Abstract

Evolutionary modeling applications are the best way to provide full information to support in-depth understanding of evaluation of organisms. These applications mainly depend on identifying the evolutionary history of existing organisms and understanding the relations between them, which is possible through the deep analysis of their biological sequences. Multiple Sequence Alignment (MSA) is considered an important tool in such applications, where it gives an accurate representation of the relations between different biological sequences. In literature, many efforts have been put into presenting a new MSA algorithm or even improving existing ones. However, little efforts on optimizing parallel MSA algorithms have been done. Nowadays, large datasets become a reality, and big data become a primary challenge in various fields, which should be also a new milestone for new bioinformatics algorithms. This survey presents four different parallel MSA algorithms, TCoffee, MAFFT, MSAProbs, and M2Align. We provide a detailed discussion of each algorithm including its strengths, weaknesses, and implementation details and the effectiveness of its parallel implementation compared to the other algorithms, taking into account the MSA accuracy on two different
datasets, BAiiBASE and OXBench.

References

17. Kazutaka Katoh and Daron M Standley. Mafft multiple sequence alignment software
Index Terms

Computer Science

Algorithms

Keywords

Bioinformatics; Multiple Sequence Alignment; Parallel Processing; Multicore Systems