

# Quantifying selection in coding sequences

1.) selectively constrained:

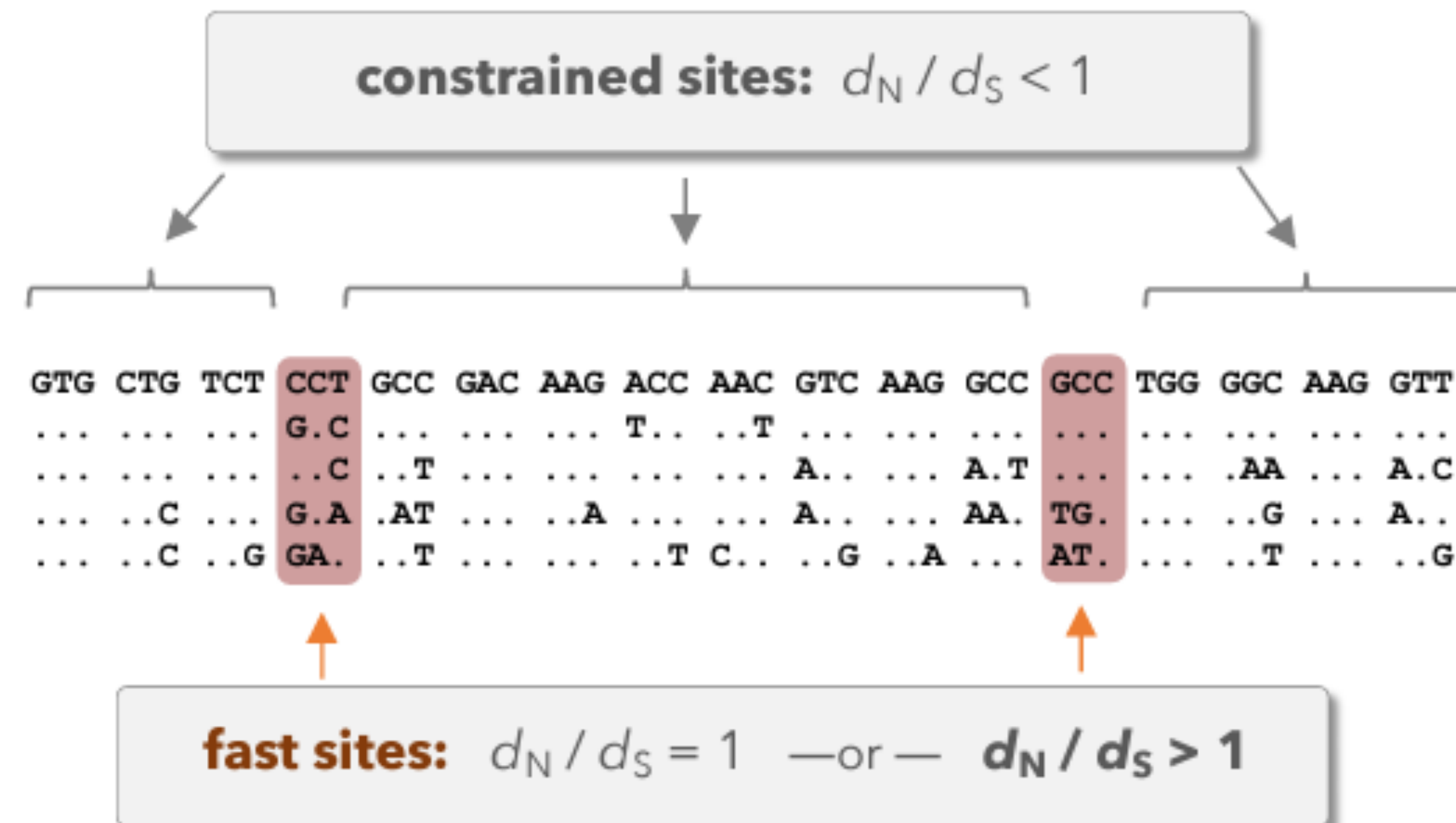
$$d_N / d_S < 1$$

2.) strictly neutral:

$$d_N / d_S = 1$$

3.) adaptive evolution:

$$d_N / d_S > 1$$



# Codon substitution models

the instantaneous rate matrix,  $Q$ , is very big:  $61 \times 61$

phenomenological codon models: just a few parameters are needed to cover the 3721 changes between codons!

From codon below:	to codon below:						
	TTT (Phe)	TTC (Phe)	TTA (Leu)	TTG (Leu)	CTT (Leu)	CTC (Leu)	GGG (Gly)
TTT (Phe)	—	$\kappa\pi_{TTC}$	$\omega\pi_{TTA}$	$\omega\pi_{TTG}$	$\omega\kappa\pi_{TTT}$	0	0
TTC (Phe)	$\kappa\pi_{TTT}$	—	$\omega\pi_{TTA}$	$\omega\pi_{TTG}$	0	$\omega\kappa\pi_{CTC}$	0
TTA (Leu)	$\omega\pi_{TTT}$	$\omega\pi_{TTC}$	—		0	0	0
TTG (Leu)	$\omega\pi_{TTT}$	$\omega\pi_{TTC}$	$\kappa\pi_{TTA}$	—	0	0	0
CTT (Leu)	$\omega\kappa\pi_{TTT}$	0	0	0	—	$\kappa\pi_{CTC}$	0
CTC (Leu)	0	$\omega\kappa\pi_{TTC}$	0	0	$\kappa\pi_{TTT}$	—	0
⋮ ↓	⋮ ↓	⋮ ↓	⋮ ↓	⋮ ↓	⋮ ↓	⋮ ↓	⋮ ↘
GGG (Gly)	0	0	0	0	0	0	—

\* This is equivalent to the codon model of Goldman and Yang (1994). Parameter  $\omega$  is the ratio  $d_N/d_S$ ,  $\kappa$  is the transition/transversion rate ratio, and  $\pi_i$  is the equilibrium frequency of the target codon ( $i$ ).