

Exploring Evolution: Unveiling the History of Flightless Birds Through Phylogenetic Trees



World map demonstrating the geographical distribution of Ostrich, Rhea, Brown kiwi, Moa, Elephant Bird, and Cassowary

Image accreditations:

Moa: "Euryapteryx," by [FunkMonk](#) (Michael B. H.) / Wikimedia Commons, licensed under CC BY-SA 3.0

(<https://creativecommons.org/licenses/by-sa/3.0/>). No modifications were made. Original Moa

Elephant bird: By Acrocynus - Acrocynus, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=3830345>

image available at: <https://commons.wikimedia.org/wiki/File:Euryapteryx.jpg>

Ostrich: By MathKnight - Own work, CC BY-SA 4.0, <https://commons.wikimedia.org/w/index.php?curid=5010729>

Rhea: By Rufus46 - Own work, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=36109606>

Kiwi: By Maungatautari Ecological Island Trust - <http://www.maungatrust.org/news/default.asp>, Public Domain, <https://commons.wikimedia.org/w/index.php?curid=2450957>

Tinamou: By Allan Drewitt - <https://www.flickr.com/photos/31003918@N05/6437912125/in/photostream>, CC BY 2.0, <https://commons.wikimedia.org/w/index.php?curid=18145163>

Map created with mapchart.net

Workshop developed by Carla H Finn¹

Acknowledgement and thanks to:

Wilda Laux² for thoughtful edits and feedback.

Peter Ritchie¹ for feedback and help with concept.

¹ Te Herenga Waka – Victoria University of Wellington

² Genomics Aotearoa

Introduction:

In this activity, students will explore the evolutionary relationships among ratites by performing two complementary activities:

1. **Manual DNA Sequence Alignment:** In the first part, students will complete a hands-on activity to manually align simplified nucleotide sequences from different ratite species. By identifying similarities and differences in the sequences, they will gain an understanding of how sequence alignments work.
2. **Computer-Based Sequence Alignment:** In the second part, students will use a computational tool (**Clustal Omega**) which aligns actual nucleotide sequences from more ratite species. This tool allows the analysis of much longer and more complex sequences quickly and accurately. From the computer-generated alignment, students will visualise a phylogenetic tree that represents the evolutionary history of the ratites.

Together, these exercises will give students a deeper appreciation of how genetic data can be used to map the evolutionary paths of species and the benefits of using bioinformatics tools in research.

Part 1 – Manual DNA Sequence Alignment

The goal of manually aligning short DNA sequences is to help students visualize how geneticists identify conserved regions and infer relationships based on genetic similarities. By counting differences between sequences, one can determine how closely related species are and which species likely share a more recent common ancestor.

Materials:

- A3 sheet of strips of portions of a nucleotide sequence ranging from 122-132bp long (with a 121bp aligned ('matching') sequence from Ostrich, Rhea, Kiwi, Moa, and Cassowary).
- Scissors
- Tape

Aims:

To manually align simplified nucleotide sequences to understand how sequence alignments work and how differences in DNA can be used to build a basic phylogenetic tree.

Methods:

Step 1: Align the DNA Sequences Manually

1. **Prepare the Sequences:**
 - Cut the five strips of DNA sequences provided.
 - Spread out the paper strips for each species so the strips are side by side.
2. **Align the Sequences:**
 - Begin aligning the sequences by matching identical nucleotides (A, T, G, C) across the species.
 - Look for regions where the sequences are the same; these are called conserved regions.
 - Paste the aligned sequences together.

Answer: (here, only the start of the DNA sequence strips are shown)

Note how the 'start' of the alignment differs among the species. Either the sequence itself (e.g., sequence evolution), or the quality of the DNA/data we have (e.g., degradation), can influence this.



Consider these example sequences for Ostrich and Rhea:

- **Ostrich:** G A T **G** T
- **Rhea:** G A T **C** T

Notice that the first three nucleotides (G, A, T) are the same, but the fourth nucleotide differs (G for Ostrich and C for Rhea). This is a single-nucleotide polymorphism, or 'SNP'.

Note: The nucleotide sequences provided in this activity are simplified and curated to reflect real phylogenetic relationships. In actual research, longer sequences or entire genomes are used for greater accuracy, but manual alignment of such sequences would be impractical.

Step 2: Count the Number of Differences and Fill in the Matrix of Differences

1. **Pairwise Comparison:**
 - For each pair of species, count the number of positions where the nucleotides differ.
2. **Record the Differences:**
 - Use the provided matrix on the next page to record the number of differences between each pair of species.

For example:

If when comparing the Ostrich and Rhea:

- **Ostrich:** G A T **G** T
- **Rhea:** G A T **C** T

There is **1 difference** at the fourth position. This counts for 1. Continue to count how many more differences there are along their DNA sequences and record this number in the matrix.

3. Continue with the other species.

Matrix of Differences:

	Ostrich	Rhea	Kiwi	Moa	Cassowary
Ostrich	0				
Rhea		0			
Kiwi			0		
Moa				0	
Cassowary					0

Answer:

	Ostrich	Rhea	Kiwi	Moa	Cassowary
Ostrich	0				
Rhea	16	0			
Kiwi	21	6	0		
Moa	22	7	1	0	
Cassowary	20	4	2	3	0

Step 3: Interpret the Matrix

1. Analyse the Results:

- Look at the numbers in your matrix.
- Identify which species pairs have the fewest differences and which have the most.

2. Consider Evolutionary Implications:

- Species with fewer differences are likely more closely related.
- Species with more differences likely diverged earlier in evolutionary history.

Questions to students:

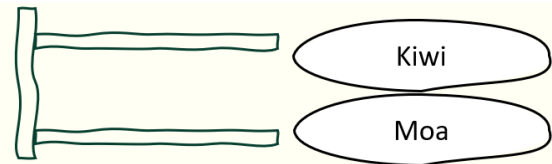
- **a.** Which species pairs have the **fewest differences**, indicating they are closely related?
(Answer: Kiwi and Moa)
- **b.** Which species has the **most differences** from the others, suggesting it diverged earlier?
(Answer: Ostrich)

Step 4: Building a Basic Phylogenetic Tree

Instructions:

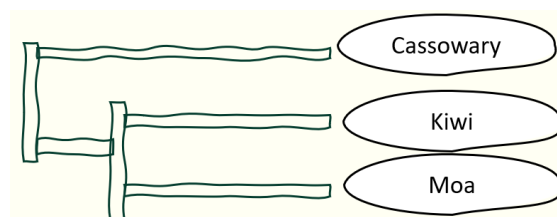
1. Create a Diagram:

- Using the data from the matrix, draw a phylogenetic tree.
- Start with the species pair that has the least difference (the kiwi and moa; 1 difference), and draw a bracket connecting them:



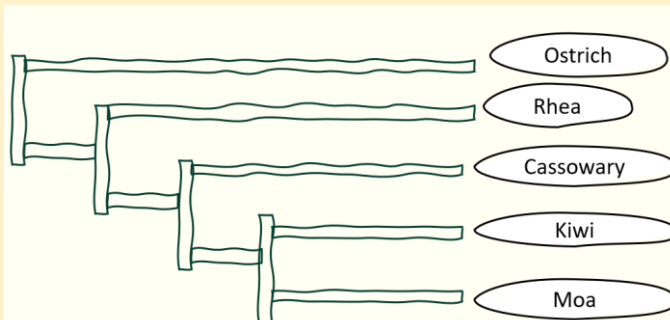
2. Determine Branching Order:

- Then, connect the next pair with fewest differences (the kiwi and cassowary; 2 differences):



- Continue adding species pairs.

Final phylogenetic tree example:



In this exercise, focus on how (nucleotide) differences between species indicate their evolutionary relationships. Fewer differences represent a closer relationship, meaning the species share a more recent common ancestor, while more differences suggest a more distant relationship and an older common ancestor.

Reflection and Discussion:

Questions:

1. Most Divergent Species:

- Which species has the most differences from the others? What might this suggest about its evolutionary history?

The **Ostrich** differs the most from the other species, suggesting that it has the longest independent evolutionary history. This means it diverged from the common ancestors of these other flightless birds much earlier.

2. Limitations of Manual Alignment:

- What limitations arise from manually aligning sequences and building a tree? (see below)
- How might computer-assisted methods differ in accuracy and ease? (see below)

Conclusion and what's next?

The above manual alignment process mirrors the approach used by a popular computational algorithm known as **Neighbor-Joining (NJ)**. This algorithm is widely used in bioinformatics for constructing phylogenetic trees because it provides a **fast and efficient method** for inferring evolutionary relationships between species.

What is Neighbor-Joining?

Neighbor-Joining (NJ) is an algorithm that estimates the evolutionary distances between species based on **genetic similarities and differences** in the sequences you align. The NJ method starts with the assumption that all species are equally related and then iteratively joins the "closest" neighbors—those with the smallest genetic distances—until a complete tree is constructed.

While Neighbor-Joining provides a **good first approximation** of a phylogenetic tree, more sophisticated statistical methods are now commonly used in research. These advanced methods, such as **Maximum Likelihood** and **Bayesian Inference**, can offer higher accuracy by accounting for more complex evolutionary processes, but they are still based on the same fundamental principles: the alignment of DNA (or protein) sequences and the assumption that **shared genetic similarities** indicate a **closer evolutionary relationship**.

Why Use Computer Algorithms for Phylogeny?

While manual alignment of short sequences helps illustrate the principles behind phylogenetic tree construction, in reality, the sequences scientists work with in research are often thousands to millions of nucleotides long. In theory, the longer the sequences, then the more information we have to confidently assess the real genetic content of an organism.

Comparing and aligning these longer sequences by hand would be impractical, if not impossible. That's where bioinformatics tools like **Clustal Omega** and algorithms like **Neighbor-Joining** come in.

- **Speed and Accuracy:** Computer algorithms can quickly calculate genetic distances and align large datasets, producing accurate trees that reflect millions of years of evolutionary history in minutes.
- **Handling Complex Data:** With modern tools, we can account for a variety of factors, such as **insertions, deletions, and mutations**, which happen over evolutionary time and complicate the relationships between species.
- **Statistical Support:** More advanced algorithms can provide **statistical support** for each branch of the tree, indicating the level of confidence in the relationships being proposed.

Table: pros and cons of manual vs computer alignment.

Manual alignment	Computer alignment (Clustal Omega)
Provides a tactile understanding of sequence alignment and phylogeny construction.	Handles much larger and more complex sequences in a fraction of the time.
Limited to simplified, shorter sequences.	Generates a more accurate and comprehensive phylogenetic tree.
Useful for grasping fundamental concepts but impractical for real-world genetic data.	Reflects the real power of bioinformatics in understanding evolutionary relationships at the molecular level.

Part 2: Computer-Based Sequence Alignment and Phylogeny Construction

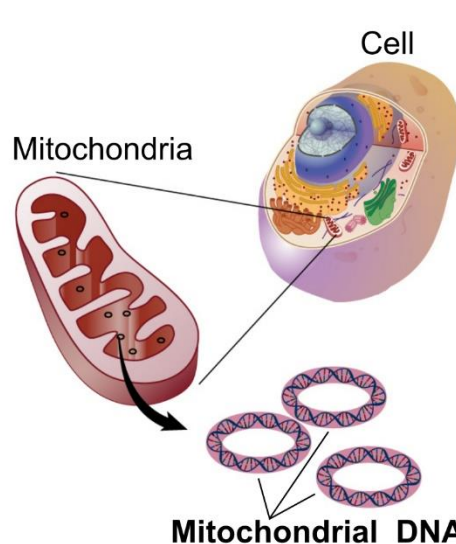
In this part, students will use **Clustal Omega**, a bioinformatics tool designed to align DNA sequences from multiple species, helping with quickly and efficiently observing the evolutionary relationships between them.

Why Use Clustal Omega?

Clustal Omega can analyse real-world datasets that are often too large or complex to handle by hand. The computer algorithm can efficiently align sequences, account for evolutionary events (e.g., insertions, deletions), and generate a detailed phylogenetic tree that reflects the evolutionary history of the species you're studying.

The sequence we're using: Cytochrome B

For this exercise, the students will use the Cytochrome B gene, which is part of the mitochondrial genome. Here's why Cytochrome B is particularly useful for phylogenetic studies:



Cytochrome B is a protein-coding gene found in the mitochondria of all animals. While its function is conserved across species, the gene accumulates mutations at a steady rate over evolutionary time, which provides the variation necessary to distinguish between different species.

This balance between conservation (because it's part of the essential mitochondrial function) and variability (due to mutations over time) makes Cytochrome B an excellent marker for studying evolutionary relationships. It allows us to track both recent divergences (as seen in ratites) and older evolutionary splits.

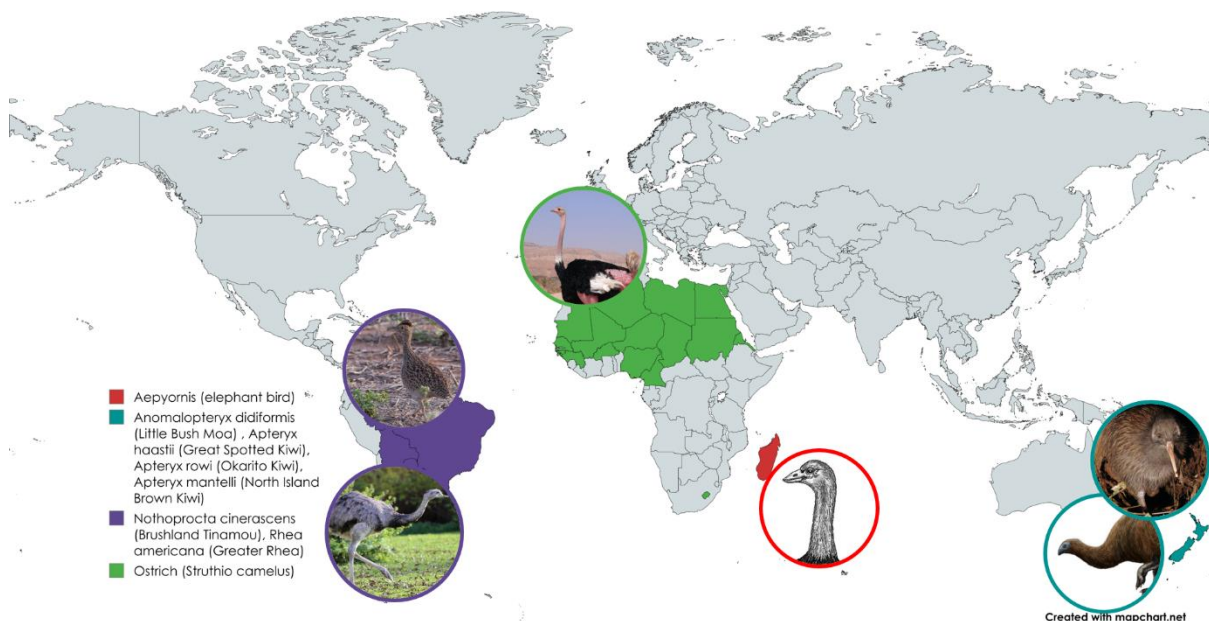
Image: By National Human Genome Research Institute - National Institutes of Health. National Human Genome Research Institute. "Talking Glossary of Genetic Terms". Retrieved November 17, 2016, from, Public Domain, <https://commons.wikimedia.org/w/index.php?curid=53240683>

Why Cytochrome B for Ratites?

Ratites (flightless birds) are a relatively young evolutionary group, and Cytochrome B evolves at a rate that is particularly useful for resolving relationships between species that diverged more recently. The variation in this gene between species provides just the right level of genetic difference to reveal how these species are related. For example, it helps us explore the evolutionary connections between the kiwi, ostrich, rhea, and other ratites.

The ratite species we're analysing today:

- *Ostrich*
- *Rhea*
- *Kiwi* (Great Spotted, North Island Brown, and Okarito)
- *Moa*
- *Elephant Bird*
- *Brushland Tinamou*



World map demonstrating the geographical distribution of the ratite species used in this activity.

Image accreditations:

Moa: "Euryapteryx," by [FunkMonk](#) (Michael B. H.) / Wikimedia Commons, licensed under CC BY-SA 3.0 (<https://creativecommons.org/licenses/by-sa/3.0/>). No modifications were made. Original Moa

Elephant bird: By Acrocynus - Acrocynus, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=3830345> image available at: <https://commons.wikimedia.org/wiki/File:Euryapteryx.jpg>

Ostrich: By MathKnight - Own work, CC BY-SA 4.0, <https://commons.wikimedia.org/w/index.php?curid=5010729>

Rhea: By Rufus46 - Own work, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=36109606>

Kiwi: By Maungatautari Ecological Island Trust - <http://www.maungatruster.org/news/default.asp>, Public Domain, <https://commons.wikimedia.org/w/index.php?curid=2450957>

Tinamou: By Allan Drewitt - <https://www.flickr.com/photos/31003918@N05/6437912125/in/photostream>, CC BY 2.0, <https://commons.wikimedia.org/w/index.php?curid=18145163>

Map created with mapchart.net

Information for each of the ratite species included in this activity.

Species	Origin	Appeared (Estimate in Mya)	Extinction (If Relevant)
Aepyornis (Elephant Bird)	Madagascar	Quaternary (~2.6)	17th century
Anomalopteryx didiformis (Little Bush Moa)	New Zealand	Miocene (~23-5)	14th century
Apteryx haastii (Great Spotted Kiwi)	New Zealand	Pleistocene (~2.6)	Extant
Apteryx rowi (Okarito Kiwi)	New Zealand	Pleistocene (~2.6)	Extant
Apteryx mantelli (North Island Brown Kiwi)	New Zealand	Pleistocene (~2.6)	Extant
Nothoprocta cinerascens (Brushland Tinamou)	South America	Cenozoic (~65)	Extant
Ostrich	Africa	Pliocene (~5)	Extant
Rhea americana (Greater Rhea)	South America	Pleistocene (~2.6)	Extant

Methods:

Step 1: Align DNA Sequences Using Clustal Omega

1. In your web browser, go to <https://www.ebi.ac.uk/jdispatcher/msa/clustalo>
2. Copy and paste the DNA sequences provided in the text file, 'ratite_sequences.fasta' into the input box.
 - Note: FASTA file stores DNA or protein sequences. Each entry starts with a header line that begins with ">", giving the name or ID of the sequence. Below the header, you'll see the sequence itself.
3. **Set Parameters:**
 - **Sequence Type:** Select **DNA**.
 - **Output Format:** Choose **ClustalW with character counts** (this helps visualize the positions better).
 - **Optional:** add a title to your alignment (e.g., "Ratite Phylogeny").
4. Click **Submit** to start the alignment process. This may take a few moments, depending on the number of sequences.

Step 2: Analyse the Alignment

1. Click on 'Alignment'.




Understanding the Output:

Each nucleotide is represented by a different colour, making it easy to spot similarities (conserved regions) and differences.

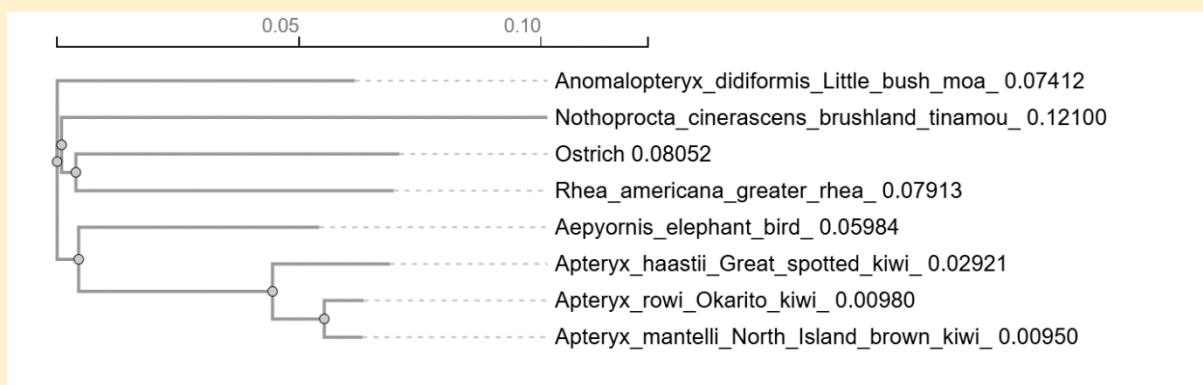
- **Conserved regions** (positions where all species have the same nucleotide) are key indicators of evolutionary conservation.
- **Variable regions** highlight differences, potentially reflecting evolutionary changes since the species diverged.

Step 3: Analyse a Phylogenetic Tree

Once you have the aligned sequences, Clustal Omega will automatically generate a phylogenetic tree based on the alignment. **Click on 'Phylogenetic Tree' to view this.** The tree shows evolutionary relationships, with the scale representing genetic distances between species.

- Click on the  to make the tree easier to read.
- Click on the  a few times to make the tree easier to read.
- Click on the  to sort the tree by the oldest/deepest clade.

The tree will look like:



Understanding the Phylogenetic Tree:

1. Interpret the Tree:

- **Branches:** Each branch represents a lineage (a species or group of species). The **length of the branches** indicates the genetic distance between species; longer branches suggest more genetic differences and earlier divergence.
- **Nodes:** A node is where two branches meet and represents a common ancestor of those species.
- **Scale Bar:** The scale bar at the bottom (e.g., 0.10) represents genetic distance. It shows the number of nucleotide differences per sequence length. For example, a distance of 0.1 means that, on average, 10% of nucleotides differ between the species.

How to Read the Tree:

- **Closely Related Species:** Species with short branches that meet at a node are closely related. For example, **Kiwi species (Great Spotted and Okarito)** should have shorter branches between them, indicating a more recent common ancestor.
- **Geographical Considerations:** Given the geographic distribution of ratites, you can hypothesize how continental drift and land connections (such as the ancient Gondwana supercontinent) might explain the patterns in the tree.

Questions: (Appendix 1, below, will provide you further material for discussion).

- Which species appear to be the most closely related, based on the tree?

The tree shows that the kiwi and the elephant bird are the most closely related species, even though they come from very distant places—New Zealand and Madagascar. This close relationship might be surprising because we often think that animals living closer together, like the kiwi and moa (both from New Zealand), would be more related. The molecular data, however, show that kiwi and elephant bird share a more recent common ancestor than the kiwi does with the moa.

- Do the molecular data support what you expected based on geography? How could this be explained?

The molecular data might not match our initial expectations based on geography. For example, you might expect that birds from the same place, like the kiwi and moa from New Zealand, would be closely related. However, the molecular data reveal that the kiwi is more closely related to the elephant bird from Madagascar, and the moa is more closely related to the tinamou from South America.

This unexpected pattern can be explained by looking at the ancient history of land movement. When Gondwana broke apart millions of years ago, groups of birds were carried to different continents, which isolated them and led to different evolutionary paths. So, the molecular data reflect these ancient separations rather than the birds' current locations.

Appendix 1: Additional teaching information

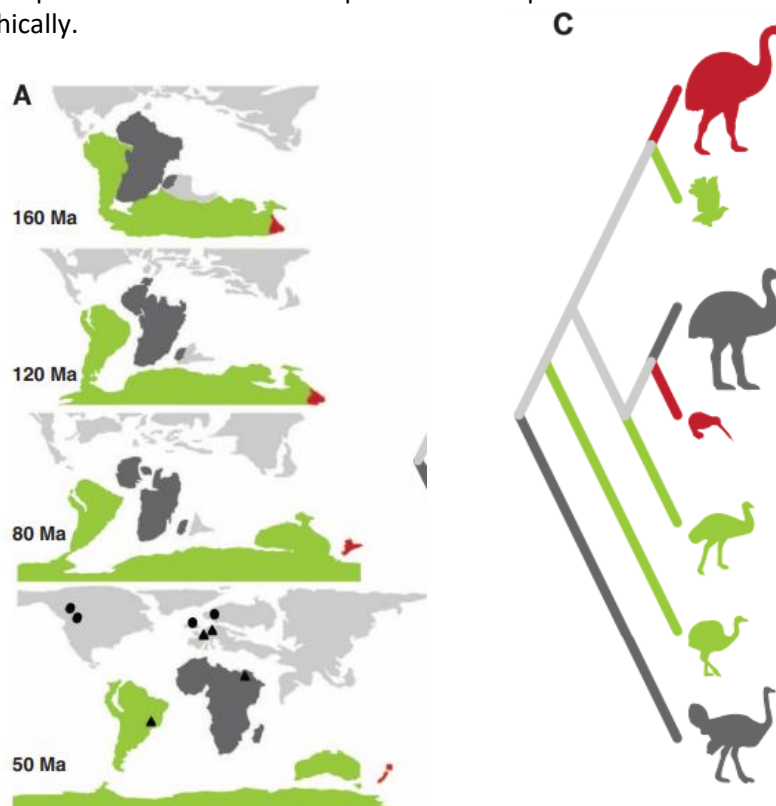
Understanding the Evolution of Ratites through Continental Drift:

The image below is from **Mitchell et al. 2014** (accessible here: https://www.science.org/doi/10.1126/science.1251981?url_ver=Z39.882003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%20%200pubmed) helps explain the evolutionary history of **ratites**—large, flightless birds such as the **kiwi**, **moa**, and **elephant bird**. Millions of years ago, all of these birds' ancestors lived on the supercontinent **Gondwana**, which was a single, joined landmass. As Gondwana began to split apart into different continents, the bird populations were carried along to these new landmasses, where they continued to evolve.

Key Points from the Image:

- **Different Continent Separation Times:** The continents separated at different times. **Africa and Madagascar** were the first to split away (about 100 to 130 million years ago), followed by **New Zealand** (around 60 to 80 million years ago), and finally **Australia, Antarctica, and South America** (about 30 to 50 million years ago).
- **Flightless vs. Flighted Birds:** The circles in the figure represent **birds that could fly**, while the triangles show **flightless birds**. As these birds' ancestors were carried to different continents, some groups, like the kiwis and moas, lost their ability to fly, while others, like the tinamous, retained it.

This figure illustrates how the movement of landmasses, combined with the evolution of bird species, helps us understand the complex relationships between birds that now live far apart geographically.



Why are the kiwi and moa more distantly related than we expect?

The reason the kiwi is more closely related to the elephant bird from Madagascar than to the moa, even though both kiwi and moa lived in New Zealand, lies in the way their ancestors were distributed across ancient landmasses. As the supercontinent Gondwana began to split up, different groups of birds ended up on separate landmasses, where they continued to evolve independently.

Initially, kiwi and elephant bird ancestors were on the same piece of Gondwana before Madagascar and New Zealand separated. This explains their close evolutionary relationship—they share a more recent common ancestor with each other than with the moa. When Madagascar separated from the rest of Gondwana, the elephant bird lineage stayed in Madagascar, while the kiwi ancestors eventually reached New Zealand, where they adapted uniquely to that environment.

Meanwhile, the moa's ancestors share a closer ancestor with the tinamou, a bird still found in South America. This reflects an ancient connection to the South American landmass that was once part of the same Gondwanan region. The tinamou retained flight, allowing some movement between landmasses, while the moa evolved to be flightless in New Zealand. Thus, continental drift and the bird populations' movement or isolation on landmasses created these surprising relationships across distant locations.

Summary of learning outcomes:

By completing this workshop, students should have gained an understanding of:

Evolution:

- **Impact of Continental Drift:** The breakup of Gondwana plays a key role in how species were distributed across the world. Understanding this helps explain why birds like the kiwi and elephant bird are related, even though they lived on distant landmasses.

Genomics:

- **DNA as the Blueprint of Life:** The workshop reinforces that DNA contains the genetic instructions for life, and small changes in the DNA sequence over time can lead to the formation of new species.
- **Genetic Similarities and Differences:** By comparing conserved regions (similar sequences) and variable regions (differences between species), students can learn how genetic variation drives evolutionary change.
- **Cytochrome B as a Molecular Marker:** By focussing on a mitochondrial gene like Cytochrome B, students are introduced to the concept of molecular markers—specific genes used to study genetic relationships across species. This gene highlights the balance between conservation (staying the same) and variability (mutations), helping students understand how scientists use genomic data to infer evolutionary relationships.

Bioinformatics

- **Power of Computational Tools:** Through the use of **Clustal Omega**, students learn how bioinformatics tools enable scientists to align large DNA sequences, which would be impossible to do manually with large datasets. This teaches the importance of bioinformatics in modern biology and its role in handling the vast amount of data generated in genomics research.
- **Phylogenetic Tree Construction:** Students learn how algorithms like **Neighbor-Joining** help visualize the evolutionary relationships between species by calculating genetic distances. This hands-on experience with bioinformatics gives students a taste of how researchers study evolutionary biology today.
- **Efficiency of Computational Methods:** While manual alignment helps students understand the basic concepts, bioinformatics tools demonstrate how modern science leverages technology to handle complex data and answer big questions, such as mapping the evolutionary history of species.