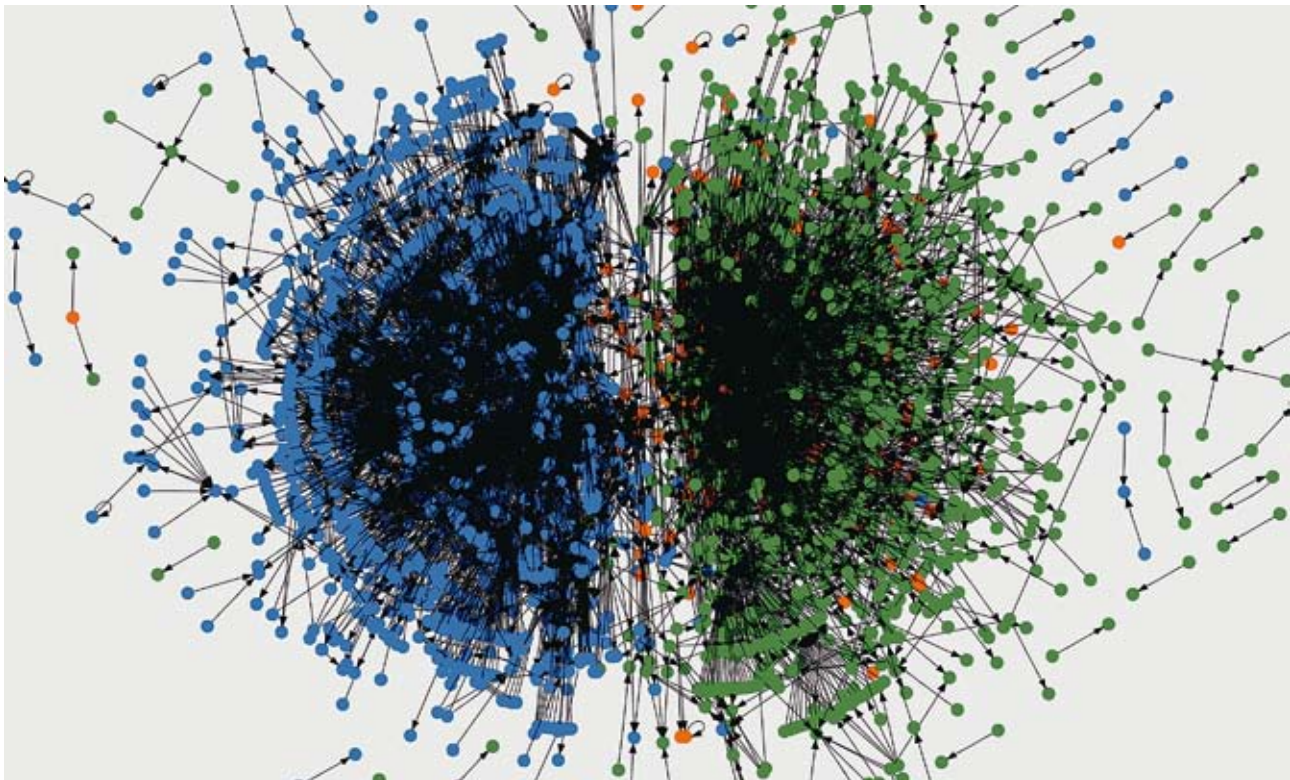


Zentrum Informatik, Statistik und Epidemiologie
Abteilung Bioinformatik
Centre for Informatics, Statistics and Epidemiology
Department of Bioinformatics



Visualisierung des Signaltransduktions- (grün) und des Stoffwechsellnetzwerkes (blau) von Säugerzellen
Visualization of the signal transduction (green) and the metabolic subnetwork (blue) in mammalian cells

Forschungsschwerpunkte Research Foci

- ▷ Analyse und Interpretation genregulatorischer Ereignisse
 - ▷ Modellierung und Analyse intra- und interzellulärer regulatorischer Netzwerke
 - ▷ CyteWalk – Datenintegration für die Systembiologie
 - ▷ Analysis and Interpretation of Gene Regulatory Events
 - ▷ Modelling and Analysing Intra- and Intercellular Regulatory Networks
 - ▷ CyteWalk – Data Integration for Systems Biology
-



Abteilungsdirektor Head of Department

Prof. Dr. rer. nat. Edgar Wingender

Kontaktdaten Contact

Abteilung Bioinformatik

UNIVERSITÄTSMEDIZIN GÖTTINGEN

Goldschmidtstraße 1, D-37077 Göttingen

Telefon +49-551 / 39-14912, Fax +49-551 / 39-14914

office@bioinf.med.uni-goettingen.de

www.bioinf.med.uni-goettingen.de

Hochschullehrer/innen Professors and Lecturers

+49-551 /

Wingender, Edgar	Prof. Dr. rer. nat.	edgar.wingender@bioinf.med.uni-goettingen.de	39-14912
------------------	---------------------	--	----------

Arbeitsgruppenleiter/innen Group Leaders

Crass, Torsten	Dr. rer. nat.	torsten.crass@bioinf.med.uni-goettingen.de	39-14919
----------------	---------------	--	----------

Haubrock, Martin	Dipl.-Inform.	martin.haubrock@bioinf.med.uni-goettingen.de	39-22764
------------------	---------------	--	----------

Michael, Holger (seit 01/2003)	Dr. rer. nat.	hom@bioinf.med.uni-goettingen.de	39-14918
-----------------------------------	---------------	----------------------------------	----------

Potapov, Antolij (seit 07/2004)	Dr. rer. nat.	anatolij.potapov@bioinf.med.uni-goettingen.de	39-14918
------------------------------------	---------------	---	----------

EINLEITUNG

Die Abteilung Bioinformatik beschäftigt sich zunehmend mit der Rekonstruktion und Analyse biologischer Netzwerke, einschließlich der Akquisition und Strukturierung des erforderlichen biologischen Wissens, als Beitrag zu systembiologischen Forschungsansätzen. Die durch diese Netzwerke repräsentierten Prozesse umfassen die Genregulation, intra- und interzelluläre Signalübertragung sowie Stoffwechselwege in multizellulären Eukaryoten. Es werden grundsätzliche theoretische Untersuchungen durchgeführt über die Eigenschaften verschiedenartiger Netzwerke im Hinblick auf deren Integration sowie ihre Anwendbarkeit bei der kausalen Interpretation entsprechender experimenteller Daten. Die aus diesen Arbeiten entstehenden Werkzeuge (Datenbanken, Programme) werden nach entsprechender Validierung als Internet-basierte Dienste zur Verfügung gestellt.

Diese Arbeiten werden durch eine Reihe von Drittmittelprojekten ergänzt: In einem integrativen Ansatz zur Modellierung zellulärer Signalwege und Kontrollprozesse wurde das p53-MDM2 Netzwerk hinsichtlich seiner topologischen Parameter charakterisiert (EU-Projekt COMBIO, bis 2007). Als Teil des Nationalen Genomforschungsnetzwerks (NGFN-2, bis 2008) wurden Arbeiten über eine integrierte Sicht auf verschiedene Arten biomolekularer Netzwerke (metabolische, Signal-, genregulatorische) durchgeführt. Brustkrebs-relevante regulatorische Netzwerke stehen im Zentrum des EU-Projektes Net2Drug (2007-2010), während das EU-geförderte Projekt LipidomicNet (2008-2011) intra- und (als Göttinger Beitrag) interzelluläre Signalübertragungen im Zusammenhang mit dem Lipidstoffwechsel bearbeitet. Die Überlagerung mehrerer regulatorischer Signale und deren Modellierung ist Gegenstand eines ETB-Projektes (GlobCell, 2008-2010), im Rahmen eines EFRE-Projektes werden neue Methoden zur Einbeziehung pathologischer Faktoren in Netzwerkmodelle entwickelt. In einem vom BMBF geförderten internationalen Kooperationsprojekt mit der Universität Nanjing in China wird die Erweiterung genregulatorischer Netzwerke um posttranskriptionelle Ereignisse (miRNAs) untersucht.

PREFACE

The Department of Bioinformatics focuses increasingly on the reconstruction and analysis of biological networks, including acquisition and structuring of the required biological knowledge, as contribution to systems biology research. The processes represented by these networks comprise gene regulation, intra- and intercellular signalling, as well as metabolic pathways in multicellular eukaryotes. Basic theoretical investigations on the properties of different kinds of networks are carried out with regard to their integration as well as their applicability on the causal interpretation of corresponding experimental data. After validation, the tools that result from these efforts (databases, programs) are made available as Internet-based services.

These activities are complemented by a number of external grant projects: In an integrative approach to modelling cellular sig-

nalling pathways and control processes, the p53-MDM2 network was characterized (EU project COMBIO, until 2007). As part of the National Genome Research Network (NGFN-2, until 2008), efforts were made to provide an integrated view on different biomolecular networks (metabolic, signalling and gene regulation). Networks relevant for breast cancer are in the focus of the EU project Net2Drug (2007-2010), the EU-funded project LipidomicNet (2008-2011) aims at intra- and (as Goettingen's contribution) intercellular signalling in the context of lipid metabolism. Superposition of several regulatory signals and their modelling is the objective of an ETB project (GlobCell, 2008-2010) and in the context of an EFRE project new methods for the inclusion of pathological factors into network models are developed. In an international collaborative project, which is funded by BMBF, the extension of gene regulatory networks by posttranscriptional factors is studied.

1. Analyse und Interpretation genregulatorischer Ereignisse

Die Analyse und Interpretation genregulatorischer Mechanismen ist ein traditioneller Forschungsschwerpunkt unserer Abteilung. Ein Organismus enthält im Allgemeinen gleichzeitig den Bauplan und das nötige Regelwerk, um alle notwendigen biologischen Prozesse ausgestalten zu können, die die Entwicklung und das Leben eines Organismus ermöglichen. Dabei spielt die Ebene der Transkription, also die Steuerung der Aktivität der Gene eines Organismus, eine entscheidende Rolle. In den Arbeiten unsere Abteilung auf diesem Gebiet konzentrieren wir uns vor allem auf die Analyse von DNA-bindenden Proteinen, den so genannten Transkriptionsfaktoren (TF), und deren Interaktionsstellen im Genom, den Transkriptionsfaktorbindestellen (TFBS). Die Vorhersage dieser TFBS ist dabei eine besondere bioinformatische Herausforderung, da bei den verschiedenen eingesetzten Methoden, die Anzahl an falschpositiven Vorhersagen sehr groß sein kann. Um die Vorhersagequalität, also die Anzahl falscher Vorhersagen, zu minimieren, versuchen wir, zusätzliche Information mit in die Vorhersage einzubeziehen. Im Wesentlichen wurden dazu bisher zwei biologische Phänomene ausgenutzt: Zum einen kann die Vorhersage von TFBS verbessert werden, indem Kontexteigenschaften mit einbezogen werden, z. B. die Vergesellschaftung mit weiteren TFBS. Zum anderen kann die Berücksichtigung von evolutionären Zusammenhängen die Vorhersage unterstützen. Diese Methode wird „phylogenetisches Footprinting“ genannt und nutzt die Konservierung der TFBS in verschiedenen Spezies aus. Sie beruht darauf, dass die Bindestellen mancher TF im Verlauf der Evolution konserviert wurden. Diese Tatsache kann in die Vorhersage mit einfließen und diese so verbessern.

Da wir bisher auf in Datenbanken gespeichertes Wissen zurückgreifen, wird einer der nächsten Schritte bei der Analyse genregulatorischer Mechanismen sein, auch potentielle, bisher unbekannte TFBS vorherzusagen. Eine weitere Herausforderung wird sein, so genannte Hochdurchsatzverfahren (wie z.B. Microarrays) mit einer qualitativ hochwertigen Promotoranalyse zu verknüpfen.

1. Analysis and Interpretation of Gene Regulatory Events

The analysis and interpretation of gene regulatory events is a traditional research area of our department. An organism usually contains the whole blueprint as well as the control equipment to complete all essential biological processes needed for developing and living. One important level in this context is transcription and its regulation. It is known that the regulation of gene expression is performed by regulatory proteins (transcription factors, TFs) that interact with specific DNA elements in regulatory regions of the genes (e.g., promoters or enhancers). These elements are called transcription factor binding sites (TFBS). We are working on high-quality prediction methods for these TFBS. Due to the high degree of degeneracy of typical TFBS and the resulting danger of false-positive predictions, this work presents a big challenge. To enhance the quality of our predictions, we aim to include additional information into our analysis steps. For example, we have attempted to include certain context features such as the association with other TFBS into the predictions. The integration of this information can help to minimize the fraction of false-positive predictions. Another method attempts to make use of phylogenetic information since it is known that for some genes, the TFBS are conserved among different species as their function is vital for their survival (phylogenetic footprinting). We are able to detect these conserved regions and can use them for better predictions.

Within our analyses we make use of collected knowledge which is stored in different databases so far. One of our next objectives is to predict potential binding sites for as yet unknown factors. Another important step is to integrate our predictions into different high-throughput technology workflows (like microarray analysis) in order to identify the combinatorial transcription logic in a set of co-regulated genes.

Arbeitsgruppenleiter/innen Group Leaders

Dipl.-Inform. Martin Haubrock

Kooperationen Cooperations

Dr. Rossner, Abteilung Neurogenetik, Max-Planck-Institut für Experimentelle Medizin, Göttingen; BIOBASE GmbH, Wolfenbüttel

Drittmittelförderung Funding

European Commission, FP6 Programme, Net2Drug, "Life sciences, genomics and biotechnology for health", LSHG-CT-2007-037590, 2007-2010

Ausgewählte Publikationen Selected Publications

Waleev T, Shtokalo D, Kononova T, Voss N, Cheremushkin E, Stegmaier P, Kel-Margoulis O, Wingender E, Kel A (2006) Composite Module Analyst: identification of transcription factor binding site combinations using genetic algorithm. *Nucleic Acids Res*, 34: W541-5.

Kel A, Kononova T, Waleev T, Cheremushkin E, Kel-Margoulis O, Wingender E (2006) Composite Module Analyst: a fitness-based tool for identification of transcription factor binding site combinations. *Bioinformatics*, 22: 1190-7.

Kel A, Voss N, Jauregui R, Kel-Margoulis O, Wingender E (2006) Beyond microarrays: Find key transcription factors controlling signal transduction pathways. *BMC Bioinformatics*, 7 Suppl. 2, S13.

Chen X, Wu J-m, Hornischer K, Kel A, Wingender E (2006) TiProD: The Tissue-specific Promoter Database. *Nucleic Acids Res*, 34: D104-7.

Sauer T, Shelest E, Wingender E (2006) Evaluating phylogenetic footprinting for human-rodent comparisons. *Bioinformatics*, 22: 430-7.

Matys V, Kel-Margoulis O, Fricke E, Liebich I, Land S, Barre-Dirrie A, Reuter I, Chekmenev D, Krull M, Hornischer K, Voss N, Stegmaier P, Lewicki-Potapov B, Saxel H, Kel AE, Wingender E (2006) TRANSFAC® and its module TRANSCOMP®: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res*, 34: D108-10.

Wingender E (2008) The TRANSFAC project as an example of framework technology that supports the analysis of genomic regulation. *Brief Bioinform*, 9: 326-32.

Kel A, Voss N, Valeev T, Stegmaier P, Kel-Margoulis O, Wingender E (2008) Explain: finding upstream drug targets in disease gene regulatory networks. *SAR QSAR Environ Res*, 19: 481-94.

2. Modellierung und Analyse intrazellulärer und interzellulärer regulatorischer Netzwerke

Im EU-Projekt COMBIO konzentrierten wir uns auf das Modellieren des p53-MDM2 Regulations-Netzwerks, das Zellproliferation, Differenzierung und Krebsentwicklung steuert. Wir rekonstruierten ein Signaltransduktions-Netzwerk für p53 sowie verschiedene Varianten davon und analysierten deren topologische Eigenschaften. Zusätzlich extrahierten wir das Transkriptionsnetzwerk von p53 und analysierten seine topologischen Eigenschaften. Diese Netzwerke wurden mit den entsprechenden ‚vollständigen‘ Signaltransduktions- und Transkriptions-Netzwerken als Referenzen verglichen. Die p53-Netzwerke erwiesen sich als grundsätzlich nicht-zufällig und inhomogen. Die Konnektivitätsverteilungen ihrer Topologien werden gut durch ein skalen-freies Modell beschrieben. Die Netzwerke zeigen eine Abhängigkeit des Clustering-Koeffizienten vom Konnektivitätsgrad der Knoten, was auf das Vorhandensein einer hierarchischen Modularität hindeutet. Vor Kurzem haben wir ein neues Maß (*pairwise disconnectivity index*) für den Einfluss einzelner Netzwerkkomponenten auf den Zusammenhalt eines Netzwerkes entwickelt, das sich bereits als extrem wertvoll erwiesen hat bei der Identifizierung von einzelnen Genen, Molekülen, oder ganzen Subnetzwerken, die eine Schlüsselfunktion bei bestimmten biologischen Prozessen haben.

Durch das Projekt EndoNet erweitern wir unsere Bemühungen, intrazelluläre Regulations-Prozesse zu modellieren, auf die Darstellung der Kommunikation zwischen Zellen und Organen. EndoNet ist eine neue Datenbank, die Informationen über Komponenten endokriner Netzwerke und ihre Beziehungen zueinander enthält. Es konzentriert sich auf endokrine Zell-Zell-Signalwege und ermöglicht die Analyse interzellulärer Regulations-Netzwerke beim Menschen. Im EndoNet-Datenmodell überspannen zwei Klassen von Komponenten einen bipartiten gerichteten Graphen. Eine Klasse bilden die Hormone (im weitesten Sinne), welche durch definierte Donorzellen sekretiert werden. Die andere Klasse besteht aus den Akzeptor- oder Zielzellen, welche die entsprechenden Hormonrezeptoren exprimieren. Deren Identität und die anatomische Umgebung aus Zellen, Geweben und Organen wird durch Verweise auf die ebenfalls von uns entwickelte Datenbank CYTOMER definiert. EndoNet strebt an, die vorhandene Lücke zwischen bekannten Genotypen und ihren molekularen und klinischen Phänotypen zu füllen, und

eine Anwendung dieses Wissens in der medizinischen Forschung zu ermöglichen.

Im Rahmen des EU-geförderten Verbundprojektes LipidomicNet adaptieren wir EndoNet auf alle interzellulären Kommunikationsvorgänge, die für den Lipidstoffwechsel eine Rolle spielen. Dies können hormonelle Regulationsprozesse sein, die den Fettstoffwechsel in den Zielzellen beeinflussen, oder auch durch Lipide vermittelte Kommunikationswege zwischen verschiedenen Zellen oder Organen.

In Zusammenarbeit mit Prof. J. Wang von der Universität Nanjing, China, haben wir ein Projekt zur Integration von microRNAs in das bei uns bereits vorhandene Transkriptions-Netzwerk begonnen, um ein erweitertes Genregulations-Netzwerk zu erhalten, dessen Eigenschaften dann untersucht werden sollen. Zu diesem Zweck machen wir Gebrauch von unseren neuesten Erkenntnissen bezüglich der Vorhersage von Transkriptionsfaktor-Bindungsstellen, wie sie sich aus dem Forschungsschwerpunkt 1 ergeben.

2. Modelling and Analysing Intracellular and Intercellular Regulatory Networks

Within the EU project 'COMBIO', we focused on modelling the p53-MDM2 regulatory network which controls cell proliferation, differentiation and cancer development. We reconstructed the mammalian signal transduction p53 network, as well as its different subvariants, and analyzed their topological properties. In addition, we extracted and studied the topological features of the p53 transcription network. These p53 networks were compared with the corresponding 'complete' signal transduction network and 'complete' transcription factor network used as reference networks. The p53 signal transduction and transcription networks have turned out to be fundamentally non-random and inhomogeneous. Their topology follows a power-law degree distribution and is best described by the scale-free model. The networks show the dependence of the clustering coefficient on the degree of a vertex, thereby indicating the presence of hierarchical modularity. More recently, we have developed a new metric for the impact of network components on the coherence of the graph (*pairwise disconnectivity index*) which has proved to be extremely useful in identifying individual genes, molecules or subnetworks, that play a key role in certain biological processes.

The project 'EndoNet' extends our efforts in modelling intracellular regulatory processes to characterizing communication between cells and organs. EndoNet is a new database concerning the components of endocrine networks and their relations among each other. It focuses on the endocrine cell-cell signalling and enables the analysis of human intercellular regulatory pathways. In the EndoNet data model, two classes of components span a bipartite directed graph. One class represents the hormones (in the broadest sense) secreted by defined donor cells. The other class consists of the acceptor or target cells expressing the corresponding hormone receptors. The identity and anatomical environment of cell types,

tissues and organs is defined through references to the CYTOMER ontology. EndoNet aims at bridging the existing gap between known genotypes and their molecular and clinical phenotypes, thus allowing the utilization of EndoNet in medical research.

In the context of the EU-funded research consortium LipidomicNet, we adapt EndoNet to all intercellular communication processes that are relevant for the lipid metabolism. These can be hormonal regulatory paths that influence the lipid metabolism in the corresponding target cells, or lipid-mediated communication paths between different cells or organs.

In collaboration with Prof. Wang (Nanjing University, China), a project has been launched to include miRNAs and their genes in the existing transcription network and to analyze the properties of the resulting graph. For this purpose, we make use of the latest state-of-the-art of TFBS prediction, as developed in research topic 1 (see above).

Arbeitsgruppenleiter/innen Group Leaders

Dr. Anatolij P. Potapov

Kooperationen Cooperations

Biobase GmbH, Wolfenbüttel; LipidomicNet-Konsortium; S. Miyano, Tokyo University; J. Wang, Nanjing University;

Drittmittelförderung Funding

European Commission, FP6 Programme, COMBIO, "Life, sciences, genomics and biotechnology for health", LSHG-CT-2004-503568, 2004-2007

BMBF und Forschungs- und Berufungspool des Niedersächsischen Ministeriums für Wissenschaft und Kultur, Kap. 06, 08 TG 74

NGFN, grant NGFN2 - Biomed - SMP - Bioinformatik, FKZ: 01GRO-480, 2005-2008

European Commission, FP7 Programme, LipidomicNet, "Life, sciences, genomics and biotechnology for health", HEALTH-F4-2008-202272, 2008-2012

Ausgewählte Publikationen Selected Publications

Dönitz J, Goemann B, Lizé M, Michael H, Sasse N, Wingender E, Potapov A. P (2008) EndoNet: An information resource about regulatory networks of cell-to-cell communication. *Nucleic Acids Res*, 36: D689-94.

Michael H, Hogan J, Kel A, Kel-Margoulis O, Schacherer F, Voss N, Wingender E (2008) Building a knowledge base for systems pathology. *Brief Bioinform*, 9: 518-31.

Potapov A. P, Goemann B, Wingender E (2008) The pairwise disconnectivity index as a new metric for the topological analysis of regulatory networks. *BMC Bioinformatics*, 9: 227.

Zubarev R A, Nielsen M L, Fung E M, Savitski M M, Kel-Margoulis O, Wingender E, Kel A (2008) Identification of dominant signaling pathways from proteomics expression data. *J Proteom*, 71: 89-96.

Krull M, Pistor S, Voss N, Kel A, Reuter I, Kronenberg D, Michael H, Schwarzer K, Potapov A, Choi C, Kel-Margoulis O, Wingender E (2006) TRANSPATH®: an Information resource for storing and visualizing signaling pathways and their pathological aberrations. *Nucleic Acids Res*, 34: D546-51.

Potapov A, Liebich I, Dönitz J, Schwarzer K, Sasse N, Schoeps T, Crass T, Wingender E (2006) EndoNet: An information resource about endocrine networks. *Nucleic Acids Res*, 34: D540-5.

Potapov A P, Wingender E (2006) Mining the genome and regulatory networks. *Genome Biol*, 7: 309.

3. Cytewalk - Datenintegration für die Systembiologie

Eine systematische Betrachtung lebender Organismen bedarf der engen Integration einer Vielzahl heterogener Informationen über die Bausteine, aus denen sich der Organismus zusammensetzt, sowie über deren Interaktionen. Als ersten Schritt in Richtung einer ver-

einheitlichenden, Ontologie-basierten Sichtweise auf die Entitäten, aus denen sich metabolische, Signaltransduktions- und genregulatorische Netzwerke zusammensetzen, sowie auf deren wechselseitige Beziehungen, haben wir ein graphenbasiertes Datenmodell („Entitäts-Interaktions-Graphen“) entwickelt, das die Semantik all dieser Typen biomolekularer Interaktionsnetzwerke wiederzugeben vermag. Unser Kooperationspartner BIOBASE GmbH (Wolfenbüttel) hat dieses Datenmodell bereits in seine TRANSPATH®-Datenbank über Signaltransduktionswege übernommen.

Die PATHWAY-Abteilung des KEGG-Datenbanksystems hingegen ist sicherlich eine der ausführlichsten Informationsressourcen zum Thema Stoffwechselwege und wird von zahlreichen Bioinformatik-Anwendungen zu Rate gezogen. Während des Versuchs, auch die KEGG/PATHWAY-Daten in unser Datenmodell zu transformieren, mussten wir jedoch feststellen, dass diese Datenbank ein implizites hierarchisches Klassifikationsschema für die enthaltenen Reaktanden annimmt, ohne dieses Schema in einer formalen Weise – etwa als Bestandteil des Datenbankschemas – anzugeben. Als Folge davon kann eine unvollständige Rekonstruktion metabolischer Netzwerke auftreten, in denen manche durchaus mögliche Reaktionssequenzen fehlen. Daher haben wir eine Methode entwickelt, um die implizite KEGG-Klassifikation chemischer Substanzen durch Vergleich der Strukturen aller möglichen Substrate eines jeden Enzyms zu erschließen. Wir konnten zeigen, dass das auf diese Weise gewonnene zusätzliche Wissen zur Verbesserung der Netzwerkrekonstruktion genutzt werden kann. Weiterhin erlaubt unser Ansatz die Ermittlung redundanter Reaktionen bei der grafischen Repräsentation von Netzwerken, was zu übersichtlicheren Darstellungen führt.

Zur Zeit erarbeiten wir eine generalisierte Vorgehensweise für die Abbildung beliebiger OWL-basierter Ontologien in objekt-relationalen Datenbanken. Diese Technologie wird verwendet zur Integration von Daten aus KEGG, TRANSPATH und anderen Informationsquellen zu biomolekularen Interaktionsnetzwerken im Rahmen einer dem Entitäts-Interaktions-Graphenmodell entsprechenden Ontologie. Sie kann aber auch in anderen Feldern, etwa der Medizin oder dem Geschäftsleben, Anwendung finden.

3. Cytewalk - Data Integration for Systems Biology

A systemic view on living organisms requires tight integration of a vast amount of heterogeneous information on the organism's building blocks as well as their interactions. As a first step towards a unified, ontology-based view on the entities comprising metabolic, signalling and gene regulatory pathways and their mutual relationships (project Cytewalk), we defined a common, graph-based data model capable of capturing the semantics of all these kinds of biomolecular interaction networks. Our cooperation partner BIOBASE GmbH (Wolfenbüttel) already has re-engineered their TRANSPATH® database on signalling pathways according to this model.

The PATHWAY section of the KEGG database system, on the other hand, is surely one of the most comprehensive informa-

tion resources on metabolic pathways and gets utilized by a large number of bioinformatics applications. However, while attempting to also transform KEGG to our data model, we realized that this database implicitly assumes reactants to be classified into a chemical compound hierarchy, but fails to provide this information in a formal manner, e. g. as intrinsic part of the database schema. As a consequence, this may lead to the reconstruction of incomplete networks, lacking certain possible reaction sequences. Hence we developed a method to make KEGG's implicit compound classification explicit through comparing the chemical structures of each enzyme's possible substrates and demonstrated that the additional knowledge thus gained can be used to improve metabolic network reconstruction. Furthermore, our approach allows to reduce the amount of redundant reactions displayed for each enzyme when rendering the network in a graphical fashion, leading to a clearer visual representation.

We are currently preparing a generalized procedure for mapping arbitrary OWL-based ontologies onto object-relational databases. This technology will be used for integrating data from KEGG, TRANSPATH and other sources of information on biomolecular interactions within an ontology corresponding to our entity-interaction graph model, but will also be applicable for many more fields of interest, e.g. medical or business applications.

Arbeitsgruppenleiter/innen Group leaders

Dr. Torsten Crass

Kooperationen Cooperations

BIOBASE GmbH, Wolfenbüttel; V. Soumelis, Institut Curie, Paris; Helios Biosciences, Paris

Ausgewählte Publikationen Selected Publications

Degenhardt J, Haubrock M, Wingender E, Crass T (2007) DEEP - A Tool for Differential Expression Effector Prediction. *Nucleic Acids Res*, 35: W619-24.

Wingender E, Crass T, Hogan J D, Kel A, Kel-Margoulis O, Potapov A P (2007) Integrative content-driven concepts for bioinformatics "beyond the cell". *J Biosci*, 32: 169-80.

Wingender E, Hogan J, Schacherer F, Potapov A P, Kel-Margoulis O (2007) Integrating pathway data for systems pathology. *In Silico Biol*, 7 S1: 03.

Seibel P N, Kruger J, Hartmeier S, Schwarzer K, Löwenthal K, Mersch H, Dandekar T, Giegerich R (2006) XML schemas for common bioinformatic data types and their application in workflow systems. *BMC Bioinformatics*, 7: 490.

Anhang Appendix

Medizinische Dissertationen (Dr. med.; Dr. med. dent.)

Doctorate Theses (Dr. med.; Dr. med. dent.)

Bleckmann A, Dr. med., Struktur-Aktivitäts-Relationen kleiner Moleküle am Beispiel von Ephothon und mGluR5-Modulatoren. Dissertation Universität Göttingen 2008.

Naturwissenschaftliche und andere Dissertationen (Dr. rer. nat. und andere)

Doctorate Theses (Dr. rer. nat. and others)

Sauer T, Dr. rer.nat., Evaluierung des phylogenetischen Footprintings und dessen Anwendung zur verbesserten Vorhersage von Transkriptionsfaktor-Bindestellen. Dissertation Georg-August-Universität Göttingen 2006.

Shelest E, Dr. rer. nat., Genetic networks of antibacterial responses of eukaryotic cells. Bioinformatics analysis and modeling. Dissertation Technische Universität Braunschweig 2006.

Diplom- und Masterarbeiten Diploma and Master Theses

Bolivar JC, Dipl.-Biol., Analyse und Bewertung der Auswirkungen von mutierten Proteinen in Signaltransduktionskaskaden auf die Topologie biomolekularer Netzwerke und resultierende Phänotypen durch Integration von Daten über pathologische Genotypen. Diplomarbeit Georg-August-Universität Göttingen 2009.

Vogler H, Dipl.-Biol., Vorhersage phänotyp-spezifischer Effektor-Moleküle durch Verknüpfung von Genexpressionsdaten mit Informationen über biomolekulare Netzwerke. Diplomarbeit Georg-August-Universität Göttingen 2008.

Riczane R, MSc, OntoScope: Interaktive Visualisierung einer anatomischen Ontologie. Masterarbeit Georg-August-Universität Göttingen 2008.

Yan D, MSc, Phylogenetic Conservation of MicroRNA Genes and Targets. Masterarbeit Georg-August-Universität Göttingen 2007.

Wissenschaftliche Tagungen Scientific Meetings

07.03.-21.03.2006, E. Wingender visiting professor at Tokyo Medical and Dental University, Tokyo, Japan

19.03.-31.03.2006, A. Potapov visiting scientist at Tokyo Medical and Dental University, Tokyo, Japan

21.03.2006, Workshop on Endocrine Networks, Tokyo, Japan (co-organization; lecture)

04.-05.04.2006, 2nd All-Wales Bioinformatics Workshop, Cardiff, Wales/UK: From TRANSFAC to EndoNet: moving ahead from promoter analysis to systems biology (E. Wingender, invited lecture)

19.09.2006, Workshop on OMICS Informatics, Tokyo, Japan (invited lecture)

30.04.-02.05.2007, Bio IT World, Boston/MA: Biology Explained: Making Sense of High-Throughput Gene Expression Data (E. Wingender, talk)

Mitgliedschaften und Mitarbeit in wissenschaftlichen Gremien und Kommissionen Memberships and Activities in Scientific Boards and Committees

Prof. Dr. Edgar Wingender

Scientific Advisory Board "CUBIC" (Cologne University Bioinformatics Center)

Herausgebertätigkeit Editorial Work

Prof. Dr. Edgar Wingender

Editor-in-Chief of In Silico Biology

Dr Holger Michael

Associated Editor, In Silico Biology

Internationale wissenschaftliche Kooperationen

International Scientific Cooperations

Dr. Xin Chen, College of Life Sciences, Beijing University, China

Dr. Jin Wang, Dept of Biochemistry, Nanjing University, China

Dr. Takako Takai-Igarashi, Tokyo Medical and Dental University, Tokyo, Japan

EU-Projekte European Research Projects

COMBIO, LSHG-CT-2004-503568, 2004-2007

Net2Drug, LSHB-CT-2007-037590

LipidomicNet, HEALTH-F4-2008-202272, 2008-2012

Gastwissenschaftler/innen Guest Scientists

Dr. Xin Chen, College of Life Sciences Peking University, Beijing 100871, P.R. China, 08/2007

Prof. Dr. Jin Wang, Nanjing University, Department of Biochemistry, Hankou Road 22th, Nanjing 210093 P.R. China, 09-10/2008,

Dr. Takako Takai-Igarashi, Tokyo Medical and Dental University, Tokyo, Japan, 10/2006

Firmenkooperationen Industrial Cooperations

BIOBASE GmbH, Wolfenbüttel

Helios Bioscience, Paris, Frankreich