

Today's class

Hardy-Weinberg

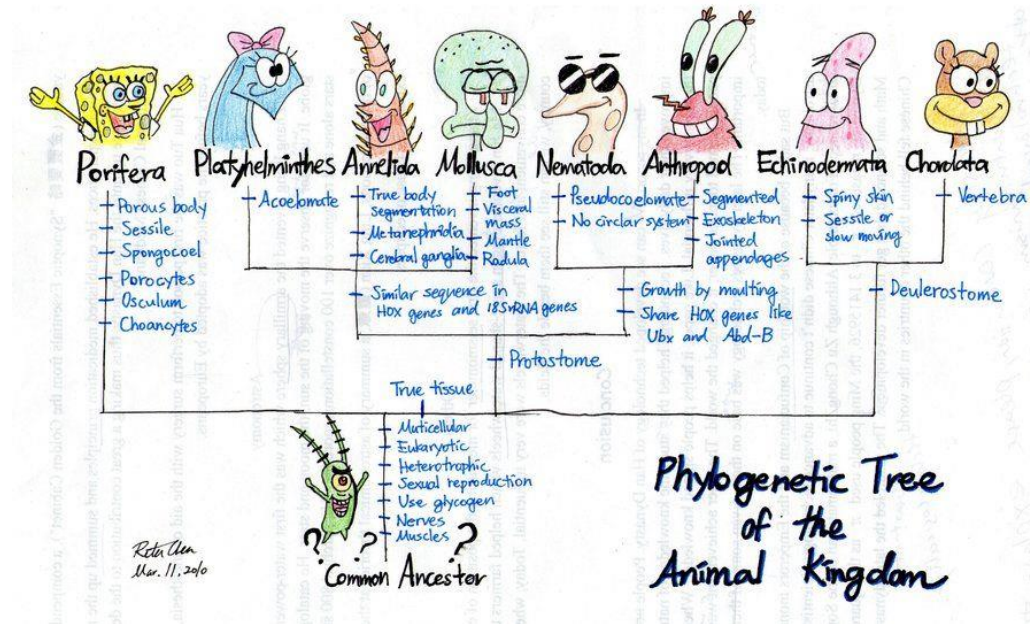
Phylogenetic Trees (start)

$$p^2 + 2pq + q^2 = 1$$

p^2 = dominant homozygous frequency (AA)

$2pq$ = heterozygous frequency (Aa)

q^2 = recessive homozygous frequency (aa)



<https://www.pinterest.ca/pin/phylogeny-with-spongebob-by-thehurricanes-on-deviantart--146437425357543433/>

Outline of today's lecture

A. Hardy-Weinberg Equilibrium (testable on midterm #2)

- How to calculate allele frequencies
- How to calculate observed genotype frequencies
- How to calculate expected genotype frequencies
- How to determine if evolution is acting on a gene in a population or not
- Assumptions that must be met for population to be in HWE for a gene.

B. Phylogenetic Trees (testable on midterm #2) – start today – finish on Thursday

- How to read a tree
 - Monophyletic, Polyphyletic and Paraphyletic groups (snip test)
 - Synapomorphies (and monophyletic groups)
- Predicting number of times trait has been gained or lost (principle of parsimony)
- Identifying homologous and analagous traits
- Identifying the most closely related taxa
- Determine if two trees show the same evolutionary relationships or not
- Common mistakes/misconceptions

Housekeeping – Midterm #2 (20%)

One week from today (20%). Evolution Unit.

- 50 minutes (it cannot be longer because Brett's classes are 50 minutes in length)
- number of questions – not yet finalized (likely 4 + bonus). I will know more by Thursday
- 4 testable topics:
 - Evidence of evolution (*i.e. that all species are descended from LUCA, species are related, evolutionary change over time*)
 - Four Mechanisms of Evolution
 - HWE
 - Phylogenetic Trees
- Species concepts (covered in class on Thursday) will not be tested until the final exam.
- Bonus question will involve an Organism of the Day (choice).
- Study sheet allowed – 8.5 x 11", two-sided, hand-written.
- Calculator allowed (needed).

Ruby & Christie will be providing a Zoom review session this week – please stay tuned for details.

Brett will also be providing a midterm study session – **Friday, March 10th 5:00-6:50pm in SCRF 100**

Due this coming Sunday @ 11:59 pm

Hardy-Weinberg (can be completed by the end of today's class)

- Quiz 7 – Evolution: Population Genetics (HWE)
- Worksheet #8 - HWE – Sea Otters

CHANGE FROM PRE-LECTURE SLIDES – Due, Sunday, March 19th (but highly recommend you attempt before midterm).

Phylogenetic Trees (can be completed by the end of Thursday's class):

- Quiz 8 – Reading Phylogenetic Trees
- Worksheet #9 – Reading Phylogenetic Trees

Population genetics

How do we know if evolution is acting on a specific gene or not in a population?

- Compared observed genotype frequencies with expected genotype frequencies (if the population is in HWE).
- Use the equation below to calculate the EXPECTED genotype frequencies (if evolution is not acting on a gene, and the population is in HWE).
- If there is a difference between observed and expected frequencies > 0.10 for even one genotype, you can conclude that evolution is acting on the gene.

$$\begin{array}{ccc} \text{AA} & \text{Aa} & \text{aa} \\ \downarrow & \downarrow & \downarrow \\ p^2 & + & 2pq & + & q^2 = 1 \end{array}$$

Assumptions

- If a population is not in HWE, then at least one of these assumptions has been violated (all related to mechanisms of evolution).
- Random mating (with respect to gene) – so no sexual selection
 - Organisms do not select mates with respect to this gene
- The population is infinitely large
 - So, no changes in allele frequencies due to sampling error (i.e., **no Genetic Drift**)
- No natural selection
- No gene flow
- No mutations

A non-testable point about HWE equation

$$.p^2 + 2pq + q^2$$

In the Genetics Unit, you used a Punnett Square to make prediction about offspring genotype and phenotypes from a cross between two individuals of known genotypes, right?

For example, if you crossed 2 heterozygotes ($Aa \times Aa$) you would make the following predictions.

Male, A/a		Female, A/a	
		A ($\frac{1}{2}$)	.a ($\frac{1}{2}$)
	A ($\frac{1}{2}$)	AA $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$	Aa $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$
	.a ($\frac{1}{2}$)	Aa $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$.aa $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

$$.p^2 + 2pq + q^2$$

In population genetics, you need to consider allele frequencies for the entire population (not just two individuals).

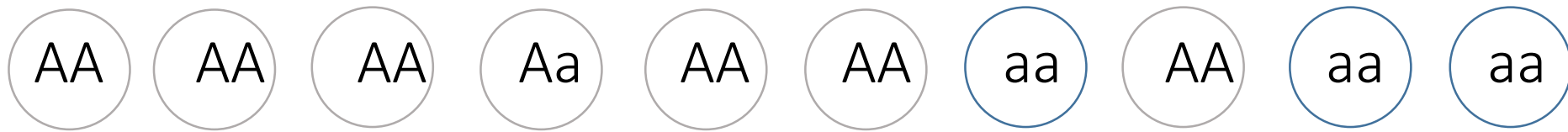
Example: If a population had allele frequencies of: $A = 0.65$, and $.a = 0.35$

- same frequencies in egg and sperm, and gametes unite at random; probability of producing each genotype:

	A (0.65)	$.a$ (0.35)
A (0.65)	AA $.p^2$ $0.65 \times 0.65 = 0.423$	Aa $.p \times q$ $0.65 \times 0.35 = 0.228$
$.a$ (0.35)	Aa $.p \times q$ $0.65 \times 0.35 = 0.228$	$.aa$ $.q^2$ $0.35 \times 0.35 = 0.123$

Step 1: Calculate the observed (real) genotype frequencies

For example: If there were 10 diploid individuals in a population



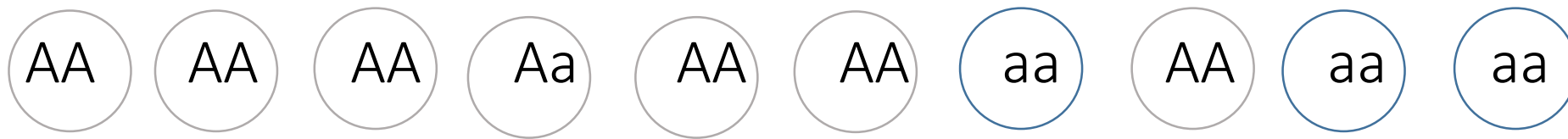
What is the frequency of AA genotype/individuals?

What is the frequency of Aa genotype/individual?

What is the frequency of aa genotype/individuals?

Step 1: Calculate observed (real) genotype frequencies

In this example there were 10 diploid individuals in a population



What is the frequency of AA individuals? $6/10 = 0.6$

What is the frequency of Aa individual? $1/10 = 0.1$

What is the frequency of aa individuals? $3/10 = 0.3$

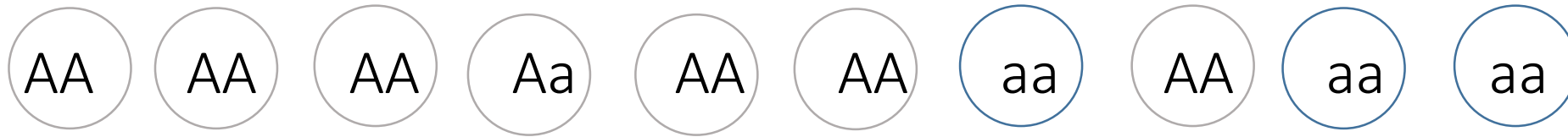
check arithmetic ($0.6 + 0.1 + 0.3 = 1$)

Step 2: Calculate the allele frequencies

p = freq. of A allele

q = frequency of a allele

$p+q=1$



Total number of alleles for the A gene in the population (diploid)?

Number of A alleles in population?

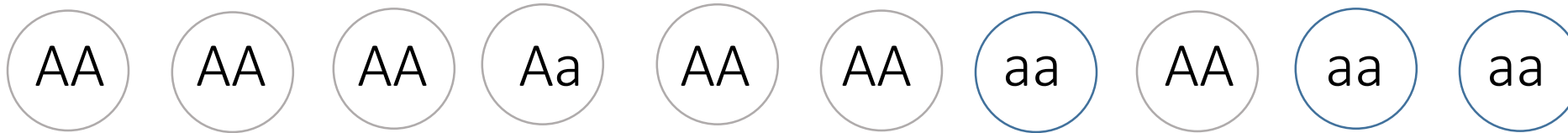
Frequency of p (A)?

Number of a alleles in population?

Frequency of q (a)?

Step 2 – calculating allele frequencies in population

Step 2: Calculate the allele frequencies in the population



Total number of alleles for the A gene in the population (diploid) = $10 \times 2 = 20$

Number of A alleles in population = 13 Frequency of A allele (p) = $(13) / 20 = 0.65$

Number of a alleles in population = 7 Frequency of a allele (q) = $(7) / 20 = 0.35$

- 6 individuals = AA $\longrightarrow 6 \times 2 = 12$ A alleles
- 1 individual = Aa $\longrightarrow 1$ A allele and 1 a allele
- 3 individuals = aa $\longrightarrow 3 \times 2 = 6$ a alleles

$$p + q = 1$$

Another way to calculate allele frequencies

Using observed genotype frequencies are:

$$AA = 0.6$$

$$Aa = 0.1$$

$$aa = 0.3$$

$$\text{Frequency of A allele (p)} = 0.6 + (0.1/2) = 0.6 + 0.05 = 0.65$$

$$\text{Frequency of a allele (q)} = 0.3 + (0.1/2) = 0.3 + 0.05 = 0.35$$

Step 3: Calculate the EXPECTED genotype frequencies

- Using allele frequencies

Genotype:	AA		Aa		aa
	↓		↓		↓
Expected Frequency:	p^2	+	$2pq$	+	$q^2 = 1$

$$p(A) = 0.65 \quad q(a) = 0.35$$

	Expected Frequencies
AA – p^2	
Aa – $2pq$	
aa – q^2	

Step 3: Calculating expected genotype frequencies

Genotypes	Expected Genotype Frequencies (if population is in HWE)
AA – p^2	$0.65^2 = 0.42$
Aa – $2pq$	$2(0.65)(0.35) = 0.46$
aa – q^2	$0.35^2 = 0.12$

$$p(A) = 0.65 \quad q(a) = 0.35$$

Check arithmetic: $0.42 + 0.46 + 0.12 = 1$
--

Step 4: Compare observed and expected values

	Expected Frequencies	Observed frequencies
$AA - p^2$	$0.65^2 = 0.42$	$6/10 = 0.60$
$Aa - 2pq$	$2(0.65)(0.35) = 0.46$	$1/10 = 0.10$
$aa - q^2$	$0.35^2 = 0.12$	$3/10 = 0.30$

Are ALL of the expected and observe genotype frequencies close to each other (i.e. ALL within 0.10)? If they are all close this gene is in HWE. Evolution is not acting on this gene.

If at least one of the expected and observed genotype frequencies are different by > 0.10 then the population is not in HWE for that gene.

iClicker Question

Is this population in HWE for the A gene?

	Expected Frequencies	Observed frequencies
AA – p^2	$0.65^2 = 0.42$	$6/10 = 0.60$
Aa – $2pq$	$2(0.65)(0.35) = 0.46$	$1/10 = 0.10$
aa – q^2	$0.35^2 = 0.12$	$3/10 = 0.30$

A. Yes

B. No

C. Not sure

Answer

Is this population in HWE for the A gene?

	Expected Frequencies	Observed frequencies
AA – p^2	$0.65^2 = 0.42$	$6/10 = 0.60$
Aa – $2pq$	$2(0.65)(0.35) = 0.46$	$1/10 = 0.10$
aa – q^2	$0.35^2 = 0.12$	$3/10 = 0.30$

- A. Yes
- B. No
- C. Not sure
- Observed more homozygous individuals (AA and aa) than expected if evolution not acting on gene and observed fewer heterozygotes (Aa) than expected.
- For BIOL121 – if the difference between an observed and expected frequency is greater than 0.10, it okay to assume the difference is big enough to conclude that population is not in HWE for that gene.*
- Scientists would do a statistical analysis (e.g. Chi-Square test)*

So, at least one of these assumptions has been violated

- Random mating (no sexual selection)
 - Organisms do not select mates with respect to this gene; i.e. no sexual selection
- Population is infinitely large
 - So, no changes in allele frequencies due to sampling error (genetic drift), chance events that are random with respect to R gene.
- No natural selection
- No gene flow
- No mutations

iClicker Question:

$$\text{If } p^2 + 2pq + q^2 = 1$$

Does that mean the population is in Hardy-Weinberg equilibrium?

A. Yes

B. No

Answer

$$p^2 + 2pq + q^2 = 1$$

Does that mean the population is in Hardy-Weinberg equilibrium?

A. Yes

B. No

It just means that the arithmetic is correct.

To determine if a population is in HWE for a specific gene or not, you must compare observed genotype frequencies and expected genotype frequencies.

iClicker Question

If $p + q = 1$, does that mean the population is in Hardy-Weinberg equilibrium?

A. Yes

B. No

Answer

If $p+q=1$,

does that mean the population is in Hardy-Weinberg equilibrium?

A. Yes

B. No

It just means that the arithmetic is correct.

Example Question

In a grassy field in Pacific Spirit Park near UBC, there is a population of isolated crickets. Crickets are small invertebrates, which are eaten by a wide variety of visual predators, including amphibians, reptiles, birds and mammals. The crickets in this grassy field have three different body colours, which are governed by a single pair of alleles with a non-dominant relationship:

Crickets that are homozygous for the R_1 allele are yellow.

Crickets that are homozygous for the R_2 allele are blue.

Heterozygotes are green.

In this cricket population, there are 400 individuals. 35 crickets are yellow (R_1R_1), 265 crickets are green (R_1R_2), and 100 crickets are blue (R_2R_2).

iClicker Question

What are the observed genotype frequencies for R1R1, R1R2 genotypes in this population of crickets?

- A. R1R1 = 0.663 R1R2 = 0.089
- B. R1R1 = 0.070 R1R2 = 0.530
- C. R1R1 = 0.088 R1R2 = 0.662
- D. Not sure

Genotype	Phenotype	# of Individuals	Observed Genotype Frequency
R1R1	Yellow	35	
R1R2	Green	265	
R2R2	Blue	100	0.250

Answer

What are the observed genotype frequencies for R1R1, R1R2 genotypes in this population of crickets?

- A. R1R1 = 0.663 R1R2 = 0.089
- B. R1R1 = 0.070 R1R2 = 0.530
- C. R1R1 = 0.088 R1R2 = 0.662
- D. Not sure

Genotype	Phenotype	# of Individuals	Observed Genotype Frequency
R1R1	Yellow	35	0.088 (35/400)
R1R2	Green	265	0.662 (265/400)
R2R2	Blue	100	0.250 (100/400)

$$0.088 + 0.662 + 0.250 = 1 \text{ (arithmetically correct)}$$

iClicker Question

What are the frequencies of the R1 (p) and R2 (q) alleles in the population?

- A. R1 (p) = 0.471 R2 (q) = 0.529
- B. R1 (p) = 0.419 R2 (q) = 0.581
- C. R1 (p) = 0.500 , R2 (q) = 0.500
- D. Not sure

Genotype	Phenotype	# of Individuals	Observed Genotype Frequency
R1R1	Yellow	35	0.088
R1R2	Green	265	0.662
R2R2	Blue	100	0.250

Answer

What are the frequencies of the R1 (p) and R2 (q) alleles in the population?

A. R1 (p) = 0.471 R2 (q) = 0.529

B. R1 (p) = 0.419 R2 (q) = 0.581

C. R1 (p) = 0.500 , R2 (q) = 0.500

D. Not sure

$$\text{Freq. of R1 (p)} = \frac{(35 \times 2) + 265}{800} = 0.419$$

$$\text{Freq. of R2 (q)} = 1 - 0.419 = 0.581$$

Genotype	Phenotype	# of Individuals	Observed Genotype Frequency
R1R1	Yellow	35	0.088
R1R2	Green	265	0.662
R2R2	Blue	100	0.250

Check arithmetic:
 $0.419 + 0.581 = 1$

Question (not iClicker)

What are the expected/predicted genotype frequencies for the R gene in this population, if this population is in HWE?

.f of R1 (p) = 0.419
.f of R2 (q) = 0.581

$$\begin{array}{ccccc} \text{R1R1} & & \text{R1R2} & & \text{R2R2} \\ p^2 & + & 2pq & + & q^2 = 1 \end{array}$$

Genotype	Phenotype	# of Individuals	Observed Genotype Frequencies	Expected Genotype Frequencies
R1R1	Yellow	35	0.088	
R1R2	Green	265	0.662	
R2R2	Blue	100	0.250	

Answer

What are the expected/predicted genotype frequencies for the R gene in this population, if this population is in HWE?

.f of R1 (p) = 0.419
.f of R2 (q) = 0.581

$$\begin{array}{ccccc} \text{R1R1} & & \text{R1R2} & & \text{R2R2} \\ p^2 & + & 2pq & + & q^2 = 1 \end{array}$$

Genotype	Phenotype	# of Individuals	Observed Genotype Frequencies	Expected Genotype Frequencies
R1R1	Yellow	35	0.088	0.175
R1R2	Green	265	0.662	0.487
R2R2	Blue	100	0.250	0.338

Check arithmetic: $0.175 + 0.487 + 0.338 = 1$

Question (not iClicker)

Is this population in HWE? Yes or no.

.f of R1 (p) = 0.419
.f of R2 (q) = 0.581

$$\begin{array}{ccccc} \text{R1R1} & & \text{R1R2} & & \text{R2R2} \\ p^2 & + & 2pq & + & q^2 = 1 \end{array}$$

Genotype	Phenotype	# of Individuals	Observed Genotype Frequencies	Expected Genotype Frequencies
R1R1	Yellow	35	0.088	0.175
R1R2	Green	265	0.662	0.487
R2R2	Blue	100	0.250	0.338

Answer (not iClicker)

Is this population in HWE? Yes or **no**.

.f of R1 (p) = 0.419
.f of R2 (q) = 0.581

- Observed more heterozygotes (green) than expected, and fewer homozygotes (yellow, blue) than expected

$$\overset{\text{R1R1}}{p^2} + \overset{\text{R1R2}}{2pq} + \overset{\text{R2R2}}{q^2} = 1$$

Genotype	Phenotype	# of Individuals	Observed Genotype Frequencies	Expected Genotype Frequencies
R1R1	Yellow	35	0.088	0.175
R1R2	Green	265	0.662	0.487
R2R2	Blue	100	0.250	0.338

iClicker Question

Which assumption has likely been violated?

- A. No natural selection (or mating is random)
- B. No natural selection
- C. No genetic drift
- D. No gene flow
- E. No mutations

In a grassy field in Pacific Spirit Park near UBC, there is a population of isolated crickets. Crickets are small invertebrates, which are eaten by a wide variety of visual predators, including amphibians, reptiles, birds and mammals. The crickets in this grassy field have three different body colours, which are governed by a single pair of alleles with a non-dominant relationship:

Crickets that are homozygous for the R1 allele are yellow.

Crickets that are homozygous for the R2 allele are blue.

Heterozygotes are green.

Question – 4 minutes

Explain why this isolated population of crickets may not be in HWE due to natural selection.

In a grassy field in Pacific Spirit Park near UBC, there is a population of isolated crickets. Crickets are small invertebrates, which are eaten by a wide variety of visual predators, including amphibians, reptiles, birds and mammals. The crickets in this grassy field have three different body colours, which are governed by a single pair of alleles with a non-dominant relationship:

Crickets that are homozygous for the R1 allele are yellow.

Crickets that are homozygous for the R2 allele are blue.

Heterozygotes are green.

Genotype	Phenotype	# of Individuals	Observed Genotype Frequencies	Expected Genotype Frequencies
R1R1	Yellow	35	0.088	0.175
R1R2	Green	265	0.662	0.487
R2R2	Blue	100	0.250	0.338

An example answer

Natural selection occurs when there is heritable variation in a trait, and individuals with a certain genotype and phenotype are more likely than individuals with other genotypes/phenotypes to survive and reproduce, and thus pass their alleles to the next generation in greater frequency.

In the cricket population there is heritable variation in body colour, which occurs in green, yellow and blue phenotypes, and is coded for by the R gene. Although no specific data was provided on differences in the reproductive success among the green, blue and yellow crickets, the green crickets would likely have a survival advantage over the blue and yellow crickets because they would be less visible to visual predators on a grassy field. Higher survivorship means a potentially longer life, and more opportunities for the green crickets to mate and pass their alleles to the next generation compared to the less well camouflaged yellow and crickets.

This would explain why the green crickets have a higher observed frequency than expected (0.663 compared to 0.487 or 265 observed green crickets compared 195 green crickets). And, why the yellow and blue crickets have lower observed frequencies than expected (yellow = 0.088 vs. 0.175 and blue = 0.250 vs 0.338).

Therefore, this cricket population may not be in HWE because natural selection may be favouring the green crickets (R_1R_2) over the yellow (R_1R_1) and blue (R_2R_2) crickets, violating one of the assumptions required for a population to be in HWE, and the population is evolving with respect to the R gene.

Hardy-Weinberg Principle – Learning Objectives

Understand/know:

- The 5 assumptions for the Hardy-Weinberg Equilibrium (HWE) Model
- How to calculate:
 - Observed genotype frequencies in a population.
 - Allele frequencies in a population from observed genotype frequencies
 - Expected (predicted) genotype frequencies, under HWE conditions (i.e. no evolution), using allele frequencies.
- How to use this information to test whether a population is in HWE for a particular gene.
- If not in HWE, be able to provide a reason why a population might not be at HWE equilibrium for a particular gene (what assumption has been violated). Be able to explain your answer using evidence and reasoning.

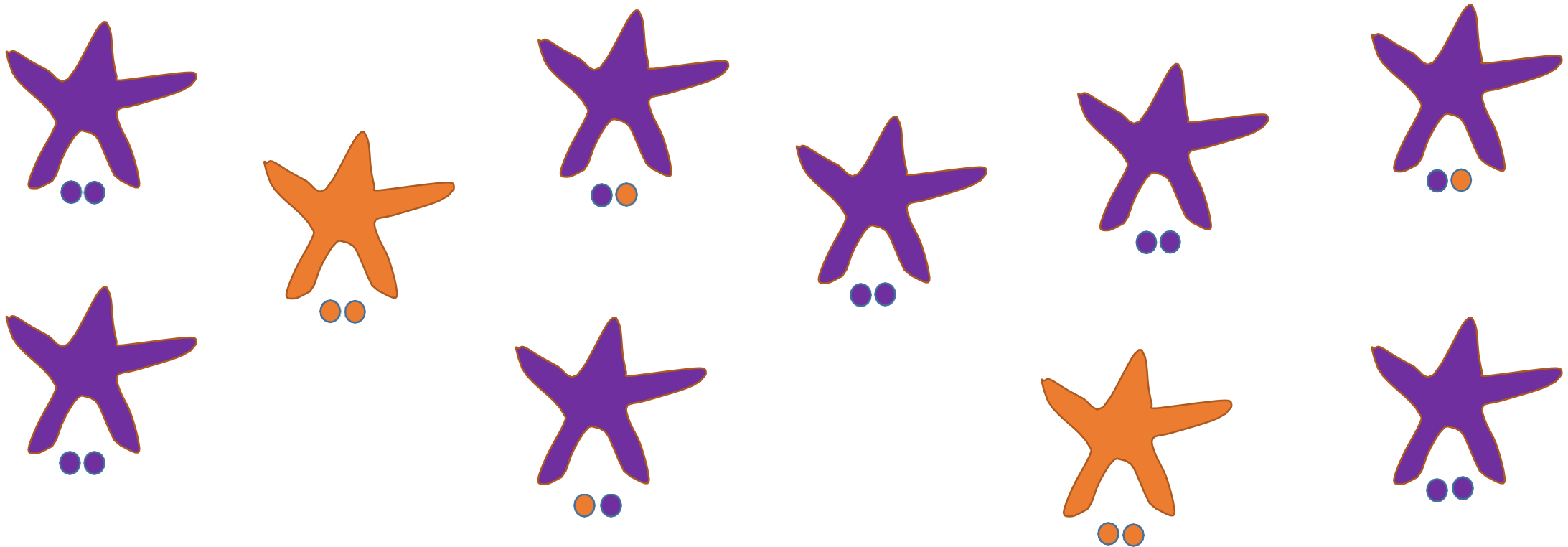
Next two questions (sea stars, fish) for practice
calculating values

Is this sea star population in HWE?

Consider a simple population of 10 sea stars with both purple and orange phenotypes. 1 gene for colour, with two alleles: C = purple and c = orange

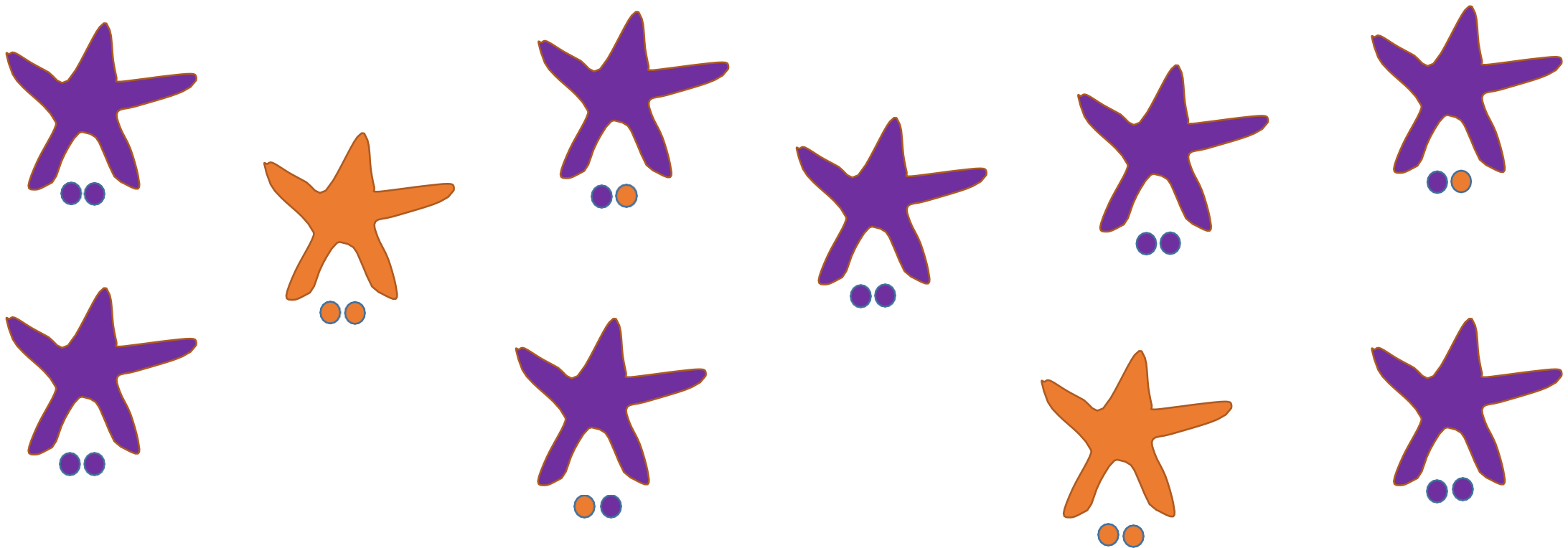
CC = purple (●●)

Cc = purple (●● or ●●) cc = orange (●●)



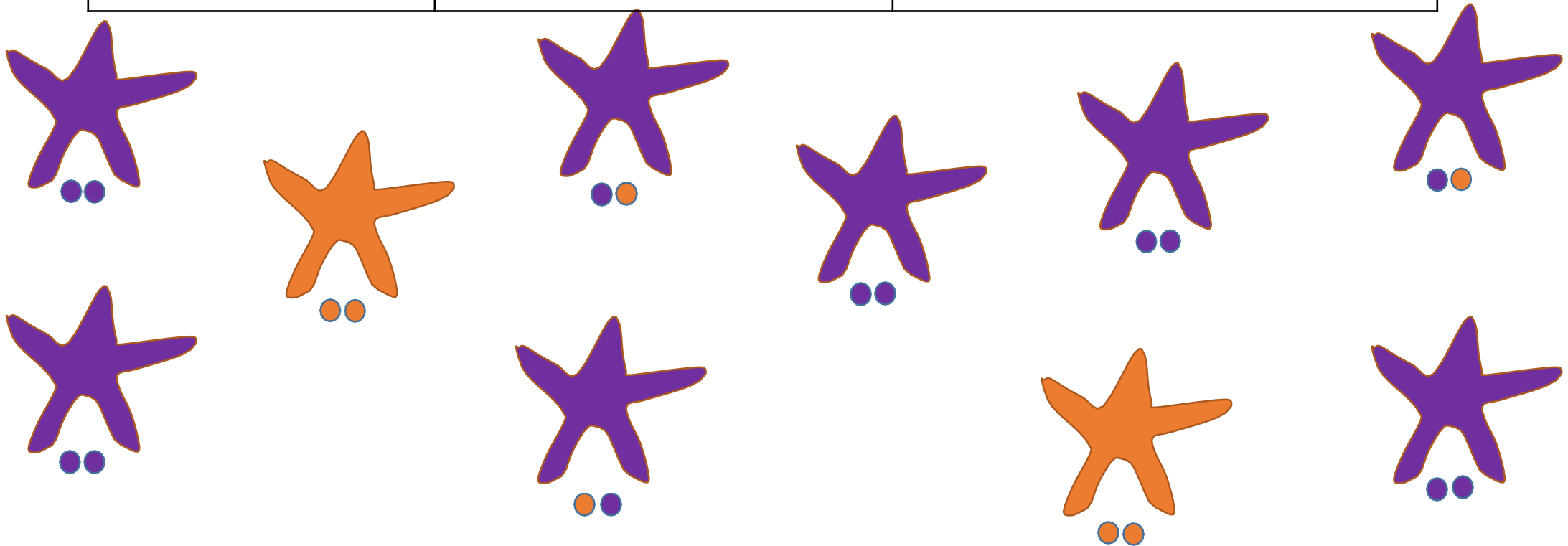
To determine if this population is in HWE...

We compare the **expected/predicted** genotype frequencies (if the population is in HWE) with the **real (or observed)** genotype frequencies



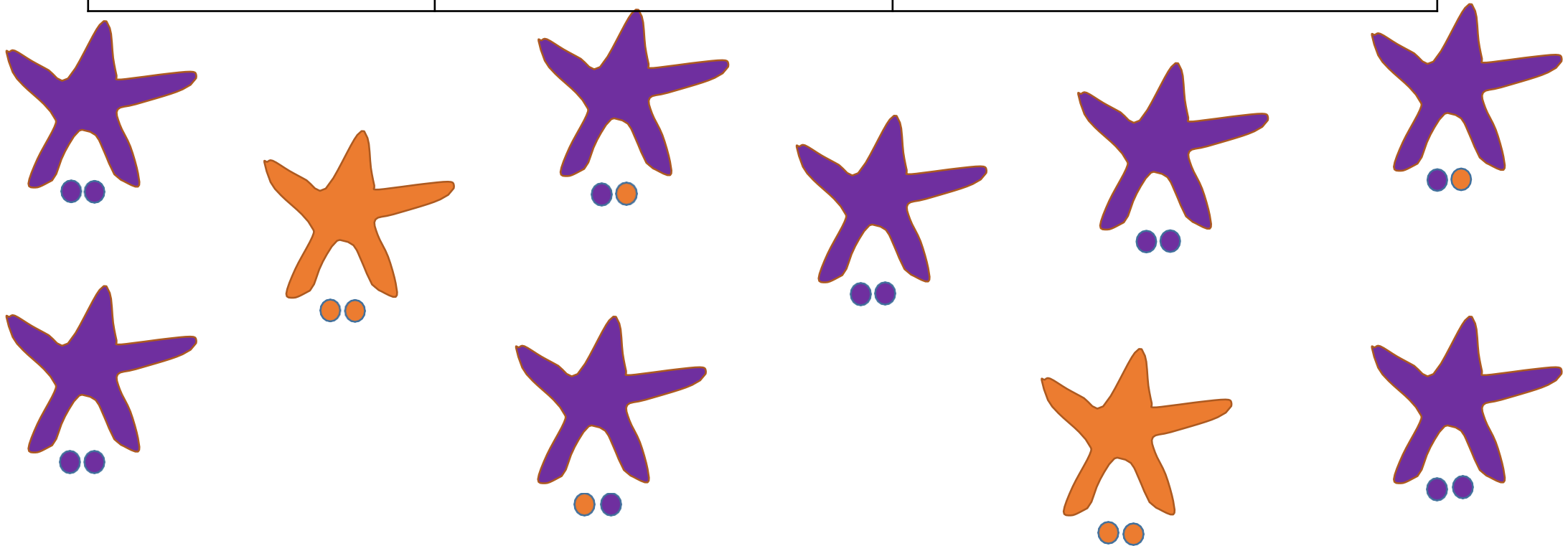
Step 1: Calculate observed genotype frequencies

Genotype	Number	Observed Genotype Frequency
CC	5	
Cc	3	
cc	2	
Total	10	



Step 1: Calculate observed genotype frequencies

Genotype	Number	Observed Genotype Frequency
CC	5	$5/10 = 0.50$
Cc	3	$3/10 = 0.30$
cc	2	$2/10 = 0.20$
Total	10	

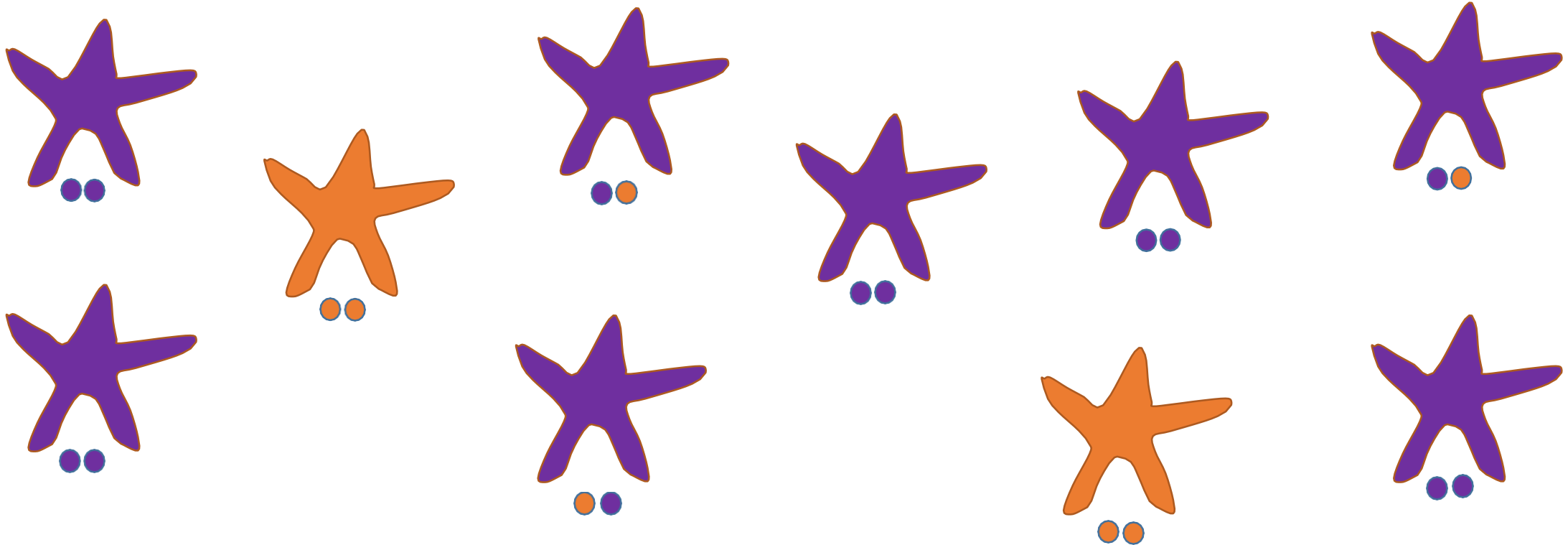


Step 2: Calculate observed allele frequencies – option #1

genotype frequency
$CC = 5/10 = 0.50$
$Cc = 3/10 = 0.30$
$cc = 2/10 = 0.20$

allele	frequency
C (p)	
c (q)	

$$p + q = 1$$

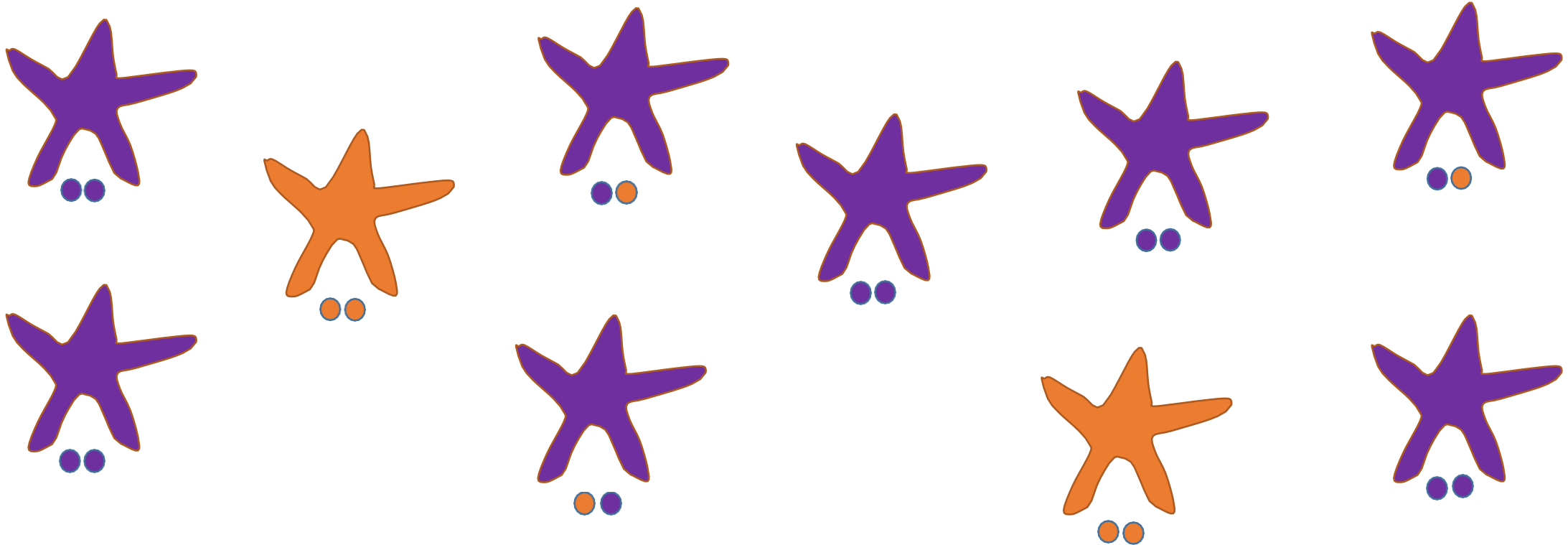


Step 2: Calculate observed allele frequencies – option #1

genotype frequency
$CC = 5/10 = 0.50$
$Cc = 3/10 = 0.30$
$cc = 2/10 = 0.20$

allele	frequency
C (p)	$[5*2 (AA) + 3*1 (Aa)] / 20 = 0.65$
c (q)	$[2*2 (aa) + 3*1 (Aa)] / 20 = 0.35$

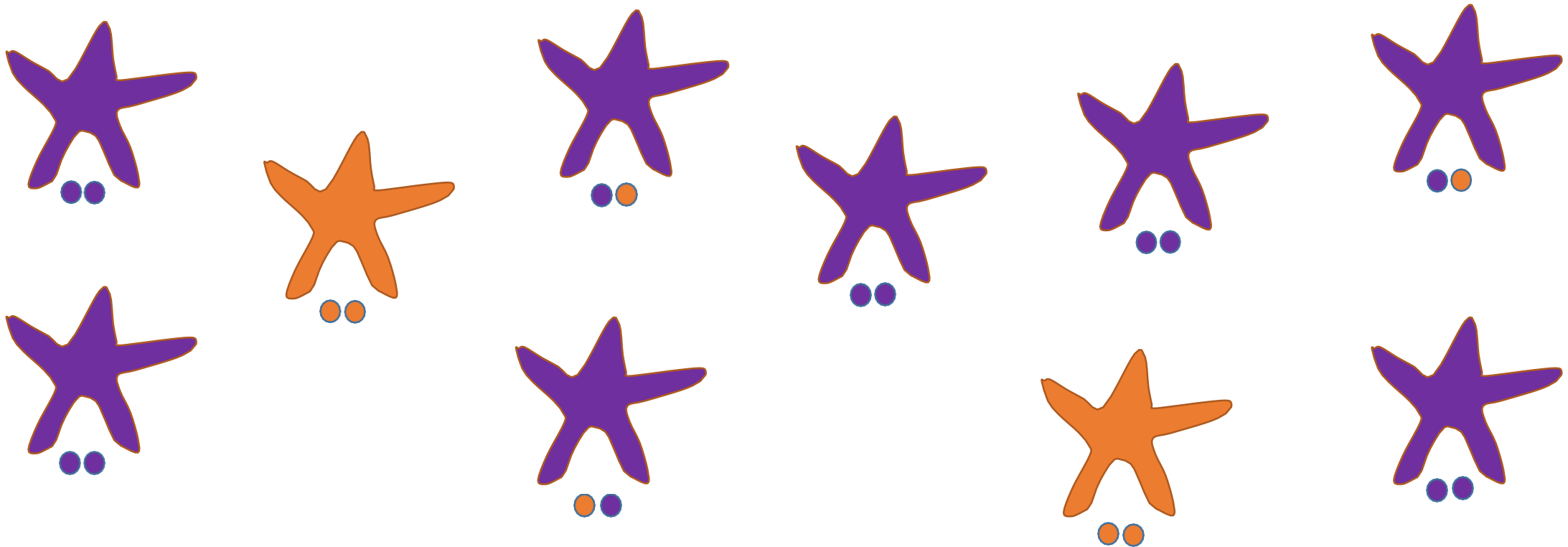
$$p + q = 1$$



Step 2: Calculate observed allele frequencies – option #2

genotype frequency
$CC = 5/10 = 0.50$
$Cc = 3/10 = 0.30$
$cc = 2/10 = 0.20$

allele	frequency
C (p)	$0.50 + (0.30/2) = 0.65$
c (q)	$0.20 + (0.30/2) = 0.35$



Step 3: Use the observed allele frequencies to calculate the expected genotype frequencies if the C gene is in HWE

Allele frequency	Genotype	Predicted genotype if in HWE
C (p) = 0.65	CC	p^2 =
c (q) = 0.35	Cc	$2pq$ =
	cc	q^2 =

Note: check the arithmetic..... $p^2 + 2pq + q^2 = 1$ (if not – arithmetical error)

Step 3: Use the observed allele frequencies to calculate the expected genotype frequencies if the C gene is in HWE

Allele frequency	Genotype	Predicted genotype if in HWE
C (p) = 0.65	CC	$p^2 = 0.65 \times 0.65 = 0.42$
c (q) = 0.35	Cc	$2pq = 2(0.65 \times 0.35) = 0.46$
	cc	$q^2 = 0.35 \times 0.35 = 0.12$

Note: check the arithmetic..... $0.42 + 0.46 + 0.12 = 1$ (if not – arithmetical error)

Step 4: Compare expected/predicted HWE genotype frequencies with real/observed genotype frequencies.

Genotype	Predicted genotype freq. if in HWE	Actual/observed genotype frequencies
CC	0.42	0.50
Cc	0.46	0.30
cc	0.12	0.20

iClicker Question – Is this sea start population in HWE for the C gene?

- A. Yes
- B. No
- C. I am not sure

Step 4: Compare expected/predicted HWE genotype frequencies with real/observed genotype frequencies.

Genotype	Predicted genotype freq. if in HWE	Actual/observed genotype frequencies
CC	0.42	0.50
Cc	0.46	0.30
cc	0.12	0.20

iClicker Question – Is this population in HWE for the C gene?

A. Yes

B. No

C. I am not sure

Observed more homozygous individuals than expected

Observed fewer heterozygotes than expected

Why might this gene (C gene) not be in Hardy-Weinberg equilibrium?

Genotype	Predicted genotype freq. if in HWE	Actual/observed genotype frequencies
CC	0.42	0.50
Cc	0.46	0.30
cc	0.12	0.20

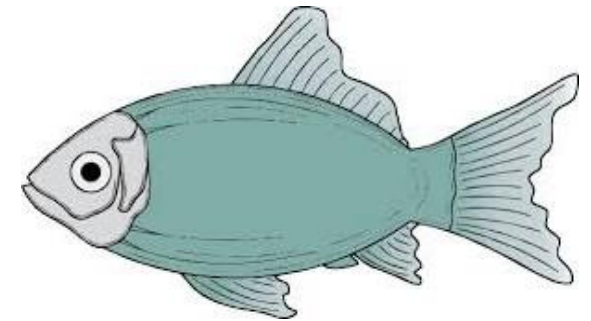
Fewer heterozygotes than expected. What assumption is violated?

iClicker question: Is this fish population in Hardy Weinberg Equilibrium for the A gene?

- A. Yes
- B. No
- C. I didn't do the calculations 😊

You've genotyped a population of 122 fish

- 28 are homozygous for the $A1$ allele
- 15 are homozygous for the $A2$ allele
- 79 are heterozygotes



$$\bullet\bullet = p^2$$

$$\bullet\circ = 2pq$$

$$\circ\circ = q^2$$

Is this fish population in Hardy-Weinberg Equilibrium for the A gene?

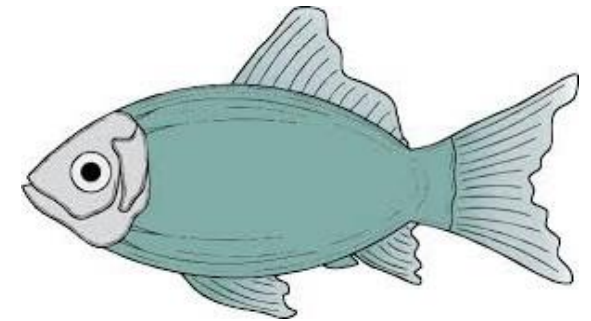
A. Yes

B. No

C. I didn't do the calculations 😊

You've genotyped a population of 122 fish

- 28 are homozygous for the $A1$ allele
- 15 are homozygous for the $A2$ allele
- 79 are heterozygotes



$$\bullet\bullet = p^2$$

$$\bullet\circ = 2pq$$

$$\circ\circ = q^2$$

Step 1: Calculate observed genotype frequencies

Genotype	Number	Observed Genotype Frequency
A1A1	28	
A1A2	79	
A2A2	15	
Total	122	

Check arithmetic:

Step 1: Calculate observed genotype frequencies

Genotype	Number	Observed Genotype Frequency
A1A1	28	$28/122 = 0.23$
A1A2	79	$79/122 = 0.65$
A2A2	15	$15/122 = 0.12$
Total	122	

Check arithmetic: $0.23 + 0.65 + 0.12 = 1$

Step 2: Calculate observed allele frequencies

One approach

of each genotype
A1A1 = 28
A1A2 = 79
A2A2 = 15
Total = 122

	allele frequency
A1 (p)	
A2 (q)	

Check arithmetic:

Step 2: Calculate observed allele frequencies

of each genotype
A1A1 = 28
A1A2 = 79
A2A2 = 15
Total = 122

genotype frequency
A1A1 = 0.23
A1A2 = 0.65
A2A2 = 0.12

	allele frequencies
A1 (p)	$[(28 * 2) + 79] / 244 = 0.55$
A2 (q)	$[(15 * 2) + 79] / 244 = 0.45$

Check arithmetic:
 $0.553 + 0.445$

Step 2: Calculate observed allele frequencies

Second approach

genotype frequency
A1A1 = 0.23
A1A2 = 0.65
A2A2 = 0.12

allele	frequency
A1 (p)	
A2 (q)	

Check arithmetic:

Step 2: Calculate observed allele frequencies

genotype frequency
A1A1 = 0.23
A1A2 = 0.65
A2A2 = 0.12

allele	frequency
A1 (p)	$0.23 + (0.65/2) = 0.55$
A2 (q)	$0.12 + (0.65/2) = 0.45$

Check arithmetic: $0.55 + 0.45 = 1$

Step 3: Use allele frequencies to calculate expected genotype frequencies

Allele frequency	Genotype	Predicted Genotype Frequency if in HWE
A1 (p) = 0.55	A1A1	
A2 (q) = 0.45	A1A2	
	A2A2	

$$p^2 + 2pq + q^2 = 1$$

Check addition:

Step 3: Use allele frequencies to calculate expected genotype frequencies

Allele frequency	Genotype	Predicted Genotype Frequencies if in HWE
A1 (p) = 0.55	A1A1	$p^2 = 0.55^2 = 0.30$
A2 (q) = 0.45	A1A2	$2pq = 2(.55)(.45) = 0.50$
	A2A2	$q^2 = 0.45^2 = 0.20$

Check addition: $0.30 + 0.50 + 0.20 = 1$

Fish Question - Step 4: Compare predicted/expected genotype frequencies (under HWE) and observed genotype frequencies.

Genotype	Predicted genotype frequencies	Observed genotype frequencies
A1A1	0.30	.23
A1A2	0.50	.65
A2A2	0.20	.12

More heterozygotes observed than expected, and a fewer homozygous individuals observed than expected.

A biologist would likely do a chi-square test to see if there is a significant difference between observed & expected. For 121, you can use greater than a 0.10 difference.

Questions?

CWPA FINALIST 2023



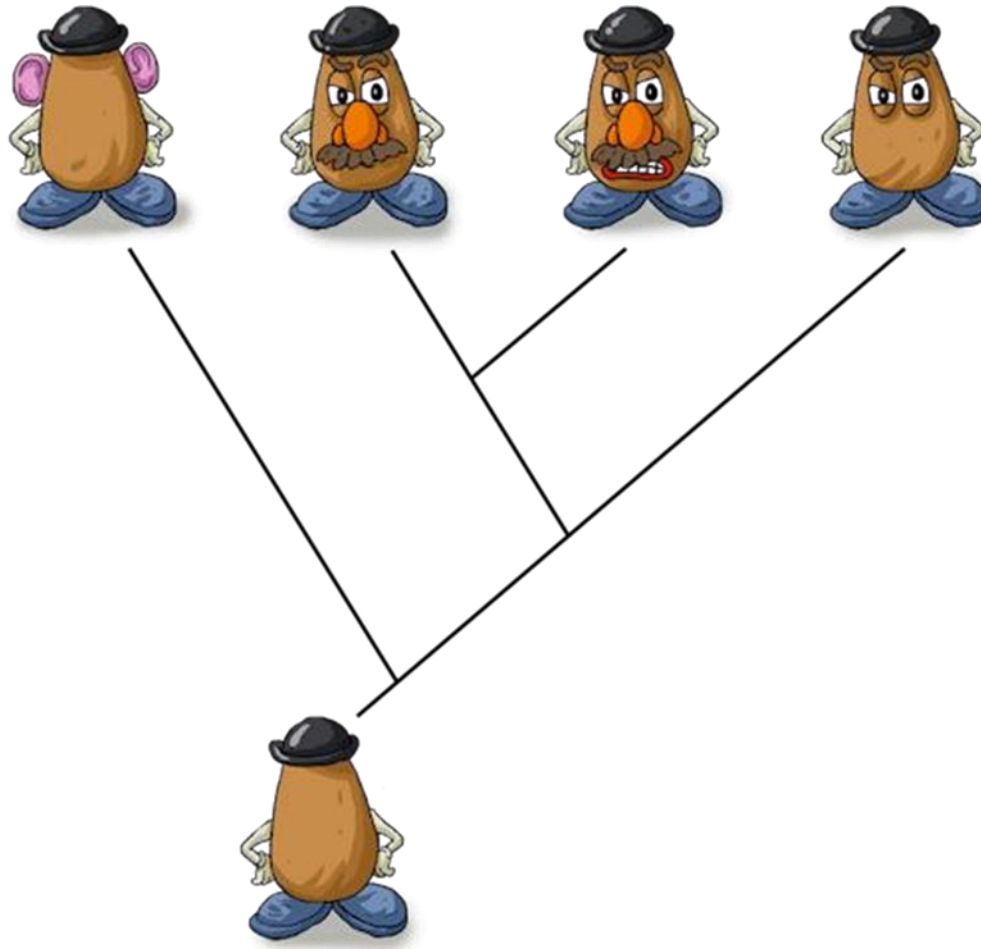
4-minute break. Find the bear in this photo.



iClicker Question

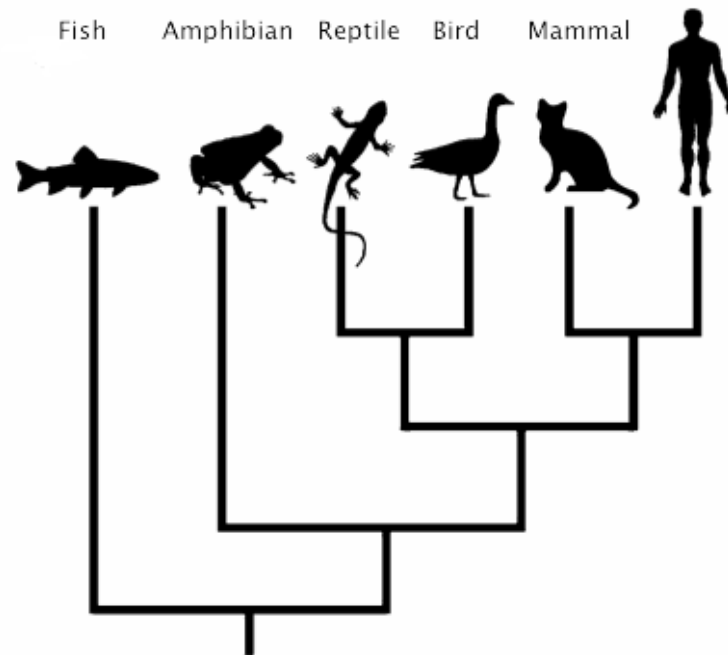
Have you had classes on phylogenetic trees previously?

- A. Yes – within the last year
- B. Yes – > than 1 years ago
- C. No but I am familiar with concept.
- D. Phylo what?



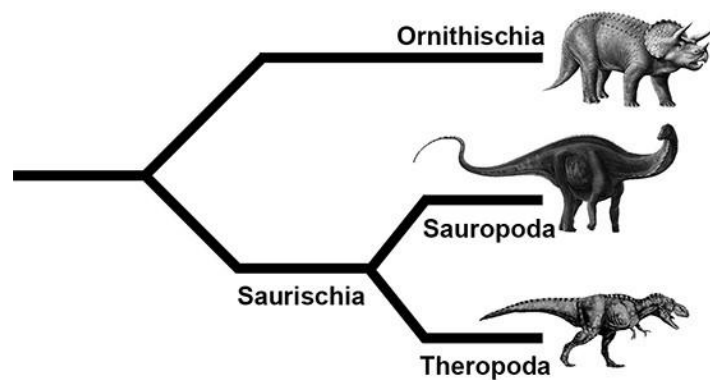
What is a phylogenetic tree?

- Phylogenetic tree – also called an evolutionary tree - is a diagram that depicts the hypothesized evolutionary relationships amongst species or other taxa (or DNA or proteins)

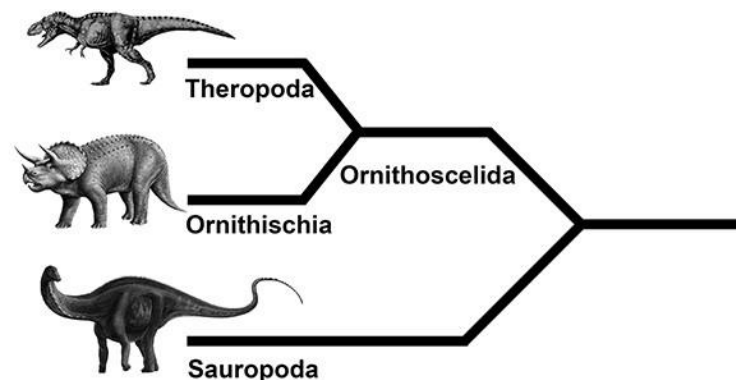


Why hypothesized?

- We cannot know evolutionary relationships with absolute certainty.
- Phylogenetic trees are always tested/revised as new data becomes available.

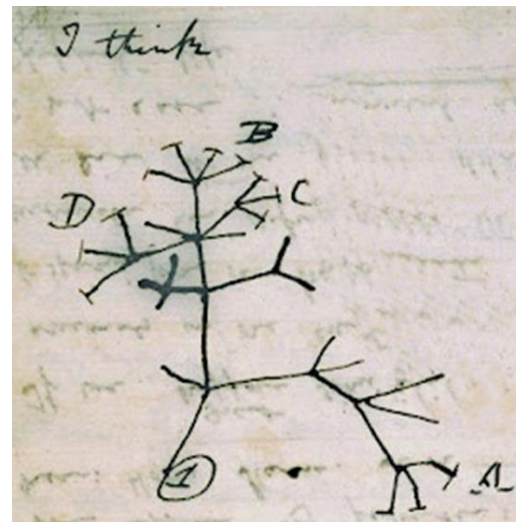
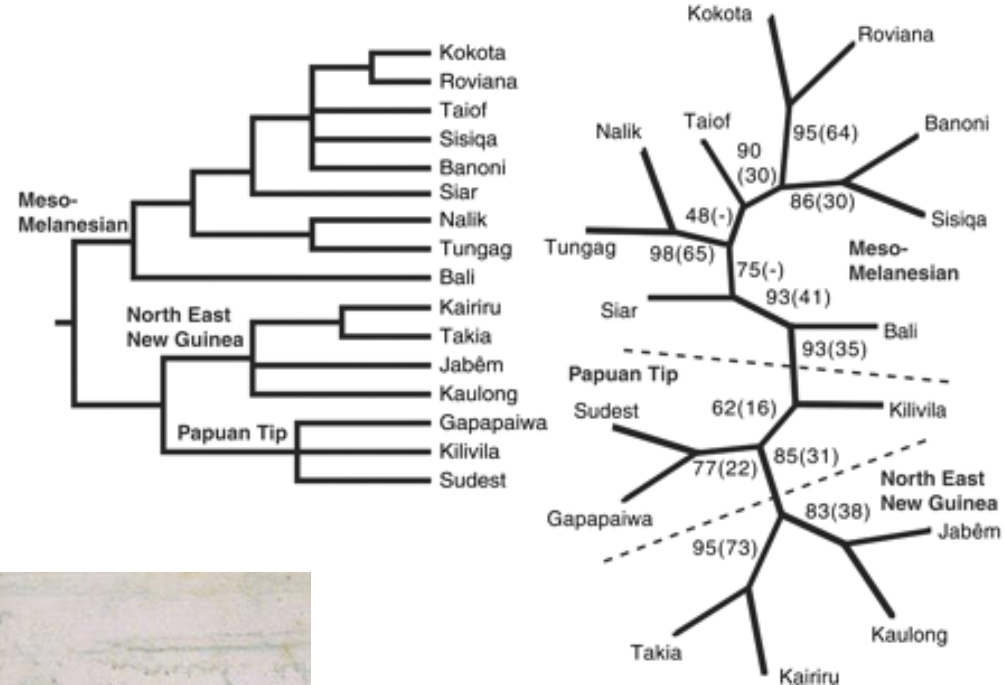
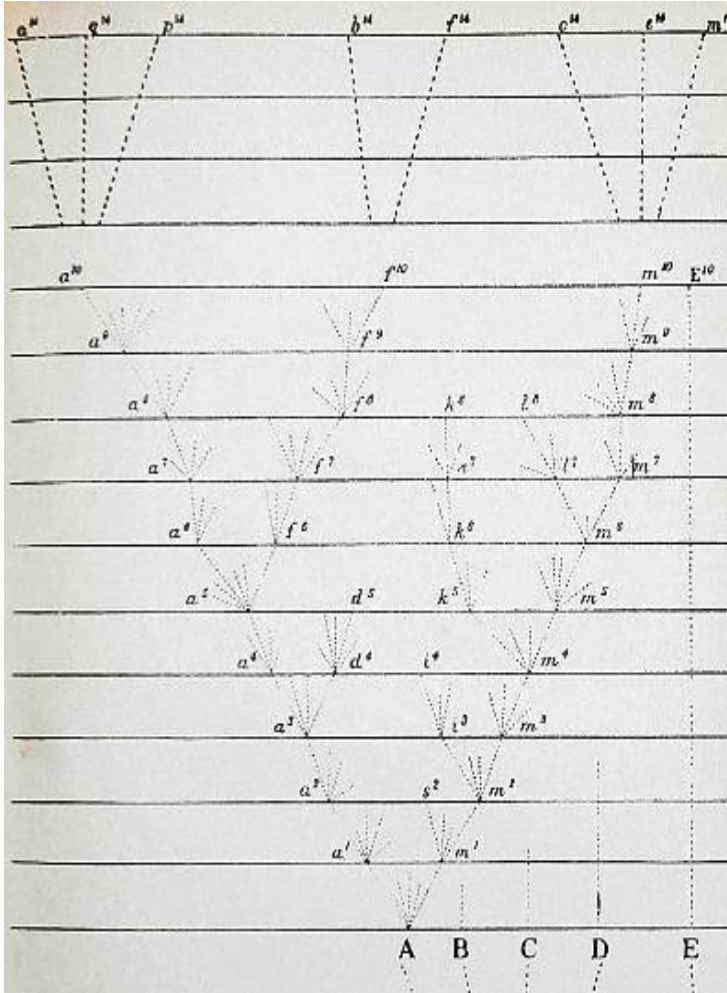


Classic interpretation



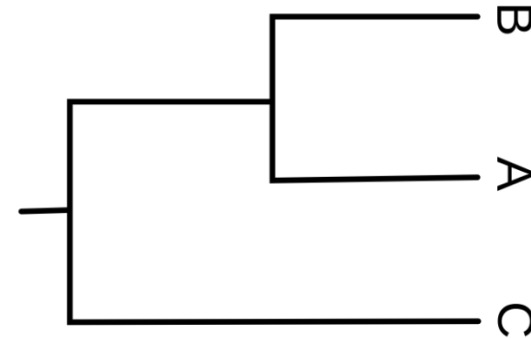
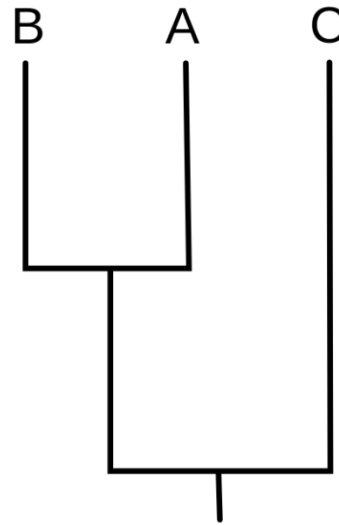
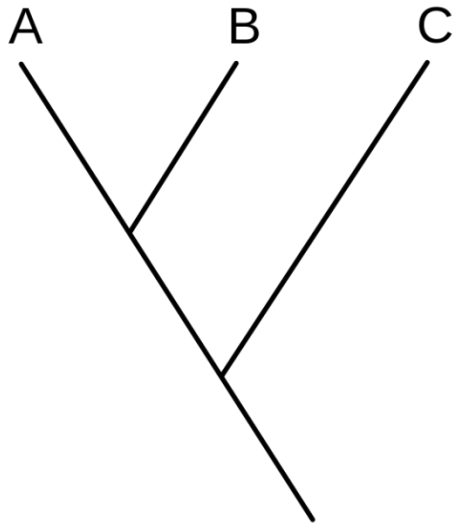
Baron et al. 2017

Phylogenetic trees – various forms



The first phylogenetic tree diagram, sketched by Charles Darwin.

Phylogenetic tree shapes used in 121



Phylogenetic Trees – Terminology

- Tips
- Branches/lineages
- Roots
- Nodes
- Character marks
- Homologous traits
 - Synapomorphies (Tuesday's class)
- Homoplasies or Analogous Traits
 - Convergent evolution

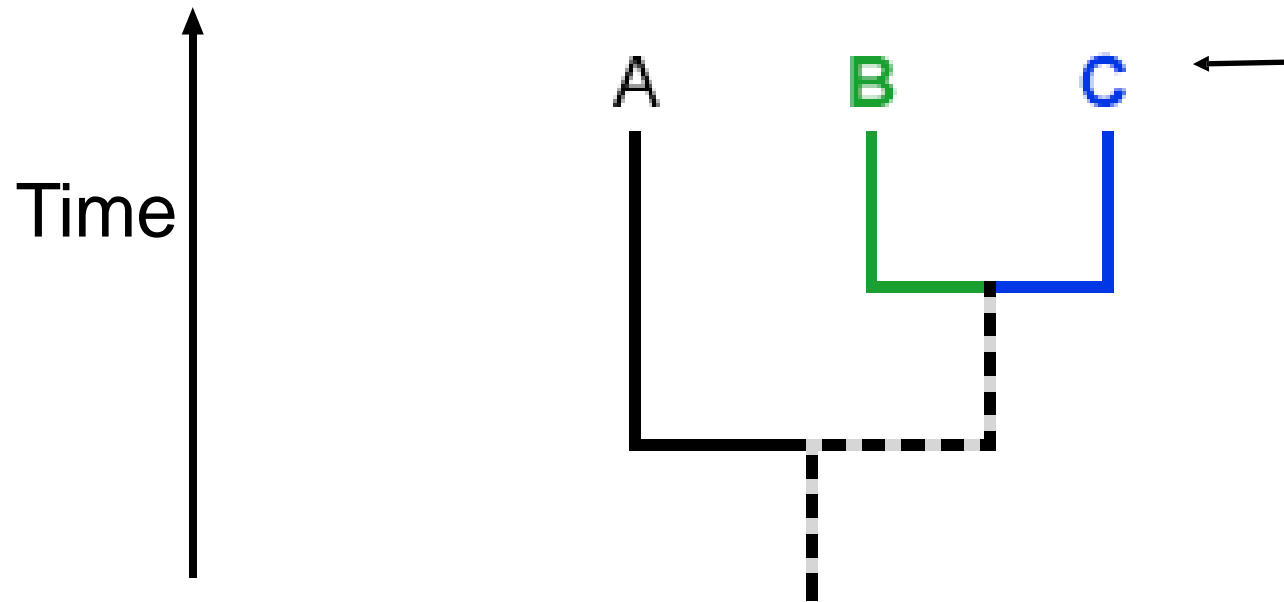
READING TREES:

- Monophyletic groups
 - Sister taxa
- Non-monophyletic groups
 - Paraphyletic Groups
 - Polyphyletic Groups
- Principle of parsimony

A helpful website:

<https://evolution.berkeley.edu/the-tree-room/evolutionary-trees-a-primer/the-anatomy-of-an-evolutionary-tree/>

Anatomy of a phylogenetic tree: Tips (terminal nodes)

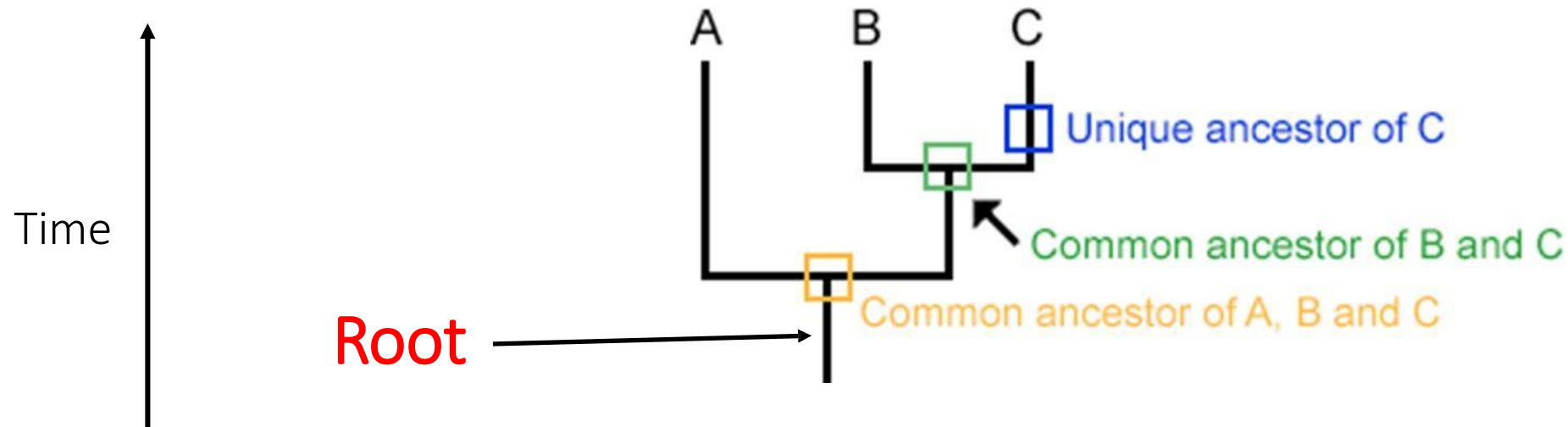


- typically represent species*, but not always (can also represent populations, larger groups – e.g. families, genes, DNA, etc.)
- tips = youngest lineage.

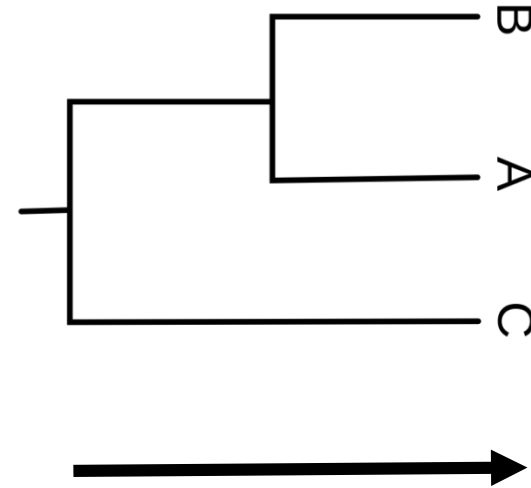
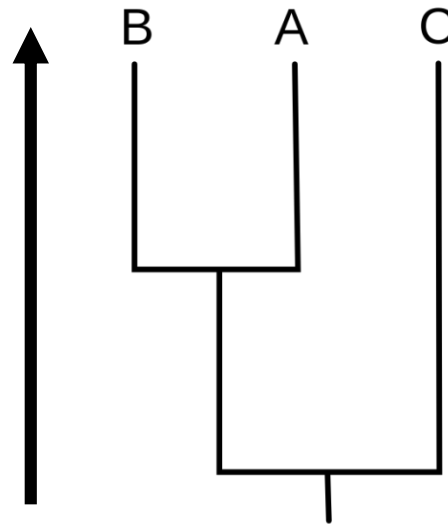
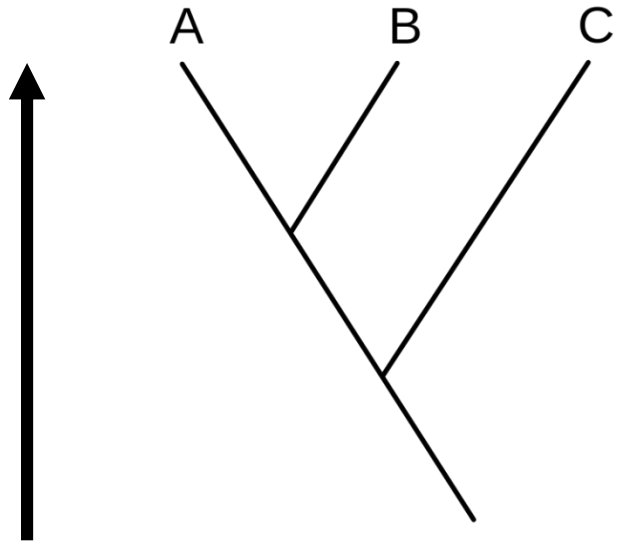
*For 121 – tips will represent extant species.

Anatomy of a phylogenetic tree: Root

- The root represents the oldest lineage on the tree. It is the lineage that is ancestral to all other organisms on the tree.



Time goes from root to tip



Arrow = direction of time

Anatomy of a phylogenetic tree: Branches

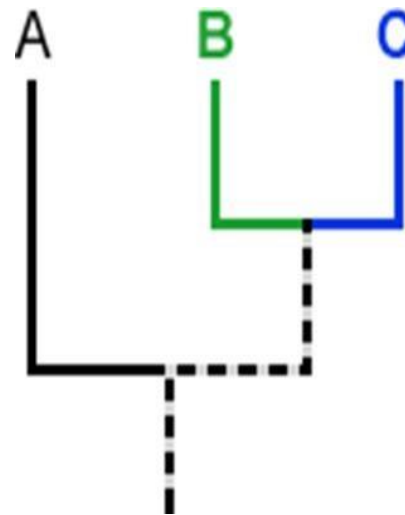
- Each branch represents a lineage: from ancestor to descendant.
- Evolution is occurring along all of the branches.

Note – in this tree horizontal lines are meaningless* – just a way to separate taxa.

In this tree - evolution occurring along vertical lines.

*Important to remember if asked to place character marks on the tree

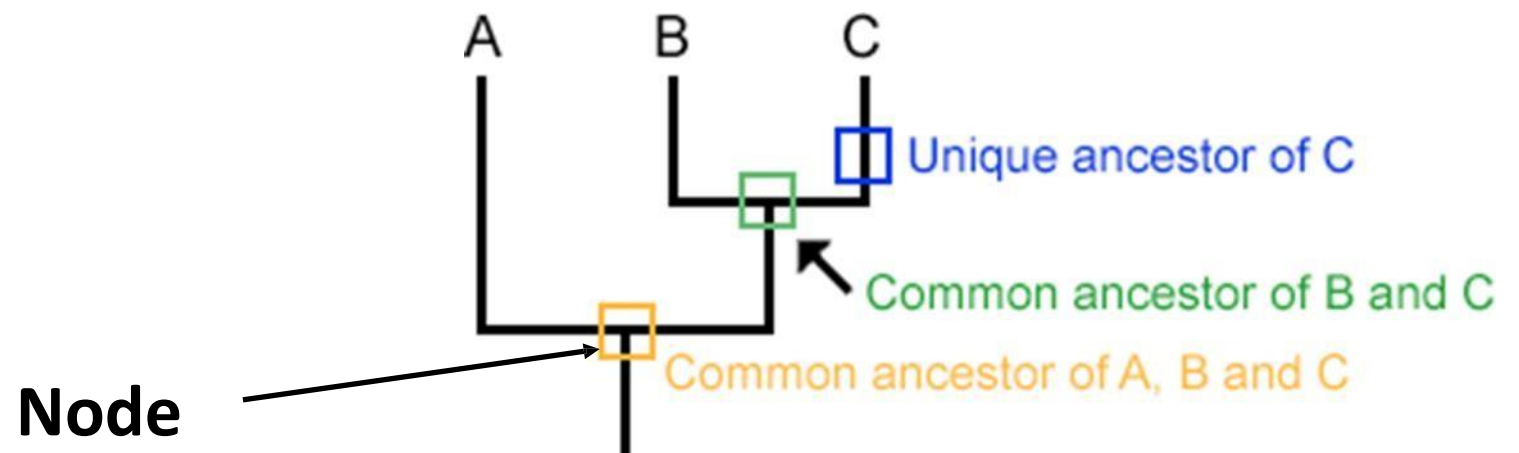
Branch length is meaningless for the trees that we will be looking at.



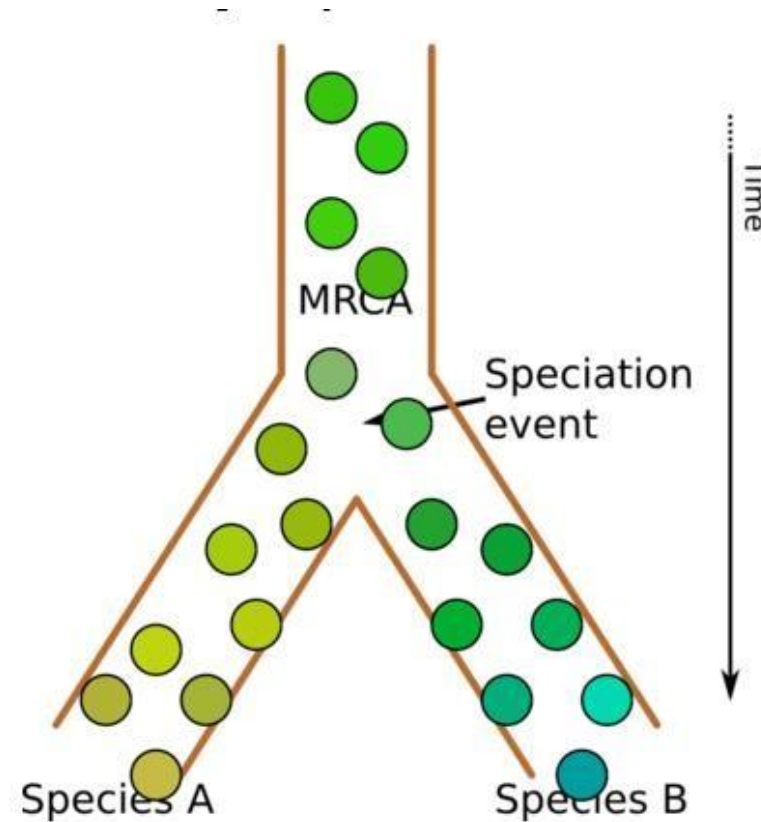
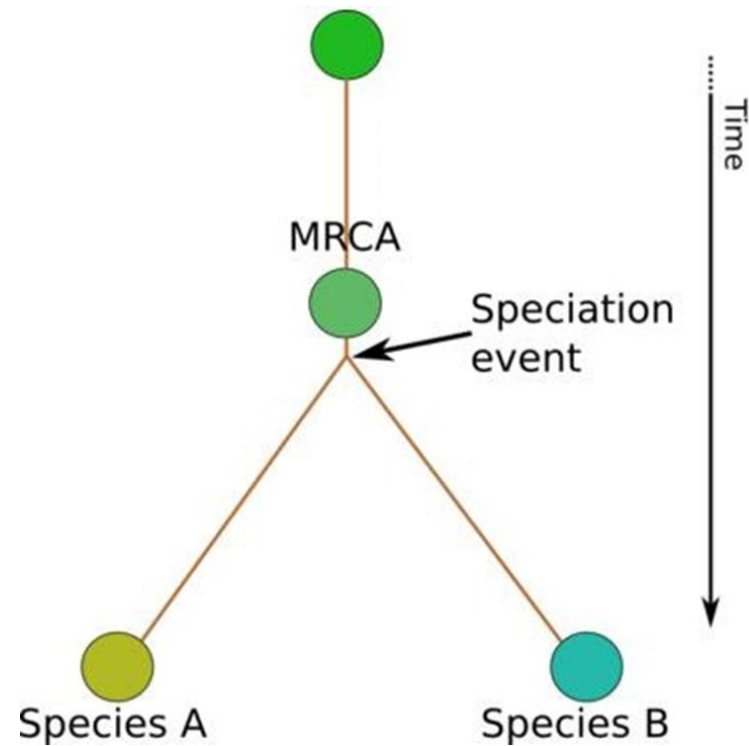
— Unique history of B
— Unique history of C
--- Shared history of B and C

Anatomy of a phylogenetic tree: internal nodes

- Branching points in a tree represent an inferred speciation event where the ancestral lineage split into 2 or more descendant lineages.
- The (internal) node represents the last common ancestor of the two (or more) lineages descended from that node.
- So, if you take any two taxa (e.g. B & C) and trace their lineage backwards (in time), the node at which their two paths meet represents the point in evolutionary time where their common ancestor split into two lineages.



Reminder: It is populations that evolve
(circles present individuals)



Anatomy – Character states/marks

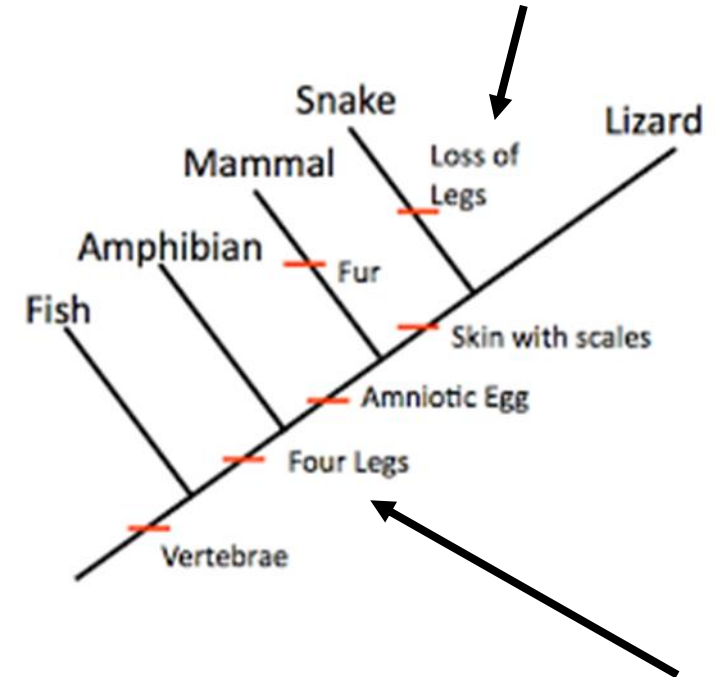
Character states (marks) are placed on trees to show when a new character state arose in a lineage (or when it was lost) – shown in red on this tree.

Character state = heritable trait.

Once a character state appears, assume that all descendants inherited that trait, unless the tree indicates that the trait was lost in a certain lineage.

For example:

- The presence of four legs arose in the common ancestor to tetrapods and was passed on to all descendants* (amphibians, mammals, crocodiles, birds).
- However, this trait was lost in snakes.



Reading a tree - Learning objective #1

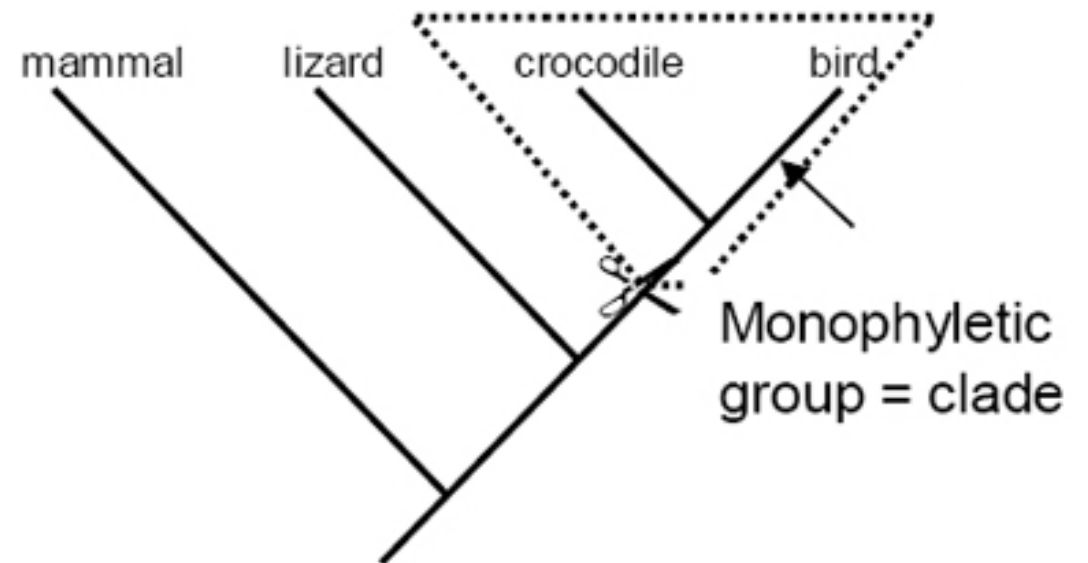
Be able to identify monophyletic, paraphyletic and polyphyletic groups in a tree.

Monophyletic groups (Clades)

- A **monophyletic group** or clade is a part of the phylogeny that includes:
(1) the ancestor (unique or common) + (2) all of its descendants of that ancestor.

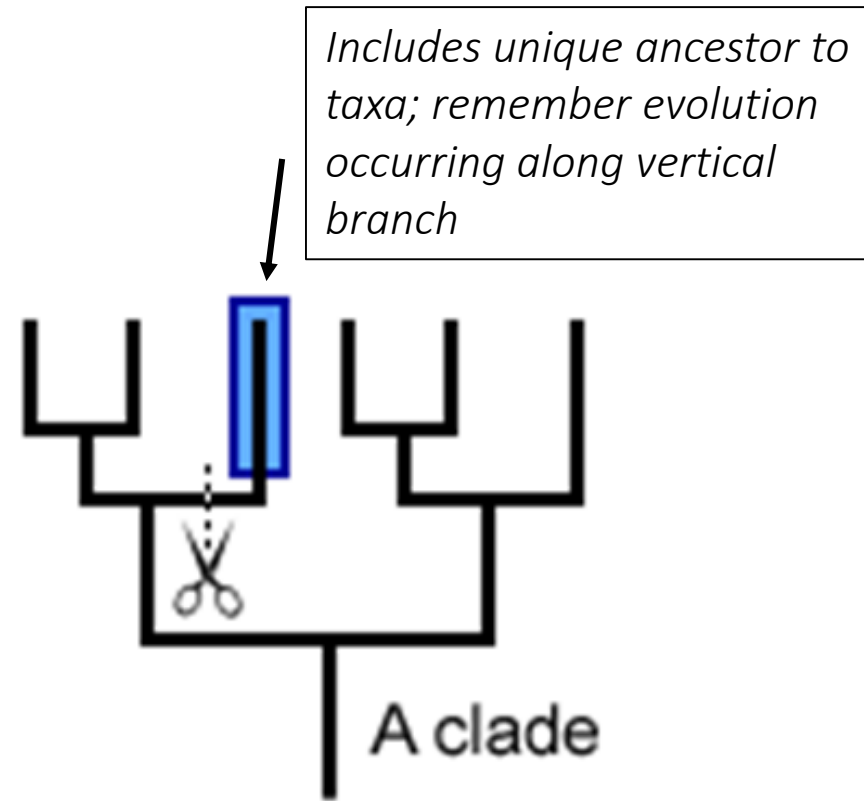
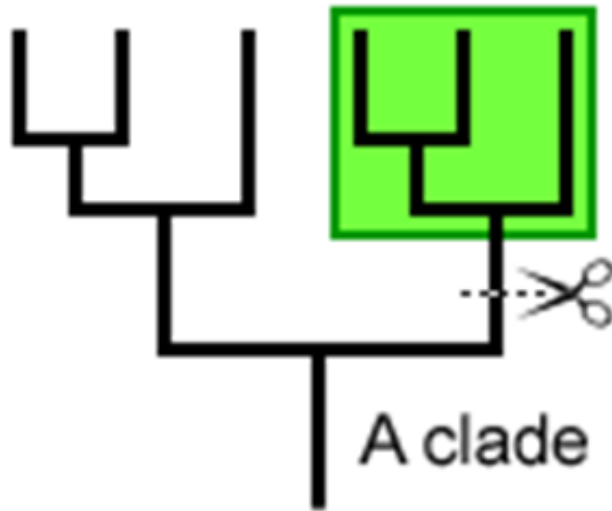
- In this example, crocodiles + birds + their common ancestor = monophyletic group
- Based on this tree, crocodiles and birds would be classified as sister taxa.

Sister taxa are two descendants that split from the same node. They are each other's closest relatives.



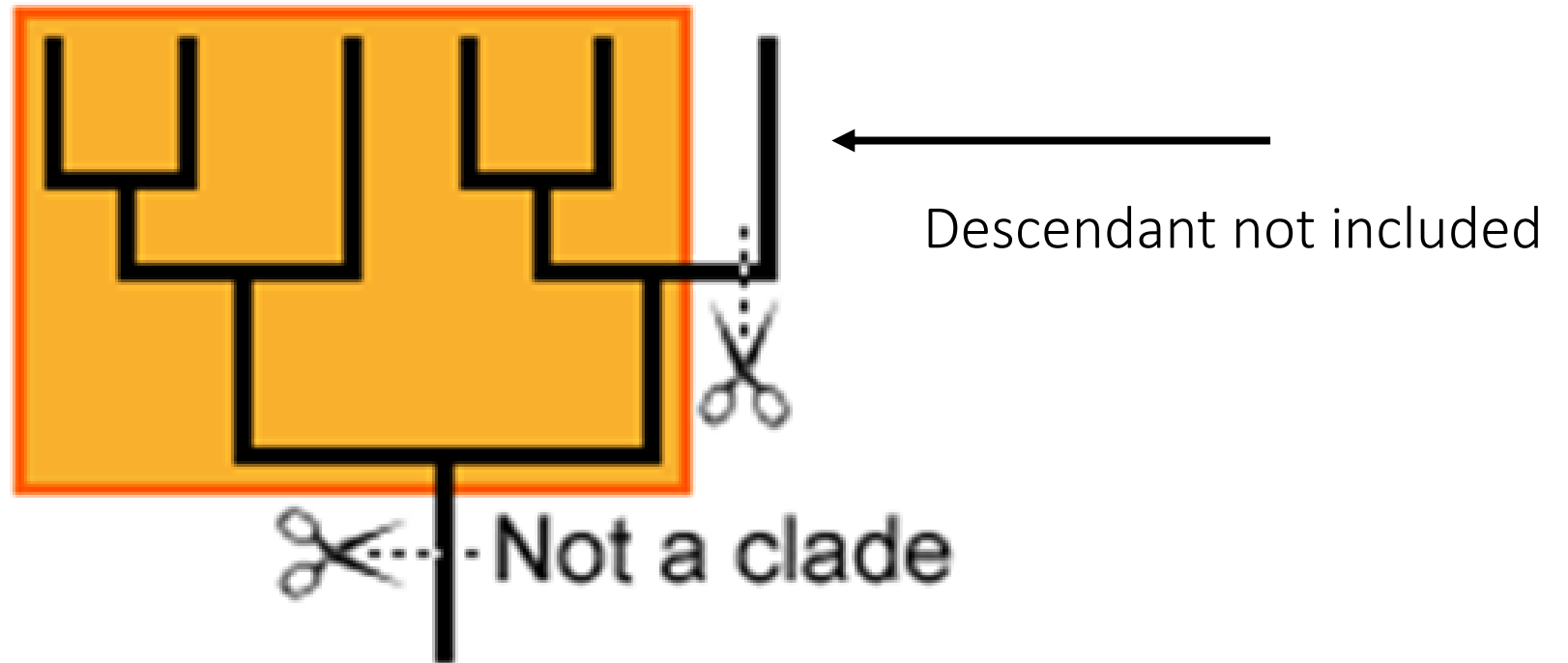
Monophyletic Groups/Clades – SNIP TEST

- If you cut a branch, anything that falls off is a clade or monophyletic group.



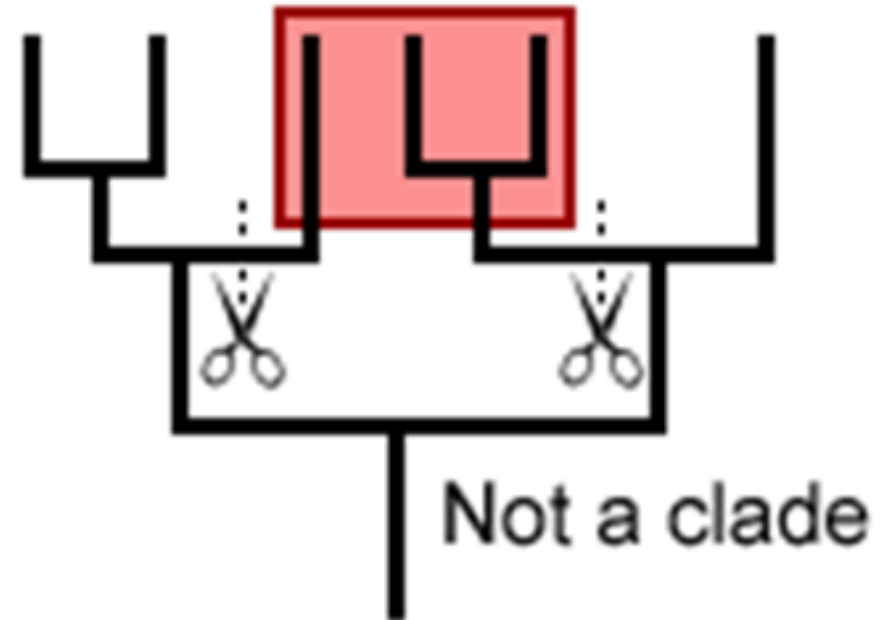
When is a group not monophyletic?

- When the group does not contain all of the descendants of the common ancestor.
- A group that contains the common ancestor, but not all descendants = a PARAPHYLETIC GROUP)



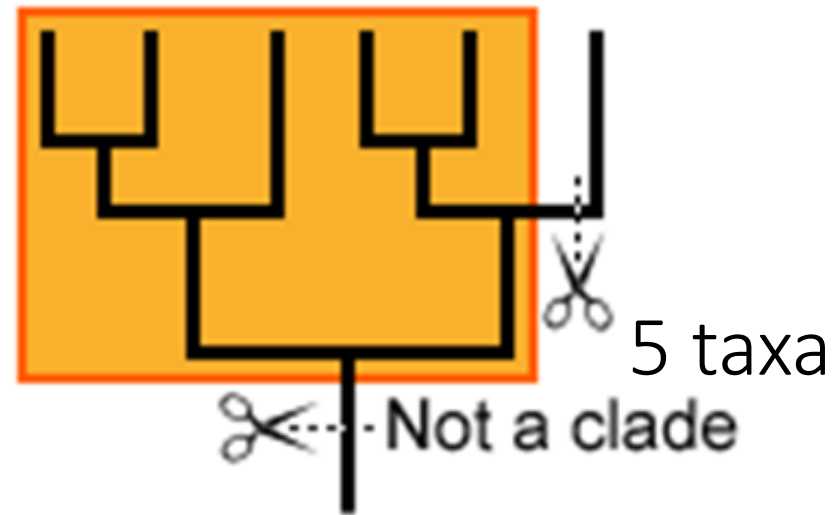
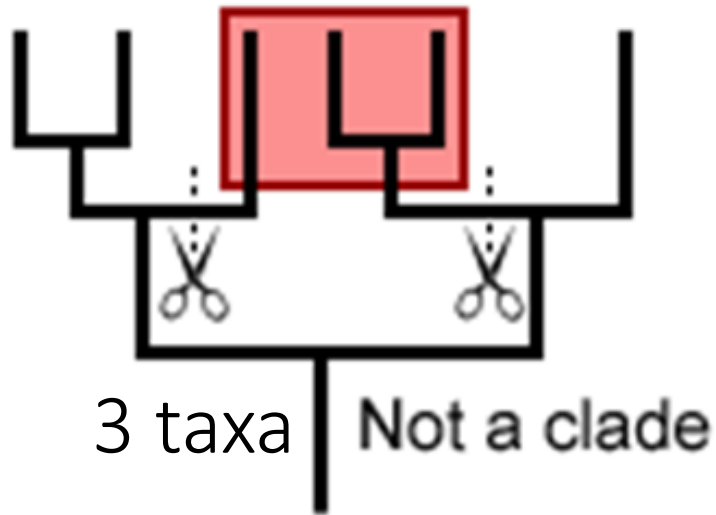
When is a group not monophyletic?

- When the group does NOT contain the most recent common ancestor to these taxa.
- A group that does NOT include the most recent common ancestor to that group is called a **POLYPHYLETIC GROUP**.
- Note - some descendants may also be missing.



Using the snip test

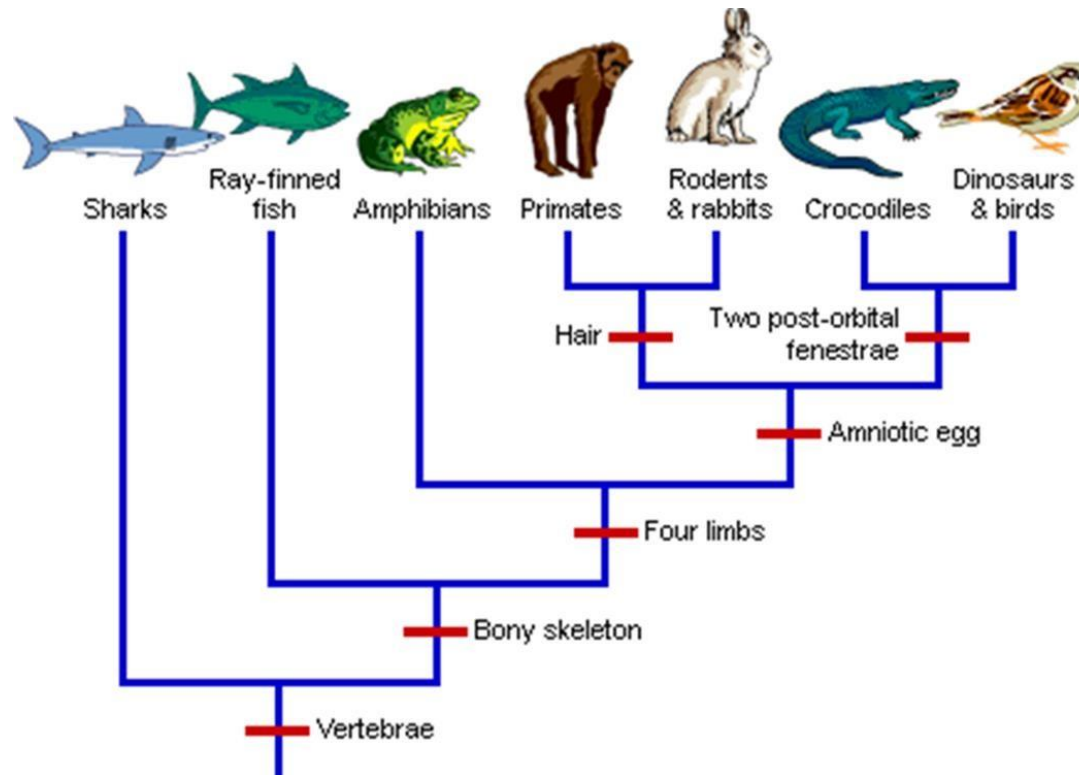
If a tree needs to be cut in two or more places to extract the taxa, then those taxa do not represent a monophyletic group or clade.



iClicker Question

- How many monophyletic groups are contained in this tree?

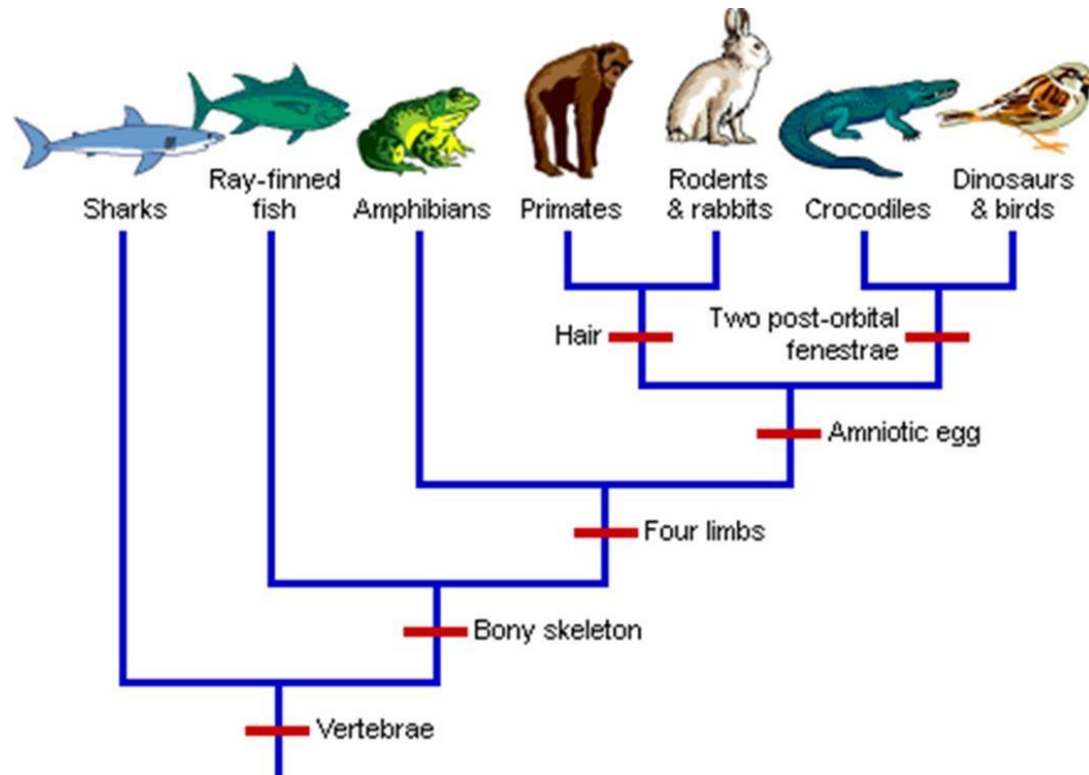
- A. 5
- B. 6
- C. 8
- D. 12
- E. 13



Answer

- How many monophyletic groups are contained in this tree?

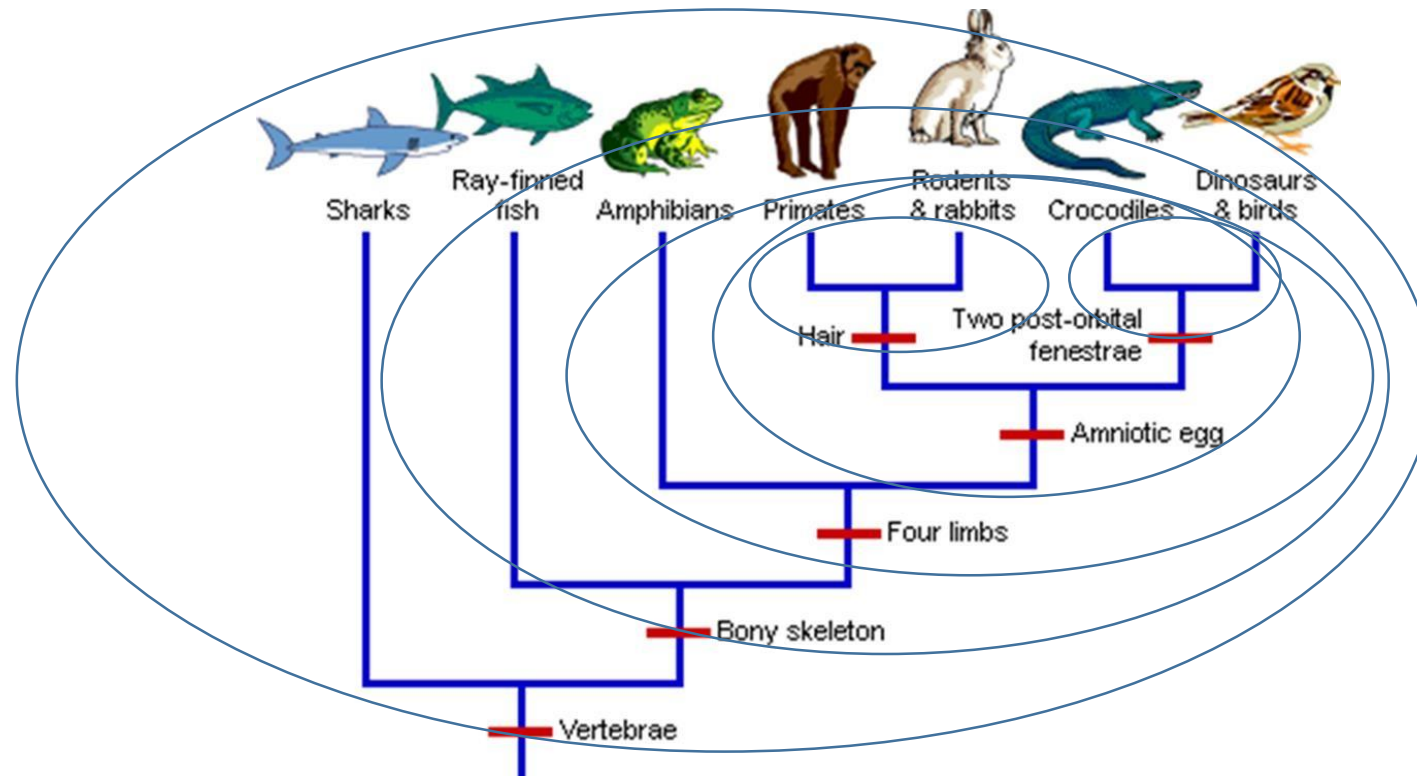
- A. 5
- B. 6
- C. 8
- D. 12
- E. 13



The monophyletic groups..

- How many monophyletic groups are contained in this tree?

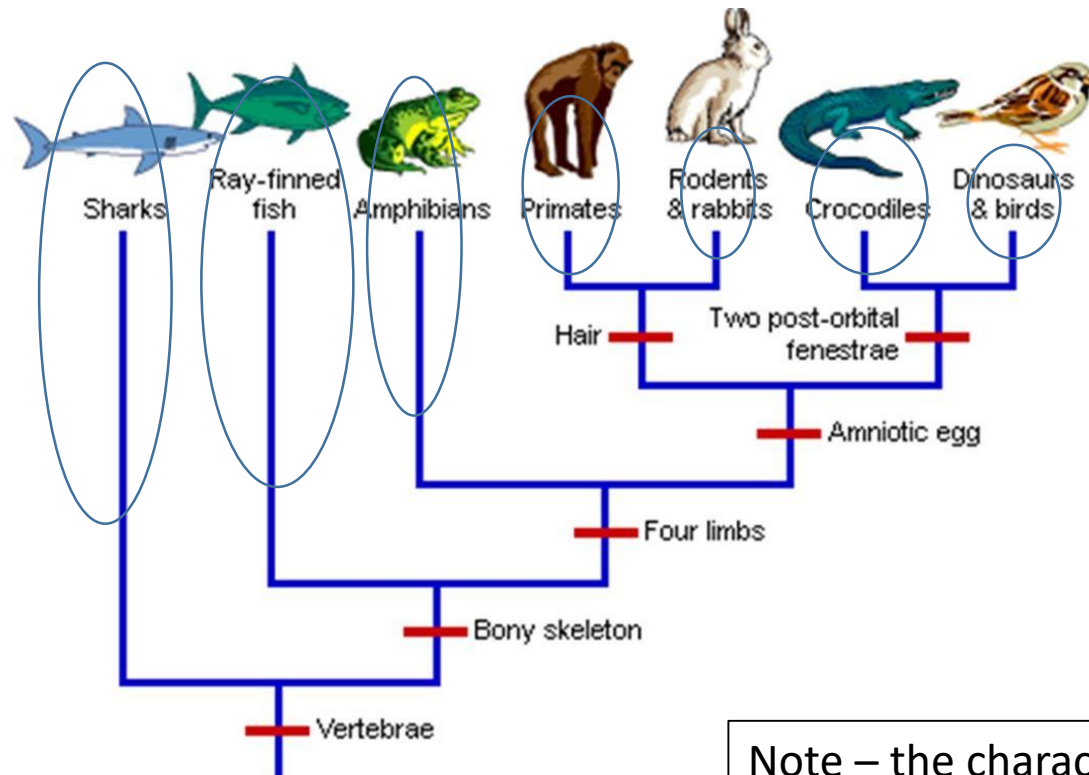
- A. 5
- B. 6
- C. 8
- D. 12
- E. 13



The monophyletic groups

- How many monophyletic groups are contained in this tree?

- A. 5
- B. 6
- C. 8
- D. 12
- E. 13



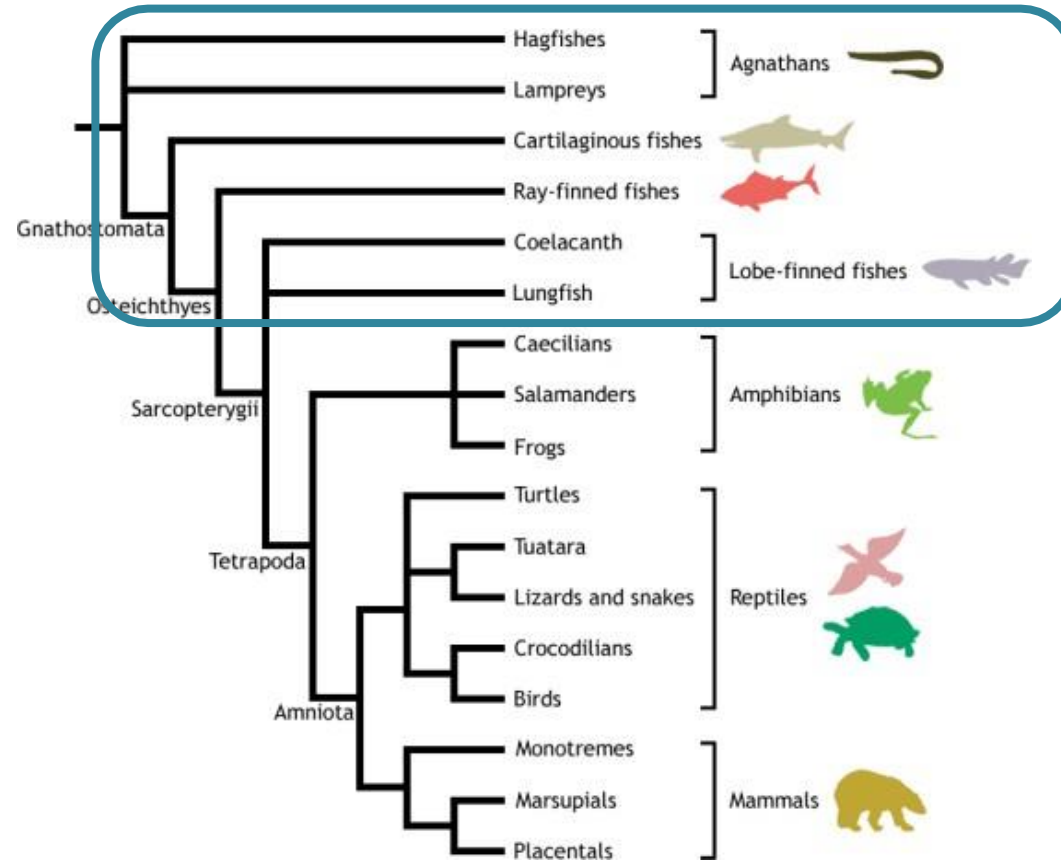
Remember the snip test

Note – the character marks are placed on vertical lines for this tree

iClicker Question

Are fish (circled in blue) an example of a monophyletic, paraphyletic or polyphyletic group?

- A. Monophyletic
- B. Paraphyletic
- C. Polyphyletic
- D. Not sure



Answer

Are fish (circled in blue) an example of a monophyletic, paraphyletic or polyphyletic group?

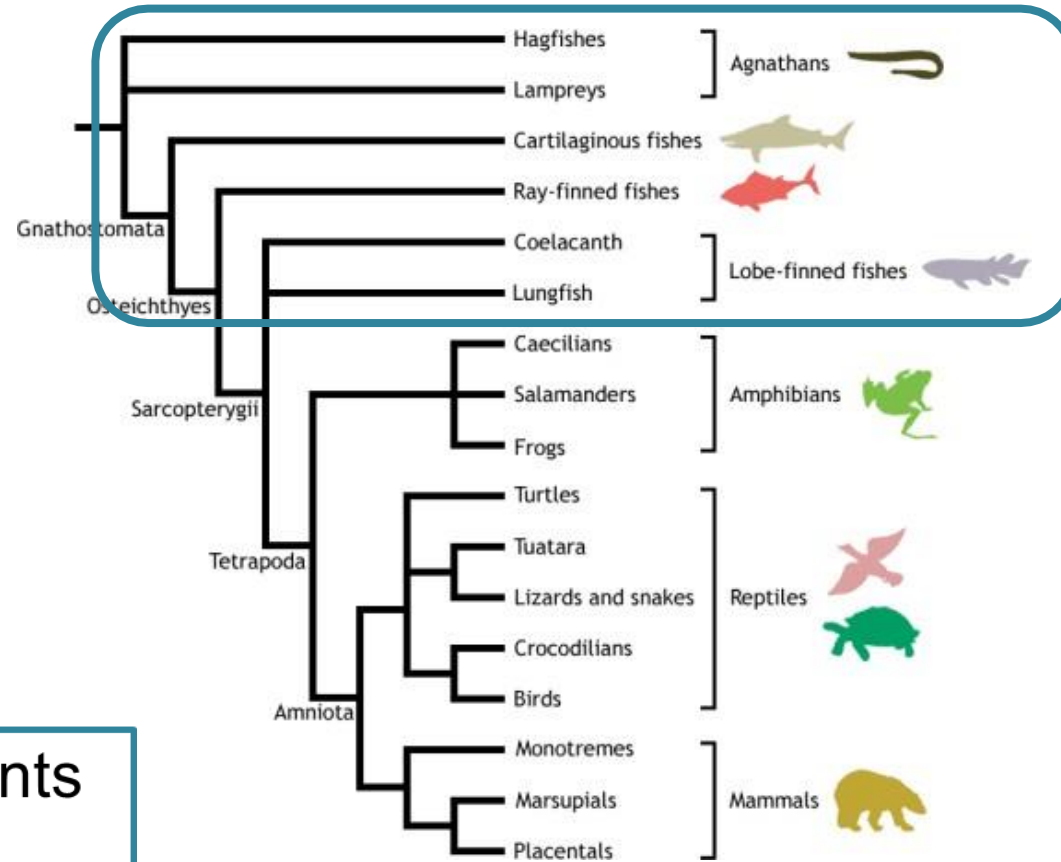
A. Monophyletic

B. Paraphyletic

C. Polyphyletic

D. Not sure

not all of the descents
of the common
ancestor included



Next class - Thursday

- Finish phylogenetic trees.
- Start Species Concepts & Speciation (not testable until final exam)

Learning goals

D2 Interpret phylogenetic trees in order to describe the evolutionary relationships among the organisms or taxa represented and infer the evolutionary history of specific traits and when they arose within the tree	<ul style="list-style-type: none">• Identify a clade, sister taxa, and most closely related taxa on a phylogeny and provide a logical justification explaining your reasoning.• Identify homologous and analogous characteristics based on a phylogeny and make predictions about the number of times a trait has evolved and where it evolved.• Determine if two trees for the same taxa show the same pattern of relationship among taxa and justify your conclusion with specific reference to trees.	<ul style="list-style-type: none">• Terms and concepts related to phylogenies (taxa, taxon, sister taxa, clades, nodes, branches, tips, common ancestor, most recent common ancestor, synapomorphy, shared derived character, homology, analogy, homoplasy, trait, character, parsimony).• Types of data use to construct trees.• Basic Linnean taxonomic groups.• Concept of monophyly as a criterion for delimiting taxonomic groups.
---	--	--

Plus add in:

- Be able to identify monophyletic, paraphyletic and polyphyletic groups (and be able to justify your conclusion).
- Be able to add character marks to a tree (using the principle of parsimony)
- Be able to predict the number of times a trait has been gained/lost (using the principle of parsimony)

Not testable:

- Basic Linnaean taxonomic groups.