Systems analysis of the liver transcriptome in adult male zebrafish exposed to the non-ionic surfactant nonylphenol

Huff, M., da Silveira, W., Starr Hazard, E., Courtney, S. M., Renaud, L., & Hardiman, G. (2019). Systems analysis of the liver transcriptome in adult male zebrafish exposed to the non-ionic surfactant nonylphenol. *General and Comparative Endocrinology*, *271*, 1-14. https://doi.org/10.1016/j.ygcen.2018.10.016

Published in:
General and Comparative Endocrinology

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Systems analysis of the liver transcriptome in adult male zebrafish exposed to the non-ionic surfactant nonylphenol

Matthew Huff[a,b], Willian da Silveira[a,c], E. Starr Hazard[a], Sean M. Courtney[a], Ludivine Renaud[d], Gary Hardiman[a,d,e,f,g,h,*]

[a] MUSC Bioinformatics, Center for Genomics Medicine, Medical University of South Carolina, Charleston, SC 29415, United States

[b] MS in Biomedical Sciences Program, Medical University of South Carolina, United States

[c] Department of Pathology and Laboratory Medicine, Medical University of South Carolina, United States

[d] Department of Medicine, Medical University of South Carolina, United States

[e] Department of Medicine, University of California San Diego, United States

[f] Department of Public Health Sciences, Medical University of South Carolina, United States

[g] Laboratory for Marine Systems Biology, Hollings Marine Laboratory, Charleston, SC 29412, United States

[h] Institute for Global Food Security, Queens University Belfast, Stranmillis Road, Belfast BT9 5AG, UK

Abstract

Nonylphenol (NP) arises from the environmental degradation of nonylphenol ethoxylates. It is a ubiquitous environmental contaminant and has been detected at levels up to 167 nM in rivers in the United States. NP is an endocrine disruptor (ED) that can act as an agonist for estrogen receptors. The Adverse Outcome Pathway (AOP) framework defines an adverse outcome as the causal result of a series of molecular initiating events (MIEs) and key events (KEs) that lead to altered phenotypes. This study examined the liver transcriptome after a 21 day exposure to NP and 17β-estradiol (E2) by exploiting the zebrafish (Danio rerio) as a systems toxicology model. The goal of this study was to tease out non-estrogenic genomic signatures associated with NP exposure using DNA microarray and RNA sequencing. Our experimental design included E2 as a positive and potent estrogenic control in order to effectively compare and contrast the 2 compounds. This approach allowed us to identify hepatic transcriptomic perturbations that could serve as MIEs for adverse health outcomes in response to NP. Our results revealed that exposure to NP was associated with differential expression (DE) of genes associated with the development of steatosis, disruption of metabolism, altered immune response, and metabolism of reactive oxygen species, further highlighting NP as a chemical of emerging concern (CEC).

1. Introduction

Since the middle of the 20th Century, more than 140,000 new chemicals have been synthesized, of which approximately 5000 are now ubiquitous in the environment (Gruber, 2018) and able to act as endocrine disruptors (EDs). These include detergents, plasticizers, pharmaceuticals, pesticides, and other consumer products. These untested and unregulated chemicals have had unforeseen impacts on the ecosystem (Wang and Zhou, 2013) and human health (Franken et al., 2017). As they possess

chemical structures similar to natural hormones (Diamanti-Kandarakis et al., 2009), EDs are able to bind and activate many receptors, including nuclear hormone receptors (Li et al., 2015, Zhang et al., 2017a), and disrupt the endocrine system (Baker and Hardiman, 2014). Since US laws do not obligate the chemical-manufacturing corporations to test new chemicals prior to releasing them in consumer products, the burden is on the scientific community to assess the environmental and health impacts of these chemicals (Murnyak et al., 2011).

A specific subset of EDCs, the xenoestrogens (XEs), are able to mimic 17β-estradiol (E2), the natural female estrogen (Paterni et al., 2016). Nonylphenol ethoxylates (NPE) are surfactants used globally in household products, including detergents, cosmetics, and PVC pipes. They are transformed in the environment by microorganisms to form more potent compounds, such as nonylphenol (NP) a well-known ED. NP is ubiquitous in the environment and detected at levels up to 167 nM in US rivers (Fernandez et al., 2007, Sharma et al., 2009). NP is persistent in marine habitats, moderately bioaccumulative, and extremely toxic to aquatic organisms (Baker et al., 2009, Lussier et al., 2000, Staples et al., 2004, Vazquez-Duhalt et al., 2005). As NP is lipophilic, it can accumulate within the adipose tissue of animals and linger in the food chain (Noorimotlagh et al., 2016).

NP's chemical structure is the basis for both its toxicity and ability to disrupt normal functioning of the endocrine system. Its molecular structure resembles estradiol (E2), allowing it to act as an agonist for estrogen receptors and disruption of the endocrine system in higher organisms (Jobling et al., 1996, Petit et al., 1999, Tollefsen et al., 2002, White et al., 1994). Studies with model organisms have shown that NP exposure cause the synthesis of vitellogenin (VTG) in the livers of the male and immature female rainbow trout (Jobling et al., 1996, Lech et al., 1996). Additionally, environmentally relevant levels of NP have been shown to decrease semen quantity (0.6 nM) and the percentage of eggs surviving to the eyed stage and to the yolk sac larvae (1.3 nM) in rainbow trout (Lahnsteiner et al., 2005). Prolonged exposure to NP is associated with chronic kidney disease (Yen et al., 2012) and with various liver related complications, including a build-up of lipid droplets (Bernabo et al., 2014, Chen et al., 2016, Yu et al., 2016, Zhang et al., 2017b). Several studies have examined the effects of NP on certain cell types (including prostate, fibroblast, and neural cells) and have highlighted that NP reduced cell viability, induced apoptosis, affected neurogenesis and stimulated cell proliferation and adipocyte formation (Gan et al., 2015, Kudo et al., 2004, Masuno et al., 2003)

In the past decade, major advances in our understanding of genomics have occurred with concurrent advances in the development of newer and refined technologies. Improvements in the sensitivity and precision of DNA microarrays coupled with the emergence of massively parallel sequencing techniques have redefined how genomic analyses are performed (Baker and Hardiman, 2014). Transitioning from current risk assessment practices to approaches more adequate for big data collection and integration requires a paradigm shift in implementation. The Adverse Outcome Pathway (AOP) is a relatively new concept that has been rapidly gaining acceptance worldwide because it provides a framework that organizes mechanistic and/or predictive relationships between initial chemical–biological interactions, pathways and networks, and adverse phenotypic outcomes (Garcia-Reyero, 2015, Villeneuve et al., 2012, Villeneuve et al., 2014). The AOP is a linear pathway composed of a Molecular Initiating Event (MIE), Key Events (KE), and an Adverse Outcome (AO) causally linked together (Supplemental Fig. S1).

We and others have previously shown that the variation of the gene expression patterns observed in the liver of NP and E2 exposed fish are quite similar, confirming that NP recapitulates the effects of E2 (Cakmak et al., 2006, Ruggeri et al., 2008). These studies also implied that NP is able to act via alternative mechanisms to that of E2, modulating the expression of the same genes but in a different

manner. The goal of this study was firstly to provide an updated analyses of the effect of NP on the hepatic transcriptome. Secondly, we wanted to contrast NP and E2, and to characterize NP's non-estrogenic signature on the hepatic transcriptome. In order to achieve these goals, adult male zebrafish were exposed to 100 nM of NP and E2 for 21 days and their liver transcriptomes were analyzed using two complementary technologies: RNA sequencing (RNAseq) and a commercial microarray platform. We noted previously that these exposure levels result in male zebrafish having much higher vitellogenin protein levels, in response to 4-nonylphenol (0.46 ± 0.11 mg/ml) and 17b-estradiol (2.56 ± 0.51 mg/ml) respectively as compared to undetectable levels in control fish. Comparison of the 4-nonylphenol and 17 β-estradiol groups revealed this to be highly significant (P value < 0.0001) (Baker et al., 2014). In this experimental design, exposure to 100 nM E2 represents the positive estrogenic control used to compare and contrast NP with in order to determine NP's non-estrogenic signature. Finally, we wanted to interpret our data in the context of the AOP framework to gain insight on NP's mode of action.

2. Methods

2.1. Experimental design

The experimental design followed an approach we described previously (Ruggeri et al., 2008) (Fig. 1). Male zebrafish (Danio rerio) were maintained in aquaria at 26–29 °C, and a light-dark cycle of 14:10 h. The pH ranged from 7.0 to 7.6 throughout the duration of the experiments. Aeration and filtration were accomplished using sponge filters. Zebrafish were fed twice a day with commercial, flaked fish food (Tetra, Germany). The fish were acclimated for one week before beginning the experiments. Three tanks (80 L/tanks) with 40 animals each were prepared for the different experimental groups: two tanks containing water with 100 nM of NP (Fig. 1: Experimental group), two tanks containing water with 100 nM of E2 as a positive estrogenic control (Positive control), and two tanks containing water without NP or E2 (Negative control). Concentrated stock solutions of E2 and 4-NP were prepared in ethanol and then diluted in water to make up working solutions from which the chemicals were added to the tank water at a final concentration of 100 nM. To minimize any effects from ETOH the negative control group received an equivalent amount of the carrier solvent which was present at percentage levels of <10−8%. Exposure lasted 21 days. The selected NP concentration (100 nM) is in the high end of the range of NP levels detected in US rivers and a 21-day exposure to 100 nM NP is intended to mimic a chronic exposure to this chemical (Fernandez et al., 2007, Sharma et al., 2009). The goal of this study was to tease out non-estrogenic genomic signatures associated with NP. For this reason we selected 100 nM E2 as a positive and potent estrogenic control in order to effectively compare and contrast the 2 compounds. E2 exhibits 10 times more potent estrogenic activity than NP (Cakmak et al., 2006, Jobling et al., 1996).

The nominal exposures utilized a continuous flow-through system. Following a 21 day exposure, the zebrafish were sacrificed and the livers were dissected out. Liver tissues were immediately frozen in liquid nitrogen and stored at −70 °C. All procedures were performed in accordance with The University of California San Diego, IACUC guidelines. All the animals were treated humanely and with regard for alleviation of suffering.

2.2. RNA Extraction and microarray and sequencing design

Isolation of total liver RNA from zebrafish liver samples was performed using TRIzol reagent (Invitrogen), and the extracted RNA were further purified using the RNeasy Mini kit (Qiagen, Valencia, California). All RNA were treated with DNase and nucleic acid concentrations were determined by absorbance readings (OD) at 260 nm using an ND-100 (Nanodrop, Wilmington, DE). RNA integrity was assessed using 6000 Nano LabChip assay from Agilent (Palo Alto, CA), with a RNA integrity number (RIN) of >8 being required for downstream genomic analyses.

2.3. Microarray analysis

For the array experiments, liver RNA from 12 fish were pooled into 2 pools (6 fish per pool) for the E2, NP and control groups respectively (Fig. 1). Of the total RNA, 100 ng were converted into fluorescently labeled Cy3 cRNA using the Low RNA Input Fluorescent Linear Amplification Kit (Agilent). Unincorporated nucleotides of fluorescent targets were removed using RNeasy (Qiagen). Absorbance (OD) at 260 nm was used to quantify cRNA concentrations, and absorbance at 550 nm was used to measure Cy3 dye incorporation. Microarray hybridization was only carried out with cRNA that had an incorporation efficiency of 9 pmol/µg or greater.

We utilized the Agilent Danio rerio Oligo Microarray 4x44K G2519F (015064), array design A-MEXP-1396 (Santa Clara, CA). Hybridization was carried out in accordance with single color Agilent hybridization protocols, as described previously (Baker et al., 2009). From each pooled sample, 1 µg of fragmented cRNA were hybridized to the array. Array data were collected using an Agilent Microarray Scanner and Feature Extraction Software (v10.5), and deposited in the ArrayExpress Database, accession number E-TABM-547 (European Bioinformatics Institute, 2013).

Though Agilent's Feature Extraction Software (v10.5) provided high quality expression reports, the data was normalized to remove background noise and other subtle biases caused by array manufacturing and hybridization conditions. Statistical analysis of the microarray experiment involved two steps: normalization and sorting of genes according to interest. All samples were normalized simultaneously using the multiple-loess technique (Sasik et al., 2004).

The data was sorted using the interest statistic, which reflects the understanding that the gene with a greater absolute fold change is potentially more interesting, described in greater detail in (Baker et al., 2009, Baker et al., 2013, Ogawa et al., 2004). The design of the interest statistic was based on ideas borrowed from the software package Focus (Cole et al., 2003).

Specifically we carried out feature level analysis of the top ranked differentially regulated probes on the array. The fold changes were determined from log2 ratios between the probe signal of each of the conditions (control, E2 and NP). Biological replicate samples were run as outlined in Fig. 1. Additionally the probes were replicated twice on the array. The log2 ratio value was calculated for each probe as the median of the 4 replicate log2 intensity ratios. The unique probes (collapsed to gene level) were subsequently sorted by their importance in descending order of the sum-squared statistic (i.e., sum of squares of log2 ratios across all fish) as described previously (Baker et al., 2009, Baker et al., 2013, Huff et al., 2018, Ogawa et al., 2004). The rationale behind this approach was that it provided a measure of change in expression values for any one or all exposed fish pools. In this manner the sum-squared statistic measured the amount of variance across any and all exposure conditions, i.e. a transcript with altered expression in one fish pool, from exposure A would be selected along with another transcript with altered expression in a separate fish pool, from exposure B. The top ranked 3000 genes were selected and used for systems level analyses. We used Ensembl BioMart to update the array annotation to GRCz10 and Ensembl homology to append a human gene

ID (where available) to a given zebrafish gene ID, in order to permit systems analysis using the richer content available for human compared to zebrafish (Baker and Hardiman, 2014).

2.4. RNA sequencing (RNAseq)

For the RNAseq experiments, liver RNA from 12 fish were pooled into 2 pools (6 fish per pool) for the E2, NP and control groups respectively (Fig. 1). To prepare RNAseq libraries using the TruSeq RNA Sample Prep Kit (Illumina, San Diego, CA), 100–200 ng of total RNA (from pooled samples in Section 2.2, (Fig. 1)) was used following the protocol described by the manufacturer. High throughput sequencing (HTS) was performed using an Illumina GAIIX with each sample sequenced to a minimum depth of ~5 million reads. A single end 50 cycle sequencing strategy was employed. Data were subjected to Illumina quality control (QC) procedures (>80% of the data yielded a Phred score of 30). RNAseq data has been submitted to the NCBI Gene Expression Omnibus, accession number GSE100369.

Secondary analysis was carried out on an OnRamp Bioinformatics Genomics Research Platform (OnRamp Bioinformatics, San Diego, CA) (Davis-Turak et al., 2017). OnRamp's advanced Genomics Analysis Engine utilized an automated RNAseq workflow to process the data including (1) data validation and quality control, (2) read alignment to the zebrafish genome (GRCz10) using TopHat2 (Kim et al., 2013), which revealed on average >80% mapping to unique genomic locations, (3) generation of gene-level count data with HTSeq (Anders et al., 2015), and (4) differential expression analysis with DEseq2 (Love et al., 2014a), which enabled the inference of differential signals with robust statistical power. Transcript count data from DESeq2 analysis of the samples were sorted according to their q-value, which is the smallest false discovery rate (FDR) at which a transcript is called significant. FDR is the expected fraction of false positive tests among significant tests and was calculated using the Benjamini-Hochberg multiple testing adjustment procedure (Davis-Turak et al., 2017, Hardiman et al., 2016, Love et al., 2014b, Trapnell et al., 2012).

2.5. Systems level analysis

For the microarray data, systems analysis was accomplished using the Top 3000 genes ranked using the interest statistic described above (Baker et al., 2009, Baker et al., 2013, Huff et al., 2018). For RNAseq data, we chose DE genes with an adjusted p-value of less than 0.4 as described (Hardiman et al., 2016, Huff et al., 2018, Irish et al., 2016). We approached our analyses from two different perspectives; firstly analysis of the datasets as zebrafish genes, and secondly analysis of their projected human homologs. We utilized the Gene Ontology Enrichment Analysis and Visualization Tool (GOrilla) to detect GO terms enriched by NP and E2 exposures (Eden et al., 2007, Eden et al., 2009). For zebrafish centric analyses. GOrilla offers two modes of analysis. This simple approach allows the user to identify terms that are most significantly enriched. The more informative approach presents the enrichment results in the context of a directed acyclic graph (DAG). We examined the DE genes using both approaches, and focused on terms enriched in the GO Biological Process category as this proved most informative. Further analyses were performed with REViGO, a summarization tool which combines redundant GO terms into a single, representative term (Supek et al., 2011).

Due to the richer annotation of the human genome relative to zebrafish and the greater depth of GO terms available for human, we carried out enrichment analysis, on the projected human homologs of

the DE zebrafish genes, using ToppFun (Chen et al., 2007, Chen et al., 2009a, Chen et al., 2009b). To ensure that only the most relevant terms of interest were considered, we filtered all results using the Bonferroni adjusted q-value with the correction.

3. Results

3.1. Microarray analysis of molecular changes in NP exposed livers revealed altered fatty acid metabolism and insulin pathway regulation

To examine the effect of NP on the adult liver transcriptome, we carried out a microarray experiment where we assessed the effects of a 21 day exposure to 100 nM of either NP or E2 relative to untreated fish. Of the top 3000 significantly ranked differentially expressed (DE) transcripts from both exposures, 1,425 were shared amongst the NP and E2 exposures (Fig. 2A). Additionally, E2 and NP exposures altered the expression of 1022 and 1079 unique mRNAs, respectively. Both treatments resulted in an estrogenic response, and induction of the following transcripts; E2, vtg5 + 2.48; vtg2 + 1.42, vtg1 + 1.39, vtg3 + 1.35, esr1 + 3.94 and esr2b + 3.9 and NP, vtg1 + 1.2, vtg3 + 2.13, vtg2 + 1.79, vtg5 + 1.57, vtg1 + 1.10, esr2b + 5.24 and esr1 + 1.2. Next we analyzed the top 3000 DE genes using the Gene Ontology (GO) enrichment analysis and visualization tool (GOrilla) (Eden et al., 2009). Exposure to E2 (Table 1, E2 Total) was associated with enrichment in metabolic processes; the organic acid metabolic process is the most significantly enriched term (q-value = 5.77E−02), along with enrichments in the carboxylic acid metabolic (q = 5.84E−02), carbohydrate metabolic (q = 1.89E−01), and alpha-amino acid metabolic processes (q = 3.73E−01) (Fig. 2C). Exposure to NP was associated with enrichments in metabolic fatty acid-related pathways, including significant enrichment of terms relating to the long-chain fatty acid metabolic (q = 4.42E−01) and biosynthetic (q = 1.99E−01) processes (Table 1 and Fig. 2B). In addition, we found enrichment in terms related to the regulation of cell proliferation (q = 1.95E−01), the insulin receptor signaling pathway (q = 2.49E−01), and antigen processing and presentation (q = 3.76E−01). We then performed enrichment analysis using gene signatures unique to each particular exposure and observed that the gene signatures unique to exposure to E2 (Table 1, E2 Unique) were associated with enrichment of the terms ruffle organization (q = 3.80E−1) and synapse assembly (q = 5.69E−01), while NP's unique signatures were associated with enrichment in endosome organization (q = 2.57E−01) (Table 1, NP Unique).

3.2. Functional enrichment analysis of array data using human gene identifiers

We mapped zebrafish genes of interest to their human orthologs using Ensembl to take advantage of the improved annotation for the human genome, as illustrated in Fig. 3 (Baker and Hardiman, 2014). Using the predicted human orthologs, we performed ToppFun enrichment analysis. Data is summarized in Table 2. Expanded lists are found in Supplemental Tables 1–4. The results of the enrichment analyses indicate a focus on metabolism with both exposures, including organic acid metabolism in NP (q = 6.26E−25) and carboxylic acid metabolism in E2 (q = 3.23E−31) (Supplemental Tables 1 and 2). Additionally, a significant connection between lipid metabolism (q = 2.21E−10) and exposure to NP was observed. E2 exposure was associated with changes in the cell's response to hormones (q = 1.87E−13). Furthermore, we noticed a trend relating to enrichment of terms relevant to the cell cycle. With both the NP and E2 exposures, cell cycle pathways were significantly enriched in the analyses based on the human orthologs of the DE genes, as opposed to the analyses based on the zebrafish DE genes (E2, mitotic cell cycle, q = 2.29E−10, NP, cell cycle, q = 4.63E−18). Analysis of

the DE genes unique to NP exposure enriched terms related to fatty acid metabolism (NP Unique, q = 1.10E−2), in addition to response to abiotic stimulus (NP Unique, q = 9.49E−07) and DNA damage (NP Unique, q = 7.89E−03) stimuli (Supplemental Table 3). The unique genes associated with exposure to E2 enriched terms related to the processing of RNA (E2 Unique, q = 3.26E−07) and non-coding RNA (E2 Unique, q = 3.85E−03), as well as translation (E2 Unique, q = 1.95E−03) (Supplemental Table 4).

In terms of co-expression, the DE genes of interest overlapped with gene signatures up-regulated in hepatoblastoma (NP Total, q = 3.62E−25) and down-regulated in response to hypoxia and the overexpression of hypoxia inducing factor 1 alpha (HIF1A) (NP Total, q = 1.81E−23) (Supplemental Table 5). DE genes in response to E2 shared overlapped with gene signatures down-regulated in fetal liver with knock-out of Krueppel-like factor 1 KLF1 (E2 Total, q = 6.97E−37) and up-regulated via activation of the mammalian target of rapamycin complex 1 (mTORC1) (E2 Total, q = 1.40E−28) (Supplemental Table 6). Analysis of DE genes unique to NP exposure identified an overlap with genes up-regulated in human liver at an advanced developmental stage (NP Unique, q = 3.60E−12) and genes with promoters bound by the MYC proto-oncogene (NP Unique, q = 1.05E−06) (Supplemental Table 7). DE genes unique to E2 were also identified as being up-regulated by knock-out of Myb-related protein B (BMYB) in zebrafish (E2 Unique, q = 5.09E−11), and genes up-regulated in response to the Ras inhibitor Salirasib in cancer cells with constant HRAS activity (E2 Unique, q = 1.19E−08) (Supplemental Table 8).

3.3. Functional enrichment analysis of RNAseq data using zebrafish gene identifiers

We carried out a high-throughput RNA sequencing analysis to further investigate the effects of these compounds on the zebrafish hepatic transcriptome. This was an independent experimental exposure of adult zebrafish to either E2 or NP. For this analysis, we selected genes with a q-value of less than or equal to 0.4. This cutoff was based on our previous work with RNAseq data sets where we noted that a more liberal FDR cutoff using biological replicates generated a larger gene list for systems analyses (Hardiman et al., 2016, Huff et al., 2018, Irish et al., 2016). In total, exposure to E2 and NP led to the differential expression of 883 and 454 genes respectively. Of these, 154 were shared between the NP and E2 exposures, as seen in Fig. 4. Exposure to E2 and NP altered the expression of 729 and 300 unique genes, respectively.

Using these DE genes lists, we performed GOrilla analysis. The initial analysis examined terms that are most significantly enriched in both the E2 and NP exposed fish. E2 exposure enriched the following GO biological process terms; response to estradiol (1.15E−12); response to estrogen (6.72E−12) and cellular response to estrogen stimulus (6.72E−12) indicating an estrogenic response (Supplemental Table 9). NP exposure enriched the following GO biological process terms; response to estradiol (7.91E−07), response to estrogen (7.91E−07) and cellular response to estrogen stimulus (7.91E−07) also indicating an estrogenic response (Supplemental Table 10). Additional analyses explored enrichment results in the context of a directed acyclic graph (DAG). Data is summarized in Table 3. Expanded lists are found in Supplemental Table 11. Of the enrichments we found in the exposure to E2 and NP, the most notable pathways observed in both exposures included lipid transport (E2 = 2.92E−03, NP = 2.74E−02) and the carboxylic acid metabolic process (E2 = 6.65E−10, NP = 2.83E−01) (Table 3 and Fig. 4B and C). Exposure to E2 was associated with over 100 enriched terms, including enriched metabolic processes and the cell's response to external stimuli, including estrogen (1.23E−07) (Table 3 and Fig. 4C). Exposure to NP enriched terms related to the immune response, including regulation of the immune system (2.96E−01) and responses to bacterium

(3.07E−01) and defense response to other organisms (3.01E−01) (Table 3 and Fig. 4B). Using only the genes associated with exposure to NP, we found significant enrichments related to the cell's response to oxygen levels (1.23E−02), particularly response to decreased oxygen levels (1.64E−02), in addition to terms associated with metabolic processes and response to stimuli (Table 3 and Fig. 4D). For the DE genes unique to E2, we noted enrichment in metabolic processes – with an emphasis in small molecule metabolic processes (1.69E−15) and carboxylic acid (9.59E−11) metabolic processes in addition to lipid metabolism (1.32E−03) (Table 3 and Fig. 4E).

3.4. Functional enrichment analysis of RNAseq data using human gene identifiers

Using the predicted human homologs of the DE genes, we performed functional enrichment analysis using ToppFun. Table 4 contains significant GO enrichment terms and co-expression signatures for the NP and E2 exposures, as well as for genes unique to NP or E2 exposure. The results of our analysis show that exposure to both E2 and NP enriches metabolic pathways (Table 4 and Supplemental Tables 12–14). E2, in particular, enriches organic acid metabolism terms, including carboxylic acid (E2 Total, q = 4.05E−32), with some terms related to lipid metabolism (E2, q = 9.87E−17) (Table 4 and Supplemental Table S13). With NP, the enriched terms include immune response (NP, q = 1.60E−08), inflammatory response (NP, q = 7.94E−04), response to oxygen containing compound (NP, q = 7.61E−06) and fatty acid metabolic process (NP, q = 7.2.61E−03) (Table 4 and Supplemental Table S12). Focusing on DE genes unique to the NP exposure, we uncovered terms related to metabolism of reactive oxygen species (NP Unique, q = 1.67E−05) and cholesterol biosynthetic process (NP Unique, q = 8.32E−03), as well as localization of proteins to the endoplasmic reticulum (NP Unique, q = 4.43E−03) (Table 4 and Supplemental Table 14). For genes unique to exposure to E2, the focus remained on metabolic processes, including cellular lipid metabolism (E2 Unique, q = 5.85E−14), as well as lipid oxidation (E2 Unique, q = 3.52E−06) and fatty acid oxidation (E2 Unique, q = 2.50E−06) (Table 4 and Supplemental Table S15).

In terms of co-expression signatures, we observed an overlap with genes in perturbed liver cells, including changes in fatty acid metabolism and changes in the immune response (Table 4 and Supplemental Tables 16–18). Some interesting co-expression signatures we found in the E2 exposed liver included genes up-regulated in relation to the zebrafish crb ("crash and burn") loss-of-function mutation in bmyb (NP, q = 7.99E−07) (Table 4 and Supplemental Table S16). Genes down-regulated in hepatoblastoma samples compared to healthy liver cells were enriched in the E2 exposed liver (E2, q = 1.56E−22) (Table 4 and Supplemental Table S17). As with the GO analysis, exposure to NP shared gene signatures with a number of inflammatory pathways, including genes regulated by NF-κB in response to TNF (NP Total, q = 1.04E−06), and genes with differential expression in hypoxic conditions (NP, q = 4.08E−06) (Table 4, Supplemental Table S16). When considering only the DE genes associated with NP exposure, we observed enrichment in terms associated with liver regeneration in mice (NP Unique, q = 1.38E−12) and gene signatures that are up-regulated in the livers of mice with reduced cytochrome p450 oxidoreductase (POR) expression (NP Unique, q = 2.49E−06) (Table 4, Supplemental Table S18). The DE genes unique to E2 exposure had similarity to co-expression signatures with genes up-regulated in hepatocellular carcinoma (HCC) cells compared with normal liver cells (E2 Unique, q = 8.69E−19) (Table 4, Supplemental Table S19).

4. Discussion

The goal of this study was to assess the effect of NP exposure on the liver transcriptome using a systems level approach. We determined from the microarray analysis that 1) exposure to NP and E2

enriched biological processes terms with a focus on metabolism, 2) NP exposure induced changes in fatty-acid metabolic processes, antigen processing, cell cycle and apoptosis related terms, 3) the DE genes shared co-expression patterns with those of liver cells with perturbed functions, particularly those associated with adverse hepatic outcomes. From the RNAseq analysis, we identified 4) enrichment in GO terms related to proteolysis, cellular response, and organic acid metabolism for both the NP and E2 exposures, 5) NP exposure induced changes in antigen processing and presentation, response to hypoxia, immune response and metabolism.

4.1. The advantages of exploiting human annotations in a zebrafish study

At this point, the zebrafish genome is not as well characterized and annotated as the human genome is. However 70% of protein-coding human genes are related to genes found in zebrafish and 84% of genes known to be associated with human disease have a zebrafish counterpart (Howe et al., 2013). For these reasons, it is valuable to consider the human orthologs of zebrafish genes for GO analyses. As shown in Fig. 2 (Baker and Hardiman, 2014), in terms of the ratio of human to zebrafish annotations, there are >5 times more non-inferred electronic and >2 times more functional annotations for human relative to the zebrafish, based on the GO database over the past two years. Therefore by projecting zebrafish genes onto their human orthologs, a richer analysis can be achieved (Fig. 2). This strategy was not utilized in our previous analysis of 100 nM of NP (Ruggeri et al., 2008). It must be noted that there are limitations to this analytical approach. Specific fish genes, such as the vitellogenins (VTGs) for example, do not have orthologs in humans. Exposure to NP is associated with higher expression of VTGs in male zebrafish, and can be used to highlight its estrogenic properties (Ruggeri et al., 2008). As VTGs lack human orthologs, analysis with human annotations will lose this information and be unable to detect any changes associated with VTG expression. However, for a comparative analysis between a model organism and humans, these limitations are outweighed by the benefits of improved annotations and a more comprehensive systems analysis.

4.2. Exposure to NP and E2 perturbed expression of genes involved in metabolic pathways and cellular response

Our data suggest that both NP and E2 dysregulated expression of genes involved in metabolic pathways and cellular response. Since the liver is the primary site of metabolism within the body that constantly filters out environmental chemicals such as NP, we were not surprised by the enrichment of metabolic pathways (Noorimotlagh et al., 2016). Among shared NP and E2 enrichments were changes to organic acid metabolism, in particular carboxylic acid metabolism, which is consistent with the expression patterns of human liver cells in response to ethanol stress (Schmidt-Heck et al., 2017).

Co-expression data from ToppFun analysis indicated that the list of DE genes overlapped with perturbation of liver cell function. A notable example of this were genes upregulated in the zebrafish crb (“crash and burn”) mutant, representing a loss-of-function for the transcriptional regulator bmyb (Shepard et al., 2005). The gene-expression signatures associated with this bmyb mutation are related to gene signatures found in human tumors, and increased cancer susceptibility in adult zebrafish heterozygotes.

Our analyses also indicated an overlap between our genes of interest and genes with promoters bound by c-MYC, the transcription factor and proto-oncogene whose over-expression is associated with the development of cancer (Zeller et al., 2003).

MYC signaling is induced by estrogen signaling, and is mediated by interactions with the estrogen receptor (ER) and activating protein 1 (AP-1) (Wang et al., 2011). In the context of the liver, expression of c-MYC is associated with the proliferation of hepatocytes during liver regeneration (Thompson et al., 1986). Furthermore, its role in the development of liver cancers has been well-defined; deregulation of c-MYC expression, even at moderate levels, has been found in hepatocellular carcinoma (HCC) cells (Xin et al., 2017). In ovarian cancer cells, exposure to NP was associated with a significant increase in the levels of c-Myc mRNA, suggesting increased expression of the c-MYC protein (Bo Lü and Zhan, 2010, Park et al., 2011). The presence of DE genes associated with c-MYC is expected in response to E2, and their presence in response to NP confirms the estrogenic effects of the surfactant. Enrichment analysis of exposure to E2 and NP identified an overlap with genes mediated by bmyb and c-MYC, two proteins whose dysregulation is associated with cancer development.

4.3. Unique effects of NP exposure relate to immune response and fatty acid metabolism in the liver

Enrichment analysis of the microarray data for zebrafish exposed to NP indicated enrichment in pathways relating to fatty acid metabolism and antigen processing and presentation, found in both GOrilla and ToppFun analyses. Analysis of the RNAseq data confirmed these findings; we found significant enrichment across a variety of immune and inflammatory pathways, as well as changes in fatty acid metabolism. We identified overlap with genes regulated by NF-κB in response to the TNF cytokine family suggesting inflammatory processes. NP's ability to induce inflammation within the liver has been previously described in mice, and was attributed to an increased presence of the cytokines TNF-α and IL-1α (Yu et al., 2016).

Our data also suggest that exposure to NP induced changes in the metabolism of fatty acids. Exposure to NP has previously been associated with the development of lipid droplets within the liver tissue and steatosis, the abnormal retention of lipids within cells (Bernabo et al., 2014, Maradonna et al., 2015). Additionally, NP exposure has been shown to increase expression of hepatic peroxisome proliferator-activated receptors (PPAR) α and β, in turn suppressing the expression of cytochrome P450 isoforms (CYP1A1 and CYP3A4) necessary for detoxification pathways (Cocci et al., 2013). The dysregulation of these processes after exposure to NP could indicate a MIE with linkage to NAFLD as the AOP. Described as chronic steatosis within the liver, NAFLD has been linked to obesity and other metabolic disorders (Neuschwander-Tetri, 2017), and is currently the most common cause of chronic liver disease in developed countries (Dyson and Day, 2014). NAFLD can ultimately progress to cirrhosis, a permanent scarring of the liver. In summary, our data indicate that exposure to NP alters fatty acid processing and the immune response, molecular perturbations that could be associated with the development of NAFLD.

4.4. Non-estrogenic effects of NP suggest changes in cell death and response to reactive oxygen species

As we analyzed NP's effects on the liver transcriptome, we assessed the effects of E2 in parallel. This allowed us to identify the estrogenic effects of NP, as well as to determine its non-estrogenic effects

by considering genes with differential expression upon NP exposure, but not E2. GO analysis of projected human orthologs of DE expressed liver transcripts revealed that exposure to NP was associated with the cell's response to DNA damage and the negative regulation of cell death, signatures not observed in the exposure to E2. The signatures associated with metabolic processes and the cell cycle that were associated with exposure to NP (all DE genes) were retained in the analysis of DE genes associated with exposure to NP and not E2.

Closely linked to these results is an enrichment in the metabolism of reactive oxidative species (ROS). ROS are normally generated during the metabolism of oxygen, but during stress are produced at higher rates, which can overwhelm the cell's antioxidant defense system (Ray et al., 2012). This oxidative stress results in damage to nucleic acids and lipids, and has been implicated in the development of cancer (Reuter et al., 2010), inflammation (Reuter et al., 2010), and cell death (Arakha et al., 2017). Exposing zebrafish embryos to NP has previously been associated with the induction of oxidative stress, resulting in increased expression of immune response genes (Xu et al., 2013). Furthermore, it has been shown that exposure to NP in mouse TM4 Sertoli cells induced apoptosis by generating ROS and activating the ERK signaling pathway (Choi et al., 2014). Linking the association with NAFLD described above, increased generation of ROS can induce inflammation in liver cells through lipid peroxidation (Day, 2002), and induces fat accumulation by inhibiting hepatocytes from secreting very low density lipoprotein (VLDL) (Polimeni et al., 2015). Analysis of the DE gene signatures unique to exposure to NP identified changes relating to the metabolism of ROS that in combination with changes in inflammation and fatty acid metabolism support previous studies suggesting NP has a role in the development of NAFLD (Kourouma et al., 2015).

4.5. Comparison of DNA microarray analyses vs. RNA sequencing

Both the microarray and RNAseq analyses of the transcriptome identified perturbations in hepatic gene expression, and allowed us to determine if these changes in gene expression are connected to changes in health outcomes. While RNAseq is quickly becoming more cost effective, with less technical noise, a greater dynamic range to quantify gene expression, and highly reproducible results, microarrays are still used due to their ease of use and lower cost (Davis-Turak et al., 2017). In general, we found that the RNAseq and microarray experiments were consistent. The RNAseq analyses provided greater sensitivity by uncovering transcriptional perturbations that were not detected with the array platform. The GOrilla (zebrafish) enrichment analysis of RNAseq data associated with exposure to E2, for instance, identified many of the same enrichment terms we observed in the microarray analysis, including organic acid and oxoacid metabolic processes, while identifying enrichments in lipid transport and proteolysis that we did not observe in our array analysis. We saw similar results in our comparison of the ToppFun (human-ortholog projected) analyses of E2 exposure. RNAseq analysis confirmed the enrichments in organic acid and lipid metabolic process terms observed with the microarray analysis, while also identifying enrichments in triglyceride metabolic process and lipid oxidation not observed with the array analysis.

The most notable difference between the array and RNAseq analyses were observed in the NP exposures: analysis of the RNAseq indicated highly significant changes in the immune response (Table 4 and Supplemental Table S10 and S14); while we did observe enrichment in some immune pathways in our analysis of the array data (Table 2 and Supplemental Tables S1 and S5), we did not observe as deep an enrichment of these terms compared to the RNAseq analysis. Furthermore, we

observed enrichment of terms relating to the cell cycle and lipid metabolism in our analysis of the array data, terms we did not observe in our analysis of the RNAseq data.

4.6. Comparison to the earlier study by Ruggeri et al. (2008)

This study is a follow-up to a previous experiment our group conducted, in which zebrafish were exposed to 100 nM of NP and liver tissue analyzed via a microarray analysis (Ruggeri et al., 2008). The previous experiment was completed over a decade ago, with a spotted oligonucleotiode microarray based on an early draft of the zebrafish genome and well before high-throughput sequencing became commonplace. In this manuscript, we revisited the effects of NP exposure using an updated zebrafish genome build, a robust commercial microarray platform and high throughput RNAseq. These newer and more sophisticated analyses permitted a deeper systems level analyses.

In our initial study, we determined that exposure to NP and E2 could significantly induce the expression of vitellogenin (VTG), a sex-related precursor of yolk proteins, in both female and male zebrafish. Both xenoestrogens changed the regulation of genes associated with energy metabolism, oxidative stress defense mechanisms, xenobiotic metabolism, and lipid metabolism. In the case of oxidative stress, the two treatments demonstrated different patterns of expression; exposure to E2 induced detoxification, while exposure to NP inhibited this mechanism. This earlier study identified non-estrogenic mechanisms of NP, while also showing the efficacy of the microarray approach on determining the expression patterns in a toxicology study. However, by today's standards, there are technological and bioinformatics limitations to this earlier work that we have addressed in this newer study.

Firstly, the technology employed was a printed oligonucleotide microarray based on an early draft of the zebrafish genome. This newer study utilizes a commercial Agilent microarray in addition to RNAseq. Secondly, the improved zebrafish genome annotation and coupled with Ensembl homology to map zebrafish genes to their human counterparts affords more sophisticated bioinformatics analyses. Our study provides a considerable update to the findings of this earlier paper; supporting its conclusions while also identifying new signatures of interest associated with both exposures. RNAseq analysis of zebrafish exposed to NP revealed enrichment of many of the same GO terms, including protein metabolism, lipid metabolism, and oxidative stress defense mechanisms. Unique to our new study were enrichments relating to the cell's response to hypoxic conditions, antigen processing and presentation, and the immune system process. Furthermore, by 'humanizing' the DE genes of interest, we discovered that NP has a significant effect on the regulation of the immune system, including response to inflammation. Co-expression analysis of our exposures identified an overlap with genes related to knock-out of BMYB in zebrafish, genes regulated by NF-kB in response to TNF, and genes up-regulated in mice with reduced expression of cytochrome P450 oxidoreductase (POR). In terms of exposure to E2, our study identified many of the same GO enrichments, mainly focusing on primary metabolism, including lipid metabolism, protein metabolism, and organic metabolism. Unique to our findings were changes in response to external stimuli, including estrogen; these were enriched in the previous study's NP exposure group, though their presence in the E2 exposure group was not as significant as it was in our experiment. This may be due to updates and improvement in the Gene Ontology database over the past decade.

Co-expression analysis identified an overlap between our perturbed genes and those up-regulated by mTORC1, up-regulated by knockout of BMYB, and down-regulated in liver tumor cells. In summary, this newer analysis supports the earlier findings, as well as identifying new processes with

differential expression in response to exposure to both NP and E2. These include the immune response, response to hypoxic conditions, and the presence of genes shared with perturbed liver cells.

5. Conclusion

The focus of this study was to examine the effects of NP on the liver using modern bioinformatics approaches to analyze microarray and sequencing data obtained from in vivo exposures of adult zebrafish. Our findings with regard to DE genes that represent non-estrogenic signatures are summarized in Fig. 5 in the context of the AOP framework. Our results indicate that exposure to NP leads to the enrichment of genes related to fatty acid metabolism, immune response, and other processes. Compared to previous studies with NP exposure, we have identified novel changes in the immune response, response to hypoxia, and a potential association with liver disorders as a result of exposure to NP, highlighting the advantages of recent advances in deep transcriptomic profiling.

Acknowledgements

We thank Drs. Bethany J. Wolf, Jeremy L. Barth and Russell A. Norris for useful discussions. We thank Drs. Roman Sasik, and Jeremy Davis-Turak for discussions on the bioinformatics analyses. GH acknowledges funding from SC EPSCoR and start-up funding from College of Medicine at the Medical University of South Carolina.

Anders, S., Pyl, P.T., Huber, W., 2015. HTSeq–a Python framework to work with highthroughput sequencing data. Bioinformatics 31, 166–169.

References

Arakha, M., Roy, J., Nayak, P.S., Mallick, B., Jha, S., 2017. Zinc oxide nanoparticle energy band gap reduction triggers the oxidative stress resulting into autophagy-mediated apoptotic cell death. Free Radic. Biol. Med.

Baker, M.E., Hardiman, G., 2014. Transcriptional analysis of endocrine disruption using zebrafish and massively parallel sequencing. J. Mol. Endocrinol. 52, R241–256.

Baker, M.E., Ruggeri, B., Sprague, L.J., Eckhardt-Ludka, C., Lapira, J., Wick, I., Soverchia, L., Ubaldi, M., Polzonetti-Magni, A.M., Vidal-Dorsch, D., Bay, S., Gully, J.R., Reyes, J.A., Kelley, K.M., Schlenk, D., Breen, E.C., Sasik, R., Hardiman, G., 2009. Analysis of endocrine disruption in Southern California coastal fish using an aquatic multispecies microarray. Environ. Health Perspect. 117, 223–230.

Baker, M.E., Vidal-Dorsch, D.E., Ribecco, C., Sprague, L.J., Angert, M., Lekmine, N.,

Ludka, C., Martella, A., Ricciardelli, E., Bay, S.M., Gully, J.R., Kelley, K.M., Schlenk, D., Carnevali, O., Sasik, R., Hardiman, G., 2013. Molecular analysis of endocrine disruption in hornyhead turbot at wastewater outfalls in southern california using a second generation multi-species microarray. PLoS One 8, e75553.

Baker, M.E., Sprague, L.J., Ribecco, C., Ruggeri, B., Lekmine, N., Ludka, C., Wick, I., Soverchia, L., Ubaldi, M., Šášik, R., Schlenk, D., Kelley, K.M., Reyes, J.A., Hardiman, G., 2014. Application of a targeted endocrine q-PCR panel to monitor the effects of pollution in southern California flatfish. Endocrine Disruptors 2, e969598.

Bernabo, I., Biasone, P., Macirella, R., Tripepi, S., Brunelli, E., 2014. Liver histology and ultrastructure of the Italian newt (Lissotriton italicus): normal structure and modifications after acute exposure to nonylphenol ethoxylates. Exp. Toxicol. Pathol. 66, 455–468.

Bo Lü, L., Zhan, P., 2010. Effects of nonylphenol on ovarian cancer cells PEO4 c-myc and p53 mRNA and protein expression. Toxicol. Environ. Chem. 92, 1303–1308.

Cakmak, G., Togan, I., Severcan, F., 2006. 17β-Estradiol induced compositional, structural and functional changes in rainbow trout liver, revealed by FT-IR spectroscopy: a comparative study with nonylphenol. Aquat. Toxicol. 77, 53–63.

Chen, J., Xu, H., Aronow, B.J., Jegga, A.G., 2007. Improved human disease candidate gene prioritization using mouse phenotype. BMC Bioinf. 8, 392.

Chen, J., Aronow, B.J., Jegga, A.G., 2009a. Disease candidate gene identification and prioritization using protein interaction networks. BMC Bioinf. 10, 73.

Chen, J., Bardes, E.E., Aronow, B.J., Jegga, A.G., 2009b. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. Nucl. Acids Res. 37, W305–311.

Chen, H., Zhang, W., Rui, B.B., Yang, S.M., Xu, W.P., Wei, W., 2016. Di(2-ethylhexyl) phthalate exacerbates non-alcoholic fatty liver in rats and its potential mechanisms. Environ. Toxicol. Pharmacol. 42, 38–44.

Choi, M.S., Park, H.J., Oh, J.H., Lee, E.H., Park, S.M., Yoon, S., 2014. Nonylphenol-induced apoptotic cell death in mouse TM4 Sertoli cells via the generation of reactive oxygen species and activation of the ERK signaling pathway. J. Appl. Toxicol. 34, 628–636.

Cocci, P., Mosconi, G., Palermo, F.A., 2013. Effects of 4-nonylphenol on hepatic gene expression of peroxisome proliferator-activated receptors and cytochrome P450 isoforms (CYP1A1 and CYP3A4) in juvenile sole (Solea solea). Chemosphere 93, 1176–1181.

Cole, S.W., Galic, Z., Zack, J.A., 2003. Controlling false-negative errors in microarray differential expression analysis: a PRIM approach. Bioinformatics 19, 1808–1816.

Davis-Turak, J., Courtney, S.M., Hazard, E.S., Glen Jr., W.B., da Silveira, W.A., Wesselman, T., Harbin, L.P., Wolf, B.J., Chung, D., Hardiman, G., 2017. Genomics pipelines and data integration: challenges and opportunities in the research setting. Expert Rev. Mol. Diagnostics 17, 225–237.

Day, C.P., 2002. Pathogenesis of steatohepatitis. Best Pract. Res. Clin. Gastroenterol. 16, 663–678.

Diamanti-Kandarakis, E., Bourguignon, J.P., Giudice, L.C., Hauser, R., Prins, G.S., Soto, A.M., Zoeller, R.T., Gore, A.C., 2009. Endocrine-disrupting chemicals: an Endocrine Society scientific statement. Endocr. Rev. 30, 293–342.

Dyson, J., Day, C., 2014. Treatment of non-alcoholic fatty liver disease. Dig Dis. 32, 597–604.

Eden, E., Lipson, D., Yogev, S., Yakhini, Z., 2007. Discovering motifs in ranked lists of DNA sequences. PLoS Comput. Biol. 3, e39.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., Yakhini, Z., 2009. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinf. 10, 48.

Fernandez, M.P., Ikonomou, M.G., Buchanan, I., 2007. An assessment of estrogenic organic contaminants in Canadian wastewaters. Sci. Total Environ. 373, 250–269.

Franken, C., Lambrechts, N., Govarts, E., Koppen, G., Den Hond, E., Ooms, D., Voorspoels, S., Bruckers, L., Loots, I., Nelen, V., Sioen, I., Nawrot, T.S., Baeyens, W., Van Larebeke, N., Schoeters, G., 2017. Phthalate-induced oxidative stress and association with asthma-related airway inflammation in adolescents. Int. J. Hyg. Environ. Health.

Gan, W., Zhou, M., Xiang, Z., Han, X., Li, D., 2015. Combined effects of nonylphenol and bisphenol A on the human prostate epithelial cell line RWPE-1. Int. J. Environ. Res. Public Health 12, 4141.

Garcia-Reyero, N., 2015. Are adverse outcome pathways here to stay? Environ. Sci. Technol. 49, 3–9.

Gruber, K., 2018. Cleaning up pollutants to protect future health. Nature 555, S20.

Hardiman, G., Savage, S.J., Hazard, E.S., Wilson, R.C., Courtney, S.M., Smith, M.T., Hollis, B.W., Halbert, C.H., Gattoni-Celli, S., 2016. Systems analysis of the prostate transcriptome in African-American men compared with European-American men. Pharmacogenomics 17, 1129–1143.

Howe, K., Clark, M.D., Torroja, C.F., Torrance, J., Berthelot, C., Muffato, M., Collins, J.E., Humphray, S., McLaren, K., Matthews, L., McLaren, S., Sealy, I., Caccamo, M., Churcher, C., Scott, C., Barrett, J.C., Koch, R., Rauch, G.J., White, S., Chow, W., Kilian, B., Quintais, L.T., Guerra-Assuncao, J.A., Zhou, Y., Gu, Y., Yen, J., Vogel, J.H., Eyre, T., Redmond, S., Banerjee, R., Chi, J., Fu, B., Langley, E., Maguire, S.F., Laird, G.K., Lloyd, D., Kenyon, E., Donaldson, S., Sehra, H., Almeida-King, J., Loveland, J., Trevanion, S., Jones, M., Quail, M., Willey, D., Hunt, A., Burton, J., Sims, S., McLay, K., Plumb, B., Davis, J., Clee, C., Oliver, K., Clark, R., Riddle, C., Elliot, D., Threadgold, G., Harden, G., Ware, D., Begum, S., Mortimore, B., Kerry, G., Heath, P., Phillimore, B., Tracey, A., Corby, N., Dunn, M., Johnson, C., Wood, J., Clark, S., Pelan, S., Griffiths, G., Smith, M., Glithero, R., Howden, P., Barker, N., Lloyd, C., Stevens, C., Harley, J., Holt, K., Panagiotidis, G., Lovell, J., Beasley, H., Henderson, C., Gordon, D., Auger, K., Wright, D., Collins, J., Raisen, C., Dyer, L., Leung, K., Robertson, L., Ambridge, K., Leongamornlert, D., McGuire, S., Gilderthorp, R., Griffiths, C., Manthravadi, D., Nichol, S., Barker, G., Whitehead, S., Kay, M., Brown, J., Murnane, C., Gray, E., Humphries, M., Sycamore, N., Barker, D., Saunders, D., Wallis, J., Babbage, A., Hammond, S., Mashreghi-Mohammadi, M., Barr, L., Martin, S., Wray, P., Ellington, A., Matthews, N., Ellwood, M., Woodmansey, R., Clark, G., Cooper, J., Tromans, A., Grafham, D., Skuce, C., Pandian, R., Andrews, R., Harrison, E., Kimberley, A., Garnett, J., Fosker, N., Hall, R., Garner, P., Kelly, D., Bird, C., Palmer, S., Gehring, I., Berger, A., Dooley, C.M., Ersan-Urun, Z., Eser, C., Geiger, H., Geisler, M., Karotki, L., Kirn, A., Konantz, J., Konantz, M., Oberlander, M., RudolphGeiger, S., Teucke, M., Lanz, C., Raddatz, G., Osoegawa, K., Zhu, B., Rapp, A., Widaa, S., Langford, C., Yang, F., Schuster, S.C., Carter, N.P., Harrow, J., Ning, Z., Herrero, J., Searle, S.M., Enright, A., Geisler, R., Plasterk, R.H., Lee, C., Westerfield, M., de Jong, P.J., Zon, L.I., Postlethwait, J.H., Nusslein-Volhard, C., Hubbard, T.J., Roest Crollius, H., Rogers, J., Stemple, D.L., 2013. The zebrafish reference genome sequence and its relationship to the human genome. Nature 496, 498–503.

Huff, M., da Silveira, W.A., Carnevali, O., Renaud, L., Hardiman, G., 2018. Systems analysis of the liver transcriptome in adult male zebrafish exposed to the plasticizer (2-Ethylhexyl) phthalate (DEHP). Sci. Rep. 8, 2118.

641 Irish, J.C., Mills, J.N., Turner-Ivey, B., Wilson, R.C., Guest, S.T., Rutkovsky, A., Dombkowski, A.,
642 Kappler, C.S., Hardiman, G., Ethier, S.P., 2016. Amplification of WHSC1L1 regulates expression and
643 estrogen-independent activation of ERalpha in SUM-44 breast cancer cells and is associated with
644 ERalpha over-expression in breast cancer. Mol. Oncol. 10, 850–865.

645 Jobling, S., Sumpter, J.P., Sheahan, D., Osborne, J.A., Matthiessen, P., 1996. Inhibition of testicular
646 growth in rainbow trout (Oncorhynchus mykiss) exposed to estrogenic alkylphenolic chemicals.
647 Environ. Toxicol. Chem. 15, 194–202.

648 Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., Salzberg, S.L., 2013. TopHat2: accurate
649 alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol.
650 14, R36.

651 Kourouma, A., Keita, H., Duan, P., Quan, C., Bilivogui, K.K., Qi, S., Christiane, N.A., Osamuyimen, A.,
652 Yang, K., 2015. Effects of 4-nonylphenol on oxidant/antioxidant balance system inducing hepatic
653 steatosis in male rat. Toxicol. Rep. 2, 1423–1433.

654 Kudo, C., Wada, K., Masuda, T., Yonemura, T., Shibuya, A., Fujimoto, Y., Nakajima, A., Niwa, H.,
655 Kamisaki, Y., 2004. Nonylphenol induces the death of neural stem cells due to activation of the
656 caspase cascade and regulation of the cell cycle. J. Neurochem. 88, 1416–1423.

657 Lahnsteiner, F., Berger, B., Grubinger, F., Weismann, T., 2005. The effect of 4-nonylphenol on semen
658 quality, viability of gametes, fertilization success, and embryo and larvae survival in rainbow trout
659 (Oncorhynchus mykiss). Aquat. Toxicol. 71, 297–306.

660 Lech, J.J., Lewis, S.K., Ren, L., 1996. In Vivo estrogenic activity of nonylphenol in rainbow trout.
661 Fundam. Appl. Toxicol. 30, 229–232.

662 Li, L., Wang, Q., Zhang, Y., Niu, Y., Yao, X., Liu, H., 2015. The molecular mechanism of bisphenol A
663 (BPA) as an endocrine disruptor by interacting with nuclear receptors: insights from molecular
664 dynamics (MD) simulations. PLoS One 10, e0120330.

665 Love, M.I., Huber, W., Anders, S., 2014b. Moderated estimation of fold change and dispersion for
666 RNA-Seq data with DESeq2.

667 Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for
668 RNA-seq data with DESeq2. Genome Biol. 15, 550.

669 Lussier, S.M., Champlin, D., LiVolsi, J., Poucher, S., Pruell, R.J., 2000. Acute toxicity of para-
670 nonylphenol to saltwater animals. Environ. Toxicol. Chem. 19, 617–621.

671 Maradonna, F., Nozzi, V., Santangeli, S., Traversi, I., Gallo, P., Fattore, E., Mita, D.G., Mandich, A.,
672 Carnevali, O., 2015. Xenobiotic-contaminated diets affect hepatic lipid metabolism: implications for
673 liver steatosis in Sparus aurata juveniles. Aquat. Toxicol. 167, 257–264.

674 Masuno, H., Okamoto, S., Iwanami, J., Honda, K., Shiosaka, T., Kidani, T., Sakayama, K., Yamamoto,
675 H., 2003. Effect of 4-nonylphenol on cell proliferation and adipocyte formation in cultures of fully
676 differentiated 3T3-L1 cells. Toxicol. Sci. 75, 314–320.

677 Murnyak, G., Vandenberg, J., Yaroschak, P.J., Williams, L., Prabhakaran, K., Hinz, J., 2011. Emerging
678 contaminants: presentations at the 2009 Toxicology and Risk Assessment Conference. Toxicol. Appl.
679 Pharmacol. 254, 167–169.

680 Neuschwander-Tetri, B.A., 2017. Non-alcoholic fatty liver disease. BMC Med. 15, 45.

Noorimotlagh, Z., Haghighi, N.J., Ahmadimoghadam, M., Rahim, F., 2016. An updated systematic review on the possible effect of nonylphenol on male fertility. Environ. Sci. Pollut. Res. Int.

Ogawa, S., Lozach, J., Jepsen, K., Sawka-Verhelle, D., Perissi, V., Sasik, R., Rose, D.W., Johnson, R.S., Rosenfeld, M.G., Glass, C.K., 2004. A nuclear receptor corepressor transcriptional checkpoint controlling activator protein 1-dependent gene networks required for macrophage activation. Proc. Natl. Acad. Sci. U.S.A. 101, 14461–14466.

Park, M.A., Hwang, K.A., Lee, H.R., Yi, B.R., Choi, K.C., 2011. Cell growth of BG-1 ovarian cancer cells was promoted by 4-Tert-octylphenol and 4-nonylphenol via downregulation of TGF-beta Receptor 2 and upregulation of c-myc. Toxicol. Res. 27, 253–259.

Paterni, I., Granchi, C., Minutolo, F., 2016. Risks and benefits related to alimentary exposure to xenoestrogens. Crit. Rev. Food Sci. Nutr.

Petit, F., Le Goff, P., Cravédi, J.-P., Kah, O., Valotaire, Y., Pakdel, F., 1999. Trout oestrogen receptor sensitivity to xenobiotics as tested by different bioassays. Aquaculture 177, 353–365.

Polimeni, L., Del Ben, M., Baratta, F., Perri, L., Albanese, F., Pastori, D., Violi, F., Angelico, F., 2015. Oxidative stress: new insights on the association of non-alcoholic fatty liver disease and atherosclerosis. World J. Hepatol. 7, 1325–1336.

Ray, P.D., Huang, B.-W., Tsuji, Y., 2012. Reactive oxygen species (ROS) homeostasis and redox regulation in cellular signaling. Cell Signal. 24, 981–990.

Reuter, S., Gupta, S.C., Chaturvedi, M.M., Aggarwal, B.B., 2010. Oxidative stress, inflammation, and cancer: How are they linked? Free Radic. Biol. Med. 49, 1603–1616.

Ruggeri, B., Ubaldi, M., Lourdusamy, A., Soverchia, L., Ciccocioppo, R., Hardiman, G., Baker, M.E., Palermo, F., Polzonetti-Magni, A.M., 2008. Variation of the genetic expression pattern after exposure to estradiol-17beta and 4-nonylphenol in male zebrafish (Danio rerio). Gen. Comp. Endocrinol. 158, 138–144.

Sasik, R., Woelk, C.H., Corbeil, J., 2004. Microarray truths and consequences. J. Mol. Endocrinol. 33, 1–9.

Schmidt-Heck, W., Wonne, E.C., Hiller, T., Menzel, U., Koczan, D., Damm, G., Seehofer, D., Knospel, F., Freyer, N., Guthke, R., Dooley, S., Zeilinger, K., 2017. Global transcriptional response of human liver cells to ethanol stress of different strength reveals hormetic behavior. Alcohol. Clin. Exp. Res.

Sharma, V.K., Anquandah, G.A., Yngard, R.A., Kim, H., Fekete, J., Bouzek, K., Ray, A.K., Golovko, D., 2009. Nonylphenol, octylphenol, and bisphenol-A in the aquatic environment: a review on occurrence, fate, and treatment. J. Environ. Sci. Health. A Tox. Hazard. Subst. Environ. Eng. 44, 423–442.

Shepard, J.L., Amatruda, J.F., Stern, H.M., Subramanian, A., Finkelstein, D., Ziai, J., Finley, K.R., Pfaff, K.L., Hersey, C., Zhou, Y., Barut, B., Freedman, M., Lee, C., Spitsbergen, J., Neuberg, D., Weber, G., Golub, T.R., Glickman, J.N., Kutok, J.L., Aster, J.C., Zon, L.I., 2005. A zebrafish bmyb mutation causes genome instability and increased cancer susceptibility. Proc. Natl. Acad. Sci. U.S.A. 102, 13194–13199.

Staples, C., Mihaich, E., Carbone, J., Woodburn, K., Klecka, G., 2004. A weight of evidence analysis of the chronic ecotoxicity of nonylphenol ethoxylates, nonylphenol ether carboxylates, and nonylphenol. Hum. Ecol. Risk Assess. 10, 999–1017.

Supek, F., Bosnjak, M., Skunca, N., Smuc, T., 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One 6, e21800.

Thompson, N.L., Mead, J.E., Braun, L., Goyette, M., Shank, P.R., Fausto, N., 1986. Sequential protooncogene expression during rat liver regeneration. Cancer Res. 46, 3111–3117.

Tollefsen, K.-E., Mathisen, R., Stenersen, J., 2002. Estrogen mimics bind with similar affinity and specificity to the hepatic estrogen receptor in Atlantic salmon (Salmo salar) and rainbow trout (Oncorhynchus mykiss). Gen. Comp. Endocrinol. 126, 14–22.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., Pachter, L., 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. 7, 562–578. Vazquez-Duhalt, R., Marquez-Rocha, F., Ponce, E., Licea, A., Viana, M., 2005. Nonylphenol, an integrated vision of a pollutant. Appl. Ecol. Environ. Res. 4, 1–25.

Villeneuve, D.L., Garcia-Reyero, N., Escalon, B.L., Jensen, K.M., Cavallin, J.E., Makynen, E.A., Durhan, E.J., Kahl, M.D., Thomas, L.M., Perkins, E.J., Ankley, G.T., 2012. Ecotoxicogenomics to support ecological risk assessment: a case study with bisphenol A in fish. Environ. Sci. Technol. 46, 51–59.

Villeneuve, D.L., Crump, D., Garcia-Reyero, N., Hecker, M., Hutchinson, T.H., LaLone, C.A., Landesmann, B., Lettieri, T., Munn, S., Nepelska, M., Ottinger, M.A., Vergauwen, L., Whelan, M., 2014. Adverse outcome pathway (AOP) development I: strategies and principles. Toxicol. Sci. 142, 312–320.

Wang, C., Mayer, J.A., Mazumdar, A., Fertuck, K., Kim, H., Brown, M., Brown, P.H., 2011. Estrogen induces c-myc gene expression via an upstream enhancer activated by the estrogen receptor and the AP-1 transcription factor. Mol. Endocrinol. 25, 1527–1538.

Wang, Y., Zhou, J., 2013. Endocrine disrupting chemicals in aquatic environments: a potential reason for organism extinction? Aquat. Ecosyst. Health Manage. 16, 88–93.

White, R., Jobling, S., Hoare, S.A., Sumpter, J.P., Parker, M.G., 1994. Environmentally persistent alkylphenolic compounds are estrogenic. Endocrinology 135, 175–182.

Xin, B., Yamamoto, M., Fujii, K., Ooshio, T., Chen, X., Okada, Y., Watanabe, K., Miyokawa, N., Furukawa, H., Nishikawa, Y., 2017. Critical role of Myc activation in mouse hepatocarcinogenesis induced by the activation of AKT and RAS pathways. Oncogene.

Xu, H., Yang, M., Qiu, W., Pan, C., Wu, M., 2013. The impact of endocrine-disrupting chemicals on oxidative stress and innate immune response in zebrafish embryos. Environ. Toxicol. Chem. 32, 1793–1799.

Yen, C.H., Lin, K.C., Leu, S., Sun, C.K., Chang, L.T., Chai, H.T., Chung, S.Y., Chang, H.W., Ko, S.F., Chen, Y.T., Yip, H.K., 2012. Chronic exposure to environmental contaminant nonylphenol exacerbates adenine-induced chronic renal insufficiency: role of signaling pathways and therapeutic impact of rosuvastatin. Eur. J. Pharm. Sci. 46, 455–467.

Yu, J., Luo, Y., Yang, X.F., Yang, M.X., Yang, J., Yang, X.S., Zhou, J., Gao, F., He, L.T., Xu, J., 2016. Effects of perinatal exposure to nonylphenol on delivery outcomes of pregnant rats and inflammatory hepatic injury in newborn rats. Braz. J. Med. Biol. Res. 49, e5647.

Zeller, K.I., Jegga, A.G., Aronow, B.J., O'Donnell, K.A., Dang, C.V., 2003. An integrated database of genes responsive to the Myc oncogenic transcription factor: identification of direct genomic targets. Genome Biol. 4, R69.

Zhang, W., Shen, X.Y., Zhang, W.W., Chen, H., Xu, W.P., Wei, W., 2017a. Di-(2-ethylhexyl) phthalate could disrupt the insulin signaling pathway in liver of SD rats and L02 cells via PPARgamma. Toxicol. Appl. Pharmacol. 316, 17–26.

Zhang, W., Shen, X.Y., Zhang, W.W., Chen, H., Xu, W.P., Wei, W., 2017b. The effects of di 2-ethyl hexyl phthalate (DEHP) on cellular lipid accumulation in HepG2 cells and its potential mechanisms in the molecular level. Toxicol. Mech. Methods 1–8.

Table 1. Microarray data analysis. Gene enrichment analysis was performed utilizing Gorilla and zebrafish gene IDs. Enriched GO: Biological Process terms are provided for NP (total), E2 (total), NP (unique to NP and not DE expressed in E2) and E2 (unique to E2 and not DE expressed in NP).

GO Term	FDR q-value
NP - Total	
ion transport	1.72E−01
positive regulation of blood circulation	1.81E−01
regulation of cell proliferation	1.95E−01
long-chain fatty acid biosynthetic process	1.99E−01
ventricular cardiac muscle cell development	1.99E−01
positive regulation of ERK1 and ERK2 cascade	2.10E−01
ventricular cardiac myofibril assembly	2.22E−01
unsaturated fatty acid metabolic process	2.25E−01
regulation of insulin receptor signaling pathway	2.49E−01
atrial cardiac myofibril assembly	2.49E−01
epoxygenase P450 pathway	2.66E−01
positive regulation of heart contraction	2.85E−01
regulation of cellular response to insulin stimulus	3.32E−01
antigen processing and presentation	3.76E−01
long-chain fatty acid metabolic process	4.42E−01
E2 - Total	
organic acid metabolic process	5.77E−02
carboxylic acid metabolic process	5.84E−02
L-serine biosynthetic process	7.77E−02
small molecule metabolic process	8.26E−02
oxoacid metabolic process	1.15E−01
carbohydrate metabolic process	1.89E−01
single-organism metabolic process	2.30E−01
Oxidation-reduction process	2.44E−01
L-serine metabolic process	2.48E−01
alpha-amino acid metabolic process	3.73E−01
NP - Unique	
endosome organization	2.57E−01
E2 - Unique	
ruffle organization	3.80E−01
synapse assembly	5.69E−01

Table 2. Microarray data analysis using projected human gene IDs: Human Entrez gene IDs were mapped to zebrafish via Ensembl protein homology. GO: Biological Process and co-expression analysis was performed utilizing the ToppGene suite's functional enrichment tool, ToppFun for DE genes. Enriched GO: Biological Process terms and c-expression signatures are provided for NP (total), E2 (total), NP (unique to NP and not DE expressed in E2) and E2 (unique to E2 and not DE expressed in NP). The most significant terms are presented. Expanded lists of enriched GO terms and co-expression signatures are found in Supplemental Tables S1–8.

GO Term	Bonferroni q-value
GO: Biological process	
NP - Total	
organic acid metabolic process	6.26E−25
cell cycle	4.63E−18
Oxidation-reduction process	1.22E−16
lipid metabolic process	2.21E−10
E2 - Total	
carboxylic acid metabolic process	3.23E−31
Oxidation-reduction process	4.53E−15
response to hormone	1.87E−13
mitotic cell cycle	2.29E−10
NP - Unique	
response to abiotic stimulus	9.49E−07
negative regulation of cell death	5.31E−03
cellular response to DNA damage stimulus	7.89E−03
fatty acid metabolic process	1.10E−02
E2 - Unique	
RNA processing	3.26E−07
response to endogenous stimulus	2.59E−05
translation	1.95E−03
ncRNA processing	3.85E−03
Co-Expression	
NP - Total	
Genes up-regulated in robust Cluster 2 (rC2) of hepatoblastoma samples compared to those in the r obust Cluster 1 (rC1)	3.62E−25
Genes down-regulated in response to both hypoxia and overexpression of an active form of HIF1A [GeneID = 3091]	1.81E−23
E2 - Total	
Genes down-regulated in erythroid progenitor cells from fetal livers of E13.5 embryos with KLF1 [GeneID = 10661] knockout compared to those from the wild type embryos	6.97E−37
Genes up-regulated through activation of mTORC1 complex	1.40E−28

NP - Unique	
Human Liver_Tzur09_1908genes	3.60E−12
Genes whose promoters are bound by MYC [GeneID = 4609], according to MYC Target Gene Database	1.05E−06
E2 - Unique	
Human orthologs of genes up-regulated in the crb ('crash and burn') zebrafish mutant that represents a loss-of-function mutation in BMYB [GeneID = 4605]	5.09E−11
Selected genes up-regulated in response to the Ras inhibitor salirasib [PubChem = 5469318] in a panel of cancer cell lines with constantly active HRAS [GeneID = 3265]	1.19E−08

Table 3. RNAseq data analysis. Gene enrichment analysis was performed utilizing Gorilla and zebrafish gene IDs. Enriched GO: Biological Process terms are provided for NP (total), E2 (total), NP (unique to NP and not DE expressed in E2) and E2 (unique to E2 and not DE expressed in NP).

GO Term	FDR q-value
NP - Total	
response to biotic stimulus	1.24E−03
response to external biotic stimulus	1.54E−03
response to other organism	1.62E−03
multi-organism process	2.99E−03
proteolysis	4.12E−03
response to stimulus	2.06E−02
antigen processing and presentation	2.70E−02
lipid transport	2.74E−02
monocarboxylic acid metabolic process	3.06E−02
single-organism metabolic process	3.73E−02
response to external stimulus	3.88E−02
response to oxygen levels	4.04E−02
response to hypoxia	4.13E−02
response to decreased oxygen levels	4.35E−02
small molecule metabolic process	9.76E−02
regulation of reactive oxygen species metabolic process	1.38E−01
microglia development	1.90E−01
oxoacid metabolic process	2.45E−01
organic acid metabolic process	2.55E−01
carboxylic acid metabolic process	2.83E−01
regulation of immune system process	2.96E−01
defense response to other organism	3.01E−01
immune system process	3.05E−01
response to bacterium	3.07E−01
E2 - Total	
single-organism metabolic process	2.62E−14
small molecule metabolic process	8.98E−14
coenzyme metabolic process	4.15E−11
carboxylic acid metabolic process	6.65E−10
organic acid metabolic process	2.17E−09
cellular response to estrogen stimulus	2.54E−09
oxoacid metabolic process	4.04E−09
oxidation-reduction process	4.09E−09
cofactor metabolic process	4.13E−09
response to estrogen	1.23E−07
cellular response to chemical stimulus	6.10E−07
nicotinamide nucleotide metabolic process	4.41E−06
pyridine nucleotide metabolic process	4.78E−06
monocarboxylic acid metabolic process	6.37E−06
pyridine-containing compound metabolic process	8.60E−06

nucleobase-containing small molecule metabolic process	2.92E−05
oxidoreduction coenzyme metabolic process	4.11E−05
response to stimulus	4.20E−05
single-organism catabolic process	4.87E−05
alpha-amino acid metabolic process	5.17E−05
single-organism biosynthetic process	8.74E−05
organic substance catabolic process	1.10E−04
organonitrogen compound metabolic process	1.32E−04
NADP metabolic process	1.39E−04
response to chemical	1.54E−04
catabolic process	1.85E−04
cellular aldehyde metabolic process	1.90E−04
small molecule catabolic process	1.92E−04
organophosphate metabolic process	2.60E−04
pyruvate metabolic process	2.85E−04
metabolic process	2.91E−04
coenzyme biosynthetic process	6.52E−04
organonitrogen compound catabolic process	8.04E−04
blood coagulation, fibrin clot formation	8.40E−04
nucleotide metabolic process	9.89E−04
proteolysis	9.92E−04
nucleoside phosphate metabolic process	1.02E−03
glucose 6-phosphate metabolic process	1.08E−03
cellular amino acid metabolic process	1.20E−03
tetrahydrofolate metabolic process	1.54E−03
pentose-phosphate shunt	1.58E−03
carboxylic acid catabolic process	1.58E−03
organic acid catabolic process	1.61E−03
glucose metabolic process	2.57E−03
folic acid-containing compound metabolic process	2.74E−03
lipid transport	2.92E−03

NP - Unique

response to oxygen levels	1.23E−02
response to decreased oxygen levels	1.64E−02
response to hypoxia	2.14E−02
single-organism metabolic process	2.91E−02
small molecule metabolic process	3.77E−02
response to external biotic stimulus	4.70E−02
response to biotic stimulus	5.51E−02
monocarboxylic acid metabolic process	9.14E−02
organic acid metabolic process	1.20E−01
oxoacid metabolic process	1.22E−01
microglia development	1.25E−01
carboxylic acid metabolic process	1.31E−01
lipid biosynthetic process	1.33E−01
response to other organism	1.38E−01
myeloid cell development	1.75E−01

multi-organism process	1.93E−01
organic anion transport	2.25E−01
hemoglobin biosynthetic process	3.05E−01
isoprenoid biosynthetic process	3.46E−01
hemoglobin metabolic process	3.64E−01

E2 - Unique

single-organism metabolic process	2.35E−16
small molecule metabolic process	1.69E−15
carboxylic acid metabolic process	9.59E−11
organic acid metabolic process	2.53E−10
coenzyme metabolic process	3.75E−10
oxoacid metabolic process	4.93E−10
oxidation-reduction process	1.00E−09
cellular response to estrogen stimulus	8.19E−08
cofactor metabolic process	1.26E−07
response to estrogen	3.13E−06
single-organism biosynthetic process	3.32E−06
single-organism catabolic process	6.52E−06
alpha-amino acid metabolic process	7.27E−06
nucleobase-containing small molecule metabolic process	1.09E−05
monocarboxylic acid metabolic process	1.70E−05
nicotinamide nucleotide metabolic process	2.12E−05
pyridine nucleotide metabolic process	2.25E−05
small molecule catabolic process	3.22E−05
pyridine-containing compound metabolic process	3.71E−05
organic substance catabolic process	3.72E−05
catabolic process	4.91E−05
organophosphate metabolic process	5.55E−05
cellular response to chemical stimulus	8.39E−05
metabolic process	1.35E−04
oxidoreduction coenzyme metabolic process	1.39E−04
cellular amino acid metabolic process	1.48E−04
carboxylic acid catabolic process	4.33E−04
cellular aldehyde metabolic process	4.36E−04
pyruvate metabolic process	4.46E−04
organic acid catabolic process	4.49E−04
blood coagulation, fibrin clot formation	4.74E−04
organonitrogen compound metabolic process	4.79E−04
NADP metabolic process	4.97E−04
small molecule biosynthetic process	6.18E−04
single-organism process	6.72E−04
nucleotide metabolic process	6.82E−04
nucleoside phosphate metabolic process	7.23E−04
tetrahydrofolate metabolic process	7.67E−04
organonitrogen compound catabolic process	8.78E−04
cellular catabolic process	1.18E−03
folic acid-containing compound metabolic process	1.22E−03

lipid metabolic process	1.32E−03
organic hydroxy compound transport	5.22E−02
single-organism carbohydrate catabolic process	5.52E−02
cofactor biosynthetic process	5.67E−02
steroid metabolic process	5.70E−02
cellular nitrogen compound catabolic process	5.92E−02
carbohydrate catabolic process	5.97E−02
cellular modified amino acid metabolic process	5.99E−02
dicarboxylic acid metabolic process	6.00E−02
ribonucleotide metabolic process	6.02E−02
cellular amino acid biosynthetic process	6.62E−02
lipid biosynthetic process	6.66E−02

Table 4. RNAseq data analysis using projected human gene IDs: Human Entrez gene IDs were mapped to zebrafish via Ensembl protein homology. GO: Biological Process and co-expression analysis was performed utilizing the ToppGene suite's functional enrichment tool, ToppFun for DE genes. Enriched GO: Biological Process terms and c-expression signatures are provided for NP (total), E2 (total), NP (unique to NP and not DE expressed in E2) and E2 (unique to E2 and not DE expressed in NP). The most significant terms are presented. Expanded lists of enriched GO terms and co-expression signatures are found in Supplemental Tables S10–16.

GO Term	Bonferroni q-value
GO: Biological Process	
NP - Total	
immune response	1.60E−08
response to oxygen-containing compound	7.61E−06
inflammatory response	7.94E−04
regulation of protein activation cascade	5.57E−04
E2 - Total	
carboxylic acid metabolic process	4.05E−32
oxidation-reduction process	3.09E−24
lipid metabolic process	9.87E−17
fatty acid metabolic process	1.97E−10
NP - Unique	
reactive oxygen species metabolic process	1.67E−05
protein localization to endoplasmic reticulum	4.43E−03
cholesterol metabolic process	8.32E−03
response to topologically incorrect protein	2.08E−02
E2 - Unique	
organic acid metabolic process	2.58E−28
cellular lipid metabolic process	5.85E−14
cellular catabolic process	1.35E−07
lipid oxidation	3.52E−06
Co-Expression	
NP - Total	
Human orthologs of genes up-regulated in the crb ('crash and burn') zebrafish mutant that represents a loss-of-function mutation in BMYB [GeneID = 4605]	7.99E−07
Genes regulated by NF-kB in response to TNF [GeneID = 7124]	1.04E−06
E2 - Total	
Liver selective genes	7.64E−37
Genes down-regulated in hepatoblastoma samples compared to normal liver tissue	1.56E−22

NP - Unique	
Mouse Liver_White05_638genes	1.38E−12
Genes up-regulated in liver from transgenic mice with reduced expression of POR [GeneID = 5447] in all tissues	2.49E−06
E2 - Unique	
Genes down-regulated in hepatocellular carcinoma (HCC) compared to normal liver samples	8.69E−19
Mouse Liver_Jeong06_492genes	2.07E−13

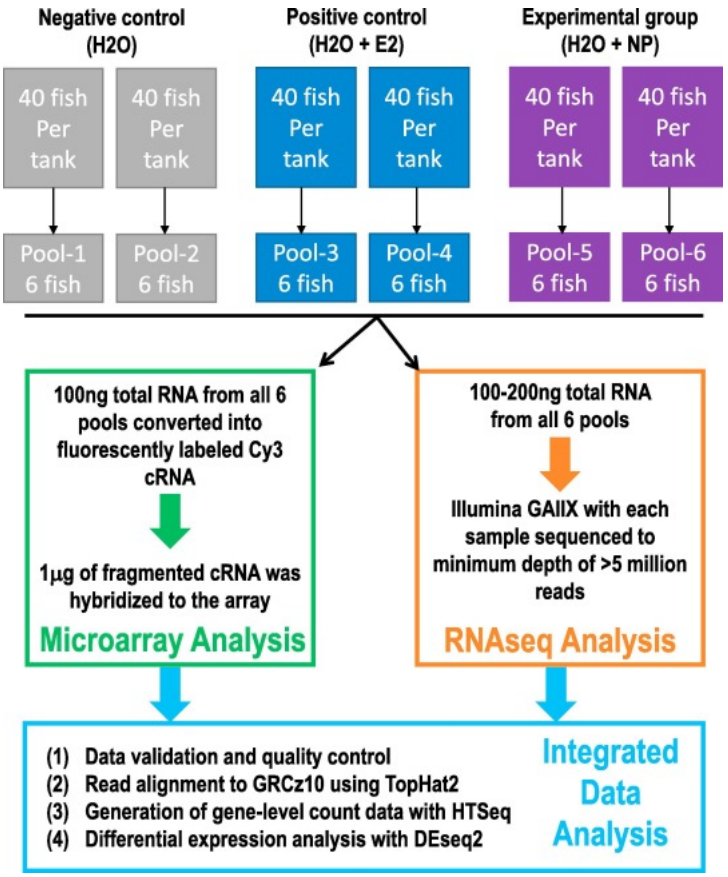
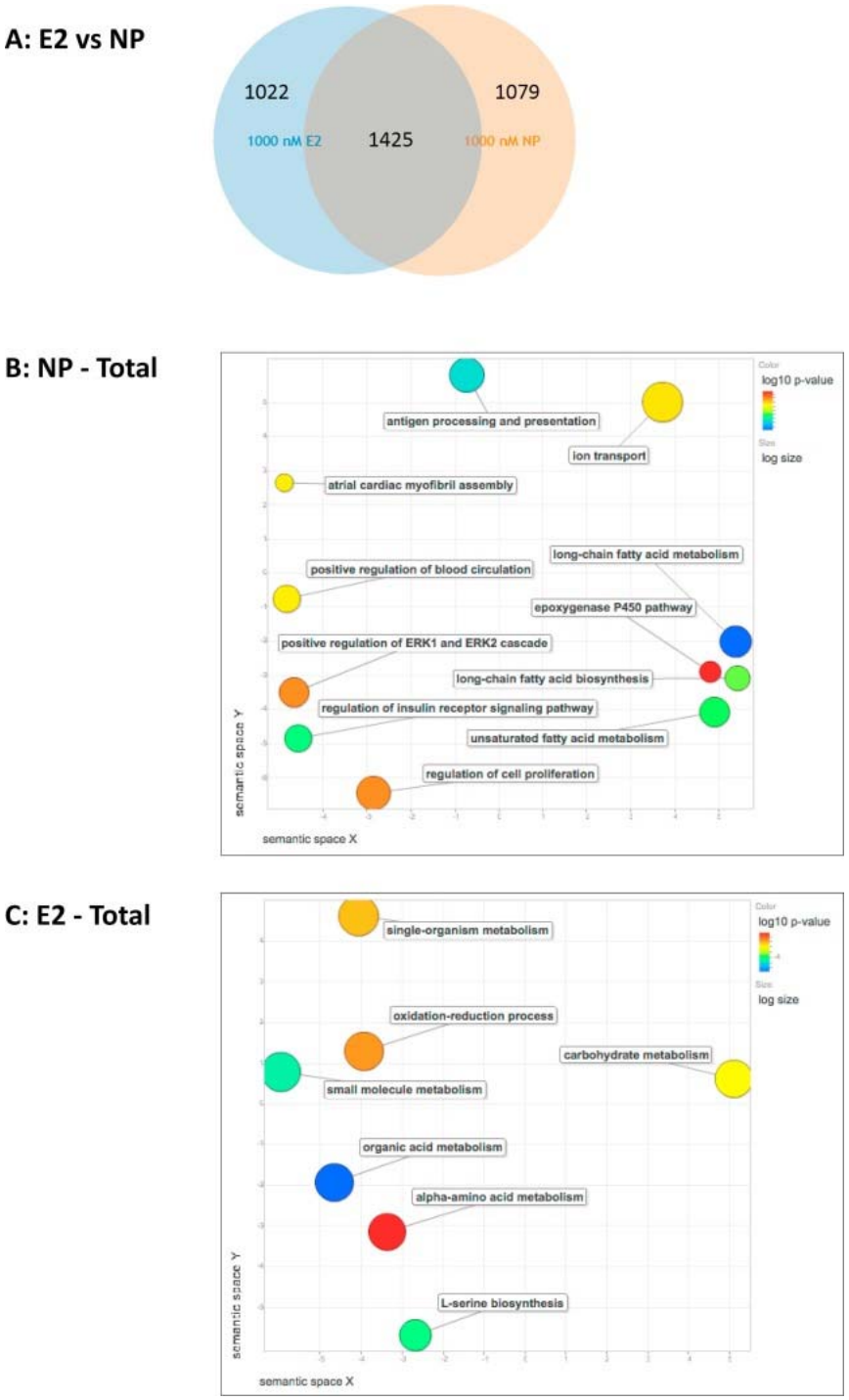


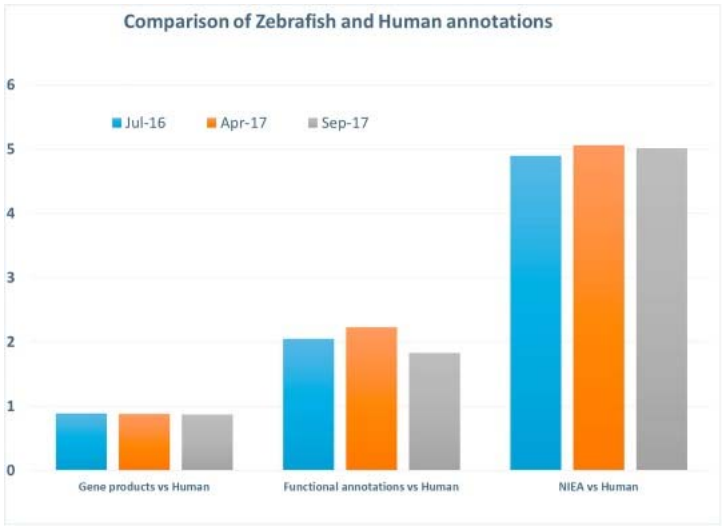
Fig. 1. Experimental workflow schematic.



904

905 Fig. 2. Functional Analyses of microarray data. (A) Overlap of the top 3000 ranked DE expressed liver
906 transcripts from 100 nM NP and 100 nM EE2 exposed adult male zebrafish relative to control fish as
907 determined by microarray analysis. (B-C) Gene Ontology Biological Process analyses: Scatterplots
908 shows the cluster representatives (i.e. terms remaining after the redundancy reduction) in a two
909 dimensional space derived by applying multidimensional scaling to a matrix of the GO terms'
910 semantic similarities. Bubble color indicates the p-value (legend in upper right-hand corner); size
911 indicates the frequency of the GO term in the underlying GOA database (bubbles of more general
912 terms are larger). GO BP analysis of DE genes in (B) NP and (C) E2 exposed livers.

913



914

915 Fig. 3. Comparison of zebrafish and human functional annotations. The x-axis presents the three
916 categories, gene products, functional annotations and non-inferred electronic annotations (NIEA).
917 The y-axis presents the data available for zebrafish relative to human. Zebrafish has larger number of
918 annotated gene products relative to human, 22,504 versus 19,473 (as of September 2017). In
919 human, however non-inferred electronic (NIEA) and functional annotations are >5 times and >2
920 times better defined respectively than they are in zebrafish, based on a query of the GO database in
921 July 2017 (blue), April 2017 (orange) and September 2017 (grey).

922

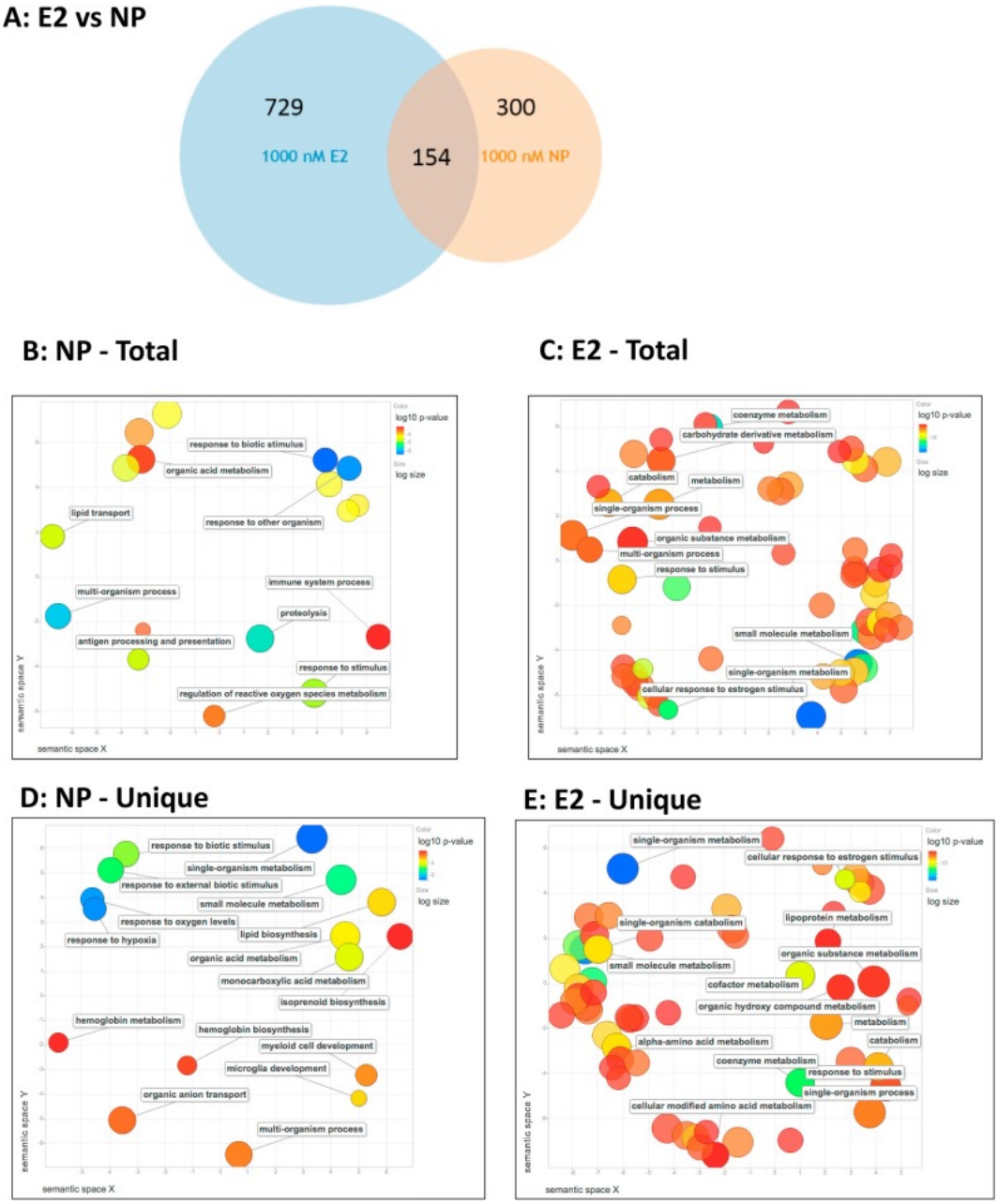
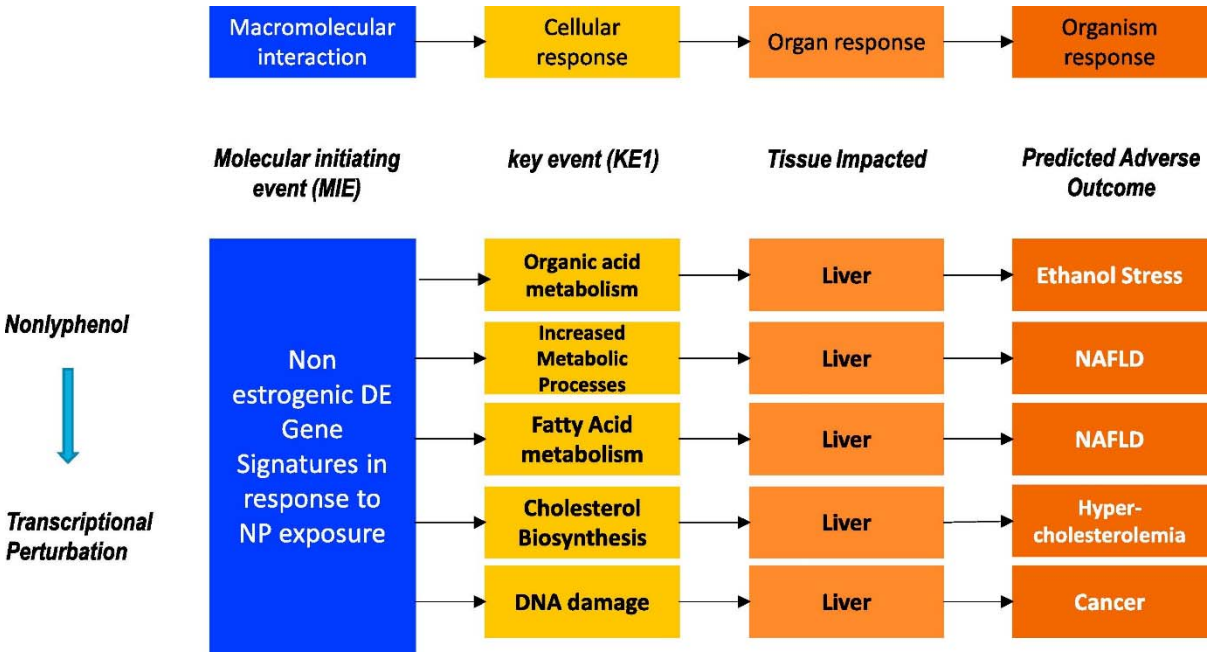


Fig. 4. Functional Analyses of RNAseq data. (A) Overlap of the significant DE expressed liver transcripts (FDR < 0.4) from 100 nM DEHP and 100 nM E2 exposed adult male zebrafish relative to control fish as determined by DESeq2. (B-D) Gene Ontology Biological Process analyses: Scatterplots shows the cluster representatives (i.e. terms remaining after the redundancy reduction) in a two dimensional space derived by applying multidimensional scaling to a matrix of the GO terms' semantic similarities. Bubble color indicates the p-value (legend in upper right-hand corner); size indicates the frequency of the GO term in the underlying GOA database (bubbles of more general terms are larger). GO BP analysis of DE genes in (B) NP and (C) E2 exposed livers. GO BP analysis of DE genes unique to NP (not DE in E2 exposed) (D). GO BP analysis of DE genes unique to E2 (not DE in NP exposed) (E).



936

937 Fig. 5. Summary of findings with regard to the non-estrogenic effects of NP using the adverse
938 outcome pathway framework. An adverse outcome pathway (AOP) is a conceptual framework
939 constructed from prior knowledge that relates exposure of an ED to molecular alterations that result
940 in an adverse health outcome in an individual or population. We summarized our findings from our
941 systems analyses using the differentially expressed genes that represent a signature unique to NP
942 and that does not overlap with E2.

943