**Investigating Evolutionary Relationships through Morphology and Molecular Analysis**

*Student Instructions Packet*

**Terminology**

|  |  |  |  |
| --- | --- | --- | --- |
| * Morphology
 | * Homologous structure
 |  |  |
| * Molecular Data
 | * Analogous structure
 |  |  |
| * Divergent evolution
 | * Cladogram
 |  |  |
| * Convergent evolution
 | * Shared derived characters
 |  |  |

**The Difficulty in Classifying Life**

Humans tend to organize things into groups to make sense out of a world of chaos. We have been organizing and classifying the world of biology for hundreds of years. Early classification relied on organism morphology for comparison. If an animal had horns it might be placed into one group, hair another, scales into yet another and so on. Many of these morphological characteristics correlate with homologies. Homologous structures are those that they share with a common ancestor but have changed or been refined over time as organisms moved into differing habitats or had differing needs.

**Homologous structures** point to **divergent evolution**. Birds, bats, whales, and humans all have similar skeletal forelimb structures. The similarities are quite clear when looking at the bones but looking at the organisms themselves they are more difficult to see. This is because these animals shared a common ancestor millions of years ago but as its descendants moved into differing habitats the selective pressures were different on each producing a wide array of refinements to the forelimbs.

Scientists can make cladograms based on analysis of common morphological features or homologies. The cladogram below was constructed by analyzing features this group has in common or **shared derived characters** which are listed at the base of the tree. These are traits that arose in the most recent common ancestor of a particular lineage and was passed along to its descendants. This means that all organisms “upstream” on the cladogram will share this trait, those behind it will be excluded because they lack this trait. This concept has been illustrated here using colored boxes. The organism at the top will share all of the derived characters “downstream” of it but be the only one to exhibit the final derived character. For example, in this cladogram the all the turtle is included because it is a vertebrate and has a backbone. It is excluded from the rest of the group due to the rest of the group sharing the derived character hair. Conversely the house cat is at the top of the cladogram as it exhibits all derived characters.

**Convergent evolution** can be a major pitfall for classification, however. This occurs when unrelated organisms living is similar habitats and occupy similar niched evolve similar features through natural selection to be successful in their respective environments. Convergent evolution leads to **analogous structures** like those of the flipper of a penguin and the flipper of a seal. Both have flippers but not due a common ancestry. The same can be said of the wings pterosaurs (flying extinct dinosaur), birds and bats. None of these animals share a common winged ancestor, therefore the wings are said to be analogous. They do share a common tetrapod (four-limbed) ancestor which makes their forelimbs homologous. In this way it becomes clear how difficult it can be to classify organisms based on morphology alone.

 There are many strange organisms that simply defied classification based on morphology. These organisms either had very few shared derived characters or too many to make a solid hypothesis. For others the scientists debated as to where to put these organisms in the tree of life was so great that they seemed to remain in limbo. These organisms were doomed to remain in their classification purgatory until the hypothesis regarding their lineages could be tested. Luckily the Human Genome Project that was completed in 2001 initiated a fever on gene sequencing and today the genomes of most known species have been sequenced. Modern technology now allows scientists to compare gene and protein sequences of organisms to determine evolutionary relatedness. The difference in sequences results from mutations that have accumulated in the DNA over time. More differences indicate that a longer amount of time has passed since the species diverged. The more alike the sequences, the more recent the common ancestor and visa-versa. Scientists can use this information to construct cladogram and support or reject current hypothesis on evolutionary relationships.

 Bioinformatics is an in-demand field that combines statistics, mathematical modeling, and computer science to analyze biological data. Using bioinformatics methods, entire genomes can be quickly compared in order to detect genetic similarities and differences. An extremely powerful bioinformatics tool is BLAST, which stands for Basic Local Alignment Search Tool. Using BLAST, you can input a gene or protein sequence of interest and search entire genomic libraries for identical or similar sequences in a matter of seconds.

During the first part of this lab cladograms will be constructed based on morphological data analysis. The second part of the lab will require the use of bioinformatics. Here we will be using the Uniprot online molecular database and blast tool. It is a free service where any person in the world can look up or compare gene sequences for almost any organisms imaginable. Organism groups will be compared using the protein cytochrome b (CYTB). This is an excellent protein to compare as it is highly **conserved** and part of the cytochrome c complex used in cellular respiration and the generation of ATP. Respirating organisms will all contain this protein.

Directed Note-Taking

Directions: Record notes containing the most important information relevant to the guiding question.

**“The Difficulty in Classifying Life”**

**Guiding Question:** How are evolutionary relationships investigated using morphological and molecular data?

 Check Relevant Categories

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Paragraph | Notes | Morphology | Homologous Structures | Analogous Structures | Divergent Evolution | Convergent Evolution | Molecular Data |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |

Collaborative Work: After completing your chart, be prepared to compare your notes with others.

Group Assignments

**Group 1: Which animal is the hyena most and least closely related to?**

 *Hyaena hyaena (*Striped hyena), *Canis lupus* (Gray wolf), *Acinonyx jubatus (****Cheetah***), *Paguma larvata* (Masked palm civet), Cryptoprocta ferox (Fossa)

**Group 2: Which animal is the echidna most related to?**

*Dasypus novemcinctus* (Nine-banded armadillo), *Manis pentadactyla* (Chinese pangolin), *Zaglossus bruijni* (Long-beaked echidna*), Choloepus didactylus* (Southern two-toed sloth)*, Ornithorhynchus anatinus* (Duckbill platypus)

**Group 3: Which animal is the okapi most and least closely related to?**

*Okapia johnstoni* (Okapi*), Syncerus caffer* (Cape buffalo), *Giraffa camelopardalis* (Giraffe), *Equus zebra hartmannae* (Hartmann's mountain zebra), *Odocoileus virginianus* (white-tailed deer)

**Group 4: Which animal is the narwhal most and least closely related to?**

*Tursiops truncatus* (Atlantic bottle-nosed dolphin) (*Delphinus truncatus*), *Galeocerdo cuvier* (Tiger shark*), Pastinachus sephen* (Cowtail stingray), *Hippopotamus amphibius* (Hippopotamus), *Monodon monoceros* (Narwhal)

**Group 5: Which animal is the Red Panda most and least closely related to?**

Ailuropoda melanoleuca (Giant panda), *Ailurus fulgens* (Lesser panda) (Red panda), *Procyon lotor* (Raccoon), *Ursus thibetanus* (Black bear)

**Group 6: Which animal is the Malayan tapir most and least closely related to?**

*Equus caballus* (Horse), *Rhinoceros unicornis* (Greater Indian rhinoceros), *Tapirus indicus* (Asiatic tapir) (Malayan tapir), *Camelus dromedarius* (Arabian camel), *Phacochoerus africanus* (Warthog*), Loxodonta africana* (African elephant)

**Group 7: Which animal is the elephant most and least closely related to?**

 *Loxodonta africana* (African elephant), *Procavia capensis* (Cape hyrax) (Rock dassie), *Hippopotamus amphibius* (Hippopotamus), *Myocastor coypus* (Coypu) (Nutria), *Dugong dugon* (Dugong)

**Group 8: Which animal is the extinct Thylacine most and least closely related to?**

*Thylacinus cynocephalus* (Tasmanian wolf), *Panthera tigris* (Tiger), Canis lupus (Gray wolf), *Macropus rufus* (Red kangaroo)

**Group 9: Which animal is the Polar Bear most and least closely related to?**

*Pagophilus groenlandicus* (Harp seal), *Lontra canadensis* (River otter), *Aplodontia rufa* (Mountain beaver), *Taxidea taxus* (Badger), *Ursus maritimus* (Polar bear)

**Group 10: Which animal is the kangaroo most and least closely related to?**

*Macropus rufus* (Red kangaroo), *Dolichotis patagonum* (Patagonian cavy), *Lepus timidus* (Mountain hare), *Phascolarctos cinereus* (Koala)

Procedure

Part 1: Relationships Based on Morphology

1. Begin by filling out Table 1 for your assigned group. Scientific name should be written in proper format (i.e., whole name italicized, genus is capitalized, species is lower case - *Homo sapiens*)
2. Table 2.
	1. Fill in the names of your organisms across the top row of the table. There may be more boxes than you have organisms - that’s okay.
	2. Your teacher may give you pictures of your organisms or you may have to look them up online. Be sure to research them using their scientific name. For pictures and videos of animals use www.Arkive.org.
	3. Examine the photos of your organisms looking for morphological similarities. Begin making a list of theses similarities - choose 5-7 and write them in the feature column of Table 2. Place check marks below each organism that each feature applies to.
3. Examine Table 2 for commonalities. The more check marks organisms have in common, the more features they share. These features can be considered **shared** **derived characters** (trait that arose in the most recent common ancestor of a particular lineage and was passed along to its descendants; all organisms “upstream” on the cladogram will share this trait, those behind it will be excluded because they lack this trait).
	1. Construct a cladogram based on the morphological similarities you have uncovered in Table 2.
	2. Be sure to place shared derived characters in the appropriate places on the diagram.
4. **Preliminary Conclusion**. Now, use the data gathered in Table 1 and represented in your cladogram in Figure 1 to generate hypotheses about the relationships among organisms and express these using the CER model.
	1. Claim (Who is your assigned animal most and least closely related to?)
		1. Note there are 2 claims here so you will go through these steps twice.
	2. Evidence (How did you decide that?)
	3. Reasoning (WHY does your evidence support your claim?)
5. Present findings to peers.

Part 2: Relationships Based on Molecular Data

1. Fill in the names of your organisms across the top row of Table 3. Use the same order to fill the names down the first column, so that the 100% in the chart is the value for same organisms.
2. Go to <http://www.uniprot.org/> to perform molecular comparisons. Under the “Search” tab at the top type the scientific name of your first organism followed by CYTB in the “Query” box. Be sure that the “Search in” box lists “Protein Knowledge Database (UniPrtoKB)”. Click search.
3. A list of organisms will be retrieved. Check the box at the left of the desired organism/protein match. The results with gold stars are the most desirable. Be sure the organism is correct, the protein is cytochrome b, and acceptable gene names are MT-CYB, COB, CYTB, MTCYB.
4. Follow step 3 again for your second organism.
5. The green bar at the bottom of the page should say “2 selected” with the entry number of the two that you selected. Click the button for “Align” on the right side of the green bar (this may take a few seconds).
6. The Alignment retrieves the amino acid sequence for cytochrome b for both organisms and denotes differences between the two sequence with either two dots ( : ) or one ( . ). Same amino acids are marked with an asterisk (\*).
7. Scroll down past the Alignment to “Job Information” and find the Identity % for the two organisms you are comparing. Fill the number in the appropriate box in the chart where the two organisms meet. This is the percentage of amino acid sequence these organisms have in common for this protein. If the organisms were the same the identity would be 100%. The different amino acids result from mutations that have accumulated in the DNA over time. More differences indicate that a longer amount of time has passed since the species diverged.

The higher the parentage, the more similar the sequences and therefore the more closely related these organisms are.

1. At the bottom of this page in the “Annotation” section UNCHECK your second organism. The green bar on the bottom should say “1 selected” now.
2. Scroll back up to the top of the page and click the “Search” tab. Type in the name of the next organism to compare the first to and click search. A distance matrix is a table that shows all the pairwise comparisons between species. Continue to make all pairwise comparisons until Table 3 is filled. Follow steps 3 through 7 for each.
3. Use Table 3 to answer questions 3.1-3.3.
4. Analyze the molecular data compiled on the Distance Matrix (Table 3) as well as your answers to questions 3.1-3.3 to construct a revised cladogram based on molecular data in Figure 3.
5. **Conclusion**. Compare your cladograms. Use the CER model to answer the following question: Is the second cladogram built from new information (molecular data) consistent with your earlier morphology based hypothesis (your first cladogram)?
6. Present findings to peers.

**Investigating Evolutionary Relationships through Morphology and Molecular Analysis**

Data Sheet

TABLE 1: List of Organisms for Analysis

|  |  |
| --- | --- |
| Common Name | Scientific Name |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

\*\*\* You will be provided figures for your assigned organisms\*\*\*

**Evolutionary Relationship Based on Morphology**

TABLE 2: Morphological comparison

|  |  |
| --- | --- |
| Feature | Organisms |
|  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

**Cladogram Based on Morphology**

Figure 2

*\*\*\*Don’t forget to include derived characters\*\*\**

|  |
| --- |
|  |

**Preliminary Conclusion**

*CER: Claim, Evidence, Reasoning*

|  |  |
| --- | --- |
| **Question** | Which organism from your list is most closely related to the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_? |
| **Claim** |  |
| **Evidence** |  |  |  |
| **Reasoning** |  |  |  |

**Evolutionary Relationship Based on Molecular Data**

*Blast Analysis*

Table 3: Distance Matrix (X’s represent repeat data and do not need to be filled in)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Write Animals Here in Same Order Down & Across 🡺🡻 |  |  |  |  |  |  |
|  | 100% |  |  |  |  |  |
|  | XXXXX | 100% |  |  |  |  |
|  | XXXXX | XXXXX | 100% |  |  |  |
|  | XXXXX | XXXXX | XXXXX | 100% |  |  |
|  | XXXXX | XXXXX | XXXXX | XXXXX | 100% |  |
|  | XXXXX | XXXXX | XXXXX | XXXXX | XXXXX | 100% |

3.1) Which two species in the above table have the most similar cytochrome b sequences?

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

3.2) Which two species in the above table have the least similar cytochrome b sequences?

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

3.3) For your comparison organism \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, make a list of species that have the most similar protein sequence to the least similar:

a. (most similar): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_ % identity

b. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_ % identity

c. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_ % identity

d. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_ % identity

e. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_ % identity

f. (least similar): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_ % identity

**Cladogram Based on Molecular Data**

Figure 3

|  |
| --- |
|  |

**Conclusion**

*CER: Claim, Evidence, Reasoning*

|  |  |
| --- | --- |
| **Question** | Is the second cladogram built from new information (molecular data) consistent with your earlier morphology-based hypothesis (your first cladogram)? |
| **Claim** |  |
| **Evidence** |  |  |  |
| **Reasoning** |  |  |  |