

References

- Bernardi, G. (2000). Isochores and the evolutionary genomics of vertebrates. *Gene*, 241, 3-17.
- Boor, J. L. (1999). Animal mitochondrial genomes. *Nucleotide Acids Research*, 27, 1767 - 1780.
- Chinnery, P. F., Howell, N., Andrews, R. M., & Turnbull, D. M. (1999). Clinical mitochondrial genetics. *J Med Genet*, 36, 425-436 .
- Daeyaert, F., Moereels, H., & Lewi, P. (1998). Classification and identification of proteins by means of common and specific amino acid n-tuples in unaligned sequences. *Compute Methods and Programs in Biomedicine*, 56, 221 - 233.
- Dai, Q., Yang, Y., & Wang, T. (2008). Markov model plus k-word distributions: a synergy that produces novel statistical measures for sequence comparison. *Bioinformatics*, 24, 2296 - 2302.
- Domazet-Lošo, & Mirjana. (2010). Algorithms for efficient alignment-free sequence comparison. (Unpubished doctoral dissertation). University of Zagreb, Zagreb.
- Domazet-Loso, M., & Haubold, B. (2009). Efficient estimation of pairwise distances between genomes. *Bioinformatics*, 25, 3221 - 3227.
- Edgar, R. C. (2004). Muscle: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, 32, 1792 - 1797.
- Haubold, B., Pfaffelhuber, P., Domazet-Loso, M., & Wiehe, T. (2009). Estimating mutation distances from unaligned genomes. *Comput Biol.*, 16, 1487-500.
- Haubold, B., Reed, F. A., & Pfaffelhuber, P. (2010). Alignment-free estimation of nucleotide diversity. *Bioinformatics*, 27, 449 - 455.
- Henikoff, S., & Henikoff, J. (1992). Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci.*, 89, 10915 - 10919.

- Huang, X., & Miller, W. (1991). A time-efficient, linear space local similarity algorithm. *Amer. Math. Soc.*, 12, 337-357.
- Jameson, D. *et al.* (2003). OGRE: a relational database for comparative analysis of mitochondrial genomes. *Nucleic Acids Research*, 31, 201-206
- Juang, W. S., & Su, S. F. (2008). Multiple sequence alignment using modified dynamic programming and particle swarm optimization. *Chinese Institute of Engineers*, 31, 659-673.
- Kantorovitz, M. R., Robinson, G. E., & Sinha, S. (2007). A statistical method for alignment-free comparison of regulatory sequences. *Bioinformatics*, 23, 249-255.
- Knight, R. *et al.* (2007) PyCogent: a toolkit for making sense from sequence. *Genome Biol.*, 8, R171
- Kong, A. D. (2002). A high resolution recombination map of the human genome. *Nat. Genet.*, 31, 241–247.
- Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., & Zody, M. C. (2001). Initial sequencing and analysis of the human genome. *Nature*, 409, 860–921.
- Li, K. B. (2003). ClustalW-MPI: ClustalW analysis using distributed and parallel computing. *Bioinformatics*, 19, 1585–1586.
- Lu, G., Wang, Z., Jones, A. M., & Moriyama, E. N. (2009). 7TMRmine: A Web server for hierarchical mining of 7TMR proteins. *BioMed Central*, 10, 275 - 290.
- Luo, H., Sun, Z., Arndt, W., Shi, J., Friedman, R., & Tang, J. (2009). Gene order phylogeny and the evolution of methanogens. *Plos One*, 4, e6069.
- Marcet-Houben, M., & Gabaldon, T. (2011). TreeKO: A duplication-aware algorithm for the comparison of phylogenetic trees. *Nucleic Acid Res.*, 39(10): e66.
- Meunier, J., & Duret, L. (2004). Recombination drives the evolution of gc-content in the human genome. *Molecular Biology and Evolution*, 21, 984 - 990.

- Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins . *J. Molecular Biology*, 48, 443-453.
- Notredame, C., Higgins, D. G., & Heringa, J. (2000). T-Coffee: a novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology*, 302, 205 - 17.
- Notredame, C., Holm, L., & Higgins, D. G. (1998). Coffee: An objective function for multiple sequence alignments. *Bioinformatics*, 14, 407-422.
- Pearson, W., & Wood, T. (2001). Statistical significance in biological sequence comparison. In D. J. Balding, M. Bishop, & C. Cannings, *Handbook of Statistical Genetics* (pp. 39-66). Virginia: Wiley.
- Reinert, G., Chew, D., Sun, F., & Waterman, M. S. (2009). Alignment-free sequence comparison (I) :Statistics and Power. *Computational Biology*, 16, 1615–1634.
- Robert, C. E., & Batzoglou, S. (2006). Multiple sequence alignment. *Current Opinion in Structural Biology* , 16, 368–373 .
- Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). Clustal w: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*, 22, 4673-4680.
- Saitou, N., & Nei, M. (1987). The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees. *Mol. Biol. Evol*, 4, 406-425.
- Sterling, T. (2001). An introduction to PC clusters for high performance computing. *Int. J. High. Perform. C.*, 15, 92-101.
- Vinga, S., & Almeida, J. (2003). Alignment-free sequence comparison—a review. *Bioinformatics*, 19, 513 - 523.
- Wang, L., & Jiang, T. (1994). On the complexity of multiple sequence alignment. *Computational Biology*, 1, 337 - 348.
- Watanabe, Y., Fujiyama, A., Ichiba, Y., Hattori, M., & Yada, Y. (2002). Chromosome-wide assessment of replication timing for human chromosomes

11q and 21q: disease-related genes in timing-switch regions. *Hum. Mol. Genet.*, 11, 13 - 21.

Zhang, X., & Kahveci, T. (2006). A new approach for alignment of multiple proteins. *Pacific Symposium on Biocomputing*, 11, 339-350.