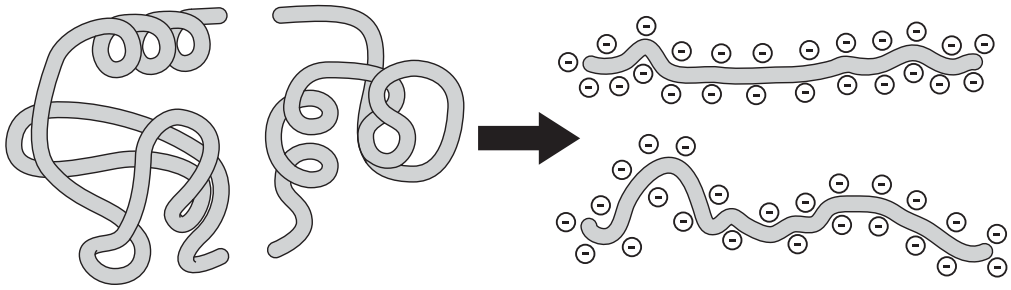


Fig. 16. The combination of heat and the detergent SDS denatures proteins for SDS-PAGE analysis.

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

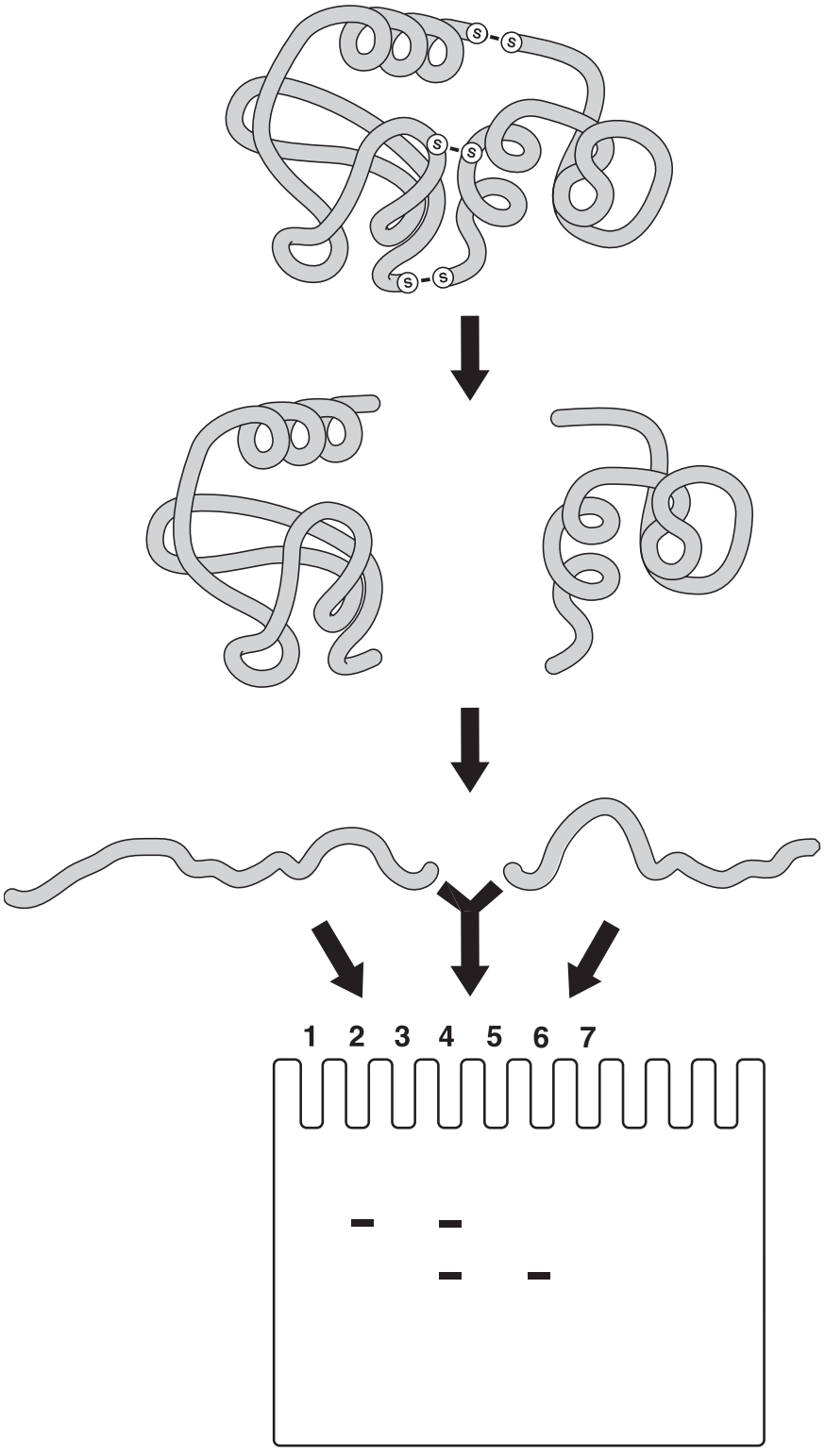


Fig. 17. A quaternary protein complex denatured with reducing agents, heat, and SDS, can be separated into individual proteins and resolved by size using SDS-PAGE.

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

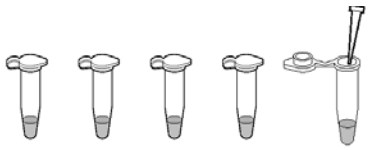
Material	Quantity
1.5 ml fliptop microtubes	5
1.5 ml screwcap microtubes	5
1 ml transfer pipet	1
Fish samples, labeled 1–5	5 species
Marking pen	1
Laemmli sample buffer	1.5 ml
Knife or scissors to cut fish samples	1

Common Workstation

Material	Quantity
Water bath set to 95°C	1

Procedure

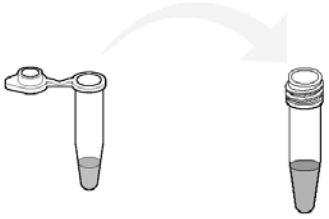
1. To make this a blind study, assign a letter (e.g., A–E) to each fish sample to be investigated. Keep a record of which fish got which number and hide their true identities until after the analysis is complete.
2. Label 1.5 ml fliptop microtubes with the number of the fish species to be analyzed. There 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 min 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 in their screwcap tubes for 5 min 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 hr, 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

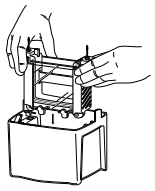
Material	Quantity
Fish protein extracts from lesson one	5 species
Actin & myosin standard, 12.5 ul	1 vial
Precision Plus Protein Kaleidoscope prestained standards, 6 µl	1 vial
4-20% Mini-PROTEAN TGX or 15% 10-well, Ready Gel precast gel	1 vial
1–20 µl adjustable-volume micropipet	1
Prot/Elec pipet tips for gel loading	7 tips
Mini-PROTEAN Tetra cell electrophoresis module	1 per 2 gels
1x Tris-glycine-SDS (TGS) running buffer	700 ml per gel box
Power supply (200 V constant) to be shared between workstations	1
Sample loading guide – for 10-well comb (optional)	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

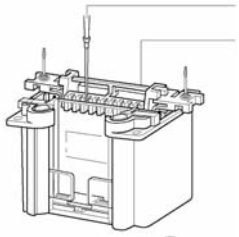
Material	Quantity
Water bath set at 95°C	1
Water for gel destaining (tap water is fine)	

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 TGX or 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.

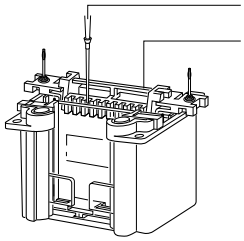


5. Record in which well of your gel you will load which of your samples in the table below:

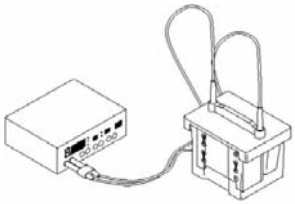
Well	Volume	Sample Name
1	empty	none
2	empty	none
3	5 µl Stds	Precision Plus Protein Kaleidoscope prestained standard (Stds)
4	*10 µl sample A	_____
5	*10 µl sample B	_____
6	*10 µl sample C	_____
7	*10 µl sample D	_____
8	*10 µl sample E	_____
9	*10 µl AM	actin & myosin standard (AM)
10	empty	none

*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.

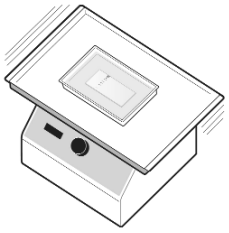
6. Load **5 µl** of Precision Plus Protein Kaleidoscope prestained standard gently into well # 3 using a thin gel loading tip. **Note**: The fine barrel of the gel loading tips means liquid is slower to go into the tip than normal tips. You must therefore release the plunger of the micropipet very slowly,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



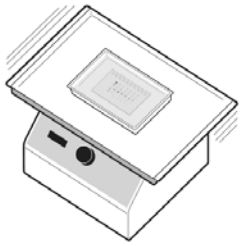
7. Using a fresh tip each time, load **10 µl** of each of your protein samples gently into the wells designated in your table above. Note: If you are going on to perform the western blot module, load 5 µl of each protein sample.
8. Using a fresh tip, load **10 µl** of the actin & myosin standard gently 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. Carefully pry apart the gel plates, using the gel opening key. 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. Rinsing the gel will improve the intensity of the protein bands.
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 at least 1 hour, with shaking if available. Gels maybe stained overnight but seal the container to reduce evaporation.
15. After at least 1 hour discard the stain and replace it with a large volume of water to destain the gel overnight with rocking action if available. Change water 2–3 times if possible. Bands will become visible after a few hours of destaining.



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 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 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.

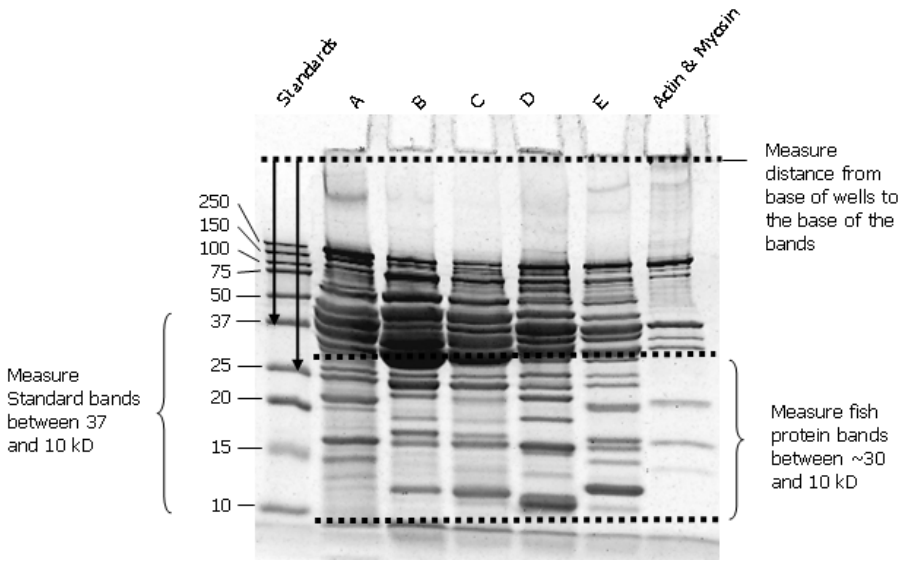


Fig. 18. Image of fish muscle proteins separated by SDS-PAGE and stained with Bio-Safe Coomassie stain. Lines illustrate measurement of bands for constructing the standard curve.

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.

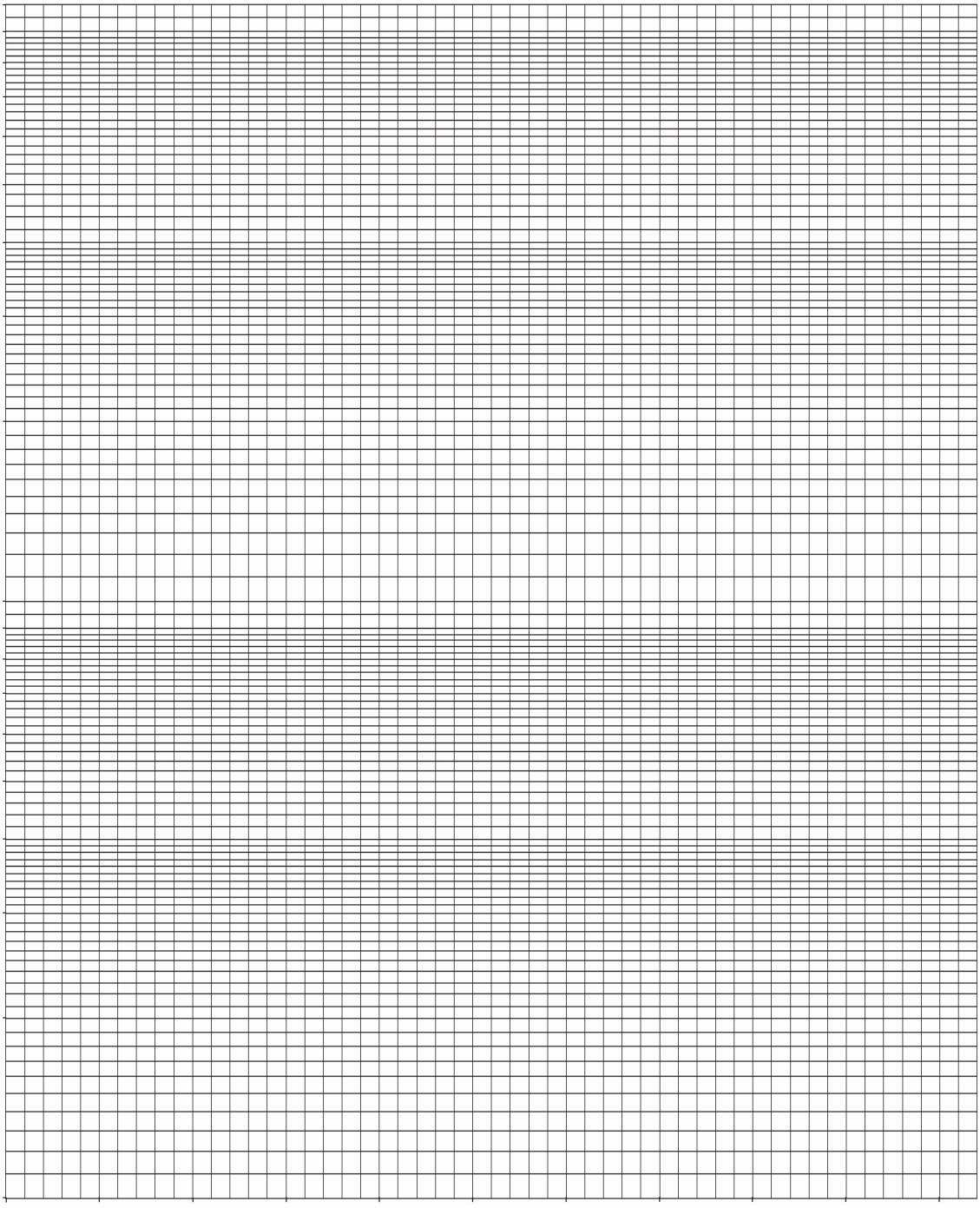


Fig. 19. Two-cycle semi-logarithmic graph paper to construct curve of the protein molecular mass against the distance migrated.

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**

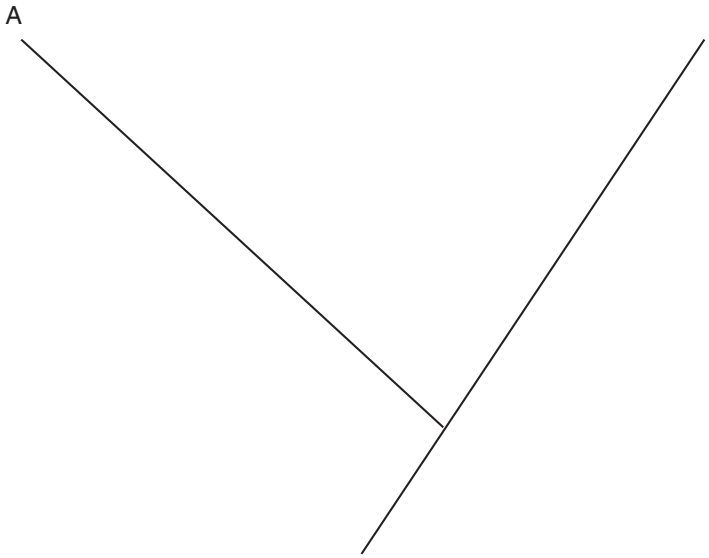
Find the number of characteristics shared by each of the fish

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.

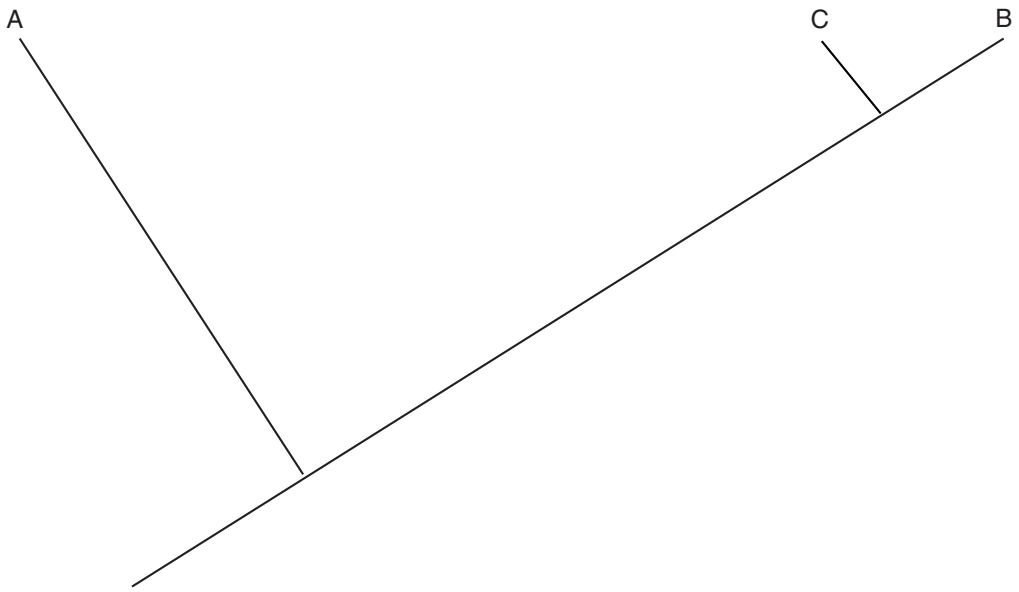
	Species A	**Species B**	**Species C**	**Species D**	**Species E**
Species A	8	2	2	2	2
Species B		10	10	5	3
Species C			13	5	4
Species D					2
Species E					12

Construct your cladogram

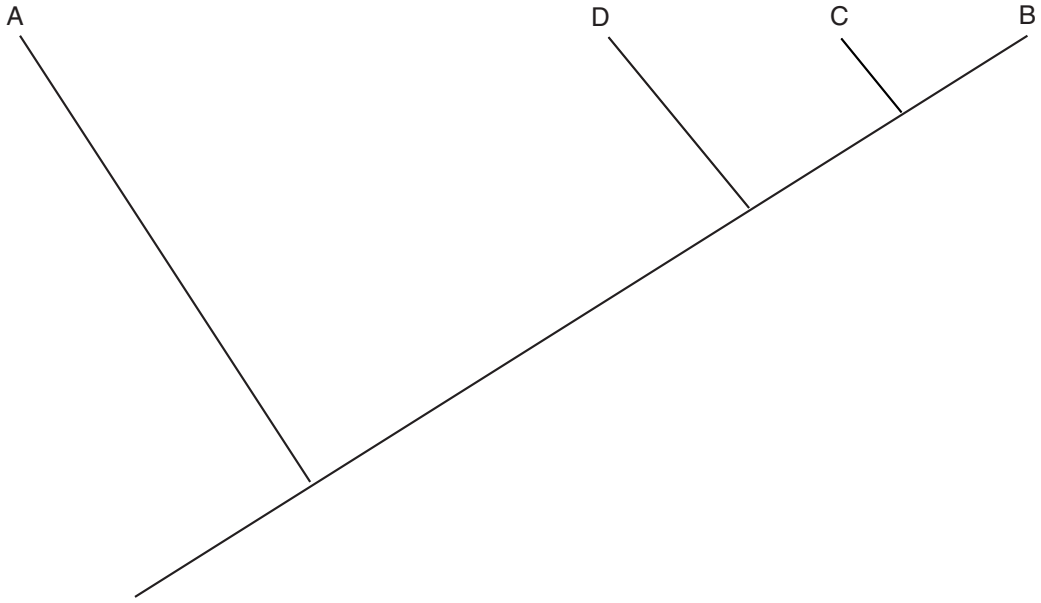
8. Now you are ready to construct your cladogram. First draw a line to form the trunk of your cladogram. Find the fish with the least bands in common. In the example above it is species A, which has only 2 bands in common with any of the other fish. Then draw a side branch off the line near the bottom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

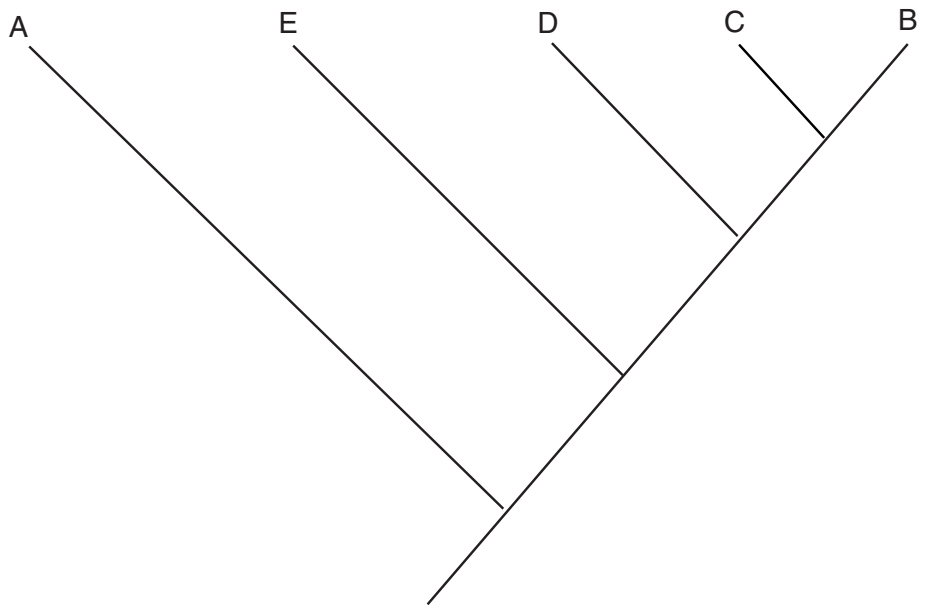


The last fish to add to the cladogram in this example is species E, which shares four bands with species C, three bands with species B, and only two bands with species A and D. This fish may seem trickier to place than the others because it shares more characteristics with species B and C than it does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