Workshop

Statistical and dynamical models in biology and medicine

October 21-22, 2010, DKFZ Heidelberg

This workshop intends to bring together researchers who are interested in modeling and analysis of biological systems, bioinformatics, statistical methods, and systems biology, with applications in biology and medicine.

Jointly organized by the Gmds/IBS Working Groups 'Statistical methods in bioinformatics' and 'Mathematical models in medicine' (Tim Beissbarth, Universität Göttingen; Julien Gagneur, EMBL Heidelberg; Nicole Radde, Universität Stuttgart; Ingo Röder, TU Dresden)

Local organization at DKFZ: Christian Bender, Mark Zapatka, Axel Benner, Benedikt Brors, Holger Sültmann

Sponsored by

International Biometrical Society

Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie

Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e.V.

Program Overview

Thursday, October 21st, 2010

Session I: Dynamical Modeling (Chair: Ingo Roeder)

12:30-13:25	***Keynote: Olaf Wolkenhauer***	A Stem Cell Niche Dominance Theorem
13:25-13:45	Ingmar Glauche	Molecular decision making in embryonic and adult stem cells
13:45-14:05	Patrick Weber	Parameter Estimation and Identifiability of Biological Networks Using Relative Data
14:05-14:25	Daniela Schittler	Switch models for cell differentiation: Bifurcation analysis reveals distinct switching properties
14:25-16:00	***Poster session I + coffee***	

Session II: Network reconstruction (Chair: Tim Beißbarth)

16:00-16:55	***Keynote: Markus Kalisch***	Extracting causal information from observational data: Achievements and Challenges
16:55-17:15	Michael Weber	Network Inference from Transcriptome Monitoring of the Response of Synovial Fibroblasts
17:15-17:35	Miriam Lohr	Extracting differential regulatory sub-networks from genome-wide Microrarray expression data
17:35-17:55	Christian Bender	Dynamic Deterministic Effects Propagation Networks for learning signalling pathways from longitudinal protein array data
17:55-18:15	Hans Kestler	Reducing the complexity of Boolean network search via binarization
	Poster session II + coffee	
20:00	***Get together (Pub – Brauhaus Vetter)***	

Friday, October 22nd, 2010

Session III: Systems Genetics (Chair: Julien Gagneur)

9:00-9:55	***Keynote: Oliver Stegle***	Association mapping of complex phenotypes using models at a systems level
9:55-10:15	Andreas Beyer	Disentangling complex regulatory traits in QTL analysis
10:15-10:35	Bernd Fischer	Mapping Gene Function through Synthetic Genetic Interaction Analysis by RNAi
10:35-10:55	Katrin Knies	Comparison and model selection for disease progression models
10:55-11:30	***coffee***	

Session IV: Open Topics (Chair: Nicole Radde)

11:30-11:50	Bartek Wilczynki	Bayesian Framework For Predicting Spatio-Temporal Expression Patterns During Development
11:50-12:10	Rodrigo Assar	Genome-wide identification of new Wnt/beta-catenin target genes in the human genome using CART method
12:10-12:30	Andreas Raue	Addressing Parameter Identifiability by Model-Based Experimentation
12:30-12:50	Achim Tresch	The dynamics of mRNA synthesis and degradation in response to osmotic stress
12:50-13:00	***Workshop closing***	

Session I: Dynamic modeling (Chair: Ingo Röder)

Keynote:

A Stem Cell Niche Dominance Theorem

Olaf Wolkenhauer (University Rostock), Darryl Shibata (University of Southern California), Mihajlo D. Mesarovic (Case Western Reserve University)
olaf.wolkenhauer@uni-rostock.de

Multilevelness is a defining characteristic of complex systems. For example, in the intestinal tissue the epithelial lining is organized into crypts that are maintained by a niche of stem cells. The behavior of the system 'as a whole' is considered to emerge from the functioning and interactions of its parts. What we are seeking here is a conceptual framework to demonstrate how the "fate" of intestinal crypts is an emergent property that inherently arises from the complex yet robust underlying biology of stem cells. Results: We establish a conceptual framework in which to formalize cross-level principles in the context of tissue organization. To this end we provide a definition for stemness, which is the propensity of a cell lineage to contribute to a tissue fate. We do not consider stemness a property of a cell but link it to the process in which a cell lineage contributes towards tissue (mal)function. We furthermore show that the only logically feasible relationship between the stemness of cell lineages and the emergent fate of their tissue, which satisfies the given criteria, is one of dominance from a particular lineage. Conclusions: The dominance theorem, conceived and proven in this paper, provides support for the concepts of niche succession and monoclonal conversion in intestinal crypts as bottom-up relations, while crypt fission is postulated to be a top-down principle.

Molecular decision making in embryonic and adult stem cells

Ingmar Glauche, Maria Herberg, Tilo Buschmann, Ingo Roeder
TU Dresden
ingmar.glauche@tu-dresden.de

Molecular interactions between transcription factors are considered to be major control mechanisms for stem cell fate decisions. By translating these interactions into an appropriate mathematical state space formulation it is possible to investigate the dynamics of cellular development on the molecular level and establish a conceptual understanding of stem cell fate decisions. In particular, we have established different mathematical models for the description of molecular switches in embryonic and haematopoietic stem cells.

It has been recently demonstrated that individual embryonic stem (ES) cells reversibly change expression level of the crucial transcription factor Nanog and that cells with a low Nanog level are more likely to undergo differentiation. We show that alternations between high and low protein concentrations can be explained by different assumptions yielding similar experimental characteristics. Based on the model results we argue that Nanog variability is a potential "gate-keeper" mechanism to the control of ES cell differentiation.

For the haematopoietic system we discuss a particular molecular switch in myeloid progenitor cells involving the antagonistic transcription factors PU.1 and Gata-1. Using an ODE approach we study the influence of different sources of noise on the transition dynamics changing the system from the co-expression state towards the dominance of one factor in each individual cell. Based on these results we analyzes the emerging heterogeneity on the population level with respect to the overall dynamics of the molecular switch.

Parameter Estimation and Identifiability of Biological Networks Using Relative Data

Patrick Weber, Jan Hasenauer, Frank Allgöwer, Nicole Radde
University of Stuttgart
patrick.weber@ist.uni-stuttgart.de

Data driven modeling of intracellular protein interaction networks requires quantitative time resolved datasets. Beside the problem that few data is available, many standard measurement methods such as Western Blotting, Enzyme-linked Immunosorbent Assay or Fluorescence activated cell sorting are semi quantitative or provide data in form of relative fluorescence or luminescence units. Instead of normalizing ODE models and loosing true physical interaction parameters we present a new method to deal with relative measurements. We focus on a maximum likelihood based approach which can directly include relative measurement data like e.g. light intensities. The resulting optimization problem has the same dimension as the problem arising from absolute data. Therefore, it is computationally only slightly more expensive as the estimation problem for absolute data. We additionally present a method to keep track of identifiability problems which may occur from parameter estimation from relative measurements. The efficiency of the methods is demonstrated with two biological examples: A basic one component overturn reaction and a network of interrelated feedback loops including six species. The second example is centered around protein
kinase D which is responsible for the secretory activity control at trans-Golgi network of mammalian cells.

Switch models for cell differentiation: Bifurcation analysis reveals distinct switching properties

Daniella Schittler, Steffen Waldherr, Frank Allgöwer
IST, University of Stuttgart
schittler@ist.uni-stuttgart.de

Cell differentiation is commonly modeled by gene regulatory networks that cause switching between stable cell states, for example in cells undergoing osteogenesis. We consider a network of three competing transcriptional regulators (TRs), which represent a progenitor, an osteogenic, and a chondrogenic cell type, respectively. From a theoretical viewpoint, there exist multiple interaction networks between these three TRs that can qualitatively reproduce the states during the differentiation process.

We compare distinct hypotheses about how the TRs are interacting, and focus in particular on the role of negative feedback. Our model comparison is based on the bifurcation properties, which correspond to switching behavior during the differentiation process, and thereby have an actual biological relevance. Preliminary results indicate that models with distinct feedback strengths predict qualitatively distinct effects of stimuli on the TR network.

Furthermore, we apply different functional terms to model the kinetic interactions. They are commonly based on specific assumptions about the molecular processes, but which are often not known in sufficient detail. We demonstrate that the choice of kinetic functions can be crucial for the overall model behavior, and that the compared models can yield even contrary predictions.

Since the development and analysis of our minimalistic models are motivated by a cooperative project with partners from cell biology, the ultimate goal is to employ data from differentiation assays for model discrimination, thereby yielding insight into the fundamental mechanisms of cell differentiation.

Session II: Network reconstruction (Chair: Tim Beißbarth)

Keynote:

Extracting causal information from observational data: Achievements and Challenges

Markus Kalisch
ETH Zürich
kalisch@stat.math.ethz.ch

Understanding cause-effect relationships between variables is of interest in many fields of science. It is a well-established scientific principle to determine the total causal effect of one variable on another via randomized controlled intervention experiments. Sometimes, however, experiments are too time consuming, expensive or unethical. We discuss an approach that aims at extracting bounds on causal effects by using observational data only. We outline the underlying theory and discuss strengths and limitations of the approach.

Network Inference from Transcriptome Monitoring of the Response of Synovial Fibroblasts from Rheumatoid Arthritis and Osteoarthritis Patients to TNF-alpha and TGF-beta

Michael Weber (HKI Jena), Reinhard Guthke (HKI Jena),
Raimund.W. Kinne (University Hospital Jena)
michael.weber@hki-jena.de

Introduction: Rheumatoid arthritis (RA) and osteoarthritis (OA), respectively, are the most common inflammatory/degenerative joint diseases in the adult population worldwide. RA is characterised by chronic inflammation, whereas OA shows degenerative changes accompanied by occasional phases of inflammation. Synovial fibroblasts are centrally involved in the pathogenesis of RA and OA. In the synovial environment, fibroblasts are affected by pro-inflammatory cytokines such as TNF-alpha and suppressive cytokines such as TGF-beta. Network inference shall reconstitute molecular interactions of disease-related proteins (cytokines) and, in addition, generate hypotheses to promote further investigations.

Methods: Fibroblasts from RA and OA patients suffering from mild or severe disease states (n=3 for each state and disease) were stimulated by TNF-alpha or TGF-beta (10 ng/ml each). Affymetrix Gene Chips HGU-133 Plus2 were employed to obtain gene expression profiles before, as well as 1 h, 2 h, 4 h and 12 h after stimulation. Candidate genes were selected using a filter combination comprised of a fold-change cut-off and a statistical test. Subsequently, KEGG signalling pathways were analysed with respect to significant representations among the candidate genes. Finally, the expression profiles of selected genes were explained by construction and simulation of ordinary differential equation (ODE) systems using the network inference tool NetGenerator.

Results and Discussion: Genes involved in the prominent Jak-STAT signalling pathway were used for network inference. The generalisability of the resulting model was confirmed by internal evaluation of model stability on the basis of data perturbation. In addition, external validation was carried out by analysing promoter sequences for potential TF binding sites. This work establishes gene regulation hypotheses carrying the potential of elucidating specific regulatory mechanisms pivotal to the abnormal state of rheumatoid fibroblasts.

Extracting differential regulatory sub-networks from genome-wide Microrarray expression data

Miriam Lohr, Patricio Godoy, Jan G. Hengstler, Jörg Rahnenführer and Marco Grzegorczyk
TU Dortmund
lohr@statistik.tu-dortmund.de

One goal in systems biology research is the reconstruction of gene regulatory networks from gene expression data. Bayesian networks have been proposed and applied in the literature, but Bayesian networks inference is computationally expensive and therefore restricted to networks consisting of at most a few dozens of genes. In some studies the focus is on the interactions among certain genes, from which it is known biologically that they are involved in a regulatory process, and Bayesian networks can be applied. In other cases genome-wide expression data have been collected, and there is no biological prior knowledge. The Gene Ontology annotation can then be used to group the genes into sub-groups of moderate sizes and the sub-networks can be inferred with Bayesian networks. Again the computational costs impose a limit on the number of sub-networks that can be analysed. Our interest is on gene interactions that differ significantly between two groups of probands. We propose a workflow based on computationally cheaper Gaussian graphical models and permutation tests and use various statistics that quantify the difference between two Gaussian graphical models to pre-select the most interesting gene sub-groups from genome-wide expression data, which subsequently can be post-processed with Bayesian networks.

Dynamic Deterministic Effects Propagation Networks for learning signalling pathways from longitudinal protein array data

Christian Bender (DKFZ), Frauke Henjes (DKFZ), Holger Fröhlich (Uni Bonn),
Stefan Wiemann (DKFZ), Ulrike Korf (DKFZ), Tim Beißbarth (Uni Heidelberg)
c.bender@dkfz.de

Network modelling in systems biology has become an important tool to study molecular interactions, especially in the medical field like cancer research. The understanding of the interplay of proteins in cellular signalling is the basis for the development of novel drugs and therapies. Here, we set up a new method for the reconstruction of signalling networks from time course protein data after external perturbation. We show how to use protein expression and phosphorylation data measured on Reverse Phase Protein Arrays to infer a signalling network among proteins of the ERBB signalling cascade in a human breast cancer cell line.

Our method models the signalling dynamics by a boolean signal propagation mechanism that defines a sequence of state transitions for a given network structure. A likelihood score is proposed that describes the probability of our measurements given a particular state transition matrix. We identify the optimal sequence of state transitions via a Hidden Markov Model. Network structure search is performed by a genetic algorithm that optimises the overall likelihood of a population of candidate networks. We test our method on simulated networks and data and show its increased performance in comparison to other Dynamical Bayesian Network approaches. The reconstruction of a network in our real data results in several known signalling chains from the ERBB network, showing the validity and usefulness of our approach.

Reducing the complexity of Boolean network search via binarization

Hans A. Kestler
University of Ulm
hans.kestler@uni-ulm.de

At the core of systems biology research lies the identification of biomolecular networks from experimental data via reverse-engineering methods. In this context, Boolean networks provide a model for analysis of gene-regulatory networks. In a Boolean net, a gene is modeled as a Boolean variable over discrete time that can attain two alternative levels, expressed (1) or not expressed (0).
Several methods have been developed for reverse-engineering these networks from 0/1 time series. However, a general problem in reconstructing these networks from time series data is the large number of different genes compared to a relatively low number of temporal measurement points. As a result, the available data are often consistent with multiple network configurations. Furthermore, high-throughput techniques like microarray analyses provide real-valued profiles, hence the data has to be binarized carefully. Yet, the noisiness of gene expression data and the low number of temporal measurement points often yield several plausible binarizations. Differences in the binarization results can have strong effects on the architecture of the resulting Boolean networks because a state change for a single gene can lead to many differences in downstream functions and gene dependencies.
To address some of these issues, we developed scale-space binarization methods to produce suitable and robust thresholds even for small numbers of data points. Additionally a measure of validity for the found thresholds is given. Incor porating this knowledge into network reconstruction allows for restricting the input of reconstruction algorithms to genes with meaningful thresholds. This reduces the complexity of network inference. The performance of our binarization algorithms was evaluated in network reconstruction experiments using artificial data as well as real-world yeast expression time series. The new approaches yield considerably improved correct network identification rates compared to other binarization techniques by effectively reducing the amount of candidate networks.

Session III: Systems Genetics (Chair: Julien Gagneur)

Keynote:

Association mapping of complex phenotypes using models at a systems level

Oliver Stegle
MPI for Biological Cybernetics
oliver.stegle@tuebingen.mpg.de

The goal of genetic association studies is to relate polymorphic genetic loci to quantitative traits that capture the phenotypic variation of interest. While this task is well studied for "simple" univariate phenotypes, recent high-dimensional phenotypes ask for more advanced modeling techniques at a systems level.

In this talk I will discuss recent machine learning techniques to address these emerging challenges and illustrate them using two case studies. First, I will focus on QTL mapping of high-dimensional microarray data in yeast. In this setting, inference of unmeasured transcription factor activations from the expression levels and pathway information allows the phenotypic variation to be dissected at a previously unavailable level of detail. Second, I will discuss mapping techniques for the treatment response of human depressed patients. In these data, the joint modeling of observations that change over time and complex confounding influence need to be taken into account to identify relevant genetic markers.

Disentangling complex regulatory traits in QTL analysis

Mathieu Clement-Ziza, Andreas Beyer
TU Dresden
andreas.beyer@biotec.tu-dresden.de

Although the technique of quantitative trait locus (QTL) analysis has helped understanding the regulation of complex traits, it has proven difficult to comprehensively determine their complex genetic architecture, partially because phenotype variation is often the result of multiple variable processes. Gene expression is a good example of such complex process: it involves several steps each of which is regulated by the cell. Expression QTL studies have been carried out using either transcript or protein abundance to monitor gene expression. However, different biological processes underlie those traits: at steady state, the transcript abundance is mainly dependent on transcription and RNA degradation whereas protein levels are also affected by post-transcriptional processes such as translation or protein degradation. In this work, we dissected gene expression traits from which we isolated the post-transcriptional component in order to better understand its specific regulation. We modeled post-transcriptional variation as the residuals after regressing on RNA levels. We integrated published data obtained from a yeast population genotyped at 1106 markers and phenotyped at the transcriptomic and proteomic levels of 137 genes. Mapping this inferred post-transcriptional contribution revealed 36 loci that post-transcriptionally affect 64 proteins. We identified regulatory hotspots that control many genes, and a candidate master regulator of amino-acid metabolism genes. Our work presents an example of how to disentangle related (yet different) complex traits in order to reveal their genetic basis.

Mapping Gene Function through Synthetic Genetic Interaction Analysis by RNAi

Bernd Fischer (EMBL), Thomas Horn, Thomas Sandmann, Wolfgang Huber (EMBL), Michael Boutros (DKFZ)
bernd.fischer@embl.de

Synthetic genetic interaction analysis provides key insights into functional relationships between genes. Large scale genetic interaction networks have recently been published for yeast. To systematically identify genetic interactions in metazoan cells, we used RNAi to simultaneously reduce the expression of thousands of genes pairs in Drosophila cell lines.

We developed methods to quantitatively assess phenotypes from double RNAi perturbations in a high-throughput format. Multi-parametric fluorescence images were recorded by automatic microscopy. An image processing pipeline, including the segmentation of cells and extraction of features, provided a multi-dimensional phentoypic feature vector for each double RNAi treatment. By statistical modeling a multi-dimensional and quantitative genetic interaction map was created.

Using unsupervised and supervised learning methods we were able to reconstruct the function of known genes with respect to different signal transduction pathways. Furthermore through supervised learning we predicted new functions for previously uncharacterized genes, including a novel regulator of Ras/MAPK sigaling, whose conserved function could be validated experimentally in vitro and in vivo.

Comparison and model selection for disease progression models

Katrin Knies and Roland Fried
TU Dortmund
knies@statistik.tu-dortmund.de

For many cancer types, tumor progression can be characterized by the accumulation of complex chromosome alterations. Gains and losses on certain chromosome arms indicate events in the course of disease. It is believed that these events do not occur at random, but follow a certain order. That means, once an event occurs, it increases the probability of other events occurring. One is interested in finding evolutionary pathways, whereby oncogenesis proceeds. There are several models which describe these pathways in a different manner, for example oncogenetic tree models, mutagenetic tree mixture models and conjunctive Bayesian networks. Every model has certain assumptions and models tumor progression in a special framework. By means of simulations we want to compare the different models with respect to the underlying data situation and the quality of the results. One aim is to answer the question if a simple model already provides adequate results or if more complex models are necessary to capture all evolutionary pathways. We also analyze how progression models work if some of their assumptions are violated. Do they still capture at least some of the correct pathways or do they provide useless results? If the true model structure is unknown one needs a criterion to decide which model to use for fitting. We present different model selection criteria, which will be applied to our simulated data to see which one provides the best results.

Session IV: Open Topics (Chair: Nicole Radde)

Addressing Parameter Identifiability by Model-Based Experimentation

A. Raue (Uni Freiburg), C. Kreutz (Uni Freiburg), U. Klingmüller (DKFZ) and J. Timmer (Uni Freiburg)
andreas.raue@fdm.uni-freiburg.de

Modeling of protein interaction networks such as signaling pathways by ordinary differential equations faces difficulties when estimating model parameters like rate constants from incomplete and noisy experimental data. Typically, networks are only partially observed meaning that not all molecular species can be measured directly. This can lead to structurally non-identifiable model parameters. Furthermore, limited amount and quality of experimental data can cause model parameters to be practically non-identifiable. Both have serious impact on model predictions.

We present an approach that exploits the profile likelihood to detect both structural and practical non-identifiabilities and derives confidence intervals, see [1]. The approach is computationally feasible for large models. By means of an illustrative example, we show how the results can be used for experimental planning, i.e. to design new experiments that efficiently resolve parameter identification issues and thereby enhance model predictability.

References:

A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller and J. Timmer. Structural and practical identifiability analysis; of partially observed dynamical models by exploiting the profile likelihood. Bioinformatics (2009), 25(15), 1923-1929.

Genome-wide identification of new Wnt/beta-catenin target genes in the human genome using CART method

Rodrigo Assar, Christian Hodar, Marcela Colombres et al.
Universität Bordeaux
Rodrigo.Assar@inria.fr

Nowadays the stored information of non-coding human genome regions is being strongly exploited to model the regulation processes. The key to explain complex biological responses, to anticipate and to treat diseases is to understand the regulation factors and the interactions with genes. I summarize the recently published work, developed by C. Hodar, myself, M. Colombres et al. It uses a statistical point of view to study the Wnt pathway that is mainly implicated in cell differentiation, and diseases such as Alzheimer`s (AD) and cancer. We obtain insights on new human genome regions being relevant to control AD or other neurodegenerative disorders. We construct a statistical method based on multiple Classification Trees (CART) to identify new Wnt/beta-catenin pathway target genes, by only using information of known Wnt target (positive) genes and it considers as decision variables the presence of transcription factor binding site motifs in the upstream region of each gene. We propose 89 new Wnt/beta-catenin pathway target genes, we found a group expressed
in brain tissue that could be involved in diverse responses to neurodegenerative diseases, like AD. These genes represent new candidates to protect cells against amyloid beta; toxicity, maybe associated to the neuroprotective role of the Wnt signaling pathway. The most relevant were calcium/calmodulin-dependent protein kinase IV (CamK4), with strong evidences for up-regulation in response to both Wnt ligands and lithium, and tropomyosin 1 (alpha) that is associated with neurofibrillary pathology of AD. Our strategy proved to be effective and robust to identify new Wnt/beta-catenin pathway targets. In silico results were biologically validated by RT-qPCR in a sample. Several of the genes represent a new group of transcriptional dependent targets of the canonical Wnt pathway, that could be involved in pathophysiology related to Alzheimer.

Bayesian Framework For Predicting Spatio-Temporal Expression Patterns During Development

Bartek Wilczynki, Ya-Hsin Liu, Zhen Xuan Yeo, Eileen EM Furlong
EMBL Heidelberg
wilczyns@embl.de

Tight regulation of gene expression patterns, both temporal and spatial, isessential to ensure reproducibility in animal development. In the process,expression pattern of each gene is controlled by the sum of action of multiplecis-regulatory modules (CRMs), each of which can be responsible for differentspatial and temporal expression of the target gene. In turn, the activity ofeach CRM is determined by time-specific binding of combinations of transcriptionfactors (TFs): proteins able to bind DNA and recruit transcriptional machineryand initiate transcription. Recently, Zinzen et al [1] showed that, at least ina particularly well studied system such as Drosophila mesoderm development, itis possible to predict CRM activity from genome wide data on TF occupancy usinga machine learning approach based on Support Vector Machines. Our aim is toextend the results of this earlier work to a model predictive of target geneexpression. We have built a Bayesian model explicitly defining the probabilitiesof target gene expression in each condition (spatial or temporal) depending onthe probabilities of CRM acticity in its vicinity. The dependence between randomvariables is represented as a Bayesian Network and the parameters are optimizedby an iterative Expectation Maximization procedure. In a cross-validationexperiment, we find that TF binding is not enough to generate reliablepredictions of gene expression, however extending the model to take into accountthe chromatin state (in the form of insulator protein binding and some histonemodifications) the results can be greatly improved. We made genome-widepredictions of gene expression using the improved model and tested them inadditional in-situ experiments finding majority of them to be correct.

[1] Zinzen RP, Girardot C, Gagneur J, Braun M, Furlong EE (2009) Combinatorialbinding predicts spatio-temporal cis-regulatory activity. Nature 462: 65-70

The dynamics of mRNA synthesis and degradation in response to osmotic stress

Achim Tresch, Björn Schwalb, Sebastian Dömcke, Benedikt Zacher, Christian Miller, Kerstin Maier, Andreas Mayer, Daniel Schulz, Dietmar Martin, Patrick Cramer
Gene Center and Department of Biochemistry, Ludwig-Maximilians-University Munich, Germany
tresch@lmb.uni-muenchen.de

Cellular transcript levels are determined by the interplay of RNA synthesis and decay, and both processes may be specifically and differentially regulated in response to environmental changes. We use a 4-thiouridine labeling technique to separate the nascent mRNA from pre-existing mRNA. The genome-wide nascent, the pre-existing, and the total mRNA transcript levels can then be quantified via microarrays of deep sequencing. We developed the bioinformatics to derive reliable estimates of mRNA synthesis and degradation, not only in the steady-state case, but also during dynamic changes such as the response to osmotic stress. The result is an increased signal-to-noise ratio for the detection of altered transcription, and a higher temporal resolution of the order of gene regulatory events, which enables the distinction between primary and secondary effects. Applying our method to examine the yeast transcriptional response to osmotic stress, we can clearly distinguish three phases, characterized by 1) an immediate transcriptional shutoff, 2) a fast induction of stress-specific genes, and 3) a delayed recovery of the overall transcriptional activity. Interestingly, we discover a temporary coupling of synthesis and degradation during phase 2, and we speculate about an indirect cause for this. Our findings are confirmed by complementary ChIP-chip measurements of RNA Polymerase II, which provide mechanistic insights into the salt stress response, suggesting that regulation of transcription is performed both at the stage of Pol II recruitment and at the transition to elongation.

Poster Session:

1:

Controlling the false coverage-statement rate of adjusted confidence intervals for the fold change - a complement to the false discovery rate

Klaus Jung and Tim Beißbarth
University Medical Center Göttingen
Klaus.Jung@ams.med.uni-goettingen.de

Comparing two distinct groups of biological samples within a DNA microarray experiment, differentially expressed genes are detected by multiple testing procedures and control of the false discovery rate. Additionally, the difference of each gene between the two groups is quantified by the log fold change. However, genes with rather small fold changes that display a small variance, often reach the same level of significance as genes with a rather large fold change that display a large variance. Especially small fold-changes might, however, not be biologically meaningfull, and this disagreement is further often confusing, especially for the non-statistician. Therefore, Benajamini and Yekuteli (2005) proposed to report adjusted confidence intervals with each fold change that coincide with the FDR-adjusted p-value. In this context, they introduced the false coverage-statement rate, which is defined as the portion of genes falsely selected by means of the confidence intervals of their log fold change. I.e., genes whose confidence interval does not cover the zero log fold change are selected as differentially expressed. A similar approach was proposed by Jung et al. (2009) who used FDR-adjusted confidence levels in building the adjusted confidence intervals. Here, we connect the algorithms for adjusting confidence intervals with the widely used linear models for microarray data (Smyth, 2004). In a simulation study, we show that this approach controls the false coverage-statement rate and that the obtained confidence intervals for the fold change coincide with the FDR-adjusted p-values, i.e. the lower (upper) limit exceed (falls below) zero when the associated p-value supports significance. In summary, this approach may facilitate biologists and physicians to select more intuitively significant genes.

References:

Benjamini, Y. and Yekuteli, D. (2005). False Discovery Rate-Adjusted Multiple Confidence Intervals for Selected Parameters. Journal of the American Statistical Association 100, 71-81.

Jung, K., Poschmann, G., Podwojski, K., Eisenacher, M., Kohl, M., Pfeiffer, K., Meyer, H.E., Stühler, K. and Stephan, C. (2009). Adjusted Confidence Intervals for the Expression Change of Proteins Observed in 2-Dimensional Difference Gel Electrophoresis. Journal of Proteomics and Bioinformatics 2, 78-87.

Smyth, G. (2004). Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments. Statistical Applications in Genetics and Molecular Biology 3, Article 3.

2:

Survival analysis from microarray data using FDR and ridge-regression.

Eugen Rempel, Jörg Rahnenführer
TU Dortmund
rempel@statistik.tu-dortmund.de

Use of microarray technology often leads to high-dimensional and low-sample size data. A variety of approaches for variable selection have been suggested in this context. We propose a filtering technique applying FDR followed by ridge-regression. Using three microarray gene expression data sets, we compare our method with another known approaches. It turns out, that our approach is competitive.

3:

Genome-wide identification of prognostic genes with bimodal expression distribution from gene array data

Birte Hellwig, Jan G. Hengstler, Marcus Schmidt and Jörg Rahnenführer
TU Dortmund
hellwig@statistik.tu-dortmund.de

Background: A major goal of the analysis of high-dimensional RNA expression data from tumor tissue is to identify prognostic signatures for discriminating patient subgroups. For this purpose genome-wide identification of bimodally expressed genes from gene array data is relevant because distinguishability of high and low expression groups is easier compared to genes with unimodal expression distributions. Recently, several methods for the identification of genes with bimodal distributions have been introduced. A straightforward approach is to cluster the expression values and score the distance between the two distributions. Other scores directly measure properties of the distribution. The kurtosis, e.g., measures divergence from a normal distribution. An alternative is the outlier-sum statistic that identifies genes with extremely high or low expression values in a subset of the samples.

Results: We compare and discuss scores for bimodality for expression data. For the genome-wide identification of bimodal genes we apply all scores to expression data from 194 patients with node-negative breast cancer. Further, we present the first comprehensive genome-wide evaluation of the prognostic relevance of bimodal genes. We first rank genes according to bimodality scores and define two patient subgroups based on expression values. Then we assess the prognostic significance of the top ranking bimodal genes by comparing the survival functions of the two patient subgroups. We also evaluate the global association between the bimodal shape of expression distributions and survival times with an enrichment type analysis. Various cluster-based methods lead to a significant overrepresentation of prognostic genes. A striking result is obtained with the outlier-sum statistic ($p<10^{-12}$). Many genes with heavy tails generate subgroups of patients with different prognosis.

Conclusions: Genes with high bimodality scores are promising candidates for defining prognostic patient subgroups from expression data. We discuss advantages and disadvantages of the different scores for prognostic purposes. The outlier-sum statistic may be particularly valuable for the identification of genes to be included in prognostic signatures. Among the genes identified as bimodal in the breast cancer data set several have not yet previously been recognized to be prognostic and bimodally expressed in breast cancer. Examples are MAGEA2 and MAGEA4 that play central roles in immune response and myosin heavy chain 7 (MYH7), which has been associated with heart disease but not with cancer.

4:

Fast and Efficient Dynamic Nested Effects Models

Holger Fröhlich (University of Bonn), Achim Tresch (Gene Center Munich)
frohlich@bit.uni-bonn.de

Targeted interventions in combination with the measurement of secondary effects can be used to computationally reverse engineer features of upstream non-transcriptional signaling cascades based on the nested structure of its effects. Nested effect models (NEMs) have been introduced first by Markowetz et al (2005) as a statistical approach to estimate the upstream signal flow from downstream nested perturbation effects. The method was substantially extended later on by several authors and successfully applied to different datasets. The connection of NEMs to Bayesian Networks and factor graph models has been highlighted. Here we introcude a computationally attractive extension of NEMs, which allows for dealing with perturbation time series. In contrast to Anchang et al. (2009) the key idea in our model is to unroll the signal flow over time. This allows for a computation showing some similarity to Dynamic Bayesian Networks and naturally extends the classical NEM formulation introduced by Markowetz et al. In our model we circumvent the need for any time consuming Gibbs sampling, which makes it also computationally attractive. Our simulations show a good competitive performance of our method with respect to network reconstruction accuracy. Applying our technique to a data set from murine stem cell development (Ivanova, 2006) we found a high accordance with current literature knowledge.

5:

Linking parallel measurements of high-throughput miRNA, gene and protein expression data

Stephan Gade (DKFZ), Harald Binder (Uni Freiburg), Tim Beißbarth (Uni Göttingen)
S.Gade@DKFZ-Heidelberg.de

The parallel measurement of different types of high throughput expression data, like gene, miRNA and protein data, gives rise to new tasks of data integration, specifically, when also the combined effect on a clinical endpoint is of interest. Different strategies have been formulated, performing the integration on different levels of the analysis process. Specifically, miRNAs are small non-coding RNAs which play an important role during cancer development. The details of the miRNA mediated regulation mechanism are not fully understood. Though it is known that the binding of a miRNA to its target transcript can lead to degradation and consequently to a measurable change in the level of the mRNA.

We proposed a graph-based work flow to integrate miRNA and gene expression data. By combining the correlation structure of the two data sets with target predictions we were able to construct a bipartite graph with connections between miRNAs and genes. Meta information were added and used to distinguish reasonable sub clusters in this graph. This structure was used to perform an over-representation analysis of targets and effected pathways. Furthermore, the bipartite graph is interpretable as a pathway linking different kinds of features. Similar to gene pathways it is suitable as an additional knowledge for classification tasks.

We applied this method to a colorectal cancer data set consisting of 97 samples and measurements of miRNA and gene expressions. With help of the graph structure we were able to identify clusters of miRNAs with significant correlations to genes in cancer related pathways like cytokine-cytokine receptor interaction and the Jak-STAT signaling. Finally, for linking the miRNA and gene expression data to a clinical endpoint, a componentwise boosting approach was employed that incorporates the graph information for fitting a multivariable model and selects a small set of prognostic molecular entities.

Combining different kind of omics data is a challenging task. With this graph based workflow we presented a method that combines the different data sets in a graph model giving rise to e.g. pathway overrepresentation analysis and classification methods for predicting clinical endpoints. Thereby, not only the features of both data sets but the coherences between these features are used.

6:

Robust Subnetworks - Computing Confidence Scores for Functional Modules

Daniela Beisser, Thomas Dandekar, Gunnar Klau, Marcus Dittrich, Tobias Müller
Department of Bioinformatics, University of Würzburg, Würzburg
daniela.beisser@biozentrum.uni-wuerzburg.de

High-throughput genomic data provides a wealth of information that is widely used in integrated network analysis. Several heuristic approaches and an exact approach based on integer linear programming by Dittrich et al. (2008) allow to identify functional modules, pathways or gene signatures containing differentially expressed genes in the context of protein-protein interaction networks. The objective of the presented study is to assess the robustness of the identified modules and explore the impact of jackknife resampling on the performance of module detection. We propose a novel concept of consensus modules for the integrated analysis of biological networks based on jackknife resampling of the integrated microarray data. This allows to assess the local robustness of all nodes and edges in the identified modules and the assignment of confidence scores to the original module. The major aim of module detection algorithms is to identify functional modules, which are not a priori known for real biological data sets. Hence, it is difficult to assess the quality of the obtained solution. We apply the presented resampling procedure to the diffuse large B-cell (DLBCL) expression data set and assess the robustness of each node and edge of the resultant optimal module. In particular we compare the different solutions obtained from the plain optimal module and the bootstrap consensus module. Furthermore, we developed a simulation framework to evaluate the performance of the proposed algorithm on known reference modules and compare it to the exact approach and other heuristic algorithms.

7:

Insertional mutagenesis in mature T cells - a mathematical modelling approach

Sebastian Gerdes, Sebastian Newrzela, Tim Heinrich, Dorothee von Laer, Ingmar Glauche, Ingo Roeder
Dresden University of Technology
sebastian.gerdes@tu-dresden.de

Leukemia caused by insertional mutagenesis is a major risk of gene therapy with HSCs. In contrast, polyclonal mature T cells have turned out to be surprisingly resistant to leukemogenic insertional mutagenesis. Newrzela et al. showed that leukemia/lymphoma did not occur upon transplantation of wild-type mature T cells with polyclonal T-cell receptors (TCR) being transduced with high copy numbers of gammaretroviral vectors encoding potent T- cell oncogenes into RAG1-1-deficient recipients [1]. However, further studies demonstrated that the transplantation of T cells from TCR-monoclonal OT1 mice that were transduced with the same protocol resulted in leukemia/lymphoma. The underlying mechanisms that prevent oncogenesis in the polyclonal situation and permit the outbreak of leukemia in the monoclonal situation are currently unclear.

Using a mathematical modelling approach, we challenge the arising hypothesis that polyclonality induces competition within the T-cell repertoire, which in turn suppresses the emergence of a leukemic clone. As a first step, we developed a mathematical model of T-cell homeostasis that is derived from a similar niche-based based model of hematopoiesis [2]. The key assumption of the novel model is that T-cell survival is critically dependant on the interaction of the clone-specific TCR with self-peptide-MHC-complexes and therefore subject to competition between different T-cell clones.

Our model consistently reproduces the polyclonal pattern observed in T-cell homeostasis, and responds in an adequate manner to perturbations of the system such as infection. Furthermore, based on our modelling results, we speculate how the deregulation of T-cell receptor affinities supports the formation of dominant clones. The modelling results underscore the possibility that clonal competition can prevent the outbreak of overt mature T-cell leukemia/lymphoma.

References:
[1] Newrzela S, Cornils K et al. Resistance of mature T cells to oncogene transformation. Blood. 2008;112(6):2278-2286.
[2] Roeder I., Loeffler M. A novel dynamic model of hematopoietic stem cell organization based on the concept of within-tissue plasticity. Experimental Hematology. 2002;30(8):853-861.

8:

Modeling of the TNF-TNF receptor signaling network

G. Wangorsch, M. Dittrich, H. Wajant, T. Dandekar
University of Würzburg
gaby.wangorsch@biozentrum.uni-wuerzburg.de

Our aim is to understand the interplay between TNF receptor 1 and 2, soluble TNF, membrane TNF and TRAF molecules. In addition, the production of endogenous TNF concerns the induction of apoptosis and the activation of NFκB. This interplay is a serious challenge to model, especially if one has later clinical applications in mind. An experimentally validated description of this network will ameliorate and underpin the development of therapy concepts resting on the modulation of the network. Quantification of signaling components and time series data were obtained through investigation of well-characterized cell lines. The cellular responses of HeLa cells, HeLa-TNFR2 transfectants and the TWEAK/Fn14 system cast light on the interplay of TNF receptor 1 and 2 signaling. Based on these data, we established a mathematical model (SBML, CellDesigner) which serves as a basis for further analysis, for investigation and prediction of the cellular signaling outcome of the TNF-TNF receptor network. We performed simulations based on Boolean modeling (SQUAD) to gain general insight into the overall network. On the basis of the model topology, SQUAD permits simulations on the continuous system even in the total absence of kinetic data. Furthermore, we fit time series data to submodels of the receptor signaling network, consisting of a set of ordinary differential equations (Matlab, PottersWheel). This enables the analysis and optimization of model parameters. By model selection we achieve parsimonious submodels of the modeled network modules. Mathematical modeling accompanied by experimental validation supports the understanding and prediction of cellular responses due to TNF induction and serves as a test platform to investigate cross-talk possibilities.

9:

Experimental Design for Parameter Identification in Cellular Interaction Networks

Andrei Kramer and Nicole Radde
IST, University of Stuttgart
kramer@ist.uni-stuttgart.de

The design of experiments is a concept of a-priori decision making concerning the setup of an experiment or the selection of a suitable experiment. It is worthwhile whenever the experiments are expensive and time consuming. We show a simple method of experiment selection for biochemical reaction networks designed for optimal parameter identification. Based on an interaction model of the participating molecular species and a model of the measurement process we estimate the average information content for a measurement procedure. This estimated average information can be compared between different experiment proposals. Information available beforehand can be incorporated as a prior distribution in a Bayesian scheme. We use the entropy of the posterior distribution as a measure of information (see [1]). The distribution entropy is obtained via sampling and normalization (described in [2]). We show results for an example ODE system for the regulation of protein secretion at the Trans Golgi Network via the protein kinase D; here we used a realistic measurement model.

Keywords: Parameter Sampling; Monte Carlo; ODE models;

References:

[1] Andrei Kramer, Nicole Radde (2010) Towards experimental design using a Bayesian framework for parameter identification in dynamic intracellular network models. Procedia Computer Science, Volume 1, Issue 1, May 2010, Pages 1639-1647

[2] A. Kramer, J. Hasenauer, F. Allgöwer and N. Radde (MSC 2010) IEEE Multi Conference on Systems and Control, Septermber 9-10, 2010, Yokohama, Japan, accepted. Computation of the posterior entropy in a Bayesian framework for parameter estimation in biological networks.

10:

Temporal Activation Profiles of Gene Groups

Andre König, Jörg Rahnenführer
TU Dortmund
akoenig@statistik.tu-dortmund.de

Time series of gene expression measurements are used to study various biological processes, e.g. immune response, cancer progression or embryonic development. Due to the costs of microarray experiments in many research projects only a few times, typically less than 8, are analyzed. Moreover, due to limited biological material or money, often none or just few replicates are considered.

We offer an approach to identify activation profiles for predefined gene sets. The use of gene sets defined by Gene Ontology (GO) is established. In distinction from other used algorithms we do neither cluster genes according to predefined expression profiles nor discriminate groups with general temporal shift.

Our bottom-up algorithm first compares the expression values at every instant with a reference distribution and identifies differentially expressed genes separately for all times. In a second step we obtain the groups whose genes are overrepresented among the differentially expressed genes per instant. This yields a characteristic and well interpretable time profile for all considered groups.

A few activation profiles show discontinuous activation over the observed time. This is for most groups hard to interpret and therefore we developed smoothing algorithms, which result for reasonable cases in a continuous activation profile.

11:

Gene Set Analysis of SNP data using Gene Ontology

Helena Janzen, Jan G. Hengstler, Katja Ickstadt
TU Dortmund
janzen@statistik.tu-dortmund.de

Gene Set Enrichment Analysis (GSEA) is a widely-used method to analyze gene expression data obtained from microarray experiments. One goal of GSEA is to find gene sets whose genes are overrepresented among the differentially expressed genes. Those groups show significant enrichment with gene expression differences between two observed population samples, e.g. cancer samples versus controls. A widespread way of gene set definition are the terms of the three ontologies of the Gene Ontology (GO) Project, which offer for example information about biological processes or molecular functions.

The idea is to transfer these methods to ordinal Single Nucleotide Polymorphism (SNP) data from genome wide association studies. For this purpose assignments of SNPs to genes and subsequently to gene sets are needed. Furthermore an adequate discrimination for the SNP distribution in the two populations is required. We apply the GSEA-SNP algorithm (Holden et al., 2008) and the classical Overrepresentation Analysis combined with GO sets to an urinary bladder cancer case-control SNP study. As additional method we adopt the SAM-GS algorithm (Dinu et al., 2007) to the cancer SNP data. Although GSEA cannot find significant GO sets in the analysis of the bladder cancer data the comparison of these three methods show some overlap of enriched GO groups.

12:

Integrative analysis of low- and high-resolution eQTL.

Salvatore Loguercio, Rupert Overall, Jacob Michaelson, Tim Wiltshire, Mathew Pletcher, Brooke Miller, John Walker, Gerd Kempermann, Andrew Su, and Andreas Beyer
BIOTEC, TU Dresden; Germany
salvatore.loguercio@biotec.tu-dresden.de

The study of expression quantitative trait loci (eQTL) is a powerful way of detecting transcriptional regulators at a genomic scale and for elucidating how natural genetic variation impacts gene expression. Power and genetic resolution are heavily affected by the study population: whereas recombinant inbred (RI) strains yield greater statistical power with low genetic resolution, using diverse inbred or outbred strains improves genetic resolution at the cost of lower power. In order to overcome the limitations of both individual approaches, we combine data from RI strains with genetically more diverse strains and analyze hippocampus eQTL data obtained from mouse RI strains (BXD) and from a panel of diverse inbred strains (MDP). We perform a systematic analysis of the consistency of eQTL independently obtained from these two populations and demonstrate that a significant fraction of eQTL can be replicated. Based on existing knowledge from pathway databases we assess different approaches for using the high-resolution MDP data for fine mapping BXD eQTL. Finally, we apply this framework to an eQTL hotspot on chromosome 1 (Qrr1), which has been implicated in a range of neurological traits. Here we present the first systematic examination of the consistency between eQTL obtained independently from the BXD and MDP populations. Our analysis of fine-mapping approaches is based on real life data as opposed to simulated data and it allows us to propose a strategy for using MDP data to fine map BXD eQTL. Application of this framework to Qrr1 reveals that this eQTL hotspot is not caused by just one (or few) master regulators, but actually by a set of polymorphic genes specific to the central nervous system.

13:

HMM length optimization improves sequence and structure modeling

Christian Kötschan, Torben Friedrich, Tobias Müller
University of Würzburg
c.koetschan@biozentrum.uni-wuerzburg.de

Hidden Markov Models (HMMs) gained an important role in computational biology and especially in sequence modeling. Although HMMs are well studied and share a broad range of application, there still is aneed for improvment regarding data sources with variable sequence lengths. Each state of an HMM is associated with a specific duration of stay, which can be correlated to a number of (emitted) symbols. By manipulating this holding time for some states of a Markov Chain, one is able to influence the underlying length distribution of a dataset. However, the more interesting approach is to create an HMM that fits to the length distribution of the dataset. Here, two different estimators are proposed. The first uses a maximum likelihood approach; the second is a fast and easy implementable moment estimator. Both estimators calculate a factor for state repetition, as well a probability for state-specific holding times. These parameters enable the adaptation of regular HMM topologies to model even bell-shaped sequence lengths by chain-linking hidden states. To demonstrate this concept, the secondary structure of the internal transcribed spacer 2 is modeled and the optimized HMM is compared to a regular HMM version.

14:

Filtering networks with gene-expression data: How much tissue-specificity is retained by including the network neighborhood?

Mathias Fuchs
University Medical Center Göttingen

mfu@bioinf.med.uni-goettingen.de

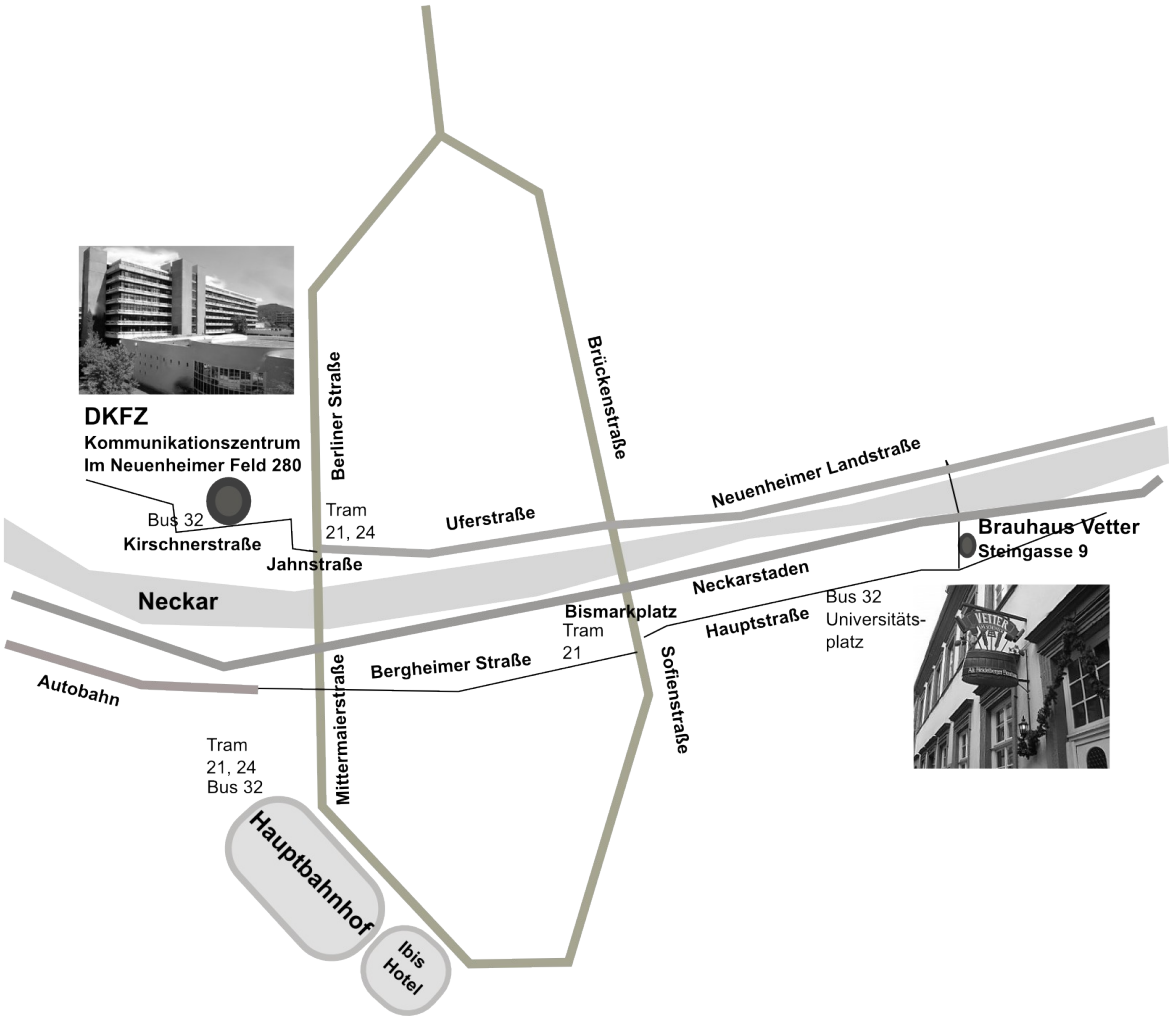
An important question in Transcriptomics is to find expression characteristics shared by samples of related tissues. While several approaches rely on annotation counts, we pursue a more quantitative approach. We define a notion of tissue-over-expression for gene sets, construct the corresponding statistical hypothesis test, and study this property rather than tissue-specificity. We filter a signal-transduction network derived from TRANSPATH database with Alzheimer's disease expression data obtaining two sub-networks, one of which reflects the expressed genes, the other of which includes their neighbors according to the signaling network. We define a gene set to be relatively tissue-over-expressed (RTOE) whenever its expression levels in a given tissue are, on average, higher than those in the remaining background set, where expression levels are normalized in such a way that all genes obtain the same weight. We apply the test on the resulting networks and show that the network's neighborhood of the expressed genes has the over-expression property for the data origin's tissues. High significance levels are attained. We show that the method applied to AD data resulted in very good RTOE-properties of the brain-related tissues. These results are not immediately expected, in particular in the 1-extended case. By construction, the strictly filtered network should have good RTOE-property, only interfered by the double mapping procedure. However, the strict RTOE-property allows to deduce that the information loss inherent to mapping is minimal. Furthermore, we check that the RTOE-property of the 1-extended network is not merely an artifact of the strict network's one.

Participants

Name	Institute	E-Mail
Ababneh, Faisal	Al-hussein bin Talal University	fmababneh@gmail.com
Artmann, Stephan	University Medical Center Göttingen	stephan.artmann@stud.uni-goettingen.de
Assar, Rodrigo	University Bordeaux, INRIA	Rodrigo.Assar@inria.fr
Barry, Joanne	BioQuant	joanne.barry@bioquant.uni-heidelberg.de
Beissbarth, Tim	University Medical Center Göttingen	tim.beissbarth@ams.med.uni-goettingen.de
Beisser, Daniela	Bioinformatics, University of Würzburg	daniela.beisser@biozentrum.uni-wuerzburg.de
Bender, Christian	DKFZ	c.bender@dkfz.de
Benner, Axel	DKFZ	a.benner@dkfz.de
Beyer, Andreas	BIOTEC, TU Dresden	andreas.beyer@biotec.tu-dresden.de
Brors, Benedikt	DKFZ	b.brors@dkfz.de
Buness, Andreas	Roche	a.buness@gmx.de
Daniela, Dossing	Deutsches Diabetes Zentrum Düsseldorf	daniela.dossing@ddz.uni-duesseldorf.de
Esins, Janina	MPI for Biological Cybernetics	janina.esins@tuebingen.mpg.de
Faelth, Maria	DKFZ	m.faelth@dkfz.de
Fellmann, Mark	Roche Diagnostics Ltd	mark.fellmann@roche.com
Fischer, Bernd	EMBL	bernd.fischer@embl.de
Fischer, Martina	DKFZ Heidelberg	martina.fischer@dkfz-heidelberg.de
Fröhlich, Holger	University Bonn	frohlich@bit.uni-bonn.de
Fuchs, Mathias	Department of Bioinformatics, UMG Göttingen	mfu@bioinf.med.uni-goettingen.de
Gade, Stephan	DKFZ	s.gade@dkfz.de
Gagneur, Julien	EMBL	gagneur@embl.de
Gerdes, Sebastian	IMB Dresden	sebastian.gerdes@tu-dresden.de
Glauche, Ingmar	TU Dresden	ingmar.glauche@tu-dresden.de
Grieb, Melanie	University Ulm	melanie.grieb@uni-ulm.de
Groß, Alexander	University Ulm	alexander.gross@uni-ulm.de
Hellwig, Birte	TU Dortmund	hellwig@statistik.tu-dortmund.de
Horn, Matthias	IMISE / University of Leipzig	matthias.horn@imise.uni-leipzig.de
Iwamoto, Nao	DKFZ	n.iwamoto@dkfz.de
Janzen, Helena	TU Dortmund	janzen@statistik.tu-dortmund.de
Jöcker, Annika	DKFZ	a.joecker@dkfz.de
Johannes, Marc	DKFZ	m.johannes@dkfz-heidelberg.de
Jung, Klaus	University Medical Center Göttingen	Klaus.Jung@ams.med.uni-goettingen.de
Kalisch, Markus	ETH Zuerich	kalisch@stat.math.ethz.ch
Kestler, Hans	University Ulm	hans.kestler@uni-ulm.de
Kiani, Narsis	University of Heidelberg	narsis.kiani@bioquant.uni-heidelberg.de
Knapp, Bettina	Uni Heidelberg	bettina.knapp@bioquant.uni-heidelberg.de
Knies, Katrin	TU Dortmund	knies@statistik.tu-dortmund.de

Name	Institute	E-Mail
Koetschan, Christian	Department of Bioinformatics / Biocenter	c.koetschan@biozentrum.uni-wuerzburg.de
König, Andre	TU Dortmund	akoenig@statistik.tu-dortmund.de
Kramer, Andrei	University of Stuttgart	kramer@ist.uni-stuttgart.de
Kramer, Frank	University Medical Center Göttingen	frank.kramer@med.uni-goettingen.de
Leha, Andreas	University Medical Center Göttingen	andreas.leha@med.uni-goettingen.de
Lenkoski, Alex	University Heidelberg	alex.lenkoski.uw@gmail.com
Loguercio, Salvatore	BIOTEC, TU Dresden	salvatore.loguercio@biotec.tu-dresden.de
Lohr, Miriam	TU Dortmund	lohr@statistik.tu-dortmund.de
Lucarelli, Philippe	DKFZ	p.lucarelli@dkfz-heidelberg.de
Madjarova, Violeta	DKFZ	v.madjarova@dkfz-heidelberg.de
Malhotra, Samta	University of Heidelberg	samta.malhotra@gmail.com
Bastias Chian, Maria Jose	University Bordeaux	cotecin@yahoo.com
Mazur, Johanna	University Heidelberg	johanna.mazur@bioquant.uni-heidelberg.de
Nietert, Manuel	Medical Statistics University Goettingen	manuel.nietert@med.uni-goettingen.de
Raddatz, Guenter	IMB Tübingen	guenter.raddatz@tuebingen.mpg.de
Radde, Nicole	University Stuttgart	radde@ist.uni-stuttgart.de
Rahnenführer, Jörg	TU Dortmund	rahnenfuehrer@statistik.tu-dortmund.de
Raue, Andreas	University of Freiburg	andreas.raue@fdm.uni-freiburg.de
Rempel, Eugen	TU Dortmund	rempel@statistik.tu-dortmund.de
Rinck, Andrea	Technical University Munich	rinck@ipt.med.tum.de
Roeder, Ingo	IMB / Dresden University of Technology	ingo.roeder@tu-dresden.de
Schittler, Daniella	University Stuttgart	schittler@ist.uni-stuttgart.de
Silvia, von, der, Heyde	University Medical Center Göttingen	silvia.heyde@med.uni-goettingen.de
Stegle, Oliver	MPI for Biological Cybernetics	oliver.stegle@tuebingen.mpg.de
Sültmann, Holger	DKFZ	h.sueltmann@dkfz.de
Tresch, Achim	Gene Center, LMU Munich	tresch@lmb.uni-muenchen.de
Wangorsch, Gaby	Department of Bioinformatics, Würzburg	gaby.wangorsch@biozentrum.uni-wuerzburg.de
Weber, Michael	HKI Jena	michael.weber@hki-jena.de
Weber, Patrick	University of Stuttgart	patrick.weber@ist.uni-stuttgart.de
Wieczorek, Jakob	TU Dortmund	Jakob.Wieczorek@uni-dortmund.de
Wilczynski, Bartek	EMBL Heidelberg	wilczyns@embl.de
Wolkenhauer, Olaf	University Rostock	olaf.wolkenhauer@uni-rostock.de

Map



The workshop is taking place in the lecture hall of the communications center of the German Cancer Research Center (Im Neuenheimer Feld 280). Participants who are interested will meet for dinner at the Brauhaus Vetter (Steingasse 9) on Thursday October 21 at 8pm. If you plan to attend, please notify the conference organizers on the conference day.

To get to the Pub you can take the bus No. 32 at “Chirurgische Klinik” (opposite DKFZ) to “Universitätsplatz” - then it is only a short walk.