Characterization of disturbed neural crest migration as mechanism of developmental toxicity of prescription drugs

Dissertation

zur Erlangung des akademischen Grades

eines Doktors der Naturwissenschaften (Dr. rer. nat.)

vorgelegt von

Giorgia Pallocca

an der

Universität Konstanz

Mathematisch-Naturwissenschaftliche Sektion

Fachbereich Biologie

Tag der mündlichen Prüfung: 15. Februar 2017

1. Referent: Prof. Marcel Leist
2. Referent: Prof. Daniel Dietrich

Oral and poster presentations

Oral presentations:

- Lush prize conference (2016, London, UK). Neural crest cell migration-based in vitro assay to study developmental toxicity. G. Pallocca, M. Leist.

Selection of given poster presentations:

- International Neurotoxicology Association (INA) conference (2013, Egmond aan Zee, The Netherlands). Changes in miRNA expression profiles in human stem cells-derived neuronal culture after exposure to methyl mercury. G. Pallocca, M. Fabbri, M.G. Sacco, L. Gribaldi, D. Pamies, I. Laurenza and A. Bal-Price

- European Society for Alternatives to Animal Testing (EUSAAT) congress (2013, Linz, Austria). Screening of a large group of medical substances and environmental pollutants in a neural crest stem-cell based functional migration assay. G. Pallocca, B. Zimmer, N. Dreser, S. Foerster, T. Waldmann, S. Julien, K.H. Krause, J. G. Henglester, S. Bosgra, M. Leist

- 9th World Congress on Alternatives and Animal Use in the Life Sciences (2014, Prague, Czech Republic). Profiling of drugs and environmental chemicals for functional impairment of neural crest migration in a novel stem cell-based test battery. G. Pallocca, B. Zimmer, M. Grinberg, N. Dreser, S. Foerster, T. Waldmann, S. Julien, K.H. Krause, J. G. Hengstler, S. Bosgra, M. Leist

- European Food Safety Authority (EFSA) 2nd scientific conference (2015, Milan, Italy). Identification of transcriptome signatures and biomarkers specific for migration-inhibiting potential developmental toxicants in human neural crest cells. G. Pallocca, M. Grinberg, T. Waldmann, J. G. Hengstler, M. Leist

Table of contents

A.Summary

In the last years, different individual human embryonic stem cell-based developmental toxicity test systems have been established and have been proven to offer new possibilities to explore toxicological hazard directly on relevant and non-transformed human cells. A further achievement has been the combination of these assays to comprehensive batteries able to predict human developmental toxicity.

In the framework of the European project ESNATS (Embryonic Stem cell-based Novel Alternative Testing Strategies), we developed a test battery which allows the inclusion of any developmental toxicity assay, and that explores the responses of such test systems to a wide range of compounds. As a first step, we selected and characterized a heterogeneous group of compounds with a wide applicability domain, which ranged from environmental pollutants to several prescription drugs. To evaluate the feasibility of the suggested test framework, we performed the initial screen in a well-characterized assay that evaluates 'migration inhibition of neural crest cells' (MINC assay), which finally resulted in the identification of 11 hits (e.g. geldanamycin, arsenite, PBDE-99).

Next, transcriptome analysis for some selected MINC hits was performed. The transcriptome changes triggered by these substances in human neural crest cells (NCC) were recorded and analyzed. Transcript profiling allowed a clear separation of different toxicants. Furthermore, a diagrammatic system was developed to visualize and compare toxicity patterns of a group of chemicals by giving a quantitative overview of altered superordinate biological processes (e.g. KEGG pathways or overrepresented gene ontology terms). Finally, the transcript data were mined for potential markers of toxicity. We found that the inclusion of transcriptome data largely increased the information from the MINC phenotypic test.

As a final step, one of the MINC-positive compounds, the prescription drug interferon-β (IFNβ), was chosen to be further characterized as potential developmental toxicity hazard. We could confirm the adverse effects of IFNβ on NCC migration in different functional assays. The analysis of transcriptome changes suggested a role of altered JAK-STAT signaling in toxicity, which was confirmed by detailed measurements of interferon effects on signaling in the presence of specific kinase inhibitors.

Zusammenfassung

In den vergangenen Jahren wurden verschiedene auf humanen embryonalen Stammzellen basierende Entwicklungstoxizitätstests entwickelt und gezeigt, dass diese neue Möglichkeiten bieten, toxikologische Risiken direkt an relevanten und nichttransformierten menschlichen Zellen zu erforschen. Eine weitere Errungenschaft war die Kombination dieser Tests zu umfassenden Testbatterien, welche erlauben die Entwicklungstoxizität von Substanzen für den Menschen vorherzusagen.

Im Rahmen des europäischen Projekts ESNATS (Embryonic Stem cell-based Novel Alternative Testing Strategies) entwickelten wir eine Testbatterie, welche die Einbindung weiterer Entwicklungstoxizitätstests sowie die Untersuchung der Reaktion dieser Tests auf eine große Bandbreite von Stoffen ermöglicht. Zuerst wurde eine heterogene Gruppe von Verbindungen mit einem breiten Anwendungsgebiet, das von Umweltschadstoffen bis zu verschreibungspflichtigen Medikamenten reicht, ausgewählt und charakterisiert. Um die Realisierbarkeit des vorgeschlagenen Testrahmens zu bewerten, führten wir den ersten Screen in einem etablierten Testsystem durch, welcher die „Migrationsinhibierung von Neuralleistenzellen“ (MINC assay, migration inhibition of neural crest cell assay) bestimmt. Dabei konnte für elf Substanzen (u.a. Geldanamycin, Arsenverbindungen, PBDE-99) ein Effekt gezeigt werden.

Im weiteren Vorgehen wurde eine Transkriptomanalyse für ausgewählte Substanzen, welche im MINC Test einen Effekt gezeigt hatten, durchgeführt, um die in humanen Neuralleistenzellen (NCC) hervorgerufenen Transkriptom-Änderungen aufzuzeichnen und zu analysieren. Diese Transkriptomprofile erlauben eine klare Unterscheidung der verschiedenen Stoffe. Des Weiteren wurde ein Visualisierungssystem entwickelt, das einen quantitativen Überblick über übergeordnete biologische Prozesse (z.B. KEGG Pathways oder überrepräsentierte „Gene Onthology Terms“) bietet, um die Toxizitätsmuster einer Gruppe von Chemikalien zu vergleichen. Abschließend wurden die Transkriptionsdaten nach potentiellen Toxizitätsmarkern untersucht. Wir haben herausgefunden, dass die Einbindung von Transkriptionsdaten den Informationsgehalt des phänotypischen MINC-Tests enorm erhöht.

Als letzter Schritt wurde eine der im MINC-Test auffälligen Substanzen, das verschreibungspflichtige Medikament Interferon β (IFNβ), hinsichtlich weiterer potenzieller entwicklungstoxischer Risiken untersucht. Wir konnten den schädigenden Effekt von IFNβ auf NCC Migration in verschiedenen funktionalen Tests bestätigen. Die Analyse der Transkriptomänderungen weist auf eine Entwicklungstoxizität aufgrund einer Veränderung der JAK-STAT Signaltransduktion hin. Dies konnte durch weitere Untersuchungen der IFNβ Signaltransduktion in Anwesenheit verschiedener spezifischer Kinaseinhibitoren bestätigt werden.

Abbreviations

ADME Absorption, Distribution, Metabolism, Excretion
AKT Protein Kinase B
AO Adverse Outcome
AOP Adverse Outcome Pathway
AraC Cytosine Arabinoside
bFGF basic Fibroblast Growth Factor
BH Benjamini - Hochberg method for p-value adjustment for multiple comparisons
BMC Benchmark Concentration
BMCL Benchmark Concentration 95% confidential interval lower Limit
BoT Biomarker of Toxicity
CC Clinical Concentration
CCR Chemokine Receptors
CM Cytokine Mix
TNFα Tumor Necrosis Factor α
CNS Central Nervous System
CSF Cerebrospinal Fluid
CYP Cytochrome P450
DEG Differentially Expressed Genes
DLHP Dorsolateral Hinge Point
Dll Distal less
DNT Developmental Neurotoxicity
DT Developmental Toxicity
EdU 5-Ethynyl-Deoxyuridine
EGF Epidermal Growth Factor
EMT Epithelial to Mesenchymal Transition
EPA Environment Protection Agency
ERK Extracellular signal Regulated Kinase
ESNATS Embryonic Stem cell-based Novel Alternative Testing Strategies
FACS Fluorescence-activated cell sorting
FAK Focal Adhesion Kinase
FC Fold Change
FCC Free Concentration
FDA Food and Drugs Administration
FDR False Discovery Rate
GA Geldanamycin
GD Gestational Day
GFP Green Fluorescent Protein

GO Gene Ontology
EC Effective Concentration
GSK3β Glycogen Synthase Kinase 3β
HDACi Histone Deacetylase inhibitor
HEK293 Human Embryonic Kidney 293 cell
hESC human Embryonic Stem Cell
hPSC human Pluripotent Stem Cell
HSP Heat Shock Protein
IFI Interferon Inducible Protein
IFN Interferon
IFNAR Interferon alpha and beta Receptor
IFN-β Interferon β
IL-1β Interleukin 1β
IRF Interferon Regulatory Factor
ISG Interferon Stimulated Gene
ISGF Interferon Stimulated Gene Factor
IVIVE *In Vitro In Vivo* extrapolation
JAK Janus Kinase
KE Key Event
KEGG Kyoto encyclopedia of genes and genomes
LOAEC Lowest Observed Adverse Effect Concentration
LOAEL Lowest Observed Adverse Effect Level
LUHMES Lund Human Mesencephalic
MEF Mouse Embryonic Fibroblast cell
mESCn murine Embryonic Stem Cell-derived neural precursor
MHC Major histocompatibility complex
MHP Medial Hinge Point
MINC Migration Inhibition of Neural Crest
MIU Million International Units
MoA Mode of Action
MPP Matrix Metalloproteinases
MS Multiple Sclerosis
MTE Mesenchymal To Epithelial
NC Neural Crest
NCC Neural Crest Cell
NOAEL No Observed Adverse Effect Level
NSPC Neural Stem and Progenitor Cells
OPC Oligodendrocyte progenitor cells
PBDE Polybrominated Diphenyl Ether
PBPK Physiology-Based Pharmacokinetic

PCA Principal Component analysis
PCB Polychlorinated Biphenyl
PoT Pathway of Toxicity
PS Probe Set
qPCR real-time Polymerase Chain Reaction
REACH Registration, Evaluation, Authorisation and Restriction of Chemicals
ROI Region Of Interest
STAT Signal Transducer and Activator of Transcription
SVM Support Vector Machine
TDF Triadimefon
TSA Trichostatin A
VPA Valproic Acid

B. General Introduction

1. The neurulation: neural tube and neural crest formation

1.1 Neurulation: the formation of the neural tube

Neurulation describes the developmental event which starts with the formation of the neural plate and ends with its transformation, and its closure, into the neural tube, the primitive structure of the brain and the spinal cord, in vertebrate embryos.

The neurulation process consists of four crucial steps, which overlap spatially and temporally (Figure 1):

(1) formation and shaping of neural plate;

(2) bending of neural plate;

(3) convergence of neural plate and;

(4) closure of neural tube.

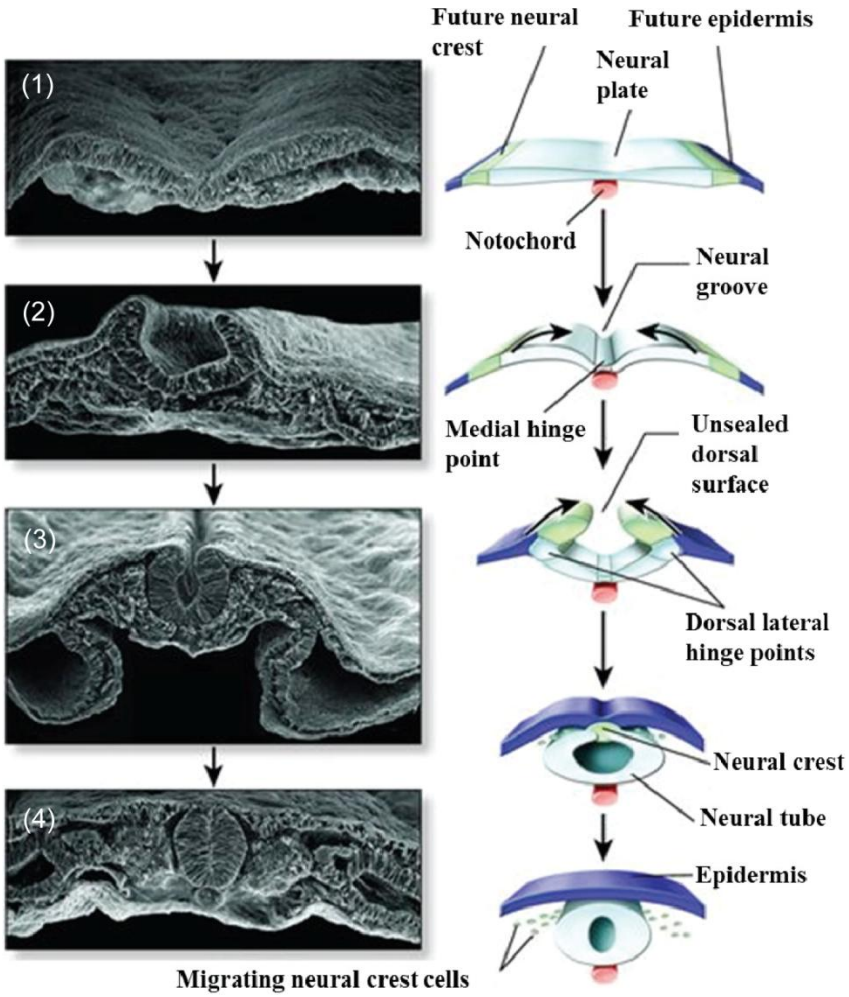


Figure 1 Representation of the different phases of the primary neurulation. Neurulation includes different steps: the formation of the neural plate (1), and its folding (2); the convergence (3) and the fusion (4) of the dorsal surfaces to form the neural tube. Scanning electron microscopy pictures of the correspondent phases taken during chicken embryo-development are shown on the left. Figure modified from biology-forum.com

In the early stages of neurulation, the embryo, called gastrula, consists of the three major germ layers, mesoderm, ectoderm and endoderm. The

neurulation process begins when approximately half of the ectodermal cells modify their morphology, elongating into columnar neural plate cells, by response to the underlying dorsal mesoderm-induced signaling. The different shape distinguishes the neural plate from the surrounding pre-epidermal cells (Gilbert 2010).

At the border of the newly formed neural plate, hinge regions starts to form. Those cells move the lateral edges of the neural plate toward its midline, which is known as medial hinge point (MHP). Next, the MHP cells anchor to the beneath notochord, and undergo further morphological changes. At the border between neural plate and the ectoderm, the dorsolateral hinge points (DLHP) form, which cells undergo similar changes as the MHP cells, reducing their height and adopting wedge-shape. The neural plate begins to bend around those hinge regions, which direct the rotation of the cells around them. Additionally, the surface ectoderm pushes toward the neural plate midline, contributing to the neural plate folding. When the peaks of the neural folds are brought closer, they adhere to each other and merge, forming the neural tube. In vertebrates, the cells at this junction give rise to the neural crest cell population. Finally, the newly formed neural tube will become the embryonic precursor of the central nervous system, which will develop into brain and spinal cord (Gilbert 2010).

1.2 Definition of neural crest

The neural crest (NC) was originally identified by His in 1868 and described as "the cord in between" because of its origin between the neural plate and non-neural ectoderm. More recently, the NC has been renamed the 4^{th} germ layer, representing a vertebrate-specific addition to the classic three germ layers, ectoderm, mesoderm and endoderm. The neural crest population is one of the most significant factors contributing to vertebrate diversity and evolution, adding features such as hinged jaw, special sense of organs and neural circuitry (Munoz and Trainor 2015).

The defining features of the neural crest cells (NCC) are their origin at the neural plate border, their multipotency and their ability to migrate and give rise to a plethora of cell types and tissues in vertebrates.

NCC execute their main developmental steps in separate regions of the embryo, and during different stages of the embryogenesis, characteristic almost exclusive of this specific cell type. NCC development can be divided in four distinct steps and events (Figure 2):

(1) pre-migratory phase, which includes induction and specification at the neural plate, beginning during gastrulation and early neurulation;

(2) delamination phase, via epithelial to mesenchymal transition (EMT) at the end of neurulation;

(3) migratory phase, from the neural plate through the embryo;

(4) differentiation phase, through the course of organogenesis and late embryogenesis.

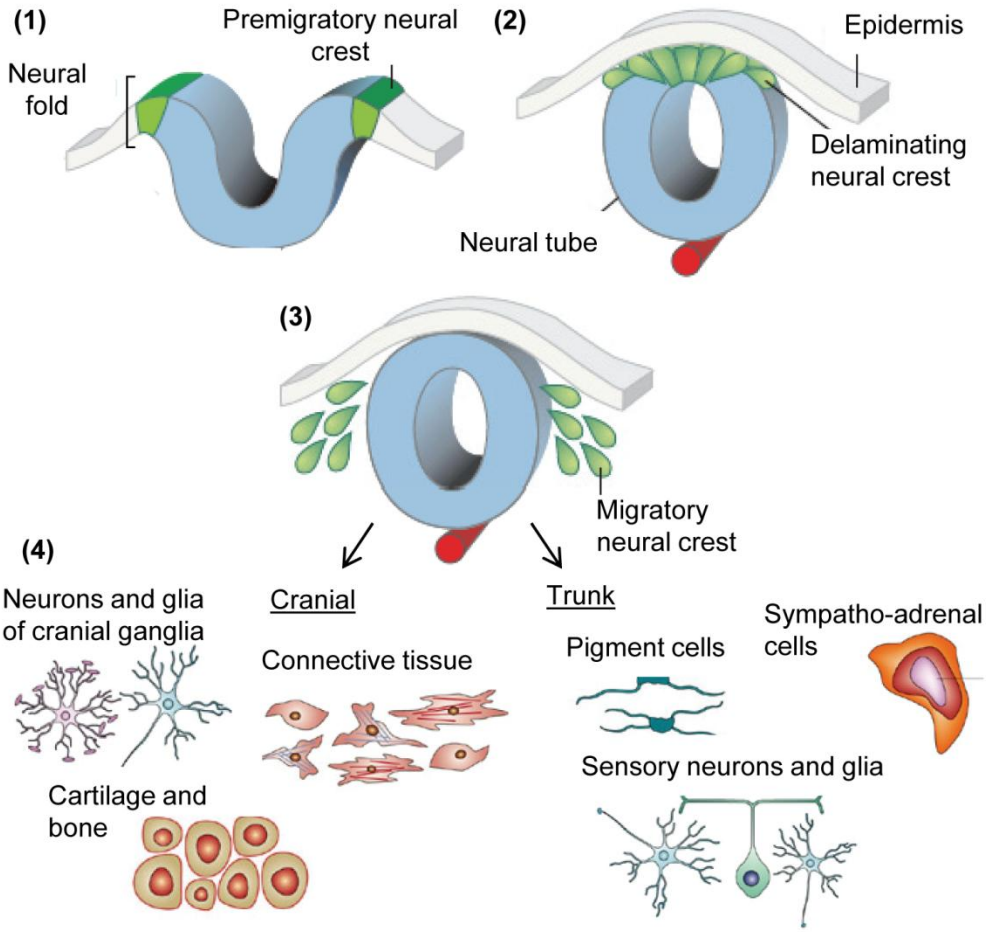


Figure 2 Representation of maturation steps of the neural crest. NCC maturation included different steps: pre-migratory, delamination, migratory and differentiation phases. Finally, NCC will differentiate, based on their position (cranial or trunk), into different celly types, such as sensory neurons, pigment cells, connective tissue, cartilage and bones. Figure modified from Green et al. 2015 and Knecht and Bronner-Fraser 2002.

In most of the species, NCC progenitors development is induced during gastrulation, with the establishment of the neural plate border. During neurulation, ectodermal, placodal, NCC and neural plate progenitors, initially dispersed, get progressively spatially ordered and located, at the beginning in the neural folds, and then in the dorsal aspect of the neural tube. The emergence of the neural crest from the neural plate border is characterized by the coordinated expression of neural crest specifier genes (e.g. TFAP2A, FOXD3, ETS1 and SNAIL1/2), which distinguish the NC from the neighboring placodal regions. They positively regulate each other and induce the following developmental steps (Simoes-Costa and Bronner 2015). At the end of the neural crest specification process, the orchestrated expression of those genes initiates drastic structural changes, resulting in the delamination of the neural crest from the neural tube. This process is named epithelial-to-mesenchymal transition (EMT) and it is a complicated mechanism which brings to extreme structural remodeling of the premigratory neural crest, including regulation of the adhesive characteristics of the cells, cytoskeletal rearrangement, degradation of basement membrane by metalloproteases and inducement of a mesenchymal phenotype which allows the NCC to separate from the neural tube and disperse through the embryo (Figure 3) (Sauka-Spengler and Bronner-Fraser 2008).

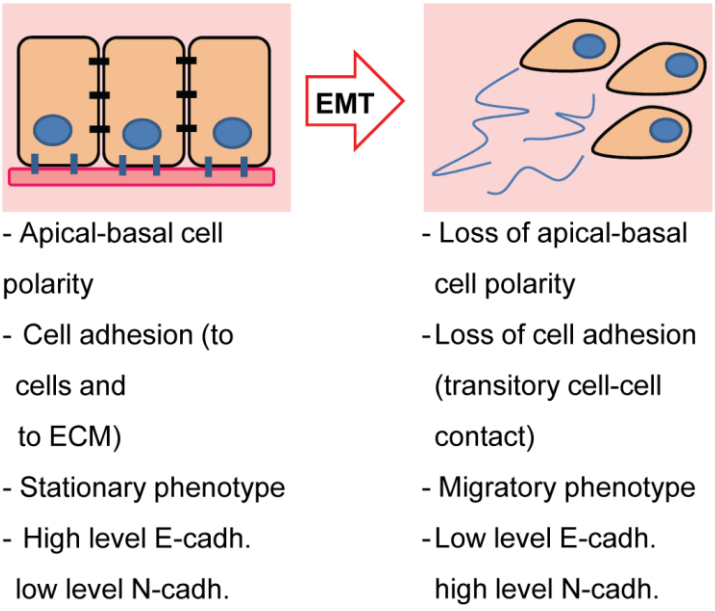


Figure 3 Epithelial to mesenchymal transition (EMT) process. Before the delamination phase, NCC undergo to the process of EMT. The characteristics of epithelial (left) and mesenchymal (right) phenotypes are listed below each graphical representation.

After delamination, migratory neural crest cells starts to colonize different areas of the embryo, responding to different kind of permissive and inhibitory stimuli, which vary along the axial position. During this process, NCC maintain a stem cell-like, multipotent state, including the capacity of self-renewal (Baggiolini et al. 2015). The process of neural crest

diversification starts with activation of different regulatory circuits in the diverse migratory subpopulations. The arising of the different derivatives depends on the combination between the regulatory state of the neural crest and the environmental signals which surround them. Once NCC arrive at their final location, they often self-aggregate during initiation of terminal differentiation (Simoes-Costa and Bronner 2015).

NCC can be divided in cranial, cardiac, vagal, trunk and sacral according to their axial position of origin. Cranial NCC give rise to most of the bone and cartilage of the facial skeleton, neurons and glia of the cranial ganglia, smooth muscle and pigment cells. Cardiac NCC form the valves, septa and outflow tract of the heart. The vagal and sacral NCC form the enteric nervous system; while the trunk NCC differentiate in melanocytes, neurons and glia of the peripheral nervous system (Motohashi and Kunisada 2015).

2 Disturbances of the developmental process

The mechanisms taking place during the developmental process must follow defined timing and spatial rules. Any kind of interferences of the physiological stages can lead to embryo-lethality or developmental defects. The Environment Protection Agency (EPA) defines the major manifestations of developmental toxicity as "death, structural abnormality, altered growth and functional deficiency of the developing organism".

2.1 Epidemiology of the developmental disorders

In the last two decades, an increase in the global incidence of developmental disorders has been registered. In 2006-2010, the prevalence rate of congenital anomalies in Europe reached 2.5%, with one out of ten specifically related to defects of the nervous system (EUROSTAT 2010). Epidemiological studies (Boyle et al. 2011) revealed an increase of the prevalence of developmental disorders of 17% in a 12-year period (1997-2008) in a US cohort. This trend can even reach higher percentages for some specific pathologies, such as autism, for which the incidence has been doubling during the last 10 years (Sullivan 2005, Rutter 2005).

The causes of developmental disorders are extremely heterogeneous. These anomalies can be a direct cause of genetic alterations in the embryo (10-30% of cases) and/or they can have an indirect origin, associated to maternal exposure to external factors, such as infections or

chemicals triggering developmental toxicity (4-13% of the cases). Mostly, developmental abnormalities are induced by a mixed etiology (20-50%) (Mattison 2010). For neuro-developmental disabilities, 3% of cases appears to result directly from toxicant exposure (e.g. lead, methylmercury, PCB or pesticides), while 25% are due to a mixed contribution of genetics-related susceptibility and chemical exposure (NRC 2000) .

2.2 NCC-derived developmental disorders

A large percentage of developmental disorders is caused by NCC deficit. This group of disorders is called neurocristopathies.

Neurocristopathies are a class of disorders that result from abnormal migration, differentiation, division or survival of neural crest cells (NCC) during embryonic development. A neurocristopathy can arise as a single-organ disease, if only one compartment is affected (e.g. Hirschsprung disease), or as a more complex pathology, in the case of different organs are affected (e.g. DiGeorge syndrome, Treacher Collins syndrome) (Mayor and Theveneau 2013). Furthermore neurocristopathies can appear as neonatal cancer (e.g. melanoma, neuroblastoma) or as neonatal malformations (e.g. piebaldism, Waardenburg syndrome, cleft palate, cardiopathies) (Marshall et al. 2014, Keyte and Hutson 2012, Trainor 2010). Malformations can overlap with neural tube closure defects (1 in 1000 births); this can be explained by the presence of common molecular origins at the root of both neural tube closure and NCC disorders (Saint-Jeannet 2006). This kind of alterations can be induced by genetic factors (Lee et al. 2009). Several genetic factors have been identified in the etiology of different neurocristopathies; germline mutations of the neural crest induction genes (e.g. PAX3, MITF, SNAI2 and SOX10), of the NCC migration genes (e.g. ephrinB), and the differentiation genes (e.g. endothelin B receptor, b-RAF and SOX10) have been found to have a direct role in the neurocristopathies etiology (Lee et al. 2009).

Besides the genetics-derived disorders, there is also evidence of neurocristopathies resulting from exposure to pharmaceuticals (e.g. valproic acid, Fuller et al. 2002) and pesticides (e.g. triadimefon, Menegola et al. 2000).

3 Models of developmental disturbances

A developmental toxicant is defined as a compound which, once exposed to a pregnant animal, is not toxic to the mother (1st generation) while it represents a hazard to the developing organism (2nd generation) (EPA 1996) (Figure 4).

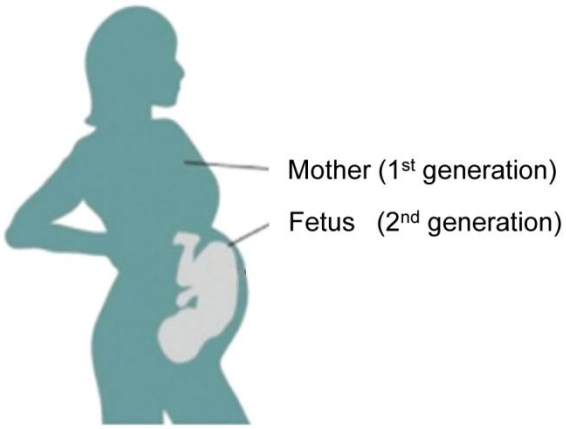


Figure 4 Developmental toxicity models. The developmental toxicity models must resemble human fetal exposure to toxicants. A substance is considered a developmental toxicant when it triggers toxicity in the fetus (2nd generation) but not in the mother (1st generation).

The classical approach utilized for developmental toxicity testing comprises the measurement of apical endpoints such as tissue malformations, mortality and growth retardation in *in vivo* two-generation studies. The animal-based testing approach has a low sensitivity and specificity with respect to human hazard prediction. The evaluation of *in vivo* developmental toxicity testing data indicates a high percentage of false positive (~40%) (Hartung 2009) and of false negative (~55%) classifications (Bremer and Hartung 2004). Furthermore, the concordance among different laboratory mammalian species is lower than 60% (Sipes et al. 2011). Species differences during development are already evident at the transcriptome level. Comparison of mouse and human RNA-seq data in 13 tissues indicated that different tissues within one species show higher similarity than the same tissue between the two species (Lin et al. 2014). These macro-differences reflect finer discrepancies: a greater grade of self-renewal activity of fetal human neocortex neural stem and progenitor cells (NSPC) has been observed when compared to the situation in mouse (Fietz et al. 2012). Different studies based on RNA-seq technology confirmed this diversity. For instance, van de Leemput and colleagues compared RNA-seq data of human and mouse corticogenesis and observed low overall correlation (van de Leemput et al. 2014); while Zeng and colleagues reported a low concordance between the expression profiles of cortical layers in human and mouse (Zeng et al. 2012).

3.1 *In vitro* models & advantage of hPSC use

The developmental process is made of the coordinated activation and repression of very specific and sensitive mechanisms, e.g. migration, differentiation, cell death and proliferation of very different cell types. These processes represent all possible targets of a developmental toxicant. *In vitro* systems offer the advantage of modeling the different mechanisms, as well as simplifying complex processes in singular systems with specific endpoints. The first *in vitro* models made use of transformed human cells, which are easy to handle and need simple culture methods. A large step forward was done with the increasing applicability of human pluripotent cell based system (hPSC); these novel stem cell-based *in vitro* test systems offer new possibilities to explore toxicological hazard directly on relevant and non-transformed human cells. Furthermore, the use of hPSC-based cell systems offers the opportunity to follow the early differentiation processes, recapitulating those particular mechanisms otherwise not available. So far, hPSC-based cell systems have been used to model many different developmental processes from multi-germ layer formation to later events as myelination (Table1).

Developmental stage	Short Description	References
Multigerm layer differentiation	Multilineage differentiation of human embryonic stem cells followed by proteomics analysis	Meganathan et al. 2012
Neuronal rosettes formation	Morphological and molecular changes upon toxic treatment of hESCs differentiate towards neural rosettes	Colleoni et al. 2011, Colleoni et al. 2012
Neuro-ectoderm differentiation	hPSC-derived neuro-epithelial precursor cells are generated. Transcriptional changes of selected marker genes are assessed after toxic treatment.	Balmer et al. 2012; Waldmann et al. 2014
Neural crest migration	hPSC-derived NCC migration is detected after treatment with chemicals.	Zimmer et al. 2012; Zimmer et al. 2014; Dreser et al. 2015
Neurite outgrowth	PNS developing neurons are obtained. Neurite outgrowth is measured after toxic treatment.	Krug et al. 2013b; Hoelting et al. 2016
Gliogenesis	Effects on oligodendrocytes differentiation from primary normal human neural progenitor is analysed under toxic treatment.	Fritsche et al. 2005; Talens-Visconti et al. 2011
Myelination	The functional identity of the hESC-derived OPCs can be best confirmed by their ability to migrate, mature and produce myelin sheaths	Hu et al. 2009; Wang et al. 2013

Table 1 hPSC-based developmental (neuro) toxicity *in vitro* models

3.2 Functional models for NCC-derived developmental disorders

Much information regarding the behavior and the toxic response of human NCC is still lacking. A recent approach to fill this gap was the establishment of an *in vitro* human NCC functional test, as described in Zimmer et al (2012). A differentiation protocol, published by Lee et al. (2007), to obtain a pure NCC population derived from human pluripotent stem cells (hPSC) was established in our group. By this method, hPSC can be induced to differentiate into neural rosettes. These are then manually picked and replated. From these neural structures, the neural crest cells arise and can be isolated and enriched by FACS sorting (Figure 5).

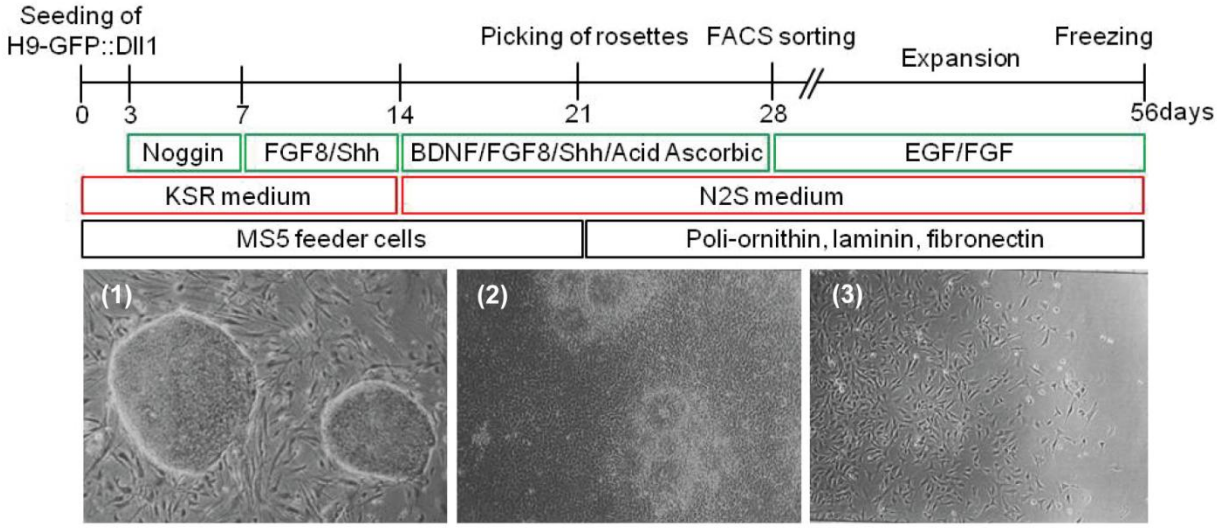


Figure 5 Differentiation protocol from human PSC to human NCC. The hESC H9-GFP::Dll1 reporter cell line is firstly seeded on feeder cells (1) and then induced to differentiate towards a neuronal stage. At this phase, the presence of neural rosettes can be observed (2); from these structures, the NCC will arise. Rosettes are manually picked and transferred into new dishes. Finally, the NCC population is enriched by FACS sorting and the cells can be expanded (3) and frozen. The green boxes indicate the medium supplements; the red boxes the different used media; and the black boxes the different coating used during the differentiation process. Phase contrast pictures for crucial stages are shown below the scheme. Figure modified from Zimmer et al 2012.

The obtained cell population has been implemented in the "migration of neural crest cell" (MINC) assay. This assay was set up as the first human stem cell-based method that is able to detect the functional effects of chemicals on one of the key events of development, the NCC migration. The assay showed good performance during its evaluation with known positive

and negative controls and allowed sensitive screening of environmental toxicants and pharmaceutical (Zimmer et al. 2012, Nyffeler et al. 2016).

3.3 New challenges for *in vitro* developmental toxicity testing: characterization of testing compounds and test systems, and harmonization of test systems into test batteries

Some of the new challenges of *in vitro* developmental toxicity testing include the identification of new developmental test compounds, and the increase of the group of positive developmental toxicants for further characterization of novel test systems.

Lack of comprehensive human DT epidemiological data and deficiency of DT guidelines for animal studies led to a scarcity of generally accepted positive controls in the field of developmental toxicity, e.g. mercury, lead and pesticides (Grandjean and Landrigan 2006).

This situation contributed to create a vicious circle between the lack of positive compounds and poor characterization of new *in vitro* models, which can exploit only a handful of positive compound groups with limited applicability domain.

The current available test systems have been characterized by the use of the few well-known positive control compounds, which mainly belong to the group of environmental toxicants or, at most, one to three drug-like compounds (Laurenza et al. 2013b, Krug et al. 2013b; Pallocca et al. 2013, Balmer et al. 2012). Therefore, it is not clear whether these test systems would be able to reliably predict DT of toxicants. An approach to change this situation is the establishment of recurrent optimization cycles involving testing of novel compounds, assay optimization, and adaptation of interpretation models, as it was done, e.g. in the field of carcinogenesis testing.

This is possible because, if on the one hand human data are hard to obtain for chemically induced-developmental toxicity, on the other hand some known specific key developmental biological processes are already known and could be defined and individually tested *in vitro*. This means that effects of a toxicant could be described as the set of alterations of endpoints in several *in vitro* test systems, which would form test batteries.

In vitro test batteries can cover many of the key biological and molecular events of interest and allow to determine the hazard of a particular test compounds (Leist et al. 2014, Rovida et al. 2015a, Bal-Price et al. 2015). It is assumed by many experts that the combination of such different tests in a battery may eventually be able to predict human developmental toxicity (Basketter et al. 2012, Piersma et al. 2013, Schenk et al. 2010).

Some test batteries have been already developed in the last years. They may consist of molecular-based assays, like in the ToxCast program (Sipes et al. 2011; Padilla et al. 2012) or cell-based assays, as, for example, in the ReProTect (Schenk et al. 2010) or ChemScreen (van der Burg et al. 2014, van der Burg et al. 2015) projects (Table 2).

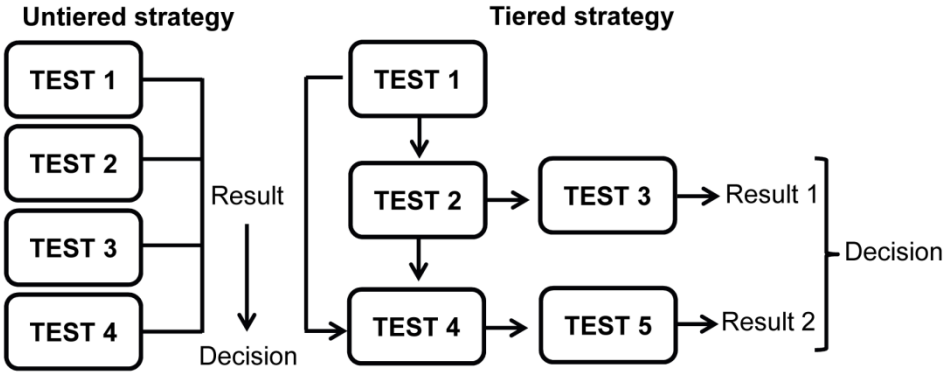


Figure 6 Test battery approaches. Test batteries can be built in a tiered or untiered way. In an untiered test battery, the test compounds can be simultaneously tested in the different systems and the collection of the data will bring to a final result and decision regarding the hazard potential of the single compound. In a tiered test battery, the systems are subordinate to each other; by this approach, separate results will be obtained, which must then be harmonized in final decision. Figure modified from Rovida et al. 2015a

Test batteries can be developed as tiered or untiered systems (Figure 6). In the first situation, all the test systems are challenged with the same list of compounds and the results coming from each of them are then taken together to evaluate the hazard of the single test compounds. In the second scenario, the test models are dependent from each other, so the results are a contribution of multi-step test- based decisions (Rovida et al. 2015a).

Test battery project	Short description	Cell-based/molecular-based	Tiered/ Untiered	Reference
Chemscreen	Combination of in silico and *in vitro* test results and pre-existing data for reproductive toxicity testing.	Cell- based	Tiered	van der Burg et al. 2015
Reprotect	Murine and human ESC or reporter gene-based tests to model interferences with fertility, implantation and prenatal development.	Cell- based	Tiered	Hareng et al. 2005
ToxCast/ Tox21	High-throughput screening. Testing of ~1000 chemicals in ~700 assays/ testing of more than 8'000 chemicals in 25 assays.	Molecular- based	Untiered	Judson et al. 2014
ESNATS	Development of hESC-based DNT testing platforms. Applicability domain range from drugs to environmentals.	Cell- based	Untiered	Rovida et al. 2014

Table 2 *In vitro* test batteries for toxicity testing

3.3.1 The ESNATS test battery

An example of cell-based untiered test battery which aimed to identify developmental (neuro-) toxicants has been the battery developed in the "embryonic stem cell-based novel alternative testing strategies" (ESNATS) framework (Krug et al. 2013c; Rovida et al. 2014, Zimmer et al. 2014).

The goal of the ESNATS test battery was to develop a novel toxicity testing platform based on (human) embryonic stem cells to accelerate drug development, to reduce related R&D costs, and to propose an alternative to animal tests in the fields of reproductive toxicity and neurotoxicity (Rovida et al. 2014). The final group of the involved test systems modeled different developmental stages: multilineage differentiation into ecto-, meso- and endoderm (Meganathan et al. 2012); neuroectodermal induction (Balmer et al. 2012); early neurogenesis, e.g. neural tube formation and transition from neural precursors to mature neurons (Stummann et al. 2009); and neural crest migration (Zimmer et al. 2012).

3.4 Transcriptomics approaches applied to *in vitro* developmental toxicity testing

The increasing advances in large-scale gene expression technologies, as microarray and RNA-seq analysis, and the consequent generation of information regarding the expression level of thousands of genes in a single assay, triggered the spreading of transcriptome approaches in various scientific fields, from evolutionary biology to human molecular pharmacology (Zhao et al. 2014).

Transcriptome studies have also been applied to the toxicology field, where the information regarding the gene expression level of several genes at the same time can be used to explore the molecular basis of pharmacological and toxicological responses (Fielden and Zacharewski 2001, Waring and Halbert 2002, Oberemm et al. 2005). The application of transcriptomics in toxicology depends on the assumptions that all toxicological relevant effects are accompanied by alterations in gene expression patterns (Farr and Dunn 1999). The ability of toxicogenomics to distinguish different compounds with different mode of action (MoA) has already been demonstrated in some particular fields, e.g. hepato-toxicology in cancer cell lines (Burczynski et al. 2000, Rempel et al. 2015)

More recently, transcriptomics was applied in more complicated toxicological fields, e.g. developmental toxicity. Several studies that address the prediction of potential developmental

toxicity of existing and novel compounds have been performed initially in mouse *in vitro* models (Robinson et al. 2010, Robinson et al. 2011, van Dartel and Piersma 2011).

Next, combination of toxicogenomics data and system biology, together with use of human stem cell based systems, has been explored (Balmer et al. 2014, Krug et al. 2013c, Krug et al. 2014; Rempel et al. 2015 ; Waldmann et al. 2014). This combined approach is expected to change the extrapolation and interpretation of human toxicological information in the next future (Hartung et al. 2012, Robinson et al. 2012a, Robinson and Piersma 2013, Waters and Fostel 2004, Wobus and Loser 2011).

4 Chemical description and clinical use of interferons class I

In the 3rd paper of this thesis, main focus will be given on the *in vitro* effects of the cytokine interferon beta on neural crest migration. This drug is included in the list of novel hits identified by the use of the MINC assay in the ESNATS test battery framework.

4.1 Chemical definition and biological description of interferons class I

Interferons (IFNs) are widely expressed polypeptides which belong to the cytokine family. Interferons have a crucial role in the antiviral cell response, and in modulating several functions of the immune systems, e.g. natural killer cells and macrophages activation and increase of antigen presentation. They are divided among three classes: interferon of class I, class II and class III. Interferons type I system is ubiquitary expressed and comprises 17 different ligands: IFNα subtypes, IFNβ, IFNκ, and IFNε. All type I IFN binds to the interferon alpha and beta receptor (IFNAR) subunits at the same location, producing structurally highly similar complexes (Piehler et al. 2012).

The hallmark of IFN type I signaling is the formation of a pSTAT1/pSTAT2 heterodimer, which in complex with IFN regulatory factor 9 forms the transcription factor IFN-stimulated gene factor 3 (ISGF3), which promotes transcription of the interferon-stimulated genes (ISG) (Figure 7) (Schreiber and Piehler 2015).

The main difference between the various subtypes lies in their binding affinity toward the receptor subunits (varying about 1000-fold

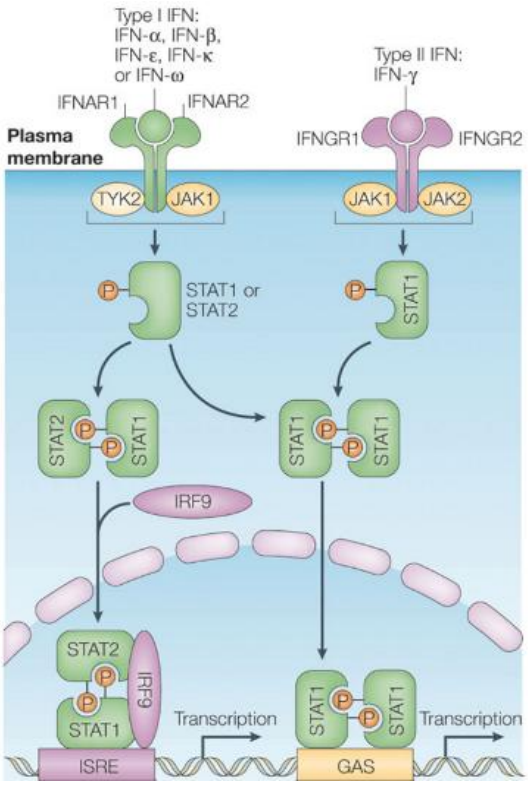


Figure 7 Activation of the JAK-STAT pathway by IFN class I and II. IFN class I binding to the receptor (IFNAR) subunits leads to activation of Jak and Tyr2. These kinases phosphorylate STAT1/2, leading to the formation of the ISG3 factor, which include the heterodimer STAT1-STAT2 and IRF9. ISG3 can then translocate in the nucleus, where it binds to the ISRE promoter, inducing the expression of several genes. IFN class II induces formation of homodimers which, in the nucleus, bind to the GAS promoter and induce the transcription of IFN class II-related genes. Figure modified from Begitt et al. 2014

among the ligands). The weakest binding to the receptor is made by IFNα1, while the tightest by IFNβ (200 nM versus 0.2 nM affinity) (Schreiber and Piehler 2015).

Interferons trigger two kinds of cellular responses:

- the robust response, which is observed similarly in all cell lines. JAK-STAT pathway is fastly activated by exposure to picomolar IFN concentrations. Genes induced by the robust response belong to the gene ontology (GO) classes of response to viruses, biotic stimulus, and MHC class I.
- the tunable response is highly cell-type specific and requires continuous receptor activation over days. It is susceptible to the cell surface receptor density. Genes induced by the tunable response belong to the GO classes of cytokines and chemokines activities, taxis and inflammatory and antiproliferative activities.

4.2 Clinical relevance & developmental toxicity evidences

The interferon family is used in clinics for different purposes. For example, IFNα is mostly used for treatment of chronic hepatitis and certain types of leukemia, while IFNγ is normally prescribed for some hereditary syndromes, as the chronic granulomatous disease. The immunomodulatory drug IFNβ has been proven effective in the treatment of relapsing-remitting multiple sclerosis (Dhib-Jalbut and Marks 2010). Multiple sclerosis (MS) is a neurodegenerative pathology which involves an auto-immune process that causes a loss of the myelin sheaths surrounding the axon of the nerve cells. Early dosage of IFNβ reduces the relapse rate and the development of brain lesions, retarding the progression of the pathology. The mode of action of the cytokine is still not fully understood. IFNβ seems to reduce the T-cell activation by MHC II downregulation and interfering with antigen processing and presenting. Furthermore, IFNβ induces an anti-inflammatory cytokine shift. Additionally, it was shown to prevent T-cell adhesion as well as their extravasation across the blood-brain-barrier (BBB) by the increasing of the soluble vascular cell adhesion molecule-1 and reducing the level of metalloproteasis (e.g. MMP9).

In vivo developmental toxicity of the cytokine has been evaluated in a study reporting the adverse effects of the exposure to IFN-β in pregnant monkeys, where a significant increased incidence of both abortions and stillbirths was observed (FDA 1999).

Furthermore, few epidemiology studies indicated that exposure to IFNβ in pregnancy is associated with lower mean birth weight, shorter mean birth length and preterm birth also in MS patients (Amato et al. 2010). For these reasons, women with MS are typically advised to discontinue the treatment before conceiving to minimize the risk of fetal harm. IFNβ is classified as risk class C drug by the FDA, indicating that "animal reproduction studies have shown an adverse effect on the fetus and there are no adequate and well controlled studies in humans, but potential benefits may warrant the use of the drug in pregnant women despite the potential risks" (Lu et al. 2012, Pozzilli and Pugliatti 2015).

Aims of the thesis

In the last years, the amount of chemicals needing to be evaluated for developmental hazard has been continuously increasing. hPSC-based *in vitro* test batteries have been proposed to increase the testing through-put.

In the next years, hPSC-based *in vitro* test batteries are assumed to be able to predict human developmental toxicity (Basketter et al. 2012, Piersma et al. 2013, Schenk et al. 2010) and to represent a valid tool for detection, prioritization, and characterization of the mechanisms of toxicity of several developmental toxicants.

The aims of this thesis were:

- to describe a novel hPSC-based test battery for DT testing and to use one the included functional system, the MINC assay, to screen a wide compound library, including environmental pollutants and prescription drugs.
- to anchor the phenotypical effects of the found hits to the induced expression changes, by transcriptome analysis.
- to combine the functional and transcriptional data in order to identify the pathway of toxicity involved in the triggered migration inhibitory effects by one the MINC-positive prescription drugs, interferon beta.

C. Results. Manuscript 1

Profiling of drugs and environmental chemicals for functional impairment of neural crest migration in a novel stem cell-based test battery

B Zimmer[1,2], G Pallocca[3*], N Dreser[3], S Foerster[3],T. Waldmann[3], J. Westerhout[4], S Julien[5], KH Krause[5], C van Thriel [7], J.G. Hengstler[7], A.Sachinidis[6], S Bosgra[4], M Leist[3]*

Affiliations:

[1]Center for Stem Cell Biology, Sloan-Kettering Institute for Cancer Research, New York City, USA; [2]Developmental Biology Program, Sloan–Kettering Institute, New York, USA; [3]Department of Biology, University of Konstanz, 78457 Konstanz, Germany; [4]Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek (TNO), 2628 VK Delft, The Netherlands; [5] Department of Pathology and Immunology, Geneva Medical Faculty, University of Geneva, 1211 Geneva, Switzerland; [6] Center of Physiology and Pathophysiology, Institute of Neurophysiology, University of Cologne, 50931 Cologne, Germany; [7]Leibniz Research Centre for Working Environment and Human Factors (IfADo), Technical University of Dortmund, 44139 Dortmund, Germany

* These authors contributed equally

Key words: test battery-based compound screening; developmental toxicity testing; hESC-based test system; neural crest migration assay

ABSTRACT

Developmental toxicity *in vitro* assays have hitherto been established as stand-alone systems, based on a limited number of toxicants. Within the ESNATS project we developed a test battery framework that allows inclusion of any developmental toxicity assay, and that explores the responses of such test systems to a wide range of drug-like compounds. We selected 28 compounds, including several biologics (e.g. erythropoietin), classical pharmaceuticals (e.g. roflumilast) and also six environmental toxicants. The chemical, toxicological and clinical data of this screen library were compiled. In order to determine a non-cytotoxic concentration range, cytotoxicity data were obtained for all compounds from HEK293 cells and from murine embryonic stem cells. Moreover, an estimate of relevant exposures was provided by literature data mining. To evaluate feasibility of the suggested test framework, we selected a well-characterized assay that evaluates 'migration inhibition of neural crest cells' (MINC). Screening at the highest non-cytotoxic concentration resulted in 11 hits (e.g. geldanamycin, abiraterone, gefitinib, chlorpromazine, cyproconazole, arsenite). These were confirmed in concentration-response studies. Subsequent pharmacokinetic modeling indicated that triadimefon exerted its effects at concentrations relevant to the *in vivo* situation, and also interferon-β and PBDE showed effects within the same order of magnitude of concentrations that may be reached in humans. In conclusion, the test battery framework can identify compounds that disturb processes relevant for human development and therefore may represent developmental toxicants. The open structure of the strategy allows rich information to be generated on both the underlying library, and on any contributing assay.

INTRODUCTION

Individual human embryonic stem cell-based developmental toxicity test systems have been established by several laboratories (Jagtap et al. 2011; Balmer et al. 2012; Stummann et al. 2009). A next step will be the combination of these and other assays to a comprehensive battery able to predict human developmental toxicities (Leist et al. 2012c; van Thriel et al. 2012). Cultures of differentiating pluripotent stem cells, such as human embryonic stem cells (hESC) or human induced pluripotent stem cells (Leist et al. 2008a; Thomson et al. 1998; Takahashi et al. 2007) offer unique possibilities of studying the very early steps of human development that lead to the formation of germ layers and primordial tissues. This opportunity was seized by the European Union research consortium for the use of 'embryonic stem cell-based novel alternative tests' (ESNATS) for the prediction of toxicity of drug candidates (www.esnats.eu). This project focused on the one hand on transcriptomics-based toxicity predictions (Krug et al. 2013c; Kuegler et al. 2010). On the other hand, several tests were established that allowed the assessment of neurochemical and cell biological cell functions (Stiegler et al. 2011; Zimmer et al. 2011b; Zimmer et al. 2012; Krug et al. 2013a) and of complex cell interactions (Preynat-Seauve et al. 2009; Kuegler et al. 2012). Moreover, concepts have been developed to compare relevant *in vitro* and *in vivo* concentrations (Bosgra et al. 2012; Krug et al. 2013a; Zimmer et al. 2011a), and to incorporate systems for metabolic activation of drugs (Godoy et al. 2013). It is assumed by many experts that the combination of such different tests in a battery may eventually be able to predict human developmental toxicity (Basketter et al. 2012; Piersma et al. 2013; Schenk et al. 2010). The hESC-based test systems of ESNATS cover different aspects of development. For instance, the UKK system (Meganathan et al. 2012) models early multi germ-layer differentiation, while the UKN1 system (Balmer et al. 2012) models specific neuroectodermal differentiation. The UKN2 system, also known as 'migration inhibition of neural crest' assay (MINC) (Zimmer et al. 2012) is a functional test probing the inhibition of neural crest cell migration by chemicals. During the initial establishment of the assays only a small number of positive and negative controls were tested. Therefore, the applicability domain of these assays and their response dynamics when faced with a broader variety of

compounds are unknown. Moreover, the information from only few compounds is not sufficient to evaluate how far the test systems are complementary, and where they may be redundant in the information they provide.

In DNT test library selection, new approaches are required (Leist et al. 2012a) to break a vicious circle between lack of sufficient tool compounds, and the inability to classically validate test systems without such compounds (Leist et al. 2012c,Leist et al. 2010). One of these would be a screening approach of hitherto little characterized compounds in multiple test systems. This would provide information on which biological processes may be targeted by the compounds. Together with mechanistic studies on the mode of action, this approach may allow to build a case for a hazard estimate independent of correlations with *in vivo* data (Kadereit et al. 2012). Moreover, characterization of the available assays would be promoted.

For the design of such a battery of different tests, experience from earlier approaches can be used as guidance. Test batteries may for instance be constructed in a tiered way to avoid redundant testing. If information on each compound from every test is desired, then non-tiered approaches are more useful. Examples from the field of reproductive toxicity testing are for instance the ReProTect feasibility study (Schenk et al. 2010) or the ChemScreen test battery (Piersma et al. 2013). Non-tiered testing is also performed in the ToxCast Program, in which hundreds of tests have been run in parallel, to use the data afterwards - in combination with pre-existing *in vivo* data - for predictions of drivers and mechanisms of reproductive toxicity (Kleinstreuer et al. 2011b; Padilla et al. 2012; Sipes et al. 2011).

Here, we defined a framework for a test battery and we provided an initial characterization of a core set of test compounds which can be expanded at later stages. To evaluate the feasibility of the suggested framework, and the usefulness of the set of compounds, we selected one well-characterized assay for a first screen. The MINC assay (Zimmer et al. 2012) was selected, as it is based on a functional endpoint, and it affords sufficient throughput to evaluate a compound battery of that size. The underlying biological rationale of the test is that disturbance of neural crest migration by toxicants leads to severe malformations in different species. Several factors (e.g. genetics and chemicals) have already been identified as causes for neural crest (NC) - related developmental defects (Di Renzo et

al. 2007; Fuller et al. 2002; Menegola et al. 2000). Identification of several hits in such a functional assay provides a good starting point for future characterization of the compounds by more phenotypic assays and for correlations of functional disturbances with e.g. transcriptome changes.

MATERIAL AND METHODS

Cell culture

The reporter hES cell line H9-Dll1 (GFP under Dll1 promoter) was provided by Mark Tomishima from the Memorial Sloan Kettering Cancer Centre (MSKCC, NY, USA). Import of the cells and all experiments were carried out according to German legislation under the license number 1710-79-1-4-27 of the Robert-Koch Institute. H9-Dll1 cells were maintained on Mouse Embryonic Fibroblasts (MEFs) in DMEM/F12 (Gibco) medium containing 20% of serum replacement, HEPES (1M, Gibco), L-glutamine (Glutamax, Gibco), non-essential amino acids (MEM NEAA, Gibco), beta-mercaptoethanol (Gibco) and basic fibroblast growth factor (10 ng/ml, Invitrogen). The murine ES cell line CGR8 was obtained from the European Collection of Cell Culture (ECACC, UK). CGR8 cells were maintained on 0.1% gelatin coated dishes in BHK21 medium, supplemented with 10% fetal calf serum, L-glutamine, non-essential amino acids, penicillin/streptomycin and leukemia inhibitory factor (LIF) (Kern et al. 2013). HEK 293 (CRL-1573, ATCC) cell line was maintained in DMEM supplemented with 10% fetal calf serum at 37°C in a humidified atmosphere containing 5% CO_2.

Neural differentiation protocols

The mESC cell line (CGR8) was differentiated towards a neural stem cell phenotype using the protocol described by Barberi et al. 2003). Briefly, CGR8 were seeded on irradiated MS5 cells and cultivated in DMEM medium containing 15% Knock-out Serum Replacement, non-essential amino acids, beta-mercaptoethanol and penicillin/streptomycin. After 4 days, cells were replated on polyornithine (15 µg/ml) coated dishes in N2 medium containing DMEM, N2 supplement, penicillin/streptomycin and 10 ng/ml of basic human fibroblast growth factor (bFGF) (Invitrogen). Differentiation of hESC into neural crest cells was initiated on Mitomycin C treated murine bone-marrow derived stromal MS5 cell line and continued as described in Zimmer et al. 2012).

Evaluation of a non-cytotoxic range by resazurin assay and bench-mark concentration (BMC) calculation

The effects of the toxic compounds on cell viability of two cell lines were evaluated by using the resazurin assay. The assay is based on the capability of viable and healthy cells to reduce resazurin to resorufin, which can be measured by a colorimetric or fluorimetric shift as described earlier (Zimmer et al. 2012). HEK293 cells and mESC-derived neural stem cells (mESCn) were exposed for 48 hours to the different substances. mESCn were exposed to test compounds after 6 days of differentiation. After this period, the cells were incubated at 37°C and 5% CO_2 with 10 µg/ml resazurin for 30 min (HEK293) or up to 5 h (neural stem cells). The background fluorescence of resazurin itself was determined by including a resazurin only control. Resazurin reduction was analyzed in cell culture medium fluorimetrically (λ_{ex} = 530 nm, λ_{em} = 590 nm). These data were used to model a concentration-response curve and to calculate the concentration corresponding to a 10% reduction of viability (BMC10). In addition, the BMC15 and the lower limit of its 95% confidence interval (BMCL15) were determined. This latter value was used as estimate for the upper boundary of the non-cytotoxic concentration range.

Cell migration analysis

Cell migration analysis was carried out using a scratch assay design as described in Lee et al. 2009) and Zimmer et al. 2012) with minor modifications. hESC-derived NCCs were grown to a confluent monolayer using 48-well plates (Corning). Right before starting the assay, each well was scratched using a 20 µl pipette tip in order to create a cell-free gap. The medium was removed and replaced by fresh medium containing the test chemicals. The width of the cell-free gap was determined right after scratching in a control plate in order to define the dimension of the region of interest (ROI) for the analysis. The cells were exposed to the toxicants for 48 hours; after this period, the general cytotoxicity was assessed by the resazurin reduction assay. Migration of NCC was evaluated by florescence microscopy analysis. In order to easily count the number of cells, incubation with fresh medium containing the DNA dye H-33342 (1 µg/ml) was performed for 30 min. After the incubation period, random images along the scratch were taken at 4 × magnification. The number of

cells with H-33342-positive nuclei within the region of interest (ROI) was automatically calculated by the use of a KNIME flowchart.

Chemical exposure during migration

hESC-derived neural crest cells were exposed to chemicals in N2 medium containing EGF (20 ng/ml) and FGF2 (20 ng/ml). For a detailed list of chemicals and their tested concentration range used in this study see Fig. 3 and Fig. S1, S2.

In vitro - in vivo comparison of toxicity data by PBPK modeling

In order to evaluate the clinical relevance of the *in vitro* concentrations found to impair the migration of the hESC-derived NCCs in this study, a three step (physiology-based) pharmacokinetic (PBPK) modeling strategy has been used, as already described in Krug et al. 2013c) and Piersma et al. 2013). Briefly, the following steps were taken: (a) choice of an appropriate absorption, distribution, metabolism, excretion (ADME) model; (b) use of this model to simulate plasma and/or target tissue concentrations in time corresponding to the exposure (dose, route of administration, interval) at which relevant toxic effects were observed in already published *in vivo* studies; (c) calculation of the nominal concentration *in vitro* that has the same unbound concentration as the toxic concentration *in vivo* (when possible).

In vitro - in vivo comparison of toxicity data for interferon β. A PBPK model for the analysis of interferon β (IFN-β) kinetics in monkeys, described by Mager et al. (2003), was implemented in the acslX software (version 3.0.2.1, Aegis Technologies) (step a). The original model was built on the basis of data from 18 cynomolgus monkeys that were exposed i.v. to single doses of 1, 3, 10 MIU/kg and then to a s.c dose of 0.3 ml/kg of IFN-β. *In vivo* developmental toxicity concentrations of the drug have been extrapolated from a study reporting the effects of the exposure of IFN-β in pregnant cynomolgus monkeys (FDA 1999) (step b).

In vitro - in vivo comparison of toxicity data for triadimefon. A PBPK model for the pesticide triadimefon and its metabolite triadimenol in rats published by Crowell et al. 2011) was reconstructed in acslX and used to predict the target tissue concentration related to the exposure scenarios leading to toxic effects on male fertility and CNS toxicity (step a).

Developmental toxicity-inducing concentrations were extrapolated from the *in vivo* study by Goetz et al. 2007), in which pregnant rats have been exposed to the pesticide. Two exposure scenarios were simulated: Dietary exposure assuming a constant intake of the entire drug dose within the first 12 h of 24 h periods; oral gavage, modeled as a bolus dose into the liver compartment (step b). The nominal *in vitro* concentrations equivalent to the concentrations predicted *in vivo* were determined correcting for the differences in albumin concentration and lipid fraction between plasma or cerebrospinal fluid and test medium, using the follow equations:

$$EC_x = EC_p \times \left\{ (1 - f_{b,p}) \times \frac{1 + K_{ow} \times VF_{L,x}}{1 + K_{ow} \times VF_{L,p}} + f_{b,p} \times \frac{P_x}{P_p} \right\} \qquad EC_{u,x} = \frac{EC_x - \frac{P_x}{P_p} \times f_{b,p} \times EC_p}{1 + K_{ow} \times VF_{L,x}}$$

Where EC represents the effective concentration; $f_{b,p}$ the plasma fraction unbound; K_{ow} the octanol:water partition coefficient; VFL the lipid fraction; P the albumin concentration; suffix u means unbound; suffix p the plasma; and suffix x the other medium (*in vitro* or CSF) (step c).

The parameters of free (unbound) fraction, octanol:water partition and blood:plasma concentration ratio were taken from the published study by (US EPA 2006) and 0.11, 912 and 0.84, respectively. Data for rat CSF (estimated as 0.5% of plasma) were taken from (Habgood et al. 1992; Koch et al. 2001) and data for MINC culture medium were calculated based on information provided by the supplier.

In vitro - in vivo comparison of toxicity data for PBDE-99. A PBPK model was constructed based on data of tissue distribution, metabolism and excretion of PBDE-99 as described by Hakk et al. 2002) and Chen et al. 2006) (step a). The PBPK model structure to describe the kinetics of PBDE-99 is shown in Fig. S4a. The model contains a gastrointestinal lumen compartment (GI), two rapid equilibrium compartments (T1 and T2), a blood compartment (B), a lipophilic tissues compartment (F) representing adipose tissue and skin, and compartments for urinary and fecal excretion (Ur and Fe). The exchange between blood and tissue compartments is described by first order rate constants k_{b1}, k_{1b}, k_{b2}, k_{2b}, k_{bf} and k_{fb} with unit h^{-1}. The compound is absorbed into T1 – containing intestinal tissues and liver, but not further specified – by a rate k_{ab}, and eliminated back into GI with rate k_{el}. Excretion occurs

from the blood compartment with rate k_{ur} and from the GI compartment with rate k_{fe}. The model was described as a set of differential equations in acslX. Concentrations were calculated from amounts by dividing compartment volumes: 0.21, 0.56 and 0.06 ml/g BW (body weight) for lipophilic tissues, rapid equilibrium tissues and blood, respectively, as reported by Brown et al. 1997). The estimated parameter values are listed in the supplemental material (Fig. S3b). The model performance was demonstrated by comparison of model predictions to *in vivo* PK data reported by Chen et al. 2006) for a single oral dose of 1 µmol/kg (Fig. S3 c,d) and an intravenous bolus dose of 1 µmol/kg (Fig. S3e, f). *In vivo* developmental toxicity concentrations were extrapolated by the data from Kuriyama et al. 2005) and Viberg et al. 2005), where neuro-developmental effects are observed in rats exposed to PBDE-99, during the gestational or the early infancy period (step b).

Statistics and data mining

For the resazurin assay, five technical replicates for HEK293 cells and four biological replicates for mESCn cells have been analyzed for each compound and concentration. For the migration assay, the number of migrated cells was automatically counted in ≥4 different images per experiment by a KNIME flowchart-based software. All data displayed are means from three independent biological experiments. Each biological experiment consisted of at least four technical replicates. Statistical differences were tested with GraphPad Prism 5.0 (Graphpad Software, La Jolla, USA) by applying ANOVA using Bonferroni's post-hoc test. Independent biological experiments (not technical replicates) were the basic unit used for statistical testing.

RESULTS

Considerations and design principles of the test battery

Several murine and human stem cell-based developmental toxicity test systems have been developed by ESNATS project partners (Balmer et al. 2012; Krug et al. 2013c; Stiegler et al. 2011; Zimmer et al. 2012; Kern et al. 2013; Jagtap et al. 2011) and others (Fritsche et al. 2011; Seiler and Spielmann 2011; Hogberg et al. 2010; Pallocca et al. 2013; Piersma et al. 2013; Suzuki et al. 2011). All these assays have been evaluated individually with positive and negative control compounds regarding their biological relevance for fundamental processes of mammalian development. However, little is known how such test systems can

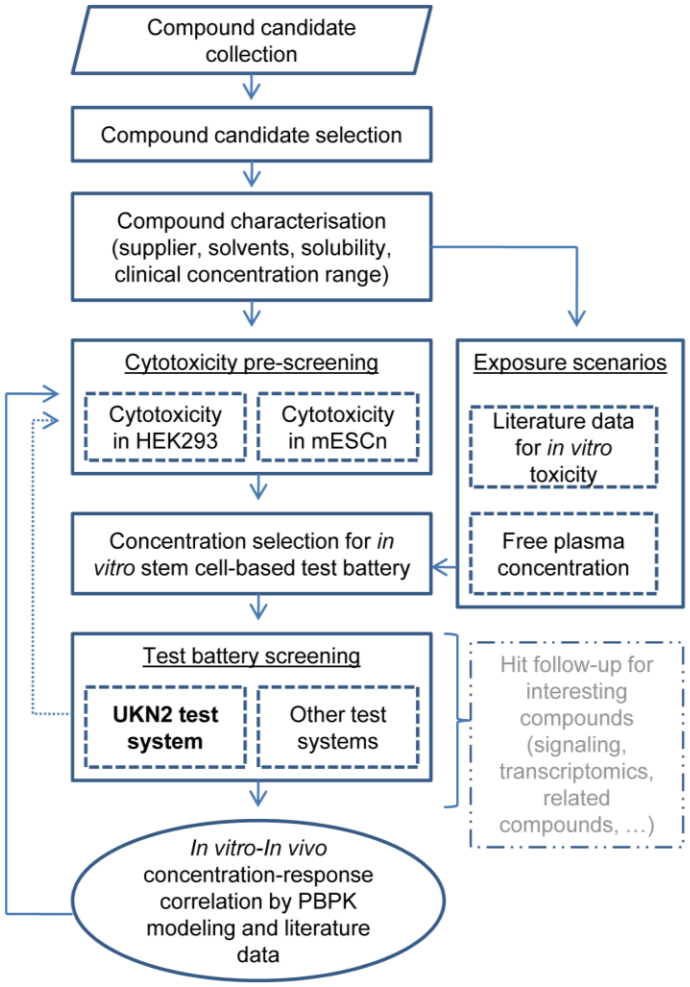


Fig.1 Overview of the ESNATS test battery. Candidate compounds were compiled based on criteria described in Fig. 2, and a final set of 28 drugs and environmental pollutants was selected. Their characteristics, including compound source, solubility, clinical concentration ranges and toxicological background information were compiled. Two approaches were chosen to determine a non-cytotoxic range for the further screening: cytotoxicity pre-screening on a transformed cell line derived from human embryonic kidney (HEK-293) and on murine embryonic stem cell-derived differentiating neural cells (mESCn). In addition, realistic exposure concentrations were estimated from literature data mining (for *in vitro* toxicity information) and pharmacokinetic (PK) prediction (free plasma concentrations). The test battery comprised initially the UKN2 test system and three other hESC-based tests, but it was designed openly for any test addition. Screening proceeded in two steps: first the highest non-toxic or relevant concentration for a given test was determined; then compounds were tested in hESC models at this concentration. Based on the results of the screening, a shortlist of compounds was selected for further characterization by physiologically-based pharmacokinetics (PBPK) modeling and for hit follow-up.

be combined to yield information on drug toxicity.

Important features of the test battery framework are the characterization of the compounds concerning general cytotoxicity, relevant *in vivo* concentrations and other necessary background data. Accessory modules for hit follow-up and *in vitro-in vivo* extrapolation should provide rich information on many of the compounds in the future. In fact, one of the initial purposes of the test battery was the pre-filtering of hits for further toxicogenomics follow-up, for instance by transcriptome profiling (Fig. 1, 2a). This will be performed, once a sufficient number of hits will be characterized in different assays.

A	Objectives of the test battery
1.	Identification and probing of test battery requirements: test redundancy, complementarity, outcome concordance.
2.	Exploring new compounds that may be used later for test system benchmarking
3.	Providing novel information about drug's developmental toxicity hazards
4.	Follow-up of hits by transcriptomics analysis and investigation of signaling pathways: providing information about pathways-of-toxicity and signatures of toxic effects triggered by the exposure to known and unknown toxicants in human test systems
5.	Screening drugs with a wide variety of expected modes-of-action (widening of scope of test systems; exploring applicability domain)
6.	Establishing a strategy for handling of data from multiple test systems

B	Compound selection criteria
1.	Belonging to the group of modern drugs (developmental toxicity hazard little characterized) or to the group of environmental pollutants (known or suspected developmental toxicants)
2.	Inclusion of the class of biologics and peptide-related small molecules
3.	Wide range of mechanisms and targets
4.	Preferentially compounds with intracellular targets
5.	Preferentially compounds showing species-specific interactions with target proteins (e.g. peptides or peptidomimetics compounds)
6.	Sufficient members of expected negatives to estimate test specificity

Fig.2 Test battery design criteria

Selection of test battery compounds

The 28 compounds were compiled according to the selection criteria outlined in Fig. 2b. The test library reflects a compromise between the different criteria. Our choice marks a deliberate and intentional departure from the use of known toxicants and endpoint-specific controls (reviewed in Kadereit et al. 2012; Crofton et al. 2011; Leist et al. 2010), and it puts emphasis on the exploration of unknown drugs. Besides the drugs, a small selection (six substances) of environmental pollutants (e.g. PCB, PBDE, arsenic) was included as likely

positive controls for many test systems. The group of drugs also included biologics (e.g. interferon-β, oxytocine) and peptide-related small molecules (e.g. sitagliptine, galnon). Some of the biologics were included as they are known to cross the blood brain barrier *in vivo* (e.g. G-CSF, erythropoietin). Finally, three compounds (sulfadiazine, chlorpromazine, amiodarone) were chosen because another drug screen (Kern et al. 2013) suggested a potential for developmental neurotoxicity. For all compounds of the test library, essential chemical and pharmacological information was compiled (Fig. S1, S2; Fig. 3). For environmental compounds with known neurotoxicity (developmental neurotoxicity) we referred to several pertinent *in vivo* and *in vitro* studies.

Compound	Solvent	Clinical conc.	Free conc.	Highest non-cytotoxic conc. in HEK293	Highest non-cytotoxic conc. in mESCn	Concentration for hit finding in NCC	LOAEC in NCC
Teriflunomide	DMSO	10.8 µM	54.0 nM	120.0 µM	3.5 µM	1 µM	
Nintedanib/ BIBF 1120	DMSO	74.2 nM	70.8 nM	42.8 nM	2.0 nM	100 µM	
Telaprevir	DMSO	5.1 µM	1.6 µM	21.1 µM	10.0 µM	10 µM	
Sitagliptin	DMSO	1.9 µM	1.2 µM	100.0 µM	> 50.0 µM	100 µM	
Abiraterone	DMSO	647.3 nM	7.8 nM	2.0 µM	970.0 nM	2 µM	102 nM
Roflumilast	DMSO	18.0 nM	0.2 nM	10.0 µM	> 50.0 µM	10 µM	
Exenatide	DMSO	44.7 nM	0.3 nM	50.0 nM	> 20.0 nM	50 nM	
Gefitinib/ Iressa	DMSO	1.2 µM	34.0 nM	24.4 µM	500.0 nM	500 nM	
Rivaroxaban	DMSO	689.6 nM	41.4 nM	10.0 µM	> 10.0 µM	1 µM	
Aliskiren	DMSO	326.5 nM	166.0 nM	100.0 µM	5.0 µM	50 µM	
Galnon	DMSO	n.a.	n.a.	10.5 µM	13.0 µM	1 µM	
Neuregulin	0.1% BSA	6.3 µM	n.a.	n.t.	> 3.0 nM	3 nM	
Erythropoietin	0.1% BSA	0.3 nM	n.a.	n.t.	> 0.3 nM	3 nM	
Geldanamycin	DMSO	0.8 µM	0.5. µM	44.0 nM	8.0 nM	16 nM	5 nM
G-CSF	0.1% BSA	1.3 pM	n.a.	n.t.	> 2.7 nM	10 nM	
IFNβ	0.1% BSA	0.4 pM -7.5 nM	0.4 pM -7.5 nM	n.t.	> 200.0 pM	500 pM	36 pM
Sildenafil	DMSO	221.0 nM	11.0 nM	50.0 µM	> 50.0 µM	50 µM	
Imatinib	DMSO	3.0 µM	224.0 nM	10.0 µM	7.0 µM	5 µM	2 µM
Oxytocin	0.1% BSA	2.0 nM	2.0 nM	> 100.0 nM	> 100.0 pM	100 nM	
Sulfadiazine	DMSO	320.0 µM	160.0 µM	100.0 µM	> 50.0 µM	50 µM	
Amiodarone	DMSO	2.3 µM	0.5 nM	10.0 µM	460.0 nM	10 µM	
Chlorpromazine	DMSO	393.1 nM	58.9 nM	50.0 µM	340.0 nM	1 µM	
Methoxyacetic acid	0.1% $NaHCO_3$	See background data in supplemental material (Fig. S4)		126.0 µM	> 50.0 µM	50 µM	
Cyproconazole	DMSO			100.0 µM	> 50.0 µM	5 µM	2 µM
Triadimefon	DMSO			50.0 µM	> 50.0 µM	50 µM	46 µM
PCB-153	DMSO			25.0 µM	10 µM	5 µM	787 nM
PBDE-99	DMSO			20.0 µM	5 µM	20 µM	4 µM
Arsenic trioxide	0.1% $NaHCO_3$	0.7µM	0.6 µM	25.0 nM	500 pM	1 µM	378 nM

◄ **Fig.3 Toxicological background data for all compounds screened in the test battery.** The solvent used for each compound is indicated. Data on clinical concentrations (maximal plasma concentration) and the free plasma concentrations are explained in greater detail in supplementary Fig. S4. The data obtained from the cytotoxicity pre-screening in HEK293 and mESCn cells are reported as the real data points closest to the mathematically-modelled highest non-cytotoxic concentrations (BMCL15). The highest non cytotoxic concentration determined in the UKN2 test system is also indicated. The last column indicates the lowest observed adverse effect concentration (LOAEC), the concentration triggering a 20% inhibition of migration of the neural crest cells. (n.a. : not available data ; n.t. : not tested).

Pre-screening of test battery compounds for general cytotoxicity

Most developmental neurotoxicity assays give reliable and specific results only when compounds are used at concentrations that do not trigger general cytotoxicity/cell death. This range has to be determined for each compound and each test system. However, most available assays allow only a relatively low throughput of samples. Therefore, it would be more efficient and economical to get some rough initial information on non-cytotoxic concentration ranges before the onset of testing. For this purpose, we used two different assays. The first was based on human HEK293 cells. Resazurin reduction was applied as viability endpoint after exposure to the test battery compounds for 48 h. The second assay made use of murine embryonic stem cells (mESCn) differentiated towards neural lineage. Also here, viability of the cells was determined by resazurin reduction after a 48 h exposure period during the initial stages of differentiation (starting on day 6). The combination of these assays was meant to cover many modes-of-action of cytotoxicants across species and cell biological functions.

To determine the non-cytotoxic concentration range, compounds were tested at multiple concentrations. We then used a mathematical procedure to determine a benchmark concentration as upper limit of the non-cytotoxic range. The procedure is displayed in detail for the example compound geldanamycin (Fig. 4). For practical purposes, we used the real data point closest to this calculated theoretical threshold as toxicity threshold (Fig. 3). In most cases (16 compounds) in which comparative data were available, minimum cytotoxic concentrations of the compounds for the two cell types were similar (< 5-fold difference). For all remaining substances (8 compounds), the embryonic stem cell system showed a

higher sensitivity than the HEK293 cell line (Fig. 3). Cytokines were only tested at pharmacological concentrations to be expected in body fluids, and they all proved to be non-cytotoxic at these test concentrations (Fig. 3). In summary, the viability data give a rough indication on good starting points, but more precise data are needed for each new experimental model and for each experimental variation within a given test system (see below).

Determination of toxicologically-relevant concentration ranges

Besides non-cytotoxicity, further criteria are important to determine reasonable test concentrations and to interpret the data. To allow decisions on concentration ranges (e.g. for cytokines) and interpretation of screen results, we compiled the clinical blood and tissue

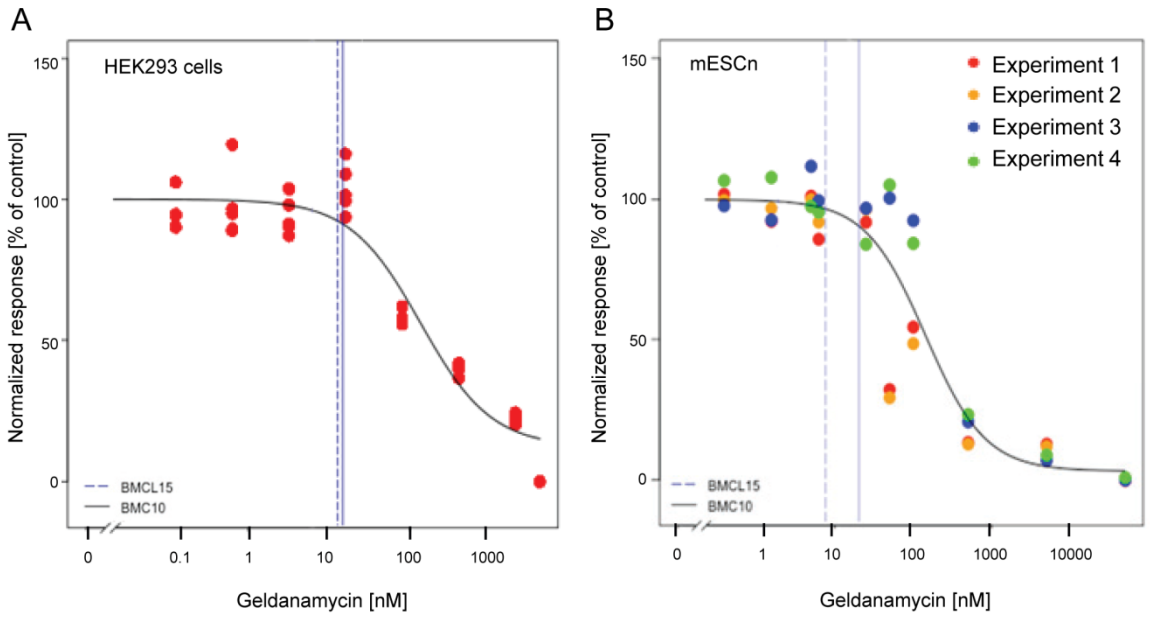


Fig.4 Determination of the highest non-cytotoxic concentration. Compounds were tested at multiple concentrations, and cell viability was determined by the resazurin assay (every replicate is represented by a single circle, different colour code for independent experiments). These data were used to model a concentration-response curve and the concentration corresponding to a 10% reduction of viability (BMC10) was calculated (solid line). In addition, the BMC15 was determined and the lower limit of its 95% confidence interval (BMCL15, dashed line) was determined. This latter value was used as estimate for the upper boundary of the non-cytotoxic concentration range. An example of BMC determination for the compound geldanamycin is shown, using data obtained from the cytotoxicity pre-screening in a HEK-293 cells and b mESC-derived differentiating neural cells (mESCn).

concentrations (CC) for most compounds. These, together with data on plasma protein binding, were used to calculate the free plasma concentrations (FCC) in patients. The latter data are important to relate *in vivo* data to *in vitro* concentrations. While an overview of the data is given in Fig. 3, more detailed information has been compiled in a supplementary table (Fig. S4).

For extrapolation of *in vivo* data to *in vitro* concentrations, often the issue arises, whether C_{max} (the peak concentration reached in clinics) or AUC (the average concentration found in a patient) should be used as anchor point for calculations. In the present study, both plasma peak (C_{max}) and average data were considered. This allows case-by-case decisions, depending on the assumed mechanism of toxicity. For instance, it may be plausible that cytokine receptors need to be triggered only for a short time (by C_{max}) to generate intracellular signals that potentially affect differentiation. In contrast to this situation, continuous cytokine signalling (in the AUC range) may be required to alter the cellular cytoskeleton and thereby to affect the cellular migration capacity. Vice versa, an environmental toxicant that inhibits a cellular pathway may need to do this continuously to exert developmental effects, i.e. an average toxic concentration (AUC) needs to be maintained to result in an average long-term inhibition. A short pulse, even at C_{max} may not be toxic. However, the situation would be different, if the compound (such as arsenite) binds irreversibly to cellular structures above a certain threshold concentration. Under such circumstances, the C_{max} values would become relevant.

Inhibition of neural crest cell migration by test battery library

As first assay to be run within the framework of the test battery, we chose the UKN2 test system that evaluates interference of potential toxicants with neural crest (NC) cell migration (Zimmer et al. 2012). For the primary screen, the library was run at a single concentration, selected according to the cytotoxicity pre-screening. At this concentration the viability of NCC was determined. If the concentration was non-toxic for NCC, then inhibition of migration was determined in the MINC assay. If the compounds were initially toxic, the UKN2 toxicity threshold was determined and then the MINC assay was performed at this concentration. When comparing cytotoxicity tests across different cell lines and models, we observed that cytotoxicity depends not only on the cell type and the medium used. Often minor experimental details affected the outcome (e.g. the plate type used). Therefore, the most relevant cell viability data were obtained when this endpoint was measured in each test run, and directly within the sample used for measuring the specific endpoint (e.g. migration). For instance, about two thirds of the test compounds showed similar cytotoxicity thresholds in the pre-testing models (HEK293; mESCn) and the MINC. For two compounds, cytotoxicity of the NCC resembled rather the mESC than HEK293; for three other compounds, NCC

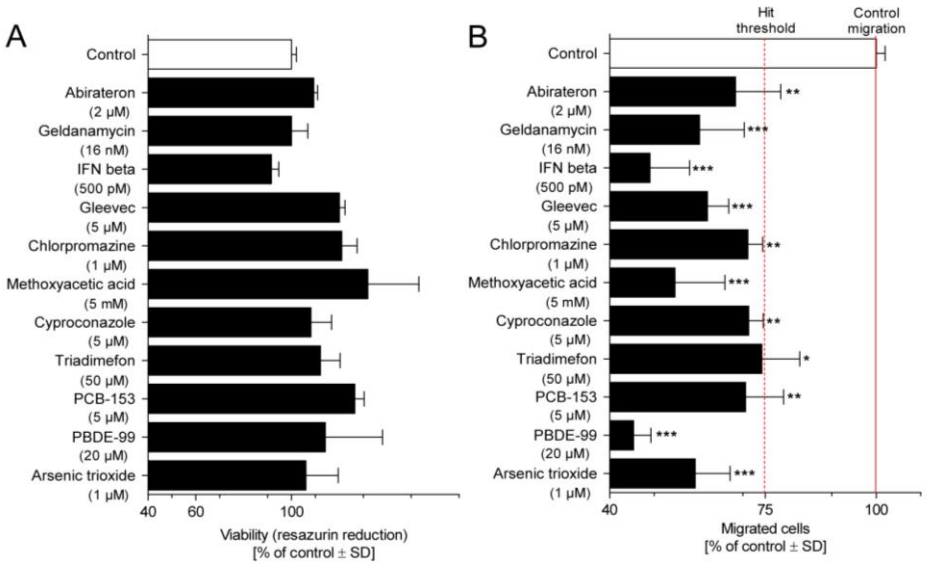


Fig.5 Overview of hits identified in the UKN2 test system. Neural crest cells were exposed to the highest non-cytotoxic concentration of test compounds for 48 h. The inhibition of cell migration and cytotoxicity induced by different compounds were measured by **a** the MINC and **b** resazurin assays. Substances leading to ≥ 25% reduction of the NCC migration activity in presence of ≤ 10% cytotoxicity (compared to the untreated controls) were considered UKN2-positive hits. Data are means ± SD of 3 independent experiments normalized to untreated controls. *: $p < 0.05$, **: $p<0.01$, ***: $p < 0.001$

sensitivity was closer to that of HEK293. After identification of adequate concentrations viability and MINC data were obtained for each compound, and experiments were repeated at this test concentration with two further cell preparations (Fig. 5, 6). Hit compounds (Fig. 5) reduced the migration for at least 25%, while they reduced cell viability in the same assay for less than 10%. Non-hits did either not affect migration strongly, or they influenced migration only above the cytotoxicity threshold (Fig. 6). Altogether 11 hits were identified, 5 of which came from the group of drugs. Within the latter group, the hit rate was 23%; amongst the environmental compounds, the hit rate was 100%.

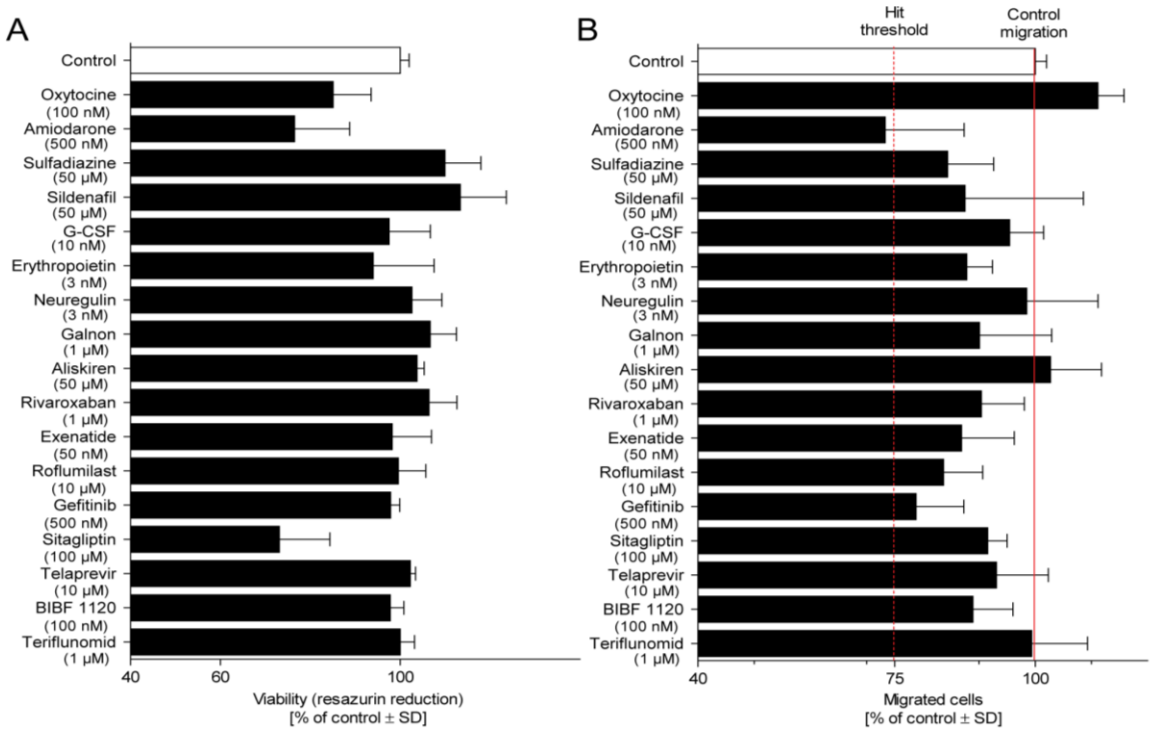


Fig.6 Overview of negative compounds in the UKN2 test system. The results of a 48 h exposure of neural crest cells to the highest non-cytotoxic concentration of test compounds are shown. The inhibition of cell migration and cytotoxicity induced by different compounds were measured by **a** the MINC and **b** resazurin assays. Substances showing reduction of cell viability (>10%) and/or reduction of migration <25% (compared to the control) were considered UKN2-negative compounds. Data are means ± SD of 3 independent experiments normalized to untreated controls. *: $p < 0.05$, **: $p<0.01$, ***: $p < 0.001$

Determination of the minimal concentration triggering developmental toxicity

To follow up on the hits from the initial screen by the MINC assay, a broader range of concentrations of these compounds was tested in the same assay. We used these data to

determine the lowest concentration inhibiting the specific test endpoint - neural crest cell migration. A non-linear regression was fitted to the concentration-effect curve (Fig. 7, 8). For this type of follow-up assay, we assumed that a reduction of migration by 20% was of biological and toxicological significance. The concentration triggering this extent of inhibition was determined and defined here as the 'lowest observed adverse effect concentration' (LOAEC). This value represents the lowest concentration at which inhibition of migration would become toxicologically relevant in our test system (Fig. 3).

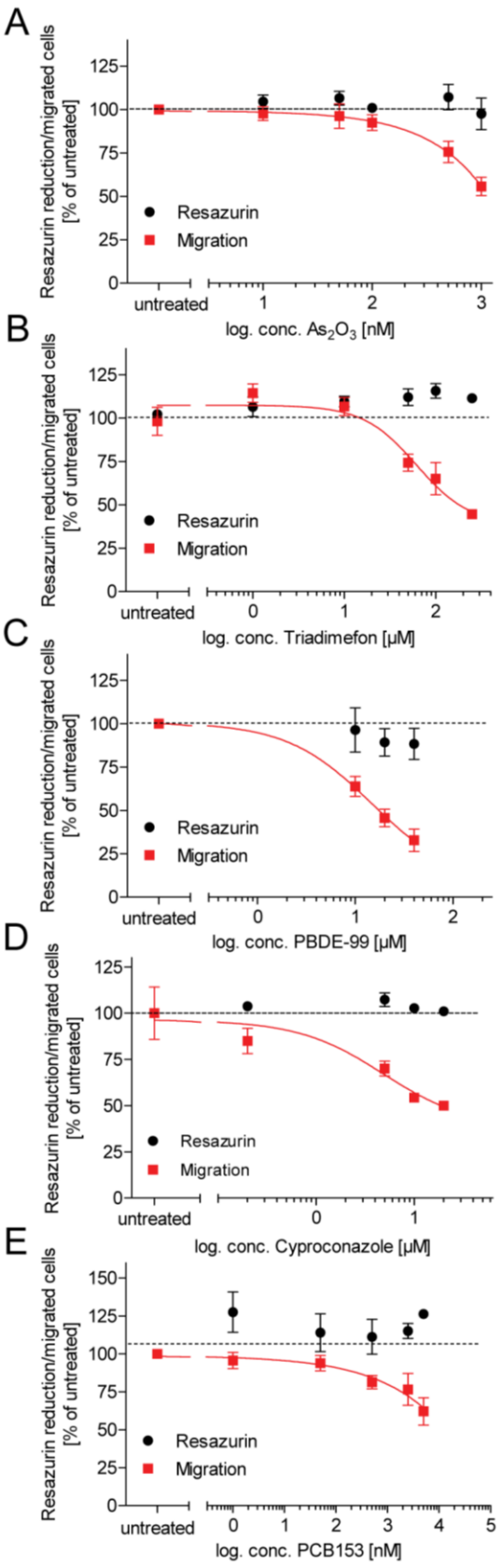


Fig.7 Concentration-response curves for UKN2-positive environmental pollutants. Neural crest cells were exposed to different concentrations of each test compound for a period of 48 h. The inhibition of cell migration (red squares, solid line) and cytotoxicity (black dots) induced by different compounds were measured by the MINC and resazurin assays. **a** As_2O_3, **b** triadimefon, **c** PBDE-99, **d** cyproconazole and **e** PCB153 showed a concentration-dependent effect on the NCC migration, in a not-cytotoxic concentration range. Data are means ± SD of 3 independent experiments normalized to untreated controls. *: $p < 0.05$, **: $p<0.01$, ***: $p < 0.001$

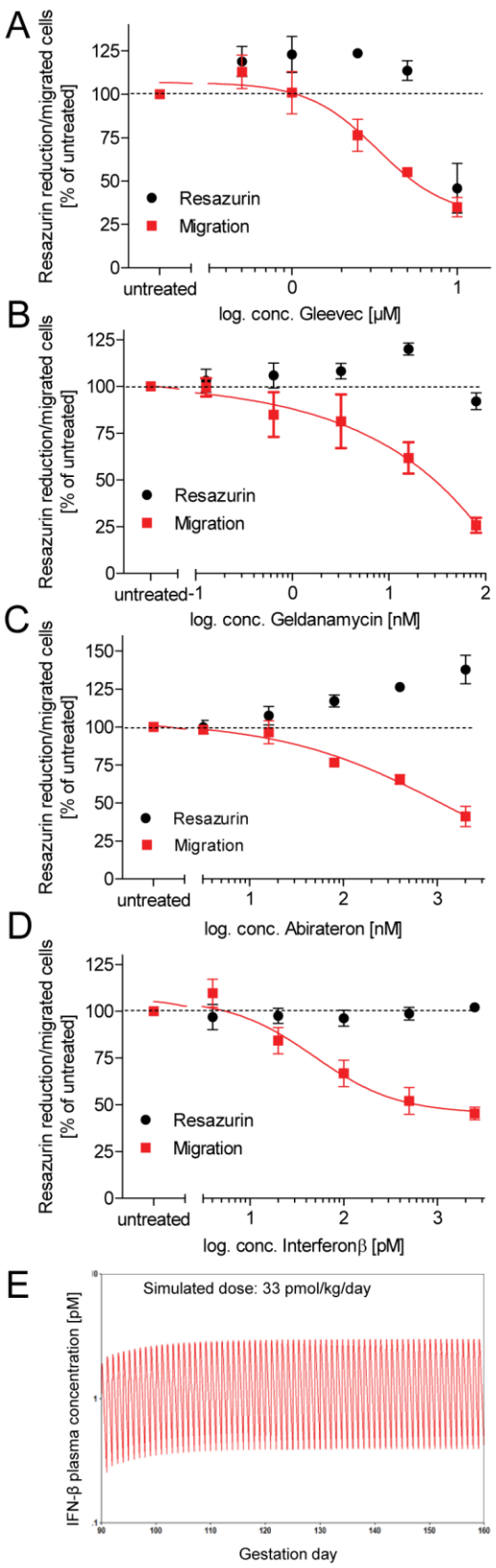


Fig.8 Concentration-response curves for drugs identified as hits in the UKN2 test system. Neural crest cells were exposed to various concentrations of test compounds for a period of 48 h. The inhibition of cell migration (red squares, solid line) and cytotoxicity (black dots) induced by different compounds were measured by the MINC and resazurin assays. **a** Gleevec, **b** geldanamycin, **c** abirateron and **d** interferon β (IFN-β) showed a concentration-dependent effect on the NCC migration, in a non-cytotoxic concentration range. Data are means ± SD of 3 independent experiments normalized to untreated controls. (*: $p < 0.05$, **: $p<0.01$, ***: $p< 0.001$). **e** Simulation of IFN-β plasma concentration induced by a sub-cutaneous dose of 33 pmol/kg/day from GD90 (gestation day 90) to term (GD160) in cynomolgus monkeys, using a PBPK model published by Mager et al. (2003) and reconstructed in acslX software.

Examples of PBPK modeling for hit compounds to relate in vitro correlations to realistic exposure scenarios

We selected three compounds for more detailed comparisons of LOAEC values and toxicologically relevant *in vivo* concentrations. First, recombinant human interferon-β (IFN-β) was examined. A PBPK model for the analysis of IFN-β kinetics in monkeys has been described by Mager et al. 2003). In this study, IFN-β plasma concentrations were measured during a period of 48 hours in order to determine the pharmacokinetic and pharmacodynamic parameters necessary for modeling. The model of Mager et al. (2003) postulates a decrease in receptor density upon repeated exposure to IFN-β, and an elimination of IFN-β (internalized into cells) depending on the receptor density. We implemented this PK model in the acslX software, and we simulated the kinetics of IFN-β after exposure to a dose known to trigger developmental toxicity in animals: this simulated toxic dose was chosen based on an unpublished report summarized by the US FDA (see Materials and Methods). In this report, developmental toxicity in pregnant monkeys exposed to IFN-β was shown for a dose of 740 ng/kg/day (= 33 pmol/kg/day, assuming a molecular weight of 22.5 kDa) given from gestational day 90 to term (GD160). Our simulations showed for a subcutaneous dose of 33 pmol IFN-β/kg/day that receptor density decreased by about one third. As a consequence, daily average plasma concentrations of IFN-β increased from 0.9 to 1.5 pM. The simulated plasma peak concentrations were in the range of 2-3 pM (Fig. 8e). This value was about 10-fold lower than our *in vitro* NOAEL (no observed adverse effects level), i.e. still within the same order of magnitude.

Triadimefon was the second compound selected for PK modeling because it represents a specific NCC toxicant (Zimmer et al. 2012; Menegola et al. 2005). Further knowledge on the mechanisms of toxicity would be of high interest. We revisited/reconstructed the PBPK model by Crowell et al. 2011) to predict the target tissue concentrations related to the exposure scenarios leading to published toxic effects on male fertility and the CNS (Goetz et al. 2007; Crofton et al. 2011), i.e. a dose of 50 mg/kg triadimefon. The maximum simulated total *in vivo* concentration (C_{max}) and its average in the 24-hour period (C_{24h}) in plasma and brain were used as relevant exposure metrics. Equal contribution of triadimefon and its metabolite triadimenol to the observed effects relative to their concentrations was assumed;

therefore the concentrations of both compounds were added. The obtained *in vivo* total concentration values were used in order to assess the equivalent *in vitro* concentrations (Fig 9a). We simulated two scenarios: dietary exposure assuming a constant intake of the entire drug dose within the first 12 h of a 24-hour period (Fig. 9b, c); oral gavage, modeled as a bolus dose into the liver compartment (Fig. 9d, e). The averaged C_{24h} simulated *in vivo* in plasma after dietary exposure was estimated to be 18 µM, with a C_{max} of 32 µM. C_{24h} and C_{max} in CSF were accordingly 15 and 27 µM. Based on the parameters listed in Fig. 9a and on the equations reported in Materials and Methods, we calculated the free concentration *in vivo* and the equivalent total concentration *in vitro* (about 1 µM in plasma, and 18-33 µM in CSF) (Fig. 9b). The same approach was used for the second analyzed scenario. A single dose of 50 mg/kg of triadimefon administrated by gavage was simulated in our model. The equivalent *in vitro* concentrations were 17-196 µM (Fig. 9d). This illustrates that *in vivo* plasma concentrations were within the range of the MINC NOAEL.

The third compound studied was PBDE-99. A new PBPK model was established for this purpose (Fig. S3). It was used to simulate concentrations in the rapid equilibrium compartment (e.g. plasma) corresponding to the lowest exposures related to neurodevelopmental toxicity described in Kuriyama et al. 2005) and Viberg et al. 2005). An oral dose of 0.11 µmol/kg (on gestational day 6, as used in the first study) was simulated to lead to a maximum concentration of 0.1 µM in our model (Fig. S5a). An oral dose of 1.4 µmol/kg (on postnatal day 10, as used by Viberg et al. 2005) led to a simulated maximum concentration of 1.3 µM, with an average concentration of 0.82 µM over the first 10 days after exposure (Fig. S5a). These concentrations were within the same order of magnitude as the MINC NOAEL.

A

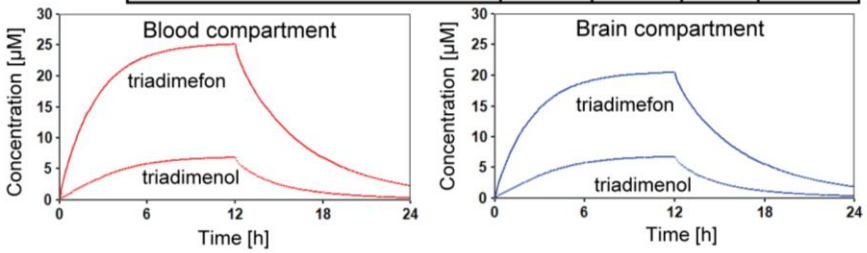
	Plasma	CSF	Medium
Lipid fraction (ml/ml)	3.6E-03	1.8E-05	2.8E-06
Albumin concentration (µM)	421	2.1	5.6

B

Dietary exposure

50 mg/kg continued exposure (for 12h) (concentrations in µM)	Plasma		CSF	
	C_{24h}	C_{max}	C_{24h}	C_{max}
Simulated total concentration in vivo	18	32	15	27
Free concentration in vivo	0.46	3.5	13	23
Equivalent total concentration in vitro	0.68	1.2	18	33

C



D

Gavage

50 mg/kg single dose exposure (concentrations in µM)	Plasma		CSF	
	C_{24h}	C_{max}	C_{24h}	C_{max}
Simulated total concentration in vivo	16	167	14	159
Free concentration in vivo	0.41	4.29	12	134
Equivalent total concentration in vitro	0.60	6.3	17	196

E

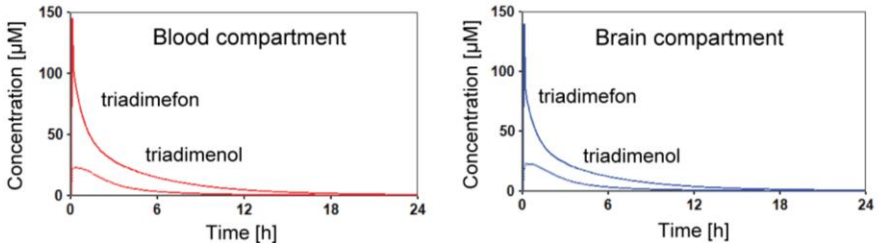


Fig.9 PBPK modeling of the pesticide triadimefon a The values of lipid fractions and albumin concentrations in rat plasma, cerebrospinal fluid (CSF) and *in vitro* cell culture medium used to calculate the equivalent *in vitro* concentrations are listed. **b-e** The toxicokinetic behavior of triadimefon and its metabolite triadimenol was simulated by using a PBPK model published by Crowell et al. 2011). Two different exposure scenarios are shown: **b,c** a dietary exposure of 50 mg/kg triadimefon (intake for the first 12 hours within a 24 h period) and **d,e** a single dose of 50 mg/kg triadimefon administered by gavage. **b** The maximum concentrations (C_{max}) and the average concentrations (C_{24h}) calculated for plasma and CSF are indicated for the first scenario: these values together with data in **a** were used to calculate the free concentration *in vivo* and the equivalent total concentration *in vitro.* **c** The simulated concentration of triadimefon and triadimenol in rat blood and brain compartment over time is shown for an oral dose of 50 mg/kg taken in the diet during the first 12 h of a 24 h period. **d** The maximum concentrations (C_{max}) and the average concentrations (C_{24h}) calculated for plasma and CSF are indicated for the second scenario: these values together with data in **a** were used to calculate the free concentration *in vivo* and the equivalent total concentration *in vitro.* **e** The simulated concentration of triadimefon and triadimenol in rat blood and brain compartment over time is shown for a single dose of 50 mg/kg.

DISCUSSION

In the present study, we defined the framework for an *in vitro* developmental toxicity test battery. Although it was originally designed for assay systems that have been developed by the ESNATS consortium (Leist et al. 2013), it may be expanded by any other robust test. Additionally, we performed a feasibility study. For this, we chose the MINC assay (Zimmer et al. 2012), because its throughput and its functional endpoint allowed us to judge the suitability of the test strategy and compound library within a reasonable time frame. The environmental chemicals with known DNT hazard potential that were included into the screen were all identified as hits. This confirmed the very high sensitivity of the assay. More importantly, new hits were identified amongst the group of medical drugs. This points to the usefulness of the test battery to provide safety information on hitherto non-characterized drugs or environmental compounds.

The field of developmental neurotoxicity relies on only a handful of generally accepted positive controls, such as mercury, lead and pesticides (Grandjean and Landrigan 2006; Kuegler et al. 2010; Kadereit et al. 2012). This situation is due to a lack of comprehensive human epidemiological data as well as a dearth of DNT guideline studies on animals (Rovida et al. 2011; Makris et al. 2009; van Thriel et al. 2012). Therefore, instead of solely relying on the compounds that have been used in the past, we applied a new approach by combining well-known positive controls with new compounds mainly belonging to the group of medical drugs. The test battery approach gives the opportunity to qualify this new group of compounds by non-tiered testing in multiple assays covering different key biological processes. We envisage that this approach will yield some extremely well characterized drug compounds, which could then be used in the future to increase the list of positive controls for *in vitro* assays. Moreover, performance of the assays would be characterized by their hit patterns and by mechanistic studies on the hits in the respective test systems.

Although we acknowledge that our approach has weaknesses concerning the definition of positive controls, it provides nevertheless a new opportunity for breaking a vicious circle between lack of sufficient reference compounds and a paucity of validated assays. We feel that this strategy is worth being explored considering also that this approach

has been successful in other fields. For instance, the area of chemically-induced carcinogenesis faces similar problems, despite more than 1000 times larger research efforts. The list of IARC (International Agency for Research on Cancer) group 1 compounds (definite human carcinogens) contains only about 4 dozen chemicals. Considering the heterogeneity of cancer, this number is far too low to validate any assay by conventional correlative statistics. Nevertheless, many assays have been developed, and many chemicals have been classified as potential carcinogens. This was done by recurrent optimisation cycles involving testing of compounds, assay optimisation, and adaptation of interpretation models. Human data are hard to obtain for chemically-induced carcinogenesis, but some key biological processes responsible for human carcinogenesis, such as mutagenesis, promotion of growth, loss of contact inhibition can be defined and tested individually *in vitro* (Adler et al. 2011). This procedure and experience may be transferred to the field of DNT. We therefore believe that increasing the size and heterogeneity of the test compound library as well as comparing the toxicity data obtained in several assays for key biological processes will be a successful strategy to overcome the lack of positive controls for DNT/DT.

By adding the new group of medical drugs to the test battery library, we also addressed a further issue: applicability domains of most existing assays are poorly defined. The currently available test systems have been characterized by the use of well-known positive control compounds, which mainly belong to the group of environmental toxicants (Klaric et al. 2013; Vojnits et al. 2012; Krug et al. 2013a; Pallocca et al. 2013; Laurenza et al. 2013a; Piersma et al. 2013, Bal-Price et al. 2012; Crofton et al. 2011; Coecke et al. 2007; Lein et al. 2007; Kadereit et al. 2012). Some others have used at most one to three drug-like compounds (Balmer et al. 2012; Stiegler et al. 2011; Kuegler et al. 2012; Meganathan et al. 2012; Jagtap et al. 2011; Falsig et al. 2004a). Therefore, it is not clear whether these test systems are able to predict DT of drugs or specific groups of environmental compounds. Apart from two screening assays developed in the context of ESNATS (Kern et al. 2013; Krug et al. 2013a) the published assays have never been challenged by a broad range of drugs, and they may therefore not have been optimized to detect their adverse effects. However, the pharmaceutical industry would need such new assays capable to predict human DT in a more reliable fashion than the currently available tests. In the past years, pharmaceutical companies have struggled more and more to develop new drugs, while at the

same time drugs already on the market had to be withdrawn due to safety issues. Between 1999 and 2011, 19% of 279 newly approved drugs in Europe were reported to have post-approval safety issues and 5 drugs had to be withdrawn from the market (Mol et al. 2013).

The strategy used here also attempts to provide a new perspective on how the field of *in vitro* DT testing may advance with regards to future validation of assay systems. The classical situation foresees the comparison of *in vitro* toxic concentrations with the known *in vivo* toxic doses, as major indication of biological relevance. This approach is not suitable for developmental toxicity testing since not enough animal data are available. As alternative, a mechanistic validation has been proposed (Leist et al. 2012c; Hartung et al. 2013a; Leist et al. 2008b). This concept is based on the assumption that toxicants disrupt key biological processes, and that test systems identify compounds that disrupt such processes (Leist et al. 2010; Crofton et al. 2011; Kadereit et al. 2012). For instance, cell migration represents such a key process (Fritsche et al. 2011; Moors et al. 2009), and inhibition of precursor cell migration in the nervous system may lead to persistent and externally visible DT. If a DNT assay is designed to identify such effects, its validity estimate would increase, as soon as it can be shown that the test is based on cell biological and signalling processes controlling precursor cell migration. Validity would increase even more, if it can be demonstrated that compounds interfering with such processes are identified as hits in the test system. Furthermore, it would be important for mechanistic validation to demonstrate that chemicals score as hits because they interrupt such defined processes, and that toxicity may be rescued by a defined, mechanistically understood counter-regulation (Krug et al. 2013a; Zimmer et al. 2011b; Poltl et al. 2012; Wang et al. 2007; Volbracht et al. 1999; Wayman et al. 2012a; Wayman et al. 2012b; Schildknecht et al. 2013). The evaluation of such a test system would thereby be independent of classical positive controls; it would solely be based on knowledge regarding toxicity pathways. Our test battery framework would allow such approaches under the following conditions: first, the same compounds will be tested in different, mechanistically defined test systems; by evaluating different endpoints and comparing the data obtained from different assays, more information will be gained on the general mode of action of those toxicants. The testing strategy needs to reach beyond the primary screening of toxicants to document adverse effects; additional mechanistic characterization in secondary assays and whole-genome transcriptomics analysis of hits is an essential part of the strategy.

This should promote our understanding of the pathways mainly affected by the exposure to the toxicants and of the pathways responsible for the toxicity.

Apart from the compound selection, we took a second unconventional approach with respect to the choice of compound test concentrations. In drug discovery, it has been general practice to screen compounds at fixed absolute concentrations. In contrast to this, we used relative concentrations. Biological activity of the compounds *in vitro* and *in vivo*, e.g. cytotoxicity, and clinical plasma concentrations were included as reference. Testing was performed relative to such reference concentrations. In most cases, the non-cytotoxic range was evaluated and chosen as starting point for the screen. However, depending on the compounds, other criteria have been considered as more appropriate; this has been the case for cytokines that were tested in concentration ranges corresponding to the levels expected in body fluids during clinical application. This strategy was designed for test systems whose throughput is limited to some extent. For assays with very high throughput, an alternative approach would be to simply screen a large number of concentrations over the entire range of compound solubility. This approach is taken for instance by the national toxicology program of the USA (Xia et al. 2008; Attene-Ramos et al. 2013; Tice et al. 2013) or the EPA ToxCast program (Judson et al. 2010; Sipes et al. 2013).

The current data are only derived from a first screening step. In the future one could imagine linking the functional disturbance of substances found in the MINC with more mechanistic data obtained in the assay itself, but also in other tests within the framework. This would allow to deepen our understanding of possible pathways of toxicity. Using the MINC assay in the first initial screen allowed us to identify 23% of compounds belonging to the drug group as potential DNT/DT toxicants. On the other hand, more than 70% of the drugs in our library did not show any specific effect, although they have potent biological activity in many other tests. This suggests that the MINC assay does not react entirely unspecifically to any drug-like compound or biologic.

Among the newly identified hits, we found three anticancer drugs (abiraterone, geldanamycin and imatinib), an anti-psychotic drug (chlorpromazine) and a cytokine, mostly used for multiple sclerosis treatment (IFN-β). Abiraterone and imatinib have been shown to promote remission of different metastatic cancers (Patel 2013). Possibly, such findings are

related to an inhibition of migration of cancer cells as well, and this may be tested in the future. Also in the case of the HSP90 inhibitor geldanamycin, other studies agree with our findings in the MINC assay. The effect of the drug on cell migration has been studied in several cancer cell lines. Activity of focal adhesion kinases (FAK), actin reorganization and integrin activation was inhibited by geldanamycin in bladder carcinoma cells (Koga et al. 2007). Furthermore, chemotactic activity in sarcoma cells (Lesko et al. 2007) as well as migration of glioma cells was reduced (Zagzag et al. 2003). Also the dopamine antagonist chlorpromazine, has been shown by other studies to inhibit migration in a pancreatic carcinoma cell line, via inhibition of k-RAS (Eisenberg et al. 2008).

Interference of IFN-β with migration has also been shown in several *in vitro* studies; the cytokine seems to modulate the activity of chemokine receptors and matrix remodeling-proteases (e.g. CCR7, MMP9) in dendritic and immune cells (Yen et al. 2010; Stuve et al. 1996). So far, these mechanistic findings have not been corroborated in an *in vivo* setting but the drug showed potential DNT activity in *in vivo* studies (US FDA, see Materials and Methods). Modeling the pharmacokinetics of the drug, we found that the toxic effect is triggered by a blood concentration of 1.9-3 pM. About 100-fold higher concentrations may be reached upon bolus injections (i.v.) (Yung et al. 1991), and the LOAEC observed in the MINC assay is about in the middle of this concentration range. The molecular targets responsible for the adverse effect of IFN-β seen in those and in our study are not known. We hope that further characterization of the compound in the MINC assay and other assays of the battery will provide more details regarding the mode of action relevant for toxicity.

For the future, we suggest that the compound library may be expanded, and data from different publications and laboratories may be stored in a common database. It cannot be stressed enough that initial hit-finding is only the first step of this strategy. Complementary information from multiple assays, pharmacokinetic modeling, and mechanistic follow-up studies are necessary components of the overall framework. For instance, transcriptome data, and measurement of other Omics and functional endpoints (Bouhifd et al. 2013; Hogberg et al. 2011; Ramirez et al. 2013; Lefew et al. 2013) will expand our knowledge regarding the toxicity mechanisms of chemicals. Moreover, such data will serve as a knowledge base concerning the biological processes and signalling pathways that need to be covered by a future DNT test battery. This will help us to identify redundancies, complementarities,

strengths, weaknesses, applicability domains and limitations of current individual assays with regards to their predictive value for DNT.

D. Results. Manuscript 2

Identification of transcriptome signatures and biomarkers specific for potential developmental toxicants inhibiting human neural crest cell migration

Giorgia Pallocca[1], Marianna Grinberg[2], Margit Henry[3], Tancred Frickey[4], Jan G. Hengstler[5], Tanja Waldmann[1], Agapios Sachinidis[4], Jörg Rahnenführer[2], Marcel Leist[1]

Affiliations:

[1] Department of in vitro toxicology and biomedicine, University of Konstanz, 78457 Konstanz, Germany;

[2] Department of Statistics, TU Dortmund University, 44139 Dortmund, Germany;

[3] Center of Physiology and Pathophysiology, Institute of Neurophysiology, University of Cologne, 50931 Cologne, Germany.

[4] Department of Bioinformatics, University of Konstanz, 78457 Konstanz, Germany

[5] Leibniz Research Centre for Working Environment and Human Factors (IfADo), Technical University of Dortmund, 44139 Dortmund, Germany;

Key words: developmental toxicity, alternative testing, transcriptome profiling, neural crest cells

ABSTRACT

The *in vitro* test battery of the European research consortium ESNATS ('novel stem cell-based test systems') has been used to screen for potential human developmental toxicants. As part of this effort, the migration of neural crest (MINC) assay has been used to evaluate chemical effects on neural crest function. It identified some drug-like compounds in addition to known environmental toxicants. The hits included the HSP90 inhibitor geldanamycin, the chemotherapeutic arsenic trioxide, the flame retardant PBDE-99, the pesticide triadimefon, and the histone deacetylase inhibitors valproic acid and trichostatin-A. Transcriptome changes triggered by these substances in human neural crest cells were recorded and analyzed here to answer three questions: i) can toxicants be individually identified based on their transcript profile; ii) how can the toxicity pattern reflected by transcript changes be compacted/dimensionality-reduced for practical regulatory use; iii) how can a reduced set of biomarkers be selected for large-scale follow up? Transcript profiling allowed clear separation of different toxicants, and the identification of toxicant types in a blinded test study. We also developed a diagrammatic system to visualize and compare toxicity patterns of a group of chemicals by giving a quantitative overview of altered superordinate biological processes (e.g. activation of KEGG pathways or overrepresentation of gene ontology terms). The transcript data were mined for potential markers of toxicity, and 39 transcripts were selected to either indicate general developmental toxicity or to distinguish compounds with different modes-of-action in read-across. In summary, we found inclusion of transcriptome data to largely increase the information from the MINC phenotypic test.

INTRODUCTION

Stem cell-based *in vitro* test systems offer new possibilities to explore toxicological hazard directly on relevant and non-transformed human cells. This novel approach particularly benefits some complex toxicological areas, such as developmental toxicity, that have not been accessible in the past to human-relevant testing.

Besides the three germ layers forming the different tissues and organs, a so called 4th layer plays a role in this complex stage: the neural crest (NC). The NC is a multipotent migratory cell population that emerges from the dorsal aspect of the neuronal tube in the early phases of development and gives rise to a multitude of different cell types, supporting the formation of cartilage, bone, connective tissue of the face, but also neurons, glial cells, melanocytes and cardiomyocytes (Huang and Saint-Jeannet 2004). A large percentage of developmental disorders (e.g. congenital heart defects, oro-facial clefts, Hirschsprung's disease) are caused by NC cell (NCC) deficit and these often correlate with neural tube defects. This kind of alterations can be induced by genetic factors (Lee et al. 2009) or exposure to pharmaceuticals (e.g. valproic acid, Fuller et al. 2002) and pesticides (e.g. triadimefon, Menegola et al. 2000).

In the field of *in vitro* DNT testing a new approach is being explored, based on the use of human cells and on the identification and modeling of distinct key biological processes representing possible targets of a toxicant. The effects of a toxicant may be described as the set of alterations of endpoints in such test systems. Examples for such test systems are the neurite outgrowth assay (Krug et al. 2013b) and the neural crest cell migration (MINC) assay (Zimmer et al. 2012, Dreser et al. 2015). These two exemplary assays are based on the quantification of a functional endpoint (neurite outgrowth and number of migrated cells) in relevant biological systems (LUHMES-derived neurons and human embryonic stem cell (hESC)-derived NCC).

Other test systems explore e.g. changes of neural differentiation (Balmer et al. 2012), of gliogenesis (Fritsche et al. 2005), of myelination (Zurich et al. 2000, Zurich et al. 2002) or synaptogenesis (Harrill et al. 2011). Usually the test systems also address defined developmental stages (Stummann et al. 2009, van Dartel et al. 2009, Zimmer et al. 2011a).

Therefore, *in vitro* testing strategies usually require a battery of tests (Leist et al. 2014, Bal-Price et al. 2015, Rovida et al. 2015a) in order to cover most of the key biological and molecular events.

Different types of test batteries have been recently presented: they may consist of molecular-based assays, like in the ToxCast program (Sipes et al. 2011, Padilla et al. 2012) or cell-based assays, as for example in the ReProTect (Schenk et al. 2010) or ChemScreen (van der Burg et al. 2014, van der Burg et al. 2015) projects.

A further step into this direction was taken by the ESNATS consortium with the establishment of a stem cell-based test battery (Zimmer et al. 2014). This testing approach was designed in a modular way to allow any interested user to join in, and to add their test system as well as the data generated from it. The ESNATS test battery has some features that distinguish it from earlier approaches, the two most important ones referring to the test chemicals selection and to the follow-up procedure of positive screen results (hits).

The usual approach of test chemical selection for new assays (Leist et al. 2010, Crofton et al. 2011, Kadereit et al. 2012) is based on the compilation of chemicals with pre-defined activity (i.e. known positives and negatives) to be used as gold standard to calibrate the assay. This approach is difficult for *in vitro* test systems of developmental toxicity for two reasons. First, only few such gold standard compounds are known from reliable *in vivo* studies or human epidemiology; second, also for known compounds, it is often not clear how they are expected to behave in an *in vitro* system. To get out of this dilemma, other approaches of assay validation and selection of initial test compound sets have been suggested (Leist et al. 2012b, Hartung et al. 2013a, Smirnova et al. 2014). In essence, such alternative strategies comprise two steps: first, the test system undergoes mechanistic validation on the basis of tool compounds that verify that the expected biochemical features and signaling pathways are represented by the system; second, once trust in the biological relevance of the systems is established, a broad set of interesting compounds can be tested. These chemicals are then classified as potentially hazardous (or not), based on the outcome of the screen (Behl et al. 2015, Pei et al. 2015). Thus, this approach takes the opposite direction from the classical approach. The advantage of this strategy is that a broad range of compounds is characterized for a potential developmental toxicity hazard, ideally across multiple test systems (test

battery). When chemicals are identified as developmental toxicity hazard, they can be used as gold standards for further test system establishment, and consequently the pool of well characterized compounds required for test system setup grows. The ESNATS test battery (Zimmer et al. 2014) and a 76 compound library of the national toxicology program of the US (Pei et al. 2015) were setup with this goal in mind.

The test compound list of the ESNATS battery comprised 28 substances, including biologics, pharmaceuticals and environmental toxicants. The background data of these compounds were extensively documented, for instance concerning general cytotoxic potency, chemical and pharmacokinetic data, as well as further relevant findings retrieved by literature data mining. In 2014, a first screening was performed using one of the assays included in the test battery, i.e. by the migration inhibition of neural crest cells (MINC) assay. This test was included in the ESNATS test battery project (as UKN2 system) to cover the developmental stage of neural crest with a cell function-specific endpoint (migration) (Zimmer et al. 2012). The MINC screening led to the identification of 11 hits, comprising all of the environmental chemicals (positive controls) and some little-characterized pharmaceuticals. In contrast to some other screening algorithms, the ESNATS test battery scheme consists of an extensive part of hit follow-up to further characterize the chemicals and their effect in the test system. This is considered an important activity towards the overall objective of identifying new potential gold standard developmental toxicants on the basis of their *in vitro* effects. One of the follow-up activities firmly anchored in the test battery scheme is the full characterization of transcriptome changes triggered by the hits. In this context it is important to note that the hits of the MINC assay are chosen because of their functional effects in the test system, i.e. because they inhibit a cell function considered to be essential for normal human development. This starting point provides thus a phenotypic anchoring of the transcriptome data to be obtained.

In the present study, we selected the six most robust and novel hits for further characterization by transcriptome profiling. They included the chemotherapeutics geldanamycin and arsenic trioxide, the flame retardant PBDE-99, the pesticide triadimefon, and the histone deacetylase inhibitors valproic acid and trichostatin-A. We performed a whole transcriptome analysis to detect changes triggered by these substances in human neural

crest cells. Three main questions were asked: i) Can transcriptome profiles of NCC be used to identify DNT compounds?; ii) How can transcriptome information be reduced to toxicological profiles?; iii) Which are the possible approaches to identify candidate biomarkers from transcriptome profiles?

To answer these questions, the present study included a blind study, probing the predictivity of compound-transcriptome pattern matching. Furthermore we developed new visualization tools to display the toxicity pattern of each substance based on transcriptome data. Finally we proposed two approaches to prioritize and select candidate biomarkers which led to the identification of 39 transcripts to be further explored as NCC toxicity indicators.

MATERIALS AND METHODS

Cell culture and neural crest differentiation

The reporter hES cell line H9-Dll1 (GFP under Dll1 promoter) was provided by Mark Tomishima from the Memorial Sloan Kettering Cancer Centre (MSKCC, NY, USA). Import of the cells and all experiments were carried out according to German legislation under the license number 1710-79-1-4-27 of the Robert-Koch Institute.

H9-Dll1 cells were maintained on Mouse Embryonic Fibroblasts (MEFs) in DMEM/F12 (Gibco) medium containing 20% of serum replacement, HEPES (1M, Gibco), L-gluthamine (Glutamax, Gibco), non-essential amino acids (MEM NEAA, Gibco), beta-mercaptoethanol (Gibco) and basic fibroblast growth factor (10 ng/ml, Invitrogen).

Differentiation of hESC into neural crest cells (NCC) was initiated on Mitomycin C treated murine bone-marrow derived stromal MS5 cell line and continued as described in Zimmer et al. 2012.

Chemical exposure during NCC migration

hESC- derived NCC were exposed for 48 hours to non-cytotoxic concentration of different NCC migration inhibiting substances in N2 medium containing EGF (20 ng/ml) and FGF2 (20 ng/ml). Six compounds were used: geldanamycin (16 nM, Selleckchem), arsenic trioxide (1µM, Sigma Aldrich), thricostatin A (TSA, 10 nM, Sigma Aldrich), valproic acid sodium salt (VPA, 250 µM, Sigma Aldrich), triadimefon (100 µM, Bayer Crop Science) and pentabromodiphenyl ether (PBDE-99, 15µM, Clickchem). Two different solvent control-groups were also produced: NCC were exposed to 0.04% DMSO or simply to N2 medium for 48 hours. Finally, a third control group (exposed to N2 medium only) was added specifically as control of arsenic trioxide, since the testing of this substance was performed not in parallel with the other compounds.

Affymetrix gene chip analysis

Samples of $\geq 5 \times 10^6$ cells were collected using RNA protect reagent from Qiagen. The RNA was quantified using a NanoDrop N-1000 spectrophotometer (NanoDrop, Wilmington, DE, USA), and the integrity of RNA was confirmed with a standard sense automated gel

electrophoresis system (Experion, Bio-Rad, Hercules, CA, USA). Analysis was then performed as described earlier (Krug et al. 2013c) using Affymetrix chip-based DNA microarray (Human genome U133 plus 2.0 arrays) with all standard quality control procedures.

Biostatistics

The microarray data analysis (extrapolation and normalization of the array sets) was performed using the statistical programming language R-version 3.1.1. as described in Waldmann (2014). For the normalization of the entire set of Affymetrix gene expression arrays, the Extrapolation Strategy (RMA+) algorithm (Harbron et al. 2007) was used that applies background correction, $\log_2$ transformation, quantile normalization, and a linear model fit to the normalized data in order to obtain a value for each probe set (PS) on each array. As reference, the normalization parameters obtained in earlier analyses (Krug et al. 2013c) were used. After normalization, the difference between gene expression and corresponding controls was calculated (paired design). Differential expression was calculated using the R package limma (Smyth et al. 2005). Here, the combined information of the complete set of genes is used by an empirical Bayes adjustment of the variance estimates of single genes. This form of a moderated t-test is abbreviated here as 'Limma t test'. The resulting p-values were multiplicity-adjusted to control the false discovery rate (FDR) by the Benjamini-Hochberg procedure (Benjamini 1995). As a result, for each compound a gene list was obtained, with corresponding estimates for log fold changes and p-values of the Limma t test (unadjusted and FDR-adjusted).

Transcripts with FDR adjusted p-values of ≤ 0.05 and fold change values of ≥ 1.8 were considered significantly deregulated and defined as differential expressed genes (DEG).

Data display: heat map and principal component analysis

The software R (version 3.1.1), was used for all calculations and display of PCA and heatmaps. Principal component analysis (PCA) plots were used to visualize expression data in two dimensions, representing the first two principal components. The percentages of the variances covered are indicated in the figures. Heatmaps were used to visualize matrices of gene expression values.

The hierarchical clustering analysis was performed as previously described (Krug et al. 2013c). Complete linkage was used as agglomeration rule for the clustering analysis. Distances based on the Euclidean distance measure were used to group together transcripts with similar expression patterns across samples (rows of the heatmap). Then, expression values within each row were normalized as Z-factors and color coded accordingly. Color encodes the magnitude of the values as z-score, ranging from blue (low) to yellow (high).

Support vector machine-based classification

A Support Vector Machine algorithm with linear kernel was used for the discrimination between two data sets: a training group composed of three biological replicates and a testing group composed of two biological replicates (with compounds blinded to the experimenter) using the same set of compounds. Both groups were normalized to the respective controls, i.e the difference between gene expression and corresponding controls was calculated (paired design). Geldanamycin , PBDE-99 and triadimefon had common controls, valproic acid (VPA) and trichostatin-A (TSA) were assigned to the same set of controls, and arsenic trioxide had its own set of controls. After subtracting controls, the number of variables was reduced to the 100 probe sets with highest variance within the training set. Then, in a second step, the hyperparameters for optimizing the decision boundary between the known training compounds were determined (using a grid search over supplied parameter ranges). These parameters were then used to generate the classification model to predict for the blinded testing sample the probabilities to belong to the known training compounds. For multiclass-classification with more than two classes, first in a 'one-against-one'-approach, all possible binary classifiers were trained and corresponding probabilities were calculated from a logistic regression as described in Rempel et al. 2015 . Then a-posteriori class probabilities for the multi-class problem are obtained using quadratic optimization.

Gene ontology (GO) and KEGG pathway enrichment analysis

The gene ontology group enrichment was performed using R-version 3.1.1 with the topGO package (Alexa et al. 2006) using Fisher’s exact test, and only results from the biological process ontology were kept. Here again the resulting p-values were corrected for multiple testing by the method of Benjamini−Hochberg (Benjamini 1995).

The KEGG pathway analysis was performed using the R package "hgu133plus2.db" (Carlson 2015). Probesets are mapped to the identifiers used by KEGG for pathways in which the genes represented by the probesets are involved. The enrichment was then performed analogous to the gene ontology group enrichment using Fisher's exact test.
Up- and down-regulated differentially expressed genes were analyzed separately for each treatment. Only GO classes and KEGG pathways with a BH-adj. p-value ≤ 0.05 were considered significant.

Toxicity pattern visualisation

ToxPi diagrams, as developed in the ToxCast project (Kleinstreuer et al. 2011a, Filer et al. 2014) were constructed using a web-based user interface (Reif et al. 2013). For this purpose, the numbers of DEG, of over-represented GO groups and KEGG pathways were normalized to the highest respective values for each category across all compounds (359 for DEG, 373 for GO classes and 17 for KEGG pathways). The ToxPi score was calculated over the six parameters used for ToxPi construction, with double weight given to KEGG and GO vs DEG.

GO classes and KEGG pathway terms were included only when ≥ 3 differentially expressed genes were found in the enriched term, or (for small groups) when $\geq 50\%$ of the genes belonging to the GO/KEGG group were found to be significantly altered in our study.

GO superordinate classes distribution

The gene ontology group enrichment was performed as described above. Up- and down-regulated differentially expressed genes were analyzed together for each treatment. Only GO classes with a BH-adj. p-value ≤ 0.05 were considered significantly enriched. Classic and elim methods (described in Alexa et al 2006) were both used: classic method was chosen for geldanamycin treatment, while elim method was used for the analysis of the other compounds. The elim algorithm iteratively removes the genes mapped to significant GO terms from more general (higher level) GO terms, whereas the classic algorithm neglects the local dependencies between GO terms in its calculations.

Enriched GOs were then assigned to superordinate cell biological processes as already described in Waldmann et al 2014 and distributed in six classes: migration/adhesion,

metabolism, differentiation, signaling, stress response and others. The migration class includes migration and adhesion related- GO classes; stress response class includes cell death- , extra-cellular stressor- , inflammation/immunity- related GO classes; signaling class consist of cell receptor activity- , second messenger (cAMP, cGMP, Ca2+) metabolism- , kinase modification- related GO classes. Metabolism class comprise all GO classes covering metabolism activity; differentiation class includes cell differentiation related- GO classes and; "other" class covers all the others: not otherwise classified GO classes.

Biomarker quality control

To evaluate how the here chosen set of 39 biomarkers (set_A) compared to random combinations of 39 genes (set_X_i; i = 1-1000000), a simple metric for the separation power (controls vs six test compounds) of biomarker sets was developed and expressed as separation units (SU). The distribution of these SU was then determined by bootstrapping (one million samples), and the relative position of set_A in this distribution was determined. The metric is based on the following procedural steps: (i) For each of the 54,675 probesets (PS) a T-score was calculated for comparison of one of the six test compounds with the control (with n = 5 for controls and test compound samples); i.e. six T-values were obtained for each PS (one for each of the compounds geldanamycin (GA), triadimefon (TDF), VPA, TSA, arsenic trioxide (AS2O3) and PBDE); (ii) only the PS with at least one compound showing a difference from control with $p < 0.05$ were considered in further steps; (iii) the T-scores of the group of PS assigned to the same gene were averaged; (iv) the genes were ranked according to their T-scores. This was done individually for each test compound; then the ranks were normalized to a ranking value ("r"; values from 0 to 1; where r = 0 corresponds to low T-score (low grade of separation between control and exposure) and r = 1 to high T-score (high grade of separation)); (v) the rank values of 39 biomarker genes (of set_X_i or set_A) were averaged to give a "biomarker separation rank" for one given compound; (vi) the six biomarker separation ranks for the toxicants were assumed to define a vector in six-dimensional space, indicating the distance of toxicants from the control. The length of this vector was used as metric for the SU:

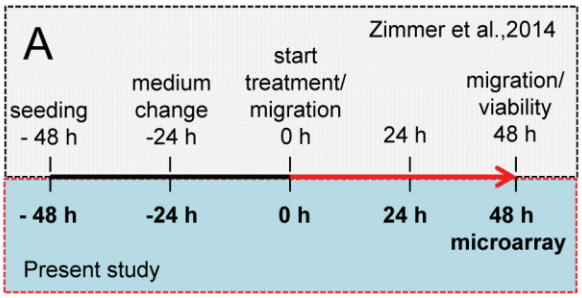
The length of this vector was in the range 0 – 2.45. The randomly chosen set_X_i had an average SU of 1.256 ± 0.081, the set_A was at 1.75 (i.e. > 6 SD larger than the average).

RESULTS

Data structure of transcriptional changes induced in neural crest cells exposed to migration – inhibiting concentrations of test battery hits

To explore the effects of neural crest toxicants on the transcriptome, six compounds were selected that had been shown earlier to inhibit migration of human neural crest cells (NCC) in the MINC assay (Zimmer et al. 2014, Dreser et al. 2015). They comprised the heat shock protein modifier and new chemotherapeutic lead compound geldanamycin (GA), the chemotherapeutic agent and environmental toxicant arsenic trioxide (As_2O_3), the brominated flame retardant PBDE-99, the triazole pesticide triadimefon (TDF), and the histone deacetylase inhibitors (HDACi) trichostatin-A (TSA) and valproic acid (VPA). To allow comparisons amongst the compounds, NCC were exposed for 48 h to each of the toxicants at their respective highest non-cytotoxic concentration and to solvent controls. Then, mRNA was prepared from three different cell lots and used for gene expression analysis by Affymetrix microarray technology (Fig.1A).

After initial data processing and quality controls, the gene expression data was used to determine the differentially expressed genes (DEG). The DEG for each toxicant were defined here as the group of microarray probe sets (PS), which differed significantly from negative controls (FDR adjusted p-value ≤ 0.05), and showed expression level changes (fold change (FC)) of ≥ 1.8 or ≤ 0.55. On this basis, the toxicant effects could be roughly categorized: strong effects on the transcriptome were observed for As_2O_3 (478 DEG) and PBDE-99 (443 DEG); a medium effect was detected for GA (93 DEG) and triadimefon (40 DEG); while no effect was observed for TSA and VPA (0 DEG). Moreover, we observed three different types of response, concerning the direction of gene regulations: a predominant upregulation (ratio between upregulated DEG (DEG up) and downregulated DEG (DEG dw) ≥ 10 (e.g. GA); a predominant downregulation, with a ratio between DEG up and DEG dw of ≤ 0.1 (e.g. triadimefon); and mixed, bi-directional regulations (e.g. As_2O_3 and PBDE-99) (Fig. 1B).



B Differentially-expressed genes (DEG)
(p-value ≤ 0.05; |FC| ≥ 1.8; n = 3)

Toxicant	*DEG up*	*DEG down*
Geldanamycin	92	1
Arsenic trioxide	137	341
PBDE-99	334	109
Triadimefon	1	39
TSA	0	0
VPA	0	0

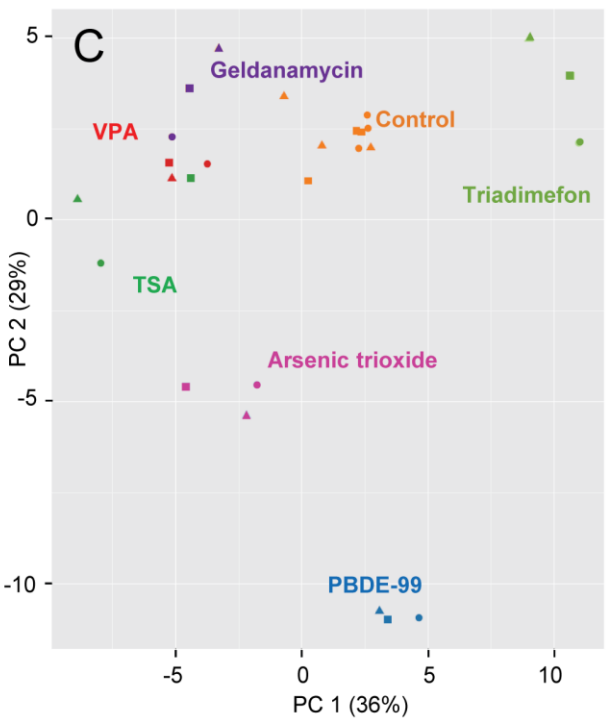


Fig.1 Experimental design and transcriptome data structure of test battery hits. A Sampling for microarray analysis was performed in neural crest cells after 48 hour-exposure (red arrow) to non-cytotoxic and migration-inhibiting concentrations of six test battery hits (geldanamycin, arsenic trioxide, PBDE-99, triadimefon, TSA and VPA), as identified by the method and the data published in Zimmer et al., 2014. B The differentially-expressed genes (DEG) were identified for each condition: the number of up- (DEG up) and down-regulated genes (DEG down) are shown in the table (details can be found in the supplementary material; p-values were FDR-corrected). C Principal component analysis (PCA) was performed, based on the 100 transcripts with highest variance, and a 2D-plot was generated to display the transcriptome data structure across compounds and experimental replicates. Each point represents one experiment (= data from one microarray), where the color coding indicates the compound and the form of the data points indicates the replicate. The percentages of the variances covered are indicated on the axes.

To visualize the different gene expression profiles across all compounds and replicates, a principal component analysis (PCA) was performed, based on the 100 PS with highest variance among the samples. Plotting of the first two principal components showed that the replicates of each compound clustered closely together, while the compounds clearly separated from the nine controls and from one another. As expected, the two "strong-effect" compounds, As_2O_3 and PBDE-99, showed the most distinct separation from controls and from one another. Notably, also HDACi (TSA and VPA) separated clearly from controls, but they could not be separated from each other (Fig. 1C). This separation effect in the PCA

(despite the absence of DEG for HDACi) was due to the combined use of an ensemble of 100 PS selected by highest variance amongst samples, instead of a gene-by-gene comparison.

In conclusion, the initial analysis of the microarray data showed that the quality of the gene expression data sets was high enough for further exploration, and that different migration inhibitors appeared to trigger distinct signatures of gene expression changes in NCC.

Computational toxicant identification, based on correlation between transcriptome data sets

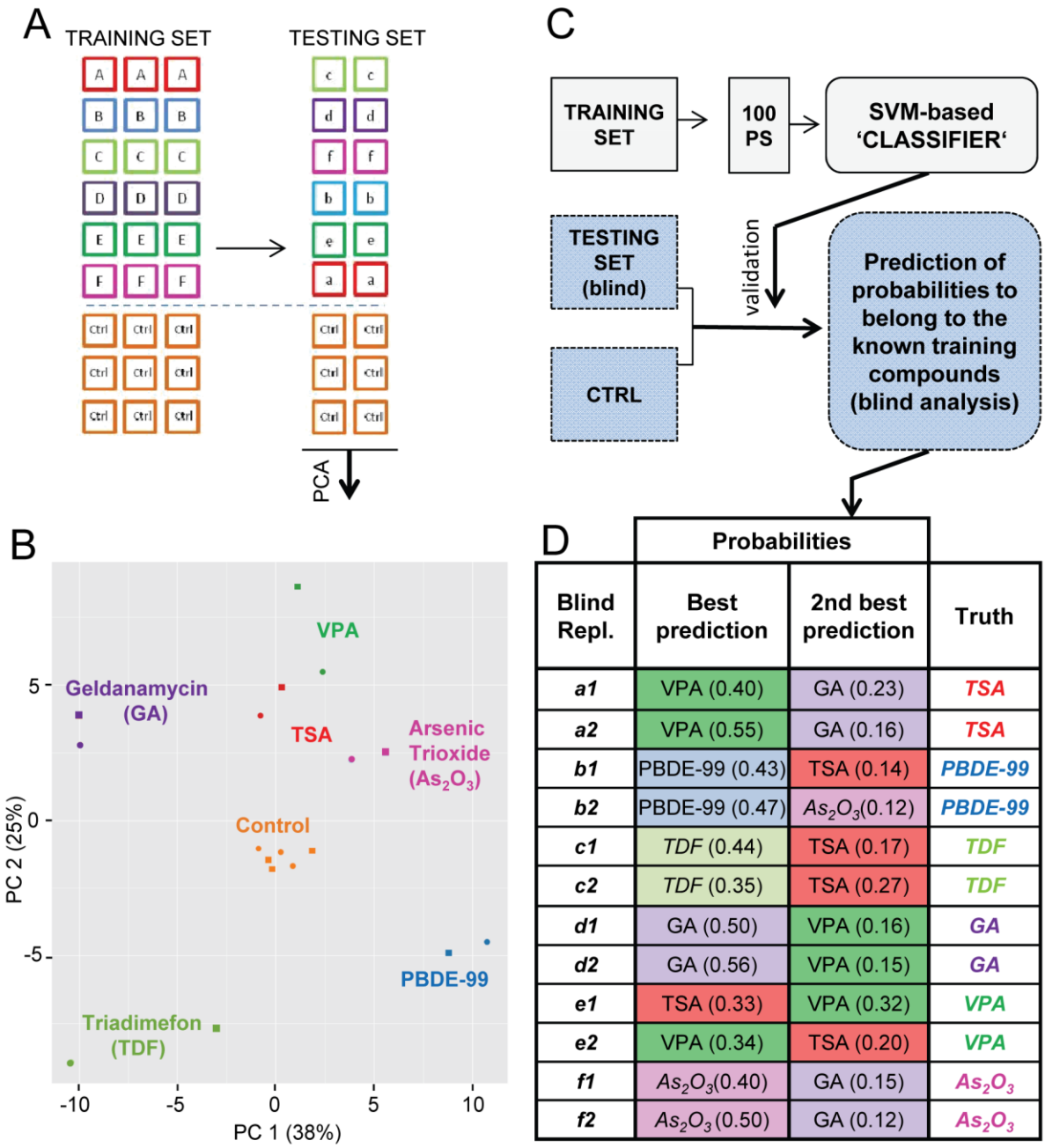
While the initial analysis suggested that a *post-hoc* separation of toxicants was possible to some extent, we were interested in the predictive power of the transcriptome analysis. For this purpose, the data sets obtained initially were used as training set of a classification rule, while new data for each compound (n = 2 additional experiments) were obtained as testing set (Fig. 2A). The 'testing set' experiments were performed in a way that it was known to the experimenter, which samples were negative controls, while the identity of the toxicants to be tested was blinded.

To obtain an overview over the new data, a PCA analysis of the testing set microarrays was performed after 'unblinding' of the data to the person doing the analysis. The results showed that the structure of the collected transcriptome data resembled the one of the training set, i.e. all toxicant samples separated 'visually' from the six controls, and the couples of toxicant replicates were closer to one another than to other data points (Fig. 2B).

In the next step, the 100 probe sets with highest variance within the training set were used for the classification analysis, in an attempt to identify the treatment of each of the blinded samples of the testing set. For this purpose, a support vector machine approach was used to discriminate between the six different compounds. The training set composed of three biological replicates was used to generate the classifier which includes 1) the selected probe sets, 2) the optimized hyperparameters of the model and 3) the classification rule for belonging to a particular compound. The classifier was then applied to obtain for each blinded testing sample the probabilities to belong to the known training compounds. The probabilities were then sorted in descending order and the training set condition yielding the

highest probability of affiliation was taken as ‘best prediction’ for the corresponding testing set data (Fig. 2C).

Nine out of twelve blind samples were correctly predicted (75% predictivity), i.e. the correct compound was assigned to the respective microarray. All the ‘wrongly’ predicted samples belonged to the HDACi group, and predictions were correct within the group. If prediction of an HDACi was accepted as correct prediction for VPA or TSA, the overall predictivity of the test was 100% (Fig. 2D).



Blind Repl.	Probabilities: Best prediction	Probabilities: 2nd best prediction	Truth
a1	VPA (0.40)	GA (0.23)	*TSA*
a2	VPA (0.55)	GA (0.16)	*TSA*
b1	PBDE-99 (0.43)	TSA (0.14)	*PBDE-99*
b2	PBDE-99 (0.47)	As_2O_3(0.12)	*PBDE-99*
c1	*TDF* (0.44)	TSA (0.17)	*TDF*
c2	*TDF* (0.35)	TSA (0.27)	*TDF*
d1	GA (0.50)	VPA (0.16)	*GA*
d2	GA (0.56)	VPA (0.15)	*GA*
e1	TSA (0.33)	VPA (0.32)	*VPA*
e2	VPA (0.34)	TSA (0.20)	*VPA*
f1	As_2O_3(0.40)	GA (0.15)	As_2O_3
f2	As_2O_3 (0.50)	GA (0.12)	As_2O_3

The predictivity of the classifier was not affected by the separation of the five microarrays per compound into training and testing set: a simulation study was performed by random selection of the replicates belonging to the training or testing data sets: 1000 different combinations were analyzed and for all compounds, with exception of the HDACi group, ~100% correct predictions were observed (Fig. S1A).

◄ **Fig.2 Correlation analysis between „training set“ and blind „testing set“ data.** A Two data sets were generated: a training set, based on the data presented in Fig.1 (boxes with capital letters, n = 3) and a testing set, based on data obtained by additional two replicates (with compounds blinded to the experimenter) using the same set of compounds (boxes with small letters, n = 2). Both groups were normalized to the respective controls (Ctrl, orange boxes). B A PCA plot based on the 100 transcripts with highest variance was generated to display the structure of the transcriptome data of the testing set along the first two principal components. C The 100 probe sets with highest variance (“100 PS”) within the training set were identified. Then, a classifier was built using the support vector machine (SVM) approach (see methods). Finally, the probabilities of the blinded testing samples to belong to the known training compounds were predicted. D The best and second best predictions, based on a support vector machine approach (indicated as relative probability in the brackets), are listed for each blind replicate (first column). The real identity of the samples (truth) is indicated in the last column. For instance, the highest probability (50% likelihood) for blind sample d1 was obtained for geldanamycin (GA), the second highest (16% likelihood) was for VPA. The unblinding of the sample revealed it to be GA.

The prediction may suffer from a skewing of our compound collection, as the two HDAC inhibitors formed a group amongst themselves and thus led to an overrepresentation of a defined toxicological mechanism within the group. For this reason, we performed a second set of analyses after exclusion of TSA from the compound set. In this new scenario, we reached 100% predictivity, with 10 of 10 correct predictions (Fig. S1B).

The outcome of this small blind-testing study suggests that microarray data may be useful to assign unknown compounds to pre-defined groups of compounds, for instance to obtain some initial toxicological information or for biology-supported read across.

Characterization of toxicant-induced transcriptome-profiles

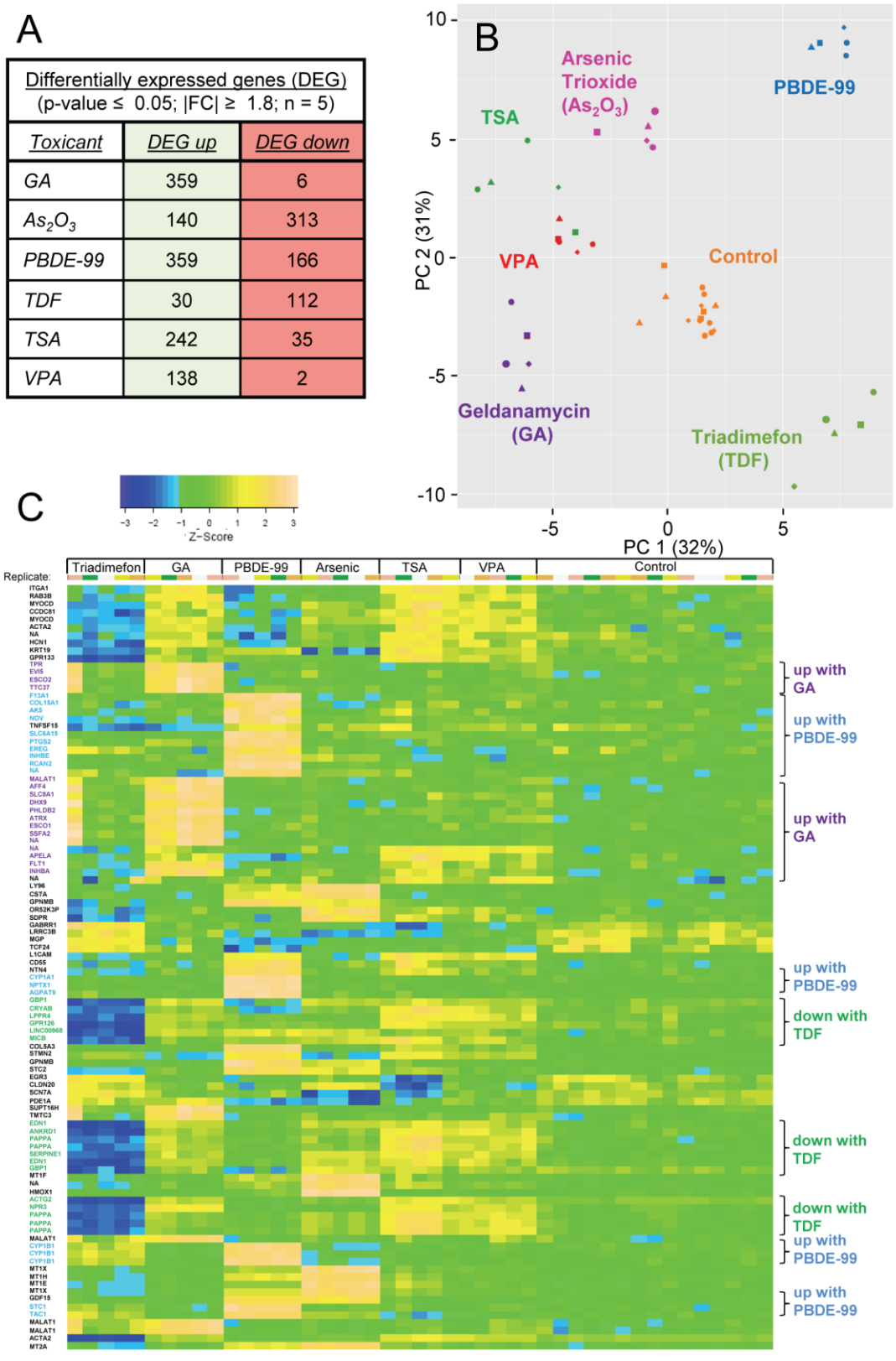
To obtain more in-depth insight into the transcriptome changes induced by each compound, the training and testing set data were combined and analysed together for DEG. The increase of the statistical power (due to the increase of replicate numbers; n = 5) allowed the detection of a higher number of DEG. For As_2O_3 (453 DEG) and PBDE-99 (525 DEG) treatments, the increase was rather moderate, for GA (365 DEG; nearly all up-regulations) and triadimefon (142 DEG; mixed regulation pattern), the increase was substantial. The increased sensitivity of the analysis was particularly evident for TSA (277 DEG) and VPA (140 DEG), which mainly led to gene up-regulation (Fig. 3A). The new PCA plot over all conditions (15 controls *plus* 6 x 5 toxicant samples) showed that all compounds separated from controls and from one another, if the HDACi TSA and VPA were considered as one group (Fig. 3B).

The differential effects of the toxicants on gene expression were also evident from a heatmap display that shows the relative expressions of the 100 PS with largest variability. This allowed some first insight on the level of individual genes. For example, a group of PS (n = 13) up-regulated by geldanamycin only comprised the three cell cycle controllers such as ESCO1, ESCO2 (N-acetyltransferases involved in establishment of sister chromatid cohesion), MALAT1 (Tripathi et al. 2013, Yang et al. 2013) and ATRX (Berube et al. 2000). Other examples are the group of PS who were specifically upregulated upon PBDE-99 treatment, such as the cytochromes CYP1A1 and CYP1B1 or tachykinin, or the 17 genes down-regulated only by triadimefon (e.g. the inflammation-related factors MICB, endothelin and GBP1) (Fig. 3C).

Fig.3 Overall transcriptome changes based on pooled data sets. A Training and testing expression data sets were analyzed together to identify overall transcriptome changes. The number of differentially-expressed genes significantly up- or down- regulated is listed in the table for each condition (detailed data are shown in supplemental material). **B** The transcriptome data structure was displayed as principal component analysis (PCA) 2D- plot showing the first two principal components. Percentages of covered variances are indicated on the axes. **C** The transcriptome data were represented as heatmap indicating the gene expression values of the top-100 probe sets with the highest variability among the compounds. Expression values of the individual genes were transformed to z-scores (along rows). On the right some exemplary gene groups are indicated for guidance (e.g. genes upregulated specifically by geldanamycin). ►

A

Differentially expressed genes (DEG) (p-value ≤ 0.05; |FC| ≥ 1.8; n = 5)		
Toxicant	*DEG up*	*DEG down*
GA	359	6
As_2O_3	140	313
PBDE-99	359	166
TDF	30	112
TSA	242	35
VPA	138	2



Encouraged by this visual exploration, the gene-expression changes were explored quantitatively on the PS level. The DEG for each toxicant were sorted according to their p-values and the top 20 up- and down-regulated genes were selected. This approach focussed on the statistically most-important regulations by each toxicant vs control conditions, without taking into account whether other toxicants affected the same gene (Fig. 4).

Arsenic trioxide affected many heavy metal-induced genes (up-regulation of different metallothioneins (MT) isotypes; 30-fold upregulation of the heme oxigenase 1), and it up-regulated the protease inhibitor cystatin-A, similar to the situation observed in human T-cells after treatment with arsenic (Shao et al. 2014). PS strongly down-regulated by arsenite comprised CXCL14, a chemokine which is known to play a role in cell migration, and LRRC3B, a gene whose cord blood leukocyte DNA methylation pattern was found altered in a cohort of newborns pre-natally exposed to arsenic in water (Rojas et al. 2015).

The exposure to geldanamycin led to a 20-fold upregulation of PHLDB2. The product of this pleckstrin homology domain gene, also known as LL5β, is a protein implicated in migration and tumor cell invasion, by stabilizing of the protrusive activity at the cell front (Astro et al. 2014). Down-regulations by geldanamycin were only moderate, and they comprised e.g. the metalloproteinase MMP1and the cell matrix proteoglycan PRG4.

Among the genes up-regulated by PBDE-99, we identified neural pentraxin (NPTX1; ~20 fold up-regulated), whose cognate protein exclusively localizes to the nervous system. Moreover, extracellular matrix factors involved in axon guidance were affected, like netrin 4 (NTN4) or the collagens, COL5A3 and COL15A1. The down-regulated genes comprised the thioredoxin interactin protein (TXNIP), whose down-regulation was also observed in HUVEC (Kawashiro et al. 2009) and H295R adrenocortical carcinoma cells (Song et al. 2009) after exposure to polybrominated diphenyl ethers. Other examples of down-regulations are contactin1 (a neuronal membrane protein with functions in cell adhesion and the formation of axon connections in the developing nervous system), or NDST4, an enzyme involved in extracellular matrix (heparin) synthesis.

In triadimefon-treated cells, upregulations were very moderate. Among the most down-regulated genes, we identified the ankyrin repeat domain' (ANKRD1) as well as α and β actin. Moreover, the down-regulted DEG comprised endothelin1, which is involved in neural

crest patterning (Pla and Larue 2003). The attenuated activity of this gene may be related to the neural crest toxicity of the pesticide (Di Renzo et al. 2011c; Menegola et al. 2005).

HDACi up-regulated e.g. myocardin, similarly as observed in forebrain precursor cells (Balmer et al. 2014), the cell adhesion molecule (L1CAM) and desmoglein, a cell-cell junction glycoprotein. Some of the top down-regulated genes (LRRC3B and cadherin-8) were similar to those downregulated by HDACi in central neural precursors (Balmer et al. 2014).

We also observed deregulation of a particular group of genes linked to neurocristopathies, a specific class of pathologies deriving from NCC deficits. Amongst the geses known to be linked to neurocristopathies, three groups were also affected by toxic chemicals: first, endothelin and its receptors (EDN1 and EDNRB) (Kurihara et al. 1994), the expression of which was altered by all toxicants except for PBDE-99; the neurofibromatosis (NF) gene family (Nakamura 1995) which was altered specifically in geldanamycin treated NCC (NF1 and NF2); and the netrins (Amiel et al. 2008), which were altered in PBDE-99 (NTN4) and TSA (NTN4 and NTNG) treated cells.

This analysis on the level of individual genes/PS suggested effects of the toxicants on very different pathways and biological processes (Fig. 4), but such more narrative descriptions are of limited use for toxicological hazard estimates and quantitative approaches. To explore options for quantifications of transcriptome changes, and for comparisons amongst compounds, we employed unbiased approaches to identify disturbed higher-order biological processes.

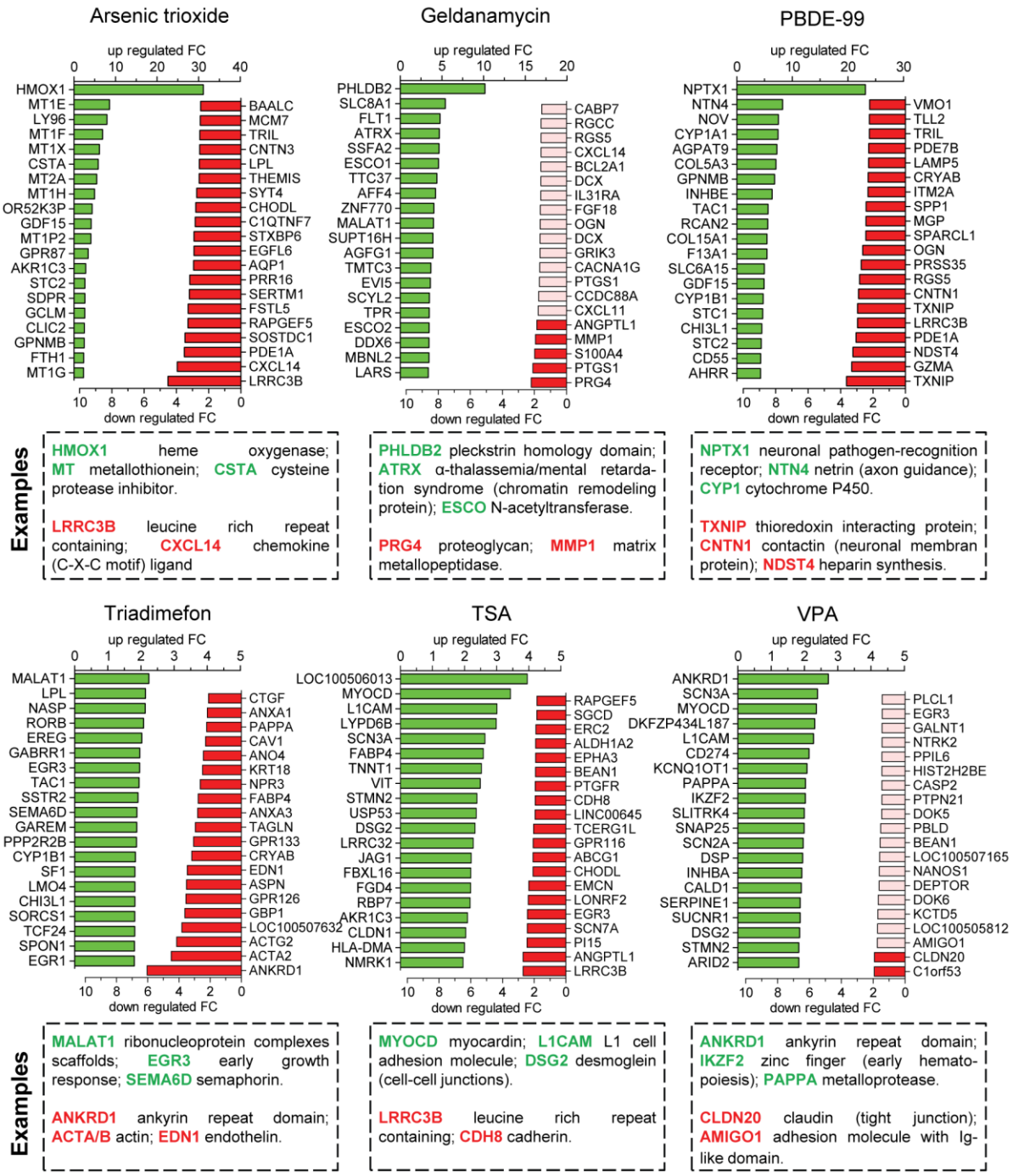


Fig.4 Selection of the 20 most significant up- and down- regulated genes for hit compounds. The differentially regulated genes were identified for six hit compounds and sorted according to their p-value. The top 20 up- (green) and down- (red) regulated genes for each condition are shown as bar graphs indicating the fold change (FC). Genes regulated with a FC below the threshold of 1.8 are indicated in light red. Few example genes were chosen according to their toxicological/pathophysiological interest level according to the literature. This biased selection is displayed only as initial rough overview and food-for-thought for later marker selection.

Visualization of toxicity patterns based on co-ordinate regulation of genes involved in joint superordinate biological processes

To compare the effects on the transcriptome regulation among the different conditions, a multi-dimensional representation was chosen, as pioneered earlier e.g. by the ToxPi approach in the ToxCast program (Reif et al. 2013), or by the use of toxicity indices developed on the basis of superordinate biological processes (Waldmann et al. 2014). Such descriptors go beyond the level of individual genes, by quantifying the regulation of entire gene ontologies (Theunissen et al. 2011, Theunissen et al. 2012, Waldmann et al. 2014), and forming aggregate measures or simplified visualizations.

In a first step, the pattern of transcriptome changes was visualized on the basis of six key parameters: the number of up- and down-regulated DEG, the number of GO terms enriched amongst up-regulated (GO up) and down-regulated DEG (GO dw), and the number of KEGG pathways enriched by up-regulated (KEGG up) and downregulated DEG (KEGG dw) (Fig.5 and 6). To facilitate comparisons, the underlying data were normalized across all compounds to the respective maximum value observed in the whole study. This procedure is similar to the one taken by EPA in their ToxPi approach, and it allows direct comparisons of the patterns observed in the radar plots used here. For instance, it becomes easily evident that TSA was characterized by overrepresented GO terms only amongst its up-regulated DEG, while the reverse was observed for triadimefon (only GO terms amongst down-regulated DEG). For information beyond an initial overview, all detailed data were compiled in tabular form (Supplementary tables S1-3).

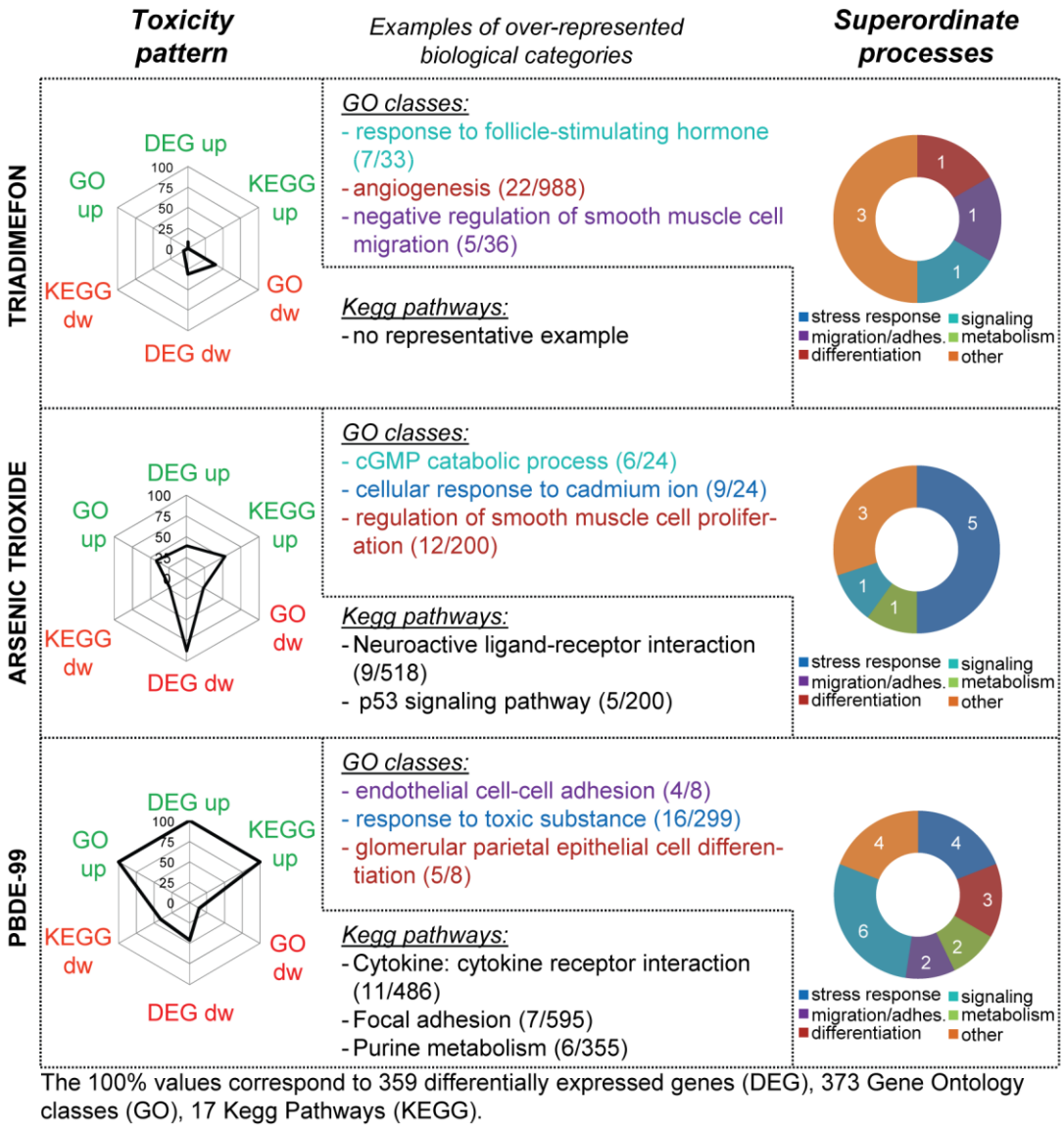


Fig.5 Alteration of superordinate biological processes by environmental pollutants. Graphical display to visualize broad information on biological changes, i.e. the toxicity patterns triggered by the compounds belonging to the class of "environmental pollutants" (triadimefon, arsenic trioxide, PBDE-99): the spider diagrams (on the left) indicate normalized numbers of up- and down- (dw) regulated differentially expressed genes (DEG); of GO terms overrepresented amongst DEG (GO); and of overrepresented KEGG pathways (KEGG). The absolute values used for normalization of each axis are indicated at the bottom of the figure. They correspond to the respective highest value for all 6 compounds. The ring diagrams on the right hand side show the relative distribution of 6 superordinate biological processes (stress response, migration/adhesion, metabolism, differentiation, signaling and other) amongst the DEG. The data are based on the counting of non-redundant over-represented GO terms (detailed table in supplemental material) within each superordinate biological process category (white numbers). Identification of overrepresented GO was done on the basis of all DEG (up- and down-regulated). In the middle, examples of over-represented biological categories (KEGG, GO) are shown (with number of regulated and total genes belonging to the specific group) and color coded according to the superordinate process.

In a second step, the types of biological processes that may be linked to the altered transcriptome of toxicant-treated cells were visualized. For this purpose, all DEG (up and down) were pooled for a given compound. Then GO term enrichment analysis was performed by a statistical method, the so-called elim alogorimth that eliminates 'redundant' GO groups, i.e. such GO groups that contain the same genes, but do not provide new information (so-called children GOs or parent GOs of a given term) (Alexa et al. 2006). The resultant unique overrepresented GO groups were assigned to five superordinate cell biological processes (Waldmann et al. 2014): stress response, signalling, migration/adhesion, metabolism, differentiation; and those that could not be assigned were grouped under 'other'. The category 'stress response' combined all GO terms related to cell death, extra-cellular stress, inflammation and immunity; the 'signalling' category comprised GO terms related to cellular receptors, second messengers, hormones/neurotransmiters and kinase modifications/regulations. The relative contribution of the different superordinate processes to the altered transcriptome response was shown in form of ring diagrams. These allow for instance a quick overview that suggests that the response to arsenite (Fig.5) is dominated by 'stress responses', while HDACi responses have a strong 'signaling' component (Fig.6). In case, more detailed information is required, then this can be retrieved from tabular compilations (Supplementary table S4).

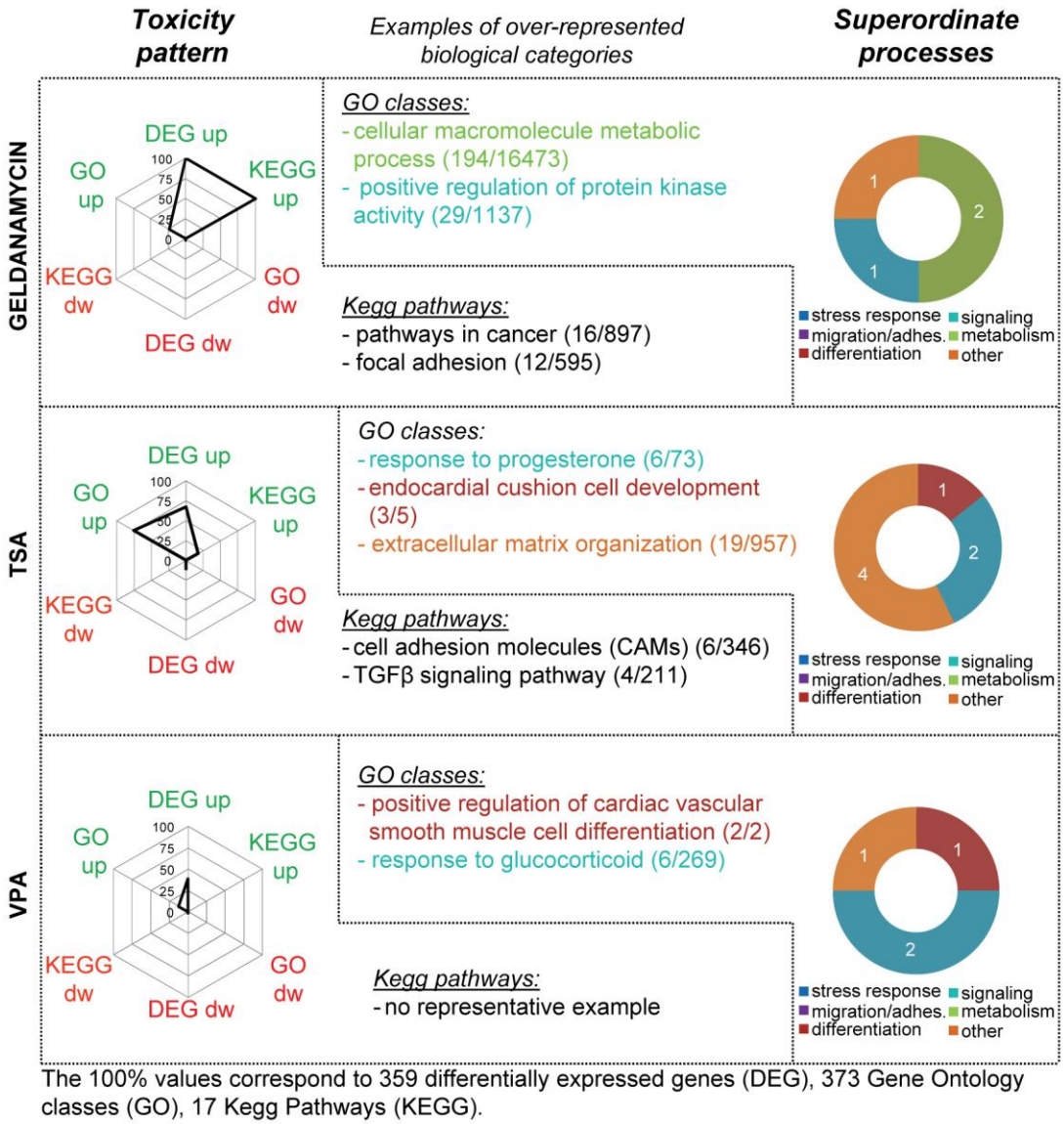


Fig.6 Alteration of superordinate biological processes by drugs. Toxicity patterns triggered by the compounds belonging to the class of "drugs" (geldanamycin, TSA, VPA) are presented as in Fig.5. The spider diagrams indicate normalized numbers of differentially expressed genes (DEG), of GO terms overrepresented amongst DEG (GO), and overrepresented KEGG pathways (KEGG). The absolute values used for normalization of each axis are indicated at the bottom of the figure. They correspond to the respective highest value for all 6 compounds. The ring diagrams show the relative distribution of 6 superordinate biological processes (stress response, migration/adhesion, metabolism, differentiation, signaling and other) amongst the DEG. In the middle, examples of over-represented biological categories (KEGG, GO) are shown (with numbers of regulated and total genes belonging to the specific group).

Altogether, these overview presentations of transcriptome changes showed at the first glance that the three environmental pollutants had very different effects on neural crest cells, although they all inhibited migration (Fig.5): triadimefon had a relatively modest influence on the transcriptome, compared to PBDE-99; and while the latter one mainly upregulated transcripts, arsenite predominantly downregulated gene-activity. However, arsenic trioxide also upregulated some transcripts and these pointed for instance to the activation of the p53 pathway, which has been implicated earlier in arsenite toxicity (van Vliet et al. 2007).

The 'toxicity patterns' obtained for the drug-like group of test chemicals were characterized by a predominant upregulation of DEG and associated biological processes (Fig.6). Geldanamycin up-regulated genes related to 'cancer' KEGG pathways, but also focal adhesion, a process that may be related to the migration-inhibitory activity of this compound in neural crest cells.

The TSA toxicity pattern was mostly characterized by upregulation. Amongst the up-regulated DEG, only few KEGG pathways were enrichment (e.g. cell adhesion molecule and TGFβ signalling pathway), but a high over-representation of GO terms was found. These included 'cardiac development' as well as 'steroid signalling', and the latter two were also found for the related compound VPA. In general, VPA triggered a much less pronounced transcriptome response. This is well in line with a relatively specific pharmacological activity of this compound, although it is used at high concentrations in clinical settings, and its mode of action is supposed to affect the chromatin structure of a large number of genes.

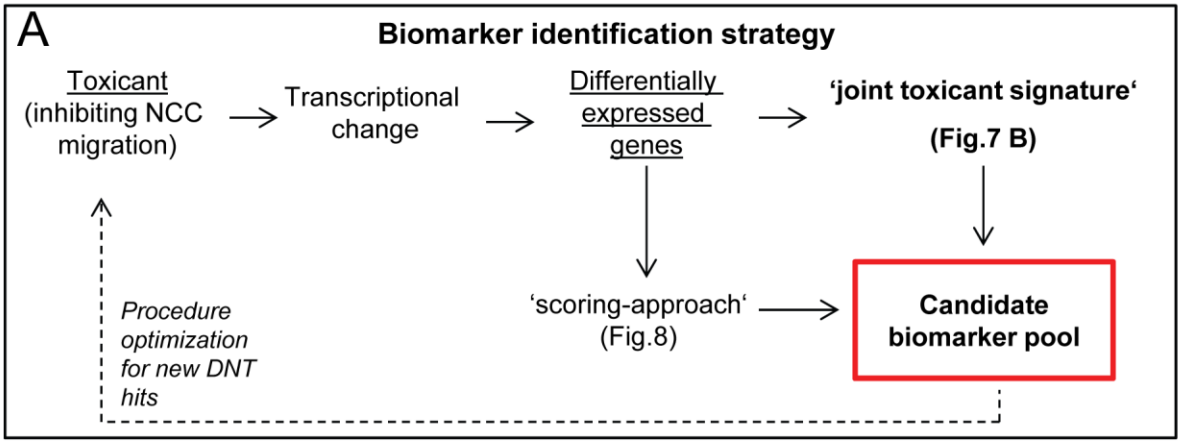
Candidate biomarker identification

Several hits from the MINC assay had been used here for transcriptome analysis, and for a proof-of concept study to blindly predict compounds within a given group. However, for the use of higher numbers of compounds, transcriptome analysis is too expensive and time consuming. Limitation to particularly informative transcripts, here called 'candidate biomarkers', would greatly facilitate alternative approaches by PCR or target-specific next-generation sequencing. Therefore, we used the DEG identified here to define biomarker candidates. We followed two different strategies to select appropriate genes (Fig.7A).

The first approach was based on the concept that DEG found for several compounds would have a broader applicability (and statistical validity). Therefore, we identified all genes, the

expression of which was affected by 4 or more toxicants. We termed this group of 12 genes "joint/general toxicant signature" (Fig.7B). Two of them (EDN1 and SERPINE1) were shared by 5 compounds, while the other 10 were altered by four toxicants. For the assignment of genes to this group, we did not consider the direction of regulation. For instance, triadimefon down-regulated most of the transcripts that were up-regulated by the other compounds, and the only of the consensus gene up-regulated by triadimefon was down-regulated by the other toxicants.

Fig.7 Biomarker identification strategy and joint toxicant signature-derived gene list. A Two different approaches were adopted to identify candidate biomarkers in this study: a "joint toxicant signature"- approach based on the overlap of DEG among the compounds; and a "scoring"- approach based on the evaluation and weighting of each gene as described in Fig.8. B The "joint toxicant signature"- approach led to the identification of 12 genes which were regulated by ≥ 4 toxicants; their expression direction is indicated by red (down-regulation) or green (up-regulation) arrows. At the bottom, information is given on how many % of the DEG regulated by each compound are represented in the list of overlapping DEG. ►



B **GENES REGULATED BY ≥ 4 TOXICANTS (direction indicated)**

Gene symbol	Gene name	As_2O_3	GA	PBDE-99	TDF	TSA	VPA
EDN1	endothelin 1	↑	↑		↓	↑	↑
SERPINE1	serine peptidase inhibitor 1	↑	↑		↓	↑	↑
LOC100507632	uncharacterized locus	↑	↑	↑	↓		
MYOCD	myocardin	↑	↑			↑	↑
EGR3	early growth response 3	↓		↓	↑	↓	
CLDN1	claudin 1	↑		↑	↓	↑	
THBS1	thrombospondin 1		↑	↑		↑	↑
CALD1	caldesmon 1		↑	↓		↑	↑
PAPPA	pregnancy-associated plasma protein		↑		↓	↑	↑
LPPR4	plasticity related gene 1		↑		↓	↑	↑
RAB3B	member RAS oncogene family		↑		↓	↑	↑
ANKRD1	ankyrin repeat domain 1		↑		↓	↑	↑
	% of total DEG	1.8	3.4	1.3	8.7	7.8	20.0

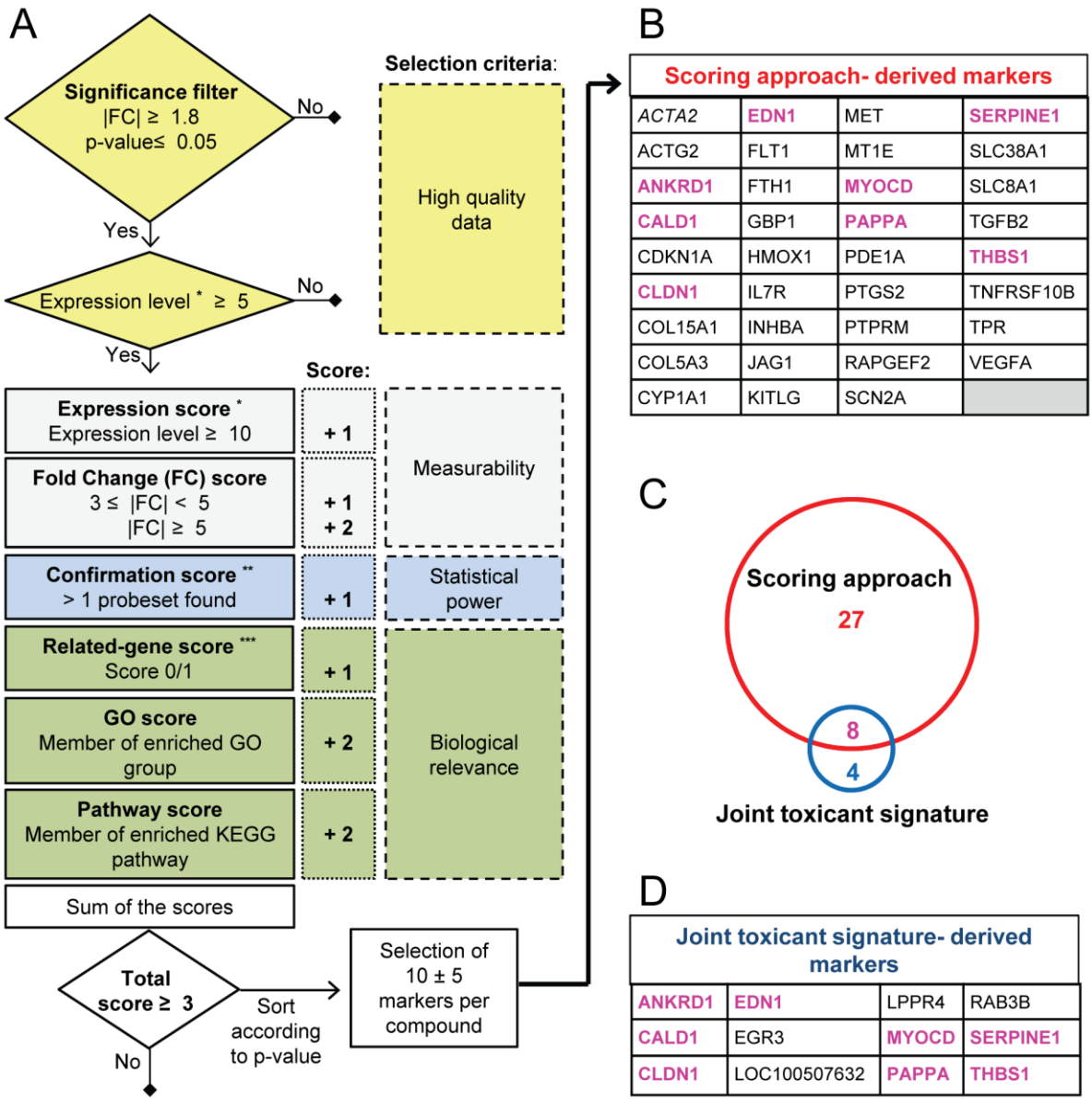
Our second approach to identify a pool of candidate biomarkers followed a "scoring approach" (Fig.7A). We considered each gene that was regulated by one of the six toxicants, and then assigned it a certain importance score according to a filtering and ranking algorithm (Fig.8A). Initially, DEG had to fulfil three minimum requirements to be considered for scoring: a $|FC| \geq 1.8$, a p-value ≤ 0.05 (corrected for false discovery rate), and a control expression level clearly above the microarray noise level, i.e. a fluorescence value of at least 5 after RMA normalization (on a log2 scale ranging from 3 to 15).

Genes were then evaluated based on three further selection criteria (measurability, statistical power and biological relevance) and scored accordingly. For instance, a scoring point was assigned to a given gene, when additional related genes (e.g. belonging to same family or

sharing same receptor) were found in the DEG list. We felt that such genes would have a higher value as biomarkers, as they reflect a regulation mechanism (co-ordinated regulation of functionally-linked genes) going on in treated cells. For the same reason, additional scores were given to genes which were members of enriched GO classes and KEGG pathways.

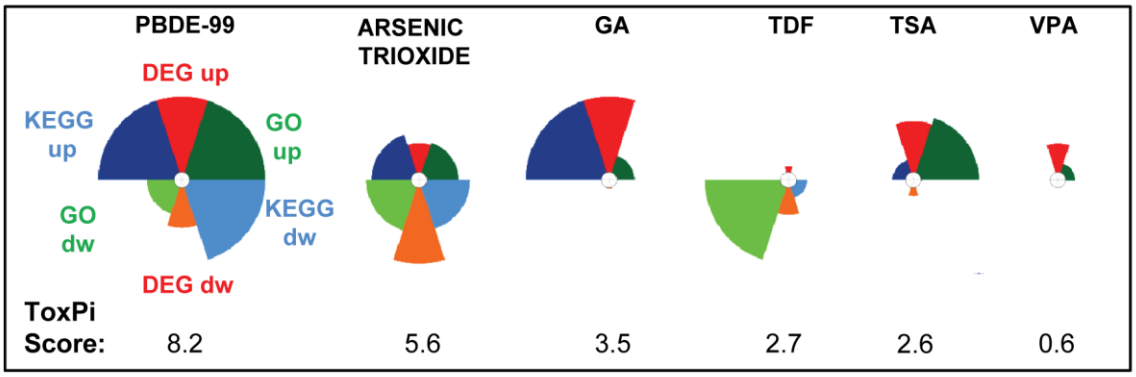
Finally, the genes with a score ≥ 3 were shortlisted, and this pool of markers was saved for potential later use (Supplementary table S5). For extraction of a reasonably small number of candidate biomarkers from this list, 8-12 genes per compound were selected manually (non-mathematical approach) in a way to i) ensure a reasonable balanced across the six toxicants, ii) giving preference to genes with high scores, and iii) favoring low p-values. This resulted in a final group of 35 candidate biomarkers (Fig. 8B). Notably, eight genes of this group overlapped with the biomarkers that were selected based on their role as "joint toxicant signature" (Fig. 8C, D).

Fig.8 Scoring flow chart and overlap between gene markers identified by the scoring and the joint toxicant signature approaches. A The algorithm of the "scoring approach" to identify candidate biomarkers was based on 4 different selection criteria groups: high quality data, measurability, statistical power and biological relevance (color coded). Genes were considered for the scoring, when they fulfilled minimum criteria (significance, expression level). Then scores were given for each gene according to criteria as indicated. Genes with a score ≥ 3 were shortlisted and used for further selection of a final list of candidates with particularly low p-values, and balanced across the six toxicants. **B** This approach led to the selection of 35 candidate biomarkers. Those among them also found by the "joint toxicant signature" approach (Fig. 7) are marked in pink. **C** The overlap of different types of candidate biomarkers is shown in a Venn diagram, **D** and, for clarity reasons, the "joint toxicant" markers are explicitly displayed, with the overlapping ones also marked in pink. * The expression level is the absolute fluorescence value of each probe set (after RMA normalization) which ranged on a scale from 3 to 15. **The "confirmation score" indicated that > 1 PS was regulated (in the same direction) for a given gene. *** The "related-gene score" was applied when, for a given gene, additional related genes were found in the DEG list (belonging to same family or sharing same receptor). ►



Thus, this final part of the study yielded altogether a pool of 39 candidate biomarkers that are of interest for further evaluation in a larger toxicant screen to predict neural crest functional toxicity, or to allow grouping of toxicants according to shared mechanisms/biomarker signatures (Fig. S2). A statistical evaluation of this biomarker set indicated that it showed differences between the toxicants, was clearly related to the cell biology of drug response, and performed much better than randomly selected sets of markers (Fig. 9).

A

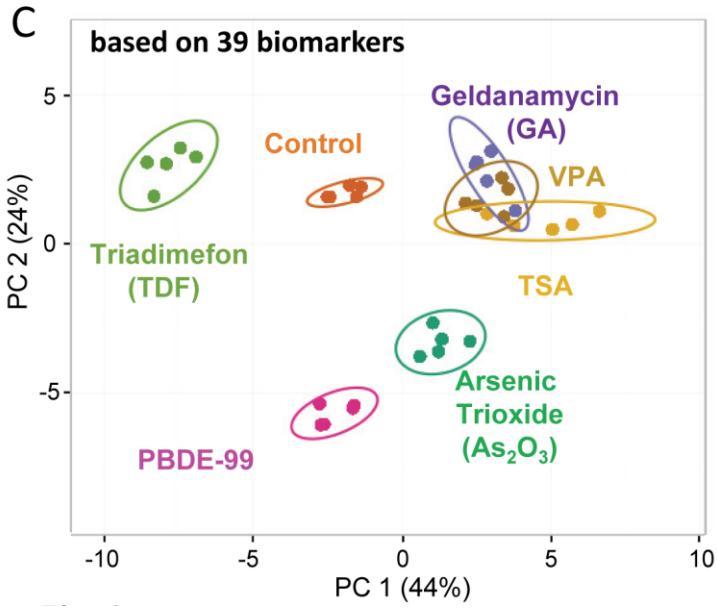


B

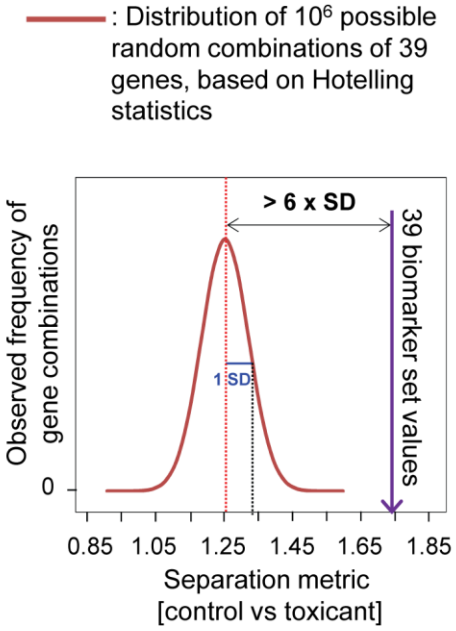
GO term	Found/Total genes	Adjusted p-value
Response to drug (GO:0042493)	9/369	5.1E-05

Found genes	CDKN1A	PTPRM	EDN1	THBS1
TGFB2	CYP1A1	PTGS2	INHBA	SLC8A1

C



D



◄ **Fig.9 Summary of observations and biomarker set characteristics.** A The ToxPi diagrams were built using the ToxPi GUI (Toxicological Priority Index graphical user interface) software tool developed in the ToxCast project (Reif et al. 2013). The numbers of up- and down-regulated DEG, enriched GO classes and KEGG pathways were normalized (range from 0 to 1) to the respective highest values (359 DEG up, 313 DEG dw, 373 GO up, 146 GO dw, 17 KEGG up and 7 KEGG dw) for each category, across all the compounds. Each slice of the diagram contains several types of information: the distance from the center, proportional to the normalized value of the composing that slice; the width (in radians) indicates the relative weight of that slice in the overall ToxPi calculation. In our analysis, double weight was given to the KEGG and GO slices (80% of the total), while DEG slices contributed with 20% of the weight. The calculated ToxPi Score is indicated for each compound (under each diagram) as parameter for potential developmental toxicity. B GO enrichment analysis of 39 biomarker candidates was perfomed using the GOelim algorithm. The most significant GO class is listed together with the genes, which contributed to its enrichment. C A 2D- PCA plot was constructed for the six toxicants, based on the expression data for the 39 candidate biomarkers. The values for each gene were calculated as the median value among the probe sets specific for that particular gene. The analysis was performed using the freely available web-tool ClustVis (http://biit.cs.ut.ee/clustvis). The ellipses show the 95%-CI. D From the pool of differentially expressed genes in NCC treated with toxicants, 39 samples (= genes) were drawn at random 1 million times. The separation strength (extent of separation of controls from the group of toxicants of these samples) was calculated (x axis; parameter ranges from 0 – 2.45), and the distribution was plotted. For comparison, the separation strength of the here selected biomarker set is indicated.

DISCUSSION

In the present study, we analyzed transcriptome changes triggered by six hits of the ESNATS test battery. The compounds examined here had been identified as potential developmental toxicants because of their ability to inhibit NCC migration at non-cytotoxic concentration, and the working procedure of the ESNATS test battery (Zimmer et al. 2014, Dreser et al. 2015) required further characterization, including transcriptome mapping. The transcriptome data obtained here allowed the separation of the toxicants on a principal component map. Moreover, they allowed the assignment of unknown samples to the known set of compounds, based on a support vector machine classification algorithm. Finally, the types of biological changes, indicated by the altered transcriptomes, were visualized in diagrams focusing on altered superordinate biological processes.

In this study, three questions were asked to explore the usefulness of transcriptome changes for a toxicant test battery. The first question addresses a key issue of toxicogenomics in general: can transcriptome profiles be used to identify hazardous compounds? Previous studies have shown that this is a highly demanding challenge, and that a lot needs to be learned on better design and evaluation of transcriptomics studies for such purposes (Thomas et al. 2013, Grinberg et al. 2014, Bourdon-Lacombe et al. 2015, El-Hachem et al. 2015). Thus, this large issue needs to be approached in smaller steps. A main issue that became evident in our study is the large heterogeneity in transcriptome responses triggered by compounds that all affect the same functional endpoint (NC migration). One reason may be that migration of cells is such a complex endpoint, that compounds with very diverse modes of action can affect it, i.e. very different molecular initiating events will affect the same adverse outcome. This precludes any simple type of analysis, such as the identification of a joint gene derangement pattern across all compounds. However, on the positive side, toxicants known to share a known mode of action, such as VPA and TSA, also showed a similar transcriptome response. This implies that it may be possible to use transcriptome responses for toxicological grouping of compounds, i.e. that an unknown compound may be assigned to a group of already known toxicants based on shared transcriptome profiles. This would be an expansion of the read-across approach, away from structure-based algorithms to

the incorporation of variable biological information (Low et al. 2013, Patlewicz et al. 2014, Bal-Price et al. 2015, Berggren et al. 2015). The basis for this was explored here in a small pilot study to see whether an unknown chemical could be assigned within a small, but diverse group of toxicants to its most related compound. The blind assignment of six compounds, to the six known compounds worked surprisingly well, given the fact that only three microarrays per compound were used to build the classifier.

Given the situation that *in vivo* testing for developmental toxicity, and especially developmental neurotoxicity has serious issues concerning species extrapolation and sensitivity (van Thriel et al. 2012, Smirnova et al. 2014), there is a large need to consider new approaches for risk assessment, or at least filtering of relevant compounds for further testing. One such approach is the consideration of key biological processes, such as neural crest migration (Bal-Price et al. 2012, Kadereit et al. 2012) that can be tested in appropriate *in vitro* systems. Compounds that affect such key biological processes could be further investigated for transcriptome changes in the respective system, and this information may then be used to better define the mode of action, but also to read across to other compounds with known in vivo toxicities. Similar approaches have been tried with promising results across largely different systems, not only based on human cells, but e.g. also using model organisms such as zebra fish (Hermsen et al. 2013). A more radical future way would be to compare toxicants based on their transcriptome profiles in well-characterized test systems rather than on their effects in animals. One condition for this is a high level of test system characterization and quality control (Leist et al. 2010, Crofton et al. 2011, Leist et al. 2012c, Coecke et al. 2005). Beyond this, a large knowledge base needs to be collected on whether toxicogenomics signatures really can predict toxicity in a given test system. Until now, only few functional developmental toxicity assays have been evaluated in this direction, and it is unclear whether e.g. neurite growth or neurite degeneration assays (Volbracht et al. 1999, Stiegler et al. 2011, Krug et al. 2014) fulfil such conditions. Also for general neurodegeneration assays, there is still little information, as transcriptome information needs to be obtained during a phase prior to cell death. Moreover, if generated from complex models, such as co-cultures (Alepee et al. 2014, Efremova et al. 2015), it needs to be distinguished from the generalized inflammatory response (Falsig et al. 2004b, Falsig et al. 2006). If such conditions are fulfilled, as recent studies have shown in the case of damage

triggered by MPP+ in human neurons, then toxicogenomics information can yield surprising new information and mechanistic insight (Krug et al. 2014) and a related *in vivo* study also revealed hitherto unsuspected genetic regulations (Maertens et al. 2015, Rahnenfuhrer and Leist 2015).

A second question addressed in this study is how primary transcriptome information, i.e. long lists of differentially-expressed genes can be reduced to a format that is easier to handle and that can be used for toxicological purposes. Classical toxicology has worked well with semi-quantitative information that is judged for its significance by experts, and that requires careful consideration of many modulatory and circumstantial factors. For instance, staining of histological slides may indicate cellular changes characterized by eosinophilia, hypertrophy and lipid droplet accumulation, and experts have to decide on the type and level of hazard this indicates or whether this is rather an insignificant or adaptive change. A similar system of several dozen to hundreds of categories is required for toxicogenomics, while dealing with 25,000 individual genes will not be possible. One composite measure is the number of DEG. It appears evident that the information content of such an endpoint is relatively low, although there is a high likelihood that compounds that do not de-regulate any gene are relatively harmless, and chemicals that de-regulate very large numbers of genes may be problematic. More information may be obtained from the examination of biologically-linked gene-networks, i.e. genes belonging to one GO group of KEGG pathway. An increasing number of overrepresented GO terms/KEGG pathways (or other biological motives) amongst the DEG would indicate a specific regulation of genes belonging to a certain cell function as opposed to random gene regulations. A summary of the changes across all study compounds can be obtained from such measures very quickly, e.g. in the form of ToxPi diagrams (Reif et al. 2013), or the associated ToxPi score (Figure 9A).

Even more information may be gained by looking exactly into which pathways (KEGG) are regulated or which specific groups the regulated genes belong to. However, a compromise has to be found between the detail of information and the simplicity of an initial toxicological statement. The chosen solution was to show only numbers of superordinate biological processes together with few examples and a rough classification. In addition to this coarse-

grained initial information layer, supplementary information can then answer details, where required.

Even for such standard approaches, some decisions have to be taken. They include the statistical criteria used to define the DEG, but also involve issues such as the separate or combined treatment of up- and downregulated genes. We decided here to identify overrepresented GO and KEGG separately in the two groups of genes. With this procedure we followed an established routine that was chosen to facilitate comparisons amongst different conditions, such as different concentrations of one compound, or one concentration of different compounds, or different exposure times of one compound (Krug et al. 2013c, Balmer et al. 2014, Waldmann et al. 2014, Rempel et al. 2015). When defining intersection between conditions, we felt that it is important to consider the direction of regulation of a gene, and also to mine the genes accordingly for overrepresented biological themes. A different situation is encountered, when no such comparisons are intended, and when different directions of regulation e.g. within a GO group make biological sense (e.g. 'positive regulation of apoptosis' involves up-regulation of apoptosis inducers (BCL-2, Caspases) and downregulation of inhibitors (BAX, IAPs, HSP70) (Latta et al. 2000, Gerhardt et al. 2001, Hansson et al. 2003). For this reason, GO analysis to identify superordinate biological processes was performed differently than for the general data exploration.

The third major question was directed to the identification of biomarker candidates that would allow a simplified approach, compared to whole genome transcript profiling. The term 'biomarker' has a wide range of implications and uses, and therefore requires some definition in the context of our study. Very strict definitions are found in the field of predictive medicine, or in toxicology in the form of a biomarker of toxicity (BoT) that is required to have a high predictive value and to show some causal relationship with the adverse outcome (Blaauboer et al. 2012) or a toxicity pathway (Leist et al. 2008b). At the other end of the spectrum, biomarkers are simply seen as any endpoint that changes in a test system upon exposure to a test compound. An approach somewhere in between these extremes is to define biomarkers as pre-selected endpoints with a certain information value concerning the study purpose (e.g. biomarker of exposure or biomarker as part of a predictive gene signature), but not necessarily linked to the mechanism of action of a compound. A good example for this type of approach is the GARD assay for skin sensitization (Johansson et al. 2013, Johansson

et al. 2014), in which first whole transcriptome data were obtained on skin sensitizers and negative controls, and then a statistics-based algorithm was applied to select the set of genes (biomarkers) that was most useful as classifier. Similar approaches have been taken to add biological information for read-across, for instance in the SEURAT-1 research project on prediction of cosmetics toxicity (Gocht et al. 2015) or based on the TG-GATES transcriptome data (Low et al. 2013). We provided here a basis for such a latter approach by selecting a small number of genes from all the DEG of the study.

These were termed here 'candidate biomarkers' as more work is required to qualify them as 'real' biomarkers at a confidence level of e.g. the GARD assay. An immediate usefulness is suggested by analysis of overrepresented GO terms amongst the 39 selected genes: the most significant enriched GO class was "response to drug" (Figure 9B). Moreover, the chosen set of genes provided a good basis for separation of the study compounds from control cells (Figure 9C, Fig.S3, Supplementary table S6). To obtain an idea on the performance of the selected biomarkers, relative to random sets of biomarkers, we use here a relatively simple and transparent approach to define 'separation strength', and we explored this separation strength of our biomarker set, when compared to one million sets, randomly chosen from the pool of all regulated probe sets of this study. The 39 biomarker set was by far superior to random sampling (99-100th percentile of the distribution, with 6 standard deviations distance to the means; $p < 10^{-6}$) (Figure 9D), which confirms the overall usefulness of our selection approach. We are aware of the relatively extreme assumptions we had to make (e.g. statistical independence of the endpoints), and that separation strength may be defined in many other ways. However, we hope that this initial attempt will trigger more work in this area, allowing to establish criteria for separation strength of a set of biomarkers (as opposed to a classifier formula). In the future, a consensus will then need to be reached on statistical approaches to judge the quality of a given biomarker set relative to randomly selected sets.

At present, the explorative approaches discussed here qualify our small set of candidate biomarkers for further exploration, and possible substitution of the full microarray approach by a cheaper and faster technology. Three major steps will have to be taken towards this objective in the future: i) the marker gens would require confirmation by the alternative analytical technology chosen (e.g. PCR) on the set of study compounds used here; ii) then, they would need to be tested for their usefulness on another set of compounds (Leist et al.

2010, Crofton et al. 2011); iii) and finally, the time and concentration-relationship of marker changes would need to be correlated with the functional toxicity (inhibited migration) in the MINC assay.

In summary, this study added information on ESNATS test battery hits and provided a case study on the predictive value of toxicogenomics by showing that chemicals can be predicted, based on their transcriptome changes, at least within a smaller group. Extension to more compounds will be necessary. Moreover, tools were developed to visualize transcriptome changes and to provide at least semi-quantitative data on the extent and type of transcriptome derangement. Finally, two different approaches were combined to pre-select biomarkers that are still able to separate the compounds and that will require further evaluation for their application in predictive toxicology.

E. Results. Manuscript 3

Impairment of human neural crest cell migration by prolonged exposure to interferon-beta

Giorgia Pallocca[1], Johanna Nyfeller[1], Xenia Dolde[1], Marianna Grinberg[2], Tanja Waldmann[1], Jörg Rahnenführer[2], Agapios Sachinidis[3] and Marcel Leist[1]

Affiliations:

[1] Department of in vitro toxicology and biomedicine, University of Konstanz, 78457 Konstanz, Germany;

[2] Department of statistics, TU Dortmund, 44139 Dortmund, Germany;

[3] Center of physiology and pathophysiology, institute of neurophysiology, University of Cologne, 50931 Cologne, Germany.

ABSTRACT

Human interferon-beta (IFNβ) was identified as a hit in the 'migration inhibition of neural crest' (MINC) assay during a screen for potential developmental toxicants amongst clinically used drugs. Because of the pronounced species-specificity of interferons, little preclinical data on the developmental toxicity hazard of IFNβ is available. We therefore followed up on the initial scratch assay screen findings, to study more closely the effects of IFNβ on human neural crest cell (NCC) function. The differentiation state and overall phenotype of NCC was maintained during an exposure to IFNβ for 48 h under migration test conditions, although an archetypical inflammatory and anti-viral response was triggered. Concentration-response studies showed that IFNβ inhibited NCC migration at concentrations as low as 20 pM. This held true also, when the test was performed in the presence of the mitotic inhibitor cytosine arabinoside to exclude any artefacts due to cytostatic interferon activity. Moreover, two other functional assays confirmed that pM concentration of IFNβ reduced the motility of NCC, while other interferons were considerably less potent. The activation of JAK kinase by IFNβ, as suggested by bioinformatics analysis of the transcriptome changes, was confirmed by biochemical methods. The degree and duration of pathway activation correlated with the extent of migration inhibition, and pharmacological block of this signaling pathway before, or up to 6 h after exposure to the cytokine prevented the effects of IFNβ on migration. Thus, the reduction of vital functions of human NCC is a hitherto unknown potential hazard of endogenous or pharmacologically applied interferons.

INTRODUCTION

Novel human cell-based *in vitro* models have been developed to explore the adverse effects of human-specific drugs and biologics on fundamental developmental processes. Any kind of external interference which alters the tempo-spatial organization of development can lead to developmental defects. This fine tuning shows differences between different species and species-specific antigenic properties and receptor interactions prevent the characterization of many side effects of cytokines, antibodies and blood factors in animal models. Therefore, human cell-based test systems play a crucial role in drug development

The approach of using *in vitro* tests is based on the strategy that a battery of assays may be assembled that covers the majority of fundamental developmental processes (Leist et al. 2014, Schmidt et al. 2016, Smirnova et al. 2014). These include for instance neural differentiation (Balmer et al. 2012), neurite outgrowth (Krug et al. 2013b), gliogenesis (Fritsche et al. 2005), myelination (Zurich et al. 2000, Zurich et al. 2002) or synaptogenesis (Harrill et al. 2011).

In this context, the migration inhibition of neural crest cells (MINC) assay was developed as the first method to test effects of toxicants on the function of the neural crest (NC) during development (Zimmer et al. 2012, Dreser et al. 2015). The importance of this developmental stage justified the inclusion of the MINC assay in the *in vitro* test battery of the European research consortium ESNATS (Embryonic Stem cell-based Novel Alternative Testing Strategies). The NC is a multipotent migratory cell population that emerges from the dorsal aspect of the neural tube in the early phases of development. NC cells (NCC) are capable of long range migration, which is finely regulated in the embryo. During migration, and at the final destination, NCC give rise to a multitude of different cell types, supporting the formation of cartilage and bone of the face, but also peripheral and enteric neurons, melanocytes and some cardiomyocytes (Huang and Saint-Jeannet 2004).

A large percentage of developmental disorders (e.g. congenital heart defects, oro-facial clefts, Hirschsprung's disease) are caused by NCC deficits (Simões-Costa and Bronner 2013), and these often correlate with neural tube defects. Such alterations can be induced by

genetic factors (Lee et al. 2009) or exposure to pharmaceuticals (e.g. valproic acid) (Fuller et al. 2002) and pesticides (e.g. triadimefon) (Menegola et al. 2000).

The MINC assay has been used to screen and evaluate the effects of many different compounds on neural crest function (Zimmer et al. 2012, Zimmer et al. 2014, Dreser et al. 2015, Nyffeler et al. 2016). In addition to known environmental toxicants, the screening identified some drug-like compounds as hits, and these included the type I interferon IFNβ (Zimmer et al. 2014). Interferons have a crucial role in the antiviral cell response, and in modulating several functions of the immune systems. Type I interferons are ubiquitously expressed and comprise IFNα and IFNβ. All type I IFNs bind the interferon alpha and beta receptor (IFNAR) subunits and produce structurally highly similar receptor-ligand complexes (Schreiber and Piehler 2015).

The hallmark of IFN signaling is the phosphorylation of STATs, which results in the formation of a pSTAT1/pSTAT2 heterodimer. This forms, in complex with IFN regulatory factor 9 (IRF9), the transcription factor complex named IFN-stimulated gene factor 3 (ISGF3). The latter promotes transcription of various interferon stimulated genes (ISG) (Schreiber and Piehler 2015).

The interferon family is clinically used for different purposes. The type II interferon IFNγ binds to different receptors and has pronounced inflammatory and immune-stimulating properties (Schroder et al. 2004). While IFNα is mainly used for treatment of certain types of leukemia and hepatitis virus infections, the immunomodulatory drug IFNβ has proven effective in the treatment of relapsing-remitting multiple sclerosis (MS) (Dhib-Jalbut and Marks 2010). The mode of action of the latter cytokine is still not fully understood, but it induces an anti-inflammatory cytokine shift and prevents T-cell adhesion and extravasation across the blood-brain barrier.

The developmental toxicity of IFNβ has been evaluated in cynomolgus monkeys, where an increased incidence of both abortions and stillbirths was observed (FDA 1999). Furthermore, epidemiological studies indicated that exposure to IFNβ in pregnancy is associated with lower mean birth weight, and preterm birth (Amato et al. 2010). For these reasons, women with MS are typically advised to discontinue the treatment before conceiving. IFNβ is classified as risk class C drug by the FDA, indicating that “animal reproduction studies have

shown an adverse effect on the fetus and there are no adequate and well controlled studies in humans, but potential benefits may warrant the use of the drug in pregnant women despite the potential risks" (Lu et al. 2012, Pozzilli and Pugliatti 2015).

In the present study, we used the MINC assay to further characterize a potential developmental toxicity hazard of IFNβ. The adverse effects of IFNβ on NCC migration and proliferation were studied in three different assays. The specificity of the effects of IFNβ on NCC migration appeared high, as cells did not show other obvious functional or structural defects. Analysis of transcriptome changes suggested a role of altered JAK-STAT signaling in toxicity. This was confirmed by detailed measurements of interferon effects on signaling and cell function in the presence of specific kinase inhibitors.

MATERIALS AND METHODS

Cell culture and neural crest differentiation

The reporter human embryonic stem cell line H9-Dll1 (GFP under Dll1 promoter) was provided by Mark Tomishima from the Memorial Sloan Kettering Cancer Centre (MSKCC, NY, USA). Import of the cells and all experiments were carried out according to German legislation under the license number 1710-79-1-4-27 of the Robert-Koch Institute.

H9-Dll1 cells were maintained on mouse embryonic fibroblasts in DMEM/F12 (Gibco™, Carlsbad, CA, USA) medium containing 20% of serum replacement, HEPES (Gibco™), L-gluthamine (Glutamax, Gibco™), non-essential amino acids (MEM NEAA, Gibco™), beta-mercaptoethanol (Gibco™) and basic fibroblast growth factor (10 ng/ml, Invitrogen™, Carlsbad, CA, USA). Differentiation of hESC into NCC was initiated on mitomycin C treated murine bone-marrow derived stromal MS5 cells and continued as described in (Zimmer et al. 2012). The human breast adenocarcinoma cell line MDA-MB-231 was maintained in DMEM Glutamax (Gibco™) containing 10% FBS and 1% penicillin/streptomycin.

Chemical exposure during NCC migration

NCC were exposed for 48 hours to non-cytotoxic concentration of different interferons (recombinant human IFN-beta 1a, recombinant human IFN-alpha 2a, recombinant human IFN-gamma) in N2 medium containing EGF (20 ng/ml) and FGF2 (20 ng/ml) (all from R&D Systems GmbH®, Minneapolis, MN, USA). In the indicated cases, cytosine arabinoside (AraC) or JAK inhibitors (ruxolitinib and tofacitinib, Selleckchem©, Munich, Germany) were added. The migration assay was performed as described in Nyffeler et al. 2016. Briefly, NCC cells were seeded (95 000 cells/cm2) in 96-well plates (Corning©, NY, USA) previously coated with 10 µg/ml poly-L-ornithine in 100 µl phosphate buffered saline (PBS) (GE Healthcare Bio-Sciences©, Pittsburgh, PA, USA) and 1 µg/ml fibronectin and 1 µg/ml laminin (Sigma-Aldrich©, St. Louis, MO, USA) in 100 µl PBS. Cells were seeded in presence of silicon stoppers (Platypus Technologies, Madison, WI, USA) to create a circular cell-free area. One day after seeding, migration into the cell-free area was initiated by manual

removal of the stoppers and the medium was replaced with medium containing the test compounds. After 48 h, NCC were stained with H-33342 and 533 nM calcein-AM (Sigma-Aldrich) and imaged 30 min later on a high content imaging microscope (Cellomics ArrayScanVTI, Thermo Fischer©, Boston, MA, USA). Viability was defined as the number of H-33342 and calcein double-positive cells, as determined by an automated algorithm described earlier (Stiegler et al. 2011, Krug et al. 2013c). For quantification of migration, a software tool (freely accessible at http://invitrotox.uni-konstanz.de/) was developed to estimate the most likely position of the previously cell-free area (covered by the silicon stopper), to set thresholds for color intensity for both dyes, and to count the number of H-33342 and calcein double-positive cells in the region of interest.

EdU staining

Cells were treated for 48 h with 10 µM 5-ethynyl-deoxyuridine (EdU). Proliferating cells incorporate EdU and can be detected using a click reaction procedure, as described in manufacturer's protocol (EdU-Click 555, PanaTecs, Heilbronn, Germany). Images were acquired using a high content imaging microscope (Cell Insight Personal Imager, Thermo Fisher©). Proliferation was defined as the percentage of EdU positive nuclei among all H-33342 positive nuclei.

Immunofluorescence staining

Cells were seeded in Lumox® multiwell plates (95 000 cells/cm²). After 42-h treatment, the middle of the well was scratched with a pipette tip to create a cell-free space in the well and cells were allowed to migrate for 6 h in presence of the treatment. Finally, the cells were fixed in 4% PFA for immunofluorescence staining. Cells were permeabilized for 10 min in 0.2% Triton and blocked with 10% FBS for 1 h. Afterwards, cells were incubated overnight with primary antibody (GM130, Abcam©, Cambridge, UK; TOM20, Santa Cruz Biotechnology™, Santa Cruz, CA, USA) or Alexa Fluor 555 phalloidin diluted in 4% FBS. After a washing step, the secondary antibody was applied for 1 h. Cell nuclei were finally stained with H-33342. Images were acquired using a point laser scanning confocal microscope Zeiss LSM 700 (Zeiss©, Oberkochen, Germany).

Western blot

Cells were seeded in 6-well plates (50 000 cells/cm²); after 24 h, cells were treated for the indicated times. Cells were then harvested in Laemmli buffer, boiled for 5 min at 95°C and purified with the NucleoSpin Filters (Macherey-Nagel GmbH, Düren, Germany). Samples were run on SDS-PAGE. Transfer on nitrocellulose membranes was performed by using iBlot™ 2 Dry Blotting System (Invitrogen). Membranes were then incubated in 5% milk in T-TBS for 1 h and overnight with the primary antibody in 5% BSA in T-TBS at 4°C (p-STAT1 Y701, Cell signaling technology; GAPDH, Invitrogen). After washing steps, membranes were incubated with secondary antibody conjugated with horseradish peroxidase (GE Healthcare Bio-Sciences©) for 1 h at room temperature. Signal was finally detected using Pierce ECL wester blotting substrate (Thermo Scientific© Boston, MA, USA) and imaged with a Fusion-SL 3500 WL device and Fusion software (Bio-Rad™, Hercules, CA, USA).

NFkB translocation

Cells were stimulated with the complete cytokine mix (CM) containing 10 ng/ml tumor necrosis factor α (TNFα), 10 ng/ml interleukin 1β (IL-1 β), and 20 ng/mL IFNγ (R&D Systems, Wiesbaden, Germany) or with 500 pM IFNβ for 1 h. For NFkB measurement, cells were fixed, permeabilized and stained with NFkB antibody (Santa Cruz Biotechnology™). NFkB translocation was measured with the high throughput device CellInsight TM CX5 High Content Screening (Thermo Scientific©) using the nuclear translocation algorithm as described previously Henn et al. 2011.

Cell tracking

Live cell imaging was performed using a Axio Observer.Z1 microscope (Zeiss©) equipped with an incubation system (37°C, 5% CO_2) and the software Zen2. Cells were imaged for a period of 48 h, taking phase contrast pictures every 15 min with a 5x objective. Images of the last 30 h migration period were loaded in Fiji ImageJ and manually tracked with the plugin "Manual track" (Schindelin et al. 2012). The tracked positions were then loaded into the freely available "Chemotaxis and Migration Tool" (Ibidi) to calculate the total distance and the cell speed as well as to create the track pictures.

Transwell assay

Cells were treated for 42 h and then detached by using Accutase (Corning©), counted and seeded (50 000 cells/transwell) in upper chamber of previously coated Transwell Permeable Support plates (0.8 µm polycarbonate membrane, Costar, Corning©) in normal medium with addition of IFNβ (500 pM) or respective control. Ten percent FBS in normal medium was added in the lower chamber of the transwell. After 6 h of incubation at 37°C, cells at the upper side of the membrane were removed and the cells attached to the lower side of the membrane were fixed and stained with crystal violet for 30 min, then washed in current water and let dry. Cells were imaged with light microscope Axio Observer.Z1 microscope (Zeiss©) using the software PALM RoboSoftware (4 fields per condition, 20x magnification) and manually counted in ImageJ.

Affymetrix gene chip analysis

Samples of $\geq 5 \times 10^6$ cells were collected using RNA protect reagent from Qiagen. The RNA was quantified using a NanoDrop N-1000 spectrophotometer (NanoDrop, Wilmington, DE, USA), and the integrity of RNA was confirmed with a standard sense automated gel electrophoresis system (Experion, Bio-Rad™, Hercules, CA, USA). Analysis was then performed as described earlier Krug et al. 2013c using Affymetrix chip-based DNA microarray (Human genome U133 plus 2.0 arrays) with all standard quality control procedures.

Biostatistics

The microarray data analysis (extrapolation and normalization of the array sets) was performed using the statistical programming language R (version 3.1.1) as described in Waldmann et al. 2014. For the normalization of the entire set of Affymetrix gene expression arrays, the Extrapolation Strategy (RMA+) algorithm Harbron et al. 2007 was used that applies background correction, $\log_2$ transformation, quantile normalization, and a linear model fit to the normalized data in order to obtain a value for each probe set (PS) on each array. As reference, the normalization parameters obtained in earlier analyzes (Krug et al. 2013c) were used. After normalization, the difference between gene expression and corresponding controls was calculated (paired design). Differential expression was calculated using the R package ‘limma’ (Smyth et al. 2005). Here, the combined information of the

complete set of genes is used by an empirical Bayes adjustment of the variance estimates of single genes. This form of a moderated t-test is abbreviated here as 'Limma t test'. The resulting p-values were multiplicity-adjusted to control the false discovery rate (FDR) by the Benjamini-Hochberg procedure Benjamini 1995. As a result, for each compound a gene list was obtained, with corresponding estimates for log fold changes and p-values of the Limma t test (unadjusted and FDR-adjusted).

Transcripts with FDR adjusted p-values of ≤ 0.05 and fold change values of ≥ 1.8 or ≤ 0.55 were considered significantly deregulated and defined as differential expressed genes (DEG).

Data display: heat map and principal component analysis

The software R (version 3.1.1), was used for all calculations and display of principal component analysis (PCA) and heatmaps. PCA plots were used to visualize expression data in two dimensions, representing the first two principal components. The percentages of the variances covered are indicated in the figures.

Gene ontology (GO) and KEGG pathway enrichment analysis

The gene ontology group enrichment was performed using R (version 3.1.1) with the topGO package (Alexa et al. 2006) using Fisher's exact test, and only results from the biological process ontology were kept. Here again the resulting p-values were corrected for multiple testing by the method of Benjamini−Hochberg (Benjamini 1995).

The KEGG pathway analysis was performed using the R package "hgu133plus2.db" Carlson 2015. Probesets were mapped to the identifiers used by KEGG for pathways in which the genes represented by the probesets are involved. The enrichment was then performed analogous to the gene ontology group enrichment using Fisher's exact test.

Up- and down-regulated differentially expressed genes were analyzed separately for each treatment. Only GO classes and KEGG pathways with a BH (Benjamini-Hochberg)-adj. p-value ≤ 0.05 were considered significant.

GO super-ordinate classes distribution

Enriched GOs were then assigned to superordinate cell biological processes as already described in Waldmann et al. 2014 and distributed in six classes: migration/adhesion, metabolism, differentiation, signaling, stress response, and others. The migration class includes migration and adhesion related-GO classes; stress response class includes cell death-, extra-cellular stressor-, inflammation/immunity-related GO classes; signaling class consist of cell receptor activity-, second messenger (cAMP, cGMP, Ca2+) metabolism-, kinase modification-related GO classes. Metabolism class comprise all GO classes covering metabolism activity; differentiation class includes cell differentiation related- GO classes; and; "other" class covers the others not otherwise classified GO classes.

qPCR

Cells were seeded in 6-well plates (50 000 cells/cm²); after 24 h, cells were treated for 48 h. Cells were then harvested and lysed in PeqGOLD Trifast™ (PEQlab©, Erlangen, Germany). Total RNA was isolated by phenol-chloroform extraction. Reverse transcription was performed using 1 µg RNA and the i-Script™ Reverse Transcription Supermix (Bio-Rad™) according to the manufacturer's protocol. Quantitative real-time PCR was performed with SsoFast EvaGreen Supermix (Bio-Rad™) using a CFX96 Real-Time PCR Detection System (Bio-Rad™). The following transcripts were analyzed: STAT1 (sense primer CCTCCTGTCACAGCTGGATGATC; antisense primer GCCTGATTAAATCTCTGGGCGTT); STAT2 (sense primer GAAGTCGCGACCAGAGCCAT; antisense primer CAGCTTCCTGCCAGTTCTGG); IRF1 (sense primer ACCTCTGCCTTCTTCCCTCTTC; antisense primer ATCCAGATGAGCCCCGGGATTT); IRF9 (sense primer GCGACAGCCTGGACAGCAAC; antisense primer GGCAGCATCCTGGTCCTCCC); SP100 (sense primer TGACAGGCTGCTCTATGACATTGTA; antisense primer GTACACCACTCTCTGTACAGGGACC); TLR3 (sense primer TTTCGAGAGTGCCGTCTATTTG; antisense primer GGCTAACAGTGCACTTGGTGGT); CXCL10 (sense primer ACCTCCAGTCTCAGCACCATGAA; antisense primer ACGTGGACAAAATTGGCTTGCAG); CXCL11 (sense primer

CTTGGCTGTGATATTGTGTGCTACA; antisense primer GGGATTTAGGCATCGTTGTCCTTTA). The cycle threshold values were determined using the Bio-Rad CFX Manager Software v2.0 (Bio-Rad™). Results were analyzed using the ΔCt method Livak and Schmittgen 2001.

RESULTS

Specific effects of IFNβ on human neural crest cell (NCC) migration

In a previous study (Zimmer et al. 2014), IFNβ was found as a positive hit in a screen of drugs for potential developmental toxicity in human NCC. To confirm the effects of the cytokine IFNβ on human NCC migration, a new automated and operator-independent test method (Nyffeler et al. 2016) was used. NCC, generated from pluripotent stem cells, were allowed to migrate in the presence or absence of IFNβ for 48 h (Fig. 1A). Picomolar concentrations of IFNβ induced a significant inhibition of migration (Fig. 1B, left graph), as indicated by the reduced number of cells found in the migration zone (Fig. 1C). The migration capacity was reduced to 75% at 20 pM (Fig. 1B, left graph), but IFNβ also slightly affected the cell viability endpoint at this concentration (reduction of the cell number by > 10%). As no dead or dying cells were observed at any time point, the apparent 'reduction of viability' was most likely due to the known cytostatic effect of the cytokine (Bekisz et al. 2010).

For interpretation of the data from the MINC assay used here, it is important background information that NCC still proliferate during the assay. This property of the test system explains why cytostatic drugs can affect the number of cells at the end of the assay (Nyffeler et al. 2016). Addition of the mitosis inhibitor cytosine arabinoside (AraC, 10 µM) blocks all proliferation (Fig. 1D), and under such test conditions, the specific effect of cytostatic drugs on migration can be quantified without interference of cell cycle effects. When the assay was repeated in the presence of AraC, IFNβ did not affect the viability endpoint at all, but it still triggered a significant inhibition of migration at 25 pM, and the threshold of 25% reduction (Zimmer et al. 2012, Zimmer et al. 2014, Nyffeler et al. 2016) was reached at 50 pM (Fig. 1B, right graph). Thus, data from different test formats corroborated the earlier finding that pM concentrations of IFNβ reduced the migration of NCC, independent of any potential effects of the drug on NCC proliferation.

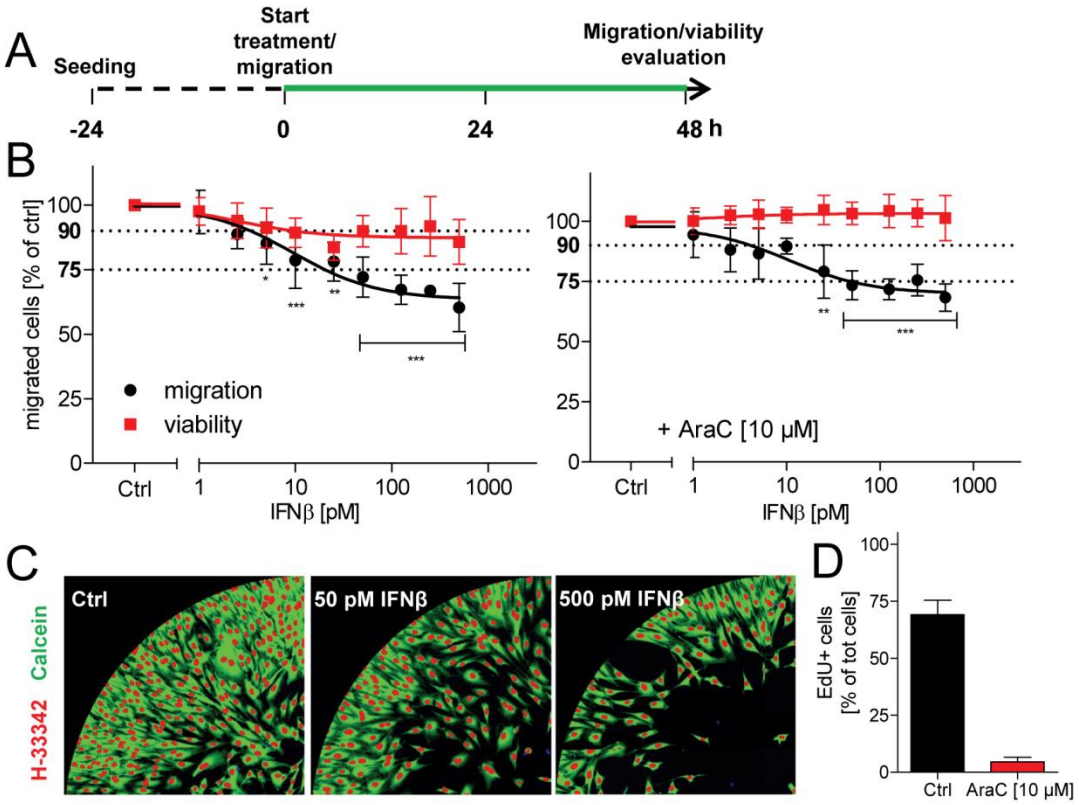
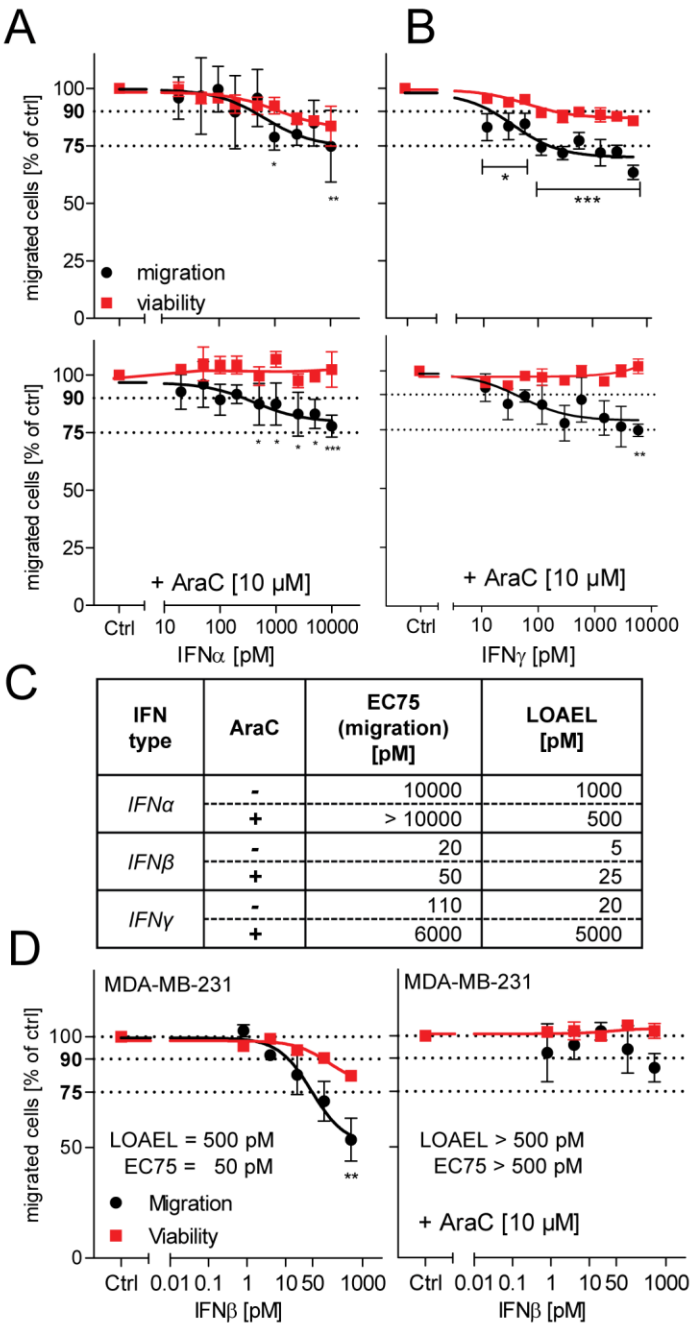


Fig.1 Impaired migration of NCC in the presence of IFNβ A NCC were allowed to attach and recover for 24 h. Then, migration was started and, 48 h later, the number of migrated cells and the viability of the cell population was quantified. In standard experiments, NCC were exposed to IFNβ (marked in green) for the entire migration period. B NCC were exposed to IFNβ at the indicated concentrations. Cell viability and the number of migrated cells are expressed relative to control cells (Ctrl, 0.1% BSA in PBS). In one series of experiments (right graph), the cell culture medium used for all conditions was supplemented with the mitosis inhibitor cytosine arabinoside (AraC, 10 µM). Data are from ≥ 3 independent experiments. Error bars indicate standard deviations (SD). Statistics was performed for each endpoint by ANOVA, followed by Dunnet's post-hoc test (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$). Viability was considered to be impaired when it dropped below 90%; migration was considered to be impaired below 75%, compared to control (dotted lines at 75 and 90% are indicated for visual support). C Representative pictures of different migration assay exposure scenarios, taken at time 48 h. Nuclei are depicted in red (H-33342), while viable cells are shown in green (calcein). D NCC were exposed to culture medium supplemented with 5-ethynyl-2′-deoxyuridine (EdU, 10 µM) and they were treated with AraC (10 µM) or the respective control (Ctrl) for 48 h. Then cells were stained with H-33342 (nuclei), and EdU-positive cells (EdU+) were quantified. Cell proliferation was expressed as percentage of EdU+ cells out of the total number of cell nuclei.

To investigate whether the effects observed were specific for IFNβ within the interferon family, related cytokines were also tested: IFNα binds to the same receptor as IFNβ, but with

lower affinity. Accordingly, it affected migration only at much higher (two orders of magnitude) concentrations (Fig. 2A). In contrast to this, the type II interferon IFNγ affected viability and the migration endpoint at low pM concentrations, similar to IFNβ (Fig. 2B). To unambiguously separate the effects on proliferation from those on migration, the interferons were tested under non-proliferating assay conditions (Fig. 2A,B), and both the 75% effective concentration (EC75), as well as the lowest effective concentration (LOAEL) were determined for the migration endpoint and compared to IFNβ (Fig. 3C). These experiments showed that IFNα inhibits migration of NCC, but that it was about two orders of magnitude less potent. The data also showed that IFNγ mainly affected the NCC proliferation at low pM concentration, while specific effects of migration were only observed in the nM range. Thus, IFNβ is the only tested interferon that affected the migration of NCC at clinically-relevant low pM concentrations.

As migration is a fundamental cell biological process, it may be assumed that the capacity of IFNβ to disturb cell movement may apply to many other cell types. However, it was shown earlier that many toxicants affect NCC migration without affecting e.g. tumor cells or other neural precursors (Zimmer et al. 2012). To study such specificity for IFNβ, we used the human breast cancer cell line MDA-MD-231, which is highly metastatic, displays a mesenchymal/fibroblastoid phenotype Prat et al. 2010, and thus migrates in a somewhat similar way as NCC. Under proliferating conditions, IFNβ showed potent effects (low pM range) on both the migration and the viability endpoints (Fig. 2D, left graph). Under test conditions that avoided any drug effects on proliferation, no inhibition of migration by IFNβ was detected (Fig. 2D, right graph). Thus, the effect of IFNβ on the migration of cells did not apply to any cell type, but may be rather specific for NCC.



IFN type	AraC	EC75 (migration) [pM]	LOAEL [pM]
IFNα	-	10000	1000
	+	> 10000	500
IFNβ	-	20	5
	+	50	25
IFNγ	-	110	20
	+	6000	5000

Fig.2 Specificity of IFNβ effects on NCC NCC were treated for 48 h with interferons while they were allowed to migrate. Then the viability and the inhibition of cell migration were measured. All assays were performed either with or without cytosine arabinoside (AraC, 10 µM) as culture medium supplement. A, B Testing of interferon-α (IFNα) and interferon-γ (IFNγ). Data are from three independent experiments. Error bars indicate standard deviations (SD). Statistics was performed for each endpoint by ANOVA, followed by Dunnet's post-hoc test (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$). C The EC_{75} and the lowest observed adverse effect level (LOAEL) were compiled for each scenario. The LOAEL was defined as the lowest concentration triggering a significant reduction of cell migration ($p \leq 0.05$) D The human breast cancer cells MDA-MD-231 were allowed to migrate for 48 h, before viability and the number of migrated cells were quantified. Cells were treated with the indicated concentrations of IFNβ for the total migration period, either with or without cytosine arabinoside (AraC, 10 µM) as culture medium supplement.

Maintenance of basic functions and cell morphology after exposure to IFNβ

One straightforward explanation for the relatively specific effects of IFNβ on NCC migration may be a change of the cell differentiation state by the cytokine. For instance a mesenchymal-to-epithelial (MTE) transition or a differentiation to a neural or other final cell type may reduce spontaneous migration of the cells. Therefore, various cell features were

studied. Typical cell type-specific mRNA markers (e.g. MSX1, SOX9; SNAIL2, NRP1, PAX3) did not change their expression level upon exposure to IFNβ (not shown), and also the characteristic cell shape was maintained. To assess the latter finding in more detail, some molecular markers of cell structure were studied. NCC were treated with IFNβ for 48 h under conditions that allowed migration. Then, the F-actin microfilaments were visualized with phalloidin. Mitochondrial structures were stained with an antibody to TOM20, and the Golgi apparatus was stained for the GM130 protein. A detailed characterization of these cell biological markers did not reveal any changes due to the exposure to IFNβ (Fig. 3A).

In a next step, various functions associated to cell adhesion were studied. The expression of 20 integrins and cadherins was studied, and none of them changed > 30% in the presence of IFNβ (data not shown). In a more functional approach, we assessed the phosphorylation (and activation) level of the focal adhesion kinase (FAK). Protein samples were prepared from NCC after 0, 6 or 24 h of adhesion, either with or without IFNβ. Adhesion of the cells led to the activation of FAK, as expected. This process was not significantly affected by exposure to IFNβ (Fig. 3B).

In a third approach, we examined whether IFNβ may affect the response of cells to external stimuli, such as inflammatory cytokines. NCC exposed to a mix of cytokines (CM: TNFα and IL1β) normally translocate the transcription factor NFkB from the cytosol to the nuclei (Fig. 3C, upper panels). IFNβ itself had no effect on NFkB translocation, but we tested whether it affected the signaling of the CM. However, cell responses did not differ in the absence or presence of the interferon (Fig. 3C, lower right panel). Thus, no obvious structural features and none of the tested signaling functions were impaired by exposure to IFNβ.

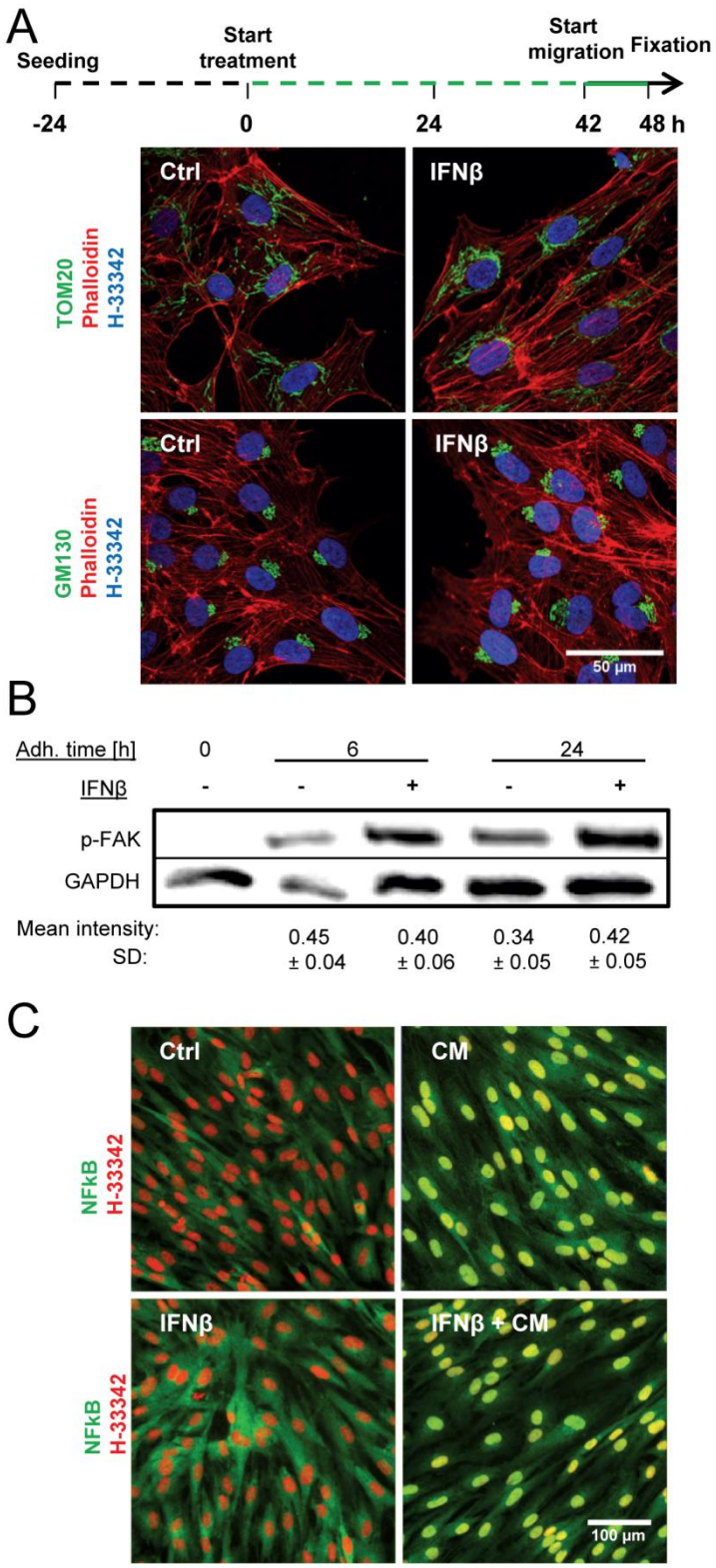


Fig.3 Maintenance of basic NCC functions and morphology in the presence of IFNβ A NCC were treated with IFNβ (500 pM) or solvent (Ctrl) for 48 h, allowed to migrate in the last 6 h of the treatment period, and finally fixed for immunofluorescence staining. The microfilament cytoskeleton was visualized by phalloidin, antibodies to TOM20 were used to visualize mitochondria and anti-GM130 for the Golgi apparatus. B NCC were seeded for 0, 6 and 24 h (adhesion time) in culture medium supplemented with IFNβ (500 pM) or solvent. Then cells were lysed and the amount of phosphorylated FAK was measured by Western blot analysis. The mean intensity of each band normalized to the respective loading control (GAPDH) ± SD is reported below each condition (n = 3). C NCC were exposed to a cytokine mix (CM, 10 ng/ml TNFα and 10 ng/ml IL1β), IFNβ (500 pM) or a combination of both for 1 h. Cells were then fixed and stained for nuclear factor kB (NFkB, green). Representative pictures for each condition are shown, with nuclei counterstained with H-33342 (red). Nuclei with translocated NFkB appear yellow, instead of red (non-translocated NFkB).

Attenuation of NCC migration speed and chemotactic behaviour by IFNβ

To confirm the impairment of NCC migration by IFNβ, and to further exclude any artefacts due to drug effects on cell proliferation or differentiation, two other functional assays were used: the first approach made use of time lapse imaging and cell tracking to study effects on cell speed. Migrating cells were incubated with IFNβ on a microscope stage in a temperature-controlled incubation system (Fig. 4A). This set-up allowed a continuous recording of phase contrast images for 30 h. The image stacks were then used to track the migrating cells (≥ 10 cells per well) (Fig. 4B). Analysis of the tracks revealed a concentration-dependent reduction of the accumulated distance travelled by IFNβ-treated NCC, compared to control cells (Fig. 4C). Thus, this method confirmed on the single cell level that IFNβ reduced the migration speed.

The second approach made use of a transwell set-up. In this system, cells can be seeded onto a porous membrane in the upper compartment. When a chemoattractant (e.g. FBS) is added to the bottom part, cells are triggered to migrate through the membrane. NCC were seeded on the transwell membrane directly or after a 42 h pre-treatment with IFNβ. At the end of the 6 h migration period, cells that had translocated through the membrane were counted (Fig. 4D). NCC treated with IFNβ for the period from 0 to 48 h showed a significant reduction of their capacity to migrate through the membrane. NCC treated only during the 6 h of the migration period (42 – 48 h) were not impaired (Fig. 4E). This set of data suggests that IFNβ also affects a short (6 h) migration test, but that the cells must be exposed for a longer time to IFNβ (> 6 h) for the inhibition of migration to become apparent. Both approaches confirmed the effects of IFNβ on NCC migration.

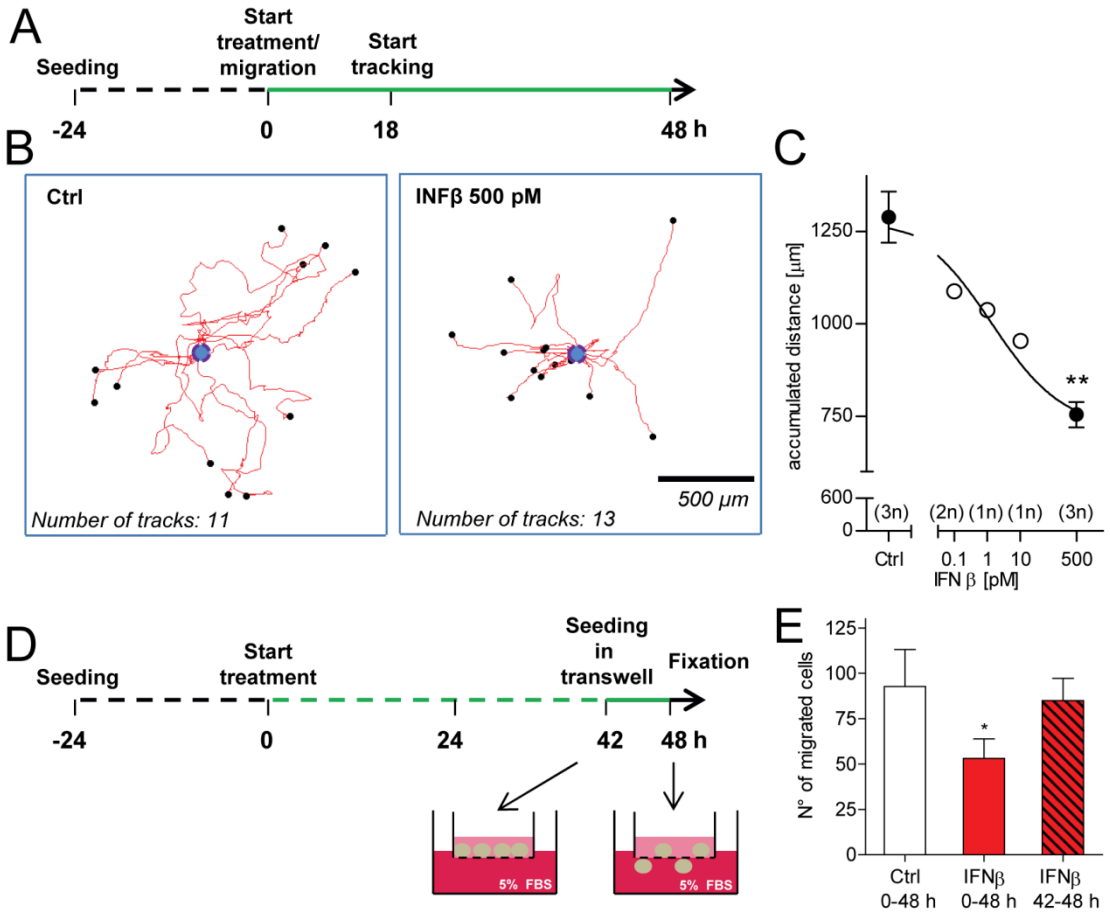


Fig.4 Confirmation of impaired NCC migration in the presence of IFNβ in other functional assays A NCC were allowed to migrate for 48 h, and they were exposed to IFNβ at the indicated concentrations for the entire migration period. Phase contrast images were taken every 15 min during the last 30 h of migration for cell tracking. B Representative migration-tracks for the control and 500 pM IFNβ-treated cells are shown. Tracks are normalized to the same starting point (blue circle); x and y dimensions are scaled similarly (see scale bar). C The averaged accumulated distance covered by the cells was then calculated for each test concentration. At least 10 cells were followed for each condition (= 1 technical replicate). The number of independent replicates for each condition is reported in the graph. Error bars indicate standard deviations (SD). Statistics is based on ANOVA, followed by Dunnet's post-hoc test (** $p \leq 0.01$). D NCC pre-treated for 42 h with IFNβ or solvent were seeded into transwells. Then the cells were induced to migrate through the transwell porous membrane by addition of 5% FBS into the lower chamber of the transwell, in the presence of IFNβ (500 pM) or respective control. After additional 6 h in the presence or absence of IFNβ, migrated cells were stained with crystal violet. Four fields per replicate were imaged, and the cells were counted. E Quantification of the number of migrated cells for the different conditions: the total IFNβ exposure period is indicated; data are from 4 independent experiments. Error bars indicate standard deviations (SD). * $p \leq 0.05$.

Transcriptome changes and alterations of superordinate biological processes induced by IFNβ in NCC

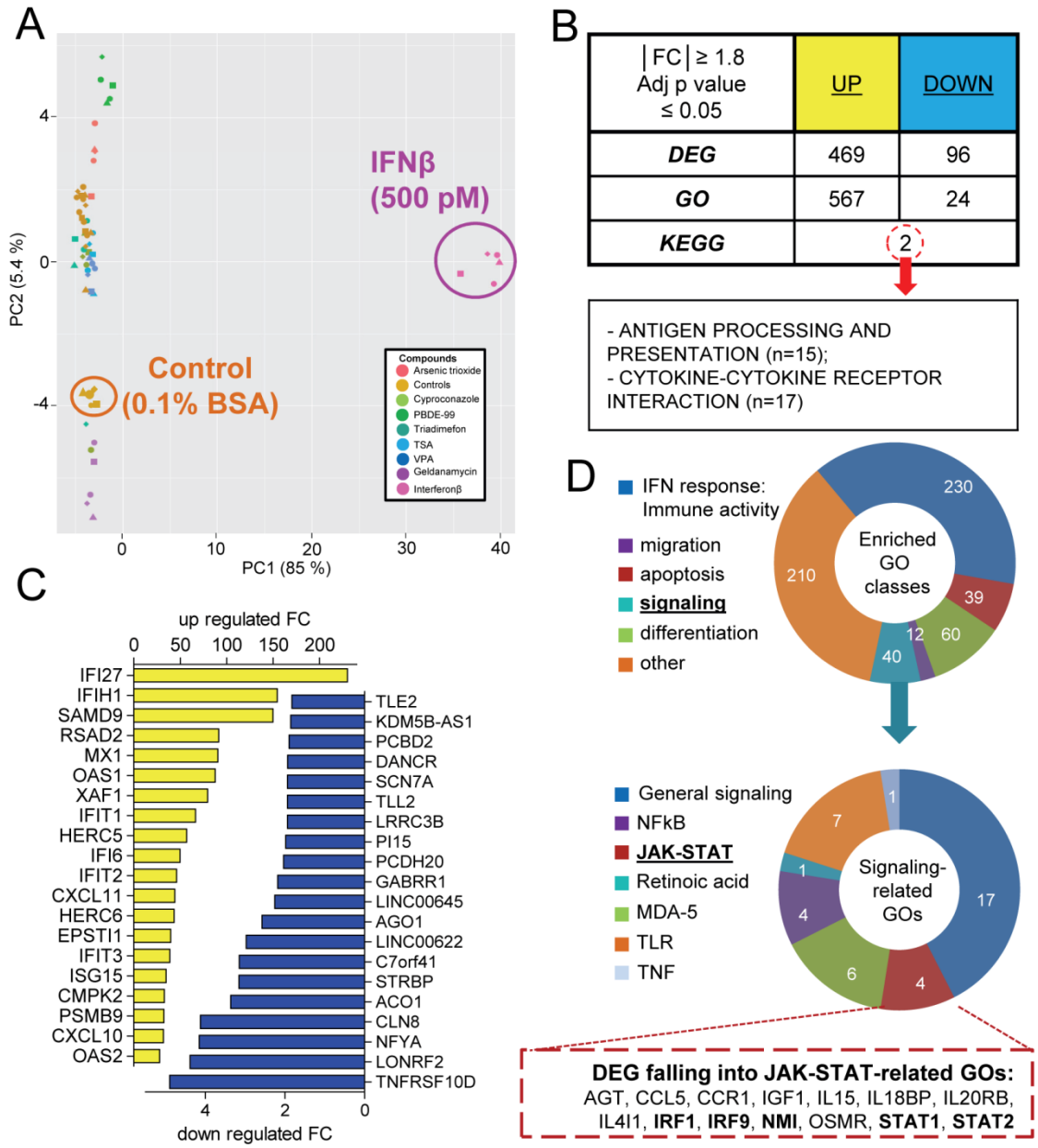
To obtain more data on the type of effects IFNβ has on NCC, cells were exposed for 48 h to the cytokine (500 pM). Then, mRNA was prepared from five different cell lots and used for gene expression analysis by Affymetrix microarray technology. The data were compared to those obtained for seven other toxicants (arsenic trioxide, cyproconazole, PBDE-99, triadimefon, TSA, VPA, geldanamycin) known to inhibit the migration of NCC (Zimmer et al. 2014, Pallocca et al. 2016). Each of the compounds was used at its respective highest non-cytotoxic concentration, and the data structure was visualized by a principal component analysis (Fig.5A). Plotting of the first two principal components showed a strong separation between IFNβ-treated samples and all other samples along the first principal component axis. This was due to the much stronger gene regulation response triggered by IFNβ, compared to small molecular weight toxicants.

The gene expression data was then used to determine the differentially expressed genes (DEG). The DEG for each toxicant were defined here as the group of microarray probe sets (PS), which differed significantly from negative controls (FDR adjusted p-value ≤ 0.05), and showed expression level changes (fold change (FC)) of ≥ 1.8 or ≤ 0.55. IFNβ triggered a much stronger up-regulation than down-regulation response, altering altogether 565 PS (Fig.5B). To get more insight into the de-regulated genes, the DEG were sorted according to their p-values and the top 20 up- and down-regulated genes were displayed. Among the most strongly up-regulated genes, there were many IFN-induced proteins (IFI), which have mostly roles in the inhibition of viral replication. Other genes were the 2' – 5' oligoadenylate synthetase family (OAS1/2), which activates the latent RNase L, and induces viral RNA degradation; another typical biomarker of IFN antiviral responses was the MX dynamin like GTPase (MX1). Additionally, we found a strong upregulation of the chemokine response (CXCL10 and 11). The group of strongly down-regulated genes was more heterogeneous. It comprised for instance the TNF receptor superfamily member TNFRSD10D, argonaute 1 (AGO1, the catalytic component of the RISC complex involved in the RNA interference process), and the zinc-dependent metalloprotease tolloid-like (TLL2) (Fig.5C).

As the analysis of individual genes did not directly indicate which migration-related pathways may be impaired, we investigated whether biologically-linked groups of genes were co-ordinately regulated. More than 500 gene ontology (GO) terms were enriched (adjusted p-value ≤ 0.05) amongst the up-regulated genes, while only 24 terms were over-represented amongst the down-regulated genes. To get a better idea on co-ordinately affected pathways, we also identified over-represented KEGG pathways amongst the sum of DEG. The two pathways found by this approach were "antigen processing and presentation" (15 genes represented) and the "cytokine-cytokine receptor interaction" pathway (17 genes represented). This was not surprising, given a main role of interferons in these processes (Fig.5B), but it did not offer an evident explanation, why migration was impaired.

To get an overview of the overrepresented GO groups, we assigned them to six superordinate cell biological processes (Waldmann et al. 2014): IFN responses, migration, apoptosis, signaling, differentiation and other. As expected, most of the enriched GO classes belonged to the "IFN response group" (230 GO term groups). We focused on the signaling-related overrepresented GO terms (40 groups) to get some hints regarding the mode of action of IFNβ on the migration capacity of NCC. Several GO in this group related to innate immunity and inflammatory signaling. Amongst the more general signal transduction pathways, only the JAK-STAT pathway was clearly over-represented (4 GO groups) with several key players of this pathway being up-regulated (Fig.5D). For this reason, we decided to examine the role of JAK-STAT signaling on NCC migration more directly.

Fig.5 Transcriptome changes triggered by IFNβ in NCC_ Sampling for microarray analysis was performed in NCC after 48 h exposure to non-cytotoxic, but migration-inhibiting, concentrations of eight test battery hits, as identified in Zimmer et al. 2014. Data are from 5 independent experiments (= data points of one colour, but different shapes). A A principal component analysis (PCA) was performed, and a 2D plot was generated to display the transcriptome data structure across compounds and experimental replicates. The positions of IFNβ-exposed samples, and the respective control, are highlighted. B The number of differentially up-regulated (UP) or downregulated (DOWN) genes (DEG, IFNβ vs control) and the corresponding biological processes (over-represented GO classes) was identified. Over-represented KEGG-pathways were searched amongst all DEG (UP and DOWN), and the only two significant pathways are indicated. C The identified DEG were sorted according to their p-value. The top 20 up- (yellow) and down- (blue) regulated genes are shown as bar graphs indicating the fold change (FC). D The ring diagrams show the relative distribution of 6 superordinate biological processes (IFN response, migration/chemotaxis, apoptosis, signaling, differentiation and other) amongst the over-represented GO classes (upper ring) and the number of the different signaling-related over-represented GO classes (lower ring). Genes with a central role in the JAK-STAT pathway are depicted in bold. ►



Correlation of IFNβ-induced inhibition of NCC migration and JAK-STAT pathway activation

As readout for JAK-STAT activation, we chose to analyze the phosphorylation level of one of the members of the STAT family, STAT1. This step is necessary for the heterodimerization of STAT proteins, which, together with IRF9 form then the transcription

factor IFN-stimulated gene factor 3 (ISGF3), and promotes IFN-induced genes transcription. Western blot analysis showed an increase of p-STAT1 in NCC after stimulation with both IFN class I interferons, with IFNα being about 100-fold less potent than IFNβ (Fig.6A, B). The extent of STAT1 phosphorylation was quantified over several experiments, and it correlated with the extent to which migration was inhibited. This data set showed clearly that p-STAT1 levels correlated with the inhibition of migration in the case of IFNβ treatment (Fig.6C). To reach 50% of STAT1 phosphorylation, NCC needed to be exposed to 2 nM of IFNα (vs 5 pM IFNβ), but also for IFNα the increase of pSTAT1 correlated with inhibition of migration (Fig.6D). To evaluate a specific role of JAK-STAT signaling, we also examined phosphorylation of extracellular signal regulated kinases (ERK1/2), of protein kinase B (AKT), and of glycogen synthase kinase-3β (GSK3β), but none of them was changed by exposure to IFNβ (data not shown).

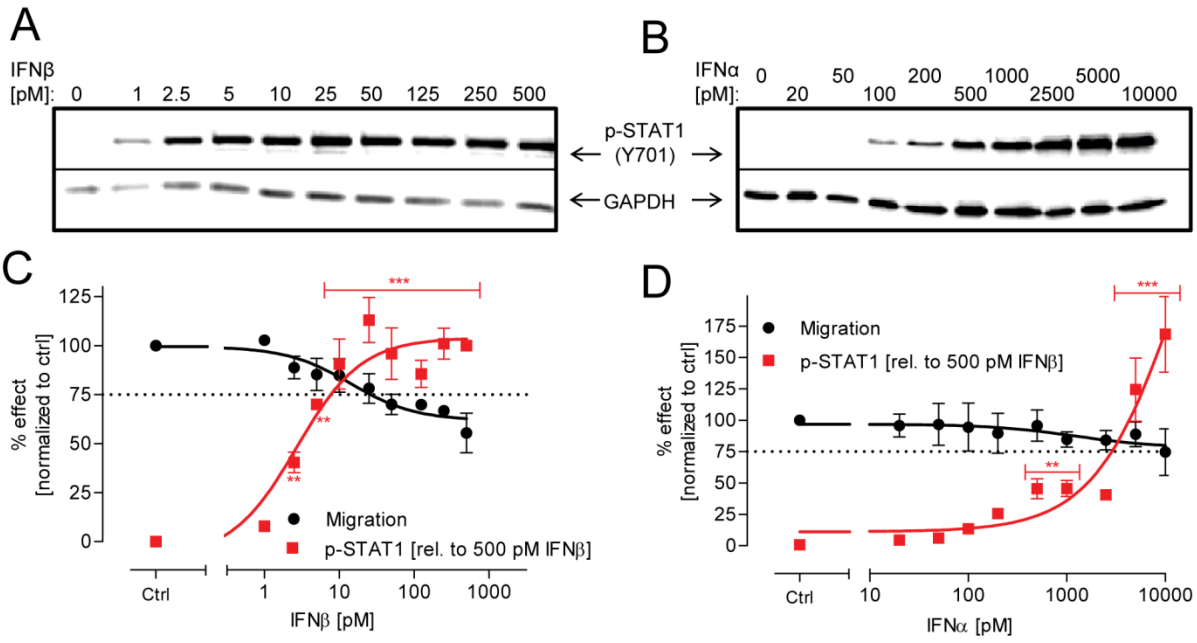
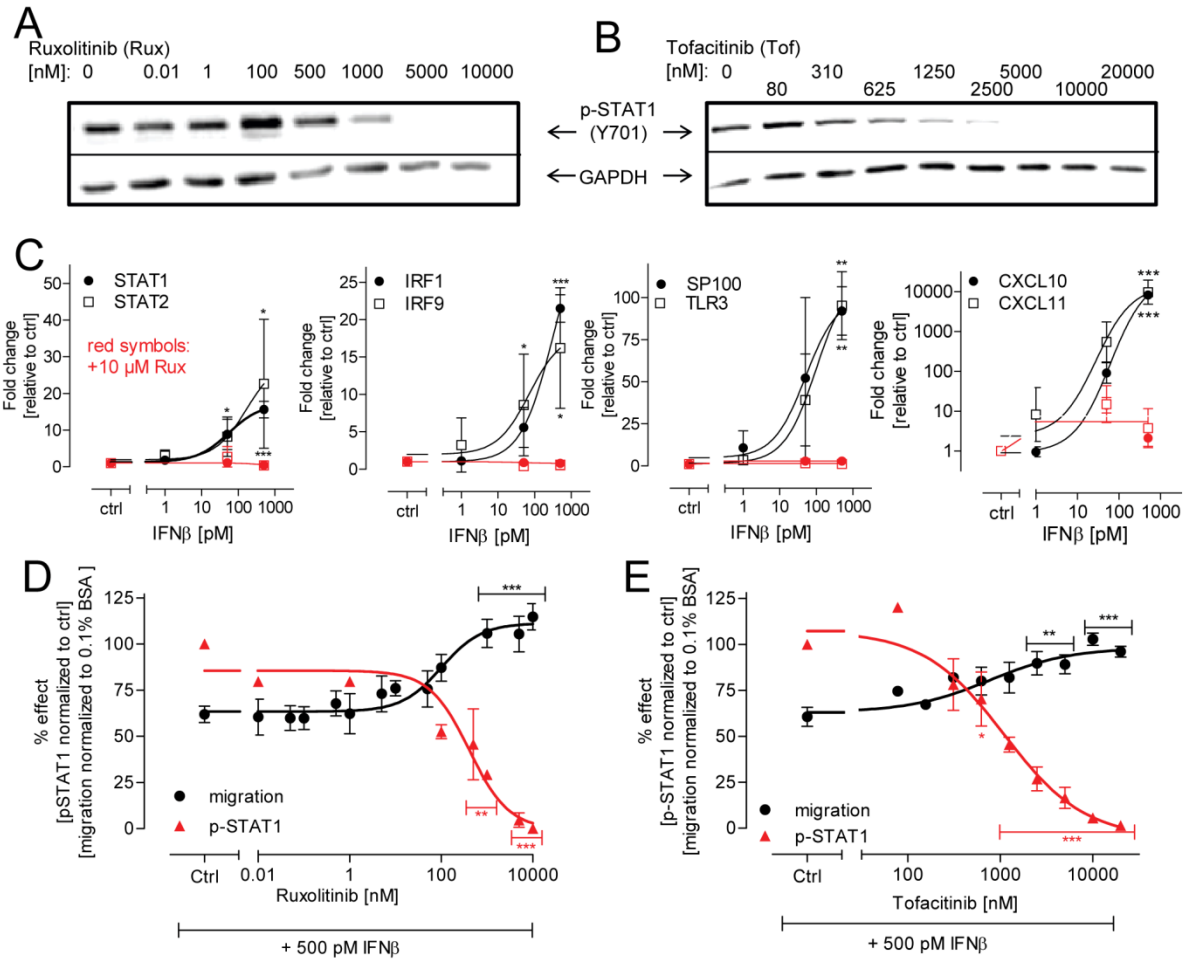


Fig.6 Correlation between JAK-STAT pathway activation and inhibition of NCC migration upon treatment with class I IFN_A, B NCC were exposed to the indicated concentrations of IFNβ and IFNα for 1 h. Then, cells were harvested and protein samples were prepared. The amount of phosphorylated STAT1 (p-Tyr701) was measured by western blot analysis (representative blots are shown). C, D The band intensity was quantified and normalized to the respective GAPDH antibody band. For better comparison, the migration inhibition data from Fig. 1B are shown in black in the same graph. Data are from three independent experiments. Error bars indicate standard deviations (SD). Statistical analysis was based on ANOVA, followed by Dunnet's post-hoc test (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$).

To obtain more causal evidence for the role of the JAK-STAT pathway in the inhibition of NCC inhibition, we used two different and specific JAK inhibitors (ruxolitinib and tofacitinib) to block the kinase signaling. Both inhibitors blocked STAT1 phosphorylation concentration-dependently (Fig.7A, B). Moreover, exposure to ruxolitinib completely abolished the activation of eight different interferon target genes in cells exposed to IFNβ (Fig.7C). A pharmacological block by any of the two inhibitors of the pathway resulted in the complete prevention of adverse effects of IFNβ. Moreover, the extent of signaling pathway inhibition correlated well with the attenuation of the cytokine effects on migration (Fig.7D, E).



◄ Fig.7 Abolishment of the effect of IFNβ on NCC migration by inhibition of the JAK-STAT pathway A, B NCC were pre-treated for 0.5 h with two different JAK inhibitors (ruxolitinib and tofacitinib) at the indicated concentrations. Then, the cells were further treated for 1 h with the inhibitors in cell culture medium supplemented with 500 pM IFNβ. Finally, cells were harvested and the amount of phosphorylated STAT1 (p-STAT1) was measured by western blot analysis (representative blots are shown). C NCC were exposed to IFNβ for 48 h at the indicated concentrations, either with or without ruxolitinib (Rux, 10 µM). Then cells were harvested, and total RNA was extracted and retro-transcribed. Effects on selected mRNAs were evaluated by qPCR. Expression levels in the presence of Rux are shown in red. D, E The band intensities were quantified for p-STAT1 (normalized to GAPDH). In a parallel set of experiments, migration (after 48 h) was evaluated in the presence of IFNβ (500 pM) plus ruxolitinib (left) or tofacitinib (right). Data are from three independent experiments. Error bars indicate standard deviations (SD). Statistics was performed by ANOVA, followed by Dunnet's post-hoc test (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$).

Requirement for continued JAK-STAT signaling for impairment of NCC migration

Having established that the adverse effects of IFNβ on the migration of NCC are signaled via the JAK-STAT pathway, we were interested in the role of short vs continued exposure to interferons. To address this question, we exploited the rescuing effect of the JAK inhibitor ruxolitinib to terminate the IFNβ signaling at defined time points.

The standard migration assay, as described in 3.1, was performed in the presence of IFNβ (or respective control). Ruxolitinib was added to the cells immediately or at various time points after exposure to interferon. The inhibition of migration was then evaluated at 48 h for the various experimental conditions (Fig.8A). The results from these experiments showed that the inhibition of migration (by IFNβ) was rescued when the inhibitor ruxolitinib was added not later than 6 h after the beginning of IFNβ treatment. When ruxolitinib was added after 10 h (or later), inhibition of migration (by at least 25%) occurred (Fig.8A). Thus, inhibition of migration by IFNβ required receptor signaling for at least 6-10 h, while activation of the JAK-STAT pathway for up to 6 h hardly affected the migration capacity of NCC (Fig.8B).

This finding was corroborated when standard migration assays were performed and IFNβ was washed out after 3 h. Under these conditions, no inhibition of cell migration was observed (data not shown).

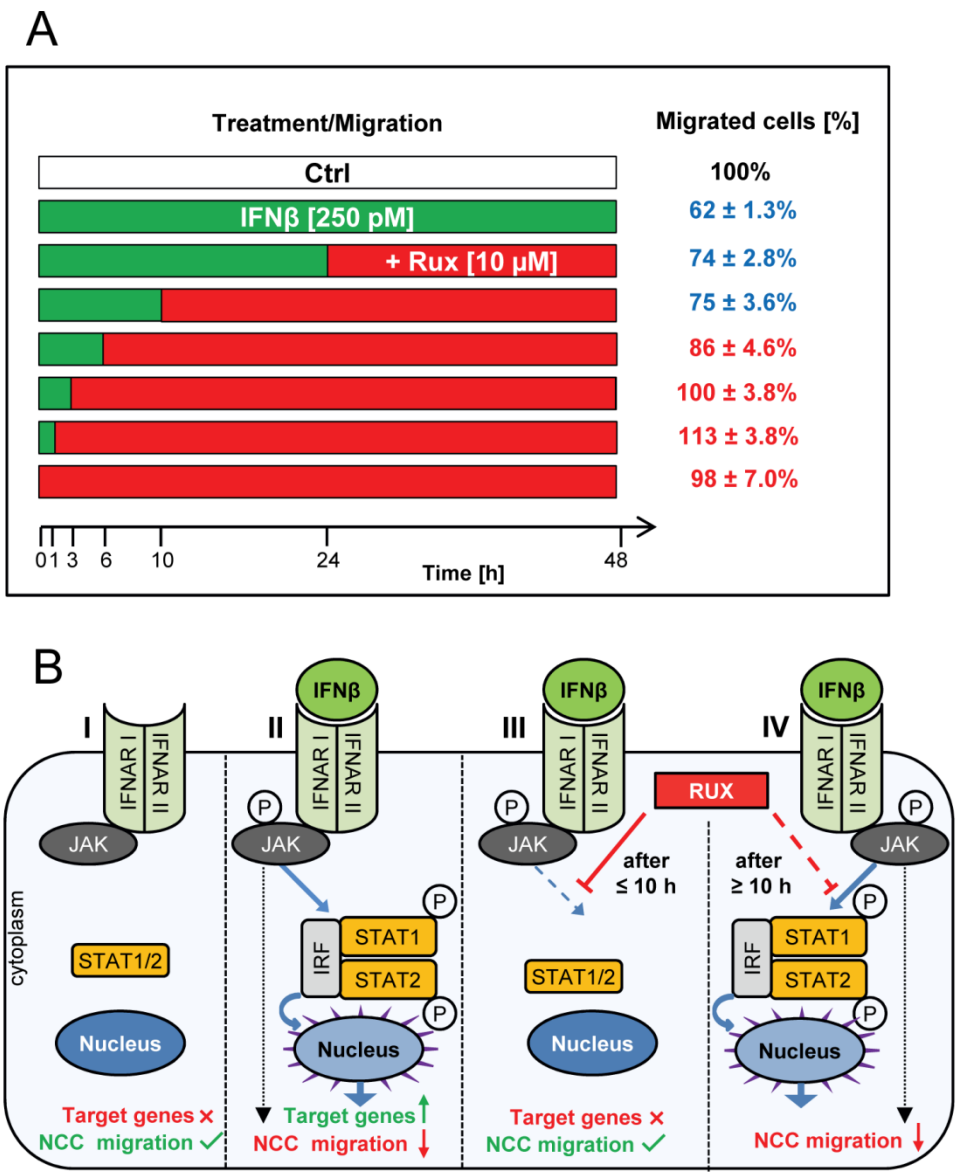


Fig.8 Requirement for continued JAK-STAT signaling to NCC migration impairment A NCC were allowed to migrate for 48 h. During this time, they were exposed to IFNβ while ruxolitinib (10 µM) was added to the culture medium at different time points after the start of the migration. The percentage of migrated cells was quantified for all conditions after 48 h. Migration was inhibited ≥ 25% in all conditions denoted in blue. Ruxolitinib without IFNβ had no effect on migration. B Graphical representation of the effects of IFNβ and JAK-STAT pathway-inhibitors on the transcription of target genes and on the functionality of NCC. NCC migration was unimpaired in the absence of IFNβ or when IFNβ was present together with ruxolitinib. NCC migration was impaired when IFNβ was present alone, or when ruxolitinib was added ≥ 10 h after IFNβ.

DISCUSSION

To summarize the present study, a potential developmental toxicity hazard of IFNβ with regard to NCC disturbance was characterized. Migration of NCC was affected at low pM concentrations by IFNβ, while IFNα was substantially (two orders of magnitude) less potent, and IFNγ did not affect migration in the pM range. The effect on migration clearly differed from the cytostatic activity of the cytokines, as (i) IFNβ and IFNα had similar cytostatic potency on the test system, and (ii) as cytostatic and migration-inhibiting properties could be experimentally separated. Moreover, the specific effect of IFNβ on migration was confirmed in assays with different protocols and endpoints. A closer cell biological characterization showed that the inhibition of migration correlated well with a prolonged activation of the JAK-STAT signaling pathway in various sets of experiments. Thus, prolonged activation of STAT1 by IFNβ in NCC provides a biochemical marker of potentially adverse effects of IFNβ on fetal development during pharmacological use.

As IFNα and IFNβ signal through the same receptor complex, the largely different potencies observed here may appear surprising. However, a more pronounced effect of IFNβ has also been observed in other test systems. For instance, only IFNβ induced JAK1 activation in human myocardial fibroblasts, and this was associated with a 120-fold higher potency to trigger an antiviral response (vs IFNα) (Grumbach et al. 1999, Heim et al. 1996); also in human vascular endothelial cells, a 2-3 log difference was observed for IFNβ vs IFNα in functional assays (da Silva et al. 2002), and IFNβ was 100-fold more potent than IFNα with respect to the inhibition of differentiation of human monocytes into osteoclasts (Coelho et al. 2005). Such observations are explained by the largely different affinities of IFNα and IFNβ for the shared receptor subunits. From this it follows that the latter cytokine is more potent in conditions with low receptor expression (Moraga et al. 2009, van Boxel-Dezaire et al. 2006, Schreiber and Piehler 2015). For cells that express high numbers of receptors, the different affinities play only a minor role and both type I interferons can trigger biological responses at about the same potency. As high and low receptor numbers are not an absolute measure, but relate to the percentage of receptors required to trigger a full biological response in a given cell type, the potency difference of IFNα and IFNβ may differ for different responses within

one given cell (Piehler et al. 2012, Zula et al. 2011). From these considerations, widely accepted in the interferon literature, we conclude that NCC express low numbers of the receptor subunits IFNAR1 and IFNAR2 on their surface. Moreover, we conclude that a higher occupancy is required to affect migration than to trigger a cytostatic effect.

To our knowledge, effects of IFNβ on the migration of neural cells or their precursors have not been reported yet. However, the migration and invasion properties of glioma cells are affected by interferon-regulated genes (Tarassishin and Lee 2013, Yu et al. 2011) and neuronal survival seems to require a basic level of interferon signaling (Ejlerskov et al. 2015). Moreover, interference of IFNβ with migration has been shown especially in the leukocyte lineage of cells (Lou et al. 1999, Staun-Ram and Miller 2011, Floris et al. 2002) and for various tumor cells (Booy et al. 2015, Rossi et al. 2015). Many of these studies suggest that cell migration may be affected by interferons through a reduced secretion of matrix metalloproteases (Stuve et al. 1997, Yen et al. 2010). This mechanism is unlikely to play a role in the experimental systems used in our study, as the penetration of extracellular matrix is not required for two-dimensional NCC migration. As IFNβ is extremely pleiotropic in its activities, great care is required to distinguish effects that are correlated with reduced migration and effects that causally lead to reduced migration.

In the present study, we found that JAK-STAT signaling is a causal factor involved in reduced migration. In parallel, we invested major efforts to avoid erroneous conclusions, based on secondary effects triggered by IFNβ. These experiments followed two main lines. First, we made sure that migration was specifically affected, and that test results were not an indirect consequence of the known cytostatic effects of interferons (Bekisz et al. 2010). These precautions are particularly important for slowly migrating cells (as NCC), while such controls are less common in the field of leukocyte migration. One taken approach was to run the assay under conditions that did not allow any proliferation (presence of a mitosis inhibitor). IFNβ showed its effect on migration also in this setup. The second approach was to use an assay with a faster readout (transwell assay), and then, going one step further, we followed individual cell migration directly by tracking. Also these experiments confirmed the inhibition of migration in a low pM range of IFNβ.

Second, we provided evidence that IFNβ treatment did not obviously alter the cell differentiation state. In theory, extracellular signals may lead to the differentiation of NCC to neurons or other progeny that are less prone to migration. This is unlikely to happen in our test system, as several differentiation markers remained unchanged, as a characterization on the organellar level did not reveal differences after treatment with IFNβ, and as key molecules relevant to adhesion and migration, like focal adhesion kinase or expressed integrins, remained unaffected. The stability of the overall NCC phenotype is remarkable, especially in the light of the transcriptome analysis, which showed that a typical (anti-viral and immunomodulatory) interferon response was triggered on the level of gene expression. Most conspicuous amongst these regulated transcripts were several chemokines and chemokine receptors (CCL5, CXCL10, CXCL11, CCR1), but the changes were observed too late to be relevant for the effects on migration observed here. Moreover, the migration of NCC was not affected in our assay by chemokine receptor antagonists (not shown).

Several sets of data presented here suggest that the extent of activation of the JAK-STAT pathway correlates with inhibited migration by IFNβ. Moreover, we also explored the temporal relationship of this pathway activation with NCC motility. It is known that cytokines, like interferons, may either trigger a hit-and-run signaling response (Brask et al. 2004), i.e. short formation of the receptor ligand complex has long-acting or irreversible effects, or they may lead to an acute cellular change reversible after termination of cytokine-receptor interaction. In other cases of cytokine action, prolonged receptor signaling is required to trigger downstream cellular processes, and once triggered the cellular alterations may remain stable. Here, we found an example of the latter response pattern: activation of the JAK-STAT pathway for less than 6 hours had no significant effect on cell migration; activation for more than 10 hours was sufficient to trigger the full response, also when the kinase pathway was then blocked by specific inhibitors. This suggests that short peak levels of the cytokine would not affect NCC migration in a fetus, but if such levels are maintained in chronic infection, or after transfer of pharmacologically applied IFNβ across a compromised placental barrier, adverse developmental effects may follow.

In this context, it is important that interferons show a high degree of species-specificity. The evaluation of the effects of human IFNβ is therefore not possible in rodent or rabbit models, and only some monkeys (e.g. rhesus monkeys) show pharmacodynamic responses to the

cytokine. Concerning the pharmacokinetic behaviour, the most relevant data come from clinical trials. IFNβ is normally given to MS patients intramuscularly at a dose of 6 million international unites (MIU) (Barbero et al. 2004). In controlled pharmacokinetic studies, 12 MIU IFNβ given intramuscularly to healthy patients led to a Cmax of 44 IU/ml, corresponding to ~9 pM Alam J 1997. A rough linear approximation would suggest that the clinical dose used in MS would lead to peak plasma levels of 4-5 pM, i.e. roughly within the range of the concentrations found to affect NCC here. In a previous study (Zimmer et al. 2014), we calculated the dose to result in developmental toxicity to monkeys (FDA 1999) to correspond to plasma levels of 540 pM. This concentration is higher than the concentrations found here to affect human NCC, but the data basis is relatively limited and the concentration estimate requires several assumptions for the PBPK modelling (Zimmer et al. 2014). Unfortunately, retrospective clinical studies do not provide a clearer set of information. The safety of IFNβ in pregnancy has been observed on several occasions, but with contrasting results (Sandberg-Wollheim et al. 2005). To date, women with MS are advised to interrupt the exposure to the drug for precautionary reasons (Lu et al. 2012, Pozzilli and Pugliatti 2015), but information on no-effect levels or critical windows of susceptibility are not available.

In the absence of clear clinical data, and in the pertinent case of highly species-specific effects of the drug, our human cell-based testing approach allowed important insights into potential toxicological hazard, and into the concentration range at which this may be expected. In this context it is important that the phenotypic adverse effect (inhibited migration) was consistent with a plausible biochemical mechanism of IFNβ signaling (JAK kinase activation), that exact information on the relevant concentration range was obtained, and that several tests and endpoints confirmed the type of potential hazard identified in initial studies. On the basis of this example study, we suggest that a combination of cellular tests, based on relevant human cell types and engineered tissue structures (e.g. placental barrier) be used more frequently to obtain information on toxicological properties of human-specific drugs.

F. Concluding Discussion

This thesis contains three experimental result chapters (C, D, E), based on two publications and one submitted manuscript. In each chapter, the main results are discussed in detail. In this concluding discussion chapter, only the most important findings will be reported and collected in a general frame. Moreover, novel approaches and challenges of *in vitro* developmental toxicity testing will be discussed (Figure 9).

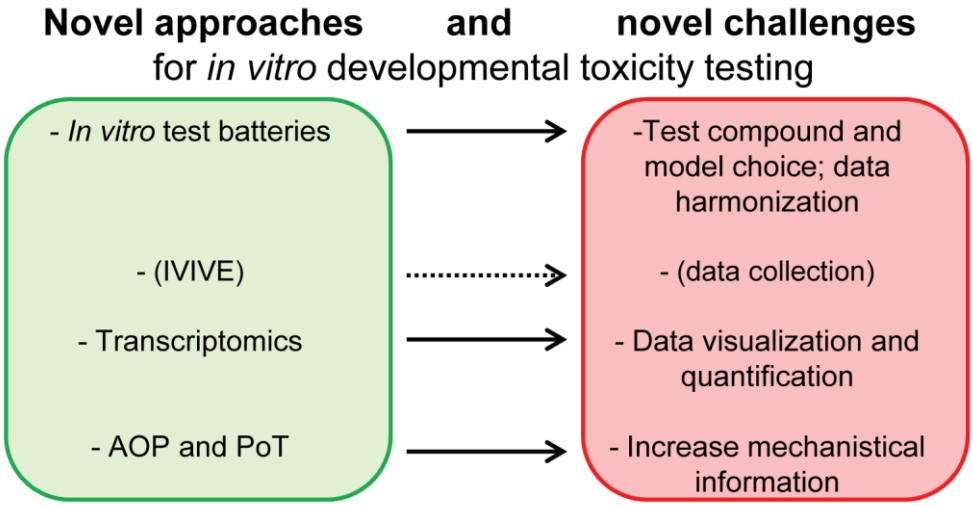


Figure 9 Novel approaches and challenges for *in vitro* developmental toxicity testing. An overview of the approaches and related issues discussed in this chapter. IVIVE: *in vivo-in vitro* extrapolation analysis; AOP: adverse outcome pathway; PoT: pathway of toxicity.

1. Novel approaches: assembling of *in vitro* DT test batteries

1.1 Advantages of *in vitro* DT testing strategies

The use of novel human *in vitro* cell-based models allows to explore the adverse effects of environmental pollutants, human-specific drugs and biologicals on the human development.

The classical developmental toxicity testing is normally performed in animals by two generations studies. This is done by exposure of pregnant animals to high dose of a substance, followed by observation of the presence of fetal toxicity, e.g. embryo-lethality or malformations.

In vivo developmental toxicity testing faces different disadvantages (Basketter et al. 2012). Firstly, this approach requires a large number of animals, since absence of toxicity must be

proved in at least two different species. Also, the used doses often far exceed the actual human exposure levels, making the data extrapolation for risk assessment more complicated (Leist et al. 2014); furthermore maternal toxicity can hardly be distinguished from embryo toxicity.

Additionally, many compounds, e.g. cytokines, may induce species-specific side effects since pharmacokinetics and maternal metabolism can widely vary between species, as well as the placental anatomy and physiology (Gundling and Wildman 2015).

Altogether these issues address the problematics of *in vivo* testing to correctly detect human DT hazards.

In addition, the high costs and relative low through-put of the DT animal testing can hardly cope with the increasing amount of chemical compounds and drugs which still need to be tested and must be registered in REACH (Registration, Evaluation, Authorisation & restriction of Chemicals), the European Union regulation for chemicals; it is estimated that more than 4'000 substances, which already underwent registration, will need additional assessment for their potential toxicity (EC 2006).

To speed-up the testing of this large amount of chemicals, *in vitro* methods have been proposed as a relevant alternative to *in viv*o testing, because of the possibility to increase the testing through-put and lowering the related costs (Judson et al. 2013).

Additionally, *in vitro* methods could address some of the presented problems, e.g. species-specific toxicity can be addressed by the use of human cell-based systems; lower concentrations, closer to the relevant human clinical concentrations can be tested; as well as maternal- and embryo-toxicity can be studied separately.

Furthermore, the molecular mechanisms underlying the observed toxicity could be easier studied in *in vitro* models. This approach would offer the opportunity to get more insight regarding the compound mode of action (MoA) and the particular involved pathways, moving to a more evidence- and mechanism- based toxicology, strategy strongly suggested and described in the report titled Toxicity Testing in the 21st Century – A Vision and a Strategy, compiled by the US National Research Council (NRC 2007, Rovida et al. 2015b)

1.2 Test batteries to model developmental toxicity processes

A relevant limitation of *in vitro* testing compared to *in vivo* studies is its lack of biological complexity. A single test model will hardly be able to fully resemble a complex toxicological response, because of the multitude of processes and mechanisms that take place during the developmental process. To address this problem a new approach has been proposed: the assessment of different relevant and complementary test models in a test battery, which will be able to model different biological and molecular events which occur at the different stages of development. This approach, in combination with sufficient information of reproductive toxic potential of chemicals, would allow to determine the hazard of a particular test compounds (Leist et al. 2014, Rovida et al. 2015a, Bal-Price et al. 2015) and to predict human developmental toxicity (Basketter et al. 2012, Piersma et al. 2013, Schenk et al. 2010).

2. DT *in vitro* test batteries challenges: selection of relevant test compound and models, and data harmonization

In the first results chapter, we addressed several issues regarding the general characteristics of an *in vitro* test battery establishment. Initially, we had to face different issues, regarding the choice of test compounds and test systems.

2.1 Characteristics of a relevant DT test compound

The field of developmental toxicity lacks of a heterogeneous training compound group that can be used as positive control for extensive characterization and validation of novel testing methods. A strategy to address this lack of positive test substances would be to choose new potentially interesting compounds by literature data mining, in particular by screening for compounds whose MoA or developmental toxicity have been studied and showed to be relevant (Kadereit et al. 2012). Additional characteristics to be taken in consideration would be the compound class: different classes of compounds should be chosen, e.g. environmental pollutants and clinical drugs, to increase the applicability domain heterogeneity of the candidate DT compounds. The newly chosen candidates, together with known positive compounds, can then be tested in relevant test systems, which should be already characterized by tool compounds that specifically interfere with the endpoint of the test

system. Once the novel compounds have been tested in many different assays and cell types, their toxicity profiles could then be compared to evaluate their relevance for developmental toxicity testing (Kadereit et al. 2012). This approach offers the opportunity of breaking the vicious circle between lack of sufficient reference compounds and a paucity of validated assays: increasing the size and heterogeneity of the test compound library as well as comparing the toxicity data obtained in the several assays could attempt to overcome the lack of DT positive controls and it would allow at the same time a more extensive characterization of novel *in vitro* methods (Figure 10).

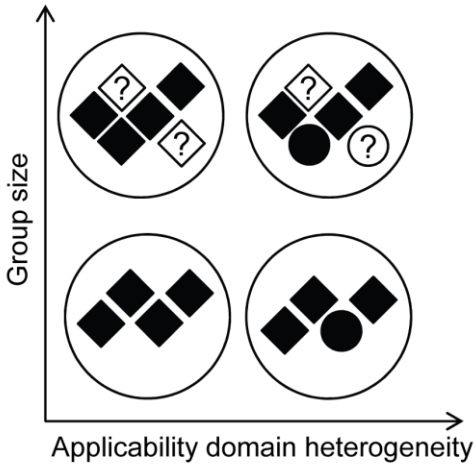


Figure 10 Test compound group composition. The substances to test can vary in applicability domain and in group size. Different applicability domain classes are depicted in different shapes. Known positive compounds are shown in black, and novel potential positive compounds in white.

A second addressed point regarded the choice of the test concentration range. This represents a crucial step as the dose is a substantial part of the toxicity of a compound. In drug discovery, it is common practice to screen compounds at fixed absolute concentrations. Instead in our screening, relative concentrations were chosen for the testing. This strategy was designed for test systems whose throughput is limited to some extent. For assays with very high throughput, an alternative approach would be to simply screen a large number of concentrations over the entire range of compound solubility. This approach was taken for instance by the USA national toxicology program (Xia et al. 2008; Attene-Ramos et al. 2013; Tice et al. 2013) or the EPA ToxCast program (Judson et al. 2010; Sipes et al. 2013).

In our screening, the initial test concentrations were chosen based on biological/mechanistic rationale. In most cases, the non-cytotoxic range was evaluated and chosen as starting point for the screen. In other situations, other criteria were chosen, e.g. in the case of cytokines we chose a test concentration that corresponds to the levels expected in body fluids during clinical application. In the follow-up studies, we tested concentrations anchored to the observed biological effect, e.g. migration inhibition.

Another possible approach to identify the initial relevant test concentrations would be to perform reverse pharmacokinetic modeling analysis a priori to the testing, in order to test *in vitro* concentrations which are relevant to the *in vivo* doses to which a person could be exposed.

2.2 Characteristics of a relevant DT test method

As further step, the characteristics that a test method of the test battery should present were defined. A test method is composed by a test system, a measurable endpoint (analytical or biological) and a prediction model. A test system can be a defined cellular model that can be exploited for measuring effects of any kind of external perturbance in a system that resembles a particular organ and tissue, or a particular developmental stage (Pamies et al. 2016). A test system was considered as relevant for the test battery testing by two criteria: being human-cell based, and biochemically and biologically well characterized. The most important aspect was the biological relevance of the model, in our case, for developmental toxicity. Furthermore, the test models must be able to be implemented in test systems which allow to measure changes of a relevant endpoint in a concentration- dependent manner and with medium/high through-put. Additionally, the test method performance, specificity and sensitivity, should be well characterized by the use of tool compounds, substances specific for interfering with the measured endpoint (Pamies et al. 2016).

Finally, the single test models must be chosen in the frame of joint test battery and properly selected. The series of tests of a test battery should be complementary to each other and allow to measure different components of a defined multi-factorial toxic effect (Hartung et al. 2013b).

2.3 Harmonization of test battery data

In the initial part of the first publication, a collection of different toxicity data from different test models for a pre-characterization of each test compound was described. Cytotoxicity studies were performed in different cell types, e.g. cancer cell lines and murine in vitro systems and the toxic concentrations were reported together with the clinical exposure data.

These data build a toxicological frame which represented a rough starting point for the following actual testing in the selected DT test models.

Several test methods were initially included in the test battery (see introduction paragraph 3.3.1). Since the test battery was built in an untiered way, it allowed temporal independency of the testing among the single tests. The MINC assay was the first functional assay used to test the complete group of compounds. This screening led to the identification of several novel DT hits. The next step would have been to harmonize the toxicological information derived from all selected test systems. Although a final collection of this information was not described, it is of extreme relevance to discuss the approaches for test battery data harmonization, in the perspective of risk assessment.

A possible and linear approach would be to choose the most sensitive of the battery models, for each compound. In this way, estimation of the safest exposure limit could be assessed. Another possibility would be to assign different weight to the different test methods depending from their characteristics and use it to calculate a probability or to assign a score to the single compound, which will indicate its toxic potential. (Hartung et al. 2013b)

3. Novel approaches: extensive hit follow-up and in vivo in vitro extrapolation analysis

While the final purpose of most screenings is limited to hit finding, one additional point, which characterized our test battery approach, was a richer hit follow-up plan of the identified NCC toxicants. We performed the additional characterization of the positive compounds in the same cell system as in the screening, anchoring the observed phenotypical changes to other mechanistic and transcriptional information. The hit follow-up included several parts*: in vitro in vivo* extrapolation (IVIVE); transcriptome profiling; hit confirmation by the use of similar compounds and different functional assays; and investigation of the underlying molecular mechanisms of the inhibition of migration for selected hits (Figure 11).

Our NCC-based system offered all the necessary characteristics to perform the different follow-up studies: defined culture medium conditions, necessary for correct *in vitro-in vivo* concentration extrapolation; high robustness in transcriptome studies, allowing to obtain

reproducible results, which are necessary for interpretation of high-information content data; and easy transferability in different functional assays.

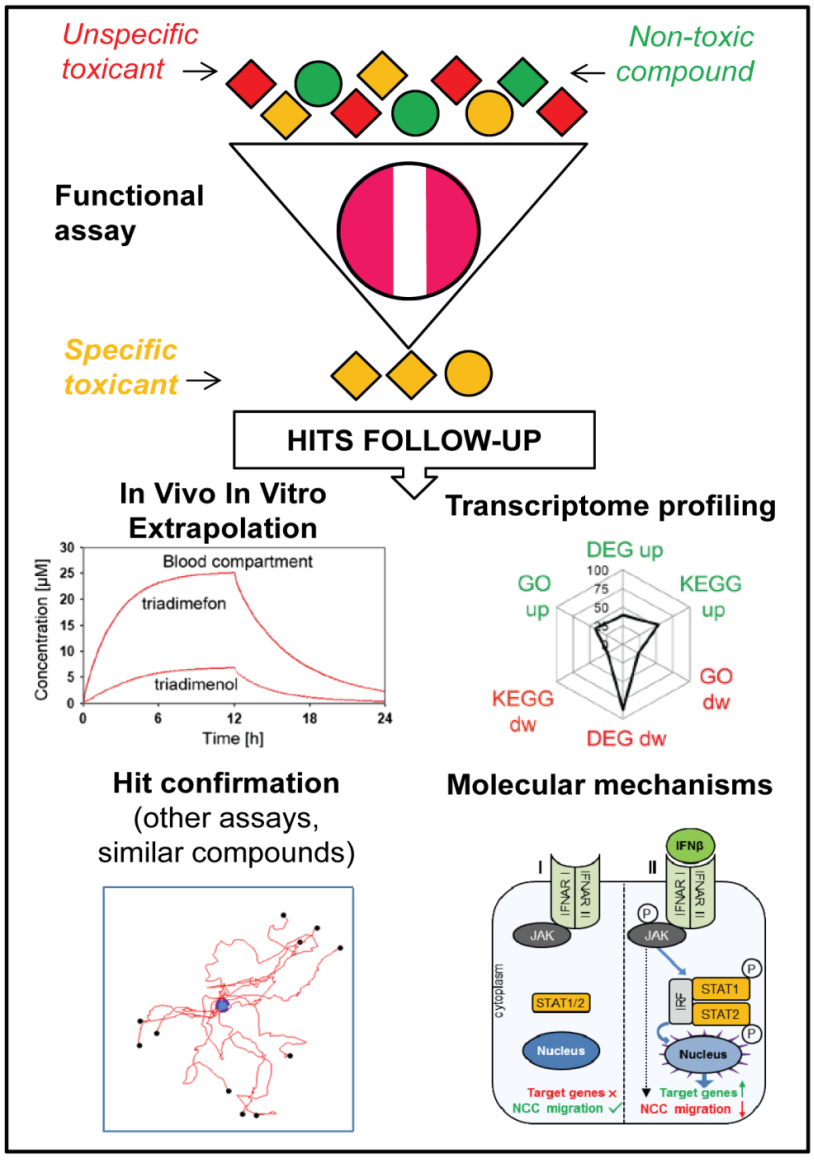


Figure 11 Hit follow-up. Prioritization of a large group of potential DT can be assessed by the use of test systems based on relevant functional assays, e.g. MINC assay. Only the compound triggering to specific toxicity will be used for the hits follow-up. This further characterization counts on different aspects: in *vivo-in vitro* extrapolation (IVIVE); transcriptome analysis; hit confirmation by the use of other assays and similar compounds; and identification of the underlying molecular mechanisms of the toxic effect.

3.1 In vitro in vivo extrapolation (IVIVE)

First step of the hit follow-up was the evaluation of the clinical relevance of the *in vitro* concentrations found to impair the NCC migration. To address this issue we performed *in vitro in vivo* extrapolation (IVIVE) and physiology-based pharmacokinetic (PBPK) analysis (Figure 12). A three-step PBPK modeling strategy was used to determine the free plasma concentrations of clinical, and when applicable, DT doses reported in literature. PBPK strategy consisted in choosing an appropriate absorption, distribution, metabolism, excretion (ADME) model; then, this model was used to simulate plasma and/or target tissue concentrations corresponding to the exposure (dose, route of administration, interval) published in *in vivo* studies. Finally, the nominal concentration *in vitro* that has the same unbound concentration as the toxic concentration *in vivo* was calculated.

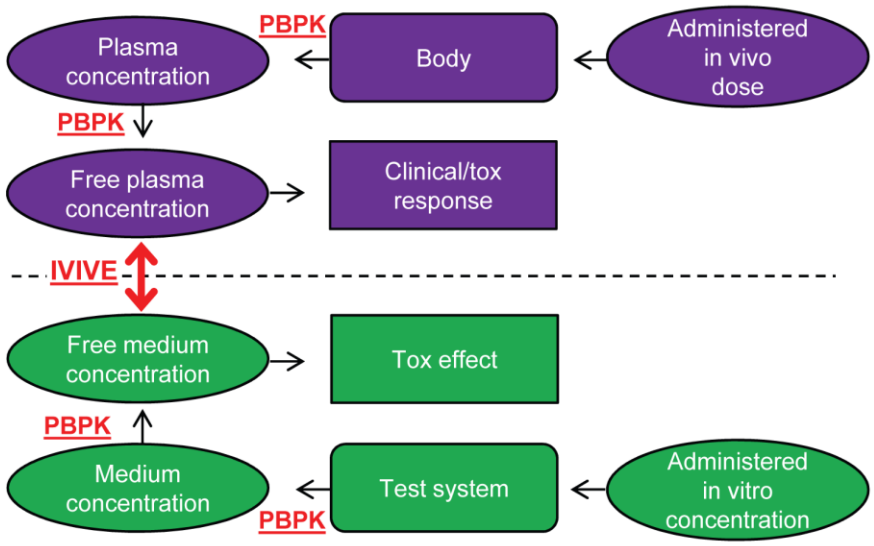


Figure 12 IVIVE, in vitro in vivo extrapolation. In vivo, the exposure to a determined dose of a drug or pollutant can lead to a clinical or toxic response. By PBPK analysis, the actual plasma and free plasma concentrations can be estimated. In a *in vitro* model, the exposure to a certain concentration of the same substance can lead to a toxic effect. PBPK analysis allows to calculate the medium and the free medium concentrations. Free plasma and free medium concentrations can then be compared.

This approach allows to consider the actual impact of a hit on risk assessment, since a compound may have toxic effects (hazard) but not be considered as a risk, because the normal level of exposure are lower than the toxic doses.

3.1.1 Challenges related to IVIVE

Challenges of IVIVE are mostly linked to the collection of already available *in vivo* and epidemiologic studies, which are often out-to-date, and may show no significant results but often only a plausible trend. Incomplete and contradictory data must be correctly weighed and critically interpreted. Furthermore, human DT is reported only in few epidemiological studies. Finally, for most compounds, we must still rely on developmental toxicity studies which are based on animal models for which the relevance on human hazard is not elsewhere proven.

4 Novel approaches: transcriptome profiling of identified DT toxicants

In the second part of the results, transcriptome analysis of the different hits identified in the manuscript 1 was performed, as part of the hit follow up strategy.

The rationale of this analysis grounds on the advantage of genomic toxicology to offer high sensitivity and comprehensive examination of the molecular changes resulting from chemical exposure. This approach relies on the assumption that toxicity is not expected to occur without alterations at transcriptional level (Farr and Dunn 1999). This step allowed to anchor a transcriptome change to a phenotypical observation, in our case, the NCC inhibition of migration.

As one approach, the transcriptomics pattern changes induced by toxic exposure can be detected and classified in a merely descriptive way, without investigating the molecular mechanisms. Application of transcriptomics data in risk assessment can otherwise focuses on exploring the MoA of a substance and to include this information in its risk assessment evaluation, as part of weight of evidence. In fact, compared to classic toxicological studies, the toxicogenomic approach is not merely hypothesis-driven (Hirabayashi and Inoue 2002) but it can rather be utilized for new MoA hypothesis generation.

Furthermore, fingerprints of toxicogenomics responses can be used to classify compounds with similar MoA and be used for read-across studies, allowing more efficient screening and prioritization.

A further approach combines transcriptome analysis with other functional assays to link expression pattern changes to mechanistic information on cellular phenotypical perturbation

and to identify biomarkers specific to a particular class of molecular damage (Oberemm et al. 2005, Wilson et al. 2013).

4.1 Transcriptome fingerprints of NCC toxicants

In our study a large heterogeneity in transcriptome responses triggered by the NCC migration- inhibiting compounds was observed. However, toxicants known to share a known mode of action, such as VPA and TSA, also showed a similar transcriptome response. This implies that it may be possible to use transcriptome responses for toxicological grouping of compounds; i.e. that an unknown compound may be assigned to a group of already known toxicants based on shared transcriptome profiles.

Transcriptome-based grouping could be used to perform non-structural-based read-across studies, moving away from the more classical structure-based algorithms to the incorporation of biological information. An approach in this direction was made in another developmental test system described by Rempel et al. 2015 ; in this study transcriptome profiles have been used to correctly classify several compounds in two groups with known and distinguished modes of action.

Classification and grouping of toxicants represents a promising strategy for toxicological risk assessment (Gocht et al. 2015). This approach is particularly convenient in developmental toxicity testing where a large number of substances must be still evaluated for developmental toxicity effects, and it would represent an attempt to solve difficulties of data interpretation.

5. Novel challenges in toxicogenomics

The use of transcriptomics approaches in toxicology faces different challenges. A crucial starting step is the data analysis of transcriptome studies. The handling of high-content data and pattern recognition algorithms has required a strong increase of informatics and statistical support, which needs to meet with the biological background for meaningful interpretation of the results. A further challenge is the development of new tools for high-dimensional data visualization (Wilson et al. 2013). Omics approaches offer a high number and level of information, which must be ordered and dimensionally reduced for easier visualization and interpretation of the results and comparison of toxic responses.

The amount of data and the related costs of the transcriptomics technology led to an additional need of information simplification. One example is the development of new algorithms for biomarker mining. In fact, a reduced number of genes as readout would be a more applicable approach concerning time and costs for safety assessment.

Finally, a critical issue is the establishment of methods for a conversion of the descriptive data to quantifiable measures to compare and predict hazards (Waldmann et al. 2014).

5.1 Visualization of transcriptome profiles

A main question addressed in this study was how primary transcriptome information, i.e. long lists of differentially expressed genes, can be reduced to a format that is easier to handle and that can be used for toxicological purposes.

One composite measure was the number of differentially expressed genes (DEG) (single-variable representation, Figure 13). It appears evident that the information content of such an endpoint is relatively low, although there is a high likelihood that compounds that do not deregulate any gene are relatively harmless, and chemicals that deregulate very large numbers of genes may be problematic (Farr and Dunn 1999).

More information may be obtained from the examination of the biologically linked gene networks, i.e. genes belonging to one gene ontology (GO) group or KEGG pathway. An

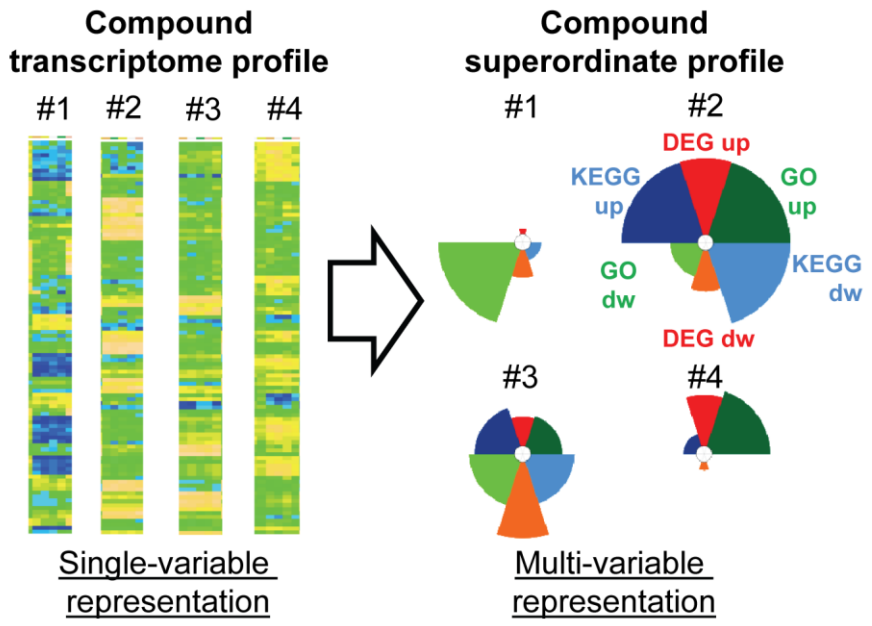


Figure 13 Visual representations of transcriptome data. Transcriptome data can be showed, e.g., as compound transcriptome profile, considering only one variable (regulated genes) or by a multi-variable representation, including also the superordinate processes which are regulated by the DEG (differentially expressed genes), such as GO class (GO) and KEGG pathways (KEGG), up- and down-regulated.

increasing number of overrepresented GO terms/KEGG pathways amongst the DEG would indicate a specific regulation of genes belonging to a certain cell function as opposed to random gene regulations. A summary of the changes across all study compounds can be obtained from such measures very quickly, e.g., in the form of ToxPi diagrams (Reif et al. 2013). In our study, the toxic profile of each compound was visualized by the number of the differentially regulated genes (DEG), together with the number of the related biological processes, such as GO classes (GO) and KEGG pathways (KEGG). (multi-variable representation, Figure 13). By this approach we could show in a compact graph the macro-effects of the different compounds on NCC transcriptome profile and compare them. Multi-variable representations easily allowed to recognize a lack of dependency among the changes of the different variables (DEG, GO, KEGG), e.g. higher number of DEG did not always correlate with an higher number of enriched GO classes or KEGG pathways. This proved the not-redundancy of the multi-variable representation and its utility in offering a more complete profile of the transcriptome effects of a particular substance.

5.2 NCC toxicity-related biomarkers: different approaches to identify relevant tox-alerts

A second major issue discussed in the second manuscript was the mining for relevant biomarkers. As already mentioned, identification of biomarkers is a very relevant process in the field of test method development since it would lead to simplification of tests, reduction of the costs and increase of the through-put.

Our approach aimed to select a defined and shorter list of biomarker genes based on some precise rules; on the one hand we were interested in those genes which were regulated by many of the compounds, and which could identify a general response of the NCC to NCC toxicants (joint- toxicant signature approach).

On the other hand, we used a filtering-ranking algorithm which takes in account different qualities of the single genes, analyzing their expression changes as well as their biological relevance within a single compound (scoring approach) (Figure 14).

An additional issue related to the biomarker identification, was their characterization. In our paper, this issue was addressed in two different ways. Firstly, the biomarker shortlist was used to repeat the principal component analysis (PCA); in this way, a clear separation of the different compounds among each other was verified to be still achievable. Secondly, a simple metric was developed for expressing the separation power (separation units, SU) of the biomarker sets compared to a distribution of randomly chosen biomarker genes. This approach allowed to show that the selected biomarker group had a higher separation power than a random group.

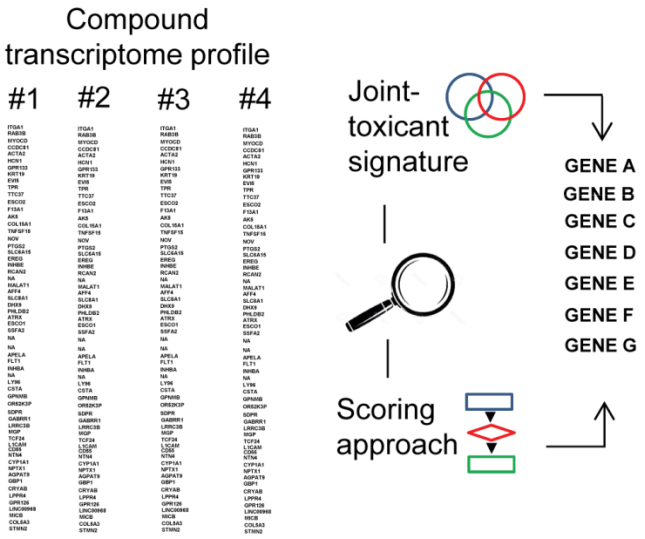


Figure 14 Biomarker mining strategies. To identify a shorter list of most relevant genes (candidate biomarkers), two different approaches have been taken. One approach was pursued by comparison of the different transcriptome profiles among the positive compounds (joint-toxicant signature). A second strategy made use of a scoring algorithm, by which assigning different weight to the single transcripts within a single transcriptome profile.

Afterwards, additional characterization was partially followed; for instance mRNA expression level of the single biomarker genes was evaluated at different concentrations and in response of additional MINC positive substances, e.g. PCB153/180, homocysteine and retinoic acid, by qPCR analysis. Although most of the analyzed gene expression levels seem to be altered in presence of the toxicants, additional analysis would be necessary to select which biomarkers might to be excluded and which ones should instead be further characterized, e.g. in other test models.

5.3 Quantification of transcriptome responses

A third issue related to the toxicogenomics approach is the quantification of the transcriptional response in dose dependent manner. Although this question was not addressed in our study, its relevance deserves to be mentioned and discussed. A major objective of toxicology and chemical risk assessment is to identify permissible exposure levels based on

data from human or experimental studies together with other relevant scientific information. (Thomas et al. 2007). For classical toxicological endpoints, such as viability, the permissible exposure levels were based on doses corresponding to LOAEL (1st dose producing a statistically significant change) or NOAEL. This approach has some relevant limitations: the dose spacing and experimental sample size can influence their evaluation. To solve these issues, the bench mark dose (BMD) approach was developed. The BMD analysis fits a statistical model to the dose response data and identifies a dose that causes a defined change in the adverse response (Figure 15).

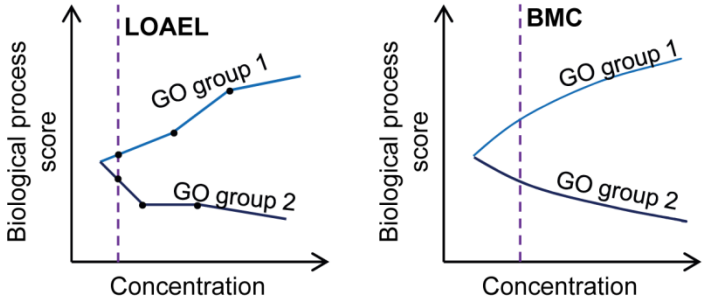


Figure 15 Benchmark approach in toxicogenomics. Representation of the BMC approach (right) compared to LOAEL evaluation (left) applied to transcriptome studies.

The application of microarray technology is useful to simultaneously measuring expression of thousands of genes and their changes in relation to toxic responses.

Transcriptome changes result from a complex mixture of primary and secondary responses to the toxic exposure. While the first reflects the direct effect on potential key events in the MoA of a chemical, the latter are caused by secondary processes which follow the initial damage. An approach for better understanding and interpreting these changes is to perform GO or pathway enrichment analysis. This type of analysis provides insights into which particular biological process is altered by a toxicant, but it has been mostly applied to studies designed to examine only a single experimental variable at a time. For multidimensional data (e.g. dose-response studies) different strategies should be followed.

One approach, described by Waldmann and colleagues (Waldmann et al. 2014), made use of biological processes (e.g. GO group) as starting point for a concentration-effect analysis. By this method, a quantitative GO activation score was assigned to each GO term and plotted on the y-axis of a concentration effect graph.

Another method was described by Yu and colleague (Yu et al. 2006). A combined average raw gene expression values (e.g., intensity or ratio) of genes associated with specific

functional categories derived from the GO database was calculated to facilitate quantitative interpretation of dose- or time-dependent genomic data.

A further step was assessed by combining the microarray technology and the GO classification analysis with the bench mark dose (BMD) method, process regularly employed by the U.S. EPA for estimating reference doses and setting standards for non-cancer human health effects (EPA 1995). This approach represents a significant step forward in applying genomic information to assessing health risks by both allowing a comprehensive survey of molecular and cellular changes associated with chemical exposure and providing the capability to identify reference doses at which cellular processes are altered (Thomas et al. 2007).

Quantitative transcriptome analysis represents a promising technology for the advancement of developmental *in vitro* testing, mostly in those tests in which an anchoring with a phenotypical or functional endpoint would allow to better understand the underlying relation between transcriptome changes and functional impairment.

6. Interferon β- induced NCC migration inhibition: proof of principle for investigation of pathway of toxicity in risk assessment

Risk assessment has been classically based on the judgment of reported observations of correlation between a toxicant exposure and an *in vivo* phenotypical alteration. This means that information regarding the molecular modes of action for many of the tested substances are not known and that this information has not been taken in consideration for decisions in risk assessment.

In parallel to the increasing establishment of *in vitro* test systems, a new risk assessment concept has been developed, the Adverse Outcome Pathway (AOP) framework. The AOP framework aims to introduce more mechanistic understanding into regulatory decision making and to facilitate a widespread systematic use of mechanistic information for regulatory safety assessment. The AOP framework rationally combines data across multiple levels of biological organization to identify correlative and causal linkages between an excessive chemical exposure and the sequence of events that lead to an adverse outcome (Ankley et al. 2010). The AOPs are built in the way to link the molecular initiating event (MIE) to the adverse outcome (AO) in an organism, by a consequent series of essential and measurable key events (KE). Central part of an AOP is the pathway of toxicity (PoT), which

starts from the molecular initiating event to the altered cellular phenotype, through the intermediary biological key events (Figure 16).

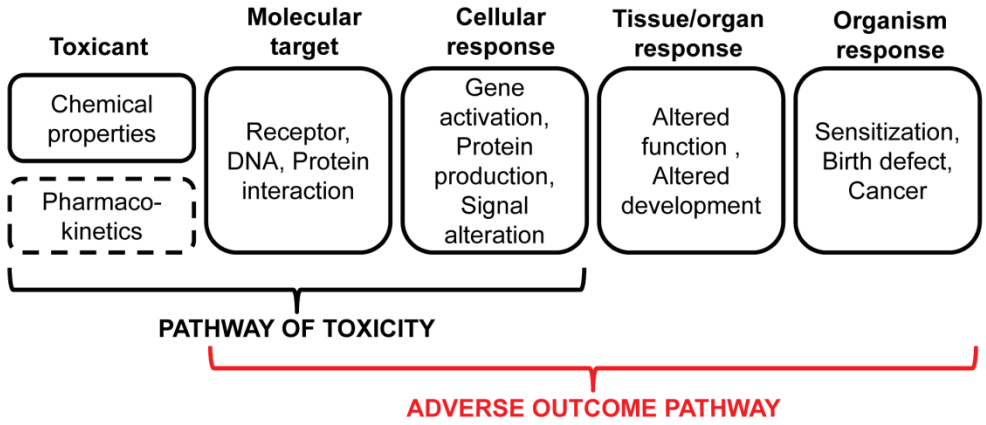


Figure 16. Structure of the Adverse Outcome Pathway. Conceptual diagram of the different steps which compose the Adverse Outcome Pathway and Pathway of Toxicity. Modified from Ankley et al. 2010

A PoT was defined by the NRC as a cellular response pathway that, when sufficiently perturbed, is expected to result in the adverse effects.

In the third manuscript of the results part, a characterization of the PoT triggering the IFNβ-induced NCC migration inhibition was attempted. We focused on the two anchoring points of the PoT: the initiating signaling events and the proximal cellular responses that can be measured and modeled *in vitro* (Figure 17).

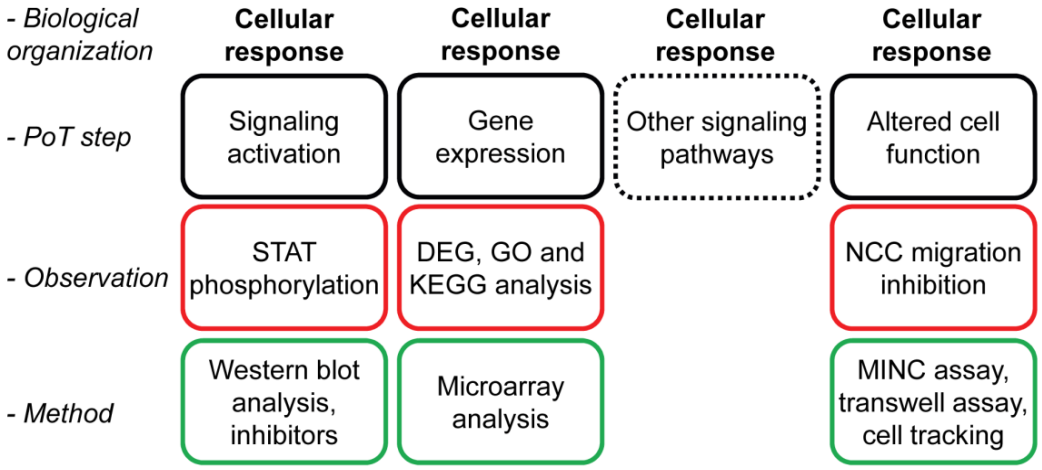


Figure 17 IFNβ- induced NCC migration inhibition pathway characterization. Schematic representation of the PoT steps that have been identified (black boxes), the respective endpoints for each point (red boxes), and the respective methods that have been used to get a quantitative measurement of each endpoint (green boxes).

Interferon β showed its migration inihibitory properties in the first MINC assay screening (Zimmer et al. 2014). To exclude the presence of artefacts, mostly due to the proliferation inhibitory effects of the cytokine, the functional alteration was then confirmed by other complementary assays, in particular, by the transwell assay, a shorter migration method, and by cell tracking.

Then, we made use of transcriptome profiling to identify potential signaling candidates. IFNβ most characteristic signaling pathway, JAK-STAT signaling pathway, was chosen among them as first attempt. We could confirm the activation of the JAK-STAT pathway and its correlation with the induced migration inhibitory effect by challenging the system with specific pathway inhibitors.

In conclusion, our human cell-based testing approach allowed important insights into potential toxicological hazard of a highly species-specific drug. The phenotypic adverse effect (inhibited migration) was consistent with a plausible biochemical mechanisms of IFNβ signalling (JAK kinase activation), confirmed by several tests and endpoints

7. Conclusions and Outlook

In this thesis, a novel approach for developmental toxicity testing has been explored. Novel DT candidates had been pre-selected and screened in a well characterized hESC-based method, the MINC assay, in the framework of the ESNATS test battery. The first publication led to the identification of 11 novel DT hits (“hit finding”, Figure 18). Next, extensive transcriptome profile characterization has been performed for some of the found hits. In the second publication, transcriptome signatures have been described and new approaches for high content information visualization have been explored. Additionally, a short list of 39 candidate biomarker genes has been selected for further characterization (“transcriptome profiling”, Figure 18).

Finally, a combined approach has been used for building the IFNβ-induced NCC migration inhibition pathway in the third manuscript. Observations obtained from functional assays and transcriptome analysis allowed to identify some of the molecular steps leading to the final phenotypical alteration (“PoT building”, Figure 18).

Each of the above mentioned studies opens additional scientific queries that would need to be further clarified. It would be of high interest to compare the toxic effects of the ESNATS test

compound group among the other systems of the test battery. This would increase the information regarding the toxic mode of actions and developmental stage-specificity. Additionally, this would add information on the complementarities and the similarities of the chosen test systems.

Further studies would be also necessary to clarify the grouping and classifying power of the transcriptome profiles, e.g. addition of more compounds with similar mode of actions and of negative controls. Moreover, extensive characterization of the found candidate biomarkers should be followed.

Finally, additional experiments would be necessary to complete the clarification of the molecular mechanisms leading to the IFNβ-induced NCC migration inhibition.

In conclusion, novel and potent technologies, such as pluripotent stem cell cultures and omics approaches, have been recently arisen and entered in the work routine of many laboratories. These techniques offer a great opportunity for the progress of developmental toxicity testing, a toxicological branch which present particular difficulties. The high potential of these approaches brings also novel challenges, which need to be addressed to finally benefit of them and improve the risk assessment of development toxic substances.

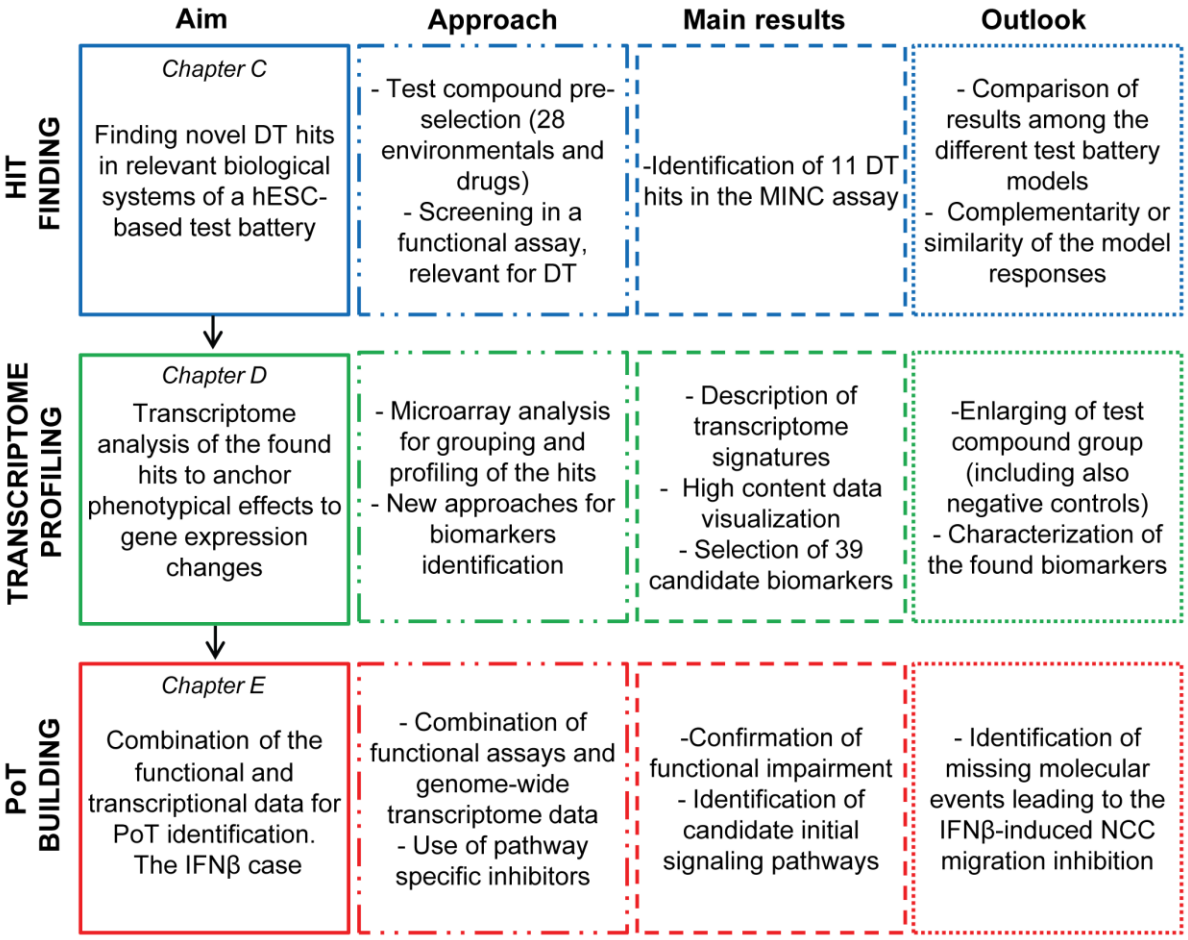


Figure 18 Schematic description of the essential points (aim, approach, main results and outlook) for the different result chapters.

G. Bibliography

Adler S, Basketter D, Creton S, Pelkonen O, van Benthem J, Zuang V, et al. (2011) Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. Arch Toxicol 85(5):367-485 doi:10.1007/s00204-011-0693-2

Alam J MA, Scaramucci J, Jones W, Rogge M (1997) Pharmacokinetics and Pharmacodynamics of Interferon Beta-la (IFNβ-1a) in Healthy Volunteers after Intravenous, Subcutaneous or Intramuscular Administration. Clin Drug Investig 14: 35 doi:10.2165/00044011-199714010-00005

Alepee N, Bahinski A, Daneshian M, De Wever B, Fritsche E, Goldberg A, et al. (2014) State-of-the-art of 3D cultures (organs-on-a-chip) in safety testing and pathophysiology. ALTEX 31(4):441-77 doi:http://dx.doi.org/10.14573/altex1406111

Alexa A, Rahnenfuhrer J, Lengauer T (2006) Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics 22(13):1600-7 doi:10.1093/bioinformatics/btl140

Alm H, Scholz B, Kultima K, Nilsson A, Andren PE, Savitski MM, et al. (2010) In vitro neurotoxicity of PBDE-99: immediate and concentration-dependent effects on protein expression in cerebral cortex cells. J Proteome Res 9(3):1226-35 doi:10.1021/pr900723c

Amato MP, Portaccio E, Ghezzi A, Hakiki B, Zipoli V, Martinelli V, et al. (2010) Pregnancy and fetal outcomes after interferon-beta exposure in multiple sclerosis. Neurology 75(20):1794-802 doi:10.1212/WNL.0b013e3181fd62bb

Amiel J, Sproat-Emison E, Garcia-Barcelo M, Lantieri F, Burzynski G, Borrego S, et al. (2008) Hirschsprung disease, associated syndromes and genetics: a review. J Med Genet 45(1):1-14 doi:10.1136/jmg.2007.053959

Ankley GT, Bennett RS, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, et al. (2010) Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. Environ Toxicol Chem 29(3):730-41 doi:10.1002/etc.34

Astro V, Chiaretti S, Magistrati E, Fivaz M, de Curtis I (2014) Liprin-alpha1, ERC1 and LL5 define polarized and dynamic structures that are implicated in cell migration. J Cell Sci 127(Pt 17):3862-76 doi:10.1242/jcs.155663

Attene-Ramos MS, Huang R, Sakamuru S, Witt KL, Beeson GC, Shou L, et al. (2013) Systematic study of mitochondrial toxicity of environmental chemicals using quantitative high throughput screening. Chem Res Toxicol 26(9):1323-32 doi:10.1021/tx4001754

Au WY, Kwong YL (2008) Arsenic trioxide: safety issues and their management. Acta Pharmacol Sin 29(3):296-304 doi:10.1111/j.1745-7254.2008.00771.x

Baggiolini A, Varum S, Mateos JM, Bettosini D, John N, Bonalli M, et al. (2015) Premigratory and migratory neural crest cells are multipotent in vivo. Cell Stem Cell 16(3):314-22 doi:10.1016/j.stem.2015.02.017

Bal-Price A, Crofton KM, Leist M, Allen S, Arand M, Buetler T, et al. (2015) International STakeholder NETwork (ISTNET): creating a developmental neurotoxicity (DNT) testing road map for regulatory purposes. Arch Toxicol 89(2):269-87 doi:10.1007/s00204-015-1464-2

Bal-Price AK, Coecke S, Costa L, Crofton KM, Fritsche E, Goldberg A, et al. (2012) Advancing the science of developmental neurotoxicity (DNT): testing for better safety evaluation. ALTEX 29(2):202-15

Balmer NV, Klima S, Rempel E, Ivanova VN, Kolde R, Weng MK, et al. (2014) From transient transcriptome responses to disturbed neurodevelopment: role of histone acetylation and methylation as epigenetic switch between reversible and irreversible drug effects. Arch Toxicol 88(7):1451-68 doi:10.1007/s00204-014-1279-6

Balmer NV, Weng MK, Zimmer B, Ivanova VN, Chambers SM, Nikolaeva E, et al. (2012) Epigenetic changes and disturbed neural development in a human embryonic stem cell-based model relating to the fetal valproate syndrome. Hum Mol Genet 21(18):4104-14 doi:10.1093/hmg/dds239

Barberi T, Klivenyi P, Calingasan NY, Lee H, Kawamata H, Loonam K, et al. (2003) Neural subtype specification of fertilization and nuclear transfer embryonic stem cells and application in parkinsonian mice. Nat Biotechnol 21(10):1200-7 doi:10.1038/nbt870

Barbero P, Verdun E, Bergui M, Pipieri A, Clerico M, Cucci A, et al. (2004) High-dose, frequently administered interferon beta therapy for relapsing-remitting multiple sclerosis must be maintained over the long term: the interferon beta dose-reduction study. J Neurol Sci 222(1-2):13-9 doi:10.1016/j.jns.2004.03.023

Basketter DA, Clewell H, Kimber I, Rossi A, Blaauboer B, Burrier R, et al. (2012) A roadmap for the development of alternative (non-animal) methods for systemic toxicity testing - t4 report*. ALTEX 29(1):3-91

Begitt A, Droescher M, Meyer T, Schmid CD, Baker M, Antunes F, et al. (2014) STAT1-cooperative DNA binding distinguishes type 1 from type 2 interferon signaling. Nat Immunol 15(2):168-76 doi:10.1038/ni.2794

Behl M, Hsieh JH, Shafer TJ, Mundy WR, Rice JR, Boyd WA, et al. (2015) Use of alternative assays to identify and prioritize organophosphorus flame retardants for potential developmental and neurotoxicity. Neurotoxicol Teratol doi:10.1016/j.ntt.2015.09.003

Bekisz J, Baron S, Balinsky C, Morrow A, Zoon KC (2010) Antiproliferative Properties of Type I and Type II Interferon. Pharmaceuticals (Basel) 3(4):994-1015 doi:10.3390/ph3040994

Benjamini YH, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Stat Methodol 57:289–300

Berggren E, Amcoff P, Benigni R, Blackburn K, Carney E, Cronin M, et al. (2015) Chemical Safety Assessment Using Read-Across: Assessing the Use of Novel Testing Methods to Strengthen the Evidence Base for Decision Making. Environ Health Perspect doi:10.1289/ehp.1409342

Berube NG, Smeenk CA, Picketts DJ (2000) Cell cycle-dependent phosphorylation of the ATRX protein correlates with changes in nuclear matrix and chromatin association. Hum Mol Genet 9(4):539-47

Blaauboer BJ, Boekelheide K, Clewell HJ, Daneshian M, Dingemans MM, Goldberg AM, et al. (2012) The use of biomarkers of toxicity for integrating in vitro hazard estimates into risk assessment for humans. ALTEX 29(4):411-25

Booy S, van Eijck CH, Janssen JA, Dogan F, van Koetsveld PM, Hofland LJ (2015) IFN-beta is a potent inhibitor of insulin and insulin like growth factor stimulated proliferation and migration in human pancreatic cancer cells. Am J Cancer Res 5(6):2035-46

Borges NC, Rezende VM, Santana JM, Moreira RP, Moreira RF, Moreno P, et al. (2011) Chlorpromazine quantification in human plasma by UPLC-electrospray ionization tandem mass spectrometry. Application to a comparative pharmacokinetic study. J Chromatogr B Analyt Technol Biomed Life Sci 879(31):3728-34 doi:10.1016/j.jchromb.2011.10.020

Bosgra S, van Eijkeren J, Bos P, Zeilmaker M, Slob W (2012) An improved model to predict physiologically based model parameters and their inter-individual variability from anthropometry. Crit Rev Toxicol 42(9):751-67 doi:10.3109/10408444.2012.709225

Bouhifd M, Hartung T, Hogberg HT, Kleensang A, Zhao L (2013) Review: toxicometabolomics. J Appl Toxicol 33(12):1365-83 doi:10.1002/jat.2874

Bourdon-Lacombe JA, Moffat ID, Deveau M, Husain M, Auerbach S, Krewski D, et al. (2015) Technical guide for applications of gene expression profiling in human health risk assessment of environmental chemicals. Regul Toxicol Pharmacol 72(2):292-309 doi:10.1016/j.yrtph.2015.04.010

Boyle CA, Boulet S, Schieve LA, Cohen RA, Blumberg SJ, Yeargin-Allsopp M, et al. (2011) Trends in the prevalence of developmental disabilities in US children, 1997-2008. Pediatrics 127(6):1034-42 doi:10.1542/peds.2010-2989

Branchi I, Capone F, Vitalone A, Madia F, Santucci D, Alleva E, et al. (2005) Early developmental exposure to BDE 99 or Aroclor 1254 affects neurobehavioural profile: interference from the administration route. Neurotoxicology 26(2):183-92 doi:10.1016/j.neuro.2004.11.005

Brask J, Kristensson K, Hill RH (2004) Exposure to interferon-gamma during synaptogenesis increases inhibitory activity after a latent period in cultured rat hippocampal neurons. Eur J Neurosci 19(12):3193-201 doi:10.1111/j.0953-816X.2004.03445.x

Bremer S, Hartung T (2004) The use of embryonic stem cells for regulatory developmental toxicity testing in vitro--the current status of test development. Curr Pharm Des 10(22):2733-47

Brown RP, Delp MD, Lindstedt SL, Rhomberg LR, Beliles RP (1997) Physiological parameter values for physiologically based pharmacokinetic models. Toxicol Ind Health 13(4):407-84

Burczynski ME, McMillian M, Ciervo J, Li L, Parker JB, Dunn RT, 2nd, et al. (2000) Toxicogenomics-based discrimination of toxic mechanism in HepG2 human hepatoma cells. Toxicol Sci 58(2):399-415

Carlson M (2015) hgu133plus2.db: Affymetrix Human Genome U133 Plus 2.0 Array annotation data (chip hgu133plus2). R package version 3.0.0. http://biocismacjp/packages/30/data/annotation/manuals/hgu133plus2db/man/hgu133plus2dbpdf

Chen LJ, Lebetkin EH, Sanders JM, Burka LT (2006) Metabolism and disposition of 2,2',4,4',5-pentabromodiphenyl ether (BDE99) following a single or repeated administration to rats or mice. Xenobiotica 36(6):515-34 doi:10.1080/00498250600674477

Coecke S, Balls M, Bowe G, Davis J, Gstraunthaler G, Hartung T, et al. (2005) Guidance on good cell culture practice. a report of the second ECVAM task force on good cell culture practice. Altern Lab Anim 33(3):261-87

Coecke S, Goldberg AM, Allen S, Buzanska L, Calamandrei G, Crofton K, et al. (2007) Workgroup report: incorporating in vitro alternative methods for developmental neurotoxicity into international hazard and risk assessment strategies. Environ Health Perspect 115(6):924-31 doi:10.1289/ehp.9427

Coelho LF, Magno de Freitas Almeida G, Mennechet FJ, Blangy A, Uze G (2005) Interferon-alpha and -beta differentially regulate osteoclastogenesis: role of differential induction of chemokine CXCL11 expression. Proc Natl Acad Sci U S A 102(33):11917-22 doi:10.1073/pnas.0502188102

Colleoni S, Galli C, Gaspar JA, Meganathan K, Jagtap S, Hescheler J, et al. (2011) Development of a neural teratogenicity test based on human embryonic stem cells: response to retinoic acid exposure. Toxicol Sci 124(2):370-7 doi:10.1093/toxsci/kfr245

Colleoni S, Galli C, Gaspar JA, Meganathan K, Jagtap S, Hescheler J, et al. (2012) Characterisation of a neural teratogenicity assay based on human ESCs differentiation following exposure to valproic acid. Curr Med Chem 19(35):6065-71

Costa LG, Giordano G, Tagliaferri S, Caglieri A, Mutti A (2008) Polybrominated diphenyl ether (PBDE) flame retardants: environmental contamination, human body burden and potential adverse health effects. Acta Biomed 79(3):172-83

Covaci A, Jorens P, Jacquemyn Y, Schepens P (2002) Distribution of PCBs and organochlorine pesticides in umbilical cord and maternal serum. Sci Total Environ 298(1-3):45-53

Crofton KM, Mundy WR, Lein PJ, Bal-Price A, Coecke S, Seiler AE, et al. (2011) Developmental neurotoxicity testing: recommendations for developing alternative methods for the screening and prioritization of chemicals. ALTEX 28(1):9-15

Crowell SR, Henderson WM, Kenneke JF, Fisher JW (2011) Development and application of a physiologically based pharmacokinetic model for triadimefon and its metabolite triadimenol in rats and humans. Toxicol Lett 205(2):154-62 doi:10.1016/j.toxlet.2011.05.1036

da Silva AJ, Brickelmaier M, Majeau GR, Lukashin AV, Peyman J, Whitty A, et al. (2002) Comparison of gene expression patterns induced by treatment of human umbilical vein endothelial cells with IFN-alpha 2b vs. IFN-beta 1a: understanding the functional relationship between distinct type I interferons that act through a common receptor. J Interferon Cytokine Res 22(2):173-88 doi:10.1089/107999002753536149

Darnerud PO (2008) Brominated flame retardants as possible endocrine disrupters. Int J Androl 31(2):152-60 doi:10.1111/j.1365-2605.2008.00869.x

Daston GP, Baines D, Yonker JE (1991) Chick embryo neural retina cell culture as a screen for developmental toxicity. Toxicol Appl Pharmacol 109(2):352-66

de Jong E, Louisse J, Verwei M, Blaauboer BJ, van de Sandt JJ, Woutersen RA, et al. (2009) Relative developmental toxicity of glycol ether alkoxy acid metabolites in the embryonic stem cell test as compared with the in vivo potency of their parent compounds. Toxicol Sci 110(1):117-24 doi:10.1093/toxsci/kfp083

de Mey C, Nassr N, Lahu G (2011) No relevant cardiac, pharmacokinetic or safety interactions between roflumilast and inhaled formoterol in healthy subjects: an open-label, randomised, actively controlled study. BMC Clin Pharmacol 11:7 doi:10.1186/1472-6904-11-7

Dhib-Jalbut S, Marks S (2010) Interferon-beta mechanisms of action in multiple sclerosis. Neurology 74 Suppl 1:S17-24 doi:10.1212/WNL.0b013e3181c97d99

Di Gion P, Kanefendt F, Lindauer A, Scheffler M, Doroshyenko O, Fuhr U, et al. (2011) Clinical pharmacokinetics of tyrosine kinase inhibitors: focus on pyrimidines, pyridines and pyrroles. Clin Pharmacokinet 50(9):551-603 doi:10.2165/11593320-000000000-00000

Di Renzo F, Bacchetta R, Sangiorgio L, Bizzo A, Menegola E (2011a) The agrochemical fungicide triadimefon induces abnormalities in Xenopus laevis embryos. Reprod Toxicol 31(4):486-93 doi:10.1016/j.reprotox.2011.01.003

Di Renzo F, Broccia ML, Giavini E, Menegola E (2007) Antifungal triazole derivative triadimefon induces ectopic maxillary cartilage by altering the morphogenesis of the first branchial arch. Birth Defects Res B Dev Reprod Toxicol 80(1):2-11 doi:10.1002/bdrb.20097

Di Renzo F, Corsini E, Broccia ML, Marinovich M, Galli CL, Giavini E, et al. (2009) Molecular mechanism of teratogenic effects induced by the fungicide triadimefon: study of the expression of TGF-beta mRNA and TGF-beta and CRABPI proteins during rat in vitro development. Toxicol Appl Pharmacol 234(1):107-16 doi:10.1016/j.taap.2008.09.025

Di Renzo F, Rossi F, Bacchetta R, Prati M, Giavini E, Menegola E (2011b) Expression analysis of some genes regulated by retinoic acid in controls and triadimefon-exposed embryos: is the amphibian Xenopus laevis a suitable model for gene-based comparative teratology? Birth Defects Res B Dev Reprod Toxicol 92(3):189-94 doi:10.1002/bdrb.20294

Di Renzo F, Rossi F, Prati M, Giavini E, Menegola E (2011c) Early genetic control of craniofacial development is affected by the in vitro exposure of rat embryos to the fungicide triadimefon. Birth Defects Res B Dev Reprod Toxicol 92(1):77-81 doi:10.1002/bdrb.20284

Dreser N, Zimmer B, Dietz C, Sugis E, Pallocca G, Nyffeler J, et al. (2015) Grouping of histone deacetylase inhibitors and other toxicants disturbing neural crest migration by transcriptional profiling. Neurotoxicology 50:56-70 doi:10.1016/j.neuro.2015.07.008

EC (2006) http://ec.europa.eu/environment/chemicals/reach/pdf/reach.pdf.

Efremova L, Schildknecht S, Adam M, Pape R, Gutbier S, Hanf B, et al. (2015) Prevention of the degeneration of human dopaminergic neurons in an astrocyte co-culture system allowing endogenous drug metabolism. Br J Pharmacol 172(16):4119-32 doi:10.1111/bph.13193

Egorin MJ, Lagattuta TF, Hamburger DR, Covey JM, White KD, Musser SM, et al. (2002) Pharmacokinetics, tissue distribution, and metabolism of 17-(dimethylaminoethylamino)-17-demethoxygeldanamycin (NSC 707545) in CD2F1 mice and Fischer 344 rats. Cancer Chemother Pharmacol 49(1):7-19

Eisenberg S, Giehl K, Henis YI, Ehrlich M (2008) Differential interference of chlorpromazine with the membrane interactions of oncogenic K-Ras and its effects on cell growth. J Biol Chem 283(40):27279-88 doi:10.1074/jbc.M804589200

Ejlerskov P, Hultberg JG, Wang J, Carlsson R, Ambjorn M, Kuss M, et al. (2015) Lack of Neuronal IFN-beta-IFNAR Causes Lewy Body- and Parkinson's Disease-like Dementia. Cell 163(2):324-39 doi:10.1016/j.cell.2015.08.069

El-Hachem N, Grossmann P, Blanchet-Cohen A, Bateman AR, Bouchard N, Archambault J, et al. (2015) Characterization of Conserved Toxicogenomic Responses in Chemically Exposed Hepatocytes across Species and Platforms. Environ Health Perspect doi:10.1289/ehp.1409157

EPA (1995) U.S. Environmental Protection Agency. The Use of the Benchmark Dose Approach in Health Risk Assessment.

EPA U (1996) https://www.epa.gov/sites/production/files/2014-11/documents/guidelines_repro_toxicity.pdf.

Eriksson P, Fischer C, Fredriksson A (2006) Polybrominated diphenyl ethers, a group of brominated flame retardants, can interact with polychlorinated biphenyls in enhancing developmental neurobehavioral defects. Toxicol Sci 94(2):302-9 doi:10.1093/toxsci/kfl109

Eriksson P, Viberg H, Jakobsson E, Orn U, Fredriksson A (2002) A brominated flame retardant, 2,2',4,4',5-pentabromodiphenyl ether: uptake, retention, and induction of neurobehavioral alterations in mice during a critical phase of neonatal brain development. Toxicol Sci 67(1):98-103

Eskenazi B, Chevrier J, Rauch SA, Kogut K, Harley KG, Johnson C, et al. (2013) In utero and childhood polybrominated diphenyl ether (PBDE) exposures and neurodevelopment in the CHAMACOS study. Environ Health Perspect 121(2):257-62 doi:10.1289/ehp.1205597

EUROSTAT (2010) EUROPEAN PERINATAL HEALTH REPORT: Health and Care of Pregnant Women and Babies in Europe in 2010.

Fabian M, Forsling ML, Jones JJ, Lee J (1969) The release, clearance and plasma protein binding of oxytocin in the anaesthetized rat. J Endocrinol 43(2):175-89

Falsig J, Latta M, Leist M (2004a) Defined inflammatory states in astrocyte cultures: correlation with susceptibility towards CD95-driven apoptosis. J Neurochem 88(1):181-93

Falsig J, Porzgen P, Lotharius J, Leist M (2004b) Specific modulation of astrocyte inflammation by inhibition of mixed lineage kinases with CEP-1347. J Immunol 173(4):2762-70

Falsig J, Porzgen P, Lund S, Schrattenholz A, Leist M (2006) The inflammatory transcriptome of reactive murine astrocytes and implications for their innate immune function. J Neurochem 96(3):893-907 doi:10.1111/j.1471-4159.2005.03622.x

Farr S, Dunn RT, 2nd (1999) Concise review: gene expression applied to toxicology. Toxicol Sci 50(1):1-9

FDA (1999) http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/TherapeuticBiologicApplications/ucm106138.pdf.

Fielden MR, Zacharewski TR (2001) Challenges and limitations of gene expression profiling in mechanistic and predictive toxicology. Toxicol Sci 60(1):6-10

Fietz SA, Lachmann R, Brandl H, Kircher M, Samusik N, Schroder R, et al. (2012) Transcriptomes of germinal zones of human and mouse fetal neocortex suggest a role of extracellular matrix in progenitor self-renewal. Proc Natl Acad Sci U S A 109(29):11836-41 doi:10.1073/pnas.1209647109

Filer D, Patisaul HB, Schug T, Reif D, Thayer K (2014) Test driving ToxCast: endocrine profiling for 1858 chemicals included in phase II. Curr Opin Pharmacol 19:145-52 doi:10.1016/j.coph.2014.09.021

Floris S, Ruuls SR, Wierinckx A, van der Pol SM, Dopp E, van der Meide PH, et al. (2002) Interferon-beta directly influences monocyte infiltration into the central nervous system. J Neuroimmunol 127(1-2):69-79

Foster PM, Creasy DM, Foster JR, Gray TJ (1984) Testicular toxicity produced by ethylene glycol monomethyl and monoethyl ethers in the rat. Environ Health Perspect 57:207-17

Fritsche E, Cline JE, Nguyen NH, Scanlan TS, Abel J (2005) Polychlorinated biphenyls disturb differentiation of normal human neural progenitor cells: clue for involvement of thyroid hormone receptors. Environ Health Perspect 113(7):871-6

Fritsche E, Gassmann K, Schreiber T (2011) Neurospheres as a model for developmental neurotoxicity testing. Methods Mol Biol 758:99-114 doi:10.1007/978-1-61779-170-3_7

Fuller LC, Cornelius SK, Murphy CW, Wiens DJ (2002) Neural crest cell motility in valproic acid. Reprod Toxicol 16(6):825-39

Gascon M, Fort M, Martinez D, Carsin AE, Forns J, Grimalt JO, et al. (2012) Polybrominated diphenyl ethers (PBDEs) in breast milk and neuropsychological development in infants. Environ Health Perspect 120(12):1760-5 doi:10.1289/ehp.1205266

Gerhardt E, Kugler S, Leist M, Beier C, Berliocchi L, Volbracht C, et al. (2001) Cascade of caspase activation in potassium-deprived cerebellar granule neurons: targets for treatment with peptide and protein inhibitors of apoptosis. Mol Cell Neurosci 17(4):717-31 doi:10.1006/mcne.2001.0962

Giavini E, Menegola E (2010) Are azole fungicides a teratogenic risk for human conceptus? Toxicol Lett 198(2):106-11 doi:10.1016/j.toxlet.2010.07.005

Gilbert SF (2010) Developmental biology, 9th edn. Sinauer Associates, Sunderland, Mass.

Gocht T, Berggren E, Ahr HJ, Cotgreave I, Cronin MT, Daston G, et al. (2015) The SEURAT-1 approach towards animal free human safety assessment. ALTEX 32(1):9-24 doi:http://dx.doi.org/10.14573/altex.1408041

Godoy P, Hewitt NJ, Albrecht U, Andersen ME, Ansari N, Bhattacharya S, et al. (2013) Recent advances in 2D and 3D in vitro systems using primary hepatocytes, alternative hepatocyte sources and non-parenchymal liver cells and their use in investigating mechanisms of hepatotoxicity, cell signaling and ADME. Arch Toxicol 87(8):1315-530 doi:10.1007/s00204-013-1078-5

Goetz AK, Ren H, Schmid JE, Blystone CR, Thillainadarajah I, Best DS, et al. (2007) Disruption of testosterone homeostasis as a mode of action for the reproductive toxicity of triazole fungicides in the male rat. Toxicol Sci 95(1):227-39 doi:10.1093/toxsci/kfl124

Golub MS, Macintosh MS, Baumrind N (1998) Developmental and reproductive toxicity of inorganic arsenic: animal studies and human concerns. J Toxicol Environ Health B Crit Rev 1(3):199-241 doi:10.1080/10937409809524552

Govarts E, Nieuwenhuijsen M, Schoeters G, Ballester F, Bloemen K, de Boer M, et al. (2012) Birth weight and prenatal exposure to polychlorinated biphenyls (PCBs) and dichlorodiphenyldichloroethylene (DDE): a meta-analysis within 12 European Birth Cohorts. Environ Health Perspect 120(2):162-70 doi:10.1289/ehp.1103767

Grandjean P, Landrigan PJ (2006) Developmental neurotoxicity of industrial chemicals. Lancet 368(9553):2167-78 doi:10.1016/S0140-6736(06)69665-7

Green SA, Simoes-Costa M, Bronner ME (2015) Evolution of vertebrates as viewed from the crest. Nature 520(7548):474-82 doi:10.1038/nature14436

Grillo JA, Zhao P, Bullock J, Booth BP, Lu M, Robie-Suh K, et al. (2012) Utility of a physiologically-based pharmacokinetic (PBPK) modeling approach to quantitatively predict a complex drug-drug-disease interaction scenario for rivaroxaban during the drug review process: implications for clinical practice. Biopharm Drug Dispos 33(2):99-110 doi:10.1002/bdd.1771

Grinberg M, Stober RM, Edlund K, Rempel E, Godoy P, Reif R, et al. (2014) Toxicogenomics directory of chemically exposed human hepatocytes. Arch Toxicol 88(12):2261-87 doi:10.1007/s00204-014-1400-x

Grumbach IM, Fish EN, Uddin S, Majchrzak B, Colamonici OR, Figulla HR, et al. (1999) Activation of the Jak-Stat pathway in cells that exhibit selective sensitivity to the antiviral effects of IFN-beta compared with IFN-alpha. J Interferon Cytokine Res 19(7):797-801 doi:10.1089/107999099313659

Gundling WE, Jr., Wildman DE (2015) A review of inter- and intraspecific variation in the eutherian placenta. Philos Trans R Soc Lond B Biol Sci 370(1663):20140072 doi:10.1098/rstb.2014.0072

Gurav S, Punde R, Farooqui J, Zainuddin M, Rajagopal S, Mullangi R (2012) Development and validation of a highly sensitive method for the determination of abiraterone in rat and human plasma by LC-MS/MS-ESI: application to a pharmacokinetic study. Biomed Chromatogr 26(6):761-8 doi:10.1002/bmc.1726

Habgood MD, Sedgwick JE, Dziegielewska KM, Saunders NR (1992) A developmentally regulated blood-cerebrospinal fluid transfer mechanism for albumin in immature rats. J Physiol 456:181-92

Hakk H, Larsen G, Klasson-Wehler E (2002) Tissue disposition, excretion and metabolism of 2,2',4,4',5-pentabromodiphenyl ether (BDE-99) in the male Sprague-Dawley rat. Xenobiotica 32(5):369-82 doi:10.1080/00498250110119117

Hansson O, Nylandsted J, Castilho RF, Leist M, Jaattela M, Brundin P (2003) Overexpression of heat shock protein 70 in R6/2 Huntington's disease mice has only modest effects on disease progression. Brain Res 970(1-2):47-57

Harbron C, Chang KM, South MC (2007) RefPlus: an R package extending the RMA Algorithm. Bioinformatics 23(18):2493-4 doi:10.1093/bioinformatics/btm357

Hareng L, Pellizzer C, Bremer S, Schwarz M, Hartung T (2005) The integrated project ReProTect: a novel approach in reproductive toxicity hazard assessment. Reprod Toxicol 20(3):441-52 doi:10.1016/j.reprotox.2005.04.003

Harrill JA, Robinette BL, Mundy WR (2011) Use of high content image analysis to detect chemical-induced changes in synaptogenesis in vitro. Toxicol In Vitro 25(1):368-87 doi:10.1016/j.tiv.2010.10.011

Hartung T (2009) Toxicology for the twenty-first century. Nature 460(7252):208-12 doi:10.1038/460208a

Hartung T, Hoffmann S, Stephens M (2013a) Mechanistic validation. ALTEX 30(2):119-30

Hartung T, Luechtefeld T, Maertens A, Kleensang A (2013b) Integrated testing strategies for safety assessments. ALTEX 30(1):3-18

Hartung T, van Vliet E, Jaworska J, Bonilla L, Skinner N, Thomas R (2012) Systems toxicology. ALTEX 29(2):119-28

He P, Wang A, Niu Q, Guo L, Xia T, Chen X (2011) Toxic effect of PBDE-47 on thyroid development, learning, and memory, and the interaction between PBDE-47 and PCB153 that enhances toxicity in rats. Toxicol Ind Health 27(3):279-88 doi:10.1177/0748233710387002

Heim A, Stille-Seigener M, Pring-Akerblom P, Grumbach I, Brehm C, Kreuzer H, et al. (1996) Recombinant Interferons beta and gamma have a higher antiviral activity than interferon-alpha in coxsackievirus B3-infected carrier state cultures of human myocardial fibroblasts. J Interferon Cytokine Res 16(4):283-7 doi:10.1089/jir.1996.16.283

Henley DV, Korach KS (2010) Physiological effects and mechanisms of action of endocrine disrupting chemicals that alter estrogen signaling. Hormones (Athens) 9(3):191-205

Henn A, Kirner S, Leist M (2011) TLR2 hypersensitivity of astrocytes as functional consequence of previous inflammatory episodes. J Immunol 186(5):3237-47 doi:10.4049/jimmunol.1002787

Herman GA, Bergman A, Liu F, Stevens C, Wang AQ, Zeng W, et al. (2006a) Pharmacokinetics and pharmacodynamic effects of the oral DPP-4 inhibitor sitagliptin in middle-aged obese subjects. J Clin Pharmacol 46(8):876-86 doi:10.1177/0091270006289850

Herman GA, Bergman A, Stevens C, Kotey P, Yi B, Zhao P, et al. (2006b) Effect of single oral doses of sitagliptin, a dipeptidyl peptidase-4 inhibitor, on incretin and plasma glucose levels after an oral glucose tolerance test in patients with type 2 diabetes. J Clin Endocrinol Metab 91(11):4612-9 doi:10.1210/jc.2006-1009

Hermsen SA, Pronk TE, van den Brandhof EJ, van der Ven LT, Piersma AH (2012) Triazole-induced gene expression changes in the zebrafish embryo. Reprod Toxicol 34(2):216-24 doi:10.1016/j.reprotox.2012.05.093

Hermsen SA, Pronk TE, van den Brandhof EJ, van der Ven LT, Piersma AH (2013) Transcriptomic analysis in the developing zebrafish embryo after compound

exposure: individual gene expression and pathway regulation. Toxicol Appl Pharmacol 272(1):161-71 doi:10.1016/j.taap.2013.05.037

Hermsen SA, van den Brandhof EJ, van der Ven LT, Piersma AH (2011) Relative embryotoxicity of two classes of chemicals in a modified zebrafish embryotoxicity test and comparison with their in vivo potencies. Toxicol In Vitro 25(3):745-53 doi:10.1016/j.tiv.2011.01.005

Heusinkveld HJ, Molendijk J, van den Berg M, Westerink RH (2013a) Azole fungicides disturb intracellular Ca2+ in an additive manner in dopaminergic PC12 cells. Toxicol Sci 134(2):374-81 doi:10.1093/toxsci/kft119

Heusinkveld HJ, Molendijk J, van den Berg M, Westerink RH (2013b) Azole Fungicides Disturb Intracellular Ca2+ in an Additive Manner in Dopaminergic PC12 Cells. Toxicol Sci doi:10.1093/toxsci/kft119

Hirabayashi Y, Inoue T (2002) [Toxicogenomics--a new paradigm of toxicology and birth of reverse toxicology]. Kokuritsu Iyakuhin Shokuhin Eisei Kenkyusho Hokoku(120):39-52

Hoelting L, Klima S, Karreman C, Grinberg M, Meisig J, Henry M, et al. (2016) Stem Cell-Derived Immature Human Dorsal Root Ganglia Neurons to Identify Peripheral Neurotoxicants. Stem Cells Transl Med 5(4):476-87 doi:10.5966/sctm.2015-0108

Hogberg HT, Kinsner-Ovaskainen A, Coecke S, Hartung T, Bal-Price AK (2010) mRNA expression is a relevant tool to identify developmental neurotoxicants using an in vitro approach. Toxicol Sci 113(1):95-115 doi:10.1093/toxsci/kfp175

Hogberg HT, Sobanski T, Novellino A, Whelan M, Weiss DG, Bal-Price AK (2011) Application of micro-electrode arrays (MEAs) as an emerging technology for developmental neurotoxicity: evaluation of domoic acid-induced effects in primary cultures of rat cortical neurons. Neurotoxicology 32(1):158-68 doi:10.1016/j.neuro.2010.10.007

Hu BY, Du ZW, Zhang SC (2009) Differentiation of human oligodendrocytes from pluripotent stem cells. Nat Protoc 4(11):1614-22 doi:10.1038/nprot.2009.186

Hua H, Qin S, Rui J, Li J (2011) Pharmacokinetics of arsenic trioxide (As(2)O(3)) in Chinese primary hepatocarcinoma patients. Asian Pac J Cancer Prev 12(1):61-5

Huang X, Saint-Jeannet JP (2004) Induction of the neural crest and the opportunities of life on the edge. Dev Biol 275(1):1-11 doi:10.1016/j.ydbio.2004.07.033

Jagtap S, Meganathan K, Gaspar J, Wagh V, Winkler J, Hescheler J, et al. (2011) Cytosine arabinoside induces ectoderm and inhibits mesoderm expression in human embryonic stem cells during multilineage differentiation. Br J Pharmacol 162(8):1743-56 doi:10.1111/j.1476-5381.2010.01197.x

Johansson C, Tofighi R, Tamm C, Goldoni M, Mutti A, Ceccatelli S (2006) Cell death mechanisms in AtT20 pituitary cells exposed to polychlorinated biphenyls (PCB 126 and PCB 153) and methylmercury. Toxicol Lett 167(3):183-90 doi:10.1016/j.toxlet.2006.09.006

Johansson H, Albrekt AS, Borrebaeck CA, Lindstedt M (2013) The GARD assay for assessment of chemical skin sensitizers. Toxicol In Vitro 27(3):1163-9 doi:10.1016/j.tiv.2012.05.019

Johansson H, Rydnert F, Kuhnl J, Schepky A, Borrebaeck C, Lindstedt M (2014) Genomic allergen rapid detection in-house validation--a proof of concept. Toxicol Sci 139(2):362-70 doi:10.1093/toxsci/kfu046

Jordan MK, Burstein AH, Rock-Kress D, Alfaro RM, Pau AK, Kovacs JA, et al. (2004) Plasma pharmacokinetics of sulfadiazine administered twice daily versus four times daily are similar in human immunodeficiency virus-infected patients. Antimicrob Agents Chemother 48(2):635-7

Judson R, Houck K, Martin M, Knudsen T, Thomas RS, Sipes N, et al. (2014) In vitro and modelling approaches to risk assessment from the U.S. Environmental Protection Agency ToxCast programme. Basic Clin Pharmacol Toxicol 115(1):69-76 doi:10.1111/bcpt.12239

Judson R, Kavlock R, Martin M, Reif D, Houck K, Knudsen T, et al. (2013) Perspectives on validation of high-throughput assays supporting 21st century toxicity testing. ALTEX 30(1):51-6

Judson RS, Houck KA, Kavlock RJ, Knudsen TB, Martin MT, Mortensen HM, et al. (2010) In vitro screening of environmental chemicals for targeted testing prioritization: the ToxCast project. Environ Health Perspect 118(4):485-92 doi:10.1289/ehp.0901392

Kadereit S, Zimmer B, van Thriel C, Hengstler JG, Leist M (2012) Compound selection for in vitro modeling of developmental neurotoxicity. Front Biosci (Landmark Ed) 17:2442-60

Kappos L, Weinshenker B, Pozzilli C, Thompson AJ, Dahlke F, Beckmann K, et al. (2004) Interferon beta-1b in secondary progressive MS: a combined analysis of the two trials. Neurology 63(10):1779-87

Kawashiro Y, Fukata H, Sato K, Aburatani H, Takigami H, Mori C (2009) Polybrominated diphenyl ethers cause oxidative stress in human umbilical vein endothelial cells. Hum Exp Toxicol 28(11):703-13 doi:10.1177/0960327109350669

Kern I, Xu R, Julien S, Suter DM, Preynat-Seauve O, Baquie M, et al. (2013) Embryonic stem cell-based screen for small molecules: cluster analysis reveals four response patterns in developing neural cells. Curr Med Chem 20(5):710-23

Keyte A, Hutson MR (2012) The neural crest in cardiac congenital anomalies. Differentiation 84(1):25-40 doi:10.1016/j.diff.2012.04.005

Klaric M, Winkler J, Vojnits K, Meganathan K, Jagtap S, Ensenat-Waser R, et al. (2013) Current status of human pluripotent stem cell based in vitro toxicity tests. Front Biosci (Schol Ed) 5:118-33

Kleinstreuer NC, Judson RS, Reif DM, Sipes NS, Singh AV, Chandler KJ, et al. (2011a) Environmental impact on vascular development predicted by high-throughput screening. Environ Health Perspect 119(11):1596-603 doi:10.1289/ehp.1103412

Kleinstreuer NC, Smith AM, West PR, Conard KR, Fontaine BR, Weir-Hauptman AM, et al. (2011b) Identifying developmental toxicity pathways for a subset of ToxCast chemicals using human embryonic stem cells and metabolomics. Toxicol Appl Pharmacol 257(1):111-21 doi:10.1016/j.taap.2011.08.025

Knecht AK, Bronner-Fraser M (2002) Induction of the neural crest: a multigene process. Nat Rev Genet 3(6):453-61 doi:10.1038/nrg819

Koch S, Donarski N, Goetze K, Kreckel M, Stuerenburg HJ, Buhmann C, et al. (2001) Characterization of four lipoprotein classes in human cerebrospinal fluid. J Lipid Res 42(7):1143-51

Koga F, Tsutsumi S, Neckers LM (2007) Low dose geldanamycin inhibits hepatocyte growth factor and hypoxia-stimulated invasion of cancer cells. Cell Cycle 6(11):1393-402

Kretz O, Weiss HM, Schumacher MM, Gross G (2004) In vitro blood distribution and plasma protein binding of the tyrosine kinase inhibitor imatinib and its active metabolite, CGP74588, in rat, mouse, dog, monkey, healthy humans and patients with acute lymphatic leukaemia. Br J Clin Pharmacol 58(2):212-6 doi:10.1111/j.1365-2125.2004.02117.x

Krug AK, Balmer NV, Matt F, Schonenberger F, Merhof D, Leist M (2013a) Evaluation of a human neurite growth assay as specific screen for developmental neurotoxicants. Arch Toxicol doi:10.1007/s00204-013-1072-y

Krug AK, Balmer NV, Matt F, Schonenberger F, Merhof D, Leist M (2013b) Evaluation of a human neurite growth assay as specific screen for developmental neurotoxicants. Arch Toxicol 87(12):2215-31 doi:10.1007/s00204-013-1072-y

Krug AK, Gutbier S, Zhao L, Poltl D, Kullmann C, Ivanova V, et al. (2014) Transcriptional and metabolic adaptation of human neurons to the mitochondrial toxicant MPP(+). Cell Death Dis 5:e1222 doi:10.1038/cddis.2014.166

Krug AK, Kolde R, Gaspar JA, Rempel E, Balmer NV, Meganathan K, et al. (2013c) Human embryonic stem cell-derived test systems for developmental neurotoxicity: a transcriptomics approach. Arch Toxicol 87(1):123-43 doi:10.1007/s00204-012-0967-3

Kuegler PB, Baumann BA, Zimmer B, Keller S, Marx A, Kadereit S, et al. (2012) GFAP-independent inflammatory competence and trophic functions of astrocytes generated from murine embryonic stem cells. Glia 60(2):218-28 doi:10.1002/glia.21257

Kuegler PB, Zimmer B, Waldmann T, Baudis B, Ilmjarv S, Hescheler J, et al. (2010) Markers of murine embryonic and neural stem cells, neurons and astrocytes: reference points for developmental neurotoxicity testing. ALTEX 27(1):17-42

Kummar S, Gutierrez ME, Gardner ER, Chen X, Figg WD, Zajac-Kaye M, et al. (2010) Phase I trial of 17-dimethylaminoethylamino-17-demethoxygeldanamycin (17-DMAG), a heat shock protein inhibitor, administered twice weekly in patients with advanced malignancies. Eur J Cancer 46(2):340-7 doi:10.1016/j.ejca.2009.10.026

Kurihara Y, Kurihara H, Suzuki H, Kodama T, Maemura K, Nagai R, et al. (1994) Elevated blood pressure and craniofacial abnormalities in mice deficient in endothelin-1. Nature 368(6473):703-10 doi:10.1038/368703a0

Kuriyama SN, Talsness CE, Grote K, Chahoud I (2005) Developmental exposure to low dose PBDE 99: effects on male fertility and neurobehavior in rat offspring. Environ Health Perspect 113(2):149-54

Laske C, Stellos K, Stransky E, Leyhe T, Gawaz M (2009) Decreased plasma levels of granulocyte-colony stimulating factor (G-CSF) in patients with early Alzheimer's disease. J Alzheimers Dis 17(1):115-23 doi:10.3233/JAD-2009-1017

Latta M, Kunstle G, Leist M, Wendel A (2000) Metabolic depletion of ATP by fructose inversely controls CD95- and tumor necrosis factor receptor 1-mediated hepatic apoptosis. J Exp Med 191(11):1975-85

Laurenza I, Pallocca G, Mennecozzi M, Scelfo B, Pamies D, Bal-Price A (2013a) A human pluripotent carcinoma stem cell-based model for in vitro developmental neurotoxicity testing: Effects of methylmercury, lead and aluminum evaluated by gene expression studies. Int J Dev Neurosci doi:10.1016/j.ijdevneu.2013.03.002

Laurenza I, Pallocca G, Mennecozzi M, Scelfo B, Pamies D, Bal-Price A (2013b) A human pluripotent carcinoma stem cell-based model for in vitro developmental neurotoxicity testing: effects of methylmercury, lead and aluminum evaluated by gene expression studies. Int J Dev Neurosci 31(7):679-91 doi:10.1016/j.ijdevneu.2013.03.002

Lee G, Papapetrou EP, Kim H, Chambers SM, Tomishima MJ, Fasano CA, et al. (2009) Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs. Nature 461(7262):402-6 doi:10.1038/nature08320

Lefew WR, McConnell ER, Crooks JL, Shafer TJ (2013) Evaluation of microelectrode array data using Bayesian modeling as an approach to screening and prioritization for neurotoxicity testing. Neurotoxicology 36:34-41 doi:10.1016/j.neuro.2013.02.006

Lein P, Locke P, Goldberg A (2007) Meeting report: alternatives for developmental neurotoxicity testing. Environ Health Perspect 115(5):764-8 doi:10.1289/ehp.9841

Leist M, Bremer S, Brundin P, Hescheler J, Kirkeby A, Krause KH, et al. (2008a) The biological and ethical basis of the use of human embryonic stem cells for in vitro test systems or cell therapy. ALTEX 25(3):163-90

Leist M, Efremova L, Karreman C (2010) Food for thought ... considerations and guidelines for basic test method descriptions in toxicology. ALTEX 27(4):309-17

Leist M, Hartung T, Nicotera P (2008b) The dawning of a new age of toxicology. ALTEX 25(2):103-14

Leist M, Hasiwa N, Daneshian M, Hartung T (2012a) Validation and quality control of replacement alternatives – current status and future challenges. Toxicology Research(1)

Leist M, Hasiwa N, Daneshian M, Hartung T (2012b) Validation and quality control of replacement alternatives – current status and future challenges. Toxicol Res 1:8-22 doi:10.1039/C2TX20011B

Leist M, Hasiwa N, Rovida C, Daneshian M, Basketter D, Kimber I, et al. (2014) Consensus report on the future of animal-free systemic toxicity testing. ALTEX 31(3):341-56 doi:http://dx.doi.org/10.14573/altex.1406091

Leist M, Lidbury BA, Yang C, Hayden PJ, Kelm JM, Ringeissen S, et al. (2012c) Novel technologies and an overall strategy to allow hazard assessment and risk prediction of chemicals, cosmetics, and drugs with animal-free methods. ALTEX 29(4):373-88

Leist M, Ringwald A, Kolde R, Bremer S, van Thriel C, Krause KH, et al. (2013) Test systems of developmental toxicity: state-of-the art and future perspectives. Arch Toxicol 87(12):2037-42 doi:10.1007/s00204-013-1154-x

Lesko E, Gozdzik J, Kijowski J, Jenner B, Wiecha O, Majka M (2007) HSP90 antagonist, geldanamycin, inhibits proliferation, induces apoptosis and blocks migration of rhabdomyosarcoma cells in vitro and seeding into bone marrow in vivo. Anticancer Drugs 18(10):1173-81 doi:10.1097/CAD.0b013e3282ef532d

Lin S, Lin Y, Nery JR, Urich MA, Breschi A, Davis CA, et al. (2014) Comparison of the transcriptional landscapes between human and mouse tissues. Proc Natl Acad Sci U S A 111(48):17224-9 doi:10.1073/pnas.1413624111

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25(4):402-8 doi:10.1006/meth.2001.1262

Lou J, Gasche Y, Zheng L, Giroud C, Morel P, Clements J, et al. (1999) Interferon-beta inhibits activated leukocyte migration through human brain microvascular endothelial cell monolayer. Lab Invest 79(8):1015-25

Low Y, Sedykh A, Fourches D, Golbraikh A, Whelan M, Rusyn I, et al. (2013) Integrative chemical-biological read-across approach for chemical hazard classification. Chem Res Toxicol 26(8):1199-208 doi:10.1021/tx400110f

Lu E, Wang BW, Guimond C, Synnes A, Sadovnick D, Tremlett H (2012) Disease-modifying drugs for multiple sclerosis in pregnancy: a systematic review. Neurology 79(11):1130-5 doi:10.1212/WNL.0b013e3182698c64

Machera K (1995) Developmental toxicity of cyproconazole, an inhibitor of fungal ergosterol biosynthesis, in the rat. Bull Environ Contam Toxicol 54(3):363-9

Madia F, Giordano G, Fattori V, Vitalone A, Branchi I, Capone F, et al. (2004) Differential in vitro neurotoxicity of the flame retardant PBDE-99 and of the PCB Aroclor 1254 in human astrocytoma cells. Toxicol Lett 154(1-2):11-21 doi:10.1016/j.toxlet.2004.06.013

Maertens A, Luechtefeld T, Kleensang A, Hartung T (2015) MPTP's pathway of toxicity indicates central role of transcription factor SP1. Arch Toxicol 89(5):743-55 doi:10.1007/s00204-015-1509-6

Mager DE, Jusko WJ (2002) Receptor-mediated pharmacokinetic/pharmacodynamic model of interferon-beta 1a in humans. Pharm Res 19(10):1537-43

Mager DE, Neuteboom B, Efthymiopoulos C, Munafo A, Jusko WJ (2003) Receptor-mediated pharmacokinetics and pharmacodynamics of interferon-beta1a in monkeys. J Pharmacol Exp Ther 306(1):262-70 doi:10.1124/jpet.103.049502

Makris SL, Raffaele K, Allen S, Bowers WJ, Hass U, Alleva E, et al. (2009) A retrospective performance assessment of the developmental neurotoxicity study in support of OECD test guideline 426. Environ Health Perspect 117(1):17-25 doi:10.1289/ehp.11447

Mannisto PT, Mantyla R, Mattila J, Nykanen S, Lamminsivu U (1982) Comparison of pharmacokinetics of sulphadiazine and sulphamethoxazole after intravenous infusion. J Antimicrob Chemother 9(6):461-70

Marshall GM, Carter DR, Cheung BB, Liu T, Mateos MK, Meyerowitz JG, et al. (2014) The prenatal origins of cancer. Nat Rev Cancer 14(4):277-89 doi:10.1038/nrc3679

Mattison DR (2010) Environmental exposures and development. Curr Opin Pediatr 22(2):208-18 doi:10.1097/MOP.0b013e32833779bf

Mayor R, Theveneau E (2013) The neural crest. Development 140(11):2247-51 doi:10.1242/dev.091751

Meganathan K, Jagtap S, Wagh V, Winkler J, Gaspar JA, Hildebrand D, et al. (2012) Identification of thalidomide-specific transcriptomics and proteomics signatures during differentiation of human embryonic stem cells. PLoS One 7(8):e44228 doi:10.1371/journal.pone.0044228

Menegola E, Broccia ML, Di Renzo F, Massa V, Giavini E (2005) Craniofacial and axial skeletal defects induced by the fungicide triadimefon in the mouse. Birth Defects Res B Dev Reprod Toxicol 74(2):185-95 doi:10.1002/bdrb.20035

Menegola E, Broccia ML, Di Renzo F, Prati M, Giavini E (2000) In vitro teratogenic potential of two antifungal triazoles: triadimefon and triadimenol. In Vitro Cell Dev Biol Anim 36(2):88-95 doi:10.1290/1071-2690(2000)036<0088:IVTPOT>2.0.CO;2

Mol PG, Arnardottir AH, Motola D, Vrijlandt PJ, Duijnhoven RG, Haaijer-Ruskamp FM, et al. (2013) Post-approval safety issues with innovative drugs: a European cohort study. Drug Saf 36(11):1105-15 doi:10.1007/s40264-013-0094-y

Moondra V, Sarma S, Buxton T, Safa R, Cote G, Storer T, et al. (2009) Serum Neuregulin-1beta as a Biomarker of Cardiovascular Fitness. Open Biomark J 2:1-5 doi:10.2174/1875318300902010001

Moors M, Rockel TD, Abel J, Cline JE, Gassmann K, Schreiber T, et al. (2009) Human neurospheres as three-dimensional cellular systems for developmental neurotoxicity testing. Environ Health Perspect 117(7):1131-8

Moraga I, Harari D, Schreiber G, Uze G, Pellegrini S (2009) Receptor density is key to the alpha2/beta interferon differential activities. Mol Cell Biol 29(17):4778-87 doi:10.1128/MCB.01808-08

Motohashi T, Kunisada T (2015) Extended multipotency of neural crest cells and neural crest-derived cells. Curr Top Dev Biol 111:69-95 doi:10.1016/bs.ctdb.2014.11.003

Mross K, Stefanic M, Gmehling D, Frost A, Baas F, Unger C, et al. (2010) Phase I study of the angiogenesis inhibitor BIBF 1120 in patients with advanced solid tumors. Clin Cancer Res 16(1):311-9 doi:10.1158/1078-0432.CCR-09-0694

Mueck W, Lensing AW, Agnelli G, Decousus H, Prandoni P, Misselwitz F (2011) Rivaroxaban: population pharmacokinetic analyses in patients treated for acute deep-vein thrombosis and exposure simulations in patients with atrial fibrillation treated for stroke prevention. Clin Pharmacokinet 50(10):675-86 doi:10.2165/11595320-000000000-00000

Munoz WA, Trainor PA (2015) Neural crest cell evolution: how and when did a neural crest cell become a neural crest cell. Curr Top Dev Biol 111:3-26 doi:10.1016/bs.ctdb.2014.11.001

Nakamura T (1995) Genetic markers and animal models of neurocristopathy. Histol Histopathol 10(3):747-59

NRC (2000) Scientific Frontiers in Developmental Toxicology and Risk Assessment. Washington (DC)

NRC (2007) National Research Council, Committee on Toxicity Testing and Assessment of Environmental Agents. Toxicity Testing in the 21st Century: A Vision and a Strategy. The National Academies Press

Nyffeler J, Karreman C, Leisner H, Kim YJ, Lee G, Waldmann T, et al. (2016) Design of a high-throughput human neural crest cell migration assay to indicate potential developmental toxicants. ALTEX doi:10.14573/altex.1605031

Oberemm A, Onyon L, Gundert-Remy U (2005) How can toxicogenomics inform risk assessment? Toxicol Appl Pharmacol 207(2 Suppl):592-8 doi:10.1016/j.taap.2005.01.044

Opacka-Juffry J, Mohiyeddini C (2012) Experience of stress in childhood negatively correlates with plasma oxytocin concentration in adult men. Stress 15(1):1-10 doi:10.3109/10253890.2011.560309

Padilla S, Corum D, Padnos B, Hunter DL, Beam A, Houck KA, et al. (2012) Zebrafish developmental screening of the ToxCast Phase I chemical library. Reprod Toxicol 33(2):174-87 doi:10.1016/j.reprotox.2011.10.018

Pallocca G, Fabbri M, Sacco MG, Gribaldo L, Pamies D, Laurenza I, et al. (2013) miRNA expression profiling in a human stem cell-based model as a tool for developmental neurotoxicity testing. Cell Biol Toxicol 29(4):239-57 doi:10.1007/s10565-013-9250-5

Pallocca G, Grinberg M, Henry M, Frickey T, Hengstler JG, Waldmann T, et al. (2016) Identification of transcriptome signatures and biomarkers specific for potential developmental toxicants inhibiting human neural crest cell migration. Arch Toxicol 90(1):159-80 doi:10.1007/s00204-015-1658-7

Pamies D, Bal-Price A, Simeonov A, Tagle D, Allen D, Gerhold D, et al. (2016) Good Cell Culture Practice for stem cells and stem-cell-derived models. ALTEX doi:10.14573/altex.1607121

Parekh JM, Vaghela RN, Sutariya DK, Sanyal M, Yadav M, Shrivastav PS (2010) Chromatographic separation and sensitive determination of teriflunomide, an active metabolite of leflunomide in human plasma by liquid chromatography-tandem mass spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci 878(24):2217-25 doi:10.1016/j.jchromb.2010.06.028

Patel S (2013) Long-term efficacy of imatinib for treatment of metastatic GIST. Cancer Chemother Pharmacol 72(2):277-86 doi:10.1007/s00280-013-2135-8

Patlewicz G, Ball N, Becker RA, Booth ED, Cronin MT, Kroese D, et al. (2014) Read-across approaches--misconceptions, promises and challenges ahead. ALTEX 31(4):387-96 doi:http://dx.doi.org/10.14573/altex.1410071

Pei Y, Peng J, Behl M, Sipes NS, Shockley KR, Rao MS, et al. (2015) Comparative neurotoxicity screening in human iPSC-derived neural stem cells, neurons and astrocytes. Brain Res doi:10.1016/j.brainres.2015.07.048

Piedrafita B, Erceg S, Cauli O, Monfort P, Felipo V (2008) Developmental exposure to polychlorinated biphenyls PCB153 or PCB126 impairs learning ability in young but not in adult rats. Eur J Neurosci 27(1):177-82 doi:10.1111/j.1460-9568.2007.5988.x

Piehler J, Thomas C, Garcia KC, Schreiber G (2012) Structural and dynamic determinants of type I interferon receptor assembly and their functional interpretation. Immunol Rev 250(1):317-34 doi:10.1111/imr.12001

Pierrehumbert B, Torrisi R, Laufer D, Halfon O, Ansermet F, Beck Popovic M (2010) Oxytocin response to an experimental psychosocial challenge in adults exposed to traumatic experiences during childhood or adolescence. Neuroscience 166(1):168-77 doi:10.1016/j.neuroscience.2009.12.016

Piersma AH, Bosgra S, van Duursen MB, Hermsen SA, Jonker LR, Kroese ED, et al. (2013) Evaluation of an alternative in vitro test battery for detecting reproductive toxicants. Reprod Toxicol 38:53-64 doi:10.1016/j.reprotox.2013.03.002

Piersma AH, Janer G, Wolterink G, Bessems JG, Hakkert BC, Slob W (2008) Quantitative extrapolation of in vitro whole embryo culture embryotoxicity data to developmental toxicity in vivo using the benchmark dose approach. Toxicol Sci 101(1):91-100 doi:10.1093/toxsci/kfm253

Pinner NA, Hamilton LA, Hughes A (2012) Roflumilast: a phosphodiesterase-4 inhibitor for the treatment of severe chronic obstructive pulmonary disease. Clin Ther 34(1):56-66 doi:10.1016/j.clinthera.2011.12.008

Pla P, Larue L (2003) Involvement of endothelin receptors in normal and pathological development of neural crest cells. Int J Dev Biol 47(5):315-25

Plum A, Jensen LB, Kristensen JB (2013) In vitro protein binding of liraglutide in human plasma determined by reiterated stepwise equilibrium dialysis. J Pharm Sci 102(8):2882-8 doi:10.1002/jps.23648

Poltl D, Schildknecht S, Karreman C, Leist M (2012) Uncoupling of ATP-depletion and cell death in human dopaminergic neurons. Neurotoxicology 33(4):769-79 doi:10.1016/j.neuro.2011.12.007

Pozzilli C, Pugliatti M (2015) An overview of pregnancy-related issues in patients with multiple sclerosis. Eur J Neurol 22 Suppl 2:34-9 doi:10.1111/ene.12797

Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, et al. (2010) Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res 12(5):R68 doi:10.1186/bcr2635

Preynat-Seauve O, Suter DM, Tirefort D, Turchi L, Virolle T, Chneiweiss H, et al. (2009) Development of human nervous tissue upon differentiation of embryonic stem cells in three-dimensional culture. Stem Cells 27(3):509-20 doi:10.1634/stemcells.2008-0600

Purvis K, Muirhead GJ, Harness JA (2002) The effects of sildenafil on human sperm function in healthy volunteers. Br J Clin Pharmacol 53 Suppl 1:53S-60S

Rahnenfuhrer J, Leist M (2015) From smoking guns to footprints: mining for critical events of toxicity pathways in transcriptome data. Arch Toxicol 89(5):813-7 doi:10.1007/s00204-015-1497-6

Ramirez T, Daneshian M, Kamp H, Bois FY, Clench MR, Coen M, et al. (2013) Metabolomics in toxicology and preclinical research. ALTEX 30(2):209-25

Reif DM, Sypa M, Lock EF, Wright FA, Wilson A, Cathey T, et al. (2013) ToxPi GUI: an interactive visualization tool for transparent integration of data from diverse sources of evidence. Bioinformatics 29(3):402-3 doi:10.1093/bioinformatics/bts686

Rempel E, Hoelting L, Waldmann T, Balmer NV, Schildknecht S, Grinberg M, et al. (2015) A transcriptome-based classifier to identify developmental toxicants by stem celltesting: design, validation, and optimization for histone deacetylase inhibitors. Arch Toxicol 89(9):1599-618 doi:10.1007/s00204-015-1573-y

Robinson JF, Pennings JL, Piersma AH (2012a) A review of toxicogenomic approaches in developmental toxicology. Methods Mol Biol 889:347-71 doi:10.1007/978-1-61779-867-2_22

Robinson JF, Piersma AH (2013) Toxicogenomic approaches in developmental toxicology testing. Methods Mol Biol 947:451-73 doi:10.1007/978-1-62703-131-8_31

Robinson JF, Tonk EC, Verhoef A, Piersma AH (2012b) Triazole induced concentration-related gene signatures in rat whole embryo culture. Reprod Toxicol 34(2):275-83 doi:10.1016/j.reprotox.2012.05.088

Robinson JF, van Beelen VA, Verhoef A, Renkens MF, Luijten M, van Herwijnen MH, et al. (2010) Embryotoxicant-specific transcriptomic responses in rat postimplantation whole-embryo culture. Toxicol Sci 118(2):675-85 doi:10.1093/toxsci/kfq292

Robinson JF, Yu X, Moreira EG, Hong S, Faustman EM (2011) Arsenic- and cadmium-induced toxicogenomic response in mouse embryos undergoing neurulation. Toxicol Appl Pharmacol 250(2):117-29 doi:10.1016/j.taap.2010.09.018

Rojas D, Rager JE, Smeester L, Bailey KA, Drobna Z, Rubio-Andrade M, et al. (2015) Prenatal arsenic exposure and the epigenome: identifying sites of 5-methylcytosine alterations that predict functional changes in gene expression in newborn cord blood and subsequent birth outcomes. Toxicol Sci 143(1):97-106 doi:10.1093/toxsci/kfu210

Rossi UA, Finocchiaro LM, Glikin GC (2015) Interferon-beta gene transfer inhibits melanoma cells adhesion and migration. Cytokine doi:10.1016/j.cyto.2015.11.012

Rovida C, Alepee N, Api AM, Basketter DA, Bois FY, Caloni F, et al. (2015a) Integrated Testing Strategies (ITS) for safety assessment. ALTEX 32(1):25-40 doi:http://dx.doi.org/10.14573/altex.1411011

Rovida C, Asakura S, Daneshian M, Hofman-Huether H, Leist M, Meunier L, et al. (2015b) Toxicity testing in the 21st century beyond environmental chemicals. ALTEX 32(3):171-81 doi:http://dx.doi.org/10.14573/altex.1506201

Rovida C, Longo F, Rabbit R (2011) How are reproductive toxicity and developmental toxicity addressed in REACH dossiers? Altex 28(4):273-94

Rovida C, Vivier M, Garthoff B, Hescheler J (2014) ESNATS conference - the use of human embryonic stem cells for novel toxicity testing approaches. Altern Lab Anim 42(2):97-113

Rutter M (2005) Incidence of autism spectrum disorders: changes over time and their meaning. Acta Paediatr 94(1):2-15

Saint-Jeannet J-P (2006) Neural crest induction and differentiation. Springer Science+Business Media ;

Landes Bioscience/Eurekah.com, New York, N.Y.

Georgetown, Tex.

Sandberg-Wollheim M, Frank D, Goodwin TM, Giesser B, Lopez-Bresnahan M, Stam-Moraga M, et al. (2005) Pregnancy outcomes during treatment with interferon beta-1a in patients with multiple sclerosis. Neurology 27(65(6)):802-6

Sauka-Spengler T, Bronner-Fraser M (2008) A gene regulatory network orchestrates neural crest formation. Nat Rev Mol Cell Biol 9(7):557-68 doi:10.1038/nrm2428

Scantamburlo G, Hansenne M, Fuchs S, Pitchot W, Marechal P, Pequeux C, et al. (2007) Plasma oxytocin levels and anxiety in patients with major depression. Psychoneuroendocrinology 32(4):407-10 doi:10.1016/j.psyneuen.2007.01.009

Scheffler M, Di Gion P, Doroshyenko O, Wolf J, Fuhr U (2011) Clinical pharmacokinetics of tyrosine kinase inhibitors: focus on 4-anilinoquinazolines. Clin Pharmacokinet 50(6):371-403 doi:10.2165/11587020-000000000-00000

Schenk B, Weimer M, Bremer S, van der Burg B, Cortvrindt R, Freyberger A, et al. (2010) The ReProTect Feasibility Study, a novel comprehensive in vitro approach to detect

reproductive toxicants. Reprod Toxicol 30(1):200-18 doi:10.1016/j.reprotox.2010.05.012

Schildknecht S, Karreman C, Poltl D, Efremova L, Kullmann C, Gutbier S, et al. (2013) Generation of genetically-modified human differentiated cells for toxicological tests and the study of neurodegenerative diseases. ALTEX 30(4):427-44

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. (2012) Fiji: an open-source platform for biological-image analysis. Nat Methods 9(7):676-82 doi:10.1038/nmeth.2019

Schmidt BZ, Lehmann M, Gutbier S, Nembo E, Noel S, Smirnova L, et al. (2016) In vitro acute and developmental neurotoxicity screening: an overview of cellular platforms and high-throughput technical possibilities. Arch Toxicol doi:10.1007/s00204-016-1805-9

Schreiber G, Piehler J (2015) The molecular basis for functional plasticity in type I interferon signaling. Trends Immunol 36(3):139-49 doi:10.1016/j.it.2015.01.002

Schreiber T, Gassmann K, Gotz C, Hubenthal U, Moors M, Krause G, et al. (2010) Polybrominated diphenyl ethers induce developmental neurotoxicity in a human in vitro model: evidence for endocrine disruption. Environ Health Perspect 118(4):572-8 doi:10.1289/ehp.0901435

Schroder K, Hertzog PJ, Ravasi T, Hume DA (2004) Interferon-gamma: an overview of signals, mechanisms and functions. J Leukoc Biol 75(2):163-89 doi:10.1189/jlb.0603252

Scott WJ, Fradkin R, Wittfoht W, Nau H (1989) Teratologic potential of 2-methoxyethanol and transplacental distribution of its metabolite, 2-methoxyacetic acid, in non-human primates. Teratology 39(4):363-73 doi:10.1002/tera.1420390408

Seiler AE, Spielmann H (2011) The validated embryonic stem cell test to predict embryotoxicity in vitro. Nat Protoc 6(7):961-78 doi:10.1038/nprot.2011.348

Shao J, Berger LF, Hendriksen PJ, Peijnenburg AA, van Loveren H, Volger OL (2014) Transcriptome-based functional classifiers for direct immunotoxicity. Arch Toxicol 88(3):673-89 doi:10.1007/s00204-013-1179-1

Shiga T, Tanaka T, Irie S, Hagiwara N, Kasanuki H (2011) Pharmacokinetics of intravenous amiodarone and its electrocardiographic effects on healthy Japanese subjects. Heart Vessels 26(3):274-81 doi:10.1007/s00380-010-0047-7

Simoes-Costa M, Bronner ME (2015) Establishing neural crest identity: a gene regulatory recipe. Development 142(2):242-57 doi:10.1242/dev.105445

Simões-Costa M, Bronner ME (2013) Insights into neural crest development and evolution from genomic analysis. Genome Res 23(7):1069-80 doi:10.1101/gr.157586.113

Sipes NS, Martin MT, Kothiya P, Reif DM, Judson RS, Richard AM, et al. (2013) Profiling 976 ToxCast chemicals across 331 enzymatic and receptor signaling assays. Chem Res Toxicol 26(6):878-95 doi:10.1021/tx400021f

Sipes NS, Martin MT, Reif DM, Kleinstreuer NC, Judson RS, Singh AV, et al. (2011) Predictive models of prenatal developmental toxicity from ToxCast high-throughput screening data. Toxicol Sci 124(1):109-27 doi:10.1093/toxsci/kfr220

Smirnova L, Hogberg HT, Leist M, Hartung T (2014) Developmental neurotoxicity - challenges in the 21st century and in vitro opportunities. ALTEX 31(2):129-56 doi:http://dx.doi.org/10.14573/altex.1403271

Smyth GK, Michaud J, Scott HS (2005) Use of within-array replicate spots for assessing differential expression in microarray experiments. Bioinformatics 21(9):2067-75 doi:10.1093/bioinformatics/bti270

Song R, Duarte TL, Almeida GM, Farmer PB, Cooke MS, Zhang W, et al. (2009) Cytotoxicity and gene expression profiling of two hydroxylated polybrominated diphenyl ethers in human H295R adrenocortical carcinoma cells. Toxicol Lett 185(1):23-31 doi:10.1016/j.toxlet.2008.11.011

Staun-Ram E, Miller A (2011) Cathepsins (S and B) and their inhibitor Cystatin C in immune cells: modulation by interferon-beta and role played in cell migration. J Neuroimmunol 232(1-2):200-6 doi:10.1016/j.jneuroim.2010.10.015

Stiegler NV, Krug AK, Matt F, Leist M (2011) Assessment of chemical-induced impairment of human neurite outgrowth by multiparametric live cell imaging in high-density cultures. Toxicol Sci 121(1):73-87 doi:10.1093/toxsci/kfr034

Stummann TC, Hareng L, Bremer S (2009) Hazard assessment of methylmercury toxicity to neuronal induction in embryogenesis using human embryonic stem cells. Toxicology 257(3):117-26 doi:10.1016/j.tox.2008.12.018

Stuve O, Chabot S, Jung SS, Williams G, Yong VW (1997) Chemokine-enhanced migration of human peripheral blood mononuclear cells is antagonized by interferon beta-1b through an effect on matrix metalloproteinase-9. J Neuroimmunol 80(1-2):38-46

Stuve O, Dooley NP, Uhm JH, Antel JP, Francis GS, Williams G, et al. (1996) Interferon beta-1b decreases the migration of T lymphocytes in vitro: effects on matrix metalloproteinase-9. Ann Neurol 40(6):853-63 doi:10.1002/ana.410400607

Sullivan (2005) Autism increase not a result of reclassification. Clin Psychiat News:68

Suzuki N, Ando S, Sumida K, Horie N, Saito K (2011) Analysis of altered gene expression specific to embryotoxic chemical treatment during embryonic stem cell differentiation into myocardiac and neural cells. J Toxicol Sci 36(5):569-85

Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, et al. (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131(5):861-72 doi:10.1016/j.cell.2007.11.019

Talens-Visconti R, Sanchez-Vera I, Kostic J, Perez-Arago MA, Erceg S, Stojkovic M, et al. (2011) Neural differentiation from human embryonic stem cells as a tool to study early brain development and the neuroteratogenic effects of ethanol. Stem Cells Dev 20(2):327-39 doi:10.1089/scd.2010.0037

Tapaninen T, Backman JT, Kurkinen KJ, Neuvonen PJ, Niemi M (2011) Itraconazole, a P-glycoprotein and CYP3A4 inhibitor, markedly raises the plasma concentrations and enhances the renin-inhibiting effect of aliskiren. J Clin Pharmacol 51(3):359-67 doi:10.1177/0091270010365885

Tarassishin L, Lee SC (2013) Interferon regulatory factor 3 alters glioma inflammatory and invasive properties. J Neurooncol 113(2):185-94 doi:10.1007/s11060-013-1109-3

Theunissen PT, Pennings JL, Robinson JF, Claessen SM, Kleinjans JC, Piersma AH (2011) Time-response evaluation by transcriptomics of methylmercury effects on neural differentiation of murine embryonic stem cells. Toxicol Sci 122(2):437-47 doi:10.1093/toxsci/kfr134

Theunissen PT, Robinson JF, Pennings JL, de Jong E, Claessen SM, Kleinjans JC, et al. (2012) Transcriptomic concentration-response evaluation of valproic acid, cyproconazole, and hexaconazole in the neural embryonic stem cell test (ESTn). Toxicol Sci 125(2):430-8 doi:10.1093/toxsci/kfr293

Thomas RS, Allen BC, Nong A, Yang L, Bermudez E, Clewell HJ, et al. (2007) A method to integrate benchmark dose estimates with genomic data to assess the functional effects of chemical exposure. Toxicol Sci 98(1):240-8

Thomas RS, Wesselkamper SC, Wang NC, Zhao QJ, Petersen DD, Lambert JC, et al. (2013) Temporal concordance between apical and transcriptional points of departure for chemical risk assessment. Toxicol Sci 134(1):180-94 doi:10.1093/toxsci/kft094

Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, et al. (1998) Embryonic stem cell lines derived from human blastocysts. Science 282(5391):1145-7

Tice RR, Austin CP, Kavlock RJ, Bucher JR (2013) Improving the human hazard characterization of chemicals: a Tox21 update. Environ Health Perspect 121(7):756-65 doi:10.1289/ehp.1205784

Trainor PA (2010) Craniofacial birth defects: The role of neural crest cells in the etiology and pathogenesis of Treacher Collins syndrome and the potential for prevention. Am J Med Genet A 152A(12):2984-94 doi:10.1002/ajmg.a.33454

Tripathi V, Shen Z, Chakraborty A, Giri S, Freier SM, Wu X, et al. (2013) Long noncoding RNA MALAT1 controls cell cycle progression by regulating the expression of oncogenic transcription factor B-MYB. PLoS Genet 9(3):e1003368 doi:10.1371/journal.pgen.1003368

Vachiery JL, Huez S, Gillies H, Layton G, Hayashi N, Gao X, et al. (2011) Safety, tolerability and pharmacokinetics of an intravenous bolus of sildenafil in patients with pulmonary arterial hypertension. Br J Clin Pharmacol 71(2):289-92 doi:10.1111/j.1365-2125.2010.03831.x

Vahidnia A, Romijn F, van der Voet GB, de Wolff FA (2008) Arsenic-induced neurotoxicity in relation to toxicokinetics: effects on sciatic nerve proteins. Chem Biol Interact 176(2-3):188-95 doi:10.1016/j.cbi.2008.07.001

Vahidnia A, van der Straaten RJ, Romijn F, van Pelt J, van der Voet GB, de Wolff FA (2007) Arsenic metabolites affect expression of the neurofilament and tau genes: an in-vitro study into the mechanism of arsenic neurotoxicity. Toxicol In Vitro 21(6):1104-12 doi:10.1016/j.tiv.2007.04.007

Vaidyanathan S, Jarugula V, Dieterich HA, Howard D, Dole WP (2008) Clinical pharmacokinetics and pharmacodynamics of aliskiren. Clin Pharmacokinet 47(8):515-31

van Boxel-Dezaire AH, Rani MR, Stark GR (2006) Complex modulation of cell type-specific signaling in response to type I interferons. Immunity 25(3):361-72 doi:10.1016/j.immuni.2006.08.014

van Dartel DA, Pennings JL, Hendriksen PJ, van Schooten FJ, Piersma AH (2009) Early gene expression changes during embryonic stem cell differentiation into cardiomyocytes and their modulation by monobutyl phthalate. Reprod Toxicol 27(2):93-102 doi:10.1016/j.reprotox.2008.12.009

van Dartel DA, Piersma AH (2011) The embryonic stem cell test combined with toxicogenomics as an alternative testing model for the assessment of developmental toxicity. Reprod Toxicol 32(2):235-44 doi:10.1016/j.reprotox.2011.04.008

van de Leemput J, Boles NC, Kiehl TR, Corneo B, Lederman P, Menon V, et al. (2014) CORTECON: a temporal transcriptome analysis of in vitro human cerebral cortex development from human embryonic stem cells. Neuron 83(1):51-68 doi:10.1016/j.neuron.2014.05.013

van der Burg B, Pieterse B, Buist H, Lewin G, van der Linden SC, Man HY, et al. (2014) A high throughput screening system for predicting chemically-induced reproductive organ deformities. Reprod Toxicol doi:10.1016/j.reprotox.2014.11.011

van der Burg B, Wedebye EB, Dietrich DR, Jaworska J, Mangelsdorf I, Paune E, et al. (2015) The ChemScreen project to design a pragmatic alternative approach to predict reproductive toxicity of chemicals. Reprod Toxicol doi:10.1016/j.reprotox.2015.01.008

van Thriel C, Westerink RH, Beste C, Bale AS, Lein PJ, Leist M (2012) Translating neurobehavioural endpoints of developmental neurotoxicity tests into in vitro assays and readouts. Neurotoxicology 33(4):911-24 doi:10.1016/j.neuro.2011.10.002

van Vliet E, Eskes C, Stingele S, Gartlon J, Price A, Farina M, et al. (2007) Development of a mechanistically-based genetically engineered PC12 cell system to detect p53-mediated cytotoxicity. Toxicol In Vitro 21(4):698-705 doi:10.1016/j.tiv.2006.12.004

Veronese ME, McLean S, Hendriks R (1988) Plasma protein binding of amiodarone in a patient population: measurement by erythrocyte partitioning and a novel glass-binding method. Br J Clin Pharmacol 26(6):721-31

Viberg H, Fredriksson A, Eriksson P (2005) Deranged spontaneous behaviour and decrease in cholinergic muscarinic receptors in hippocampus in the adult rat, after neonatal

exposure to the brominated flame-retardant, 2,2',4,4',5-pentabromodiphenyl ether (PBDE 99). Environ Toxicol Pharmacol 20(2):283-8 doi:10.1016/j.etap.2005.02.004

Vojnits K, Ensenat-Waser R, Gaspar JA, Meganathan K, Jagtap S, Hescheler J, et al. (2012) A tanscriptomics study to elucidate the toxicological mechanism of methylmercury chloride in a human stem cell based in vitro test. Curr Med Chem 19(36):6224-32

Volbracht C, Leist M, Nicotera P (1999) ATP controls neuronal apoptosis triggered by microtubule breakdown or potassium deprivation. Mol Med 5(7):477-89

Waldmann T, Rempel E, Balmer NV, Konig A, Kolde R, Gaspar JA, et al. (2014) Design principles of concentration-dependent transcriptome deviations in drug-exposed differentiating stem cells. Chem Res Toxicol 27(3):408-20 doi:10.1021/tx400402j

Wang L, Zhang ZG, Gregg SR, Zhang RL, Jiao Z, LeTourneau Y, et al. (2007) The Sonic hedgehog pathway mediates carbamylated erythropoietin-enhanced proliferation and differentiation of adult neural progenitor cells. J Biol Chem 282(44):32462-70 doi:10.1074/jbc.M706880200

Wang S, Bates J, Li X, Schanz S, Chandler-Militello D, Levine C, et al. (2013) Human iPSC-derived oligodendrocyte progenitor cells can myelinate and rescue a mouse model of congenital hypomyelination. Cell Stem Cell 12(2):252-64 doi:10.1016/j.stem.2012.12.002

Waring JF, Halbert DN (2002) The promise of toxicogenomics. Curr Opin Mol Ther 4(3):229-35

Waters MD, Fostel JM (2004) Toxicogenomics and systems toxicology: aims and prospects. Nat Rev Genet 5(12):936-48 doi:10.1038/nrg1493

Wayman GA, Bose DD, Yang D, Lesiak A, Bruun D, Impey S, et al. (2012a) PCB-95 modulates the calcium-dependent signaling pathway responsible for activity-dependent dendritic growth. Environ Health Perspect 120(7):1003-9 doi:10.1289/ehp.1104833

Wayman GA, Yang D, Bose DD, Lesiak A, Ledoux V, Bruun D, et al. (2012b) PCB-95 promotes dendritic growth via ryanodine receptor-dependent mechanisms. Environ Health Perspect 120(7):997-1002 doi:10.1289/ehp.1104832

Welsch F (2005) The mechanism of ethylene glycol ether reproductive and developmental toxicity and evidence for adverse effects in humans. Toxicol Lett 156(1):13-28 doi:10.1016/j.toxlet.2003.08.010

Wilson VS, Keshava N, Hester S, Segal D, Chiu W, Thompson CM, et al. (2013) Utilizing toxicogenomic data to understand chemical mechanism of action in risk assessment. Toxicol Appl Pharmacol 271(3):299-308 doi:10.1016/j.taap.2011.01.017

Wobus AM, Loser P (2011) Present state and future perspectives of using pluripotent stem cells in toxicology research. Arch Toxicol 85(2):79-117 doi:10.1007/s00204-010-0641-6

Xia M, Huang R, Witt KL, Southall N, Fostel J, Cho MH, et al. (2008) Compound cytotoxicity profiling using quantitative high-throughput screening. Environ Health Perspect 116(3):284-91 doi:10.1289/ehp.10727

Xuereb F, Chaignepain S, Breilh D, Godde F, Saux MC, Lenz C, et al. (2011) Quantitative analysis of erythropoietin in human plasma by tandem mass spectrometry. Anal Bioanal Chem 400(7):2073-84 doi:10.1007/s00216-011-4897-3

Yamada I, Suzuki F, Kamiya N, Aoki K, Sakurai Y, Kano M, et al. (2012) Safety, pharmacokinetics and resistant variants of telaprevir alone for 12 weeks in hepatitis C virus genotype 1b infection. J Viral Hepat 19(2):e112-9 doi:10.1111/j.1365-2893.2011.01514.x

Yang F, Yi F, Han X, Du Q, Liang Z (2013) MALAT-1 interacts with hnRNP C in cell cycle regulation. FEBS Lett 587(19):3175-81 doi:10.1016/j.febslet.2013.07.048

Yen JH, Kong W, Ganea D (2010) IFN-beta inhibits dendritic cell migration through STAT-1-mediated transcriptional suppression of CCR7 and matrix metalloproteinase 9. J Immunol 184(7):3478-86 doi:10.4049/jimmunol.0902542

Yeung PK, Hubbard JW, Korchinski ED, Midha KK (1993) Pharmacokinetics of chlorpromazine and key metabolites. Eur J Clin Pharmacol 45(6):563-9

Yu F, Ng SS, Chow BK, Sze J, Lu G, Poon WS, et al. (2011) Knockdown of interferon-induced transmembrane protein 1 (IFITM1) inhibits proliferation, migration, and invasion of glioma cells. J Neurooncol 103(2):187-95 doi:10.1007/s11060-010-0377-4

Yu X, Griffith WC, Hanspers K, Dillman JF, Ong H, Vredevoogd MA, et al. (2006) A system-based approach to interpret dose- and time-dependent microarray data: quantitative integration of gene ontology analysis for risk assessment. Toxicol Sci 92(2):560-77

Yung WK, Prados M, Levin VA, Fetell MR, Bennett J, Mahaley MS, et al. (1991) Intravenous recombinant interferon beta in patients with recurrent malignant gliomas: a phase I/II study. J Clin Oncol 9(11):1945-9

Zagzag D, Nomura M, Friedlander DR, Blanco CY, Gagner JP, Nomura N, et al. (2003) Geldanamycin inhibits migration of glioma cells in vitro: a potential role for hypoxia-inducible factor (HIF-1alpha) in glioma cell invasion. J Cell Physiol 196(2):394-402 doi:10.1002/jcp.10306

Zeng H, Shen EH, Hohmann JG, Oh SW, Bernard A, Royall JJ, et al. (2012) Large-scale cellular-resolution gene profiling in human neocortex reveals species-specific molecular signatures. Cell 149(2):483-96 doi:10.1016/j.cell.2012.02.052

Zhang X, Cornelis R, de Kimpe J, Mees L, Lameire N (1998) Study of arsenic-protein binding in serum of patients on continuous ambulatory peritoneal dialysis. Clin Chem 44(1):141-7

Zhao S, Fung-Leung WP, Bittner A, Ngo K, Liu X (2014) Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. PLoS One 9(1):e78644 doi:10.1371/journal.pone.0078644

Zimmer B, Kuegler PB, Baudis B, Genewsky A, Tanavde V, Koh W, et al. (2011a) Coordinated waves of gene expression during neuronal differentiation of embryonic stem cells as basis for novel approaches to developmental neurotoxicity testing. Cell Death Differ 18(3):383-95 doi:10.1038/cdd.2010.109

Zimmer B, Lee G, Balmer NV, Meganathan K, Sachinidis A, Studer L, et al. (2012) Evaluation of developmental toxicants and signaling pathways in a functional test based on the migration of human neural crest cells. Environ Health Perspect 120(8):1116-22 doi:10.1289/ehp.1104489

Zimmer B, Pallocca G, Dreser N, Foerster S, Waldmann T, Westerhout J, et al. (2014) Profiling of drugs and environmental chemicals for functional impairment of neural crest migration in a novel stem cell-based test battery. Arch Toxicol 88(5):1109-26 doi:10.1007/s00204-014-1231-9

Zimmer B, Schildknecht S, Kuegler PB, Tanavde V, Kadereit S, Leist M (2011b) Sensitivity of dopaminergic neuron differentiation from stem cells to chronic low-dose methylmercury exposure. Toxicol Sci 121(2):357-67 doi:10.1093/toxsci/kfr054

Zula JA, Green HC, Ransohoff RM, Rudick RA, Stark GR, van Boxel-Dezaire AH (2011) The role of cell type-specific responses in IFN-beta therapy of multiple sclerosis. Proc Natl Acad Sci U S A 108(49):19689-94 doi:10.1073/pnas.1117347108

Zurich MG, Eskes C, Honegger P, Berode M, Monnet-Tschudi F (2002) Maturation-dependent neurotoxicity of lead acetate in vitro: implication of glial reactions. J Neurosci Res 70(1):108-16 doi:10.1002/jnr.10367

Zurich MG, Honegger P, Schilter B, Costa LG, Monnet-Tschudi F (2000) Use of aggregating brain cell cultures to study developmental effects of organophosphorus insecticides. Neurotoxicology 21(4):599-605

List of publications

Publications, integrated in this thesis:

- Profiling of drugs and environmental chemicals for functional impairment of neural crest migration in a novel stem cell-based test battery. Zimmer B, Pallocca G, Dreser N, Foerster S, Waldmann T, Westerhout J, Julien S, Krause KH, van Thriel C, Hengstler JG, Sachinidis A, Bosgra S, Leist M. Arch Toxicol. 2014 May;88(5):1109-26. doi: 10.1007/s00204-014-1231-9.
- Identification of transcriptome signatures and biomarkers specific for potential developmental toxicants inhibiting human neural crest cell migration. Pallocca G, Grinberg M, Henry M, Frickey T, Hengstler JG,Waldmann T, Sachinidis A, Rahnenführer J, Leist M. Arch Toxicol. 2016 Jan;90(1):159-80. doi: 10.1007/s00204-015-1658-7.
- Impairment of human neural crest cell migration by prolonged exposure to interferon-beta. Pallocca G, Nyfeller J, Dolde X, Grinberg M, Waldmann T, Rahnenführer J, Sachinidis A and Leist M. Submitted to *Biochemical Pharmacology* journal.

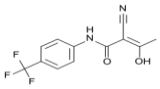
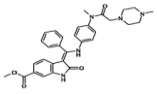
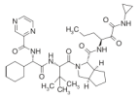
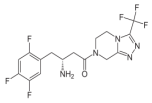
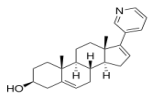
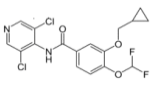
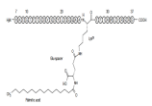
Publications, not integrated in this thesis:

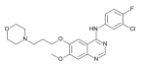
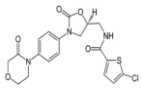
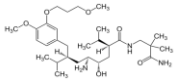
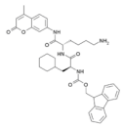
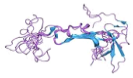
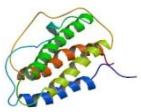
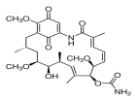
- Grouping of histone deacetylase inhibitors and other toxicants disturbing neural crest migration by transcriptional profiling. Dreser N, Zimmer B, Dietz C, Sügis E, Pallocca G, Nyffeler J, Meisig J, Blüthgen N, Berthold MR, Waldmann T, Leist M. Neurotoxicology. 2015 Sep;50:56-70. doi: 10.1016/j.neuro.2015.07.008.
- Fingerprinting of neurotoxic compounds using a mouse embryonic stem cell dual luminescence reporter assay. Colaianna M, Ilmjärv S, Peterson H, Kern I,Julien S, Baquié M, Pallocca G, Bosgra S, Sachinidis A, Hengstler JG, Leist M, Krause KH. Arch Toxicol. 2016 Mar 25

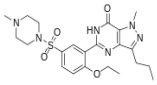
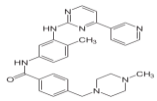
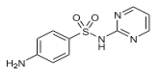
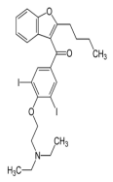
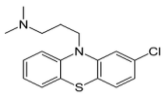
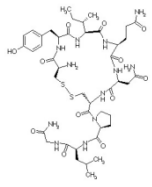
Supplemental material manuscript 1

Supplemental Material, Figure S1:

Chemical and pharmacological characteristics of the group of medical drugs

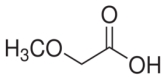
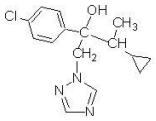
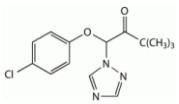
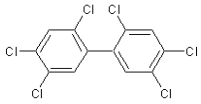
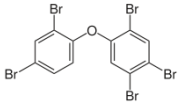
Compound	**Pharmacological characteristics**	**Chemical characteristics**	**Supplier/ Catalog No.**
Teriflunomide 	Immunomodulatory drug, inhibiting pyrimidine de novo synthesis by blocking the enzyme dihydroorotate dehydrogenase	Amide	Enzo Life sciences/ ALX-430-096-M005
Nintedanib  (BIBF1120;Vergatef)	Tyrosin kinase inhibitor developed for tumor therapy, inhibits signaling of three growth factor receptors involved in angiogenesis (VEGFR, PDGFR and FGFR)	Nucleotide mimetic	Selleckchem/ S1010
Telaprevir 	Antiviral drug for the treatment of hepatitis C; protease inhibitors	Peptide mimetic	Selleckchem/ S1538
Sitagliptin 	Oral anti-hyperglycemic (antidiabetic) of the dipeptidyl peptidase-4 (DPP-4) inhibitor class	Peptide mimetic	Selleckchem/ S4002
Abiraterone 	Antiandrogen; inhibits 17 α-hydroxylase/C17,20 lyase (CYP17A1), an enzyme which is involved in steroid (testosterone) synthesis	Steroid	Selleckchem/ S1123
Roflumilast 	Selective, long-acting inhibitor of PDE-4; isoform of phosphodiesterases	Benzamide	Selleckchem/ S2131
Exenatide 	Glucagon- like peptide-1 receptor agonist (GLP-1 mimetic); treatment of diabetes mellitus type 2	Peptide/Protein	Prospec/ HOR-246

Gefitinib  (Iressa)	Inhibitor of epidermal growth factor receptor's (EGFR) tyrosine kinase domain	Nucleotide mimetic	Selleckchem/ S1025
Rivaroxaban 	Oral anticoagulant; inhibition of the factor Xa protease	Peptide mimetic	Selleckchem/ S3002
Aliskiren 	Inhibitor of renin protease; treatment of essential (primary) hypertension (preventing of the conversion of angiotensinogen to angiotensin I)	Peptide mimetic	Selleckchem/ S2199
Galnon 	Selective agonist at the galanin receptors GALR. Anticonvul sant, anxiolytic, anorectic and amnestic effects in animal models	Peptide mimetic	Sigma-Aldrich/ G4419
Neuregulin 	Endogenous agonist of the erbB family of tyrosine kinase receptors; plays multiple essential roles in neuronal development and disease	Peptide/Protein	R&D Systems/ 378-SM-025
Erythropoietin 	Glycoprotein hormone that acts as agonist of EpoR and controls erythropoiesis and neurogenesis	Peptide/Protein	R&D Systems/ 287-TC-500
Geldanamycin 	Benzoquinone ansamycin antibiotic th at binds to Hsp90 (heat shock protein 90) and inhibits its function. Antitumor effects by acting on v-Src, mutant p53 proteins, Raf-1 and EGFR signalling	Amide	Selleckchem/ S2713
G-CSF 	Protein hormone that stimulates granulopoiesis and with neuroprotective effects	Peptide/Protein	R&D Systems/ 214-CS-005

IFNβ	Protein hormone (cytokine); multiple sclerosis treatment	Peptide/Protein	R&D Systems/ 11415-1
Sildenafil 	Inhibitor of the PDE5 isoform of phospodiesterases; used to treat erectile dysfunction and pulmonary arterial hypertension (PAH).	Nucleotide mimetic	Sigma-Aldrich/ PZ-0003
Imatinib 	Tyrosin-kinase inhibitor used in the treatment of multiple cancers; inihibits c-kit and PDGF-R (platelet-derived growth factor receptor) signaling	Peptide mimetic	Selleckchem/ S1026
Sulfadiazine 	Sulfonamide antibiotic; stops the production of folic acid in parasites	Benzene sulfonamide	Sigma-Aldrich/ S8626
Amiodarone 	Class III of antiarrhythmic agents; blocks sodium channels	Tertiary amine	Sigma-Aldrich/ A8123
Chlorpromazine 	Dopamine antagonist possessing additional antiadrenergic, antiserotonergic, antic holinergic and antihistaminergic prope rties used to treat schizophrenia	Tertiary amine	Sigma-Aldrich/ C8138
Oxytocin 	Peptide hormone; stimulates uterine contraction and lactation	Peptide/ Protein	R&D Systems/ 1910

Supplemental Material, Figure S2:

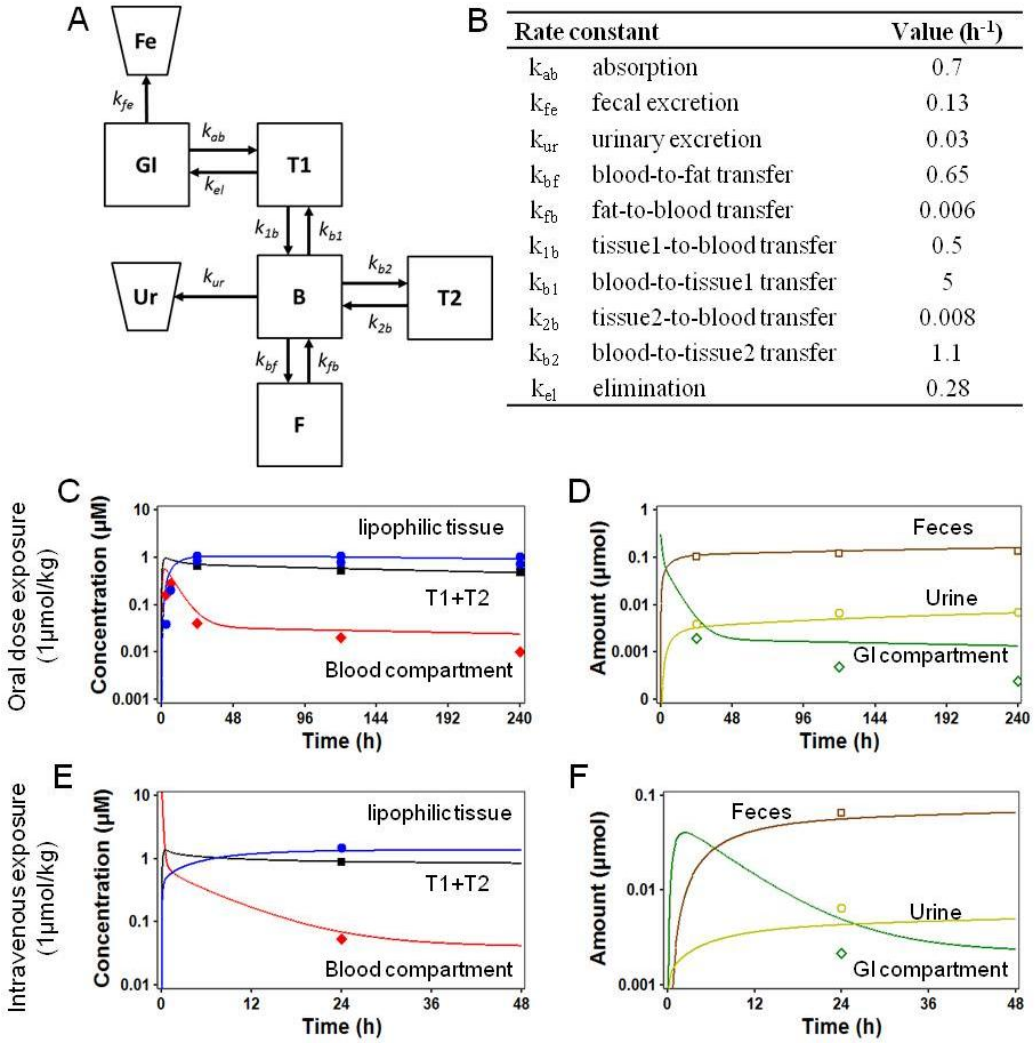
Chemical and pharmacological characteristics of the group of environmental pollutants

Compound	Pharmacological characteristics	Chemical characteristics	Developmental toxicity/Neurotoxicity evidences		Supplier info
			In vitro	In vivo	
Methoxyacetic acid 	Phthalate ester (DMEP) and methoxyethanol metabolite formed rapidly formed in vivo from industrial solvents	Carboxylic acid	(1) (2)	(3) (4) (5)	Sigma-Aldrich/ 194557
Cyproconazole 	Pesticide, showing various teratogenic effects, via inhibition of cyp enzymes and reduction of steroidogenesis	Triazole	(6) (7) (8)	(9)	Sigma-Aldrich/ 46068
Triadimefon 	Pesticide; various teratogenic effects observed in rats	Triazole	(10) (11)	(12) (13)	Bayer Crop Science
PCB-153 	Non-planar polychlorinated biphenyl (environmental toxicant); acts on Ca^{2+} homeostasis and is teratogenic/ neurotoxic	Polychlorinated biphenyl	(14)	(15) (16)	
PBDE-99 	Flame retardant which belongs to the group of polybrominated diphenyl ethers (PBDEs); acts on Ca^{2+} homeostasis	Polybrominated diphenyl ether	(17) (18) (19)	(20) (21) (22) (23) (24)	

Arsenic trioxide 	Targets cellular SH groups, and has multiple toxic effects; used for the treatment of certain leukemias.	Arsenite (As_2O_3)	(25)	(26)	Sigma-Aldrich/ 11099

(1) Daston et al. 1991; (2) Robinson et al. 2010; (3) Hermsen et al. 2011; (4) Scott et al. 1989; (5) Welsch 2005; (6) Robinson et al. 2012b; (7) Theunissen et al. 2012; (8) Heusinkveld et al. 2013a; (9) Machera 1995; (10) Di Renzo et al. 2011c; (11) Zimmer et al. 2012; (12) Menegola et al. 2005; (13) Di Renzo et al. 2011b; (14) Johansson et al. 2006; (15) Piedrafita et al. 2008; (16) He et al. 2011; (17) Madia et al. 2004; (18) Schreiber et al. 2010; (19) Alm et al. 2010; (20) Darnerud 2008; (21) Costa et al. 2008; (22) Eriksson et al. 2002; (23) Eriksson et al. 2006; (24) Branchi et al. 2005; (25) Vahidnia et al. 2007; (26) Golub et al. 1998

Supplemental Material, Figure S3:



Supplemental material, Figure S3. PBPK modeling for the polybrominated diphenyl ethers PBDE-99. a Schematic representation of the PBPK model for PBDE-99, which was constructed based on data on tissue distribution, metabolism and excretion of PBDE-99 as described by Hakk et al. (2002) and Chen et al. (2006). The model contains a gastrointestinal lumen compartment (GI), two rapid equilibrium compartments (T1 and T2), a blood compartment (B), a lipophilic tissues compartment (F) representing adipose tissue and skin, and compartments for urinary and fecal excretion (Ur and Fe). **b** The exchange between blood and tissue compartments is described by first order rate constants which are reported with their parameter estimates. **c-f** Comparison between the PBPK model simulations and the values reported from Chen et al. (2006) is shown: upper panels show simulated (curves) and observed (symbols) **c** concentrations of PBDE-99 after a single oral dose of 1 µmol/kg, in blood, lipophilic and rapid equilibrium compartments and **d** the simulated and observed amounts of PBDE-99 in gut, feces and urine. Lower panels show simulated and observed **e** concentrations and **f** amounts of PBDE-99 after an intravenous dose of 1 µmol/kg, in blood, lipophilic and rapid equilibrium compartments and in gut, feces and urine, respectively.

Supplemental Material, Figure S4:

Information about clinical concentration ranges and free concentration ranges for chemicals belonging to the test battery compound list

For each compound tested in the test battery, clinical and epidemiological studies were chosen in order to determine a realistic clinical concentration range (CC) of the substance (total drug concentration in plasma). The free concentration range (FCC) was then calculated based on the plasma protein binding value. In case of environmental pollutants, it was not always possible to obtain data from human exposure studies; in these cases, in vivo and in vitro animal studies reporting developmental toxicity effects have been used to calculate the exposure range of interest.

Teriflunomide

MW: 270	CC: 10.8 µM	FCC: 54.0 nM

In Parekh et al. 2010, the mean pharmacokinetic parameters of teriflunomide are calculated following oral administration of 20 mg leflunomide tablet formulation in 12 healthy human subjects. The clinical concentration range extrapolated from the average maximal concentration (C_{max}) observed was of 2543 – 3299 ng/ml (9.4 – 12.2 µM).

The pharmacologically active metabolite is reported being extensively bound to plasma proteins (>99.3), primarily to albumin, with almost constant portion (0.5%) of free teriflunomide.

Based on these data, the free concentration range calculated was of 12.7 – 16.5 ng/ml (47.0 – 61.1 nM).

Nintedanib (BIBF 1120)

MW: 539	CC: 74.2 nM	FCC: 70.8 nM

In Mross et al. 2010, sixty-one patients with advanced cancers received BIBF 1120 in successive cohorts. Twenty-five subjects received 50 to 450 mg once daily and 36 received 150 to 300 mg twice daily in 4-week treatment courses interspersed by 1 week of washout.

The clinical concentration range extrapolated from the average C_{max} observed was 30 - 50 ng/ml (55.6 – 92.8 nM).

The portion of free nintedanib in plasma is predicted being 95.5%.

Based on these data, the free concentration range calculated was 28.6 – 47.7 ng/ml (53.1 – 88.5 nM).

Telaprevir

MW: 681	CC: 5.1 µM	FCC: 1.6 µM

In a phase 1b study, Yamada et al. 2012 examine safety, tolerability, pharmacokinetics of telaprevir in 10 patients infected with hepatitis C virus genotype 1b with high viral load (> 5 log10 IU/mL) and receiving 750 mg telaprevir every 8 h for 12 weeks.

The clinical concentration range observed, basing on the C_{max} values, is: 2 – 5 µg/ml (2.9 – 7.3 µM).

Telaprevir binds primarily to alpha 1-acid glycoprotein and albumin, in a concentration dependent manner. The drug is approximately 59% to 76% bound to human plasma proteins (compound data sheet).

Based on these values, the free concentration range calculated was 0.6 – 1.6 µg/ml (0.9 – 2.3 µM).

Sitagliptin

MW: 407	CC: 1.9 µM	FCC: 1.2 µM

In Herman et al. 2006b, the pharmacodynamics, pharmacokinetics, and tolerability of sitagliptin are examined after administration of single oral doses (25 or 200 mg) in a cohort of 58 patients with type 2 diabetes, not exposed to anti-hyperglycemic agents.

The clinical concentration range calculated basing on observed C_{max} was: 1.3 – 2.6 µM.

The level of binding to plasma proteins is of 38%. as reported in the pharmacokinetic study by Herman et al. 2006a.

The free concentration range was calculated, basing on the reported values and was of 0.8 - 1.6 µM.

Abiraterone

MW: 349	CC: 647.3 nM	FCC: 7.8 nM

Abiraterone is a novel potent, selective, irreversible inhibitor of CYP 17a-hydroxylase/C17,20-lyase enzyme. In Gurav et al. 2012, metastatic castration-resistant prostate cancer patients were followed after oral administration of arbirateron acetate (4*250 mg tablet).

The clinical concentration range calculated basing on the observed C_{max} was of 48 - 404 ng/ml (137.5 – 1157.6 nM). The ART is highly bound (>98.8%) to plasma proteins.

The free concentration range calculated from the reported values was 0.6 – 4.8 ng/ml (1.7 – 13.9 nM).

Roflumilast

MW: 403	CC: 18.0 nM	FCC: 0.2 nM

In de Mey et al. 2011, healthy men are treated with 500 µg tablet roflumilast once daily, for 16 days. The steady-state plasma pharmacokinetics of roflumilast as well as pharmacodynamics is evaluated on day 11.

Basing on the C_{max} observed, the clinical concentration range is of 4.94 – 9.64 ng/ml (12.2 – 23.9 nM).

The plasma protein binding of Roflumilast is 99% (Pinner et al. 2012).

Using the reported values, the free concentration range was calculated and was 0.0494 – 0.0964 ng/ml (0.12 – 0.24 nM).

Exenatide

MW: 4187	CC: 44.7 nM	FCC: 0.3 nM

In the assessment report compiled by the European Medicines Agency (EMEA) for the evaluation of medicines for human use, the maximal recommended human dose is 1.8 mg/day, which gives rise to a C_{max} of 44.7 nM.

In the study of Plum et al. 2013, *in vitro* protein binding of the drug in human plasma is evaluated and estimated being of 99.4%.

Based on these evidences, the calculated free concentration range was of 0.3 nM.

Gefitinib

MW: 446	CC: 1.2 µM	FCC: 34.0 nM

In Scheffler et al. 2011, maximum plasma concentration of gefitinib is evaluated after administration of multiple oral doses of 250 mg in patients with solid tumors.

The clinical concentration range reported from the obtained C_{max} is 265 – 814 ng/ml (594.1 – 1825.1 nM). In whole blood from cancer patients, 2.8% of free drug is observed.

Based on these evidences, the free concentration range calculated was of 7.4 – 22.8 ng/ml (16.6 – 51 nM).

Rivaroxaban

MW: 435	CC: 689.6 nM	FCC: 41.4 nM

In Mueck et al. 2011, a population pharmacokinetic model is developed using plasma samples from patients with acute deep-vein thrombosis. The clinical relevant peak concentration range is 200 – 400 ng/ml (459.8 – 919.5 nM).

In Grillo et al. 2012, a PBPK model is developed to simulate rivaroxaban pharmacokinetics in young (20 – 45 years) or older (55 – 65 years) subjects with normal renal function, mild, moderate and severe renal impairment, with or without concomitant use of the combined P-glycoprotein and moderate CYP3A4 inhibitor, erythromycin. The unbound fraction in plasma calculated in this study is of 6%.

Based on these values, the free concentration range calculated was of 12 – 24 ng/ml (27.6 – 55.2 nM).

Aliskiren

MW: 551	CC: 326.5 nM	FCC: 166 nM

Aliskiren pharmacokinetic and pharmodynamic parameters have been analyzed in the research of Tapaninen et al. 2011. In a randomized crossover study, 100 mg of the antifungal drug itraconazole, a P-glycoprotein and CYP-3A4 inhibitor, or placebo is given to 11 healthy volunteers twice daily for 5 days. On day 3, they ingest a single 150-mg dose of aliskiren, a renin inhibitor used in the treatment of hypertension. The extrapolation of the results from the average C_{max} showed a clinical concentration range of: 60 – 300 ng/ml (109.0 – 544.5 nM).

Protein binding of Aliskiren is reported to be moderate, in the range of 47-51% (Vaidyanathan et al. 2008). The free concentration range calculated using the reported values was 30.6 – 153 ng/ml (55.5 – 277 nM).

Galnon

MW: 679	CC: n.a.	FCC: n.a.

Neuregulin

MW: 40000	CC: 6.3 µM	FCC: n.a.

In Moondra et al. 2009, the mean serum neuregulin-1β levels in 9 healthy men range from 32 to 473 ng/ml (0.8 nM – 11.8 µM).

Erythropoietin

MW: 34000	CC: 0.3 nM	FCC: n.a.

In Xuereb et al. 2011, level of recombinant human erythropoietin is analyzed in whole blood from anonymous healthy volunteers. The clinical concentration range calculated was of 0.1 – 0.5 nM.

Geldanamycin

MW: 616 (17-DMAG)	CC: 0.8 µM	FCC: 0.5 µM

In Kummar et al. 2010, phase I dose-escalation study is used to determine the toxicity and maximum tolerated dose of 17-dimethylaminoethylamino-17-demethoxygeldanamycin (17-DMAG), a geldanamycin derivative, administered on a twice weekly schedule in patients with advanced cancer.

The clinical concentration range was calculated from the C_{max} value and it was 225 – 773 ng/ml (0.4 – 1.2 µM).

In Egorin et al. 2002, mice or rats were exposed to 17-DMAG i.v. bolus doses of 33.3, 50, and 75 mg/kg in order to perform pharmacokinetic studies.

From this study the level of 17-DMAG bound to plasma proteins is calculated as 30 – 45%.

The free concentration range predicted by using the reported values was: 140.6 – 483.1 ng/ml (0.2 – 0.8 µM).

G-CSF

MW: 19600	CC: 1.3 pM	FCC: n.a.

Laske et al. 2009 show that G-CSF (granulocyte colony-stimulating factor) plasma concentrations range from 10-40 pg/ml (average ~20 pg/ml) in 50 patients affected of Alzheimer's disease, compared to 5-80 pg/ml (average ~28 pg/ml) in 50 healthy controls.

Clinically relevant concentration calculated from the reported evidences was of 25 pg/ml (1.3 pM).

Interferon-β

MW: 22500	CC: 0.4 pM – 7.5 nM	FCC: 0.4 pM – 7.5 nM

The therapeutic concentration of IFN-β depends on the indication. Yung et al. 1991 demonstrate activity of IFN-β in patients with recurrent malignant glioma upon intravenous dose of 90 MIU three times per week, increasing the dose to 180 MIU after two weeks.

Neurotoxicity is dose-limiting, with adverse events noted at and above this active dosing pattern.

Kappos et al. 2004 recommend treatment with 8 MIU subcutaneous on alternating days for patients affected by multiple sclerosis (MS). Though not significant, a higher rate of spontaneous abortion is noted in presence of *in utero* IFN-β exposure. Similar effects during pregnancy are also noted in cynomolgus monkey in an unpublished study reported by the US FDA (1999). In this study pregnant cynomolgus monkeys exposed to intramuscular doses of 0.2 MIU or 0.033 nmol IFN-β/kg/day from gestation day 90 through term show an increase in spontaneous abortions and/or fetal loss. Although the effects show no apparent dose-response, they can be considered treatment-related in view of the reported abortifacient effects of other interferons.

We estimated the concentrations corresponding to these clinical and/or neurotoxic IFN-β1a doses using the PK-PD models for cynomolgus monkey and human published by Mager et al. (Mager and Jusko 2002, Mager et al. 2003). Simulated exposure to intravenous dose reported to be used in glioma patients by Yung et al. (1991) lead to a maximum free IFN-β plasma concentration of 7.5 nM. Simulation of the MS therapeutic dosing regimen of s.c. 8 MIU- IFN- β resulted in a maximum free plasma concentrations of 0.44 pM. Simulation of the subcutaneous dose of 0.033 nmol IFN-β/kg/day from GD 90 to term (GD 160) related to spontaneous abortions resulted in maximum plasma concentrations of IFN-β increasing from 1.9 on the first day to 3 pM at term.

Since aspecific binding is considered negligible, the free concentration range was considered corresponding to the clinical one.

Sildenafil

MW: 475	CC: 221.0 nM	FCC: 11.0 nM

In the study by Vachiery and collegueas (Vachiery et al. 2011), the pharmacokinetics and pharmacodynamics of a 10 mg intravenous sildenafil bolus is assessed in pulmonary arterial hypertension patients. The clinical concentration range observed is 30 – 180 ng/ml (63.1 – 378.9 nM).

In Purvis et al. 2002), a total of 17 healthy male volunteers are randomized to receive a single 100 mg dose of sildenafil for two periods and a single dose of placebo for two periods, with each period separated by a minimum of 5 – 7 days. Blood samples are collected before each dose and at different time points after dose for measurement of sildenafil and metabolite concentrations. In this study the plasma protein binding is reported being as approximately 95%.

Using the reported values, the calculated free concentration range was of 1.5 – 9.0 ng/ml (3.2 – 18.9 nM).

Imatinib

MW: 494	CC: 3.0 µM	FCC: 224.0 nM

In Di Gion et al. 2011, the reported clinical concentration range is 1000 – 2000 ng/ml (2.0 – 4.0 µM).

In Kretz et al. 2004 the blood distribution and protein binding of imatinib are determined *in vitro* using ^{14}C labelled compounds. Blood samples are taken from healthy males exposed to 300-500 ng/ml, 5000 ng/ml, 12000 ng/ml and 26000 ng/ml of imatinib. The average level of unbound fraction of compound in plasma calculated is 7.4 %.

The free concentration range calculated using the reported values was of 74 – 148 ng/ml (149.7 – 299.6 nM).

Oxytocin

MW: 1007	CC: 2.0 nM	FCC: 2.0 nM

Evaluating different studies (Opacka-Juffry and Mohiyeddini 2012; Pierrehumbert et al. 2010; Scantamburlo et al. 2007), the average clinical concentration range was estimated as 10 – 4000 pg/ml (0.01 – 4.0 nM).

In Fabian et al. 1969 study, it is stated that oxytocin in plasma remains unbound.

The free concentration range calculated basing on the reported studies was 0.01 – 4.0 nM (2.0 nM).

Sulfadiazine

MW: 250	CC: 320.0 µM	FCC: 160.0 µM

In Jordan et al. 2004, the pharmacokinetics of a dose of 2000 mg of sulfadiazine administered twice daily versus those of 1000 mg administered four times every day are compared in eight human immunodeficiency virus-infected patients. Serial blood samples are collected following administration of the morning dose on the fifth day after the initiation of each new regimen. Plasma samples are collected over 48 h and assayed by a validated high-performance liquid chromatography method.

The average C_{max} showed a clinical concentration range of 50 – 110 µg/ml (200.0 – 440.0 µM).

In Mannisto et al. 1982, the pharmacokinetic of sulphadiazine/trimethoprim (1000 mg/320 mg) is evaluated after intravenous infusion of the drugs over a 60 min period to six young, healthy volunteers. In this study the degree of protein binding is evaluated as 50%.

Based on the reported studies, the free concentration range calculated was 25 – 55 µg/ml (100.0 – 220.0 µM).

Amiodarone

MW: 645	CC: 2.3 µM	FCC: 0.5 nM

In the study of Shiga et al. 2011, thirty-two healthy Japanese male volunteers (20–32 years) are randomized to three single-dose groups (1.25, 2.5 and 5.0 mg/kg) of intravenous amiodarone. The average value of plasma concentration range is of 1 – 2 µg/ml (1.5 – 3.1 µM).

In Veronese et al. 1988, the plasma protein binding of amiodarone is measured by erythrocyte partitioning, and found to be the same in six healthy subjects and eight patients being treated for cardiac arrhythmias; the average value is of 99.98%.

Using the values reported from these studies, the free concentration range calculated was 0.2 – 0.4 ng/ml (0.31 – 0.62 nM).

Chlorpromazine

MW: 318	CC: 393.1 nM	FCC: 58.9 nM

In Borges et al. 2011, the values of plasma concentration of chlorpromazine is evaluated in 72 healthy volunteers of both sexes aged between 18-50 years, exposed to a single oral dose of 100 mg chlorpromazine. The clinically relevant concentration reported is of 100 – 150 ng/ml (314.5 – 471.7 nM).

In Yeung et al. 1993, the pharmacokinetics of chlorpromazine is investigated in 11 healthy young men after a bolus intravenous (i.v.) dose (10 mg) and three single oral doses (25, 50 and 100 mg), with a washout period of two weeks between doses. The drug is revealed being bound to the plasma protein albumin with a percentage around 85%.

The free concentration range calculated from the reported values was 15 – 22.5 ng/ml (47.2 – 70.7 nM).

Methoxyacetic acid

MW: 90	Human exposure concentration: 60 µM	DT concentration: 7.5 mM

Methoxyacetic acid (MAA) is the major metabolite of ethyl glycol monomethyl ether (EGME) a compound, widely used in the working environment (Henley and Korach 2010). Like valproic acid, chemically related to MAA, it inhibits histone deacetylates and alters gene expression via histone hyperacetylation. MAA is an endocrine disruptor and supposed

to be responsible for the reproductive toxicity of EGME (Foster et al. 1984). The mode of action is mainly via perturbations of estrogen signaling. a mechanism that could be used to link this compound to DNT. Estimation of human blood concentration for MAA due to EGME permitted occupational exposure is approximately 60 µM (Welsch 2005). PBPK simulation of *in vivo* adverse effects doses (in rats and mice) show a maternal plasma MAA maximal concentration of 6 – 9 mM (Welsch 2005).

In vitro studies confirm this concentration range; in Piersma et al. 2008, the BMC_5 of MAA in the whole embryo culture morphological assay is reported being of 1.6 mM, while the BMC_{50} in the embryonic stem cell differentiation assay of the Embryonic Stem cell Test (EST) is of 2.4 mM (de Jong et al. 2009).

Cyproconazole

MW: 292	Human exposure concentration: n.a.	DT concentration: 65-118 µM

Cyproconazole is an azole fungicide that inhibits ergosterol biosynthesis in fungi. The toxic mode of action of azole fungicides in mainly based on the inhibition of mammalian cytochrome P450 enzymes with a possible impact on retinoic acid (RA) concentration. No human data are available on teratogenic effects of cyproconazole (Giavini and Menegola 2010). In a rat study, cleft palate and hydrocepathy are identified as the most common malformations of the fetuses exposed to the fungicide, with a suggested NOEL of 20 mg/kg (Machera 1995). In an *in vivo* study by Hermsen et al. 2012, the exposure of zebrafish embryos to 65 µM cyproconazole show an increase in teratogenic effects such as pericardial edema and malformations of head and heart. In Robinson et al. 2012b, the teratogenic effects of cyproconazole are showed in rat whole embryo culture; the exposure to concentrations of 1 and 1.7 mM for 48 hours leads, respectively, to 50% and the 100% of malformations. In another *in vitro* study by Heusinkveld et al. 2013b, the neurotoxicity of cyproconazole is analyzed in dopaminergic PC12 cells by cell viability and intracellular calcium concentration level analysis; the inhibitory concentration of the azole fungicide on depolarization-evoked intracellular calcium concentration is 65 µM. In an *in vitro* study by Theunissen et al. 2012, neural embryonic stem cell test (ESTn) is used to evaluate the transcriptomic concentration-response to cyproconazole; concentration reducing cell viability to 80% is of 117.6 µM; the fungicide also showed to regulate genes within the neuron development GO-term (dendrite microtubule formation, early ectodermal development transcription factors).

Triadimefon

MW: 293.75	Human exposure concentration: n.a.	DT concentration: 27 – 500 µM

Triadimefon is a fungicide belonging to the conazole class, together with cyproconazole and its metabolite triadimenol. These pesticides are known to disturb steroid homeostasis in mammals (Crowell et al. 2011). A range of toxic effects is observed in rats at doses around 50 mg/kg/day, including effects on male fertility and CNS toxicity. Goetz et al. 2007 report increased testosterone levels and related effects (reduced testes weight, increased anogenital distance) in male offspring of Wistar rats perinatally exposed to 47 mg/kg/day.

Crofton et al. 2011 report that adult male Long-Evans rats exposed to a single dose of 50 mg/kg triadimefon or triadimenol by oral gavage show transient CNS toxicity as indicated by hyperactivity and stereotyped behavior. The same neurotoxic symptoms are observed upon chronic dietary exposure of male Wistar rats to 54.6 mg/kg/day (US EPA (2006)). Our simulation of this exposure suggested a corresponding central nervous system C_{max} of 27 µM and 196 µM (for dietary and acute exposure, respectively) (see in the text). Di Renzo et al. 2011a) observe the induction of abnormalities (craniofacial defects, bent forebrain and abnormal hindbrain segmentation) in Xenopus laevis embryos treated during early neurulation phases with 500 µM triadimefon. In vitro studies in rat embryos confirmed this toxic concentration range (Menegola et al. 2000, Di Renzo et al. 2009).

PCB-153

MW: 361	Human exposure concentration: 0.3 –1.8 nM	DT concentration: 388 pM

The range of human exposure to the PCB-153 is showed by Covaci et al. 2002 where the level of PCB-153 in maternal serum concentrations of 670 ± 350 pg/ml (0.3 - 1.8 nM).

In a study by Govarts et al. 2012, the PCB-153 concentration in maternal and cord blood is measured in a cohort of 7990 women. The median concentration of cords serum PCB-153 is 140 ng/l (388 pM). Linear regression of birth weight on estimated of cord serum concentrations show that low-level exposure to PCB impairs fetal growth, with an inverse association corresponding to a 150 g reduction per 1 µg/l increase in cord serum PCB-153.

PBDE-99

MW: 568	Human exposure concentration: 5 – 97 pM	DT concentration: 0.1 – 1.3 µM

The toxicity of PBDE-99 is reviewed by US EPA in 2008 (US EPA (2008)); this document collects diverse epidemiological studies, reporting the PBDE-99 concentration in different human biological samples. The concentration range in serum is between 0.6 ng/g lw and 11 ng/g lw. Under the assumption that there is 0.5% lipid in blood, total blood concentrations can be calculated; the resulting concentration range was between 3 ng/l and 55 ng/l (5 pM-97 pM).

In the study by Eskenazi et al. 2013, the combined effect of PBDEs on neurobehavioral readouts is investigated; especially high exposed children show pronounced deficits. The geometric mean of PBDE-99 concentration in maternal serum is of 4.5 ng/g, corresponding to 25 ng/l (44 pM). After PBDE-47, this congener yields the highest concentration in maternal serum and in child serum samples collected at 7 years of age. In Gascon et al. 2012, PBDE-99 concentrations in breast milk samples are provided; the average concentration is 0.27 ng/g lw, corresponding to 1 ng/l (1.8 pM). In both human developmental neurotoxicity studies, the behavioral effects are caused by mixed exposures and the specific contribution of this congener to the effects is unclear

A simpler exposure situation has been reproduced in *in vivo* studies; effects on neurodevelopment and fertility in adult rodents after perinatal exposure to low doses of PBDE-99 have been observed. Kuriyama et al. 2005 find hyperactivity and impaired spermatogenesis in Wistar rat offspring after exposure of the dams to a single oral dose of 0.11 µmol/kg on gestational day 6. Viberg et al. 2005 observe effects on spontaneous behavior of Sprague-Dawley rats exposed to doses between 1.4 and 14 µmol/kg on postnatal day 10. The simulation of this exposure suggested a corresponding C_{max} of 0.10 µM and 1.3 µM (respectively for the two studies) in the rapid equilibrium compartment (see in the text).

Arsenic Trioxide

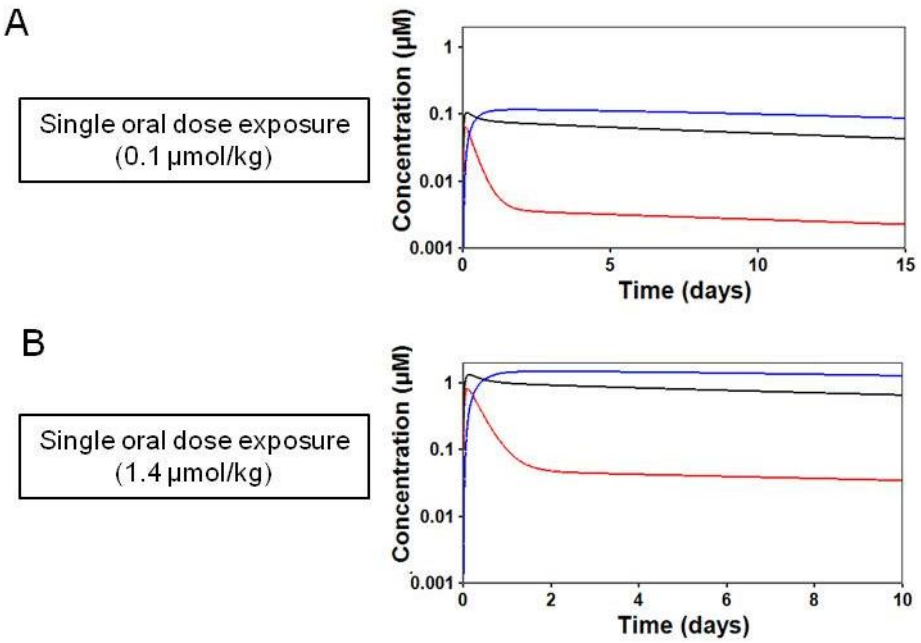
MW: 198	Human exposure concentration: 0.2 –1.1 µM	DT concentration: 0.2 –1.1 µM

Many case studies on arsenic poisoning have been reported, but they include the contribution of mixtures of various inorganic arsenic compounds. Inorganic arsenic compounds differ markedly with respect to their water solubility and arsenic trioxide belongs to the more water-soluble compounds, normally associated to a lower accumulation in soft tissues.

Arsenic is also used for its therapeutic effects in the treatment of acute promyielocytic leukaemia (APL); as side effects, many APL patients show development of neuropatological syntoms after injection with arsenic trioxide (up to 0.15 µg/kg daily) (Vahidnia et al. 2008). The toxicity mode of action of this drug is still unknown; in an *in vivo* study in rats, induction of axonal degeneration has been hypothesized as a possible toxicity-inducing mechanism (Vahidnia et al. 2008).

In the study by Hua et al. 2011, 10 mg arsenic trioxide is administered intravenously over a 4 hour period to Chinese primary hepatocarcinoma patients. Arsenic peak plasma concentrations are reported to be 47 – 225.8 ng/ml; (237.0 nM - 1140.4 nM). These doses have also been reported to be toxic. However, the therapeutic effect (anticancer) outweighs the toxic side effects, to a certain extent (Au and Kwong 2008). Zhang et al. 1998 show a ratio of 5.6% of total arsenic bound to serum proteins. Free plasma protein concentration range can be calculated as 233.7 – 1078.8 nM.

Supplemental Material, Figure S5:



Supplemental material, Figure S5. Simulation of toxic exposure to PBDE-99 by a PBPK modeling approach. Simulation of PBDE concentration in blood (red line), lipophilic (blue line) and rapid equilibrium (black line) compartment induced by a single oral dose of a 0.1 µmol/kg or b 1.4 µmol/kg in rats, using the PBPK model described in Fig. S3.

Supplemental material manuscript 2

Supplemental Material, Figure S1:

A

Truth \ Predicted as	As_2O_3	GA	PBDE-99	TDF	TSA	VPA
	Prediction frequency [%]					
As_2O_3	100.00	0.00	0.00	0.00	0.00	0.00
GA	0.00	99.95	0.00	0.00	0.05	0.00
PBDE-99	0.00	0.00	100.00	0.00	0.00	0.00
TDF	0.00	0.00	0.00	100.00	0.00	0.00
TSA	0.00	0.00	0.00	0.00	25.60	74.40
VPA	0.00	0.00	0.00	0.00	29.95	70.05

B

Blind Repl.	Best prediction	2nd best prediction	Truth
	Probabilities		
b1	PBDE-99 (0.50)	VPA (0.18)	*PBDE-99*
b2	PBDE-99 (0.54)	As_2O_3 (0.13)	*PBDE-99*
c1	*TDF* (0.49)	VPA (0.18)	*TDF*
c2	*TDF* (0.45)	VPA (0.25)	*TDF*
d1	GA (0.67)	As_2O_3 (0.10)	*GA*
d2	GA (0.70)	As_2O_3 (0.10)	*GA*
e1	VPA (0.56)	GA (0.13)	*VPA*
e2	VPA (0.42)	GA (0.31)	*VPA*
f1	As_2O_3 (0.40)	VPA (0.18)	As_2O_3
f2	As_2O_3 (0.50)	VPA (0.16)	As_2O_3

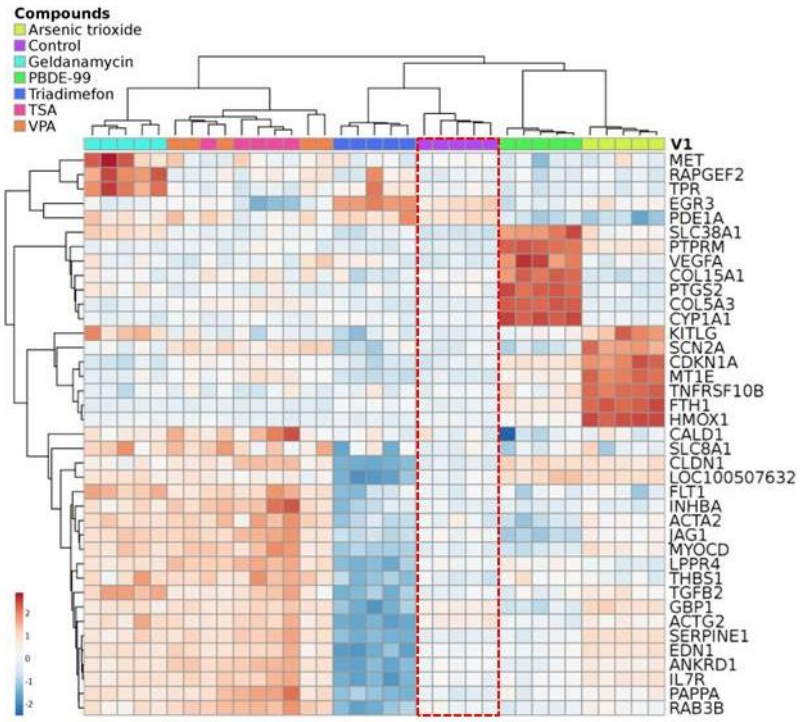
Supplemental material, Fig.S1. Evaluation of the SVM- based classifier. The support vector machine (SVM)-based classifier as described in Fig.2 was used to predict different scenarios. A A simulation study of the SVM-based classifier (Fig.2) was performed: three replicates (of 5) per compound were randomly chosen to form the training set and to build the classifier. The identity of the remaining 2 replicates (testing set) was then predicted. The procedure was reiterated for 1000 times. Finally, the best predictions were summed (considering together the replicates of the same compound) and normalized (n° prediction/ 2000 * 100). B The compound TSA was excluded from the training and testing set. The 100 probe sets with highest variance ("100 PS") within the training set were newly identified, and used to build a new classifier. The best and second best predictions, based on a support vector machine approach (indicated as relative probability in the brackets), are listed for each blind replicate (first column). The real identity of the samples (truth) is indicated in the last column.

Supplemental Material, Figure S2:

	Score						
	TDF	PBDE-99	TSA	VPA	As_2O_3	GA	TOT
SERPINE1	5	nd	3.5	2	5	6	21.5
CLDN1	2	6	5	nd	3	nd	16
THBS1	nd	5	5	2	nd	4	16
EDN1	3.5	nd	3	2	3	3	14.5
PTGS2	nd	7	nd	nd	7	nd	14
TGFB2	5	nd	5	nd	nd	4	14
TPR	nd	7	nd	nd	nd	7	14
ACTA2	5	5	nd	3	nd	nd	13
MT1E	nd	3	0	3	6	nd	12
MYOCD	nd	nd	2	3	2	5	12
PDE1A	nd	6	nd	nd	6	nd	12
ANKRD1	4	nd	3	2	nd	2	11
PAPPA	3	2	4	nd	nd	2	11
RAB3B	3	nd	3	2	nd	3	11
SLC8A1	nd	2	nd	nd	nd	8	10
CALD1	nd	5	0	3	nd	1	9
FLT1	nd	nd	3	nd	nd	6	9
INHBA	nd	nd	3.5	2	nd	3	8.5
COL5A3	nd	8	nd	nd	nd	nd	8
IL7R	3	nd	5	nd	nd	nd	8
LPPR4	3	nd	2	2	nd	1	8
COL15A1	nd	7	nd	nd	nd	nd	7
CYP1A1	nd	7	nd	nd	nd	nd	7
GBP1	4	3	nd	nd	nd	nd	7
HMOX1	nd	nd	nd	nd	7	nd	7
PTPRM	nd	7	nd	nd	nd	nd	7
SCN2A	nd	nd	nd	3	4	nd	7
VEGFA	nd	7	nd	nd	nd	nd	7
CDKN1A	nd	nd	nd	nd	6	nd	6
FTH1	nd	nd	nd	nd	6	nd	6
MET	nd	nd	nd	nd	nd	6	6
RAPGEF2	nd	nd	0	nd	nd	6	6
SLC38A1	nd	nd	nd	nd	nd	6	6
JAG1	nd	2	3.5	nd	nd	nd	5.5
KITLG	nd	nd	nd	nd	5.5	nd	5.5
TNFRSF10B	nd	nd	nd	nd	5.5	nd	5.5
EGR3	1	2	0	nd	2	nd	5
ACTG2	4	nd	nd	nd	nd	nd	4
LOC100507	nd	3	nd	nd	0	1	4

Supplemental material, Fig.S2. "Scoring approach" applied to the candidate biomarker genes. List of the candidate biomarker genes and their respective scores calculated per each compound using the "scoring"- algorithm showed in Fig.8A.

Supplemental Material, Figure S3:



Supplemental material, Fig.S3. Expression pattern of the candidate biomarker genes among the different exposure conditions. Heat map showing the expression values (expressed as median among the probesets of an individual gene) of each candidate biomarker gene in each exposure condition. Control group indicated by red line.

Record of contributions

Results manuscript 1

The experiments were designed by Bastian Zimmer, Tanja Waldmann and Marcel Leist. The experiments were performed by Bastian Zimmer and Nadine Dreser. Sunniva Förster contributed to the benchmark concentration (BMC) calculation. Joost Westerhout and Sieto Bosgra performed the PBPK analysis. Stephanie Julien and Karl-Heinz Krause provided the toxicity data in murine ESC system. Jan Hengstler, Agapios Sachidinis and Cristoph van Thriel contributed to the discussion of the paper. Marcel Leist and I designed the figures and wrote the manuscript in collaboration with Bastian Zimmer.
The manuscript is published in *Archives of Toxicology*.

Results manuscript 2

I conceived and designed the experiments in collaboration with Tanja Waldmann and Marcel Leist. I performed the major part of the experiments. Marianna Grinberg and Jörg Rahnenführer performed the statistical analysis of the gene array data. Margit Henry and Agapios Sachidinis performed the microarray analysis at the University of Cologne. Tancred Frickey contributed with the statistical analysis of the biomarkers selection. Marcel Leist and I designed the figures and wrote the manuscript in collaboration with Jan Hengstler.
The manuscript is published in *Archives of Toxicology*.

Results manuscript 3

I conceived and designed the experiments in collaboration with Tanja Waldmann. I performed the experiments. Johanna Nyffeler and Xenia Dolde gave experimental support for the establishment of the migration assay protocols and collaborated to the correction of the manuscript. Marianna Grinberg and Jörg Rahnenführer performed the statistical analysis of the gene array data, generated by the group of Agapios Sachidinis. I designed the figures and wrote the manuscript in collaboration with Marcel Leist.
The manuscript has been submitted to *Biochemical Pharmacology*.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Prof. Marcel Leist for the continuous scientific and personal support over my doctoral studies.

I would like to thank Prof. Daniel Dietrich and Prof. Thomas Brunner for having readily accepted to be part of my thesis committee.

Great thanks also goes to my supervisor Dr. Tanja Waldmann for her feedbacks to my thesis and her help in the critical points of my studies, together with all the rest of the UKN2 group, Johanna, Heidrun and Xenia.

I would like to thank all the colleagues and students that work or have been working in the AG Leist group, for the familiar atmosphere they have created in the lab. A particular thank goes to my 4-year long office-mates, Christiaan and Hanne, for their support and for all the early-lunch conversations about any imaginable topic.

Huge thanks goes to Steffi, Nadine and Simon, for their friendship and for welcoming me in their houses and lives since my very first day in Konstanz.

Many warm thanks to my volleyball team, the Seehasen, the best German family I could have ever hoped for.

I would like to greatly thank my family, without which I would have never reached this achievement. Thanks to my brother and sister for being always my first supporters and to my mum and dad for passing me down the values of ethics and animal sensitivity, which formed the roots of my studies.

Finally, huge thanks to my team-mate and life-mate, Malte, for lifting me up every time I felt down.