

Universal genetic code

This genetic code has 61 sense (non-termination) codons

Substitution types

| | Synonymous | | | Non-synonymous | | | To a stop codon |
|---------------|-------------|---------------|-------|----------------|---------------|-------|-----------------|
| | Transitions | Transversions | Total | Transitions | Transversions | Total | Total |
| 1st position: | 8 | 0 | 8 | 140 | 26 | 166 | 9 |
| 2nd position: | 0 | 0 | 0 | 148 | 28 | 176 | 7 |
| 3rd position: | 58 | 68 | 126 | 2 | 48 | 50 | 7 |
| ----- | | | | | | | |
| Total | 66 | 68 | 134 | 290 | 102 | 392 | 23 |

- Approximately 3:1 (392 N : 134 S) ratio when mutations are generated and **fixed** completely at random
- Non-random distribution over codon positions
 - **All** second position mutations are non-synonymous
 - **Most** (but not all) synonymous mutations are confined to the third position

Neutral expectation

- A random mutation is **~3 times more likely to be non-synonymous than 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.