Estimating dS and dN

Consider two **aligned homologous** sequences

	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6
DNA	ACA	ATA	A**T**C	TTT	AA**T**	CAA
AA	T	I	**I**	F	N	Q
DNA	ACA	ATA	A**C**C	TTT	AA**C**	CAA
AA	T	I	**T**	F	N	Q

Can one claim that `dN/dS` = 1, because there is **one** synonymous and **one** non-synonymous substitution?

Estimating dS and dN

Consider two **aligned homologous** sequences

	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6
DNA	ACA	ATA	A**T**C	TTT	AA**T**	CAA
AA	T	I	**I**	F	N	Q
DNA	ACA	ATA	A**C**C	TTT	AA**C**	CAA
AA	T	I	**T**	F	N	Q

Can one claim that dN/dS = 1, because there is **one** synonymous and **one** non-synonymous substitution?

```
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