

Bioinformatics

Course Code

Chirayu Gupta

1 PAM Paper

- Relationship between sequences are important to study evolutionary parameters
- Accepted point mutation is a replacement of one amino acid with another accepted by natural selection
- There are 20 amino acids in nature(I think)
- Replacement of X by Y is same as Y by X
- Non transitive nature of accepted point mutations
- Pam matrices value $M_{ij} = \frac{\lambda m_j A_{ij}}{\sum_i A_{ij}}$
- λ is prop constant. A_{ij} is element of the point mutation matrix m_j is the mutability of jth amino acid
- Total must sum to 1 so diagonal elements are $M_{jj} = 1 - \lambda m_j$
- PAM describes a unit of evolutionary change
- Applying PAM matrix N times gives matrix after N such evolutionary changes
- f_i is the probability that i will occur in the second sequence by chance(I think its the sum of elements in the corresponding row along with some normalization)
- Percentage of amino acids observed to change on average $100(1 - \sum_i f_i M_{ii})$
- Relatedness Odds Matrix $R_{ij} = \frac{M_{ij}}{f_i}$ is a matrix giving probability of replacement per occurrence of i per occurrence of j
- PAM 250 is relatedness odds matrix. In log scale

2 Dynamic Programming Overview

Why need an algorithm?

Complexity of scoring all possible sequences is $O(2^{2N})$.