Lecture 3 – Characters: Homology, Morphology

I. Introduction – Nearly all methods of phylogenetic analysis rely on **characters** as the source of data.

A. Character variation is coded into a **character-by-taxon matrix**.

	Taxon A	Taxon B	Taxon C	Taxon D
Character 1	0	0	1	1
Character 2	1	1	0	1
Character 3	0	0	0	1
Character 4	1	1	1	0
Character 5	1	1	0	0
Character 6	1	2	2	0
Character 7	G	A	C	C
Character 8	*	*	&	&
Character 9	15	10	10	25

Even in modern studies that use genetic distances to estimate phylogeny, those distances most often summarize character data. The **exceptions** to this are some methods that estimate genetic distances indirectly. In addition, there's a fair amount of interest in genomic distances that are calculated using an alignment-free approach rather than from discrete, putatively homologous characters.

B. Definition – A character is a trait, feature, or attribute of an organism; these are typically discreet.

1. In the **most common usage**, a **character is composed of a number of states**. Character states are the manifestation of the character in particular taxon.

   For example, **Eye Color** is a character and its states are Blue, Brown, & Green

2. Another usage (that many cladists have employed) of the term character is equivalent to the above definition of character state.

   In this usage, there is a single transformation series (TS), and the various manifestations of the TS are the characters. So, Eye Color would be the TS and Blue, Brown & Green are the characters. We will not use the term in this manner, although we need to be aware of this usage.

Assumptions are that characters are homologous and independent from each other.

3. Selection of Characters - In principle, any heritable independent character may be used in phylogeny analysis. Efforts are usually made to focus only on characters that exhibit levels of variation that are appropriate in the context of the study at hand.

For example, Eye Color wouldn't be considered an appropriate character for a study of mammalian relationships because it exhibits **far too much variation**. Conversely, the number of heart chambers wouldn't be an appropriate character in the same study, because it is invariant across mammals and, therefore, uninformative.

Therefore, the variation shown by a character must match the **level of universality**.

Furthermore, the status of a state as primitive versus derived is also a dependent on the level of universality. Hair is primitive if we're working at the level of mammalian orders. It's derived and unites mammals if we're working at the level of Vertebrata.

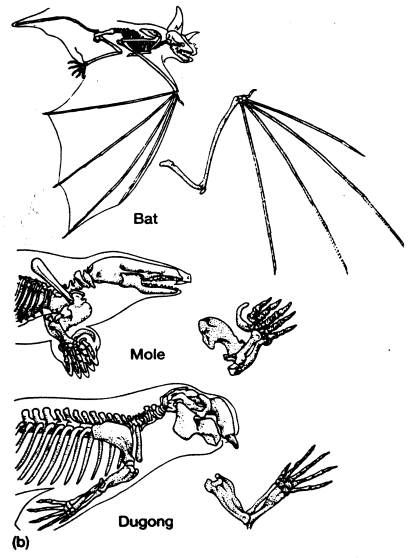
The selection of characters on which to focus prior to collection of data is an extremely important consideration. It also represents an enormous source of subjectivity in phylogenetic analysis that is usually not acknowledged.

Again, characters used for phylogenetics must be **homologous**.

II. Homology – The concept of homology is central to phylogenetic inference (and indeed fundamental to evolutionary biology). Only homologous characters can be compared (although there is a growing dissenting view).

A. Original Definition – Although it was used earlier, the concept attained prominence in ~1848 when Richard Owen defined it as "The same organ in different organisms in all its varied forms."

Richard Owen was a very influential vertebrate morphologist of Darwin's day, and one of Darwin's biggest critics (although not a creationist). As a comparative anatomist, he recognized amazing skeletal similarities across a broad range of vertebrates, exemplified by front limbs below.



B. A more appropriate definition of homology would be something like this.

"Possession by two or more species of a trait inherited from a common ancestor, either with or without modification."

This contrasts with **homoplasy**, which is similarity due to reversal or convergence.

C. Criteria for establishing homology in choosing morphological characters.

Prior to analysis, **hypotheses of character homology must be established**. We'll talk about molecular characters later; for now, I want to focus on morphological characters and exemplify the extreme care that the best morphological cladists exhibit.

Also, I want you all to be able to explain to a random lay-person why it's not circular that evolution is central to the definition of homology and yet homologous structures provide such compelling evidence for evolution.

Certainly, one of the fundamental texts on the issue in morphological phylogenetics is Wiley's 1981 book *Phylogenetics: The Theory and Practice of Phylogenetic Systematics*. We have this in our library. (The second edition –Wiley and Lieberman – was published in 2011).

1. Morphological Criteria – Any of these can be used to **propose** hypotheses of homology.

a. Similarity of position – This includes overall position (i.e., humerus occurs in the front limb) It also includes position relative to other structures (scapula, ulna)

b. Special similarity – This is a rather vague criterion, but it refers to either similarity on a finer scale (like scanning EM), or perhaps in developmental pathways.

c. Continuance through intermediate forms (either intermediate fossils or extant forms that illustrate intermediate conditions).

2. Phylogenetic Criterion – Once we've conducted a phylogenetic analysis, including this character in our analysis, we have **tested** the hypothesis of homology further.

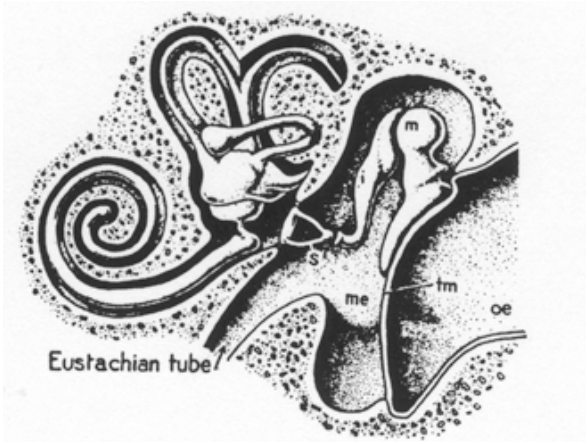
Are the putatively identical character states we've identified in generating our character by taxon matrix indeed synapomorphies that unite the groups that possess them?

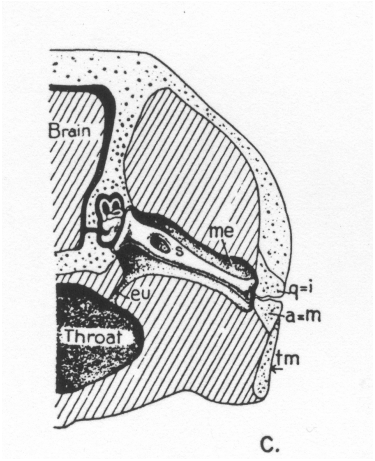
Character congruence either supports or refutes these hypothesized homologies.

Example – Mammalian ear ossicles.

There are three bones in the middle ears of mammals: malleus, incus, and stapes. These function in hearing and are encased in the middle ear.

These function to transmit sound, as manifest by vibrations of the tympanum, to the inner ear, where the sensory cells for hearing reside.





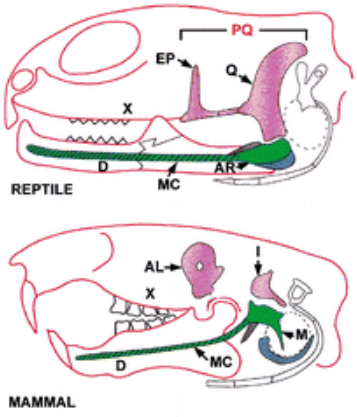
Other vertebrates have a single ear ossicle, the stapes.

In many non-mammalian vertebrate groups, the stapes articulates with the upper jawbone, the quadrate. Thus, based on similarity of position, and common connections, we might hypothesize that the quadrate is homologous to the mammalian incus; mammals lack a quadrate on the upper jaw.

Hypothesis – quadrate is homologous to the incus (derived by similarity of position) and articular is homologous to malleus.

We can test this hypothesis using the other criteria.

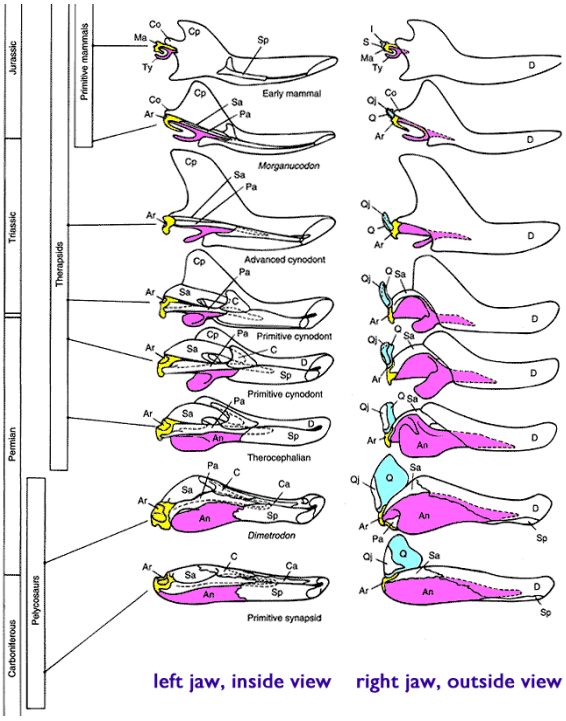
Special similarity. If we look at the development, we see that the incus in mammals and the quadrate in non-mammalian vertebrates share a common developmental precursor, the palatoquadrate cartilage. Also, both articular and malleus ossify from Meckel's cartilage.



So, the homology hypotheses are corroborated by development, or the criterion of special similarity.

We can use the third criterion, **continuance through intermediate forms** to test the hypothesis still further.

As it turns out, there is fabulous series of transitional fossils (supported by a myriad of fossil intermediates) that document the transition from one morphology to the next.



Thus**, transitional fossils corroborated the homology hypothesis further by demonstrating continuance through intermediate forms**.

In addition, we have fossils that have both ancestral and derived jaw joints.

Thus, one could feel extremely confident using the homology of the mammalian ear ossicles with the jaw bones of non-mammalian vertebrates as a phylogenetic character.

This represents the appeal (i.e., "logical clarity") of the Hennigian approach to defining monophyletic groups by discovery of synapomorphies.

It also, critically, illustrates that statements of homology represent very well-tested and corroborated hypotheses.

III. Coding: We need to transform our putative hypotheses about character-state variation into our character-by-taxon matrix that we can then subject to phylogenetic analysis.

This is of fundamental importance in phylogenetics, because the matrix is actually what is analyzed.

We need to make decisions regarding whether to accept multi-state characters, and whether to order the character states or leave them unordered.

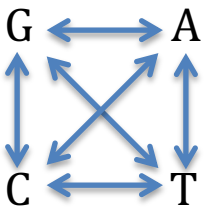
Simplest are qualitative binary characters. For example, is the first upper premolar present or absent?

0 <---> 1

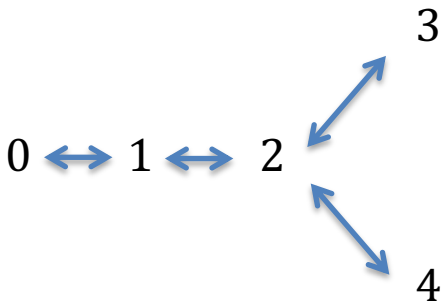
Also simple are **linear, ordered muiti-state characters**.

0 <----> 1 <----> 2 <----> 3

Unordered multistate characters are similarly easy to deal with. These are very common, e.g., molecular sequence data.



More difficult to deal with are **non-linear, ordered, multistate characters**. These often become necessary when dealing with complex morphological characters. These require a complex character-state tree that must be reflected in the character coding.



Frequently, these are decomposed into a series of binary characters:

0 → 0 0 0 0
1 → 1 0 0 0
2 → 1 1 0 0
3 → 1 1 1 0
4 → 1 1 0 1

These all represent decisions that have to be made in order to convert your understanding of character variation into a character-by-taxon matrix that can be used to infer phylogenies.

IV. Polarity – Recall that the Hennigian approach focuses only on synapomorphies, not on shared primitive characters (Symplesiomorphies).

This led most early cladists to conduct an assessment of character polarity prior to erecting a matrix.

There was a period in which a number of criteria were proposed to establish polarity (i.e., determine the primitive state), including common=primitive (not useful), ontogeny (following von Baer's Law), and evidence from fossils (as we've already exemplified).

By far, the most important criterion is **outgroup analysis**.

An **outgroup** is a taxon (or taxa) included in the analysis that is thought, **based on independent data**, to be more distantly related to the focal taxa, which are called the **ingroup**.

Another way of saying this is that the ingroup must be monophyletic relative to the outgroup.

A character state that is **shared between the outgroup and at least one member of the ingroup is deemed to be the primitive state** for that character.

For several years, the only acceptable approach was to polarize characters *a priori*, that is, before a phylogenetic analysis.

It became apparent during the early 1980's though that simply including an outgroup(s) in an analysis would allow *a posteriori* polarity determination, and that these are equivalent (Maddison et al., 1984). We'll talk more about this when we discuss rooted vs. unrooted trees.

Some references for the material covered here are:

Wiley, E. O., and B. S. Lieberman. 2011. *Phylogenetics: The Theory and Practice of Phylogenetic Systematics*. 2nd ed. John Wiley and Sons, Inc., NY.

Wiley, E. O., D. Siegel-Causey,, D. R. Brooks, and V. Funk. 1991. The Compleat Cladist: A Primer of Phylogenetic Procedures. University of Kansas Museum of Natural History, Specieal Publication 19. This is out of print, but a pdf is available from http://www.nhm.ukans.edu/cc.html.

Maddison, W. P., M.J. Donoghue, and D. R. Maddison. 1984. Outgroup analysis and parsimony. Systematic Zoology, 33:83-103.