Phylogenetic Biology Week 3

Biology 1425

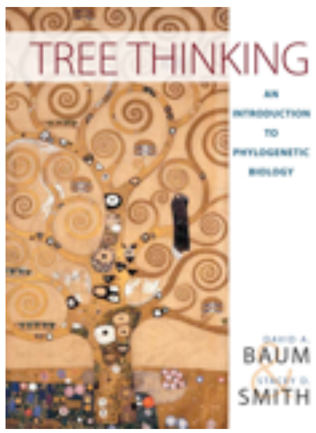
Professor: Casey Dunn, dunnlab.org
Brown University

2013

Front matter...

Sources

Some non-original content is drawn from:



Baum, D and S. Smith (2012) Tree Thinking: and Introduction to Phylogenetic Biology. Roberts and Company Publishers. ISBN 9781936221165

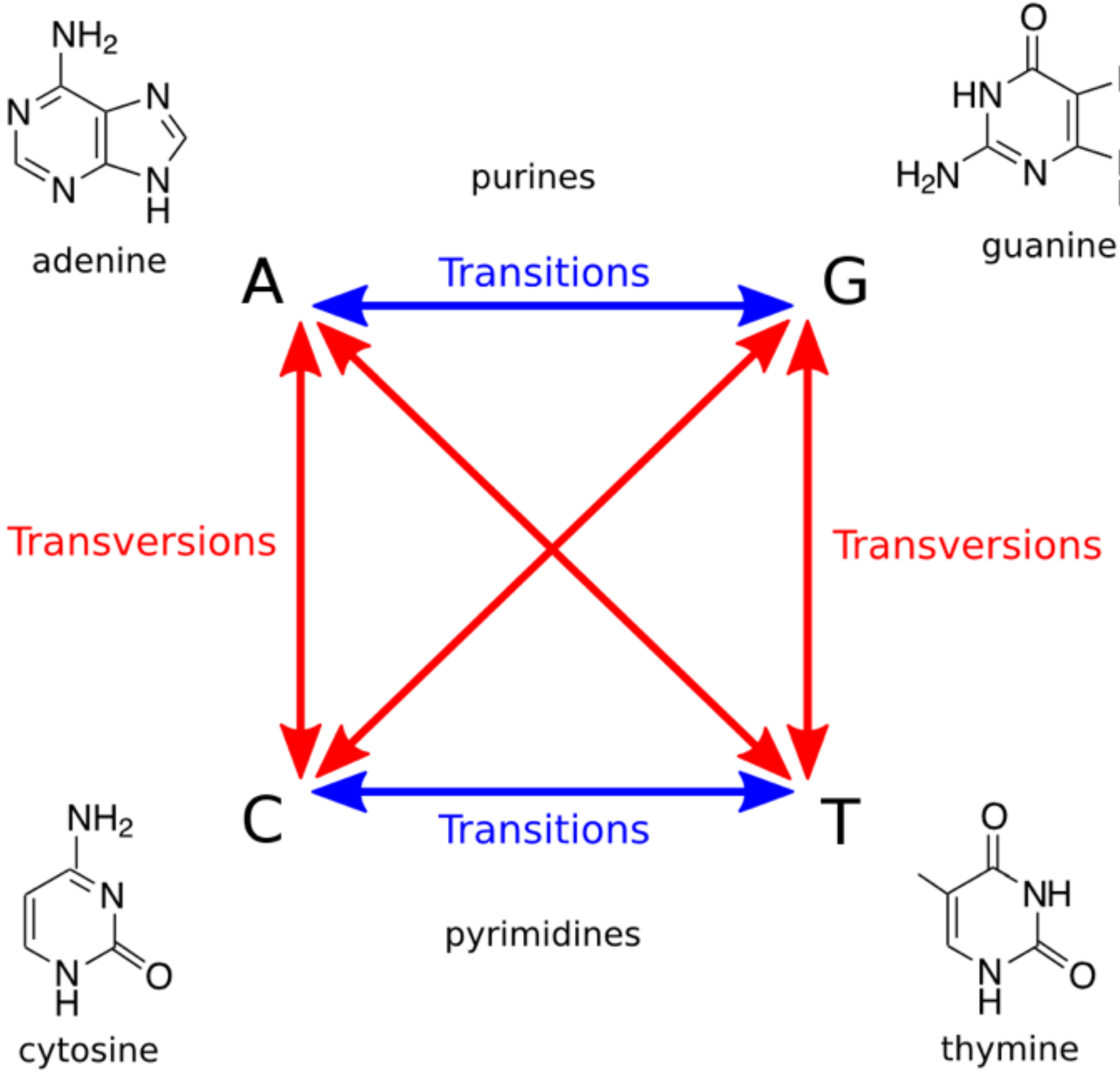
Other non-original content is referenced by url.

Other resources

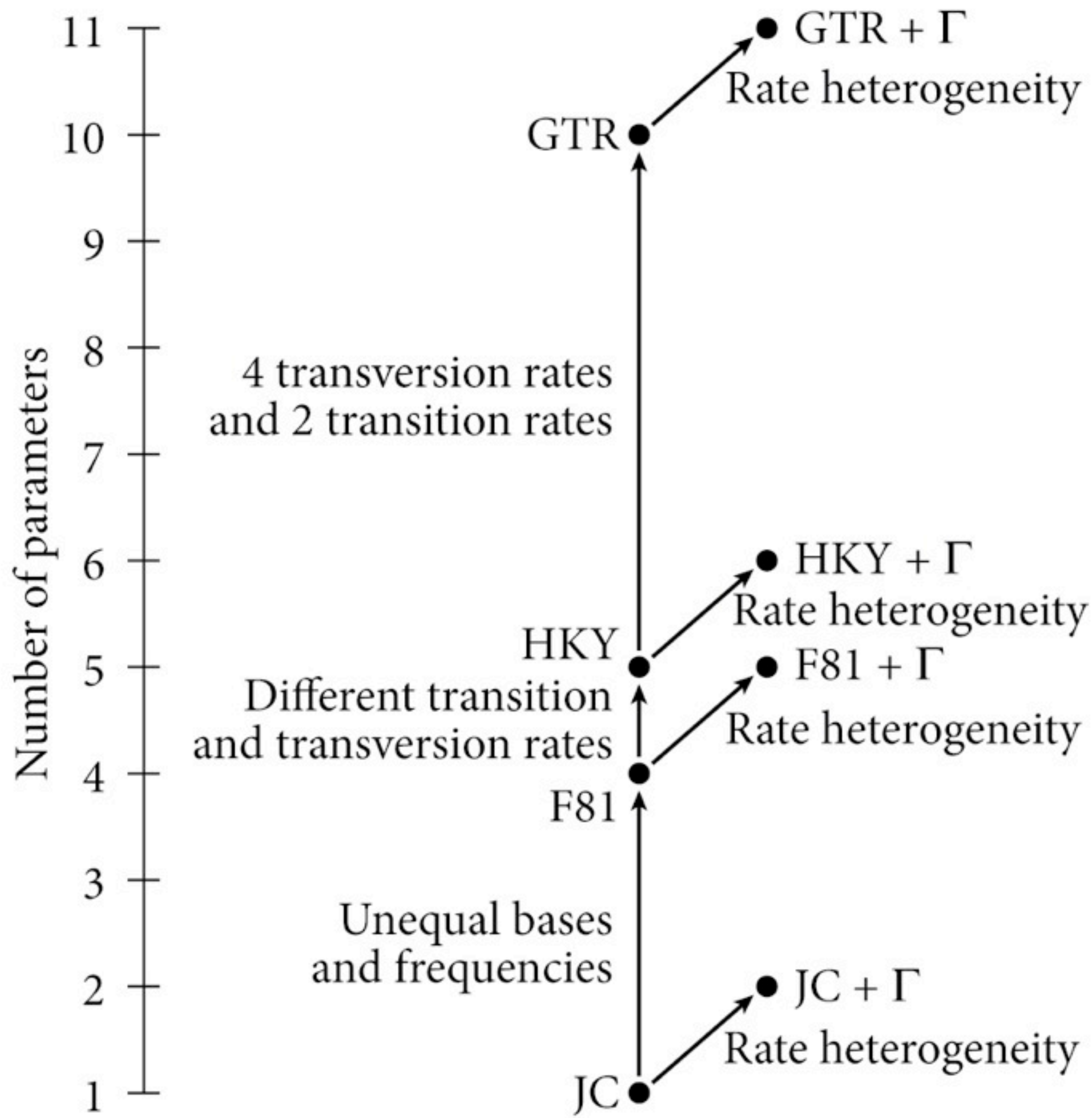
These slides supplement the following excellent presentations from the Wood's Hole Workshop in Molecular Evolution:

Paul Lewis - http://www.eeb.uconn.edu/people/plewis/downloads/wh2012/Likelihood_WoodsHole_24July2012_1-per-page.pdf

John Huelsenbeck - https://molevol.mbl.edu/wiki/images/3/37/WoodsHoleHandout.pdf



http://upload.wikimedia.org/wikipedia/commons/3/35/Transitions-transversions-v3.png



Baum and Smith 2012, Figure 8.10

Rate matrix

The instantaneous rate of a given substitutions

Q - Rate matrix

Substitution probability matrix

The probability of a given substitution occurring in a given interval (branch length). Because of reversals, there are an infinite number of histories that could have given rise to the particular substitution. Can be derived from the rate matrix.

P - Substitution probability matrix

Substitution probability matrix

Substitution probability matrix

Rate matrix

$$\boldsymbol{P}(v) = e^{\boldsymbol{Q}v}$$

Branch length

This is called matrix exponentiation

F81 model

Q - Rate matrix

		To:			
		A (freq = π_A)	C (freq = π_C)	G (freq = π_G)	T (freq = π_T)
From:	A (freq = π_A)	$-m(\pi_C + \pi_G + \pi_T)$	$\pi_C m$	$\pi_G m$	$\pi_T m$
	C (freq = π_C)	$\pi_A m$	$-m(\pi_A + \pi_G + \pi_T)$	$\pi_G m$	$\pi_T m$
	G (freq = π_G)	$\pi_A m$	$\pi_C m$	$-m(\pi_A + \pi_C + \pi_T)$	$\pi_T m$
	T (freq = π_T)	$\pi_A m$	$\pi_C m$	$\pi_G m$	$-m(\pi_A + \pi_C + \pi_G)$

P - Substitution probability matrix

		To:			
		A	C	G	T
From:	A	$\pi_A + (1 - \pi_A)e^{-mt}$	$\pi_C(1 - e^{-mt})$	$\pi_G(1 - e^{-mt})$	$\pi_T(1 - e^{-mt})$
	C	$\pi_A(1 - e^{-mt})$	$\pi_C + (1 - \pi_C)e^{-mt}$	$\pi_G(1 - e^{-mt})$	$\pi_T(1 - e^{-mt})$
	G	$\pi_A(1 - e^{-mt})$	$\pi_C(1 - e^{-mt})$	$\pi_G + (1 - \pi_G)e^{-mt}$	$\pi_T(1 - e^{-mt})$
	T	$\pi_A(1 - e^{-mt})$	$\pi_C(1 - e^{-mt})$	$\pi_G(1 - e^{-mt})$	$\pi_T + (1 - \pi_T)e^{-mt}$

Baum and Smith 2012, Figures 8.7, 8.8

F81 model

As the branch length goes to 0, **P** becomes a diagonal matrix

	A	C	G	T
A	1	0	0	0
C	0	1	0	0
G	0	0	1	0
T	0	0	0	1

As the branch length goes to infinity, the rows become the equilibrium base frequencies

P

		To:			
		A	C	G	T
From:	A	$\pi_A + (1 - \pi_A)e^{-mt}$	$\pi_C(1 - e^{-mt})$	$\pi_G(1 - e^{-mt})$	$\pi_T(1 - e^{-mt})$
	C	$\pi_A(1 - e^{-mt})$	$\pi_C + (1 - \pi_C)e^{-mt}$	$\pi_G(1 - e^{-mt})$	$\pi_T(1 - e^{-mt})$
	G	$\pi_A(1 - e^{-mt})$	$\pi_C(1 - e^{-mt})$	$\pi_G + (1 - \pi_G)e^{-mt}$	$\pi_T(1 - e^{-mt})$
	T	$\pi_A(1 - e^{-mt})$	$\pi_C(1 - e^{-mt})$	$\pi_G(1 - e^{-mt})$	$\pi_T + (1 - \pi_T)e^{-mt}$

Baum and Smith 2012, Figure 8.8