

The concepts of statistical alignment and local alignment has not been combined at present. Doing this could have several practical advantages.

Substitutions, insertions and deletions are random evolutionary events that should be modelled by a stochastic process. For a long time this was totally ignored for insertions and deletions, which were treated by an ad hoc method. Using a stochastic model for these events defines a class of methods called statistical alignment.

The first model for statistical alignment was proposed for in 1986 by Bishop and Thompson. In 1991 the model was proposed that provides the basis for most statistical alignment today by Thorne, Kishino and Felsenstein.

Although statistical alignment for a long time had the reputation of being very difficult, it can actually be solved by an algorithm that is very similar to the traditional optimisation alignment.

The most famous method for local alignment were published by Smith and Waterman in 1981 (SW81). It is based on a similarity maximizing method proposed by Needleman and Wunch in 1970 that aligned complete sequences. Smith and Waterman added the possibility of initiating and stopping the regions of interest within a longer sequence by a simple modification to the Needleman-Wunch recursions. SW81 is too slow for database searches and genomic comparisons and has for practical purposes been superseded by the BLAST family of programs (Altschul et al., 1990).

The advantage of a statistical alignment approach to local alignment would be that it would give a probability distribution of the borders of the homologous regions that would depend on the content of the sequences. If the homologous sequences were very similar, then the borders would be very well determined, if they were very different the border positions would get quite a wide distribution.