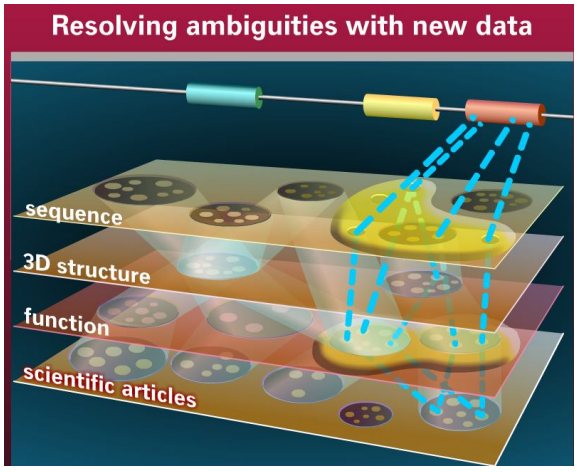
Next generation bioinformatics for biomedical research

While traditional bioinformatics has evolved from simple data management to data-interpretation, the emphasis today has shifted to high-throughput data collection, personal medicine and the analysis of complex systems. This tendency is accompanied by an unprecedented development of new computer architectures and cloud computing that bring the power of supercomputers within arm's reach of bench scientists and clinical practicioners. At the same time, bionic devices and on-line diagnostic tools open up new areas of applications.



In this fast evolving scene of new technologies, integrating heterogeneous bioinformatics data is perhaps one of the most challenging tasks. Databases increase both in volume and in complexity, and public resources available on the Internet can not cope with a growing number of user groups, especially medical and industrial users concerned with data confidentiality. On the other hand many, if not most biomolecular mechanisms that translate the human genomic information into phenotypes are not known and as a consequence, most of the molecular and cellular data cannot be interpreted in terms of biomedically relevant conclusions. While personalized diagnostics and cures are likely to remain a dominant trend, the temperate view suggests biomedical applications relying on the comparison of biomolecular sequences and/or on the already known biomolecular mechanisms may have even greater chances to enter clinical practice. Developing stand-alone tools for genome annotation, personalized medicine and high throughput technologies is especially important in the analysis of complex diseases such as neurological and psychiatric disorders.

Project Participants:

Zsolt Gelencsér, PhD. student
Áron Erdei, BSc student
Prof. Sándor Pongor, PI

Collaborators

Prof. Mária Judit Molnár, Clinical and Research Centre for Molecular Neurology Semmelweiss University, Budapest, Hungary
Prof. Frank Eisenhaber, Bioinformatics Institute A*STAR Singapore
Dr. Michael P. Myers, International Centre for Genetic Engineering and Biotechnology, Trieste, Italy,
Roberto Vera, PhD student International Centre for Genetic Engineering and Biotechnology, Trieste, Italy

References

265. Kuznetsov, V.; Lee, H.K.; Maurer-Stroh, S.;Molnár, M.J.; Pongor, S.; Eisenhaber, B.; Eisenhaber, F. (2013). How bioinformatics influences health informatics: Usage of biomolecular sequences, expression profiles and automated microscopic image analyses for clinical needs and public health. *Health Information Science and Systems*. 1.

263. Vera, R.; Perez-Riverol, Y.; Perez, S.; Ligeti, B.; Kertész-Farkas, A.; Pongor, S. (2013). JBioWH: an open-source Java framework for bioinformatics data integration. *Database*. 2013.

231. Busa-Fekete, R.; Kertész-Farkas, A.; Kocsor, A.; Pongor, S. (2008). Balanced ROC (BaROC) analysis for portien classification. *Journal of Biochemical and Biophysical Methods*. 70(6), 1210-1214.

230. Busa-Fekete, R.; Kocsor, A.; Pongor, S. (2008). Tree-Based Algorithms for Protein Classification. *Computational Intelligence in Bioinformatics*, 165-182.

228. Kertesz-Farkas, A.; Dhir, S.; Sonego, P.; Pacurar, M.; Netoteia, S.; Nijveen, H.; Kuzniar, A.; Leunissen, J.A.; Kocsor, A.; Pongor, S. (2008). Benchmarking protein classification algorithms via supervised cross-validation. *Journal of Biochemical and Biophysical Methods*. 70(6), 1215-1223.

227. Kertész-Farkas, A.; Kocsor, A.; Pongor, S. (2008). The application of Data Compression-based Distances to Biological Sequences. *In: Frank Emmert-Streib. (Ed.) Information Theory and Statistical Learning,*, 73-88.

226. Kocsor, A.; Busa-Fekete, R.; Pongor, S. (2008). Protein classification based on propagation of unrooted binary trees. *Protein Pept Lett*. 15(5), 428-434.

225. Kuzniar, A.; Van Ham, R.C.; Pongor, S.; Leunissen, J.A. (2008). The quest for orthologs: finding the corresponding gene across genomes. *Trends Genet*. 24(11), 539-551.

223. Sonego, P.; Kocsor, A.; Pongor, S. (2008). ROC analysis: applications to the classification of biological sequences and 3D structures. *Brief Bioinform*. 9(3), 198-209.