Alignment-wide estimates

- Using standard MLE approaches it is straightforward to obtain point estimates of `dN/dS := ` β/α

- Can also easily test whether or not `dN/dS > 1`, or `< 1` using the likelihood ratio test (LRT)

- Codon models also support the concepts of synonymous and non-synonymous distances between sequences using standard properties of Markov processes (exponentially distributed waiting times)

$$E[subs] = -\sum_i \pi_i \hat{q}_{ii}, \quad E[subs] = E[syn] + E[nonsyn] = -\sum_i \pi_i \hat{q}_{ii}^{s} - \sum_i \pi_i \hat{q}_{ii}^{ns}.$$

Three example datasets

- **West Nile Virus NS3 protein**

  - An interesting case study of how positive selection detection methods lead to testable hypotheses for function discovery

  - Brault et al 2007, *A single positively selected West Nile viral mutation confers increased virogenesis in American crows*

- **HIV-1 transmission pair**

  - Partial *env* sequences from two epidemiologically linked individuals

  - An example of multiple selective environments (source, recipient, transmission)

- **SARS-CoV-2 Spike**

  - Full length spike sequences chosen to represent viral diversity (circa mid 2021)

  - Good example for analyzing selection in population samples with many "dead-end" intra-host variants