Aquatic Toxicology 186 (2017) 28–39

Contents lists available at ScienceDirect

Aquatic Toxicology

journal homepage: www.elsevier.com/locate/aquatox

Characterisation of the transcriptome of male and female wild-type guppy brains with RNA-Seq and consequences of exposure to the pharmaceutical pollutant, 17α-ethinyl estradiol

Minna Saaristo[a,b,*], Bob B.M. Wong[a], Laura Mincarelli[c], Allison Craig[c], Christopher P. Johnstone[a], Mayumi Allinson[d], Kai Lindström[b], John A. Craft[c]

[a] School of Biological Sciences, Monash University, 3800 Victoria, Australia
[b] Environmental and Marine Biology, Åbo Akademi University, 20500 Turku, Finland
[c] Life Sciences, Glasgow Caledonian University, Glasgow G4 0BA, UK
[d] Victorian Centre for Aquatic Pollution Identification and Management (CAPIM), School of Chemistry, 3010 Victoria, Australia

ARTICLE INFO

Article history:
Received 5 September 2016
Received in revised form 9 February 2017
Accepted 14 February 2017
Available online 16 February 2017

Keywords:
Brain
EE2
Endocrine disrupting chemicals
Next generation sequencing
Poecilia reticulata
Transcriptome

ABSTRACT

Waterways are increasingly being contaminated by chemical compounds that can disrupt the endocrinology of organisms. One such compound is 17α-ethinyl estradiol (EE2), a synthetic estrogen used in the contraceptive pill. Despite considerable research interest in the effects of EE2 on reproduction and gene expression, surprisingly, only a few studies have capitalised on technologies, such as next-generation sequencing (NGS), to uncover the molecular pathways related to EE2 exposure. Accordingly, using high-throughput sequencing technologies, the aim of our study was to explore the effects of EE2 on brain transcriptome in wild-type male and female guppy (*Poecilia reticulata*). We conducted two sets of experiments, where fish were exposed to EE2 (measured concentrations: 8 ng/L and 38 ng/L) in a flow-through system for 21 days. The effects on the brain transcriptome on both males and females were assessed using Illumina sequencing (MiSeq and HiSeq) platform followed by bioinformatics analysis (edgeR, DESeq2). Here, we report that exposure to EE2 caused both up- and downregulation of specific transcript abundances, and affected transcript abundance in a sex-specific manner. Specifically, we found 773 transcripts, of which 60 were male-specific, 61 female-specific and 285 treatment-specific. EE2 affected expression of 165 transcripts in males, with 88 downregulated and 77 upregulated, while in females, 120 transcripts were affected with 62 downregulated and 58 upregulated. Finally, RT-qPCR validation demonstrated that expression of transcripts related to transposable elements, neuroserpin and heat shock protein were significantly affected by EE2-exposure. Our study is the first to report brain transcriptome libraries for guppies exposed to EE2. Not only does our study provide a valuable resource, it offers insights into the mechanisms underlying the feminizing effects on the brains of organisms exposed to environmentally realistic concentrations of EE2.

1. Introduction

Aquatic pollution is a global problem. Waterways are increasingly being contaminated by chemical compounds, either natural or synthetic, that are known to disrupt the endocrine function of organisms, including humans. For example, over 800 chemicals in current commercial use, such as pharmaceuticals, pesticides, plasticizers and dioxins, are listed as endocrine disrupting chemicals (EDCs) due to their ability to interfere with hormone receptors, synthesis or conversion (Bergman et al., 2012). EDCs enter the environment via wastewater from domestic, industrial and agricultural sources. Even if they are present in the environment at very low levels (ng/L), laboratory and field studies have shown that EDCs can cause a wide range of physiological and morphological changes to exposed organisms (Colborn et al., 1993; Kidd et al., 2007, 2014) as well as population-level effects (Arnold et al., 2014). In humans, EDCs have been linked to the rise of endocrine-related disorders and diseases, such as prostate cancer (Skakkebaek et al., 2001; Diamanti-Kandarakis et al., 2009; Kortenkamp et al., 2011) Tools to detect early-warning signs of EDCs are therefore crucially needed. New technologies, such as next-generation sequencing (NGS), offer

* Corresponding author at: School of Biological Sciences, Monash University, Clayton campus, Bld 18, 3800 Victoria, Australia.
E-mail address: minna.saaristo@monash.edu (M. Saaristo).

http://dx.doi.org/10.1016/j.aquatox.2017.02.016
0166-445X/

promising avenues for investigating how wildlife – at the genome level – is affected. NGS can lead to pivotal insights into the effects of EDC exposure at the molecular and cellular levels, and help in building hypotheses for more detailed studies at the functional level (Davey et al., 2011; Fraser et al., 2011; Mehinto et al., 2012; Sharma et al., 2014).

To date, the impacts of estrogenic EDCs on gene expression have been well documented. For example, we know how EE2 affects gene expression in the gonad (Santos et al., 2007; Filby et al., 2007; Gardner et al., 2012; Miller et al., 2012), liver (Hoffmann et al., 2006; Filby et al., 2007; De Wit et al., 2010; Hultman et al., 2015) and brain (Martyniuk et al., 2006, 2010; Vosges et al., 2010; Harding et al., 2013) in fish. Although studies have demonstrated that the neuroendocrine transcriptome can be affected by estrogens (e.g. Martyniuk et al., 2006, 2010; Vosges et al., 2010), most of the studies published so far have used a microarray approach. Microarrays have several potential drawbacks. They are known for 'design bias' (Ioannidis et al., 2009), are heavily dependent on *a priori* knowledge of the genome (Hurd and Nelson, 2009) and, due to high signal to noise ratios, have poor detection of low-abundance sequences (Hurd and Nelson, 2009). By contrast, high-throughput cDNA sequencing (RNA-Seq) offers sequencing at the full-genome depth of the transcriptome without requiring any *a priori* knowledge about sequences, allowing for the analysis to detect rare sequences, alternative splicing events, and noncoding RNAs, in complex genomic samples (Pan et al., 2008; Hurd and Nelson, 2009; Anders et al., 2012).

One common pharmaceutical of global concern is 17α-ethinyl estradiol (EE2), a key active ingredient in the contraceptive pill that has been detected from sewage effluent and surface waters (<1–15 ng/L) around the world (Baronti et al., 2000; Muller et al., 2008; Miege et al., 2009; Valdes et al., 2015). Over 99% of the estrogenic activity in waterways is attributable to the presence of EE2 and 17β-estradiol (E2) due to insufficient removal of these chemical compounds during wastewater treatment (Clouzot et al., 2008; Rojas et al., 2013). EE2 exposure at concentrations of only a few ng/L has been shown to cause severe impairments in the reproduction of fish, such as sex reversal (Lange et al., 2009), reduced survival and fecundity through transgenerational effects (Schwindt et al., 2014).

Guppies (*Poecilia reticulata*) are a promising non-model species for studying genomic responses to aquatic pollutants, such as EE2. The guppy is a small, sexually dimorphic livebearing fish, native to north-eastern South America. Guppies are a well-studied model species in behavioural ecology, but most importantly, the species has been introduced into tropical regions around the world and can be found in a variety of habitats, including environments close to human habitation that can be heavily polluted by anthropogenic chemicals (Widianarko et al., 2000; Araújo et al., 2009).

Little is known about how EDC-induced changes in the brain transcriptome correlate with changes at the translational level in the brain. Even fewer studies have used next generation sequencing to unravel impacts of EDCs on brain tissues of non-model organisms. To fill this knowledge gap, our aim was to provide a database-description of the brain transcriptome of EE2-exposed and control fish of both sexes in wild-type guppies. Specifically, our objectives in this work were threefold: (1) to create four reference transcriptome assemblies for the brain of both sexes for control animals and those exposed to an acute level of EE2; (2) to identify sex-specific and treatment-specific gene profile differences with an independent group of fish exposed to a lower, environmentally relevant concentration of EE2 to unravel the different mechanisms of action of EE2 in a fish brain, and (3) to design gene-specific qPCR assays for independent study of differential expression of sex- and treatment-specific genes of the brain. In achieving these aims, an overarching goal of the study was to provide a valuable genomic resource to explore the molecular underpinnings of organismal responses to human-induced environmental change.

2. Material and methods

2.1. Collection and housing

The guppies used in our study came from a laboratory population (Prof John Endler, Deakin University, Australia), descended from wild caught fish inhabiting Alligator Creek (19°26′17.94"S, 146°57′1.09"E), Queensland, Australia. Fish were acclimated to laboratory conditions (+26 °C, 12:12 h light:dark regime) for 2 months in 54 L tanks (20 fish per tank) prior to exposure. Fish were fed ad libitum once daily with commercial fish pellets (Otohime Hirame).

2.2. Exposure design

After acclimation, fish were exposed to EE2 for 21 days, as previous experiments have shown that exposure periods ranging from 4 to 28 days are sufficient to induce molecular, including hepatic transcriptomic, physiological and behavioural changes (Bayley et al., 1999; Bell 2001; Martinovic et al., 2007; Saaristo et al., 2009a, 2009b; Katsiadaki et al., 2010; Prokkola et al., 2015).

In the first experiment, guppies were exposed to EE2 via a flow-through system using the design of Saaristo et al. (2013) in July 2013. Briefly, this system included 8 exposure tanks, and was made up of 4 control tanks (a solvent control: 0.00004% ethanol) and 4 exposure tanks (EE2: measured mean concentration of 37 ng/L, SE = 4.05, n = 11 water samples). During the exposure period, water samples were taken weekly from the exposure tanks and measurement of EE2 was conducted in accordance with the manufacturer's instructions (Ecologiena, EE2 ELISA Kit, Tokiwa Chemical Industries, Japan) with minor modifications (Saaristo et al., 2013). A total of 224 fish (112 males, 112 females) were randomly taken from the holding tanks and placed into the exposure tanks, with the sexes kept separately (60 × 30 × 24 cm; 28 fish per tank). All exposure tanks were maintained at a temperature of 25–27 °C, on a 12:12 h light:dark regime, with fish fed ad libitum once a day during the exposure period.

To identify sex-specific and treatment-specific gene profile differences, we conducted a second experiment, where an independent group of wild caught fish was exposed to an environmentally relevant concentration of EE2. This second exposure was conducted as above and included 16 exposure tanks, which were made up of 8 control tanks (a solvent control: 0.00004% ethanol) and 8 exposure tanks (EE2: measured mean concentration of 8 ng/L, SE = 1.13, n = 22). A total of 448 fish (224 males, 224 females) were randomly taken from the holding tanks and placed into the exposure tanks, with the sexes kept separately (60 × 30 × 24 cm; 28 fish per tank). As with the first experiment, all exposure tanks were maintained at a temperature of 25–27 °C on a 12:12 h light:dark regime, and fish were fed ad libitum once a day during the exposure period.

After 21 days, males and females from both experimental exposures were anesthetized using clove oil. The weight and length of the individuals were measured. Fish were killed by cutting the spinal cord, and brain tissue was excised under a stereomicroscope (Motic SMZ-168). The dissected brain tissue was placed in RNALater (Sigma Aldrich), stored at room temperature for the first 24 h, then at − 4 °C for 24 h, and finally placed in a − 80 °C freezer until RNA isolation.

2.3. RNA isolation and quality assessment

Previous work reported that the brain tissue weight of male and female guppies are 0.55 – 0.75 mg and 0.30–0.50 mg, respectively (Kotrschal et al., 2013). Total RNA was isolated from the whole brain

tissue using a combination of the Trizol method and the NucleoSpin RNA II kit (Marchery-Nagel, Germany) following the protocol of Saaristo et al. (2009a, 2010) with slight modifications. RNA isolation and quality were determined (Nanodrop ND-1000, the Agilent Bioanalyzer) and the concentrations varied between 15 and 90 ng/uL, 260/280 ratios were >1.8, and RIN were >7.

2.4. Preparation of normalised cDNA libraries

To maximise the chances of including transcripts with low expression in the transcriptome, normalised cDNA libraries were prepared for future sequencing campaigns. To achieve this, we used fish from the first experiment (high EE2 exposure level to ensure an effect) and took six RNA samples from each treatment combination (i.e. control female, control male, EE2-exposed female, EE2-exposed male) and pooled them for constructing the cDNA libraries. Specifically, a pool of total RNA for each group was prepared by taking equal amounts of RNA from six individuals from within each treatment group. Each of the four pools was then used for cDNA synthesis and preparation of normalised libraries conducted by Evrogen (Moscow, Russia) (Zhu et al., 2001; Shagin et al., 2002; Zhulidov et al., 2004). Total RNA was used for double-stranded cDNA synthesis using the Spliceosome-Mediated RNA *trans*-splicing (SMaRT) approach (Zhu et al., 2001). SMaRT-prepared amplified cDNA samples were then normalized using the duplex-specific nuclease (DSN) normalization method (Zhulidov et al., 2004). Normalization included cDNA denaturation/reassociation, treatment by duplex-specific nuclease (Shagin et al., 2002) and amplification of the normalized fraction by PCR. The following acronyms will be used to describe the four libraries: FC = control female, FE = EE2-exposed female, MC = control male, ME = EE2-exposed male.

2.5. RNA-Seq library construction and sequencing

For the first experiment, sequencing of the normalised cDNA libraries was conducted at GATC Gmbh (Konstanz, Germany). Sequencing library preparation was performed with TruSeq mRNA-Seq Sample Prep kit (Illumina) on normalised cDNA (see details above). Sequencing was performed on the Illumina MiSeq 300 bp paired end (read length = 2 × 300 bp) mode using MiSeq Reagent Kit v3 and software version MCS 2.4.1. Samples were barcoded using the recommended Illumina index sequences, pooled and sequenced in a single run.

For the second experiment, sequencing of the RNA samples, the library preparation and sequencing was conducted at LGC Genomics (Berlin, Germany). Specifically, we used RNA samples from individual fish (3x control females, 3x EE2-females, 3x control males and 3x EE2-males), to create biological replicates rather than pooling the samples as in the first experiment. Library preparation was performed with the TruSeq Stranded mRNA Library Preparation kit (Illumina), using oligodT mRNA enrichment. In addition, individual libraries were bar-coded. Sequencing was carried out on the Illumina HiSeq 2000 platform with 100 bp paired-end (read length = 2 × 100 bp) strand-specific reads and 6 samples per channel.

2.6. Illumina data analysis

For the first experiment, Illumina MiSeq raw data (paired end) were processed through the Casava Pipeline from Illumina, where the data was demultiplexed and quality filtered with default settings, and reads stitched by FLASH (v1.2.6 software). The trimmed reads from four original cDNA libraries (control female, EE2-exposed female, control male, EE2-exposed male) were assembled and mapping statistics generated *de novo* at GATC using CLCbio pipelines (CLCbio, Qiagen). The resulting contiguous sequences (contigs) were annotated, for each library, using FastAnnotator (Chen et al., 2012). Annotated transcripts, generated by FastAnnotator, were further analysed in Blast2Go (Conesa et al., 2005). The data were analysed using a Fisher's Exact Test comparing males with females, EE2-exposed males with control males, and EE2-exposed females with control female.

For the second experiment, the four separate transcriptome assemblies corresponding to the four experimental groups as produced by the CLCbio pipeline were unified into a single reference assembly using the bowtie alignment tools in Trinity (v2.0.6) (Grabherr et al., 2011). This assembly was used to identify differentially expressed transcripts as a result of the second experiment.

2.7. Differential gene expression analysis

To identify genes that might be expressed in a sex- or treatment-specific manner, we used the Differential Gene Expression Analysis (DE) tools in Trinity (v2.0.6) (Grabherr et al., 2011). Specifically, the abundance of reads in the fish from the second experiment (individual fish samples), were determined against the unified transcriptome assembly obtained from the first exposure (pooled samples). Read counts for the demultiplexed and quality trimmed data analysed for DE are presented in the Supplementary Table 1. The transcripts that differed significantly between the groups were identified with edgeR (Robinson et al., 2010), with settings of a minimum of 1.5-fold change in abundance and FDR at $p < 0.05$ for pair-wise comparisons. edgeR assesses abundance by comparing one group to another and then identifies if an individual transcript is at higher or lower abundance in one relative to the other. Transcripts thus identified were extracted from the unified assembly and annotated with BLAST2GO (Conesa et al., 2005). Independently, differentially expressed transcripts were identified with a second R package, DESeq2 (Anders and Huber 2010), with the same filter criteria as used for edgeR. This conservative approach – only the genes that are identified by both edgeR and DESeq2 were further analyzed – was used to further ensure the exclusion of false positives (Yendrek et al., 2012; Gunter et al., 2013). To find out the degree of association between the DE transcripts in each group, transcripts were further explored with hierarchical clustering (Eisen et al., 1998) and clusters with similar expression profiles were extracted with Trinity.

2.8. Quantitative real-time PCR

This part of the project was done in parallel with the sequencing campaign of the first experiment. Accordingly, a separate group of individual fish (n = 8 for each sex and treatment: n = 8 control females, n = 8 EE2-exposed female, n = 8 control male, n = 8 EE2-exposed male; total = 32 fish) from the first experiment was used in the final qPCR. Based on previous research, genes that were likely to be targets of EE2 exposure, such as estrogen receptors, hormone metabolism and neurotransmitters and their receptors, were chosen as candidate genes (Martyniuk et al., 2006; Vosges et al., 2010; Harding et al., 2013; Nikoleris et al., 2016). A group of 10 candidate genes spanning transcripts associated with non-reproductive to reproductive behaviours, steroid metabolism, transcription factors, and house-keeping genes were selected (Supplementary Table 2). The sequence of these genes was derived from our normalized library transcriptome, and compared to homologues (blastx) to ensure the annotation was correct and then used to design potential primers for qPCR. The primers were all tested by an end-point PCR and the chosen 10 candidate genes were used for final assessment of transcript abundance in the guppy brain.

2.9. Statistical analysis

2.9.1. Differential expression analysis

For each pair of the comparisons, two treatment groups (sex or treatment) were first analyzed and then we investigated which of the gene transcripts showed increased levels in abundance and which decreased. For example, pair comparison of MCvsME compared control male (MC) gene transcripts to EE2-exposed male (ME) transcripts. Further, when the differential expression analysis gave an outcome of MCvsME-MC+ for a specific gene transcript, it meant that the abundance of that gene was repressed in the EE2-exposed male group relative to the control male group. On the other hand, an outcome of MCvME-ME+ indicated that the gene transcript abundance was induced in the EE2-exposed male group.

2.9.2. Quantitative real-time PCR

Specifically, the level of expression for each target gene analyzed by qPCR was determined by the comparative quantification method $2^{-\Delta\Delta CT}$(Livak and Schmittgen, 2001), with a modification of the presentation of real-time PCR data as normalized individual data points (Schmittgen and Livak, 2008). For each fish and each gene assay, the expression levels of target genes were average values of the three replicate wells, and normalized with the expressions of one of the housekeeping genes (n = 8–10 for each individual RNA sample for each sex, each exposure treatment (i.e. control versus EE2-exposed) and each gene assay). First, the count data, which was the average value of three technical replicates/sample, was normalized, with the expression of the housekeeping genes checked for normality and heterogeneity of variance. We then used multiple models (structural equation modelling (SEM) (using R library 'sem'), random forest ('cforest' in R library 'party'), GLM ('glm' in R core)), to test for differences in gene expression among the treatment groups. Multiple models were used because the different methods have strengths and weaknesses, and we wished to look for areas of agreement among the analyses. Initially both length and weight were included in the models. Because length and weight (or interaction terms) had no significant effect on the response variables, they were excluded from all final models. We used Mann-Whitney *U* tests to compare treatment differences. Use of a non-parametric test instead of *t*-tests can slightly increase the risk of Type II error, but we consider this acceptable, as non-parametric tests represent a more cautious approach to significance testing overall (Krzywinski and Altman, 2014). Finally, in order to identify the gene expression profiles of the treatment groups, the differently expressed genes were analyzed using principal component analysis (PCA) ('prcomp' in R core). Here it is prudent to acknowledge that PCAs on datasets of this nature – the number of observations and/or variables is relatively small – need to be treated with care. All statistical analyses were performed using RStudio (0.99.467).

2.10. Ethical statement

The methods for animal housing, handling and experimental protocols were assessed and approved by the Biological Sciences Animal Ethics Committee at Monash University (permit number: BSCI/2011/07) and complies with all relevant State and Federal laws of Australia.

2.11. Statement on data availability

The Illumina-MiSeq reads reported in this study have been deposited in the National Center for Biotechnology Information Short Reads Archive, www.ncbi.nlm.nih.gov/sra (SRA accession ID: SRP076083) and The Illumina-HiSeq reads have been deposited in the GeneBank (accession ID: DDBJ/EMBL/GenBank/GFHH01000000).

3. Results

3.1. Illumina RNA-Seq analysis

For the first experiment, we obtained 1.2-1.9 M quality-trimmed Illumina MiSeq reads for each library from the pooled brain tissues (four libraries: FC = control female, FE = EE2-exposed female, MC = control male, ME = EE2-exposed male) (Supplementary Table 1). The *de novo* assembly of all separate libraries consisted of 147,106,348 bp and 254,666 contigs (N50 values 832–864 bp), which was reduced to 166,749 contigs following clustering at 99% identity with CD-HIT. All four libraries were annotated through FastAnnotator using blastx and further annotated with Gene Ontology (GO) terms for biological processes, cellular component and molecular function. These descriptive terms and their relative frequency at level two have been visualized in Supplementary Fig. 1. No significant differences are apparent from these terms for each library or the scores recorded for those annotations (Fisher's Exact Test: $p > 0.05$).

For the second experiment, we obtained 342 M quality-trimmed Illumina HiSeq paired reads from the individual brain tissues (3x control female, 3x EE2-exposed female, 3x control male, 3x EE2-exposed male). A unified reference assembly from the first experiment was produced with Trinity from the four individual assemblies (control female, EE2-exposed female, control male, EE2-exposed male). This produced 134,405 transcripts (N50 value 1295 bp) and this was subsequently used as the reference for the DE transcript analysis.

3.2. Differential gene expression analysis

Differential gene expression analysis tools edgeR and DESeq2 both identified the same set of genes and with similar levels of fold-change. Specifically, with the pair-wise comparisons for six combinations (control female vs EE2-exposed female (FC vs FE); control male vs EE2-exposed male (MC vs ME); control female vs control male (FC vs MC); EE2-exposed female vs EE2-exposed male (FE vs ME); control female vs EE2-exposed male (FC vs ME); EE2-exposed female vs control male (FE vs MC)), we identified 773 gene transcripts with a fold change of at least 1.5 fold with confidence for FDR at $p < 0.05$. Included in this total were 60 transcripts that were male-specific (MC > FC), 61 that were female-specific (FC > MC), and 285 that were treatment-specific. Specifically, in females, EE2 affected expression of 120 transcripts, with 62 downregulated and 58 upregulated (Fig. 1A). In males, on the other hand, 165 transcripts were affected, with 88 downregulated and 77 upregulated (Fig. 1B). The other 367 transcripts were identified in comparisons including sex and treatment. Furthermore, out of the 773 transcripts, 49 transcripts were found to be common to two or more of the possible six combinations (Supplementary Table 3). Thus, it is apparent that the EE2-treatment caused both up- and down-regulation of specific transcript abundances and that EE2 appears to affect transcript abundance in a sex-specific manner (Supplementary Table 3).

To more effectively visualize the effect of EE2-exposure, PCA was performed to examine the level of similarity/dissimilarity in the gene expression profiles of different groups. PCA explained ~43% of the variability in the first two factors (Eigen values >186) and demonstrated a relatively high level of dissimilarity between the treatment groups (Fig. 2). Specifically, there was a clear shift of EE2-exposed females away from control females and towards control and EE2-exposed males on PCA2, while gene transcripts of EE2-

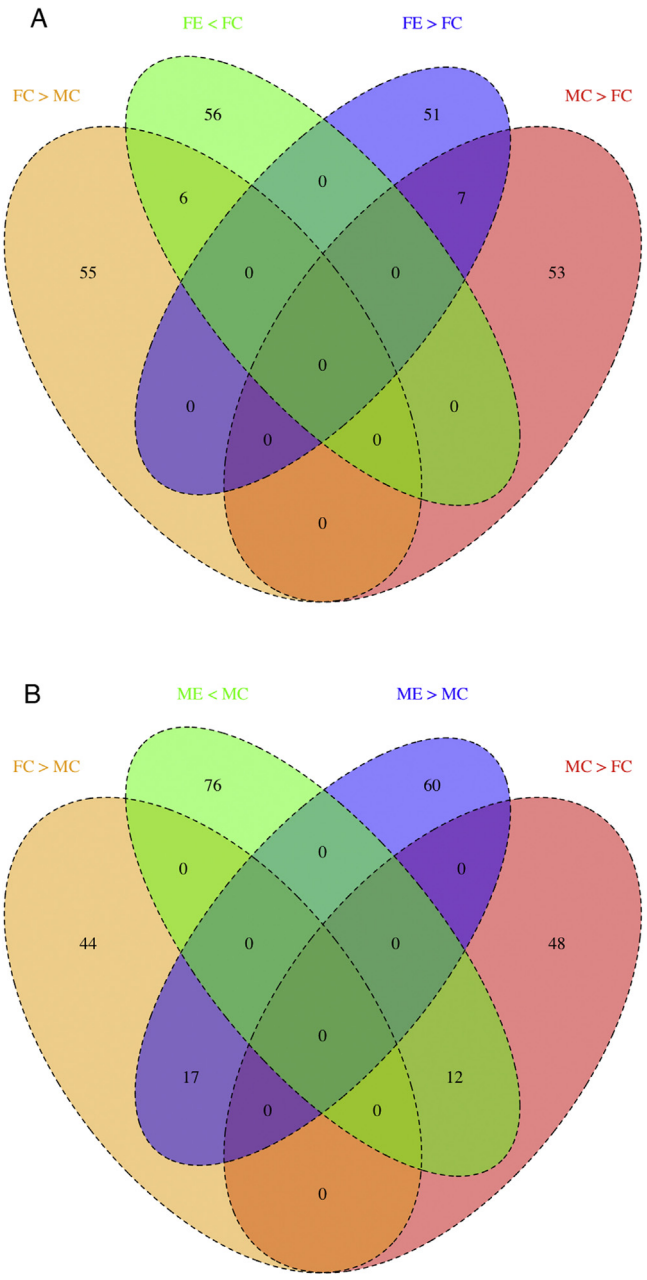

Fig. 1. EE2-exposure induced differential transcript abundance in guppy brain. (A) Responses of females to EE2-exposure with male and female controls for comparison. FC > MC = transcripts that were more abundant in control females than control males. MC > FC = transcripts that were more abundant in control males than control females. FE < FC = transcripts that were less abundant in EE2-exposed females than control females. FE > FC = transcripts that were more abundant in EE2-exposed females than control females; (B) Responses of males to EE2-exposure with male and female controls for comparison. FC > MC = transcripts that were more abundant in control females than control males. MC > FC = transcripts that were more abundant in control males than control females. ME < MC = transcripts that were less abundant in EE2-exposed males than control males. ME > MC = transcripts that were more abundant in EE2-exposed males than control males. Note that both diagrams are based only on the subset of transcripts and comparisons such as EE2-exposed females to EE2-exposed males have been excluded.

males had shifted away from control males and appear to have 'over-shot' control female expression on PCA1 (Fig. 2). This segregation of gene transcripts into clusters by treatment groups (FC, FE, MC, ME) was further confirmed with hierarchical cluster analysis (Fig. 3). The resulting heat-map revealed distinctive patterns of transcript abundance for each of the four groups of fish (Fig. 3). Specifically, the hierarchical tree consisted of 24 distinct clusters of which 12 contained 76% of the total transcripts (>25 transcripts per cluster) (Supplementary Table 4). Visual inspection of the heatmap (Fig. 3) and alignment of the transcripts in each cluster with the DE results identified groups of transcripts associated with specific responses to EE2 and sex specificity. For instance, transcripts in cluster 1 were associated with downregulation of abundance in male, in cluster 2 with downregulation in both male and female, while in cluster 3 transcripts were upregulated in both male and female.

3.3. Annotation results from the normalized libraries (first experiment)

Our model compound, EE2, is known to bind to estrogen receptors (ER) with a greater affinity than natural estrogen 17β-estradiol and, by so doing, control the transcription of ER-responsive genes. Annotated genes in this study contained relatively few contigs (5) for ER in the different libraries, however, genes encoding many of the enzymes required for the synthesis of steroid hormones were present (Supplementary Table 5). In addition, the annotated transcripts included those of genes encoding receptors for neurotransmitters, but mainly receptors associated with excitatory circuits, such as glutamate receptors (69 contigs) (Supplementary Table 5). Finally, the presence of an abundance of annotated contigs described as "reverse transcriptase" prompted a search of each library for targets relating to transposable elements. The search identified 144 contigs with representatives found in each of the libraries.

3.4. qPCR analysis

According to structural equation modelling, four target genes out of the 10 tested (listed in Methods) were differentially expressed between the treatment groups using each of the house keeping genes as reference. These genes were: a gene encoding a heat shock protein, transcripts related to transposable elements (2 contigs) and the gene encoding neuroserpin. In exposed fish, expression of the gene encoding heat shock protein (Mann-Whitney: $W = 16$, $p = 0.001$, $n = 32$) (Fig. 4A) and transcripts related to transposable elements (RT1: Mann-Whitney: $W = 25$, $p = 0.011$, $n = 32$, RT3: Mann-Whitney: $W = 29$, $p = 0.022$, $n = 32$) were upregulated compared to control fish (Fig. 4B,C). The gene encoding neuroserpin, however, was down-regulated in EE2-exposed fish compared to controls (Mann-Whitney: $W = 102$, $p = 0.027$, $n = 32$)(Fig. 4D). Interestingly, transcripts related to transposable elements were the only target genes, which showed significant sex-specific effects. Expression of these genes was higher in females compared to males (RT1: Mann-Whitney: $W = 100$, $p = 0.037$, $n = 32$; RT3: Mann-Whitney: $W = 109$, $p = 0.007$, $n = 32$) (Fig. 4B,C).

With the selected four candidate genes, we ran a PCA analysis. The first two factors (PC1, PC2) explained ~50% of the variability (Eigen values >1). The first axis was contributed by all gene expressions, but was dominated by transcripts related to transposable elements (loading = 0.461), neuroserpin (NEU) (loading = −0.432) and heat shock protein (hsp70) (loading = 0.436) (Fig. 5). Thus, expression of transcripts related to transposable elements and heat shock protein were positively correlated, and both were negatively correlated with neuroserpin. Males of both treatments had lower values of PC1 than did females. Control males had a lower average PC1 than exposed males and the same pattern was true for females (Fig. 5).

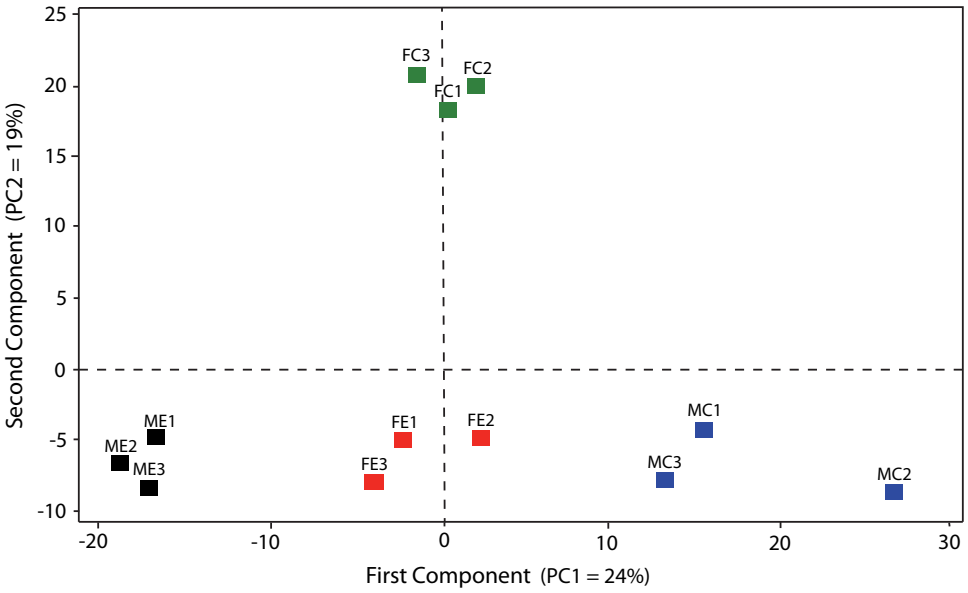

Fig. 2. Principal component analysis (PCA) of the 773 gene transcript pairs identified with Differential Expression analysis in the fish from the second exposure. The abundance of reads in these fish was determined against the normalized transcriptome assembly and PCA ran to more effectively visualize the effect of EE2-exposure. Each square represents the mean score per fish. FC1 = control female 1, FC2 = control female 2, FC3 = control female 3, FE1 = EE2-exposed male 1, FE2 = EE2-exposed female 2, FE3 = EE2-exposed female 3, MC1 = control male 1, MC2 = control male 2, MC3 = control male 3, ME1 = EE2-exposed male 1, ME2 = EE2-exposed male 2, ME3 = EE2-exposed male 3.

4. Discussion

The effects of the pharmaceutical EE2 on reproduction in mammals and fish have been well documented. Only a limited number of studies, however, have used next generation sequencing to uncover the molecular pathways related to exposure to EE2. Here, we report the results of the first transcriptome-wide gene expression analysis of brain tissue of wild-type guppies after exposure to EE2. Specifically, we conducted two independent experiments and exposed both male and female guppies to two levels of EE2 (low = 8 ng/L and high = 37 ng/L). The first experiment produced four brain transcriptome libraries (control female (FC), EE2-exposed female (FE), control male (MC), EE2-exposed male (ME)) based on pooled samples, and the second experiment enabled us to validate the transcriptome libraries against control and EE2-exposed individual fish from both sexes. Using this approach (pooled vs individual fish samples), we were able to assess the differential transcript abundance of transcripts not previously studied. To go further in our interpretation (using DE analysis), we explored how EE2 exposure impacted all six different permutations (FC vs FE, MC vs ME, FC vs MC, FC vs ME, FE vs MC, FE vs ME) and found exposure to EE2 caused both up- and downregulation of specific transcripts and affects transcript abundance in a sex-specific manner. Results of our study provide evidence that EE2, even after a short-term exposure, affects the gene expression relevant for steroid metabolism and signalling pathways of brain, but most importantly, shows that EE2 is feminizing the male brain transcriptome.

4.1. Differential expression by sex and EE2-exposure

Toxicology due to EE2 will result from a disruption of the normal interactions between gonadally produced E2 and the brain influencing behaviour and reproductive processes. Based on the DE analysis, we found both treatment and sex-specific effects. Specifically, we discovered gene transcripts that were either induced or repressed due to EE2-exposure. For example, in EE2-exposed females, exposure resulted in upregulation of genes encoding for proteins with functions related to cytochrome P450, T cell differentiation, stress-induced aggregation of blood plasma proteins and steroidogenesis, while genes encoding functions linked to E2 signalling and organization of synaptic active zones were downregulated. In EE2-exposed males, on the other hand, exposure caused upregulation of genes encoding oxidative stress resistance, ion channel proteins and pre-B-cell leukemia transcription factors, while down-regulating E3-ubiquitin-protein ligase activity, histone methyltransferase activity and transmembrane receptor protein tyrosine activity. We also found gene transcripts that were sex-specific. In females, for example, we discovered upregulation, relative to males, of genes that encode cell proteins regulating synaptic signal transmission, that contribute to oxidative stress resistance, as well as genes that encode proteins that play an important role in cascades of cellular responses evoked by extracellular stimuli such as physical stress. In males, on the other hand, we found genes that encode putative steroid 6-beta hydroxylase, genes that mediate molecular chaperones in response to stress, and genes that play a role in regulating cell growth, cell motility, immune response and axon guidance. Next, we highlight selected key areas of impact due to EE2-exposure based on the gene transcripts achieved with DE analysis.

4.1.1. EE2 effects on steroid metabolism

Cytochromes P450 (CYP) are the major enzymes involved in detoxification of a wide range of exogenous compounds and metabolism of endogenous substrates (Nebert et al., 2004). In particular, subfamily CYP3A has been shown to catalyse the breakdown of lipophilic xenobiotics, such as pharmaceuticals and endogenous steroids (Thummel and Wilkinson, 1998; Guengerich, 1999; Nebert and Russell, 2002). In fish, expression of the CYP3A gene has been reported in several organs, including brain (Hegelund and Celander, 2003; Arukwe, 2005). In the present study, we found upregulation of one isoform of *CYP3A* (the gene encoding cytochrome P450 3A56) in EE2-exposed females, which was also upregulated in control males relative to control females. Another isoform of *CYP3A*, on the other hand, was more abundant in control females relative to control males. Recent studies investigating the impact of EE2 on *CYP3A* gene expression have yielded contrasting results.

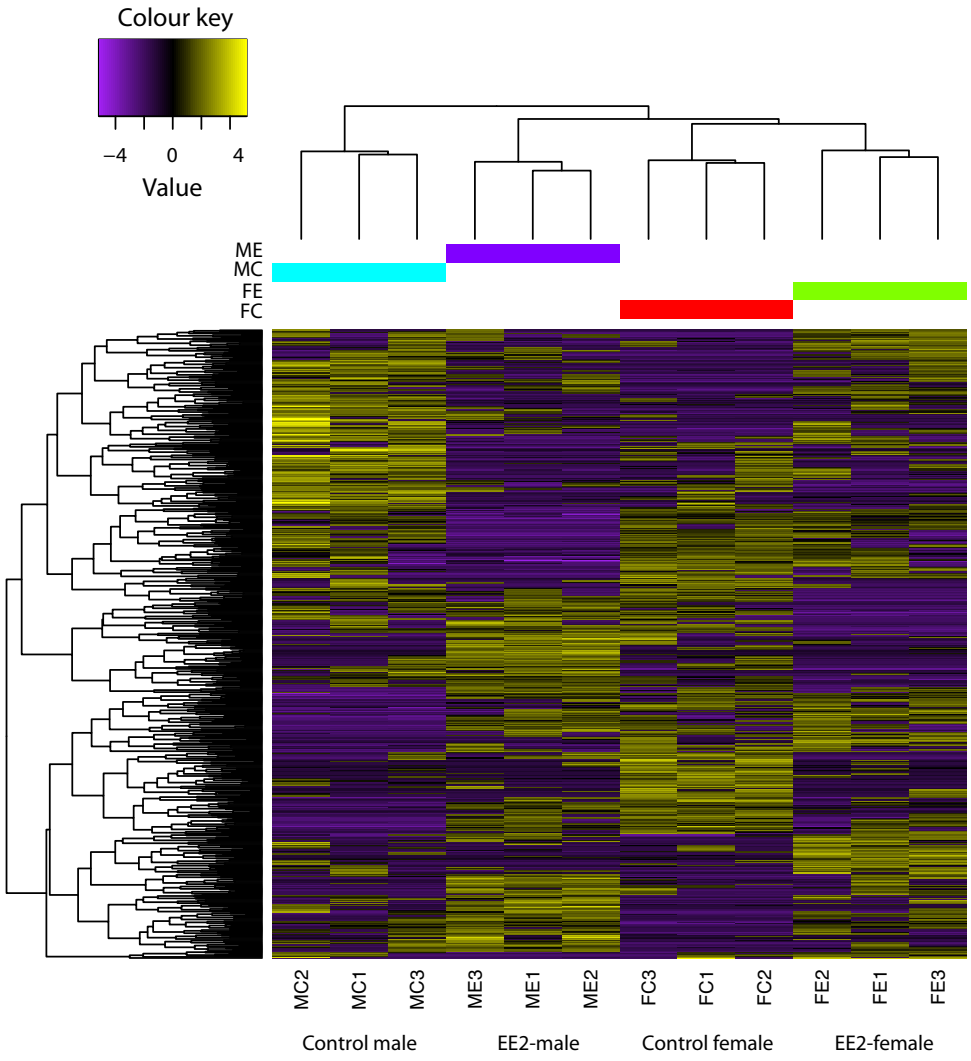

Fig. 3. Clustering analysis of differentially expressed transcripts. The heat map and dendrogram presents 773 gene transcript pairs present in the fish from the second exposure based on the cluster analysis. The rows represent expression levels of single transcripts and the columns represent treatment groups (FC = control female, FE = EE2-exposed female, MC = control male, ME = EE2-exposed male). The colour key presents the level of fold change between the individual gene transcript pairs. Relative expression of transcripts (rows) is displayed across individuals (columns) with yellow representing upregulation and purple representing downregulation. FC1 = control female 1, FC2 = control female 2, FC3 = control female 3, FE1 = EE2-exposed female 1, FE2 = EE2-exposed female 2, FE3 = EE2-exposed female 3, MC1 = control male 1, MC2 = control male 2, MC3 = control male 3, ME1 = EE2-exposed male 1, ME2 = EE2-exposed male 2, ME3 = EE2-exposed male 3.

For example, short-term exposure to EE2 suppressed CYP3A production in some studies (medaka, *Oryzias latipes*: Kashiwada et al., 2007; atlantic cod, *Gadus morhua*: Bizarro et al., 2016), but not in others (carp, *Cyprinus carpio*: Solé et al., 2003). There are several potential reasons for these differences. First, studies have quantified CYP3A production from different tissues (liver: Kashiwada et al., 2007; Bizarro et al., 2016), but also from plasma samples (Solé et al., 2003). Another possibility is that quantification at the transcription (mRNA) and translation (protein) level is very likely to produce different results. It is not clear which isoform(s) were being measured in these studies since our results show that these were differentially expressed and affected. Finally, the exposure route (injection vs water exposure) combined with different exposure concentrations is inclined to trigger different responses. Even if CYP plays an important role in the detoxification of a wider range of pharmaceuticals and protecting the body against the adverse effects of these chemicals, significant induction of CYP enzymes, as seen in our study, may lead to alterations of the metabolism of endogenous substances with consequences for normal homeostasis (Stegeman and Hahn, 1994; Nelson et al., 1996; Guengerich, 1999).

4.1.2. EE2 effects on genes encoding glutamate and nuclear receptors

Because glutamate receptors are the most abundant excitatory neurotransmitter receptors of the central nervous system, and particularly important in the brain (Niswender and Conn, 2010; Alix and Domingues, 2011), it was expected that we would find them in our libraries. Glutamate receptors interact with ERs at the membrane level and are known to play a significant role in E2 signalling by altering cell signalling pathways and via epigenetic alterations (Frick et al., 2015). In the present study, we found transcripts coding for *glutamate receptor ionotropic NMDA 2B-like isoform X2* (*GRIN2B*) to be downregulated by EE2-exposure in females. This result is consistent with evidence from a recent work by Arini et al. (2016) which reported that a 4-day cage-exposure to WWTP effluents was enough to cause significant decrease of glutamate receptor N-methyl-D-aspartate (NMDA) binding in fathead minnow females. In the current study, in addition to repressed effects, we also found EE2-exposure to upregulate gene expression. One such gene transcript was *nuclear receptor corepressor 2* (*NCOR2*), which was present in females and upregulated by EE2 exposure. This finding is concordant with evidence from a recent RNA-Seq

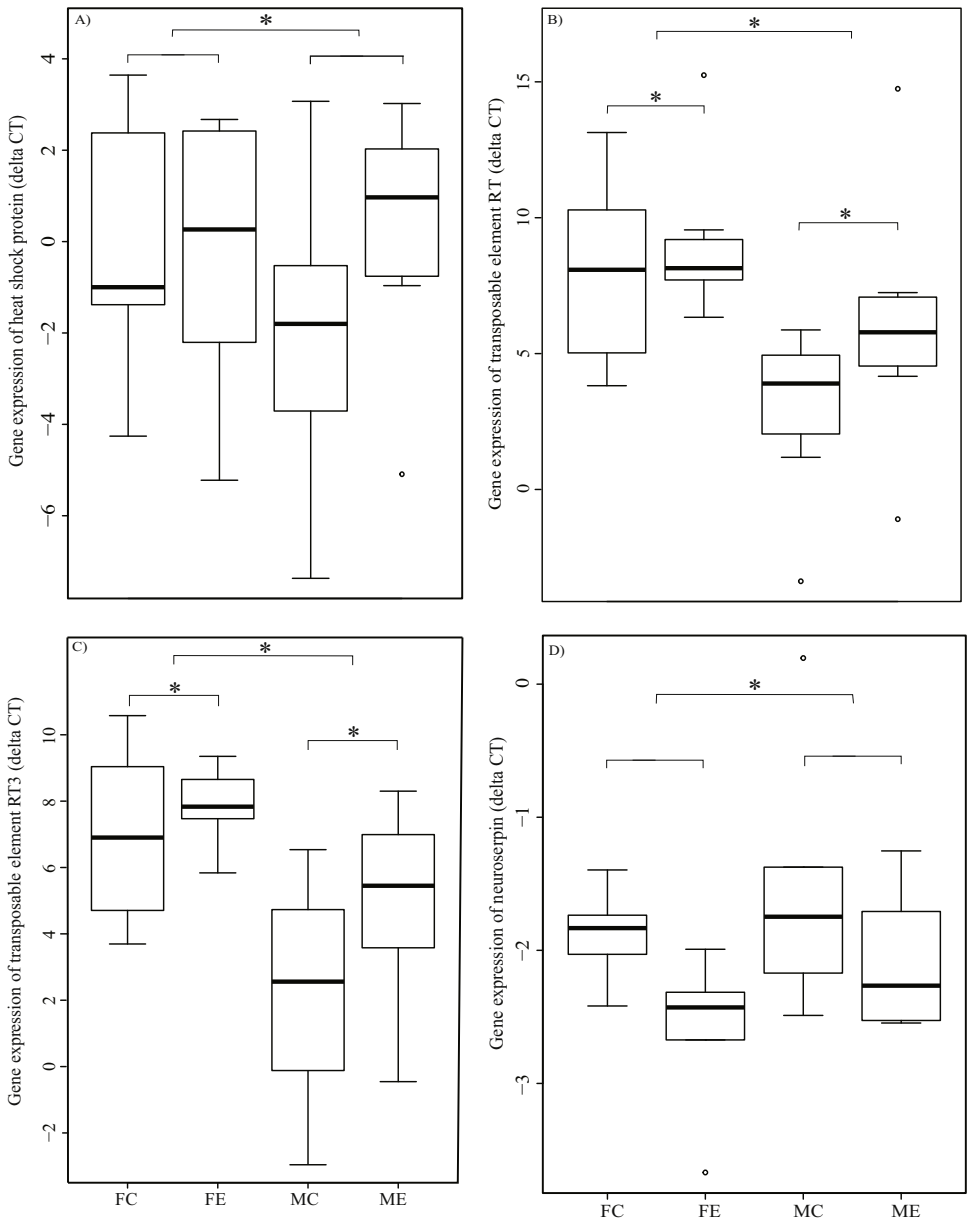

Fig. 4. Effect of 17α-ethinyl estradiol (EE2) treatment on guppy female and male mRNA levels in the brain tissue. A separate group of individual fish from the first exposure were used in the qPCR. Quantitative real-time PCR analysis of the chosen target genes: (A) *heat shock protein (hsp70)*, (B, C) *retrotransposons (RT1, RT3)* and (D) *neuroserpin (NEUS)*. The average expression levels (delta CT) of target genes present normalised data points (n = 8 fish per treatment/sex). FC = control female, FE = EE2-exposed female, MC = control male, ME = EE2-exposed male. Significant differences ($p < 0.05$) are indicated by asterisk (*).

study, where coho salmon (*Oncorhynchus kisutch*) exposed to EE2 (12 ng/L) over 6 weeks showed significantly altered expression of NCOR2 (pituitary gland: Harding et al., 2013). NCOR2 is best known for its role in suppressing the activity of ERα and interacting with ERα in the presence of an antagonist, such as the breast cancer medication tamoxifen (Smith et al., 1997; Lavinsky et al., 1998; Varlakhanova et al., 2010; Zhang et al., 2013).

4.1.3. EE2 effects on transcripts related to transposable elements

It was surprising to find that a large number of the assembled contigs annotated by blast search with FastAnnotator were described as "reverse transcriptase". Closer inspection revealed that these contigs shared low sequence similarity to each other but represented distinct transcripts of the guppy brain transcriptome rather than being a failure of the library normalisation procedures. DE analysis revealed reverse transcriptase transcripts to be at higher abundance in female than male brains but also to be significantly induced in males. This was supported by qPCR quantification, which demonstrated both treatment and sex differences that were statistically significant.

4.1.4. Cluster analysis

Cluster analysis revealed interesting clues about the complex interactions of EE2 in male and female brains, and further, differences between males and females in the brain transcriptome. For instance, the majority of the members of cluster 1 were found to be associated with downregulation of transcripts in males due to EE2-exposure. Only one transcript in this cluster was affected in female. In males, among the downregulated genes was the gene encoding male-specific lethal 3 homolog (MSL3). MSL3 has been described in organisms as diverse as yeast and humans and expressed in all eukaryotic organisms (Marín and Baker, 2000; Zhang et al.,

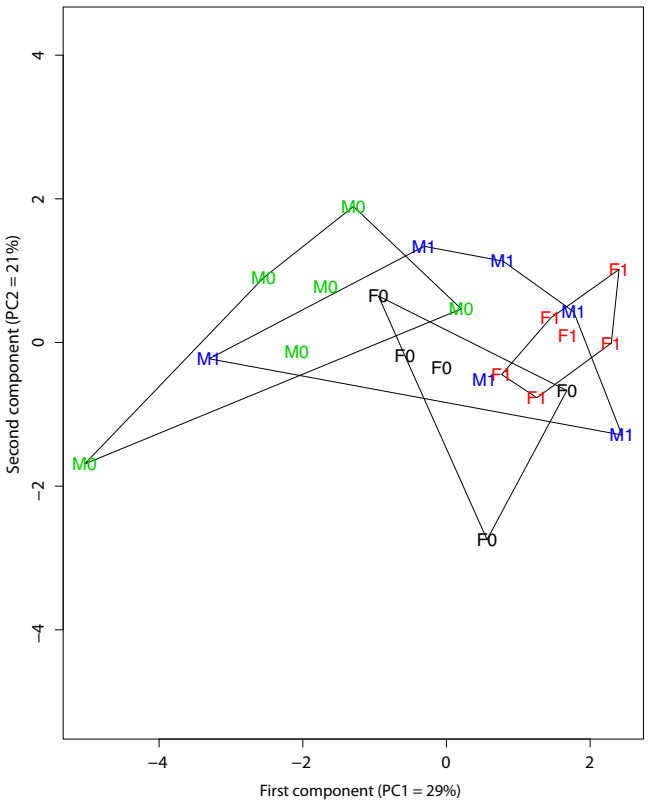

Fig. 5. Principal component analysis (PCA) of the chosen target genes, which were *retrotransposons (RT1, RT3), neuroserpin (NEUS)* and *heat shock protein (hsp70)*. A separate group of individual fish from the first exposure were used (n = 8 per treatment/sex). Four libraries: F0 = control female, F1 = EE2-exposed female, M0 = male control, M1 = EE2-exposed male.

2015). The MSL3 protein plays a role in a dosage-compensation pathway that results in equal levels of gene expression from the X-chromosome in males, and females by a mechanism involving chromatin remodelling and regulation of transcription through histone H4 lysine-16 acetylation (Akhtar and Becker, 2000). Thus, repression of this protein is likely to have adverse effects in the male brain. It is interesting to note that, in the current study, the other downregulated transcripts – piggyBac transposable element-derived protein 4-like (PGBD4) and corepressor interacting with RBPJ 1 (CIR1) – were present in cluster 1 and known to be also involved in chromatin organisation (Hsieh et al., 1999; Li et al., 2013; Yusa, 2015). Our study gives further support that EE2 is altering chromatin and transcription in the male brain, which could have serious consequences at the chromosome level.

Our study revealed transcripts that were quite distinct between the sexes. For example, cluster analyses indicated a group of genes upregulated in females more than in males following EE2 exposure. Furthermore, the identification of clusters of genes with similar expression profiles suggests shared regulatory mechanisms within each cluster but distinct mechanisms between clusters. Our study suggests that EE2 shifts the brain transcriptome in the male to become more like the female. A more "feminized" brain transcriptome was discovered by Stiver et al. (2015), when the authors compared, using a microarray approach, the brain transcriptome of female ocellated wrasse (*Symphodus ocellatus*) with males that engage in an alternative mating strategy (i.e. those that engage in parasitic spawnings). In the current study, feminization of the male brain after EE2 exposure was also indicated by inspection of the DE expression groups that share common transcripts. For example, in the group that is female specific (control female vs control male) there is a large percentage (12/49 = 24%) of transcripts, which are common with transcripts that are induced in the male by EE2. Furthermore, the PCA analysis based on DE, strongly suggests a hyper-feminizing effect: gene transcripts of EE2-males had shifted away from control males and appeared to have 'over-shot' control female expression on Axis 1.

4.2. qPCR validation: EE2 effects on candidate genes

We found both treatment- and sex-specific effects in the chosen candidate genes (a gene encoding a heat-shock protein, transcripts related to transposable elements and the gene encoding neuroserpin). A straightforward interpretation of the PCA would be that exposure to EE2 caused a shift in the gene expression profiles to the right on principal component axis 1, and this could be characterized as a 'feminisation' of males and a 'hyper-feminisation' of females, at least in terms of their gene expression.

4.2.1. Heat-shock protein (hsp70)

Under normal conditions, heat-shock proteins are present at low levels, but when faced with a stressor, such as hypoxia or chemical pollution, their expression increases rapidly (Airaksinen et al., 1998; Basu et al., 2002). In our study, the gene encoding one particular heat-shock protein, hsp70, was upregulated in EE2-exposed fish. Previous studies have shown elevated levels of hsp70 in fish exposed to heavy metals (Misra et al., 1989), kraft pulp mill (Janz et al., 1997) and the surfactant nonylphenol (Maradonna and Carnevali, 2007), but not when exposed to EE2 (Chandra et al., 2012; Maltais and Roy, 2014). A possible reason could be that we measured hsp70 gene expression in brain tissues, while earlier studies measured expression in the liver. Indeed, hsp70 gene expression in the brain of juvenile sole (*Solea solea*) was increased in response to exposure to the estrogen mimic nonylphenol (Palermo et al., 2012). Due to its sensitivity, hsp70, despite being a non-discriminating marker of stress, is widely used in environmental monitoring as a biomarker of adverse biological effects of contamination (Carnevali and Maradonna, 2003; Weber, 2004; Migliarini et al., 2005; Lee et al., 2006). Taken together, our results and those of previous studies highlight the importance, and differences, of chosen test tissue.

4.2.2. Transcripts related to transposable elements

In our study, transcripts related to LINE transposable elements showed both treatment- and sex-dependent effects. EE2-exposed fish had higher abundance of transcripts related to transposable elements than control fish, and females had more transcripts than males. To our knowledge there are no other studies that have looked at how EE2 affects transposable elements. In the present study, we found transposases to be upregulated by EE2 in both male and female brain. Our findings, in this regard, are consistent with studies conducted on other chemical contaminants. For example, exposure to β-naphthoflavone activated transcription of DNA transposons (Krasnov et al., 2005: rainbow trout (*Oncorhynchus mykiss*)), with the authors suggesting that genes transcribing transposable elements could be stimulated under acute conditions. Also, environmental stressors (e.g. hyperthermal, hypoxic, hyposmotic) can regulate transposable element transcription, as shown in black tiger shrimp (*Penaeus monodon*) (De la Vega et al., 2007). Thus, our study gives further support that transposable elements could be used as potential biomarkers of stress.

The higher abundance of transcripts in our female libraries could reflect a higher extent of global transcription in females, compared to males, and may therefore be a non-gene specific effect. In mammals, sex-specific expression of placenta-related retroviral genes encoding syncytin is essential for normal fetal development (Luo et al., 2009) but this is not relevant in fish. The observation that expression of TEs (including LINEs) is important in memory (Maag

et al., 2015) warrants further investigation to determine if similar genome remodelling occurs in the guppy, and if it is more prevalent in females.

4.2.3. Neuroserpin

Neuroserpin is a neuronal protease inhibitor and is known to play a role in neural plasticity (Krueger et al., 1997). In our study, the gene encoding neuroserpin was down-regulated in fish exposed to EE2. Previous studies have shown that overall over- and under-expression of this gene is associated with behavioural changes in mice (i.e. anxiety, reduced locomotion, neophobia) (Madani et al., 2003), but is also switched on in females interacting with attractive males (Cummings et al., 2008). Although we did not look at behavioural responses in the current study, another study by Hallgren et al. (2011) showed that guppies exposed to EE2 increased anxiogenic behaviours (freezing and bottom-dwelling behaviour), which could have been partially due to decreased neuroserpin production. Further research in this area is needed to get a better understanding of the underlying molecular pathway.

5. Conclusions

Our study is the first to establish a brain transcriptome of male and female guppies exposed to a pharmaceutical contaminant. We created four treatment- and sex-specific, brain transcript, normalized libraries and generated transcriptome assemblies for each group. By using an independent group of fish exposed to an environmentally relevant concentration of EE2, we discovered that exposure to EE2 caused both up- and downregulation of specific transcripts, and affected transcript abundance in a sex-specific manner. Finally, RT-qPCR validation demonstrated that expression of transcripts related to transposable elements, neuroserpin and heat shock protein were significantly affected by EE2-exposure. More broadly, not only does our study provide a valuable resource for investigating how wildlife, at the genome level, may be affected by chemical pollutants, it also offers insights into the mechanisms underlying the feminizing of the brains of organisms exposed to environmentally realistic concentrations of EE2.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We thank Rowan Jacques-Hamilton for assistance in the exposure experiments, and John Endler and his research group for supplying fish for this study. We also thank Dr Richard Martin (Strathclyde University) for assistance with the Trinity run using the Engineering and Physical Sciences Research Council (EPSRC grant no. EP/K000586/1)-funded ARCHIE-WeSt High Performance Computer (www.archie-west.ac.uk). This work was supported by an Academy of Finland Postdoctoral Researcher Fellowship (265629) (to MS) and Discovery Grants from the Australian Research Council (DP130100385 and DP160100372) (to BBMW). At the initiation of this study, the Centre for Aquatic Pollution, Identification and Management (CAPIM) received funding from the Victorian Science Agenda Investment Fund managed by the Department of Business and Innovation (DBI), with additional funding from Melbourne Water and the Department of Environment and Primary Industries (Victoria).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.aquatox.2017.02.016.

References

Airaksinen, S., Råbergh, C.M., Sistonen, L., Nikinmaa, M., 1998. Effects of heat shock and hypoxia on protein synthesis in rainbow trout (*Onchorhynchus mykiss*) cells. J. Exp. Biol. 201, 2543–2551.
Alix, J.J., Domingues, A.M., 2011. White matter synapses: form, function, and dysfuntion. Neurology 76, 397–404.
Akhtar, A., Becker, P.B., 2000. Activation of transcription through histone H4 acetylation by MOF: an acetyl transferase essential for dosage compensation in Drosophila. Mol. Cell 5, 367–375.
Anders, S., Reyes, A., Huber, W., 2012. Detecting differential usage of exons from RNA-seq data. Gen. Res. 22, 2008–2017.
Anders, S., Huber, W., 2010. Differential expression analysis for sequence count data. Gen. Biol. 11, R106.
Araújo, F.G., Peixoto, M.G., Pinto, B.C.T., Teixeira, T.P., 2009. Distribution of guppies *Poecilia reticulata* (Peters, 1860) and *Phalloceros caudimaculatus* (Hensel 1868) along a polluted stretch of the Paraíba do Sul River, Brazil. Braz. J. Biol. 69, 41–48.
Arini, A., Cavallin, J.E., Berninger, J.P., Vega, R.M., Mills, M., Villeneuve, D.L., Basu, N., 2016. In vivo and in vitro neurochemical-based assessments of wastewater effluents from the Maumee River area of concern. Environ. Pollut. 211, 9–19.
Arnold, K.E., Brown, A.R., Ankley, G.T., Sumpter, J.P., 2014. Medicating the environment: assessing risks of pharmaceuticals to wildlife and ecosystems. Philos. Trans. R. Soc. B. 369 (2013.0569).
Arukwe, A., 2005. Modulation of brain steroidogenesis by affecting transcriptional changes of steroidogenic acute regulatory (StAR) protein and cholesterol side chain cleavage (P450scc) in juvenile Atlantic salmon (*Salmo salar*) is a novel aspect of nonylphenol toxicity. Environ. Sci. Technol. 39, 9791–9798.
Bayley, M., Nielsen, J.R., Baatrup, E., 1999. Guppy sexual behavior as an effect biomarker of estrogen mimics. Ecotoxicol. Environ. Saf. 43, 68–73.
Baronti, C., Curini, R., D'Ascenzo, G., Di Corcia, A., Gentili, A., Samperi, R., 2000. Monitoring natural and synthetic estrogens at activated sludge sewage treatment plants and in a receiving river water. Environ. Sci. Technol. 34, 5059–5066.
Basu, N., Todgham, A.E., Ackerman, P.A., Bibeau, M.R., Nakano, K., Schulte, P.M., Iwama, G.K., 2002. Heat shock protein genes and their functional significance in fish. Gene 295, 173–183.
Bell, A.M., 2001. Effects of an endocrine disrupter on courtship and aggressive behaviour of male three-spined stickleback, *Gasterosteus aculeatus*. Anim. Behav. 62, 775–780.
Bergman, A., Heindel, J.J., Jobling, S., Kidd, K.A., Zoeller, R.T., 2012. WHO (World Health Organization)/UNEP (United Nations Environment Programme) Global Assessment of the State-of-the-science of Endocrine Disruptors. http://www.who.int/ipcs/publications/new_issues/endocrine_disruptors/en.
Bizarro, C., Eide, M., Hitchcock, D.J., Goksøyr, A., Ortiz-Zarragoitia, M., 2016. Single and mixture effects of aquatic micropollutants studied inprecision-cut liver slices of Atlantic cod (*Gadus morhua*). Aquat. Toxicol. 177, 395–404.
Carnevali, O., Maradonna, F., 2003. Exposure to xenobiotic compounds: looking for new biomarkers. Gen. Comp. Endocrinol. 131, 203–209.
Chandra, K., Bosker, T., Hogan, N., Lister, A., MacLatchy, D., Currie, S., 2012. Sustained high temperature increases the vitellogenin response to 17α-ethynylestradiol in mummichog (*Fundulus heteroclitus*). Aquat. Toxicol. 118–119, 130–140.
Chen, T.W., Gan, R.C., Wu, T.H., Huang, P.J., Lee, C.Y., Chen, Y.Y., et al., 2012. FastAnnotator- an efficient transcript annotation web tool. BMC Genomics 13 (Suppl. 7), S9.
Clouzot, L., Marrot, B., Doumenq, P., Roche, N., 2008. 17α-ethinylestradiol: an endocrine disrupter of great concern. Analytical methods and removal processes applied to water purification. A review. Environ. Prog. 27, 383–396.
Colborn, T., vom Saal, F.S., Soto, A.M., 1993. Developmental effects of endocrine-disrupting chemicals in wildlife and humans. Environ. Health Perspect. 101, 378–384.
Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M., Robles, M., 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21, 3674–3676.
Cummings, M.E., Larkins-Ford, J., Reilly, C.R.L., Wong, R.Y., Ramsey, M., Hoffman, H.A., 2008. Sexual and social stimuli elicit rapid and contrasting genomic response. Proc. R. Soc. B 275, 393–402.
Davey, J.W., Hohenlohe, P.A., Etter, P.D., Boone, J.Q., Catchen, J.M., Blaxter, M.L., 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat. Rev. Gen. 12, 499–510.
De la Vega, E., Degnan, B.M., Hall, M.R., Wilson, K.J., 2007. Differential expression of immune-related genes and transposable elements in black tiger shrimp (*Penaeus monodon*) exposed to a range of environmental stressors. Fish Shellfish Immunol. 23, 1072–1088.
De Wit, M., Keil, D., van der Ven, K., Vandamme, S., Witters, E., De Coen, W., 2010. An integrated transcriptomic and proteomic approach characterizing

estrogenic and metabolic effects of 17α-ethinylestradiol in zebrafish (Danio rerio). Gen. Comp. Endocrinol. 167, 190–201.
Diamanti-Kandarakis, E., Bourguignon, J.P., Giudice, L.C., Hauser, R., Prins, G.S., Soto, A.M., et al., 2009. Endocrine-disrupting chemicals: an endocrine society scientific statement. Endocrinol. Rev. 30, 293–342.
Eisen, M.B., Spellman, P.T., Brown, P.O., Botstei, D., 1998. Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. U. S. A. 95, 14863–14868.
Filby, A.L., Thorpe, K.L., Maack, G., Tyler, C.R., 2007. Gene expression profiles revealing the mechanisms of anti-androgen- and estrogen-induced feminization in fish. Aquat. Toxicol. 81, 219–231.
Fraser, B.A., Weadick, C.J., Janozitz, I., Rodd, F.H., Hughes, K.A., 2011. Sequencing and characterization of the guppy (*Poecilia reticulata*) transcriptome. BMC Genomics 12, 202.
Frick, K.M., Kim, J., Tuscher, J.J., Fortress, A.M., 2015. Sex steroid hormones matter for learning and memory: estrogenic regulation of hippocampal function in male and female rodents. Learn. Mem. 22, 472–493.
Gardner, L.D., Jayasundara, N., Castilho, P.C., Block, B., 2012. Microarray gene expression profiles from mature gonad tissues of Atlantic bluefin tuna, Thunnus thynnus in the Gulf of Mexico. BMC Genomics 13, 530.
Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., et al., 2011. A full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652.
Guengerich, F.P., 1999. Cytochrome P-450 3A4: Regulation and role in drug metabolism. Annu. Rev. Pharmacol. Toxicol. 39, 1–17.
Gunter, H.M., Fan, S., Xiong, F., Franchini, P., Fruciano, C., Meyer, A., 2013. Shaping development through mechanical strain: the transcriptional basis of diet-induced phenotypic plasticity in a cichlid fish. Mol. Ecol. 22, 4516–4531.
Hallgren, S., Volkova, K., Reyhanian, N., Olsen, K.H., Hällström, I.P., 2011. Anxiogenic behaviour induced by 17α-ethynylestradiol in male guppies (*Poecilia reticulata*). Fish Physiol. Biochem. 37, 911–918.
Harding, L.B., Schultz, I.R., Goetz, G.W., Luckenbach, J.A., Young, G., Goetz, F.W., Swanson, P., 2013. High-throughput sequencing and pathway analysis reveal alteration of the pituitary transcriptome by 17 alpha-ethynylestradiol (EE2) in female coho salmon, *Oncorhynchus kisutch*. Aquat. Toxicol. 142, 146–163.
Hegelund, T., Celander, M.C., 2003. Hepatic versus extrahepatic expression of CYP3A30 and CYP3A56 in adult killifish (*Fundulus heteroclitus*). Aquat. Toxicol. 64, 277–291.
Hoffmann, J.L., Torontali, S.P., Thomason, R.G., Lee, D.M., Brill, J.L., Price, B.B., Carr, G.J., Versteeg, D.J., 2006. Hepatic gene expression profiling using Genechips in zebrafish exposed to 17α-ethynylestradiol. Aquat. Toxicol. 79, 233–246.
Hsieh, J.J., Zhou, S., Chen, L., Young, D.B., Hayward, S.D., 1999. CIR, a corepressor linking the DNA binding factor CBF1 to the histone deacetylase complex. Proc. Natl. Acad. Sci. U. S. A. 96, 23–28.
Hultman, M.T., Song, Y., Tollefsen, K.E., 2015. 17α-Ethinylestradiol (EE2) effect on global gene expression in primary rainbow trout (*Oncorhynchus mykiss*) hepatocytes. Aquat. Toxicol. 169, 90–104.
Hurd, P.J., Nelson, C.J., 2009. Advantages of next-generation sequencing versus the microarray in epigenetic research. Brief Funct. Gen. Proteomics 8, 174–183.
Ioannidis, J.P., Allison, D.B., Ball, C.A., Coulibaly, I., Cui, X., Culhane, A.C., et al., 2009. Repeatability of published microarray gene expression analyses. Nat. Genet. 41, 149–155.
Janz, D.M., McMaster, M.E., Munkittrick, K.R., Van Der Kraak, G., 1997. Elevated ovarian follicular apoptosis and heat shock protein-70 expression in white sucker exposed to bleached kraft pulp mill effluent. Toxicol. Appl. Pharmacol. 147, 391–398.
Kashiwada, S., Kameshiro, M., Tatsuta, H., Sugaya, Y., Kullman, S.W., Hinton, D.E., Goka, K., 2007. Estrogenic modulation of CYP3A38, CYP3A40, and CYP19 in mature male medaka (*Oryzias latipes*). Comp. Biochem. Physiol. C 145, 370–378.
Katsiadaki, I., Williams, T.D., Ball, J.S., Bean, T.P., Sanders, M.B., Wu, H., et al., 2010. Hepatic transcriptomic and metabolomic responses in the Stickleback (*Gasterosteus aculeatus*) exposed to ethinyl-estradiol. Aquat. Toxicol. 97, 174–187.
Kidd, K.A., Blanchfield, P.J., Mills, K.H., Palace, V.P., Evans, R.E., et al., 2007. Collapse of a fish population after exposure to a synthetic estrogen. Proc. Natl. Acad. Sci. U. S. A. 104, 8897–8901.
Kidd, K.A., Paterson, M.J., Rennie, M.D., Podmski, C.L., Findlay, D.L., et al., 2014. Direct and indirect responses of a freshwater food web to a potent synthetic oestrogen. Phil. Trans. R. Soc. B 369, 20130578.
Kortenkamp, A., Martin, O., Faust, M., Evans, R., McKinlay, R. *et al.,* 2011. State of the art assessment of endocrine disrupters. European Commission, Directorate-General for the Environment. Final report. Project Contract No. 070307/2009/550687/SER/D3.
Kotrschal, A., Rogell, B., Bundsen, A., Svensson, B., Zajitschek, S., et al., 2013. Artificial selection on relative brain size in the guppy reveals costs and benefits of evolving a larger brain. Curr. Biol. 23, 168–171.
Krasnov, A., Koskinen, H., Afanasyev, S., Mölsä, H., 2005. Transcribed Tc1-like transposons in salmonid fish. BMC Genomics 6, 107.
Krueger, S.R., Ghisu, G.P., Cinelli, P., Gschwend, T.P., Osterwalder, T., et al., 1997. Expression of neuroserpin, an inhibitor of tissue plasminogen activator, in the developing and adult nervous system of the mouse. J. Neurosci. 17, 8984–8996.
Krzywinski, M., Altman, Naomi, 2014. Points of significance: nonparametric tests. Nat. Methods 11, 467–468.
Lange, A., Paull, G.C., Coe, T.S., Katsu, Y., Urushitani, H., et al., 2009. Sexual reprogramming and estrogenic sensitization in wild fish exposed to ethinylestradiol. Environ. Sci. Technol. 43, 1219–1225.
Lavinsky, R.M., Jepsen, K., Heinzel, T., Torchia, J., Mullen, T.M., Schiff, R., et al., 1998. Diverse signaling pathways modulate nuclear receptor recruitment of N-CoR and SMRT complexes. Proc. Natl. Acad. Sci. U. S. A. 95, 2920–2925.
Lee, S.M., Lee, S.B., Park, C.H., Choi, J., 2006. Expression of heat shock protein and hemoglobin genes in *Chironomus tentans* (Diptera, chironomidae) larvae exposed to various environmental pollutants: a potential biomarker of freshwater monitoring. Chemosphere 65, 1074–1081.
Li, X.H., Burnight, E.R., Cooney, A.L., Malani, N., Brady, T., Sander, J.D., et al., 2013. piggyBac transposase tools for genome engineering. Proc. Natl. Acad. Sci. U. S. A. 110, E2279–E2287.
Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods 25, 402–408.
Luo, S.S., Ishibashi, O., Ishikawa, G., Ishikawa, T., Katayama, A., et al., 2009. Human villous trophoblasts express and secrete placenta-specific microRNAs into maternal circulation via exosomes. Biol. Reprod. 81, 717–729.
Maag, J.L., Panja, D., Sporild, I., Patil, S., Kaczorowski, D.C., et al., 2015. Dynamic expression of long noncoding RNAs and repeat elements in synaptic plasticity. Front. Neurosci. 9, 351.
Madani, R., Kozlov, S., Akhmedov, A., Cinelli, P., Kinter, J., et al., 2003. Impaired explorative behavior and neophobia in genetically modified mice lacking or overexpressing the extracellular serine protease inhibitor neuroserpin. Mol. Cell. Neurosci. 23, 473–494.
Maltais, D., Roy, R.L., 2014. Effects of nonylphenol and ethinylestradiol on copper redhorse (*Moxostoma hubbsi*): an endangered species. Ecotoxicol. Environ. Saf. 108, 168–178.
Maradonna, F., Carnevali, O., 2007. Vitellogenin, zona radiata protein, cathepsin D and heat shock protein 70 as biomarkers of exposure to xenobiotics. Biomarkers 12, 240–255.
Marín, I., Baker, B.S., 2000. Origin and evolution of the regulatory gene male-specific lethal-3. Mol. Biol. Evol. 17, 1240–1250.
Martyniuk, C.J., Xiong, H., Crump, K., Chiu, S., Sardana, R., Nadler, A., Gerrie, E.R., Xia, X., Trudeau, V.L., 2006. Gene expression profiling in the neuroendocrine brain of male goldfish (*Carassius auratus*) exposed to 17α-ethinylestradiol. Physiol. Genomics 27, 328–336.
Martyniuk, C.J., Kroll, K.J., Doperalski, N.J., Barber, D.S., Denslow, N.D., 2010. Environmentally relevant exposure to 17α-ethinylestradiol affects the telencephalic proteome of male fathead minnows. Aquat. Toxicol. 98, 344–353.
Martinovic, D., Hogarth, W.T., Jones, R.E., Sorensen, P.W., 2007. Environmental estrogens suppress hormones, behavior, and reproductive fitness in male fathead minnows. Environ. Toxicol. Chem. 26, 271–278.
Mehinto, A.C., Martyniuk, C.J., Spade, D.J., Denslow, N.D., 2012. Applications for next-generation sequencing in fish ecotoxicogenomics. Front. Genet. 3, 62.
Miege, C., Choubert, J.M., Ribeiro, L., Eusebe, M., Coquery, M., 2009. Fate of pharmaceuticals and personal care products in wastewater treatment plants: conception of a database and first results. Environ. Pollut. 157, 1721–1726.
Miller, H.D., Clark, B.W., Hinton, D.E., Whitehead, A., Martin, S., Kwok, K.W., Kullman, S.W., 2012. Anchoring ethinylestradiol induced gene expression changes with testicular morphology and reproductive function in the medaka. PLoS One 7, e52479.
Migliarini, B., Campisi, A.M., Maradonna, F., Truzzi, C., Annibaldi, A., et al., 2005. Effects of cadmium exposure on testis apoptosis in the marine teleost *Gobius niger*. Gen. Comp. Endocrinol. 142, 241–247.
Misra, S., Zafarullah, M., Price-Haughey, J., Gedamu, L., 1989. Analysis of stress-induced gene expression in fish cell lines exposed to heavy meals and heat shock. Biochim. Biophys. Acta 1007, 325–333.
Muller, M., Rabenoelina, F., Balaguer, P., Patureau, D., Lemenach, K., et al., 2008. Chemical and biological analysis of endocrine-disrupting hormones and estrogenic activity in an advanced sewage treatment plant. Environ. Toxicol. Chem. 27, 1649–1658.
Nebert, D.W., Russell, D.W., 2002. Clinical importance of the cytochromes P450. Lancet 360, 1155–1162.
Nebert, D.W., Dalton, T.P., Okey, A.B., Gonzalez, F.J., 2004. Role of aryl hydrocarbon receptor-mediated induction of the CYP1 enzymes in environmental toxicity and cancer. J. Biol. Chem. 279, 23847–23850.
Nelson, D.R., Koymans, L., Kamataki, T., Stegeman, J.J., Feyereisen, R., Waxman, D.J., et al., 1996. P450 superfamily: update on new sequences gene mapping, accession numbers and nomenclature. Pharmacogenetics 6, 1–42.
Nikoleris, L., Hultin, C.L., Hallgren, P., Hansson, M.C., 2016. 17α-Ethinylestradiol (EE2) treatment of wild roach (*Rutilus rutilus*) during early life development disrupts expression of genes directly involved in the feedback cycle of estrogen. Comp. Biochem. Physiol. C 180, 56–64.
Niswender, C.M., Conn, P.J., 2010. Metabotropic glutamate receptors: physiology, pharmacology, and disease. Annu. Rev. Pharmacol. Toxicol. 50, 295–322.
Palermo, F.A., Cocci, P., Nabissi, M., Polzonetti-Magni, A., Mosconi, G., 2012. Cortisol response to waterborne 4-nonylphenol exposure leads to increased brain POMC and HSP70 mRNA expressions and reduced total antioxidant capacity in juvenile sole (*Solea solea*). Comp. Biochem. Physiol. C Toxicol. Pharmacol. 156, 135–139.
Pan, Q., Shai, O., Lee, L.J., Frey, B.J., Blencowe, B.J., 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat. Genet. 40, 1413–1415.
Prokkola, J.M., Katsiadaki, I., Sebire, M., Elphinstone-Davis, J., Pausio, S., Nikinmaa, M., Leder, E.H., 2015. Microarray analysis of di-n-butyl phthalate and 17α ethinyl-oestradiol responses in three-spined stickleback testes reveals novel candidate genes for endocrine disruption. Ecotoxicol. Environ. Saf. 124, 96–104.

Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140.
Rojas, M.R., Leung, C., Bonk, F., Zhu, Y., Edwards, L., et al., 2013. Assessment of the effectiveness of secondary wastewater treatment technologies to remove trace chemicals of emerging concern. Crit. Rev. Environ. Sci. Technol. 43, 1281–1314.
Saaristo, M., Tomkins, P., Allinson, M., Allinson, G., Wong, B.B.M., 2013. An androgenic agricultural contaminant impairs female reproductive behaviour in a freshwater fish. PLoS One 8, e62782.
Saaristo, M., Craft, J.A., Lehtonen, K.K., Lindström, K., 2009a. Sand goby (*Pomatoschistus minutus*) males exposed to an endocrine disrupting chemical fail in nest and mate competition. Horm. Behav. 56, 315–321.
Saaristo, M., Craft, J.A., Lehtonen, K.K., Björk, H., Lindström, K., 2009b. Disruption of sexual selection in sand gobies (*Pomatoschistus minutus*) by 17alpha-ethinyl estradiol, an endocrine disruptor. Horm. Behav. 55, 530–537.
Saaristo, M., Craft, J.A., Lehtonen, K.K., Lindström, K., 2010. Exposure to 17α-ethinyl estradiol impairs courtship and aggressive behaviour of male sand gobies (*Pomatoschistus minutus*). Chemosphere 79, 541–546.
Santos, E.M., Paull, G.C., Van Look, K.J.W., Workman, V.L., Holt, W.V., van Aerle, R., Kille, P., Tyler, C.R., 2007. Gonadal transcriptome responses and physiological consequences of exposure to oestrogen in breeding zebrafish (*Danio rerio*). Aquat. Toxicol. 83, 134–142.
Schmittgen, T.D., Livak, K.J., 2008. Analyzing real-time PCR data by the comparative CT method. Nat. Protoc. 3, 1101–1108.
Schwindt, A.R., Winkelman, D.L., Keteles, K., Murphy, M., Vajda, A.M., 2014. An environmental oestrogen disrupts fish population dynamics through direct and transgenerational effects on survival and fecundity. J. App. Ecol. 51, 582–591.
Shagin, D.A., Rebrikov, D.V., Kozhemyako, B., Altshuler, I.M., Shcheglov, A.S., et al., 2002. A novel method for SNP detection using a new duplex-specific nuclease from crab hepatopancreas. Genome Res. 12, 1935–1942.
Sharma, E., Kunstner, A., Fraser, B.A., Zipprich, G., Kottler, V.A., et al., 2014. Transcriptome assemblies for studying sex-biased gene expression in the guppy, *Poecilia reticulata*. BMC Genomics 15, 400.
Skakkebaek, N.E., Rajpert-De Meyts, E., Main, K.M., 2001. Testicular dysgenesis syndrome: an increasingly common developmental disorder with environmental aspects. Hum. Reprod. 16, 972–978.
Smith, C.L., Nawaz, Z., O'Malley, B.W., 1997. Coactivator and corepressor regulation of the agonist/antagonist activity of the mixed antiestrogen, 4-hydroxytamoxifen. Mol. Endocrinol. 11, 657–666.
Solé, M., Raldua, D., Piferrer, F., Barceló, D., Porte, C., 2003. Feminization of wild carp, *Cyprinus carpio*, in a polluted environment: plasma steroid hormones, gonadal morphology and xenobiotic metabolizing system. Comp. Biochem. Physiol. C Toxicol. Pharmacol. 136, 145–156.
Stegeman, J.J., Hahn, M.E., 1994. Biochemistry and molecular biology of monooxygenases: current perspectives on forms, functions and regulation of Cytochrome P450 in aquatic species. In: Malins, D.C., Ostrander, G.K. (Eds.), Aquatic Toxicology: Molecular, Biochemical, and Cellular Perspectives. Lewis Publishers Inc., Boca Raton, FL, pp. 87–204.
Stiver, K.A., Harris, R.M., Townsend, J.P., Hofmann, H.A., Alonzo, S.H., 2015. Neural gene expression profiles and androgen levels underlie alternative reproductive tactics in the ocellated wrasse, *Symphodus ocellatus*. Ethology 121, 152–167.
Thummel, K.E., Wilkinson, G.R., 1998. In vitro and in vivo drug interactions involving human CYP3A. Annu. Rev. Pharmacol. Toxicol. 38, 389–430.
Valdes, M.E., Marino, D.J., Wunderlin, D.A., Somoza, G.M., Ronco, A.E., et al., 2015. Screening concentration of E1, E2 and EE2 in sewage effluents and surface waters of the "Pampas" region and the "Río de la Plata" estuary (Argentina). Bull. Environ. Contam. Toxicol. 94, 29–33.
Varlakhanova, N., Snyder, C., Jose, S., Hahm, J.B., Privalsky, M.L., 2010. Estrogen receptors recruit SMRT and N-CoR corepressors through newly recognized contacts between the corepressor N terminus and the receptor DNA binding domain. Mol. Cell Biol. 30, 1434–1445.
Vosges, M., Le Page, Y., Chung, B.C., Combarnous, Y., Porcher, J.M., Kah, O., Brion, F., 2010. 17α-Ethinylestradiol disrupts the ontogeny of the forebrain GnRH system and the expression of brain aromatase during early development of zebrafish. Aquat. Toxicol. 99, 479–791.
Weber, L.P., 2004. Acute ex vivo effects of dimethylbenz[a]anthracene and heat shock on rat tail artery. Cardiovasc. Toxicol. 4, 355–362.
Widianarko, B., Van Gestel, C.A.M., Verweij, R.A., Van Straalen, N.M., 2000. Associations between trace metals in sediment, water, and guppy, Poecilia reticulata (Peters), from urban streams of Semarang, Indonesia. Ecotoxicol. Environ. Safe. 46, 101–107.
Zhang, L.D., Gong, C., Lau, S.L.Y., Yang, N., Wong, O.G.W., Cheung, A.N.Y., Tsang, J.W.H., Chan, K.Y.K., Khoo, U.S., 2013. SpliceArray profiling of breast cancer reveals a novel variant of NCOR2/SMRT that is associated with tamoxifen resistance and control of ER alpha transcriptional activity. Cancer Res. 73, 246–255.
Zhang, Y.P., Sun, S.M., Fu, H.T., Ge, X.P., Qiao, H., Zhang, W.Y., et al., 2015. Characterization of the male-specific lethal 3 gene in the oriental river prawn, *Macrobrachium nipponense*. Gen. Mol. Res. 14, 3106–3120.
Zhu, Y.Y., Machleder, E.M., Chenchik, A., Li, R., Siebert, P.D., 2001. Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction. Biotechniques 30, 892–897.
Zhulidov, P.A., Bogdanova, E.A., Shcheglov, A.S., Vagner, L.L., Khaspekov, G.L., et al., 2004. Simple cDNA normalization using kamchatka crab duplex-specific nuclease. Nucleic Acids Res. 32, e37.
Yendrek, C.R., Aisnworth, E.A., Thimmapuram, J., 2012. The bench scientist's guide to statistical analysis of RNA-Seq data. BMC Res. Notes 5, 506.
Yusa, K., 2015. PiggyBac transposon. Microbiol. Spectr. 3 (MDNA3-0028-2014).