

Evolutionary analyses in bioinformatics

The light of evolution and the lab in a computer

Saioa Manzano-Morales

Toni Gabaldón

Comparative Genomics

Institute for Research in Biomedicine

A bit about me

A bit about me



Saioa Manzano-Morales
Comparative Genomics Lab

 saioa.manzano@bsc.es

 saimanzano

 MicroBioBits

A bit about me



2015

HS

Saioa Manzano-Morales
Comparative Genomics Lab



saioa.manzano@bsc.es



saimanzano



MicroBioBits

A bit about me



2015

HS

Now what?

Saioa Manzano-Morales
Comparative Genomics Lab



saioa.manzano@bsc.es



saimanzano



MicroBioBits

A bit about me

- I wanted to do **science**



2015

HS

Saioa Manzano-Morales
Comparative Genomics Lab



saihoa.manzano@bsc.es



[saimanzano](#)



[MicroBioBits](#)

A bit about me



2015

HS

- I wanted to do **science**
- I didn't like engineering (or maths)

Saioa Manzano-Morales
Comparative Genomics Lab



saihoa.manzano@bsc.es



[saimanzano](#)



[MicroBioBits](#)

A bit about me



2015

HS

- I wanted to do **science**
- I didn't like engineering (or maths)
- I was good at bio-things?

Saioa Manzano-Morales
Comparative Genomics Lab



saihoa.manzano@bsc.es



[saimanzano](https://github.com/saimanzano)



[MicroBioBits](#)

A bit about me



2015

HS

- I wanted to do **science**
- I didn't like engineering (or maths)
- I was good at bio-things?
- I didn't want to do Medicine

Saioa Manzano-Morales
Comparative Genomics Lab



saihoa.manzano@bsc.es



[saimanzano](https://github.com/saimanzano)



[MicroBioBits](#)

A bit about me



Saioa Manzano-Morales
Comparative Genomics Lab

-  saioa.manzano@bsc.es
-  saimanzano
-  MicroBioBits

2015

2019

HS

Bachelor

Biochemistry and
Molecular Biology



Universidad
del País Vasco
Euskal Herriko
Unibertsitatea

A bit about me



Saioa Manzano-Morales
Comparative Genomics Lab

-  saioa.manzano@bsc.es
-  saimanzano
-  MicroBioBits

2015

HS

2019

Bachelor

Biochemistry and
Molecular Biology

Now what? (x2)



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

A bit about me



Saioa Manzano-Morales
Comparative Genomics Lab

-  saioa.manzano@bsc.es
-  saimanzano
-  MicroBioBits

2015

2019

HS

Bachelor

Biochemistry and
Molecular Biology



Universidad
del País Vasco
Euskal Herriko
Unibertsitatea

- I enjoyed **research**

A bit about me



Saioa Manzano-Morales
Comparative Genomics Lab

- saioa.manzano@bsc.es
- saimanzano
- MicroBioBits

2015

HS

2019

Bachelor

Biochemistry and
Molecular Biology



Universidad
del País Vasco
Euskal Herriko
Unibertsitatea

- I enjoyed **research**
- I didn't like wet-lab

A bit about me



Saioa Manzano-Morales
Comparative Genomics Lab

- saioa.manzano@bsc.es
- saimanzano
- MicroBioBits

2015

HS

Bachelor

Biochemistry and
Molecular Biology



Universidad
del País Vasco
Euskal Herriko
Unibertsitatea

2019 2020

2021

MSc

Computational
Biology



POLITÉCNICA

A bit about me



Saioa Manzano-Morales
Comparative Genomics Lab

- saioa.manzano@bsc.es
- saimanzano
- MicroBioBits

2015

HS

Bachelor

Biochemistry and
Molecular Biology



Universidad
del País Vasco
Euskal Herriko
Unibertsitatea

2019 2020

2021

MSc

Computational
Biology



Looking for a PhD

A bit about me



Saioa Manzano-Morales
Comparative Genomics Lab

- saioa.manzano@bsc.es
- saimanzano
- MicroBioBits

2015

HS

Bachelor

Biochemistry and
Molecular Biology



2019 2020

2021

2022

MSc

JAE

Computational
Biology JAE Intro
fellow



A bit about me



Saioa Manzano-Morales
Comparative Genomics Lab

- saioa.manzano@bsc.es
- saimanzano
- MicroBioBits

2015

HS

Bachelor

Biochemistry and
Molecular Biology



2019 2020

MSc

JAE

Computational
Biology

JAE Intro
fellow

PhD Genetics
Comparative Genomics



Now

PhD

A bit about you

Instructions

Go to

www.menti.com

Enter the code

1200 9311



Or use QR code

A bit about what I do

A bit about what I do - my group

Comparative Genomics

Toni Gabaldón



Barcelona
Supercomputing
Center

Centro Nacional de Supercomputación



A bit about what I do - my group



Mecanismos de las Enfermedades Comparative Genomics



Dr. Toni Gabaldon
AFFILIATED GROUP LEADER

Telephone: +34 93 40 21077
Email: toni.gabaldon@irbbarcelona.org



A bit about what I do - my group



**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación



 @BSC-CNS
@Bioinfo4women



 @IRBBarcelona

A bit about what I do - my group



**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación



- High-speed computing



@BSC-CNS
@Bioinfo4women

A bit about what I do - my group



**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación



- High-speed computing
- 1h on MN = 46 yrs on a laptop



@BSC-CNS
@Bioinfo4women

A bit about what I do - my group



- Experimental part



 @IRBBarcelona

A bit about what I do - my group



- Experimental part
- The building you will get to see



 @IRBBarcelona

A bit about what I do - my group

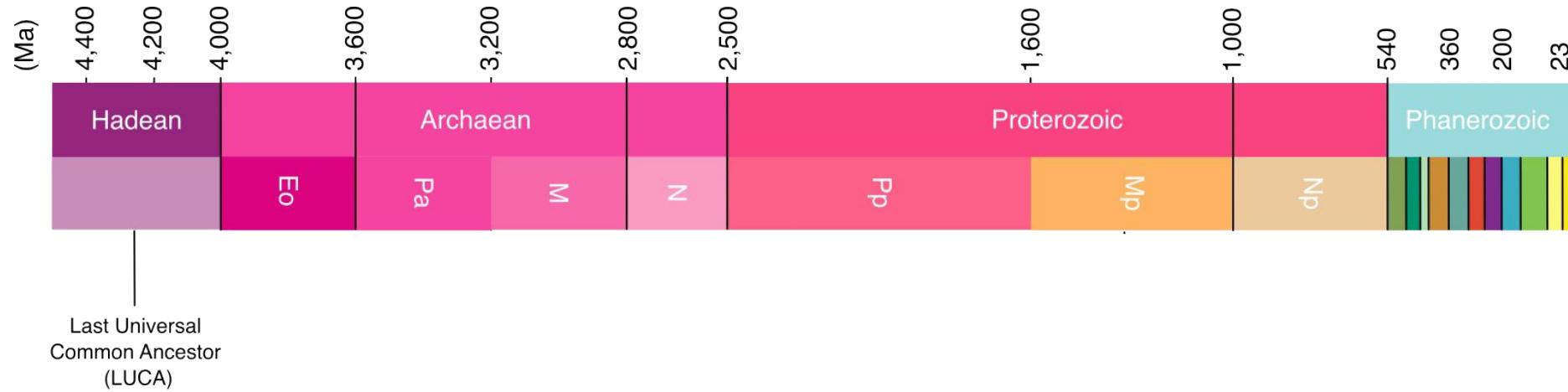


- Experimental part
- The building you will get to see
- My colleagues will probably show the labs better than me :)



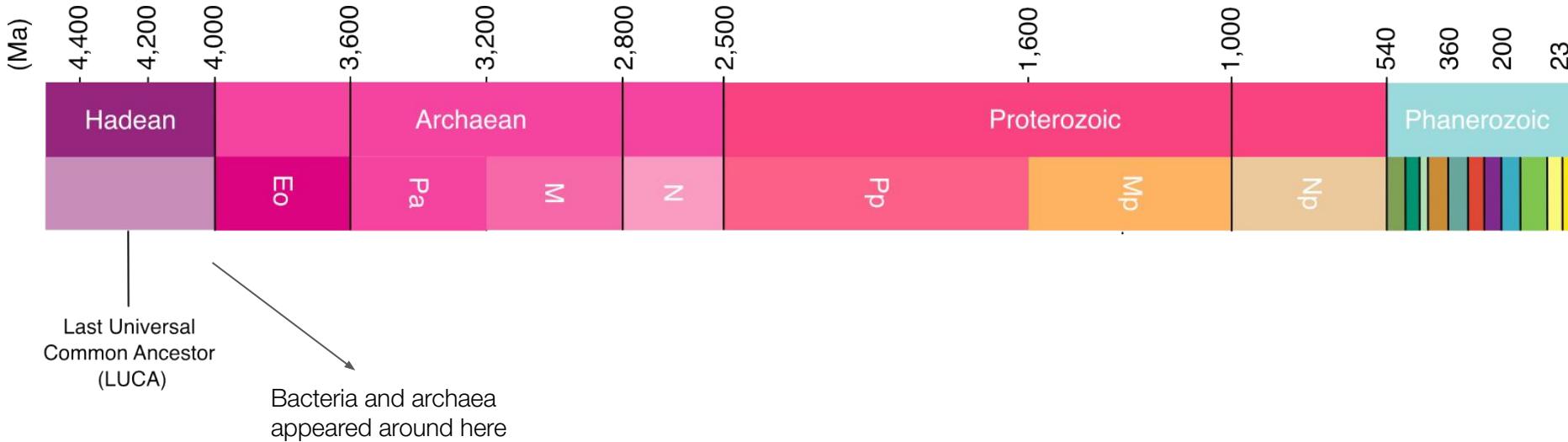
 @IRBBarcelona

A bit about what I do - my PhD

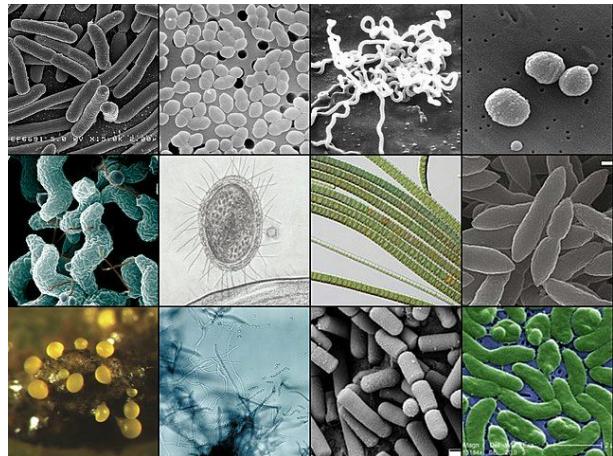


Start of life as we know it

A bit about what I do - my PhD

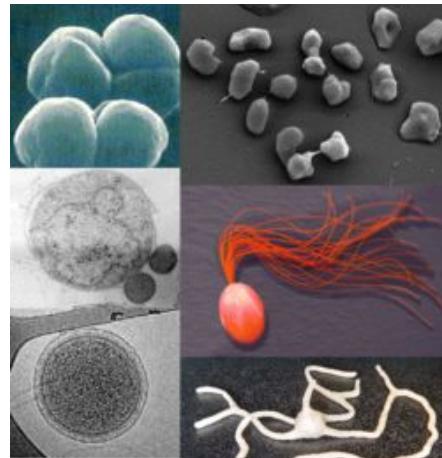


A bit about what I do - my PhD



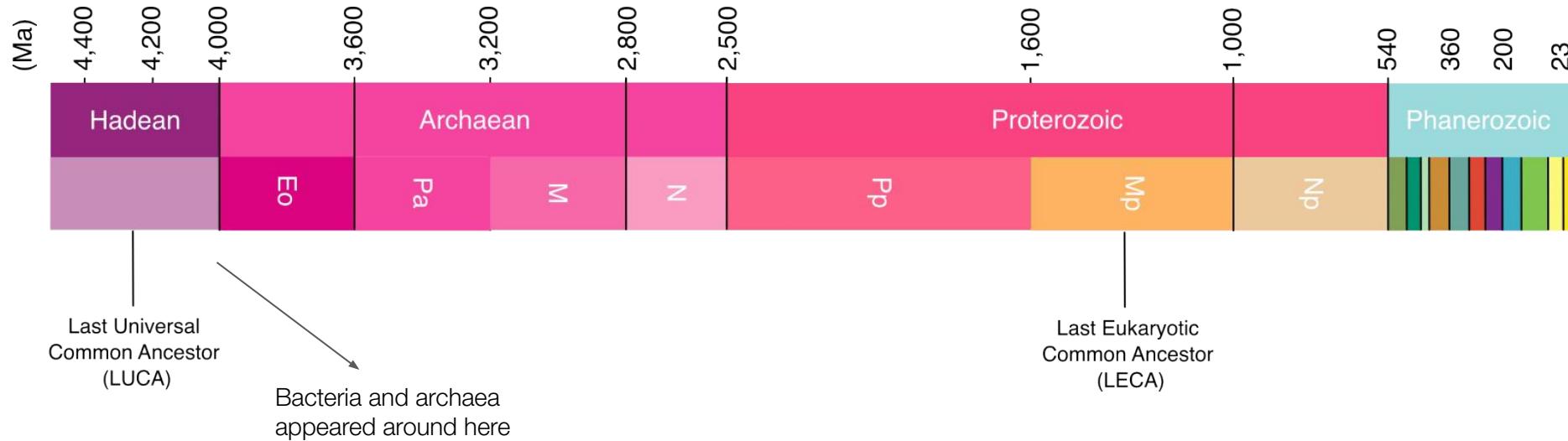
Bacteria

Prokaryotes

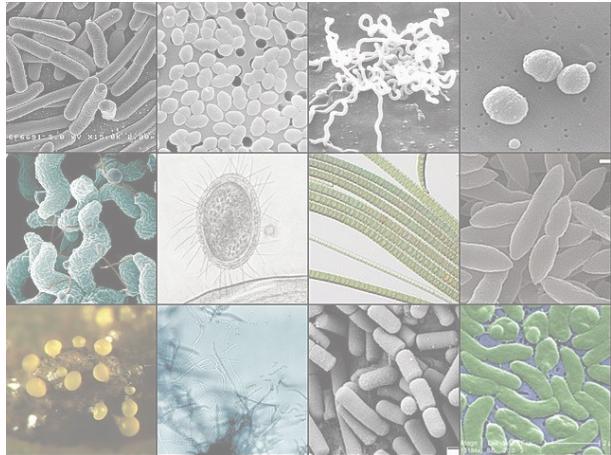


Archaea

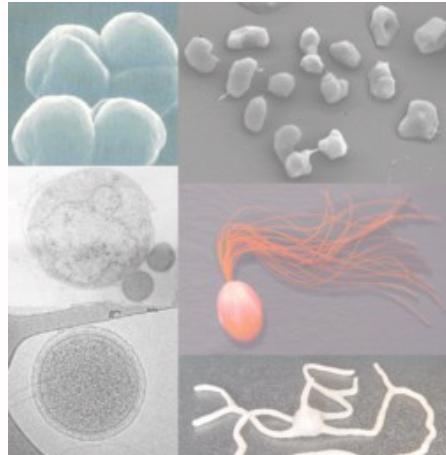
A bit about what I do - my PhD



A bit about what I do - my PhD



Bacteria

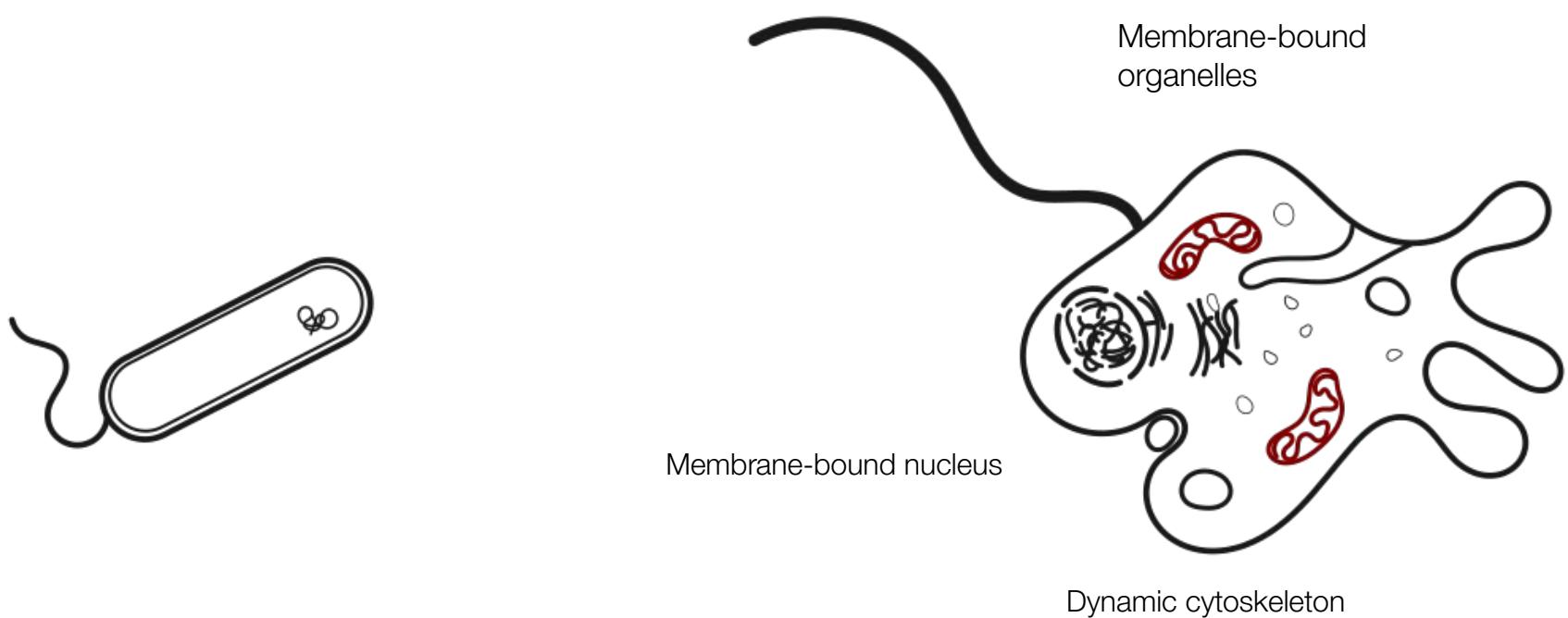


Archaea

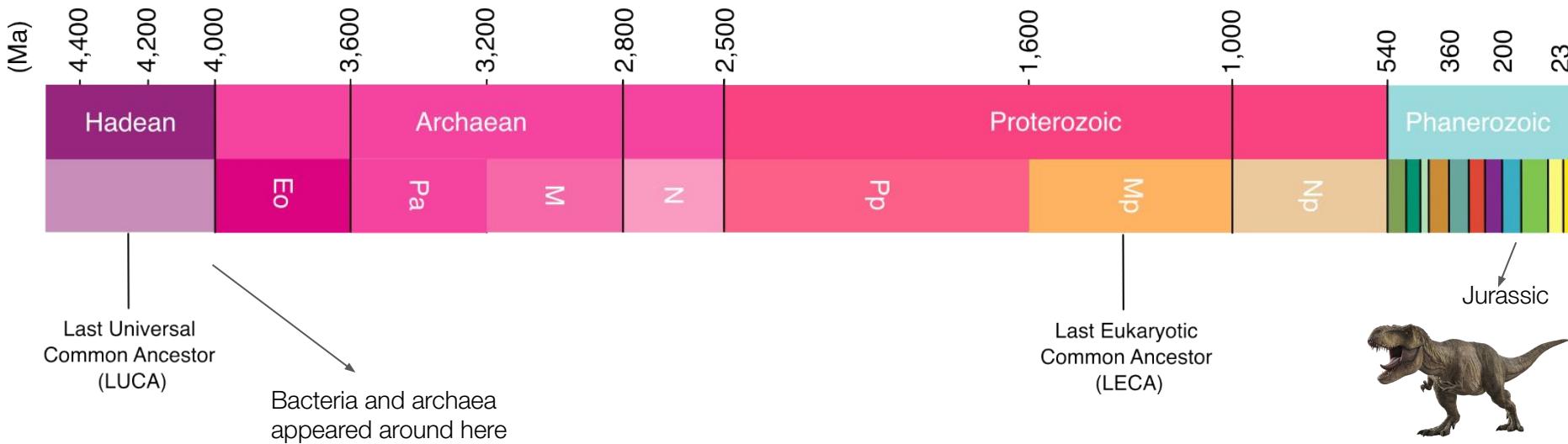


Eukaryota

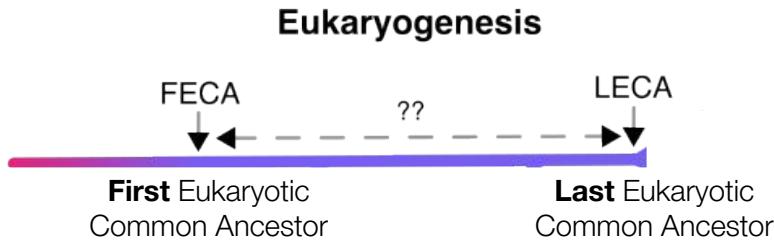
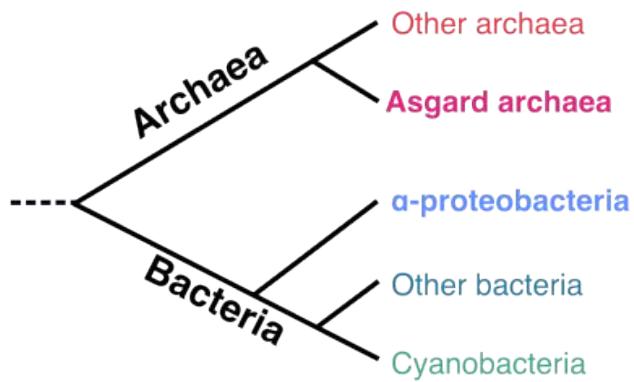
A bit about what I do - my PhD



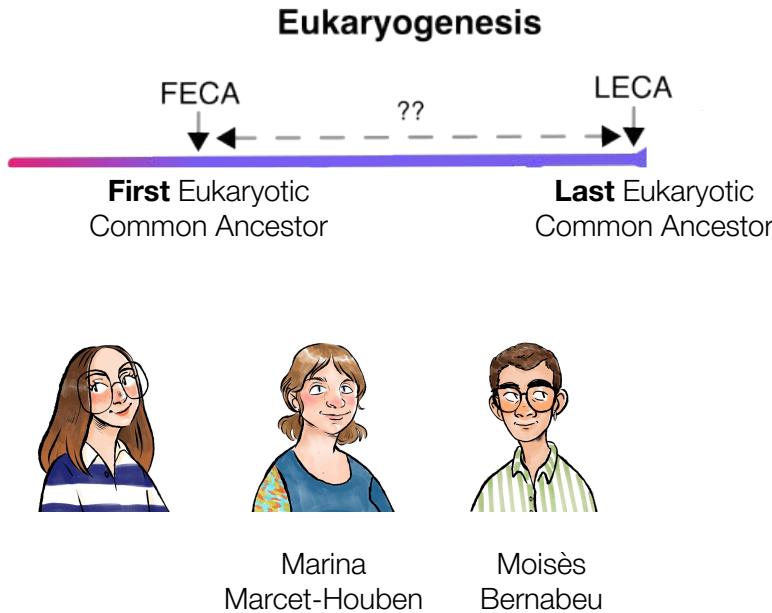
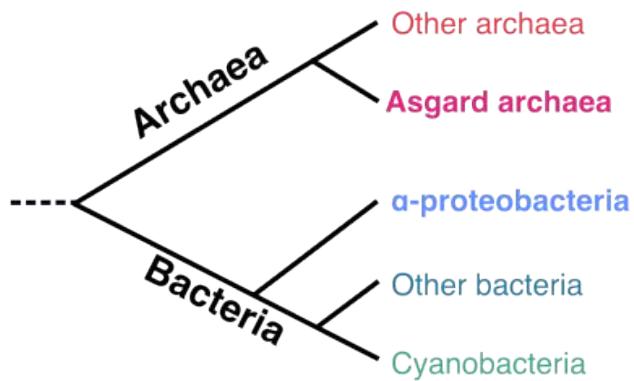
A bit about what I do - my PhD



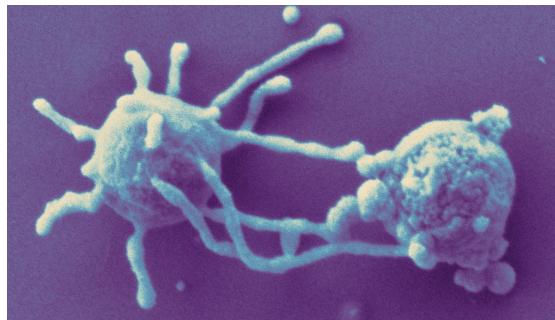
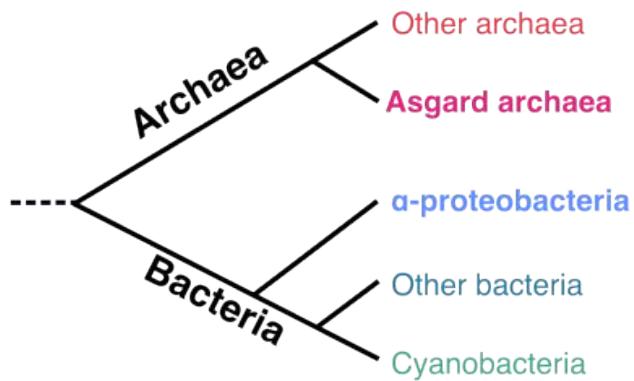
A bit about what I do - my PhD



A bit about what I do - my PhD



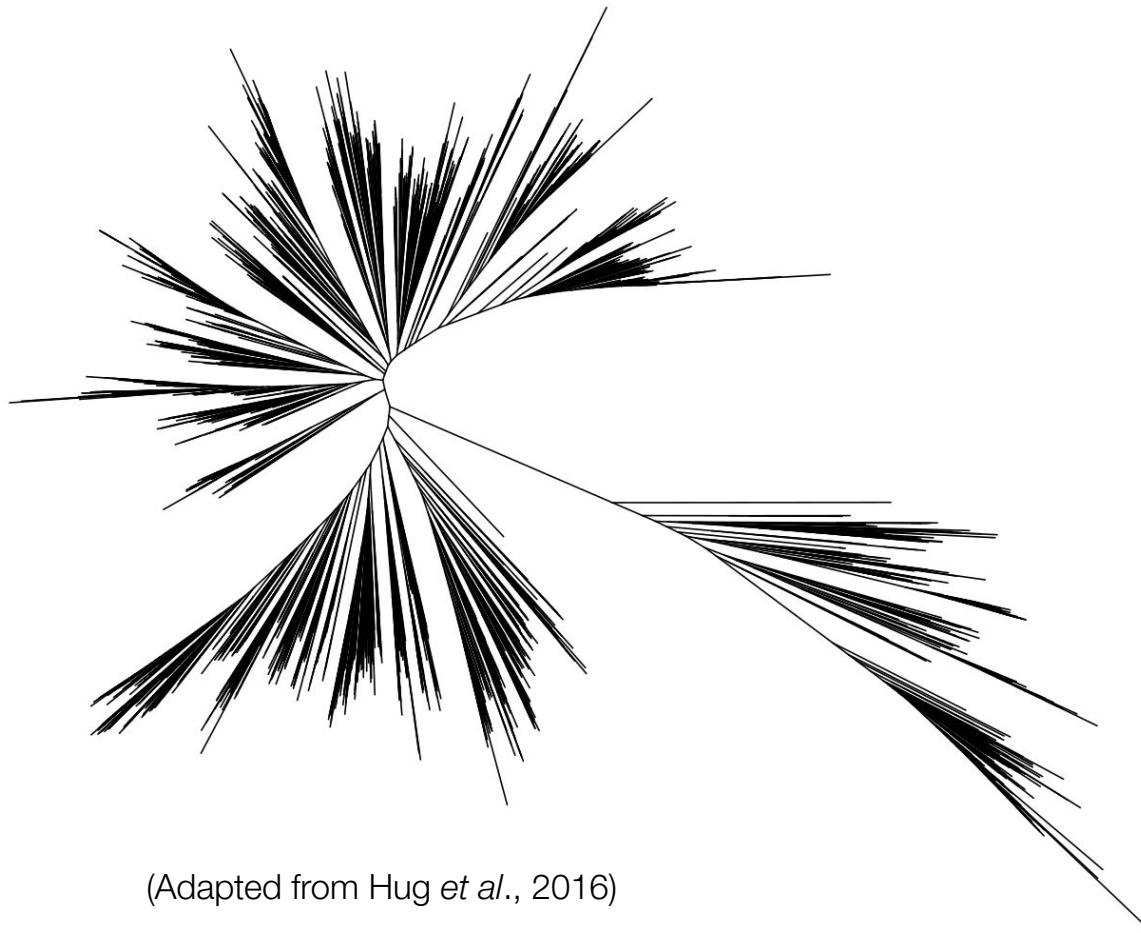
A bit about what I do - my PhD



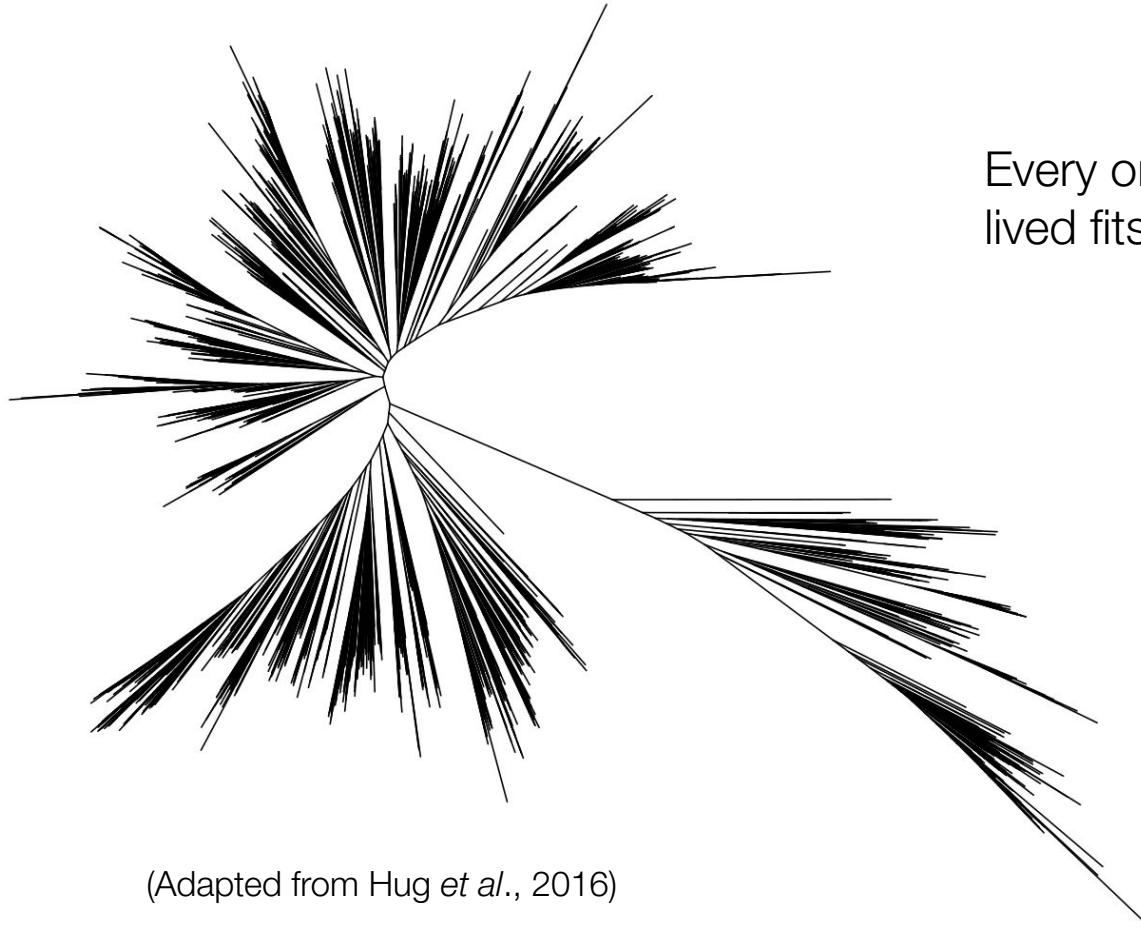
A bit about what we will do

ACT I: THE LIGHT OF EVOLUTION





(Adapted from Hug *et al.*, 2016)



Every organism that has ever lived fits here

(Adapted from Hug *et al.*, 2016)

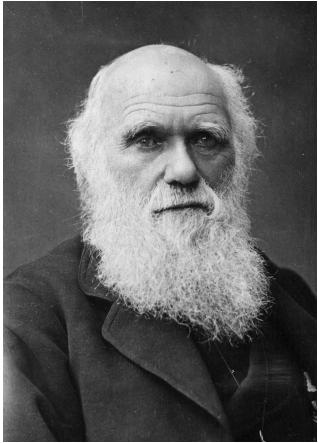
Evolution is the source of biodiversity on Earth



Evolution is the source of biodiversity on Earth

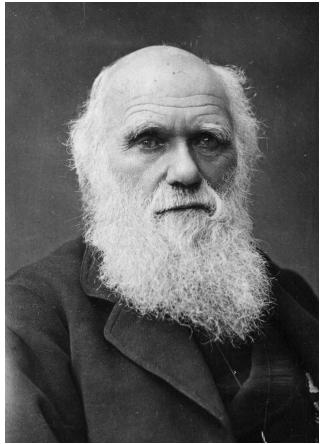


All species share a common ancestor and have changed through time to the variety of shapes and lifestyles we see today.



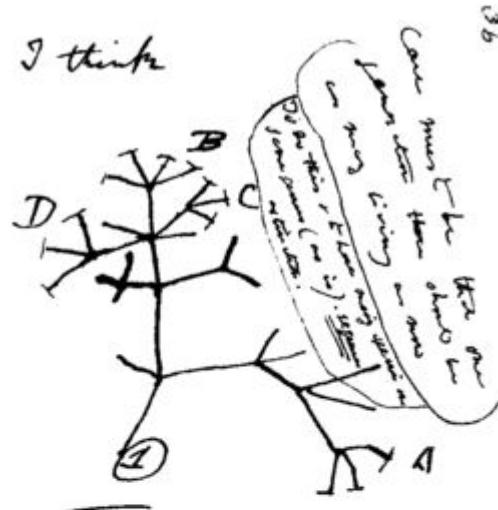
- This concept was proposed by Charles Darwin in 1859 in his book “On the Origin of Species”

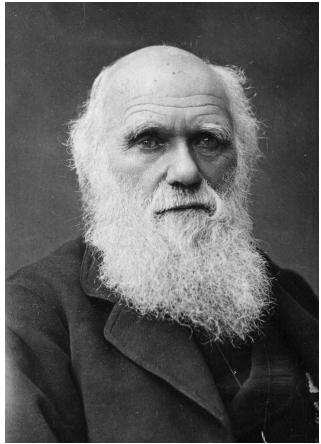
Charles Darwin



Charles Darwin

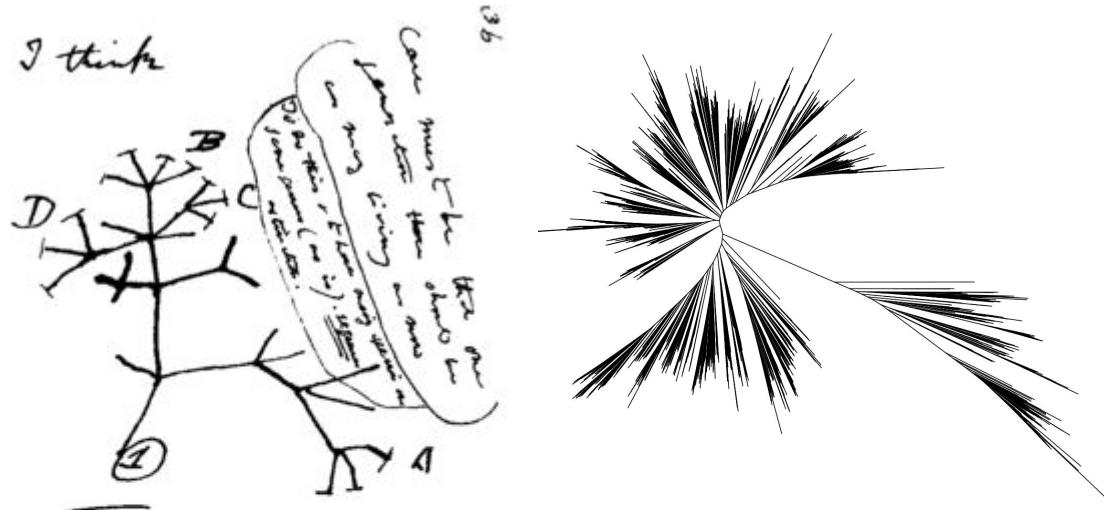
- This concept was proposed by Charles Darwin in 1859 in his book “On the Origin of Species”.
- He proposed that:
 - All species are related to each other through their ancestors

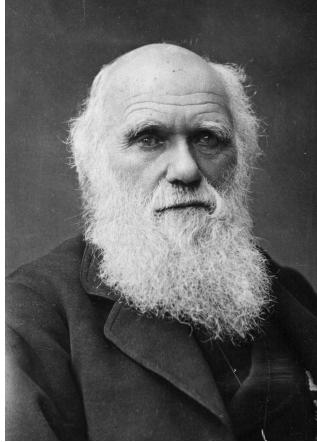




Charles Darwin

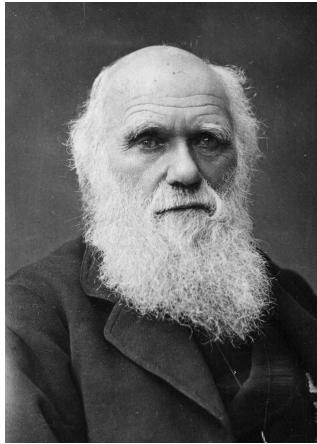
- This concept was proposed by Charles Darwin in 1859 in his book “On the Origin of Species”.
- He proposed that:
 - All species are related to each other through their ancestors





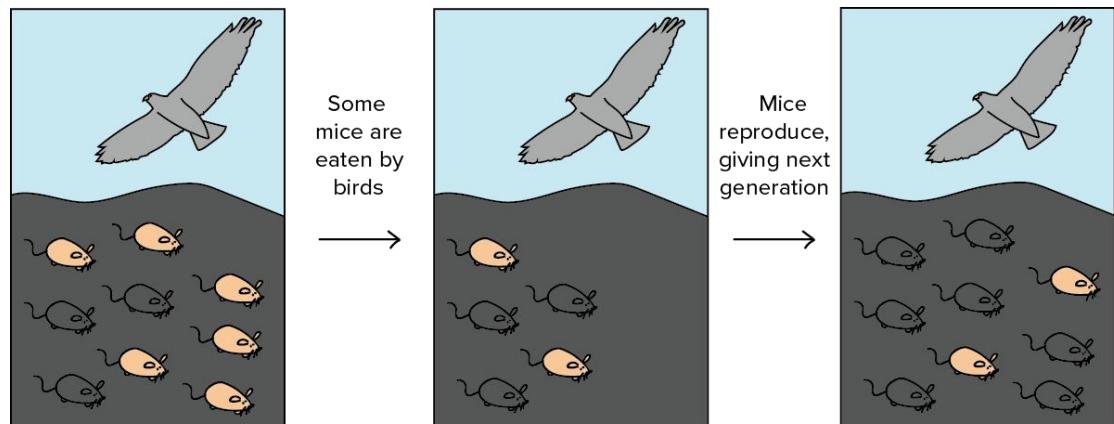
- This concept was proposed by Charles Darwin in 1859 in his book “On the Origin of Species”.
- He proposed that:
 - All species are related to each other through their ancestors
 - Species evolve through natural selection to adapt to their environment

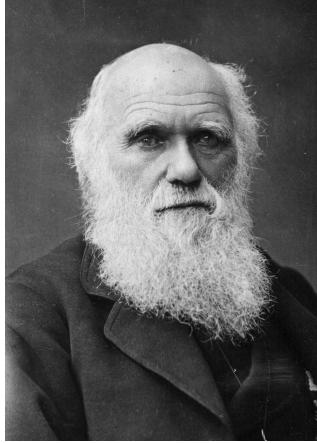
Charles Darwin



Charles Darwin

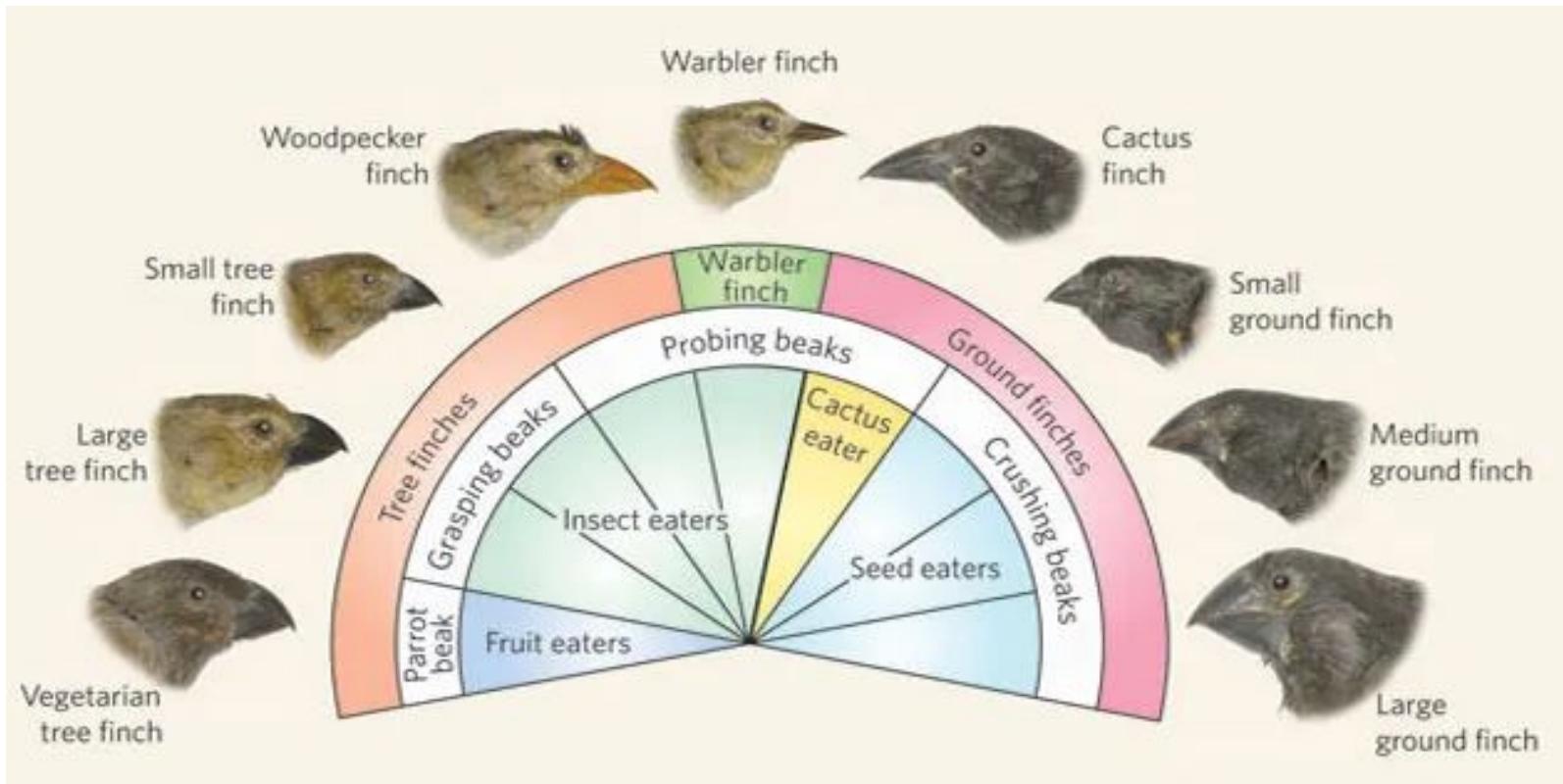
- This concept was proposed by Charles Darwin in 1859 in his book “On the Origin of Species”.
- He proposed that:
 - All species are related to each other through their ancestors
 - Species evolve through natural selection to adapt to their environment





Charles Darwin

- This concept was proposed by Charles Darwin in 1859 in his book “On the Origin of Species”.
- He proposed that:
 - All species are related to each other through their ancestors
 - Species evolve through natural selection to adapt to their environment
- In Darwin’s time, this was studied through morphological data

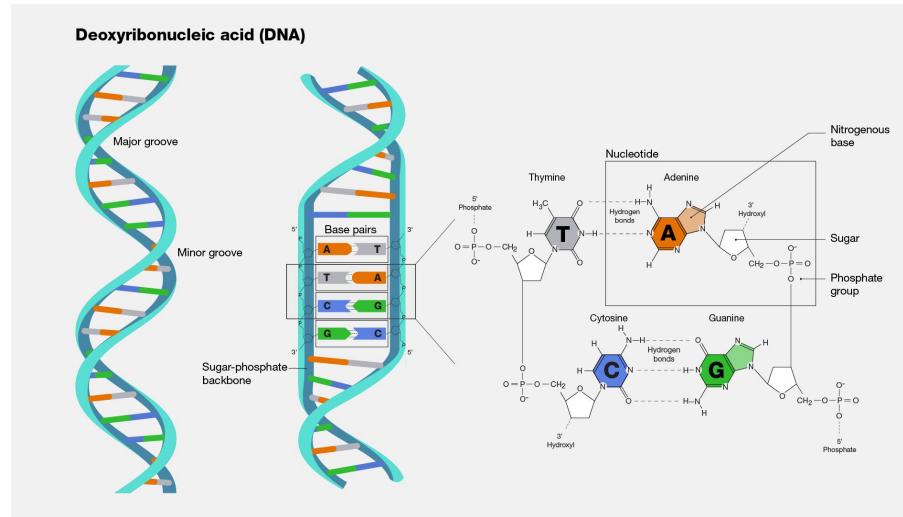


Evolution and phylogenetics

- We use **phylogenetics** to study the evolutionary history of life with molecular information

Evolution and phylogenetics

- We use **phylogenetics** to study the evolutionary history of life with molecular information
 - The set of instructions to build a cell (**and the whole organism!**) is stored in DNA



Evolution and phylogenetics

- We use **phylogenetics** to study the evolutionary history of life with molecular information
 - The set of instructions to build a cell (**and the whole organism!**) is stored in DNA

```
ATTAAGGTTTACCTTCCCAGGTAACAAACCAA  
CCAACTTTCGATCTCTGTAGATCTGTTCTCTAA  
CGAACCTTAAATCTGTGTGGCTGTCACTCGGCTG  
CATGCTTAGTGCACTCACGCAGTATAATTAAAC  
TAATTACTGTCGTTGACAGGACACGAGTAACCTGT  
CTATCTCTGCAGGCTGCTTACGGTTTCGTCCGTG  
TTGCAGCCGATCATCAGCACATCTAGGTTCTGTCC  
GGGTGTGACCGAAAGGTAAGATGGAGAGCCTTGTC  
CCTGGTTCAACGAGAAAACACACGTCCAACTCAG  
TTTGCCTGTTTACAGGTTCGCGACGTGCTCGTAC
```

DNA is a complex molecule, but for us, we can imagine it as a string of text

Evolution and phylogenetics

- We use **phylogenetics** to study the evolutionary history of life with molecular information
 - The set of instructions to build a cell (**and the whole organism!**) is stored in DNA



“Many years later, as he faced the firing squad, Colonel Aureliano Buendía was to remember that distant afternoon when his father took him to discover ice.”

Like books are organized into chapters, DNA is organized into genes

Like books are organized into chapters, DNA is organized into genes

Nelson Family Recipe Book

Table Of Contents

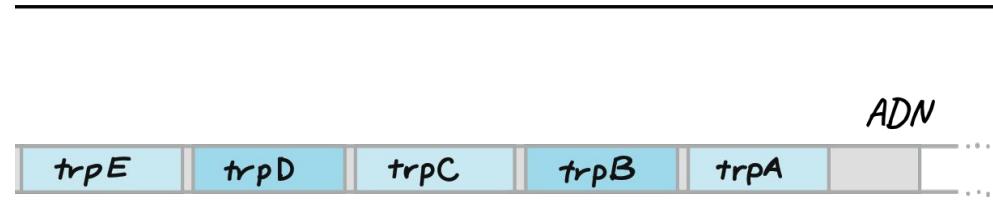
| | |
|--------------------------------------|-----|
| Breads | 3 |
| Soups | 15 |
| Vegetables | 23 |
| Salads | 29 |
| Main Dishes | 39 |
| Cake & Frostings | 65 |
| Cookies, Bars, and Lefse | 75 |
| Pies & Desserts | 93 |
| Beverages, Snacks, Jam, & Pickles | 103 |
| Index | 113 |

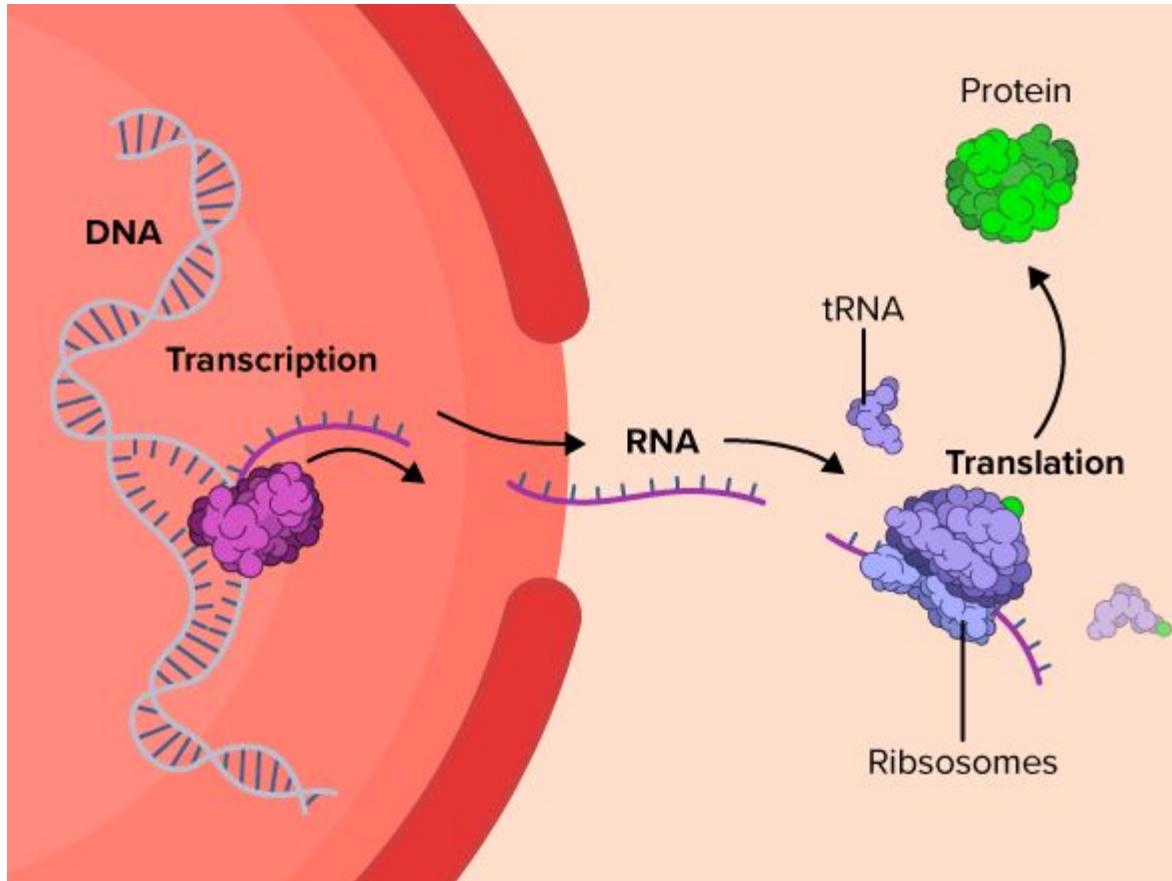
Like books are organized into chapters, DNA is organized into genes

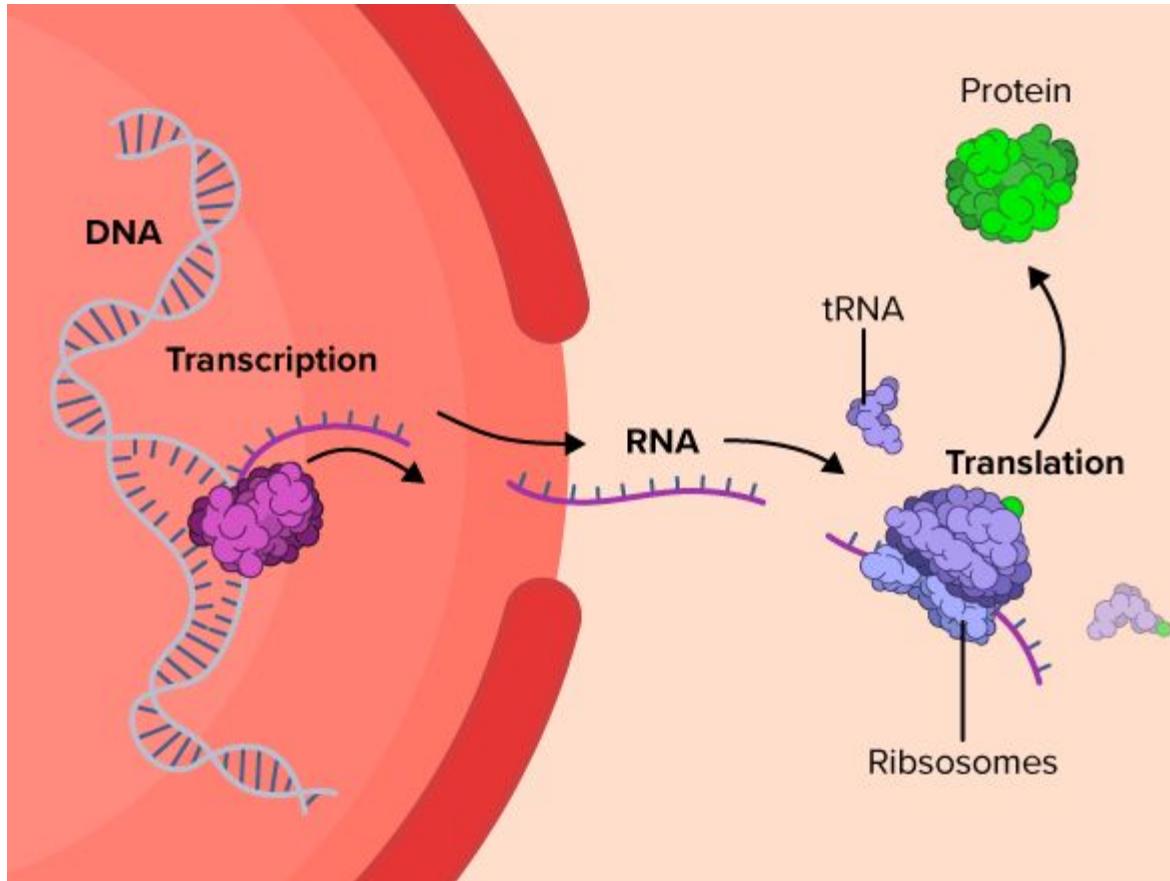
Nelson Family Recipe Book

Table Of Contents

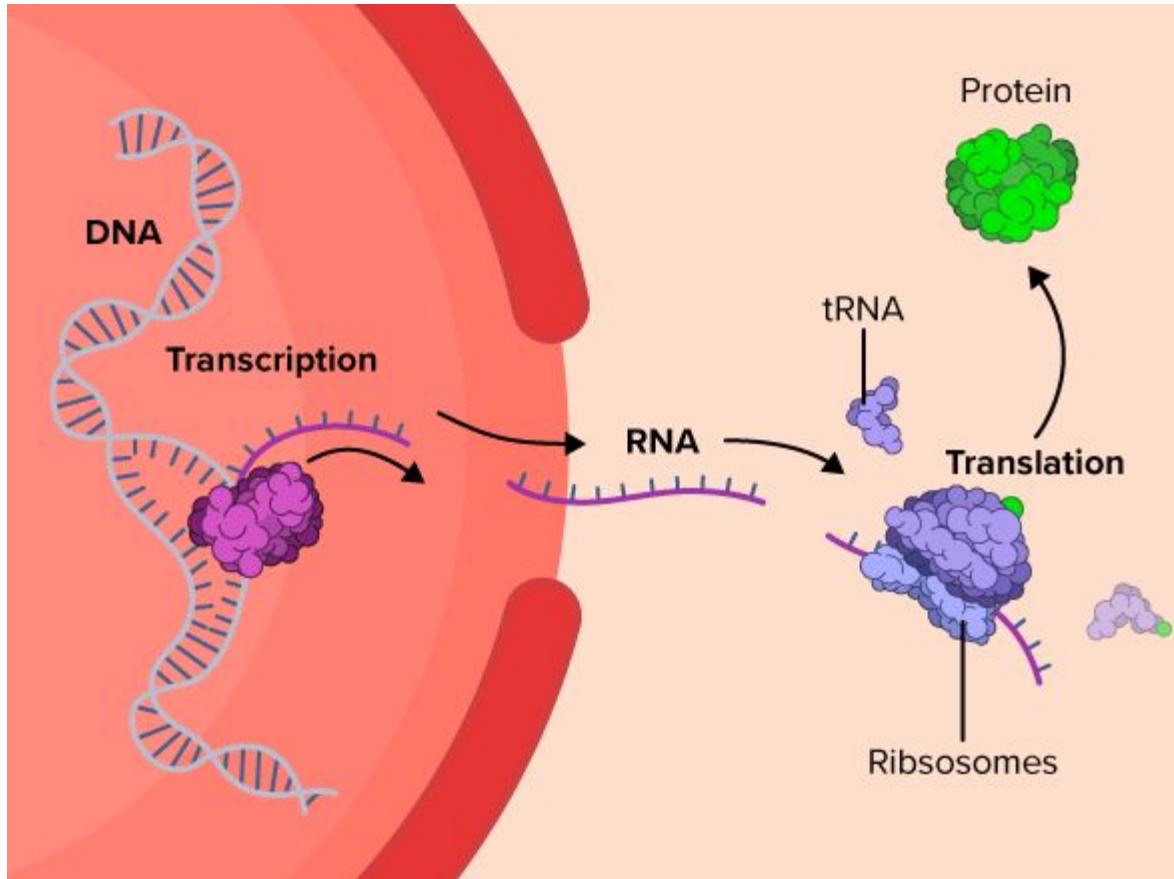
| | |
|--------------------------------------|-----|
| Breads | 3 |
| Soups | 15 |
| Vegetables | 23 |
| Salads | 29 |
| Main Dishes | 39 |
| Cake & Frostings | 65 |
| Cookies, Bars, and Lefse | 75 |
| Pies & Desserts | 93 |
| Beverages, Snacks, Jam, & Pickles | 103 |
| Index | 113 |



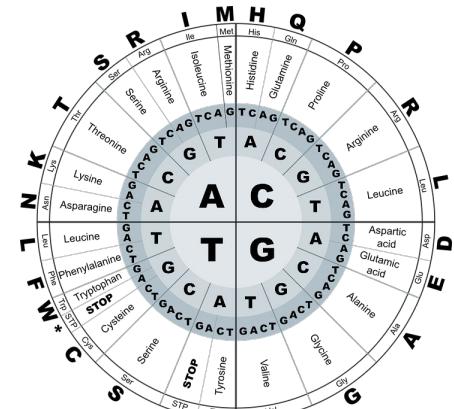




Proteins are the building blocks of cells, and the end products of protein-coding genes

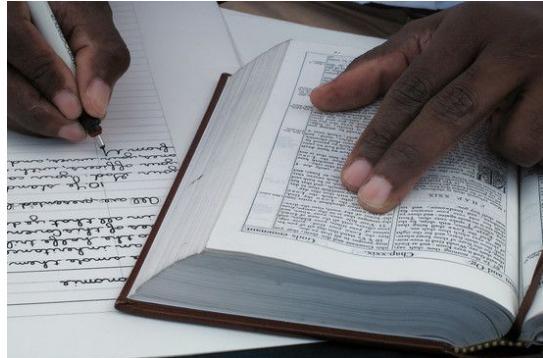


Proteins are the building blocks of cells, and the end products of protein-coding genes



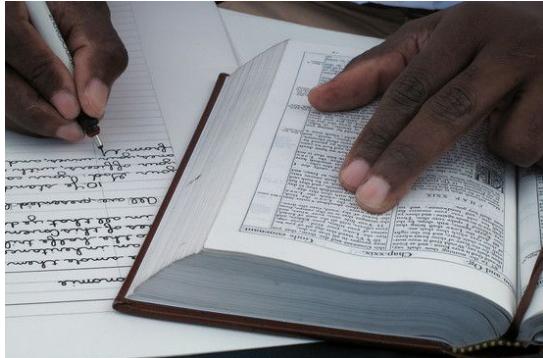
Evolution and phylogenetics

- We use **phylogenetics** to study the evolutionary history of life with molecular information
 - The set of instructions to build a cell (**and the whole organism!**) is stored in DNA
 - This set of instructions must be copied every time the cell replicates



Evolution and phylogenetics

- We use **phylogenetics** to study the evolutionary history of life with molecular information
 - The set of instructions to build a cell (**and the whole organism!**) is stored in DNA
 - This set of instructions must be copied every time the cell replicates



"Paul," she asked softly. "Paul, why are you holding your hands like that?"
He began to cry. It was guilt he cried from, and he hated that most of all: **in addition to everything else** that this monstrous woman had done to him, she had made him feel guilty as well. So he cried from guilt . . . but also from simple childish

original

```
ATTAAGGTTACCTCCAGGTAAACAAACAA  
CCAACTTCGATCTCTGTAGATCTGTTCTAAA  
CGAACTTAAATCTGTGGCTGTCACTCGGCTG  
CATGCTTAGTGCACTCACGCAGTATAATTAAAC  
TAATTACTGTCGTTGACAGGACACGAGTAACCTGT  
CTATCTCTGCAGGCTGCTTACGGTTTCGTCCTG  
TTGCAGCCGATCATCAGCACATCTAGGTTCTGCC  
GGGTGTGACCGAAAGGTAAGATGGAGAGCCTTGTC  
CCTGGTTCAACGAGAAAAACACACGTCCAACCTCAG  
TTGCCTGTTACAGGTTCGCGACGTGCTCGTAC
```

1. substitution

T → **A**

2. deletion

CTTACGGTT → **CT**

3. insertion

TC → **TTTC**

copy

```
ATTAAGGTTACCTCCAGGTAAACAAACAA  
CCAACATCTCTGTAGATCTGTTCTAAA  
CGAACTTAAATCTGTGGCTGTCACTCGGCTG  
CATGCTTAGTGCACTCACGCAGTATAATTAAAC  
TAATTACTGTCGTTGACAGGAGAGTAACCTGT  
CTATCTCTGCAGGCTGCTTTGT  
TTGCAGCCGATCATCAGCACATAGGTTCTGCC  
GGGTGTGACCGAAAGGTAAGATGGAGAGCCTTGTC  
CCTGGTTCAACGAGAAAAACACACGTACTCAG  
TTGCCTGTTACAGGTTTTCGTGCTCG  
TAC
```

Evolution and phylogenetics

- We use **phylogenetics** to study the evolutionary history of life with molecular information
 - The set of instructions to build a cell (**and the whole organism!**) is stored in DNA
 - This set of instructions must be copied every time the cell replicates
 - Imperfect copies cause **mutations**

Evolution and phylogenetics

Colour

Color

Evolution and phylogenetics

Colour

Silent mutation

Color

Evolution and phylogenetics

Colour



Silent mutation

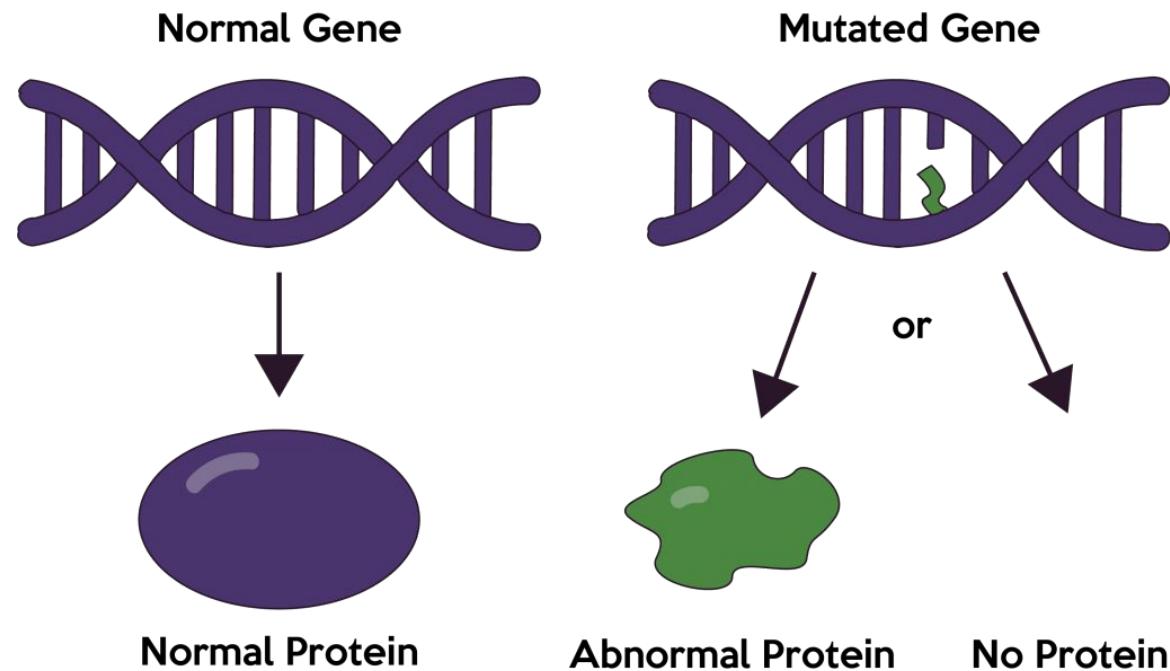
Color

Cat

Rat

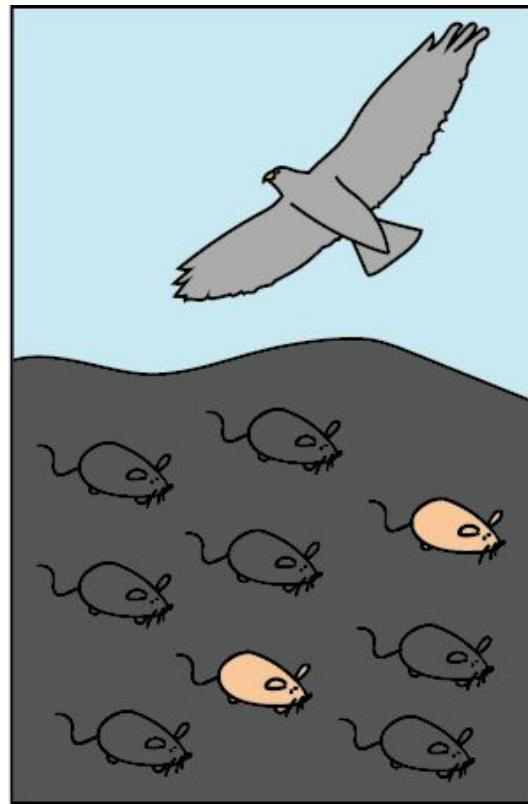


Evolution and phylogenetics



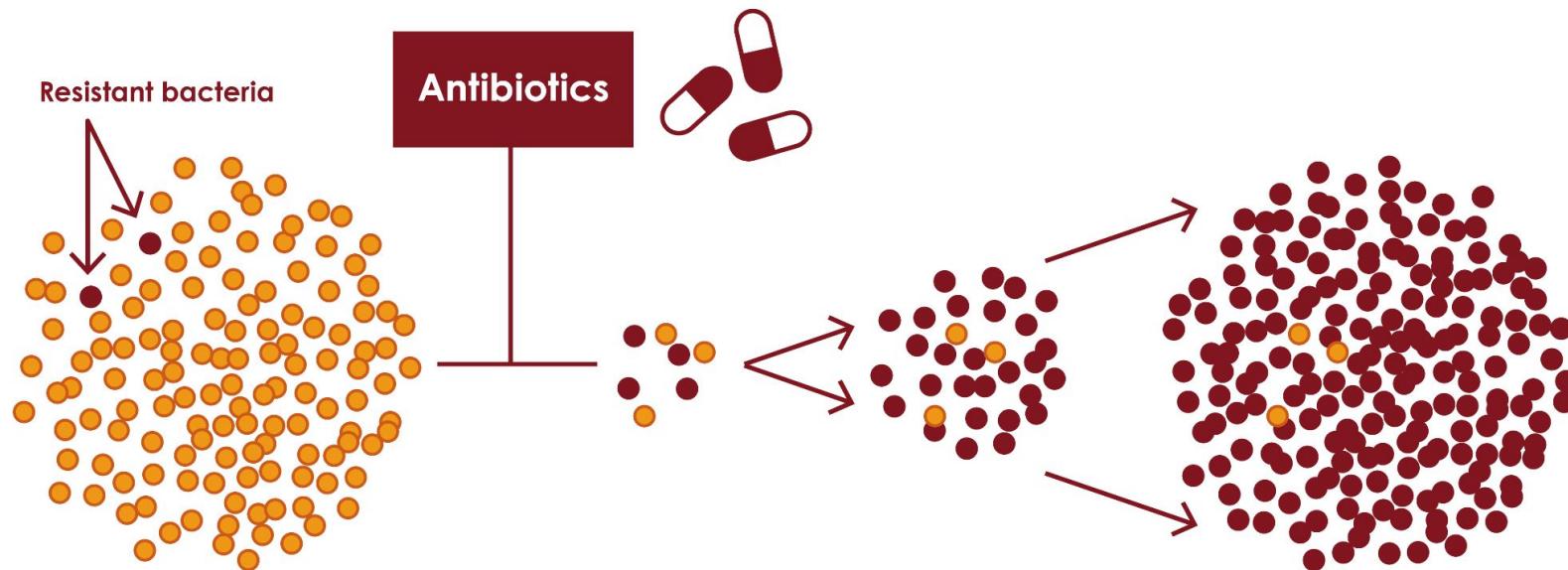
Evolution and phylogenetics

- We use **phylogenetics** to study the evolutionary history of life with molecular information
 - The set of instructions to build a cell (**and the whole organism!**) is stored in DNA
 - This set of instructions must be copied every time the cell replicates
 - Imperfect copies cause **mutations**
 - When mutations have an impact, it can be **positive** or **negative**



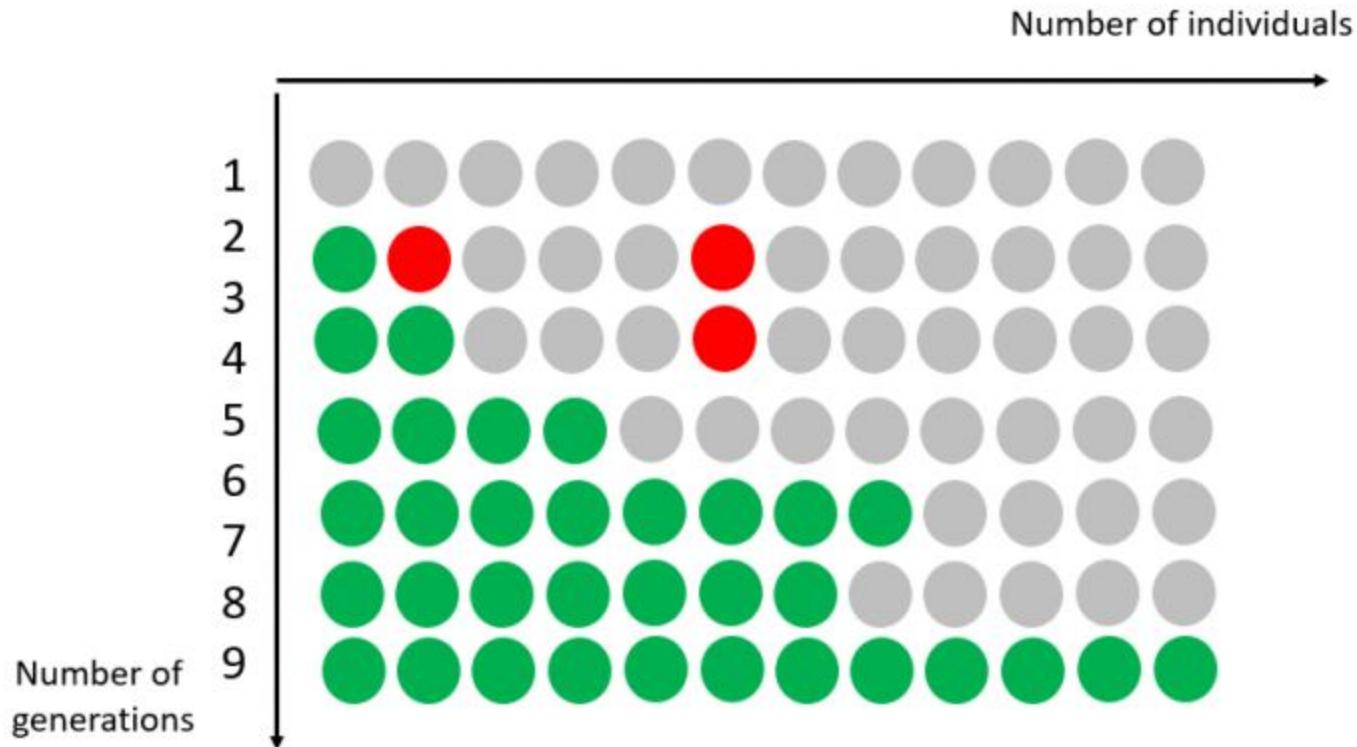
Evolution and phylogenetics

- We use **phylogenetics** to study the evolutionary history of life with molecular information
 - The set of instructions to build a cell (**and the whole organism!**) is stored in DNA
 - This set of instructions must be copied every time the cell replicates
 - Imperfect copies cause **mutations**
 - When mutations have an impact, it can be **positive** or **negative**
 - Positive mutations carry on in the population due to **natural selection**



Population of mainly
susceptible bacteria

Population of mainly
resistant bacteria



Evolution and phylogenetics

- We use **phylogenetics** to study the evolutionary history of life with molecular information
 - The set of instructions to build a cell (**and the whole organism!**) is stored in DNA
 - This set of instructions must be copied every time the cell replicates
 - Imperfect copies cause **mutations**
 - When mutations have an impact, it can be **positive** or **negative**
 - Positive mutations carry on in the population due to **natural selection**
 - Over time, populations that are too different become new species

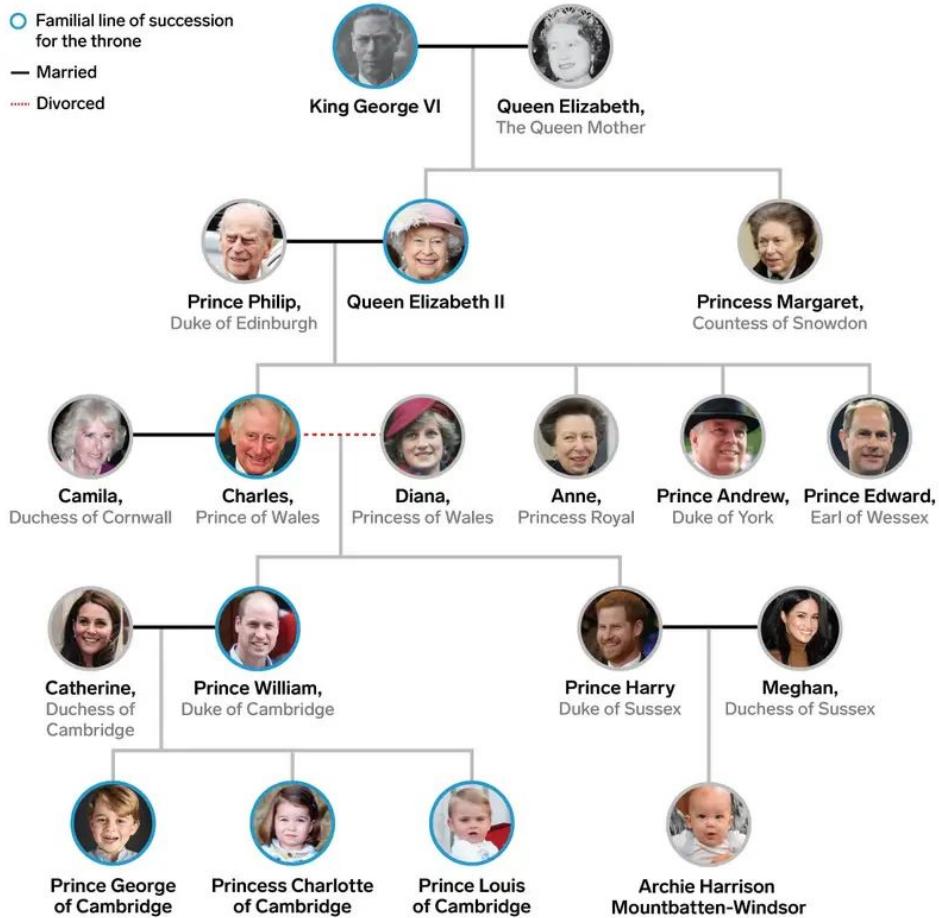
Evolution and phylogenetics

- We use **phylogenetics** to study the evolutionary history of life with molecular information
- The information of the relationships between species (or other entities) is represented through a **phylogenetic tree**

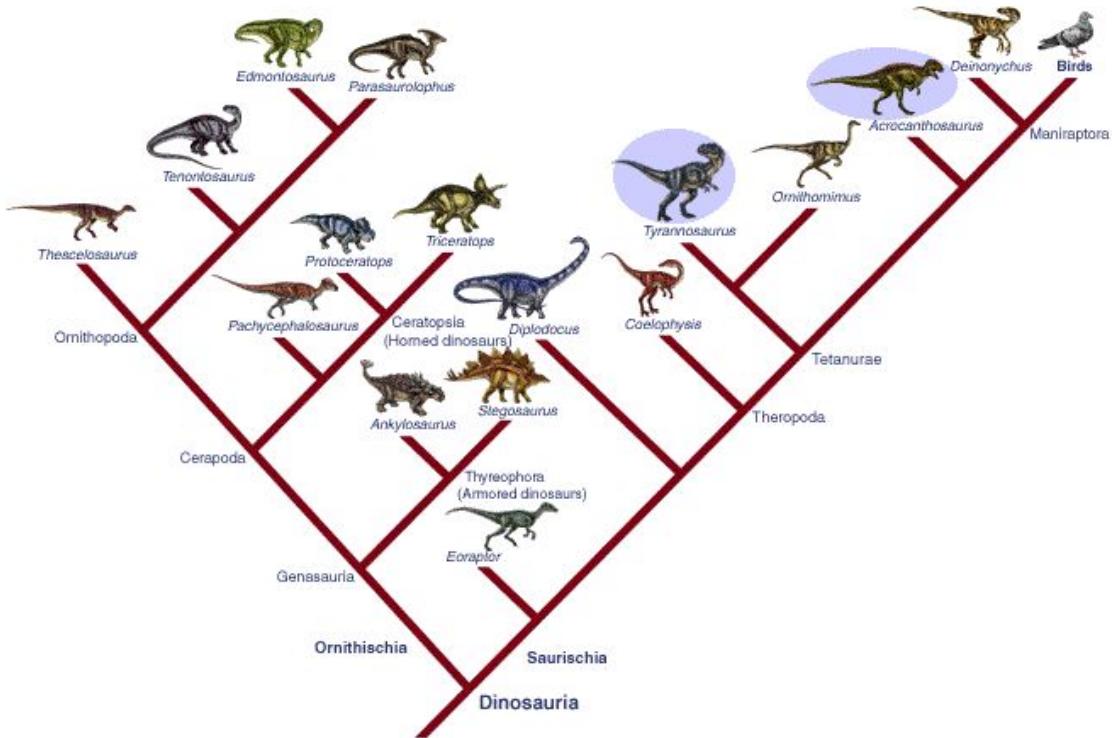
○ Familial line of succession
for the throne

— Married

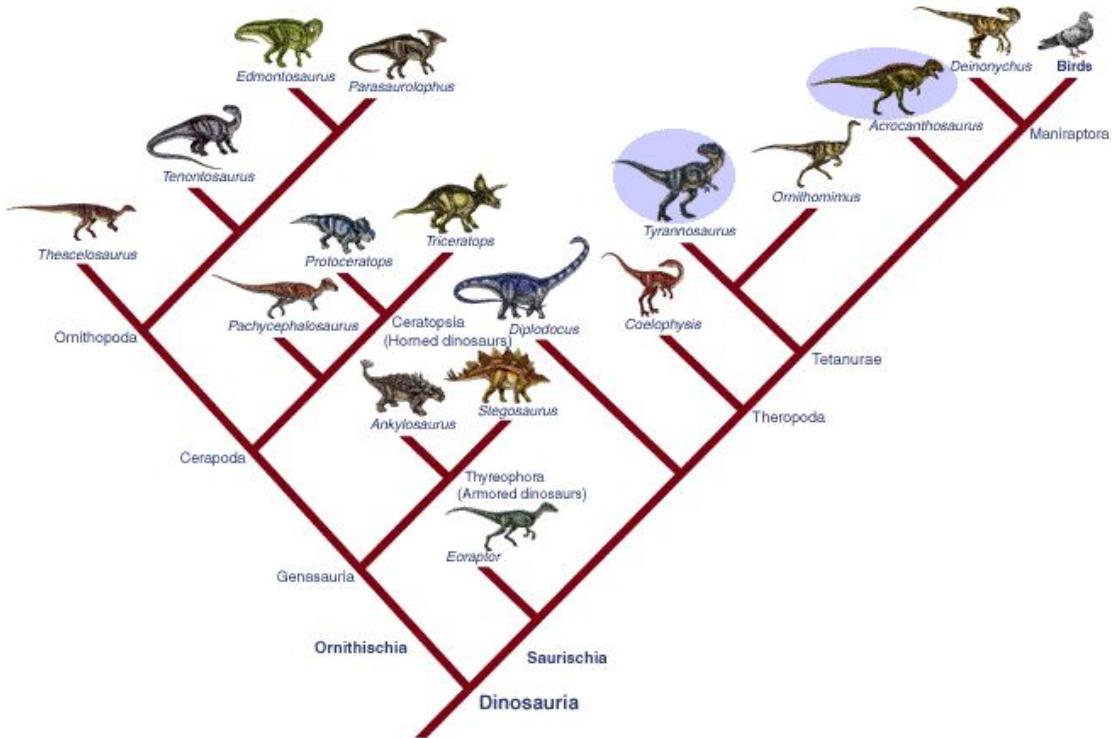
···· Divorced



Phylogenetic trees are family
trees for evolution

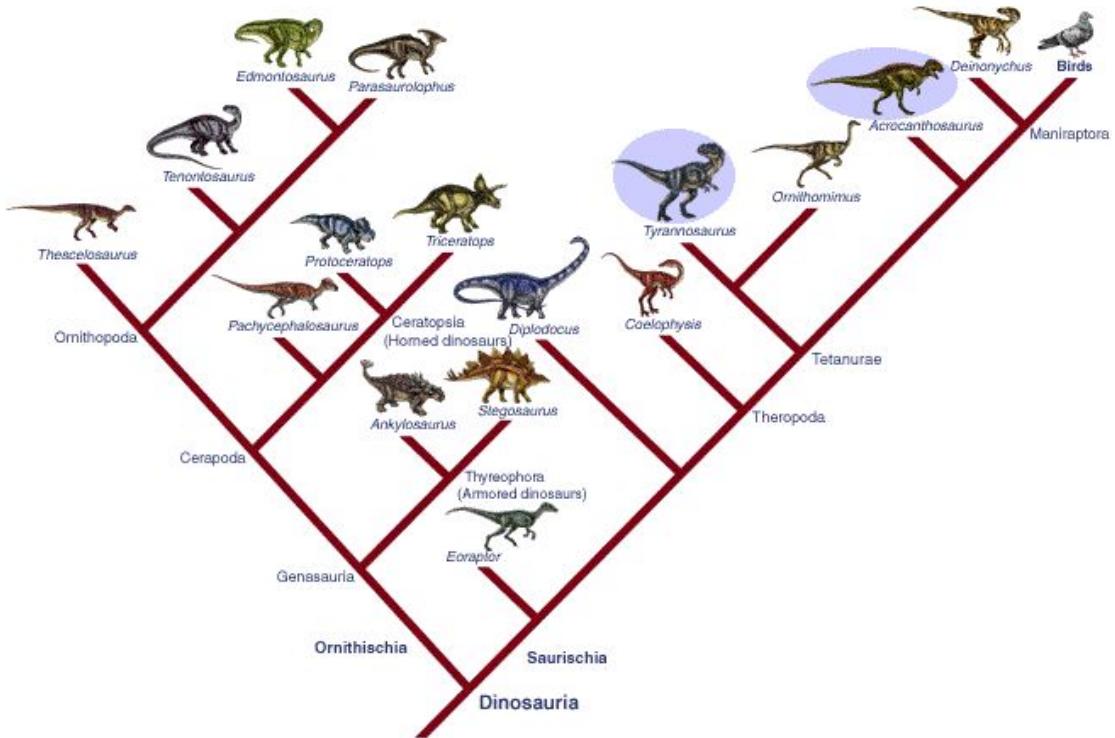


Phylogenetic trees are family trees for evolution

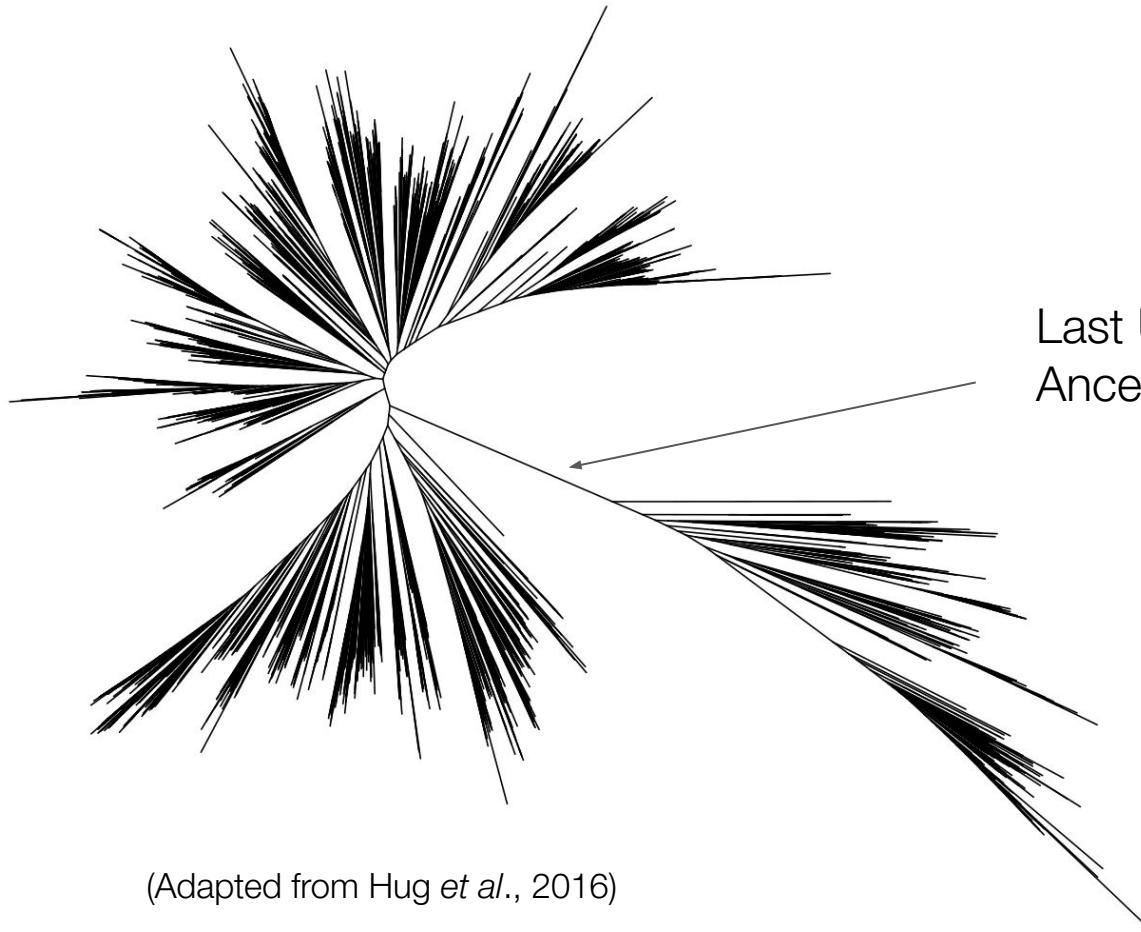


Phylogenetic trees are family trees for evolution



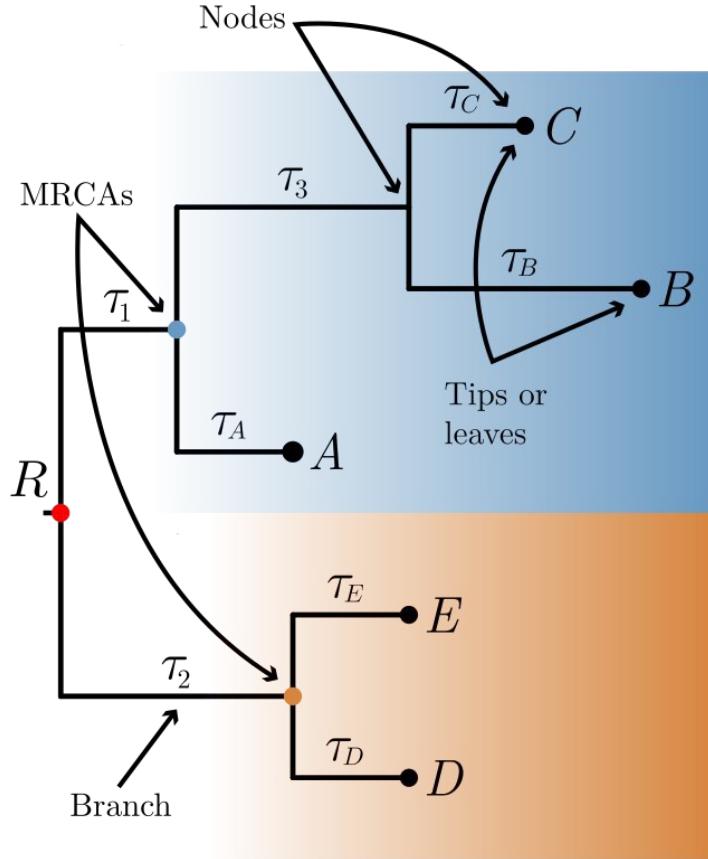


Phylogenetic trees are family trees for evolution

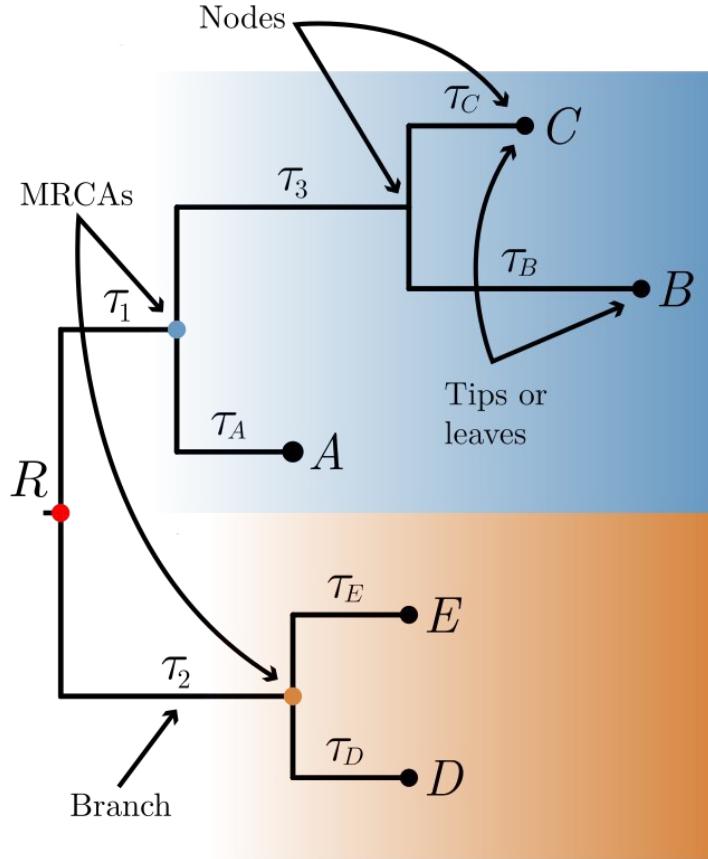


Last Universal Common
Ancestor

(Adapted from Hug *et al.*, 2016)

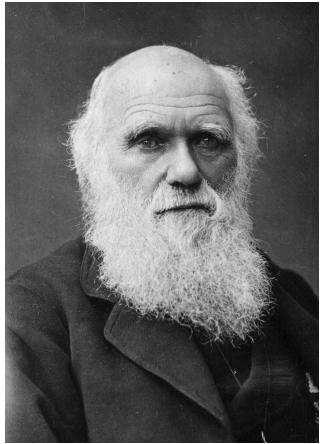


In phylogenetic trees, nodes represent the entities (species): ones that exist and ones that have existed in the past



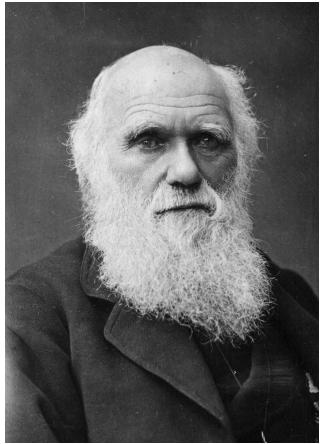
In phylogenetic trees, nodes represent the entities (species): ones that exist and ones that have existed in the past

The branches represent the relationship between them



- Darwin spent years studying the finches in the Galapagos to be able to understand how they evolved

Charles Darwin



- Darwin spent years studying the finches in the Galapagos to be able to understand how they evolved
- There are 5-10 million species on Earth

Charles Darwin

ACT II: THE LAB IN A COMPUTER



Computers allow us to work at bigger scales, faster

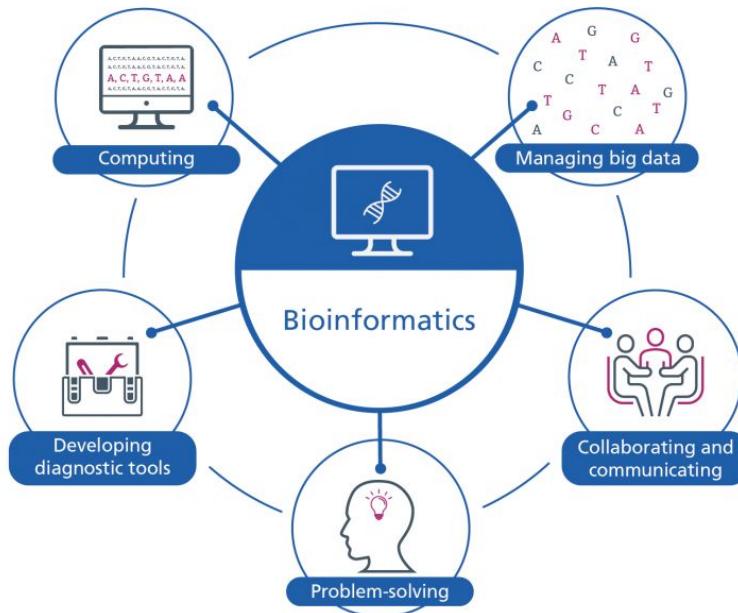
Bioinformatics

Bioinformatics

- Use of computers to work on biological problems

Bioinformatics

- Use of computers to work on biological problems



Bioinformatics

- Use of computers to work on biological problems
- Computers are better and faster than us at working through similarities and differences in molecular data

Bioinformatics

- Use of computers to work on biological problems
 - Computers are better and faster than us at working through similarities and differences in molecular data

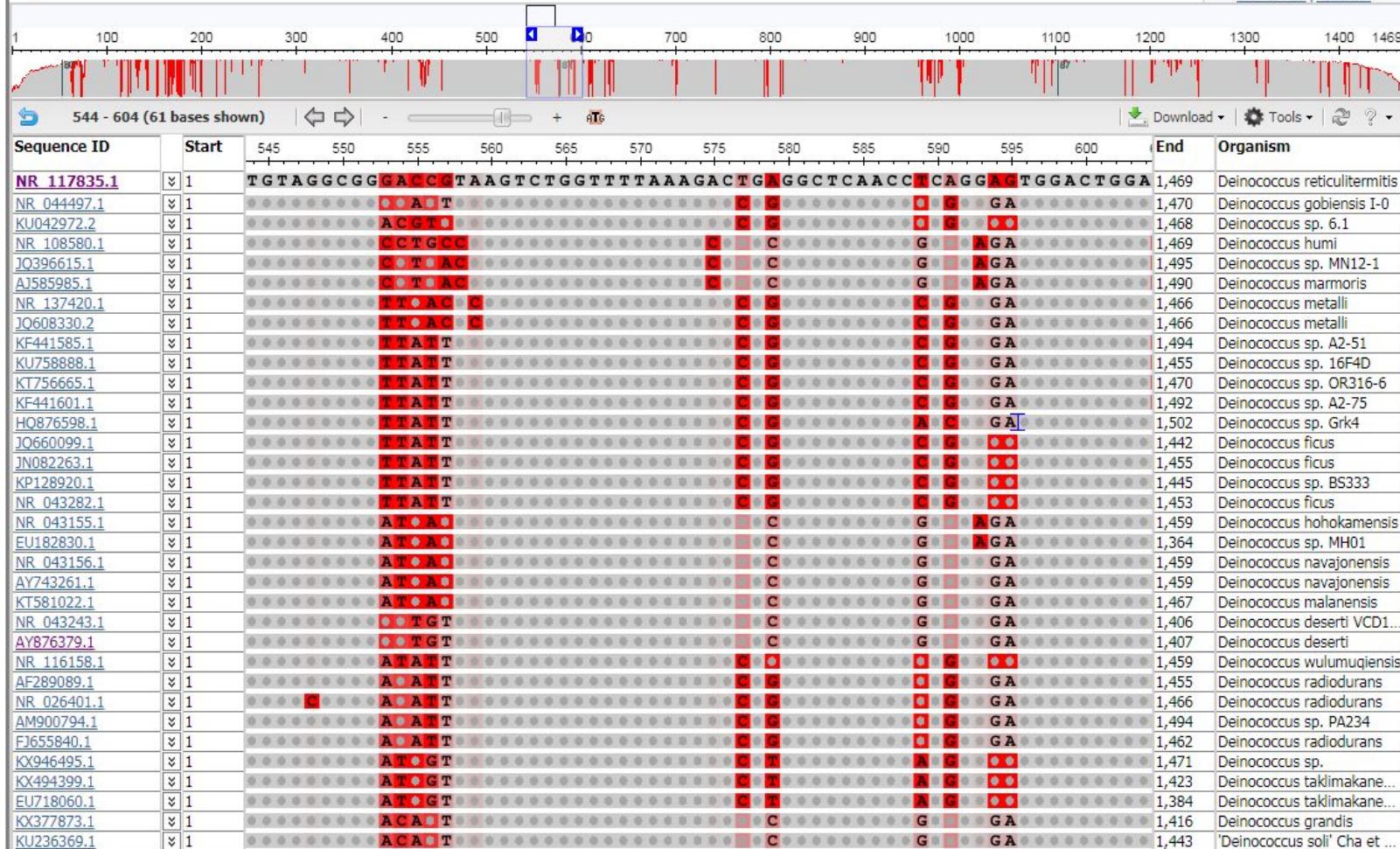
Bioinformatics

- Use of computers to work on biological problems
- Computers are better and faster than us at working through similarities and differences in molecular data

| | | | | | | | |
|---------------|-----|-----|-----|-----|-----|-----|-----|
| RNA_seq_1 | GUA | AAA | AUA | GGU | GAU | GAC | CUC |
| Protein_seq_1 | | | | | | | |
| RNA_seq_2 | GUA | AAA | AUA | GGU | GAC | GAC | CUC |
| Protein_seq_2 | | | | | | | |
| RNA_seq_3 | GUA | AAA | AUA | GCU | GAU | GAC | CUC |
| Protein_seq_3 | | | | | | | |

Is there any difference? Where?

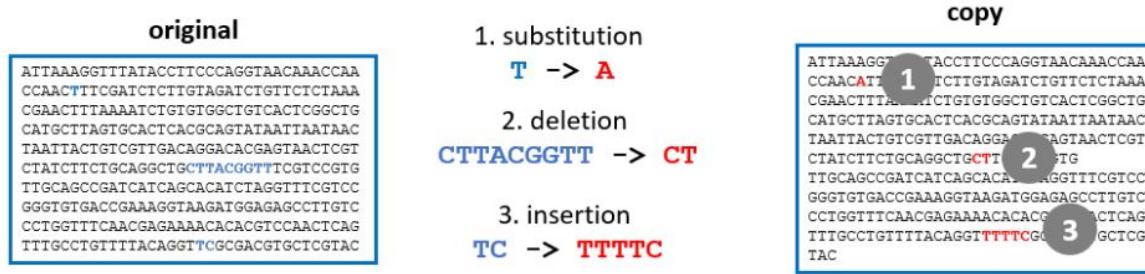
Alignment

[Link To View](#) | [Feedback](#)

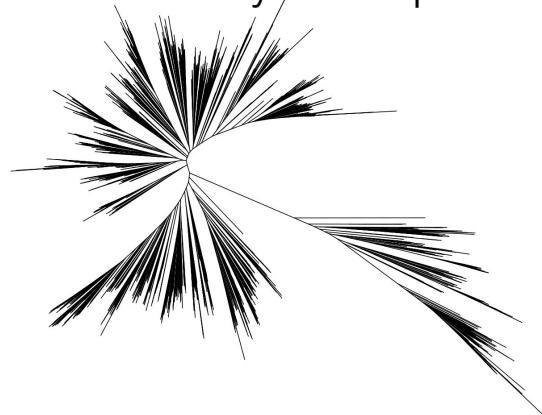
Bioinformatics

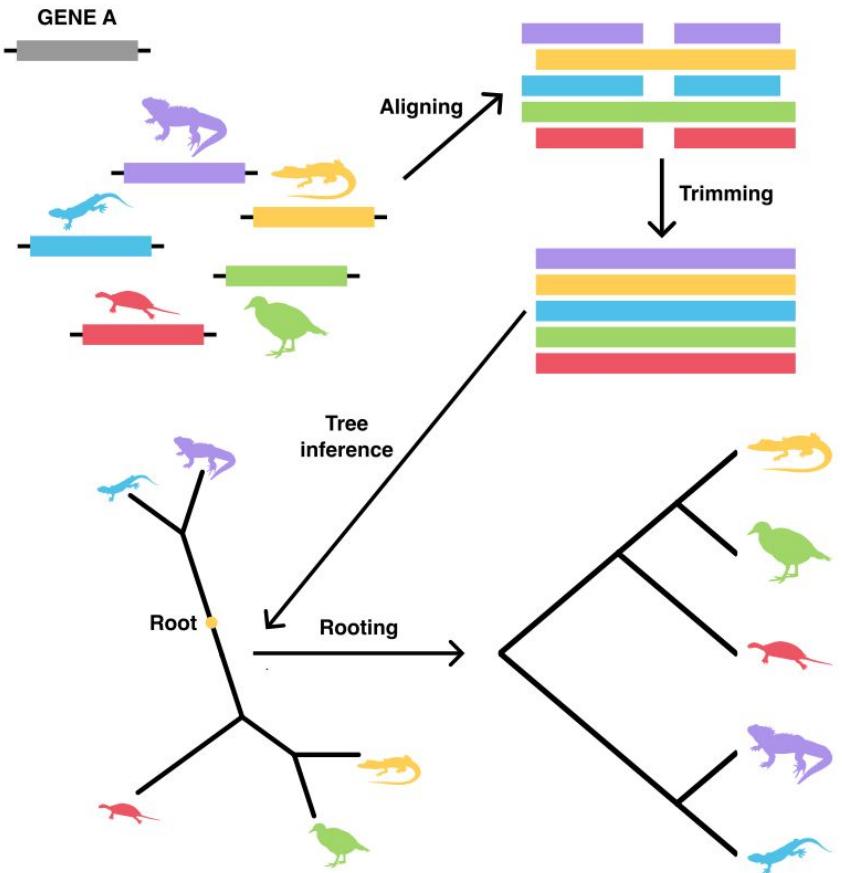
- Use of computers to work on biological problems
- Computers are better and faster than us at working through similarities and differences in molecular data
- They are also faster at turning these differences into phylogenetic trees

Bridge the molecular concept of **mutations**



With the evolutionary concept of **trees**





PH4H_Homo_sapiens

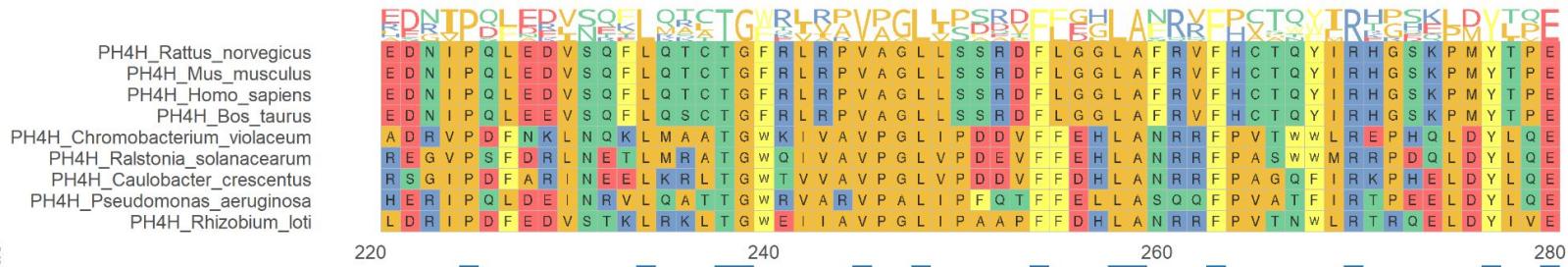
The sequence logo displays the following amino acid distribution across the sequence:

- Position 1: D (red)
- Position 2: N (light green)
- Position 3: P (light orange)
- Position 4: Q (light green)
- Position 5: L (light orange)
- Position 6: E (blue)
- Position 7: D (red)
- Position 8: V (light orange)
- Position 9: S (light green)
- Position 10: Q (light green)
- Position 11: F (yellow)
- Position 12: L (light orange)
- Position 13: Q (light green)
- Position 14: T (light green)
- Position 15: C (orange)
- Position 16: G (light green)
- Position 17: F (yellow)
- Position 18: R (dark blue)
- Position 19: L (light orange)
- Position 20: R (dark blue)
- Position 21: P (light orange)
- Position 22: V (light orange)
- Position 23: A (green)
- Position 24: G (light green)
- Position 25: L (light orange)
- Position 26: S (light green)
- Position 27: S (light green)
- Position 28: R (dark blue)
- Position 29: D (red)
- Position 30: F (yellow)
- Position 31: L (light orange)
- Position 32: G (light green)
- Position 33: G (light green)
- Position 34: L (light orange)
- Position 35: A (green)

PH4H_*Rattus_norvegicus*
PH4H_*Mus_musculus*
PH4H_*Homo_sapiens*
PH4H_*Bos_taurus*
PH4H_*Chromobacterium_violaceum*
PH4H_*Ralstonia_solanacearum*
PH4H_*Caulobacter_crescentus*
PH4H_*Pseudomonas_aeruginosa*
PH4H_*Rhizobium_loti*

Sequence logo showing the conservation of amino acids at positions 220, 240, 260, and 280 across various species. The x-axis marks positions 220, 240, 260, and 280. The y-axis lists species: PH4H_Rattus_norvegicus, PH4H_Mus_musculus, PH4H_Homo_sapiens, PH4H_Bos_taurus, PH4H_ChromobacteriumViolaceum, PH4H_Ralstonia_solanacearum, PH4H_Caulobacter_crescentus, PH4H_Pseudomonas_aeruginosa, and PH4H_Rhizobium_loti.

Multiple Sequence Alignment



Bioinformatics

- Use of computers to work on biological problems
- Computers are better and faster than us at working through similarities and differences in molecular data
- They are also faster at turning these differences into phylogenetic trees
- There are multiple repositories of molecular data that are public



Biological Databases



TACACAATCAGTTAGTTCCACCGACAGTCCGCAGAAACCATTGACGGC
GTGGCAATCCGTAAGATAGCCAAATATTATTATTGTCAGATACTCACT
AGCAAGACAAACTGCAGATCAGATGAGTGTCAAATCAGTGAAATTC
TAAACTTCAACGAAATTCAGGAAATGAAATGAAATGAAATGAAATGAA
ATCGAACGTAAAGAUATGCGATTCATAGATGAAATGAAATGAAATGAA
AATAATAAAAAACAAACAAACAGTGCACAAACAGCCGGGGCATCTTCATAGAT
AACTTCTGCCTGCACTTGGTATATGTACTTACACATAGACATATATATA

FlyBase



Log in

sation

mation (CDC) | Research information (NIH) | SARS-CoV-2 data (NCBI) | Prevention and treatment information (HHS) | Español

BLAST®

Home Recent Results Saved Strategies Help

Basic Local Alignment Search Tool

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance. [Learn more](#)

NEWS

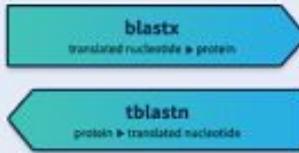
ElasticBLAST is here!

ElasticBLAST is a new cloud based tool to run your BLAST searches faster and make you more effective.

Mon, 07 Feb 2022 12:00:00 EST

[More BLAST news...](#)

Web BLAST



BLAST Genomes

Enter organism common name, scientific name, or tax id:

search

Human Mouse Rat Microbes

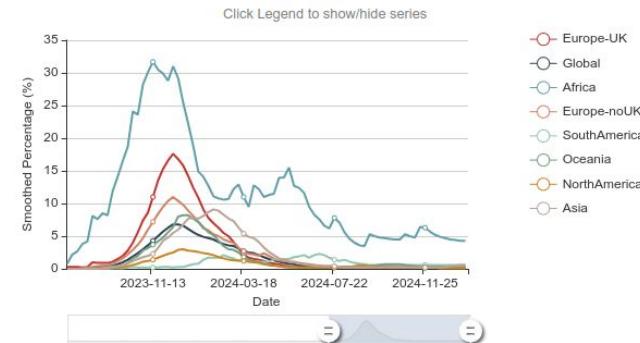
Tracking of hCoV-19 Variants

VOI GRA (BA.2.86+BA.2.86.*) first detected in Denmark/Israel/USA

As of 24 January 2025 - 0751UTC, 110 countries shared 25,360 GRA (BA.2.86+BA.2.86.*) genome sequences with unprecedented speed from sample collection to making these data publicly accessible via GISAID EpiCoV, in some cases within less than 24 hours.



Relative Variant Genome Frequency per Region (exponentially smoothed alpha=0.3)



Most recent submission per country

| Country | Virus Name | Submitted |
|---------------|-----------------------------------|------------|
| 110 countries | | |
| Kenya | hCoV-19/Kenya/2023084003/2024 | 2025-01-23 |
| Argentina | hCoV-19/Argentina/INF1127032/2024 | 2025-01-21 |

Country Submission Count

| Country | Total #GRA (BA.2.86+BA.2.86.*) | #GRA (BA.2.86+BA.2.86.*) in past 4 weeks | %GRA (BA.2.86+BA.2.86.*) in past 4 weeks |
|----------------|---------------------------------|---|---|
| United Kingdom | 3,327 | 0 | 0.0% |
| USA | 3,120 | 1 | 0.0% |

Bioinformatics

- Use of computers to work on biological problems
- Computers are better and faster than us at working through similarities and differences in molecular data
- They are also faster at turning these differences into phylogenetic trees
- There are multiple repositories of molecular data that are public
- Some bioinformatic tools can be used through the web, but some use the **command line**

GUI



CLI

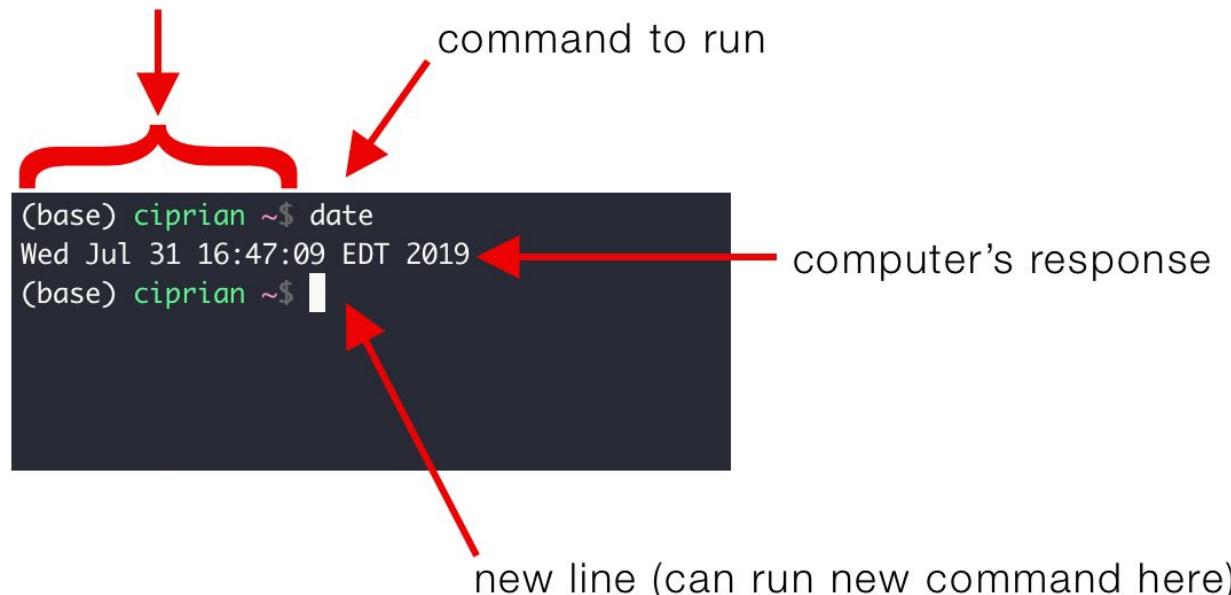


Bioinformatics

- Use of computers to work on biological problems
- Computers are better and faster than us at working through similarities and differences in molecular data
- They are also faster at turning these differences into phylogenetic trees
- There are multiple repositories of molecular data that are public
- Some bioinformatic tools can be used through the web, but some use the **command line**
 - Usually faster
 - Automatizable
 - Handles big data

command line prompt

command to run



BASH
THE BOURNE-AGAIN SHELL

ACT III: OUR STUDY

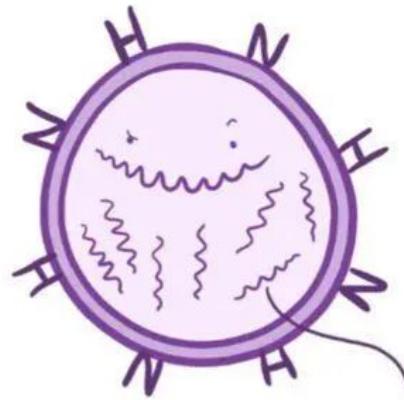
A bit about influenza

- **Influenza** (flu) is caused by 4 viral genera within Orthomyxoviridae

A bit about influenza

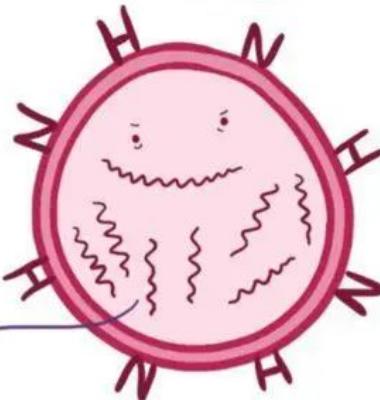
- **Influenza** (flu) is caused by 4 viral genera within Orthomyxoviridae
 - Only 3 of them are known to infect humans

TYPE A



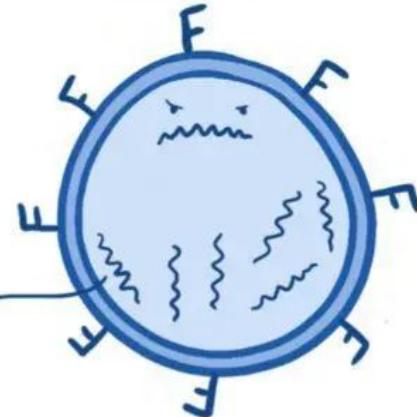
8 RNA SEGMENTS

TYPE B



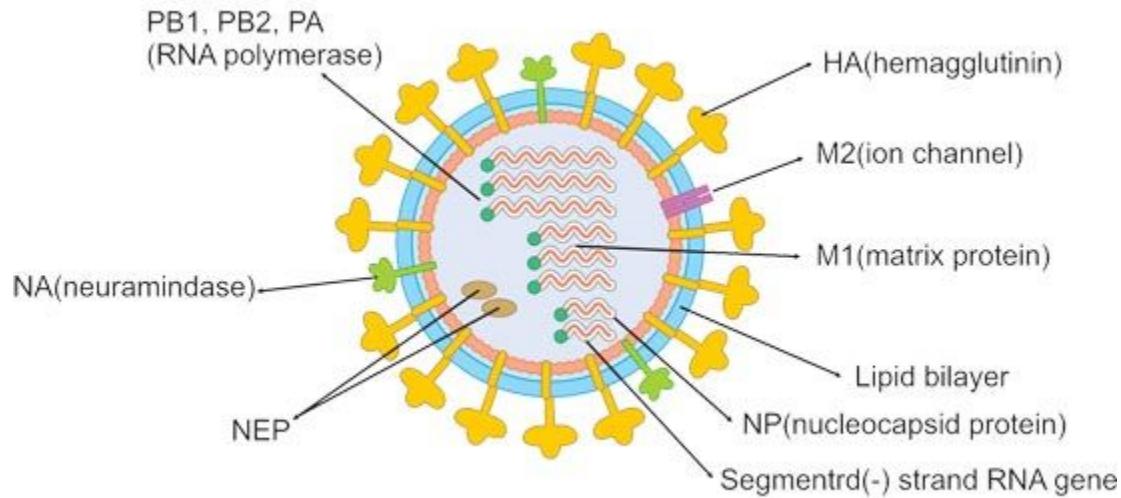
7 RNA SEGMENTS

TYPE C



A bit about influenza

- **Influenza** (flu) is caused by 4 viral genera within Orthomyxoviridae
 - Only 3 of them are known to infect humans
- Within **Influenza A** viruses, we can identify subtypes according to their surface proteins



A bit about influenza

- **Influenza** (flu) is caused by 4 viral genera within Orthomyxoviridae
 - Only 3 of them are known to infect humans
- Within **Influenza A** viruses, we can identify subtypes according to their surface proteins
- Viruses like influenza virus mutate **very quickly**

A bit about influenza

- **Influenza** (flu) is caused by 4 viral genera within Orthomyxoviridae
 - Only 3 of them are known to infect humans
- Within **Influenza A** viruses, we can identify subtypes according to their surface proteins
- Viruses like influenza virus mutate **very quickly**
 - This allows us to use phylogenetic methods to track its evolution in real time

A bit about what we will do

- Understand the importance of genome databases

A bit about what we will do

- Understand the importance of genome databases
- Understand the value of bioinformatics and comparative genomics

A bit about what we will do

- Understand the importance of genome databases
- Understand the value of bioinformatics and comparative genomics
- Understand the use of phylogenomics for biomedicine

A bit about what we will do

- Understand the importance of genome databases
- Understand the value of bioinformatics and comparative genomics
- Understand the use of phylogenomics for biomedicine
- Learn how to access and process biological sequence data from publicly available repositories

A bit about what we will do

- Understand the importance of genome databases
- Understand the value of bioinformatics and comparative genomics
- Understand the use of phylogenomics for biomedicine
- Learn how to access and process biological sequence data from publicly available repositories
- Learn the basics of programming in bash

A bit about the workshop

A bit about the workshop

- We will use **real influenza virus sequences** to track the spread of Influenza H1N1 outbreak in USA

A bit about the workshop

- We will use **real influenza virus sequences** to track the spread of Influenza H1N1 outbreak in USA
- We will reconstruct phylogenetic trees to understand the origin and spread of the outbreak

A bit about the workshop - what you will need

- A laptop
- A Google Account (to use Google Colab)

A bit about the workshop - what you will need

- A laptop
- A Google Account (to use Google Colab)
- Energy to learn!

Questions?

Evolutionary analyses in bioinformatics

The light of evolution and the lab in a computer

Saioa Manzano-Morales

Toni Gabaldón

Comparative Genomics

Institute for Research in Biomedicine