

Mammalogy Laboratory 5 - Systematics

I. Goal

The goal of this lab is to teach you a little about the field of systematics, especially as it pertains to mammalogy. Specifically, the exercises in this lab will familiarize you with how mammalian species are delimited and how phylogenies are inferred. The methodologies you will be using are not restricted to mammalian systematics, but they have been chosen for inclusion in this lab because they have been particularly helpful as we continue to uncover interesting aspects of mammalian diversity.

II. Species Limits

A. Species Concepts.

One of the most active fields of mammalian systematics has been the circumscription of species limits. However, species concepts have changed dramatically over the last 20 years, and there is no current consensus as to what a species actually is. The problem stems from the following. There is relatively little difficulty in recognizing the “packaging of variation” that we see in nature at a particular location (say Moscow Mountain). That is, all the mammals we’ll find at that site fall into more or less discrete morphologies and species are easy to differentiate (at least with training). These boundaries, however, break down as we increase the scope of our attention both geographically and temporally (i.e., back into the fossil record). The discrete packages that we observe on a local scale start to merge.

The classic species definition that we learn as budding biologists is that of Ernst Mayr, the **Biological Species Concept (BSC)**. This has been the standard species concept in mammalogy, and many mammalian systematists still subscribe to it, although most other systematists don’t. Under this concept, species are the largest groups of interbreeding (or potentially interbreeding) populations. If there is reproductive isolation between two sets of populations, then under this species concept they are considered different species. This isolation can take on many guises and the study of isolating mechanisms has been a very active field for over 80 years. There are several problems with this concept, however. Allopatric populations (populations that are geographically isolated) are problematic because there is no possibility to test the criterion of reproductive isolation. We must therefore use a morphospecies concept and infer the presence/absence of reproductive isolation based on the degree of differentiation. Thus, application of this species concept leads to subjective decisions. A second problem with this concept is that hybridization occurs in many groups. Furthermore, there actually is a continuum in the degree of hybridization; there is relatively little among mammalian species, moderate amount among fish species, and relatively more among some plant groups. A third problem is that many groups don’t rely on sexual reproduction. Although this is not true for mammals, there are many vertebrate taxa in which the males have absolutely no genetic contribution to the subsequent generation. These last two situations render the criterion of reproductive isolation entirely subjective; there is a continuum of reproductive isolation in nature.

Several alternative species concepts have been proposed as a result of the above problems with the BSC. Many systematists have espoused the **Phylogenetic Species Concept (PSC)**. Under this concept, species are defined as sets of populations that exhibit fixed differences from other sets of populations. That is, this concept emphasizes diagnosability as its criterion. This has the advantages of being objective, applicable to non-sexually reproducing organisms, and the issue of allopatry/sympatry is irrelevant. However, this criterion will tend to over represent diversity, and a recent study suggested that application of the criterion of diagnosability would result in essentially doubling the number of bird species recognized.

A second alternative species concept is the **Genealogical Concordance Species Concept (GCSC)**. This has been proposed by several authors, including John Avise (a vertebrate zoologist), David Baum (a botanist), and Kerry Shaw (an entomologist). This is based on the PSC, in that diagnosability is a major criterion, but is more restrictive in that many character systems (i.e., morphology, allozymes and DNA sequences) must all exhibit similar patterns of divergence (that is, be concordant). It has the disadvantage of under-representing diversity, in that if any single character system is discordant, then the groups fail to meet the criteria. As we have learned more about molecular data (such as DNA sequences, for example), it has become obvious that polymorphisms may persist for many millions of years.

As mentioned above, there is no current consensus as to the most widely applicable species concept, and there are other concepts I haven't mentioned. At this point, systematists are left having to choose a concept, defend it, and apply it to a particular taxonomic question.

B. Types of Data.

Morphological Data. Without a doubt, studies of morphological variation have contributed the lion's share of our knowledge of mammalian species. This field of inquiry, morphometrics, has relied on cranial characters more than any other character system. The next two pages define many of the measurements that are typically taken. In general, the approach has been to take as many of these measurements as possible for crania of 10 or more individuals from two or more sets of populations and look for discontinuities.

The following descriptive statistics are commonly employed:

Mean - an indication of central tendency.

Arithmetic Mean:

$$\bar{X} = \frac{\sum X}{N}$$

Where X is the observation and N is the number of measurements taken.

Range = Largest measurement and smallest measurement.

Variance - one measure of the dispersion of points around the mean.

$$s^2 = \frac{\sum (\bar{X} - X)^2}{N}$$

Standard Deviation (s) = the square root of the variance.

Classically, these were compared visually in the form of Dice-Leraas Diagrams (Dice-grams), and taxonomic decisions were made on the basis of non-overlapping standard deviation. However, a more rigorous treatment of these univariate statistics can be accomplished by means of t -tests. This traditional statistical test is designed to compare two samples to determine if they represent different samples from the same underlying distribution

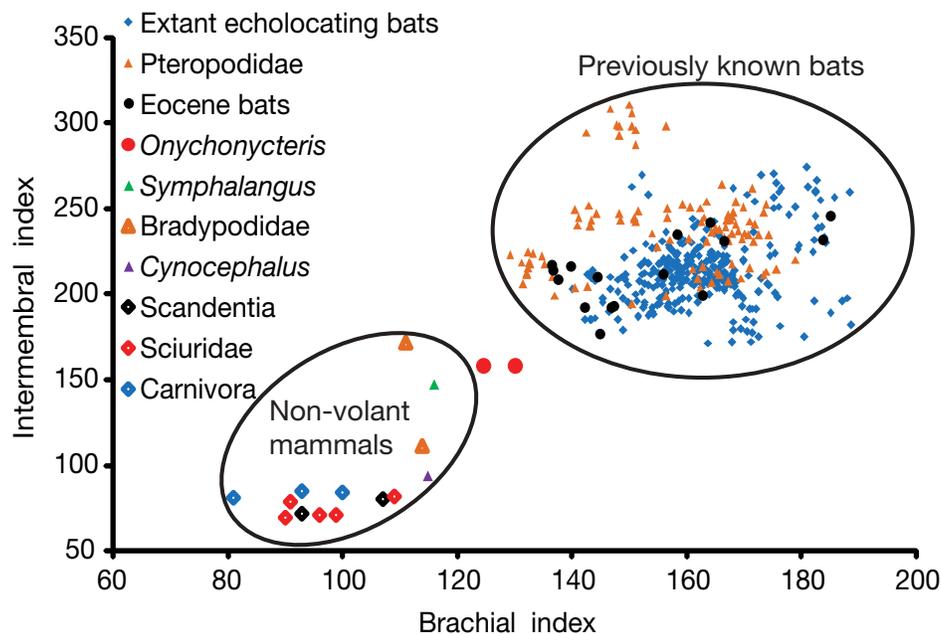
(it is really a special case of the more common ANOVA). That is, this method examines whether or not the observed difference in the two sample means is the result of sampling error.

The test statistic is calculated as follows:

$$t = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{s_1^2 + s_2^2}{N}}}$$

This test statistic has $2(N-1)$ degrees of freedom and can then be compared to a table of critical values of the t -distribution. If the test statistic is higher than the critical value for the appropriate degrees of freedom at some (arbitrary) alpha value (usually $\alpha = 0.05$), then we can reject the null hypothesis of no significant difference in between the two means.

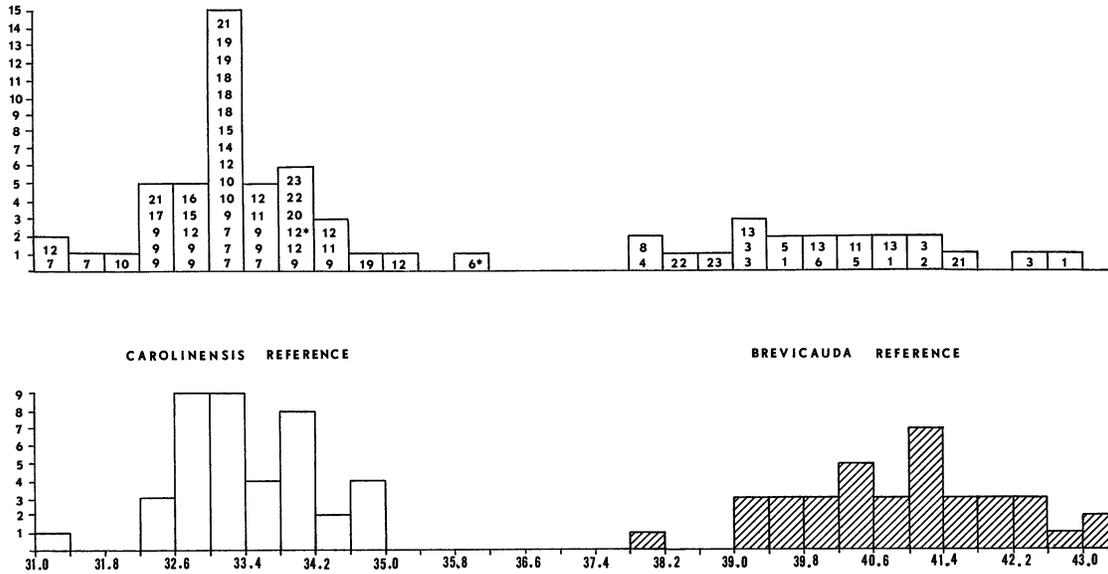
Another way we can examine these measurements is by use of scatter plots, in which the value of one character is plotted on one axis (the abscissa) and another is plotted on the other axis (the ordinate). An example is presented below (from Simmons et al., 2008).



An example of a bivariate scatter plot.

More frequently, however, researchers are using multivariate statistics to analyze morphological variation. Two commonly employed techniques are Principle Components Analysis (PCA) and Discriminant Function Analysis (DFA). The first of these, PCA, is used to cluster samples in multivariate space, the idea being to represent that n -dimensional space (n is the number of variates, or characters) in three dimensions (or components). DFA is a multivariate method that assigns a single index to a specimen. Specimens of known taxonomic affiliations are used as a

reference (training) collection, a number of measurements are taken for each individual and an equation (the discriminant function) is derived to calculate an index that will identify an unknown specimen. The example below illustrates how DFA works. The lower histogram represents the distribution of discriminant scores calculated for a number of known individuals of *Blarina brevicauda* and *Blarina carolinensis*, two very closely related species of short-tailed shrews. The upper histogram represents the distribution of scores of unknown affinity; each specimen can be assigned to its appropriate species based on its discriminant score. The important assumption of DFA is that the reference samples have been properly classified on the basis of some independent grounds.



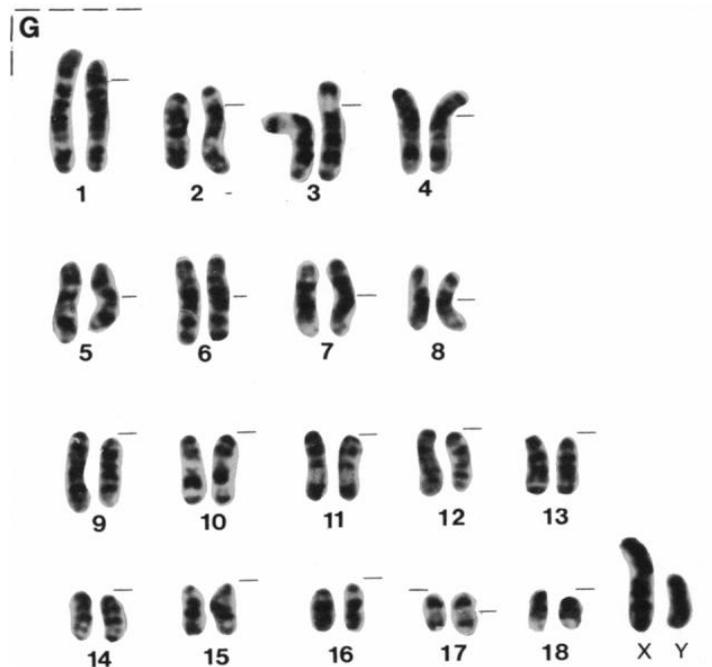
Exercise 1. Examine the two sets of skulls in the lab and take the following six cranial measurements for each of these: greatest length of skull, maxillary tooth row, interorbital width, zygomatic breadth, two other measurements (from among the possibilities on the figures). Calculate the descriptive statistics of mean, range, and standard deviation. Present these in the form of a table. A separate handout illustrates these measurements.

Exercise 2. Use a bivariate scatter plot to illustrate two cranial characters that differentiate these two taxa.

Exercise 3. Use a *t*-test to determine if the differences between these two sets of skulls is significant (use the same two characters you used for the scatter plot - so do two different *t*-tests). Use $\alpha = 0.05$, and the appropriate degrees of freedom to look up the critical value in the table provided. If you have access to a computer package (e.g., R, Excel) that will do these tests, fine, otherwise, calculate the test statistics by hand using the formulae above.

Chromosomal Data. The variation that is exhibited in the chromosomal complement within and among species has also proven to be an important tool for exploring species limits. The chromosomal complement is usually displayed as a karyotype, and there are two attributes that

are readily discernible from a standard karyotype; the diploid number ($2N$), which is simply the number of chromosomes present, and the fundamental number (FN). The FN is the number of arms in the karyotype, excluding the sex chromosomes. Each chromosome has either one arm, if the centromere is terminal (such chromosomes are termed acrocentric chromosomes), or two arms if the centromere is not terminal (such chromosomes are termed biarmed chromosomes). The technique involves stopping the growth of cells at metaphase (the portion of the cell cycle during which the DNA is most condensed on the chromosomes and the chromosomes are therefore most easily visualized microscopically), disrupting the cells, fixing the disrupted cells onto a microscope slide, and staining the DNA contained on the chromosomes. Slides are then examined under a microscope and the $2N$ and FN are counted. Many species are differentiated from close relatives by differences in these attributes.



Karyotype of *Mastomys natalensis*, a murid rodent from Africa (Britton-Davidian et al., 1995).

Allozyme Data. The application of protein electrophoresis to systematics began in the early 1970's. This technique assesses genetic variation in the form of charge-changing amino-acid substitutions and has uncovered otherwise cryptic species. The technique has been immensely important in assessing the geographic structure of genetic variation in a number of mammalian species. Typically, allele frequencies are calculated for a number of populations, and these are used both to estimate genetic distances among those populations and to assess the degree of population differentiation, (i.e. though F -statistics).

DNA Data. More recently, the advances in biotechnology have resulted in the ability to assess genetic variation more directly, through restriction site analyses and DNA sequencing. We'll discuss examples during lecture later in the semester.

III. Phylogenetic Analyses.

Essentially the same types of data that are used in exploration of species limits are used in the estimation of phylogenies, tree-like graphs that represent ancestor-descendent relationships.

A. Types of Analysis.

There are several approaches to phylogenetic inference, and there is considerable debate (often very contentious) as to the relative merits of each type of data. At opposite extremes, there are those who eschew all molecular data as irrelevant, and those who view molecular data as the only legitimate source of phylogenetic data. In addition to arguments regarding the relative merits of different data types, the method of data analysis is the subject of often vehement debates that sometimes degenerate into personal attacks.

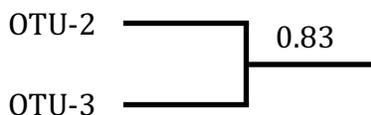
Regarding phylogenetic methodology, there are three classical schools of thought, and one newly emerging school. The oldest approach is called **Evolutionary Taxonomy**, and practitioners of this approach hold that the only way to understand the evolutionary history of a particular group is to spend a lifetime studying that group. This approach has been criticized as having no objective methodology and therefore no reproducibility. A second school of thought, **Phenetics**, holds that there is no way to infer ancestor-descendent relationships, and we therefore should group organisms on the basis of overall similarity. All data are considered relevant in estimating overall similarity and data are often summarized in a single index. The morphometrics techniques above are examples of phenetic methods. A third school of thought, **Cladistics**, holds that overall similarity is a poor criterion for grouping taxa, and that common ancestry actually can in fact be inferred. Under this view, derived characters that are uniquely shared by a group are *prima facie* evidence that those taxa share a common ancestor not shared by other taxa. That is, those taxa form a clade (hence the name). The only method that is permissible to a cladist is **parsimony analysis**. Under this methodology, the tree that requires the fewest number of evolutionary steps has the most explanatory power, and therefore is favored. With the application of DNA sequencing to phylogenetics, a new school is emerging. This new school has taken the name **Statistical Phylogenetics**, and **maximum likelihood** is the method of choice. This school views phylogenetic inference as a statistical endeavor. The idea is that, given that there is one true phylogeny, the goal phylogenetic inference is to estimate that phylogeny. Like all statistical estimates, a phylogeny is estimated with uncertainty and statistical phylogeneticists are interested in quantifying that uncertainty. The most acrimonious debates in systematics are currently between cladists and statistical phylogeneticists.

B. Examples of Two Simple Methods.

UPGMA - Unweighted Pair-Group Method using Arithmetic means (UPGMA) is perhaps the simplest method of inferring a dendrogram. This method is actually a phenetic method, and therefore was not intended by its developers (Sneath and Sokal) as producing phylogenies. Basically, it works from a matrix, the cells of which contain an index of similarity (e.g. Nei's genetic identity). Data from Marinkovic et al. (1978) from several species of *Drosophila*.

	OTU-1	OTU-2	OTU-3	OTU-4	OTU-5
OTU-1	-----	0.48	0.36	0.35	0.27
OTU-2		-----	0.83	0.12	0.03
OTU-3			-----	0.05	0.01
OTU-4				-----	0.53
OTU-5					-----

1. The first two OTU's (Operational Taxonomic Unit) to be clustered are the most similar (those with the highest genetic similarity). In this example, it is OTU-2 and OTU-3.

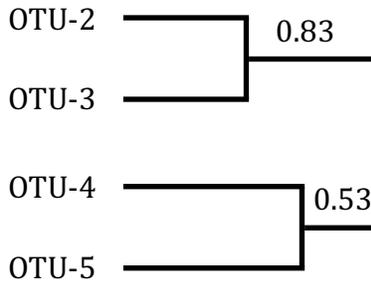


2. The OTU's are now combined as a single OTU-2/3 and a new matrix of similarities is calculated. The similarity of OTU-2/3 to any other OTU is the mean of the similarity between OTU-2 to that OTU and OTU-3 to that OTU. Thus, the similarity of OTU-1 to OTU-2/3 is: $(0.48 + 0.36)/2 = 0.42$

Thus, the new matrix is:

	OTU-1	OTU-2/3	OTU-4	OTU-5
OTU-1	-----	0.42	0.35	0.27
OTU-2/3		-----	0.09	0.02
OTU-4			-----	0.53
OTU-5				-----

3. The next most similar pair of OTU's is joined in the dendrogram. These are OTU's 4 & 5.

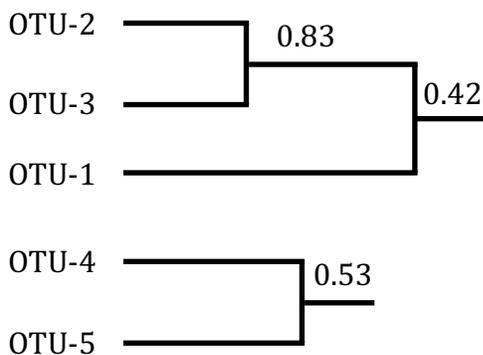


4. These now form OTU-4/5, and the matrix is again recalculate as before. Again, the similarity between a composite OTU and another is the mean of all the members of the composite in the **original** matrix to that OTU. Thus, the similarity of OTU-2/3 to OTU-4/5 is the average of the similarities between 2 & 4, 2 & 5, 3 & 4, and 3 & 5.

So at this point, the matrix becomes:

	OTU-1	OUT-2/3	OTU-4/5
OTU-1	-----	0.42	0.31
OTU-2/3		-----	0.05
OTU-4/5			-----

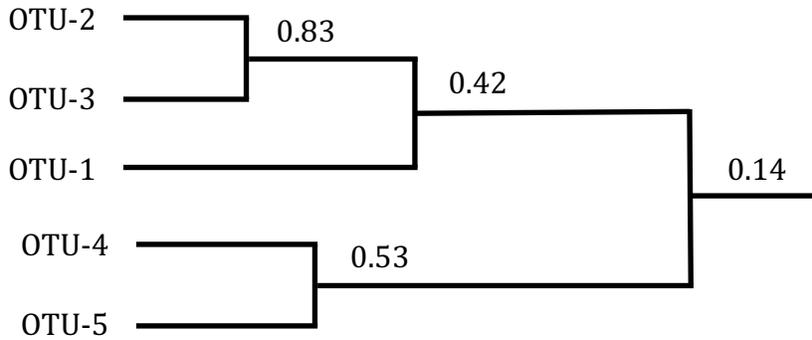
5. The calculations are continued in a cyclical fashion. OTU-1 is one combined with OTU-2/3.



Now, there is only one cell left in the matrix, the similarity between OTU-1/2/3 and OTU-4/5.

Again, the similarity value for this level in the tree is calculated as the average of the following pairwise similarity values from the original matrix: 1 & 4, 1 & 5, 2 & 4, 2 & 5, 3 & 4, and 3 & 5.

So the two clusters of taxa are joined at a similarity value of 0.14.



Exercise 4. Construct a phenogram from the following matrix of genetic identities using the UPGMA approach outlined above.

	OTU-1	OTU-2	OTU-3	OTU-4	OTU-5
OTU-1	-----	0.47	0.41	0.50	0.22
OTU-2		-----	0.76	0.56	0.15
OTU-3			-----	0.50	0.15
OTU-4				-----	0.33
OTU-5					-----

Parsimony Analysis. The goal of parsimony is to build a cladogram, a branching diagram in which the branching pattern is based on the inferred historical connections among the taxa. These inferences are based on shared derived (synapomorphic) characters. (Shared primitive, or plesiomorphic, characters have no significance).

Some definitions:

Homoplasy: A character found in two or more species that was not present in the common ancestor. Homoplasies can arise by convergent evolution.

Homology: The same character found in different groups is homologous if that character has been inherited from a common ancestor.

Ingroup: The set of taxa whose relationships are to be inferred.

Outgroup: The reference taxa (or taxon). The ingroup must be monophyletic relative to the outgroup.

Because cladistics relies on shared derived characters as evidence of common ancestry, many (strict) cladists insist on determining primitive versus derived character states prior to analysis. A character is an attribute of a group, such as the zygomatic arch, and the character state is the manifestation of that character. Character states for the zygomatic arch might be complete versus incomplete. Most often, cladists differentiate between primitive and derived character states (that is they polarize characters) by comparing states between the ingroup and the outgroup. Character states shared between the outgroup and any ingroup taxon are determined to be primitive (plesiomorphic) and the alternative states for that character are considered to be derived (apomorphic). Plesiomorphies do not imply exclusive common ancestry, but shared apomorphies (synapomorphies) do.

Consider the following data set. The author (Bugge 1978) examined cephalic artery characteristics among some carnivorans.

Taxon	Anastomoses					X	Y	MM
	a ₁	a ₂	a ₃	a ₄	a ₅			
<i>Selenarctos</i> (a bear, outgroup)	+	+	+	-	-	-	-	1
<i>Martes</i> (martin, fisher)	+	+	+	+	-	+	+	2
<i>Mustela</i> (weasels)	+	+	+	-	-	+	-	1
<i>Meles</i> (Honey badger)	+	+	+	+	-	+	+	1
<i>Lutra</i> (Eurasian otters)	+	+	+	-	-	+	+	2

1. Under classical cladistic analyses, the first step is to polarize characters by reference to the outgroup.

The following characters are uninformative because they are invariant: a₁ a₂ a₃ & a₅.

Character X is also uninformative in that no primitive (plesiomorphic) state can be determined; no state is shared between the ingroup and the outgroup.

a₄ – absence (-) is plesiomorphic, presence (+) is apomorphic.

Y – absence is plesiomorphic, presence is apomorphic.

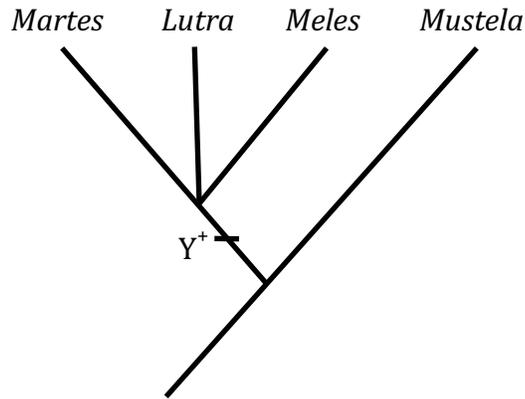
MM – 1 is plesiomorphic, 2 is apomorphic.

2. The next step is to construct a cladogram based on synapomorphic character states.

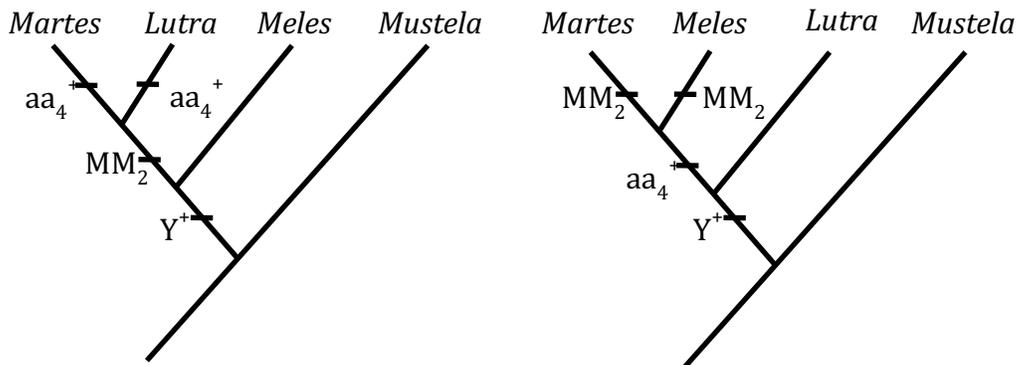
Technically, we should consider all possible trees that relate these taxa (for 5 ingroup taxa

there are 105 possible unique trees to consider) and we would favor the one(s) that requires the fewest homoplasies; for a data set of this size, we can examine it manually.

Note that character Y is uniquely shared by *Martes*, *Meles*, and *Lutra*. These are the only taxa that share the derived state for this character. Because there are no characters that contradict this grouping, we may begin building the cladogram this way:



Now there are only two characters left that are potentially informative. Unfortunately, they are conflicting; MM_2 unites *Martes* and *Lutra*, to the exclusion of *Meles*, and aa_4 unites *Martes* and *Meles* to the exclusion of *Lutra*. Such character conflicts are common. In this case there will be two equally parsimonious trees (each requires 4 steps), and no way to choose between them:



Thus, each tree explains the data equally well, and neither can be preferred.

Exercise 5. Construct a cladogram from the following data set. Include decisions on character polarity and intermediate steps.

Taxa	Characters						
	1	2	3	4	5	6	7
<i>Didelphis</i> (outgroup)	-	-	-	-	-	-	-
<i>Elephas</i> (Elephant)	+	-	-	-	+	+	+
<i>Trichechus</i> (Manatee)	-	+	-	-	+	+	+
<i>Procavia</i> (Hyrax)	-	-	+	-	-	+	+
<i>Bradypus</i> (sloth)	-	-	-	+	-	-	+

Characters

- 1 - Graviportal limbs.
- 2 - Front limbs modified for swimming.
- 3 - Plantar surfaces with concave, suction generating surface.
- 4 - Xenarthrous processes on vertebrae.
- 5 - Cheek teeth replaced from back to front.
- 6 - Arose in Africa in early Tertiary.
- 7 - Chorioallantoic placenta.

Literature Cited

- Britten-Davidian, J., J. Catalan, L. Granjon, and J.-M. Duplantier. 1995. Chromosomal phylogeny and evolution in the genus *Mastomys* (Mammalia: Rodentia). *J. Mammal.*, 76:248-262.
- Bugge, J. 1978. The cephalic arterial system in carnivores, with special reference to the systematic classification. *Acta. Anat.*, 101:45-61.
- Genoways, H. H. and J. R. Choate. 1972. A multivariate analysis of systematic relationships among populations of the short-tailed shrew (Genus *Blarina*) in Nebraska. *Syst. Zool.*, 21:106-116.
- Marinkovic, D. F. J. Ayala, and M. Andjelkovic. 1978. Genetic polymorphism and phylogeny of *Drosophila subobscura*. *Evolution*, 32:164-173.
- Simmons, N. B., K. L. Seymour, J. Habersetzer, and G. F. Gunnell. 2008. Primitive Early Eocene bat from Wyoming and the evolution of flight and echolocation. *Nature*, 451:818-821.