The Chemical Defensome of Fish: Conservation and Divergence of Genes Involved in Sensing and Responding to Pollutants Among Five Model Teleosts

Marta Eide
University of Bergen
Xiaokang Zhang
Oslo University Hospital
Odd André Karlsen
University of Bergen
Jared V. Goldstone
Woods Hole Oceanographic Institution
John Stegeman
Woods Hole Oceanographic Institution
Inge Jonassen
University of Bergen
Anders Goksøyr (anders.goksoyr@uib.no)
University of Bergen

Research Article

Keywords: Chemical defensome, environmental contaminants, detoxification, nuclear receptors, biotransformation, antioxidant proteins, heat shock proteins, model species, toxicology

Posted Date: February 9th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-175531/v1

The chemical defensome of fish: conservation and divergence of genes involved in sensing and responding to pollutants among five model teleosts.

Marta Eide[1*], Xiaokang Zhang[2,3*], Odd André Karlsen[1], Jared V. Goldstone[4], John Stegeman[4], Inge Jonassen[2], Anders Goksøyr[1§]

1. Department of Biological Sciences, University of Bergen, Norway
2. Computational Biology Unit, Department of Informatics, University of Bergen, Norway
3. Department of Molecular Oncology, Institute for Cancer Research, Oslo University Hospital-Radiumhospitalet, Norway
4. Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA, USA

[*] The authors contributed equally to the study
[§] Corresponding author: anders.goksoyr@uib.no

Abstract

How an organism copes with chemicals is largely determined by the genes and proteins that collectively function to defend against, detoxify and eliminate chemical stressors. This integrative network includes receptors and transcription factors, biotransformation enzymes, transporters, antioxidants, and metal- and heat-responsive genes, and is collectively known as the *chemical defensome*. Although the types of defensome genes are generally conserved in animals, there are important differences in the complement and function of specific genes between species. Teleost fish is the largest group of vertebrate species and can provide valuable insights into the evolution and functional diversity of defensome genes.

In this study, we compared the genes comprising the chemical defensome of five fish species that span the teleosteii evolutionary branch often used as model species in toxicological studies and environmental monitoring programs: zebrafish (*Danio rerio*), Atlantic cod (*Gadus morhua*), medaka (*Oryzias latipes*), Atlantic killifish (*Fundulus heteroclitus*) and three-spined stickleback (*Gasterosteus aculeatus*). Genome mining revealed evolved differences in the number and composition of defensome genes that can have implication for how these species sense and respond to environmental pollutants. The results indicate that knowledge regarding the diversity and function of the defensome will be important for toxicological testing and risk assessment studies.

Keywords

Chemical defensome; environmental contaminants; detoxification; nuclear receptors; biotransformation; antioxidant proteins; heat shock proteins; model species; toxicology

1. Background

The aquatic environment is a sink for anthropogenic compounds, and aquatic animals are particularly vulnerable to chemical stressors in their natural habitats. Many of these chemicals may profoundly influence organism health, including viability, growth, performance, and reproductive abilities. Aquatic species are also widely used as model organisms to assess responses to environmental pollutants. In the OECD Guidelines for the Testing of Chemicals, 13 tests for toxic properties of chemicals use fish in general, and often specific fish species such as zebrafish, medaka, Atlantic killifish and three-spined stickleback.

The intrinsic defense against toxic chemicals largely depends on a set of genes and proteins collectively known as the chemical defensome [1]. The chemical defensome include a wide range of transcription factors, enzymes, transporters, and antioxidant enzymes that together function to detoxify and eliminate harmful compounds, including xenobiotic and endobiotic chemicals. Thus, the composition of genes comprising a species' chemical defensome will affect the species overall responsiveness and sensitivity towards chemicals stressors. The recent years of sequencing efforts have produced high quality genome assemblies from a wide range of species, facilitating genome-wide mapping and annotation of genes. The chemical defensome was first described in the invertebrates sea urchin and sea anemone [1,2], and was later mapped in zebrafish (*Danio rerio*), coral, arthropods, and partly in tunicates [3-6]. Although these reports show that the overall metabolic pathways involved in the chemical defensome are largely evolutionarily conserved, the detailed comparison of defensome gene composition in different teleost fish species is not studied.

The diversity in both presence and number of gene homologs can vary substantially between fish species due to the two whole genome duplication (WGD) events in early vertebrate evolution [7] and a third fish-specific WGD event [8], in addition to other evolutionary mechanisms such as gene loss, inversions and neo- and subfunctionalizations. For example, we have previously shown that several losses of the pregnane x receptor (*pxr*[*], or *nr1i2*) have occurred independently across teleost evolution [9]. PXRis an important xenosensor and as a ligand-activated transcription factor one of the key regulators of the chemical defensome [10,11]. The importance of PXR in response to chemical stressors in vertebrates, raises questions of how some fish species cope without this gene.

Thus, the objective of this study was to compare the chemical defensome of zebrafish (*Danio rerio*), medaka (*Oryzias latipes*), and Atlantic killifish (*Fundulus heteroclitus*), which are species that have retained a *pxr* gene, to Atlantic cod (*Gadus morhua*) and three-spined stickleback (*Gasterosteus aculeatus*), that have lost this gene by independent mechanisms. Zebrafish, medaka, killifish, and stickleback are established laboratory and environmental model species in both developmental and toxicological studies [12-15]. Atlantic cod is an ecologically and economically important species in the North Atlantic Ocean, and has commonly been used as a bioindicator species in environmental monitoring programs [16-18]. The genome of Atlantic cod was published in 2011 [19], which has facilitated its increased use as model in toxicological studies [20-24]. Although these fish are all benthopelagic species, their natural habitats range from freshwater to marine environments, from tropical to temperate conditions, and reach sizes ranging from less than five cm (zebrafish and medaka) and 15 cm (killifish and stickleback) to 200 cm (Atlantic cod).

[*] In this paper, the nomenclature is in line with the ZFIN Zebrafish Nomenclature Conventions (https://wiki.zfin.org/display/general/ZFIN+Zebrafish+Nomenclature+Conventions). Thus, fish genes are written in lowercase italic and fish proteins in non-italic and first letter uppercase.

Our results showed that putative orthologs of all genes comprising the chemical defensome were retained in these five fish species, except for the absence of *pxr* in Atlantic cod and stickleback. We found that the number of homologs in some gene families can vary greatly between fish species, which could result in differences in the corresponding defense pathway. However, these variations appeared to be randomly distributed in the defensome gene families and were not unique to known target genes of Pxr or other xenosensors. Furthermore, we found that many of chemical defensome genes are not transcribed in early development of zebrafish and stickleback until after hatching. The consequential lack of transcriptional response in zebrafish embryo compared to larvae were further demonstrated following exposure to the model polycyclic aromatic hydrocarbon contaminant, benzo(a)pyrene.

In conclusion, our study represents the first interspecies comparison of the full complement of chemical defensome genes in teleost model species. We found that although most defensome genes have been retained in the teleost genomes over millions of years, there are distinct differences between the species. Based on our results, we suggest a holistic approach to analyze omics datasets from toxicogenomic studies, where differences in the chemical defensome gene complement are taken into consideration.

2. Material and methods

2.1. Sequence resources

The mapping of chemical defensome genes were performed in the most recently published fish genomes available in public databases (**Supplementary Table 1**, available at FAIRDOMhub: https://doi.org/10.15490/fairdomhub.1.document.872.1). For zebrafish (*Danio rerio,* GRCz11), three-spined stickleback (*Gasterosteus aculatus,* BROAD S1), Atlantic killifish (*Fundulus heteroclitus,* Fundulus_heteroclitus-3.0.2), and Japanese medaka HdrR (*Oryzias latipes,* ASM223467v1), we used the genome assemblies and annotations available in ENSEMBL. For Atlantic cod (*Gadus morhua*), we used the recent gadMor3 genome assembly available in NCBI (GCA_902167405). For all fish, we focused on the protein coding genes and transcripts.

2.2. Identification of chemical defensome genes

Two main approaches were used to identify the genes related to the chemical defensome of the fish species. First, using gene names listed in previous publications on the chemical defensome [1-3] (available at FAIRDOMHub: https://doi.org/10.15490/fairdomhub.1.datafile.3957.1), we searched the current annotations in NCBI for Atlantic cod or ENSEMBL for zebrafish, stickleback, killifish, and medaka. For the well-annotated zebrafish genome, this approach successfully identified the genes that are part of the chemical defensome, as previously mapped by Stegeman, et al. [3]. However, only relying on annotations will not identify all defensome genes in the other less characterized fish genomes. Thus, secondly, we also performed hidden Markov model (HMM) searches using HMMER and Pfam profiles representing protein families that are part of the chemical defensome (available at FAIDOMHub: https://doi.org/10.15490/fairdomhub.1.datafile.3956.1) in the genomes of the remaining four fish species.

Putative orthologs of the retrieved protein sequences were identified using reciprocal best hit BLAST searches against the well-annotated zebrafish proteome. To capture any species-specific duplications in the fish genomes compared to the zebrafish reference genome, hits from one-way BLAST hits were also included. The identified peptide sequence IDs were subsequently converted to their related gene IDs using the BioMart tool on ENSEMBL (https://m.ensembl.org/biomart/martview) and R package "mygene" (https://doi.org/doi:10.18129/B9.bioc.mygene). Finally, the resulting gene lists were then refined to contain only members of gene families and subfamilies related to the chemical defensome, using the same defensome gene lists as in the first approach (https://doi.org/10.15490/fairdomhub.1.datafile.3957.1).

2.3. Transcription of chemical defensome genes in early development

RNA-Seq datasets of early developmental stages of zebrafish (expression values in Transcripts Per Million from ArrayExpress: E-ERAD-475) and stickleback (sequencing reads from NCBI BioProject: PRJNA395155) were previously published by White, et al. [25] and Kaitetzidou, et al. [26], respectively.

For stickleback, embryos were sampled at early morula, late morula, mid-gastrula, early organogenesis, and 24 hours post hatching (hph). The sequencing data was processed and analyzed following the automatic pipeline RASflow [27]. Briefly, the reads were mapped to the stickleback genome downloaded from ENSEMBL of version release-100. HISAT2 [28] was

used as aligner and featureCounts [29] was used to count the reads. The library sizes were normalized using Trimmed Mean of M values (TMM) [30] and the Counts Per Million (CPM) were calculated using R package edgeR [31]. The source code and relevant files can be found on GitHub: https://github.com/zhxiaokang/fishDefensome/tree/main/developmentalStages/stickleback/RASflow.

The zebrafish dataset included 18 time points from one cell to five days post fertilization (dpf). In order to best compare to the available stickleback developmental data, we chose to include the following time points in this study: Cleavage_2 cell (early morula), blastula_1k cell (late morula), mid-gastrula, segmentation_1-4-somites (early organogenesis), and larval_protruding_mouth (24 hph).

2.4 Exposure response on defensome genes in zebrafish early development

RNA-Seq datasets of zebrafish exposed to benzo(a)pyrene (B(a)P) (gene counts from NCBI GEO: GSE64198) were previously published by Fang, et al. [32]. Briefly, the datasets are results from the following *in vivo* experiments: Adult zebrafish were exposed to B(a)P for 7-11 days before their eggs were collected and further exposed until 3.3 (embryonic state) and 96 (larvae state) hours post fertilization (hpf). 200 embryos and 10 larvae were pooled for each group, giving three replicate groups of control and exposed at 3.3 hpf and two replicate groups of control and exposed at 96 hpf. The gene counts were then normalized followed by differential expression analysis using edgeR [31].

3. Results and discussion

3.1 Chemical defensome genes present in model fish genomes.

The full complement of chemical defensome genes in zebrafish, killifish, medaka, stickleback and Atlantic cod are available at the FAIRDOMHub (https://doi.org/10.15490/fairdomhub.1.datafile.3958.1). In short, genome analyses of the selected fish species shows that the number of chemical defensome genes range from 446 in stickleback to 510 in zebrafish (**Figure 1**). Although the number of putative homologous genes in each subfamily varies, we found that all gene subfamilies of the chemical defensome is represented in each species, except for the absence of *pxr* in stickleback and Atlantic cod.

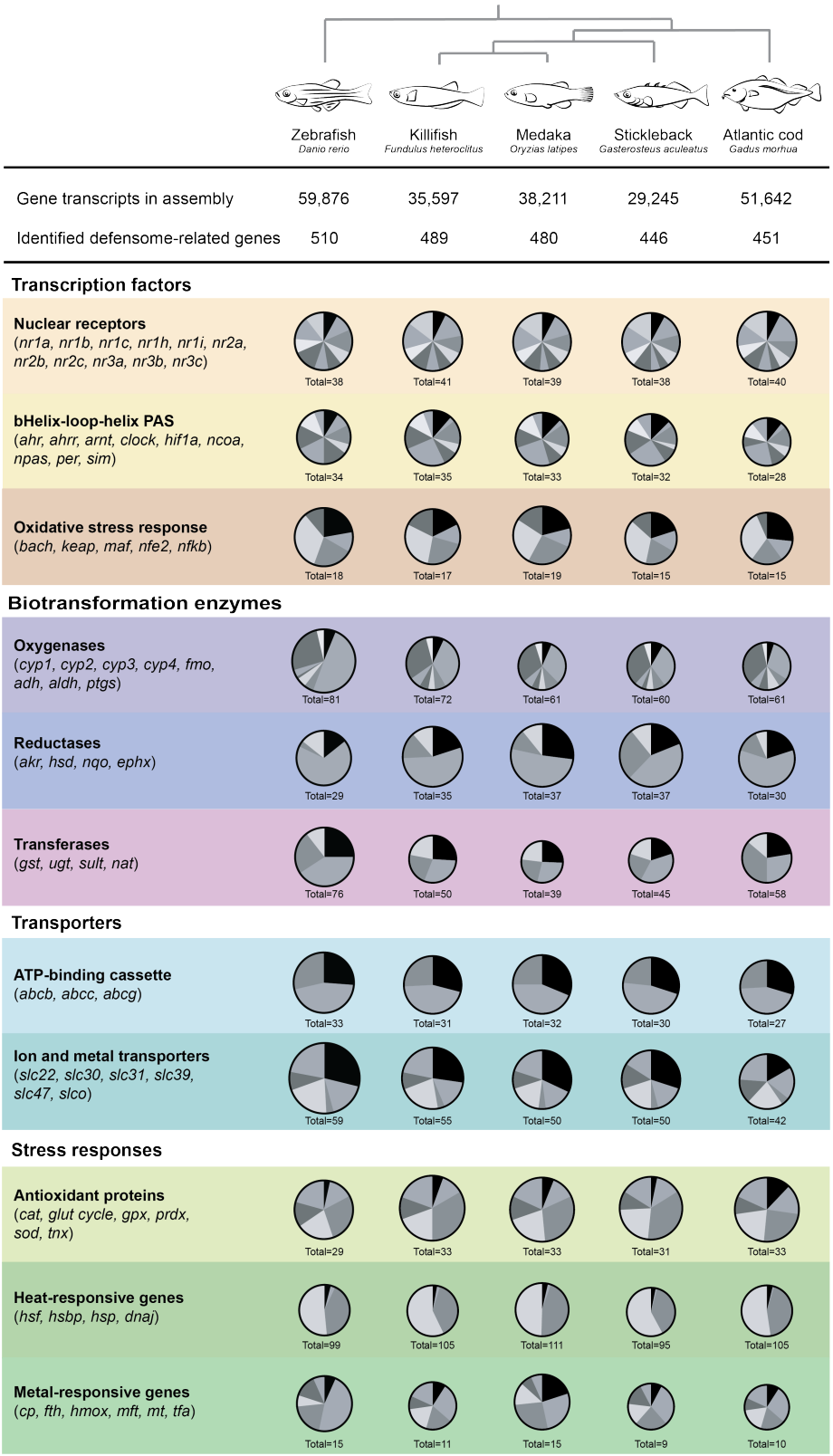

Figure 1: Chemical defensome genes in five model fish species. The genes were identified by searching gene names and using HMMER searches with Pfam profiles, followed by reciprocal or best-hit blast searches towards the zebrafish proteome. The gene families are organized in categories following Gene Ontology annotations and grouped by their role in the chemical defensome. The size of the disk represents the relative number of genes in the different fish genomes within each group, with the number of genes in a specific gene family as slices.

3.2.1 Soluble receptors and transcription factors

Stress-activated transcription factors serve as important first responders to many chemicals, and in turn regulate the transcription of other parts of the chemical defensome. **Nuclear receptors** (NRs) are a superfamily of structurally similar, ligand-activated transcription factors, where members of subfamilies NR1A, B, C, H, and I (such as retinoid acid receptors, peroxisomal proliferator-activated receptors, and liver X receptor), NR2A and B (hepatocyte nuclear factors and retinoid x receptors), and NR3 (such as estrogen receptors and androgen receptor) are involved in the chemical defense [33-36]. All NR subfamilies were found in the five fish genomes, except for the *nr1i2* gene.

NR1I2, or pregnane x receptor (PXR) is considered an important xenosensor responsible for the transcription of many genes involved in the biotransformation of xenobiotics [10,37]. We have previously shown that loss of the *pxr* gene has occurred multiple times in teleost fish evolution [9], including in Atlantic cod and stickleback. Interestingly, our searches did not reveal a *pxr* gene in the ENSEMBL genome assembly of Japanese medaka HdRr, which is considered the reference medaka strain [38,39]. In contrast, a *pxr* gene is annotated in the ENSEMBL genomes of the closely related medaka strains HSOK and HNI. Our previous study identified a sequence similar to zebrafish Pxr in the MEDAKA1 (ENSEMBL release 93) genome [9], and a partial coding sequence (cds) of *pxr* is cloned from medaka genome [40]. However, the specific strain of these resources is not disclosed.

To assess the possible absence of pxr in the medaka HdRr strain, we also performed synteny analysis. In vertebrate species, including fish, *pxr* is flanked by the genes *maats1* and *gsk3b*. These genes are also annotated in medaka HdRr, but the specific gene region has a very low %GC and low sequence quality. Thus, we suspect that the absence of *pxr* in the Japanese medaka HdRr genome is due to a sequencing or assembly error. However, until more evidence of its absence can be presented, we chose to include the medaka *pxr* gene (UniProt ID A8DD90_ORYLA) in our resulting list of medaka chemical defensome genes.

Other important transcription factors are the **basic helix-loop-helix Per-Arnt-Sim** (bHLH/PAS) proteins and the **oxidative stress-activated transcription factors**. bHLH/PAS proteins are involved in circadian rhythms (such as *clock* and *arntl*), hypoxia response (such as *hif1a* and *ncoa*), development (such as *sim*), and the aryl hydrocarbon receptor pathway (*ahr* and *arnt*) (as reviewed by Gu, et al. [41] and Kewley, et al. [42]), while oxidative stress-activated transcription factors respond to changes in the cellular redox status and promote transcription of antioxidant enzymes [43,44]. The latter protein family includes, nuclear factor erythroid-derived 2 (NFE2), NFE2-like (NFE2L, also known as NRFs) 1, 2, and 3, BACH, the dimerization partners small-Mafs (MafF, MafG and MafK), and the inhibitor Kelch-like-ECH-associated protein 1 (KEAP1) [45]. Putative orthologs for all these subfamilies were identified in all five fish genomes.

3.2.2 Biotransformation enzymes

In the first phase of xenobiotic biotransformation, a set of enzymes modifies substrates to more hydrophilic and reactive products. The most important gene family of **oxygenases** is the cytochrome P450 enzymes (CYPs, EC 1.14.-.-), a large superfamily of heme-proteins that initiate the biotransformation of numerous xenobiotic compounds through their monooxygenase activity [46]. The subfamilies considered to be involved in xenobiotic transformation is Cyp1, Cyp2, Cyp3, and Cyp4, and genes of these families were found in all fish species. The number of genes in each subfamily slightly differs from previous mappings

of the CYPome of zebrafish and cod [47,48] (**Supplementary Table 2,** available at FAIRDOMhub: https://doi.org/10.15490/fairdomhub.1.document.872.1). This is likely explained by the sequence and annotation improvement in latest genome assemblies.

Other oxygenases include flavin-dependent monooxygenases (FMOs, EC 1.14.13.8), aldehyde dehydrogenases (ALDH, EC 1.2.1.3), alcohol dehydrogenases (ADHs, EC 1.1.1.1), and prostaglandin-endoperoxide synthases (PTGS, also known as cyclooxygenases, EC 1.14.99.1). Of these, *aldh* represented the largest family in our study, with number of putative gene orthologs ranging from 19 to 22 genes. In comparison, *fmo* had only one gene in zebrafish and four in killifish and cod.

Furthermore, **reductases** modify chemicals by reducing the number of electrons. Reductases include aldo-keto reductases (AKRs, EC 1.1.1), hydroxysteroid dehydrogenases (HSDs, EC 1.1.1), epoxide hydrolases (EPHXs, EC 3.3.2.9 and EC 3.3.2.10), and the NAD(P)H:quinone oxidoreductases (NQOs, EC 1.6.5.2). Interestingly, the number of putative orthologous genes in the *nqo* reductase families varied greatly between the fish species, ranging from one in zebrafish to ten in stickleback. A phylogenetic analysis of the evolutionary relationship of the sequences (**Figure 2**), shows that all fish species have a Nqo1 annotated gene. In addition, medaka, killifish, stickleback and cod have three to nine other closely related genes. The endogenous functions of the different *nqo* genes found in fish, and thus the consequences of their putative evolutionary gain in teleost fish, remains unknown and should be studied further.

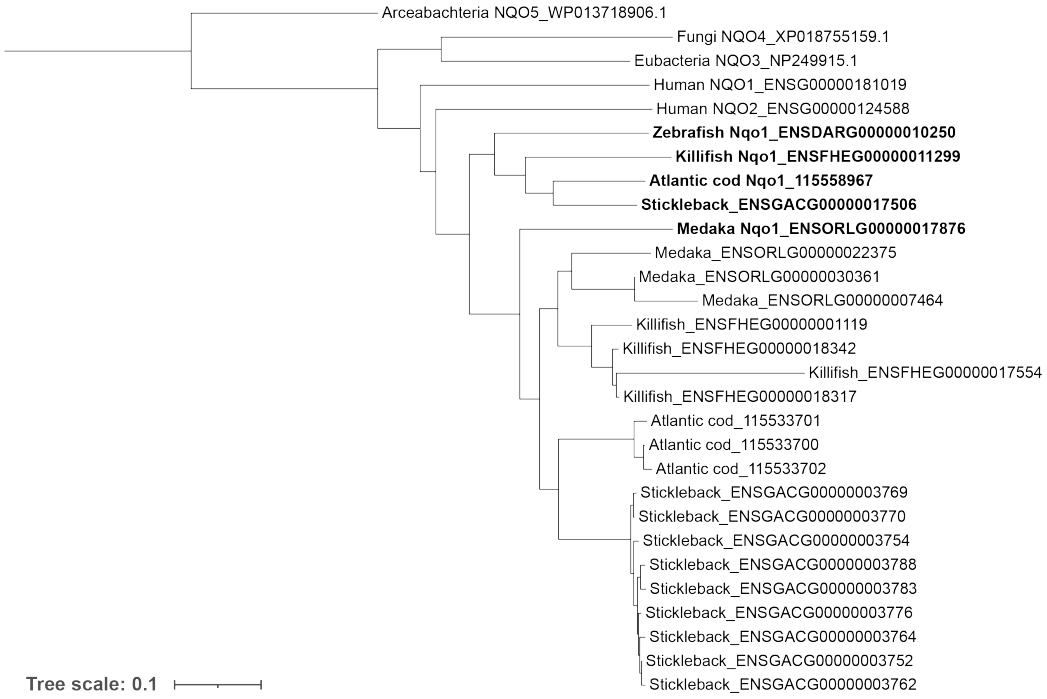

Figure 2: Phylogenetic tree of NAD(P)H:quinone oxidoreductases (NQO), also known as DT-diaphorase (DTD). Multiple sequence alignment and phylogenetic tree was built using Clustal Omega[49] with standard settings. The tree was drawn using iTol[50], and rooted with the archaebacterial NQO5.

In the second phase of biotransformation, endogenous polar molecules are covalently attached to xenobiotic compounds by **transferases**, thus facilitating the generation of more water-soluble products that can be excreted from the cells and the organism. An important

class of such conjugating enzymes are glutathione S-transferases (GSTs, EC 2.5.1.18), which are divided into three superfamilies: the cytosolic GSTs (divided in six subfamilies designated alpha through zeta), the mitochondrial GST (GST kappa) and the membrane-associated GST (designated MAPEG) [51,52]. Other classes of conjugating enzymes in vertebrates include cytosolic sulfotransferases (SULTs, EC 2.8.2), UDP-glucuronosyl transferases (UGTs, EC 2.4.1.17), N-acetyltransferases (NATs, EC 2.3.1.5), and arylamine NATs (aa-NAT).

Our searches identified a *gstp* gene in zebrafish and cod genomes, but not in medaka, killifish and stickleback. Gstp is previously identified as the major Gst isoenzyme in livers of marine salmonid species [53]. Although the specificity of GSTP is not fully understood, its activity seems related to oxidative stress [54]. Furthermore, the total number of *gst* encoding genes was substantially higher in zebrafish (19 genes) compared to the other fish species (9 in stickleback, 10 in medaka, and 13 in killifish and in cod).

Similarly, the number of *ugt* encoding genes is considerably higher in zebrafish compared to the other fish genomes. Whereas 31 *ugt* genes were found annotated in zebrafish, we only identified 15 in killifish, 11 genes in medaka, 17 in stickleback, and 16 in cod. All Ugt subfamilies (ugt1, -2, -5, and -8) are represented in the different fish genomes but include a varying number of homologs. Previous publications indicate that the number of zebrafish *ugt* encoding genes are as high as 45, with the ugt5 subfamily only existing in teleost and amphibian species [55]. One study has found cooperation of NQO1 and UGT in detoxification of vitamin K3 in HEK293 cell line [56]. However, it is not known if there is any correlation between the high number of *ugt* genes and low number of *nqo* genes in zebrafish, relative to the other fish species.

3.2.3 Transporter proteins

Energy-dependent *efflux* transport of compounds across both extra- and intracellular membranes is facilitated by **ATP-binding cassette** (ABC) transporters. In humans, the ABCs are organized into seven subfamilies, named ABC A through G, where proteins of B (also known as MDR1), C and G are known to be involved in multidrug resistance (MDR) [57]. A separate group of proteins is called the Solute Carrier (SLC) 'superfamily,' which consists of diverse non-homologous groups of **ion and metal transporting** membrane proteins that facilitate passive transport [58]. Relevant solute carrier proteins include the drug transporting SLC22 and SLC47, the zinc transporting SLC30 and SLC39, the copper transporting SLC31, and the organic anion transporting SLCO [59].

We found that the number of *abcb*, *abcc*, *abcg*, *slc*, and *slco* genes was similar between the fish species, with zebrafish holding a slightly higher number of homologs. MDR and P-glycoproteins have been relatively understudied in fish [60-62]. A clade of *abch* transporters related to *abcg* is found in some fish species, including zebrafish, but not in Japanese medaka, stickleback and cod [63]. However, as the endogenous function of these genes are not determined, we have not included them specifically into this study.

3.2.4 Antioxidant proteins

Antioxidant proteins protect against harmful reactive oxygen species (ROS), such as superoxide anions, hydrogen peroxide and hydroxyl radicals that are formed as by-products in many physiological processes [64,65]. The enzyme superoxide dismutase (SOD, EC 1.15.1.1) catalyze the conversion of superoxide, one of the most abundant ROS species, to hydrogen peroxide [66]. The further detoxification of hydrogen peroxide can be performed by catalases (CAT, EC 1.11.1.6) and glutathione peroxidases (GPXs, EC 1.11.1.9) [64]. The antioxidants also

include the glutathione (GSH) system, where GSH is supplied by reduction of gluthatione disulphide by glutathione reductase (GSR, EC 1.8.1.7), or by *de novo* synthesis via glutamate cysteine ligase (made up by the subunits GCLC and GCLM, EC 6.3.2.2), and glutathione synthase (GSS, EC 6.3.2.3).

Together with xenobiotic metabolizing enzymes, induction of genes and enzymatic activity involved in antioxidant defense has long been recognized as a gold standard in the biomarker approach to environmental studies [67]. Putative orthologs for all antioxidant genes were clearly identified in the five fish genomes examined.

3.2.5 Heat-responsive genes

Heat-responsive genes represents the largest functional group of genes in the chemical defensome and act in response to a wide range of endogenous and exogenous stressors, such as temperature-shock and heavy metal exposure [68]. In response to stressors, heat shock factors (HSFs) regulate transcription of heat shock proteins (HSPs) [69]. HSPs are divided into families based on their molecular size, and each subfamily has various cellular tasks, including cytoskeleton modulation, protein folding, and chaperone functioning [70,71].

Not much is known about the heat shock protein expression in fish [68]. We found that all fish genomes hold putative orthologs of *hsf*, heat shock binding proteins (*hsbp*) and *hsp*. The number of putative orthologs of *hsp* and *dnaj* (formerly known as *hsp40*) is high in all fish genomes, with the highest number in killifish with 40 and 60 genes, respectively.

3.2.6 Metal-responsive genes

In response to heavy metals such as zinc, cadmium, and copper, metal-responsive transcription factors (MTFs) induce expression of metal-binding proteins, such as metallothioneins (MT), ferritin (ferritin heavy subunit *fth*/*fthl*); heme oxygenases (*hmox,* EC 1.14.99.3,*),* transferrins (*tfa*), and ferroxidase (also known as ceruloplasmin, *cp*; EC 1.16.3.1) [72].

In our study, we found putative orthologs for all gene families, except a *mt* encoding gene in stickleback or cod genome assemblies. Metallothioneins are cysteine-rich, low molecular weight proteins, and can thus be lost due to low-quality sequence and subsequent assembly. In discrepancy with the genome data, Mts are previously described in both Atlantic cod (Hylland et al 1994) and stickleback (Uren Webster 2017), and these protein IDs were included into our overview. Similarly, only one or two *mt* genes were found in zebrafish, killifish, and medaka, and this low number is in line with previous findings on metallothioneins in fish [73].

3.3 Expression of defensome genes in early development of fish

The developmental stage at which a chemical exposure event occurs greatly impacts the effect on fish. In general, chemical exposures during early developmental stages of fish cause the most adverse and detrimental effects. Based on data from the ECETOC Aquatic Toxicity database, fish larvae are more sensitive to substances than embryos and juveniles [74]. However, it is not known how the sensitivity is correlated to the expression of the chemical defensome. As examples in this study, we mapped the expression of the full complement of chemical defensome genes during early development using transcriptomics data from zebrafish and stickleback (**Figure 3,** relevant data available on FAIRDOMHub: https://doi.org/10.15490/fairdomhub.1.assay.1379.1).

Our results showed that there are many defensome genes that are not expressed until after hatching in both species. The delayed genes belonged to all functional categories but were especially prominent where there are several paralogs within the same gene subfamily, for example the transporters and the transferases (**Figure 3**). Other genes were highly expressed at the early developmental stages, before gradually decreasing. Glutathione-related genes, such as *gclc*, were previously shown to be highly expressed in early development of zebrafish due to maternal loading [75], and this was supported in our findings in both zebrafish and stickleback.

Moreover, we found patterns of clustered transcriptional regulation of oxygenase and transferase enzymes in both zebrafish and stickleback. For example, the Ahr target genes *cyp1a*, *gsta.1*, *ugt1a1/7* were transcribed at the early morula stage in both fish species, before the levels decreased at late morula and mid gastrula, before again increasing post hatching (**Figure 3**). In both zebrafish and stickleback, *ahr2* is continuously expressed throughout development, with the highest levels at the 24 hph stage. It has been demonstrated that genes regulated by common transcription factors tend to be located spatially close in the genome sequence and thus facilitate a concerted gene expression [76], and our findings could be supporting such an arrangement.

Finally, there were some genes that were transcribed at very high levels at similar developmental stages in both fish. The heat shock proteins *hspa8* and *hsp90ab1* were highly transcribed at all stages in both zebrafish and stickleback (**Figure 3**). *Hspa8* is a constitutively expressed member of the Hsp70 subfamily, which is previously known as important in rodent embryogenesis [77]. The role of *Hsp90ab1* in development is less known [78]. Furthermore, the ferritin genes fth1a and fth1b were expressed at high levels in both zebrafish and stickleback, respectively (**Figure 3**). However, although ferritin mRNA was found present throughout early development of brown trout, the translated protein was only present after hatching [79]. Importantly, this suggests that there are additional mechanisms that regulate the expression of chemical defensome genes.

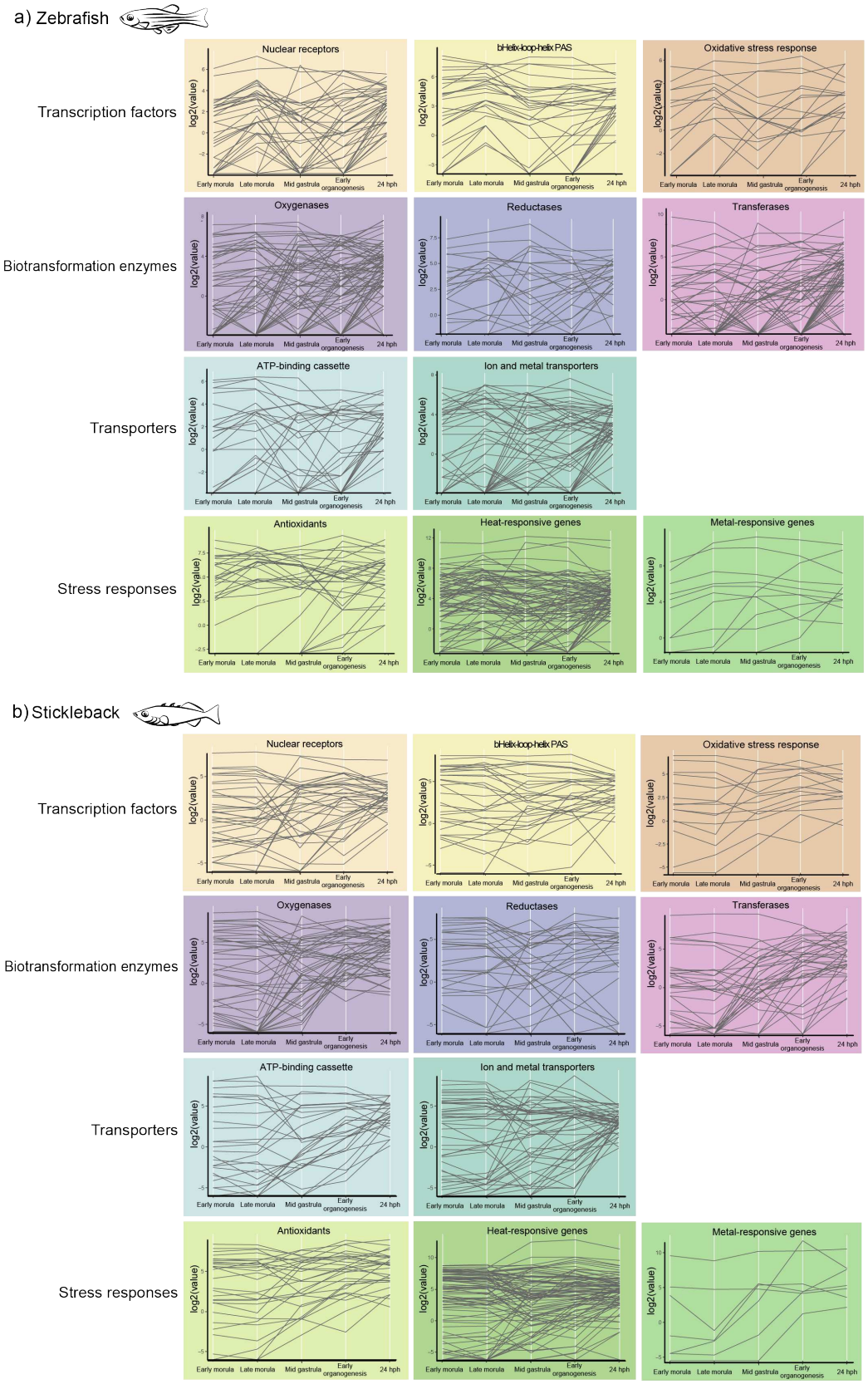

Figure 3: Transcription of chemical defensome genes in early development of a) zebrafish (*Danio rerio*) and b) stickleback (*Gasterasteus aculeatus*). Absolute transcription values (log2 scale) of defensome genes, grouped into their functional category, are shown at early morula, late morula, mid gastrula, early organogenesis, and 24 hours post hatching (hph).

3.4 Exposure response of the defensome genes

Next, we studied the transcriptional effect of a well-known Ahr agonist, benzo(a)pyrene (BaP), on embryonic and larval stages of zebrafish (relevant data available on FAIRDOMHub: https://doi.org/10.15490/fairdomhub.1.datafile.3961.1). At the embryonic stage (3.3 hours post fertilization (hpf), the exposure led to a strong upregulation of *cyp2aa9* (3.87 fold) and an increased transcription of single genes such as *ahr2*, *nfe2*, *aanat1*, and *hspb1* (**Figure 4a**). *Cyp1a*, which is an established biomarker of exposure to BaP and other polycyclic aromatic hydrocarbons (PAH) in fish [80-82], was not induced at this stage. However, as described in the original study [32], we found a strong induction of *cyp1a* (5.67 fold) at the larvae developmental stage (**Figure 4b**). Induction of zebrafish *cyp1a* is previously shown from 24 hpf following exposure to the Ahr model-agonist TCDD[83]. Following exposure at the larval stage, we found a trend of clustered regulation of functionally grouped genes. In general, transcription factors were downregulated, whereas biotransformation enzymes were upregulated. However, the BaP xenosensor, *ahr2*, and the oxidative stress-responsive transcription factor *nfe2l1a*, were both upregulated (0.53 and 1.28 fold, respectively). The crosstalk between these transcription factors following exposure to chemical stressors is previously studied in zebrafish [84,85].

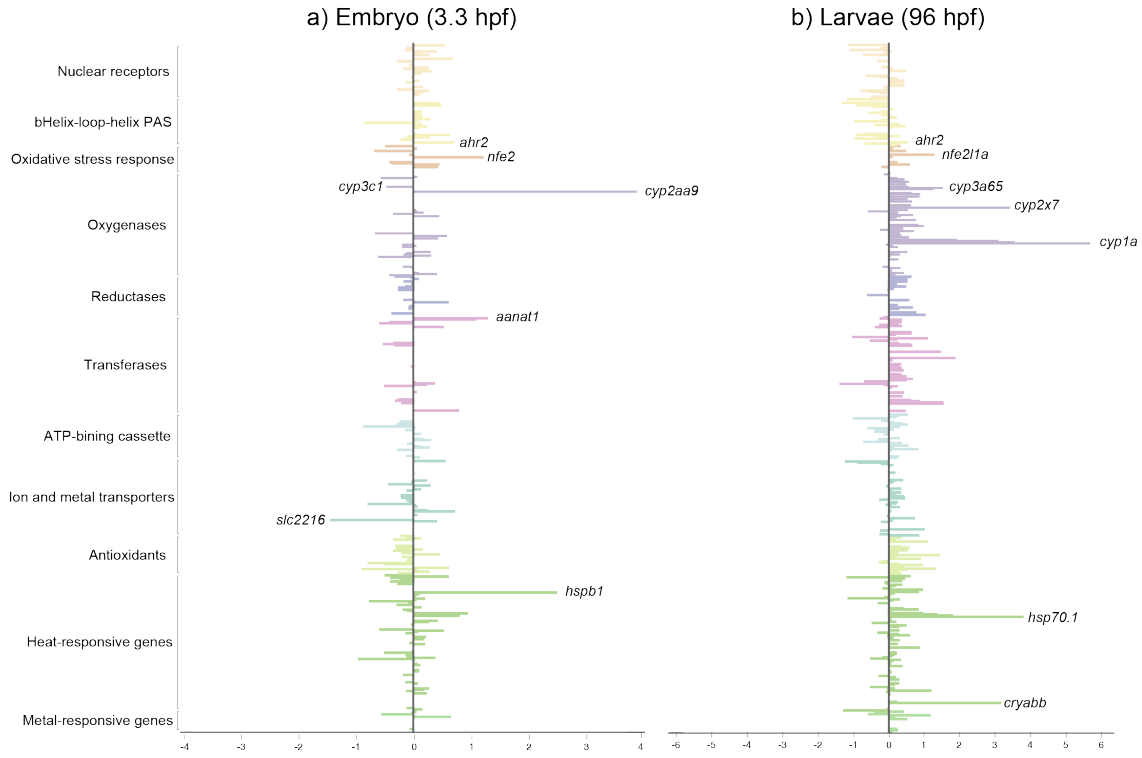

Figure 4: Transcriptional responses on chemical defensome genes in a) zebrafish embryo (3.3 hours post fertilization (hpf)) and b) zebrafish larvae (96 hpf) following exposure to benzo(a)pyrene. The transcription is shown as log2 fold change between exposed and control group at each timepoint. The genes are grouped into their functional categories in the chemical defensome and the name of some genes are indicated for clarity.

4. Summary and perspectives

The chemical defensome is essential for detoxification and subsequent clearance of xenobiotic compounds, and the composition of the defensome can determine the toxicological responses to many chemicals. Our results showed that the number of chemical defensome genes ranged from 446 in three-spined stickleback to 510 in zebrafish, due to a varying number of gene homologs in the evolutionarily conserved modules. Of the five fish included in this study, zebrafish has the highest number of gene homologs in most gene families, with the interesting exception of the *nqo* reductases where medaka, killifish, cod, and especially stickleback, had retained a higher number of homologs compared to only one in zebrafish.

We have previously shown that the stress-activated receptor *pxr* gene has been lost in stickleback and cod, but is retained in zebrafish, Atlantic killifish and medaka (Eide et al., 2018). Still, no differences in the pattern of other defensome genes could be observed linked to this important difference.

Furthermore, we analyzed the transcriptional levels of the defensome genes in early development of zebrafish and stickleback. Importantly, the full complement of defensome genes was not transcribed until after hatching. This was further demonstrated when comparing the transcriptional effects of BaP exposure in two developmental stages of zebrafish, where the larvae had a stronger response that involved more components of the chemical defensome compared to the embryos.

This study presents characterization of the chemical defensome in five different fish species and at different developmental stages as a way of illustrating and understanding inherit interspecies and stage-dependent differences in sensitivity and response to chemical stressors. One aspect not included in the present study is the role of intraspecies, strain-dependent variants in defensome genes. Several studies have identified defensome gene variants linked to pollution tolerance in fish populations, e.g. in the Ahr pathway [86-89]. Lille-Langøy, et al. [90] showed that single-nucleotide polymorphisms (SNPs) in the zebrafish *pxr* gene affect ligand activation patterns. Thus, strain- or population-dependent differences in toxicological responses also play an important role.

Traditionally, studying single molecular biomarkers of exposure has proven very useful in toxicological studies [67,91]. Now, the recent advances in omics technologies enable a more holistic view of toxicological responses, including gene set enrichment analysis and pathway analysis approaches [92-95]. However, these analyses can be challenging when working with less studied and annotated species, such as marine teleosts. As seen from our results, studying the full gene complement of the chemical defense system can identify trends of grouped responses that can provide a better understanding of the overall orchestrated effects to chemical stressors. Such insights will be highly useful in chemical toxicity testing and environmental risk assessment.

Acknowledgements

The work was supported by the Norwegian Research Council as part of the iCOD and iCOD 2.0 projects (grant no. 192441/I30 and 244654/E40), and the dCod 1.0 project (grant no. 248840) which is part of Centre for Digital Life Norway. The American collaborators were funded by the National Institute of Health (USA) NIH P42ES007381 (Boston University Superfund Center to JJS and JVG), NIH R21HD073805 (JVG) and NHI R01ES029917 (JVG) grants. The Ocean

Outlook exchange program funded the trans-Atlantic collaboration. The authors are thankful for the constructive comments from Adele Mennerat in the writing process.

Author contribution

ME and XZ contributed equally to the study design and execution, and wrote the main draft of the manuscript. OAK, JVG, JJS, IJ and AG contributed to the design of the study, discussion of results, and writing of the manuscript. All authors read and approved the submitted article.

4. References

1 Goldstone, J. V. *et al.* The chemical defensome: Environmental sensing and response genes in the Strongylocentrotus purpuratus genome. *Dev. Biol.* **300**, 366-384, doi:Doi 10.1016/J.Ydbio.2006.08.066 (2006).

2 Goldstone, J. V. Environmental sensing and response genes in cnidaria: the chemical defensome in the sea anemone Nematostella vectensis. *Cell Biol. Toxicol.* **24**, 483-502, doi:Doi 10.1007/S10565-008-9107-5 (2008).

3 Stegeman, J. J., Goldstone, J. V. & Hahn, M. E. in *Fish physiology: Zebrafish* Ch. 10, (Elsevier, 2010).

4 Shinzato, C., Hamada, M., Shoguchi, E., Kawashima, T. & Satoh, N. The repertoire of chemical defense genes in the coral Acropora digitifera genome. *Zoolog. Sci.* **29**, 510-517, doi:10.2108/zsj.29.510 (2012).

5 Yadetie, F. *et al.* Conservation and divergence of chemical defense system in the tunicate Oikopleura dioica revealed by genome wide response to two xenobiotics. *BMC Genomics* **13**, 55, doi:10.1186/1471-2164-13-55 (2012).

6 De Marco, L. *et al.* The choreography of the chemical defensome response to insecticide stress: insights into the Anopheles stephensi transcriptome using RNA-Seq. *Sci. Rep.* **7**, doi:ARTN 41312

10.1038/srep41312 (2017).

7 Dehal, P. & Boore, J. L. Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol.* **3**, 1700-1708, doi:ARTN e314

10.1371/journal.pbio.0030314 (2005).

8 Meyer, A. & Van de Peer, Y. From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). *Bioessays* **27**, 937-945, doi:10.1002/bies.20293 (2005).

9 Eide, M. *et al.* Independent losses of a xenobiotic receptor across teleost evolution. *Sci. Rep.* **8**, doi:ARTN 10404

10.1038/s41598-018-28498-4 (2018).

10 Blumberg, B. *et al.* SXR, a novel steroid and xenobiotic-sensing nuclear receptor. *Genes Dev.* **12**, 3195-3205, doi:DOI 10.1101/gad.12.20.3195 (1998).

11 Kliewer, S. A., Goodwin, B. & Willson, T. M. The nuclear pregnane X receptor: A key regulator of xenobiotic metabolism. *Endocr. Rev.* **23**, 687-702, doi:Doi 10.1210/Er.2001.0038 (2002).

12 Hill, A. J., Teraoka, H., Heideman, W. & Peterson, R. E. Zebrafish as a model vertebrate for investigating chemical toxicity. *Toxicol. Sci.* **86**, 6-19, doi:10.1093/toxsci/kfi110 (2005).

13 Embry, M. R. *et al.* The fish embryo toxicity test as an animal alternative method in hazard and risk assessment and scientific research. *Aquat. Toxicol.* **97**, 79-87, doi:10.1016/j.aquatox.2009.12.008 (2010).

14 Burnett, K. G. *et al.* Fundulus as the premier teleost model in environmental biology: Opportunities for new insights using genomics. *Comp Biochem Phys D* **2**, 257-286, doi:10.1016/j.cbd.2007.09.001 (2007).

15 Katsiadaki, I., Scott, A. P. & Mayer, I. The potential of the three-spined stickleback (Gasterosteus aculeatus L.) as a combined biomarker for oestrogens and androgens in European waters. *Mar. Environ. Res.* **54**, 725-728, doi:Pii S0141-1136(02)00110-1

Doi 10.1016/S0141-1136(02)00110-1 (2002).

16 Hylland, K. *et al.* Water column monitoring near oil installations in the North Sea 2001-2004. *Mar Pollut Bull* **56**, 414-429, doi:10.1016/j.marpolbul.2007.11.004 (2008).

17 Brooks, S. J. *et al.* Water Column Monitoring of the Biological Effects of Produced Water from the Ekofisk Offshore Oil Installation from 2006 to 2009. *J Toxicol Env Heal A* **74**, 582-604, doi:Pii 934599436

Doi 10.1080/15287394.2011.550566 (2011).

18 Holth, T. F., Beylich, B. A., Camus, L., Klobucar, G. I. & Hylland, K. Repeated sampling of Atlantic cod (Gadus morhua) for monitoring of nondestructive parameters during exposure to a synthetic produced water. *J. Toxicol. Environ. Health A* **74**, 555-568, doi:10.1080/15287394.2011.550564 (2011).

19 Star, B. *et al.* The genome sequence of Atlantic cod reveals a unique immune system. *Nature* **477**, 207-210, doi:Doi 10.1038/Nature10342 (2011).

20 Yadetie, F. *et al.* RNA-Seq analysis of transcriptome responses in Atlantic cod (Gadus morhua) precision-cut liver slices exposed to benzo[a]pyrene and 17 alpha-ethynylestradiol. *Aquat. Toxicol.* **201**, 174-186, doi:10.1016/j.aquatox.2018.06.003 (2018).
21 Yadetie, F., Karlsen, O. A., Eide, M., Hogstrand, C. & Goksoyr, A. Liver transcriptome analysis of Atlantic cod (Gadus morhua) exposed to PCB 153 indicates effects on cell cycle regulation and lipid metabolism. *BMC Genomics* **15**, doi:Artn 481
10.1186/1471-2164-15-481 (2014).
22 Bizarro, C., Eide, M., Hitchcock, D. J., Goksoyr, A. & Ortiz-Zarragoitia, M. Single and mixture effects of aquatic micropollutants studied in precision-cut liver slices of Atlantic cod (Gadus morhua). *Aquat. Toxicol.* **177**, 395-404, doi:10.1016/j.aquatox.2016.06.013 (2016).
23 Hansen, B. H. *et al.* Embryonic exposure to produced water can cause cardiac toxicity and deformations in Atlantic cod (Gadus morhua) and haddock (Melanogrammus aeglefinus) larvae. *Mar. Environ. Res.* **148**, 81-86, doi:10.1016/j.marenvres.2019.05.009 (2019).
24 Eide, M., Karlsen, O. A., Kryvi, H., Olsvik, P. A. & Goksøyr, A. Precision-cut liver slices of Atlantic cod (Gadus morhua): An in vitro system for studying the effects of environmental contaminants. *Aquat. Toxicol.* **153**, 110-115 (2014).
25 White, R. J. *et al.* A high-resolution mRNA expression time course of embryonic development in zebrafish. *Elife* **6**, doi:ARTN e30860
10.7554/eLife.30860.001 (2017).
26 Kaitetzidou, E., Katsiadaki, I., Lagnel, J., Antonopoulou, E. & Sarropoulou, E. Unravelling paralogous gene expression dynamics during three-spined stickleback embryogenesis. *Sci. Rep.* **9**, doi:ARTN 3752
10.1038/s41598-019-40127-2 (2019).
27 Zhang, X. K. & Jonassen, I. RASflow: an RNA-Seq analysis workflow with Snakemake. *BMC Bioinformatics* **21**, doi:ARTN 110
10.1186/s12859-020-3433-x (2020).
28 Kim, D., Landmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12**, 357-U121, doi:10.1038/Nmeth.3317 (2015).
29 Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930, doi:10.1093/bioinformatics/btt656 (2014).
30 Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25, doi:10.1186/gb-2010-11-3-r25 (2010).
31 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140, doi:10.1093/bioinformatics/btp616 (2010).
32 Fang, X. F. *et al.* Transcriptomic Changes in Zebrafish Embryos and Larvae Following Benzo[a]pyrene Exposure. *Toxicol. Sci.* **146**, 395-411, doi:10.1093/toxsci/kfv105 (2015).
33 Mangelsdorf, D. J. *et al.* The Nuclear Receptor Superfamily - the 2nd Decade. *Cell* **83**, 835-839 (1995).
34 Jin, L. H. & Li, Y. Structural and functional insights into nuclear receptor signaling. *Adv Drug Deliver Rev* **62**, 1218-1226, doi:10.1016/j.addr.2010.08.007 (2010).
35 Aranda, A. & Pascual, A. Nuclear hormone receptors and gene expression. *Physiol. Rev.* **81**, 1269-1304 (2001).
36 Bertrand, S. *et al.* Unexpected novel relational links uncovered by extensive developmental profiling of nuclear receptor expression. *Plos Genet* **3**, 2085-2100, doi:ARTN e188
DOI 10.1371/journal.pgen.0030188 (2007).
37 Kretschmer, X. C. & Baldwin, W. S. CAR and PXR: xenosensors of endocrine disrupters? *Chem. Biol. Interact.* **155**, 111-128, doi:10.1016/j.cbi.2005.06.003 (2005).
38 Spivakov, M. *et al.* Genomic and phenotypic characterization of a wild medaka population: towards the establishment of an isogenic population genetic resource in fish. *G3 (Bethesda)* **4**, 433-445, doi:10.1534/g3.113.008722 (2014).
39 Kirchmaier, S., Naruse, K., Wittbrodt, J. & Loosli, F. The genomic and genetic toolbox of the teleost medaka (Oryzias latipes). *Genetics* **199**, 905-918, doi:10.1534/genetics.114.173849 (2015).
40 Milnes, M. R. *et al.* Activation of steroid and xenobiotic receptor (SXR, NR1I2) and its orthologs in laboratory, toxicologic, and genome model species. *Environ. Health Perspect.* **116**, 880-885, doi:Doi 10.1289/Ehp.10853 (2008).

41 Gu, Y. Z., Hogenesch, J. B. & Bradfield, C. A. The PAS superfamily: Sensors of environmental and developmental signals. *Annu. Rev. Pharmacol. Toxicol.* **40**, 519-561, doi:Doi 10.1146/Annurev.Pharmtox.40.1.519 (2000).
42 Kewley, R. J., Whitelaw, M. L. & Chapman-Smith, A. The mammalian basic helix-loop-helix/PAS family of transcriptional regulators. *Int. J. Biochem. Cell Biol.* **36**, 189-204, doi:Doi 10.1016/S1357-2725(03)00211-5 (2004).
43 Nguyen, T., Sherratt, P. J. & Pickett, C. B. Regulatory mechanisms controlling gene expression mediated by the antioxidant response element. *Annu. Rev. Pharmacol. Toxicol.* **43**, 233-260, doi:Doi 10.1146/Annurev.Pharmtox.43.100901.140229 (2003).
44 Oyake, T. *et al.* Bach proteins belong to a novel family of BTB-basic leucine zipper transcription factors that interact with MafK and regulate transcription through the NF-E2 site. *Mol. Cell. Biol.* **16**, 6083-6095 (1996).
45 Nguyen, T., Nioi, P. & Pickett, C. B. The Nrf2-Antioxidant Response Element Signaling Pathway and Its Activation by Oxidative Stress. *J. Biol. Chem.* **284**, 13291-13295, doi:Doi 10.1074/Jbc.R900010200 (2009).
46 Nelson, D. R. *et al.* Comparison of cytochrome P450 (CYP) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudogenes and alternative-splice variants. *Pharmacogenetics* **14**, 1-18, doi:Doi 10.1097/01.Fpc.0000054151.92680.31 (2004).
47 Karlsen, O. A., Puntervoll, P. & Goksøyr, A. Mass spectrometric analyses of microsomal cytochrome P450 isozymes isolated from beta-naphthoflavone-treated Atlantic cod (Gadus morhua) liver reveal insights into the cod CYPome. *Aquat. Toxicol.* **108**, 2-10, doi:Doi 10.1016/J.Aquatox.2011.08.018 (2012).
48 Goldstone, J. V. *et al.* Identification and developmental expression of the full complement of Cytochrome P450 genes in Zebrafish. *BMC Genomics* **11**, doi:Artn 643
Doi 10.1186/1471-2164-11-643 (2010).
49 Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636-W641, doi:10.1093/nar/gkz268 (2019).
50 Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* **47**, W256-W259, doi:10.1093/nar/gkz239 (2019).
51 Hayes, J. D., Flanagan, J. U. & Jowsey, I. R. Glutathione transferases. *Annu. Rev. Pharmacol. Toxicol.* **45**, 51-88, doi:10.1146/annurev.pharmtox.45.120403.095857 (2005).
52 Nebert, D. W. & Vasiliou, V. Analysis of the glutathione S-transferase (GST) gene family. *Human genomics* **1**, 460-464 (2004).
53 Dominey, R. J., Nimmo, I. A., Cronshaw, A. D. & Hayes, J. D. The Major Glutathione-S-Transferase in Salmonid Fish Livers Is Homologous to the Mammalian Pi-Class Gst. *Comp Biochem Phys B* **100**, 93-98, doi:Doi 10.1016/0305-0491(91)90090-Z (1991).
54 Tew, K. D. *et al.* The role of glutathione S-transferase P in signaling pathways and S-glutathionylation in cancer. *Free Radic. Biol. Med.* **51**, 299-313, doi:10.1016/j.freeradbiomed.2011.04.013 (2011).
55 Huang, H. Y. & Wu, Q. Cloning and Comparative Analyses of the Zebrafish Ugt Repertoire Reveal Its Evolutionary Diversity. *PLoS One* **5**, doi:ARTN e9144
10.1371/journal.pone.0009144 (2010).
56 Nishiyama, T. *et al.* Cooperation of NAD(P)H:quinone oxidoreductase 1 and UDP-glucuronosyltransferases reduces menadione cytotoxicity in HEK293 cells. *Biochem. Biophys. Res. Commun.* **394**, 459-463, doi:10.1016/j.bbrc.2009.12.113 (2010).
57 Dean, M., Hamon, Y. & Chimini, G. The human ATP-binding cassette (ABC) transporter superfamily. *J. Lipid Res.* **42**, 1007-1017 (2001).
58 Hediger, M. A. *et al.* The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteins - Introduction. *Pflug Arch Eur J Phy* **447**, 465-468, doi:Doi 10.1007/S00424-003-1192-Y (2004).
59 Roth, M., Obaidat, A. & Hagenbuch, B. OATPs, OATs and OCTs: the organic anion and cation transporters of the SLCO and SLC22A gene superfamilies. *Br. J. Pharmacol.* **165**, 1260-1287, doi:Doi 10.1111/J.1476-5381.2011.01724.X (2012).
60 Fischer, S. *et al.* Abcb4 acts as multixenobiotic transporter and active barrier against chemical uptake in zebrafish (Danio rerio) embryos. *BMC Biol.* **11**, doi:Artn 69
10.1186/1741-7007-11-69 (2013).

61 Jackson, J. S. & Kennedy, C. J. Regulation of hepatic abcb4 and cyp3a65 gene expression and multidrug/multixenobiotic resistance (MDR/MXR) functional activity in the model teleost, Danio rerio (zebrafish). *Comp Biochem Phys C* **200**, 34-41, doi:10.1016/j.cbpc.2017.06.004 (2017).
62 Ferreira, M., Costa, J. & Reis-Henriques, M. A. ABC transporters in fish species: a review. *Front. Physiol.* **5**, doi:ARTN 266
10.3389/fphys.2014.00266 (2014).
63 Jeong, C. B. *et al.* Marine medaka ATP-binding cassette (ABC) superfamily and new insight into teleost Abch nomenclature. *Sci. Rep.* **5**, doi:ARTN 15409
10.1038/srep15409 (2015).
64 Droge, W. Free radicals in the physiological control of cell function. *Physiol. Rev.* **82**, 47-95 (2002).
65 Dupre-Crochet, S., Erard, M. & Nuss, O. ROS production in phagocytes: why, when, and where? *J. Leukoc. Biol.* **94**, 657-670, doi:10.1189/jlb.1012544 (2013).
66 Deby, C. & Goutier, R. New Perspectives on the Biochemistry of Superoxide Anion and the Efficiency of Superoxide Dismutases. *Biochem. Pharmacol.* **39**, 399-405, doi:Doi 10.1016/0006-2952(90)90043-K (1990).
67 van der Oost, R., Beyer, J. & Vermeulen, N. P. E. Fish bioaccumulation and biomarkers in environmental risk assessment: a review. *Environ. Toxicol. Pharmacol.* **13**, 57-149 (2003).
68 Iwama, G. K., Thomas, P. T., Forsyth, R. H. B. & Vijayan, M. M. Heat shock protein expression in fish. *Rev. Fish Biol. Fish.* **8**, 35-56, doi:Doi 10.1023/A:1008812500650 (1998).
69 Morimoto, R. I. Cells in Stress - Transcriptional Activation of Heat-Shock Genes. *Science* **259**, 1409-1410, doi:Doi 10.1126/Science.8451637 (1993).
70 Basu, N. *et al.* Heat shock protein genes and their functional significance in fish. *Gene* **295**, 173-183, doi:Pii S0378-1119(02)00687-X
Doi 10.1016/S0378-1119(02)00687-X (2002).
71 Kampinga, H. H. *et al.* Guidelines for the nomenclature of the human heat shock proteins. *Cell Stress Chaperones* **14**, 105-111, doi:Doi 10.1007/S12192-008-0068-7 (2009).
72 Brugnera, E. *et al.* Cloning, Chromosomal Mapping and Characterization of the Human Metal-Regulatory Transcription Factor Mtf-1. *Nucleic Acids Res.* **22**, 3167-3173, doi:Doi 10.1093/Nar/22.15.3167 (1994).
73 Smirnov, L. P., Sukhovskaya, I. V. & Nemova, N. N. Effects of environmental factors on low-molecular-weight peptides of fishes: A review. *Russ J Ecol+* **36**, 41-47, doi:DOI 10.1007/s11184-005-0007-0 (2005).
74 Hutchinson, T. H., Solbe, J. & Kloepper-Sams, P. J. Analysis of the ECETOC aquatic toxicity (EAT) database - III - Comparative toxicity of chemical substances to different life stages of aquatic organisms. *Chemosphere* **36**, 129-142, doi:Doi 10.1016/S0045-6535(97)10025-X (1998).
75 Timme-Laragy, A. R. *et al.* Glutathione redox dynamics and expression of glutathione-related genes in the developing embryo. *Free Radic. Biol. Med.* **65**, 89-101, doi:10.1016/j.freeradbiomed.2013.06.011 (2013).
76 Zhang, J. *et al.* Spatial clustering and common regulatory elements correlate with coordinated gene expression. *PLoS Comput. Biol.* **15**, e1006786, doi:10.1371/journal.pcbi.1006786 (2019).
77 Luft, J. C. & Dix, D. J. Hsp70 expression and function during embryogenesis. *Cell Stress Chaperones* **4**, 162-170, doi:Doi 10.1379/1466-1268(1999)004<0162:Heafde>2.3.Co;2 (1999).
78 Haase, M. & Fitze, G. HSP90AB1: Helping the good and the bad. *Gene* **575**, 171-186, doi:10.1016/j.gene.2015.08.063 (2016).
79 Andersen, O. Accumulation of waterborne iron and expression of ferritin and transferrin in early developmental stages of brown trout (Salmo trutta). *Fish Physiol. Biochem.* **16**, 223-231, doi:Doi 10.1023/A:1007729900376 (1997).
80 Goksøyr, A. Purification of hepatic microsomal cytochromes P-450 from ß-naphthoflavone-treated Atlantic cod *(Gadus morhua)*, a marine teleost fish. *Biochim. Biophys. Acta* **840**, 409-417 (1985).
81 Goksøyr, A. & Förlin, L. The cytochrome P450 system in fish, aquatic toxicology, and environmental monitoring. *Aquat. Toxicol.* **22**, 287-312 (1992).
82 Stegeman, J. J. & Lech, J. J. Cytochrome-P-450 Monooxygenase Systems in Aquatic Species - Carcinogen Metabolism and Biomarkers for Carcinogen and Pollutant Exposure. *Environ. Health Perspect.* **90**, 101-109, doi:Doi 10.2307/3430851 (1991).

83 Andreasen, E. A. *et al.* Tissue-specific expression of AHR2, ARNT2, and CYP1A in zebrafish embryos and larvae: Effects of developmental stage and 2,3,7,8-tetrachlorodibenzo-p-dioxin exposure. *Toxicol. Sci.* **68**, 403-419, doi:DOI 10.1093/toxsci/68.2.403 (2002).

84 Hahn, M. E., Timme-Laragy, A. R., Karchner, S. I. & Stegeman, J. J. Nrf2 and Nrf2-related proteins in development and developmental toxicity: Insights from studies in zebrafish (Danio rerio). *Free Radic. Biol. Med.* **88**, 275-289, doi:10.1016/j.freeradbiomed.2015.06.022 (2015).

85 Rousseau, M. E. *et al.* Regulation of Ahr signaling by Nrf2 during development: Effects of Nrf2a deficiency on PCB126 embryotoxicity in zebrafish (Danio rerio). *Aquat. Toxicol.* **167**, 157-171, doi:10.1016/j.aquatox.2015.08.002 (2015).

86 Wirgin, I. *et al.* Mechanistic basis of resistance to PCBs in Atlantic tomcod from the Hudson River. *Science* **331**, 1322-1325, doi:10.1126/science.1197296 (2011).

87 Oziolor, E. M., Bigorgne, E., Aguilar, L., Usenko, S. & Matson, C. W. Evolved resistance to PCB- and PAH-induced cardiac teratogenesis, and reduced CYP1A activity in Gulf killifish (Fundulus grandis) populations from the Houston Ship Channel, Texas. *Aquat. Toxicol.* **150**, 210-219, doi:10.1016/j.aquatox.2014.03.012 (2014).

88 Williams, L. M. & Oleksiak, M. F. Ecologically and evolutionarily important SNPs identified in natural populations. *Mol. Biol. Evol.* **28**, 1817-1826, doi:10.1093/molbev/msr004 (2011).

89 Reid, N. M. *et al.* The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. *Science* **354**, 1305-1308, doi:10.1126/science.aah4993 (2016).

90 Lille-Langøy, R. *et al.* Sequence Variations in pxr (nr1i2) From Zebrafish (Danio rerio) Strains Affect Nuclear Receptor Function. *Toxicol. Sci.* **168**, 28-39, doi:10.1093/toxsci/kfy269 (2019).

91 Peakall, D. B. The Role of Biomarkers in Environmental Assessment .1. Introduction. *Ecotoxicology* **3**, 157-160, doi:Doi 10.1007/Bf00117080 (1994).

92 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545-15550, doi:10.1073/pnas.0506580102 (2005).

93 Fabregat, A. *et al.* Reactome pathway analysis: a high-performance in-memory approach. *BMC Bioinformatics* **18**, 142, doi:10.1186/s12859-017-1559-2 (2017).

94 Martins, C., Dreij, K. & Costa, P. M. The State-of-the Art of Environmental Toxicogenomics: Challenges and Perspectives of "Omics" Approaches Directed to Toxicant Mixtures. *Int J Environ Res Public Health* **16**, doi:10.3390/ijerph16234718 (2019).

95 Brooks, B. W. *et al.* Toxicology Advances for 21st Century Chemical Pollution. *One Earth* **2**, 312-316, doi:10.1016/j.oneear.2020.04.007 (2020).

Figures

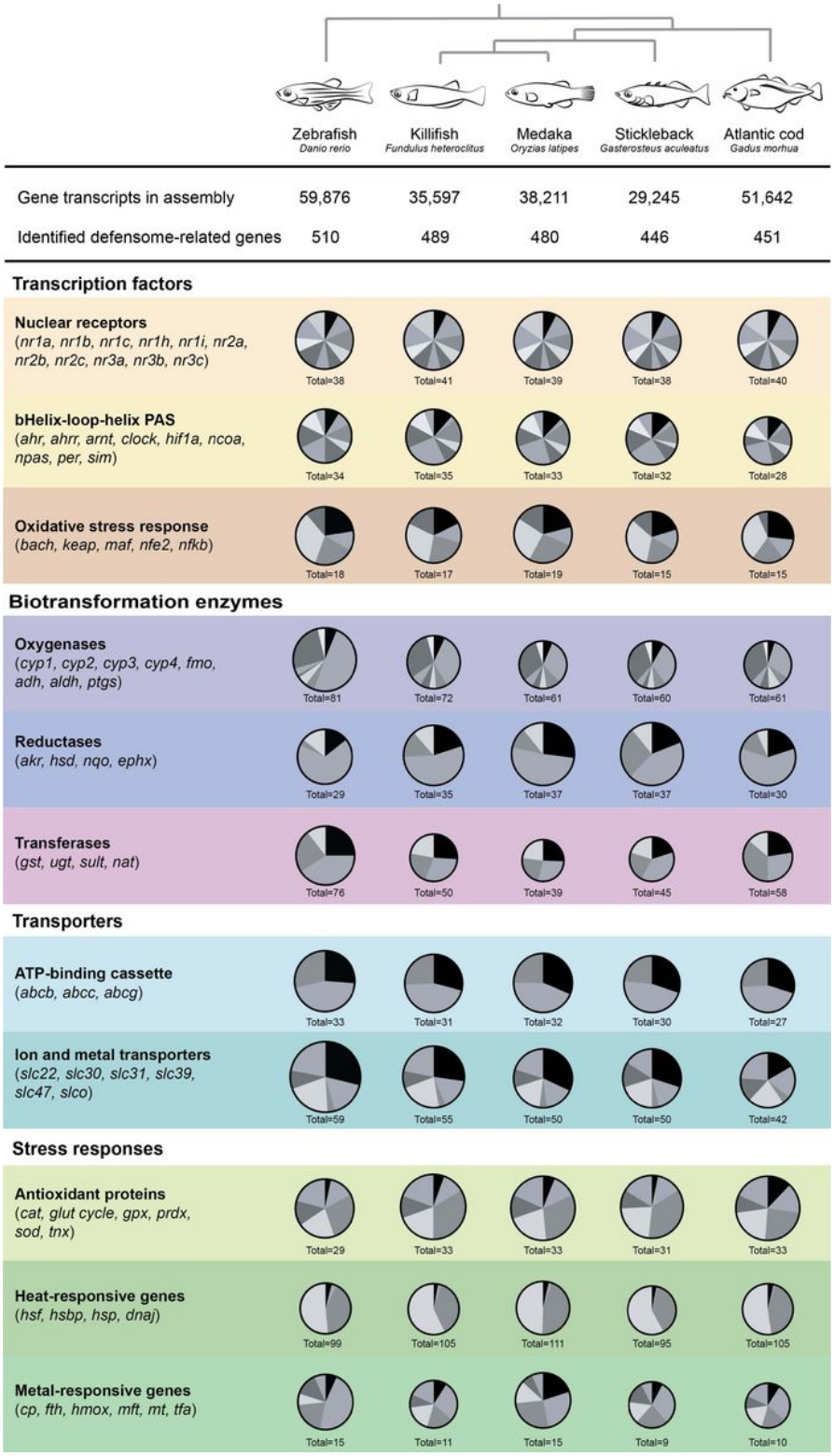

Figure 1

Chemical defensome genes in five model fish species. The genes were identified by searching gene names and using HMMER searches with Pfam profiles, followed by reciprocal or best-hit blast searches towards the zebrafish proteome. The gene families are organized in categories following Gene Ontology

annotations and grouped by their role in the chemical defensome. The size of the disk represents the relative number of genes in the different fish genomes within each group, with the number of genes in a specific gene family as slices.

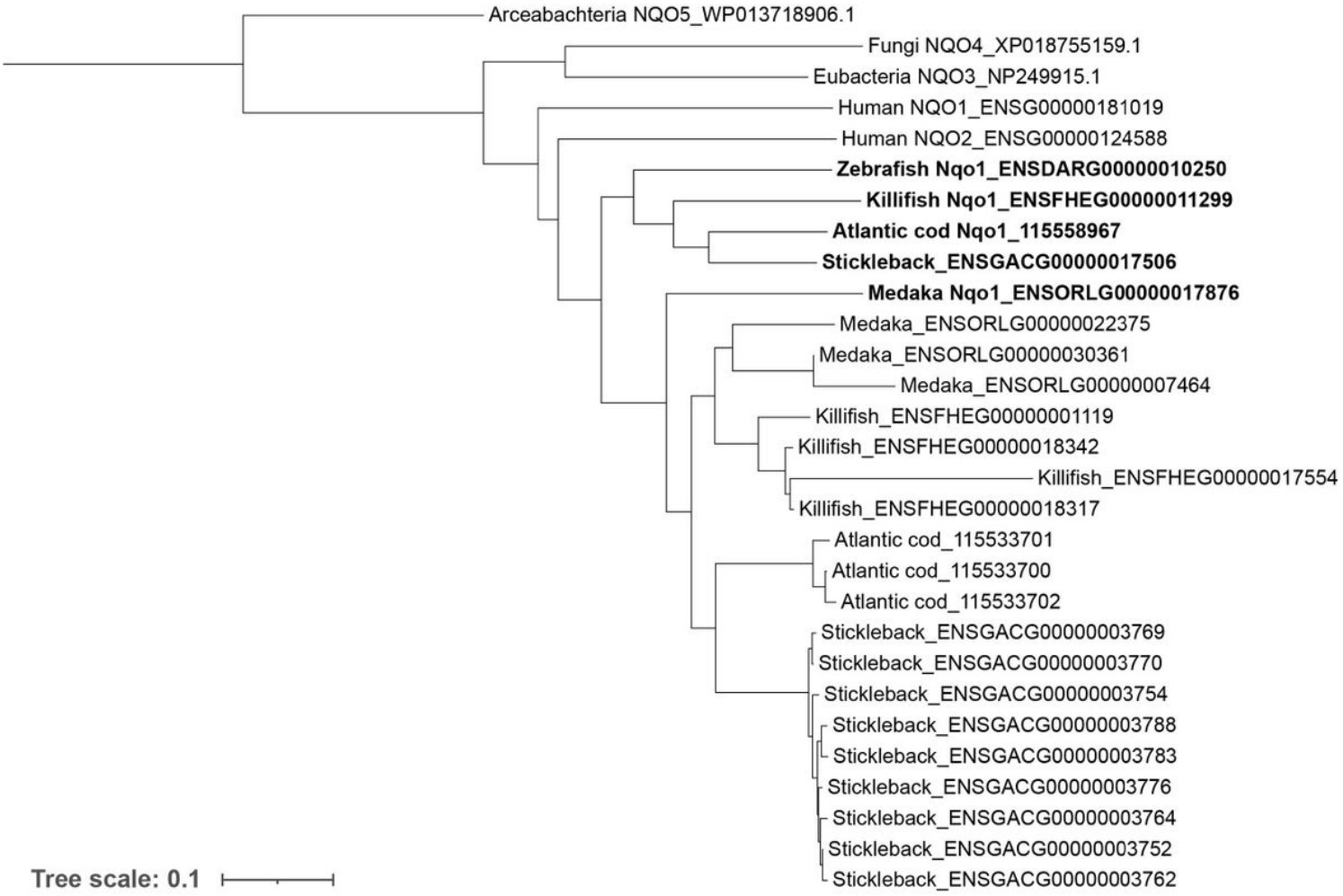

Figure 2

Phylogenetic tree of NAD(P)H:quinone oxidoreductases (NQO), also known as DT-diaphorase (DTD). Multiple sequence alignment and phylogenetic tree was built using Clustal Omega49 with standard settings. The tree was drawn using iTol50, and rooted with the archaebacterial NQO5.

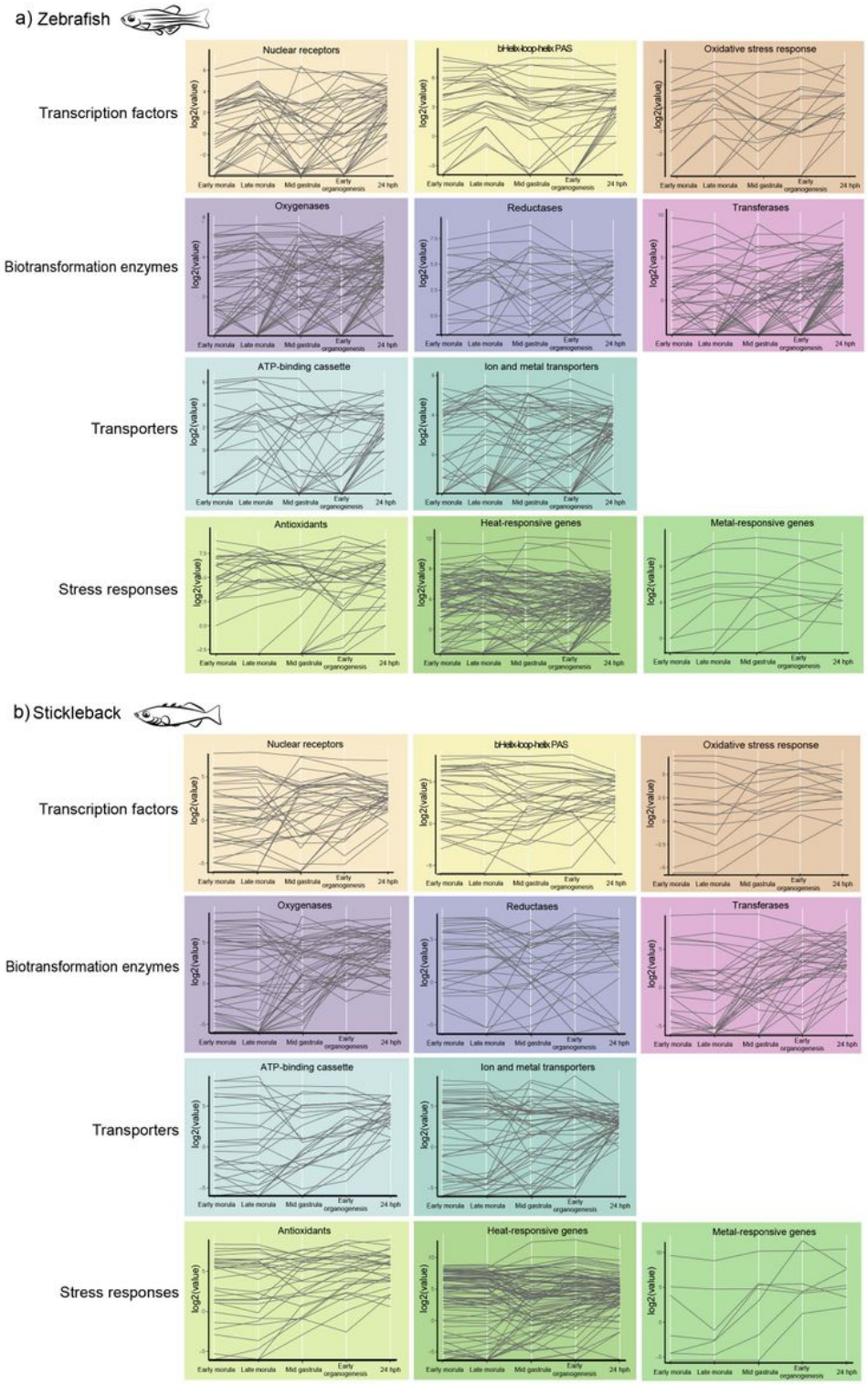

Figure 3

Transcription of chemical defensome genes in early development of a) zebrafish (Danio rerio) and b) stickleback (Gasterasteus aculeatus). Absolute transcription values (log2 scale) of defensome genes, grouped into their functional category, are shown at early morula, late morula, mid gastrula, early organogenesis, and 24 hours post hatching (hph).

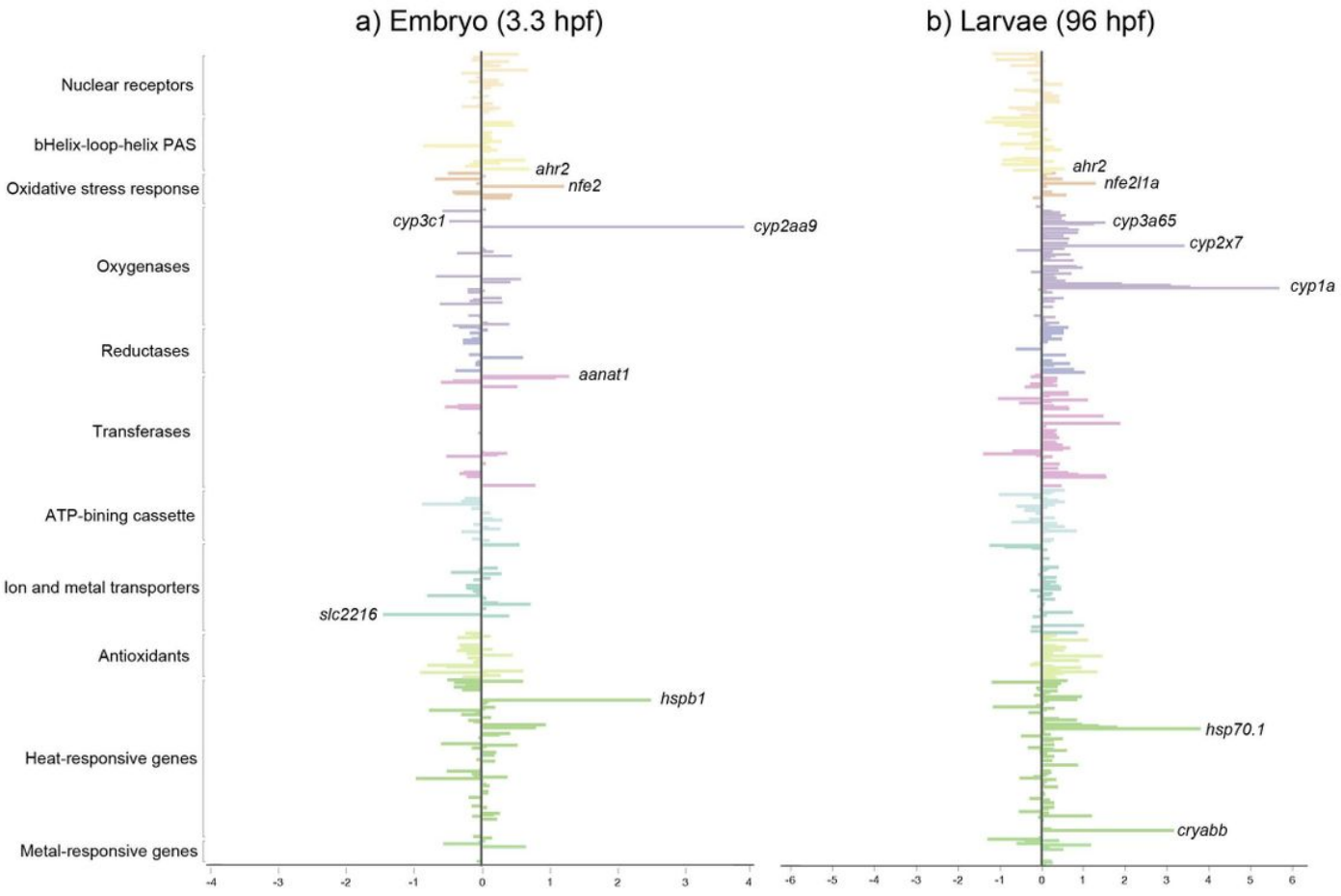

Figure 4

Transcriptional responses on chemical defensome genes in a) zebrafish embryo (3.3 hours post fertilization (hpf)) and b) zebrafish larvae (96 hpf) following exposure to benzo(a)pyrene. The transcription is shown as log2 fold change between exposed and control group at each timepoint. The genes are grouped into their functional categories in the chemical defensome and the name of some genes are indicated for clarity.