Neutral expectation

- A random mutation is **~3 times more likely to be non-synonymous that synonymous**, depending on the variety of factors, such as codon composition, transition/transversion ratios, etc.
- We need to **estimate** the proportion of random mutations that are synonymous, and use it as a reference to compute **dS**.
- In early literature, these quantities were codified as synonymous and non-synonymous “sites” and/or mutational opportunity.
- As a very crude approximation (assuming that third positions ~ synonymous), each codon has 1 synonymous and 2 non-synonymous sites.

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**