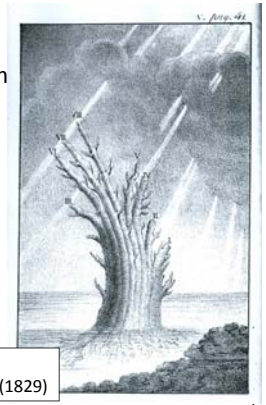
03-327/727 Phylogenetics

Dannie Durand
- Office hours: Tuesday 5pm – 6:30pm
- MI 646

Pieter Spealman
- TBA



Eichwald,
Zoologia specialis (1829)

From "Trees and networks before and after Darwin", Mark A Ragan Biology Direct, 2009

1

03-327/727 Phylogenetics

Course website:
http://www.cs.cmu.edu/~durand/Phylogenetics

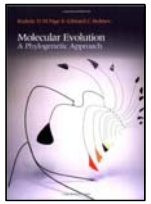
Password required to access some material when logging in from off campus
id: compbio
password: genomics

2

Textbook

Textbook: *Molecular Evolution: A phylogenetic approach,* Page and Holmes

This book is available, free, in pdf format on the course website.



Additional readings from journal articles and scanned textbooks will be provided on the syllabus page.

3

03-327/727 Phylogenetics: Course work, homework, grading

- 6-8 homework assignments
- 4 data analysis practica
- Two in-class exams
- Final exam
- Project for 727

4

03-327/727 Phylogenetics: Course work, homework, grading

- 03-327:
  Homework: 25 pts
  Practica: 30 pts
  In class exams: 15 pts
  Final: 30 pts
  Total: 100 pts
- 03-727:
  Homework: 25 pts
  Practica: 30 pts
  In class exams: 15 pts
  Final: 30 pts
  Project: 25 pts
  Total: 125 pts

5

03-727 Project: Phylogenetic analysis of a molecular data set

Milestones

- Formulate a phylogenetic question involving a sequence family of interest
- Collect appropriate sequences
- Multiple sequence alignment
- Phylogeny estimation
- Assess the reliability of your tree
- Interpret results

6

03-727 Project: Phylogenetic analysis of a molecular data set

Deliverables

- Project plan
- Final paper
- In-class presentation
- 3-5 pages (plus appendices)
  - Introduction
  - Methods
  - Results/Discussion
  - Appendix including alignments, trees, etc.

Project can be carried out by a 2-person team

7

Topics

- Introduction:
  - a phylogeny is an evolutionary tree that represents a hypothesis concerning evolutionary history
- The underlying evolutionary processes
  - Species evolution
  - Gene and genome evolution
  - Gene trees versus species trees
- How to interpret an evolutionary tree
  - Tree terminology
  - Evolution of traits on a tree
- Molecular phylogenetics: Building evolutionary trees using sequences
  - Finding and aligning sequences
  - Models of sequence evolution
  - Tree reconstruction methods
  - Sources of error, assessing accuracy

8

What is a phylogeny?

Phylogeny:

An evolutionary tree that represents a *hypothesis* concerning the evolutionary history of a group of taxa and their ancestors.

Taxon:

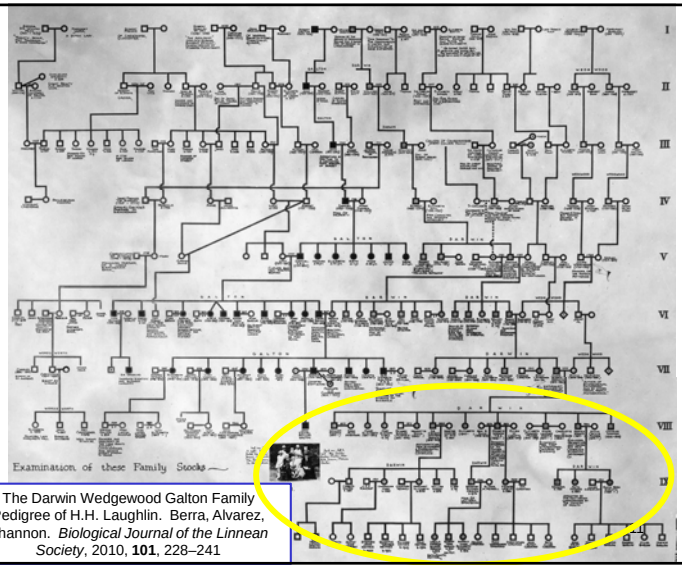
- a unit of classification
- strain, species, individual, gene

A phylogeny is (sort of) like a pedigree or family tree….

9

Two motivating examples

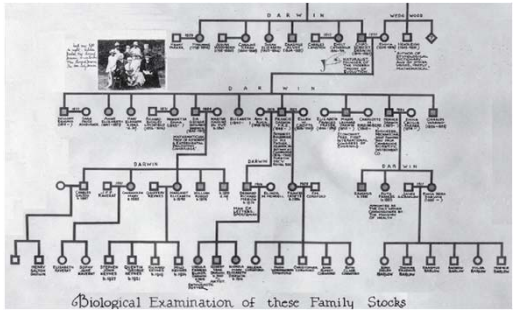
- Darwin's family and Darwin's finches
  - Similarities and differences between family trees and evolutionary trees
- Color vision in Primates
  - Why is it useful to think about the evolutionary history of genes independent of the evolutionary history of species?



The Darwin Wedgewood Galton Family Pedigree of H.H. Laughlin. Berra, Alvarez, Shannon. *Biological Journal of the Linnean Society*, 2010, **101**, 228–241

A pedigree…

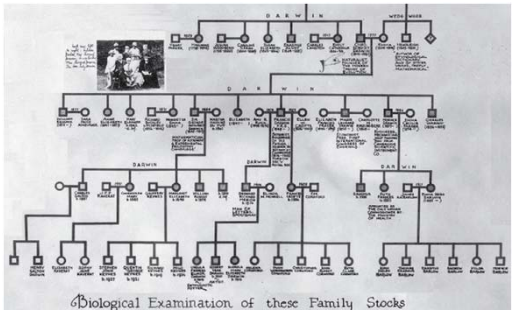
- represents a branching process where
- an individual gives rise to zero, one, or more than one offspring



12

Family relationships

- Parent versus child, ancestor versus descendant
- Siblings versus cousins

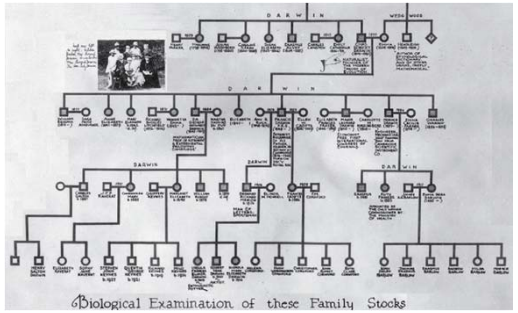


- You are more closely related to your siblings than to your cousins. Why?

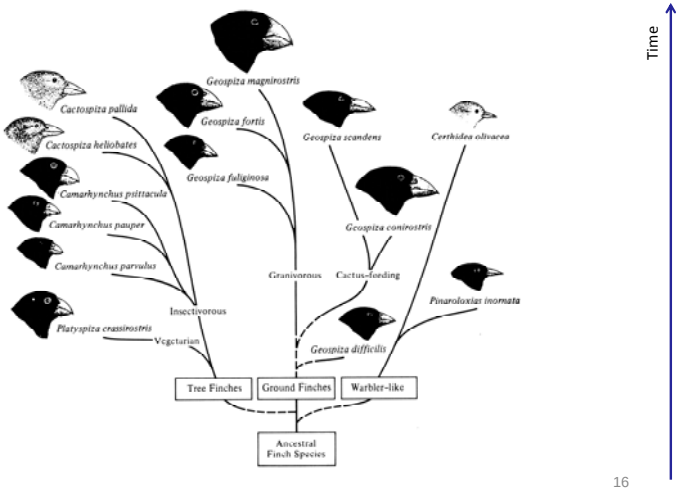
13

Some temporal relationships

- A pedigree tells us that your parent was born before you were born, but not who was born first, you or your cousin

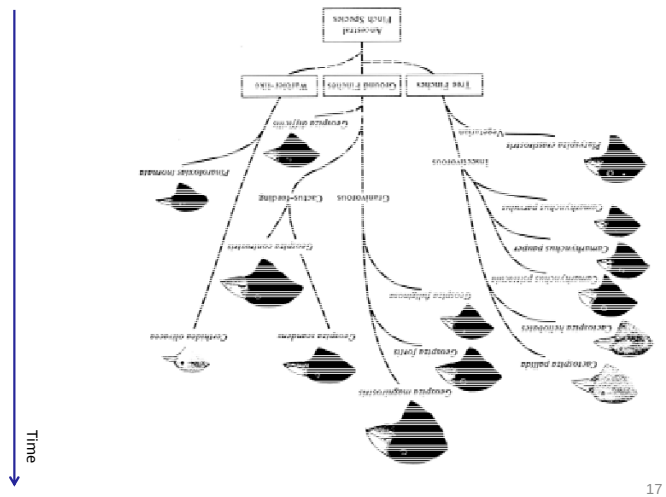


14



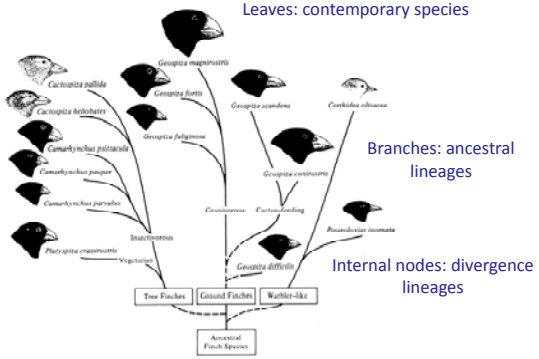
16

Pianka, Evolutionary Ecology, 6th or 7th ed. Phylogeny of the Galápagos finches. Adapted from Lack (1947) and Grant (1986)



17

Pianka, Evolutionary Ecology, 6th or 7th ed. Phylogeny of the Galápagos finches. Adapted from Lack (1947) and Grant (1986)



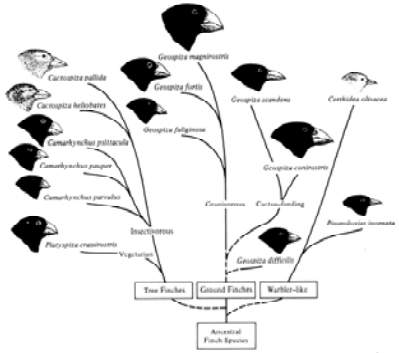
18

Pianka, Evolutionary Ecology, 6th or 7th ed. Phylogeny of the Galápagos finches. Adapted from Lack (1947) and Grant (1986)

A phylogeny…

represents a branching process where

a species gives rise to zero, one, or more than one species

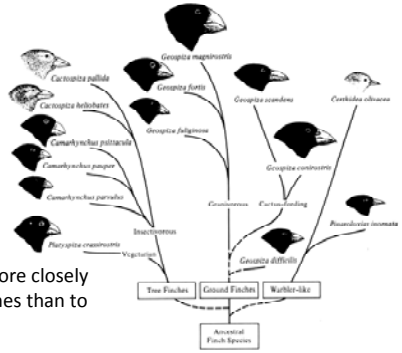


19

Pianka, Evolutionary Ecology, 6th or 7th ed. Phylogeny of the Galápagos finches. Adapted from Lack (1947) and Grant (1986)

Family relationships

- Parent/child
- Ancestor/descendant
- Sister taxa (siblings)



Cactus feeding finches are more closely related to granivorous finches than to warbler like finches

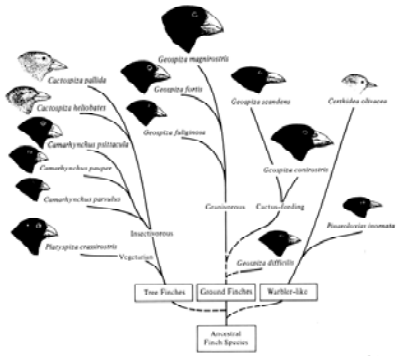
20

Pianka, Evolutionary Ecology, 6th or 7th ed. Phylogeny of the Galápagos finches. Adapted from Lack (1947) and Grant (1986)

Some temporal relationships

A phylogeny tells us that the common ancestor of ground finches predated the common ancestor of granivorous (grain eating) finches

But not whether the common ancestor of cactus-eating finches live before or after the common ancestor of granivorous finches



21

Pianka, Evolutionary Ecology, 6th or 7th ed. Phylogeny of the Galápagos finches. Adapted from Lack (1947) and Grant (1986)

Differences between family trees and phylogenies

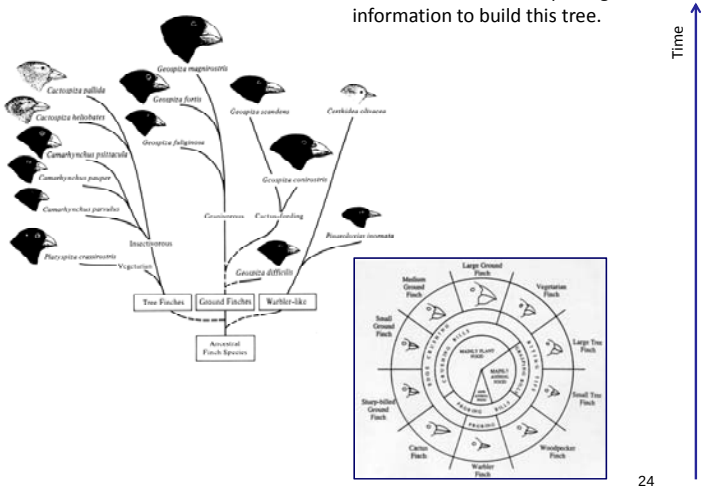
Family trees

- Each node corresponds to a single individual
- The history is derived from the historical record
- Every node has two parents
- Not binary: A couple can have many children

Species trees

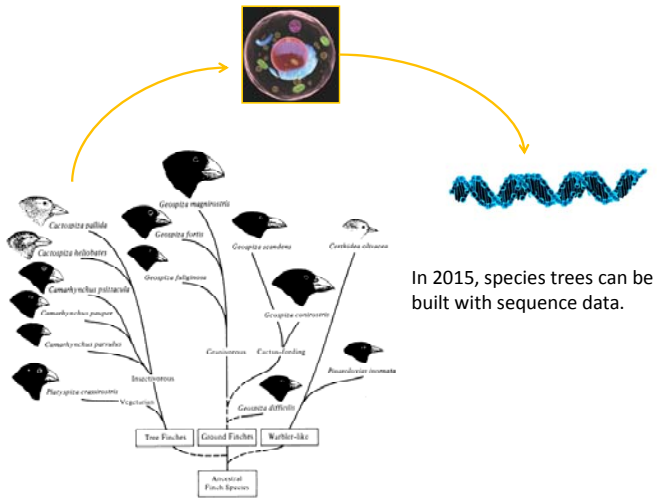
- Each node corresponds to a species
- The history is inferred from information about present-day species
- Each node has one "parent"
- Binary: every node has two children, except for leaf nodes

In 1945, Lack used morphological information to build this tree.



24

Pianka, Evolutionary Ecology, 6th or 7th ed. Phylogeny of the Galápagos finches. Adapted from Lack (1947) and Grant (1986)



In 2015, species trees can be built with sequence data.

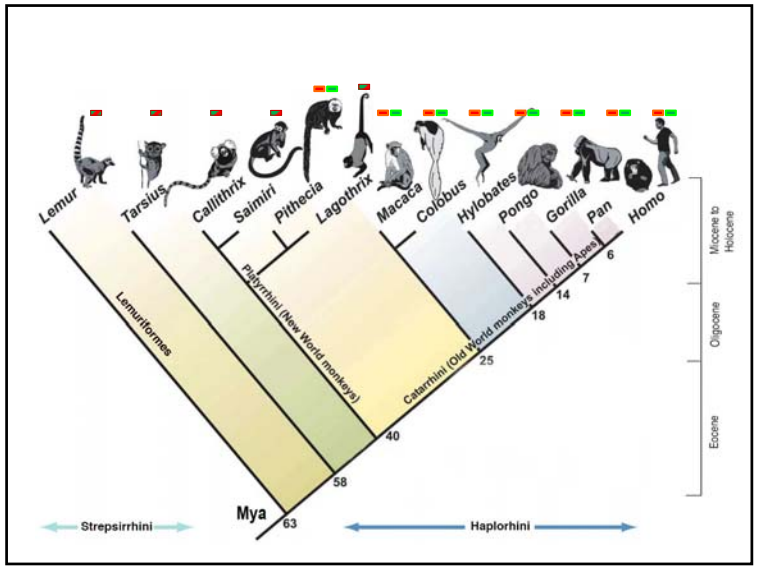
25

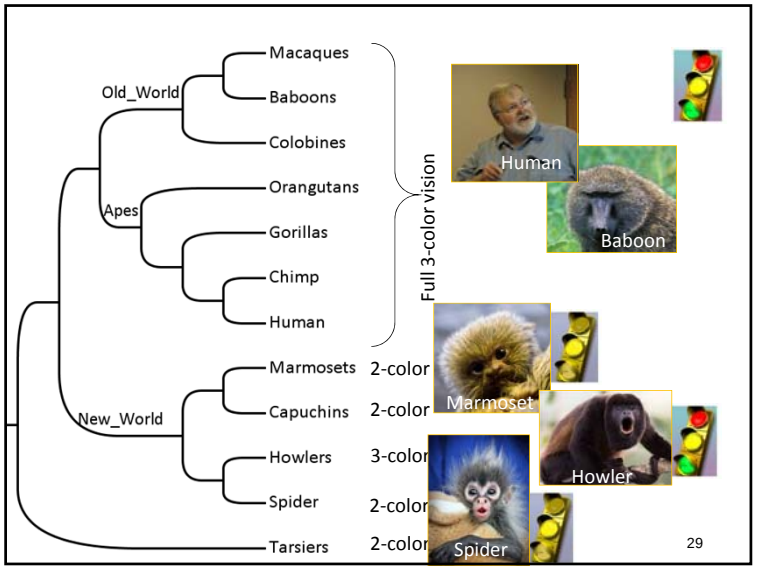
Pianka, Evolutionary Ecology, 6th or 7th ed. Phylogeny of the Galápagos finches. Adapted from Lack (1947) and Grant (1986)

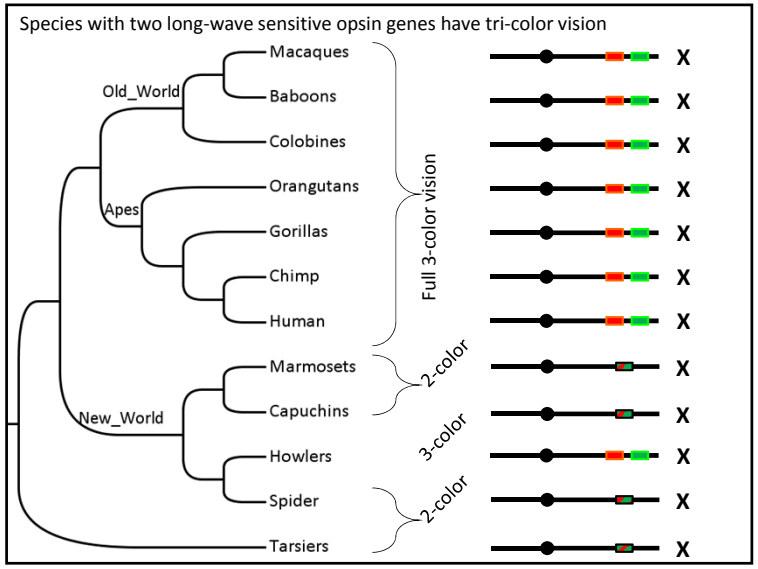
Two motivating examples

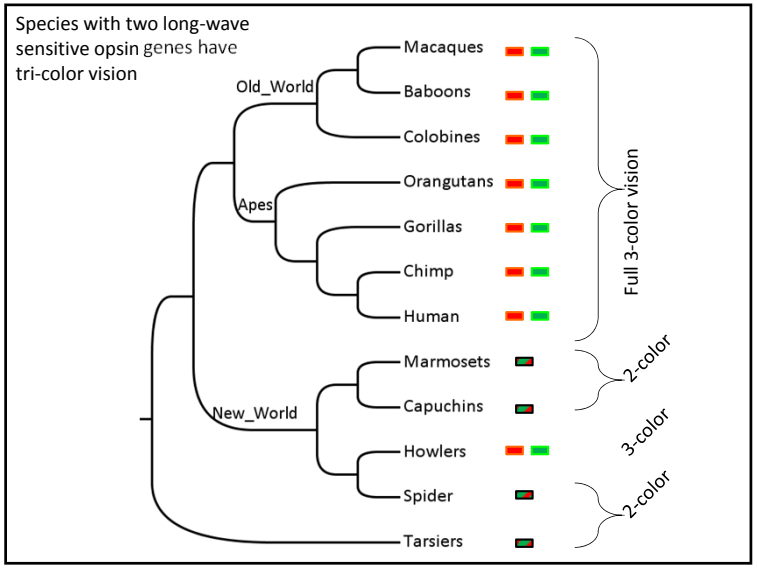
- Darwin's family and Darwin's finches
  - Similarities and differences between family trees and evolutionary trees

➢ Color vision in Primates
  - Why is it useful to think about the evolutionary history of genes independent of the evolutionary history of species?


Lemur
Tarsius
Callithrix
Saimiri
Pithecia
Lagothrix
Macaca
Colobus
Hylobates
Pongo
Gorilla
Pan
Homo
Lemuriformes
Platyrrhini (New World monkeys)
Catarrhini (Old World monkeys including Apes)
58
40
25
18
14
7
6
Mya
63
Strepsirrhini
Haplorhini
Miocene to Holocene
Oligocene
Eocene


Macaques
Baboons
Colobines
Orangutans
Gorillas
Chimp
Human
Old_World
Apes
New_World
Marmosets
Capuchins
Howlers
Spider
Tarsiers
Full 3-color vision
2-color
2-color
3-color
2-color
2-color
Human
Baboon
Marmoset
Howler
Spider
29

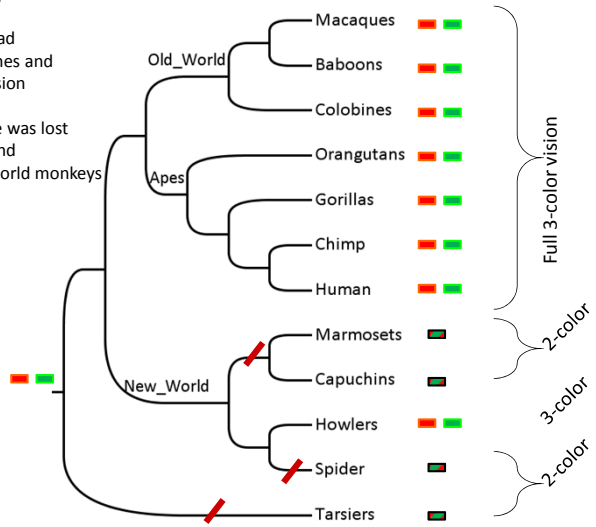

Species with two long-wave sensitive opsin genes have tri-color vision
Macaques
Baboons
Colobines
Orangutans
Gorillas
Chimp
Human
Old_World
Apes
New_World
Marmosets
Capuchins
Howlers
Spider
Tarsiers
Full 3-color vision
2-color
3-color
2-color
X


Species with two long-wave sensitive opsin genes have tri-color vision
Macaques
Baboons
Colobines
Orangutans
Gorillas
Chimp
Human
Old_World
Apes
New_World
Marmosets
Capuchins
Howlers
Spider
Tarsiers
Full 3-color vision
2-color
3-color
2-color

Hypothesis 1:

The ancestor had
- 2 opsin genes and
- tri-color vision

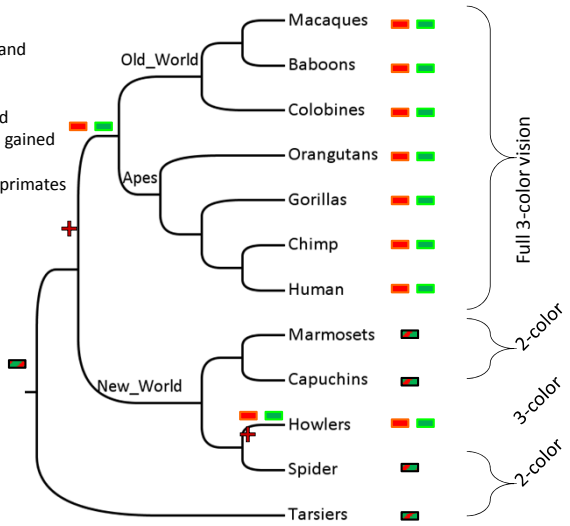
One opsin gene was lost
- in Tarsiers and
- most new world monkeys



Hypothesis 2:

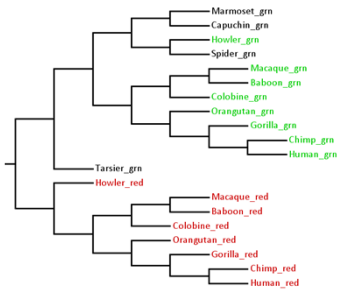
The ancestor had
- 1 opsin genes and
- bi-color vision

One opsin gene and tri-color vision was gained
- in Howlers and
- most old world primates

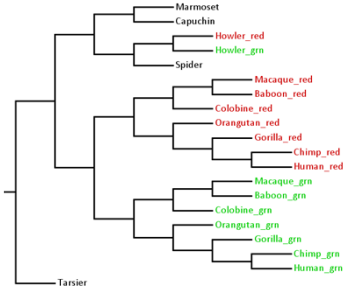


We can test these hypotheses using evolutionary trees for the long-wave opsin gene family

Hypothesis 1



Hypothesis 2

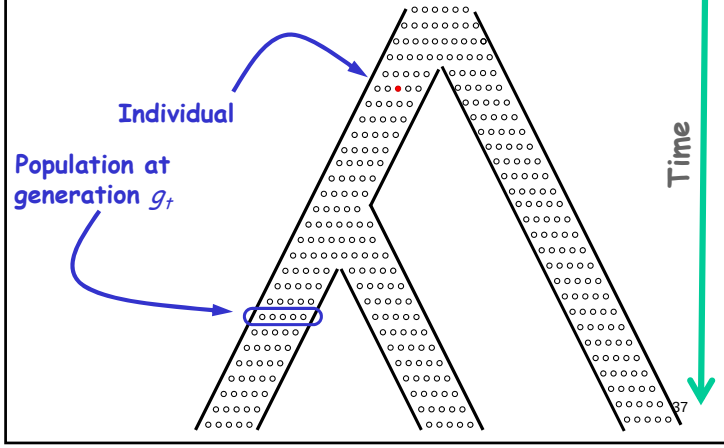


Evolutionary tree reconstruction

What are the processes we are trying to reconstruct?

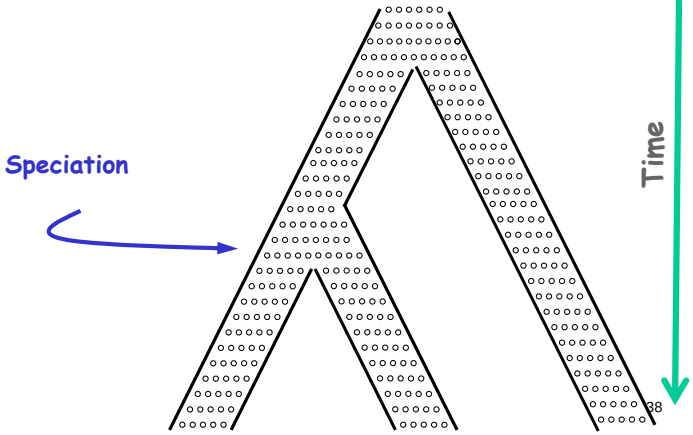
- ➢ Species evolution
- Sequence evolution
- Gene family evolution

A species tree represents the evolution of populations



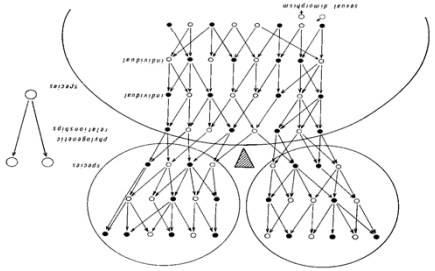
37

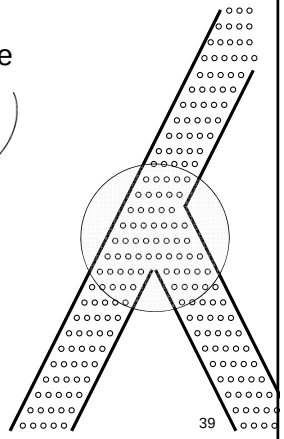
A species tree represents the evolution of populations



38

Close-up view of divergence





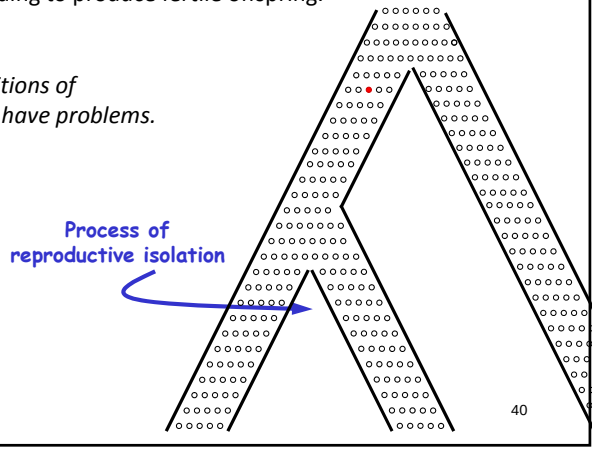
Modified from Hennig, W. (1966) *Phylogenetic Systematics*

39

Biological Species Concept

- A species is a group of individuals capable of interbreeding to produce fertile offspring.

Many definitions of species. All have problems.



40

One Problem in the Biological Species Concept

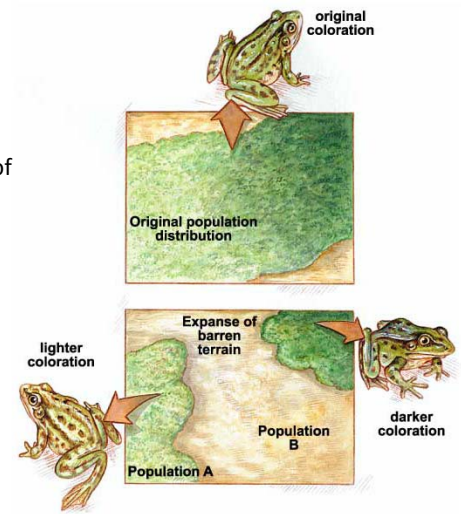
For asexually-reproducing organisms, like these bacteria, what constitutes a species?

How Do Species Arise?

The key to speciation is reproductive isolation of populations.

There are extrinsic and intrinsic reproductive isolating mechanisms.

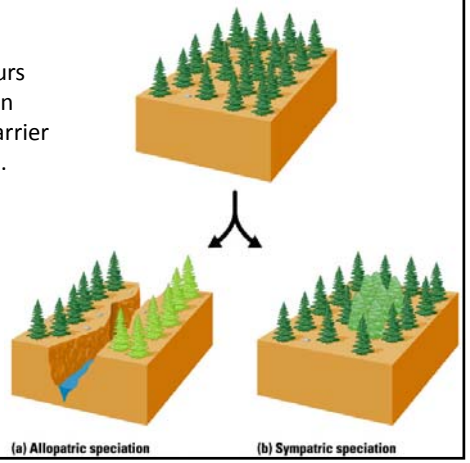
Geographic isolation is the primary extrinsic reproductive isolating mechanism.



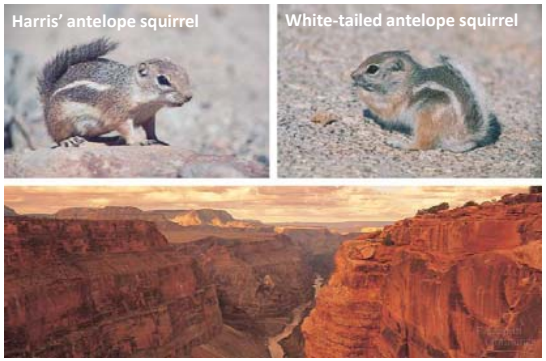
Reproductive Isolation May Occur With or Without Geographic Isolation

Allopatric speciation occurs when geographic isolation creates a reproductive barrier (an extrinsic mechanism).

Sympatric speciation occurs when a reproductive barrier is created by something other than geographic isolation (intrinsic mechanisms).



Allopatric Speciation



Two species of ground squirrel are postulated to have descended from a common ancestral population that was separated by formation of the Grand Canyon.

Many Intrinsic Reproductive Isolating Mechanisms Drive Speciation

Ecological Isolation:

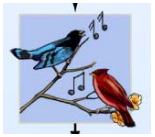
Individuals from separate species do not mate because they do not come into contact (e.g., different habitats within an overlapping range).

Temporal isolation:

Individuals from separate species do not mate because they breed at different times.

Many Intrinsic Reproductive Isolating Mechanisms Drive Speciation

Behavioral isolation:



Individuals from separate species do not mate because they are not attracted to one another.

Courtship rituals are

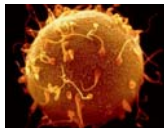
- critical for mating within a species
- ineffective for attracting members of other species.

Many Intrinsic Reproductive Isolating Mechanisms Drive Speciation

Mechanical isolation:

Individuals from separate species do not mate because they are not physically compatible.

Gametic isolation:



Even if they are physically compatible, an embryo will not form if the egg and sperm do not fuse properly.

Many Intrinsic Reproductive Isolating Mechanisms Drive Speciation

Hybrid inviability or infertility:

Even if fertilization occurs successfully, the offspring may not survive, or be infertile (e.g., mule).

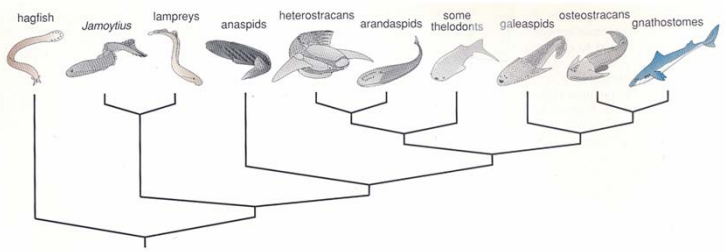
In this course,

- we will consider speciation as a separation of populations with reproductive isolation;
- species concepts and process of speciation are not the focus.

Why phylogeny reconstruction?

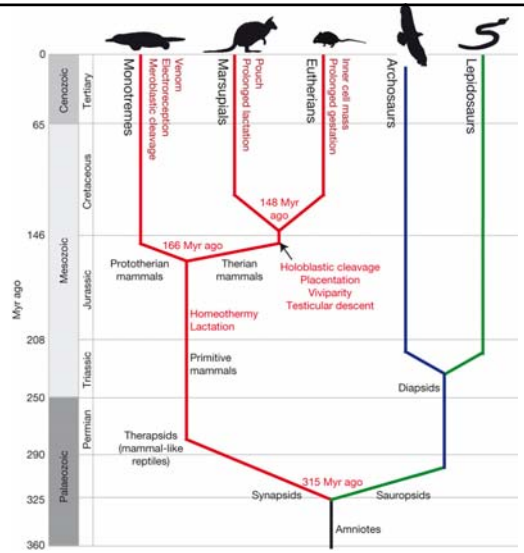
Some applications

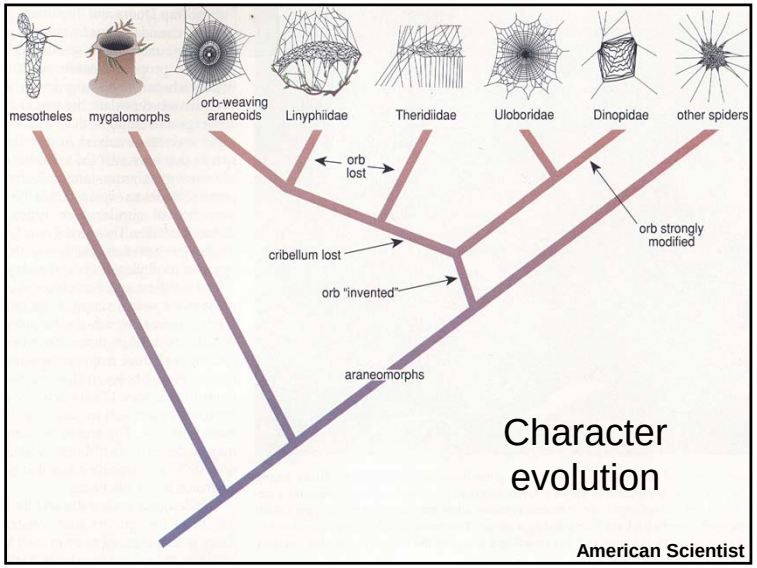
Species relationships



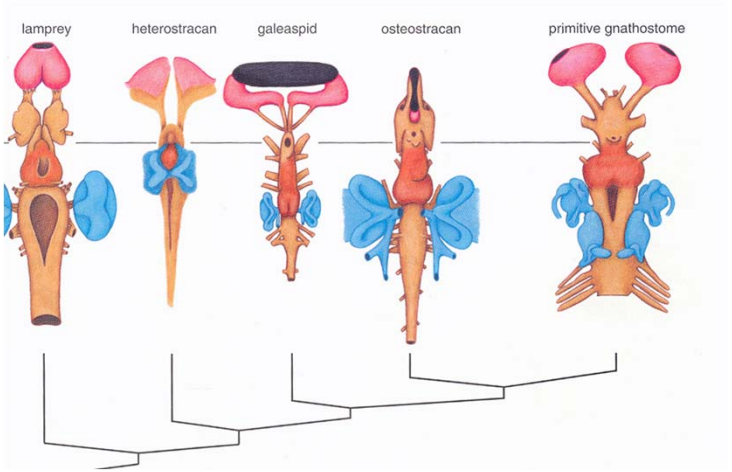
American Scientist

Genome analysis of the platypus reveals unique signatures of evolution
Nature 453, 175-183(8 May 2008)



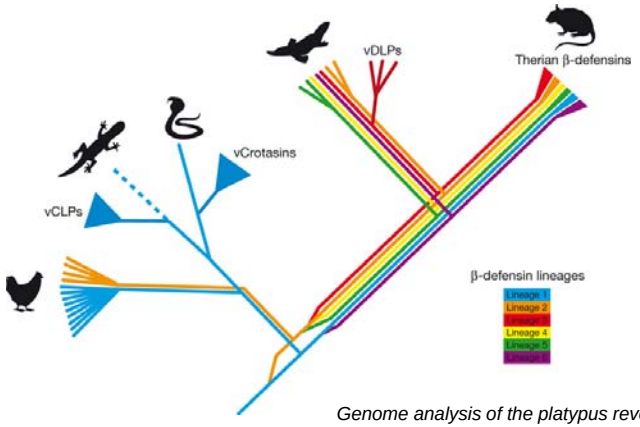


American Scientist



American Scientist

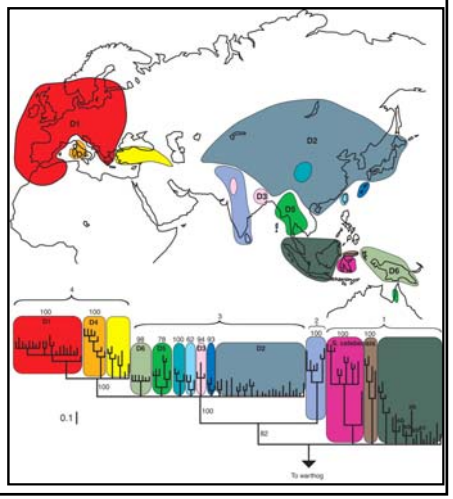
Novel function in a gene family: platypus venom



Genome analysis of the platypus reveals unique signatures of evolution
Nature 453, 175-183(8 May 2008)

Example of phylogeography

Independent sites of pig domestication



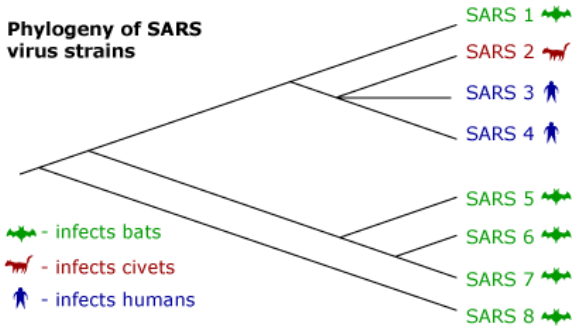
Source: Greger Larson, et al. 2005. Worldwide Phylogeography of Wild Boar Reveals Multiple Centers of Pig Domestication. Science 11 March 2005. Vol. 307. no. 5715, pp. 1618 - 1621. DOI: 10.1126/science.1106927.

Example of timescales on phylogenies

Timing the evolution of sociality in sweat bees to a warm period in geologic history

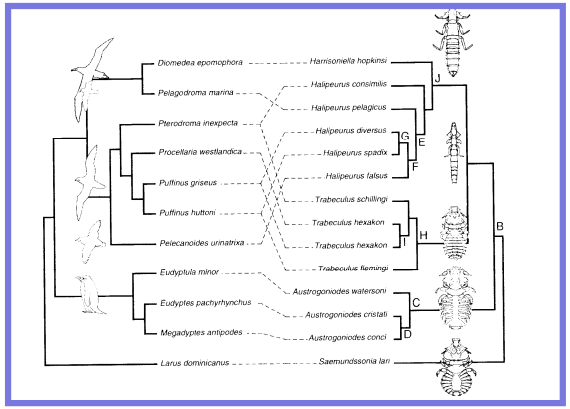
Source: Brady, S.G., S. Sipes, A Pearson, B. Danforth. 2006. Recent an simultaneous origins of eusociality in halictid bees. Proceedings of the Royal Society B. doi:10.1098/rspb.2006.3496

Example of disease phylogeny



Source: Understanding Evolution: Tracking SARS back to it's source. http://evolution.berkeley.edu/evolibrary/news/060101_batsars
After: Wendong Li, et al. 2005. Bats Are Natural Reservoirs of SARS-Like Coronaviruses. Science 28 October 2005: Vol. 310. no. 5748, pp. 676 - 679. DOI: 10.1126/science.1118391

Example of host-parasite phylogeny



Source: Page, R.D.M., Cruickshank, R.H., Dickens, M., Furness, R.W., Kennedy, M., Palma, R.L., Smith, V.S. 2004. Phylogeny of "*Philoceanus* complex" seabird lice (Phthiraptera: Ischnocera) inferred from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, 30: 633-652.