Predictive toxicogenomics for the identification of chemical carcinogens : application to human hepatic cell lines

Citation for published version (APA):

Christina Magkoufopoulou, C. (2011). *Predictive toxicogenomics for the identification of chemical carcinogens : application to human hepatic cell lines*. [Doctoral Thesis, Maastricht University]. Datawyse / Universitaire Pers Maastricht. <https://doi.org/10.26481/dis.20111208cc>

Document status and date:

Published: 01/01/2011

DOI:

[10.26481/dis.20111208cc](https://doi.org/10.26481/dis.20111208cc)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

**Predictive toxicogenomics for the
identification of chemical carcinogens:
*Application to human hepatic cell lines***

Christina Magkoufopoulou

Christina Magkoufopoulou, Maastricht 2011

ISBN: 978 94 6159 096 1

Layout: Lia Rozou

Printing: Datawyse/ Universitaire Pers Maastricht, The Netherlands

Predictive toxicogenomics for the identification of chemical carcinogens: *Application to human hepatic cell lines*

DISSERTATION

to obtain the degree of Doctor at Maastricht University,
on the authority of the Rector Magnificus, Prof. dr. G.P.M.F Mols,
in accordance with the decision of the Board of Deans,
to be defended in public on Thursday 8th December 2011, at 10:00

by

Christina Magkoufopoulou

Supervisor

Prof. dr. J.C.S. Kleinjans

Co-supervisor

Dr. J. van Delft

Assessment Committee

Prof. dr. F.C.S. Ramaekers (chairman)

Prof. dr. D. Kirkland

Prof. dr. P. Lambin

Prof. dr. L.H.F. Mullenders

Prof. dr. H. Schmidt

This research project was supported by grants from the Dutch Technology Foundation STW, applied science division of NWO and the Technology Program of the Ministry of Economic Affairs (STW MFA6809).

Contents

<i>Chapter 1</i>	General introduction	7
<i>Chapter 2</i>	Comparison of HepG2 and HepaRG by whole-genome gene expression analysis for the purpose of chemical hazard identification	31
<i>Chapter 3</i>	Comparison of phenotypic and transcriptomic effects of false positive genotoxins, true genotoxins and non-genotoxins using HepG2 cells	53
<i>Chapter 4</i>	A transcriptomics-based <i>in vitro</i> assay for predicting chemical genotoxicity <i>in vivo</i>	77
<i>Chapter 5</i>	Characterization of the importance of NR0B2 as a classifying gene for genotoxicity prediction: a silencing approach prior and after genotoxic treatment	97
<i>Chapter 6</i>	Transcriptomics-based evaluation in HepG2 cells of human carcinogenic potential of known rodent non-genotoxic carcinogens	119
<i>Chapter 7</i>	Summary and general discussion	143
<i>Chapter 8</i>	Samenvatting en algemene discussie	157
	Acknowledgements	173
	Curriculum vitae and list of publications	177

Chapter 1

**General
introduction**

Molecular biology of cancer

With cancer being one of the leading causes of death in the western world, accounting for approximately 25% of total deaths (1, 2), a lot of effort is made for the understanding of the molecular events that precede carcinogenesis. The main hypothesis for the explanation of cancer development refers to the occurrence of gene mutations. Whereas several types of cancer have been associated with the presence of specific gene mutations, these are not sufficient or required to induce cancer. For instance, although the inheritance of specific mutations of the *BRCA1* or *BRCA2* genes predispose for breast cancer; different studies result in different average risk for breast cancer in *BRCA1* mutation carriers varying from 57% to 90% (3). In a similar manner, whereas mutations in the *APC* gene are responsible for the development of familial adenomatous polyposis, a hereditary disease leading to colorectal cancer, a significant number of patients diagnosed with this disease do not carry these mutations (4). In general, hereditary cancers account for only 5%-10% of all diagnosed cancers, whereas the remaining 90%-95% of cancer incidences, which are characterized as sporadic, are due to environmental exposures, dietary factors, hormones, normal aging and other influences (5, 6).

Independently from the type of cancer, whether it is inherited or sporadic, it is generally acknowledged that carcinogenesis is a very complicated process that comprises several alterations in the cells required for the progressive conversion of normal cells into cancer cells (7, 8). These alterations, at the molecular level, mainly refer to changes in the genome of a particular cell and include point mutations leading to an altered protein product; chromosomal imbalance or instability resulting in an altered expression of a particular gene; chromosomal breakage and rearrangement resulting to loss of a gene or its fusion with another gene leading to the production of a chimeric protein with altered function; and epigenetic modifications to DNA causing gene silencing. Occurrence of mutations to oncogenes or tumor suppressor genes is considered the initial event of the process of carcinogenesis. Oncogenes are defined as genes whose products positively regulate cell proliferation. Mutations to these genes usually result to an increased activity of the proteins that they code for. On the contrary, tumor suppressor genes code for proteins that negatively regulate cell proliferation and, in most cases, mutations to these genes inactivate the protein function (6). Usually, a mutation to one copy of the oncogene is sufficient to alter the function of the protein, whereas for the tumor suppressor genes both copies need to be mutated for the function of the protein to change. Therefore, mutations in oncogenes are called dominant, while mutations in tumor suppressor genes are called recessive (9).

At the cellular level, the alterations that occur provide the cells with growth advantages and can be divided into 8 different types: sustaining proliferative signaling, evading growth suppressors, resisting cell death, enabling replicative immortality, reprogramming energy metabolism, evading immune destruction, inducing angiogenesis and activating invasion and metastasis. Besides the genomic instability described above, another cause that can lead to the acquisition of these properties from the cancer cells is the presence of inflammatory cells in the tumor microenvironment, as inflammatory cells can release

chemicals that are mutagenic for the nearby cancer cells and therefore can accelerate their genomic instability (10).

Chemical carcinogenesis

Although the alterations occurring in a cell that can lead to its transformation into a tumor cell as described above, can happen spontaneously, many times these alterations are caused by exogenous factors, such as chemical compounds. Exposure to these compounds may lead to carcinogenesis. This process of chemical carcinogenesis has been divided in three distinct stages: initiation, promotion and progression (Figure 1) (8).

Figure 1: Chemical Carcinogenesis. After exposure to chemicals the process of chemical carcinogenesis is initiated. The cells can repair the induced DNA damage and return to the normal state or proliferate with damaged DNA entering the promotion phase of chemical carcinogenesis. The cells with DNA adducts as well as the initiated cells may either induce apoptotic mechanisms or mechanisms of toxicity that will lead to cell death or further proliferate entering the progression phase of chemical carcinogenesis. The cells that are in the progression phase will be transformed to cancer cells. [Figure modified from (8)]

Initiation is the first stage of carcinogenesis where the cells accumulate genomic changes caused by the chemical agents. These genomic changes are due to the direct interaction of the chemical or its metabolites with the DNA, via the formation of DNA-adducts (11). At this stage, the cells may avoid tumor transformation by either repairing the DNA-damage or by activating apoptotic mechanisms (12, 13). However, if the cells proceed to cell division while carrying these genomic changes, these changes are irreversible and transmitted to the daughter cells. The chemicals that can cause DNA damage directly or through their metabolites and thus initiate carcinogenesis, are characterized as genotoxic (GTX) (14).

Promotion is the stage of carcinogenesis where the initiated cells show an increased proliferation rate (15) by becoming insensitive to growth constraints (16). Several chemicals have been identified to act as tumor promoters. Although the modes of action of tumor promoters vary significantly and are not yet fully elucidated, it is generally accepted that these chemicals or their metabolites do not damage DNA and thus they are characterized as non-genotoxic (NGTX). For several of these chemicals it has been proven that they act through binding to nuclear receptors, resulting to alterations in gene expressions and/or via the production of reactive oxygen species (ROS) that can affect DNA integrity (17, 18). In addition in the presence of these chemicals apoptosis and intercellular interactions can be inhibited, allowing the initiated cells to escape growth regulatory signals of the surrounding normal

cells or can result to the stimulation of cellular proliferation (19). Other mechanisms through which NGTX carcinogens act include immunosuppression and endocrine disruption. The effects of tumor promoters are not irreversible and carcinogenesis can be inhibited if the exposure to the chemical is discontinued (18).

Progression is the last and irreversible stage of carcinogenesis and encompasses also the processes of angiogenesis, invasion and metastasis. In this phase cellular alterations occur regarding the biochemical, metabolic and morphological characteristics of the tumor cells, transforming the tumor from benign to malignant (8).

Genotoxicity and carcinogenicity assessment of chemicals

Taking into consideration the contribution of chemical exposure to the development of cancer, it is very important to evaluate the genotoxic and carcinogenic properties of chemicals to which humans are or will be exposed to. Therefore, regulatory authorities require the performance of specific assays for the evaluation of these properties of chemicals depending on their application.

For instance the International Conference on Harmonization (ICH) has produced guidelines for the safety testing of pharmaceuticals intended for human use (20, 21). The guidelines for genotoxicity testing proposes either the performance of two *in vitro* assays, one in bacterial for mutations and one in mammalian cells for the inductions of mutations or chromosomal damage, in combination with an *in vivo* assay for the induction of chromosomal damage, or the performance of one *in vitro* bacterial mutation assay in combination with two *in vivo* assays using two tissues. The carcinogenicity assessment, according to ICH, should only be performed when several conditions are met, among which is the induction of genotoxic effects by the pharmaceutical under investigation (22). When these conditions are met, the proposed carcinogenicity assessment includes either a long-term carcinogenicity study in two rodent species or a long-term rodent carcinogenicity study in combination with a short or medium-term study such as the initiation-promotion model in rodents, the use of transgenic mouse assays or the neonatal rodent tumorigenicity model (Figure 2a) (20).

In contrast to the obligatory evaluation of hazardous properties of pharmaceutical chemicals, a large number of other substances has been manufactured and placed on the European market for many years without sufficient information on their hazardous effects for human health and for the environment. Therefore, in order to improve the protection of human health and the environment, by collecting additional information on the chemicals, the European Regulation for the Registration, Evaluation, Authorisation and Restriction of Chemical substances (REACH) is in effect since June 2007 (EC 1907/2006). According to the Regulation, depending on the amount of the chemical substance that is imported or manufactured in Europe, different genotoxicity and carcinogenicity assays should be carried out. For chemicals that are imported / manufactured in quantities more than one tonne, a bacterial mutation assay is sufficient for genotoxicity evaluation. However, in case of a positive result further studies should be conducted (EC 1907/2006 ANNEX VII). For those chem-

icals for which their imported / manufactured amounts exceed 10 tonne, an additional *in vitro* assay, using mammalian cells, for chromosomal damage is required. In case that the results of the bacterial mutation assay or the *in vitro* mammalian assay for chromosomal damage are negative, then an additional *in vitro* mammalian mutation assay is required. *In vivo* testing will also be necessary in case of positive outcomes of the *in vitro* mammalian assays (EC 1907/2006 ANNEX VIII). When the chemical is imported / manufactured in quantities more than 100 tonne and is positive *in vitro* in genotoxicity assays, additional *in vivo* genotoxicity testing is required (EC 1907/2006 ANNEX IX). Finally, for chemicals that are imported / manufactured in quantities that exceed 1000 tonne a carcinogenicity study may be required if there is evidence of frequent or long-term human exposure or inconclusive results regarding their genotoxic properties. However, in cases that the *in vitro* and *in vivo* genotoxicity assays are positive then the default presumption is that this chemical is carcinogenic via a genotoxic mechanism and therefore no carcinogenicity assessment is required (EC 1907/2006 ANNEX X) (Figure 2b).

A. Genotoxicity and carcinogenicity assessment for pharmaceuticals

B. Genotoxicity and carcinogenicity assessment for industrial chemicals

Figure 2: Guidelines for the testing of chemicals. Simplified flow chart of the genotoxicity and carcinogenicity testing of pharmaceutical chemicals according to ICH (2a) and of chemicals imported or manufactured in EU according to the REACH legislation (2b).

Similarly, substances that are intended for use in food products have to be evaluated for their hazardous properties. In the past the required testing for genotoxicity included a battery of three *in vitro* genotoxicity assays – bacterial and mammalian mutagenicity assays and a mammalian assay for chromosomal damage. Positive results in any of the above *in vitro* tests normally required further assessment of genotoxicity *in vivo*. Regarding carcinogenicity, the substance should be tested in two species, usually rats and mice (23). However, since 2002 the European Food Safety Authority (EFSA) provides scientific advice on the hazardous properties of chemicals on a case-by-case approach, therefore no specific testing requirements exist for new food ingredients (EC 178/2002). Cosmetics products and ingredients also need to be safe for human use. However, European Commission imposed a ban on animal testing of finished cosmetic products, since 2004, and ingredients or combinations, since 2009 (Directive 2003/15/EC). Therefore, for the evaluation of the safety of cosmetic ingredients and products only *in vitro* testing is accepted. Regarding genotoxicity testing, a list of approved *in vitro* assays is provided in Annex V of Directive 67/548/EC. However, the fact that animal testing is prohibited and there are no *in vitro* assays for carcinogenicity testing that will prove the safety of the cosmetic ingredients, implies that no new ingredients can be used in cosmetic products; and therefore no testing can be conducted.

With all these chemicals that need to be evaluated and registered regarding their toxic effects, the assays that are used for this evaluation should be standardized. Therefore the Organisation for Economic Co-operation and Development (OECD) has created the Guidelines for the Testing of Chemicals of the most relevant internationally agreed testing methods that are used by government, industry and independent laboratories for the identification and characterization of potential hazards of new and existing chemical substances, chemical preparations and chemical mixtures.

Assays for genotoxicity assessment of chemicals

The most commonly used assays for genotoxicity assessment for which OECD has provided with guidelines will be shortly discussed here in relation to the endpoints that they measure and their limitations.

Ames bacterial mutation assay

As mentioned above, independently of the reasons for the evaluation of genotoxic properties of chemicals, a bacterial mutation assay is always required. Although the initial assay developed by B.N. Ames (24) was using *Salmonella typhimurium* strains, the current OECD guidelines refer also to the use of *Escherichia coli* strains. The bacterial strains that are used for this evaluation carry certain mutations that do not allow them to grow in the absence of specific amino acids. The principle of this bacterial reverse mutation test is that in the presence of a mutagen, mutations will occur that will revert the mutations present in the test strains and therefore restore the functional capability of the bacteria to synthesize the essential amino acid. Thus the revertant bacteria will be detected by their ability to grow in the absence of the amino acid required by the parent test strain (25).

Although in principle this assay can detect direct mutagens, the fact that it uses prokaryotic cells, which differ significantly in uptake, metabolism, chromosome structure and DNA repair processes from mammalian cells, results to several shortcomings when the data are to be used for assessment of human health risk. Especially due to the limitations and differences in metabolic activity of prokaryotic with mammalian cells, an exogenous source of metabolic activation has to be used in order to evaluate the mutagenic properties of the metabolites of the chemical under investigation (24).

***In vitro* mammalian cell gene mutation assay**

An *in vitro* method for the identification of mutagenic and/or carcinogenic chemicals which is more relevant to human health is the mammalian cell gene mutation assay. For this assay, cells expressing thymidine kinase (TK) are being used. This enzyme phosphorylates thymidine to thymidine monophosphate (TMP), which serves as a regulator for DNA synthesis. If TMP is replaced by other lethal analogues, like trifluorothymidine (TFT), the cells will not survive due to the inhibition of cellular metabolism and cell division arrest. In the presence of a mutagen, it is expected that mutations will occur at the TK locus resulting to thymidine kinase deficient cells. The mutant cells therefore will be insensitive to TFT and will be able to proliferate in its presence (26). Similarly, cells proficient in hypoxanthine-guanine phosphoribosyl transferase (HPRT) or a transgene of xanthine-guanine phosphoribosyl transferase (XPRT) are used that are sensitive to 6-thioguanine (TG) or 8-azaguanine (AG), respectively. In the presence of a mutagen, these cells will become resistant to the lethal effects of TG or AG (27). For these assays the most suitable test systems include mouse lymphoma cells, chinese hamster cells and human lymphoblastoid cells (28).

Although the results of these assays appear to be more relevant to human health risk assessment and are capable of identifying chemicals that can cause both point mutations and small deletions, several studies show that these assays are over-sensitive resulting to positive results for numerous compounds that do not appear to be genotoxic *in vivo* (29, 30). Furthermore, even in the mammalian *in vitro* test systems, enzymes involved in the metabolism of xenobiotics are lacking or are expressed to a limited extent. Therefore, the *in vitro* assays are performed in both the absence and the presence of an exogenous metabolic system, usually the rat liver S9-mix. However, studies have shown that in the *in vitro* systems the addition of S9-mix, still does not fully result in the same metabolism of compounds as *in vivo* (31).

***Clastogenicity* assays**

For the initiation of chemical carcinogenesis, as mentioned earlier, genomic instability plays a central role. Hence, besides the detection of induced point mutations or deletions by a chemical agent, it is important to identify those chemicals that can cause chromosomal damage. OECD provides guidelines for two such assays that can be performed both *in vitro* and *in vivo*; namely the micronucleus test and the chromosome aberration test (32-35).

The micronucleus assay detects the formation of micronuclei in dividing cells during mitosis that can be due to numerical or structural chromosomal changes. In more detail,

aneugens, chemicals that result in the loss or gain of whole chromosomes, may increase chromosome mis-segregation to the spindle. The chromosomes that will stay behind during anaphase will be then surrounded by the nuclear envelope and form micronuclei. Similarly, clastogens, agents that cause chromosomal breakage, will result to the formation of acentric fragments that will not be segregated along the spindle and thus will not be integrated in the daughter cells but will be left behind during mitosis, thus forming micronuclei (36).

The chromosome aberration test is used for the detection of structural chromosomal aberrations induced by test compounds and may be of two types, chromosome or chromatid aberrations. These aberrations are the result of breaks in the DNA strand and the subsequent recombination of the broken extremes. However double-strand breaks can take place either directly or through the process of DNA repair or synthesis (37, 38).

For the *in vitro* assays, OECD proposes the use of several cell types (32, 34). Besides the obvious limitations of the use of an *in vitro* system, the use of these assays has also been proven to yield a high number of positive results that are not reproducible *in vivo*. These positive results occur without direct damage to the DNA and, in most of the cases, under conditions that would not be relevant *in vivo* (39, 40). *In vivo* the use of bone marrow or peripheral blood of mice or rats is recommended (33, 35). This limitation for the use of only those specific targets, may lead to the wrong evaluation of genotoxic properties of chemicals that target other tissues or organs.

Unscheduled DNA synthesis assays

Although DNA damage is one of the most important initiative events for tumorigenesis, as mentioned previously, cells have the ability to repair the induced damage caused by chemical agents (12). Therefore, several assays have been developed that detect the activation of DNA repair mechanisms as a consequence of exposure to genotoxic / carcinogenic agents. The unscheduled DNA synthesis (UDS) test aims to measure the activation of the base and/or nucleotide excision repair process in cells that do not undergo DNA replication, and therefore they are not in the S-phase of the cell cycle. In more details, at the DNA damaged sites, the nucleotides or bases will be excised and replaced. By using radio-labeled nucleotides, such as tritium-labelled thymidine (^3H -TdR), and autoradiography the amounts of labeled nucleotides incorporated in the DNA can be measured, indicating that the repair process has been activated (41).

OECD provides guidelines for the performance of this assay both *in vitro* and *in vivo* (42, 43). *In vitro*, several cell types can be used for the UDS test; however for the *in vivo* assays liver is strongly recommended since it is the organ that is responsible for the metabolism of most chemicals. However, although UDS test appears to be a useful tool for the evaluation of DNA damage responses, there is a high possibility that the DNA damage caused by the chemical agent will not be repaired leading to false negative results regarding the genotoxic effects of the chemicals.

Assays for carcinogenicity assessment of chemicals

Initially the above mentioned *in vitro* assays for genotoxicity assessment have been con-

sidered as a possible tool for the evaluation of carcinogenic potential of chemicals (44). However, soon it was known that particular classes of chemicals are capable of causing tumorigenic effects without directly damaging the DNA (19). In addition, the studies of the National Toxicology Program (NTP), which aims to evaluate the toxicological effects of chemicals of public health concern that are used in the United States, show that several chemicals that appear to be negative in the *in vitro* assays are indeed rodent carcinogens. It has been shown that the best approach for the detection of rodent carcinogens *in vitro* would be the combination of the Ames mutagenicity assay and the micronuclei assay (45); however both Ames and micronuclei assay, as well as most of the other *in vitro* genotoxicity assays, result in high numbers of false positive outcomes for rodent carcinogenicity, reaching even 50% (46). The above mentioned explains why the only accepted assays for carcinogenicity assessment of chemicals are currently the *in vivo* rodent assays.

For several years, the long-term rodent bio-assays have thus been considered as the golden standard for carcinogenicity assessment. According to the OECD guidelines, the test substance should be administered daily for a period of 24 months to animals of both sexes. In addition, at least three different doses have to be tested. For statistical analysis, the use of at least 50 animals per dose group and control group is required. The number of experimental animals is further increased in cases that interim kills have to be conducted. In those cases 10 additional animals have to be used per each dose and control group for each sex. By the two-year rodent bioassay several types of information can be acquired. Firstly the target organs of carcinogenicity, the time of appearance of the neoplasms and the tumorigenic dose-response relationship, can be identified. Secondly for non-genotoxic carcinogens, the no-observed-adverse-effect level (NOAEL) can be characterized. Further, additional data regarding mode of actions can be acquired. All these in combination will eventually lead to the extrapolation of carcinogenic effects to low dose human exposure levels (47). For most of the carcinogens, that is the genotoxic carcinogens, however, there is no level of exposure that does not pose a small, but finite, probability of generating a carcinogenic / toxic response. Those carcinogens are characterized as non-threshold carcinogens. In those cases, different models can be used for the extrapolation of the expected toxic responses in humans from the responses observed for experimental animals. Those models are based on tolerance distributions and/or mechanistic assumptions. The most commonly used models are the linearised multistage (LMS) model, which reflects the theory of carcinogenesis that for a normal cell to express a malignant phenotype a series of irreversible changes or stages is required; the benchmark model, that uses linear extrapolation downwards from a benchmark dose; and the dose descriptor T25 that based on the chronic dose rate which will give tumors to 25% of the animals the corresponding human dose descriptor will be calculated by appropriate scaling for interspecies differences on metabolic rates (48).

The rodent bio-assays, however, are time-consuming and very expensive. Therefore several efforts have been made for the development of medium- and short-term bioassays and some of them have been accepted by regulatory authorities, such as the initiation/promotion model and the transgenic mice models (20).

The initiation/promotion model for carcinogenicity testing lasts for 8 weeks and is based on the hypothesis that some chemicals are able of initiating carcinogenesis whereas other chemicals can only promote carcinogenesis. To evaluate the potential of a chemical to act as a tumor promoter, the animals are initially treated with a known carcinogen, such as diethylnitrosamine (DEN), to initiate carcinogenesis and after a 2-week recovery period they receive the test compound. In addition, to evaluate the potential of a chemical to initiate carcinogenesis, the animals are initially injected with saline instead of DEN solution and after a 2-week period they are subjected to the administration of the test compound. The carcinogenic potential of the chemicals can, after an additional period of 6-weeks, be detected by the measurement of altered liver foci, which are early indicators for carcinogenicity (49, 50).

Several transgenic mice models have been developed for the evaluation of carcinogenic potential of chemicals. The most commonly used include the *Trp53*^{+/-}, Tg.AC, and *RasH2* transgenic models (51) and the double knockdown mice *p53*^{+/-}, *XPA*^{+/-} (52). The first model is developed based on the deregulation of the *Trp53* tumor-suppressor gene. The product of this gene is the p53 transcription factor that regulates the activity of a variety of genes involved in cell cycle arrest, apoptosis, anti-angiogenesis, differentiation, DNA repair, and genomic stability. This protein suppresses cancer development in humans and rodents and mutations on this gene have been identified in more than 50% of all cancers (53). The next two models are based on the deregulation and over-expression of the *ras* oncogene. The *H-ras* protein plays an important role to cell proliferation through signaling by growth factors mediated via the mitogen-activated protein kinase kinase (MAPKK) pathway. Similarly to the *Trp53* tumor-suppressor gene, mutations at the *ras* oncogene are often observed in both human and rodent cancers (54). The use of these models results to the appearance of neoplasms after relatively short exposures and therefore the duration of these assays is reduced to approximately 26 weeks. Several studies using the transgenic mice models have shown that they do not fully reproduce the results obtained from the long-term two-year rodent bioassay (51, 55, 56). The double knock-down mouse model takes advantage of the deregulation of two genes, the *Trp53* tumor-suppressor gene and the *XPA* gene that its product is involved in DNA repair, and therefore may be more sensitive in identifying both genotoxic and non-genotoxic carcinogens (52).

However, the main reason for performing rodent bioassays always remains the human health risk assessment. The relevance of experimental bioassays to the assessment of human health risk is based on the hypothesis that because rodents and humans are mammals, they should share many basic genetic, pharmacologic, toxicologic, and carcinogenic responses. Furthermore, there is the fact that almost all known human carcinogens, for which appropriate *in vivo* testing has been conducted, are also carcinogenic in rodents (57).

Yet, there is still no agreement on the interpretation of the data resulting from the rodent bioassays. For some scientists a chemical is considered carcinogenic to humans when it can cause cancer at any rodent group tested, whereas for others a chemical can be considered as human carcinogen when it can cause cancer to both mice and rats. For the ten known human carcinogens that were tested among hundreds of chemicals by NTP, only six were tested in both species, from which only three were positive in both, while one was

negative in both species. If we accept that a chemical is a human carcinogen when it causes cancer in either rats or mice, then the sensitivity appears to be 90%, but 10% of the human carcinogens are still false negatives. If we define a human carcinogen as a chemical that can cause cancer in both mice and rats then the sensitivity is only 50%. Specificity, on the other hand, cannot be estimated directly, since epidemiologic studies only give an upper limit on potency, but they do not prove lack of carcinogenicity. However, according to Ennever and Lave, the specificity in the first case would be 78% whereas in the second case it would be 93% (58). Regardless of which approach is used for the estimation of the human cancer risk based on rodent bioassays, it is clear that they do not fully represent the true human risks. The extrapolation of human carcinogenic risks from rodent bioassays has been further challenged in regard to the differences between rodents and humans in relation to the metabolism of the chemicals, the susceptibility to the chemicals and the molecular events occurring after exposure to the chemicals (59-62). These differences make the assessment of human health risk based on rodent bioassays even more complicated (63).

Alternative assays for genotoxicity and carcinogenicity assessment of chemicals

All these limitations of the current testing system for the evaluation of the genotoxic and carcinogenic properties of chemicals have several consequences. Firstly for the confirmation of the *in vitro* genotoxic effects of chemicals, *in vivo* assays are required. Taking into consideration the high false positive results of the standard *in vitro* test battery, additional animal experiments are required that would be avoided if a more accurate *in vitro* test system would be available. Further, the increased use of experimental animals results to the loss of significant investments in costs and time that would again be avoided by the availability of an accurate *in vitro* genotoxicity testing system. However, in the development of new chemicals, for instance in the pharmaceutical industry, a lot of compounds are discarded in the pre-selection process after positive outcomes of the *in vitro* genotoxicity assays. In view of the fact of the high false positive rates, this implies that many promising chemicals are discarded based on false premises. Finally and most importantly, the available assays do not represent the true human health risks.

Taking into consideration the consequences of the previously mentioned limitations of the currently genotoxicity and carcinogenicity testing system together with the current legislations that aim to the reduction, refinement and replacement of animal experimentation (EU Directive 2010/63/EU), alternative, more accurate and more reliable *in vitro* methods for the genotoxicity and carcinogenicity assessment of chemicals are urgently required. Therefore, several organizations, among which the European Center for the Validation of Alternative Methods (ECVAM), aim to contribute to and support the development, validation and acceptance of alternative approaches that could reduce, refine or replace the use of laboratory animals and could provide the same or higher levels of information as those obtained from experiments using animals (EU Directive 2010/63/EU).

Several laboratories, therefore, are currently focusing on the development of such alternative *in vitro* methods for genotoxicity and carcinogenicity testing. Among these methods, the most valuable for genotoxicity testing appear to be the *in vitro* reconstructed skin micronucleus (RSMN) assay, the Comet Assay and the *GADD45a-GFP* genotoxicity assay (64-66), whereas for carcinogenicity evaluation the *in silico* methods and the BALB/c 3T3 cell transformation assay appear to be able to replace the rodent bioassays.

The RSMN assay is based on the measurement of micronuclei formation, similarly to the *in vitro* micronuclei assay described earlier; however it is using human reconstructed skin models therefore providing a more relevant to human health target for the evaluation of the genotoxic properties of chemicals. In addition, this assay can be very useful for the evaluation of genotoxic properties of cosmetics ingredients since the *in vivo* genotoxicity testing for cosmetics marketed in Europe was banned in March 2009 (EC-Directive 2003/15/EC); however, it cannot be applied for the evaluation of genotoxic properties of other chemicals such as pharmaceuticals or food additives. Furthermore, although a successful study has been conducted for the evaluation of the reproducibility of the results of this assay among different laboratories, the number of chemicals tested in this assay was very limited. In particular only one non-genotoxic and three genotoxic agents were used (64).

The Comet assay, that can be performed both *in vitro* and *in vivo*, is based on the assessment of the migration of DNA in an agarose matrix under electrophoretic conditions, as an indication for induced DNA damage. In particular, a cell with damaged DNA will have the appearance of a comet, with a head (the nuclear region) and a tail containing DNA fragments or strands. A limitation of the comet assay is that it fails to detect aneugenic effects (65).

The *GADD45a-GFP* genotoxicity assay is based on the induced transcription of the *GADD45a* gene by a genotoxic compound and the consequent synthesis of green fluorescent protein (GFP) using a human lymphoblastoid cell line. This assay has been proven to perform better in relation to both sensitivity and specificity when compared to the currently available *in vitro* assays based on a study using 20 non-genotoxic and 32 genotoxic agents, resulting to a possible alternative to the current *in vitro* genotoxicity assays. However, this assay does not distinguish between direct and indirect acting genotoxins. In particular chemicals that result in the production of Reactive Oxygen Species that may damage DNA give positive results in the *GADD45a-GFP* assay (67).

In silico methods for carcinogenicity assessment are based on the concept of structure–activity relationships (SAR) and on the hypothesis that chemicals sharing similar structures may cause similar toxicological effects. These substructures or functional groups that are mechanistically linked to the carcinogenic activity of the chemicals are called structural alerts (SAs). Only the presence of SAs, though, is not sufficient for predicting the carcinogenic properties of a chemical since the potential carcinogenic effect of the SA may be diminished or completely eliminated depending of the characteristics of the rest of the molecule. Therefore, an improvised approach includes the use of quantitative structure–activity relationships (QSARs), which takes also into consideration the combination of physical, chemical and structural properties in order to produce a mathematical model of the

chemical determinants of the biological activity (44, 68). However, the application of these predictive models for carcinogenicity is limited only for chemicals of known carcinogenic classes, and therefore cannot be applied for novel compounds that do not share structural similarities with these classes.

On the contrary to the *in silico* methods, the BALB/c 3T3 cell transformation assay measures morphological changes induced in embryonic mouse fibroblasts *in vitro* by a chemical agent. It is based on the hypothesis that the *in vitro* cell transformation represents the various stages of *in vivo* carcinogenesis (69). However, a study applying 168 chemicals for the evaluation of this assay to predict the carcinogenic properties of chemicals resulted in a very low specificity (60%) in regard to the results of the rodent bioassay (70) and currently efforts are made to improve the protocol (69), while ECVAM has recently conducted a pre-validation study for this assay for the development of standardised protocols. The pre-validation study showed that the assay is transferable between laboratories, and reproducible within- and between-laboratories. Furthermore, it is concluded that the assay is valuable to detect rodent carcinogens (71). Nevertheless, this assay still applies the use of rodent cells, raising again the question of the biological relevance of this test system to human health.

In conclusion, the above alternative assays have not yet been fully proven to correctly evaluate the genotoxic and carcinogenic properties of chemicals. In view of their limitations, that resemble those of the currently used testing systems, still a better alternative approach has to be developed.

Predictive toxicogenomics

In contrast to the alternative methods described above, for the evaluation of genotoxic and carcinogenic properties of chemicals that are based mainly on the evaluation of phenotypic endpoints, the field of toxicogenomics has been considered as a promising tool for class discrimination of chemicals. Toxicogenomics is defined as the measurement of the gene expression pattern elicited by toxicant exposure. These gene expression patterns are expected to facilitate the identification of the mechanisms of action of chemicals (72). Therefore, in predictive toxicogenomics it is expected that chemicals which share the same mechanism of action, will also share the same gene expression patterns. The use of these gene expression patterns, is expected to further result to the selection of a subset of commonly regulated genes that can be used as biomarkers to predict modes of action (Figure 3) (73).

In an effort to confirm the hypothesis of the commonly shared gene expression patterns among compounds of the same toxicity class, several studies have focused on the comparison of transcriptomics responses after exposure to genotoxic and / or carcinogenic compounds. For these studies different models and different compounds have been selected. Most of these studies have been conducted for short term *in vivo* exposures of rats, where RNA was isolated from the liver of rats treated either for a single dose or after repeated daily administration of the test compounds for up to 28 days. Predictive toxicogenomics approaches have been further applied for the generation of gene expression signatures for a variety of compound

Figure 3: Principle of predictive toxicogenomics. Gene expression signatures are developed for each group of known agents with different modes of action. The gene expression profile of an unknown agent, the “suspected toxicant”, is compared with the gene expression signatures of the known groups for finding a match. In this example the suspected toxicant belongs in Group A. [Figure modified from (72)]

classes including rat hepatocarcinogens, genotoxic and non-genotoxic rat carcinogens, non-genotoxic carcinogens and non-carcinogens, as well as for compounds that share the same mode of action, such as peroxisome proliferators or enzyme inducers. All these studies result in the selection of genes that are known to play a central role in the development of tumor formation and are involved in important cellular processes, such as responses to oxidative stress, cell cycle, cell proliferation, apoptosis, metabolic pathways and others. Furthermore, class prediction using these approaches *in vivo* results in high accuracy, reaching even 95%, confirming the hypothesis of predictive toxicogenomics *in vivo* (74-79). Further, an interlaboratory comparison was conducted for two of these studies that aimed in the prediction of non-genotoxic carcinogens resulting in an accuracy of 60-65%, indicating that validation of the signatures is required prior to their application as alternative methods (80).

However, according to the current legislations, the reduction of animal experimentations is a necessity (EU Directive 2010/63/EU). The accurate results of the *in vivo* toxicogenomics approaches for class prediction, creates new opportunities to investigate the use of *in vitro* toxicogenomics approaches for this purpose. Therefore, toxicogenomics studies for the purpose of class prediction have also been conducted using primary hepatocytes. In a study using rat primary hepatocytes for the selection of genes that can discriminate between compounds of different modes of action, i.e. peroxisome proliferators from AhR agonists, fourteen of the fifteen compounds used were assigned to the correct class. In addition, half of the genes identified by the toxicogenomics analyses have been associated with these specific modes of action (81). Another study applying toxicogenomics approaches for the discrimination between genotoxic and non-genotoxic carcinogens by using mice primary

hepatocytes, resulted also in correct classification of all the 6 chemicals used. Furthermore, the developed gene expression profiles included genes known to be involved in biologically relevant pathways of genotoxicity such as transcription, DNA damage, cell cycle, apoptosis and others. Additionally, this study showed that not all *in vitro* exposure durations result in correct class discrimination, and therefore the selection of the exposure duration plays a crucial role for *in vitro* toxicogenomics-based class prediction (82). Further, studies have been conducted for comparing the toxicogenomics results obtained from *in vivo* studies and the use of primary hepatocytes *in vitro*. One of these studies compared the gene expression profiles resulting after exposure of rats *in vivo* and primary rat hepatocytes *in vitro* to PPAR α agonists. Several differences were observed, including the down-regulation of several genes involved mainly in proliferation, apoptosis, immune response, transcription, cell adhesion, cellular morphogenesis and blood coagulation only after *in vivo* exposure. On the contrary, a lot of genes, mainly involved in metabolic processes, were up-regulated for both *in vivo* and *in vitro* exposures. These genes were further used as markers for identifying PPAR α agonists among a set of 32 chemicals for both *in vivo* and *in vitro* exposures. The obtained results from the *in vitro* exposures strongly represented those obtained *in vivo*, indicating that gene expression profiles *in vitro* may be used for the identification of the *in vivo* modes of action of chemicals (83).

Taking into consideration that transcriptomics responses *in vitro* represent those *in vivo*, the use of cell lines for predictive toxicogenomics has also been explored. A study using a rat hepatoma cell line showed that gene expression profiling can be used for the *in vitro* identification of rodent carcinogens, with an accuracy of approximately 90%. Most of the selected genes were involved in metabolism, translation and transcription, while some of the genes were involved in regulation of cell proliferation and signaling pathways that are known to be involved in carcinogenesis (84). The mouse fibroblast Balb/c 3T3 cell line was used for obtaining gene expression profiles after exposure to four carcinogens. From the four compounds used, only three shared several gene expression alterations of genes mainly involved in immune response, cell cycle, transcription, metabolism, cell adhesion, proliferation and apoptosis (85). The use of L5178Y Tk⁺ mouse lymphoma cells for the identification of gene signatures after exposure to three genotoxic agents resulted in only 19 genes affected by all three chemicals. For most of these genes, their function is still unknown. Furthermore, the expression of DNA damage and cell-cycle-related genes was not altered as expected after treatment of these cells to the genotoxic chemicals (86). On the contrary, another study using the same cell line and four genotoxic agents, two direct and two indirect, resulted in the selection of more than one thousand genes involved in a variety of biological processes. While the use of a cell line of rat origin seems promising for the identification of rodent hepatocarcinogens, the studies using cell lines of mice used only a small number of chemicals and therefore their application as alternative accurate method cannot be fully evaluated.

Taking into consideration the interspecies differences between humans and rodents, as described earlier, the use of human hepatic cell lines for predictive toxicogenomics ap-

proaches appears more promising (87). In particular, a predictive toxicogenomics approach for the discrimination of genotoxic from non-genotoxic carcinogens using data from HepG2 cells exposed to 22 compounds, resulted in an accuracy of higher than 80%, whereas the selected genes were mainly involved in apoptosis and cell death (87).

In combination these studies confirmed the hypothesis of specific gene expression patterns for different classes of carcinogens with different modes of actions. In addition, they show that predictive toxicogenomics approaches can be effectively applied using both *in vivo* and *in vitro* systems. However, for the *in vitro* systems, several factors have to be taken into consideration including the duration of exposure with the test compounds, as well as the selection of the appropriate and most relevant *in vitro* cellular system.

Human hepatic cell lines as *in vitro* systems for the application of toxicogenomic approaches

Given the fact that toxicogenomic approaches provide a promising alternative for the better prediction of the carcinogenic and genotoxic effects of chemicals, the most appropriate model for their application has to be identified that will comply with the current legislations for the reduction, refinement and replacement of animal experimentation (EU Directive 2010/63/EU) but will also be biologically relevant for human health.

Since many compounds need to be metabolized for their genotoxic or carcinogenic effects to take place, a cellular system that has the ability to metabolize these compounds should be used. The liver is the primary site where xenobiotic metabolism takes place, although other organs have some metabolic capacities. In addition, the liver is the most common target site for carcinogens in rodents, since more than 50% of the rodent carcinogens induce liver tumors (88). Therefore a, preferably *in vitro*, hepatic system with active biotransformation enzymes should be applied for toxicological studies. To achieve biological relevance to human health effects, a human *in vitro* hepatic system would be the best option.

Currently used *in vitro* liver systems can be divided in cultures of primary hepatocytes, precision-cut tissue slices or immortalized hepatic cell lines. Although human primary hepatocytes and precision-cut liver slices resemble the *in vivo* situation best, several limitations exist regarding their availability since they are primarily isolated from livers that are not suited for transplantation, and liver fragments from primary or secondary tumors (89). In addition, they exhibit high variability regarding the activity of liver-specific enzymes among individuals and their metabolic capabilities are lost with time (90, 91).

In contrast, from a practical point of view, hepatic cell lines, although they share fewer similarities to primary hepatocytes, are easier to handle and they retain their characteristics for longer periods. HepaRG and HepG2 cell lines are the most promising for toxicity studies among the currently available human hepatic cell lines. HepaRG cells, a recently developed cell line, have similar metabolic functions as the primary hepatocytes and have been successfully used for the evaluation of genotoxic properties of a limited number of chemicals (92, 93). HepG2 cells, on the other hand, have been frequently used in toxicological and

pharmacological studies. Although their limitation is the lower expression of drug metabolizing enzymes in comparison to primary hepatocytes, several studies show that HepG2 cells over-express these enzymes especially after exposure to toxins (94-99). In addition, when the basal gene expression profiles obtained from primary hepatocytes of three independent donors were compared to the gene expression profile from HepG2 cells, it was observed that the expression patterns of the different donors varied significantly, whereas the expression pattern of HepG2 cells had major similarities with those of all three donors (100). Furthermore, HepG2 have been proven a useful tool for genotoxicity studies, since genotoxic effects of chemicals that require metabolic activation can be detected by means of several assays (95, 101-103). Therefore, HepaRG and HepG2 cell lines appear to be the most suitable *in vitro* system for toxicological studies. In combination with toxicogenomic approaches, these cell lines may lead to the development of an improved *in vitro* method for the assessment of the *in vivo* genotoxic and carcinogenic properties of chemicals, especially since it is necessary to replace, reduce or refine testing on vertebrate animals (Directive 86/609/EEC). For instance, in the case of the REACH Regulation, its implementation should be based on the use of alternative test methods, wherever possible, and the use of animals should be avoided by recourse to validated or recognized alternative methods that provide the information required under the Regulation (EC 1907/2006).

Aim and outline of the thesis

In summary, the investigation of the genotoxic and carcinogenic potential of chemicals that humans are exposed to (e.g. cosmetics, pharmaceuticals and industrial chemicals) is of major importance for the evaluation of human cancer risk. Therefore several *in vitro* and *in vivo* methods have been developed for serving this purpose. However, the application of these assays encounters several limitations. The currently available *in vitro* methods, such as micronuclei, chromosomal aberrations and Ames mutagenicity assay, are limited in detecting only genotoxic compounds, but their major limitation is in their high false positive rate. The *in vivo* rodent assays for the evaluation of genotoxicity, on the other hand, are more accurate, since they evaluate the most relevant endpoint; however the use of experimental animals raises ethical issues and are very costly and time-consuming. Furthermore *in vivo* rodent assays are the only tool available for the evaluation of the carcinogenic potential of chemicals, namely the rodent bioassays, that in addition to the ethical issues raised by the use of experimental animals, these assays are also time consuming and have a limited relevance to human health, mainly due to the differences between rodents and humans in relation to the bioactivation of chemicals, the susceptibility to the chemicals and the activated molecular pathways after exposure to the chemicals. Due to these limitations, a more relevant, accurate and fast assay is required for the evaluation of the genotoxic and carcinogenic potential of chemicals in relation to human risk. Toxicogenomics approaches have been proven a valuable tool in the efforts for the development of alternative *in vitro* methods for genotoxicity and carcinogenicity prediction mainly since gene expression analysis can provide additional mechanistic information.

This thesis is based on the hypothesis that toxicogenomics approaches can be applied in *in vitro* human hepatic cell systems for genotoxicity and carcinogenicity assessment of chemicals. Therefore, the aim of this thesis is to firstly develop, by means of toxicogenomics analyses, alternative *in vitro* tests for the prediction of *in vivo* genotoxicity and carcinogenicity of chemicals that will overcome the limitations of the currently used test systems; and secondly, to elucidate effects and modes-of-actions of chemical carcinogens based on toxicogenomics approaches.

To this end, the most appropriate *in vitro* cell system should be selected. Taking into consideration that the liver is the organ where the metabolism of chemicals into reactive intermediates takes place which probably explains why the liver is the main target organ for carcinogenicity as has become apparent from the rodent cancer studies, an *in vitro* hepatic system should be used. Therefore, as described in **Chapter 2**, two human hepatic cell lines were compared by means of gene expression analyses to select the most promising *in vitro* cell system that can be used for the purpose of class prediction.

Based on the results of Chapter 2, in **Chapter 3** the HepG2 cell line was used to compare the effects of false positive genotoxic compounds to true genotoxins and non-genotoxins by means of phenotypic and transcriptomics analyses to identify the induced molecular pathways after exposure to the different compounds. This study further aimed in the selection of the most appropriate endpoint, as measured in HepG2 cells, which can be used for the better discrimination between true and false positive genotoxins.

In **Chapter 4**, several toxicogenomics-based approaches were explored in order to identify a set of genes that can be used for the development of a more accurate *in vitro* method for the evaluation of *in vivo* genotoxicity.

In **Chapter 5**, we investigated the biological role of one of the genes that was used in the prediction method described in Chapter 4. To this end, the selected gene was silenced using siRNA techniques and phenotypic and transcriptomics responses of HepG2 cells after exposure to genotoxic and non-genotoxic chemicals were again analyzed. The transcriptomics responses were further used for the classification of the chemicals as genotoxic or non-genotoxic by means of the method developed in Chapter 4, based on the hypothesis that the chemicals should be classified wrongly due to the silencing of an important gene for genotoxicity prediction.

Taking into consideration that genotoxic chemicals are carcinogenic, while non-genotoxic chemicals may pose a cancer risk or not, in the final study in **Chapter 6**, the aim was to evaluate the ability of transcriptomics responses in HepG2 cells to discriminate between non-genotoxic carcinogens and non-genotoxic non-carcinogens using available data from the previous studies. Taking into consideration the limitations of the rodent bioassays in regard to the evaluation of human health risk (i.e. different modes of action, metabolism etc), we aimed in the evaluation of the potential of known non-genotoxic rodent carcinogens to induce cancer in humans based on a transcriptomics comparison of HepG2 responses induced by known human non-genotoxic carcinogens, non-carcinogens and non-genotoxic rodent carcinogens.

This thesis ends with **Chapter 7**, where the work presented in the previous chapters is summarized and the results of this work and future research perspectives are discussed.

References

1. Jemal A, Siegel R, Xu J, & Ward E (2010) Cancer statistics, 2010. *CA Cancer J Clin* 60(5):277-300.
2. OECD (2010) Mortality from Cancer, in OECD, *Health at a Glance: Europe 2010* OECD Publishing:34-35.
3. Petrucelli N, Daly MB, & Feldman GL (2010) Hereditary breast and ovarian cancer due to mutations in BRCA1 and BRCA2. *Genet Med* 12(5):245-259.
4. Half E, Bercovich D, & Rozen P (2009) Familial adenomatous polyposis. *Orphanet J Rare Dis* 4:22.
5. Foulkes WD (2008) Inherited susceptibility to common cancers. *N Engl J Med* 359(20):2143-2153.
6. Bertram JS (2000) The molecular biology of cancer. *Mol Aspects Med* 21(6):167-223.
7. Hanahan D & Weinberg RA (2000) The hallmarks of cancer. *Cell* 100(1):57-70.
8. Oliveira PA, et al. (2007) Chemical carcinogenesis. *An Acad Bras Cienc* 79(4):593-616.
9. Cordon-Cardo C & Sheinfeld J (1997) Molecular and immunopathology studies of oncogenes and tumor-suppressor genes in bladder cancer. *World J Urol* 15(2):112-119.
10. Hanahan D & Weinberg RA (2011) Hallmarks of cancer: the next generation. *Cell* 144(5):646-674.
11. Preston RJ & Williams GM (2005) DNA-reactive carcinogens: mode of action and human cancer hazard. *Crit Rev Toxicol* 35(8-9):673-683.
12. Smith J, Tho LM, Xu N, & Gillespie DA (2010) The ATM-Chk2 and ATR-Chk1 pathways in DNA damage signaling and cancer. *Adv Cancer Res* 108:73-112.
13. Engelmann D & Putzer BM (2010) Translating DNA damage into cancer cell death-A roadmap for E2F1 apoptotic signalling and opportunities for new drug combinations to overcome chemoresistance. *Drug Resist Updat* 13(4-5):119-131.
14. Druckrey H (1973) Specific Carcinogenic and Teratogenic Effects of 'Indirect' Alkylating Methyl and Ethyl compounds, and their Dependency on Stages of Ontogenic Developments. *Xenobiotica* 3(5):271-303.
15. Hennings H, et al. (1993) Critical aspects of initiation, promotion, and progression in multi-stage epidermal carcinogenesis. *Proc Soc Exp Biol Med* 202(1):1-8.
16. Vincent TL & Gatenby RA (2008) An evolutionary model for initiation, promotion, and progression in carcinogenesis. *Int J Oncol* 32(4):729-737.
17. Ralph SJ, Rodriguez-Enriquez S, Neuzil J, Saavedra E, & Moreno-Sanchez R (2010) The causes of cancer revisited: "mitochondrial malignancy" and ROS-induced oncogenic transformation - why mitochondria are targets for cancer therapy. *Mol Aspects Med* 31(2):145-170.
18. Kobliakov VA (2010) Mechanisms of tumor promotion by reactive oxygen species. *Biochemistry (Mosc)* 75(6):675-685.
19. Hernandez LG, van Steeg H, Luijten M, & van Benthem J (2009) Mechanisms of non-genotoxic carcinogens and importance of a weight of evidence approach. *Mutat Res* 682(2-3):94-109.
20. ICH (1997) S1B TESTING FOR CARCINOGENICITY OF PHARMACEUTICALS.
21. ICH (2008) Guidance on genotoxicity testing and data interpretation for pharmaceuticals intended for human use S2(R1).
22. ICH (1995) GUIDELINE ON THE NEED FOR CARCINOGENICITY STUDIES OF PHARMACEUTICALS - S1A.
23. European Commission (2001) GUIDANCE ON SUBMISSIONS FOR FOOD ADDITIVE EVALUATIONS BY THE SCIENTIFIC COMMITTEE ON FOOD.
24. Ames BN, Lee FD, & Durston WE (1973) An improved bacterial test system for the detection and classification of mutagens and carcinogens. *Proc Natl Acad Sci U S A* 70(3):782-786.
25. OECD (1997) OECD GUIDELINE FOR THE TESTING OF CHEMICALS: Test No. 471: Bacterial Reverse Mutation Test.
26. Hozier J, Applegate M, & Moore MM (1992) In vitro mammalian mutagenesis as a model for genetic lesions in human cancer. *Mutat Res* 270(2):201-209.
27. Aaron CS & Stankowski LF, Jr. (1989) Comparison of the AS52/XPRT and the CHO/HPRT assays: evaluation of 6 drug candidates. *Mutat Res* 223(2):121-128.
28. OECD (1997) OECD GUIDELINE FOR THE TESTING OF CHEMICALS: Test No. 476: In vitro Mammalian Cell Gene Mutation Test.
29. Li AP, et al. (1991) An evaluation of the roles of mammalian cell mutation assays in the testing of chemical genotoxicity. *Regul Toxicol Pharmacol* 14(1):24-40.
30. Parry JM, Parry E, Phrakonkham P, & Corvi R (2010) Analysis of published data for top concentration considerations in mammalian cell genotoxicity testing. *Mutagenesis* 25(6):531-538.
31. Brambilla G & Martelli A (2004) Failure of the standard battery of short-term tests in detecting some rodent and human genotoxic carcinogens. *Toxicology* 196(1-2):1-19.

32. OECD (2010) OECD GUIDELINE FOR THE TESTING OF CHEMICALS: Test No. 487: In vitro mammalian cell micronucleus test.
33. OECD (1997) OECD GUIDELINE FOR THE TESTING OF CHEMICALS: Test No. 474: Mammalian Erythrocyte Micronucleus Test.
34. OECD (1997) OECD GUIDELINE FOR THE TESTING OF CHEMICALS: Test No. 473: In Vitro Mammalian Chromosome Aberration Test.
35. OECD (1997) OECD GUIDELINE FOR THE TESTING OF CHEMICALS: Test No. 475: Mammalian Bone Marrow Chromosome Aberration Test.
36. Terradas M, Martin M, Tusell L, & Genesca A (2010) Genetic activities in micronuclei: is the DNA entrapped in micronuclei lost for the cell? *Mutat Res* 705(1):60-67.
37. Natarajan AT, Kihlman BA, & Obe G (1980) Use of the 5-bromodeoxyuridine-labelling technique for exploring mechanisms involved in the formation of chromosomal aberrations. *Mutat Res* 73(2):307-317.
38. Natarajan AT & Obe G (1978) Molecular mechanisms involved in the production of chromosomal aberrations. I. Utilization of *Neurospora* endonuclease for the study of aberration production in G2 stage of the cell cycle. *Mutat Res* 52(1):137-149.
39. Galloway SM (2000) Cytotoxicity and chromosome aberrations in vitro: experience in industry and the case for an upper limit on toxicity in the aberration assay. *Environ Mol Mutagen* 35(3):191-201.
40. Hilliard C, et al. (2007) Chromosome aberrations in Chinese hamster and human cells: a comparison using compounds with various genotoxicity profiles. *Mutat Res* 616(1-2):103-118.
41. Working PK & Butterworth BE (1984) An assay to detect chemically induced DNA repair in rat spermatocytes. *Environ Mutagen* 6(3):273-286.
42. OECD (1986) OECD GUIDELINE FOR THE TESTING OF CHEMICALS: Test No. 482: Genetic Toxicology: DNA Damage and Repair, Unscheduled DNA Synthesis in Mammalian Cells in vitro.
43. OECD (1997) OECD GUIDELINE FOR THE TESTING OF CHEMICALS: Test No. 486: Unscheduled DNA Synthesis (UDS) Test with Mammalian Liver Cells in vivo.
44. Benigni R, Bossa C, Tcheremenskaia O, & Giuliani A (2010) Alternatives to the carcinogenicity bioassay: in silico methods, and the in vitro and in vivo mutagenicity assays. *Expert Opin Drug Metab Toxicol* 6(7):809-819.
45. Kirkland D, Reeve L, Gatehouse D, & Vanparys P (2011) A core in vitro genotoxicity battery comprising the Ames test plus the in vitro micronucleus test is sufficient to detect rodent carcinogens and in vivo genotoxins. *Mutat Res* 721(1):27-73.
46. Kirkland D, Aardema M, Henderson L, & Muller L (2005) Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens I. Sensitivity, specificity and relative predictivity. *Mutat Res* 584(1-2):1-256.
47. OECD (2009) OECD GUIDELINE FOR THE TESTING OF CHEMICALS: Test No. 451: Carcinogenicity Studies.
48. Sanner T, Dybing E, Willems MI, & Kroese ED (2001) A simple method for quantitative risk assessment of non-threshold carcinogens based on the dose descriptor T25. (Translated from eng) *Pharmacol Toxicol* 88(6):331-341 (in eng).
49. Ito N, Tamano S, & Shirai T (2003) A medium-term rat liver bioassay for rapid in vivo detection of carcinogenic potential of chemicals. *Cancer Sci* 94(1):3-8.
50. Bannasch P (1986) Preneoplastic lesions as end points in carcinogenicity testing. I. Hepatic preneoplasia. *Carcinogenesis* 7(5):689-695.
51. Pritchard JB, French JE, Davis BJ, & Haseman JK (2003) The role of transgenic mouse models in carcinogen identification. *Environ Health Perspect* 111(4):444-454.
52. van Steeg H, et al. (2001) DNA repair-deficient Xpa and Xpa/p53+/- knock-out mice: nature of the models. (Translated from eng) *Toxicol Pathol* 29 Suppl:109-116 (in eng).
53. Olivier M, Hollstein M, & Hainaut P (2010) TP53 mutations in human cancers: origins, consequences, and clinical use. *Cold Spring Harb Perspect Biol* 2(1):a001008.
54. Corominas M, et al. (1991) ras activation in human tumors and in animal model systems. *Environ Health Perspect* 93:19-25.
55. Jacobson-Kram D, Sistare FD, & Jacobs AC (2004) Use of transgenic mice in carcinogenicity hazard assessment. *Toxicol Pathol* 32 Suppl 1:49-52.
56. Wells MY & Williams ES (2009) The transgenic mouse assay as an alternative test method for regulatory carcinogenicity studies--implications for REACH. *Regul Toxicol Pharmacol* 53(2):150-155.
57. Huff J, Jacobson MF, & Davis DL (2008) The limits of two-year bioassay exposure regimens for identifying chemical carcinogens. *Environ Health Perspect* 116(11):1439-1442.

58. Ennever FK & Lave LB (2003) Implications of the lack of accuracy of the lifetime rodent bioassay for predicting human carcinogenicity. *Regul Toxicol Pharmacol* 38(1):52-57.
59. Hoenerhoff MJ, Hong HH, Ton TV, Lahousse SA, & Sills RC (2009) A review of the molecular mechanisms of chemically induced neoplasia in rat and mouse models in National Toxicology Program bioassays and their relevance to human cancer. *Toxicol Pathol* 37(7):835-848.
60. Hard GC & Khan KN (2004) A contemporary overview of chronic progressive nephropathy in the laboratory rat, and its significance for human risk assessment. *Toxicol Pathol* 32(2):171-180.
61. Thayer KA & Foster PM (2007) Workgroup report: National Toxicology Program workshop on Hormonally Induced Reproductive Tumors - Relevance of Rodent Bioassays. *Environ Health Perspect* 115(9):1351-1356.
62. Caldwell DJ (1999) Review of mononuclear cell leukemia in F-344 rat bioassays and its significance to human cancer risk: A case study using alkyl phthalates. *Regul Toxicol Pharmacol* 30(1):45-53.
63. Ito Y, et al. (2005) Species differences in the metabolism of di(2-ethylhexyl) phthalate (DEHP) in several organs of mice, rats, and marmosets. *Arch Toxicol* 79(3):147-154.
64. Aardema MJ, et al. (2010) International prevalidation studies of the EpiDerm 3D human reconstructed skin micronucleus (RSMN) assay: transferability and reproducibility. *Mutat Res* 701(2):123-131.
65. Witte I, Plappert U, de Wall H, & Hartmann A (2007) Genetic toxicity assessment: employing the best science for human safety evaluation part III: the comet assay as an alternative to in vitro clastogenicity tests for early drug candidate selection. *Toxicol Sci* 97(1):21-26.
66. Billinton N, et al. (2008) Interlaboratory assessment of the GreenScreen HC GADD45a-GFP genotoxicity screening assay: an enabling study for independent validation as an alternative method. *Mutat Res* 653(1-2):23-33.
67. Hastwell PW, et al. (2006) High-specificity and high-sensitivity genotoxicity assessment in a human cell line: validation of the GreenScreen HC GADD45a-GFP genotoxicity assay. *Mutat Res* 607(2):160-175.
68. Benfenati E, et al. (2009) Predictive models for carcinogenicity and mutagenicity: frameworks, state-of-the-art, and perspectives. *J Environ Sci Health C Environ Carcinog Ecotoxicol Rev* 27(2):57-90.
69. Mascolo MG, et al. (2010) BALB/c 3T3 cell transformation assay for the prediction of carcinogenic potential of chemicals and environmental mixtures. *Toxicol In Vitro* 24(4):1292-1300.
70. Matthews EJ, Spalding JW, & Tennant RW (1993) Transformation of BALB/c-3T3 cells: V. Transformation responses of 168 chemicals compared with mutagenicity in Salmonella and carcinogenicity in rodent bioassays. *Environ Health Perspect* 101 Suppl 2:347-482.
71. Vanparys P. CR, Aardema M., Gribaldo L., Hayashi M., Hoffmann S. and Schechtman L. (2011) ECVAM Prevalidation of Three Cell Transformation Assays. *ALTEX* 28(1):56-59.
72. Nuwaysir EF, Bittner M, Trent J, Barrett JC, & Afshari CA (1999) Microarrays and toxicology: the advent of toxicogenomics. *Mol Carcinog* 24(3):153-159.
73. Fielden MR & Zacharewski TR (2001) Challenges and limitations of gene expression profiling in mechanistic and predictive toxicology. *Toxicol Sci* 60(1):6-10.
74. Ellinger-Ziegelbauer H, Gmuender H, Bandenburg A, & Ahr HJ (2008) Prediction of a carcinogenic potential of rat hepatocarcinogens using toxicogenomics analysis of short-term in vivo studies. *Mutat Res* 637(1-2):23-39.
75. Ellinger-Ziegelbauer H, Stuart B, Wahle B, Bomann W, & Ahr HJ (2005) Comparison of the expression profiles induced by genotoxic and nongenotoxic carcinogens in rat liver. *Mutat Res* 575(1-2):61-84.
76. Hamadeh HK, et al. (2002) Gene expression analysis reveals chemical-specific profiles. *Toxicol Sci* 67(2):219-231.
77. Nakayama K, et al. (2006) Differences in gene expression profiles in the liver between carcinogenic and non-carcinogenic isomers of compounds given to rats in a 28-day repeat-dose toxicity study. *Toxicol Appl Pharmacol* 217(3):299-307.
78. Nie AY, et al. (2006) Predictive toxicogenomics approaches reveal underlying molecular mechanisms of nongenotoxic carcinogenicity. *Mol Carcinog* 45(12):914-933.
79. Uehara T, et al. (2008) A toxicogenomics approach for early assessment of potential nongenotoxic hepatocarcinogenicity of chemicals in rats. *Toxicology* 250(1):15-26.
80. Fielden MR, et al. (2008) Interlaboratory evaluation of genomic signatures for predicting carcinogenicity in the rat. (Translated from eng) *Toxicol Sci* 103(1):28-34 (in eng).
81. Yang Y, Abel SJ, Ciurlionis R, & Waring JF (2006) Development of a toxicogenomics in vitro assay for the efficient characterization of compounds. *Pharmacogenomics* 7(2):177-186.

82. Mathijs K, et al. (2009) Discrimination for genotoxic and nongenotoxic carcinogens by gene expression profiling in primary mouse hepatocytes improves with exposure time. *Toxicol Sci* 112(2):374-384.
83. Tamura K, Ono A, Miyagishima T, Nagao T, & Urushidani T (2006) Profiling of gene expression in rat liver and rat primary cultured hepatocytes treated with peroxisome proliferators. *J Toxicol Sci* 31(5):471-490.
84. Tsujimura K, et al. (2006) Prediction of carcinogenic potential by a toxicogenomic approach using rat hepatoma cells. (Translated from eng) *Cancer Sci* 97(10):1002-1010 (in eng).
85. Rohrbeck A, et al. (2010) Toxicogenomics applied to in vitro carcinogenicity testing with Balb/c 3T3 cells revealed a gene signature predictive of chemical carcinogens. *Toxicol Sci* 118(1):31-41.
86. Lee M, et al. (2003) cDNA microarray gene expression profiling of hydroxyurea, paclitaxel, and p-anisidine, genotoxic compounds with differing tumorigenicity results. *Environ Mol Mutagen* 42(2):91-97.
87. van Delft JH, et al. (2005) Comparison of supervised clustering methods to discriminate genotoxic from non-genotoxic carcinogens by gene expression profiling. *Mutat Res* 575(1-2):17-33.
88. Gold LS, Slone TH, Stern BR, & Bernstein L (1993) Comparison of target organs of carcinogenicity for mutagenic and non-mutagenic chemicals. *Mutat Res* 286(1):75-100.
89. Guguen-Guillouzo C & Guillouzo A (2010) General review on in vitro hepatocyte models and their applications. *Methods Mol Biol* 640:1-40.
90. Hewitt NJ, et al. (2007) Primary hepatocytes: current understanding of the regulation of metabolic enzymes and transporter proteins, and pharmaceutical practice for the use of hepatocytes in metabolism, enzyme induction, transporter, clearance, and hepatotoxicity studies. *Drug Metab Rev* 39(1):159-234.
91. Vermeir M, et al. (2005) Cell-based models to study hepatic drug metabolism and enzyme induction in humans. *Expert Opin Drug Metab Toxicol* 1(1):75-90.
92. Guillouzo A, et al. (2007) The human hepatoma HepaRG cells: a highly differentiated model for studies of liver metabolism and toxicity of xenobiotics. *Chem Biol Interact* 168(1):66-73.
93. Le Hégarat L, et al. (2010) Assessment of the genotoxic potential of indirect chemical mutagens in HepaRG cells by the comet and the cytokinesis-block micronucleus assays. *Mutagenesis* 25(6):555-560.
94. Guo L, et al. (2010) Similarities and Differences in the Expression of Drug Metabolizing Enzymes between Human Hepatic Cell Lines and Primary Human Hepatocytes. *Drug Metab Dispos*.
95. Knasmüller S, et al. (2004) Use of human-derived liver cell lines for the detection of environmental and dietary genotoxicants; current state of knowledge. *Toxicology* 198(1-3):315-328.
96. Dvorak Z, Vrzal R, Starha P, Klanicova A, & Travnicek Z (2010) Effects of dinuclear copper(II) complexes with 6-(benzylamino)purine derivatives on AhR and PXR dependent expression of cytochromes P450 CYP1A2 and CYP3A4 genes in primary cultures of human hepatocytes. *guk^olu^oQ^o?ψ^ordlr r8lr avd□□□*
97. Chen Y, Huang C, Zhou T, & Chen G (2008) Genistein induction of human sulfotransferases in HepG2 and Caco-2 cells. *Basic Clin Pharmacol Toxicol* 103(6):553-559.
98. Liguori MJ, Blomme EA, & Waring JF (2008) Trovafloxacin-induced gene expression changes in liver-derived in vitro systems: comparison of primary human hepatocytes to HepG2 cells. *Drug Metab Dispos* 36(2):223-233.
99. Lancon A, et al. (2007) Resveratrol in human hepatoma HepG2 cells: metabolism and inducibility of detoxifying enzymes. *Drug Metab Dispos* 35(5):699-703.
100. Harris AJ, Dial SL, & Casciano DA (2004) Comparison of basal gene expression profiles and effects of hepatocarcinogens on gene expression in cultured primary human hepatocytes and HepG2 cells. *Mutat Res* 549(1-2):79-99.
101. Natarajan AT & Darroudi F (1991) Use of human hepatoma cells for in vitro metabolic activation of chemical mutagens/carcinogens. *Mutagenesis* 6(5):399-403.
102. Winter HK, et al. (2008) Use of four new human-derived liver-cell lines for the detection of genotoxic compounds in the single-cell gel electrophoresis (SCGE) assay. *Mutation research* 657(2):133-139.
103. Majer BJ, et al. (2004) Genotoxic effects of dietary and lifestyle related carcinogens in human derived hepatoma (HepG2, Hep3B) cells. *Mutation research* 551(1-2):153-166.

Chapter 2

Comparison of HepG2 and HepaRG by whole-genome gene expression analysis for the purpose of chemical hazard identification

Toxicol Sci. 2010 May;115(1):66-79

D.G.J. Jennen, C. Magkoufopoulou,
H.B. Ketelslegers, H.M. van Herwijnen,
J.C.S. Kleinjans, J.H.M. van Delft

Abstract

Direct comparison of the hepatoma cell lines HepG2 and HepaRG has previously been performed by only evaluating a limited set of genes or proteins. In this study, we examined the whole-genome gene expression of both cell lines before and after exposure to the genotoxic (GTX) carcinogens aflatoxin B1 and benzo[a]pyrene and the nongenotoxic (NGTX) carcinogens cyclosporin A, 17-beta Estradiol, and 2,3,7,8-tetrachlorodibenzo-para-dioxin for 12h and 48h. Before exposure, this analysis revealed an extensive network of genes and pathways, which were regulated differentially for each cell line. The comparison of the basal gene expression between HepG2, HepaRG, primary human hepatocytes (PHH) and liver clearly showed that HepaRG resembles PHH and liver the most. After exposure to the GTX and NGTX carcinogens, for both cell lines, common pathways were found that are important in carcinogenesis, for example, cell cycle regulation and apoptosis. However, also clear differences between exposed HepG2 and HepaRG were observed and these are related to common metabolic processes, immune response and transcription processes. Furthermore, HepG2 performs better in discriminating between GTX and NGTX carcinogens. In conclusion, these results have shown that HepaRG is a more suited *in vitro* liver model for biological interpretations of the effects of exposure to chemicals, whereas HepG2 is a more promising *in vitro* liver model for classification studies using the toxicogenomics approach. Although, it should be noted that only five carcinogens were used in this study.

Introduction

The quest for toxicogenomics-based tests assessing the toxicity, for example the carcinogenicity, of chemicals as an alternative to current rodent bioassays has resulted in the investigation of multiple *in vitro* models. Hepatic models are frequently used, since the liver plays a major role in the metabolism of many compounds and also represents an important target organ in systemic toxicity (1). A preferred model is the human hepatoma cell line HepG2. The HepG2 cells are able to carry out biotransformation of xenobiotic compounds, are capable of activating mutagens and carcinogens and carry no p53 mutations, which enables the cells to activate DNA damage response, induce growth arrest and initiate apoptosis (2-4). This cell line has frequently been applied in toxicogenomics studies (5-9), as HepG2 cells are an easy to handle cell model compared to, for example, primary human hepatocytes (PHH). However, basal gene expression levels of phase I and II biotransformation enzymes in HepG2 are generally lower compared to PHH, which might result in an underestimation of toxicity for particular compounds (4,10-12). Furthermore, studies comparing HepG2 and PHH at the transcriptome level show a substantial number of basal gene expression differences (10, 13, 14), that is, transcripts showing a higher expression in HepG2 compared to PHH were involved in cell cycle regulation, DNA, RNA and nucleotide metabolism, transcription, transport and signal transduction, whereas lower transcription levels were associated with cell death, lipid metabolism and xenobiotic metabolism.

Recently, a novel human hepatoma cell line, HepaRG, was introduced. Under specific culture conditions, HepaRG cells can become highly differentiated and quiescent. In contrast to HepG2 cells, the differentiated HepaRG cells express various cytochrome P450 enzymes (CYP450s) (e.g., CYP1A2, CYP2B6, CYP2C9, CYP2E1, and CYP3A4) at a higher level and exhibit other metabolic functions (e.g., phase II enzymes and membrane transporters) normally found in liver, which make HepaRG more comparable to cultured PHH (15-17). In addition, induction as well as inhibition of several CYP450s were evident in HepaRG after exposure to prototypical CYP-selective inducers and inhibitors and were comparable to PHH (16, 17). Furthermore, a comparison at the transcriptome level of HepaRG and PHH exposed to phenobarbital showed a close resemblance between these two cellular models with regard to the genomic response (18). Although HepaRG and HepG2 have been compared for their gene expression levels and metabolic capacities, this was done on a limited set of genes and proteins.

Our study elucidates the correspondences and differences between these cell lines using whole-genome gene expression analysis, with the aim of selecting the most promising model for *in vitro* liver toxicogenomics studies. Whole-genome transcriptomics enables us to investigate a broad spectrum of activities within both cell lines. Therefore, in the present study, the genomic response of HepG2 and HepaRG cells before and after exposure to the well known genotoxic (GTX) carcinogens aflatoxin B1 (AFB1) and benzo[a]pyrene (BaP) and nongenotoxic (NGTX) carcinogens cyclosporin A (CsA), 17-beta Estradiol (E2), and 2,3,7,8-tetrachlorodibenzo-para-dioxin (TCDD) was investigated by microarray analysis. AFB1, a mycotoxin, and BaP, a ligand of the aryl hydrocarbon receptor (AHR), require meta-

bolic activation before being able to exert their GTX properties. Both are metabolized by CYP450s and their metabolites detoxified by several glutathione S-transferases (GSTs) (19-22). The NGTX carcinogens do not damage the DNA directly but can trigger numerous cellular processes, which may eventually lead to carcinogenesis. CsA is a widely used immunosuppressive drug, which has been shown to affect the steroid and lipid metabolism in liver cell lines including HepG2 (23-25). Also E2 is involved in the steroid metabolism as it is part of this metabolism and can be further metabolized by several CYP450s to estrogenically inactive metabolites (26). TCDD, also an AHR ligand resulting in the expression of various CYP450s, induces a wide spectrum of biological responses involved in the carcinogenic process (27).

In addition to the biological interpretation of the effect of the five carcinogens on the gene expression in HepG2 and HepaRG, these expression profiles were also used to compare both cell lines with respect to their capability of class discrimination. The results of this study substantiate the selection of the best model for genotoxicity testing in future experiments that will generate data required for adequate carcinogenic hazard identification and risk assessment.

Material and methods

Cell culture and treatment

In this study, 20 experimental groups were created by exposing the two cell lines (HepG2 and HepaRG) to five compounds (AFB1, BaP, CsA, E2, and TCDD) for 12h and 48h, at which we expected the GTX compounds to be partially and fully metabolized, respectively. These time points were selected based on results from an earlier time series experiment, wherein HepG2 cells were exposed to BaP for 3 – 60 h. In that study, the presence of BaP in the culture medium was measured, and this showed that after 30 – 36 h, 1% was remaining in the medium and that after 48h BaP was not detectable anymore.

HepG2 cell line

HepG2 cells were cultured in six-well plates in the presence of minimal essential medium supplemented with 1% nonessential amino acids, 1% sodium pyruvate, 2% penicillin/streptomycin, and 10% fetal bovine serum (all from Gibco BRL, Breda, The Netherlands). The cells were incubated at 37°C and 5% CO₂. When the cells were 80% confluent, the medium was replaced with fresh medium containing 1µM AFB1 (CAS no. 1162-65-8; Sigma-Aldrich, Zwijndrecht, The Netherlands), 2µM BaP (CAS no. 50-32-8; Sigma-Aldrich), 3µM CsA (CAS no. 59865-13-3; Sigma-Aldrich), 30µM E2 (CAS no. 50-28-2; Steraloids Inc., Newport), 10nM TCDD (CAS no. 1746-01-6; Cerilliant, Round Rock, TX), or 0.5% dimethyl sulfoxide (DMSO) (CAS no. 67-68-5; Sigma-Aldrich) as a vehicle control. Doses have been determined based on an 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) cytotoxicity assay with 80% viability for the highest dose (MTT test) in 96-well plates following a 72h treatment (Figure 1A). Treatment with the compounds was terminated after 12h or 48h by removing the culture medium and immediately adding TRIZOL (Gibco BRL). Three independent biological replicates were conducted.

HepaRG cells

Differentiated HepaRG cells were provided by Biopredic International (Rennes, France) in six-well plates or T25 flasks. Upon arrival, fresh low DMSO-containing medium (after shipment medium, MIL475, Biopredic) was added to the cells for recovery (2 ml per well for the six-well plates and 5 ml per T25 flask). Cells were kept in this medium for 7 days. For maintenance and use of the cells, 2 ml (per well of a six-well plate) or 5 ml (per T25 flask) DMSO-containing medium (MIL375, Biopredic) was added to the cells; this upregulates hepatocyte-specific functions to higher levels for maximal metabolic capacity of HepaRG cells. Cells were cultured at 37°C in a 95/5% air/CO₂ incubator. For chemical treatment, medium was replaced with DMSO-containing medium (MIL375, Biopredic) supplemented with 3µM AFB1, 10µM BaP, 300µM CsA, 30µM E2, 10nM TCDD, or 0.5% DMSO (vehicle control). Doses have been determined by MTT test as described for HepG2 (Figure 1B). Treatment with the compounds was terminated after 12h or 48h by removing the culture medium and immediately adding TRIZOL (Gibco, BRL). Five biological replicates were performed using three six-well plates and two T25 flasks.

Figure 1: Viability plots of (A) HepG2 and (B) HepaRG cells exposed to the test compounds for 72h expressed as percentage of that for the negative (vehicle) controls as determined by the MTT method. The plots are in order of the compounds: AFB1, BaP, CsA, E2 and TCDD. The average viability of two assays for each compound per cell line is shown and SD bars are indicated.

Total RNA isolation

Total RNA was extracted using 0.5 ml TRIZOL according to the manufacturer's instructions. RNeasy Mini Kits (Qiagen, Westburg bv, The Netherlands) were used to purify total RNA from salts and residual DNA. From each sample, RNA quantity was measured by means of a spectrophotometer and integrity was determined using the Bioanalyzer (Agilent Technologies, Amstelveen, The Netherlands). Extracted RNA was stored at -80 °C until it was used as template for complementary DNA (cDNA) synthesis. The quantity of the RNA extracted from the first three experimental samples with HepaRG cells on six-well plates appeared too low for further analysis, and therefore these RNA samples were pooled.

Microarray hybridization

Target preparation

cDNA was prepared using the Affymetrix One-Cycle cDNA synthesis kit (Affymetrix, Santa Clara). cDNA synthesis and labeling was performed according to the manufacturer's procedures. Subsequent labeling of the samples was conducted by synthesis of Biotin-labeled complementary RNA (cRNA) using the GeneChip IVT labeling kit (Affymetrix). Purified cRNA was quantified using a spectrophotometer, and unfragmented samples were checked on the Bioanalyzer. Subsequently, cRNA samples were fragmented for target preparation according to the Affymetrix manual and checked on the Bioanalyzer. Samples were stored at -20 °C until ready to perform hybridization.

Hybridization

cRNA targets were hybridized on high-density oligonucleotide gene chips (Affymetrix Human Genome U133 Plus 2.0 GeneChip arrays) according to the Affymetrix Eukaryotic Target Hybridization manual. The gene chips were washed and stained using the Affymetrix Fluidics Station 450/250 and Genechip Operating Software and scanned by means of an Affymetrix GeneArray scanner. Normalization quality controls, including scaling factors, average intensities, present calls, background intensities, noise, and raw Q values, were within acceptable limits for all chips except two, which contained samples from HepaRG exposed to BaP for 12h and from HepaRG exposed to E2 for 48h. These were therefore excluded for further analysis. Hybridization controls BioB, BioC, BioD, and CreX were called present on all chips and yielded the expected increases in intensities.

Microarray data analyses

Re-annotation, normalization, and data filtering

Seventy data sets were obtained, which were re-annotated to the Molecular and Behavioural Neuroscience Institute, University of Michigan, Custom CDF-files (http://brainarray.mbni.med.umich.edu/Brainarray/Database/CustomCDF/genomic_curated_CDF.asp) (28) and Robust Multi-array Average (RMA) normalized (29) using the NuGOExpressionFileCreator, an enhanced version of the standard ExpressionFileCreator module that is present in GenePattern (30). The resulting 17788 probe sets represent 17726 unique genes and 62 internal controls.

Next to re-annotation and RMA normalization, presence and absence calls were computed using the NuGOExpressionFileCreator. Each gene was analyzed for the number of absence calls in the three replicates per treatment. Genes that contained two or more absence calls within the three replicates for all the treatments as well as vehicle controls were omitted from the data. This resulted in filtered data sets per time point consisting of 12535 unique genes at 12h and 12624 unique genes at 48h, of which approximately 97% were in common. These two filtered data sets were used for further analyses. The data in this publication have been deposited in EBI's ArrayExpress (<http://www.ebi.ac.uk/microarray-as/ae/>) and are accessible through ArrayExpress accession number E-MEXP-2458.

Principal component analysis

Principal component analysis (PCA) was applied to identify data patterns and to highlight data similarity and differences between the treated and untreated cell lines at different time points. Therefore, the log₂-transformed intensities of the filtered data sets were uploaded into the PCA module of ArrayTrack (31, 32).

Basal gene expression analysis

The basal expression profiles of DMSO-treated HepG2 cells and HepaRG cells were examined at 12h and 48h by calculating the log₂ differences between the HepG2 and HepaRG cell lines based on the average log₂ data of the three replicates for each cell line. Genes with a log₂ difference of < -0.58 or > 0.58 were considered differentially expressed between the two cell lines.

The basal gene expression levels of the HepG2 and HepaRG cell lines were further compared with basal gene expression data from PHH and human liver (collected postmortem) obtained from the NCBI GEO database (<http://www.ncbi.nlm.nih.gov/geo/>). Expression data for PHH were collected from two experiments, both using cells from four donors (accession numbers GSE10410 and GSE11942), and expression data for liver were collected from one experiment with samples from five donors (accession number GSE13471). These experiments used the same microarray platform as described in this study, therefore, enabling the same analytical procedure. For the comparison, hierarchical clustering analysis (HCA) in GenePattern (Euclidean correlation and pairwise average linkage) (33) was performed.

Selection of Differentially Expressed Genes

The intensities of the filtered data sets were log₂ transformed, and subsequently, log₂ differences of treated versus vehicle controls were calculated. Differentially expressed genes (DEGs) for each experimental group were selected using the following criteria: (1) a log₂ difference of < -0.58 or > 0.58 (i.e., absolute fold change of 1.5) for the average of the three replicates within the experimental group, (2) same direction of the log ratio for all these replicates, and (3) a log₂ difference of < -0.26 or > 0.26 (i.e., absolute fold change of 1.2) for at least two replicates.

Pathway analyses

In order to increase specificity of the transcriptomics changes, DEGs were uploaded onto MetaCore (GeneGo, San Diego, CA) for identifying the involvement of these genes in

specific cellular pathway maps by overrepresentation analyses compared to the total amount of genes involved in the particular maps. In the analysis, the filtered data sets were used as background list, and thereafter, pathway maps with a p-value < 0.05 were considered significantly modulated. Gene Ontology (GO) processes for individual as well as groups of genes were also determined using MetaCore.

Class discrimination

Prediction Analysis for Microarrays 2.13 (PAM, Stanford University Software) analysis was conducted for each cell line per time point to select genes that can discriminate between GTX and NGTX compounds. PAM uses gene expression data to identify a subset of genes that best characterize each class by using the method of “nearest shrunken centroids” (34). As input for the PAM analysis, a combined gene list based on the DEGs of the treatments with BaP, AFB1, CsA and TCDD for each time point and each cell line was used. The gene expression data of these treatments were used as a training set and the data of E2 were included in the test set. Following a 10-fold cross-validation, an average misclassification error was calculated for each threshold. An optimum threshold was selected based on the following criteria: (1) lowest misclassification error, (2) highest predictability, (3) highest average probability, which was calculated as the average of the probabilities of each treatment to be discriminated to its true class and (4) lowest number of discriminators. This threshold was used for the prediction of the test set, which included all the treatments.

Results

In total, 70 raw data sets were retrieved, which were re-annotated, normalized and filtered for presence and absence calls, resulting in two filtered data sets per time point consisting of 12535 unique genes at 12h and 12624 unique genes at 48h. Between these two data sets, an overlap of 12162 genes (approximately 97%) was found. The intensities of the filtered data sets were log2 transformed and were used in PCA to visualize cell line and time-dependent effects on gene expression before and after exposure to the compounds (Figure 2). Only the log2 data of the overlapping genes of the filtered data sets were used in this analysis. The cell lines were separated by principal component (PC) #1, containing most of the variance (76.93%). PC #2 (5.84%) and PC #3 (5.08%) separated the different time points showing a larger exposure time effect for the HepG2 cell line compared to the HepaRG cell line. Within each experimental group, little variance is found for the exposed HepG2 cells, whereas the exposed HepaRG cells show more variance. Furthermore, in both cell lines, the samples from cells exposed to CsA were located separately from the other samples.

The PCA analysis clearly indicates differences between the two cell lines, which were further analyzed based on the basal gene expression level, on the DEGs' level and on the discrimination capacity for the two cell lines.

Expression profiles of unexposed HepG2 versus HepaRG

The basal expression profiles of DMSO-treated HepG2 cells and HepaRG cells were ex-

Figure 2: PCA of the log₂ - transformed intensities for the whole filtered data set of 12162 genes before and after exposure of HepG2 and HepaRG to DMSO (solvent), AFB1, BaP, CsA, E2 and TCDD.

amined to establish the overall basal concordances and differences between the two cell lines. Therefore, average log₂ differences between the HepG2 and HepaRG cell lines were calculated and DEGs selected, resulting in 6061 genes at 12h and 6132 genes at 48h. Of these DEGs, 3270 were over- and 2791 were underexpressed in HepG2 relative to HepaRG at 12h and 3202 were over- and 2930 were underexpressed in HepG2 relative to HepaRG at 48h. The expression level of most CYP450 genes was lower in HepG2 cells relative to HepaRG cells regardless of time. Also other phase I genes (e.g., coding for alcohol dehydrogenases, aldehyde dehydrogenases and flavin-containing monooxygenases) as well as several transcription factors involved in regulation of xenobiotic response (e.g., pregnane X receptor, constitutive androstane receptor, AHR and farnesoid X receptor) showed a lower expression level in HepG2 relative to HepaRG. With respect to the phase II genes, the number of overexpressed genes did not differ from the number of underexpressed genes within the different groups of phase II genes for HepG2 relative to HepaRG. A complete list of DEGs between DMSO-treated HepG2 and HepaRG is included as Supplementary table 1.

Pathway analysis using MetaCore (Supplementary table 2) showed for the differentially overexpressed genes at both time points significantly modulated processes (p -value < 0.05) associated with cell cycle regulation and checkpoint control, regulation of transcription, DNA,

RNA and nucleotide metabolism in the HepG2 cell line compared to the HepaRG cell line, whereas for the differentially underexpressed genes at both time points, significantly modulated processes (p -value < 0.05) were associated with common metabolic pathways, mainly amino acid, lipid, and steroid metabolism.

In addition to the direct comparison of the basal gene expression between the HepG2 and HepaRG cell lines, these expression data were also compared with basal gene expression data from PHH and human liver (collected postmortem). HCA was performed with the obtained data in order to show the relationship between the different liver models (Figure 3). A close relationship between HepaRG and PHH and to some lesser extent between HepaRG and liver was found for phase I and II genes (Figure 3A) compared to HepG2. Regarding the whole filtered data set (Figure 3B), HepG2, HepaRG and PHH share the same relationship with liver. Still, HepaRG and PHH are more closely related compared to HepG2. Genes involved in DNA, RNA, and protein metabolic processes, cell cycle-related processes, DNA repair and regulation of apoptosis showed lower expression in liver compared to PHH, HepaRG and HepG2, whereas genes involved in response to stimulus, immune response, cell communication and common metabolic processes showed higher gene expression. From these processes, HepG2 showed lower gene expression compared to PHH and HepaRG for response to stimulus, common metabolic processes and regulation of apoptosis, but higher gene expression was found for DNA, RNA and protein metabolic processes and cell cycle-related processes.

Figure 3: HCA (Euclidean correlation and pairwise average linkage) of the log2-transformed intensities for the basal gene expression in liver, PHH, HepG2 and HepaRG based on (A) 179 phase I and II genes and (B) the whole filtered data set of 12162 genes.

Gene expression changes in GTX- and NGTX-treated HepG2 and HepaRG

Gene expression changes, either up- or downregulated, were lower in HepG2 cells exposed to AFB1, CsA and E2 compared to the HepaRG cell line at both time points, whereas the opposite is observed for gene expression changes in BaP and TCDD-exposed HepG2. The DEGs of the experimental groups show a similar pattern as the overall gene expression (Table 1). Average log2 differences of the complete gene list, with the DEGs highlighted, of the exposed HepG2 and HepaRG cells at 12h and 48h are included in Supplementary table 3.

HCA (Pearson correlation and pairwise complete linkage) of combined gene lists containing expression data based on DEGs from at least one experimental group after 12h (3864 genes) and 48h (4721 genes) exposure were used to compare the exposed cell lines. In Figure 4, GO processes, determined by MetaCore, are indicated for the main clusters, giving

Table 1: Total, unique and common number of differentially expressed genes (DEGs) per treatment, time point and cell line

Compound Exposure time	AFB1		BaP		CsA		E2		TCDD	
	12h	48h	12h	48h	12h	48h	12h	48h	12h	48h
HepG2										
total no. of DEGs	30	239	808	483	314	415	37	63	342	1028
up-regulated	21	162	300	304	195	148	9	26	222	598
down-regulated	9	77	508	179	119	267	28	37	120	430
total no. of unique DEGs ^a	14	145	737	409	87	238	28	60	318	929
up-regulated	11	92	262	238	35	60	6	26	202	527
down-regulated	3	53	475	171	52	178	22	34	116	402
HepaRG										
total no. of DEGs	413	1225	268	347	2951	3468	411	113	108	369
up-regulated	186	636	153	166	1338	1843	193	33	69	149
down-regulated	227	589	115	181	1613	1625	218	80	39	220
total no. of unique DEGs ^a	397	1131	197	273	2724	3291	402	110	84	270
up-regulated	176	570	113	107	1184	1760	189	33	48	90
down-regulated	221	561	84	166	1540	1531	213	77	36	180
HepG2 & HepaRG										
total no. of common DEGs ^b	16	94	71	74	227	117	9	3	24	99

^aUnique DEGs per treatment and time-point are only expressed in one cell line and not in the other cell line.^bCommon DEGs per treatment and timepoint are expressed in both cell lines.

an overall impression of the similarities and differences between exposed HepG2 and HepaRG. At the 48h time point, seemingly the biggest differences between the cell lines were observed, but this is mainly influenced by the large amount of DEGs from the CsA-exposed HepaRG and the TCDD-exposed HepG2 (Table 1).

Therefore, in a more in-depth analysis, DEGs in common between the HepG2 and HepaRG as well as unique DEGs for each cell line (Table 1) were examined per compound at 12h and 48h. For these DEGs, GO processes were determined by MetaCore. The main GO processes for the common and unique genes are listed per compound in Table 2. The GTX compounds show common genes between HepG2 and HepaRG involved in apoptosis and the NGTX mainly in common metabolic processes (e.g., amino acid, carbohydrate, lipid and steroid metabolic processes). The processes that the unique DEGs are involved in differ for HepG2 exposed to each compound. The GTX- and NGTX-exposed HepaRG, on the other hand, have unique DEGs involved in the same processes for each class, especially after 12h of exposure, despite the limited overlap of these unique genes within each class.

Furthermore, MetaCore was also used to determine the pathway maps for all DEGs from each experimental group at 12h and 48h. Cellular processes with five or more significantly modulated (p -value < 0.05) pathway maps are listed in Table 3. A complete list of significantly modulated maps can be found as Supplementary data 4 and 5 for the 12h and 48h time

point, respectively. The number of maps in common between HepG2 and HepaRG are limited and mainly involved in transcription and common metabolic processes (Table 3).

Figure 4: HCA (Pearson correlation and pairwise complete linkage) of gene lists containing DEGs from at least one experimental group after (A) 12h (3864 genes) and (B) 48h (4721 genes) exposure. For the main clusters, the following GO processes are indicated: (1) Transport-related processes; (2) Cellular response to stimulus; (3) Common metabolic processes; (3a) Amino acid related; (3b) Carbohydrate related; (3c) Lipid related; (3d) Steroid related; (4) Cell cycle-related processes, DNA replication, and DNA repair; (5) Cell development, proliferation, and differentiation; (6) Protein modification processes; (7) Communication; (8) Apoptosis; (9) Protein metabolic processes and translation-related processes; (10) DNA metabolic processes; (11) RNA metabolic processes and translation related processes; (12) Secretion.

Table 2: GO-processes from MetaCore analysis of unique and common differentially expressed genes (DEGs) of the exposed HepG2 and HepaRG at 12h and 48h

Aflatoxin B1		
Exposure time	DEGs	GO-processes
12h	Unique DEGs of HepG2	cell cycle regulation; cell cycle arrest; apoptosis; transcription regulation; cell adhesion; lipid metabolism; immune response; response to drug; ion binding; transport
	Unique DEGs of HepaRG	cellular response to stimulus; immune response; G-protein coupled receptor protein signaling; lipid and steroid metabolic processes; exocytosis; transport
	Common	cell cycle regulation; apoptosis related processes
48h	Unique DEGs of HepG2	cellular developmental process; regulation of cellular metabolic processes; regulation of hormone secretion; riboflavin metabolism; cellular component assembly; chemical homeostasis
	Unique DEGs of HepaRG	cell cycle regulation; cellular response to stimulus; immune response; G-protein coupled receptor protein signaling; common metabolic processes ^a
	Common	common metabolic processes ^a ; apoptosis related processes

Benzo[a]pyrene		
Exposure time	DEGs	GO-processes
12h	Unique DEGs of HepG2	regulation of apoptosis; regulation of cellular processes; cellular response to stimulus; G-protein coupled receptor protein signaling; immune response; cell adhesion
	Unique DEGs of HepaRG	cellular response to stimulus; immune response; G-protein coupled receptor protein signaling; lipid and steroid metabolic processes; exocytosis; transport
	Common	immune and stress response processes
48h	Unique DEGs of HepG2	cellular response to stimulus; apoptosis; regulation of cell proliferation; lipid and steroid metabolic processes
	Unique DEGs of HepaRG	cell cycle related processes; common metabolic processes ^a
	Common	apoptosis and DNA damage related processes; cellular developmental processes; common metabolic processes ^a
Cyclosporine A		
Exposure time	DEGs	GO-processes
12h	Unique DEGs of HepG2	cellular response to stimulus; regulation of kinase activity; negative regulation of nucleotide related processes; negative regulation of growth; cell cycle regulation
	Unique DEGs of HepaRG	immune response; cellular response to stimulus; apoptosis; cystic fibrosis; G-protein coupled receptor protein signaling; common metabolic processes ^a
	Common	cell cycle regulation; common metabolic processes ^a
48h	Unique DEGs of HepG2	cellular component assembly; cellular developmental process; cellular response to stimulus; cell cycle related processes
	Unique DEGs of HepaRG	apoptosis, cystic fibrosis; cellular metabolic processes; common metabolic processes ^a
	Common	cell cycle regulation; common metabolic processes ^a
17-beta Estradiol		
Exposure time	DEGs	GO-processes
12h	Unique DEGs of HepG2	E2 metabolism; anti-apoptosis; metal ion binding transport; negative regulation of cellular processes
	Unique DEGs of HepaRG	apoptosis; cystic fibrosis; cellular metabolic processes; common metabolic processes ^a
	Common	metal ion binding; steroid metabolic processes; response to stimulus; transcription; transport; cell division; chromosome partitioning
48h	Unique DEGs of HepG2	cellular component assembly; chemical homeostasis; response to hormone stimulus
	Unique DEGs of HepaRG	cell cycle regulation; common metabolic processes ^a
	Common	amino acid and steroid metabolic processes; cell growth regulation
2,3,7,8-tetrachlorodibenzo-para-dioxin		
Exposure time	DEGs	GO-processes
12h	Unique DEGs of HepG2	G-protein coupled receptor protein signaling; regulation of growth; regulation of cellular processes; cellular response to stimulus; steroid metabolic processes
	Unique DEGs of HepaRG	G-protein coupled receptor protein signaling; response to hormone stimulus; cystic fibrosis
	Common	steroid metabolic processes; G-protein coupled receptor protein signaling
48h	Unique DEGs of HepG2	cellular response to stimulus; immune response; lipid metabolic processes; amino acid and RNA metabolic processes.
	Unique DEGs of HepaRG	G-protein coupled receptor protein signaling; response to hormone stimulus; common metabolic pathways ^a
	Common	common metabolic processes ^a ; cellular response to stimulus

^aincludes amino acid, lipid and lipoprotein, vitamin and cofactor, carbohydrates, steroid and nucleotide metabolism.

Table 3: Cellular processes and pathways from MetaCore analysis of differentially expressed genes of the exposed HepG2 and HepaRG at 12h and 48h

Aflatoxin B1			
Exposure time	Cell line	total no. of pathways	Cellular process ^a (number of pathways)
12h	HepG2	2	Apoptosis Transcription
	HepaRG	71	G-protein coupled receptor protein signaling (21) Immune response (18) Common metabolic pathways ^b (11)
	Common	1	Transcription
48h	HepG2	18	Common metabolic pathways ^b (12)
	HepaRG	64	Common metabolic pathways ^b (44) Cell cycle (7)
	Common	9	Common metabolic pathways ^b (8) Transcription (1)
Benzo[a]pyrene			
Exposure time	Cell line	total no. of pathways	Cellular process ^a (number of pathways)
12h	HepG2	75	G-protein coupled receptor protein signaling (27) Common metabolic pathways ^b (14) Cell adhesion (7) Immune response (7) Transcription (5)
	HepaRG	82	Immune response (18) Common metabolic pathways ^b (14) G-protein coupled receptor protein signaling (13) Intracellular receptor-mediated signaling pathway (13) Transcription (6)
	Common	9	G-protein coupled receptor protein signaling (3) Transcription (3) Common metabolic pathways ^b (2) Response to stimulus (1)
48h	HepG2	24	Common metabolic pathways ^b (9) G-protein coupled receptor protein signaling (6)
	HepaRG	31	Common metabolic pathways ^b (23)
	Common	1	Common metabolic pathways ^b
Cyclosporine A			
Exposure time	Cell line	total no. of pathways	Cellular process ^a (number of pathways)
12h	HepG2	27	Common metabolic pathways ^b (14) Cell cycle (8)
	HepaRG	56	Common metabolic pathways ^b (40) Cell cycle (7)
	Common	11	Common metabolic pathways ^b (5) Cell cycle (4) Cystic fibrosis (2)
48h	HepG2	31	Common metabolic pathways ^b (15) Cell cycle (14)
	HepaRG	70	Common metabolic pathways ^b (55)
	Common	6	Common metabolic pathways ^b (6)
17-beta Estradiol			
Exposure time	Cell line	total no. of pathways	Cellular process ^a (number of pathways)
12h	HepG2	2	Apoptosis (2)
	HepaRG	50	Common metabolic pathways ^b (27) Immune response (5) Cystic fibrosis (5) Apoptosis (5)
	Common	0	-
48h	HepG2	41	Common metabolic pathways ^b (14) G-protein coupled receptor protein signaling (5)
	HepaRG	27	Common metabolic pathways ^b (21)
	Common	9	Common metabolic pathways ^b (8) Cell cycle (1)

2,3,7,8-tetrachlorodibenzo-para-dioxin			
Exposure time	Cell line	total no. of pathways	Cellular process ^a (number of pathways)
12h	HepG2	55	Common metabolic pathways ^b (20) G-protein coupled receptor protein signaling (19) Cell adhesion (6)
	HepaRG	17	Common metabolic pathways ^b (12)
	Common	4	Common metabolic pathways ^b (2) G-protein coupled receptor protein signaling (1) Response to stimulus (1)
48h	HepG2	53	Common metabolic pathways ^b (43) Transcription (5)
	HepaRG	56	Common metabolic pathways ^b (35) Transcription (8)
	Common	20	Common metabolic pathways ^b (18) Transcription (2)

^aonly cellular processes with 5 or more pathways are listed, when applicable.

^bincludes amino acid, lipid and lipoprotein, vitamin and cofactor, carbohydrates, steroid and nucleotide metabolism.

Class discrimination

The possibility to use the two cell lines to discriminate between GTX and NGTX compounds was evaluated by PAM analysis. The analysis was performed using gene lists consisting of the union of DEGs from the AFB1-, BaP-, CsA- and TCDD-treated cell line per time point. This resulted in 1247 and 1613 DEGs for HepG2 at 12h and 48h, respectively, and in 3225 and 3977 DEGs for HepaRG at 12h and 48h, respectively. These DEGs were used for the identification of the discriminators by selecting an optimum threshold from an average misclassification error rate curve based on a 10-fold cross-validation (Figure 5). For HepG2 cells treated for 12h and 48h the optimum threshold resulted in 54 and 18 discriminators, respectively. When these discriminators were used for testing the different treatments, the prediction was 100% correct, with an average probability of 100% (Figure 6 A-B). For HepaRG cells treated for 12h, the selected threshold led to the identification of 27 discriminators that resulted in an 100% correct predictability with an average probability of 89%, whereas for HepaRG cells treated for 48h, 35 discriminators were identified using the optimum threshold, which resulted in an 86% correct predictability with an average probability of 100% (Figure 6 C-D). In the latter discrimination, E2 was placed in the wrong class together with AFB1 and BaP.

MetaCore analysis of the discriminators revealed mostly different GO processes per cell line and time point. The discriminators for HepG2 cells treated for 12h were involved in response to stimulus, apoptosis, cell cycle and protein modification processes and for the 48h treatment also response to stimulus was found as well as lipid metabolic processes. The discriminators for HepaRG had more processes in common. Both exposure times had the GO process cellular homeostasis and cell proliferation in common, whereas cell communication was found for 12h and response to stimulus, cell differentiation and development for 48h.

Figure 5: Cross-Validation (CV) plots for PAM analysis of (A) HepG2 at 12h, (B) HepG2 at 48h, (C) HepaRG at 12h, and (D) HepaRG at 48h. Misclassification error for each threshold is the average of a 10-fold CV. Numbers above each threshold represent the number of discriminators at that given threshold. Arrows indicate the optimal threshold.

Figure 6: Test probability plots for PAM analysis of (A) HepG2 at 12h, (B) HepG2 at 48h, (C) HepaRG at 12h, and (D) HepaRG at 48h. PAM probability results for the optimal threshold indicated in Figure 5 are shown. Black squares and white squares represent the probability of the treatments to be classified as GTX and NGTX, respectively.

Discussion

In this study, we examined the whole-genome gene expression of HepG2 and HepaRG before and after exposure to GTX and NGTX carcinogens for 12h and 48h. Correspondences and differences between HepG2 and HepaRG were evaluated as well as the usage of both cell lines to discriminate between the two classes, for selecting the most promising model for *in vitro* liver toxicogenomics studies.

Approximately 6100 differences in basal gene expression levels were evident between the two cell lines at 12h and 48h, which may explain the phenotypic differences between HepG2 and HepaRG. In HepG2, a dividing cell line, many genes associated with cell cycle and checkpoint control were overexpressed, whereas genes involved in common metabolic processes were underexpressed in comparison to HepaRG. Similar results were found in a comparison between HepG2 and PHH (10, 13, 14).

Furthermore, most CYP450s and several other phase I genes were underexpressed in HepG2 compared to HepaRG as previously observed (15, 16). For the phase II gene groups, an equal number of genes were over- and underexpressed. So far, a limited number of phase II genes have been investigated for their expression in HepG2 and HepaRG. Kanebratt and Andersson (16) showed that GSTA1 and sulfotransferase (SULT2A1) were underexpressed in HepG2 compared to HepaRG, which was also observed in our study. Other studies compared the gene expression of phase II genes in HepG2 to PHH. For example, Wilkening *et al* (4) showed underexpression of epoxide hydrolase (EPH)X1 and UDP glucuronosyltransferase (UGT) 1A6 and overexpression of N-acetyltransferase (NAT1), NAD(P)H:quinone oxidoreductase (NQO1), and SULT1A1 in HepG2 compared to PHH. In our study, EPHX1, UGT1A6, and NAT1 were underexpressed and NQO1 and SULT1A1 overexpressed in HepG2 compared to HepaRG. These findings imply that HepaRG and PHH are more alike. However, this implication is based on an indirect comparison on the transcriptome level or on a direct comparison using a limited set of genes. Therefore, we directly compared HepG2, HepaRG, PHH and liver on whole-genome gene expression level using HCA. This analysis confirmed that HepaRG and PHH are closely related than HepG2 is to either one. Furthermore, both cell lines and PHH share the same relationship with liver. However, regarding the phase I and II genes, liver is more closely related to HepaRG and PHH as to HepG2, which corresponds with the observed lower expression of these genes in HepG2 (15, 16).

The whole-genome gene expression and pathway analyses of the exposed cell lines resulted in several common but mostly unique DEGs and pathways for the different experimental groups. The comparison between the AFB1- and BaP-exposed cell lines revealed common DEGs involved in processes related to, for example, cell cycle regulation, apoptosis, immune and stress response, and common metabolic processes. These findings were supported by MetaCore analysis, revealing common pathways involved in G protein-coupled receptors (GPCR) signaling, apoptosis, cell cycle, DNA repair, and immune response in both cell lines. Most of these pathways are the direct result of the exposure to the compounds. For example, GPCRs are involved in the regulation of cellular proliferation, specif-

ically in the progression from G1- to S-phase of the cell cycle (35). MetaCore analysis showed that GPCR signaling and cell cycle regulation takes place at 12h, but not at 48h. This corresponds well with the observed cell cycle arrest at the S-phase after 24h until 48h of BaP-exposed HepG2 cells (36). Before 24h, the cells are able to progress from G1- to S-phase, and after 48h, the cells can enter G2/M-phase.

Common DEGs of the CsA-, E2 and TCDD exposed cell lines also showed similar responses in HepG2 and HepaRG that were supported by the pathway analysis. Each NGTX compound had mainly different pathways in common between HepG2 and HepaRG that could be the result of the large differences in the number of DEGs among these compounds. In addition to the pathways commonly affected in HepG2 and HepaRG after exposure to the five compounds, clear differences were seen regarding common metabolic processes, immune response, and transcription processes. These differences may be due to differences in basal gene expression as described above. The underlying genes of the observed pathways, however, did not always show expression in the opposite direction, but absolute expression levels were higher in HepaRG.

Regarding the unique DEGs of the different experimental groups, a comparison between the AFB1- and BaP-exposed HepG2 showed many differences but also common processes for apoptosis at 12h and for cellular development at 48h. HepaRG, on the other hand, had more processes in common between the AFB1- and BaP-treated cells, but different processes were triggered at different time points. This change over time could be a generic response to GTX carcinogens. Common pathways among the NGTX compounds were only found for HepaRG.

Differences between exposed HepG2 and HepaRG cells were also evident for the phase I and II genes, showing a weak upregulation of them in HepG2 and a strong downregulation of many of them in HepaRG. Nevertheless, both cell lines showed the prerequisite expression of CYP450 and phase II genes for metabolizing and detoxifying AFB1 and BaP. From the CYP450s (i.e., CYP1A2, 1B1, 3A4, 3A5, and 3A7) that can metabolize AFB1 (37, 38), CYP1B1 was upregulated at both time points in AFB1-treated HepaRG, whereas in AFB1-treated HepG2 cells, CYP3A5 and 3A7 were upregulated at 48h. Major players in the metabolism of BaP, that is CYP1A1, CYP1A2, and CYP1B1 (19, 22), were upregulated in both cell lines at both time points. Important phase II genes in the detoxification of AFB1 and BaP include EPHs, aldo-keto reductases, and GSTs (19, 21, 22). Several of these genes were upregulated in HepG2 and HepaRG.

NGTX compounds CsA and E2 are also metabolized by CYP450s. CsA metabolism is mainly catalyzed by members of the CYP3A subfamily (39, 40). In this study, however, several of the CYP3A genes appear downregulated in HepG2 and HepaRG. E2 can be metabolized by CYP1A1, CYP1A2 or CYP3A4 to 2-hydroxyestradiol or by CYP1B1 to 4-hydroxyestradiol (26). CYP1A1 is upregulated in E2-treated HepG2 at both time points, whereas HepaRG showed upregulation of CYP1B1 at 12h. Further metabolism of E2 metabolites does not seem to occur as the genes involved, for example, catechol-O-methyltransferase, (26) are downregulated.

The third NGTX compound TCDD is not metabolized. However, it shows resemblance to BaP as it is an AHR ligand (27). Therefore, in both cell lines, the same CYP450s, for example, CYP1A1, were upregulated with similar expression values after exposure to both compounds for 12h. After 48h, TCDD exposure resulted in higher CYP1A1 expression than BaP exposure. The same effect was also observed for other CYP450s and phase II genes. This is because TCDD, in contrast to BaP, remains longer in the cells, resulting in different cellular responses to these compounds.

Furthermore, the two cell lines were examined for their ability to discriminate between the GTX and NGTX classes. PAM analysis showed that HepG2 performed better at both time points. Correct discrimination using HepaRG was only achieved at 12h, whereas at 48h, E2 was grouped with the GTX class. These results show that HepG2 cells can be used for the discrimination between GTX and NGTX compounds.

In conclusion, we have intensively compared the whole genome gene expression profiles of the hepatoma cell lines, HepG2 and HepaRG, before and after exposure to GTX and NGTX carcinogens. Before exposure, the results revealed an extensive network of genes and pathways differentially regulated for each cell line. The comparison of the basal gene expression between HepG2, HepaRG, PHH and liver clearly showed that HepaRG resembles PHH and liver the most.

After exposure, the common pathways for both cell lines are considered as a result of the direct impact of the compounds and their metabolites. For both cell lines, common pathways were found that are important in carcinogenesis. Hence, the results of this study could be used to improve the biological understanding of the underlying mechanisms of the exposure to GTX and NGTX compounds, especially because of the phenotypic differences of the HepG2 and HepaRG cell lines. However, HepaRG showed a larger in common response among the GTX and NGTX carcinogens compared to HepG2. The comparison of the discrimination capacities of the two cell lines showed that HepG2 can perform better. Therefore, these results have shown that HepaRG is a more suited *in vitro* liver model for biological interpretations of the effects of exposure to chemicals, whereas HepG2 is a more promising *in vitro* liver model for predictive toxicogenomics studies in the area of chemical carcinogenesis. Although, it should be noted that only five carcinogens were used in this study.

Supplementary data

Supplementary data are available online at <http://toxsci.oxfordjournals.org/>

Funding

This research was supported by the Dutch Technology Foundation STW, applied science division of The Netherlands Organisation for Scientific Research, Technology Program of the Ministry of Economic Affairs (STW MFA6809); The Netherlands Genomics Initiative; and the European Union FP6 project carcinoGENOMICS (LSHB-CT-2006-037712).

References

1. Davila JC, Rodriguez RJ, Melchert RB, & Acosta D, Jr. (1998) Predictive value of in vitro model systems in toxicology. *Annu Rev Pharmacol Toxicol* 38:63-96.
2. Hsu IC, et al. (1993) p53 gene mutation and integrated hepatitis B viral DNA sequences in human liver cancer cell lines. *Carcinogenesis* 14(5):987-992.
3. Knasmüller S, et al. (1998) Use of metabolically competent human hepatoma cells for the detection of mutagens and antimutagens. *Mutat Res* 402(1-2):185-202.
4. Wilkening S, Stahl F, & Bader A (2003) Comparison of primary human hepatocytes and hepatoma cell line HepG2 with regard to their biotransformation properties. *Drug Metab Dispos* 31(8):1035-1042.
5. Burczynski ME, et al. (2000) Toxicogenomics-based discrimination of toxic mechanism in HepG2 human hepatoma cells. *Toxicol Sci* 58(2):399-415.
6. Harries HM, Fletcher ST, Duggan CM, & Baker VA (2001) The use of genomics technology to investigate gene expression changes in cultured human liver cells. *Toxicol In Vitro* 15(4-5):399-405.
7. Hong Y, Muller UR, & Lai F (2003) Discriminating two classes of toxicants through expression analysis of HepG2 cells with DNA arrays. *Toxicol In Vitro* 17(1):85-92.
8. van Delft JH, et al. (2004) Discrimination of genotoxic from non-genotoxic carcinogens by gene expression profiling. *Carcinogenesis* 25(7):1265-1276.
9. van Delft JH, et al. (2005) Comparison of supervised clustering methods to discriminate genotoxic from non-genotoxic carcinogens by gene expression profiling. *Mutat Res* 575(1-2):17-33.
10. Olsavsky KM, et al. (2007) Gene expression profiling and differentiation assessment in primary human hepatocyte cultures, established hepatoma cell lines, and human liver tissues. *Toxicol Appl Pharmacol* 222(1):42-56.
11. Westerink WM & Schoonen WG (2007) Phase II enzyme levels in HepG2 cells and cryopreserved primary human hepatocytes and their induction in HepG2 cells. *Toxicol In Vitro* 21(8):1592-1602.
12. Westerink WM & Schoonen WG (2007) Cytochrome P450 enzyme levels in HepG2 cells and cryopreserved primary human hepatocytes and their induction in HepG2 cells. *Toxicol In Vitro* 21(8):1581-1591.
13. Harris AJ, Dial SL, & Casciano DA (2004) Comparison of basal gene expression profiles and effects of hepatocarcinogens on gene expression in cultured primary human hepatocytes and HepG2 cells. *Mutat Res* 549(1-2):79-99.
14. Liguori MJ, Blomme EA, & Waring JF (2008) Trovafloxacin-induced gene expression changes in liver-derived in vitro systems: comparison of primary human hepatocytes to HepG2 cells. *Drug Metab Dispos* 36(2):223-233.
15. Guillouzo A, et al. (2007) The human hepatoma HepaRG cells: a highly differentiated model for studies of liver metabolism and toxicity of xenobiotics. *Chem Biol Interact* 168(1):66-73.
16. Kanebratt KP & Andersson TB (2008) Evaluation of HepaRG cells as an in vitro model for human drug metabolism studies. *Drug Metab Dispos* 36(7):1444-1452.
17. Turpeinen M, et al. (2009) Functional expression, inhibition and induction of CYP enzymes in HepaRG cells. *Toxicol In Vitro* 23(4):748-753.
18. Lambert CB, Spire C, Claude N, & Guillouzo A (2009) Dose- and time-dependent effects of phenobarbital on gene expression profiling in human hepatoma HepaRG cells. *Toxicol Appl Pharmacol* 234(3):345-360.
19. Endo K, et al. (2008) Inhibition of aryl hydrocarbon receptor transactivation and DNA adduct formation by CYP1 isoform-selective metabolic deactivation of benzo[a]pyrene. *Toxicol Appl Pharmacol* 230(2):135-143.
20. Gelboin HV (1980) Benzo[alpha]pyrene metabolism, activation and carcinogenesis: role and regulation of mixed-function oxidases and related enzymes. *Physiol Rev* 60(4):1107-1166.
21. Guengerich FP (2008) Cytochrome p450 and chemical toxicology. *Chem Res Toxicol* 21(1):70-83.
22. Jiang H, et al. (2007) Metabolism of benzo[a]pyrene in human bronchoalveolar H358 cells using liquid chromatography-mass spectrometry. *Chem Res Toxicol* 20(9):1331-1341.
23. Backman L, Appelkvist EL, Sundberg A, Teclebrhan H, & Brunk U (1991) Modulation of metabolism in HepG2 cells upon treatment with cyclosporin A and Nva2-cyclosporin. *Exp Mol Pathol* 54(3):242-254.

24. Gueguen Y, et al. (2007) Compared effect of immunosuppressive drugs cyclosporine A and rapamycin on cholesterol homeostasis key enzymes CYP27A1 and HMG-CoA reductase. *Basic Clin Pharmacol Toxicol* 100(6):392-397.
25. Levy J, Budai K, & Javitt NB (1994) Bile acid synthesis in HepG2 cells: effect of cyclosporin. *J Lipid Res* 35(10):1795-1800.
26. Tsuchiya Y, Nakajima M, & Yokoi T (2005) Cytochrome P450-mediated metabolism of estrogens and its regulation in human. *Cancer Lett* 227(2):115-124.
27. NTP (2005) NTP 11th Report on Carcinogens. *Rep Carcinog* (11):1-A32.
28. Dai M, et al. (2005) Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res* 33(20):e175.
29. Irizarry RA, et al. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4(2):249-264.
30. De Groot PJ, Reiff C, Mayer C, & Muller M (2008) NuGO contributions to GenePattern. *Genes Nutr* 3(3-4):143-146.
31. Tong W, et al. (2003) ArrayTrack--supporting toxicogenomic research at the U.S. Food and Drug Administration National Center for Toxicological Research. *Environ Health Perspect* 111(15):1819-1826.
32. Tong W, et al. (2004) Development of public toxicogenomics software for microarray data management and analysis. *Mutat Res* 549(1-2):241-253.
33. Reich M, et al. (2006) GenePattern 2.0. *Nat Genet* 38(5):500-501.
34. Tibshirani R, Hastie T, Narasimhan B, & Chu G (2002) Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci U S A* 99(10):6567-6572.
35. New DC & Wong YH (2007) Molecular mechanisms mediating the G protein-coupled receptor regulation of cell cycle progression. *J Mol Signal* 2:2.
36. Hockley SL, Arlt VM, Brewer D, Giddings I, & Phillips DH (2006) Time- and concentration-dependent changes in gene expression induced by benzo(a)pyrene in two human cell lines, MCF-7 and HepG2. *BMC Genomics* 7:260.
37. Kamdem LK, Meineke I, Godtel-Armbrust U, Brockmoller J, & Wojnowski L (2006) Dominant contribution of P450 3A4 to the hepatic carcinogenic activation of aflatoxin B1. *Chem Res Toxicol* 19(4):577-586.
38. Murray GI, Melvin WT, Greenlee WF, & Burke MD (2001) Regulation, function, and tissue-specific expression of cytochrome P450 CYP1B1. *Annu Rev Pharmacol Toxicol* 41:297-316.
39. Shaw PM, et al. (1989) Purification and characterization of an anticonvulsant-induced human cytochrome P-450 catalysing cyclosporin metabolism. *Biochem J* 263(3):653-663.
40. Watkins PB (1990) The role of cytochromes P-450 in cyclosporine metabolism. *J Am Acad Dermatol* 23(6 Pt 2):1301-1309; discussion 1301-1311.

Chapter 3

Comparison of phenotypic and transcriptomic effects of false positive genotoxins, true genotoxins and non-genotoxins using HepG2 cells

Mutagenesis. 2011 Sep;26(5):593-604

C. Magkoufopoulou,
S.M. Claessen, D.G.J. Jennen,
J.C.S. Kleinjans, J.H.M. van Delft

Abstract

The conventional *in vitro* assays for genotoxicity assessment of chemicals are characterized by a high false positive rate, thus failing to correctly predict their *in vivo* genotoxic effects. This study aimed to identify the cellular mechanisms induced by the false positive genotoxins Quercetin, 8-Hydroxyquinoline and 17-beta Estradiol in comparison to true genotoxins and non-genotoxins, by combining *in vitro* phenotypic parameters with transcriptomics data from HepG2 cells. The effects of these compounds on the phosphorylation of H2AX, cell cycle distribution and whole genome gene expression following treatment for 12h, 24h and 48h were compared with the effects of true genotoxins (Benzo[a]pyrene and Aflatoxin B1) and non-genotoxins (2,3,7,8-Tetrachlorodibenzo-para-dioxin, Cyclosporin A and Ampicillin C). Quercetin induced similar phenotypic effects as true genotoxins and to some extent similar gene expression alterations. Different gene expression changes were also observed, including the up-regulation of DNA-repair related genes. 8-hydroxyquinoline and 17-beta Estradiol showed no similarities to the true genotoxins at both the phenotypic and the transcriptomic level. In a classification approach, classifiers were selected to discriminate between genotoxins and non-genotoxins. Subsequent analysis for the false positive compounds showed Quercetin to be predicted as genotoxic and 8-hydroxyquinoline and 17-beta Estradiol as non-genotoxic. Our results support that transcriptomics analysis of compound-effects in HepG2 leads to similar results with phenotypic analysis and provides additional mechanistic information. Therefore combined evaluation of gene expression alterations and relevant functional endpoints using HepG2 cells may contribute to the better understanding of modes-of-action of chemicals and the correct evaluation of their genotoxic properties.

Introduction

The conventional *in vitro* assays for evaluating DNA-damaging properties of chemicals often fail to correctly assess genotoxic effects of compounds *in vivo* (1). The currently approved assays by regulatory authorities for the *in vitro* genotoxicity assessment are: the bacterial mutation assay (Ames assay), the mammalian micronuclei assay, the chromosomal aberration assay and the mouse lymphoma assay (2). All these assays have high rates of false positive results (1, 3). Therefore several efforts are made for the reduction of the false positive rates of genotoxicity testing. These include the application of mutant cell lines (4) or the reduction of the selected dose for the conventional assays (5) as well as the development of new alternative *in vitro* methods (6-9). However, none of the above studies is focusing on the mechanisms of the compounds that are falsely classified.

Whereas the key mechanism-of-action of true genotoxins (GTX) is the induction, direct or after metabolic activation, of DNA damage, false positive genotoxins (FPGTX) show positive results in *in vitro* genotoxicity assays, by causing either mutagenicity or clastogenicity, but show negative results in *in vivo* genotoxicity assays. Generally the failure of the standard *in vitro* test battery to correctly evaluate the *in vivo* genotoxic properties of chemicals may frequently be due to the lack of the ability of *in vitro* systems to fully mimic the *in vivo* situation of the correct target cells or tissue. This can be caused by metabolic activation differences, since *in vitro* an artificial metabolic activation system of rat liver S9-mix is used; target cell, tissue or species differences; differences at cell-cell interactions and interplays; or differences at cellular receptors and transporters between *in vitro* and *in vivo* situations (10).

In order to improve *in vitro* evaluation of genotoxic properties of chemicals, a better understanding is required of the mechanisms-of-action for false positive chemicals, ideally to be obtained with *in vitro* methods (11). Toxicogenomics – the evaluation of gene expression alterations in cells or tissues after treatment with toxic substances – is proposed as a tool to better understand the mechanisms-of-actions of chemicals (12). Several studies support this (13-15). Therefore, here we studied whether the combination of toxicogenomics with phenotypic effects provides relevant information for understanding *in vitro* genotoxic mechanisms.

We aimed to identify the molecular pathways modified by FPGTX in HepG2 cells, by combining phenotypic changes with transcriptomic data. Despite the obvious limitations with regard to similarities to primary human hepatocytes, the human hepatic cell line HepG2 is frequently used in toxicology and pharmacology. HepG2 have been repeatedly shown to overexpress many drug metabolizing, phase I and II, enzymes especially after exposure to toxins (16-21). Additionally, transcriptomics analysis of HepG2 confirmed activated biotransformation pathways already at base-line levels (22). HepG2 have been proven a useful tool for genotoxicity studies since genotoxicity induced by chemicals that require metabolic activation is detected in HepG2 by means of micronuclei, sister chromatid exchange assays and single cell gel electrophoresis (7, 17, 23, 24), lack mutations in p53, which is a major factor in DNA damage response (25-27) and they have been successfully used to discriminate GTX from non-genotoxins (NGTX) using transcriptomics-based classification tools (28-31) as well for the identification of hepatotoxic compounds (32, 33).

In this study, the known FPGTX compounds Quercetin (Que), 8-Hydroxyquinoline (8HQ) and 17-beta Estradiol (E2) were examined and compared with the GTX compounds Benzo[a]pyrene (BaP) and Aflatoxin B1 (AFB1) and the NGTX compounds 2,3,7,8-Tetrachlorodibenzo-para-dioxin (TCDD), Cyclosporin A (CsA) and Ampicillin C (AmpC). Several *in vitro* genotoxicity studies resulted in positive outcomes for Que, 8HQ and E2 (34-36). All three chemicals, however, showed negative results in *in vivo* genotoxicity assays (Table 1 and references there in). Que and 8HQ also show negative results in carcinogenicity studies, whereas E2 is considered a human carcinogen (37).

For the phenotypic analysis, firstly, the phosphorylation levels of the histone variant H2AX at Ser-139 (γ H2AX) were evaluated in order to confirm the presence/absence of the genotoxic potential of the selected compounds in HepG2 cells. γ H2AX is a new marker of genotoxicity (38) and an immediate response to the occurrence of double DNA strand breaks (39). Secondly, the cell distribution in the different cell cycle phases was evaluated. First, as an additional confirmation of the genotoxic effects of the selected compounds in view of the fact that DNA damage can activate cell cycle checkpoints and inhibit DNA synthesis; thus genotoxic treatments often alter the cell cycle distribution and lead to an accumulation of cells in the S-phase (40). Secondly, non-genotoxic compounds may cause other effects on the cell cycle. For instance CsA and TCDD have been reported to cause G1 phase arrest (41, 42). The transcriptomics data were used to identify genes correlated with the induced levels of DNA damage and cell cycle distributions, for functional annotation analysis of gene expression alterations per chemical as well as for selection of genes with the potential to discriminate GTX from NGTX compounds. Both the phenotypic and transcriptomics analysis aimed to identify similarities and differences in the induced molecular mechanistic pathways in HepG2 cells between FPGTX and GTX, and FPGTX and NGTX.

Material and methods

Cell culture and treatment

HepG2 cells were cultured in 6-well plates as previously described (28). When the cells were 80% confluent, medium was replaced with fresh medium containing the corresponding dose of each compound diluted in 0.5% DMSO or 0.5% DMSO as a vehicle control (Table 1). All chemicals were obtained from Sigma-Aldrich, except E2 (Steraloids Inc) and TCDD (Cerilliant). For obtaining comparable results all doses were selected based on a MTT cytotoxicity assay resulting in 80% viability after 72h treatment (28). Cells were exposed for 12h, 24h or 48h. These exposure periods were selected based on several reasons. Firstly, the phenotypic induction of DNA damage is detected after relatively short exposure times, whereas the differences in gene expression changes induced by GTX and NGTX are more predominant after longer exposure periods (38, 43). Other reasons are the period that is needed to metabolize the GTX compounds (28) and the cell cycle duration of HepG2 cells (approximately 20h) (26).

Thereafter cells were either methanol-fixated and stored at -20°C for flow cytometric

Table 1: Information on compounds used in this study

Compound	Abbreviation	CAS	Dose (µM)	Mode of action	genotoxicity in vitro			genotoxicity in vivo		Carcinogenicity	References			
					Ames ^a	Result		MLA ^b	MN ^c			CA ^d	IARC Group ^e	
						Presence (+) of metabolic activation system	absent (-)							
Aflatoxin B1	AFB1	1162-65-8	1	DNA adduct formation	+	+	+	-	N/A	+	+	1	(70, 86-88)	
Benzo[a]pyrene	BaP	50-32-8	2	DNA adduct formation	+	+	+	+	+	N/A	+	+	2A	(70, 89, 90)
2,3,7,8-tetra-chloro-dibenzo-para-dioxin	TCDD	1746-01-6	0.01	oxidative damage Ahr-ligand	-	+	+	-	-	N/A	-	N/A	1	(70, 91)
Cyclosporin A	CsA	59865-13-3	3	calcineurin phosphatase inhibitor	-	+	+	-	-	N/A	-	-	1	(70, 92)
Ampicillin	AmpC	7177-48-2	250	beta-lactamase inhibitor	-	+	+	-	-	N/A	-	-	3	(70, 86, 93-95)
8-hydroxy-quinoline	8HQ	148-24-3	15	proteasome inhibitor	+	+	+	-	+	N/A	+	-	3	(70, 96)
Quercetin	Que	117-39-5	50	scavenger of reactive oxygen species	+	+	+	-	+	N/A	+	-	3	(70, 97-100)
17-beta Estradiol	E2	50-28-2	30	ER-ligand DNA adduct formation	-	+	+	-	+	-	+	-	1	(37,70,101,102)

^aAmes mutagenicity assay; ^bMouse Lymphoma Assay; ^cMicronucleus assay; ^dChromosomal Aberration Assay;
^eIARC group 1: carcinogenic to humans, group 2A: probably carcinogenic to humans, group 3: not classifiable as to its carcinogenicity to humans

analysis (FACS) or the culture medium was replaced by TRIZOL (Gibco/BRL) for RNA isolation. Three independent biological replicates were conducted.

Flow cytometric analysis for γ H2AX and cell cycle

Methanol-fixated cells were washed with phosphate-buffered saline containing 1 mg/ml bovine serum albumin (PBS/BSA) and incubated for 1h with the primary antibody mouse anti-phospho-histone H2AX (Ser-139) Clone JBW 301 (Upstate) diluted 1:1000 in PBS/BSA. Cells were washed and incubated for 1h with the secondary antibody Alexa Fluor 488-conjugated anti-mouse IgG, (Invitrogen). Cells were re-suspended in 0.5 ml phosphate-buffered saline containing 20 μ g/ml propidium iodide (Sigma-Aldrich) and 100 μ g/ml RNase A (Sigma-Aldrich). As positive and negative controls, cells were incubated with 5 μ M etoposide (Sigma-Aldrich, CAS: 33419-42-0) and DMSO respectively (38).

FACSort (Becton Dickinson, Sunnyvale, CA) was used for flow cytometric analysis as previously described (44). Cells with significant amounts of anti-phospho-Histone H2AX positive signals were displayed as a percentage of total cells. Cell cycle was analyzed using ModFit LT for Mac (version 2.0). Cells in the G0/G1-, S- or G2/M-phase were expressed as a percentage of the total cells.

Values for both γ H2AX and cell cycle were considered significant in comparison to control cells when p-value < 0.05 (paired t-test).

Total RNA isolation and microarray experiments

Total RNA was extracted using 0.5 ml TRIZOL according to the manufacturer's instructions and purified using RNeasy® Mini Kits (Qiagen). Sample preparation, hybridization, washing, staining and scanning of the Affymetrix Human Genome U133 Plus 2.0 GeneChip arrays were conducted according to the manufacturer's manual as previously described (28). Quality controls were within acceptable limits. Hybridization controls were called present on all arrays and yielded the expected increases in intensities. The data in this publication are accessible in EBI's ArrayExpress (<http://www.ebi.ac.uk/microarray-as/ae/>) through the accession number E-MEXP-2574.

Annotation and normalization of microarray data

The obtained data sets were re-annotated to the MBNI Custom CDF-files (http://brainarray.mbni.med.umich.edu/Brainarray/Database/CustomCDF/genomic_curated_CDF.asp) (45) and RMA normalized (46) using the NuGOExpressionFileCreator in GenePattern (47). MAS5 normalization was conducted for the retrieval of presence and absence calls. Data filtering was conducted for each treatment and time point separately. Genes with more than four absence calls within the three replicates of each treatment and the three corresponding control replicates were omitted from the data. Only the filtered data sets (approximately 10000/treatment) were further used.

Selection of Differentially Expressed Genes

Log2 ratios of treatments to the corresponding control treatments were calculated in Microsoft Excel. Differentially expressed genes (DEGs) per compound and time point were selected using the criteria: absolute average fold change > 1.5 in comparison to control, ab-

solute fold change >1.2 for at least two replicates and same direction of expression for all three replicates. No statistical criteria were used since it has been reported that reproducibility of microarray data is higher when only fold change criteria are used (48). A combination of the DEGs resulting after treatment with GTX and NGTX were used for Hierarchical Clustering Analysis (HCA; Euclidean distance, pairwise complete distance) in Genepattern v.3.2.1. and for Principal Component Analysis (PCA) by ArrayTrack (49).

Correlation analysis

Spearman's correlation coefficient was calculated in GEPAS 4.0 (<http://gepas.bioinfo.cipf.es/>) for: i) the levels of induced DNA damage, expressed as log₂ ratio of the percentage of cells with significant levels of γ H2AX to the control cells, and the log₂ ratio of the mRNA levels as resulted from the whole genome gene expression analysis, and ii) the number of cells in the different cell cycle phases, expressed as log₂ ratio of the percentage of cells in each cell cycle phase to the control cells and the log₂ ratio of the mRNA levels as resulted for the whole genome gene expression analysis. For the correlation analysis, only the GTX and the NGTX treatments for all time points (12h, 24h and 48h) were used. Significantly associated genes were considered those for which the mRNA levels showed a correlation coefficient with a false discovery rate (FDR) lower than 0.05.

The expression values of the significantly associated genes (log₂ ratios) for all treatments were used for PCA in ArrayTrack and HCA (Euclidean distance, pairwise complete-linkage) in GenePattern v.3.2.1. Relative, discrete color scheme was used for the HCA images.

Prediction analysis

PAM (Prediction Analysis for Microarrays 2.13, Stanford University Software) analysis was conducted to select genes that can discriminate between GTX and NGTX compounds. For that reason the DEGs resulting from each compound and incubation time were combined. As input for PAM analysis the expression values of this union for each treatment and time point combination were used (50). As a training set the log₂ ratios of the GTX and the NGTX treatments for each time point were used, whereas as a test set the log₂ ratios of the GTX, NGTX and FPGTX were used. The threshold was set to 4.2 since it fulfilled the criteria of i) highest average probability, which was calculated as the average of the probabilities of each treatment to be discriminated to its true class and ii) the lowest number of genes used for the classification (28).

Network and functional annotation analysis

Network Analysis was conducted in MetaCore Network Tools (GeneGo, San Diego, CA), using the auto-expand option and limit of 15 nodes. Only positive and negative interactions were used for the image. Functional Annotation based on Gene Ontology (GO) Biological Processes was done using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (51). Functional annotation was conducted separately for the up- and down-regulated DEGs for each treatment and time point. For each gene list, a corresponding background list resulting from the data filtering was used. GO terms were considered significant when $EASE < 0.1$ and $p\text{-value} < 0.05$. The same criteria were applied for the selection of significant GO terms of the correlated genes.

Results

Analysis of phenotypic endpoints of genotoxicity

DNA damage

The levels of induced DNA damage were evaluated in order to verify the *in vitro* genotoxic properties of true GTX and the absence of *in vitro* genotoxicity for the NGTX used in this study, as well as to evaluate the *in vitro* genotoxic effects of the FPGTX in HepG2 cells. For that purpose the cellular levels of γ H2AX foci after treatment with the different compounds was quantified (Figure 1A). As expected, both GTX significantly induced γ H2AX levels and all three NGTX did not cause DNA damage. Regarding the FPGTX, Que caused significant levels of γ H2AX after 12h and 24h and 8HQ resulted in a small but significant increase of γ H2AX-positive cells only after 24h, whereas E2 did not cause any DNA damage.

Cell cycle distribution

Since DNA damage can activate different cell cycle checkpoints and lead to cell cycle arrest (40), we further investigated the effect of the different compounds on the cell cycle distribution. For that purpose we quantified the percentage of cells in the different cell cycle phases (Figure 1B). Both GTX resulted in an accumulation of cells in S-phase and in a decreased number of cells in G0/G1-phase after 24h. After 48h, BaP-treated cells showed a similar pattern to the control cells, while AFB1-treated cells showed a significant increase of the number of cells in G2/M-phase.

From the NGTX group, none caused an increase in S-phase cells. CsA, however, caused a significant increase of the number of cells in G0/G1-phase after 12h and 48h and a decrease in G2/M-phase after 24h and 48h.

None of the FPGTX caused any significant effect on the cell cycle after 12h or 24h. After 48h, 8HQ and E2 induced small but significant increases of the number of cells in S-phase.

Whole genome gene expression analysis

Initially the DEGs resulting after treatment with the different compounds at various time points were identified (Table 2, supplementary data 2). To reveal the similarities / differences of the expression changes induced by FPGTX to that of GTX and NGTX, only the DEGs after GTX- and NGTX treatments were used for HCA and PCA (Figure 2 A-C).

In the PCA plots for each time point, Que appears distinct from any other treatment whereas 8HQ and E2 appear closely related to CsA and AmpC. In addition, the dendrograms for the three individual time points show a major distance between Que and the rest of the treatments and a very close relation of 8HQ with AmpC, whereas for E2 a closer relation to AFB1 is observed at 12h, but at later time points E2 appears closely related to AmpC.

Association of the phenotypic endpoints of genotoxicity with gene expression

Next we identified the genes significantly associated over time with levels of DNA damage or cell cycle distribution after GTX and NGTX treatments and the GO processes that they were involved in (supplementary data 1). These genes were also used for HCA and PCA to reveal the similarities / differences of the expression changes of those genes induced by the FPGTX with those induced by GTX and NGTX (Figure 3).

Figure 1: Phenotypic endpoints of genotoxicity. A. Percentage of the cell population with γ H2AX after treatment with the different compounds after 12h (white bars), 24h (grey bars) and 48h (black bars). B. Percentage of the cell population in G0-G1 phase (grey bars), in S-phase (white bars) and G2-M-phase (black bars) after treatment with the different compounds for (a) 12h, (b) 24h and (c) 48h. For both A. and B * indicates $p < 0.05$, ** indicates p -value < 0.01 , *** indicates p -value < 0.001 in comparison to the control.

Figure 2: HCA / PCA of all the treatments using the DEGs of GTX and NGTX treatments. A. 12h (1228 DEGs), B. 24h (1323 DEGs), C. 48h (1616 DEGs). In the PCA plots, red circles indicate GTX-treatments, blue NGTX-treatments and green FPGTX-treatments.

Figure 3: HCA / PCA of all the treatments using genes with significant correlation with A. γ H2AX levels and B. cell cycle distribution of GTX and NGTX treatments. In the PCA plot, light red circles indicate GTX-treatments of 12h, red GTX-treatments of 24h, dark red GTX-treatments of 48h, light blue NGTX-treatments of 12h, blue NGTX-treatments of 24h, dark blue NGTX-treatments of 48h, light green FPGTX-treatments of 12h, green FPGTX-treatments of 24h and dark green FPGTX-treatments of 48h.

DNA damage

Correlation analysis revealed 115 genes associated with the levels of γ H2AX over time. Functional annotation analysis revealed several of these genes to be involved in metabolic processes (supplementary data 1). In the HCA of all treatments based on expression levels of the γ H2AX-associated genes (Figure 3A), two main clusters are observed; one containing the GTX and Que, and the second containing the NGTX, 8HQ and E2. This reveals a similar mRNA expression pattern for the DNA-damage associated genes for Que to that induced by GTX, and a similar mRNA expression pattern for these genes for 8HQ and E2 to the NGTX. In the PCA plot of all the treatments using the γ H2AX-associated genes, however, only the earliest exposure time of Que appears to be similar to the GTX, whereas 8HQ and E2 are again closer to the NGTX.

Cell cycle

Spearman's correlation analysis resulted in 154 genes significantly associated with the cell population in G0/G1-phase, 8 with the cell population in S-phase, that were also correlated with G0/G1 phase, and only one gene with the cell population in G2/M-phase. The functional annotation analysis of these genes revealed several GO-terms related to cell cycle and DNA repair (Supplementary data 1). Figure 3B shows the HCA/PCA for all treatments based on cell cycle-associated genes. Clear differences are observed for the expression patterns of these genes for GTX and NGTX. Correspondingly to the DNA damage-associated genes, Que resulted in expression patterns similar to the GTX-treatments whereas 8HQ and E2 profiles equal to the NGTX. In the PCA plot, however, Que appears distant from the GTX, whereas 8HQ and E2 are again close to the NGTX.

Functional annotation analysis

To explore the underlying mechanisms that might shed light on the FPGTX mechanisms, the DEGs resulted after treatment with these compounds were functionally annotated by DAVID (Table 2). The resulting significant GO-terms that appeared to be strongly represented among the up- or down-regulated DEGs after FPGTX-treatments were further compared with the resulting union of the significant GO terms of the GTX treatments and the union of the significant GO-terms of the NGTX treatments to identify A) processes that might

Table 2: Number of DEGs and GO-terms identified for each compound, the union of GTX and the union of NGTX

		Que	8HQ	E2	AFB1	BaP	union GTX	AmpC	TCDD	CsA	union NGTX
12h	(+) ^a DEGs	449	23	6	18	291		16	212	191	
	GO-terms	30	-	2	21	116	120	1	160	97	238
	(-) ^b DEGs	886	74	23	6	491		7	117	108	
	GO-terms	39	-	11	-	30	30	-	35	73	108
24h	(+) ^a DEGs	365	11	8	134	472		5	341	167	
	GO-terms	137	2	12	49	178	188	6	143	57	188
	(-) ^b DEGs	391	36	10	27	330		2	151	117	
	GO-terms	150	14	3	12	79	86	-	45	73	102
48h	(+) ^a DEGs	682	19	29	143	283		4	568	138	
	GO-terms	248	2	27	44	100	106	3	208	53	225
	(-) ^b DEGs	602	48	37	75	184		9	458	274	
	GO-terms	234	17	63	40	42	66	1	48	139	188

^a up-regulated DEGs and resulting GO terms; ^b down-regulated DEGs and resulting GO-terms.

explain the *in vitro* genotoxic effect of these compounds, B) processes that might explain the *in vivo* non-genotoxic effect of these compounds, and C) processes related to each of the FPGTX compounds (Table 3). All the affected GO-terms for all the compounds can be found in supplementary data 3. For all the compounds, multiple processes related to responses to a variety of stimuli were represented by both the up- and down-regulated DEGs.

Quercetin

Processes that appeared to be commonly affected by the FPGTX Que and the true GTX included: apoptosis/cell death for the up-regulated DEGs throughout time, cell cycle arrest and positive regulation of biosynthetic processes for the up-regulated DEGs of 24h, biosynthetic and catabolic processes for the down-regulated DEGs of 24h and 48h as well as homeostatic and transport processes for the down-regulated DEGs at 48h. In addition multicellular organismal processes were commonly affected by Que and GTX for all time points.

Processes that were commonly affected by Que and the NGTX included: cell division and mitosis for the down-regulated DEGs of 12h; signal transduction for the up-regulated DEGs of both 24h and 48h; homeostasis and cell communication for the down-regulated DEGs of 48h; biosynthetic and catabolic processes for the down-regulated DEGs of 48h. Multicellular organismal processes were also commonly affected by Que and the NGTX at 24h and 48h.

Processes that appeared to be affected only by Que included: several metabolic, transport-related and multicellular organismal processes throughout time; maintenance of location for the up-regulated DEGs of 12h and 24h; homeostasis and immune response for the down-regulated DEGs of 24h and 48h; DNA repair and transcription for the up-regulated DEGs of 24h and 48h; cell cycle checkpoint and DNA replication for the up-regulated DEGs of 48h; inflammatory response and regulation of enzymatic activity for the down-regulated DEGs of 48h. Lastly, up-regulated DEGs of 24h were involved in JAK-STAT signaling, whereas down-regulated DEGs of 48h were involved in CDC42 signaling cascade.

8-Hydroxyquinoline

Functional Annotation Analysis of DEGs after treatment with 8HQ did not reveal any GO-terms at 12h. Processes commonly affected by 8HQ and GTX were response to chemical stimulus for the down-regulated DEGs of 24h and sterol biosynthetic process for the down-regulated DEGs of 48h. Two processes for the up-regulated DEGs of 48h were affected by 8HQ and NGTX and both were related to regulation of cell size. Processes affected only by 8HQ included monocarboxylic acid metabolic process and digestion for the up-regulated DEGs of 24h, aging for the down-regulated DEGs of 24h, metabolic and transport related processes for the down-regulated DEGs of 48h.

Estradiol

Processes commonly affected by E2 and GTX included: metabolic processes at 12h and 24h; ion transport for the down-regulated DEGs of 48h and multicellular organismal processes at 48h. Processes commonly affected by E2 and NGTX included steroid metabolic process for the up-regulated DEGs of 24h; epithelium development for the up-regulated DEGs of 48h; organic acid catabolic process and anion transport for the down-regulated DEGs of 48h.

Table 3: Number of GO-terms appearing in common between each FPGTX compound and the union of GTX, the union of NGTX and specific for each FPGTX

	12h				24h				48h					
	In common with GTX		In common with NGTX		In common with GTX		In common with NGTX		In common with GTX		In common with NGTX		In common with NGTX	
	In common with GTX	In common with NGTX	FPGTX specific	FPGTX specific	In common with GTX	In common with NGTX	FPGTX specific	FPGTX specific	In common with GTX	In common with NGTX	FPGTX specific	FPGTX specific	In common with GTX	In common with NGTX
(+) ^a	6	4	(muscle contraction; multicellular organismal process)	14	32	(apoptosis/cell death; cell cycle arrest; positive regulation of biosynthetic process; positive regulation of development; reproduction)	14	55	21	48	134	(DNA repair; transcription; cell cycle checkpoint; DNA replication; signaling; system development; smooth muscle contraction; immune system process)	48	(signal transduction; homeostasis; cell communication; circulatory system process; positive regulation of development; reproduction)

^aGO terms resulting from up-regulated DEGs; ^bGO-terms resulting from down-regulated DEGs

down-regulated DEGs of 48h.

Prediction analysis

shows a connection with the DNA damage response transcription factors: p53 and SP1.

Figure 4: Genotoxicity prediction analysis and network analysis for selected classifiers. A Results of the prediction analysis. The probability of each compound to be GTX is indicated with grey squares and the probability to be NGTX is indicated with black squares for 12h (left), 24h (middle) and 48h (right). For the prediction of the compound at each time point the average of the log2 ratios of the three replicates for the selected genes was used. B. Auto-expand Network for selected genes by PAM. A legend is provided indicating the functional information for each of the symbols. The type of interaction between the genes is indicated by the different colors (green: activation, red: inhibition). Symbols surrounded by circles indicate genes used for the prediction. The checkbox circles next to the symbols indicate different expression (up- or down-regulation) by the different compounds.

Discussion

In this study we aimed at identifying the similarities and the differences between the effects of FPGTX and true GTX – NGTX compounds in HepG2 cells. For that reason we compared the effects of Que, 8HQ and E2 at the phenotypic as well as at the transcriptomic level with the effects of known GTX and NGTX.

The GTX selected for this study are BaP and AFB1. Both require metabolic transformation to exert their genotoxic properties. BaP needs to be metabolized to Benzo[a]pyrene 7,8-diol 9,10-epoxide (BPDE) and AFB1 to AFB1-8,9-exo-epoxide, which covalently binds to DNA thereby forming DNA adducts (52, 53). Here, we confirm the genotoxic effect of both chemicals in HepG2 cells as they increase the phosphorylation of the histone variant H2AX, a specific marker for DNA double-strand breaks (38), while the NGTX remain ineffective. Furthermore, both BaP and AFB1 induced S-phase arrest after 24h of treatment, as it has been previously reported (54, 55).

The NGTX used in this study were TCDD, CsA and AmpC. The toxic effect of TCDD is considered to be mediated by its binding to the Aryl Hydrocarbon Receptor (AhR) and the subsequent induction of gene expression, e.g. of CYP1A1 (56). In our study, CYP1A1 is highly up-regulated after TCDD-exposure for all exposure periods. It has been reported that both TCDD and CsA can cause G1 phase arrest (41, 42). However, in our study, only CsA results in a cell-accumulation in G1-phase whereas TCDD does not alter the cell cycle distribution.

Que is the only FPGTX examined in our study, which phenotypically resembled the effects of GTX compounds. Que induced similar levels of DNA damage and resulted in a similar cell cycle distribution as GTX. Previous studies have also demonstrated the ability of Que to induce the phosphorylation of the histone variant H2AX (57) and to alter the cell cycle distribution by causing either G0/G1 arrest or G2/M arrest in different cell types (58–62). In contrast to these, we found Que to induce S-phase arrest in HepG2 cells, similarly to both GTX. However Que induced different patterns than GTX compounds on the expression levels of the cell cycle- and γ H2AX-correlated genes at 24h and 48h, some of which are involved in p53-mediated DNA repair signaling (Figure 3; supplementary data 1). This indicates that at early (12h) exposure durations Que and GTX may activate similar mechanisms for DNA repair; however different pathways are activated after longer exposure periods.

By the PCA (Figure 2) we observe that Que treatments result in different whole genome gene expression alterations compared to either the GTX or NGTX compounds. In addition, functional annotation of the DEGs after Que-treatment revealed mostly differences in comparison to either the GTX or the NGTX compounds. Several signaling pathways were affected only by Que (e.g. regulation of I-kappaB kinase/NF-kappaB cascade at 48h). The effect of Que on the transcriptional activity of NF-kappaB was previously reported, supporting its anti-inflammatory activity (63). Among the similarly affected processes by Que and GTX compounds, apoptosis and cell death were induced at all time points, representing effects on DNA damage response pathways but no alteration of the viability of the cells was observed. However, in other cell types, Que can cause cell death (58, 60).

Que can form transient covalent DNA adducts *in vitro* (64) and can be transformed to quinone methides, which are DNA-reactive intermediates, whereas *in vivo* Que is enzymatically metabolized to the non-DNA reactive 3γ-O-methylquercetin in the absorptive cells of the intestinal epithelium. Thus, the lack of *in vivo* genotoxicity of Que is currently explained by the fast metabolization of Que to non-DNA reactive metabolites (34).

DNA repair processes appeared to be affected at 24h and 48h by Que. At 24h Que induces similar DNA repair mechanisms to GTX compounds which may be p53-dependent and that may result in similar cell distribution at 24h. At 48h, however, Que-treatments induce the up-regulation of genes involved in nucleotide excision repair mechanisms (i.e., DDB2, ERCC1, GTF2H1, GTF2H4, POLD4, PNKP, PCNA, RFC4 and XPC). This DNA repair mechanism is responsible for the removal of bulky DNA adducts (65). This in combination with the observations for the expression changes for the γH2AX- and cell cycle-correlated genes further suggests that Que induces DNA repair mechanisms that are not affected by GTX-treatments, especially at later time points.

In summary, our study supports the induction of genotoxic effects by Que *in vitro* and the activation of nucleotide excision repair. This activation of DNA repair enzymes may counteract the DNA damage induced by Que and may further explain why co-treatments of Que with GTX chemicals leads to the reduction of the DNA-damage induced by the GTX chemical, and may further explain its anti-mutagenic activity (66-68).

The other two FPGTX compounds examined in this study, 8HQ and E2, showed no similarities with any GTX compound, neither at the phenotypic nor at the transcriptomic level. A proposed mechanism for the *in vitro* genotoxicity of 8HQ refers to its property as a chelating agent to bind iron and thereby causing DNA strand breaks (69). *In vivo* studies show that 8HQ is not GTX (70). With respect to the transcriptomic responses to 8HQ, only a limited number of DEGs could be identified, which are similar to those induced by the NGTX. In particular, 8HQ seems closely related to the non-carcinogenic compound AmpC (71). This observation is in agreement with the fact that 8HQ is non-carcinogenic (35). In our study, we observe the genotoxic effects of 8HQ only at a phenotypic level; initially by the induction of the phosphorylation of H2AX at 24h and then by a delayed increase of the cell fraction in S-phase at 48h. Transcriptomic analyses show no similarities of 8HQ with the true *in vivo* GTX compounds. This confirms that also the H2AX assay gives a false positive result for this compound, whereas the genomic analyses provide a correct assessment.

Regarding E2, in contrast to previous findings in MCF-7 cells (72), it did not induce the phosphorylation of H2AX in HepG2 cells. E2-treatments resulted in a small increase in the number of cells in S-phase only after the longest exposure period, whereas both GTX induced an S-phase arrest at 24h. In MCF-7 cells it has been described that E2 increases the cell fraction in S-phase only after short exposure times, whereas for longer exposure periods E2 leads to G2/M arrest (73). The phenotypic effect of E2 on the cell cycle in our study is also supported by the transcriptomics analysis, where we observe the induction of the expression of genes involved in cell cycle regulation only at 48h. Furthermore, the transcriptomic changes induced by E2 are similar to those induced by NGTX. In particular, E2

appears close to the carcinogenic compound CsA (74). The functional annotation analysis of the DEGs after E2 treatment revealed several negatively affected processes. Among them, blood vessel development was identified at the shortest exposure period. In agreement with this, it has been reported that E2 promotes angiogenesis (75) whereas the E2 metabolite 2-methoxyestradiol inhibits angiogenesis (76).

E2 needs to be metabolized in order to exert its genotoxic activity *in vitro* and to lead to DNA-adducts formation (36). Particularly, E2 is transformed to either 4-hydroxyestradiol by CYP1B1 or to 2-hydroxyestradiol by CYP1A1, CYP1A2 or CYP3A4 (77). In humans and animals 2-hydroxyestradiol is further detoxified, among others, by COMT and SULT1A1 (78). The activity of 2-hydroxyestradiol for the formation of DNA adducts and the production of reactive oxygen species has been proven to be lower in comparison to the activity of 4-hydroxyestradiol (79, 80). In our study, E2 induces the expression of CYP1A1 (log2 ratios of 1.45, 1.16 and 0.92 for 12h, 24h and 48h respectively) but not of CYP1B1 throughout time, and thus likely increases the CYP1A1 enzymatic activity. This together with the results of the functional annotation analysis suggest that HepG2 cells metabolize E2 to 2-hydroxyestradiol instead of to 4-hydroxyestradiol, and this may explain why no significant levels of DNA damage are observed. In summary, E2 is probably metabolized to 2-hydroxyestradiol. Since no expression changes are observed for the enzymes responsible for the further metabolism of 2-hydroxyestradiol, we hypothesize that it may generate free radicals, that although they do not induce significant levels of H2AX phosphorylation, they may be responsible for the cell cycle effect that is observed at 48h.

Further, the transcriptomics data were used for the selection of a gene-set with the potential to predict correctly the genotoxic effects of the FPGTX compounds. The resulting gene-set led to the prediction of Que as GTX and the prediction of 8HQ and E2 as NGTX (Figure 4) confirming the results of the phenotypic analyses. This gene set appears to be regulated by the DNA-damage related transcription factors p53 and SP1 (81). These results also suggest that biologically relevant genes for genotoxicity evaluation can be obtained by means of transcriptomics without the association with phenotypic endpoints. Whether this gene set can predict genotoxicity in general, requires the inclusion of many more compounds tested in the same microarray platform and using the same cell line. So far only one study has been conducted using Affymetrix microarrays to measure compound-induced transcriptomics changes in HepG2 cells, but this study is limited to the analysis of Ethanol induced effects (82). Several other studies exist that analyze the compound-induced gene expression alterations in HepG2 using, however, other platforms for the analysis (20, 29, 83-85). Therefore, a more extended study including more compounds using the same platform for the evaluation of the expression alterations induced in HepG2 cells has to be conducted for the evaluation of the ability of transcriptomics to evaluate genotoxic properties. Besides the evaluation of transcriptomics changes induced by different chemicals in HepG2, it is also very important to further evaluate the ability of HepG2 cells to metabolize compounds for which no information is available. Although HepG2 are metabolically competent, the expression of some CYP450 enzymes is not detected or very low (16). In our study, the

expression of the CYP450 enzymes involved in the metabolism of the compounds used, as mentioned above, is detected. In case other compounds are tested in HepG2 cells, characterization of their metabolites/metabolism may be needed. For those compounds, however, that are metabolized by enzymes absent in HepG2 cells, the application of HepG2 cells over-expressing the absent enzymes should be considered.

In conclusion, from the three FPGTX examined here, only Que showed clear similarities to the GTX used in this study both at the phenotypic and the transcriptomic level and was also predicted as a GTX-compound by PAM analysis; hence pointing out that Que is also for HepG2 cells a true *in vitro* genotoxin. At the transcriptomic level, however, Que also shows very distinct effects from both BaP and AFB1, such as genes involved in nucleotide excision repair. This may explain the dual nature of Que, namely as an *in vitro* GTX compound and as an anti-mutagen. Both 8HQ and E2 induced mostly similar responses at both the phenotypic and transcriptomic level as the NGTX and were predicted as NGTX compounds, thus being *in vitro* non-genotoxins in HepG2 cells. Due to the small numbers of DEGs induced by 8HQ and E2, the functional annotation analysis revealed only a small number of affected processes, and therefore does not shed further light on the possible mechanisms that can explain their FPGTX nature. In this study, however, we reveal new mechanisms and support existing hypotheses of the mode-of-actions of Que, 8HQ and E2. Our results support that transcriptomics analysis can be used for the prediction of genotoxic properties. However the integration of gene expression alterations to relevant functional endpoints may contribute to understanding the discrepancies between *in vitro* and *in vivo* genotoxicity test outcomes, and provide relevant mechanistic information.

Acknowledgments

We thank Mieke Henfling (Molecular Cell Biology, Maastricht University) for assisting with the settings of FACS analysis for γ H2AX measurements.

Supplementary data

Supplementary Data are available online (<http://toxicogenomics-um.nl/d/3818147>)

References

1. Kirkland D, Aardema M, Henderson L, & Möller L (2005) Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens I. Sensitivity, specificity and relative predictivity. *Mutat Res*. 584(1-2):1-256.
2. ICH (2008) Guidance on genotoxicity testing and data interpretation for pharmaceuticals intended for human use S2(R1).
3. Pfuhler S, *et al.* (2010) A tiered approach to the use of alternatives to animal testing for the safety assessment of cosmetics: genotoxicity. A COLIPA analysis. *Regul Toxicol Pharmacol* 57(2-3):315-324.
4. Evans TJ, Yamamoto KN, Hirota K, & Takeda S (2010) Mutant cells defective in DNA repair pathways provide a sensitive high-throughput assay for genotoxicity. *DNA Repair (Amst)* 9(12):1292-1298.
5. Parry JM, Parry E, Phrakonkham P, & Corvi R (2010) Analysis of published data for top concentration considerations in mammalian cell genotoxicity testing. *Mutagenesis* 25(6):531-538.
6. Birrell L, Cahill P, Hughes C, Tate M, & Walmsley RM (GADD45a-GFP GreenScreen HC assay results for the ECVAM recommended lists of genotoxic and non-genotoxic chemicals for assessment of new genotoxicity tests. *Mutat Res* 695(1-2):87-95.
7. Winter HK, *et al.* (2008) Use of four new human-derived liver-cell lines for the detection of genotoxic compounds in the single-cell gel electrophoresis (SCGE) assay. *Mutation research* 657(2):133-139.
8. Liu X, *et al.* (2008) Development of a highthroughput yeast-based assay for detection of metabolically activated genotoxins. *Mutat Res* 653(1-2):63-69.
9. Fontaine FR, *et al.* (2004) Optimisation of the comet genotoxicity assay in freshly isolated murine hepatocytes: detection of strong in vitro DNA damaging properties for styrene. *Toxicol In Vitro* 18(3):343-350.
10. Brambilla G & Martelli A (2004) Failure of the standard battery of short-term tests in detecting some rodent and human genotoxic carcinogens. *Toxicology* 196(1-2):1-19.
11. Ohno Y (2002) ICH guidelines—implementation of the 3Rs (refinement, reduction, and replacement): incorporating best scientific practices into the regulatory process. *Ilar J* 43 Suppl:S95-98.
12. Nuwaysir EF, Bittner M, Trent J, Barrett JC, & Afshari CA (1999) Microarrays and toxicology: the advent of toxicogenomics. *Mol Carcinog* 24(3):153-159.
13. Hamadeh HK, *et al.* (2002) Gene Expression Analysis Reveals Chemical-Specific Profiles *Toxicol Sci* 67(2):219-231.
14. Cui Y & Paules RS (2010) Use of transcriptomics in understanding mechanisms of drug-induced toxicity. *Pharmacogenomics* 11(4):573-585.
15. Ellinger-Ziegelbauer H, Aubrecht J, Kleinjans JC, & Ahr HJ (2009) Application of toxicogenomics to study mechanisms of genotoxicity and carcinogenicity. *Toxicol Lett* 186(1):36-44.
16. Guo L, *et al.* (2010) Similarities and Differences in the Expression of Drug Metabolizing Enzymes between Human Hepatic Cell Lines and Primary Human Hepatocytes. *Drug Metab Dispos*.
17. Knasmüller S, *et al.* (2004) Use of human-derived liver cell lines for the detection of environmental and dietary genotoxicants; current state of knowledge. *Toxicology* 198(1-3):315-328.
18. Dvorak Z, Vrzal R, Starha P, Klanicova A, & Travnicek Z (2010) Effects of dinuclear copper(II) complexes with 6-(benzylamino)purine derivatives on AhR and PXR dependent expression of cytochromes P450 CYP1A2 and CYP3A4 genes in primary cultures of human hepatocytes. *Toxicol In Vitro* 24(2):425-429.
19. Chen Y, Huang C, Zhou T, & Chen G (2008) Genistein induction of human sulfotransferases in HepG2 and Caco-2 cells. *Basic Clin Pharmacol Toxicol* 103(6):553-559.
20. Liguori MJ, Blomme EA, & Waring JF (2008) Trovafloxacin-induced gene expression changes in liver-derived in vitro systems: comparison of primary human hepatocytes to HepG2 cells. *Drug Metab Dispos* 36(2):223-233.
21. Lancon A, *et al.* (2007) Resveratrol in human hepatoma HepG2 cells: metabolism and inducibility of detoxifying enzymes. *Drug Metab Dispos* 35(5):699-703.
22. Jennen DG, *et al.* (2010) Biotransformation pathway maps in WikiPathways enable direct visualization of drug metabolism related expression changes. *Drug Discov Today* 15(19-20):851-858.
23. Natarajan AT & Darroudi F (1991) Use of human hepatoma cells for in vitro metabolic activation of chemical mutagens/carcinogens. *Mutagenesis* 6(5):399-403.
24. Majer BJ, *et al.* (2004) Genotoxic effects of dietary and lifestyle related carcinogens in human derived hepatoma (HepG2, Hep3B) cells. *Mutation research* 551(1-2):153-166.

25. Hsu I, *et al.* (1993) p53 gene mutation and integrated hepatitis B viral DNA sequences in human liver cancer cell lines. *Carcinogenesis*. 14(5):987-992.
26. Knasmöller S, *et al.* (1998) Use of metabolically competent human hepatoma cells for the detection of mutagens and antimutagens. *Mutat Res.* 402(1-2):185-202.
27. Wilkening S, Stahl F, & Bader A (2003) Comparison of primary human hepatocytes and hepatoma cell line HepG2 with regard to their biotransformation properties. *Drug Metab Dispos* 31(8):1035-1042.
28. Jennen DG, *et al.* (2010) Comparison of HepG2 and HepaRG by whole-genome gene expression analysis for the purpose of chemical hazard identification. *Toxicol Sci* 115(1):66-79.
29. van Delft JH, *et al.* (2004) Discrimination of genotoxic from non-genotoxic carcinogens by gene expression profiling. *Carcinogenesis* 25(7):1265-1276.
30. van Delft JH, *et al.* (2005) Comparison of supervised clustering methods to discriminate genotoxic from non-genotoxic carcinogens by gene expression profiling. *Mutation research* 575(1-2):17-33.
31. Westerink WM, Stevenson JC, Horbach GJ, & Schoonen WG (The development of RAD51C, Cystatin A, p53 and Nrf2 luciferase-reporter assays in metabolically competent HepG2 cells for the assessment of mechanism-based genotoxicity and of oxidative stress in the early research phase of drug development. *Mutation research* 696(1):21-40.
32. O'Brien PJ, *et al.* (2006) High concordance of drug-induced human hepatotoxicity with in vitro cytotoxicity measured in a novel cell-based model using high content screening. *Arch Toxicol* 80(9):580-604.
33. Schoonen WG, de Roos JA, Westerink WM, & Debiton E (2005) Cytotoxic effects of 110 reference compounds on HepG2 cells and for 60 compounds on HeLa, ECC-1 and CHO cells. II mechanistic assays on NAD(P)H, ATP and DNA contents. *Toxicol In Vitro* 19(4):491-503.
34. Harwood M, *et al.* (2007) A critical review of the data related to the safety of quercetin and lack of evidence of in vivo toxicity, including lack of genotoxic/carcinogenic properties. *Food Chem Toxicol* 45(11):2179-2205.
35. Andersen A (2006) Final amended report on the safety assessment of oxyquinoline and oxyquinoline sulfate as used in cosmetics. *Int J Toxicol* 25 Suppl 1:1-9.
36. Cavalieri E, Frenkel K, Liehr JG, Rogan E, & Roy D (2000) Estrogens as endogenous genotoxic agents – DNA adducts and mutations. *J Natl Cancer Inst Monogr* (27):75-93.
37. IARC (1987) Monographs on the Evaluation of Carcinogenic Risks to Humans, Suppl.7: steroidal estrogens. Summaries and evaluations, IARC. .
38. Watters GP, Smart DJ, Harvey JS, & Austin CA (2009) H2AX phosphorylation as a genotoxicity endpoint. *Mutation research* 679(1-2):50-58.
39. Dickey JS, *et al.* (2009) H2AX: functional roles and potential applications. *Chromosoma* 118(6):683-692.
40. Skladanowski A, Bozko P, & Sabisz M (2009) DNA structure and integrity checkpoints during the cell cycle and their role in drug targeting and sensitivity of tumor cells to anticancer treatment. *Chem Rev* 109(7):2951-2973.
41. Seki Y, *et al.* (2005) In vitro effect of cyclosporin A, mitomycin C and prednisolone on cell kinetics in cultured human umbilical vein endothelial cells. *Thromb Res* 115(3):219-228.
42. Barhooover MA, Hall JM, Greenlee WF, & Thomas RS (2010) Aryl hydrocarbon receptor regulates cell cycle progression in human breast cancer cells via a functional interaction with cyclin-dependent kinase 4. *Mol Pharmacol* 77(2):195-201.
43. Mathijs K, *et al.* (2009) Discrimination for genotoxic and non-genotoxic carcinogens by gene expression profiling in primary mouse hepatocytes improves with exposure time. *Toxicol Sci*.
44. Staal YC, *et al.* (2007) Binary PAH mixtures cause additive or antagonistic effects on gene expression but synergistic effects on DNA adduct formation. *Carcinogenesis* 28(12):2632-2640.
45. Dai M, *et al.* (2005) Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res* 33(20):e175.
46. Irizarry RA, *et al.* (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4(2):249-264.
47. De Groot PJ, Reiff C, Mayer C, & Muller M (2008) NuGO contributions to GenePattern. *Genes Nutr* 3(3-4):143-146.
48. Guo L, *et al.* (2006) Rat toxicogenomic study reveals analytical consistency across microarray platforms. *Nat Biotechnol* 24(9):1162-1169.
49. Tong W, *et al.* (2003) ArrayTrack – supporting toxicogenomic research at the U.S. Food and Drug Administration National Center for Toxicological Research. *Environ Health Perspect* 111(15):1819-1826.

50. Tibshirani R, Hastie T, Narasimhan B, & Chu G (2002) Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci U S A* 99(10):6567-6572.
51. Dennis G, Jr., et al. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 4(5):P3.
52. Baird WM, Hooven LA, & Mahadevan B (2005) Carcinogenic polycyclic aromatic hydrocarbon-DNA adducts and mechanism of action. *Environ Mol Mutagen* 45(2-3):106-114.
53. Bedard LL & Massey TE (2006) Aflatoxin B1-induced DNA damage and its repair. *Cancer Lett* 241(2):174-183.
54. Andrysik Z, et al. (2007) The aryl hydrocarbon receptor-dependent deregulation of cell cycle control induced by polycyclic aromatic hydrocarbons in rat liver epithelial cells. *Mutation research* 615(1-2):87-97.
55. Ricordy R, Gensabella G, Cacci E, & Augusti-Tocco G (2002) Impairment of cell cycle progression by aflatoxin B1 in human cell lines. *Mutagenesis* 17(3):241-249.
56. Mandal PK (2005) Dioxin: a review of its environmental effects and its aryl hydrocarbon receptor biology. *J Comp Physiol B* 175(4):221-230.
57. Ye R, et al. (2004) The isoflavonoids genistein and quercetin activate different stress signaling pathways as shown by analysis of site-specific phosphorylation of ATM, p53 and histone H2AX. *DNA Repair (Amst)* 3(3):235-244.
58. Choi EJ, Bae SM, & Ahn WS (2008) Antiproliferative effects of quercetin through cell cycle arrest and apoptosis in human breast cancer MDA-MB-453 cells. *Arch Pharm Res* 31(10):1281-1285.
59. Shan BE, Wang MX, & Li RQ (2009) Quercetin inhibit human SW480 colon cancer growth in association with inhibition of cyclin D1 and survivin expression through Wnt/beta-catenin signaling pathway. *Cancer Invest* 27(6):604-612.
60. Tan J, Wang B, & Zhu L (2009) Regulation of survivin and Bcl-2 in HepG2 cell apoptosis induced by quercetin. *Chem Biodivers* 6(7):1101-1110.
61. Zhang Q, Zhao XH, & Wang ZJ (2009) Cytotoxicity of flavones and flavonols to a human esophageal squamous cell carcinoma cell line (KYSE-510) by induction of G2/M arrest and apoptosis. *Toxicol In Vitro* 23(5):797-807.
62. Tanigawa S, Fujii M, & Hou DX (2008) Stabilization of p53 is involved in quercetin-induced cell cycle arrest and apoptosis in HepG2 cells. *Biosci Biotechnol Biochem* 72(3):797-804.
63. Ruiz PA, Braune A, Holzwimmer G, Quintanilla-Fend L, & Haller D (2007) Quercetin inhibits TNF-induced NF-kappaB transcription factor recruitment to proinflammatory gene promoters in murine intestinal epithelial cells. (Translated from eng) *The Journal of nutrition* 137(5):1208-1215 (in eng).
64. van der Woude H, et al. (2005) Formation of transient covalent protein and DNA adducts by quercetin in cells with and without oxidative enzyme activity. (Translated from eng) *Chem Res Toxicol* 18(12):1907-1916 (in eng).
65. Shuck SC, Short EA, & Turchi JJ (2008) Eukaryotic nucleotide excision repair: from understanding mechanisms to influencing biology. (Translated from eng) *Cell Res* 18(1):64-72 (in eng).
66. Gupta C, Vikram A, Tripathi DN, Ramarao P, & Jena GB (Antioxidant and antimutagenic effect of quercetin against DEN induced hepatotoxicity in rat. *Phytother Res* 24(1):119-128.
67. Ramos AA, Lima CF, Pereira ML, Fernandes-Ferreira M, & Pereira-Wilson C (2008) Antigenotoxic effects of quercetin, rutin and ursolic acid on HepG2 cells: evaluation by the comet assay. *Toxicol Lett* 177(1):66-73.
68. Edenharder R, Krieg H, Kottgen V, & Platt KL (2003) Inhibition of clastogenicity of benzo[a]pyrene and of its trans-7,8-dihydrodiol in mice in vivo by fruits, vegetables, and flavonoids. *Mutation research* 537(2):169-181.
69. Leanderson P & Tagesson C (1996) Iron bound to the lipophilic iron chelator, 8-hydroxyquinoline, causes DNA strand breakage in cultured lung cells. *Carcinogenesis* 17(3):545-550.
70. NTP (NTP: <http://ntp-server.niehs.nih.gov>).
71. IARC (1990) Summaries & Evaluations - Vol.50: Ampicillin.
72. Iso T, Watanabe T, Iwamoto T, Shimamoto A, & Furuichi Y (2006) DNA damage caused by bisphenol A and estradiol through estrogenic activity. *Biol Pharm Bull* 29(2):206-210.
73. Liu H, et al. (2003) Apoptotic action of 17-beta Estradiol in raloxifene-resistant MCF-7 cells in vitro and in vivo. (Translated from eng) *Journal of the National Cancer Institute* 95(21):1586-1597 (in eng).
74. IARC (1997) Monographs on the Evaluation of Carcinogenic Risks to Humans Vol.50: Pharmaceutical Drugs.
75. Rubanyi GM, Johns A, & Kauser K (2002) Effect of estrogen on endothelial function and angiogenesis. *Vascul Pharmacol* 38(2):89-98.

76. Fotsis T, *et al.* (1994) The endogenous oestrogen metabolite 2-methoxyoestradiol inhibits angiogenesis and suppresses tumour growth. *Nature* 368(6468):237-239.
77. Tsuchiya Y, Nakajima M, & Yokoi T (2005) Cytochrome P450-mediated metabolism of estrogens and its regulation in human. *Cancer Lett* 227(2):115-124.
78. Shatalova EG, Walther SE, Favorova OO, Rebbeck TR, & Blanchard RL (2005) Genetic polymorphisms in human SULT1A1 and UGT1A1 genes associate with breast tumor characteristics: a case-series study. *Breast Cancer Res* 7(6):R909-921.
79. Zahid M, Kohli E, Saeed M, Rogan E, & Cavalieri E (2006) The greater reactivity of estradiol-3,4-quinone vs estradiol-2,3-quinone with DNA in the formation of depurinating adducts: implications for tumor-initiating activity. *Chem Res Toxicol* 19(1):164-172.
80. Salama SA, *et al.* (2008) Catecholestrogens induce oxidative stress and malignant transformation in human endometrial glandular cells: protective effect of catechol-O-methyltransferase. *Int J Cancer* 123(6):1246-1254.
81. Iwahori S, *et al.* (2008) Identification of phosphorylation sites on transcription factor Sp1 in response to DNA damage and its accumulation at damaged sites. *Cell Signal* 20(10):1795-1803.
82. Castaneda F, Rosin-Steiner S, & Jung K (2007) Functional genomics analysis of low concentration of ethanol in human hepatocellular carcinoma (HepG2) cells. Role of genes involved in transcriptional and translational processes. *Int J Med Sci* 4(1):28-35.
83. Harris AJ, Dial SL, & Casciano DA (2004) Comparison of basal gene expression profiles and effects of hepatocarcinogens on gene expression in cultured primary human hepatocytes and HepG2 cells. *Mutation research* 549(1-2):79-99.
84. Hockley SL, *et al.* (2009) Interlaboratory and interplatform comparison of microarray gene expression analysis of HepG2 cells exposed to benzo(a)pyrene. *Omics* 13(2):115-125.
85. Kawata K, Shimazaki R, & Okabe S (2009) Comparison of gene expression profiles in HepG2 cells exposed to arsenic, cadmium, nickel, and three model carcinogens for investigating the mechanisms of metal carcinogenesis. *Environ Mol Mutagen* 50(1):46-59.
86. Kirkland D, Kasper P, Muller L, Corvi R, & Speit G (2008) Recommended lists of genotoxic and non-genotoxic chemicals for assessment of the performance of new or improved genotoxicity tests: a follow-up to an ECVAM workshop. *Mutation research* 653(1-2):99-108.
87. Corvi R, *et al.* (2008) ECVAM retrospective validation of in vitro micronucleus test (MNT). *Mutagenesis* 23(4):271-283.
88. IARC (2002) Monographs on the Evaluation of Carcinogenic Risks to Humans Vol.82: Some Traditional Herbal Medicines, Some Mycotoxins, Naphthalene and Styrene.
89. Muller-Tegethoff K, Kasper P, & Muller L (1995) Evaluation studies on the in vitro rat hepatocyte micronucleus assay. *Mutation research* 335(3):293-307.
90. IARC (1998) Monographs on the Evaluation of Carcinogenic Risks to Humans Vol.32: Polynuclear Aromatic Compounds, Part 1, Chemical, Environmental and Experimental Data.
91. Knerr S & Schrenk D (2006) Carcinogenicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin in experimental models. *Mol Nutr Food Res* 50(10):897-907.
92. Ozbay A, Karamperis N, & Jorgensen KA (2007) A review of the immunosuppressive activity of cyclosporine metabolites: new insights into an old issue. *Curr Clin Pharmacol* 2(3):244-248.
93. Mitchell AD, *et al.* (1997) The L5178Y/tk+/- mouse lymphoma specific gene and chromosomal mutation assay a phase III report of the U.S. Environmental Protection Agency Gene-Tox Program. *Mutation research* 394(1-3):177-303.
94. Westerink WM, *et al.* (2009) Evaluation of the Vitotox and RadarScreen assays for the rapid assessment of genotoxicity in the early research phase of drug development. *Mutation research* 676(1-2):113-130.
95. Drawz SM & Bonomo RA (2010) Three decades of beta-lactamase inhibitors. *Clin Microbiol Rev* 23(1):160-201.
96. Daniel KG, Kuhn DJ, Kazi A, & Dou QP (2005) Anti-angiogenic and anti-tumor properties of proteasome inhibitors. *Curr Cancer Drug Targets* 5(7):529-541.
97. Cahill PA, *et al.* (2004) The GreenScreen genotoxicity assay: a screening validation programme. *Mutagenesis* 19(2):105-119.
98. Kirkland D & Speit G (2008) Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens III. Appropriate follow-up testing in vivo. *Mutation research* 654(2):114-132.
99. IARC (1999) Monographs on the Evaluation of Carcinogenic Risks to Humans Vol.73: Some Chemicals that Cause Tumours of the Kidney or Urinary Bladder in Rodents and Some Other Substances.

100. Bischoff SC (2008) Quercetin: potentials in the prevention and therapy of disease. *Curr Opin Clin Nutr Metab Care* 11(6):733-740.
101. Kirkland D, Aardema M, Muller L, & Makoto H (2006) Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens II. Further analysis of mammalian cell results, relative predictivity and tumour profiles. *Mutation research* 608(1):29-42.
102. Nadal A, Alonso-Magdalena P, Ripoll C, & Fuentes E (2005) Disentangling the molecular mechanisms of action of endogenous and environmental estrogens. *Plugers Arch* 449(4):335-343.

Chapter 4

A transcriptomics-based *in vitro* assay for predicting chemical genotoxicity *in vivo*

Carcinogenesis, submitted

C. Magkoufopoulou, S.M.H. Claessen,
M. Tsamou, D.G.J. Jennen,
J.C.S. Kleinjans, J.H.M. van Delft

Abstract

The lack of accurate *in vitro* assays for predicting *in vivo* toxicity of chemicals together with new legislations demanding replacement and reduction of animal testing has triggered the development of alternative methods. The present study aimed at developing a transcriptomics-based *in vitro* prediction assay for *in vivo* genotoxicity. Transcriptomics changes induced in the human liver cell line HepG2 by 34 compounds after treatment for 12h, 24h and 48h were used for the selection of gene-sets that are capable of discriminating between *in vivo* genotoxins (GTX) and *in vivo* non-genotoxins (NGTX). By combining transcriptomics with publicly available results for these chemicals from standard *in vitro* genotoxicity studies, we developed several prediction models. These models were validated by using an additional set of 28 chemicals. The best prediction was achieved after stratification of chemicals according to results from the Ames bacterial gene mutation assay (Ames-test) prior to transcriptomics evaluation after 24h of treatment. 33 genes were selected for discriminating GTX from NGTX for Ames-positive chemicals and 22 for Ames-negative chemicals. Overall, this method resulted in 89% accuracy and 91% specificity, thereby clearly outperforming the standard *in vitro* test battery. Transcription factor network analysis revealed HNF3 α , HNF4 α , HNF6, androgen receptor and SP1 as main factors regulating the expression of classifiers for Ames-positive chemicals. Thus, *in vitro* transcriptomics in HepG2 in combination with the classical bacterial gene mutation assay is proposed as an upgraded *in vitro* approach for predicting *in vivo* genotoxicity of chemicals holding a great promise for reducing animal experimentations on genotoxicity.

Introduction

Although cancer is ranked the second cause of death in USA and EU accounting for approximately 25% of total deaths (1, 2), this disease remains very complicated and is not yet fully understood. However, two main causal factors for its development are appreciated: first, the presence of specific gene mutations genetically inherited or endogenously induced, e.g. BRCA1 and BRCA2 mutations are considered responsible for breast cancer (3); and second, exposure to exogenous carcinogenic factors, such as tobacco smoke that may lead to lung cancer (4). The molecular mechanism of tumor formation after carcinogenic exposure frequently comprises the induction of DNA mutations by the carcinogen or its metabolites (5). Cellular transformation to a tumor cell may also be promoted through a variety of mechanisms which do not necessarily involve DNA damage. Consequently, carcinogens are classified as genotoxic (GTX) or non-genotoxic (NGTX) (5). Since almost all GTX compounds are carcinogenic, it is important, in particular for regulatory purposes, to evaluate the genotoxic potential of chemicals to which humans are exposed, and thereby to discriminate GTX from NGTX compounds.

The most commonly used assay for evaluating genotoxic properties of chemicals *in vitro* is the bacterial gene mutation assay (Ames-test) (6). Several *in vitro* genotoxicity assays using mammalian cellular models have also been developed, namely the mammalian micronuclei (MN), chromosomal aberration (CA) and mouse lymphoma assays (MLA). However, the conventional *in vitro* test battery often fails to correctly predict the *in vivo* genotoxic and carcinogenic potential of compounds, reaching a 50% false positive rate (7). Especially for chemicals that are genotoxic *in vitro*, regulatory authorities may require the *in vivo* evaluation of genotoxic properties in rodents (EC 1907/2006). Consequently the high false positive rate of *in vitro* genotoxicity assays results in a relatively high number of unnecessary animal experiments. Further, the use of experimental animals also inflicts considerable costs and raises ethical issues. Thus, a more reliable *in vitro* assay for predicting *in vivo* genotoxicity is urgently required. Currently, many efforts are internationally made for the replacement, reduction and refinement (3Rs) of animal studies (8). The socioeconomic necessity to reduce animal experiments inspired the search for alternative, more accurate *in vitro* methods thereby exploring novel technological approaches. In this context, toxicogenomics-based alternative approaches are also considered (9). Based on the hypothesis that compounds of the same toxicity class cause similar effects on gene expression, toxicogenomics-based approaches have been explored for the development of classification tools, including genotoxicity showing that toxicogenomics methods are indeed capable of discriminating GTX from NGTX compounds (10-13).

In this respect, the aim of the present study was to improve prediction of *in vivo* genotoxicity using an *in vitro* transcriptomics-based method in HepG2 cells. The human hepatic cell line HepG2 was selected as it expresses many drug metabolizing enzymes that their expression appears to be regulated by toxins (14-18); lacks mutations in p53, a major factor in DNA damage response (19) and has been successfully applied in the

phenotypic detection of genotoxic effects of chemicals that require metabolic activation (20-22), in predictive transcriptomics studies (23) and in high through-put screening tools for genotoxicity (24). Lastly, HepG2-based assays are also used by the EPA ToxCast™ project that aims to develop a cost-effective approach for efficiently prioritizing toxicity testing of chemicals (25).

Here, gene expression alterations following HepG2 treatments for 12h, 24h and 48h with 34 chemicals (12 GTX and 22 NGTX) were compared to select gene signatures that can be used for the prediction of *in vivo* genotoxicity. Further, we investigated whether a *priori* stratification of the compounds based on results of various standard genotoxicity tests improves the transcriptomics-based prediction. Validation was thereupon conducted using an additional set of 28 chemicals, 11 GTX and 17 NGTX. For the biological interpretation of the acquired gene signature, functional annotation and transcriptional regulation network analyses of the selected classifiers were performed.

Materials and methods

Chemicals

Table 1 shows the doses for the 34 chemicals used for the development of the prediction model (A) and for the 28 chemicals used for the validation (B) and provides information on the stratification of the compounds based on the Ames mutagenicity assay, *in vitro* and *in vivo* genotoxicity. More detailed information on the results of the conventional genotoxicity assays can be found in Supplementary Data 1. A chemical is considered *in vitro* genotoxic when at least one *in vitro* genotoxicity assay (MN, CA, MLA) showed clear positive results, since these three assays are currently considered equally appropriate and therefore interchangeable when used together with other genotoxicity tests in a standard battery for testing of pharmaceuticals (26); a chemical is defined genotoxic *in vivo* when at least one *in vivo* genotoxicity assay (MN, CA) showed positive results; since negative results in both the previous mentioned *in vivo* assays, with adequate justification, are sufficient to demonstrate absence of genotoxicity (26).

Cell culture and treatment

HepG2 cells were cultured in 6-well plates as previously described (23). When the cells were 80% confluent, medium was replaced with fresh medium containing the corresponding dose of each compound or with the corresponding control treatment (DMSO, EtOH, or PBS 0.5%) (Table 1).

All doses were selected based on a MTT assay resulting to 80% viability at 72h incubation or a maximum dose of 2mM when no cytotoxicity was observed (23). Cells were exposed for 12h, 24h or 48h for the training set of chemicals and for 24h and 48h for the validation set. Thereafter the culture medium was replaced by TRIZOL (Gibco/BRL) for RNA isolation. Three independent biological replicates were conducted.

Table 1: Chemicals used in this study, selected doses and information on *in vitro* and *in vivo* genotoxicity data

A. Training set of chemicals							
Compound	Abbreviation	CAS no.	Dose	Solvent	Ames ^a	Genotoxicity	
						in vitro ^b	in vivo ^c
2-acetyl aminofluorene	2AAF	53-96-3	50 µM	DMSO	+	+	+
Aflatoxin B1	AFB1	1162-65-8	1 µM	DMSO	+	+	+
Benzo[a]pyrene	BaP	50-32-8	2 µM	DMSO	+	+	+
7,12-dimethylbenz [a]anthracene	DMBA	57-97-6	5 µM	DMSO	+	+	+
Dimethyl nitrosamine	DMN	62-75-9	2 mM	DMSO	+	+	+
Mitomycine C	MMC	50-07-7	200 nM	DMSO	+	+	+
Para-cresidine	pCres	120-71-8	2 mM	EtOH	+	+	+
2-(chloromethyl) pyridine.HCl	2CMP	6959-47-3	300 µM	DMSO	+	+	-
4-acetyl aminofluorene	4AAF	28322-02-3	100 nM	DMSO	+	+	-
4-Nitro-o- phenylenediamine	NPD	99-56-9	2 mM	DMSO	+	+	-
8-quinolinol	8HQ	148-24-3	15 µM	DMSO	+	+	-
Quercetin	Que	117-39-5	50 µM	DMSO	+	+	-
Phenobarbital	PhB	50-06-6	1 mM	DMSO	+	+	-
Acetaminophen	APAP	103-90-2	100 µM	PBS	-	+	+
Diethylstilbestrol	DES	56-53-1	5 µM	EtOH	-	+	+
Phenol	Ph	108-95-2	2 mM	DMSO	-	+	+
Tributyltin oxide	TBTO	56-35-9	0.02 nM	EtOH	-	+	+
Curcumin	Cur	458-37-7	1 µM	DMSO	-	+	-
o-anthranilic acid	AnAc	118-92-3	2 mM	DMSO	-	+	-
Resorcinol	RR	108-46-3	2 mM	EtOH	-	+	-
Sulfisoxazole	Sulfi	127-69-5	5 µM	DMSO	-	+	-
17-beta Estradiol	E2	50-28-2	30 µM	DMSO	-	+	-
Ethylacrylate	EtAc	140-88-5	1 mM	EtOH	-	+	-
Phenacetin	Phen	62-44-2	1 mM	EtOH	-	+	-
L-ascorbic acid	VitC	50-81-7	2 mM	PBS	-	-	+
Ampicillin trihydrate	AmpC	7177-48-2	250 µM	DMSO	-	-	-
Diclofenac	Diclo	15307-86-5	100 µM	PBS	-	-	-
D-mannitol	Dman	69-65-8	250 µM	PBS	-	-	-
Cyclosporine A	CsA	59865-13-3	3 µM	DMSO	-	-	-
di(2-ethylhexyl) phthalate	DEHP	117-81-7	10 mM	DMSO	-	-	-
Reserpine	Res	50-55-5	12.5 µM	DMSO	-	-	-
2,3,7,8-tetrachloro dibenzo-p-dioxin	TCDD	1746-01-6	10 nM	DMSO	-	-	-
Tetradecanoyl phorbol acetate	TPA	16561-29-8	500 nM	DMSO	-	-	-
Wy 14643	Wy	50892-23-4	200 µM	DMSO	-	-	-

Table 1 (continued)

B. Validation Set							
Compound	Abbreviation	CAS no.	Dose	Solvent	Ames ^a	Genotoxicity	
						<i>in vitro</i> ^b	<i>in vivo</i> ^c
4-aminobiphenyl	ABP	92-67-1	80 µM	DMSO	+	+	+
Azathioprine	AZA	446-86-6	250 µM	DMSO	+	+	+
Benzidine	BZ	92-87-5	1 mM	DMSO	+	+	+
Chlorambucil	Cb	305-03-3	20 µM	DMSO	+	+	+
Cisplatin	cisPt	15663-27-1	20 µM	PBS	+	+	+
Cyclophosphamide	CP	6055-19-2	2 mM	PBS	+	+	+
Diethylnitrosamine	DEN	55-18-5	500 µM	DMSO	+	+	+
1-ethyl-1-nitrosourea	ENU	759-73-9	1 mM	DMSO	+	+	+
Furan	Fu	110-00-9	2 mM	DMSO	+	+	+
2-amino-3-methylimidazo (4,5-f) quinoline	IQ	76180-96-6	800 µM	DMSO	+	+	+
4,4'-methylenebis (2-chloroaniline)	MOCA	101-14-4	60 µM	DMSO	+	+	+
2-chloroethanol	2-Cl	107-07-3	2 mM	DMSO	+	+	-
p-anisidine	Anis	104-94-9	60 µM	DMSO	+	+	-
Bromodichloro- methane	BDCM	75-27-4	2 mM	DMSO	-	+	-
Carbon tetrachloride	CCl ₄	56-23-5	2 mM	DMSO	-	+	-
Ethylbenzene	EthylB	100-41-4	800 µM	DMSO	-	+	-
Eugenol	EuG	97-53-0	500 µM	DMSO	-	+	-
Nitrobenzene	NBZ	98-95-3	2 mM	DMSO	-	-	-
1,1,1-trichloro-2,2-di- (4-chlorophenyl) ethane	DDT	50-29-3	80 µM	DMSO	-	-	-
Pentachlorophenol	PCP	87-86-5	10 µM	EtOH	-	-	-
Progesterone	Prog	57-83-0	6 µM	EtOH	-	-	-
Tetrachloroethylene	TCE	127-18-4	2 mM	EtOH	-	-	-
Lindane	γ-HCH	58-89-9	2 mM	DMSO	-	-	-
Acesulfame-K	ASK	55589-62-3	2 mM	DMSO	-	-	-
Caprolactam	CAP	105-60-2	2 mM	DMSO	-	-	-
Coumaphos	COU	56-72-4	250 µM	DMSO	-	-	-
Diazinon	DZN	333-41-5	250 µM	DMSO	-	-	-
Simazine	Sim	122-34-9	50 µM	DMSO	-	-	-

^aAmes results based on NTP data;^b*in vitro* genotoxicity is considered positive when at least one *in vitro* genotoxicity assay (Ames, MN, CA, MLA) showed positive results; ^c*in vivo* genotoxicity is considered positive when at least one *in vivo* genotoxicity assay (MN, CA) showed positive results. Equivocal *in vivo* data are considered negative. (References for *in vitro* and *in vivo* genotoxicity data are available in supplementary Data1)

Total RNA isolation and microarray experiments

Total RNA was extracted using 0.5 ml TRIzol according to the manufacturer's instructions and purified using RNeasy® Mini Kits (Qiagen). Sample preparation, hybridization, washing, staining and scanning of the Affymetrix Human Genome U133 Plus 2.0 GeneChip arrays were conducted according to the manufacturer's manual as previously described (23). Quality controls were within acceptable limits. Hybridization controls were called present on all arrays and yielded the expected increases in intensities. The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (27) and are accessible through GEO Series accession number GSE28878. (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE28878>).

Annotation and normalization of microarray data

The obtained data sets were re-annotated to the MBNI Custom CDF-files (http://brainarray.mbni.med.umich.edu/Brainarray/Database/CustomCDF/genomic_curated_CDF.asp) (28) and RMA normalized (29) using the NuGOExpressionFileCreator in GenePattern (30). Log2 ratios were calculated for each replicate to the corresponding control treatment. For the training set of chemicals, data filtering was conducted for each time point separately as previously described (23). Only the filtered data sets were further used.

Selection of classifiers for genotoxicity (Figure 1)

Method 1: Genes with significantly different expression values ($p < 0.01$) between *in vivo* GTX chemicals and *in vivo* NGTX chemicals were selected from the filtered data sets based on a series of 34 t-tests. For each t-test the three replicates of one of the 34 chemicals were removed (leave-one-out procedure). The significant genes that were present in all t-tests (the intersection of all lists) were selected as classifiers.

Method 2: The 34 chemicals were stratified into two groups based on the results of the Ames mutagenicity assay (Table 1) and consequently assigned to Ames-positive and Ames-negative. Within each group both *in vivo* GTX and *in vivo* NGTX chemicals are present. For the Ames-positive group, 13 t-tests were performed to select classifiers for discriminating *in vivo* GTX compounds from *in vivo* NGTX compounds, as mentioned for Method 1. For the Ames-negative group 21 t-tests were performed to select classifiers for discriminating *in vivo* GTX from *in vivo* NGTX chemicals.

Method 3: The 34 chemicals were stratified into two groups based on the results of the *in vitro* genotoxicity assays (Table 1) into *in vitro* GTX and *in vitro* NGTX compounds. For the *in vitro* GTX group, 24 t-tests were performed to select classifiers that can discriminate *in vivo* GTX from *in vivo* NGTX compounds, as mentioned for Method 1. For the *in vitro* NGTX group no t-tests were performed due to the presence of only one *in vivo* GTX compound.

Figure 1: Developed methods for genotoxicity prediction. Schematic representation of the Methods used for predicting *in vivo* genotoxicity. No stratification of the compounds was used for Method 1. For Method 2 the compounds were stratified based on the results of the Ames Assay and for Method 3 based on the *in vitro* GTX assessment. The number of chemicals, for the training and validation sets, within each group is indicated. The number of classifiers resulting from the T-tests between *in vivo* GTX and *in vivo* NGTX compounds is indicated.

Class prediction of the training set of chemicals

PAM (Prediction Analysis for Microarrays 2.13, Stanford University Software) analysis (31) was conducted for each of the selected lists of classifiers for class prediction (threshold: 0). The gene expression data of the three replicates of all compounds except one were used as training set (leave-one-out approach) to generate a predictive model, whereafter that model was tested with the left out compound. The leave-one-out approach was conducted for all the compounds. A compound was predicted to be *in vivo* GTX or *in vivo* NGTX when at least two out of the three replicates were assigned to one class.

The accuracy was calculated as the percentage of the correctly classified chemicals to the total number of tested chemicals; the sensitivity as the percentage of the correctly classified GTX to the total number of tested GTX compounds and the specificity as the percentage of the correctly classified NGTX to the total number of tested NGTX compounds. For method 2, the prediction results of all the 34 compounds from both sub-groups were taken into consideration for calculating accuracy, sensitivity and specificity. For method 3 only the results of the *in vitro* GTX group (24 compounds) were used for calculating accuracy, sensitivity and specificity.

Validation of prediction methods

For the validation of the prediction methods, a set of 28 chemicals was used (Table 1B). PAM analysis was conducted using as training set the gene expression values (log2 ratios) of the selected classifiers after treatment with the 34 chemicals used for the development of the methods (threshold: 0).

Selection of Differentially Expressed Genes

Differentially expressed genes (DEGs) per compound were selected using the criteria: absolute average fold change > 1.5 in comparison to control, absolute fold change > 1.2 for at least two replicates and same direction of expression for all three replicates, as previously described (23).

Functional annotation and transcription regulation analyses

Functional Annotation based on Gene Ontology (GO) Biological Processes was done

using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (32). For each gene list, a background list resulting from the data filtering was used. GO terms were considered significant when $EASE < 0.1$ and $p\text{-value} < 0.05$ and included more than four genes. Identification of the transcription factors regulating the expression of the selected classifiers was conducted in MetaCore Network Tools (GeneGo, San Diego, CA) using the Transcription Regulation Network option. Transcription regulation networks were taken into account when at least three of the classifiers were included in the network.

Results

Genotoxicity prediction

Three methods were applied for the selection of gene signatures with the ability to predict genotoxicity *in vivo* (Figure 1) by using 34 chemicals (Table 1A). The methods were further validated by predicting genotoxicity for an additional set of 28 chemicals (Table 1B). The accuracy, sensitivity and specificity of all methods are presented in Table 2A for the training set of chemicals, Table 2B for the validation set and Table 2C for all chemicals. The selected classifiers from all methods and the predicted classes for each of the chemicals after the application of each method, for both the training and the validation set, are available in Supplementary Data 2.

Class prediction based on the transcriptomics data only (Method 1) resulted in approximately 80% accuracy for all exposure durations, with the highest sensitivity (67%) achieved at 48h and the highest specificity (95%) obtained at 24h (Table 2A).

Thereupon, we investigated whether stratifying the compounds based on conventional *in vitro* genotoxicity tests prior to the selection of genes improves the transcriptomics-based prediction.

In Method 2, classifiers were selected upon stratification of the chemicals solely based on the results of the bacterial mutagenicity assay, as it has the highest specificity among the traditional *in vitro* genotoxicity assays (7). With these genes, a clear improvement in comparison to Method 1 with respect to the accuracy and the sensitivity of class prediction was observed in particular at 24h of exposure (92%).

Method 3 applied *a priori* stratification of chemicals based on the integrated outcome of the established *in vitro* genotoxicity test battery. Since hardly any chemicals exist which are NGTX *in vitro* but GTX *in vivo* (only one compound in our study), this analysis had to be restricted only to compounds which are positive in *in vitro* genotoxicity testing. The use of the classifiers selected after this stratification resulted in approximately 80% accuracy for all exposure durations, with high sensitivity at both 24h and 48h of treatment (82%) whereas the highest specificity occurred at 24h (85%).

We, thus, concluded that exposure periods of 24h and 48h generate the most accurate prediction models. Therefore, HepG2 cells were exposed for 24h and 48h to an additional set of 28 chemicals (Table 1B) to validate the capability of all methods to predict *in vivo* genotoxicity. This validation clearly shows that the highest accuracy (91%) is achieved after 24h of exposure upon stratification of the chemicals based on the results of the Ames assay. Table 3 shows the class prediction results for all chemicals using Method 2 after 24h of treat-

Table 2: Results of genotoxicity prediction methods

A. Training Set (34 chemicals)									
	Method 1^a			Method 2^b			Method 3^c		
	12h	24h	48h	12h	24h	48h	12h	24h	48h
Accuracy	76%	79%	79%	82%	91%	79%	75%	83%	79%
Sensitivity	58%	50%	67%	83%	92%	83%	73%	82%	82%
False negative rate	42%	50%	33%	17%	8%	17%	27%	18%	18%
Specificity	86%	95%	86%	82%	91%	77%	77%	85%	77%
False positive rate	14%	5%	14%	18%	9%	23%	23%	15%	23%
B. Validation Set (28 chemicals)									
	Method 1^a			Method 2^b			Method 3^c		
	12h	24h	48h	12h	24h	48h	12h	24h	48h
Accuracy	-	57%	61%	-	86%	75%	-	41%	41%
Sensitivity	-	36%	36%	-	91%	73%	-	46%	64%
False negative rate	-	64%	64%	-	9%	27%	-	54%	36%
Specificity	-	71%	76%	-	82%	82%	-	33%	0%
False positive rate	-	29%	24%	-	18%	18%	-	67%	100%
C. All 62 chemicals									
	Method 1^a			Method 2^b			Method 3^c		
	12h	24h	48h	12h	24h	48h	12h	24h	48h
Accuracy	-	69%	71%	-	89%	79%	-	66%	63%
Sensitivity	-	43%	52%	-	91%	78%	-	64%	73%
False negative rate	-	57%	48%	-	9%	22%	-	36%	27%
Specificity	-	85%	82%	-	87%	79%	-	68%	53%
False positive rate	-	15%	18%	-	13%	21%	-	32%	47%

^aPrediction based on transcriptomics data only. ^bPrediction based on transcriptomics data after stratification of chemicals to Ames-positive and Ames-negative. ^cPrediction based on transcriptomics data after stratification of chemicals to GTX *in vitro* and NGTX *in vitro* based on the results of Ames-test, Micronuclei and Chromosomal Aberrations *in vitro* assays.

ment. A comparison of the results obtained by transcriptomics-based prediction after 24h of exposure by Method 2 to the outcomes of the conventional *in vitro* assays for the chemicals used in the present study is presented in Table 4.

The developed method resulted in the misclassification of the Ames-negative chemicals Sim, RR and APAP and of the Ames-positive chemicals, PhB, 2-Cl, Anis and AZA. To further understand possible causes of these misclassifications, differentially expressed genes (DEGs) after treatment with these compounds were selected and used for functional annotation analysis (Supplementary Data 3). The functional annotation analysis for the DEGs after Sim-treatments revealed four genes involved in DNA repair process. Regarding RR, no DEGs were involved in any DNA damage-related processes; however several genes appeared involved in metabolic processes, transcription, cell death/apoptosis and cell cycle-related processes. APAP-treatment resulted in only one gene differentially expressed (RRN3P2), thus no functional annotation analysis was possible. None of the misclassified NGTX Ames-positive compounds (2Cl, PhB and Anis) resulted in expression alterations of genes involved in DNA-damage related processes. Although DEGs resulting from PhB treatments were not significantly involved in any biological process, 2Cl and Anis affected the expression of several genes involved mainly in metabolic processes. Conversely, AZA, a misclassified GTX Ames-positive compound, altered the expression of more than 2500 genes, from which 28 are involved in double-strand-break repair, 23 in nucleotide excision repair and 11 in base excision repair.

Table 3: Class prediction results using the classifiers selected by Method 2 after 24h of treatment

Training Set		Validation Set	
Compound	Prediction	Compound	Prediction
2-CMP	NGTX	ABP	GTX
8HQ	NGTX*	AZA	<u>NGTX</u>
NPD	NGTX	BZ	GTX*
ANAC	NGTX	Cb	GTX
RR	<u>GTX</u>	cisPt	GTX
4-AAF	NGTX*	CP	GTX*
Cur	NGTX	DEN	GTX*
Que	NGTX	ENU	GTX
d-man	NGTX	Fu	GTX*
diclo	NGTX	IQ	GTX
Sulfi	NGTX	MOCA	GTX
AmpC	NGTX	2-Cl	<u>GTX*</u>
TBTO	GTX	Anis	<u>GTX</u>
VitC	GTX	BDCM	NGTX
Ph	GTX*	CCl4	NGTX*
Phen	NGTX	EthylB	NGTX*
E2	NGTX	EuG	NGTX
EtAc	NGTX*	NBZ	NGTX*
PhB	<u>GTX*</u>	DDT	NGTX*
DEHP	NGTX	PCP	NGTX
Res	NGTX	Prog	NGTX
TCDD	NGTX	TCE	NGTX
CsA	NGTX	γ-HCH	NGTX
TPA	NGTX	ASK	NGTX
Wy	NGTX	CAP	NGTX
DMN	GTX*	COU	NGTX
2-AAF	GTX*	DZN	NGTX
DMBA	GTX	Sim	<u>GTX*</u>
BaP	GTX		
MMC	GTX		
DES	GTX*		
AFB1	GTX		
APAP	<u>NGTX</u>		
pCres	GTX		

GTX: the compound is predicted genotoxic; NGTX: the compound is predicted non-genotoxic; Results indicated with bold and underlined letters indicate misclassification, results labeled * indicate that two of the three replicates were classified in the indicated class.

Table 4: Comparison of the performance of Ames test, MLA, MN/CA and the combination of Ames test with other *in vitro* genotoxicity assays for predicting *in vivo* genotoxicity with prediction based on gene expression profiles of 24h exposure after stratification of chemicals based on Ames results^a

	Ames ^b	Ames ⁺ GE ^b	Ames ⁺ MLA ^c	Ames ⁺ MLA ^d	Ames ⁺ + GE ^{c,d}	Ames ⁺ MN/CA ^e	Ames ⁺ MN/CA ^f	Ames ⁺ + GE ^{e,f}	Ames ⁺ MLA/MN/CA ^g	Ames ⁺ + GE ^g
Accuracy	77%	89%	60%	60%	91%	63%	62%	88%	68%	89%
Sensitivity	78%	91%	94%	94%	100%	96%	96%	91%	96%	91%
False negative rate	22%	9%	6%	6%	0%	4%	4%	9%	4%	9%
Specificity	77%	87%	42%	42%	97%	46%	40%	86%	51%	87%
False Positive rate	23%	13%	58%	58%	3%	54%	60%	14%	49%	13%

MLA: Mouse lymphoma assay, **GE:** Gene expression, **MN:** Micronuclei Assay, **CA:** Chromosomal Aberrations ^a: the chemicals from both the training and the validation set were used; ^b: based on 62 compounds with available Ames results; ^c: based on 47 compounds with available MLA results; ^d: based on 47 compounds with available Ames and MLA results; ^e: based on 60 compounds with available MN or CA (or both results); ^f: based on 60 compounds with available Ames and MN or CA (or both) results; ^g: based on 62 compounds with data in at least one of the four *in vitro* assays.

Table 5: Classifiers for genotoxicity prediction of chemicals resulted by Method 2 after 24h of treatment & Ames stratification

A. For Ames positive chemicals				
EntrezGene ID	Gene Symbol	Function^a	GTX^b	NGTX^c
634	CEACAM1	Receptor ligand	0,428	-0,069
1183	CLCN4	Voltage-gated ion-channel	0,073	-0,101
2009	EML1	Generic phosphatase	-0,101	-0,264
170394	PWWP2B	Generic protein	0,026	-0,169
7325	UBE2E2	Generic enzyme	-0,243	0,255
8975	USP13	Generic protease	0,212	-0,120
9535	GMFG	Generic binding protein	0,181	-0,206
11212	PROSC	Generic protein	-0,179	0,090
7276	TTR	Generic binding protein	0,241	-0,734
8431	NR0B2	Transcription factor	0,438	-0,636
9027	NAT8	Generic enzyme	0,283	-0,817
11030	RBPMS	Generic binding protein	0,147	-0,246
11067	C10orf10	Generic protein	0,667	-0,596
23158	TBC1D9	Generic protein	0,016	-0,342
29916	SNX11	Generic binding protein	-0,189	0,074
54538	ROBO4	Generic receptor	0,221	-0,013
54880	BCOR	Generic binding protein	0,029	-0,312
6092	ROBO2	Generic receptor	0,026	0,136
6725	SRMS	Protein kinase	0,025	-0,180
26539	OR10H1	GPCR	-0,067	0,158
27010	TPK1	Generic kinase	-0,096	0,096
64115	C10orf54	Generic receptor	0,116	-0,035
319103	SNORD8	RNA	-0,182	0,161
414918	FAM116B	Generic protein	0,141	-0,010
100131914	LOC100131914	-	0,250	-0,359
55937	APOM	Transporter	0,108	-0,434
56675	NRIP3	Generic binding protein	-0,031	0,124
57095	C1orf128	Generic protein	-0,137	0,368
57099	AVEN	Generic binding protein	-0,069	0,365
60677	BRUNOL6	Generic binding protein	-0,066	0,238
84133	ZNRF3	Generic binding protein	-0,003	-0,634
146227	BEAN	Generic binding protein	0,000	0,238
376497	SLC27A1	Generic enzyme	0,091	-0,165
B. For Ames negative chemicals				
EntrezGene ID	Gene Symbol	Function^a	GTX^b	NGTX^c
309	ANXA6	Generic binding protein	0,310	0,021
337	APOA4	Receptor ligand	0,111	-0,111
686	BTD	Generic enzyme	0,176	-0,102
1939	LGTN	Generic receptor	0,124	-0,069
3267	AGFG1	Generic binding protein	-0,169	-0,003
4705	NDUFA10	Generic enzyme	0,146	-0,070
4775	NFATC3	Transcription factor	0,297	0,021
9373	PLAA	Generic binding protein	-0,138	0,024
22909	MTMR15	Generic binding protein	0,208	-0,057
30061	SLC40A1	Transporter	0,585	0,073
51433	ANAPC5	Generic enzyme	0,123	-0,070
64969	MRPS5	Generic binding protein	0,170	-0,001
79970	ZNF767	Generic protein	0,267	-0,070
84968	PNMA6A	Generic protein	0,467	0,034
373156	GSTK1	Generic enzyme	0,151	-0,080
112817	DHDPSL	Generic enzyme	0,305	-0,014
2243	FGA	Generic binding protein	0,271	-0,312
6446	SGK1	Protein kinase	0,502	-0,107
6532	SLC6A4	Transporter	0,270	-0,163
90507	SCRN2	Generic protease	0,217	-0,136
200014	CC2D1B	Generic protein	-0,096	0,129
648921	LOC648921	-	-0,182	0,086

^aInformation on the function of the classifiers was retrieved from MetaCore; Average gene expression (log2 ratios) for ^b*in vivo* GTX treatments and ^c*in vivo* NGTX treatments.

Functional annotation and transcriptional regulation network analyses of classifiers

The classifiers selected after 24h of treatment by Method 2 are shown in Table 5. Table 6 includes the biological processes that are significantly represented by these genes (6A) and the transcription factors which regulate their expression (6B).

The functional annotation analysis for the 33 classifiers for Ames-positive compounds did not reveal any Gene Ontology (GO) terms. However five transcription factors, HNF4 α , SP1, HNF3 α , Androgen Receptor (AR) and HNF6 were identified, which appear to regulate the expression of some of these classifiers (5, 4, 2, 2 and 3 respectively) (Table 6B) and which are involved in response to stimulus, metabolic and developmental processes. NR0B2 (SHP), one of the classifiers, appears to negatively regulate the function of four of these transcription factors. A network showing the direct interactions among these transcription factors and the identified classifiers is presented in Figure 2.

Out of the 22 classifiers for Ames-negative compounds, 8 are involved in GO-terms for transport and localization and 6 in response to stress. Most of these genes showed higher expression levels for GTX than for NGTX compounds. Within this set of 22 genes, no significant transcriptional regulation networks were identified.

Table 6: Biological Interpretation of selected classifiers for genotoxicity by Method 2 after 24h of treatment

A. Functional annotation analysis				
	Gene ontology terms ^a		Classifiers ^b	
Ames positive compounds	-	-		
Ames negative compounds	transport / localization	AGFG1, NDUFA10, ANXA6, APOA4, LGTN, SGK1, SLC40A1, SLC6A4		
	response to stress	APOA4, FGA, MTMR15, NFATC3, PLAA, SGK1		
B. Transcriptional regulation network analysis				
	Transcription factor ^c	Gene ontology terms ^d	p-value ^e	Classifiers ^f
Ames positive compounds	HNF4α	response to glucose / hexose stimulus, regulation to peptide secretion/transport	5,09e-20	TTR, CEACAM1, NAT8, SNX11, C10orf10 <i>NR0B2</i>
	SP1	specification of axis polarity, phosphatidylserine biosynthetic process	1,10e-16	ROBO4, GMFG, SLC27A1, TPK1 <i>BCOR</i>
	HNF3α	thiamin diphosphate metabolic process, hormone metabolic process	4,10e-10	TTR, TPK1 <i>NR0B2</i>
	Androgen receptor	androgen receptor signaling pathway, steroid hormone receptor signaling pathway, specification of axis polarity	4,10e-10	BCOR, RBPMS <i>NR0B2</i>
	HNF6	alcohol metabolic process, thyroid hormone metabolic process, spleen/pancreas development	4,10e-10	TTR, CLCN4, NR0B2 <i>NR0B2</i>
Ames negative compounds	-	-	-	-

^a: Gene ontology terms significantly represented in the total classifiers list; ^b: classifiers involved in the indicated Gene ontology term; ^c: Significant transcription factor for the regulation of the expression of the classifiers; ^d: Gene ontology terms in which the transcription network is involved; ^e: p-value referring to the significance of the developed network; ^f: classifiers that their expression is affected by the indicated transcription factor, classifiers indicated with italic letters regulate the function of the transcription factors.

Figure 2: Direct interactions network for transcription factors and their targets among the classifiers of Method 2 for Ames-positive compounds. The type of interaction between the genes is indicated by the different colors (green: activation, red: inhibition, grey: unspecified). The circles next to the symbols indicate the expression of the genes for GTX and NGTX compounds (average log2 ratios). Red circles indicate up-regulation and blue circles down-regulation after both GTX and NGTX treatments; the “checkerboard” color indicates mixed expression between GTX and NGTX compounds.

Discussion

In this study we aimed at developing an alternative *in vitro* transcriptomics-based method applying HepG2 cells for predicting *in vivo* genotoxic properties of chemicals. The best accuracy was achieved after exposure for 24h and when the compounds were initially stratified based on the results of the Ames assay (Method 2).

This novel approach for the prediction of *in vivo* genotoxicity results in an improved accuracy when compared to the conventional *in vitro* genotoxicity assays (Table 4). Consequently, for this series of compounds, our proposed method results in a clear improvement of the specificity and thus a clear reduction of the false positive results (13%) in comparison to the outcome of all *in vitro* assays (Table 4). This false positive rate is based on the misclassification of 5 NGTX compounds, namely Sim, PhB, RR, 2CI and Anis.

Functional annotation analysis revealed genes associated with DNA-damage responses only after Sim-treatment, suggesting that DNA-damage may have been induced. Sim has given negative results in *in vivo* MN studies (33, 34), whereas in a SCGE assay Sim showed weak signs of genotoxicity (35). Since these studies were performed for bone marrow or blood cells but not for other target tissues, it cannot be excluded that Sim may have genotoxic effects to other cell types, and the genotoxic effect of Sim to liver cells may be represented in our assay.

Similarly, PhB appeared NGTX in the *in vivo* MN assay, but positive results were obtained for the SCGE assay for short exposures (36), showing that the genotoxic effects of PhB *in vivo* are not fully elucidated yet. PhB is known to induce CAR transcriptional activity leading to induction of the expression of CYP2B (37, 38). In our study, PhB resulted in limited transcriptomics alterations in HepG2, which did not include the overexpression of CYP2B enzymes, suggesting that PhB did not cause the activation of CAR in HepG2 cells. This may have led to the activation of other mechanisms that resulted to the classification of PhB as GTX.

RR-treatments altered the gene expression levels in HepG2 of more than 2000 genes, several of which appeared to be involved in apoptosis/cell death related processes. We hypothesize that this massive transcriptomics response of HepG2 cells to RR is due to the applied high dose (2mM). Despite that this dose did not induce cytotoxicity as measured by the MTT test; it resulted in a massive deregulation of the expression of genes associated to apoptosis/cell death and, thus, indirectly to its classification as GTX.

Functional annotation analysis for the DEGs after 2CI and Anis treatments mainly included metabolic processes. *In vivo*, 2CI has been tested in mice by applying the MN, CA and SCE assays showing negative results, whereas in a CA study in rats 2CI was found to be genotoxic (39). Anis was only tested in the micronucleus assay resulting in a negative outcome (Table 1, Supplementary Data 1), indicating that the available data may not have been sufficient for the assignment of these two chemicals in the non-genotoxic class possible explaining their classification by the transcriptomics-based method as GTX.

Furthermore, our method results in a lower false negative rate (9%) than the Ames mutagenicity test (22%) and in similar false negative rate as the other classical *in vitro* genotoxicity assays and combinations thereof (Table 4). This false negative rate refers to the misclassification of only two chemicals, namely AZA and APAP. Although the IARC evaluation for both AZA and APAP presents positive data for *in vivo* genotoxicity testing (33, 40), their ability to induce genotoxic and carcinogenic effects in humans is doubtful (41, 42). Thus, the assignment of these compounds to the GTX class may have been incorrect, and their prediction as NGTX by our method may better represent their effects to humans. AZA altered the expression of more than 2500 genes, several of which were involved in DNA-repair processes. Functionally, this suggests that AZA may indeed be genotoxic to HepG2 cells; nevertheless, its genotoxic activity was not detected by the selected classifiers. With respect to APAP, it has been reported that this compound induces genotoxicity *in vitro* and *in vivo* only at high, non-physiological doses (42, 43). In addition APAP-induced genotoxicity *in vivo* has been attributed to the induced cytotoxicity at high dose (43). In our study, treatment with 100 μ M APAP, corresponding to 20% cytotoxicity, resulted in only one gene differentially expressed. Based on the small transcriptomics responses, APAP was classified as NGTX. These observations suggest that the selected dose was not sufficient for APAP to exert its cytotoxic and genotoxic effect which is in concordance with the studies that link its genotoxicity to high cytotoxicity.

For the biological understanding of the selected classifiers obtained by Method 2 upon 24h of exposure, functional annotation and transcriptional regulation analyses were per-

formed (Table 5). AR, SP1, HNF3 α , HNF4 α and HNF6 were identified as the transcription factors that regulate the expression of the classifiers for the Ames-positive compounds. AR is a ligand-dependent transcription factor and its hyper-activation is linked to the development of prostate cancer (44). For SP1, evidence exists for its involvement in DNA damage responses (45). HNF3 α , HNF4 α and HNF6 regulate the expression of liver-specific genes (46). In particular HNF4 α is responsible for the regulation of expression of genes involved, among others, in drug and xenobiotic metabolism (47). Furthermore, NR0B2, one of the selected classifiers, was identified by the transcriptional regulation network analysis to negatively regulate the function of HNF3 α , HNF4 α , HNF6 and SP1 (Figure 2). NR0B2 is a known inhibitor of transcription and has been proposed as a tumor suppressor gene for hepatocellular carcinoma (48). The increased expression levels of NR0B2 after GTX treatments in this study suggest a possible role of this gene in the development of chemically induced genotoxicity by compounds that cause mutagenicity to both bacterial and mammalian cells. These functionalities thus subscribe to the biological plausibility of the identified gene signature for predicting genotoxicity *in vivo*.

For the classifiers that predict genotoxicity *in vivo* for Ames-negative compounds, no significant transcriptional regulation networks were identified. Overrepresentation analyses based on GO-terms, however, indicated enrichment of classifiers in cellular stress responses. Among these, SGK1 was identified by PAM as the gene with the highest contribution to the prediction model. SGK1 is a kinase that plays an important role in stress response and cell survival and has been proposed as an oncogene related to breast cancer development (49, 50). The high expression values of SGK1 after GTX treatments in this study suggest a possible involvement of this gene in genotoxicity induction by compounds that cause mutagenicity to mammalian but not to bacterial cells.

The approach for *in vivo* genotoxicity prediction developed here may be used in several applications in order to avoid unnecessary experiments on animals. For instance, it may serve the purposes of the EU chemical policy program REACH, for which it has been estimated that approximately 400000 rodents may be used for testing genotoxicity *in vivo* (51), many of which will be based on false premises from *in vitro* studies. Further, it may improve hazard identification of existing industrial chemicals, and may also be applied for assessing genotoxic properties of novel cosmetics, since in the EU, for cosmetic ingredients, animal testing is prohibited since 2009 (EC Regulation 1223/2009). Moreover, this approach may be effective in drug development, by significantly avoiding false positive results of the standard *in vitro* genotoxicity test battery, implying that promising lead compounds will no longer be eliminated due to wrong assumptions on their genotoxic properties and that rodents would not be unnecessarily sacrificed in costly follow-up experiments. Due to the small number of genes required for this transcriptomics-based prediction of *in vivo* genotoxicity – 33 for the Ames-positive compounds and 22 for the Ames-negative compounds – several technological options may be considered for developing fast high-throughput assays.

In conclusion, with this study, we demonstrate that the evaluation of specific gene expression alterations in HepG2 cells after 24h of treatment with chemicals, in combination

with results from the Ames mutagenicity test, can successfully predict genotoxic hazard of chemicals *in vivo*. We identified genes that play an important role in chemically induced genotoxicity, whose expression appears to be regulated by the DNA damage response-related transcription factor SP1 and the liver-specific transcription factors HNF3 α , HNF4 α and HNF6. Our prediction model in combination with analysis of compound specific transcriptomics responses in HepG2 cells, thus, provides additional relevant mechanistic information. Our method has a greatly improved specificity compared to other *in vitro* tests, while maintaining a similar high sensitivity. Eventually, our method may contribute to the reduction of animal experimentations that are required for labeling of compounds, and thus complies with 3R strategies.

Supplementary data

Supplementary data are available online at <http://toxicogenomics-um.nl/d/9665491>.

References

1. Jemal A, Siegel R, Xu J, & Ward E (2010) Cancer statistics, 2010. *CA Cancer J Clin* 60(5):277-300.
2. OECD (2010) Mortality from Cancer, in OECD, Health at a Glance: Europe 2010 *OECD Publishing*:34-35.
3. Petrucelli N, Daly MB, & Feldman GL (2010) Hereditary breast and ovarian cancer due to mutations in BRCA1 and BRCA2. *Genet Med* 12(5):245-259.
4. Clapp RW, Jacobs MM, & Loechler EL (2008) Environmental and occupational causes of cancer: new evidence 2005-2007. *Rev Environ Health* 23(1):1-37.
5. Oliveira PA, *et al.* (2007) Chemical carcinogenesis. *An Acad Bras Cienc* 79(4):593-616.
6. Ames BN, Lee FD, & Durston WE (1973) An improved bacterial test system for the detection and classification of mutagens and carcinogens. *Proc Natl Acad Sci U S A* 70(3):782-786.
7. Kirkland D, Aardema M, Henderson L, & Möller L (2005) Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens I. Sensitivity, specificity and relative predictivity. *Mutat Res.* 584(1-2):1-256.
8. Workman P, *et al.* (2010) Guidelines for the welfare and use of animals in cancer research. *Br J Cancer* 102(11):1555-1577.
9. Hamadeh HK, *et al.* (2002) Gene Expression Analysis Reveals Chemical-Specific Profiles *Toxicol Sci* 67(2):219-231.
10. Ellinger-Ziegelbauer H, Stuart B, Wahle B, Bomann W, & Ahr HJ (2005) Comparison of the expression profiles induced by genotoxic and nongenotoxic carcinogens in rat liver. *Mutat Res.* 575(1-2):61-84.
11. van Delft JH, *et al.* (2004) Discrimination of genotoxic from non-genotoxic carcinogens by gene expression profiling. *Carcinogenesis* 25(7):1265-1276.
12. Mathijs K, *et al.* (2009) Discrimination for genotoxic and non-genotoxic carcinogens by gene expression profiling in primary mouse hepatocytes improves with exposure time. *Toxicol Sci.*
13. Dickinson DA, *et al.* (2004) Differentiation of DNA reactive and non-reactive genotoxic mechanisms using gene expression profile analysis. *Mutat Res* 549(1-2):29-41.
14. Chen Y, Huang C, Zhou T, & Chen G (2008) Genistein induction of human sulfotransferases in HepG2 and Caco-2 cells. *Basic Clin Pharmacol Toxicol* 103(6):553-559.
15. Guo L, *et al.* (2010) Similarities and Differences in the Expression of Drug Metabolizing Enzymes between Human Hepatic Cell Lines and Primary Human Hepatocytes. *Drug Metab Dispos.*
16. Lancon A, *et al.* (2007) Resveratrol in human hepatoma HepG2 cells: metabolism and inducibility of detoxifying enzymes. *Drug Metab Dispos* 35(5):699-703.
17. Liguori MJ, Blomme EA, & Waring JF (2008) Trovafloxacin-induced gene expression changes in liver-derived in vitro systems: comparison of primary human hepatocytes to HepG2 cells. *Drug Metab Dispos* 36(2):223-233.
18. Jennen DG, *et al.* (2010) Biotransformation pathway maps in WikiPathways enable direct visualization of drug metabolism related expression changes. *Drug Discov Today* 15(19-20):851-858.
19. Hsu I, *et al.* (1993) p53 gene mutation and integrated hepatitis B viral DNA sequences in human liver cancer cell lines. *Carcinogenesis.* 14(5):987-992.
20. Natarajan AT & Darroudi F (1991) Use of human hepatoma cells for in vitro metabolic activation of chemical mutagens/carcinogens. *Mutagenesis* 6(5):399-403.
21. Knasmüller S, *et al.* (2004) Use of human-derived liver cell lines for the detection of environmental and dietary genotoxicants; current state of knowledge. *Toxicology* 198(1-3):315-328.
22. Winter HK, *et al.* (2008) Use of four new human-derived liver-cell lines for the detection of genotoxic compounds in the single-cell gel electrophoresis (SCGE) assay. *Mutation research* 657(2):133-139.
23. Jennen DG, *et al.* (2010) Comparison of HepG2 and HepaRG by whole genome gene expression analysis for the purpose of chemical hazard identification. *Toxicol Sci.*
24. Westerink WM, Stevenson JC, Horbach GJ, & Schoonen WG (2010) The development of RAD51C, Cystatin A, p53 and Nrf2 luciferase-reporter assays in metabolically competent HepG2 cells for the assessment of mechanism-based genotoxicity and of oxidative stress in the early research phase of drug development. *Mutat Res* 696(1):21-40.
25. Knight AW, *et al.* (2009) Evaluation of high-throughput genotoxicity assays used in profiling the US EPA ToxCast chemicals. *Regul Toxicol Pharmacol* 55(2):188-199.
26. ICH (2008) Guidance on genotoxicity testing and data interpretation for pharmaceuticals intended for human use S2(R1).
27. Edgar R, Domrachev M, & Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30(1):207-210.

28. Dai M, *et al.* (2005) Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res* 33(20):e175.
29. Irizarry RA, *et al.* (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4(2):249-264.
30. De Groot PJ, Reiff C, Mayer C, & Muller M (2008) NuGO contributions to GenePattern. *Genes Nutr* 3(3-4):143-146.
31. Tibshirani R, Hastie T, Narasimhan B, & Chu G (2002) Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci U S A*. 99(10):6567-6572.
32. Dennis G, Jr., *et al.* (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 4(5):P3.
33. IARC (1999) IARC Monographs on the evaluation of carcinogenic risks to humans. Volume 73: Some Chemicals that Cause Tumours of the Kidney or Urinary Bladder in Rodents and Some Other Substances.
34. Kligerman AD, Doerr CL, Tennant AH, & Peng B (2000) Cytogenetic studies of three triazine herbicides. II. In vivo micronucleus studies in mouse bone marrow. *Mutat Res* 471(1-2):107-112.
35. Tennant AH, Peng B, & Kligerman AD (2001) Genotoxicity studies of three triazine herbicides: in vivo studies using the alkaline single cell gel (SCG) assay. *Mutat Res* 493(1-2):1-10.
36. Rothfuss A, *et al.* (2010) Collaborative study on fifteen compounds in the rat-liver Comet assay integrated into 2- and 4-week repeat-dose studies. *Mutat Res* 702(1):40-69.
37. Czekaj P (2000) Phenobarbital-induced expression of cytochrome P450 genes. *Acta Biochim Pol* 47(4):1093-1105.
38. Rotroff DM, *et al.* (2010) Xenobiotic-metabolizing enzyme and transporter gene expression in primary cultures of human hepatocytes modulated by ToxCast chemicals. *J Toxicol Environ Health B Crit Rev* 13(2-4):329-346.
39. NTP (1985) NTP TR 275: NTP TECHNICAL REPORT ON THE TOXICOLOGY AND CARCINOGENESIS STUDIES OF 2-CHLOROETHANOL (ETHYLENE CHLOROHYDRIN) (CAS NO. 107-07-3) IN F344/N RATS AND SWISS CD-1 MICE (DERMAL STUDIES).
40. IARC (1987) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Supplement 7: Overall Evaluations of Carcinogenicity: An Updating of IARC Monographs Volumes 1 to 42.
41. Voogd CE (1989) Azathioprine, a genotoxic agent to be considered non-genotoxic in man. *Mutat Res* 221(2):133-152.
42. Bergman K, Muller L, & Teigen SW (1996) Series: current issues in mutagenesis and carcinogenesis. No. 65. The genotoxicity and carcinogenicity of paracetamol: a regulatory (re)view. *Mutat Res* 349(2):263-288.
43. Oshida K, Iwanaga E, Miyamoto-Kuramitsu K, & Miyamoto Y (2008) An in vivo comet assay of multiple organs (liver, kidney and bone marrow) in mice treated with methyl methanesulfonate and acetaminophen accompanied by hematology and/or blood chemistry. *J Toxicol Sci* 33(5):515-524.
44. Knudsen KE & Penning TM (2010) Partners in crime: deregulation of AR activity and androgen synthesis in prostate cancer. *Trends Endocrinol Metab* 21(5):315-324.
45. Iwahori S, *et al.* (2008) Identification of phosphorylation sites on transcription factor Sp1 in response to DNA damage and its accumulation at damaged sites. *Cell Signal* 20(10):1795-1803.
46. Limaye PB, *et al.* (2008) Expression of specific hepatocyte and cholangiocyte transcription factors in human liver disease and embryonic development. *Lab Invest* 88(8):865-872.
47. Gonzalez FJ (2008) Regulation of hepatocyte nuclear factor 4 alpha-mediated transcription. *Drug Metab Pharmacokinet* 23(1):2-7.
48. He N, *et al.* (2008) Epigenetic inhibition of nuclear receptor small heterodimer partner is associated with and regulates hepatocellular carcinoma growth. *Gastroenterology* 134(3):793-802.
49. You H, *et al.* (2004) p53-dependent inhibition of FKHL1 in response to DNA damage through protein kinase SGK1. *Proc Natl Acad Sci U S A* 101(39):14057-14062.
50. Zhang L, Cui R, Cheng X, & Du J (2005) Antiapoptotic effect of serum and glucocorticoid-inducible protein kinase is mediated by novel mechanism activating I{kappa}B kinase. *Cancer Res* 65(2):457-464.
51. van der Jagt K, Munn S, Tırslıvı J, & Bruijn Jd (2004) Alternative Approaches can reduce the use of test animals under REACH: Addendum to the report "Assessment of additional testing needs under REACH. Effects of (Q)SARS, risk based testing and voluntary industry initiatives". EUROPEAN COMMISSION, DIRECTORATE GENERAL JRC, JOINT RESEARCH CENTRE, Institute for Health and Consumer Protection.

Chapter 5

Characterization of the importance of NR0B2 as a classifying gene for genotoxicity prediction: a silencing approach prior and after genotoxic treatment

In preparation

C. Magkoufopoulou, L. Timmermans
S.M.H. Claessen, D.G.J. Jennen
J.C.S. Kleinjans, J.H.M van Delft

Abstract

We recently demonstrated that stratification of chemicals based on the results of the Salmonella mutagenicity test (Ames assay) in combination with transcriptomics changes in HepG2 cells upon 24h of treatment reliably predicts *in vivo* genotoxicity. For chemicals that show positive results in the Ames assay, an expression profile is used comprising 33 genes. However the biological relevance of several of these genes in DNA damage responses is not yet known. Among those is NR0B2; a nuclear receptor that acts as a repressor of transcription affecting several biological processes and was identified as one of the genes with the highest contribution in the prediction model. The aim of this study therefore was to characterize the significance of NR0B2 in the prediction model and its involvement in DNA damage responses. For that purpose small interference RNA (siRNA) technology was applied for suppressing the expression of NR0B2 in HepG2. Subsequently HepG2 cells were exposed for 24h to two Ames-positive chemicals, Benzo[a]pyrene (an established genotoxic carcinogen) and 2Chloromethylpyridine (a non-genotoxic irritant). Transcriptomics responses were analyzed and the prediction model was applied for genotoxicity evaluation. Further, cell cycle distribution, induction of apoptosis and induction of DNA double strand breaks were analyzed. Both Benzo[a]pyrene and 2Chloromethylpyridine were correctly classified as genotoxic and non-genotoxic agents, respectively. Both compounds altered the cell cycle distribution of HepG2 cells. Our results thus indicate that knockdown of NR0B2 is not sufficient to affect genotoxicity prediction by transcriptomics-profiling. Regarding the role of NR0B2, transcriptomics analyses show that NR0B2 appears to be important in several biological processes of HepG2 cells, including metabolic and signaling pathways, while phenotypic analyses indicate that after chemical treatment NR0B2 plays a dual role in the S- to G2/M-phase transition and in the induction of apoptosis depending on the external stimulus. In conclusion, with this study we prove the biological relevance of NR0B2 in genotoxic responses and explain its selection as one of the classifying genes.

Introduction

With cancer being one of the principal causes of death worldwide (World Health Organization) and the second cause of death accounting for approximately 25% of total deaths in the western world (1, 2), a lot of effort is made in understanding its development. Cancer is a very complicated and yet not fully understood disease; however, two causal factors for its development are appreciated. The first is the presence of specific gene mutations genetically inherited or endogenously induced; for instance the inheritance of BRCA1 and BRCA2 mutations are considered responsible for increased breast cancer risk (3). The second causal factor for cancer development is the exposure to exogenous carcinogenic factors, such as tobacco smoke that has been associated with lung cancer development (4). The molecular mechanism of tumor formation after exposure to carcinogenic factors frequently comprises the induction of DNA mutations by the carcinogen or its metabolites. If mutations occur within genes responsible for cell proliferation or survival, the cells may become malignant (5). Cellular transformation to a tumor cell may also be promoted through a variety of chemically induced mechanisms (production of reactive oxygen species, immunosuppression, peroxisome proliferation etc.) which do not necessarily involve DNA damage. Consequently, carcinogens are classified as genotoxic (GTX) or non-genotoxic (NGTX) (5).

The most commonly used assay, the bacterial gene mutation assay, for evaluating genotoxic properties of chemicals *in vitro* was developed in 1975 by Bruce N. Ames (6). Subsequently, several *in vitro* assays were developed, aiming at assessing genotoxic properties of chemicals in mammalian cellular models which have been accepted by the regulatory authorities. These include the mammalian micronuclei (MN), chromosomal aberration (CA) and mouse lymphoma assays (MLA). However, this conventional *in vitro* test battery often fails to correctly predict the *in vivo* genotoxic and carcinogenic potential of compounds, reaching a 50% false positive rate in some cases (7). Recently, we demonstrated that the evaluation of specific gene expression alterations in HepG2 cells after 24h of treatment with chemicals, after stratification of the chemicals based on their effects in the standard Salmonella mutagenicity test according to B.N. Ames, can successfully predict their genotoxic hazard *in vivo* (Chapter 4). This method has a greatly improved specificity compared to other *in vitro* tests, while maintaining a similarly high sensitivity. In applying this method the expression alterations are measured of 33 classifying genes for chemicals that show positive results in the standard Ames assay. However the biological role of several of these genes in DNA damage responses has not yet been studied. One of these genes, with no known association with DNA damage responses is NR0B2. NR0B2 showed increased expression levels after GTX treatments and was ranked first based on the t-test selection of genes. Furthermore, NR0B2 appeared to regulate several transcription networks of identified classifiers.

NR0B2 is a member of the nuclear receptor superfamily. Nuclear receptors are transcription factors that their transcriptional activity depends on the binding of a ligand. However, NR0B2 lacks the DNA-binding conserved domain, while it still contains a putative ligand-binding domain. Therefore, it regulates target gene transcription without directly bind-

ing to target promoter sequences. Rather, it directly interacts with other nuclear receptors and represses their transcriptional activity. Inhibitory effects of NR0B2 has been demonstrated for the transcriptional activity of Estrogen Receptors (8), Estrogen Related Receptors (9), Farnesoid X Receptor (10) Retinoid X Receptor, Hepatocyte Nuclear Factor 4 (11) and several other known ligand-dependent nuclear receptors (12). Three possible mechanisms of the transcriptional repression by NR0B2 have been proposed and these include (a) co-activator competition, (b) blocking of DNA binding and (c) active repression. NR0B2 has been further reported to play a role in a variety of physiological processes including lipid, cholesterol and glucose metabolism, homeostatic regulation and detoxification and drug metabolism in liver [reviewed in (13)]. Furthermore, NR0B2 has been proposed as a tumor suppressor gene based on its reduced expression in hepatocellular carcinoma but also in other types of cancer (12,14).

We, consequently, hypothesized that the expression of NR0B2 is important for the correct classification of GTX agents by transcriptomics-profiling and therefore may play a role in DNA damage responses. Thus, the aim of this study was to initially characterize the impact of NR0B2 expression on the genotoxicity evaluation model that we previously developed and to further characterize its involvement in DNA damage responses. For that purpose NR0B2 was silenced in HepG2 cells using small interference RNA (siRNA). Since this gene was selected as a classifier for Ames positive chemicals, after silencing HepG2 cells were exposed to the Ames positive GTX chemical, Benzo[a]pyrene and the Ames positive NGTX chemical 2Chloromethylpyridine for 24h, in order to initially investigate the effect of NR0B2 silencing on genotoxicity prediction using the previously identified gene expression profile of 33 genes. Further, we used the resulting whole genome transcriptomics responses for the characterization of the biological role of the silenced gene prior and after chemical treatment. Lastly, phenotypic analyses, including apoptosis, cell cycle distribution and induction of DNA double strand breaks, were conducted to demonstrate the role of NR0B2 in these processes that are considered to be phenotypic endpoints of genotoxicity (15,16).

Material and methods

Cell culture

HepG2 cells were cultured in six-well plates in the presence of minimal essential medium supplemented with 1% nonessential amino acids, 1% sodium pyruvate, 2% penicillin/streptomycin and 10% fetal bovine serum (all from Gibco BRL, Breda, The Netherlands). The cells were incubated at 37°C and 5% CO₂ (17).

siRNA transfection

ON-TARGETplus non-targeting control siRNA pool and NR0B2 SMARTpool siRNA were purchased from ThermoScientific. Cells were transfected with 100 nM of non-targeting or NR0B2 siRNA using DharmaFECT-4 transfection reagent (Thermoscientific) according to the manufacturer's protocol.

Chemical treatment

24h after transfection with non-targeting or NR0B2 siRNA, the medium was replaced with fresh medium containing 2 μ M Benzo[a]Pyrene (BaP) (CAS no.50-32-8; Sigma-Aldrich), 300 μ M 2-ChloromethylPyridene (2CMP) (CAs no. 6959-47-3; Fluka) or 0.5% dimethyl sulfoxide (DMSO) (CAS no. 67-68-5; Sigma-Aldrich) as a vehicle control. Cells were exposed for 24h. Thereafter the culture medium was replaced by TRIZOL (Gibco/BRL) for RNA isolation or the cells were methanol fixated and stored at -20°C for flow cytometric analysis. Three independent biological replicates were conducted.

Total RNA isolation and microarray experiments

Total RNA was extracted using 0.5 ml TRIZOL according to the manufacturer's instructions and purified using RNeasy® Mini Kits (Qiagen). Sample preparation, hybridization, washing, staining and scanning of the Affymetrix Human Genome U133 Plus 2.0 GeneChip arrays were conducted according to the manufacturer's manual as previously described (17). Quality controls were within acceptable limits. Hybridization controls were called present on all arrays and yielded the expected increases in intensities. Available microarray experiments of BaP- and 2CMP-exposed HepG2 cells for 24h were obtained from NCBI's Gene Expression Omnibus with GEO Series accession number GSE28878 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE28878>) (Chapter 4) and used as reference for chemically induced gene expression alterations.

Annotation, normalization and filtering of microarray data

The obtained data sets were re-annotated to the MBNI Custom CDF-files (http://brainarray.mbni.med.umich.edu/Brainarray/Database/CustomCDF/genomic_curated_CDF.asp) (18) and RMA normalized (19) using the NuGOExpressionFileCreator in GenePattern (20). Filtering of microarray data were conducted per experimental condition as previously described (21).

Class prediction of chemical treatments

The developed prediction model based on the expression levels of 33 genes (Chapter 4) was applied for the classification of BaP and 2CMP-treatments in the presence of NR0B2 or non-targeting siRNA. For this the log₂ ratios of the treatments to the corresponding DMSO-controls were calculated.

Selection of Differentially Expressed Genes

The selection of DEGs for DMSO-treated HepG2 cells after NR0B2 siRNA transfection was performed by calculating the log₂ ratios of intensities to DMSO-treated cells after non-targeting siRNA transfection. The selection of DEGs for BaP and 2CMP treatments, in the presence of NR0B2 siRNA, non-targeting siRNA or in the absence of siRNA, was done by calculating the log₂ ratios of intensities to the corresponding DMSO treatments. DEGs per treatment were selected using the criteria: (a) absolute average log₂ ratio for the three replicates > 0.58, (b) same direction of the log ratio for all the three replicates, and (c) absolute log₂ ratio of > 0.26 for at least two replicates in comparison to control.

Functional annotation and network analyses

Functional Annotation based on Gene Ontology (GO) Biological Processes (GOTERM_BP_FAT)

was conducted for the up- and down-regulated DEGs of each treatment separately, using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (22) in order to identify processes that are significantly affected by NR0B2 silencing. For each gene list, a background list resulting from the data filtering was used. GO terms were considered significant when $EASE < 0.1$ and $p\text{-value} < 0.05$ and included more than four genes.

Direct interaction network analyses were conducted in MetaCore Network Tools (GeneGo, San Diego, CA) for each treatment separately for the identification of direct interacting genes among the DEGs for each treatment. The following filters were used: species – Homo sapiens, interaction types (effects) – activation and inhibition.

Flow cytometric analyses

Methanol-fixated cells were washed with phosphate-buffered saline containing 1 mg/ml bovine serum albumin (PBS/BSA) and incubated for 1h with the primary antibody mouse anti-phospho-histone H2AX (Ser-139) Clone JBW 301 (Upstate) diluted 1:1000 in PBS/BSA for the evaluation of DNA double strand breaks, and M30 cytodeath antibody (Peviva) diluted 1:1000 in PBS/BSA for the evaluation of apoptosis. Cells were washed and incubated for 1h with the secondary antibody Alexa Fluor 488-conjugated anti-mouse IgG, (Invitrogen). Cells were re-suspended in 0.5 ml phosphate-buffered saline containing 20 $\mu\text{g/ml}$ propidium iodide (Sigma-Aldrich) and 100 $\mu\text{g/ml}$ RNase A (Sigma-Aldrich) (15).

FACSort (Becton Dickinson, Sunnyvale, CA) was used for flow cytometric analysis as previously described (23). Cells with significant amounts of anti-phospho-Histone H2AX and M30 positive signals were displayed as a percentage of total cells. Cell cycle was analyzed using ModFit LT for Mac (version 2.0). Cells in the G0/G1-, S- or G2/M-phase were expressed as a percentage of the total cells. Values for γH2AX , M30 and cell cycle were considered significant in comparison to control cells when $p\text{-value} < 0.05$ (paired t-test).

Results

Transcriptomics analysis

The Hierarchical Cluster Analysis (Figure 1A) of the whole genome transcriptomics responses for all the compound and control treatments without or in combination with NR0B2 siRNA or non-targeting siRNA results in two main clusters. The first one includes the transcriptomics responses of HepG2 cells that have been transfected with siRNA (NR0B2 or non-targeting). Further, a clear separation is observed between NR0B2 siRNA transfected cells and non-targeting siRNA transfected cells; while BaP and DMSO treatments are not clearly separated.

The second main cluster includes the transcriptomics responses of compound and control treated HepG2 in the absence of siRNA. Within this cluster the treatments are separated based on the compound. Two different groups of DMSO treatments were conducted; the first as control to BaP exposures (DMSO_A) and the second as control to 2CMP exposures (DMSO_B). These two groups of DMSO treatments are not clustered together; nevertheless each one is clustered closer to the compound treatments for which they served as control, suggesting a technical variation among the experiments.

Figure 1: Whole genome gene expression comparison. A. HCA (Arraytrack v.3.5.0) of BaP-, 2CMP- and DMSO-treatments using log2 intensities after RMA normalization (NR0B2: treatment in the presence of NR0B2 siRNA; NT: treatment in the presence of non-targeting siRNA; numbers after the treatments indicate the biological replicate) B.C.D. Scatter plots for DMSO treatments using the median of log2 intensities. (A_DMSO, B_DMSO: DMSO treatments in the absence of siRNA conducted as control for BaP and 2CMP-treatments, respectively; DMSO_NT: DMSO treatment in the presence of non-targeting siRNA).

Figure 2: Direct Interaction Network of Differentially Expressed Genes by NR0B2 silencing in control (DMSO) treatments. The type of interaction between the genes is indicated by the different colors (green: activation, red: inhibition). The circles next to the symbols indicate the expression of the genes after NR0B2 silencing. Red circles indicate up-regulation and blue circles down-regulation. NR0B2 is highlighted by a red circle around it. Detailed information on the symbols can be found at http://www.genego.com/pdf/MC_legend.pdf.

To ensure that the transfection process would not significantly alter the basal gene expressions of HepG2 cells, we performed a comparison between DMSO treatments in the absence of siRNA and DMSO treatments in the presence of non-targeting siRNA (Figure 1C and 1D). While there is a small number of genes that appears to be affected by the transfection process (see Figure 1B for comparison), the variation between the DMSO-treatments in the presence of non-targeting siRNA and the DMSO-treatments in the absence of non-targeting siRNA is similar to that observed among the two DMSO-treatments in the absence of siRNA. This is further supported by the Pearson correlation coefficient that is 0.98 in all three comparisons, indicating that there is no major effect of the transfection on basal gene expression levels of HepG2 cells.

Effect of NR0B2 silencing on genotoxicity prediction

In order to evaluate the contribution of NR0B2 in genotoxicity prediction we used the previously developed algorithm (Chapter 4) for the classification of BaP and 2CMP exposures in combination with the presence of NR0B2 siRNA or non-targeting siRNA. BaP treatments were classified as GTX with a 99% probability for both siRNA transfections, whereas 2CMP-treatments were classified as NGTX with a probability of 98% and 99% for NR0B2 siRNA transfection and non-targeting siRNA transfection, respectively. Thus, our results indicate that NR0B2 silencing does not affect the correct classification of chemicals based on the expression values of the remaining 32 classifying genes.

Transcriptomics effects of NR0B2 silencing prior to chemical treatment

For a better understanding of the biological role of NR0B2, we analyzed the transcriptomics changes induced by NR0B2 in DMSO-treated HepG2 cells. Silencing of NR0B2 resulted in the up-regulation of 133 genes and the down-regulation of 281 genes in comparison to cells expressing NR0B2 but were transfected with non-targeting siRNA. Network Analysis revealed direct interactions among 73 of these genes (Figure 2), including NR0B2 that appears down-regulated. In addition genes that are known to be negatively regulated by NR0B2, show high expression values after NR0B2 silencing, as expected (Figure 2).

Functional Annotation analysis based on Gene Ontology biological processes revealed several processes that are positively or negatively affected by NR0B2 silencing (Table 1). In particular up-regulated DEGs after NR0B2 silencing are mainly involved in metabolic, transport and homeostasis related processes, indicating a negative effect of NR0B2 in these biological processes. Down-regulated DEGs after NR0B2 silencing are involved in more processes, including signaling, response to a variety of stimuli, metabolic processes, proliferation, transport, homeostasis, development, transcription, cell cycle and apoptosis. These results indicate an important role of NR0B2 for the regular function of HepG2 cells since its absence alters several biological processes.

Transcriptomics effects of NR0B2 silencing in combination with chemical treatment

In order to understand the role of NR0B2 in chemical responses, we further analyzed transcriptomics effects of BaP and 2CMP exposures in NR0B2-silenced HepG2 cells. In order to reveal the transcriptomics effects due to the silencing of NR0B2, the transcriptomics effects caused by chemical treatment in NR0B2-silenced cells were compared with the transcriptomics effects by chemical treatment in cells with non-targeting siRNA transfection.

Table 1: GO terms affected by NR0B2 silencing in control treatments

A. Resulting from up-regulated genes	
Metabolic processes	oxidation reduction (12)*, lipid/glycerolipid catabolic process (8), lipid biosynthetic process (8), triglyceride/neutral lipid/glycerol metabolic process (6), cellular amino acid derivative metabolic process (6), hormone metabolic process (5), regulation of lipid metabolic process (5), regulation of fatty acid metabolic process (4), regulation of cellular catabolic process (4), lipoprotein metabolic process (4)
Homeostatic processes	chemical homeostasis (9), lipid homeostasis (4)
Transport processes	lipid transport/localization (7), carboxylic acid/organic acid transport (5)
Other	regulation of hormone levels (7)
B. Resulting from down-regulated genes	
Signaling	cell surface receptor linked signal transduction (36), small GTPase mediated signal transduction (14), G-protein coupled receptor protein signaling pathway (13), enzyme linked receptor protein signaling pathway (13), Ras protein signal transduction (8), transmembrane receptor protein serine/threonine kinase signaling pathway (7), transforming growth factor beta receptor signaling pathway (5)
Response to stimuli	response to organic substance (30), response to wounding (22), response to endogenous stimulus (19), defense response (19), response to hormone stimulus (16), response to extracellular stimulus (13), response to nutrient levels (12), regulation of response to external stimulus (11), positive regulation of response to stimulus (9), inflammatory response (9), response to hypoxia (8), response to protein stimulus (8), response to bacterium (7), response to vitamin (5), response to osmotic stress (4), response to amine stimulus (4)
Metabolic processes	positive regulation of macromolecule metabolic process (25), positive regulation of biosynthetic process (22), positive regulation of nitrogen compound metabolic process (20), regulation of phosphate metabolic process (17), regulation of lipid metabolic process (7), regulation of cellular catabolic process (5), regulation of nitric oxide biosynthetic process (4)
Proliferation / growth	regulation of cell proliferation (28), cell proliferation (17), regulation of growth (14), regulation of cellular component size (11), regulation of cell size (9), regulation of smooth muscle cell proliferation (5)
Transport	ion transport (17), carboxylic acid / organic acid transport (10), Golgi vesicle transport (9), amino acid transport (7)
Homeostatic processes	homeostatic process (29), cation homeostasis (12)
Development / differentiation	vasculature development (13), skeletal system development (11), positive regulation of developmental process (11), angiogenesis (9), negative regulation of cell differentiation (9), regulation of cell morphogenesis involved in differentiation (5)
Locomotion	cell motion (17), regulation of cell migration (15), regulation of locomotion (15), regulation of leukocyte migration (4)
Enzyme regulation	phosphorylation (16), kinase activity (14), transferase activity (14), MAP kinase activity (8), proteolysis (5)
Transcription	positive regulation of transcription, DNA-dependent (15), regulation of transcription factor activity (7)
Organization	membrane organization (16), positive regulation of cellular component organization (9)
Other	negative regulation of apoptosis/cell death (14), behavior (11), regulation of body fluid levels (8), chemotaxis (7), wound healing (10), coagulation (7), hemostasis (7), positive regulation of cell cycle (5), regulation of membrane potential (7), elevation of cytosolic calcium ion concentration (4), regulation of muscle contraction (4)

*: Number in brackets indicate the number of genes involved in the indicated process

BaP exposure after NR0B2 silencing resulted in only 44 up-regulated and 23 down-regulated DEGs (Figure 3A). Low numbers of DEGs were also observed for BaP-exposed HepG2 cells after transfection with non-targeting siRNA (39 up-regulated and 15 down-regulated DEGs). When HepG2 cells were exposed to BaP in the absence of any siRNA, major transcriptomics changes were observed. BaP changed the expression of 847 genes in total (Figure 3A). The HCA for the commonly affected genes after BaP exposure in the presence or in the absence of siRNA (Figure 3B), clearly shows a stronger response of HepG2 to BaP in the absence of siRNA when compared to BaP-treated cells in the presence of siRNA. These observations indicate that the transfection process as such inhibits the response of HepG2 cells to BaP. Nevertheless, 8 genes were identified to show altered expression only in BaP-treated HepG2 cells in the presence of NR0B2 siRNA. OSMR, TGFB1I1, KIAA119, MBL2 and APOBEC3B were up-regulated indicating a negative relation of the expression of these genes with the expression of NR0B2 after BaP-treatment; whereas PLA2G1B, CYP7A1 and HPX were down-regulated indicating a positive relation of the expression of these genes with the expression of NR0B2 after BaP-treatment (Table 2). These genes do not show any direct interactions by Metacore Network Analysis; however, the expression of only CYP7A1 is known to be regulated by NR0B2. It further appears that these genes are involved in immune responses and signaling, suggesting a role for NR0B2 in these processes after genotoxic treatments.

Figure 3: Differentially expressed genes after BaP and 2CMP treatments. A. Venn Diagrams (Venny: <http://bioinformatics.csic.es/tools/venny/index.html>) of the up-regulated (left) or the down-regulated (right) DEGs after BaP (up) or 2CMP (down) treatments in the presence or absence of siRNA. Numbers in brackets indicate the total number of up- or down-regulated DEGs for the different treatments. B. HCA (Arraytrack v.3.5.0) of common DEGs among all BaP treatments. C. HCA (Arraytrack v.3.5.0) of common DEGs among all 2CMP treatments. (NR0B2: treatment in combination with NR0B2 siRNA transfection; NT: treatment in combination with non-targeting siRNA transfection).

Similarly, siRNA transfection (NR0B2 and non-targeting) of HepG2 cells prior to 2CMP treatment resulted in a reduced number of DEGs (Figure 3A) and decreased levels of expression for the genes that were commonly affected in siRNA-transfected and siRNA-non-transfected cells (Figure 3B). However, the expression of 70 genes appeared to be specifically affected by NR0B2 silencing. Whereas neither direct interactions among these

Table 2: DEGs specifically affected after BaP exposure and NR0B2 silencing

EntrezGene ID	Symbol	AVG ¹	GO-terms ²
9180*	OSMR	0,76711	cell proliferation, cell surface receptor linked signaling pathway, positive regulation of acute inflammatory response, response to cytokine stimulus
7041	TGFB11	0,67059	Wnt receptor/androgen receptor signaling pathway, cell adhesion, cell fate commitment, epithelial cell differentiation, negative regulation of cell proliferation, regulation of transforming growth factor beta receptor signaling pathway, positive regulation of transcription, ubiquitin-dependent SMAD protein catabolic process
57214	KIAA1199	0,62479	-
4153*	MBL2	0,60367	acute-phase response, complement activation, defense response to Gram-positive bacterium, innate immune response, response to oxidative stress
9582	APOBEC3B	0,58901	-
5319**	PLA2G1B	-0,59661	actin filament organization, activation of MAPK activity, signal transduction, arachidonic acid secretion, cellular response to insulin stimulus, fatty acid/leukotriene biosynthetic process, glucose transport, interleukin-8 production, lipid catabolic process, neutrophil chemotaxis, phosphatidylcholine/phospholipid metabolic process, positive regulation of DNA replication/immune response/NF-kappaB transcription factor activity/calcium ion transport/protein secretion/transcription
1581	CYP7A1	-0,69510	xenobiotic metabolic process, bile acid biosynthetic/metabolic process, lipid metabolic process, cholesterol catabolic process, cholesterol homeostasis, cellular response to glucose stimulus/cholesterol
3263*	HPX	-0,69841	cellular iron ion homeostasis, heme/hemoglobin metabolic process, heme transport, positive regulation of humoral immune response/immunoglobulin production/interferon-gamma-mediated signaling pathway/JAK-STAT cascade

¹: Average log2 ratios of three independent biological replicates.

²: Gene Ontology Biological Process that the genes are known to be involved in.

*: the expression of the indicated genes is statistically (t-test, $p < 0.05$) significant in comparison to their expression in cells expressing NR0B2.

**: the expression of the indicated genes is not detected in cells expressing NR0B2.

genes nor significantly represented GO-biological processes were identified, several of these genes appear to be involved in signaling, metabolic processes and transcription, indicating a role of NR0B2 in these processes after exposure of HepG2 cells to 2CMP (Table 3).

Phenotypic effects of NR0B2 silencing

To further evaluate the involvement of NR0B2 in genotoxic responses, after NR0B2 silencing we measured phenotypic endpoints of genotoxicity – cell cycle, apoptosis and induction of H2AX phosphorylation – for control, BaP- and 2CMP-treated cells (Figure 4).

No significant effect of the NR0B2 silencing is observed for any of the phenotypic endpoints in the control treatments.

Table 3: GO-terms for the individual DEGs after 2CMP-treatment and NR0B2 silencing

EntrezGene ID	Symbol	AVG ¹	GO-terms ²
2730	GCLM	0,9717	cellular homeostasis; blood circulation; glutathione BP; cysteine, glutamate, xenobiotic, peptide, sulfur MP; membrane potential (R); cell death (NR); mitochondrial depolarization (NR); response to stress
145389*	SLC38A6	0,8018	amino acid, ion, carboxylic acid transport
4141	MARS	0,7706	gene expression; methionyl-tRNA aminoacylation
57016*	AKR1B10	0,7394	aldehyde, steroid MP; digestion; oxidation reduction
221662	RBM24	0,7282	cell differentiation, mRNA stability (R); myotube differentiation (R)
1591	CYP24A1	0,7229	intracellular signaling cascade; intracellular receptor-mediated, vitamin D receptor signaling pathway; hormone BP; osteoblast differentiation; oxidation reduction; ossification; skeletal system, bone development; response to vitamin D, nutrient, extracellular stimulus; steroid, vitamin D, xenobiotic MP; steroid, vitamin, lipid CP
122953	JDP2	0,7180	transcription (NR)
80267	EDEM3	0,7178	protein, glycoprotein MP; glycoprotein, ubiquitin-dependent protein CP; post-translational protein modification; protein folding; proteolysis; response to organic substance, protein stimulus
8560	DEGS1	0,7021	oxidation reduction; sphingolipid, fatty acid MP; unsaturated fatty acid, lipid BP
570*	BAAT	0,7006	acyl-CoA, bile acid, fatty acid, glycine, lipid, steroid MP; bile acid and bile salt transport; digestion
4189*	DNAJB9	0,6852	protein folding
1978	EIF4EBP1	0,6810	posttranscriptional regulation of gene expression; translation; G1/S transition of mitotic cell cycle; mitotic cell cycle (PR); TOR, insulin receptor signaling cascade; cell surface receptor linked signal transduction; cellular protein MP; gene expression; enzyme linked receptor protein, transmembrane receptor protein tyrosine kinase, insulin receptor signaling pathway; response to endogenous stimulus, hormone stimulus, organic substance, insulin stimulus; protein complex assembly (NR); translational initiation (NR); biosynthetic process (NR); translation (NR)
10797	MTHFD2	0,6719	folic acid-containing compound BP; one-carbon, tetrahydrofolate MP; oxidation reduction
3958	LGALS3	0,6706	skeletal system development; extracellular matrix organization
27032*	ATP2C1	0,6695	ion transport, ion homeostasis; cytoskeleton organization; cell adhesion; ectoderm, epidermis development; signal transduction (PR); cell communication (PR); protein kinase cascade (R); actin filament-based process; actin cytoskeleton organization; purine nucleotide MP; ATP BP
283219	KCTD21	0,6627	ion transport
2617	GARS	0,6437	tRNA MP; translation; cell death; vesicle-mediated transport; secretion
81671	VMP1	0,6353	autophagy; cell adhesion
64747	MFSD1	0,6336	transmembrane transport
90*	ACVR1	0,6317	G1/S transition of mitotic cell cycle; mitotic cell cycle; interphase; angiogenesis; gamete generation; mesoderm formation; gastrulation; pattern specification process; response to wounding; determination of symmetry; phosphorylation; phosphorus, phosphate MP; cell motion; cell migration; reproductive cellular process; cell motility; localization of cell; reproductive developmental process; vasculature, urogenital system, in utero embryonic, germ cell, mesoderm, heart, pharyngeal system, mesenchyme, tube, neural crest cell development; acute inflammatory, defense response; cell surface receptor linked signal transduction; enzyme linked receptor protein, transmembrane receptor protein serine/threonine kinase, transforming growth factor beta receptor signaling pathway; mesenchymal cell, muscle cell, neural crest cell differentiation; blood vessel, embryonic, mesoderm, tube, tissue morphogenesis; muscle development (R); biomineral formation (R/PR); transcription (R/PR); osteoblast differentiation (R/PR); apoptosis (R/NR); cell death (R/NR); ossification (R); bone mineralization (R/PR); activin receptor signaling pathway (R/NR); cell communication (NR); signal transduction (NR); developmental process (PR); ossification (PR); cell differentiation (PR); macromolecule MP (PR)
1192	CLIC1	0,6235	ion transport; defense response; response to unfolded protein, nutrient, extracellular stimulus, organic substance, protein stimulus
83941	TM2D1	0,6234	apoptosis; induction of apoptosis; cell death; cell death (R/PR); apoptosis (R/PR)

Table 3: (Continued)

EntrezGene ID	Symbol	AVG ¹	GO-terms ²
84376	HOOK3	0,6211	establishment of protein localization; organelle, Golgi, protein localization; microtubule cytoskeleton, endosome, vacuole, lysosome organization; vacuolar, lysosomal, endosome to lysosome, protein, vesicle-mediated, endosome intracellular transport
8412*	BCAR3	0,6146	intracellular signaling cascade; small GTPase mediated signal transduction; response to drug
7169	TPM2	0,6019	ATPase activity (R); hydrolase activity (R)
57561	ARRDC3	0,5953	signal transduction
29940	DSE	0,5916	polysaccharide BP; polysaccharide MP
55171	TBCCD1	0,5905	cell morphogenesis; maintenance of Golgi, centrosome location; cell migration (R); cell shape (R)
10201	NME6	0,5895	nucleotide, nucleoside MP; apoptosis; anti-apoptosis; cell death; cell death (R/NR); apoptosis (R/NR)
2584**	GALK1	-0,5850	phosphorylation; carbohydrate phosphorylation; monosaccharide, galactose, hexitol, phosphorus, phosphate, hexose, alditol, galactitol, polyol MP
3978	LIG1	-0,5855	somatic diversification of immune receptors; immune system development; DNA replication; DNA repair; nucleotide-excision repair; response to DNA damage stimulus; cell cycle; cell division
8352**	HIST1H3C	-0,5865	DNA packaging; chromatin, chromosome, nucleosome organization; chromatin assembly or disassembly; nucleosome assembly; macromolecular complex subunit organization
6909*	TBX2	-0,6130	cell fate determination; cell fate commitment; cell aging; heart morphogenesis; heart, muscle organ, striated muscle tissue, mammary gland, cardiac muscle tissue development; transcription; transcription (R/NR), cell proliferation (R); biosynthetic process (NR)
6678	SPARC	-0,6207	skeletal system, bone development; ossification; cell proliferation (R)
55753	OGDHL	-0,6262	generation of precursor metabolites and energy; oxidation reduction; glycolysis; monosaccharide, glucose, hexose MP; glucose, carbohydrate, hexose, alcohol, monosaccharide CP
7342	UBP1	-0,6315	angiogenesis; viral reproduction; viral infectious cycle; blood vessel morphogenesis; vasculature development; transcription; transcription (R/NR); biosynthetic process (NR)
64241	ABCG8	-0,6341	lipid, phospholipid, sterol, cholesterol transport; cholesterol, chemical, lipid, sterol homeostasis; lipid localization; digestion; excretion; secretion; intestinal cholesterol absorption; cholesterol efflux; lipid digestion
84662*	GLIS2	-0,6344	transcription; transcription (R); biosynthetic process (R); macromolecule MP (R)
54877	ZCCHCH2	-0,6360	cell communication
7274**	TTPA	-0,6403	vitamin MP; female pregnancy; embryo implantation; response to nutrient, pH, abiotic stimulus, extracellular stimulus
93663	ARGHGAP18	0,6410	signal transduction; small GTPase mediated signal transduction (R); GTPase activity (PR)
147837	ZNF563	-0,6578	transcription
6119	RPA3	-0,6623	nucleotide-excision repair; DNA replication; DNA repair; response to DNA damage stimulus
5948**	RBP2	-0,6672	isoprenoid, vitamin, vitamin A MP; ectoderm, epidermis development
2161*	F12	-0,6861	kinin cascade; acute inflammatory, defense, immune, innate immune response; proteolysis; coagulation; hemostasis; protein processing; zymogen activation; wound healing; fibrinolysis; protein maturation; response to wounding, organic substance, misfolded protein, protein stimulus; regulation of acute inflammatory response (R); plasminogen activation (R/PR); proteolysis (R); coagulation (R); response to external stimulus (R); cellular protein MP (R); body fluid levels (R); fibrinolysis (R); macromolecule MP (PR); proteolysis (PR)
120071	GYLTL1B	-0,6869	homeostatic process; muscle maintenance
7448*	VTN	-0,6878	immune response; cell, cell-matrix, cell-substrate adhesion; cell, cell-substrate adhesion (R); extracellular matrix organization
2165*	F13B	-0,6926	coagulation; hemostasis; response to wounding; wound healing; body fluid levels (R)
25797	QPCT	-0,7109	proteolysis

Table 3: (Continued)

EntrezGene ID	Symbol	AVG ¹	GO-terms ²
8202	NCOA3	-0,7159	transcription; transcription (R/PR); steroid hormone receptor signaling pathway (R/PR); estrogen receptor signaling pathway (R/PR); transcription factor activity (R/PR); intracellular signaling cascade; steroid hormone receptor-, androgen receptor-, intracellular receptor-mediated signaling pathway; biosynthetic process (PR); signal transduction (PR); cell communication (PR)
8339*	HIST1H2BG	-0,7210	DNA packaging; chromatin, nucleosome, chromosome organization; chromatin assembly or disassembly; nucleosome assembly; defense response; response to bacterium
5493*	PPL	-0,7419	ectoderm, epidermis, epithelium development; epidermal cell, keratinocyte, epithelial cell differentiation; keratinization
2633	GBP1	-0,7443	immune response
55224*	ETNK2	-0,7526	growth; in utero embryonic, placenta, post-embryonic development; cellular amino acid derivative, biogenic amine, ethanolamine, phospholipid, glycerophospholipid, organophosphate MP; phosphatidylethanolamine, lipid, phospholipid, amine, nitrogen compound, glycerolipid, ethanolamine BP
220359**	TIGD3	-0,7567	transcription (R)
732	C8B	-0,7725	protein maturation; immune effector process; activation of immune response; proteolysis; cell death; response to wounding; protein processing; cytolysis; complement activation; adaptive immune, humoral immune, innate immune, acute inflammatory, defense response; leukocyte-, lymphocyte-, B-cell-, immunoglobulin-mediated immunity; response to stimulus (PR); immune response (PR); immune system process (PR)
84634*	KISS1R	-0,8934	MAPKKK cascade; phosphorus, phosphate MP; lipid, monocarboxylic acid, organic acid, fatty acid transport; cell surface receptor linked signal transduction; G-protein coupled receptor protein signaling pathway; intracellular signaling cascade; calcium-, second-messenger-mediated signaling; behavior; lipid localization; secretion; cell migration (R); neurological system process (R); actin filament-based process (R); organelle organization (R/PR); locomotion (R); cell proliferation (R); cellular component biogenesis and organization (R); synaptic transmission (R); phosphorus metabolic process (R); cell motion (R); cytoskeleton organization (R); transmission of nerve impulse (R); cellular localization (R); stress fiber formation (R); cell communication (PR)
6660**	SOX5	-0,9002	skeletal system development; neuron differentiation; cell fate commitment; transcription; development (R); neurogenesis (R); gene expression (R/PR); biosynthetic process (PR)
8694*	DGAT1	-0,9130	peptide, insulin secretion; hormone levels (R); very-low-density lipoprotein particle assembly; cell-cell signaling; hormone, peptide transport; neutral lipid, acylglycerol, triglyceride, glycerol ether, organic ether, glycerolipid MP; lipid, triglyceride, glycerol ether BP
6539**	SLC6A12	-0,9503	neurotransmitter, amino acid, amine, carboxylic acid transport
3263*	HPX	-1,3113	cellular iron ion homeostasis; heme, hemoglobin MP; heme transport; humoral immune response (PR); immunoglobulin production (PR); interferon-gamma-mediated signaling pathway (PR); JAK-STAT cascade (PR)
843323**	DYDC2	0,6834	
80336 ³	PABPC1L	0,6065	
400629 ³	TEX19	0,5929	
201931 ³	TMEM192	0,5855	
387640*	C10orf140	-0,6572	
2033283**	SUSD3	-0,7187	
677838 ³	SNORA61	-0,7806	
134548 ³	ANKRD43	-1,0231	
115669 ³ *	TTC6	-1,1452	

¹: Average log2 ratios of three independent biological replicates; ²: Gene ontology biological processes that the genes are known to be involved in; ³: No gene ontology biological processes are identified for the indicated genes; R: regulation of; NR: negative regulation of; PR: positive regulation of; BP: biosynthetic process; MP: metabolic process; CP: catabolic process; *: the expression of the indicated genes is statistically (t-test, $p < 0.05$) significant in comparison to their expression in cells expressing NR0B2; **: the expression of the indicated genes is not detected in cells expressing NR0B2

BaP treatment of HepG2 cells after NR0B2 silencing results in a different cell cycle distribution in comparison to BaP treatment after non-targeting siRNA transfection (Figure 4A). In particular, BaP treatment in the presence of non-targeting siRNA results in a decrease of the number of cells in G1-phase and in increase of the number of cells in S-phase. BaP treatment in the presence of NR0B2 siRNA also results in a decrease of the number of cells in G1-phase and an increase of the number of cells in S-phase in comparison to the control-treatments. However, a decrease of the number of cells in S-phase can be observed in comparison to the BaP-treated cells in the presence of non-targeting siRNA. Additionally, BaP treatments of NR0B2-silenced cells result in an increase of the number of cells in G2/M-phase.

Furthermore, while significant levels of DNA damage are observed for BaP treatments in both NR0B2 siRNA transfected and non-targeting siRNA transfected cells in comparison to the control cells, no significant difference is observed between them (Figure 4B).

Lastly, significant levels of apoptosis are observed for BaP-treated cells in the presence of NR0B2 siRNA and in the presence of non-targeting siRNA. However, the levels of apoptosis induced by BaP in NR0B2 siRNA transfected cells are significantly lower than the levels of apoptosis induced by BaP in non-targeting siRNA transfected cells (Figure 4B).

All in all, these observations suggest that NR0B2 plays an important role in cell cycle progress and induction of apoptosis after genotoxic stress.

Figure 4: Phenotypic effects of NR0B2 silencing. A. Percentage of the cell population in G0–G1 phase, in S-phase and G2–M-phase. B. Percentage of the cell population with phosphorylated H2AX (DNA damage) or cleaved CK18 (apoptosis) after BaP-or 2CMP-treatment in the presence of non-targeting siRNA (NT) or NR0B2 siRNA (NR0B2). *: indicates $p < 0.05$ in comparison to the corresponding control; **: indicates $p < 0.05$ in comparison to the corresponding treatment in the presence of non-targeting siRNA.

2CMP exposure of NR0B2 siRNA-transfected cells also alters the cell cycle distribution in comparison to 2CMP exposures of non-targeting siRNA transfected cells (Figure 4A). In particular, while both 2CMP exposures of transfected cells result in a decrease of the percentage of cells in G1 phase, 2CMP exposure in combination with NR0B2 silencing further results in an increase of the percentage of cells in S-phase and a decrease of the percentage of cells in G2/M-phase. Although 2CMP is a non-genotoxic compound, 2CMP-treatments of transfected cells resulted in a significant increase of DNA damage as measured by the phosphorylation of H2AX. However, NR0B2 silencing did not significantly alter the levels of induced apoptosis or DNA damage after 2CMP treatment. These data further support the involvement of NR0B2 in cell cycle progress also after treatments by chemicals with no apparent *in vivo* genotoxic properties.

Discussion

In this study we aimed at the understanding of the importance of NR0B2 in the previously developed transcriptomics-based prediction model for Ames-positive chemicals in HepG2 cells (Chapter 4) and of its role in DNA damage responses. To this end, a siRNA silencing approach was applied in HepG2 cells in combination with and without chemical treatments while simultaneously, transcriptomics effects, as well phenotypic responses were evaluated. For a better understanding of the role of NR0B2, initially the induced transcriptomics effects of NR0B2 silencing in HepG2 cells growing under normal condition were evaluated. Further, the contribution of NR0B2 in the class prediction model, as derived from transcriptomics analysis, was studied. Lastly, the phenotypic effects after chemical treatment of NR0B2-silenced HepG2 cells were analyzed in order to evaluate the association of NR0B2 with DNA damage and known endpoints of genotoxicity.

Transcriptomics analysis of control treatments after NR0B2 silencing revealed several altered gene expressions involved in a variety of biological processes, including metabolic processes, signaling pathways, developmental processes and response to stimuli. Previous studies showed that overexpression of NR0B2 results in the inhibition of bile acid biosynthesis, by mainly inhibiting the expression of CYP7A1, the enzyme that initiates the biosynthesis of bile acids by cholesterol (10, 24, 25). In our study, the functional annotation analysis did not result in the identification of bile acid biosynthesis as an enriched biological process; however a four-fold increase of CYP7A1 was induced by NR0B2 silencing.

Furthermore, up-regulated DEGs after NR0B2 silencing were involved in lipid metabolism, proposing an inhibiting effect of NR0B2 in lipid metabolism. These findings are in concordance with a previous study using transgenic mice constitutively expressing NR0B2, resulting in the down-regulation of enzymes involved in lipid metabolism (24).

Down-regulated DEGs after NR0B2 silencing in HepG2 cells were involved in signaling pathways, including the TGF-beta receptor signaling pathway. These observations suggest a possible activating effect of NR0B2 in TGF-beta signaling. However, these findings are in contrast with previously reported results. In particular studies using rat renal cells over-expressing NR0B2 showed an inhibition of the TGF-beta-mediated expression of several genes (26). Similar findings were also reported for HepG2 and primary mice hepatocytes, where TGF-beta1 exposure in combination with over-expression of NR0B2 resulted in suppression of TGF-beta-mediated transcription, while knockdown of NR0B2 resulted in an increase of TGF-beta-mediated transcription after TGF-beta1 exposure (27). In our study, five genes (TGF-beta1, GDF15, TGF-beta-R3, BAMBI and ID1) involved in the TGF-beta receptor signaling pathway were down-regulated after NR0B2 silencing. GDF15, similarly to TGF-beta1, can activate a variety of serine/threonine protein kinase receptors including TGF-beta receptors (28). BAMBI, on the other hand, is a negative regulator of TGF-beta signaling, acting as a pseudoreceptor lacking the intracellular serine/threonine kinase domain (29). The only gene that is regulated by TGF-beta and appears to be down-regulated by NR0B2 silencing is ID1 (30, 31). Thus, our results do not directly contradict previous findings that indicate an in-

hibitory effect of NR0B2 to TGF-beta-mediated transcription; however our data neither support that NR0B2 inhibits TGF-beta mediated transcription. Conversely, our results propose an additional role of NR0B2 in TGF-beta receptor signaling. In particular, it appears that NR0B2 regulates the transcription of ligands as well receptors of the TGF-beta signaling pathway.

In contrast to the several changes that were observed in the transcriptomics level, no significant effects of NR0B2 silencing were observed for apoptotic and DNA damage levels as well as for the cell cycle distribution, as evaluated by phenotypic analyses. These data suggest that at base-line NR0B2 is not involved in the maintenance of DNA integrity or the cell cycle progress under normal conditions in the absence of genotoxic stress.

NR0B2 silencing in combination with chemical treatment dramatically altered the transcriptomics responses of HepG2 cells in comparison to previously conducted chemical treatments of cells expressing NR0B2. Furthermore, the altered gene expressions induced by chemical exposures of non-targeting siRNA transfected cells are also noticeably reduced in comparison to those of untransfected cells (Figure 3), suggesting that the transfection process as such inhibits the response of HepG2 cells to BaP and to a lesser extent to 2CMP. Furthermore, the transfection process does not seem to have a major effect on the transcriptomics responses of cells under non-chemical treated conditions (Figure 1C, 1D). Thus, it appears that the transfection process mainly interferes and inhibits the transcriptomics responses of HepG2 cells to chemicals.

Nevertheless, both BaP and 2CMP upon challenging NR0B2-silenced HepG2 cells were again correctly classified as genotoxic and non-genotoxic compounds, respectively, by the transcriptomics-based prediction model, indicating that NR0B2 has no collective effect on the expression levels of the other classifying genes. This further indicates that while 33 genes were selected for the classifying gene-set, there might be no biological interplay among them at the level of gene expression regulated by a pivotal role of NR0B2.

Due to the effect of the transfection process on the responses of HepG2 cells to chemical treatments, the loss of many DEGs in response to BaP and 2CMP after NR0B2 silencing cannot be directly assigned to the loss of NR0B2 expression. Therefore, gene expression changes directly linked to NR0B2 silencing can be considered only those that are present in chemical-treated cells upon NR0B2 siRNA transfection. These include 8 and 70 altered gene expressions after BaP and 2CMP treatments of HepG2 cells, respectively.

Due to the small number of identified genes whose expression is altered by chemical-exposure in HepG2 cells not expressing NR0B2, in combination with the observation that there are no known interactions between these genes and NR0B2, we propose that these gene expression changes may not be important or not actually linked to NR0B2 silencing. Therefore, we further evaluated which of the DEGs appear to have statistically significant different expression values in chemical-treated NR0B2 knockdown HepG2 cells and chemical-treated NR0B2 expressing HepG2 cells. Among the DEGs identified for BaP-treated NR0B2 knockdown cells, only the expression of PLA2G1B was undetectable in the BaP-treated HepG2 cells expressing NR0B2. This gene codes for one of the isoforms of phospholipase A2 that has been linked to a variety of cellular processes including cell proliferation, hormone

release and cytokine production (32). Among the rest of the genes, only three – hemopexin, oncostatin M receptor and Mannose-binding lectin 2 (MBL2) – showed statistically significantly ($p < 0.05$) altered expression after BaP-treatment of NR0B2 knockdown cells in comparison to BaP-treatment of NR0B2 expressing cells (Table 2). Hemopexin is known to protect cells from oxidative stress (33), while OSMR and MBL2 are involved, among others, in immune responses (34, 35). These observations thus suggest that most likely the expression of PLA2G1B, hemopexin, OSMR and MBL2 is due to the silencing of NR0B2.

Similarly, for 2CMP-treated cells, the expression levels of 9 genes were not detected in cells expressing NR0B2; while 20 genes showed statistically significant different expression values ($p < 0.05$) between NR0B2 knockdown and NR0B2 expressing cells (Table 3). Therefore, we suggest that the expression changes observed for these 29 genes may indeed be due to the silencing of NR0B2 prior to exposure to 2CMP. This is further supported by the fact that most of these genes are involved in transport, metabolic processes and transcription, processes that are known to be affected by NR0B2 (13).

In contrast to the limited transcriptomics effects of BaP-exposure after NR0B2 silencing, significant alterations were observed regarding the progression through the cell cycle and the apoptotic levels of BaP-treated cells (Figure 4). In particular, BaP-treatment of NR0B2 knockdown cells resulted in a decrease of cells in S-phase and in an increase of cells in G2/M phase as well as in a reduction of the apoptotic cells compared to BaP-treatment in NR0B2 expressing HepG2 cells. However, the BaP-induced DNA damage levels were similar in NR0B2 knockdown and NR0B2 expressed cells. Taking all these in consideration, we propose that under genotoxic stress NR0B2 may play a role in inhibiting the transition from S- to G2/M-phase, and therefore may be important for the maintenance of the cells in S-phase in order to repair the induced DNA damage as well as in the induction of apoptosis when the damage is not repaired. However, NR0B2 does not play a role in the protection of cells from DNA damage. Previous studies on the role of NR0B2 in induction of apoptosis show contradicting results, with two studies supporting a protective effect of NR0B2 against Fas-mediated and etoposide-induced apoptosis (36, 37), while another study supports that NR0B2 is necessary for the induction of apoptosis by two known inducers of apoptosis of malignant cells (38). Therefore, NR0B2 may have different roles in apoptosis depending on the causes that induced apoptosis.

While no significant alterations were observed for the apoptotic and the DNA damage levels, different cell cycle distribution patterns were detected for 2CMP-exposed cells not expressing NR0B2 in comparison to 2CMP-exposed cells expressing NR0B2. In particular, 2CMP exposures of cells not expressing NR0B2 resulted in an additional increase of cells in S-phase and a decrease of cells in G2/M-phase, suggesting an opposite effect of NR0B2 silencing than the one observed for BaP exposures. These results further support a role of NR0B2 in the S- to G2/M transition. However, while after genotoxic treatment NR0B2 appears to delay this transition, after NGTX treatment it appears that NR0B2 promotes this transition. This further underlines that the role of NR0B2 is dependent on the external stimulus, as described for its role in apoptosis induction.

In summary, in this study we show that knockdown of NR0B2 is not sufficient to affect the genotoxicity prediction by the transcriptomics-based classification approach that was recently developed. However, NR0B2 appears to play an important role in several biological processes of HepG2 cells. Furthermore, NR0B2 knockdown alters the responses of HepG2 cells after genotoxic stress and appears to play a dual role in the cell cycle progress and in the induction of apoptosis depending on the external stimulus, as evaluated by phenotypic analysis.

In conclusion, we show that one of the genes that were selected based on statistical methods for genotoxicity prediction (Chapter 4), is indeed important for genotoxic responses. However, the prediction model can also perform accurately in its absence. Apparently, excluding only one gene is not sufficient for altering the classification results. This may also suggest that other genes from the classifying gene set may have a higher contribution to the prediction model and thus knockdown of their expression may lead to incorrect classification.

References

1. Jemal A, Siegel R, Xu J, & Ward E (2010) Cancer statistics, 2010. *CA Cancer J Clin* 60(5):277-300.
2. OECD (2010) Mortality from Cancer, in OECD, Health at a Glance: Europe 2010 *OECD Publishing*:34-35.
3. Petrucelli N, Daly MB, & Feldman GL (2010) Hereditary breast and ovarian cancer due to mutations in BRCA1 and BRCA2. *Genet Med* 12(5):245-259.
4. Clapp RW, Jacobs MM, & Loechler EL (2008) Environmental and occupational causes of cancer: new evidence 2005-2007. *Rev Environ Health* 23(1):1-37.
5. Oliveira PA, *et al.* (2007) Chemical carcinogenesis. *An Acad Bras Cienc* 79(4):593-616.
6. Ames BN, Lee FD, & Durston WE (1973) An improved bacterial test system for the detection and classification of mutagens and carcinogens. *Proc Natl Acad Sci U S A* 70(3):782-786.
7. Kirkland D, Aardema M, Henderson L, & Möller L (2005) Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens I. Sensitivity, specificity and relative predictivity. *Mutat Res*. 584(1-2):1-256.
8. Johansson L, *et al.* (2000) The orphan nuclear receptor SHP utilizes conserved LXXLL-related motifs for interactions with ligand-activated estrogen receptors. *Mol Cell Biol* 20(4):1124-1133.
9. Sanyal S, *et al.* (2002) Differential regulation of the orphan nuclear receptor small heterodimer partner (SHP) gene promoter by orphan nuclear receptor ERR isoforms. *J Biol Chem* 277(3):1739-1748.
10. Goodwin B, *et al.* (2000) A regulatory cascade of the nuclear receptors FXR, SHP-1, and LXR-1 represses bile acid biosynthesis. *Mol Cell* 6(3):517-526.
11. Lee YK, Dell H, Dowhan DH, Hadzopoulou-Cladaras M, & Moore DD (2000) The orphan nuclear receptor SHP inhibits hepatocyte nuclear factor 4 and retinoid X receptor transactivation: two mechanisms for repression. *Mol Cell Biol* 20(1):187-195.
12. Zhang Y, Hagedorn CH, & Wang L (2011) Role of nuclear receptor SHP in metabolism and cancer. *Biochim Biophys Acta* 1812(8):893-908.
13. Lee YS, Chanda D, Sim J, Park YY, & Choi HS (2007) Structure and function of the atypical orphan nuclear receptor small heterodimer partner. *Int Rev Cytol* 261:117-158.
14. He N, *et al.* (2008) Epigenetic inhibition of nuclear receptor small heterodimer partner is associated with and regulates hepatocellular carcinoma growth. *Gastroenterology* 134(3):793-802.
15. Watters GP, Smart DJ, Harvey JS, & Austin CA (2009) H2AX phosphorylation as a genotoxicity endpoint. *Mutat Res* 679(1-2):50-58.
16. Skladanowski A, Bozko P, & Sabisz M (2009) DNA structure and integrity checkpoints during the cell cycle and their role in drug targeting and sensitivity of tumor cells to anticancer treatment. *Chem Rev* 109(7):2951-2973.
17. Jennen DG, *et al.* (2010) Comparison of HepG2 and HepaRG by whole genome gene expression analysis for the purpose of chemical hazard identification. *Toxicol Sci*.
18. Dai M, *et al.* (2005) Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res* 33(20):e175.
19. Irizarry RA, *et al.* (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4(2):249-264.
20. De Groot PJ, Reiff C, Mayer C, & Muller M (2008) NuGO contributions to GenePattern. *Genes Nutr* 3(3-4):143-146.
21. Magkoufopoulou C, Claessen SM, Jennen DG, Kleinjans JC, & van Delft JH (2011) Comparison of phenotypic and transcriptomic effects of false-positive genotoxins, true genotoxins and non-genotoxins using HepG2 cells. *Mutagenesis*.
22. Dennis G, Jr., *et al.* (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 4(5):P3.
23. Staal YC, *et al.* (2007) Binary PAH mixtures cause additive or antagonistic effects on gene expression but synergistic effects on DNA adduct formation. *Carcinogenesis* 28(12):2632-2640.
24. Boulias K, *et al.* (2005) Regulation of hepatic metabolic pathways by the orphan nuclear receptor SHP. *Embo J* 24(14):2624-2633.
25. Lu TT, *et al.* (2000) Molecular basis for feedback regulation of bile acid synthesis by nuclear receptors. *Mol Cell* 6(3):507-515.
26. Jung GS, *et al.* (2009) The orphan nuclear receptor SHP attenuates renal fibrosis. *J Am Soc Nephrol* 20(10):2162-2170.
27. Suh JH, *et al.* (2006) Orphan nuclear receptor small heterodimer partner inhibits transforming growth factor-beta signaling by repressing SMAD3 transactivation. *J Biol Chem* 281(51):39169-39178.

28. Mimeault M & Batra SK (2010) Divergent molecular mechanisms underlying the pleiotropic functions of macrophage inhibitory cytokine-1 in cancer. *J Cell Physiol* 224(3):626-635.
29. Onichtchouk D, *et al.* (1999) Silencing of TGF-beta signalling by the pseudoreceptor BAMBI. *Nature* 401(6752):480-485.
30. Liang YY, Brunicardi FC, & Lin X (2009) Smad3 mediates immediate early induction of Id1 by TGF-beta. *Cell Res* 19(1):140-148.
31. Spender LC & Inman GJ (2009) TGF-beta induces growth arrest in Burkitt lymphoma cells via transcriptional repression of E2F-1. *J Biol Chem* 284(3):1435-1442.
32. Hanasaki K & Arita H (2002) Phospholipase A2 receptor: a regulator of biological functions of secretory phospholipase A2. *Prostaglandins Other Lipid Mediat* 68-69:71-82.
33. Tolosano E & Altruda F (2002) Hemopexin: structure, function, and regulation. *DNA Cell Biol* 21(4):297-306.
34. Dommett RM, Klein N, & Turner MW (2006) Mannose-binding lectin in innate immunity: past, present and future. *Tissue Antigens* 68(3):193-209.
35. Hummel M, *et al.* (2009) Intra-graft TNF receptor signaling contributes to activation of innate and adaptive immunity in a renal allograft model. *Transplantation* 87(2):178-188.
36. Yeo MG, Yoo YG, Choi HS, Pak YK, & Lee MO (2005) Negative cross-talk between Nur77 and small heterodimer partner and its role in apoptotic cell death of hepatoma cells. *Mol Endocrinol* 19(4):950-963.
37. Kim K, Choi YH, Kim HH, & Cheong J (2009) The orphan nuclear receptor SHP inhibits apoptosis during the monocytic differentiation by inducing p21WAF1. *Exp Mol Med* 41(6):429-439.
38. Farhana L, *et al.* (2007) Adamantyl-substituted retinoid-related molecules bind small heterodimer partner and modulate the Sin3A repressor. *Cancer Res* 67(1):318-325.

Chapter 6

Transcriptomics-based evaluation in HepG2 cells of human carcinogenic potential of known rodent non-genotoxic carcinogens

In preparation

C. Magkoufopoulou,
M. Tsamou, D.G.J. Jennen,
J.C.S. Kleijnans, J.H.M van Delft

Abstract

The lack of available methods for the identification of the carcinogenic potential of non-genotoxic chemicals for humans, leads to the testing of chemicals in the *in vivo* rodent bioassays or in the recently developed *in vitro* cell transformation assay on embryonic mouse fibroblasts. These assays have limited relevance to human health risk; thereby emphasizing that an alternative, more relevant to human health, *in vitro* carcinogenicity assessment assay is needed. Previously, we developed a transcriptomics-based method using HepG2 cells for *in vivo* genotoxicity assessment. Thus, we now hypothesize that transcriptomics responses in HepG2 cells can be used to assess the carcinogenic potential of non-genotoxic chemicals for humans.

This study, thus, aimed at evaluating the cancer risks for humans of rodent non-genotoxic carcinogens by means of investigating transcriptomics responses in HepG2 cells. Our analyses included four approaches: sscMap and PAM, two statistical tools, and functional annotation as well transcriptional regulation analyses based on the biological interpretation of the transcriptomics responses. Gene expression changes induced in HepG2 cells after 24h and 48h of exposure to seven rodent carcinogens were compared to those induced by eight human carcinogens and six non-carcinogens. Although the statistical tools gave clear results regarding the potential of rodent carcinogens to be a risk for humans, the results were contradicting between the two approaches. Using the biological based analyses, inconclusive results were obtained for several rodent carcinogens. Therefore, these approaches did not allow us to evaluate the cancer risk for humans of the rodent carcinogens in a consistent manner. Nevertheless, the transcriptomics-based approaches in HepG2 cells confirmed the mode of actions of some known human carcinogens.

Introduction

With cancer being one of the most common causes of death worldwide, accounting for approximately 25% of total deaths in western world (1, 2), a lot of effort is made to prevent its occurrence. While several types of cancer are inherited and their occurrence cannot be avoided, such as breast cancer due to inheritance of BRCA1 and BRCA2 mutations (3), the cause of most cancer cases is the exposure to exogenous carcinogenic factors. Such as tobacco smoke that can lead to the development of lung cancer (4). Thus, by correctly identifying the carcinogenic chemicals and by preventing exposure to those chemicals, several cancer cases may be prevented (4).

More than 90% of the chemicals classified by the International Agency for Research on Cancer (IARC) as human carcinogens result, directly or through their metabolites, in the induction of DNA damage that may lead to mutations within genes responsible for cell proliferation or survival, and therewith to the malignant transformation of cells (5). The remaining 10% of carcinogenic chemicals, however, act through a variety of mechanisms that do not induce DNA damage, such as endocrine modulation, tumor promotion, induction of tissue-specific toxicity, inflammatory responses, immunosuppression, or inhibition of gap junction intercellular communication (GJIC) (6). Consequently, chemical carcinogens are classified as genotoxic (GTX) and non-genotoxic (NGTX) (5).

For the identification of GTX compounds several *in vitro* assays have been developed, including the bacterial mutagenicity assay (Ames assay) (7), mammalian micronuclei (MN), chromosomal aberration (CA) and mouse lymphoma (MLA) assays (8-10). Recently, we developed a transcriptomics-based assay that results in a higher accuracy of the prediction of *in vivo* genotoxic properties in comparison to the conventionally used genotoxicity assays (Chapter 4). Since most *in vivo* GTX compounds are also carcinogenic, the confirmation of their *in vivo* genotoxic properties is sufficient to classify these chemicals as potent carcinogens (EC 1907/2006). However, the existence of NGTX carcinogens that cause cancer through a variety of mechanisms, makes the prediction of carcinogenic properties of non-genotoxic chemicals solely by means of *in vitro* or *in vivo* genotoxicity assays unsuccessful (11). Therefore, long-term rodent bioassays are required for the evaluation of the carcinogenic properties of chemicals (EC 1907/2006), (12).

The characterization of the mode-of-action of non-genotoxic chemicals is considered to significantly contribute in the understanding of the carcinogenic properties of a chemical and whether it may or may not pose a cancer risk to humans (6). The analysis of transcriptomics responses after exposure to toxic agents has been proposed that may assist the characterization of the putative mechanism of action of chemicals (13). Therefore, the field of toxicogenomics has been extensively explored for this purpose. Several studies, both *in vitro* and *in vivo*, indicate that discrimination of non-genotoxic chemicals with different modes of actions is achievable via toxicogenomics approaches (14-21). The majority of these studies, though, uses rodents or rodent cells. This is due to the fact that current guidelines suggest the use of either mice or rats in the classical rodent carcinogenicity bioassay for the carcinogenic evaluation of chemicals.

However, whereas some chemicals can be carcinogenic to both mice and rats, the majority of chemicals tested in rodents cause cancer only in one of the two species. When known human carcinogens were tested in rodents, only 50% were able to induce cancer in both species with a corresponding specificity of 93% to detect human carcinogens, while 90% of known human carcinogens caused tumors in at least one species with a corresponding specificity of 78% (22). This suggests that the use of rodents for carcinogenic evaluation of chemicals cannot fully predict the human cancer risks. In order to overcome the ethical issues raised by the use of experimental animals, the *in vitro* BALB/c 3T3 cell transformation assay has been developed for evaluating the carcinogenic properties of chemicals. This assay measures morphological changes induced in embryonic mouse fibroblasts *in vitro* (23), and it has been proposed to be valuable for the detection of rodent carcinogens (24).

The relevance of rodent based assays for the evaluation of human risk has been challenged in view of the differences between rodents and humans in relation to the metabolism of chemicals. Furthermore, several modes of actions have been considered to be rodent specific, leading to the overestimation of human risks (25-28). Therefore a distinction has to be made between human carcinogens and rodent carcinogens.

Based on the above mentioned, this study aims to identify among rodent non-genotoxic carcinogens (RC) those that may induce cancer risk for humans by analysis of transcriptomics responses. HepG2 cells were selected as an appropriate cellular system since the liver is considered to be the major target organ of chemical carcinogens (29), and since we previously showed that transcriptomics responses of HepG2 cells after chemical exposures can be used for class discrimination (Chapter 4). For the purpose of the current study, a comparison was conducted between the induced transcriptomics responses by RC to those induced by non-genotoxic human carcinogens (HC) (i.e. chemicals for which available epidemiological data exist supporting increase cancer risk to humans after exposure) and non-carcinogens (NC). To this end, two different approaches that use statistical algorithms and two different approaches that study the underlying biological mechanisms were applied.

Material and methods

Microarray data selection and processing

Microarray data from HepG2 cells after 24h and 48h of exposures to 63 chemicals are accessible at NCBI GEO database (30), accession GSE28878 (Chapter 4). For these exposures three biological replicates of a single dose per compound and their vehicle controls are available.

From these data, we selected 21 non-genotoxic compounds (Table 1), namely those with clear negative results in the micronucleus or chromosomal aberration *in vivo* assays, and divided them in three groups based on the following criteria:

Group 1 - Human Carcinogens (HC): For the chemicals included in this group, epidemiological data indicating increased human cancer risk according to IARC (www.iarc.fr) is available. (Table 1)

Table 1: Chemicals and information on their genotoxic and carcinogenic properties

Compound				GTX <i>in vivo</i> ¹			IARC ²		Types of tumors induced			References	
Name	Abbr.	CAS no.	Dose	Solvent	MN	CA	Human		Mice		Rats	References	
HC	2,3,7,8-tetrachloro dibenzo-para-dioxin	TCDD	1746-01-6	10 nM	DMSO	-	-	1	combined sites	multiple sites	multiple sites	(61, 65)	
	Cyclosporin A	CsA	59865-13-3	3 μ M	DMSO	-	-	1	lymphomas, connective tissue, squamous cell carcinoma	leukemia	-	(66)	
	17-beta Estradiol	E2	50-28-2	3 μ M	DMSO	-	NT	1	breast, endometrial cancers	Mammary, pituitary, uterine, cervical, vaginal, lymphoid, testicular tumors	Mammary, pituitary tumors	(64, 67)	
	Tetrachloro ethylene	TCE	127-18-4	2 mM	EtOH	NT	-(humans)	2A	oesophageal, cervical cancer, non-Hodgkin's lymphoma	hepatocellular carcinomas	leukemia	(68)	
	Pentachlorophenol	PCP	87-86-5	10 μ M	EtOH	-(amphibian)	-(humans)	2B	soft-tissue sarcoma, non-Hodgkin's lymphoma	NT	mesotheliomas of the tunica vaginalis	(69)	
	1,1,1-trichloro-2,2-di-(4-chlorophenyl)ethane	DDT	50-29-3	2 mM	DMSO	NT	-	2B	lung, non-Hodgkin's lymphoma	liver cell tumors, lung carcinomas, lymphomas	liver tumors	(63)	
	Lindane	HCH	58-99-9	2 mM	DMSO	NT	-(humans,mice)	2B	leukemia, aplastic anaemia	liver tumors	thyroid tumors	(67)	
	Phenobarbital	PhB	50-06-6	1 mM	DMSO	-	-	2B	lung, ovary, gall-bladder cancers	hepatocellular adenomas and carcinomas	hepatocellular adenomas	(56)	
	Bromodichloro-methane	BDCM	75-27-4	2 mM	DMSO	-	NT	2B	ND	kidney, hepatocellular carcinomas	kidney, large intestine	(69)	
	Ethylacrylate	EtAc	140-88-5	1 mM	EtOH	-	-	2B	ND	forestomach (squamous cells)	forestomach (squamous cells)	(67)	
RC	Ethylbenzene	EthylB	100-41-4	800 μ M	DMSO	-	NT	2B	ND	lung, liver	kidney	(70)	
	Nitrobenzene	NBZ	98-95-3	2 mM	DMSO	NT	-	2B	ND	alveolar-bronchiolar neoplasms, thyroid tumors	hepatocellular neoplasms, kidney tumors, endometrial tumors	(71)	
	Carbon Tetrachloride	CCl4	56-23-5	2 mM	DMSO	-	NT	2B	ND	liver tumors	liver tumors, mammary adenomas	(69)	
	di(2-ethylhexyl) phthalate	DEHP	117-81-7	10 mM	DMSO	-	-	3	ND	hepatocellular tumors	hepatocellular tumors	(70)	
	Ampicillin	AmpC	7177-48-2	250 μ M	DMSO	-	NT	3	ND	benign lung tumors	leukemia,phaeochromocytomas	(66)	
	Caprolactam	CAP	105-60-2	2 mM	DMSO	-	-	4	ND	non-carcinogenic	non-carcinogenic	(69)	
	Resorcinol	RR	108-46-3	2 mM	EtOH	-	NT	3	ND	non-carcinogenic	non-carcinogenic	(69)	
	4-Nitro-1,2-benzene-diamine	NPD	99-56-9	2 mM	DMSO	-	-	3	ND	non-carcinogenic	non-carcinogenic	(72)(NTP)	
	Sulfisoxazole	Sulfi	127-69-5	5 μ M	DMSO	-	-	NA	ND	non-carcinogenic	non-carcinogenic	(NTP)	
	D-mannitol	Dman	69-65-8	250 μ M	PBS	-	-	NA	ND	non-carcinogenic	non-carcinogenic	(NTP)	
NC	2-(chloromethyl)pyridine.HCl	2CMP	6959-47-3	300 μ M	DMSO	-	-	NA	ND	non-carcinogenic	non-carcinogenic	(NTP)	

¹ GTX *in vivo*: data from *in vivo* genotoxicity testing, unless mentioned otherwise the data are from studies conducted in rodents; ² IARC: Classification of chemicals according to IARC: Group 1: Carcinogenic to humans, Group 2A: probably carcinogenic to humans, Group 2B: possibly carcinogenic to humans, Group 3: not classifiable as to its carcinogenicity to humans; Group 4: probably not carcinogenic to humans. NT: not tested; ND: no data available indicating carcinogenic potential for this chemical.

Group 2 - Rodent Carcinogens (RC): The chemicals in this group induce tumors to both rodent species (mice and rats) based on data from IARC. When no IARC monographs were available, data from NTP were taken into consideration. For these chemicals human epidemiological data are not available.

Group 3 - Non Carcinogens (NC): Chemicals included in this group are selected based on data from IARC showing no carcinogenicity to both mice and rats. When no IARC monographs were available, data from NTP were taken into consideration. Also for these chemicals there are no human epidemiological data available.

Filtering of microarray data

Microarray data were processed and filtered as previously described (31). For the selection of genes by t-tests for PAM and connectivity analysis the union of the filtered datasets was used (11855 genes for 24h and 12480 genes for 48h). For functional annotation and transcriptional regulation analysis the filtered datasets per compound were used.

Selection of differentially expressed genes (DEGs)

Differentially expressed genes (DEGs) per compound were selected from the filtered datasets for each compound using the criteria: absolute average fold change > 1.5 , absolute fold change > 1.2 for at least two replicates in comparison to control treatments and same direction of expression for all three replicates (31).

Carcinogenicity evaluation of RC

Four different approaches were applied as described below.

Approach 1: Classification analysis

PAM (Prediction Analysis for Microarrays 2.13, Stanford University Software) (32) was used for classification as previously described (Chapter 4). As classifiers, genes with significantly different expression values ($p < 0.01$) between HC and NC exposures were selected following a leave-one-out procedure. The data from the HC and NC exposures were used as training set to generate a predictive model that was used to classify all the compounds. A compound was predicted to be HC or NC when at least two out of the three replicates were assigned to that class.

Approach 2: Connectivity analysis

A connectivity mapping approach (33) was done to investigate similarities in gene expression profiles between the various compounds, using the sscMap program (34). The union of the DEGs after 24h or 48h of exposure to the HC was compared with the union of the DEGs after 24h or 48h of exposure to NC, respectively. The commonly affected DEGs were excluded. The DEGs that were specifically affected by HC or NC exposures were combined and used to construct a custom-reference expression database. Using sscMAP, connections for each of the 21 compounds to all the reference compounds were calculated. The connection strength between each compound to each of the reference compounds is regulation status-dependent and ranges from -100% (absolute negative connectivity) to +100% (absolute positive connectivity). For a connectivity to be considered statistically significant a p-value threshold of 0.01 was used. For the connection between two compounds

to be considered significant, the connectivity in both directions should be significant and both connectivities should be either positive or negative. For visualization of the results, Cytoscape 2.8.0 (www.cytoscape.org) was used.

Approach 3: Functional annotation analysis

Functional Annotation based on Gene Ontology (GO) Biological Processes (GOTERM_BP_FAT) was done using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (35) in order to identify which processes are positively or negatively affected by the different compounds. For each compound the Functional Annotation was conducted for the up-regulated and the down-regulated DEGs separately, using the corresponding filtered datasets as background. GO-terms that appeared in common between the up- and down-regulated DEGs of a compound were excluded. GO terms were considered significant when $EASE < 0.1$ and $p\text{-value} < 0.05$ and included more than four genes. The union of GO-terms after exposures to HC was compared with the union of those identified after exposures to NC for both the up- and down-regulated genes separately. Biological processes affected by both HC and NC compounds were excluded in order to identify those that are specifically affected by either HC or NC. The GO-terms resulting from exposure to RC were compared with the identified GO-terms after HC or NC exposures in order to reveal similarities at the level of affected biological processes.

Approach 4: Activation of nuclear receptors and transcriptional regulation analysis

For the identification of nuclear receptors that are activated after exposure of HepG2 cells to the different compounds after 24h or 48h of exposure, Transcription Regulation Network Analysis was conducted in the MetaCore Network Tools (GeneGo, San Diego, CA) for the up- and down-regulated DEGs after exposure to each compound separately. Nuclear Receptors that appeared with the analyses of both up- and down-regulated DEGs after exposure to one chemical were excluded. Transcription regulation networks were taken into account when at least three DEGs were included in the network. Nuclear receptors identified for both HC and NC exposures were excluded in order to identify those that are specifically affected by either HC or NC compounds. The nuclear receptors activated by exposure to RC were compared with those activated after exposure to HC and NC in order to reveal similarities at the level of activated nuclear receptors.

For the identification of active transcription factors, the Transcription Regulation Network Analysis was conducted following the same procedure as described for the activation of nuclear receptors. The down-regulated DEGs were also analyzed for the identification of inhibited transcription factors.

Results

Approach 1: Classification analysis

243 and 413 genes were selected for class prediction after 24h and 48h of exposure, respectively. The prediction models developed in PAM resulted in the correct classification of 7 out of 8 HC for both 24h and 48h of exposure, with PhB misclassified at both time points. The classification of NC varied notably for the different time points. After 24h of exposure,

half of the NC – CAP, Sulfi and Dman – were misclassified as HC, whereas after 48h only CAP was misclassified as a HC.

Among the RC, BDCM, EthylB, NBZ and CCl₄ were predicted as HC at both time points, whereas EtAc was predicted as NC in both time points. Different results were obtained for DEHP and AmpC for the different time points, with DEHP being predicted as HC after 24h of exposure and AmpC after 48h of exposure (Table 2).

Table 2: Class Prediction Results

HC			NC			RC		
Compound	24h	48h	Compound	24h	48h	Compound	24h	48h
TCDD	HC	HC	CAP	HC*	HC*	BDCM	HC*	HC
CsA	HC	HC	RR	NC	NC	EtAc	NC	NC
E2	HC	HC	NPD	NC	NC	EthylB	HC*	HC
TCE	HC*	HC	Sulfi	HC*	NC	NBZ	HC*	HC
PCP	HC	HC*	Dman	HC*	NC*	CCl ₄	HC*	HC
DDT	HC	HC	2CMP	NC	NC*	DEHP	HC*	NC*
HCH	HC	HC				AmpC	NC*	HC
PhB	NC*	NC*						

*: indicates that only 2 from the three replicates of these compound were classified to the indicated class; Results in bold indicate misclassification

Approach 2: Connectivity analysis

For a better understanding of the connections of the transcriptomics responses in HepG2 cells after exposure to the different compounds, the sscMap tool was used (Figure 1). For 24h of exposure no connections could be identified for the NC compounds Dman and Sulfi due to the fact that Dman exposures altered the expression of only one gene, while no DEGs were identified after exposures to Sulfi. Regarding the RC compounds, CCl₄ and AmpC show no significant connections with any of the HC or the NC compounds. These may be due to the limited number of DEGs after 24h treatments with these compounds (5 and 4, respectively). BDCM appears to be positively connected with the NC compound CAP, while EtAc is positively connected with the NC compound RR and negatively connected with four HC compounds. Therefore BDCM and EtAc probably do not act via a similar mode of action with the HC compounds used in this study. DEHP appears to be strongly connected with the HC compounds DDT and HCH and thus may share a similar mode of action with these HC. Finally EthylB appears to be strongly connected only with the RC compound NBZ, and therefore no conclusion can be derived; while NBZ shows positive connectivity with both HC and NC and no conclusion can be made.

In Figure 1B all the significant connections after 48h of treatment are represented. The HC compound TCE resulted in the altered expression of only one gene in HepG2 cells after 48h of treatment, thus no connections were identified for this compound. Despite the 27 DEGs identified after exposures to the HC compound PCP, no significant connections were identified with any of the other compounds. Regarding the RC compounds, strong positive connections have been identified for EthylB, CCl₄, NBZ, BDCM and the NC compound CAP, indicating a different mode of action of these compounds to the HC used in this study. While the RC compound EtAc is negatively connected with

CAP and the HC compounds CsA, DDT, TCDD and HCH, indicating that also EtAc does not share a similar mode of action with these HC. DEHP shows positive connectivity with the HC compounds HCH, DDT and PhB but also with the NC compounds 2CMP, NPD and Sulfi. However, 2CMP and Sulfi are also positively connected with the HC compound HCH. Thus, it is not possible to make any conclusion whether DEHP may be carcinogenic for humans. Lastly, AmpC did not show any significant connectivity with any of the compounds, that may be due to the low number of DEGs (15) after exposure for 48h and thus no conclusion can be made.

Figure 1: Connections among the different compounds. All significant dual connections are presented. Positive connections are indicated with blue arrows and negative with red arrows. The highest the connectivity scores, the thickest the line of the arrows. The class of the different compounds is indicated by the different shapes: HC: square, NC: circle and RC: hexagon. A. After 24h of exposure. B. After 48h of exposure.

Approach 3: Functional annotation analysis

Table 3 and Table 4 show the number of differentially expressed genes involved in the biological processes specifically affected after treatment with either HC or NC, as well as which RC affect the same processes for 24h and 48h of exposure respectively.

After 24h of exposure the HC compounds E2 and PhB as well as the NC compounds Sulfi and Dman did not significantly affect any biological processes. Processes identified to be specifically affected by HC include cell adhesion, signaling, homeostasis and leukocyte activation for the up-regulated genes and DNA damage and cell cycle for the down-regulated genes (Table 3A). Among the RC exposures, only DEHP and NBZ affect several of these processes, while EtAc only affects placenta development similarly to exposures to the HC compound HCH. Processes affected specifically by exposures to NC compounds include signaling, metabolic processes and apoptosis and were mainly identified after exposures to RR (Table 3B). These processes are also affected by exposures to EtAc. DEHP exposures for 24h also alter the expression of genes involved in response to abiotic stimulus and signaling similarly to NC compounds, while BDCM shares only one process (transcription) in common with the NC compound RR.

The functional annotation analysis of transcriptomics responses after 24h of exposure, therefore, indicates that NBZ may act via a similar mode of action with the selected HC com-

pounds, while BDCM only affects processes that are specifically affected by NC and thus may be non carcinogenic to humans. EtAc and DEHP affect processes that were identified for both HC and NC exposures thus no conclusion can be made for these compounds. Lastly, no results were obtained for the RC compounds EthylB, CCl4 and AmpC since no biological processes were identified to be affected by these chemicals.

Table 3A: GO-terms identified to be specifically affected by HC after 24h of exposure compared to RC*

HC(+)								
Gene Ontology - Biological process	HC					RC		
	TCDD	CsA	PCP	DDT	HCH	EtAc	NBZ	DEHP
cell adhesion	19	-	5	-	41	-	-	14
cell motion	-	-	-	11	37	-	-	11
leukocyte activation	-	-	-	-	23	-	-	6
transmembrane receptor protein tyrosine kinase signaling pathway	-	-	4	-	17	-	-	6
cell morphogenesis involved in differentiation	-	-	-	-	16	-	-	7
negative regulation of transport	-	-	-	-	12	-	-	6
cytosolic calcium ion homeostasis	-	-	-	-	9	-	-	4
negative regulation of signal transduction	-	-	-	8	-	-	-	6
protein homooligomerization	-	-	-	5	-	-	-	5
placenta development	-	-	-	-	8	7	-	-

HC(-)								
Gene Ontology - Biological process	HC					RC		
	TCDD	CsA	PCP	DDT	HCH	EtAc	NBZ	DEHP
cell cycle	-	25	-	61	179	-	26	-
DNA metabolic process	13	23	-	20	128	-	11	-
M phase	-	17	-	42	118	-	20	-
nuclear division	-	14	-	35	90	-	16	-
response to DNA damage stimulus	-	11	-	18	84	-	10	-
microtubule-based process	-	8	-	22	52	-	8	-
DNA repair	-	-	-	-	69	-	8	-
cell cycle checkpoint	-	-	-	11	31	-	5	-
DNA recombination	-	-	-	-	26	-	7	-

Table 3B: GO-terms identified to be specifically affected by NC after 24h of exposure compared to RC*

NC(+)						
Gene Ontology - Biological process	NC			RC		
	RR	NPD	2CMP	BDCM	EtAc	DEHP
response to abiotic stimulus	-	16	35	-	-	8
G-protein coupled receptor protein signaling pathway	20	-	-	-	-	9
mRNA metabolic process	38	-	-	-	43	-
M phase	33	-	-	-	33	-
microtubule-based process	26	-	-	-	27	-
phosphoinositide-mediated signaling	10	-	-	-	8	-

NC(-)						
Gene Ontology - Biological process	NC			RC		
	RR	NPD	2CMP	BDCM	EtAc	DEHP
glycerol metabolic process	-	-	6	-	7	-
regulation of transcription	240	-	-	11	238	-
protein kinase cascade	35	-	-	-	35	-
response to unfolded protein	17	-	-	-	20	-
activation of pro-apoptotic gene products	9	-	-	-	7	-

*: Compounds that did not affect any of the indicated processes were not included in the table; Only GO-terms that are affected by at least one RC compound are presented; numbers indicate the number of DEGs involved in the corresponding treatment and GO-term. -: the indicated GO-term was not affected by the corresponding treatment; (+): positively affected biological processes; (-): negatively affected biological processes.

At 48h TCE-exposures resulted in the up-regulation of only one gene. Therefore no biological processes were identified to be affected after TCE exposure. TCDD, DDT and HCH-exposures showed the highest overlap for positively affected biological processes (Table 4A). Positively affected biological processes after exposure to HC include regulation of axonogenesis, blood vessel morphogenesis, cellular metal ion homeostasis and regulation of cell morphogenesis involved in differentiation. From the RC, only DEHP exposure positively affected the same processes. Gene expressions involved in macromolecular complex subunit organization were altered by exposures to the HC compound PhB and the RC compound EtAc. The functional annotation analysis of the down-regulated DEGs did not result in the identification of any biological processes after PCP exposure. The highest overlap was observed among CsA, DDT and HCH exposures, mainly for cell division and DNA repair processes (Table 4A). These processes were also negatively affected by the RC compound BDCM. Negatively affected processes related to response to stimulus were identified after exposures to the HC compounds E2 and PhB and the RC compound EtAc. Processes that were positively affected by exposures to NC compounds included cell cycle and DNA repair and were also affected by the RC compound EtAc. Only one biological process, regulation of cellular catabolic process, was identified to be negatively affected by exposures to NC compounds. This process was also affected by exposures to EtAc (Table 4B).

The functional annotation analysis of transcriptomics responses after 48h of exposure, thus, indicates that the RC compound DEHP and BDCM share a similar mode of action with TCDD, CsA, DDT and HCH, whereas EtAc may share a similar mode of action with E2 and PhB. However, EtAc affected several processes that were also affected by NC compounds and thus no conclusion can be made regarding the carcinogenic properties of this compound. Lack of identified biological processes for AmpC, NBZ as well lack of any similarities of EthylB and CCl4 to HC or NC does not allow any conclusion to be made.

Table 4A: GO-terms identified to be specifically affected by HC after 48h of exposure compared to RC*

Gene Ontology - Biological process	HC(+)								
	HC						RC		
	TCDD	CsA	E2	DDT	HCH	PhB	BDCM	EtAc	DEHP
regulation of axonogenesis	7	-	-	8	9	-	-	-	6
blood vessel morphogenesis	14	-	-	30	22	-	-	-	19
cellular metal ion homeostasis	-	5	-	16	14	-	-	-	10
regulation of cell morphogenesis involved in differentiation	9	-	-	13	13	-	-	-	8
macromolecular complex subunit organization	-	-	-	-	-	12	-	43	-
Gene Ontology - Biological process	HC(-)								
	HC						RC		
	TCDD	CsA	E2	DDT	HCH	PhB	BDCM	EtAc	DEHP
cell cycle checkpoint	-	17	-	25	24	-	5	-	-
response to DNA damage stimulus	-	26	-	60	68	-	7	-	-
DNA repair	-	20	-	51	56	-	6	-	-
cellular response to stress	-	30	-	74	79	-	7	-	-
response to endogenous stimulus	-	-	6	-	-	9	-	18	-
response to hormone stimulus	-	-	5	-	-	9	-	17	-
negative regulation of cell proliferation	-	-	-	-	-	6	-	15	-
ion transport	-	-	7	-	-	-	-	25	-

Table 4B: GO-terms identified to be specifically affected by NC after 48h of exposure compared to RC*

	NC(+)					
	NC					RC
Gene Ontology - Biological process	RR	NPD	Sulfi	Dman	2CMP	EtAc
cell cycle	183	107	-	-	-	119
cytoskeleton organization	47	-	44	-	-	39
nucleotide-excision repair	17	11	-	-	-	13
protein polymerization	11	9	-	-	-	9
DNA replication	74	41	-	-	-	51
DNA integrity checkpoint	18	12	-	-	-	7
double-strand break repair	20	16	-	-	-	10
DNA recombination	27	24	-	-	-	12
neural tube development	-	-	-	6	12	6
cell proliferation	53	44	-	-	-	27
	NC(-)					
	NC					RC
Gene Ontology - Biological process	RR	NPD	Sulfi	Dman	2CMP	EtAc
regulation of cellular catabolic process	-	-	7	-	12	7

*: Compounds that did not affect any of the indicated processes were not included in the table; Only GO-terms that are present in at least one RC compound are presented. Numbers indicate the number of genes involved in the corresponding treatment and GO-term; -: the indicated GO-term was not affected by the corresponding treatment; (+) positively affected biological processes; (-) negatively affected biological processes.

Approach 4: Activation of nuclear receptors and transcription regulation network analysis

Taking into consideration that several non-genotoxic carcinogens act via ligand-activated nuclear receptors leading to the transcription of their target genes, the next step in our analysis was to identify those nuclear receptors that are activated after exposure to the different chemicals. For that purpose, transcriptional regulation networks were created in Metacore for DEGs after 24h and 48h of exposure. After 24h of exposure no nuclear receptor networks were identified to be specifically affected by any of the NC compounds, while the activated Nuclear Receptors by HC were not activated by any RC and thus no conclusion can be made for any of the RC compounds.

For 48h of exposure transcriptional regulation analysis for TCE-exposures was not possible. E2 did not result in the activation of any nuclear receptors. CsA, PCP and PhB were the HC compounds that activated nuclear receptors that were not activated by any of the NC (Table 5). Among those were estrogen related receptors. From the RC compounds, EtAc, EthylB and CCl₄ activated nuclear receptors that were also activated by HC. From the NC compounds, Sulfi, RR, NPD and 2CMP did not result in the specific activation of any nuclear receptors. Dman resulted in the activation of the thyroid hormone receptor, whereas CAP activated three nuclear receptors (RAR α , Progesterone Receptor and REV-ERB α) (Table 5). Nuclear receptors that were activated by NC were also activated by the RC compounds EtAc, EthylB, CCl₄ and AmpC. From this analysis no conclusion can be made about the carcinogenic potential of the RC compounds EtAc, EthylB and CCl₄ since they activate nuclear receptors that are activated by both HC and NC. No conclusion can be made also for BDCM, DEHP and NBZ since they do not activate any nuclear receptor specifically affected by either HC or NC. AmpC is the only RC compound that activates a nuclear receptor that

is only activated by the NC compound CAP, and thus AmpC probably acts via a different mechanism than the selected HC compound used in this study and therefore may be non-carcinogenic to humans.

Table 5: Nuclear Receptors mediated transcription after 48h of exposure*

	HC			NC		RC			
	CsA	PCP	PhB	CAP	Dman	EtAc	EthylB	CCI4	AmpC
CAR	-	-	4	-	-	-	3	-	-
VDR	-	-	4	-	-	40	-	-	-
PXR	-	-	4	-	-	-	-	-	-
ROR α	-	3	-	-	-	-	-	-	-
HNF4 α	-	6	24	-	-	-	11	11	-
ESR2	3	-	-	-	-	-	-	-	-
ERR1	-	3	-	-	-	-	4	-	-
ERR3	-	-	3	-	-	-	-	-	-
ESR1	-	-	-	-	-	-	8	-	-
RAR α	-	-	-	4	-	42	3	3	3
Progesterone Receptor	-	-	-	4	-	-	-	-	-
REV-ERB α	-	-	-	4	-	-	5	4	-
TR α	-	-	-	-	13	-	3	-	-
COUP-TF1	-	-	-	-	-	-	-	3	-

*: Compounds that did not induce the transcription of target genes of the nuclear receptors are not included; Numbers indicate the number of up-regulated genes that their expression after the indicated exposure is affected by the indicated nuclear receptor.

Additionally to the regulation of transcription by ligand-activated nuclear receptors, we further identified those transcription factors that may be activated (i.e. their target genes are up-regulated) or inhibited (i.e. their target genes are down-regulated) by exposures to HC or NC. After 24h of exposure three and six transcription factors were identified to be specifically activated or inhibited, respectively, by HC (Table 6A). Altered expression of target genes of these transcription factors was also induced after exposures to the RC compounds NBZ and DEHP. The HC compound PhB did not result in the activation or inhibition of any transcription factor, whereas the identified transcription factors after exposure to the HC compound E2 were not affected by any of the RC compound. From the NC compounds, Sulfi and Dman exposures did not result in any transcriptional regulation networks. Transcription factors specifically activated by NC exposures include AP2-A, HSF1, which were also identified after exposures to HC but with an opposite effect, and NANOG. These transcription factors appear to be activated also after exposures to the RC compounds EtAc and DEHP (Table 6B). Three transcription factors were inhibited after exposure to NC and their target genes were also down-regulated after exposures to the RC compounds BDCM, EthylB and NBZ (Table 6B). The altered transcription regulation networks after exposure to the NC compound NPD were not affected by any of the RC.

Thus, after 24h of exposure the use of transcriptional regulation networks does not allow any conclusion to be made for NBZ and DEHP as they alter networks identified for both HC and NC, while BDCM, EtAc and EthylB affect only transcription factors that are specifically affected by NC compounds and thus may not act via a same mode of action with the selected HC. Finally, no conclusion can be made for AmpC and CCI4 since they do not alter any transcriptional regulation network.

Table 6A: Transcription Regulation Networks by HC and RC after 24h of exposure*

	HC(+)							
	HC						RC	
Transcription Factor	TCDD	CsA	TCE	PCP	DDT	HCH	NBZ	DEHP
AP-1	41	-	4	5	32	-	-	20
SMAD4	-	-	-	-	-	73	-	14
p300	18	-	-	-	-	-	-	13

	HC(-)							
	HC						RC	
Transcription Factor	TCDD	CsA	TCE	PCP	DDT	HCH	NBZ	DEHP
E2F1	-	29	-	-	44	142	12	3
HSF1	-	8	4	4	-	-	8	3
AP-2A	12	8	-	5	-	74	-	3
YY1	-	-	-	-	-	80	-	4
HIF1A	-	-	-	6	39	-	9	-
EGR1	-	-	-	4	-	-	6	-

Table 6B: Transcription Regulation Networks by NC and RC after 24h of exposure*

Table S21 Transcription Regulation Networks by NC and RC after 2 hr of exposure

Transcription Factor	NC(+)							
	NC			RC				
	CAP	RR	2CMP	BDCM	EtAc	EthylB	NBZ	DEHP
AP-2A	-	70	75	-	45	-	-	-
HSF1	-	-	71	-	51	-	-	-
NANOG	-	69	-	-	-	-	-	9
Transcription Factor	NC(-)							
	NC			RC				
	CAP	RR	2CMP	BDCM	EtAc	EthylB	NBZ	DEHP
AP-1	-	88	72	-	-	-	7	-
SP1	7	-	-	8	-	-	-	-
c-Myc	4	-	-	6	-	3	33	-

*: Compounds that did not alter the expression of target genes of the transcription factors are not included; Numbers indicate the number of target genes of the indicated transcription factor that their expression after the indicated exposure is affected; -: the indicated transcription network was not affected by the corresponding treatment; (+) activated transcription factors; (-) inhibited transcription factors.

After 48h of exposure only two transcription factors appeared to be activated only by HC compounds. Both transcription networks were also identified after exposure to the RC compound, DEHP. Three transcription factors appeared to be inhibited only by HC compounds (E2F4, FOXO3A and CREB1). E2F4 transcriptional activity was also inhibited by the RC compound DEHP. The other two transcription factors were also inhibited by the RC compounds BDCM and EthylB (Table 7A).

For the NC compounds six transcription factors appeared to affect the up-regulation of gene expression (Table 7B). Target genes of these transcription factors were also up-regulated by the RC compounds BDCM, EtAc, EthylB, NBZ, CCl4 and DEHP. No transcription factors were identified to affect only the down-regulation of gene expression after NC exposures.

From this analysis no conclusion can be made for the carcinogenic potential of BDCM, EthylB and DEHP due to the fact that they affect transcription factors that are affected by both HC and NC. No conclusion can be made also for AmpC since no transcriptional regulation networks were identified. The fact that EtAc, NBZ and CCl4 exposures resulted in the identification of transcription regulation networks affected also by NC compounds may suggest that these compounds do not pose a cancer risk for humans.

Table 7A: Transcription Regulation Networks after 48h of exposure with HC and NC*

Transcription Factor	HC(+)						RC		
	HC						RC		
	TCDD	CsA	E2	DDT	HCH	PhB	BDCM	EthylB	DEHP
c-Rel	-	5	-	72	69	-	-	-	44
EGFR	41	4	-	-	-	-	-	-	51

Transcription Factor	HC(-)						RC		
	HC						RC		
	TCDD	CsA	E2	DDT	HCH	PhB	BDCM	EthylB	DEHP
E2F4	-	-	-	133	132	3	-	-	10
FOXO3A	-	-	3	-	-	6	3	6	-
CREB1	-	47	10	-	-	-	7	-	-

Table 7B: Transcription Regulation Networks after 48h of exposure with NC and RC*

Transcription Factor	NC(+)						RC					
	NC						RC					
	CAP	RR	Sulfi	Dman	2CMP		BDCM	EtAc	EthylB	NBZ	CCI4	DEHP
c-Jun	7	-	66	14	-		-	-	-	4	-	-
HNF6	-	90	72	-	-		-	-	6	-	-	-
BMAL1	3	-	-	13	-		-	-	4	-	6	-
HIF1A	8	-	-	14	-		7	57	10	-	12	-
NF-Y	-	133	-	-	84		-	-	-	-	-	48
E2F4	-	135	69	-	-		-	84	-	-	-	-

*: Compounds that did not alter the expression of target genes of the transcription factors are not included; Numbers indicate the number of target genes of the indicated transcription factor that their expression after the indicated exposure is affected; -: the indicated transcription network was not affected by the corresponding treatment; (+) activated transcription factors; (-) inhibited transcription factors.

Discussion

In this study we aimed by means of *in vitro* transcriptomics analyses to identify among non-genotoxic RC compounds, those that may pose a cancer risk for humans. Our analyses included two approaches that by means of statistical algorithms show whether a RC compound induces similar transcriptomics responses with HC or NC and two approaches that by means of gene set analysis would reveal similarities on the level of biological effects of a RC compound to HC or NC. PAM and sscMAP, the two statistical tools used in this study, have been proven capable for *in vitro* class discrimination and identification of compounds with similar modes of action, respectively (Chapter 4), (36, 37). These tools, however, provide indirect evidence of the molecular and cellular effects of the compounds under investigation, whereas the use of functional annotation and transcriptional regulation analytical tools provides a more in-depth analysis that can directly link the transcriptomics responses with cellular and biological changes induced by the compounds under investigation (35).

It is apparent that different approaches for the evaluation of the potential cancer risk for humans after exposure to RC give different results (Table 8). In particular, we observe that for both time points of exposure the statistical tools (Approach 1 and Approach 2) give clear results in comparison to the evaluations based on cellular and molecular effects (Approach 3 and 4). Furthermore, using the biology-based approaches inconclusive results were obtained for several compounds. By taking into consideration over all approaches only the conclusive results, we can provide estimations regarding the cancer risk for humans for exposures to five RC.

Table 8: Summary of evaluation of carcinogenic potential of RC to humans based on transcriptomics responses induced in HepG2 cells

After 24h of exposure					
RC	Approach 1	Approach 2	Approach 3	Approach 4 (NR)	Approach 4 (TF)
BDCM	HC	NC	NC	-	NC
EtAc	NC	NC	inc.	-	NC
EthylB	HC	inc.	-	-	NC
NBZ	HC	inc.	HC	-	inc.
CCl4	HC	-	-	-	-
DEHP	HC	HC	inc.	-	inc.
AmpC	NC	-	-	-	-

After 48h of exposure					
RC	Approach 1	Approach 2	Approach 3	Approach 4 (NR)	Approach 4 (TF)
BDCM	HC	NC	HC	-	inc.
EtAc	NC	inc.	inc.	inc.	NC
EthylB	HC	NC	-	inc.	inc.
NBZ	HC	NC	-	NC	NC
CCl4	HC	NC	-	inc.	NC
DEHP	NC	inc.	HC	-	inc.
AmpC	HC	-	-	NC	-

HC: the results of the analysis for this compound indicate a similar mode of action as the HC.

NC: the results of the analysis for this compound indicate similarities with the transcriptomics responses of NC.

inc: for the indicated analysis the compound shows similarities with both HC and NC compounds.

-: for the indicated analysis no result was obtained for the corresponding compound.

In particular, EtAc appears to be non-carcinogenic based on the results of both 24h and 48h of exposure. AmpC appears to be non-carcinogenic based only on the results of PAM analysis for 24h of exposure. For EtAc, it has been proposed, based on histopathological data, that its carcinogenic mechanism is not relevant to humans (38); while AmpC is widely used as an antibiotic agent with no reported observations on cancer risk for humans (39), and our data are supported by these.

NBZ, CCl4 and DEHP appear to share a similar mode of action with HC based on the results of the different approaches at 24h of exposure, thus these compounds may pose a cancer risk for humans. The carcinogenic mode of action for NBZ has been proposed to be multifactorial, including, among others, the production of free radicals. CCl4-induced rodent tumorigenesis is mediated by its reactive metabolite, trichloromethyl radical, that can bind to DNA and react with oxygen leading also to the production of ROS and thus initiate cancer (40). The induction of oxidative stress is considered a common mechanism of action of non-genotoxic carcinogens. Six of the HC compounds used in this study (TCDD, DDT, HCH, E2, PCP and PhB) result in the production of Reactive Oxygen Species (ROS) and oxidative stress (41-45). That may explain why transcriptomics responses of CCl4 and NBZ indicate that it may be carcinogenic to humans. However, while TCDD, DDT, E2 and PhB exposures led to the activation of the transcription factor Nrf2, that is generally considered as the mediator of oxidative stress (46), neither CCl4 nor NBZ exposures induced its activation in HepG2 cells.

The rodent carcinogenic properties of DEHP were considered to be mediated by the activation of PPAR α thereby leading to peroxisome proliferation. However, recent studies show that DEHP can induce cancer in rodents also via a PPAR α independent mechanism

(47). While peroxisome proliferation has been considered as a key event for the development of liver cancer in rodents, the relevance of this mechanism for human carcinogenesis has been challenged especially due to lack of sufficient evidence that can link PPAR α activation to cancers of humans (48). Therefore, it has been proposed that other pathways that have not yet been elucidated may be responsible for the induced cancers by peroxisome proliferators in humans, and these pathways may be the underlying reasons for DEHP being predicted as HC by the statistical approaches in this study. The biological based analyses indicate that DEHP induces similar effects with the HC compound HCH, and therefore these compounds may share a similar mode of action.

For the remaining two non-genotoxic RC, however, we cannot provide a conclusion, as no clear similarities are to be seen with either HC or NC. The main reason for that may have been the low number of non-genotoxic chemicals with proven carcinogenic risks for humans (6), or the lack of non-genotoxic chemicals with proven absence of carcinogenic effects to humans. For this reason, non-carcinogenic compounds for rodents which are generally considered safe for humans were used for the purposes of this study. The lack of sufficient HC as a reference, therefore resulted in the identification of only limited non-genotoxic modes-of-action that may have not included the mode-of-actions of the RC.

Using transcriptomics responses, however, we were able to identify biological effects related to the mode-of-actions of five of the HC compounds (TCDD, HCH, DDT, CsA and PhB). TCDD induces carcinogenicity by activating the AhR receptor that deregulates the expression of several genes leading to the promotion of cell growth and inhibition of apoptosis (49, 50), as well as to increased expression of CYP450 enzymes (mainly CYP1A1) that may lead to oxidative stress (41). In agreement with this, TCDD-exposure of HepG2 cells resulted in the activation of AhR, as indicated by transcriptional regulation analysis, in a >300-fold increase of CYP1A1 mRNA levels after 48h of exposure, and in the up-regulation of genes involved in the positive regulation of anti-apoptosis for both 24h and 48h of exposure. TCDD further acts as an endocrine disruptor, meaning that it may deregulate hormone synthesis, secretion, transport, binding or elimination, as well as hormone-mediated processes such as maintenance of homeostasis, reproduction, development, and/or behavior (51, 52). DDT and HCH are also known endocrine disruptors (52, 53). In concordance to this, exposures of HepG2 cells to these compounds resulted in the deregulation of genes involved in hormone metabolic process and in cellular homeostasis.

Another potential mechanism for DDT induced carcinogenicity is the inhibition of GJIC, but without altering the expression of connexins (54). This may explain why the functional annotation analysis of DEGs after DDT exposures did not identify GJIC as an affected biological process. However, up-regulated DEGs after DDT treatment were involved in the regulation of cellular communication for both exposure durations.

CsA induced carcinogenicity is considered to be mediated through its immunosuppressive properties. Furthermore, CsA increases the risk for skin cancer and this is believed to result from DNA-damage induced after exposure to ultraviolet radiation. However, the cells of the skin contain mechanisms to prevent this DNA damage from leading to skin carcino-

genesis; therefore the increased risk of skin cancer after CsA therapy may be due to the suppression of DNA repair mechanisms (55). CsA-exposure of HepG2 cells did not alter the expression of genes involved in immune related processes; however, several DNA-damage repair processes and cell cycle processes appeared to be down-regulated, supporting the hypothesis explaining the increased risk of skin cancer for immunosuppressed patients.

Lastly, PhB activates the Constitutive Androstane Receptor (CAR) that leads to the transcription of CYP2B1 and CYP2B2 (56). Our analysis showed that also in HepG2 cells PhB activates CAR only after 48h of exposure; however neither CYP2B1 nor CYP2B2 are up-regulated, but increased expression of CYP2B5 is observed. In agreement with studies showing that PhB enhances cell proliferation, PhB-exposure results in down-regulation of genes involved in the negative regulation of cell proliferation after 48h of exposure. Furthermore, the effects of PhB on cell proliferation are considered to be mediated by CAR and Pregnane X Receptor (PXR) (57) and our analysis shows both nuclear receptors to be activated at 48h of exposure.

While for the above mentioned HC compounds, we were able to identify responses related to their mode-of-action, for E2, TCE and PCP that was not feasible. Whereas E2 is the natural ligand of Estrogen Receptor (ESR), this ligand-dependent transcription factor was not activated in HepG2 after exposure to E2. E2 is further known to mediate the following physiological events: development of reproductive tissues, bone growth, lipid homeostasis and cell proliferation (58). Functional annotation of DEGs revealed that E2 altered the expression of genes involved in development and proliferation, mainly after 48h. Thus, we were only partly able to detect effects related to its mechanism-of-action. The mechanism of TCE induced carcinogenesis has not yet been fully established, but *in vitro* data indicate that it may be the inhibition of GJIC (59). In our study, TCE exposures resulted in limited transcriptomics responses, 22 and 1 DEGs for 24h and 48h of exposures respectively, and therefore no information was obtained for its putative mode of action. PCP is considered to act as a promoter of carcinogenesis enhancing cell proliferation (60). In our study, the use of statistical approaches revealed connections of these compounds with other HC. All three chemicals were classified as HC by PAM. Further, E2 appeared significantly connected with CsA after 24h and with TCDD after both 24h and 48h of exposure. This may be due to the fact that exposures to E2 and TCDD can lead to production of ROS (41, 44). PCP exposures also result in oxidative stress (45, 60), and that may be the reason that PCP appears also significantly connected with TCDD after 24h of exposure. Lastly, connectivity analysis revealed a strong connection between TCE and PCP at 24h of exposure.

The reason that we were unable to identify transcriptomics responses related to the mode-of-action for the above mentioned HC may lie in the transcriptomics analysis itself or even in the selected cell system. Transcriptomics analysis may not have been an appropriate endpoint for the evaluation of carcinogenic properties of non-genotoxic chemicals, as it can be derived by the lack of transcriptomics responses after PCP and TCE exposures, as well as for some of the RC. While, the selection of HepG2 cells for this study was based on the fact that the liver is the major target organ of chemical carcinogens in rodent bioassays (29), only TCDD is known to cause liver cancer in humans (61), and only one epidemiological study supports

an association for liver cancer and CsA exposure (62). However, epidemiological studies show no association with liver cancer for E2, PhB and DDT exposures (56, 63, 64).

By combining the above, we can conclude that the use of transcriptomics responses after non-genotoxic treatments of HepG2 was successful in correctly identifying biological effects related to the mode of actions of five of the HC compounds whereas statistical approaches failed to identify only one HC compound. Furthermore, it appears that different exposure times are required for the biological effects of different compounds to take place. Since several of the HC used in this study did not result in the identification of specific transcriptomics responses, the lack of similar transcriptomics responses of some of the RC compounds with the HC cannot exclude the possibility that these RC may be carcinogenic to humans. Since the selected transcriptomics-approaches used in this study were unsuccessful in evaluating the risks for humans of known non-genotoxic rodent carcinogens in a consistent manner, other endpoints or combinations thereof such as metabolomics, proteomics or epigenomics analyses, may further facilitate the identification and characterization of non-genotoxic carcinogens. These in combination with a more relevant cell system for each of the different cancer types, e.g. lymphomas, breast cancer, liver cancer etc., may be needed in further studies for the purpose of the development of an *in vitro* carcinogenicity prediction method.

References

1. Jemal A, Siegel R, Xu J, & Ward E (2010) Cancer statistics, 2010. *CA Cancer J Clin* 60(5):277-300.
2. OECD (2010) Mortality from Cancer, in OECD, Health at a Glance: Europe 2010 *OECD Publishing*:34-35.
3. Petrucelli N, Daly MB, & Feldman GL (2010) Hereditary breast and ovarian cancer due to mutations in BRCA1 and BRCA2. *Genet Med* 12(5):245-259.
4. Clapp RW, Jacobs MM, & Loechler EL (2008) Environmental and occupational causes of cancer: new evidence 2005-2007. *Rev Environ Health* 23(1):1-37.
5. Oliveira PA, *et al.* (2007) Chemical carcinogenesis. *An Acad Bras Cienc* 79(4):593-616.
6. Hernandez LG, van Steeg H, Luijten M, & van Benthem J (2009) Mechanisms of non-genotoxic carcinogens and importance of a weight of evidence approach. *Mutat Res* 682(2-3):94-109.
7. Ames BN, Lee FD, & Durston WE (1973) An improved bacterial test system for the detection and classification of mutagens and carcinogens. *Proc Natl Acad Sci U S A* 70(3):782-786.
8. OECD (1997) OECD GUIDELINE FOR THE TESTING OF CHEMICALS: Test No. 476: In vitro Mammalian Cell Gene Mutation Test.
9. OECD (1997) OECD GUIDELINE FOR THE TESTING OF CHEMICALS: Test No. 473: In Vitro Mammalian Chromosome Aberration Test.
10. OECD (2010) OECD GUIDELINE FOR THE TESTING OF CHEMICALS: Test No. 487: In vitro mammalian cell micronucleus test.
11. Kirkland D, Aardema M, Henderson L, & Möller L (2005) Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens I. Sensitivity, specificity and relative predictivity. *Mutat Res*. 584(1-2):1-256.
12. ICH (1997) S1B TESTING FOR CARCINOGENICITY OF PHARMACEUTICALS.
13. Nuwaysir EF, Bittner M, Trent J, Barrett JC, & Afshari CA (1999) Microarrays and toxicology: the advent of toxicogenomics. *Mol Carcinog* 24(3):153-159.
14. Ellinger-Ziegelbauer H, Stuart B, Wahle B, Bomann W, & Ahr HJ (2005) Comparison of the expression profiles induced by genotoxic and nongenotoxic carcinogens in rat liver. *Mutat Res* 575(1-2):61-84.
15. Waring JF, Ciurlionis R, Jolly RA, Heindel M, & Ulrich RG (2001) Microarray analysis of hepatotoxins in vitro reveals a correlation between gene expression profiles and mechanisms of toxicity. *Toxicol Lett*. 120(1-3):359-368.
16. Uehara T, *et al.* (2008) A toxicogenomics approach for early assessment of potential non-genotoxic hepatocarcinogenicity of chemicals in rats. *Toxicology* 250(1):15-26.
17. Hamadeh HK, *et al.* (2002) Gene expression analysis reveals chemical-specific profiles. *Toxicol Sci* 67(2):219-231.
18. Aniagu SO, Williams TD, & Chipman JK (2009) Changes in gene expression and assessment of DNA methylation in primary human hepatocytes and HepG2 cells exposed to the environmental contaminants-Hexabromocyclododecane and 17-beta oestradiol. *Toxicology* 256(3):143-151.
19. Maeshima H, Ohno K, Tanaka-Azuma Y, Nakano S, & Yamada T (2009) Identification of tumor promotion marker genes for predicting tumor promoting potential of chemicals in BALB/c 3T3 cells. *Toxicol In Vitro* 23(1):148-157.
20. Jonker MJ, *et al.* (2009) Finding transcriptomics biomarkers for in vivo identification of (non-)genotoxic carcinogens using wild-type and Xpa/p53 mutant mouse models. *Carcinogenesis* 30(10):1805-1812.
21. Nioi P, Pardo ID, Sherratt PJ, & Snyder RD (2008) Prediction of non-genotoxic carcinogenesis in rats using changes in gene expression following acute dosing. *Chem Biol Interact* 172(3):206-215.
22. Ennever FK & Lave LB (2003) Implications of the lack of accuracy of the lifetime rodent bioassay for predicting human carcinogenicity. *Regul Toxicol Pharmacol* 38(1):52-57.
23. Mascolo MG, *et al.* (2010) BALB/c 3T3 cell transformation assay for the prediction of carcinogenic potential of chemicals and environmental mixtures. *Toxicol In Vitro* 24(4):1292-1300.
24. Vanparys P. CR, Aardema M., Gribaldo L., Hayashi M., Hoffmann S. and Schechtman L. (2011) ECVAM Prevalidation of Three Cell Transformation Assays. *ALTEX* 28(1):56-59.
25. Hoenerhoff MJ, Hong HH, Ton TV, Lahousse SA, & Sills RC (2009) A review of the molecular mechanisms of chemically induced neoplasia in rat and mouse models in National Toxicology Program bioassays and their relevance to human cancer. *Toxicol Pathol* 37(7):835-848.
26. Hard GC & Khan KN (2004) A contemporary overview of chronic progressive nephropathy in the laboratory rat, and its significance for human risk assessment. *Toxicol Pathol* 32(2):171-180.

27. Thayer KA & Foster PM (2007) Workgroup report: National Toxicology Program workshop on Hormonally Induced Reproductive Tumors - Relevance of Rodent Bioassays. *Environ Health Perspect* 115(9):1351-1356.
28. Caldwell DJ (1999) Review of mononuclear cell leukemia in F-344 rat bioassays and its significance to human cancer risk: A case study using alkyl phthalates. *Regul Toxicol Pharmacol* 30(1):45-53.
29. Gold LS, Slone TH, Stern BR, & Bernstein L (1993) Comparison of target organs of carcinogenicity for mutagenic and non-mutagenic chemicals. *Mutat Res* 286(1):75-100.
30. Edgar R, Domrachev M, & Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30(1):207-210.
31. Jennen DG, et al. (2010) Comparison of HepG2 and HepaRG by whole genome gene expression analysis for the purpose of chemical hazard identification. *Toxicol Sci*.
32. Tibshirani R, Hastie T, Narasimhan B, & Chu G (2002) Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci U S A*. 99(10):6567-6572.
33. Lamb J (2007) The Connectivity Map: a new tool for biomedical research. (Translated from eng) *Nat Rev Cancer* 7(1):54-60 (in eng).
34. Zhang SD & Gant TW (2009) sscMap: an extensible Java application for connecting small-molecule drugs using gene-expression signatures. *BMC Bioinformatics* 10:236.
35. Dennis G, Jr., et al. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 4(5):P3.
36. Tibshirani R, Hastie T, Narasimhan B, & Chu G (2002) Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci U S A*. 99(10):6567-6572.
37. Lamb J, et al. (2006) The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313(5795):1929-1935.
38. Williams GM & Iatropoulos MJ (2009) Evaluation of potential human carcinogenicity of the synthetic monomer ethyl acrylate. *Regul Toxicol Pharmacol* 53(1):6-15.
39. Drawz SM & Bonomo RA (2010) Three decades of beta-lactamase inhibitors. (Translated from eng) *Clin Microbiol Rev* 23(1):160-201 (in eng).
40. Weber LW, Boll M, & Stampfl A (2003) Hepatotoxicity and mechanism of action of haloalkanes: carbon tetrachloride as a toxicological model. *Crit Rev Toxicol* 33(2):105-136.
41. Yoshida R & Ogawa Y (2000) Oxidative stress induced by 2,3,7,8-tetrachlorodibenzo-p-dioxin: an application of oxidative stress markers to cancer risk assessment of dioxins. *Ind Health* 38(1):5-14.
42. Perez-Maldonado IN, et al. (2005) DDT-induced oxidative damage in human blood mononuclear cells. *Environ Res* 98(2):177-184.
43. Piskac-Collier AL & Smith MA (2009) Lindane-induced generation of reactive oxygen species and depletion of glutathione do not result in necrosis in renal distal tubule cells. *J Toxicol Environ Health A* 72(19):1160-1167.
44. Yu FL (2002) 17-beta Estradiol epoxidation as the molecular basis for breast cancer initiation and prevention. *Asia Pac J Clin Nutr* 11 Suppl 7:S460-466.
45. Zhu BZ & Shan GQ (2009) Potential mechanism for pentachlorophenol-induced carcinogenicity: a novel mechanism for metal-independent production of hydroxyl radicals. *Chem Res Toxicol* 22(6):969-977.
46. Singh S, Vrishni S, Singh BK, Rahman I, & Kakkar P (2010) Nrf2-ARE stress response mechanism: a control point in oxidative stress-mediated dysfunctions and chronic inflammatory diseases. (Translated from eng) *Free Radic Res* 44(11):1267-1288 (in eng).
47. Ito Y & Nakajima T (2008) PPARalpha- and DEHP-Induced Cancers. *PPAR Res* 2008:759716.
48. Guyton KZ, et al. (2009) A reexamination of the PPAR-alpha activation mode of action as a basis for assessing human cancer risks of environmental contaminants. (Translated from eng) *Environ Health Perspect* 117(11):1664-1672 (in eng).
49. Yang JH & Rhim JS (1995) 2,3,7,8-Tetrachlorodibenzo-p-dioxin: molecular mechanism of carcinogenesis and its implication in human in vitro model. *Crit Rev Oncol Hematol* 18(2):111-127.
50. Chopra M, Gahrs M, Haben M, Michels C, & Schrenk D (2010) Inhibition of apoptosis by 2,3,7,8-tetrachlorodibenzo-p-dioxin depends on protein biosynthesis. *Cell Biol Toxicol* 26(4):391-401.
51. Mocarelli P, et al. (2008) Dioxin exposure, from infancy through puberty, produces endocrine disruption and affects human semen quality. *Environ Health Perspect* 116(1):70-77.
52. Crisp TM, et al. (1998) Environmental endocrine disruption: an effects assessment and analysis. *Environ Health Perspect* 106 Suppl 1:11-56.

53. Di Consiglio E, De Angelis G, Traina ME, Urbani E, & Testai E (2009) Effect of lindane on CYP-mediated steroid hormone metabolism in male mice following in utero exposure. *J Appl Toxicol* 29(8):648-655.
54. Cowles C, Mally A, & Chipman JK (2007) Different mechanisms of modulation of gap junction communication by non-genotoxic carcinogens in rat liver in vivo. *Toxicology* 238(1):49-59.
55. Bugelski PJ, *et al.* (2010) Critical review of preclinical approaches to evaluate the potential of immunosuppressive drugs to influence human neoplasia. *Int J Toxicol* 29(5):435-466.
56. IARC (2001) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 79: Some Thyrotropic Agents.
57. Ross J, *et al.* (2010) Human constitutive androstane receptor (CAR) and pregnane X receptor (PXR) support the hypertrophic but not the hyperplastic response to the murine nongenotoxic hepatocarcinogens phenobarbital and chlordane in vivo. *Toxicol Sci* 116(2):452-466.
58. JECFA (2001) Toxicological evaluation of certain veterinary drug residues in food. *WHO Food Additives Series, No. 49* nos 1041-1052 on INCHEM. .
59. Benane SG, Blackman CF, & House DE (1996) Effect of perchloroethylene and its metabolites on intercellular communication in clone 9 rat liver cells. *J Toxicol Environ Health* 48(5):427-437.
60. Umemura T, *et al.* (1999) Pentachlorophenol (PCP) produces liver oxidative stress and promotes but does not initiate hepatocarcinogenesis in B6C3F1 mice. *Carcinogenesis* 20(6):1115-1120.
61. IARC (1997) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 69: Polychlorinated Dibenzo-para-Dioxins and Polychlorinated Dibenzofurans.
62. IARC (2011) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 100A: A Review of Human Carcinogens: Pharmaceuticals.
63. IARC (1991) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 53: Occupational Exposures in Insecticide Application, and Some Pesticides
64. IARC (1999) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 72: Hormonal Contraception and Post-menopausal Hormonal Therapy.
65. Huff JE, Salmon AG, Hooper NK, & Zeise L (1991) Long-term carcinogenesis studies on 2,3,7,8-tetrachlorodibenzo-p-dioxin and hexachlorodibenzo-p-dioxins. *Cell Biol Toxicol* 7(1):67-94.
66. IARC (1990) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 50: Pharmaceutical Drugs.
67. IARC (1987) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Supplement 7: Overall Evaluations of Carcinogenicity: An Updating of IARC Monographs Volumes 1 to 42.
68. IARC (1995) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 63: Dry Cleaning, Some Chlorinated Solvents and Other Industrial Chemicals
69. IARC (1999) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 71: Re-evaluation of Some Organic Chemicals, Hydrazine and Hydrogen Peroxide.
70. IARC (2000) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 77: Some Industrial Chemicals.
71. IARC (1996) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 66: Printing Processes and Printing Inks, Carbon Black and Some Nitro Compounds.
72. IARC (1978) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 16: Some Aromatic Amines and Related Nitro Compounds.

Chapter 7

Summary and general discussion

This thesis sought to investigate the application of toxicogenomics approaches *in vitro* for the purpose of identification of genotoxic and carcinogenic properties of chemicals. The fundamental triggers of this research were the major limitations of the currently used *in vitro* and *in vivo* systems for genotoxicity and carcinogenicity prediction. Specifically, these limitations refer to the high false positive results of the *in vitro* assays (1). Consequently, where positive results from *in vitro* tests lead to animal experimentation there are concerns with respect to both ethical issues and the relevance of animal-based models for human health (2).

This research, therefore, aimed at the development of *in vitro* assays, by means of toxicogenomics analyses, for the prediction of *in vivo* genotoxicity and carcinogenicity of chemicals that will overcome the limitations of the currently available test systems. Alongside, this research aimed at further elucidating modes-of-actions of chemical carcinogens based on the toxicogenomics results.

Selection of the most appropriate *in vitro* cellular system

The first challenge was to identify the most appropriate *in vitro* cell system that can be used for realizing these research aims. Taking into consideration that the liver is the main organ where the metabolism of chemical carcinogens into reactive intermediates takes place (3), the use of an *in vitro* hepatic system appeared to be the most appropriate. Therefore, as described in **Chapter 2**, two human hepatic cell lines were compared by means of gene expression analyses, for selecting the most promising *in vitro* cellular system that can be used for the purpose of transcriptomics-based class prediction.

The first cell line, HepaRG cells, originates from a liver tumor of a female patient suffering from hepatocarcinoma (4). Seeded at low density, HepaRG cells emerge as two cell types when they reach confluence. The cells of the first type resemble hepatocytes while the cells of the second type surround the hepatocyte-like cells and are flattened with a clear cytoplasm. Under specific differentiation conditions, the hepatocyte-like cells further differentiate into more granular cells, closely resembling typical adult primary hepatocytes. The hepatocyte-like cells represent around 50 to 55% of the total cell population (5). In addition to these morphological similarities to adult primary hepatocytes, differentiated HepaRG cells have also been proven to express various cytochrome P450 enzymes and to exhibit other metabolic functions normally found in liver (6, 7). Whole genome gene expression analysis as conducted for the purposes of this thesis further revealed under normal culture conditions a strong resemblance of differentiated HepaRG cells to primary human hepatocytes and postmortem collected human liver.

The second cell line, HepG2 cells, was originally derived from a 15-year old male patient suffering from a well-differentiated hepatocellular carcinoma. HepG2 cells have been proven to possess cancer stem cells properties (8, 9). Cancer stem cells represent a small fraction of cancer cells with properties of embryonic stem cells and are considered responsible for the initiation, progression, local and distant recurrence / metastasis of tumors, as well as for the failure of chemo- and radiotherapy (10, 11). In contrast to HepaRG

cells, HepG2 cells show limited expression of metabolizing enzymes under normal conditions. Whole genome gene expression analysis further revealed limited similarities of HepG2 cells to primary hepatocytes, as described in this thesis. Nevertheless, activated biotransformation pathways and high expression levels of CYP450 enzymes are detected after chemical exposures (12-14), reaching similar levels to that observed in HepaRG cells under similar conditions.

The comparison between HepG2 and HepaRG cell lines by whole genome gene expression analysis revealed several differences, at a gene- as well as at a pathway-level, under normal cultural conditions; furthermore, the two cell lines showed differences but also similarities after chemical exposures. The similar responses after chemical exposure comprise pathways that are important in chemical carcinogenesis, including cell cycle and apoptotic pathways.

In the initial attempt to perform a class discrimination for the purpose of identifying the most appropriate cellular system for this application, two genotoxic and two non-genotoxic carcinogens were selected for the eventual development of a prediction model, that was further validated by a third non-genotoxic carcinogen with some evidence for genotoxic *in vitro* properties. This analysis, described in **Chapter 2**, clearly shows that HepG2 cells outperform HepaRG cells in terms of genotoxicity evaluation.

Thus, HepG2 cells were selected as the most appropriate system for further studies: firstly, because of their ability to discriminate genotoxic from non-genotoxic carcinogens more accurately than HepaRG cells; and secondly, in view of their capability to generate transcriptomics responses to carcinogens that included alterations in important pathways of carcinogenesis. However, extrapolation of the obtained data from HepG2 cells to the human *in vivo* situation should be done with precaution due to the limited resemblance that they show to primary human hepatocytes and human liver. Nevertheless, due to their similarities to cancer stem cells, HepG2 cells may be a more appropriate system for the study of carcinogenesis.

Suitability and limitations of HepG2 for genotoxicity evaluation

Next efforts were focused on testing the capability of HepG2 cells to indeed discriminate between true *in vivo* genotoxic compounds and false positive genotoxic compounds (i.e. chemicals that appear genotoxic *in vitro*, but do not induce genotoxicity *in vivo*). To this end, in **Chapter 3**, the effects of three false positive genotoxic compounds at the phenotypic and the transcriptomic level were analyzed in HepG2 cells and compared with the effects of true genotoxins and non-genotoxins. From the three false positive genotoxic compounds under investigation, 8-hydroxyquinoline showed similarities to the analyzed true genotoxic compounds only at the phenotypic level, thus providing evidence that for evaluating the genotoxic properties of chemicals the analysis of transcriptomics responses is a more accurate method than the analysis of phenotypic endpoints. This may relate to the high false positive rate of the currently available *in vitro* genotoxicity assays as these only measure phenotypic

effects. Only one false positive genotoxic compound, Quercetin, was shown to induce similar transcriptomics responses as the used genotoxic compounds. This study therefore provides first proof of the principle that analysis of transcriptomics responses in HepG2 cells is suitable for genotoxicity evaluation generating a reduced false positive rate in comparison to the currently used *in vitro* test systems for evaluating genotoxicity *in vivo*. Furthermore, the observed similarities between Quercetin and true genotoxins were more evident in the early phase of exposure, which thus indicates that the duration of exposure is an important factor for the accurate classification of genotoxic properties of chemicals.

The next aim was to develop an accurate prediction model for *in vivo* genotoxicity. Since time-dependent gene expression changes have been extensively demonstrated for a variety of chemicals the role of exposure durations on chemical classification was simultaneously investigated. To achieve this in a statistically reliable manner, a high number of chemicals is required. Therefore, in **Chapter 4**, transcriptomics changes induced in HepG2 cells by 34 compounds after treatment for 12h, 24h and 48h were used for the selection of gene-sets that are capable of discriminating between *in vivo* genotoxins and *in vivo* non-genotoxins. For the selection of such gene-sets different approaches were applied that led to the development of different prediction models. The first approach was based only on transcriptomics data generated by this study, the second further incorporated publicly available results on these test compounds from the Ames bacterial mutagenicity assay, while the last incorporated the integrated outcome of standard *in vitro* genotoxicity studies for the different chemicals. The application of these approaches led to the identification of different gene-sets for each of the exposure durations. This analysis further showed that gene-sets generated after 12h of exposure led, in general, to the poorest prediction, and thus 12h was excluded as exposure duration for genotoxicity evaluation. Gene-sets generated after 48h of exposure performed well for the different genotoxicity prediction models. Clearly, the most promising gene-set was generated after stratification of chemicals based on the publicly available results of the Ames mutagenicity assay and particularly after 24h of exposure. In order to confirm this observation, validation of these results was conducted for 24h and 48h of exposure using an additional set of 28 chemicals. This confirmed the initial observations that the best prediction is achieved after 24h of exposure using the gene-sets that were selected after stratification of chemicals according to results from the Ames bacterial gene mutation assay. In particular, this method resulted in 89% accuracy and 91% specificity, thereby clearly outperforming the standard *in vitro* test battery. Thus, the results of **Chapter 4** confirm that the analysis of transcriptomics responses in HepG2 cells after chemical exposures can be successfully used as an upgraded *in vitro* approach for predicting *in vivo* genotoxicity when combined with the results of the Ames bacterial gene mutation assay. Therefore, the use of this method does not only overcome the limitations of the currently used *in vitro* genotoxicity assays but also holds great promise for reducing animal experimentations on genotoxicity testing. Some possible applications of this method for the purpose of reducing animal experimentations will be shortly discussed within the section *Future Perspectives and Recommendations*.

Yet, results obtained by this method still have to be interpreted in a careful and cautious manner, especially with respect to the extrapolation to human genotoxic hazards, mostly due to the fact that rodent genotoxic chemicals were primarily used for the development of this assay. This was mainly due to lack of sufficient information on genotoxic effects of the selected chemicals to humans. Further, several considerations arise regarding the classification of the chemicals to the genotoxic and non-genotoxic classes. In exploiting available information from rodent *in vivo* genotoxicity studies, we often encountered difficulties in classifying chemicals as genotoxic or non-genotoxic, mostly due to inconclusive results from the *in vivo* studies as reported in the public literature. Therefore only clearly positive results were taken into consideration for the assignment of the chemicals into the genotoxic class.

The second uncertainty related to this method, refers to the limited metabolic capacity of HepG2 cells as mentioned earlier. Although our studies showed that HepG2 cells express the required enzymes for the metabolism of several of the examined compounds (**Chapter 2** and **Chapter 3**), expression of some CYP450 enzymes is not detected or very low in HepG2 cells (12). However, almost all genotoxic chemicals used in this thesis, even those that need metabolic activation were correctly classified. Nevertheless, the results obtained from this assay for chemicals of which the metabolic fate is unknown may need further investigation in order to prove that these chemicals are indeed metabolized and to additionally support the obtained results from the genotoxicity prediction method.

Suitability and limitations of HepG2 for carcinogenicity evaluation

While more than 90% of the chemicals classified by the International Agency for Research on Cancer (IARC) as human carcinogens, act through genotoxic mechanisms, the remaining 10% of these chemicals act through a variety of non-genotoxic mechanisms, thus without inducing DNA damage (16). As almost all established genotoxic compounds are also carcinogenic, the application of our upgraded method for predicting genotoxicity *in vivo* can, thus, successfully be used for the identification of the majority of carcinogenic compounds. However, the carcinogenic potential of non-genotoxic chemicals cannot be evaluated by this method. Therefore, the research described in **Chapter 6** sought to analyze transcriptomics responses in HepG2 cells for the purpose of identification of non-genotoxic carcinogens. Taking into consideration the major limitation of rodent bioassays for carcinogenicity evaluation, namely their limited relevance to human health risks (2, 17, 18), this research specifically aimed at assessing the cancer risk of rodent non-genotoxic carcinogens for humans. This study was unavoidably limited by the low number of non-genotoxic chemicals with proven carcinogenic risks for humans (16) as well as by the lack of non-genotoxic chemicals with proven absence of carcinogenic effects to humans (19). For this reason, non-carcinogenic compounds for rodents which are generally considered safe for humans were used for the purposes of this study. Overall, this lack of sufficient test compounds probably provides the main reasons that efforts to classify rodent non-genotoxic carcinogens as human carcinogens or human non-carcinogens have failed, as mostly inconclusive results

were obtained for the rodent carcinogens that were investigated. This may be firstly due to the fact that not all possible non-genotoxic carcinogenic mode-of-actions were covered by the selected compounds and secondly, due to the fact that transcriptomics responses are not sufficient discriminative for non-genotoxic carcinogens.

Nevertheless, the analysis of transcriptomics responses of HepG2 cells after exposure to some of the human non-genotoxic carcinogens confirmed their modes of action, as it will be discussed in the next section, suggesting that HepG2 cells may be capable of identifying some of the human non-genotoxic carcinogens. Therefore, to exploit this potential of the HepG2 model, alternative approaches to those used in **Chapter 6** need to be considered for increasing the accuracy of predicting human carcinogenicity. Some of these options will be further discussed under the section *Future Perspectives and Recommendations*.

Suitability and limitations of HepG2 for mechanistic studies

The selection of toxicogenomics as a tool for the purposes of this thesis was mainly based on its potential for class discrimination (20). However, toxicogenomics studies also hold great promise for identifying mechanisms-of-action of toxicants (21). Therefore, the secondary aim of this thesis was to apply toxicogenomics tools, in combination with phenotypic analyses, in HepG2 cells for the collection of information related to the mode-of-actions of chemicals.

Most in depth studied chemicals comprise the well-known compounds: Benzo[a]pyrene, a genotoxic carcinogen, 17-beta Estradiol, a non-genotoxic carcinogen with evidence of *in vitro* genotoxic effects (false positive genotoxic) and 2,3,7,8-tetrachlorodibenzo-para-dioxin (TCDD), a non-genotoxic carcinogen.

Benzo[a]pyrene is known to form DNA adducts, after being metabolized, that may lead either to cell cycle arrest or apoptosis (22, 23). In HepG2 cells, Benzo[a]pyrene exposure led to the alteration of the expression levels of genes involved in apoptosis, DNA damage and cell cycle progress regulation as described in **Chapter 2** and **Chapter 3**, as early as 12h after exposure. The studies conducted in **Chapter 3** and **Chapter 5** confirmed these effects also at the phenotypic level, showing however a reduction of the induced DNA damage at later time points. This demonstrates that HepG2 cells are capable of metabolizing Benzo[a]pyrene to its active metabolites thus enabling the detection of Benzo[a]pyrene-induced effects by both transcriptomics and phenotypic analysis. This further suggests that HepG2 is a suitable system for studying genotoxic mechanisms.

However, when the effects of false-positive genotoxic chemicals were analyzed, it appeared that they also induce genotoxic effects in HepG2 cells. For instance, as described in **Chapter 3** both Quercetin and 8-Hydroxyquinoline induced DNA damage. DNA damage in HepG2 was also detected by phenotypic analysis after exposure to the false-positive genotoxic compound 2-Chloromethylpyridine, as described in **Chapter 5**. Nonetheless, only Quercetin induced the expression of genes involved in DNA-damage repair processes, suggesting that this actually represents an *in vitro* genotoxic compound (**Chapter 3**). Nevertheless, the developed assay for predicting genotoxicity *in vivo* in **Chapter 4**, classified

Quercetin correctly as an *in vivo* non-genotoxic compound. The other two false positive compounds were also correctly identified as *in vivo* non-genotoxins by the prediction model developed in **Chapter 4**.

Correct prediction, as non-genotoxic, was also achieved for the well known false positive genotoxic carcinogen, 17-beta Estradiol, in all classification approaches described in **Chapter 2**, **Chapter 3** and **Chapter 4**. However, the absence of evidence of genotoxicity is not sufficient for the elucidation of its mode of action. While 17-beta Estradiol is the natural ligand of the Estrogen Receptor – a ligand-activated transcription factor –, the transcriptional regulation analysis conducted in **Chapter 6** failed to identify its activation. 17-beta Estradiol is further known to mediate the following physiological events: development of reproductive tissue, bone growth, lipid homeostasis and cell proliferation (24). Analysis of transcriptomics effects at the pathway level revealed that 17-beta Estradiol altered the expression of genes involved in development and proliferation. Thus, transcriptomics analysis of 17-beta Estradiol induced effects in HepG2 cells was partly able to detect effects related to the mechanism-of-action of this non-genotoxic carcinogen. These effects were more prominent at later time points as described in **Chapter 3**.

In contrast to the partial identification of 17-beta Estradiols mode-of-action, clear results were obtained for the mode-of-action of the well-known non-genotoxic carcinogen, TCDD. TCDD is known to activate the Aryl hydrocarbon receptor that deregulates the expression of several genes, including the CYP450 enzyme CYP1A1. The altered expression of several genes further leads to the promotion of cell growth and inhibition of apoptosis as well as to endocrine disruption (25-27). In HepG2 cells, TCDD exposure indeed activated the Aryl hydrocarbon receptor resulting in the increased expression of CYP1A1, as described in **Chapter 2** and **Chapter 6**, and affected the expression of several genes involved in cell growth and homeostasis and hormone-related processes, as described in the Supplementary Data of **Chapter 3** and in **Chapter 6**. These effects, in contrast to the effects of 17-beta Estradiol, were identified at all exposure durations, suggesting that for non-genotoxic chemicals different exposure durations are necessary for the identification of their modes-of-action by toxicogenomics analyses. This is further supported by the observations related to the activation of the Constitutive Androstane Receptor by Phenobarbital. Phenobarbital, which is a non-genotoxic carcinogen, acts via the activation of this ligand-dependent transcription factor (28). The transcriptomics effects of Phenobarbital in HepG2 cells after 24h of exposure did not support the activation of this receptor, as described in **Chapter 4**; however, analysis of the transcriptomics responses of HepG2 cells after exposure to Phenobarbital for 48h, as described in **Chapter 6**, identified the activation of Constitutive Androstane Receptor, suggesting that longer exposures durations are necessary for the identification of the effects of Phenobarbital in HepG2 cells.

Thus, where this thesis proves that HepG2 cells is a suitable model for studies of genotoxic mechanisms, it is however also apparent that HepG2 cells are not able to detect the mode-of-action of all non-genotoxic carcinogens. Further, the absence of transcriptomics responses in HepG2 cells after exposures to novel chemicals should not be misinterpreted

as absence of carcinogenicity, as this is clearly not applicable for Tetrachloroethylene and Pentachlorophenol that did not induce any transcriptomics responses in HepG2 although they are known human carcinogens (**Chapter 6**).

Lastly, this thesis investigated the suitability of HepG2 cells for mechanistic studies related to the function of genes. Especially the function of NR0B2 was investigated, as described in **Chapter 5**, in order to investigate the biological relevance of this gene to DNA damage responses, as it was one of the classifying genes identified in **Chapter 4**. An RNA silencing approach was applied in HepG2 to reveal the role of NR0B2 prior and after genotoxic treatment. NR0B2 is a nuclear receptor that acts as a repressor of transcription and was identified as one of the genes with the highest contribution in the prediction model. The inhibition of its expression in combination with chemical exposure, however, did not affect the classification of model compounds, indicating that NR0B2 is not the determining factor for the classification of these compounds. That further suggests that other genes from the classifying gene-set may have a higher contribution to the prediction model and / or that inhibiting the expression of only one gene may not be sufficient for drastically altering the classification model. For this to be proven however, additional experiments using higher numbers of test compounds, in order to increase statistical power, would be necessary.

NR0B2 is further known to affect a variety of physiological processes such as lipid, cholesterol and glucose metabolism, homeostatic regulation, detoxification and drug metabolism in liver (29). These processes were identified to be altered in HepG2 cells growing under normal conditions after silencing of NR0B2, as described in **Chapter 5**. Therefore, HepG2 cells, although they have a hepatocellular carcinoma origin, appear to sustain functions of non-cancerous cells, since the NR0B2 network appears to remain unaffected in cells expressing NR0B2. The silencing approach described in **Chapter 5**, further revealed a significant function of NR0B2 in chemical responses, and especially in genotoxic stress responses as derived from phenotypic analysis. Thus, these observations support that mechanistic information regarding the functionality of genes can be obtained by the use of HepG2 cells.

In summary, the results described in **Chapter 2**, **Chapter 3** and **Chapter 5** prove the suitability of HepG2 cells for the investigation of genotoxic mechanisms of chemical carcinogens. Further, the results discussed in **Chapter 2**, **Chapter 3**, **Chapter 4** and **Chapter 6** indicate that HepG2 cells may be a useful system for studying some non-genotoxic modes-of-actions of chemical carcinogens. Lastly, the results described in **Chapter 5** prove that HepG2 cells can be used for the evaluation of functionalities of genes.

Conclusion

In conclusion, this thesis identified the HepG2 cell line as a suitable *in vitro* liver model for the purpose of genotoxicity evaluation by means of toxicogenomics analysis. Further, for the discrimination of true from false positive genotoxic compounds, the application of toxicogenomics analysis was more successful than the evaluation of phenotypic endpoints in HepG2 cells. Additionally, this thesis presents the development of a genotoxicity predic-

tion method by evaluating a limited number of classifying genes that outperforms the conventional *in vitro* test battery for assessing genotoxicity. For one of these genes, this thesis demonstrates its biological relevance to genotoxicity responses. Thus, it has been shown that HepG2 cells can be used for the prediction of the genotoxic properties of chemicals as well as for the study of genotoxic mechanisms of chemical carcinogens. Lastly, this thesis provides some evidence that the application of toxicogenomics approaches using HepG2 cells may provide mechanistic information about the mode-of-action of non-genotoxic chemicals, but also that this still is not sufficient for carcinogenicity prediction of non-genotoxic chemicals.

Future perspectives and recommendations

As mentioned above, this thesis deals with the application of toxicogenomics approaches in HepG2 cells for two major aims, the prediction of genotoxic properties and the evaluation of the carcinogenic potential of chemicals.

Due to the lower false-positive rate with equally high sensitivity of the developed genotoxicity prediction method in **Chapter 4** in comparison to the currently used *in vitro* test battery, this method is recommended for several applications that may serve the purpose of reducing the number of animal toxicity tests. These applications include the early phases of drug development and the purposes of the EU chemical policy program REACH. For evaluating intrinsic genotoxic properties of chemicals in drug development, a tiered approach for genotoxicity testing is currently used starting with *in vitro* testing followed by *in vivo* testing. Thus, this transcriptomics-based method may facilitate the discontinuation of true genotoxic compounds after *in vitro* genotoxicity testing and the prioritization of promising candidates which have been established as true non-genotoxic compounds, thereby avoiding, animal experiments on false positive genotoxic compounds. The EU chemical policy program REACH aims at collecting sufficient toxicological information of chemicals that are produced in or imported into the EU. It has been estimated that approximately 400000 rodents may be used for testing genotoxicity *in vivo* in order to fulfill the requirements of the REACH policy (30). Thus, the developed method described in **Chapter 4** may be used as an initial evaluation for this purpose and therefore avoid the animal experimentations that would be based on false premises from the conventional *in vitro* test battery.

Moreover, the developed method is recommended as a tool for genotoxicity assessment of novel cosmetics. In this case, the method developed in **Chapter 4** may replace the use of animals for genotoxicity evaluation of cosmetic ingredients, which is particularly relevant in view of the prohibition of animal testing for this purpose in EU (EC Regulation 1223/2009).

However, for this method to be generally accepted in chemical safety assessment, recognition by regulatory authorities is required. Thus, future efforts should aim at this recognition. The first steps into this direction have already been made by applying for the validation of this novel method by the European Centre for the Validation of Alternative Methods (ECVAM). This validation should aim at both intra- and inter-laboratory comparative studies,

in order to evaluate the robustness and the reproducibility of this assay. Additionally, the development of standardized procedures will be required in order to achieve recognition by regulatory authorities. To this end, the development of dedicated analytical platforms, based on RT-PCR arrays or customized microarrays, as well the development of software for the analysis of the data may be useful.

The reasons that toxicogenomics approaches could not provide a successful method for evaluating the carcinogenic potential of non-genotoxic chemicals may be found in the limited number of chemicals that were used for this purpose. Inclusion of more chemicals with known carcinogenic potential in humans may better serve the purpose of developing a carcinogenicity prediction model by covering more modes-of-action. Another reason could lie in the selected cell system or in the selected prediction approaches. For instance, while the selection of HepG2 cells for this study was based on the fact that the liver is the major target organ of chemical carcinogens in rodent bioassays (3), from the selected human carcinogens that were used in **Chapter 6**, only one is known to cause liver cancer in humans, while the majority of carcinogens are known to cause lymphomas. Thus, a more relevant cell system that will represent the target tissues of known human carcinogens may be needed in further studies for the identification of carcinogenic properties of chemicals. Finally, although for the majority of human non-genotoxic carcinogens, transcriptomics analysis was able to identify responses linked to their mode-of-action, for some chemicals this appeared not feasible. This may suggest that transcriptomics analysis is not always an appropriate endpoint for the evaluation of carcinogenic properties of non-genotoxic chemicals and thus other endpoints or combinations thereof need to be considered, such as metabolomics, proteomics or epigenomics analyses. Thus, within the domain of non-genotoxic carcinogenesis further research is recommended in order to identify a more appropriate *in vitro* cell system. A specific cell system for each of the different cancer types, e.g. lymphomas, breast cancer, liver cancer etc., may be considered for the evaluation of several endpoints, as described above, for the purpose of the development of an *in vitro* carcinogenicity prediction method.

References

1. Kirkland D, Aardema M, Henderson L, & Móller L (2005) Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens I. Sensitivity, specificity and relative predictivity. *Mutat Res* 584(1-2):1-256.
2. Caldwell DJ (1999) Review of mononuclear cell leukemia in F-344 rat bioassays and its significance to human cancer risk: A case study using alkyl phthalates. *Regul Toxicol Pharmacol* 30(1):45-53.
3. Gold LS, Slone TH, Stern BR, & Bernstein L (1993) Comparison of target organs of carcinogenicity for mutagenic and non-mutagenic chemicals. *Mutat Res* 286(1):75-100.
4. Gripon P, et al. (2002) Infection of a human hepatoma cell line by hepatitis B virus. *Proc Natl Acad Sci U S A* 99(24):15655-15660.
5. Cerec V, et al. (2007) Transdifferentiation of hepatocyte-like cells from the human hepatoma HepaRG cell line through bipotent progenitor. *Hepatology* 45(4):957-967.
6. Guillouzo A, et al. (2007) The human hepatoma HepaRG cells: a highly differentiated model for studies of liver metabolism and toxicity of xenobiotics. *Chem Biol Interact* 168(1):66-73.
7. Kanebratt KP & Andersson TB (2008) Evaluation of HepaRG cells as an in vitro model for human drug metabolism studies. *Drug Metab Dispos* 36(7):1444-1452.
8. Uchida Y, et al. (2010) Analogy between sphere forming ability and stemness of human hepatoma cells. *Oncol Rep* 24(5):1147-1151.
9. Cao L, et al. (2011) Sphere-forming cell subpopulations with cancer stem cell properties in human hepatoma cell lines. *BMC Gastroenterol* 11(1):71.
10. Reya T, Morrison SJ, Clarke MF, & Weissman IL (2001) Stem cells, cancer, and cancer stem cells. *Nature* 414(6859):105-111.
11. Sell S & Leffert HL (2008) Liver cancer stem cells. *J Clin Oncol* 26(17):2800-2805.
12. Guo L, et al. (2011) Similarities and Differences in the Expression of Drug Metabolizing Enzymes between Human Hepatic Cell Lines and Primary Human Hepatocytes. *Drug Metab Dispos*.
13. Jennen DG, et al. (2010) Biotransformation pathway maps in WikiPathways enable direct visualization of drug metabolism related expression changes. *Drug Discov Today* 15(19-20):851-858.
14. Knasmüller S, et al. (2004) Use of human-derived liver cell lines for the detection of environmental and dietary genotoxicants; current state of knowledge. *Toxicology* 198(1-3):315-328.
15. Mathijs K, et al. (2009) Discrimination for genotoxic and non-genotoxic carcinogens by gene expression profiling in primary mouse hepatocytes improves with exposure time. *Toxicol Sci*.
16. Hernandez LG, van Steeg H, Luijten M, & van Benthem J (2009) Mechanisms of non-genotoxic carcinogens and importance of a weight of evidence approach. *Mutat Res* 682(2-3):94-109.
17. Hoenerhoff MJ, Hong HH, Ton TV, Lahousse SA, & Sills RC (2009) A review of the molecular mechanisms of chemically induced neoplasia in rat and mouse models in National Toxicology Program bioassays and their relevance to human cancer. *Toxicol Pathol* 37(7):835-848.
18. Thayer KA & Foster PM (2007) Workgroup report: National Toxicology Program workshop on Hormonally Induced Reproductive Tumors - Relevance of Rodent Bioassays. *Environ Health Perspect* 115(9):1351-1356.
19. Ennever FK & Lave LB (2003) Implications of the lack of accuracy of the lifetime rodent bioassay for predicting human carcinogenicity. *Regul Toxicol Pharmacol* 38(1):52-57.
20. Fielden MR & Zacharewski TR (2001) Challenges and limitations of gene expression profiling in mechanistic and predictive toxicology. *Toxicol Sci* 60(1):6-10.
21. Nuwaysir EF, Bittner M, Trent J, Barrett JC, & Afshari CA (1999) Microarrays and toxicology: the advent of toxicogenomics. *Mol Carcinog* 24(3):153-159.
22. Baird WM, Hooven LA, & Mahadevan B (2005) Carcinogenic polycyclic aromatic hydrocarbon-DNA adducts and mechanism of action. *Environ Mol Mutagen* 45(2-3):106-114.
23. Andrysik Z, et al. (2007) The aryl hydrocarbon receptor-dependent deregulation of cell cycle control induced by polycyclic aromatic hydrocarbons in rat liver epithelial cells. *Mutat Res* 615(1-2):87-97.
24. JECFA (2001) Toxicological evaluation of certain veterinary drug residues in food. *WHO Food Additives Series, No. 49* nos 1041-1052 on INCHEM. .
25. Yang JH & Rhim JS (1995) 2,3,7,8-Tetrachlorodibenzo-p-dioxin: molecular mechanism of carcinogenesis and its implication in human in vitro model. *Crit Rev Oncol Hematol* 18(2):111-127.
26. Chopra M, Gahrs M, Haben M, Michels C, & Schrenk D (2010) Inhibition of apoptosis by 2,3,7,8-tetrachlorodibenzo-p-dioxin depends on protein biosynthesis. *Cell Biol Toxicol* 26(4):391-401.

27. Mocarelli P, *et al.* (2008) Dioxin exposure, from infancy through puberty, produces endocrine disruption and affects human semen quality. *Environ Health Perspect* 116(1):70-77.
28. IARC (2001) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 79: Some Thyrotropic Agents.
29. Lee YS, Chanda D, Sim J, Park YY, & Choi HS (2007) Structure and function of the atypical orphan nuclear receptor small heterodimer partner. *Int Rev Cytol* 261:117-158.
30. van derJagt K, Munn S, Tψrslψv J, & Bruijn Jd (2004) Alternative Approaches can reduce the use of test animals under REACH: Addendum to the report "Assessment of additional testing needs under REACH. Effects of (Q)SARS, risk based testing and voluntary industry initiatives". EUROPEAN COMMISSION, DIRECTORATE GENERAL JRC, JOINT RESEARCH CENTRE, *Institute for Health and Consumer Protection*.

Chapter 8

**Samenvatting en
algemene discussie**

Dit proefschrift heeft getracht de applicatie van toxicogenomics methoden te onderzoeken voor de *in vitro* identificatie van genotoxische en carcinogene eigenschappen van chemische stoffen. De belangrijkste redenen van dit onderzoek waren de grote beperkingen van de op dit moment gebruikte *in vitro* en *in vivo* systemen om genotoxiciteit en carcinogeniteit te voorspellen. In het bijzonder betreffen deze beperkingen de hoge vals-positieve resultaten van de *in vitro* testen (1). Aangezien de positieve resultaten van *in vitro* testen leiden tot dierexperimenten, zijn er zorgen met betrekking tot ethische aspecten. Tevens is de relevantie van proefdier-gebaseerde modellen voor de menselijke gezondheid vaak onduidelijk (2).

Het onderzoek in dit proefschrift is daarom gericht op de ontwikkeling van op toxicogenomics analyses gebaseerde *in vitro* testen, voor de voorspelling van *in vivo* genotoxiciteit en carcinogeniteit van chemische stoffen, waarbij de beperkingen van de momenteel beschikbare testsystemen zal worden teniet gedaan. Daarnaast is het onderzoek gericht op het verder ophelderen van werkingsmechanismen van chemische carcinogenen op basis van de toxicogenomics resultaten.

Selectie van het meest geschikte *in vitro* cellulaire systeem

De eerste uitdaging was om het meest geschikte *in vitro* cellulaire systeem te identificeren dat gebruikt kan worden voor het realiseren van het onderzoeksdoel. Rekening houdend met het feit dat de lever het voornaamste orgaan is waar het metabolisme van chemische carcinogenen in reactieve tussenproducten plaats vindt (3), leek het gebruik van een *in vitro* hepatisch systeem het meest geschikt. Daarom, zoals beschreven in **Hoofdstuk 2**, werden twee menselijke lever cellijnen vergeleken door middel van genexpressie analyse, met als doel om de meest belovende *in vitro* cellulaire systeem te selecteren voor een op transcriptomics gebaseerde classificatie.

De eerste cellijn, HepaRG, komt uit een levercarcinoom van een vrouwelijke patiënt (4). Wanneer uitgezaaid met een lage dichtheid, ontstaan uit HepaRG cellen bij het bereiken van confluente twee soorten cellen. De cellen van het eerste type lijken op hepatocyten, terwijl de cellen van het tweede type de hepatocyte-achtige cellen omringen en afgeplat zijn met een duidelijke cytoplasma. Onder specifieke differentiatie omstandigheden, worden de hepatocyte-achtige cellen verder gedifferentieerd in meer granulaire cellen, sterk lijkend op typische volwassene primaire hepatocyten. De hepatocyte-achtige cellen vertegenwoordigen ongeveer 50 tot 55% van de totale celpopulatie (5). Naast deze morfologische overeenkomsten met volwassen primaire hepatocyten, is voor gedifferentieerde HepaRG cellen ook bewezen dat verschillende cytochroom P450 enzymen tot expressie komen en vertonen ze ook andere metabole functies die normaal in de lever gevonden worden (6, 7). Genoombrede genexpressie analyse zoals uitgevoerd in dit proefschrift onthulde verder onder normale kweekomstandigheden een sterke gelijkens tussen gedifferentieerde HepaRG cellen en primaire menselijke hepatocyten, alsook met post-mortem verzamelde menselijke lever.

De tweede cellijn, HepG2, is oorspronkelijk afkomstig van een 15-jarige mannelijke patiënt met een goed gedifferentieerde hepatocellulair carcinoom. Voor HepG2 cellen is be-

wezen dat ze kankerstamcel eigenschappen bezitten (8, 9). Kankerstamcellen vertegenwoordigen een klein deel van de kankercellen en hebben eigenschappen van embryonale stamcellen en worden verantwoordelijk geacht voor de initiatie, progressie, lokaal en op afstand metastase van tumoren, maar ook voor het mislukken van chemo- en radiotherapie (10, 11). In tegenstelling tot HepaRG cellen, vertonen HepG2-cellen een beperkte expressie van cytochroom P450 enzymen onder normale kweek omstandigheden. Uit genoom-brede analyse van genexpressie bleek verder dat HepG2 cellen een beperkte gelijkenis tonen met primaire hepatocyten, zoals beschreven in dit proefschrift. Toch worden er geactiveerde biotransformatie pathways en een hoge expressie van CYP450-enzymen gevonden na behandeling met chemische stoffen (12-14), die onder vergelijkbare omstandigheden een vergelijkbaar niveau bereiken met dat gevonden in HepaRG cellen.

De vergelijking tussen HepG2 en HepaRG cellijnen door genoom-brede genexpressie analyse hebben een aantal verschillen getoond, zowel op gen- als op pathway-niveau. Echter, de twee cellijnen toonden ook overeenkomsten na behandeling met chemische stoffen. De vergelijkbare responsen na behandeling met chemische stoffen bestaan uit pathways die van belang zijn in chemische carcinogenese, waaronder cel cyclus en apoptose pathways.

In een eerste poging om een classificatie uit te voeren en het meest geschikte cellulaire systeem voor deze applicatie te identificeren, werden twee genotoxische en twee niet-genotoxische carcinogenen geselecteerd voor de ontwikkeling van een voorspellingsmodel, dat vervolgens werd gevalideerd door een derde niet-genotoxisch carcinogeen. Deze analyse, beschreven in **Hoofdstuk 2**, laat duidelijk zien dat HepG2 cellen HepaRG cellen overtreffen op het gebied van evaluatie van genotoxiciteit.

HepG2 cellen werden geselecteerd als het meest geschikte systeem voor het verdere onderzoek: ten eerste, vanwege hun vermogen om nauwkeuriger genotoxische van niet-genotoxische carcinogenen dan HepaRG cellen te onderscheiden, en ten tweede omdat ze transcriptomics responsen genereren op carcinogenen in belangrijke pathways van carcinogenese. De extrapolatie van de data verkregen met HepG2 cellen naar de menselijk *in vivo* situatie moeten met voorzichtigheid gedaan worden vanwege de beperkte gelijkenis die ze tonen met primaire menselijke hepatocyten en menselijke lever. Echter, als gevolg van hun overeenkomsten met kankerstamcellen, kunnen HepG2 cellen een meer passend model zijn voor het carcinogenese onderzoek.

Geschiktheid en beperkingen van HepG2 voor genotoxiciteit evaluatie

De volgende inspanningen waren gericht op het testen van de geschiktheid van HepG2 cellen om echte *in vivo* genotoxische stoffen te kunnen discrimineren van vals-positieve genotoxische stoffen (dat wil zeggen stoffen die worden weergegeven als genotoxisch *in vitro*, maar als niet-genotoxisch *in vivo*). Daartoe werden in **Hoofdstuk 3** de effecten van drie vals-positieve genotoxische stoffen op het fenotypische en het transcriptomics niveau geanalyseerd in HepG2 cellen en vergeleken met de effecten van echte *in vivo* genotoxinen en niet-genotoxinen. Van de drie onderzochte vals-positieve genotoxische stoffen toonde 8-

hydroxyquinoline alleen gelijkenissen met de geanalyseerde echte genotoxische stoffen op het fenotypisch niveau, en daarmee van bewijs leverend dat voor de beoordeling van de genotoxische eigenschappen van chemische stoffen de analyse van transcriptomics responsen een nauwkeurigere methode is dan de analyse van fenotypische eindpunten. Dit kan van belang zijn voor het hoge percentage vals-positieve van de momenteel beschikbare *in vitro* genotoxiciteitstesten, omdat deze alleen fenotypische effecten evalueren. Slechts één vals-positieve genotoxisch stof, Quercetin, toonde soortgelijke transcriptomics responsen als de gebruikte genotoxische stoffen. Dit onderzoek levert dus het eerst bewijs van het principe dat de analyse van transcriptomics responsen in HepG2 cellen geschikt is voor de evaluatie van genotoxiciteit, en dat het tot een verlaagd vals-positieve percentage leidt in vergelijking met de momenteel gebruikte *in vitro* testsystemen voor de evaluatie van genotoxiciteit *in vivo*. Tevens waren de gevonden overeenkomsten tussen Quercetin en echte genotoxinen duidelijk in de vroege fase van de blootstelling, wat dus aangeeft dat de duur van de behandeling een belangrijke factor is voor de nauwkeurige classificatie van genotoxische eigenschappen van chemische stoffen.

Het volgende doel was om een nauwkeurig voorspellingsmodel voor *in vivo* genotoxiciteit te ontwikkelen. Aangezien tijdsafhankelijke genexpressie veranderingen voor verschillende chemicaliën uitvoerig gedemonstreerd zijn (15), is tegelijk de rol van de behandelingsduur met chemische stoffen voor de classificatie onderzocht. Om dit te bereiken op een statistisch betrouwbare niveau, zijn een groot aantal chemische stoffen nodig. Daarom werden in **Hoofdstuk 4** transcriptomics veranderingen geïnduceerd in HepG2 cellen door 34 chemicaliën na een behandeling van 12u, 24u en 48u gebruikt voor de selectie van genensets die geschikt zijn voor de discriminatie van *in vivo* genotoxinen en *in vivo* niet-genotoxinen. Voor de selectie van dergelijke genensets werden verschillende methoden toegepast die leidden tot de ontwikkeling van verschillende voorspellingsmodellen. De eerste methode was alleen gebaseerd op transcriptomics gegevens, de tweede integreerde ook de beschikbare resultaten voor deze teststoffen van de Ames bacteriële mutageniciteitstest, terwijl de laatste het geïntegreerde resultaat van de standaard *in vitro* genotoxiciteitsstudies voor de verschillende chemische stoffen mee nam. De toepassing van deze methoden heeft geleid tot de identificatie van verschillende genensets voor elk van de blootstellingstijden. Deze analyse toonde aan dat de genensets gegenereerd na 12u blootstelling in het algemeen leidden tot de slechtste voorspelling; hiermee werd 12 uur uitgesloten als de behandelingsduur voor evaluatie van genotoxiciteit. Genensets gegenereerd na 48u blootstelling presteerde goed voor de verschillende genotoxiciteit voorspellingsmodellen. Echter, de meest veelbelovende genenset werd gegenereerd na stratificatie van chemicaliën op basis van de publiekelijk beschikbare resultaten van de Ames mutageniciteitstest en vooral na 24 uur behandeling. Om deze observaties te bevestigen, werd de validatie van deze resultaten uitgevoerd voor 24u en 48u blootstelling met behulp van een extra set van 28 chemicaliën. Dit bevestigde de eerste waarnemingen dat de beste voorspelling wordt bereikt na 24u blootstelling via de genensets die werden geselecteerd na stratificatie van chemicaliën op basis van de resultaten van de Ames bacteriële genmutatietest. In het bijzonder,

deze methode resulteerde in 89% nauwkeurigheid en 91% specificiteit, waardoor het duidelijk de standaard *in vitro* testbatterij overtreft. De resultaten van **Hoofdstuk 4** bevestigen dat de analyse van transcriptomics responsen in HepG2 cellen na behandeling met chemische stoffen met succes kunnen worden gebruikt als een verbeterde *in vitro* benadering voor het voorspellen van *in vivo* genotoxiciteit, zeker wanneer ze worden gebruikt in combinatie met de resultaten van de Ames bacteriële genmutatie test. Het gebruik van deze methode doet daarom niet alleen de beperkingen teniet van de momenteel gebruikte *in vitro* genotoxiciteitstesten, maar houdt ook een grote belofte voor het verminderen van dierexperimenten voor genotoxiciteitstesten. Een aantal mogelijke toepassingen van deze methode voor wat betreft de vermindering van dierexperimenten zal hieronder worden behandeld in de sectie *Toekomst perspectieven en aanbevelingen*.

Desalniettemin moeten resultaten verkregen met deze methode nog op een zorgvuldige en voorzichtige manier worden geïnterpreteerd, in het bijzonder met betrekking tot de extrapolatie naar de genotoxische gevaren voor de mens, vooral vanwege het feit dat voor de ontwikkeling van deze test vooral chemicaliën werden gebruikt die genotoxisch zijn in knaagdieren. Dit was voornamelijk te wijten aan het gebrek aan voldoende informatie over genotoxische effecten van de geselecteerde stoffen voor mensen. Verder, zijn er een aantal bedenkingen met betrekking tot de indeling van chemicaliën in de genotoxische en niet-genotoxische klassen. Bij de verwerking van beschikbare informatie uit *in vivo* genotoxiciteitsstudies, ondervonden wij vaak moeilijkheden bij de classificatie van chemische stoffen als genotoxisch of niet-genotoxische. Dit was vooral te wijten aan de niet eenduidige resultaten van de *in vivo* studies zoals die in de openbare literatuur werden gerapporteerd. Om deze reden werden alleen duidelijk positieve resultaten in aanmerking genomen voor de indeling van chemische stoffen in de genotoxische klasse.

De tweede onzekerheid met betrekking tot deze methode, verwijst naar de eerder vermelde beperkte metabole capaciteit van HepG2 cellen. Hoewel onze studies hebben aangetoond dat HepG2 cellen de noodzakelijke enzymen hebben voor het metabolisme van een aantal van de onderzochte stoffen (**Hoofdstuk 2** en **Hoofdstuk 3**), is de expressie van bepaalde CYP450-enzymen zeer laag of niet detecteerbaar in HepG2-cellen (12). Bijna alle genotoxische chemicaliën die werden gebruikt in dit proefschrift correct zijn geclassificeerd, ook de velen die moeten worden gemetaboliseerd. Ondanks dit, moeten de resultaten van deze test voor chemicaliën waarvan het metabolisme niet bekend is, verder worden onderzocht om te bewijzen dat deze chemicaliën inderdaad worden gemetaboliseerd, en bovendien om de verkregen resultaten van de genotoxiciteitsvoorspelling te ondersteunen.

Geschiktheid en beperkingen van HepG2 voor carcinogeniteit evaluatie

Terwijl meer dan 90% van de chemische stoffen die door het Internationaal Agentschap voor Kankeronderzoek (IARC) als carcinogeen voor mensen zijn geclassificeerd via genotoxische mechanismen werken, werken de resterende 10% van deze chemicaliën via een

verscheidenheid van niet-genotoxische mechanismen, dus zonder het induceren van DNA-schade (16). Omdat bijna alle bekende genotoxische stoffen ook kankerverwekkend zijn, kan de toepassing van onze verbeterde methode voor het voorspellen van genotoxiciteit *in vivo* met succes worden gebruikt voor de identificatie van de meerderheid van carcinogene stoffen. Echter, het carcinogene potentieel van niet-genotoxische chemicaliën kan niet worden bepaald door onze methode. Daarom is geprobeerd met het onderzoek beschreven in **Hoofdstuk 6** om de transcriptomics responsen in HepG2 cellen te analyseren voor de identificatie van niet-genotoxische carcinogenen. Rekening houdend met de belangrijkste beperking van de proefdiertesten voor carcinogeniteit evaluatie, namelijk hun beperkte relevantie voor de menselijke gezondheidsrisico's (2, 17, 18), was dit onderzoek specifiek gericht op de beoordeling van het risico op kanker voor mensen van de voor proefdieren niet-genotoxische carcinogenen. De onvermijdelijke beperking van deze studie was het geringe aantal niet-genotoxische stoffen met bewezen kankerverwekkende risico's voor de mens (16) en het feit dat er geen niet-genotoxische chemicaliën bestaan waarvan bewezen is dat ze geen kankerverwekkende effecten op mensen hebben (19). Om deze reden werden stoffen die niet-carcinogeen zijn voor proefdieren, en die over het algemeen worden beschouwd als veilig voor de mens, voor deze studie gebruikt. De pogingen om voor proefdieren niet-genotoxische carcinogenen te classificeren als menselijke carcinogenen of menselijke niet-carcinogenen lukte niet, aangezien meestal niet duidelijke resultaten werden verkregen. Dit kan zijn, ten eerste, vanwege het feit dat niet alle mogelijke niet-genotoxische carcinogene werkingsmechanismen werden gedekt door de geselecteerde stoffen vanwege het gebrek aan voldoende teststoffen, en ten tweede, als gevolg van het feit dat transcriptomics responsen niet voldoende zijn voor de discriminatie van niet-genotoxische carcinogenen.

Desalniettemin, de analyse van transcriptomics responsen van HepG2 cellen na behandeling met een aantal van de menselijke niet-genotoxische carcinogenen bevestigt hun werkingsmechanismen, zoals zal worden besproken in de volgende sectie. Dit suggereert dat HepG2 cellen een aantal van de menselijke niet-genotoxische carcinogenen kunnen identificeren. Om dit potentieel van de HepG2 model te benutten, is het nodig dat alternatieve benaderingen dan die gebruikt in **Hoofdstuk 6** in ogenschouw worden genomen voor het verhogen van de nauwkeurigheid van het voorspellen van de menselijke carcinogeniteit. Sommige van deze opties zullen verder worden besproken in de sectie *Toekomst perspectieven en aanbevelingen*.

Geschiktheid en beperkingen van HepG2 voor mechanistische studies

De selectie van toxicogenomics als een instrument voor het doel van dit proefschrift is voornamelijk gebaseerd op zijn potentieel voor classificatie (20). Echter, toxicogenomics studies houden ook een grote belofte voor het identificeren van werkingsmechanismen van toxische stoffen (21). Daarom was het secundaire doel van dit proefschrift om toxicogenomics tools, in combinatie met fenotypische analyses, toe te passen in HepG2 cellen informatie te verkrijgen met betrekking tot de werkingsmechanismen van chemicaliën.

De chemicaliën die het meest diepgaand werden onderzocht, zijn deze bekende stoffen: Benzo[a]pyreen, een genotoxisch carcinogeen, 17-beta Estradiol, een niet-genotoxisch carcinogeen met bewijs van *in vitro* genotoxische effecten (vals-positief genotoxisch) en 2,3,7,8-tetrachloordibenzo-para-dioxine (TCDD), een niet-genotoxisch carcinogeen.

Benzo[a]pyreen is bekend voor het vormen van DNA adducten, na metabolisatie, die kan leiden tot celcyclusstilstand of apoptose (22, 23). In HepG2-cellen, leidde Benzo[a]pyreen blootstelling tot de wijziging van de expressie van genen die betrokken zijn bij apoptose, DNA-schade en de celcyclus regulering, zoals beschreven in **Hoofdstuk 2** en **Hoofdstuk 3**, al vanaf 12u na de blootstelling. De studies in **Hoofdstuk 3** en **Hoofdstuk 5** bevestigden deze effecten ook op fenotypisch niveau, echter met een reductie van DNA-schade op de latere tijdstippen. Dit alles toont aan dat HepG2 cellen in staat zijn om Benzo[a]pyreen te metaboliseren naar zijn actieve metabolieten, waardoor de detectie van Benzo[a]pyreen-geïnduceerde effecten door zowel transcriptomics als fenotypische analyse mogelijk is. Dit wijst er verder op dat HepG2 een geschikt systeem is voor het bestuderen van genotoxische mechanismen.

Echter, wanneer de effecten van vals-positieve genotoxische chemicaliën worden geanalyseerd, bleek dat ze ook genotoxische effecten kunnen induceren in HepG2 cellen. Bijvoorbeeld, zoals beschreven in **Hoofdstuk 3**, induceerden zowel Quercetin en 8-hydroxyquinoline DNA-schade. DNA-schade in HepG2 werd ook ontdekt door fenotypische analyse na behandeling met de vals-positieve genotoxische stof 2-Chloromethylpyridine, zoals beschreven in **Hoofdstuk 5**. Echter, alleen Quercetin heeft de expressie van genen die betrokken zijn bij DNA-schade herstelprocessen geïnduceerd. Dit suggereert dat Quercetin inderdaad een *in vitro* genotoxische stof (**Hoofdstuk 3**) vertegenwoordigt. Desalniettemin, classificeerde de test ontwikkeld voor het voorspellen van genotoxiciteit *in vivo* in **Hoofdstuk 4**, Quercetin correct als een *in vivo* niet-genotoxische stof. De andere twee vals-positieve stoffen werden ook correct geïdentificeerd als *in vivo* niet-genotoxinen door het voorspellingsmodel ontwikkeld in **Hoofdstuk 4**.

De correcte voorspelling als niet-genotoxisch werd ook bereikt voor de bekende vals-positieve genotoxisch carcinogeen 17-beta Estradiol in alle classificatie benaderingen beschreven in **Hoofdstuk 2**, **Hoofdstuk 3** en **Hoofdstuk 4**. Echter, de afwezigheid van bewijs van genotoxiciteit is niet voldoende voor de opheldering van het werkingsmechanisme. Terwijl 17-beta Estradiol de natuurlijke ligand van de oestrogeen receptor is – een ligand-geactiveerde transcriptie factor –, de transcriptionele regulatie analyse in **Hoofdstuk 6** faalde om deze activering te identificeren. Van 17-beta Estradiol is verder bekend dat het de volgende fysiologische gebeurtenissen beïnvloedt: ontwikkeling van reproductieve weefsels, botgroei, lipide homeostase en celproliferatie (24). Analyse van de transcriptomics effecten op pathway-niveau toont aan dat 17-beta Estradiol de expressie van genen betrokken bij de ontwikkeling en proliferatie veranderde. Dus, transcriptomics analyse van door 17-beta Estradiol geïnduceerde effecten in HepG2-cellen was deels in staat om effecten in verband met het werkingsmechanisme van dit niet-genotoxisch carcinogeen te detecteren. Deze effecten waren meer zichtbaar op latere tijdstippen, zoals beschreven in **Hoofdstuk 3**.

In tegenstelling tot de gedeeltelijke identificatie van het werkingsmechanisme voor 17-beta Estradiol, werden duidelijke resultaten verkregen voor het werkingsmechanisme van de bekende niet-genotoxisch carcinogeen TCDD. TCDD is bekend voor de activering van de Aryl hydrocarbon receptor die de expressie van verschillende genen, waaronder het CYP450 enzym CYP1A1, regelt. De veranderde expressie van verschillende genen leidt vervolgens naar de bevordering van de celgroei en inhibitie van apoptose, alsook naar endocriene verstoring (25-27). In HepG2-cellen, heeft TCDD blootstelling inderdaad de Aryl hydrocarbon receptor geactiveerd, resulterend in de verhoogde expressie van CYP1A1, zoals beschreven in **Hoofdstuk 2** en **Hoofdstuk 6**, en heeft het invloed op de expressie van verschillende genen die betrokken zijn bij celgroei en homeostase en hormoon-gerelateerde processen, zoals beschreven in de aanvullende gegevens van **Hoofdstuk 3** en in **Hoofdstuk 6**. Deze effecten, in tegenstelling tot de effecten van 17-beta Estradiol, werden geïdentificeerd bij alle blootstellingtijden. Dit suggereert dat voor niet-genotoxische stoffen verschillende blootstellingtijden nodig zijn voor de identificatie van werkingsmechanisme door toxicogenomics analyses. Dit wordt verder ondersteund door de waarnemingen met betrekking tot de activering van de Constitutive Androstane Receptor door fenobarbital. Fenobarbital, dat een niet-genotoxisch carcinogeen is, werkt via de activering van deze ligand-afhankelijke transcriptie factor (28). De transcriptomics effecten van Fenobarbital in HepG2 cellen na 24u blootstelling kon de activatie van deze receptor niet ondersteunen, zoals beschreven in **Hoofdstuk 4**. Echter een analyse van de transcriptomics responsen van HepG2 cellen na behandeling met Fenobarbital voor 48u, zoals beschreven in **Hoofdstuk 6**, identificeerde de activering van Constitutive Androstane Receptor. Dit suggereert dat langere looptijden noodzakelijk zijn voor de identificatie van de effecten van Fenobarbital in HepG2 cellen.

Dus, waar dit proefschrift bewijst dat HepG2 een geschikt model is voor de studie van genotoxische mechanismen, is het echter ook duidelijk dat HepG2 cellen niet in staat zijn om de werkingsmechanismen van alle niet-genotoxische carcinogenen te detecteren. Verder dient te worden opgemerkt, dat het ontbreken van transcriptomics responsen in HepG2 cellen na behandeling met nieuwe chemicaliën niet verkeerd mag worden geïnterpreteerd als afwezigheid van carcinogeniteit, omdat deze duidelijk niet van toepassing is voor Tetrachloorethyleen en Pentachloorfenol. Beiden veroorzaken geen enkele transcriptomics respons in HepG2, hoewel ze bekende carcinogeen voor mensen zijn (**Hoofdstuk 6**).

Tot slot, werd in dit proefschrift de geschiktheid van HepG2 cellen onderzocht voor mechanistische studies met betrekking tot de functie van genen. Vooral de functie van NR0B2 werd onderzocht, zoals beschreven in **Hoofdstuk 5**, met het doel om de biologische relevantie van dit gen te onderzoeken op DNA-schade responsen, aangezien het een van de genen was die geïdentificeerd werd in **Hoofdstuk 4** voor de classificatie. Een RNA-silencing benadering werd toegepast in HepG2 om de rol van NR0B2 te onthullen voor en na de genotoxische behandeling. NR0B2 is een nucleaire receptor die werkt als een repressor van transcriptie en werd geïdentificeerd als een van de genen met de hoogste bijdrage in het voorspellingsmodel. De inhibitie van de expressie ervan, had evenwel geen invloed op de classificatie van de chemische modelstoffen, wat aangeeft dat NR0B2 niet

de bepalende factor voor de classificatie van deze stoffen is. Dit suggereert verder dat andere genen uit de genenset voor de classificatie een hogere bijdrage aan de voorspelling model kunnen hebben en / of dat inhibitie van de expressie van slechts één gen niet voldoende is voor drastische veranderen van het classificatiemodel. Om dit te bewezen, zouden aanvullende experimenten met hogere aantallen teststoffen nodig zijn, om daarmee de statistische kracht te vergroten.

NR0B2 is verder bekend om een verscheidenheid van fysiologische processen te beïnvloeden, zoals vet, cholesterol en glucose metabolisme, homeostatische regulering en ontgifting en drugs metabolisme in de lever (29). Deze processen werden geïdentificeerd te worden gewijzigd in HepG2-cellen die onder normale omstandigheden groeien na het onderdrukken van NR0B2, zoals beschreven in **Hoofdstuk 5**. Aangezien het NR0B2 netwerk niet beïnvloed lijkt te worden in cellen die NR0B2 tot expressie brengen, lijken HepG2 cellen, alhoewel ze een hepatocellulair carcinoom oorsprong hebben, de functies van niet-kanker-cellen te behouden. De silencing benadering beschreven in **Hoofdstuk 5**, onthulde verder een belangrijke functie van NR0B2 in chemische responsen, vooral in genotoxische stress responsen zoals afgeleid van fenotypische analyses. Deze observaties ondersteunen dat mechanistische informatie met betrekking tot de functionaliteit van genen kan worden verkregen door het gebruik van HepG2 cellen.

Samenvattend, bewijzen de resultaten beschreven in **Hoofdstuk 2**, **Hoofdstuk 3** en **Hoofdstuk 5** de geschiktheid van HepG2-cellen voor het onderzoek naar genotoxische mechanismen van chemische carcinogenen. Verder geven de resultaten beschreven in **Hoofdstuk 2**, **Hoofdstuk 3**, **Hoofdstuk 4** en **Hoofdstuk 6** aan dat HepG2 cellen een nuttig model voor het bestuderen van een aantal niet-genotoxische werkingsmechanismen van chemische carcinogenen kunnen zijn. Ten slotte hebben de resultaten beschreven in **Hoofdstuk 5** bewezen dat HepG2 cellen kunnen worden gebruikt voor de evaluatie van de functionaliteit van genen.

Conclusie

Ter afsluiting, dit proefschrift laat zien dat HepG2 een geschikte *in vitro* lever model kan zijn voor de evaluatie van genotoxiciteit door middel van toxicogenomics analyses. Voor de discriminatie van echte en vals-positieve genotoxische stoffen, was de toepassing van toxicogenomics analyses meer succesvol dan de evaluatie van de fenotypische eindpunten in HepG2 cellen. Dit proefschrift presenteert ook de ontwikkeling van een genotoxiciteit voorspellingsmethode door het evalueren van een beperkt aantal genen, welke beter is dan de conventionele *in vitro* testbatterij voor de beoordeling van *in vivo* genotoxiciteit. Voor een van deze genen toont dit proefschrift zijn biologische relevantie voor genotoxiciteit responsen. Hiermee is aangetoond dat HepG2 cellen kunnen worden gebruikt voor de voorspelling van de genotoxische eigenschappen van chemische stoffen, alsook voor de studie van genotoxische mechanismen van chemische carcinogenen. Tot slot, geeft dit proefschrift een aantal aanwijzingen dat de toepassing van toxicogenomics benaderingen met behulp van

HepG2 cellen mechanistische informatie over de werkingsmechanismen van niet-genotoxische stoffen kunnen bieden, maar ook dat dit nog niet voldoende is voor de carcinogeniteitsvoorspelling van niet-genotoxische stoffen.

Toekomst perspectieven en aanbevelingen

Zoals eerder vermeld, gaat dit proefschrift over de toepassing van toxicogenomics benaderingen in HepG2 cellen voor twee belangrijke doelen: de voorspelling van genotoxische eigenschappen en de evaluatie van het carcinogeen potentieel van chemicaliën.

Dankzij het lagere vals-positieve percentage met even hoge gevoeligheid van de ontwikkelde methode voor genotoxiciteitsvoorspelling in **Hoofdstuk 4** in vergelijking met de momenteel gebruikte *in vitro* testbatterij, wordt deze methode aanbevolen voor een aantal toepassingen om te leiden tot de vermindering van het aantal dieren in toxiciteitstesten. Deze toepassingen omvatten de vroege fasen van de ontwikkeling van geneesmiddelen en de toepassing in het EU-beleid inzake chemische stoffen, het REACH-programma. Voor het evalueren van intrinsieke genotoxische eigenschappen van chemische stoffen in de ontwikkeling van geneesmiddelen, wordt momenteel een trapsgewijze benadering gebruikt voor genotoxiciteit testen die beginnen met *in vitro* testen, gevolgd door *in vivo* testen. De transcriptomics-gebaseerde methode kan de stopzetting van de echte genotoxische stoffen na *in vitro* genotoxiciteit testen vergemakkelijken, en helpen bij de prioritering van veelbelovende kandidaten die zijn vastgesteld als echte niet-genotoxische stoffen. Op deze manier worden dierenexperimenten op vals-positieve genotoxische stoffen vermeden. Het EU-beleid inzake chemische stoffen, het REACH-programma, is gericht op het verzamelen van voldoende toxicologische informatie van chemische stoffen die worden geproduceerd of geïmporteerd in de EU. Er wordt geschat dat ongeveer 400.000 proefdieren kunnen worden gebruikt voor het testen van genotoxiciteit *in vivo* om aan de voorschriften van de REACH-beleid te vervullen (30). De ontwikkelde methode zoals beschreven in **Hoofdstuk 4** kan worden gebruikt als een eerste beoordeling voor dit doel, waarmee dierexperimenten die gebaseerd zouden zijn op foute uitkomsten van de conventionele *in vitro* testbatterij vermeden kunnen worden.

De ontwikkelde methode kan bovendien ook worden aanbevolen als een instrument voor genotoxiciteitsbeoordeling van nieuwe cosmetica. In dit geval kan de ontwikkelde methode in **Hoofdstuk 4** het gebruik van dieren voor genotoxiciteit evaluatie van cosmetische ingrediënten vervangen. Dit is vooral van belang in het licht van het verbod op dierexperimenten voor cosmetische ingrediënten in de EU (EG-verordening 1223/2009).

Echter, om voor deze methode algemeen aanvaard te krijgen voor de chemische veiligheidsbeoordeling, is erkenning door de regelgevende autoriteiten nodig. Daarom moeten toekomstige inspanningen gericht zijn op deze erkenning. De eerste stappen in deze richting zijn al gemaakt door deze nieuwe methode voor te dragen voor validatie door het Europees Centrum voor de validatie van alternatieve methoden (ECVAM). Deze validatie moet gericht zijn op zowel intra- en inter-laboratorium vergelijkende studies, om de robuustheid en de

reproduceerbaarheid van de test te evalueren. Daarnaast zal de ontwikkeling van gestandaardiseerde procedures nodig zijn om erkenning te bereiken van de regelgevende autoriteiten. Voor dat doel kan de ontwikkeling van specifieke analytische platformen, op basis van RT-PCR arrays of aangepaste microarrays, evenals de ontwikkeling van software voor de analyse van de gegevens, nuttig zijn.

De redenen dat de toxicogenomics benaderingen geen succesvolle methode voor het beoordelen van de carcinogene potentie van niet-genotoxische stoffen opleverde, kan worden gevonden in het beperkte aantal chemicaliën die werden gebruikt voor dit doel. Inclusie van meer stoffen met bekende carcinogene potentie bij mensen kan beter zijn, omdat dan meer werkingsmechanismen kunnen worden beoordeeld. Een andere reden zou kunnen liggen in het geselecteerde celmodel of in de geselecteerde voorspellingsbenaderingen. Bijvoorbeeld, terwijl de selectie van HepG2 cellen voor deze studie was gebaseerd op het feit dat de lever het belangrijkste doelorgaan van chemische carcinogenen in proefdier studies is (3), is het bekend dat van de geselecteerde carcinogenen voor mensen die werden gebruikt in **Hoofdstuk 6** slechts één leverkanker kan veroorzaken bij mensen, terwijl de meerderheid van de carcinogenen lymfomen kan veroorzaken. Een meer relevant celmodel dat de doelweefsels van bekende carcinogenen voor mensen beter vertegenwoordigt, zal dus relevant zijn voor verdere studies naar de identificatie van carcinogene eigenschappen van chemische stoffen. Tot slot, alhoewel analyse transcriptomics responsen in staat was om voor het merendeel van de menselijke niet-genotoxische carcinogenen hun werkmechanisme te identificeren, bleek dat voor sommige chemicaliën niet haalbaar. Dit kan erop wijzen dat transcriptomics analyse niet altijd een geschikt eindpunt voor de evaluatie van carcinogene eigenschappen van niet-genotoxische stoffen is en dus andere eindpunten of combinaties daarvan moeten worden beschouwd, zoals de metabolomics, proteomics of epigenomics analyses. Zo is binnen het domein van de niet-genotoxische carcinogenese nader onderzoek aanbevolen om een meer geschikte *in vitro* celmodel te identificeren. Een specifiek celmodel voor elk van de verschillende soorten kanker, zoals lymfomen, borstkanker, leverkanker, enz., kan worden overwogen voor de evaluatie van de verschillende eindpunten, zoals hierboven beschreven, in het kader van de ontwikkeling van een *in vitro* carcinogeniteit voorspellingsmethode.

Referenties

1. Jemal A, Siegel R, Xu J, & Ward E (2010) Cancer statistics, 2010. *CA Cancer J Clin* 60(5):277-300.
2. OECD (2010) Mortality from Cancer, in OECD, Health at a Glance: Europe 2010 *OECD Publishing*:34-35.
3. Petrucelli N, Daly MB, & Feldman GL (2010) Hereditary breast and ovarian cancer due to mutations in BRCA1 and BRCA2. *Genet Med* 12(5):245-259.
4. Clapp RW, Jacobs MM, & Loechler EL (2008) Environmental and occupational causes of cancer: new evidence 2005-2007. *Rev Environ Health* 23(1):1-37.
5. Oliveira PA, *et al.* (2007) Chemical carcinogenesis. *An Acad Bras Cienc* 79(4):593-616.
6. Hernandez LG, van Steeg H, Luijten M, & van Benthem J (2009) Mechanisms of non-genotoxic carcinogens and importance of a weight of evidence approach. *Mutat Res* 682(2-3):94-109.
7. Ames BN, Lee FD, & Durston WE (1973) An improved bacterial test system for the detection and classification of mutagens and carcinogens. *Proc Natl Acad Sci U S A* 70(3):782-786.
8. OECD (1997) OECD GUIDELINE FOR THE TESTING OF CHEMICALS: Test No. 476: In vitro Mammalian Cell Gene Mutation Test.
9. OECD (1997) OECD GUIDELINE FOR THE TESTING OF CHEMICALS: Test No. 473: In Vitro Mammalian Chromosome Aberration Test.
10. OECD (2010) OECD GUIDELINE FOR THE TESTING OF CHEMICALS: Test No. 487: In vitro mammalian cell micronucleus test.
11. Kirkland D, Aardema M, Henderson L, & Möller L (2005) Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens I. Sensitivity, specificity and relative predictivity. *Mutat Res.* 584(1-2):1-256.
12. ICH (1997) S1B TESTING FOR CARCINOGENICITY OF PHARMACEUTICALS.
13. Nuwaysir EF, Bittner M, Trent J, Barrett JC, & Afshari CA (1999) Microarrays and toxicology: the advent of toxicogenomics. *Mol Carcinog* 24(3):153-159.
14. Ellinger-Ziegelbauer H, Stuart B, Wahle B, Bomann W, & Ahr HJ (2005) Comparison of the expression profiles induced by genotoxic and nongenotoxic carcinogens in rat liver. *Mutat Res* 575(1-2):61-84.
15. Waring JF, Ciurlionis R, Jolly RA, Heindel M, & Ulrich RG (2001) Microarray analysis of hepatotoxins in vitro reveals a correlation between gene expression profiles and mechanisms of toxicity. *Toxicol Lett.* 120(1-3):359-368.
16. Uehara T, *et al.* (2008) A toxicogenomics approach for early assessment of potential non-genotoxic hepatocarcinogenicity of chemicals in rats. *Toxicology* 250(1):15-26.
17. Hamadeh HK, *et al.* (2002) Gene expression analysis reveals chemical-specific profiles. *Toxicol Sci* 67(2):219-231.
18. Aniagu SO, Williams TD, & Chipman JK (2009) Changes in gene expression and assessment of DNA methylation in primary human hepatocytes and HepG2 cells exposed to the environmental contaminants-Hexabromocyclododecane and 17-beta oestradiol. *Toxicology* 256(3):143-151.
19. Maeshima H, Ohno K, Tanaka-Azuma Y, Nakano S, & Yamada T (2009) Identification of tumor promotion marker genes for predicting tumor promoting potential of chemicals in BALB/c 3T3 cells. *Toxicol In Vitro* 23(1):148-157.
20. Jonker MJ, *et al.* (2009) Finding transcriptomics biomarkers for in vivo identification of (non-)genotoxic carcinogens using wild-type and Xpa/p53 mutant mouse models. *Carcinogenesis* 30(10):1805-1812.
21. Nioi P, Pardo ID, Sherratt PJ, & Snyder RD (2008) Prediction of non-genotoxic carcinogenesis in rats using changes in gene expression following acute dosing. *Chem Biol Interact* 172(3):206-215.
22. Ennever FK & Lave LB (2003) Implications of the lack of accuracy of the lifetime rodent bioassay for predicting human carcinogenicity. *Regul Toxicol Pharmacol* 38(1):52-57.
23. Mascolo MG, *et al.* (2010) BALB/c 3T3 cell transformation assay for the prediction of carcinogenic potential of chemicals and environmental mixtures. *Toxicol In Vitro* 24(4):1292-1300.
24. Vanparys P. CR, Aardema M., Gribaldo L., Hayashi M., Hoffmann S. and Schechtman L. (2011) ECVAM Prevalidation of Three Cell Transformation Assays. *ALTEX* 28(1):56-59.
25. Hoenerhoff MJ, Hong HH, Ton TV, Lahousse SA, & Sills RC (2009) A review of the molecular mechanisms of chemically induced neoplasia in rat and mouse models in National Toxicology Program bioassays and their relevance to human cancer. *Toxicol Pathol* 37(7):835-848.
26. Hard GC & Khan KN (2004) A contemporary overview of chronic progressive nephropathy in the laboratory rat, and its significance for human risk assessment. *Toxicol Pathol* 32(2):171-180.

27. Thayer KA & Foster PM (2007) Workgroup report: National Toxicology Program workshop on Hormonally Induced Reproductive Tumors - Relevance of Rodent Bioassays. *Environ Health Perspect* 115(9):1351-1356.
28. Caldwell DJ (1999) Review of mononuclear cell leukemia in F-344 rat bioassays and its significance to human cancer risk: A case study using alkyl phthalates. *Regul Toxicol Pharmacol* 30(1):45-53.
29. Gold LS, Slone TH, Stern BR, & Bernstein L (1993) Comparison of target organs of carcinogenicity for mutagenic and non-mutagenic chemicals. *Mutat Res* 286(1):75-100.
30. Edgar R, Domrachev M, & Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30(1):207-210.
31. Jennen DG, *et al.* (2010) Comparison of HepG2 and HepaRG by whole genome gene expression analysis for the purpose of chemical hazard identification. *Toxicol Sci*.
32. Tibshirani R, Hastie T, Narasimhan B, & Chu G (2002) Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci U S A*. 99(10):6567-6572.
33. Lamb J (2007) The Connectivity Map: a new tool for biomedical research. (Translated from eng) *Nat Rev Cancer* 7(1):54-60 (in eng).
34. Zhang SD & Gant TW (2009) sscMap: an extensible Java application for connecting small-molecule drugs using gene-expression signatures. *BMC Bioinformatics* 10:236.
35. Dennis G, Jr., *et al.* (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 4(5):P3.
36. Tibshirani R, Hastie T, Narasimhan B, & Chu G (2002) Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci U S A*. 99(10):6567-6572.
37. Lamb J, *et al.* (2006) The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313(5795):1929-1935.
38. Williams GM & Iatropoulos MJ (2009) Evaluation of potential human carcinogenicity of the synthetic monomer ethyl acrylate. *Regul Toxicol Pharmacol* 53(1):6-15.
39. Drawz SM & Bonomo RA (2010) Three decades of beta-lactamase inhibitors. (Translated from eng) *Clin Microbiol Rev* 23(1):160-201 (in eng).
40. Weber LW, Boll M, & Stampfl A (2003) Hepatotoxicity and mechanism of action of haloalkanes: carbon tetrachloride as a toxicological model. *Crit Rev Toxicol* 33(2):105-136.
41. Yoshida R & Ogawa Y (2000) Oxidative stress induced by 2,3,7,8-tetrachlorodibenzo-p-dioxin: an application of oxidative stress markers to cancer risk assessment of dioxins. *Ind Health* 38(1):5-14.
42. Perez-Maldonado IN, *et al.* (2005) DDT-induced oxidative damage in human blood mononuclear cells. *Environ Res* 98(2):177-184.
43. Piskac-Collier AL & Smith MA (2009) Lindane-induced generation of reactive oxygen species and depletion of glutathione do not result in necrosis in renal distal tubule cells. *J Toxicol Environ Health A* 72(19):1160-1167.
44. Yu FL (2002) 17-beta Estradiol epoxidation as the molecular basis for breast cancer initiation and prevention. *Asia Pac J Clin Nutr* 11 Suppl 7:S460-466.
45. Zhu BZ & Shan GQ (2009) Potential mechanism for pentachlorophenol-induced carcinogenicity: a novel mechanism for metal-independent production of hydroxyl radicals. *Chem Res Toxicol* 22(6):969-977.
46. Singh S, Vrishni S, Singh BK, Rahman I, & Kakkar P (2010) Nrf2-ARE stress response mechanism: a control point in oxidative stress-mediated dysfunctions and chronic inflammatory diseases. (Translated from eng) *Free Radic Res* 44(11):1267-1288 (in eng).
47. Ito Y & Nakajima T (2008) PPARalpha- and DEHP-Induced Cancers. *PPAR Res* 2008:759716.
48. Guyton KZ, *et al.* (2009) A reexamination of the PPAR-alpha activation mode of action as a basis for assessing human cancer risks of environmental contaminants. (Translated from eng) *Environ Health Perspect* 117(11):1664-1672 (in eng).
49. Yang JH & Rhim JS (1995) 2,3,7,8-Tetrachlorodibenzo-p-dioxin: molecular mechanism of carcinogenesis and its implication in human in vitro model. *Crit Rev Oncol Hematol* 18(2):111-127.
50. Chopra M, Gahrs M, Haben M, Michels C, & Schrenk D (2010) Inhibition of apoptosis by 2,3,7,8-tetrachlorodibenzo-p-dioxin depends on protein biosynthesis. *Cell Biol Toxicol* 26(4):391-401.
51. Mocarelli P, *et al.* (2008) Dioxin exposure, from infancy through puberty, produces endocrine disruption and affects human semen quality. *Environ Health Perspect* 116(1):70-77.
52. Crisp TM, *et al.* (1998) Environmental endocrine disruption: an effects assessment and analysis. *Environ Health Perspect* 106 Suppl 1:11-56.

53. Di Consiglio E, De Angelis G, Traina ME, Urbani E, & Testai E (2009) Effect of lindane on CYP-mediated steroid hormone metabolism in male mice following in utero exposure. *J Appl Toxicol* 29(8):648-655.
54. Cowles C, Mally A, & Chipman JK (2007) Different mechanisms of modulation of gap junction communication by non-genotoxic carcinogens in rat liver in vivo. *Toxicology* 238(1):49-59.
55. Bugelski PJ, *et al.* (2010) Critical review of preclinical approaches to evaluate the potential of immunosuppressive drugs to influence human neoplasia. *Int J Toxicol* 29(5):435-466.
56. IARC (2001) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 79: Some Thyrotropic Agents.
57. Ross J, *et al.* (2010) Human constitutive androstane receptor (CAR) and pregnane X receptor (PXR) support the hypertrophic but not the hyperplastic response to the murine nongenotoxic hepatocarcinogens phenobarbital and chlordane in vivo. *Toxicol Sci* 116(2):452-466.
58. JECFA (2001) Toxicological evaluation of certain veterinary drug residues in food. *WHO Food Additives Series, No. 49* nos 1041-1052 on INCHEM. .
59. Benane SG, Blackman CF, & House DE (1996) Effect of perchloroethylene and its metabolites on intercellular communication in clone 9 rat liver cells. *J Toxicol Environ Health* 48(5):427-437.
60. Umemura T, *et al.* (1999) Pentachlorophenol (PCP) produces liver oxidative stress and promotes but does not initiate hepatocarcinogenesis in B6C3F1 mice. *Carcinogenesis* 20(6):1115-1120.
61. IARC (1997) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 69: Polychlorinated Dibenzo-para-Dioxins and Polychlorinated Dibenzofurans.
62. IARC (2011) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 100A: A Review of Human Carcinogens: Pharmaceuticals.
63. IARC (1991) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 53: Occupational Exposures in Insecticide Application, and Some Pesticides
64. IARC (1999) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 72: Hormonal Contraception and Post-menopausal Hormonal Therapy.
65. Huff JE, Salmon AG, Hooper NK, & Zeise L (1991) Long-term carcinogenesis studies on 2,3,7,8-tetrachlorodibenzo-p-dioxin and hexachlorodibenzo-p-dioxins. *Cell Biol Toxicol* 7(1):67-94.
66. IARC (1990) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 50: Pharmaceutical Drugs.
67. IARC (1987) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Supplement 7: Overall Evaluations of Carcinogenicity: An Updating of IARC Monographs Volumes 1 to 42.
68. IARC (1995) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 63: Dry Cleaning, Some Chlorinated Solvents and Other Industrial Chemicals
69. IARC (1999) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 71: Re-evaluation of Some Organic Chemicals, Hydrazine and Hydrogen Peroxide.
70. IARC (2000) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 77: Some Industrial Chemicals.
71. IARC (1996) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 66: Printing Processes and Printing Inks, Carbon Black and Some Nitro Compounds.
72. IARC (1978) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 16: Some Aromatic Amines and Related Nitro Compounds.

Acknowledgements

This thesis is over and together a big chapter of my life that lasted for a little more than four years. During these years a lot of people supported and helped me, each one with their own way, so that today I can write this last chapter.

I came to Maastricht in May 2007, to start working on this project and for that I would like to thank my promotor, Jos Kleinjans and my co-promotor, Joost van Delft. As every PhD student, soon after the project started, I had to get in the lab and start my experiments. That, of course, had to be done under the appropriate guidance for which Marcel van Herwijnen was responsible. As the time was passing by, more people joined the group and eventually helped me with this project. Firstly, Danyel Jennen joined the team as my daily supervisor. Later on, Sandra Claessen also joined the team and her contribution was major in technical and experimental issues. Towards the end of the project more people got involved and their contribution was equally important for finalizing this project. Maria Tsamou's contribution was essential regarding the selection of chemicals. Leen Timmermans provided her knowledge on siRNA that led to the realization of chapter 5. Finally, Stan Gaj and Charly John completed the team by providing their bioinformatics expertise.

These four years however, would be much more difficult, without a nice and collaborative environment that was provided from all the other colleagues from both the GRAT and the later TGX department and for that reason, I am grateful to all of them. Especially to my roommates, Kevin, Martin and Joep, that could handle listening to Greek music almost all day long.

However, during these four years I also had a lot of support from people outside the doors of the GRAT/TGX department, friends and family members that added a lot of fun in my life and often distracted me when things were not going as planned. Some of them were Daneida Lizzarraga, Cinzia Caolo, Mattia Marino, Ineke Erwich, Nina Lyrarakis, Babrak Ibrahimy, Sofia Magkouloufopoulou and of course my parents Mary and Nikos Magkouloufopoulos.

The last thing I had to do in order to finalize my PhD was this book and I would like to thank Lia for the lay out, Babrak and Joost for the dutch translation.

Indeed, four years and a little more, passed with ups and downs... but most importantly with great companions, whose contribution will always be appreciated.

Christina Magkouloufopoulou

Curriculum vitae and list of publications

Curriculum vitae

Christina Magkoufopoulou was born on November 5th 1984, in Thessaloniki, Greece. In 2002, after finishing high school at the 19th High School of Thessaloniki, she studied Molecular Biology and Genetics at the Democritus University of Thrace, in Greece. She conducted research for her Diploma Thesis at the Department of Molecular Biology, Adolf Butenandt Institute, Ludwig-Maximilians University, in Munich, Germany, where she investigated the substrates of Jil1 kinase in the Drosophila Dosage Compensation Complex. She obtained her University Diploma in March 2007.

In May 2007 she was appointed as a PhD student at the Department of Health Risk Analysis and Toxicology (GRAT) in Maastricht University, the Netherlands. In January 2011 she continued working as a PhD student in the newly developed Toxicogenomics (TGX) department, Maastricht University. This thesis describes the research conducted from May 2007 until August 2011 under the supervision of Prof. Dr. Jos Kleinjans and Dr. Joost van Delft.

List of publications

Jennen DGJ, Magkoufopoulou C, Ketelslegers HB, van Herwijnen MH, Kleinjans JCS, van Delft JHM. **Comparison of HepG2 and HepaRG by whole-genome gene expression analysis for the purpose of chemical hazard identification.** Toxicol Sci. 2010 May;115(1):66-79.

Magkoufopoulou C, Claessen SMH, Jennen DGJ, Kleinjans JCS, van Delft JHM. **Comparison of phenotypic and transcriptomic effects of false positive genotoxins, true genotoxins and non-genotoxins using HepG2 cells.** Mutagenesis. 2011 Sep;26(5):593-604.

Magkoufopoulou C, Claessen SHM, Tsamou M, Jennen DGJ, Kleinjans JCS, van Delft JHM. **An *in vitro* transcriptomics-based assay for predicting genotoxicity *in vivo*.** Submitted.

Jennen DGJ, Ruiz-Aracama A, Magkoufopoulou C, Peijnenburg A, Lommen A, van Delft JHM, Kleinjans JCS. **Integrating transcriptomics and metabonomics to unravel modes-of-action of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in HepG2 cells.** BMC Syst Biol. 2011 Aug 31;5:139.

Magkoufopoulou C, Timmermans L, Claessen SMH, Jennen DGJ, Kleinjans JCS, van Delft JHM. **Characterization of the importance of NR0B2 as a classifying gene for genotoxicity prediction: a silencing approach prior and after genotoxicity treatment.** In preparation.

Magkoufopoulou C, Tsamou M, Jennen DGJ, Kleinjans JCS, van Delft JHM. **Transcriptomics based evaluation in HepG2 cells of human carcinogenic potential of known rodent non-genotoxic carcinogens.** In preparation.