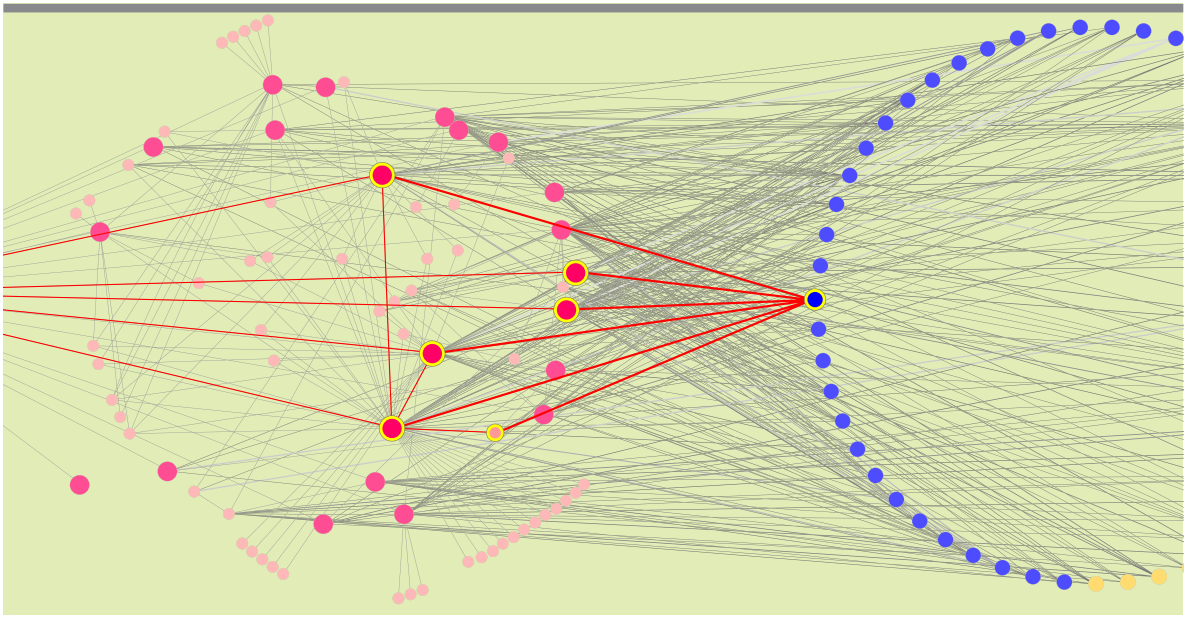
# Integrative Systems Biology Applied to Toxicology



**Kristine Grønning Kongsbak**
**PhD Thesis**
**January 2015**

DTU Food
National Food Institute

# Integrative Systems Biology Applied to Toxicology

Kristine Grønning Kongsbak

Søborg 2015
FOOD-PHD-2015

PhD Thesis

2015

**Supervisors**
Professor Anne Marie Vinggaard
Senior Scientist Niels Hadrup
Division of Toxicology and Risk Assessment
National Food Institute
Technical University of Denmark

Associate Professor Aron Charles Eklund
Center for Biological Sequence Analysis
Department for Systems Biology
Technical University of Denmark

Associate Professor Karine Audouze
Molécules Thérapeutiques In Silico
Paris Diderot University

**Funding**
This project was supported financially by the Ministry of Food, Agriculture and Fisheries of Denmark and the Technical University of Denmark.


FOOD-PHD: ISBN 978-87-93109-30-8

Division of Toxicology and Risk Assessment
National Food Institute
Technical University of Denmark
DK-2860 Søborg, Denmark
www.food.dtu.dk

# Summary

Humans are exposed to various chemical agents through food, cosmetics, pharmaceuticals and other sources. Exposure to chemicals is suspected of playing a main role in the development of some adverse health effects in humans. Additionally, European regulatory authorities have recognized the risk associated with combined exposure to multiple chemicals. Testing all possible combinations of the tens of thousands environmental chemicals is impractical. This PhD project was launched to apply existing computational systems biology methods to toxicological research.

In this thesis, I present in three projects three different approaches to using computational toxicology to aid classical toxicological investigations. In project I, we predicted human health effects of five pesticides using publicly available data. We obtained a grouping of the chemical according to their potential human health effects that were in concordance with their effects in experimental animals. In project II, I profiled the effects on rat liver gene expression levels following exposure to a 14-chemical mixture $\pm$ the presence of an endocrine disrupting chemical. This project helped us shed light on the mechanism of action of the 14-chemical mixture and the endocrine disrupting chemical. In project III, I modeled a predictive signature for an *in vivo* endpoint that is sensitive to endocrine disruption. I used publicly available data generated for the purpose of modeling predictive signatures for various *in vivo* endpoints. From this modeling effort, I have suggested a mechanism of action for a subset of the chemicals that has not previously been associated with endocrine disruption.

The use of computational methods in toxicology can aid the classical toxicological tests by suggesting interactions between separate components of a system thereby suggesting new ways of thinking specific toxicological endpoints. Furthermore, computational methods can serve as valuable input for the hypothesis generating phase of the preparations of a research project.

# Resume

Mennesker bliver eksponeret for adskillige kemiske stoffer gennem fødevarer, kosmetik, lægemidler og andre kilder. Eksponering for kemikalier er mistænkt for at spille en vigtig rolle i udviklingen af nogle uønskede sundhedseffecter hos mennesker. Derudover har de europæiske tilsynsmyndigheder anerkendt risikoen forbundet med kombineret eksponering for flere kemikalier. Test af samtlige mulige kombinationer af de titusinder tilgængelige miljø- kemikalier er upraktisk. Dette Ph.d.-projekt blev iværksat for at anvende eksisterende systembiologiske computermetoder i toksikologisk forskning.

I denne afhandling præsenterer jeg i tre projekter tre forskellige tilgange til at bruge computaterbaseret toksikologi til at støtte klassiske toksikologiske undersøgelser. I projekt I, forudsagde vi sundhedsvirkningerne af fem pesticider ved brug af offentligt tilgængelige data. Vi opnåede en gruppering af kemikalierne i henhold til deres potentielle indvirkninger på menneskers sundhed, der var i overensstemmelse med deres effekter i forsøgsdyr. I projekt II, profilerede jeg effekterne af en kemisk blanding med 14 kemikalier med og uden tilstedeværelse af et hormonforstyrrende stof på genekspressionsniveauerne i rottelevere. Dette projekt hjalp os med at belyse virkningsmekanismen af den kemiske blanding og det hormonforstyrrende stof. I projektet III, modellerede jeg en prædiktiv signatur for et *in vivo* effektmål, der er følsomt over for hormonforstyrrende stoffer. Jeg brugte offentligt tilgængelige data, der er genereret med henblik på modellering af prædiktive signaturer til forskellige *in vivo* effektmål. Fra denne modelleringsindsats, har jeg foreslået en virkningsmekanisme for en gruppe af de undersøgte kemikalier, som ikke tidligere

har været forbundet med hormonforstyrrende stoffer.

Brugen af computerbaserede metoder i toksikologi kan støtte de klassiske toksikologiske undersøgelser ved at foreslå interaktioner mellem separate komponenter i et system og dermed foreslå nye måde at opfatte særlige toksikologiske effektmål. Desuden kan computerbaserede metoder give værdifuldt input til den hypotesegenererende fase af forberedelsen af et forskningsprojekt.

# Preface

This thesis was prepared at the National Food Institute and Center for Biological Sequence Analysis, Department of Systems Biology, the Technical University of Denmark in partial fulfillment of the requirements for acquiring the PhD degree in toxicology. During my work I was supervised by Anne Marie Vinggaard and Niels Hadrup from the National Food Institute, and Karine Audouze and Aron Eklund from the Center for Biological Sequence Analysis.

The thesis deals with different aspects of computational toxicology using various methods and data sources. Collectively, my work has been focused on integrative data analyses using either existing data from the literature or *in vivo* and *in vitro* data generated for my and other studies.

The thesis consists of a summary report and a collection of four research papers/manuscripts written during the period 2011–2014, and published elsewhere along with preliminary data for a modeling project.

Søborg, October 2014

Kristine Grønning Kongsbak

# Acknowledgments

All my supervisors are greatly acknowledged for their patience and supervision during my studies. With their varied skill sets they have all contributed to my development towards becoming an independent and critically thinking scientist. They have all greeted me with an open-door policy, supported me on the rainiest days, and celebrated the successes with me. I would like to express my greatest gratitude to my colleagues at the National Food Institute, especially my fellow PhD students for making a very supportive and caring working environment. Anna, in particular, has been my go-to person and my "partner in crime".

Dr. Richard Judson is thanked for accepting me as as his mentee at the National Center for Computational Toxicology at the U.S. Environmental Protection Agency. He and his colleagues made my six months in North Carolina very memorable. They all welcomed me to their professional yet warm place to enhance/improve my professional as well as personal skills.

I would like to thank my friends, especially Sabine, Maria, Vicki, and Jill for always being there. Lastly, I would like to thank my family and partner for creating an escape for my thoughts to regenerate.

# Papers and manuscripts included in the thesis

Manuscript 1 **Kongsbak K**, Hadrup N, Audouze K, Vinggaard AM. **Applicability of Computational Systems Biology in Toxicology.** BCPT. 2014;115(1):45–9.

Manuscript 2 **Kongsbak K**, Vinggaard AM, Hadrup N, Audouze K. **A computational approach to mechanistic and predictive toxicology of pesticides.** ALTEX. 2013;1–17.

Manuscript 3 Hadrup N, Pedersen M, Skov K, Hansen NL, Berthelsen LO, **Kongsbak K**, Boberg J, Dybdahl M, Hass U, Frandsen H, Vinggaard AM. **Perfluorononanoic acid in combination with 14 chemicals exerts low dose mixture effects in rats.** Arch Toxicol. doi: 10.1007/s00204-015-1452-6

Manuscript 4 Skov K, **Kongsbak K**[1], Hadrup N, Frandsen H, Smedsgaard J, Audouze K, Eklund A, Vinggaard AM. **Metabolite and gene expression profiling of rats exposed to perfluorononanoic acid combined with a low-dose mixture of 14 human relevant compounds.** *Submitted*

[1]Skov K and Kongsbak K contributed equally to the work.

# Acronyms

**AA** aggregated assay set value.

**$AC_{50}$** half maximal active concentration.

**ACToR** Aggregated Computational Toxicology Resource.

**AGD** anogenital distance.

**AR** androgen receptor.

**BA** balanced accuracy.

**CAR** constitutive androstane receptor.

**cDNA** complementary strand deoxyribonucleic acid (DNA).

**CTD** Comparative Toxicogenomics Database.

**CYP** cytochrome P450.

**DNA** deoxyribonucleic acid.

**DTU** Technical University of Denmark.

**EDC** endocrine disrupting chemical.

**ER** estrogen receptor.

**EU** European Union.

**FDR** false discovery rate.

**GEO** Gene Expression Omnibus.

**H295** human adrenocortical carcinoma cell line.

**HTS** high-throughput screening.

**iCSS** interactive Chemical Safety for Sustainability.

**KEGG** Kyoto Encyclopedia of Genes and Genomes.

**LDA** linear discriminant analysis.

**LOAEL** lowest observed adverse effect level.

**LXR** liver X receptor.

**mg/kg BW/day** mg/kg body weight/day.

**MGR** multi-generational reproductive study.

**Mix** 14-compound mixture.

**mRNA** messenger-ribonucleic acid (RNA).

**NOAEL** no observed adverse effect level.

**OMIM** Online Mendelian Inheritance in Man.

**PBR** peripheral benzodiazepine receptor.

**PFC** perfluorinated compound.

**PFHxS** perfluorohexanoic acid.

**PFNA** perfluorononanoic acid.

**PFOA** perfluorooctanoic acid.

**PFOS** perfluorooctanesulphonic acid.

**PPAR** peroxisome proliferator-activated receptors.

**PPI** protein-protein interaction.

**PXR** pregnane X receptor.

**QC** quality control.

**REACH** Registration, Evaluation, Authorisation and Restriction of Chemicals.

**RNA** ribonucleic acid.

**RT-qPCR** real-time quantitative polymerase chain reaction.

**TDS** testicular dysgenesis syndrome.

**TGCC** testicular germ cell cancer.

**ToxRefDB** Toxicity Reference Database.

**US EPA** U. S. Environmental Protection Agency.

**VDR** vitamin D receptor.

**WHO** World Health Organisation.

# Contents

# Part I

# Introduction

Chapter 1

# Prelude

## 1.1 Motivation and context

In our industrial society, humans are exposed to chemical agents through e.g. food, cosmetics, pharmaceuticals, and air inhalation [1]. Among these, food is considered the predominant source. Chemical exposure is suspected of playing a main role in the development of some adverse health effects in humans, e.g. various male and female reproductive disorders [2–5]. Recently, European regulatory authorities have recognized the risks associated with combined exposure to multiple chemicals [6]. However, using traditional methods to test for every potentially harmful combination of chemicals is impractical; therefore we need new ways of thinking.

In 2007, the U.S. National Academy of Sciences published its report, *Toxicity Testing in the 21st Century: A Vision and a Strategy* [7–9]. This report envisions a paradigm shift in toxicity testing moving from traditional animal testing to a paradigm that 1) covers chemicals, chemical mixtures, outcomes, and life stages, broadly, 2) reduces the cost and time of testing, 3) uses fewer animals and reduces the suffering in the experimental animals, and 4) develops a more robust basis for assessing health effects of environmental chemicals. Meanwhile, the Danish government reached Food Settlement I[1] with the parliament concerning the control

[1] Fødevareforlig I

of food. In 2010, Food Settlement II[2] was reached with the aim that Denmark has the best food safety within the European Union. Part of the effort towards achieving this aim includes increased monetary resources for universities for research on chemicals in food [10].

The Cocktail Project was a four-year project that ran in 2011-2014 and was funded with resources from Food Settlement II with the aim to establish a broad basis for risk assessment of chemical cocktails[3]. To achieve this aim, several research projects aiming at generating cocktail data and developing practical tools for risk assessment were initiated. This PhD project was initiated as part of the Cocktail Project and in accordance with the above-mentioned vision and strategy for toxicity testing, it applies computational systems biology approaches to toxicological investigations of single chemicals and chemical cocktails. More specifically, I worked from the hypothesis that computational systems biology can aid classical toxicological investigations elucidating mechanisms of action and potential diseases caused by environmental substances.

## 1.2 Organization of this Thesis

Part I of the thesis gives an introduction to the topics of toxicology and systems biology, touching upon the technical platforms of the data used for modeling purposes in this thesis work.

Part II describes the projects of the thesis in separate chapters. The papers from each of the projects are included in their respective chapters.

Part III contains the summarizing perspectives and suggests a few possible future directions for the projects in this thesis work.

[2] Fødevareforlig II

[3] "chemical cocktails" and "chemical mixtures" are used interchangeably throughout this thesis.

Chapter 2

# Toxicology

Toxicology is the field within biology studying potential unwanted effects of chemical substances on living organisms. It is a multidisciplinary field touching upon biochemistry, chemistry, physiology, and pathology and applying many of the same theories and methods as pharmacology [11]. Toxicology in the etymological sense (the science of poisons) has existed since ancient times; however, toxicology as a research field was founded by Paracelsus (1493-1541), also referred to as the "father" of toxicology. He articulated the famous words:

*"All things are poison and nothing is without poison; only the dose makes a thing not a poison"*

which are often condensed to *"The dose makes the poison"*. The French toxicologist, Mathieu Orfila (1787-1853), continuing the work of Paracelsus, is accredited for founding the modern science of toxicology through his analytical work in forensic toxicology. Throughout the 1800s, individual chemicals such as caffeine, nitroglycerin, cocaine, and saccharin were identified, and many more followed in the 1900s. This effort also resulted in the use of chemicals in warfare. During World War I, the Germans were the first to use chemical weapons when they released chlorine gas over a battlefield. World War II accelerated the chemical revolution, which resulted in e.g. the development of powerful nerve gasses. Following World War

II, an array of pesticides was developed and the foundation for an enormous global chemical industry was laid. However, compared to the fields of pharmacology and biochemistry, traditionally, toxicology has a more regulatory basis and a more descriptive scientific character, possibly resulting from the fact that toxicology has been seen as a practical art [11]. Nonetheless, the interest in toxicology is increasing as regulators, scientists, and the public urge more focus on human and animal health and hence safety of chemicals in our surroundings.

## 2.1 Endocrine Disrupting Chemicals

A basal understanding of the endocrine system is essential to understand potential implications of endocrine disrupting chemicals (EDCs) on human health. The endocrine system acts via a coordinated response by signaling from the hormone-producing tissue to a hormone-sensitive target tissue. For most systems, the primary goal is to maintain homeostasis, balancing out changes in hormone levels and responses, which would otherwise have serious effects on the body. One example is the regulation of blood glucose levels by secretion of the hormone insulin from pancreas. Insulin targets various cells with the response to increase uptake of glucose from the blood. Too much or too little insulin causes potentially fatal disturbances in blood glucose levels. Therefore, close regulation of insulin secretion is crucial for homeostasis [12].

EDCs are chemicals with the ability to disrupt any endocrine system in humans or wildlife [13]. EDCs are often thought of as chemicals that affect hormones at the receptor level; however, chemicals interfering with any other part of the hormone signaling pathways, including but not limited to synthesis, degradation, and transport of the hormones, should also be considered as EDCs. Endocrine disruption in hormone-sensitive tissues might lead to adverse effects; however, the timing of exposure might be critical for the effects. As discussed in IPCS/WHO [13], the following points are critical in considering the potential impact of EDCs on bodily functions:

1. Exposure in adulthood may be compensated for by normal homeostatic mechanisms and may therefore not result in any significant or detectable effect.
2. Exposure during the period when programming of the endocrine system is in progress may result in a permanent change of function or sensitivity to stimulatory/inhibitory signals.
3. Exposure to the same level of an endocrine signal at different stages in the life history or in different seasons may produce different effects.

4. Because of cross talk between different endocrine systems, effects may occur unpredictably in endocrine systems other than the system predicted to be affected. This is true for each of the situations in (1) through (3) above.

5. In view of (4), considerable caution should be exercised in extrapolating *in vitro* measures of hormonal activity to the situation *in vivo*.

Examples of effects caused by EDCs include learning disabilities following disruption of thyroid hormone signaling, sex hormone-dependent cancers, and problems with sexual development for males or females following fetal or postnatal exposure to an array of EDCs [14].

### 2.1.1 Male Reproductive Disorders

Regional differences in semen quality within Europe [15] and the U. S. [16] have been observed indicating environmental factors to play a part. Whether or not there is a decreasing trend over time is currently controversial [17, 18]; however, a series of coordinated studies in Germany, Denmark, Sweden, Norway, Finland, Estonia, and Lithuania have concluded that the average sperm count in the included healthy males (aged 18-25 years) was remarkably low [19–22] but not declining [23]. Furthermore, Jø rgensen et al. [20] reported that approximately 20% of young men from Norway and Denmark had sperm concentrations below the World Health Organisation (WHO) reference level for sperm count ($20 \times 10^6$ spermatozoa/ml) and as many as 40% had a sperm count of $< 40 \times 10^6$ spermatozoa/ml. This is worrisome as sperm counts below this figure along with a reduced fraction of morphologically normal sperm are associated with decreased fecundity [24].

The incidence of other male reproductive disorders such as testicular germ cell cancer (TGCC) has increased rapidly over the past 20 years. There is great geographical variation in the incidence with Denmark having the highest incidence within the countries in the the European Union (EU) [25]. Because of the geographical variation the environment is thought to be involved in determining the risk of developing this form of cancer. To support this hypothesis, studies have shown that people from a country with a low incidence of TGCC (e.g. Finland) who migrate to a country with a higher incidence (e.g. Denmark) retain the risk of developing TGCC as their countrymen; however, second generation immigrants (i.e. boys of Finnish parents who immigrated to Denmark) have a similar risk of developing TGCC as native Danes [26].

The predominant male reproductive disorder is cryptorchidism which covers cases where one or both testes are undescended at birth. The reported prevalence varies

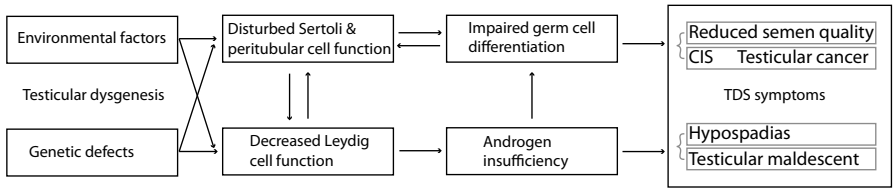



Figure 2.1: Illustration of the components of testicular dysgenesis syndrome (TDS). Modified from [27, 28].

with source of the data (1%-9%) [14]. The presence of environmental chemicals in maternal breast milk [29, 30] but not placenta [31] has been associated with higher incidence of cryptorchidism. Due to differences in procedures and registration of cryptorchidism at hospitals in different countries, registry data are often unreliable. Therefore the basis for assessing whether the incidence rates of cryptorchidism are changing are not very good. Studies have, however, indicated an increasing trend [14, 25]. This male reproductive disorder has also been associated with environmental factors, e.g. documented as a higher incidence of cryptorchid newborns in spring than summer [32], or variation in incidence according to geographical location [33].

The last male reproductive disorder to be mentioned in this work is the congenital malformation known as hypospadias. Hypospadias covers penile malformations where the urethral opening is abnormally placed anywhere on the ventral side of the penis rather than at the tip of glans penis. The prevalence of hypospadias varies with geography in a pattern similar to that of testicular cancer and cryptorchidism [14]. Additionally, increasing trends have been observed in England and Wales, Hungary, Sweden, Norway, and Denmark, whereas no trends have been reported in Finland, Spain, New Zealand, Australia or the former Czechoslovakia [34].

Studies have indicated co-occurrence of the four male reproductive health problems mentioned above; testicular cancer, low sperm quality, cryptorchidism, and hypospadias (Figure 2.1). All conditions are thought to have a common etiology collectively termed the TDS [27, 35]. The components are inter-connected and involve disturbed function of the Sertoli cells, which are the cells in testis that nurse the sperm cells throughout spermatogenesis as well as disturbed function of the Leydig cells, which are the site for steroid hormone production (also termed steroidogenesis). As indicated in Figure 2.2, the severity of TDS is determined by the number of conditions present at one time. Environmental factors such as chemical exposure have been suggested as possible causes of the increasing TDS

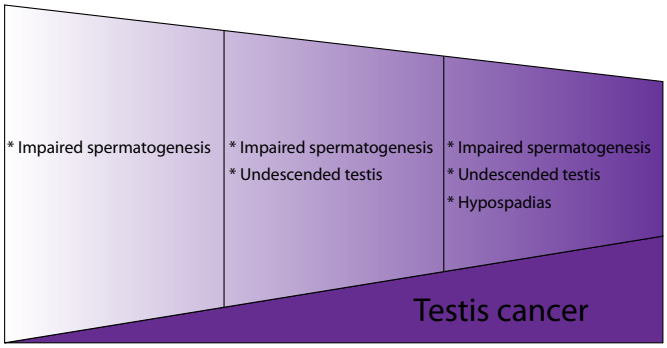



Figure 2.2: Illustration of the relationship between the relative frequency of various symptoms of the TDS. Modified from [27].

incidence, as the increase in incidence has been too rapid to be explained by genetic factors alone [29, 36, 37].

## 2.2 Current Challenges in Toxicology

Regulatory toxicology is facing large challenges due to the fact that the chemical space in our environment/surroundings is so large that classical toxicity testing in animals of all available chemicals is practically impossible. As called for in the before-mentioned vision and strategy [7], a paradigm shift within toxicology is necessary to handle today's challenges. Alongside, recent technological developments have prompted an ocean of data. Handling these amounts pose another challenge in modern toxicology, which calls for tools to handle them.

Aside from the large amount of chemicals queued for risk assessment, humans are simultaneously exposed to multiple chemicals making up so-called chemical cocktails. The toxicology of such cocktails is a complex topic, which is particularly relevant in the area of food and feed safety due to the many sources and routes of exposure to chemicals. Currently, chemical risk assessment is based on the hazard and exposure associated with single chemicals; however, there are concerns that this approach does not provide sufficient security as chemicals might act together to exert effects though they are present at levels below their no observed adverse effect levels (NOAELs). All possible combinations of the tens of thousands of chemicals currently available in our environment makes up a number approximating infinity. Therefore, as mentioned in Section 1.1, we are in urgent need of new ways of

thinking when considering the risk assessment of chemicals as well as chemical mixtures. As outlined by Kortenkamp et al. [38], experimental evidence exists that the combined effects of chemicals in a mixture might be adequately calculated using mathematical modeling. Furthermore, the authors point out that only chemical mixtures that humans are actually exposed to require risk assessment. Examples of such mixtures could be intentionally designed chemical mixtures, chemicals present in cosmetics (e.g. chemical sun filters and parabens) or chemicals known to occur simultaneously in the same body compartment. Furthermore, the authors call for investigations of real-world mixtures such as those of waste water or engine exhaust.

Previous efforts in the 'mixture toxicology' community have aimed to 1) "evaluate and quantify the overall toxicity of complex environmental samples (whole mixture approach)" and 2) "explain the joint action of selected pure compounds in terms of their individual effects (component-based approach)" [38]. In the 'whole mixture approach' a complex environmental sample such as engine exhaust, waste water, or human blood is used as the chemical mixture. This approach is feasible where little is known about the constituents of the sample; however, inference to higher or lower doses of the mixture constituents are very difficult. The 'component-based approach' uses e.g. chemical exposure data as basis for composing the mixture. Using this approach allows for mathematical modeling of the effects of the mixture rather than testing of all possible combinations of chemicals in the mixture [38]. In the Experimental Studies of a Chemical Mixture project the 'component-based approach' was taken using human exposure levels for known EDCs.

Chapter 3

# Computational Systems Biology

Traditionally, there are four accepted fields within biology: genetics and molecular biology, cell biology, biochemistry, and evolutionary biology. Systems biology, however, represents a fifth field, based on the idea that all life forms require the interactions of genes and macro molecules, or cells at a higher scale to exist [39]. Systems biology is often defined in antithesis to the so-called reductionist approach. The reductionist approach includes breaking systems into bits and pieces and then understand the components separately. From that, the idea is to reconstruct the system physically or intellectually [40]. Though the reductionist approach has identified many individual components of biological systems, it lacks the interconnections between the components. Systems biology, on the contrary, is integrating information about all individual components, creating a holistic view of the biological system (Figure 3.1). Therefore, systems biology methods may prove useful for studies of how interactions between the individual components of a biological system give rise to certain behaviors or functions. It is important to notice that the approaches are complementary and they are both of great value to the scientific community.

Some approaches within systems biology could be considered data-driven (Figure 3.2). Studies are data-driven in cases where unbiased experiments and data anal-

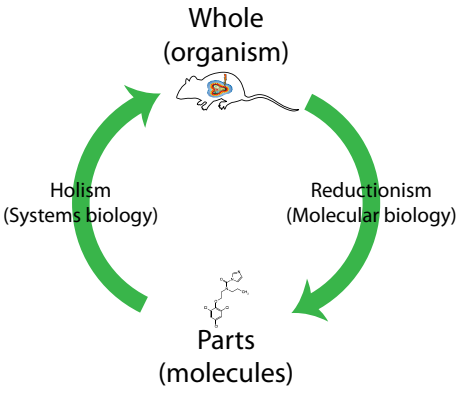



Figure 3.1: Relationship between holistic and reductionist approaches to research. One cannot exist without the other if in-depth knowledge about the system in question is the ultimate goal. Adapted from Kell and Oliver [40].

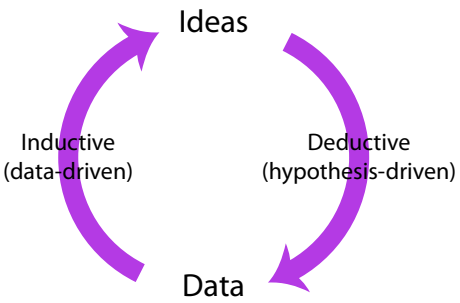



Figure 3.2: The cycle of knowledge, illustrating how ideas and data are interdependent. Ideas lead to data lead to new ideas lead to new data etc. Adapted from Kell and Oliver [40].

yses are carried out without prior knowledge of or theories about the outcome(s), result(s) and/or endpoint(s). This stands in contrast to the hypothesis-driven or bottom-up approach, where the research starts with hypotheses about the behavior of a system, and the experiment aims to elucidate whether the hypotheses holds true or not. One approach often feed into the other, as ideas lead to data, which after interpretation leads to new or refined ideas then leading to new experiments with new data etc. [40].

This PhD project as a whole is based on the hypothesis that computational systems biology can aid toxicological investigations. This also holds true for the individual sub-projects; however, the projects are of a more data-driven character.

In the following sections, I will briefly touch upon some of the data forms used in my PhD project. I will outline the nature of the data forms, and their advantages for use in toxicological research.

## 3.1 Gene Expression Profiling

'Transcriptomics' covers the high-throughput technologies supporting studies of the entire transcriptome on a global scale. The transcriptome can, however, be divided into two types of RNAs: protein coding- and non-coding RNAs. Non-coding RNAs are not translated into specific proteins but maintain structural and regulatory roles.[1] This thesis does not cover the field of non-coding-RNAs but only protein-coding RNA.

Application of gene expression microarrays aims to identify gene sets or pathways that are altered or perturbed upon a change in the environmental conditions (e.g. a disease state or exposure to an environmental chemical). DNA microarrays are collections of thousands of spots, each spot corresponding to a specific gene. At each spot, DNA oligomers are attached to a solid surface. microarray experiments aim to quantify the amount of DNA hybridized to the oligomers thereby measuring gene expression levels for multiple genes simultaneously. microarrays are very rich in features or genes included on the array, however, the cost of microarray experiments is relatively high, which often forces the researchers to keep the sample size relatively small [41, Chapter 7]. When monetary resources are sparse publicly available gene expression datasets can be accessed through databases for use in the initial steps of a research project. The largest repositories of publicly available gene expression data are the Gene Expression Omnibus (GEO) [42, 43] and ArrayExpress [44].

DNA microarray analysis builds on standard molecular biology in that it permits the analysis of gene expression levels similar to real-time quantitative polymerase chain reaction (RT-qPCR), however in a high-throughput fashion. This technology allows the researcher to obtain quantitative gene expression information for a multitude of genes in many samples.

High-throughput microarrays have existed in the scientific community since the 1990s, where the first microarray was developed [45]. Since then, multiple microarray technologies have been developed by competing vendors. The basic concept of the microarray experiment is the same across platforms, but they might differ on areas such as the technique of RNA isolation and labeling, probe design and selection of probes, density of probes on the array, and the technique used to attach the probes to the array [46]. I used Agilent Whole Rat Genome Oligo Microarrays[2] in the Integrative 'OMICs' analysis project. Therefore, I will focus on that technology in the following.

[1] examples of non-coding RNAs include transfer-RNA, micro-RNA, ribosomal-RNA to mention but a few.

[2] SurePrint G3 Rat GE 8x60K Microarray Kit

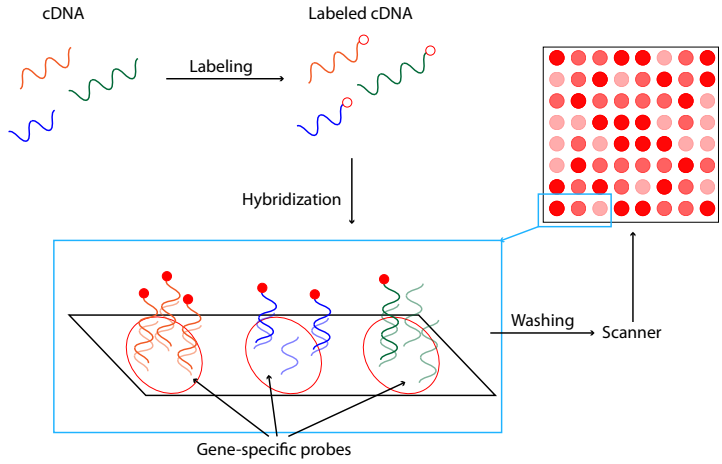



Figure 3.3: Outline of the major steps in a microarray experiment. complementary strand DNA (cDNA) is synthesized from the messenger-RNA (mRNA) in the sample, labeled with fluorescent dye, hybridized onto the array, where gene-specific oligonucleotides are attached. After hybridization the fluorescent probes are excited with laser light and thus emits fluorescent light. After wash-off of un-hybridized labeled cDNA the intensity of the fluorescence is scanned, and the scanner outputs intensity reads for each spot of gene-specific probes.

### 3.1.1 The Microarray Platform

The Agilent arrays used in the Integrative 'OMICs' analysis project utilize nucleic acid hybridization of fluorophore-labeled cDNA targets to match the sense-strand, 60-mer oligonucleotide probes on the array (Figure 3.3). The arrays are manufactured using an industrial inkjet printing process enabling *in situ* synthesis of 60-mer oligonucleotide probes being printed one nucleobase at a time according to digital information. To manufacture these microarrays, Agilent uses phophoramidite chemistry in which drops of activated nucleobases are spotted on the microarrays according to the DNA sequences[3]. This manufacturing process contrasts the one used in deposition microarrays, where prefabricated oligonucleotides are deposited onto the surface of the microarrray[4]. The obtainable spot density when using *in situ* synthesis by spotting results in microarrays with lower spot density compared to similar *in situ* synthesis techniques using e.g. light to activate the nucleobases

[3] `http://www.genomics.agilent.com/article.jsp?pageId=2011`

[4] `http://www.chem.agilent.com/Library/technicaloverviews/Public/5988-8171en.pdf`

for incorporation as is the case with photolithography [47, Chapter 5].

After hybridization of the fluorophore-labeled targets to the probes, un-hybridized cDNA is washed off, the fluorescent probes are exited using laser light and the subsequently emitted fluorescent light is recorded by a scanner 3.3. The scanner then outputs fluorescent light intensity reads for each gene-specific spot on the array.

### 3.1.2 Data Analysis

Prior to statistical analyses of microarray data the arrays from the experiment must undergo quality control (QC) and pre-processing. QC includes a visual inspection of the raw microarray images to detect irregularities and significant flaws, detection of control and background features, and an assessment of the distribution of the signal intensities. This is a critical step as the lack of QC can potentially introduce noise to the data [48, Chapter 1].

Normalizing data between arrays is a step that is undertaken to reduce potential technical variance in the data (i.e. variance that is not caused by the biological differences). There are multiple ways to perform this normalization. The method applied in the Integrative 'OMICs' analysis project is quantile normalization aiming at making the distribution of probe intensities for each array in a set of arrays the same. This method assumes that the majority of features on the arrays are not differentially expressed. Hence, equalizing the distribution of probe intensities across arrays should not suppress the biological variance to a large extent. Other normalization methods include the other 'complete data methods', cyclic loess and contrast normalization, and methods making use of a baseline array, scaling and non-linear, among others [49].

### 3.1.3 Statistical Testing

The statistical method used in the Integrative 'OMICs' analysis project is the moderated $t$ statistic described by Smyth [50]. The moderated $t$ statistic is, as inferred by its name, a moderation over the ordinary $t$ statistic, where the variance and the degrees of freedom are estimated based on the entire dataset; thereby, large statistics are less likely to arise merely from under-estimated sample variances [50].

Microarray experiments pose a major multiple testing problem, given that at least one statistical test is performed for each of the tens of thousands of genes. Gene expression levels are usually not normally distributed, and the expression of one gene is likely dependent on the expression of another gene. Therefore, methods have

been developed to deal with these kinds of challenges. When performing multiple statistical comparisons the number of false discoveries is increased according to theory. Several methods for controlling this increase have been developed. The Bonferroni correction of the familywise error rate (FWER) is considered the most conservative of the methods and can be obtained by correcting the significance level, $\alpha$, by division with the total number of statistical comparisons performed thereby obtaining a more strict $\alpha$. Another way to perform the Bonferroni correction is by correcting all obtained *p*-values by multiplying them with the number of tested hypotheses. Controlling the FWER is important when a conclusion from one hypothesis is likely to be erroneous when at least one other hypothesis is. In other terms, controlling the FWER seeks to reduce the probability of even one false discovery. This method for controlling the FWER was used in the Integrative Systems Biology project [51]. A less stringent method is the Benjamini-Hochberg procedure for controlling the false discovery rate (FDR) [51]. As with the Bonferroni method, either the $\alpha$ or the obtained *p*-values can be corrected to account for the increased false discovery rate following the multiple hypotheses. In contrast to the Bonferroni correction method the Benjamini-Hochberg procedure applies a weight to the correction factor that is inversely correlated to the position in the list of sorted *p*-values (increasing order). Rather than reducing the probability of committing even one type I error, the Benjamini-Hochberg procedure seeks to control the rate at which type I errors occur. Hereby, statistical power is maintained at the cost of type I errors. The Benjamini-Hochberg procedure was applied in the Integrative 'OMICs' analysis project as it has previously proven a good procedure for controlling FDR in microarray experiments [52].

## 3.2 ToxCast™ Data

With its ToxCast program, the U. S. Environmental Protection Agency (US EPA) runs at the forefront in computational toxicology research. In 2007, the ToxCast effort was launched with the aim "*to develop an ability to forecast toxicity based on bioactivity profiling. Ultimately, ToxCast's purpose is to develop methods of prioritizing chemicals for further screening and testing to assist EPA programs in the management and regulation of environmental contaminants* [53]."

### 3.2.1 Phases

The ToxCast program has been conducted in several phases (Figure 3.4). Phase I was the 'Proof of Concept' phase and was completed in 2009. Around 300 data-rich

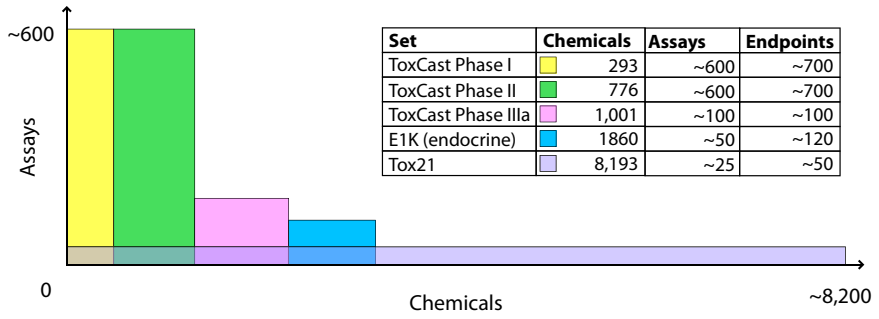



Figure 3.4: Illustration of the number of assays and chemicals in the various datasets included in the ToxCast program. Adapted from Richard Judson, US EPA.

pesticide active ingredients were screened in ∼600 *in vitro* assays totaling ∼700 assay endpoints. ToxCast phase II includes a more diverse set of chemicals comprising ∼800 chemicals from various sources, such as industrial and consumer products, food additives, nanomaterials, and failed pharmaceuticals. These chemicals were evaluated in the same assays as the Phase I chemicals with a few exceptions.

ToxCast phase IIIa was initiated recently and comprise ∼1000 additional chemicals of environmental concern. Additional chemical sets include ∼900 chemicals being screened for potential endocrine disruption (E1K chemicals) in the relevant subset of endocrine-related assays, and Tox21 currently comprising ∼8,000 chemicals being screened in subsets of the ToxCast assays every year.

### 3.2.2 Data formats and origin

The assays included in the ToxCast high-throughput screening effort were run by commercial laboratories, the National Chemical Genomics Center, at the U.S. National Institutes of Health, and at the US EPA. The assay types include biochemical assays (e.g. nuclear receptor binding, enzyme inhibition), cell-based assays (e.g. cytotoxicity profiles, reporter gene assays), complex culture systems (e.g. embryonic stem cell differentiation, inflammatory/angiogenic signals), and small animal models (e.g. zebrafish embryo development) [54]. Most assays target human proteins and are run in human primary cells or transformed cell lines. Most of the commercial laboratories that stated their interest in contributing to the ToxCast program were supporting the pharmaceutical/drug discovery community. Therefore, the applied assays tended to be directed at targets and pathways relevant for pathological processes of interest for the pharmaceutical industry [55]. Most

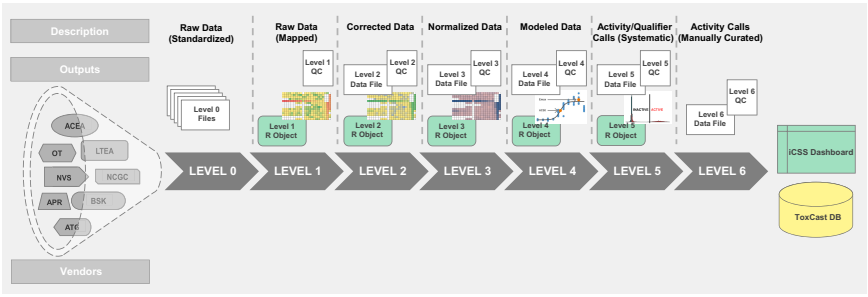



Figure 3.5: The ToxCast data processing workflow. Courtesy of Dayne Filer, US EPA.

assay platforms have an associated descriptive publication [56–70]. With few exceptions, all assays have been run in concentration-response for all chemicals. For the exceptions, chemicals were screened at a high concentration initially to prioritize chemicals for concentration-response testing[71].

In parallel to the *in vitro* ToxCast data, the US EPA have collected available *in vivo* data, mainly on the phase I chemicals, in the Toxicity Reference Database (ToxRefDB). These data comprise publicly available literature data as well as scientific study reports intended for use by the US EPA chemical risk assessors. The data are manually curated and the quality of each study has been assessed for its compliance with current guidelines. These data are available in browsable form via the Aggregated Computational Toxicology Resource (ACToR)[5] and available for download via the ToxCast web-page[6].

### 3.2.3 Data processing

The US EPA have customized a pipeline for analyzing the data (Figure 3.5). This pipeline handles data retrieved from multiple suppliers and thus in multiple formats. To reduce the risk of bias, the chemicals are blinded for the assay providers. The data processing pipeline begins with mapping well-level data followed by chemical identification and concentration, curve-fitting (4-parameter Hill function), and hit-calling. Hit-calling is where a decision is made about whether there is a statistically significant concentration-response. The pipeline handles outliers and considers general toxicity (e.g. cell viability) when determining the hit calls. Furthermore,

[5] `http://actor.epa.gov/toxrefdb/`
[6] `http://epa.gov/ncct/toxcast/data.html`

plate replicates and chemical replicates are used as a means of quality control [54]. All data are available as concentration-response plots via the interactive Chemical Safety for Sustainability (iCSS) Dashboard[7] and the raw and preprocessed data are available via the ToxCast web-page along with further information and documentation about the data[8].

A specific example of the use of the ToxCast data are provided in Section 6.1.

## 3.3 Outlook

In recent decades, advances in technologies have provided the life-science community with large amounts of biological data. This poses an array of challenges regarding storage, transfer, search, sharing etc. but the main issue is the need for improved technological solutions to deal with the data.

In Part II, my three projects including background, hypotheses, overall conclusion, and published peer-reviewed articles or submitted manuscripts will follow. In all three projects I apply methods for handling large amounts of biological data in a toxicological context.

[7] http://actor.epa.gov/dashboard
[8] http://epa.gov/ncct/toxcast/data.html

# Part II

# Projects

Chapter 4

# Integrative Systems Biology

Integrative systems biology often applies statistical modeling to handle and integrate large datasets of different origins from e.g. 'omics' technologies. One example of such data integration is disease enrichment of a set of differentially expressed genes identified by microarray analysis. In such a case, the datasets for integration are the list of significantly differentially expressed genes, and a collection of gene-disease relationships available in databases. A statistical analysis would reveal diseases in which the gene list in question is over-represented compared to random sampling. Systems biology methods may thus prove useful in studies of how interactions between the individual components of a biological system give rise to certain behaviors or functions.

In this chapter, I will summarize the hypotheses and methods used in the Integrative Systems Biology project. Two papers have resulted from this project; a mini-review found on page 25 and an original article found on page 31, respectively. A brief summary of the results and concluding remarks will follow on page 46.

## 4.1 Aims

In addition to providing information about interactions between individual components, using databases and integrating information from various sources can help overview published results. From this overview, one can hypothesize on potential harmful health effects for further investigation. When hypothesizing on health effects of several chemicals data integration might also serve to group chemicals according to their potential modes of action.

Here we describe an example of how and why to apply integrative systems biology methods in the hypothesis-generating phase of toxicological research. The aim of this project was 1) to elucidate modes or mechanisms of action of five pesticidal chemicals (prochloraz, tebuconazole, epoxiconazole, procymidone, and mancozeb) with the aim of grouping them according to their mechanism of action, and 2) to generate hypotheses on the effects of the chemicals on human health.

## 4.2 Methods

For this project, I found inspiration in the recent study by Audouze and Grandjean [72]. In this study, the authors 1) extracted database information about an environmental chemical and its metabolites giving a network of proteins interacting with the chemicals in question, 2) enriched the set of retrieved proteins with known protein-protein interactions (PPIs) resulting in a PPI network for the chemicals in question, 3) enriched the PPI network with known diseases, thus ending up with a list of diseases potentially linked to the chemicals in question. Each of the steps will be touched upon in the sections below and thorough descriptions of the applied methods are available in the papers on page 25 and 31, respectively.

### 4.2.1 Chemical Biology

Chemical biology is the discipline of the effects of (small) molecules on biological systems. These small molecules may be endogenous compounds such as hormones, neurotransmitters, and other signaling molecules, or exogenous compounds such as drug compounds and environmental pollutants. Chemical biology as a research discipline thus covers pharmacology, toxicology, chemistry, and physics.

Several freely available databases contain information about chemicals and their targets. Those used in this project are the Comparative Toxicogenomics Database

(CTD)[1] [73], and the ChemProt database[2] [74]. The information in CTD is manually curated from peer-reviewed scientific literature. ChemProt is a Technical University of Denmark (DTU)-developed resource of annotated and predicted chemical-protein-disease interactions retrieved from a set of databases. For this project, the predicted associations were not retrieved.

For each of the five chemicals, we retrieved lists of target genes/proteins serving as input for subsequent analyses.

### 4.2.2 Protein Interactomics

Biological systems consist of sets of organic compounds ranging from small signaling molecules over peptides and proteins to large, complex, fibrous tissues such as connective tissue or bone. Therefore, when a chemical disrupts the normal function of a protein or signaling molecule it may seem like a subtle effect but if the disrupted protein is part of a large protein cluster it may have severe down-stream effects.

Center for Biological Sequence Analysis, DTU (CBS) has its own collection of PPI networks (InWeb) [75] retrieved from multiple databases with PPI information. InWeb allows the user to retrieve a network of interacting genes/proteins based on an input gene list.

### 4.2.3 Disease Enrichment

The output gene/protein network can subsequently enter disease or pathway enrichment analyses by the use of gene-disease collections such as the Online Mendelian Inheritance in Man (OMIM) database[3] [76] and GeneCards[4] [77] or gene-pathway collections such as Kyoto Encyclopedia of Genes and Genomes (KEGG)[5] [78, 79] and Reactome[6] [80].

The enrichment analysis applied in our experiment is an over-representation analysis based on a hypergeometric distribution.

[1] http://ctdbase.org
[2] http://www.cbs.dtu.dk/services/ChemProt-2.0/
[3] http://www.omim.org
[4] http://www.genecards.org
[5] http://www.genome.jp/kegg/
[6] http://www.reactome.org/PathwayBrowser/

BCPT Basic & Clinical Pharmacology & Toxicology

*Basic & Clinical Pharmacology & Toxicology*, 2014, **115**, 45–49 Doi: 10.1111/bcpt.12216

***MiniReview***

# Applicability of Computational Systems Biology in Toxicology

**Kristine Kongsbak[1,2], Niels Hadrup[1], Karine Audouze[2] and Anne Marie Vinggaard[1]**

[1]Division of Toxicology and Risk Assessment, National Food Institute, Technical University of Denmark, Søborg, Denmark and [2]Department for Systems Biology, Centre for Biological Sequence Analysis, Technical University of Denmark, Kgs. Lyngby, Denmark

(Received 19 November 2013; Accepted 5 February 2014)

*Abstract:* Systems biology as a research field has emerged within the last few decades. Systems biology, often defined as the antithesis of the reductionist approach, integrates information about individual components of a biological system. In integrative systems biology, large data sets from various sources and databases are used to model and predict effects of chemicals on, for instance, human health. In toxicology, computational systems biology enables identification of important pathways and molecules from large data sets; tasks that can be extremely laborious when performed by a classical literature search. However, computational systems biology offers more advantages than providing a high-throughput literature search; it may form the basis for establishment of hypotheses on potential links between environmental chemicals and human diseases, which would be very difficult to establish experimentally. This is possible due to the existence of comprehensive databases containing information on networks of human protein–protein interactions and protein–disease associations. Experimentally determined targets of the specific chemical of interest can be fed into these networks to obtain additional information that can be used to establish hypotheses on links between the chemical and human diseases. Such information can also be applied for designing more intelligent animal/cell experiments that can test the established hypotheses. Here, we describe how and why to apply an integrative systems biology method in the hypothesis-generating phase of toxicological research.

## What is Computational Systems Biology?

Systems biology is often defined as the antithesis of the reductionist approach. Although the reductionist approach has identified many important individual components of biological systems, it often fails to integrate the interconnections between the individual components. Systems biology, on the other hand, deals with the integration of information from a wealth of individual components, creating a holistic view of the biological system [1]. Each component is made up by larger or smaller data sets. For instance, analysis of microarray experiments gives quantitative estimates of changes in expression of a whole-organism genome. High-throughput screening and high-content screening are procedures giving biological activity data on a large number of chemicals (e.g. the U.S. Environmental Protection Agency's ToxCast programme [2]). Additionally, low-/medium-throughput single-target assays provide high-quality measures of the biological function of a single or more targets after chemical exposure. Combining such complementary data types may add value in creating realistic models of potential toxic or adverse effects of chemicals [3,4]. A key strategy to handle multiple data sources is data integration, the use of which has been demonstrated in several applications as reviewed by Mitra *et al.* [5], for example for the identification of new biomarkers for Alzheimer's disease [6].

Computational toxicology integrates molecular biology and chemistry of toxicological interest with mathematical modelling and computational science [7,8] and can therefore be considered a separate branch within computational systems biology. The use of computational toxicology has, for example, proven useful in the development of predictive signatures for various *in vivo* end-points by integration of the ToxCast data with *in vivo* data [9–13] and for the proposal of an adverse outcome pathway for disruption of embryonic vascular development [14].

In addition to providing information about interactions between individual components, using databases and integrating information from various sources can help overview published results. From such an overview, one can suggest potential harmful health effects for further investigation. This has, for example, proven useful in the elucidation of the role of the pesticides rotenone and paraquat in the development of Parkinson's disease [15].

## How to Perform a Network Analysis?

Network analyses can be carried out as previously described [16,17]. When a set of chemicals of toxicological interest has been selected, statistical analyses are carried out using data from various sources. These sources can be (1) databases that contain chemical–target information (chemical biology), (2)

Author for correspondence: Anne Marie Vinggaard, Division of Toxicology and Risk Assessment, National Food Institute, Technical University of Denmark, Mørkhøj Bygade 19, Søborg DK-2860, Denmark (e-mail annv@food.dtu.dk).

databases that contain target–target information such as protein–protein interactions and (3) databases that contain target–human disease information. Below are examples of how to access these data sources.

*Chemical biology.*
Several freely available databases contain information about chemicals and their targets. For example, the Comparative Toxicogenomics Database (CTD, http://ctdbase.org) is a manually curated database with 903.357 (as of 22 January 2014) chemical–gene/protein associations taken from peer-reviewed scientific literature [18]; the database is updated on a monthly basis and contains both experimentally determined and inferred chemical-disease associations. The inferred chemical–disease associations are based on data from various species, which can be considered either an advantage or a disadvantage depending on the aim of the study. In cases where the investigated chemical targets a protein that is not human relevant/not expressed in human beings inclusion of data from other species may add noise to the outcome. On the contrary, when data on human targets or from human tissues are missing, inferring associations to human beings from other species might serve as useful predictions.

An alternative database is the ChemProt database [19]. ChemProt (http://www.cbs.dtu.dk/services/ChemProt-2.0/) is a resource of annotated and predicted chemical–protein–disease interactions. The information is based on integration of data from various resources of experimentally determined chemical–protein interactions and knowledge about proteins involved in disease (retrieved in June 2012) as stated by Kim Kjaerulff *et al.* [19] and reviewed by Panagiotou and Tabourreau [20]; from the web server, the user can select which of the included data resources to use for the analysis. With the latest release, ChemProt contains data for >1,100,000 unique chemicals acting on >15,000 proteins. The predictions are based on the structural features of chemicals. This means that data and predictions for structurally similar compounds can be presented along with the queried compound upon request. Also, the other way around, the queried chemical target predictions are made based on the target(s) of structurally similar compound(s). It has a visual web interface that supports user-friendly navigation through biological activity data and inferred human–disease associations. In these databases, the user can retrieve a list of genes/proteins associated with the query chemical and then use the list in a subsequent analysis for determination of disease links.

HExpoChem, another freely available web resource (http://www.cbs.dtu.dk/services/HExpoChem-1.0/) [21] aims at exploring the exposure and effects of chemicals on human beings by combining chemical–target information, and disease and pathway annotations. HExpoChem takes a different approach as the chemical–protein predictions are based on protein complexes aiming at mimicking the true biological organization as proteins tend to function in groups. Furthermore, HExpoChem identifies proteins interacting with chemicals and consequently suggests chemical–chemical associations.

Other freely available databases worth mentioning in this context are STITCH [22] and ChEMBL [23]. STITCH provides chemical–protein relationships from databases and experiments, and from text mining the co-occurrence of a chemical term and a protein (gene) term in MEDLINE abstracts. ChEMBL contains binding, functional and ADMET (Absorption, Distribution, Metabolism, Elimination, Toxicity) information for a large number of compounds of which the majority are drugs or drug-like compounds.

*Protein–protein interactions.*
As proteins often act in large protein clusters or in concerted pathways, disruption of the function of a protein or signalling molecule by a chemical may seem like a subtle effect; however, if the disrupted protein is part of a large protein cluster, it may have severe downstream effects.

To gather information about physically interacting human proteins in biological systems, several data sets exist. The first example of such a human protein–protein interaction network is Rual's protein–protein interaction network that is based on a high-throughput yeast two-hybrid system applied to test human binary protein–protein interactions [24]. Another such approach is InWeb developed by Lage *et al.* [25–27] with data compiled from various sources. InWeb allows the user to retrieve and predict a network of interacting genes/proteins based on an input gene list.

STRING [28] is a freely available database of experimentally determined and predicted protein–protein interactions. The interactions cover both direct, physical interactions and indirect (functional) associations determined experimentally or mined from literature. The database currently holds information about ~5,200,000 proteins from 1133 organisms.

*Disease and pathway enrichment.*
To obtain information about diseases or pathways possibly perturbed by the investigated chemical, a list of genes/proteins associated with the chemical can enter disease or pathway enrichment analyses. Freely available collections of gene–disease data such as the Online Mendelian Inheritance in Man (OMIM) database (http://www.omim.org) [29] and GeneCards (http://www.genecards.org) [30] or gene-pathway data such as the Kyoto Encyclopedia of Genes and Genomes (KEGG) (http://www.genome.jp/kegg/) [31,32] and Reactome (http://www.reactome.org/PathwayBrowser/) [33,34] can provide the data for such an analysis.

Online Mendelian Inheritance in Man (OMIM) provides information about all known Mendelian disorders and >12,000 human genes. The database is freely available via the web server and is updated on a daily basis. GeneCards is a database of human genes providing comprehensive information on all known and predicted human genes. Among this information is disease information covering heritable disorders but also diseases caused by new mutations or changes to the DNA. GeneCards currently holds information about 19,231 disease-related genes and is freely available for academic non-profit institutions, whereas other users can obtain a commercial licence.

The original concept of KEGG was to create a reference knowledge base of metabolism and other cellular processes, which has now been expanded to contain organismal systems, human diseases and data from various 'omics' platforms. Reactome is a resource of human pathways for basic research, genome analysis, pathway modelling and systems biology.

### A Case Study: Computational Systems Biology Applied on Prochloraz

Using the integrative systems biology approach described above with the pesticide prochloraz as an example, data from various sources were integrated to study the interconnections between (1) prochloraz and its experimentally determined human targets (fig. 1), (2) these human protein/genes and other proteins/genes (fig. 1), and finally (3) these 'secondary' protein/genes and human diseases (fig. 1). Prochloraz is a widely used pesticide that has been demonstrated to have endocrine activity. In animals, it has been shown to cause adverse reproductive outcomes in the male offspring after exposure during foetal life [35–38]. Prochloraz is known to demasculinize male rat foetuses and virilize female foetuses; however, as clinical studies of environmental chemicals are not carried out, little is known about human effects of this compound.

We collected data from the CTD and ChemProt web servers on human tissues giving 22 associations between prochloraz and human gene/protein targets (fig. 1). These targets fall into the categories 'steroidogenesis' (10 targets), 'nuclear receptors' (10 targets) and 'metabolism' (two targets). These initial findings are in line with effects of prochloraz seen in *in vivo* rat studies and *in vitro* studies [38–40]. To widen the basis for hypothesis generation, we expanded the list of 22 associated gene/protein targets to include proteins known to interact with the 22 targets. The reasoning for this operation is that many proteins act in clusters of proteins, and while a direct target of a chemical may not be involved in a given disease, the proteins interacting with the target may be. Therefore, a chemical may cause a disease/adverse outcome through an indirect action on a protein. Using the 22 genes/proteins associated with prochloraz and first-order protein–protein interactions from the high-confidence InWeb data set,

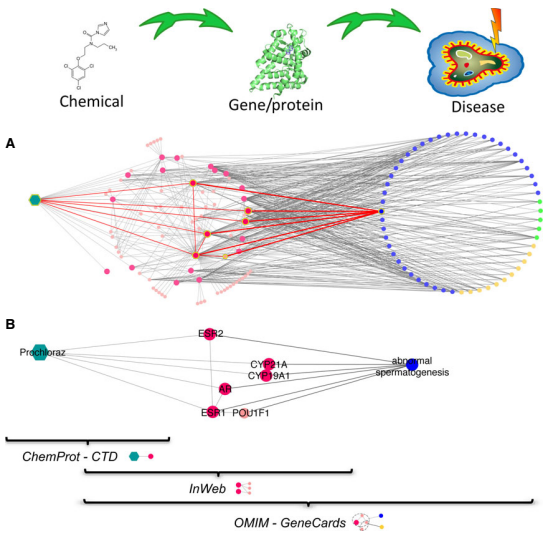



Fig 1. Overview of the integrative systems biology approach employed previously [16,17]. (A) Association between the chemical (dark cyan hexagon) and genes/proteins (magenta circles) was retrieved from the databases ChemProt and Comparative Toxicogenomics Database (CTD). The retrieved list of associated genes/proteins was subsequently expanded using the resource InWeb to include proteins with known physical interactions (pink circles) to the genes/proteins in question. As a final step, we performed an enrichment analysis to identify diseases (blue: reproductive disorders, yellow: adrenal disorders and green: other disorders) in which the genes/proteins were overrepresented using the Online Mendelian Inheritance in Man (OMIM) and GeneCards databases, and a specific data set of genes related to reproductive disorders. (B) A subnetwork of the initial network showing how prochloraz is potentially linked to the development of abnormal spermatogenesis in human beings through interaction with the oestrogen receptors (ESR1 and ESR2), the androgen receptor (AR), two cytochrome P450 enzymes (CYP21A and CYP19A1) and a transcription factor (POU1F1). The brackets indicate the databases from which the data were retrieved.

an additional 56 interacting proteins were added to the network (fig. 1).

The final step of the workflow was to perform a disease and pathway enrichment of the obtained protein network, to generate hypotheses on potential modes of action of the investigated chemicals. From the databases OMIM and GeneCards, we retrieved associations between genes/proteins and diseases. Furthermore, based on information from a review by Matzuk and Lamb [41], information on specific links between genes/proteins and diseases related to male reproduction was used as a separate data set. To determine whether our total set of genes/proteins for prochloraz was likely associated to any diseases, we performed a statistical test based on a hypergeometric distribution. For prochloraz, this data integration resulted in 55 significantly associated disease outcome (39 reproductive disorders, 11 adrenal disorders and five other diseases; fig. 1) out of a total of several thousand diseases. These predictions for effects of prochloraz on human health are in line with experimental data from rats indicating that the main adverse outcomes after exposure to prochloraz are diseases related to reproductive health [35–38]. Another pesticide, mancozeb, underwent the same network analysis and was found to be associated with, for example, inflammatory targets and had no targets in common with prochloraz or other tested pesticides [20]. This outcome is also in line with our experimental data [42] and gives confidence in applying this approach on chemicals for which no animal data exist.

## Conclusion

Applying systems biology methods provides several advantages for toxicologists aiming at investigating potential human effects of a chemical. One advantage is that the scientist will be able to get an overview of potentially harmful effects of a chemical and to generate hypotheses on human adverse outcomes relatively fast compared with a manual literature search; however, the extent of predictions is dependent on the amount of available and published data on the chemical and knowledge of at least one target molecule. Furthermore, where most toxicological data for non-drug compounds are generated from *in vitro* or *in vivo* animal studies and not from human studies, this approach poses an opportunity to suggest human effects of environmental chemicals.

Acknowledgements

The Ministry of Food, Agriculture and Fisheries of Denmark and the Danish Veterinary and Food Administration are acknowledged for their financial support. The Villum Kann Rasmussen Foundation contributed financially to development of the systems toxicology approach.

## References

1 Chandra N. Computational systems approach for drug target discovery. Expert Opin Drug Discov 2009;**4**:1221–36.
2 Dix DJ, Houck KA, Martin MT, Richard AM, Setzer RW, Kavlock RJ. The ToxCast program for prioritizing toxicity testing of environmental chemicals. Toxicol Sci 2007;**95**:5–12.
3 Kitano H. Computational systems biology. Nature 2002;**420**:206–10.
4 Nigsch F, Macaluso NJM, Mitchell JBO, Zmuidinavicius D. Computational toxicology: an overview of the sources of data and of modelling methods. Expert Opin Drug Metab Toxicol 2009;**5**:1–14.
5 Mitra K, Carvunis A-R, Ramesh SK, Ideker T. Integrative approaches for finding modular structure in biological networks. Nat Rev Genet 2013;**14**:719–32.
6 Ma X, Lee H, Wang L, Sun F. CGI: a new approach for prioritizing genes by combining gene expression and protein-protein interaction data. Bioinformatics 2007;**23**:215–21.
7 Kavlock RJ, Ankley G, Blancato J, Breen M, Conolly R, Dix D *et al.* Computational toxicology–a state of the science mini review. Toxicol Sci 2008;**103**:14–27.
8 Reisfeld B, Mayeno AN. What is computational toxicology? Reisfeld B, Mayeno AN, editors. Methods Mol Biol 2012;**929**:3–7.
9 Kleinstreuer NC, Dix DJ, Houck KA, Kavlock RJ, Knudsen TB, Martin MT *et al. In vitro* perturbations of targets in cancer hallmark processes predict rodent chemical carcinogenesis. Toxicol Sci 2013;**131**:40–55.
10 Kleinstreuer NC, Judson RS, Reif DM, Sipes NS, Singh AV, Chandler KJ *et al.* Environmental impact on vascular development predicted by high-throughput screening. Environ Health Perspect 2011;**119**:1596–603.
11 Kleinstreuer N, Dix D, Rountree M, Baker N, Sipes N, Reif D *et al.* A computational model predicting disruption of blood vessel development. PLoS Comput Biol 2013;**9**:e1002996.
12 Sipes NS, Martin MT, Reif DM, Kleinstreuer NC, Judson RS, Singh AV *et al.* Predictive models of prenatal developmental toxicity from ToxCast high-throughput screening data. Toxicol Sci 2011;**124**:109–27.
13 Martin MT, Knudsen TB, Reif DM, Houck KA, Judson RS, Kavlock RJ *et al.* Predictive model of rat reproductive toxicity from ToxCast high throughput screening. Biol Reprod 2011;**85**:327–39.
14 Knudsen TB, Kleinstreuer NC. Disruption of embryonic vascular development in predictive toxicology. Birth Defects Res C Embryo Today 2011;**93**:312–23.
15 Qi Z, Miller GW, Voit EO. Rotenone and paraquat perturb dopamine metabolism: a computational analysis of pesticide toxicity. Toxicology 2014;**315**:92–101.
16 Audouze K, Grandjean P. Application of computational systems biology to explore environmental toxicity hazards. Environ Health Perspect 2011;**119**:1754–9.
17 Kongsbak K, Vinggaard AM, Hadrup N, Audouze K. A computational approach to mechanistic and predictive toxicology of pesticides. ALTEX 2013;**14**:1–17.
18 Davis AP, Murphy CG, Johnson R, Lay JM, Lennon-Hopkins K, Saraceni-Richards C *et al.* The comparative toxicogenomics database: update 2013. Nucleic Acids Res 2013;**41**(Database issue): D1104–14.
19 Kim Kjærulff S, Wich L, Kringelum J, Jacobsen UP, Kouskoumvekaki I, Audouze K *et al.* ChemProt-2.0: visual navigation in a disease chemical biology database. Nucleic Acids Res 2013;**41**(Database issue):D464–9.
20 Panagiotou G, Taboureau O. The impact of network biology in pharmacology and toxicology. SAR QSAR Environ Res 2012; **23**:221–35.
21 Taboureau O, Jacobsen UP, Kalhauge C, Edsgärd D, Rigina O, Gupta R *et al.* HExpoChem: a systems biology resource to explore human exposure to chemicals. Bioinformatics 2013;**29**: 1231–2.
22 Kuhn M, Szklarczyk D, Franceschini A, Campillos M, von Mering C, Jensen LJ *et al.* STITCH 2: an interaction network database for small molecules and proteins. Nucleic Acids Res 2010;**38**(Database issue):D552–6.

23 Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A *et al.* ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res 2012;**40**(Database Issue):D1100–7.

24 Rual J-F, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A, Li N *et al.* Towards a proteome-scale map of the human protein-protein interaction network. Nature 2005;**437**:1173–8.

25 Lage K, Møllgård K, Greenway S, Wakimoto H, Gorham JM, Workman CT *et al.* Dissecting spatio-temporal protein networks driving human heart development and related disorders. Mol Syst Biol 2010;**6**:381.

26 Lage K, Karlberg EO, Størling ZM, Olason PI, Pedersen AG, Rigina O *et al.* A human phenome-interactome network of protein complexes implicated in genetic disorders. Nat Biotechnol 2007;**25**:309–16.

27 Lage K, Hansen NT, Karlberg EO, Eklund AC, Roque FS, Donahoe PK *et al.* A large-scale analysis of tissue-specific pathology and gene expression of human disease genes and complexes. Proc Natl Acad Sci 2008;**105**:20870–5.

28 Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A *et al.* STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res 2013;**41**(Database Issue):D808–15.

29 Online Mendelian Inheritance in Man, OMIM®. McKusick-Nathans Institute of Genetic Medicine Johns Hopkins University. 2012.

30 Safran M, Dalah I, Alexander J, Rosen N, Iny Stein T, Shmoish M *et al.* GeneCards Version 3: the human gene integrator. Database 2010;**2010**:baq020.

31 Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000;**28**:27–30.

32 Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res 2014;**42**:D199–205.

33 Milacic M, Haw R, Rothfels K, Wu G, Croft D, Hermjakob H *et al.* Annotating cancer variants and anti-cancer therapeutics in reactome. Cancers 2012;**4**:1180–211.

34 Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G *et al.* The Reactome pathway knowledgebase. Nucleic Acids Res 2014;**42**:D472–7.

35 Vinggaard AM, Nellemann C, Dalgaard M, Jørgensen EB, Andersen HR. Antiandrogenic effects *in vitro* and *in vivo* of the fungicide prochloraz. Toxicol Sci 2002;**69**:344–53.

36 Vinggaard AM, Christiansen S, Laier P, Poulsen ME, Breinholt V, Jarfelt K *et al.* Perinatal exposure to the fungicide prochloraz feminizes the male rat offspring. Toxicol Sci 2005;**85**:886–97.

37 Laier P, Metzdorff SB, Borch J, Hagen ML, Hass U, Christiansen S *et al.* Mechanisms of action underlying the antiandrogenic effects of the fungicide prochloraz. Toxicol Appl Pharmacol 2006;**213**:160–71.

38 Noriega NC, Ostby J, Lambright C, Wilson VS, Gray LE. Late gestational exposure to the fungicide prochloraz delays the onset of parturition and causes reproductive malformations in male but not female rat offspring. Biol Reprod 2005;**72**:1324–35.

39 Dreisig K, Taxvig C, Birkhøj Kjærstad M, Nellemann C, Hass U, Vinggaard AM. Predictive value of cell assays for developmental toxicity and embryotoxicity of conazole fungicides. ALTEX 2013;**30**:319–30.

40 Vinggaard AM, Hass U, Dalgaard M, Andersen HR, Bonefeld-Jørgensen E, Christiansen S *et al.* Prochloraz: an imidazole fungicide with multiple mechanisms of action. Int J Androl 2006;**29**:186–92.

41 Matzuk MM, Lamb DJ. The biology of infertility: research advances and clinical challenges. Nat Med 2008;**14**:1197–213.

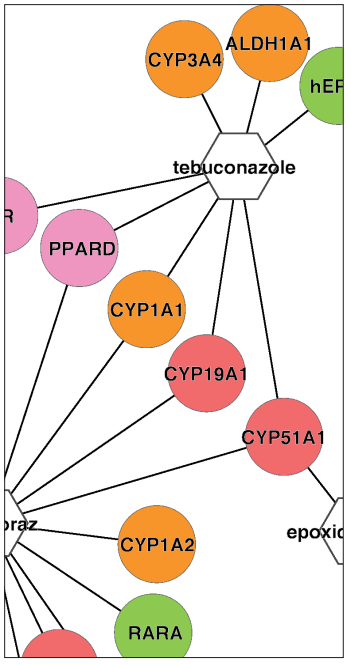
42 Axelstad M, Boberg J, Nellemann C, Kiersgaard M, Jacobsen PR, Christiansen S *et al.* Exposure to the widely used fungicide mancozeb causes thyroid hormone disruption in rat dams but no behavioral effects in the offspring. Toxicol Sci 2011;**120**:439–46.

1/2014
Volume 31, No. 1
ISSN 1868-596X
1-104 (2014)

# ALTEX

ALTERNATIVES TO ANIMAL EXPERIMENTATION








Springer Spektrum

# A Computational Approach to Mechanistic and Predictive Toxicology of Pesticides

*Kristine Kongsbak [1,2], Anne Marie Vinggaard [2], Niels Hadrup [2], and Karine Audouze [1]*

[1]Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, Kongens Lyngby, Denmark; [2]Division of Toxicology and Risk Assessment, National Food Institute, Technical University of Denmark, Søborg, Denmark

**Summary**

*Emerging challenges of managing and interpreting large amounts of complex biological data have given rise to the growing field of computational biology. We investigated the applicability of an integrated systems toxicology approach on five selected pesticides to get an overview of their modes of action in humans, to group them according to their modes of action, and to hypothesize on their potential effects on human health.*

*We extracted human proteins associated with prochloraz, tebuconazole, epoxiconazole, procymidone, and mancozeb and enriched each protein set by using a high confidence human protein interactome. Then we explored modes of action of the chemicals by integrating protein-disease information into the resulting protein networks. The dominating resulting human adverse effects were reproductive disorders followed by adrenal diseases.*

*Our results indicated that prochloraz, tebuconazole, and procymidone exert their effects mainly via interference with steroidogenesis and nuclear receptors. Prochloraz was associated with a large number of human diseases and, together with tebuconazole, showed several significant associations with testicular dysgenesis syndrome. Mancozeb showed a differential mode of action involving inflammatory processes. This method provides an efficient way of overviewing data and grouping chemicals according to their mode of action and potential human adverse effects. Such information is valuable when dealing with predictions of mixture effects of chemicals and may contribute to the development of adverse outcome pathways.*

*Keywords: computational tool, pesticides, grouping, systems biology, testicular dysgenesis syndrome*

## 1 Introduction[1]

Increasing evidence shows that adverse human health effects like male and female reproductive disorders are increasing. Around 8% of all children are currently conceived by *in vitro* fertilization and up to 9% of newborn boys are born with malformed reproductive organs like cryptorchidism in Denmark (Wohlfahrt-Veje et al., 2009). Chemical exposure may contribute to the increasing prevalence of these health disorders (Wohlfahrt-Veje et al., 2009; Weidner et al., 1998; Swan et al., 2005). Previous epidemiological research has indicated associations between pesticide exposure and endocrine disruption leading to, e.g., poor semen quality (Swan et al., 2003) and increased incidence of cryptorchidism (Weidner et al., 1998; Damgaard et al., 2006).

Other studies have shown the simultaneous occurrence of various endocrine active chemicals in human body fluids, such as urine (Swan et al., 2005) and breast milk (Damgaard et al., 2006; Krysiak-Baltyn et al., 2012). These findings have been correlated with incidences of malformed reproductive organs such as decreased anogenital distance and cryptorchidism in boys of exposed mothers, respectively. These epidemiological studies indicate that prenatal chemical exposure might be a risk factor for congenital reproductive abnormalities. Experimental studies have shown that mixtures of chemicals can result in substantial effects on various endpoints in spite of doses of the individual chemicals being too low to exert effects on their own. This has been demonstrated both *in vitro* (Rajapakse et al., 2002; Silva et al., 2002) and *in vivo* (Hass et al., 2007, 2012; Metzdorff et al., 2007; Christiansen et al., 2008, 2009; Jacobsen

Received April 24, 2013; accepted in revised form September 6, 2013; Epub September 17, 2013; http://dx.doi.org/10.14573/altex.1304241

[1] **Abbreviations:** CTD, Comparative Toxicogenomics Database; CYP, cytochrome P450; ODS, ovarian dysgenesis syndrome; OMIM, Online Mendelian Inheritance in Man; PPI, protein-protein interaction; TDS, testicular dysgenesis syndrome

et al., 2012). It is often a challenge within epidemiological research to find significant associations between human diseases and exposure to single chemicals. This may be partly due to the fact that humans are seldom exposed to one single chemical at a given time point, but are most likely exposed to complex mixtures of chemicals (Krysiak-Baltyn et al., 2010).

Today's toxicologists are facing several challenges. One challenge is how to handle risk assessment of mixtures. Risk assessment of chemicals has traditionally focused on the effects of individual chemical exposures. Currently, risk assessment of chemicals, including pesticides, is based on the no observed adverse effect levels (NOAELs) for effects of single compounds. Based on results from animal studies, exposure to single environmental chemicals generally does not cause major concern for adverse reproductive effects in humans due to low exposure levels. Humans are, however, exposed to a mixture of several chemicals and, during the past decade, scientific and regulatory focus has gradually begun shifting towards examining the effects of mixtures. Since 2005, the European Union member states have for example been obliged to evaluate and, if possible, refine existing methodologies in order to take combined actions of pesticides into account during risk assessment and, especially, when establishing maximum residue levels for pesticides (European Commission, 2005). Due to the high number of possible chemical combinations, the option of basing human risk assessment of mixtures on experimental data of each possible chemical combination is not practically feasible. Therefore, development of new approaches is needed for handling the challenge of human risk assessment of mixtures. The European Food Safety Authority (EFSA) has just released the report, "International Framework Dealing with Human Risk Assessment of Combined Exposure to Multiple Chemicals" presenting considerations on how to evaluate and handle mixture effects (European Food Safety Authority, 2013). The current thinking is that human risk for mixture exposure can be predicted by selecting a reasonable mathematical model (like the concentration-addition model (Hermens et al., 1984; Konemann, 1981)), and quantitative measures of toxic effects (e.g., NOAEL), as well as exposure levels for the individual chemicals (Kortenkamp et al., 2009; Hass et al., 2007; Hadrup et al., 2013). Based on this line of thought, one future challenge in risk assessment of mixtures is potentially lack of sufficient homogenous toxicological information on single compounds and their mechanisms of action. A second challenge for the toxicologist is the exponentially growing information load on chemicals in scientific databases, which poses the challenge of how to digest and utilize all existing information. A third challenge is how to predict adverse effects in humans based on toxicological information usually obtained from other species.

A potential solution to some of these challenges may be the application of systems toxicology. Systems toxicology has previously been employed as a means to explore potential adverse effects of environmental chemicals (Audouze and Grandjean, 2011). One of the strengths of the procedure is that focus can be directed exclusively towards human effects, avoiding the challenge of inter-species extrapolation.

In this paper, we present an evaluation of the procedure for grouping of chemicals according to their effects and modes of

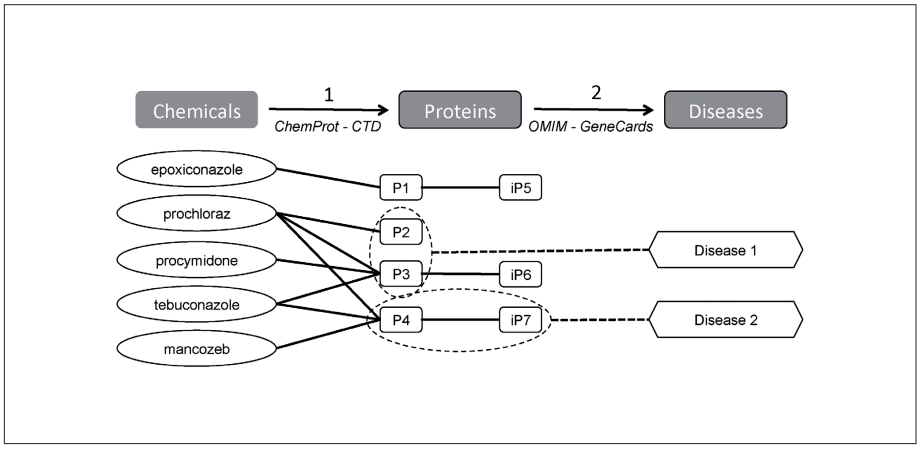



**Fig. 1: Workflow of the proposed multi-step systems chemical biology approach**
1. Proteins (P) known to be biologically associated with the five chemicals are extracted using two different databases (ChemProt and CTD). For each chemical, the set of proteins has been enriched with high confidence human protein-protein interaction data (iP).
2. Integration of disease information using three databases (GeneCards, OMIM, and a specific dataset related to male infertility) to each of the protein complexes. Diseases are statistically ranked to help link the chemicals to potential phenotypic outcomes.

action that may be of benefit for a mixture risk assessment of chemicals for which homogenous toxicological information is lacking. We investigated the effects of five pesticides using a recently developed systems toxicology approach. This study served multiple purposes: 1) to apply systems toxicology to investigate modes of action of individual chemicals, 2) to explore the value of the approach for grouping of chemicals according to their mode of action, and 3) to hypothesize on new, unexpected effects of the chemicals in humans.

## 2 Methods

Procymidone, mancozeb, and the azole fungicides epoxiconazole, prochloraz, and tebuconazole were investigated using the multi-step data integration workflow (Fig. 1). In the first step, we used existing knowledge from disease chemical biology databases to extract information on specific human chemical-protein associations. We queried for associations of the selected chemicals with genes and/or proteins in several sources of experimental data. Secondly, these chemical-protein associations were expanded to protein complexes. By using a high-confidence set of experimental protein-protein interactions, we identified protein complexes associated with the chemicals (Lage et al., 2007). Finally, protein-disease annotations were integrated into these protein complexes in order to statistically rank the chemicals' relation with diseases.

*Chemical-protein associations*
The first step was to extract available information for the five chemicals of interest. Only human related data was taken into consideration. Known chemical-protein associations were compiled from two publicly available databases, the Comparative Toxicogenomics Database (CTD) (Davis et al., 2011) and ChemProt 2.0 (Kim Kjærulff et al., 2013), accessed on October 8, 2012). The CTD is a database of manually curated chemical-gene associations mined from peer-reviewed scientific literature. The Batch Query search method was used with default settings, searching for curated chemical-gene interactions using each chemical as input. ChemProt 2.0 is a newly established disease chemical biology database containing chemical-protein annotation resources for more than 1,100,000 unique chemicals and more than 15,000 proteins.

*Protein-protein interactions*
As proteins tend to function in protein complexes, we expanded each of the five obtained protein lists to contain known protein-protein interactions (PPIs). We used a high confidence human protein interactome (Lage et al., 2007) (version 3.0), which is based on experimental data from humans and data inferred from model organisms. The version that was used (3.0) contains 507,142 unique PPIs and 22,997 genes, as of October 2012. The protein-protein interaction network used for our analyses is a scored network relying on network topology and reliability of each individual interaction. For further details on the construction and application of the protein-protein interaction network, please see Lage et al. (2007, 2008, 2010). The human protein interactome was used to expand the initial networks by adding to the network proteins known to physically interact with the genes/proteins from the initial network.

*Disease enrichment*
To identify diseases possibly related to each individual chemical, we integrated disease information from three sources in each of the five protein complexes. The Online Mendelian Inheritance in Man (OMIM) database (McKusick-Nathans Institute of Genetic Medicine), the GeneCards database (Safran et al., 2010), and a specific dataset related to male infertility based on a review (Matzuk and Lamb, 2008) were used for data integration. The OMIM database is a comprehensive compendium of relationships between human genotypes and phenotypes. GeneCards is a searchable, integrated database providing information on human genes and selected gene-related information, such as functional and disease information.

The protein complexes for each of the five chemicals were individually tested for significant disease associations using a test based on a hypergeometric distribution. A significance level of 0.05 after Bonferroni correction for multiple testing of $p$-values was used to select the most relevant associations. The gene list for prochloraz is used here as an example in the test for an association with the disease adrenal hyperplasia: 46 out of 6778 genes in the GeneCards database were associated with the disease, and of the 27 genes associated with prochloraz, 10 were associated with the disease. This gave a $p$-value of 5.7e-16, and the Bonferroni corrected $p$-value, when accounting for 10,428 diseases in the database, was 5.9e-12, which is below the chosen significance level of 0.05.

To investigate unexpected chemical-disease associations, we included non-significant associations for chemicals if the disease in question was significantly associated with at least one other chemical.

## 3 Results

*Protein lists*
Using the ChemProt 2.0 and CTD databases, we extracted one relevant human protein target for epoxiconazole, 22 for prochloraz, 5 for procymidone, 8 for tebuconazole, and 11 for mancozeb (Fig. 2). As 7 of the genes are shared between two or more of the chemicals, the complete gene list contains 38 different genes. Figure 3 shows a unified view of the respective associations. This figure demonstrates that there are a few shared gene associations for the chemicals. The androgen receptor (AR) is a shared target for prochloraz, procymidone, and tebuconazole. The cytochrome P450 (CYP) enzyme lanosterol-14-α-demethylase (CYP51A1) is a shared target for all three azole fungicides in accordance with their desired mechanism of action as fungicides. Prochloraz affects many proteins in the steroidogenic pathway and only CYP19A1 is a shared target with tebuconazole. Also, peroxisome proliferator-activated receptor δ (PPARδ) is a common target for prochloraz and tebuconazole, and PPARδ is affected by prochloraz. In particular, mancozeb appears to have very distinct associations compared with the

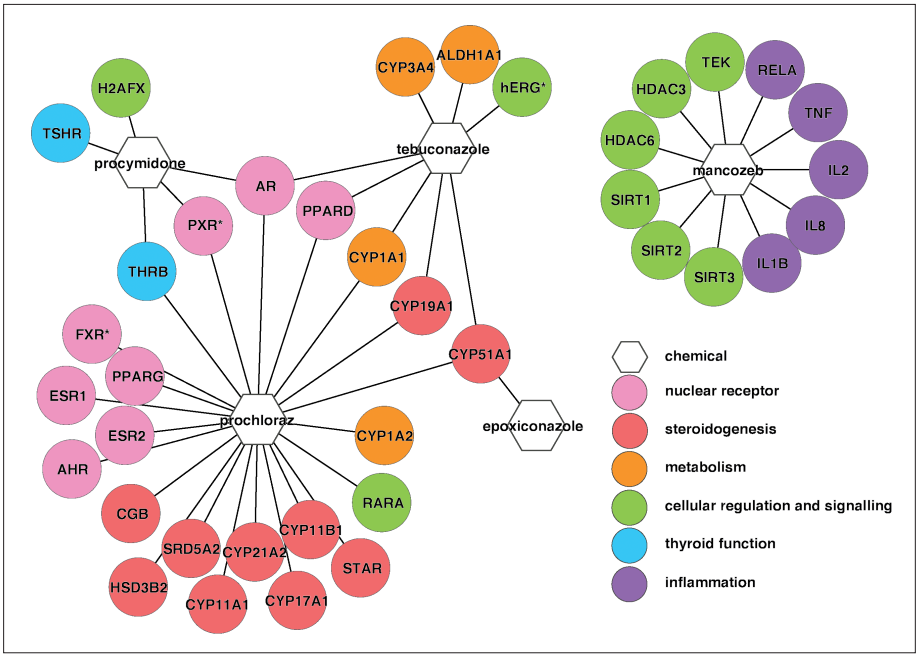



**Fig. 2: Chemical-protein association network**
Network view of the five chemicals and their associated genes denoted by HUGO gene symbol. Asterisks (*) indicate that the notation refers to the gene product rather than the gene itself. The colors of the genes correlate with primary function of the protein encoded by the genes. Red: transcription factor, orange: steroidogenesis, yellow: metabolism, green: cellular regulation and signaling, blue: thyroid function, and purple: inflammation. Chemicals are denoted as white hexagons. Mancozeb does not share any associations with the four other chemicals, and only a few genes connect the remaining chemicals.

remaining four chemicals. It targets mainly proteins involved in inflammatory and transcriptional processes. The thyroid receptor (THRB) is suggested as a target of both procymidone and prochloraz, whereas only procymidone affects the thyroid-stimulating-hormone receptor (TSHR).

*Protein-protein interactions*
For each chemical we generated individual PPI networks by determining PPI partners for the proteins associated with the chemicals. To assess overlapping proteins, we generated a merged network. In this network, the chemicals were connected to their respective associated proteins retrieved from the initial step. The PPI networks were then viewed simultaneously (see Fig. S1 in supplementary data at http://www.altex-edition.org). As shown in Table 1, no proteins were identified in the epoxiconazole network, 81 in the prochloraz network, 18 in the procymidone network, 31 in the mancozeb network, and 25 in the tebuconazole network. This adds up to a total of 142 different

**Tab. 1: Number of human proteins associated with the five chemicals within the different steps of the systems biology procedure**
Numbers in parentheses indicate the number of connections between proteins in the network.

| Chemical name | Number of proteins | |
|---|---|---|
| | ChemProt, CTD | Proteins (connections) |
| Epoxiconazole | 1 | 1 (0) |
| Prochloraz | 22 | 81 (123) |
| Procymidone | 5 | 18 (16) |
| Mancozeb | 11 | 31 (28) |
| Tebuconazole | 8 | 25 (22) |
| Mixture | 38 | 142 |

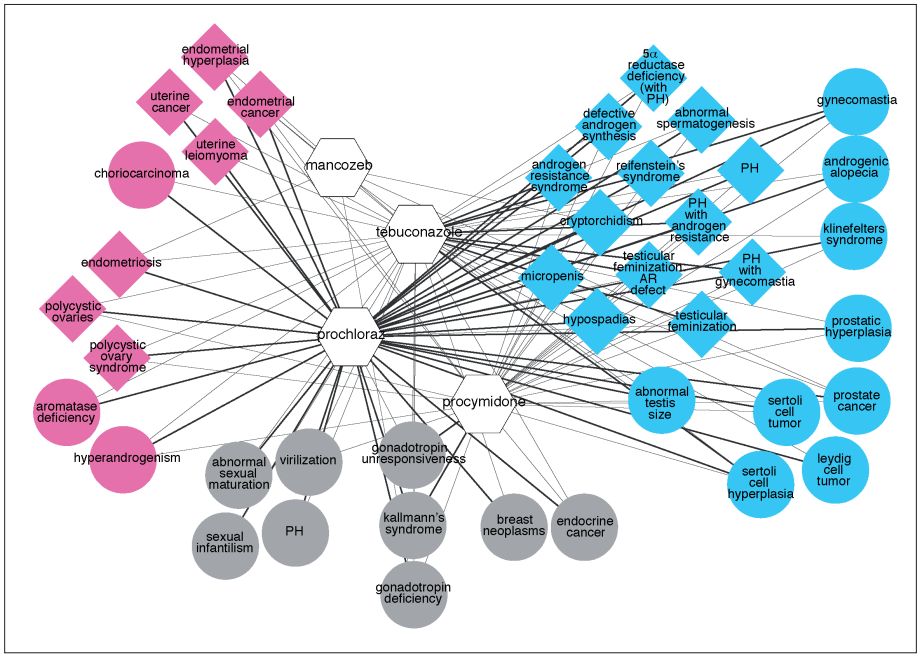



**Fig. 3: Network showing the reproductive diseases that are associated with the four chemicals**
The chemicals are the white hexagons, whereas female reproductive diseases are pink, male reproductive diseases are blue, and diseases affecting both genders or unclear phenotypes are grey. Blue diamonds indicate diseases that may be involved in testicular dysgenesis syndrome (TDS), whereas red diamonds indicate diseases that may be involved in ovarian dysgenesis syndrome (ODS). Black lines indicate significant ($p$-value ≤0.05) associations, whereas grey lines indicate non-significant associations ($p$-values ≥0.05). PH, pseudo hermaphroditism; AR, androgen receptor

proteins. The sub-network associated with mancozeb is completely isolated from other protein complexes. Considering the above results, epoxiconazole was not included in further calculations.

*Translation into disease associations*

To identify diseases associated with the individual chemicals, each PPI network was enriched using OMIM, GeneCards, and the male infertility dataset. This procedure provides indications of which chemicals might be associated with human disease(s). Instead of looking exclusively at statistically significant diseases, we increased the number of associations by including non-significant associations between diseases and chemicals if the disease in question was retrieved on the basis of a low p-value for at least one other chemical out of the five. One such example is the association between the chemicals and the disease term "micropenis". Prochloraz is the only chemical significantly associated with this disease term with a $p$-value corrected for multiple testing of 0.002. Among the 81 proteins in the prochloraz network, only 41 were found in the GeneCards database. Among these 41 proteins, 5 were known to be associated with the disease "micropenis" as compared to a total of 9 known genes involved in the disease out of 5515 genes in the database. Tebuconazole and procymidone also showed associations with the condition "micropenis", however these associations were both non-significant due to the low number of associated genes. Despite the lack of significance, the association was extracted for further analyses to indicate potential associations with hypothetical adverse outcome pathways.

Three clusters of diseases appeared: One containing 41 reproductive diseases, a second one with seven disorders related to the adrenal gland, and a last one of six other diseases including adenoma, Antley-Bixler syndrome, Cushing syndrome, granulosa cell tumor, isolated deficiency of pituitary hormone, and xanthomatosis cerebrotendinous (see Fig. S2 in supplementary data at http://www.altex-edition.org).

The large cluster of reproductive disorders was divided into three sub-clusters: male, female, and unisex (Fig. 3). The male sub-group contains 20 diseases. Among them, 13 seem to be associated with testicular dysgenesis syndrome (TDS), which covers a range of male reproductive diseases or disorders like malformed reproductive organs, cryptorchidism, poor sperm quality, and testicular cancer. These disorders are believed to reflect various stages of TDS arising during gestation (Skakkebaek et al., 2001). The 13 disorders covered by this analysis included androgen insensitivity syndrome, various types of pseudohermaphroditism, malformed or wrongly positioned testes, malformations of the penis including hypospadias, and underlying mechanisms for these disorders, such as androgen resistance syndrome and defective biosynthesis of testicular androgen. Additionally, various types of hyperplasia and prostate disease, and the terms gynecomastia, androgenic alopecia, Klinefelters syndrome, and abnormal testis size are associated with one or more of the chemicals through the chemical-gene associations (Fig. 3 and Tab. 2).

The female sub-cluster contains 10 diseases, of which seven diseases are associated with the ovarian dysgenesis syndrome (ODS), a recently suggested hypothesis parallel to the TDS paradigm, suggesting a shared etiology for endometriosis, ovarian cancers, and other diseases related to female fecundity (Louis et al., 2011). In this cluster, three subgroups of diseases exist: 1) disease terms related to various cancers within the female reproductive system, 2) polycystic ovaries and endometriosis, and 3) diseases related to abnormal hormone levels. All of these diseases are significantly associated with prochloraz and none of the other investigated chemicals. *P*-values after Bonferroni correction for multiple testing and the genes describing the chemical-disease associations are summarized in Table 2.

**Tab. 2: Diseases extracted for the respective chemicals**
The diseases are grouped according to gender specificity. Bonferroni corrected *p*-values and the genes associated with the diseases and involved in the respective diseases are listed by HUGO gene symbol. Statistically significant *p*-values are shown in bold format. N/A values indicate lack of associations. N/S indicates Bonferroni corrected *p*-values ≥1.

| | | prochloraz | | procymidone | | tebuconazole | | mancozeb | |
|---|---|---|---|---|---|---|---|---|---|
| | disease | Genes | bonf. *p*-val | genes | bonf. *p*-val | genes | bonf. *p*-val | genes | bonf. *p*-val |
| Male | gynecomastia | CGB; CYP17A1; CYP19A1; CYP21A2; ESR1; HSD3B1; HSD3B2; HSD17B3; AR; SRD5A2 | **5.E-11** | AR | 0.79 | CYP19A1; AR | **0.03** | NONE | N/A |
| | hypospadias | CYP1A1; CYP17A1; CYP21A2; ESR1; ESR2; HSD3B1; HSD3B2; HSD17B3; AR; SRD5A2 | **4.E-08** | AR | N/S | AR | N/S | NONE | N/A |
| | PH with gynecomastia | CYP19A1; HSD3B2; HSD17B3; AR; SRD5A2 | **5.E-08** | AR | N/S | CYP19A1; AR | **0.01** | NONE | N/A |
| | PH with androgen resistance | CYP11B1; CYP19A1; HSD3B2; AR | **3.E-05** | AR | N/S | CYP19A1; AR | **0.02** | NONE | N/A |
| | reifenstein's syndrome | CYP11B1; CYP19A1; HSD3B2; AR | **3.E-05** | AR | N/S | CYP19A1; AR | **0.02** | NONE | N/A |
| | leydig cell tumor | CYP11A1; CYP17A1; CYP19A1; CYP21A2; ESR1; ESR2; STAR | **9.E-05** | NONE | N/A | CYP19A1 | N/S | NONE | N/A |
| | 5α reductase deficiency (with PH) | HSD17B3; AR; SRD5A2 | **1.E-04** | AR | 0.63 | AR | N/S | NONE | N/A |
| | PH | CYP17A1; HSD3B2; HSD17B3; AR; SRD5A2 | **3.E-04** | AR | N/S | AR | N/S | NONE | N/A |
| | prostate cancer | CGB; CYP1A1; CYP1A2; CYP11A1; CYP17A1; CYP19A1; AHR; ESR1; ESR2; SPDEF; HSD3B1; HSD3B2; HSD17B3; AR; PPARD; PPARG; PMEPA1; RARA; SRD5A2; NR1I2; EBAG9 | **1.E-03** | SPDEF; AR; PMEPA1 | N/S | CYP1A1; CYP3A4; CYP19A1; SPDEF; AR; KCNH2; PPARD; PMEPA1 | N/S | ANGPT2; RELA; TEK | N/S |

| | | prochloraz | | procymidone | | tebuconazole | | mancozeb | |
|---|---|---|---|---|---|---|---|---|---|
| | disease | Genes | bonf. *p*-val | genes | bonf. *p*-val | genes | bonf. *p*-val | genes | bonf. *p*-val |
| Male | abnormal testis size | CYP19A1; ESR1; ESR2; AR; PPARG; STAR | **2.E-03** | AR | 0.40 | AR | N/S | NONE | N/A |
| | cryptorchidism | CYP17A1; CYP19A1; ESR1; AR; PPARG; STAR | **2.E-03** | AR | N/S | CYP19A1; AR | N/S | IL1B | N/S |
| | micropenis | CYP17A1; AHR; HSD3B2; AR; SRD5A2 | **2.E-03** | AR | N/S | AR | N/S | NONE | N/A |
| | androgen resistance syndrome | CYP19A1; ESR1; HSD3B2; AR; PPARG; SRD5A2 | **2.E-03** | AR | N/S | CYP19A1; AR | **0.03** | NONE | N/A |
| | defective androgen synthesis | CYP17A1; CYP21A2; AR | **3.E-03** | AR | 0.13 | AR | 0.40 | NONE | N/A |
| | androgenetic alopecia | CYP19A1; CYP21A2; AR; SRD5A2 | **4.E-03** | AR | N/S | CYP19A1; AR | N/S | NONE | N/A |
| | prostatic hyperplasia | CYP17A1; CYP19A1; ESR2; AR | **0.01** | AR | N/S | CYP19A1; AR | N/S | NONE | N/A |
| | abnormal spermatogenesis | CYP19A1; CYP21A2; ESR1; ESR2; AR; POU1F1 | **0.02** | AR | 0.53 | AR | N/S | NONE | N/A |
| | klinefelters syndrome | CGB; CYP19A1; CYP21A2; AR | **0.05** | AR | N/S | CYP19A1; AR | N/S | NONE | N/A |
| | sertoli cell hyperplasia | CYP11A1; CYP19A1; ESR1; ESR2; AR; STAR | **2.E-06** | AR | 0.40 | CYP19A1; AR | **0.05** | TNF | 0.56 |
| | sertoli cell tumor | CYP19A1; ESR1; ESR2; AR | **0.02** | AR | N/S | CYP19A1; AR | 0.79 | NONE | N/A |
| | testicular feminization: AR defect | CYP19A1; AR | 0.11 | AR | N/S | CYP19A1; AR | **0.01** | NONE | N/A |
| | testicular feminization | AR | N/S | AR | 0.13 | CYP19A1; ALDH1A1; AR | **2.E-03** | NONE | N/A |
| Female | polycystic ovary syndrome | CYP11A1; CYP17A1; CYP19A1; CYP21A2; ESR2; HSD3B2; AR; POR; PPARG; STAR | **3.E-07** | AR | N/S | CYP19A1; AR | N/S | NONE | N/A |
| | hyperandro-genism | CYP11A1; CYP17A1; CYP19A1; CYP21A2; HSD3B2; AR; PPARG | **4.E-05** | AR | N/S | CYP19A1 | N/S | NONE | N/A |
| | polycystic ovaries | CYP11A1; CYP17A1; CYP19A1; CYP21A2; AR; STAR | **4.E-04** | AR | N/S | CYP19A1; AR | N/S | NONE | N/A |
| | endometrial cancer | CYP1A1; CYP1A2; CYP17A1; CYP19A1; AHR; ESR1; ESR2; AR; NR1I2; EBAG9 | **7.E-04** | AR; NR1I2 | N/S | CYP1A1; CYP3A4; CYP19A1; AR; KCNH2 | N/S | ANGPT2 | N/S |
| | choriocarcinoma | CGB; CYP11A1; CYP19A1; AHR; ESR2; FDX1; POU1F1; PPARG; RARA; STAR | **7.E-04** | NONE | N/A | CYP19A1 | N/S | NONE | N/A |
| | endometriosis | CYP1A1; CYP11A1; CYP17A1; CYP19A1; AHR; ESR1; ESR2; STAR; EBAG9 | **1.E-03** | NONE | N/A | CYP1A1; CYP19A1 | N/S | ANGPT2; IL1B; IL8; TEK | 0.63 |
| | endometrial hyperplasia | CYP17A1; CYP19A1; ESR1; ESR2; AR; RARA | **1.E-03** | AR | N/S | CYP19A1; AR | N/S | NONE | N/A |
| | aromatase deficiency | CYP11A1; CYP19A1; CYP21A2; ESR1 | **3.E-03** | NONE | N/A | CYP19A1 | 0.79 | NONE | N/A |

| | | prochloraz | | procymidone | | tebuconazole | | mancozeb | |
|---|---|---|---|---|---|---|---|---|---|
| | disease | Genes | bonf. *p*-val | genes | bonf. *p*-val | genes | bonf. *p*-val | genes | bonf. *p*-val |
| **Female** | uterine leiomyoma | CYP17A1; CYP19A1; CYP21A2; ESR1; ESR2 | **0.03** | NONE | N/A | CYP19A1 | N/S | NONE | N/A |
| | uterine cancer | CYP19A1; ESR1; ESR2 | **0.04** | NONE | N/A | CYP19A1 | N/S | NONE | N/A |
| **Unisex** | virilization | CYP11A1; CYP11B1; CYP17A1; CYP19A1; CYP21A2; HSD3B2; AR; POR; SRD5A2 | **2.E-10** | AR | N/S | CYP19A1; AR | N/S | NONE | N/A |
| | kallmann's syndrome | CYP19A1; CYP21A2; ESR1; AR; POU1F1; PPARG; STAR; THRB | **4.E-05** | AR; THRB; TSHB | **0.03** | CYP19A1; AR | N/S | NONE | N/A |
| | PH | CYP17A1; CYP19A1; HSD3B2; SRD5A2 | **3.E-04** | NONE | N/A | CYP19A1 | N/S | NONE | N/A |
| | gonadotropin deficiency | CYP19A1; ESR1; AR; POU1F1; STAR | **2.E-03** | AR; TSHB | 0.17 | CYP19A1; AR | 0.47 | IL1B | N/S |
| | breast neoplasms | CYP19A1; ESR1; ESR2 | **7.E-03** | NONE | N/A | CYP19A1 | N/S | NONE | N/A |
| | sexual infantilism | CYP17A1; CYP19A1 | **0.01** | NONE | N/A | CYP19A1 | 0.40 | NONE | N/A |
| | abnormal sexual maturation | CYP17A1; CYP19A1; CYP21A2; ESR1 | **0.01** | NONE | N/A | CYP19A1 | N/S | NONE | N/A |
| | endocrine cancer | CYP19A1; ESR1; AR | 0.01 | AR | N/S | CYP19A1; AR | 0.23 | NONE | N/A |
| | gonadotropin unresponsiveness | AR | N/S | AR; TSHR | **2.E-03** | AR | N/S | NONE | N/A |

## 4 Discussion

Due to recent advances in toxicogenomics, the existence of high confidence PPIs, and various types of "omics" information, a reasonable basis for developing models to predict associations between chemical exposures and subsequent human health effects has arisen.

We retrieved a set of genes linked to the query chemicals and these chemical-gene associations, reported in experimental studies, constitute the raw data. This method constitutes an efficient means for retrieval of known information compared to manual search in the published literature for chemical-gene/protein associations. Such a manual literature search is feasible in less investigated areas but for well-investigated disorders and common chemicals obtaining such a literature overview manually is laborious. Thus, the current approach creates a valuable overview of existing data on the mechanisms of action of chemicals, allowing grouping of chemicals according to their mechanism of action. This will be helpful when predicting mixture effects. Furthermore, this approach provides a method that may contribute to developing candidate adverse outcome pathways. For chemicals that are known to cause similar adverse effects, applying this approach and thus setting up hypotheses on potential molecular initiating events for specific adverse effects can help elucidate similarities in mode of action of the chemicals (Ankley et al., 2010). However, since the data from the initial data sources may have been tested under complex conditions, there is a risk that the obtained gene lists contain false positives. CTD, as an example, contains data from microarray experiments and several biological factors are detrimental for gene expression levels. Hence, misregulation of genes is not always a direct response to chemicals affecting expression levels of genes but may be due to a stress condition of the exposed organism induced by the chemical, e.g., via disruption of the electron transport chain in mitochondria. The results of our study serve as a guideline, and careful interpretation is an essential part of the workflow as false positives might be present.

Furthermore, one must take into account the so-called "Matthew effect", resulting in maintained research interests regarding already well-investigated chemicals, effects, and/or targets, and thus a larger amount of available data for these chemicals and effects (Grandjean et al., 2011). This skews the findings towards pathways and diseases already being intensely investigated but also highlights the areas or chemicals needing further attention, like epoxiconazole in this study. The relatively high and low number of chemical-gene/protein associations for, e.g., prochloraz and epoxiconazole, respectively, may be a result of

this effect, as the amount of available data for the two chemicals differs (by a factor of 2.7 based on a PubMed search in July 2013). Hence, the lack of associations between, e.g., epoxiconazole or procymidone and the diseases predicted to be associated with prochloraz might be the result of less available data for the chemicals rather than lack of effects.

For risk assessment of mixtures, it has been proposed that being able to group the chemicals according to their mechanism or mode of action might be a great advantage (Kortenkamp et al., 2009). Previously, we found that mixture effects can be predicted on the basis of endpoints of varying molecular complexity, such as anatomical, morphological, or genomic output (Metzdorff et al., 2007). In a comprehensive *in vivo* developmental rat study, additivity of some anti-androgenic chemicals was predicted irrespective of the level at which the adverse effects were seen; be it adverse morphological changes or changes in gene expression in the male reproductive organs. This finding supports the hypothesis that molecular data can be applied for predicting mixture effects (Metzdorff et al., 2007). Viewing the chemical-gene associations for all chemicals simultaneously provides an overview of chemicals that share pharmacological or toxicological targets. This is valuable information in cases where grouping of chemicals is needed. According to our results, mancozeb has a mode of action that is distinct from that of prochloraz, procymidone, and tebuconazole. The latter three, on the other hand, seem to share targets. Therefore, rough grouping of the five chemicals results in two groups, mancozeb alone, and the above-mentioned three chemicals together. Epoxiconazole, having a single gene association, is only vaguely grouped with prochloraz, procymidone, and tebuconazole via its association with CYP51. This is most likely due to lack of data for this chemical.

The results presented in Figure 3 (and Fig. S2) indicate involvement of the chemicals in human diseases related to reproductive disorders. Prochloraz has the highest number of significant disease associations compared to the other chemicals. The disorders affected by this chemical have both masculinizing effects in females and feminizing effects in males. Several findings in the literature support the feminizing effects in males suggested here. Prochloraz is the most widely studied chemical and has been demonstrated to affect fetal rat testes, resulting in disrupted steroidogenesis during prenatal development (Laier et al., 2006). The mechanism underlying this disruption includes inhibition of the cytochrome P450 enzymes CYP17 and CYP19 (Andersen et al., 2002). Prochloraz has been demonstrated to be associated with decreased estradiol and testosterone production, increased progesterone production, and aromatase inhibition *in vitro* (Dreisig et al., 2013). In several instances it has been demonstrated that chemicals having such an *in vitro* profile cause feminizing effects, such as decreased nipple retention and decreased anogenital distance *in vivo*. This has been demonstrated for prochloraz exposure during fetal development (Vinggaard et al., 2005). The predicted involvement of prochloraz in TDS (cryptorchidism, hypospadias, poor sperm quality, and carcinoma *in situ* testis) is of particular concern. In the past four decades the incidence of male reproductive disorders has increased, and to a larger extent in developed and industrialized countries compared to developing countries (Giwercman et al., 1993; Skakkebaek et al., 2001; Jorgensen et al., 2006). It has been suggested that occupational exposure to pesticides contributes to the increased incidence of male reproductive disorders, including TDS (Weidner et al., 1998). Similarly, a recent paper suggested that prenatal exposure of human male fetuses to various non-persistent pesticides was associated with smaller genitals in boys aged 6-11 years compared to boys without prenatal pesticide exposure (Wohlfahrt-Veje et al., 2012). Furthermore, the results also point to an association between exposure to prochloraz and female reproductive diseases, including diseases of ODS. Previous studies have indicated the existence of a causal relationship between exposure to (organochlorine) pesticides and the development of endometriosis (Porpora et al., 2009; Cooney et al., 2010). Together, these findings point towards the importance of investigating female reproductive effects as well.

The findings of prochloraz on feminizing and masculinizing effects are paralleled by tebuconazole, although less pronounced. According to the results presented in Figure 3, tebuconazole has five significant disease associations in common with prochloraz. It is known that tebuconazole has a mode of action like prochloraz, as both affect sex hormone synthesis (Dreisig et al., 2013; Kjaerstad et al., 2010).

Procymidone has one significant disease association in common with prochloraz (Kallmann's syndrome, *p*-values 2.53e-2 and 4.21e-5, respectively). Despite the few significant associations, it is reasonable to infer a possible association with reproductive dysfunction, since procymidone acts as an antagonist of the androgen receptor. This is known from *in vitro* and *in vivo* experiments (Ostby et al., 1999; Hosokawa et al., 1993; Nellemann et al., 2003) that reported cryptorchidism (Wolf et al., 1999), decreased anogenital distance, and decreased nipple retention in prenatally exposed rats (Hass et al., 2012; Wolf et al., 1999).

Our data indicate that the effects of mancozeb are not a consequence of interference with human sex hormone synthesis or sex hormone receptor activity as opposed to the other chemicals investigated in this study. This is consistent with previous experimental work in rodents showing no effect of mancozeb on male reproductive endpoints (Axelstad et al., 2011). Mancozeb might, however, contribute to the manifestation of reproductive diseases through other mechanisms. It is established that mancozeb induces oxidative stress in rodent cells (Domico et al., 2007). Persistent oxidative stress might lead to inflammation, and this may be the reason why mancozeb, according to our model, is associated with several genes involved in inflammation. It is, however, surprising not to see any genes related to thyroid function, as animal studies in rodents indicate a strong relationship between exposure to mancozeb and decreased levels of the thyroid hormone $T_4$ in dams and hence a possible role of mancozeb in thyroid disruption in humans (Axelstad et al., 2011). Species differences are not expected to explain why our model does not predict a link between mancozeb and thyroid function (US EPA, 2005). Since epidemiological studies have

shown an association between human mancozeb exposure and diagnosed thyroid disease (Goldner et al., 2010; Steenland et al., 1997), the lack of associations between mancozeb and any thyroid-related genes in our study might be caused by a lack of thyroid tissue data in the queried databases. The epidemiological studies, however, do not provide a suggestion as to which proteins or genes are targeted by mancozeb. One alternative explanation might be that the thyroid effects are secondary to inflammatory changes.

Using this computational approach to assess the effect of chemicals provides an overview of which chemicals in a mixture might interfere with identical proteins, and which chemicals might result in similar disease phenotypes despite dissimilar modes of action. The method might prove useful if hypotheses on similar and dissimilar modes of action are needed to determine chemicals more prone to act on the same target and thus needing more attention. In addition, this methodology may prove valuable for generating hypotheses on the linkage between chemical exposure and human disease, providing knowledge on which chemicals to prioritize for further testing. This methodology provides a unique opportunity to get closer to the potential adverse effects on human health of chemicals.

## 5 Conclusion

We suggest that this computational technology is valuable for achieving an overview of existing targets related to a certain chemical. The value lies in the relevance for human health. Information about human health and exposure to environmental chemicals is rarely available, and exists as results from epidemiological studies, where several confounding factors might interfere with the measured endpoints. Therefore, this method opens up new possibilities for generation of hypotheses linking chemical exposure and human diseases.

Application of the method resulted in grouping of the three pesticides, prochloraz, tebuconazole, and procymidone. Male and female reproductive disorders like TDS and ODS and some adrenal diseases were primarily found to be associated with these pesticides. Epoxiconazole is expected to belong to this group, but lack of data did not allow this conclusion. Mancozeb had a differential mode of action involving inflammatory processes not shared by any of the other pesticides. These findings seem plausible based on comparisons with actual knowledge from cellular and rodent studies, giving confidence in the approach when less well-known chemicals are evaluated in the future.

## References

Andersen, H. R., Vinggaard, A. M., Rasmussen, T. H., et al. (2002). Effects of currently used pesticides in assays for estrogenicity, androgenicity, and aromatase activity in vitro. *Toxicol Appl Pharmacol 179*, 1-12.

Ankley, G. T., Bennett, R. S., Erickson, R. J., et al. (2010). Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. *Environ Toxicol Chem 29*, 730-741.

Audouze, K. and Grandjean, P. (2011). Application of computational systems biology to explore environmental toxicity hazards. Environ Health Perspect 119, 1754-1759.

Axelstad, M., Boberg, J., Nellemann, C., et al. (2011). Exposure to the widely used fungicide mancozeb causes thyroid hormone disruption in rat dams but no behavioral effects in the offspring. *Toxicol Sci 120*, 439-446.

Christiansen, S., Scholze, M., Axelstad, M., et al. (2008). Combined exposure to anti-androgens causes markedly increased frequencies of hypospadias in the rat. *Int J Androl 31*, 241-248.

Christiansen, S., Scholze, M., Dalgaard, M., et al. (2009). Synergistic disruption of external male sex organ development by a mixture of four antiandrogens. *Environ Health Perspect 117*, 1839-1846.

Cooney, M. A., Buck Louis, G. M., Hediger, M. L., et al. (2010). Organochlorine pesticides and endometriosis. *Reprod Toxicol 30*, 365-369.

Damgaard, I. N., Skakkebaek, N. E., Toppari, J., et al. (2006). Persistent pesticides in human breast milk and cryptorchidism. *Environ Health Perspect 114*, 1133-1138.

Davis, A. P., King, B. L., Mockus, S., et al. (2011). The Comparative Toxicogenomics Database: update 2011. *Nucleic Acids Res 39*, D1067-D1072.

Domico, L. M., Cooper, K. R., Bernard, L. P., et al. (2007). Reactive oxygen species generation by the ethylene-bis-dithiocarbamate (EBDC) fungicide mancozeb and its contribution to neuronal toxicity in mesencephalic cells. *Neurotoxicology 28*, 1079-1091.

Dreisig, K., Taxvig, C., Birkhoj Kjaerstad, M., et al. (2013). Predictive value of cell assays for developmental toxicity and embryotoxicity of conazole fungicides. *ALTEX 30*, 319-330.

EC – European Commision (2005). Regulation (EC) No 396/2005 of the European Parliament and of the Council of 23 February 2005 on maximum residue levels of pesticides in or on food and feed of plant and animal origin and amending Council Directive 91/414/EECText with EEA relevance. *Official J EU 396/2005*, 1-16.

EFSA – European Food Safety Authority (2013). International framework dealing with human risk assessment of combined exposure to multiple chemicals. *EFSA Journal 11*, 1-69.

Giwercman, A., Carlsen, E., Keiding, N., et al. (1993). Evidence for increasing incidence of abnormalities of the human testis: a review. *Environ Health Perspect 101, Suppl 2*, 65-71.

Goldner, W. S., Sandler, D. P., Yu, F., et al. (2010). Pesticide use and thyroid disease among women in the Agricultural Health Study. *Am J Epidemiol 171*, 455-464.

Grandjean, P., Eriksen, M. L., Ellegaard, O., et al. (2011). The Matthew effect in environmental science publication: a bibliometric analysis of chemical substances in journal articles. *Environ Health 10*, 96.

Hadrup, N., Taxvig, C., Pedersen, M., et al. (2013). Concentration addition, independent action and generalized concentration addition models for mixture effect prediction of sex hormone synthesis in vitro. *PLoS One 8*, e70490.

Hass, U., Scholze, M., Christiansen, S., et al. (2007). Combined exposure to anti-androgens exacerbates disruption of sexual differentiation in the rat. *Environ Health Perspect 115, Suppl 1*, 122-128.

Hass, U., Boberg, J., Christiansen, S., et al. (2012). Adverse effects on sexual development in rat offspring after low dose exposure to a mixture of endocrine disrupting pesticides. *Reprod Toxicol 34*, 261-274.

Hermens, J., Leeuwangh, P., and Musch, A. (1984). Quantitative structure-activity relationships and mixture toxicity studies of chloro- and alkylanilines at an acute lethal toxicity level to the guppy (Poecilia reticulata). *Ecotoxicol Environ Safety 8*, 388-394.

Hosokawa, S., Murakami, M., Ineyama, M., et al. (1993). The affinity of procymidone to androgen receptor in rats and mice. *J Toxicol Sci 18*, 83-93.

Jacobsen, P. R., Axelstad, M., Boberg, J., et al. (2012). Persistent developmental toxicity in rat offspring after low dose exposure to a mixture of endocrine disrupting pesticides. *Reprod Toxicol 34*, 237-250.

Jorgensen, N., Asklund, C., Carlsen, E., et al. (2006). Coordinated European investigations of semen quality: results from studies of Scandinavian young men is a matter of concern. *Int J Androl 29*, 54-61; discussion 105-108.

Kim Kjærulff, S., Wich, L., Kringelum, J., et al. (2013). ChemProt-2.0: visual navigation in a disease chemical biology database. *Nucleic Acids Res 41*, D464-469.

Kjaerstad, M. B., Taxvig, C., Andersen, H. R., et al. (2010). Mixture effects of endocrine disrupting compounds in vitro. *Int J Androl 33*, 425-433.

Knudsen, T., Martin, M., Chandler, K., et al. (2013). Predictive models and computational toxicology. *Methods Mol Biol 947*, 343-374.

Konemann, H. (1981). Fish toxicity tests with mixtures of more than two chemicals: a proposal for a quantitative approach and experimental results. *Toxicology 19*, 229-238.

Kortenkamp, A., Backhaus, T., and Faust, M. (2009). State of the Art Review of Mixture Toxicity. http://ec.europa.eu/environment/chemicals/effects/pdf/report_mixture_toxicity.pdf

Krysiak-Baltyn, K., Toppari, J., Skakkebaek, N. E., et al. (2010). Country-specific chemical signatures of persistent environmental compounds in breast milk. *Int J Androl 33*, 270-278.

Krysiak-Baltyn, K., Toppari, J., Skakkebaek, N. E., et al. (2012). Association between chemical pattern in breast milk and congenital cryptorchidism: modelling of complex human exposures. *Int J Androl 35*, 294-302.

Lage, K., Karlberg, E. O., Storling, Z. M., et al. (2007). A human phenome-interactome network of protein complexes implicated in genetic disorders. *Nat Biotechnol 25*, 309-316.

Lage, K., Hansen, N. T., Karlberg, E. O., et al. (2008). A large-scale analysis of tissue-specific pathology and gene expression of human disease genes and complexes. *Proc Natl Acad Sci U S A 105*, 20870-20875.

Lage, K., Mollgard, K., Greenway, S., et al. (2010). Dissecting spatio-temporal protein networks driving human heart development and related disorders. *Mol Syst Biol 6*, 381.

Laier, P., Metzdorff, S. B., Borch, J., et al. (2006). Mechanisms of action underlying the antiandrogenic effects of the fungicide prochloraz. *Toxicol Appl Pharmacol 213*, 160-171.

Louis, G. M. B., Cooney, M. A., and Peterson, C. M. (2011). The ovarian dysgenesis syndrome. *J Dev Orig Health Dis 2*, 25-35.

Matzuk, M. M. and Lamb, D. J. (2008). The biology of infertility: research advances and clinical challenges. *Nature Med 14*, 1197-1213.

Metzdorff, S. B., Dalgaard, M., Christiansen, S., et al. (2007). Dysgenesis and histological changes of genitals and perturbations of gene expression in male rats after in utero exposure to antiandrogen mixtures. *Toxicol Sci 98*, 87-98.

Nellemann, C., Dalgaard, M., Lam, H. R., et al. (2003). The combined effects of vinclozolin and procymidone do not deviate from expected additivity in vitro and in vivo. *Toxicol Sci 71*, 251-262.

Ostby, J., Kelce, W. R., Lambright, C., et al. (1999). The fungicide procymidone alters sexual differentiation in the male rat by acting as an androgen-receptor antagonist in vivo and in vitro. *Toxicol Ind Health 15*, 80-93.

Porpora, M. G., Medda, E., Abballe, A., et al. (2009). Endometriosis and organochlorinated environmental pollutants: a case-control study on Italian women of reproductive age. *Environ Health Perspect 117*, 1070-1075.

Rajapakse, N., Silva, E., and Kortenkamp, A. (2002). Combining xenoestrogens at levels below individual No-observed-effect concentrations dramatically enhances steroid hormone action. *Environ Health Perspect 110*, 917-921.

Safran, M., Dalah, I., Alexander, J., et al. (2010). GeneCards Version 3: the human gene integrator. *Database 2010*, baq020.

Silva, E., Rajapakse, N., and Kortenkamp, A. (2002). Something from "nothing" – eight weak estrogenic chemicals combined at concentrations below NOECs produce significant mixture effects. *Environ Sci Technol 36*, 1751-1756.

Skakkebaek, N. E., Rajpert-De Meyts, E., and Main, K. M. (2001). Testicular dysgenesis syndrome: an increasingly common developmental disorder with environmental aspects. *Hum Reprod 16*, 972-978.

Steenland, K., Cedillo, L., Tucker, J., et al. (1997). Thyroid hormones and cytogenetic outcomes in backpack sprayers using ethylenebis(dithiocarbamate) (EBDC) fungicides in Mexico. *Environ Health Perspect 105*, 1126-1130.

Swan, S. H., Kruse, R. L., Liu, F., et al. (2003). Semen quality in relation to biomarkers of pesticide exposure. *Environ Health Perspect 111*, 1478-1484.

Swan, S. H., Main, K. M., Liu, F., et al. (2005). Decrease in anogenital distance among male infants with prenatal phthalate exposure. *Environ Health Perspect 113*, 1056-1061.

US EPA (2005). EPA Mancozeb Facts. Prevention Pesticides

and Toxic Substances. http://www.epa.gov/oppsrrd1/REDs/factsheets/mancozeb_fact.pdf
Vinggaard, A. M., Christiansen, S., Laier, P., et al. (2005). Perinatal exposure to the fungicide prochloraz feminizes the male rat offspring. *Toxicol Sci 85*, 886-897.
Weidner, I. S., Moller, H., Jensen, T. K., et al. (1998). Cryptorchidism and hypospadias in sons of gardeners and farmers. *Environ Health Perspect 106*, 793-796.
Wohlfahrt-Veje, C., Boisen, K. A., Boas, M., et al. (2009). Acquired cryptorchidism is frequent in infancy and childhood. *Int J Androl 32*, 423-428.
Wohlfahrt-Veje, C., Andersen, H. R., Jensen, T. K., et al. (2012). Smaller genitals at school age in boys whose mothers were exposed to non-persistent pesticides in early pregnancy. *Int J Androl 35*, 265-272.
Wolf, C., Jr., Lambright, C., Mann, P., et al. (1999). Administration of potentially antiandrogenic pesticides (procymidone, linuron, iprodione, chlozolinate, p,p'-DDE, and ketoconazole) and toxic substances (dibutyl- and diethylhexyl phthalate, PCB 169, and ethane dimethane sulphonate) during sexual differentiation produces diverse profiles of reproductive malformations in the male rat. *Toxicol Ind Health 15*, 94-118.

**Acknowledgements**
The Ministry of Food, Agriculture and Fisheries of Denmark and the Danish Veterinary and Food Administration are acknowledged for their financial support. The Villum Kann Rasmussen Foundation contributed financially to development of the systems toxicology approach.

**Financial interests declaration**
The authors declare they have no actual or potential competing financial interests.

**Correspondence address**
Kristine Kongsbak
Division of Toxicology and Risk Assessment,
National Food Institute,
Technical University of Denmark
Mørkhøj Bygade 19
Building G, Room 118
2860 Søborg, Denmark
e-mail: kgko@food.dtu.dk

Kongsbak et al.:

# A Computational Approach to Mechanistic and Predictive Toxicology of Pesticides

## Supplementary Data

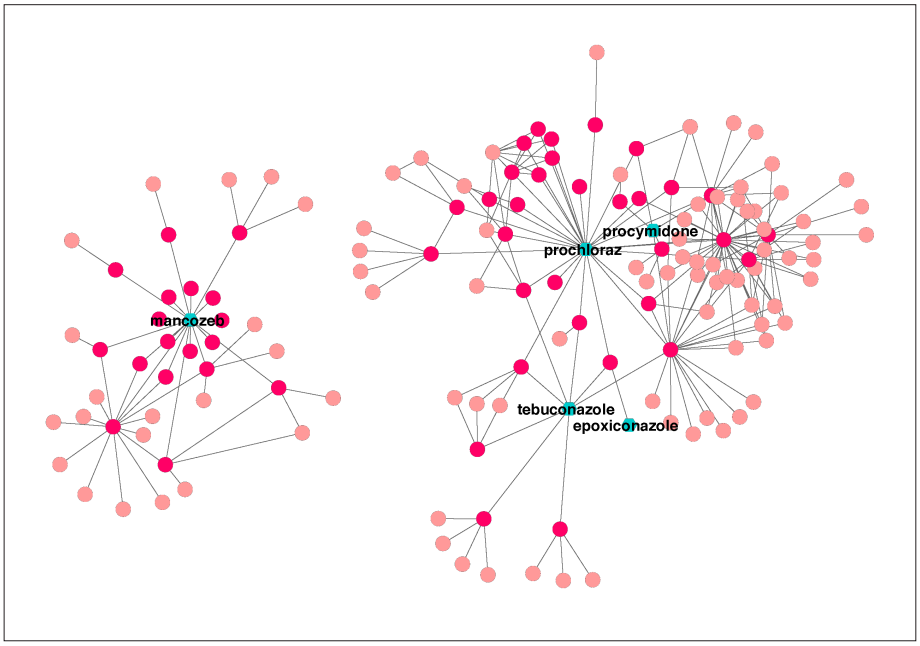



**Fig. S1: Network view of the protein-protein interaction (PPI) network of the five chemicals**
Chemicals are represented by blue hexagons, genes/proteins from the raw network – based on experimental data – are the dark pink circles, and genes/proteins retrieved via known PPIs are the pink circles.

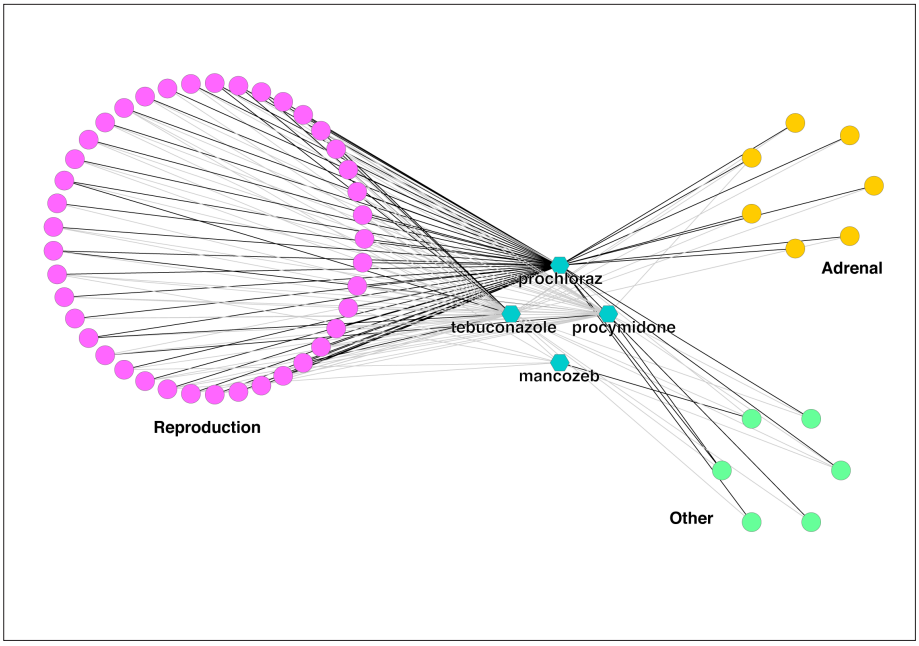



**Fig. S2: Network showing the dominating groups of diseases associated with the four chemicals**
The groups are color coded: reproductive diseases, magenta; adrenal diseases, yellow; other diseases, green. The four chemicals are represented by the blue hexagons. Black lines indicate significant ($p$-value ≤0.05) associations after Bonferroni correction of $p$-values.

## 4.5 Concluding remarks

As concluded in the papers (Sections 4.3 and 4.4), the approach applied in this project provides a means to, relatively quickly, obtaining an overview of existing data regarding a chemical of interest. The approach further enables generation of hypotheses on the mode of action or potential harmful effects of the investigated chemical. This is particularly relevant for environmental chemicals, where testing in humans are otherwise unethical.

For the investigated chemicals, the analysis resulted in a grouping of the chemicals with the azole-fungicides (tebuconazole and prochloraz) and procymidone in one group and mancozeb alone, as it had no shared targets with the other pesticides. Grouping of epoxiconazole was vague as very little data existed in the queried databases for this pesticide. For the other chemicals, the diseases with the strongest associations to the azole-pesticides were related to reproductive diseases. This correlates well with results from animal studies with these chemicals.

Despite the above-mentioned benefits of this approach, there is one obvious short-coming; that it is highly reliant on existing data. Therefore, this approach is not applicable in cases where little to no information on a chemical is available. In our project, this was evident with epoxiconazole, for which only one human gene-target was retrieved in the initial step. Subsequent analyses were not possible for this chemical despite the fact that data from animal studies indicate epoxiconazole to have a similar mode of action as the other azole-fungicides (tebuconazole and prochloraz) and procymidone. The fact that little to no information is available for epoxiconazole might indicate either that the chemical has no effects on human targets, or that the chemical has not been investigated exhaustively. The latter possibility might be taken as an indication that further studies on epoxiconazole are needed. One has to bear in mind, that positive findings spurs interest in investigating those findings further. This might skew the data availability towards pathways and diseases already being intensely investigated (the so-called Matthew effect). Therefore, this approach might also prove useful in the planning phase of a study in the process of determining which chemical(s) to investigate.

An additional short-coming of the approach is that the nature and direction of the initially retrieved chemical-gene/protein associations are not taken into account, as agonists and antagonists are not distinguished, and neither is up- and down-regulation. This might introduce bias to the downstream analysis.

Chapter 5

# Experimental Studies of a Chemical Mixture

This chapter provides a summary of the hypotheses and methods used in the studies with a 14-compound mixture (Mix) in combination with three doses of perfluorononanoic acid (PFNA). Two manuscript have resulted from this project. Both have been submitted but not yet accepted for publication. The first manuscript (page 52) describes the endocrine-related effects related to the interactions between the EDC, PFNA, and Mix, whereas the second manuscript is a profiling study, where changes in the plasma metabolome and the liver transcriptome upon Mix and PFNA exposure have been analyzed and described (page 68). A brief summary of the results and concluding remarks will follow on page 103.

## 5.1 Background

As described in Section 2.2, humans are simultaneously exposed to multiple chemicals. Previous studies have indicated combination effects of environmental chemicals at estimated safe levels [81, 82], environmental background [83], or human acceptable daily intake levels [84]. For example, Naville et al. [82] present data indicating significantly altered hepatic metabolism following exposure to a mixture

Table 5.1: The 14 chemicals composing the Mix, which were tested along with three doses of PFNA in adolescent rats for 14 days.

| CASRN | Chemical name | Source/use | Ratio (weight) | Rat dose mg/kg BW/day |
|---|---|---|---|---|
| 7380-40-7 | bergamottin | grapefruit constituent | 0.08 | 0.2 |
| 59870-68-7 | glabridin | liquorice constituent | 0.12 | 0.3 |
| 80-05-7 | bisphenol A | plastic additive | 0.004 | 0.01 |
| 94-26-8 | butyl paraben | preservative | 0.21 | 0.52 |
| 84-74-2 | DBP | plasticizer | 0.02 | 0.06 |
| 117-81-7 | DEHP | plasticizer | 0.03 | 0.09 |
| 36861-47-9 | 4-MBC | sun filter | 0.15 | 0.38 |
| 5466-77-3 | OMC | sun filter | 0.27 | 0.68 |
| 72-55-9 | *p,p*'-DDE | pesticide | 0.002 | 0.006 |
| 133855-98-8 | epoxiconazole | pesticide | 0.02 | 0.05 |
| 330-55-2 | linuron | pesticide | 0.002 | 0.004 |
| 67747-09-5 | prochloraz | pesticide | 0.025 | 0.06 |
| 32809-16-8 | procymidone | pesticide | 0.035 | 0.09 |
| 50471-44-8 | vinclozolin | pesticide | 0.021 | 0.05 |
| In total | 14 chemicals | | 1.0 | 2.5 |

Abbreviations: DBP - dibutylphthalate, DEHP - bis(2-ethylhexyl)phthalate, 4-MBC - 4-methylbenzylidene camphor, OMC - 2-ethylhexyl-2-methoxycinnamate, p,p'-DDE - dichlorodiphenyldichloroethylene.

of four persistent organic pollutants at doses grossly corresponding to the tolerated daily intake levels promulgated by regulatory agencies.

In this project, we investigated the effects on adolescent rats of Mix composed of 12 environmental chemicals previously shown to affect fetal development in rats [85], two food constituents known to affect cytochrome P450 (CYP) activity (Table 5.1), and three doses of the recently classified EDC, PFNA. The 12 previously investigated chemicals indicated endocrine disruption measured as e.g. increased nipple retention in exposed pubs, an androgen-sensitive endpoint [85]. However, the effects were observed at a dose approximately 24 times higher than the dose of the Mix in the current study, since their lowest dose was 150 times the human high-end exposure levels and our dose of the Mix corresponded to the human high-end exposure levels corrected for different body surface areas of human and rat.

Perfluorinated compounds (PFCs) are organic compounds in which all hydrogen atoms of the carbon chain have been replaced by fluorine atoms (Figure 5.1) [86].[1]

[1] 'Perfluorinated compound' covers perfluorinated alkyl compounds of varying carbon chain length and with varying functional groups.

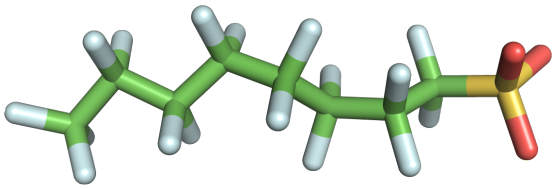


Figure 5.1: Three-dimensional structure of the eight-carbon perfluorinated sulphonate, PFOS. The figure illustrates the conformation of the compound when bound to serum albumin. Source: PDB ID 4E99 [103].

Due to their physical-chemical properties, PFCs are widely used as surfactants [87] and are often formulated along with proteins in bioscience, cosmetics applications, and medicine [86].

PFNA is a nine-carbon PFC used as a surfactant and emulsifier [88]. It is among the four most frequently detected polyfluorinated compounds in human serum[2] [89]. Furthermore, over a time-course of 11 years (1999-2010), a three-fold increase in serum PFNA concentrations in the US population has been observed [89–91]. Currently, global efforts are made to monitor human exposure to per- and polyfluorinated compounds in various subpopulations including pregnant [92], breastfeeding [93], and other women [94, 95], and children [96–99], elderly [100], and various other age groups [89–91, 101, 102].

Polyfluorinated compounds are known to compete with naturally occurring fatty acids for binding to the hepatic fatty acid-binding proteins [104–106]; however, due to their physical-chemical properties, they do not participate as substrates in biochemical reactions requiring fatty acids. Due to their surfactant effects, PFCs have been shown to alter the mitochondrial and cell membrane potentials, and to change the membrane fluidity and function [107–109]. Other studies have shown that PFCs inhibit gap junctional intercellular communication [110–112]. Evidence exist that longer-chained PFCs are more toxic than shorter-chained PFCs [107, 108, 110, 113].

Activation of the peroxisome proliferator-activated receptors (PPAR) is a hallmark PFC effect. Therefore, PFCs are currently under investigation for their potential involvement in the activation/functioning of other nuclear receptors, such as the constitutive androstane receptor (CAR) and the pregnane X receptor (PXR) involved in regulating xenobiotic metabolism, and the liver X receptor (LXR) involved in regulating cholesterol and lipid metabolism [88].

[2]The other most frequently detected PFCs are perfluorooctanoic acid (PFOA), perfluorooctanesulphonic acid (PFOS), and perfluorohexanoic acid (PFHxS).

PFCs are well-known initiators of hepatocytic hypertrophy in laboratory animals [88]. This effect is likely related to peroxisome proliferation and PPARα activation. Prolonged exposure to PFCs leads to hepatic accumulation of lipid droplets, an effect likely resulting from altered lipid metabolism and transport. In turn, this might result in hepatic steatosis - a state of retention of lipids in the liver.

Recently, Rosenmai et al. [114] demonstrated the effects of PFOA (and other polyfluorinated compounds) on the steroid synthesis *in vitro*. Consistently, previous *in vivo* studies have indicated potentials of PFCs to decrease plasma testosterone concentrations in male rats and affects male reproductive organs [115–117].

We conducted an animal study in which adolescent male rats were treated with Mix ± PFNA for 14 days. In the following sections I will present the two papers resulting from the animal study, and the papers will follow in Section 5.4 and 5.5. Specific aims and hypotheses for the two sub-projects will follow below (Section 5.2). Concluding remarks covering both projects will follow at page 103.

## 5.2 Aims and Hypotheses

### 5.2.1 Endocrine Disrupting Effects

This EDC sub-project was designed as a hypothesis-driven study. The aim was to test the hypothesis that a mixture of food chemicals in doses approaching high-end human exposure levels cause adverse mixture effects. The plasma levels of steroid hormones were used as effect measures along with pituitary hormone levels, classical toxicological parameters and mechanistic markers in testis, liver, adipose tissue, and kidney.

As my main role in the mixture project was to perform the microarray analyses, my primary responsibilities in this sub-project included selection of genes for quantification using RT-qPCR, interpretation of data, and critical editing of the manuscript prior to submission.

### 5.2.2 Integrative 'OMICs' analysis

This study builds on top of the before-mentioned EDC study. However, where the EDC study was hypothesis driven, this study had a character as a data-driven project. With this study, we aimed at profiling the plasma metabolome and the liver transcriptome following the chemical exposure described in section 5.3. As is the nature of profiling studies there is no specific hypothesis [40]; however, we worked

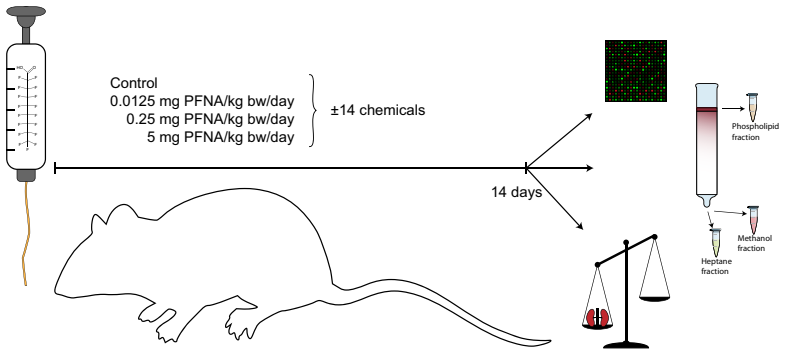



Figure 5.2: Experimental design of the animal study. 6 weeks old male rats were treated orally with three doses of PFNA (6-8 animals/group) ± a Mix via gavage for 14 days. After euthanization organs (liver, kidney, testes) were collected for weighing, livers were prepared for histology and microarray analyses, plasma was collected for metabolomics analyses.

from the assumption that two complementary 'omics' approaches (metabolomics and transcriptomics) would provide additional insight to the mode of action of the Mix and PFNA.

As mentioned above, I was responsible for conducting the transcriptomics analysis. In this sub-project, MSc. Kasper Skov and I contributed equally to interpreting the results from the metabolomics and transcriptomics analyses and to writing the manuscript.

## 5.3 Methods

Detailed descriptions about the applied methods are provided in the respective manuscripts (page 52-68). In brief, six weeks old male Wistar rats arrived at our facilities one week prior to study initiation to allow for acclimatization to the new environment. As outlined in Figure 5.2 the experimental animals were treated with chemical(s) (Table 5.1 ± three doses of PFNA[3]) for 14 days prior to euthanization.

Upon study completion, neck blood was collected and plasma was isolated for hormone and metabolite analyses. Body weight and organ weights (liver, kidney,

[3]0.0125 (Low PFNA), 0.25 (Mid PFNA) and 5 (High PFNA) mg/kg body weight/day (mg/kg BW/day)

testes) were recorded and subsequently stored in RNAlater until mRNA measurements. Parts of the livers were fixed and processed for histopathological examination; whole kidneys were stored frozen until measurements of levels of organic anion transporters.

### 5.3.1 Metabolomics

Plasma from all treatment groups were analyzed using a recently established platform for the analysis of metabolites in plasma [118]. In brief, phospholipids, lipophilic, and polar metabolites were separated using a phospholipid solid-phase extraction column[4]. The lipid and polar fractions were extracted from the eluate with heptane and methanol, respectively. Lastly, phospholipids were eluted from the column using $NH_4OH$ in methanol. The fractions were subsequently analyzed in an HPLC system combined with a maxis qTime-of-flight mass spectrometer. Detailed information about feature extraction and identification is given in Manuscript IV.

### 5.3.2 Transcriptomics

We extracted the hepatic mRNA from the animals from 4 treatment groups; control, Low PFNA + Mix, Mid PFNA + Mix, and Mid PFNA. Labeled cDNA was applied to Agilent Whole Rat genome Oligo Microarrays according to the manufacturer's protocol. Microarrays were scanned and features extracted according to the manufacturer's protocol. Extracted data were loaded into R [119], and data were analyzed using the limma software package [50, 120]. I performed pathway analyses of the differentially expressed transcripts using QIAGEN's Ingenuity® Pathway Analysis (IPA®, QIAGEN Redwood City[5]) More detailed information about the microarray analyses are provided in section 3.1.

[4] Phospholipids adsorb to the column leaving other metabolites in the eluate.
[5] `www.qiagen.com/ingenuity`

Arch Toxicol
DOI 10.1007/s00204-015-1452-6

ORGAN TOXICITY AND MECHANISMS

# Perfluorononanoic acid in combination with 14 chemicals exerts low-dose mixture effects in rats

**Niels Hadrup · Mikael Pedersen · Kasper Skov · Niels Lund Hansen · Line Olrik Berthelsen · Kristine Kongsbak · Julie Boberg · Marianne Dybdahl · Ulla Hass · Henrik Frandsen · Anne Marie Vinggaard**

Received: 3 October 2014 / Accepted: 6 January 2015


**Abstract** Humans are simultaneously exposed to several chemicals that act jointly to induce mixture effects. At doses close to or higher than no-observed adverse effect levels, chemicals usually act additively in experimental studies. However, we are lacking knowledge on the importance of exposure to complex real-world mixtures at more relevant human exposure levels. We hypothesised that adverse mixture effects occur at doses approaching high-end human exposure levels. A mixture (Mix) of 14 chemicals at a combined dose of 2.5 mg/kg bw/day was tested in combination with perfluorononanoic acid (PFNA) at doses of 0.0125 (Low PFNA), 0.25 (Mid PFNA) and 5 (High PFNA) mg/kg bw/day by oral administration for 14 days in juvenile male rats. Indication of a toxicokinetic interaction was found, as simultaneous exposure to PFNA and the Mix caused a 2.8-fold increase in plasma PFNA concentrations at Low PFNA. An increase in testosterone and dihydrotestosterone plasma concentrations was observed for Low PFNA + Mix. This effect was considered non-monotonic, as higher doses did not cause this effect. Reduced LH plasma concentrations together with increased androgen concentrations indicate a disturbed pituitary-testis axis caused by the 15-chemical mixture. Low PFNA by itself increased the corticosterone plasma concentration, an effect which was normalised after simultaneous exposure to Mix. This combined with affected ACTH plasma concentrations and down-regulation of 11β HSD mRNA in livers indicates a disturbed pituitary-adrenal axis. In conclusion, our data suggest that mixtures of environmental chemicals at doses approaching high-end human exposure levels can cause a hormonal imbalance and disturb steroid hormones and their regulation. These effects may be non-monotonic and were observed at low doses. Whether this reflects a more general phenomenon that should be taken into consideration when predicting human mixture effects or represents a rarer phenomenon remains to be shown.

**Keywords** Mixture toxicology · Steroidogenesis · Testosterone · Corticosterone · Pituitary hormones · Perfluorononanoic acid (PFNA)

N. Hadrup · N. L. Hansen · L. O. Berthelsen · K. Kongsbak · J. Boberg · M. Dybdahl · U. Hass · A. M. Vinggaard (✉)
Division of Toxicology and Risk Assessment, National Food Institute, Technical University of Denmark, Mørkhøj Bygade 19, 2860 Søborg, Denmark
e-mail: annv@food.dtu.dk

M. Pedersen · K. Skov · H. Frandsen
Division of Food Chemistry, National Food Institute, Technical University of Denmark, Søborg, Denmark

## Introduction

Humans are concomitantly exposed to several chemicals that can exert mixture effects. At doses close to no-observed adverse effect levels or higher, chemicals usually act additively in experimental studies (Kortenkamp 2014). However, we are lacking knowledge on the importance of exposure to complex real-world mixtures at more relevant human exposure levels. Only few in vivo studies performed with low-human relevant exposure levels have been reported. Wade et al. investigated a mixture containing 18 persistent contaminants at a dose of 1x the estimated safe level. This dose was based on the minimum risk levels (MRLs) or the tolerable daily intakes and was dosed orally to male rats for 70 days. At this dose, increased natural killer cell lytic activity was observed (Wade et al. 2002). Crofton et al. investigated whether deviations from additivity could be detected at low doses. They studied

Published online: 15 January 2015

thyroid disruptors in doses ranging from environmental background to 100x environmental background in rats for effects on serum total thyroxine. No deviation from additivity was found at the lowest doses (<0.2 mg/kg bw/day), but at higher doses (>0.67 mg/kg bw/day), synergy was found (Crofton et al. 2005). Stanko et al. found an increase in prostate inflammation with 0.09 mg/kg bw of an atrazine metabolite mixture following prenatal exposure of male rats (Stanko et al. 2010). Merhi et al. exposed mice to a mixture of six pesticides at levels derived from the human acceptable daily intake levels. Changes in blood cell counts were observed (Merhi et al. 2010). Demur et al. (2013) found dietary exposure to a low atrazine, endosulfan and chlorpyrifos mixture dose (25, 30 and 50 μg/kg food corresponding to approximately 4, 5, and 8 μg/bw/day) to decrease the red blood cell count and haemoglobin levels.

To further address exposure to human relevant mixtures of chemicals, we designed an experiment to test whether chemicals in food approaching high-end human exposure cause adverse mixture effects. We hypothesised that effects could occur via toxicokinetic metabolism effects comparable to those described for food–drug interactions, e.g., as has been described for grapefruit juice and $Ca^{2+}$ channel blockers (Bailey et al. 1989). For our investigation, we used steroid hormone metabolism as an effect measure. Steroid hormones are involved in multiple important male developmental processes and thus alterations in its plasma concentration may potentially be accompanied by adverse effects. We designed an experiment in which a single chemical known to affect the steroid hormone testosterone on its own was tested at increasing doses in the presence or absence of a fixed dose of a 'background' chemical mixture (Mix). On basis of its human toxicological importance and on basis of its ability to increase testosterone at a dose of 1 mg/kg bw/day to rats (Feng et al. 2009), we selected the fluorosurfactant perfluorononanoic acid (PFNA) as the variably dosed chemical in this investigation. We exposed the rats to a fixed background dose of 14 chemicals—12 environmental chemicals representing typical endocrine disrupting chemicals and two food ingredients. For the design of this Mix, we took advantage of the fact that testosterone is metabolized in the liver by cytochrome P450 3A4 (CYP3A4 homologous to CYP3A23/3A1 in rat) and CYP2C9 (homologous to CYP2C11 in rat) (Cheng and Schenkman 1983; Guengerich 1999; Kenworthy et al. 1999; Martignoni et al. 2006). CYP3A4 and CYP2C9 are inhibited by the grapefruit constituent bergamottin and the liquorice constituent glabridin (Bailey et al. 2003; Foti and Wahlstrom 2008; Kent et al. 2002; Lim et al. 2005; Tassaneeyakul et al. 2000; Uesawa and Mohri 2006; Wen et al. 2002). These substances were included in the Mix along with twelve food contaminants (described in Table 2), each reported to exert endocrine disrupting effects (Christiansen et al. 2012). The doses of these twelve contaminants were selected to reflect a high-end exposure level to the European human population. Thus, to test whether chemicals in food at human relevant doses caused adverse mixture effects, we administered PFNA at doses 0.0125 (Low PFNA), 0.25 (Mid PFNA) and 5 (High PFNA) mg/kg bw/day in the presence or absence of Mix at a total dose of 2.5 mg/kg bw/day (Table 2) to male rats for 14 days. Steroid hormone plasma concentrations, pituitary hormone levels as well as classical toxicological parameters and mechanistic markers in testis, liver, adipose tissue and kidney were investigated.

## Methods

### Dose selection and chemicals

We aimed at selecting a PFNA dose to cover human internal exposure of the combined exposure to perfluorinated compounds such as PFOA, PFOS and PFNA. Lau et al. (2007) gathered human exposure data from a large number of studies. The PFOS mean plasma concentrations in single studies were 10–73 ng/mL (see Table 1 for an overview of the data in table format), for PFOA the values were 2.1–354 ng/mL (second highest value was 41 ng/mL), and for PFNA 0.8–2.2 ng/mL. This gives a possibility of a combined human exposure to these three perfluorinated compounds of 13–429 ng/mL (or 13–117 ng/mL if only the second highest PFOA value is used) (Lau et al. 2007). For comparison, the high-end exposure (95th percentile) in the American population as determined by the National Health and Nutrition Examination Survey (NHANES) is 67.6 ng/mL for PFNA, PFOA and PFOS combined (CDC 2009). The PFNA dose employed in the present investigation was chosen on basis of a study by Tatum Gibbs et al. (Tatum-Gibbs et al. 2011) in which an oral dose of 1 mg/kg bw to rats gave a serum concentration of approximately 10,000 ng/mL. From that we chose a dose of 0.0125 mg/kg bw/day for a 14-day investigation (Abbreviated: Low PFNA), predicted giving rise to a PFNA plasma concentration of 125 ng/mL. For a medium- and high-dose level, we included 0.25 (Mid PFNA) and 5 (High PFNA) mg/kg bw/day.

We aimed to administer bergamottin in an amount corresponding to the content of one grapefruit. A grapefruit contains 25 μM bergamottin (Bailey et al. 2003) corresponding to 1,700 μg bergamottin per grapefruit. Adjusting according to body surface area as described (Reagan-Shaw et al. 2008), a rat dose corresponding to a human intake of 100 mL grapefruit juice was calculated to be 0.2 mg/kg bw/day. For glabridin, an amount corresponding to an intake of 100-g liquorice candy per day was determined in the following way. The extraction of glycyrrhizic acid and glabridin from Chinese (raw) liquorice gives 2.39 mg/g

**Table 1** Comparison of reported human perfluorinated compound levels with measured levels in rats

| | Reported human exposure levels to PFCs (ng/mL) | | |
|---|---|---|---|
| | NHANES (geometric means) (CDC 2009) | NHANES 95th percentile (CDC 2009) | Mean value ranges from literature reviewed in Lau et al. (2007) |
| PFNA in serum | 1.0 | 3.2 | 0.8–2.2 ($n = 6$) |
| PFOA in serum | 4.0 | 9.8 | 2.1–354 ($n = 9$) |
| PFOS in serum | 20.7 | 54.6 | 10–73.2 ($n = 11$) |
| Total PFC in serum | 25.7 | 67.6 | 32.8–429 |

| | Measured rat exposure levels (ng/mL) | Comparison human versus rat |
|---|---|---|
| Measured PFNA plasma conc. in rats at low dose | 396 | 0.9- to 15-fold human exposure to total PFC<br>Sixfold human 95th percentile NHANES exposure |
| Measured PFNA plasma conc. in rats at mid dose | 29,950 | 70- to 1,200-fold human exposure to total PFC<br>440-fold human 95th percentile NHANES exposure |

*NHANES* National Health and Nutrition Examination Survey, *PFC* perfluorinated compound

glycyrrhizic acid and 0.92 mg/g glabridin (Tian et al. 2008). This taken together with a reported 150 mg glycyrrhizic acid in 100 g of sweet liquorice (candy) (Sigurjonsdottir et al. 2001), suggests an amount of 58 mg glabridin in 100 g of sweet liquorice, suggesting that a 70-kg person takes in 0.8 mg/kg bw/day of glabridin when ingesting 100 g of sweet liquorice. This suggests—by a body surface area conversion (Reagan-Shaw et al. 2008)—that we should give a dose of 4 mg/kg bw/day to rats. However, Furhman et al. previously reported that a glabridin dose to rats corresponding to the selected human intake was 0.3 mg/kg bw/day (Fuhrman et al. 1997) and to take a conservative approach we settled for this dose. Regarding the remaining chemicals in the mixture, a previously designed mixture based on exposure levels of the European human population was used (Christiansen et al. 2012; Hadrup et al. 2013). The dose reported to be a realistic 'high-end human exposure level' (Christiansen et al. 2012) was 0.32 mg/kg bw/day. By conversion to rat dose via body surface area, this yields 2 mg/kg bw/day. The chemicals and their ratio in Mix are described in Table 2. Chemicals were purchased as follows: Bergamottin, glabridin, bisphenol A, butyl paraben and 4-methylbenzylidene camphor (4-MBC) were purchased from Sigma-Aldrich, Brøndby, Denmark. Dibutylphthalate (DBP), bis(2-ethylhexyl)phthalate (DEHP), 4-MBC, 2-ethylhexyl-4-methoxycinnamate (OMC), dichlorodiphenyldichloroethylene (DDE), epoxiconazole, linuron, prochloraz, procymidone and vinclozolin were purchased from VWR, Bie & Berntsen, Herlev, Denmark.

## Animals and pathology

Male Wistar Hannover Galas rats, 6 weeks of age with specific pathogen-free health status, were obtained from Taconic M&B (Lille Skensved, Denmark) and allowed to acclimatise for 1 week. The animals were housed two per cage (Macrolon, Buguggiate, Italy) with light on from 7 a.m. to 7 p.m. Room temperature and relative humidity were 22 ± 1 °C and 55 ± 5 %, respectively. Rats were given ad libitum access to citric acid acidified tap water and standard diet (prod. no. 1324 Altromin, Brogården, Gentofte, Denmark). The animals were administered test substances once a day orally by gavage for 14 days with corn oil (VWR—Bie & Berntsen, Herlev, Denmark) as vehicle. The dosing volume was 1 mL/100 g of body weight (bw). In total, 70 male rats were randomly placed into eight groups, i.e., vehicle control ($n = 10$), PFNA 0.0125 mg/kg/day (Low PFNA) ($n = 10$), PFNA 0.25 mg/kg/day (Mid PFNA) (n = 8), PFNA 5 mg/kg/day (High PFNA) ($n = 8$), Mix + PFNA 0.0125 mg/kg/day (Low PFNA + Mix) ($n = 10$), Mix + PFNA 0.25 mg/kg/day (Mid PFNA + Mix) ($n = 8$), Mix + PFNA 5 mg/kg/day (High PFNA + Mix) ($n = 8$), Mix ($n = 8$). Animals were evaluated clinically and subjected to necropsy in four different sets each starting 1 day after the prior. Each set comprised animals of each dosing group. For the euthanisation, the animals were anaesthetised in $CO_2/O_2$ and decapitated. Neck blood was collected in heparinised tubes, and plasma was isolated by centrifugation at 1,000×*g*, 4 °C for 10 min. Plasma was stored at −80 °C. To avoid bias, e.g., due to stress in animals, the sectioning of animals were randomised according to groups. Animals were given the last dose in the time span of 1 h and 15 min to 1 h and 45 min before euthanisation. Body weight and organ weights (liver, kidney, testes) were recorded, and livers were fixed and processed for histopathological examination as previously described (Hadrup et al. 2012). Liver haematoxylin and eosin stained sections were evaluated by a pathologist blinded to treatment groups. Changes were described qualitatively, and in addition, selected parameters were scored in the following manner: Cell borders (score

**Table 2** Fourteen chemicals in fixed ratio mixture

| CAS registry number | Chemical name | Source/use | Ratio in mixture (weight) | Rat dose (mg/kg bw/day) |
|---|---|---|---|---|
| 7380-40-7 | Bergamottin | Grapefruit constituent | 0.08 | 0.2 |
| 59870-68-7 | Glabridin | Liquorice constituent | 0.12 | 0.3 |
| 80-05-7 | Bisphenol A | Plastic additive | 0.004 | 0.01 |
| 94-26-8 | Butyl paraben | Preservative | 0.21 | 0.52 |
| 84-74-2 | Dibutylphthalate (DBP) | Plasticiser | 0.02 | 0.06 |
| 117-81-7 | Bis(2-ethylhexyl)phthalate (DEHP) | Plasticiser | 0.03 | 0.09 |
| 36861-47-9 | 4-Methylbenzylidene camphor (4-MBC) | Sun filter | 0.15 | 0.38 |
| 5466-77-3 | 2-Ethylhexyl-4-methoxycinnamate (OMC) | Sun filter | 0.27 | 0.68 |
| 72-55-9 | Dichlorodiphenyldichloroethylene (p,p′-DDE) | Pesticide | 0.002 | 0.006 |
| 133855-98-8 | Epoxiconazole | Pesticide | 0.02 | 0.05 |
| 330-55-2 | Linuron | Pesticide | 0.002 | 0.004 |
| 67747-09-5 | Prochloraz | Pesticide | 0.025 | 0.06 |
| 32809-16-8 | Procymidone | Pesticide | 0.035 | 0.09 |
| 50471-44-8 | Vinclozolin | Pesticide | 0.021 | 0.05 |
| In total | 14 Chemicals | | 1.0 | 2.5 |

0 = not visible, score 1 = not clear, score 2 = clear) and cell size (score 1 = small/normal, score 2 = slight increase, score 3 = marked increase).

PFNA plasma concentration measurement

For each PFNA dose, plasma concentrations were measured. To 20 µl plasma, 60 µl ice-cold acetonitrile was added. The sample was incubated at −20 °C for 20 min and centrifuged at 10,000×*g* for 7 min. The supernatant was used for analysis. A matrix-assisted standard curve was established using PFNA and a standard plasma sample (Precinom U, Roche Diagnostics, Hvidovre, Denmark). Samples from the 0.0125 mg/kg bw/day administration group were not diluted, 0.25 mg/kg bw/day samples were diluted tenfold, and the 5 mg/kg bw/day samples were diluted 100-fold. The analysis was conducted using a high-resolution maxis qTOF instrument (Bruker Daltonics, Bremen, Germany) coupled to an Agilent 1200 (Agilent Technologies, USA) with a Supelco C8 (100 × 2.0 mm, 1.7 µm) run in negative mode ionisation with a mass scan from 50 to 800 m/z. The gradient was 0 % B, 0 min—5 % B, 2 min—100 % B, 10 min—100 % B, 12 min—0 % B, 12.1 min—0 % B, 14 min. To test the repeatability of the method, the Low PFNA groups were analysed on two separate days. These measurements gave similar results. In order to investigate the metabolism of PFNA, a targeted approach was used to identify possible metabolites. The metabolites searched for were PFNA + glucuronic acid (+176 m/z value) and PFNA + glycine (+75 m/z value). The targeted approach was conducted using the Bruker Daltonics software: Target Analysis (Bruker Daltonics, Bremen, Germany).

Hormone measurements

Plasma samples were added internal standards (testosterone-d2 and methyltestosterone-d3), deproteinised with addition of acetonitrile and ultracentrifugation, and steroid hormones were extracted using a C18 end-capped solid-phase extraction cartridge (500 mg, 3 ml) (Merck, Darmstadt, Germany). Impurities were removed from the cartridge with demineralised water followed by elution of steroid hormones from the cartridge with methanol. The extract was evaporated to dryness with nitrogen and re-suspended in 40 % acetonitrile. Steroid hormones were separated, detected and quantified using the LC–MS/MS method previously described (Mortensen and Pedersen 2007). However, to accommodate more hormones, minor modifications were made. An Ascentis Express C8 column (2.1 × 100 mm, 2.7 µm) (Supelco) was added to the LC system (Agilent 1100). Steroids were measured with an injection volume of 10 µL in ESI+ mode with acetonitrile and water/0.2 % formic acid as the mobile phases (flow rate 0.25 mL/min, gradient method). The MS was a Quattro Ultima Triple Quadropole Instrument (Waters Corp., Milford, MA, USA). For quantification, external calibration standards were run before and after the samples at levels of 0.1, 0.5, 1.0, 2.0, 5.0 and 10 ng/mL (with 4.0 ng/mL of internal standards). The absolute recoveries of the hormones in plasma samples were estimated to be 42–94 %, based on the absolute recoveries of the internal standards in >30 experiments. The limit of quantification (LOQ) of each of the hormones in the plasma samples was estimated as the concentrations corresponding to six times signal-to-noise and was <100 pg/mL for testosterone (α- and β-isomer), progesterone, corticosterone and hydroxycortisol, <200

**Table 3** Primer/probe sets for mRNA measurements

| Gene | Prod no./sequence | Tissues tested |
|---|---|---|
| Aldo–keto reductase family 1 member C1 (AKR1C1) | Prod. No. Rn01487552_m1, Life Technologies Europe BV, Nærum, Denmark | Liver |
| Benzodiazepine receptor (BZRP) | Forward, 5′-TGG TTC CCT TGG GTC TCT ACA CT-3′ | Testes |
| | Reverse, 5′-CAC CCC ACT GAC AAG CAT GA-3′ | |
| | Probe: 5′-FAM-AAA GCC CAG CCC ATC T-MGB-3′ | |
| Cytochrome P450 (CYP) 1A1 | Prod. No. Rn00487218_m1, Life Technologies Europe BV, Nærum, Denmark | Liver |
| CYP1A2 | Prod. No. Rn00561082_m1, Life Technologies Europe BV, Nærum, Denmark | Liver |
| CYP2B6 | Prod. No. Rn00597739_m1, Life Technologies Europe BV, Nærum, Denmark | Liver |
| CYP2C11 | Prod. No. Rn01502203_m1, Life Technologies Europe BV, Nærum, Denmark | Liver |
| CYP3A23/3A1 | Prod. No. Rn03062228_m1, Life Technologies Europe BV, Nærum, Denmark | Testes |
| CYP11A | Forward 5′-ACG ACC TCC ATG ACT CTG CAA T-3′ | Testes/adipose tissue |
| | Reverse: 5′-CTT CAG CCC GCA GCA TCT-3′ | |
| | Probe: 5′- FAM-CCT TTA TGA AAT GGC ACA CAA CTT GAA GGT CA-TAMRA-3′ | |
| CYP17 | Forward: 5′-GCC ACG GGC GAC AGA A-3′ | Testes/adipose tissue |
| | Reverse: 5′-CCA AGC CTT TGT TGG GAA-3′ | |
| | Probe: 5′-FAM-CGT CAA CCA TGG GAA TAT GTC CAC CAG A-TMARA-3′ | |
| CYP19 | Forward: 5′-AGAACGGTCCGCCCTTTCT-3′ | Testes/adipose tissue |
| | Reverse: 5′-TGGATTCCACACAGACTTCTACCA-3′ | |
| | Probe: 5′-FAM-AGCTCTGACGGGCCCTGGTCTTATTC-TAMRA-3′ | |
| 3β-hydroxysteroid dehydrogenase (3β HSD) | Prod. No. Rn01789220_m1, Life Technologies Europe BV, Nærum, Denmark | Liver |
| 11β HSD | Prod. No. Rn00567167_m1, Life Technologies Europe BV, Nærum, Denmark | Liver |
| 17β HSD | Prod. No. Rn00588942_m1, Life Technologies Europe BV, Nærum, Denmark | Testes/adipose tissue |
| 5α reductase | Prod. No. Rn00575595_m1, Life Technologies Europe BV, Nærum, Denmark | Testes/adipose tissue |
| 18 s-ribosomal RNA | Forward:, 5′-GCC GCT AGA GGT GAA ATT CTT G-3′ | Liver/testes/adipose tissue |
| | Reverse: 5′-GAA AAC ATT CTT GGC AAA TGC TT-3′ | |
| | Probe: 5′-FAM-ACC GGC GCA AGA CGA ACC AGA G-TAMRA-3′ | |
| Steroidogenic acute regulatory protein (StAR) | Forward, 5′-CCC TTG TTT GAA AAG GTC AAG TG-3′ | Testes |
| | Reverse, 5′-TGA AAC GGG AAT GCT GTA GCT-3′ | |
| | 5′-FAM-CTG ACT CCT CTA ACT CCT GTC TGC CTA CAT GGT-TAMRA-3′ | |
| UDP-glucuronosyltransferase 2B15 (UGT2B15) | Prod. No. Rn00755925_m1, Life Technologies Europe BV, Nærum, Denmark | Liver |
| UGT2B17 | Prod. No. Rn01790037_g1, Life Technologies Europe BV, Nærum, Denmark | Liver |

pg/mL for androstenedione and hydroxyprogesterone, and <2,000 for dihydrotestosterone and hydroxytestosterone. Pituitary hormone plasma concentration measurements were done by use of the Milliplex Map Rat Pituitary Kit (Prod no. RPT86K, Millipore Corporation, St. Charles, MO, USA). This was done according to the protocol of the manufacturer by use of a Luminex 100 apparatus (Bio-Rad, Hercules, CA, USA).

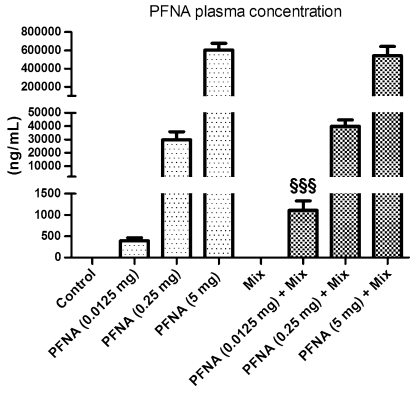



**Fig. 1** Plasma concentration of PFNA in the rats. Rats were administered PFNA at 0.0125, 0.25 or 5 mg/kg bw/day for 14 days in the presence or absence of a mixture of 14 chemicals (Mix) at a total dose of 2.5 mg/kg bw/day. The PFNA plasma concentration was measured in blood from all rats at the end of the experiment using qTOF. PFNA was found in all PFNA treatment groups. At 0.0125 mg/kg bw/day (Low) PFNA, the plasma level was 396 ng/mL for Low PFNA and 1,111 ng/mL for Low PFNA + Mix. At 0.25 mg/kg bw/day, the values were 29,950 and 39,880 ng/mL for Mid PFNA and Mid PFNA + Mix, respectively. For 5 mg/kg bw/day, the values were 602,000 and 541,700 ng/mL for High PFNA and High PFNA + Mix, respectively. Data are mean plus SEM. *N* was 8 except for control and Low PFNA + Mix where *n* was 10. A *t* test was applied with $^{§§§}p < 0.001$ for Low PFNA versus Low PFNA + Mix

### mRNA measurements

Organs were stored in RNAlater at −20 °C until purification by use of the RNAeasy Mini Kit (Qiagen, Hilden, Germany) and cDNA synthesis by use of the Omniscript RT kit (Life Technologies Europe BV, Nærum, Denmark). mRNA levels were next measured by quantitative (q)PCR using specific primer pairs in combination with TaqMan probes (sequences and investigated organs are specified in Table 3). Samples, primers and probes were added TaqMan Fast Universal PCR Master Mix (Life Technologies Europe BV, Nærum, Denmark) and run on a TaqMan 7900 HT qPCR machine (Applied Biosystems, Nærum, Denmark). Quantification was done by use of the comparative threshold cycle (Ct) method, where the Ct value is the cycle number at which the fluorescence signal of the amplified target reaches a defined threshold (Schmittgen and Livak 2008). Ct data on the transcripts of interest were normalised by subtraction of the Ct value of 18 s ribosomal RNA to obtain the ΔCt value. To obtain normal distributed data, $2^{-\Delta Ct}$ values were used for statistical analysis.

### Kidney transporter protein-level measurements

Frozen whole kidneys were thawed on ice and homogenised with a Yellowline DI25 Basic Homogeniser (Bie & Berntsen, Glostrup, Denmark) in an ice-cold buffer containing 300 mM sucrose, 25 mM imidazole, 1 mM EDTA, and 1:200 Calbiochem Protease inhibitor cocktail set II, EDTA-Free (prod. no. 539134, Calbiochem, Darmstadt, Germany). Protein determination and Western blotting were conducted as previously described (Hadrup et al. 2007) except that the employed primary antibodies were anti-organic anion transporter 1 (OAT1) (prod. no. ABIN653184, Antibodies-online, Aachen, Germany), anti-organic anion-transporting polypeptide 1/3 (OATP1/3) (prod. no. Sc-47265, Santa Cruz Biotechnologies, La Jolla, CA) and anti-organic anion-transporting polypeptide 4C1 (OATP4C1) (prod. no. Sc-136775, Santa Cruz Biotechnologies, La Jolla, CA). All antibodies were used at a dilution of 1: 2,000. Moreover, the employed camera was a Gel Doc 2000 (Bio-Rad, Hercules, CA). To accommodate up to 36 samples, a total of three blots were run for each investigated PFNA dose level (0.0125 mg/kg bw/day (PFNA Low) and 0.25 mg/kg bw/day (PFNA Mid) and on each blot a standard sample with a pooled volume of eight randomly selected samples was included, and all samples on each blot were normalised to this sample before data analysis.

### Statistics

Regarding data on mechanisms of toxicity (hormone concentrations, mRNA data, protein-level data), only the Low and Mid PFNA doses were included in the statistical evaluation (and on the graphs). The High PFNA groups were excluded due to the severe toxicity observed on body weight and pathology. Data were analysed for normal distribution by use of the D'Agostino & Pearson omnibus normality test. In case of a lack of normal distribution, data were transformed using the logarithm function and again tested for normal distribution. For normally distributed data, one-way ANOVA with Dunnett's post-test was employed to assess effects of PFNA and of PFNA + Mix, respectively. In case of lack of normality, the nonparametric Kruskal–Wallis test with Dunn's post-test (exposed groups vs. controls) was employed. To assess differences between individual data points, a *t* test was employed. In case of lack of normality, differences between two groups were assessed by use of the Mann–Whitney test. In case of comparison with a group in which all data were below the limit of detection, one sample *t* test was employed. A *p* value of <0.05 was considered significant. The statistical software package was Graph Pad Prism (Graph Pad Software, La Jolla, USA).

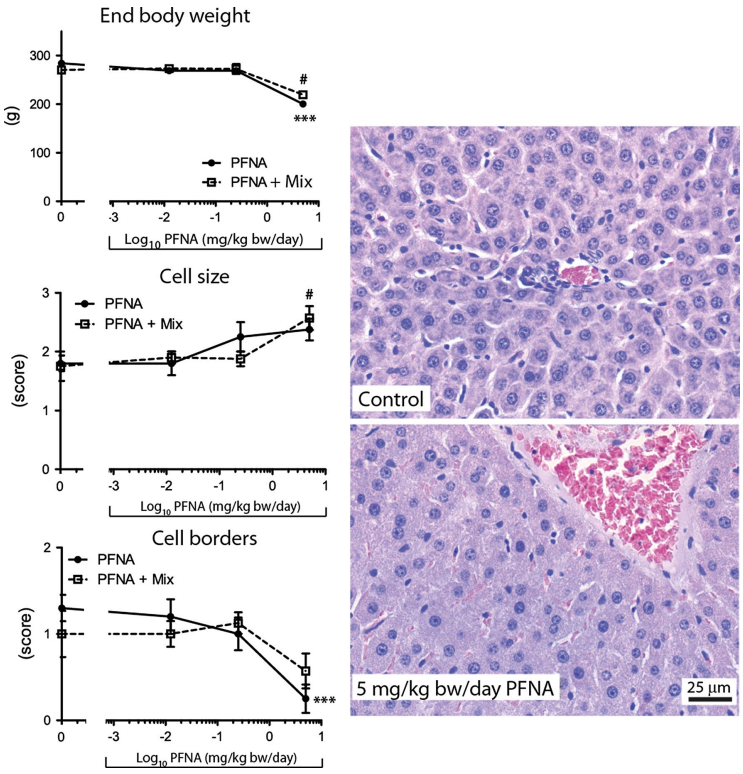



**Fig. 2** At high-dose, PFNA induces toxicity as measured by body weight and pathology. Rats were administered PFNA at 0.0125, 0.25 or 5 mg/kg bw/day for 14 days in the presence or absence of a mixture of 14 chemicals (Mix) at a total dose of 2.5 mg/kg bw/day. A piece of liver was fixed in paraformaldehyde and processed into paraffin-embedded sections and stained with haematoxylin and eosin. Sections were then blinded to the observer and scored by use of microscopy. *Upper* image shows a section from a control rat, *lower* image shows a section from a rat administered 5 mg/kg bw/day PFNA. The *upper* graph shows that the end body weight is increased at High PFNA ± Mix. The *lower* graphs show scores of cells size and cell borders. Cell size is increased by High PFNA ± Mix, and cell borders are decreased by High PFNA ± Mix. Data are mean *plus* or *minus* SEM. *N* was 10 except for Mid PFNA, Mix and Mid PFNA + Mix where *n* was 8. Data are analysed by one-way ANOVA or Kruskal–Wallis test. $^{***}p < 0.001$ by Dunnett's ANOVA post-test for PFNA versus control; and $^{\#}p < 0.05$ by Dunn's Kruskal–Wallis post-test for PFNA + Mix versus control

## Results

### PFNA plasma concentration

PFNA plasma concentrations following administration of 0.0125 mg/kg bw/day were 396 ng/mL for Low PFNA and 1,111 ng/mL for Low PFNA + Mix—a difference that was statistically significant ($p = 0.0007$). Following dosage with 0.25 mg/kg bw/day, the plasma concentrations were 29,950 ng/mL (Mid PFNA) and 39,880 ng/mL for Mid PFNA + Mix. Following dosage with 5 mg/kg bw/day, the values were 602,000 ng/mL (High PFNA) and 541,700 ng/mL for High PFNA + Mix (Fig. 1). Thus, at the low PFNA dose, the Mix caused a pronounced increase

in PFNA plasma levels, but this did not happen at higher PFNA doses.

Body weight, organ weights and pathology

General toxicity was observed with High PFNA with and without Mix. The end body weights were decreased in High PFNA and High PFNA + Mix groups (Fig. 2). Macroscopic pathological examination showed steatotic livers and congestive hearts (pictures not shown) that were observed for High PFNA and High PFNA + Mix. Microscopically, increased size of liver cells (hypertrophy) was observed with increasing doses of PFNA, reaching statistical significance in the High PFNA + Mix group (Fig. 2). Cell borders were less apparent in liver sections of rats receiving increasing doses of PFNA being statistically significant in the High PFNA group.

Hormone plasma concentrations

The corticosterone plasma concentration was increased at the Low PFNA dose level without Mix (twofold). In the presence of the Mix, corticosterone was not increased at the Low and Mid PFNA groups (Fig. 3). For androstenedione, testosterone and dihydrotestosterone increases were found in the Low PFNA + Mix group as compared to Low Mix (Fig. 3). For androstenedione and testosterone, a decrease was found both in the High PFNA and the High PFNA + Mix groups (data not shown). This effect was not detected for dihydrotestosterone for which many measurements were below the level of quantification (data not shown). Regarding pituitary hormones (Fig. 4), Mid PFNA + Mix decreased the plasma concentration of luteinising hormone (LH). The concentration of follicle stimulating hormone (FSH) was decreased by Low PFNA + Mix and adrenocorticotropic hormone (ACTH) was decreased in the Mid PFNA + Mix as compared to Mid PFNA. For prolactin and brain-derived neurotrophic factor, PFNA showed higher concentrations as compared to PFNA + Mix at the Mid dose.

mRNA levels in testis, liver and fatty tissue

In the testes, 17β HSD was down-regulated at both Low and Mid PFNA + Mix as compared to control. This effect was not observed with PFNA alone (Fig. 5). For all other genes measured in testes, there were no significant effects at the mRNA levels of genes involved in regulation of steroid metabolism when considering Low PFNA and Mid PFNA with or without Mix (data not shown). At the High PFNA, where severe toxicity was found, mRNA levels of steroidogenic acute regulatory protein (StAR), benzodiazepine receptor (BZRP), CYP11A, CYP17 and 17β-hydroxysteroid dehydrogenase (17β HSD) were all found to be down-regulated by PFNA (data not shown).

**Fig. 3** PFNA at a dose of 0.0125 mg/kg bw/day plus Mix induces ▶ an effect on androgens. Rats were administered PFNA at 0.0125or 0.25 mg/kg bw/day for 14 days in the presence or absence of a mixture of 14 chemicals (Mix) at a total dose of 2.5 mg/kg bw/day. Steroid hormones were measured in plasma by use of LC–MS/MS. Low PFNA + Mix increased testosterone and dihydrotestosterone. A normalising effect was seen on corticosterone with Low PFNA + Mix. Data are mean plus or minus SEM. *N* was 10 except for Mid PFNA, Mix and Mid PFNA + Mix where *n* was 8. Data were analysed by one-way ANOVA or Kruskal–Wallis test. $^{**}p < 0.01$ by Dunnett's ANOVA post-test for PFNA versus control. $^{\S}p < 0.05$ and $^{\S\S}p < 0.01$ by *t* test for PFNA versus PFNA + Mix. For dihydrotestosterone, a one sample *t* test was applied versus the detection limit of this hormone. $^{\S\S}p < 0.01$

In the liver, 11β HSD was down-regulated with PFNA + Mix at both Low and Mid doses. This effect was not observed for PFNA alone (Fig. 5). For all other genes, no effects were found at the Low and Mid PFNA groups. At the High PFNA groups, where severe toxicity occurred, the aldo–keto reductase family 1 member C1 (AKR1C1), UDP-glucuronosyltransferase 2B15 (UGT2B15), CYP2C11, CYP1A2 and CYP2B6 were all found to be down-regulated, whereas CYP3A23/3A1 was found to be up-regulated (data for the high-dose groups are not shown on the graphs).

In adipose tissue, CYP19 mRNA was up-regulated for Low PFNA + Mix as compared to Low PFNA (Fig. 5). At the toxic High PFNA, CYP11 was up-regulated (data not shown). There were no effects on the other steroid metabolism enzymes: CYP17, 17β HSD and 5α reductase (data not shown).

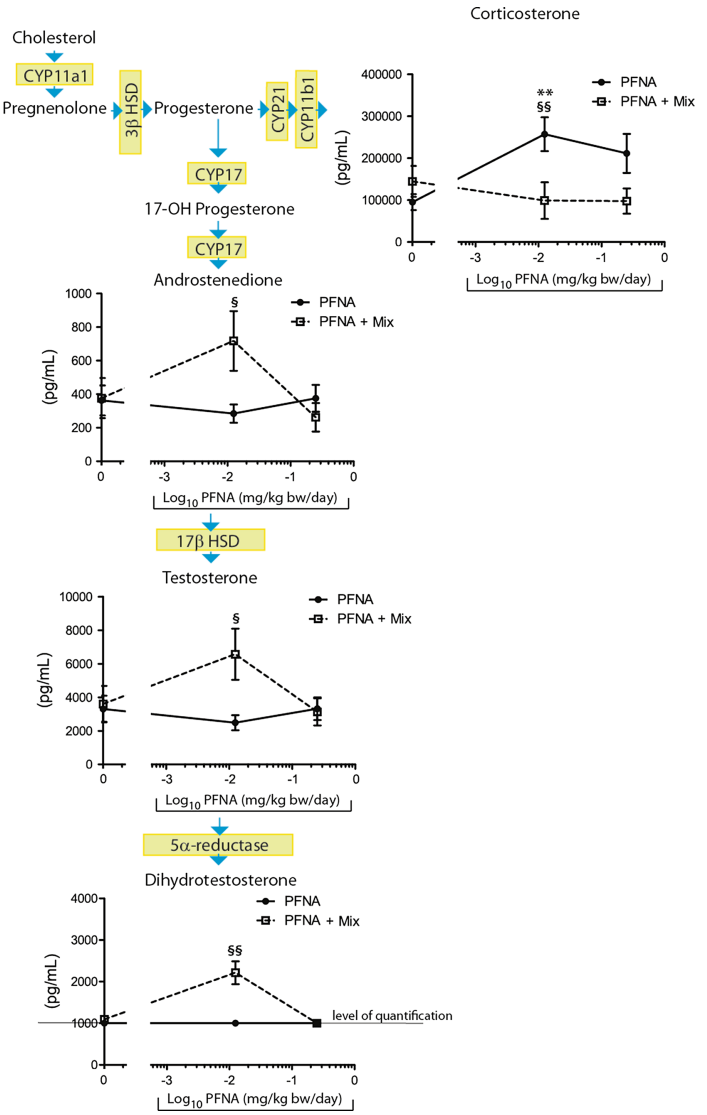
Kidney transporter protein levels

OAT1 was increased in whole kidney homogenates when Low PFNA was combined with Mix (Fig. 6). OATP4C1 was decreased by Low PFNA alone. This effect was normalised by addition of Mix (Fig. 6). At the Mid dose level, no significant effects were found. OATP1/3 was not affected at any of the investigated dose levels (data not shown).

## Discussion

Relevance of the employed PFNA doses in relation to human exposure levels

The hypothesis of the current study was that adverse mixture effects occur at doses approaching high-end human exposure levels of food chemicals. Our aim was to reach a plasma level of PFNA covering the combined human


Cholesterol
CYP11a1
Pregnenolone
3β HSD
Progesterone
CYP21
CYP11b1
CYP17
17-OH Progesterone
CYP17
Androstenedione
17β HSD
Testosterone
5α-reductase
Dihydrotestosterone
Corticosterone
PFNA
PFNA + Mix
(pg/mL)
Log10 PFNA (mg/kg bw/day)
§
§§
**
level of quantification

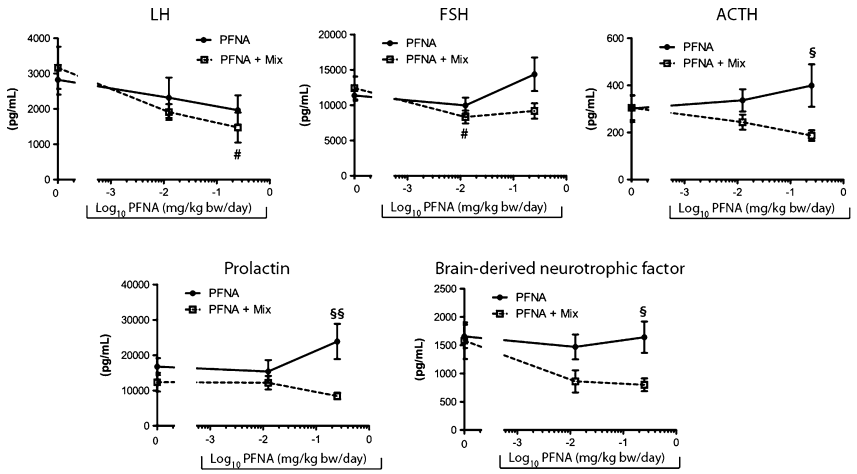



**Fig. 4** Effects of dose PFNA and Mix on pituitary hormone levels. Rats were administered PFNA at 0.0125 or 0.25 mg/kg bw/day for 14 days in the presence or absence of a mixture of 14 chemicals (Mix) at a total dose of 2.5 mg/kg bw/day. Hormones were measured by use of a Milliplex Map Rat Pituitary Luminex Kit. LH was decreased by Mid PFNA + Mix. FSH was decreased by Low PFNA + Mix. For ACTH, prolactin and brain-derived neurotrophic factor Mid PFNA + Mix was lower than Mid PFNA alone. Data are mean *plus* or *minus* SEM. N was 8 except for control ($n = 10$) and Low PFNA ($n = 9$). Data were analysed by one-way ANOVA. $^{\#}p < 0.05$ by Dunnett's ANOVA post-test for PFNA + Mix versus control. $^{\S}p < 0.05$ and $^{\S\S}p < 0.01$ by $t$ test for PFNA versus PFNA + Mix

plasma concentrations of perfluorinated compounds such as PFOA, PFOS and PFNA. According to Lau et al. (2007), human exposure means obtained from a substantial number of investigations gave a combined value of these three perfluorinated compounds of 13–429 ng/mL (or 13–117 ng/mL if the second highest PFOA value is used). For comparison, the 95th percentile obtained from the NHANES study was 67.6 ng/mL for PFOA, PFNA and PFOS combined (CDC 2009) which is in line with the Lau et al. data. A PFNA plasma concentration of 396 ng/mL was found in the Low PFNA group, indicating that PFNA in this study may be up to a factor of six higher than a high-end human exposure level (Table 1). However, this depends on the choice of a 'high-end' exposure for which a fixed, true value do not exist. In a study, not reported in the review by Lau et al., Emmett et al. found a mean of 423 ng/mL PFOA for persons not subjected to occupational exposure (median 329, $n = 312$) and a mean of 824 ng/mL in persons with substantial occupational exposure (median 775, $n = 18$) (Emmett et al. 2006). Taking this into account, the present PFNA exposure in the rats was close to a combined human PFNA, PFOA and PFOS exposure. We made the assumption that the applied PFNA exposure should represent exposure to other perfluorinated congeners, although it can be argued that the difference in structures of these molecules affects potency and binding specificity to different target molecules in the mammalian body. It is noted that a range of congeners different from PFNA, PFOA and PFOS exist, and these likely also contribute to the combined effects of perfluorinated compounds in humans.

## Observed mixture effects and discussion on the presence of interactions

The aim of the present investigation was to test whether a relevant mixture of food chemicals at doses approaching high-end human exposure levels exerted adverse effects. The study was designed to target testosterone, and we found this steroid as well as its downstream metabolite, dihydrotestosterone and its precursor, androstenedione, to be affected by the mixture. There were no significant effects of Low PFNA or of Mix alone, whereas androgen levels increased considerably after exposure to Low PFNA in combination with Mix; thus for testosterone, dihydrotestosterone, and androstenedione a non-monotonic effect seemed to have occurred. Previously, additive mixture effects have

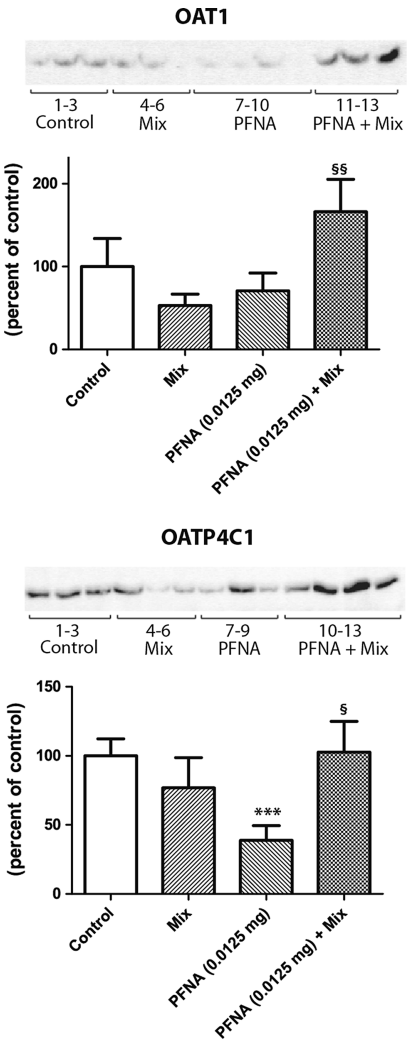



**Fig. 6** Effects of low-dose PFNA and Mix on protein levels of organic anion transporter 1 and organic anion-transporting polypeptide. Rats were administered PFNA at 0.0125 mg/kg bw/day for 14 days in the presence or absence of a mixture of 14 chemicals (Mix) at a total dose of 2.5 mg/kg bw/day. Protein levels of OAT1 (organic anion transporter 1) and OATP4C1 (organic anion-transporting polypeptide-4C1) were measured by Western blotting. All samples were measured in a total number of three blots per protein. Only one blot is shown on the figure. A standard sample on each blot was used for normalisation to obtain a graph representing all samples ($n = 9$ for control, $n = 8$ for Mix, $n = 10$ for Low PFNA (0.0125 mg/kg bw/day) and $n = 10$ for Low PFNA + Mix). OAT1 was increased at Low PFNA + Mix as compared to Low PFNA. OATP4C1 was decreased at Low PFNA as compared to control, and this effect was normalised at Low PFNA + Mix. Data are mean plus or minus SEM. Data were not normal distributed and were therefore tested by use of the nonparametric Mann–Whitney test. $^{***}p < 0.001$ for PFNA versus control. $^{\S}p < 0.05$ and $^{\S\S}p < 0.01$ for PFNA versus PFNA + Mix

been found for chemicals each present at doses for which measurement techniques are not sensitive enough to detect their individual effects (Hass et al. 2007; Silva et al. 2002). Feng et al. have previously demonstrated a non-monotonic dose–response curve of PFNA on testosterone in rats with an increase at 1 mg/kg bw/day and a decrease at 5 mg/kg bw/day. This supports the presence of such a non-monotonic relationship although this was seen at higher-dose levels. Also Wade et al. (2002) found a non-monotonic dose–response effect on natural killer cell lytic activity following exposure to a mixture containing 18 persistent contaminants at doses ranging from 1× to 100× the estimated safe level.

Looking at the PFNA plasma concentration, a toxicokinetic interaction was found in that the PFNA plasma concentration was increased at Low PFNA + Mix as compared to Low PFNA. PFNA could not be detected in animals only given Mix, indicating that no background PFNA levels, e.g., from the feed were present in the animals. Thus, the 14-chemical mixture is able to increase the PFNA plasma levels at lower doses possibly by interference with ADME issues for PFNA. For corticosterone plasma levels, Low PFNA + Mix normalised an increase in corticosterone observed for Low PFNA. Also on the kidney OATP4C1 transporter, the addition of Mix to Low PFNA normalised a Low PFNA-induced decrease in the protein level. These data suggest that in addition to reported food–drug interactions (Bailey et al. 1989), food–environment chemical interactions may occur at the toxicokinetic level at doses approaching high-end human exposure levels. The effect of Mix on PFNA plasma levels is due to a toxicokinetic interaction between Mix and PFNA, but whether the effect on steroid hormone levels is due to kinetic or dynamic interferences is unknown. It is striking that we have detected potential non-monotonic low-dose effects. This gives food for thought concerning the extrapolation from high to low doses typically done in toxicological studies and concerning future human risk assessment of chemicals.

In this study, possible effects of a chemical mixture were investigated in juvenile male rats 7 weeks of age at onset. Male Wistar rats are considered sexually mature when they

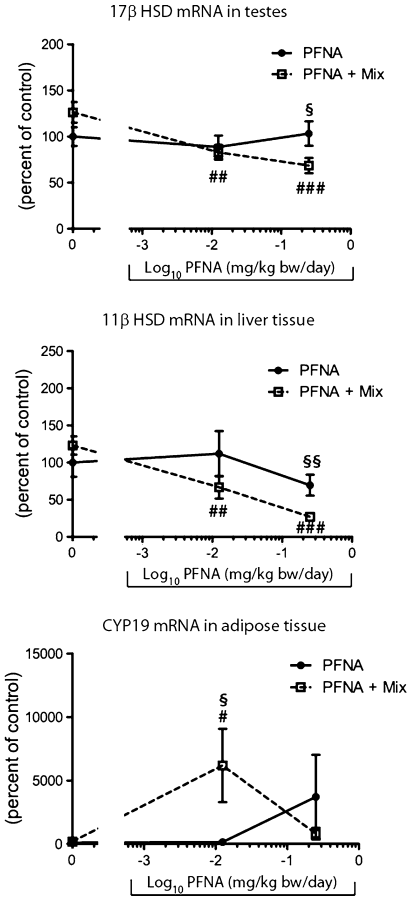



**Fig. 5** mRNA levels of 17β HSD in testes, 11β HSD in liver tissue and CYP19 in adipose tissue. Rats were administered PFNA at 0.0125 or 0.25 mg/kg bw/day for 14 days in the presence or absence of a mixture of 14 chemicals (Mix) at a total dose of 2.5 mg/kg bw/day. mRNA levels were measured by use of qPCR. 11β and 17β HSD were decreased by Low and Mid PFNA + Mix. CYP19 was higher for Low PFNA + Mix as compared to Low PFNA. Data are mean *plus* or *minus* SEM. N was 10 except for Mid PFNA, Mix and Mid PFNA + Mix where n was 8. Data were analysed by one-way ANOVA. $^{\#}p < 0.05$, $^{\#\#}p < 0.01$ and $^{\#\#\#}p < 0.001$ by Dunnett's ANOVA post-test for PFNA + Mix versus control. $^{\S}p < 0.05$ and $^{\S\S}p < 0.01$ by *t* test for PFNA versus PFNA + Mix

are around 50 days old; thus, we have investigated effects in rats just post-puberty. In our view, marked hormone changes may be considered as adverse. Increased testosterone in males has for example been linked to increased aggression (McGinnis 2004); and androgenic anabolic steroid abuse has been linked to cardiac disease (Ahlgrim and Guglin 2009).

## Possible mechanisms involved in the mixture effects on androgens

The possible mechanisms underlying the observed mixture effects could provide a deeper understanding of how such effects develop in the mammalian body. Regarding the mixture effects on androgens, mechanisms involved could be related to (a) increased production or release of steroid hormones either via hormonal factors or locally induced (b) decreased tissue deposition of hormones (c) decreased metabolism of androgens and/or (d) decreased excretion of hormones. The pituitary hormone LH is involved in the regulation of testosterone production in the testes, and LH was decreased by Mid PFNA + Mix, and this may reflect a secondary negative feedback on the pituitary caused by disturbed steroid hormone levels.

In the testes, there were no findings to explain the effects seen on androgens except that a disturbed 17β HSD mRNA regulation may be part of the explanation. In adipose tissue, CYP19 was increased at Low PFNA + Mix as compared to Low PFNA. This suggests increased conversion of testosterone to estradiol when Mix is present. Among the chemicals in Mix, only DDE has shown a link to CYP19. Daughters of Michigan fish-eaters had increased gene expression of CYP19 in blood leucocytes that was correlated to prenatal DDE levels (Karmaus et al. 2011). The fact that CYP19 was increased in the current study indicates that a compensatory mechanism to eliminate the excess testosterone might be in play.

In the liver, CYP expression was investigated because CYP3A4 and CYP2C9 (CYP3A23/A1 and 2C11in rats) are involved in the hydroxylation of testosterone, and these enzymes are known to be inhibited by bergamottin and glabridin constituents of the Mix (Cheng and Schenkman 1983; Guengerich 1999; Kenworthy et al. 1999; Martignoni et al. 2006; Tassaneeyakul et al. 2000; Yamazaki and Shimada 1997). Moreover, in the ToxCast (high throughput in vitro screening project) several perfluorinated carboxylic acids including PFNA as well as BPA and prochloraz (constituents of the Mix) have been found in vitro to inhibit CYP2C9 (US_Environmental_Protection_Agency 2014). However, no effects on the mRNA levels of CYP3A23/1 or CYP 2C11 for Low PFNA in combination with Mix were found. Also, on the steroid glucuronidation pathway, there were no findings to explain the effects seen on androgens.

To study interaction via excretion, we investigated organic anion transporters in the kidney because these transporters have been suggested to play a role in PFNA and testosterone excretion. OAT1 as well as OATP4C1 transport organic anions from the blood into the kidney tubular cells for subsequent secretion into the pre-urine (Han et al. 2012). With Low PFNA + Mix, an increase in protein level of OAT1 was found (Fig. 6) but neither Mix nor PFNA individually exerted this effect. It has been shown that testosterone stimulates OAT1 (Cerrutti et al. 2002; Ljubojevic et al. 2004) thus the increased testosterone level seen at Low PFNA + Mix correlates well with an increase in this transporter. This up-regulation, however, might also reflect that OAT1 could transport PFNA and that this transporter is up-regulated as a compensatory mechanism in order to excrete the increased amounts of PFNA. For OATP4C1, the Mix normalised the Low PFNA-induced decrease in the protein level, suggesting that the Mix protects the body by allowing an increased transport of PFNA (Fig. 6). As for OAT1, OATP4C1 is increased by testosterone (Lu et al. 1996), again suggesting the increased testosterone level to be causing the normalisation of the OATP4C1 expression. Taken together, the data on kidney transport do not readily explain the mixture effect on androgens at Low PFNA + Mix and may rather reflect downstream effects of the increased androgen levels.

Possible mechanisms underlying the increased PFNA concentration caused by Mix

The increase in the PFNA plasma concentration following Low PFNA + Mix as compared to Low PFNA may be explained by (1) increased PFNA absorption, (2) decreased metabolism, (3) decreased tissue deposition, or (4) decreased excretion. We looked for metabolites of PFNA but did not find any significant metabolism, and this is in accordance with the literature (Lau et al. 2007). Regarding proposed reabsorption of perfluorinated compounds, the OATP1 is located on the luminal side of the tubular cells (Han et al. 2012) and has been demonstrated to have the ability to transport perfluorocarboxylates (Yang et al. 2009b). OATP1 was not found being affected by Low PFNA alone or in combination with Mix; thus, testosterone-stimulated reabsorption of PFNA via this transporter seems not to be underlying the increased plasma concentration of PFNA. However, it should be noted that we did not measure two other OATs involved in perfluorocarboxylate transport, namely OAT2 and OAT3 (Kudo et al. 2002). Whether the PFNA plasma-level effect is a cause or a consequence of (or not related to) the androgen levels is unknown. However, when we depict the effect on androgen levels as a function of the internal PFNA plasma levels (graph not shown), we still observe a pronounced increase in androgen levels, indicating that there is more to the androgen effect than just affected toxicokinetics by PFNA alone.

Possible mechanisms underlying the normalisation of corticosterone

The mechanism underlying the normalisation of the corticosterone level by the Mix is likely explained by the diminished plasma ACTH observed when Mix was administered along with PFNA. Thus, a central effect on the pituitary may be evident. The enzymes responsible for inter-conversion of corticosterone to the physiologically inactive 11-dehydrocorticosterone in rats are 11β HSD (Thomson et al. 1998). We found PFNA + Mix to down-regulate this enzyme at the mRNA level. A differential regulation of this enzyme may be an explanation of the observed differential effect on corticosterone caused by PFNA with or without Mix.

Mixture-independent effects of PFNA suggest non-monotonic dose–response relationships

At Low PFNA dose, an increased corticosteroid effect as well as a diminished effect of the OATP4C1 protein level in kidney was observed (Fig. 6), suggesting that this chemical exerts low-dose effects on its own. The PFNA effect on corticosterone plasma levels could involve regulation of 11β HSD. PFNA has shown a potential to activate liver X receptor α (LXRα) (US_Environmental_Protection_Agency 2014), and LXR negatively regulates expression of 11β HSD (Vogeli et al. 2013). At High PFNA, macroscopical and microscopical pathology of the liver as well as decreased body and organ weights showed that this dose is highly toxic to the animals. From that point of view, it is not surprising that the mRNA levels of several enzymes are down-regulated as seen for testes and liver, along with the testosterone and androstenedione plasma concentration being down-regulated at the high dose. Several enzymes like StAR, BZRP and CYP11A in the testes were down-regulated and provide a suggestion for the decrease in testosterone and androstenedione at high doses. Steatosis, obscure hepatic cell borders and a decrease in testosterone have been shown for perfluorinated compounds by others previously (Fang et al. 2010; Feng et al. 2009; Yang et al. 2009a). Notably, when these toxic effects occur, CYP1A2, CYP2B6 and CYP2C11 were down-regulated by PFNA, whereas CYP3A23/3A1 corresponding to the human CYP3A4 was up-regulated suggesting that this CYP is up-regulated to increase elimination of multiple toxic metabolites and confirming the broad substrate specificity of this enzyme. Our findings highlight the challenge of extrapolating from high-dose to low-dose effects, as is the ordinary

practice within toxicology and suggest that lower doses should be employed as well in toxicological studies.

## Conclusion

We found that a human relevant mixture of fifteen chemicals given to rats at doses approaching human realistic high-end exposure levels disturbed several plasma steroid and pituitary hormone levels. Androgen levels were non-monotonically increased by >100 %, and corticosterone levels were decreased by 60 %. Moreover, a toxicokinetic interaction may have occurred as the Mix caused a markedly increased PFNA plasma concentration. Our data suggests that mixture effects of chemicals may be non-monotonic and may occur even at doses approaching high-end human exposure levels. Further studies are warranted to determine whether this reflects a general phenomenon that should be taken into consideration when predicting human mixture toxicities.

**Acknowledgments** Excellent technical assistance was provided by Anne Ørngreen, Maja Danielsen, Eva Ferdinansen, Elise E. Navntoft, Eigil V. Frank, Kenneth R. Worm, Kitt Lademann, Lis Abildgaard Andersen, Birgitte Møller Plesning, Liljana Petrevska, Heidi Letting and Dorte Lykkegaard Korsbech. The Ministry of Food, Agriculture and Fisheries of Denmark and the Danish Veterinary and Food Administration are acknowledged for their financial support.

**Conflict of interest** The authors declare that there are no conflicts of interest.

## References

Ahlgrim C, Guglin M (2009) Anabolics and cardiomyopathy in a bodybuilder: case report and literature review. J Cardiac Fail 15(6):496–500. doi:10.1016/j.cardfail.2008.12.014

Bailey DG, Dresser GK, Bend JR (2003) Bergamottin, lime juice, and red wine as inhibitors of cytochrome P450 3A4 activity: comparison with grapefruit juice. Clin Pharmacol Ther 73(6):529–537. doi:10.1016/S0009-9236(03)00051-1

Bailey DG, Spence JD, Edgar B, Bayliff CD, Arnold JM (1989) Ethanol enhances the hemodynamic effects of felodipine. Clin Invest Med 12(6):357–362

CDC (2009) Fourth national report on human exposure to environmental chemicals. Department of health and human services centers for disease control and prevention

Cerrutti JA, Brandoni A, Quaglia NB, Torres AM (2002) Sex differences in p-aminohippuric acid transport in rat kidney: role of membrane fluidity and expression of OAT1. Mol Cell Biochem 233(1–2):175–179. doi:10.1023/A:1015563021602

Cheng KC, Schenkman JB (1983) Testosterone metabolism by cytochrome P-450 isozymes RLM3 and RLM5 and by microsomes. Metabolite identification. J Biol Chem 258(19):11738–11744

Christiansen S, Kortenkamp A, Axelstad M et al (2012) Mixtures of endocrine disrupting contaminants modelled on human high end exposures: an exploratory study in rats. Int J Androl 35(3):303–316. doi:10.1111/j.1365-2605.2011.01242.x

Crofton KM, Craft ES, Hedge JM et al (2005) Thyroid-hormone-disrupting chemicals: evidence for dose-dependent additivity or synergism. Environ Health Perspect 113(11):1549–1554

Demur C, Metais B, Canlet C et al (2013) Dietary exposure to a low dose of pesticides alone or as a mixture: the biological metabolic fingerprint and impact on hematopoiesis. Toxicology 308:74–87. doi:10.1016/j.tox.2013.03.004

Emmett EA, Shofer FS, Zhang H, Freeman D, Desai C, Shaw LM (2006) Community exposure to perfluorooctanoate: relationships between serum concentrations and exposure sources. J Occup Environ Med/American College of Occupational and Environmental Medicine 48(8):759–770. doi:10.1097/01.jom.0000232486.07658.74

Fang X, Feng Y, Wang J, Dai J (2010) Perfluorononanoic acid-induced apoptosis in rat spleen involves oxidative stress and the activation of caspase-independent death pathway. Toxicology 267(1–3):54–59. doi:10.1016/j.tox.2009.10.020

Feng Y, Shi Z, Fang X, Xu M, Dai J (2009) Perfluorononanoic acid induces apoptosis involving the Fas death receptor signaling pathway in rat testis. Toxicol Lett 190(2):224–230. doi:10.1016/j.toxlet.2009.07.020

Foti RS, Wahlstrom JL (2008) The role of dietary supplements in cytochrome P450-mediated drug interactions. Bol Latinoam Caribe 7(2):66–84

Fuhrman B, Buch S, Vaya J et al (1997) Licorice extract and its major polyphenol glabridin protect low-density lipoprotein against lipid peroxidation: in vitro and ex vivo studies in humans and in atherosclerotic apolipoprotein E-deficient mice. Am J Clin Nutr 66(2):267–275

Guengerich FP (1999) Cytochrome P-450 3A4: regulation and role in drug metabolism. Annu Rev Pharmacol Toxicol 39:1–17. doi:10.1146/annurev.pharmtox.39.1.1

Hadrup N, Loeschner K, Bergstrom A et al (2012) Subacute oral toxicity investigation of nanoparticulate and ionic silver in rats. Arch Toxicol 86(4):543–551

Hadrup N, Petersen JS, Windfeld S et al (2007) Differential down-regulation of aquaporin-2 in rat kidney zones by peripheral nociceptin/orphanin FQ receptor agonism and vasopressin type-2 receptor antagonism. J Pharmacol Exp Ther 323(2):516–524

Hadrup N, Taxvig C, Pedersen M, Nellemann C, Hass U, Vinggaard AM (2013) Concentration addition, independent action and generalized concentration addition models for mixture effect prediction of sex hormone synthesis in vitro. PLoS One 8(8):e70490. doi:10.1371/journal.pone.0070490

Han X, Nabb DL, Russell MH, Kennedy GL, Rickard RW (2012) Renal elimination of perfluorocarboxylates (PFCAs). Chem Res Toxicol 25(1):35–46. doi:10.1021/Tx200363w

Hass U, Scholze M, Christiansen S et al (2007) Combined exposure to anti-androgens exacerbates disruption of sexual differentiation in the rat. Environ Health Perspect 115(Suppl 1):122–128

Karmaus W, Osuch JR, Landgraf J, Taffe B, Mikucki D, Haan P (2011) Prenatal and concurrent exposure to halogenated organic compounds and gene expression of CYP17A1, CYP19A1, and oestrogen receptor alpha and beta genes. Occup Environ Med 68(6):430–437. doi:10.1136/oem.2009.053249

Kent UM, Aviram M, Rosenblat M, Hollenberg PF (2002) The licorice root derived isoflavan glabridin inhibits the activities of human cytochrome P450S 3A4, 2B6, and 2C9. Drug Metab Dispos 30(6):709–715

Kenworthy KE, Bloomer JC, Clarke SE, Houston JB (1999) CYP3A4 drug interactions: correlation of 10 in vitro probe substrates. Br J Clin Pharmacol 48(5):716–727

Kortenkamp A (2014) Low dose mixture effects of endocrine disrupters and their implications for regulatory thresholds in chemical risk assessment. Curr Opin Pharmacol 19C:105–111. doi:10.1016/j.coph.2014.08.006

Kudo N, Katakura M, Sato Y, Kawashima Y (2002) Sex hormone-regulated renal transport of perfluorooctanoic acid. Chem-Biol Interact 139(3):301–316. doi:10.1016/S0009-2797(02)00006-6

Lau C, Anitole K, Hodes C, Lai D, Pfahles-Hutchens A, Seed J (2007) Perfluoroalkyl acids: a review of monitoring and toxicological findings. Toxicol Sci 99(2):366–394. doi:10.1093/toxsci/kfm128

Lim HK, Duczak N Jr, Brougham L, Elliot M, Patel K, Chan K (2005) Automated screening with confirmation of mechanism-based inactivation of CYP3A4, CYP2C9, CYP2C19, CYP2D6, and CYP1A2 in pooled human liver microsomes. Drug Metab Dispos 33(8):1211–1219. doi:10.1124/dmd.104.003475

Ljubojevic M, Herak-Kramberger CM, Hagos Y et al (2004) Rat renal cortical OAT1 and OAT3 exhibit gender differences determined by both androgen stimulation and estrogen inhibition. Am J Physiol Renal 287(1):F124–F138. doi:10.1152/ajprenal.00029.2004

Lu R, Kanai N, Bao Y, Wolkoff AW, Schuster VL (1996) Regulation of renal oatp mRNA expression by testosterone. Am J Physiol Renal Fluid Electrolyte Physiol 270(2):F332–F337

Martignoni M, Groothuis GM, de Kanter R (2006) Species differences between mouse, rat, dog, monkey and human CYP-mediated drug metabolism, inhibition and induction. Expert Opin Drug Metab Toxicol 2(6):875–894. doi:10.1517/17425255.2.6.875

McGinnis MY (2004) Anabolic androgenic steroids and aggression: studies using animal models. Ann N Y Acad Sci 1036:399–415. doi:10.1196/annals.1330.024

Merhi M, Demur C, Racaud-Sultan C et al (2010) Gender-linked haematopoietic and metabolic disturbances induced by a pesticide mixture administered at low dose to mice. Toxicology 267(1–3):80–90. doi:10.1016/j.tox.2009.10.024

Mortensen SK, Pedersen M (2007) Confirmatory analysis of acetylgestagens in plasma using liquid chromatography-tandem mass spectrometry. Anal Chim Acta 586(1–2):217–222

Reagan-Shaw S, Nihal M, Ahmad N (2008) Dose translation from animal to human studies revisited. FASEB J 22(3):659–661. doi:10.1096/fj.07-9574LSF

Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc 3(6):1101–1108

Sigurjonsdottir HA, Franzson L, Manhem K, Ragnarsson J, Sigurdsson G, Wallerstedt S (2001) Liquorice-induced rise in blood pressure: a linear dose-response relationship. J Hum Hypertens 15(8):549–552. doi:10.1038/sj.jhh.1001215

Silva E, Rajapakse N, Kortenkamp A (2002) Something from "nothing"–eight weak estrogenic chemicals combined at concentrations below NOECs produce significant mixture effects. Environ Sci Technol 36(8):1751–1756

Stanko JP, Enoch RR, Rayner JL et al (2010) Effects of prenatal exposure to a low dose atrazine metabolite mixture on pubertal timing and prostate development of male Long-Evans rats. Reprod Toxicol 30(4):540–549. doi:10.1016/j.reprotox.2010.07.006

Tassaneeyakul W, Guo LQ, Fukuda K, Ohta T, Yamazoe Y (2000) Inhibition selectivity of grapefruit juice components on human cytochromes P450. Arch Biochem Biophys 378(2):356–363. doi:10.1006/abbi.2000.1835

Tatum-Gibbs K, Wambaugh JF, Das KP et al (2011) Comparative pharmacokinetics of perfluorononanoic acid in rat and mouse. Toxicology 281(1–3):48–55. doi:10.1016/j.tox.2011.01.003

Thomson LM, Raven PW, Smith KE, Hinson JP (1998) Effects of metyrapone on hepatic cortisone-cortisol conversion in the rat. Endocr Res 24(3–4):607–611

Tian M, Yan H, Row KH (2008) Extraction of glycyrrhizic acid and glabridin from licorice. Int J Mol Sci 9(4):571–577

Uesawa Y, Mohri K (2006) The use of heat treatment to eliminate drug interactions due to grapefruit juice. Biol Pharm Bull 29(11):2274–2278

US_Environmental_Protection_Agency (2014) Interactive chemical safety for sustainability (iCSS) dashboard. http://actor.epa.gov/dashboard/. Accessed 20 March 2014

Vogeli I, Jung HH, Dick B et al (2013) Evidence for a role of sterol 27-hydroxylase in glucocorticoid metabolism in vivo. J Endocrinol 219(2):119–129. doi:10.1530/JOE-13-0141

Wade MG, Foster WG, Younglai EV et al (2002) Effects of subchronic exposure to a complex mixture of persistent contaminants in male rats: systemic, immune, and revoductive effects. Toxicol Sci 67(1):131–143. doi:10.1093/toxsci/67.1.131

Wen YH, Sahi J, Urda E et al (2002) Effects of bergamottin on human and monkey drug-metabolizing enzymes in primary cultured hepatocytes. Drug Metab Dispos 30(9):977–984

Yamazaki H, Shimada T (1997) Progesterone and testosterone hydroxylation by cytochromes P450 2C19, 2C9, and 3A4 in human liver microsomes. Arch Biochem Biophys 346(1):161–169. doi:10.1006/abbi.1997.0302

Yang C, Tan YS, Harkema JR, Haslam SZ (2009a) Differential effects of peripubertal exposure to perfluorooctanoic acid on mammary gland development in C57Bl/6 and Balb/c mouse strains. Reprod Toxicol 27(3–4):299–306. doi:10.1016/j.reprotox.2008.10.003

Yang CH, Glover KP, Han X (2009b) Organic anion transporting polypeptide (Oatp) 1a1-mediated perfluorooctanoate transport and evidence for a renal reabsorption mechanism of Oatp1a1 in renal elimination of perfluorocarboxylates in rats. Toxicol Lett 190(2):163–171. doi:10.1016/j.toxlet.2009.07.011

## 5.5 Manuscript IV

### 5.5.1 Abstract

Study background: Humans are simultaneously exposed to multiple compounds, many of which can be detected in human body fluids; however, the consequences of low dose exposure to complex mixtures are poorly understood. By use of two 'omics' methods, metabolomics and transcriptomics, we have profiled the effects on rats caused by exposure to a 14-compound mixture (Mix) ± perfluorononanoic acid (PFNA).

Methods: Adult male rats were dosed orally for 14 days with PFNA at 0.0125, 0.25 and 5 mg/kg/day ± Mix at 2.5 mg/kg/day. Vehicle control and Mix alone were used as reference groups. The plasma metabolomes were investigated using an HPLC system combined with a maxis qTime-of-flight mass spectrometer. Liver gene expression levels were analyzed using Agilent Whole Rat Genome Oligo Microarrays.

Results: Sixty-three and 64 metabolites were significantly changed upon exposure to Mix and PFNA + Mix, respectively. Twelve identified metabolites were changed in both settings. The affected metabolites were mainly lipids; however, various lipid classes were affected differentially among the study groups. In the livers, 182 and 203 genes mainly related to energy homeostasis and lipid metabolism were differentially expressed upon exposure to PFNA ± Mix, respectively.

Conclusion: The applied technologies, metabolomics and transcriptomics, provided complimentary information allowing for a detailed analysis of the affected signaling pathways. Mix alone caused effects on lipid metabolism evident in plasma. The hepatic effects on lipid metabolism were mainly driven by PFNA. This study verifies that a chemical mixture given at high-end human exposure levels can affect lipid homeostasis. (*submitted*).

# Metabolite and gene expression profiling of rats exposed to perfluorononanoic acid combined with a low-dose mixture of 14 human relevant compounds

Kasper Skov[1*♯], Kristine Kongsbak[1*], Niels Hadrup[1], Henrik Lauritz Frandsen[1], Jørn Smedsgaard[1], Karine Audouze[2,3], Aron Charles Eklund[2], Anne Marie Vinggaard[1]

1 National Food Institute, Technical University of Denmark, DK-2860 Søborg, Denmark

2 Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

3 Inserm UMR-S 973, Molécules Thérapeutiques In Silico, Paris Diderot University, Paris, France

* These authors contributed equally to the work.

♯ Corresponding author:

Mørkhøj Bygade 19
DK-2860 Søborg
Denmark
Email: kasko@food.dtu.dk
Phone: +45 35887624
Fax: +45 35887448

Abbreviated title: Metabolomics and transcriptomics toxicity profiling of PFNA exposure

Financial acknowledgements: The Ministry of Food, Agriculture and Fisheries of Denmark and the Danish Veterinary and Food Administration are acknowledged for their financial support.

## Introduction

Humans, especially those living in industrialized countries, are exposed to multiple compounds every day through e.g. foods, cosmetics, pharmaceuticals, and air inhalation (Monosson 2005). Biomonitoring and epidemiological studies have, additionally, shown that the exposure to multiple compounds (human relevant mixtures) results in detectable levels of multiple compounds in human body fluids (Calafat et al 2007; NHANES 2013). Previously, animal studies have shown effects of human relevant mixtures of environmental compounds given at doses close to No Observed Adverse Effect levels (NOAELs) for the single compounds (Christiansen et al 2008; Christiansen et al 2012; Axelstad et al 2014). In epidemiological studies, associations between compound mixtures and e.g. reproductive diseases have been proposed (Krysiak-Baltyn et al 2012; Taylor et al 2014).

The traditional toxicology approach aims at understanding the effect(s) of a single toxic compound on biological systems (e.g. activity at a specific receptor, activity at cell or organ level, or activity on a whole organism/animal model). The compound is typically investigated at varying doses in order to obtain information on e.g. the NOAEL. A few studies of rats exposed to chemicals or chemical mixtures at doses representing human exposure levels have been reported recently (Moser et al 2006; Chen et al 2014). These studies suggest that even low doses of compounds exert effects on e.g. the plasma metabolite composition. However, no studies have, to our knowledge, investigated the effects of chemical mixtures on the metabolome or the gene expression levels. “Omics” techniques have revealed new approaches to evaluate the effects of toxic compounds. Fluorinated compounds have been shown to change the lipid metabolism and affect both rats and zebra fish (Fang et al 2012a; Zhang et al 2012) and reproductive toxicants such as bisphenol A have also shown a change in the metabolome even at doses far below the NOAEL (Chen et al 2014). In general, changes in the transcriptome, proteome and metabolome are detectable even at low doses of compounds, which indicates that these methods might help in understanding how the compounds affect the organism (Chen and Kim 2013). Furthermore, it has been shown that a mixture of low dose compounds exerted a marked mixture effect even if individual compounds by themselves had no effect (Silva et al 2002).

We wanted to investigate the effects of a compound mixture (Mix) alone and together with increasing doses of the perfluorocarboxylic acid, perfluorononanoic acid (PFNA) as well as the effects of PFNA alone. Mix was composed of 12 environmentally relevant compounds representing typical endocrine disrupting compounds at high-end human exposure levels (Christiansen et al 2012) and two food ingredients in doses corresponding to a high-end daily intake of grape fruit and licorice. We hypothesized that no adverse effects would be observed with Mix alone or combined with the low dose PFNA. By use of two complementary omics methods, metabolomics and transcriptomics, we have profiled the effects in plasma and livers of exposed rats and compared them to adverse effects observed at higher PFNA doses.

## Materials and methods

### Compounds & doses

Compounds were selected as previously described (In preparation), The animals were dosed with a human relevant mixture of 12 environmental compounds known for their endocrine disrupting activities (bisphenol A, butyl paraben, dibutylphthalate (DBP), bis(2-ethylhexyl)phthalate (DEHP), 4-methylbenzylidene camphor (4-MBC), octyl methoxycinnamate (OMC), dichlorodiphenyldichloroethylene (p,p'-DDE), epoxiconazole, linuron, prochloraz, procymidone, vinclozolin) and described elsewhere (Christiansen et al 2012; Hadrup et al 2013) along with the two food components, glabridin from licorice and bergamottin from grape fruit (In preparation). The total dose of the 14-component mixture was 2.5 mg/kg/day and the ratio of the compounds in Mix is presented in Supplementary Table 1.

In addition to Mix, the animals were treated with three doses of PFNA, chosen such that the lowest dose corresponds to a high-end human exposure level (Lau et al 2007). The chosen doses of PFNA were 0.0125 mg/kg/day, 0.25 mg/kg/day, and 5 mg/kg/day.

### Animal study

The animal study has been described previously (In preparation). In brief, we obtained male Wistar Hannover Galas rats at six weeks of age and allowed one week of acclimatization prior to initiation of the study. The animals were housed two per cage with a 12-hour light/dark cycle and ad libitum access to citric acid acidified tap water and standard diet. The animals received test substances (PFNA ± Mix) once daily by gavage for 14 days with corn oil as vehicle. Eighty-two rats were randomly assigned into eight groups: vehicle control (n=10), Mix 2.5 mg/kg/day (n=8), PFNA 0.0125 mg/kg/day + Mix (Low PFNA + Mix) (n=10), PFNA 0.0125 mg/kg/day (Low PFNA) (n=10), PFNA 0.25 mg/kg/day + Mix (Mid PFNA + Mix) (n=8), PFNA 0.25 mg/kg/day (Mid PFNA) (n=8) PFNA 5 mg/kg/day + Mix (High PFNA + Mix) (n=8), and PFNA 5 mg/kg/day (High PFNA) (n=8). The last dose was given to each animal in the range of 1 hr 15 min to 1 hr 45 min before euthanization. The rats were anaesthetized in $CO_2/O_2$ prior to decapitation. Plasma was isolated from heparinized neck blood by centrifugation at 1,000 x g at 4 °C for 10 min and subsequently stored at -80 °C. Livers were weighed and frozen in liquid nitrogen.

### Metabolomics

The applied procedure is described in (In preparation). In brief, phospholipids were adsorbed on a phospholipid SPE column (Supelco, Sigma-Aldrich). The eluate was collected, dried and extracted using first 200 µl heptane to isolate the lipids followed by 200 µl methanol to extract the more polar compounds. The phospholipids were eluted from the SPE column using 300 µl 10 % $NH_4OH$ in methanol. The phospholipid, lipid and the polar fractions were analyzed by an HPLC system combined with a maxis qTime-of-flight mass spectrometer (Bruker Daltonics, Bremen, Germany).

The data were analysed in Profile Analysis 2.1 (Bruker Daltonics, Bremen, Germany). Data was extracted using the "find molecular features" algorithm in a mass range from 50-1100 m/z value. The noise was reduced using R (R Core Team 2012) by removing peaks that were present in < 50% of the samples among all treatment groups and at the same time had a peak intensity of ≤ 3000. The data were then uploaded to metaboanalyst.com (Wishart group, Alberta, Edmonton Canada) and analysed with t-test, principal component analysis, and partial least squares discriminant analysis. The accurate masses of significantly different metabolites were searched for in databases such as human metabolome database (HMDB) (Wishart et al 2009). The identities of the compounds were verified by comparison of MS/MS patterns with data from the databases HMDB, metlin (www.metlin.script.edu)(Smith et al 2005), lipidmaps (www.lipidmaps.org)(Sud et al 2007) and massbank (www.massbank.jp)(Horai et al 2010).

### Statistical analysis

Initial analyses of the metabolome data were performed by upload of the data to the Metaboanalyst server (Xia et al 2012). Here, a one-way analysis of variance (ANOVA) comparing PFNA-treated animals to control animals and PFNA + Mix-treated animals to control, respectively, formed the basis for initial selection of significantly altered metabolites. All *p*-values were corrected using false discovery rate (FDR) according to the protocol implemented in the Metaboanalyst workflow, and 0.05 was used as cut-off for statistical significance. The statistically significantly altered metabolites were subsequently analyzed and plotted using GraphPad Prism version 5.00 for Windows, GraphPad Software, La Jolla, California, USA, www.graphpad.com as follows. The D'Agostino and Pearson omnibus normally test was used to test for normality of the data. If data were normally distributed, an ANOVA was performed using Dunnett's multiple comparisons test to correct the *p*-values. If data were not normally distributed, data were log transformed and, if normally distributed after transformation, analyzed by ANOVA. If not normally distributed, a non-parametric Kruskal-Wallis test followed by a Dunn's multiple comparisons test was conducted. The criteria for statistical significance was $p<0.05$, $p<0.01$ and $p<0.001$ leading to marking *, ** and ***, respectively. Statistical comparisons not applicable to ANOVA tests were carried out using a Student's t-test.

### Transcriptomics

Total RNA from six rat livers from each of the vehicle control, Low PFNA + Mix, and Mid PFNA ± Mix groups were separately converted into labeled cRNA and applied to the One-Color Microarray-Based Gene Expression Analysis (Low Input Quick Amp Labeling) version 6.5, May 2010 (Agilent Technologies, Santa Clara, Ca). Labeled cRNA from each rat was hybridized to Agilent Whole Rat Genome Oligo Microarrays (G4122F) for 17 h at 65°C. The hybridized microarrays were scanned using an Agilent DNA Microarray Scanner and evaluated using the Feature Extraction software version 10.7.3.1 according to protocol GE1_107_Sep09 (Agilent Technologies) to generate feature extraction files for further analysis. Reads were quality controlled by the software prior to release of the data. Arrays that did not pass quality control were removed from the data set. Based on the quality control reports, two of the six microarrays

from the Mid PFNA + Mix group were excluded from further analysis. The remaining arrays, six from each of control, Low PFNA + Mix, and Mid PFNA and the remaining four from Mid PFNA + Mix were found to be of high quality.

Extracted data were analyzed using the limma software package (Smyth 2004; Smyth 2005) in R (R Core Team 2012). Data were background corrected using the "normexp" method (Ritchie et al 2007) and normalized between arrays using quantile normalization (Smyth and Speed 2003) prior to statistical analyses. Within-array replicate probes were replaced with the average expression level. To identify treatment-specific gene effects, we fitted a linear model for each gene and applied empirical Bayes statistics (Smyth 2004) for each relevant two-group comparison. The FDR was controlled using the Benjamini-Hochberg method (Benjamini and Hochberg 1995). Reported *p*-values for the transcriptomics analysis are all corrected, and corrected *p*-values $\leq 0.05$ were considered statistically significant

### Pathway analysis

Transcription data were analyzed through the use of QIAGEN's Ingenuity Pathway Analysis (IPA, QIAGEN Redwood City, www.qiagen.com/ingenuity). IPA was used to map significantly differentially expressed gene (DEG) probes to genes, and expression values were used for prediction of the involvement of the DEGs in functional networks, pathways, and diseases, and for graphical representations. Using the Fisher's Exact Test, we analyzed the overlap between the DEGs in our dataset and genes known to be involved in disease networks, pathways, and diseases available in the Ingenuity Knowledgebase.

## Results and discussion

Since low-dose effects of toxicologically relevant compounds is a fairly new and much discussed topic, little is known about the effects of those low doses on human or animal health, and even less so on the plasma metabolome or the liver gene expression levels. Our results indicate that even short-term exposure (14 days) to a low dose of Mix and a relatively low dose of PFNA ± Mix interferes with lipid homeostasis.

### Pathological investigation

Findings in the livers of the High PFNA ± Mix animals indicated hepatic steatosis – a state of retention of lipids in the liver. No pathological effects were observed in any of the other dosing groups. The specific data on the animals are described elsewhere (In preparation). Since we expected the toxicity in the High PFNA dose groups to disturb the mechanistic investigation of biomarkers, these data were not analyzed in this study.

### Effects of the 14-compound mixture (Mix) alone

The plasma samples and standards were analyzed by HPLC connected to a high-resolution qTOF mass spectrometer, and extracted ion chromatograms of the accurate mass of PFNA ± 2 mDA were created. After 14 days of exposure, the concentration of PFNA in plasma was determined to be 1.1, 30 and 40 µg/ml for Low PFNA + Mix, Mid PFNA + Mix

and High PFNA + Mix, respectively. The PFNA plasma concentrations without Mix were in the range of or up to six times the high-end human exposure levels (Emmett et al 2006; Lau et al 2007; NHANES 2013; In preparation). As described elsewhere (In preparation), the addition of Mix significantly increased the plasma levels of PFNA (2.8-fold) thereby indicating altered ADME (administration, distribution, metabolism, excretion) properties of PFNA when Mix is co-administered. The dose of Mix is calculated from a high-end exposure level in the European population (Christiansen et al 2012) and corrected for body surface area of a rat compared to a human (In preparation). None of the compounds in Mix or their likely metabolites were found when extracted ion chromatograms were created based on their accurate mass. This could be due to rapid metabolism or levels below the limit of detection in the LC-MS analysis.

## Metabolomics

### *Mix alone*

A comparison between animals exposed to Mix alone and controls was conducted. Mix significantly affected 63 molecular features out of 882 molecular features in total. The affected phospholipids were lyso-phosphatidylcholine (lyso-PC)(20:4) and lyso-PC(18:2), while the neutral lipids were diacylglycerols (DG) e.g. DG(18:1/18:3) and DG(18:2/16:0). Representative metabolite profiles across treatment groups are shown in Figure 1; whereas plots of all significantly altered metabolites are shown in Supplementary Figure 1 and Supplementary Figure 2. Animal studies with other compounds (Chen and Kim 2013; Zhang et al 2013; Androutsopoulos et al 2013) have indicated disturbed energy metabolism as a consequence of insults to various environmental compounds. This includes the organochlorine pesticide, *p-p'*-DDE (Androutsopoulos et al 2013), which is a constituent of Mix. Also, complex mixtures such as those of wastewater treatment plant effluents, give rise to altered lipid metabolism (Zhang et al 2014). Despite the significance of the changes in the lipidome upon exposure to compounds, no studies have, to our knowledge, been able to clarify specific mechanisms of action.

### *Low PFNA*

To identify the features causing the most variance in the data from the Low PFNA dose group we performed Partial least square-Discriminant Analyses (PLS-DA) of the lipid fractions. As illustrated in Figure 2 for Low PFNA, the first two components of the analysis separated the Mix-exposed animals from the non-Mix groups and there was a difference between Low PFNA and control. However, it was not possible to identify the metabolites responsible for this difference. We obtained similar PLS-DA plots for the phospholipid and polar fractions (In preparation). There were 30 metabolites responsible for separating Mix-treated animals from non-Mix-treated animals (Supplementary Figure 1). Mix alone altered 22 of these metabolites. For six out of 22 metabolites the concomitant exposure to PFNA enhanced the effect on the metabolites, none of which are identified. The effect of Low PFNA + Mix exposure was decreased concentrations of plasma lipids, primarily DGs and a single phospholipid (Figure 3a-c). The only metabolite directly

affected by Low PFNA but not by Mix was the steroid hormone, corticosterone (Figure 3a). None of the metabolites that were altered in Mix and not in Low PFNA + Mix were identified.

### *Mid PFNA*

For the Mid PFNA dose with or without Mix, we conducted the same analysis as for the Low PFNA dose. We identified 24 metabolites that were significantly changed (Supplementary Figure 2); among these, Mix alone altered 17. Figure 3d-f shows three representative significantly changed metabolites. As seen in Figure 3d there was an increase in the level of lyso-PC(20:4) by Mid PFNA, but not by Mix. Mid PFNA + Mix seemed to affect diacylated phospholipids such as PC(18:0/16:1) (Figure 3e).

### *Effects of Mix compared to PFNA ± Mix*

To investigate the effect of Mix on the PFNA-induced effects, we compared the significant metabolites resulting from statistical analysis of Mix vs. control (t-test) and PFNA + Mix vs. control (ANOVA). Twenty-four metabolites were affected by all treatments and 12 of the 24 metabolites were identified (Supplementary Figure 3). The 12 metabolites consisted of neutral short-chained lipids such as monoacylglycerols (MG) and phospholipids. Mix alone affected the DGs, whereas the PFNA + Mix groups affected the diacylated phospholipids.

Over the last decades, there has been a worldwide increase in disorders related to diet. One disorder in particular is the metabolic syndrome, which increases the risk of developing diabetes and heart failure. Studies have revealed decreased phospholipid levels in diabetics (Wang et al 2005; Liu et al 2013). In particular, changes in the ratio between lyso-PC and PC concentrations indicates a systemic change potentially related to diabetes (Altmaier et al 2008). The PCs are involved in the biosynthesis of multiple compounds in the endoplasmic reticulum. Specifically, decreased PC levels reflects an increased biosynthesis of lipids such as triacylglycerols (Lagace and Ridgway 2013). The changes in the metabolome obtained from exposure to Mix alone and PFNA + Mix are similar to the differences observed in patients with metabolic syndrome as compared to healthy subjects (Ferrannini et al 2013); however, data from Chen *et al.* show that the lyso-PCs are increased in metabolic syndrome patients compared to healthy controls (Chen et al 2011). This corresponds to findings observed for Mid PFNA alone but not for Mix and PFNA + Mix.

A cholesterol derivative was found significantly decreased by Mix, Low PFNA + Mix and Mid PFNA + Mix, though from the accurate mass (429.380 m/z value) an unambiguous metabolite was not identified. Three possible cholesterol metabolites matching the accurate mass were suggested; 4α-Formyl-4β-methyl-5α-cholesta-8-en-3β-ol, 4α-Hydroxymethyl-4β-methyl-5α-cholesta-8,24-dien-3beta-ol, and cholesteryl acetate. We analyzed a standard of cholesteryl acetate and based on the MS/MS pattern and retention time this metabolite was excluded from the list. The retention time and the accurate mass indicate that the cholesterol derivative could be either of the other two, but definitive identification was not possible.

### Transcriptomics

#### *Low*

Statistical analysis of the transcriptome of Low PFNA + Mix vs. control resulted in no differentially expressed genes (DEGs) (corrected $p \leq 0.05$); however, at a borderline significance level (corrected $p$-value $\leq 0.1$), 31 unique genes were differentially expressed (Supplementary Table 2). The pathway analysis of these genes indicated pathways related to the immune system to be affected by the compound treatment, mainly due to a slight downregulation of the major histocompatibility complex II (RT1-Ba) ($\log_2$(fold change)=-0.244, $p$-value = 0.098). Other borderline DEGs are reported in Supplementary table 3.

#### *Mid*

For Mid PFNA ± Mix groups the significantly DEGs include the set of annotated genes contained in the Ingenuity Pathway Analysis database and exclude most expressed sequence tags and genes without adequate literature-based information for use in the analysis. Statistical analysis of the Mid PFNA ± Mix groups resulted in 206 and 182 DEGs, with or without Mix, respectively (Supplementary Table 2).

In the PFNA ± Mix dosing groups, fatty acid metabolism was suggested as the main biochemical function affected by the DEGs (Table 1). This included upregulation of genes involved in peroxisomal fatty acid β-oxidation (such as Crot, Crat, Acox1, Ehhadh, Hadha, Hadhb, Decr2, Eci2, Ech1), and mitochondrial β-oxidation (Cpt2, Slc25a20, Acad11, Acadl, Acadm, Acads, Acadvl). Additionally, genes associated with lipid transport, fatty acid activation, and peroxisomal transport (Apoa2, Abcd3, Cd36, Slc27a2) were upregulated. Fatty acid binding protein 5 (Fabp5) was downregulated. Similar effects of perfluorinated compounds on lipid homeostasis have been presented previously (Guruge et al 2006; Rosen et al 2007; Fang et al 2012a; Fang et al 2012b; Fang et al 2012c).

Whereas the effects on lipid metabolism are very similar for Mid PFNA and Mid PFNA + Mix, less pronounced profiles are seen on the aerobic respiration for Mid PFNA + Mix compared to Mid PFNA (Table 1). The citric acid cycle enzymes, aconitase 2 (Aco2), isocitrate dehydrogenase 3 (NAD+) beta (Idh3b), and succinate-CoA ligase (Sucla2 and Suclg1) were significantly upregulated by Mid PFNA, whereas this was only the case for Aco2 and Suclg1 for Mid PFNA + Mix. Similarly, for enzymes in the electron transport chain (Ndufa10, Ndufab1, Ndufs3 (complex I), and Sdhb (complex II)) only Ndufa10 is significantly upregulated in both groups. The other enzymes are upregulated by Mid PFNA, exclusively. This could indicate, that constituent(s) of Mix counteract the effects imposed by PFNA alone. Expression of the genes encoding enzymes involved in glucose metabolism disorders were mainly disrupted by Mid PFNA + Mix, except for a few genes such as phosphofructokinase that was only significant in Mid PFNA and carbonic anhydrase VII, Heat-shock protein 40, plus STEAP family member 4 that were differentially expressed by both groups. The profiles of Mid PFNA ± Mix are thus very similar and only subtle differences in the exact DEGs cause the

differences in z-scores and p-values. Additionally, a statistical comparison between Mid PFNA and Mid PFNA + Mix did not yield any significant DEGs.

IPA has a function to predict the activity of potential upstream regulators based on the DEGs. With our dataset, several transcription factors and other central proteins were predicted activated (Table 2). Downstream of the DEGs, IPA predicts regulatory effects, such as diseases and functions likely to be perturbed due to the changes in gene expression. As illustrated in Figure 4, activation of the upstream regulators PPARα, PPARγ, PPARδ, RXRα, PPARgc1a, and MED1 results in gene expression changes (Table 1) that subsequently leads to increased oxidation of lipids and in particular fatty acids. This, in turn, protects against accumulation of lipids in the liver and therefore also against hepatic steatosis. The PPARs are known regulators of fatty acid β-oxidation (Kanehisa and Goto 2000). Increased β-oxidation results in decreased plasma lipid concentration (Lau 2012). Furthermore, perfluorinated alkyl acids are known PPAR-activators (Vanden Heuvel et al 2006; Lau 2012; U. S. Environmental Protection Agency 2014). The Mix contains the two phthalates DBP and DEHP, which are both known PPAR activators (Desvergne et al 2009). Several of the other compounds in Mix also activate PPAR (U. S. Environmental Protection Agency 2014).

The total lists of DEGs for the three treatment groups are reported in Supplementary Table 2, and a summary of the functions of the Mid PFNA ± Mix is given in Supplementary Table 4. The effects on lipid homeostasis observed for Mix with and without PFNA could be ascribed to the presence of glabridin. It is a naturally occurring isoflavonoid from licorice that has been shown to inhibit the activity of the cytochrome P450 (CYP) CYP2C9 in the liver (Kent et al 2002). Studies have shown that CYP2C9 is involved in the regulation of lipid metabolism in the body (Kent et al 2002; Kirchheiner et al 2003). Furthermore, people with a polymorphism in CYP2C9 have a tendency to have lower levels of the lipid carriers LDL and HDL (Kirchheiner et al 2003). Thus, a decrease in LDL and HDL may result in decreased levels of plasma lipids, including cholesterol and cholesterol derivatives, as these are carrying lipids around the body (Kent et al 2002). However, our data does not suggest downregulation of the rat CYP2C9 homologue, cyp2c11. On the other hand, the HDL-constituent, apolipoprotein A-II (encoded by apoa2) is upregulated by PFNA but not by PFNA + Mix. The effects seen for Mix and PFNA + Mix might therefore be caused by a combination of decreased CYP2C9 activity and PPAR activation.

## Conclusion

We profiled the plasma metabolome and the liver transcriptome in rats after exposure to environmental chemicals. As the organism has a diverse biochemical response, it is a challenge to determine the exact mechanism of action for the 14-chemical Mix. All Mix-containing treatments significantly affected the plasma metabolome, whereas changes in the liver transcriptome became evident at a higher PFNA dose with or without Mix. Therefore, the hepatic effects seemed

to be driven by PFNA exposure, whereas Mix drove the effects on the plasma metabolome. Within Mix, two CYP inhibitors were present. These might be part of the cause of the changes in lipid homeostasis observed in the plasma metabolome. Based on these results, we suggest that PPAR activation and/or CYP2C9 inhibition might explain the observed changes in the metabolome. The study verifies that even at low dose levels, a chemical mixture can affect the metabolome and cause disturbed lipid homeostasis.

## Acknowledgements

A special thanks to Mike Wilson, University of Alberta, for assisting in conducting the R script for the metabolomics matrix reduction. Many thanks to Dr. Christoffer Clemmensen, Helmholtz Zentrum Munich, for fruitful discussions.

## Compliance with ethical requirements

Conflict of interest: Kasper Skov, Kristine Kongsbak, Niels Hadrup, Henrik Lauritz Frandsen, Jørn Smedsgaard, Karine Audouze, Aron Charles Eklund, and Anne Marie Vinggaard declare that they have no conflict of interest.

All applicable institutional guidelines for the care and use of animals were followed. All procedures performed in the studies involving animals were in accordance with the ethical standards of the institution at which the studies were conducted. The animal study was approved by the Danish Animal Experiments Inspectorate. The authorization number given is 2012/561-188. The National Food Institute's in-house Animal Welfare Committee for animal care and use supervised the experiments.

## References

Altmaier E, Ramsay SL, Graber A, et al (2008) Bioinformatics analysis of targeted metabolomics--uncovering old and new tales of diabetic mice under medication. Endocrinology 149:3478–89. doi: 10.1210/en.2007-1747

Androutsopoulos VP, Hernandez AF, Liesivuori J, Tsatsakis AM (2013) A mechanistic overview of health associated effects of low levels of organochlorine and organophosphorous pesticides. Toxicology 307:89–94. doi: 10.1016/j.tox.2012.09.011

Axelstad M, Christiansen S, Boberg J, et al (2014) Mixtures of endocrine-disrupting contaminants induce adverse developmental effects in preweaning rats. Reproduction 147:489–501. doi: 10.1530/REP-13-0447

Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B 57:289 – 300. doi: 10.2307/2346101

Calafat AM, Wong L-Y, Kuklenyik Z, et al (2007) Polyfluoroalkyl chemicals in the U.S. population: data from the National Health and Nutrition Examination Survey (NHANES) 2003-2004 and comparisons with NHANES 1999-2000. Environ Health Perspect 115:1596–602. doi: 10.1289/ehp.10598

Chen C, Kim S (2013) LC-MS-based Metabolomics of Xenobiotic-induced Toxicities. Comput Struct Biotechnol J 4:e201301008. doi: 10.5936/csbj.201301008

Chen M, Zhou K, Chen X, et al (2014) Metabolomic analysis reveals metabolic changes caused by bisphenol A in rats. Toxicol Sci 138:256–67. doi: 10.1093/toxsci/kfu016

Chen S, Chu Y, Zhao X, et al (2011) HPLC-MS-Based Metabonomics Reveals Disordered Lipid Metabolism in Patients with Metabolic Syndrome. J Anal Sci Technol 2:A173–A178. doi: 10.5355/JAST.2011.A173

Christiansen S, Kortenkamp A, Axelstad M, et al (2012) Mixtures of endocrine disrupting contaminants modelled on human high end exposures: an exploratory study in rats. Int J Androl 35:303–316. doi: 10.1111/j.1365-2605.2011.01242.x

Christiansen S, Scholze M, Axelstad M, et al (2008) Combined exposure to anti-androgens causes markedly increased frequencies of hypospadias in the rat. Int J Androl 31:241–248. doi: 10.1111/j.1365-2605.2008.00866.x

Desvergne B, Feige JN, Casals-Casas C (2009) PPAR-mediated activity of phthalates: A link to the obesity epidemic? Mol Cell Endocrinol 304:43–8. doi: 10.1016/j.mce.2009.02.017

Emmett EA, Zhang H, Shofer FS, et al (2006) Community exposure to perfluorooctanoate: relationships between serum levels and certain health parameters. J Occup Environ Med 48:771–9. doi: 10.1097/01.jom.0000233380.13087.37

Fang X, Gao G, Xue H, et al (2012a) Exposure of perfluorononanoic acid suppresses the hepatic insulin signal pathway and increases serum glucose in rats. Toxicology 294:109–15. doi: 10.1016/j.tox.2012.02.008

Fang X, Gao G, Xue H, et al (2012b) In vitro and in vivo studies of the toxic effects of perfluorononanoic acid on rat hepatocytes and Kupffer cells. Environ Toxicol Pharmacol 34:484–94. doi: 10.1016/j.etap.2012.06.011

Fang X, Zou S, Zhao Y, et al (2012c) Kupffer cells suppress perfluorononanoic acid-induced hepatic peroxisome proliferator-activated receptor α expression by releasing cytokines. Arch Toxicol 86:1515–25. doi: 10.1007/s00204-012-0877-4

Ferrannini E, Natali A, Camastra S, et al (2013) Early metabolic markers of the development of dysglycemia and type 2 diabetes and their physiological significance. Diabetes 62:1730–7. doi: 10.2337/db12-0707

Guruge KS, Yeung LWY, Yamanaka N, et al (2006) Gene expression profiles in rat liver treated with perfluorooctanoic acid (PFOA). Toxicol Sci 89:93–107. doi: 10.1093/toxsci/kfj011

Hadrup N, Taxvig C, Pedersen M, et al (2013) Concentration addition, independent action and generalized concentration addition models for mixture effect prediction of sex hormone synthesis in vitro. PLoS One 8:e70490. doi: 10.1371/journal.pone.0070490

Horai H, Arita M, Kanaya S, et al (2010) MassBank: a public repository for sharing mass spectral data for life sciences. J Mass Spectrom 45:703–14. doi: 10.1002/jms.1777

Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28:27–30.

Kent UM, Aviram M, Rosenblat M, Hollenberg PF (2002) The licorice root derived isoflavan glabridin inhibits the activities of human cytochrome P450S 3A4, 2B6, and 2C9. Drug Metab Dispos 30:709–15.

Kirchheiner J, Kudlicz D, Meisel C, et al (2003) Influence of CYP2C9 polymorphisms on the pharmacokinetics and cholesterol-lowering activity of (-)-3S,5R-fluvastatin and (+)-3R,5S-fluvastatin in healthy volunteers. Clin Pharmacol Ther 74:186–94. doi: 10.1016/S0009-9236(03)00121-8

Krysiak-Baltyn K, Toppari J, Skakkebaek NE, et al (2012) Association between chemical pattern in breast milk and congenital cryptorchidism: modelling of complex human exposures. Int J Androl 35:294–302. doi: 10.1111/j.1365-2605.2012.01268.x

Lagace TA, Ridgway ND (2013) The role of phospholipids in the biological activity and structure of the endoplasmic reticulum. Biochim Biophys Acta 1833:2499–510. doi: 10.1016/j.bbamcr.2013.05.018

Lau C (2012) Perfluorinated compounds. EXS 101:47–86. doi: 10.1007/978-3-7643-8340-4_3

Lau C, Anitole K, Hodes C, et al (2007) Perfluoroalkyl acids: a review of monitoring and toxicological findings. Toxicol Sci 99:366–94. doi: 10.1093/toxsci/kfm128

Liu L, Wang M, Yang X, et al (2013) Fasting serum lipid and dehydroepiandrosterone sulfate as important metabolites for detecting isolated postchallenge diabetes: serum metabolomics via ultra-high-performance LC-MS. Clin Chem 59:1338–48. doi: 10.1373/clinchem.2012.200527

Monosson E (2005) Chemical mixtures: considering the evolution of toxicology and chemical assessment. Environ Health Perspect 113:383–390.

Moser VC, Simmons JE, Gennings C (2006) Neurotoxicological interactions of a five-pesticide mixture in preweanling rats. Toxicol Sci 92:235–45. doi: 10.1093/toxsci/kfj189

NHANES (2013) Fourth National Report on Human Exposure to Environmental Chemicals Updated Tables, March, 2013.

R Core Team (2012) R: A Language and Environment for Statistical Computing. R Found Stat Comput. doi: 10.1007/978-3-540-74686-7

Ritchie ME, Silver J, Oshlack A, et al (2007) A comparison of background correction methods for two-colour microarrays. Bioinformatics 23:2700–7. doi: 10.1093/bioinformatics/btm412

Rosen MB, Thibodeaux JR, Wood CR, et al (2007) Gene expression profiling in the lung and liver of PFOA-exposed mouse fetuses. Toxicology 239:15–33. doi: 10.1016/j.tox.2007.06.095

Silva E, Rajapakse N, Kortenkamp A (2002) Something from “nothing”--eight weak estrogenic chemicals combined at concentrations below NOECs produce significant mixture effects. Environ Sci Technol 36:1751–1756.

Smith C a, O’Maille G, Want EJ, et al (2005) METLIN: a metabolite mass spectral database. Ther Drug Monit 27:747–51.

Smyth GK (2005) limma: Linear Models for Microarray Data. Bioinforma. Comput. Biol. Solut. Using R Bioconductor. pp 397–420

Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 3:Article3. doi: 10.2202/1544-6115.1027

Smyth GK, Speed T (2003) Normalization of cDNA microarray data. Methods 31:265–273. doi: 10.1016/S1046-2023(03)00155-5

Sud M, Fahy E, Cotter D, et al (2007) LMSD: LIPID MAPS structure database. Nucleic Acids Res 35:D527–32. doi: 10.1093/nar/gkl838

Taylor KW, Hoffman K, Thayer K a, Daniels JL (2014) Polyfluoroalkyl Chemicals and Menopause among Women 20-65 Years of Age (NHANES). Environ Health Perspect 122:145–50. doi: 10.1289/ehp.1306707

U. S. Environmental Protection Agency (2014) Interactive Chemical Safety for Sustainability (iCSS) Dashboard. http://actor.epa.gov/dashboard/.

Vanden Heuvel JP, Thompson JT, Frame SR, Gillies PJ (2006) Differential activation of nuclear receptors by perfluorinated fatty acid analogs and natural fatty acids: a comparison of human, mouse, and rat peroxisome proliferator-activated receptor-alpha, -beta, and -gamma, liver X receptor-beta, and retinoid X rec. Toxicol Sci 92:476–89. doi: 10.1093/toxsci/kfl014

Wang C, Kong H, Guan Y, et al (2005) Plasma phospholipid metabolic profiling and biomarkers of type 2 diabetes mellitus based on high-performance liquid chromatography/electrospray mass spectrometry and multivariate statistical analysis. Anal Chem 77:4108–16. doi: 10.1021/ac0481001

Wishart DS, Knox C, Guo AC, et al (2009) HMDB: a knowledgebase for the human metabolome. Nucleic Acids Res 37:D603–10. doi: 10.1093/nar/gkn810

Xia J, Mandal R, Sinelnikov I V, et al (2012) MetaboAnalyst 2.0--a comprehensive server for metabolomic data analysis. Nucleic Acids Res 40:W127–33. doi: 10.1093/nar/gks374

Zhang W, Liu Y, Zhang H, Dai J (2012) Proteomic analysis of male zebrafish livers chronically exposed to perfluorononanoic acid. Environ Int 42:20–30. doi: 10.1016/j.envint.2011.03.002

Zhang Y, Deng Y, Zhao Y, Ren H (2014) Using combined bio-omics methods to evaluate the complicated toxic effects of mixed chemical wastewater and its treated effluent. J Hazard Mater 272:52–8. doi: 10.1016/j.jhazmat.2014.02.041

Zhang Y, Zhang Z, Zhao Y, et al (2013) Identifying health effects of exposure to trichloroacetamide using transcriptomics and metabonomics in mice ( Mus musculus ). Environ Sci Technol 47:2918–24. doi: 10.1021/es3048976

**Table 1. Energy homeostasis genes significantly altered by Mid PFNA + Mix. Grey fields denote non-significant changes.**

| Symbol | Entrez ID | Entrez Gene Name | log2(fold change) | | | Biological process |
|---|---|---|---|---|---|---|
| | | | Low PFNA + Mix | Mid PFNA | Mid PFNA + Mix | |
| Aco1 | 50655 | aconitase 1, soluble | 0,12 | 0,337 | 0,544 | Aerobic metabolism |
| Aco2 | 79250 | aconitase 2, mitochondrial | -0,065 | 0,428 | 0,448 | Aerobic metabolism |
| Cox11 | 690300 | cytochrome c oxidase assembly homolog 11 (yeast) | 0,24 | 0,593 | 0,235 | Aerobic metabolism |
| Idh3b | 94173 | isocitrate dehydrogenase 3 (NAD+) beta | 0,003 | 0,325 | 0,247 | Aerobic metabolism |
| Ndufa10 | 678759\| 316632 | NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 10, 42kDa | 0,042 | 0,3 | 0,275 | Aerobic metabolism |
| Ndufab1 | 293453 | NADH dehydrogenase (ubiquinone) 1, alpha/beta subcomplex, 1, 8kDa | 0,051 | 0,269 | 0,116 | Aerobic metabolism |
| Ndufs3 | 295923 | NADH dehydrogenase (ubiquinone) Fe-S protein 3, 30kDa (NADH-coenzyme Q reductase) | 0,085 | 0,336 | 0,28 | Aerobic metabolism |
| Sdhb | 298596 | succinate dehydrogenase complex, subunit B, iron sulfur (Ip) | 0,043 | 0,244 | 0,17 | Aerobic metabolism |
| Sucla2 | 361071 | succinate-CoA ligase, ADP-forming, beta subunit | 0,145 | 0,43 | 0,357 | Aerobic metabolism |
| Suclg1 | 114597 | succinate-CoA ligase, alpha subunit | 0,024 | 0,275 | 0,284 | Aerobic metabolism |
| Acaa1 | 501072\| 24157 | acetyl-CoA acyltransferase 1 | 0,113 | 1,008 | 0,985 | Lipid metabolism |
| Acaa2 | 170465 | acetyl-CoA acyltransferase 2 | 0,394 | 0,556 | 0,969 | Lipid metabolism |
| Acot1 | 314304 | acyl-CoA thioesterase 1 | 0,386 | 0,984 | 1,427 | Lipid metabolism |
| Acot12 | 170570 | acyl-CoA thioesterase 12 | 0,245 | 0,557 | 0,666 | Lipid metabolism |
| Acot2 | 192272 | acyl-CoA thioesterase 2 | 0,241 | 1,312 | 1,497 | Lipid metabolism |
| Acot4 | 681337 | acyl-CoA thioesterase 4 | 0,264 | 1,115 | 1,585 | Lipid metabolism |
| Acot7 | 26759 | acyl-CoA thioesterase 7 | 0,019 | 0,431 | 0,593 | Lipid metabolism |
| Aldh1a1 | 24188 | aldehyde dehydrogenase 1 family, member A1 | 0,377 | 2,077 | 2,523 | Lipid metabolism |
| Cyp2b6 | 361523\| 24300 | cytochrome P450, family 2, subfamily B, polypeptide 6 | 0,971* | 2,12 | 2,561 | Lipid metabolism |
| Cyp2c19 | 293989 | cytochrome P450, family 2, subfamily C, polypeptide 19 | 0,128 | 0,686 | 0,756 | Lipid metabolism |
| Cyp2j2 | 65210 | cytochrome P450, family 2, subfamily J, polypeptide 2 | 0,237 | 1,018 | 1,187 | Lipid metabolism |
| Cyp4a11 | 50549 | cytochrome P450, family 4, subfamily A, polypeptide 11 | 0,643 | 2,419 | 2,568 | Lipid metabolism |
| Cyp4a14 | 298423\| 24306 | cytochrome P450, family 4, subfamily a, polypeptide 14 | 0,333 | 1,201 | 1,335 | Lipid metabolism |
| Decr1 | 117543 | 2,4-dienoyl CoA reductase 1, mitochondrial | 0,234 | 1,364 | 1,344 | Lipid metabolism |
| Ephx2 | 65030 | epoxide hydrolase 2, cytoplasmic | 0,243 | 1,201 | 1,267 | Lipid metabolism |
| Gcdh | 364975 | glutaryl-CoA dehydrogenase | 0,175 | 0,576 | 0,638 | Lipid metabolism |
| Hadh | 113965 | hydroxyacyl-CoA dehydrogenase | 0,056 | 0,418 | 0,4 | Lipid metabolism |
| Hadha | 170670 | hydroxyacyl-CoA dehydrogenase/3-ketoacyl-CoA thiolase/enoyl-CoA hydratase (trifunctional protein), alpha subunit | 0,144 | 0,659 | 0,746 | Lipid metabolism |
| Hadhb | 171155 | hydroxyacyl-CoA dehydrogenase/3-ketoacyl-CoA thiolase/enoyl-CoA hydratase (trifunctional protein), beta subunit | 0,194 | 0,8 | 0,817 | Lipid metabolism |
| Hsd11b1 | 25116 | hydroxysteroid (11-beta) dehydrogenase 1 | -0,261 | -0,446 | -1,082 | Lipid metabolism |
| Immt | 312444 | inner membrane protein, mitochondrial | -0,011 | 0,345 | 0,376 | Lipid metabolism |
| Mlycd | 85239 | malonyl-CoA decarboxylase | 0,093 | 0,316 | 0,325 | Lipid metabolism |
| Ncam1 | 24586 | neural cell adhesion molecule 1 | -0,163 | -0,05 | -0,27 | Lipid metabolism |

| | | | | | | |
|---|---|---|---|---|---|---|
| Phgdh | 58835 | phosphoglycerate dehydrogenase | -0,308 | -1,271 | -0,976 | Lipid metabolism |
| Plin5 | 501283 | perilipin 5 | 0,198 | 0,595 | 0,61 | Lipid metabolism |
| Ptk2b | 50646 | protein tyrosine kinase 2 beta | 0,175 | 0,607 | 0,47 | Lipid metabolism |
| Sirt4 | 304539 | sirtuin 4 | 0,162 | 0,291 | 0,461 | Lipid metabolism |
| Acad11 | 315973 | acyl-CoA dehydrogenase family, member 11 | 0,115 | 0,62 | 0,656 | Lipid metabolism - mitochondrial b-oxidation |
| Acadl | 25287 | acyl-CoA dehydrogenase, long chain | 0,05 | 0,442 | 0,419 | Lipid metabolism - mitochondrial b-oxidation |
| Acadm | 24158 | acyl-CoA dehydrogenase, C-4 to C-12 straight chain | 0,15 | 0,47 | 0,47 | Lipid metabolism - mitochondrial b-oxidation |
| Acads | 64304 | acyl-CoA dehydrogenase, C-2 to C-3 short chain | 0,056 | 0,326 | 0,416 | Lipid metabolism - mitochondrial b-oxidation |
| Acadvl | 25363 | acyl-CoA dehydrogenase, very long chain | 0,047 | 0,413 | 0,482 | Lipid metabolism - mitochondrial b-oxidation |
| Cpt2 | 25413 | carnitine palmitoyltransferase 2 | 0,163 | 0,565 | 0,786 | Lipid metabolism - mitochondrial b-oxidation |
| Slc25a20 | 117035 | solute carrier family 25 (carnitine/acylcarnitine translocase), member 20 | 0,098 | 0,668 | 0,692 | Lipid metabolism - mitochondrial b-oxidation |
| Acox1 | 50681 | acyl-CoA oxidase 1, palmitoyl | 0,074 | 0,601 | 0,745 | Lipid metabolism - proxisomal b-oxidation |
| Crat | 311849 | carnitine O-acetyltransferase | 0,372 | 1,55 | 1,71 | Lipid metabolism - proxisomal b-oxidation |
| Crot | 83842 | carnitine O-octanoyltransferase | 0,19 | 1,115 | 1,219 | Lipid metabolism - proxisomal b-oxidation |
| Decr2 | 64461 | 2,4-dienoyl CoA reductase 2, peroxisomal | 0,275 | 0,831 | 0,944 | Lipid metabolism - proxisomal b-oxidation |
| Ech1 | 64526 | enoyl CoA hydratase 1, peroxisomal | 0,459 | 1,503 | 1,856 | Lipid metabolism - proxisomal b-oxidation |
| Eci2 | 291075 | enoyl-CoA delta isomerase 2 | 0,117 | 0,577 | 0,548 | Lipid metabolism - proxisomal b-oxidation |
| Ehhadh | 171142 | enoyl-CoA, hydratase/3-hydroxyacyl CoA dehydrogenase | 0,184 | 1,702 | 1,86 | Lipid metabolism - proxisomal b-oxidation |
| Hsd17b4 | 79244 | hydroxysteroid (17-beta) dehydrogenase 4 | 0,113 | 0,295 | 0,278 | Lipid metabolism - proxisomal b-oxidation |
| Apoa2 | 25649 | apolipoprotein A-II | 0,321 | 0,504 | 0,474 | Lipid metabolism - transport |
| Abcd3 | 25270 | ATP-binding cassette, sub-family D (ALD), member 3 | 0,093 | 0,729 | 0,685 | Lipid metabolism - transport |
| Cd36 | 29184 | CD36 molecule (thrombospondin receptor) | 0,158 | 1,276 | 1,867 | Lipid metabolism - transport |
| Fabp5 | 140868 | Lipid binding protein 5 (psoriasis-associated) | -0,385 | -1,392 | -1,508 | Lipid metabolism - transport |
| Slc27a2 | 65192 | solute carrier family 27 (Lipid transporter), member 2 | 0,215 | 0,543 | 0,57 | Lipid metabolism - transport |
| Fbp2 | 114508 | fructose-1,6-bisphosphatase 2 | 0,272 | 0,642 | 1,081 | Glucose metabolism disorders |
| Pfkm | 65152 | phosphofructokinase, muscle | -0,388 | -0,487 | -0,35 | Glucose metabolism disorders |
| Adrb3 | 25645 | adrenoceptor beta 3 | -0,279 | -0,134 | -0,366 | Glucose metabolism disorders |
| Ca1 | 310218 | carbonic anhydrase I | -0,003 | -0,78 | -1,132 | Glucose metabolism disorders |
| Ca4 | 29242 | carbonic anhydrase IV | -0,266* | -0,009 | -0,236 | Glucose metabolism disorders |
| Ca7 | 291819 | carbonic anhydrase VII | -0,06 | 0,817 | 0,643 | Glucose metabolism disorders |
| Calcr | 116506 | calcitonin receptor | 0,225 | 0,114 | 0,392 | Glucose metabolism disorders |

| | | | | | | |
|---|---|---|---|---|---|---|
| Dnajc3 | 63880 | DnaJ (Hsp40) homolog, subfamily C, member 3 | -0,113 | -0,279 | -0,362 | Glucose metabolism disorders |
| Hdc | 24443 | histidine decarboxylase | 0,267 | 1,03 | 1,309 | Glucose metabolism disorders |
| Hipk1 | 365895 | homeodomain interacting protein kinase 1 | 0,189 | 0,011 | 0,296 | Glucose metabolism disorders |
| Rt1-Ba | 309621 | major histocompatibility complex, class II, DQ alpha 1 | -0,244* | -0,139 | -0,302 | Glucose metabolism disorders |
| Lef1 | 161452 | lymphoid enhancer-binding factor 1 | 0,1 | 0,36 | 0,573 | Glucose metabolism disorders |
| Mgat2 | 94273 | mannosyl (alpha-1,6-)-glycoprotein beta-1,2-N-acetylglucosaminyltransferase | -0,135 | -0,436 | -0,544 | Glucose metabolism disorders |
| Prox1 | 305066 | prospero homeobox 1 | -0,09 | -0,114 | -0,336 | Glucose metabolism disorders |
| Psmc3 | 29677 | proteasome (prosome, macropain) 26S subunit, ATPase, 3 | 0,102 | 0,287 | 0,326 | Glucose metabolism disorders |
| Psmd8 | 292766 | proteasome (prosome, macropain) 26S subunit, non-ATPase, 8 | 0,103 | 0,179 | 0,238 | Glucose metabolism disorders |
| Ramp1 | 58965 | receptor (G protein-coupled) activity modifying protein 1 | -0,133 | -0,213 | -0,299 | Glucose metabolism disorders |
| Steap4 | 499991 | STEAP family member 4 | -0,15 | -0,332 | -0,404 | Glucose metabolism disorders |
| Hmgcs2 | 24450 | 3-hydroxy-3-methylglutaryl-CoA synthase 2 (mitochondrial) | 0,118 | 0,469 | 0,471 | Ketogenesis |
| Pex11a | 85249 | peroxisomal biogenesis factor 11 alpha | 0,209 | 1,368 | 1,387 | Peroxisome biogenesis |
| Pex16 | 311203 | peroxisomal biogenesis factor 16 | 0,208 | 0,405 | 0,439 | Peroxisome biogenesis |
| Pex5 | 312703 | peroxisomal biogenesis factor 5 | 0,077 | 0,326 | 0,284 | Peroxisome biogenesis |
| Dhrs4 | 266686 | dehydrogenase/reductase (SDR family) member 4 | 0,005 | 0,363 | 0,294 | Retinol metabolism |

*significant at $p \leq 0.1$

**Table 2. Main upstream regulators predicted to regulate effects caused by Mid PFNA ± Mix. An arrow indicating up- or downregulation marks upstream regulators that are differentially expressed genes (DEG). Negative z-scores indicate that the upstream regulator is inactivated, whereas positive z-scores indicate activation.**

| Upstream regulator | Mid PFNA + Mix | | Mid PFNA | | Genes |
|---|---|---|---|---|---|
| | z-score | p-value | z-score | p-value | |
| PPARα<br>Peroxisome proliferator-activated receptor α | 5.3 | 2e-24 | 5.1 | 5e-28 | Abcd3, Acaa1, Acaa2*, Acadl, Acadm, Acads, Acadvl, Acat1, Acot1*, Acot2, Acox1, Adtrp, Cd36, Cpt2, Crot, Cyp2b6, Cyp4a11, Cyp4a14, Decr1, Decr2, Ech1, Eci2, Ehhadh, Fabp5, Gpd1, Hadh, Hadha, Hadhb, Rt1-Ba*, Hmgcs2, Mlycd, Pex11a, Plin5, Slc25a20, Slc27a2, Vnn1, Aadac+, Apoa2+, Cfh+, Ftcd+, H2afz+ |
| KLF15<br>Krüppel-like factor 5 | 3.2 | 3e-18 | 3.0 | 5e-17 | Acadl, Acadm, Acadvl, Acot1*, Acox1, Cd36, Cpt2, Decr1, Ehhadh, Fabp5, Hadha, Hadhb, Mlycd, Slc25a20 |
| ACOX1 ↑<br>Peroxisomal acyl-coenzyme A oxidase 1 | -2.4 | 5e-15 | -2.474 | 3e-13 | Abcd3, Acaa1, Acadl, Acadm, Acadvl, Acot2, Acox1, Aig1, Cd36, Crat, Cyp4a11, Cyp4a14, Ehhadh, Rt1-Ba*, Hsd11b1*, Pex11a, Slc27a2, Tnfrsf10a*, Hspa5+ |
| PPARγ<br>Peroxisome proliferator-activated receptor α | 3.9 | 8e-14 | 3.3 | 1e-13 | Acaa1, Acaa2*, Acadl, Acadm, Acads, Acox1, Adrb3*, Cd36, Cpt2, Crat, Cyp4a11, Cyp4a14, Ehhadh, Fabp5, Fbp2*, Gpd1, Hadha, Hadhb, Hmgcs2, Mlycd, Pepd, Pex11a, Plin5, Slc25a20, Vnn1, Apoa2+, Sdc1+ |
| EHHADH ↑<br>Enoyl-Coenzyme A, Hydratase/3-Hydroxyacyl Coenzyme A Dehydrogenase | -2.8 | 1e-12 | -2.8 | 6e-13 | Abcd3, Acaa1, Acot2, Acox1, Cd36, Cyp4a11, Cyp4a14, Pex11a |
| HSD17B4 ↑<br>Peroxisomal multifunctional enzyme type 2 | -3.0 | 1e-12 | -2.8 | 4e-11 | Abcd3, Acaa1, Acot2, Acox1, Cd36, Cyp4a11, Cyp4a14, Hoxd13*, Pex11a |

*DEG only present in Mid PFNA + Mix
+DEG only present in Mid PFNA

## Figures

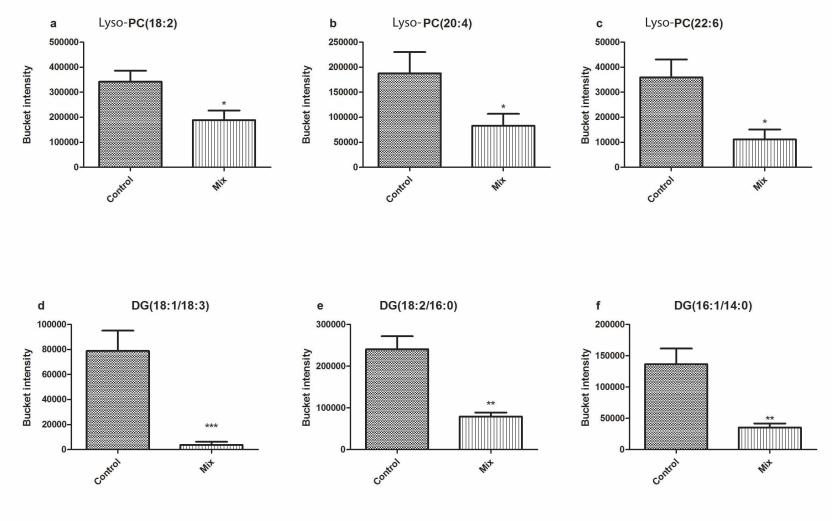



**Fig.1 The six most significantly different metabolites in the Mix group compared to the control group. In total, 63 metabolites were found to be significantly different between the two groups (t-test). The plotted metabolites belong to two metabolite classes, diglycerides (DG) and lyso-phosphotidylcholines (lyso-PC)**

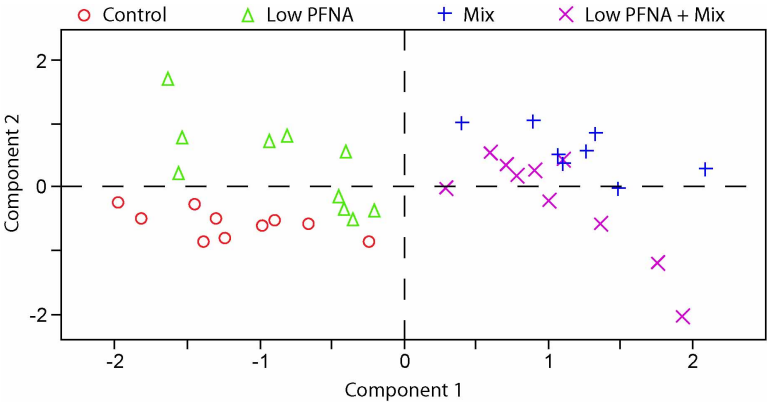



**Fig. 2 PLS-DA plot of the heptane fraction correlating the four groups Control (n=10), Low PFNA (n=10), Low PFNA + Mix (n=10) and Mix (n=8). The plot illustrates that Mix accounts for a major part of the variance in the dataset**

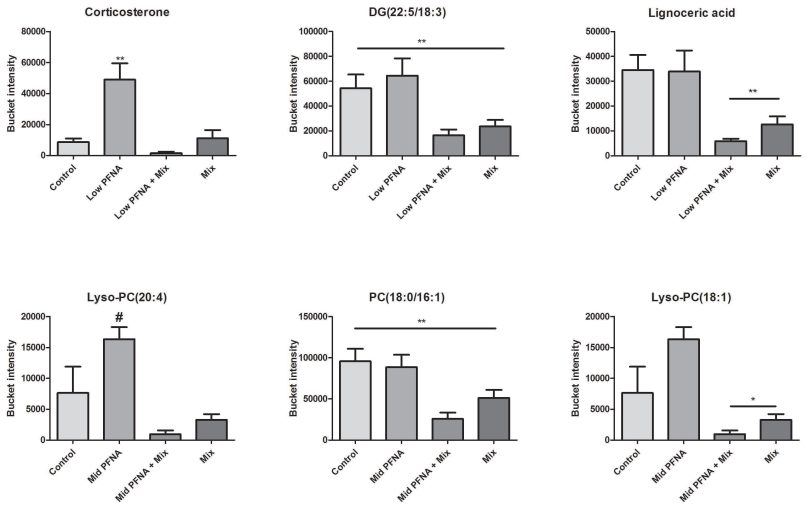



**Fig. 3 Representative metabolite changes in the Low PFNA + Mix group. Metabolite levels were analyzed statistically for each group by comparison to control (*t*-test). Figure a and d illustrates the case where PFNA alone causes an increase in metabolite level, whereas the combination of PFNA and Mix causes a decrease in metabolite level. Only in c and f is the effect by PFNA+Mix stronger than Mix alone (# = *p*-value of 0.053). The plotted metabolites belong to two metabolite classes PC (phosphatidylcholine) and DG (diacylglycerol)**

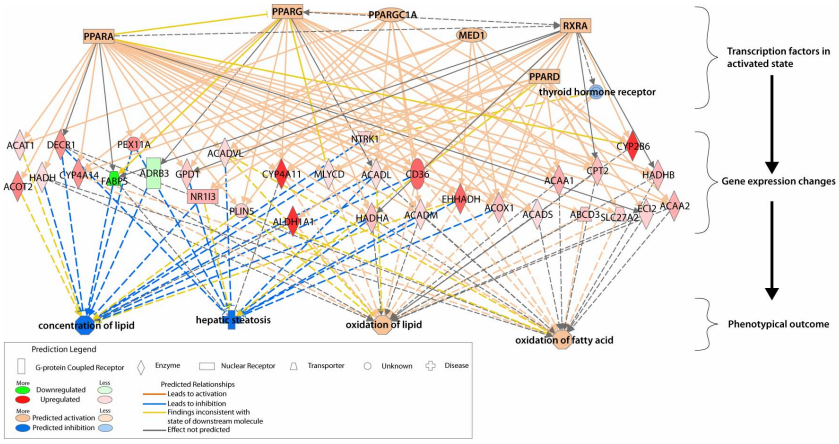



**Fig. 4 Illustration of potential hepatic signaling network from PPAR (and other transcriptional regulators) activation through gene expression to development of phenotypical outcomes. Orange indicates predicted activation and blue indicates predicted inhibition. Shades of red and green indicates level of increased and decreased gene expression, respectively. The network was generated through the use of QIAGEN's Ingenuity Pathway Analysis (IPA, QIAGEN Redwood City, www.qiagen.com/ingenuity)**

**Supplementary Table 1. The composition of Mix is based on the presented ratios that were estimated from human exposure levels** (Christiansen et al. 2012)**. These ratios formed the basis for calculation of the rat doses.**

| CAS registry number | Compound name | Source/use | Ratio in mixture (weight) | Rat dose (mg/kg bw/day) |
|---|---|---|---|---|
| 7380-40-7 | bergamottin | grapefruit constituent | 0.08 | 0.2 |
| 59870-68-7 | glabridin | liquorice constituent | 0.12 | 0.3 |
| 80-05-7 | bisphenol A | plastic additive | 0.004 | 0.01 |
| 94-26-8 | butyl paraben | preservative | 0.21 | 0.52 |
| 84-74-2 | dibutylphtalate (DBP) | plasticizer | 0.02 | 0.06 |
| 117-81-7 | bis(2-ethylhexyl)phthalate (DEHP) | plasticizer | 0.03 | 0.09 |
| 36861-47-9 | 4-methylbenzylidene camphor (4-MBC) | sun filter | 0.15 | 0.38 |
| 5466-77-3 | 2-Ethylhexyl-4-methoxycinnamate (OMC) | sun filter | 0.27 | 0.68 |
| 72-55-9 | dichlorodiphenyldichloroethylene (p,p´-DDE) | pesticide | 0.002 | 0.006 |
| 133855-98-8 | epoxiconazole | pesticide | 0.02 | 0.05 |
| 330-55-2 | linuron | pesticide | 0.002 | 0.004 |
| 67747-09-5 | prochloraz | pesticide | 0.025 | 0.06 |
| 32809-16-8 | procymidone | pesticide | 0.035 | 0.09 |
| 50471-44-8 | vinclozolin | pesticide | 0.021 | 0.05 |
| Total dose | | | | 2.50 |

**Supplementary Table 2. Number of differentially expressed genes up- or downregulated by Mid PFNA with or without Mix ($p \leq 0.05$) based on a statistical comparison to the control group. These genes were included in the pathway analysis.**

| Direction of regulation | Treatment group | |
|---|---|---|
| | Mid PFNA | Mid PFNA + Mix |
| Up | 140 | 136 |
| Down | 42 | 67 |
| Total | 182 | 203 |

**Supplementary Table 3. Significantly differentially expressed genes ($p \leq 0.05$ and $p \leq 0.1$) in Mid PFNA ± Mix and Low PFNA + Mix, respectively.**

| Symbol | Entrez ID | Entrez Gene Name | $Log_2$(fold change) | | | Functional network |
|---|---|---|---|---|---|---|
| | | | Low PFNA Mix | Mid PFNA | Mid PFNA Mix | |
| | | | Upregulated | | | |
| Cyp4a11 | 50549 | cytochrome P450, family 4, subfamily A, polypeptide 11 | 0,643 | 2,419 | 2,568 | Energy Production, Lipid Metabolism |
| Cyp2b6 | 361523\| 24300 | cytochrome P450, family 2, subfamily B, polypeptide 6 | 0,971 | 2,12 | 2,561 | Energy Production, Lipid Metabolism |
| Aldh1a1 | 24188 | aldehyde dehydrogenase 1 family, member A1 | 0,377 | 2,077 | 2,523 | Energy Production, Lipid Metabolism |
| Vnn1 | 29142 | vanin 1 | 0,287 | 2,036 | 2,059 | Energy Production, Lipid Metabolism |
| Acot1 | 50559 | acyl-CoA thioesterase 1 | 1,118 | 1,278 | 2,052 | Energy Production, Lipid Metabolism |
| Colq | 29755 | collagen-like tail subunit (single strand of homotrimer) of asymmetric acetylcholinesterase | 0,528 | 1,514 | 1,954 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Aig1 | 292486 | androgen-induced 1 | 0,199 | 1,788 | 1,908 | Energy Production, Lipid Metabolism |
| Cd36 | 29184 | CD36 molecule (thrombospondin receptor) | 0,158 | 1,276 | 1,867 | Energy Production, Lipid Metabolism |
| Ehhadh | 171142 | enoyl-CoA, hydratase/3-hydroxyacyl CoA dehydrogenase | 0,184 | 1,702 | 1,86 | Energy Production, Lipid Metabolism |
| Ech1 | 64526 | enoyl CoA hydratase 1, peroxisomal | 0,459 | 1,503 | 1,856 | Energy Production, Lipid Metabolism |
| Crat | 311849 | carnitine O-acetyltransferase | 0,372 | 1,55 | 1,71 | Energy Production, Lipid Metabolism |
| Acot4 | 681337 | acyl-CoA thioesterase 4 | 0,264 | 1,115 | 1,585 | Lipid Metabolism, Small Molecule Biochemistry, Nucleic Acid Metabolism |
| Slc39a5 | 362812 | solute carrier family 39 (zinc transporter), member 5 | 0,362 | 0,911 | 1,562 | Energy Production, Lipid Metabolism |
| Acsm5 | 361637 | acyl-CoA synthetase medium-chain family member 5 | 0,764 | 1,547 | 1,554 | Energy Production, Lipid Metabolism |
| Acot2 | 192272 | acyl-CoA thioesterase 2 | 0,241 | 1,312 | 1,497 | Energy Production, Lipid Metabolism |
| Cyp2b23 | 292728 | cytochrome P450, family 2, subfamily b, polypeptide 23 | 0,354 | 1,021 | 1,44 | Energy Production, Lipid Metabolism |
| Acot1 | 314304 | acyl-CoA thioesterase 1 | 0,386 | 0,984 | 1,427 | Lipid Metabolism, Small Molecule Biochemistry, Nucleic Acid Metabolism |
| Pex11a | 85249 | peroxisomal biogenesis factor 11 alpha | 0,209 | 1,368 | 1,387 | Energy Production, Lipid Metabolism |
| Decr1 | 117543 | 2,4-dienoyl CoA reductase 1, mitochondrial | 0,234 | 1,364 | 1,344 | Energy Production, Lipid Metabolism |
| Cyp4a14 | 298423\| 24306 | cytochrome P450, family 4, subfamily a, polypeptide 14 | 0,333 | 1,201 | 1,335 | Energy Production, Lipid Metabolism |
| Ces2e | 192257 | carboxylesterase 2E | 0,731 | 0,557 | 1,325 | Energy Production, Lipid Metabolism |
| Hdc | 24443 | histidine decarboxylase | 0,267 | 1,03 | 1,309 | Energy Production, Lipid Metabolism |
| Ephx2 | 65030 | epoxide hydrolase 2, cytoplasmic | 0,243 | 1,201 | 1,267 | Energy Production, Lipid Metabolism |
| Cyp17a1 | 25146 | cytochrome P450, family 17, subfamily A, polypeptide 1 | 0,337 | 1,358 | 1,254 | Energy Production, Lipid Metabolism |
| Crot | 83842 | carnitine O-octanoyltransferase | 0,19 | 1,115 | 1,219 | Energy Production, Lipid Metabolism |
| Cyp2j2 | 65210 | cytochrome P450, family 2, subfamily J, polypeptide 2 | 0,237 | 1,018 | 1,187 | Energy Production, Lipid Metabolism |
| Lppr1 | 298062 | lipid phosphate phosphatase-related protein type 1 | 0,29 | 1,236 | 1,091 | Other |
| Fbp2 | 114508 | fructose-1,6-bisphosphatase 2 | 0,272 | 0,642 | 1,081 | Energy Production, Lipid Metabolism |
| Hsdl2 | 313200 | hydroxysteroid dehydrogenase like 2 | 0,095 | 0,989 | 1,038 | Energy Production, Lipid Metabolism |
| Ces2c (includes others) | 100910144\| 100910040\| 171118 | carboxylesterase 2C | 0,282 | 0,705 | 1,029 | Energy Production, Lipid Metabolism |
| Acaa1 | 501072\| 24157 | acetyl-CoA acyltransferase 1 | 0,113 | 1,008 | 0,985 | Energy Production, Lipid Metabolism |
| H19 | 309122 | H19, imprinted maternally expressed transcript (non-protein coding) | 0,393 | 1,142 | 0,981 | Energy Production, Lipid Metabolism |
| Acaa2 | 170465 | acetyl-CoA acyltransferase 2 | 0,394 | 0,556 | 0,969 | Energy Production, Lipid Metabolism |
| Decr2 | 64461 | 2,4-dienoyl CoA reductase 2, peroxisomal | 0,275 | 0,831 | 0,944 | Energy Production, Lipid Metabolism |
| Nr1i3 | 65035 | nuclear receptor subfamily 1, group I, member 3 | 0,226 | 0,787 | 0,935 | Energy Production, Lipid Metabolism |
| Lgals2 | 171134 | lectin, galactoside-binding, soluble, 2 | 0,279 | 0,829 | 0,916 | Energy Production, Lipid Metabolism |
| Scpep1 | 114861 | serine carboxypeptidase 1 | 0,346 | 1,013 | 0,912 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Prrg4 | 499847 | proline rich Gla (G-carboxyglutamic acid) 4 (transmembrane) | 0,224 | 0,893 | 0,877 | Energy Production, Lipid Metabolism |
| Hadhb | 171155 | hydroxyacyl-CoA dehydrogenase/3-ketoacyl-CoA thiolase/enoyl-CoA | 0,194 | 0,8 | 0,817 | Energy Production, Lipid Metabolism |

| Symbol | Entrez ID | Entrez Gene Name | $Log_2$(fold change) Low PFNA Mix | Mid PFNA | Mid PFNA Mix | Functional network |
|---|---|---|---|---|---|---|
| | | Upregulated | | | | |
| | | hydratase (trifunctional protein), beta subunit | | | | |
| Tex36 | 499279 | testis expressed 36 | 0,119 | 1,017 | 0,812 | Cell Cycle, Cellular Movement, Cellular Development |
| Tmem14c | 171432\| 100361993 | transmembrane protein 14C | 0,281 | 0,828 | 0,808 | Energy Production, Lipid Metabolism |
| Gde1 | 60418 | glycerophosphodiester phosphodiesterase 1 | 0,268 | 0,859 | 0,792 | Cellular Compromise, Cellular Function and Maintenance, Lipid Metabolism |
| Cpt2 | 25413 | carnitine palmitoyltransferase 2 | 0,163 | 0,565 | 0,786 | Energy Production, Lipid Metabolism |
| Reep6 | 362835 | receptor accessory protein 6 | 0,192 | 0,658 | 0,784 | Energy Production, Lipid Metabolism |
| Cyp2e1 | 25086 | cytochrome P450, family 2, subfamily E, polypeptide 1 | 0,339 | 0,736 | 0,779 | Other |
| Sorl1 | 300652 | sortilin-related receptor, L(DLR class) A repeats containing | 0,309 | 0,64 | 0,761 | Energy Production, Lipid Metabolism |
| Cyp2c19 | 293989 | cytochrome P450, family 2, subfamily C, polypeptide 19 | 0,128 | 0,686 | 0,756 | Energy Production, Lipid Metabolism |
| Hadha | 170670 | hydroxyacyl-CoA dehydrogenase/3-ketoacyl-CoA thiolase/enoyl-CoA hydratase (trifunctional protein), alpha subunit | 0,144 | 0,659 | 0,746 | Energy Production, Lipid Metabolism |
| Acox1 | 50681 | acyl-CoA oxidase 1, palmitoyl | 0,074 | 0,601 | 0,745 | Energy Production, Lipid Metabolism |
| Lipi | 288322 | lipase, member I | 0,056 | -0,04 | 0,731 | Energy Production, Lipid Metabolism |
| Nudt7 | 361413 | nudix (nucleoside diphosphate linked moiety X)-type motif 7 | 0,163 | 0,622 | 0,731 | Other |
| Acy3 | 293653 | aspartoacylase (aminocyclase) 3 | 0,393 | 0,821 | 0,721 | Hepatocellular Peroxisome Proliferation, Cellular Assembly and Organization, Cellular Function and Maintenance |
| Dnhd1 | 690115 | dynein heavy chain domain 1 | 0,177 | 0,384 | 0,706 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Fam195a | 685545 | family with sequence similarity 195, member A | 0,165 | 0,656 | 0,705 | Energy Production, Lipid Metabolism |
| Slc25a20 | 117035 | solute carrier family 25 (carnitine/acylcarnitine translocase), member 20 | 0,098 | 0,668 | 0,692 | Energy Production, Lipid Metabolism |
| Pcsk4 | 171085 | proprotein convertase subtilisin/kexin type 4 | 0,029 | 0,509 | 0,686 | Molecular Transport, Protein Trafficking, Cellular Assembly and Organization |
| Abcd3 | 25270 | ATP-binding cassette, sub-family D (ALD), member 3 | 0,093 | 0,729 | 0,685 | Energy Production, Lipid Metabolism |
| Cpne4 | 367160 | copine IV | 0,297 | 0,782 | 0,673 | Energy Production, Lipid Metabolism |
| Acot12 | 170570 | acyl-CoA thioesterase 12 | 0,245 | 0,557 | 0,666 | Energy Production, Lipid Metabolism |
| Naprt1 | 315085 | nicotinate phosphoribosyltransferase domain containing 1 | 0,238 | 0,579 | 0,661 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Rrm1 | 685579 | ribonucleotide reductase M1 | 0,154 | 0,006 | 0,66 | Molecular Transport, Protein Trafficking, Cellular Assembly and Organization |
| Wfdc3 | 296366 | WAP four-disulfide core domain 3 | 0,203 | 0,714 | 0,656 | Energy Production, Lipid Metabolism |
| Acad11 | 315973 | acyl-CoA dehydrogenase family, member 11 | 0,115 | 0,62 | 0,656 | Energy Production, Lipid Metabolism |
| Ca7 | 291819 | carbonic anhydrase VII | -0,06 | 0,817 | 0,643 | Energy Production, Lipid Metabolism |
| Gpd1 | 60666 | glycerol-3-phosphate dehydrogenase 1 (soluble) | 0,165 | 0,591 | 0,642 | Energy Production, Lipid Metabolism |
| Gcdh | 364975 | glutaryl-CoA dehydrogenase | 0,175 | 0,576 | 0,638 | Energy Production, Lipid Metabolism |
| Spink1/spink 3 | 266602\| 24833 | serine peptidase inhibitor, Kazal type 3 | 0,321 | 1,319 | 0,633 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Aqp11 | 286758 | aquaporin 11 | 0,082 | 0,508 | 0,632 | Energy Production, Lipid Metabolism |
| Pdzk1 | 65144 | PDZ domain containing 1 | 0,167 | 0,542 | 0,63 | Lipid Metabolism, Small Molecule Biochemistry, Nucleic Acid Metabolism |
| Plin5 | 501283 | perilipin 5 | 0,198 | 0,595 | 0,61 | Energy Production, Lipid Metabolism |
| Acot7 | 26759 | acyl-CoA thioesterase 7 | 0,019 | 0,431 | 0,593 | Energy Production, Lipid Metabolism |
| Adtrp | 361228 | androgen-dependent TFPI-regulating protein | 0,174 | 0,602 | 0,586 | Energy Production, Lipid Metabolism |
| Orc3 | 313138 | origin recognition complex, subunit 3 | 0,258 | 0,198 | 0,575 | Lipid Metabolism, Small Molecule Biochemistry, Nucleic Acid Metabolism |
| Lef1 | 161452 | lymphoid enhancer-binding factor 1 | 0,1 | 0,36 | 0,573 | Energy Production, Lipid Metabolism |
| Slc27a2 | 65192 | solute carrier family 27 (fatty acid transporter), member 2 | 0,215 | 0,543 | 0,57 | Energy Production, Lipid Metabolism |
| Galk1 | 287835 | galactokinase 1 | 0,146 | 0,38 | 0,569 | Energy Production, Lipid Metabolism |
| Eci2 | 291075 | enoyl-CoA delta isomerase 2 | 0,117 | 0,577 | 0,548 | Energy Production, Lipid Metabolism |
| Aco1 | 50655 | aconitase 1, soluble | 0,12 | 0,337 | 0,544 | Drug Metabolism, Small Molecule |

| Symbol | Entrez ID | Entrez Gene Name | $Log_2$(fold change) Low PFNA Mix | Mid PFNA | Mid PFNA Mix | Functional network |
|---|---|---|---|---|---|---|
| | | Upregulated | | | | |
| | | | | | | Biochemistry, Cell-To-Cell Signaling and Interaction |
| Ces2a/ces2i | 246252\| 292152 | carboxylesterase 2A | 0,32 | 0,676 | 0,54 | Other |
| Gnat3 | 286924 | guanine nucleotide binding protein, alpha transducing 3 | 0,001 | 0,534 | 0,539 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Dtwd2 | 361326 | DTW domain containing 2 | 0,245 | 0,078 | 0,53 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Fn3k | 498034 | fructosamine 3 kinase | 0,155 | 0,711 | 0,528 | Carbohydrate Metabolism, Nucleic Acid Metabolism, Small Molecule Biochemistry |
| Ntrk1 | 59109 | neurotrophic tyrosine kinase, receptor, type 1 | 0,131 | 0,562 | 0,528 | Cell Morphology, Cellular Assembly and Organization, Cellular Development |
| Ccdc185 | 289334 | coiled-coil domain containing 185 | 0,158 | -0,18 | 0,521 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Vill | 301057 | villin-like | 0,166 | 0,521 | 0,52 | Hepatocellular Peroxisome Proliferation, Cellular Assembly and Organization, Cellular Function and Maintenance |
| Asb5 | 361187 | ankyrin repeat and SOCS box containing 5 | 0,088 | 0,778 | 0,513 | Hepatocellular Peroxisome Proliferation, Cellular Assembly and Organization, Cellular Function and Maintenance |
| Nagk | 297393 | N-acetylglucosamine kinase | 0,136 | 0,407 | 0,508 | Energy Production, Lipid Metabolism |
| Pepd | 292808 | peptidase D | 0,15 | 0,48 | 0,506 | Energy Production, Lipid Metabolism |
| Mecr | 29470 | mitochondrial trans-2-enoyl-CoA reductase | 0,229 | 0,634 | 0,504 | Energy Production, Lipid Metabolism |
| Tmem120a | 288591 | transmembrane protein 120A | 0,133 | 0,364 | 0,502 | Energy Production, Lipid Metabolism |
| Acat1 | 25014 | acetyl-CoA acetyltransferase 1 | -0,111 | 0,591 | 0,497 | Energy Production, Lipid Metabolism |
| Ccbl1 | 311844 | cysteine conjugate-beta lyase, cytoplasmic | 0,142 | 0,268 | 0,49 | Energy Production, Lipid Metabolism |
| Acadvl | 25363 | acyl-CoA dehydrogenase, very long chain | 0,047 | 0,413 | 0,482 | Energy Production, Lipid Metabolism |
| Elovl2 | 498728 | ELOVL fatty acid elongase 2 | 0,352 | 0,518 | 0,478 | Cellular Compromise, Cellular Function and Maintenance, Lipid Metabolism |
| Map4k2 | 293694 | mitogen-activated protein kinase kinase kinase 2 | 0,09 | 0,393 | 0,476 | Cell Morphology, Cellular Assembly and Organization, Cellular Development |
| Rnase9 | 364301 | ribonuclease, RNase A family, 9 (non-active) | 0,266 | 0,135 | 0,475 | Energy Production, Lipid Metabolism |
| Apoa2 | 25649 | apolipoprotein A-II | 0,321 | 0,504 | 0,474 | Energy Production, Lipid Metabolism |
| Spdef | 689210 | SAM pointed domain containing ETS transcription factor | 0,187 | -0,032 | 0,471 | Energy Production, Lipid Metabolism |
| Hmgcs2 | 24450 | 3-hydroxy-3-methylglutaryl-CoA synthase 2 (mitochondrial) | 0,118 | 0,469 | 0,471 | Energy Production, Lipid Metabolism |
| Ptk2b | 50646 | protein tyrosine kinase 2 beta | 0,175 | 0,607 | 0,47 | Cell Morphology, Cellular Assembly and Organization, Cellular Development |
| Acadm | 24158 | acyl-CoA dehydrogenase, C-4 to C-12 straight chain | 0,15 | 0,47 | 0,47 | Energy Production, Lipid Metabolism |
| Prodh2 | 361538 | proline dehydrogenase (oxidase) 2 | 0,197 | 0,417 | 0,47 | Molecular Transport, Protein Trafficking, Cellular Assembly and Organization |
| Impa2 | 282636 | inositol(myo)-1(or 4)-monophosphatase 2 | 0,389 | 0,326 | 0,469 | Energy Production, Lipid Metabolism |
| Pink1 | 298575 | PTEN induced putative kinase 1 | 0,19 | 0,375 | 0,467 | Energy Production, Lipid Metabolism |
| Iba57 | 363611 | IBA57, iron-sulfur cluster assembly homolog (S. cerevisiae) | 0,044 | 0,43 | 0,463 | Energy Production, Lipid Metabolism |
| Sirt4 | 304539 | sirtuin 4 | 0,162 | 0,291 | 0,461 | Energy Production, Lipid Metabolism |
| Psrc1 | 691380 | proline/serine-rich coiled-coil 1 | 0,235 | 0,104 | 0,451 | Drug Metabolism, Small Molecule Biochemistry, Cell-To-Cell Signaling and Interaction |
| Cyp2a1 | 24894 | cytochrome P450, family 2, subfamily a, polypeptide 1 | 0,085 | 0,419 | 0,45 | Molecular Transport, Protein Trafficking, Cellular Assembly and Organization |
| Aco2 | 79250 | aconitase 2, mitochondrial | -0,065 | 0,428 | 0,448 | Energy Production, Lipid Metabolism |
| Depdc1 | 295538 | DEP domain containing 1 | 0,166 | -0,04 | 0,444 | Energy Production, Lipid Metabolism |
| Pex16 | 311203 | peroxisomal biogenesis factor 16 | 0,208 | 0,405 | 0,439 | Molecular Transport, Protein Trafficking, Cellular Assembly and Organization |
| Plcl2 | 301173 | phospholipase C-like 2 | 0,157 | 0,429 | 0,436 | Energy Production, Lipid Metabolism |
| Tssk1b | 288358 | testis-specific serine kinase 1B | 0,107 | -0,011 | 0,433 | Cell Morphology, Cellular Assembly and Organization, Cellular Development |
| Slc39a4 | 300051 | solute carrier family 39 (zinc transporter), member 4 | -0,035 | 0,658 | 0,428 | Energy Production, Lipid Metabolism |

| Symbol | Entrez ID | Entrez Gene Name | $Log_2$(fold change) Low PFNA Mix | Mid PFNA | Mid PFNA Mix | Functional network |
|---|---|---|---|---|---|---|
| | | Upregulated | | | | |
| Acadl | 25287 | acyl-CoA dehydrogenase, long chain | 0,05 | 0,442 | 0,419 | Energy Production, Lipid Metabolism |
| Acads | 64304 | acyl-CoA dehydrogenase, C-2 to C-3 short chain | 0,056 | 0,326 | 0,416 | Energy Production, Lipid Metabolism |
| Hadh | 113965 | hydroxyacyl-CoA dehydrogenase | 0,056 | 0,418 | 0,4 | Energy Production, Lipid Metabolism |
| Kiaa1033 | 314690 | KIAA1033 | 0,246 | 0,078 | 0,393 | Energy Production, Lipid Metabolism |
| Nckap1 | 58823 | NCK-associated protein 1 | 0,169 | 0,26 | 0,393 | Lipid Metabolism, Small Molecule Biochemistry, Nucleic Acid Metabolism |
| Calcr | 116506 | calcitonin receptor | 0,225 | 0,114 | 0,392 | Lipid Metabolism, Small Molecule Biochemistry, Nucleic Acid Metabolism |
| Mnd1 | 295160 | meiotic nuclear divisions 1 homolog (S. cerevisiae) | 0,089 | 0,505 | 0,377 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Immt | 312444 | inner membrane protein, mitochondrial | -0,011 | 0,345 | 0,376 | Energy Production, Lipid Metabolism |
| Iqsec2 | 685244 | IQ motif and Sec7 domain 2 | 0,005 | 0,331 | 0,375 | Molecular Transport, Protein Trafficking, Cellular Assembly and Organization |
| Apmap | 366227 | adipocyte plasma membrane associated protein | 0,168 | 0,496 | 0,369 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Fgf1 | 25317 | fibroblast growth factor 1 (acidic) | 0,142 | 0,291 | 0,368 | Cell Death and Survival, Visual System Development and Function, Cell-To-Cell Signaling and Interaction |
| Lactb | 300803 | lactamase, beta | -0,048 | 0,486 | 0,36 | Hepatocellular Peroxisome Proliferation, Cellular Assembly and Organization, Cellular Function and Maintenance |
| Sucla2 | 361071 | succinate-CoA ligase, ADP-forming, beta subunit | 0,145 | 0,43 | 0,357 | Energy Production, Lipid Metabolism |
| Blvrb | 292737 | biliverdin reductase B (flavin reductase (NADPH)) | 0,148 | 0,325 | 0,354 | Energy Production, Lipid Metabolism |
| Cabp1 | 171051 | calcium binding protein 1 | 0,05 | 0,13 | 0,344 | Cell Morphology, Cellular Assembly and Organization, Cellular Development |
| Hoxd13 | 288154 | homeobox D13 | 0,208 | 0,003 | 0,338 | Energy Production, Lipid Metabolism |
| Meiob | 685099 | meiosis specific with OB domains | -0,052 | 0,626 | 0,334 | Post-Translational Modification, Cancer, Organismal Injury and Abnormalities |
| Dqx1 | 680434 | DEAQ box RNA-dependent ATPase 1 | 0,033 | 0,188 | 0,334 | Cell Morphology, Cellular Assembly and Organization, Cellular Development |
| Rgs5 | 54294 | regulator of G-protein signaling 5 | 0,185 | 0,498 | 0,332 | Cellular Compromise, Cellular Function and Maintenance, Lipid Metabolism |
| Chchd10 | 361824 | coiled-coil-helix-coiled-coil-helix domain containing 10 | 0,055 | 0,337 | 0,327 | Energy Production, Lipid Metabolism |
| Psmc3 | 29677 | proteasome (prosome, macropain) 26S subunit, ATPase, 3 | 0,102 | 0,287 | 0,326 | Cell Morphology, Cellular Assembly and Organization, Cellular Development |
| Mlycd | 85239 | malonyl-CoA decarboxylase | 0,093 | 0,316 | 0,325 | Energy Production, Lipid Metabolism |
| Pet112 | 361974 | PET112 homolog (yeast) | -0,027 | 0,39 | 0,313 | Hepatocellular Peroxisome Proliferation, Cellular Assembly and Organization, Cellular Function and Maintenance |
| Ldha/RGD1562690 | 500965\|24533 | lactate dehydrogenase A | 0,092 | 0,393 | 0,313 | Energy Production, Lipid Metabolism |
| Trim10 | 294210 | tripartite motif containing 10 | 0,328* | 0,108 | 0,312 | Cell Death and Survival, Visual System Development and Function, Cell-To-Cell Signaling and Interaction |
| Asap1 | 314961 | ArfGAP with SH3 domain, ankyrin repeat and PH domain 1 | 0,099 | 0,122 | 0,31 | Cell Morphology, Cellular Assembly and Organization, Cellular Development |
| Perp | 292949 | PERP, TP53 apoptosis effector | -0,045 | 0,388 | 0,305 | Hepatocellular Peroxisome Proliferation, Cellular Assembly and Organization, Cellular Function and Maintenance |
| Ppfia4 | 140592 | protein tyrosine phosphatase, receptor type, f polypeptide (PTPRF), interacting protein (liprin), alpha 4 | 0,223 | 0,07 | 0,304 | Energy Production, Lipid Metabolism |
| Fam83c | 683506 | family with sequence similarity 83, member C | 0,14 | 0,09 | 0,301 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Hipk1 | 365895 | homeodomain interacting protein kinase 1 | 0,189 | 0,011 | 0,296 | Energy Production, Lipid Metabolism |
| Dhrs4 | 266686 | dehydrogenase/reductase (SDR family) member 4 | 0,005 | 0,363 | 0,294 | Energy Production, Lipid Metabolism |
| Ftcd | 89833 | formimidoyltransferase cyclodeaminase | 0,111 | 0,288 | 0,291 | Energy Production, Lipid Metabolism |
| Tcea2 | 29575 | transcription elongation factor A (SII), 2 | 0,054 | 0,25 | 0,288 | Energy Production, Lipid Metabolism |
| Abhd12 | 499913 | abhydrolase domain containing 12 | 0,051 | 0,356 | 0,286 | Energy Production, Lipid Metabolism |
| Txn2 | 79462 | thioredoxin 2 | 0,119 | 0,358 | 0,285 | Energy Production, Lipid Metabolism |
| Pex5 | 312703 | peroxisomal biogenesis factor 5 | 0,077 | 0,326 | 0,284 | Energy Production, Lipid Metabolism |

| Symbol | Entrez ID | Entrez Gene Name | $Log_2$(fold change) Low PFNA Mix | Mid PFNA | Mid PFNA Mix | Functional network |
|---|---|---|---|---|---|---|
| Upregulated | | | | | | |
| Etfdh | 295143 | electron-transferring-flavoprotein dehydrogenase | 0,031 | 0,354 | 0,284 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Suclg1 | 114597 | succinate-CoA ligase, alpha subunit | 0,024 | 0,275 | 0,284 | Energy Production, Lipid Metabolism |
| Ndufs3 | 295923 | NADH dehydrogenase (ubiquinone) Fe-S protein 3, 30kDa (NADH-coenzyme Q reductase) | 0,085 | 0,336 | 0,28 | Cellular Compromise, Cellular Function and Maintenance, Lipid Metabolism |
| Hsd17b4 | 79244 | hydroxysteroid (17-beta) dehydrogenase 4 | 0,113 | 0,295 | 0,278 | Other |
| Ndufa10 | 678759\| 316632 | NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 10, 42kDa | 0,042 | 0,3 | 0,275 | Energy Production, Lipid Metabolism |
| Wdr13 | 317370 | WD repeat domain 13 | 0,15 | 0,072 | 0,273 | Molecular Transport, Protein Trafficking, Cellular Assembly and Organization |
| Rsph3 | 361476 | radial spoke 3 homolog (Chlamydomonas) | 0,202 | 0,42 | 0,269 | Energy Production, Lipid Metabolism |
| Lalba | 24528 | lactalbumin, alpha- | 0,381 | 0,08 | 0,268 | Molecular Transport, Behavior, Post-Translational Modification |
| Tert | 301965 | telomerase reverse transcriptase | 0,098 | 0,122 | 0,263 | Other |
| Tnfrsf10a | 364420 | tumor necrosis factor receptor superfamily, member 10a | 0,042 | 0,138 | 0,255 | Drug Metabolism, Small Molecule Biochemistry, Cell-To-Cell Signaling and Interaction |
| Aadac | 57300 | arylacetamide deacetylase | 0,088 | 0,389 | 0,254 | Energy Production, Lipid Metabolism |
| Lrpprc | 313867 | leucine-rich pentatricopeptide repeat containing | 0,082 | 0,279 | 0,253 | Carbohydrate Metabolism, Nucleic Acid Metabolism, Small Molecule Biochemistry |
| Etfa | 300726 | electron-transfer-flavoprotein, alpha polypeptide | -0,012 | 0,372 | 0,252 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Idh3b | 94173 | isocitrate dehydrogenase 3 (NAD+) beta | 0,003 | 0,325 | 0,247 | Carbohydrate Metabolism, Nucleic Acid Metabolism, Small Molecule Biochemistry |
| Mfn2 | 100911485\| 64476 | mitofusin 2 | 0,023 | 0,264 | 0,24 | Energy Production, Lipid Metabolism |
| Psmd8 | 292766 | proteasome (prosome, macropain) 26S subunit, non-ATPase, 8 | 0,103 | 0,179 | 0,238 | Cell Morphology, Cellular Assembly and Organization, Cellular Development |
| Cox11 | 690300 | cytochrome c oxidase assembly homolog 11 (yeast) | 0,24 | 0,593 | 0,235 | Hepatocellular Peroxisome Proliferation, Cellular Assembly and Organization, Cellular Function and Maintenance |
| Micu2 | 171433 | mitochondrial calcium uptake 2 | 0,13 | 0,295 | 0,232 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Or10a4 | 293375 | olfactory receptor, family 10, subfamily A, member 4 | 0,136 | 0,021 | 0,222 | Nervous System Development and Function, Cell Morphology, Cellular Movement |
| Glrx5 | 362776 | glutaredoxin 5 | 0,039 | 0,431 | 0,221 | Energy Production, Lipid Metabolism |
| C6orf136 | 294231 | chromosome 6 open reading frame 136 | 0,16 | 0,336 | 0,203 | Energy Production, Lipid Metabolism |
| Smim20 | 501923 | small integral membrane protein 20 | 0,108 | 0,319 | 0,198 | Hepatocellular Peroxisome Proliferation, Cellular Assembly and Organization, Cellular Function and Maintenance |
| Vps26a | 361846 | vacuolar protein sorting 26 homolog A (S. pombe) | 0,12 | 0,315 | 0,195 | Energy Production, Lipid Metabolism |
| Fbxl16 | 494223 | F-box and leucine-rich repeat protein 16 | 0,23* | 0,03 | 0,182 | Cell Death and Survival, Visual System Development and Function, Cell-To-Cell Signaling and Interaction |
| Ndufb9 | 299954 | NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 9, 22kDa | 0,039 | 0,214 | 0,18 | Energy Production, Lipid Metabolism |
| Tgm5 | 691932 | transglutaminase 5 | 0,381* | 0,098 | 0,177 | Cell Death and Survival, Visual System Development and Function, Cell-To-Cell Signaling and Interaction |
| Olfr1305 | 404796 | olfactory receptor 1305 | 0,287 | 0,091 | 0,172 | Other |
| Sdhb | 298596 | succinate dehydrogenase complex, subunit B, iron sulfur (Ip) | 0,043 | 0,244 | 0,17 | Energy Production, Lipid Metabolism |
| Arhgap24 | 305156 | Rho GTPase activating protein 24 | 0,451* | -0,014 | 0,164 | Cell Death and Survival, Visual System Development and Function, Cell-To-Cell Signaling and Interaction |
| Endod1 | 363015 | endonuclease domain containing 1 | 0,167 | 0,269 | 0,162 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Nsmce1 | 361645 | non-SMC element 1 homolog (S. cerevisiae) | -0,009 | 0,218 | 0,161 | Energy Production, Lipid Metabolism |
| Prl3b1 | 24283 | prolactin family 3, subfamily b, member 1 | 0,224 | -0,042 | 0,156 | Molecular Transport, Behavior, Post-Translational Modification |
| Mrps18a | 301249 | mitochondrial ribosomal protein S18A | 0,03 | 0,259 | 0,123 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Ndufab1 | 293453 | NADH dehydrogenase (ubiquinone) 1, alpha/beta subcomplex, 1, 8kDa | 0,051 | 0,269 | 0,116 | Cellular Compromise, Cellular Function and Maintenance, Lipid Metabolism |

| Symbol | Entrez ID | Entrez Gene Name | $Log_2$(fold change) | | | Functional network |
|---|---|---|---|---|---|---|
| | | | Low PFNA Mix | Mid PFNA | Mid PFNA Mix | |
| Upregulated | | | | | | |
| Downregulated | | | | | | |
| Fabp5 | 140868 | fatty acid binding protein 5 (psoriasis-associated) | -0,385 | -1,392 | -1,508 | Cell Morphology, Cellular Assembly and Organization, Cellular Development |
| Erich5 | 681820 | glutamate-rich 5 | -0,895 | -0,956 | -1,372 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Cxcl13 | 498335 | chemokine (C-X-C motif) ligand 13 | -0,39 | -0,817 | -1,185 | Drug Metabolism, Small Molecule Biochemistry, Cell-To-Cell Signaling and Interaction |
| Ca1 | 310218 | carbonic anhydrase I | -0,003 | -0,78 | -1,132 | Energy Production, Lipid Metabolism |
| Hsd11b1 | 25116 | hydroxysteroid (11-beta) dehydrogenase 1 | -0,261 | -0,446 | -1,082 | Drug Metabolism, Small Molecule Biochemistry, Cell-To-Cell Signaling and Interaction |
| Phgdh | 58835 | phosphoglycerate dehydrogenase | -0,308 | -1,271 | -0,976 | Energy Production, Lipid Metabolism |
| Mon1b | 307868 | MON1 secretory trafficking family member B | -0,255 | -0,317 | -0,877 | Energy Production, Lipid Metabolism |
| Isyna1 | 290651 | inositol-3-phosphate synthase 1 | -0,1 | -0,906 | -0,825 | Energy Production, Lipid Metabolism |
| Ddc | 24311 | dopa decarboxylase (aromatic L-amino acid decarboxylase) | -0,355 | -0,591 | -0,634 | Energy Production, Lipid Metabolism |
| Nfic | 29228 | nuclear factor I/C (CCAAT-binding transcription factor) | -0,341 | -0,262 | -0,631 | Drug Metabolism, Small Molecule Biochemistry, Cell-To-Cell Signaling and Interaction |
| Ptpn13 | 498331 | protein tyrosine phosphatase, non-receptor type 13 (APO-1/CD95 (Fas)-associated phosphatase) | -0,361 | -0,411 | -0,574 | Lipid Metabolism, Small Molecule Biochemistry, Nucleic Acid Metabolism |
| Kcnn2 | 54262 | potassium intermediate/small conductance calcium-activated channel, subfamily N, member 2 | -0,271 | -0,197 | -0,552 | Lipid Metabolism, Small Molecule Biochemistry, Nucleic Acid Metabolism |
| Xrcc2 | 499966 | X-ray repair complementing defective repair in Chinese hamster cells 2 | -0,485 | -0,378 | -0,547 | Energy Production, Lipid Metabolism |
| Spcs2 | 293142 | signal peptidase complex subunit 2 homolog (S. cerevisiae) | -0,311 | -0,345 | -0,544 | Molecular Transport, Protein Trafficking, Cellular Assembly and Organization |
| Mgat2 | 94273 | mannosyl (alpha-1,6-)-glycoprotein beta-1,2-N-acetylglucosaminyltransferase | -0,135 | -0,436 | -0,544 | Energy Production, Lipid Metabolism |
| Mpc1 | 171087 | mitochondrial pyruvate carrier 1 | -0,185 | -0,371 | -0,539 | Energy Production, Lipid Metabolism |
| Dap | 64322 | death-associated protein | -0,066 | -0,778 | -0,529 | Energy Production, Lipid Metabolism |
| Cacna1h | 114862 | calcium channel, voltage-dependent, T type, alpha 1H subunit | -0,42 | -0,288 | -0,527 | Lipid Metabolism, Small Molecule Biochemistry, Nucleic Acid Metabolism |
| Adora1 | 29290 | adenosine A1 receptor | -0,155 | -0,916 | -0,519 | Amino Acid Metabolism, Drug Metabolism, Endocrine System Development and Function |
| March9 | 679272 | membrane-associated ring finger (C3HC4) 9 | -0,115 | -0,454 | -0,518 | Other |
| S1pr1 | 29733 | sphingosine-1-phosphate receptor 1 | -0,09 | -0,272 | -0,493 | Cell Morphology, Cellular Assembly and Organization, Cellular Development |
| Cks1b | 499655 | CDC28 protein kinase regulatory subunit 1B | -0,185 | -0,465 | -0,491 | Energy Production, Lipid Metabolism |
| Znf25 | 100158232 | zinc finger protein 25 | -0,302 | -0,29 | -0,48 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Lhx6 | 311901 | LIM homeobox 6 | -0,151 | -0,462 | -0,478 | Energy Production, Lipid Metabolism |
| Sdf2l1 | 680945 | stromal cell-derived factor 2-like 1 | -0,194 | -0,638 | -0,477 | Cellular Compromise, Cellular Function and Maintenance, Lipid Metabolism |
| Nipal3 | 502990 | NIPA-like domain containing 3 | -0,364 | -0,32 | -0,474 | Cell Morphology, Cellular Assembly and Organization, Cellular Development |
| Sec11c | 266758 | SEC11 homolog C (S. cerevisiae) | -0,274 | -0,444 | -0,455 | Molecular Transport, Protein Trafficking, Cellular Assembly and Organization |
| Tmem123 | 363013 | transmembrane protein 123 | -0,158 | -0,257 | -0,447 | Molecular Transport, Protein Trafficking, Cellular Assembly and Organization |
| Map4k4 | 301363 | mitogen-activated protein kinase kinase kinase kinase 4 | -0,26 | -0,338 | -0,44 | Other |
| Synpo2 | 499702 | synaptopodin 2 | -0,317 | -0,117 | -0,417 | Energy Production, Lipid Metabolism |
| Dopey2 | 304077 | dopey family member 2 | -0,135 | -0,453 | -0,416 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Steap4 | 499991 | STEAP family member 4 | -0,15 | -0,332 | -0,404 | Energy Production, Lipid Metabolism |
| Yipf5 | 361315 | Yip1 domain family, member 5 | -0,25 | -0,458 | -0,394 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Manf | 315989 | mesencephalic astrocyte-derived neurotrophic factor | -0,263 | -0,425 | -0,392 | Energy Production, Lipid Metabolism |

| Symbol | Entrez ID | Entrez Gene Name | $Log_2$(fold change) Low PFNA Mix | Mid PFNA | Mid PFNA Mix | Functional network |
|---|---|---|---|---|---|---|
| Upregulated | | | | | | |
| Col14a1 | 314981 | collagen, type XIV, alpha 1 | -0,275 | -0,233 | -0,381 | Energy Production, Lipid Metabolism |
| Olfr572 | 684365 | olfactory receptor 572 | -0,348 | -0,102 | -0,379 | Other |
| Ccnk | 500715 | cyclin K | -0,292 | -0,079 | -0,377 | Energy Production, Lipid Metabolism |
| Nek7 | 360850 | NIMA-related kinase 7 | -0,111 | -0,122 | -0,376 | Energy Production, Lipid Metabolism |
| Cfh | 155012 | complement factor H | -0,062 | -0,354 | -0,373 | Energy Production, Lipid Metabolism |
| Adrb3 | 25645 | adrenoceptor beta 3 | -0,279 | -0,134 | -0,366 | Lipid Metabolism, Small Molecule Biochemistry, Nucleic Acid Metabolism |
| Ccdc28b | 682445 | coiled-coil domain containing 28B | -0,151 | -0,716 | -0,362 | Energy Production, Lipid Metabolism |
| Dnajc3 | 63880 | DnaJ (Hsp40) homolog, subfamily C, member 3 | -0,113 | -0,279 | -0,362 | Energy Production, Lipid Metabolism |
| Mybph | 83708 | myosin binding protein H | -0,281 | -0,231 | -0,361 | Cell Morphology, Cellular Assembly and Organization, Cellular Development |
| Chpf | 316533 | chondroitin polymerizing factor | 0,004 | -0,524 | -0,36 | Energy Production, Lipid Metabolism |
| Isg20 | 293052 | interferon stimulated exonuclease gene 20kDa | 0,066 | -0,533 | -0,36 | Energy Production, Lipid Metabolism |
| Hook1 | 313370 | hook microtubule-tethering protein 1 | -0,09 | -0,282 | -0,356 | Cell Morphology, Cellular Assembly and Organization, Cellular Development |
| Dynlt3 | 363448 | dynein, light chain, Tctex-type 3 | -0,395* | -0,297 | -0,355 | Cell Death and Survival, Visual System Development and Function, Cell-To-Cell Signaling and Interaction |
| Kiaa1919 | 499462 | KIAA1919 | -0,237 | -0,171 | -0,352 | Humoral Immune Response, Inflammatory Response, Cell Cycle |
| Pfkm | 65152 | phosphofructokinase, muscle | -0,388 | -0,487 | -0,35 | Energy Production, Lipid Metabolism |
| Kpnb1 | 24917 | karyopherin (importin) beta 1 | -0,232 | -0,226 | -0,35 | Molecular Transport, Protein Trafficking, Cellular Assembly and Organization |
| Nsmf | 117536 | NMDA receptor synaptonuclear signaling and neuronal migration factor | -0,119 | -0,605 | -0,345 | Cellular Compromise, Cellular Function and Maintenance, Lipid Metabolism |
| Prox1 | 305066 | prospero homeobox 1 | -0,09 | -0,114 | -0,336 | Drug Metabolism, Small Molecule Biochemistry, Cell-To-Cell Signaling and Interaction |
| LOC685792 (includes others) | 363306* | similar to Discs large homolog 5 (Placenta and prostate DLG) (Discs large protein P-dlg) | -0,323 | 0,036 | -0,334 | Energy Production, Lipid Metabolism |
| Lnx2 | 360761 | ligand of numb-protein X 2 | -0,05 | -0,245 | -0,332 | Energy Production, Lipid Metabolism |
| C5orf30 | 501195 | chromosome 5 open reading frame 30 | -0,207 | -0,173 | -0,328 | Molecular Transport, Protein Trafficking, Cellular Assembly and Organization |
| Olfr1321 | 302821 | olfactory receptor 1321 | -0,433 | -0,243 | -0,319 | Other |
| Arl4d | 303559 | ADP-ribosylation factor-like 4D | -0,247 | -0,323 | -0,316 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Pld3 | 361527 | phospholipase D family, member 3 | -0,127 | -0,391 | -0,312 | Energy Production, Lipid Metabolism |
| Myo1d | 25485 | myosin ID | -0,157 | -0,601 | -0,312 | Hepatocellular Peroxisome Proliferation, Cellular Assembly and Organization, Cellular Function and Maintenance |
| Dner | 316573 | delta/notch-like EGF repeat containing | -0,256 | -0,086 | -0,309 | Lipid Metabolism, Small Molecule Biochemistry, Nucleic Acid Metabolism |
| Mybpc1 | 362867 | myosin binding protein C, slow type | -0,149 | -0,085 | -0,305 | Drug Metabolism, Small Molecule Biochemistry, Cell-To-Cell Signaling and Interaction |
| Grifin | 117130 | galectin-related inter-fiber protein | -0,248 | -0,236 | -0,304 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Rt1-ba | 309621 | major histocompatibility complex, class II, DQ alpha 1 | -0,244 | -0,139 | -0,302 | Energy Production, Lipid Metabolism |
| Ramp1 | 58965 | receptor (G protein-coupled) activity modifying protein 1 | -0,133 | -0,213 | -0,299 | Lipid Metabolism, Small Molecule Biochemistry, Nucleic Acid Metabolism |
| Gja3 | 79217 | gap junction protein, alpha 3, 46kDa | -0,192 | -0,15 | -0,292 | Energy Production, Lipid Metabolism |
| Hspa5 | 25617 | heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa) | -0,081 | -0,32 | -0,288 | Cellular Compromise, Cellular Function and Maintenance, Lipid Metabolism |
| H2afz | 58940 | H2A histone family, member Z | -0,007 | -0,456 | -0,284 | Energy Production, Lipid Metabolism |
| Krtcap2 | 295243 | keratinocyte associated protein 2 | -0,11 | -0,242 | -0,28 | Energy Production, Lipid Metabolism |
| Slc6a18 | 29323 | solute carrier family 6 (neutral amino acid transporter), member 18 | -0,379 | -0,229 | -0,279 | Other |
| Ssr3 | 81784 | signal sequence receptor, gamma (translocon-associated protein gamma) | -0,075 | -0,191 | -0,274 | Energy Production, Lipid Metabolism |
| Parp6 | 300759 | poly (ADP-ribose) polymerase family, member 6 | -0,062 | -0,594 | -0,272 | Amino Acid Metabolism, Drug Metabolism, Endocrine System Development and Function |

| Symbol | Entrez ID | Entrez Gene Name | $Log_2$(fold change) | | | Functional network |
|---|---|---|---|---|---|---|
| | | | Low PFNA Mix | Mid PFNA | Mid PFNA Mix | |
| Upregulated | | | | | | |
| Ipo7 | 308939 | importin 7 | -0,1 | -0,196 | -0,272 | Molecular Transport, Protein Trafficking, Cellular Assembly and Organization |
| Ncam1 | 24586 | neural cell adhesion molecule 1 | -0,163 | -0,05 | -0,27 | Cell Morphology, Cellular Assembly and Organization, Cellular Development |
| Eml4 | 313861 | echinoderm microtubule associated protein like 4 | -0,249 | -0,075 | -0,265 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Erp29 | 117030 | endoplasmic reticulum protein 29 | -0,165 | -0,323 | -0,25 | Energy Production, Lipid Metabolism |
| Timd2 | 287222 | T cell immunoglobulin and mucin domain containing 2 | -0,074 | -0,365 | -0,249 | Amino Acid Metabolism, Drug Metabolism, Endocrine System Development and Function |
| Mta1 | 64520 | metastasis associated 1 | -0,1 | -0,237 | -0,242 | Energy Production, Lipid Metabolism |
| Polr2l | 502374 | polymerase (RNA) II (DNA directed) polypeptide L | -0,03 | -0,23 | -0,237 | Energy Production, Lipid Metabolism |
| Ca4 | 29242 | carbonic anhydrase IV | -0,266 | -0,009 | -0,236 | Energy Production, Lipid Metabolism |
| Stmn2 | 84510 | stathmin 2 | -0,099 | -0,234 | -0,235 | Energy Production, Lipid Metabolism |
| Rrm2 | 362720\| 100359539 | ribonucleotide reductase M2 | -0,609 | -1,162 | -0,226 | Energy Production, Lipid Metabolism |
| Dync1li1 | 252902 | dynein, cytoplasmic 1, light intermediate chain 1 | -0,172 | -0,068 | -0,223 | Cell Death and Survival, Visual System Development and Function, Cell-To-Cell Signaling and Interaction |
| Ube2ql1 | 679949 | ubiquitin-conjugating enzyme E2Q family-like 1 | -0,287* | -0,04 | -0,217 | Cell Death and Survival, Visual System Development and Function, Cell-To-Cell Signaling and Interaction |
| Sdc1 | 25216 | syndecan 1 | -0,062 | -0,396 | -0,18 | Energy Production, Lipid Metabolism |
| Abcb9 | 63886 | ATP-binding cassette, sub-family B (MDR/TAP), member 9 | -0,326* | -0,062 | -0,177 | Cell Death and Survival, Visual System Development and Function, Cell-To-Cell Signaling and Interaction |
| Atpif1 | 25392 | ATPase inhibitory factor 1 | -0,122 | -0,208 | -0,173 | Energy Production, Lipid Metabolism |
| Ubap2 | 313169 | ubiquitin associated protein 2 | -0,089 | -0,301 | -0,173 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Zc3h7b | 315158 | zinc finger CCCH-type containing 7B | -0,233* | -0,115 | -0,17 | Cell Death and Survival, Visual System Development and Function, Cell-To-Cell Signaling and Interaction |
| Agrn | 25592 | agrin | -0,017 | -0,304 | -0,17 | Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Development |
| Or10h1 | 299553\| 299576 | olfactory receptor, family 10, subfamily H, member 1 | -0,289 | 0,034 | -0,144 | Other |
| Impg2 | 245919 | interphotoreceptor matrix proteoglycan 2 | -0,253* | -0,075 | -0,14 | Cell Death and Survival, Visual System Development and Function, Cell-To-Cell Signaling and Interaction |
| Ndnf | 312401 | neuron-derived neurotrophic factor | -0,395* | -0,026 | -0,138 | Cell Death and Survival, Visual System Development and Function, Cell-To-Cell Signaling and Interaction |
| Rpl31 | 64298 | ribosomal protein L31 | -0,077 | -0,201 | -0,13 | Energy Production, Lipid Metabolism |

**Supplementary Table 4. Top molecular and cellular functions, canonical pathways, molecules, and hepatotoxic functions for Mid PFNA + Mix. Values for Mid PFNA are presented for comparison.**

Molecular and cellular functions

| | Mid PFNA + Mix | | Mid PFNA | |
|---|---|---|---|---|
| Name | *p*-value | # molecules | *p*-value | # molecules |
| Energy production | 2E-23 - 2E-2 | 30 | 2E-24 - 3E-02 | 33 |
| Lipid metabolism | 2E-23 - 2E-2 | 58 | 2E-24 - 3E-02 | 59 |
| Small molecule biochemistry | 2E-23 - 2E-2 | 66 | 2E-24 - 3E-02 | 71 |
| Nucleic acid metabolism | 7E-11 - 2E-2 | 26 | 1E-6- 3E-02 | 26 |
| Molecular Transport | 3E-6 - 2E-2 | 55 | 2E-6 - 3E-02 | 26 |

Top canonical pathways

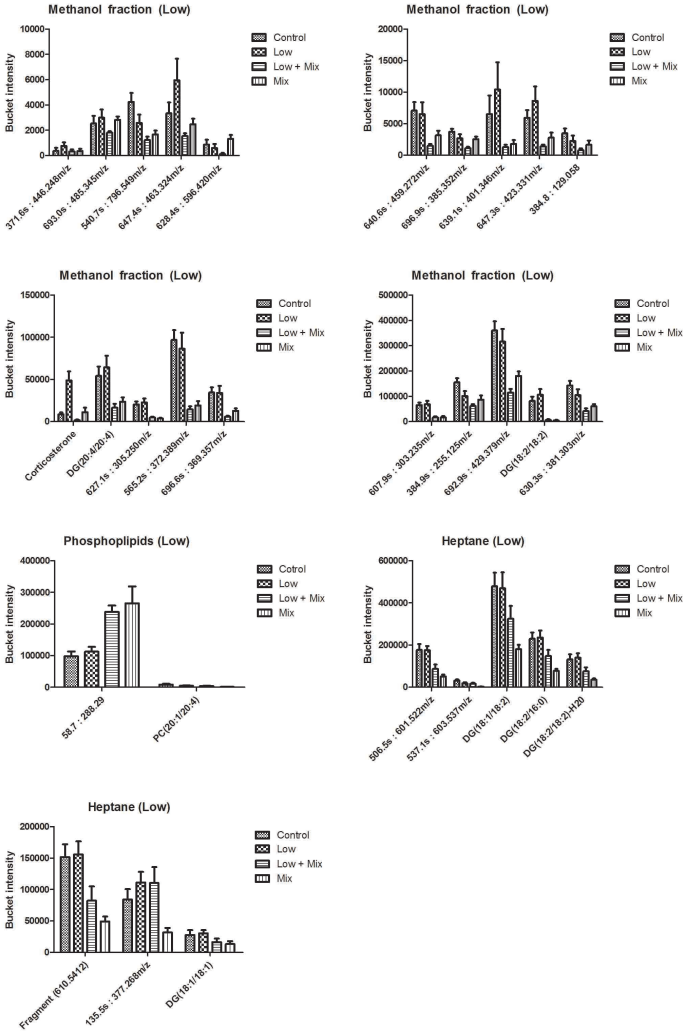
| | Mid PFNA + Mix | | Mid PFNA | |
|---|---|---|---|---|
| Name | *p*-value | Ratio | *p*-value | Ratio |
| Fatty acid β-oxidation I | 1E-11 | 9/29 | 3E-10 | 8/29 |
| Glutaryl-CoA degradation | 7E-10 | 6/11 | 4E-10 | 6/11 |
| Tryptophan Degradation III | 4E-8 | 6/19 | 2E-8 | 6/19 |
| LPS/IL-1 Mediated Inhibition of RXR Function | 2E-6 | 12/203 | 1E-6 | 12/203 |
| Ketogenesis | 3E-6 | 4/10 | 2E-6 | 4/10 |

Top molecules

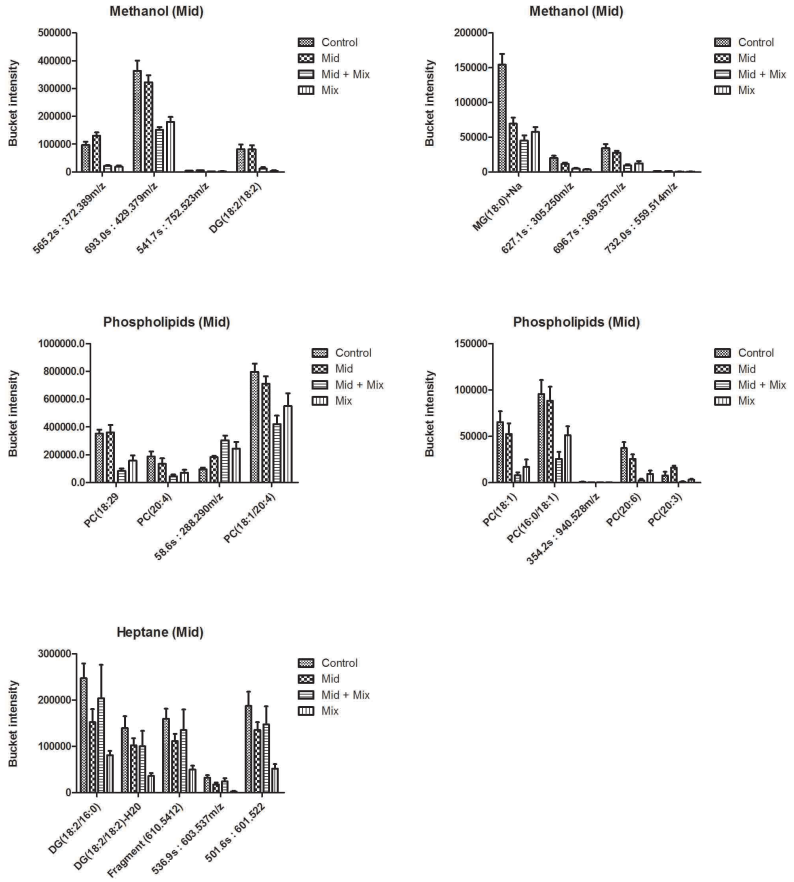
| | Log2 fold change up-/down-regulated | |
|---|---|---|
| Gene, description | Mid PFNA + Mix | Mid PFNA |
| Cyp4a11 | ↑2.568 | ↑2.419 |
| Cyp2b6 | ↑2.561 | ↑2.120 |
| Aldh1a1 | ↑2.532 | ↑2.077 |
| Vnn1 | ↑2.059 | ↑2.036 |
| Acot1 | ↑2.052 | ↑1.278 |
| Colq | ↑1.954 | ↑1.514 |
| Aig1 | ↑1.908 | ↑1.788 |
| Cd36 | ↑1.867 | ↑1.276 |
| Ehhadh | ↑1.860 | ↑1.702 |
| Ech1 | ↑1.856 | ↑1.512 |
| Fabp5 | ↓-1.508 | ↓-1.392 |
| Erich5 | ↓-1.372 | ↓-0.956 |
| Cxcl13 | ↓-1.173 | ↓-0.817 |
| Ca1 | ↓-1.132 | ↓-0.782 |
| Hsd11b1 | ↓-1.082 | ↓-0.446 |
| Phgdh | ↓-0.976 | ↓-1.271 |
| Mon1b | ↓-0.877 | ↓-0.317 |
| Isyna1 | ↓-0.825 | ↓-0.906 |
| Nfic | ↓-0.631 | ↓-0.262 |

| Top hepatotoxic functions | | | | |
|---|---|---|---|---|
| | Mid PFNA + Mix | | Mid PFNA | |
| Name | *p*-value | # molecules | *p*-value | # molecules |
| Liver steatosis | 3E-8 - 2E-1 | 16 | 1E-9 - 4E-2 | 17 |
| Hepatocellular peroxisome proliferation | 1E-4 - 7E-2 | 4 | 1E-04 - 7E-2 | 4 |
| Liver necrosis/cell death | 1E-2 - 4E-1 | 2 | 5E-1 | 2 |
| Liver proliferation | 2E-2 - 1E0 | 2 | NA | 0 |
| Liver hypertrophy | 3E-2 | 1 | 3E-2 | 1 |

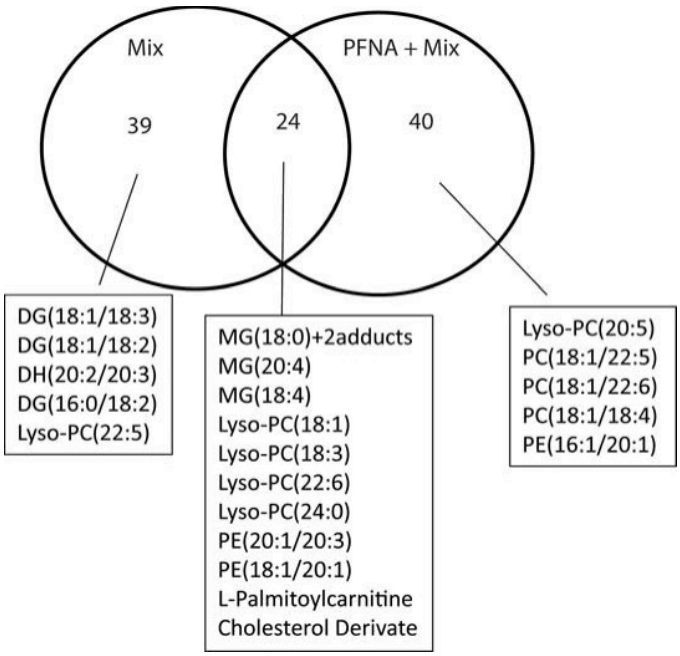
## Supplementary figures




**Supplementary Fig. 1 All significantly changed metabolites in the Low dose group. All metabolites are described by their retention time and m/z value**




**Supplementary Fig. 2 All significantly changed metabolites in the Mid dose group. All metabolites are described by the retention time and m/z value**




**Supplementary Fig. 3 Venn diagram showing the number of significant metabolites in the Mix group (t-test) and the PFNA + Mix groups (1-way ANOVA)($p \leq 0.05$). The boxes present the identified metabolites. 63 metabolites were significantly altered by Mix and 64 were significantly altered by PFNA + Mix. Twenty-four metabolites were altered in both groups; 12 were identified. The described metabolites are MG (Monoacylglycerol), PC (Phosphatidylcholine), PE (Phosphatidylethanolamine) and DG (diacylglycerol)**

## 5.6 Concluding remarks

In summary,

- Low PFNA administered alone affected plasma corticosterone levels; however, this effect was normalized upon co-administration of the Mix.
- Co-administration of Mix caused a 2.8-fold increase in plasma PFNA concentration at Low PFNA.
- Low PFNA ± Mix increased plasma levels of testosterone and dihydrotestosterone.
- High PFNA ± Mix caused steatosis.
- Changes on the metabolome mainly concerned lipids. Different lipids were affected with and without the presence of the Mix.
- Transcriptional changes were related to lipid metabolism, and the effects were driven by PFNA.

### 5.6.1 The Chemical Mixture

The investigations of the effects exerted by Mix outlined that statistically significant disturbances of hormone and metabolite levels occurred as a consequence of co-administration of Mix. This might be of concern, as the doses of the chemicals in Mix approached human high-end exposure levels. We do not have liver transcriptomics data from the animals treated with Mix alone and thus do not know which effects the Mix alone had on the livers. However, we observed borderline significant effects on the liver transcriptome after exposure to Mix + Low PFNA. Therefore, we might infer that the liver transcriptome was not readily affected by the Mix. This is consistent with the liver data presented in Paper II (Section 5.4); however, results on the metabolomics analysis (Manuscript IV) indicated effects on plasma phospholipids upon exposure to Mix alone.

The effects on androgen levels at Low PFNA seems critical as they may lead to morphological and/or functional changes in the adult male rat.

### 5.6.2 Integrated Data Analysis

Profiling of the liver transcriptome and the plasma metabolome has provided us with valuable insight into potential mechanisms through which PFNA and Mix act. The

methods were complementary in that they highlight biological changes at different levels (metabolite and transcriptional) and thereby provides insight into different aspects of the mode of action of PFNA and Mix. Our results indicate upregulation of hepatic genes involved in fatty acid β-oxidation and decreased plasma levels of various lipids. The metabolomics analysis revealed changes in the metabolome even at Low PFNA and also upon exposure to Mix, whereas the transcriptomics analysis appeared less sensitive, as no genes were significantly differentially misregulated at Low PFNA + Mix but only at higher PFNA doses. Therefore, metabolomics might be a more sensitive method for detecting changes caused by chemical exposure. On the other hand, the plasma metabolome changes rapidly in response to intake of food, stress and various other factors, whereas changes in the transcriptome occur more slowly. Therefore, without the results from the transcriptomics analysis, we would be less certain that the observed effects on the plasma lipid balance was a direct consequence of the chemical treatment or merely a result of the fed state of the animals.

In this study, we applied two 'omics' methods on two different body compartments (metabolomics on plasma and transcriptomics on liver tissue). Performing various 'omics' analyses on the same body compartment might provide even more insight into the mechanism underlying the observed effects compared to the setup in this project. For example, it would have been very useful to perform a metabolomics analysis on liver samples from the same animals as we had transcriptomics data for. This would provide a better foundation to understand mechanisms of action for PFNA, Mix, and the combination of the two.

Chapter 6

# Modeling Anogenital Distance from ToxCast Data

As described in Section 3.2, the main aim of the ToxCast program is to aid chemical safety evaluations by easing prioritization of chemicals for animal testing.

## 6.1 Background

For the phase I chemicals, several initial predictive models for reproductive [121] and developmental [122, 123] end-points have been built. With balanced accuracy (mean of sensitivity and specificity) (BA) in the range of 71% - 90%, the authors conclude that the ToxCast data can be used to build predictive models for defined toxicity endpoints for subsequent prioritization of chemicals for further testing.

As more data arrive with the completion of phase II, phase III, E1K, and phases of Tox21, the data analysis pipeline is being modified to fit data from all assay platforms. Therefore, as data undergoes slight modifications and more data arrive, the models need refinement as well. Nonetheless, successful modeling attempts have laid the ground for approaches for modeling *in vivo* toxicity endpoints using *in vitro* high-throughput screening (HTS) ToxCast data.

Drawing on expertise at the US EPA, this project aims at building a predictive model for a single reproductive endpoint, the so-called anogenital distance (AGD), using all phase I and II, E1K, and Tox21 chemicals and assays. As outlined in Section 2.1 there is increasing concern about human fertility, primarily based on the generally low human fertility and the increased incidence of TDS. Sufficient androgen action must occur within the time frame where male programing of the fetus occurs (the so-called male programing window) to ensure normal development of all male reproductive organs [124]. AGD has been recognized as a sensitive endpoint for assessment of exposure to EDCs during fetal life [125, 126]. In rodents, disruption of androgen-driven masculinization causing hypospadias and cryptorchidism has been correlated inversely to AGD [126]. Androgen-mediated regulation of AGD has also been indicated in humans [127–129]. Furthermore, AGD has been shown to be significantly reduced in boys with hypospadias [130], and moderately reduced in boys with cryptorchidism [4, 131], conditions that are caused by disrupted androgen action during fetal development. Additionally, studies in adult men have found AGD to be a predictor of testicular function, fertility, sperm quality, and prostate cancer risk [132, 133].

In 1998, AGD was included as an endpoint in the US EPA test guideline for the multi-generational reproductive study (MGR) [134] and was recently included in the OECD guideline for the extended one-generation reproductive toxicity study [135].

## 6.2 Aims and hypothesis

I took a data-driven approach to build a predictive model for reduced male AGD using ToxRefDB and ToxCast data, working from the hypothesis that AGD could be predicted from ToxCast data with sufficient accuracy, sensitivity, and specificity. The data-driven approach opened for the possibility that other assay endpoint(s) than those expected to predict AGD might add value to the predictive model.

## 6.3 Methods

This project applied a workflow similar to that described by Martin et al. [121], Sipes et al. [122], however slightly modified. Detailed descriptions of the applied methods (Figure 6.1), and the obtained results are given in the following.

All calculations were carried out in R [119] using the 'MASS' package [136] (for linear discriminant analysis (LDA)), and the 'gplots' package [137] for generating

heatmaps.

### 6.3.1 ToxRefDB

ToxRefDB was used to extract a positive and a negative set of chemicals. The set of AGD-positive chemicals was defined as chemicals in ToxRefDB being able to induce reduced AGD in male offspring. Chemicals that were tested in an MGR acceptable[1] guideline study conducted after 1998 but lacking effects on AGD were classified as AGD-negative chemicals (Supplementary Table 6.6).

### 6.3.2 Univariate Associations for Anogenital Distance

We performed a univariate analysis for each of the available assay readouts (from ToxCast phase I & II, E1K, and Tox21)[2] as an initial step in generating a list of assays or assay targets, that describes the *in vivo* classification sufficiently.

In a univariate analysis each variable (in this case each assay) is explored separately with respect to patterns of chemical activity in that variable/assay in relation to the *in vivo* activity for reduced AGD of the chemicals. We applied Fisher's exact test on a 2×2 contingency table to compare chemicals being positive/negative *in vivo* and *in vitro*. Chemicals that had a half maximal active concentration ($AC_{50}$) i.e. with positive hit calls according to ToxCast definitions (Section 3.2.3) were considered positive *in vitro*. Assays with a *p*-value $\leq 0.1$ and $\geq 3$ true positive chemicals were included for further modeling.

### 6.3.3 Assay Set Aggregation and Reduction

As described by Martin et al. [121], aggregating multiple related assays into a single composite assay value yields a more balanced and stable model. Therefore, we condensed assays based on common gene target and type of effect if at least two assays with a common target were retrieved from the univariate analysis. For instance, if PXR trans-activation was associated with reduced AGD we aggregated all assays for PXR activity. We calculated an aggregated assay set value (AA) for each chemical-assay set pair:

$$AA = \frac{1}{n} times \sum_{i=1}^{n} - \log_3 \frac{AC_{50i}}{1000}$$

[1] According to the staff at US EPA that curated the information in ToxRefDB (see Section 3.2.2)

[2] Accessed in Spring 2013.

where *n* is the number of assays pr. assay set, *i* represents each assay in the assay set, and $AC_{50}$ is the individual $AC_{50}$ value for each chemical-assay pair. A $log_3$ transformation was used over a $log_{10}$ transformation, which has been done in previous publications of ToxCast results [56], to enhance the scoring range between high- and low-potency active chemicals and to decrease the distance between active (i.e. achieving an $AC_{50}$ and defined as a hit in the assay) and inactive chemicals. Using this composite assay set, we developed a multivariate model using LDA.

### 6.3.4 Linear Discriminant Analysis

LDA is a statistical modeling method that gives a linear combination of features/variables which separates two or more classes of objects. In this case the features/variables are the composite assay sets and the classes are AGD-positive and -negative. Thus, the method looks for linear combinations of the variables that best explain the data.

### 6.3.5 External Validation Set

To test the predictability of the full model we searched the literature for chemicals causing reduced male AGD to obtain a set of AGD-positive and -negative chemicals not included in the training set (i.e. not available in ToxRefDB) but having been tested in the ToxCast battery of HTS assays.

## 6.4 Results and Discussion

The positive and negative sets, respectively, consisted of 13 chemicals with determined lowest observed adverse effect levels (LOAELs) for reduced male AGD (3 chemicals with LOAEL above 500 mg/kg BW/day). The negative set consisted of 71 chemicals that had been tested in an MGR acceptable guideline study post 1998 and did not cause reduced male AGD. All 84 chemicals had been tested in the ToxCast/Tox21 battery of *in vitro* assays (Supplementary Table 6.6). These data were extracted for use in the model.

### 6.4.1 Univariate Analysis

The univariate analysis resulted in 14 different assay readouts with $p < 0.1$ and $geq3$ true positive chemicals, where some were readouts for the same endpoint (e.g. PXR

transactivation and activation of transcription via PXR response element) (Figure 6.2). The endpoints represented in the list of assays most significantly associated to reduced AGD are androgen receptor (AR), estrogen receptor (ER), peripheral benzodiazepine receptor (PBR)[3], PPARγ, PXR, vitamin D receptor (VDR), and cell viability/cytotoxicity (Table 6.5). Cell viability was left out of the model since this is a very unspecific endpoint and including it worsened the model performance (data not shown).

We plotted the sensitivity (true positive rate) against the specificity (true negative rate) to visually assess the relevance of the univariate associations (Supplementary Figure 6.4). Data points located farthest away from the diagonal are considered more predictive for reduced AGD. Notably, all sensitivity values lie below 50% except for three assays targeting PXR. This indicates that with the exception of the PXR assays, no assay is better at predicting the positives than random sampling. The assay being most strongly associated to reduced AGD (lowest association *p*-value) is a PXR trans-activation assay ($p = 7 \times 10^5$, sensitivity = 77%), and the least strongly associated assays (highest association *p*-values, yet $leq$ 0.1) are two ERα dimerization assays with a sensitivity of 31% ($p = 0.10$).

These poor univariate associations might be a result of 1) the unbalanced dataset (13 AGD-positive and 71 AGD-negative chemicals), 2) the small size of the dataset, 3) that the assays currently available in ToxCast do not represent pathways through which the AGD-positive chemicals exert their action *in vivo* (e.g. effects on steroidogenesis), or 4) that the chemicals require metabolic activation to exert their *in vitro* effect, and most ToxCast assays are deficient of metabolic capacity. One of the AGD positive chemicals (monobenzyl phthalate) was inactive in all of the assays resulting from the univariate analysis. This is, however, contradictory to another study showing activity of this phathalate on PPARα, PPARγ, and PPARδ [138].

### 6.4.2 Linear Discriminant Analysis

We generated the composite assay sets by aggregating similar assays (Table 6.1). Hereby, we allow biologically meaningful assays to be included in the full model, despite a potentially low number of chemicals being active in the individual assays. We distinguished between agonist, antagonist, and binding assays where applicable. Assay endpoints related to cytotoxicity (e.g. cell cycle arrest assays) were not included in the aggregated assay sets.

The aggregated assay sets comprise VDR activation, AR and ER antagonist, activation/agonist, and binding, PXR activation/binding, PBR binding, and PPARγ

[3]Formally known as translocator protein (TSPO)

activation/binding assays (Figure 6.3). As in the assays selected from the univariate analysis, monobenzyl phthalate was inactive in all assay sets. VDR activation and AR antagonism received the highest positive correlation weight factors in the LDA model on the full dataset, whereas ER binding and ER antagonism were the strongest negatively associated assay sets. The model thereby suggests, that the AGD positive chemicals exert their action via VDR activation and/or AR antagonism and chemicals binding to the ER and/or acting as antagonists on ER are less likely to cause reduced AGD.

Indeed, VDR has been suggested to affect male reproduction [139]. If not via a direct mechanism, it is biologically reasonable to suggest an indirect action of vitamin D on testosterone production via calcium and phosphate homeostasis due to their role as $2^{nd}$ messengers.

The role of the other targets selected by univariate analysis (PXR, PBR, and PPAR$\gamma$) in male reproduction remain unresolved. However, it seems reasonable to expect PBR to play a role in development of reduced AGD via altered transport of cholesterol into mitochondria. Following transport of cholesterol into mitochondria the initial and rate-limiting step of steroidogenesis (conversion of cholesterol to pregnenolone) takes place. However, in mice, the absence of PBR has been shown not to affect steroid hormone biosynthesis [140]. Testosterone is inactivated by CYP3A4 and CYP2C9 [141] and expression of those enzymes is in part regulated by PXR. Therefore, chemicals activating PXR-regulated transcription might cause reduced testosterone levels [142]. Corton and Lapinskas [143] discuss the involvement of PPARs on the development of the male reproductive tract. There is little evidence to support the expression of PPAR$\gamma$ in testis, whereas the $\alpha$ and $\beta$ subtypes are expressed in neonatal or adult Sertoli and Leydig cells [143]. Based on this observation it could be of value to include the other PPAR subtypes in the AGD model. Despite its lack of expression in testis, evidence that PPAR$\gamma$ is involved in the regulation of steroidogenesis and thus the levels of sex hormones exists from *in vitro* studies of porcine ovarian follicles [144].

Table 6.1: The 10 composite assay sets composing the AGD predictive signature. There are 6 positive and 4 negative predictors, and each assay set is composed of minimum three assays.

| Assay set | Individual assay | Weight |
|---|---|---|
| | Positive predictors | |
| VDR | ATG_VDRE_CIS<br>ATG_VDR_TRANS | 0.85 |

| Assay set | Individual assay | Weight |
|---|---|---|
| AR antagonist | Tox21_AR_BLA_Antagonist_ch1<br>Tox21_AR_BLA_Antagonist_ch2<br>Tox21_AR_BLA_Antagonist_ratio<br>Tox21_AR_LUC_MDAKB2_Antagonist | 0.46 |
| AR binding | NVS_NR_cAR<br>NVS_NR_hAR<br>NVS_NR_rAR | 0.21 |
| AR agonist | ATG_AR_TRANS<br>ATG_AR_TRANS_perc<br>OT_AR_ARE_LUC_Agonist_1440<br>OT_AR_ARSRC1_0480<br>OT_AR_ARSRC1_0960<br>Tox21_AR_BLA_Agonist_ch1<br>Tox21_AR_BLA_Agonist_ch2<br>Tox21_AR_BLA_Agonist_ratio<br>Tox21_AR_LUC_MDAKB2_Agonist | 0.20 |
| ER agonist | ATG_ERa_TRANS<br>ATG_ERa_TRANS_perc<br>ATG_ERE_CIS<br>ATG_ERE_CIS_perc<br>OT_ER_ERaERa_0480<br>OT_ER_ERaERa_1440<br>OT_ER_ERaERb_0480<br>OT_ER_ERaERb_1440<br>OT_ER_ERbERb_0480<br>OT_ER_ERbERb_1440<br>OT_ERa_ERE_LUC_Agonist_1440<br>OT_ERa_GFPERaERE_0120<br>OT_ERa_GFPERaERE_0480<br>Tox21_ERa_BLA_Agonist_ch1<br>Tox21_ERa_BLA_Agonist_ch2<br>Tox21_ERa_BLA_Agonist_ratio<br>Tox21_ERa_LUC_BG1_Agonist | 0.11 |
| PXR | ATG_PXR_TRANS<br>ATG_PXR_TRANS_perc<br>ATG_PXRE_CIS<br>ATG_PXRE_CIS_perc<br>NVS_NR_hPXR | 0.046 |

| Assay set | Individual assay | Weight |
|---|---|---|
| | Negative predictors | |
| ER binding | NVS_NR_bER<br>NVS_NR_hER<br>NVS_NR_mERa | -0.29 |
| ER antagonist | OT_ERa_ERE_LUC_Antagonist_1440<br>OT_ERb_ERE_LUC_Antagonist_1440<br>Tox21_ERa_BLA_Antagonist_ch1<br>Tox21_ERa_BLA_Antagonist_ch2<br>Tox21_ERa_BLA_Antagonist_ratio<br>Tox21_ERa_LUC_BG1_Antagonist | -0.12 |
| PBR | NVS_MP_hPBR<br>NVS_MP_rPBR | -0.095 |
| PPARγ | ATG_PPARg_TRANS<br>ATG_PPARg_TRANS_perc<br>NVS_NR_hPPARg<br>OT_PPARg_PPARgSRC1_1440<br>Tox21_PPARg_BLA_Agonist_ch1<br>Tox21_PPARg_BLA_Agonist_ch2<br>Tox21_PPARg_BLA_Agonist_positive<br>Tox21_PPARg_BLA_Agonist_ratio | -0.012 |

Modeling the data using the 10 assay sets yielded a training BA of 78% with a *p*-value of $1 \times 10^{-6}$. The sensitivity and specificity were 62% and 94%, respectively. Validating the model with an external dataset retrieved from the public literature and not part of ToxRefDB (Table 6.3) resulted in a sensitivity and specificity of 33% and 63%, respectively (Table 6.2). The sensitivity should, preferably, be above 60% as this measure indicates how good the model is at predicting true positives. Thus, a sensitivity $< 50\%$ means that the model is worse than random sampling at predicting the true positives. The drop in model performance following validation indicates that the model has been over-fitted to the training data. This is likely to happen when the size of the training data is small or when the data is imbalanced, which are both the case in this model. Furthermore, chemicals with poor correlations between the *in vivo* and *in vitro* effects introduce uncertainty to the model. Also, as mentioned earlier, the assays currently available in the ToxCast data might not fully represent the pathways through which some of the chemicals cause reduced AGD. This is likely to be the case, since seven of the 13 AGD-positive chemicals are phthalates, and phthalates are thought to cause reduced AGD by inhibiting fetal testosterone synthesis [145]. We did, indeed, attempt

Table 6.2: Statistics from full model and prediction of an external validation set.

| | Model statistics | |
|---|---|---|
| | Full model | Validation |
| Sensitivity | 62 % | 33 % |
| Specificity | 94 % | 63 % |
| Balanced accuracy | 78 % | 48 % |
| Accuracy | 89 % | 50 % |
| Precision | 67 % | 40 % |
| $p$-value | $1 \times 10^{-6}$ | NS |

NS: Not significant

to add the recent addition to the ToxCast data, the HTS human adrenocortical carcinoma cell line (H295R) assay[4] to the model. Unfortunately, only half of the chemicals in the AGD dataset had been tested; however, as more chemicals get tested in the H295R assay modeling should be attempted again. Despite the lack of sensitivity of the model, this model could still prove useful in giving potential AGD negative chemicals low priority for further testing. This seems particularly relevant when facing thousands of chemicals potentially requiring animal testing for their potential to cause reproductive toxicity.

Additionally, Martin et al. [121] defined chemicals without a LOAEL for reproductive toxicity or LOAEL $> 500$ mg/kg BW/day as negative, whereas I did not incorporate that cut-off in this model. Furthermore, where the previous models and this model all use univariate analysis for assay selection we use different cut-offs for which assays to use for assay aggregation.

The method used to generate the model presented in this chapter deviates from the models presented by Martin et al. [121], Sipes et al. [122] in various aspects, which leaves this AGD predictive signature at a premature stage. The signature presented by Martin et al. [121] excluded chemicals in the dataset that were broadly inactive in the ToxCast assays. The rationale for excluding these chemicals was that several reasons for their lack of *in vitro* might exist (e.g. chemical degradation, aqueous insolubility, lack of metabolic activation, or volatility) which renders them bad predictors of their *in vivo* activity. Of the chemicals in the AGD training set, 40 (approx. 48%) fell below this criteria. Excluding those 40 chemicals would leave behind a very small dataset.

[4] Unpublished data

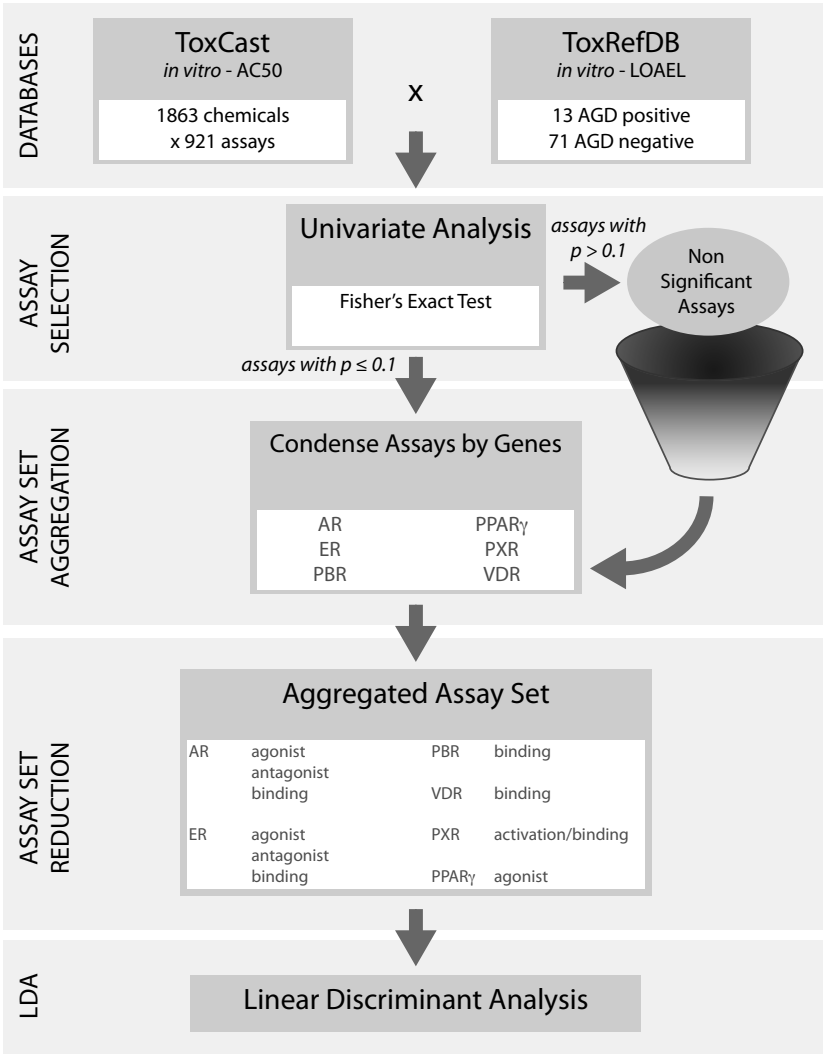



Figure 6.1: Workflow for generation of the predictive model for reduced male AGD. ToxCast *in vitro* HTS data consisted of 1863 chemicals tested in 921 assays. ToxRefDB entries for reduced male AGD consisted of 13 positive and 71 negative chemicals. Univariate analysis was performed to obtain an initial assay selection, which was followed by assay aggregation by condensing assays by gene. Aggregated assay set values were calculated for each assay set/chemical pair. These were used as input for LDA. Modified from [121, 122]

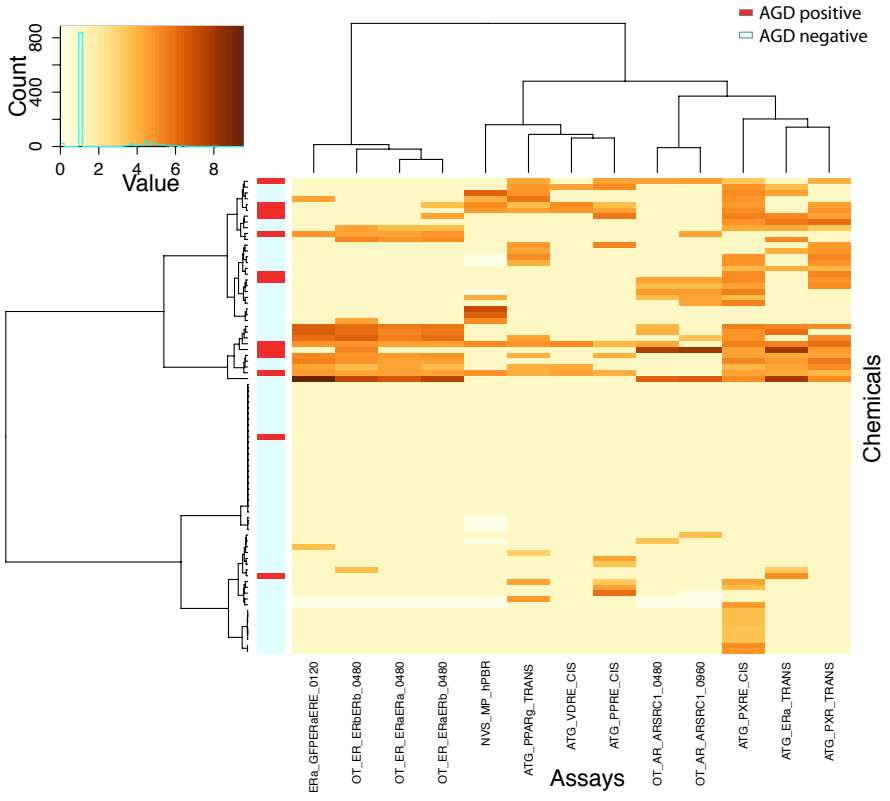



Figure 6.2: Heatmap showing the correlations between *in vivo* and *in vitro* effects for AGD. Red and blue bars indicate AGD positive and negative chemicals, respectively. The assays in the plot were significant at $\alpha < 0.1$ and were associated with $\geq 3$ AGD positive chemicals.

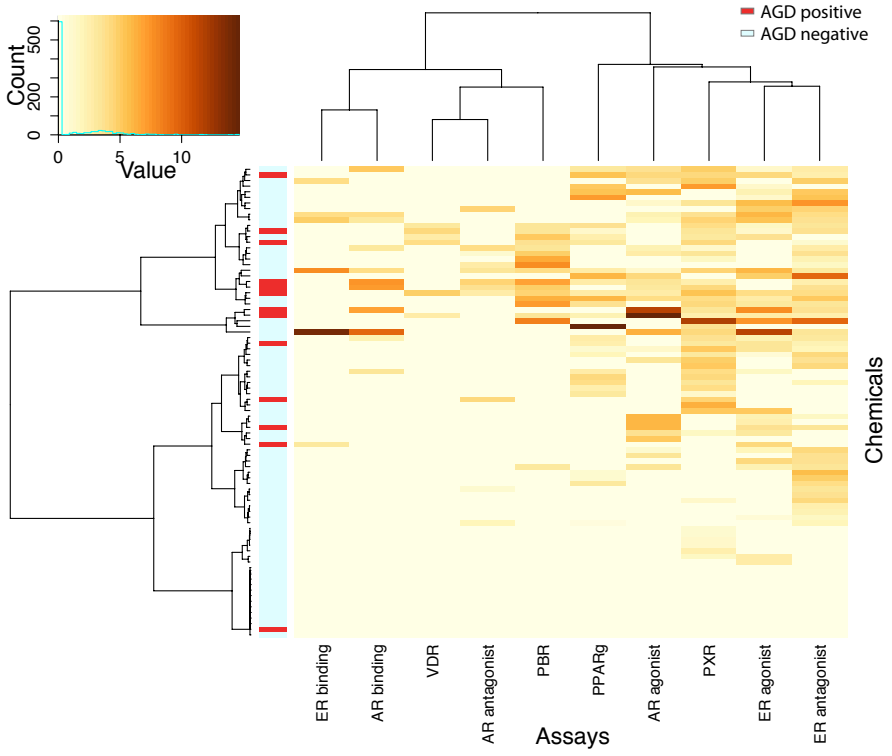



Figure 6.3: Heatmap showing the activities of the chemicals in the training data in terms of the aggregated assay sets. Red and blue bars indicate AGD positive and negative chemicals, respectively.

Table 6.3: Chemicals comprising the external validation set retrieved from literature. LOAELs are provided where applicable. The predictions are given along with the classification of the chemicals.

| CASRN | Chemical name | LOAEL | Reference | Evidence of reproductive toxicity | Predicted reproductive toxicant |
|---|---|---|---|---|---|
| 106325-08-0 | Epoxiconazole | 15 mg/kg/day | [146] | Yes | No |
| 65277-42-1 | Ketoconazole | 50 mg/kg/day | [147] | Yes | Yes |
| 446-72-0 | Genistein | 5 ppm | [148] | Yes | No |
| 330-55-2 | Linuron | 75 mg/kg/day | [149] | Yes | Yes |
| 67747-09-5 | Prochloraz | 50 mg/kg/day | [150] | Yes | No |
| 103-90-2 | Paracetamol | 150 mg/kg/day | [151] | Yes | No |
| 60207-90-1 | Propiconazole | 2500 ppm | [147, 152] | No | Yes |
| 107534-96-3 | Tebuconazole | NA | [146, 147] | No | Yes |
| 117-84-0 | Di-n-octyl phthalate | NA | [153] | No | No |
| 120-47-8 | Ethylparaben | NA | [154] | No | No |
| 76-44-8 | Heptachlor | NA | [155] | No | No |
| 58-89-9 | Lindane | NA | [156] | No | Yes |
| 556-67-2 | Octamethylcyclotetrasiloxane | NA | [157] | No | No |
| 131-11-3 | Dimethyl phthalate | NA | [158] | No | No |

## 6.5 Concluding remarks

In this project, I trained a predictive model for 'reduced male AGD'. The predictive model correctly predicted 62% of the AGD-positive chemicals used to train the model, however, when testing the predictability on an external validation set the model performance was substantially reduced. As discussed above there might be several reasons why this model performs inadequately. Including more assay endpoints as they get added to the ToxCast assay set might improve the model. Furthermore, adding more chemicals and/or expanding the *in vivo* endpoint being predicted might also improve the model performance. In the event of planning a study where reduced AGD is a 'desired' effect obtained from fetal exposure to the chemical in rats this predictive model can aid in the prioritization of a long list of chemicals. Given the poor model performance, however, there is a 50% chance (accuracy) of getting a wrong prediction (false negative or false positive) when predicting the effects of chemicals not included in the training set. Theoretically, this is no better than random sampling given a balanced dataset. However, when prioritizing thousands of chemicals for animal testing in reproductive toxicity tests predictions of AGD inactive chemicals can be used for assigning low priority to those chemicals.

## 6.6 Supplementary Information

Table 6.4: AGD-active and -inactive chemicals included in the predictive signature for AGD.

| CASRN | Chemical name | LOAEL (mg/kg BW/day) |
|---|---|---|
| 51-52-5 | 6-Propyl-2-thiouracil | 0.1 |
| 13311-84-7 | Flutamide | 2 |
| 32809-16-8 | Procymidone | 13 |
| 50471-44-8 | Vinclozolin | 15 |
| 84-61-7 | Dicyclohexyl phthalate | 90 |
| 85-68-7 | Butyl benzyl phthalate | 100 |
| 36734-19-7 | Iprodione | 120 |
| 84-75-3 | Dihexyl phthalate | 250 |
| 2528-16-7 | Monobenzyl phthalate | 250 |
| 117-81-7 | Di(2-ethylhexyl) phthalate | 390 |
| 84-74-2 | Dibutyl phthalate | 530 |
| 84-69-5 | Diisobutyl phthalate | 600 |
| 57-85-2 | Testosterone propionate | 5000[5] |
| 208465-21-8 | Mesosulfuron-methyl | Inactive |
| 80-05-7 | Bisphenol A | Inactive |
| 149877-41-8 | Bifenazate | Inactive |
| 173159-57-4 | Foramsulfuron | Inactive |
| 144550-36-7 | Iodosulfuron-methyl-sodium | Inactive |
| 361377-29-9 | Fluoxastrobin | Inactive |
| 148477-71-8 | Spirodiclofen | Inactive |
| 134605-64-4 | Butafenacil | Inactive |
| 135410-20-7 | Acetamiprid | Inactive |
| 175013-18-0 | Pyraclostrobin | Inactive |
| 120116-88-3 | Cyazofamid | Inactive |
| 161326-34-7 | Fenamidone | Inactive |
| 210880-92-5 | Clothianidin | Inactive |
| 2310-17-0 | Phosalone | Inactive |
| 5598-13-0 | Chlorpyrifos-methyl | Inactive |
| 60-51-5 | Dimethoate | Inactive |
| 709-98-8 | Propanil | Inactive |
| 51707-55-2 | Thidiazuron | Inactive |

[5]parts per million

| CASRN | Chemical name | LOAEL (mg/kg BW/day) |
|---|---|---|
| 153719-23-4 | Thiamethoxam | Inactive |
| 188425-85-6 | Boscalid | Inactive |
| 25606-41-1 | Propamocarb hydrochloride | Inactive |
| 116714-46-6 | Novaluron | Inactive |
| 181274-15-7 | Propoxycarbazone-sodium | Inactive |
| 199119-58-9 | Trifloxysulfuron-sodium | Inactive |
| 219714-96-2 | Penoxsulam | Inactive |
| 10265-92-6 | Methamidophos | Inactive |
| 188489-07-8 | Flufenpyr-ethyl | Inactive |
| 156052-68-5 | Zoxamide | Inactive |
| 63-25-2 | Carbaryl | Inactive |
| 283594-90-1 | Spiromesifen | Inactive |
| 163520-33-0 | Isoxadifen-ethyl | Inactive |
| 99-99-0 | 4-Nitrotoluene | Inactive |
| 87-86-5 | Pentachlorophenol | Inactive |
| 51-28-5 | 2,4-Dinitrophenol | Inactive |
| 68-12-2 | N,N-Dimethylformamide | Inactive |
| 120-83-2 | 2,4-Dichlorophenol | Inactive |
| 108-95-2 | Phenol | Inactive |
| 119-61-9 | Benzophenone | Inactive |
| 50-29-3 | p,p'-DDT | Inactive |
| 140-66-9 | 4-(1,1,3,3-Tetramethylbutyl)phenol | Inactive |
| 52-51-7 | Bronopol | Inactive |
| 108-46-3 | Resorcinol | Inactive |
| 1806-26-4 | 4-Octylphenol | Inactive |
| 822-06-0 | 1,6-Diisocyanatohexane | Inactive |
| 3871-99-6 | Potassium perfluorohexanesulfonate | Inactive |
| 84-66-2 | Diethyl phthalate | Inactive |
| 26172-55-4 | 5-Chloro-2-methyl-3(2H)-isothiazolone | Inactive |
| 1445-75-6 | Diisopropyl methylphosphonate | Inactive |
| 51229-78-8 | 1-(cis-3-Chloroallyl)-3,5,7-triaza-1-azoniaadamantane chloride | Inactive |
| 2634-33-5 | 1,2-Benzisothiazolin-3-one | Inactive |
| 7487-94-7 | Mercuric chloride | Inactive |
| 104-51-8 | Butylbenzene | Inactive |
| 50-28-2 | 17beta-Estradiol | Inactive |
| 1461-22-9 | Tributyltin chloride | Inactive |
| 97-54-1 | Isoeugenol | Inactive |

| CASRN | Chemical name | LOAEL (mg/kg BW/day) |
|---|---|---|
| 2795-39-3 | Potassium perfluorooctanesulfonate | Inactive |
| 7747-35-5 | 5-Ethyl-1-aza-3,7-dioxabicyclo[3.3.0]octane | Inactive |
| 556-67-2 | Octamethylcyclotetrasiloxane | Inactive |
| 29420-49-3 | Potassium nonafluoro-1-butanesulfonate | Inactive |
| 1934-21-0 | FD&C yellow 5 | Inactive |
| 111-30-8 | Glutaraldehyde | Inactive |
| 26530-20-1 | Octhilinone | Inactive |
| 84852-15-3 | 4-Nonylphenol, branched | Inactive |
| 30516-87-1 | 3'-Azido-3'-deoxythymidine | Inactive |
| 6422-86-2 | Bis(2-ethylhexyl) terephthalate | Inactive |
| 3825-26-1 | Ammonium perfluorooctanoate | Inactive |
| 64359-81-5 | 4,5-Dichloro-2-octyl-3(2H)-isothiazolone | Inactive |
| 319-85-7 | beta-1,2,3,4,5,6-Hexachlorocyclohexane | Inactive |
| 181274-17-9 | Flucarbazone-sodium | Inactive |
| 210631-68-8 | Topramezone | Inactive |
| 165252-70-0 | Dinotefuran | Inactive |

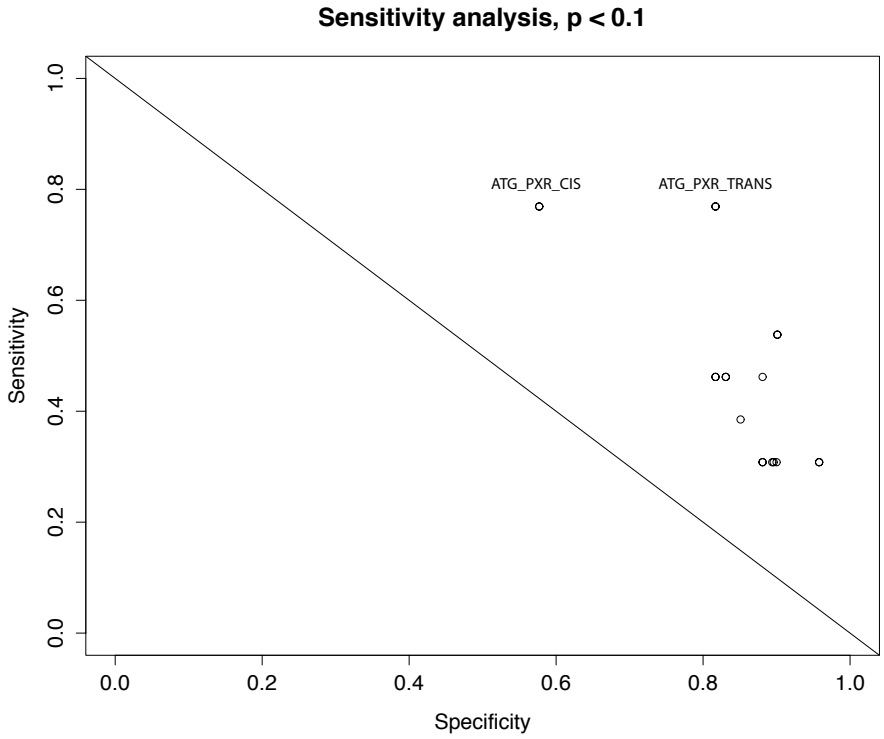



Figure 6.4: Univariate associations between ToxRefDB AGD reductions and ToxCast *in vitro* HTS assay data. Data points in the top right corner have greater diagnostic value. The line indicates random sampling.

Table 6.5: Univariate associations between AGD retrieved from ToxRefDB and the ToxCast HTS *in vitro* data. SENS = sensitivity, SPEC = specificity, BA = balanced accuracy (the mean of the sensitivity and specificity), ACC = accuracy, and OR = the odds ratio. $p$-values results from Fisher's exact test and $p$-values $\leq 0.1$ are presented.

| | | % | | | | | |
|---|---|---|---|---|---|---|---|
| Biological target | Assay | SENS | SPEC | BA | ACC | OR | $p$-value |
| AR | OT_AR_ARSRC1_0480 | 31 | 90 | 60 | 80 | 3.8 | 0.07 |
| | OT_AR_ARSRC1_0960 | 31 | 90 | 60 | 80 | 3.8 | 0.08 |
| CellCycle | APR_CellCycleArrest_24hr_dn | 25 | 94 | 59 | 83 | 5 | 0.07 |
| ER | OT_ER_ERaERb_0480 | 46 | 88 | 67 | 81 | 6.3 | 0.009 |
| | OT_ER_ERbERb_0480 | 39 | 85 | 62 | 78 | 3.6 | 0.06 |
| | ATG_ERa_TRANS | 46 | 82 | 64 | 76 | 3.8 | 0.06 |
| | OT_ER_ERaERa_0480 | 31 | 88 | 59 | 79 | 3.3 | 0.1 |
| | OT_ERa_GFPERaERE_0120 | 31 | 88 | 60 | 79 | 3.3 | 0.1 |
| PBR | NVS_MP_hPBR | 31 | 90 | 60 | 80 | 4 | 0.07 |
| PPARγ | ATG_PPRE_CIS | 54 | 90 | 72 | 85 | 11 | 0.0008 |
| | ATG_PPARg_TRANS | 46 | 83 | 65 | 77 | 4.2 | 0.03 |
| PXR | ATG_PXR_TRANS | 77 | 82 | 79 | 81 | 15 | $7 \times 10^{-5}$ |
| | ATG_PXRE_CIS | 77 | 58 | 67 | 61 | 4.6 | 0.03 |
| VDR | ATG_VDRE_CIS | 31 | 96 | 63 | 86 | 10 | 0.01 |

# Part III

# Epilogue

Chapter 7

# Concluding Remarks

## 7.1 Summary/Conclusion

The overall focus of this thesis was to apply existing computational systems biology methods for evaluation and prediction of human effects of chemicals and to group the chemicals according to their proposed mechanism of action. The aim was to evaluate if this would complement classical toxicological (*in vitro* and *in vivo*) investigations.

In project I (Integrative Systems Biology), we applied a computational approach published elsewhere [72]. With that approach we predicted human health effects of exposure to the five pesticides under investigation. These predictions further facilitated a grouping of the chemicals that was in concordance with effects of the chemicals on experimental animals. This approach is relatively fast to carry out and is cheap in that it does not require any further experimental work. It provides a means to obtain predictions of human health effects of chemicals. This contrasts the classical toxicological *in vivo* investigations, where inferences of effects are made from animal to man. Thereby, it complements traditional toxicological investigations and sheds light on potentially interesting endpoints to keep in mind for further research. The approach serves as a good method for 1) obtaining an overview of existing data on the chemical(s) of interest and 2) generating hypothe-

ses on mechanisms and modes of action of chemicals for subsequent validation in e.g. animal studies and 3) ideas of which chemicals to monitor in biomonitoring studies.

In project II (Experimental Studies of a Chemical Mixture), we investigated the effects of a chemical mixture of 14 environmental chemicals and food constituents at high-end human exposure levels with and without the addition of PFNA in a sub-acute animal study with rats. The project was divided into two sub-projects; one hypothesis-driven and one data-driven each resulting in a paper. The paper on the hypothesis-driven sub-project describes pathological findings along with gene expression on a selected set of genes in various tissues and the levels of transcribed protein for a small number of proteins aiming at highlighting potential mechanisms of action of the chemical mixture as well as PFNA. The data-driven project has character of a profiling study in which changes in the plasma metabolome and the liver transcriptome were described in detail. From this project, we learned that the 'omics' methods used in this study can 1) shed light on the mechanism(s) of action for the investigated chemicals by being complementary to one another, and 2) aid the traditional toxicological investigation, which has a more targeted character.

Project III (Modeling Anogenital Distance from ToxCast Data) describes the work that I did in collaboration with the US EPA. In this project, I applied already developed approaches [121, 122] to build a predictive signature for reduced AGD - a measure for endocrine disruption during fetal life. The approach is data-driven in that all ToxCast data are used in an unbiased manner for the initial assay selection. The approach further allows for inclusion of assays hitting targets know to be associated with the endpoint in question, thereby making the approach a combination of hypothesis- and data-driven. Despite the dubious performance of the model on the external validation set I believe that the approach is still valuable for highlighting potential mechanisms through which the chemicals cause the endpoint in question. This is exemplified by the phthalates. Their mechanism of action for causing reduced AGD is not entirely clear. However, this modeling effort points to PXR and PPAR$\gamma$ activation as potential mechanisms through which the phthalates cause reduced AGD.

The methods and approaches applied to toxicology in this PhD project have proven valuable as a means to supplement classical toxicological investigations. Different approaches are good for different purposes but common for all the applied approaches is that they feed into the idea-generating phase of the planning process of further studies (Figure 3.2) and complements evaluation of classical toxicological *in vivo* and *in vitro* studies.

## 7.2 Limitations of the Applied Methods

### 7.2.1 Lack of Data

Computational systems biology is highly reliant on existing data. As exemplified with epoxiconazole in the Integrative Systems Biology project, predictions of potential health effects of that chemical was not possible due to very limited data. Whether the lack of data reflects that the chemical is inactive in biological systems or that it has not been fully investigated is not clear from the analysis. Lack of data might also be part of the explanation why the predictive signature for reduced AGD performed sub-optimally. It is generally accepted that reduced AGD could be mediated by disrupted androgen signaling during fetal development. One way of disrupting androgen signaling is decreasing the androgen levels. Therefore, it is reasonable to expect improved performance of the model by inclusion of data from the H295R steroidogenesis assay. For this project, however, that was not possible as the overlap between chemicals tested in the H295R assay and the chemicals classified as positive or negative *in vivo* was too small to base a model upon. Therefore, lack of data forms a limitation on the applied methods that 1) reduces the applicability domain of the generated models, and 2) reduces the reliability in the model performances. However, the quantity of data is not the sole limitation of the applied methods.

### 7.2.2 Data Quality

In computer science, it is well known that computers will unquestioningly process nonsensical data (garbage in) and produce nonsensical output (garbage out). This is also phrased "garbage in, garbage out". When working with computational toxicology/systems biology it is important to keep the quality of the input data in mind. In the Integrative Systems Biology project we relied on the quality control procedure of the databases from which we retrieved data. The data in CTD are manually curated, which provides a means of assuring the quality of the data. The other database used in that project, ChemProt, retrieves data from various databases and has a built-in quality assurance protocol. In the Experimental Studies of a Chemical Mixture project, I assured good quality of the microarray data by reviewing the quality control reports that came with the microarray data. Leaving out this step could potentially introduce bias to the statistical and pathway analyses. The subsequent pathway analyses relied on canonical pathway and disease information. Lastly, in the Modeling Anogenital Distance from ToxCast Data project, I relied on the quality assurance protocol implemented in the ToxCast data analysis workflow

and the ToxRefDB data review process. However, whether the means taken to assure the quality of the data are sufficient remains a matter of debate.

## 7.3 Perspectives

As highlighted above, the methods applied in this thesis have proven valuable supplements to classical toxicological investigations. Given the challenges that regulatory toxicology is currently facing (Section 2.2) application of computational systems biology methods might prove useful when deciding which chemicals to test thoroughly in classical toxicity tests and which endpoints to investigate. Developing tools for prioritizing chemicals for further testing is the aim of the US EPA ToxCast program. An expansion of that program to face the challenges of chemical mixtures would, in my perspective, ease the process of investigating the potential risk(s) associated with exposure to relevant chemical mixtures.

Despite the usefulness and broad applicability of computational methods in toxicology it is unlikely to become the sole method for assessing the risk of chemical exposure to human health and the environment. Only in the hypothetical case where all biological processes and their interactions are fully understood will it be possible to predict the effects of chemicals on human health and the environment.

# Bibliography


[1] Monosson E. Chemical Mixtures: Considering the Evolution of Toxicology and Chemical Assessment. *Environmental Health Perspectives*, 113(4):383–390, October 2004.

[2] Wohlfahrt-Veje C, Boisen K. A, Boas M, Damgaard I. N, Kai C. M, et al. Acquired cryptorchidism is frequent in infancy and childhood. *International journal of andrology*, 32(4):423–8, August 2009.

[3] Wohlfahrt-Veje C, Andersen H. R, Schmidt I. M, Aksglaede L, Sø rensen K, et al. Early breast development in girls after prenatal exposure to non-persistent pesticides. *International journal of andrology*, 35(3):273–82, June 2012.

[4] Swan S. H, Main K. M, Liu F, Stewart S. L, Kruse R. L, et al. Decrease in anogenital distance among male infants with prenatal phthalate exposure. *Environmental health perspectives*, 113(8):1056–61, August 2005.

[5] Weidner I. S, Mø ller H, Jensen T. K, and Skakkebaek N. E. Cryptorchidism and hypospadias in sons of gardeners and farmers. *Environmental health perspectives*, 106(12):793–6, December 1998.

[6] European Food Safety Authority. International frameworks dealing with human risk assessment of combined exposure to multiple chemicals. *EFSA Journal*, 11(7):1–69, 2013.

[7] National Research Council. *Toxicity Testing in the 21st Century: A Vision and a Strategy*. The National Academies Press, 2007. ISBN 9780309151733.

[8] Andersen M. E and Krewski D. Toxicity testing in the 21st century: bringing the vision to life. *Toxicological sciences*, 107(2):324–30, February 2009.

[9] Krewski D, Acosta D, Andersen M, Anderson H, Bailar J. C, et al. Toxicity testing in the 21st century: a vision and a strategy. *Journal of toxicology and environmental health. Part B, Critical reviews*, 13(2-4):51–138, February 2010.

[10] FVM. Fødevareforlig 2.0, 2010. URL `http://fvm.dk/fileadmin/user_upload/FVM.dk/Dokumenter/Foedevarer/Kontrol/foedevareforlig20_juni2010.pdf`.

[11] Timbrell J. *Principles of Biochemical Toxicology*. CRC Press, third edition, 2000. ISBN 0-7484-0736-7.

[12] Berg J. M, Tymoczko J. L, and Stryer L. *Biochemistry*. W. H. Freeman and Company, New York, 6th edition, 2006. ISBN 978-0-7167-8724-2.

[13] IPCS/WHO. Global assessment of the state-of-the-science of endocrine disruptors. Technical report, International Programme on Chemical Safety, World Health Organization and United Nations Environment Programme, Geneva, 2002. URL `http://www.who.int/ipcs/publications/new_issues/endocrine_disruptors/en/`.

[14] Toppari J, Adamsson A, Boas M, Juul A, Main K. M, et al. Endocrine disrupters and child health Possible developmental early effects of endocrine disrupters on child health. Technical report, World Health Organisation, 2012. URL `http://www.who.int/ceh/publications/endocrine_disrupters_child/en/`.

[15] Jø rgensen N, Andersen A. G, Eustache F, Irvine D. S, Suominen J, et al. Regional differences in semen quality in Europe. *Human reproduction*, 16(5): 1012–9, May 2001.

[16] Swan S. H, Brazil C, Drobnis E. Z, Liu F, Kruse R. L, et al. Geographic Differences in Semen Quality of Fertile U.S. Males. *Environmental Health Perspectives*, 111(4):414–420, November 2002.

[17] Pastuszak A. W and Lipshultz L. I. Words of wisdom. Re: Human semen quality in the new millennium: a prospective cross-sectional population-based study of 4867 men. *European urology*, 62(6):1197–8, December 2012.

[18] Jø rgensen N, Joensen U. N, Jensen T. K, Jensen M. B, Almstrup K, et al. Human semen quality in the new millennium: a prospective cross-sectional population-based study of 4867 men. *BMJ open*, 2(4):1–14, January 2012.

[19] Jø rgensen N, Carlsen E, Nermoen I, Punab M, Suominen J, et al. East-West gradient in semen quality in the Nordic-Baltic area: a study of men from the general population in Denmark, Norway, Estonia and Finland. *Human reproduction*, 17(8):2199–208, August 2002.

[20] Jø rgensen N, Asklund C, Carlsen E, and Skakkebæ k N. E. Coordinated European investigations of semen quality: results from studies of Scandinavian young men is a matter of concern. *International journal of andrology*, 29(1): 54–61; discussion 105–8, February 2006.

[21] Paasch U, Salzbrunn A, Glander H. J, Plambeck K, Salzbrunn H, et al. Semen quality in sub-fertile range for a significant proportion of young men from the general German population: a co-ordinated, controlled study of 791 men from Hamburg and Leipzig. *International journal of andrology*, 31(2): 93–102, April 2008.

[22] Richthoff J, Rylander L, Hagmar L, Malm J, and Giwercman A. Higher sperm counts in Southern Sweden compared with Denmark. *Human reproduction*, 17(9):2468–73, September 2002.

[23] Carlsen E, Swan S. H, Petersen J. r. H, and Skakkebaek N. E. Longitudinal changes in semen parameters in young Danish men from the Copenhagen area. *Human reproduction*, 20(4):942–9, April 2005.

[24] Bonde J. P, Ernst E, Jensen T. K, Hjollund N. H, Kolstad H, et al. Relation between semen quality and fertility: a population-based study of 430 first-pregnancy planners. *Lancet*, 352(9135):1172–7, October 1998.

[25] Sharpe R. M. Male Reproductive Health Disorders and the Potential Role of Exposure to Environmental Chemicals. Technical report, CHEM Trust, Edinburgh, 2011. URL `http://www.chemtrust.org.uk/wp-content/uploads/ProfRSHARPE-MaleReproductiveHealth-CHEMTrust09-1.pdf`.

[26] Myrup C, Westergaard T, Schnack T, Oudin A, Ritz C, et al. Testicular cancer risk in first- and second-generation immigrants to Denmark. *Journal of the National Cancer Institute*, 100(1):41–7, January 2008.

[27] Skakkebaek N. E, Rajpert-De Meyts E, and Main K. M. Testicular dysgenesis syndrome: an increasingly common developmental disorder with environmental aspects. *Human reproduction*, 16(5):972–8, May 2001.

[28] Sonne S. B, Kristensen D. M. b, Novotny G. W, Olesen I. A, Nielsen J. E, et al. Testicular dysgenesis syndrome and the origin of carcinoma in situ testis. *International journal of andrology*, 31(2):275–87, April 2008.

[29] Krysiak-Baltyn K, Toppari J, Skakkebaek N. E, Jensen T. S, Virtanen H. E, et al. Association between chemical pattern in breast milk and congenital cryptorchidism: modelling of complex human exposures. *International journal of andrology*, 35(3):294–302, June 2012.

[30] Damgaard I. N, Skakkebaek N. E, Toppari J, Virtanen H. E, Shen H, et al. Persistent pesticides in human breast milk and cryptorchidism. *Environmental health perspectives*, 114(7):1133–8, July 2006.

[31] Virtanen H. E, Koskenniemi J. J, Sundqvist E, Main K. M, Kiviranta H, et al. Associations between congenital cryptorchidism in newborn boys and levels of dioxins and PCBs in placenta. *International journal of andrology*, 35(3): 283–93, June 2012.

[32] Kaleva M, Virtanen H. E, Haavisto A.-M, Main K. M, Reunanen M, et al. Circannual rhythm in the incidence of cryptorchidism in Finland. *International journal of andrology*, 28(1):53–7, February 2005.

[33] García-Rodríguez J, García-Martín M, Nogueras-Ocaña M, de Dios Luna-del-Castillo J, Espigares García M, et al. Exposure to pesticides and cryptorchidism: geographical evidence of a possible association. *Environmental health perspectives*, 104(10):1090–5, October 1996.

[34] Toppari J, Larsen J. C, Christiansen P, Giwercman A, Grandjean P, et al. Male reproductive health and environmental xenoestrogens. *Environmental health perspectives*, 104 Suppl(August):741–803, August 1996.

[35] Giwercman A, Carlsen E, Keiding N, and Skakkebaek N. E. Evidence for increasing incidence of abnormalities of the human testis: a review. *Environmental health perspectives*, 101 Suppl:65–71, July 1993.

[36] Fernandez M. F, Olmos B. n, Granada A, López-Espinosa M. J, Molina-Molina J.-M, et al. Human exposure to endocrine-disrupting chemicals and prenatal risk factors for cryptorchidism and hypospadias: a nested case-control study. *Environmental health perspectives*, 115 Suppl:8–14, December 2007.

[37] Juul A, Almstrup K, Andersson A.-M, Jensen T. K, Jø rgensen N, et al. Possible fetal determinants of male infertility. *Nature reviews. Endocrinology*, 10(9):553–62, September 2014.

[38] Kortenkamp A, Backhaus T, and Faust M. State of the Art Review of Mixture Toxicity. Technical Report 070307/2007/485103/ETU/D.1, European Commission, 2009. URL `http://ec.europa.eu/environment/chemicals/effects/pdf/report_mixture_toxicity.pdf`.

[39] Vidal M. A unifying view of 21st century systems biology. *FEBS letters*, 583 (24):3891–4, December 2009.

[40] Kell D. B and Oliver S. G. Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. *BioEssays*, 26(1):99–105, January 2004.

[41] Reisfeld B and Mayeno A. N. *Computational Toxicology*, volume 930 of *Methods in Molecular Biology*. Humana Press, Totowa, NJ, 2013. ISBN 978-1-62703-058-8.

[42] Edgar R, Domrachev M, and Lash A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic acids research*, 30(1):207–10, January 2002.

[43] Barrett T, Wilhite S. E, Ledoux P, Evangelista C, Kim I. F, et al. NCBI GEO: archive for functional genomics data sets–update. *Nucleic acids research*, 41 (Database issue):D991–5, January 2013.

[44] Rustici G, Kolesnikov N, Brandizi M, Burdett T, Dylag M, et al. ArrayExpress update–trends in database growth and links to data analysis tools. *Nucleic acids research*, 41(Database issue):D987–90, January 2013.

[45] Schena M, Shalon D, Davis R. W, and Brown P. O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science (New York, N.Y.)*, 270(5235):467–70, October 1995.

[46] Stoughton R. B. Applications of DNA microarrays in biology. *Annual review of biochemistry*, 74:53–82, January 2005.

[47] Dufva M. Fabrication of DNA microarray. *Methods in molecular biology*, 529:63–79, January 2009.

[48] Yakovlev A. Y, Klebanov L, and Gaile D. *Statistical Methods for Microarray Data Analysis*. Humana Press, New York, NY, 2013. ISBN 9781603273367.

[49] Bolstad B. M, Irizarry R. A, Astrand M, and Speed T. P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*, 19(2):185–93, January 2003.

[50] Smyth G. K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical applications in genetics and molecular biology*, 3:Article3, January 2004.

[51] Benjamini Y and Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1):289–300, January 1995.

[52] Reiner A, Yekutieli D, and Benjamini Y. Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics*, 19 (3):368–375, February 2003.

[53] Dix D. J, Houck K. A, Martin M. T, Richard A. M, Setzer R. W, et al. The ToxCast program for prioritizing toxicity testing of environmental chemicals. *Toxicological sciences*, 95(1):5–12, January 2007.

[54] Knudsen T, Martin M, Chandler K, Kleinstreuer N, Judson R, et al. Predictive models and computational toxicology. *Methods in molecular biology*, 947: 343–74, January 2013.

[55] Kavlock R, Chandler K, Houck K, Hunter S, Judson R, et al. Update on EPA's ToxCast program: providing high throughput decision support tools for chemical risk management. *Chemical research in toxicology*, 25(7):1287–302, July 2012.

[56] Judson R. S, Houck K. A, Kavlock R. J, Knudsen T. B, Martin M. T, et al. In vitro screening of environmental chemicals for targeted testing prioritization: the ToxCast project. *Environmental health perspectives*, 118(4): 485–92, April 2010.

[57] Martin M. T, Dix D. J, Judson R. S, Kavlock R. J, Reif D. M, et al. Impact of environmental chemicals on key transcription regulators and correlation to toxicity end points within EPA's ToxCast program. *Chemical research in toxicology*, 23(3):578–90, March 2010.

[58] Houck K. A, Dix D. J, Judson R. S, Kavlock R. J, Yang J, et al. Profiling bioactivity of the ToxCast chemical library using BioMAP primary human cell systems. *Journal of biomolecular screening*, 14(9):1054–66, October 2009.

[59] Kleinstreuer N. C, Yang J, Berg E. L, Knudsen T. B, Richard A. M, et al. Phenotypic screening of the ToxCast chemical library to classify toxic and therapeutic mechanisms. *Nature biotechnology*, 32(6):583–91, June 2014.

[60] Rotroff D. M, Beam A. L, Dix D. J, Farmer A, Freeman K. M, et al. Xenobiotic-metabolizing enzyme and transporter gene expression in primary cultures of human hepatocytes modulated by ToxCast chemicals. *Journal of toxicology and environmental health. Part B, Critical reviews*, 13(2-4): 329–46, February 2010.

[61] Chandler K. J, Barrier M, Jeffay S, Nichols H. P, Kleinstreuer N. C, et al. Evaluation of 309 environmental chemicals using a mouse embryonic stem cell adherent cell differentiation and cytotoxicity assay. *PloS one*, 6(6):e18540, January 2011.

[62] Knight A. W, Little S, Houck K, Dix D, Judson R, et al. Evaluation of high-throughput genotoxicity assays used in profiling the US EPA ToxCast chemicals. *Regulatory toxicology and pharmacology*, 55(2):188–99, November 2009.

[63] Huang R, Xia M, Cho M.-H, Sakamuru S, Shinn P, et al. Chemical genomics profiling of environmental chemical modulation of human nuclear receptors. *Environmental health perspectives*, 119(8):1142–8, August 2011.

[64] Knudsen T. B, Houck K. A, Sipes N. S, Singh A. V, Judson R. S, et al. Activity profiles of 309 ToxCast™ chemicals evaluated across 292 biochemical targets. *Toxicology*, 282(1-2):1–15, March 2011.

[65] Sipes N. S, Martin M. T, Kothiya P, Reif D. M, Judson R. S, et al. Profiling 976 ToxCast chemicals across 331 enzymatic and receptor signaling assays. *Chemical research in toxicology*, 26(6):878–95, June 2013.

[66] MacDonald M. L, Lamerdin J, Owens S, Keon B. H, Bilter G. K, et al. Identifying off-target effects and hidden phenotypes of drugs in human cells. *Nature chemical biology*, 2(6):329–37, June 2006.

[67] McDonough P. M, Agustin R. M, Ingermanson R. S, Loy P. a, Buehrer B. M, et al. Quantification of lipid droplets and associated proteins in cellular models of obesity via high-content/high-throughput microscopy and automated image analysis. *Assay and drug development technologies*, 7(5):440–60, October 2009.

[68] McDonough P. M, Ingermanson R. S, Loy P. a, Koon E. D, Whittaker R, et al. Quantification of hormone sensitive lipase phosphorylation and colocalization with lipid droplets in murine 3T3L1 and human subcutaneous adipocytes via automated digital microscopy and high-content analysis. *Assay and drug development technologies*, 9(3):262–80, June 2011.

[69] Padilla S, Corum D, Padnos B, Hunter D. L, Beam A, et al. Zebrafish developmental screening of the ToxCast™ Phase I chemical library. *Reproductive toxicology*, 33(2):174–87, April 2012.

[70] Rotroff D. M, Dix D. J, Houck K. A, Kavlock R. J, Knudsen T. B, et al. Real-time growth kinetics measuring hormone mimicry for ToxCast chemicals in T-47D human ductal carcinoma cells. *Chemical research in toxicology*, 26 (7):1097–107, July 2013.

[71] Judson R, Houck K, Martin M, Knudsen T, Thomas R. S, et al. In vitro and modelling approaches to risk assessment from the U.S. Environmental Protection Agency ToxCast programme. *Basic & clinical pharmacology & toxicology*, 115(1):69–76, July 2014.

[72] Audouze K and Grandjean P. Application of computational systems biology to explore environmental toxicity hazards. *Environmental health perspectives*, 119(12):1754–9, December 2011.

[73] Davis A. P, Murphy C. G, Johnson R, Lay J. M, Lennon-Hopkins K, et al. The Comparative Toxicogenomics Database: update 2013. *Nucleic acids research*, 41(Database issue):D1104–14, January 2013.

[74] Kim Kjærulff S, Wich L, Kringelum J, Jacobsen U. P, Kouskoumvekaki I, et al. ChemProt-2.0: visual navigation in a disease chemical biology database. *Nucleic acids research*, 41(Database issue):D464–9, January 2013.

[75] Lage K, Mø llgård K, Greenway S, Wakimoto H, Gorham J. M, et al. Dissecting spatio-temporal protein networks driving human heart development and related disorders. *Molecular systems biology*, 6:381, June 2010.

[76] McKusick-Nathans Institute of Genetic Medicine Johns Hopkins University (Baltimore MD). Online Mendelian Inheritance in Man, OMIM®, 2012. URL `http://omim.org/`.

[77] Safran M, Dalah I, Alexander J, Rosen N, Iny Stein T, et al. GeneCards Version 3: the human gene integrator. *Database*, 2010:baq020, January 2010.

[78] Kanehisa M and Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research*, 28(1):27–30, January 2000.

[79] Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, et al. Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic acids research*, 42(Database issue):D199–205, January 2014.

[80] Croft D, Mundo A. F, Haw R, Milacic M, Weiser J, et al. The Reactome pathway knowledgebase. *Nucleic acids research*, 42(Database issue):D472–7, January 2014.

[81] Wade M. G, Foster W. G, Younglai E. V, McMahon A, Leingartner K, et al. Effects of subchronic exposure to a complex mixture of persistent contaminants in male rats: systemic, immune, and reproductive effects. *Toxicological sciences*, 67(1):131–43, May 2002.

[82] Naville D, Pinteur C, Vega N, Menade Y, Vigier M, et al. Low-dose food contaminants trigger sex-specific, hepatic metabolic changes in the progeny of obese mice. *FASEB journal*, 27(9):3860–70, September 2013.

[83] Crofton K. M, Craft E. S, Hedge J. M, Gennings C, Simmons J. E, et al. Thyroid-Hormone–Disrupting Chemicals: Evidence for Dose-Dependent Additivity or Synergism. *Environmental Health Perspectives*, 113(11):1549–1554, July 2005.

[84] Merhi M, Demur C, Racaud-Sultan C, Bertrand J, Canlet C, et al. Gender-linked haematopoietic and metabolic disturbances induced by a pesticide mixture administered at low dose to mice. *Toxicology*, 267(1-3):80–90, January 2010.

[85] Christiansen S, Kortenkamp A, Axelstad M, Boberg J, Scholze M, et al. Mixtures of endocrine disrupting contaminants modelled on human high end exposures: an exploratory study in rats. *International journal of andrology*, 35(3):303–16, June 2012.

[86] Lu R.-C, Cao A.-N, Lai L.-H, and Xiao J.-X. Protein-surfactant interaction: differences between fluorinated and hydrogenated surfactants. *Colloids and surfaces. B, Biointerfaces*, 64(1):98–103, June 2008.

[87] Lau C, Anitole K, Hodes C, Lai D, Pfahles-Hutchens A, et al. Perfluoroalkyl acids: a review of monitoring and toxicological findings. *Toxicological sciences*, 99(2):366–94, October 2007.

[88] Lau C. Perfluorinated compounds. *EXS*, 101:47–86, January 2012.

[89] Calafat A. M, Wong L.-Y, Kuklenyik Z, Reidy J. A, and Needham L. L. Polyfluoroalkyl chemicals in the U.S. population: data from the National Health and Nutrition Examination Survey (NHANES) 2003-2004 and comparisons with NHANES 1999-2000. *Environmental health perspectives*, 115 (11):1596–602, November 2007.

[90] NHANES. Fourth National Report on Human Exposure to Environmental Chemicals Updated Tables, August, 2014. Technical report, U. S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2014. URL `http://www.cdc.gov/exposurereport`.

[91] Calafat A. M, Kuklenyik Z, Reidy J. A, Caudill S. P, Tully J. S, et al. Serum concentrations of 11 polyfluoroalkyl compounds in the u.s. population: data from the national health and nutrition examination survey (NHANES). *Environmental science & technology*, 41(7):2237–42, April 2007.

[92] Starling A. P, Engel S. M, Whitworth K. W, Richardson D. B, Stuebe A. M, et al. Perfluoroalkyl substances and lipid concentrations in plasma during pregnancy among women in the Norwegian Mother and Child Cohort Study. *Environment international*, 62:104–12, January 2014.

[93] Mondal D, Weldon R. H, Armstrong B. G, Gibson L. J, Lopez-Espinosa M.-J, et al. Breastfeeding: a potential excretion route for mothers and implications for infant exposure to perfluoroalkyl acids. *Environmental health perspectives*, 122(2):187–92, February 2014.

[94] Axmon A, Axelsson J, Jakobsson K, Lindh C. H, and Jönsson B. a. G. Time trends between 1987 and 2007 for perfluoroalkyl acids in plasma from Swedish women. *Chemosphere*, 102:61–7, May 2014.

[95] Taylor K. W, Hoffman K, Thayer K. a, and Daniels J. L. Polyfluoroalkyl chemicals and menopause among women 20-65 years of age (NHANES). *Environmental health perspectives*, 122(2):145–50, February 2014.

[96] Okada E, Sasaki S, Kashino I, Matsuura H, Miyashita C, et al. Prenatal exposure to perfluoroalkyl acids and allergic diseases in early childhood. *Environment international*, 65:127–34, April 2014.

[97] Kato K, Calafat A. M, Wong L.-Y. Y, Wanigatunga A. A, Caudill S. P, et al. Polyfluoroalkyl compounds in pooled sera from children participating in the National Health and Nutrition Examination Survey 2001-2002. *Environmental science & technology*, 43(7):2641–2647, April 2009.

[98] Klenow S, Heinemeyer G, Brambilla G, Dellatte E, Herzke D, et al. Dietary exposure to selected perfluoroalkyl acids (PFAAs) in four European regions. *Food additives & contaminants: Part A*, 30(12):2141–51, January 2013.

[99] Granum B, Haug L. S, Namork E, Stø levik S. B, Thomsen C, et al. Prenatal exposure to perfluoroalkyl substances may be associated with altered

vaccine antibody levels and immune-related health outcomes in early childhood. *Journal of immunotoxicology*, 10(4):373–9, 2013.

[100] Lind L, Zethelius B, Salihovic S, van Bavel B, and Lind P. M. Circulating levels of perfluoroalkyl substances and prevalent diabetes in the elderly. *Diabetologia*, 57(3):473–9, March 2014.

[101] Yamada a, Bemrah N, Veyrand B, Pollono C, Merlo M, et al. Dietary exposure to perfluoroalkyl acids of specific French adult sub-populations: high seafood consumers, high freshwater fish consumers and pregnant women. *The Science of the total environment*, 491-492:170–5, September 2014.

[102] Yeung L. W. Y, Guruge K. S, Taniyasu S, Yamashita N, Angus P. W, et al. Profiles of perfluoroalkyl substances in the liver and serum of patients with liver cancer and cirrhosis in Australia. *Ecotoxicology and environmental safety*, 96:139–46, October 2013.

[103] Luo Z, Shi X, Hu Q, Zhao B, and Huang M. Structural evidence of perfluorooctane sulfonate transport by human serum albumin. *Chemical research in toxicology*, 25(5):990–2, May 2012.

[104] Vanden Heuvel J. P, Sterchele P. F, Nesbit D. J, and Peterson R. E. Coordinate induction of acyl-CoA binding protein, fatty acid binding protein and peroxisomal beta-oxidation by peroxisome proliferators. *Biochimica et biophysica acta*, 1177(2):183–90, June 1993.

[105] Woodcroft M. W, Ellis D. A, Rafferty S. P, Burns D. C, March R. E, et al. Experimental characterization of the mechanism of perfluorocarboxylic acids' liver protein bioaccumulation: the key role of the neutral species. *Environmental toxicology and chemistry*, 29(8):1669–77, August 2010.

[106] Luebker D. J, Hansen K. J, Bass N. M, Butenhoff J. L, and Seacat A. M. Interactions of fluorochemicals with rat liver fatty acid-binding protein. *Toxicology*, 176(3):175–85, July 2002.

[107] Kleszczynski K, Stepnowski P, and Skladanowski A. C. Mechanism of cytotoxic action of perfluorinated acids II. Disruption of mitochondrial bioenergetics. *Toxicology and applied pharmacology*, 235(2):182–90, March 2009.

[108] Kleszczynski K and Skladanowski A. C. Mechanism of cytotoxic action of perfluorinated acids. I. alteration in plasma membrane potential and intracellular pH level. *Toxicology and applied pharmacology*, 234(3):300–5, February 2009.

[109] Kleszczynski K and Skladanowski A. C. Mechanism of cytotoxic action of perfluorinated acids. III. Disturbance in $Ca^2+$ homeostasis. *Toxicology and applied pharmacology*, 251(2):163–8, March 2011.

[110] Upham B. L, Deocampo N. D, Wurl B, and Trosko J. E. Inhibition of gap junctional intercellular communication by perfluorinated fatty acids is dependent on the chain length of the fluorinated tail. *International journal of cancer*, 78(4):491–5, November 1998.

[111] Hu W, Jones P. D, Upham B. L, Trosko J. E, Lau C, et al. Inhibition of gap junctional intercellular communication by perfluorinated compounds in rat liver and dolphin kidney epithelial cell lines in vitro and Sprague-Dawley rats in vivo. *Toxicological sciences*, 68(2):429–36, August 2002.

[112] Hu W. Y, Jones P. D, DeCoen W, King L, Fraker P, et al. Alterations in cell membrane properties caused by perfluorinated compounds. *Comparative biochemistry and physiology. Toxicology & pharmacology : CBP*, 135(1): 77–88, May 2003.

[113] Kudo N, Suzuki-Nakajima E, Mitsumoto A, and Kawashima Y. Responses of the liver to perfluorinated fatty acids with different carbon chain length in male and female mice:in relation to induction of hepatomegaly, peroxisomal beta-oxidation and microsomal 1-acylglycerophosphocholine acyltransferase. *Biological & pharmaceutical bulletin*, 29(9):1952–7, September 2006.

[114] Rosenmai a. K, Nielsen F. K, Pedersen M, Hadrup N, Trier X, et al. Fluorochemicals used in food packaging inhibit male sex hormone synthesis. *Toxicology and applied pharmacology*, 266(1):132–42, January 2013.

[115] Shi Z, Ding L, Zhang H, Feng Y, Xu M, et al. Chronic exposure to perfluorododecanoic acid disrupts testicular steroidogenesis and the expression of related genes in male rats. *Toxicology letters*, 188(3):192–200, August 2009.

[116] Shi Z, Zhang H, Liu Y, Xu M, and Dai J. Alterations in gene expression and testosterone synthesis in the testes of male rats exposed to perfluorododecanoic acid. *Toxicological sciences*, 98(1):206–15, July 2007.

[117] Feng Y, Shi Z, Fang X, Xu M, and Dai J. Perfluorononanoic acid induces apoptosis involving the Fas death receptor signaling pathway in rat testis. *Toxicology letters*, 190(2):224–30, October 2009.

[118] Skov K, Hadrup N, Smedsgaard J. r, and Frandsen H. LC-MS analysis of the plasma metabolome-A novel sample preparation strategy. *Journal of chromatography B*, 978-979:83–8, January 2015.

[119] R Core Team. R: A Language and Environment for Statistical Computing, 2012. ISSN 16000706. URL http://www.r-project.org.

[120] Smyth G. limma: linear models for microarray data. In Gentleman R, Carey V. J, Huber W, Irizarry R. A, and Dudoit S, editors, *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, chapter 23, pages 397–420. Springer New York, New York, NY, 2005. URL http://link.springer.com/chapter/10.1007/0-387-29362-0_23.

[121] Martin M. T, Knudsen T. B, Reif D. M, Houck K. A, Judson R. S, et al. Predictive model of rat reproductive toxicity from ToxCast high throughput screening. *Biology of reproduction*, 85(2):327–39, August 2011.

[122] Sipes N. S, Martin M. T, Reif D. M, Kleinstreuer N. C, Judson R. S, et al. Predictive models of prenatal developmental toxicity from ToxCast high-throughput screening data. *Toxicological sciences*, 124(1):109–27, November 2011.

[123] Kleinstreuer N. C, Judson R. S, Reif D. M, Sipes N. S, Singh A. V, et al. Environmental impact on vascular development predicted by high-throughput screening. *Environmental health perspectives*, 119(11):1596–603, November 2011.

[124] Dean A and Sharpe R. M. Clinical review: Anogenital distance or digit length ratio as measures of fetal androgen exposure: relationship to male reproductive development and its disorders. *The Journal of clinical endocrinology and metabolism*, 98(6):2230–8, June 2013.

[125] Wolf C, Lambright C, Mann P, Price M, Cooper R. L, et al. Administration of potentially antiandrogenic pesticides (procymidone, linuron, iprodione, chlozolinate, p,p'-DDE, and ketoconazole) and toxic substances (dibutyl- and diethylhexyl phthalate, PCB 169, and ethane dimethane sulphonate) during sexual differen. *Toxicology and industrial health*, 15(1-2):94–118, 1999.

[126] Welsh M, Saunders P. T. K, Fisken M, Scott H. M, Hutchison G. R, et al. Identification in rats of a programming window for reproductive tract masculinization, disruption of which leads to hypospadias and cryptorchidism. *The Journal of clinical investigation*, 118(4):1479–90, April 2008.

[127] Liu C, Xu X, and Huo X. Anogenital distance and its application in environmental health research. *Environmental science and pollution research international*, 21(8):5457–64, April 2014.

[128] Thankamony A, Lek N, Carroll D, Williams M, Dunger D. B, et al. Anogenital distance and penile length in infants with hypospadias or cryptorchidism: comparison with normative data. *Environmental health perspectives*, 122 (2):207–11, February 2014.

[129] Jain V. G and Singal A. K. Shorter anogenital distance correlates with undescended testis: a detailed genital anthropometric analysis in human newborns. *Human reproduction*, 28(9):2343–9, September 2013.

[130] Hsieh M. H, Eisenberg M. L, Hittelman A. B, Wilson J. M, Tasian G. E, et al. Caucasian male infants and boys with hypospadias exhibit reduced anogenital distance. *Human reproduction*, 27(6):1577–80, June 2012.

[131] Hsieh M. H, Breyer B. N, Eisenberg M. L, and Baskin L. S. Associations among hypospadias, cryptorchidism, anogenital distance, and endocrine disruption. *Current Urology Reports*, 9(2):137–142, May 2008.

[132] Eisenberg M. L, Shy M, Walters R. C, and Lipshultz L. I. The relationship between anogenital distance and azoospermia in adult men. *International journal of andrology*, 35(5):726–30, October 2012.

[133] Eisenberg M. L, Jensen T. K, Walters R. C, Skakkebaek N. E, and Lipshultz L. I. The relationship between anogenital distance and reproductive hormone levels in adult men. *The Journal of urology*, 187(2):594–8, February 2012.

[134] U. S. Environmental Protection Agency. Health Effects Test Guidelines, OPPTS 870.3800, Reproduction and Fertility Effects. Technical Report August, U. S. Environmental Protection Agency, 1998. URL `http://ntp.niehs.nih.gov/iccvam/suppdocs/feddocs/epa/epa_870_3800.pdf`.

[135] OECD. *Test No. 443: Extended One-Generation Reproductive Toxicity Study*. Number July in OECD Guidelines for the Testing of Chemicals, Section 4. OECD Publishing, July 2011.

[136] Venables W. N and Ripley B. D. *Modern Applied Statistics with S*. Springer, New York, fourth edition, 2002.

[137] Warnes G. R, Bolker B, Bonebakker L, Gentleman R, Liaw W. H. A, et al. *gplots: Various R programming tools for plotting data*, 2012. URL `http://cran.r-project.org/package=gplots`.

[138] Lampen A, Zimnik S, and Nau H. Teratogenic phthalate esters and metabolites activate the nuclear receptors PPARs and induce differentiation of F9 cells. *Toxicology and Applied Pharmacology*, 188(1):14–23, April 2003.

[139] Blomberg Jensen M. Vitamin D and male reproduction. *Nature reviews. Endocrinology*, 10(3):175–86, March 2014.

[140] Morohaku K, Pelton S. H, Daugherty D. J, Butler W. R, Deng W, et al. Translocator protein/peripheral benzodiazepine receptor is not required for steroid hormone biosynthesis. *Endocrinology*, 155(1):89–97, January 2014.

[141] Cheng K. C and Schenkman J. B. Testosterone metabolism by cytochrome P-450 isozymes RLM3 and RLM5 and by microsomes. Metabolite identification. *The Journal of biological chemistry*, 258(19):11738–44, October 1983.

[142] Guengerich F. P. Cytochrome P-450 3A4: regulation and role in drug metabolism. *Annual review of pharmacology and toxicology*, 39:1–17, January 1999.

[143] Corton J. C and Lapinskas P. J. Peroxisome proliferator-activated receptors: mediators of phthalate ester-induced effects in the male reproductive tract? *Toxicological sciences : an official journal of the Society of Toxicology*, 83 (1):4–17, January 2005.

[144] Rak-Mardyla A and Karpeta A. Rosiglitazone stimulates peroxisome proliferator-activated receptor gamma expression and directly affects in vitro steroidogenesis in porcine ovarian follicles. *Theriogenology*, 82(1):1–9, July 2014.

[145] Johnson K. J, Heger N. E, and Boekelheide K. Of mice and men (and rats): phthalate-induced fetal testis endocrine disruption is species-dependent. *Toxicological sciences*, 129(2):235–48, October 2012.

[146] Taxvig C, Hass U, Axelstad M, Dalgaard M, Boberg J, et al. Endocrine-disrupting activities in vivo of the fungicides tebuconazole and epoxiconazole. *Toxicological sciences*, 100(2):464–73, December 2007.

[147] Taxvig C, Vinggaard A. M, Hass U, Axelstad M, Metzdorff S, et al. Endocrine-disrupting properties in vivo of widely used azole fungicides. *International journal of andrology*, 31(2):170–7, April 2008.

[148] Ball E. R, Caniglia M. K, Wilcox J. L, Overton K. a, Burr M. J, et al. Effects of genistein in the maternal diet on reproductive development and spatial learning in male rats. *Hormones and behavior*, 57(3):313–22, March 2010.

[149] Hotchkiss a. K, Parks-Saldutti L. G, Ostby J. S, Lambright C, Furr J, et al. A mixture of the "antiandrogens" linuron and butyl benzyl phthalate alters sexual differentiation of the male rat in a cumulative fashion. *Biology of reproduction*, 71(6):1852–61, December 2004.

[150] Laier P, Metzdorff S. B, Borch J, Hagen M. L, Hass U, et al. Mechanisms of action underlying the antiandrogenic effects of the fungicide prochloraz. *Toxicology and applied pharmacology*, 213(2):160–71, June 2006.

[151] Kristensen D. M. b, Hass U, Lesné L, Lottrup G, Jacobsen P. R, et al. Intrauterine exposure to mild analgesics is a risk factor for development of male reproductive disorders in human and rat. *Human reproduction*, 26(1): 235–44, January 2011.

[152] Goetz A. K, Ren H, Schmid J. E, Blystone C. R, Thillainadarajah I, et al. Disruption of testosterone homeostasis as a mode of action for the reproductive toxicity of triazole fungicides in the male rat. *Toxicological sciences*, 95 (1):227–39, January 2007.

[153] Saillenfait A.-M, Roudot A.-C, Gallissot F, and Sabaté J.-P. Prenatal developmental toxicity studies on di-n-heptyl and di-n-octyl phthalates in Sprague-Dawley rats. *Reproductive toxicology*, 32(3):268–76, November 2011.

[154] Taxvig C, Vinggaard A. M, Hass U, Axelstad M, Boberg J, et al. Do parabens have the ability to interfere with steroidogenesis? *Toxicological sciences*, 106 (1):206–13, November 2008.

[155] Lawson G and Luderer U. Gestational and lactational exposure to heptachlor does not alter reproductive system development in rats. *Veterinary and human toxicology*, 46(3):113–8, June 2004.

[156] Matsuura I, Saitoh T, Tani E, Wako Y, Iwata H, et al. Evaluation of a two-generation reproduction toxicity study adding endpoints to detect endocrine disrupting activity using lindane. *The Journal of toxicological sciences*, 30 Spec No(Special):135–161, December 2005.

[157] Siddiqui W. H, Stump D. G, Plotzke K. P, Holson J. F, and Meeks R. G. A two-generation reproductive toxicity study of octamethylcyclotetrasiloxane (D4) in rats exposed by whole-body vapor inhalation. *Reproductive toxicology*, 23(2):202–15, February 2007.

[158] Gray L. E, Ostby J, Furr J, Price M, Veeramachaneni D. N, et al. Perinatal exposure to the phthalates DEHP, BBP, and DINP, but not DEP, DMP, or DOTP, alters sexual differentiation of the male rat. *Toxicological sciences*, 58(2):350–65, December 2000.

National Food Institute
Technical University of Denmark
Mørkhøj Bygade 19
DK-2860 Søborg

Tlf. 35 88 70 00
Fax 35 88 70 01

www.food.dtu.dk

ISBN: 978-87-93109-30-8