BP 21: High-Throughput Data and their Analysis

Time: Wednesday 18:15–19:15 Location: H43

BP 21.1 Wed 18:15 H43

Competitive DNA Hybridization: Experimental Results and Conclusions for Microarray Experiments — •Timo Mai[1], Wolfgang Michel[1], Philipp Baaske[2], Thomas Naiser[1], and Albrecht Ott[1] — [1]Physikalisches Institut, Universität Bayreuth, 95440 Bayreuth — [2]present address: Lehrstuhl für Experimentelle Physik - Biophysik, LMU München, 80539 München

We experimentally approach the complex multi-component hybridization to microarrays by a simple system: A two-component mixture in solution consisting of a perfect match (PM) and a mismatch (MM) oligonucleotide competing for a immobilized probe sequence. We first characterize the binding of PM and MM separately in single hybridization experiments. We deduce rate constants and confirm them with values extracted from analysis of the temperature dependent hybridization signal. From this we predict the time course of hybridization of PM and MM when in competition for surface binding sites and compare with our competitive hybridization experiments.

This is useful for understanding the process of MM displacement by the PM and attaining an estimate for the specificity and the detection limit of microarray experiments.

BP 21.2 Wed 18:30 H43

Optical Study of DNA surface hybridization reveals DNA surface density as a key parameter for interpretation of microarray data — •Wolfgang Michel, Timo Mai, Thomas Naiser, and Albrecht Ott — Experimental Physics 1, University Bayreuth, Germany

We investigate the kinetics of DNA hybridization reactions on glass substrates, where one 22mer strand (bound-DNA) is immobilized via phenylene-diisothiocyanate linker molecule on the substrate, the dye-labeled (Cy3) complementary strand (free-DNA) is in solution in a reaction chamber. We use total internal reflection fluorescence (TIRF) for surface detection of hybridization. As a new feature we perform a simultaneous real-time measurement of the change of free-DNA concentration in bulk parallel to the TIRF measurement. We observe that the free-DNA concentration decreases considerably during hybridization. We show how the standard Langmuir kinetics needs to be extended to take into account the change in bulk concentration and explain our experimental results. Connecting both measurements we can estimate the surface density of accessible, immobilized bound-DNA. We observe that the fluorescent signal from the surface ceases to be proportional to the number of dye-labeled molecules on the surface for surface-densities of hybridized molecules above 5*10ˆ11 molecules/cmˆ2. We discuss the implications with respect to DNA microarray detection.

BP 21.3 Wed 18:45 H43

Decomposing gene expression profiles using sparseness and nonnegativity via genetic optimization — Kurt Stadlthanner[1], •Elmar Lang[1], Ana-Maria Tomé[2], Carlos Puntonet[3], and Fabian Theis[4] — [1]Institute of Biophysics, University of Regensburg, Germany — [2]DETI/IEETA, Universidade de Aveiro, Portugal — [3]Dep. Arqitectura y Tecnología de Computadores, Universidad de Granada, Spain — [4]MPI for Dynamics and Self-Organisation, Göttingen, Germany

Nonnegative matrix factorization (NMF) has proven to be a useful tool for the analysis of nonnegative multivariate data. Gene expression profiles naturally conform to assumptions about data formats raised by NMF. However, its cost function is known to have a rather high indeterminacy concerning the component signals extracted. Hence we consider an extension of the NMF algorithm that provides unique solutions whenever the underlying component signals are sufficiently sparse. However, the resulting fitness function is discontinuous and exhibits many local minima, hence we use a genetic algorithm for its optimization. The algorithm is first applied to toy data in order to investigate its statistical properties. Application to a microarray data set related to Pseudo-Xanthoma Elasticum (PXE) then shows that the proposed algorithm performs superior when compared to standard methods with respect to the estimated PXE-related gene clusters.

BP 21.4 Wed 19:00 H43

Efficient dimension reduction of large-scale biomedical time-series — •Fabian Theis — MPI for Dynamics and Self-Organisation, Göttingen, Germany

Dimension reduction considers the question of removing a noise subspace from a larger multivariate signal. It is a key preprocessing step in contemporary biomedical data analysis for example of EEG or fMRI. Classically, a signal is differentiated from noise by having a higher variance, and algorithms such as principal component analysis (PCA) remove the low-variance components thereby failing to capture signals that are deteriorated by noise of similar or stronger power.

In order to perform noisy dimension reduction, either higher-order statistics of the data or additional information such as temporal structure may be used. The former methods assume i.i.d. signals, whereas the latter deal with the more realistic assumption of a multivariate 'colored' (in contrast to white) stochastic process. The proposed method, denoted by colored subspace analysis (CSA), distinguishes signal from noise by nontrivial autocovariances. The goal of CSA is to find such a projection onto a signal subspace of minimal dimension. We can prove that the signal subspace is unique, and an efficient algorithm can be proposed. Its complexity is in the order of twice the order of PCA, with an optional accuracy factor. The feasibility of the algorithm is illustrated when applied to fMRI. Independent of data dimension, a task-related subspace is robustly identified. In contrast, the necessary dimension of the PCA-based signal subspace is not invariant under increasing number of captured MRI frames.