

Evolutionary Thinking 2022

TA session

week 2 – substitution models

Jilong Ma
aujilongm@birc.au.dk

Outline

1. Recap of last week's contents (5 minutes)

2. The learning outcome of this week (20 minutes)

Substitution Models – Using them for what?

Basic models – JC and Kimura

Distance-based phylogeny tree building algorithms

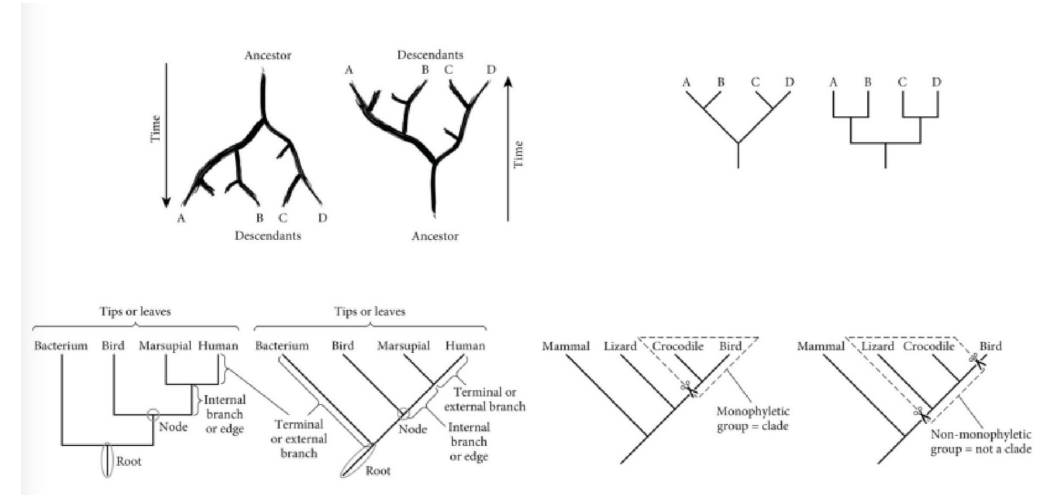
UPGMA, NJ tree

3. Working on paper discussion Modeltest (40 minutes)

Going through solutions (20 minutes)

Recap of last week

1. Reading phylogeny trees (Wednesday)



2. Sequence alignment, distance matrix, and public database (Friday)

		C	A	T	T	A	T	C	G
	0	0	0	0	0	0	0	0	0
C	0	5	1	0	-1	-2	-3	5	1
G	0	1	2	-2	-3	-4	-5	1	10
T	0	0	-2	7	3	2	1	0	6
T	0	-1	-3	3	12	8	7	6	5
T	0	-2	-4	2	8	9	13	9	8
C	0	5	1	1	7	5	9	18	14
G	0	1	2	0	6	4	8	14	23

C	A	T	T	A	T	C	G
C	G	T	T	-	T	C	G

	Human	Chimp	Goril
Human			
Chimp	30		
Goril	31	8	
Mouse	152	139	133

Learning outcome of this week

1. Substitution Models (Wednesday)

Why we need substitution models/ What can we use it for ?
Examples - JC model and Kimura 2 parameter model
Nested models and testing for best fit models (Paper)

2. Tree building methods (Friday)

Distance based – UPGMA and NJ
Maximum Parsimony
Maximum Likelihood
MEGA practice

Learning outcome of today

1. Substitution Models (Wednesday)

Why we need substitution models/ What can we use it for ?
Examples - JC model and Kimura 2 parameter model
Nested models and testing for best fit models (Paper)

Discussion (5 minutes)

1. Substitution Models (Wednesday)

Distance between two nucleotide sequence

Pairwise Distance

Number of Substitutions

What are these two terms

What is the difference between them?

Which one is uncertain?

Which one is larger?

How can we build connection between them



Discussion

1. Substitution Models (Wednesday)

Distance between two nucleotide sequence

Pairwise Distance

Number of Substitutions

What are these two terms

What is the difference between them?

Which one is uncertain?

Which one is larger?

How can we build connection between them

████████████████████

Substitution models correct for the multiple substitutions



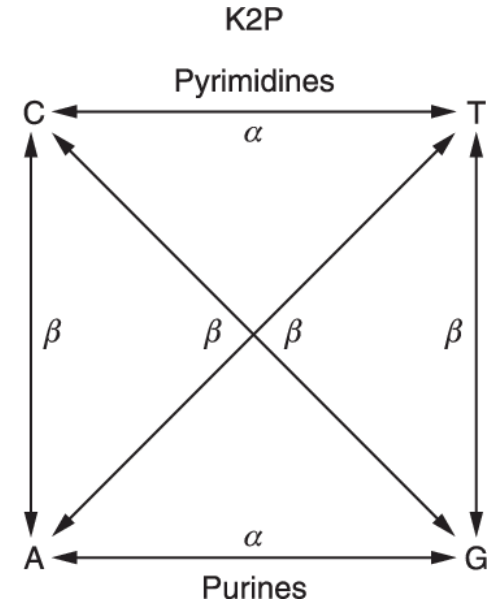
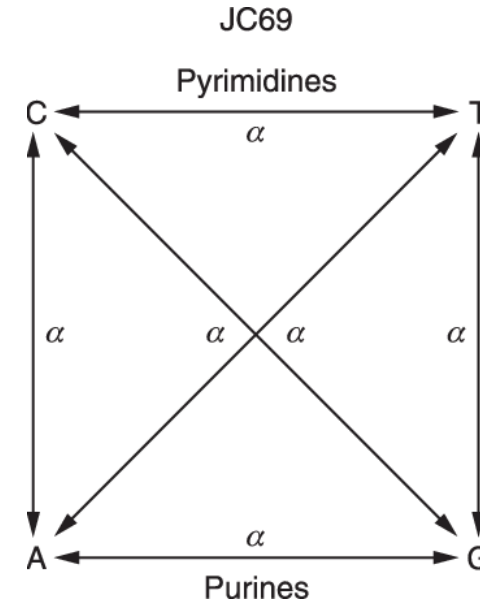
Basic substitution models (5 minutes)

1. Substitution Models (Wednesday)

What is alpha/beta

What interested values can we calculate by using the model ?

What are the values calculated can be used for ?



Basic substitution models

1. Substitution Models (Wednesday)

What is alpha/beta

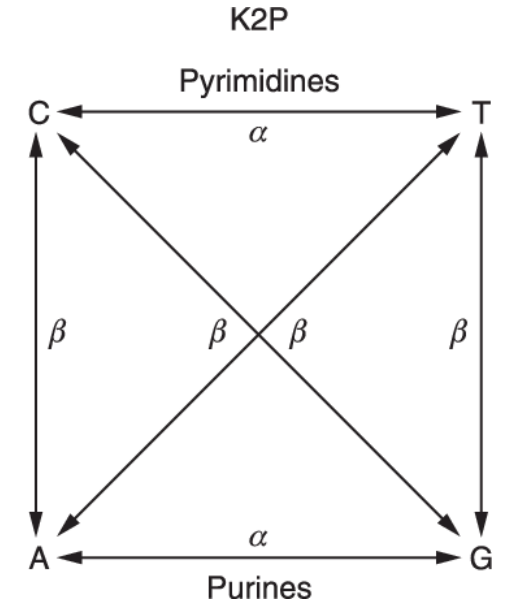
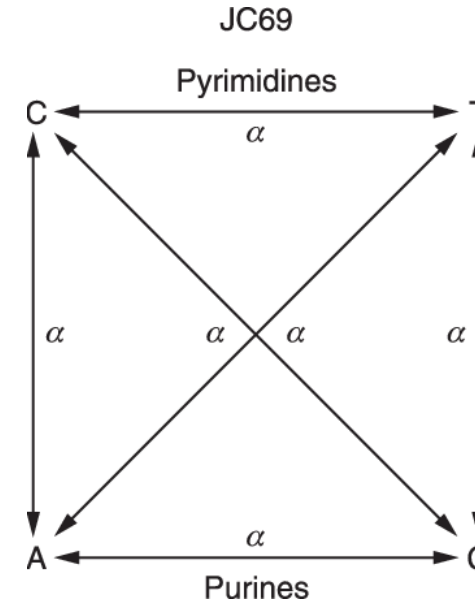
Probability of substitution (A→C) per unit time/generation

What interesting values can we calculate by using the model?

Expected number of substitutions since divergence (K on Dan Graur P85)

Probability of observing some sequence data given some phylogeny tree.

What are the values calculated can be used for ?



Correct for a more precise evolutionary distance.
Maximum likelihood
Building Trees! (Friday)

Paper Discussion (40 minutes)

1. Go to the Github page week36/Wednesday
2. Discussion the questions together with your group mates
3. Write down your discussed answer in the shared google doc
Q1:
 Group 1:
 Group 2:



AARHUS
UNIVERSITY