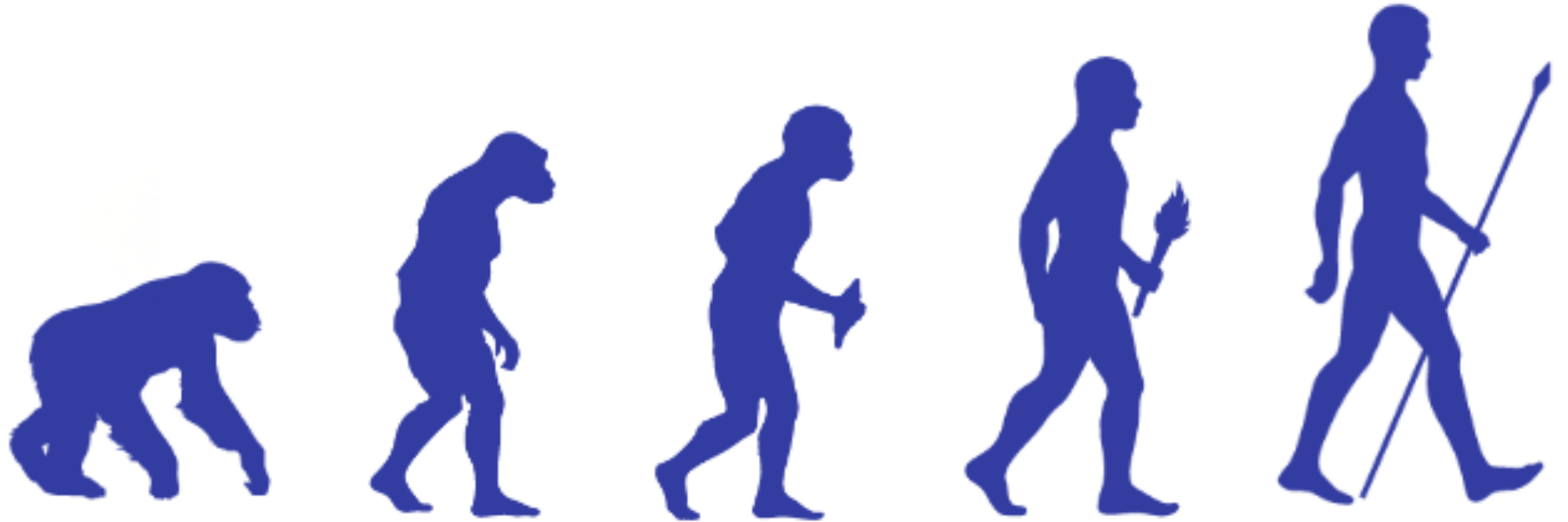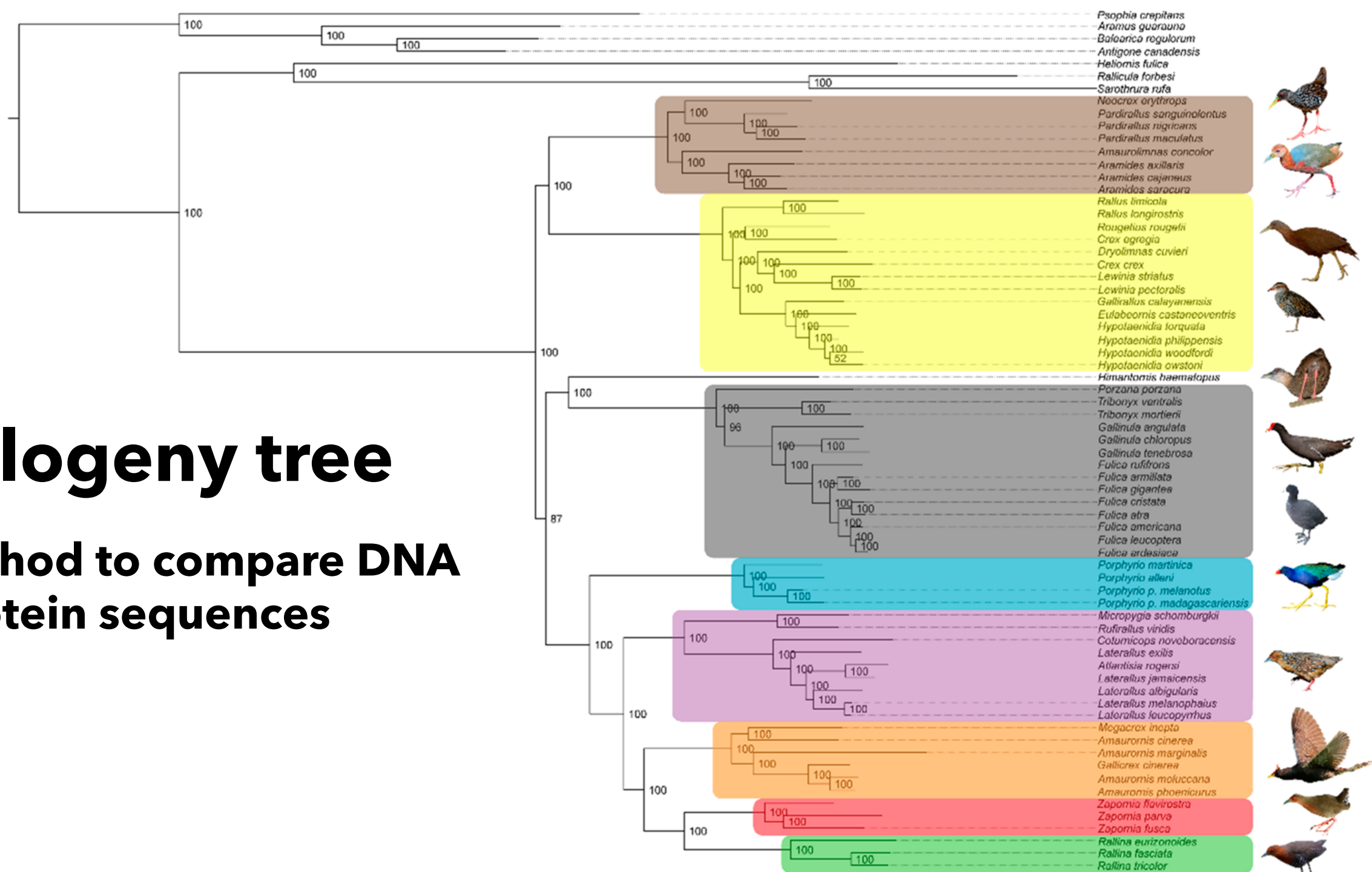# Phylogenomics

Haeun Lee and Amber Diederichs

What is phylogenomics?

Using genomic data to study evolution

Phylogeny tree

A method to compare DNA or protein sequences

Psophia crepitans
Aramus guarauna
Balearica regulorum
Antigone canadensis
Heliornis fulica
Rallicula forbesi
Sarothrura rufa
Neocrex erythrops
Pardirallus sanguinolentus
Pardirallus nigricans
Pardirallus maculatus
Amaurolimnas concolor
Aramides axillaris
Aramides cajaneus
Aramides saracura
Rallus limicola
Rallus longirostris
Rougetius rougetii
Crex egregia
Dryolimnas cuvieri
Crex crex
Lewinia striatus
Lewinia pectoralis
Gallirallus calayanensis
Eulabeornis castaneoventris
Hypotaenidia torquata
Hypotaenidia philippensis
Hypotaenidia woodfordi
Hypotaenidia owstoni
Himantornis haematopus
Porzana porzana
Tribonyx ventralis
Tribonyx mortierii
Gallinula angulata
Gallinula chloropus
Gallinula tenebrosa
Fulica rufifrons
Fulica armillata
Fulica gigantea
Fulica cristata
Fulica atra
Fulica americana
Fulica leucoptera
Fulica ardesiaca
Porphyrio martinica
Porphyrio alleni
Porphyrio p. melanotus
Porphyrio p. madagascariensis
Micropygia schomburgkii
Rufirallus viridis
Coturnicops noveboracensis
Laterallus exilis
Atlantisia rogersi
Laterallus jamaicensis
Laterallus albigularis
Laterallus melanophaius
Laterallus leucopyrrhus
Megacrex inepta
Amaurornis cinerea
Amaurornis marginalis
Gallicrex cinerea
Amaurornis moluccana
Amaurornis phoenicurus
Zapornia flavirostra
Zapornia parva
Zapornia fusca
Rallina eurizonoides
Rallina fasciata
Rallina tricolor

Why is phylogenomics important?
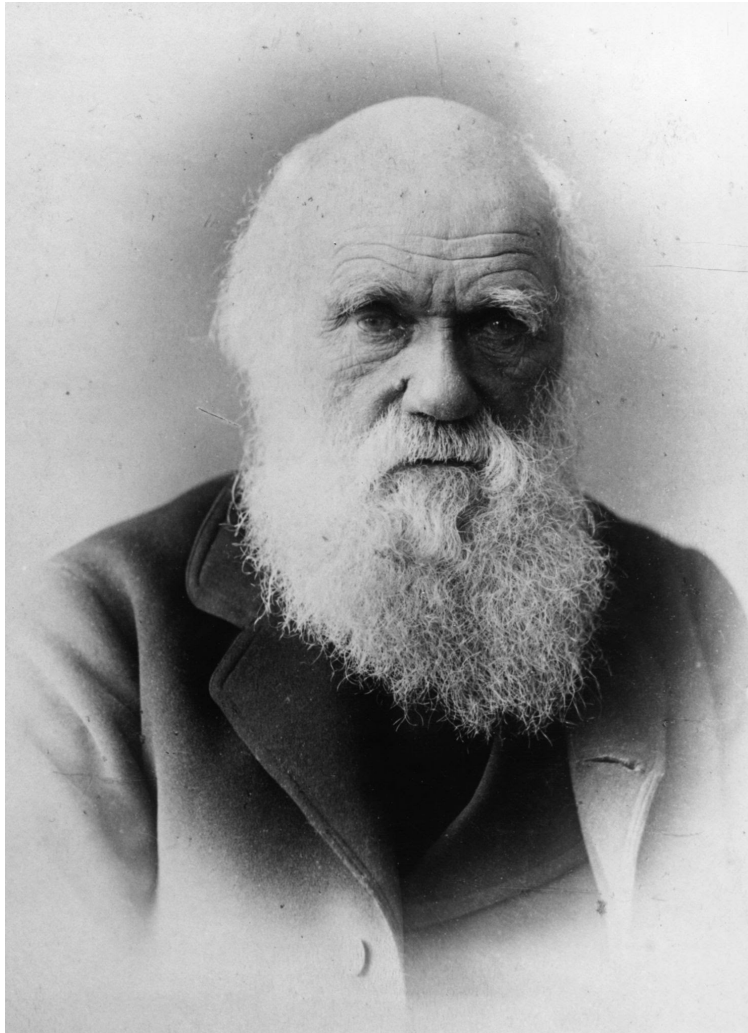

Emergence of new genes

Morphology

Molecular adaptations
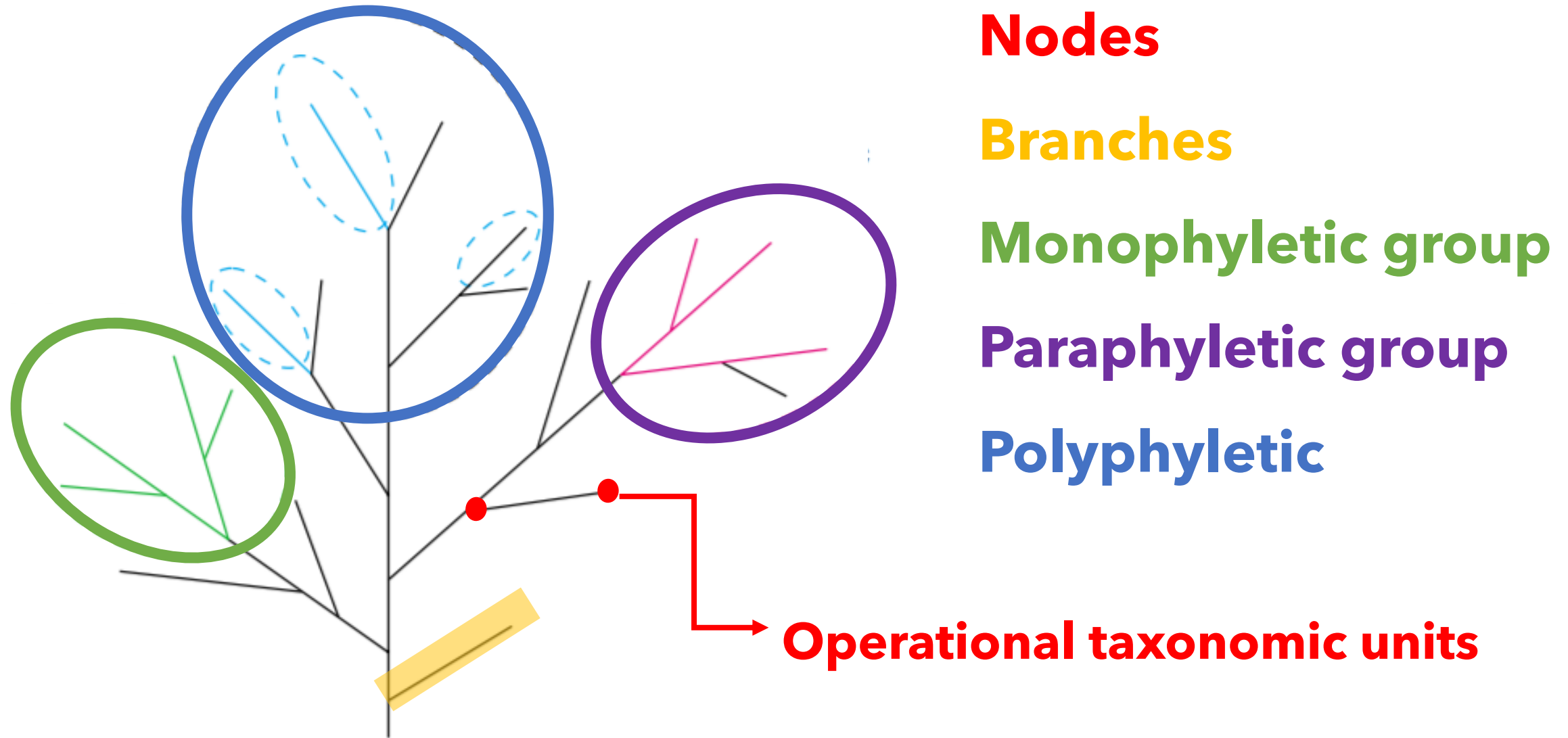
Understanding major transitions in evolution

What is the history of phylogenomics?
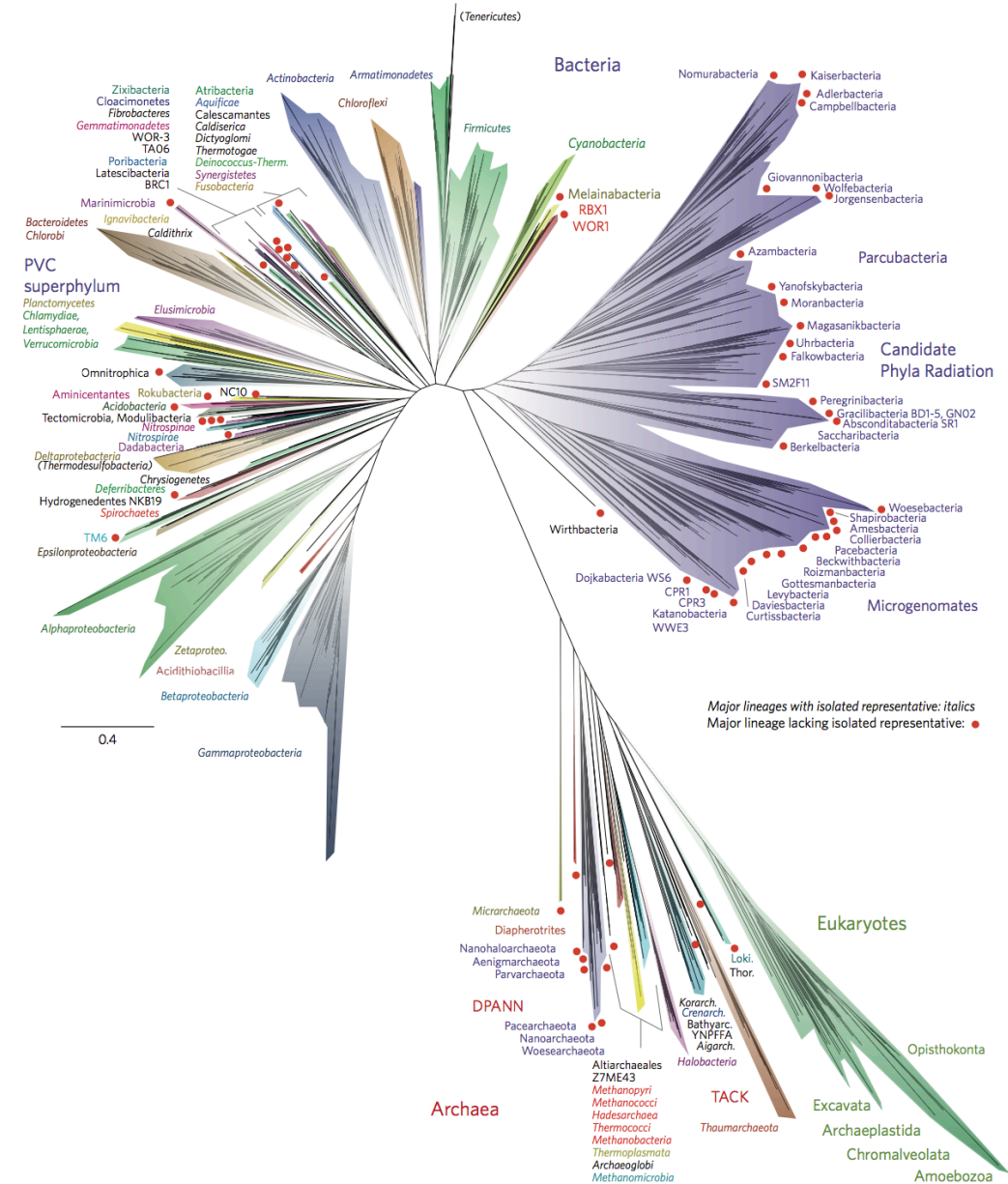
How do you read phylogeny trees?

Nodes
Branches
Monophyletic group
Paraphyletic group
Polyphyletic

Operational taxonomic units

How do you read a phylogeny tree?

Homologs

Paralogs

Orthologs | Orthologs

Frog α   Chick α   Mouse α   Mouse β   Chick β   Frog β

α –Chain gene   β –Chain gene

Carly globin gene
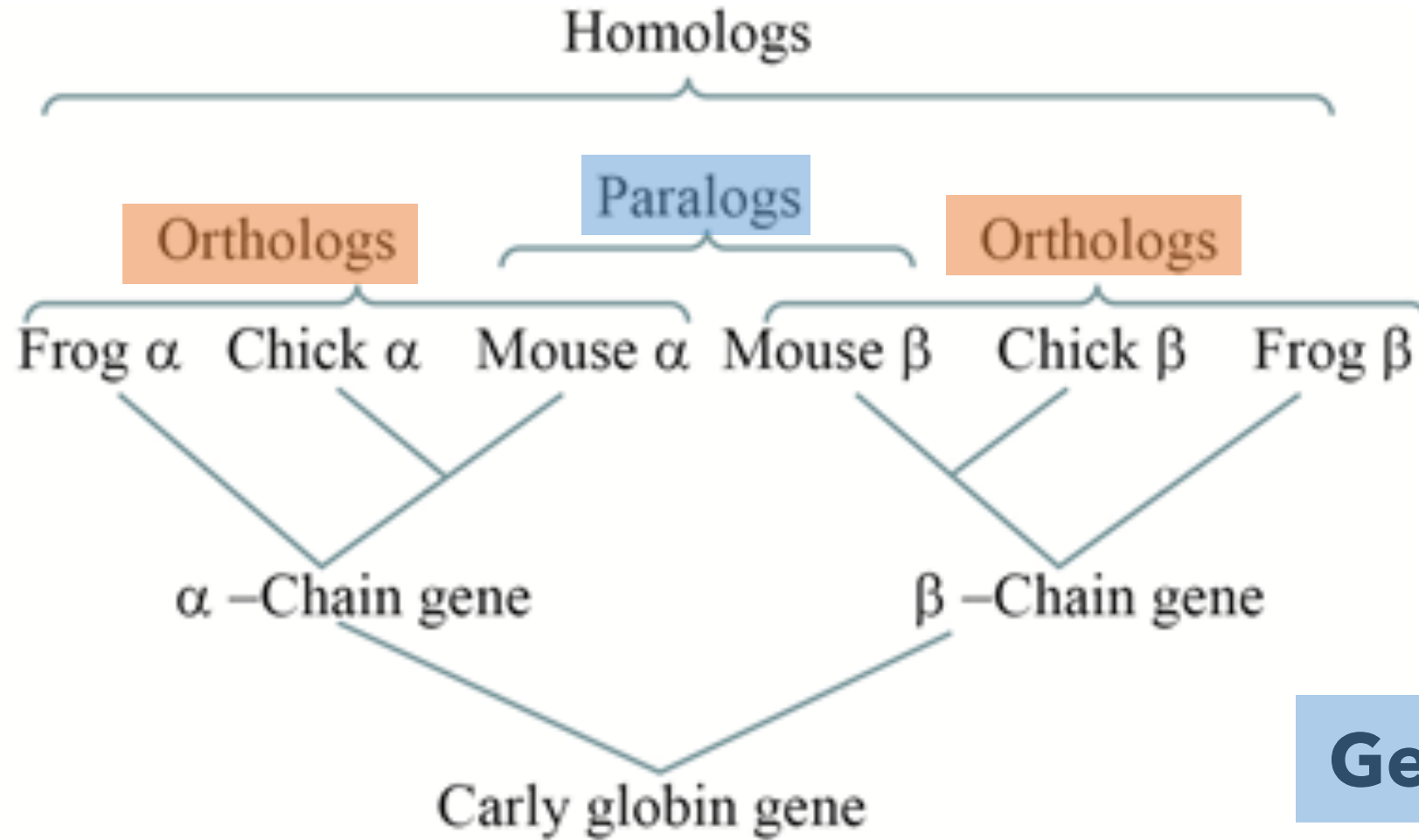
Gene duplication event

Parent to offspring duplication

How do you create a phylogeny tree?

How do you assemble data?
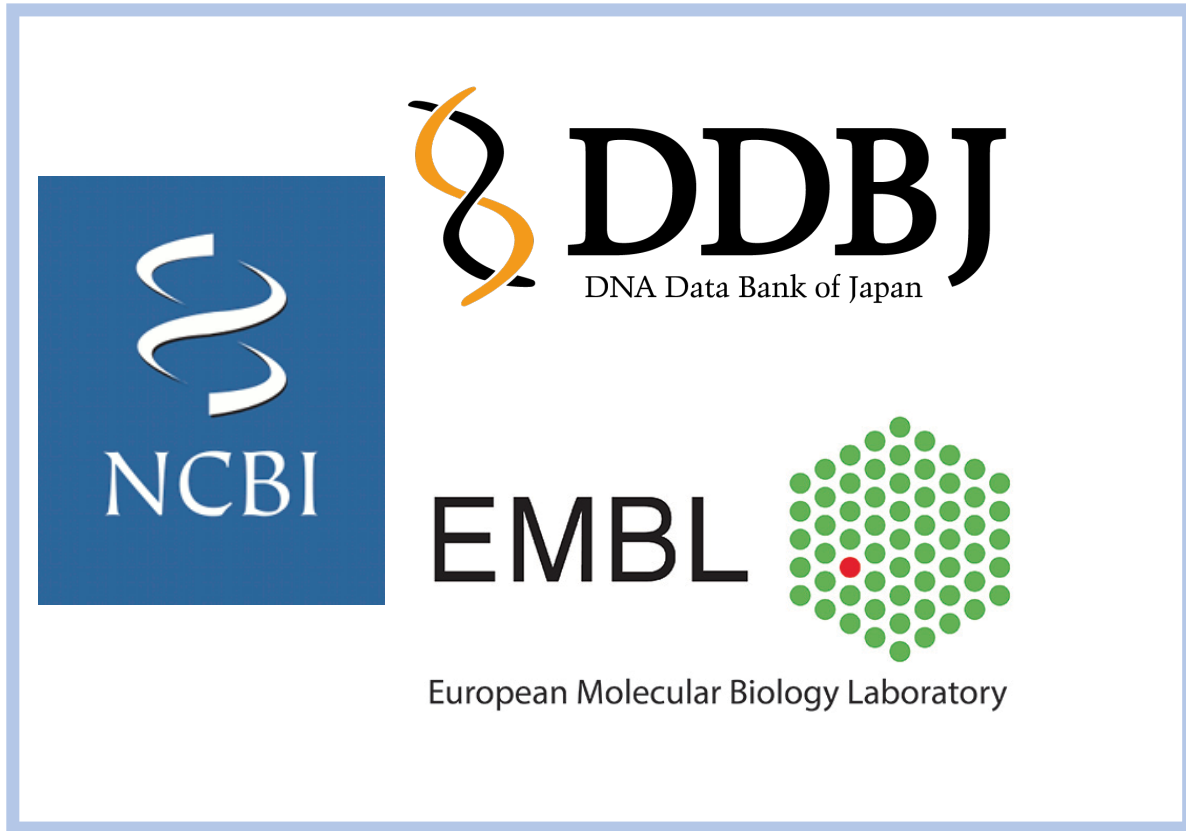
Data is obtained through searching public domains

How do you do multiple sequence alignment?
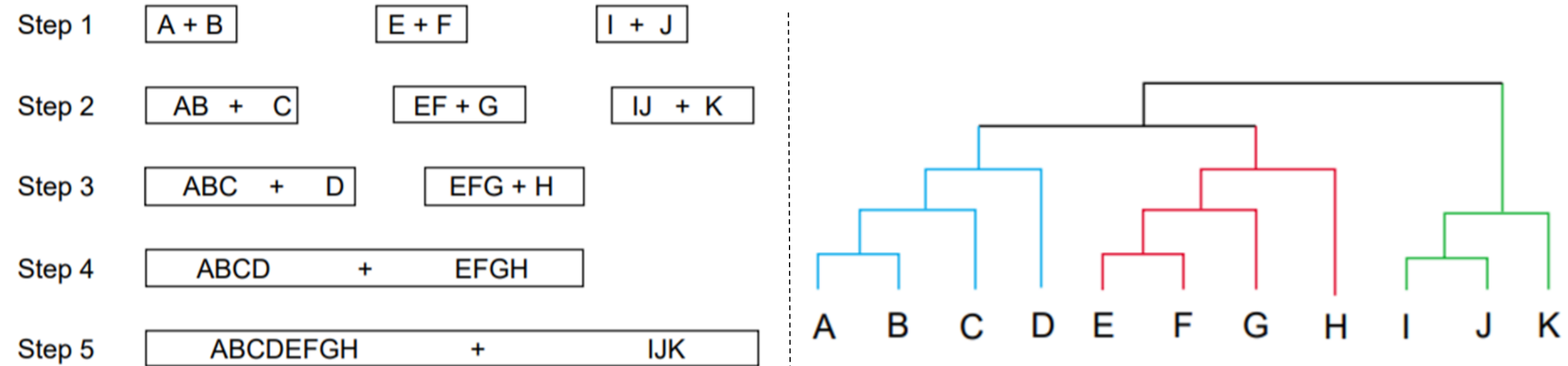
Progressive sequence alignment

How do you do multiple sequence alignment?

BioEdit Sequence Alignment Editor

How do you build the tree?

Sequence alignment

X: AAT TTG TCC
Y: AAT TTG TAA
Z: AAT TAA GGG

Pairwise genetic distance matrix

	X	Y	Z
X	–		
Y	2	–	
Z	4	4	–

Distance tree

2 — Z
1 — X
1 — Y

Distance matrix method
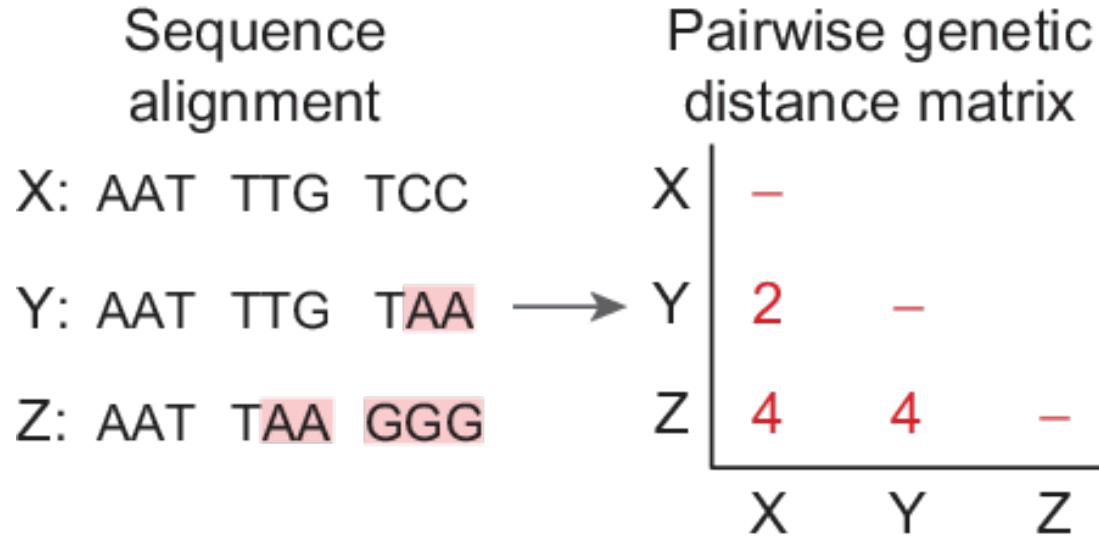
characters

taxon	shape	color
A	□	yellow
B	○	yellow
C	□	blue
D	○	yellow

B D A C — 2 changes
B C D A — 3 changes
A B C D — 3 changes

Discrete data method

Dataset

```
        0123456789
seqA    ACCGTTCGGT
seqB    ATGGTTCAGA
seqC    ATCGATCGGA
```

Replicate 1

```
        1562314951
seqA    CTCCGCTTTC
seqB    TTCGGTTATT
seqC    TTCCGTAATT
```

Replicate 2

```
        5234924418
seqA    TCGTTCTTCG
seqB    TGGTAGTTTG
seqC    TCGAACAATG
```

Replicate 3

```
        5607718907
seqA    TCAGGCGTAG
seqB    TCAAATGAAA
seqC    TCAGGTGAAG
```

Tree 1
seqA
seqB
seqC

Tree 2
seqA
seqC
seqB

Tree 3
seqA
seqB
seqC

67%
seqA
seqB
seqC

Bootstrap consensus tree

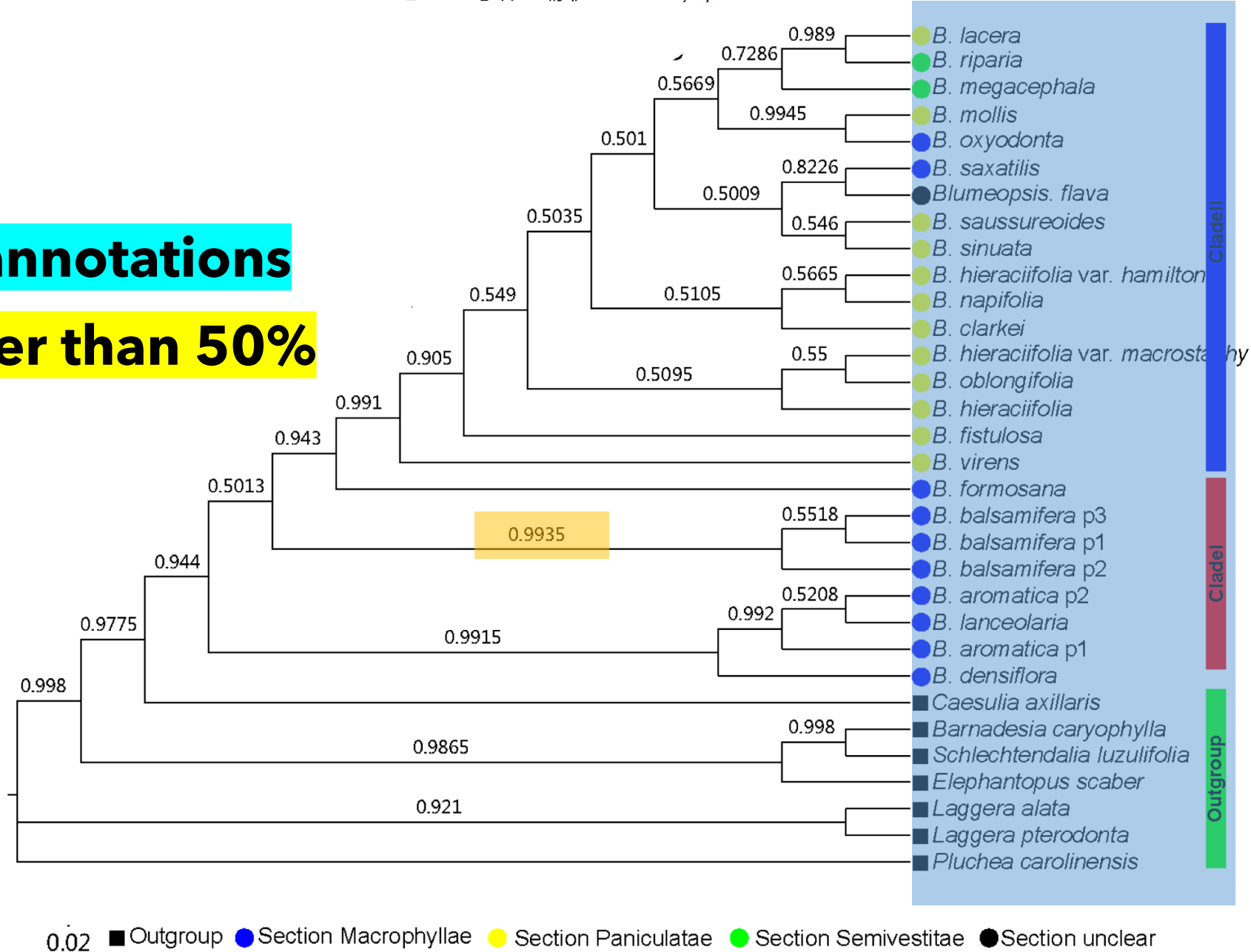# How do you test the tree?

Bootstrapping
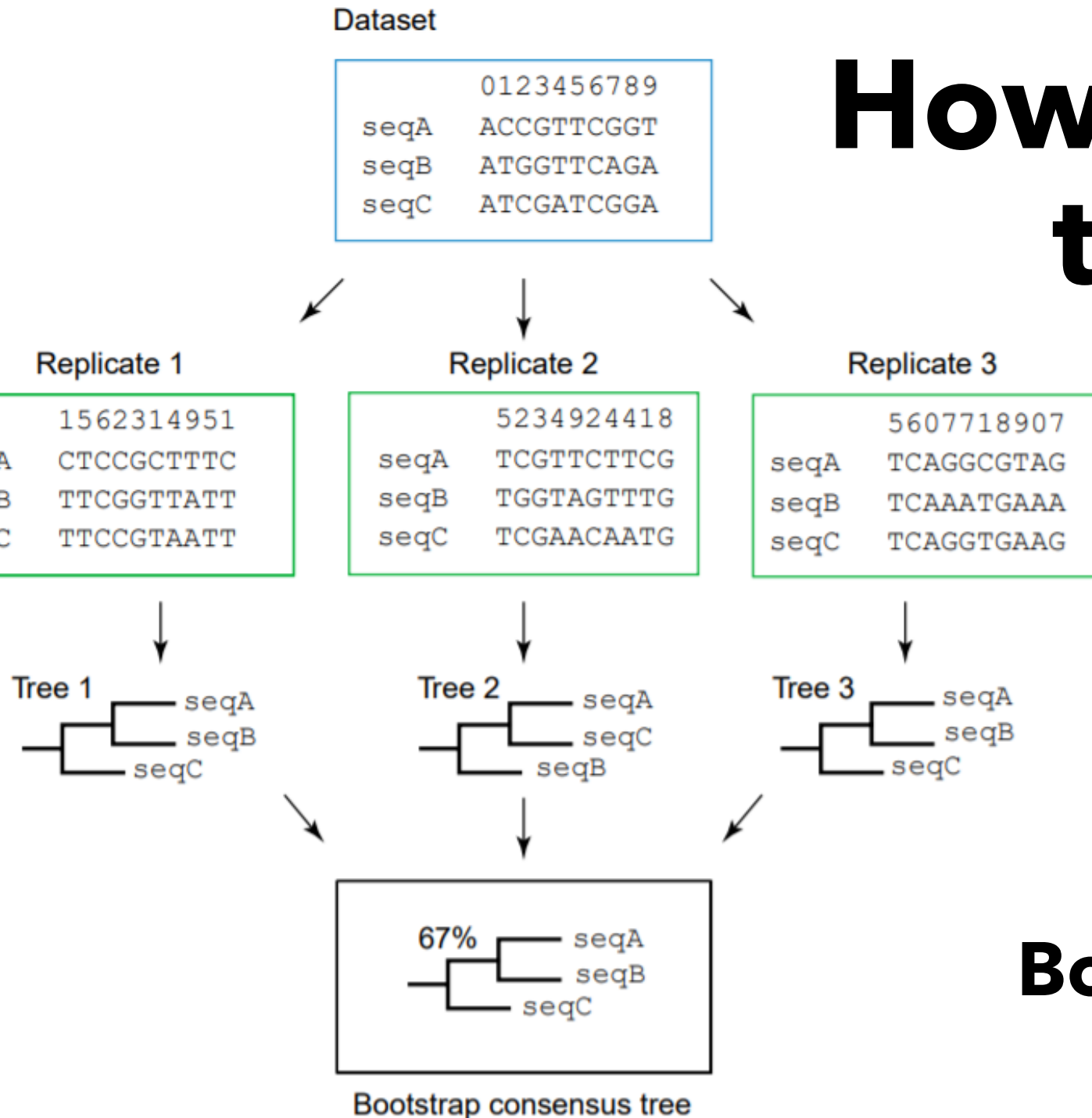
Bootstrap value of 70% and up are reliable

What should you include in the tree?

Proper labeling and annotations
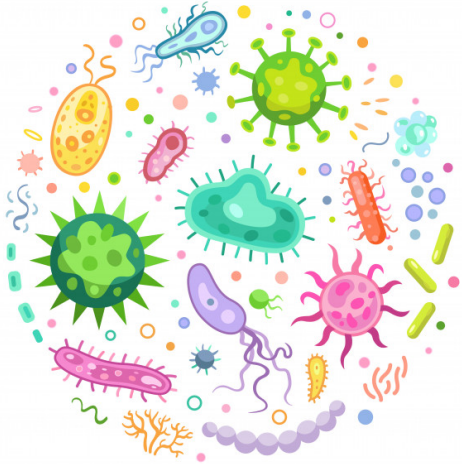
Bootstrap values higher than 50%

What are the disadvantages of phylogeny trees?

True tree → Inferred tree

long branch attraction

Phylogeny can contain errors

What are the advantages of phylogeny trees?

Extensive introgression in a malaria vector species complex revealed by phylogenomics

Michael C. Fontaine[1,2,*,†], James B. Pease[3,*], Aaron Steele[4], Robert M. Waterhouse[5,6,7,8], Daniel E. Neafsey[6], Igor V. Sharak...

+ See all authors and affiliations

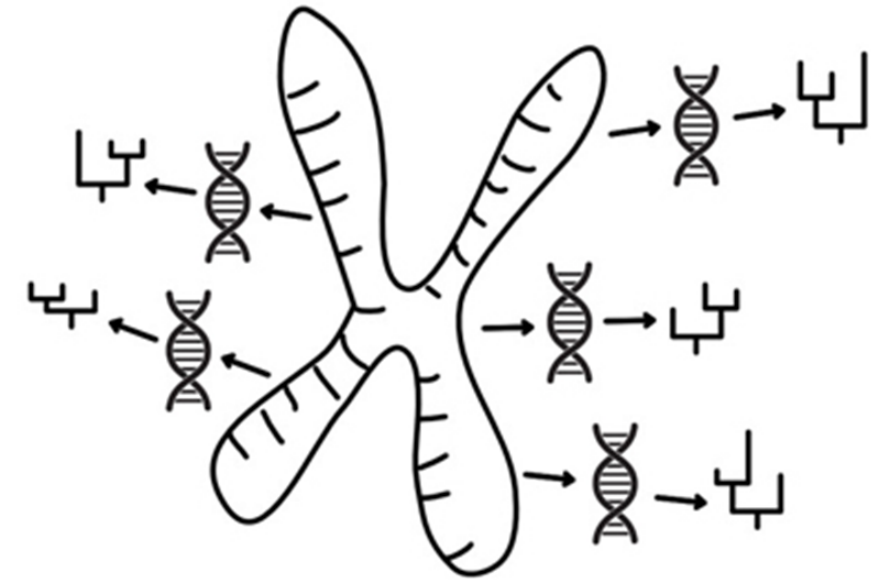## Evolutionary prediction of medicinal properties in the genus *Euphorbia* L.

Madeleine Ernst ✉, C. Haris Saslis-Lagoudakis, Olwen M. Grace, Niclas Nilsson, Henrik Toft Simonsen, James W. Horn & Nina Rønsted

Summary

© Jeremy M Brown

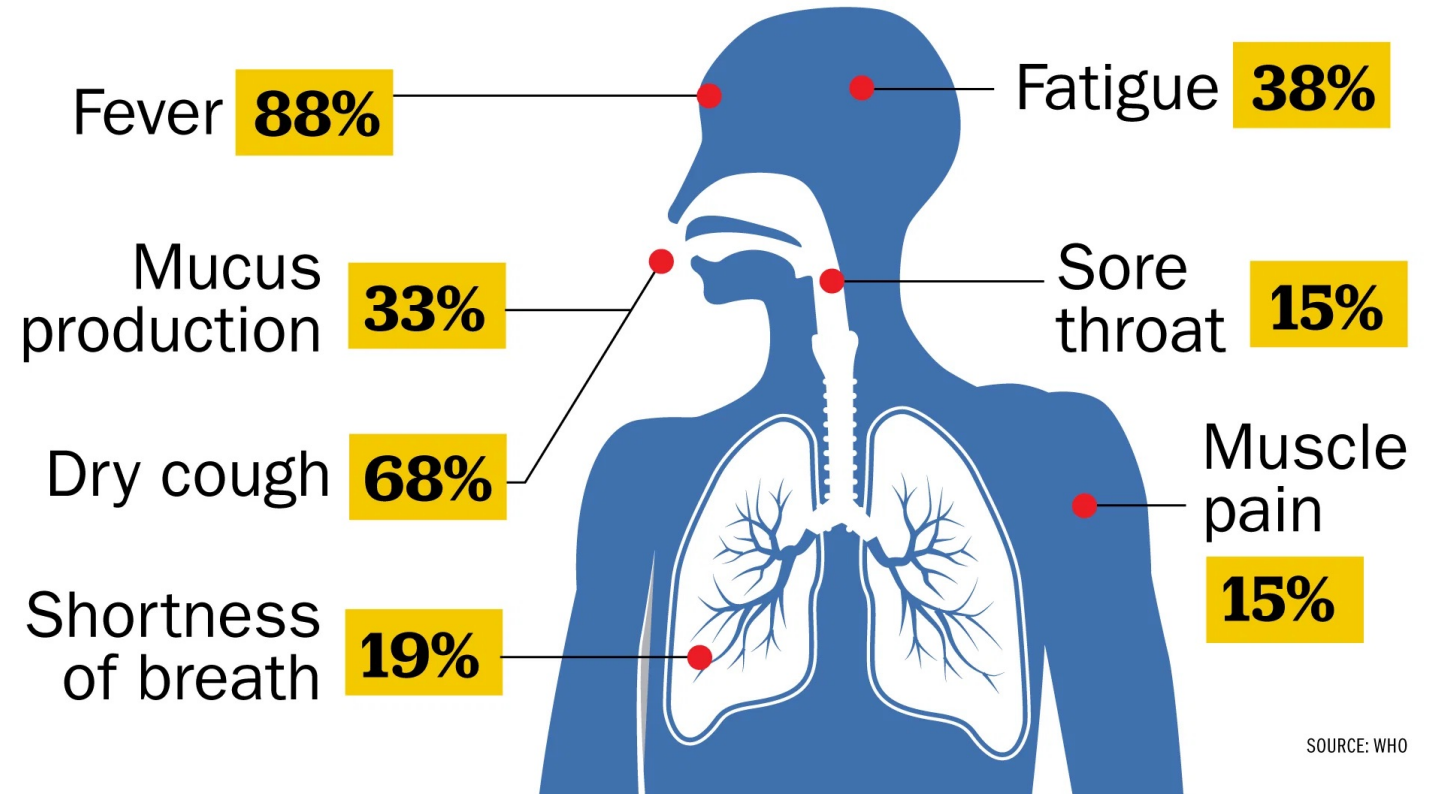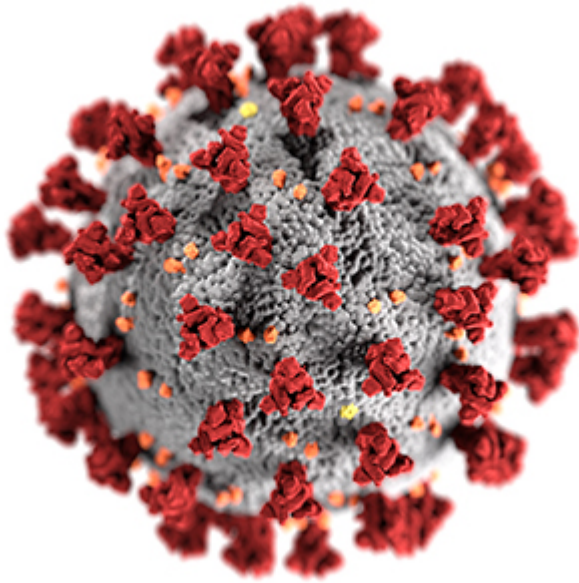**Phylogenomics reconstruct evolutionary pathways of a group of organisms**

Phylogeny trees are made through sequence comparisons, multiple sequence alignment

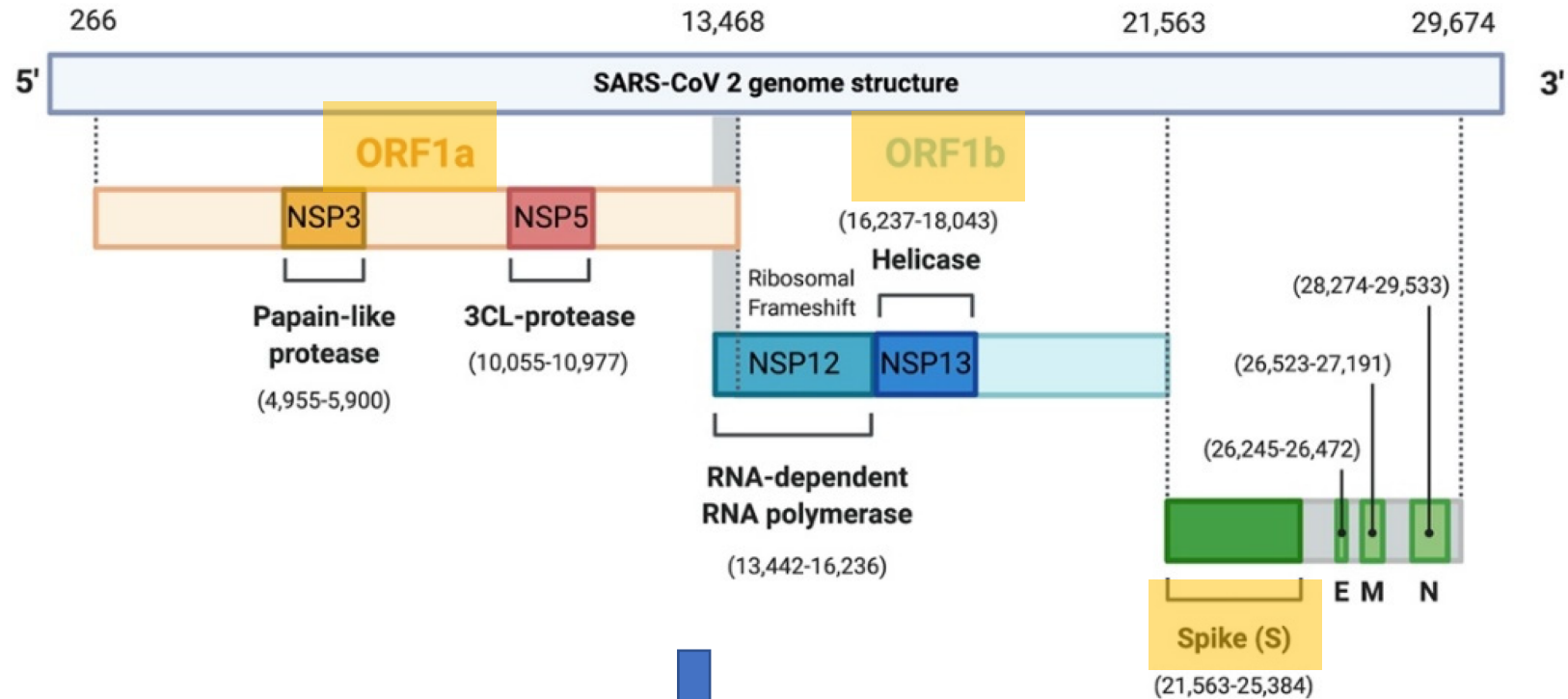**Many methods and tests allow accurate representations of phylogeny trees**
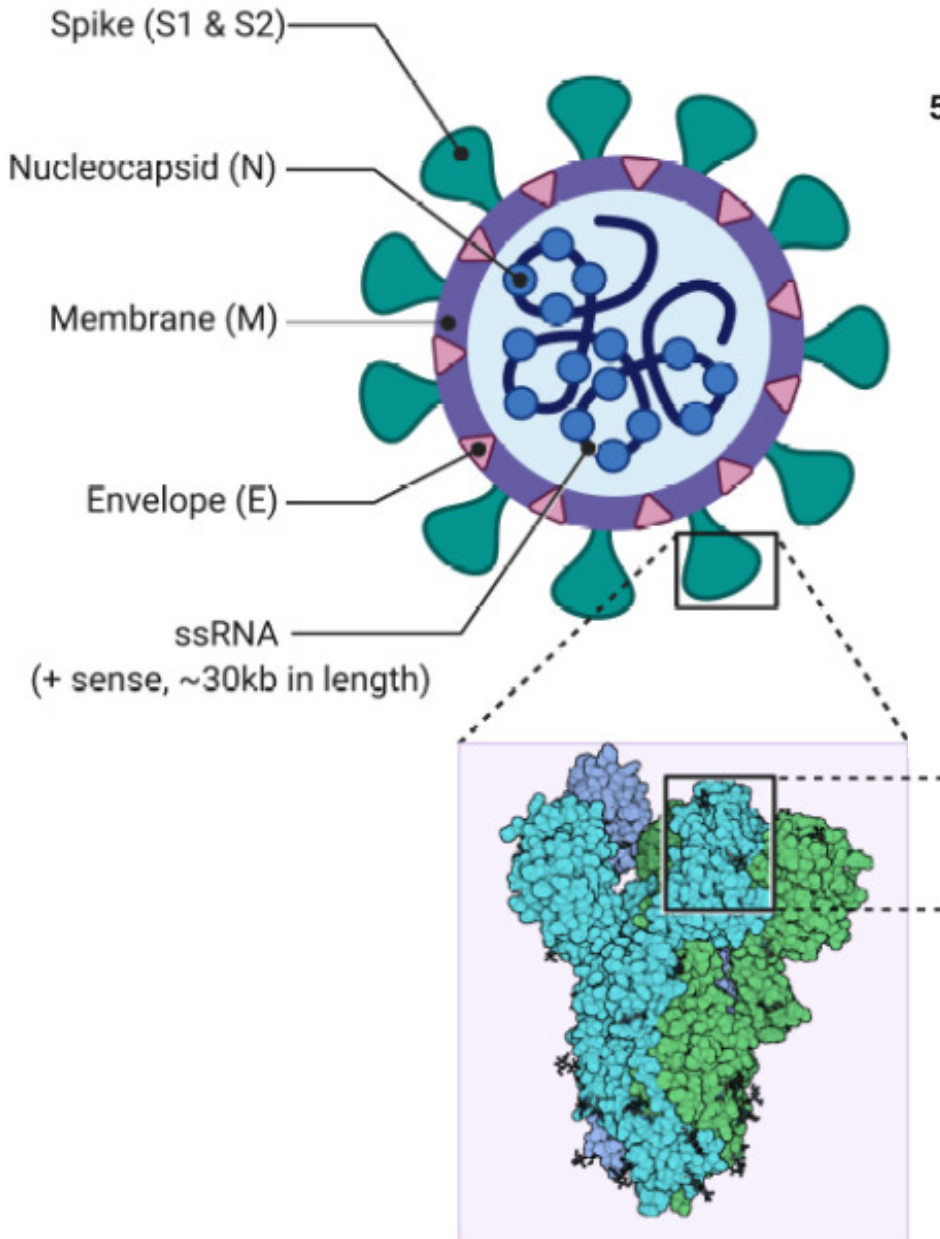
What is SARS-CoV-2?

Fever **88%**

Fatigue **38%**

Mucus production **33%**

Sore throat **15%**

Dry cough **68%**

Muscle pain **15%**

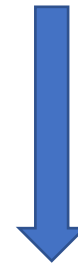Shortness of breath **19%**

SOURCE: WHO

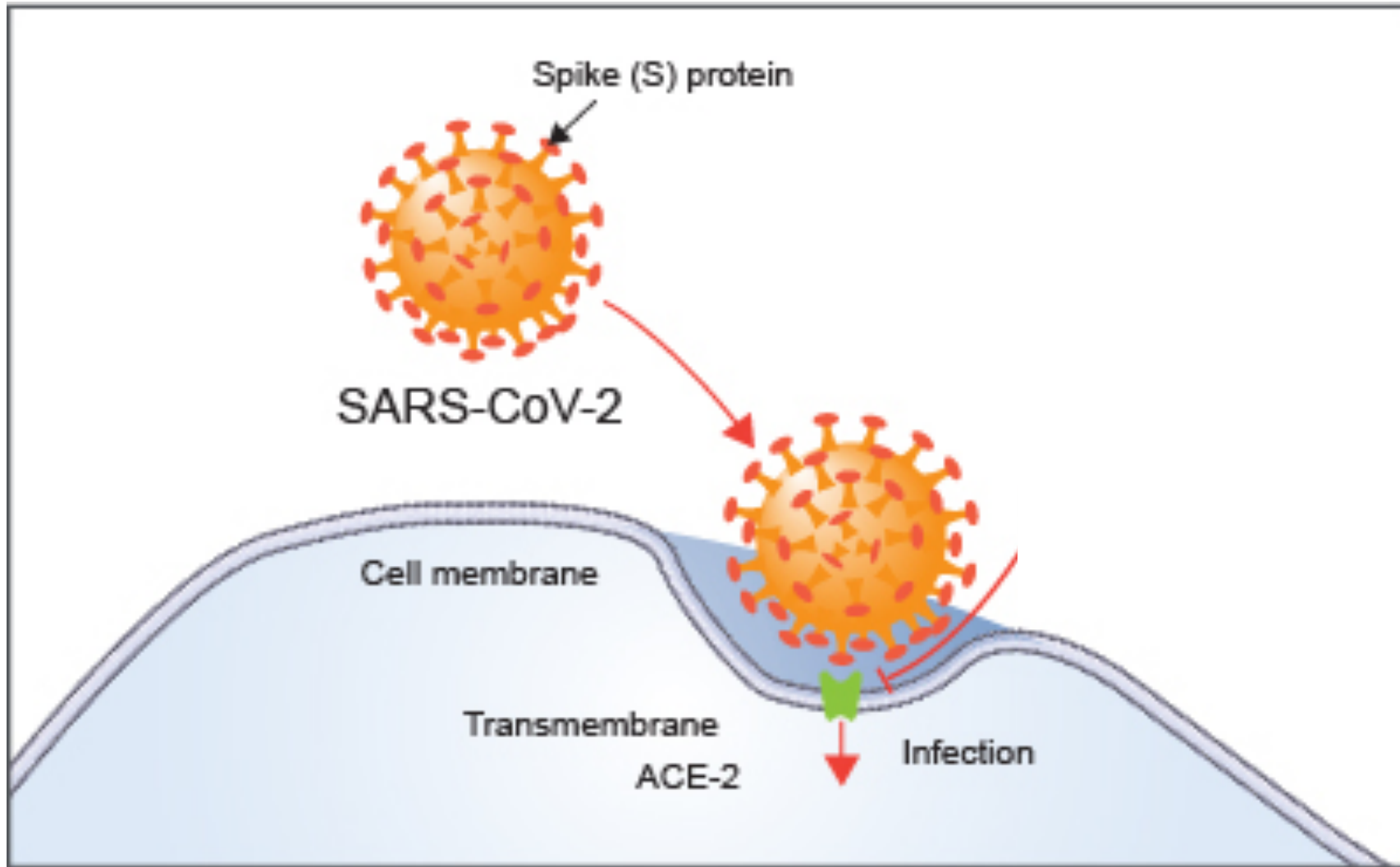# What does the SARS-CoV-2 genome look like?

Binding domain

RNA synthesis and replication

How does SARS-CoV-2 infect respiratory cells?

Spike protein binds to the ACE2 receptors on cell membranes

Feared South African Covid-19 Mutant Discovered In California As It Surpasses NY As State With Most Virus-Related Deaths In US

By Tom Tapp
February 10, 2021 1:18pm

COVID-19 virus variants are evolving 'escape' mutations, and immune-compromised patients provide a breeding ground
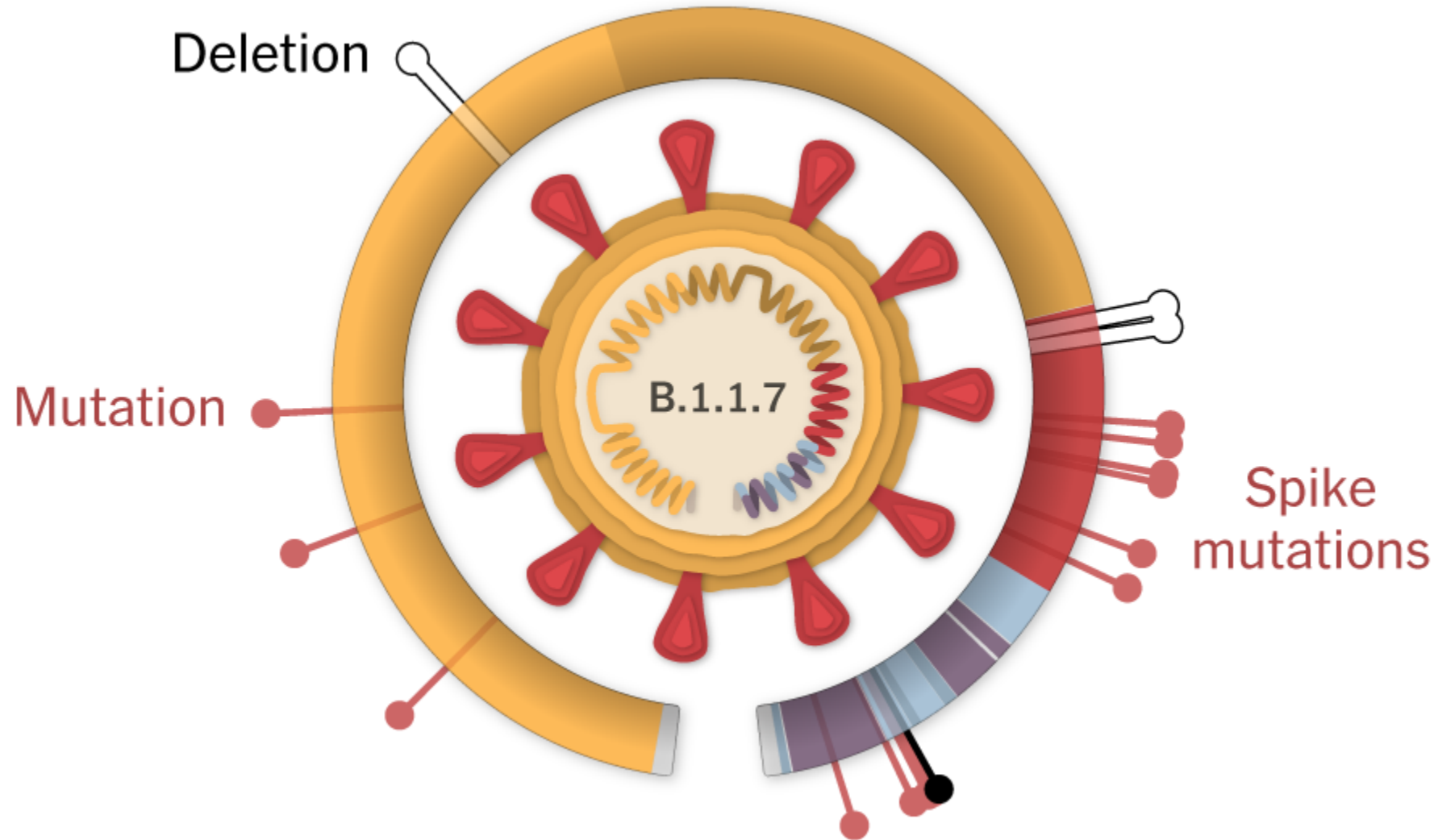
by Marie McCullough, Posted: February 12, 2021

The Mutated Virus Is a Ticking Time Bomb

There is much we don't know about the new COVID-19 variant—but everything we know so far suggests a huge danger.
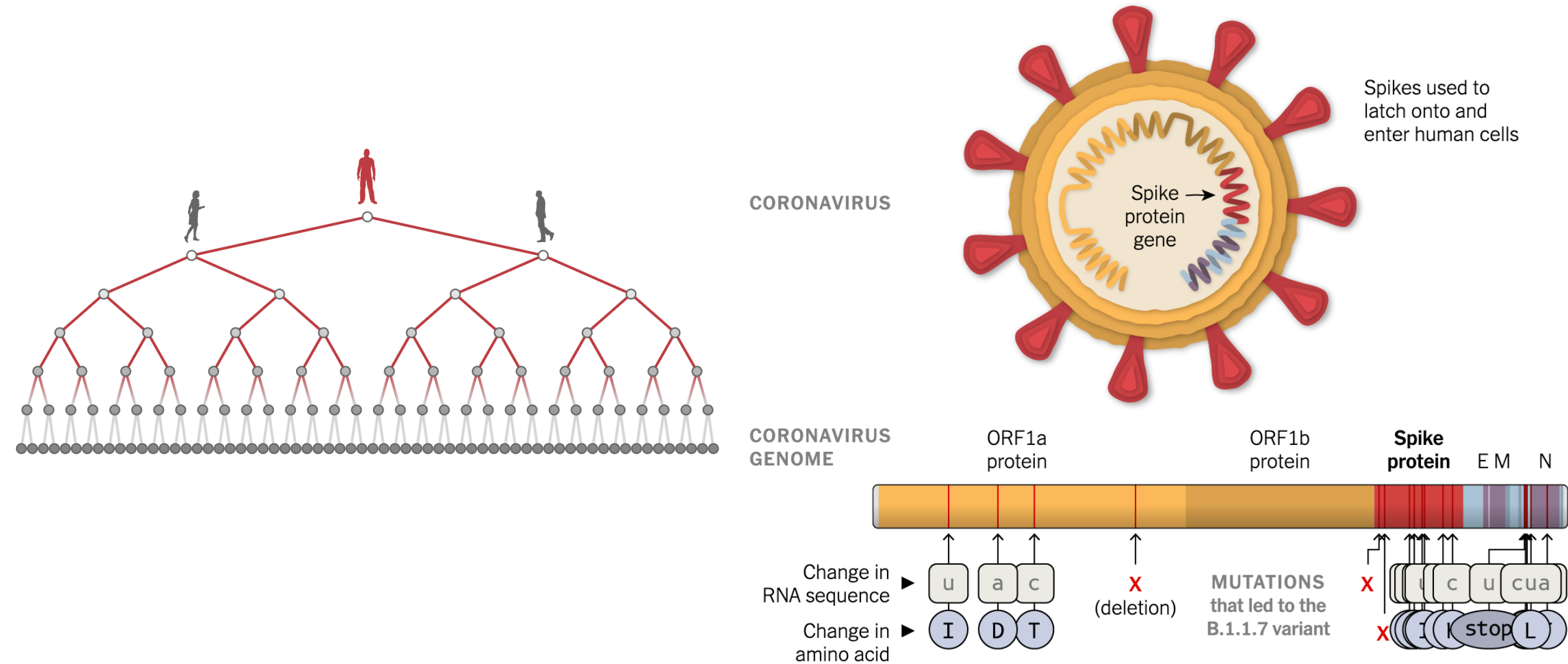
ZEYNEP TUFEKCI   DECEMBER 31, 2020

What parts of SARS-CoV-2 are mutating?

Spike protein mutations can enhance transmission

How are researchers tracking the SARS-CoV-2 variants?

CORONAVIRUS

Spikes used to latch onto and enter human cells

Spike protein gene

CORONAVIRUS GENOME

ORF1a protein

ORF1b protein

Spike protein E M N

Change in RNA sequence ▶ u a c

Change in amino acid ▶ I D T

X (deletion)

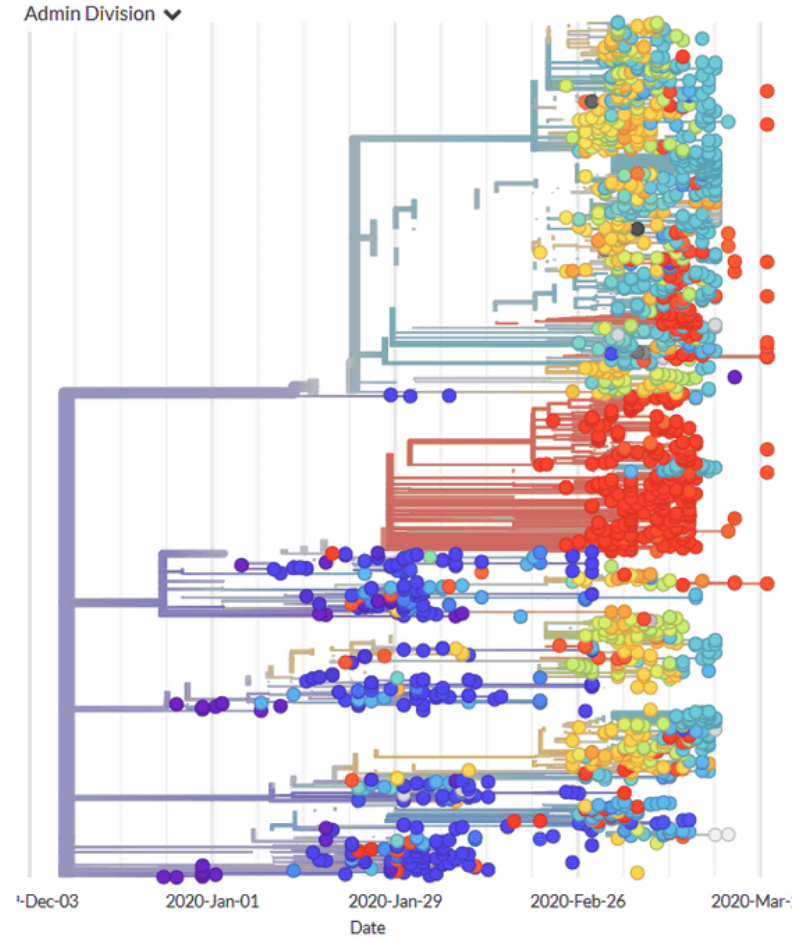MUTATIONS that led to the B.1.1.7 variant

X

X

u c u cua

I K stop L

They track the changes in the RNA sequence in a population

Who is Dr. Trevor Bedford?

Bedford Lab
Fred Hutchinson Cancer Research Center,
Vaccine and Infectious Disease Division

https://nextstrain.org/

References

- Baldauf SL. Phylogeny for the faint of heart: a tutorial. *Trends Genet*. 2003;19(6):345-351. doi:10.1016/S0168-9525(03)00112-4
- Kapli, P., Yang, Z., Telford, M. Phylogenic tree building in the genomic age. *Nature Reviews Genetics*. May 18th, 2020; 21:428-444
- Bawono, P., Heringa, J. Volume 6: Bioinformatics. *Comprehensive Biomedical Physics*. Elsevier; 2014.
- Som, A. Causes, consequences and solutions of phylogenetic incongruence. *Briefing in Bioinformatics*. 2015; 16(3):536-648
- Chowdhury, J., Scarr, S., Wardell, J. How the novel coronavirus has evolved. *Reuters Graphics*. December 10th, 2020. Accessed February 13th, 2021. https://graphics.reuters.com/HEALTH-CORONAVIRUS/EVOLUTION/yxmpjqkdzvr/