



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

Aminoacid	Codons	Redundancy
Alanine	GC*	4
Cysteine	TGC,TGT	2
Aspartic Acid	GAC,GAT	2
Glutamic Acid	GAA,GAG	2
Phenylalanine	TTC,TTT	2
Glycine	GG*	4
Histidine	CAC,CAT	2
Isoleucine	ATA,ATC,ATT	3
Lysine	AAA,AAG	2
Leucine	CT*,TTA,TTG	6
Methionine	ATG	1
Aspargine	AAC,AAT	2
Proline	CC*	4
Glutamine	CAA,CAG	2
Arginine	AGA,AGG,CG*	6
Serine	AGC,AGT,TC*	6
Threonine	AC*	4
Valine	GT*	4
Tryptophan	TGG	1
Tyrosine	TAC,TAT	2
Stop	TAA,TAG,TGA	3

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

**1/3 synonymous sites**

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

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

Start codon:	G	A	A
Site/Change to	1	2	3
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

Aminoacid	Codons	Redundancy
Alanine	GC*	4
Cysteine	TGC,TGT	2
Aspartic Acid	GAC,GAT	2
Glutamic Acid	GAA,GAG	2
Phenylalanine	TTC,TTT	2
Glycine	GG*	4
Histidine	CAC,CAT	2
Isoleucine	ATA,ATC,ATT	3
Lysine	AAA,AAG	2
Leucine	CT*,TTA,TTG	6
Methionine	ATG	1
Asparagine	AAC,AAT	2
Proline	CC*	4
Glutamine	CAA,CAG	2
Arginine	AGA,AGG,CG*	6
Serine	AGC,AGT,TC*	6
Threonine	AC*	4
Valine	GT*	4
Tryptophan	TGG	1
Tyrosine	TAC,TAT	2
Stop	TAA,TAG,TGA	3

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

# Nei-Gojobori dN/dS estimate (NG86)

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

M. Nei and T. Gojobori

*Mol. Biol. Evol.* **3** 418–426 (1986)

- 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***