Hazard assessment of chemical carcinogenicity *in vitro* – replacing animal testing in toxicology

Ralf Herwig
Max Planck Institute for Molecular Genetics
Ihnestr. 73, 14195 Berlin
herwig@molgen.mpg.de
• Introduction

• Toxicogenomics for carcinogenic hazard assessment – recent results (FP6 carcinoGENOMICS project)

• Computational network analysis – interpreting genomic data at the network level (FP7 diXa project)

• Outlook and summary
• Introduction

• Toxicogenomics for carcinogenic hazard assessment – recent results (FP6 carcinoGENOMICS project)

• Computational network analysis – interpreting genomic data at the network level (FP7 diXa project)

• Outlook and summary
Research Interests

- Participation in the sequencing of the human – and other – genomes
- High-throughput sequencing – individualisation of genomic information (1000 Genomes Project, Int. Cancer Genome Cons.)
- Epigenetics – identifying modifiers of tumour formation and progression
- Systems biology – network analysis and model development
- Predictive toxicology
- Animal genetics
Toxicity testing

• REACH (June 2007) compliance necessitates the assessment of toxicity for a huge number of chemicals (30,000-68,000) by 2018

• involve a large number of animal testing (13-54 million)

• tests in animals involve a rather long period of time (e.g. two-year rodent bioassay for carcinogenesis)

• toxicity pathways might be different in rodents and human leading, for example, to false positive predictions (several FDA approved drugs are rodent carcinogens)

• use of alternative *in vitro* assays is under development in many laboratories that are hopefully faster and closer to the human *in vivo* situation
1. Develop robust computational methods for analyzing genome data (biomarkers, patterns)

2. Cross-validate the different bits of genomic information and integrate them in order to focus on cancer-relevant pathways

3. Connect these data with computational models that are able to predict the toxic effects of chemicals, ideally from human in vitro assays
The overall aim of the carcinogenomics project is to develop in vitro methods for assessing the carcinogenic potential of compounds, as an alternative to current rodent bioassays for genotoxicity and carcinogenicity. The major goal is to develop a battery of mechanism-based in vitro tests accounting for various modes of carcinogenic action. These tests will be designed to cover major target organs for carcinogenic action e.g. the liver, the lung, and the kidney.

- EU project with 20 academic and industrial partners
- Goal: replacement of animal testing of carcinogenicity of chemical compounds with in vitro systems (liver, lung, kidney, ES-cell derived hepatocytes)
- test three classes of compounds (genotoxic carcinogens, non-genotoxic carcinogens and non-carcinogens) on in vitro systems using Affymetrix microarrays
diXa – Europe-wide data integration system for chemical safety

- EU project with 7 academic and industrial partners
- development of a joined platform
- data collection and pathway-based analysis (ToxDB)
- liver toxicity modelling
- international regulatory activities

Figure 2: the diXa network
• Introduction

• Toxicogenomics for carcinogenic hazard assessment – recent results (FP6 carcinoGENOMICS project)

• Computational network analysis – interpreting genomic data at the network level (FP7 diXa project)

• Outlook and summary
Comparison of in vitro models - liver

Carcinogenesis Advance Access published March 8, 2013

Transcriptional responses generated by hepatocarcinogenesis in a battery of liver-based
in vitro models

Tatyana Y. Doktorova1,2,3, Reha Yildirimman2,3, Mathieu Vinken1, Mireia Vilardell1, Tamara Vanvackel1, Hans Gruenend1, Roque Borr2,4, Gabriela Broder2, Sylvestre Holmgren5, Ranya I. Christophes5, Eos van der Bilt5, Jos Kleijn5, Jos Castell5,6, Peter Bjorjorset5, Ralf Herwitz5 and Vera Rogiers1,2

1Department of Toxicology, Center for Pharmaceutical Research, Vrije Universiteit Brussel, Laarbeeklaan 103, B 1090 Brussels, Belgium
2Department of Vertebrate Genomics, Max Planck Institute for Molecular Genetics, D-14195 Berlin, Germany; Venedig AG, 4085 Basel, Switzerland
3Unit of Experimental Haematology, University Hospital La Fe Valencia, Spain; Facultad de Medicina, Department of Biochemistry and Molecular Biology, University of Valencia, E-46001 Valencia, Spain; 4Collectif, 54-413 46 Gotemburg, Sweden; 5System Biology Research Center, School of Life Sciences, University of Twente, Enschede, The Netherlands
6To whom correspondence should be addressed. Tel: +32 24774372; Fax: +32 24774362; Email: tatyana.y.doktorova@vub.ac.be

As the conventional approach to assess the potential of a chemical to cause cancer in humans still includes the 2-year rodent carcinogenicity bioassay, development of alternative methodologies is needed. In the present study, the transcriptional responses following exposure to genotoxic (GTX) and non-genotoxic (NGTX) hepatocarcinogens and non-carcinogens (NC) in five liver-based in vitro models, namely conventional and enzymatically stabilized cultures of primary rat hepatocytes, the human hepatoma-derived cell lines HepG2 and Hep2 and human embryonic stem cell-derived hepatocyte-like cells (hES-Hep) and two cell lines, HepG2 and HepRG. These models are exposed to 15 prototypical compounds that have been carefully chosen (R). The chemical selection was based on the diversity and selectivity, biochemical and biophysical properties and availability of toxicological information of the compounds. Detailed information can be found in Vinken et al. (8). The selected compounds belong to three toxic classes: (1) genotoxic (GTX) carcinogens (4-aminobiphenyl (4-ABP), 4-nitrofluoranthene (4-NF)); (2) non-genotoxic (NGTX) carcinogens (3,4-dimethylimidazol(4,5-C)-pyridyl)1,3-butanone (NNK); 2-nitrofluorene (2NF); benzo[a]pyrene (BaP); cyclophosphamide (Cyclophosphamide); (3) non-genotoxic (NGTX) anticarcinogens (piretroxide (Piretroxide); Wy-14643 (Wyeth) ; phenobarbital sodium (PBS); 1,2-diethylcarbonylethanol-acetate (TCEA)); and (4) non-carcinogens (NC) (salidroside (SNF), chlorofen (CNF), levimisol (MAN), tolbutamide (TOL), dichlofenac sodium (DSF)). For the data analysis, several biostatistical approaches have been applied.

Materials and methods

Compounds, cell culture and treatment

Five selected liver-based in vitro models (HepG2, HepES-Hep, HepG2 and HepRG) were exposed to 15 compounds, i.e., 5 GTXs (NNK, BaP, AF1 and CYCLO), 5 NGTXs (2NF, 4ABP, SNF, PBS, SNF, TCEA, WY14643, PIRE, MAN, TOL, and DSF) and 5 NCs (SNF, CND, MAN, TOL, and DSF) at inhibitory concentration (IC50) concentrations (reducing cell viability by 50%) allowing to measure specific toxicological effects. Exposure conditions were the same as in Supplementary Table 1 and available at Carcinogenesis Online, showing how.

Pathway responses following GTX exposure

Pathway responses following NGTX exposure

42. Seminar über Versuchstiere und Tierversuche, 11.09.2013, BfR, Berlin
Introducing a stem cell-based hepatocyte system to toxicology

Experimental design:

Hepatic marker ICC screens

a1-AT  AFP
Albumin  CK18
HNF4a  CYP1A2

hES-Hep™002 cell system (hESC cell line SA002)

P Björquist, G Brolén, Cellectis AB Gothenburg, Sweden.

Metabolic competence

Expression of phase I-III genes:

<table>
<thead>
<tr>
<th>Gene</th>
<th>Symbol</th>
<th>Expression in hES-Hep</th>
<th>Expression in hpTep</th>
<th>p-value</th>
<th>Fold</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP1A1</td>
<td></td>
<td>0.006, fold = 3.57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CYP1B1</td>
<td></td>
<td>0.0004, fold = 6.30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AHR</td>
<td></td>
<td>0.002, fold = 1.93</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- variable expression of CYP family genes
- 4148 of 8745 (47%) expressed genes are associated with "metabolic process"
- e.g. biotransformation of BAP is visible in hES-Hep:
Discriminative response genes

- GC-RMA normalization
- compute log2-ratios within each experiment (treated vs untreated)
- built an ANOVA model
- identified 592 genes (P<0.05) discriminative for the tox classes
- 234 genes (40%) associated with metabolic processes
- 87 genes (15%) associated with transcriptional regulation
- tumor suppressors (PDCD4, BCL2, SMAD3, FHIT, ATM, TCHP, ITGB5, RPL10) and oncogenes (RAB17, RRAS, FAS, MDM2, GNA15)


Table 1 Over-representation of pathways with respect to the response gene set

<table>
<thead>
<tr>
<th>Pathway</th>
<th>#Genes</th>
<th>#Overlap</th>
<th>P-value</th>
<th>Source</th>
<th>Response genes in pathway</th>
</tr>
</thead>
<tbody>
<tr>
<td>wnt Ip6 signaling</td>
<td>7 (7)</td>
<td>3.00349</td>
<td></td>
<td>BioCarta</td>
<td>DKK1, FZD1, WNT8A</td>
</tr>
<tr>
<td>atm signaling pathway</td>
<td>20 (18)</td>
<td>4.000991</td>
<td></td>
<td>BioCarta</td>
<td>ATM, GA45A, MDM2, P73</td>
</tr>
<tr>
<td>hop pathway in cardiac development</td>
<td>4 (4)</td>
<td>2.0134</td>
<td></td>
<td>BioCarta</td>
<td>NIK25, SRF</td>
</tr>
<tr>
<td>Aurora A signaling</td>
<td>33 (31)</td>
<td>6.00161</td>
<td></td>
<td>PID</td>
<td>GA45A, MDM2, MIPF2, OAZ1, RASA1</td>
</tr>
<tr>
<td>Stabilization of p53</td>
<td>5 (5)</td>
<td>2.0215</td>
<td></td>
<td>Reactome</td>
<td>ATM, MDM2</td>
</tr>
<tr>
<td>Glucose breakdown (glycolysis)</td>
<td>14 (14)</td>
<td>3.0281</td>
<td></td>
<td>Reactome</td>
<td>GDE, PHKG1, PHKG2</td>
</tr>
<tr>
<td>cysteine biosynthesis II</td>
<td>7 (7)</td>
<td>2.0313</td>
<td></td>
<td>HumanCyc</td>
<td>SERA, SERB</td>
</tr>
<tr>
<td>CREB phosphorylation</td>
<td>7 (7)</td>
<td>2.0424</td>
<td></td>
<td>Reactome</td>
<td>KSR63, KSR6A</td>
</tr>
<tr>
<td>p53 signaling pathway</td>
<td>69 (68)</td>
<td>7.047</td>
<td></td>
<td>Reactome</td>
<td>ATM, GA45A, MDM2, P73, PPM10, SESN1, TNR6</td>
</tr>
<tr>
<td>Branched-chain amino acid catabolism</td>
<td>18 (17)</td>
<td>3.0471</td>
<td></td>
<td>Reactome</td>
<td>AUHM, HBCH, ODBB</td>
</tr>
<tr>
<td>p53 signaling pathway</td>
<td>17 (17)</td>
<td>3.0471</td>
<td></td>
<td>BioCarta</td>
<td>BCL2, GA45A, MDM2</td>
</tr>
</tbody>
</table>

42. Seminar über Versuchstiere und Tierversuche, 11.09.2013, BfR, Berlin
Dominant GTX responses at gene level

- ANOVA judges whether one of the three groups is different from the others
- this is in most cases the GTX group
- GTX response is dominating the outcome of response genes
• Introduction

• Toxicogenomics for carcinogenic hazard assessment – recent results (FP6 carcinoGENOMICS project)

• Computational network analysis – interpreting genomic data at the network level (FP7 diXa project)

• Outlook and summary
Scoring pathway response with gene expression data

- relation between pathways and compounds is measured through gene expression vectors

\[
\begin{align*}
\text{Pathways (K = 1695)} & \quad \text{Genes (N = 18,394)} & \quad \text{Pathways (K = 1695)} \\
\text{Pathway 1} & \quad \text{Gene 1} & \quad \text{Pathway 1} \\
\text{Pathway 2} & \quad \text{Gene 2} & \quad \text{Pathway 2} \\
\text{Pathway K} & \quad \text{Gene N} & \quad \text{Pathway K} \\
\end{align*}
\]

- \( x_{ij} = \frac{1}{\text{Pathway}_i} \)
- \( y_{ij} = |\log_2(R)| \times |\log_{10}(P)| \)
- \( z_{ij} = \frac{1}{\text{Pathway}_i} \times \sum_n |\log_2(R_{nj})| \times |\log_{10}(P_{nj})| \)
Pathway response in hES-Hep

- Pathway response scores can be normalized in order to compare different compounds
- Pathway response can help interpreting effects of single compounds
- Compounds can be correlated at the pathway level

Pathway response can discriminate specific groups of compounds or toxicity classes
Pathway scoring reflects dose
Tox class specific pathway responses

- apoptotic response through FASL higher in GTX than in NGTX (P=0.005)

- PPAR signalling higher in NGTX than in GTX (P=0.009)

- e.g. WYE is a PPAR-a ligand prototypical for peroxisome proliferators

- mechanisms of PPAR-mediated hepatocarcinogenesis have been shown in rodents and there is an ongoing debate whether this mechanism is also operative in human.

Pathway patterns are more stable than gene patterns

Patterns at the pathway level exert a stabilizing effect on gene expression variation and, thus, improve discrimination performance.

N=37 pathway patterns


N=592 gene patterns

42. Seminar über Versuchstiere und Tierversuche, 11.09.2013, BfR, Berlin
ConsensusPathDB - why interaction integration?

Overlap in genes associated with apoptosis in four most commonly used interaction databases
ConsensusPathDB – a meta-database for human functional interactions

- integrated 30 heterogeneous databases on human functional interactions
- ~155,000 molecular species, >360,000 interactions, 4,700 pre-defined human pathways
- heterogeneous interactions
- origin of interaction is preserved
- enables graph searches in the integrated network (e.g. shortest paths, functional modules)
- visualisation features
- enables over-representation and gene enrichment analyses


http://cpdb.molgen.mpg.de

42. Seminar über Versuchstiere und Tierversuche, 11.09.2013, BfR, Berlin
Network analysis functionality update

Molecular concept visualization
- over-representation analysis
- integration of different concepts such as pathways, network neighborhoods, complexes
- overlap and node size features


Computation of induced network modules
- over-representation analysis
- Berger et al., BMC Bioinformatics, 2007
- significance scoring of enriched network modules
- parameter for intermediate nodes

42. Seminar über Versuchstiere und Tierversuche, 11.09.2013, BfR, Berlin
IntScore – improving interaction networks

• implements six different methods for scoring interactions based on network topology or on annotation (+ one aggregate method)

• can be used to construct high-quality interaction networks


http://intscore.molgen.mpg.de
• robust and reproducible protocols for omics data processing and analysis are needed and are generated

• molecular interaction databases (e.g. ConsensusPathDB) are useful tools that can be utilized to improve predictive power of assay systems

• network-based analysis of high-throughput data is robustifying the analysis and improves data interpretation

• stem cell-based assays can be used for hazard identification of chemicals
Structure

• Introduction

• Toxicogenomics for carcinogenic hazard assessment – recent results (carcinoGENOMICS and diXa)

• Computational network analysis – interpreting genomic data at the network level

• Outlook and summary
Toxicogenomics and high-throughput sequencing
New genetic basis for cancer therapy (and also toxicology?)

Catalogues of somatic variations


Catalogues of germline variations in human populations

**Table 1: Summary of 1000 Genomes Project phase I data**

| Samples | Autosomes | Chromosomes | GDRDC0002 sample
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1,092</td>
<td>1,092</td>
<td>1,092</td>
</tr>
<tr>
<td>Total DNA bases (Gb)</td>
<td>150.049</td>
<td>155.088</td>
<td>150.049</td>
</tr>
<tr>
<td>Mean reaped depth (x)</td>
<td>5.1</td>
<td>5.9</td>
<td>5.1</td>
</tr>
<tr>
<td>SRRS</td>
<td>30.7 M</td>
<td>5 M</td>
<td>498 K</td>
</tr>
<tr>
<td>Novelty rate</td>
<td>38%</td>
<td>87%</td>
<td>30 K</td>
</tr>
<tr>
<td>No. by single nucleotide polymorphism (SNPs)</td>
<td>NR</td>
<td>4,766,007</td>
<td>109,239,451</td>
</tr>
<tr>
<td>Average no. SNPs per sample</td>
<td>3.1 W</td>
<td>105 K</td>
<td>24,008</td>
</tr>
<tr>
<td>Indels</td>
<td>1.18 M</td>
<td>59 K</td>
<td>1,357</td>
</tr>
<tr>
<td>Novelty rate</td>
<td>62%</td>
<td>78%</td>
<td>54 K</td>
</tr>
<tr>
<td>No. by single nucleotide polymorphism (SNPs)</td>
<td>NR</td>
<td>191,146</td>
<td>719,090</td>
</tr>
<tr>
<td>Average no. indels per sample</td>
<td>93 K</td>
<td>13 K</td>
<td>40 K</td>
</tr>
<tr>
<td>Genotype relative frequencies</td>
<td>1.18 M</td>
<td>432</td>
<td>847</td>
</tr>
<tr>
<td>Novelty rate</td>
<td>56%</td>
<td>54%</td>
<td>90 K</td>
</tr>
<tr>
<td>Average no. variants per sample</td>
<td>717</td>
<td>26</td>
<td>9</td>
</tr>
</tbody>
</table>

*Not net samples.

High-throughput sequencing and carcinogenicity assessment

• new information (e.g. isoforms, mutations, structural variations) allows a better characterization of the assay system

• new application allow system-wide characterization of compound responses

Network module analysis

-> pathways are natural „data integrators“ and build the opportunity to analyze multiple „omics“ data in a correlated way


-> the biological network contains subnetworks that respond to specific classes of chemical treatments

-> these subnetworks define toxicity read-outs that can be utilized to construct predictive computer models for toxicity and, in general, human health decisions
Thanks for Your attention !