

6. BIOINFORMATICS LABORATORY

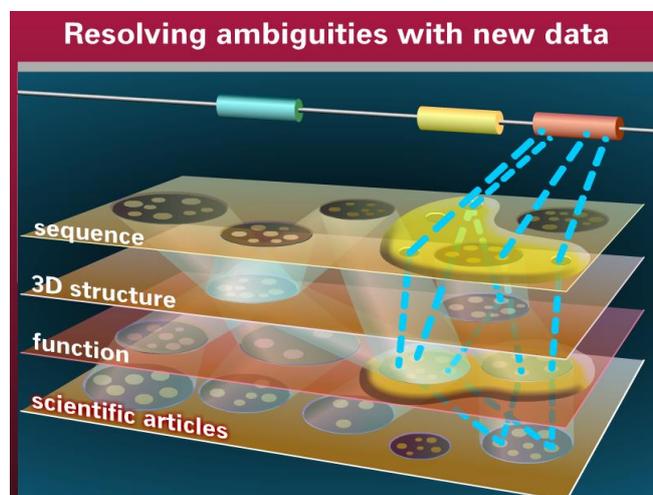
SÁNDOR PONGOR, Professor

Graduate students: BALÁZS LIGETI, JÁNOS JUHÁSZ, JÓZSEF HEGEDŰS, DÁNIEL VARGA

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 practitioners.

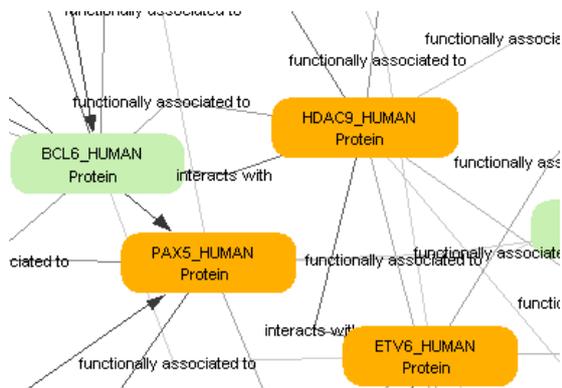
BIOINFORMATICS OF NEXT GENERATION SEQUENCING DATA

Interpretation of next generation sequencing data is a highly challenging task, because extracting information useful for medical researchers and practitioners requires advanced data mining methods, simultaneously applied to high throughput data linked to a continuously updated network of molecular databases and medical publications. In this fast evolving scene of new technologies, integrating heterogeneous data is perhaps one of the



most challenging tasks. Our current interest includes development of marker databases, new algorithms for analyzing metagenomics data, including medical hypothesis generation and prediction of useful drug combinations. . In this project we are concerned with developing new, hardware-accelerated pipelines for these purposes, making use of the in house knowledge of new computer architectures such as FPGA and GPU, which allow us to increase computer power by orders of magnitude. We are also collaborating with the Technical University of Graz where one of the strongest FPGA-based bioinformatics infrastructure of Central Europe is being built.

HIGH THROUGHPUT DATA/MINING OF PROTEOMICS DATA

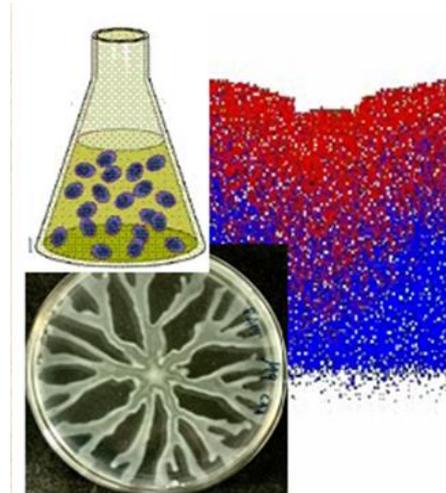


High throughput mass spectrometry analysis produces large amounts of noisy data that have to be filtered and preprocessed with computational tools before subjected to detailed analysis and interpretation. Our strategy uses principles borrowed from cognitive psychology for identifying network patterns in mass spectra. Namely, the human mind is able to capture holistic features in complex sensory inputs, and we trust that similar principles can be applied to

abstract data structures. The bioinformatics support of proteomics research is a central theme in our projects. We develop new tools capable of filtering and processing large data streams characteristic of high throughput analysis workflows.

MICROBIOME BIOINFORMATICS: BACTERIAL COMMUNICATION NETWORKS

Microbial communities play fundamental roles in health and disease as well as the stability of the ecosystem. A better understanding of these systems may provide insights into the mechanisms of infections, epidemics as well as environmental and social processes. Our group uses bioinformatics tools as well as agent based models to understand how signals contribute to colonization and infection. We showed that signal sharing allows several bacterial species to cross barriers that the single species can not which may provide important clues to polymicrobial diseases.



COLLABORATIONS:

- Prof. Balázs Györfy, Semmelweis University and Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest, Hungary
- Prof. Attila Kertész-Farkas, National Research University Higher School of Economics, Moscow, Russia
- Prof. Ines Mandic-Mulec, University of Ljubljana, Ljubljana, Slovenia

- Prof. Dóra Szabó, Semmelweis University, Budapest, Hungary
- Prof. Christoph Sensen, Graz University of Technology, Graz, Austria
- Prof. Vittorio Venturi, International Centre for Genetic Engineering and Biotechnology, Trieste, Italy

REFERENCES

- [1] Marini, B.; Kertész-Farkas, A.; Ali, H.; Lucic, B.; Lisek, K.; Manganaro, L.; Pongor, S.; Luzzati, R.; Recchia, A.; Mavilio, F.; Giacca, M.; Lusic, M. (2015). Nuclear architecture favors HIV-1 integration into a subset of active genes at the periphery of CD4+ T-cell nuclei. *Nature*. 14:227-31
- [2] Juhász, J.; Kertész-Farkas, A.; Szabó, D.; Pongor, S. (2014). Emergence of collective territorial defense in bacterial communities: Horizontal gene transfer can stabilize microbiomes. *PlosOne*. 9 (4), e95511.
- [3] Kertész-Farkas, A.; Reiz, B.; Vera, R.; Myers, M.P.; Pongor, S. (2013). PTMTreeSearch: a Novel Two-Stage Tree Search Algorithm with Pruning Rules for the Identification of Post-Translational Modification of Proteins in MS/MS Spectra. *Bioinformatics*. 30 (2), 234-41
- [4] 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, bat051
- [5] 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
- [6] Csermely, P.; Agoston, V.; Pongor, S. (2005). The efficiency of multi-target drugs: the network approach might help drug design. *Trends Pharmacol Sci*. 26(4), 178-182.