Today: models

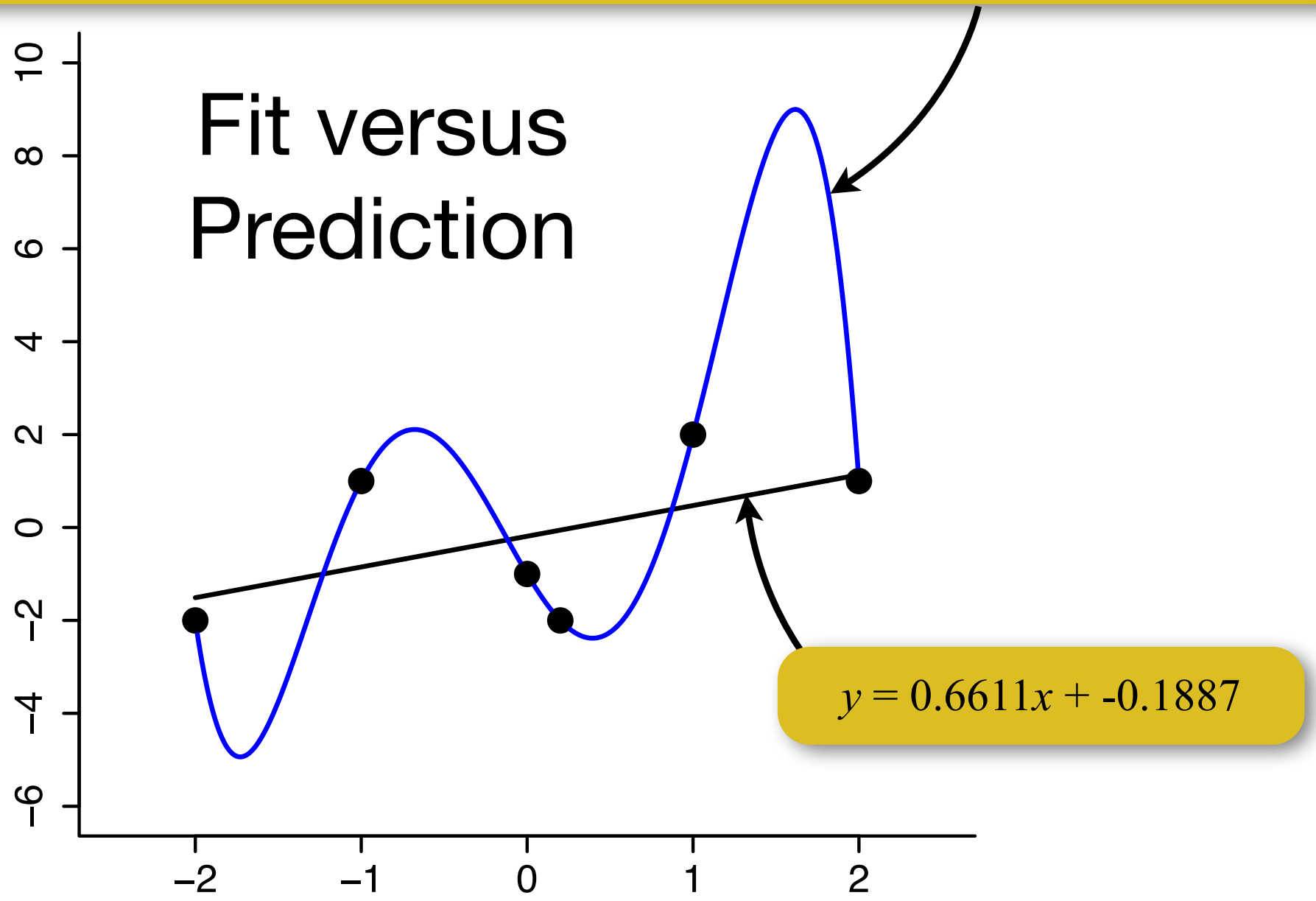
- More detail may mean less useful
- Dimension of a model
- Transition probabilities
- Jukes and Cantor (1969) model
- Long run behavior of a model
- Meet the other members of the family (F81, K80, F84, HKY, GTR)

A very *practical* MBTA subway map

A very *realistic* MBTA subway map

Which is more useful?

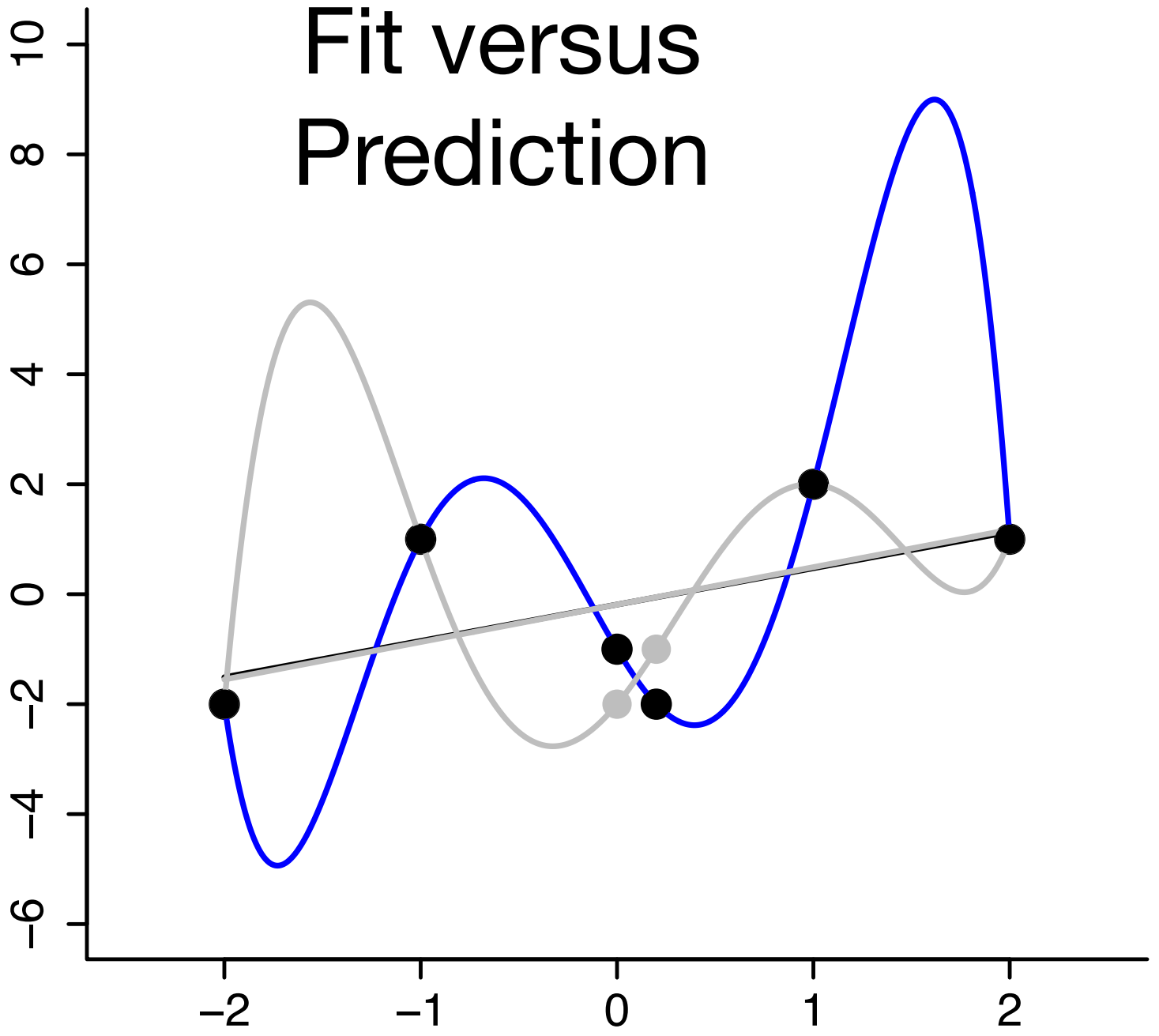
$y = -1.5972\,x^5 + -0.7917\,x^4 + 8.0694\,x^3 + 3.2917\,x^2 + -5.9722\,x + -1.0$



goodness-of-fit

prediction

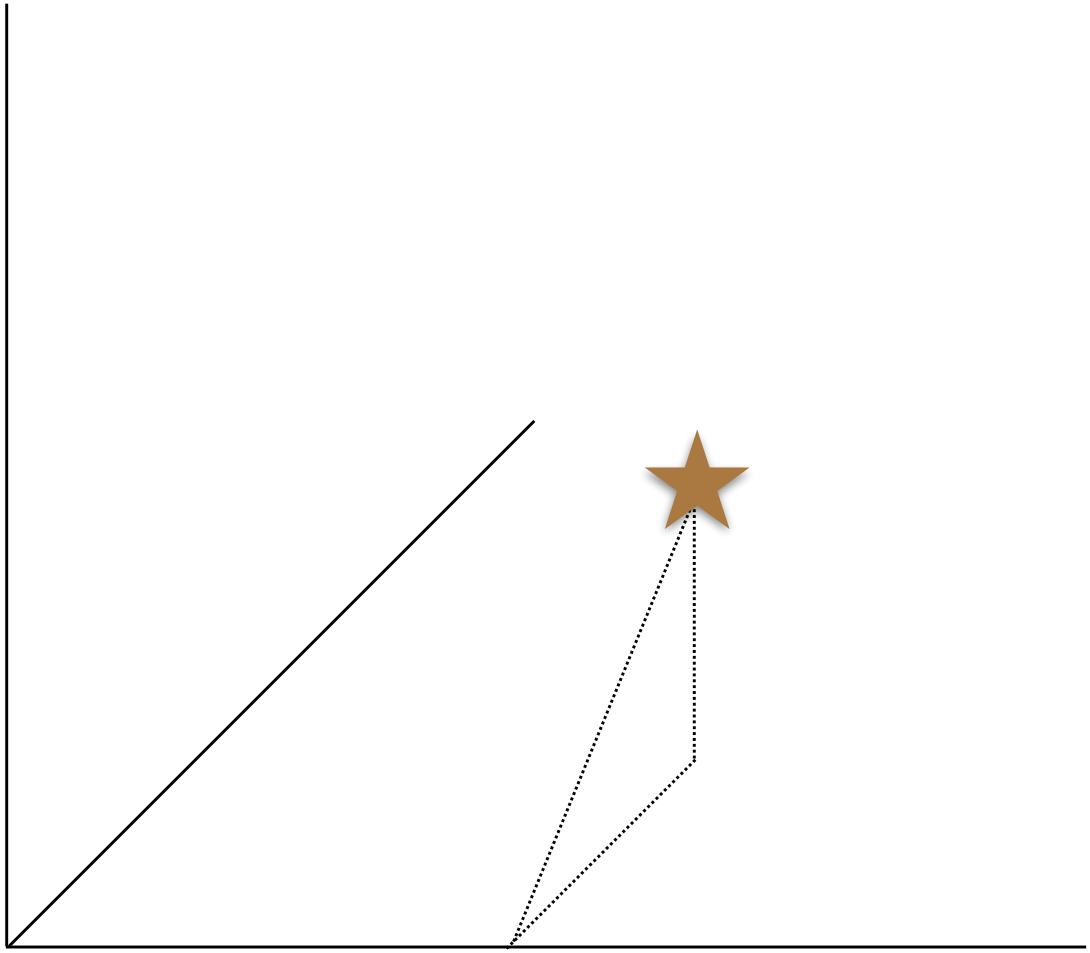
(Thanks to David Swofford for pointing out this great example to me.)



complex models

simple models

Model dimensions

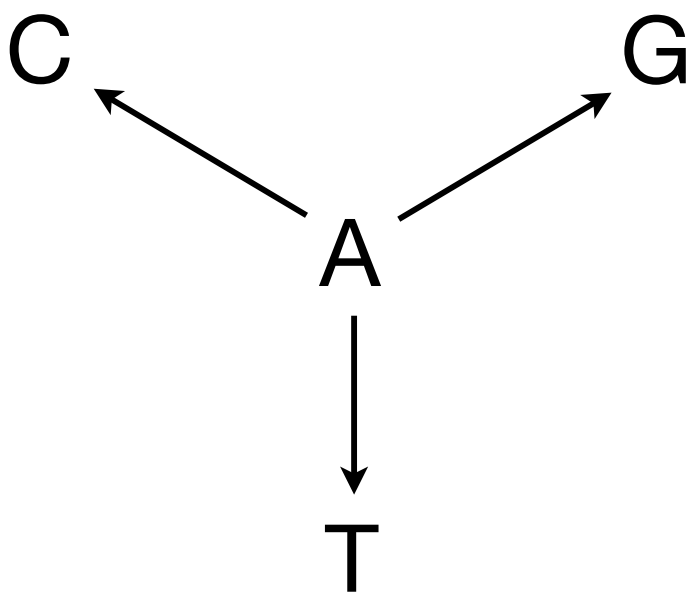


1-parameter model: 2.83

2-parameter model: 2.00

3-parameter model: 0.00

JC69 model



Jukes & Cantor (1969)

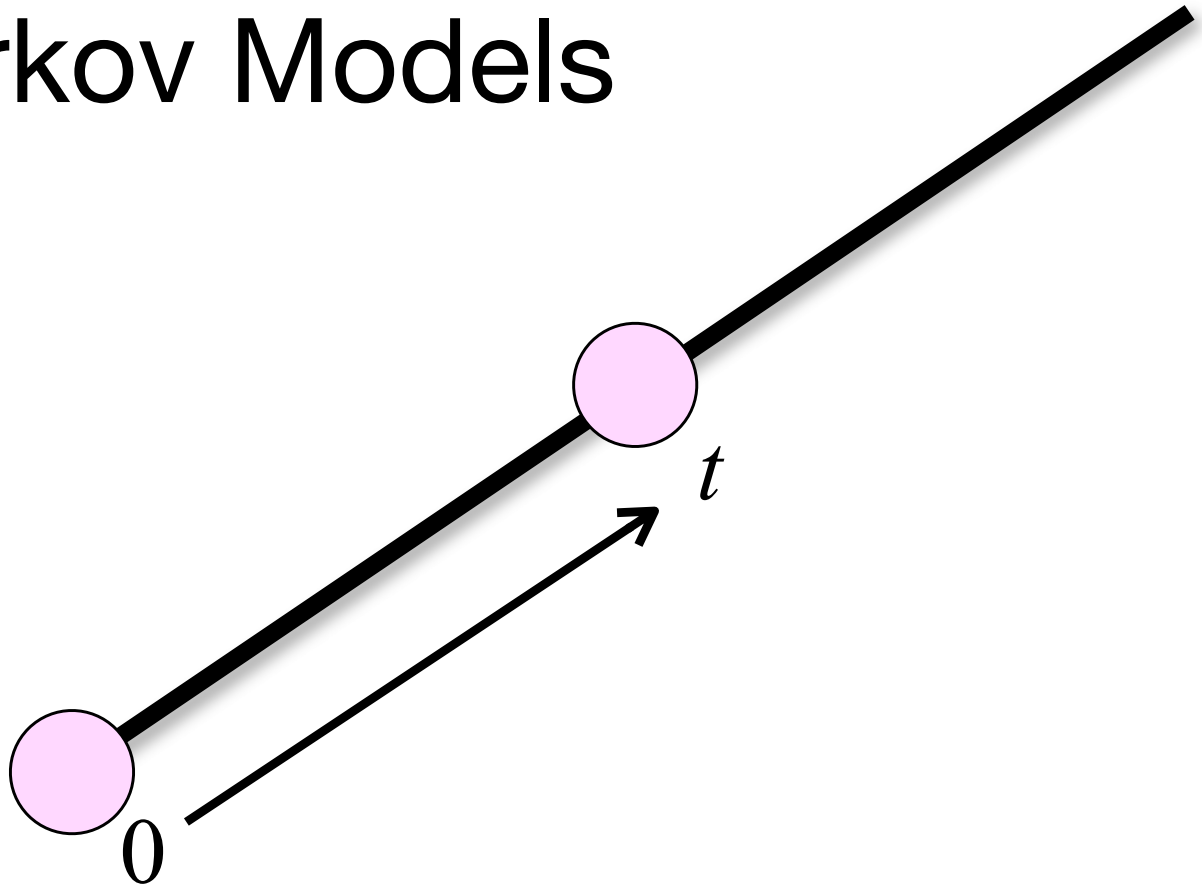
number of substitutions
substitution rate
total substitution rate

Edge lengths

edge length parameters

long edge lengths means...

Markov Models



Markov property

transition probability

conditional probability

multiple hits

Evolutionary Distance

JC69 Transition Probability

$$P_{TG}(t) = \frac{1}{4}\left(1 - e^{-4\beta t}\right)$$

e = 2.718281828459045...

achnow rate vs. substitution rate

poisson distribution

JC69 Transition Probability

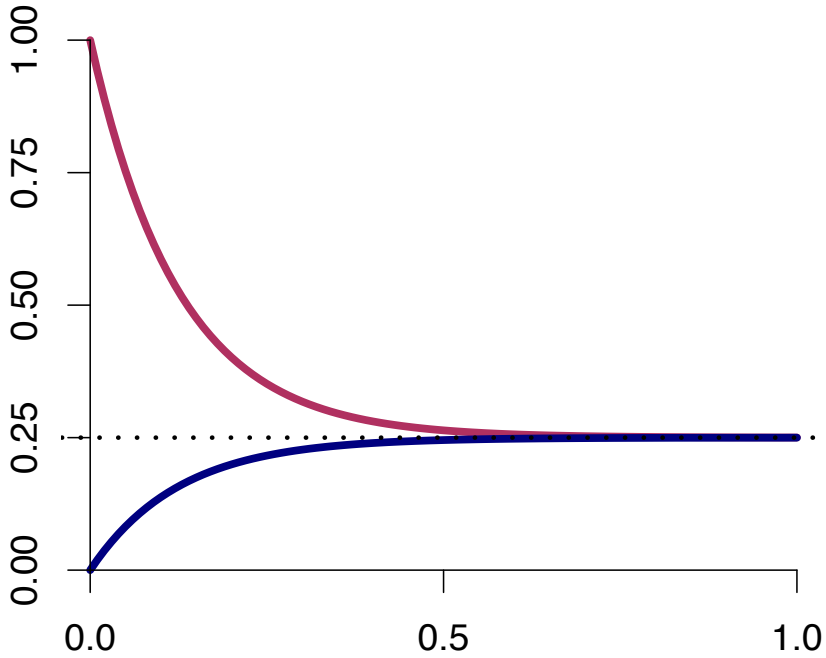
$$P_{TA}(t) = \frac{1}{4} - \frac{1}{4}e^{-4\beta t}$$

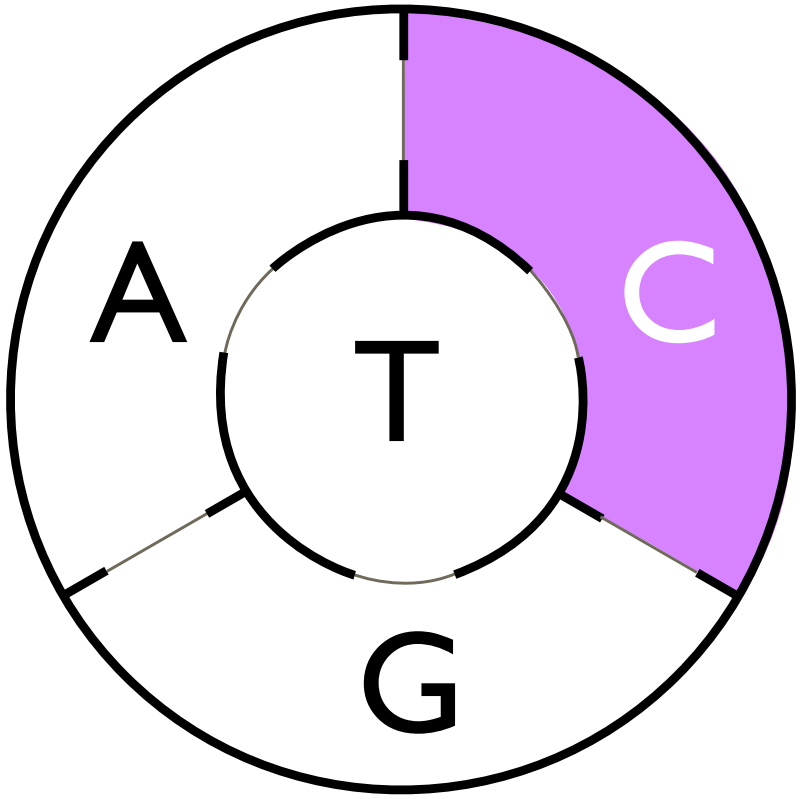
$$P_{TC}(t) = \frac{1}{4} - \frac{1}{4}e^{-4\beta t}$$

$$P_{TG}(t) = \frac{1}{4} - \frac{1}{4}e^{-4\beta t}$$

$$P_{TT}(t) = \frac{1}{4} - \frac{1}{4}e^{-4\beta t}$$

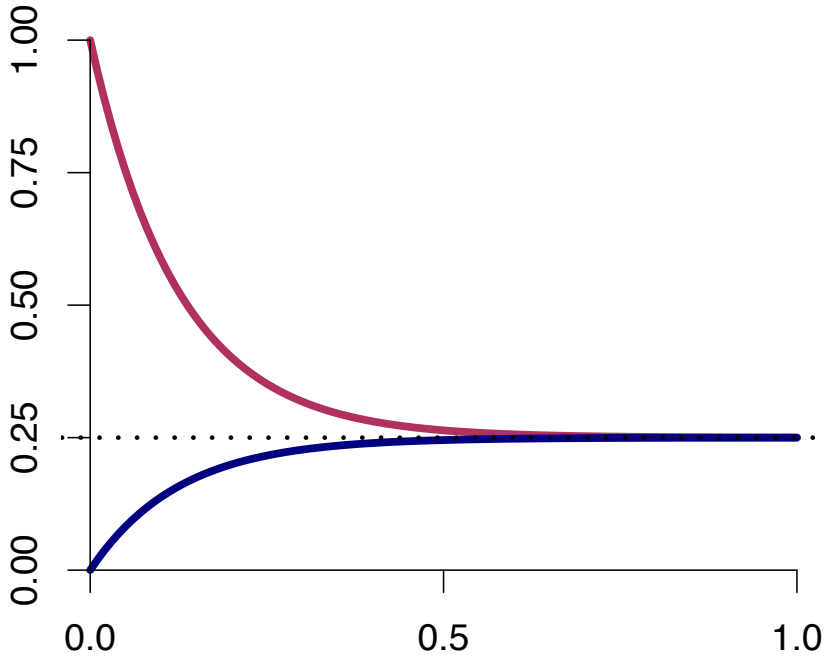
Equilibrium Frequencies

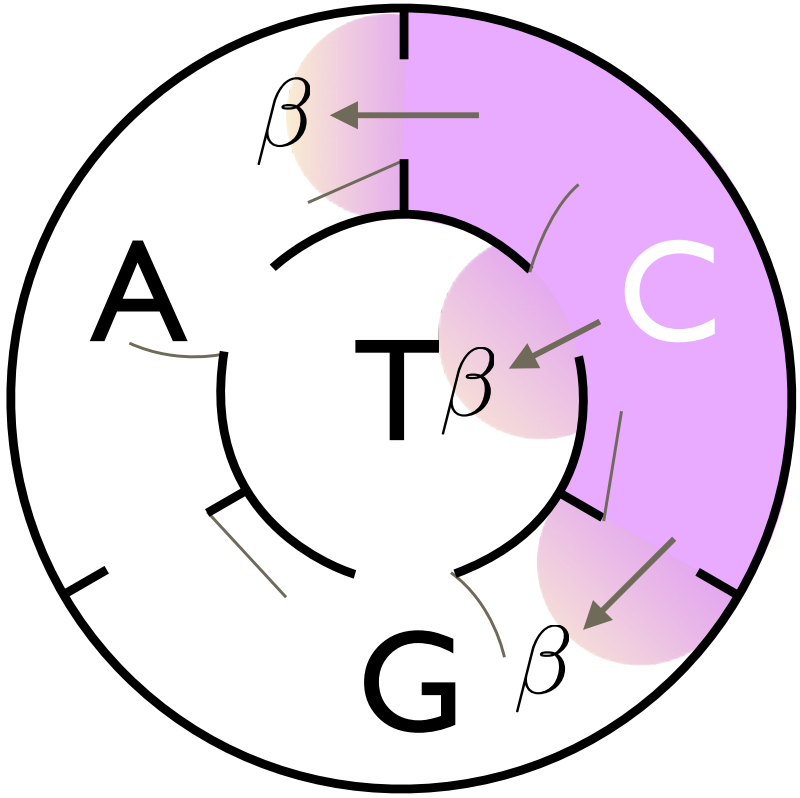




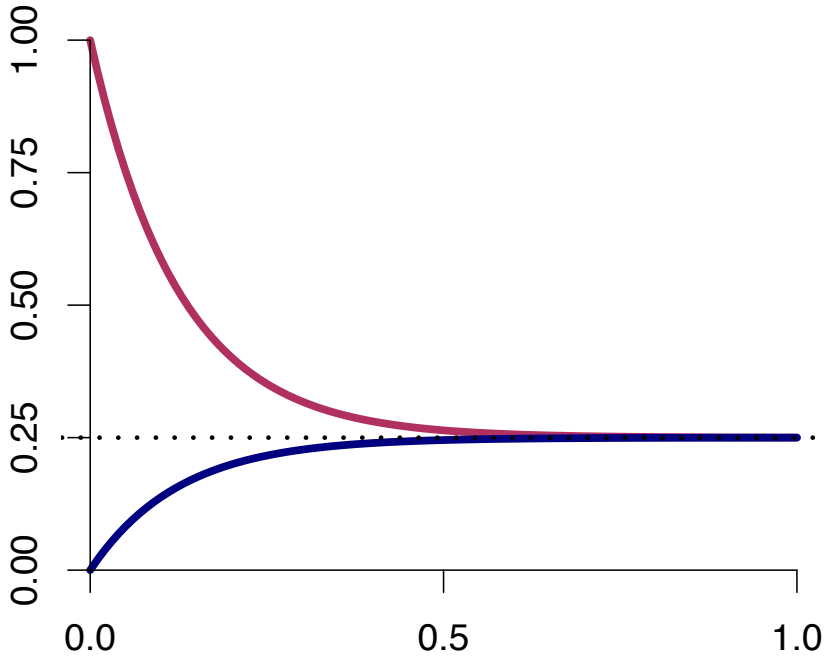
(architect: Joe Bielawski)

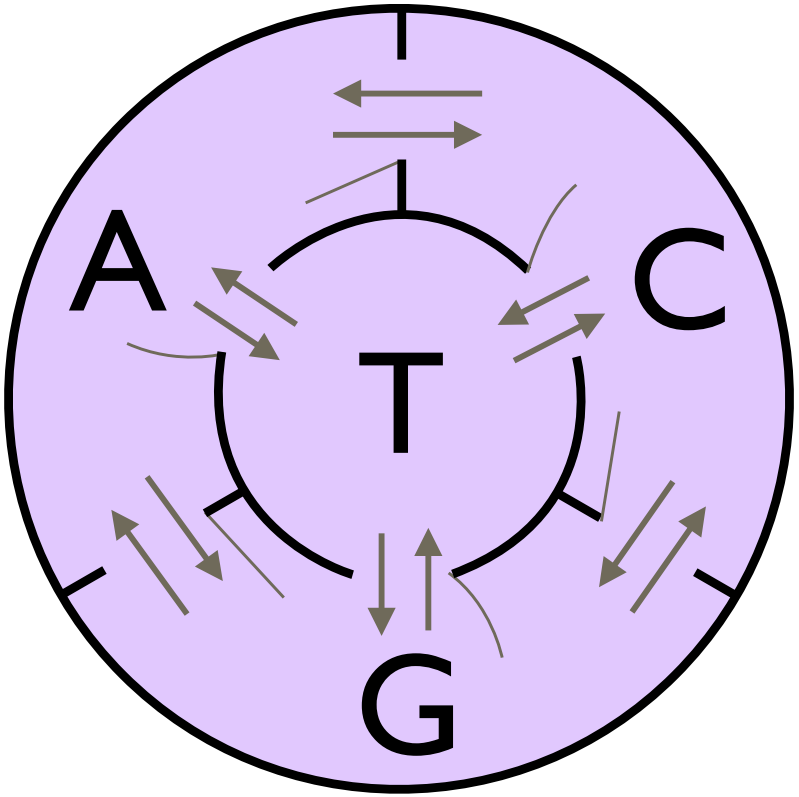
Equilibrium Frequencies





Equilibrium Frequencies





$$\pi_A = \pi_C = \pi_G = \pi_T = \frac{1}{4}$$

JC69 instantaneous rate matrix

		"To" state			
		A	C	G	T
"From" state	A				
	C				
	G				
	T				

Jukes & Cantor (1969)

parameters?

K80 (K2P) rate matrix

“From” state	“To” state: A	C	G	T
A				
C				
G				
T				

rate of transversions

transition/transversion rate ratio

parameters?

equivalence to JC69

Kimura (1980)