



**Computing synonymous and non-synonymous  
sites for GAA (Glutamic Acid)**

**8/3 non-synonymous sites (or 7/3 + 1/3 “stop” site)**

**1/3 synonymous sites**

	1	2	3
Start codon:	G	A	A
A	AAA Lysine	*	*
C	CAA Glutamine	GCA Alanine	GAC Aspartic Acid
G	*	GGA Glycine	GAG Glutamic Acid
T	TAA Stop	GTA Valine	GAT Aspartic Acid
Synonymous changes	0	0	1
Non-synonymous changes	3	3	2
Synonymous sites	0	0	1/3
Non-synonymous sites	1	1	2/3

# Computing synonymous and non-synonymous sites for GAA (Glutamic Acid)

	1	2	3
Start codon:	G	A	A
A	AAA Lysine	*	*
C	CAA Glutamine	GCA Alanine	GAC Aspartic Acid
G	*	GGA Glycine	GAG Glutamic Acid
T	TAA Stop	GTA Valine	GAT Aspartic Acid
Synonymous changes	0	0	1
Non-synonymous changes	3	3	2
Synonymous sites	0	0	1/3
Non-synonymous sites	1	1	2/3

8/3 non-synonymous sites (or 7/3 + 1/3 “stop” site)  
1/3 synonymous sites

# Nei-Gojobori dN/dS estimate (NG86)

- For each codon  $C$  we define  $ES(C)$  and  $EN(C)$  - the numbers of synonymous and non-synonymous *sites* of a codon
  - e.g.,  $ES(GAA) = 1/3$ ,  $EN(GAA) = 8/3$ .
- May also define them as fractions of substitutions that do not lead to stop codons,
  - e.g.,  $ES(GAA) = 1/3$ ,  $EN(GAA) = 7/3$ .
- The sum of  $ES$  and  $EN$  over all codons in a sequence gives an estimate of expected synonymous and non-synonymous **sites** in a sequence.
- For two sequences (the target of the original method), we average  $ES(C)$  and  $EN(C)$  at each site.
- $EN/ES$  is thus the ***expected ratio of non-synonymous to synonymous substitutions counts under neutral evolution***

Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions

M. Nei and T. Gojobori

Mol. Biol. Evol. 3 418--426 (1986)

>5,300 citations