Application of Systems Biology in Toxicology - Past Experiences, Present Applications and Future Perspectives

Thomas Steger-Hartmann, PhD
Head of Laboratory Diagnostics, Genetic and Ecotoxicology
Schering AG, Berlin

Hans Gmuender, PhD
Head of Scientific Consulting, Genedata AG, Switzerland

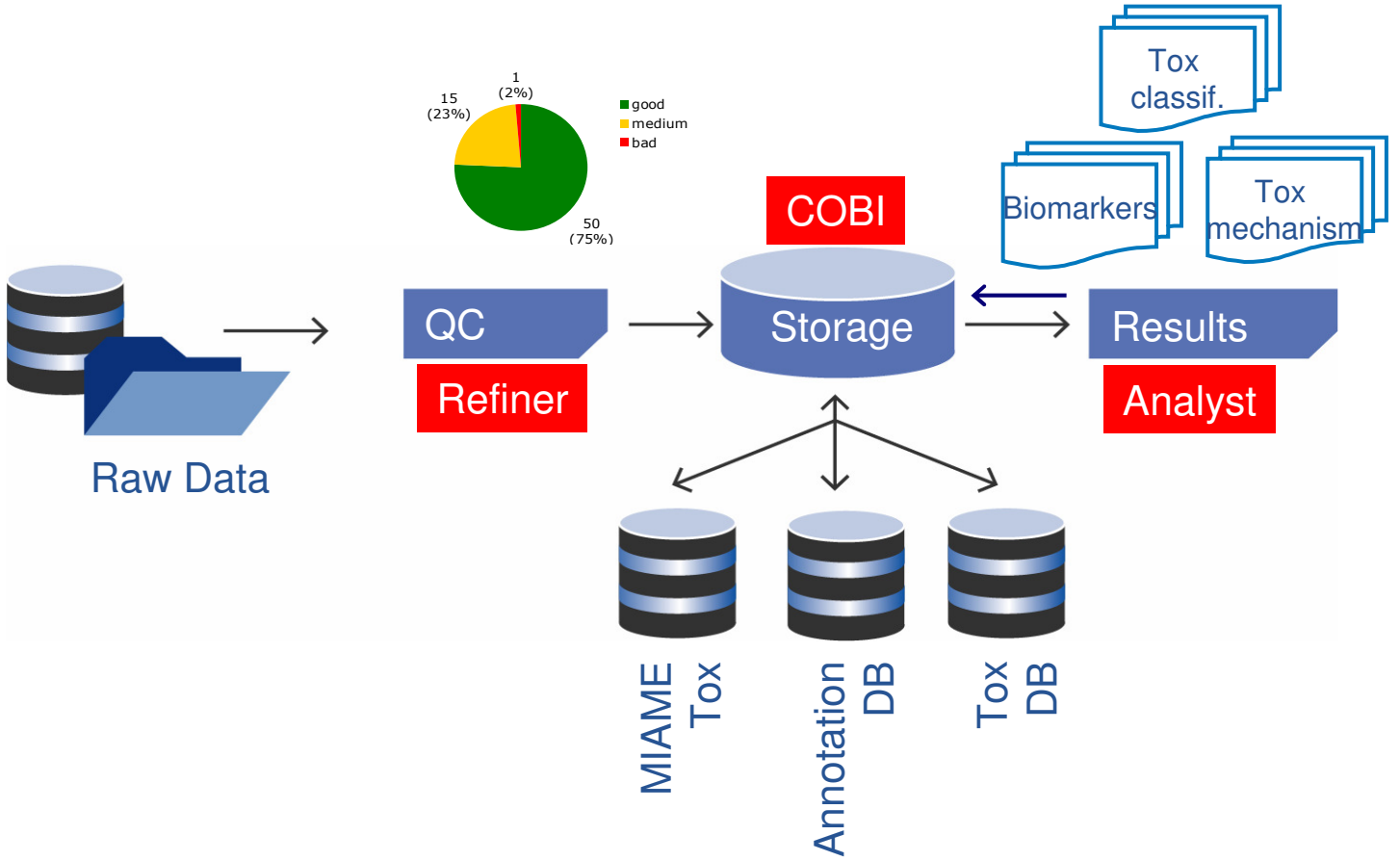
Schering AG - Genedata Collaboration

- Schering AG, a mid-sized globally orientated pharma company (gynecology & andrology, oncology, diagnostic imaging, specialized therapeutics)
- Genedata, a bioinformatics company with software and services for functional genomics and HTS (www.genedata.com)
- Long-standing, ongoing collaboration in gene expression analysis, particularly toxicogenomics

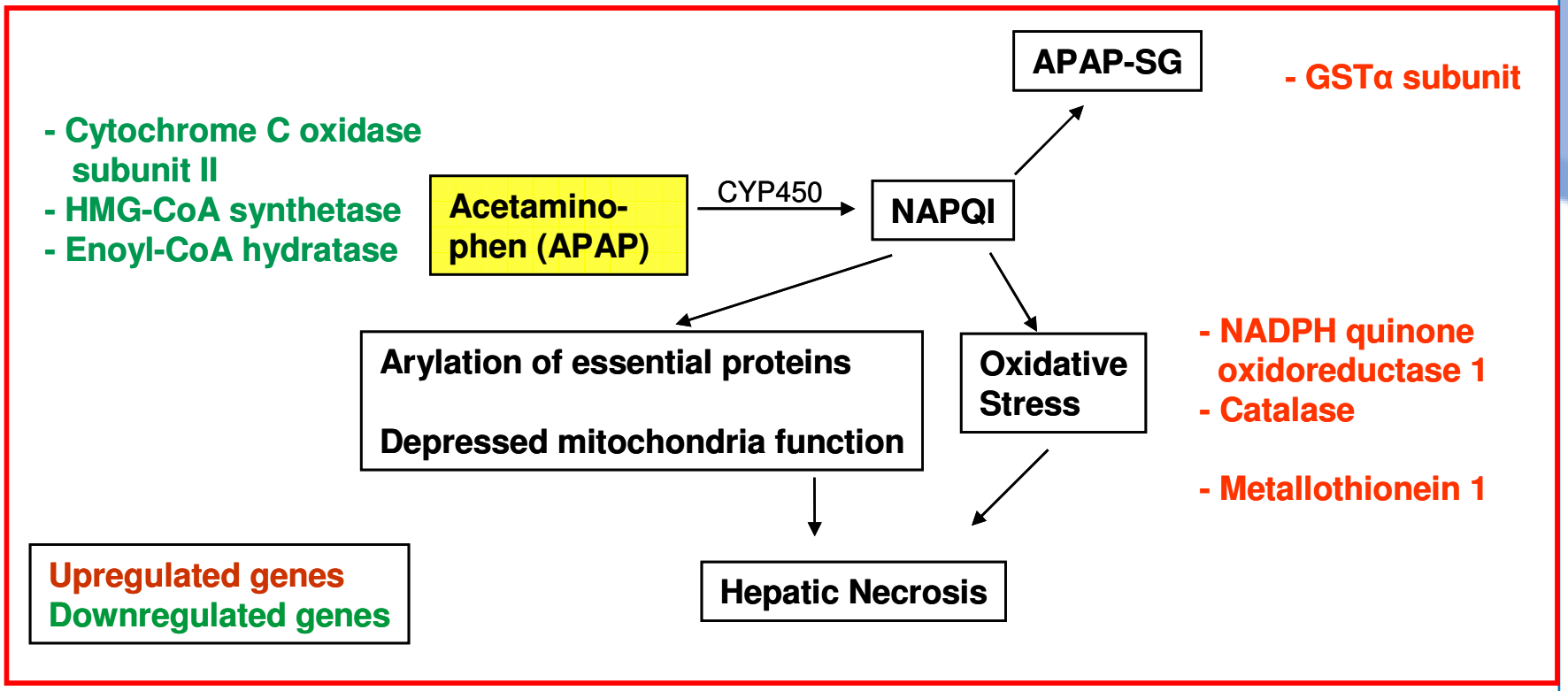
Toxicogenomics - Our Questions

- Can we elucidate the ***mode of toxic action (MotA)*** from gene expression data?
- Can we ***predict*** toxicological effects (histopathology scores) based on in vivo expression profiles stored in a commercial data base?
- Can we derive ***marker genes of toxicity***?

Genedata Expressionist®, Tox Edition

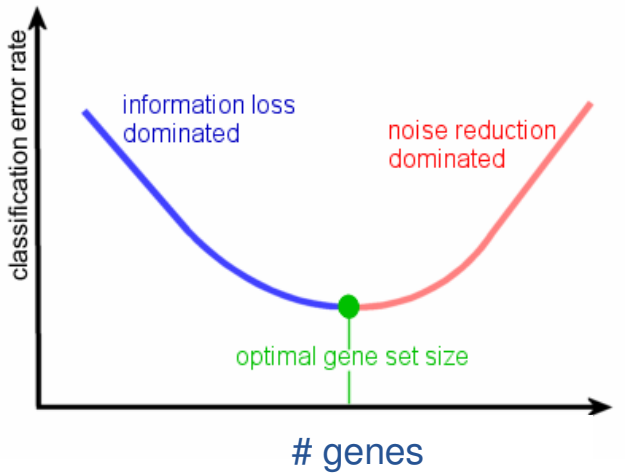


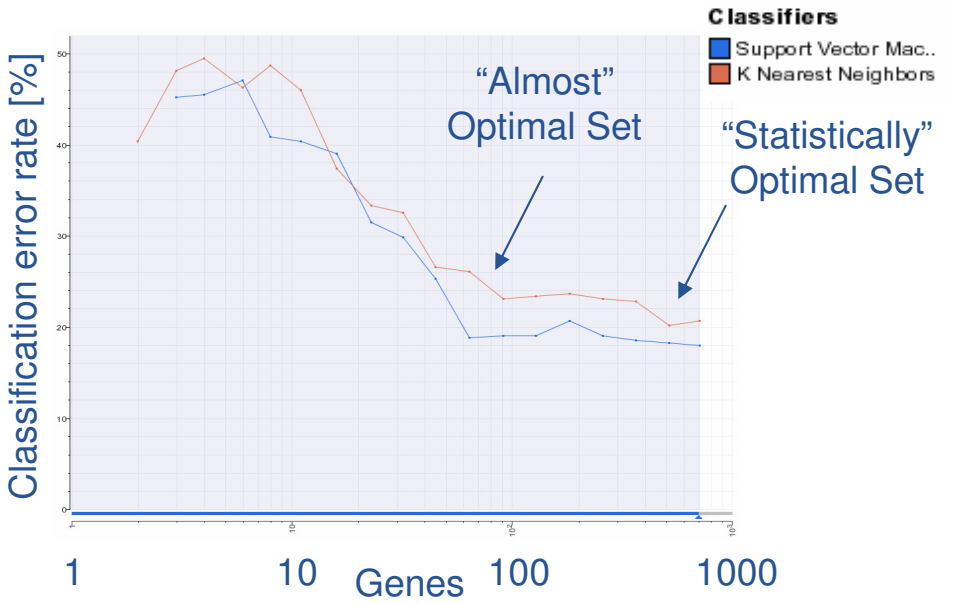
Toxicogenomics – Mode of Toxic Action (MotA) of APAP



MotA of reference compound is generally reflected in gene expression profiles

Toxicogenomics - Markers Genes for Toxicity





*Liver necrosis versus no liver toxicity

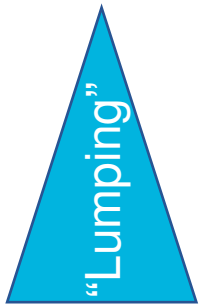
No small set (n<20) of marker genes for liver necrosis

Prediction of Liver Toxicity

Phase-1 ToxBank:

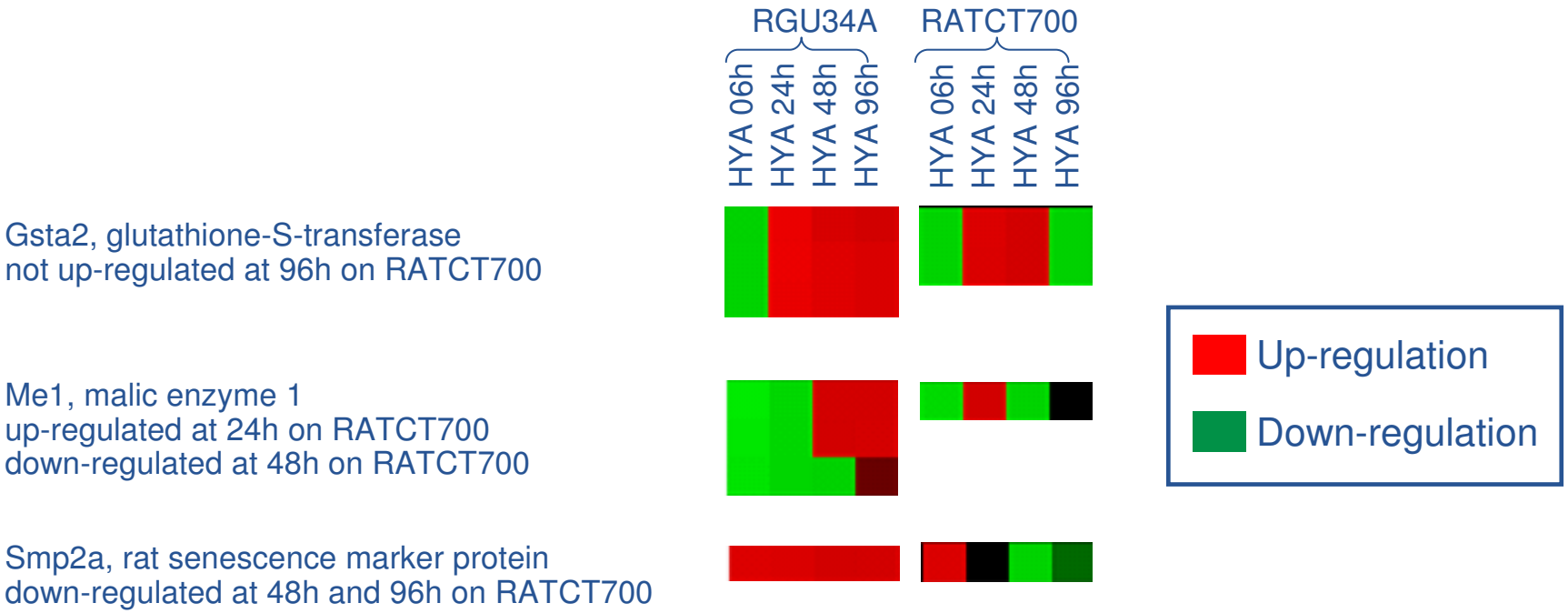
- Expression profiles from the liver of rats treated once with 52 different toxicants at different concentrations and time points (Phase-1 Toxchips)
- Assignment of Histo-Scores, SVP for prediction model

Grouping of histopathological scores	Prediction error
Individual histopathological scores	44 %
0b-3: liver necrosis *vs.* 0+0a: no liver necrosis	13 %



- Grouping histopathological scores reduces the prediction error but is paralleled by a loss of information
- Classification of own Acetaminophen expression data predict no liver necrosis (contrast to microscopical observations)

Toxicogenomics - Platform Variability



Up- and down-regulation of genes that occur on both arrays is not consistent among platforms.
Thus, platform variability may limit the use of the data base.

Toxicogenomics - Our Results and Answers

"Mechanistic" Toxicogenomics

- Gene expression profiles in general reflect the reported mechanism of toxic action of the investigated hepatotoxicant.

"Predictive" Toxicogenomics

- A competent partner for datamining is a prerequisite
- Prediction of toxicity using a commercially available data set in general is technically possible.
- Identification of a small set of marker genes for the endpoint liver

Early predictive toxicogenomics presently remains a promise!

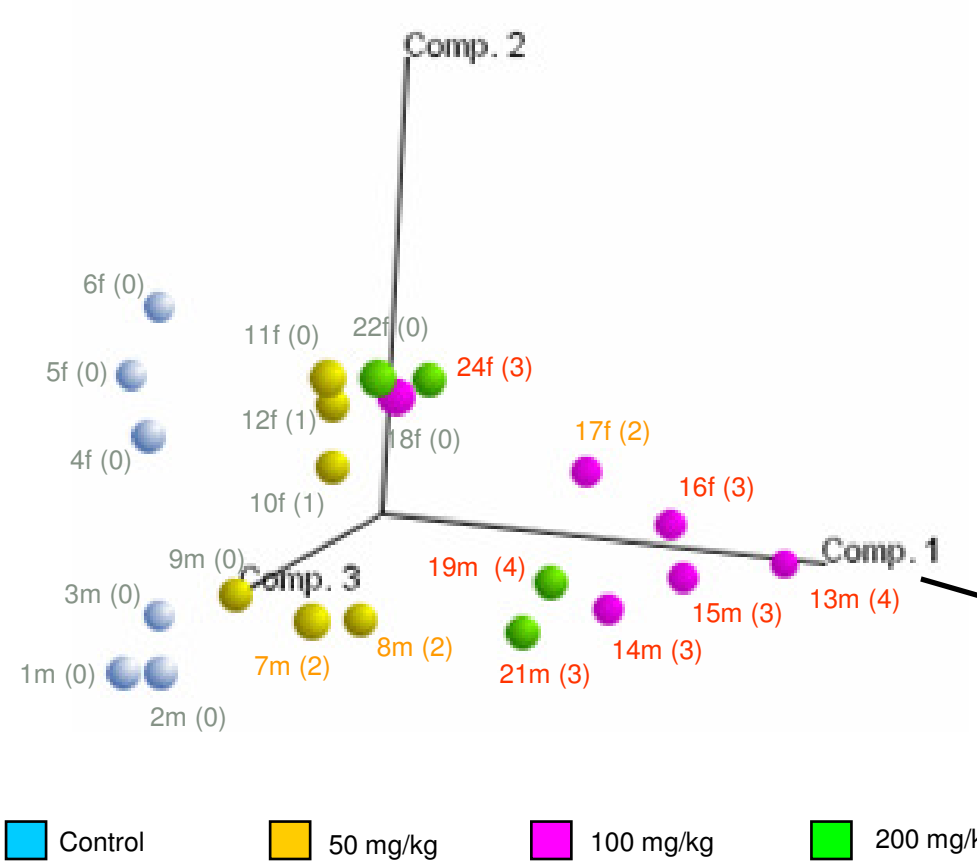
Towards Systems Biology - A Case Study

Findings in systemic rat studies (7-days, 26-weeks) after oral administration of a new multi-target tumor growth inhibitor (MTGI):

- Toxicogenomics data
- Metabonomics data
- Traditional endpoints

MTGI and Toxicogenomics: Analysis of Kidney

RG_U34A-Arrays (7 day rat study) - Principal Component Analysis



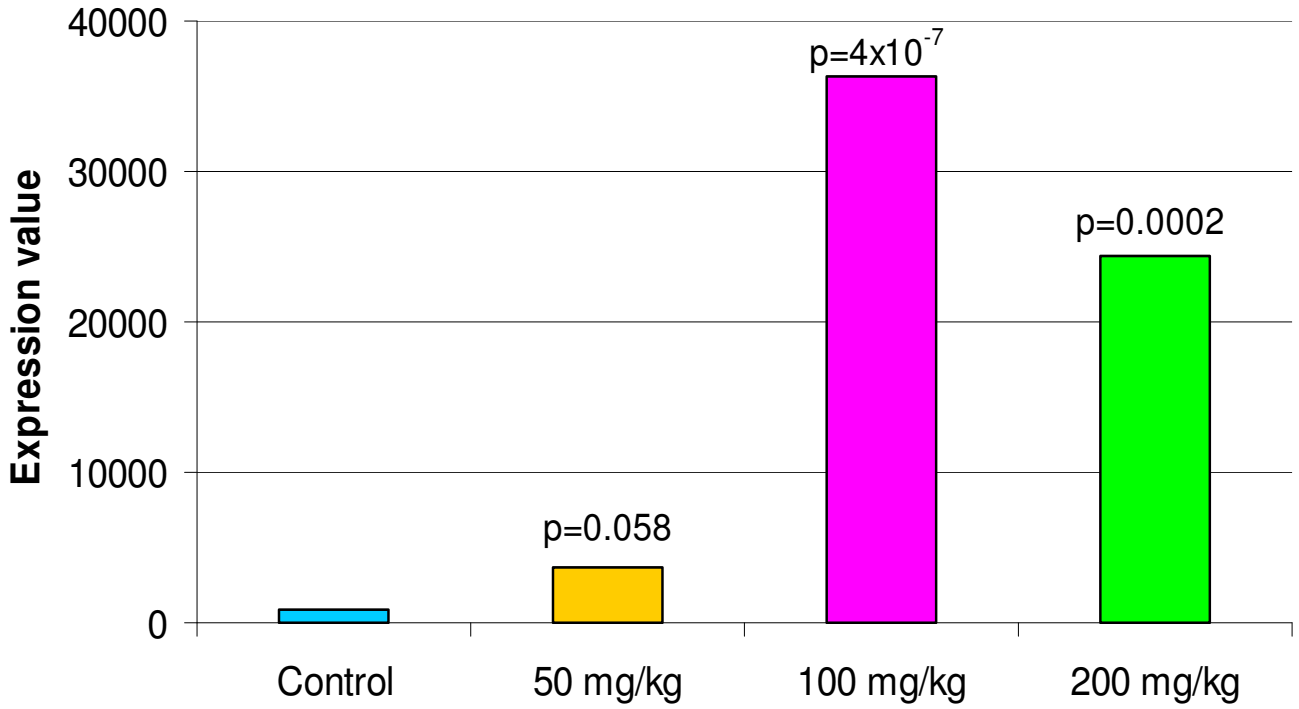
- Clear separation between control and treated animals
- Animals with histograding > 2 separate clearly from all other animals
- Differences in expression profiles between male and female animals

Eigengene can be thought as a major expression pattern in the data set which cause the highest variability in the direction of the respective vector.

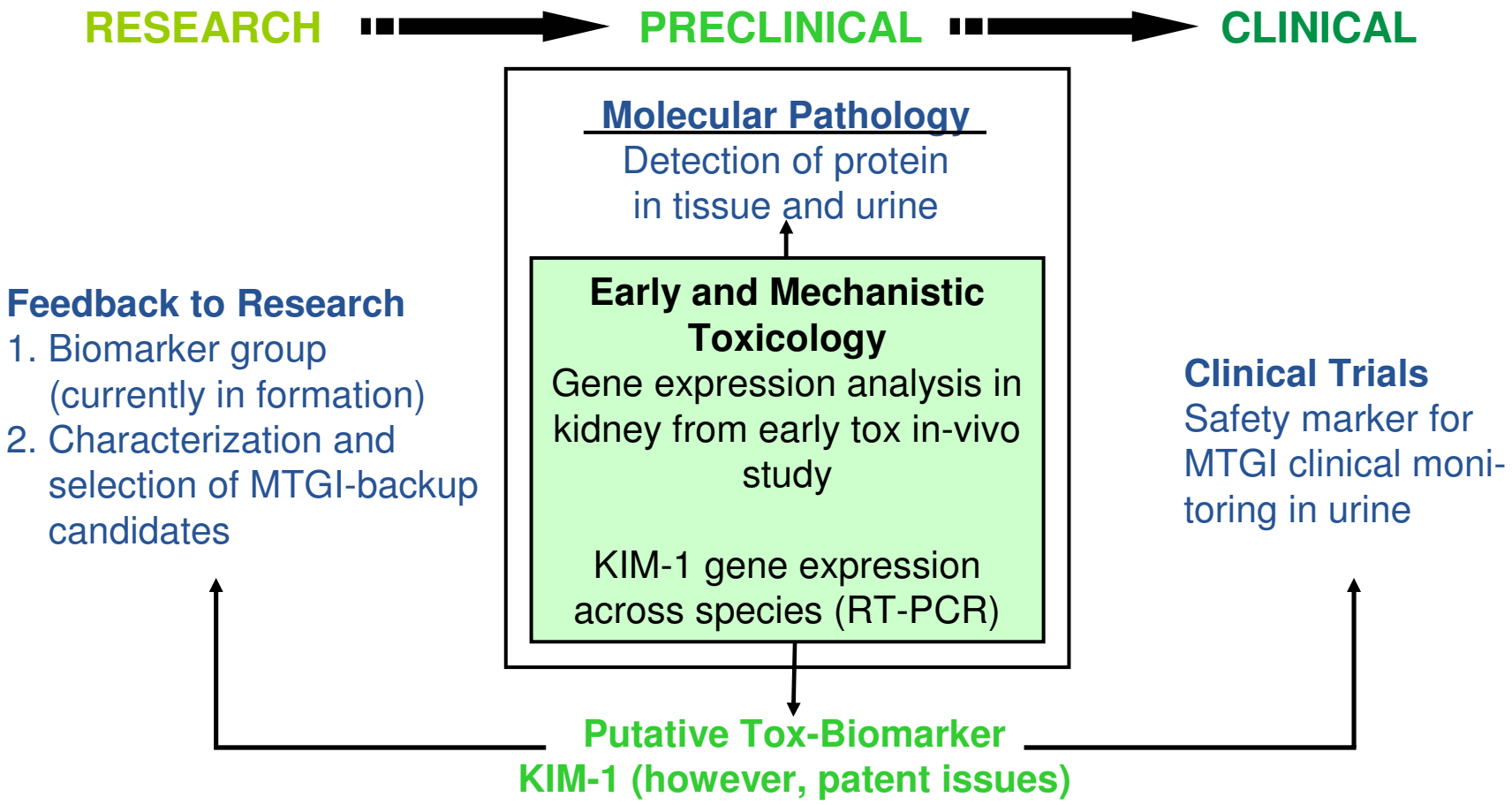
MTGI and Toxicogenomics: Analysis of Kidney

RG_U34A-Arrays (7 day rat study) - Principal Component Analysis

Kidney Injury Molecule-1 (KIM-1) is the highest upregulated gene in all dose groups

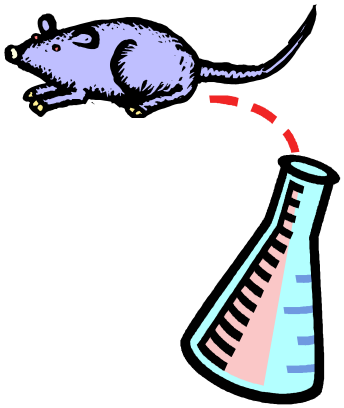


KIM-1 as Biomarker of Safety



The Metabonomics Technology

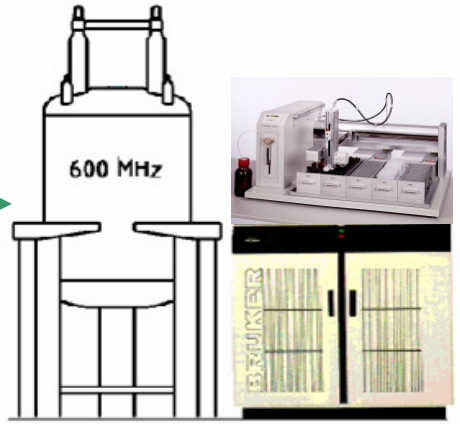
Animal experiment



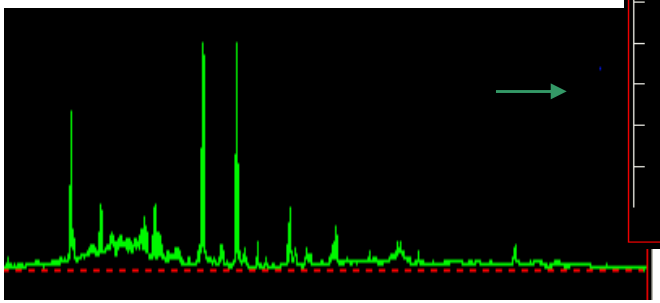
Sample storage

Sample Preparation

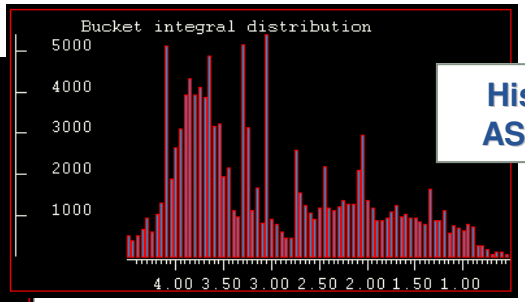
Analytics with NMR-Spectroscopy



NMR-Spectra



Bucketing



Histogram + ASCII-output

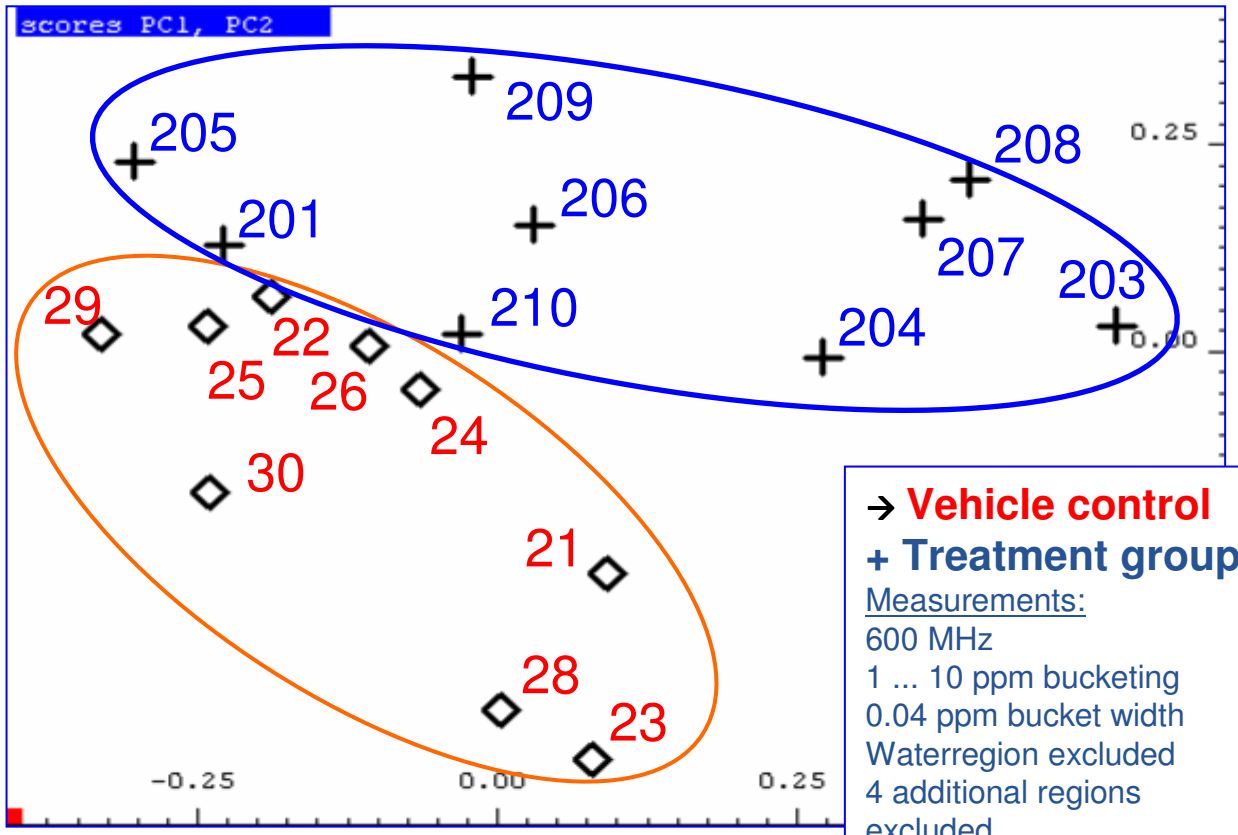
Pattern recognition techniques

(e.g. PCA)

Adapted from M. Spraul, Bruker BioSpin GmbH

MTGI - Metabolomics: Bruker BioSpin NMR-Measurements

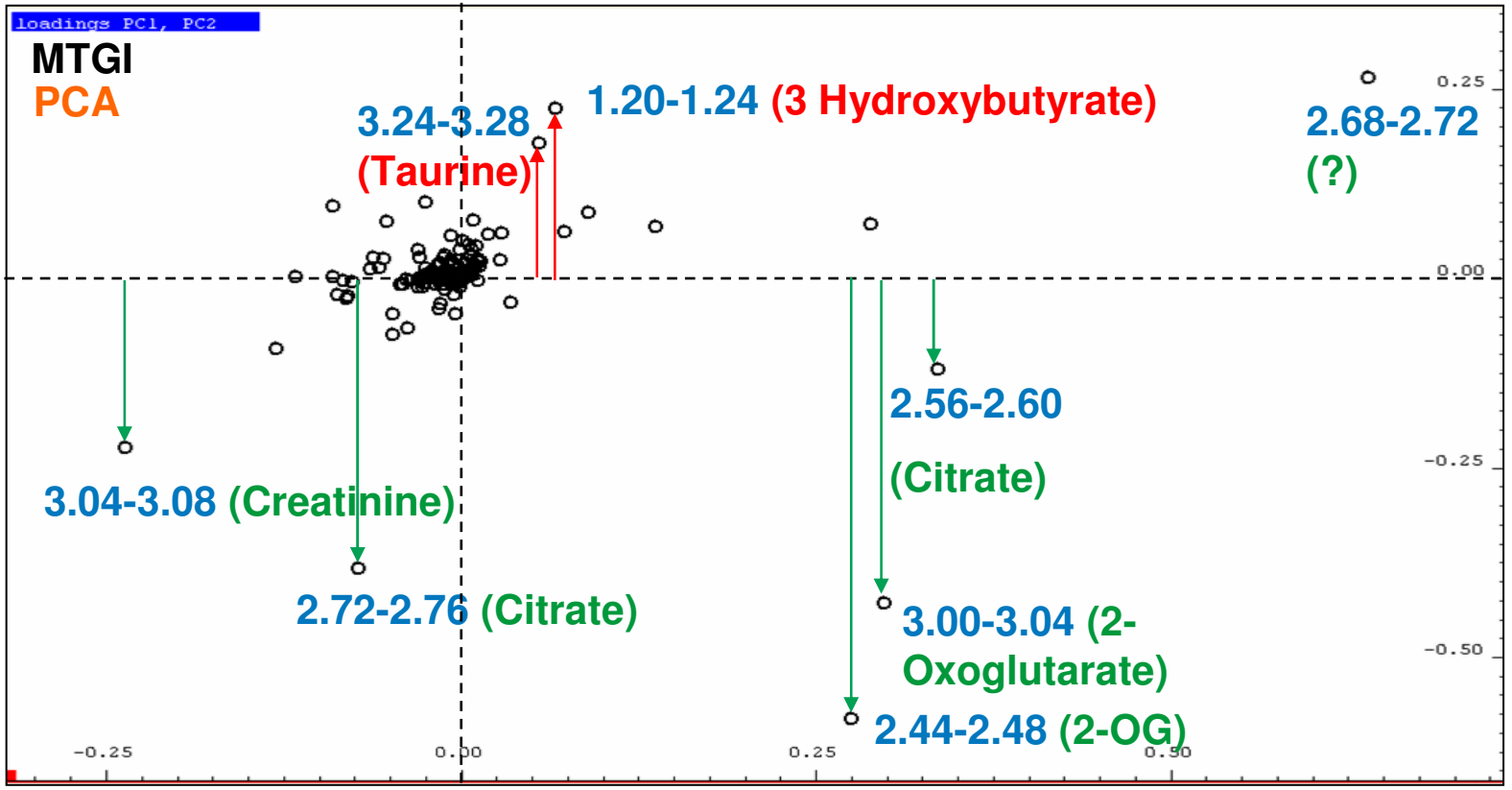
- 26 week i.g. rat study with MTGI
- Satellite animals of high dose group vs. Vehicle control (n= 9)
- 24 h in metabolic cages for urine collection (fasting)



→ **Vehicle control**
+ Treatment group

Measurements:
600 MHz
1 ... 10 ppm bucketing
0.04 ppm bucket width
Waterregion excluded
4 additional regions excluded

MTGI - Metabolomics: Loading Plots of the 26-W Rat Study



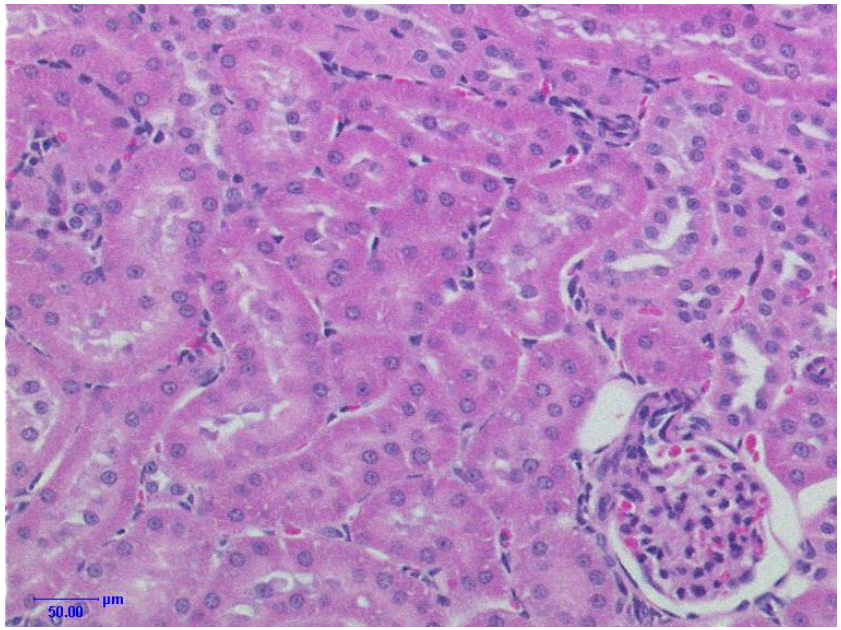
MTGI - Metabolomics Findings

- Excretion of 3-hydroxybutyrate is a potential marker for proximal tubular damage (*Anthony et al., Arch. Toxicol, 1994*)
- Taurineurea correlates with hepatotox (*Sequeira et al., J. Pharm Biomed Anal, 1990*)
- Decrease in creatinine, pointing towards liver toxicity (Anthony et al., Arch. Toxicol. 1002)
- Decreases in citrate, 2-oxoglutarate, indicating inhibition of carbonic anhydrase (*Anthony et al., Arch.Toxicol, 1994*)

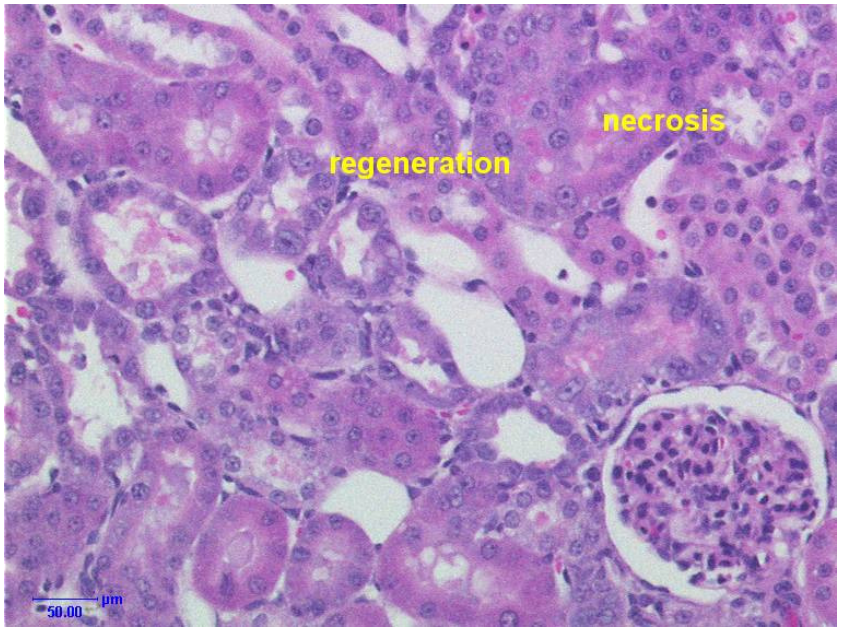
MTGI - Histopathology of 7-D Rat Study

Observation: Tubular necrosis with regeneration in the kidney at all tested doses, but no significant increase in serum creatinine.

Control



Male rat, 200 mg/kg MTGI



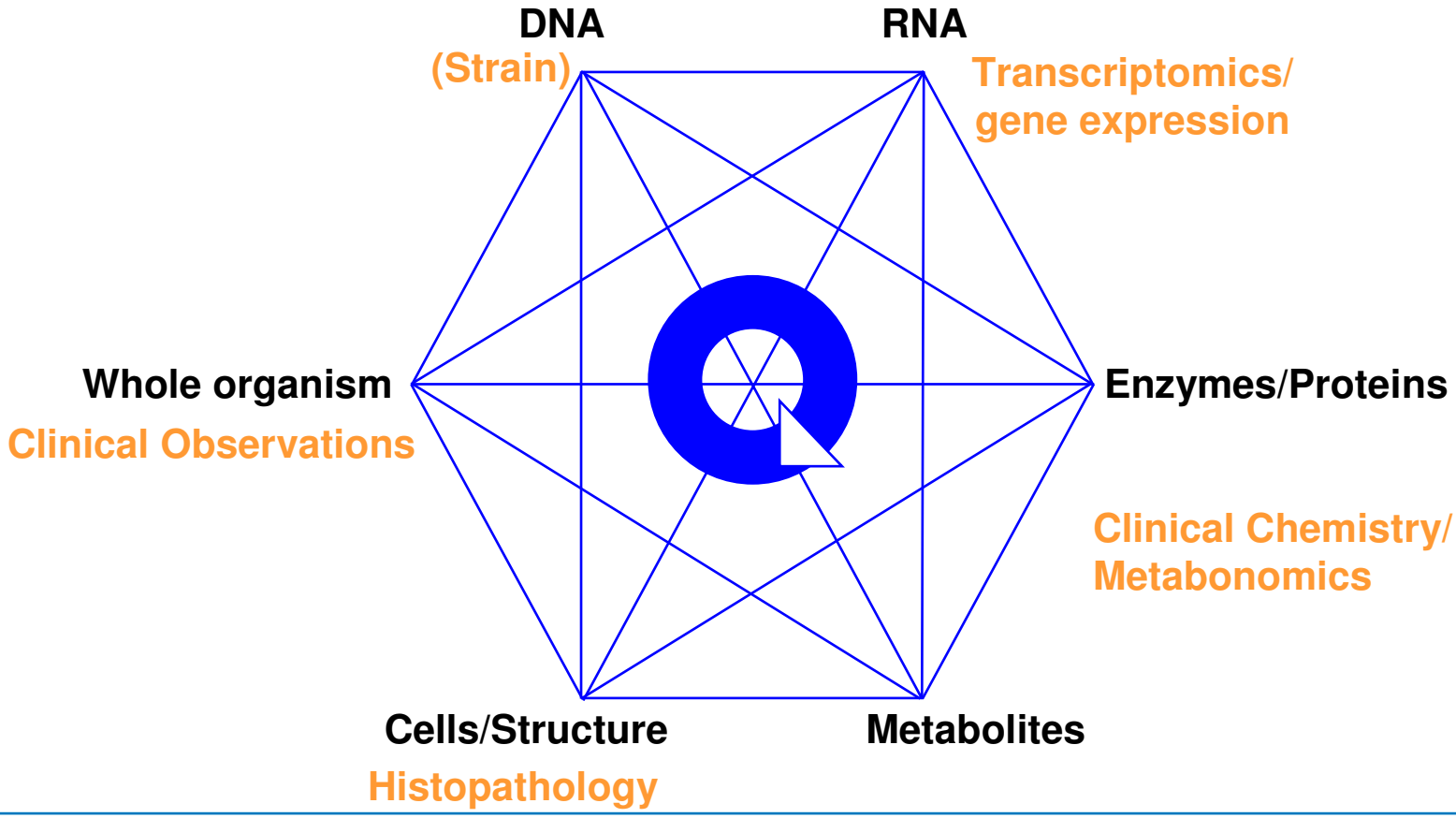
Systems Biology? - Putting the Pieces together

- 7-day study: **Kidneys:** KIM-1 increased, no increase of serum creatinine
- 26-week study: **Liver:** decreases in total serum protein and serum cholesterol, changes in protein fractions and blood coagulation parameters, decreased creatinine
  Kidne
  specifi
  NAG a

The conventional comparison is time-consuming, incomplete and not very effective!

Kidneys are the primary target organ in the rat, KIM and 3- hydroxybutyrate point towards a damage similar to postischemia (oxidative stress?), carbonic anhydrase inhibition evident by clinical chemistry and metabolomics (drug binds to CA!!!), liver is the secondary target organ.

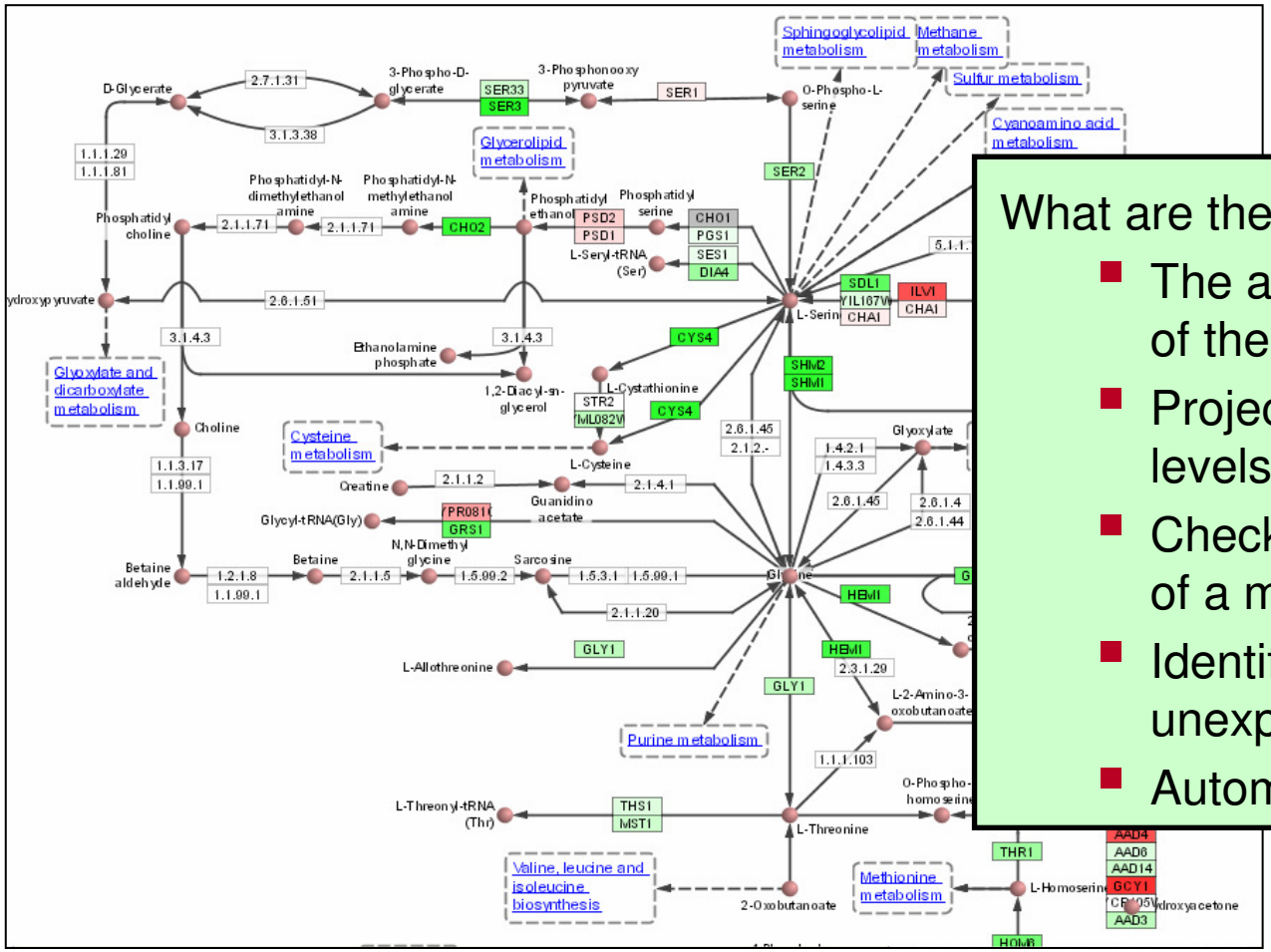
Systems Biology – The Effort of Integrating Data from Different Levels



Integration of Cross-level Data

- Advantages
  - A handle on the effective action of the genes
  - Elucidation of biological pathways and mechanisms
  - A deeper understanding of the (molecular) biology of the transitions
- Challenges
  - The dynamic ranges are not constant over the different levels
  - How do we express our knowledge in terms of probabilities?
  - Strong one-to-one correlations are the exception
  - Vast amount of interactions within and between ‘omics’ levels, plus feedback loops => the combinatorics become enormous

The Core of Integration - the Pathway Map



What are the challenges?

- The actual reconstruction of the pathway
- Projection of expression levels onto the pathways
- Checking the consistency of a map
- Identifying new or unexpected things in a map
- Automation

Predictive -omics - The Need for a Reference Database ("populating the platform")

The prediction of toxic effects based on -omics data requires high quality data residing in a reference data base

Option 1: Use of a commercial data
study-wise approach); ex
available

Option 2: Build a proprietary database (consortium approach with Genedata); less expensive, dependence on consortium members; mid-term availability

Option 2 is clearly favored!

Starting a Toxicogenomics Consortium

- Building a joint toxicogenomics reference database from GeneChip hybridizations
- Partners: Altana Pharma AG, Merck KGaA, and Schering AG
- Genedata guides the consortium, performs data analysis services and provides a platform for hosting toxicogenomics data
- Phases of the compendium establishment:
  - Standardization of study protocols and data analyses
  - Build a reference compendium with a core set of compounds
- Activities may merge into *InnoMed Predictive Toxicology Consortium*

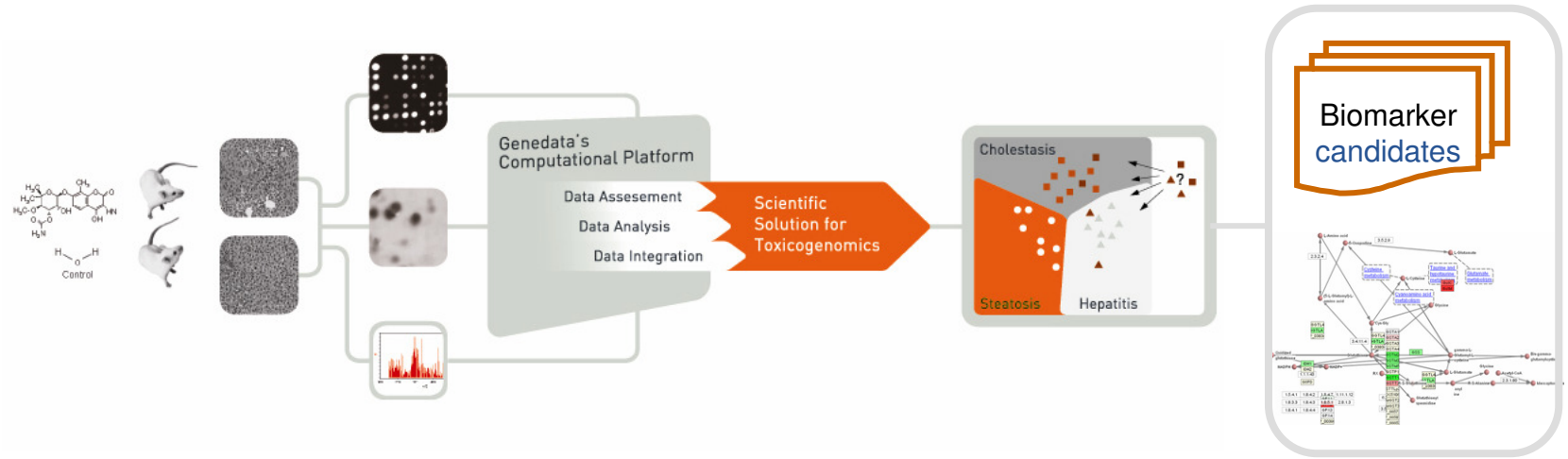
The InnoMed Predictive Toxicology Consortium

- Application of an EU grant FP6/FP7 (“Innovative Medicines for Europe - Project Predictive Toxicology”)
- Participants: 13 EU Pharma companies, 3 universities, 1 IT provider (Genedata)
- Study design: in vivo studies in Wistar rats with known toxins
- Endpoints: traditional endpoints, transcriptomics (Affymetrix array), proteomics (2-D gels, SELDI), metabonomics (NMR, LC-MS)
- Aim: establishment of an integrated shared database used by industry, academia and regulators for prediction and decision making

The Genedata Platform for Systems Biology

Predictive Toxicology

Mode of Action



Summary

- Diagnostic tools for investigations of the different levels in Systems Biology (-omics technologies) are available
- There is a lack of reference data for the new technologies
- First software approaches for data integration are in place, but await optimisation

Acknowledgements

- Kirstin Meyer, Head of Early and Mechanistic Toxicology
- Björn Riefke; Head of Laboratory Diagnostic
- Jakob Walter, Senior Pathologist

Back-ups

Agenda:

1. -omics (“Toxic-omics”) at Schering AG - Evaluating the technologies
2. Towards systems biology - A case study
3. Future perspectives, needs and developments

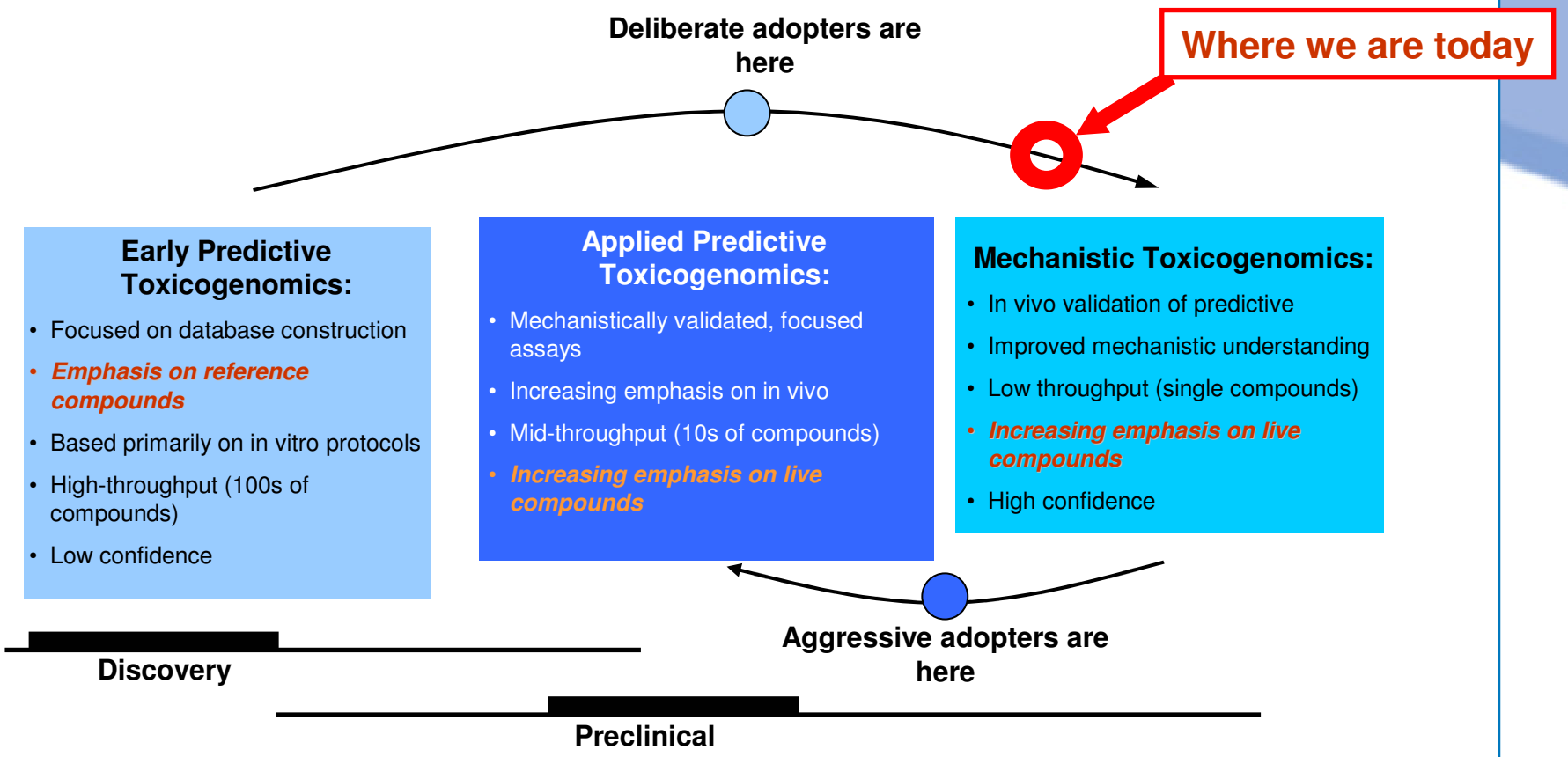
The Need for a Consortium Approach in Systems Biology for Toxicology

A systematic effort to perform systems biology in toxicology requires enormous resources. It would be better to join forces with other pharmaceutical companies

Requirements:

- Standard protocols and high quality data
- Data transferable between consortium partners
- Application of different technologies + traditional endpoints
- Sophisticated data analysis and bioinformatics IT
- Search tools across endpoints
- Integrated metabolic/physiological interpretation tools

Where are we on the map



CHA: A Benchmark of Current Practices and Future Expectations, April 2004

Applications and examples of NMR-based metabonomics

Application	Examples	Matrix	Reference (examples)
Classification of toxicity	Nephrotoxicity	Urine	JK Nicholson et al., Mol. Pharmacol., 1985, **27**, 644.
	Hepatotoxicity	Urine, serum	DG Robertson et al., Toxicol. Sci., 2000, **57**, 326.
	Phospholipidosis	Urine, serum	AW Nicholls et al., Biomarkers, 2000, 5, **410**.
	Testicular toxicity	Urine	JK Nicholson et al., Mol. Pharmacol., 1994, **46,** 199.
	Mitochondrial toxicity	Urine	E Holmes et al., Mol. Pharmacol., 1992, 42, 922.
Classification of disease	Inborn errors of metabolism	Urine	E Holmes et al., Anal. Biochem., 1994, **220**, 284.
	Cancer (prostatic, brain, renal, etc)	Tissue specimen	D Moka et a l., J. Pharm. Biomed. Anal., 1998, **17**, 125.
	Renal disease	Urine	GH Neild et al., Nephrol Dial Transplant., 1997, **12**, 404.
	Diabetes	Urine	H Antti et al., J Chemom., 2002, **16**, 461.
	Muscular dystrophy	Tissue extracts	JL Griffin et al., Anal. Biochem., 2001, **293**, 16.
	Atherosclerosis	Serum	JT Brindle et al., Nature Medicine, 2002, **8**, 1439.
	Alzheimers	Cerbrospinal fluid	FYK Ghauri et al., NMR Biomed., 1997, **10**, 99.
Investigation of physiological status	Diurnal variation	Urine	ME Bollard et al., Anal. Biochem., 2001, **295**, 194.
	Hormonal variation	Urine	ME Bollard et al., Anal. Biochem., 2001, **295**, 194.
	Dietary effects	Urine	CL Gavaghan et al., Anal. Biochem., 2001, **291(2)**, 245.
Monitoring efficacy of therapeutic intervention	Renal transplantation (cyclosporin)	Urine	GH Neild et al., Nephrol Dial Transplant., 1997, **12**, 404.
Functional genomics	Assessment of strain differences in animal models	Urine	CL Gavaghan et al., FEBS Lett, 2000, **484**, 169.
Characterisation of natural products	Assessment of batch variation	Plant extracts	NJ Bailey etal., Planta Med., 2002, **68(8)**, 734

Facts about Kidney Injury Molecule-1 (KIM-1)

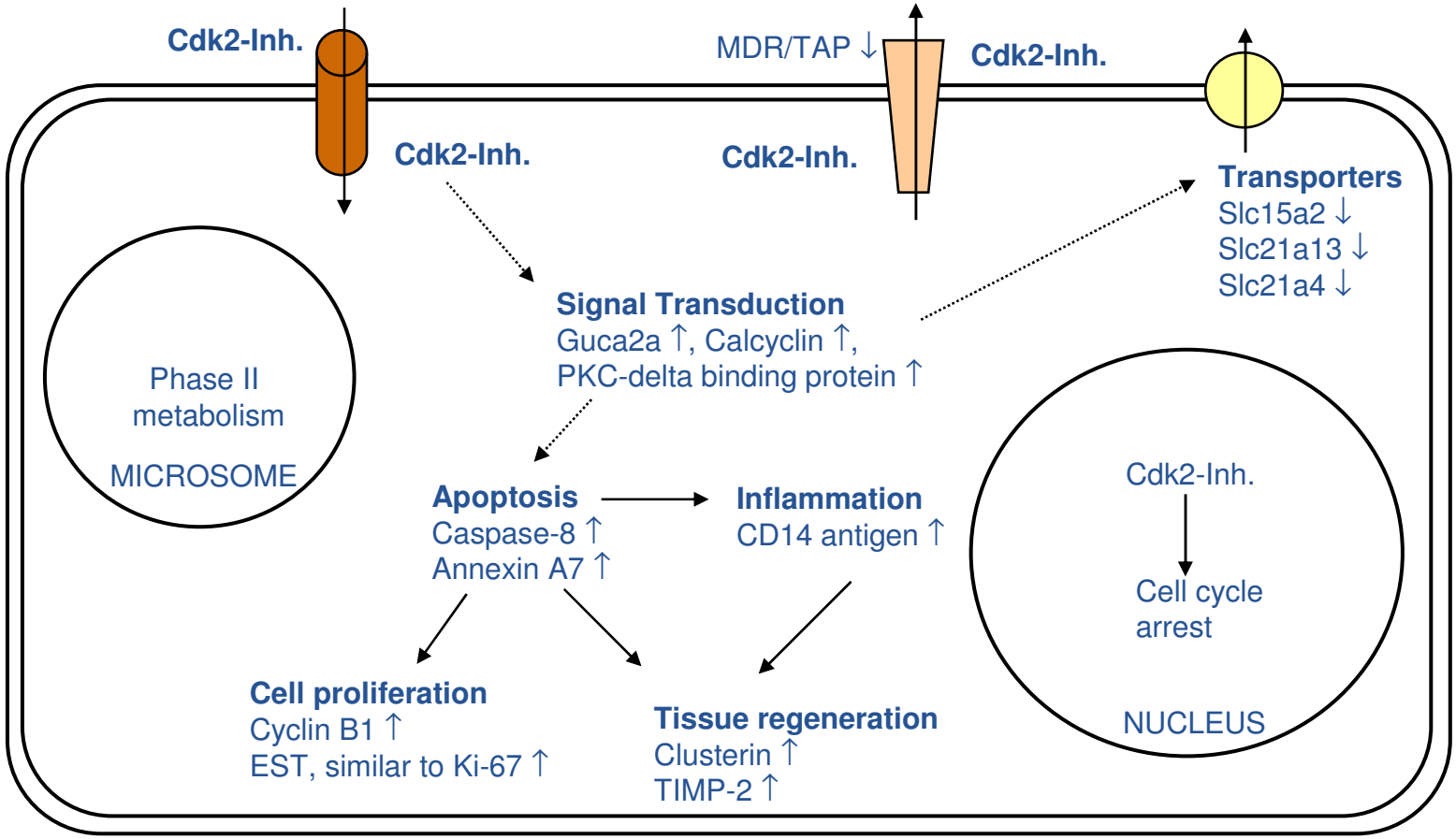
KIM-1:
- Membrane protein with extracellular immunglobulin and mucin domains
- Expression dramatically increased after injury in proximal tubule epithelial cells, i.e. in the postischemic rat kidney

Rodents: Kidney injury molecule-1 (Kim-1): a tissue and urinary biomarker for nephrotoxicant induced renal injury.
(Ichimura et al. Am. J. Phyiol. Renal Physiol 286: F552-F563, 2004)

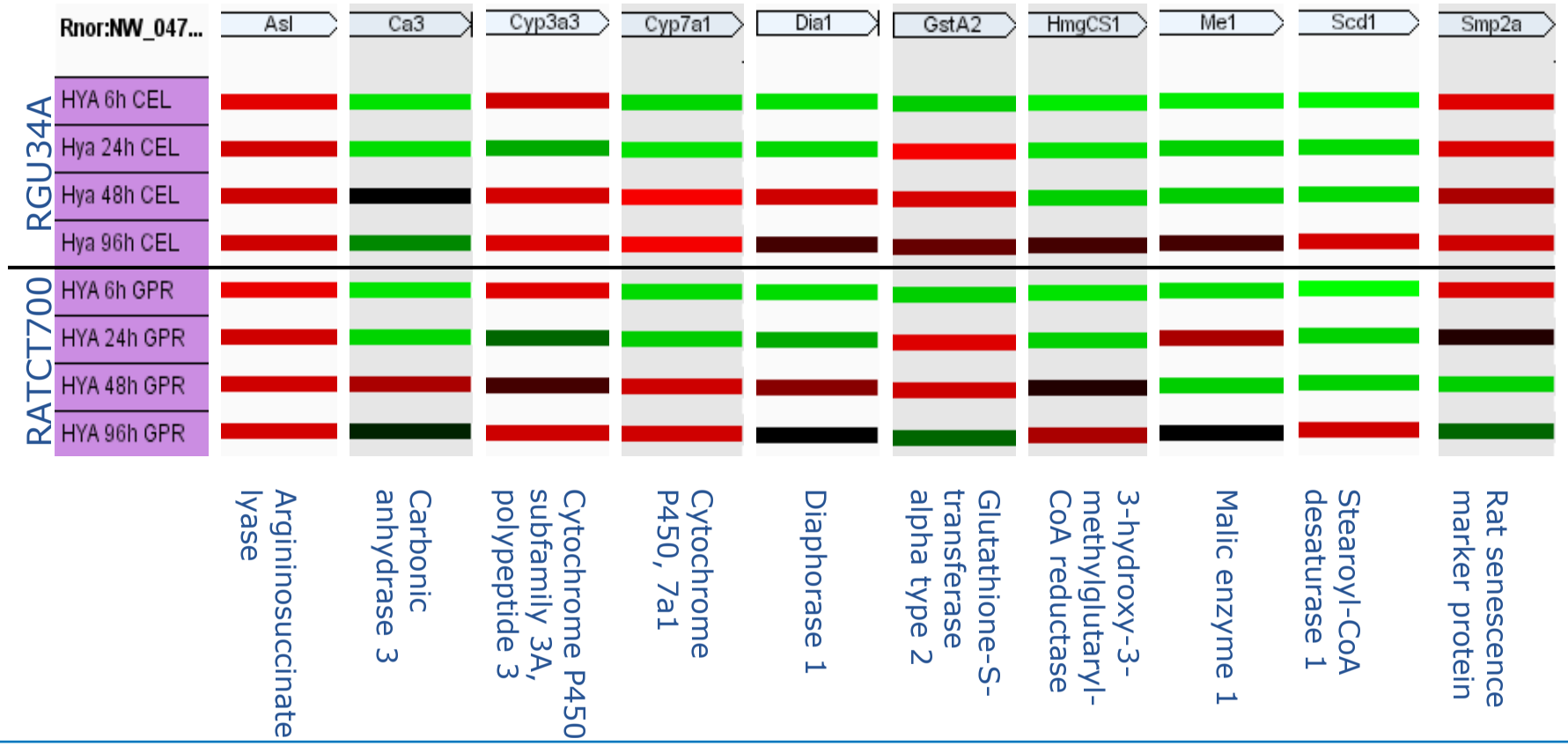
Non-human primates: Qantitative gene expression analysis in a nonhuman primate model of antibiotic-induced nephrotoxicity.
(*Davis II et al. Toxicol. Appl. Pharmacol. 200: 16-26, 2004)*

Humans: Kidney Injury Molecule-1 (KIM-1): a novel biomarker for human renal proximal tubule injury.
(*Han et al. Kidney Int. 62: 237-244, 2002)*

Putative Cellular Pathways Involved in Cdk2-mediated Nephrotoxicity

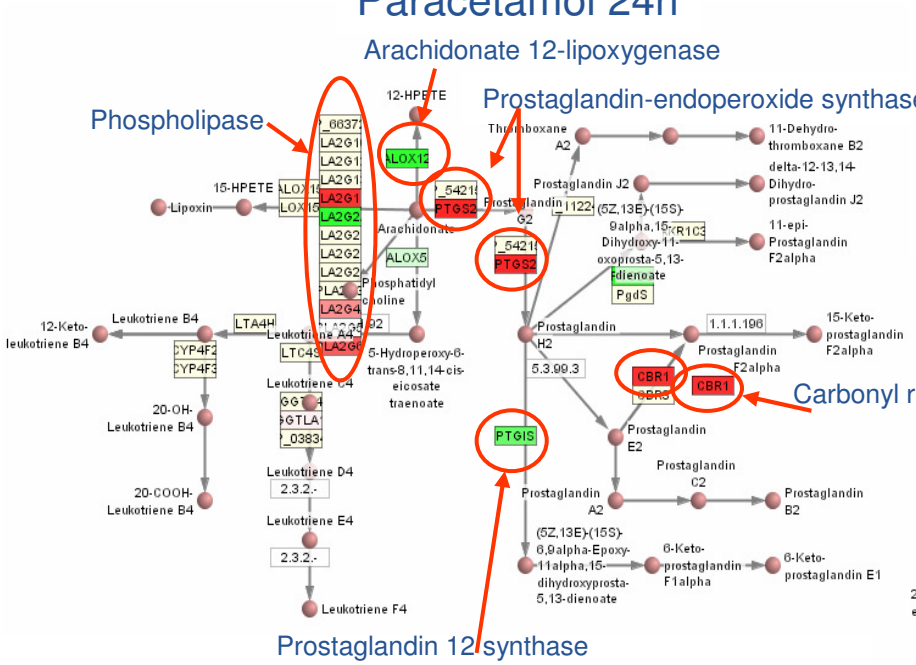


Mapping of Paracetamol Induced Gene Expression Values onto the Rat Genome

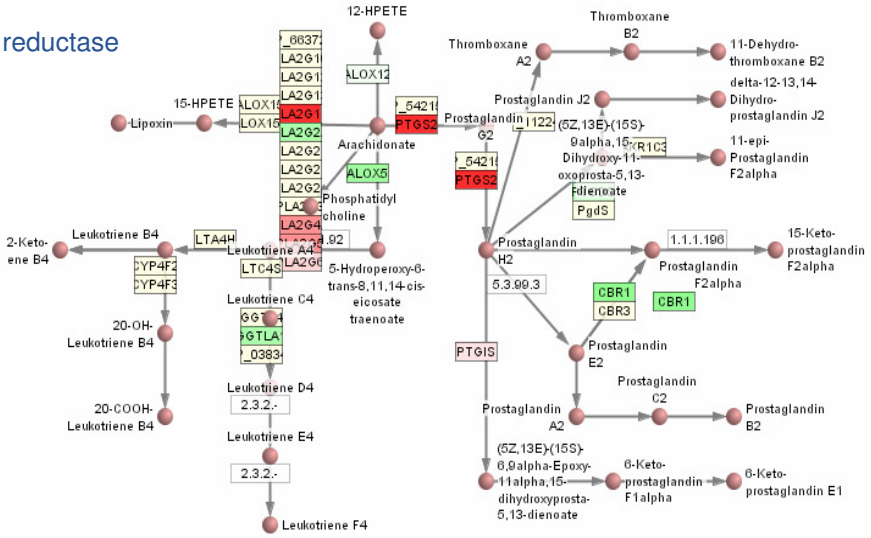


Influence of Paracetamol Treatment on the Prostaglandine and Leukotriene Metabolism

Paracetamol 24h



Paracetamol 48h



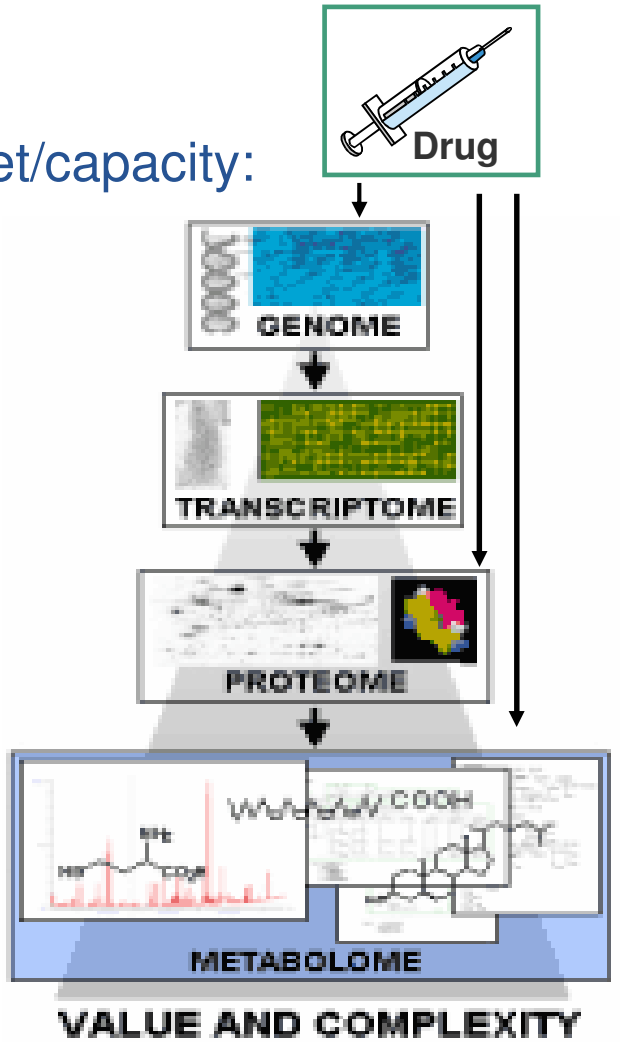
Expression decrease from 24h to 48h for:
- Carbonyl reductase

Expression increase from 24h to 48h for:
- Arachidonate 12-lipoxygenase
- Prostaglandin 12 synthase

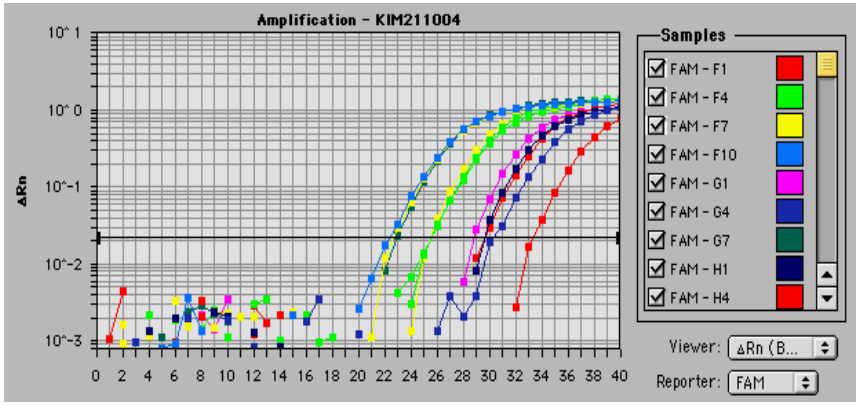
Metabonomics/Metabolomics

Pursuing a new technology with minimal budget/capacity:

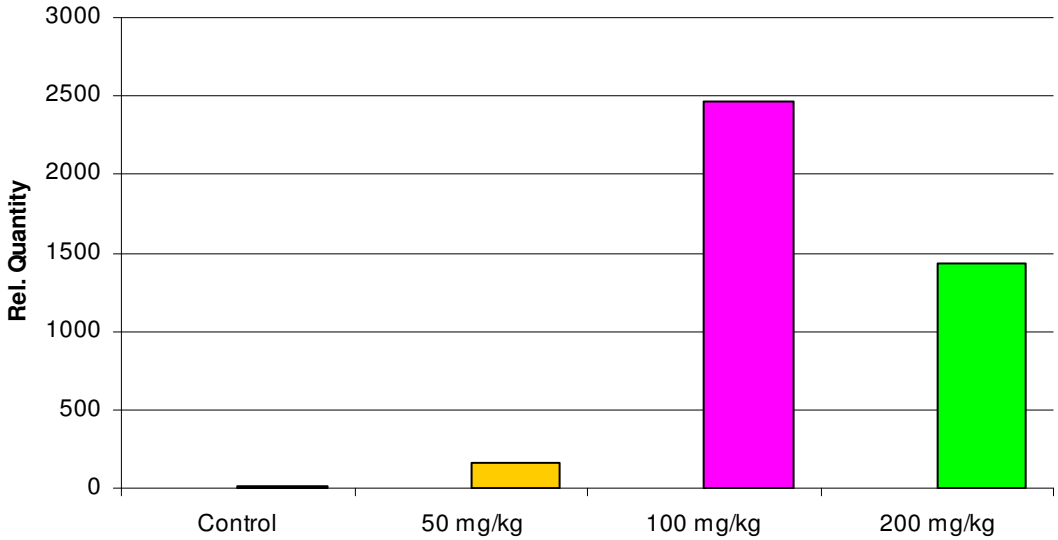
- Method: NMR analysis of body fluids (mainly urine of rats)
- Goal: non-invasive detection of effects during in-life phase; detection of sets of biomarkers based on metabolic response
- Best-NMR (high resolution) in place in Department of Structural Analysis
- Experiences in data processing and evaluation from toxicogenomics can be transferred to this new field (e.g. PCA)
- Feasibility study with Metanomics Health GmbH, Berlin



Verification of KIM-1 Expression Using RT-PCR (TaqMan® assay)

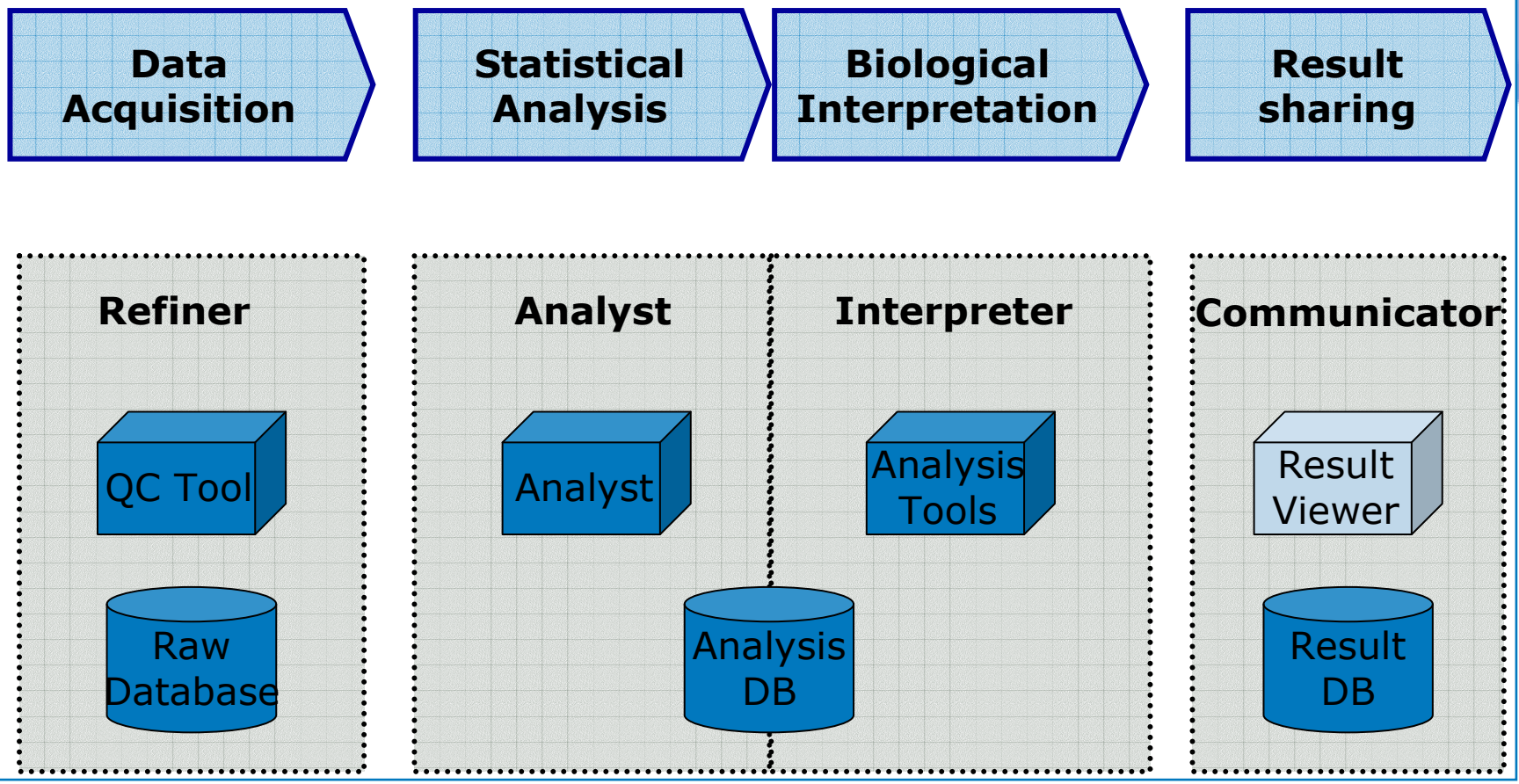


- 18S served as endogeneous control to normalize the data
- Standard curve for each target gene was generated using pooled RNA
- Quantisation of gene expression was determined using the standard nethod



The TaqMan® assay yielded results that were in agreement with the gene array results.

Genedata's Scientific Computing Platform



Results

- Using the Affymetrix platform, reagents and protocols, data from three companies can be compared
- Conclusion: the consortium approach can be scaled

Prerequisites of a Future Systems Biology Data Base

- Highly quality data (unified protocols)
- Appropriate data storage capabilities
- Search tools across endpoints
- Integrated metabolic/physiological interpretation tools
- ....

→ Bioinformatics has to be considered from the very beginning

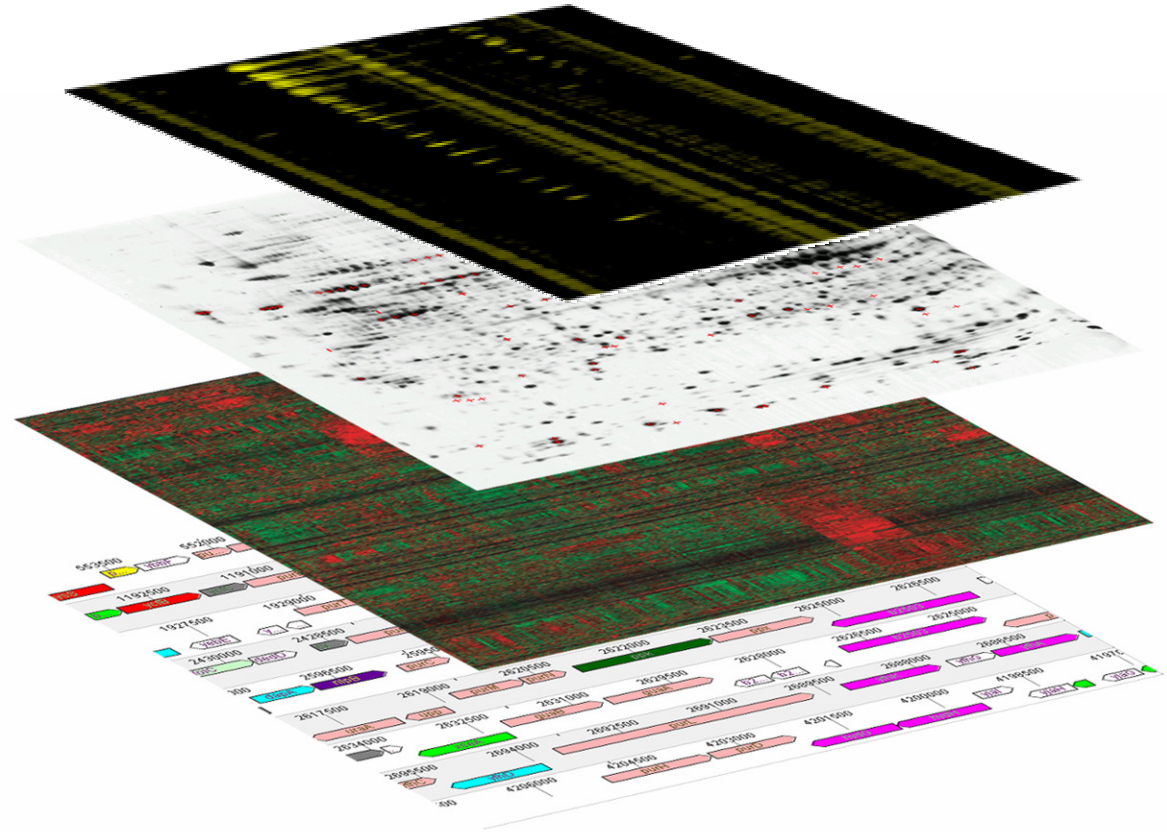
Overview Innomed

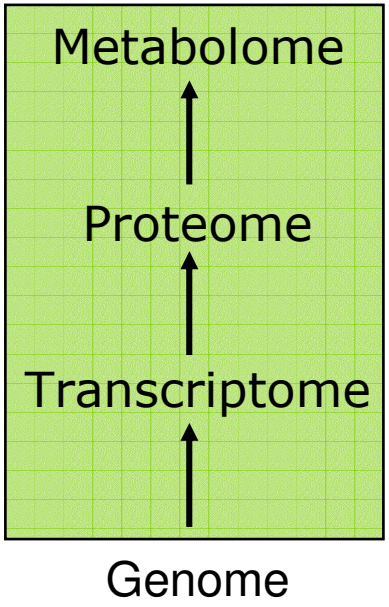
- Consortium approach
- Contributions from all major European pharma companies expected
- Data base open to all contributors
- Includes data for several -omics endpoints
- ....

→ Database may set the stage for predictive systems biology in toxicology

The '-omics' Levels

Expression analysis is performed on three levels





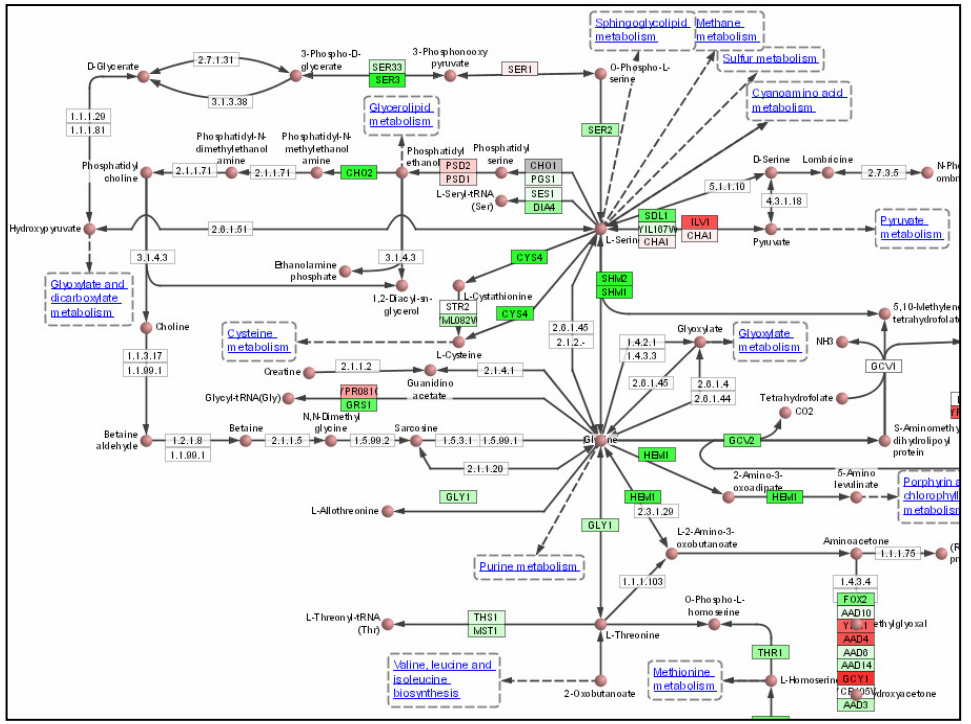
Integration of Cross-level Data

- **What is gained by integrating cross-level data?**
  - A handle on the effective action of the genes
  - Elucidation of biological pathways
  - A deeper understanding of the molecular biology of the [genes → proteins → metabolites] transitions
- **What are the computational challenges?**
  - The dynamical ranges are not constant over the three levels
  - How do we express our knowledge in terms of probabilities?
  - Strong one-to-one correlations are rare
  - Many-to-many interactions within and between 'omics' levels, plus feedback loops → the combinatorics becomes enormous
- **Our approach:**
  - The QA and the statistical and biological data analysis takes place in one single system for all three levels → consistency & integration of data
  - An efficient use of QA, to filter out redundant and faulty data
  - Integration of expression data and pathway maps
  - An overview of the sometimes intimidating amount of data is needed → strong cross-level visualization tools

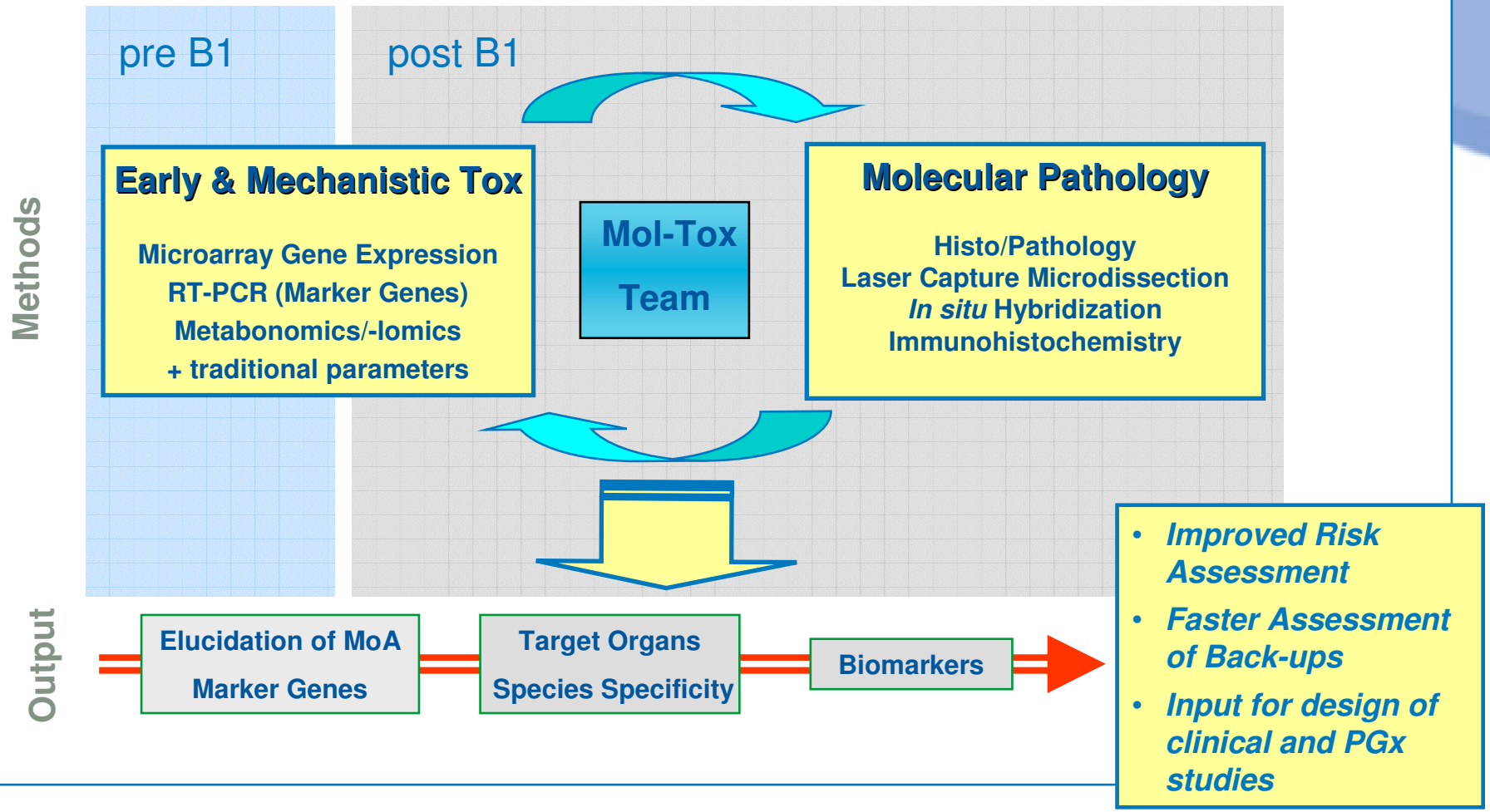
The Core of Integration - the Pathway Map

- At the core of integrating ‘omics’:
  - The maps provide the visual and biological bridge between the different ‘omics’ levels
  - Provides a powerful tool to address the question

- What are the challenges?
  - The actual reconstruction of the pathway
  - Projection of expression levels onto the pathways
  - Checking the consistency of a map
  - Identifying new or unexpected things in a map
  - Automation



The Integrated View



Prediction of Liver Toxicity

Score	Description
0	no liver toxicity
0a	no liver necrosis observed, but at this dose or other doses other forms of liver toxicity observed or expected
0b	no liver necrosis observed, but at other dose or time point expected
1	minimal liver necrosis
2	slight liver necrosis
3	moderate liver necrosis

Cross validation analysis to verify grouping of experiments according to histopathological scores and to check the suitability of the dataset to be used as reference compendium for classification of unknown expression profiles

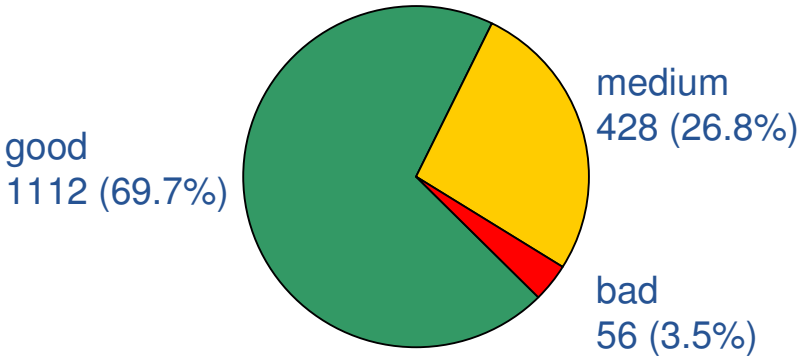
Data-Mining the Phase-1 Database

Phase-1 ToxBank:

- Gene expression profiles from the liver of rats treated once with 52 different toxicants at different concentrations and time points
- Expression profiles measured using Phase-1 Toxchips

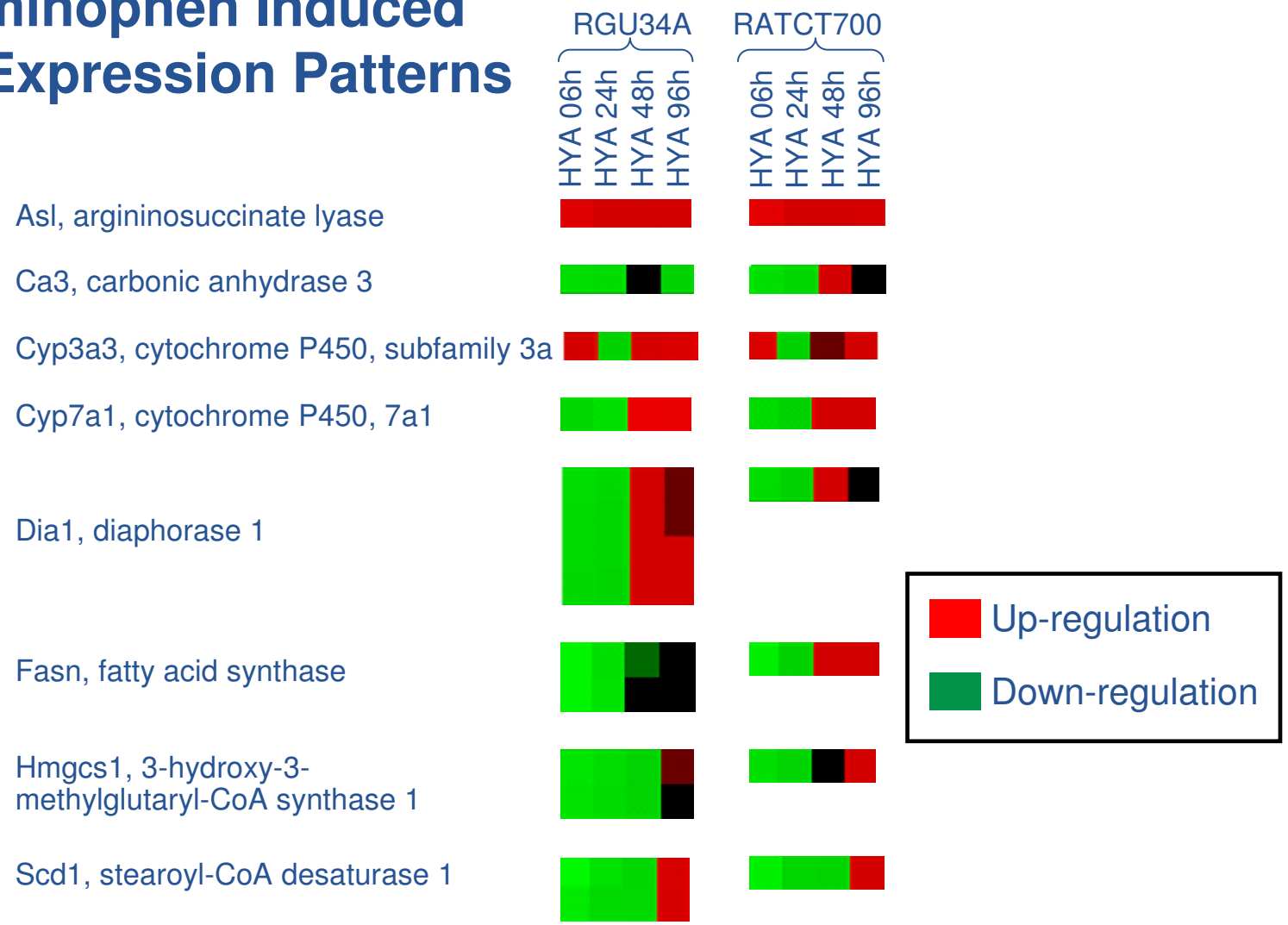
Results and Conclusions

Data quality



→ The Phase-1 ToxBank data set is of good quality.

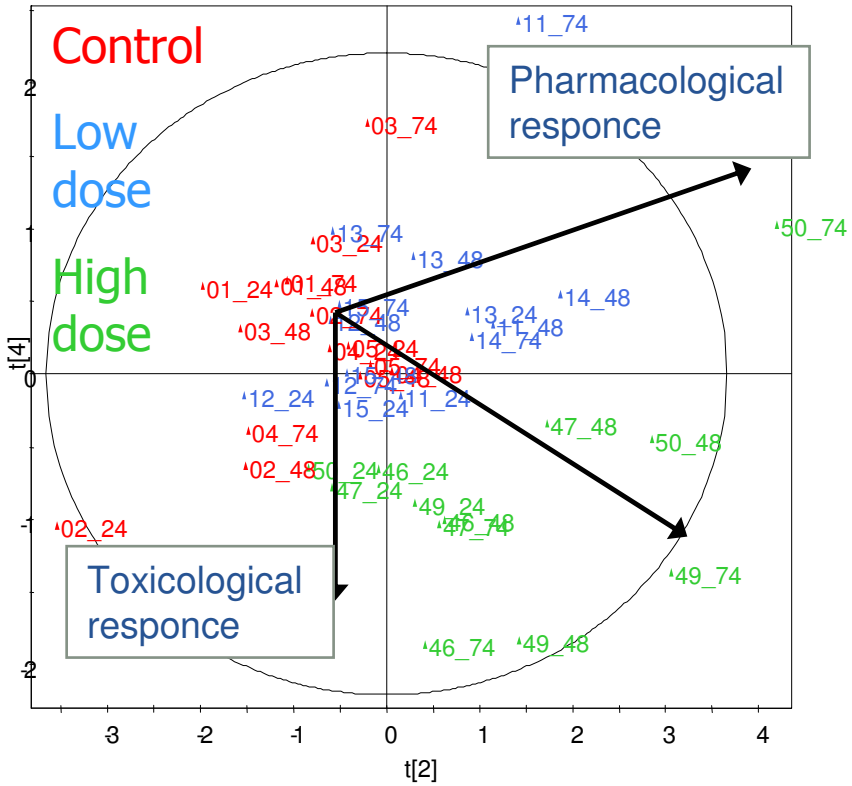
Acetaminophen Induced Gene Expression Patterns

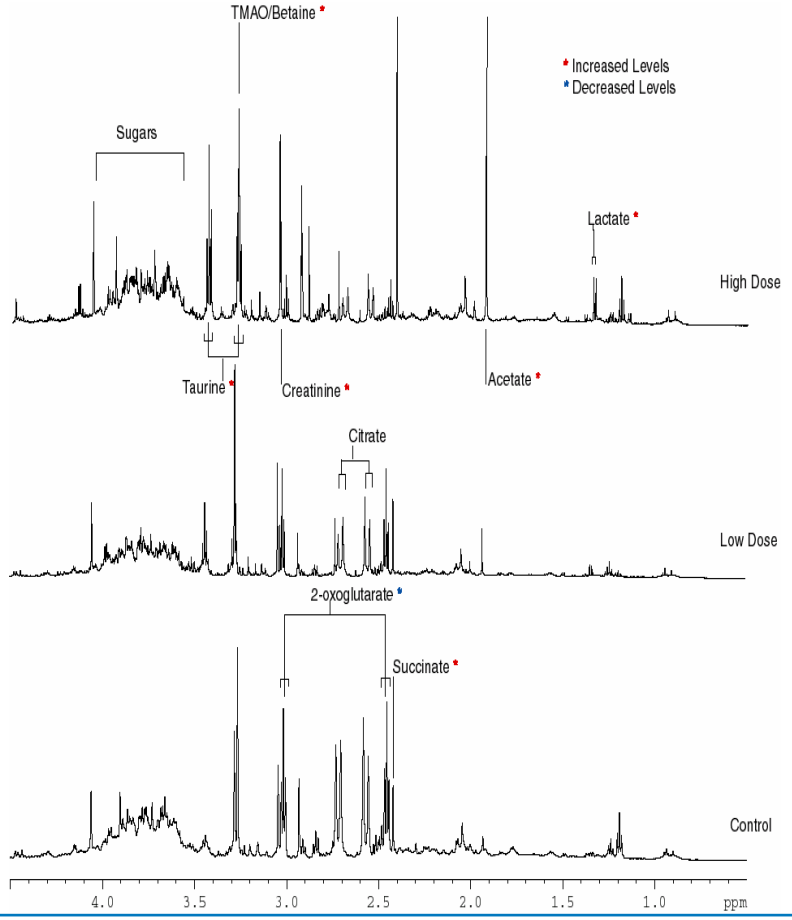


Examples

Methapyrilene (anti-histamine) -removed from market due to hepatotoxicity

^{1}H Urine Spectra at 24, 48 and 74 h





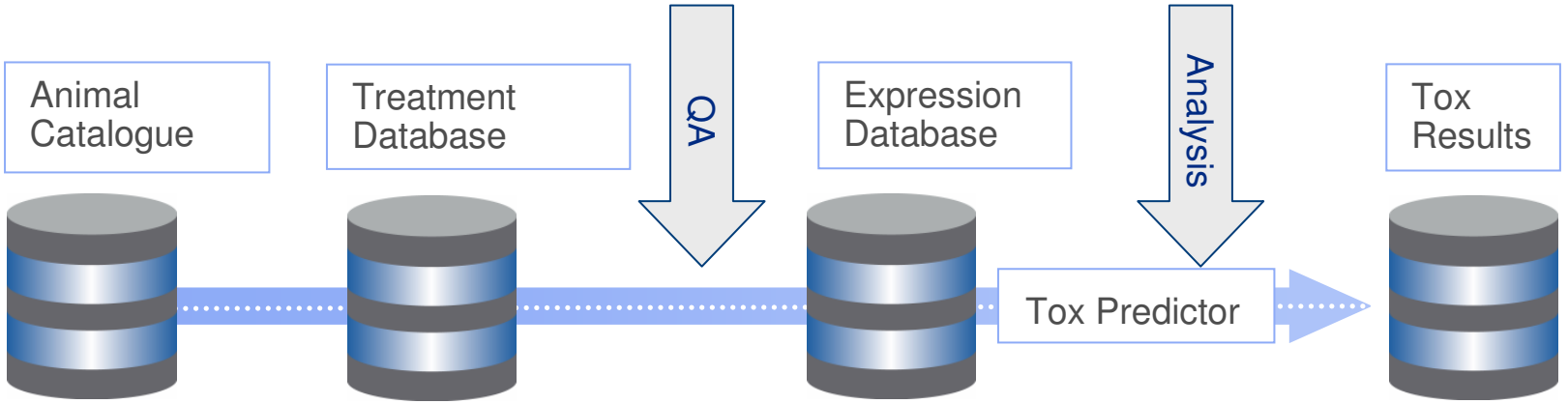
E.Holmes, Imperial College

Systems Biology ? - Putting the Pieces together

- 3-hydroxybutyrate is fuel of respiration esp. in the renal cortex
- 3-Hydroxybutyrate dehydrogenase (BDH) is a lipid requiring mitochondrial enzyme
- high concentrations of DL-3-hydroxybutyrate reduce myocardial infarction size and apoptosis induced by ischemia-reperfusion (Zou et al., 2002) and protects against ischemic brain damage (Suzuki et al., 2002)
- KIM-1, which encodes a type I cell membrane glycoprotein, was originally cloned from post-ischemic rat kidneys (Davis et al., 2004)
- KIM is not detectable in normal kidney tissue but is expressed at high levels in human and rodent kidneys with dedifferentiated proximal tubule epithelial cells after ischemic or toxic injury
- KIM-1 may play an impo
  function to postischemic
- We propose that the she
  constitutes an active mechanism allowing dedifferentiated regenerating cells to scatter on denuded patches of the basement membrane and reconstitute a continuous epithelial layer (Bailly et al., 2002)
- ***"oxidative stress" in kidneys????***
- ***Correlation with other endpoints???***

The classical way is time-consuming, incomplete and not very effective!

Toxicogenomics Database



Species
Strain
Sex
Age
Tissue
Cell line
Histopathology

Compound
Concentration
Treatment time
Dosing route
Dosing frequency

Array type
Labeling
Hybridization
Replicate
Image Analysis
Scanner settings
Raw data

Experimental values
p-Values
Gene Names
Experiment Annotation

Tox classification
Biomarkers
Tox Mechanism

Generic items
Toxicogenomics items