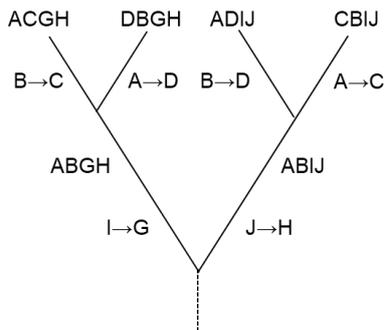


### W3.2: PAM の求め方

シトクロムやグロビンなど近縁のタンパク質を集め、分子進化系統樹を作成し、置換の数をカウントする。



1PAM をマルコフ過程の 1 ステップとし、その間にアミノ酸  $i$  がアミノ酸  $j$  に置換された数を  $m_{ij}$  とすると、アミノ酸  $i$  が他のアミノ酸に置換された数は

$$m_i = \sum_{j \neq i} m_{ij}$$

アミノ酸  $i$  が別なアミノ酸に置換される場合の、アミノ酸  $i$  がアミノ酸  $j$  に置換される割合は

$$a_{ij} = \frac{m_{ij}}{m_i}$$

アミノ酸  $i$  が別なアミノ酸に置換される割合は、 $i$  によらないとして、 $c$  とおく。

アミノ酸  $i$  がアミノ酸  $j$  に置換される確率は、

$$M_{ij} = ca_{ij}$$

アミノ酸  $i$  の出現確率を  $q_i$  とするとき、1PAM で置換されるアミノ酸の割合

$$\sum_i q_i \sum_{j \neq i} M_{ij} = c \sum_i \sum_{j \neq i} q_i a_{ij} = 0.01$$

これを満たす  $c$  の値は以下のようになる。

$$c = \frac{0.01}{\sum_i \sum_{j \neq i} q_i a_{ij}}$$

1PAM の推移確率行列を  $M = (M_{ij})$  とすると、 $k$ PAM の推移確率行列は  $M^k = (M^k_{ij})$

PAM $k$  マトリックスのスコアは以下の式で与えられる。

$$s(i, j) = \log \left( \frac{p_{ij}}{q_i q_j} \right)$$

ここで、 $p_{ij}$  はアミノ酸  $i$  があってアミノ酸  $j$  に置換される確率なので  $p_{ij} = q_i M^k_{ij}$