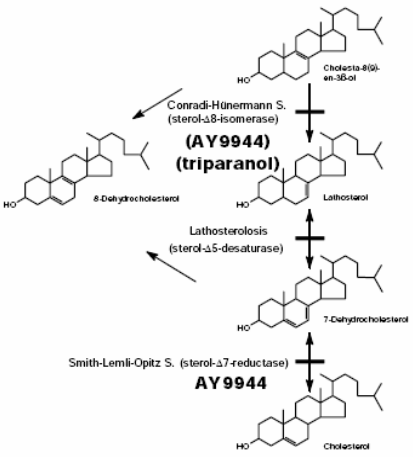
Analysis of gene expression data on metabolic networks

Anna-Lena Kranz[1], Marcus Oswald[3], Thorsten Bonato[3], Hanna Seitz[3], Gerhard Reinelt[3], Heiko Runz[4], Johannes Zschocke[4], Roland Eils[1,2] and Rainer König[1,2]

[1]Department of Bioinformatics and Functional Genomics, Institute for Pharmacy and Molecular Biotechnology, University of Heidelberg, 69120 Heidelberg, Germany. [2]Theoretical Bioinformatics, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany. [3] Institute of Computer Science, University of Heidelberg, 69120 Heidelberg, Germany. [4]Institute for Human Genetics, University of Heidelberg, 69120 Heidelberg, Germany.

When analysing gene expression data it is often not enough to examine single genes but rather to evaluate groups of genes [1]. As the modular structure of complex networks plays a critical role in functionality, the use of metabolic networks provides a more structural approach to the analysis of gene expression data. Determining significant expression patterns of topologically *associated* genes enables the identification of functionally relevant central components in the network with respect to different conditions of interest.

We invented a novel technique: The identification of sub-graphs in the metabolic network and thereby the grouping of reactions into parts with their major connections is achieved with clustering procedures on the network [2, 3]. Several clustering heuristics are existent that have been applied to the clustering of metabolic networks by our group, i.e. simulated annealing [4], a greedy and a consecutive ones heuristic [5, 6]. The overall modularity of the clustering as a quality control is determined and optimised [7].
Once the clusters are identified, the gene expression data is mapped onto the corresponding enzymatic reactions of its metabolic network. It is then possible to extract expression patterns for each cluster using a combinatorial approach that has been developed in our group. Thereby, values for every possible expression pattern of genes within a cluster are calculated. These values show essential differences between samples of different conditions and identify regions with a varying pattern between different states.



As a case study our approach is applied to gene expression data of HeLa cells under different concentrations of cholesterol. This has high clinical relevance as impaired sterol biosynthesis can cause severe human diseases (see Figure 1) [8]. Our results promise new insights into sterol biosynthesis when applied to the human metabolic network. With our method it is not only possible to detect broken enzymes but also to discover crucial imbalances of affected pathways in the cell.

Figure 1: Section of the cholesterol biosynthesis pathway showing interrupts causing the genetic diseases Lathosterolosis, Conradi-Hünemann Syndrome and Smith-Lemli-Opitz Syndrome. [8]

1. Rapaport, F., et al., *Classification of microarray data using gene networks.* BMC Bioinformatics, 2007. **8**: p. 35.
2. König, R., et al., *Discovering functional gene expression patterns in the metabolic network of Escherichia coli with wavelets transforms.* BMC Bioinformatics, 2006. **7**: p. 119.
3. König, R. and R. Eils, *Gene expression analysis on biochemical networks using the Potts spin model.* Bioinformatics, 2004. **20**(10): p. 1500-5.
4. Guimera, R. and L.A. Nunes Amaral, *Functional cartography of complex metabolic networks.* Nature, 2005. **433**(7028): p. 895-900.
5. Christof, T., M. Oswald, and G. Reinelt, *Consecutive Ones and a Betweenness Problem in Computational Biology*, in *Proceedings of the 6th International IPCO Conference on Integer Programming and Combinatorial Optimization*. 1998, Springer-Verlag.
6. Oswald, M. and G. Reinelt, *Polyhedral Aspects of the Consecutive Ones Problem*, in *Proceedings of the 6th Annual International Conference on Computing and Combinatorics*. 2000, Springer-Verlag.
7. Newman, M.E. and M. Girvan, *Finding and evaluating community structure in networks.* Phys Rev E Stat Nonlin Soft Matter Phys, 2004. **69**(2 Pt 2): p. 026113.
8. Haas, D., et al., *Abnormal sterol metabolism in holoprosencephaly: Studies in cultured lymphoblasts.* J Med Genet, 2007.