Identification of conserved hepatic transcriptomic responses to 17β-estradiol using high-throughput sequencing in brown trout

Tamsyn M. Uren Webster, Janice A. Shears, Karen Moore, and Eduarda M. Santos

Biosciences, College of Life & Environmental Sciences, University of Exeter, Exeter, United Kingdom

Submitted 2 December 2014; accepted in final form 8 June 2015

Uren Webster TM, Shears JA, Moore K, Santos EM. Identification of conserved hepatic transcriptomic responses to 17β-estradiol using high-throughput sequencing in brown trout. Physiol Genomics 47: 420–431, 2015. First published June 16, 2015; doi:10.1152/physiolgenomics.00123.2014.—Estrogenic chemicals are major contaminants of surface waters and can threaten the sustainability of natural fish populations. Characterization of the global molecular mechanisms of toxicity of environmental contaminants has been conducted primarily in model species rather than species with limited existing transcriptomic or genomic sequence information. We aimed to investigate the global mechanisms of toxicity of an endocrine disrupting chemical of environmental concern [17β-estradiol (E2)] using high-throughput RNA sequencing (RNA-Seq) in an environmentally relevant species, brown trout (Salmo trutta). We exposed mature males to measured concentrations of 1.94, 18.06, and 34.38 ng E2/l for 4 days and sequenced three individual liver samples per treatment using an Illumina HiSeq 2500 platform. Exposure to 34.4 ng E2/L resulted in 2,113 differentially regulated transcripts (FDR < 0.05). Functional analysis revealed upregulation of processes associated with vitellogenesis, including lipid metabolism, cellular proliferation, and ribosome biogenesis, together with a downregulation of carbohydrate metabolism. Using real-time quantitative PCR, we validated the expression of eight target genes and identified significant differences in the regulation of several known estrogen-responsive transcripts in fish exposed to the lower treatment concentrations (including esr1 and zp2.5). We successfully used RNA-Seq to identify highly conserved responses to estrogen and also identified some estrogen-responsive transcripts that have been less well characterized, including nts and tgm2.1. These results demonstrate the potential application of RNA-Seq as a valuable tool for assessing mechanistic effects of pollutants in ecologically relevant species for which little genomic information is available.

RNA-Seq; Illumina; transcriptomics; sequencing; salmonid
response and its role as a valuable and robust tool in ecotoxicology.

MATERIALS AND METHODS

Fish maintenance. All experiments were conducted under approved protocols according to the UK Home Office regulations for use of animals in scientific procedures.

A mixed-sex population of brown trout (2 yr old) including mature and immature fish of both sexes were obtained from a local aquaculture facility (Hooke Springs Trout Farm, Dorset, UK) in late September, to correspond with the latter stages of reproductive maturation in this species, and maintained in 215-liter tanks to allow for acclimation to laboratory conditions for 3 wk prior to the start of the exposure, at the University of Exeter, UK. Each tank was aerated, supplied with 430 l/day dechlorinated tap water, and maintained at 12 ± 0.2°C, pH 7.5. Fish were kept under a 16:8 h light-dark cycle (with 30 min dawn/dusk transitional periods) and fed with pellet feed (8 mm; Biomar, Grangemouth, UK) at a rate of 2% body weight per day. To exclude sexually mature females, which would be naturally excreting estrogens that may have influenced the exposure, we measured plasma calcium concentrations in all fish prior to the start of the exposure. Concentration of plasma calcium is known to be a good indicator of vitellogenin and maturity status of female fish (40). Fish used in the exposure experiment included mature males and immature fish of both sexes.

Chemical exposures and sampling. Chemical exposure was conducted via a flow-through system for a period of 4 days. This exposure duration was chosen because short-term estrogen exposures are known to induce considerable and rapid transcriptional change (19) and to limit any possible confounding secondary effects of estrogen exposure (38). Four days can also be expected to simulate pulses of estrogenic compounds in the environment (dawn/dusk transitional periods) and fed with pellet feed (8 mm; Biomar, Grangemouth, UK) at a rate of 2% body weight per day. To exclude sexually mature females, which would be naturally excreting estrogens that may have influenced the exposure, we measured plasma calcium concentrations in all fish prior to the start of the exposure. Concentration of plasma calcium is known to be a good indicator of vitellogenin and maturity status of female fish (40). Fish used in the exposure experiment included mature males and immature fish of both sexes.

Chemical exposures and sampling. Chemical exposure was conducted via a flow-through system for a period of 4 days. This exposure duration was chosen because short-term estrogen exposures are known to induce considerable and rapid transcriptional change (19) and to limit any possible confounding secondary effects of estrogen exposure (38). Four days can also be expected to simulate pulses of environmental estrogenic exposure associated with agricultural pollution (30, 51). Fish were exposed to three nominal concentrations, 2.5, 25, and 250 ng E2/l (17β-estradiol ≥98% purity, Sigma), or a dilution water control. The lowest concentration is in the range of assay detection (0.02–2 ng/l) and measured in parallel with exposure water samples and were 85, 77, and 98% for testosterone and other sex steroids. The measured concentrations of E2 in the water were 1.94, 18.06, and 34.38 ng E2/l.

Transcriptomic analysis. Raw sequence reads from individual samples were mapped back against the assembled transcripts using SigmaStat (version 12.0). All morphometric data met assumptions of normality (verified by the Shapiro-Wilk test) and equal variance, and were analyzed by single-factor one-way analysis of variance (ANOVA).

RNA extraction, library preparation, and sequencing. Transcript profiling was conducted in the livers of three sexually mature males per treatment group. RNA was extracted from livers with TRI reagent (Sigma-Aldrich) according to the manufacturer’s instructions and then further purified and treated with DNase on RNeasy Mini extraction columns (Qiagen). The concentration, purity, and integrity of RNA were determined using a NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies) and an Agilent 2100 Bioanalyzer (Agilent Technologies). All RNA input to library construction was of high quality with 260/280 and 260/230 ratios >1.8 and RNA integrity number scores >8. External RNA Controls Consortium (ERCC) spike-in control mixes (Ambion) were added to all individual RNA samples, according to the manufacturer’s instructions. cDNA libraries from all 15 samples were then prepared using the Illumina TruSeq RNA Sample Preparation kit, multiplexed with 24 samples per lane (together with samples from another project) and sequenced using an Illumina HiSeq 2500 to generate 100 bp paired-end reads, according to the manufacture’s instructions.

Transcriptome assembly and annotation. To maximize sequence coverage depth and assemble an optimized male liver transcriptome for brown trout, we combined sequence reads from all samples from the current study with those from another project. Transcriptome assembly and annotation were conducted as described previously (57), and this transcriptome was then used as a basis for expression analysis in both projects. Contaminating Illumina adaptor sequences were removed, and the first 12 bp of all raw sequence reads were trimmed to remove 5′ bias caused by random hexamer priming using the FASTXToolkit (http://hannonlab.cshl.edu/fastx_toolkit; July 2013). 3′ sliding window quality trimming was performed (http://wiki.bioinformatics.ucdavis.edu/index.php/Trim.slidingWindow.pl, July 2013) and all reads where <90% bases had a Phred quality score >20 and those shorter than 15 bp were discarded. Digital normalization was performed to remove highly duplicated reads using the normalize-by median.py script part of the khmer package described by Brown et al. (7), with the recommended k-mer value of 20 and a coverage threshold of 200. This process reduces the computer memory requirements of transcriptome assembly and also reduces the risk of potential sequencing error accumulation in abundant transcripts. All retained reads were then paired and separated into forward and reverse fastq files before de novo transcriptome assembly using Trinity (version r2013-02-25; Ref. 15, using the default parameters and specifying a minimum contig length of 200 bp). All transcripts were annotated using Blastx against Ensembl peptide databases (Release 71; April 2013) using an e-value cut off <1e−15 and assigned in the following preferential order: zebrafish (Danio rerio); human (Homo sapiens) and mouse (Mus musculus); stickleback (Gasterosteus aculeatus), medaka (Oryzias latipes), tilapia (Oreochromis niloticus), and cod (Gadus morhua)). Additional annotation of previously unannotated differentially expressed transcripts was performed with Blast (e-value <1e−15) against NCBI RefSeq, nr, and nt databases.

Transcriptomic analysis. Raw sequence reads from individual samples were mapped back against the assembled transcripts using Bowtie2 (version 2.1.0, Ref. 32), using the -k 1 parameter to report a single best hit for each read and limit ambiguous mapping to redund
dant transcripts. Raw count data for each transcript were extracted using idstats in samtools (version 0.1.18, Ref. 35) and input into edgeR (45) for differential expression analysis. A criterion of at least one count in a minimum of three individual samples (corresponding to the number of individuals per treatment group) was imposed, and tagwise dispersion was applied with the recommended prior.df = 10. Comparisons were initially conducted between the two control groups to ensure that our analysis did not identify differential expression as a result of random variation between groups. Following this initial analysis, comparisons were conducted between the six individual fish from the combined control groups and three individuals from each of the other treatment groups. Transcripts were considered differentially expressed with an FDR < 0.05 (Benjamini-Hochberg correction). Hierarchical clustering was performed on all differentially expressed transcripts for all samples by an Euclidean distance metric, using the Pheatmap package for R. Functional analysis was then performed for differentially expressed genes from each treatment using the Database for Annotation, Visualization and Integrated Discovery (DAVID v6.7, Ref. 23), with the newly assembled brown trout male liver transcriptome as a background. KEGG pathways and Gene Ontology (GO) terms for biological processes, cellular components, and molecular functions were considered significantly overrepresented when P < 0.05. Canonical pathway and network analysis was conducted using Ingenuity Pathways Analysis (IPA; Ingenuity Systems, http://www.ingenuity.com) based on the list of differentially expressed transcripts.

The raw sequence data, and processed results from the expression analysis have been deposited in the National Center for Biotechnology Information’s Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo), and are available via the GEO series accession number GSE57490.

**Results**

**Morphological parameters.** Visual examination of the gonads revealed there were 17 mature males, defined by the presence of large white testis and milt, across all treatment groups (n = 3–7 fish per treatment). The remaining fish were sexually immature males and females. The mean mass and length of all mature males were 472.3 ± 9.1 g and 34.3 ± 0.2 cm, of immature males were 430.5 ± 44.5 g and 32.6 ± 0.9 cm, and of females were 420.0 ± 15.9 g and 33.5 ± 0.5 cm. The mean condition factors, HSI, and GSI were, respectively, 1.16 ± 0.01, 1.11 ± 0.03, and 3.95 ± 0.32 for mature males, 1.24 ± 0.12, 0.85 ± 0.04, and 0.06 ± 0.01 for immature males, and 1.11 ± 0.02, 1.00 ± 0.03, and 0.31 ± 0.02 for immature females, and there were no significant differences for these parameters between treatment groups. Additionally, we observed no alteration of general health or behavior during the exposure period.

**Sequencing and transcriptome assembly.** In total, we sequenced 225.3 million paired 100 bp reads from male brown trout liver samples, and 208.1 million (92.4%) of these were retained after processing and quality filtering. As described in Uren Webster et al. (57), highly duplicated reads were then removed by digital normalization, and 46.73 million paired reads were retained for input into the de novo transcriptome assembly. The final transcriptome assembly consisted of 172,688 transcripts (107,095 loci) with a mean length of 767.5 bp and an N50 of 1,292 bp. Of these, 62,236 transcripts were annotated using Blastx (e-value < 1e-15) against Ensembl peptide databases, and these included representation of 16,121 unique zebrafish transcripts.

<table>
<thead>
<tr>
<th>Target Gene</th>
<th>Symbol</th>
<th>Forward Primer (5’-3’)</th>
<th>Reverse Primer (5’-3’)</th>
<th>Product Size, bp</th>
<th>Ta, °C</th>
<th>PCR Efficiency, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vesicle-associated membrane</td>
<td>vapa</td>
<td>CACTGAAACATCTCAACATCT</td>
<td>TGAACATTGGATACAGG</td>
<td>118</td>
<td>59.5</td>
<td>95.5</td>
</tr>
<tr>
<td>Estrogen receptor 1</td>
<td>esr1</td>
<td>GCAACAACTCCACACAGATT</td>
<td>ATCCAATATAGCAAGCGAA</td>
<td>126</td>
<td>59.5</td>
<td>101.9</td>
</tr>
<tr>
<td>Carnitine O-octanoyltransferase</td>
<td>crot</td>
<td>GTTGATTAATGCTGGGTTG</td>
<td>ATCAAGATATACGTCTTAG</td>
<td>83</td>
<td>53.5</td>
<td>105.0</td>
</tr>
<tr>
<td>Nothepsin</td>
<td>nots</td>
<td>AGTATACAGAAGAGCTGA</td>
<td>AGAGGGGAAAGAAGAAG</td>
<td>86</td>
<td>58.0</td>
<td>114.0</td>
</tr>
<tr>
<td>Tyrosine aminotransferase</td>
<td>tat</td>
<td>AGCATCTGATTAGCTAGGAA</td>
<td>TCAGACACCGACCAAGAT</td>
<td>83</td>
<td>56.0</td>
<td>95.7</td>
</tr>
<tr>
<td>Transglutaminase 2 like</td>
<td>tgt2l</td>
<td>CTGCGACCTAAGCAACAA</td>
<td>ATGACAAGACCTTGGACAC</td>
<td>75</td>
<td>56.0</td>
<td>99.9</td>
</tr>
<tr>
<td>Vitellogenin 1</td>
<td>vtg1</td>
<td>AACTTGTTGATCTTATGAG</td>
<td>TATATCTACTCTCTCTGA</td>
<td>132</td>
<td>55.0</td>
<td>111.5</td>
</tr>
<tr>
<td>Zona pellucida glycoprotein 2.5</td>
<td>zp2.5</td>
<td>ATGAAATACCAGAACAGG</td>
<td>ATGAGGACACAGGAAATAG</td>
<td>75</td>
<td>55.0</td>
<td>101.2</td>
</tr>
<tr>
<td>Zona pellucida glycoprotein 3a.2</td>
<td>zp3a.2</td>
<td>AACTGACATCAGATATC</td>
<td>CACATGCTAATCTCAGTGC</td>
<td>86</td>
<td>54.5</td>
<td>112.6</td>
</tr>
</tbody>
</table>

RT-QPCR, real-time quantitative PCR; Ta, annealing temperature.
Transcript expression analysis. A total of 137.6 million raw reads were obtained from the libraries generated from liver samples of E2-exposed and control male fish, averaging 9.2 million reads per individual sample, and 83.1% of these were remapped against the transcriptome assembly. Differential expression analysis between the control groups revealed only three differentially regulated transcripts. Comparisons between each E2 treatment group and the combined control group were conducted and resulted in only four (bcl6a, spns1, NM_001124310, uncharacterized transcript) and two (NM_001124310, uncharacterized transcript) differentially expressed transcripts for fish exposed to 1.9 and 18.1 ng E2/l, respectively. This may correspond to changes associated with tank effects, given that similar numbers of differentially expressed transcripts were found between the two control tanks, and no treatment associated trends were observed for any of these transcripts. Exposure to 34.4 ng E2/l, however, resulted in 2,113 differentially expressed transcripts (Fig. 1A, Supplemental Table S1), including 808 unique annotations. Multidimensional scaling plots and Euclidean cluster analysis based on all differentially regulated transcripts show that all three individual fish exposed to 34.4 ng E2/l have a very similar and consistent expression profile, clearly distinct from all other fish, whereas the control fish and those exposed to the lower concentrations of E2 cluster together (Fig. 1, B and C).

A list of the 20 most up- and downregulated transcripts following exposure to 34.4 ng E2/l is shown in Table 2. The greatest fold-changes in expression were associated with up-

---

1 The online version of this article contains supplemental material.
regulated transcripts and were dominated by well-characterized estrogen-responsive genes, including a number of vitellogenin transcripts (vtg1, vtg1l, vtg2, vtg3, vtg6, vtg7), of which vtg1 was the most highly expressed. Additionally, a transcript encoding nothepsin (not) was similarly strongly induced in fish exposed to 34.4 ng E2/l. Transcripts encoding zona pellucida proteins (zp2.2, zp2.5, zp3a.1, zp3a.2) were also upregulated (up to 70- to 230-fold) and esr1 was upregulated by up to 27-fold. There was a trend toward upregulation of esr1, zp3a.2, and zp2.5 (2- to 4-fold) in the lower treatment groups compared with the control, but these results were not statistically significant, likely due to the low number of replicates in the RNA-Seq dataset.

Analysis of ERCC spike-in control data was conducted to determine the accuracy and dynamic range of the transcript expression measurements in this study. For all individual samples, there was a strong correlation between the calculated FPKM (fragments per kilobase of transcript per million mapped reads) values and the expected concentration of control transcripts (mean $R^2 = 0.902 \pm 0.005$). The dynamic range was calculated for all samples individually, using the control transcripts that were detected in a minimum of three libraries as the lower cut-off limit. The mean dynamic range in expression level for all 15 libraries was 26,753 FPKM. There was also a good correlation between the calculated and expected changes in transcript expression level between samples spiked with ERCC mix 1 and mix 2 ($R^2 = 0.58$). Together, these results provide strong technical validation for the quantitative expression profiling conducted in this study.

### Functional analysis.
Enriched GO terms and KEGG pathways among up- and downregulated transcripts following exposure to 34.4 ng E2/l are illustrated in Fig. 2. GO terms including translation, ribosome, lipid metabolic processes, and growth factor binding were overrepresented in the list of differentially regulated transcripts. Regulated transcripts within these GOs include RNA polymerases (polr1a, polr3a) for transcription; translation initiation factors (eif1ad, eif3s10, eif4a2) for translation; and ribosomal components and binding proteins (rp5l5a, rpl12, rpl15, rpl36a, rpl39, rplp0, rps21, rps2, rps9, rps23, rps3a, rrbp1a) for ribosome. Within lipid metabolism, differentially regulated transcripts included apolipoproteins (apoB, apoB, apoF, apoE), lipoprotein receptor (lpn), glycolipid transfer proteins (gptd2), and transcripts involved in PPAR signaling (ppardb, acoxl). In addition, insulin-like growth

### Table 2. List of the 20 most upregulated and downregulated transcripts in fish exposed to 34.4 ng E2/l

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Namea</th>
<th>Fold Changeb</th>
<th>FDR</th>
<th>Symbol</th>
<th>Namea</th>
<th>Fold Changeb</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>vtg1</td>
<td>vitellogenin 1</td>
<td>$\geq 5438$</td>
<td>4.6E-11</td>
<td>tat</td>
<td>tyrosine aminotransferase</td>
<td>↓ 4.6</td>
<td>8.9E-9</td>
</tr>
<tr>
<td>nots</td>
<td>nothepsin</td>
<td>$\geq 4475$</td>
<td>5.4E-10</td>
<td>tgm21</td>
<td>transglutaminase 2, like</td>
<td>↓ 186.1</td>
<td>3.3E-8</td>
</tr>
<tr>
<td>vtg6</td>
<td>vitellogenin 6</td>
<td>$\geq 2000$</td>
<td>1.3E-10</td>
<td>cabln8</td>
<td>cerebellin 8</td>
<td>↓ 5.7</td>
<td>2.6E-7</td>
</tr>
<tr>
<td>vtg2</td>
<td>vitellogenin 2</td>
<td>$\geq 1100$</td>
<td>7.3E-92</td>
<td>hsd3b7</td>
<td>hydroxy-Δ5-steroid dehydrogenase, 3beta- and steroid delta-isomerase</td>
<td>↓ 5.4</td>
<td>2.8E-7</td>
</tr>
<tr>
<td>si:dkey-4c23.3</td>
<td>vitellogenin 1-1</td>
<td>$\geq 220$</td>
<td>9.6E-59</td>
<td>efrf1</td>
<td>ERBB receptor feedback inhibitor 1</td>
<td>↓ 6.3</td>
<td>5.5E-7</td>
</tr>
<tr>
<td>vtg3</td>
<td>vitellogenin 3</td>
<td>$\geq 825$</td>
<td>1.4E-56</td>
<td>igtg1a</td>
<td>insulin-like growth factor binding protein 1a</td>
<td>↓ 50.7</td>
<td>6.2E-7</td>
</tr>
<tr>
<td>zp3a.2</td>
<td>zona pellucida 3a.2</td>
<td>149</td>
<td>1.6E-53</td>
<td>slc3a2a</td>
<td>solute carrier family 3, member 2a</td>
<td>↓ 10.4</td>
<td>6.8E-7</td>
</tr>
<tr>
<td>zp2.5</td>
<td>zona pellucida 2.5</td>
<td>$\geq 177.6$</td>
<td>1.1E-42</td>
<td>fxd2a</td>
<td>chromosome 5 open reading frame 4</td>
<td>↓ 4.8</td>
<td>1.8E-6</td>
</tr>
<tr>
<td>zp3a.1</td>
<td>zona pellucida 3a.1</td>
<td>$\geq 161.5$</td>
<td>1.1E-42</td>
<td>epha8</td>
<td>eph receptor A8</td>
<td>↓ 21</td>
<td>7.5E-6</td>
</tr>
<tr>
<td>vtg7</td>
<td>vitellogenin 7</td>
<td>$\geq 107$</td>
<td>8.3E-42</td>
<td>pphfb1</td>
<td>6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 1</td>
<td>↓ 30.1</td>
<td>8.1E-6</td>
</tr>
<tr>
<td>crot</td>
<td>carmine o-octanoyltransferase</td>
<td>54.4</td>
<td>4.9E-40</td>
<td>ppc7a</td>
<td>PTC7 protein phosphatase homolog a</td>
<td>↓ 6.3</td>
<td>9.2E-6</td>
</tr>
<tr>
<td>esr1</td>
<td>estrogen receptor 1</td>
<td>25.7</td>
<td>1.6E-37</td>
<td>si:dkey-238ol3.4</td>
<td>↓ 4.5</td>
<td>1.8E-5</td>
<td></td>
</tr>
<tr>
<td>zp2.2</td>
<td>zona pellucida 2.2</td>
<td>160.7</td>
<td>5.1E-31</td>
<td>sl3gal3b</td>
<td>ST3 beta-galactoside alpha-2,3-sialyltransferase 3b</td>
<td>↓ 3.8</td>
<td>2.4E-5</td>
</tr>
<tr>
<td>aqp12</td>
<td>aquaporin 12</td>
<td>28.6</td>
<td>8.4E-31</td>
<td>ret</td>
<td>ret proto-oncogene receptor tyrosine kinase</td>
<td>↓ 8.5</td>
<td>4.3E-5</td>
</tr>
<tr>
<td>lirc5kb</td>
<td>leucine rich repeat containing 58b</td>
<td>20.6</td>
<td>1.0E-30</td>
<td>ntg2a</td>
<td>netrin g2a</td>
<td>↓ 4.9</td>
<td>4.9E-5</td>
</tr>
<tr>
<td>igb5a</td>
<td>insulin-like growth factor binding protein 5a</td>
<td>$\geq 49$</td>
<td>2.6E-30</td>
<td>ulk1a</td>
<td>unc-51-like kinase 1a</td>
<td>↓ 4.7</td>
<td>4.9E-5</td>
</tr>
<tr>
<td>rdh10a</td>
<td>retinol dehydrogenase 10a</td>
<td>$\geq 108$</td>
<td>8.8E-28</td>
<td>erg</td>
<td>growth factor receptor-bound protein 7</td>
<td>↓ 26.3</td>
<td>1.2E-4</td>
</tr>
<tr>
<td>slc7a11</td>
<td>solute carrier family 7, member 11</td>
<td>$\geq 51$</td>
<td>8.9E-27</td>
<td>cldn1a</td>
<td>claudin 11a</td>
<td>↓ 6.5</td>
<td>1.3E-4</td>
</tr>
<tr>
<td>lpgat1</td>
<td>lysophosphatidylglycerol acyltransferase 1</td>
<td>26.0</td>
<td>3.0E-25</td>
<td>slc2a29</td>
<td>solute carrier family 25, member 29</td>
<td>↓ 4.8</td>
<td>1.4E-4</td>
</tr>
</tbody>
</table>

*aWhere there were multiple differentially regulated transcripts assigned the same annotation, only the most significantly regulated transcript is included in this list.
*bFor transcripts where no read counts were detected in any of the individuals in one of the groups, a nominal value of 1 count was given to each individual in that group to calculate a fold change value, for visualization purposes, in this table. E2, 17β-estradiol; FDR, false discovery rate.
factor (IGF) signaling was also affected and transcripts encoding IGF binding proteins (IGFBPs) were upregulated in some cases (igfbp5a, igfbp2a, igfbp2b) and downregulated in others (igfbp1a, igfbp1b).

For downregulated transcripts, the most overrepresented GO terms related to amino acid metabolism and biosynthesis and associated processes including organic acid biosynthesis, transaminase activity, and pyridoxal phosphate binding. Of note, a number of processes involved in cysteine and methionine metabolic pathways were enriched, whereby differentially regulated transcripts included betaine-homocysteine methyltransferase (bhmt), S-adenosylmethionine synthase (sash1), methionine adenosyltransferase (mat2aa), and cysteine dioxygenase (cdo1). Apoptosis and programmed cell death were also overrepresented in the list of downregulated transcripts.

IPA identified a gene network involved in the response to E2 with functions relating to amino acid metabolism, cell death and survival, endocrine system development, and small molecule biochemistry, and with esr1 and the myelocytomatosis oncogene (myc) as central nodes (Fig. 3). In particular, a number of genes and processes in this network can be associated with transcription and translation including histones, RNA polymerase, and several translation elongation factors.

RT-QPCR validation. RT-QPCR analysis, performed for eight transcripts using all of the male fish, fully validated the results of the RNA-Seq expression analysis (Fig. 4A). Five upregulated transcripts (vtg1, nots, esr1, zp2.5, zp3a.2) and two downregulated transcripts (tat, tgm2l) were confirmed as being significantly differentially expressed in fish exposed to 34.4 ng/l E2 compared with those in the control group. Furthermore, the expression of esr1 was also significantly increased in both the 1.9 and 18.1 ng/l groups, while zp2.5 and tgm2l were significantly up- and downregulated, respectively, following exposure to 18.1 ng/l, confirming the trends identified in the RNA-Seq dataset. For crot, there was a significant increase in expression of this transcript in male fish exposed to 1.9 ng/l E2 and increasing, but nonsignificant, trends in expression in the two higher treatment groups. RT-QPCR analysis was also performed on the same transcripts in the immature female fish and identified very similar patterns of expression including significant upregulation of vtg1,nots, zp2.5, zp3a.2, crot in the highest treatment group, and significant upregulation of esr1 in fish exposed to both 1.9 and 34.4 ng/l E2 (females were not present in the group of fish exposed to 18.1 ng/l; Fig. 4B). tgm2l expression was detected by RT-QPCR only in sexually mature males (n = 15 across all groups) and not in any immature males (n = 6) or females (n = 10).

**DISCUSSION**

Despite the ecological and economic importance of brown trout, relatively little is known about the responses of this species to key stressors affecting its freshwater habitat, including endocrine disrupting chemicals. Here, we have conducted global transcriptional profiling using RNA-Seq in the liver of sexually mature males exposed to E2 and identified very significant transcriptional changes at the highest concentration tested (34.4 ng E2/l) that were very consistent between the
three individual fish in this group. In contrast, concentrations of 
up to 18.1 ng E2/l did not induce significant changes following 
4 days of exposure. RT-QPCR analysis, using a greater number 
of individual fish per treatment, fully validated the results of 
the RNA-Seq analysis for eight target genes (including both 
up- and downregulated transcripts) and also identified signifi-
cant differences in expression of several of these in the 1.9 ng 
E2/l group (esr1 and crot) and 18.1 ng E2/L group (esr1, zp2.5, 
and tgm2l). Importantly, this demonstrates that short-term 
exposures to low, environmentally relevant concentrations of
E2 induce significant changes in the expression of some of the most sensitive estrogen-responsive genes in brown trout. Although 34.4 ng/l is far higher than EEQ concentrations regularly reported in surface waters, it is within a range reported to occur in treated sewage effluent (16, 26) and may be associated with short-term peaks of E2 contamination in streams inhabited by this species that occur as a result of agricultural pollution (51). Therefore, the extent of transcriptional change found in this study after a 4-day exposure to 34.4 ng E2/l may also be of interest for evaluating the potential impact of estrogens on populations of brown trout in the most contaminated environments.

Conserved estrogen-responsive transcripts. Transcripts encoding six vitellogenin isoforms were very strongly induced in males exposed to 34.4 ng E2/l but were not detected by RNA-Seq in fish exposed to 18.1 ng/l and below. RT-QPCR quantified very low levels of vtg1 expression in males in the control group and confirmed a very significant upregulation in males exposed to 34.4 ng E2/l (by 66,000-fold), but no significant difference in expression in the two lower treatment groups. These results are similar to those reported in previous transcriptomic studies where vtg transcripts are generally the most strongly upregulated following estrogen exposure (e.g., 34, 60). The threshold for induction of vitellogenin transcript expression in mature males in this study was slightly higher than previously reported values for transcript and protein induction in salmonids. In juvenile rainbow trout exposed to E2 for 14 days, the median effective treatment concentration for plasma Vtg protein induction was in the range of 19–26 ng/l (55), while the lowest effective concentration for both plasma Vtg protein and hepatic vtg1 transcript induction was found to be 14 ng/l (54). In juvenile brown trout, the median EC50 for plasma Vtg protein induction following 7-day E2 exposure was 15 ng/l (5). Here, this relatively higher threshold level for vitellogenin induction compared with previous reports for...
E2 exposure may reflect the shorter exposure period in this study of only four days and/or differences in maturity status.

Transcripts encoding four zona pellucida proteins and esr1 were also among the most upregulated transcripts in fish exposed to 34.4 ng E2/l, similar to previous reports showing strong upregulation in vitellogenic females and induction by E2 in males (18, 34, 60). There were also nonsignificant trends toward upregulation of zps and esr1 in the lower treatment groups (by 2- to 4-fold), and RT-QPCR confirmed significant upregulation of esr1 (in males and females) from 1.9 ng E2/l, and of zp2.5 (males only) from 18.1 ng E2/l. This suggests that, compared with vitellogenin, zona pellucida protein and esr1 transcription is particularly sensitive to estrogen exposure in brown trout, as has been previously reported in other species (18, 30, 54). This corresponds with previous reports that, in the liver, esr1 is the most responsive estrogen receptor to estrogen exposure and is associated with vitellogenesis (13, 30). There are reports that esr2α and/or esr2β have a major role in regulating vitellogenesis (33, 52), but our results showed no significant difference in the expression of either of these transcripts or any apparent trends toward this, suggesting that within our experiment, the potential involvement of these genes in vitellogenesis was not regulated by E2-induced changes at the transcriptional level.

Novel estrogen-responsive transcripts. The second most significantly upregulated transcript in fish exposed to 34.4 ng E2/l encoded nots. This is a liver-specific aspartic proteinase normally exclusively expressed in the livers of females and has been linked to the proteolytic cleavage of the vitellogenin precursor (44). Although there are several reports of an increase in nethepsin expression following estrogen exposure in fish (44, 60), it has not been widely considered as a particularly estrogen-responsive gene. The degree of induction observed with RNA-Seq in all fish treated with 34.4 ng E2/l and confirmed in all individuals by RT-QPCR (36,000 fold increase for males) is extraordinary and is in the same order of magnitude as the increase in vgfl expression. This suggests that nethepsin could serve as a useful indicator of estrogen exposure. However, similar to vgfl, there was no increase in nots expression quantified by RT-QPCR or RNA-Seq in the lower treatment groups, suggesting nethepsin may be less sensitive to lower estrogen exposure concentrations compared with esr1 and zona pellucida proteins, at least following this short-term 4-day exposure.

The most downregulated transcript (by 186-fold) in fish exposed to 34.4 ng/l encoded transglutaminase 2-like (tgml2). This strong decrease in expression was confirmed by RT-QPCR (calculated as 250-fold), and tgml2 was also found to be significantly downregulated in mature males exposed to 18.1 ng/l. Transglutaminases are a family of enzymes responsible for a diverse range of posttranslational protein modifications by catalyzing the formation of isopeptide bonds (17). Tgml2 has only been characterized in fish species, and its specific function is not well defined. Here, the striking expression of hepatic tgml2 exclusively in sexually mature male trout suggests a role of this gene in male reproductive function. It is possible that, for example, this role may be similar to that of mammalian prostate transglutaminase tgml4, which is important in the formation and function of seminal fluid and subsequently influences male fertility (9), but this gene has not been characterized in the majority of fish species. Furthermore, human transglutaminases tgml2 and tgml4 are known to have upstream androgen regulatory elements and to be regulated by androgen treatment (12, 25), while gonadal tgml2 was also upregulated following androgen exposure in juvenile female rainbow trout (3). Recently, we also found reduced expression of tgml2 in mature male brown trout exposed to the antiandrogen, linuron (57). Together, this suggests that tgml2 may be regulated by androgens and therefore susceptible to disruption by estrogens, which disrupt the androgen-to-estrogen ratio, highlighting a potentially important endocrine biomarker.

Estrogen-regulated hepatic processes. A number of signaling pathways and processes enriched in the list of differentially regulated transcripts can be broadly related to vitellogenesis. Functional analysis revealed enrichment of lipid transport, and also differential regulation of many other transcripts involved in lipid, fatty acid, and cholesterol metabolism. These processes have been previously associated with vitellogenesis in females and E2 exposure in male fish (34, 60) and are likely to reflect the incorporation of lipids into vitellogenins as they are synthesized in the liver.

Additionally, we found evidence of altered regulation of cellular signaling pathways involved in the regulation of hepatic cellular growth and proliferation, in particular IGF signaling and myelocytomatosis oncogene (MYC) signaling. Previously, vitellogenesis in maturing females and estrogen-exposed males has been extensively linked with cellular growth and proliferation in the liver. Transcripts encoding igfbp types 2 and 5 were upregulated, while those encoding igfbp type 1 were downregulated in fish exposed to 34.4 ng E2/l, suggesting regulation of the transport and bioavailability of IGF1 to bind to its receptors at target cells (24). Cross talk between IGF and estrogen signaling pathways has also been previously demonstrated, and the transcription of igfbps is known to be directly regulated by E2 (20, 29). MYC signaling has been proposed as the dominant regulator of estrogen-induced cellular growth, and estrogen exposure induces myc transcription via upstream enhancer activation (39). We observed an upregulation of myc by up to 10-fold following exposure to 34.4 ng E2/l. Additionally, pathway analysis highlighted its role as a central regulator, alongside esr1, of other differentially expressed genes involved in cell proliferation. Tissue homeostasis depends on a balance between cell death and cell survival, growth and proliferation, which are often controlled by the same interacting signaling pathways, including regulation by MYC and IGFs (41). In parallel, apoptosis was among the downregulated cellular processes, suggesting that E2 exposure induced liver growth and proliferation and suppressed apoptosis.

Differential regulation of processes and transcripts involved in methionine and cysteine metabolism were also observed. This pathway plays an important role in regulating DNA methylation, whereby S-adenosylmethionine (SAM) acts as the key methyl group donor. Modulation of DNA methylation has been implicated in cell proliferation and tumorigenesis and reported to be altered by estrogen exposure (2, 37). Additionally, studies in human cell lines have shown that reactive estrogen metabolites (quinones) bind homocysteine, which is a key intermediate in methionine and cysteine metabolism (14). Plasma concentrations of free homocysteine are also regulated by estrogen and are lower in women of reproductive age (11). Therefore, a reduction in homocysteine might contribute to the
observed differential regulation of these associated metabolic enzymes.

Exposure to E2 also resulted in upregulation of a number of transcripts with roles in transcription and translation, as well as an overrepresentation processes and pathways involved in their regulation. Furthermore, ribosome and endoplasmic reticulum were among the most enriched GO and KEGG pathway terms. Ribosome biogenesis in response to estrogen exposure has been previously linked to increased cell growth and proliferation, reflecting a general upregulation of translation (39, 60). In fish, the observed induction of transcription and translation machinery is also likely to reflect the very significant increase in the synthesis and posttranslational modification of vitellogenins and zona pellucida proteins. Ribosomal constituent overexpression has been previously reported in male zebrafish exposed to E2 (47) and in female vitellogenic livers (60). Together, induction of the expression of growth regulators, and of transcription and translation pathways, illustrates the very significant stimulatory effect of E2 on cell proliferation and protein synthesis in the livers of male brown trout.

A number of processes involved in carbohydrate and amino acid metabolism dominate the overrepresented GO terms in the list of downregulated transcripts. In particular, there was a striking downregulation of transcripts associated with gluconeogenesis, including tyrosine, alanine, aspartate, and glutamate metabolic pathways. Transcripts encoding a key gluconeogenic enzyme, tyrosine aminotransferase (tut), were amongst the most significantly downregulated transcripts in the 34.4 ng E2/l-exposed fish, and this was also confirmed by RT-QPCR analysis. Additionally, a large number of other transcripts involved in gluconeogenesis were also downregulated, including a transcript encoding the rate limiting enzyme phosphoenolpyruvate carboxylase 1 (pck1). A decrease in liver glucose concentration has previously been reported in vitellogenic female fish, as well as immature trout treated with estrogen and has been shown to be primarily associated with reduced gluconeogenesis rather than increased glucose utilization (59). In mammalian studies, estrogen signaling has been widely shown to regulate energy metabolism, including through both glucose and lipid metabolic pathways (8, 22, 46). In the present dataset, there was also evidence of upregulation of lipid synthesis and transport, possibly suggesting a shift toward lipid metabolism as a preferable energy source for vitellogenesis.

Application of RNA-Seq in ecotoxicology. Together with a number of previous studies [see review by Qian et al. (42)], the present study provides evidence that RNA-Seq has very significant potential for mechanistic analysis of chemical exposures in (non-) model organisms and also offers a number of technical advantages over other global methodologies to measure global transcript expression. Here, we successfully identified highly conserved responses to estrogen, compared with other species, together with several more novel estrogen-responsive transcripts. This highlights the potential for RNA-Seq to investigate mechanisms of toxicity for less-studied chemical pollutants and also demonstrates the feasibility of conducting global gene expression profiling in species of environmental interest for which previous sequence information is limited, without the investment required to develop a specific microarray or the need to use surrogate model species.

Analysis of spike-in controls provided strong technical validation for the accuracy of the expression analysis, and the mean calculated dynamic range in expression measured in our experimental data based on the quantified expression of control transcripts was 26,753, which far exceeds that typically found in microarray experiments (up to several hundred fold)(6). The very small number of differentially expressed transcripts between control treatments and also between the control group and groups exposed to 1.9 and 18.1 ng E2/l reflects the ability of the technique to avoid false positives. This is consistent with other studies that have reported that RNA-Seq is accurate and reproducible. However, there is also some evidence that the stringent statistical thresholds imposed during RNA-Seq analysis can reduce its relative sensitivity, particularly for rare transcripts or those with small fold-changes between treatment groups (36, 43, 49). Here, we found evidence that RNA-Seq was less sensitive than RT-QPCR (using more individual fish per treatment) for detecting significant changes in expression of transcripts in the lower concentration groups, although clear trends toward differential regulation were apparent in some cases in the RNA-Seq dataset. It is important to note that we used only three individual fish per treatment group for the RNA-Seq analysis, and this is likely to have contributed to the low sensitivity of the technique in this study.

Using individual fish within an RNA-Seq experimental design provides a considerable advantage for statistical power compared with the use of pooled replicates, and maximizing the number of biological replicates is essential to reduce the impact of biological variation between individuals and likely to increase sensitivity for detection of transcripts with small fold-changes in expression. In the present study we analyzed three individual fish per treatment group, and we would expect sensitivity to improve considerably with the use of more replicates. Maximizing the sequence coverage depth per sample is also likely to increase the sensitivity of RNA-Seq, particularly for rare transcripts, and is therefore equally important to consider. The feasibility of maximizing both the number of replicates and coverage in RNA-Seq experiments is rapidly improving with developments in sequencing technology and is likely to considerably improve sensitivity in the future. Overall, our data highlight the potential of RNA-Seq as a valuable tool in mechanistic ecotoxicology that, crucially, is not reliant on pre-existing genomic resources for the species of interest.

ACKNOWLEDGMENTS

We thank Ronny van Aerle for advice on the bioinformatics and Audrey Farbos and Konrad Paszkiewicz for facilitating the sequencing experiments. We also thank Charles Tyler, Rod Wilson, Erin Reardon, and Jo Rabineau for critical discussions and contributions toward setting up the fish exposure experiments.

GRANTS

This work was supported by a Natural Environment Research Council CASE PhD studentship (Grant NE/I528326/1) and the Salmon & Trout Association. Karen Moore was supported by a Wellcome Trust Institutional Strategic Support Award (WT097835/MF).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: T.M.U.W. and E.M.S. conception and design of research; T.M.U.W. and E.M.S. interpreted results of experiments; T.M.U.W.
REFERENCES


