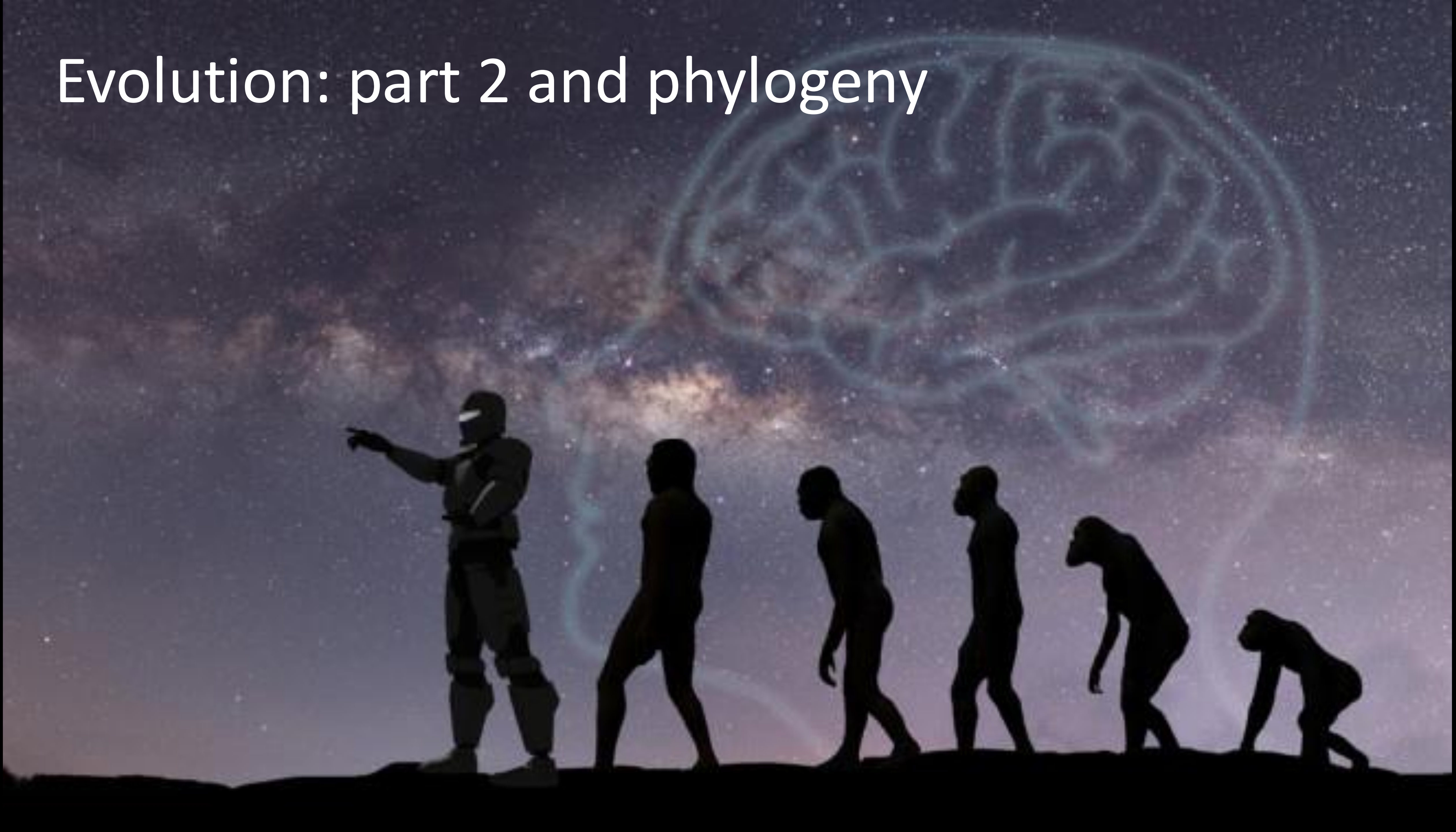
Evolution: part 2 and phylogeny



Evolution recap:

- History of evolution: Aristoteles, Lamarck, Darwin…
- Natural selection
- Speciation
- Reproductive barriers

- Sources of genetic diversity within a population
- Mechanisms for changes in allele frequency in the population
- Phylogeny and tree interpretation

- History of Earth and diversification of life

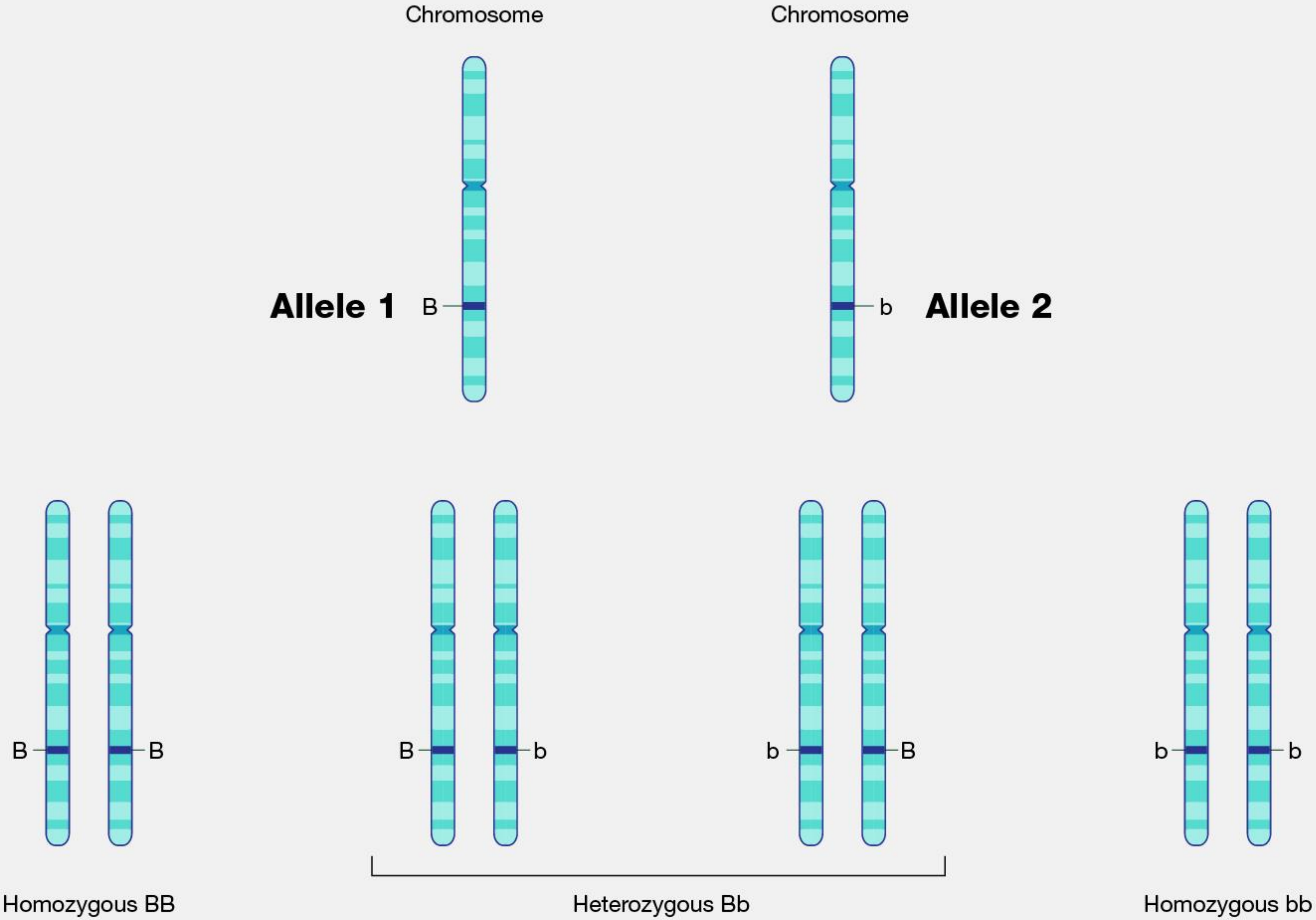
Microevolution:

- Refers to small-scale changes in gene frequencies within populations over relatively short periods.
- Typically occurs over generations or within a few hundred to thousands of years.
- Examples: Changes in allele frequencies, adaptations to local environments, variations in traits within a population.

Macroevolution:

- Macroevolution involves large-scale changes resulting in the formation of new species or higher taxonomic groups over geological time scales.
- Occurs over millions of years and leads to the diversification of life forms.
- Examples: Speciation events, evolutionary novelties, divergence of major taxa.

Reminder:


Chromosome
Chromosome
Allele 1
B
b
Allele 2
B
B
B
b
b
B
b
b
Homozygous BB
Heterozygous Bb
Homozygous bb

Reminder: the Smallest Unit of Evolution are populations

- A common misconception is that individual organisms evolve
- Natural selection acts on individuals, but only populations evolve
- Consider, for example, a population of medium ground finches on Daphne Major Island
  - During a drought, large-beaked birds were more likely to crack large seeds and survive

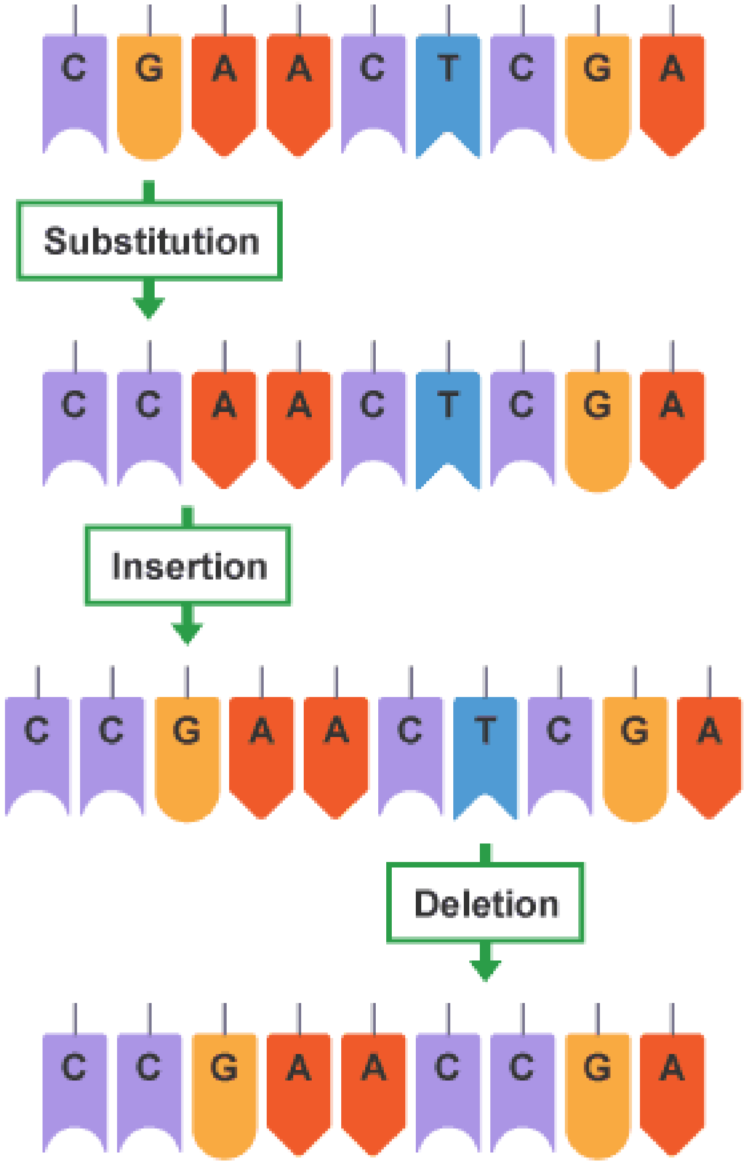
Genetic Variation in Populations

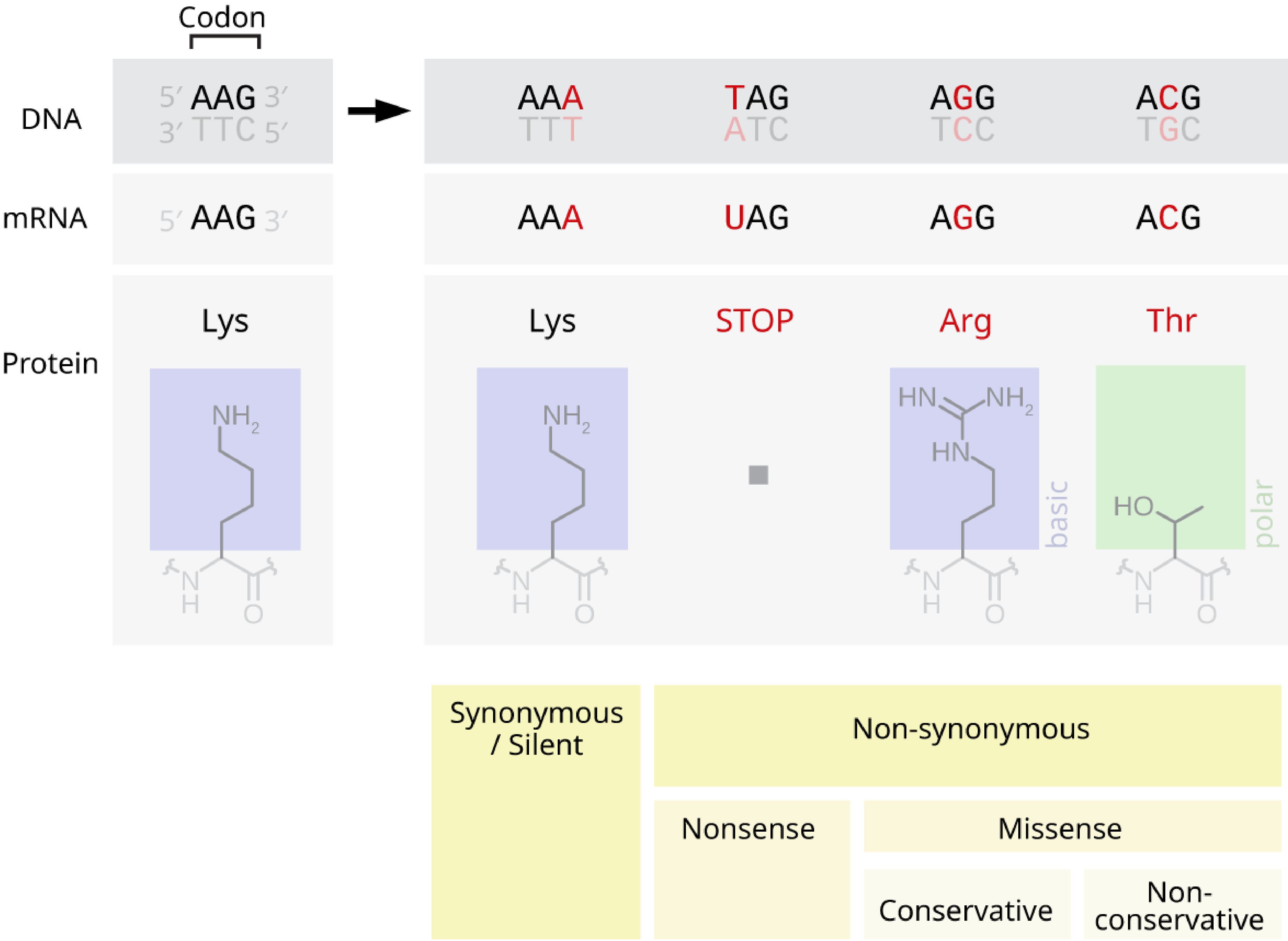
- Genetic variation in heritable traits is the basis on which evolution acts, providing the diversity necessary for populations to adapt to changing environments.
- **Sources of genetic variation**:
  - Mutations
  - Genetic recombination, sexual reproduction
  - Gene flow (migration)

- **Genetic variation** among individuals is caused by differences in genes or other DNA segments
- Variation in traits within populations can lead to differential survival and reproductive success, driving evolutionary change over time.

Types of mutations (nucleotides)

- Only mutations in cells that produce gametes can be passed to offspring
- A point mutation is a change in a single nucleotide in a DNA sequence





	Codon	→				
DNA	5′ AAG 3′ 3′ TTC 5′	→	AAA TTT	TAG ATC	AGG TCC	ACG TGC
mRNA	5′ AAG 3′		AAA	UAG	AGG	ACG
Protein	Lys		Lys	STOP	Arg	Thr

Synonymous / Silent

Non-synonymous

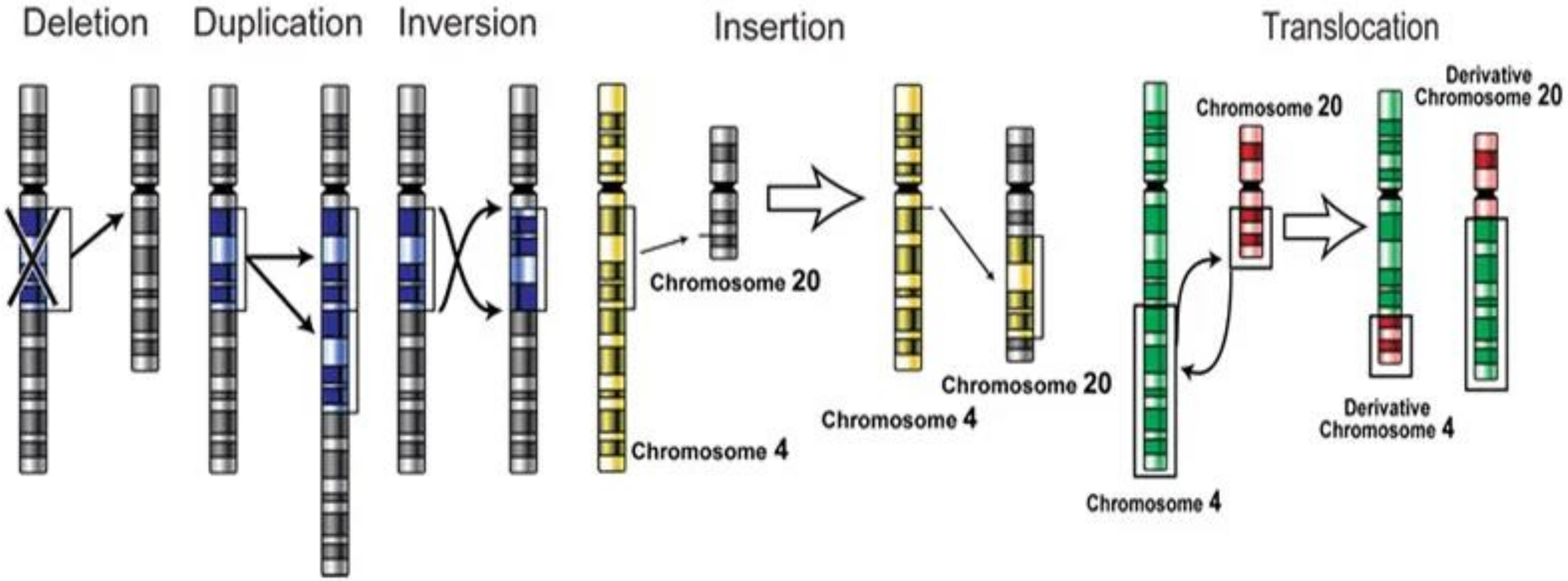
Nonsense

Missense

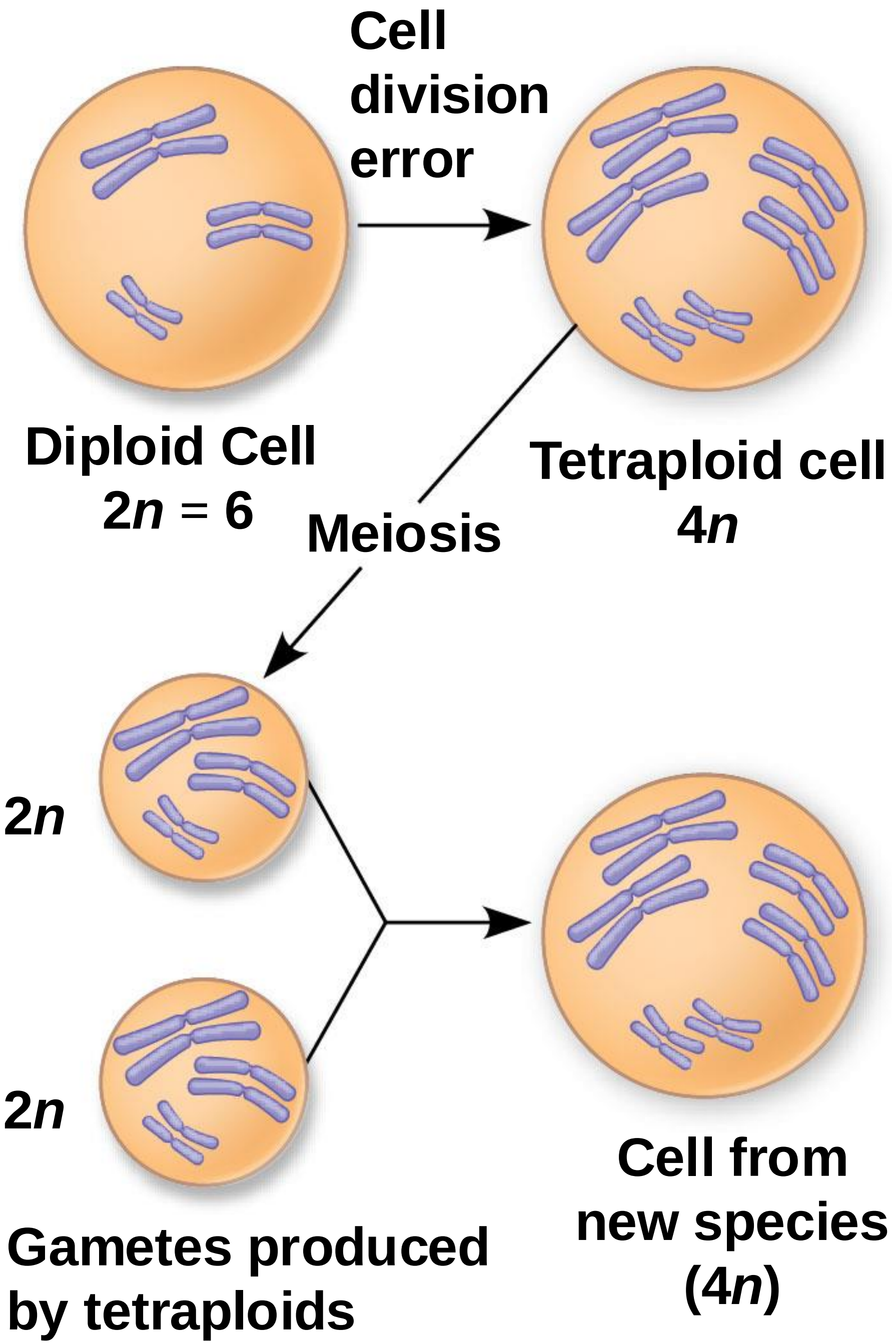
Conservative

Non-conservative

Types of chromosome variations



Repetition: Sympatric speciation by autopolyploidy

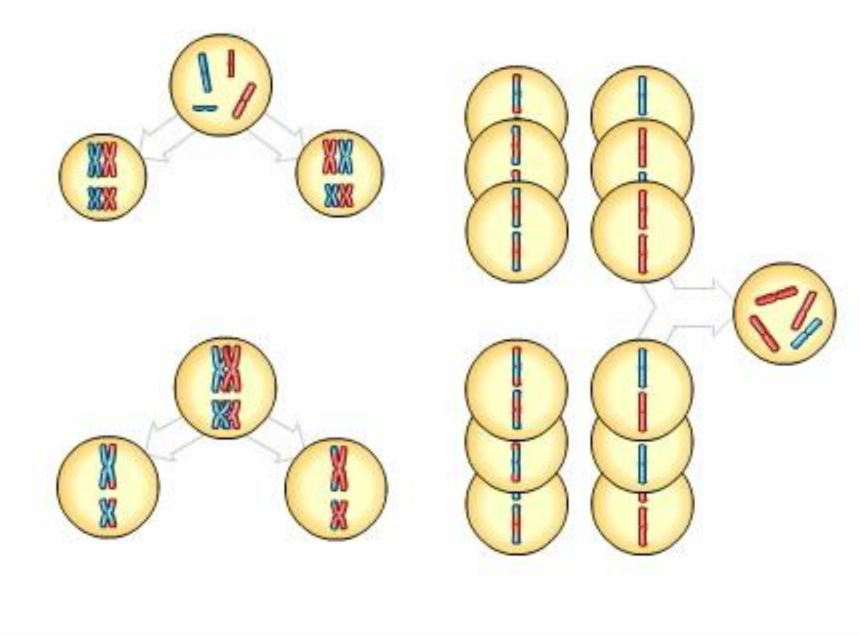


Nonheritable variation due to the environment

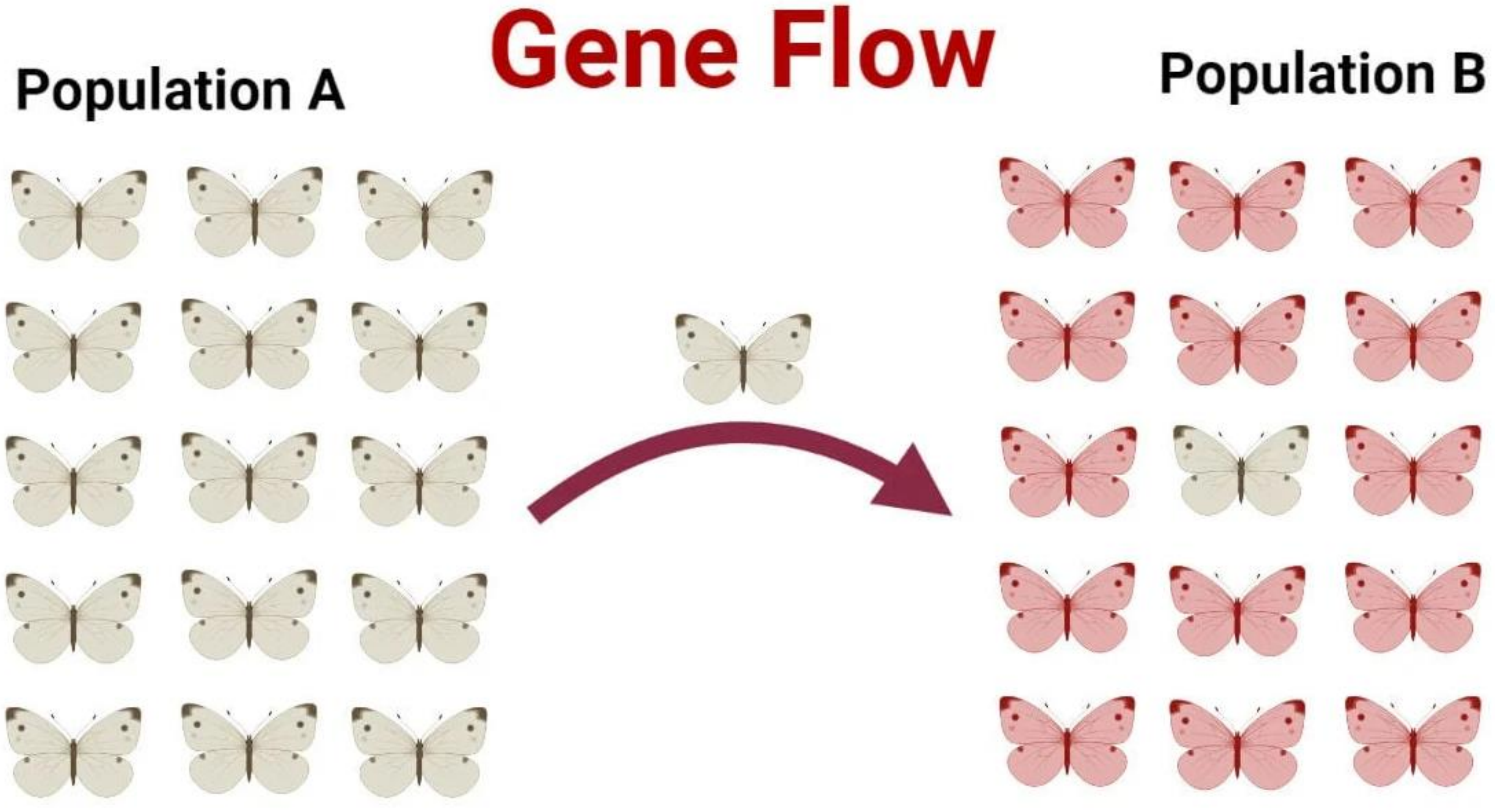
- Some phenotypic variation does not result from genetic differences among individuals, but rather from environmental influences
- Only genetically determined variation can have evolutionary consequences

Nonheritable variation due to the environment

Animation: Genetic Variation from Sexual Recombination



Source of genetic variation: migration

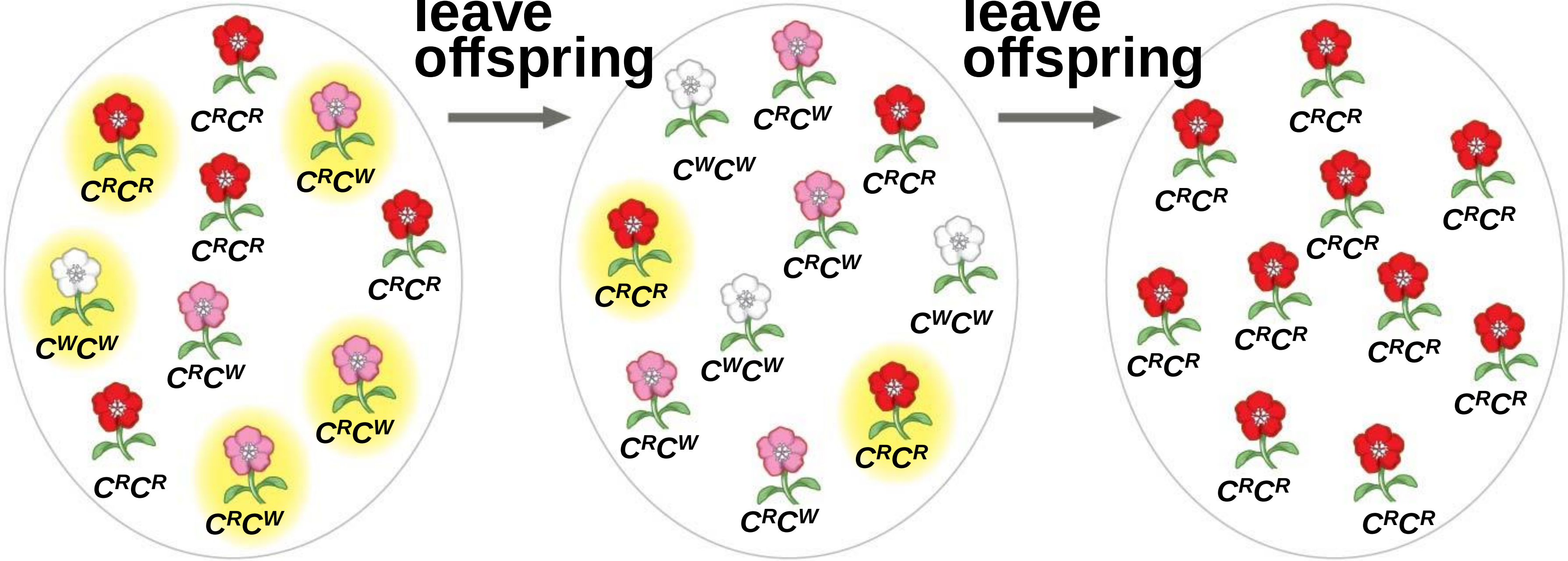


Processes influencing allele frequency in a population

Genetic Drift

- The smaller a sample, the greater the chance of random deviation from a predicted result
- **Genetic drift** describes how allele frequencies fluctuate unpredictably from one generation to the next
- Genetic drift tends to reduce genetic variation through the random loss of alleles

Genetic drift



5 plants leave offspring

2 plants leave offspring

Generation 1

p (frequency of C^R) = 0.7
q (frequency of C^W) = 0.3

Generation 2

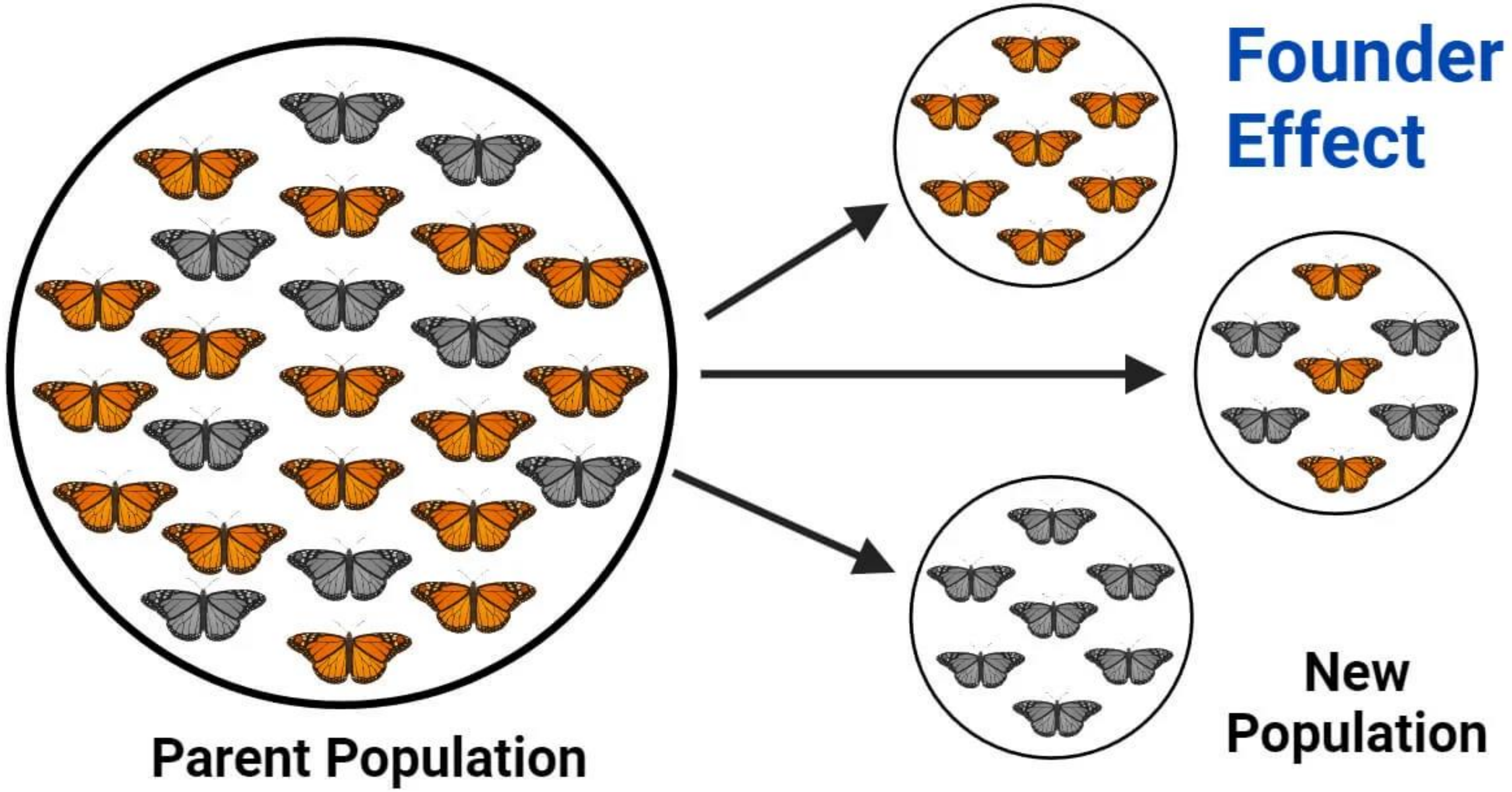
p = 0.5
q = 0.5

Generation 3

p = 1.0
q = 0.0

The Founder Effect

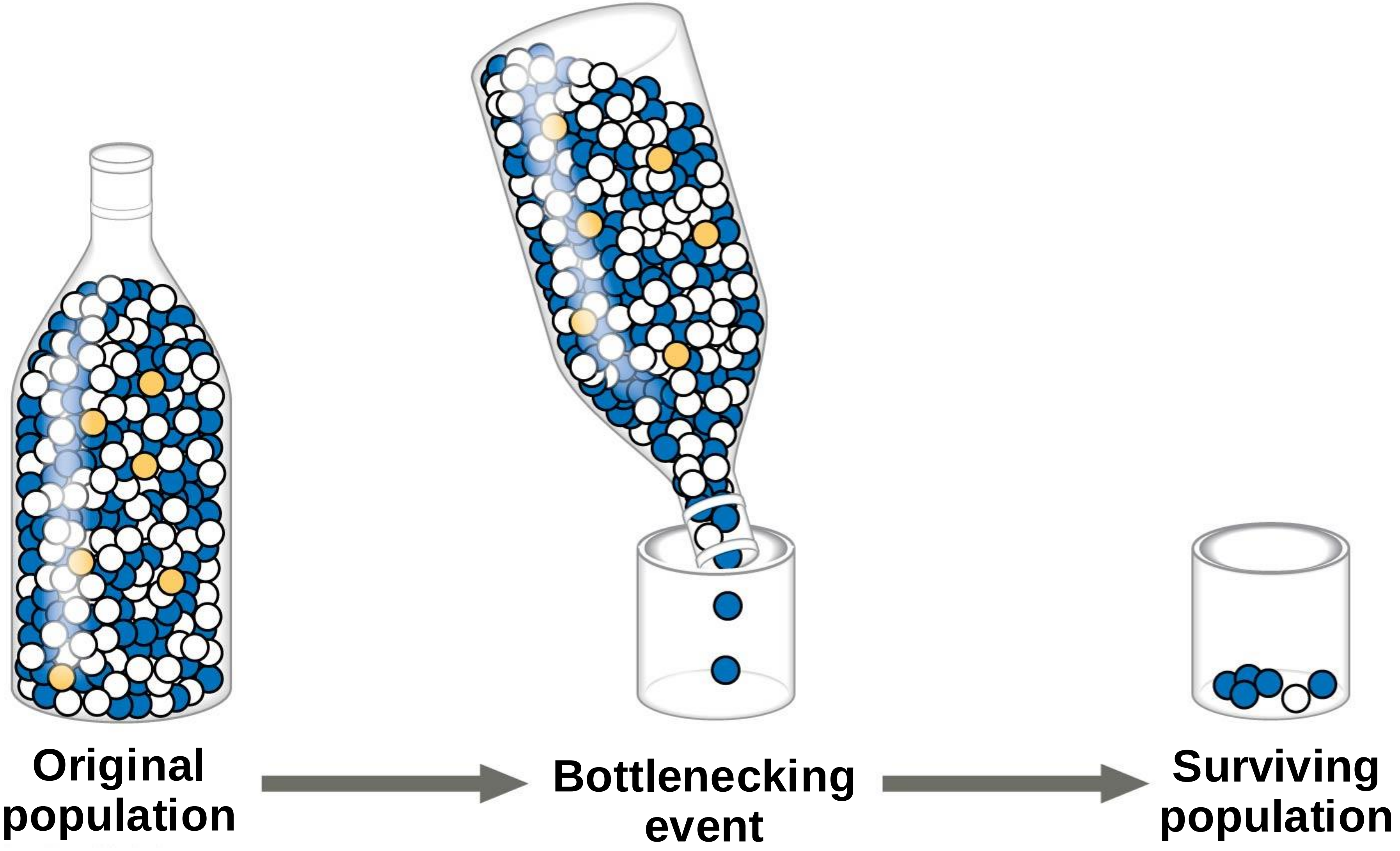
- The **founder effect** occurs when a few individuals become isolated from a larger population
- Allele frequencies in the small founder population can be different from those in the larger parent population



The Bottleneck Effect

- The **bottleneck effect** occurs when there is a drastic reduction in population size due to a sudden change in the environment
- The resulting gene pool may no longer be reflective of the original population’s gene pool
- If the population remains small, it may be further affected by genetic drift

The Bottleneck Effect

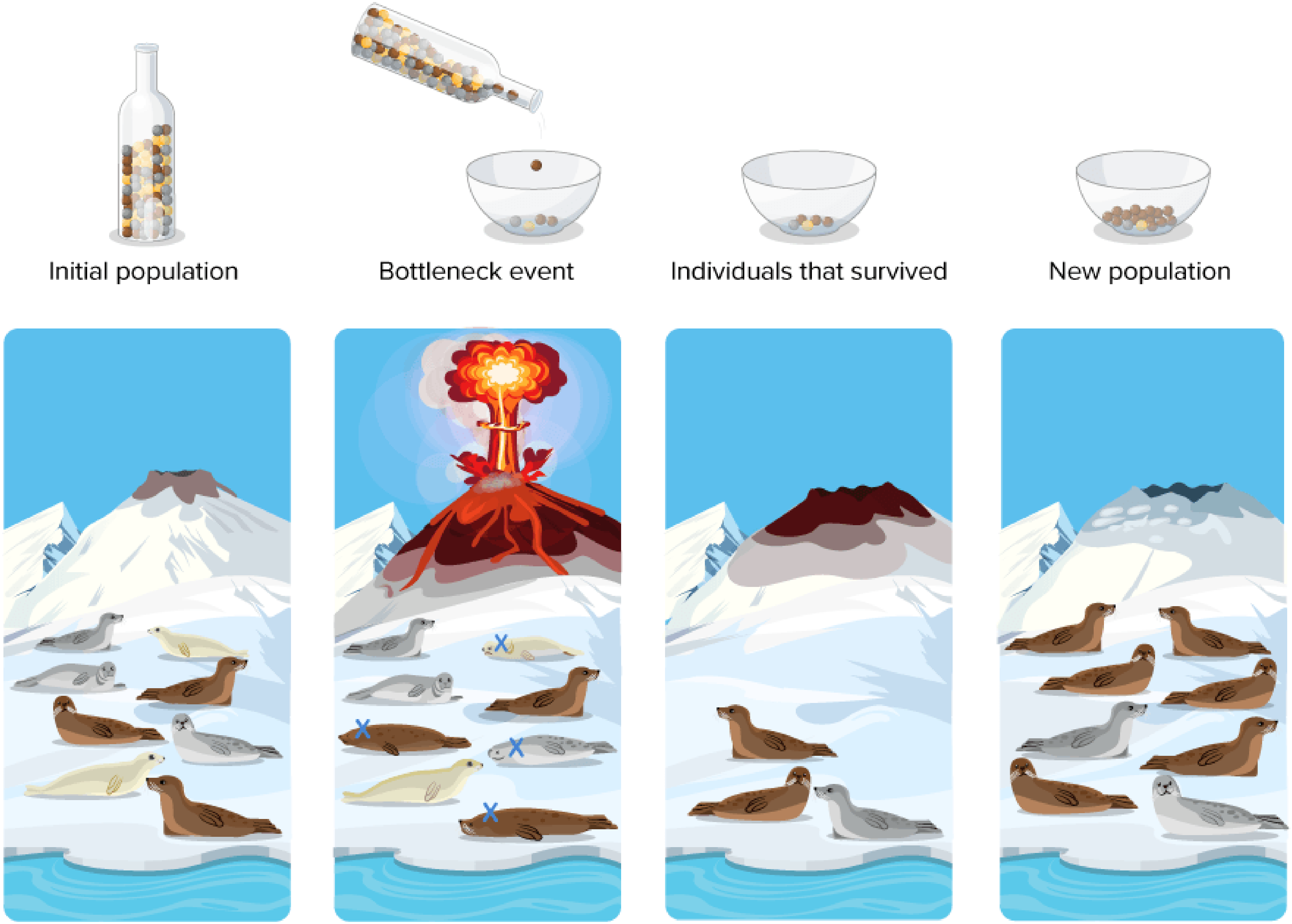


Original population

Bottlenecking event

Surviving population

Example:


Initial population
Bottleneck event
Individuals that survived
New population

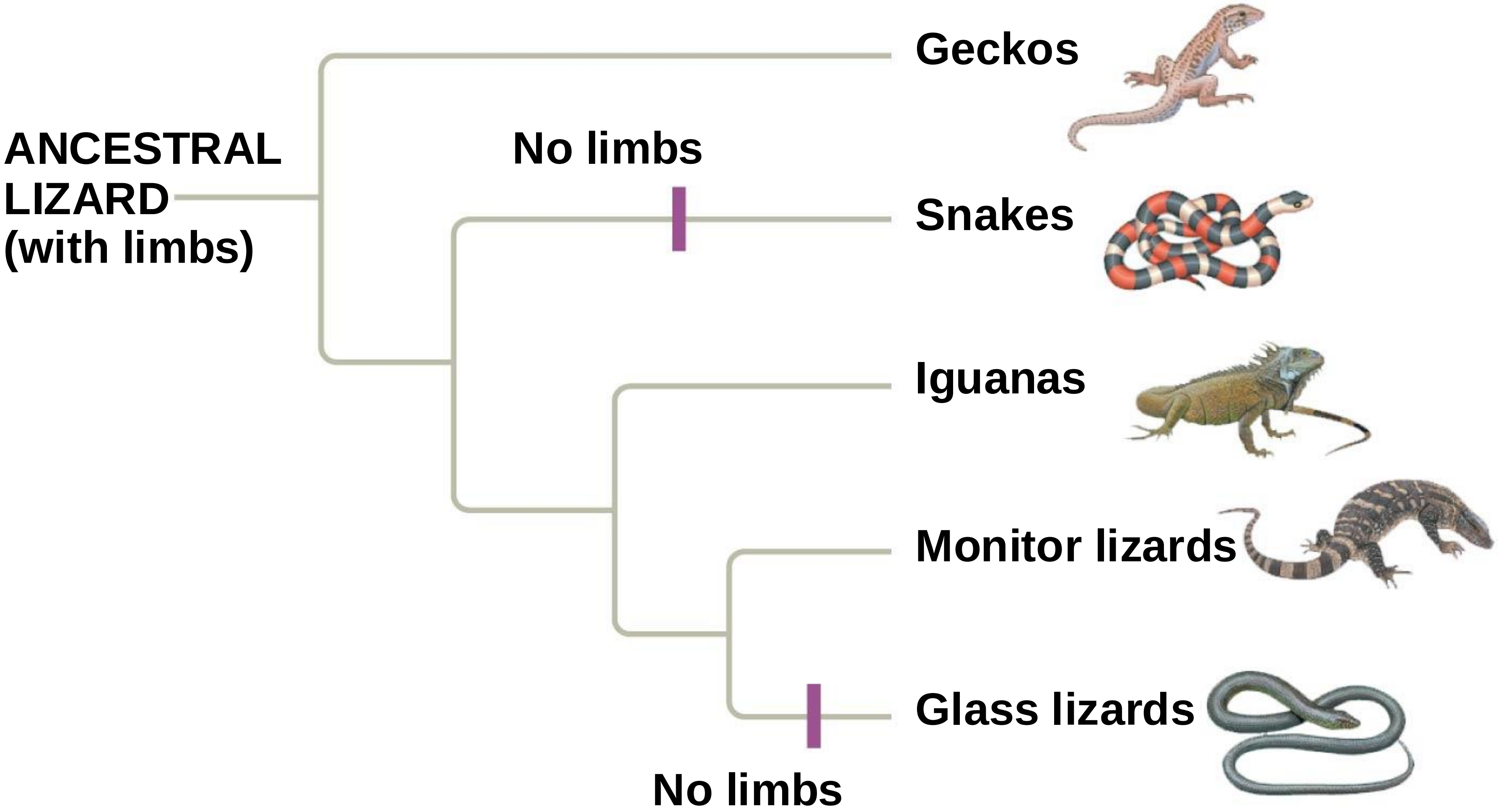
Phylogenetics

Investigating the Tree of Life

- **Phylogeny** is the evolutionary history of a species or group of related species
    - For example, a phylogeny shows that legless lizards and snakes evolved from different lineages of legged lizards
- The discipline of **systematics** classifies organisms and determines their evolutionary relationships

What kind of organism is this?

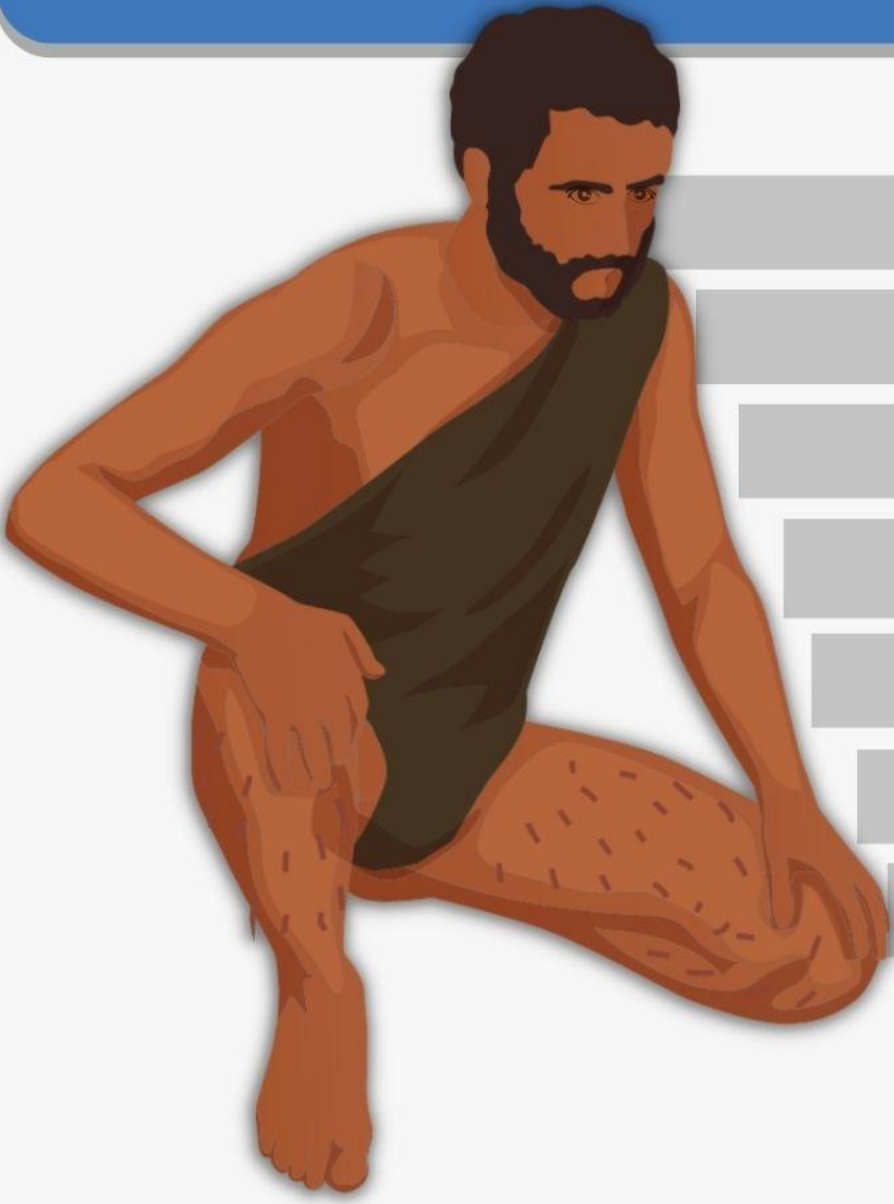
Convergent evolution of limbless bodies



Binomial Nomenclature

- In the 18th century, Carolus Linnaeus published a system of taxonomy based on resemblances
- Two key features of his system remain useful today: two-part names for species and hierarchical classification Eg. *Homo sapiens*

Binomial nomenclature



DOMAIN	Eukarya
KINGDOM	Animalia
PHYLUM	Chordates
CLASS	Mammalia
ORDER	Primates
FAMILY	Hominids
GENUS	*Homo*
SPECIES	*Homo sapiens*

Binomial nomenclature is a binomial system of naming a species.

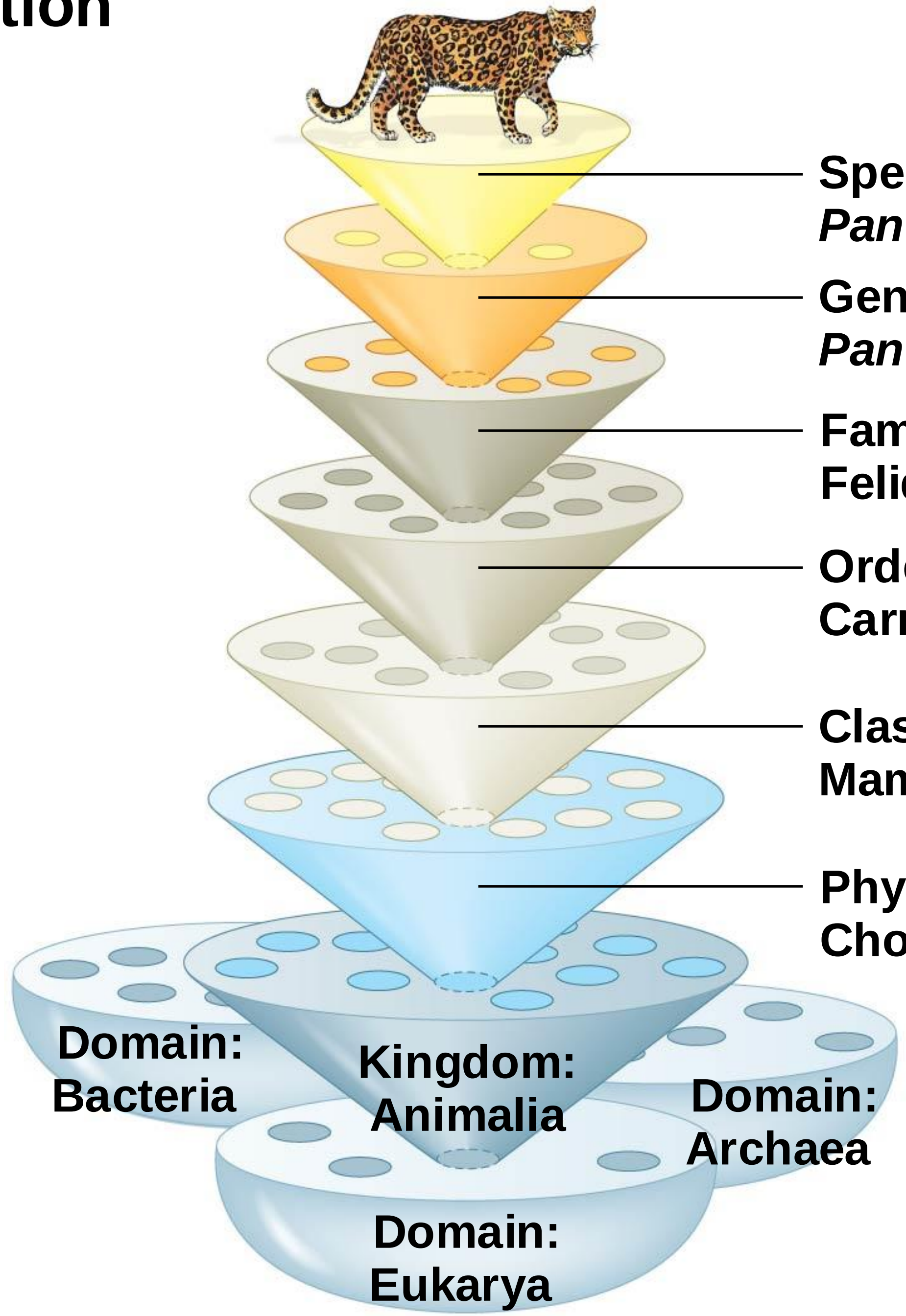
Example: Homo sapiens (scientific name of humans) consists of two parts:

(1) Generic name
(2) Specific name

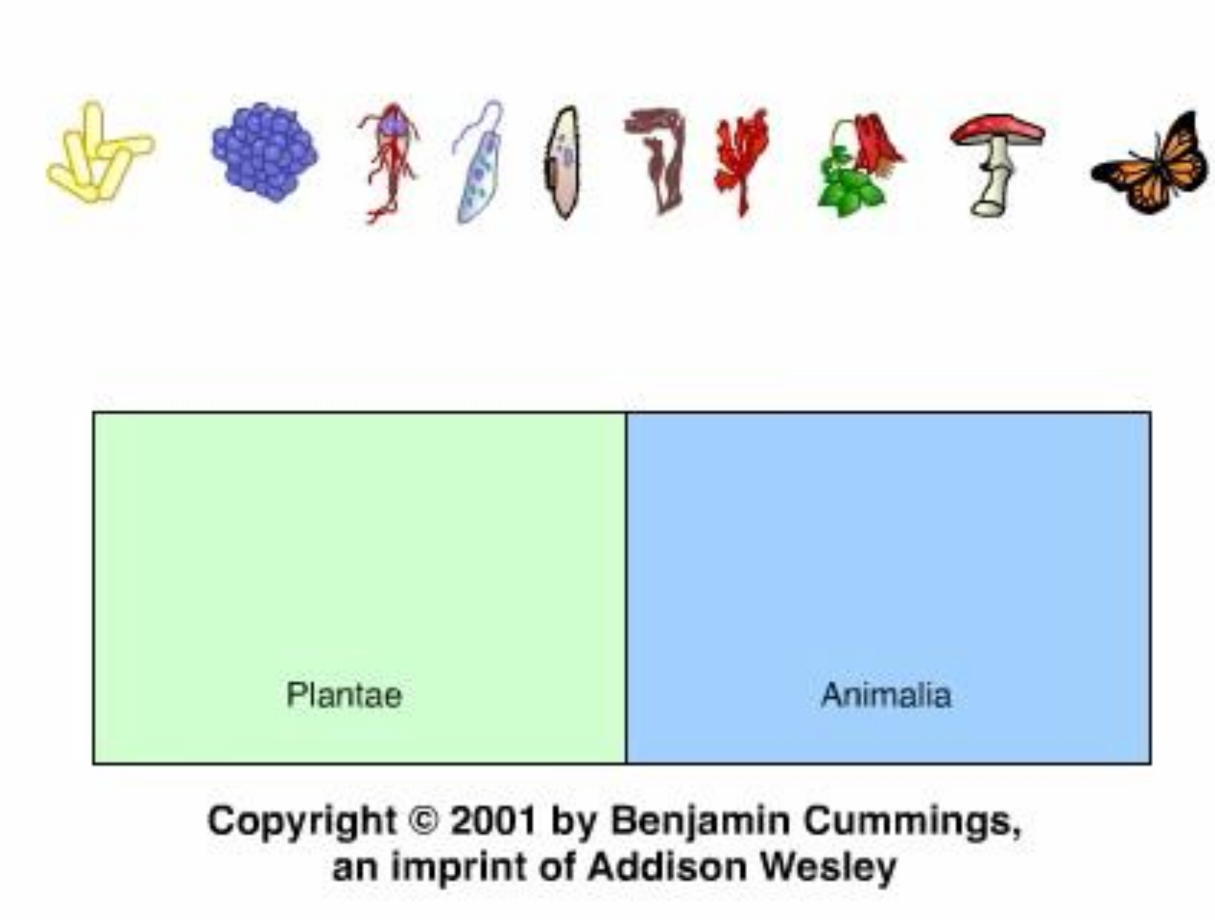
Hierarchical Classification

- Linnaeus introduced a system for grouping species in increasingly inclusive categories
- The taxonomic groups from broad to narrow are **domain**, **kingdom**, **phylum**, **class**, **order**, **family**, **genus**, and **species**
- A taxonomic unit at any level of hierarchy is called a **taxon**
- The broader taxa are not comparable between lineages
    - For example, an order of snails has less genetic diversity than an order of mammals

Linnaean classification

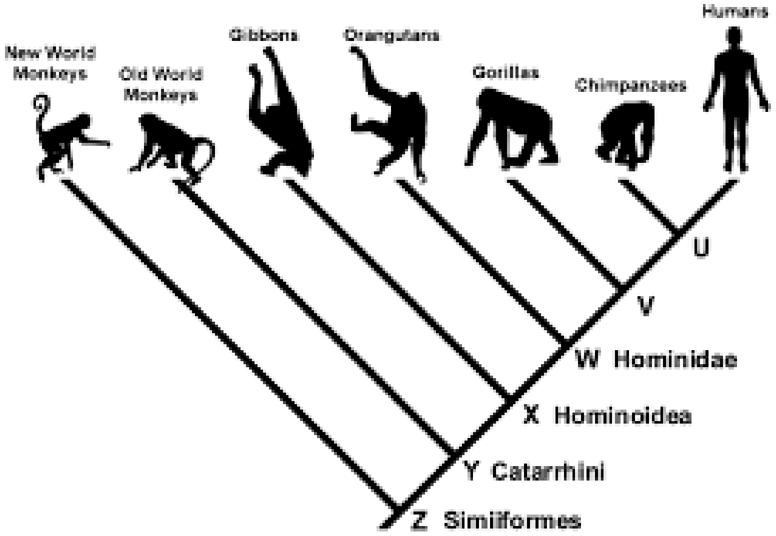


Animation: Classification Schemes

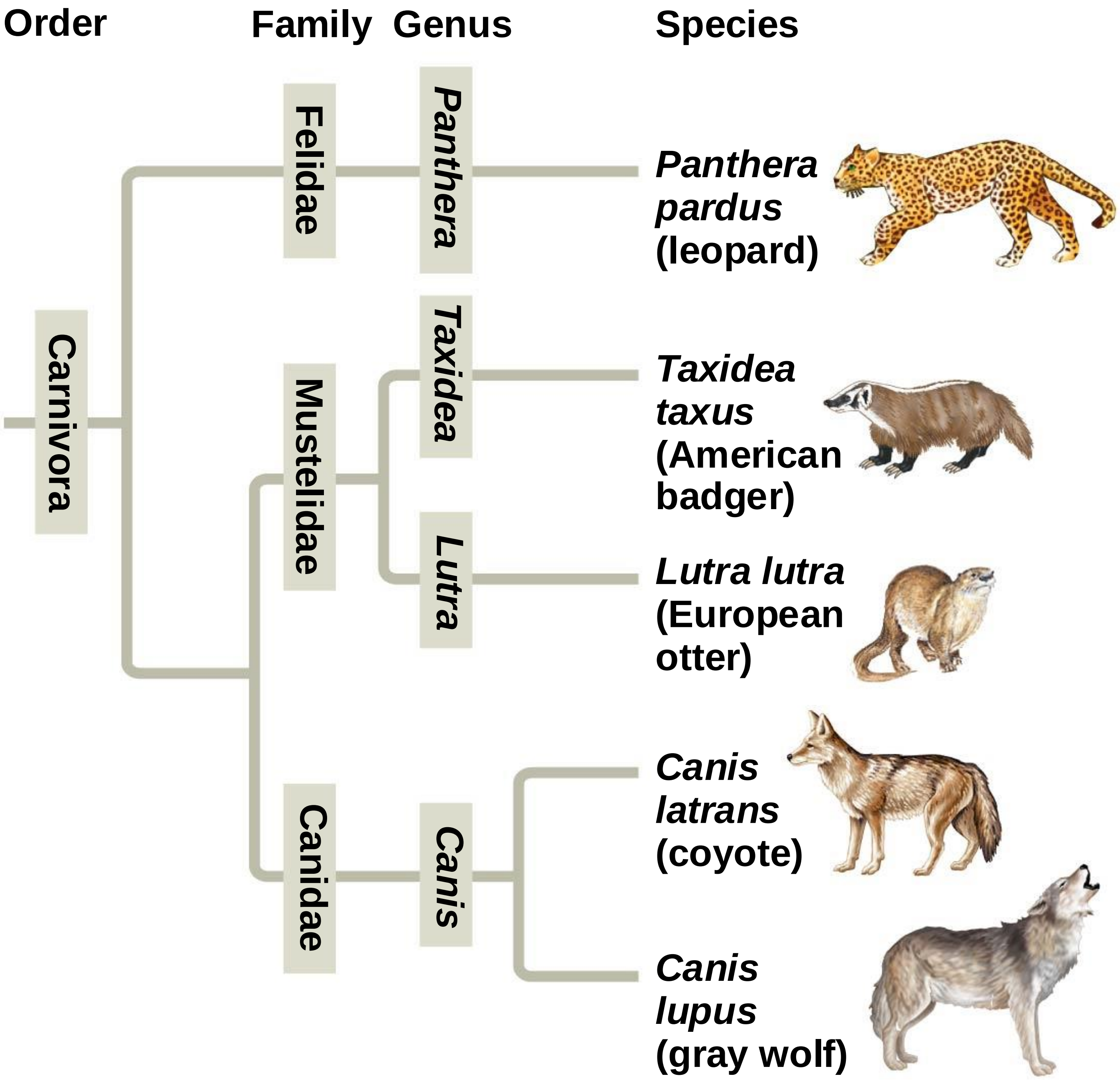


Linking Classification and Phylogeny

- The evolutionary history of a group of organisms can be represented in a branching **phylogenetic tree**
- Linnaean classification and phylogeny can differ from each other
- Systematists have proposed a classification system that would recognize only groups that include a common ancestor and all its descendants



The connection between classification and phylogeny



Order

Family

Genus

Species

Carnivora

Felidae

Mustelidae

Canidae

Panthera

Taxidea

Lutra

Canis

Panthera pardus
(leopard)

Taxidea taxus
(American badger)

Lutra lutra
(European otter)

Canis latrans
(coyote)

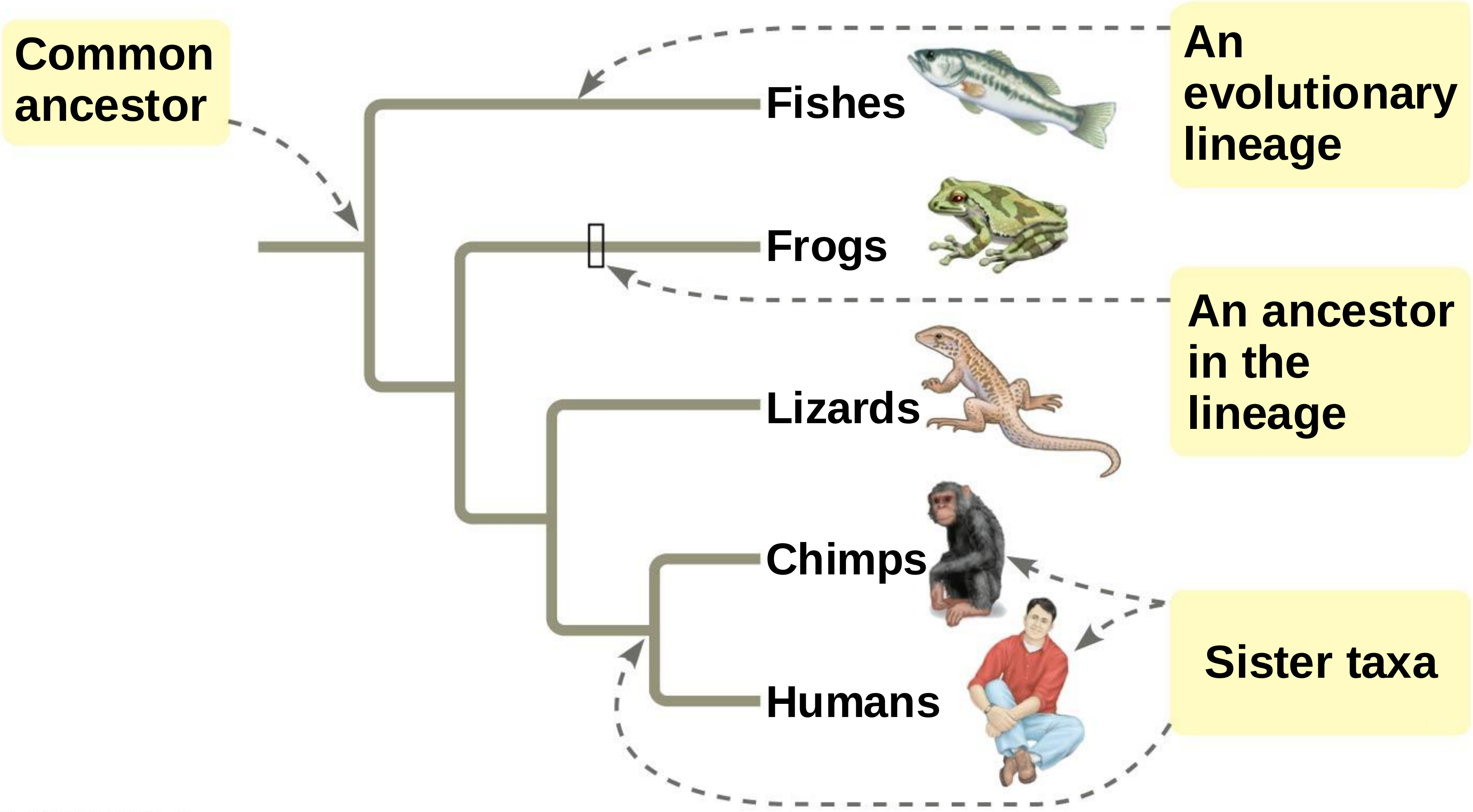
Canis lupus
(gray wolf)

What We Can and Cannot Learn from Phylogenetic Trees

- A phylogenetic tree represents a hypothesis about evolutionary relationships
- Each **branch point** represents the divergence of two evolutionary lineages from a common ancestor
- **Sister taxa** are groups that share an immediate common ancestor that is not shared by any other group

Visualizing phylogenetic relationships

Parts of a Tree



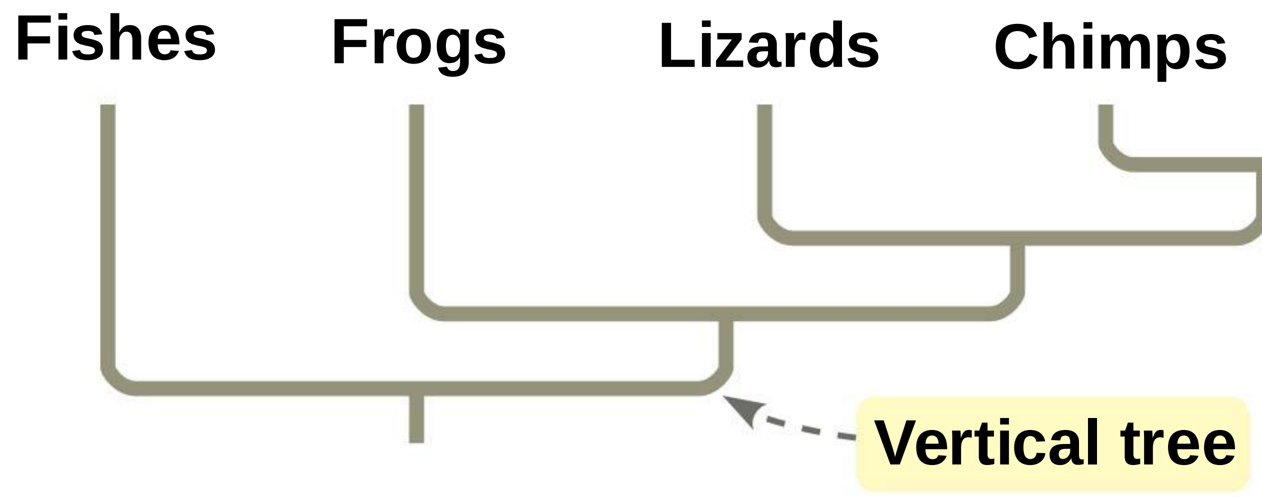
Common ancestor

An evolutionary lineage

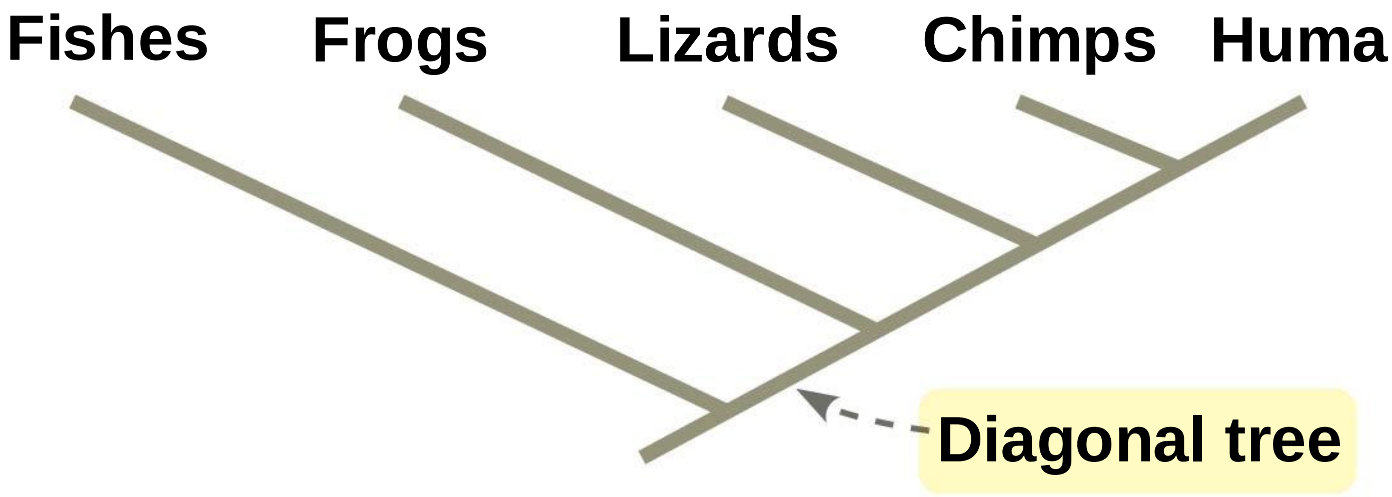
An ancestor in the lineage

Sister taxa

Alternative Forms of Tree Diagrams

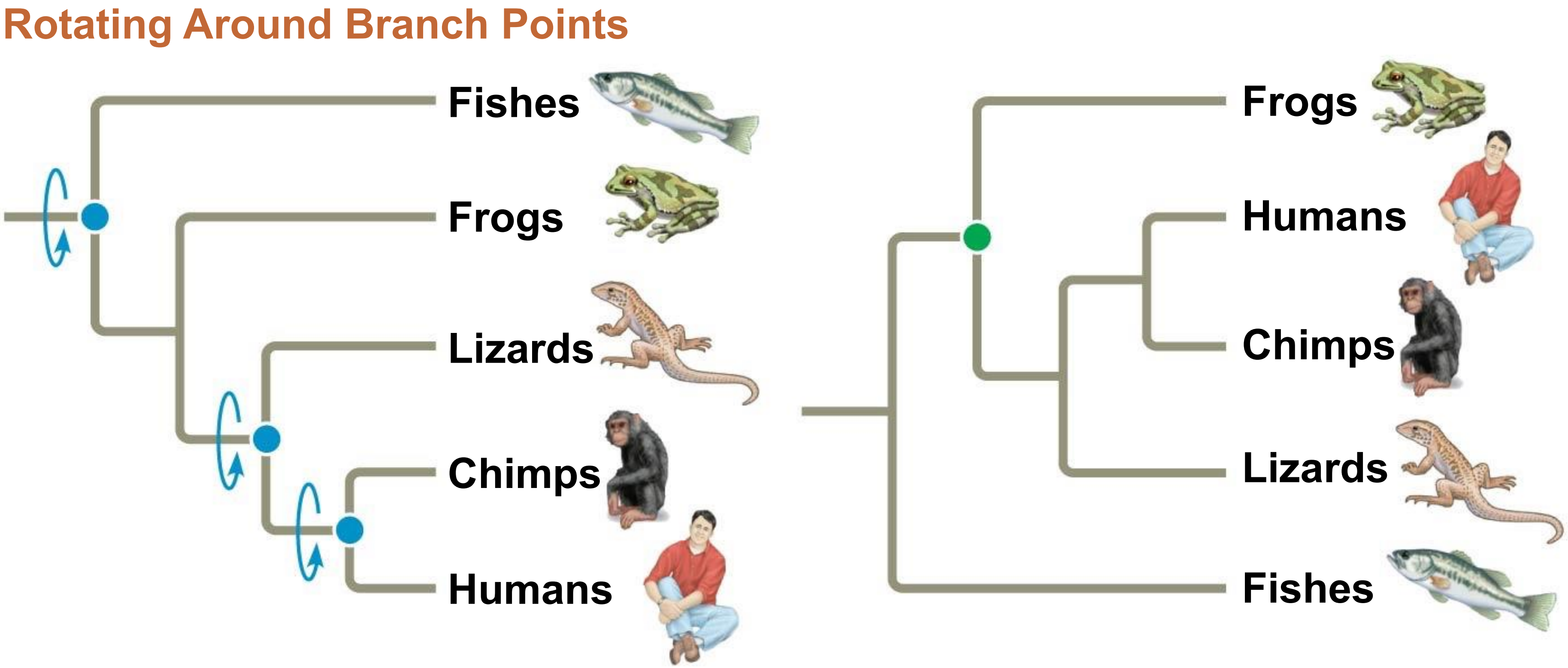


Vertical tree



Diagonal tree

- Tree branches can be rotated around a branch point without changing the evolutionary relationships



- A **rooted** tree includes a branch to represent the most recent common ancestor of all taxa in the tree
- A **basal taxon** diverges early in the history of a group and originates near the common ancestor of the group

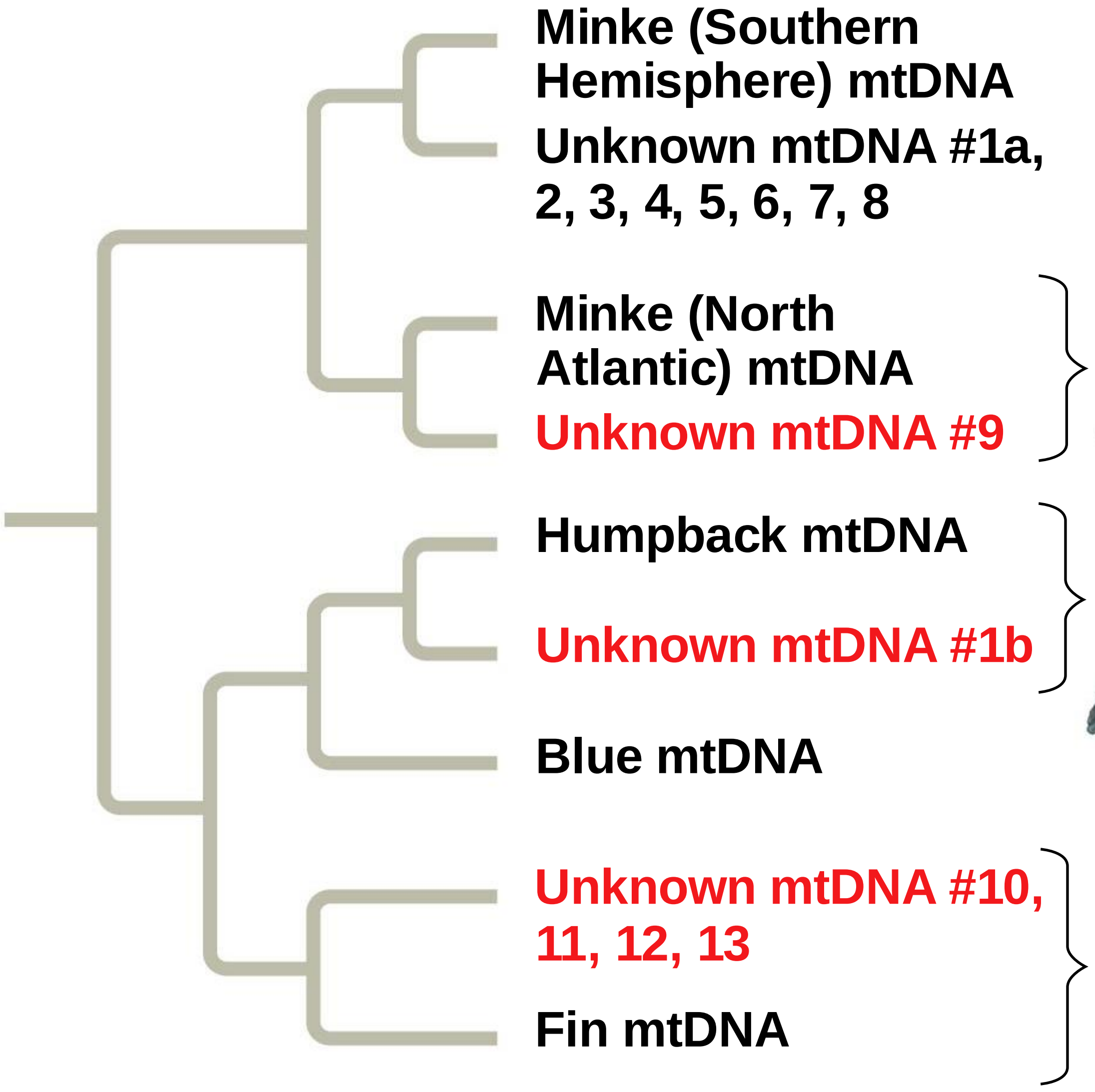
- Phylogenetic trees show patterns of descent, not phenotypic similarity
- Phylogenetic trees do not indicate when species evolved or how much change occurred in a lineage
- It should not be assumed that a taxon evolved from the taxon next to it

Applying Phylogenies

- Phylogeny provides important information about similar characteristics in closely related species
- Case study: a phylogeny was used to identify the species of whale from which “whale meat” originated to discover if the whale was harvested illegally

What is the species identity of food being sold as whale meat?

Results



Minke (Southern Hemisphere) mtDNA

Unknown mtDNA #1a, 2, 3, 4, 5, 6, 7, 8

Minke (North Atlantic) mtDNA

Unknown mtDNA #9

Humpback mtDNA

Unknown mtDNA #1b

Blue mtDNA

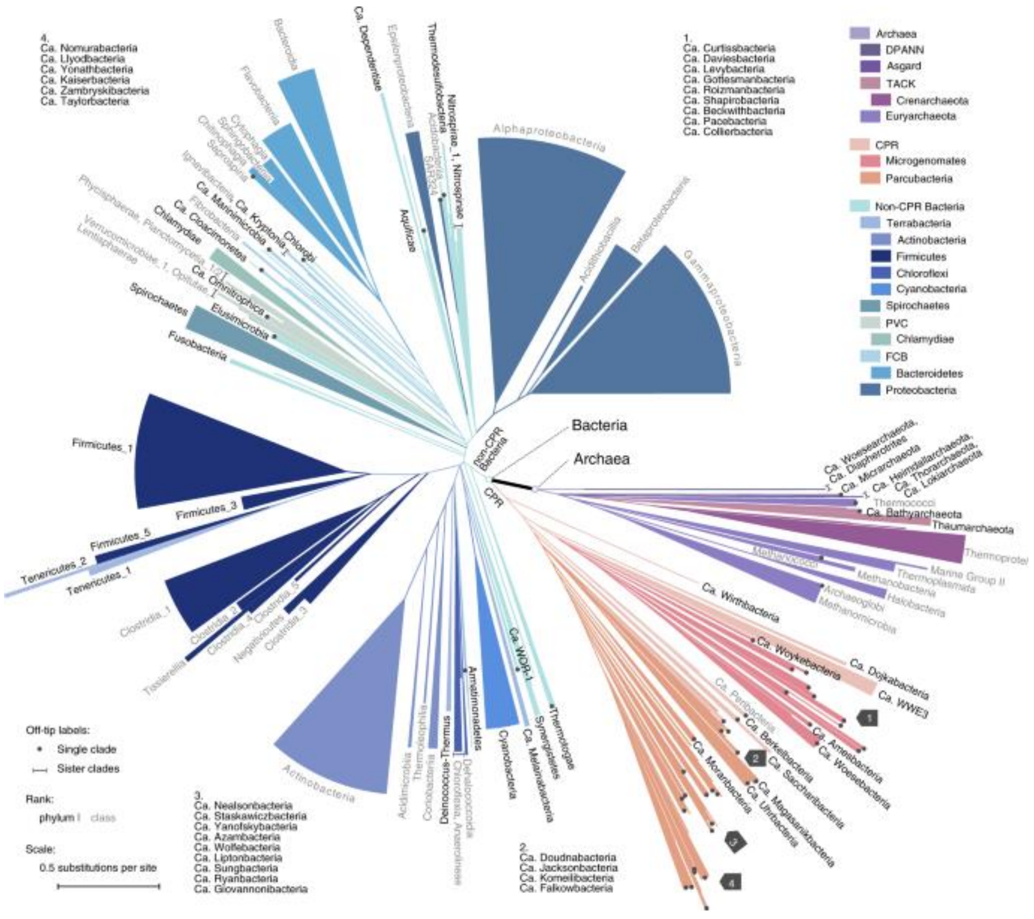
Unknown mtDNA #10, 11, 12, 13

Fin mtDNA

Data from C. S. Baker and S. R. Palumbi, Which whales are hunted? A molecular genetic approach to monitoring whaling, *Science* 265:1538–1539 (1994). Reprinted with permission from AAAS.

Phylogenies can be inferred from morphological and molecular data

- To infer phylogenies, systematists gather information about the morphologies, genes, and biochemistry of living organisms
- Can be highly complex!



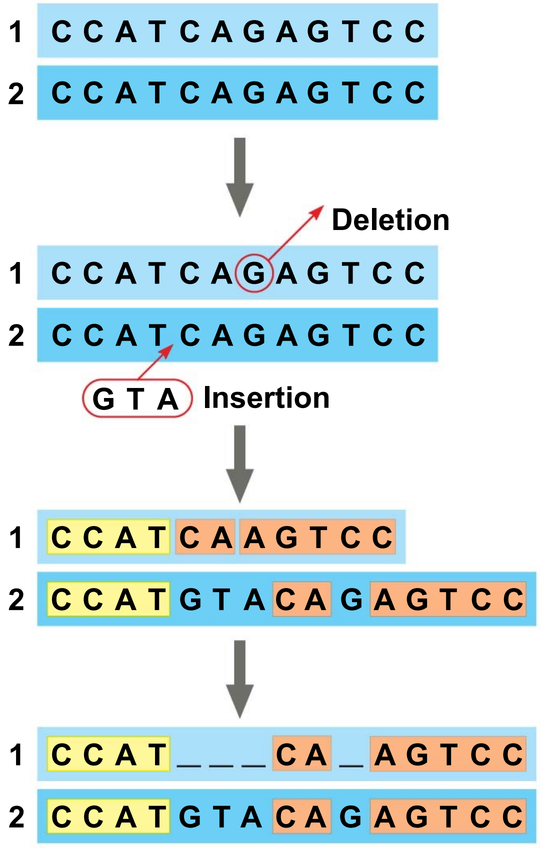
Morphological and Molecular Homologies

- Phenotypic and genetic similarities due to shared ancestry are called **homologies**
- Organisms with similar morphologies or DNA sequences are likely to be more closely related than organisms with different structures or sequences
    - Evolutionary proximity
    - Geographical proximity
    - Temporal “proximity”

Evaluating Molecular Homologies

- Systematists use computer programs and mathematical tools when analyzing comparable DNA segments from different organisms

Aligning segments of DNA

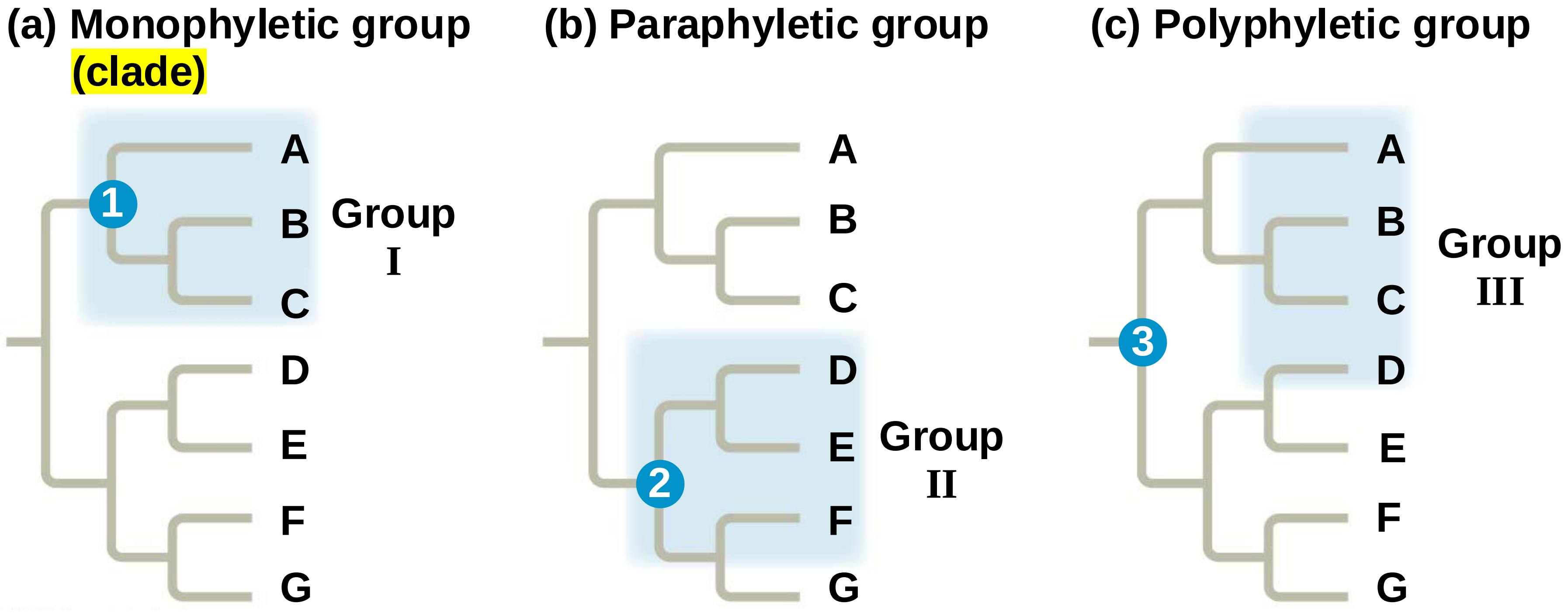


Cladistics

- **Cladistics** groups organisms by common ancestry
- A **clade** is a group of species that includes an ancestral species and all its descendants
- Clades can be nested in larger clades, but not all groupings of organisms qualify as clades

- A valid clade is **monophyletic**, signifying that it consists of the ancestor species and all its descendants
- A **paraphyletic** grouping consists of an ancestral species and some, but not all, of the descendants
- A **polyphyletic** grouping includes distantly related species but does not include their most recent common ancestor

Monophyletic, paraphyletic, and polyphyletic groups

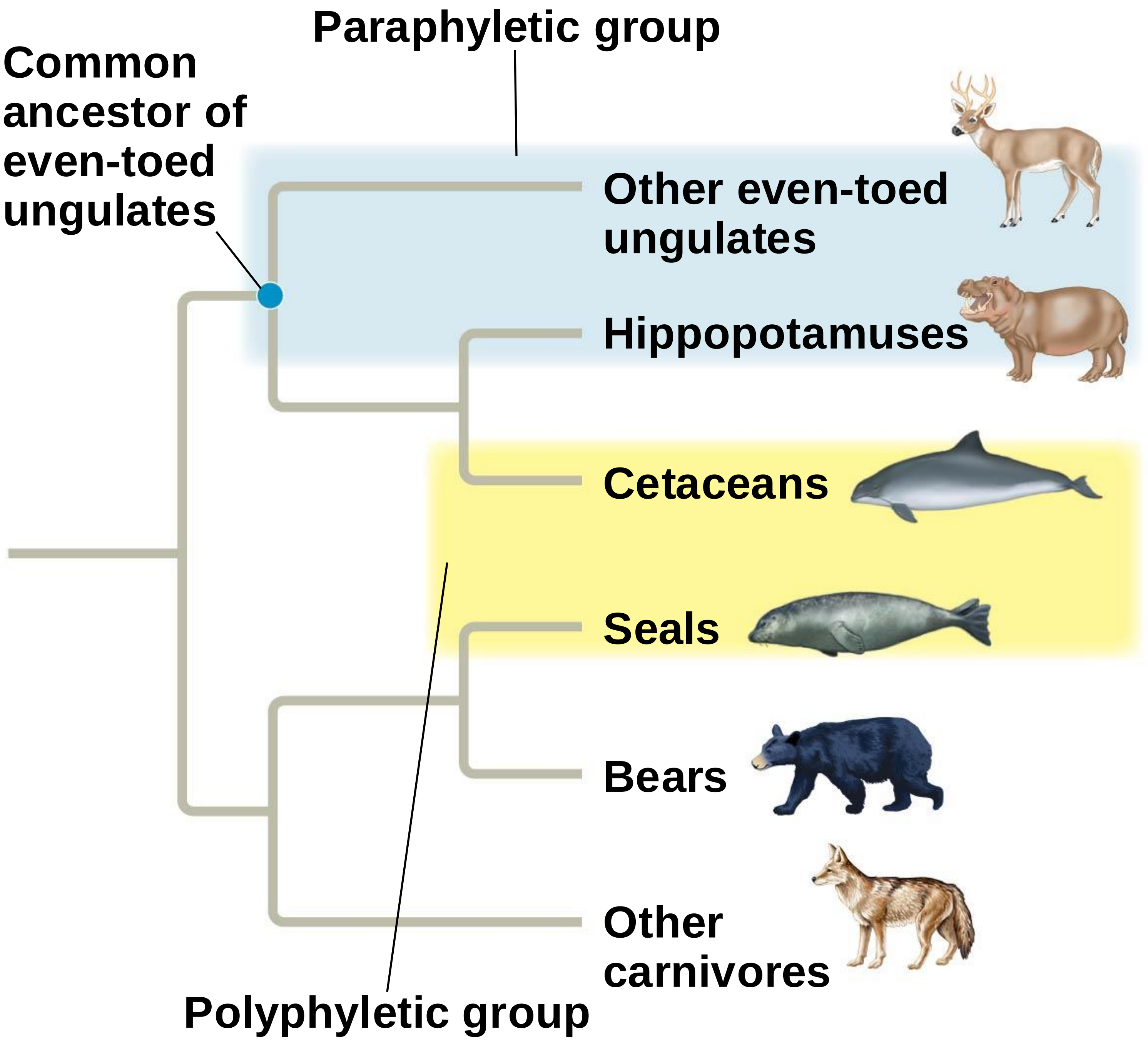


(a) Monophyletic group (clade)

(b) Paraphyletic group

(c) Polyphyletic group

- Polyphyletic groups are distinguished from paraphyletic groups by the fact that they do not include the most recent common ancestor
- Biologists avoid defining polyphyletic groups and instead reclassify organisms if evidence suggests they are polyphyletic

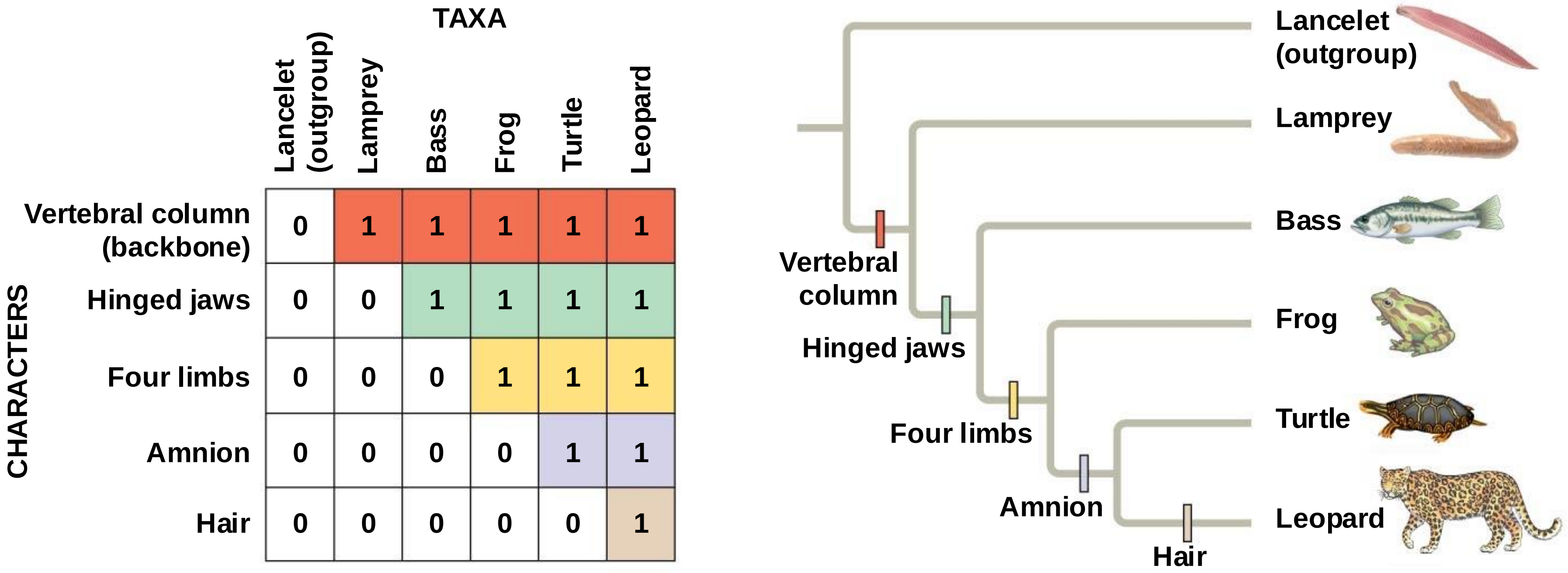

Paraphyletic group
Common ancestor of even-toed ungulates
Other even-toed ungulates
Hippopotamuses
Cetaceans
Seals
Bears
Other carnivores
Polyphyletic group

Shared Ancestral and Shared Derived Characters

- In comparison with its ancestor, an organism has both shared and different characters
- A **shared ancestral character** is a character that originated in an ancestor of the taxon
- A **shared derived character** is an evolutionary novelty unique to a particular clade
- A character can be both ancestral and derived, depending on the context

Using derived characters to infer phylogeny

When inferring evolutionary relationships, it is useful to know in which clade a shared derived character first appeared



TAXA

CHARACTERS	Lancelet (outgroup)	Lamprey	Bass	Frog	Turtle	Leopard
Vertebral column (backbone)	0	1	1	1	1	1
Hinged jaws	0	0	1	1	1	1
Four limbs	0	0	0	1	1	1
Amnion	0	0	0	0	1	1
Hair	0	0	0	0	0	1

(a) Character table

(b) Phylogenetic tree

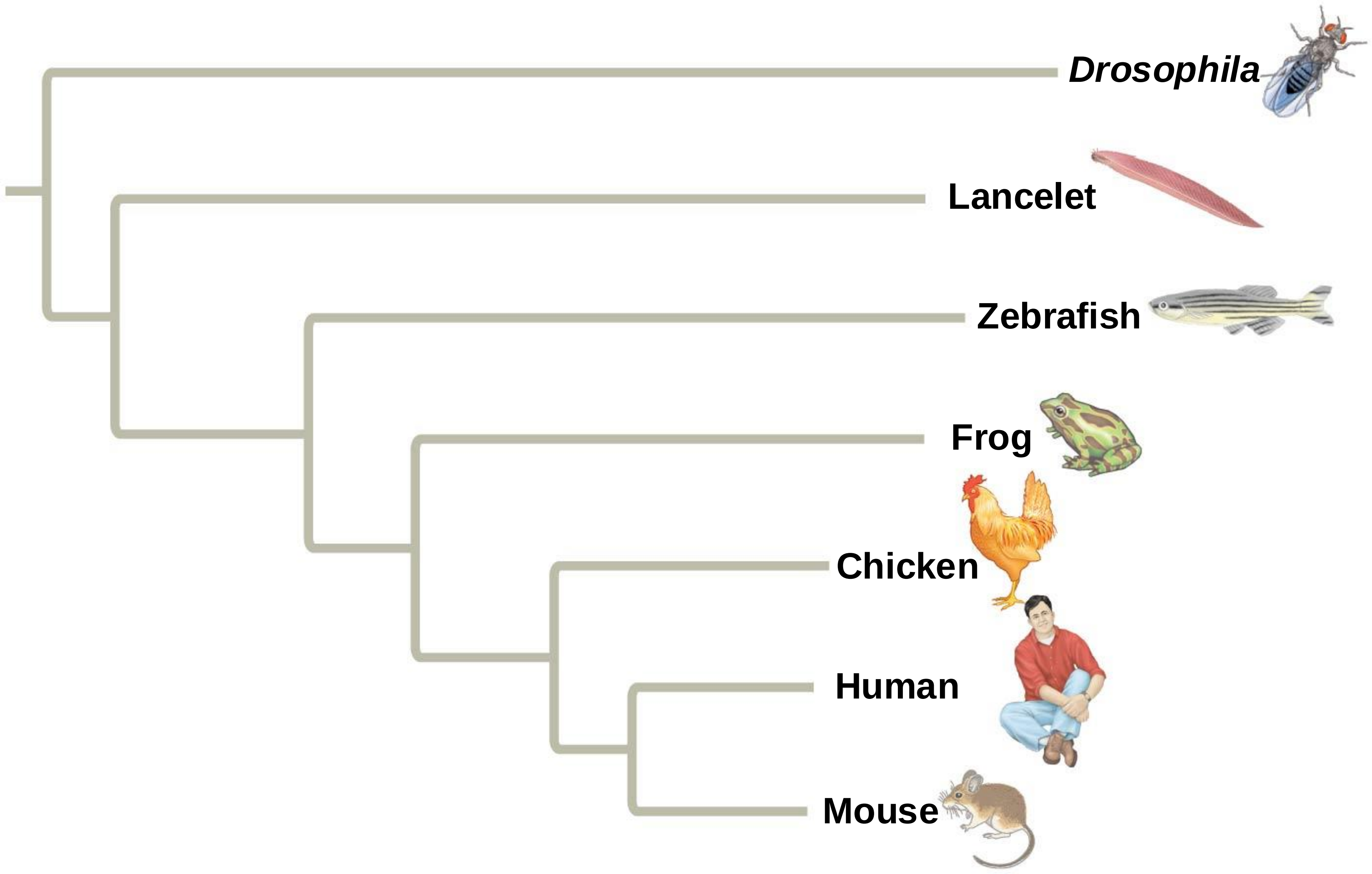
Outgroup and ingroup

- An **outgroup** is a species or group of species that is closely related to the **ingroup**, the various species being studied
- The outgroup is a group that has diverged before the ingroup
- Systematists compare each ingroup species with the outgroup to differentiate between shared derived and shared ancestral characters
- Characters shared by the outgroup and ingroup are ancestral characters that predate the divergence of both groups from a common ancestor

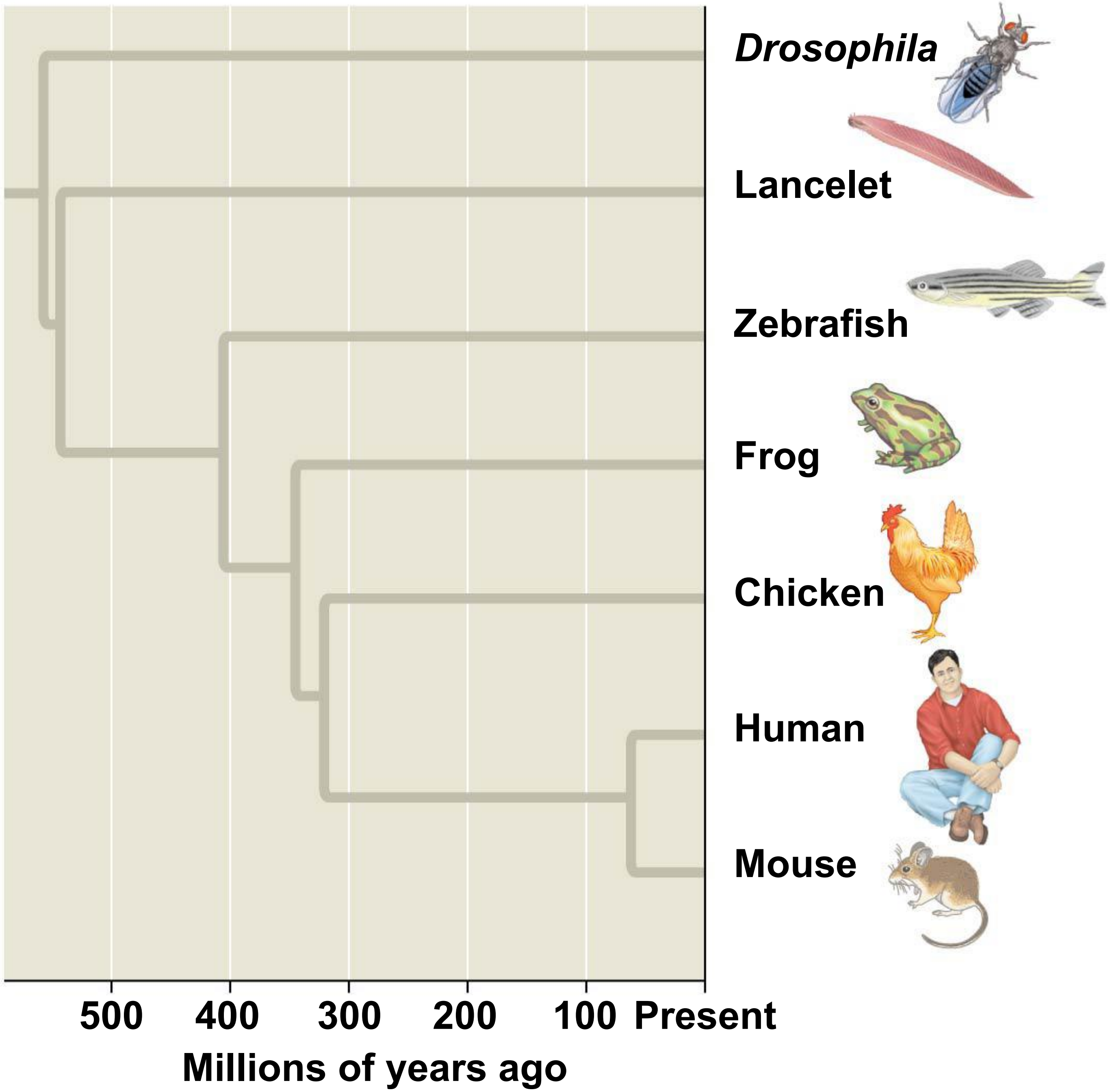
Phylogenetic Trees with Proportional Branch Lengths

- In some trees, the length of a branch can reflect the number of genetic changes that have taken place in a particular DNA sequence in that lineage

Branch lengths can represent genetic change

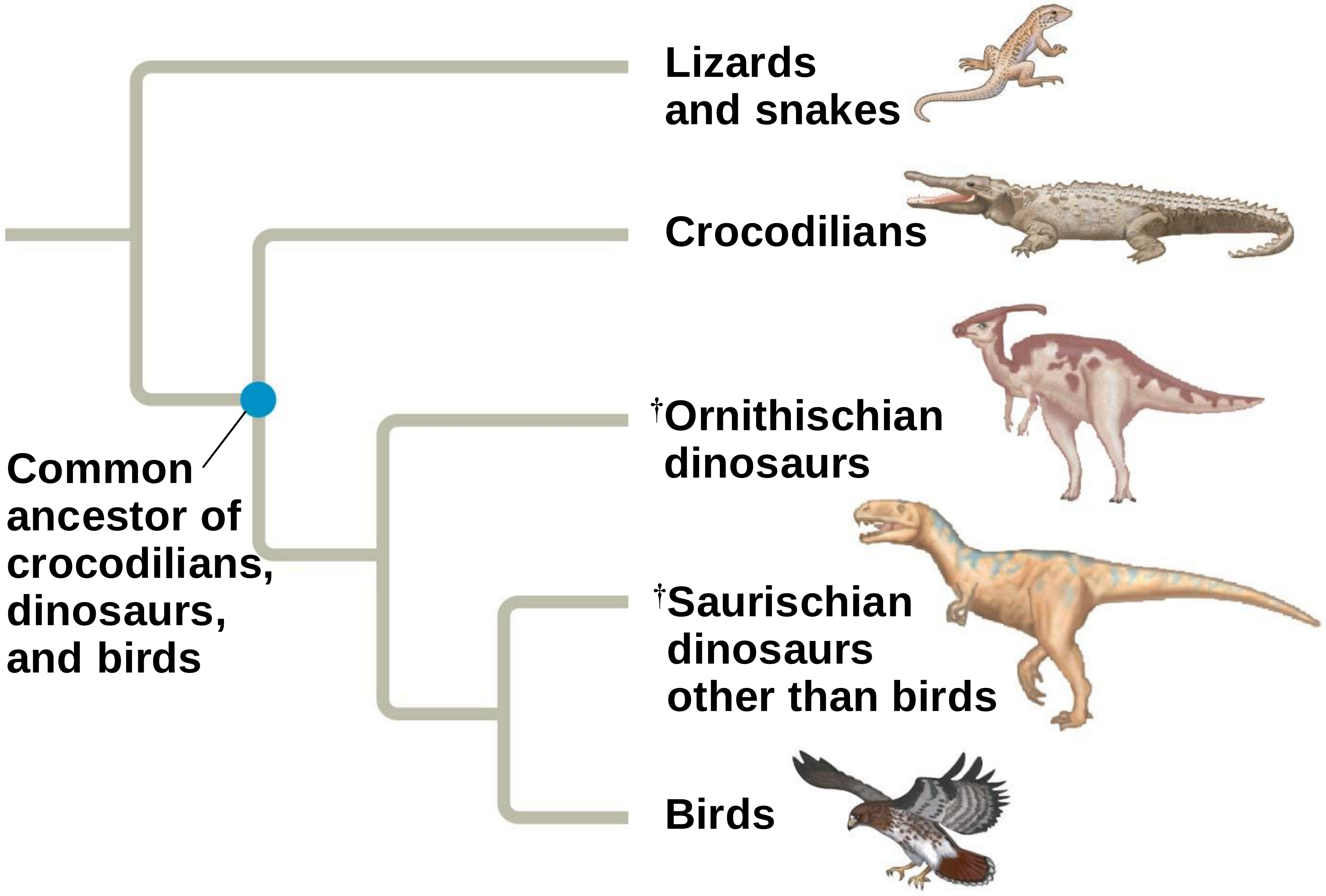


- In other trees, branch length can be proportional to time, and branching points can be determined from the fossil record



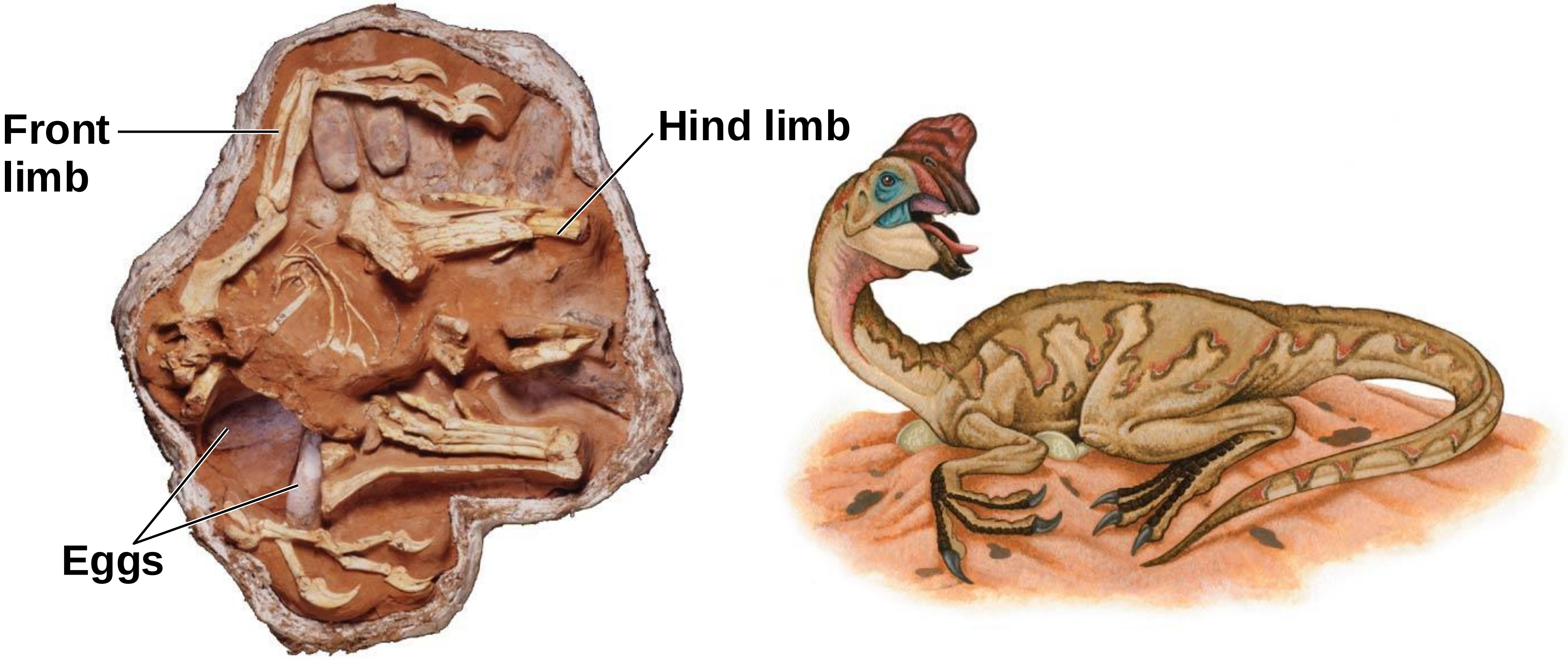
Phylogenetic Trees as Hypotheses

- The best hypothesis for a phylogenetic tree fits the most data: morphological, molecular, and fossil
- Phylogenetic bracketing allows us to predict features of an ancestor from features of its descendants
  - For example, phylogenetic bracketing allows us to infer characteristics of dinosaurs


Lizards
and snakes
Crocodilians
†Ornithischian
dinosaurs
†Saurischian
dinosaurs
other than birds
Birds
Common
ancestor of
crocodilians,
dinosaurs,
and birds

- Birds and crocodiles share several features: four-chambered hearts, song, nest building, and brooding
- These features likely evolved in a common ancestor and were shared by all of its descendants, including dinosaurs
- The fossil record supports nest building and brooding in dinosaurs

Fossil support for a phylogenetic prediction: Dinosaurs built nests and brooded their eggs



(a) Fossil remains of *Oviraptor* and eggs

(b) Artist’s reconstruction of the dinosaur’s posture based on the fossil findings

An organism’s evolutionary history is documented in its genome

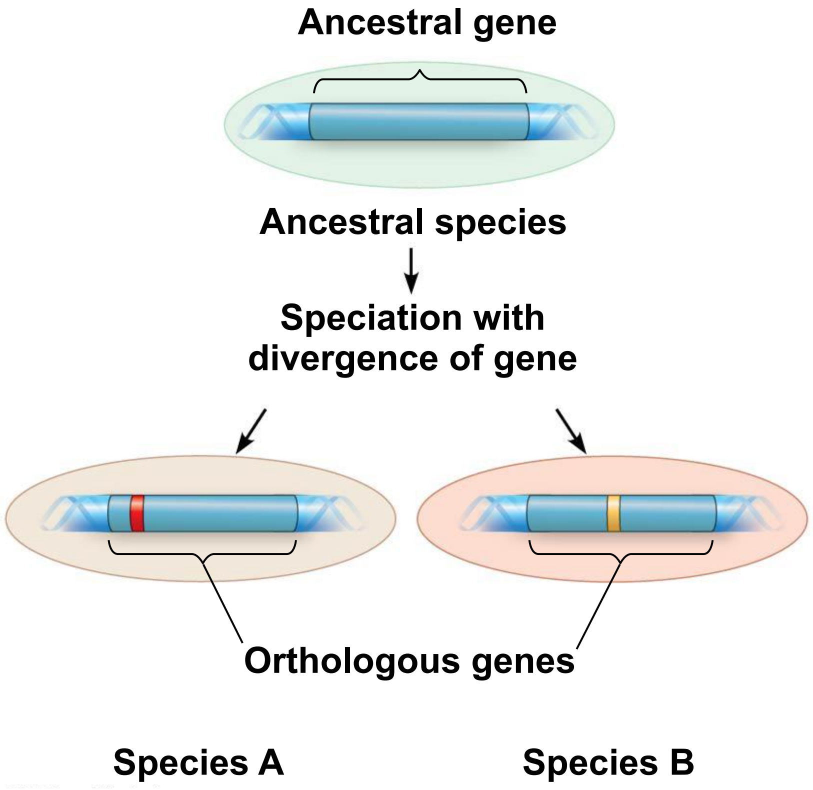
- Comparing nucleic acids or other molecules to infer relatedness is a valuable approach for tracing organisms’ evolutionary history
- DNA that codes for rRNA changes relatively slowly and is useful for investigating branching points that diverged hundreds of millions of years ago
- mtDNA evolves rapidly and can be used to explore recent evolutionary events

Gene Duplications and Gene Families

- Gene duplication increases the number of genes in the genome, providing more opportunities for evolutionary changes
- Repeated gene duplications result in gene families
- Like homologous genes, duplicated genes can be traced to a common ancestor

- **Orthologous genes** are found in a single copy in the genome and are homologous between species
- They can diverge only after speciation occurs

(a) Formation of orthologous genes: a product of speciation



Ancestral gene

Ancestral species

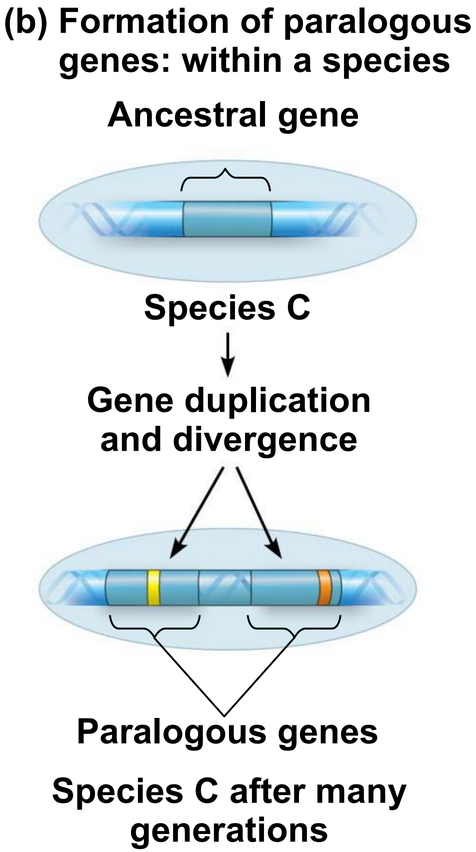
Speciation with divergence of gene

Orthologous genes

Species A

Species B

- **Paralogous genes** result from gene duplication so are found in more than one copy in the genome
- They can diverge within the species that carries them and often evolve new functions

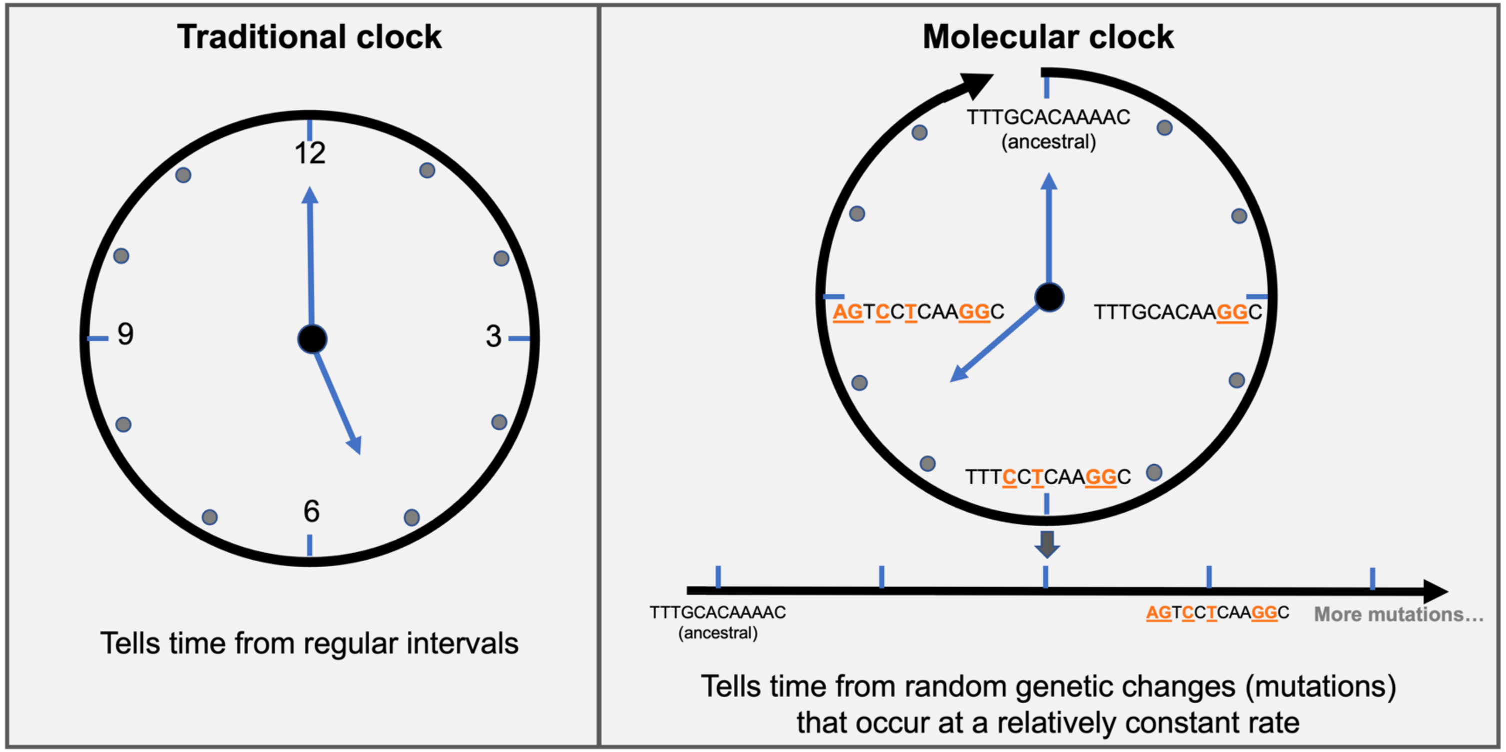


Genome Evolution

- Orthologous genes are widespread and extend across many widely varied species
  - For example, humans and mice diverged about 65 million years ago, and 99% of our genes are orthologous
- Gene number and the complexity of an organism are not strongly linked
  - For example, humans have only four times as many genes as yeast, a single-celled eukaryote
- Genes in complex organisms appear to be very versatile, and each gene can encode multiple proteins that perform many different functions

Molecular clocks help track evolutionary time

- To extend phylogenies beyond the fossil record, we must make an assumption about how molecular change occurs over time



Traditional clock

Tells time from regular intervals

Molecular clock

Tells time from random genetic changes (mutations) that occur at a relatively constant rate

Molecular Clocks

- A **molecular clock** uses constant rates of evolution in some genes to estimate the absolute time of evolutionary change
- In orthologous genes, nucleotide substitutions are assumed to be proportional to the time since they last shared a common ancestor
- In paralogous genes, nucleotide substitutions are proportional to the time since the genes became duplicated

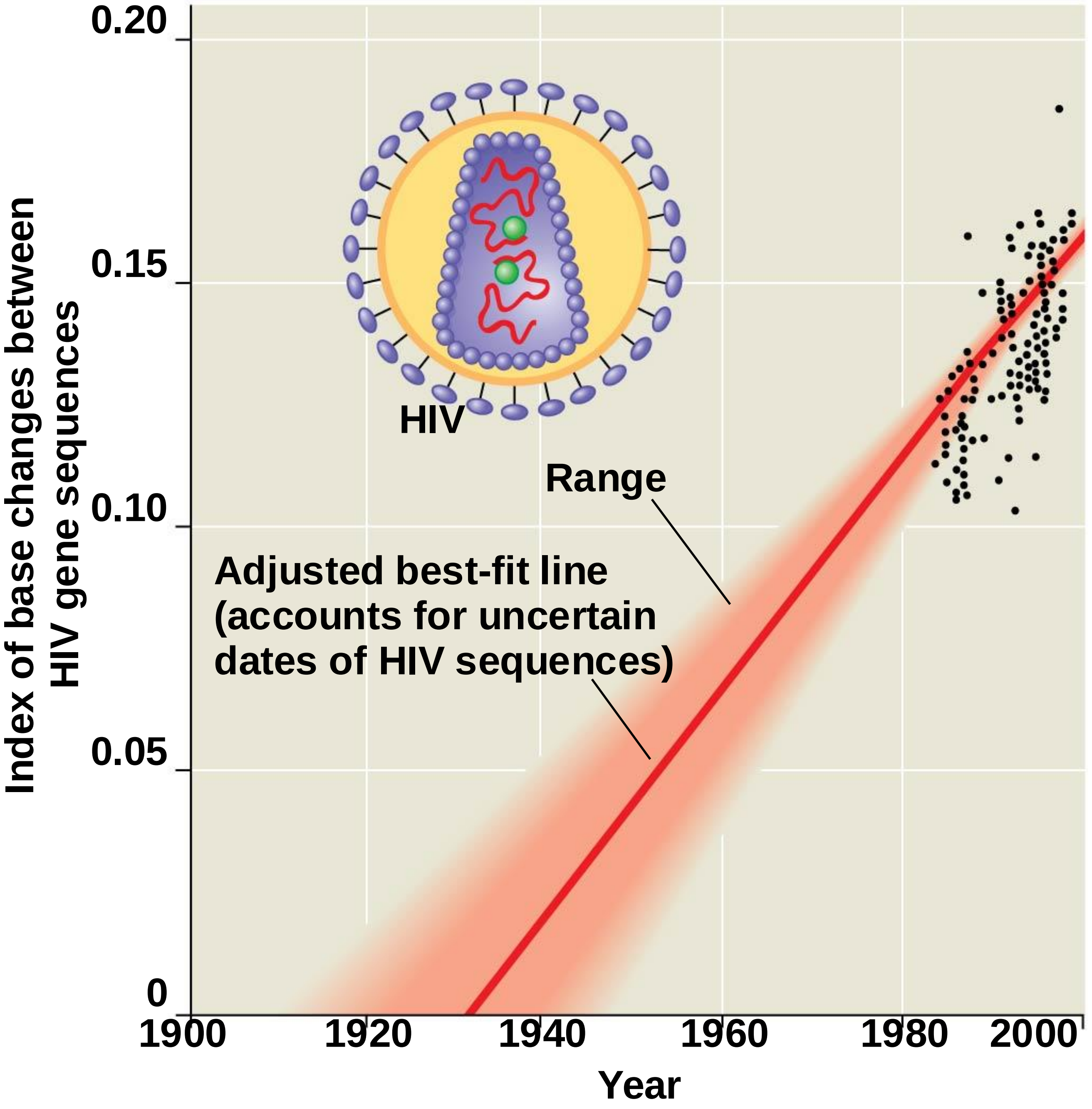
Differences in Clock Speed

- If most of the evolutionary change in genes and proteins has no effect on fitness, then the rate of molecular change should be regular, like a clock
- Differences in clock rate for different genes are a function of the importance of the gene and how critical the specific amino acid is to protein function

Potential Problems with Molecular Clocks

- The molecular clock doesn't always tick at a constant rate because natural selection favors some mutations over others, making older evolutionary estimates uncertain, but using multiple or diverse genes can help improve accuracy.

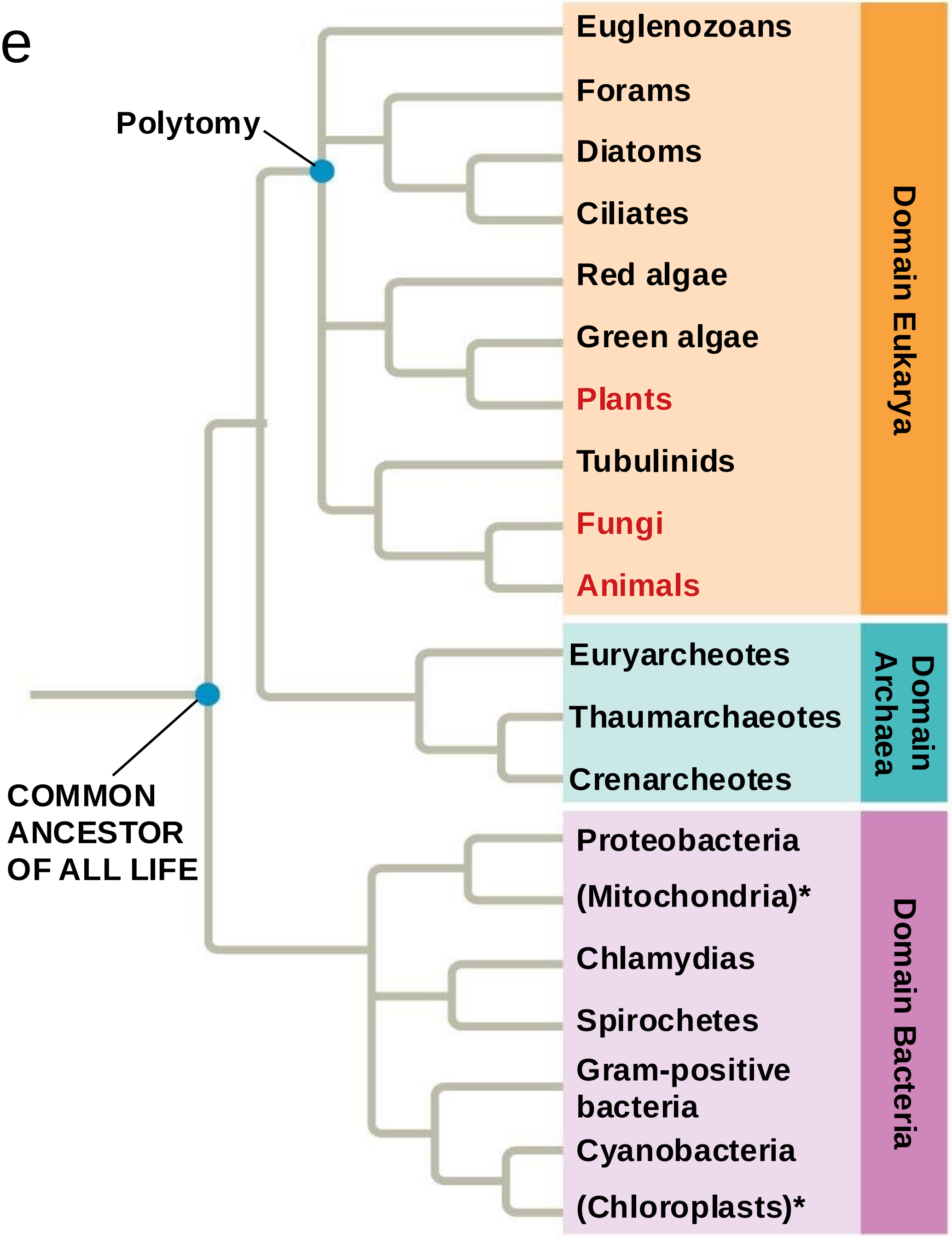
Dating the origin of HIV-1 M



From Two Kingdoms to Three Domains

- Our understanding of the tree of life continues to change based on new data. Recently, we have gained insight into the very deepest branches of the tree of life through molecular systematics.
- Early taxonomists classified all species as either plants or animals
- Later, five kingdoms were recognized: Monera (prokaryotes), Protista, Plantae, Fungi, and Animalia
- More recently, the three-domain system has been adopted: Bacteria, Archaea, and Eukarya
- The three-domain system is supported by data from many sequenced genomes

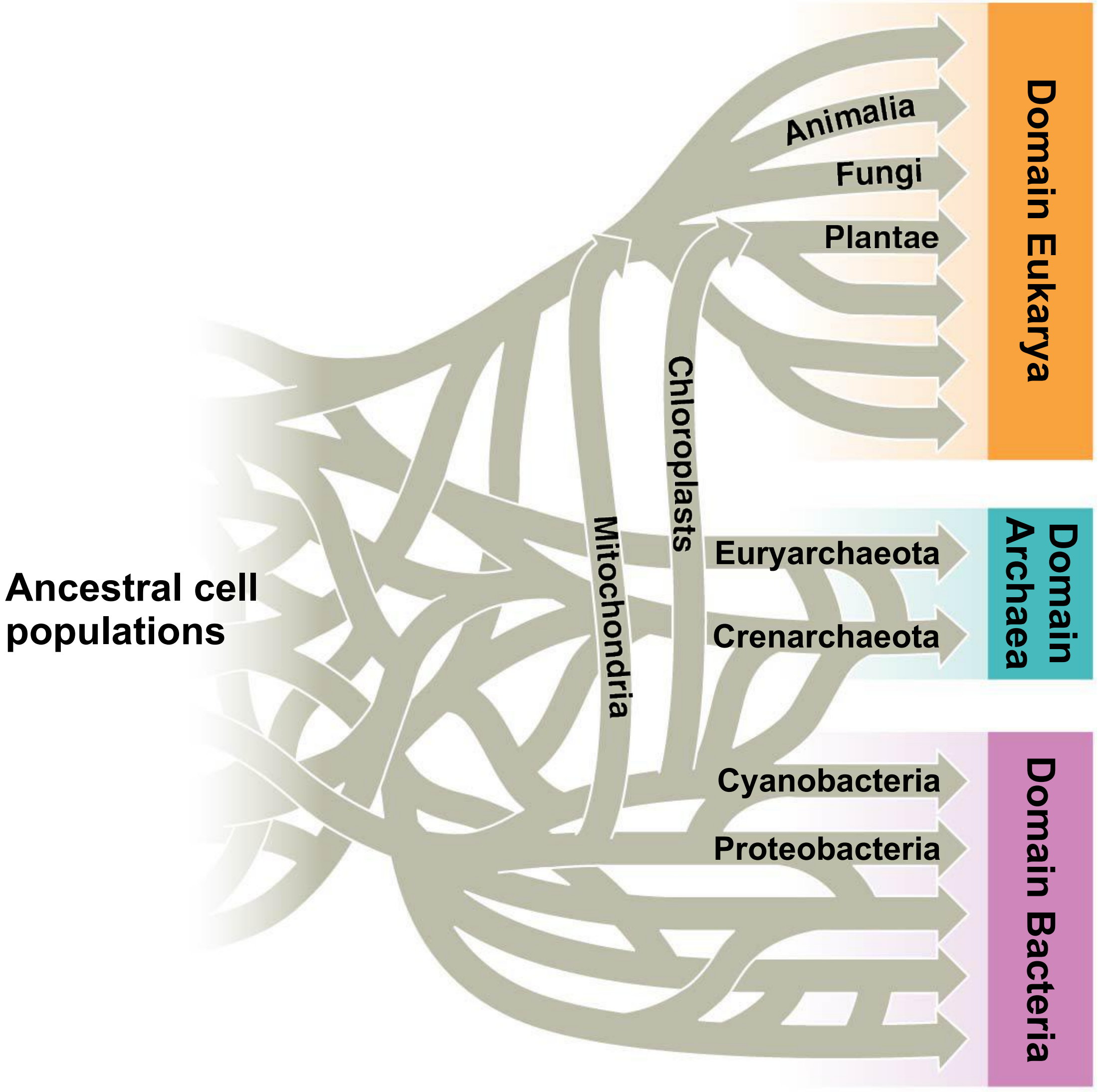
The three domains of life



Horizontal Gene Transfer and the tree of life

- The tree of life shows that eukaryotes are more closely related to archaea than to bacteria, based mainly on rRNA genes.
- However, gene trees sometimes differ due to **horizontal gene transfer (HGT)**—the movement of genes between unrelated organisms.
- HGT occurs via plasmids, viruses, transposable elements, and even cell fusion.
- These gene exchanges blur evolutionary relationships and have shaped both prokaryotic and eukaryotic genomes.
- Example: The alga *Galdieria sulphuraria* gained ~5% of its genes from bacteria and archaea.

- Some biologists argue that horizontal gene transfer was so common that the early history of life should be represented as a tangled network of connected branches



LUCA - Last Universal Common Ancestor

- Originated billions of years ago
- Hypothetical ancestor of all life on Earth
- Simple single-celled organism
- Common ancestor of all living organisms
- Every living organism on Earth, from microbes to plants to animals, can trace its ancestry back to LUCA.
- Crucial for understanding life's origins and evolution