Abstract

Database sequencing applications such as sequence comparison process large size of sequences and considered to be high consumers of computation time. Heuristic algorithms have the problem of sensitivity since they trim the search and miss unexpected but important homologies. Traditional optimal methods apply these applications on the whole database to find the most matched sequences but this consumes very high computation time. We introduce novel and efficient technique which optimizes the performance of the database sequencing applications by reducing the computation time of finding the optimal matched sequence in a large database. Our technique uses our new similarity functions which are based on the mathematical parameters: frequency and mean of the codes of each sequence in the database. Using our technique, we explicitly accelerate the database sequencing applications by 60% in comparison to the traditional known methods.

References

Journal of Molecular Biology. Pages 196197, 1981
- Talal Bonny, M. A. Z. and Salama, K. N. An adaptive hybrid multiprocessor technique for bioinformatics sequence alignment. In the 5th Cairo International Conference on Biomedical Engineering. pages 112115, 2010

- DNA Data Bank of Japan (ddbj): http://www.ddbj.nig.ac.jp/

**Index Terms**

Computer Science

Databases

**Keywords**

Database sequence comparison Sequence Analysis