---

## NUCLEOTIDE SUBSTITUTIONS (2)

**Jukes-Cantor 2-Parameter Model**

$$A \xleftrightarrow{\ \alpha\ } G$$

A — α → G, β vertical, β diagonal, C ↔ T

---
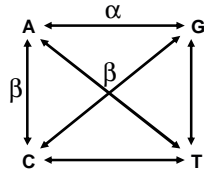
**Jukes-Cantor 2-Parameter Model**

$P_{AA(t)} = P(A, t \mid A, 0)$

$P_{AA(1)} = 1 - \alpha - 2\beta$

$P_{AA(2)} = (1 - \alpha - 2\beta)P_{AA(1)} + \beta P_{TA(1)} + \beta P_{CA(1)} + \alpha P_{GA(1)}$

$P_{AA(t+1)} = (1 - \alpha - 2\beta)P_{AA(t)} + \beta P_{TA(t)} + \beta P_{CA(t)} + \alpha P_{GA(t)}$

$\Delta P_{AA(t)} = -(\alpha + 2\beta) P_{AA(t)} + \beta P_{TA(t)} + \beta P_{CA(t)} + \alpha P_{GA(t)}$

$P_{AA(t)} = (1 + e^{-4\beta t} + 2e^{-2(\alpha + \beta)t}) / 4$

$X_{(t)} = (1 + e^{-4\beta t} + 2e^{-2(\alpha + \beta)t}) / 4$

$Y_{(t)} = (1 + e^{-4\beta t} - 2e^{-2(\alpha + \beta)t}) / 4$

$Z_{(t)} = (1 - e^{-4\beta t}) / 4$

---

## NUCLEOTIDE SUBSTITUTIONS BETWEEN SEQUENCES

**usually is inferred from pairwise comparisons**

**Hamming Distance**
  **divergence**
  **n / N**
  **[nucleotide site$^{-1}$]**
  **if small,**
      **then observed = actual**
  **if large,**
      **then observed < actual**

---

**Noncoding Sequences**

$I_{(t)} = P^2_{AA(t)} + P^2_{AT(t)} + P^2_{AC(t)} + P^2_{AG(t)}$

$I_{(t)} = (1 + 3e^{-8\alpha t}) / 4$

$\rho = 1 - I_{(t)} = 3 (1 - e^{-8\alpha t}) / 4$

$8\alpha t = -Log[1 - (4\rho / 3)]$

$K = -3Log[1 - (4\rho / 3)] / 4$

$K = Log[(1 - 2P - 2Q)^{-1}] / 2 + Log[(1 - 2Q)^{-1}] / 4$