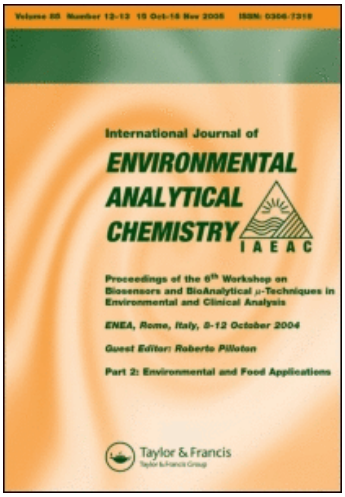

International Journal of Environmental Analytical Chemistry

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713640455

Gene-expression profiling in gill and liver of zebrafish exposed to produced water

Pål A. Olsvik[a]; Kai K. Lie[a]; Anne-Kristin Stavrum[b]; Sonnich Meier[c]
[a] National Institute of Nutrition and Seafood Research, N-5517 Bergen, Norway [b] Department of Informatics, Norwegian Microarray Consortium, Center for Medical Genetics and Molecular Medicine, University of Bergen, N-5008 Bergen, Norway [c] Institute of Marine Research, N-5817 Bergen, Norway

To cite this Article Olsvik, Pål A. , Lie, Kai K. , Stavrum, Anne-Kristin and Meier, Sonnich(2007) 'Gene-expression profiling in gill and liver of zebrafish exposed to produced water', International Journal of Environmental Analytical Chemistry, 87: 3, 195 — 210
To link to this Article: DOI: 10.1080/03067310600959164
URL: http://dx.doi.org/10.1080/03067310600959164

Intern. J. Environ. Anal. Chem.
Vol. 87, No. 3, 15 March 2007, 195–210

Taylor & Francis
Taylor & Francis Group

Gene-expression profiling in gill and liver of zebrafish exposed to produced water

PÅL A. OLSVIK*†, KAI K. LIE†, ANNE-KRISTIN STAVRUM‡ and SONNICH MEIER§

†National Institute of Nutrition and Seafood Research, N-5517 Bergen, Norway
‡Department of Informatics, University of Bergen, Norwegian Microarray Consortium, Center for Medical Genetics and Molecular Medicine, N-5008 Bergen, Norway
§Institute of Marine Research, N-5817 Bergen, Norway

(*Received 17 March 2006; in final form 9 August 2006*)

The effects of produced water (a by-product of oil and gas extraction) on gene expression were studied in gills and liver tissues of zebrafish. Adult, non-breeding zebrafish were exposed to control (freshwater mixed with 5% seawater) or produced water (freshwater mixed with 5% of produced water from the Oceberg C Oil Platform in the North Sea). A zebrafish library was used to make a microarray that consisted of 15,806 unique genes. The results indicate that 27 genes in the gills and 55 genes in the liver show significantly altered expression (greater than two-fold change). More than 70% of these gene sequences have not been annotated in the Gene Ontology (GO) database, making it difficult to characterize the affected genes. CYP1A displayed the greatest upregulation in the gills (eightfold, verified with quantitative real-time PCR). This study illustrates the utility of microarray approaches in investigations of environmental effects of toxicants.

Keywords: Gene expression; Microarray; Produced water; Toxicogenomics; Zebrafish; CYP1A

1. Introduction

Produced water is the largest volume of wastewater arising from offshore oil and gas production with an estimated discharge in Norwegian waters of about 170 million m^3 in 2006 [1]. The discharge comes up with oil and gas from the wells. Produced waters may include inorganic salts, heavy metals, solids, production chemicals, hydrocarbons, benzene, polyaromatic hydrocarbons (PAHs), and, on occasion, naturally occurring radioactive material [2]. The environmental impacts of these compounds depend on the location of the discharge, season, wind conditions, currents, and receiving ecosystems.

*Corresponding author. Email: pal.olsvik@nifes.no

International Journal of Environmental and Analytical Chemistry
ISSN 0306-7319 print/ISSN 1029-0397 online © 2007 Taylor & Francis
DOI: 10.1080/03067310600959164

Toxic compounds present in produced water can generally be found at low concentrations in the marine environment, due to the high dilution rate of offshore discharge. The most important effects of produced water with regard to the environment are PAHs and alkylphenols. PAHs are organic compounds composed of two or more benzene rings, and some are strong carcinogens in humans and fish [3]. The induction of cytochrome P450 is one of the best-characterized biomarkers of PAH exposure [4].

Alkylphenols in produced water have been of special concern, owing to their oestrogenic effects on fish, causing endocrine disruption [5, 6]. With respect to environmental effects, most focus has been on nonylphenols and octylphenols, which are acutely toxic to fish, invertebrates, and algae [7–9]. Hasselberg *et al*. [8] have shown that alkylphenols affect the redox status in first spawning Atlantic cod (*Gadus morhua*) in a sex-dependent manner. Reported responses were increased hepatic total glutathione concentrations in female fish exposed to 0.02 ppm alkylphenol mixture for one week, increased glutathione reductase activities in both male and female fish, and altered glutathione *S*-transferase activities in male fish exposed to a 0.02 ppm alkylphenol mixture for 4 weeks. Arukwe *et al*. [5] studied the effects of nonylphenol on juvenile rainbow trout (*Oncorhynchus mykiss*) and reported transcriptional responses on the oestrogen receptor, vitellogenin, and eggshell zona radiata protein. These and other studies clearly suggest that produced water may have effects on aquatic organisms in areas adjacent to oil- and gas-production platforms around the world.

Microarray technology is now widely used to gain an insight into the genetic basis of many biological mechanisms in fish [10–14]. A genome-wide analysis can reveal new knowledge on the molecular effects of multiple stressors like produced water that contains a number of toxicants. Recent findings suggest that microarray data are more reproducible and therefore more reliable than previously thought [15–17], making this technique an even more promising toxicogenomic tool. As long as no microarray platforms are available for fish species living in areas close to the offshore oil production installations in the North Sea, we decided to use the zebrafish (*Danio rerio*) as a model organism in this preliminary study. A large number of genomic resources are available for this species [18]. The zebrafish is an important model organism for the analysis of developmentally regulated genes [19] but has also been used to study transcriptional effects of environmental stress [20–22].

The aim of this preliminary work was to use a high-throughput microarray technology to screen for transcriptional biomarkers of produced water exposure in fish based on the current information in genomic databases. Until a microarray platform is available for the Atlantic cod *Gadus morhua*, living next to the oil installations in the North Sea, the zebrafish was selected as an experimental species in the search for transcription markers of exposure to produced water. Zebrafish were exposed to a 5% mixture of produced water obtained from one of the largest production installations in the North Sea, the Oceberg C platform. Gill and liver tissues were sampled from exposed and control fish, and extracted total RNA labelled for hybridization to a 16k oligonucleotide array (Compugen Sigma-Genosys oligonucleotides). Three biological replicates were run for exposed and control samples from each tissue, and two-fold or greater regulated genes listed to provide biomarkers for further toxicological studies.

2. Experimental

2.1 *Animals and experimental design*

Zebrafish were obtained from a commercial supplier in Bergen, Norway and kept in the laboratory for several weeks for acclimation before the start of the experiment. Conditioned tap water was used throughout the experiment. Adult fish were maintained in a non-breeding environment at 26°C, and males and females were combined throughout. A semi-static system was used to expose zebrafish to produced water for 96 h. Sixty individuals were transferred to a total of six beakers, each containing 2 L of water oxygenated with air bubbling. Thirty individuals were exposed to a 5% mixture of produced water. The produced water was obtained from the Oceberg C oil platform in the North Sea. As controls, 30 individuals were kept in freshwater. Five per cent pure seawater was added to the control group water, to mimic the salinity of the exposure water. The chemical composition of produced water from the Oceberg C oil production platform is shown in table 1. The water in all six beakers was changed daily. The fish were not fed during the experiment. No test fish died during the experiment. Gill and liver tissues from five fish were pooled to obtain sufficient total RNA for each hybridization reaction (5–10 μg of total RNA was needed). A recent statistical examination of pooling in microarray experiments indicates that pooling can decrease

Table 1. Composition of produced water from the Oceberg C Oil Platform in the North Sea, adapted from Utvik [2].

Component	Unit	Oceberg C
Sum BTEX[a]	$mg\,L^{-1}$	5.8
Sum NPD[b]	$mg\,L^{-1}$	1.60
Naphthalenes	$mg\,L^{-1}$	1.06
Phenanthrenes	$\mu g\,L^{-1}$	76.3
Sum organic acids	$mg\,L^{-1}$	717
Acenaphtylene	$\mu g\,L^{-1}$	5.1
Fluorene	$\mu g\,L^{-1}$	2.7
Fluoranthene	$\mu g\,L^{-1}$	7.8
Pyrene	$\mu g\,L^{-1}$	8.6
Chrysene	$\mu g\,L^{-1}$	0.4
Benzo(a)pyrene	$\mu g\,L^{-1}$	0.1
Benzo(a)anthracene	$\mu g\,L^{-1}$	1.9
Benzo(ghi)perylene	$\mu g\,L^{-1}$	0.1
Benzo(k)fluoranthene	$\mu g\,L^{-1}$	0.2
Sum phenols	$mg\,L^{-1}$	10.96
Barium	$mg\,L^{-1}$	142
Iron	$mg\,L^{-1}$	7.7
Mercury	$ng\,L^{-1}$	26
Zinc	$\mu g\,L^{-1}$	340
^{214}Pb	$Bql\,L^{-1}$	7
^{214}Bi	$Bql\,L^{-1}$	6
^{228}Ac	$Bql\,L^{-1}$	<2
^{212}Bi	$Bql\,L^{-1}$	<2
^{212}Pb	$Bql\,L^{-1}$	<2
^{226}Ra	$Bql\,L^{-1}$	7

[a]Benzene, toluene, ethylbenzene, and xylene.
[b]Naphthalenes, phenanthrenes, and dibenzo-thiophenes.

the number of arrays required without loss of precision [23]. Three control samples and three exposed samples (biological replicates) were analysed from both tissues.

Due to the limited amounts of total RNA available, parallel samples had to be used for the qRT-PCR validation analysis; G1, G4, G5 (control gill), G10, G11, G12 (exposed gill), L1, L4, L6 (control liver), L8, L10, and L12 (exposed liver). Samples G2, G3, G6 (control gill), G7, G8, G9 (gill exposed), L2, L3, L5 (control liver), L7, L9, and L11 (exposed liver) were used for microarray analysis. Because of this procedure, we did not expect to obtain exactly the same results with qRT-PCR as with microarray analysis. However, the expression of the greatest up- and down-regulated genes should be in general agreement, even if they are measured in parallel biological samples.

2.2 *Tissue sampling and RNA extraction*

Pooled tissues from five animals were used for RNA extraction. Samples from gills and liver were dissected out and immediately transferred to RNA later (Ambion, Austin, TX) and stored at −20°C before RNA extraction. Total RNA was extracted from tissues using Trizol reagent (Invitrogen, Life Technologies, Carlsbad, CA), according to the manufacturer's instructions and stored in 50 μL of RNase-free MilliQ H_2O. Genomic DNA was eliminated from the samples by a DNase kit from Ambion (Cat. 1906) according to the manufacturer's descriptions. The RNA was then stored at −80°C before further processing. The quality of the RNA was assessed with the NanoDrop® ND-1000 UV-Vis spectrophotometer (NanoDrop Technologies, Wilmington, DE) and the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). A 260/280 nm absorbance ratio of 1.8–2.0 indicates a pure RNA sample. The RNA 6000 Nano LabChip® kit (Agilent Technologies) was used to evaluate the integrity of the RNA.

2.3 *Microarray design*

The zebrafish library obtained from Compugen/Sigma-Genosys (Jamesburg, NJ) consisted of 16,399 oligos (65 nt) representing 15,806 unique genes (LEADS clusters) as indicated by the most recent Compugen annotation of December 2005 (http://www.labonweb.com/oligo/). According to Mathavan *et al*. [19], about 99% of the probes have non-redundant GenBank entries, and the rest of the GenBank entries are duplicated. A description based on GO (Gene Ontology) terms are given for about 2300 probes, and the remainder are indicated as 'GO unknown' [19]. The oligos were modified 5′ amino C6 linker. The library included 172 zebrafish β-actin internal control oligos distributed over the entire library (approximately 4/384 well plate) used as calibration spots. The oligo probes were spotted and immobilized on Amersham CodeLink™ activated glass slides that were prepared using a hydrophilic polymer contain *N*-hydroxysuccinimide ester groups, and this polymeric coating was attached covalently to the silane basecoat by the vendor.

The oligos were resuspended in 50 mM phosphate buffer pH 8.0 at 20 mM concentration. The individual oligo-probe was printed on slides under 45% humidity using a GeneMachine OmniGrid™ 100 Microarrayer in 4×12 pin configuration and 20×20 spot configuration of each subarray. The spot diameter was 100 μm, and the distance from centre to centre was 200 μm. The printed zebrafish microarrays were

further chemically covalently coupled at 70% humidity overnight. The microarrays were ready for sample hybridization after additional 100 mM taurine/bicine blocking and 4× SSC/0.1% SDS washing steps.

2.4 *cRNA preparation, hybridization, and scanning*

The biotin-labelled cRNA target was prepared by a linear amplification method. Poly(A) RNA from 5 μg of total RNA was primed for reverse transcription by a DNA oligonucleotide containing a T7 RNA polymerase promoter 5′ to a d(T)24 sequence. After second-strand cDNA synthesis, the cDNA serves as the template in an *in vitro* transcription (IVT) reaction to produce the target cRNA. The IVT was performed in the presence of biotinylated nucleotides to label the target cRNA.

First-strand cDNA was synthesized by incubation of 5 μg of total RNA with T7 primer at 70°C for 10 min. Superscript reverse transcriptase was added to the reaction mix and incubated at 37°C for 1 h. Second-strand cDNA was made by adding first-strand cDNA to the polymerase mix and incubation in a Tropicooler incubator at 16°C for 2 h. Double-stranded cDNA was precipitated in isopropanol with glycogen and ammonium acetate at −80°C for 20 min and centrifuged at 12,000 × *g* for 20 min. The supernatant was removed and the pellet air-dried for 30 min. cRNA was synthesized by *in vitro* transcription using the Ambion MEGAscript T7 kit. Biotinylated UTP, dNTP, and enzyme reaction mix were added to dried cDNA and incubated at 37°C overnight. Biotin-labelled cRNA was purified using the Qiagen RNeasy kit. The quality and quantity of cRNA was assessed by an Eppendorf Biophotometer. The target labelling procedure resulted in a 50–200-fold amplification of the input poly(A) RNA, and an A260/A280 ratio between 1.8 and 2.0.

Biotin-labelled cRNA, 10 μg of each target was used for hybridization on each 16K zebrafish oligo expression microarray on a TECAN HS4800 Hybridization Station. The microarrays were hybridized in 6× SSPE buffer with 50% formamide at 37°C for 20 h, washed in 0.75× TNT buffer at 46°C for 1 h, followed by blocking in TNB buffer at RT for 30 min, and then processed by using direct detection method of the biotin-containing transcripts by Streptavidin-Alexa 647 conjugate in TNB buffer (1 : 500) at RT for 30 min. Stained chips were washed in 1× TNT wash buffer at room temperature for 60 min and changed with fresh TNT buffer every 15 min.

Processed chips were scanned with a Perkin Elmer ScanArray® XL5000 Scanner, software version 3.1 with the laser set to 635 nm, at a power of 90 and PMT 70-50 setting, and a scan resolution of 10 μm. Images were quantified using Perkin Elmer QuantArray® Software 3.0. Total intensities were quantified with the fixed circle method. The confidence calculation was the weighted average. The gene-expression technology applied has the sensitivity of one to three copies per cell and a dynamic range of 2.5–3 orders of magnitude. The technology specificity is 90–94%, allowing any highly homologous genes within a gene family to be distinguished. cRNA preparation, hybridization, and scanning were done at the Kimmel Cancer Center Microarray Core Facility at the Thomas Jefferson University, Philadelphia, and vendors are given at their web page.

2.5 *Microarray data analysis*

For the pre-processing steps, we removed all control spots. In-array replicates were combined, and logs 2 of background-subtracted signals were used. The data were normalized either by global mean normalization or by quantile normalization [24]. The mean normalization divides all values for an array by the mean value for the same array. This means that the mean value is the same for all arrays. β-Actin has proven unreliable as an endogenous control in gene-expression experiments [25], but since all genes were used to calculate the mean signal, the contribution from β-actin is minimal. Quantile normalization makes the intensity distribution the same across all arrays. Both methods are used to make the arrays comparable. Lists of genes showing different levels of expression between control groups and exposed groups were produced for both gills and liver. The UniGene names of the top BLASTN hits (E value cutoff 10^{-4}) and biological function of the top BLAST hits are presented in tables as given by the most recent annotations. To ensure we selected genes that display significantly different gene-expression levels between the two treatments, while the in-group variance was low, we used two different selection criteria; fold change and p-values. The p-values were created by first calculating the two sample t-scores for each gene and then permuting of the sample labels, recalculating the t-scores for each permutation. Counting the number of t-scores from the permutations with a higher score than that which we had originally gives us the p-value for each gene. We used a significance level of 0.05 when looking for differentially expressed genes. The genes in the final lists were subjected to a BLAST search and also analysed using GO. The resulting gene lists have at least a two-fold difference in expression between exposed and control groups and a p-value below 0.05. The analysis using mean normalization was done using the software Genespring from SiliconGenetics (Agilent Technologies, Palo Alto, CA), and the analysis using quantile normalization was done using the software J-Express Pro from MolMine AS (MolMine AS, Bergen, Norway). The resulting gene lists contain the genes that were significantly up- or down-regulated irrespective of the normalization method.

2.6 *Real-time PCR*

Eight transcripts that were up- or down-regulated more than two-fold by microarray analysis in gills or liver were selected for real-time qRT-PCR verification analysis. qRT-PCR assays were designed using Primer Express 2.0 software (Applied Biosystems, Foster City, CA) and sequences from known zebrafish genes. The PCR primers that were used to quantify the transcript levels of the genes of interest were obtained from Invitrogen and are listed in table 2. In addition, a qRT-PCR assay for β-actin was designed for normalization of the transcription data. Accession numbers and amplicon sizes of the PCR products are shown in table 2. A two-step real-time RT-PCR protocol was developed to measure the mRNA levels of the studied genes in gill and liver tissues of zebrafish. Two-fold serial dilution curves of total RNA were used for RT and PCR efficiency calculations. The five serial dilutions and all samples for each gene were run on the same 96-well plate. The RT reactions were run in triplicate on 96-well reaction plates with the GeneAmp PCR 9700 machine (Applied Biosystems, Foster City, CA) using TaqMan Reverse Transcription Reagent containing Multiscribe Reverse Transcriptase ($50\,U\,\mu L^{-1}$). Reverse transcription was performed at 48°C for 60 min by using oligo dT primers (2.5 μM) in a total volume of 30 μL.

Table 2. PCR primer sequences used to amplify a selected number of zebrafish transcripts found to be more than two-fold up- or down-regulated in gill or liver tissues with microarray examination (real-time PCR verification data are shown).

					Gills		Liver	
Gene	Accession no.	Forward primer	Reverse primer	Amplicon size (bp)	Array	qRT-PCR	Array	qRT-PCR
CYP1A1 (EST 1)	AF057713	TGTTATCTGACAGGGCGAACCT	TGACGGAAGATCTCCAGAATGA	61	8.6	11.1		
CYP1A1 (EST 2)	AW342687	CACTCTTCATCCTGATCTGTGCTT	GCATGGCCTTGTCCTTTCAA	91	7.2	12.7		
Vitellogenin 3 precursor	AF254638	AAGCAAAACCCCTTGGCTATG	TGTCGTCTTTTTGGGCAGTTG	63	−2.9	−1.9		
MAP kinase-interacting serine/threonine kinase 2	BI839784	CACACGCCAAAGGTCTAATACCT	CGTGGAAAAAGCACAAAAAGG	63			−3.1	−1.8
DNA repair protein RAD51	AW184428	CGTCTCGCTGATGAGTTTGG	CCACCTGTGCTACAACCTGGTT	61			2.7	−2.1
Zinc finger protein	BM095323	GCCTCTCGCCTCTGTGCTTA	ACGCAAGCGCAAGAAGAGAT	101			2.7	−1.9
Sestrin 1	BI878036	CCTGCCATCATCAGAATGCA	TTCGAGTGAGCGTGTGTGTTC	106			−3.0	−1.7
Arginase	BF717769	ACCTCTTGGCAAGGAAGCAA	TGTAGGTCAGTCCGCCGTTAA	111			−2.6	−1.5
β-Actin	AF025305	TGACCGAGAGAGGCTACAGCTT	CCTTGATGTCACGGACAATTTC	63				

The input RNA concentration was 250 ng in each reaction. The final concentration of the other chemicals in each RT reaction was: $MgCl_2$ (5.5 mM), dNTP (500 μM of each), 10× TaqMan RT buffer (1×), RNase inhibitor (0.4 U μL^{-1}) and Multiscribe Reverse Transcriptase (1.67 U μL^{-1}). cDNA from each RT reaction (0.5 μL) was transferred to a new 96-well reaction plate, and the real-time PCR run on the ABI Prism 7000 Sequence Detection System from Applied Biosystems (Applied Biosystems, Foster City, CA). Real-time PCR was performed by using QuantiTect SYBR Green PCR Master Mix (Qiagen, Chatsworth, CA), according to the manufacturer's instructions.

The baseline and threshold for Ct calculation were set automatically or manually whenever necessary with the ABI Prism 7000 SDS software version 1.1. (Applied Biosystems, Foster City, CA), and the mean normalized expression was calculated using the Microsoft Excel-based software Q-Gene. The Q-Gene tool was developed to manage and expedite the entire experimental process of quantitative real-time RT-PCR and is available free from the BioTechniques Software Library [26]. β-Actin was used as an endogenous control in the final calculations of mean normalized expression.

3. Results and discussion

3.1 *General findings*

The main purpose of most microarray experiments is to find contrasting gene-expression levels across tissues and treatments of a chosen subset of the genome, and this is the most cost-effective method for monitoring the relative levels of expression of multiple genes in parallel. The approach replaces hypothesis-driven research with discovery-driven research. The ultimate goal of this study was to elucidate molecular mechanisms involved in cellular responses to stress caused by produced water; to identify toxicant-specific and adverse-effect-specific patterns of gene expression and to develop gene-expression-based biomarkers of this kind of stress in teleosts. Using a significance level of 0.05, we found genes to be significantly differentially expressed with fold change values down to 1.65. With the combined use of two normalization strategies and the described filtering method, only 27 transcripts in gills and 55 transcripts in liver were found to have a two-fold or greater altered expression (tables 3 and 4). Approximately the same number of transcripts were found differentially expressed in both tissues applying only one single normalizing technique (either global mean normalizing or quantile normalizing, data not shown). Even with a less stringent filtering method (*t*-test *p*-value less than 0.05, data not shown), the experiment clearly suggests that only a few of the studied transcripts were affected by the treatment. More than 70% of the differentially expressed genes have currently not been annotated, making it difficult to obtain meaningful gene ontology (GO) results from this experiment. This will most likely be a common problem in contemporary teleostean fish microarray experiments, until more fish genes are annotated. The GO has three organizing levels, molecular function, biological process, and cellular component (see http://www.geneontology.org). A gene product can have one or more molecular functions and be used in one or more biological processes; it might be associated with one or more cellular components. We used the most recent putative annotations for the clones in the array provided from a database set up by the Genome Institute of

Table 3. Transcripts greater than two-fold up- or down-regulated in gills of zebrafish exposed to 5% produced water mixture.[a]

EST	UniGene name of top BLAST hit	Biological function of top BLAST (N) hit	Mean fold change
AF057713	Cytochrome P450, subfamily I, polypeptide 1	Electron transport, response to chemical stimulus	8.5
AW342687	Cytochrome P450, subfamily I, polypeptide 1	Electron transport, response to chemical stimulus	7.2
AW567349	Wu:fc49d01		6.9
AW184205	Wu:fj10e08		3.6
AW777978	Transcribed locus		2.9
BM183152	CDNA clone IMAGE:7177046		2.9
AW184695	Wu:fj17g10		2.8
AW232249	Transcribed locus		2.7
AI584347	Wu:fb92d12		2.6
BM024762	H2A histone family, member X	DNA repair, DNA recombination, nucleosome assembly, response to DNA damage stimulus, chromosome organization and biogenesis, cell cycle, meiosis	2.4
AW059068	Wu:fe24f11		2.4
BM005368	Transcribed locus		2.4
AI397328	CDNA clone IMAGE:7228590, containing frame-shift errors		2.3
AW154269	Wu:fa20f04		2.3
AW566603	Zgc:77469	Transcription, regulation of transcription, DNA-dependent	2.2
BG727181	Zgc:77038	Intracellular signaling cascade	2.2
BI891793	Si:rp71-46j2.8		2.1
AI584199	Ankyrin repeat domain 13		2.1
AW232738	Ladinin		−2.0
AI544475	Wu:fb75d10		−2.0
AF295407	Alcohol dehydrogenase 8a	Alcohol metabolism, ethanol metabolism	−2.1
AI330535	Wu:fa92h10		−2.1
AW077961	Annexin A1b		−2.3
AW280088	Wu:fj49c06		−2.3
AF246162	Immunoglobulin light lota variable 2, s1		−2.5
BG727389	Wu:fc18h11		−2.9
AF254638	Vitellogenin 3, phosvitinless	Lipid transport	−2.9

[a]UniGene names are given with a BLAST cutoff of $E < 10^{-4}$. The biological function of the top BLAST hit is given according to biological processes reported in the Gene Ontology database.

Singapore and described by Mathavan *et al.* [19] (http://giscompute.gis.a-star.edu.sg/~govind/zebrafish/index.html). Obtaining the latest (February 2006) terms using the Annotation Database http://www.genetools.no/, we found five cellular components (GO:0005575), seven molecular functions (GO:0003674), and seven biological processes (GO:0008150) for the 27 genes affected by produced water treatment in the gills. Twenty of the 27 genes in gills had no GO terms associated with them. In liver, the corresponding values were 11 cellular components, 15 molecular functions, and 14 biological processes. For 40 of the 55 altered genes in liver, no GO terms were found. Thus, this experiment yielded less information on biological processes in zebrafish affected by produced water than anticipated with the selected two-fold change cutoff limit.

Table 4. Transcripts greater than two-fold up- or down-regulated in liver of zebrafish exposed to 5% produced water mixture.[a]

EST	UniGene name of top BLAST hit	Biological function of top BLAST (N) hit	Mean fold change
BI867265	Transcribed locus		2.8
AI721316	Transcribed locus, weakly similar to XP_514061.1 PREDICTED: hypothetical protein XP_514061 (*Pan troglodytes*)		2.8
AW184428	RAD51 homologue (RecA homologue, *E. coli*) (*S. cerevisiae*)	DNA metabolism, DNA repair, response to methylmercury	2.7
BE017642	Wu:fa04c02		2.7
BM095323	Transcribed locus, weakly similar to XP_134736.4 PREDICTED: zinc finger protein 26 (*Mus musculus*)		2.7
BG304241	Spindle assembly 6 homologue (*C. elegans*)	Lipid transport, cell cycle, lipoprotein metabolism	2.5
BM153986	Wu:fb54a03		2.5
BM181758	Ankyrin repeat and SOCS box-containing 3	Intracellular signalling cascade	2.5
AW154091	Wu:fi22d01		2.5
AW117161	Sb:cb541		2.4
BI980240	Transcribed locus		2.4
AI878452	Wu:fb74d05		2.4
BI877478	Zgc:86634	Protein modification	2.4
AW466858	Zgc:55868	Lipid metabolism-catabolism, signal transduction, intracellular signalling cascade	2.4
BG985689	ADP-ribosylation factor interacting protein 1 (arfaptin 1)		2.3
BM156974	Zgc:55580	Regulation of translational initiation	2.3
AW343858	Wu:fb93c06		2.3
AW343880	Wu:fi69e09		2.2
BI878798	Wu:fc47g09		2.1
BI877690	CDNA clone IMAGE:7265391		2.1
BI673416	Acidic (leucine-rich) nuclear phosphoprotein 32 family, member B		2.1
BG303462	Dipeptidylpeptidase 3	Proteolysis	2.1
BM095379	Wu:fb18f07		2.1
BI845475	Transcribed locus		2.1
AW305657	Proteasome (prosome, macropain) 26S subunit, ATPase, 1a	Protein catabolism	2.1
AI793812	Wu:fc54d09		2.0
BE201450	Fructose-1,6-bisphosphatase 1, like	Carbohydrate metabolism	−2.0
AA495032	Transcribed locus	Nicotinate phosphoribosyltransferase activity	−2.1
AI958534	Wu:fa97h07		−2.2
BI866461	Does not exist in UniGene		−2.2
BI842159	Does not exist in UniGene		−2.2
BG728432	Sulfotransferase family 1, cytosolic sulfotransferase 4		−2.2
BI534277	Transcribed locus		−2.2
AW279660	Does not exist in UniGene		−2.3
BM185913	Wu:fk57a03		−2.3
AI957474	Transcribed locus		−2.3
AW154478	Wu:fc16e03		−2.3

(*Continued*)

Table 4. Continued.

EST	UniGene name of top BLAST hit	Biological function of top BLAST (N) hit	Mean fold change
AW077599	Does not exist in UniGene		−2.3
BG727449	Transcribed locus, moderately similar to NP_659189.1 urocanase domain containing 1 (*Mus musculus*)		−2.3
AW421062	Zgc:86658		−2.4
BG728516	Wu:fa55d05		−2.4
AW077452	Does not exist in UniGene		−2.4
BG305916	Transcribed locus, weakly similar to XP_413787.1 PREDICTED: similar to aquaporin 9 (*Gallus gallus*)		−2.5
BF717769	Arginase, type II	Arginine catabolism, catalytic activity	−2.6
BI886811	Does not exist in UniGene		−2.6
BI877640	Does not exist in UniGene		−2.6
AW826859	Transcribed locus		−2.8
BG727600	Transcribed locus		−2.9
BM005448	Does not exist in UniGene		−3.0
AW019245	Transcribed locus		−3.0
BI982104	Retinol binding protein 4, plasma	Transport	−3.0
BI885475	Transcribed locus, weakly similar to NP_001009668.1 alpha-ETF (*Rattus norvegicus*)		−3.0
BI878036	Zgc:91970	Cell-cycle arrest	−3.0
BI839784	MAP kinase-interacting serine/threonine kinase 2	Protein amino-acid phosphorylation	−3.1
BI878503	Wu:fj65h10		−3.2

[a]UniGene names are given with a BLAST cut off of $E < 10^{-4}$. The biological function of top BLAST hit is given according to biological processes reported in the Gene Ontology database.

We were unable to determine the variation between individual samples, because mRNA had to be pooled from five animals to provide sufficient material for each hybridization. Evidence appears to be conflicting over the effects of pooling [27, 28]. The variation induced by produced water may be masked by non-stimuli-related differences in the individuals biological state. Kendziorski *et al*. [23] studied the effects of mRNA pooling on estimates of gene expression and suggested that a pooling strategy can decrease the number of arrays required in an experiment without loss of precision.

To search for new genes or patterns, we included all genes with a significant *p*-value less than 0.05. This gave us a list of genes with fold change values down to 1.65. Tanaka *et al*. [29] found that significant differences were reproducible for many differences in the range of 1.2–2-fold, whereas Mootha *et al*. [30] introduced the Gene Set Enrichment Analysis, developed to detect modest but coordinate changes in expression of groups of functionally related genes. By lowering the fold limit, 12 cellular components, 22 molecular functions, and 20 biological processes were affected in the gills by the treatment with the most recent annotation. The corresponding numbers for liver were 63 cellular components, 75 molecular functions, and 76 biological processes. Even with this lower fold expression change limit, only one KEGG pathway was affected in

gills (tryptophan metabolism, lysine degradation, and benzoate degradation via CoA ligation) and one in liver (phosphatidylinositol signalling system). This finding underscores the fact that many genes sequenced in fishes are currently unknown. table 5 provides an overview of affected clusters in gills and liver based on biological processes, and some of the most interesting affected processes. With a minimum cluster size of 5, many biological processes were affected, especially in the liver. Detoxification (increased glutathione *S*-transferase activity) seems to have been heavily affected by the treatment in liver tissue.

3.2 *Differentially expressed genes in gills*

Produced water, wastewater originating from the reservoirs during oil production, contains a number of toxicants, both heavy metals and organic chemicals (table 3), known to affect aquatic organisms [5, 6, 8, 9, 31]. CYP1A, a gene encoding a P450 monooxygenase protein that catalyses the oxidation of a number of organic chemicals [4, 7], displayed greatest upregulation in gills. The array contained two different oligonucleotide sequences for this gene, designed from GenBank accession nos AF057713 and AW342687. The mean fold upregulation for these two sequences was 8.6 and 7.2, respectively. Produced water contains high levels of polyaromatic hydrocarbons (PAHs) and alkylphenols, known inducers of CYP1A, thus verifying that the exposure experiment profoundly affected the fish in the predicted way.

Except for one unknown gene that was 6.9 times up-regulated in gills (AW567349), only minor changes (less than 3.6 times changed expression) were found for the other genes on the list. Vitellogenin 3 was significantly down-regulated in gills of exposed zebrafish. It was surprising to find that a gene encoding a vitellogenin is expressed at all in gill tissue. Wang *et al*. [32] examined the expression of vitellogenin 3 in eight tissues of adult zebrafish. They found this gene to be expressed mainly in liver, but also weakly in the intestine. No expression was found in gill tissue with the applied Northern blot technique. According to a UniGene search performed in July 2006, vitellogenin 3 has been found expressed in adult intestinal, skin and kidney tissue of zebrafish, as well as in liver tissue (http://www.ncbi.nlm.nih.gov). It is unlikely that the expression stems from contamination from other tissues, since this gene was found to be evenly present in a total of 12 independent gill samples.

Considering molecular functions, several interesting genes with differentially expression were found with a lower fold change cutoff (e.g. 1.65 as described in the previous section). In gills, a gene encoding a protein that interacts selectively with heat-shock protein (AF295376) was found to be up-regulated in the exposed group, suggesting that the exposure had triggered a general stress response. Three genes encoding ATP-binding proteins were also found to be up-regulated in the gills of the exposed group: MAP kinase-interacting serine/threonine kinase 2 (AI657551), hypoxia up-regulated 1 (BI876732) and MCM5 minichromosome maintenance deficient 5 (AW058902). In addition, two genes encoding ion-binding proteins, manganese: arginase, type-II (BF217769) and zinc: alcohol dehydrogenase 8a (AF295407) were found to be up-regulated in the gills.

Table 5. Biological processes affected by produced water treatment in gills and liver as found by clustering (minimum cluster size 5, unknown clusters omitted), analysed with a 1.65-fold change limit, with zebrafish GO annotation as of February 2006.

Gill clusters	Biological process	Liver clusters	Biological process
Nucleus	Small nucleolar ribonucleoprotein complex	Integral to membrane	
DNA binding	Damaged DNA binding	Membrane	Apical membrane complex
			Gap junction
		Cytoplasm	Mitochondrial inner membrane
			Nucleus
		Nucleus	
		Intracellular	Microtubule
			Myosin
			Actin filament
		Cellular component unknown	
		Molecular function unknown	
		ATP binding	
		GTP binding	
		DNA binding	Damaged DNA binding
		Structural molecule activity	Structural constituent of cytoskeleton
		Transferase activity	Beta DNA polymerase activity
			Protein-tyrosine kinase activity
			Cyclin-dependent protein kinase activity
			Panthothenate kinase activity
			Kinase activity
			Transaminase activity
			Glutathione transferase activity
			Nicotinate phosphoribosyltransferase activity
			Glycogen synthase activity
		Electron transport	
		Metabolism	Lipid metabolism
			Regulation of transcription, DNA dependent
			Regulation of transcription
			Nucleotide-sugar metabolism
			Biosynthesis
			DNA recombination
			DNA replication initiation
			Nuclear mRNA splicing, via spliceosome

3.3 *Differentially expressed genes in liver*

In liver, fold change differences ranged from −3.2 to 2.8 (table 4). Using the putative annotations for the clones in the array provided from the database set up by the Genome Institute of Singapore [19], very few of the listed genes are annotated. Worth mentioning from a biomarker perspective is the DNA repair gene RAD51, which was found to be up-regulated 2.7-fold. However, we were unable to confirm this up-regulation by qRT-PCR analysis, which suggests that there are false positives in the array data (an alpha level of 0.05 indicates that up to 5% of the genes identified as being regulated are not). Several of the other annotated listed genes suggest a reduced

metabolism in the liver of exposed fish. It is also worth mentioning that CYP1A was not up-regulated in liver after 96 h of water-borne exposure to substances that up-regulated this gene in the gills approximately 10-fold. One possible explanation for this discrepancy between gill and liver response may be that there was enough CYP1A protein in liver to catalyse the oxidation of the organic chemicals present in produced water, and no new transcription was necessary. Another explanation may be that the observed differences in effects on gills and liver are due to a time-dependent response. Further studies should therefore include time series to evaluate the kinetics and temporal effects of produced water on teleosts.

The following interesting genes were up-regulated in liver tissue in the exposed group as discovered with a lower fold change cutoff (accession number in parentheses): mitogen-activated protein kinase 14a (AB030897), transcription factor activity regulating gene Zgc:85857 (BM101515), translation initiation factor activity regulating gene Zgc:5580 (BM156974), three calcium-binding genes interacting selectively with Ca^{2+}, member 12 (BM182849), Zgc:55868 (AW466858) and met proto-oncogene (hepatocyte growth factor receptor) (AB255680), one iron-metal-binding gene interacting selectively with Fe ions (AI957754), glutathione *S*-transferase gene Zgc:101897 (BI979918), five oxidoreductase activity regulating genes; cysteine dioxygenase, type I (AI957754), cytochrome P450, family 2, subfamily J, polypeptide 2, A (AW059252), cytochrome-c oxidase activity gene Zgc:73355 (BI879411), phosphoglycerate dehydrogenase activity gene Zgc:65956 (AI883922), mono-oxygenase activity gene Zgc:66494 (AW116567) and finally two ubiquitin-conjugating enzyme acting genes; ubiquitin-conjugating enzyme E2L 3, like (AI444365) and ubiquitin-conjugating enzyme E2G 2 (AW175187). Overall, these findings suggest that exposure to produced water induced genes known to be affected by both inorganic and organic toxicants, clearly indicating that both metals and organic chemicals in the produced water can potentially harm fish. Since all of the genes mentioned in this section can be expected to be affected by exposure to produced water, this finding clearly illustrates that it can be useful to set a lower fold expression change cutoff than the generally accepted two-fold limit in toxicogenomic examinations.

3.4 *Real-time PCR validation*

Eight transcripts were selected for qRT-PCR confirmation, in addition to β-actin, which was used for calculation of mean normalized expression. Preference was given to annotated transcripts with a relative high fold difference in expression between exposed and non-exposed gill and liver samples. In general, the qRT-PCR data and the array data were in line regarding up- or down-regulation (gills: CYP1A EST1↑, CYP1A EST2↑, vitellogenin 3↓; and liver: MAP kinase↓, sestrin 1↓ and arginase↓) (table 2). For RAD51 and Zinc Finger Protein (ZFP), the qRT-PCR data suggested down-regulation contrary to up-regulation, as suggested by the array data. The two ESTs for CYP1A displayed the greatest up-regulation in the gills (11× and 12×, respectively), compared to 8× and 7× as suggested by the microarray analysis. qRT-PCR also verified the array results for vitellogenin 3 in gills. The qRT-PCR results therefore suggest that vitellogenin 3 might be expressed in gill tissue, and that the array findings are not false positives resulting from oligo cross-reaction with another mRNA. The minor differences in fold change seen for six of the genes are probably due to the

fact that we had to analyse parallel, but not the same, biological replicates with the qRT-PCR technique. Several studies have reported that qRT-PCR is more sensitive than microarray analysis and detect greater changes, but also results in higher variability among biological replicates [33]. In most microarray experiments, it is typical to find discrepancies between qRT-PCR data and array data for a limited number of genes [14].

4. Conclusions

Twenty-seven transcripts in gills and 55 transcripts in liver were found to be differentially expressed in zebrafish exposed to a 5% mixture of produced water for 96 h compared to control fish using a two-fold expression change limit. This was fewer than expected after exposure to a water-borne mixture containing a number of known toxicants. Not surprisingly, CYP1A was up-regulated to the highest degree in gills of zebrafish exposed to produced water. This study illustrates the utility of microarray technology in investigations of environmental effects of toxicants, but also highlights the biological limitation of this technology due to the high number of sequences that have no homology to known sequences or have unknown function in fishes.

Acknowledgements

We would like to thank Jennifer H. Ma, and the Microarray Core Facility, Kimmel Cancer Center, Thomas Jefferson University, US, for help with array hybridization and crude-data analysis. We acknowledge the data-analysis support provided by the Norwegian Microarray Consortium (NMC), a national FUGE technology platform, funded by the FUGE programme of the Research Council of Norway. This work was funded by the National Institute of Nutrition and Seafood Research, Bergen, Norway.

References

[1] Norwegian Ministry of Petroleum and Energy. *Report*: *Ministry of Petroleum and Energy 2005*. Oslo, **1** (2005).
[2] T.I.R. Utvik. *Chemosphere*, **39**, 2593 (1999).
[3] M.S. Myers, L.L. Johnson, O.P. Olson, C.M. Stehr, B.H. Horness, T.K. Collier, B.B. McCain. *Mar. Pollut. Bull.*, **37**, 92 (1998).
[4] T.D. Bucheli, K. Fent. *Crit. Rev. Environ. Sci. Technol.*, **25**, 201 (1995).
[5] A. Arukwe, S.W. Kullman, D.E. Hinton. *Comp. Biochem. Physiol.*, **129C**, 1 (2001).
[6] L. Hasselberg, S. Meier, A. Svardal, T. Hegelund, M.C. Celander. *Aquat. Toxicol.*, **67**, 303 (2004).
[7] M.R. Servos. *Water Qual. Res. J. Can.*, **34**, 123 (1999).
[8] L. Hasselberg, S. Meier, A. Svardal. *Aquat. Toxicol.*, **69**, 95 (2004).
[9] J.-P. Cravedi, D. Zalko. Metabolic fate of nonylphenols and related phenolic compounds in fish. In *Environmental Toxicology*, T.P. Mommsen and T.W. Moon (Eds), pp. 153–169, Elsevier, Amsterdam (2005).
[10] A.Y. Gracey, J.V. Troll, G.N. Somero. *Proc. Natl. Acad. Sci.*, **98**, 1993 (2001).
[11] T.D. Williams, K. Gensberg, S.D. Minchin, J.K. Chipman. *Aquat. Toxicol.*, **65**, 141 (2003).
[12] P. Larkin, I. Knoebl, N.D. Denslow. *Comp. Biochem. Physiol.*, **136B**, 149 (2003).

[13] J.E.S. Podrabsky, G.N. Somero. *J. Exp. Biol.*, **207**, 2237 (2004).
[14] S.E. Hook, A.D. Skillman, J.A. Small, I.R. Schultz. *Aquat. Toxicol.*, **77**, 372 (2006).
[15] T. Bammler, R.P. Beyer, S. Bhattacharya, G.A. Boorman, A. Boyles, B.U. Bradford, R.E. Bumgarner, P.R. Bushel, K. Chaturvedi, D. Choi, M.L. Cunningham, S. Deng, H.K. Dressman, R.D. Fannin, F.M. Farin, J.H. Freedman, R.C. Fry, A. Harper, M.C. Humble, P. Hurban, T.J. Kavanagh, W.K. Kaufmann, K.F. Kerr, L. Jing, J.A. Lapidus, M.R. Lasarev, J. Li, Y.J. Li, E.K. Lobenhofer, X. Lu, R.L. Malek, S. Milton, S.R. Nagalla, J.P. O'malley, V.S. Palmer, P. Pattee, R.S. Paules, C.M. Perou, K. Phillips, L.X. Qin, Y. Qiu, S.D. Quigley, M. Rodland, I. Rusyn, L.D. Samson, D.A. Schwartz, Y. Shi, J.L. Shin, S.O. Sieber, S. Slifer, M.C. Speer, P.S. Spencer, D.I. Sproles, J.A. Swenberg, W.A. Suk, R.C. Sullivan, R. Tian, R.W. Tennant, S.A. Todd, C.J. Tucker, B. Van Houten, B.K. Weis, S. Xuan, H. Zarbl. *Nat. Meth.*, **2**, 351 (2005).
[16] R.A. Irizarry, D. Warren, F. Spencer, I.F. Kim, S. Biswal, B.C. Frank, E. Gabrielson, J.G. Garcia, J. Geoghegan, G. Germino, C. Griffin, S.C. Hilmer, E. Hoffman, A.E. Jedlicka, E. Kawasaki, F. Martinez-Murillo, L. Morsberger, H. Lee, D. Petersen, J. Quackenbush, A. Scott, M. Wilson, Y. Yang, S.Q. Ye, W. Yu. *Nat. Meth.*, **2**, 345 (2005).
[17] J.E. Larkin, B.C. Frank, H. Gavras, R. Sultana, J. Quackenbush. *Nat. Meth.*, **2**, 337 (2005).
[18] C. Teh, S. Parinov, V. Korzh. *Biotechniques*, **38**, 897 (2005).
[19] S. Mathavan, S.G. Lee, A. Mak, L.D. Miller, K.R. Murthy, K.R. Govindarajan, Y. Tong, Y.L. Wu, S.H. Lam, H. Yang, Y. Ruan, V. Korzh, Z. Gong, E.T. Liu, T. Lufkin. *PLOS Genet.*, **1**, e29 (2005).
[20] C. Ton, D. Stamatiou, C.C. Liew. *Physiol. Genom.*, **13**, 97 (2003).
[21] R.L. Malek, H. Sajadi, J. Abraham, M.A. Grundy, G.S. Gerhard. *Comp. Biochem. Physiol.*, **138C**, 363 (2004).
[22] K. van der Ven, M. De Wit, D. Keil, L. Moens, K. Van Leemput, B. Naudts, W. De Coen. *Comp. Biochem. Physiol.*, **141B**, 408 (2005).
[23] C.M. Kendziorski, Y. Zhang, H. Lan, A.D. Attie. *Biostatistics*, **4**, 465 (2003).
[24] B.M. Bolstad, R.A. Irizarry, M. Åstrand, T.P. Speed. *Bioinformatics*, **19**, 185 (2003).
[25] J. Huggett, K. Dheda, S. Bustin, A. Zumla. *Genes Immun.*, **6**, 279 (2005).
[26] P.Y. Muller, H. Janovjak, A.R. Miserez, Z. Dobbie. *BioTechniques*, **32**, 1372 (2002).
[27] J.P. Novak, R. Sladek, T.J. Hudson. *Genomics*, **79**, 104 (2002).
[28] M. Bakay, Y.W. Chen, R. Borup, P. Zhao, K. Nagaraju, E.P. Hoffman. *BMC Bioinform.*, **3**, 4 (2002).
[29] T.S. Tanaka, S.A. Jaradat, M.K. Lim, G.J. Kargul, X. Wang, M.J. Grahovac, S. Pantano, Y. Sano, Y. Piao, R. Nagaraja, H. Doi, W.H. Wood, K.G. Becker, M.S.H. Ko. *Proc. Natl. Acad. Sci.*, **97**, 9127 (2000).
[30] V.K. Mootha, C.M. Lindgren, K.F. Eriksson, A. Subramanian, S. Sihag, J. Lehar, P. Puigserver, E. Carlsson, M. Ridderstrale, E. Laurila, N. Houstis, M.J. Daly, N. Patterson, J.P. Mesirov, T.R. Golub, P. Tamayo, B. Spiegelman, E.S. Lander, J.N. Hirchhorn, D. Altshuler, L.C. Groop. *Nat. Genet.*, **34**, 267 (2003).
[31] K.L. Willett, S.J. McDonald, M.A. Steinberg, K.B. Beatty, M.C. Kennicutt, S.H. Safe. *Environ. Toxicol. Chem.*, **16**, 1472 (1997).
[32] H. Wang, T. Yan, J.T.T. Tan, Z. Gong. *Gene*, **256**, 303 (2000).
[33] S.C. Tilton, L.G. Gerwick, J.D. Hendricks, C.S. Rosato, G. Corley-Smith, S.A. Givan, G.S. Bailey, C.J. Bayne, D.E. Williams. *Toxicol. Sci.*, Doi: 10.1093/toxsci/kfi309 (2005).