

Genomic Evolution

Jutta Marzillier, PhD.

Lehigh University

April 12th 2011

Where do we come from, we humans and the
staggering variety of other organisms with
which we share our planet?

Study conducted by the PEW Research Center (published in 2005)

“Life on Earth has....

....existed in its present form since the beginning of time”

....evolved over time”

.... don’t know”

Study conducted by the PEW Research Center (published in 2005)

“Life on Earth has....

....existed in its present form since the beginning of time” (42%)

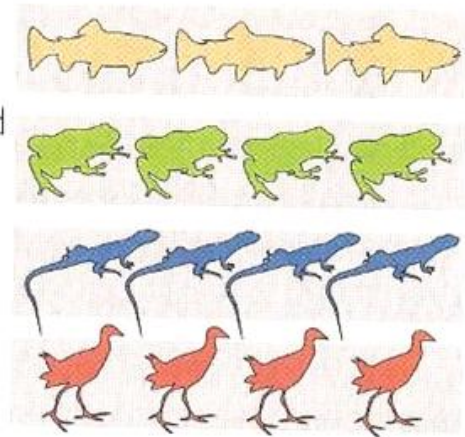
....evolved over time” (48%)

.... don’t know” (10%)

Two views of History of Life

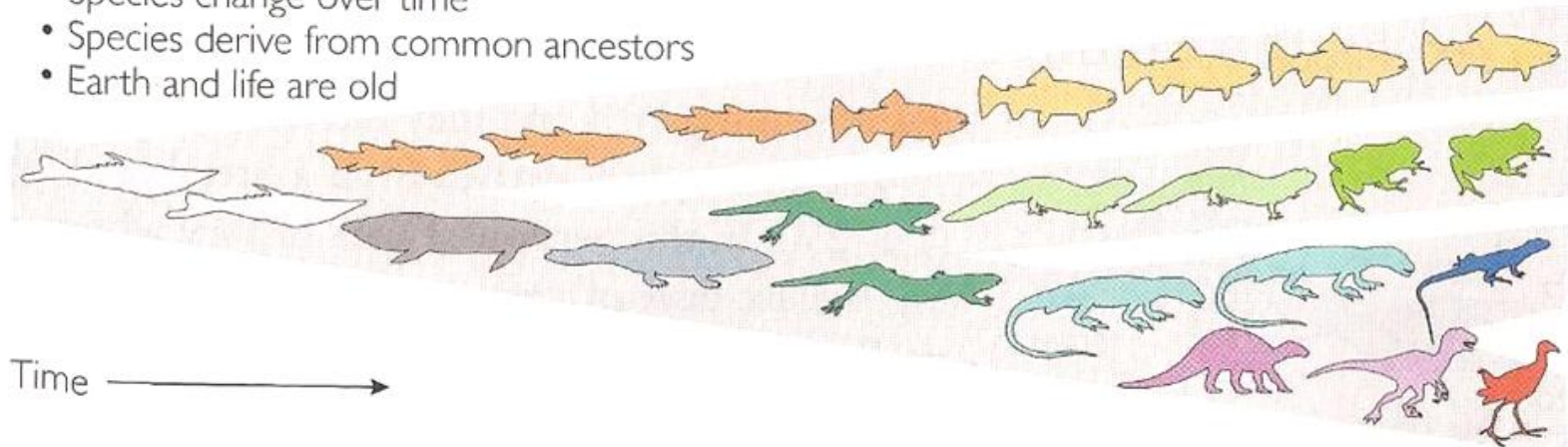
Special Creation

- Species do not change
- Each species separately created
- Earth and life are young

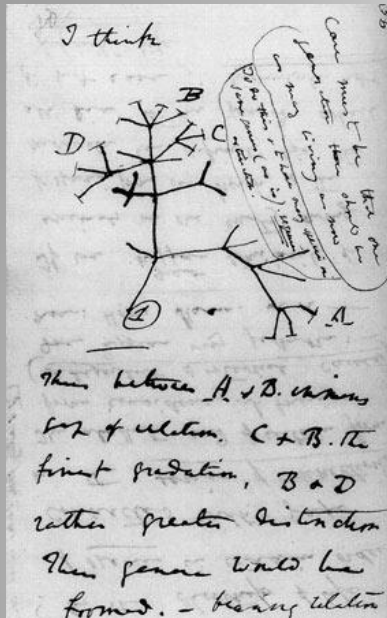


Descent with Modification

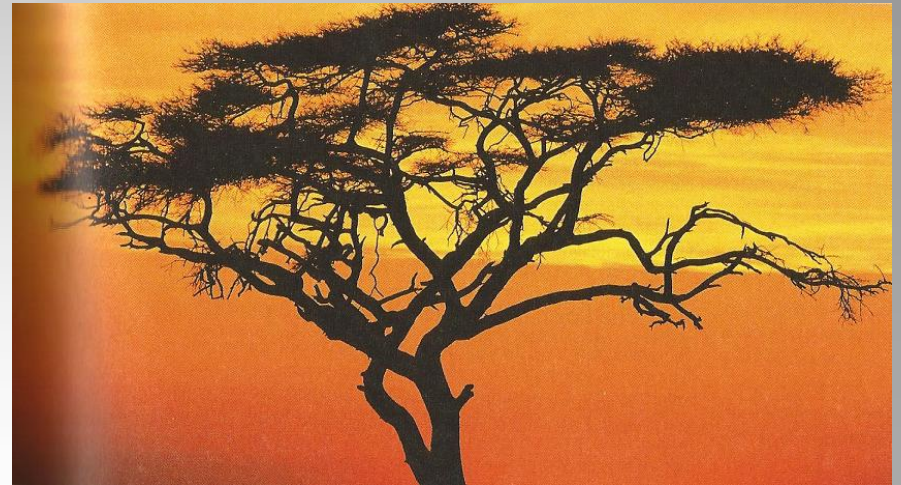
- Species change over time
- Species derive from common ancestors
- Earth and life are old



Darwin's 'Tree of Life'



The first-known sketch by Charles Darwin of an evolutionary tree describing the relationships among groups of organisms



Homology

shared characteristics that have been inherited from a common ancestor

To represent the idea that all living things, despite their diversity, share a common ancestor, Darwin used the metaphor of a "tree of life".

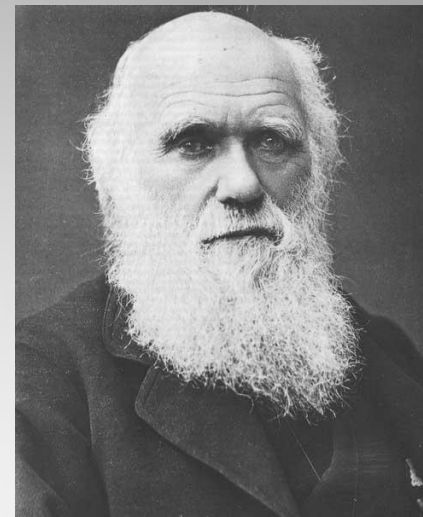
p. 484....."I should infer.....that probably all the organic beings which have ever lived on this earth have descended from some one primordial form, into which life was first breathed."

Evidence of Change through Time

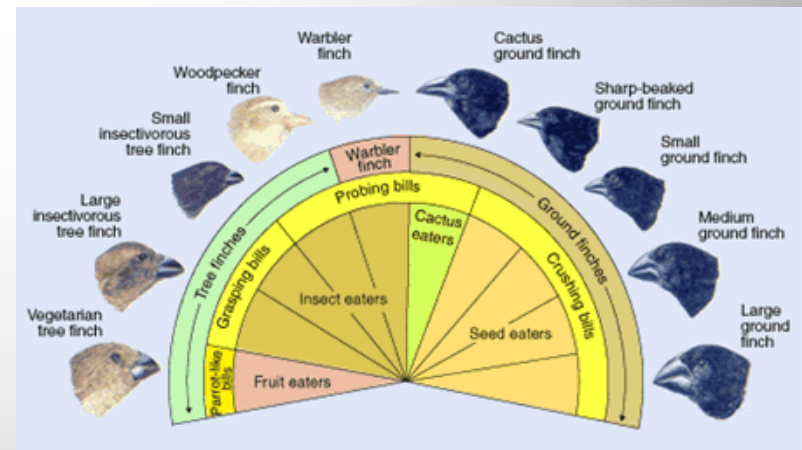
Populations of organisms change through time

Evidence from living species as well extinct forms:
Direct Observation

- Galapagos finches
- Peppered moth
- Vestigial Organs
- Fossils

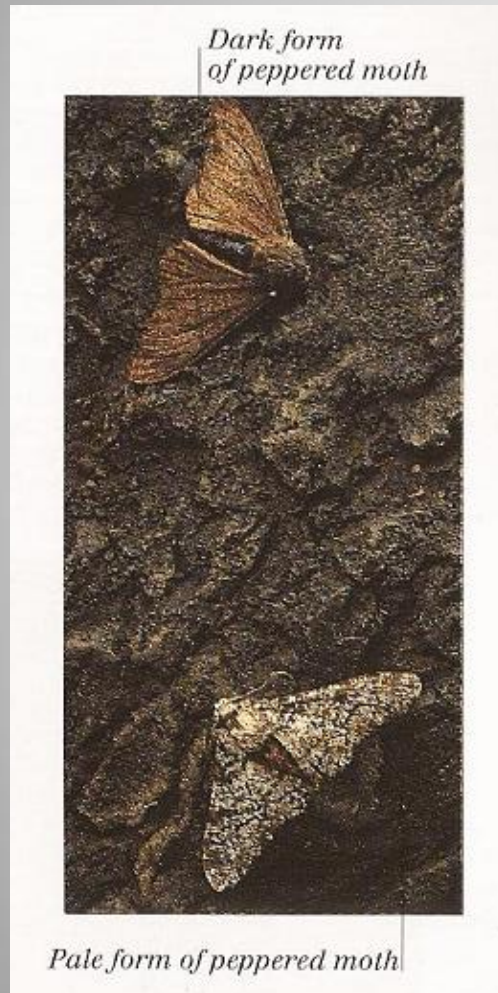


Charles Darwin (1809-1882)



Darwin concluded that all Galapagos finches had evolved from a single, ancestral South American species.

Natural Selection



Is the principle by which each slight variation, if useful, is preserved (Charles Darwin).

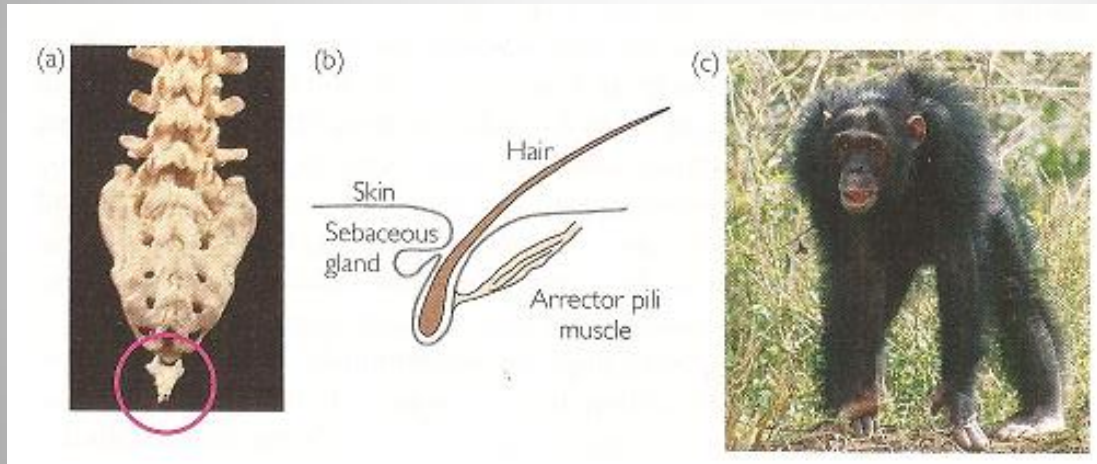
Pale peppered moth rests on lightly colored Lichen covered tree trunks

In mid of 19th century air pollution killed the Lichen blackened trees with soot

Dark forms of the moth who appeared as a Result of natural variation were now better Camouflaged

Human Vestigial Traits

Useless or rudimentary version of a body part that has an important function in other, closely allied, species



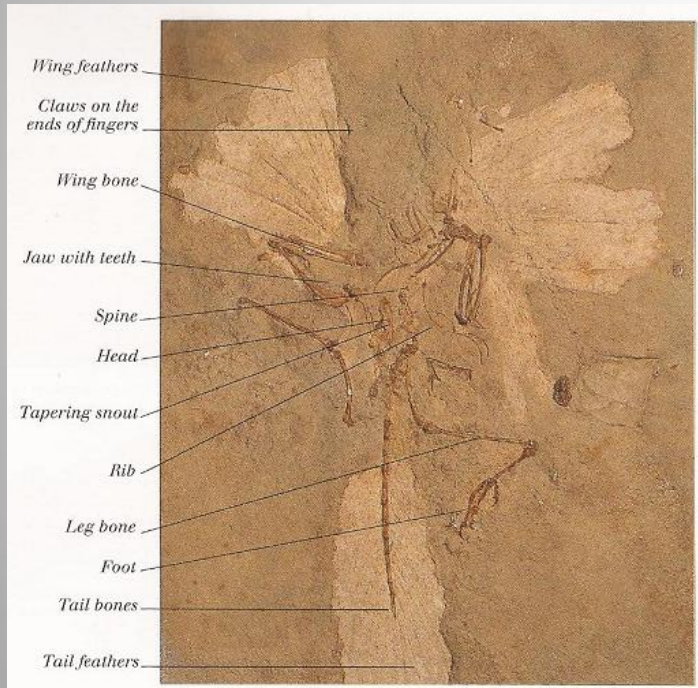
Rudimentary
tailbone

Arrector pili muscle at base
of each hair follicle

Implies that humans are descendants from ancestors with tails or fur, respectively

Fossil Records

Archaeopteryx



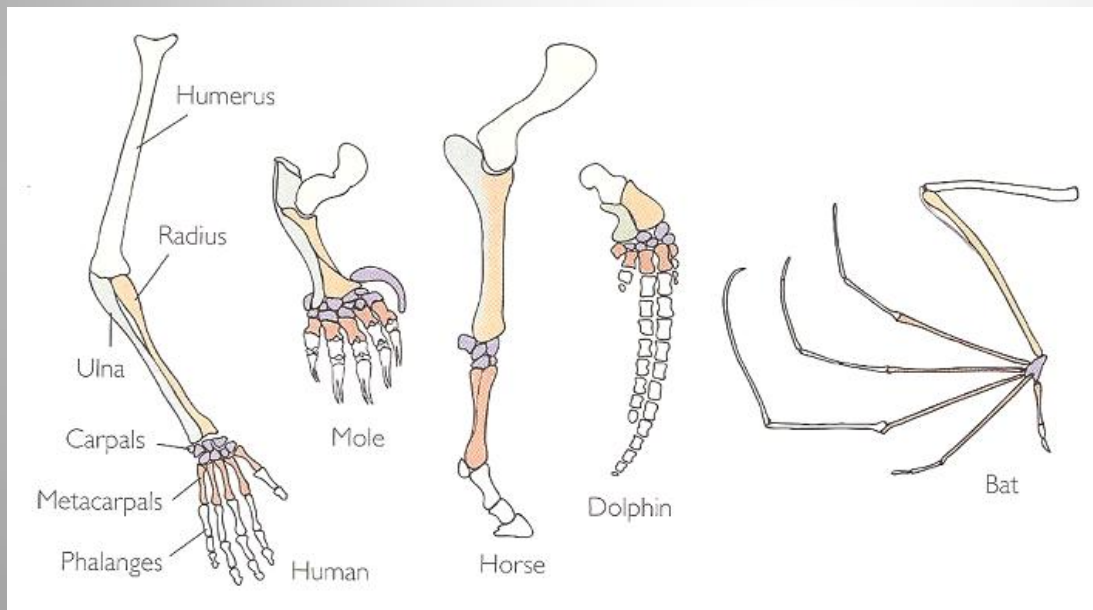
Provided a “missing link”

Long bony tail, jaws with teeth and claws on fingers resembled small dinosaurs called theropods

It had feathers and a forelimb adapted as a wing

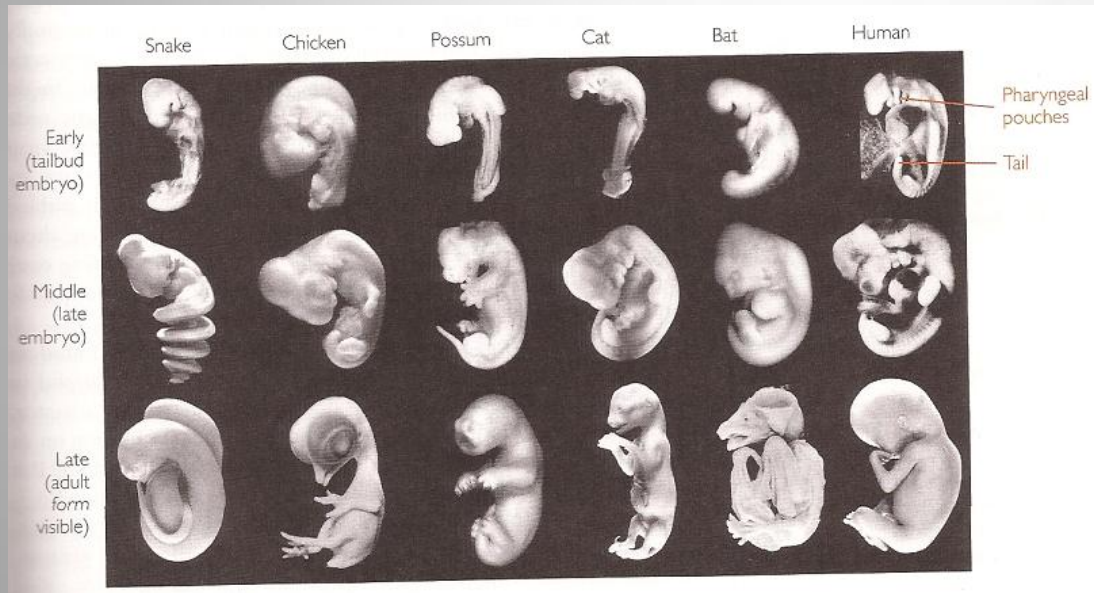
Structural Homology

Vertebrate forelimbs are used for different functions, but have the same sequence and alignments of bones



Developmental Homology

Embryos from different vertebrates show striking similarities early in development



Embryos are similar because all vertebrates evolved from the same common ancestor

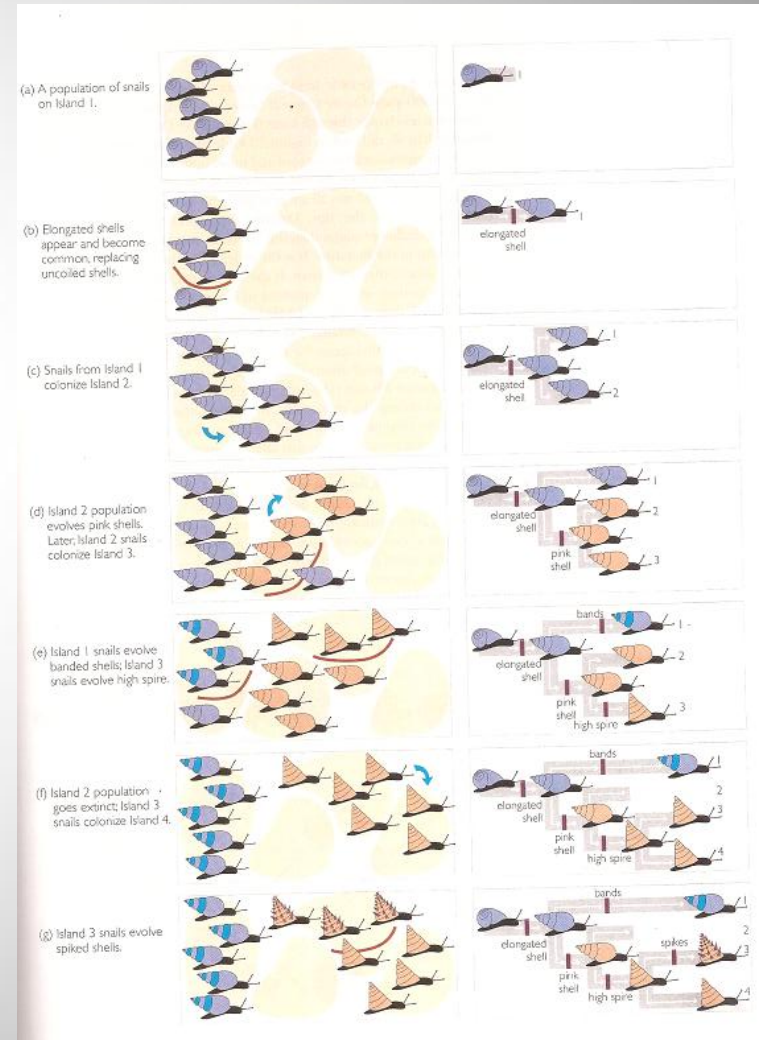
Evidence of Common Ancestry

Species are not independent, but are connected by descent from a shared ancestor

Phylogenetic trees are a visual representation of descent with modification from a common ancestor

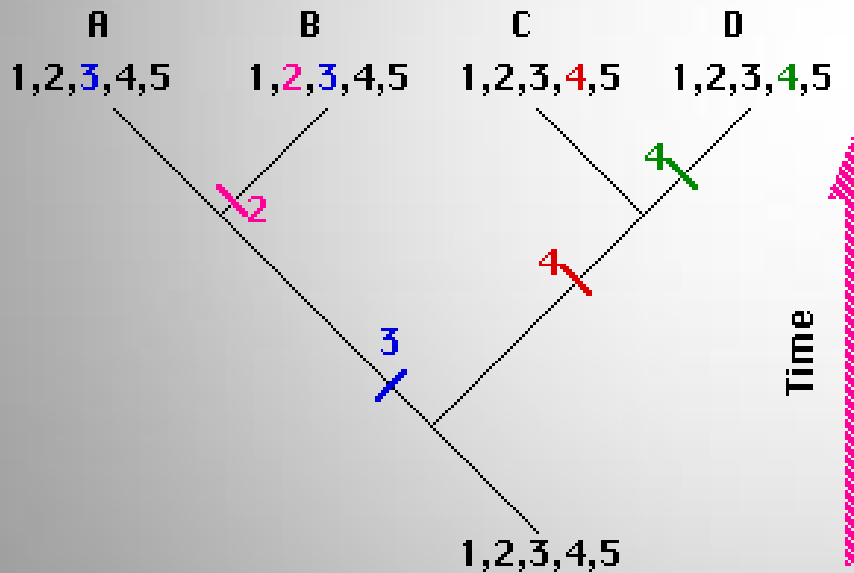
History of a population of imaginary snails

Growing evolutionary tree



Evolutionary Relationships of 4 hypothetical Species depicted in a Phylogenetic tree or Cladogram

- Depicts five ancestral (“original”) traits (**plesiomorphy**) (black # 1–5)
- Three speciation events occurred producing branches
- Several of ancestral traits evolved into a modified/“changed” or derived form (**apomorphy**) (colored numbers)



A **synapomorphy** is a shared trait found among two or more taxa and their most recent common ancestor, whose ancestor in turn does not possess the trait

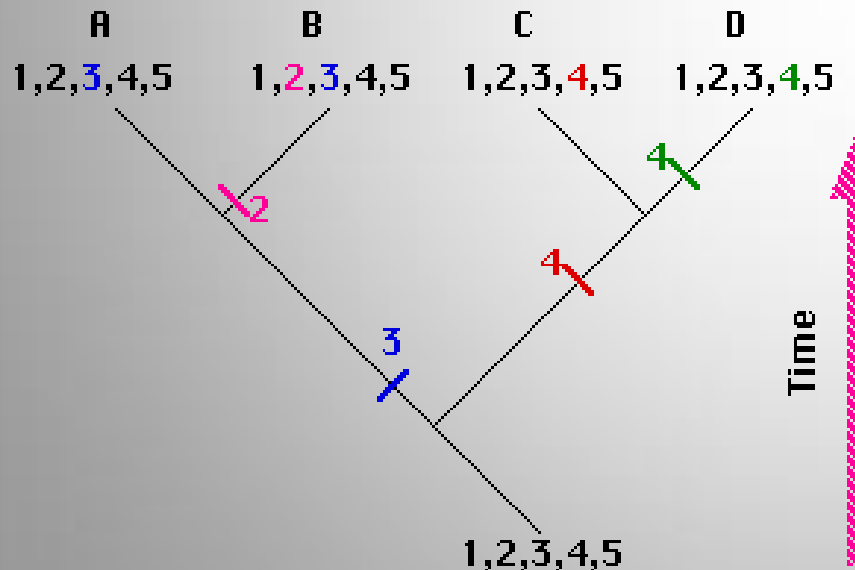
Here species A and B share the synapomorphic trait designated with a blue 3 .

Synapomorphies Identify Evolutionary Relationships

syn = in company with
apo = away from
morphe = shape.

Synapomorphies identify evolutionary branch points

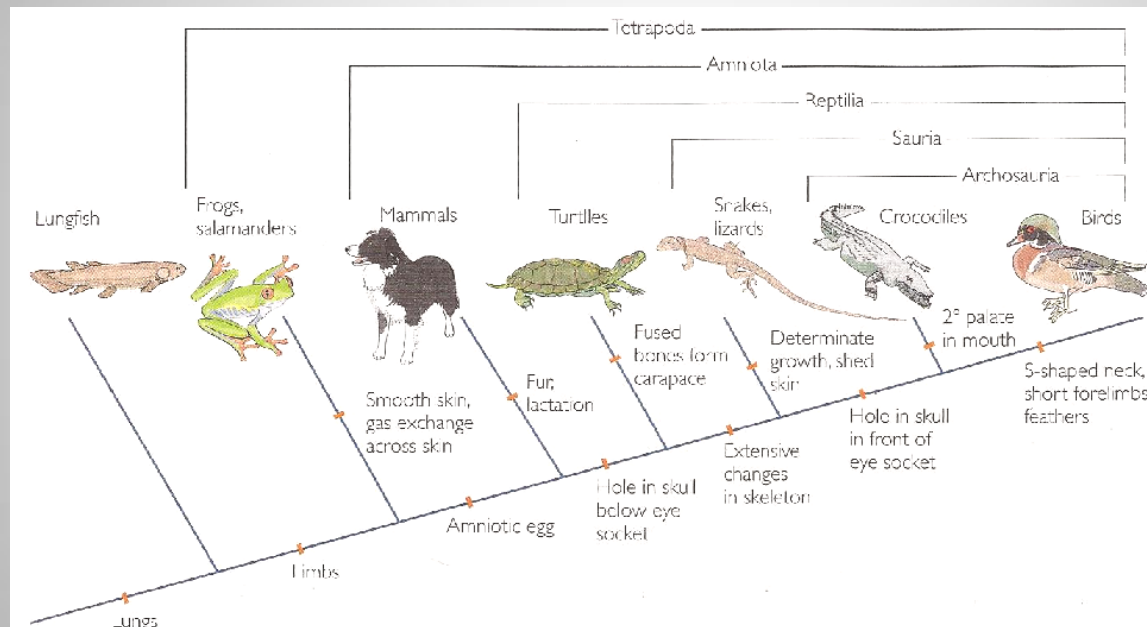
Synapomorphies create a nested hierarchy



Two species that recently separated from one another will share a number of unique features not found in a third species that branched off much earlier

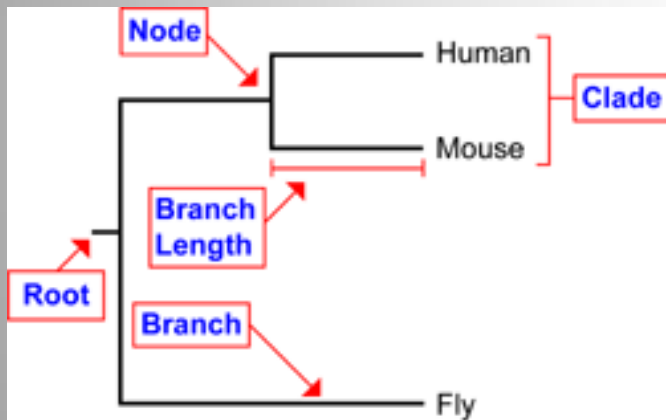
Phylogenetic relationships among species are determined through analysis of their shared similarities

Synapomorphies reveal the Relationships among Tetrapods



Traits that are labeled at each hash mark are synapomorphies shared by the descendant species above that point

Building of Phylogenetic Trees to Present Evolutionary Relationships



<http://www.ncbi.nlm.nih.gov/About/primer/phylo.html>

Node: represents a taxonomic unit (species, populations, individuals)

Branch: defines the relationship between the taxonomic units in terms of descent and ancestry

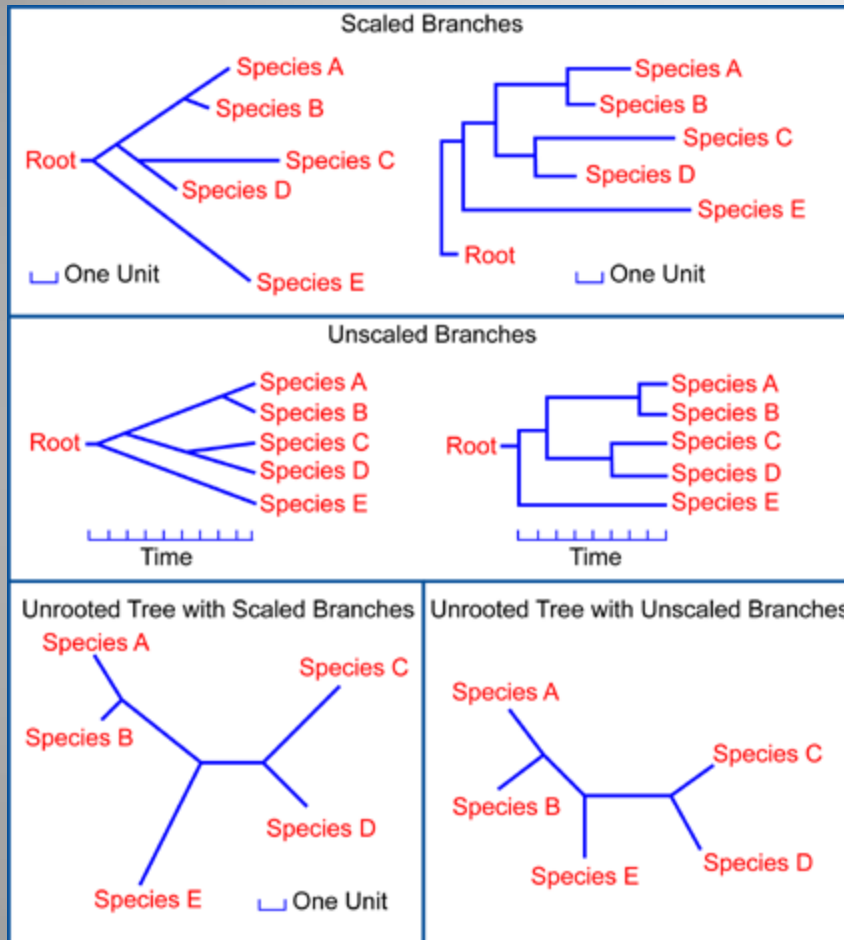
Branch length: represents number of changes

Scaled branches: trees calibrated to represent the passage of time

Root: the common ancestor of all taxa

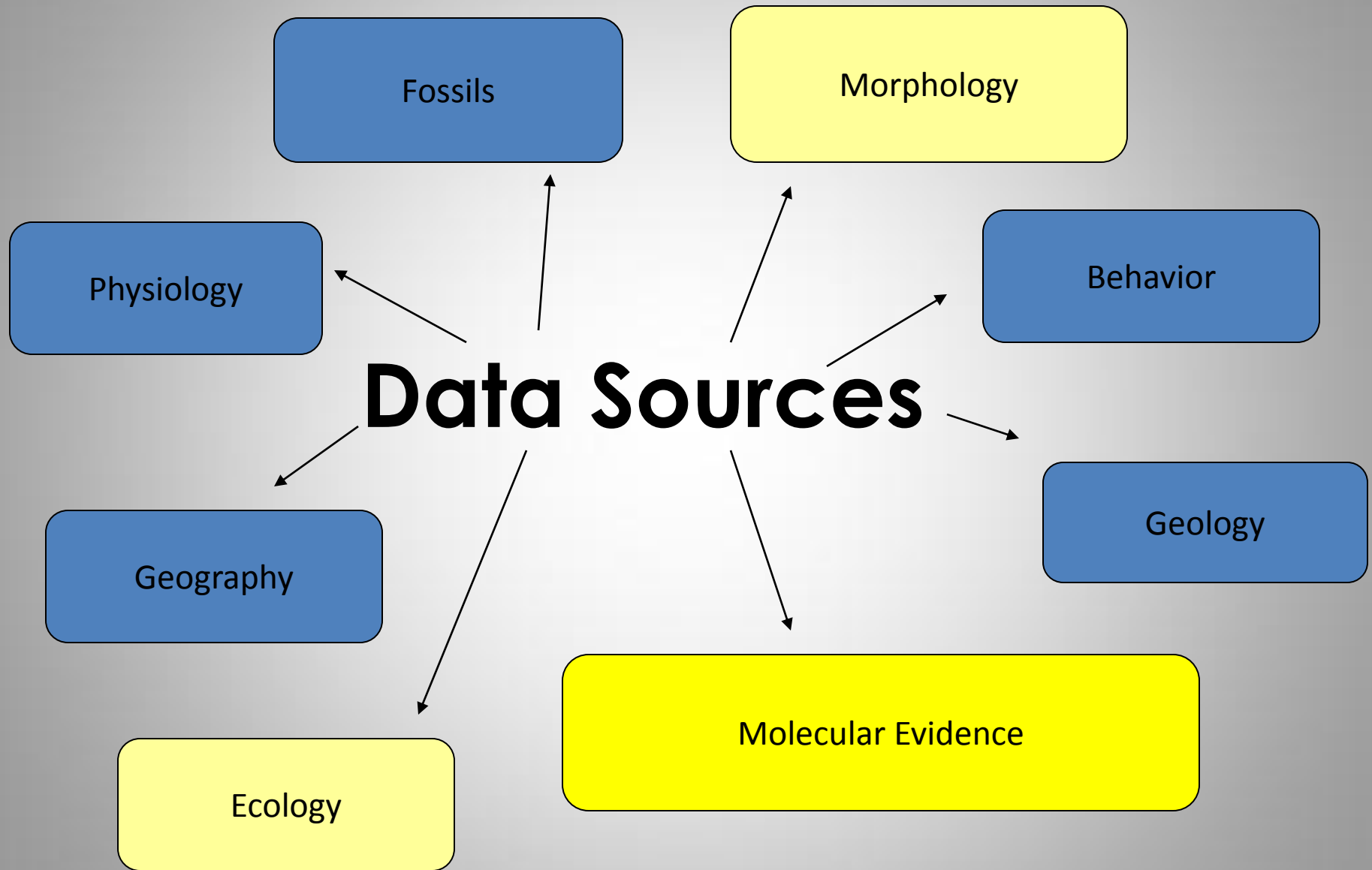
Clade: a group of two or more taxa or DNA sequences that includes both their common ancestor and all of their descendants

Possible Ways of Drawing a Phylogenetic Tree



Homology is the Guiding Principle behind most Biomedical Research

Researchers use study – or ‘model’ organisms based on the degree of homology required to study a particular process or disease



Understanding the Evolutionary Process on Molecular Level

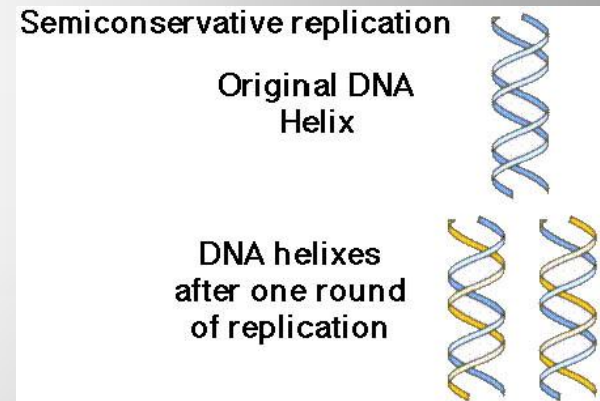
Genetic Variation: Changes in the Gene Pool

Genome: contains all of the biological information needed to construct and maintain an organism

Stored in form of DNA

Genetic information is passed on

First make copy of complete genome (DNA replication)



intranet.canacad.ac.jp

Mutations occur from errors in DNA replication or exposure to DNA damaging agents

Mutations

To have an effect on the next generation:

Must happen *in germ line cell (sperm or egg cell)*, not somatic cell

can be silent: - degenerated Genetic Code

- non-coding stretch of DNA

Mutations that have evolutionary effect:

Loss-of-Function or **Gain-of-Function** mutations

Important:

Randomness with which mutations can occur

Changes in the characteristics of an organism occur by chance and are not influenced by the environment

		second base in codon				
		T	C	A	G	
first base in codon	T	TTT Phe	TCT Ser	TAT Tyr	TGT Cys	T
		TTC Phe	TCC Ser	TAC Tyr	TGC Cys	C
		TTA Leu	TCA Ser	TAA stop	TGA stop	A
		TTG Leu	TCG Ser	TAG stop	TGG Trp	G
C		CTT Leu	CCT Pro	CAT His	CGT Arg	T
		CTC Leu	CCC Pro	CAC His	CGC Arg	C
		CTA Leu	CCA Pro	CAA Gln	CGA Arg	A
		CTG Leu	CCG Pro	CAG Gln	CGG Arg	G
A		ATT Ile	ACT Thr	AAT Asn	AGT Ser	T
		ATC Ile	ACC Thr	AAC Asn	AGC Ser	C
		ATA Ile	ACA Thr	AAA Lys	AGA Arg	A
		ATG Met	ACG Thr	AAG Lys	AGG Arg	G
G		GTT Val	GCT Ala	GAT Asp	GGT Gly	T
		GTC Val	GCC Ala	GAC Asp	GGC Gly	C
		GTA Val	GCA Ala	GAA Glu	GGA Gly	A
		GTG Val	GCG Ala	GAG Glu	GGG Gly	G

Evolution on Molecular Level

(a) Mutations can create synapomorphies

DNA sequences in descendants:

TGCTATT TGCTTTT TGCTTTT

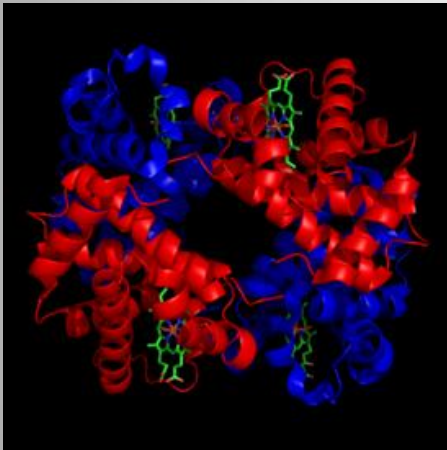


Changes in nucleotide sequence due to mutation

Change in the 5th position of this DNA sequence creates a *shared derived character* (**synapomorphy**) in the descendant populations

New species are assumed to arise primarily through **bifurcation**, or separation, of the ancestral lineage

Structure of Human Hemoglobin



the iron-containing oxygen-transport metalloprotein in the red blood cells of all vertebrates

Hetero-tetramer with two alpha and two beta subunits

Protein and DNA Alignment of the Human beta Globin Gene for Phylogenetic Tree Construction

Genomes evolve by gradual accumulation of mutation

Human		M	V	H	L	T	P	E	E	K	S	A	V				
Chimpanzee		M	V	H	L	T	P	E	E	K	S	A	V				
Mouse		M	V	H	L	T	▼	▼	E	K	S	A	V				
Donkey		M	V	H	L	T	▼	A	E	E	K	▼	V				
							A	E	E	K	S	L	V				
Human	5'	AACAGACACC	ATG	GTG	CAT	CTG	ACT	CCT	GAG	GAG	AAG	TCT	GCC	GTT	3'		
Chimpanzee	5'	AACAGACACC	ATG	GTG	CAC	CTG	ACT	CCT	GAG	GAG	AAG	TCT	GCC	GTT	3'		
Mouse	5'	AACAGACATC	ATG	GTG	CAC	CTG	ACT	▼	▼	GAT	GCT	GAG	AAG	TCT	GCT	GTC	3'
Donkey	5'	AACAGACACC	ATG	GTG	CAT	CTG	ACT	▼	GCT	GAA	GAG	AAG	AGT	CTT	GTC	3'	
			↑														
			1	2	3	4	5	6	7	8	9	10	11	12			

Human and chimpanzee DNA sequences almost identical , identical on Amino Acid level

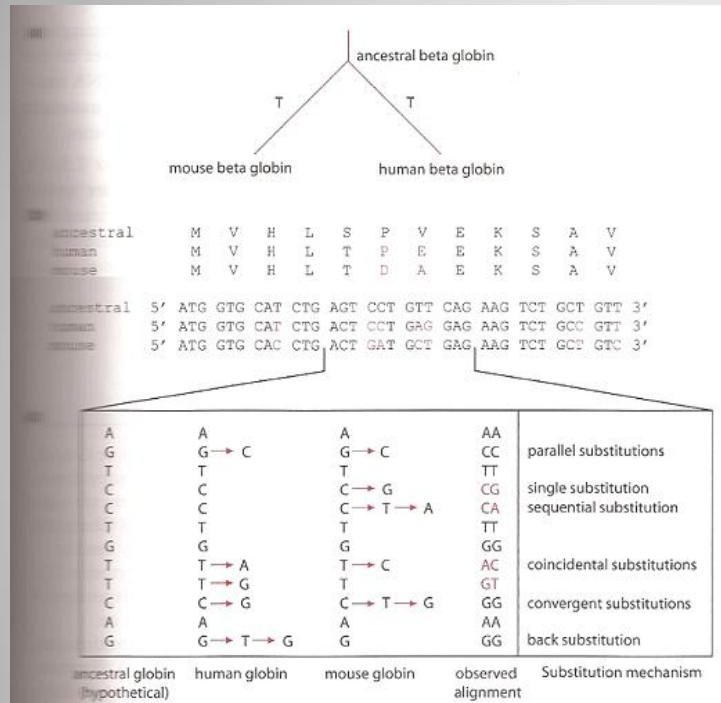
Variant amino acid residues at position 6, 7, and 11

Eight synonymous nucleotide changes that do not cause aa changes (red nucleotides)

Seven non- synonymous nucleotide changes that cause aa changes (red arrows)

Two genomes that derived in the recent past should have fewer differences than two genomes whose common ancestor is more ancient

Construction of Phylogenetic Relationships using DNA/ Protein Sequence



Species Tree: node represents a speciation event, e.g. last common ancestor of humans and mice is about 80 million years ago

Gene/Protein Tree: internal node represents the divergence of ancestral gene/protein into new genes/proteins with distinct sequences

Two observed mismatches at amino acid level

Seven observed mismatches at nucleotide level

Alignment of the N-terminal 22 amino acid residues of human Cytochrome c with corresponding sequences from six other organisms

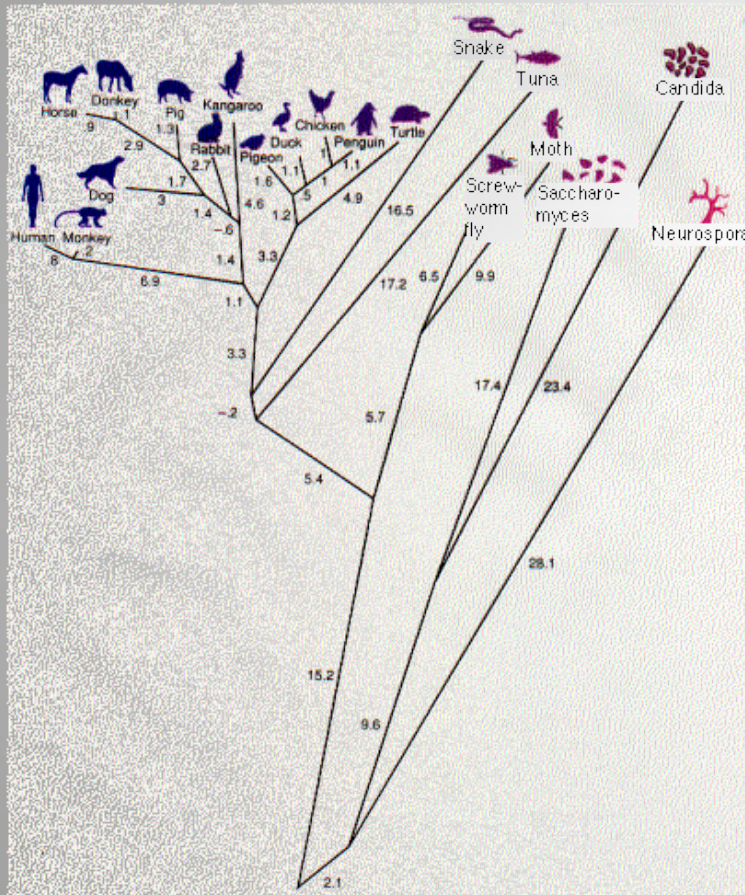


Component of electron transport chain in mitochondria
Highly conserved among species
The heme group is attached via Cys-14 and Cys-17

Molecular homology of cytochrome c

		1					6				10				14			17	18		20		
Human		Gly	Asp	Val	Glu	Lys	Gly	Lys	Lys	Ile	Phe	Ile	Met	Lys	Cys	Ser	Gln	Cys	His	Thr	Val	Glu	Lys
Pig		-	-	-	-	-	-	-	-	-	-	Val	Gln	-	-	Ala	-	-	-	-	-	-	-
Chicken		-	-	Ile	-	-	-	-	-	-	-	Val	Gln	-	-	-	-	-	-	-	-	-	-
Dogfish		-	-	-	-	-	-	-	-	Val	-	Val	Gln	-	-	Ala	-	-	-	-	-	-	Asn
Drosophila	<<<<	-	-	-	-	-	-	-	-	Leu		Val	Gln	Arg		Ala	-	-	-	-	-	-	Ala
Wheat	<<<<	-	Asn	Pro	Asp	Ala	-	Ala	-	-	-	Lys	Thr	-	-	Ala	-	-	-	-	-	Asp	Ala
Yeast	<<<<	-	Ser	Ala	Lys	-	-	Ala	Thr	Leu	-	Lys	Thr	Arg	-	Glu	Leu	-	-	-	-	-	-

First Cytochrome c phylogenetic Tree suggested by Fitch and Margoliash



numbers represent the minimum number of nucleotide substitutions in the gene for cytochrome c needed to produce these 20 proteins from a series of hypothetical ancestral genes at the various branching points (nodes)

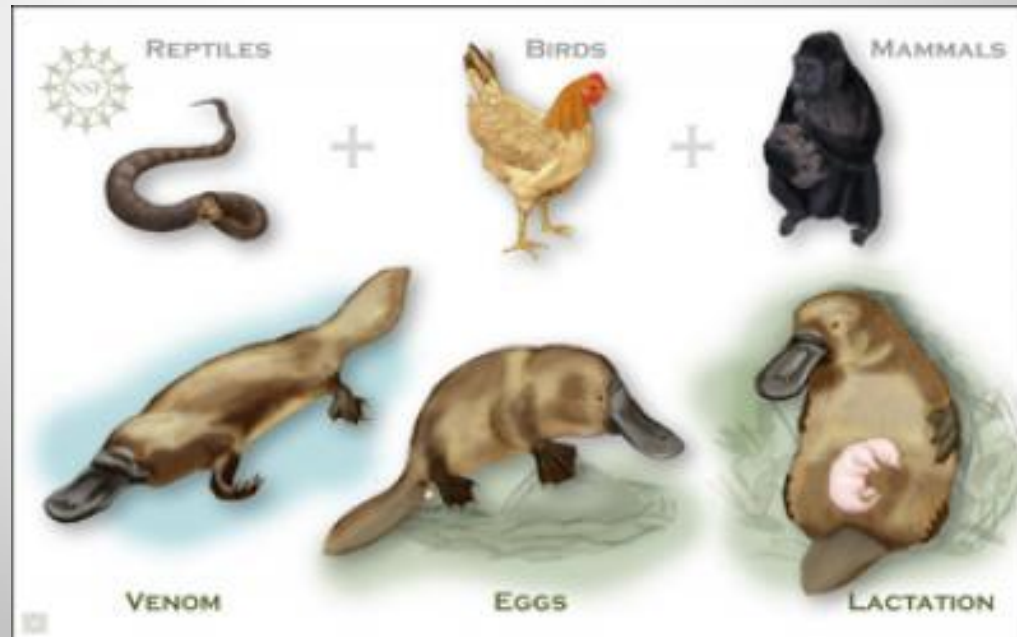
It indicates, for example, that the primates (humans and monkeys) split off before the split separating the kangaroo, a marsupial, from the other placental mammals

Construction of Phylogenetic Trees
•Walter M. Fitch and
•Emanuel Margoliash
Science 20 January 1967: 279-284.

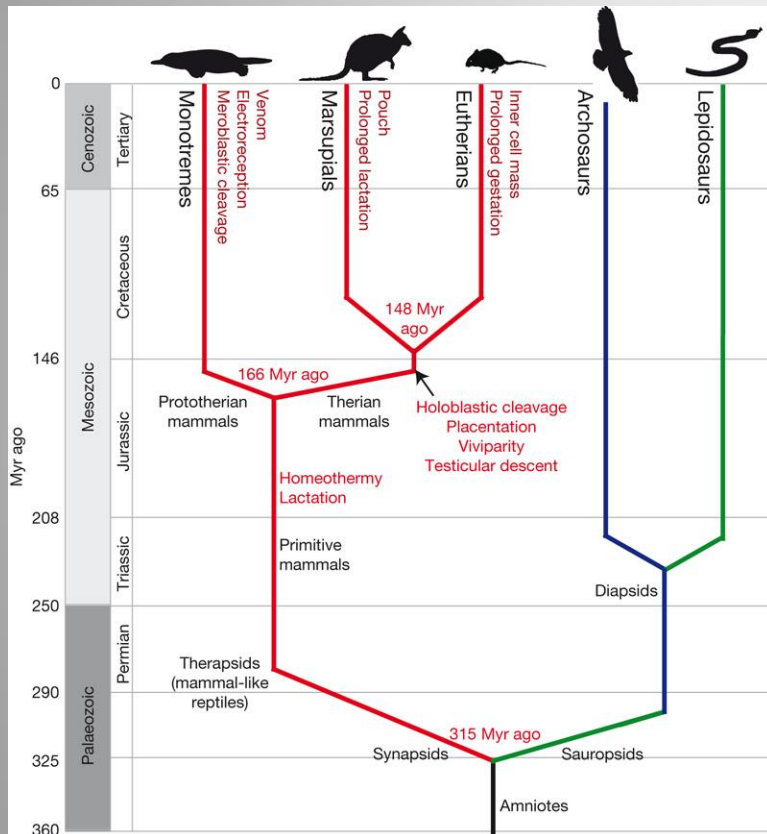
More Analysis needed!!



Is the Platypus a freak combination
of nature:
Part bird, part Reptile, part Mammal?



Platypus Phylogenetic Tree



Nature, 453:175, 2008

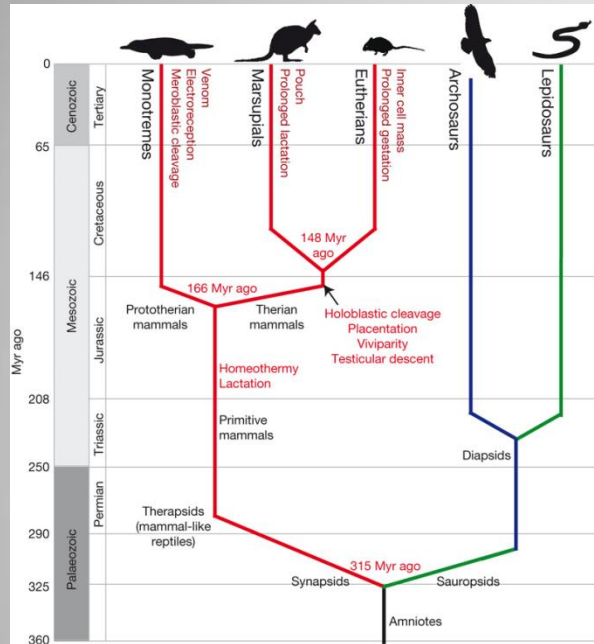
If particular gene (e.g. egg laying) is present in platypus and chicken genomes, but not other mammal groups we conclude that this gene was present in a shared ancestor of all these groups, but lost in the line leading to marsupials or placental mammals

If gene is present in platypus, but not in other mammals or birds, this gene was not present in shared ancestor, but newly derived (e.g. venom)

Alternatively, this gene could have been present in common ancestor, but lost in birds marsupials and placental animals

Parsimony: applied to phylogeny prefers the tree that minimizes the total amount of evolutionary change

The Platypus is part of the group of Mammals



Kept the ancestral , egg –laying trait from reptile-like ancestors of all mammals

The duck bill, webbed feet and venomous spines are newly derived traits. These are examples of convergent evolution.

By making the right comparisons, we can figure out which genes were in the shared ancestors of birds, reptiles, and mammals, but which have been lost in the line leading to marsupial and placental mammals

OUTLOOK:

Challenges Drawing Phylogenetic Trees

Horizontal Gene Transfer (esp. bacteria)

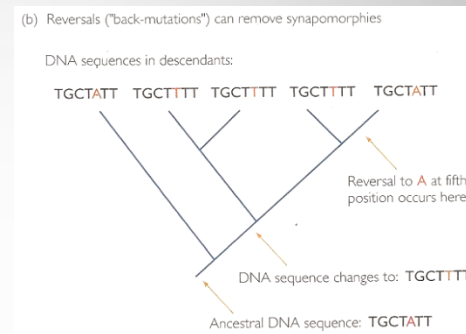
Back mutations

Convergence development

There is no “one-size-fits-all” recipe for phylogenetic analysis of molecular data

Reliability and practicality of different algorithms and computer programs depends on size and structure of dataset under analysis.

User needs to understand what the method is actually doing with the data



Bibliography:

Freeman, Herron. Evolutionary Analysis, 4th edition, 2007, Published by Pearson Education, Inc. ISBN: 0-13-227584-8

Pevsner. Bioinformatics and Functional Genomics, 2nd edition. 2009, Published by John Wiley & Sons. ISBN: 978-0-470-08585-1

<http://www.ncbi.nlm.nih.gov/About/primer/phylo.html>

Warren et al., Genome Analysis of the platypus reveals unique signatures of Evolution. Nature, Vol 453, 8 May 2008. 175-183

Walter M. Fitch and Emanuel Margoliash . Construction of Phylogenetic Trees Science 20 January 1967: 279-284.

<http://www.science20.com/print/28389>

[http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/T/Taxonomy.](http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/T/Taxonomy.html#DNA_DNAHybridization)

html#DNA_DNAHybridization

<http://people-press.org/reports/pdf/254.pdf>