



Integrated transcriptomics and metabolomics reveal signatures of lipid metabolism dysregulation in HepaRG liver cells exposed to PCB 126

Robin Mesnage¹ · Martina Biserni¹ · Sucharitha Balu² · Clément Frainay³ · Nathalie Poupin³ · Fabien Jourdan³ · Eva Wozniak⁴ · Theodoros Xenakis⁴ · Charles A. Mein⁴ · Michael N. Antoniou¹

Received: 8 February 2018 / Accepted: 4 June 2018 / Published online: 14 June 2018
© The Author(s) 2018

Abstract

Chemical pollutant exposure is a risk factor contributing to the growing epidemic of non-alcoholic fatty liver disease (NAFLD) affecting human populations that consume a western diet. Although it is recognized that intoxication by chemical pollutants can lead to NAFLD, there is limited information available regarding the mechanism by which typical environmental levels of exposure can contribute to the onset of this disease. Here, we describe the alterations in gene expression profiles and metabolite levels in the human HepaRG liver cell line, a validated model for cellular steatosis, exposed to the polychlorinated biphenyl (PCB) 126, one of the most potent chemical pollutants that can induce NAFLD. Sparse partial least squares classification of the molecular profiles revealed that exposure to PCB 126 provoked a decrease in polyunsaturated fatty acids as well as an increase in sphingolipid levels, concomitant with a decrease in the activity of genes involved in lipid metabolism. This was associated with an increased oxidative stress reflected by marked disturbances in taurine metabolism. A gene ontology analysis showed hallmarks of an activation of the AhR receptor by dioxin-like compounds. These changes in metabolome and transcriptome profiles were observed even at the lowest concentration (100 pM) of PCB 126 tested. A decrease in docosatrienoate levels was the most sensitive biomarker. Overall, our integrated multi-omics analysis provides mechanistic insight into how this class of chemical pollutant can cause NAFLD. Our study lays the foundation for the development of molecular signatures of toxic effects of chemicals causing fatty liver diseases to move away from a chemical risk assessment based on *in vivo* animal experiments.

Keywords NAFLD · Transcriptome · Metabolome · HepaRG · Liver · PCB

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00204-018-2235-7>) contains supplementary material, which is available to authorized users.

✉ Michael N. Antoniou
michael.antoniou@kcl.ac.uk

- ¹ Gene Expression and Therapy Group, King's College London, Faculty of Life Sciences and Medicine, Department of Medical and Molecular Genetics, 8th Floor, Tower Wing, Guy's Hospital, Great Maze Pond, London SE1 9RT, UK
- ² Genomics Centre, King's College London, Waterloo Campus, 150 Stamford Street, London SE1 9NH, UK
- ³ INRA UMR1331, Toxalim, Research Centre in Food Toxicology, Toulouse, France
- ⁴ Genome Centre, Barts and the London School of Medicine and Dentistry, John Vane Science Centre, London EC1M 6BQ, UK

Introduction

A growing epidemic of non-alcoholic fatty liver disease (NAFLD) is affecting human populations that consume a western diet (Argyrou et al. 2017). The spectrum of NAFLD ranges from fatty infiltration of the liver (steatosis), to more advanced stages characterized by liver inflammation and fibrosis (non-alcoholic steatohepatitis, NASH), and subsequent cirrhosis, and hepatocellular carcinoma (Michelotti et al. 2013; Vernon et al. 2011). NAFLD currently affects 25% of the US population, and approximately 2–5% have NASH (Vernon et al. 2011) with concomitant massive clinical and economic burden (Younossi et al. 2016). Annual direct medical costs have been estimated to be approximately \$103 billion (\$1613 per patient) in the US. The occurrence of NAFLD in Europe is also very high and ranges from 20 to 30% with annual medical costs

currently running at €35 billion (from €354 to €1163 per patient) in countries such as Germany, France, Italy, and the United Kingdom (Younossi et al. 2016). Risk factors identified to date for the development of NAFLD include being overweight or obese, having diabetes and high cholesterol or high triglycerides in the blood. Rapid weight loss and poor eating habits can also lead to NAFLD. The heritability of NAFLD was estimated at ~50% (Sookoian and Pirola 2017). However, some individuals develop NAFLD even if they do not have any of these risk factors (Foulds et al. 2017) and with exposure to physiologically active environmental pollutants being a recognized additional determinant of disease establishment.

Chemical exposure has long been known to be a risk factor for NAFLD/NAFLD (Petraakis et al. 2017; Wahlang et al. 2013), which is also known as toxicant-associated fatty liver disease (TAFLD) and its more severe form, toxicant-associated steatohepatitis (TASH). TAFLD has been commonly associated with acute chemical exposures at the workplace (Cave et al. 2010b), or after industrial accidents (Yu et al. 1997). However, there is also evidence showing that environmental pollutants have the ability to cause TAFLD at environmental levels of exposure (Cave et al. 2010a). The ability of environmental pollutants to disturb metabolic processes at low environmental levels of exposure generally results from their ability to interfere with the activity of natural hormones by binding to their receptors (Heindel et al. 2017). Different structural classes of environmental pollutants have been shown to modulate liver nuclear receptors, such as dioxins acting on the aryl hydrocarbon receptor (AhR), organotins on peroxisome proliferator-activated receptor γ (PPAR γ), or organochlorines on the constitutive androstane receptor (CAR) (Foulds et al. 2017). Some environmental pollutants also promote NAFLD because of their hormone mimicking properties on distant organs. For example, bisphenol A acts on the pancreas by stimulating the production and secretion of insulin which, in turn, increases the storage of fat by stimulating lipogenesis in the liver (Fabricio et al. 2016). Commonly defined as obesogens, some other endocrine disrupting chemicals (EDCs) have been found to promote adipose cell differentiation and lipid accumulation (Janesick and Blumberg 2016) with some studies establishing a direct link between human exposures to some EDCs and risk of obesity (Tang-Peronard et al. 2014). Recent advances in human microbial ecology has also revealed that alterations of the gut microbiota by chemicals at early stages of life accelerate hepatic lipid metabolism (Ba et al. 2017). However, TAFLD is not only a hormone receptor-mediated disease but can also arise from metabolic imbalances. For instance, fatty acid breakdown can be altered following a direct inhibition of mitochondrial β -oxidation by, for example, the insecticide chlordecone (Kaiser et al. 2012). Mitochondrial β -oxidation can also be indirectly inhibited by an

alteration in NAD⁺ cofactor metabolism by compounds such as vinyl chloride.

Overall, biological mechanisms underlying the development of steatosis are grouped into four biological processes affecting the metabolism of fatty acids: an increased uptake, a decreased efflux, an increased synthesis, or defects in their oxidative metabolism (Angrish et al. 2017). All these processes are interrelated with a need for more studies providing mechanistic insight into the alteration of hepatic lipid metabolism, which in turn induces the development of NAFLD (Foulds et al. 2017). This is a priority for regulatory toxicology programs and includes a move away from a chemical risk assessment based on *in vivo* animal experiments (Merrick et al. 2015). This can be achieved by detailed molecular profiling studies describing the effects of known toxicants on validated tissue culture model systems, such as the human liver cell line HepaRG (Angrish et al. 2017).

We describe here the alterations in the gene expression profile and metabolite levels of HepaRG cells exposed to the polychlorinated biphenyl (PCB) congener 126, one of the most potent chemicals associated with development of TAFLD (Al-Eryani et al. 2015). Our investigation included a metabolomics analysis, which identified 802 metabolites, which were analyzed with MetExplore (Cottret et al. 2010), to extract the metabolic sub-network involved in the biological response to low environmental levels of PCB 126 exposure. This was coupled with a transcriptomics analysis allowing identification of gene networks involved in the response to PCB 126. Our study is the first in-depth investigation of the molecular profiles underlying the toxic effects of PCB 126 by integrating the transcriptome and metabolome of HepaRG cells in response to this compound and which ultimately highlighted TAFLD-associated sensitive biomarkers of exposure to this class of pollutant.

Materials and methods

Reagents

All reagents and chemicals, unless otherwise specified, were of analytical grade and were purchased from Sigma-Aldrich (Gillingham, Dorset, UK). The PCB 126 (98.5% purity, CAS Number 57465-28-8) was purchased from LGC Standards GmbH (Wesel, Germany). Stock solutions of PCB 126 were made by dissolving in DMSO. The Williams' E medium + GlutaMAXTM were purchased from Gibco (Thermo Fisher Scientific, Loughborough, UK). The supplement ADD670, as well as the DMSO, was provided by Biopredic International (Rennes, France).

HepaRG cell culture

HepaRG™ cells (HPR 116) were purchased from Biopredic International (Rennes, France). Cells were thawed, suspended and plated in the general purpose medium (Williams' E medium + GlutaMAX™) containing the ADD670 supplement. A total of 72,000 and 2,000,000 cells were plated in collagen-coated 96-well plates (Greiner Bio-One, Germany) and 6-well plates (Biopredic), respectively. All cells were cultured at 37 °C and a 5% CO₂ atmosphere. The medium was refreshed at days 2, 5 and 7 after initial plating. The cells were kept in the general purpose medium until day 8, when the culture becomes well organized and includes well-delineated trabeculae and many canaliculi-like structures. At this time, the culture is composed of primitive biliary epithelial cells and mature hepatocytes with basal metabolic activities similar to fresh hepatocyte cells. From day 8 to day 14, cells were switched to the test medium composed of Williams' E medium + GlutaMAX™ supplemented with 2% fetal bovine serum (FBS; GE Healthcare Life Sciences, Buckinghamshire, UK), 2 mM L-glutamine (GE Healthcare Life Sciences), 10 µg/mL penicillin/streptomycin (Life technologies) and 1% DMSO, as well as different concentrations of PCB 126 or the solvent control. Effects of PCB 126 was assessed at three concentrations to cover a wide range of biological activities, starting from a low environmental concentration (100 pM) to a high concentration of (1 µM), with an intermediate concentration (10 nM).

Mass spectrometry-based metabolomics

Approximately 5,000,000 cells per sample were harvested from the 6-well plate cultures to obtain a sufficient amount of material to perform the metabolomics analysis. Cells were detached using 0.05% trypsin EDTA (Thermo Fisher Scientific, Loughborough, UK), and collected by centrifugation to eliminate trypsin residues and cell pellets frozen. The frozen cell pellets were then processed by Metabolon Inc. (Durham, NC, USA) for metabolomics analysis. Samples were prepared using the automated MicroLab STAR® system from Hamilton Company (Reno, NV, USA). Proteins were precipitated with methanol under vigorous shaking for 2 min using the Glen Mills GenoGrinder 2000 (Glen Mills, NJ, USA) followed by centrifugation. Samples were placed briefly on a TurboVap® (Zymark, Hopkinton, MA, USA) to remove the organic solvent. The sample extracts were stored overnight under nitrogen before preparation for analysis. The resulting extract was analyzed on four independent instrument platforms: two different separate reverse-phase ultrahigh performance liquid chromatography–tandem mass spectroscopy analysis (RP/UPLC–MS/MS) with positive ion mode electrospray ionization (ESI), a RP/UPLC–MS/MS with negative ion mode ESI, as well as a by hydrophilic

interaction chromatography (HILIC)/UPLC–MS/MS with negative ion mode ESI.

All UPLC–MS/MS methods utilized a Waters ACQUITY ultraperformance liquid chromatography (UPLC) (Waters Corporation, Milford, MA, USA) and a Thermo Scientific Q-Exactive high-resolution/accurate mass spectrometer (Thermo Fisher Scientific Inc., Waltham, MA USA) interfaced with a heated electrospray ionization (HESI-II) source and Orbitrap mass analyzer operated at 35,000 mass resolution. The sample extract was dried then reconstituted in solvents compatible to each of four methods used. Each reconstitution solvent contained a series of standards at fixed concentrations to ensure injection and chromatographic consistency. One aliquot was analyzed using acidic positive ion conditions, chromatographically optimized for more hydrophilic compounds. In this method, the extract was gradient eluted from a C18 column (Waters UPLC BEH C18-2.1 × 100 mm, 1.7 µm) using water and methanol, containing 0.05% perfluoropentanoic acid (PFPA) and 0.1% formic acid (FA). Another aliquot was also analyzed using acidic positive ion conditions, chromatographically optimized for more hydrophobic compounds. In this method, the extract was gradient eluted from the same aforementioned C18 column using methanol, acetonitrile, water, 0.05% PFPA and 0.01% FA, and was operated at an overall higher organic content. Another aliquot was analyzed using basic negative ion optimized conditions using a separate dedicated C18 column. The basic extracts were gradient eluted from the column using methanol and water, with 6.5 mM ammonium bicarbonate at pH 8. The fourth aliquot was analyzed via negative ionization following elution from a HILIC column (Waters UPLC BEH Amide 2.1 × 150 mm, 1.7 µm) using a gradient consisting of water and acetonitrile with 10 mM ammonium formate, pH 10.8. The MS analysis alternated between MS and data-dependent MSn scans using dynamic exclusion. The scan range varied slightly between methods but covered 70–1000 *m/z*.

Instrument variability was 5%. This was determined by calculating the median relative standard deviation (RSD) for the standards that were added to each sample prior to injection into the mass spectrometers. Overall process variability was 10%. This was determined by calculating the median RSD for all endogenous metabolites (that is, non-instrument standards) present in 100% of the pooled matrix samples.

Raw data were extracted, peak-identified and QC processed using hardware and software of Metabolon Inc. Compounds were identified by comparison with library entries of purified standards or recurrent unknown entities. Biochemical identifications are based on three criteria: retention index within a narrow retention time/index (RI) window of the proposed identification, accurate mass match to the library \pm 10 ppm, and the MS/MS forward and reverse scores

between the experimental data and authentic standards. While there may be similarities between these molecules based on one of these factors, the use of all three data points can be utilized to distinguish and differentiate biochemicals. Peaks were quantified using the area-under-the-curve method.

RNA sequencing (RNA-seq)

RNA extraction and double-stranded cDNA library preparation

RNA extraction was performed using the Qiagen RNeasy kit (Qiagen, Manchester, UK) according to the manufacturer's instructions. The samples were checked for RNA quality using the Agilent 2100 Bioanalyzer (Agilent Technologies UK Limited, Craven Arms, UK) and quantified using a Nanodrop instrument (ND-1000 Spectrophotometer; Thermo Fisher Scientific Inc.). The extracted RNA was subsequently subjected to treatment with DNase I (Thermo Fisher Scientific Cat. no: AM222) followed by purification with Agencourt Ampure XP beads (Thermo Fisher Scientific Cat. no: A63881). A 20 ng aliquot of intact high-quality total RNA (RIN > 9) from each sample was then used as input to generate libraries for RNA-seq using the NEBNext Ultra II Directional RNA kit (NEB, Hitchin Herts, UK; Cat. no: E7760S) following the manufacturer's recommendations. This protocol involved an initial step of rRNA depletion, followed by fragmentation prior to first-strand cDNA synthesis and barcoding of the second-strand cDNA synthesised with indices for Illumina platform sequencing for final library amplification. The resulting libraries were assessed on the Bioanalyzer 2100 for purity. Sequencing of the resultant double-stranded (ds) cDNA libraries was performed by applying Illumina sequencing by synthesis technology.

Sequencing of ds-cDNA libraries

Out of a total of 40 cDNA libraries generated, 15 were randomly selected and assessed for quality and fragment size distribution using the Agilent 2200 TapeStation (Agilent Technologies UK Limited) prior to sequencing. All samples showed libraries of good quality, with minimal adapter-dimer contamination, and an average fragment size of 345 bp. All 40 libraries were quantified using the Qubit 2.0 spectrophotometer and the high-sensitivity dsDNA Qubit reagent kit (Life Technologies, Thermo Fisher Scientific, Loughborough, UK). An equimolar quantity of each library was pooled for sequencing, and the resulting complete pool was quality checked using the Agilent 2200 TapeStation. Sequencing consisted of generating 75-bp

paired-end reads for the final pool of 40 libraries using the Illumina NextSeq[®]500 instrument in conjunction with the NextSeq[®]500 v2 High-output 150-cycle kit (Illumina Inc., Cambridge, UK).

Statistical analysis

The metabolome data analysis was performed using MetaboAnalyst 3.0 (Xia and Wishart 2016). We removed variables with more than 50% of missing values. We also replaced remaining missing values with a small value (half of the minimum positive values in the original data) assuming to be the detection limit. Furthermore, variables that show low repeatability as measured using the relative standard deviation were removed. This step is recommended for untargeted metabolomics datasets with a large number of variables to remove baseline noise (Hackstadt and Hess 2009). Data were then median scaled, log transformed, and normalized to protein concentration, as determined using the Bradford reagent, from the cell pellets. A principal component analysis (PCA) was first performed to inspect the data variance and see how the samples are related to each other. This was then followed by an orthogonal projection to latent structures discriminant analysis (OPLS-DA) to identify the source of variation between control and treated groups (Worley and Powers 2016). OPLS-DA is an extension of the PLS-DA method, which incorporates an orthogonal component distinguishing the variability corresponding to the experimental perturbation (here the effect of PCB 126) from the portion of the data which is orthogonal, that is, independent from the experimental perturbation. Although PCA is an unsupervised method, the OPLS-DA is supervised. The separation between the different classes is calculated by maximizing the covariance between the metabolome matrix (x) and the experimental group labels (y). It is thus prone to overfitting, which we addressed by performing permutation testing. Statistically significant alterations in the levels of metabolites were identified based on the significance after analysis with a one-way ANOVA test adjusted for multiple comparisons with Fisher's least significant difference ($q < 0.05$). The metabolome network visualization was performed with MetExplore as described (Cottret et al. 2010). Briefly, the sub-network was obtained by keeping reactions belonging to shortest paths up to a length of 4 between pairs of metabolites in the fingerprint. Network was modelled based on Recon2 (Thiele et al. 2013) human metabolic network using atom conservation graph (Frainay and Jourdan 2017) where two metabolites are connected if they are substrate and product of the same reaction and share at least one carbon (excluding carbon dioxide from the network).

The RNA-seq data analysis was performed using the latest version of the Tuxedo protocol with HISAT, StringTie and Ballgown (Pertea et al. 2016). First, we analyzed

the quality scores and other metrics using FASTQC (Andrews 2010). Contamination from rRNA was measured by aligning the sequences of human rRNA to the FASTQ files (<http://genomespot.blogspot.co.uk/2015/08/scree-n-for-rRNA-contamination-in-rna.html>). Sequences were then aligned to the human genome (Supplementary File 1) using the hierarchical indexing for spliced alignment of transcripts program HISAT2 (Kim et al. 2015). For this purpose, we used prebuilt indexes (*H. sapiens*, GRCh38) downloaded from the HISAT2 website. Alignment rates ranged from 85.12 to 95.53% (average of 91.63%). The number of reads per samples ranged from 9,467,558 to 35,623,876 (average of 21,575,770 reads). Then, StringTie was used to assemble and quantify the transcripts in each sample using the *Homo_sapiens.GRCh38.89* database (Pertea et al. 2016). Finally, the differential gene expression analysis was made with Ballgown (Frazee et al. 2015) in R environment (Team 2017). Low abundance transcripts with a variance across samples of less than one were filtered. A standard linear model-based comparison of transcript abundance was performed without adjusting for other covariates to identify differentially expressed transcripts ($q < 0.05$).

Functional implications of the alteration in gene expression profiles were analyzed using ClueGO (Bindea et al. 2009) and CluePedia plugins in Cytoscape (version 3.5.1) (Shannon et al. 2003). The GO biological process database (23.02.2017) and the KEGG annotation database (01.03.2017) were used. The analysis was conducted using a two-sided hypergeometric test for enrichment using a p value threshold of 0.05 after its adjustment by the Benjamini–Hochberg procedure (Benjamini and Hochberg 1995). GO term fusion was employed to integrate GO categories, minimize the output, and create a functionally organized GO category network. Our network displayed GO terms found in the levels 5–10 of the GO hierarchy, to avoid meaningless high-level hierarchy terms. Since it is well known that gene-annotation enrichment tools can give different results, we corroborated our interpretations by undertaking an additional KEGG functional analysis using the Wallenius non-central hypergeometric distribution in the Bioconductor package GOSep (Young et al. 2010). This analysis has the advantage of correcting for the gene selection bias due to differences in transcript length present in RNA-seq datasets. KEGG annotation enrichment profiles were found to be very similar to the results obtained with ClueGO. These RNA-seq data have been submitted to Gene Omnibus and are accessible through accession number GSE109565.

The metabolome and transcriptome datasets were then integrated by a sparse partial least squares regression (sPLS) performed with the MixOmics package. PLS is a supervised method, which selects correlated variables (genes,

metabolites) across the same samples by maximizing the covariance between the datasets. The sparse version of PLS (sPLS) was used to select the most correlated variables using LASSO penalization on the pair of loading vectors as described by Le Cao et al. (2008).

Results

We determined transcriptome and metabolome signatures of the effect of a 6-day exposure to PCB 126 in differentiated HepaRG human liver cells (Fig. 1). The cells presented no visible signs of aging, or phenotypic alterations caused by PCB 126, at the time point used to evaluate effects. This was verified by measuring lipid accumulation using the AdipoRed assay (Supplementary File 5). PCB 126 treatment of HepaRG cells gave rise to a non-significant increase in lipid accumulation at the two lowest concentrations (100 pM, 10 nM) tested. This validates our approach since the first aim of our experiments was to identify the primary effects of PCB 126 and not secondary effects due to the accumulation of lipids.

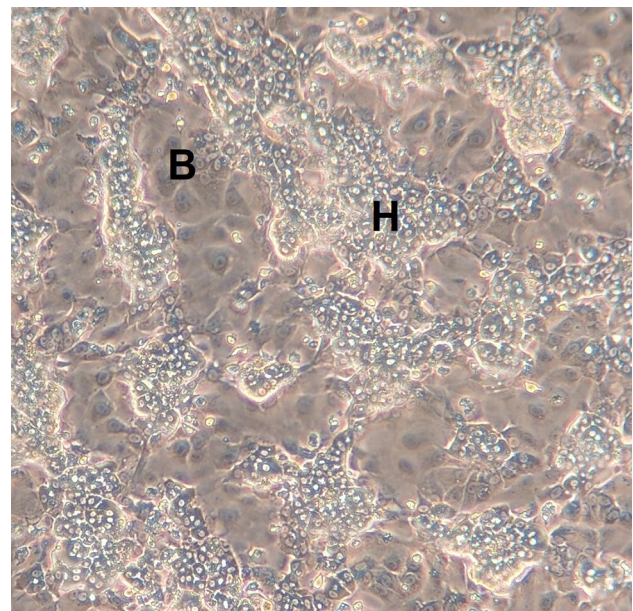


Fig. 1 Morphology of HepaRG cells. After undergoing a complete programme of hepatocyte differentiation, HepaRG cells display the phenotype reflective of normal human liver cells including binuclear hepatocytes and forming bile canaliculus-like structures. A mixed population of two types of cells is visible, namely, hepatocyte-like colonies (H) surrounded by clear epithelial cells corresponding to primitive biliary cells (B)

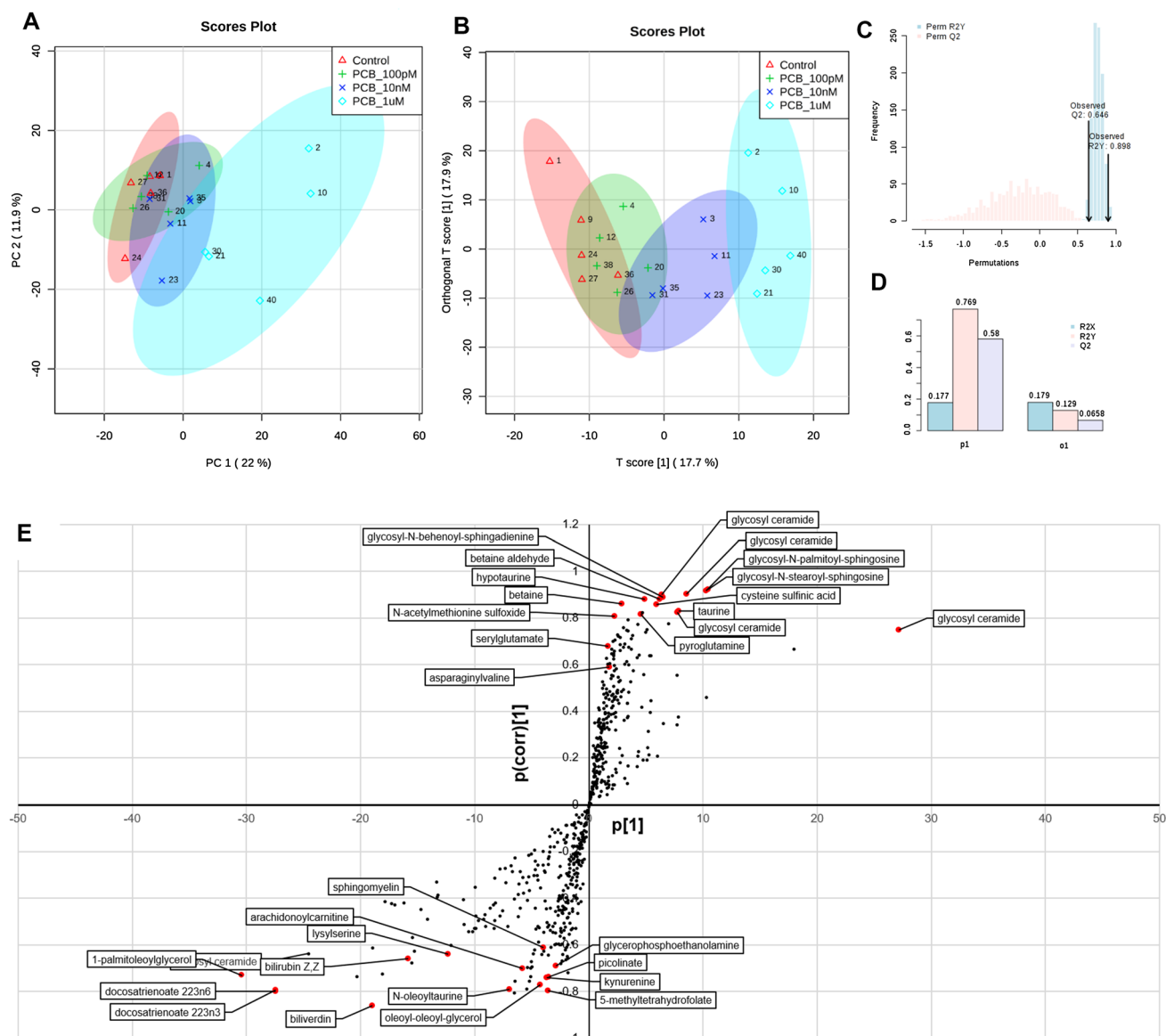


Fig. 2 Multivariate analysis of HepaRG cell metabolome following treatment with PCB 126 shows alterations in lipid metabolism. **a** Principal component analysis of metabolome profiles separate the group of samples treated with the PCB126 from the control group. As the dose of PCB 126 increases, the groups become more clustered. **b** Orthogonal projection to latent structures discriminant analysis (OPLS-DA) properly classified all samples ($R2X=0.177$, $R2Y=0.769$, and $Q2=0.58$). The 95% confidence regions are displayed by shaded ellipses. **c** A 1000-time permutation test shows that the observed statistics is not part of the distribution formed by

the statistics from the permuted data ($R2Y p=0.022$; $Q2 p<0.001$). **d** Cross-validation parameters, $R2$ and $Q2$, representing the quality of the model. **e** The S-plot visualizes the variable influence in the OPLS-DA model. Significantly disturbed metabolites towards the separation in OPLS-DA models (red dots) were selected based on the significance threshold of $q<0.05$ after analysis with one-way ANOVA test adjusted for multiple comparisons with Fisher's least significant difference. A total of 30 metabolites had their levels disturbed by the PCB 126 treatments

Metabolomics analysis

The metabolome platform detected 802 metabolites. We removed 7 variables with more than 50% missing values, and 199 variables that showed low repeatability. The effects of the PCB 126 were first visualized by plotting each sample as a point in space defined by the two principal components

from a PCA, to reduce the 596-dimensional space defined by the variations in the metabolites levels (Fig. 2a). The components separated the group of samples treated with the PCB 126 from the control group. As the dose of PCB 126 increases, the groups become more clustered. We then built an OPLS-DA model (Fig. 2b) on the basis of the PCA results. The OPLS-DA model for PCB 126 ($R2X=0.177$,

R2Y = 0.769, and Q2 = 0.58) appropriately classified all samples (Fig. 2d). This model explains 17.7% of the variation in metabolite levels (R2X) and 76.9% of the variations between the different groups (R2Y). The average prediction capability (Q2) was 58%. The difference between R2Y and Q2 was less than 0.2 and the Q2 value was greater than 50%, revealing an excellent predictive capability (Robotti and Marengo 2016). A 1000-time permutation test was conducted to further validate the OPLS-DA model (Fig. 2c). The empirical *p* values for R2Y (*p* = 0.022) and Q2 (*p* < 0.001) indicate that the observed statistic (based on our data) is not part of the distribution formed by those from the permuted data. Thus, the dose-dependent clustering of the data is statistically significant.

An S-plot was then constructed to visualize the loading of the PCB 126 OPLS-DA model (Fig. 2e), and determine

which variables are the best discriminators between the different treatment groups. This plot combines the modelled covariance (p_1) and the modelled correlation ($p(\text{corr})$) from the OPLS-DA model in a scatter plot. The highest magnitude and reliability scores were attributable to polyunsaturated fatty acids (docosatrienoate 22:3n3 and 22:3n6) and ceramides. However, because OPLS-DA loading is difficult to interpret with more than two classes, we used the OPLS-DA method for score overview but for significantly disturbed metabolites (red dots, Fig. 2e) selected based on the results of a one-way ANOVA test. A total of 30 metabolites had their levels disturbed by the PCB 126 treatments (Table 1). Most of the changes were dose dependent (Fig. 3). The most pronounced metabolic differences were attributed to a change in lipid metabolism. A total of four glycosylceramides and three glycosyl-sphingosines had their levels

Table 1 Metabolome disturbances provoked by exposure to PCB 126 in HepaRG cells

Biochemical	Super_pathway	Sub_pathway
Pyroglutamine	Amino acid	Glutamate metabolism
Betaine	Amino acid	Glycine, serine and threonine metabolism
Betaine aldehyde	Amino acid	Glycine, serine and threonine metabolism
Cysteine sulfinic acid	Amino acid	Methionine, cysteine, SAM and taurine metabolism
Hypotaurine	Amino acid	Methionine, cysteine, SAM and taurine metabolism
<i>N</i> -Acetylmethionine sulfoxide	Amino acid	Methionine, cysteine, SAM and taurine metabolism
Taurine	Amino acid	Methionine, cysteine, SAM and taurine metabolism
Kynurenine	Amino acid	Tryptophan metabolism
Picolinate	Amino acid	Tryptophan metabolism
5-Methyltetrahydrofolate (5MeTHF)	Cofactors and vitamins	Folate metabolism
Bilirubin (<i>Z,Z</i>)	Cofactors and vitamins	Hemoglobin and porphyrin metabolism
Biliverdin	Cofactors and vitamins	Hemoglobin and porphyrin metabolism
Glycosyl ceramide (d16:1/24:1, d18:1/22:1)	Lipid	Ceramides
Glycosyl ceramide (d18:1/20:0, d16:1/22:0)	Lipid	Ceramides
Glycosyl ceramide (d18:1/23:1, d17:1/24:1)	Lipid	Ceramides
Glycosyl ceramide (d18:2/24:1, d18:1/24:2)	Lipid	Ceramides
Glycosyl- <i>N</i> -behenoyl-sphingadienine (d18:2/22:0)	Lipid	Ceramides
Glycosyl- <i>N</i> -palmitoyl-sphingosine (d18:1/16:0)	Lipid	Ceramides
Glycosyl- <i>N</i> -stearoyl-sphingosine (d18:1/18:0)	Lipid	Ceramides
Oleoyl-oleoyl-glycerol (18:1/18:1)	Lipid	Diacylglycerol
<i>N</i> -Oleoyl taurine	Lipid	Endocannabinoid
Arachidonoyl carnitine (C20:4)	Lipid	Fatty acid metabolism (acyl carnitine)
1-Palmitoleoylglycerol (16:1)	Lipid	Monoacylglycerol
Glycerophosphoethanolamine	Lipid	Phospholipid metabolism
Docosatrienoate (22:3n3)	Lipid	Polyunsaturated fatty acid (n3 and n6)
Docosatrienoate (22:3n6)	Lipid	Polyunsaturated fatty acid (n3 and n6)
Sphingomyelin (d18:1/21:0, d17:1/22:0, d16:1/23:0)	Lipid	Sphingolipid metabolism
Asparaginyl valine	Peptide	Dipeptide
Lysylserine	Peptide	Dipeptide
Serylglutamate	Peptide	Dipeptide

All the metabolites displayed have their levels significantly altered ($q < 0.05$) after analysis with one-way ANOVA test adjusted for multiple comparisons with Fisher's least significant difference

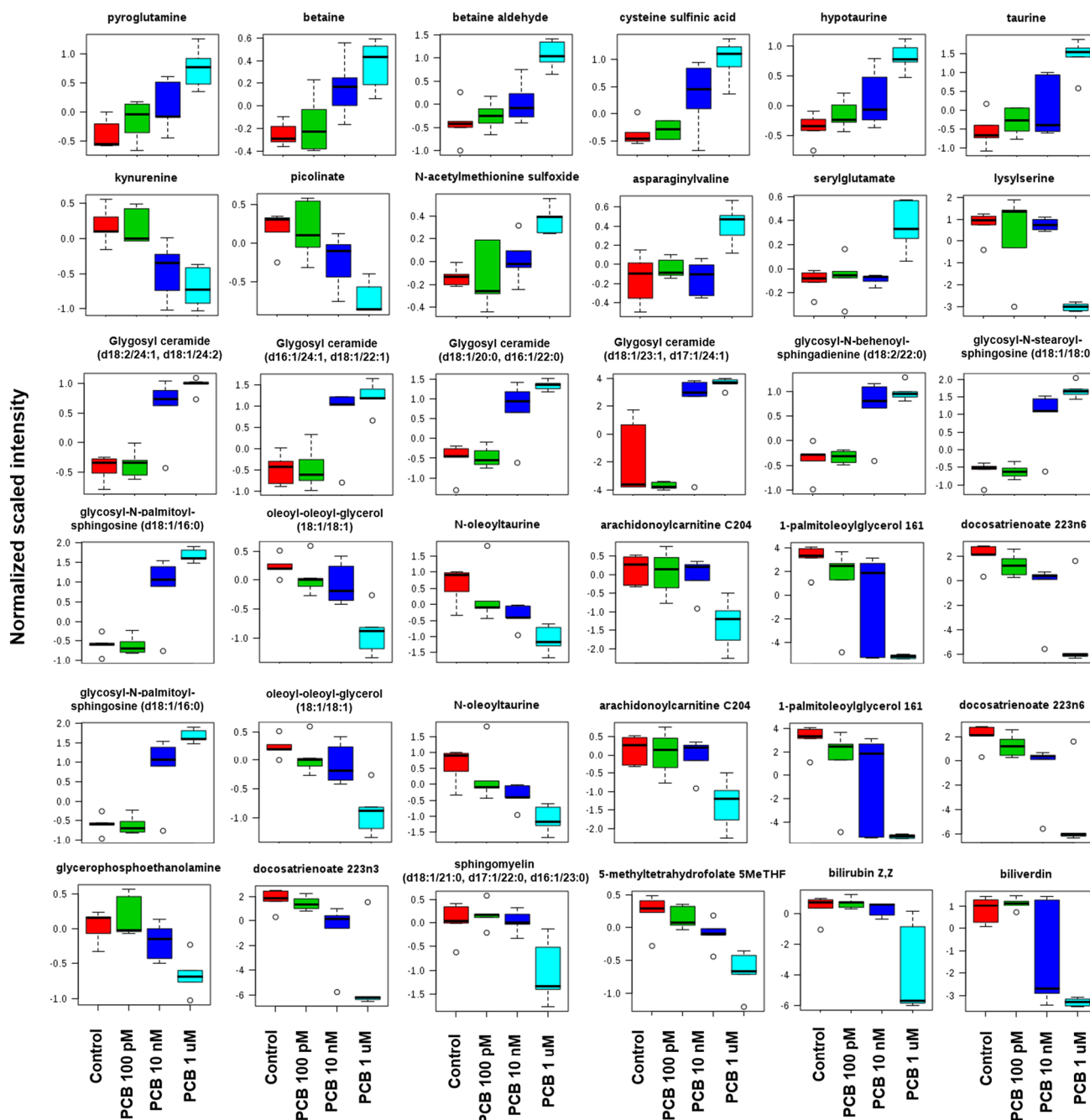


Fig. 3 Box plot of metabolome changes associated with the exposure to PCB 126 to HepaRG cells shows a dose-dependent effect. All the metabolites displayed have their levels significantly altered

($q < 0.05$) after analysis with one-way ANOVA test adjusted for multiple comparisons with Fisher's least significant difference. Most of the changes caused by the PCB treatment were dose dependent

increased [fold change (FC) up to 13.2 for glycosyl ceramide d181/231, d171/241], while polyunsaturated fatty acids such as docosatrienoate 22:3n3 (FC ranging from -1.3 to -5.4), and docosatrienoate 22:3n6 (FC from -1.6 to -6.5) and 1-palmitoleoylglycerol (FC from -1.8 to -370) had their levels decreased. These lipids are common constituents of cell membranes, and thus changes in their levels suggest that

membrane integrity and maintenance is adversely affected by the PCB 126 treatment.

Our results also show an increased activity within detoxification pathways. Multiple metabolites, including cysteine sulfinic acid, taurine, hypotaurine, *N*-oleoyltaurine, and *N*-acetylmethionine sulfoxide are significantly elevated in a dose-dependent manner. The predominant

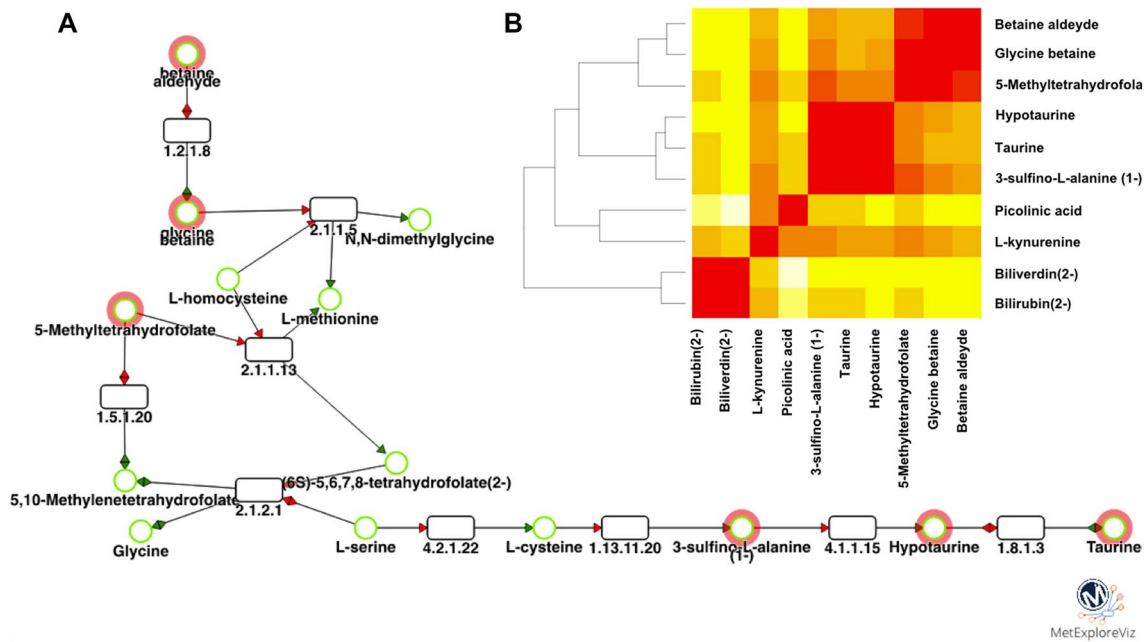


Fig. 4 Network analysis of metabolome profile alterations demonstrates a role of taurine and hypotaurine in oxidative stress induced by PCB 126 in HepaRG cells. **a** Sub-network. Circles are metabolites and rectangles are reactions. Reaction labels are EC numbers and

metabolite names are from Recon2 model. Red circle metabolites are the ones from the fingerprint. **b** Distance matrix between metabolites belonging to the network. Red corresponds to shorter distance (0) and white to longer distances (12 reactions between nodes)

pathway elicited as a response to oxidative stress seems to be taurine metabolism as indicated by the network analysis performed using MetExplore (Fig. 4a). Our network analysis reveals that bilirubin and biliverdin are further away from most metabolites in the fingerprint (Fig. 4b). Thus, the network extraction focused on the first cluster which contains more metabolites (Fig. 4a). The core part of the taurine and hypotaurine metabolism seems to be modulated (disregarding bilirubin and biliverdin) as a consequence of the oxidative stress induced by PCB 126.

Multiple gamma-glutamyl amino acids such as gamma-glutamylglutamine and gamma-glutamylcysteine are also elevated. This is possibly in response to oxidative stress being caused by the PCB 126 exposure. Several activators of aryl hydrocarbon receptor (AhR)-mediated transcription such as the bile pigments biliverdin and bilirubin, as well as tryptophan metabolites kynurenine and picolinate, had their levels decreased by the PCB 126 treatment. Other intermediates in tryptophan metabolism such as quinolinate, kynurenate and pyridoxate were also decreased but did not reach statistical significance.

Transcriptomics analysis

Transcriptome profiles of HepaRG cells were then analyzed using Illumina-based RNA sequencing. An unsupervised visualization of the variance structure by a PCA showed

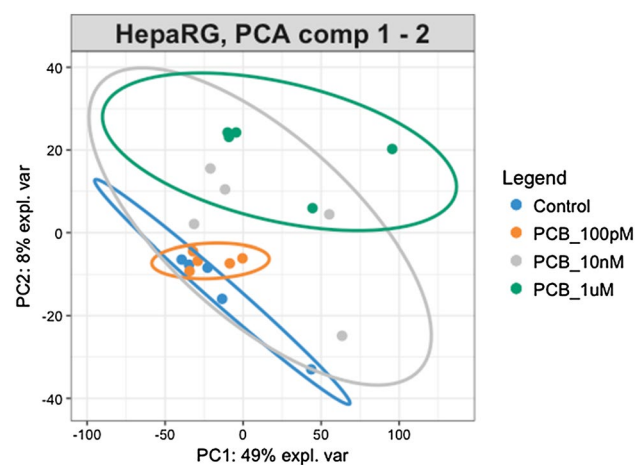


Fig. 5 Principal component analysis of transcriptome profile alterations provoked by exposure of HepaRG cells to PCB 126. Transcript abundances were assessed using Stringtie. The PCA was performed using log₂ transformed FPKM measurements of transcripts across samples assessed with Ballgown. The groups become more clustered as the dose of PCB 126 increases. The 95% confidence regions are displayed by ellipses

comparable results to those of the metabolome profiles. The control group separates from the PCB 126 groups in a dose-dependent manner on the second component (Fig. 5). The source of variation causing the clustering on the first component remains unknown as no batch effects was identified

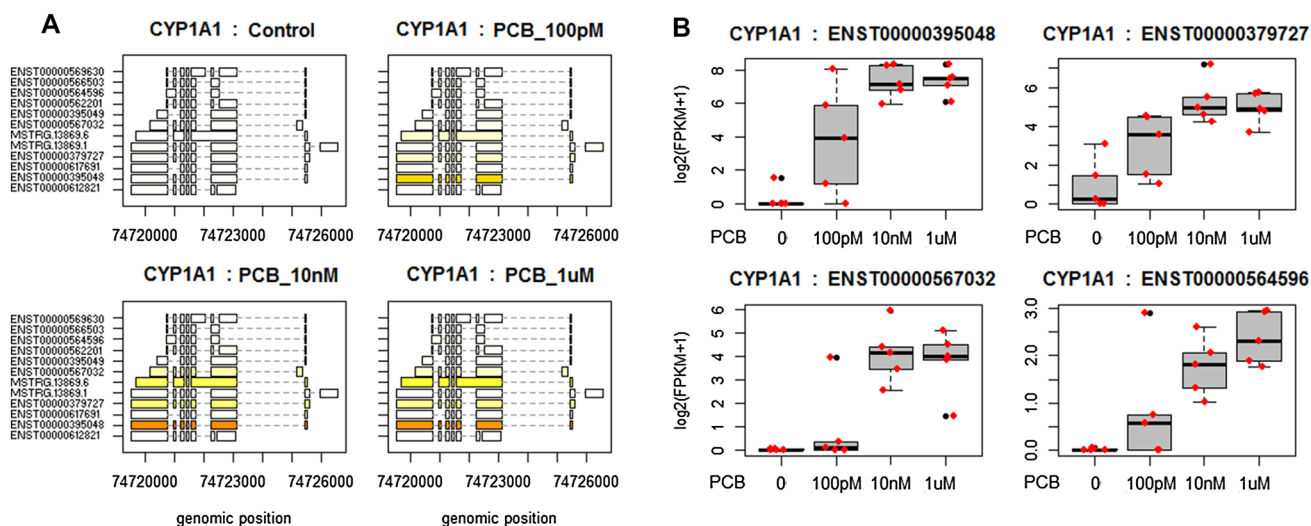


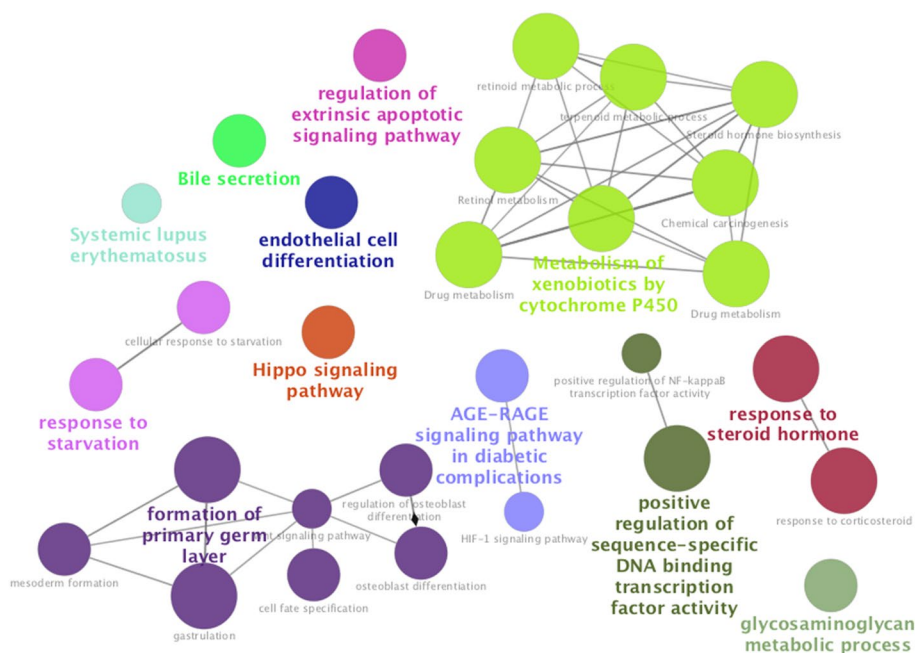
Fig. 6 Differential *CYP1A1* expression analysis using RNA-seq in HepaRG cells exposed to three concentrations of PCB 126. **a** Structure and expression levels of 12 distinct isoforms of *CYP1A1* across the different treatment groups. Differences in expression levels are displayed in varying shades of yellow. The ENST00000395048 isoform of *CYP1A1* is expressed at a much higher level than the others,

as indicated by the dark orange color. **b** FPKM distributions of four *CYP1A1* transcripts displayed as box-and-whiskers plots. All four isoforms have their expression significantly altered ($q < 0.05$) by exposure to PCB 126 as measured by a standard linear model-based comparison in Ballgown

in the experimental procedure. As a result, the standard linear model-based comparison of transcript abundance was performed without adjusting for other covariates. A total of 264 transcripts had their levels altered by the PCB 126. We present a detailed analysis of *CYP1A1* expression as a representative of the changes we observed (Fig. 6). *CYP1A1* is known to be highly expressed following exposure to dioxins and PCBs as its promoter contains seven dioxin-responsive

elements, which bind the aryl hydrocarbon receptor (Li et al. 2014). The evaluation of expression levels among 12 distinct isoforms of *CYP1A1* shows that some transcripts are expressed at a much higher level than others (Fig. 6a). Four of these 12 isoforms had their expression increased from the lowest concentration ($q < 0.05$). A scatter plot of the expression levels for the 24 most differentially expressed transcripts shows that alterations were dose dependent and

Fig. 7 Pathway enrichment analysis in the transcriptome of HepaRG cells exposed to PCB 126 shows an activation of xenobiotic metabolism by dioxin-like compounds. Gene functions were studied using ClueGO and CluePedia plugins in Cytoscape (version 3.5.1). The analysis was conducted using a two-sided hypergeometric test for enrichment using a p value threshold of 0.05 after its adjustment by the Benjamini–Hochberg procedure



detectable at concentrations as low as 100 pM (Supplementary File 2).

We then performed a gene-annotation enrichment analysis to determine the biological processes affected by the exposure to PCB 126 (Fig. 7, Supplementary File 3). Unsurprisingly, the most important networks affected by PCB 126 were related to the metabolism of xenobiotics by P450 cytochromes (CYP1A1, CYP1A2, CYP1B1) and UDP glucuronosyltransferase (UGT1A1, UGT1A3, UGT2A3, UGT2B11), both being well-known PCB metabolic pathways. In addition, the activity of the aldehyde dehydrogenases (ALDH1A3, ALDH3A1), which catalyze oxidation of aromatic aldehydes produced by-products of mono-oxygenation reactions, were also upregulated. These are all hallmarks of an activation of xenobiotic metabolism by dioxin-like compounds (Grimm et al. 2015). However, surprisingly genes involved in the glycosaminoglycan metabolic process were the most enriched category (*AP2A1*, *CEMIP*, *CHSY1*, *FUCA1*, *GALNT5*, *GPC6*, *LUM*). This can reflect an increase in glucuronidation metabolism as the UGT group of enzymes are central players in glycosylation processes. Perhaps one of the most surprising enriched pathways were the systemic lupus erythematosus (SLE) KEGG pathway containing histone genes (*C5*, *H2AFV*, *HIST1H2AB*, *HIST1H2AE*, *HIST1H2AG*, *HIST1H2BO*). The genes for the AGE-RAGE signalling pathway associated with diabetic complications (*AGTRI*, *CDKN1B*, *COL3A1*, *MMP2*, *PRKCA*, *SERPINE1*) were also enriched. The same genes are found in the HIF-1 signalling pathway, which was also enriched. Taken together, these results point to a role of hypoxia in the toxic effects generated by the exposure to PCB 126.

We then integrated the results from the transcriptome and the metabolome by performing a sPLS analysis, with two components being sufficient to model the data (Q2 of 0.1 and 0.17). The performance of the sPLS analysis was weak because of the relatively small number of samples. However, the different groups were well separated (Fig. 8a), with the samples from the two highest concentrations of PCB 126 treatments being clustered together and separate from the control and the lowest dose samples. We then investigated the correlation between variables (Fig. 8b, c). Sphingolipids and polyunsaturated fatty acids (PUFA) were negatively and positively correlated, respectively, with the expression of genes involved in triglyceride metabolism (such as *APOA2*, *CPS1*, *HMGCS2*, *LPL*). That is, the decrease in PUFA levels, as well as the increase in sphingolipid levels, is concomitant to a decrease in the activity of genes involved in lipid metabolism.

Discussion

With the steadily increasing number of synthetic chemicals being released in the environment, the ability to use human cell lines to evaluate toxicity of pollutants and move away from more time-consuming and expensive risk assessment based on in vivo animal experiments, has become one of the challenges of the twenty-first century (Hartung 2009). We present here an in-depth investigation of the toxic effects of the PCB 126, a model chemical for the induction of TAFLD, on the transcriptome and metabolome of human liver HepaRG cells (Merrick et al. 2015; Mueller et al. 2014). We envisage that our investigation would be particularly useful for the establishment of high-throughput assays using the HepaRG cell line, which has been found to be a good model system to predict the results of acute and repeated dose toxicity (Mueller et al. 2014) and to study human xenobiotic metabolism (Guillouzo et al. 2007). HepaRG is one of the best characterized liver cell lines. The most recent application of metabolomics analytical methods on HepaRG cells reported coverage of 2200 metabolites (Cuykx et al. 2017). HepaRG cells are also metabolically competent and more stable than other hepatic cell lines, producing more reproducible data than primary human hepatocytes (Guillouzo et al. 2007). They are thus a reliable system to obtain information on networks of pathways adversely affected by pollutants that could lead to hepatic steatosis (Angrish et al. 2017).

Our study provides interesting insights into the development of sensitive biomarkers for TASH development. The most significant effect of PCB 126 exposure was a decrease in free long-chain fatty acids (LCFAs) and PUFAs. Our results show that docosatrienoate 22:3n3 and 22:3n6 are the most sensitive biomarkers of lipid metabolism disturbances in HepaRG cells (Fig. 1). This is in accord with the results obtained from a birth cohort from a fishing community that linked a high seafood intake, and thus high PCB exposure, to alterations in lipid profiles (Grandjean and Weihe 2003). Other experimental studies suggest that PCBs may affect the utilization of PUFAs by inhibition of $\Delta 5$ - and $\Delta 6$ -desaturation (Matsusue et al. 1999). Such enzyme inhibition in fishing communities lead to deficient formation of arachidonic acid from its precursor, linoleic acid, attenuating the beneficial effects of essential lipids contained in seafood (Grandjean and Weihe 2003).

The analysis of gene expression profiles revealed alterations in pathways reflecting PCB-induced metabolic diseases such as the activation of xenobiotic metabolism by dioxin-like compounds. This can guide the development of gene expression biomarkers, which have been advocated by the US Environmental Protection Agency (EPA) as promising tools to reduce the cost and time of toxicity testing (EPA

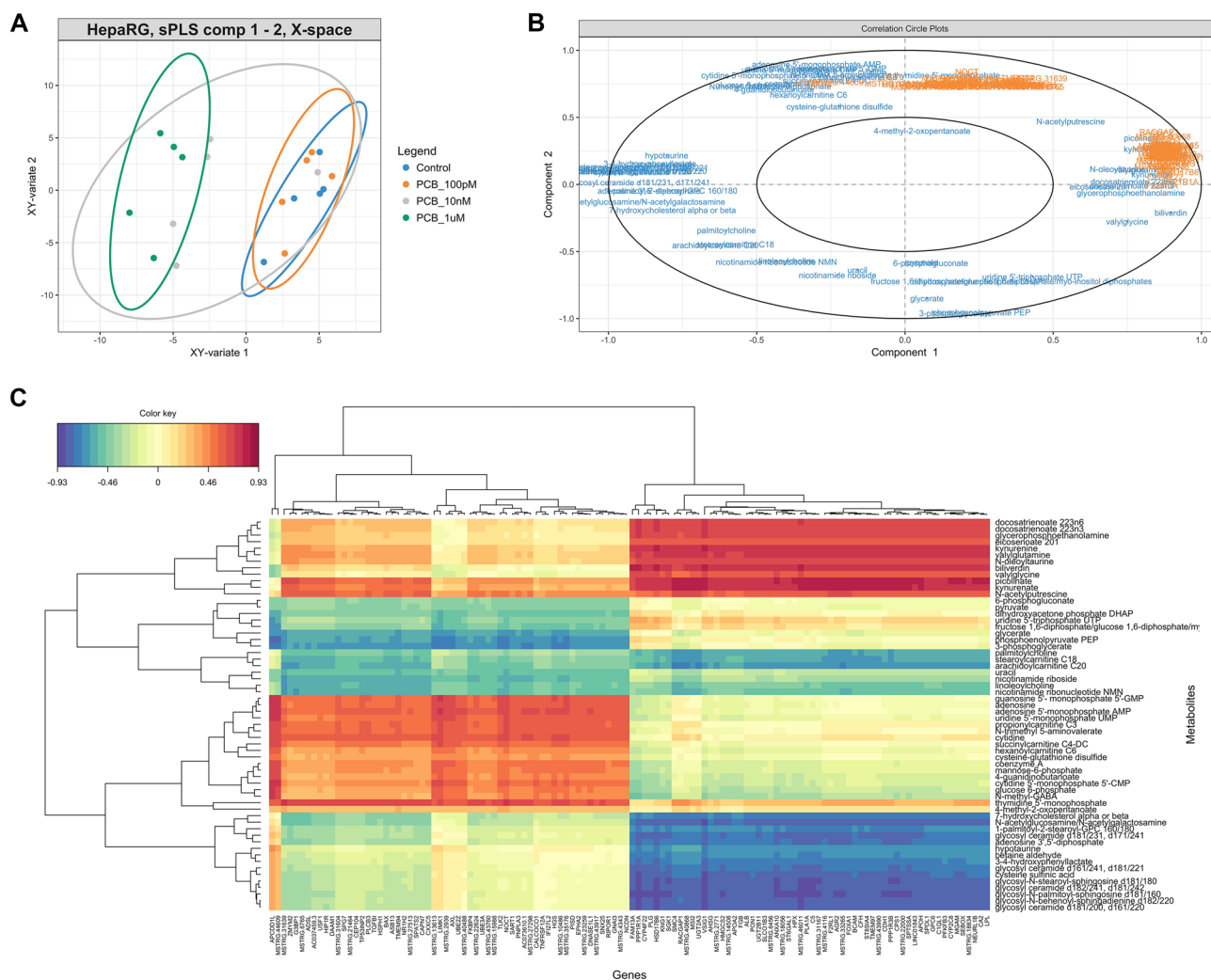


Fig. 8 Sparse partial least squares regression (sPLS) integration of the metabolome and transcriptome profiles of HepaRG cells exposed to PCB 126 shows that alteration in sphingolipid levels is concomitant to a decrease in the activity of genes involved in lipid metabolism. **a** Individual plots displaying the covariance between the metabolome and the transcriptome datasets. **b** The variables selected by the sPLS are projected on a correlation circle plot to display the clusters of correlated variables. In this plot, the angle defined by the coordinates of the variables on the axis defined by the components give an indication on the nature of the correlation. If the angle is sharp and

the variables cluster together, the correlation is positive. If the angle is obtuse and the variables are not clustered together, the variables are negatively correlated. Perpendicular angles represent uncorrelated variables. **c** A clustered image map visualizes correlations between the metabolites and the genes by a color gradient on a two-dimensional colored image. The negatively correlated variables (blue) are represented along the positively correlated variables (red). Dendrograms are added to represent the clusters produced by the hierarchical clustering

2015), such as for the detection of potential EDCs acting as estrogen receptor agonists (Mesnage et al. 2017). Other projects have implemented online tools to provide an easy access to gene expression signatures allowing the detection of novel associations between genes, diseases and drugs (Wang et al. 2016). We provide here putative biomarkers of metabolic disturbances induced by PCB 126, which are known to correlate with the mechanisms by which this class of pollutants induce TASH (Al-Eryani et al. 2015). The expected hallmarks of an activation of xenobiotic metabolism by dioxin-like compounds were detected in the

molecular profiles of HepaRG cells exposed to the PCB 126 (Fig. 7). A limitation of this study is that the changes induced by the PCB 126 were measured at a late time point and may not reflect the first wave of transcriptional changes induced by this compound.

Other pathways altered in our study correlate with other less well-known health effects of PCB compounds. For instance, the SLE KEGG pathway was one of the most altered. This in all likelihood is not coincidental as the exposure to PCBs has been linked to an increased incidence of SLE (Kamen 2014). Histone gene expression was the most

disturbed of the SLE pathways in our study (Supplementary Material 3). This correlates with the fact that patients with SLE often carry anti-histone antibodies (Sun et al. 2008). A link with SLE is further corroborated by laboratory experiments showing that docosahexaenoic acid consumption attenuates the autoimmune response in a mouse model of SLE (Bates et al. 2016). Another autoimmune disease observed with higher levels of PCB and organochlorine pesticide residue exposure is amyotrophic lateral sclerosis (ALS), which was detected in a case–control study measuring chemical pollutants in the cerebrospinal fluid of 38 ALS patients and 38 control individuals (Vinceti et al. 2017). It may be argued that the detection of this gene expression signature in a hepatic cell line is of no direct relevance to SLE. This signature could be an artefact due to the limitations of gene ontology analysis, which could give different results depending on the software, its parameters, and the nucleic acid sequencing technology employed (Timmons et al. 2015). However, comparable responses have been found in other cell types equipped with AhR signalling (Vorrink et al. 2014). Another category of genes whose levels we found altered were those involved in the glycosaminoglycan metabolic process. This correlates with findings obtained from the analysis of the gene ontology enrichment in peripheral blood mononuclear cells of children chronically exposed to PCBs (Dutta et al. 2012). Generally, it is not surprising that we found alterations in gene expression pathways related to neurological processes since PCBs are known neurotoxic compounds with a gene–environment interaction capability (Docea et al. 2017; Ribas-Fito et al. 2001).

Our study also provides mechanistic insights into perturbations provoked by very low environmental concentrations of PCB 126. Although the exposure to persistent organic pollutants such as PCBs is decreasing over time as they are progressively being banned, they are still found in human tissue at levels capable of causing pathologies. PCB 126 was found at concentrations of 209 and 80.4 ng/kg in the breast milk of Inuit and Caucasian women, respectively, in a survey performed in 1989–1990 (Dewailly et al. 1994). The analysis of the National Health and Nutrition Examination Survey (NHANES) data (1999–2002) reveals associations between PCB serum levels and adverse cognitive effects in older US adults (Przybyla et al. 2017). In more recent biomonitoring surveys, PCB 126 was found at a mean concentration of 6.5 µg/L in umbilical cord serum of populations living near waste dump sites in Italy (Grumetto et al. 2015).

Although our study provides extensive insight into the mechanisms of toxicity of a single molecule, the human population is never exposed to single compounds but to complex mixtures (Hernandez et al. 2013). These complex effects cannot be assessed within the current research framework and new methodologies are urgently needed (Tsatsakis et al. 2017). It is generally considered that the effects of chemical

mixtures that act as AhR agonists can be approximately deduced using additive mathematical models (Brennerova et al. 2016). Although it can be assumed that the effects detected in this study may not be specific to PCB 126 but generalizable to other AhR agonists, it should be noted that some non-dioxin-like PCB congeners cause non-additive effects by altering AhR activation by PCB mixtures (Brennerova et al. 2016).

The transcriptome and metabolome signatures that we have identified in response to PCB 126 in HepaRG cells are likely to be similar to what occurs in cases of NAFLD caused by other factors. It is known from metabolome profiles that hepatic PUFAs are lower in NAFLD patients (Arendt et al. 2015). This suggests that the exposure to environmental pollutants could also be an aggravating factor in the development of NAFLD, amplifying the effects of over-nutrition or a sedentary lifestyle. It has been proposed that PCB exposure can act as a second hit in NAFLD, driving progression of steatosis to NASH as has been found in laboratory animals exposed to PCB and a high-fat diet (Wahlang et al. 2014).

Conclusion

Although it is recognized that occupational exposures to chemical pollutants can lead to NAFLD, there is limited information available regarding the mechanism by which typical environmental levels of exposure can contribute to the onset of this disease. Importantly, the alterations in metabolome and transcriptome profiles in response to PCB 126 were observed even at the lowest concentration (100 pM) tested corresponding to what may be found in humans. Overall, our integrated multi-omics analysis provides mechanistic insight into how this class of chemical pollutant can cause NAFLD as demonstrated in other studies (Al-Eryani et al. 2015). In addition, we confirm the HepaRG cell line as a reliable model for molecular profiling investigations of TAFLD. Metabolome profiling reveals docosatrienoate levels as the most reliable marker of the exposure to PCB 126. Hallmarks of activation of the AhR receptor by dioxin-like compounds were detected in the transcriptome profiles. Our study provides the foundation for the development of molecular signatures of fatty liver diseases to rapidly assess chronic toxic effects from exposure to compounds having a direct effect on the liver in a reliable hepatotoxicity model system. The availability of such an *in vitro* model system will lessen the burden of chemical risk assessment based on *in vivo* animal experiments.

Acknowledgements This work was funded by the Sustainable Food Alliance (USA) whose support is gratefully acknowledged.

Author contributions RM participated in the cell culture experiment, performed the metabolome and transcriptome data analysis, and drafted the manuscript. MB performed the cell culture experiment and the RNA extraction. SB performed the library preparation. CF, NP and FJ performed the network analysis of the metabolome data. EW, TX and CAM performed the RNA-seq and assisted for the transcriptome data analysis. MNA coordinated the investigation and assisted with the drafting of the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare they have no competing interests.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Al-Eryani L, Wahlang B, Falkner KC, Guardiola JJ, Clair HB, Prough RA, Cave M (2015) Identification of environmental chemicals associated with the development of toxicant-associated fatty liver disease in rodents. *Toxicol Pathol* 43:482–497. <https://doi.org/10.1177/0192623314549960>
- Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/doi:citeulike-article-id:11583827>
- Angrish MM, McQueen CA, Cohen-Hubal E, Bruno M, Ge Y, Chorley BN (2017) Editor's highlight: mechanistic toxicity tests based on an adverse outcome pathway network for hepatic steatosis. *Toxicol Sci* 159:159–169. <https://doi.org/10.1093/toxsci/kfx121>
- Arendt BM et al (2015) Altered hepatic gene expression in non-alcoholic fatty liver disease is associated with lower hepatic n-3 and n-6 polyunsaturated fatty acids. *Hepatology* 61:1565–1578. <https://doi.org/10.1002/hep.27695>
- Argyrou C, Moris D, Vernadakis S (2017) Hepatocellular carcinoma development in non-alcoholic fatty liver disease and non-alcoholic steatohepatitis. Is it going to be the “Plague” of the 21st century? A literature review focusing on pathogenesis, prevention and treatment. *J BUON* 22:6–20
- Ba Q et al (2017) Sex-dependent effects of cadmium exposure in early life on gut microbiota and fat accumulation in mice. *Environ Health Perspect* 125:437–446. <https://doi.org/10.1289/EHP360>
- Bates MA et al (2016) Silica-triggered autoimmunity in lupus-prone mice blocked by docosahexaenoic acid consumption. *PLoS One* 11:e0160622. <https://doi.org/10.1371/journal.pone.0160622>
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B (Methodological)* 57:289–300
- Bindea G et al (2009) ClueGO: a cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 25:1091–1093. <https://doi.org/10.1093/bioinformatics/btp101>
- Brenerova P, Hamers T, Kamstra JH, Vondracek J, Strapacova S, Andersson PL, Machala M (2016) Pure non-dioxin-like PCB congeners suppress induction of AhR-dependent endpoints in rat liver cells. *Environ Sci Pollut Res Int* 23:2099–2107. <https://doi.org/10.1007/s11356-015-4819-6>
- Cave M, Appana S, Patel M, Falkner KC, McClain CJ, Brock G (2010a) Polychlorinated biphenyls, lead, and mercury are associated with liver disease in American adults: NHANES 2003–2004. *Environ Health Perspect* 118:1735–1742. <https://doi.org/10.1289/ehp.1002720>
- Cave M et al (2010b) Toxicant-associated steatohepatitis in vinyl chloride workers. *Hepatology* 51:474–481. <https://doi.org/10.1002/hep.23321>
- Cottret L, Wildridge D, Vinson F, Barrett MP, Charles H, Sagot MF, Jourdan F (2010) MetExplore: a web server to link metabolomic experiments and genome-scale metabolic networks. *Nucleic Acids Res* 38:W132–W137. <https://doi.org/10.1093/nar/gkq312>
- Cuykx M et al (2017) Tailored liquid chromatography-mass spectrometry analysis improves the coverage of the intracellular metabolome of HepaRG cells. *J Chromatogr A* 1487:168–178. <https://doi.org/10.1016/j.chroma.2017.01.050>
- Dewailly E, Ryan JJ, Laliberte C, Bruneau S, Weber JP, Gingras S, Carrier G (1994) Exposure of remote maritime populations to coplanar PCBs. *Environ Health Perspect* 102(Suppl 1):205–209
- Docea AO et al (2017) CYP polymorphisms and pathological conditions related to chronic exposure to organochlorine pesticides. *Toxicol Rep* 4:335–341. <https://doi.org/10.1016/j.toxrep.2017.05.007>
- Dutta SK et al (2012) Differential gene expression and a functional analysis of PCB-exposed children: understanding disease and disorder development. *Environ Int* 40:143–154. <https://doi.org/10.1016/j.envint.2011.07.008>
- EPA US (2015) Endocrine disruptor screening program: use of high throughput assays and computational tools. Notice. Federal Register (2015–15182, Jun 19, 2015) (FRL-9928-69)
- Fabricio G, Malta A, Chango A, De Freitas Mathias PC (2016) Environmental contaminants and pancreatic beta-cells. *J Clin Res Pediatr Endocrinol* 8:257–263. <https://doi.org/10.4274/jcrpe.2812>
- Foulds CE, Trevino LS, York B, Walker CL (2017) Endocrine-disrupting chemicals and fatty liver disease. *Nat Rev Endocrinol* 13:445–457. <https://doi.org/10.1038/nrendo.2017.42>
- Frainay C, Jourdan F (2017) Computational methods to identify metabolic sub-networks based on metabolomic profiles. *Brief Bioinform* 18:43–56. <https://doi.org/10.1093/bib/bbv115>
- Frazeo AC, Perteu G, Jaffe AE, Langmead B, Salzberg SL, Leek JT (2015) Ballgown bridges the gap between transcriptome assembly and expression analysis. *Nat Biotechnol* 33:243–246. <https://doi.org/10.1038/nbt.3172>
- Grandjean P, Weihe P (2003) Arachidonic acid status during pregnancy is associated with polychlorinated biphenyl exposure. *Am J Clin Nutr* 77:715–719
- Grimm FA et al (2015) Metabolism and metabolites of polychlorinated biphenyls. *Crit Rev Toxicol* 45:245–272. <https://doi.org/10.3109/10408444.2014.999365>
- Grumetto L, Ortosecco G, Russo G, Guida M, Ferranti P, Nasi A, Barbato F (2015) Dioxin-like PCB levels in maternal and umbilical cord sera of people living near dump sites in southern Italy: a pilot study of biomonitoring. *Environ Monit Assess* 187:88. <https://doi.org/10.1007/s10661-015-4267-2>
- Guillouzo A, Corlu A, Aninat C, Glaise D, Morel F, Guguen-Guillouzo C (2007) The human hepatoma HepaRG cells: a highly differentiated model for studies of liver metabolism and toxicity of xenobiotics. *Chem Biol Interact* 168:66–73. <https://doi.org/10.1016/j.cbi.2006.12.003>
- Hackstadt AJ, Hess AM (2009) Filtering for increased power for microarray data analysis. *BMC Bioinform* 10:11. <https://doi.org/10.1186/1471-2105-10-11>
- Hartung T (2009) Toxicology for the twenty-first century. *Nature* 460:208–212. <https://doi.org/10.1038/460208a>
- Heindel JJ et al (2017) Metabolism disrupting chemicals and metabolic disorders. *Reprod Toxicol* 68:3–33. <https://doi.org/10.1016/j.reprotox.2016.10.001>

- Hernandez AF, Parron T, Tsatsakis AM, Requena M, Alarcon R, Lopez-Guarnido O (2013) Toxic effects of pesticide mixtures at a molecular level: their relevance to human health. *Toxicology* 307:136–145. <https://doi.org/10.1016/j.tox.2012.06.009>
- Janesick AS, Blumberg B (2016) Obesogens: an emerging threat to public health. *Am J Obstet Gynecol* 214:559–565. <https://doi.org/10.1016/j.ajog.2016.01.182>
- Kaiser JP, Lipscomb JC, Wesselkamper SC (2012) Putative mechanisms of environmental chemical-induced steatosis. *Int J Toxicol* 31:551–563. <https://doi.org/10.1177/1091581812466418>
- Kamen DL (2014) Environmental influences on systemic lupus erythematosus expression. *Rheum Dis Clin N Am* 40:401–412. <https://doi.org/10.1016/j.rdc.2014.05.003>
- Kim D, Langmead B, Salzberg SL (2015) HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* 12:357–360. <https://doi.org/10.1038/nmeth.3317>
- Le Cao KA, Rossouw D, Robert-Granic C, Besse P (2008) A sparse PLS for variable selection when integrating omics data. *Stat Appl Genet Mol Biol* 7:35. <https://doi.org/10.2202/1544-6115.1390>
- Li S, Pei X, Zhang W, Xie HQ, Zhao B (2014) Functional analysis of the dioxin response elements (DREs) of the murine CYP1A1 gene promoter: beyond the core DRE sequence. *Int J Mol Sci* 15:6475–6487. <https://doi.org/10.3390/ijms15046475>
- Matsusue K, Ishii Y, Ariyoshi N, Oguri K (1999) A highly toxic coplanar polychlorinated biphenyl compound suppresses Delta5 and Delta6 desaturase activities which play key roles in arachidonic acid synthesis in rat liver. *Chem Res Toxicol* 12:1158–1165
- Merrick BA, Paules RS, Tice RR (2015) Intersection of toxicogenomics and high throughput screening in the Tox21 program: an NIEHS perspective. *Int J Biotechnol* 14:7–27. <https://doi.org/10.1504/IJBT.2015.074797>
- Mesnager R, Pheonon A, Arno M, Balu S, Corton JC, Antoniou MN (2017) Editor's highlight: transcriptome profiling reveals bisphenol a alternatives activate estrogen receptor alpha in human breast cancer cells. *Toxicol Sci* 158:431–443. <https://doi.org/10.1093/toxsci/kfx101>
- Michelotti GA, Machado MV, Diehl AM (2013) NAFLD, NASH and liver cancer. *Nat Rev Gastroenterol Hepatol* 10:656–665. <https://doi.org/10.1038/rgastro.2013.183>
- Mueller D, Kramer L, Hoffmann E, Klein S, Noor F (2014) 3D organotypic HepaRG cultures as in vitro model for acute and repeated dose toxicity studies. *Toxicol In Vitro* 28:104–112. <https://doi.org/10.1016/j.tiv.2013.06.024>
- Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL (2016) Transcript-level expression analysis of RNA-seq experiments with HISAT StringTie and Ballgown. *Nat Protoc* 11:1650–1667. <https://doi.org/10.1038/nprot.2016.095>
- Petrakis D et al (2017) Endocrine disruptors leading to obesity and related diseases. *Int J Environ Res Public Health* 14:1282. <https://doi.org/10.3390/ijerph14101282>
- Przybyla J, Houseman EA, Smit E, Kile ML (2017) A path analysis of multiple neurotoxic chemicals and cognitive functioning in older US adults (NHANES 1999–2002). *Environ Health* 16:19. <https://doi.org/10.1186/s12940-017-0227-3>
- RC Team (2017) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <https://www.r-project.org/>
- Ribas-Fito N, Sala M, Kogevinas M, Sunyer J (2001) Polychlorinated biphenyls (PCBs) and neurological development in children: a systematic review. *J Epidemiol Commun Health* 55:537–546
- Robotti E, Marengo E (2016) Chemometric multivariate tools for candidate biomarker identification: LDA, PLS-DA, SIMCA, Ranking-PCA. In: Marengo E, Robotti E (eds) 2-D PAGE map analysis: methods and protocols. Springer, New York, pp 237–267. https://doi.org/10.1007/978-1-4939-3255-9_14
- Shannon P et al (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504. <https://doi.org/10.1101/gr.1239303>
- Sookoian S, Pirola CJ (2017) Genetic predisposition in nonalcoholic fatty liver disease. *Clin Mol Hepatol* 23:1–12. <https://doi.org/10.3350/cmh.2016.0109>
- Sun XY, Shi J, Han L, Su Y, Li ZG (2008) Anti-histones antibodies in systemic lupus erythematosus: prevalence and frequency in neuropsychiatric lupus. *J Clin Lab Anal* 22:271–277. <https://doi.org/10.1002/jcla.20248>
- Tang-Peronard JL, Heitmann BL, Andersen HR, Steuerwald U, Grandjean P, Weihe P, Jensen TK (2014) Association between prenatal polychlorinated biphenyl exposure and obesity development at ages 5 and 7 year: a prospective cohort study of 656 children from the Faroe Islands. *Am J Clin Nutr* 99:5–13. <https://doi.org/10.3945/ajcn.113.066720>
- Thiele I et al (2013) A community-driven global reconstruction of human metabolism. *Nat Biotechnol* 31:419–425. <https://doi.org/10.1038/nbt.2488>
- Timmons JA, Szkop KJ, Gallagher IJ (2015) Multiple sources of bias confound functional enrichment analysis of global-omics data. *Genome Biol* 16:186. <https://doi.org/10.1186/s13059-015-0761-7>
- Tsatsakis AM et al (2017) Simulating real-life exposures to uncover possible risks to human health: a proposed consensus for a novel methodological approach. *Hum Exp Toxicol* 36:554–564. <https://doi.org/10.1177/0960327116681652>
- Vernon G, Baranova A, Younossi ZM (2011) Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. *Aliment Pharmacol Ther* 34:274–285. <https://doi.org/10.1111/j.1365-2036.2011.04724.x>
- Vinceti M et al (2017) Pesticides, polychlorinated biphenyls and polycyclic aromatic hydrocarbons in cerebrospinal fluid of amyotrophic lateral sclerosis patients: a case-control study. *Environ Res* 155:261–267. <https://doi.org/10.1016/j.envres.2017.02.025>
- Vorriink SU, Severson PL, Kulak MV, Futscher BW, Domann FE (2014) Hypoxia perturbs aryl hydrocarbon receptor signaling and CYP1A1 expression induced by PCB 126 in human skin and liver-derived cell lines. *Toxicol Appl Pharmacol* 274:408–416. <https://doi.org/10.1016/j.taap.2013.12.002>
- Wahlang B, Beier JI, Clair HB, Bellis-Jones HJ, Falkner KC, McClain CJ, Cave MC (2013) Toxicant-associated steatohepatitis. *Toxicol Pathol* 41:343–360. <https://doi.org/10.1177/0192623312468517>
- Wahlang B et al (2014) Evaluation of Aroclor 1260 exposure in a mouse model of diet-induced obesity and non-alcoholic fatty liver disease. *Toxicol Appl Pharmacol* 279:380–390. <https://doi.org/10.1016/j.taap.2014.06.019>
- Wang Z et al (2016) Extraction and analysis of signatures from the Gene Expression Omnibus by the crowd. *Nat Commun* 7:12846. <https://doi.org/10.1038/ncomms12846>
- Worley B, Powers R (2016) PCA as a practical indicator of OPLS-DA model reliability. *Curr Metabolom* 4:97–103. <https://doi.org/10.2174/2213235X04666160613122429>
- Xia J, Wishart DS (2016) Using MetaboAnalyst 3.0 for comprehensive metabolomics data analysis. *Curr Protoc Bioinform* 55(14):11–14. <https://doi.org/10.1002/cpbi.11091>
- Young MD, Wakefield MJ, Smyth GK, Oshlack A (2010) Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol* 11:R14. <https://doi.org/10.1186/gb-2010-11-2-r14>
- Younossi ZM et al (2016) The economic and clinical burden of nonalcoholic fatty liver disease in the United States and Europe. *Hepatology* 64:1577–1586. <https://doi.org/10.1002/hep.28785>
- Yu ML, Guo YL, Hsu CC, Rogan WJ (1997) Increased mortality from chronic liver disease and cirrhosis 13 years after the Taiwan “yucheng” (“oil disease”) incident. *Am J Ind Med* 31:172–175