

Systems biology

A biased perspective



Maik Kschischo

RheinAhrCampus Remagen

University of Applied Sciences Koblenz

[*kschischo@rheinahrcampus.de*](mailto:kschischo@rheinahrcampus.de)

Current address:

Cancer Research UK, London Research
Institute, Gene Expression Laboratory



RheinAhrCampus Remagen (UAS Koblenz)

- Department of **Mathematics and Technology**
- **Biomathematics (B.Sc.)**
- **Mathematics in Finance / Life Science (M.Sc.)**

www.rheinahrcampus.de



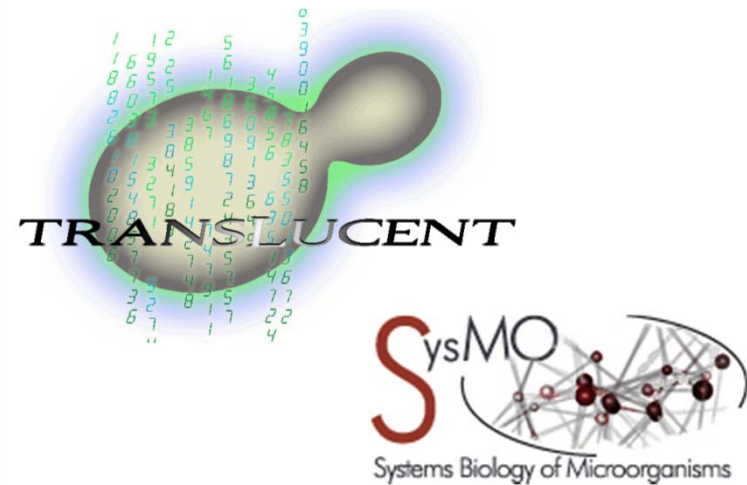
Current research projects (Kschischo)

Cancer systems biology

- in collaboration with **Charles Swanton** and **Almut Schulze** (CRUK London)
- Genetic instability in cancer
- Prediction of drug resistance
- Cancer metabolism

Mathematical modelling of ion homeostasis in yeast (*S.c*)

- Era-Net



<http://ptjapps.fz-juelich.de/sysmo>

Outline

I. Why systems biology?

II. The role of mathematical models

- **Statistical models**
- **Mechanistic models**

III. Data management

- **The need for integrated biological and medical data**
- **Data standards and data management**

IV. Some links and literature

Systems biology

- **A paradigm**

*"Systems biology...is about **putting together rather than taking apart**, integration rather than reduction. It requires that we develop ways of thinking about integration that are as rigorous as our reductionist programmes, but different....It means changing our philosophy, in the full sense of the term"*

Denis Noble (2006). *The Music of Life: Biology beyond the genome*. Oxford University Press. p21

- **A way of doing research**

- **Systems biology circle** between experimentation and modelling

The goal

- Make **useful predictions**
- Guide better understanding

Example:

- Oncotype DX (Breast)
 - Gene signature (21 genes)
 - Predicts **response to chemotherapy** in combination with traditional diagnostic scores like node status, ER status, tumor grade,...
- MammaPrint 70-gene signature

Systems biology and synthetic biology

Synthetic biology:

- **Design and fabrication of biological components** and systems that do not already exist in the natural world
- Re-design and fabrication of existing biological systems
- **Engineer a biological system** to produce a certain substance or to fulfill a certain function

Systems biology:

- Studies **natural systems**

The role of mathematical models

Is math useful at all?

Modelling

George Edward Pelham Box:

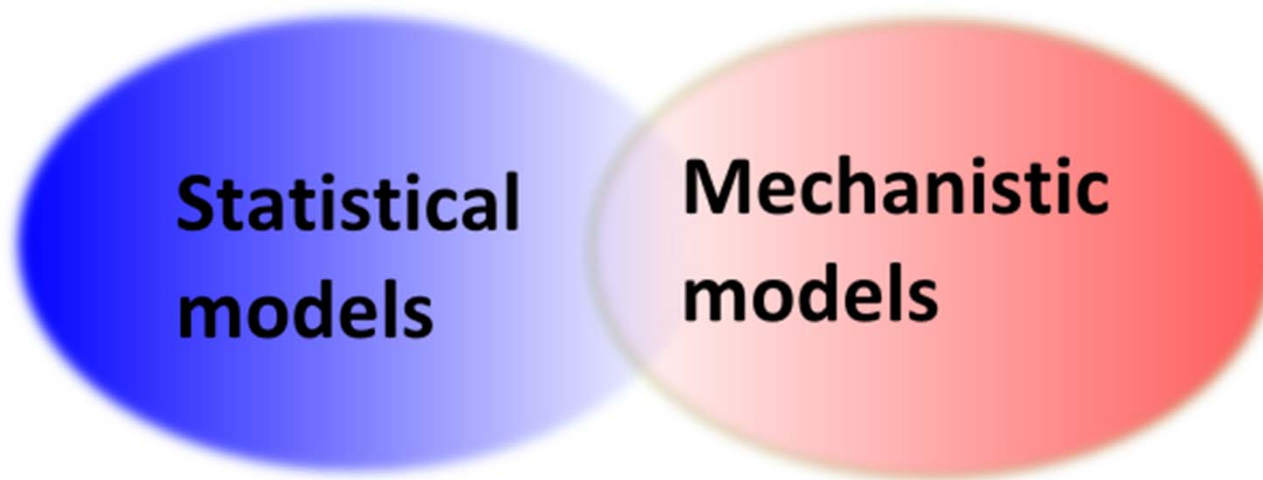
“Essentially, all models are wrong, but some are useful.”

Box, George E. P.; Norman R. Draper (1987). Empirical Model-Building and Response Surfaces. Wiley. pp. 688, p. 424.



Which type of model?

- It depends on the
 - Biological question
 - Data



Example 1: Gene expression signatures and statistical models

Biological question:

- Which genes can be used for predicting a certain phenotype (e.g. survival or response to therapy)?

Data:

Patient	Gene 1	Gene 2	...	Gene 10000	Survival	Node	ER+	...
1	1.3	7.6	...	0.9	3	positive	Yes	...
2	1.7	1.9	...	2.1	1	negative	No	...
3	3.8	4.2	...	3.7	censored	positive	Yes	...
...

Statistical model

Survival = $f(\text{gene1}, \dots, \text{gene1000}, \text{node}, \text{ER}+, \dots) + \text{noise}$

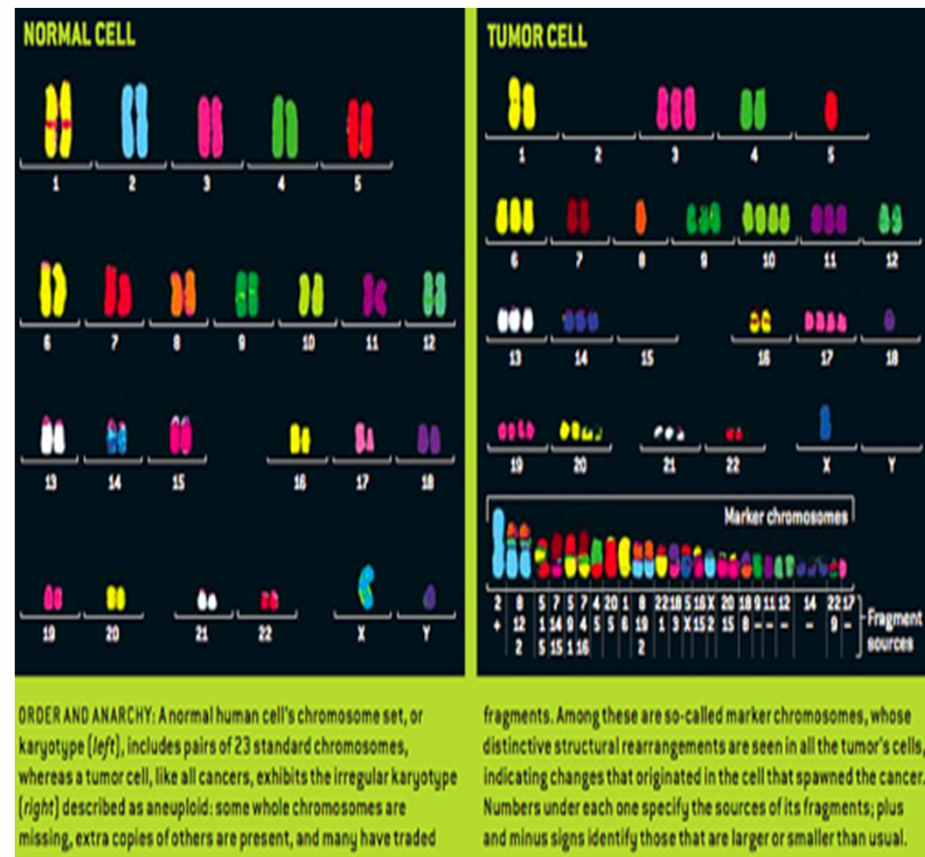
Challenges:

- Small sample size, large number of variables
(small n, large p)
- Need an independent training set and a test set

Example 2 from own research: Taxane response signature

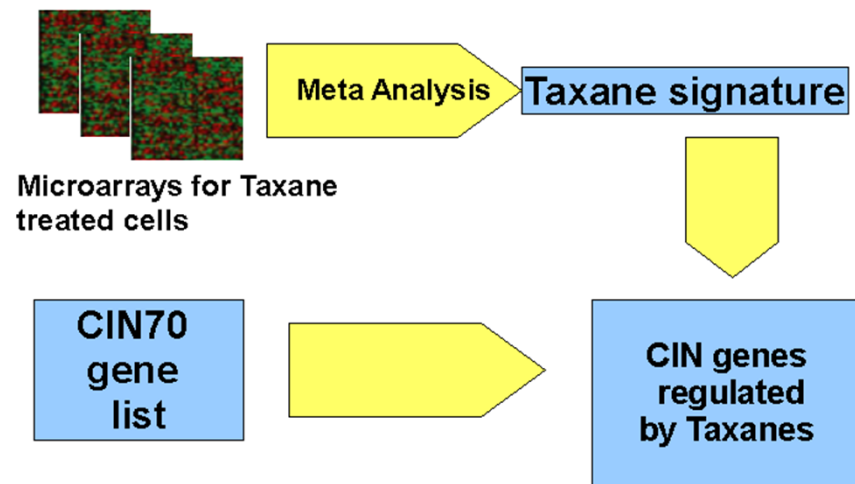
Swanton et al. Proc Natl Acad. Sci. USA, 106, 2009

- Many cancer cells are aneuploid and chromosomally instable (CIN)
- CIN: Increased rate gains/losses
- CIN cells are often more resistant against drugs
- Why?

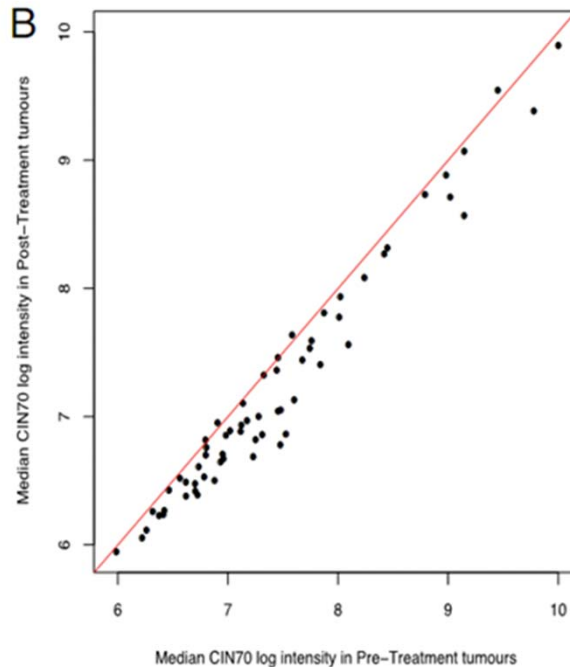


Strategy for finding the gene signature

- Generate a list of genes affected by taxanes (Taxane response gene signature)
- Analyze the relationship of genes in the Taxane response signature to genes involved in CIN
- Use the CIN70 gene signature of genes overexpressed in CIN cells from *Carter et al., Nat Genet, (2006)*



Main results

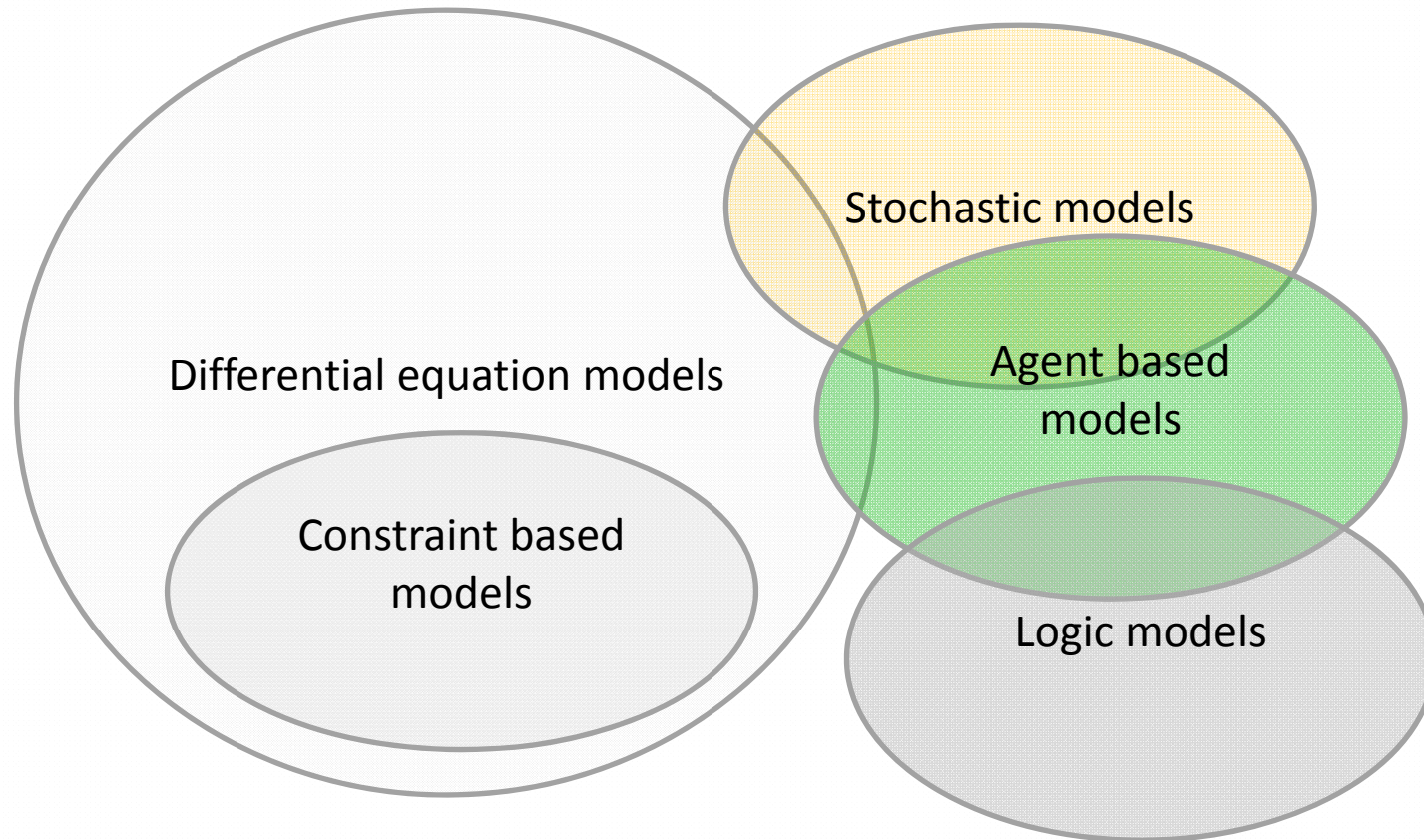


**Expression of taxane signature
in ovarian carcinomas before
and after 3 cycles of paclitaxel
treatment**

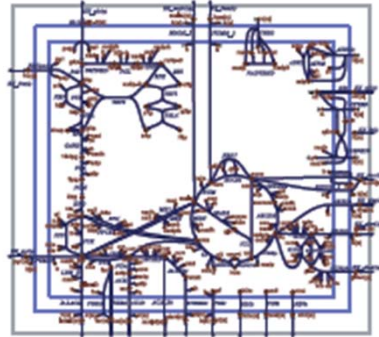
(Independent test set!!!)

- Genes overexpressed in CIN cells are downregulated by taxanes
- Genes with overexpressed CIN70 signature are more resistant to taxane treatment

Some mechanistic models



Deterministic models for reaction networks



Reaction 1



Reaction 2

...

Reaction n

*Orth et al.
Nat. Biotech,
2010*

Ordinary differential equations (ODEs)

Predictions:

- Time courses of concentrations

Requirements

- Time course data for many reactants (e.g. proteins)
- Enzyme kinetic data

Applications:

- Small reaction networks

Constraint based models (COBRA)

Predictions:

- *Stationary reaction fluxes (metabolic networks)*

Requirements

- *Constraints on certain reaction rates*

Applications:

- Genome scale metabolic networks

A nice example for a constraint based model

Network-based prediction of human tissue-specific metabolism

Tomer Shlomi^{1,4}, Moran N Cabili^{1,4}, Markus J Herrgård², Bernhard Ø Palsson² & Eytan Ruppin^{1,3}
Nat Biotech, 2008

Data

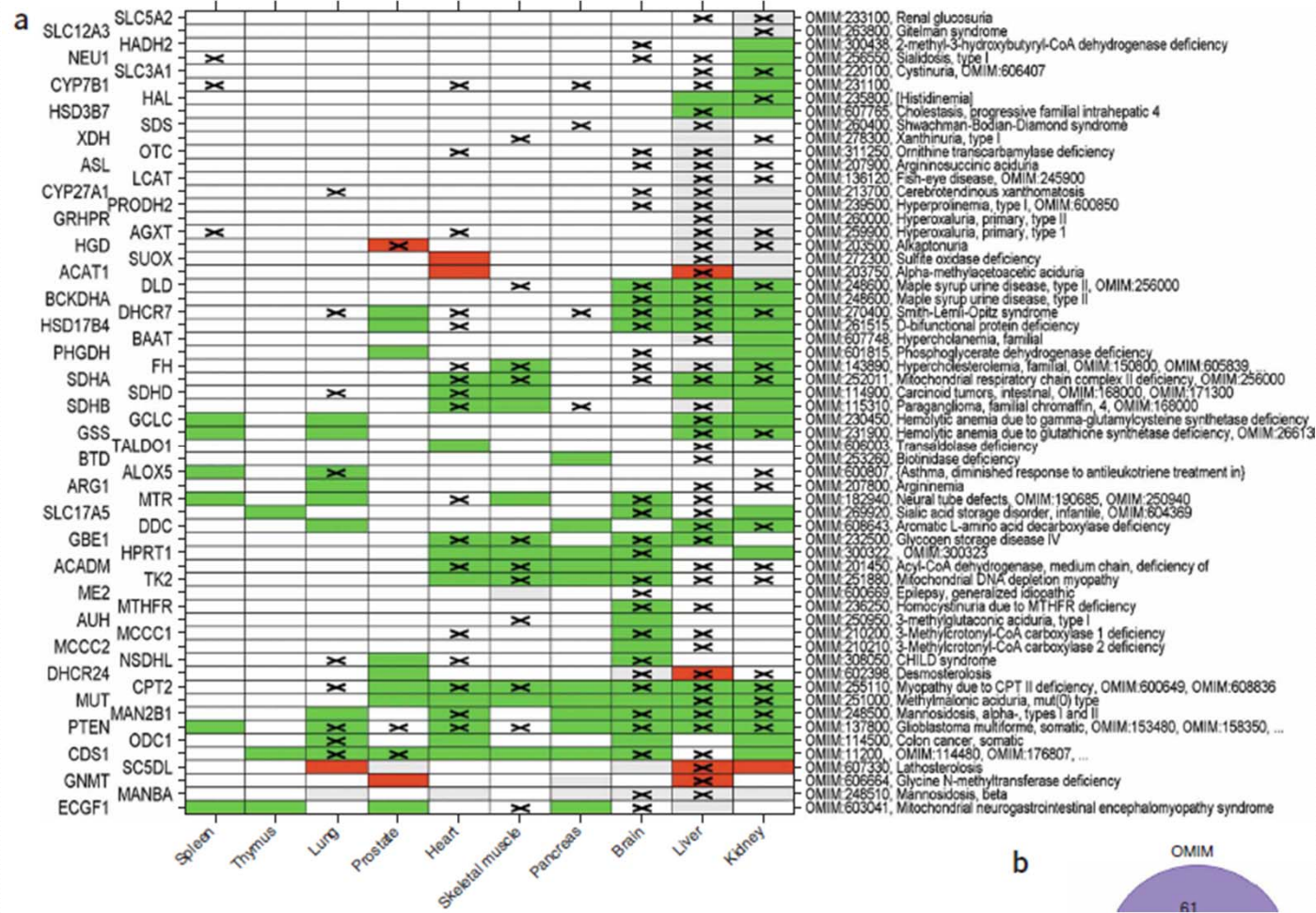
- Reconstructed human metabolic network (*Duarte et al. Proc. Natl. Acad. Sci. 2007*)
- Gene and protein expression data for different tissue types

Goal:

- Predict and compare tissue specific activity of disease related genes

Comparison of predicted and measured gene activity of disease related genes

Shlomi et al, 2008



Data management

**Its a big mess! But it gets better.
Some Examples.**

Problems in biological data management and acquisition

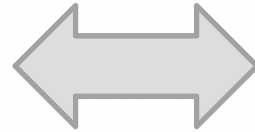
- Heterogeneity and complexity of the data
- Lack of well established data standards
- **Link to medical data**

Simple example: Meta analysis of breast cancer microarrays

- Used 10 large data sets (more than 1000 patients)
- Only 8 studies report the ER status

My wish list for medical systems biology

Biological data



Medical data
(Patients)

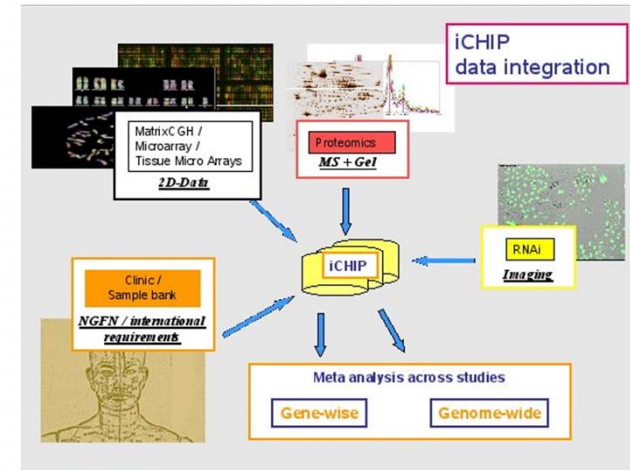
Matching of high throughput data and clinical
diagnostics or outcome

NCBI is on the starting blocks

The screenshot shows the NCBI website in a Mozilla Firefox browser window. The browser's address bar displays the URL <http://www.ncbi.nlm.nih.gov/>. The page features a search bar at the top with the text "Search All Databases" and buttons for "Search" and "Clear". On the left side, there is a vertical navigation menu with the following items: NCBI Home, Site Map (A-Z), All Resources, Chemicals & Bioassays, Data & Software, DNA & RNA, Domains & Structures, Genes & Expression, Genetics & Medicine, Genomes & Maps, Homology, Literature, Proteins, Sequence Analysis, Taxonomy, Training & Tutorials, and Variation. The main content area is titled "Welcome to NCBI" and includes a paragraph: "The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information." Below this, there are links for "About the NCBI | Mission | Organization | Research | RSS Feeds". A "Get Started" section lists several resources: Tools, Downloads, How-To's, and Submissions. To the right, there are two sidebars. The "Popular Resources" sidebar lists: BLAST, Bookshelf, Gene, Genome, Nucleotide, OMIM, Protein, PubChem, PubMed, PubMed Central, and SNP. The "NCBI News" sidebar features two news items: "New NCBI News Issue" dated 28 Mar 2011 and "Retirement of Peptidome, SRA & Trace Archive" dated 16 Feb 2011. At the bottom of the page, the word "Fertig" is visible.

Proprietary data integration activities

- Data warehouses
 - Example: iCHIP
 - www.ichip.de
 - DKFZ, Bioquant, Uni Heidelberg
 - Integrates Affymetrix, cDNA, tissue microarray, array CGH, mass spectrometry, gel electrophoreses, RNAi,...)



www.ichip.de

Minimal information data standards

- Started with MIAME (**M**inimal **I**nformation **A**bout a **M**icroarray **E**xperiment)
- Other minimal information standards followed and were integrated in the MIBBI project



Project website: <http://mibbi.org/>

32 projects committed to date including:

MIAME (microarrays), MIAPE (proteomics), CIMR (metabolomics)

MIGS, MINSEQE & MINIMESS (genomics, sequencing)

MIGen & MIQAS (genotyping), MIARE (RNAi), MISFISHIE (in situ)

A MIBBI project

Projects/GIATE - MIBBI - Mozilla Firefox

http://mibbi.org/index.php/Projects/GIATE

Meistbesuchte Seiten Erste Schritte Aktuelle Nachrichten GEO Accession viewer

MIBBI

WEB-SUCHE

Projects/GIATE - MIBBI

Log in / create account

article discussion view source history

Projects/GIATE

Guidelines for Information About Therapy Experiments

1 General features		
1.1	Domain	Cancer therapy experiments
1.2	Document Type	Primary checklist
1.3	Group	UCL
1.4	Main Website	http://www.antibodysociety.org/data/datastandards.php https://www.ucl.ac.uk/csm/projects/giate/
1.5	MI Checklist's Name	Guidelines for Information About Therapy Experiments
1.6	MI Checklist's Acronym	GIATE
1.7	Current Version Designation	1.0
1.8	Release Date for Current Version	n/a

2 Contact Person		
2.1	Full Name	Alejandra Gonzalez Beltran
2.2	Email Address	a.gonzalezbeltran[@]cs.ucl.ac.uk

2 Contact Person		
2.1	Full Name	May Yong
2.2	Email Address	m.yong[@]ucl.ac.uk

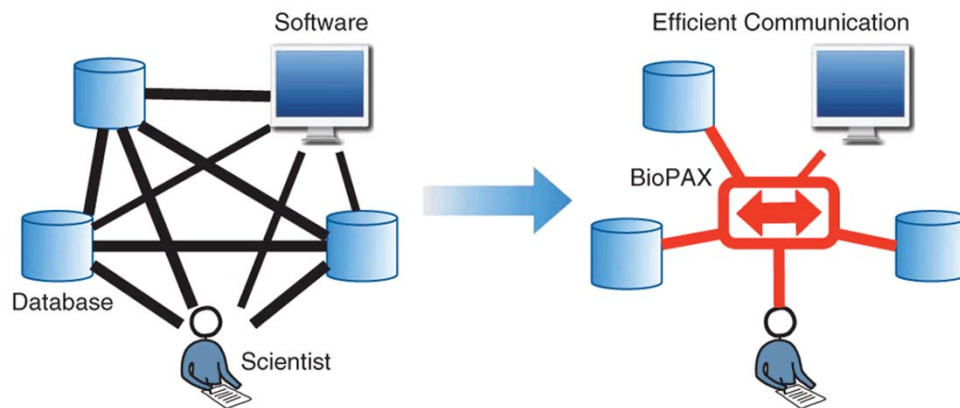
Fertig

Desktop DE 100% 20:15 26.03.2011

BioPAX

Demir et al. Nature Biotechnology 28 , 935-942 (2010)

- is a standard language for the integration, exchange, visualization and analysis of **biological pathway data**
- is coordinated with other pathway standards e.g. **SBML** and **CellML**



BioPAX Level 3 covers metabolic pathways, molecular interactions, signaling pathways, gene regulation and genetic interactions.

Data management systems in systems biology

- **Spreadsheets**
 - Template spreadsheets, e.g. ISA-TAB (<http://isatab.sourceforge.net/>)
- **Web-based document sharing tools**
 - Groupware software (e.g. eGroupware) or more specialized tools (e.g. DaMaSys, pyMantis)
 - Exchange of spreadsheets, SOPs, meta data, ...
- **Workflow systems based on web services**
 - Construct and run an analysis workflow on the grid

A web based document sharing tool: pyMantis (Falco Krause, HUB, SysMO-Translucent)

The screenshot shows the pyMantis web application running in a Mozilla Firefox browser. The browser's address bar displays the URL <https://translucent-network.org/pyMantis/isa/index>. The page title is "pyMantis - pyMantis".

The main content area features a large heading "Investigation Study Assay" with the letters "I", "S", and "A" highlighted in blue, pink, and blue respectively. Below the heading are two buttons: "Collapse All" and "Expand All".

The central part of the page displays a hierarchical list of research topics, each with a colored bar and a small icon:

- Metal Stress in *Saccharomyces cerevisiae*
- NHA1 Transport System in *Saccharomyces cerevisiae*
 - Current-voltage relations
 - Current voltage relation for different external KCl
 - SOP BCN-RNA001
 - SOP SOP microplate test Bonn
 - Current-voltage relation for different internal potassium
- Potassium uptake in *Saccharomyces cerevisiae*
 - External concentration changes
 - Potassium changes
 - Proton changes
 - Ion Flux Changes of Potassium Uptake Investigation
 - Proton fluxes
 - Additional fluxes
- Role of 14-3-3 proteins in *Saccharomyces cerevisiae*
- TRK1,2 Transport Systems of *Saccharomyces cerevisiae*

The left sidebar contains a search bar with the text "pyMantis" and a "go" button. Below the search bar are several menu items: "Maik Kschischo", "Logout", "Edit Profile", "Report a Bug", "Web Folder", "Wiki", "Members", "My Groups", and "RESTfull API".

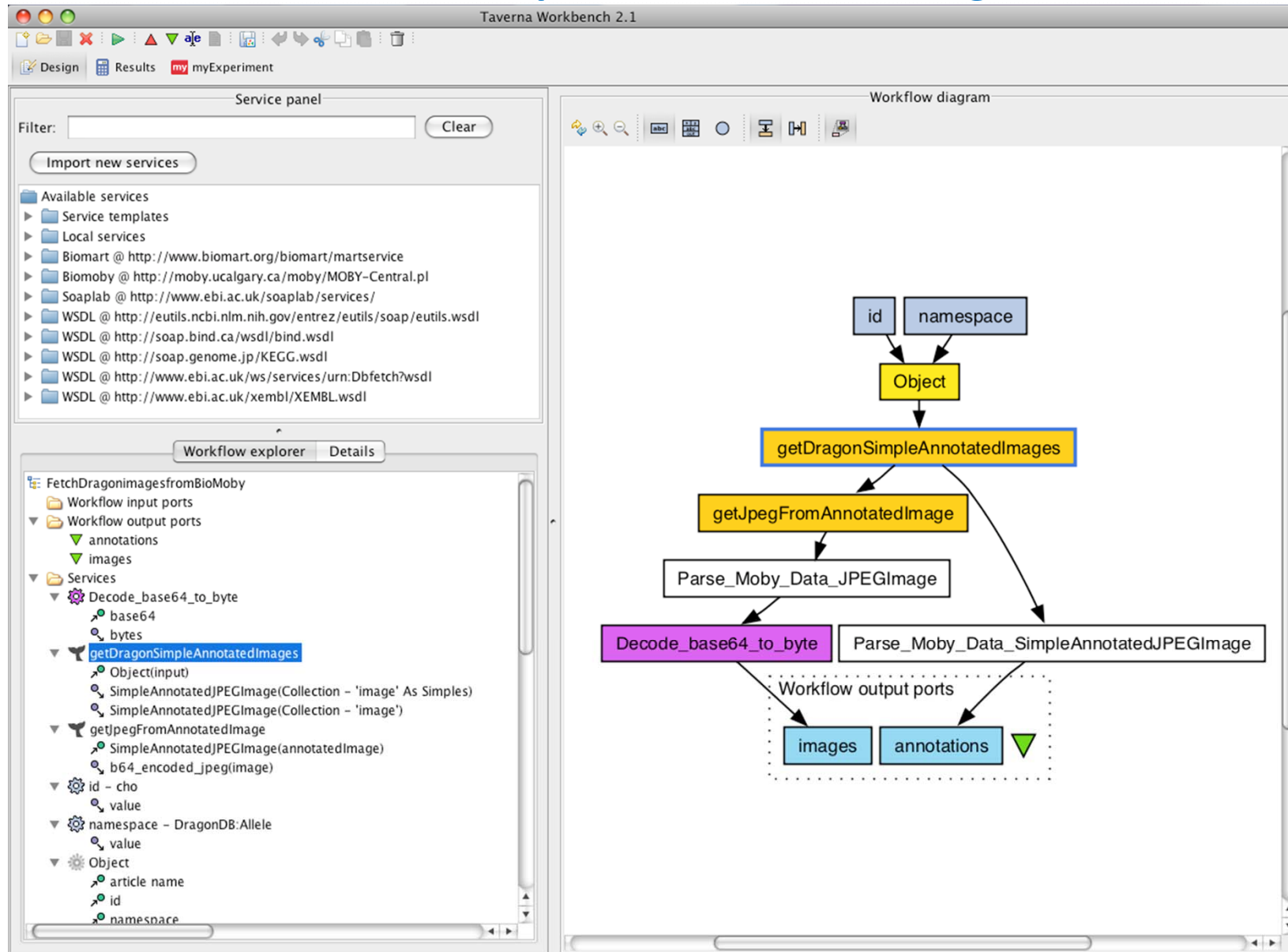
At the bottom of the sidebar, there is a section titled "Translucent 2" with a sub-section "I S A" and a list of categories: "MATERIALS", "SOP (26)", and "Strain (38)".

The bottom of the browser window shows a chat bar with "Chat" and "0 online". The Windows taskbar at the very bottom displays the system tray with the date "26.03.2011" and time "21:09".

Taverna work flow system

<http://www.taverna.org.uk>

Courtesy:
Katy Wolstencroft



Organisations and institutions

Many institutions and organisations, only good starting points

National:

- Helmholtz-Allianz-Systembiologie
www.helmholtz.de/systembiologie
- FORSYS
www.forsys.net
- Portal for systems biology
research in Germany
www.systembiologie.de

International:

- Systems Biology portal
www.systemsbiology.org

My key questions?

- How to **integrate biological data** (genomic, proteomic, *omic) **with medical data**?
- **Maintenance of data bases**?
- How can systems biology contribute to translational medicine?

Journals and conferences

Journals

- **Molecular systems biology**
- **BMC systems biology**
- **IET Systems Biology**
- **Bioinformatics**

and many other biology journals
including

- Nature Biotechnology
- PNAS
- Journal of Biological Chemistry
- ...

Conferences

- **International Conference on Systems Biology (ICSB)**
- Conference on Constraint-based Reconstruction and Analysis
- International Congress on Molecular Systems Biology
- Data Integration in the Life Sciences
- ...

Selected Literature

Reviews and Books:

1. Barnes, D.J. and Chu, D. "Introduction to Modeling for Biosciences", Springer, 2010.
2. Klipp, E., Liebermeister, L., Wierling, C., Kowald, A., Lehrach, H. & Herwig, R. Systems Biology. A Textbook. Wiley-Blackwell, Weinheim, 2009.
3. Sauer, U. et al.. "Getting Closer to the Whole Picture". *Science* **316**, 2007.
4. Way, J.C. and Silver, P.A. "Systems engineering without an engineer: Why we need systems biology: Essays and Commentaries" *Complexity - Complex Systems Engineering* **13**, 2007.

Data Management:

1. Lambrix, P. and Kemp, G.J.L. (Eds.): Lecture Notes in Computer Science **6254** Springer, 2010.
2. Mayer, G. Data management in systems biology 1: Overview and bibliography. [arXiv:0908.0411v3](https://arxiv.org/abs/0908.0411v3) [cs.DB].