Leibniz Association

LEIBNIZ RESEARCH CENTRE FOR WORKING ENVIRONMENT AND HUMAN FACTORS

Transcriptomics-based identification of teratogens in differentiating hiPSCs: A novel in vitro test system

Dissertation

zur Erlangung des akademischen Grades des Doktors der Naturwissenschaften (Dr. rer. nat.) an der Fakultät für Chemie und Chemische Biologie der Technischen Universität Dortmund

vorgelegt von
Florian Seidel, M. Sc.

Dortmund, July 2021

1. Gutachter: Prof. Dr. Jan G. Hengstler
2. Gutachter: Prof. Dr. Carsten Watzl

Veröffentlicht als Dissertationsschrift an der Fakultät für Chemie und chemische Biologie der TU Dortmund.

Eidesstattliche Versicherung (Affidavit)

Name, Vorname
(Surname, first name)

Matrikel-Nr.
(Enrolment number)

Belehrung:

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden, § 63 Abs. 5 Hochschulgesetz NRW.

Die Abgabe einer falschen Versicherung an Eides statt ist strafbar.

Wer vorsätzlich eine falsche Versicherung an Eides statt abgibt, kann mit einer Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft werden, § 156 StGB. Die fahrlässige Abgabe einer falschen Versicherung an Eides statt kann mit einer Freiheitsstrafe bis zu einem Jahr oder Geldstrafe bestraft werden, § 161 StGB.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Official notification:

Any person who intentionally breaches any regulation of university examination regulations relating to deception in examination performance is acting improperly. This offence can be punished with a fine of up to EUR 50,000.00. The competent administrative authority for the pursuit and prosecution of offences of this type is the chancellor of the TU Dortmund University. In the case of multiple or other serious attempts at deception, the candidate can also be unenrolled, Section 63, paragraph 5 of the Universities Act of North Rhine-Westphalia.

The submission of a false affidavit is punishable.

Any person who intentionally submits a false affidavit can be punished with a prison sentence of up to three years or a fine, Section 156 of the Criminal Code. The negligent submission of a false affidavit can be punished with a prison sentence of up to one year or a fine, Section 161 of the Criminal Code.

I have taken note of the above official notification.

Ort, Datum
(Place, date)

Unterschrift
(Signature)

Titel der Dissertation:
(Title of the thesis):

Ich versichere hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel selbstständig und ohne unzulässige fremde Hilfe angefertigt habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht.
Die Arbeit hat in gegenwärtiger oder in einer anderen Fassung weder der TU Dortmund noch einer anderen Hochschule im Zusammenhang mit einer staatlichen oder akademischen Prüfung vorgelegen.

I hereby swear that I have completed the present dissertation independently and without inadmissible external support. I have not used any sources or tools other than those indicated and have identified literal and analogous quotations.

The thesis in its current version or another version has not been presented to the TU Dortmund University or another university in connection with a state or academic examination.*

*Please be aware that solely the German version of the affidavit ("Eidesstattliche Versicherung") for the PhD thesis is the official and legally binding version.

Ort, Datum
(Place, date)

Unterschrift
(Signature)

Table of contents

Table of contents

Summary

The knowledge about the developmental toxicity potential of substances is a matter of high concern in toxicology and medicine, especially if pregnant women need to take potentially teratogenic drugs to treat severe infections or diseases. The regulatory risk assessment of such substances is, however, very difficult, costly and time-consuming and relies currently, due to a lack of appropriate alternatives, solely on animal studies. During the last decades, much effort was put into the development and establishment of trustworthy in vitro-based test systems for the risk assessment of teratogenic substances, but not even one was approved for this purpose yet. In this PhD project, a new attempt was undertaken to develop such a system. The here presented novel system is based on the UKN1 assay, in which pluripotent human embryonic stem cells were differentiated to neuroepithelial precursor cells, and made considerable innovations to it by utilizing human induced pluripotent stem cells (hiPSCs), whole-transcriptome analyses and a thoroughly chosen set of 39 teratogenic and non-teratogenic substances that each were applied at two on human in vivo data-based concentrations to differentiating hiPSCs in this “UKN1 6-day” assay. In contrast to the non-teratogenic compounds, the teratogens induced either significant gene expression alterations or a high toxicity in almost all cases, so that this test system was able to predict substances with a high accuracy of 90 %, a robust sensitivity of 83 % and an outstanding specificity of 100 %, what makes it therefore to a very high performing in vitro test system. Even a variation of the system, the UKN1 one-day assay, which shortened the total time requirements of the assay and replaced the whole-transcriptome analyses by a four-biomarker-based RT-qPCR, could predict the compounds with an accuracy, sensitivity and specificity of 77 %, 63 % and 100 %, respectively. In addition to the core set of 39 substances, the UKN1 6-day assay was used to determine the health risks of parabens which are widely used as preservatives in cosmetics. No significant gene alterations were found in differentiated, paraben-exposed hiPSCs, and the obtained negative in vitro results could now support the existing in vivo data of paraben-exposed animals in accordance with a read-across-approach for ethylparaben, so that no further animal experiments would be needed for the regulatory risk assessment of ethylparaben.
In the near future, steps for a further optimization of the test system will be investigated like the application of targeted instead of whole-transcriptome analyses and multiple concentrations instead of just two. Furthermore, the number of tested compounds shall be increased. Although there is still a long way to go and many hurdles to overcome until this or another test system will be approved for the regulatory risk assessment of teratogens, the UKN1 6-day test system can already provide a tool for an in vitro teratogenicity screening of compounds and aid authorities and companies to assess the health risks of potentially teratogenic substances.

Zusammenfassung

Das Wissen über das entwicklungstoxische Potenzial von Substanzen ist in der Toxikologie und Medizin von großer Bedeutung, insbesondere wenn schwangere Frauen potenziell teratogene Medikamente zur Behandlung schwerer Infektionen oder Krankheiten einnehmen müssen. Die regulatorische Risikobewertung solcher Substanzen ist jedoch sehr schwierig, kostspielig und zeitaufwendig und stützt sich derzeit, mangels geeigneter Alternativen, ausschließlich auf Tierversuche. In den letzten Jahrzehnten wurde viel Aufwand in die Entwicklung und Etablierung von vertrauenswürdigen in vitro-basierten Testsystemen für die Risikobewertung von teratogenen Substanzen gesteckt, jedoch wurde bisher kein einziges für diesen Zweck zugelassen. In diesem Dissertationsprojekt wurde ein neuer Versuch unternommen, ein solches System zu entwickeln. Das hier vorgestellte, neuartige „UKN1-6-Tage"-System basierte auf dem UKN1-Assay, in welchem pluripotente humane embryonale Stammzellen zu neuroepithelialen Vorläuferzellen differenziert wurden, und verbesserte dieses erheblich durch die Nutzung von humanen induzierten pluripotenten Stammzellen (hiPSCs), transkriptomweiten Analysen und einem sorgfältig ausgewählten Set von 39 teratogenen und nicht-teratogenen Substanzen, die jeweils in zwei in auf humanen in vivo Daten-basierten Konzentrationen eingesetzt und mit denen differenzierende hiPSCs inkubiert wurden. Im Gegensatz zu den nicht-teratogenen Substanzen induzierten die Teratogene in fast allen Fällen entweder signifikante Genexpressionsveränderungen oder eine hohe Toxizität, so dass dieses Testsystem in der Lage war, Substanzen mit einer hohen Genauigkeit von 90 %, einer robusten Sensitivität von 83 % und einer herausragenden Spezifität von 100 % vorherzusagen, was es somit zu einem überaus leistungsfähigen in vitro Testsystem macht. Auch eine Variation des Systems, der UKN1-Ein-Tages-Assay, bei dem der gesamte Zeitaufwand des Assays verkürzt und die Transkriptomanalyse durch eine Vier-Biomarker-basierte RT-qPCR ersetzt wurde, konnte die Substanzen mit einer Genauigkeit, Sensitivität und Spezifität von 77 %, 63 % bzw. 100 % vorhersagen. Zusätzlich zum Kernsatz von 39 Substanzen wurde der UKN1 6-Tage Assay verwendet, um die Gesundheitsrisiken von Parabenen zu bestimmen, die als Konservierungsmittel in Kosmetika weit verbreitet sind. Es wurden keine signifikanten Genexpressionsveränderungen in differenzierten, Paraben-exponierten hiPSCs gefunden, und die erhaltenen negativen in vitro-Ergebnisse könnten nun die bestehenden in vivo-Daten von Paraben-exponierten Tieren gemäß einem Read-Across-Ansatz für Ethylparaben unterstützen, so dass keine weiteren Tierversuche für die regulatorische Risikobewertung von Ethylparaben erforderlich wären.
In naher Zukunft werden Schritte zur weiteren Optimierung des Testsystems unternommen, wie z. B. die Anwendung gezielter statt ungezielter Transkriptomanalysen und von mehr Konzentrationen statt nur zwei. Außerdem soll die Anzahl der getesteten Stoffe erhöht werden. Obwohl es noch ein weiter Weg ist und viele Hürden zu überwinden sind, bis dieses oder ein anderes Testsystem für die behördliche Risikobewertung von Teratogenen zugelassen wird, kann das UKN1 6-Tage Testsystem bereits jetzt ein Werkzeug für ein In-vitro-Teratogenitäts-Screening von Substanzen bereitstellen und Behörden und Unternehmen bei der Bewertung der Gesundheitsrisiken von potenziell teratogenen Stoffen unterstützen.

Abbreviations & units

Table 1: Abbreviations and their explanations

Abbreviation	Meaning
AUC	Area under curve
BMP	Bone morphogenic protein
cDNA	Complementary deoxyribonucleic acid
C_{max}	Therapeutic plasma/blood concentration
CNS	Central nervous system
CO_2	Carbon dioxide
cRNA	Complementary ribonucleic acid
C_T	Threshold cycle
DAPI	4',6-diamidino-2-phenylindole
DART	Developmental and reproductive toxicology / toxicity
DEPC	Diethylpyrocarbonate
DHODH	Dihydroorotate dehydrogenase,
D_i	Developmental index
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DNT	Developmental neurotoxicology /-toxicity
DoD	Day of differentiation
D_p	Developmental potency
DPBS (+/+)	Dulbecco's phosphate-buffered saline with calcium and magnesium
DPBS (-/-)	Dulbecco's phosphate-buffered saline without calcium and magnesium
DT	Developmental toxicology / toxicity
E6/8	Essential 6/8 medium
E8-S	Essential 8 supplement
EC	European commission
EC50	Effective concentration that causes a distinct effect (here: toxicity/death) to 50 % of the cells
ECHA	European Chemicals Agency
ECVAM	European Centre for the Validation of Alternative Methods
EDC	Endocrine disrupting chemical
EMX2	Empty spiracles homeobox 2
EPA	Environmental Protection Agency
EU	European Union
FC	Fold-change
FDA	Food and Drug Administration
FDR	False discovery rate
FN	False negative
FOXA2	Forkhead box A2
FOXG1	Forkhead box G1
FP	False positive

Abbreviation	Meaning
fRMA	Frozen robust multiarray analysis
g	Gravitational acceleration
GF	Growth factor
HCS	High-content screening
HDACi	Histone deacetylase inhibitor
hESC	Human embryonic stem cell
hiPSC	Human induced pluripotent stem cell
HMG-CoA	3-hydroxy-3-methylglutaryl-coenzyme A
IM	Intramuscular
iPSC	Induced pluripotent stem cell
IV	Intravenous
IVT	In vitro transcription
LiCl	Lithium chloride
LOOCV	Leave-One-Out-Cross-Validation
MeHg	Methyl mercury
mEST	mouse embryonic stem cell test
$MgCl_2$	Magnesium chloride
MgOH	Magnesium hydroxide
$MgSO_4$	Magnesium sulfate
MoA	Mechanism of action
mRNA	Messenger ribonucleic acid
MEIS2	Meis homeobox 2
MSX1	Msh homeobox 1
NANOG	Nanog homeobox
NEP	Neuroepithelial progenitor (cell)
NH	Null hypothesis
NOAEL	No Observed Adverse Effect Level
NPC	Neural progenitor/precursor cell
NSC	Neural stem cell
NTD	Neural tube defect
OCT4A	also known as POU5F1 (POU class 5 homeobox 1) alpha
OECD	Organisation for Economic Co-operation and Development
OTX2	Orthodenticle homeobox 2
PAX6	Paired box 6
PBPK	Physiologically-based pharmacokinetic (modelling)
PCA	Principal component analysis
PCR	Poly chain reaction
PFA	Paraformaldehyde
PO	Peroral
PPPC	Predicted probability for positive class
PSC	Pluripotent stem cell
qPCR	quantitative real-time poly-chain reaction
RAR	Retinoic acid receptor
RDA	Recommended daily allowance

Abbreviations & units

Abbreviation	Meaning
REACH	Registration, Evaluation, Authorisation and Restriction of Chemicals
RNA	Ribonucleic acid
RNase	Ribonuclease
Rocki	Rho-kinase inhibitor
ROS	Reactive oxygen-species
RT	Reverse transcription
RT-qPCR	Reverse transcription quantitative real-time poly-chain reaction
RXR	Retinoid x receptor
SMAD	Nomenclature refers to the human homologue of Caenorhabditis elegans protein SMA and MAD genes in Drosophila
SOX1/2	SRY (sex determining region Y)-box transcription factor 1/2
SPS	Significant probe set
SSEA4	Stage-specific embryonic antigen-4
TBP	TATA-binding protein
TFAP2B	Transcription factor activating enhancer binding protein 2 beta
TG	Test guideline
TGA	Therapeutic Goods Administration
TGF	Transforming growth factor
TN	True negative
TP	True positive
TRA-1-60/81	T cell receptor alpha locus 1-60/81
TSA	Trichostatin A
VPA	Valproic acid

Table 2: Physical quantities, belonging units and their abbreviations

Physical quantity	Unit(s)	Abbreviation(s)
Time	Second, minute, hour, day	s, min, h, d
Volume	Liter	l
Mass	Gram	g
Molar concentration	Moles per liter	mol/l or M
	Prefix: pico- (10^{-12}), nano- (10^{-9}), micro- (10^{-6}), milli- (10^{-3}), centi- (10^{-2}), dezi- (10^{-1}) (x = liter / gram / moles per liter)	px, nx, µx, mx, cx, dx
Acceleration	Revolution per minute, relative centrifugal force	rpm, rcf
	delta (“difference”)	Δ

1. Introduction

1.1 Developmental (neuro)toxicology & teratogenicity

Developmental toxicology (DT) is a research field within toxicology that investigates adverse effects of substances on the development of animals and humans (Klaassen 2018; Ujházy et al. 2012). It deals with all kinds of anatomical malformations, physiological disfunctions and mental disorders, which can be manifested at any age, but are all caused by a substance during the organismic development, meaning both prenatally between fertilization and delivery, and postnatally until the end of adolescence. However, the majority of DT-studies and -articles focus on effects of substances that take place during the prenatal phase as this period is considered to be much more susceptible to toxicants than the postnatal phase, shown by history and several different substances (Selevan et al. 2000). In this phase, developmental toxicity is also called teratogenicity, and teratogenicity-inducing agents are called teratogens. Among these substances, the most infamous one is thalidomide, in Germany well-known under the name Contergan, which was responsible for the worst drug-induced disaster in history (Vargesson 2015). In the late 1950's, it was given as an over-the-counter drug to pregnant women as a sedative or to treat morning sickness and insomnia, with a catastrophical result: Over 10,000 newborns suffered from numerous damages and malformations, with deformed limbs as the most common damage, but also internal organs defects or even damages to the central nervous system (CNS). However, thalidomide does not remain the only potent teratogen. Valproic acid (VPA), widely used as an anti-convulsant and mood-stabilizer, acts mainly as a neurodevelopmental teratogen and is associated with spina bifida, neural tube defects, autism and other defects like organ damages as well as the "fetal valproate syndrome" in humans (Ornoy 2009), and related effects could also be found in mice (Roullet et al. 2010). Another substance is isotretinoin, which is used to treat acne topically and found to be highly teratogenic in humans and animals (Browne et al. 2014; Nau 2001). Two more examples of well-known teratogens are methotrexate (Verberne et al. 2019; Hyoun et al. 2012) and carbamazepine (Matalon et al. 2002).

Due to its relevance, the specialized area "developmental neurotoxicology" (DNT) emerged within the general DT. It focuses on the mental impairments after birth, which were caused by toxicant-induced damages on the neural network's structure and functionality (William Slikker, JR et al. 2018). Like in general DT, neurodevelopmental damage occurs typically, but not exclusively, during early stages of development, which are fundamentally more susceptible than the adult nervous system (Bondy and Campbell 2005; Rodier 1995; Rice and Barone 2000; Bennett et al. 2019). The manifestation of mental defects caused by such damage, however, can be very subtle and delayed by many years, underlining the importance of DNT-studies. To date, compounds of many different classes have been identified to contribute to an increased risk for cognitive, behavioral or social impairments, or even to more specific disorders like attention deficit hyperactivity disorder or autism in humans (Grandjean and Landrigan 2006; 2014; Lanphear 2015). Usually, substances, which are known to cause DNT, are of environmental nature, and some prominent ones are the metals lead (Landrigan et al. 1975; Needleman et al. 1979; Lanphear et al. 2005; Eubig et al. 2010), mercury (Karagas et al. 2012; Grandjean et al. 1997; Harada 1995; EPA 2012, p. 3-85 to 3-93) and manganese (Khan et al. 2012; Riojas-Rodríguez et al. 2010; Lucchini et al. 2012; Moreno et al. 2009), organophosphate pesticides like chlorpyrifos and dichlorodiphenyltrichloroethane (Bouchard et al. 2011; Rauh et al. 2011; 2012; Torres-Sánchez et al. 2013), the flame retardants polybrominated diphenyl ethers (Chen et al. 2014; Cowell et al. 2015; Eskenazi et al. 2013) and their structural relatives and plasticizers polychlorinated biphenyls (Jacobson and Jacobson 1996; Schantz et al. 2003).

1.2 Risk assessment

1.2.1 Identification of possible in vivo teratogens and DNT-substances

Teratogens and neurodevelopmental substances are identified in two ways: Accidental incidences in humans (see thalidomide and mercury) or in vivo animal studies. These studies, mostly utilizing rodents and other small species like rabbits, are regularly used in toxicology, pharmacology and other research fields to assess the risk of many substances, both novel and established ones, and especially those with suspected detrimental effects, in order to avoid any harm for human beings (Meigs et al. 2018; Fielden and Kolaja 2008). The test strategies for regulatory purposes are given by specific guidelines, and for assessing the risk of potential teratogens and DNT-substances these are mainly the OECD test guideline (TG) 414 for prenatal developmental toxicity studies (OECD 2018a), TGs 415 and 416, both combining developmental and reproductive toxicity (DART) in a one- and two-generation testing procedure, respectively (OECD 1983, 2018b), TGs 421 and 422 for a more rapid DART-screening (OECD 2016, 1996) and TGs 426 (DNT-study) and 443 (extended one-generation DART-study), taking also neurodevelopmental damage into account (OECD 2007, 2012). A recent example of extensive in vivo testing on DART is the regulation for "Registration, Evaluation, Authorisation and Restriction of Chemicals" (REACH), which was drafted by the "European Chemicals Agency" (ECHA) and came into force on the 1st July 2007. According to REACH, Annex VIII, a substance has to be tested on DART if it is marketed or produced in quantities above 10 tonnages / year (ECHA 2006). Rovida et al. 2011 estimated, that this will require almost 1,000,000 animals and 141,000,000 € costs.

1.2.2 Parabens

Due to REACH, a special group of compounds came into focus: Parahydroxybenzoates, esters of the 4-hydroxybenzoic acid, also called "Parabens", which were widely used as preservatives in food, drugs and especially cosmetics due to their high antimicrobial and antifungal effectiveness (Aalto et al. 1953). Although parabens didn't show any evidence for adverse effects in animals by 2008 (Cosmetic ingredient review 2008) and they were chosen as the (non)allergen of the year by the North American Contact Dermatitis Group (Fransway et al. 2019), they had to be tested on their DART- and DNT-potential in order to fulfil the requests of REACH. Besides that, they arouse attention as potential endocrine disrupting chemicals (EDC) in the last couple of years and are therefore still considered to corrupt human health (Nowak et al. 2018). A recent review from Matwiejczuk et al. 2020 objectively summarized an extensive amount of publications and concluded that the use of single paraben-containing cosmetics can be regarded as safe and harmless. However, the exposure to parabens may exceed a safe range when cosmetics are used excessively, and their capability to have detrimental effects on the human health, in particular in their proposed action as EDCs, needs to be addressed by large-scale and long-term scientific studies.

In the last decade, authorities of the European Union (EU) already enacted important regulations to restrict the use of parabens and other chemicals found in cosmetics. In 2009, the parabens' use in cosmetics was limited to 0.4 % and 0.8 % as a single component or as a mix, respectively, by the Regulation (EC) No 1223/2009 of the European Parliament, which set new rules for the marketing of cosmetics to "ensuring a high level of protection of human health" (EC 2009). In April 2014, the European Commission (EC) banned all parabens from further use except for methyl-, ethyl-, propyl- and butyl-4-hydroxybenzoate (EC 2014) as these were the only parabens which could provide enough data for a reassessment. Only one year later, in April 2015, the application of propyl- and butylparaben was even more reduced to a maximum of 0.14 % per single compound or mix and banned completely for distinct children-related applications (European Commission 2014), leaving methyl- and ethylparaben as the only parabens which can be used in higher concentrations. Recently, in June 2020,

butylparaben was listed as a substance of very high concern, aiming at its soon phase-out and replacement by safer alternatives (European Chemicals Agency 2020).
From the three remaining parabens, methyl- and propyl-4-hydroxybenzoate are very well characterized. Studies regarding their toxicity, including such to assess the developmental toxicity, which were provided for the registration in ECHA, weren't able to find any evidence for teratogenicity, thus setting the No Observed Adverse Effect Level (NOAEL) to the very high dosis of 1000 mg/kg bw/day (ECHA n.d.c, n.d.b). However, ethyl-4-hydroxybenzoate is less characterized with no available data about its DT (ECHA n.d.a), so to date extensive animal experiments for it are required as well.

1.3 In vitro methods for teratogenicity- and DNT-testing

High costs for animal in vivo studies as well as concerns about ethics and animal welfare lead very early to the development of the 3R-principle. Firstly mentioned in 1959, the authors William Russel and Rex Burch defined in their book "The Principles of Humane Experimental Technique" these three R's, i.e. the Reduction of animals used in experiments, the Refinement of experiments to minimize the animals' suffering and the overall Replacement of animal experiments (Russell W and Burch R 1959). Today, the 3R-pinciple is covered by law in many countries, for example in the USA by the Animal Welfare Act (United States Code 2014; Code of Federal Regulations 2014) or in Germany by the Animal Protection Act and the Animal Protection Act for Test Animals based on the European Directive 2010/63/EU (EU 2010; Federal Ministry of Justice and Consumer Protection 2015, 2020).
A fundamental help in the reduction of animal numbers used in research and risk assessment in the context of the 3R's is the utilization of in vitro methods. Those methods usually use established cell lines and are not in need of living animals, so they are advantageous in regard of ethics and costs. The EpiDerm model for example uses an in vitro 3D epidermal model, is compliant with the OECD TG 439 and replaces rabbits for skin irritation tests (OECD 2015; Kandárová et al. 2009). Although not a single in vitro test has been approved for the use in DT-testing yet, myriads of different protocols are under development and recent advances are promising. Basically, in vitro methods for DT-assessment feature three aspects:

1.3.1 1st feature: Model organisms

Pluripotent stem cells

Current protocols favorite three different organisms: Stem cells, zebrafish and, especially for DNT-studies, neural cells. Stem cell-based DT-methods represent the majority of alternative approaches to assess DT. The biggest advantage of these methods is underlain by the use of pluripotent stem cells (PSC), which can in principle differentiate to all cell types and therefore offer a powerful in vitro model for early embryonic development (Wu and Izpisua Belmonte 2015). Plenty of protocols involving PSCs are known, but the most famous one is the mouse embryonic stem cell test (mEST), which was validated by the European Centre for the Validation of Alternative Methods (ECVAM) in 2002 (Genschow et al. 2002). It had been used to test the DT-potential of metals (Stummann et al. 2007, 2008), pharmaceuticals (Paquette et al. 2008; Eckardt and Stahlmann 2010), environmental pollutants (Kamelia et al. 2017; Zhou et al. 2017) and more. However, development of pluripotent stem cell models like the mEST took a long way. In 1981, Evans and Kaufman established the first in vitro culture of pluripotent mouse embryonic stem cells (mESC). In 1998, almost 20 years later, Thomson et al. isolated and established five human embryonic stem cell lines (hESC) from in-vitro-fertilized embryos, enabling researchers to work with human-derived stem cells for the first time. One of those cell lines, H9, is well-known and was widely used in all kinds of scientific and medical research. An important

milestone was achieved in 2006 and 2007 when Yamanaka and his fellow-scientists discovered that adult cells can be reprogrammed to pluripotent stem cells by defined factors (Takahashi and Yamanaka 2006; Takahashi et al. 2007). The so-called induced pluripotent stem cells (iPSC), particularly the human-derived iPSCs (hiPSC), overcame the hitherto given issues of ethics and limited access to hESCs and furthermore opened the doors for the generation of patient- and disease-specific stem cells in medical research. The recognition of these significant findings was finally rewarded with the Nobel Prize for Medicine in 2012, awarded to Shinya Yamanaka and John Gurdon.

Zebrafish

The second organism, zebrafish, is actually a living animal and per definition in vivo, but its ethical and monetary concerns are comparable to classical in vitro methods and thus zebrafish is mentioned here as an encouraging alternative to classical rodent in vivo experiments. Zebrafish as an in vitro- / in-vivo-hybrid-model organism has a lot of strengths which makes it suitable for toxicological tests (Horzmann and Freeman 2018). Due to their small size, the husbandry is easy and cost-effective. They are fertilized ex vivo, and their development is well-described with all relevant steps completed after 72 h. Additionally, in early developmental steps embryos are translucent (Kimmel et al. 1995). These features for example allow to simultaneously screen for numerous water-soluble toxicants in zebrafish-containing multiwell-plates and an optical outcome-assessment, as shown by Ali et al. 2014.. In 2018 and 2019, 123 polycyclic aromatic hydrocarbons were screened for their developmental toxicity, with 16 of them deeply characterized (Geier et al. 2018; Shankar et al. 2019). In 2020, Gaballah et al. identified unique toxicity signatures among distinct per- and polyfluoroalkyl substances in zebrafish. Moreover, the zebrafish's genome is known (Howe et al. 2013), which set the basis for genetic editing (Varshney et al. 2015) and thus further expanding the animal's applicability. It is also suitable for the investigation of behavioral effects during DNT-studies (Bailey et al. 2013), showing the versatileness of zebrafish as a model organism in DT research.

Neural cells

Neural cells as the third important model organism in DT are regularly used to study the occurrence of substance-induced neurodevelopmental damage. The specific cell types, which are applied in this context, include neural progenitor cells (NPC), neuronal cells and glial cells, with neuronal and glial cells usually derived from the first-mentioned NPCs. NPCs itself can be derived from primary neural tissue or PSCs. Human brain fetal tissue-derived primary NPCs can recapitulate relevant developmental processes of the brain, e.g. proliferation, differentiation, migration and apoptosis, when they are cultivated as neurospheres in vitro (Fritsche et al. 2011; Schmuck et al. 2017), but also PSC-derived NPCs are suitable for DNT-testing, though the distinct source yields some differences regarding the differentiation outcomes of the respective NPCs (Hofrichter et al. 2017). NPCs can be used to study their migration capacities upon substance exposure (Moors et al. 2007), but most laboratories focus on the generation of NPC-derived neuronal and glial cells for their DNT-experiments, either before or during the substance exposure (Fritsche et al. 2005; Pistollato et al. 2014; 2017; Druwe et al. 2016; Ryan et al. 2016).

Taken together, the three shown organisms act currently as a kind of gold-standard in the development of new in vitro methods for DT-testing. Human iPSC (hiPSC)-based methods offer a powerful tool in the in vitro examination of both DT and DNT on the molecular level, and the recruitment of zebrafish fills a gap between classical in vitro and in vivo experiments by providing a more complex model for an organism-wide DT-assessment as well as for behavioral DNT-studies.

1.3.2 2nd feature: Read-Outs

The effects of substances which organisms are exposed to can be detected in many different ways. An overview of important read-outs in DT-testing is given below:

Cytotoxicity

Toxicity of compounds onto cells is called cytotoxicity and it is the easiest read-out among all. Basically, established methods determine the amount of living cells in both substance-exposed and control conditions and compare them to calculate the cytotoxic effect of the used substance. The best-known protocol to assess the viability of the cells utilizes the blue redox-dye resazurin which is able to measure the metabolic activity of cells (O'Brien et al. 2000). Resazurin is taken up by living cells and converted to the highly fluorescent, pink resorufin. Cytotoxic substances, which are adversely affecting or even killing cells, reduce the overall metabolic activity of the cells and the corresponding fluorescence signal emitted by resorufin. A strong reduction of fluorescence therefore means a strong cytotoxic effect of the substance at the given concentration. The term “half maximal effective concentration” or “EC_{50}”, originating from pharmacology, for example indicates in toxicology a substance’s concentration, at which a 50 % reduction of viable cells in the given system could be observed (Neubig et al. 2003).

Histology, Immunostaining & High-Content-Screening (HCS)

In classical in vivo studies, a common way to determine the DT of a compound is to examine the offspring of the substance-exposed mother animal histologically at the tissue-level on morphological abnormalities and malformations, as long as the substance causes such defects. For example, Sucov et al. 1995 showed, that an in utero exposure to retinoic acid was responsible for limb defects in wild type mice, but not in RXR$\alpha^{-/-}$ mice. Like in that study, histological assays are normally supported by distinct stainings like Alcian blue or Alizarin red, which are used to stain cartilage and bone, respectively, and to facilitate conclusions drawn from histological findings (McLeod 1980). Another type of staining is immunostaining, which was firstly described by Coons et al. 1941 and which developed to an important technique since then, especially if small-scale imaging on the cellular level is needed. Briefly, immunohistochemistry and related methods utilize specialized antibodies which detect distinct antigens on a surface or within tissues and cells (Ramos-Vara 2005). The used antibodies are further tagged with dyes or other molecules which enable the specific staining of the antigen. Doing so, cellular substance-induced effects can be observed, as for example the impact of VPA on the differentiation of spiral ganglion neural stem cells (NSCs) (Moon et al. 2018) or on hair cell regeneration in zebrafish (He et al. 2014), leading to a better understanding of the specific molecular mechanisms that underlay organism-wide substance effects.
In high-content screening (HCS) approaches, imaging of macro- or microscopic pictures like stained tissues or cells is automated with high-throughput imaging devices and combined with an automated, quantitative, software-driven image analysis (Li and Xia 2019). With HCS, the otherwise time-expensive taking and analysis of a high number of images becomes applicable for toxicological and pharmacological approaches, and recent studies have already successfully implemented HCS for their experiments. For example, in 2019, Teixidó et al. established a HCS-method to investigate the effects of substances on the developmental outcome of zebrafish embryos, and in 2020, Dreser et al. developed a stem cell-based method to differentiate and screen so-called “neural rosettes”.

Introduction

Metabolomics

Inside of cells, many processes are continuously happening which fulfil important functions like homeostasis, energy production, etc. This metabolic activity involves, produces and consumes a variety of small molecules, referred to as metabolites, and the study of them is called metabolomics. If a substance affects cellular processes, it will probably also have an impact on the presence of involved metabolites, which can be measured with mass spectrometric means. In the last decade, a remarkable approach was developed, starting with the work of Cezar et al. 2007 and the discovery of metabolites that where distinctly deregulated in hESCs in response to VPA. Follow-up studies in 2010, 2011, 2013 and 2017 extended the knowledge about metabolites perturbated by teratogens in PSCs, leading to a simple but highly predictive model for DT-testing (West et al. 2010; Kleinstreuer et al. 2011; Palmer et al. 2013; 2017). Simultaneously, the company Stemina Biomarker Discovery was founded in 2006/07, selling the only commercially available PSC-based DT-testing service on the market (Stemina Biomarker Discovery 2020). Most recently, the devTOX quickPredict platform was used to predict the developmental toxicity of 1065 ToxCast phase I and II chemicals, demonstrating the powerfulness of this method (Zurlinden et al. 2020).

Transcriptomics

Analysis of the transcriptome, i.e. the sum of all RNA transcripts of a cell, provides, similar to metabolomics, a fingerprint-like information about the on-going cellular processes at the measured time point (Futschik et al. 2018). Among all transcripts, the messenger RNAs (mRNA) are of special interest. They are transcribed from well-documented genes and serve as blueprints for proteins, which in turn can affect the activity of other genes or cellular processes, so that mRNA molecules are also linked to genomics, proteomics and metabolomics. Nowadays, the whole human mRNA transcriptome is known and can be analyzed with different techniques like gene chips or RNA-sequencing, and even the analysis of single cell-transcriptomics is possible. In the context of DT-testing, it was already shown that the transcriptional activity of differentiating hESCs is disturbed by substances like thalidomide (Meganathan et al. 2012) or VPA (Krug et al. 2013). Especially the test system used in the latter study, the so named UKN1 system, considerably progressed in the last years. It was developed by Chambers et al. 2009 and differentiated hESCs to neuroepithelial progenitor (NEP) cells by inhibiting SMAD signaling twice with the two complementary acting factors SB431542 and noggin (Figure **1**). Balmer et al. 2012 modified the protocol, e.g. by adding the third SMAD inhibiting molecule dorsomorphin, and found distinct gene expression patterns of NEPs in control conditions as well as reproducible pattern-alterations after VPA exposure. In 2013 then, Krug et al. utilized the modified protocol to investigate the impact of the teratogens VPA and methylmercury (MeHg) on genome-wide gene expression with transcriptomic means and named it UKN1. The establishment of UKN1 as a well-known protocol progressed when Rempel et al. 2015 investigated the transcriptomic changes caused by histone deacetylase inhibitors (HDACi) and mercury-derived compounds ("mercurials"). They found that both groups affect neuroepithelial differentiation and induce distinct gene expression alterations, facilitating a classifier-based, group-specific separation of the compounds. Utilizing UKN1 and a second system, Shinde et al. 2017 introduced two new indices, called "developmental potency" (D_p) and "developmental index" (D_i) to quantitatively determine teratogenicity based on gene expression alterations.

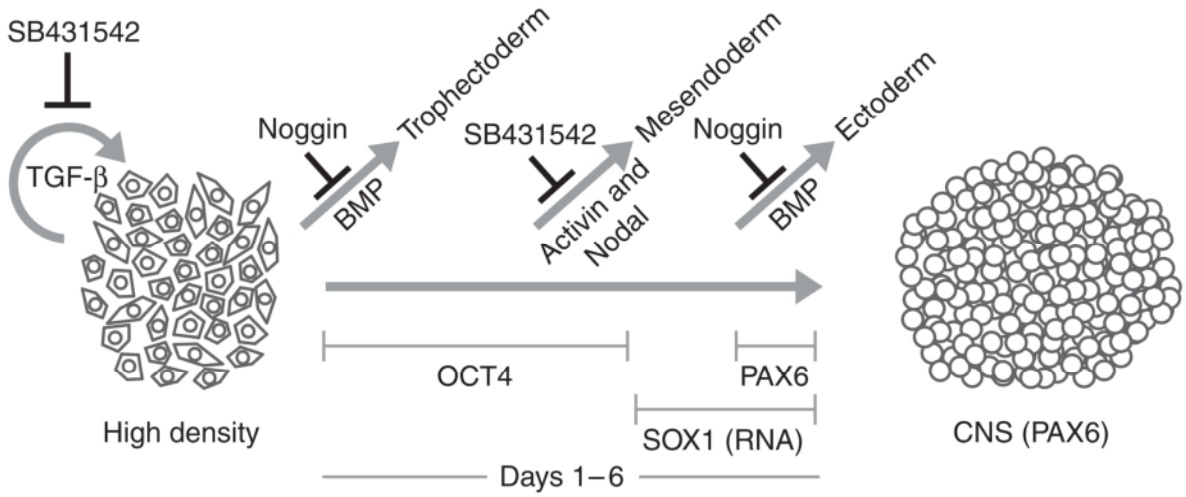


Figure 1: *Proposed mechanisms happening during the differentiation of pluripotent stem cells towards NEPs (here referred to as CNS). The small molecule SB431542 inhibits the Transforming Growth Factor-beta (TGF-β)-mediated self-renewal networks of PSCs and forces them to enter the differentiation process. Furthermore, it inhibits Activin and Nodal and blocks the mesendodermal lineage. Noggin supports the differentiation to NEPs by suppressing Bone Morphogenic Protein (BMP)-signaling, preventing differentiation into (tropho-)ectodermal cells. Source: Chambers et al. 2009*

1.3.3 3rd feature: Test system evaluation

According to Paracelsus (1493-1541), the guiding principle in toxicology, i.e. that the toxicity of a substance depends on its dose, is of utmost importance and plays a central role in the risk assessment of every substance. Therefore, in vitro methods in DT must be able to discriminate between three concentration ranges where the applied substance is a) non-cytotoxic and non-teratogenic, b) non-cytotoxic but teratogenic and c) cytotoxic. Non-cytotoxic and cytotoxic conditions are relatively easy to determine, but the identification of teratogenicity in between is difficult. Nevertheless, Waldmann et al. 2014 and Palmer et al. 2013 showed that a deregulation in the transcriptomic and metabolomics markers, respectively was detectable prior to the onset of cytotoxicity. Once the concentration ranges have been determined by a test system, its performance has to be evaluated. A common way is the testing of known teratogens and non-teratogens like thalidomide or VPA and dietary supplements, respectively, and comparing all results, in particular the obtained concentrations, to the true in vivo situation. The better the performance metrics, which are typically stated as accuracy, sensitivity and selectivity, the better the test system (Baratloo et al. 2015). Conversely, the precision of the test also depends on the quality of the in vivo data. If a test aims at the detection of DT for humans, then high-quality data originate from clinical human studies, where healthy patients received substances, mostly drugs, at therapeutic doses and the subsequent pharmacokinetics were measured, including blood or plasma concentrations. Even better are data from neonates at childbirth, whose mothers took substances right before delivery, but, understandably, very few of such data are existing and restricted to some indispensable substances, e.g. drugs for treatment of distinct diseases or dietary supplements. Unfortunately, the use of non-human-derived in vivo data for the evaluation of the test system is connected to the major problem of interspecies-differences, meaning that test conditions and results observed in animals cannot be transferred one-to-one to the human situation (Nau 1986). An interesting approach to overcome this issue and make best use of the existing and costly obtained data was recently presented by Zurlinden et al. 2020. In their study, they found 432 animal data sets from a total of 1065 tested chemicals to provide sufficient DT data, classified these into different groups in respect of their evidence for teratogenicity and finally used this classification for the test system evaluation. Another very promising approach to transfer results from one situation to another is

physiologically-based pharmacokinetic (PBPK) modeling (Zhuang and Lu 2016). On the basis of existing substance-specific pharmacokinetic data, the pharmacokinetics of that substance for the new situation are predicted. So, results could be transferred from male to female, young to old, healthy to diseased, mouse to man or even in vitro to vivo.

1.4 Aim of this work: Development of a novel in vitro test system for teratogenicity

A broad diversity of protocols for the identification of teratogens in vitro exists as outlined in the previous chapters, yet only a few are validated and accepted for regulatory risk assessments by the authorities so far. Therefore, the major aim of this work was to develop an innovative assay that can reliably detect teratogens in vitro and that maybe will be recognized by authorities as a sound test system for regulatory risk assessment in the future. For that purpose, the new system which based on the UKN1 protocol combined hiPSCs, transcriptomics and a well-chosen set of substances with high-quality in vivo data. Thus, the presented work will show, how the novel system further advanced the UKN1 system by (i) replacing hESCs with hiPSCs, (ii) screening a set of 39 compounds, which covered different molecular mechanisms and included well-known teratogens as well as proven non-teratogens and (iii) anchoring the concentrations of the applied substances to human in vivo relevant therapeutic concentrations. The ability of each compound to affect neuroepithelial differentiation was assessed by substance-induced cytotoxicity and genome-wide gene expression analyses. The in vitro outcome was then compared to the in vivo situation to evaluate the performance of the new “UKN1 6-day test system”.

In the second part of this thesis, the UKN1 6d assay was used to analyze the developmental toxicity potential of parabens in vitro, especially of ethylparaben. The aim was to facilitate a read-across approach for the DT risk assessment of ethylparaben in the context of a highly topical discussion about the DT-effects of parabens, so that in vivo experiments would not be necessary and harm to hundreds or thousands of animals could be prevented. For this purpose, methyl-, ethyl-, propyl- and butylparaben were tested over a broad concentration range that covered non-cytotoxic, cytotoxic and therapeutic in vivo relevant concentrations. Outcomes were determined by viability assays, quantitative real-time PCR and genome-wide gene expression analysis.

The aim of the third and last part of this work was to optimize the UKN1 6-day protocol. Therefore, it was explored if a reduction of the differentiation and substance-exposure time in combination with the gene expression analysis of only four biomarker-genes instead of the whole transcriptome yielded similar results as the UKN1 6-day test. This assay was called “UKN1 one-day test”.

2. Material & methods

2.1 Material

2.1.1 Technical equipment

Table 3: Technical equipment in the laboratory

Device	Company
Balance PB602	Mettler Toledo
Bright Field Microscope Axio Vert. A1 FL-LED + Camera Axiocam 305 + Software ZEN	Zeiss
Bright Field Microscope Eclipse TS100 + Camera Digital Sight DS-2MBW + Software NIS Elements	Nikon
Centrifuge 3-30K	Sigma
Centrifuge 5415 D	Eppendorf
Centrifuge Megastar 1.6R	VWR
Centrifuge MiniSpin plus	Eppendorf
CO2 Incubator C170 (E3)	Binder
Freezing Container Mr. Frosty	Nalgene
Hemocytometer Neubauer improved	Marienfeld
Laminar Flow Hood Lamin Air HBB2472	Heraeus
Magnetic stirrer IKAMAG RCT	IKA
pH meter Lab 850	SI Analytics
Pipetteboy Pipetus®	Hirschmann Laborgeräte
Pipettes Research, Reference, Xplorer and Xplorer Plus	Eppendorf
Plate reader Infinite M200 Pro	Tecan
Precision Balance ALJ	Kern
Real Time PCR System ABI 7500	Applied Biosystems
Shaker HS501	IKA
Sonicator Laboson 200	Bender & Hobein
Sonicator SONOPULS GM Mini20	Bandelin
Spectrometer NanoDrop 2000	Thermo Fisher Scientific
Thermo cycler Tgradient	Biometra
Thermomixer 5436	Eppendorf
Vacuum pump N022 AN.18	Neuberger
Vortex	VWR
Water purification system Milli-Q	Merck
Waterbath	Memmert

2.1.2 Consumables

Table 4: Plastic ware consumed in experiments

Product	Cat.	Company
Biosphere® Filtered Tip 100 µl	70760212	Sarstedt
Biosphere® Filtered Tip 1000 µl	703050255	Sarstedt
Biosphere® Filtered Tip 2.5 µl	701130212	Sarstedt
Biosphere® Filtered Tip 20 µl	701116210	Sarstedt
Biosphere® Filtered Tip 200 µl	70760211	Sarstedt
Canula Gr. 1, 20 G x 1 1/2"	4657519	BBraun
Cell scraper 25 cm	831830	Sarstedt
Cellstar® 96-Well cell culture microplate black	655079	Greiner Bio-One
CryoPure microtube 1.0 ml	72377992	Sarstedt
epT.I.P.S.® Standard, 5 ml	0030000650	Eppendorf
MicroAmp® Optical 96-Well Reaction Plate	N8010560	Thermo Fisher Scientific
MicroAmp® Optical Adhesive Film	4311971	Thermo Fisher Scientific
Parafilm®	H951.1	Carl Roth
Pipette tip 10 µl	701130	Sarstedt
Pipette tip 1000 µl	703050020	Sarstedt
Pipette tip 20 µl	701116	Sarstedt
Pipette tip 200 µl	70760012	Sarstedt
RNase-free Microfuge Tubes 1.5 mL	AM12400	Ambion
Rnase-free PCR microtube 0.2 ml	72991002	Sarstedt
SafeSeal microtube 0.5 ml	72704	Sarstedt
SafeSeal microtube 1.5 ml	72706	Sarstedt
SafeSeal microtube 2.0 ml	72695500	Sarstedt
SafeSeal microtube 5.0 ml	72701	Sarstedt
Serological Pipette 10 ml	861254001	Sarstedt
Serological Pipette 25 ml	861685001	Sarstedt
Serological Pipette 5 ml	861253001	Sarstedt
Serological Pipette 50 ml	861256001	Sarstedt
Syringe 10 ml	4616103V	BBraun
Syringe 3 ml	4617022V	BBraun
Syringe 50 ml	4616502F	BBraun
Syringe filter 0.2 µm	831826001	Sarstedt
Tissue culture flask 175 cm²	833912002	Sarstedt
Tissue culture flask 25 cm²	833910002	Sarstedt
Tissue culture flask 75 cm²	833911002	Sarstedt
Tissue culture plate, 12-Well	833921	Sarstedt
Tissue culture plate, 6-Well	833920	Sarstedt

Product	Cat.	Company
Tissue culture plate, 96-Well	833924	Sarstedt
Tubes 15 ml	62554502	Sarstedt
Tubes 50 ml	62547254	Sarstedt
Vacuum Filtration Unit 250 ml, 0,2 µm	833940001	Sarstedt

2.1.3 Chemicals, substances and solutions

Table 5: Chemicals, substances and solutions used in experiments

Product	Cat.	Company
2-Propanol	6752.5	Carl Roth
3,3',5-Triiodo-L-thyronine sodium salt	T6397	Sigma Aldrich
5,5-Diphenylhydantoin sodium salt	sc-214337	Santa Cruz
9-cis-Retinoic acid	14587	Cayman
Acitretin	PHR1523	Merck
Actinomycin D	BVT-0089	Adipogen Life Sciences
Ampicillin anhydrous	A9393	Sigma Aldrich
Ascorbic acid	A0278	Sigma Aldrich
Atorvastatin Calcium	PHR1422	Sigma Aldrich
Buspirone hydrochloride	B7148	Sigma Aldrich
Carbamazepine	C4024	Merck
Cefotiam dihydrochloride	1098005	Sigma Aldrich
Chlorpheniramine maleate salt	C3025	Sigma Aldrich
Clomiphene citrate salt	C6272	Sigma Aldrich
Dextromethorphan HBr	PHR1018	Sigma Aldrich
Diethylpyrocarbonate treated (DEPC) water	750024	Thermo Fisher Scientific
Dimethyl sulfoxide (DMSO)	A994.1	Carl Roth
Diphenhydramine	sc-204729	Santa Cruz Biotechnoolgy
DNAZap® DNA and RNA Decontamination Solution	AM9890	Thermo Fisher Scientific
Doxorubicin hydrochloride	D2975000	Sigma Aldrich
Doxylamine succinate	D3775	Sigma Aldrich
DPBS, calcium, magnesium	14040091	ThermoFisher Scientific
DPBS, no calcium, no magnesium	14190094	ThermoFisher Scientific
Entinostat (MS-275)	Cay13284	Cayman Chemicals
Ethanol, 70 % denatured	2202	ChemSolute
Ethanol, absolute	100983	Merck
Famotidine	F6889	Sigma Aldrich
Favipiravir	HY14768	Hycultec
Finasteride	14938	Cayman Chemical
Folic acid	F7876	Sigma Aldrich
Isotretinoin	PHR1188	Sigma Aldrich

Product	Cat.	Company
Leflunomide	PHR1378	Merck
Levothyroxine	PHR1613	Sigma Aldrich
Lithium chloride	L4408	Sigma Aldrich
Magnesium chloride anhydrous	8147330500	Merck
Medroxyprogesterone	24908	Cayman Chemical
Methicillin sodium salt monohydrate	1410002	Sigma Aldrich
Methotrexate	PHR1396	Sigma Aldrich
Methylmercury	33368	Sigma Aldrich
Misoprostol	13820	Cayman Chemical
Panobinostat	Cay13280	Cayman Chemicals
Paraformaldehyde 4% (PFA)	P087.5	Carl Roth
Paroxetine hydrochloride	PHR1804	Sigma Aldrich
Ranitidine hydrochloride	R101-1G	Sigma Aldrich
Retinol	17772	Merck
RNaseZap® RNase Decontamination Solution	AM9780/AM9782	Thermo Fisher Scientific
Sucralose	PHR1342	Sigma Aldrich
Teriflunomide (A-771726)	HY15405	Hycultec
Thalidomide	T144	Sigma Aldrich
Trichostatin A	T1952	Sigma Aldrich
TritonX-100	T8787	Sigma-Aldrich
Tween20	P2287	Sigma-Aldrich
Valproic acid	PHR1061	Sigma-Aldrich
Vinblastine sulfate salt	V1377	Sigma Aldrich
Vismodegib	HY10440	Hycultec
Vorinostat (SAHA)	Cay10009929	Cayman Chemicals

2.1.4 Media, supplements & cell culture reagents (commercially purchasable)

Table 6: Media, supplements and reagents specifically needed for cell culturing

Product	Cat.	Company
2-mercaptoethanol	31350010	ThermoFisher Scientific
Albumin from human serum, fatty acid free, globulin free	A3782	Sigma Aldrich
apo-Transferrin human	T2036	Sigma Aldrich
Cellartis® DEF-CS™ 500 Culture System	Y30010	Takara Bio
D(+) Glucosemonohydrate	49159	Sigma Aldrich
DMEM/F12, no glutamine	21331-020	ThermoFisher Scientific
Dorsomorphin Dihydrochloride	SM03	Cell guidance systems
Essential 6™ Medium	A1516401	ThermoFisher Scientific
Essential 8™ Medium	A1517001	ThermoFisher Scientific
GlutaMax Supplement (100x)	35050038	ThermoFisher Scientific

Product	Cat.	Company
Insulin solution human (ca. 10 mg/ml)	I9278	Sigma Aldrich
KnockOut DMEM	10829-018	ThermoFisher Scientific
KnockOut Serum Replacement	10828-010	ThermoFisher Scientific
Laminin-521, human recombinant	LN521-05	BioLamina
MEM non-essential amino acids solution (100x)	11140035	ThermoFisher Scientific
Noggin, human recombinant	6057-NG	R&D Systems
Progesterone	P7556	Sigma Aldrich
Putrescine dihydrochloride	P5780	Sigma Aldrich
SB431542	SM33	Cell guidance systems
Sodium selenite	S5261	Sigma Aldrich
Stem-Cellbanker GMP-Grade	11890	amsbio
Trypan blue stain (0.4 %)	15250061	Thermo Fisher Scientifc
TrypLE™ Select	12563011	ThermoFisher Scientific
Vitronectin, truncated, human recombinant	A31804	ThermoFisher Scientific
Y-27632	SM02	Cell guidance systems

2.1.5 Self-made cell culture media

Table 7: Composition of KSR-S and N2-S

KSR-S	500 ml
KnockOut DMEM	415 ml
KnockOut Serum Replacement	75 ml
GlutaMax Supplement (100x)	5 ml
MEM non-essential amino acids solution (100x)	5 ml
2-Mercaptoethanol (50 mM)	500 µl
N2-S (filtered with 0.2 µm filter)	**250 ml**
DMEM/F12, no glutamine	247 ml
Apo-transferrin human	25 mg
D(+) Glucosemonohydrate	426.5 mg
GlutaMax Supplement (100x)	2.5 ml
Insulin solution human (ca. 10 mg/ml)	625 µl
Putrescine dihydrochloride (1 M)	25 µl
Sodium Selenite (500 µM)	15 µl
Progesterone (100 µM)	50 µl

2.1.6 Kits, assays and arrays

Table 8: Proprietary kits, assays and arrays for diverse purposes

Product	Cat.	Company
CellTiter-Blue Cell Viability Assay	G8081	Promega
High-Capacity cDNA Reverse Transcription Kit	4368813	Thermo Fisher Scientific
ExtractMe Total RNA-Kit	EM09.1	Blirt
QuantiFast SYBR Green PCR Kit	204057	Qiagen
GeneChip® Human Genome U133 Plus 2.0 array	900466	ThermoFisher Scientific
GeneChip® 3' IVT PLUS Reagent Kit	902416	ThermoFisher Scientific

2.1.7 Quantitative real-time PCR primers

Table 9: Primers (manufactured by Eurofins Scientific) with known sequence used in qPCR experiments

Target Gene (human)	Type	Sequence (5'--> 3')
EMX2 (Empty spiracles homeobox 2)	Forward	CCAAGGGAACGACACTAGCC
EMX2 (Empty spiracles homeobox 2)	Reverse	CCATACTTTTACCTGAGTTTCCGTG
MSX1 (Msh homeobox 1)	Forward	GGATCAGACTTCGGAGAGTGAACT
MSX1 (Msh homeobox 1)	Reverse	GCCTTCCCTTTAACCCTCACA
NANOG (Nanog homeobox)	Forward	GGTGAAGACCTGGTTCCAGAAC
NANOG (Nanog homeobox)	Reverse	CATCCCTGGTGGTAGGAAGAGTAAAG
OCT4A (POU class 5 homeobox 1 alpha)	Forward	GCAAAGCAGAAACCCTCGTGC
OCT4A (POU class 5 homeobox 1 alpha)	Reverse	ACACTCGGACCACATCCTTCTCG
OTX2 (Orthodenticle homeobox 2)	Forward	GTTCAGAGTCCTTGGTGGGT
OTX2 (Orthodenticle homeobox 2)	Reverse	CCCTCACTCGCCACATCTAC
PAX6 (Paired box 6)	Forward	CCGCCTATGCCCAGCTTCAC
PAX6 (Paired box 6)	Reverse	AAGTGGTGCCCGAGGTGCCC
TBP (TATA-binding protein)	Forward	GGGCACCACTCCACTGTATC
TBP (TATA-binding protein)	Reverse	GCAGCAAACCGCTTGGGATTATATTCG
TFAP2B (Transcription factor activating enhancer binding protein 2 beta)	Forward	GGGGAGATCTTTGCGAGAAAGG
TFAP2B (Transcription factor activating enhancer binding protein 2 beta)	Reverse	CTGTGTGCTGCCGGTTCAAATA

Table 10: Proprietary Quantitect primer assays (Qiagen) used in qPCR experiments

Target Gene (human)	Cat.
FOXA2 (Forkhead box A2)	QT00212786
MEIS2 (Meis homeobox 2)	QT00077315
OTX2 (Orthodenticle homeobox 2)	QT00213129

2.1.8 Antibodies & Staining reagents

Primary antibodies

Table 11: Primary antibodies for use in immunostainings

Target antigen	Host	Cat#/Company
NANOG	Rabbit	StemLight Pluripotency Kit #9656, Cell Signaling Technology
OCT4A	Rabbit	
SOX2 (SRY (sex determining region Y)-box transcription factor 2)	Rabbit	
SSEA4 (Stage-specific embryonic antigen-4)	Mouse	
TRA-1-60 (T cell receptor alpha locus 1-60)	Mouse	
TRA-1-81 (T cell receptor alpha locus 1-81)	Mouse	

Secondary antibodies

Table 12: Secondary antibodies applied in immunostainings

Target antigen	Host	Dye	Cat#	Company
Rabbit	Donkey	AlexaFluor 488	#711545152	Jackson ImmunoResearch
Mouse	Donkey	AlexaFluor 488	#715546150	Jackson ImmunoResearch

Other reagents

Table 13: DAPI and serum needed for immunostainings

Product	Cat#	Company
4',6-diamidino-2-phenylindole (DAPI)	#D1306	Life Technologies
Normal Donkey Serum	#017000121	Jackson ImmunoResearch

2.1.9 Cell line

For this project, SBAD2 human induced pluripotent stem cells (hiPSCs) were obtained from the Chair of in-vitro Toxicology and Biomedicine Dept inaugurated by the Doerenkamp-Zbinden foundation at the University of Constanz. Originally, SBAD2 hiPSCs had been supplied by the StemBANCC project (http://stembancc.org) (Morrison et al. 2015) and had been created by reprogramming dermal fibroblasts from a 51-year old male Caucasian (Lonza, Switzerland) with the non-integrating Sendai virus CytoTune™-iPS reprogramming kit (ThermoFisher Scientific, USA) (Fusaki et al. 2009; Beers et al. 2015; Schlaeger et al. 2015). The identity of the obtained SBAD2 hiPSCs was confirmed by STR (short tandem repeat) profiling performed at the Leibniz-Institute DSMZ (German Collection of Microorganisms and Cell Cultures).

2.2 Methods

2.2.1 hiPSC cultivation

SBAD2 cells were cultured under sterile conditions, i.e. all working steps were performed aseptically under a laminar flow hood and only aseptical materials and solutions were used. SBAD2 cells were incubated at 37 °C, the atmosphere was humidified and enriched with 5 % CO_2. Moreover, every medium and solution was prewarmed to 37 °C before it got in contact with the cells to prevent a temperature-related shock and subsequent death of cells.
To grow the cells and maintain their pluripotent state, two different culture systems were used (Table 14). In the first one, cells were cultured in Essential 8 medium (E8) in laminin 521-coated tissue culture flasks. Essential 8 medium consisted of Essential 8 basal medium and 1x Essential 8 supplement. During passaging, freezing and thawing of cells, medium was additionally spiked with 10 µM Rho-kinase inhibitor Y-27632 (Rocki).
The second system, the proprietary Cellartis® DEF-CS culture system, included basal medium, growth factors 1-3 (GF1, GF2, GF3) and the coating stock-solution Coat-1. Cells were cultured in basal medium spiked with 1:333 GF-1 and 1:1000 GF-2 in Coat-1-coated tissue culture flasks. During passaging, freezing and thawing of cells, medium was additionally spiked with 1:1000 GF-3.
Every day, the cells were feeded with freshly prepared medium. To save the costly medium, the amount of given medium was adjusted to the confluency of the cells, so that low-confluent cells received 0.2 ml/cm² and high-confluent cells up to 0.5 ml/cm² medium. Furthermore, cells were checked daily via light-microscopy for areas of spontaneous differentiation. If too many of those areas had been formed in a flask, that flask was discarded.

Table 14: Composition of media used in the two culture systems

System	Basal medium	Supplements	Additional supplement for passaging, freezing and thawing cells
E8	E8 basal medium	1x E8 supplement (E8-S)	10 µM Rocki
DEF-CS	DEF-CS basal medium,	1:333 GF-1, 1:1000 GF-2	1:1000 GF-3

2.2.2 Passaging cells

Following a three- or four-day feeding protocol, i.e. cells were confluent three or four days after the last passage, cells were passaged twice a week. First, tissue culture flasks were coated with Laminin-521 or the proprietary Coat-1-coating. For that purpose, both stock solutions were diluted 1:20 in DPBS (+/+), which resulted in a Laminin-521 protein concentration of 5 µg/ml and an unknown protein concentration of Coat-1. Flasks were covered with 100 µl/cm² coating solution and incubated for >30 min at 37 °C. Then, old medium was aspirated from the cells, they were washed twice with DPBS (-/-) and covered with 20 µl/cm² TrypLE Select. After 5 min incubation at 37 °C and 5 % CO2, cells were detached by bumping firmly several times against the closed tissue culture flask. The growing surface was rinsed twice with freshly prepared passaging medium (12 times the volume of TrypLE) to detach strong adhesive cells. The cell suspension was aliquoted in volumes of 8 ml in 15 ml Falcon tubes. In order to separate remaining cell clumps into single cells, each aliquot of the cell suspension was taken up with a 10 ml serological pipette. Then, the opening of the serological pipette was pressed against the inner flask wall opposing to the flask's opening and the cell suspension was blown out under pressure. This was done twice with every aliquot. Afterwards, aliquots were centrifuged for 5 min at

300 rcf. The supernatant was aspirated, each pellet was resuspended in 1 ml passaging medium and the aliquots were merged. Cells were counted with a Neubauer hemocytometer in the presence of trypan blue to determine their viability. Finally, cells were seeded in Coat-1-coated flasks in passaging medium at densities of usually 12000 cells/cm² and 20000 cell/cm² for a four-day and three-day feeding protocol, respectively.

2.2.3 Freezing cells

SBAD2 cells were cultivated until confluency, a sample of the spend media was taken for a mycoplasma test and the cells were processed as described in "Passaging cells". After cell counting, cells were centrifuged for 5 min at 300 rcf. The supernatant was aspirated and the cell pellet was resuspended in ice-cold stem cellbanker. The cells were aliquoted á 1.5 million cells per 1 ml cryo-tube and stored for at least 24 h at -80 °C in a Mr. Frosty freezing container, which was filled with 2-propanol to provide a linear and gentle cooling-rate of -1 °C / min. After a negative mycoplasm test, the cells could be transferred into a liquid nitrogen tank for long-term storage.

2.2.4 Thawing cells

A single 1 ml vial of 1.5 million SBAD2 cells was taken from a liquid nitrogen storage tank and thawed quickly within 2 min in a 37 °C water bath. The cell suspension was transferred to a 15 ml Falcon tube and 10 ml passaging medium were added. To prevent an osmotic shock of the cells, the last was done slowly and dropwise. Afterwards, cells were centrifuged 5 min at 300 rcf. The supernatant was then aspirated and the pellet was resuspended in 5 ml passaging medium. The cells were seeded in a coated T25 tissue culture flask.

2.2.5 Immunostaining

The SBAD2 hiPSCs' pluripotency state was checked by looking immunohistochemically at the expression of the nuclear proteins OCT4A, NANOG and SOX2 and the surface proteins SSEA4, TRA-1-60 and TRA-1-81. For that purpose, SBAD2 cells were seeded in 96-well-plates and grown to ca. 80 % confluency. An incubation in 4% PFA for 15 min preserved the cells for the following staining procedures by crosslinking cellular proteins (Tayri-Wilk et al. 2020). The PFA was removed and the cells were rinsed three times with DPBS. Then, cells were incubated in DPBS containing 5 % donkey serum (blocking solution) for 60 min to block unspecific binding-sites for primary and secondary antibodies (Ramos-Vara 2005). If cells were planned to be stained for the nuclear proteins OCT4A, NANOG and SOX2, the blocking solution was spiked with 0.3 % TritonX-100 which permeabilized the cells so that the antibodies could access the intracellular antigens. If cells were planned to be stained for the surface proteins SSEA4, TRA-1-60 and TRA-1-81, TritoX-100 was not added to prevent a denaturation of the surface proteins. After blocking, the blocking solution was aspirated and the cells were incubated overnight at 4 °C with primary antibodies for the surface and the nuclear proteins diluted 1:200 in DPBS or DPBS/TritonX-100, respectively. On the next day, cells were washed three times in DPBS for 5 min each before adding the secondary antibodies. For the surface proteins, anti-mouse conjugated Alexa Fluor 488 antibody was diluted 1:500 in DPBS. For the nuclear proteins, anti-rabbit conjugated Alexa Fluor 488 antibody was diluted 1:500 in DPBS/TritonX-100. Cells were incubated with secondary antibodies for 1-2 h in the dark to avoid photobleaching of the fluorophores, as well as all following steps were performed under the exclusion of light. Next, cells were washed three times in DPBS for 5 min each and incubated with DAPI diluted 1:10000 in DPBS for 10 min. Cells were washed again three times in DPBS for 5 min each. Finally, cells were covered with DPBS, the plates were wrapped with Parafilm and stored in the dark at 4°C.

2.2.6 Fluorescence imaging

For assessment of the SBAD2 hiPSCs' pluripotency state, relevant biomarkers were labeled as described in the preceding chapter. Imaging of the dyes Alexa Fluor 488 and DAPI was done at the excitation (Ex) and emission (Em) wavelengths 490Ex/525Em nm and 350Ex/470Em nm, respectively, with a Zeiss Axio Vert.A1 fluorescence microscope. For the histological controls, cells had been stained without primary antibodies.

2.2.7 Differentiation of hiPSCs

The differentiation of SBAD2 hiPSCs to neuroepithelial precursor cells was performed via a strict protocol in two versions A and B. In version A, cells had been cultured in E8 previous to the application in the differentiation protocol. On the first day of this protocol, these cells were passaged as described in the chapter "2.2.2 Passaging cells" and seeded in 1 ml E8 per well on vitronectin-coated 12-well-plates at a density of 24,000 cells/cm². This day was called "Day of Differentiation (DoD) -3". On DoD -2 and DoD -1, E8 was refreshed. On DoD 0, 1 and 2, medium was changed to the differentiation medium Essential 6 (E6) which was spiked with 21.6 µM SB431542, 0.64 µM dorsomorphin, 35 ng/ml noggin and 0.1 % DMSO* to induce neural differentiation. Additionally, from DoD 0 onwards cells were exposed to various compounds at different concentrations for a total of 96 h. On DoD 4, medium was changed to a mixed differentiation medium containing 75 % E6, 25 % N2-S (**Table 7**) and previous supplements. On DoD 6, cells were lysed with Rlys buffer from the Blirt® RNA extraction kit and harvested for gene expression analysis. In order to enhance the disruption of the cells, they were sonicated (Table **15**) before they were stored at -80 °C.
In version B of the protocol, Essential 8 was replaced by DEF-CS and E6 was replaced by KSR-S (**Table 7**). Furthermore, on DoD -3 cells were seeded at a density of 12,000 cells/cm² on laminin-521-coated 12-well-plates.
In a shortened one-day-differentiation protocol based on version B, cells were exposed to substances for 24 h from DoD 0 to DoD 1 only. On that very same DoD 1, cells were lysed and harvested for gene expression analysis, sonicated and stored at -80 °C.

*The compounds leflunomide and teriflunomide were tested at a DMSO concentration of 0.5 %.

Table 15: Sonication parameters (Sonopuls GM mini20)

Parameter	Setting
Amplitude	50 %
Time	30 s (per sample)
Pulse	5 s
Pause (between two pulses)	2 s

2.2.8 Preparation of compound stock-solutions

Upon arrival, the compounds, which were intended to be exposed to the differentiating hiPSCs, were stored at appropriate conditions. For an easier and more efficient handling in the media preparation step (see the following chapter), compounds were solved and prepared as stock solutions in high concentrations in 100% DMSO, if soluble. Alternatively, substances were solved in high concentrations in water. Stock solutions were stored at -20 °C.

2.2.9 Preparation of compound-spiked differentiation media

hiPSCs were exposed to various substances at different concentrations during the differentiation process. Usually, substance-spiked differentiation media E6 or KSR-S were prepared on DoD -1, i.e. the day before differentiation was induced, by diluting the solved substances from the stock-solutions into the highest desired concentrations. When stock-solutions were not available for specific compounds, the solid powder form of the compounds was taken, weighed, solved into the differentiation media and filtrated at the highest desired concentration. Media were vortexed immediately to avoid delayed precipitation of specific compounds. If a substance was not completely soluble at the first attempt, the solubilization was supported by heating and sonication of the medium in a sonication water bath. Lower concentrated media were prepared in a dilution row manner unless a compound was not soluble at the first attempt at higher concentrations. Then, stock-solutions were diluted and used for lower concentrated media, or less powder was weighed, solved and filtrated. All media were stored at 4 °C.

2.2.10 Viability test

The CellTiter-Blue® cell viability assay (Promega), which is basically a resazurin-solution, was used to measure the cell viability of cells which were exposed to parabens. This measurement took place on DoD 6 of the differentiation process, immediately before the cell harvest for the RNA extraction. As resazurin and resorufin are light-sensitive, all steps were performed at low-light conditions. In a first step, a medium-resazurin-solution had to be prepared by mixing 40 % Essential 6 medium, 40 % N2-S and 20 % resazurin-solution, spiked with 21.6 µM SB431542, 0.64 µM dorsomorphin and 35 ng/ml noggin. Cells were washed three times with DPBS (-/-) before adding 1 ml/well medium-resazurin-solution. Cell-plates were incubated at 37 °C and 5 % CO_2 until a color-shift from blue to purple of the medium-resazurin-solution could be observed, which usually happened after 20 min. The resorufin-containing supernatant was then transferred to a black 96-well-plate and its fluorescent signal was measured in triplicates at $579_{Ex}/584_{Em}$ nm in an Infinite M200 Pro plate reader (Tecan). Finally, the raw measurements were background-corrected by subtracting the signal that was emitted from unmetabolized medium-resazurin-solution which had not been in contact with cells.

2.2.11 Drawing of cytotoxicity curves, curve fitting and calculation of EC50 values

The data obtained from the viability test were used to draw cytotoxicity curves which in turn based the calculation of the EC50 values. The drawing and curve fitting was performed according to the method described in Albrecht et al. 2019 in the chapter “Curve fitting and calculation of EC values”. Deviating from the therein described method, no goodness-of-fit was calculated for the parabens’ cytotoxicity curves.

2.2.12 Isolation of RNA

To avoid degradation of the ribonucleic acid (RNA) caused by ribonucleases (RNases), all working steps were performed under RNase-free conditions, i.e. tables were cleaned with RNaseZap® and only certified RNase-free pipette filter tips and reaction tubes were used. This applied also for the following chapters.
Total RNA was isolated from the cell lysates with the “ExtractMe Total RNA-Kit” (blirt) which uses membrane-based mini spin columns that specifically bind RNA and separate it from other components of the cell lysate. To enable RNA-binding to the membrane, cell lysates were mixed 1:1 with 70 % clean ethanol. Up to 700 µl of that mix was given on a RNA-column, centrifuged for 15 s at 12000 rcf and the flow-through was discarded. If there was still mix left, these steps were repeated for the same mix on

the same column. Then, membranes were washed with 500 µl RW2 buffer, i.e. the buffer was given on the column, followed by centrifugation for 15 s at 12000 rcf and discard of the flow-through. As the membrane's RNA binding is based on affinity to nucleic acids, the membrane might also bind genomic DNA. To get rid of potentially bound genomic DNA, the membrane was treated with deoxyribonuclease I (DNase I) (9 µl DNase I + 81 µl DNase I reaction buffer) for 5-15 min. Afterwards, the membrane was washed once with 600 µl RW1 buffer and twice with 500 µl RW2 buffer. In order to dry the column, it was placed in a clean collection tube and centrifuged for 1 min at 12000 rcf. For the elution of the RNA, the column was placed in a 1.5 ml reaction tube, 50-100 µl diethylpyrocarbonate (DEPC) treated water was added and it was centrifuged for 1 min at 12000 rcf. The concentration and purity of the RNA were determined with a NanoDrop2000 (ThermoFisher Scientific) and RNA was stored at -80 °C.

2.2.13 Reverse transcription of RNA into cDNA

RNA was transcripted into complementary deoxyribonucleic acid (cDNA) with the "High-Capacity cDNA Reverse Transcription Kit" (ThermoFisher Scientific). For each RNA sample, up to 2 µg/10 µl was added to an RNase-free microtube and filled up to 10 µl with DEPC treated water. Another 4.2 µl DEPC treated water as well as 2 µl 10x RT-buffer, 2 µl 10x RT-Primer and 0.8 µl dNTPs were added and mixed well. Then, 1 µl reverse transcriptase was added and mixed gently. The microtubes were centrifuged shortly, placed in a thermo cycler (Biometra) and the RNA was transcribed into cDNA following the thermo cycler program given in **Table 16**. Afterwards, cDNA was diluted with DEPC treated water to 10 ng/µl and stored at -20 °C.

Table 16: Thermo cycler program for cDNA transcription

	Step 1: Incubation	**Step 2: Reverse transcription**	**Step 3: Inactivation**	**Step 4: Cooling**
Temperature	25 °C	37 °C	85 °C	4 °C
Time	10 min	120 min	5 min	∞ (Pause)

2.2.14 RT-qPCR Analysis

In a standard reverse-transcription quantitative polymerase chain reaction (RT-qPCR), RNA or cDNA (the RNA's complementary transcribed form) is amplified over multiple cycles and their quantity is measured by a fluorescent dye, for example SYBR Green I, which is fluorescent upon DNA-binding (Higuchi et al. 1993; Arya et al. 2005). The fluorescent signal corresponds to the amount of cDNA, so it increases like the cDNA with every cycle and can be illustrated on a graph, where it is given as a function of the cycle number. When the fluorescent signal reaches a defined intensity threshold, the corresponding cycle number is called C_T-value. To ensure a sound C_T-value, the threshold has to be defined within the exponential phase, where the amplification is stable.
To assess C_T-values of genes in the cDNA samples, an ABI 7500 real-time PCR system (Applied Biosystems) was used. For one reaction in a MicroAmp® optical 96-well reaction plate, 50 ng/5 µl cDNA was mixed with 2.5 µl RNase-free water, 12.5 µl Quantifast SYBR® Green PCR master mix and 2.5 µl of each a forward and a reverse primer. Plates were centrifuged for 1 min at 1000 rpm (Centrifuge 3-30K, Sigma) and placed in the ABI 7500. The amplification program, which was used to detect the gene expression levels, is given in Table 17. Succeeding to the 40 amplification cycles, a melt curve analysis was performed to assess the "purity" of the amplification, i.e. if the detected fluorescent signal for one gene was caused by the targeted gene or if contaminating genes contributed to the signal intensity.

For the quantitative comparison of C_T-values in order to determine gene expression alterations, for example of the same gene at two different time points in a stem cell differentiation process, the so-called $2^{-\Delta\Delta CT}$-method was utilized (Livak and Schmittgen 2001).. At first, C_T-values of every target gene had to be normalized to a reference or "house-keeping" gene, which was expressed in the same sample as the target gene. Hallmarks of house-keeping genes are their stable gene expression levels that are not affected by changing conditions like substance treatment or different time points in an experiment. The normalized ΔC_T-values were calculated by subtracting the house-keeping gene's C_T-value from each target gene's C_T-value ($\Delta C_{T(Target)} = C_{T(Target)} - C_{T(House\text{-}keeper)}$). By comparing ΔC_T-values, gene expression changes of a target gene at different time points or between different conditions (e.g. control and substance-exposed cells) can be assessed ($\Delta\Delta C_{T2} = \Delta C_{T2} - \Delta C_{T1}$). Under the assumption that the exponential cDNA amplification is performed with ca. 100 % efficiency for every amplicon, the exponentialization of a negative $\Delta\Delta C_T$-value to the basis of two results in a fold-change value (= $2^{-\Delta\Delta CT2}$) which answers the question how much the expression of the target gene had changed from the first time point to the second time point or how much a substance-treatment affected the gene expression.

Table 17: ABI 7500 amplification program

Step	Temperature	Time
Holding stage	95 °C	3 min
Cycling stage (40 cycles in total):		
Step 1	95 °C	5 s
Step 2	60 °C	30 s
Melt curve stage:		
Step 1	95 °C	15 s
Step 2	60 °C	20 s
Step 3	95 °C	15 s
Step 4	60 °C	15 s

2.2.15 Statistical analysis of RT-qPCR results

Data for differential gene expression in RT-qPCR measurements were analyzed on their statistical significance in one sample t-tests with the Origin Pro 2021 software (Origin Labs).
The p-value for significance in each t-test was set as $p<0.05$. The following two null hypotheses were investigated:

1) Mean-$\Delta\Delta C_T$-value = 0 (equals to a mean-fold-change = 1)
2) Mean-$\Delta\Delta C_T$-value <= |1| (equals to a mean-fold-change <= |2|)

2.2.16 Gene arrays

Genome wide expression levels were obtained with Affimetrix' "GeneChip® Human Genome U133 Plus 2.0" arrays (ThermoFisher Scientific). For that purpose, RNA samples were sent to the Gene Expression Affymetrix Facility of the Center for Molecular medicine in cologne, headed by Prof. A. Sachinidis. Sample processing was done with Affymetrix' "GeneChip® 3' IVT PLUS Reagent Kit" following the manufacturer's instructions. Briefly, RNA was transcribed to double-stranded cDNA first, which was subsequently used to synthesize and amplify single-stranded and biotinylated complementary RNA (cRNA) by in vitro transcription (IVT) using T7 RNA polymerase (van Gelder et al. 1990). Labeled cRNA was then purified, fragmented and hybridized to the array that was composed of 1,300,000 unique

oligonucleotides. After hybridization, the array was washed, stained and scanned. The scanner captured the fluorescent signals emitted by the stained cRNA-oligonucleotide hybrids, which corresponded to the expression levels of 54,675 probe sets, that code for approximately 39,000 human genes.

2.2.17 Statistical analysis of gene expression data, principal component analysis, volcano plots

Gene expression levels measured with Affymetrix gene arrays were analyzed by the statisticians Prof. J. Rahnenführer and M. Sc. F. Kappenberg, TU Dortmund. Using the frozen robust multiarray analysis (fRMA) method (McCall et al. 2010), the raw gene expression values were normalized. The principal component analysis based on the 100 probe sets with the highest variance across all samples or on all 54675 probe sets. The mathematical algorithms for the calculation of the PCA data used procedure described in James et al. 2015, 374 ff.. Differential gene expression was assessed using the limma package for the software R (Smyth 2005; Ritchie et al. 2015) and comparing/normalizing gene expression levels of compound-exposed NEPs to the corresponding untreated controls. Volcano-plots were drawn using the fold-changes and p-values from that "limma"-analyses on logarithmic axes. Moreover, the false discovery rate (FDR) method developed by Benjamini and Hochberg was carried out to adjust p-values for multiple testing (Benjamini and Hochberg 1995). "Significant" probe sets (SPS) with both an FDR-adjusted p-value of <0.05 and an absolute deregulation of more than two-fold were highlighted in red in the volcano plots.

2.2.18 Classifier „Cytotox 1000" and „Cytotox SPS"

The classifier "Cytotox 1000" was based on a "Leave-One-Out"-cross-validation (LOOCV)-approach and the genome-wide expression analyses of all samples. The gene expression data used for this approach were first normalized via fRMA, and in a second step these data were normalized with respect to the controls. Hence, each sample provided normalized, differential gene expression data for 54,675 probe sets. The computation using that data involved several steps which were conducted iteratively for each substance of interest:

1) Assignment of "training set" and "test set":

The training set consisted of all samples except those belonging to the substance of interest. The test set consisted analogously only of the samples of the substance of interest.

2) Calculation of probe set variances within the training set:

For each of the 54,675 probe sets, a cumulative "variance" value was calculated. The variance of one probe set was a sum of all normalized gene expression values of that probe set across the samples of the training set. The variance was unique for the training set in this iteration.

3) Test set-specific classifier:

The 1000 probe sets with the highest variance across the samples of the training set were determined and used to build the classifier for the test set.

4) Classification of the test set:

Among all probe sets of the test set samples, only such that were identical to those 1000 "top" probe sets were relevant for the classification. The gene expression values of these relevant probe sets calculated the predicted probability for positive class in a complex mathematical algorithm for each test set sample. Final probabilities were based on the mean-values of biological replicates.

After the iterative conduction of these steps, each substance had its own classifier which were merged into one plot. Additionally, cytotoxic conditions were plotted with a predicted probability for positive class of 1 (=100 %). The "Area Under Curve" (AUC)-value was computed by an algorithm described in Albrecht et al. 2019.

The classifier “Cytotox SPS” based on the number of significantly deregulated probe sets as described in the previous subchapter (see chapter 2.2.17). Additionally, cytotoxic conditions were plotted with the same number of SPS as the non-cytotoxic condition, which showed the highest number of SPS across all non-cytotoxic conditions. As for the classifier “Cytotox 1000”, values for AUC, accuracy, sensitivity and specificity were calculated, but in contrast to “Cytotox 1000”, the AUC-value was not cross-validated.

The box plots were each made up of a box, a thick line within the box and whiskers on both sides of the box. The thick line represented the median of all data points. The ends of the box were given by the 0.25 and the 0.75 quartile of the data and the difference between the 0.75 and 0.25 quartile (the length of the box) was called the interquartile range (IQR). The antennas reached to the smallest value (lower antenna) or to the largest value (upper antenna) across all data points, but not further than 1.5*IQR from the corresponding end of the box.

2.2.19 Threshold, accuracy, sensitivity and specificity

Thresholds for the separation and definition of negative and positive in vitro test results were set in a way so that the values for accuracy, sensitivity and specificity were maximized. Accuracy, sensitivity and specificity were based on the determination of true negatives (TN), true positives (TP), false positives (FP) and false negatives (FN), which were defined as

- True negative: An in vivo negative (non-teratogen) showed a negative in vitro test result,
- True positive: An in vivo positive (teratogen) showed a positive in vitro test result,
- False negative: An in vivo positive (teratogen) showed a negative in vitro test result,
- False positive: An in vivo negative (non-teratogen) showed a positive in vitro test result.

The values for accuracy, sensitivity and specificity were calculated as

- Accuracy = (TN+TP)/(TN+TP+FN+FP),
- Sensitivity = TP/(TP+FN),
- Specificity = TN/(TN+FP).

3. Results

3.1 Characterisation of pluripotent stem cells

In the UKN1 6-day test system the human induced pluripotent stem cell line “SBAD2” was cultivated for the first time in the two different cell culture systems “Essential 8” and “DEF-CS”. In order to investigate if both systems were able to maintain the pluripotency of the SBAD2 cells, the morphological appearance as well as the expression of pluripotency-specific immunohistochemical markers of SBAD2 cells was analyzed.

3.1.1 Morphological appearance of SBAD2 cells

For the morpholigcal characterization, phase-contrast images at 24 h, 48 h and 72 h after cell seeding were taken. To investigate the impact of different seeding densities and passages, low and high seeding densities as well as low and high passages were used.
24 h after cells have been seeded as single cells in Essential 8, each cell showed an angular shape and sharp boundaries (**Figure 2** A+D). Single cells connected to each other and bridged the intercellular space via dendrite-like structures, whereas groups of cells started to form colonies. Within the colonies, single cells were separated by a whitish intercellular space. After 48 h, the formation of the colonies proceeded (**Figure 2** B+E). The edges of the colonies were smooth, defined and highlighted. When more cells were present, colonies connected to each other and formed a network-like structure. The sharp boundaries of each single cell disappeared so that an idenfitcation of single cells within a colony was not possible at this stage. Cells which had not formed colonies died, leaving large spaces between the colonies. After 72 h, cells formed homogeneous flat cell layers which were in distinct areas so high in density, that no space between single cells was left. At this stage, single cells within the colonies could be identified as small, round cells which had a comparable big nucleus (**Figure 2** C+F). No major differences could be observed between the low and the high passage or between the low and the high intial seeding densitiy except for the overall cell density.
SBAD2 cells which had been cultivated in DEF-CS also showed sharp boundaries but where less angular than in Essential 8 after 24 h (**Figure 3** A+D). Formation of colonies started as well and single cells were separated by a whitish intercellular space. Some cells did connect to each other via dendrite-like structures. After 48 h, colonies had grown (**Figure 3** B+E). In areas of higher cell densities, the whitish intercellular space disappeared and single cells could be identified as small, round cells which had a comparable big nucleus. After 72 h, the number of high density-areas increased and formed a homogenous cell layer (**Figure 3** C+F). No major differences could be observed between the low and the high passage or between the low and the high intial seeding densitiy except for the overall cell density.
Although SBAD2 cells showed differences in their appearance which were related to the specific cell culture system, both systems supported distinct morphological key elements of a pluripotent stem cell culture. These were the formation of colonies, the “compaction” (i.e. the formation of high density areas of small, round cells) and the overall homogenous morphological appearance. In total, the Essential 8 as well as the DEF-CS culture system was able to maintain a pluripotent morphology of SBAD2 cells.

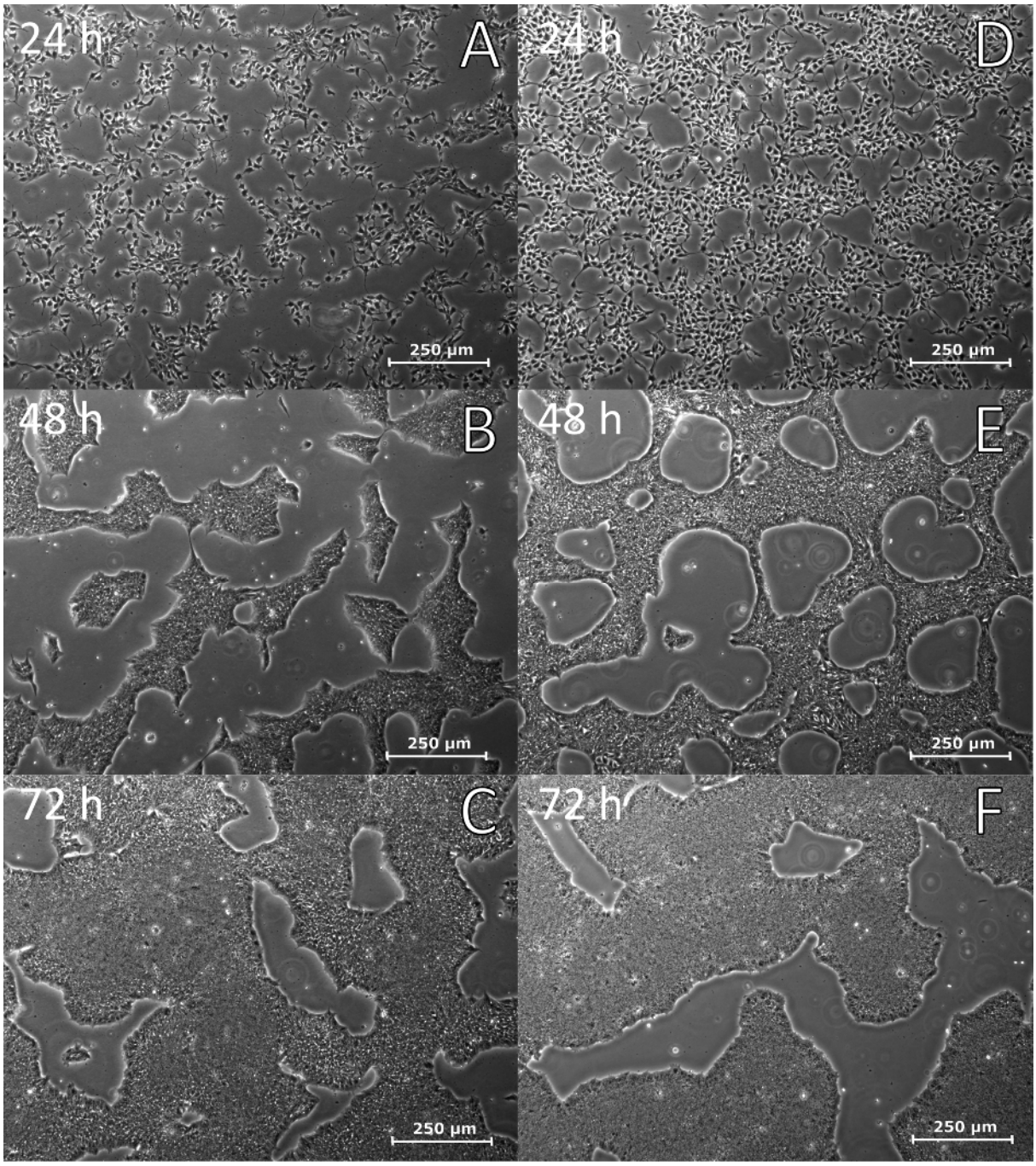


Figure 2: *Phase-contrast pictures of SBAD2 cells cultured in Essential 8 (ThermoFisher Scientific). The pictures show cells of passage 11 (A-C) and passage 26 (D-F) at 24 h, 48 h and 72 h after they have been seeded as single cells at densities of 12000 cells/cm² and 28000 cells/cm², respectively, at the beginning of these passages.*

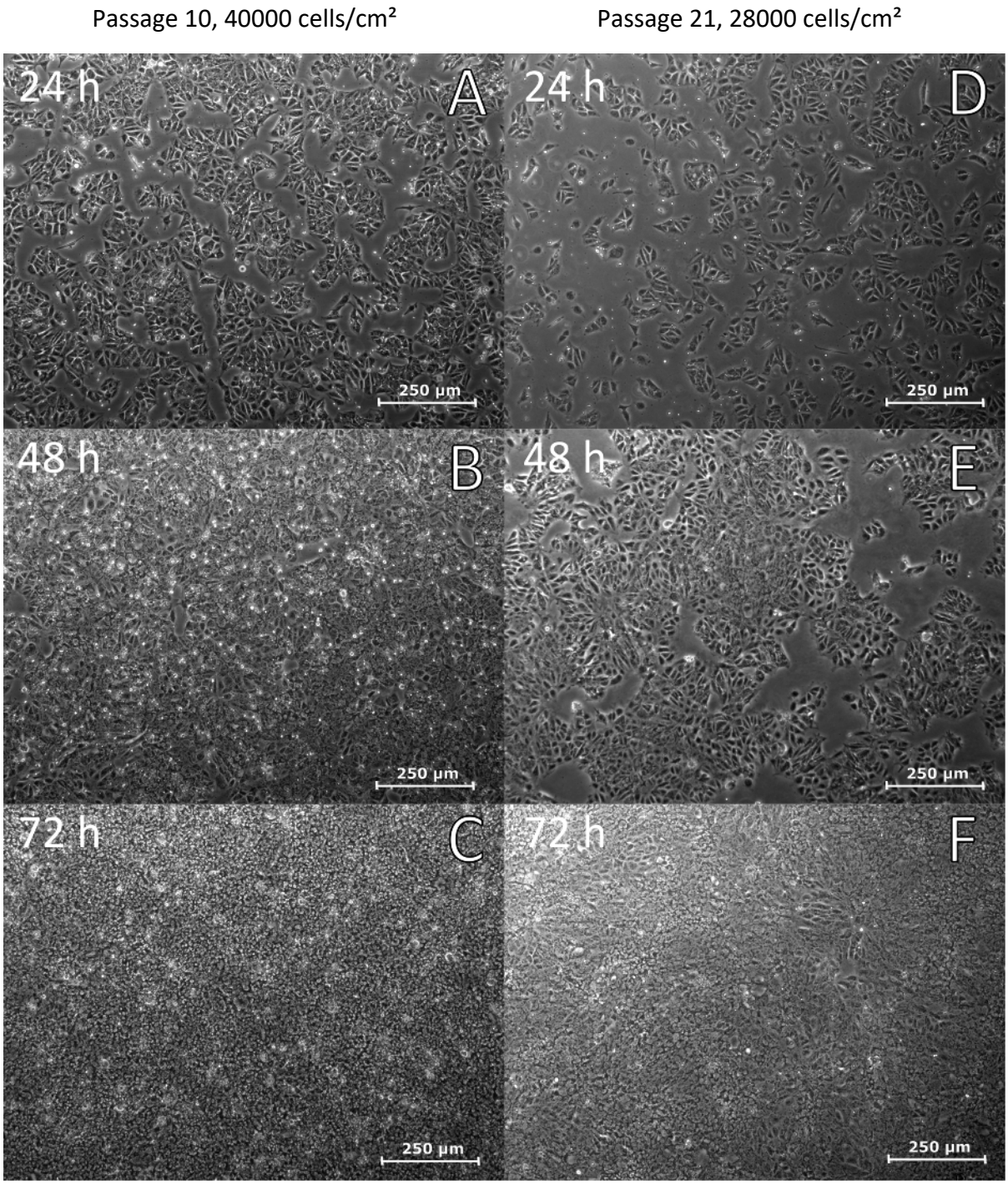


Figure 3: *Phase-contrast pictures of SBAD2 cells cultured in DEF-CS (Cellartis). The pictures show cells of passage 10 (A-C) and passage 21 (D-F) at 24 h, 48 h and 72 h after they have been seeded as single cells at densities of 40000 cells/cm² and 28000 cells/cm², respectively, at the beginning of these passages.*

3.1.2 Immunohistochemical analysis of SBAD2 cells

To investigate the pluripotent state of SBAD2 cells, it was analyzed via immunohistochemistry if the cells expressed the pluripotent markers OCT4A, SOX2, NANOG, SSEA4, TRA-1-60 and TRA-1-81. These antigens were targeted with marker-specific primary antibodies and visualized with Alexa Fluor 488-conjugated secondary antibodies. DAPI proved the existence of cells. To verifiy the specificity of the secondary antibodies, histological controls were stained without primary antibodies.
In Essential 8, the nuclear markers OCT4A, NANOG and SOX2 were strongly expressed, even though the Alexa Fluor 488-conjugated secondary antibody showed a high background signal in the histological control (**Figure 4**,
Supplementary figure 1, **Supplementary figure 2**). The surface markers SSEA4, TRA-1-60 and TRA-1-81 were not as strongly expressed as the nuclear markers. TRA-1-81 had the strongest signal (**Figure 4**, **Supplementary figure 4**), followed by TRA-1-60 and SSEA4 (**Figure 4**, **Supplementary figure 3**). The cobblestone-like expression patterns were very heterogeneous, i.e. some cells demonstrated a strong expression whereas other cells seemed to lack the expression of markers completely, but the overall expression was still detectable and distinguishable from the high background signal of the histological control (**Supplementary figure 4**). It has to be noted, that the pictures were made from 20 months old stainings and the markers' signal intensity probable had been diminished during that time.
Like Essential 8, the DEF-CS culture system supported a strong expression of the nuclear markers OCT4A, SOX2 and NANOG (**Figure 4**, **Supplementary figure 5**, **Supplementary figure 6**). A background signal from the histological control was completely absent (**Supplementary figure 6**). The surface markers SSEA4, TRA-1-60 and TRA-1-81 were again not as strongly expressed as the nuclear markers, but notably stronger than in Essential 8. The expression of SSEA4 (**Figure 4**, **Supplementary figure 7**) and TRA-1-81 (**Figure 4**,**Supplementary figure 8**) was also heterogeneous and formed cobblestone-like patterns, but was strong in most cells. Expression of TRA-1-60 (**Figure 4**, **Supplementary figure 7**) was comparable weak, but distinguishable from the histological control (**Supplementary figure 8**).
Overall, all immunohistochemical markers that were relevant for stem cell-pluripotency were expressed in SBAD2 cells, demonstrating the pluripotent state of the cells and the ability of both cell cultures systems Essential 8 and DEF-CS to maintain it.

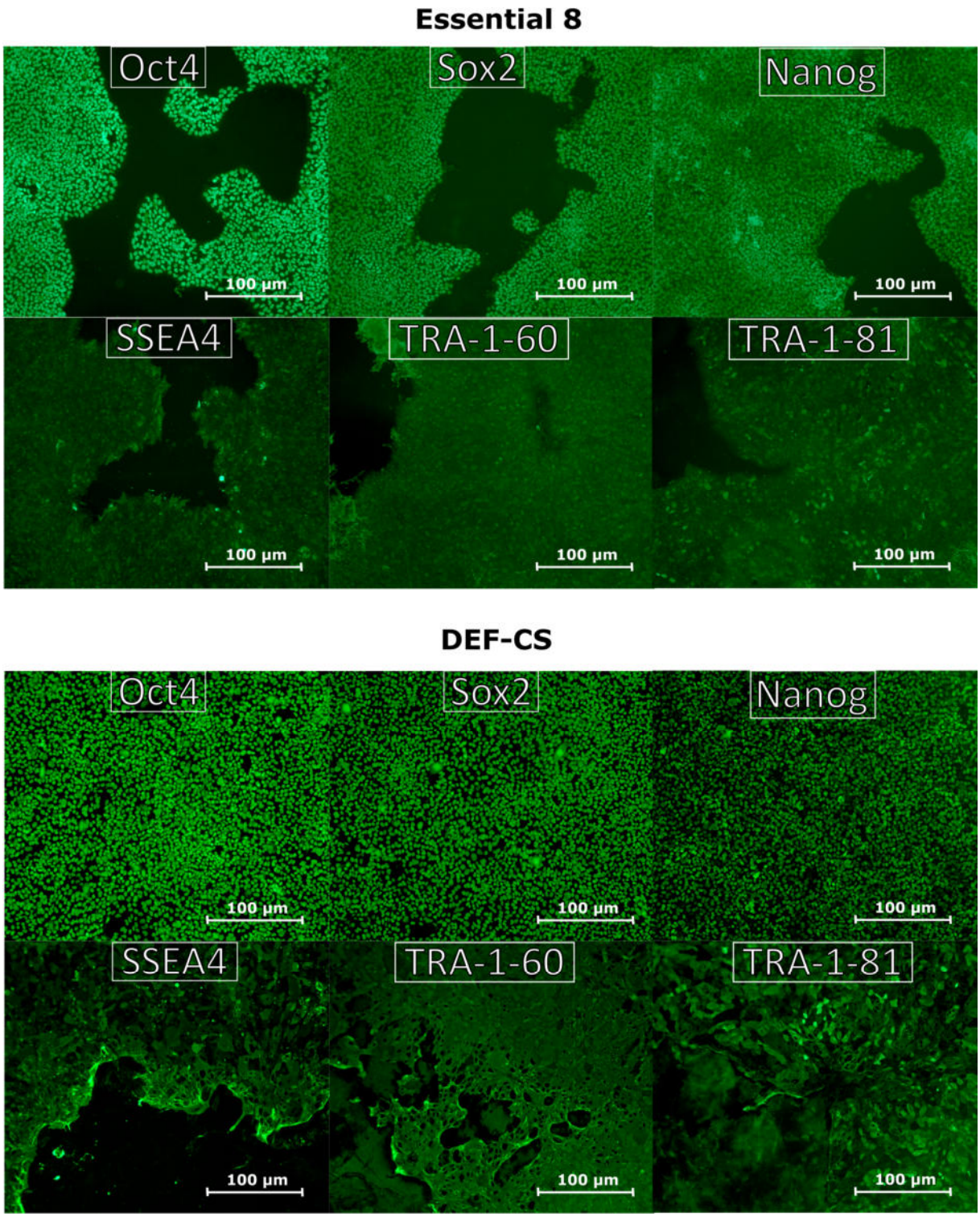


Figure 4: *Immunohistochemical staining of the stem cell markers OCT4A, SOX2, NANOG, SSEA4, TRA-1-60 and TRA-1-81 expressed in SBAD2 cells cultured in Essential 8 and DEF-CS. The markers were visualized by Alexa Fluor 488-conjugated secondary antibodies.*

3.2 Establishment of the UKN1 6-day differentiation protocol

At the beginnings of the thesis, the UKN1 6-day protocol had to be established. For that purpose, the protocol (**Figure 5**) was performed several times to analyze if the pluripotent stem cells, which were cultivated in the two culture systems Essential 8 and DEF-CS, could reproducibly differentiated to neuralepithelial progenitor cells. The quality and long-term stability of the protocol was investigated by controls of the morphology and gene expression analysis of NEPs as described in more detail in the following chapters.

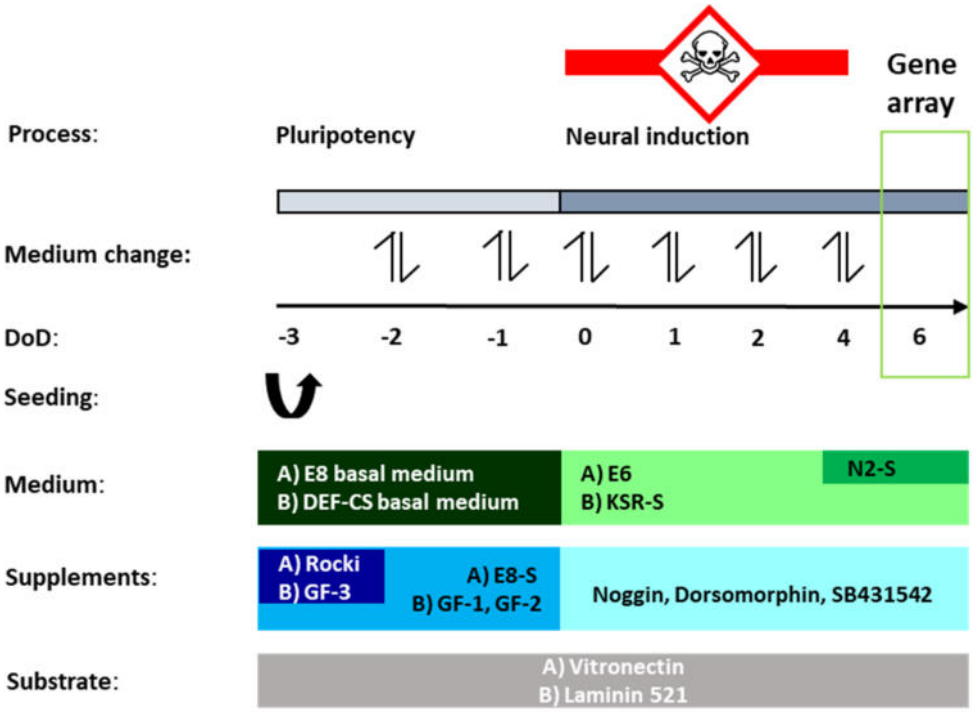


***Figure 5**: The scheme depicts the differentiation protocol in the two versions A and B from DoD-3 to DoD 6. In the pluripotency phase, the applied E8 or DEF-CS basal media and corresponding supplements E8-S or GF-1 and GF-2 maintained the pluripotent state of the hiPSCs. The supplements Rocki or GF-3 additionally given on the day of seeding (DoD -3) supported the survival of hiPSCs which were seeded as single cells on vitronectin or laminin 521. From DoD 0 onwards, the change to E6 or KSR-S spiked with noggin, dorsomorphin and SB431542 initiated the neural differentiation of the cells. Simultaneously, cells were exposed to potential (non-)developmental toxic substances for a total of 96 h. On DoD 4, substances were withdrawn and addition of 25 % N2-S further enhanced the neural differentiation. On DoD 6, NEPs were harvested for gene array analysis. Media changes were conducted as indicated.*

3.2.1 Quality control of the differentiation process via cell morphology

Each differentiation was monitored carefully to control the vitality and homogeneous differentiation of the cells. For that purpose, cells were checked morphologically on every day with a medium change and on the last day before they were harvested for gene expression analysis.
Pluripotent SBAD2 cells which were seeded on DoD -3 were homogenously distributed on DoD -2 and grown to a confluency of ca. 80 % until DoD 0 in both protocol versions A and B (**Figure 6**, **Figure 7**) like it was described in more detail in a previous chapter (3.1.1). Cells under control conditions, which were not incubated with substances, grew to a very dense cell layer until DoD 6. Cell boundaries of single cells within the dense cell layer were present on DoD 0, 1 and 2, but begin to disappear on DoD 4. On DoD 6, cells were not distinguishable from each other within the cell layer.
So, both versions A and B supported the vitality and constant growth as well as a homogeneous differentiation of SBAD2 cells along the UKN1 6-day protocol.

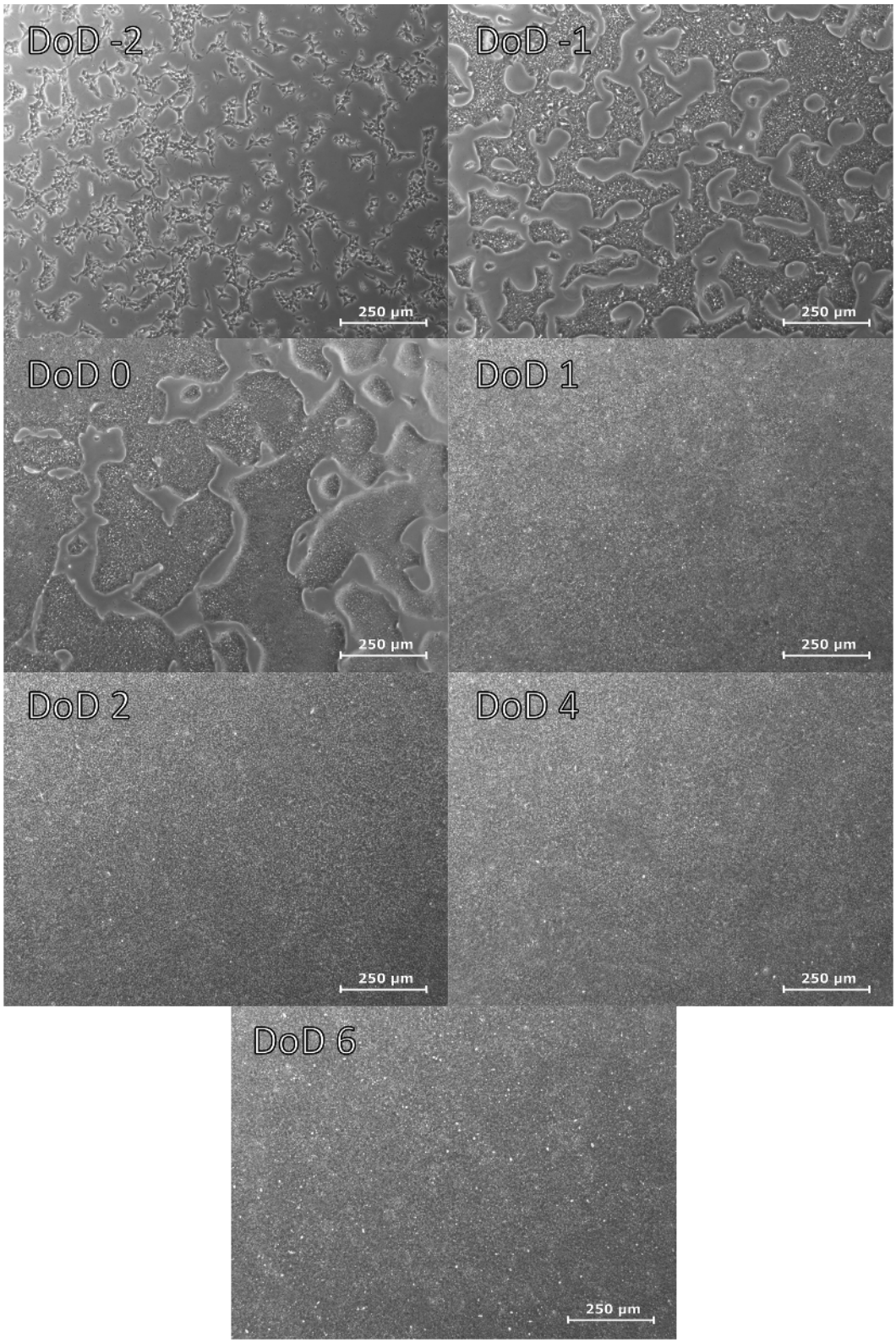


Figure 6: *Light microscopy images of SBAD2 cells (passage 19) in the pluripotency and the neural induction phase of a differentiation (protocol version A). Cells had been cultivated in Essential 8 and were seeded at 24,000 cells/cm². DoD -2 shows the cells 24 h after seeding. DoD -1 and DoD 0 shows growing cells in their pluripotent state. On DoD 0, at approximately 80 % confluency, neural differentiation was induced until DoD 6.*

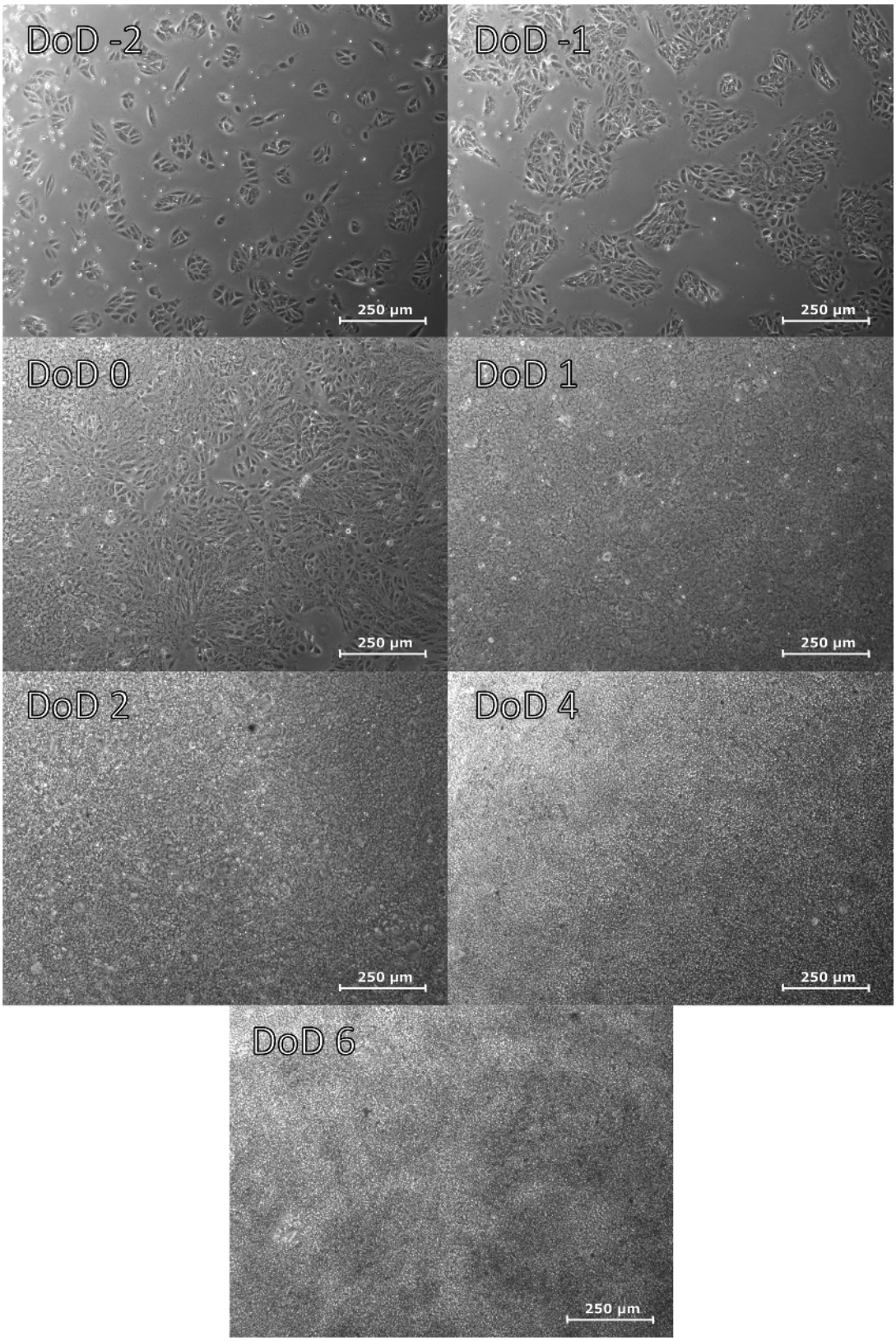


Figure 7: *Light microscopy images of SBAD2 cells (passage 9) in the pluripotency and the neural induction phase of a differentiation (protocol version B). Cells had been cultivated in DEF-CS and were seeded at 12,000 cells/cm². DoD -2 shows the cells 24 h after seeding. DoD -1 and DoD 0 shows growing cells in their pluripotent state. On DoD 0, at approximately 80 % confluency, neural differentiation was induced until DoD 6.*

3.2.2 Quality control of neuroepithelial differentiation via biomarkers in RT-qPCR

In order to verify the success of a differentiation, the gene expression changes of seven biomarker-genes that took place in cells which differentiated from hiPSCs to NEPs were analyzed with a reverse transcription quantitative real-time poly-chain reaction (RT-qPCR).
In both versions A and B of the UKN1 6-day protocol (**Figure 5**), untreated hiPSCs-derived NEPs showed a significant downregulation of the genes NANOG and OCT4A and an upregulation of the genes OTX2, PAX6, TFAP2 and EMX2, whereas expression of the gene MSX1 was not altered compared to hiPSCs (**Figure 8**).

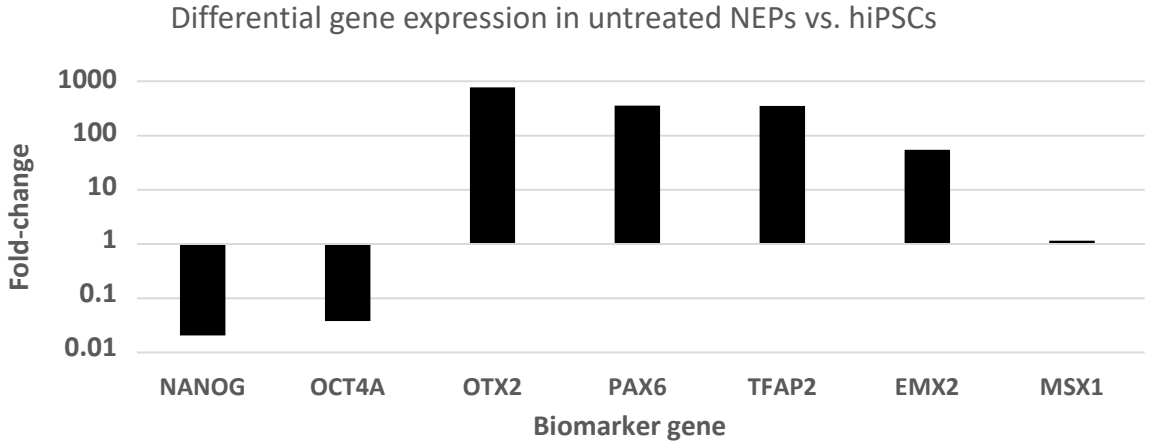


Figure 8: *Gene expression changes of seven biomarker-genes in untreated NEPs compared to hiPSCs.*

Moreover, to assess if cells of a differentiation and the same biological replicate were generally susceptible and not resilient to substances, each differentiation included conditions where cells were exposed to VPA at relevant concentrations of 0.6 mM and 1.0 mM (see also the chapter 3.3) and which thus served as positive controls. Compared to untreated NEPs, VPA induced a significant downregulation of OTX2, PAX6 and EMX2 and an upregulation of TFAP2B and MSX1. (**Figure 9**).

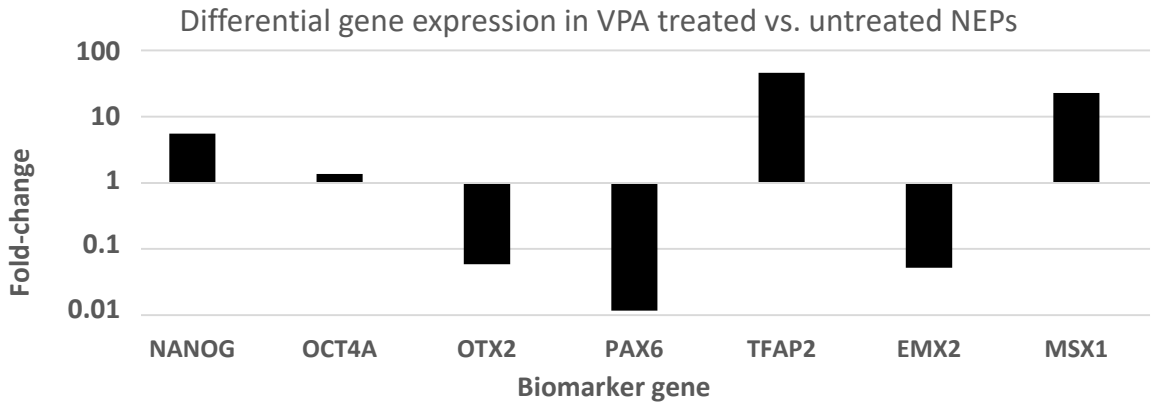


Figure 9: *VPA-induced gene expression changes of seven biomarkers in NEPs compared to untreated NEPs.*

All in all, the expression of the seven biomarker-genes was completely in line with the original UKN1 protocol (Balmer et al. 2012; Rempel et al. 2015; Dreser et al. 2020) and the differentiation of SBAD2-hiPSCs to NEPs was possible with both versions A and B. Consequently, the establishment of the UKN1 6-day protocol was successful.

3.3 Search and selection of representative non-teratogenic and teratogenic compounds

In the UKN1 6-day test system, differentiating cells were exposed to a variety of teratogenic and non-teratogenic compounds to figure out if they had particular effects on the cells (**Figure 5**). In order to find the best candidates for both classes that could be applied in the test system, an extensive literature search was conducted that focused on compounds that were well-documented and offered lots of information, especially about their teratogenic potential and pharmacokinetics in humans. In this context, the "pregnancy class" letter categorization systems from the North American authority "Food and Drug Administration" (FDA) and the Australian authority "Therapeutic Goods Administration" (TGA) (Ramoz and Patel-Shori 2014; TGA n.d.) proved to be very helpful. Although these systems were old-fashioned and the US FDA had already updated their categorization system since 2015, they still offered valuable information. All compounds that have been once approved and classified by these systems usually were rich in history, and well-founded knowledge about their medical use was available. Similarly, both systems had five main categories A, B, C, D and X to classifiy the risk during pregnancy:

- Category A: Compounds were safe in the use during pregnancy, proven by well-controlled studies in humans or plenty data from pregnant women; there is no or just a negligible risk of teratogenicity
- Category B: Compounds were also considered to be safe, but they lack sufficient human data
- Categories C and D: Compounds showed little or some evidence for teratogenicity in humans or animals
- Category X: Compounds with known teratogenic activity in humans or with a suspected high teratogenic potential based on animal experiments

The non-teratogenic and teratogenic compounds that were found in the course of the research are presented and described in the following two subschapters.

3.3.1 Non-teratogenic compounds and their characterstics

In order to find non-teratogenic compounds for the application in the UKN1 6-day system, the literature and internet search looked for compounds that had not shown any evidence for teratogenicity so far and were therefore considered to be harmless to the human fetus. Of special interest were all FDA and TGA pregnancy classes A and B compounds. In total, sixteen compounds with such characteristics were selected (**Table 18**). Thirteen of them were classified as pregnancy class A and B compounds and prevalently used as antibiotics, medicines or dietary supplements. The remaining three compounds magnesium (chloride), retinol and sucralose were not classified but commonly considered to be safe during pregnancy.

Table 18: Non-teratogenic substances in the UKN1 6-day test system. Details about their classification within the US FDA and AU TGA pregnancy categories and about their (medical) applications are given. Information was obtained from www.drugs.com (accessed in November 2020) if not stated otherwise. RDA = Recommended daily allowance. n/a = not available; [a]doses >RDA; [b]Prolonged high dosing (in constipation or pre-eclampsia); [c]exempt from classification; [d]approved as a sweetener in foods (FDA 2019)

Non-teratogenic compound	Pregnancy category	(Medical) Application
Ampicillin	A / B	Antibiotic
Ascorbic acid	A / C[a]	Dietary supplement
Buspirone	B	Treatment of anxiety
Chlorpheniramine	B	Antihistamine; used to treat allergy symptoms
Dextromethorphan	A	Cough suppressant, affects the signals in the brain that trigger cough reflex
Diphenhydramine	A / B	Antihistamine, treatment of allergy symptoms, motion sickness, insomnia and cold
Doxylamine	A	Antihistamine, used to treat allergy symptoms or short-termly insomnia Primary used in combination with pyridoxine to treat nausea and vomiting of pregnancy (Madjunkova et al. 2014)
Famotidine	B	Histamine-2 blocker, decreases the amount of acid the stomach produces, treatment of ulcers in the stomach and intestines and gastroesophageal reflux disease
Folic acid	A	Dietary supplement during pregnancy
Levothyroxine	A	Thyroid hormone thyroxine (T_4) replacement for hypothyroidism
Liothyronine	A	Thyroid hormone triiodothyronine (T_3) replacement for hypothyroidism
Magnesium (chloride)	n/a; C[b] / D[b]	Dietary supplement; treatment of constipation, dyspepsia, hypomagnesemia, pre-eclampsia
Methicillin	B	Antibiotic
Ranitidine	B	Histamine-2 blocker, decreases the amount of acid the stomach produces, treatment of ulcers in the stomach and intestines and gastroesophageal reflux disease
Retinol	n/a[c]	Vitamin, essential nutrient; retinoid
Sucralose	n/a[d]	Artificial sweetener and sugar substitute in food

As non-teratogenic substances are known to not induce developmental defects to unborn children at therapeutic doses, a proof of their ability to reach the children's biosystems was needed. Therefore, information about the transplacental passage was gathered to verify whether the non-teratogens could take effect in vivo at all. Of the sixteen non-teratogenic substances, ten were known to cross the human placenta at different rates (**Table 19**) and ascorbic acid and folic acid even accumulate in the fetus after administration to the mother. For diphenhydramine and doxylamine, only animal data from sheep or rhesus monkey, respectively, could be found. In the cases of buspirone, chlorpheniramine, dextromethorphan and sucralose, data about the transplacental passage could neither be found for humans nor for animals.

In total, sixteen compounds were found to be appropriate non-teratogenic candidates for the application in the UKN1 6-day assay.

Table 19: Data on the placental transfer of non-teratogenic substances

Non-teratogenic compound	Placental transfer and ratio of fetal:maternal compound concentration in blood or plasma	Reference
Ampicillin	Yes (1:1)	Pacifici 2006
Ascorbic acid	Yes Ascorbic acid enriches in the fetus 2- to 4-fold compared to maternal plasma / blood concentrations	Briggs et al. 2017, p. 1555 ff.
Buspirone	Unknown	No information online at 18.11.2020, no information in Briggs et al. 2017
Chlorpheniramine	Unknown	No information online at 18.11.2020, no information in Briggs et al. 2017
Dextromethorphan	Unknown	No information online at 18.11.2020, no information in Briggs et al. 2017
Diphenhydramine	In humans: Unknown In sheep: Yes	Yoo et al. 1986
Doxylamine	In humans: Unknown In rhesus monkey: Yes, occurs at a rate of 30-60 %	Slikker, Jr. et al. 1987
Famotidine	Yes (Fetal to maternal ratio of compound in blood/plasma: 0.64 ± 0.13)	Doi et al. 1991
Folic acid	Yes Folic acid enriches in the fetus 2- to 4-fold compared to maternal plasma / blood concentrations	Henderson et al. 1995; Hutson et al. 2012
Levothyroxine	Yes, occurs at a rate of estimated 30 %	Vulsma et al. 1989; Carr et al. 1959; Burrow et al. 1994
Liothyronine	Yes	Dussault et al. 1969
Magnesium (chloride)	Yes, almost 1:1 (Fetal to maternal ratio of compound in blood/plasma 0.94 ± 0.15)	Brookfield et al. 2016
Methicillin	Yes (1:1)	Pacifici 2006; Depp et al. 1970
Ranitidine	Yes (Fetal to maternal ratio of compound in blood/plasma - after peroral application: 0.38 - after intravenous application: 0.9)	McAuley et al. 1983
Retinol	Yes	Sklan et al. 1985; Dancis et al. 1992; Torma and Vahlquist 1986
Sucralose	Unknown	No information online at 18.11.2020, no information in Briggs et al. 2017

3.3.2 Teratogenic compounds and their characteristics

For the finding of suitable teratogenic substances, a literature and internet search for all compounds which had shown in vivo evidence for teratogenicity, especially in humans, was done. This time, FDA and TGA class D and X compounds were in focus. 23 compounds that were known or highly suspected to cause developmental defects in humans were selected. They were classified as D and X compounds, except for entinostat, favipiravir, methylmercury and trichostatin A (TSA) (**Table 20**), and they were prevalently used as antibiotics and medicines in the treatment of various diseases.

Table 20: Teratogenic substances in the UKN1 6-day test system. Details about their classification within the US FDA and AU TGA pregnancy categories and about their (medical) applications are given. Information was obtained from www.drugs.com (accessed in November 2020) if not stated otherwise. n/a = not available

Teratogenic compound	Pregnancy category	(Medical) Application
9-cis-Retinoic acid (Alitretinoin)	D	Topical treatment of skin lesions caused by AIDS-related Kaposi's sarcoma
Acitretin	X	Treatment of severe psoriasis
Actinomycin D	D	Antineoplastic, treatment of different types of cancers
Atorvastatin	X (Pfizer Ireland Pharmaceuticals 2009; Lecarpentier et al. 2012)	Antilipemic, treatment of high cholesterol levels, i.e. decreases the amount of low density lipoprotein cholesterol and triglycerids in the blood and increases high density lipoprotein cholesterol
Carbamazepine	D	Anticonvulsant
Doxorubicin	D	Antineoplastic, cancer treatment
Entinostat (MS-275)	n/a	Clinical trial candidate for cancer treatment
Favipiravir	n/a	Antiviral drug
Isotretinoin	X	Treatment of severe acne
Leflunomide	X	Anti-inflammatory; treatment of rheumatoid arthritis
Lithium (chloride)	D	Mood stabilizer; treatment of symptoms of bipolar disorder
Methotrexate	D / X	Treatment of leukemia, certain types of cancer and autoimmune diseases like severe psoriasis and rheumatoid arthritis in juveniles and adults
Methylmercury	n/a	Bioaccumulative environmental toxicant
Panobinostat	Approved, but not assigned (Recommendation: D)	Treatment of multiple myeloma
Paroxetine	D	Antidepressant; treatment of depression, anxiety and other disorders
Phenytoin (5,5-Diphenylhydantoin)	D	Anticonvulsant
Teriflunomide (A-771726)	X	Treatment of multiple sclerosis
Thalidomide	X	Treatment of multiple myeloma and leprosy
Trichostatin A (TSA)	n/a	Antifungal antibiotic
Valproic acid (VPA)	D / X (FDA 2013)	Anticonvulsant, treatment if bipolar disorder, epilepsy and migraine
Vinblastine	D	Antineoplastic, cancer treatment
Vismodegib	X	Treatment of basal-cell carcinoma
Vorinostat	D	Treatment of cutaneous T-cell lymphoma

The substances covered several different mechanisms of action (MoAs) (**Table 21**) and represented an interesting selection to investigate their effects on the differentiating hiPSCs. Some of the compounds were selected to distinct features and reasons beside the classification of the FDA and TGA, which are explained below:

- **9-cis-Retinoic acid, acitretin and isotretinoin:** These compounds are derived from retinol and are so-called "retinoids". They share a similar molecular structure, MoA and medical application and should therefore cause a similar effect in the test system.
- **Entinostat, panobinostat, trichostatin A (TSA) and vorinostat:** They belong to a class called "Histone deacetylase inhibitors" (HDACi) which are known to influence gene transcription and are therefore perfect candidates for a test system which analyzes transcriptional changes. As they have a common MoA, they should act similarly in the test system. VPA, which also has HDAC inhibitory properties, is not regarded to be solely a HDACi due to its multiple molecular targets.
- **Methylmercury** is a known neurodevelopmental toxicant and could give insights about the detectability of such compounds in the system. Furthermore, it was used along with HDACis in a former study (Rempel et al. 2015).
- **Leflunomide and teriflunomide:** Leflunomide is activated to its main metabolite teriflunomide in vivo. Its effect in the in vitro test system and differences to teriflunomide could give insights about the metabolic activity of the cells.
- **Favipiravir** was selected because of its promising antiviral potential during the highly topical COVID-19 pandemic in 2020.

In order to verify the teratogenic status of the compounds, the literature was searched for evidence of their teratogenicity in vivo. For all of the 23 teratogenic compounds except for entinostat and TSA, evidence for teratogenicity could be found in either humans or animals (**Table 21**). Information about the ability to cross the placenta was not additionally researched for teratogenic compounds as this information was indirectly included in the data about the teratogenicity: Only compounds that can cross the placenta can cause teratogenicity.
In sum, 23 compounds were found to be representative teratogenic substances for the use in the UKN1 6-day test system, and together with the 16 non-teratogens, a total of 39 proper substances were selected for the application in the in vitro test. Although the data for some were not complete, they belonged to the best-known and -characterized non-teratogenic and teratogenic substances that were available at that time and represented a well-chosen set of substances for the in vitro testing of teratogenicity.

Table 21: Mechanisms of action and teratogenic evidence of teratogenic compounds

Teratogenic compound	Mechanism of action (MoA)	Evidence for teratogenicity
9-cis-Retinoic acid (Alitretinoin)	Binds at and acts over retinoic acid receptors (RARs) and retinoid x receptors (RXRs) (Comptour et al. 2016; Allenby et al. 1993; Idres et al. 2002)	Short trunk, limb reduction defects, ectrodactyly, cleftpalate, micrognathia, and other craniofacial defects in mice (Kochhar et al. 1995)
Acitretin	Activates, but does not bind to RARs (Pilkington and Brogden 1992). Binding to RXRs is not known	Retinoid acid related defects (see isotretinoin) in few human case reports and animal studies (Briggs et al. 2017, p. 17–19)
Actinomycin D	Binds to DNA and inactivates RNA synthesis (Hollstein 1974)	Malformations of CNS, viscera and skeleton in animals (rat, hamster, rabbit) (Tuchmann-Duplessis et al. 1973). Human data do not show evidence for teratogenicity (Briggs et al. 2017, p. 364–366)
Atorvastatin	3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase inhibitor (Lecarpentier et al. 2012)	CNS, skeleton and cardiovascular malformations in few human and animal studies in which different statins were used. The biochemical mechanism indicates a potential embryo-fetal risk. However, in a lot of studies, statins didn't induce malformations (Briggs et al. 2017, p. 111–112; Lecarpentier et al. 2012)
Carbamazepine	Blocks voltage-gated sodium channels (Rogawski et al. 2016)	Association with an increased incidence of major and minor malformations, including neural tube defects (NTDs), cardiovascular and urinary tract defects, and cleft palate, supported by different human studies and case reports (Briggs et al. 2017, p. 210–216; Matalon et al. 2002)
Doxorubicin	Among others, interaction with and alteration of DNA and related proteins; production of free radicals and reactive oxygen-species (ROS) (Varela-López et al. 2019)	Structural anomalies and death in rats and rabbits; association with similar outcomes in humans after exposure during organogenesis (Briggs et al. 2017, p. 452–454)
Entinostat (MS-275)	HDACi (Connolly et al. 2017)	No data available, refer to other HDACis
Favipiravir	Selective inhibitor of the RNA-dependent RNA polymerase of influenza virus (Furuta et al. 2013)	Teratogenic and embryotoxic effects in animal studies. No human data available (Pharmaceutical and Food Safety Bureau 2014, p. 45 ff.)
Isotretinoin	Binds at and acts over retinoid acid receptors (RARs) but not retinoid X receptors (RXRs) (Comptour et al. 2016; Allenby et al. 1993; Idres et al. 2002)	Craniofacial, CNS, cardiovascular and thymic defects, well-documented in animal studies and humans (Nau 2001; Browne et al. 2014; Briggs et al. 2017, p. 779–781)
Leflunomide	Is metabolized to its active metabolite teriflunomide; inhibits dihydroorotate dehydrogenase (DHODH), a key player in de novo pyrimidine synthesis (Fox et al. 1999)	Teratogenic in animals at human relevant doses. Human data do not show evidence for teratogenicity (Briggs et al. 2017, p. 818–820)
Lithium (chloride)	Many cellular targets, e.g. inhibition of glycogen synthase kinase 3 (GSK-3) (Malhi et al. 2013)	Although evidence is weak, lithium may lead to an increased incidence of cardiovascular defects and Ebstein's anomaly (Briggs et al. 2017, p. 848–852)

Teratogenic compound	Mechanism of action (MoA)	Evidence for teratogenicity
Methotrexate	Multiple targets; best known is its antifolate effect as dihydrofolate reductase inhibitor (Cronstein and Aune 2020)	Multiple defects in animals and humans, e.g. limb malformations, CNS damages and craniofacial defects (Briggs et al. 2017, p. 927–932; Verberne et al. 2019; Hyoun et al. 2012)
Methylmercury	Exact mechanism is unknown; among others, the mostly hypothesized is the production of ROS (Fretham et al. 2012; Farina et al. 2011; Unoki et al. 2018)	Malformations in animals, neural and neurodevelopmental defects in animals and humans (EPA 2012, p. 3-85 to 3-93; Harada 1995; Grandjean et al. 1997; Karagas et al. 2012)
Panobinostat	HDACi (Novartis Pharmaceuticals 2016)	Teratogenic, embryo-fetal toxic and mutagenic in animals. Human data are not available (Briggs et al. 2017, p. 1105; Novartis Pharmaceuticals 2016)
Paroxetine	Selective serotonin reuptake inhibitor (SSRI) (Davis et al. 2016)	Different major and minor malformations and defects in some human studies and case reports, e.g. cardiac defects; withdrawal symptoms in neonates (Briggs et al. 2017, p. 1110-1120)
Phenytoin (5,5-Diphenylhydantoin)	Blocks voltage-gated sodium channels (Rogawski et al. 2016)	Major and minor congenital abnormalities like craniofacial and limb defects; fetal hydantoin syndrome (FHS); association with neurodevelopmental defects; evidence is supported by different human studies and case reports (Briggs et al. 2017, 1161 ff.)
Teriflunomide (A-771726)	Active metabolite of leflunomide; inhibits dihydroorotate dehydrogenase (DHODH), a key player in de novo pyrimidine synthesis (Fox et al. 1999)	Teratogenic in animals at human relevant doses. Human data do not show evidence for teratogenicity (Briggs et al. 2017, 1395 ff.)
Thalidomide	Multiple possible, though mostly unknown targets; most prominent are its antiangiogenic and TNF-α inhibitory properties and binding to Cereblon (Vargesson 2015)	Severe defects of limbs, skeleton, head & face, CNS and organ system in numerous human cases (Briggs et al. 2017, p. 1405-1413; Vargesson 2015)
Trichostatin A (TSA)	HDACi (Xu et al. 2007)	No data available, refer to other HDACis
Valproic acid (VPA)	Multiple possible targets and actions: Activator of glutamic acid decarboxylase (GAD), voltage-gated sodium channel blocker, reduction of folic acid levels, HDACi (Ornoy 2009; Rogawski et al. 2016)	Neural tube defects (NTDs), spina bifida, fetal valproate syndrome, defects of the head, face, limbs, urogenital tract and neurodevelopment in animals and humans (Briggs et al. 2017, p. 1516-1523; Ornoy 2009)
Vinblastine	Inhibits mitosis and arrests cell cycle by affecting microtubule dynamics (Dhamodharan et al. 1995)	Evidence is weak: Malformations in hamster embryos in one study (Ferm 1963), unspecified details in other animal studies, no data on human pregnancies (Briggs et al. 2017, p. 1543-1544)
Vismodegib	Hedgehog pathway inhibitor (Genentech 2012)	Congenital malformations in rats at human relevant doses. No human data available (Morinello et al. 2014; Genentech 2012)
Vorinostat	HDACi (Merck Sharp & Dohme 2018)	Teratogenic in animals at human relevant doses. Human data are not available (Briggs et al. 2017, p. 1567-1568; Merck Sharp & Dohme 2018)

3.4 Determination of relevant in vivo concentrations

Before the non-teratogenic and teratogenic substances that were researched and described in the previous chapter 3.3 were used in the UKN1 6-day test system, it was necessary to know which concentrations would be most meaningful for the in vitro application. Therefore, a labor-intensive recherché about substance concentrations that were usually found in humans after the application of therapeutic doses was conducted.

For all substances except for TSA, information from human subjects were found and summarized for the non-teratogenics (**Table 22**) and teratogenics (**Table 23**). Common therapeutic doses were mainly collected from the US-webpage www.drugs.com, which provided a lot of information about substances that were used in medicine. Information about actual plasma and blood concentrations in humans were collected from clinical studies for most substances. In such studies, human subjects received therapeutic doses in distinct dosing schemes under medical supervision and the resulting substance concentrations in blood or plasma were carefully monitored, thus providing reliable high-quality data. Moreover, for most substances at least two of such studies could be found. In rare cases, e.g. famotidine, even clinical information about blood concentrations in neonates, where the mother received therapeutic doses shortly before the delivery, could be gathered.

Overall, a collection of high-quality data about human in vivo relevant concentrations of all 39 non-teratogenic and teratogenic substances (except for TSA) could be gathered.

3.4.1 Therapeutic doses & plasma/blood concentrations of non-teratogenic compounds

Table 22: Non-teratogenic compounds and their human relevant plasma or blood concentrations at therapeutic doses. Common therapeutic doses and dosing schemes were collected from www.drugs.com if not stated otherwise. Routes of application were as follows: PO= Peroral, IV=Intravenously, IM=Intramuscular or subcutaneously. RDA = Recommended daily allowance.

Non-teratogenic compound	Common therapeutic doses and dosing schemes in humans	Clinical study design and other resources of information	Plasma peak or blood concentrations	Reference
Ampicillin	Endocarditis: 2 g IV every 4 h Meningitis: 150 to 200 mg/kg/day IV in equally divided doses every 3 to 4 h Other infections: 500 mg PO, IM, IV every 6 h	a) 400 mg/kg/day in six doses IV, healthy children (n = 17) b) 1.5 g sultamicillin (containing the equivalent of 882 mg ampicillin) with probenecid (1.0 g) (n=8, men)	a) After 30 min: 37.3 ± 19 µg/ml (107 ± 55 µM) After 3 h: 11 ± 10.2 µg/ml (32 ± 29 µM) b) 23.1 µg/ml (66 µM)	a) Giachetto et al. 2004 b) Emmerson et al. 1983
Ascorbic acid	RDA: 80 mg 50 - 200 mg/day PO, IM, IV, or subcutaneously as dietary supplement	a) 30 - 2500 mg PO daily (n= 17, 7 men, 10 women) b) 30 - 2500 mg PO (n=7, men)	a) 200 mg: 90 µM (predicted) 1250 mg: 187 µM (predicted) 3000 mg: 206 µM (predicted) b) 60 mg: 20 µM (steady-state) 200 mg: 60 µM (steady-state) 2500 mg: 90 µM (steady-state)	a) Padayatty et al. 2004 b) Levine et al. 1996
Buspirone	15-60 mg in divided doses	a) 20 mg PO b) 20mg PO, healthy men (n=12) c) 30mg PO (n=9, healthy adults)	a) 2.5 ng/ml (5.9 nM) b) 1.15 ± 0.77ng/ml (0.9 - 4.55 nM) c) 6.6 ± 3.7 ng/ml (15.6 - 24.4 nM)	a) Mahmood and Sahajwalla 1999 b) Dalhoff et al. 1987 c) Lamberg et al. 1998
Chlor-pheniramine	4 mg PO every 4 to 6 hours	a) 8 mg chlorpheniramine maleate PO (4 male, 1 female healthy adults) b) (+)Chlorpheniramine 2 mg and 6 mg PO, 12 healthy men	a) 11.9 - 35.6 ng/ml (30.4 nM - 91 nM), mean 17.9 ng/ml (46 nM) b) 2mg: 3.4–7.4 ng/ml, mean 5.0 ± 1.1 ng/ml 6 mg: 2.0–14.3 ng/ml, mean 7.9 ± 3.2 ng/ml	a) Huang et al. 1982 b) Tagawa et al. 2002
Dextro-methorphan	30-120 mg / day	a) 30 mg (Dextrometorphan Hbr) (n=9) b) 30-60 mg PO, with or without quinidine (n=121, healthy adults)	a) 0.12 ± 0.03 µM b) max. 7.7 ± 7.0 ng/ml (28 ± 26 nM) with quinidine: max. 232 ± 96 ng/ml (856 ± 354 nM)	a) Schadel et al. 1995 b) Pope et al. 2004

Non-teratogenic compound	Common therapeutic doses and dosing schemes in humans	Clinical study design and other resources of information	Plasma peak or blood concentrations	Reference
Diphen-hydramine	12.5 - 76 mg / dose, PO, IV, IM, max. 400 mg / day	a) Single 50 mg oral dose, healthy women (n=25) b) 50 mg p.o once (healthy humans) c) 6.25 - 50 mg PO (allergic children) d) Overdose (humans)	a) 0.3 µM (74 ng/ml) b) 37 - 83 ng/ml c) 48 - 93 ng/ml d) 0.1 – 4.7 µg/ml (lethal)	a) Luna et al. 1989 b) Carruthers et al. 1978 c) Gelotte et al. 2018 d) Köppel et al. 1987
Doxylamine	10-40 mg daily	a) 25 mg PO once (n=25, healthy women) b) 12.5 mg PO once (n=12, 3 men, 9 women, healthy) c) 25 mg PO once (n=12, 3 men, 9 women, healthy)	a) 0.38 µM b) 0.23 µM c) 0.46 µM	a) Luna et al. 1989 b-c) Videla et al. 2013
Famotidine	20-40 mg PO daily	a) 40 mg PO (n=7, healthy adults) b) 20-40 mg, different studies c) 20 mg IM, pregnant women (n=34)	a) 17 to 139 ng/ml (50.3 - 411.9 nM) b)20 mg: 4 - 137 ng/ml (11.8 - 406 nM), mean 71 ng/ml (210 nM) 40 mg: 15 - 358 ng/ml (44.4 - 1060 nM) , mean 132 ng/ml (391 nM) c) Maternal: approx. 140 ng/ml (418 nM) Fetal: approx. 90 ng/ml (267 nM)	a) Morgan et al. 1990 b) Chremos 1987 c) Doi et al. 1991
Folic acid	400 µg daily 500 µg during Lactation 600 µg during Pregnancy	a) 1.1 mg or 5 mg PO once (n=6, non-pregnant healthy women) b) 400 µg PO daily during whole pregnancy (n=10, pregnant women)	a) Baseline: 0.011 – 0.034 µM (5 to 15 ng/mL) 1.1 mg: 0.014 µM (6.2 +/- 2.2 ng/mL) steady-state (estimated) 5.0 mg: 0.094 µM (41.6 +/- 11.3 ng/mL) steady-state (estimated) b) Umbilical cord blood: Up to 90.5 nM total folate	a) Nguyen et al. 2008 b) Obeid et al. 2010
Levothyroxine	max. 200 - 300 µg daily	a) 600 µg PO, 18 men and 18 women b) 54 - 300 µg 2 men, 4 women c) 600 µg (n=84) d) 115.2 ± 38.5 µg/day (n=10, 9 female, 1 male) e) Baseline, neonates cord blood (n=40)	a) 76.64 ± 16.48 ng/ml (99 nM ± 21.2 nM) b) max. 1.25±0.51 ng/dl c) 71.4 ± 16.0 ng/ml d) 106 ± 22 nM e) Cord blood: 35 - 70 nM	a) Tanguay et al. 2019 b) Kashiwagura et al. 2014 c) Yue et al. 2012 d) Celi et al. 2010 e) Vulsma et al. 1989
Liothyronine	25-75 µg daily	a) 40.3 ± 11.3 µg/day (n=10, 9 female, 1 male) b) 50 µg PO, (n=12, 4 females and 8 males) c) 50 - 300 µg / day (n=15, pregnant women)	a) 2.84 ± 1.54 nM b) 346 ng/dl (5.3 nM) c) Baseline: Maternal: 4.1 ± 1.1 µg/l (6.3 ± 1.7 nM) Baseline Fetal: 3.2 ± 1.3 µg/l Treated: Maternal: 8.5 ± 1.7 µg/l Treated: Fetal: 5.6 ± 1.7 µg/l	a) Celi et al. 2010 b) Jonklaas et al. 2015 c) Dussault et al. 1969

Non-teratogenic compound	Common therapeutic doses and dosing schemes in humans	Clinical study design and other resources of information	Plasma peak or blood concentrations	Reference
Magnesium (chloride)	As dietary supplement / RDA: Up to 400 mg (16.5 mmol) daily for pregnant women Constipation & Dyspepsia: Up to 4.8 g daily over max. 7 days (as MgOH) Hypomagnesemia: Up to 5 g IV in 3 hours (as $MgSO_4$) Pre-eclampsia: Max. daily dose: 40 g IV (as $MgSO_4$)	a) 4 g loading dose + 2 g/h maintenance dose IV (n=111, pre-eclampsia treatment of pregnant women b) MgOH 3 x 360 mg PO (45 mmol Mg) (n=10, healthy men) $MgSO_4$ 2 g IV (8 mmol Mg) (n=10, healthy men) c) 12 and 24 mmol Mg PO (n=8, healthy men) 4 and 8 mmol Mg IV (n=8, healthy men)	a) 7.2 mg/dl (3 mM) steady-state (Baseline: 2.0 mg/dl (0.82 mM)) b) MgOH: 0.95 mM $MgSO_4$: 1.2 mM c) Baseline: 0.78 - 0.83 mM 12 mmol PO: 3.1 % increase 24 mmol PO: 4.6 % increase 4 mmol IV: 9.5 % increase 8 mmol IV: 16.1 % increase	a) Brookfield et al. 2016 b) Dolberg et al. 2017 c) Wilimzig et al. 1996
Methicillin	Adults: 1 g IM every 4 h Infants: 0.5 g IM every 4 h (Simon and Rantz 1962)	a) 2 g IV once, (n=5, male and female adults) b) 500 mg IV, pregnant women (n=105)	a) Immediately after administration: ~ 140 µM 2 h after injection: 13-20 µg/ml (34 - 53 µM) b) 30 min after injection: 10.5 ± 4.3 µg/ml 2 h after injection: 3.2 ± 1.9 µg/ml	a) Sørensen et al. 1982 b) Depp et al. 1970
Ranitidine	150 - 300 mg PO daily	a) 150 mg b) 37.5 mg PO (Ranitidine HCl) (n=9, children) c) 50 mg IV (n=20, pregnant women) 150 mg PO (n=80, pregnant women) d) 2.4 mg/kg IV bolus (n=27, infants)	a) 360 - 650 ng/ml (1.14 -2.07 µM) b) 53.9 to 492 ng/ml (0.17 - 1.56 µM), mean 117 ng/ml (0.37 µM) c) Cord blood: 50 mg IV: At delivery approx. 127 ng/ml 12 h after delivery 10-38 ng/ml 150 mg PO: approx. max. 200 ng/ml d) 1520 ng/ml	a) Richards 1983) b) Blumer et al. 1985 c) McAuley et al. 1983 d) Fontana et al. 1993
Retinol	RDA: 800 µg	a) Baseline b) Cord blood measurements (n=145) c) N/A d) Estimated intake: 0.5 - 1.5 mg/day (n > 1000 men)	a) 0.2 - 0.3 µg/ml (0.67 - 1.0 µM) b) 0.7 - 1.3 µM c) 0.3 - 0.7 µg/ml (1.0 - 2.3 µM) d) approx. 1.7 - 3.0 µM	a) Sklan et al. 1985 b) Manolescu et al. 2010 c) Repetto and Repetto 1998 d) Michaëlsson et al. 2003
Sucralose	RDA: 5 mg/kg/day Max. dose: 15 mg/kg (SCF 2000)	a) 1 mg/kg PO (adult men) b) 5-10 mg/kg/day (adults) c) 250 mg PO (adult men)	a) 140 - 460 ng/ml (350 - 1157 nM) b) 5 mg/kg: 100 - 190 ng/ml; 10 mg/kg: max. 900 ng/ml (2.26 µM) c) 1560 ng/ml (3.9 µM)	a) Roberts et al. 2000 b) Baird et al. 2000 c) Sylvetsky et al. 2017

3.4.2 Therapeutic doses & plasma/blood concentrations of teratogenic compounds

Table 23: Teratogenic compounds and their human relevant plasma or blood concentrations at therapeutic doses. Common therapeutic doses and dosing schemes were collected from www.drugs.com if not stated otherwise. Routes of application were as follows: PO= Peroral, IV=Intravenously, IM=Intramuscular or subcutaneously. RDA = Recommended daily allowance. RDA = Recommended daily allowance. EPA = Environmental protection agency

Teratogenic compound	Common therapeutic doses and dosing schemes in humans	Clinical study design and other resources of information	Plasma (peak) or blood concentrations	Reference
9-cis-Retinoic acid (Alitretinoin)	Topical application of 0.1 % alitretinoin containing gel up to four times daily (Panretin®)	Single oral dose of a) 40 mg (men) b) 40 mg daily + contraceptives (women)	a) 146 ng/ml - 272 ng/ml (0.49 - 0.91 µM) b) 143 ng/ml	a-b) Schmitt-Hoffmann et al. 2012
Acitretin	25 to 50 mg PO once a day	Single dose of a) 50 mg b) 50 mg (men) c) 40 mg	a) 500 ng/ml b) 400 ng/ml c) 70 - 420 ng/ml	a) Wiegand and Chou 1998 b) Brindley 1989 c) Heath et al. 2018
Actinomycin D	45 µg/kg or 1250 µg/m2 IV over 10 to 15 minutes once every 3 weeks	a) 0.2 -2 mg IV once (children) b) 0.7-1.5 mg/m² IV bolus (n=31, children and adolescents)	a) 5-186 ng/ml (4 - 148 nM) b) 3-100 ng/ml (2.4 - 80 nM)	a) Hill et al. 2014 b) Veal et al. 2005
Atorvastatin	10-80 mg PO daily	a) 80 mg PO once daily over 14 days (n=8, healthy adults) b) 120 mg PO once (n=1) (Upper dose-limit) c) In rats: 225 mg/kg daily	a) 252 ng/ml (0.45 µM) b) 300 ng/ml (0.54 µM) c) 5.8 µg/ml (10.44 µM)	a) Cilla et al. 1996 b) Posvar et al. 1996 c) Henck et al. 1998
Carba-mazepine	200 - 1600 mg/day	a) 400 - 1000 mg once PO (pregnant women) b) Up to 8 mg/kg (pregnant women)	a) Fetal cord blood: Up to 0.5 - 4.4 µg/ml b) Fetal cord blood: Up to 4.5 µg/ml	a) Nau et al. 1982 b) Pynnönen et al. 1977
Doxorubicin	Up to 75 mg/m² IV over 10 min each 21 days, 4 cycles in total; Lifetime cumulative dosis max. 550 mg/m²	a) 60 mg/m² IV over 40 min (female adults, adjuvant therapy with cyclophosphamide) b) 4-day infusion 75 mg/m² low-dose 25 mg/m² bolus IV high-dose 75 mg/m² bolus IV c) 75 mg/m² IV oder 15 min	a) 0.63 µg/ml b) 4-day IV: 0.1 µg/ml Low-dose: 2.3 µg/ml High-dose: 3.7 µg/ml c) 0.1 µM (0.05 µg/ml)	a) Barpe et al. 2010 b) Twelves et al. 1991 c) Greene et al. 1983
Entinostat (MS-275)	n/a	a) In vitro benchmark concentration b) Max.8 mg/week	a) 208 nM b) 5 ng/ml mean (0.12 µM), max.150 ng/ml (0,36 µM)	a) Rempel et al. 2015 b) Connolly et al. 2017

Teratogenic compound	Common therapeutic doses and dosing schemes in humans	Clinical study design and other resources of information	Plasma (peak) or blood concentrations	Reference
Favipiravir	1600 mg PO twice on first day, 600 mg PO twice following days (Pharmaceutical and Food Safety Bureau 2014)	a) 1st study (JP101): 30 - 1600 mg PO once (n=36, men) 2nd study (JP103): 400-600 mg PO, multiple dosing (n=18, men) 3rd study (JP111): 1200 single dose followed by 400-600 mg, multiple dosing (n=12, men) b) 1800 mg twice on first day, followed by 800 mg twice/day (adults) c) 600 mg PO twice/day (adults)	a) 1st study: 1.39 - 78 .6 µg/ml 2nd study: 17.2 - 43.8 µg/ml 3rd study: 40.5 -61.5 µg/ml b) 35-50 µg/ml c) Steady-state: 61.5 µg/ml	a) Pharmaceutical and Food Safety Bureau 2014 b) Hayden and Shindo 2019 c) Madelain et al. 2016
Isotretinoin	0.25 to 0.5 mg/kg PO twice a day Maximum dose: Up to 2 mg/kg/day	a) 80 mg oral dose b) 100 mg oral dose in man	a) 300 ng/ml (1 µM) b) 74 to 511 ng/ml (1.7 µM)	a) Wiegand and Chou 1998 b) Colburn et al. 1983
Leflunomide	20 mg PO daily	5-100 mg PO, different dosing regimens (adults)	Steady-State: 4 -152 µg/ml	Rozman 2002
Lithium (chloride)	900-1800 mg PO daily	a) Recommended therapeutic plasma levels based on multiple studies b) 600 mg c) 1000 mg	a) 0.6-0,75 mM b) 0.6 mM Therapeutic 0.4 -1.1 mM, Mid toxic 1.5 -2 mM, Toxic >2 mM c) 0.4 – 0.8 mM	a) Severus et al. 2008 b) Apotex Inc. 2013 c) Wesseloo et al. 2017
Methotrexate	Acute Lymphoblastic Leukemia: -Induction: 3.3 mg/m2/day orally for 4 to 6 weeks -Maintenance dose during remission: 30 mg/m2 orally or IM twice a week -Alternate maintenance dose during remission: 2.5 mg/kg IV every 14 days Psoriasis: 10 to 30 mg/week PO, IM, IV or subcutaneously Rheumatoid Arthritis: 7.5 -20 mg PO or subcutaneously once a week Osteosarcoma: 12-15 g/m2 IV as a 4-hour infusion	2.1 - 36 mg/m2 PO (n= 297, chilrden)	0.1 - 4.3 µM, mean 1.1 µM	Balis et al. 1998

Teratogenic compound	Common therapeutic doses and dosing schemes in humans	Clinical study design and other resources of information	Plasma (peak) or blood concentrations	Reference
Methylmercury	RDA: 0.1 µg/kg/day	a) Baseline concentration in Korean Population b) RDA intake and EPA reference dose corresponding concentration c) In vitro benchmark concentration	a) 4.5 µg/l (17,9 nM) b) 5.8 µg/l c) 1.5 µM	a) Kim et al. 2012 b) Mortensen et al. 2014 c) Rempel et al. 2015
Panobinostat	20 mg PO every other day (3 doses per week)	a) In vitro benchmark concentration b) 20 mg once	a) 4 nM b) 20 ng/ml mean (57 nM)	a) Rempel et al. 2015 b) van Veggel et al. 2018
Paroxetine	20-60 mg PO daily	a) 20-50 mg PO single dose (n=29) b) 30 mg/day PO over 30 days (n=15) c) 40 mg PO daily, chronic dosing (n=30)	a) 2.5-65.1 ng/ml (7- 198 nM) b) 8.6-105 ng/ml (26- 319 nM) c) 1.7-407.0 ng/ml (7.9- 1237 nM)	a-c) Kaye et al. 1989
Phenytoin (5,5-Diphenyl-hydantoin)	300 mg PO daily Status epilepticus: max. 50 mg/min IV in 4 h	a) 100-700 mg IV (pregnant women) b) 1.7 - 6.0 mg/kg (n=13, pregnant women)	a) Fetal cord blood: 0.4-13 µg/ml b) Maternal blood = cord blood at delivery: 2.2 - 20 µg/ml	a) Nau et al. 1982 b) Koch et al. 1996
Teriflunomide (A-771726)	7-14 mg PO daily	refer to leflunomide	refer to leflunomide	refer to leflunomide
Thalidomide	100-400 mg PO daily	200 mg PO (n =45, adult men)	1-2 mg/l (3,9 µM)	Teo et al. 2004
Trichostatin A	n/a	In vitro benchmark concentration	10 nM	Rempel et al. 2015
Valproic acid (VPA)	Maximum dose: 60 mg/kg daily Therapeutic plasma levels: 50 - 100 µg/ml (347 - 693 µM)	50-1350 mg / day (pregnant women)	Maternal blood: 15.6 - 85 µg/ml Fetal blood: 15.8 - 90 µg/ml	Nau et al. 1982
Vinblastine	up to 18.5 mg/m2 IV weekly	a) 7.5 mg/m² IV once b) 7-11 mg/m² IV once c) 3 mg/m² (coinjection cis-Platin) IV once d) Mice: 12-35 mg/kg IV once e) Mice	a) 10 -20 ng/ml (12,3 - 24,6 nM) b) 20-40 ng/ml (24,6 - 59,2 nM) c) 4,8 ng/ml (5,9 nM) d) 10 ng/ml (12,3 nM) e) 10 ng/ml (effective teratogenic fetal blood concentration in mice)	a) Lu et al. 1983 b) Nelson et al. 1980 c) Links et al. 1999 d) van Tellingen et al. 1993 e) van Calsteren et al. 2009
Vismodegib	150 mg PO daily	a) 150 mg PO once (Adults with hepatic disfunctions) b) 150 mg PO once or multiple dosing (healthy females)	a) 17 - 27 µM b) Single dose: 6 µM Multiple doses (steady-state): 16.4 µM	a) Abou-Alfa et al. 2017 b) Graham et al. 2012
Vorinostat	400 mg PO daily	a) 400 mg PO once daily over 2 to 480+ days b) In vitro benchmark concentration	a) 0.4-3 µM b) 140 nM	a) Iwamoto et al. 2013 b) Rempel et al. 2015

3.5 Translation of in vivo concentrations into in vitro concentrations

To apply biologically relevant concentrations to the UKN1 6-day test system which represent the human situation, the researched therapeutic, in vivo plasma and blood concentrations described in the previous chapter 3.4 had to be translated into meaningful in vitro concentrations for all 39 compounds.
Basically, for every substance two concentrations were calculated: The 1x and the 20x C_{max} (**Table 24**). The 1x C_{max} usually represented the highest plasma or blood concentration that was found in clinical studies after an application of a substance's therapeutic dose. If such information were not available, other resources of information that provided credible information about compound concentrations were considered (**Table 22** & **Table 23**). However, as the in vivo concentrations of each compound in many cases varied over a wide range which depended on many factors like the exact dose or interindividuality, the 1x C_{max} was fitted to all available data of the corresponding compound as best as possible to get the most reasonable in vitro concentration. Thus, the 1x C_{max} was either as high as the highest observed concentration in vivo or a kind of mean-/median-value of the existing data. In case of ascorbic and folic acid, the two- to four-fold accumulation of the compounds in the fetus was also taken into consideration (**Table 19**). The 20x C_{max} had to clearly exceed the highest found concentration in literature and would fulfill two distinct functions: If neither the 1x nor the 20x C_{max} of a non-teratogenic substance affected differentiating cells, the substance would be clearly non-teratogenic and underline the robust specificity of the system. On the other hand, the 20x C_{max} of a teratogenic substance would provide a "second chance" to induce effects at a higher concentration, if the 1x C_{max} did not so.
For all 39 substances, the in vivo concentrations could be translated into biologically relevant 1x C_{max} in vitro concentrations, and for the majority of substances also a 20x C_{max} could be applied.

Table 24: 1x and 20x C_{max} of non-teratogenic and teratogenic compounds in the UKN1 6-day test system. *Due to limited solubility, carbamazepine was tested at 10x instead of 20x, and leflunomide, phenytoin, teriflunomide and vismodegib were tested at 1x only. **Due to the known cytotoxicity of high VPA concentrations from previous experiments, VPA was tested at 1000 µM as the highest concentration.

Non-teratogenic compound	Absolute concentration [µM]		Teratogenic compound	Absolute concentration [µM]	
	1x C_{max}	20x C_{max}		1x C_{max}	20x C_{max}
Ampicillin	107	2140	9-cis-Retinoic acid	1	20
Ascorbic acid	200	4000	Acitretin	1.2	24
Buspirone	0.0244	0.488	Actinomycin D	0.1	2
Chlorpheniramine	0.0304	0.608	Atorvastatin	0.54	10.8
Dextromethorphan	0.15	3	Carbamazepine*	19	10x: 190
Diphenhydramine	0.3	6	Doxorubicin	1.84	36.8
Doxylamine	0.38	7.6	Entinostat	0.2	4
Famotidine	1.06	21.2	Favipiravir	382	7600
Folic acid	0.38	7.6	Isotretinoin	1.7	34
Levothyroxine	0.077	1.54	Leflunomide*	370	--
Liothyronine	0.00307	0.06145	Lithium chloride	1000	20000
Magnesium chloride	1200	24000	Methotrexate	1	20
Methicillin	140	2800	Methylmercury	0.020	0.4
Ranitidine	0.8	16	Panobinostat	0.06	1.2
Retinol	1	20	Paroxetine	1.2	24
Sucralose	2.5	50	Phenytoin*	20	--
			Teriflunomide*	370	---

Non-teratogenic compound	Absolute concentration [µM] 1x C_{max}	20x C_{max}

Teratogenic compound	Absolute concentration [µM] 1x C_{max}	20x C_{max}
Thalidomide	3.9	78
Trichostatin A	0.01	0.2
Valproic acid**	600	1.67x: 1000
Vinblastine	0.0247	0.494
Vismodegib*	20	--
Vorinostat	3	60

3.6 Cytotoxic effects of teratogenic compounds

The non-teratogenic and teratogenic compounds described in the preceding chapters all had the potential to induce cytotoxicity to the cells of the UKN1 6-day test. The documentation of the cytotoxicity was a first important step to evaluate the developmental toxicitiy potential of the compounds. In order to assess the cytotoxicity of compounds, the cytotoxic effect was estimated qualitatively by cell culture pictures and reduced RNA-yields.
Compounds which demonstrated a clear reduction of viable cells or low RNA-yields compared to untreated control cells were called “Strong” cytotoxic. If the effect was less obvious, but still distinguishable from the untreated control, the compound was called “Little” cytotoxic. If no effect on cells could be observed at all, the compound had “No effect”. For non-teratogenic compounds, no cytotoxicity was observed and only teratogens showed cytotoxic effects (**Table 25**). Most of them were not cytotoxic at the 1x C_{max} but became little or strong cytotoxic at the 20x C_{max}, and six compounds were strong cytotoxic at both concentrations.
Under the line, non-teratogens were not cytotoxic, but some of teratogens, especially at the 20x C_{max}.

Table 25: Cytotoxicity of teratogenic compounds. A strong effect means a cytotoxicity of almost 100 % whereas a little effect describes all conditions in which at least some reduction of cell viability could be observed. *Carbamazepine and valproic acid were applied at 10x and 1.67x C_{max}, respectively, instead of 20x C_{max}. Leflunomide, phenytoin, teriflunomide and vismodegib were only applied at 1x C_{max}.

	Cytotoxic effect	
C_{max} / Compound	**1x**	**20x**
9-cis-Retinoic acid	No effect	No effect
Acitretin	No effect	No effect
Actinomycin D	Strong	Strong
Atorvastatin	Strong	Strong
Carbamazepin	No effect	10x: Little*
Doxorubicin	Strong	Strong
Entinostat	No effect	Strong
Favipiravir	No effect	Little
Isotretinoin	No effect	Strong
Leflunomide	No effect	-----*
Lithium	No effect	Strong
Methotrexate	No effect	Little
Methylmercury	No effect	No effect
Panobinostat	Strong	Strong
Paroxetine	No effect	Strong
Phenytoin	No effect	-----*
Teriflunomide	Little	-----*
Thalidomide	No effect	No effect
Trichostatin A	No effect	Strong
Valproic acid	No effect	1.67x: Little*
Vinblastine	Strong	Strong
Vismodegib	No effect	-----*
Vorinostat	Strong	Strong

3.7 Transcriptomic responses to substances-exposure

In order to figure out if teratogenic and non-teratogenic compounds had an effect on the transcriptome of cells and if they could be distinguished by their gene expression profiles, the RNA of substance-exposed cells of all non- and little cytotoxic conditions in the UKN1 6-day assay was analyzed with Affymetrix gene arrays on differential gene expression.

3.7.1 Principal component analysis

To gain a first and quick overview of the substance-induced effects on gene expression, a principal component analysis (PCA) of the gene expression data was performed (**Figure 10**). In the PCA, the most meaningful parts of the gene expression data from the gene arrays were merged into two main components and the individual compounds and conditions were positioned along these two main components on the basis of the gene expression changes they induced in NEPs. Thus, the geometric distances between the compounds correlated with their genetic differences. In addition, the percentages stated at the main components specified, which proportion of the genetic difference was explained by the particular component with respect to the total genetic variance between the compounds and conditions in the given data set.
In the shown PCAs, two data sets were analyzed: The first PCA was based on the expression of all 54,675 analyzed probe sets, the second PCA only on the 100 probe sets with the highest variance across all samples. Although some differences could be observed in the positioning of individual substances between the two plots, the overall distribution and relative distances of substances to each other were similar. The majority of non-teratogenic substances formed a main cluster near the zero point of both axis which also included many teratogenic substances at the 1x C_{max}. In the Top-100-PCA, but not in the PCA of all probe sets, thalidomide at the 20x C_{max} was also included in the cluster. Laterally to the cluster, levothyroxine and liothyronine at the 20x C_{max} were positioned. A little bit distant from the cluster, ascorbic acid at the 20x C_{max} was found in both PCAs, and in the Top-100-PCA, it formed a second small sub-cluster with both methotrexate samples. Around the half of teratogenic samples were positioned in a considerable distance from the main cluster, especially the retinoids 9-cis-retinoic acid, acitretin and retinol (the latter only at the 20x C_{max}). Moreover, the high concentrations of all teratogenic substances were found in a higher distance from the main cluster than the low concentrations. Interestingly, although the relative positioning of the substances was similar in both PCAs, the main components 1 and 2 explained together 84 % of the variances between the substances in the Top-100-PCA, whereas in the PCA of all 54,675 probe sets they could only explain 33 % of the variances. This indicated that the major genetic differences were caused by only a few genes and that it could be beneficial to consider only a well-chosen subset of probe sets instead the entirety of the gene expression data for the distinction of teratogens and non-teratogens.
So, the PCA was already able to show differences between the gene expression data of substance-exposed cells, especially between teratogens and non-teratogens, and that a small subset of genes was responsible for the major differences.

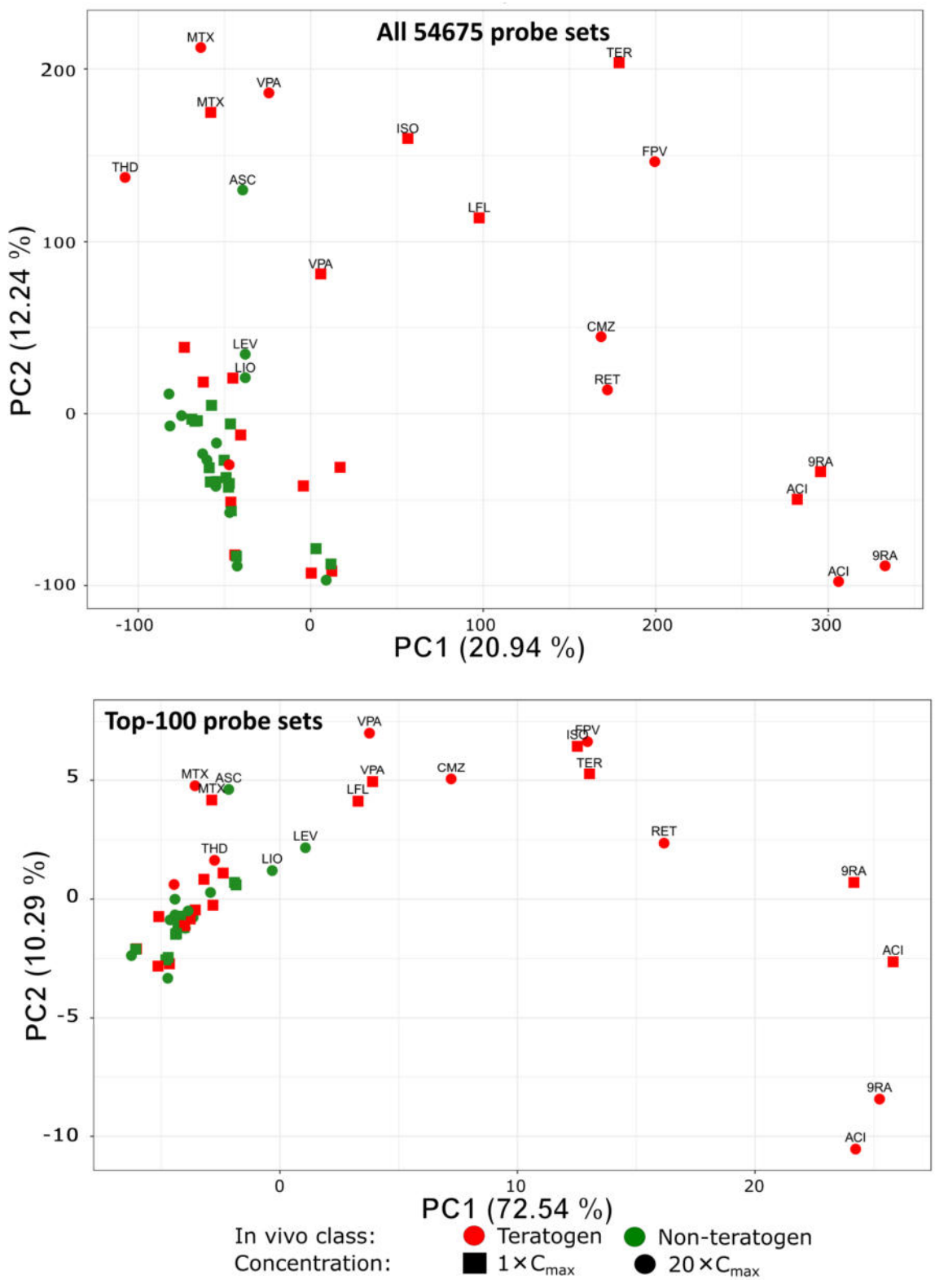


***Figure 10**: Principal component analyses of all substances at low and high concentrations with respect of their gene expression profiles of all 54675 probe sets and of the 100 probe sets with the highest variance across all samples. Green and red tags represent in vivo non-teratogenic and teratogenic substances, respectively, and mean-values of n=3-4 biological replicates. Low concentrations (1x C_{max}) were indicated by squares and high concentrations (1.67x, 10x and 20x C_{max}) by circles. The distribution of the data points on the x-axis is given by the principal component (PC) 1 and on the y-axis by PC2. Abbreviations: 9RA = 9-cis-Retinoic acid, ACI = Acitretin, ASC = Ascorbic acid, CMZ = Carbamazepine, FPV = Favipiravir, ISO = Isotretinoin, LEV = Levothyroxine, LFL = Leflunomide, LIO = Liothyronine, MTX = Methotrexate, RET = Retinol (at 20x C_{max}), TER = Teriflunomide, THD = Thalidomide, VPA = Valproic acid*

3.7.2 Significantly deregulated genes in NEPs exposed to non-teratogenic compounds

In order to find and filter for significant alterations within all 54,675 probe sets of the gene arrays, the substance-induced gene expression changes in NEPs of all samples were systematically analyzed. For that purpose, the data were in a first step processed by several algorithm-based statistical multiarray methods and plotted in V-shaped volcano plots (**Figure 11** - **Figure 18**). A black dot represented a single probe set, and its substance-induced fold-change compared to unexposed-NEPs was given on the x-axis in a logarithmic form. On the y-axis, the corresponding "limma"-p-value was given, also in a logarithmic form. Probe sets, which demonstrated an FDR-adjusted p-value <0.05 and simultaneously a deregulation of more than two-fold up or ½-down were considered to be both statistically significant and biologically relevant and therefore highlighted in red.
For all non-teratogenic conditions, a genome-wide gene expression analysis was conducted. In NEPs that were exposed to non-teratogenic compounds a lot of probe sets were deregulated. The majority of them ranged between a $\log_2$(FC) of 0 and 0.5 and many were also deregulated up to a $\log_2$(FC) of 1. Only few exceeded a $\log_2$(FC) of 1 except for ascorbic acid and retinol at the 20x C_{max} where a lot of probe sets were deregulated up to a $\log_2$(FC) or 2 or even above. Limma-p-values < 0.05 could be observed for many deregulated probe sets in all conditions. When looking for significant probe sets with an FDR-adjusted p-value <0.05 and an absolute fold-change > 2, most conditions did not show a significant gene expression change and the number of significantly up- or downregulated probe sets was zero (**Figure 11** - **Figure 14**). Small substance effects could be observed for levothyroxine at the 1x C_{max} and liothyronine, $MgCl_2$ and methicillin at the 20x C_{max}, where the numbers of up-/downregulated probe sets were $18_{up}/20_{down}$, $27_{up}/57_{down}$, $13_{up}/3_{down}$ and $0_{up}/2_{down}$, respectively. Strong substance effects could only be noted for retinol at the 20x C_{max}, were 1032 up- and 936 downregulated probe sets were found. A condensed overview of all significantly deregulated probe sets is given in **Table 26**. Overall, non-teratogenic substances did not induce significant gene expression changes in NEPs except for a few alterations in some samples and retinol at the 20x C_{max}, where almost 2000 probe sets were significantly deregulated.

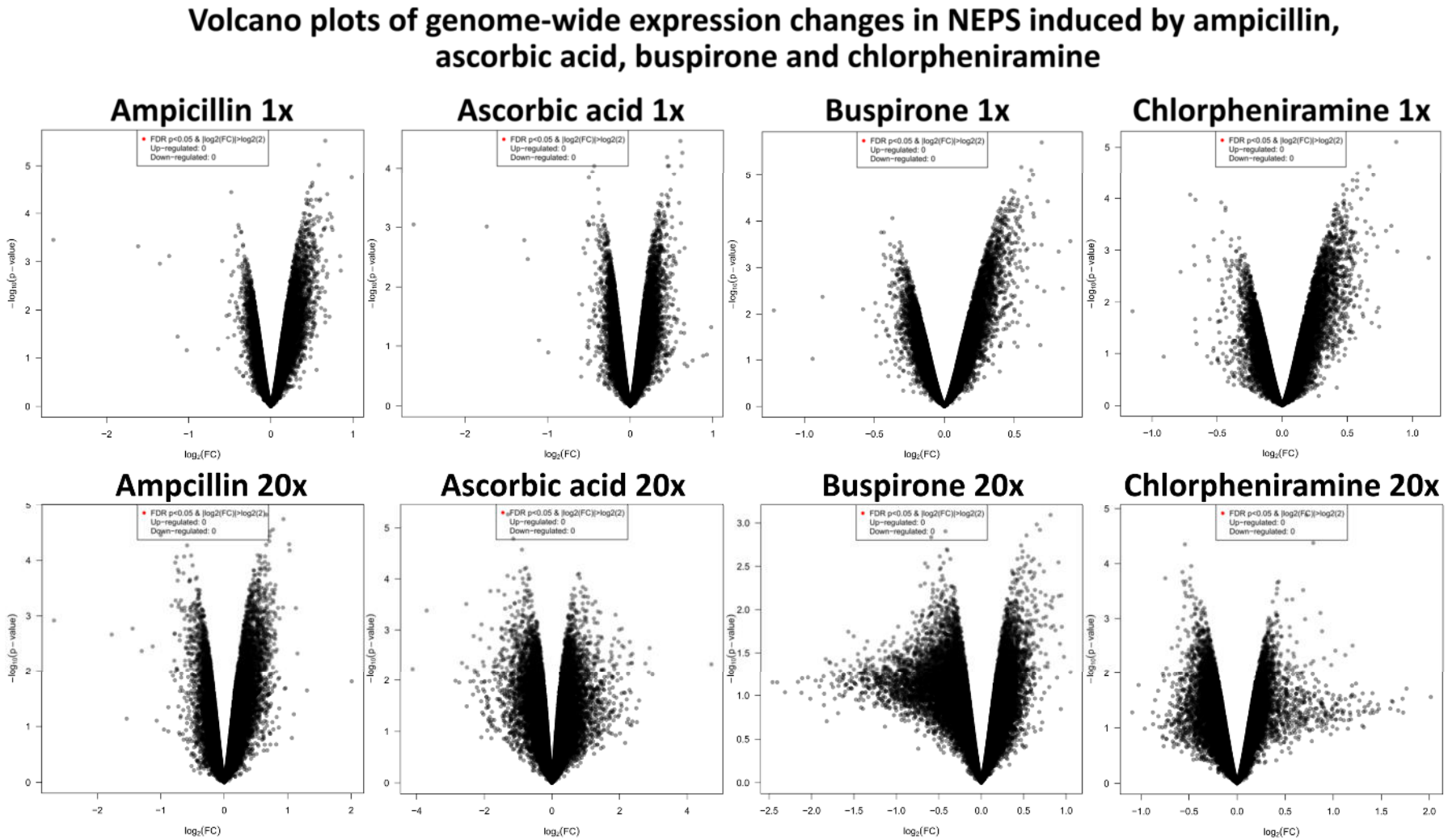


Figure 11: *Genome-wide gene expression changes in NEPs caused by ampicillin, ascorbic acid, buspirone and chlorpheniramine at the 1x and 20x C_{max}. Imaged is the fold-change on the x-axis in its logarithmic form ($log_2(FC)$) and the corresponding, not FDR-adjusted "limma"-p-value on the y-axis in its negative, logarithmic form ($-log_{10}$(p-value)) of each probe set-mean-value from n=3-4 independent biological experiments. Highlighted in red are all probe sets, which had an FDR-adjusted p-value of <0.05 and were at least two-fold up or ½-fold down regulated.*

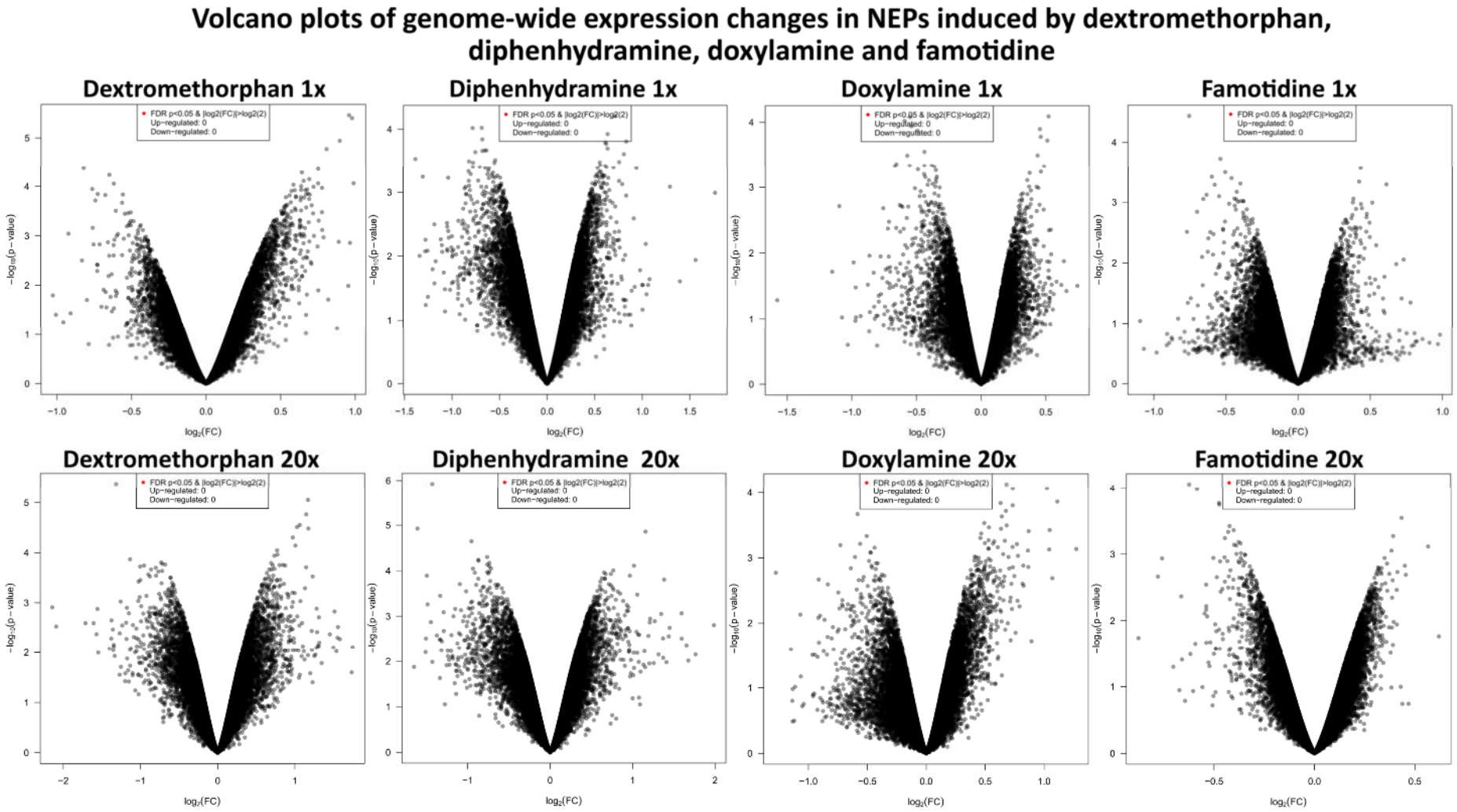


***Figure 12**: Genome-wide gene expression changes in NEPs caused by dextromethorphan, diphenhydramine, doxylamine and famotidine at the 1x and 20x C_{max}. Imaged is the fold-change on the x-axis in its logarithmic form ($log_2(FC)$) and the corresponding, not FDR-adjusted "limma"-p-value on the y-axis in its negative, logarithmic form ($-log_{10}$(p-value)) of each probe set-mean-value from n=3-4 independent biological experiments. Highlighted in red are all probe sets, which had an FDR-adjusted p-value of <0.05 and were at least two-fold up or ½-fold down regulated.*

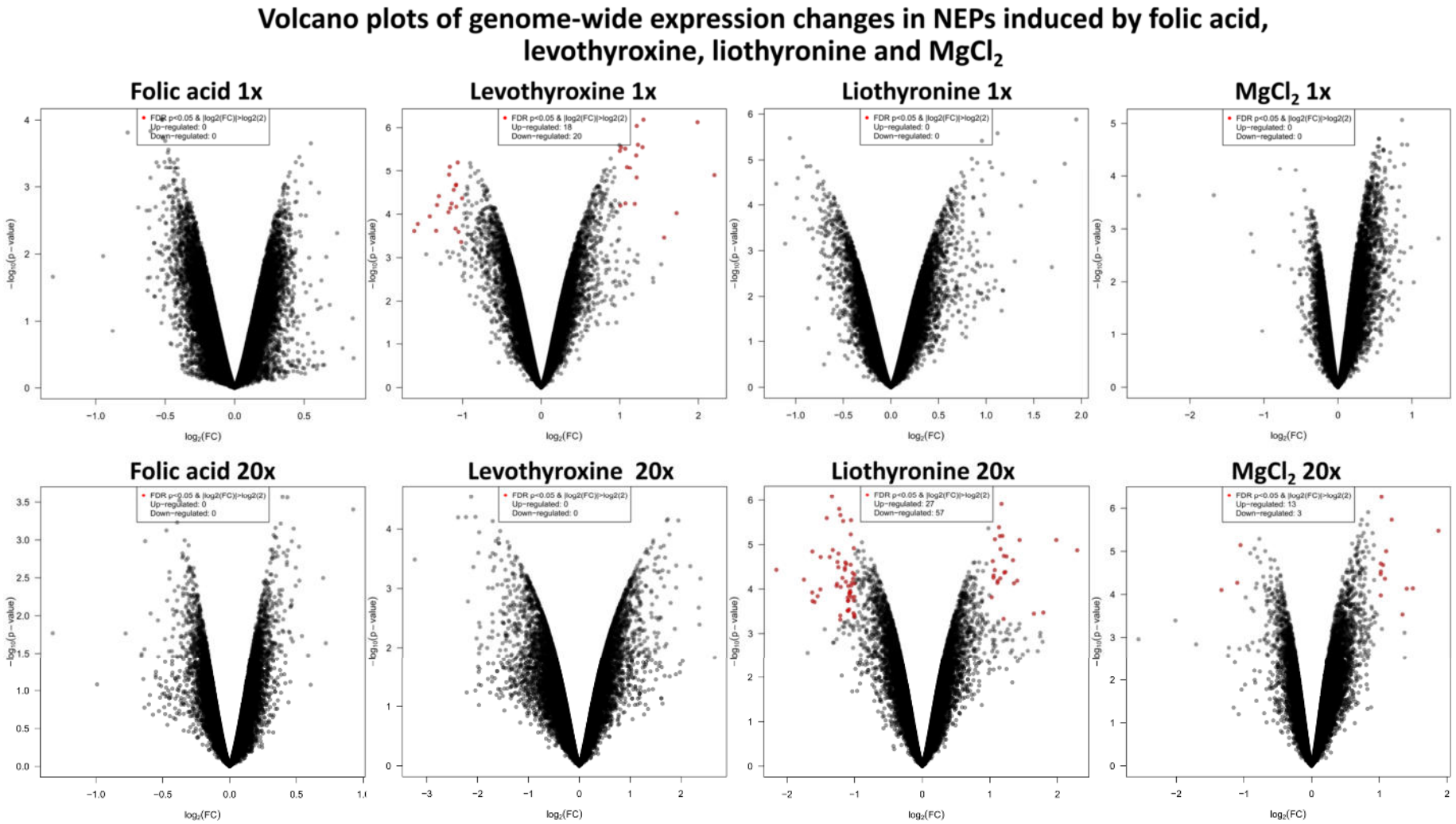


Figure 13: *Genome-wide gene expression changes in NEPs caused by folic acid, levothyroxine, liothyronine and $MgCl_2$ at the 1x and 20x C_{max}. Imaged is the fold-change on the x-axis in its logarithmic form ($log_2(FC)$) and the corresponding, not FDR-adjusted "limma"-p-value on the y-axis in its negative, logarithmic form ($-log_{10}$(p-value)) of each probe set-mean-value from n=3-4 independent biological experiments. Highlighted in red are all probe sets, which had an FDR-adjusted p-value of <0.05 and were at least two-fold up or ½-fold down regulated.*

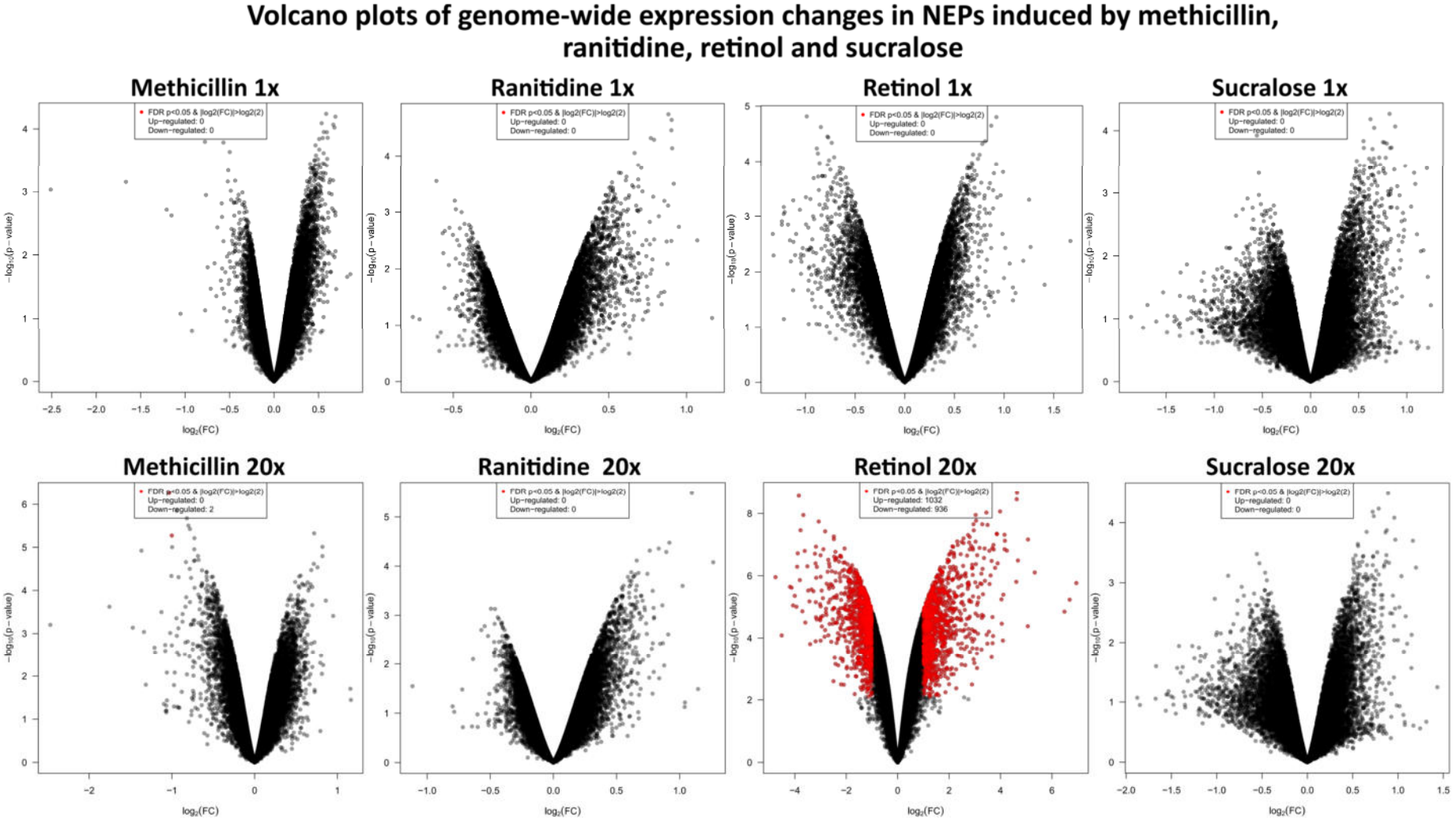


***Figure 14**: Genome-wide gene expression changes in NEPs caused by methicillin, ranitidine, retinol and sucralose at the 1x and 20x C_{max}. Imaged is the fold-change on the x-axis in its logarithmic form (log_2(FC)) and the corresponding, not FDR-adjusted "limma"-p-value on the y-axis in its negative, logarithmic form (-log_{10}(p-value)) of each probe set-mean-value from n=3-4 independent biological experiments. Highlighted in red are all probe sets, which had an FDR-adjusted p-value of <0.05 and were at least two-fold up or ½-fold down regulated.*

3.7.3 Significantly deregulated genes in NEPs exposed to teratogenic compounds

To find significant gene expression alterations in teratogen-exposed NEPs, the gene array data were analyzed and plotted in the same way as for the non-teratogens in the previous chapter.
For all teratogenic conditions except the ones in which a strong cytotoxic effect was observed, a genome-wide gene expression analysis was conducted. Consequently, no measurements could be performed for actinomycin D, atorvastatin, doxorubicin, panobinostat, vinblastine and vorinostat (**Table 25**). Eight teratogenic substances were measured at both the 1x and the 20x C_{max}. Four of them were able to induce a strong transcriptomic response at both concentrations in NEPs: At the 1x C_{max}, treatment with 9-cis-retinoic acid , acitretin, methotrexate and VPA led to a deregulation of $1956_{up}/1785_{down}$, $1803_{up}/1604_{down}$, $435_{up}/209_{down}$ and $630_{up}/364_{down}$ significant probe sets, respectively, and at the 20x C_{max} even to $2426_{up}/1887_{down}$, $2309_{up}/1795_{down}$, $687_{up}/622_{down}$ and $878_{up}/685_{down}$ significantly deregulated probe sets, respectively (**Figure 15**, **Figure 17**, **Figure 18**). Although carbamazepine and favipiravir did not induce significant gene expression changes at the 1x C_{max}, they deregulated $910_{up}/591_{down}$ probe sets at 10x C_{max} and $1551_{up}/1290_{down}$ probe sets at 20x C_{max}, respectively (**Figure 15** and **Figure 16**). Notably, these aforementioned six compounds had a concentration dependent effect on the transcriptome and the higher concentrations deregulated more significant probe sets than the 1x C_{max}. In addition, the number of upregulated probe sets was constantly higher than the number of downregulated ones. Thalidomide at the 1x C_{max} affected only $41_{up}/21_{down}$ probe sets significantly, but zero probe sets at the 20x C_{max}, and MeHg did not significantly affect any probe set at neither the 1x nor the 20x C_{max}. From the remaining, not yet mentioned nine teratogenic substances, only the 1x C_{max} samples were available for transcriptomic analyses, either due to the cytotoxicity of the 20x C_{max} or to limited solubility of the compound at 20x C_{max}: Isotretinoin, leflunomide and teriflunomide induced more than 1000 significant deregulated probe sets each, in particular $1029_{up}/703_{down}$, $614_{up}/415_{down}$ and $1829_{up}/1394_{down}$, respectively (**Figure 16** and **Figure 17**). In contrast, entinostat and TSA could only affect $48_{up}/30_{down}$ and $4_{up}/0_{down}$ probe sets significantly, respectively (**Figure 16** and **Figure 18**), and LiCl, paroxetine, phenytoin and vismodegib had no significant effect on gene expression at all (**Figure 16** - **Figure 18**).
Basically, in conditions which did induce none or only very few significant deregulated probe sets in NEPs, the overall genetic response was comparable to the non-teratogenic compounds. Accordingly, $\log_2$(FC)-values were mostly seen to range between 0 and 0.5 and rarely exceeded 1, but limma-p-values were <0.05 for many probe sets. A condensed overview of all significantly deregulated probe sets is given in **Table 26**.
In total, many teratogens induced a remarkable number of significantly deregulated probe sets, and the ones that did not induce significant changes showed results which were comparable to the non-teratogens.

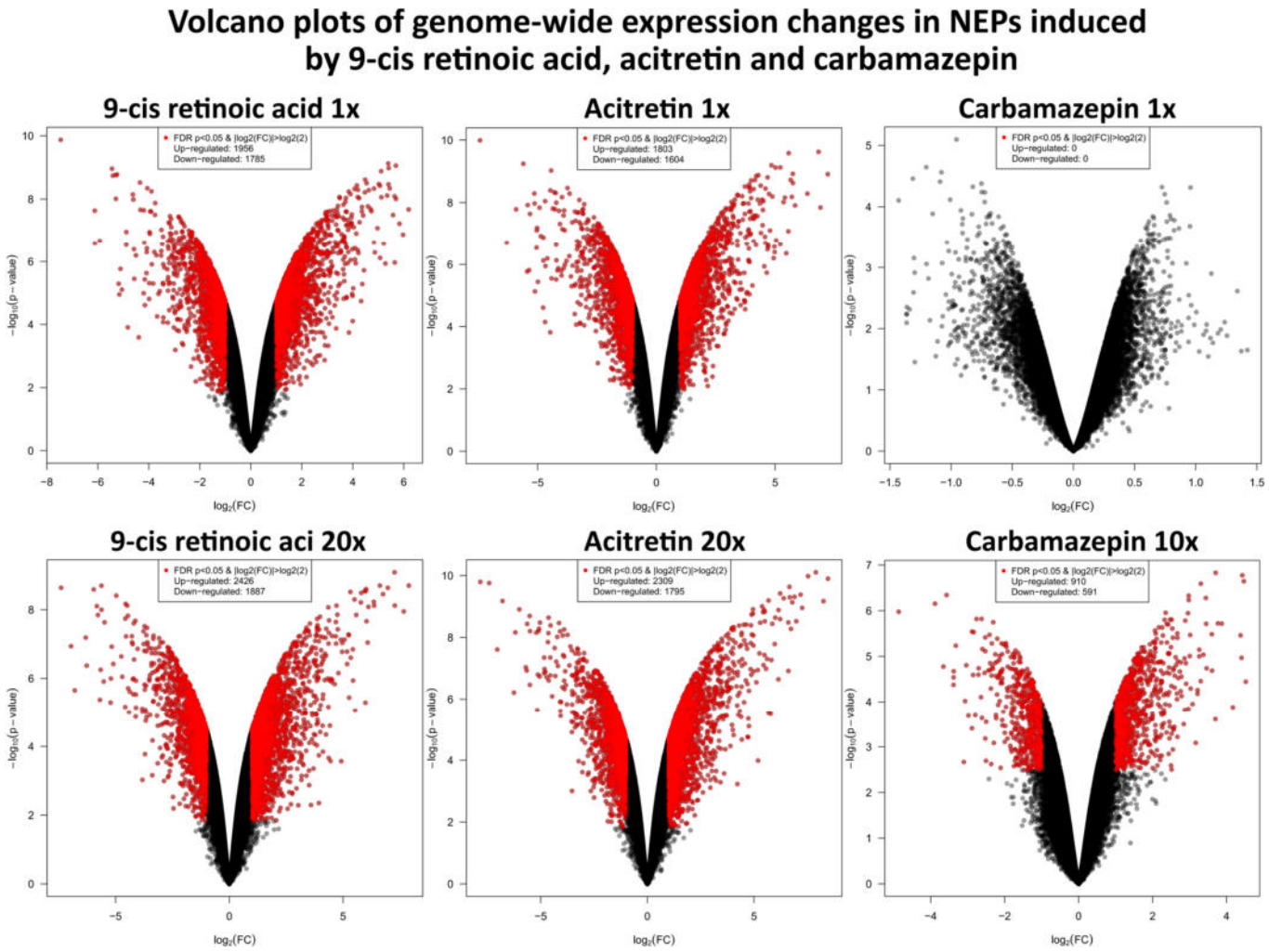


Figure 15: *Genome-wide gene expression changes in NEPs caused by methicillin, ranitidine, retinol and sucralose at the 1x and 20x C_{max}. Imaged is the fold-change on the x-axis in its logarithmic form ($log_2(FC)$) and the corresponding, not FDR-adjusted "limma"-p-value on the y-axis in its negative, logarithmic form ($-log_{10}$(p-value)) of each probe set-mean-value from n=3-4 independent biological experiments. Highlighted in red are all probe sets, which had an FDR-adjusted p-value of <0.05 and were at least two-fold up or ½-fold down regulated.*

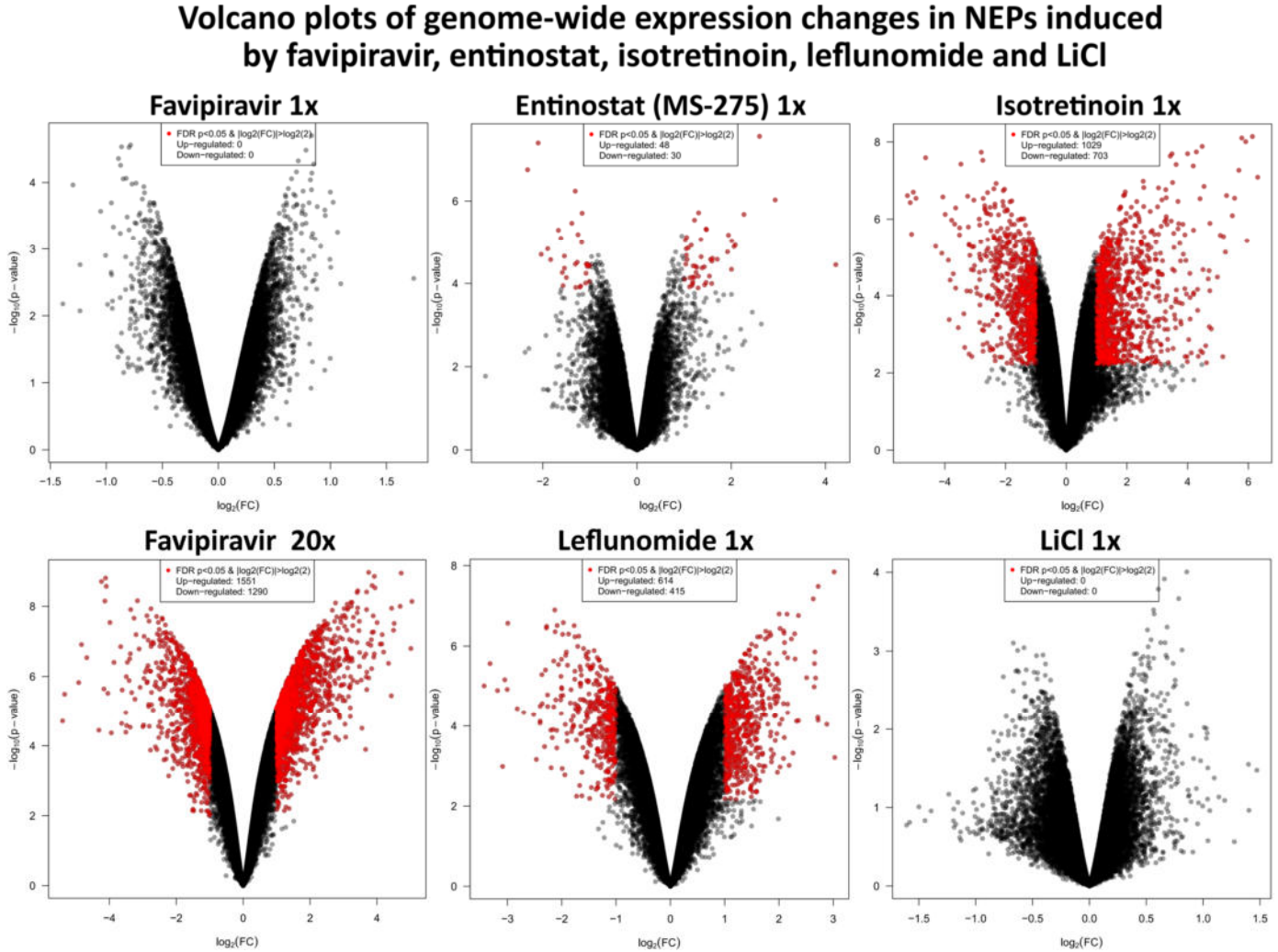


Figure 16: *Genome-wide gene expression changes in NEPs caused by methicillin, ranitidine, retinol and sucralose at the 1x and 20x C_{max}. Imaged is the fold-change on the x-axis in its logarithmic form ($log_2(FC)$) and the corresponding, not FDR-adjusted "limma"-p-value on the y-axis in its negative, logarithmic form ($-log_{10}$(p-value)) of each probe set-mean-value from n=3-4 independent biological experiments. Highlighted in red are all probe sets, which had an FDR-adjusted p-value of <0.05 and were at least two-fold up or ½-fold down regulated.*

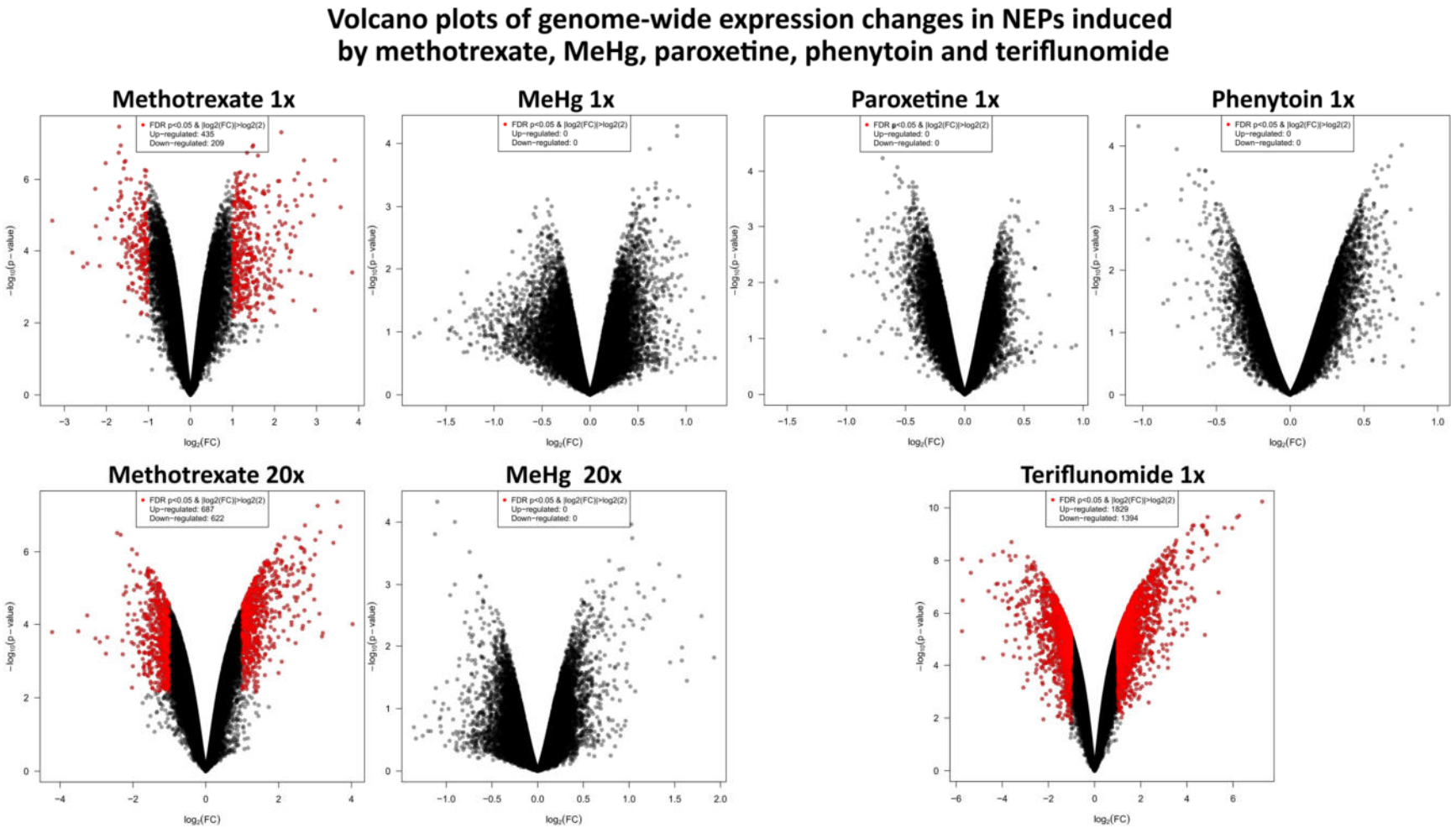


***Figure 17**: Genome-wide gene expression changes in NEPs caused by methicillin, ranitidine, retinol and sucralose at the 1x and 20x C_{max}. Imaged is the fold-change on the x-axis in its logarithmic form ($log_2(FC)$) and the corresponding, not FDR-adjusted "limma"-p-value on the y-axis in its negative, logarithmic form ($-log_{10}(p\text{-value})$) of each probe set-mean-value from n=3-4 independent biological experiments. Highlighted in red are all probe sets, which had an FDR-adjusted p-value of <0.05 and were at least two-fold up or ½-fold down regulated.*

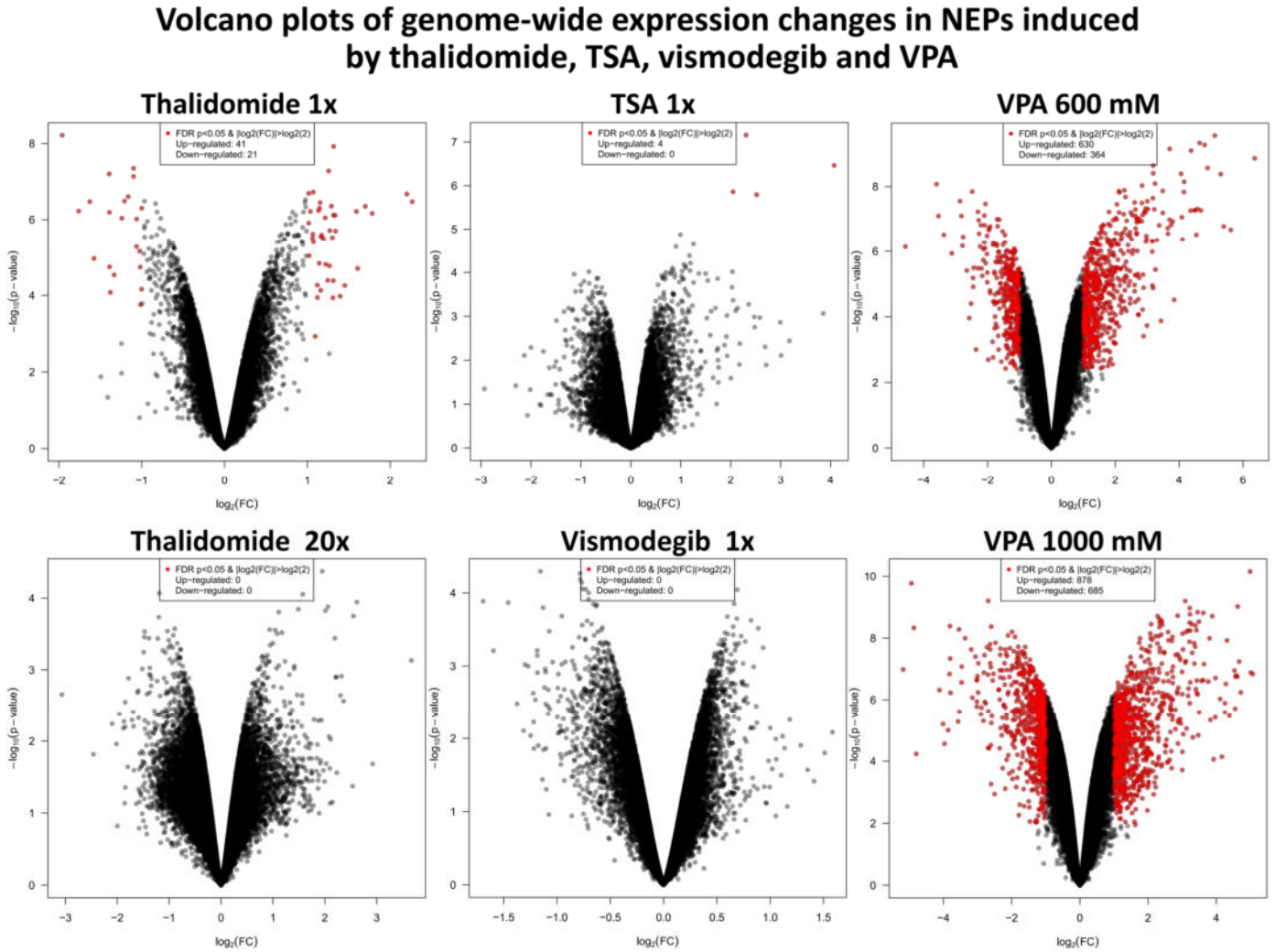


Figure 18: *Genome-wide gene expression changes in NEPs caused by methicillin, ranitidine, retinol and sucralose at the 1x and 20x C_{max}. Imaged is the fold-change on the x-axis in its logarithmic form (log_2(FC)) and the corresponding, not FDR-adjusted "limma"-p-value on the y-axis in its negative, logarithmic form (-log_{10}(p-value)) of each probe set-mean-value from n=3-4 independent biological experiments. Highlighted in red are all probe sets, which had an FDR-adjusted p-value of <0.05 and were at least two-fold up or ½-fold down regulated.*

Table 26: Number of significantly up- and downregulated probe sets with an FDR-adjusted p-value < 0.05 and an absolute fold-change > 2 at 1x and 20x C_{max}. Toxic conditions could not be measured due to an insufficient amount of RNA. *Carbamazepine and valproic acid were tested at 10x and 1.67 C_{max}, respectively, instead of 20x C_{max}. Leflunomide, phenytoin, teriflunomide and vismodegib were only tested at 1x C_{max}.

Non-teratogens	Number of up-/ downregulated probe sets				Teratogens	Number of up-/ downregulated probe sets			
	1x C_{max}		20x C_{max} *			1x C_{max}		20x C_{max} *	
	Up	Down	Up	Down		Up	Down	Up	Down
Ampicillin	0	0	0	0	9-cis-Retinoic acid	1956	1785	2426	1887
Ascorbic acid	0	0	0	0	Acitretin	1803	1604	2309	1795
Buspirone	0	0	0	0	Actinomycin D	toxic	toxic	toxic	toxic
Chlorpheniramine	0	0	0	0	Atorvastatin	toxic	toxic	toxic	toxic
Dextromethorphan	0	0	0	0	Carbamazepine	0	0	910	591
Diphenhydramine	0	0	0	0	Doxorubicin	toxic	toxic	toxic	toxic
Doxylamine	0	0	0	0	Entinostat	48	30	toxic	toxic
Famotidine	0	0	0	0	Favipiravir	0	0	1551	1290
Folic acid	0	0	0	0	Isotretinoin	1029	703	toxic	toxic
Levothyroxine	18	20	0	0	Leflunomide	614	415	--	--
Liothyronine	0	0	27	57	Lithium chloride	0	0	toxic	toxic
Magnesium chloride	0	0	13	3	Methotrexate	435	209	687	622
Methicillin	0	0	0	2	Methylmercury	0	0	0	0
Ranitidine	0	0	0	0	Panobinostat	toxic	toxic	toxic	toxic
Retinol	0	0	1032	936	Paroxetine	0	0	toxic	toxic
Sucralose	0	0	0	0	Phenytoin	0	0	--	--
					Teriflunomide	1829	1394	--	--
					Thalidomide	41	21	0	0
					Trichostatin A	4	0	toxic	toxic
					Valproic acid	630	364	878	685
					Vinblastine	toxic	toxic	toxic	toxic
					Vismodegib	0	0	--	--
					Vorinostat	toxic	toxic	toxic	toxic

3.8 Prediction models & assessment of the test system performance

In the previous chapters 3.6 and 3.7 the effects of substances on the gene expression and viability of NEPs were described in detail. In order to classify the results, to compare the in vitro to the in vivo situation and to evaluate the performance of the whole test system, these effects were used to compute prediction models in the form of two classifiers: The "Cytotox 1000" classifier and the "Cytotox SPS" classifier.

3.8.1 Cytotox 1000 classifier

To build the first classifier "Cytotox 1000", a mathematical algorithm called "Leave-One-Out"-cross-validation (LOOCV) used the results for the compound-induced cytotoxicity as well as the gene expression results to predict the probabilities for each compound to be a "positive" (i.e. teratogen) in the in vitro system. The "predicted probability for positive class" (PPPC) of each non-cytotoxic condition based on the genomde-wide gene expression across all samples and was calculated with a value between 0 and 1, whereas the PPPC of all cytotoxic conditions was automatically set as 1. The PPPCs of all conditions are summarized in **Table 27** and are plotted in **Figure 19**, which offers a quick overview of the calculated PPPC of the in vitro results with respect to the in vivo situation. Here, the majority of non-teratogenic substances had PPPCs with low values of max. 25 % whereas the majority of teratogenic substances had PPPCs with high values of min. 75 %. However, some non-teratogenic and teratogenic samples showed also high or low PPPC values, respectively, so that samples of both classes could be found across almost the whole PPPC range from 0 to 1. The three non-teratogenic samples which showed the highest PPPCs were sucralose with PPPCs of 75 % for the 1x C_{max} and 82 % for the 20x C_{max} as well as retinol with 100 % PPPC for the 20x C_{max}. The latter one was switched to the positive in vivo class due to its biological background as a retinoid and its impact on gene expression in the test system which were similar to those of the other in vivo teratogenic retinoids. Strong concentration-dependent differences, where the 20x C_{max} of the same substance had a much higher PPPC than the 1x C_{max} and that based on gene expression data only, were observed for ascorbic acid, retinol, carbamazepine and favipiravir. For other substances, such an effect was not observed or the 20x C_{max} had even lower PPPC than the 1x C_{max}. Interestingly, among the teratogenic conditions, a distinct "gap" could be observed between lithium chloride at the 1x C_{max} and methotrexate at 20x C_{max}. Such a gap was not observed for the non-teratogenics.

Table 27: Predicted probabilities for positive class of all conditions in the classifier "Cytotox 1000" at 1x and 20x C_{max}. Toxic conditions were always considered to be positive with a probability of 1. *Carbamazepine and valproic acid were tested at 10x and 1.67 C_{max}, respectively, instead of 20x C_{max}. Leflunomide, phenytoin, teriflunomide and vismodegib were only tested at 1x C_{max}.

Non-teratogens	Probability		Teratogens	Probability	
	1x C_{max}	20x C_{max}*		1x C_{max}	20x C_{max}*
Ampicillin	0.09	0.05	9-cis-Retinoic acid	0.99	0.99
Ascorbic acid	0.18	0.60	Acitretin	0.92	0.72
Buspirone	0.08	0.05	Actinomycin D	1.00	1.00
Chlorpheniramine	0.04	0.06	Atorvastatin	1.00	1.00
Dextromethorphan	0.41	0.52	Carbamazepine	0.14	0.96
Diphenhydramine	0.53	0.43	Doxorubicin	1.00	1.00
Doxylamine	0.17	0.13	Entinostat	0.93	1.00
Famotidine	0.16	0.19	Favipiravir	0.44	0.77
Folic acid	0.12	0.10	Isotretinoin	1.00	1.00
Levothyroxine	0.09	0.14	Leflunomide	0.93	--

Non-teratogens	Probability		Teratogens	Probability	
	1x C_{max}	20x C_{max}*		1x C_{max}	20x C_{max}*
Liothyronine	0.08	0.05	Lithium chloride	0.46	1.00
Magnesium chloride	0.12	0.20	Methotrexate	0.84	0.69
Methicillin	0.08	0.12	Methylmercury	0.22	0.43
Ranitidine	0.39	0.26	Panobinostat	1.00	1.00
Retinol	0.43	1.00	Paroxetine	0.04	1.00
Sucralose	0.75	0.82	Phenytoin	0.36	--
			Teriflunomide	0.86	--
			Thalidomide	0.33	0.41
			Trichostatin A	0.89	1.00
			Valproic acid	0.87	0.99
			Vinblastine	1.00	1.00
			Vismodegib	0.24	--
			Vorinostat	1.00	1.00

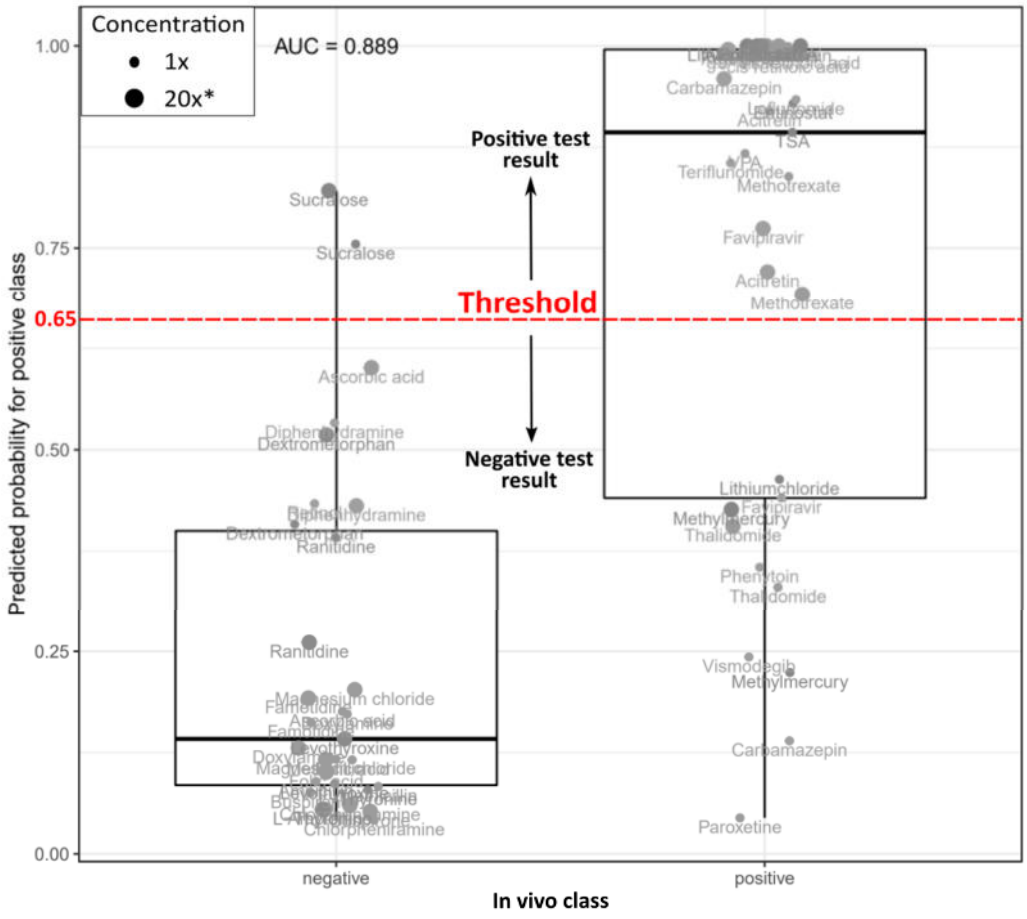


***Figure 19**: The classifier "Cytotox 1000" predicted the probability for positive class of every condition based on the genome-wide gene expression data in mean-values from n=3-4 biological replicates. The probability is given on the y-axis and the x-axis marks the in vivo class for negative (non-teratogen) and positive (teratogen) conditions. Cytotoxic conditions were always considered to be 100 % positive. The threshold at a probability of 65 % separated negative and positive test results. The AUC-value described the precision of the prediction (AUC=1 would mean a perfect prediction). The box plots represent the distribution of the samples as described in chapter 2.2.18. *10x for carbamazepine and 1.67x for VPA. Retinol at 20x was set as "in vivo class = positive"*

Next, the conditions had to be separated into "negative" and "positive" in vitro results with respect to their PPPC to be able to define true negative (TN), true positive (TP), false negative (FN) and false positive (FP) in vitro results. This was done by a threshold: Every condition above a PPPC of 65 % was considered as a positive in vitro test result, and if, for instance, such a condition was also an in vivo class positive, i.e a teratogen, it was called a TP. Consequently, the threshold between ascorbic acid at 20x C_{max} and methotrexate at 20x C_{max} allowed to determine the number of TNs, TPs, FPs and FNs and the calculation of the test system's accuracy, sensitivity and specificity (**Table 28**). Moreover, in order to investigate if the used concentrations had an effect on the performance of the test system, the different outcomes if only 1x C_{max}, only 20x C_{max} (including all 1x C_{max} samples of compounds that were only tested at 1x C_{max}) or all samples were used are shown. In total, Cytotox 1000 classified with a high specificity of 93-94 %, regardless of the concentration, but the use of 20x C_{max} samples instead of 1x C_{max} samples considerably improved the accuracy and sensitivity from 77 % to 87 % and from 65 % to 83 %, respectively. The use of all samples together performed equally or worse than the use of 20x C_{max} samples only. Finally, a LOOCV-algorithm across all samples calculated the the overall ability of the classifier to separate non-teratogens from teratogens. The outcome of that algorithm was an "Area Under Curve" (AUC)-value which indicated that Cytotox 1000 was able to classify the in vitro results with a correctness of 89 %.
In summary, the classifier "Cytotox 1000" predicted the in vitro results in a way that was close to the true in vivo situation when the substances were used at the 20x C_{max}, so that high values for the accuracy, sensitivity and specificity of the UKN1 6-day test system could be reached.

Table 28: Predictions and accuracy, sensitivity and specificity of the "Cytotox 1000" classifier at 1x, 20x and both combined C_{max} with respect to the applied threshold-value of 65 % PPPC. [a]Leflunomide, phenytoin, teriflunomide and vismodegib at 1x C_{max} were included in the calculations for the 20x C_{max} etc. [b]Retinol at 20x C_{max} was regarded as an in vivo teratogenic. [c]10x C_{max} for carbamazepine and 1.67x C_{max} for VPA.

Threshold	C_{max}	In vivo class / In vitro result	Non-teratogenic	Teratogenic	Accuracy	Sensitivity	Specificity	AUC
65% PPPC	1x	Negative	15 (TN)	8 (FN)	0.77	0.65	0.94	0.89
		Positive	1 (FP)	15 (TP)				
	20x[c]	Negative	14 (TN)	4[a] (FN)	0.87	0.83	0.93	
		Positive	1 (FP)	20[a,b] (TP)				
	Combined	Negative	29 (TN)	10 (FN)	0.84	0.77	0.94	
		Positive	2 (FP)	33[b] (TP)				

3.8.2 Cytotox SPS classifier

To build the classifier „Cytotox SPS“, the information of the compound-induced cytotoxicity (**Table 25**) was combined with the number of significant probe sets (SPS) (**Table 26**). Therefore, the classifier used the genome-wide gene expression data in another way than the “Cytotox 1000” classifier. The cytotoxicity was included by considering cytotoxic conditions like conditions that theoretically induced high numbers of SPS. The number of SPS that was defined for these conditions were taken from the highest number observed among the non-cytotoxic conditions, i.e. 4313 SPS as for 9-cis retinoic acid at the 20x C_{max}.

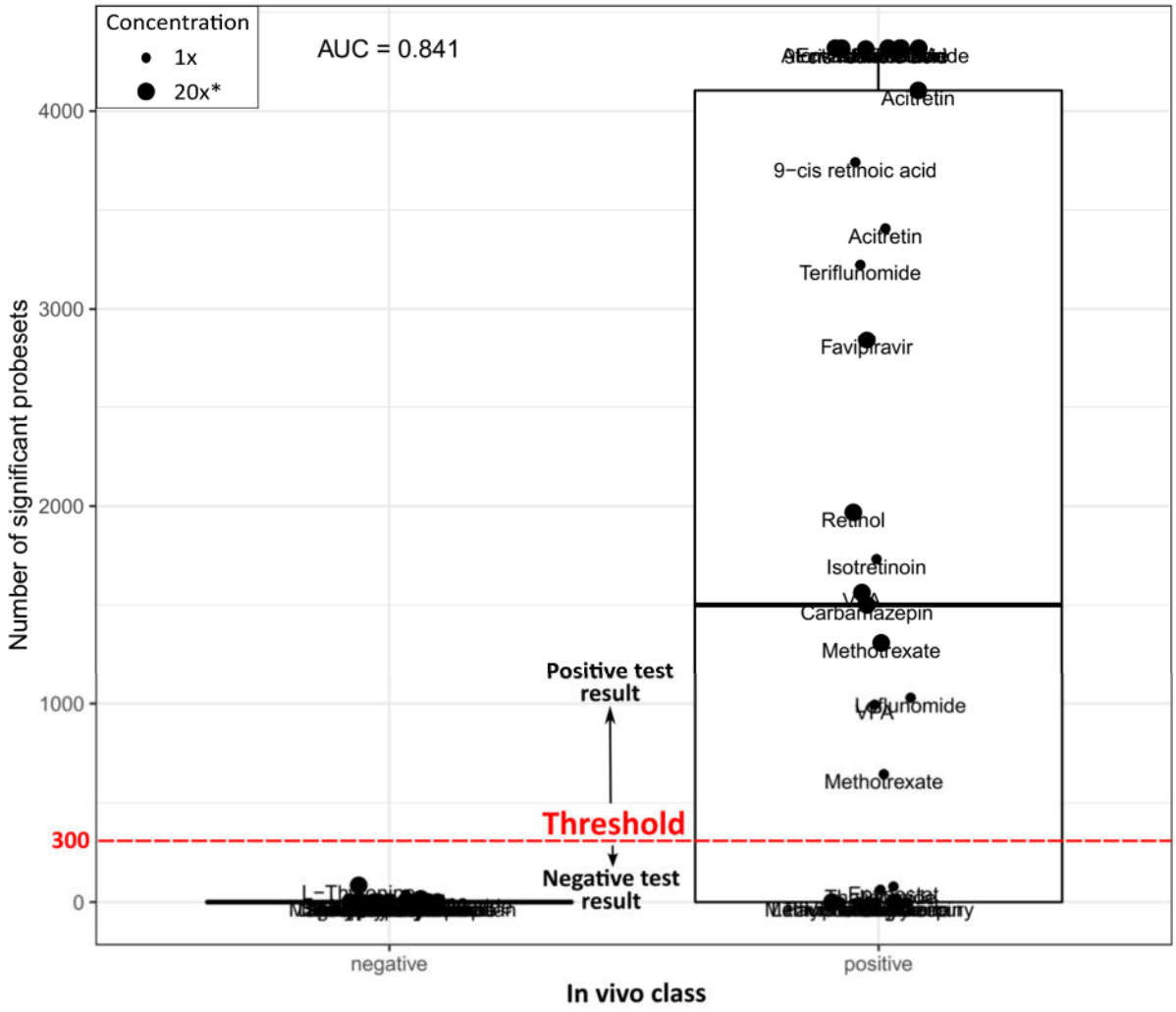


***Figure 20**: The classifier “Cytotox SPS” compares the number of significant probe sets (**Table 26**) to the in vivo class. The number of SPS is given in the y-axis and the x-axis marks the in vivo class for negative (non-teratogen) and positive (teratogen) conditions. Cytotoxic conditions were considered to have 4313 SPS, the highest number observed among the non-cytotoxic conditions. The threshold at 300 SPS separated negative and positive test results. The AUC-value described the precision of the prediction (AUC= 1 would mean a perfect prediction). The box plots represent the distribution of the samples as described in chapter 2.2.18. *10x C_{max} for carbamazepine and 1.67x C_{max} for VPA. Retinol at 20x was set as “in vivo class = positive”*

As already observed in chapter 3.7.2, non-teratogenic conditions did not induce SPS at all except for retinol at the 20x C_{max} and were therefore plotted at the bottom line of the classifier (**Figure 20**). For the classification, retinol at the 20x C_{max} was switched to the positive in vivo class due to its biological background as a retinoid and its high number of SPS which were similar to those of the other in vivo teratogenic retinoids. In contrast to the non-teratogens, in vivo teratogens were plotted along the whole width of the classifier. Many of the in vivo teratogenic conditions induced high numbers of SPS or were cytotoxic, so that they were plotted at the upper end of the classifier. Many teratogens were also found between ca. 500 and 4000 SPS. However, a considerable amount also did not induce any SPS or only a few and were plotted at the bottom line like the non-teratogens. Interestingly, two "gaps" could be observed in the plot where no conditions were plotted: The first one between methotrexate at the 1x C_{max} and the bottom line, and the second one between retinol and favipiravir, both at the 20x C_{max}.

In order to define in vitro positive and negative results, a threshold of 300 SPS was set in the "gap" between methotrexate at the 1x C_{max} and the bottom line-conditions. Every condition that induced more than 300 SPS was considered as a positive in vitro test result. A comparison to the "true" in vivo class then led to the identification of TNs, FNs, TPs and FPs and the calculation of the AUC-value, accuracy, sensitivity and specificity (Table 29). The utilization of only the 20x C_{max} conditions (including all 1x C_{max} samples of compounds that were only tested at 1x C_{max}) demonstrated the highest values for accuracy, sensitivity and specificity of 90 %, 83 % and 100 %, respectively. The AUC-value across all samples was 84 %, but was not cross-validated like it was done for the "Cytotox 1000" classifier.

In all, "Cytotox SPS" classified the in vitro results of the UKN1 6-day test with very high values for accuracy, sensitivity and specificity, especially when only the 20x C_{max} samples were used. It performed slightly better than the "Cytotox 1000", with a similar sensitivity and higher accuracy, which was caused by its outstanding specificity of 1.

Table 29: Predictions and accuracy, sensitivity and specificity of the "Cytotox SPS" classifier at 1x, 20x and both combined C_{max} with respect to the applied thresholds of 300 SPS. [a]Leflunomide, phenytoin, teriflunomide and vismodegib at 1x C_{max} were included in the calculations for the 20x C_{max} etc. [b]Retinol at 20x C_{max} was regarded as an in vivo teratogenic. [c]10x C_{max} for carbamazepine and 1.67x C_{max} for VPA.

Threshold	C_{max}	In vivo class / In vitro result	Non-teratogenic	Teratogenic	Accuracy	Sensitivity	Specificity	AUC
300 SPS	1x	Negative	16 (TN)	10 (FN)	0.74	0.57	1.00	0.84
		Positive	0 (FP)	13 (TP)				
	20x[c]	Negative	15 (TN)	4[a] (FN)	0.90	0.83	1.00	
		Positive	0 (FP)	20[a,b] (TP)				
	Combined	Negative	31 (TN)	12 (FN)	0.84	0.72	1.00	
		Positive	0 (FP)	31[b] (TP)				

3.9 Parabens in the UKN1 6-day protocol

The four parabens methyl-, ethyl-, propyl- and butylparaben were applied in the UKN1 6-day differentiation protocol to investigate their developmental toxic potential, especially that of ethylparaben, and to support a read-across approach for the regulatory risk assessment of ethylparaben. Basically, all steps that were done in the preceding chapters for the 39 non-teratogens and teratogens were also performed for the parabens. This means that (i) parabens were tested at in vitro concentrations that based on human in vivo concentrations, (ii) the cytotoxic effects of parabens were tested and (iii) the effects of parabens on gene expression were analyzed.

3.9.1 In vivo- based in vitro concentrations of parabens

To obtain relevant in vitro concentrations for the UKN1 6-day assay of parabens, a literature search was conducted to find therapeutic in vivo plasma or blood concentrations. **Table 30** summarizes the concentrations found in neonates or umbilical cords from three studies, thus providing optimal data for the calculation of appropriate in vitro concentrations that could be applied in the UKN1 6-day protocol.

Table 30: Therapeutic plasma and blood concentrations of parabens in neonates and umbilical cords

Substance	Therapeutic plasma (peak) or blood concentration [µg/l]	Blood and plasma levels [µM]	Reference
Methylparaben	a) 0.89–210 (mean 54.8) b) Median: 12-15, maximum: 311-874	0.006-5.74	a) Pycke et al. 2015 b) Mulla et al. 2015
Ethylparaben	a) 0.15–3.82 (mean 0.7) b) 0.01–9.95 (mean 0.68)	0.00006-0.06	a) Pycke et al. 2015 b) Geer et al. 2017
Propylparaben	a) 0.27–31.8 (mean 4.28) b) 0.03 – 78.12 (mean 5.59) c) Median: n/d, maximum 134-147	0.00017-0.82	a) Pycke et al. 2015 b) Geer et al. 2017 c) Mulla et al. 2015
Butylparaben	a) 0.09–0.26 (mean 0.09) b) 0.01–0.35 (mean 0.07)	0.00005-0.0018	a) Pycke et al. 2015 b) Geer et al. 2017

In contrast to the 39 previously tested compounds which were tested at the 1x and 20x C_{max}, parabens were tested at eight concentrations in a range from 0.316 – 1000 µM. In this way, the concentrations covered therapeutic concentrations as well as high concentrations which should induce cytotoxicity. Higher concentrations than 1000 µM could not be prepared due to solubility limits of the parabens. Under the line, in vivo-based in vitro concentrations were found for all parabens, so that parabens could be tested in a concentration-dependend manner.

3.9.2 Cytotoxicity and EC50 values of parabens

In order to get information about the cytotoxic potential of parabens in the UKN1 6-day assay, parabens were tested in a concentration-dependend manner between 0.316 – 1000 µM and the cell viability at DoD 6 was assessed with a cell titer blue assay for each concentration and paraben. They were visualized in cytotoxicity curves which in turn enabled the calculation of EC50 values, i.e. the concentration at which 50 % of the cells had died, for each paraben (**Figure 21**, **Figure 22**).
Methyl-, ethyl- and propylparaben were non-cytotoxic up to 100 µM and cell viability only dropped at the two highest concentrations of 316 and 1000 µM (**Figure 21** A and B, **Figure 22** A). The resulting EC50 values were calculated as 1022, 712 and 220 µM, respectively, demonstrating an increasing cytotoxic effect in response to an increasing length of the alkyl-chains in parabens. So, butylparaben was the most cytotoxic paraben, with a stable non-cytotoxic concentration range up to only 10 µM and an EC50 value of 67 µM (**Figure 22** B). Although calculations underlay a strong batch effect between the biological replicates, each paraben showed a distinct cytotoxic strength which was dependent on the size of its alkyl-group. Most important, no paraben was cytotoxic at therapeutic concentrations; the difference between therapeutic and cytotoxic concentrations was at least one magnitude of order.
Summed up, parabens were only cytotoxic at high concentrations which were at least one magnitude of order higher than the therapeutic in vivo concentration, and the cytotoxic potential of each paraben increased with the size of the alkyl-group.

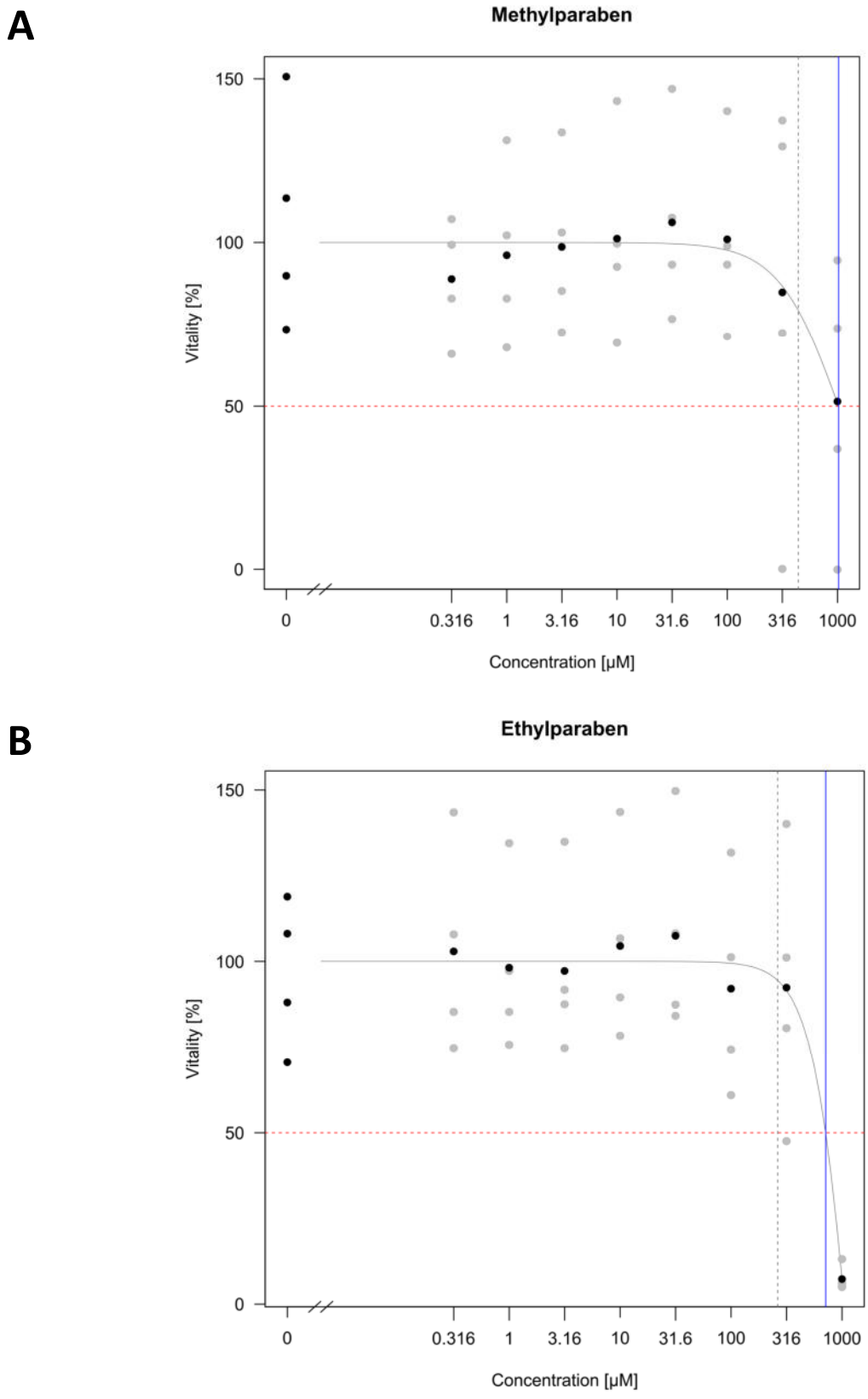


Figure 21: *The vitality of neuroepithelial precursor cells is given in dependence of the applied A) methylparaben and B) ethylparaben concentrations. Grey data points represent vitality data of an independent biological replicate and black data points mean-values of three/four of these replicates. The curves were fitted to 100 % vitality of untreated control cells. The vertical blue line marks the EC50 values on the x-axis. The dashed lines mark the 50 % vitality level (horizontal line) and the confidence intervals (vertical lines).*

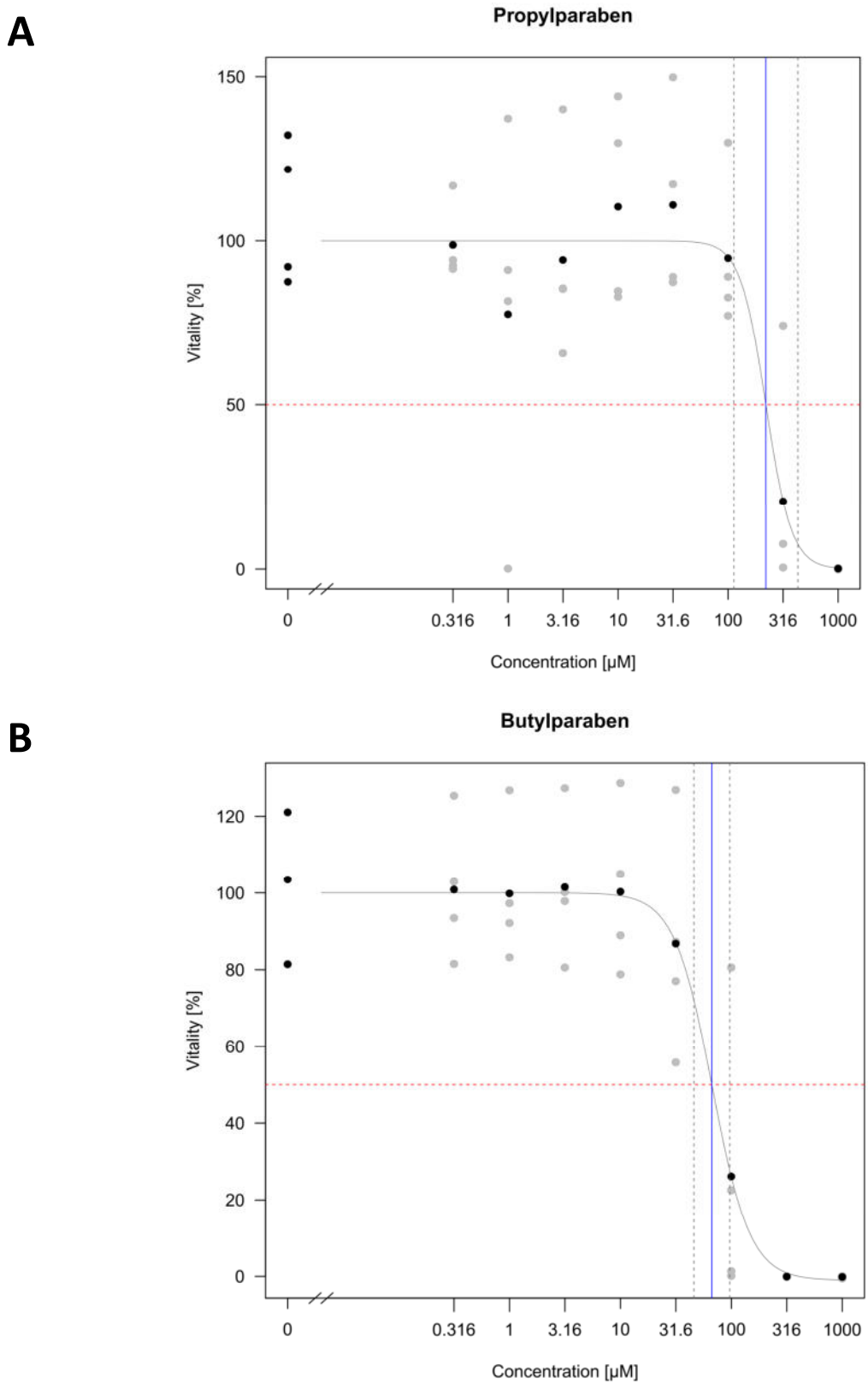


Figure 22: *The vitality of neuroepithelial precursor cells is given in dependence of the applied A) propylparaben and B) butylparaben concentrations. Grey data points represent vitality data of an independent biological replicate and black data points mean-values of three/four of these replicates. The curves were fitted to 100 % vitality of untreated control cells. The vertical blue line marks the EC50 values on the x-axis. The dashed lines mark the 50 % vitality level (horizontal line) and the confidence intervals (vertical lines).*

3.9.3 OCT4A and TFAP2B gene expression levels in parabens

To get a first and low-cost impression of paraben-induced gene expression changes, the two biomarkers OCT4A and TFAP2B in dependence of the applied concentrations were analyzed by RT-qPCR (**Figure 23**). Strikingly, a deregulation of at least two-fold up or down was not found at low and therapeutic concentrations for neither OCT4A nor TFAP2B, but at the highest applied concentrations some deregulation for TFAP2B was observed. However, most of these were not statistically significant upregulated > 2-fold. Only VPA at 0.6 mM induced significant alterations in TFAP2B, but not in OCT4A, and methylparaben at 1000 µM significantly deregulated both genes. Remarkably, the effect of VPA on TFAP2B was approximately 2-fold as strong as of methylparaben.
In total, parabens did not significantly change TFAP2B or OCT4A expressions except methylparaben at the highest applied concentration, and the effect was notable weaker than that induced by VPA.

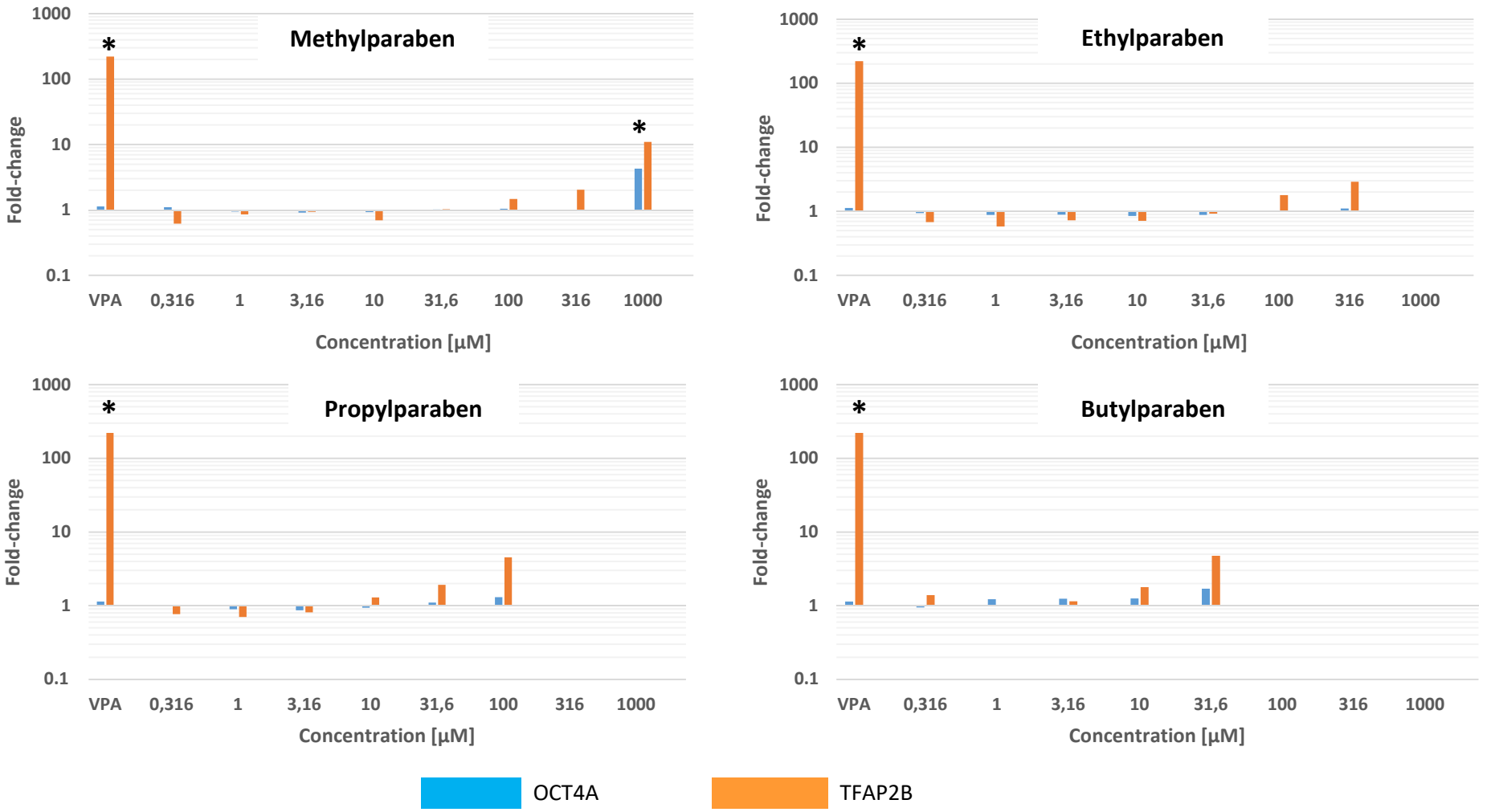


Figure 23: *Concentration-dependent gene expression changes of OCT4A and TFAP2B caused by VPA and methyl-, ethyl-, propyl-, and butylparaben in the UKN1 6-day protocol. The fold-change of the gene expression is given on a logarithmic scale. Mean-values were calculated from n=2-4 independent biological experiments. *The induced gene expression change was significantly higher than two-fold compared to the untreated control.*

3.9.4 Genome-wide gene expression changes in paraben-exposed NEPs

In order to determine the effect of parabens on the whole transcriptome, RNA of paraben-exposed NEPs on DoD 6 of the UKN1 6-day assay was measured with Affymetrix gene arrays like it was done for the 39 compounds in chapter 3.7.

Figure 24 shows the alterations caused by 600 µM VPA, which was used as positive control along with all paraben involving experiments. From all 54,675 analyzed probe sets, 298 and 301 were significantly up- and downregulated, respectively. In contrast, none of the paraben-exposed cells showed significant changes except for two conditions. In 0.316 µM methylparaben- and 100 µM ethylparaben-exposed cells, a total of 6 and 195 probe sets, respectively, were significantly deregulated. (**Figure 25**-**Figure 28**).

So, parabens did not have significant effects on the genome-wide gene expression except for a few significant probe sets in 0.316 µM methylparaben- and 100 µM ethylparaben-exposed cells.

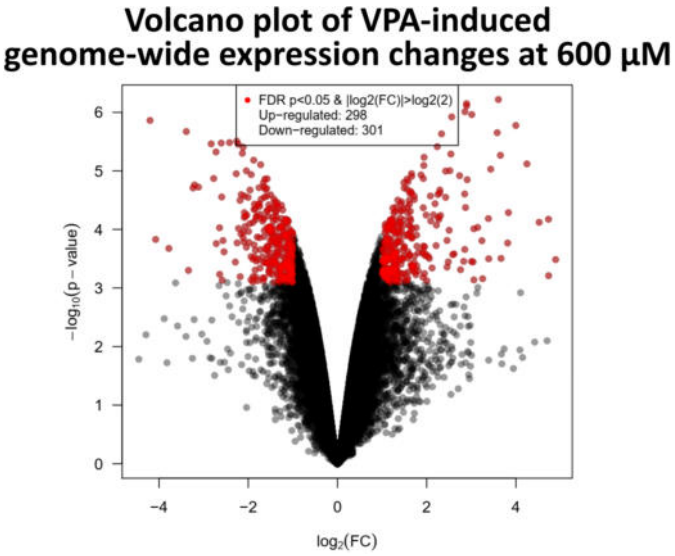


Figure 24: *Genome-wide gene expression changes in NEPs caused by 600 µM VPA that was exposed along with the parabens to the cells. Imaged is the fold-change on the x-axis in its logarithmic form ($log_2(FC)$) and the corresponding, not FDR-adjusted p-value on the y-axis in its negative, logarithmic form ($-log_{10}(p\text{-value})$) of each probe set-mean-value from n=4 independent biological experiments. Highlighted in red are all probe sets, which had an FDR-adjusted p-value of <0.05 and were at least two-fold up or down regulated.*

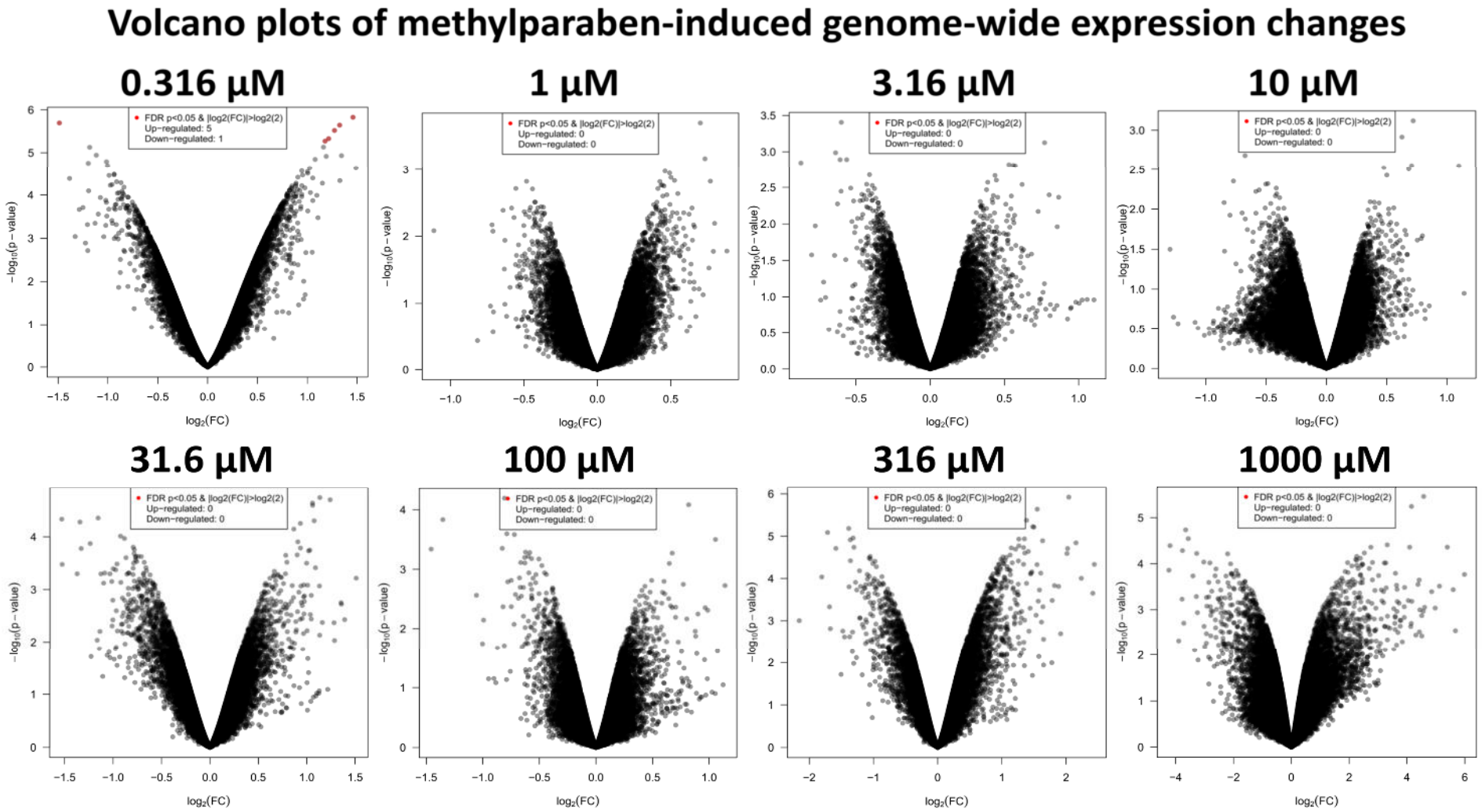


Figure 25: *Genome-wide gene expression changes in NEPs caused by methylparaben at concentrations from 0.316 to 1000 µM. Imaged is the fold-change on the x-axis in its logarithmic form ($log_2(FC)$) and the corresponding, not FDR-adjusted p-value on the y-axis in its negative, logarithmic form ($-log_{10}(p\text{-value})$) of each probe set-mean-value from n=3-4 (n=2 for 0.316 µM) independent biological experiments. Highlighted in red are all probe sets, which had an FDR-adjusted p-value of <0.05 and were at least two-fold up or down regulated.*

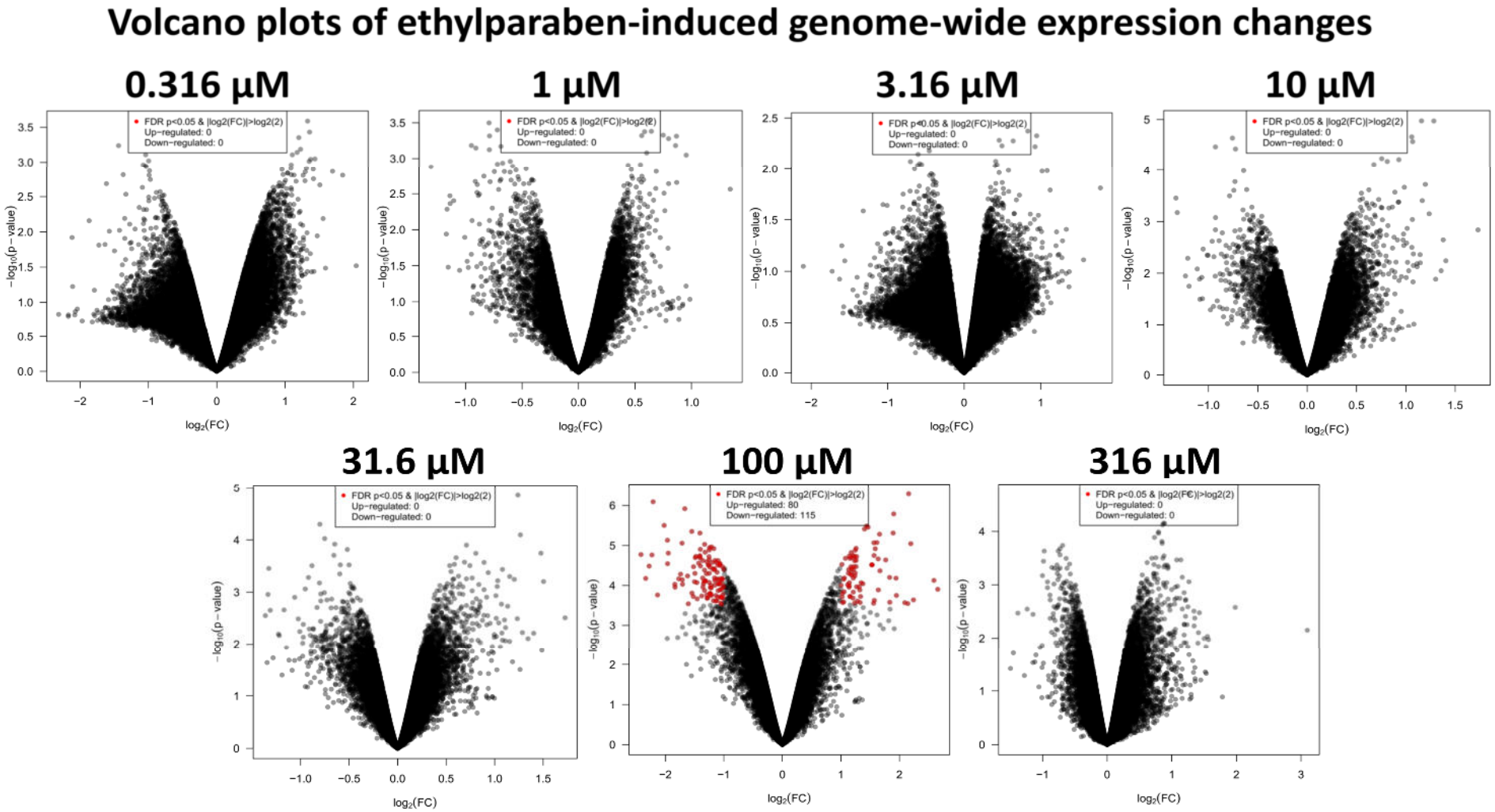


Figure 26: *Genome-wide gene expression changes in NEPs caused by ethylparaben at concentrations from 0.316 to 316 µM. Imaged is the fold-change on the x-axis in its logarithmic form ($log_2(FC)$) and the corresponding, not FDR-adjusted p-value on the y-axis in its negative, logarithmic form ($-log_{10}(p\text{-value})$) of each probe set-mean-value from n=3-4 independent biological experiments. Highlighted in red are all probe sets, which had an FDR-adjusted p-value of <0.05 and were at least two-fold up or down regulated.*

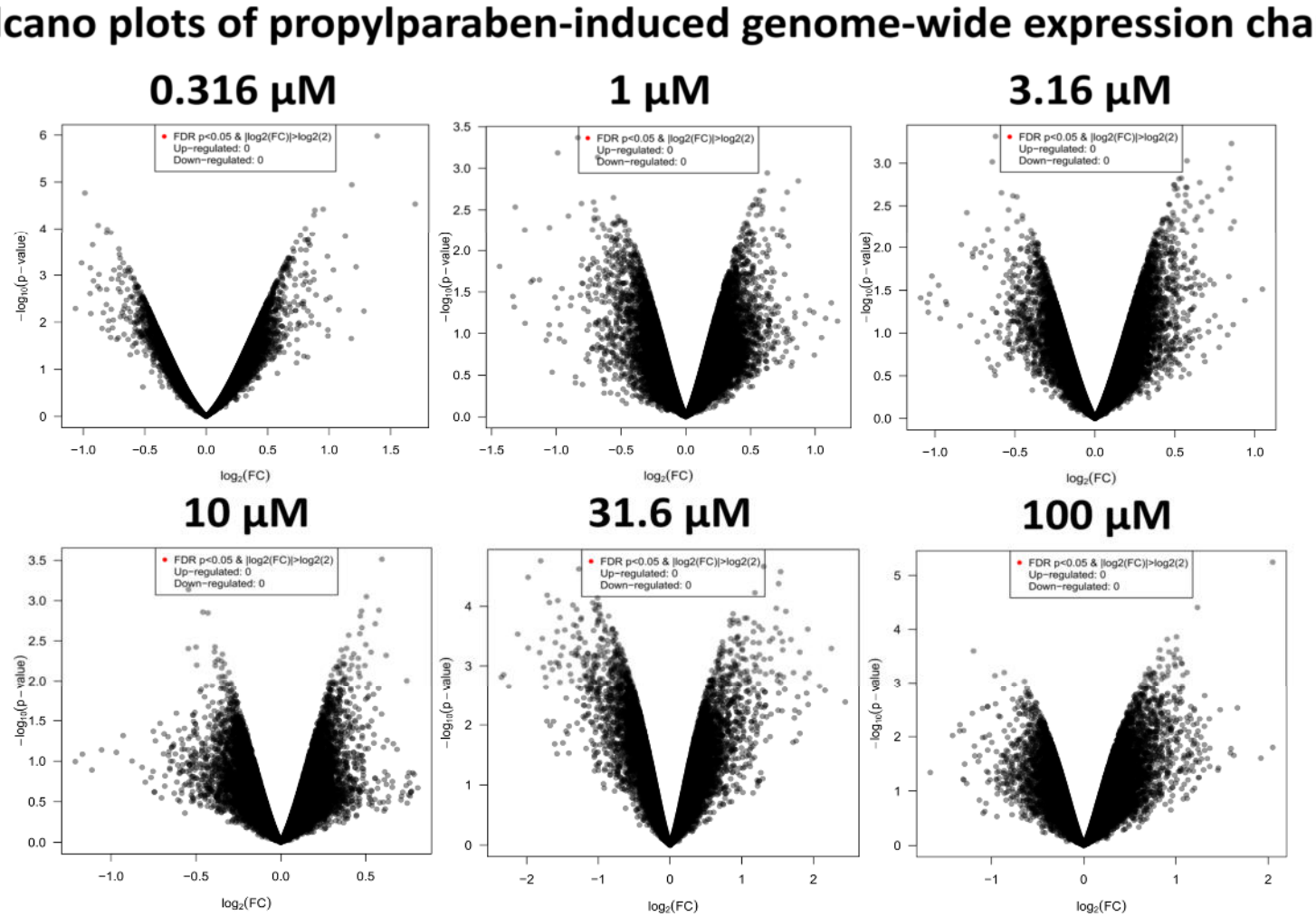


Figure 27: *Genome-wide gene expression changes in NEPs caused by propylparaben at concentrations from 0.316 to 100 µM. Imaged is the fold-change on the x-axis in its logarithmic form ($log_2(FC)$) and the corresponding, not FDR-adjusted p-value on the y-axis in its negative, logarithmic form ($-log_{10}$(p-value)) of each probe set-mean-value from n=3-4 independent biological experiments. Highlighted in red are all probe sets, which had an FDR-adjusted p-value of <0.05 and were at least two-fold up or down regulated.*

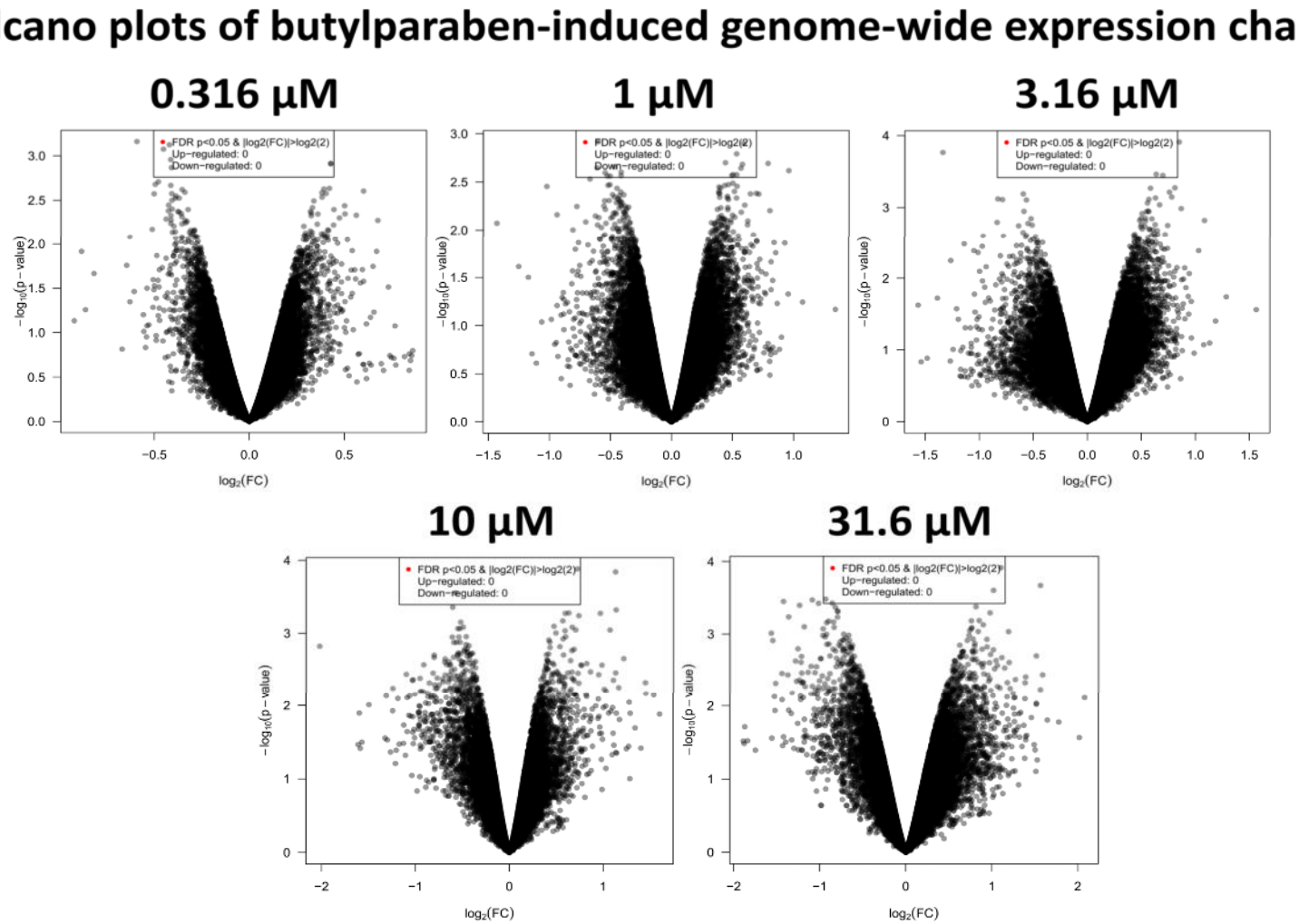


Figure 28: *Genome-wide gene expression changes in NEPs caused by buytlparaben at concentrations from 0.316 to 31.6 µM. Imaged is the fold-change on the x-axis in its logarithmic form ($log_2(FC)$) and the corresponding, not FDR-adjusted p-value on the y-axis in its negative, logarithmic form ($-log_{10}$(p-value)) of each probe set-mean-value from n=3-4 independent biological experiments. Highlighted in red are all probe sets, which had an FDR-adjusted p-value of <0.05 and were at least two-fold up or down regulated.*

3.10 Test optimization: UKN1 one-day protocol

During the progression of the main project, the idea of a further optimization of the UKN1 protocol was born. Thus, the already established UKN1 6-day protocol was reduced by five days, so that the pluripotent stem cells were only differentiated and incubated with substances for 24 h (**Figure 29**). All compounds, which had previously been applied to the UKN1 6-day protocol (see chapter 3.3), were also investigated in the UKN1 one-day protocol on substance-induced effects on cell viability and gene expression, which were described in the following chapters.

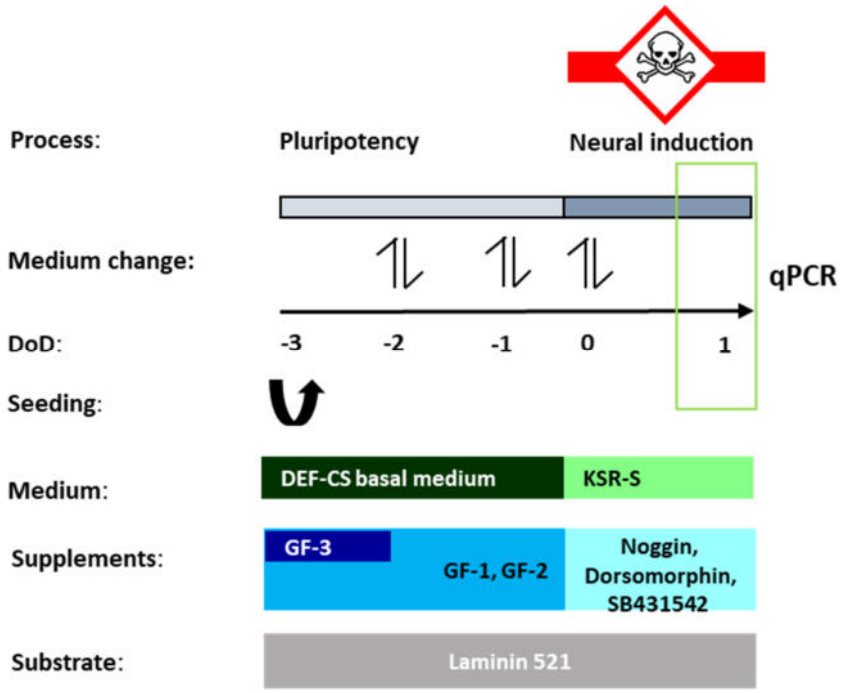


Figure 29: *A shortened one-day-assay utilized the DEF-CS system for the maintenance of the hiPSCs and laminin 521 as the growth surface protein. After 24 h of neural induction in KSR-S supplemented with noggin, dorsomorphin and SB431542, as well as an exposure to substances, cells were harvested for qPCR analysis on DoD 1. Medium was changed daily.*

3.10.1 Cytotoxic effects of teratogenic compounds

To detect the cytotoxic effects of the compounds to the compound-exposed cells, the viability of the cells was estimated qualitatively by cell culture pictures like it was done in the UKN1 6-day system (compare to 3.6). Hence, compounds which demonstrated a clear reduction of viable cells or low RNA-yields compared to untreated control cells were called “Strong” cytotoxic. If the effect was less obvious, but still distinguishable from the untreated control, the compound was called “Little” cytotoxic. If no effect on cells could be observed at all, the compound had “No effect”.
As in the UKN1 6-day test, none of the non-teratogenic compounds showed any cytotoxic effect in the UKN1 one-day test (**Table 31**). Most of the teratogenic substances also showed only few cytotoxic effects. At the 1x C_{max}, only actinomycin D, doxorubicin and vorinostat were strong cytotoxic. At the 20x C_{max}, these three and additionally panobinostat and paroxetine were strong cytotoxic.
In short, non-teratogens were not cytotoxic, and only some of the teratogens, especially at the 20x C_{max}, showed cytotoxicity after 24 h of incubation.

Table 31: Cytotoxicity of teratogenic compounds in the UKN1 one-day protocol. A strong effect means a clear observable cytotoxicity whereas a little effect describes all conditions in which at least some reduction of cell viability could be observed.

	Cytotoxic effect	
C_{max} / Compound	1x	20x
9-cis-Retinoic acid	No effect	No effect
Acitretin	No effect	No effect
Actinomycin D	Strong	Strong
Atorvastatin	No effect	No effect
Carbamazepine	No effect	10x: Little
Doxorubicin	Strong	Strong
Entinostat	No effect	Little
Favipiravir	No effect	No effect
Isotretinoin	No effect	No effect
Leflunomide	No effect	-----
Lithium	No effect	Little
Methotrexate	Little	Little
Methylmercury	No effect	No effect
Panobinostat	Little	Strong
Paroxetine	No effect	Strong
Phenytoin	No effect	-----
Teriflunomide	No effect	-----
Thalidomide	No effect	No effect
Trichostatin A	Little	Little
Valproic acid	No effect	No effect
Vinblastine	Little	Little
Vismodegib	No effect	No effect
Vorinostat	Strong	Strong

3.10.2 Biomarker-based RT-qPCR screening

To assess the impact of substance exposure on the gene expression of the cells, the four biomarker-genes FOXA2, MEIS2, OTX2 and MSX1 were analyzed for significant changes with RT-qPCR.
Overall, non-teratogenic substances induced only minor gene expression changes, and most of them were not significant (**Figure 30** and **Figure 31**). Only ascorbic acid, magnesium chloride and retinol at high concentrations were able to significantly change the expression level of at least one biomarker under the first null hypothesis NH1 (mean-fold-change = 1), meaning that gene expression after substance exposure was indifferent to 1 and didn't change compared to the untreated control. However, after the application of a more stringent, second null hypothesis NH2 (mean-fold-change <= |2|), compounds had to induce "strong" significant gene expression changes of at least two-fold for the upregulation and ½-fold for the downregulation of genes. Here, only retinol at the high concentration induced a significant deregulation of at least one biomarker. In contrast to the non-teratogenic compounds, most of the teratogenics significantly altered the gene expression (**Figure 32**, **Figure 33** and **Figure 34**). In total, 20 from 23 teratogenic compounds induced at least one differentially expressed biomarker under NH1, either at one or both of the applied concentrations, though 8 compounds did so only at the high concentration. NH2 reduced the number down to 15 teratogenic compounds which had a significant effect. Interestingly, substances of the same class induced similar gene expression changes. In particular, retinoids (retinol, 9-cis-retinoic acid, acitretin and isotretinoin) caused a downregulation of OTX2 and an upregulation of FOXA2 and MEIS2 while not affecting MSX1 at all. HDACis (entinostat, panobinostat, TSA, vorinostat) except for VPA downregulated OTX2 and upregulated FOXA2, MEIS2 and MSX1. A third gene expression pattern across different substances showed a downregulation of OTX2 and MEIS2 and an upregulation of FOXA2 and MSX1, as clearly demonstrated by actinomycin D, doxorubicin and paroxetine and slightly demonstrated by lithium chloride and vinblastine at their high concentrations.
Taken together, with few exceptions, non-teratogens had no significant impact on the selected biomarker-genes in RT-qPCR, whereas teratogens induced many significant gene expression alterations.

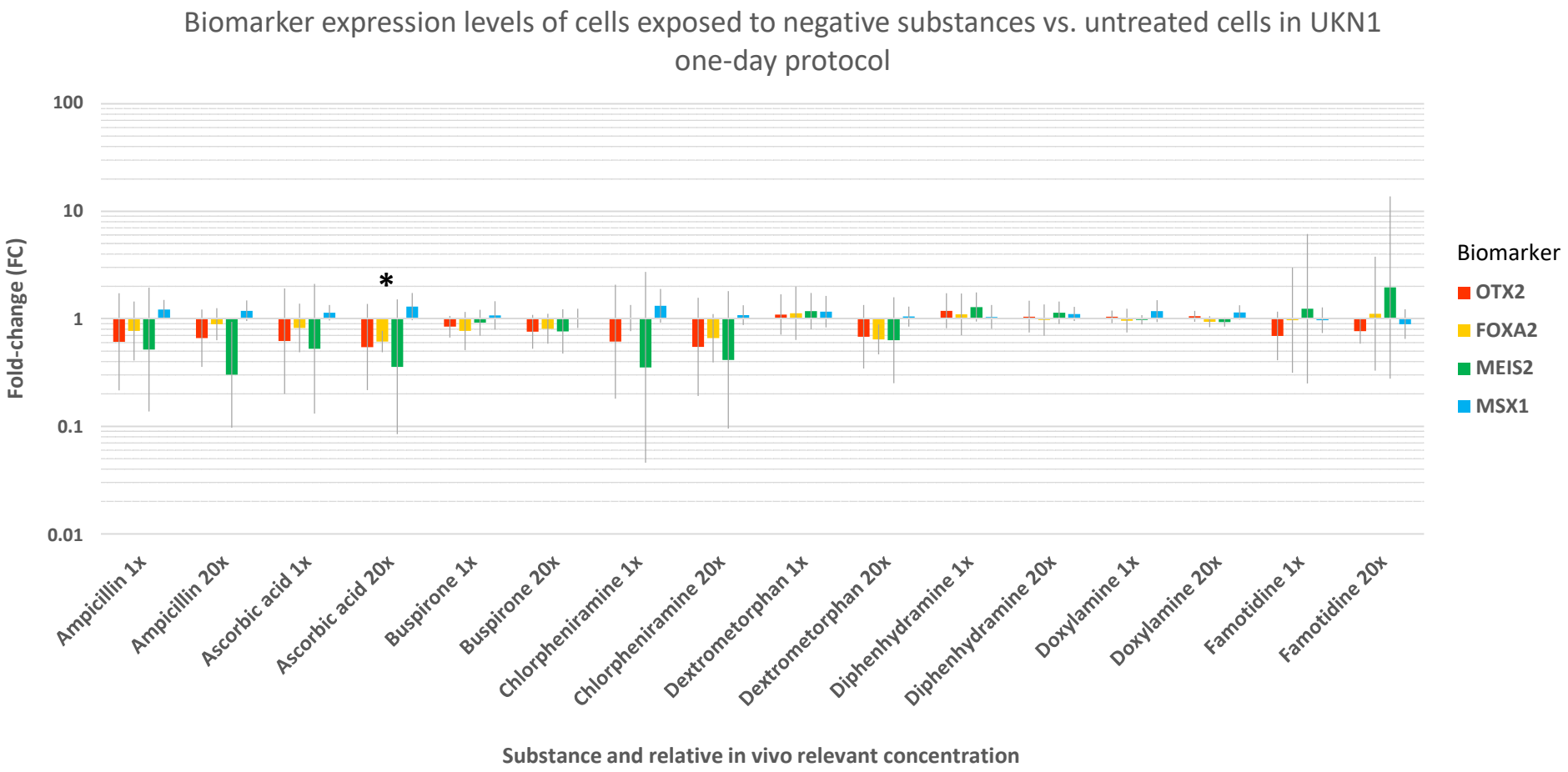


Figure 30: *Gene expression changes of the four biomarkers OTX2, FOXA2, MEIS2 and MSX1 caused by non-teratogenic substances at 1x and 20x C_{max} in the UKN1 one-day protocol. The fold-change of the gene expression is given on a logarithmic scale. Mean-values were calculated from n=3 or n=4 independent biological experiments. Error bars express the standard deviation. *The condition induced at least one significantly (p-value < 0.05) altered biomarker compared to the untreated control (first null hypothesis).*

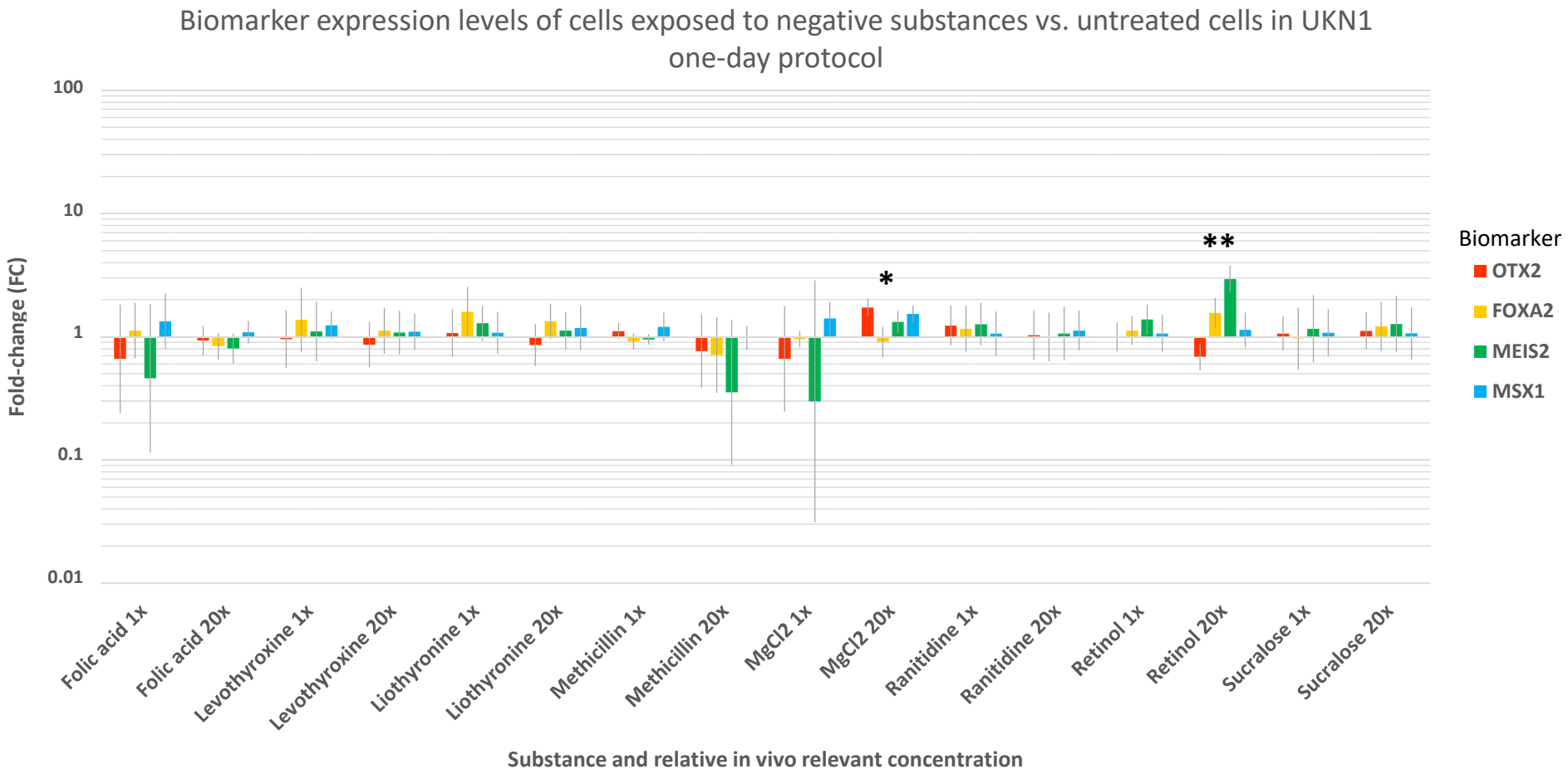


Figure 31: *Gene expression changes of the four biomarkers OTX2, FOXA2, MEIS2 and MSX1 caused by non-teratogenic substances at 1x and 20x C_{max} in the UKN1 one-day protocol. The fold-change of the gene expression is given on a logarithmic scale. Mean-values were calculated from n=3 or n=4 independent biological experiments. Error bars express the standard deviation. *The condition induced at least one significantly (p-value < 0.05) altered biomarker compared to the untreated control (first null hypothesis). **The condition induced at least one altered biomarker which was significantly (p-value < 0.05) higher expressed than two-fold or significantly lower expressed than ½-fold (second null hypothesis).*

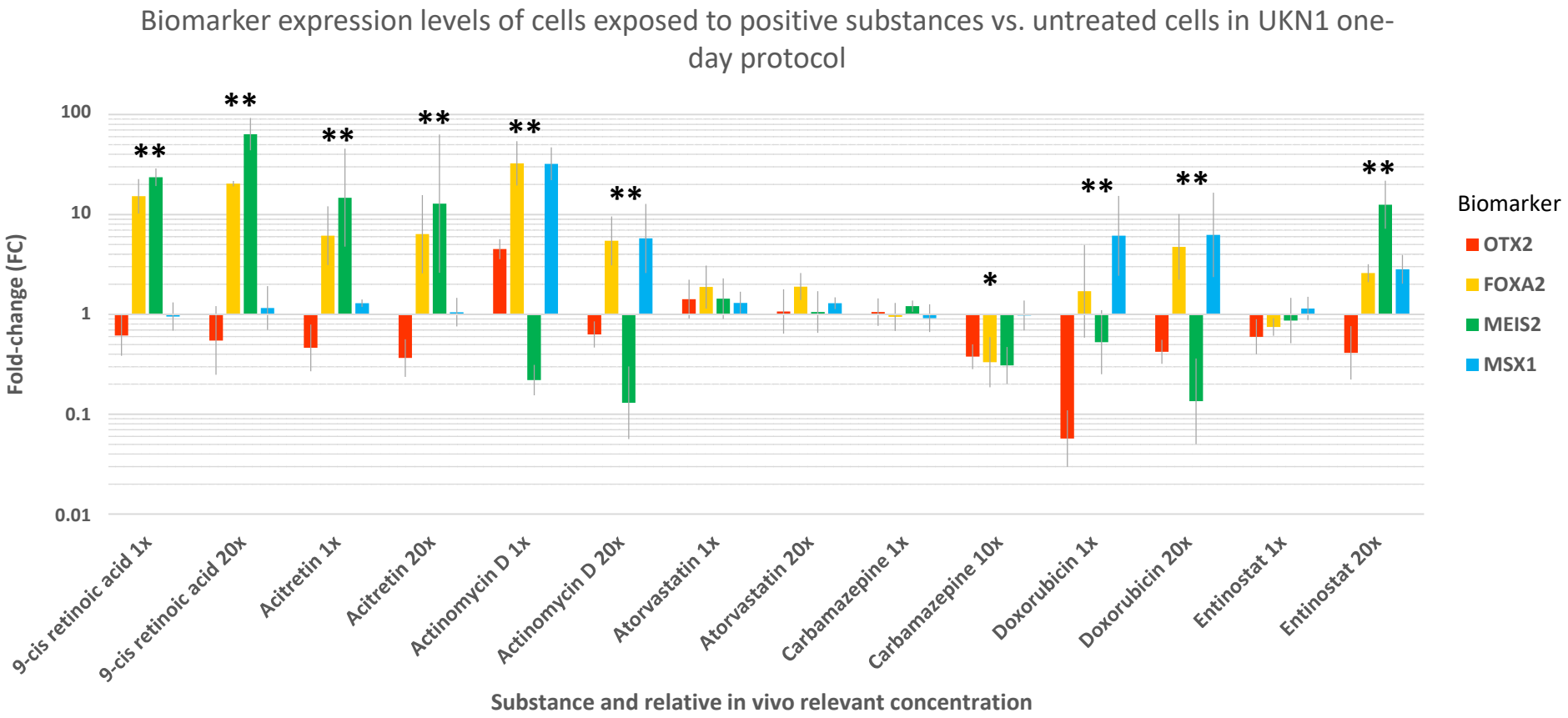


Figure 32: *Gene expression changes of the four biomarkers OTX2, FOXA2, MEIS2 and MSX1 caused by teratogenic substances at 1x and 20x C_{max} in the UKN1 one-day protocol. The fold-change of the gene expression is given on a logarithmic scale. Mean-values were calculated from n=3 or n=4 independent biological experiments. Error bars express the standard deviation. *The condition induced at least one significantly (p-value < 0.05) altered biomarker compared to the untreated control (first null hypothesis). **The condition induced at least one altered biomarker which was significantly (p-value < 0.05) higher expressed than two-fold or significantly lower expressed than ½-fold (second null hypothesis).*

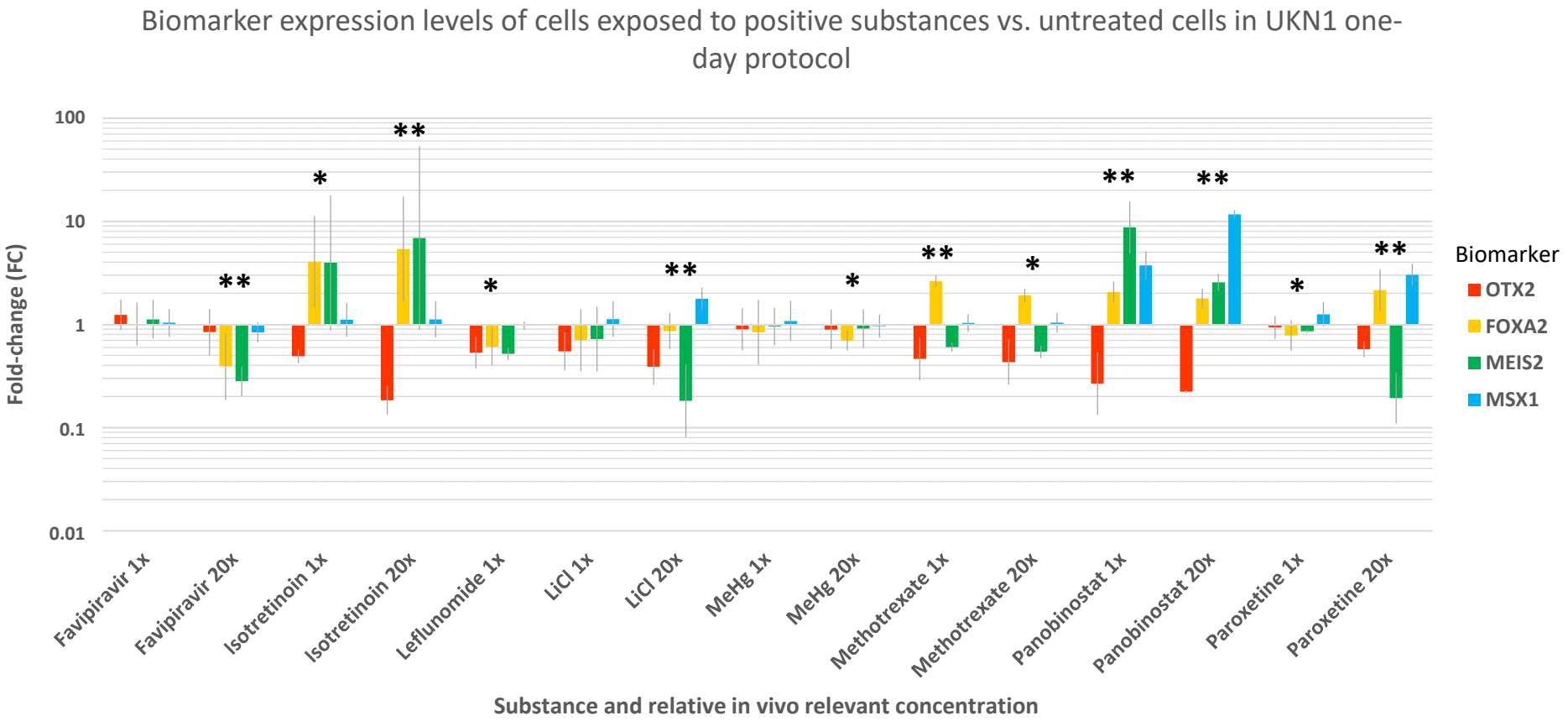


Figure 33: *Gene expression changes of the four biomarkers OTX2, FOXA2, MEIS2 and MSX1 caused by teratogenic substances at 1x and 20x C_{max} in the UKN1 one-day protocol (except for leflunomide which was only tested at 1x C_{max}). The fold-change of the gene expression is given on a logarithmic scale. Mean-values were calculated from n=3 or n=4 independent biological experiments. Error bars express the standard deviation. *The condition induced at least one significantly (p-value < 0.05) altered biomarker compared to the untreated control (first null hypothesis). **The condition induced at least one altered biomarker which was significantly (p-value < 0.05) higher expressed than two-fold or significantly lower expressed than ½-fold (second null hypothesis).*

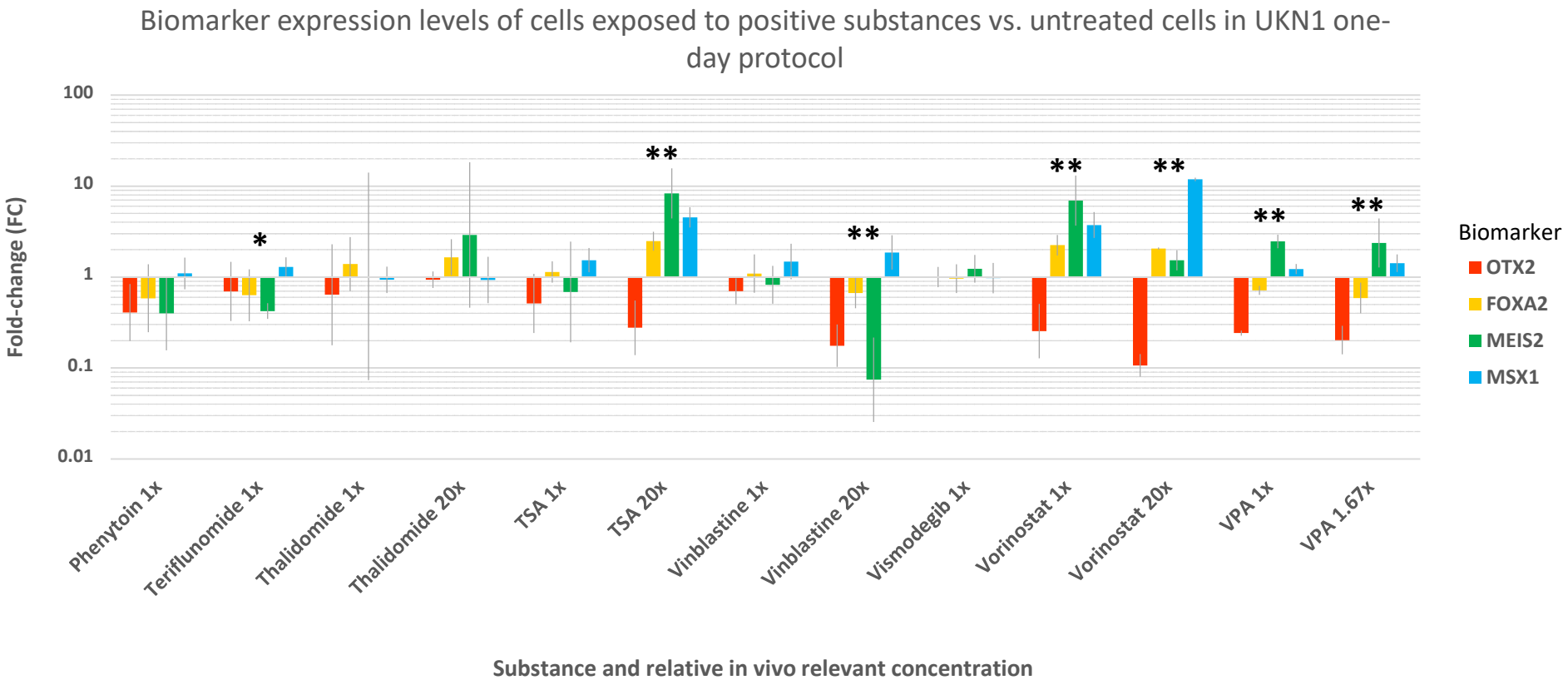


Figure 34: *Gene expression changes of the four biomarkers OTX2, FOXA2, MEIS2 and MSX1 caused by teratogenic substances at 1x and 20x C_{max} in the UKN1 one-day protocol (except for VPA, which was tested at 1x and 1.67x C_{max}, and phenytoin, teriflunomide and vismodegib, which were tested only at 1x C_{max}). The fold-change of the gene expression is given on a logarithmic scale. Mean-values were calculated from n=3 or n=4 independent biological experiments. Error bars express the standard deviation (first null hypothesis). *The condition (here only teriflunomide) induced at least one significantly (p-value < 0.05) altered biomarker compared to the untreated control. **The condition induced at least one altered biomarker which was significantly (p-value < 0.05) higher expressed than two-fold or significantly lower expressed than ½-fold (second null hypothesis).*

3.10.3 Performance of the UKN1 one-day assay

The effects of non-teratogenic and teratogenic compounds on the gene expression levels of the four biomarker-genes FOXA2, MEIS2, OTX2 and MSX1 were used to calculate the accuracy, sensitivity and specificity, so that the performance of the UKN1 one-day assay could be compared to that of the UKN1 6-day assay.

Whenever a compound significantly deregulated at least one biomarker in either the low or the high concentration, it was regarded as a positive test result. Negative and positive in vitro results were compared to the "true" in vivo situation, thus allowing the determination of TNs, FNs, TPs and FPs and the calculation of the assay's accuracy, sensitivity and specificity (**Table 32**). The best performances were achieved when only 20x C_{max} samples were used (including all 1x C_{max} samples of compounds that were only tested at 1x C_{max}). When focusing on accuracy and sensitivity, the use of NH1 was beneficial and led to an accuracy, sensitivity and specificity of 85 %, 83 % and 87 %, respectively. Instead, when focusing on specificity, NH2 performed better and led to an accuracy, sensitivity and specificity of 77 %, 63 % and 100 %, respectively.

All in all, the prediction and classification of compounds with only four biomarker-genes were successful and the UKN1 one-day system performed surprisingly well, although the Cytotox 1000 or Cytotox SPS classifiers still yielded a better prediction.

Table 32: Predictions and accuracy, sensitivity and specificity of the UKN1 24 h assay at 1x, 20x and both combined C_{max} with respect to the applied null hypothesis. [a]Leflunomide, phenytoin, teriflunomide and vismodegib at 1x C_{max} were included in the calculations for the 20x C_{max} etc. [b]Retinol at 20x C_{max} was regarded as an in vivo teratogenic. [c]10x C_{max} for carbamazepine and 1.67x C_{max} for VPA.

NH #	Null hypothesis	C_{max}	In vivo class / In vitro result	Non-teratogenic	Teratogenic	Accuracy	Sensitivity	Specificity
NH1	Mean-ΔΔCT-value = 0	1x	Negative	16 (TN)	11 (FN)	0.72	0.52	1.00
			Positive	0 (FP)	12 (TP)			
		20x[c]	Negative	13 (TN)	4[a] (FN)	0.85	0.83	0.87
			Positive	2 (FP)	20[a,b] (TP)			
		Combined	Negative	29 (TN)	13 (FN)	0.80	0.70	0.94
			Positive	2 (FP)	30[b] (TP)			
NH2	Mean-ΔΔCT-value <= |1|	1x	Negative	16 (TN)	15 (FN)	0.62	0.35	1.00
			Positive	0 (FP)	8 (TP)			
		20x[c]	Negative	15 (TN)	9[a] (FN)	0.77	0.63	1.00
			Positive	0 (FP)	15[a,b] (TP)			
		Combined	Negative	31 (TN)	20 (FN)	0.73	0.53	1.00
			Positive	0 (FP)	23[b] (TP)			

4. Discussion

The regulatory safety evaluation and risk assessment for human health of developmental toxic compounds is very challenging and requires a lot of time, money and animals. Therefore, authorities and scientists world-wide aim at the establishment of proper in vitro methods to reduce or even replace in vivo experiments, not only to diminish the high costs, but also the harm inflicted to the animals. However, up to date only three methods have been validated by the ECVAM (Genschow et al. 2002) and none is approved for the regulatory risk assessment of DT (Bal-Price et al. 2018). The development of new and better in vitro test systems is thus an important issue. In this context, the combination of induced pluripotent stem cells and transcriptomics seems to be a promising approach. iPSCs on the one hand offer a reliable, unlimited source of human cells which are ethically unproblematic and which can simulate developmental processes due to their differentiation capabilities (Wu and Izpisua Belmonte 2015). Transcriptomics on the other hand offer insights into developmental processes which go hand in hand with transcriptional changes (Cardoso-Moreira et al. 2019), so supposing that teratogens which affect development also influence genetic transcription is only logical. Such an assumption is also supported by Meganathan et al. 2012, Krug et al. 2013 and Rempel et al. 2015 who showed that thalidomide, VPA and HDACis cause gene expression alterations in differentiating cells in vitro, respectively, and they were the first who gave rise to in vitro test systems that makes use of both transcriptomics and stem cells.

In this PhD thesis, one of these systems, the so-called UKN1 test system, was optimized and analyzed for its ability to assess the risk of developmental toxic compounds. Therefore, human iPSCs were differentiated to neuroepithelial precursor cells and simultaneously exposed to 39 different substances at human in vivo relevant therapeutic concentrations. Substance-induced, genome-wide gene expression changes were measured with Affymetrix gene arrays and used in combination with the cytotoxic effects of compounds to determine the performance of the optimized “UKN1 6-day” test system. Two distinct classifiers compared the data of the compounds for in vivo teratogenicity to the experimental in vitro results in two different ways and calculated accuracies, sensitivities and specificities of up to 87%, 83% and 93% for the first classifier, respectively, and up to 90 %, 83 % and 100 % for the second classifier, respectively.
Additionally, the system was applied for a read-across-approach in terms of a regulatory risk assessment for parabens, especially for ethylparaben. In our system, all four tested parabens were without any effect on NEPs at human in vivo relevant therapeutic concentrations.
Finally, after the successful implementation of the UKN1 6-day system, first steps towards a further optimization were undertaken by reducing the differentiation and substance-exposure time and replacing the costly transcriptome analysis by a biomarker-based RT-qPCR. Still, substances could be detected with an accuracy, sensitivity and specificity of 77 %, 63 % and 100 %, respectively.

Now, in the following chapters, the results of this work will be discussed in detail.

4.1 Establishment and quality management of the stem cell culture and the UKN1 6-day test system

4.1.1 Comparison of the two culture systems Essential 8 and DEF-CS for the maintenance of pluripotent SBAD2 cells

The handling of pluripotent stem cells needs a high amount of care and accuracy. Their potential to differentiate to almost all kinds of cells is a powerful feature but includes also the risk of unwanted spontaneous differentiation. Therefore, a constant monitoring of their health and pluripotent status is of utmost important. The hiPSC-line SBAD2 was cultured in the two different culture systems Essential 8 and DEF-CS for the first time and their pluripotent status was monitored by a daily check of their morphology and periodical immunohistological analyses of the expression of nuclear and surface pluripotency markers.
Cells cultured in E8 quickly formed flat colonies after they had been seeded as single cells, which were tightly packed after 72 h of cell growth and presented smooth, defined and highlighted edges. Also, cells had a high nucleus-to-cytoplasm-ratio, i.e. a big nucleus and small cytoplasmic volume. In total, the appearance of the hiPSCs conformed to the description in the manufacturer's manual of the Essential 8 system and to what is known in the scientific community (Amit and Itskovitz-Eldor 2012). Moreover, no passage-dependent variances could be observed in the cell culture, meaning that even the highest passage used had no detrimental effect on the appearance of the cells. Just the seeding density showed that a higher number of initially seeded cells formed dense colonies faster. The immunohistological analyses visualized a strong expression of the nuclear markers OCT4A, SOX2, and NANOG and a comparatively weak expression of the surface markers SSEA4, TRA-1-60 and TRA-1-81, but all were clearly distinguishable from the background signal. The expression of these markers is related to pluripotency (Pesce and Schöler 2001; Pan and Thomson 2007; Looijenga et al. 2003; Henderson et al. 2002; Draper et al. 2002) and confirmed the pluripotent status of the SBAD2 cells.
Cells cultured in DEF-CS presented a different morphology than those that had been cultured in E8. Their distribution in the cell culture dish was much more homogenous than in E8, so that they formed highly dense cell areas later despite a higher initial cell number. Maybe cells also grew slower than in E8, but that comparison was not performed so that this aspect remains unclear. The colonies that were formed by the cells in DEF-CS did not have smooth, defined and highlighted edges like in E8 and do not present the typical PSC morphology. However, when cells were high in numbers and tightly packed, they showed the typical PSC feature of small, round and flat cells with comparatively huge nuclei. Also, their overall appearance in the cultivation process matched to the manufacturer's description. Like in E8, no passage-dependent variances could be observed in the cell culture, and a higher cell number in the beginning only accelerated the compaction of cells. The immunohistological expression of the nuclear markers OCT4, SOX2 and NANOG was strong and comparable to the cells cultivated in E8, but the expression of the surface markers SSEA4, TRA-1-60 and TRA-1-81 was notably stronger, maybe due to an imaging of the cells right after labeling which had not been done for the E8 cells.
Taken together, both stem cell culture systems E8 and DEF-CS demonstrated that they were able to maintain SBAD2-hiPSCs in their pluripotent state.

4.1.2 Biomarkers of successful neuroepithelial differentiations

The capabilities of SBAD2 hiPSCs to differentiate into cells of the neural fate was demonstrated by several laboratories (Terryn et al. 2018; EU-ToxRisk 2019). Here in this project, pluripotent SBAD2-hiPSCs cells were differentiated into NEPs via the UKN1 6-day protocol, which was adopted and slightly modified from the original protocol described by Krug et al. 2013. The neural induction of hiPSC was based on the use of noggin, dorsomorphin and SB431542, which blocked SMAD-signaling and

facilitated the directed differentiation to NEPs (Chambers et al. 2009). Under control conditions, NEPs grew to a homogenous and very dense cell layer and showed no morphological differences between the two culture protocols E8 and DEF-CS except those described in the previous subchapter. The genetical quality and neural identity of the cells was controlled by investigating the gene expression levels of seven distinct biomarker-genes with RT-qPCR in hiPSCs, untreated NEPs and VPA-exposed NEPs. In untreated NEPs compared to hiPSCs, NANOG and OCT4A, both famous markers for pluripotency (Pan and Thomson 2007; Looijenga et al. 2003; Pesce and Schöler 2001), were strongly downregulated, showing that the cells lost their pluripotency. At the same time, OTX2, PAX6, TFAP2 and EMX2, all genes which play important roles in early neural development, were upregulated, proving the success of the neural induction. In particular, OTX2 influences neural patterning at multiple points (Puelles et al. 2004), PAX6 and EMX2 are linked to the formation of neuroepithelial cells (Osumi et al. 2008; Frowein et al. 2006) and TFAP2B induces neural crest cells (Dooley et al. 2019; Eckert et al. 2005). The expression of MSX1, which also induces neural crest cells (Osumi et al. 2008), was however not changed. This gene expression pattern was remarkably altered after VPA exposure and the cell fate was switched from neuroepithelium to neural crest: Compared to untreated NEPs, in VPA-treated cells OTX2 was downregulated as well as both neuroepithelial markers PAX6 and EMX2, and neural crest markers TFAP2B and MSX1 were upregulated. These results were reproducible in virtually all experiments and regardless of the used culture system E8 or DEF-CS, underlining the specificity of this directed differentiation. Also, the gene expression patterns were in line with the biomarker expressions observed in previous studies from which the UKN1 6-day protocol was adopted (Balmer et al. 2012; Rempel et al. 2015; Dreser et al. 2020), thus demonstrating the successful establishment of the protocol in our lab, the correct identity of differentiated NEPs and their susceptibility to VPA.

4.2 Substance effects on viability and gene expression of NEPs

The functionality of the UKN1 6-day test system was investigated by exposing differentiating SBAD2 hiPSCs to a selection of 39 non-teratogenic and teratogenic substances and measuring their effects on cell viability and gene expression. Cytotoxicity was determined qualitatively via microscopy, and gene expression changes were obtained by utilizing Affymetrix chips which measured expression levels of 54,675 probe sets. A multivariate analysis revealed differences in gene expression levels that were regarded as both statistical and biological relevant and significant if they had an FDR-adjusted p-value < 0.05 and either a two-fold upregulation or a ½-fold downregulation.
In two PCA-plots which were based on the gene expression analyses of all non-cytotoxic conditions, a diverging genetic impact of non-teratogenic and teratogenic substances could be observed. In both PCA-plots which were based either on all probe sets or only on the top 100 probe sets, the distribution of all conditions was similar: One main cluster was formed of almost all non-teratogenic substances, whereas most of the teratogenic substances were found distinctly apart from that cluster. Thus, a quick view on the PCA-plots already revealed that there were distinct differences between the non-teratogenic and teratogenic substances and their ability to affect gene expression, and that teratogen-induced deregulated genes might be similar.

4.2.1 Effects of non-teratogenic substances

16 out of the selected 39 substances were known to be non-teratogenic in humans at therapeutic concentrations and had therefore been approved by authorities for the use in humans (3.3.1. The classification of the FDA or TGA into classes A and B proved their safe use during pregnancy. For therapeutic concentrations, an extensive literature research was conducted (**Table 22**). Non-teratogenic compounds were applied to the UKN1 6-day test system and exposed to differentiating hiPSCs at 1x and 20x C_{max} (**Table 25**). None of the non-teratogenic compounds induced considerable cytotoxicity, neither at the 1x nor at the 20x C_{max}. Although substance-induced effects on gene expression were detectable, most of them were not significant. Four compounds, namely levothyroxine at the 1x C_{max} and liothyronine, $MgCl_2$ and methicillin at the 20x C_{max}, induced 38, 84, 26 and 2 significantly deregulated probe sets, respectively. Compared to the genetic effects caused by teratogenic substances (see next chapter 4.2.2), which induced several hundreds or thousands significant changes, the effects of these four compounds were very small and considered to result rather from the chosen multivariate analysis and threshold-criteria than from a “real” effect of the compounds. However, one compound indeed had a remarkable effect on gene expression: Retinol at the 20x, but not at the 1x C_{max}, generated 1968 significant differentially expressed probe sets, indicating a teratogenic condition. Although retinol at therapeutic concentrations is regarded as safe and non-teratogenic, its teratogenicity at high doses is known and derived from its active forms all-trans-, 9-cis- and 13-cis-retinoic acid after metabolization in vivo (Comptour et al. 2016; Williams et al. 2020). Thus, a switch of retinol at the high concentration from a true in vivo non-teratogen to a teratogen for the later performed analysis was justified.
In total, non-teratogenic substances did not affect neuroepithelial differentiation except for retinol at a high concentration, which was completely in line with the literature and our expectations.

4.2.2 Effects of teratogenic substances

23 out of the selected 39 substances were known or suspected from human and animal data to be teratogenic in humans at therapeutic concentrations (3.3.2). Their use during pregnancy is contraindicated and the risk is underlined by their classification of the FDA or TGA into pregnancy classes D and X. Additionally, the evidence for teratogenicity of each substance was verified and the underlying mechanisms of action were stated as far as they were known to the scientific community (**Table 21**). As well as for the non-teratogenic substances, an extensive literature research was conducted for the correct determination of human therapeutic concentrations (**Table 23**). The majority of teratogenic compounds were applied at 1x and 20x C_{max} in the UKN1 6-day test system (Table 24). In contrast to non-teratogenic compounds, many teratogenic compounds were cytotoxic at either both the 1x and 20x C_{max} or only at the 20x C_{max} (**Table 25**). Interestingly, the strongest cytotoxic compounds that killed all cells already at the 1x C_{max}, were the antineoplastic agents actinomycin D, doxorubicin and vinblastine which interact with DNA, related proteins or microtubule, as well as atorvastatin, an HMG-CoA reductase inhibitor, and two of the HDAC inhibitors, namely panobinostat and vorinostat. So, the differentiating stem cells were especially susceptible and sensitive to compounds, which interfere with transcription or target directly basic cellular functions which are essential for cell survival and growth. Maybe this is due to the nature of differentiating stem cells, which grow fast and where numerous transcription events take place. That at least panobinostat and vorinostat can also have an significant effect on gene expression if applied at lower concentrations was shown in Rempel et al. 2015, where they were used in approx. a 20-fold lower concentration than the 1x C_{max}, indicating that the cytotoxicity in our system was a concentration-dependent effect.
The genome-wide transcriptome analyses of all little- or non-cytotoxic conditions revealed that many teratogenic compounds were able to induce considerable gene expression alterations compared to unexposed NEPs and that the effects depended strongly on the applied substance as well as on the concentration. In total, nine out of 23 teratogenic compounds, namely 9-cis-retinoic acid, acitretin, carbamazepine, favipiravir, isotretinoin, leflunomide, methotrexate, teriflunomide and VPA induced significant changes, and except for carbamazepine and favipiravir, these changes happened already at therapeutic concentrations. The most important compound among them was VPA due to its status as a well-known teratogen at therapeutic concentrations and its known impact on gene expression of hiPSCs in the UKN1 differentiation protocol based on the experiences from former studies. In Krug et al. 2013 and Rempel et al. 2015, VPA at 0.6 mM induced $2164_{up}/1533_{down}$ and $2004_{up}/1458_{down}$ significantly deregulated probe sets, respectively, and in our UKN1 6-day test system it deregulated $630_{up}/364_{down}$ probe sets. The differences in the exact amounts of deregulated probe sets could be explained due to variant thresholds for the biological relevance and biological variations overall, like the use of dissimilar cell lines. So, VPA as the gold-standard compound validated the basic functionality of our UKN1 6-day test system to detect genome-wide, teratogen-induced transcriptional changes. Furthermore, the vulnerability of SBAD2 hiPSCs to VPA means that at least one of the molecular targets of VPA, which are numerous, had to be present in the cells. As methotrexate as another teratogenic compound also induced a lot of significant changes in SBAD2 cells and both methotrexate and VPA have antifolate effects, it can be hypothesized that the molecular target of them in SBAD2 cells was indeed folate. To further elucidate this assumption in a context of basic research, a “folate-rescue” experiment by a simultaneous application of methotrexate or VPA with folic acid could be an idea.
The retinoids 9-cis-retinoic acid, acitretin and isotretinoin altered between $1029_{up}/703_{down}$ and $2426_{up}/1887_{down}$ probe sets and even retinol at the 20x C_{max} deregulated $1032_{up}/936_{down}$ probe sets, indicating that the retinoid receptors, which are the molecular targets of retinoids and involved pathways were present in SBAD2 hiPSCs. The cells responded to 9-cis-retinoic acid, acitretin and isotretinoin already at the therapeutic concentrations with lots of genetic changes that even increased at the higher concentration, but the extent depended on the substance and retinol had such an effect

only at the 20x C_{max}, although the applied concentrations were almost the same for the retinoids. This showed that the specific structure of each compound had a substantial effect on the effective concentration and overall effectivity on SBAD2 cells. It can therefore be assumed, that all retinoids and substances, which have similar molecular targets and pathways, would induce gene expression changes in NEPs as well, but these changes would occur only at substance-specific teratogenic concentrations when a "teratogenic threshold" was surpassed.

Leflunomide and teriflunomide, both DHODH-inhibitors, altered up to $1829_{up}/1394_{down}$ probe sets significantly, meaning that SBAD2 hiPSCs likely expressed DHODH which was affected. Interestingly, teriflunomide, which is the active metabolite of leflunomide, induced much more significant probe set alterations than its prodrug leflunomide at the same concentration. This indicates that teriflunomide indeed was the active form which was responsible for the genetic impact and that SBAD2 cells were able to metabolize leflunomide to teriflunomide.

Carbamazepine and phenytoin, both teratogenic anticonvulsants that block voltage-gated sodium channels, did not induce significant changes at therapeutic concentrations. Both were used at concentrations at the upper limits of their solubility and the highest applied concentrations were 1x C_{max} for phenytoin and 10x C_{max} for carbamazepine. Although the latter condition was able to have significant effects on gene expression with $910_{up}/591_{down}$ deregulated genes, SBAD2 cells were obviously less sensitive to this kind of substance class than to other classes, which was further underlined by the fact, that no cytotoxic concentration could be reached.

Favipiravir, an antiviral agent, and vismodegib, a hedgehog pathway inhibitor, did not induce significant changes in the gene expression of NEPs at the 1x C_{max}, but favipiravir did so at the 20x C_{max} with $1551_{up}/1290_{down}$ probe sets. This effect of favipiravir was very interesting as it normally targets specifically the RNA machinery in influenza virus, so probably favipiravir addressed other pathways and molecular targets in UKN1 than that. Although human in vivo data do not exist, it is teratogenic in animals. In contrast, vismodegib was as insoluble at higher concentrations and as effectless as phenytoin, indicating an insensitiveness of SBAD2 cells towards hedgehog pathway inhibitors.

For entinostat, lithium (LiCl), paroxetine and TSA, no or just negligible significant effects on gene expression at the 1x C_{max} were observed, but all compounds were highly cytotoxic at the 20x C_{max} and could therefore be clearly distinguished from the non-teratogenic compounds. Moreover, there is evidence that these compounds would have altered the gene expression in the UKN1 6-day test if they had been applied in the correct concentrations. It was shown in Rempel et al. 2015 that entinostat and TSA can induce gene expression changes at the applied 1x C_{max} in UKN1. However, the used cells and the media were different than in our study, what probably had an influence on the effective concentrations. The effect of lithium on the gene expression of cells is evident (Roux and Dosseto 2017) and could also be shown recently for embryonic stem cells (Mnatsakanyan et al. 2020). In the latter example, lithium was used in a concentration of 10 mM und would correspond a 10x C_{max} in our test system, underlining the dose-dependent effect and the assumption that a concentration which is in between the 1x and 20x C_{max} could induce gene expression alterations in SBAD2 cells. Paroxetine showed developmental toxic effects to iPSC-derived-brainspheres (Zhong et al. 2020) and changed gene expression and DNA methylation in brain tissues of rodents and zebrafish (Glover et al. 2019; Sillaber et al. 2008; Huang et al. 2020), indicating that paroxetine probably would have affected gene expression in our system at concentrations higher that 1x C_{max}, but lower than 20x C_{max}.

Two teratogenic compounds that were soluble at both the 1x and the 20x C_{max} but which did not have any effect on cells, neither on viability nor on gene expression, were methylmercury and thalidomide. The non-detection of them could have different reasons. The most obvious are that the selected compounds were not hazardous at all, that the concentrations were wrong or that the test system was insensitive to the treatment under the given conditions. In the case of methylmercury, the chosen concentration probably was too low to induce effects. Although the compound is known to cause developmental defects, the chosen concentration was linked to therapeutic doses in humans and

resulted in 20 nM and 400 nM for the 1x and 20x C_{max}, respectively, whereas hazardous effects in humans were reported to occur at blood levels of >800 nM (LOAEL) (Robert A. Young 1992). However, Karagas et al. 2012 reported to have found evidence for neurocognitive defects in children caused by low-level methylmercury exposure, so adverse effects of low concentrated methylmercury cannot be ruled out completely. That methylmercury can induce effects was shown by Rempel et al. 2015. At a distinct higher concentration of 1500 nM, methylmercury induced significant gene expression changes in the UKN1 system. Therefore, methylmercury could actually be a non-teratogenic compound at the used concentrations, which showed no effects, but this possibility needs further investigation.
In contrast, thalidomide is known to act teratogenic at therapeutic doses (Speirs 1962), so the applied concentrations were not a matter to change, especially as even the 20x C_{max} had no effect. A thalidomide-induced transcriptomic response in vitro was already demonstrated by Meganathan et al. 2012, also at therapeutic concentrations, but the test system was different, with longer exposure- and differentiation time windows. Thus, it seems that our test system indeed was insensitive to thalidomide. A last explanation could be that thalidomide was hydrolyzed into several metabolites, whose teratogenic potentials seem to be low or absent compared to the mother-compound (Lepper et al. 2006), before it could take action. In unbuffered aqueous solutions it has a short half-life of approximately seven hours in contrast to 44 days in plasma (Teo et al. 2004). The possibility of a fast hydrolysis of thalidomide and other compounds was not considered during the experiments as a general application of freshly prepared medium every day in the multi-compound and -dose testing regime was not feasible, but it could be worth to retest for specific compounds.
Taken together, most of the applied teratogenic substances could clearly be distinguished from non-teratogenic substances and identified as potential teratogens due to their ability to induce remarkable gene expression changes in differentiating SBAD2-hiPSCs or due to their cytotoxicity. The effects and the numbers of significant gene expression alterations were very substance-specific as well as concentration-dependent, and the “ideal” concentration range for some substances, i.e. the range in which they would have been non-cytotoxic but potential teratogenic by differentially expressed genes, was very narrow. For instance, 9-cis-retinoic acid was non-cytotoxic, but teratogenic at 1x and 20x C_{max}, whereas LiCl was non-cytotoxic and non-teratogenic at 1x, but cytotoxic at 20x, so that no gene expression could be measured. Some substances like doxorubicin were already cytotoxic at the 1x C_{max}. And VPA, the gold-standard, was teratogenic at both the applied 1x and 1.67x C_{max}, where it became little cytotoxic und would probably been have 100 % cytotoxic at a 3x C_{max}. So far, the in vivo therapeutic concentrations offered the most reasonable and effective basis for the applied concentrations in the UKN1 6-day test system, but in order to assess the teratogenic potential completely on gene expression alterations, it would be necessary to utilize a test schedule which is more flexible and applies concentrations with respect to each compounds cytotoxicity, for example by the use of so-called benchmark concentrations (Rempel et al. 2015).

4.3 Performance of the UKN1 6-day test system

4.3.1 Classifier “Cytotox 1000” vs. “Cytotox SPS”

The in vitro results from the UKN1 6-day test system were used to build the two classifiers “Cytotox 1000” and “Cytotox SPS” in order to compare the in vitro results to the in vivo situation and to assess the performance of the system. Although both classifiers utilized the compound-induced cytotoxicity in their prediction models, they made different use of the gene expression results: The first classifier looked at the compound-induced gene expression alterations of the 54,675 probe sets in a leave-one-out-cross-validation to calculate the probability of each sample to be a teratogenic compound. The second classifier utilized the numbers of significantly deregulated probe sets to separate positive from negative test results. The performance of the classifiers was determined by AUC-values, accuracies, sensitivities and specificities, whereas higher values meant a better match of the in vitro and in vivo data and all values could reach a theoretical maximum of 1 = 100 %. The AUC-values, which were computed with a novel mathematical algorithm described in Albrecht et al. 2019 and corresponded in a way with the classical accuracy, calculated the performance of the classifiers in a single value and were determined to be 0.889 for “Cytotox 1000” and 0.841 for “Cytotox SPS”. Accuracies, sensitivities and specificities were determined as up to 87 %, 83 % and 93% for “Cytotox 1000” (**Table 28**) and 90 %, 83 % and 100 % for “Cytotox SPS” (**Table 29**), respectively, showing that both classifiers were very similar in their assessments of the in vitro test results and that the UKN1 6-day test system did a pretty good job in predicting in vivo teratogens and non-teratogens. The best prediction in both classifiers was achieved when only the results from the higher concentrations were used instead of the 1x C_{max}. This is due to the fact that some teratogenics were detected only at the higher concentration, whereas all non-teratogenics, except for retinol and ascorbic acid, showed no or only minor differences between the 1x and 20x C_{max}, and underlined the outstanding specificity of the test system even at high concentrations. In a biological context, the absolute specific separation of the classifier “Cytotox SPS” allows the interpretation, that every condition which induces a considerable number of significantly deregulated genes in the UKN1 6-day test is a true teratogen in the human in vivo situation. However, the biological interpretation of the results of the classifier “Cytotox 1000” is not so clear-cut. Many in vivo non-teratogens showed high probabilities for the positive class and particularly the positive prediction of sucralose with a probability of 75 – 82 % was much unexpected. Especially in comparison to the high concentrated teratogens acitretin and methotrexate, which were predicted with lower probabilities than sucralose, the high PPPCs for some non-teratogens were irrational. The best explanation might be, that the high PPPCs were based on shortcomings of the underlying mathematical approach rather than on an actual effect. However, the biological relevance cannot be ruled out completely. For instance, sucralose increased adipogenesis and the accumulation of ROS in MSCs (Kundu et al. 2020), altered gut microbiota and induced hepatic steatosis (Liu et al. 2019; Dai et al. 2020), and also overdosed ascorbic acid might be disadvantageous (NRC and IOM 2000).

Altogether, both classifiers attestet a high performance to the UKN1 6-day test system and showed only slight differences in the calculated values for the AUC, accuracy, sensitivity and specificity. Nevertheless, the classifier “Cytotox SPS” demonstrated a better performance and was more reasonable in a biological context, although the algorithm-based AUC was lower than for “Cytotox 1000”. Therefore, the approach of “Cytotox SPS” to focus on the number of SPS as a kind of value to determine the “magnitude” of genetic deregulation which was caused specifically by an individual substance seemed to be more beneficial than the approach of “Cytotox 1000”, which searched for genetic similarities across all substances and all probe sets, although the latter one was much more sophisticated with respect to the underlying statistical and mathematical means.

4.3.2 Comparison to other test systems

Based on the classifier "Cytotox SPS", the performance of the UKN1 6-day test system was assessed with an accuracy, sensitivity and specificity of 90 %, 83 % and 100 %, respectively, for 39 tested compounds. In **Table 33**, an overview of other test systems for developmental toxicity is given which were mainly reviewed and described in detail by Luz and Tokar 2018. All of the shown assays shared the idea to detect developmental toxicity in vitro in cells which were capable to differentiate to other cell types. The eldest assays used mouse or rat embryonic cells, whereas more recent ones switched to human embryonic cells to overcome interspecies variations before finally using human iPSCs as the best applicable and most versatile cell type. However, differences between the assays were manifold and involved the used cells, the exact protocol, the number of tested compounds, the read-out and the way, in which the accuracy, sensitivity and specificity were calculated, so that a direct comparison between the different assays was difficult. Nevertheless, with respect to the given data for each assay, the UKN1 6-day test was one of the best performing. The eldest assays mEST, MM and WEC demonstrated the worst performance of all with accuracies of 70-80 %, no information about sensitivity and specifity and a small set of compounds, but in contrast to the other assays, they were validated by the ECVAM. The variations of the mEST, i.e. FACS-, transgenic-, molecular-EST and EBT, applied different novel and promising read-outs and techniques, but could not provide overall convincing results: The FACS-EST had an accuracy of 100 %, but only a set of 10 compounds; the transgenic EST had a high accuracy and sensitivity, but only a limited specificity and also a small substance set; the molecular EST tested many compounds, but the performance was weak; and the EBT had a small substance set and missed to state the sensitivity and specificity. The ReProGlo assay, which had as a high specificity as the UKN1 6-day assay, tested only 17 compounds and were additionally less accurate and sensitive than the UKN1 6-day assay. Indeed, only the two assays which used targeted metabolomics or high-throughput imaging could compete with the performance of the UKN1 6-day assay and were therefore of special interest and worth to be discussed in detail in the following lines. Interestingly, these assays were the only two in the comparison which used human instead of rodent cells, and as mouse and human embryonic stem cells have some differences in their molecular signaling, the better outcomes maybe related to that (Schnerch et al. 2010; Gabdoulline et al. 2015).

The targeted metabolomics assay was established in 2013 and performed almost identical to the UKN1 6-day system. The number of tested compounds as well as the composition of the substance set was very similar and the applied concentrations were also anchored to therapeutic in vivo concentrations in humans. Moreover, the assay was 100 % specific, and the in vivo teratogens, which were used in both systems, were also predicted similarly, except for thalidomide, which was not detected by the UKN1 6-day system. However, a closer look at the metabolomics assay, which measured only two amino acids to detect teratogenicity, gave the impression that the results of the test were overfitted with respect to its sensitivity. For instance, the therapeutic and teratogenic concentration of all-trans retinoic acid was stated with 1.2 µM in the test description, but the test itself detected the substance as teratogenic already at 0.00004 µM, meaning a relative sensitivity increase of 30,000-fold. Such inconsistensies were observed for multiple substances. Moreover, although in the metabolomics assay many biological and technical replicates were measured, the in vitro results had remarkable error bars, indicating a high biological variability of the results. In 2014, the assay was sent for validation to the ECVAM, but in October 2015, the process was already stopped and the assay was not considered for further validation studies due to a PARERE consultation opinion, stating that even professionals were not convinced of the assay's functionality (ECVAM 2015). Under the line, the reliability and robustness of the whole assay is questionable.

Table 33: Overview of test systems for developmental toxicity and their performance. Modified after Luz and Tokar 2018

Assay / read-out	Cells	Tested compounds	Accu-racy	Sensi-tivity	Speci-ficity	Reference
UKN1 6-day	hiPSC	39	90	83	100	
UKN1 one-day			77	63	100	
Targeted metabolomics	hESC	36	89	79	100	Palmer et al. 2013
	hESC	1065	79-82 %	<67 %	>84 %	Zurlinden et al. 2020
High-throughput imaging	hESC	71	94	97	92	Kameoka et al. 2014
mEST	mESC & fibroblasts	14-20	78 %	--	--	Genschow et al. 2002
FACS-EST		10	100 %			Buesen et al. 2009
Transgenic EST		24	83-92	93	63-87	Suzuki et al. 2011
Molecular EST	mESC	65	72	76	69	Panzica-Kelly et al. 2013
EBT	mESC & fibroblasts	21	90.5	--	--	Kang et al. 2017
ReProGlo	mESC	17	76	71	100	Uibel et al. 2010
MM	Rat embryo limb bud	14-20	70 %	--	--	Genschow et al. 2002
WEC	Whole rat embryo	14-20	80 %	--	--	Genschow et al. 2002

A recent publication from 2020 used the metabolomics assay to screen 1065 ToxCast substances. Most of the substances could not be anchored to therapeutic in vivo concentrations and the authors utilized complex computer and data modeling strategies to assess the performance of the system, so that this approach could hardly be compared to classical tests. Despite the shortcomings of the test, the authors made a sophisticated suggestion how thousands of samples could be tested and evaluated in a test system and tried to offer a solution for the upcoming problem to test the myriad of compounds in the progressively increasing substance libraries of companies.

The second very interesting test system, namely the high-throughput imaging assay, demonstrated a big set of compounds, an impressive performance with 97 % sensitivity, and the easiest read-out by an immunohistochemical measurement of just the single marker SOX17. A detailed analysis, however, revealed quickly a lot of shortcomings: At first, the substance set was poor. 50 % of it was composed of internal substances, and they represented the majority of the non-teratogenic substances, meaning that only a few known non-teratogens were used. This issue raised the question if non-teratogenic FDA pregnancy class A compounds like diphenhydramine or doxylamine, which were included in our UKN1 6-day test as well as in the metabolomic assay, were not tested on purpose to improve the test result. Also, these internal substances prevented a comparison to publicly existing data. The composition of the teratogenic compounds was better and included some of the best-known teratogens. Another major shortcoming was the absence of an anchoring to human therapeutic concentrations and that the in vitro test results were solely compared to animal in vivo data, what increased the risk of a misclassification. Lastly, the prediction model was too sensitive and overfitted for some compounds. For example, VPA was predicted in the test to be teratogenic at a concentration of 2.6 µM, although its therapeutic and teratogenic concentration is 600 µM, a 250-fold increase of sensitivity. So, the significance and predictivity of the chosen marker and the test system was doubtful.

Altogether, the UKN1 6-day assay demonstrated the best performance among the discussed test systems. The sophisticated approach resulted in a highly robust, specific and predictive test system, whereas other assays revealed many weak spots like a bad performance, small substance sets or a too drastical reduction of method complexity, like the measurement of just one or two biological markers, which led to overfitting and misclassifications.

4.4 Risk assessment of parabens

The UKN1 6-day system, which successfully predicted the in vivo teratogenic potential of 35 out of 39 investigated substances was also used to assess the teratogenic potential of methyl-, ethyl-, propyl- and butylparaben. The applied concentrations of 0.316 – 1000 µM covered therapeutic as well as cytotoxic concentrations and the effect on NEPs was measured by viability assays, RT-qPCR and genome-wide gene expression analysis. High concentrated parabens were found to be cytotoxic and the cytotoxicity increased with the size of the alkyl-side-chain of the parabens, so that methylparaben was the weakest and butylparaben the strongest cytotoxic compound. None of the parabens induced any significant gene expression changes in NEPs in contrast to VPA, which was used along with the experiments as a positive control. Although ethylparaben at 100 µM induced some significant gene expression alterations in NEPs, it was considered as an outlier because the number of altered probe sets was relatively low and the higher concentration of 316 µM did not induce significant gene alterations.

Unexpectedly, methylparaben at 1000 µM, which was cytotoxic to 50 % of the cells and whose gene expression could still be measured, showed also no significant gene expression alterations. This finding allows two different interpretations: (i) Cytotoxicity in UKN1 does not cause significant gene expression changes. In Shinde et al. 2017 and Waldmann et al. 2017, the authors tried to discriminate, whether differentially expressed genes were caused by teratogenicity or cytotoxicity, but they missed to include non-teratogenic compounds in their analyses, so the findings could possibly solely related to teratogenicity. Furthermore, they declared that other transcriptomic data than theirs about pluripotent stem cell cytotoxicity were rare, which is still the case. So, to investigate if cytotoxicity indeed can induce gene expression alterations in the UKN1 6-day assay, it will be necessary to test non-teratogenic compounds, which did not show significantly altered genes in the applied concentrations, in increased, cytotoxic concentrations and measure the gene expression alterations subsequently. (ii) The measurement and analyses were erroneous or the biological variance among samples was so high that no significance in the gene expression data could be find. A hint for this possibility is given by the volcano plots of the limma results of methylparaben (1000 µM) which showed many probe sets with fold-changes>2, a pattern which was usually only observed in teratogenic conditions. Another hint was observed in RT-qPCR expression analyses where elevated levels of OCT4A and TFAP2B in conditions with slight or strong cytotoxicity were found, especially for methylparaben at 1000 µM. Either a repetition of the experiments or the already mentioned measurement of cytotoxic non-teratogens could shed a light on this issue.

In sum, the tested parabens showed no teratogenic potential in the UKN1 6-day test system which is in line with our expectations and supports the idea of a read-across-approach for the risk assessment of ethylparaben. The necessary in vivo data for this approach were recently provided by Fayyaz et al. 2021 who also discussed the data with focus on the endocrine system and concluded that the investigated parabens had no endocrine disrupting properties, in contrast to the opinion of many other scientists. Unfortunately, the UKN1 6-day system cannot provide any further information on that since its response to EDC-exposure is unknown yet as well as the endocrine function of hiPSCs. To clear this up, other EDCs which are structurally unrelated to parabens would have to be tested in the UKN1 6-day system.

4.5 Substance effects in a shortened UKN1 one-day protocol

All substances that were tested in the UKN1 6-day test system were reused and applied in a shortened UKN1 one-day protocol in which cells were exposed to substances for only 24 h. The effects were detected by the measurements of four selected biomarkers in RT-qPCR instead of an Affymetrix gene chip measurements of 54,675 probe sets. In contrast to a substance-exposure of 144 h, cells of all conditions after 24 h substance-exposure were viable enough to get a sufficient amount of RNA for the RT-qPCR analyses, so that the assessment of the performance of the system was solely based on gene expression changes, although some conditions were also cytotoxic after 24 h.

The up- and downregulation of each of the four markers FOXA2, MEIS2, OTX2 and MSX1 was substance-specific and a higher concentration had increasing as well as decreasing effects on the fold-change of an altered biomarker, but not on the direction of the deregulation. Also, distinct patterns of gene alterations could be observed for substances of the same class like the retinoids or the HDAC inhibitors as well as for non-HDACi-cytotoxic substances and underlined the MoA-specific effects of the substances on gene expression. Strikingly, the results in the UKN1 one-day protocol were very comparable to those of the UKN1 6-day protocol and "Cytotox SPS", meaning that (i) all conditions that had been cytotoxic after 144 h showed remarkable gene expression alterations of the four biomarkers after 24 h, except for atorvastatin at 1x and 20x C_{max} and vinblastine at 20x C_{max}, and could therefore easily be detected, (ii) the UKN1 one-day protocol was highly specific, (iii) the majority of teratogenics could be detected and (iv) the performance improved with use of only 20x C_{max} samples instead of only 1x C_{max}. The exact classification of each condition and the subsequent outcome for the overall performance depended on the two applied null hypotheses (NH) for significance: NH1 and NH2 determined a significant gene expression if it was different from a fold-change = 1 or if it was greater than an absolute fold-change = 2, respectively. For the following comparison, only 20x C_{max} samples including leflunomide, phenytoin, teriflunomide and vismodegib at the 1x C_{max} were considered: NH2 was the more stringent one and had a high similarity to the "Cytotox SPS" classifier. It classified the substances almost in the same way except for atorvastatin, carbamazepine, leflunomide and teriflunomide, which were not detected as positives and classified as false negatives. These misclassifications were the reason for the reduced performance compared to "Cytotox SPS". NH1 in contrast identified carbamazepine, leflunomide and teriflunomide as teratogenics, but also ascorbic acid, $MgCl_2$ and MeHg at the 20x C_{max} as well as paroxetine at 1x C_{max} were detected to express significant gene expression alterations, which had no significant probe sets in the UKN1 6-day protocol. An estimation, which of the null hypothesis provided the more correct classification of the conditions in a biological context is hard to say, because paroxetine at the 1x C_{max} is a known in vivo teratogen, MeHg is said to be a potential teratogen and also ascorbic acid and $MgCl_2$ may be hazardous in high concentrations (NRC and IOM 2000; Cascella and Vaqar 2021). None of the null hypotheses-related results were completely in line with the detected significant probe sets of the UKN1 6-day protocol, so that comparison also is not helpful. Nevertheless, it is logical to assume that (i) substances are more cytotoxic to cells and have a greater impact on the gene expression of cells when they are incubated for 96 h instead of 24 h and that (ii) a whole-transcriptome analysis is more sensitive than four biomarkers. In this sense it is rather plausible that e.g. teriflunomide was a false negative after 24 h and a true positive after 144 h than that paroxetine was a true positive after 24 h but a false negative after 144 h. Thus, the more reasonable null hypothesis was the more stringent NH2 which led to more false negatives and an accuracy, sensitivity and specificity of 77 %, 63 % and 100 %, respectively. In comparison to the UKN1 6-day protocol, the inferior performance could be related to either the reduced complexity of the read-out (i.e. Affymetrix gene chips vs. four biomarkers) or to the shortened exposure and differentiation time, in which substances could take effect on the cells.

4.6 Conclusion & future perspectives

The UKN1 6-day test system showed a high accuracy in the in vitro prediction of teratogenic and non-teratogenic substances as well as a reliable sensitivity and an outstanding specificity. The system confirmed that teratogenic substances at in vivo therapeutic concentrations indeed cause remarkable gene expression changes in differentiating pluripotent stem cells, indicating that this is possibly also happening in developing embryos under similar conditions and providing an explanation for the hazardous in vivo action of teratogens. To date, the UKN1 6-day protocol is probably the best test system to detect the teratogenic potential in vitro and it is therefore worth to extend the number of tested compounds in the system to assess its value for regulatory risk assessment. Along that way, a few specific points will have to be taken care of and followed with certain awareness: (i) The sensitivity of the system, especially in the context of false negatives, (ii) the classification of substances whose human therapeutic concentrations are unknown or which are tested in a blinded manner and (iii) the expensiveness of whole-transcriptome analyses. The best way to investigate the sensitivity of the system is to test as many different compounds as possible. In addition, the testing of specific compound classes like retinoids could verify the response of the sysem towards that specific class, and this could also be a way to address the unsensitivity of the system towards thalidomide. The assessment of unknown or blinded compounds necessitate a multiple-dose, concentration-dependent testing schedule with a subsequent sophisticated method to classify the results. In such an approach, the applied concentrations of a compound could be anchored to its cytotoxicity, if therapeutic concentrations were not available. The data of the four parabens that were already measured in that way during this PhD thesis can now facilitate an in vitro-based read-across-approach for the regulatory risk assessment of ethylparaben, thus underlining the indispensability of such a strategy. The reduction of the costs could be achieved by either future advances in methods to measure transcriptomics more efficiently than with Affymetrix gene chips or by reducing the complexity of the read-out, for example with targeted transcriptomics. The UKN1 one-day protocol was such an approach and demonstrated an impressive predictivity, although the accuracy and sensitivity were not as good as of the UKN1 6-day protocol. An optimized biomarker selection or the combination with other low-cost read-outs like metabolomics could possibly improve the performance of the UKN1 one-day protocol. Alternatively, it could be combined with the UKN1 6-day assay in the following two-step model: At first, the UKN1 one-day protocols identifies all substances which cause significant changes of the biomarkers as teratogens and leaves over the remaining substances for whole-transcriptome analyses, thereby reducing the number of samples and costs of the latter one. Another idea is to use not the UKN1 one-day protocol but the UKN1 6-day assay with targeted transcriptomics, as on the one hand the biomarkers could be optimally selected based on the known results from the whole-transcriptome analyses, and on the other hand the effects of teratogenic substances are more present due to a longer exposure- and differentiation time, thus improving the final outcome. Last but not least, it would be worth to investigate the reproducibility of the results in other hiPSC-lines or labs to support and underline the applicability of the system for the regulatory risk assessment of developmental toxic substances.

5. References

Aalto, T. R.; Firman, M. C.; Rigler, N. E. (1953): p-Hydroxybenzoic Acid Esters as Preservatives*. In *Journal of the American Pharmaceutical Association (Scientific ed.)* 42 (8), pp. 449–457. DOI: 10.1002/jps.3030420802.

Abou-Alfa, Ghassan K.; Lewis, Lionel D.; LoRusso, Patricia; Maitland, Michael; Chandra, Priya; Cheeti, Sravanthi et al. (2017): Pharmacokinetics and safety of vismodegib in patients with advanced solid malignancies and hepatic impairment. In *Cancer chemotherapy and pharmacology* 80 (1), pp. 29–36. DOI: 10.1007/s00280-017-3315-8.

Albrecht, Wiebke; Kappenberg, Franziska; Brecklinghaus, Tim; Stoeber, Regina; Marchan, Rosemarie; Zhang, Mian et al. (2019): Prediction of human drug-induced liver injury (DILI) in relation to oral doses and blood concentrations. In *Arch Toxicol* 93 (6), pp. 1609–1637. DOI: 10.1007/s00204-019-02492-9.

Ali, Shaukat; Aalders, Jeffrey; Richardson, Michael K. (2014): Teratological effects of a panel of sixty water-soluble toxicants on zebrafish development. In *Zebrafish* 11 (2), pp. 129–141. DOI: 10.1089/zeb.2013.0901.

Allenby, G.; Bocquel, M. T.; Saunders, M.; Kazmer, S.; Speck, J.; Rosenberger, M. et al. (1993): Retinoic acid receptors and retinoid X receptors: interactions with endogenous retinoic acids. In *Proceedings of the National Academy of Sciences of the United States of America* 90 (1), pp. 30–34. DOI: 10.1073/pnas.90.1.30.

Amit, Michal; Itskovitz-Eldor, Joseph (2012): Atlas of Human Pluripotent Stem Cells. Derivation and Culturing. Totowa, NJ, s.l.: Springer; Humana Press (Stem Cell Biology and Regenerative Medicine). Available online at http://gbv.eblib.com/patron/FullRecord.aspx?p=884439.

Apotex Inc. (2013): Product Monograph Apo-Lithium Carbonate. Apotex Inc. Available online at https://pdf.hres.ca/dpd_pm/00023004.PDF, checked on 11/27/2020.

Arya, Manit; Shergill, Iqbal S.; Williamson, Magali; Gommersall, Lyndon; Arya, Neehar; Patel, Hitendra R. H. (2005): Basic principles of real-time quantitative PCR. In *Expert review of molecular diagnostics* 5 (2), pp. 209–219. DOI: 10.1586/14737159.5.2.209.

Bailey, Jordan; Oliveri, Anthony; Levin, Edward D. (2013): Zebrafish model systems for developmental neurobehavioral toxicology. In *Birth defects research. Part C, Embryo today : reviews* 99 (1), pp. 14–23. DOI: 10.1002/bdrc.21027.

Baird, I.McLean; Shephard, N. W.; Merritt, R. J.; Hildick-Smith, G. (2000): Repeated dose study of sucralose tolerance in human subjects. In *Food and Chemical Toxicology* 38, pp. 123–129. DOI: 10.1016/S0278-6915(00)00035-1.

Balis, F. M.; Holcenberg, J. S.; Poplack, D. G.; Ge, J.; Sather, H. N.; Murphy, R. F. et al. (1998): Pharmacokinetics and pharmacodynamics of oral methotrexate and mercaptopurine in children with lower risk acute lymphoblastic leukemia: a joint children's cancer group and pediatric oncology branch study. In *Blood* 92 (10), pp. 3569–3577.

Balmer, Nina V.; Weng, Matthias K.; Zimmer, Bastian; Ivanova, Violeta N.; Chambers, Stuart M.; Nikolaeva, Elena et al. (2012): Epigenetic changes and disturbed neural development in a human embryonic stem cell-based model relating to the fetal valproate syndrome. In *Human molecular genetics* 21 (18), pp. 4104–4114. DOI: 10.1093/hmg/dds239.

Bal-Price, Anna; Pistollato, Francesca; Sachana, Magdalini; Bopp, Stephanie K.; Munn, Sharon; Worth, Andrew (2018): Strategies to improve the regulatory assessment of developmental neurotoxicity (DNT) using in vitro methods. Toxicology and Applied Pharmacology, 354, 7-18. In *Toxicology and Applied Pharmacology* 354, pp. 7–18. DOI: 10.1016/J.TAAP.2018.02.008.

Baratloo, Alireza; Hosseini, Mostafa; Negida, Ahmed; El Ashal, Gehad (2015): Part 1: Simple Definition and Calculation of Accuracy, Sensitivity and Specificity. In *Emergency* 3 (2), pp. 48–49.

Barpe, Deise Raquel; Rosa, Daniela Dornelles; Froehlich, Pedro Eduardo (2010): Pharmacokinetic evaluation of doxorubicin plasma levels in normal and overweight patients with breast cancer and simulation of dose adjustment by different indexes of body mass. In *European journal of pharmaceutical sciences : official journal of the European Federation for Pharmaceutical Sciences* 41 (3-4), pp. 458–463. DOI: 10.1016/j.ejps.2010.07.015.

Beers, Jeanette; Linask, Kaari L.; Chen, Jane A.; Siniscalchi, Lauren I.; Lin, Yongshun; Zheng, Wei et al. (2015): A cost-effective and efficient reprogramming platform for large-scale production of integration-free human induced pluripotent stem cells in chemically defined culture. In *Sci Rep* 5 (1), p. 11319. DOI: 10.1038/srep11319.

Benjamini, Yoav; Hochberg, Yosef (1995): Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. In *Journal of the Royal Statistical Society. Series B (Methodological)* 57 (1), pp. 289–300. Available online at http://www.jstor.org/stable/2346101.

Bennett, Kristin P.; Brown, Elisabeth M.; Santos, Hannah De Los; Poegel, Matthew; Kiehl, Thomas R.; Patton, Evan W. et al. (2019): Identifying Windows of Susceptibility by Temporal Gene Analysis. In *Sci Rep* 9 (1), p. 2740. DOI: 10.1038/s41598-019-39318-8.

References

Blumer, Jeffrey L.; Rothstein, Fred C.; Kaplan, Barbara S.; Yamashita, Toyoko S.; Eshelman, Fred N.; Myers, Carolyn M.; Reed, Michael D. (1985): Pharmacokinetic determination of ranitidine pharmacodynamics in pediatric ulcer disease. In *The Journal of pediatrics* 107 (2), pp. 301–306. DOI: 10.1016/s0022-3476(85)80156-6.

Bondy, Stephen C.; Campbell, Arezoo (2005): Developmental neurotoxicology. In *Journal of neuroscience research* 81 (5), pp. 605–612. DOI: 10.1002/jnr.20589.

Bouchard, Maryse F.; Chevrier, Jonathan; Harley, Kim G.; Kogut, Katherine; Vedar, Michelle; Calderon, Norma et al. (2011): Prenatal exposure to organophosphate pesticides and IQ in 7-year-old children. In *Environmental health perspectives* 119 (8), pp. 1189–1195. DOI: 10.1289/ehp.1003185.

Briggs, Gerald G.; Towers, Craig V.; Freeman, Roger K. (2017): Drugs in pregnancy and lactation. A reference guide to fetal and neonatal risk. Eleventh edition. Philadelphia, PA: Wolters Kluwer. Available online at http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&AN=1857599.

Brindley, C. J. (1989): Overview of recent clinical pharmacokinetic studies with acitretin (Ro 10-1670, etretin). In *Dermatologica* 178 (2), pp. 79–87. DOI: 10.1159/000248397.

Brookfield, Kathleen F.; Su, Felice; Elkomy, Mohammed H.; Drover, David R.; Lyell, Deirdre J.; Carvalho, Brendan (2016): Pharmacokinetics and placental transfer of magnesium sulfate in pregnant women. In *American Journal of Obstetrics and Gynecology* 214 (6), 737.e1-9. DOI: 10.1016/j.ajog.2015.12.060.

Browne, Hannah; Mason, Gerald; Tang, Thomas (2014): Retinoids and pregnancy: an update. In *Obstet Gynecol* 16 (1), pp. 7–11. DOI: 10.1111/tog.12075.

Buesen, Roland; Genschow, Elke; Slawik, Birgitta; Visan, Anke; Spielmann, Horst; Luch, Andreas; Seiler, Andrea (2009): Embryonic stem cell test remastered: comparison between the validated EST and the new molecular FACS-EST for assessing developmental toxicity in vitro. In *Toxicological Sciences* 108 (2), pp. 389–400. DOI: 10.1093/toxsci/kfp012.

Burrow, G. N.; Fisher, D. A.; Larsen, P. R. (1994): Maternal and fetal thyroid function. In *The New England journal of medicine* 331 (16), pp. 1072–1078. DOI: 10.1056/NEJM199410203311608.

Cardoso-Moreira, Margarida; Halbert, Jean; Valloton, Delphine; Velten, Britta; Chen, Chunyan; Shao, Yi et al. (2019): Gene expression across mammalian organ development. In *Nature* 571 (7766), pp. 505–509. DOI: 10.1038/s41586-019-1338-5.

Carr, E. A.; Beierwaltes, W. H.; Raman, G.; Dodson, V. N.; Tanton, J.; Betts, J. S.; Stambaugh, R. A. (1959): The effect of maternal thyroid function of fetal thyroid function and development. In *The Journal of clinical endocrinology and metabolism* 19 (1), pp. 1–18. DOI: 10.1210/jcem-19-1-1.

Carruthers, S. G.; Shoeman, D. W.; Hignite, C. E.; Azarnoff, D. L. (1978): Correlation between plasma diphenhydramine level and sedative and antihistamine effects. In *Clinical Pharmacology and Therapeutics* 23 (4), pp. 375–382. DOI: 10.1002/cpt1978234375.

Cascella, Marco; Vaqar, Sarosh (2021): StatPearls. Hypermagnesemia. Treasure Island (FL).

Celi, Francesco S.; Zemskova, Marina; Linderman, Joyce D.; Babar, Nabeel I.; Skarulis, Monica C.; Csako, Gyorgy et al. (2010): The pharmacodynamic equivalence of levothyroxine and liothyronine: a randomized, double blind, cross-over study in thyroidectomized patients. In *Clinical endocrinology* 72 (5), pp. 709–715. DOI: 10.1111/j.1365-2265.2009.03700.x.

Cezar, Gabriela G.; Quam, Jessica A.; Smith, Alan M.; Rosa, Guilherme J. M.; Piekarczyk, Marian S.; Brown, James F. et al. (2007): Identification of small molecules from human embryonic stem cells using metabolomics. In *Stem Cells and Development* 16 (6), pp. 869–882. DOI: 10.1089/scd.2007.0022.

Chambers, Stuart M.; Fasano, Christopher A.; Papapetrou, Eirini P.; Tomishima, Mark; Sadelain, Michel; Studer, Lorenz (2009): Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. In *Nat Biotechnol* 27 (3), pp. 275–280. DOI: 10.1038/nbt.1529.

Chen, Aimin; Yolton, Kimberly; Rauch, Stephen A.; Webster, Glenys M.; Hornung, Richard; Sjödin, Andreas et al. (2014): Prenatal polybrominated diphenyl ether exposures and neurodevelopment in U.S. children through 5 years of age: the HOME study. In *Environmental health perspectives* 122 (8), pp. 856–862. DOI: 10.1289/ehp.1307562.

Chremos, A. N. (1987): Clinical pharmacology of famotidine: a summary. In *Journal of clinical gastroenterology* 9 Suppl 2, pp. 7–12. DOI: 10.1097/00004836-198707002-00003.

Cilla, Donald D.; Whitfield, Lloyd R.; Gibson, Donald M.; Sedman, Allen J.; Posvar, Edward L. (1996): Multiple-dose pharmacokinetics, pharmacodynamics, and safety of atorvastatin, an inhibitor of HMG-CoA reductase, in healthy subjects. In *Clin Pharmacol Ther* 60 (6), pp. 687–695. DOI: 10.1016/S0009-9236(96)90218-0.

Code of Federal Regulations (2014): Animal Welfare Act 2014, Title 9, chapter 1, subchapter A, parts 1–4.

Colburn, W. A.; Gibson, D. M.; Wiens, R. E.; Hanigan, J. J. (1983): Food increases the bioavailability of isotretinoin. In *Journal of clinical pharmacology* 23 (11-12), pp. 534–539. DOI: 10.1002/j.1552-4604.1983.tb01800.x.

Comptour, Aurélie; Rouzaire, Marion; Belville, Corinne; Bouvier, Damien; Gallot, Denis; Blanchon, Loïc; Sapin, Vincent (2016): Nuclear retinoid receptors and pregnancy: placental transfer, functions, and pharmacological aspects. In *Cellular and molecular life sciences : CMLS* 73 (20), pp. 3823–3837. DOI: 10.1007/s00018-016-2332-9.

Connolly, Roisin M.; Rudek, Michelle A.; Piekarz, Richard (2017): Entinostat: a promising treatment option for patients with advanced breast cancer. In *Future oncology (London, England)* 13 (13), pp. 1137–1148. DOI: 10.2217/fon-2016-0526.

Coons, A. H.; Creech, H. J.; Jones, R. N. (1941): Immunological Properties of an Antibody Containing a Fluorescent Group. In *Experimental Biology and Medicine* 47 (2), pp. 200–202. DOI: 10.3181/00379727-47-13084P.

Cosmetic ingredient review (2008): Final amended report on the safety assessment of Methylparaben, Ethylparaben, Propylparaben, Isopropylparaben, Butylparaben, Isobutylparaben, and Benzylparaben as used in cosmetic products 27 Suppl 4, pp. 1–82. DOI: 10.1080/10915810802548359.

Cowell, Whitney J.; Lederman, Sally A.; Sjödin, Andreas; Jones, Richard; Wang, Shuang; Perera, Frederica P. et al. (2015): Prenatal exposure to polybrominated diphenyl ethers and child attention problems at 3-7 years. In *Neurotoxicology and Teratology* 52 (Pt B), pp. 143–150. DOI: 10.1016/j.ntt.2015.08.009.

Cronstein, Bruce N.; Aune, Thomas M. (2020): Methotrexate and its mechanisms of action in inflammatory arthritis. In *Nature reviews. Rheumatology* 16 (3), pp. 145–154. DOI: 10.1038/s41584-020-0373-9.

Dai, Xin; Guo, Zixuan; Chen, Danfeng; Li, Lu; Song, Xueli; Liu, Tianyu et al. (2020): Maternal sucralose intake alters gut microbiota of offspring and exacerbates hepatic steatosis in adulthood. In *Gut microbes* 11 (4), pp. 1043–1063. DOI: 10.1080/19490976.2020.1738187.

Dalhoff, K.; Poulsen, H. E.; Garred, P.; Placchi, M.; Gammans, R. E.; Mayol, R. F.; Pfeffer, M. (1987): Buspirone pharmacokinetics in patients with cirrhosis. In *British Journal of Clinical Pharmacology* 24 (4), pp. 547–550. DOI: 10.1111/j.1365-2125.1987.tb03210.x.

Dancis, J.; Levitz, M.; Katz, J.; Wilson, D.; Blaner, W. S.; Piantedosi, R.; Goodman, D. S. (1992): Transfer and metabolism of retinol by the perfused human placenta. In *Pediatric research* 32 (2), pp. 195–199. DOI: 10.1203/00006450-199208000-00014.

Davis, Bruce A.; Nagarajan, Anu; Forrest, Lucy R.; Singh, Satinder K. (2016): Mechanism of Paroxetine (Paxil) Inhibition of the Serotonin Transporter. In *Sci Rep* 6, p. 23789. DOI: 10.1038/srep23789.

Depp, Richard; Kind, Allan C.; Kirby, William M.M.; Johnson, Wayne L. (1970): Transplacental passage of methicillin and dicloxacillin into the fetus and amniotic fluid. In *American Journal of Obstetrics and Gynecology* 107 (7), pp. 1054–1057. DOI: 10.1016/0002-9378(70)90628-9.

Dhamodharan, R.; Jordan, M. A.; Thrower, D.; Wilson, L.; Wadsworth, P. (1995): Vinblastine suppresses dynamics of individual microtubules in living interphase cells. In *Molecular biology of the cell* 6 (9), pp. 1215–1229. DOI: 10.1091/mbc.6.9.1215.

Doi, H.; Maruta, H.; Kudoh, I.; Takahashi, Y.; Takano, O.; Ogawa, K. et al. (1991): Placental transfer and effects of famotidine on neonates. In *J Anesth* 5 (3), pp. 276–280. DOI: 10.1007/s0054010050276.

Dolberg, Mette Konow Bøgebjerg; Nielsen, Lars Peter; Dahl, Ronald (2017): Pharmacokinetic Profile of Oral Magnesium Hydroxide. In *Basic & clinical pharmacology & toxicology* 120 (3), pp. 264–269. DOI: 10.1111/bcpt.12642.

Dooley, Christopher M.; Wali, Neha; Sealy, Ian M.; White, Richard J.; Stemple, Derek L.; Collins, John E.; Busch-Nentwich, Elisabeth M. (2019): The gene regulatory basis of genetic compensation during neural crest induction. In *PLoS genetics* 15 (6), e1008213. DOI: 10.1371/journal.pgen.1008213.

Draper, Jonathan S.; Pigott, Christine; Thomson, James A.; Andrews, Peter W. (2002): Surface antigens of human embryonic stem cells: changes upon differentiation in culture. In *Journal of anatomy* 200 (Pt 3), pp. 249–258. DOI: 10.1046/j.1469-7580.2002.00030.x.

Dreser, Nadine; Madjar, Katrin; Holzer, Anna-Katharina; Kapitza, Marion; Scholz, Christopher; Kranaster, Petra et al. (2020): Development of a neural rosette formation assay (RoFA) to identify neurodevelopmental toxicants and to characterize their transcriptome disturbances. In *Arch Toxicol* 94 (1), pp. 151–171. DOI: 10.1007/s00204-019-02612-5.

Druwe, Ingrid; Freudenrich, Theresa M.; Wallace, Kathleen; Shafer, Timothy J.; Mundy, William R. (2016): Comparison of Human Induced Pluripotent Stem Cell-Derived Neurons and Rat Primary Cortical Neurons as In Vitro Models of Neurite Outgrowth. In *Applied In Vitro Toxicology* 2 (1), pp. 26–36. DOI: 10.1089/aivt.2015.0025.

Dussault, J.; Row, V. V.; Lickrish, G.; Volpé, R. (1969): Studies of serum triiodothyronine concentration in maternal and cord blood: transfer of triiodothyronine across the human placenta. In *The Journal of clinical endocrinology and metabolism* 29 (4), pp. 595–603. DOI: 10.1210/jcem-29-4-595.

EC (2009): Regulation (EC) No 1223/2009 of the European Parliament and of the Council of 30 November 2009 on cosmetic products (Text with EEA relevance). Available online at http://data.europa.eu/eli/reg/2009/1223/oj, checked on 10/29/2020.

References

EC (2014a): Commission Regulation (EU) No 358/2014 of 9 April 2014 amending Annexes II and V to Regulation (EC) No 1223/2009 of the European Parliament and of the Council on cosmetic products Text with EEA relevance. Available online at http://data.europa.eu/eli/reg/2014/358/oj, checked on 10/29/2020.

EC (2014b): Consumers: Commission improves safety of cosmetics. European Commission. Available online at https://ec.europa.eu/commission/presscorner/detail/en/IP_14_1051, checked on 10/29/2020.

ECHA (n.d.): Registration dossier 13843. Ethyl 4-hydroxybenzoate. European Chemicals Agency. Available online at https://echa.europa.eu/registration-dossier/-/registered-dossier/13843/1, checked on 10/30/2020.

ECHA (n.d.): Registration dossier 13890. Propyl 4-hydroxybenzoate. European Chemicals Agency. Available online at https://echa.europa.eu/registration-dossier/-/registered-dossier/13890/1, checked on 10/30/2020.

ECHA (n.d.): Registration dossier 14310. Methyl 4-hydroxybenzoate. European Chemicals Agency. Available online at https://echa.europa.eu/registration-dossier/-/registered-dossier/14310/1, checked on 10/30/2020.

ECHA (2006): Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC (Text with EEA relevance)Text with EEA relevance. REACH.

ECHA (2020): Candidate List update: Four new hazardous chemicals to be phased out. ECHA/PR/20/05. European Chemicals Agency. Available online at https://echa.europa.eu/-/candidate-list-update-four-new-hazardous-chemicals-to-be-phased-out, checked on 10/30/2020.

Eckardt, Kathrin; Stahlmann, Ralf (2010): Use of two validated in vitro tests to assess the embryotoxic potential of mycophenolic acid. In *Arch Toxicol* 84 (1), pp. 37–43. DOI: 10.1007/s00204-009-0476-1.

Eckert, Dawid; Buhl, Sandra; Weber, Susanne; Jäger, Richard; Schorle, Hubert (2005): The AP-2 family of transcription factors. In *Genome Biology* 6 (13), p. 246. DOI: 10.1186/gb-2005-6-13-246.

ECVAM (2015): devTOX quick Predict Validation Process. European Centre for the Validation of Alternative Methods. Available online at https://tsar.jrc.ec.europa.eu/test-method/tm2014-04, checked on 3/29/2021.

Emmerson, A. M.; Cox, D. A.; Lees, L. J. (1983): Pharmacokinetics of sulbactam and ampicillin following oral administration of sultamicillin with probenecid. In *European journal of clinical microbiology* 2 (4), pp. 340–344. DOI: 10.1007/BF02019464.

EPA (2012): Mercury Study Report to Congress. Volume V: Health Effects of Mercury and Mercury Compounds. Environmental Protection Agency. Available online at https://www.epa.gov/mercury/mercury-study-report-congress, checked on 1/22/2021.

Eskenazi, Brenda; Chevrier, Jonathan; Rauch, Stephen A.; Kogut, Katherine; Harley, Kim G.; Johnson, Caroline et al. (2013): In utero and childhood polybrominated diphenyl ether (PBDE) exposures and neurodevelopment in the CHAMACOS study. In *Environmental health perspectives* 121 (2), pp. 257–262. DOI: 10.1289/ehp.1205597.

EU (2010): European Directive 2010/63/EU.

Eubig, Paul A.; Aguiar, Andréa; Schantz, Susan L. (2010): Lead and PCBs as risk factors for attention deficit/hyperactivity disorder. In *Environmental health perspectives* 118 (12), pp. 1654–1667. DOI: 10.1289/ehp.0901852.

EU-ToxRisk (2019): DB-ALM Protocol n° 215: Expansion and standard maintenance of SBAD2 human induced Pluripotent Stem Cells (hiPSC). Edited by European Commission, Joint Research Centre (2019): EURL ECVAM dataset on alternative methods to animal experimentation (DB-ALM). Available online at http://cidportal.jrc.ec.europa.eu/ftp/jrc-opendata/EURL-ECVAM/datasets/DBALM/LATEST/online/DBALM_docs/215_P_iPSC_culture.pdf, checked on 9/10/2020.

Evans, M. J.; Kaufman, M. H. (1981): Establishment in culture of pluripotential cells from mouse embryos. In *Nature* 292 (5819), pp. 154–156. DOI: 10.1038/292154a0.

Farina, Marcelo; Rocha, João B. T.; Aschner, Michael (2011): Mechanisms of methylmercury-induced neurotoxicity: evidence from experimental studies. In *Life sciences* 89 (15-16), pp. 555–563. DOI: 10.1016/j.lfs.2011.05.019.

Fayyaz, Susann; Kreiling, Reinhard; Sauer, Ursula G. (2021): Application of grouping and read-across for the evaluation of parabens of different chain lengths with a particular focus on endocrine properties. In *Arch Toxicol* 95 (3), pp. 853–881. DOI: 10.1007/s00204-020-02967-0.

FDA (2013): FDA Drug Safety Communication: Valproate Anti-seizure Products Contraindicated for Migraine Prevention in Pregnant Women due to Decreased IQ Scores in Exposed Children. Food and Drug Administration. Available online at https://www.fda.gov/drugs/drug-safety-and-availability/fda-drug-safety-communication-valproate-anti-seizure-products-contraindicated-migraine-prevention, checked on 11/25/2020.

FDA (2019): Food and Drug Administration, Code of Federal Regulations Title 21. 21CFR172.831, revised 4/1/2019. Available online at https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/cfrsearch.cfm?fr=172.831, checked on 11/24/2020.

Federal Ministry of Justice and Consumer Protection (2015): Tierschutz-Versuchstierordnung. TierSchVersV. Source: BGBl. 2013 Part I, S. 3125, 3126; Revised 31.08.2015, BGBl. 2015 Part I, S. 1474.

Federal Ministry of Justice and Consumer Protection (2020): Tierschutzgesetz. TierSchG. Source: BGBl. 2006 Part I, S. 1206, 1313; Revised 19.06.2020, BGBl. 2020 Part I, S. 1328.

Ferm, V. H. (1963): Congenital malformations in hamster embryos after treatment with vinblastine and vincristine. In *Science (New York, N.Y.)* 141 (3579), p. 426. DOI: 10.1126/science.141.3579.426.

Fielden, Mark R.; Kolaja, Kyle L. (2008): The role of early in vivo toxicity testing in drug discovery toxicology. In *Expert opinion on drug safety* 7 (2), pp. 107–110. DOI: 10.1517/14740338.7.2.107.

Fontana, M.; Massironi, E.; Rossi, A.; Vaglia, P.; Gancia, G. P.; Tagliabue, P.; Principi, N. (1993): Ranitidine pharmacokinetics in newborn infants. In *Archives of disease in childhood* 68 (5 Spec No), pp. 602–603. DOI: 10.1136/adc.68.5_spec_no.602.

Fox, R. I.; Herrmann, M. L.; Frangou, C. G.; Wahl, G. M.; Morris, R. E.; Strand, V.; Kirschbaum, B. J. (1999): Mechanism of action for leflunomide in rheumatoid arthritis. In *Clinical immunology (Orlando, Fla.)* 93 (3), pp. 198–208. DOI: 10.1006/clim.1999.4777.

Fransway, Anthony F.; Fransway, Paulina J.; Belsito, Donald V.; Warshaw, Erin M.; Sasseville, Denis; Fowler, Joseph F. et al. (2019): Parabens. In *Dermatitis* 30 (1), pp. 3–31. DOI: 10.1097/DER.0000000000000429.

Fretham, Stephanie Jb; Caito, Samuel; Martinez-Finley, Ebany J.; Aschner, Michael (2012): Mechanisms and Modifiers of Methylmercury-Induced Neurotoxicity. In *Toxicology Research* 1 (1), pp. 32–38. DOI: 10.1039/C2TX20010D.

Fritsche, Ellen; Cline, Jason E.; Nguyen, Ngoc-Ha; Scanlan, Thomas S.; Abel, Josef (2005): Polychlorinated biphenyls disturb differentiation of normal human neural progenitor cells: clue for involvement of thyroid hormone receptors. In *Environmental health perspectives* 113 (7), pp. 871–876. DOI: 10.1289/ehp.7793.

Fritsche, Ellen; Gassmann, Kathrin; Schreiber, Timm (2011): Neurospheres as a model for developmental neurotoxicity testing. In *Methods in molecular biology (Clifton, N.J.)* 758, pp. 99–114. DOI: 10.1007/978-1-61779-170-3_7.

Frowein, Julia von; Wizenmann, Andrea; Götz, Magdalena (2006): The transcription factors Emx1 and Emx2 suppress choroid plexus development and promote neuroepithelial cell fate. In *Developmental biology* 296 (1), pp. 239–252. DOI: 10.1016/j.ydbio.2006.04.461.

Furuta, Yousuke; Gowen, Brian B.; Takahashi, Kazumi; Shiraki, Kimiyasu; Smee, Donald F.; Barnard, Dale L. (2013): Favipiravir (T-705), a novel viral RNA polymerase inhibitor. In *Antiviral research* 100 (2), pp. 446–454. DOI: 10.1016/j.antiviral.2013.09.015.

Fusaki, Noemi; Ban, Hiroshi; Nishiyama, Akiyo; Saeki, Koichi; Hasegawa, Mamoru (2009): Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. In *Proc. Jpn. Acad., Ser. B* 85 (8), pp. 348–362. DOI: 10.2183/pjab.85.348.

Futschik, Matthias E.; Morkel, Markus; Schäfer, Reinhold; Sers, Christine (2018): The Human Transcriptome. In William B. Coleman, Gregory J. Tsongalis (Eds.): Molecular pathology. The molecular basis of human disease. Second edition. London: Academic Press, pp. 135–164.

Gaballah, Shaza; Swank, Adam; Sobus, Jon R.; Howey, Xia Meng; Schmid, Judith; Catron, Tara et al. (2020): Evaluation of Developmental Toxicity, Developmental Neurotoxicity, and Tissue Dose in Zebrafish Exposed to GenX and Other PFAS. In *Environmental health perspectives* 128 (4), p. 47005. DOI: 10.1289/EHP5843.

Gabdoulline, R.; Kaisers, W.; Gaspar, A.; Meganathan, K.; Doss, M. X.; Jagtap, S. et al. (2015): Differences in the Early Development of Human and Mouse Embryonic Stem Cells. In *PLoS ONE* 10 (10), e0140803. DOI: 10.1371/journal.pone.0140803.

Geer, Laura A.; Pycke, Benny F. G.; Waxenbaum, Joshua; Sherer, David M.; Abulafia, Ovadia; Halden, Rolf U. (2017): Association of birth outcomes with fetal exposure to parabens, triclosan and triclocarban in an immigrant population in Brooklyn, New York. In *Journal of hazardous materials* 323 (Pt A), pp. 177–183. DOI: 10.1016/j.jhazmat.2016.03.028.

Geier, Mitra C.; Chlebowski, Anna C.; Truong, Lisa; Massey Simonich, Staci L.; Anderson, Kim A.; Tanguay, Robert L. (2018): Comparative developmental toxicity of a comprehensive suite of polycyclic aromatic hydrocarbons. In *Arch Toxicol* 92 (2), pp. 571–586. DOI: 10.1007/s00204-017-2068-9.

Gelotte, Cathy K.; Zimmerman, Brenda A.; Thompson, Gary A. (2018): Single-Dose Pharmacokinetic Study of Diphenhydramine HCl in Children and Adolescents. In *Clinical pharmacology in drug development* 7 (4), pp. 400–407. DOI: 10.1002/cpdd.391.

Genentech (2012): Erivedge Product Information. Available online at https://www.accessdata.fda.gov/drugsatfda_docs/label/2012/203388lbl.pdf, checked on 1/21/2021.

References

Genschow, Elke; Spielmann, Horst; Scholz, Gabriele; Seiler, Andrea; Brown, Nigel; Piersma, Aldert et al. (2002): The ECVAM international validation study on in vitro embryotoxicity tests: results of the definitive phase and evaluation of prediction models. European Centre for the Validation of Alternative Methods. In *Alternatives to laboratory animals : ATLA* 30 (2), pp. 151–176. DOI: 10.1177/026119290203000204.

Giachetto, Gustavo; Pirez, María Catalina; Nanni, Luciana; Martínez, Adriana; Montano, Alicia; Algorta, Gabriela et al. (2004): Ampicillin and penicillin concentration in serum and pleural fluid of hospitalized children with community-acquired pneumonia. In *The Pediatric infectious disease journal* 23 (7), pp. 625–629. DOI: 10.1097/01.inf.0000128783.11218.c9.

Glover, Matthew E.; McCoy, Chelsea R.; Shupe, Elizabeth A.; Unroe, Keaton A.; Jackson, Nateka L.; Clinton, Sarah M. (2019): Perinatal exposure to the SSRI paroxetine alters the methylome landscape of the developing dentate gyrus. In *The European journal of neuroscience* 50 (2), pp. 1843–1870. DOI: 10.1111/ejn.14315.

Graham, Richard A.; Hop, Cornelis E. C. A.; Borin, Marie T.; Lum, Bert L.; Colburn, Dawn; Chang, Ilsung et al. (2012): Single and multiple dose intravenous and oral pharmacokinetics of the hedgehog pathway inhibitor vismodegib in healthy female subjects. In *British Journal of Clinical Pharmacology* 74 (5), pp. 788–796. DOI: 10.1111/j.1365-2125.2012.04281.x.

Grandjean, P.; Landrigan, P. J. (2006): Developmental neurotoxicity of industrial chemicals. In *The Lancet* 368 (9553), pp. 2167–2178. DOI: 10.1016/S0140-6736(06)69665-7.

Grandjean, Philippe; Landrigan, Philip J. (2014): Neurobehavioural effects of developmental toxicity. The Lancet Neurology, 13(3), 330-338. In *The Lancet Neurology* 13 (3), pp. 330–338. DOI: 10.1016/S1474-4422(13)70278-3.

Grandjean, Philippe; Weihe, Pal; White, Roberta F.; Debes, Frodi; Araki, Shunichi; Yokoyama, Kazuhito et al. (1997): Cognitive Deficit in 7-Year-Old Children with Prenatal Exposure to Methylmercury. In *Neurotoxicology and Teratology* 19 (6), pp. 417–428. DOI: 10.1016/S0892-0362(97)00097-4.

Greene, R. F.; Collins, J. M.; Jenkins, J. F.; Speyer, J. L.; Myers, C. E. (1983): Plasma pharmacokinetics of adriamycin and adriamycinol: implications for the design of in vitro experiments and treatment protocols. In *Cancer research* 43 (7), pp. 3417–3421.

Harada, M. (1995): Minamata disease: methylmercury poisoning in Japan caused by environmental pollution. In *Critical reviews in toxicology* 25 (1), pp. 1–24. DOI: 10.3109/10408449509089885.

Hayden, Frederick G.; Shindo, Nahoko (2019): Influenza virus polymerase inhibitors in clinical development. In *Current opinion in infectious diseases* 32 (2), pp. 176–186. DOI: 10.1097/QCO.0000000000000532.

He, Yingzi; Cai, Chengfu; Tang, Dongmei; Sun, Shan; Li, Huawei (2014): Effect of histone deacetylase inhibitors trichostatin A and valproic acid on hair cell regeneration in zebrafish lateral line neuromasts. In *Frontiers in cellular neuroscience* 8, p. 382. DOI: 10.3389/fncel.2014.00382.

Heath, Michael S.; Sahni, Dev R.; Curry, Zachary A.; Feldman, Steven R. (2018): Pharmacokinetics of tazarotene and acitretin in psoriasis. In *Expert opinion on drug metabolism & toxicology* 14 (9), pp. 919–927. DOI: 10.1080/17425255.2018.1515198.

Henck, J. W.; Craft, W. R.; Black, A.; Colgin, J.; Anderson, J. A. (1998): Pre- and postnatal toxicity of the HMG-CoA reductase inhibitor atorvastatin in rats. In *Toxicological sciences : an official journal of the Society of Toxicology* 41 (1), pp. 88–99. DOI: 10.1006/toxs.1997.2400.

Henderson, G. I.; Perez, T.; Schenker, S.; Mackins, J.; Antony, A. C. (1995): Maternal-to-fetal transfer of 5-methyltetrahydrofolate by the perfused human placental cotyledon: evidence for a concentrative role by placental folate receptors in fetal folate delivery. In *The Journal of laboratory and clinical medicine* 126 (2), pp. 184–203.

Henderson, J. K.; Draper, J. S.; Baillie, H. S.; Fishel, S.; Thomson, J. A.; Moore, H.; Andrews, P. W. (2002): Preimplantation human embryos and embryonic stem cells show comparable expression of stage-specific embryonic antigens. In *Stem cells (Dayton, Ohio)* 20 (4), pp. 329–337. DOI: 10.1634/stemcells.20-4-329.

Higuchi, R.; Fockler, C.; Dollinger, G.; Watson, R. (1993): Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. In *Bio/technology (Nature Publishing Company)* 11 (9), pp. 1026–1030. DOI: 10.1038/nbt0993-1026.

Hill, Christopher R.; Cole, Michael; Errington, Julie; Malik, Ghada; Boddy, Alan V.; Veal, Gareth J. (2014): Characterisation of the clinical pharmacokinetics of actinomycin D and the influence of ABCB1 pharmacogenetic variation on actinomycin D disposition in children with cancer. In *Clin Pharmacokinet* 53 (8), pp. 741–751. DOI: 10.1007/s40262-014-0153-2.

Hofrichter, Maxi; Nimtz, Laura; Tigges, Julia; Kabiri, Yaschar; Schröter, Friederike; Royer-Pokora, Brigitte et al. (2017): Comparative performance analysis of human iPSC-derived and primary neural progenitor cells (NPC) grown as neurospheres in vitro. In *Stem Cell Research* 25, pp. 72–82. DOI: 10.1016/j.scr.2017.10.013.

Hollstein, Ulrich (1974): Actinomycin. Chemistry and mechanism of action. In *Chem. Rev.* 74 (6), pp. 625–652. DOI: 10.1021/cr60292a002.

Horzmann, Katharine A.; Freeman, Jennifer L. (2018): Making Waves: New Developments in Toxicology With the Zebrafish. In *Toxicological Sciences* 163 (1), pp. 5–12. DOI: 10.1093/toxsci/kfy044.

Howe, Kerstin; Clark, Matthew D.; Torroja, Carlos F.; Torrance, James; Berthelot, Camille; Muffato, Matthieu et al. (2013): The zebrafish reference genome sequence and its relationship to the human genome. In *Nature* 496 (7446), pp. 498–503. DOI: 10.1038/nature12111.

Huang, Irvin J.; Dheilly, Nolwenn M.; Sirotkin, Howard I.; McElroy, Anne E. (2020): Comparative transcriptomics implicate mitochondrial and neurodevelopmental impairments in larval zebrafish (Danio rerio) exposed to two selective serotonin reuptake inhibitors (SSRIs). In *Ecotoxicology and environmental safety* 203, p. 110934. DOI: 10.1016/j.ecoenv.2020.110934.

Huang, S. M.; Athanikar, N. K.; Sridhar, K.; Huang, Y. C.; Chiou, W. L. (1982): Pharmacokinetics of chlorpheniramine after intravenous and oral administration in normal adults. In *European journal of clinical pharmacology* 22 (4), pp. 359–365. DOI: 10.1007/BF00548406.

Hutson, Janine R.; Stade, Brenda; Lehotay, Denis C.; Collier, Christine P.; Kapur, Bhushan M. (2012): Folic acid transport to the human fetus is decreased in pregnancies with chronic alcohol exposure. In *PLoS ONE* 7 (5), e38057. DOI: 10.1371/journal.pone.0038057.

Hyoun, Sara C.; Običan, Sarah G.; Scialli, Anthony R. (2012): Teratogen update: methotrexate. In *Birth defects research. Part A, Clinical and molecular teratology* 94 (4), pp. 187–207. DOI: 10.1002/bdra.23003.

Idres, Nadia; Marill, Julie; Flexor, Maria A.; Chabot, Guy G. (2002): Activation of retinoic acid receptor-dependent transcription by all-trans-retinoic acid metabolites and isomers. In *The Journal of biological chemistry* 277 (35), pp. 31491–31498. DOI: 10.1074/jbc.M205016200.

Iwamoto, Marian; Friedman, Evan J.; Sandhu, Punam; Agrawal, Nancy G. B.; Rubin, Eric H.; Wagner, John A. (2013): Clinical pharmacology profile of vorinostat, a histone deacetylase inhibitor. In *Cancer chemotherapy and pharmacology* 72 (3), pp. 493–508. DOI: 10.1007/s00280-013-2220-z.

Jacobson, J. L.; Jacobson, S. W. (1996): Intellectual impairment in children exposed to polychlorinated biphenyls in utero. In *The New England journal of medicine* 335 (11), pp. 783–789. DOI: 10.1056/NEJM199609123351104.

James, Gareth; Witten, Daniela; Hastie, Trevor; Tibshirani, Robert (2015): An Introduction to Statistical Learning. With Applications in R. New York, NY: Springer (Springer Texts in Statistics, 103).

Jonklaas, Jacqueline; Burman, Kenneth D.; Wang, Hong; Latham, Keith R. (2015): Single-dose T3 administration: kinetics and effects on biochemical and physiological parameters. In *Therapeutic drug monitoring* 37 (1), pp. 110–118. DOI: 10.1097/FTD.0000000000000113.

Kamelia, Lenny; Louisse, Jochem; Haan, Laura de; Rietjens, Ivonne M. C. M.; Boogaard, Peter J. (2017): Prenatal developmental toxicity testing of petroleum substances: Application of the mouse embryonic stem cell test (EST) to compare in vitro potencies with potencies observed in vivo. In *Toxicology in vitro : an international journal published in association with BIBRA* 44, pp. 303–312. DOI: 10.1016/j.tiv.2017.07.018.

Kameoka, Sei; Babiarz, Joshua; Kolaja, Kyle; Chiao, Eric (2014): A high-throughput screen for teratogens using human pluripotent stem cells. In *Toxicological Sciences* 137 (1), pp. 76–90. DOI: 10.1093/toxsci/kft239.

Kandárová, Helena; Hayden, Patrick; Klausner, Mitchell; Kubilus, Joseph; Sheasgreen, John (2009): An in vitro skin irritation test (SIT) using the EpiDerm reconstructed human epidermal (RHE) model. In *JoVE (Journal of Visualized Experiments)* (29), e1366. DOI: 10.3791/1366.

Kang, Hee Young; Choi, Young-Kwon; Jo, Na Rae; Lee, Jae-Hwan; Ahn, Changhwan; Ahn, Il Young et al. (2017): Advanced developmental toxicity test method based on embryoid body's area. In *Reproductive toxicology (Elmsford, N.Y.)* 72, pp. 74–85. DOI: 10.1016/j.reprotox.2017.06.185.

Karagas, Margaret R.; Choi, Anna L.; Oken, Emily; Horvat, Milena; Schoeny, Rita; Kamai, Elizabeth et al. (2012): Evidence on the human health effects of low-level methylmercury exposure. In *Environmental health perspectives* 120 (6), pp. 799–806. DOI: 10.1289/ehp.1104494.

Kashiwagura, Yasuharu; Uchida, Shinya; Tanaka, Shimako; Watanabe, Hiroshi; Masuzawa, Masahiro; Sasaki, Tadanori; Namiki, Noriyuki (2014): Clinical efficacy and pharmacokinetics of levothyroxine suppository in patients with hypothyroidism. In *Biological & pharmaceutical bulletin* 37 (4), pp. 666–670. DOI: 10.1248/bpb.b13-00998.

Kaye, C. M.; Haddock, R. E.; Langley, P. F.; Mellows, G.; Tasker, T. C.; Zussman, B. D.; Greb, W. H. (1989): A review of the metabolism and pharmacokinetics of paroxetine in man. In *Acta psychiatrica Scandinavica. Supplementum* 350, pp. 60–75. DOI: 10.1111/j.1600-0447.1989.tb07176.x.

Khan, Khalid; Wasserman, Gail A.; Liu, Xinhua; Ahmed, Ershad; Parvez, Faruque; Slavkovich, Vesna et al. (2012): Manganese exposure from drinking water and children's academic achievement. In *NeuroToxicology* 33 (1), pp. 91–97. DOI: 10.1016/j.neuro.2011.12.002.

Kim, Byoung-Gwon; Jo, Eun-Mi; Kim, Gyeong-Yeon; Kim, Dae-Seon; Kim, Yu-Mi; Kim, Rock-Bum et al. (2012): Analysis of methylmercury concentration in the blood of Koreans by using cold vapor atomic fluorescence spectrophotometry. In *Annals of laboratory medicine* 32 (1), pp. 31–37. DOI: 10.3343/alm.2012.32.1.31.

References

Kimmel, C. B.; Ballard, W. W.; Kimmel, S. R.; Ullmann, B.; Schilling, T. F. (1995): Stages of embryonic development of the zebrafish. In *Developmental Dynamics* 203 (3), pp. 253–310. DOI: 10.1002/aja.1002030302.

Klaassen, Curtis D. (Ed.) (2018): Toxicology : The Basic Science of Poisons. The basic science of poisons. 9th edition. New York: McGraw-Hill Education.

Kleinstreuer, N. C.; Smith, A. M.; West, P. R.; Conard, K. R.; Fontaine, B. R.; Weir-Hauptman, A. M. et al. (2011): Identifying developmental toxicity pathways for a subset of ToxCast chemicals using human embryonic stem cells and metabolomics. In *Toxicology and Applied Pharmacology* 257 (1), pp. 111–121. DOI: 10.1016/j.taap.2011.08.025.

Koch, S.; Jäger-Roman, E.; Lösche, G.; Nau, H.; Rating, D.; Helge, H. (1996): Antiepileptic drug treatment in pregnancy: drug side effects in the neonate and neurological outcome. In *Acta paediatrica (Oslo, Norway : 1992)* 85 (6), pp. 739–746. DOI: 10.1111/j.1651-2227.1996.tb14137.x.

Kochhar, D. M.; Jiang, H.; Penner, J. D.; Heyman, R. A. (1995): Placental transfer and developmental effects of 9-cis retinoic acid in mice. In *Teratology* 51 (4), pp. 257–265. DOI: 10.1002/tera.1420510411.

Köppel, C.; Ibe, K.; Tenczer, J. (1987): Clinical symptomatology of diphenhydramine overdose: an evaluation of 136 cases in 1982 to 1985. In *Journal of toxicology. Clinical toxicology* 25 (1-2), pp. 53–70. DOI: 10.3109/15563658708992613.

Krug, Anne K.; Kolde, Raivo; Gaspar, John A.; Rempel, Eugen; Balmer, Nina V.; Meganathan, Kesavan et al. (2013): Human embryonic stem cell-derived test systems for developmental neurotoxicity: a transcriptomics approach. In *Arch Toxicol* 87 (1), pp. 123–143. DOI: 10.1007/s00204-012-0967-3.

Kundu, Nabanita; Domingues, Cleyton C.; Patel, Jay; Aljishi, Mohammed; Ahmadi, Neeki; Fakhri, Mona et al. (2020): Sucralose promotes accumulation of reactive oxygen species (ROS) and adipogenesis in mesenchymal stromal cells. In *Stem Cell Research & Therapy* 11 (1), p. 250. DOI: 10.1186/s13287-020-01753-0.

Lamberg, T. S.; Kivistö, K. T.; Neuvonen, P. J. (1998): Concentrations and effects of buspirone are considerably reduced by rifampicin. In *British Journal of Clinical Pharmacology* 45 (4), pp. 381–385. DOI: 10.1046/j.1365-2125.1998.t01-1-00698.x.

Landrigan, PhilipJ; Baloh, RobertW; Barthel, WilliamF; Whitworth, RandolphH; Staehling, NormanW; Rosenblum, BernardF (1975): NEUROPSYCHOLOGICAL DYSFUNCTION IN CHILDREN WITH CHRONIC LOW-LEVEL LEAD ABSORPTION. In *The Lancet* 305 (7909), pp. 708–712. DOI: 10.1016/S0140-6736(75)91627-X.

Lanphear, Bruce P. (2015): The impact of toxins on the developing brain. In *Annual Review of Public Health* 36 (1), pp. 211–230. DOI: 10.1146/annurev-publhealth-031912-114413.

Lanphear, Bruce P.; Hornung, Richard; Khoury, Jane; Yolton, Kimberly; Baghurst, Peter; Bellinger, David C. et al. (2005): Low-level environmental lead exposure and children's intellectual function: an international pooled analysis. In *Environmental health perspectives* 113 (7), pp. 894–899. DOI: 10.1289/ehp.7688.

Lecarpentier, Edouard; Morel, Olivier; Fournier, Thierry; Elefant, Elisabeth; Chavatte-Palmer, Pascale; Tsatsaris, Vassilis (2012): Statins and pregnancy: between supposed risks and theoretical benefits. In *Drugs* 72 (6), pp. 773–788. DOI: 10.2165/11632010-000000000-00000.

Lepper, Erin R.; Smith, Nicola F.; Cox, Michael C.; Scripture, Charity D.; Figg, William D. (2006): Thalidomide metabolism and hydrolysis: mechanisms and implications. In *CDM* 7 (6), pp. 677–685. DOI: 10.2174/138920006778017777.

Levine, M.; Conry-Cantilena, C.; Wang, Y.; Welch, R. W.; Washko, P. W.; Dhariwal, K. R. et al. (1996): Vitamin C pharmacokinetics in healthy volunteers: evidence for a recommended dietary allowance. In *Proceedings of the National Academy of Sciences of the United States of America* 93 (8), pp. 3704–3709. DOI: 10.1073/pnas.93.8.3704.

Li, Shuaizhang; Xia, Menghang (2019): Review of high-content screening applications in toxicology. In *Arch Toxicol* 93 (12), pp. 3387–3396. DOI: 10.1007/s00204-019-02593-5.

Links, M.; Watson, S.; Lethlean, K.; Aherne, W.; Kirsten, F.; Clarke, S. et al. (1999): Vinblastine pharmacokinetics in patients with non-small cell lung cancer given cisplatin. In *Cancer investigation* 17 (7), pp. 479–485. DOI: 10.3109/07357909909032857.

Liu, Chih-Wei; Chi, Liang; Tu, Pengcheng; Xue, Jingchuan; Ru, Hongyu; Lu, Kun (2019): Quantitative proteomics reveals systematic dysregulations of liver protein metabolism in sucralose-treated mice. In *Journal of proteomics* 196, pp. 1–10. DOI: 10.1016/j.jprot.2019.01.011.

Livak, K. J.; Schmittgen, T. D. (2001): Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. In *Methods (San Diego, Calif.)* 25 (4), pp. 402–408. DOI: 10.1006/meth.2001.1262.

Looijenga, Leendert H. J.; Stoop, Hans; Leeuw, Hubert P. J. C. de; Gouveia Brazao, Carlos A. de; Gillis, Ad J. M.; van Roozendaal, Kees E. P. et al. (2003): POU5F1 (OCT3/4) identifies cells with pluripotent potential in human germ cell tumors. In *Cancer research* 63 (9), pp. 2244–2250.

Lu, K.; Yap, H. Y.; Loo, T. L. (1983): Clinical pharmacokinetics of vinblastine by continuous intravenous infusion. In *Cancer research* 43 (3), pp. 1405–1408.

Lucchini, Roberto G.; Guazzetti, Stefano; Zoni, Silvia; Donna, Filippo; Peter, Stephanie; Zacco, Annalisa et al. (2012): Tremor, olfactory and motor changes in Italian adolescents exposed to historical ferro-manganese emission. In *NeuroToxicology* 33 (4), pp. 687–696. DOI: 10.1016/j.neuro.2012.01.005.

Luna, B. G.; Scavone, J. M.; Greenblatt, D. J. (1989): Doxylamine and diphenhydramine pharmacokinetics in women on low-dose estrogen oral contraceptives. In *Journal of clinical pharmacology* 29 (3), pp. 257–260. DOI: 10.1002/j.1552-4604.1989.tb03323.x.

Luz, Anthony L.; Tokar, Erik J. (2018): Pluripotent Stem Cells in Developmental Toxicity Testing: A Review of Methodological Advances. In *Toxicological Sciences* 165 (1), pp. 31–39. DOI: 10.1093/toxsci/kfy174.

Madelain, Vincent; Nguyen, Thi Huyen Tram; Olivo, Anaelle; Lamballerie, Xavier de; Guedj, Jérémie; Taburet, Anne-Marie; Mentré, France (2016): Ebola Virus Infection: Review of the Pharmacokinetic and Pharmacodynamic Properties of Drugs Considered for Testing in Human Efficacy Trials. In *Clin Pharmacokinet* 55 (8), pp. 907–923. DOI: 10.1007/s40262-015-0364-1.

Madjunkova, Svetlana; Maltepe, Caroline; Koren, Gideon (2014): The delayed-release combination of doxylamine and pyridoxine (Diclegis®/Diclectin ®) for the treatment of nausea and vomiting of pregnancy. In *Pediatr Drugs* 16 (3), pp. 199–211. DOI: 10.1007/s40272-014-0065-5.

Mahmood, I.; Sahajwalla, C. (1999): Clinical pharmacokinetics and pharmacodynamics of buspirone, an anxiolytic drug. In *Clinical Pharmacokinetics* 36 (4), pp. 277–287. DOI: 10.2165/00003088-199936040-00003.

Malhi, Gin S.; Tanious, Michelle; Das, Pritha; Coulston, Carissa M.; Berk, Michael (2013): Potential mechanisms of action of lithium in bipolar disorder. Current understanding. In *CNS drugs* 27 (2), pp. 135–153. DOI: 10.1007/s40263-013-0039-0.

Manolescu, Daniel C.; El-Kares, Reyhan; Lakhal-Chaieb, Lajmi; Montpetit, Alexandre; Bhat, Pangala V.; Goodyer, Paul (2010): Newborn serum retinoic acid level is associated with variants of genes in the retinol metabolism pathway. In *Pediatric research* 67 (6), pp. 598–602. DOI: 10.1203/PDR.0b013e3181dcf18a.

Matalon, S.; Schechtman, S.; Goldzweig, G.; Ornoy, A. (2002): The teratogenic effect of carbamazepine: a meta-analysis of 1255 exposures. In *Reproductive Toxicology* 16 (1), pp. 9–17. DOI: 10.1016/S0890-6238(01)00199-X.

Matwiejczuk, Natalia; Galicka, Anna; Brzóska, Małgorzata M. (2020): Review of the safety of application of cosmetic products containing parabens. In *Journal of Applied Toxicology* 40 (1), pp. 176–210. DOI: 10.1002/jat.3917.

McAuley, D. M.; Moore, J.; McCaughey, W.; Donnelly, B. D.; Dundee, J. W. (1983): Ranitidine as an antacid before elective Caesarean section. In *Anaesthesia* 38 (2), pp. 108–114. DOI: 10.1111/j.1365-2044.1983.tb13927.x.

McCall, Matthew N.; Bolstad, Benjamin M.; Irizarry, Rafael A. (2010): Frozen robust multiarray analysis (fRMA). In *Biostatistics (Oxford, England)* 11 (2), pp. 242–253. DOI: 10.1093/biostatistics/kxp059.

McLeod, M. J. (1980): Differential staining of cartilage and bone in whole mouse fetuses by alcian blue and alizarin red S. In *Teratology* 22 (3), pp. 299–301. DOI: 10.1002/tera.1420220306.

Meganathan, Kesavan; Jagtap, Smita; Wagh, Vilas; Winkler, Johannes; Gaspar, John Antonydas; Hildebrand, Diana et al. (2012): Identification of thalidomide-specific transcriptomics and proteomics signatures during differentiation of human embryonic stem cells. In *PLoS ONE* 7 (8), e44228. DOI: 10.1371/journal.pone.0044228.

Meigs, Lucy; Smirnova, Lena; Rovida, Costanza; Leist, Marcel; Hartung, Thomas (2018): Animal testing and its alternatives - the most important omics is economics. In *ALTEX* 35 (3), pp. 275–305. DOI: 10.14573/altex.1807041.

Merck Sharp & Dohme (2018): Zolinza Prescribing Information. Available online at https://www.merck.com/product/usa/pi_circulars/z/zolinza/zolinza_pi.pdf, checked on 1/20/2021.

Michaëlsson, Karl; Lithell, Hans; Vessby, Bengt; Melhus, Håkan (2003): Serum retinol levels and the risk of fracture. In *New England Journal of Medicine* 348 (4), pp. 287–294. DOI: 10.1056/NEJMoa021171.

Mnatsakanyan, Hayk; Salmerón-Sánchez, Manuel; Rico, Patricia (2020): Lithium Directs Embryonic Stem Cell Differentiation into Hemangioblast-Like Cells. In *SSRN Electronic Journal*. DOI: 10.2139/ssrn.3701258.

Moon, Byoung-San; Lu, Wange; Park, Hong Ju (2018): Valproic acid promotes the neuronal differentiation of spiral ganglion neural stem cells with robust axonal growth. In *Biochemical and Biophysical Research Communications* 503 (4), pp. 2728–2735. DOI: 10.1016/j.bbrc.2018.08.032.

Moors, Michaela; Cline, Jason E.; Abel, Josef; Fritsche, Ellen (2007): ERK-dependent and -independent pathways trigger human neural progenitor cell migration. In *Toxicology and Applied Pharmacology* 221 (1), pp. 57–67. DOI: 10.1016/j.taap.2007.02.018.

Moreno, Julie A.; Yeomans, Elizabeth C.; Streifel, Karin M.; Brattin, Bryan L.; Taylor, Robert J.; Tjalkens, Ronald B. (2009): Age-dependent susceptibility to manganese-induced neurological dysfunction. In *Toxicological Sciences* 112 (2), pp. 394–404. DOI: 10.1093/toxsci/kfp220.

References

Morgan, M. Y.; Stambuk, D.; Cottrell, J.; Mann, S. G. (1990): Pharmacokinetics of famotidine in normal subjects and in patients with chronic liver disease. In *Alimentary pharmacology & therapeutics* 4 (1), pp. 83–96. DOI: 10.1111/j.1365-2036.1990.tb00452.x.

Morinello, Eric; Pignatello, Michael; Villabruna, Loris; Goelzer, Petra; Bürgin, Heinrich (2014): Embryofetal development study of vismodegib, a hedgehog pathway inhibitor, in rats. In *Birth defects research. Part B, Developmental and reproductive toxicology* 101 (2), pp. 135–143. DOI: 10.1002/bdrb.21093.

Morrison, Michael; Klein, Christine; Clemann, Nicole; Collier, David A.; Hardy, John; Heisserer, Barbara et al. (2015): StemBANCC: Governing Access to Material and Data in a Large Stem Cell Research Consortium. In *Stem Cell Rev and Rep* 11 (5), pp. 681–687. DOI: 10.1007/s12015-015-9599-3.

Mortensen, Mary E.; Caudill, Samuel P.; Caldwell, Kathleen L.; Ward, Cynthia D.; Jones, Robert L. (2014): Total and methyl mercury in whole blood measured for the first time in the U.S. population: NHANES 2011-2012. In *Environmental Research* 134, pp. 257–264. DOI: 10.1016/j.envres.2014.07.019.

Mulla, H.; Yakkundi, S.; McElnay, J.; Lutsar, I.; Metsvaht, T.; Varendi, H. et al. (2015): An observational study of blood concentrations and kinetics of methyl- and propyl-parabens in neonates. In *Pharmaceutical research* 32 (3), pp. 1084–1093. DOI: 10.1007/s11095-014-1520-2.

Nau, H. (1986): Species differences in pharmacokinetics and drug teratogenesis. In *Environmental health perspectives* 70, pp. 113–129. DOI: 10.1289/ehp.8670113.

Nau, H. (2001): Teratogenicity of isotretinoin revisited: species variation and the role of all-trans-retinoic acid. In *Journal of the American Academy of Dermatology* 45 (5), S183-7. DOI: 10.1067/mjd.2001.113720.

Nau, H.; Kuhnz, W.; Egger, H. J.; Rating, D.; Helge, H. (1982): Anticonvulsants during pregnancy and lactation. Transplacental, maternal and neonatal pharmacokinetics. In *Clinical Pharmacokinetics* 7 (6), pp. 508–543. DOI: 10.2165/00003088-198207060-00003.

Needleman, H. L.; Gunnoe, C.; Leviton, A.; Reed, R.; Peresie, H.; Maher, C.; Barrett, P. (1979): Deficits in psychologic and classroom performance of children with elevated dentine lead levels. In *The New England journal of medicine* 300 (13), pp. 689–695. DOI: 10.1056/NEJM197903293001301.

Nelson, R. L.; Dyke, R. W.; Root, M. A. (1980): Comparative pharmacokinetics of vindesine, vincristine and vinblastine in patients with cancer. In *Cancer Treatment Reviews* 7, pp. 17–24. DOI: 10.1016/S0305-7372(80)80003-X.

Neubig, Richard R.; Spedding, Michael; Kenakin, Terry; Christopoulos, Arthur (2003): International Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification. XXXVIII. Update on terms and symbols in quantitative pharmacology. In *Pharmacological reviews* 55 (4), pp. 597–606. DOI: 10.1124/pr.55.4.4.

Nguyen, Patricia; Boskovic, Rada; Yazdani, Parvaneh; Kapur, Bhushan; Vandenberghe, Hilde; Koren, Gideon (2008): Comparing folic acid pharmacokinetics among women of childbearing age: single dose ingestion of 1.1 versus 5 MG folic acid. In *The Canadian journal of clinical pharmacology = Journal canadien de pharmacologie clinique* 15 (2), e314-22.

Novartis Pharmaceuticals (2016): Farydak Prescribing Information. Available online at https://www.accessdata.fda.gov/drugsatfda_docs/label/2015/205353s000lbl.pdf, checked on 1/20/2021.

Nowak, Karolina; Ratajczak-Wrona, Wioletta; Górska, Maria; Jabłońska, Ewa (2018): Parabens and their effects on the endocrine system. In *Molecular and Cellular Endocrinology* 474, pp. 238–251. DOI: 10.1016/j.mce.2018.03.014.

NRC; IOM (2000): Dietary reference intakes for vitamin C, vitamin E, selenium, and carotenoids. A report of the Panel on Dietary Antioxidants and Related Compounds, Subcommittees on Upper Reference Levels of Nutrients and of Interpretation and Use of Dietary Reference Intakes, and the Standing Committee on the Scientific Evaluation of Dietary Reference Intakes, Food and Nutrition Board, Institute of Medicine. Washington, D.C: National Academy Press. Available online at http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=112083.

Obeid, Rima; Kasoha, Mariz; Kirsch, Susanne H.; Munz, Winfried; Herrmann, Wolfgang (2010): Concentrations of unmetabolized folic acid and primary folate forms in pregnant women at delivery and in umbilical cord blood. In *The American journal of clinical nutrition* 92 (6), pp. 1416–1422. DOI: 10.3945/ajcn.2010.29361.

O'Brien, J.; Wilson, I.; Orton, T.; Pognan, F. (2000): Investigation of the Alamar Blue (resazurin) fluorescent dye for the assessment of mammalian cell cytotoxicity. In *European journal of biochemistry* 267 (17), pp. 5421–5426. DOI: 10.1046/j.1432-1327.2000.01606.x.

OECD (1983): Test No. 415: One-Generation Reproduction Toxicity Study. Paris: OECD Publishing (OECD Guidelines for the Testing of Chemicals, Section 4).

OECD (1996): Test No. 422: Combined Repeated Dose Toxicity Study with the Reproduction/Developmental Toxicity Screening Test: OECD.

OECD (2007): Test No. 426: Developmental Neurotoxicity Study. Paris: OECD Publishing (OECD Guidelines for the Testing of Chemicals, Section 4).

OECD (2012): Test No. 443. Extended One-Generation Reproductive Toxicity Study. [Place of publication not identified]: [publisher not identified].

OECD (2015): Test No. 439: In Vitro Skin Irritation: Reconstructed Human Epidermis Test Method. Paris: OECD Publishing (OECD Guidelines for the Testing of Chemicals, Section 4).

OECD (2016): Test No. 421: Reproduction/Developmental Toxicity Screening Test. Paris: OECD Publishing (OECD Guidelines for the Testing of Chemicals, Section 4).

OECD (2018a): Test No. 414: Prenatal Development Toxicity Study. Paris: OECD Publishing (OECD Guidelines for the Testing of Chemicals, Section 4).

OECD (2018b): Two-Generation Reproduction Toxicity Study (OECD TG 416).

Ornoy, Asher (2009): Valproic acid in pregnancy: how much are we endangering the embryo and fetus? In *Reproductive toxicology (Elmsford, N.Y.)* 28 (1), pp. 1–10. DOI: 10.1016/j.reprotox.2009.02.014.

Osumi, Noriko; Shinohara, Hiroshi; Numayama-Tsuruta, Keiko; Maekawa, Motoko (2008): Concise review: Pax6 transcription factor contributes to both embryonic and adult neurogenesis as a multifunctional regulator. In *STEM CELLS* 26 (7), pp. 1663–1672. DOI: 10.1634/stemcells.2007-0884.

Pacifici, G. M. (2006): Placental transfer of antibiotics administered to the mother: a review. In *International journal of clinical pharmacology and therapeutics* 44 (2), pp. 57–63. DOI: 10.5414/cpp44057.

Padayatty, Sebastian J.; Sun, He; Wang, Yaohui; Riordan, Hugh D.; Hewitt, Stephen M.; Katz, Arie et al. (2004): Vitamin C pharmacokinetics: implications for oral and intravenous use. In *Annals of internal medicine* 140 (7), pp. 533–537. DOI: 10.7326/0003-4819-140-7-200404060-00010.

Palmer, Jessica A.; Smith, Alan M.; Egnash, Laura A.; Colwell, Michael R.; Donley, Elizabeth L. R.; Kirchner, Fred R.; Burrier, Robert E. (2017): A human induced pluripotent stem cell-based in vitro assay predicts developmental toxicity through a retinoic acid receptor-mediated pathway for a series of related retinoid analogues. In *Reproductive toxicology (Elmsford, N.Y.)* 73, pp. 350–361. DOI: 10.1016/j.reprotox.2017.07.011.

Palmer, Jessica A.; Smith, Alan M.; Egnash, Laura A.; Conard, Kevin R.; West, Paul R.; Burrier, Robert E. et al. (2013): Establishment and assessment of a new human embryonic stem cell-based biomarker assay for developmental toxicity screening. In *Birth defects research. Part B, Developmental and reproductive toxicology* 98 (4), pp. 343–363. DOI: 10.1002/bdrb.21078.

Pan, Guangjin; Thomson, James A. (2007): Nanog and transcriptional networks in embryonic stem cell pluripotency. In *Cell research* 17 (1), pp. 42–49. DOI: 10.1038/sj.cr.7310125.

Panretin®, Stiefel Laboratories: Panretin(R) product information. Available online at https://www.accessdata.fda.gov/drugsatfda_docs/label/1999/20886lbl.pdf, checked on 11/25/2020.

Panzica-Kelly, Julieta M.; Brannen, Kimberly C.; Ma, Yan; Zhang, Cindy X.; Flint, Oliver P.; Lehman-McKeeman, Lois D.; Augustine-Rauch, Karen A. (2013): Establishment of a molecular embryonic stem cell developmental toxicity assay. In *Toxicological Sciences* 131 (2), pp. 447–457. DOI: 10.1093/toxsci/kfs293.

Paquette, Jennifer A.; Kumpf, Steven W.; Streck, Randal D.; Thomson, Jason J.; Chapin, Robert E.; Stedman, Donald B. (2008): Assessment of the Embryonic Stem Cell Test and application and use in the pharmaceutical industry. In *Birth defects research. Part B, Developmental and reproductive toxicology* 83 (2), pp. 104–111. DOI: 10.1002/bdrb.20148.

Pesce, M.; Schöler, H. R. (2001): Oct-4: gatekeeper in the beginnings of mammalian development. In *Stem cells (Dayton, Ohio)* 19 (4), pp. 271–278. DOI: 10.1634/stemcells.19-4-271.

Pfizer Ireland Pharmaceuticals (2009): Lipitor prescribing information. Available online at https://www.accessdata.fda.gov/drugsatfda_docs/label/2009/020702s056lbl.pdf, checked on 11/25/2020.

Pharmaceutical and Food Safety Bureau (2014): Report on the Deliberation Results for Avigan Tablet 200 mg. Available online at https://www.pmda.go.jp/files/000210319.pdf, checked on 11/27/2020.

Pilkington, Tania; Brogden, Rex N. (1992): Acitretin. In *Drugs* 43 (4), pp. 597–627. DOI: 10.2165/00003495-199243040-00010.

Pistollato, Francesca; Canovas-Jorda, David; Zagoura, Dimitra; Price, Anna (2017): Protocol for the Differentiation of Human Induced Pluripotent Stem Cells into Mixed Cultures of Neurons and Glia for Neurotoxicity Testing. In *JoVE (Journal of Visualized Experiments)* (124). DOI: 10.3791/55702.

Pistollato, Francesca; Louisse, Jochem; Scelfo, Bibiana; Mennecozzi, Milena; Accordi, Benedetta; Basso, Giuseppe et al. (2014): Development of a pluripotent stem cell derived neuronal model to identify chemically induced pathway perturbations in relation to neurotoxicity: effects of CREB pathway inhibition. In *Toxicology and Applied Pharmacology* 280 (2), pp. 378–388. DOI: 10.1016/j.taap.2014.08.007.

References

Pope, Laura E.; Khalil, M. H.; Berg, James E.; Stiles, Mark; Yakatan, Gerald J.; Sellers, Edward M. (2004): Pharmacokinetics of dextromethorphan after single or multiple dosing in combination with quinidine in extensive and poor metabolizers. In *Journal of clinical pharmacology* 44 (10), pp. 1132–1142. DOI: 10.1177/0091270004269521.

Posvar, E. L.; Radulovic, L. L.; Cilla, D. D.; Whitfield, L. R.; Sedman, A. J. (1996): Tolerance and pharmacokinetics of single-dose atorvastatin, a potent inhibitor of HMG-CoA reductase, in healthy subjects. In *Journal of clinical pharmacology* 36 (8), pp. 728–731. DOI: 10.1002/j.1552-4604.1996.tb04242.x.

Puelles, Eduardo; Annino, Alessandro; Tuorto, Francesca; Usiello, Alessandro; Acampora, Dario; Czerny, Thomas et al. (2004): Otx2 regulates the extent, identity and fate of neuronal progenitor domains in the ventral midbrain. In *Development (Cambridge, England)* 131 (9), pp. 2037–2048. DOI: 10.1242/dev.01107.

Pycke, Benny F. G.; Geer, Laura A.; Dalloul, Mudar; Abulafia, Ovadia; Halden, Rolf U. (2015): Maternal and fetal exposure to parabens in a multiethnic urban U.S. population. In *Environment International* 84, pp. 193–200. DOI: 10.1016/j.envint.2015.08.012.

Pynnönen, S.; Kanto, J.; Sillanpää, M.; Erkkola, R. (1977): Carbamazepine: placental transport, tissue concentrations in foetus and newborn, and level in milk. In *Acta pharmacologica et toxicologica* 41 (3), pp. 244–253. DOI: 10.1111/j.1600-0773.1977.tb02145.x.

Ramos-Vara, J. A. (2005): Technical aspects of immunohistochemistry. In *Veterinary pathology* 42 (4), pp. 405–426. DOI: 10.1354/vp.42-4-405.

Ramoz, Leda L.; Patel-Shori, Nima M. (2014): Recent changes in pregnancy and lactation labeling: retirement of risk categories. In *Pharmacotherapy* 34 (4), pp. 389–395. DOI: 10.1002/phar.1385.

Rauh, Virginia; Arunajadai, Srikesh; Horton, Megan; Perera, Frederica; Hoepner, Lori; Barr, Dana B.; Whyatt, Robin (2011): Seven-year neurodevelopmental scores and prenatal exposure to chlorpyrifos, a common agricultural pesticide. In *Environmental health perspectives* 119 (8), pp. 1196–1201. DOI: 10.1289/ehp.1003160.

Rauh, Virginia A.; Perera, Frederica P.; Horton, Megan K.; Whyatt, Robin M.; Bansal, Ravi; Hao, Xuejun et al. (2012): Brain anomalies in children exposed prenatally to a common organophosphate pesticide. In *PNAS* 109 (20), pp. 7871–7876. DOI: 10.1073/pnas.1203396109.

Rempel, Eugen; Hoelting, Lisa; Waldmann, Tanja; Balmer, Nina V.; Schildknecht, Stefan; Grinberg, Marianna et al. (2015): A transcriptome-based classifier to identify developmental toxicants by stem cell testing: design, validation and optimization for histone deacetylase inhibitors. In *Arch Toxicol* 89 (9), pp. 1599–1618. DOI: 10.1007/s00204-015-1573-y.

Repetto, M. R.; Repetto, M. (1998): Therapeutic, toxic, and lethal concentrations of 73 drugs affecting respiratory system in human fluids. In *Journal of toxicology. Clinical toxicology* 36 (4), pp. 287–293. DOI: 10.3109/15563659809028023.

Rice, D.; Barone, S. (2000): Critical periods of vulnerability for the developing nervous system: evidence from humans and animal models. In *Environmental health perspectives* 108 Suppl 3 (Suppl 3), pp. 511–533. DOI: 10.1289/ehp.00108s3511.

Richards, D. A. (1983): Comparative pharmacodynamics and pharmacokinetics of cimetidine and ranitidine. In *Journal of clinical gastroenterology* 5 Suppl 1, pp. 81–90. DOI: 10.1097/00004836-198312001-00008.

Riojas-Rodríguez, Horacio; Solís-Vivanco, Rodolfo; Schilmann, Astrid; Montes, Sergio; Rodríguez, Sandra; Ríos, Camilo; Rodríguez-Agudelo, Yaneth (2010): Intellectual function in Mexican children living in a mining area and environmentally exposed to manganese. In *Environmental health perspectives* 118 (10), pp. 1465–1470. DOI: 10.1289/ehp.0901229.

Ritchie, Matthew E.; Phipson, Belinda; Di Wu; Hu, Yifang; Law, Charity W.; Shi, Wei; Smyth, Gordon K. (2015): limma powers differential expression analyses for RNA-sequencing and microarray studies. In *Nucleic Acids Research* 43 (7), e47. DOI: 10.1093/nar/gkv007.

Robert A. Young (1992): Formal Toxicity Summary for methyl mercury. Available online at https://rais.ornl.gov/tox/profiles/methyl_mercury_f_V1.html, updated on 1998, checked on 2/5/2021.

Roberts, A.; Renwick, A. G.; Sims, J.; Snodin, D. J. (2000): Sucralose metabolism and pharmacokinetics in man. In *Food and Chemical Toxicology* 38, pp. 31–41. DOI: 10.1016/S0278-6915(00)00026-0.

Rodier, P. M. (1995): Developing brain as a target of toxicity. In *Environmental health perspectives* 103 Suppl 6, pp. 73–76. DOI: 10.1289/ehp.95103s673.

Rogawski, Michael A.; Löscher, Wolfgang; Rho, Jong M. (2016): Mechanisms of Action of Antiseizure Drugs and the Ketogenic Diet. In *Cold Spring Harbor perspectives in medicine* 6 (5). DOI: 10.1101/cshperspect.a022780.

Roullet, F. I.; Wollaston, L.; Decatanzaro, D.; Foster, J. A. (2010): Behavioral and molecular changes in the mouse in response to prenatal exposure to the anti-epileptic drug valproic acid. In *Neuroscience* 170 (2), pp. 514–522. DOI: 10.1016/j.neuroscience.2010.06.069.

Roux, Magali; Dosseto, Anthony (2017): From direct to indirect lithium targets: a comprehensive review of omics data. In *Metallomics : integrated biometal science* 9 (10), pp. 1326–1351. DOI: 10.1039/c7mt00203c.

Rovida, Costanza; Longo, Fabiola; Rabbit, Richard R. (2011): How are reproductive toxicity and developmental toxicity addressed in REACH dossiers? In *ALTEX* 28 (4), pp. 273–294. DOI: 10.14573/altex.2011.4.273.

Rozman, Blaz (2002): Clinical pharmacokinetics of leflunomide. In *Clinical Pharmacokinetics* 41 (6), pp. 421–430. DOI: 10.2165/00003088-200241060-00003.

Russell W and Burch R (1959): The principles of humane experimental technique. London: Methuen & Co.

Ryan, Kristen R.; Sirenko, Oksana; Parham, Fred; Hsieh, Jui-Hua; Cromwell, Evan F.; Tice, Raymond R.; Behl, Mamta (2016): Neurite outgrowth in human induced pluripotent stem cell-derived neurons as a high-throughput screen for developmental neurotoxicity or neurotoxicity. In *NeuroToxicology* 53, pp. 271–281. DOI: 10.1016/j.neuro.2016.02.003.

SCF (2000): Opinion of the Scientific Committee on Food on sucralose (Adopted by the SCF on 7 September 2000). SCF/CS/ADDS/EDUL/190 Final. Scientific Committee on Food (SCF). Available online at https://ec.europa.eu/food/sites/food/files/safety/docs/sci-com_scf_out68_en.pdf, updated on 12/9/2000, checked on 11/26/2020.

Schadel, M.; Wu, D.; Otton, S. V.; Kalow, W.; Sellers, E. M. (1995): Pharmacokinetics of dextromethorphan and metabolites in humans: influence of the CYP2D6 phenotype and quinidine inhibition. In *Journal of clinical psychopharmacology* 15 (4), pp. 263–269. DOI: 10.1097/00004714-199508000-00005.

Schantz, Susan L.; Widholm, John J.; Rice, Deborah C. (2003): Effects of PCB exposure on neuropsychological function in children. In *Environmental health perspectives* 111 (3), pp. 357–576. DOI: 10.1289/ehp.5461.

Schlaeger, Thorsten M.; Daheron, Laurence; Brickler, Thomas R.; Entwisle, Samuel; Chan, Karrie; Cianci, Amelia et al. (2015): A comparison of non-integrating reprogramming methods. In *Nat Biotechnol* 33 (1), pp. 58–63. DOI: 10.1038/nbt.3070.

Schmitt-Hoffmann, A. H.; Roos, B.; Schoetzau, A.; Leese, P. T.; Meyer, I.; van de Wetering, J.; Kovacs, P. (2012): Oral alitretinoin: a review of the clinical pharmacokinetics and pharmacodynamics. In *Expert review of clinical pharmacology* 5 (4), pp. 373–388. DOI: 10.1586/ECP.12.26.

Schmuck, Martin R.; Temme, Thomas; Dach, Katharina; Boer, Denise de; Barenys, Marta; Bendt, Farina et al. (2017): Omnisphero: a high-content image analysis (HCA) approach for phenotypic developmental neurotoxicity (DNT) screenings of organoid neurosphere cultures in vitro. In *Arch Toxicol* 91 (4), pp. 2017–2028. DOI: 10.1007/s00204-016-1852-2.

Schnerch, Angelique; Cerdan, Chantal; Bhatia, Mickie (2010): Distinguishing between mouse and human pluripotent stem cell regulation: the best laid plans of mice and men. In *STEM CELLS* 28 (3), pp. 419–430. DOI: 10.1002/stem.298.

Selevan, S. G.; Kimmel, C. A.; Mendola, P. (2000): Identifying critical windows of exposure for children's health. In *Environmental health perspectives* 108 Suppl 3 (Suppl 3), pp. 451–455. DOI: 10.1289/ehp.00108s3451.

Severus, W. E.; Kleindienst, N.; Seemüller, F.; Frangou, S.; Möller, H. J.; Greil, W. (2008): What is the optimal serum lithium level in the long-term treatment of bipolar disorder--a review? In *Bipolar disorders* 10 (2), pp. 231–237. DOI: 10.1111/j.1399-5618.2007.00475.x.

Shankar, Prarthana; Geier, Mitra C.; Truong, Lisa; McClure, Ryan S.; Pande, Paritosh; Waters, Katrina M.; Tanguay, Robert L. (2019): Coupling Genome-wide Transcriptomics and Developmental Toxicity Profiles in Zebrafish to Characterize Polycyclic Aromatic Hydrocarbon (PAH) Hazard. In *International journal of molecular sciences* 20 (10). DOI: 10.3390/ijms20102570.

Shinde, Vaibhav; Hoelting, Lisa; Srinivasan, Sureshkumar Perumal; Meisig, Johannes; Meganathan, Kesavan; Jagtap, Smita et al. (2017): Definition of transcriptome-based indices for quantitative characterization of chemically disturbed stem cell development: introduction of the STOP-Toxukn and STOP-Toxukk tests. In *Arch Toxicol* 91 (2), pp. 839–864. DOI: 10.1007/s00204-016-1741-8.

Sillaber, Inge; Panhuysen, Markus; Henniger, Markus S. H.; Ohl, Frauke; Kühne, Claudia; Pütz, Benno et al. (2008): Profiling of behavioral changes and hippocampal gene expression in mice chronically treated with the SSRI paroxetine. In *Psychopharmacology* 200 (4), pp. 557–572. DOI: 10.1007/s00213-008-1232-6.

Simon, H. J.; Rantz, L. A. (1962): The newer penicillins. II. Clinical experiences with methicillin and oxacillin. In *Annals of internal medicine* 57, pp. 344–362. DOI: 10.7326/0003-4819-57-3-344.

Sklan, D.; Shalit, I.; Lasebnik, N.; Spirer, Z.; Weisman, Y. (1985): Retinol transport proteins and concentrations in human amniotic fluid, placenta, and fetal and maternal sera. In *The British journal of nutrition* 54 (3), pp. 577–583. DOI: 10.1079/bjn19850144.

Slikker, W., Jr.; Bailey, J.R.; Holder, C. L.; Lipe, G. L. (1987): Transplacental disposition of doxylamine succinate in the late-term rhesus monkey. In Heinz Nau (Ed.): Pharmacokinetics in teratogenesis, vol. 1. Boca Raton, Fl.: CRC Pr, pp. 193–202.

Smyth, G. K. (2005): limma: Linear Models for Microarray Data. In Robert Gentleman, Vincent J. Carey, Sandrine Dudoit, Wolfgang Huber, Rafael A. Irizarry (Eds.): Bioinformatics and Computational Biology Solutions Using R and Bioconductor. New York, NY: Springer Science+Business Media Inc (Statistics for Biology and Health), pp. 397–420.

References

Sørensen, T. Sandberg; Petersen, O.; Heerfordt, J. (1982): The Pharmacokinetics of Methicillin and Dicloxacillin in Wound Fluid Following Internal Fixation of Trochanteric Fractures. In *Acta Orthopaedica Scandinavica* 53 (4), pp. 535–539. DOI: 10.3109/17453678208992253.

Speirs, A.L (1962): THALIDOMIDE AND CONGENITAL ABNORMALITIES. In *The Lancet* 279 (7224), pp. 303–305. DOI: 10.1016/S0140-6736(62)91248-5.

Stemina Biomarker Discovery (2020): In vitro Toxicity Testing of Drugs, Food, and Products | stemina.com. Available online at https://stemina.com/, updated on 10/23/2020, checked on 10/23/2020.

Stummann, T. C.; Hareng, L.; Bremer, S. (2007): Embryotoxicity hazard assessment of methylmercury and chromium using embryonic stem cells. In *Toxicology* 242 (1-3), pp. 130–143. DOI: 10.1016/j.tox.2007.09.022.

Stummann, T. C.; Hareng, L.; Bremer, S. (2008): Embryotoxicity hazard assessment of cadmium and arsenic compounds using embryonic stem cells. In *Toxicology* 252 (1-3), pp. 118–122. DOI: 10.1016/j.tox.2008.08.001.

Sucov, H. M.; Izpisúa-Belmonte, J. C.; Gañan, Y.; Evans, R. M. (1995): Mouse embryos lacking RXR alpha are resistant to retinoic-acid-induced limb defects. In *Development (Cambridge, England)* 121 (12), pp. 3997–4003.

Suzuki, Noriyuki; Ando, Satoshi; Yamashita, Norihisa; Horie, Nobuyuki; Saito, Koichi (2011): Evaluation of novel high-throughput embryonic stem cell tests with new molecular markers for screening embryotoxic chemicals in vitro. In *Toxicological Sciences* 124 (2), pp. 460–471. DOI: 10.1093/toxsci/kfr250.

Sylvetsky, Allison C.; Bauman, Viviana; Blau, Jenny E.; Garraffo, H. Martin; Walter, Peter J.; Rother, Kristina I. (2017): Plasma concentrations of sucralose in children and adults. In *Toxicological and environmental chemistry* 99 (3), pp. 535–542. DOI: 10.1080/02772248.2016.1234754.

Tagawa, Masaaki; Kano, Michiko; Okamura, Nobuyuki; Higuchi, Makoto; Matsuda, Michiaki; Mizuki, Yasuyuki et al. (2002): Differential cognitive effects of ebastine and (+)-chlorpheniramine in healthy subjects: correlation between cognitive impairment and plasma drug concentration. In *British Journal of Clinical Pharmacology* 53 (3), pp. 296–304. DOI: 10.1046/j.0306-5251.2001.01183.x.

Takahashi, Kazutoshi; Tanabe, Koji; Ohnuki, Mari; Narita, Megumi; Ichisaka, Tomoko; Tomoda, Kiichiro; Yamanaka, Shinya (2007): Induction of pluripotent stem cells from adult human fibroblasts by defined factors. In *Cell* 131 (5), pp. 861–872. DOI: 10.1016/j.cell.2007.11.019.

Takahashi, Kazutoshi; Yamanaka, Shinya (2006): Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. In *Cell* 126 (4), pp. 663–676. DOI: 10.1016/j.cell.2006.07.024.

Tanguay, Mario; Girard, Johanne; Scarsi, Claudia; Mautone, Giuseppe; Larouche, Richard (2019): Pharmacokinetics and Comparative Bioavailability of a Levothyroxine Sodium Oral Solution and Soft Capsule. In *Clinical pharmacology in drug development* 8 (4), pp. 521–528. DOI: 10.1002/cpdd.608.

Tayri-Wilk, Tamar; Slavin, Moriya; Zamel, Joanna; Blass, Ayelet; Cohen, Shon; Motzik, Alex et al. (2020): Mass spectrometry reveals the chemistry of formaldehyde cross-linking in structured proteins. In *Nature Communications* 11 (1), p. 3128. DOI: 10.1038/s41467-020-16935-w.

Teixidó, Elisabet; Kießling, Tobias R.; Krupp, Eckart; Quevedo, Celia; Muriana, Arantza; Scholz, Stefan (2019): Automated Morphological Feature Assessment for Zebrafish Embryo Developmental Toxicity Screens. In *Toxicological Sciences* 167 (2), pp. 438–449. DOI: 10.1093/toxsci/kfy250.

Teo, Steve K.; Colburn, Wayne A.; Tracewell, William G.; Kook, Karin A.; Stirling, David I.; Jaworsky, Markian S. et al. (2004): Clinical pharmacokinetics of thalidomide. In *Clinical Pharmacokinetics* 43 (5), pp. 311–327. DOI: 10.2165/00003088-200443050-00004.

Terryn, Joke; Welkenhuysen, Marleen; Krylychkina, Olga; Firrincieli, Andrea; Andrei, Alexandru; Reumers, Veerle et al. (2018): Topographical Guidance of PSC-Derived Cortical Neurons. Journal of Nanomaterials, 2018, 1-10. In *Journal of Nanomaterials* 2018, pp. 1–10. DOI: 10.1155/2018/5238901.

TGA (n.d.): Australian categorisation system for prescribing medicines in pregnancy. Australian Therapeutic Goods Administration. Available online at https://www.tga.gov.au/node/4013, checked on 11/12/2020.

Thomson, J. A.; Itskovitz-Eldor, J.; Shapiro, S. S.; Waknitz, M. A.; Swiergiel, J. J.; Marshall, V. S.; Jones, J. M. (1998): Embryonic stem cell lines derived from human blastocysts. In *Science* 282 (5391), pp. 1145–1147. DOI: 10.1126/science.282.5391.1145.

Torma, H.; Vahlquist, A. (1986): Uptake of vitamin A and retinol-binding protein by human placenta in vitro. In *Placenta* 7 (4), pp. 295–305. DOI: 10.1016/S0143-4004(86)80147-3.

Torres-Sánchez, Luisa; Schnaas, Lourdes; Rothenberg, Stephen J.; Cebrián, Mariano E.; Osorio-Valencia, Erika; Hernández, María del Carmen et al. (2013): Prenatal p,p´-DDE exposure and neurodevelopment among children 3.5-5 years of age. In *Environmental health perspectives* 121 (2), pp. 263–268. DOI: 10.1289/ehp.1205034.

Tuchmann-Duplessis, H.; Hiss, D.; Mottot, G.; Rosner, I. (1973): Embryotoxic and teratogenic effect of actinomycin D in the syrian hamster. In *Toxicology* 1 (2), pp. 131–133. DOI: 10.1016/0300-483X(73)90025-5.

Twelves, C. J.; Dobbs, N. A.; Aldhous, M.; Harper, P. G.; Rubens, R. D.; Richards, M. A. (1991): Comparative pharmacokinetics of doxorubicin given by three different schedules with equal dose intensity in patients with breast cancer. In *Cancer chemotherapy and pharmacology* 28 (4), pp. 302–307. DOI: 10.1007/BF00685539.

Uibel, Frederik; Mühleisen, Annette; Köhle, Christoph; Weimer, Marc; Stummann, Tina C.; Bremer, Susanne; Schwarz, Michael (2010): ReProGlo: a new stem cell-based reporter assay aimed to predict embryotoxic potential of drugs and chemicals. In *Reproductive toxicology (Elmsford, N.Y.)* 30 (1), pp. 103–112. DOI: 10.1016/j.reprotox.2009.12.002.

Ujházy, Eduard; Mach, Mojmír; Navarová, Jana; Brucknerová, Ingrid; Dubovický, Michal (2012): Teratology - past, present and future. In *Interdisciplinary toxicology* 5 (4), pp. 163–168. DOI: 10.2478/v10102-012-0027-0.

United States Code (2014): Animal Welfare Act 2014, Title 7, chapter 54, sections 2131–2159.

Unoki, Takamitsu; Akiyama, Masahiro; Kumagai, Yoshito; Gonçalves, Filipe Marques; Farina, Marcelo; Da Rocha, João Batista Teixeira; Aschner, Michael (2018): Molecular Pathways Associated With Methylmercury-Induced Nrf2 Modulation. In *Frontiers in genetics* 9, p. 373. DOI: 10.3389/fgene.2018.00373.

van Calsteren, Kristel; Hartmann, Dieter; van Aerschot, Leen; Verbesselt, Rene; van Bree, Rieta; D'Hooge, Rudi; Amant, Frédéric (2009): Vinblastine and doxorubicin administration to pregnant mice affects brain development and behaviour in the offspring. In *NeuroToxicology* 30 (4), pp. 647–657. DOI: 10.1016/j.neuro.2009.04.009.

van Gelder, R. N.; Zastrow, M. E. von; Yool, A.; Dement, W. C.; Barchas, J. D.; Eberwine, J. H. (1990): Amplified RNA synthesized from limited quantities of heterogeneous cDNA. In *Proceedings of the National Academy of Sciences of the United States of America* 87 (5), pp. 1663–1667. DOI: 10.1073/pnas.87.5.1663.

van Tellingen, O.; Beijnen, J. H.; Nooijen, W. J.; Bult, A. (1993): Plasma pharmacokinetics of vinblastine and the investigational Vinca alkaloid N-(deacetyl-O-4-vinblastoyl-23)-L-ethyl isoleucinate in mice as determined by high-performance liquid chromatography. In *Cancer research* 53 (9), pp. 2061–2065.

van Veggel, Mathilde; Westerman, Elsbeth; Hamberg, Paul (2018): Clinical Pharmacokinetics and Pharmacodynamics of Panobinostat. In *Clin Pharmacokinet* 57 (1), pp. 21–29. DOI: 10.1007/s40262-017-0565-x.

Varela-López, Alfonso; Battino, Maurizio; Navarro-Hortal, María D.; Giampieri, Francesca; Forbes-Hernández, Tamara Y.; Romero-Márquez, José M. et al. (2019): An update on the mechanisms related to cell death and toxicity of doxorubicin and the protective role of nutrients. In *Food and chemical toxicology : an international journal published for the British Industrial Biological Research Association* 134, p. 110834. DOI: 10.1016/j.fct.2019.110834.

Vargesson, Neil (2015): Thalidomide-induced teratogenesis: history and mechanisms. In *Birth defects research. Part C, Embryo today : reviews* 105 (2), pp. 140–156. DOI: 10.1002/bdrc.21096.

Varshney, Gaurav K.; Sood, Raman; Burgess, Shawn M. (2015): Understanding and Editing the Zebrafish Genome: Elsevier (Advances in Genetics, 92). In *Advances in genetics*, pp. 1–52.

Veal, Gareth J.; Cole, Michael; Errington, Julie; Parry, Annie; Hale, Juliet; Pearson, Andrew D. J. et al. (2005): Pharmacokinetics of dactinomycin in a pediatric patient population: a United Kingdom Children's Cancer Study Group Study. In *Clinical cancer research : an official journal of the American Association for Cancer Research* 11 (16), pp. 5893–5899. DOI: 10.1158/1078-0432.CCR-04-2546.

Verberne, Eline A.; Haan, Emma de; van Tintelen, J. Peter; Lindhout, Dick; van Haelst, Mieke M. (2019): Fetal methotrexate syndrome: A systematic review of case reports. In *Reproductive toxicology (Elmsford, N.Y.)* 87, pp. 125–139. DOI: 10.1016/j.reprotox.2019.05.066.

Videla, Sebastián; Cebrecos, Jesús; Lahjou, Mounia; Wagner, France; Guibord, Pascal; Xu, Zhengguo et al. (2013): Pharmacokinetic dose proportionality between two strengths (12.5 mg and 25 mg) of doxylamine hydrogen succinate film-coated tablets in fasting state: a single-dose, randomized, two-period crossover study in healthy volunteers. In *Drugs in R&D* 13 (2), pp. 129–135. DOI: 10.1007/s40268-013-0015-7.

Vulsma, T.; Gons, M. H.; Vijlder, J. J. de (1989): Maternal-fetal transfer of thyroxine in congenital hypothyroidism due to a total organification defect or thyroid agenesis. In *The New England journal of medicine* 321 (1), pp. 13–16. DOI: 10.1056/NEJM198907063210103.

Waldmann, Tanja; Grinberg, Marianna; König, André; Rempel, Eugen; Schildknecht, Stefan; Henry, Margit et al. (2017): Stem Cell Transcriptome Responses and Corresponding Biomarkers That Indicate the Transition from Adaptive Responses to Cytotoxicity. In *Chemical research in toxicology* 30 (4), pp. 905–922. DOI: 10.1021/acs.chemrestox.6b00259.

Waldmann, Tanja; Rempel, Eugen; Balmer, Nina V.; König, André; Kolde, Raivo; Gaspar, John Antonydas et al. (2014): Design principles of concentration-dependent transcriptome deviations in drug-exposed differentiating stem cells. In *Chemical research in toxicology* 27 (3), pp. 408–420. DOI: 10.1021/tx400402j.

References

Wesseloo, Richard; Wierdsma, André I.; van Kamp, Inge L.; Munk-Olsen, Trine; Hoogendijk, Witte J. G.; Kushner, Steven A.; Bergink, Veerle (2017): Lithium dosing strategies during pregnancy and the postpartum period. In *The British journal of psychiatry : the journal of mental science* 211 (1), pp. 31–36. DOI: 10.1192/bjp.bp.116.192799.

West, Paul R.; Weir, April M.; Smith, Alan M.; Donley, Elizabeth L. R.; Cezar, Gabriela G. (2010): Predicting human developmental toxicity of pharmaceuticals using human embryonic stem cells and metabolomics. In *Toxicology and Applied Pharmacology* 247 (1), pp. 18–27. DOI: 10.1016/j.taap.2010.05.007.

Wiegand, Ulf-W.; Chou, Ruby C. (1998): Pharmacokinetics of acitretin and etretinate. In *Journal of the American Academy of Dermatology* 39 (2), S25-S33. DOI: 10.1016/s0190-9622(98)70441-4.

Wilimzig, C.; Latz, R.; Vierling, W.; Mutschler, E.; Trnovec, T.; Nyulassy, S. (1996): Increase in magnesium plasma level after orally administered trimagnesium dicitrate. In *European journal of clinical pharmacology* 49 (4), pp. 317–323. DOI: 10.1007/BF00226334.

William Slikker, JR; Paule, Merle G.; Wang, Cheng (2018): Handbook of Developmental Neurotoxicology. 2nd ed. San Diego: Elsevier Science. Available online at https://ebookcentral.proquest.com/lib/gbv/detail.action?docID=5214862.

Williams, Amy Lavin; Pace, Nelson D.; DeSesso, John M. (2020): Teratogen update: Topical use and third-generation retinoids. In *Birth defects research* 112 (15), pp. 1105–1114. DOI: 10.1002/bdr2.1745.

Wu, Jun; Izpisua Belmonte, Juan Carlos (2015): Dynamic Pluripotent Stem Cell States and Their Applications. In *Cell stem cell* 17 (5), pp. 509–525. DOI: 10.1016/j.stem.2015.10.009.

Xu, W. S.; Parmigiani, R. B.; Marks, P. A. (2007): Histone deacetylase inhibitors: molecular mechanisms of action. In *Oncogene* 26 (37), pp. 5541–5552. DOI: 10.1038/sj.onc.1210620.

Yoo, S. D.; Axelson, J. E.; Taylor, S. M.; Rurak, D. W. (1986): Placental transfer of diphenhydramine in chronically instrumented pregnant sheep. In *Journal of pharmaceutical sciences* 75 (7), pp. 685–687. DOI: 10.1002/jps.2600750714.

Yue, C. S.; Scarsi, C.; Ducharme, M. P. (2012): Pharmacokinetics and potential advantages of a new oral solution of levothyroxine vs. other available dosage forms. In *Arzneimittel-Forschung* 62 (12), pp. 631–636. DOI: 10.1055/s-0032-1329951.

Zhong, Xiali; Harris, Georgina; Smirnova, Lena; Zufferey, Valentin; Da Sá, Rita de Cássia Silveira E.; Baldino Russo, Fabiele et al. (2020): Antidepressant Paroxetine Exerts Developmental Neurotoxicity in an iPSC-Derived 3D Human Brain Model. In *Frontiers in cellular neuroscience* 14, p. 25. DOI: 10.3389/fncel.2020.00025.

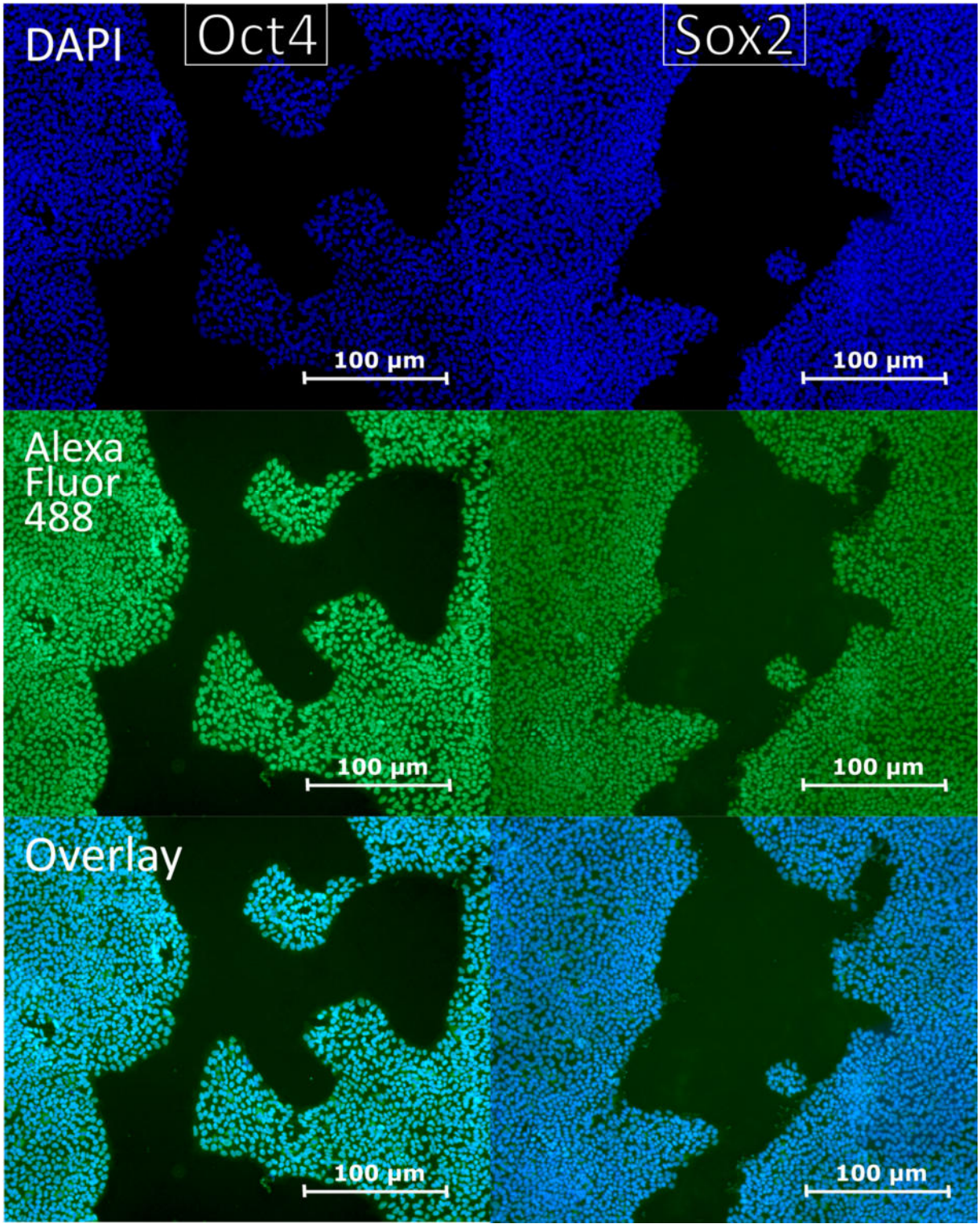
Zhou, Ren; Cheng, Wei; Feng, Yan; Wei, Hongying; Liang, Fan; Wang, Yan (2017): Interactions between three typical endocrine-disrupting chemicals (EDCs) in binary mixtures exposure on myocardial differentiation of mouse embryonic stem cell. In *Chemosphere* 178, pp. 378–383. DOI: 10.1016/j.chemosphere.2017.03.040.

Zhuang, Xiaomei; Lu, Chuang (2016): PBPK modeling and simulation in drug research and development. In *Acta Pharmaceutica Sinica B* 6 (5), pp. 430–440. DOI: 10.1016/j.apsb.2016.04.004.

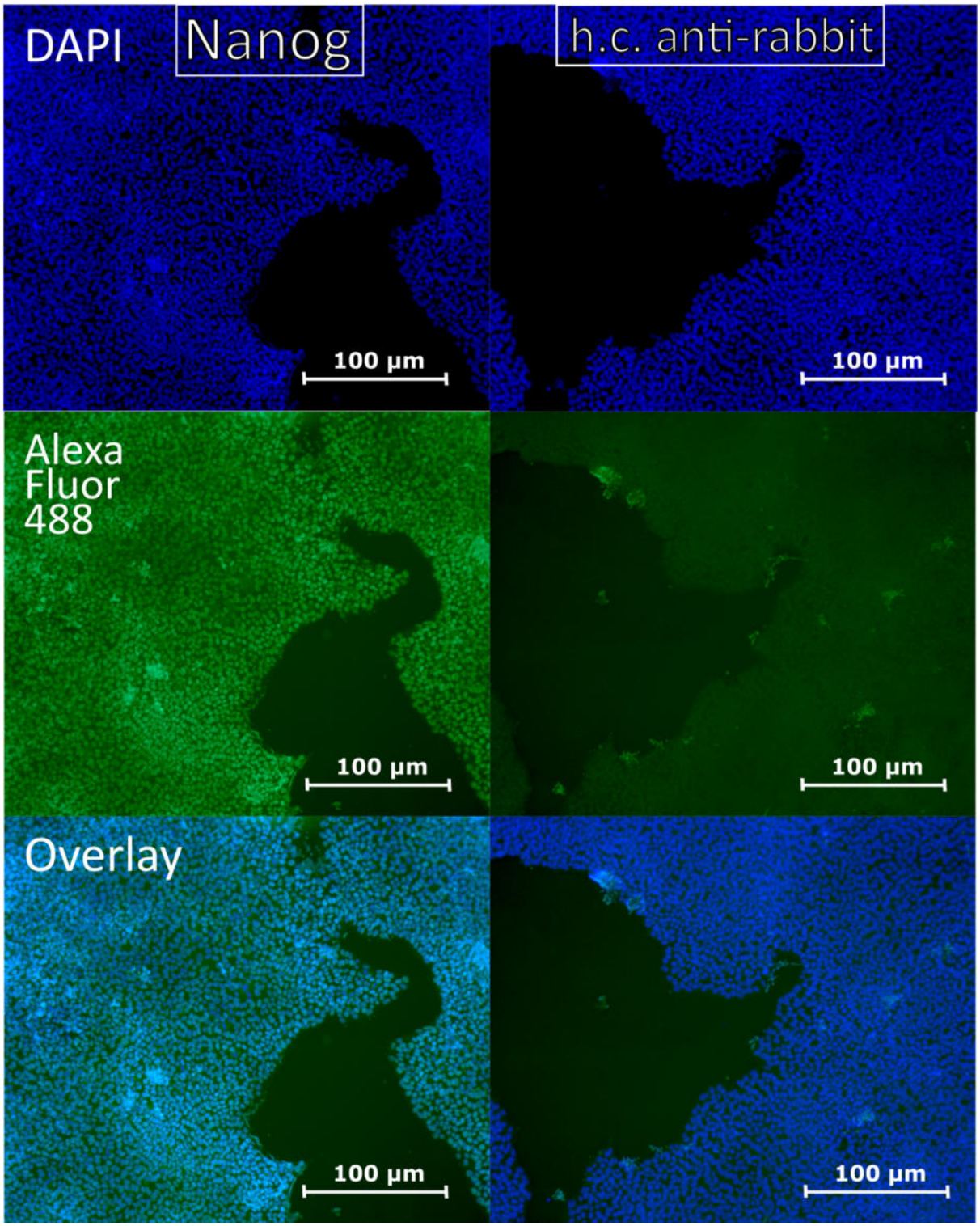
Zurlinden, Todd J.; Saili, Katerine S.; Rush, Nathaniel; Kothiya, Parth; Judson, Richard S.; Houck, Keith A. et al. (2020): Profiling the ToxCast Library With a Pluripotent Human (H9) Stem Cell Line-Based Biomarker Assay for Developmental Toxicity. In *Toxicological Sciences* 174 (2), pp. 189–209. DOI: 10.1093/toxsci/kfaa014.

6. Appendix

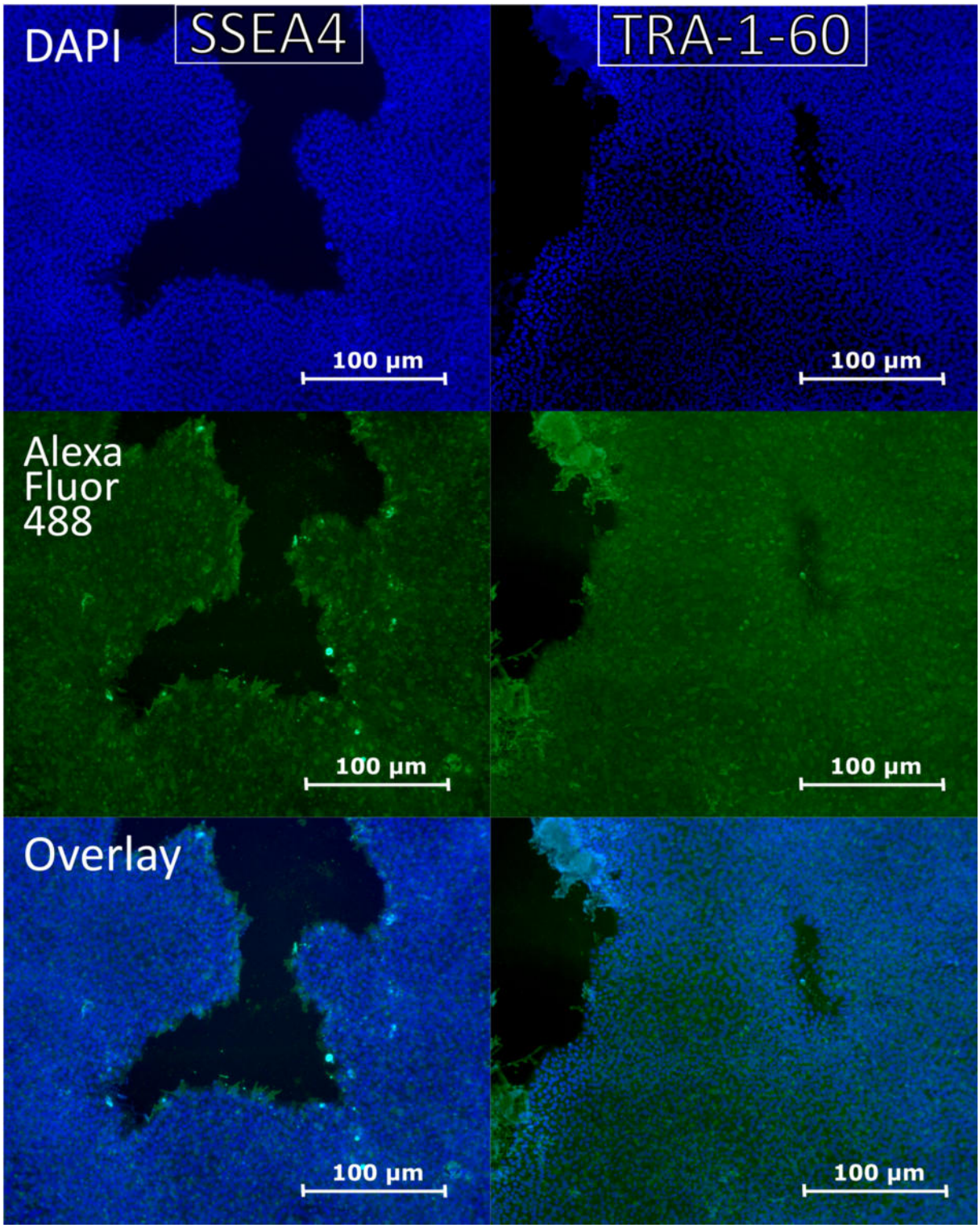
6.1 Supplements



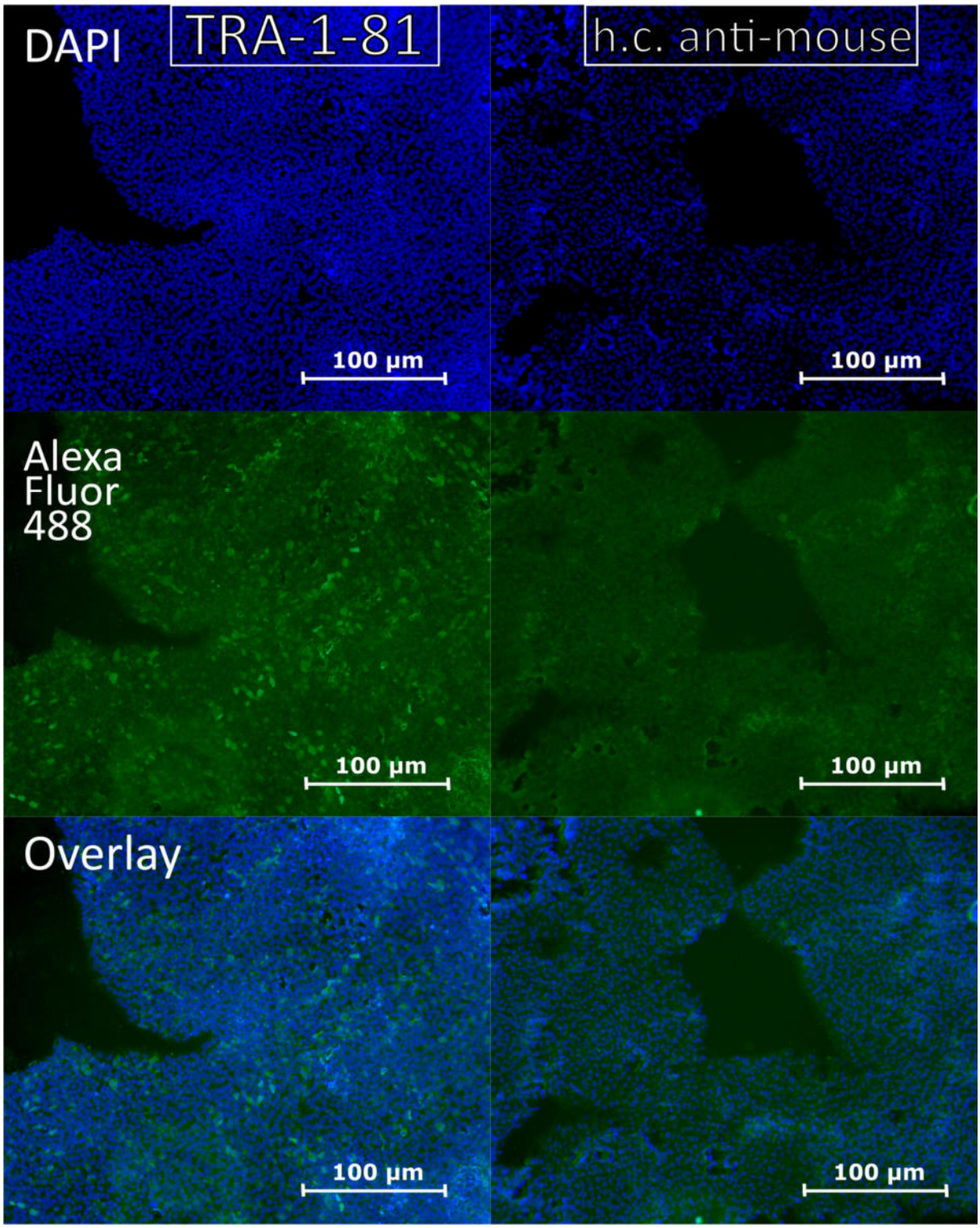
***Supplementary figure 1**: Immunohistochemical staining of the nuclear stem cell markers OCT4A and SOX2 expressed in SBAD2 cells (passage 11) cultured in Essential 8. DAPI stained the cell nuclei and Alexa Fluor 488 the factors OCT4A and SOX2. The overlay proved the simultaneous presence of both dyes in cells.*



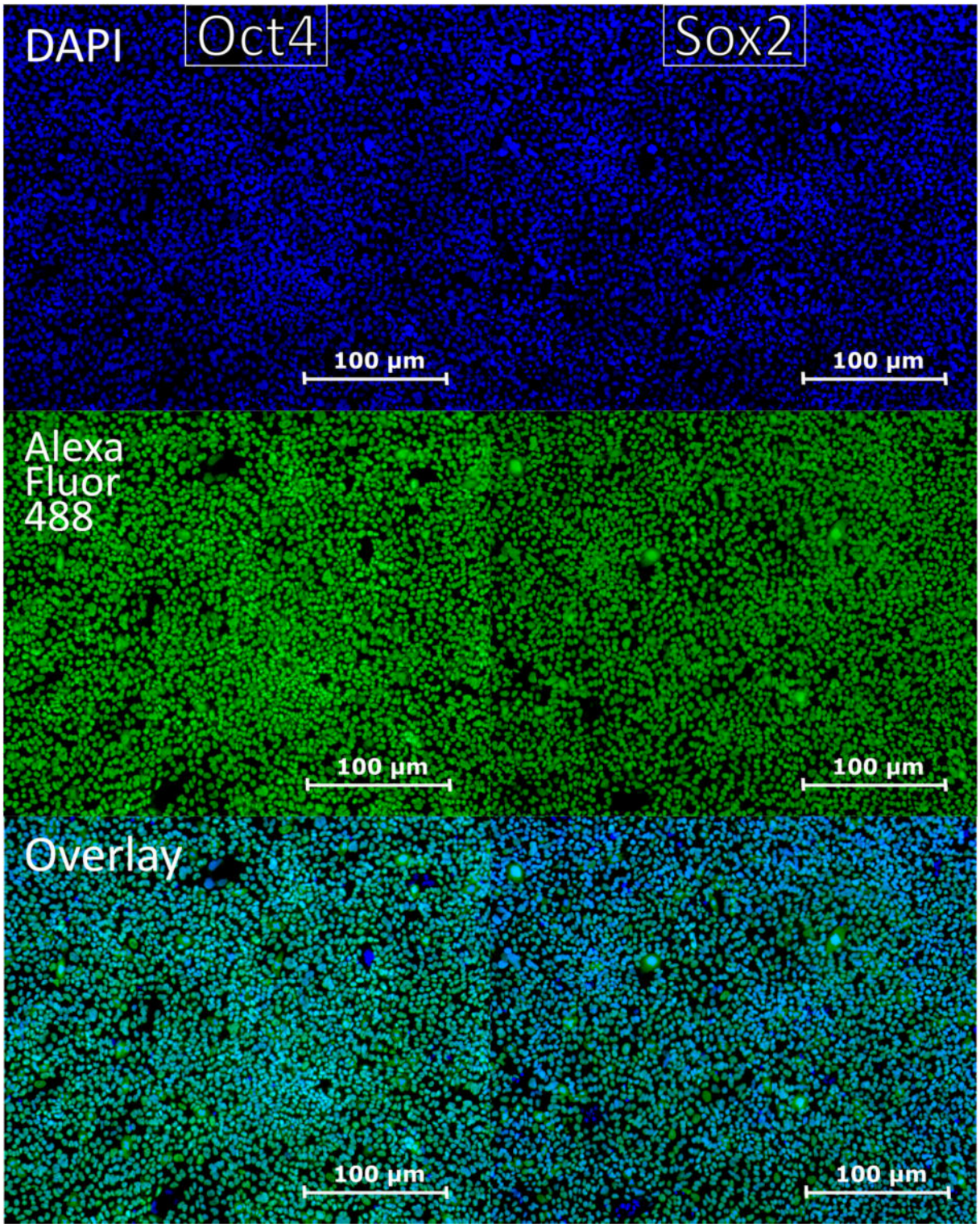
***Supplementary figure 2**: Immunohistochemical staining of the nuclear stem cell marker Nanog expressed in SBAD2 cells (passage 11) cultured in Essential 8 and the histological control of the secondary anti-rabbit antibody. DAPI stained the cell nuclei and Alexa Fluor 488 the factor Nanog. The overlay proved the simultaneous presence of both dyes in cells.*



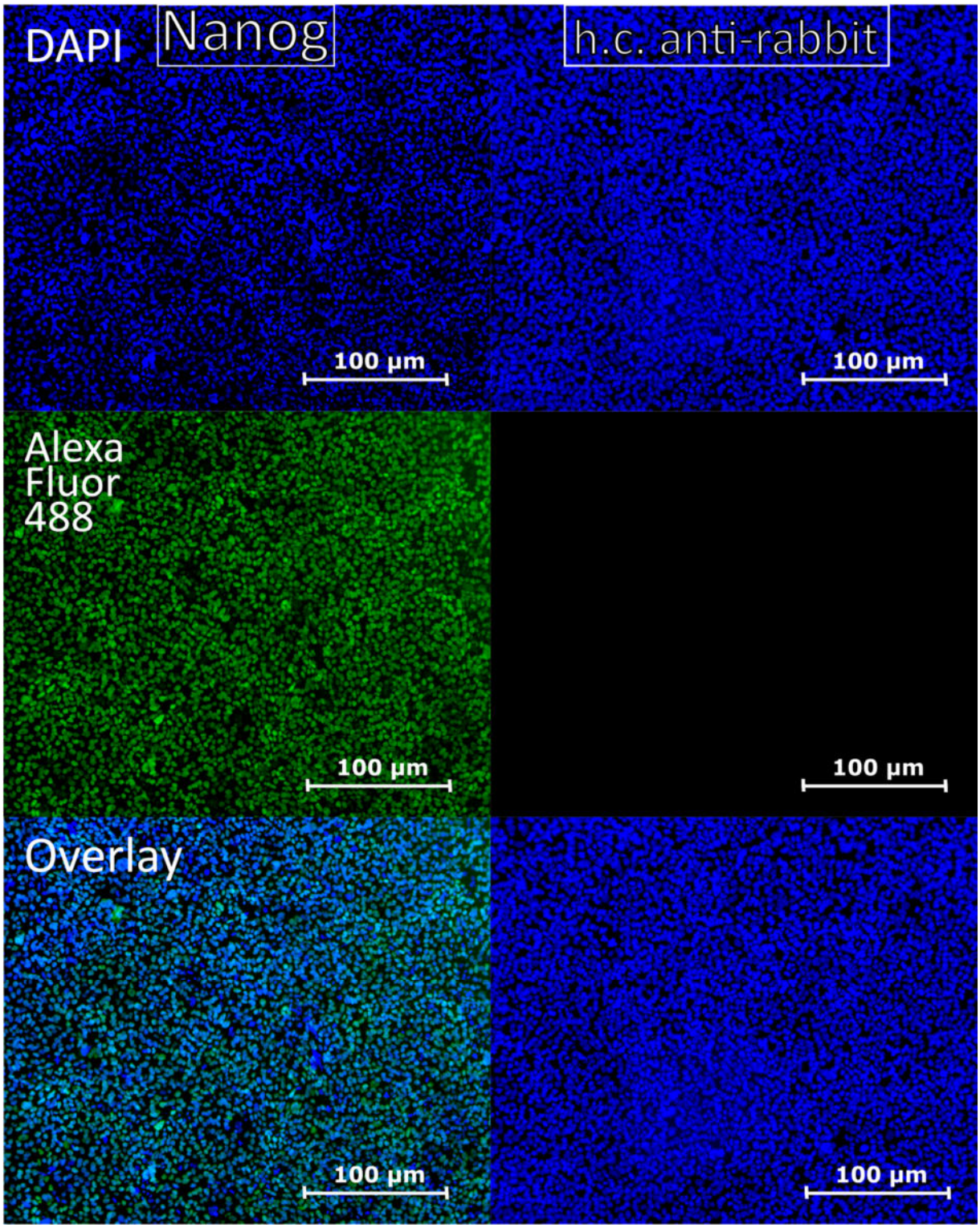
***Supplementary figure 3**: Immunohistochemical staining of the stem cell markers SSEA4 and TRA-1-60 expressed on the surface of SBAD2 cells (passage 11) cultured in Essential 8. DAPI stained the cell nuclei and Alexa Fluor 488 the markers SSEA4 and TRA-1-60. The overlay proved the simultaneous presence of both dyes in and on cells.*



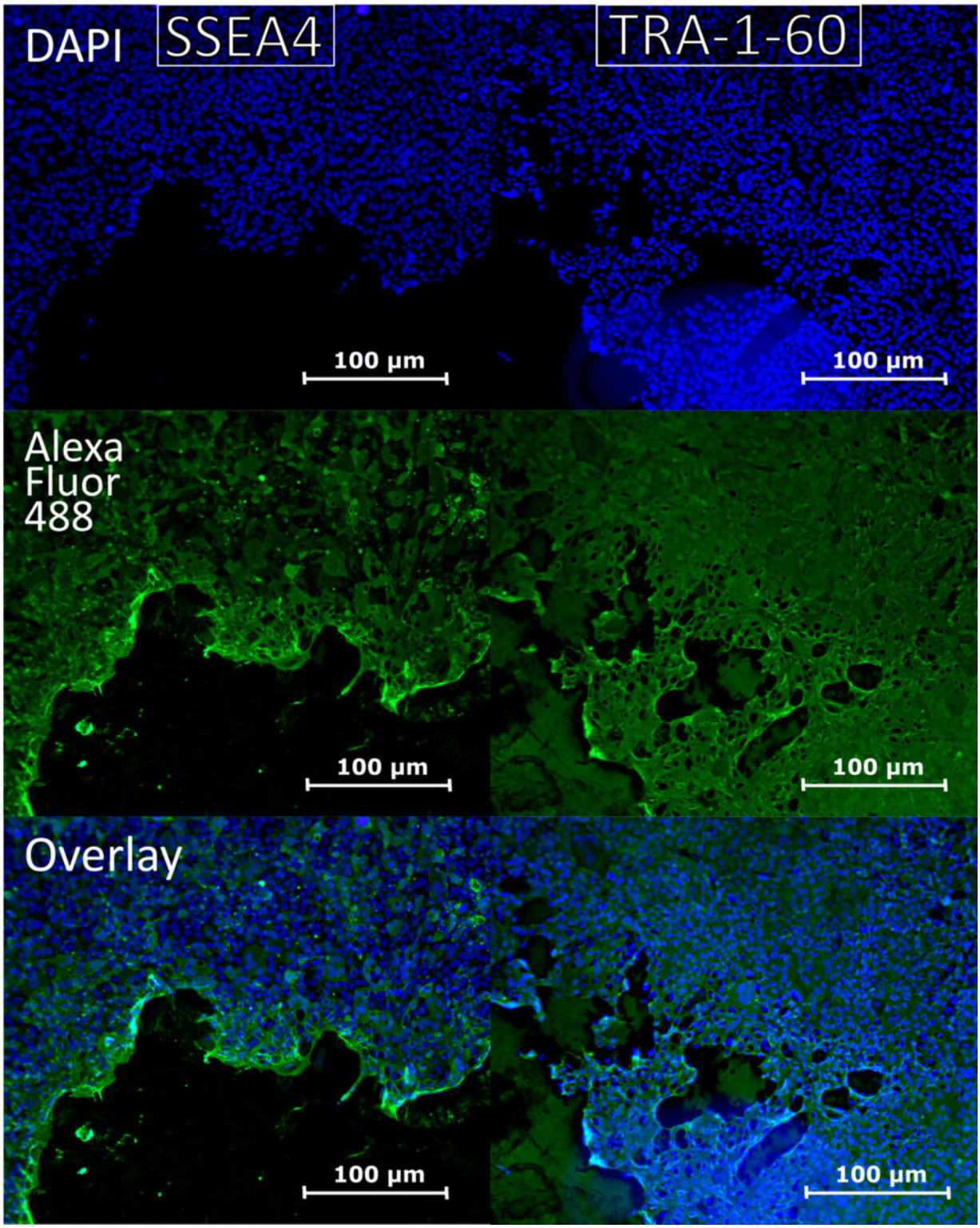
Supplementary figure 4: *Immunohistochemical staining of the stem cell marker TRA-1-81 expressed on the surface of SBAD2 cells (passage 11) cultured in Essential 8 and the histological control of the secondary anti-mouse antibody. DAPI stained the cell nuclei and Alexa Fluor 488 the marker TRA-1-81. The overlay proved the simultaneous presence of both dyes in and on cells.*



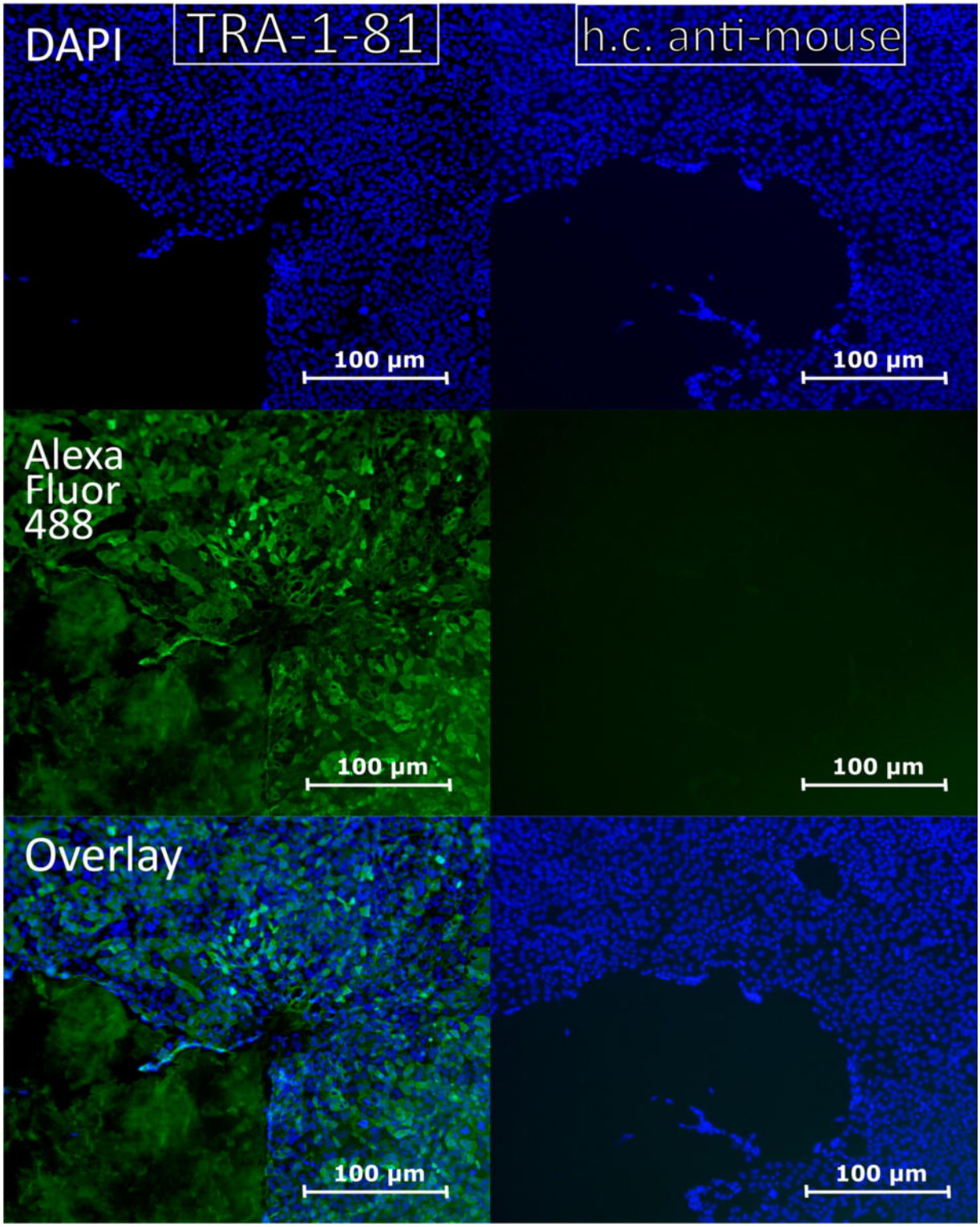
Supplementary figure 5: *Immunohistochemical staining of the nuclear stem cell markers OCT4A and SOX2 expressed in SBAD2 cells (passage 10) cultured in DEF-CS. DAPI stained the cell nuclei and Alexa Fluor 488 the factors OCT4A and SOX2. The overlay proved the simultaneous presence of both dyes in cells.*



Supplementary figure 6: *Immunohistochemical staining of the nuclear stem cell marker Nanog expressed in SBAD2 cells (passage 10) cultured in DEF-CS and the histological control of the secondary anti-rabbit antibody. DAPI stained the cell nuclei and Alexa Fluor 488 the factor Nanog. The overlay proved the simultaneous presence of both dyes in cells.*



Supplementary figure 7: *Immunohistochemical staining of the stem cell markers SSEA4 and TRA-1-60 expressed on the surface of SBAD2 cells (passage 10) cultured in DEF-CS. DAPI stained the cell nuclei and Alexa Fluor 488 the markers SSEA4 and TRA-1-60. The overlay proved the simultaneous presence of both dyes in and on cells.*



Supplementary figure 8: *Immunohistochemical staining of the stem cell marker TRA-1-81 expressed on the surface of SBAD2 cells (passage 10) cultured in DEF-CS and the histological control of the secondary anti-mouse antibody. DAPI stained the cell nuclei and Alexa Fluor 488 the marker TRA-1-81. The overlay proved the simultaneous presence of both dyes in and on cells.*

6.2 Raw data

The raw data of the Affymetrix gene array measurements, the cell titer blue measurements and the RT-qPCR measurements can be found in digital files which were handed in along with this work.

6.3 List of figures

Appendix

Appendix

6.4 List of tables

Appendix

6.5 Acknowledgements

First and foremost, I want to thank Prof. Jan Hengstler for the great opportunity to work at the IfADo in this very interesting project and for all the trust he placed in me through these challenging last years. His supervision and outstanding expertise were an important help in many difficult situations and decisions and guided my way in becoming a better scientist.

I give my thanks to all the collaboration partners that were involved in this project, namely Prof. Agapios Sachinidis and Anna Cherianidou from the the department of neurophysiology at the University of Cologne, Prof. Marcel Leist, Dr. Nadine Dreser, Jonathan Blum and Dr. Tanja Waldmann (former member) from the department of in vitro toxicology and biomedicine at the University of Konstanz, Prof. Jörg Rahnenführer, Franziska Kappenberg and Dr. Katrin Madjar (former member) from the faculty of statistics at the TU Dortmund and Prof. Nils Blüthgen and Dr. Johannes Meisig from the Charité at the University of Berlin. We had very fruitful and enjoyable meetings and discussions, and our united efforts could lead the project to a successful end.

A big gratitude goes to all members of the department of toxicology at the IfADo. You all were great and friendly colleagues with a helping hand whenever one was necessary, and spending time with you was always a joy. Special thanks go to Tim Brecklinghaus for being such a good-humored, communicative and helpful office colleague every day, Patrick Nell for his knowledged support with the stem cells, especially at the beginning of my work, and Miriam Marta, who contributed to this work during her master thesis. Thank you all!

Mein größter Dank gebührt zum einen meinen Eltern, die immer für mich da waren, sich mit offenen Ohren alle meine Sorgen angehört und mit gutem Ratschlag, viel Herz und Zuversicht immer meinen Rücken gestärkt haben. Zum anderen meiner Freundin, mit der ich seit vielen Jahren meinen Lebensweg teile, die mich aber immer noch nicht satthat und trotz der fordernden Arbeit, die besonders gegen Ende sehr stressig war, ertragen und mich stets unterstützt hat. Ebenso danke ich auch dem Rest meiner Familie für all die Unterstützung, die sie mir entgegengebracht hat. Last but not least, bedanke ich mich auch bei allen meinen Freunden, die einen immer wieder ins Gedächtnis gerufen haben, dass es auch noch ein Leben außerhalb des Labors gab. Ich danke euch allen von Herzen!