SEQUENCE ANALYSIS OF DNA H1N1 VIRUS USING SUPER PAIRWISE ALIGNMENT METHODS

By: Alfi Yusrotis Zakiyyah
Student Identity Number: 12092010101
Supervisor: Prof. Dr. M. Isa Irawan, MT
Co-Supervisor: Dr. rer. nat. Ir. Maya Shovitri, M.Si

ABSTRACT

The most basic sequence biology analysis task is to know relation between sequences. There are three kinds of sequence biology that DNA sequence, RNA sequence and protein sequence. The sequence analysis of pairwise sequences is usually done by first aligning the sequences (or parts of them) and then deciding whether that alignment is more likely to have occurred because the sequences are related, or just by chance. The key issues are what sorts of alignment should be considered, the scoring system used to rank alignments and the algorithm used to find optimal (or good) scoring alignments. The other problem is the statistical methods used to evaluate the significance of an alignment score.

One of methods in sequence alignment is dynamic programming. The widely used alignment, dynamic programming though generating optimal alignment, takes too much due to its high computation complexity $O(N^2)$. A majority of sequence alignment software utilizes dynamic programming, such as global Needleman Wunsch and local alignment use Smith Waterman. Both methods were a classical algorithm in sequence alignment and it has disadvantage of which one is the speed of computation. To solve this problem, Shen et al found a new methods that is a Super Pairwise alignment. This methods combines the analysis of the methods combinatorics and probability (Shen et al, 2002).

Identification of new virus can be analyzed by homologous. On the problem of the identification of disease, DNA virus mutates so that it may give rise to new viruses. The case H1N1 is one example of mutation. The H1N1 virus mutates quickly enough. Hemaglutinin virus transmission from aspartic acid (D) to Glynine (G) on the line 222. Super Pairwise Alignment can be applied to get
optimal alignment from this study found that sequences GU451262 and GU451280 have 99.56% homolog.

Keyword: DNA H1N1, Sequence Analysis, Super Pairwise Alignment