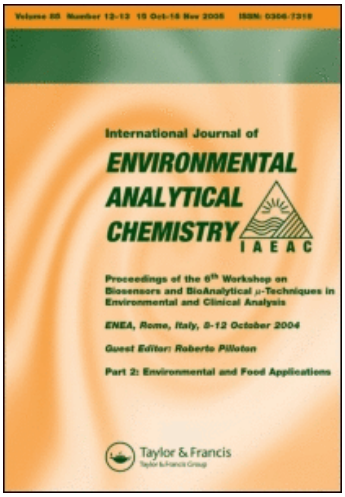

International Journal of Environmental Analytical Chemistry
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713640455

Gene expression patterns - a tool for bioanalysis

Martin Alberti[a]; Ulf Kausch[a]; Stefanie Haindl[a]; Robert Leibiger[a]; Jan Budczies[b]; Martin Seifert[a]; Bertold Hock[a]
[a] Technische Universitaet Muenchen, Center of Life and Food Sciences Weihenstephan, Chair of Cell Biology, Alte Akademie 12, 85354 Freising, Germany [b] Institute for Bioinformatics, GSF-National Research Center for Environment and Health, Ingoldstaedter Landstrasse 1, 85758 Neuherberg, Germany

To cite this Article Alberti, Martin , Kausch, Ulf , Haindl, Stefanie , Leibiger, Robert , Budczies, Jan , Seifert, Martin and Hock, Bertold(2005) 'Gene expression patterns - a tool for bioanalysis', International Journal of Environmental Analytical Chemistry, 85: 9, 589 — 608
To link to this Article: DOI: 10.1080/03067310500145930
URL: http://dx.doi.org/10.1080/03067310500145930

Intern. J. Environ. Anal. Chem.
Vol. 85, Nos. 9–11, 10 August–15 September 2005, 589–608

Taylor & Francis
Taylor & Francis Group

Gene expression patterns – a tool for bioanalysis

MARTIN ALBERTI†*, ULF KAUSCH†, STEFANIE HAINDL†,
ROBERT LEIBIGER†, JAN BUDCZIES‡,
MARTIN SEIFERT† and BERTOLD HOCK†

†Technische Universitaet Muenchen, Center of Life and Food Sciences Weihenstephan,
Chair of Cell Biology, Alte Akademie 12, 85354 Freising, Germany
‡Institute for Bioinformatics, GSF–National Research Center for Environment and Health,
Ingoldstaedter Landstrasse 1, 85758 Neuherberg, Germany

(*Received 7 December 2004; in final form 1 February 2005*)

Gene expression patterns are efficient tools to elucidate responses to hormones as well as endocrine disruptors at the molecular level. Male zebrafish (*Danio rerio*) were exposed to 17β-estradiol (E2). Transcriptome analysis was first carried out by quantitative polymerase chain reaction (PCR) of the gene coding for vitellogenin (*vtg1*) in relation to the housekeeping gene *ef1*α. A significant increase in the expression of *vtg1* was found at concentrations $>100\,\mathrm{ng\,L^{-1}}$ E2. Then DNA microarray experiments were performed to examine the expression levels of the zebrafish genome. A total of 210 sequences were regulated in male zebrafish exposed to $500\,\mathrm{ng\,L^{-1}}$ E2. They include genes involved in the estrogen pathway, such as the gene coding for the estrogen receptor, and in embryonic development, such as homeo box genes. A complete list of up- and down-regulated genes and sequences is provided in this article.

Keywords: Gene expression; Estrogen; Zebrafish

1. Introduction

Estrogenic effects are not only caused by endogenous hormones, but also mimicked by synthetic estrogens, phyto- and mycoestrogens as well as a great number of endocrine disruptors (EDCs).

Two complementary approaches are available for the detection of estrogenic substances. Chemical analysis is used for identifying analytes, already known to have estrogenic effects, and in determining their concentrations. In contrast, bioanalysis is able to examine the presence of even unknown estrogenic substances and to determine effect concentrations. But the identification is not within the scope of this approach. Bioeffects can be observed at different levels of complexity and therefore exploited for bioanalysis: behavioral and developmental responses, and effects at the sub-organismic and molecular levels. Expression of suitable biomarkers provides a convenient tool to detect exposure to EDCs.

*Corresponding author. Fax: +49-8161-4403. Email: alberti@wzw.tum.de

International Journal of Environmental Analytical Chemistry
ISSN 0306-7319 print: ISSN 1029-0397 online © 2005 Taylor & Francis
DOI: 10.1080/03067310500145930

Vitellogenin (Vtg), an egg yolk precurser, can be used as a biomarker for examining exposure to estrogenic substances [1] in fish. Its synthesis is normally controlled by estrogens such as 17β-estradiol (E2) [2].

In this study the zebrafish (*Danio rerio*) was applied as a test organism. This popular model organism is used throughout the world in numerous laboratories. Research directed at pharmacological [3] and genetic studies [4] are a few examples used for important applications. Zebrafish do not grow larger than 4 cm in length and can therefore be kept in relatively small tanks. Additionally, this species produces a large number of offspring under good breeding conditions. Its genomic sequence is known and available on the internet [5].

Earlier approaches to biomonitoring included the observation of whole populations or species and the detection of effects of pollutants at the molecular level, concentrating on single responses. These concepts do not take into account the complexity of pathways or the diversity of responses. Therefore, it would be desirable to examine as many responses as possible to environmental pollutants at the same time.

DNA microarray techniques offer the possibility of measuring multiple effects of pollutants at the level of gene expression, giving better insights into effects and their interactions. A great number of possible target genes can be tested on a small glass slide. In order to find out relevant concentrations of pollutants, quantitative PCR (qPCR) is used. The combination of qPCR and the microarray technique allows us to detect exposure to estrogens and to gain better knowledge about their mode of action. By this approach, the current physiological state of tissue is elucidated.

In this article a survey of genes and sequences is presented that are regulated by E2 in zebrafish. In addition to genes directly involved in the E2 pathway leading to Vtg, a great number of genes related to cell proliferation and embryonic development were identified.

2. Experimental

2.1 *Fish breeding*

Zebrafish were kept in 5 L tanks to obtain a large number of eggs. Four female and two male fish were kept at a time for 48 h at 24°C. The bottom was covered with glass beads to prevent the fish from feeding on their eggs. Hatching took place after 48 h and juvenile fish were fed with Micron (Sera, Heinsberg, Germany).

2.2 *Exposure of fish*

Figure 1 shows a scheme of a typical exposure unit providing stable concentrations and conditions.

Zebrafish were exposed to E2 in 20 L tanks, being part of a flow-through system. Tap water was running through an activated charcoal filter into a 500 L tank where it was heated to 24°C prior to use.

Exposure experiments were carried out in 20 L tanks, connected to the heating tank. Water flow and dosing with E2 were controlled by peristaltic precision pumps. A heater kept the water temperature at 24°C in each tank and air pumps circulated the water. The flow through rate was set to 21 mL min^{-1}. In this way, the tank volume was

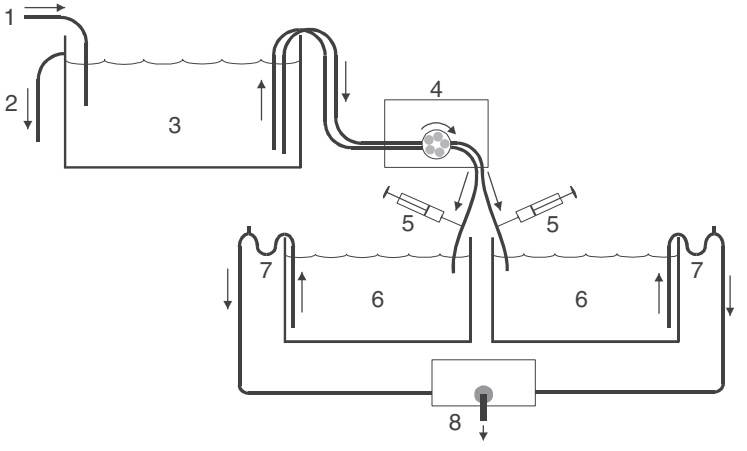

Figure 1. Scheme of an exposure unit. (1) Fresh water; (2) overflow; (3) heating tank (500 L); (4) peristaltic pump; (5) dosing pump; (6) exposure tanks (20 L); (7) overflow; (8) activated charcoal filter.

changed 1.5 times per 24 h. Further peristaltic pumps were used to adjust the concentration of E2. Concentrations in the tanks were measured by ELRA [6]. Stock solutions were kept in a refrigerator at 4°C.

Up to nine adult fish were kept in each tank. Eight tanks were used for exposures with the following concentrations: 0, 1, 10, 100, 200, 300, 400 and 500 ng L^{-1} E2. Four adult female fish were kept in a further tank without E2. They were used as a calibrator to render different qPCR experiments comparable. The daylight period was set to 12 h, controlled by a timer. The duration of an exposure experiment was eleven days. Two replicates were carried out in each case.

2.3 *Quantitative PCR*

Liver tissue of exposed and control fish were dissected. RNA was extracted according to the TRI REAGENT™ RNA Isolation System protocol (Sigma #T 9424), based on phenol/chloroform extraction. RNA was dissolved in DEPC-H_2O at 65°C and 0.5 µL RNase Inhibitor (Protector, Roche #3335399) was added. The RNA concentration was measured photometrically. Two micrograms were reversely transcribed to cDNA according to the 1st Strand cDNA Synthesis Kit for RT-PCR (AMV) protocol (Roche #1483188). The resulting cDNA was diluted 20-fold in DEPC-H_2O. A master mix was prepared according to the protocol LightCycler® FastStart DNA Master^{PLUS} SYBR Green I (Roche #03515885). The genes coding for vitellogenin 1 (*vtg1*) and elongation factor 1 α (*ef1α*) as a housekeeping gene were examined in the LightCycler experiments.

Table 1 shows the applied primers. They were designed according to the data of NCBI [5] and obtained from MWG Biotech (Ebersberg, Germany).

Sixteen microlitres of the master mix and 4 µL of the diluted cDNA were pipetted in a glass capillary. Each sample of a defined E2 concentration was measured three times.

The amplificates of several PCRs with different annealing temperatures were electrophoresed in a 1% agarose gel to determine the optimal temperature for LightCycler experiments.

Table 1. Primers used for the qPCR experiments.

Type of primer	Sequence
vtg1 forward	5′-GCC AAA AAG CTG GGT AAA CA-3′
vtg1 reverse	5′-AGT TCC GTC TGG ATT GAT GG-3′
ef1α forward	5′-CAG CTG ATC GTT GGA GTC AA-3′
ef1α reverse	5′-TCT TCC ATC CCT TGA ACC AG-3′

Table 2. LightCycler program.

Step	Temperature	Duration
Preincubation	95°C	10 min
40 Cycles		
Denaturation	95°C	10 s
Annealing	56°C	5 s
Elongation	72°C	9 s

Table 2 shows the program used for all LightCycler runs.

Crossing points (CPs) of the fluorescence curves were calculated by LightCycler Software 3.5 (Roche) with the second derivation maximum method. CP describes the cycle number at which the fluorescence signal of a sample exceeds the background signal. Increasing template cDNA in a sample leads to decreasing CP.

In addition melting curves were established for the amplificates. This was done by denaturing PCR products at 95°C and measuring the decrease of fluorescence, using LightCycler Software 3.5 (Roche).

For quantification of gene expression levels a calibrator-normalized quantification with efficiency correction was carried out. The pooled cDNA of four non-exposed female zebrafish served as calibrator and was used for normalization of different LightCycler runs. For relative quantification of differential expression between different samples the CPs of the target gene *vtg1* were related to a reference gene, the housekeeping gene *ef1α*. An efficiency test with different dilutions of calibrator cDNA (1 : 5, 1 : 10, 1 : 100, 1 : 1000 and 1 : 10 000) was performed to examine the quality of PCR reaction and primers. Data analysis of CPs was carried out with REST Software according to the method of Pfaffl [7] and the delta–delta method. The delta–delta method assumes an identical efficiency of amplification of target gene (*vtg1*) and reference gene (*ef1α*):

$$\Delta\text{CP value} = \text{CP value(target)} - \text{CP value(reference)} = \text{constant}$$

The ΔΔCP value considers the influence of differential expression caused by exposure to E2:

$$\Delta\Delta\text{CP value} = \Delta\text{CP value(exposed)} - \Delta\text{CP value(control)}$$

The comparison of differential expression between exposed and control group is done by calculating the ratio:

$$\text{ratio} = 2^{-\Delta\Delta\text{CP}}$$

The number 2 in the formula refers to the efficiency of amplification.

The method of Pfaffl [7] provides information about the level of expression:

$$R = \frac{(E_{\text{target}})^{\Delta \text{CPtarget (MEANcontrol}-\text{MEANsample)}}}{(E_{\text{reference}})^{\Delta \text{CPreference (MEANcontrol}-\text{MEANsample)}}}$$

whereas R = ratio, E = efficiency, reference = reference gene (*ef1α*), and target = target gene (*vtg1*). The calculation using this formula is done by REST Software.

Significance analysis between CPs of different exposure conditions is performed by Tukey HSD (honestly significant difference) test. This *post hoc* test (multiple comparison test) can be used to determine the significant differences between group mean values in an analysis of variance setting. The comparison of the calculated q value with the differences of mean values between exposure conditions allows conclusions about significant differentiation in gene expression.

2.4 *Microarrays*

For microarray experiments, RNA was extracted as described above. RNA concentrations were measured photometrically. Samples of 45 μg RNA were reversely transcribed to cDNA according to the CyScribe First-Strand cDNA Labeling System protocol (Amersham Biosciences #RPN6202X). cDNA was labeled alternatively with fluorescent cy3 or, respectively, cy5, depending on the protocol. Hybridization between labeled samples and DNA target sequences on 2 × 7 k or 14 k zebrafish arrays (MWG Biotech AG #2260-000000) was carried out according to the array application guide of MWG, Ebersberg, Germany.

The dataset included 20 hybridizations and a total number of 14,067 probes. Processing of the expression data was performed using the statistical language *R* [8], shown to be particular suitable for the analysis of microarray data [9]. The expression values were calculated as difference between mean signal and mean background values. Values were passed to the log-2 scale. In doing so NAs (not applicable) were generated in cases where the background value exceeded the signal value. Statistics about the number of NAs per sequence was executed. All sequences with more than four NAs were removed, leaving a number of 14,045 probes. After this filtering the missing values were imputed using the *transcan* algorithm [10].

On-chip normalization was performed by a non-linear method based on the scatterplot smother *loess* [11]. This was followed by a linear normalization between chips, i.e. a scaling of all channels to equal mean and variance.

For each of the probes two measures of differential expression were calculated, the p-value of Student's 1-sample t-test and the mean fold change between exposed and control samples. A combination of p-value and fold change thresholding was used to select sequences that responded to E2 exposure. A permutation procedure was applied to validate the results of different selection procedures. This was done by calculating the number of selected sequences for all $0.5 \times \binom{10}{5} = 126$ informative permutations of the samples. The specificity of the procedure was estimated by comparison of the mean number of selected sequences n_{exp} with the observed number of responding sequences n_{obs}:

$$\text{Specificity} = \frac{n_{\text{obs}} - n_{\text{exp}}}{n_{\text{obs}}}$$

In addition the statistical significance of selecting n_{obs} or more sequences was estimated from the null distribution generated by the permutations. A gene and sequence table corresponding to p-value threshold 0.01 and fold change threshold 1.50 was provided. The selected sequences were visualized in a MA plot, where M is the fold change between exposed and control fish, $M = \log_2(x_{exposed}/x_{control})$, and A is the average expression value in exposed and control fish, $A = 0.5 \log_2(x_{exposed} \times x_{control})$.

3. Results

3.1 *Zebrafish exposure*

Adult zebrafish were kept in 20 L glass tanks for exposure experiments with E2. The spiked concentrations in the exposure tanks were controlled by ELRA, a non-radioactive receptor assay based on the human estrogen receptor α. The detection range of E2 for this assay was from 70 to 500 ng L^{-1}. The tanks containing 100, 200, 300, 400 and 500 ng L^{-1} were tested. Figure 2 shows the results.

The data show a good correspondence between the measured and expected concentrations.

3.2 *Quantitative transcriptome analysis of E2 responsive genes*

In order to define relevant E2 concentrations affecting gene expression in male zebrafish, quantitative PCR (qPCR) was carried out by LightCycler. The *vtg1* expression was compared with the *ef1α* as an internal control, which is not affected by E2 (data not shown). Primers were constructed that specifically amplify segments of the regions coding for *vtg1* and *ef1α* of the cDNA (table 1). Stringent conditions of PCR were obtained by gradient PCR. The amplificates for *ef1α* and *vtg1* were visualized by gel

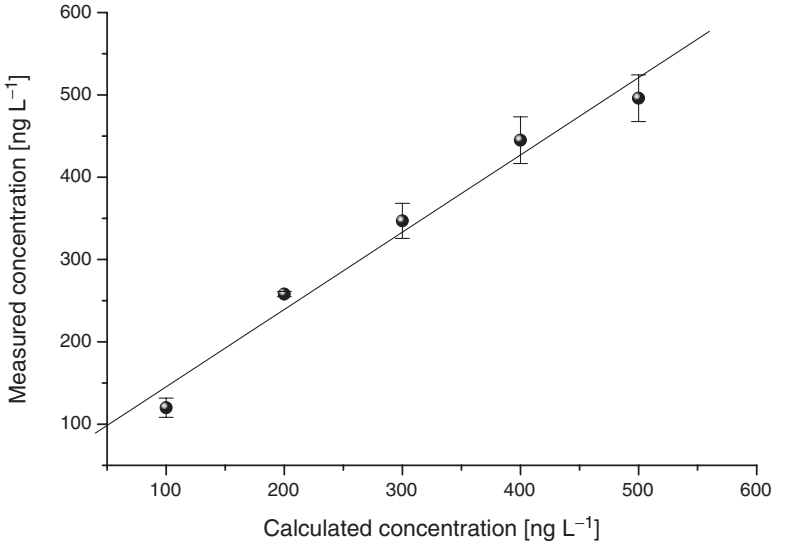

Figure 2. Exposure tank E2 concentrations measured by ELRA.

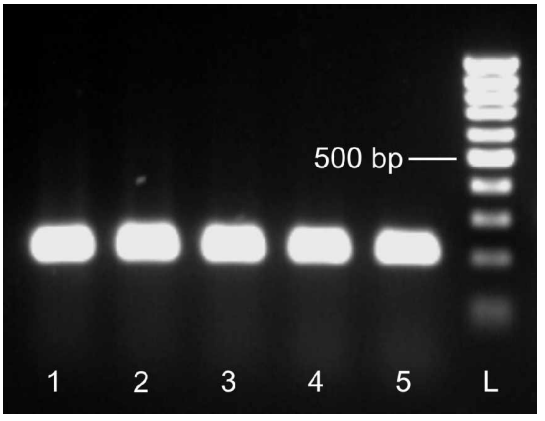

Figure 3. Agarose gel of gradient PCR products. Amplificates of *vtg1* (210 bp) at different annealing temperatures. (1) 54°C, (2) 56°C, (3) 58°C, (4) 60°C, (5) 62°C, (L) ladder.

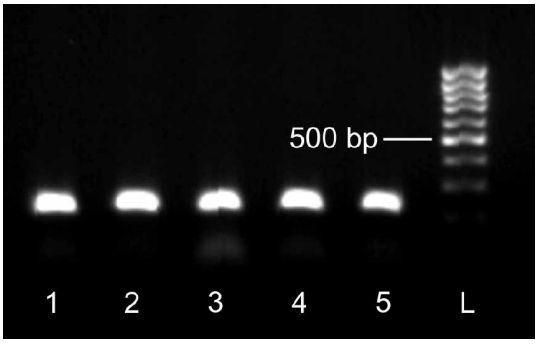

Figure 4. Agarose gel of gradient PCR products. Amplificates of *ef1α* (208 bp) at different annealing temperatures. (1) 54°C, (2) 56°C, (3) 58°C, (4) 60°C, (5) 62°C, (L) ladder.

electrophoresis yielding the expected sizes of 208 bp for *ef1α* and 210 bp for *vtg1*, as shown in figures 3 and 4.

These data provide evidence of the high quality of amplificates.

To determine suitable template quantities for qPCR, different calibrator dilutions were applied in the LightCycler experiment. A typical LightCycler run is shown in figure 5 for several dilutions of calibrator cDNA samples. Table 3 reveals the according CPs. The increasing fluorescence follows a sigmoid function and results from the binding of the fluorescent dye. A dilution of 1 : 20 was selected for measuring samples.

The crossing and melting points of the products were calculated by the LightCycler Software 3.5 (Roche). Melting points of the amplificates are 84.0°C for *vtg1* and 87.4°C for *ef1α*.

Figure 6 shows the melting curves of the amplificates. The different dilutions coincide in each case at a single peak, showing that the amplificates are of equal size.

In addition an efficiency test with calibrator cDNA in different concentrations was performed. The efficiencies of the qPCR amplifying segments of *vtg1* and *ef1α* are the same. Data analysis of crossing points was carried out according to the delta–delta method and the method of Pfaffl [7] with REST Software.

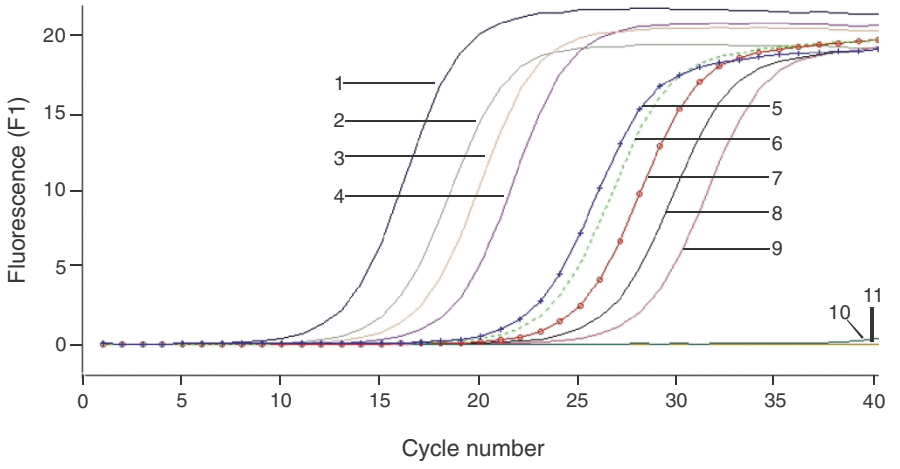

Figure 5. LightCycler experiment. Curves: (1) *vtg1*, 1 : 5 dilution; (2) *vtg1*, 1 : 20 dilution; (3) *vtg1*, 1 : 50 dilution; (4) *vtg1*, 1 : 100 dilution; (5) *ef1α*, 1 : 5 dilution; (6) *ef1α*, 1 : 10 dilution; (7) *ef1α*, 1 : 20 dilution; (8) *ef1α*, 1 : 50 dilution; (9) *ef1α*, 1 : 100 dilution; (10) *ef1α*, PCR master mix without template; (11) *vtg1* PCR master mix without template.

Table 3. Crossing points of LightCycler experiment with different dilutions of *vtg1* and *ef1α* samples.

Sample	Crossing points
ef1α 1 : 5	22.47
ef1α 1 : 10	23.20
ef1α 1 : 20	24.68
ef1α 1 : 50	26.29
ef1α 1 : 100	28.00
vtg1 1 : 5	12.94
vtg1 1 : 20	15.16
vtg1 1 : 50	16.71
vtg1 1 : 100	18.42
Master mix *ef1α* without template	> 36
Master mix *vtg1* without template	> 36

Significance analysis between CPs of different exposure conditions was performed by the Tukey HSD test. In doing so a q value of 4.28 was calculated. The comparison of the q value with the differences of mean values between exposure conditions resulted in a differentiation in gene expression for zebrafish exposed to 200, 300, 400 and 500 ng L^{-1} E2.

Figure 7 shows the *vtg1* expression related to the calibrator. Male fish exposed to E2 for 11 days yielded a significant vitellogenin induction in liver tissue starting at a concentration of 200 ng L^{-1} E2.

3.3 *Microarray approach*

Ten male fish were exposed to 500 ng L^{-1} E2. Liver tissue of the exposed fish and ten control fish were analyzed by microarray experiments. Gene expression profiles were obtained on 2×7 k as well as on 14 k zebrafish arrays of MWG Biotech AG.

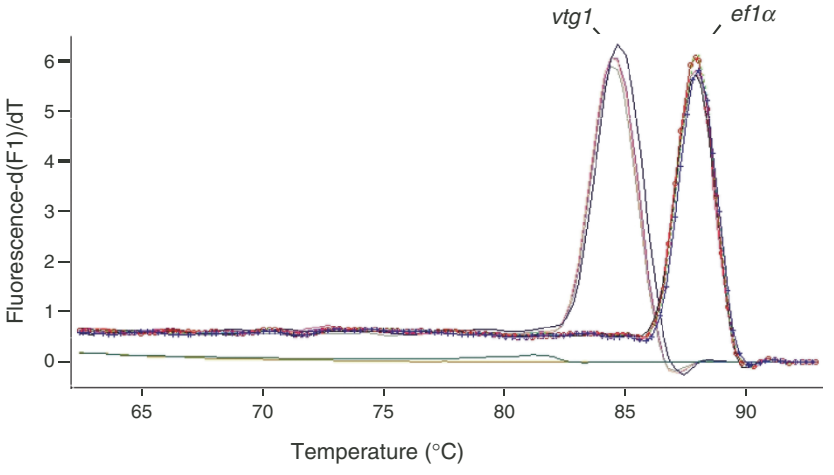

Figure 6. Melting points of *vtg1* (84.0°C) and *ef1α* (87.4°C) amplificates at different concentrations, (cf figure 5).

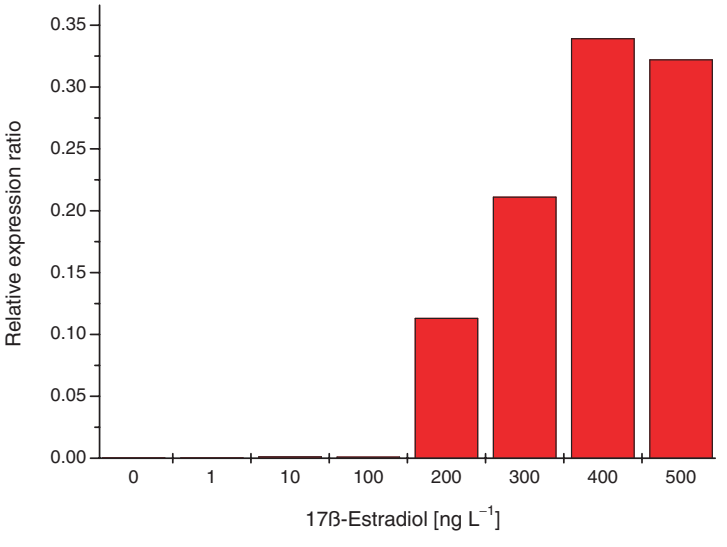

Figure 7. Relative expression of *vtg1* at different E2 concentrations.

The zebrafish oligonucleotide array contained 14 067 gene spots. Figure 8 shows an overlay of cy3 and cy5 fluorescently labeled samples on an array after hybridization. Ten hybridization experiments were carried out. For five hybridizations sample cDNA of control fish was labeled with cy3-dCTP (green) and sample cDNA of exposed fish with cy5-dCTP (red). Intensive red signals indicate sequences that were significantly up-regulated. Green signals refer to down-regulated sequences. Yellow signals represent no difference in expression between control and exposed fish. For five additional hybridizations dyes for labeling were flipped.

The intensities of cy3 and cy5 labeled oligonucleotides were measured and background signals were taken into account. Expression values were calculated as difference between mean signal and mean background values. On-chip and inter-chip

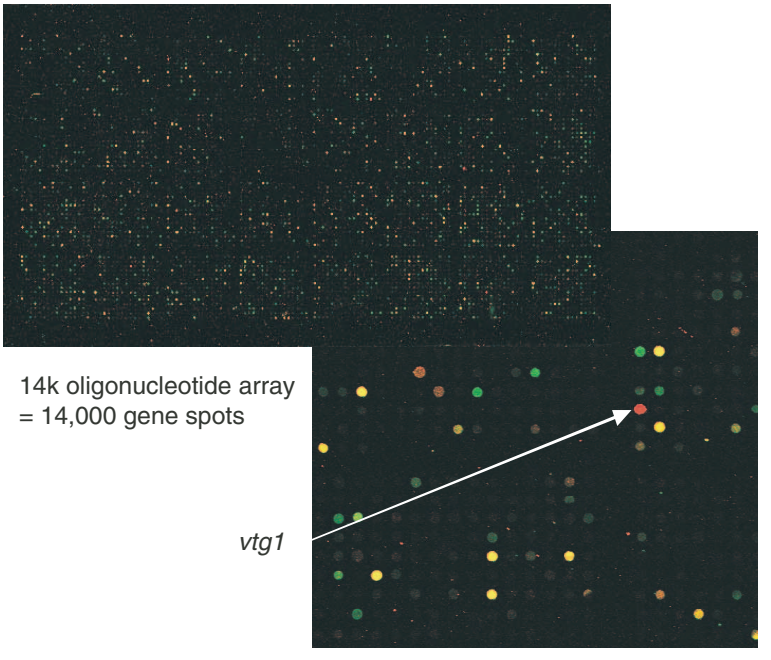

Figure 8. Overlay of 14k zebrafish oligonucleotide array after hybridization.

Table 4. Comparison of different procedures for the detection of E2-responding sequences. The number of expected differentially expressed sequences n_{exp} is compared to the number n_{obs} of observed responding sequences.

Set maximum *p*-value	Set minimum fold change	n_{exp}	n_{obs}	Estimated specifity	Significance of n_{obs}
1.0000	1.00	14045.0	14045	0.00	1.000
1.0000	1.25	2805.4	3516	0.20	0.183
1.0000	1.50	631.2	1148	0.45	0.079
1.0000	2.00	113.1	350	0.68	0.016
0.0100	1.00	52.4	274	0.81	0.008
0.0100	1.25	48.8	269	0.82	0.008
0.0100	1.50	23.1	210	0.89	0.008
0.0100	2.00	6.1	144	0.96	0.008
0.0010	1.00	4.3	86	0.95	0.008
0.0010	1.25	4.1	86	0.95	0.008
0.0010	1.50	2.5	85	0.97	0.008
0.0010	2.00	1.0	72	0.99	0.008
0.0001	1.00	0.7	42	0.98	0.008
0.0001	1.25	0.7	42	0.98	0.008
0.0001	1.50	0.5	42	0.99	0.008
0.0001	2.00	0.4	40	0.99	0.008

normalization were performed as described in the material and methods section. Statistical analyses were carried out to select sequences that respond to E2 exposure. Responding sequences were selected by a combined *p*-value and fold change thresholding. A permutation procedure was applied to validate the results of different selection procedures. Different sequence selection procedures are compared in table 4.

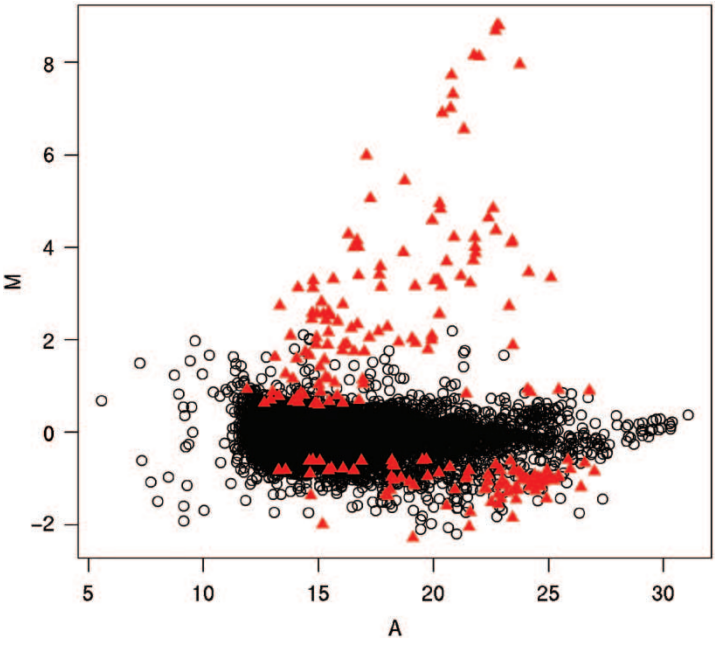

Figure 9. Fold change (M) between exposed and control fish as a function of average expression (A) in exposed and control fish (log-2 scale). A total of 210 E2-responding probes (triangles) were selected by a combined p-value and fold change threshold.

Selecting probes by p-values <0.01 and fold changes >1.50 (cf table 4, row 7) yielded a number of 210 E2-responding probes at a specificity of 89%. This number turned out to be significant ($p=0.008$) with respect to the null distribution generated by the permutations. A pronounced asymmetry between up-regulation (121 probes) and down-regulation (89 probes) is obvious from a MA plot (figure 9).

The 210 selected probes were collected in a list and annotated due to information from the chip producer (table 5).

The *vtg1* shows high up-regulation up to 129-fold. In addition, further zebrafish genes exhibit a strong up-regulation. Among them are the ones coding for the homeo box a3a (31-fold), vitellogenin 3 (about 16-fold), nothepsin (about 7-fold) and estrogen receptor (about 7-fold).

Other genes were found to be down-regulated by E2, like the ones coding for pleiotrophin 1 (2.7-fold), ephrin b2 (1.9-fold) or POU domain gene 50 (1.8-fold).

4. Discussion

Presently about five million chemicals are known. Eighty thousand substances are in use and 500–1000 additional ones appear every year [12]. The identification of potential estrogenic compounds is an important issue in environmental analysis. Several studies on different species in aquatic systems such as African catfish (*Clarias gariepinus*) [13], broad-nosed caiman (*Caiman latirostris*) [14] and zebrafish (*Danio rerio*) [15] were carried out recently. Aquatic organisms exposed to estrogenic substances showed

Table 5. List of zebrafish genes and sequences regulated after exposure to 500 ng L^{-1} E2. *p*, *p*-value, fc, fold change, DRFF, DRFRAMEFINDER by MWG Biotech AG, Ebersberg, Germany.

Zebrafish gene or sequence	Accession reference	*p*	fc
Up-regulated genes			
vitellogenin 1	AF406784_1	3.18E−07	128.9
homeo box a3a; hoxa3a	NM_131534_1	1.51E−06	31.0
vitellogenin 3 precursor; vg3	AF254639_1	4.43E−06	15.9
vitellogenin 3 precursor; vg3	AF254638_1	3.43E−04	9.8
nothepsin; nots	NM_131804_1	2.97E−05	6.8
estrogen receptor; er	AB037185_1	3.59E−04	6.7
rev-erb beta 2	AF342942_1	6.70E−05	6.1
tarama; tarama	NM_130990_1	3.12E−03	5.9
putative mature peptide; dvr1	NM_130948_1	3.30E−04	5.7
mesoderm posterior b; mespb	NM_131552_1	8.05E−04	3.4
kallmann syndrome 1b sequence; kal1b	NM_131379_1	8.63E−04	1.9
Down-regulated genes			
pleiotrophin 1; plei1	NM_131070_1	9.63E−03	−2.7
carboxypeptidase a	AF376130_1	4.87E−03	−2.4
zg2	U18312_1	1.57E−04	−2.1
ephrin b2a; efnb2a	NM_131023_1	4.34E−04	−1.9
pou domain gene 50; pou50	NM_131161_1	9.68E−03	−1.8
caspase 8; casp8	NM_131510_1	5.16E−03	−1.8
leukocyte cell derived chemotaxin 1; lect1	NM_131835_1	4.64E−03	−1.7
ppara	U93475_1	1.02E−03	−1.5
Up-regulated sequences similar to GenPept			
|x92804|x92804_1 vitellogenin; vtg1 – *Oncorhynchus mykiss*	DRFF12001	1.51E−08	445.3
|af406784|af406784_1 vitellogenin 1 – *Danio rerio*	DRFF05666	9.23E−09	442.4
|af406784|af406784_1 vitellogenin 1 – *Danio rerio*	DRFF10974	4.21E−09	410.5
|af406784|af406784_1 vitellogenin 1 – *Danio rerio*	DRFF10141	2.25E−08	283.5
|af406784|af406784_1 vitellogenin 1 – *Danio rerio*	DRFF08615	1.53E−08	279.2
|af406784|af406784_1 vitellogenin 1 – *Danio rerio*	DRFF07803	1.45E−08	248.5
|ay049952|ay049952_1 vitellogenin; vtg – *Salmo salar*	DRFF09685	4.61E−07	211.3
|x92804|x92804_1 vitellogenin; vtg1 – *Oncorhynchus mykiss*	DRFF01482	1.11E−06	159.2
|af406784|af406784_1 vitellogenin 1 – *Danio rerio*	DRFF06844	1.87E−08	119.9
|z49204|z49204_1 nadp transhydrogenase – *Mus musculus*	DRFF09851	7.72E−07	94.1
|af068286|af068286_1 hdcmd38p – *Homo sapiens*	DRFF10861	8.95E−05	63.6
|bc016428|bc016428_1 phenylalanine-trna synthetase-like – *Mus musculus*	DRFF04483	3.58E−06	43.5
|d86970|d86970_1 kiaa0216 – *Homo sapiens*	DRFF00141	1.30E−04	33.4
|ak017839|ak017839_1 sap domain containing protein data source:pfam, source key:pf02037, evidence:iss putative – *Mus musculus*	DRFF06664	9.68E−06	28.8
|u97682|u97682_1 regeneration induced 2′,3′-cyclic-nucleotide 3′-phosphodiesterase isoform 2 grich70 – *Carassius auratus*	DRFF07101	2.76E−04	28.4
|u41744|u41744_1 pdgf associated protein – *Rattus norvegicus*	DRFF06048	1.89E−05	24.9
|ab032971|ab032971_1 kiaa1145 protein; kiaa1145 – *Homo sapiens*	DRFF12264	2.96E−06	24.1
|ab071402|ab071402_1 chondroitin synthase; chsy – *Homo sapiens*	DRFF08404	9.86E−05	20.7
|af406784|af406784_1 vitellogenin 1 – *Danio rerio*	DRFF04616	9.39E−05	19.5
similar to pir-nref|nf00522037 similar to nuclear receptor subfamily 1, group d, member 1 – *Mus musculus*	DRFF07335	4.42E−04	18.6
|m12105|m12105_1 aspartate aminotransferase precursor – *Gallus gallus*	DRFF03594	8.98E−05	18.5
|m77013|m77013_1 beta-catenin; beta-catenin – *Xenopus laevis*	DRFF03552	3.60E−06	17.8
|u80741|u80741_1 cagh44; cagh44 – *Homo sapiens*	DRFF04902	4.88E−05	17.7

(*Continued*)

Table 5. Continued.

Zebrafish gene or sequence	Accession reference	*p*	fc
|u67156|u67156_1 mitogen-activated kinase kinase kinase 5; mapkkk5 – *Homo sapiens*	DRFF10702	4.35E−05	17.1
|al030996|al030996_1 dj1189b24.4 novel putative protein similar to hypothetical proteins s. pombe c22f3.14c and C. elegans c16a3.8; dj1189b24.4 – *Homo sapiens*	DRFF01505	5.15E−04	16.1
|u12979|u12979_1 pc4 – *Homo sapiens*	DRFF01918	6.82E−04	16.1
|al136564|al136564_1 hypothetical protein; dkfzp761i141 – *Homo sapiens*	DRFF06094	4.79E−06	14.9
|af026527|af026527_1 zipcode-binding protein; zbp1 – *Gallus gallus*	DRFF05819	4.31E−04	14.6
|af250920|af250920_1 12cc4 – *Homo sapiens*	DRFF10375	1.75E−04	13.1
|m31899|m31899_1 ercc3 – *Homo sapiens*	DRFF09928	5.59E−04	13.0
|af406784|af406784_1 vitellogenin 1 – *Danio rerio*	DRFF01826	4.24E−04	12.0
|u54559|u54559_1 translation initiation factor eif3 p40 subunit – *Homo sapiens*	DRFF07720	3.31E−04	11.0
|l16507|l16507_1 formiminotransferase-cyclodeaminase – *Sus scrofa*	DRFF09121	8.78E−05	10.7
|x65371|x65371_1 polypirimidine tract binding protein; ptb-2 – *Homo sapiens*	DRFF11081	4.06E−03	10.5
|af084461|af084461_1 type i cytokeratin; cyt1 – *Danio rerio*	DRFF01319	1.59E−05	10.2
|u53922|u53922_1 dnaj-like protein; rdj1 – *Rattus norvegicus*	DRFF05454	9.08E−05	9.9
|af130354|af130354_1 vitellogenin precursor; vtg – *Pimephales promelas*	DRFF11751	7.67E−05	9.8
|ab067489|ab067489_1 kiaa1902 protein; kiaa1902 – *Homo sapiens*	DRFF05659	5.22E−03	8.9
|af390018|af390018_1 putative protein kinase wnk4; prkwnk4 – *Homo sapiens*	DRFF01845	4.60E−03	8.9
|af406784|af406784_1 vitellogenin 1 – *Danio rerio*	DRFF11696	5.08E−04	8.8
|ab050460|ab050460_1 cyclin-dependent kinase regulatory subunit cyclin b2 – *Oryzias curvinotus*	DRFF12983	2.19E−05	8.7
|af406784|af406784_1 vitellogenin 1 – *Danio rerio*	DRFF01323	9.67E−04	8.7
|af406784|af406784_1 vitellogenin 1 – *Danio rerio*	DRFF01465	9.64E−05	7.1
|ak026904|ak026904_1 unnamed protein product – *Homo sapiens*	DRFF07966	1.69E−04	6.7
|ak011356|ak011356_1 data source:sptr, source key:q9p0n9, evidence:iss homolog to hspc239 putative – *Mus musculus*	DRFF07548	3.16E−03	6.0
|d17296|d17296_1 polyubiquitin; ubc – *Rattus norvegicus*	DRFF11489	1.44E−04	5.9
|af414432|af414432_1 vitellogenin – *Cyprinus carpio*	DRFF06954	2.51E−03	5.9
|bc011131|bc011131_1 riken cdna 2400009b11 gene – *Mus musculus*	DRFF03877	4.39E−03	5.5
|af406784|af406784_1 vitellogenin 1 – *Danio rerio*	DRFF08349	2.19E−04	5.4
|af414432|af414432_1 vitellogenin – *Cyprinus carpio*	DRFF06304	3.82E−03	5.3
|af224337|af224337_1 beta-1 integrin – *Ictalurus punctatus*	DRFF09925	1.83E−05	5.1
|af406784|af406784_1 vitellogenin 1 – *Danio rerio*	DRFF11852	7.76E−05	4.9
|y08565|y08565_1 udp-galnac:polypeptide n-acetylgalactosaminyltransferase; galnac-t6 – *Homo sapiens*	DRFF00373	2.28E−03	4.8
|j02942|j02942_1 calcium/calmodulin-dependent protein kinase – *Rattus norvegicus*	DRFF03352	1.16E−03	4.5
|ak023794|ak023794_1 unnamed protein product – *Homo sapiens*	DRFF03665	8.57E−03	4.5
|y10258|y10258_1 ta2 ket alpha; p63 – *Rattus norvegicus*	DRFF04073	8.44E−03	4.3
|z14253|z14253_1 xlcl1 – *Xenopus laevis*	DRFF10891	9.53E−06	4.3
|af406784|af406784_1 vitellogenin 1 – *Danio rerio*	DRFF07942	1.34E−03	4.1
|u37012|u37012_1 cleavage and polyadenylation specificity factor – *Homo sapiens*	DRFF08875	1.65E−05	4.1

(*Continued*)

Table 5. Continued.

Zebrafish gene or sequence	Accession reference	*p*	fc
|bc015922|bc015922_1 unknown protein for mgc:20472 – *Homo sapiens*	DRFF06098	1.51E−03	4.1
|af077613|af077613_1 s100-like calcium binding protein; s100 – *Salvelinus fontinalis*	DRFF01742	7.95E−03	4.0
|x86969|x86969_1 beta 1 subunit of heterotrimeric gtp-binding protein; xgbeta1 – *Xenopus laevis*	DRFF12499	8.89E−05	3.9
|d38441|d38441_1 acylamino acid-releasing enzyme – *Homo sapiens*	DRFF01245	3.69E−03	3.8
|af227986|af227986_1 hsc71 – *Rivulus marmoratus*	DRFF06571	5.17E−03	3.8
|ab067489|ab067489_1 kiaa1902 protein; kiaa1902 – *Homo sapiens*	DRFF01979	1.88E−04	3.7
|al136943|al136943_1 hypothetical protein; dkfzp586g1024 – *Homo sapiens*	DRFF12575	7.28E−07	3.7
|m86381|m86381_1 secreted hyaluronate binding protein – *Oryctolagus cuniculus*	DRFF00762	1.56E−03	3.7
|af314544|af314544_1 nuclear receptor co-repressor/hdac3 complex subunit tblr1; tblr1 – *Homo sapiens*	DRFF07696	8.26E−04	3.7
|ab026129|ab026129_1 alpha-2-macroglobulin-2; a2m-2 – *Cyprinus carpio*	DRFF01214	7.06E−03	3.4
|al162062|al162062_1 hypothetical protein; dkfzp762b245 – *Homo sapiens*	DRFF08468	4.44E−04	3.4
|u17606|u17606_1 chs-rex-b – *Gallus gallus*	DRFF11202	1.07E−03	3.3
|u03272|u03272_1 fibrillin-2 – *Homo sapiens*	DRFF03268	6.42E−04	3.2
|al136794|al136794_1 hypothetical protein; dkfzp434c011 – *Homo sapiens*	DRFF09985	2.68E−03	3.1
|x54240|x54240_1 p97 subunit of 15s mg2+ - atpase – *Xenopus laevis*	DRFF08587	3.18E−04	3.0
|al162049|al162049_1 hypothetical protein; dkfzp762e1712 – *Homo sapiens*	DRFF11613	3.40E−04	3.0
|bc011993|bc011993_1 similar to riken cdna 5730568a12 gene – *Homo sapiens*	DRFF05865	3.76E−03	2.7
|af406784|af406784_1 vitellogenin 1 – *Danio rerio*	DRFF00638	4.92E−03	2.4
|af286160|af286160_1 tandem ph domain containing protein-1 – *Homo sapiens*	DRFF08332	1.11E−03	2.3
|af406784|af406784_1 vitellogenin 1 – *Danio rerio*	DRFF06450	2.51E−03	2.2
|af178534|af178534_1 talin; tln – *Homo sapiens*	DRFF03193	6.36E−03	2.2
|aj002078|aj002078_1 syntaxin 6 – *Homo sapiens*	DRFF05078	3.40E−03	2.2
|af266217|af266217_1 peptidyl-proline isomerase – *Gillichthys mirabilis*	DRFF08368	5.87E−03	2.1
|bc016073|bc016073_1 unknown protein for mgc:27555 – *Mus musculus*	DRFF09486	8.59E−03	2.1
|bc007195|bc007195_1 unknown protein for mgc:10780 – *Homo sapiens*	DRFF03017	1.21E−06	2.1
|x57115|x57115_1 serine/threonine specific protein phosphatase; calcineurin a gene – *Rattus norvegicus*	DRFF00841	7.25E−03	1.9
|af401585|af401585_1 ribosomal protein l30 – *Ictalurus punctatus*	DRFF11533	2.09E−04	1.9
|af401553|af401553_1 ribosomal protein p2 – *Ictalurus punctatus*	DRFF12895	6.31E−03	1.9
|u09564|u09564_1 serine kinase – *Homo sapiens*	DRFF12200	2.93E−03	1.8
|bc003996|bc003996_1 unknown protein for mgc:7535 – *Mus musculus*	DRFF11675	7.15E−03	1.8
|ab029946|ab029946_1 p32 subunit of splicing factor sf2; sf2 – *Gallus gallus*	DRFF12671	8.62E−03	1.8
|ak005032|ak005032_1 data source:sptr, source key:q9nwr0, evidence:iss homolog to cdna flj20668 fis, clone kaia585 putative – *Mus musculus*	DRFF00142	5.71E−03	1.8

(*Continued*)

Table 5. Continued.

Zebrafish gene or sequence	Accession reference	*p*	fc
|af330045|af330045_1 lim domain only 7; lmo7 – *Homo sapiens*	DRFF11366	1.65E−03	1.8
|u14391|u14391_1 myosin-ic – *Homo sapiens*	DRFF06074	7.96E−03	1.8
|x82829|x82829_1 nuclear dna helicase ii; ndh2 – *Bos taurus*	DRFF09687	4.09E−03	1.7
|ab027414|ab027414_1 cmp-sialic acid synthetase – *Oncorhynchus mykiss*	DRFF10173	2.06E−03	1.7
|af317517|af317517_1 limb bud and heart; lbh – *Mus musculus*	DRFF04165	7.50E−03	1.7
|aj306399|aj306399_1 selenoprotein n; sepn1 – *Homo sapiens*	DRFF09619	6.96E−03	1.6
|d64010|d64010_1 seizure-related gene product 6 type 3 precursor – *Mus musculus*	DRFF11050	9.82E−06	1.6
|af307107|af307107_1 stat4 – *Tetraodon fluviatilis*	DRFF05905	9.77E−03	1.6
|aj245569|aj245569_1 hypothetical protein; orf37 – *Mus musculus*	DRFF07241	2.72E−05	1.6
|ab051475|ab051475_1 kiaa1688 protein; kiaa1688 – *Homo sapiens*	DRFF11388	4.02E−03	1.6
|x61588|x61588_1 gtpase; rhog – *Cricetus cricetus*	DRFF06938	5.78E−03	1.6
|l10911|l10911_1 splicing factor; cc1.4 – *Homo sapiens*	DRFF00658	2.33E−04	1.5
|d13645|d13645_1 kiaa0020 – *Homo sapiens*	DRFF01283	1.99E−03	1.5
|u32107|u32107_1 type vii collagen; col7a1 – *Mus musculus*	DRFF10106	1.31E−04	1.5
Down-regulated sequences similar to GenPept			
|ab052623|ab052623_1 warm-temperature-acclimation-related-65 kda-protein; wap65 – *Cyprinus carpio*	DRFF02571	5.04E−04	−4.8
|x82278|x82278_1 fructose-bisphosphate aldolase; aldolase b – *Sparus aurata*	DRFF09029	1.81E−03	−4.1
|y00718|y00718_1 enolase aa 1-434 – *Xenopus laevis*	DRFF11091	1.94E−03	−4.0
|x82278|x82278_1 fructose-bisphosphate aldolase; aldolase b – *Sparus aurata*	DRFF04254	3.50E−03	−3.6
|ay014272|ay014272_1 fksg30; fksg30 – *Homo sapiens*	DRFF07019	8.63E−03	−3.3
|af117754|af117754_1 thyroid hormone receptor-associated protein complex component trap240 – *Homo sapiens*	DRFF06121	6.38E−03	−3.0
|ab066373|ab066373_1 glyceraldehyde 3-phosphate dehydrogenase – *Oncorhynchus mykiss*	DRFF10121	5.52E−03	−2.9
|bc004595|bc004595_1 riken cdna 0610008n23 gene – *Mus musculus*	DRFF08382	1.93E−03	−2.9
|ab011087|ab011087_1 kiaa0515 protein; kiaa0515 – *Homo sapiens*	DRFF00726	6.77E−03	−2.7
|ab037850|ab037850_1 kiaa1429 protein; kiaa1429 – *Homo sapiens*	DRFF01129	3.63E−03	−2.7
|u70439|u70439_1 silver-stainable protein ssp29 – *Homo sapiens*	DRFF01152	7.85E−03	−2.6
|af283813|af283813_1 cytochrome p450 monooxygenase cyp2k6 – *Danio rerio*	DRFF05847	3.00E−04	−2.6
|l07578|l07578_1 casein kinase i delta – *Rattus norvegicus*	DRFF01871	7.92E−03	−2.5
|af274054|af274054_1 glutathione s-transferase – *Pimephales promelas*	DRFF08801	4.20E−03	−2.5
|bc009002|bc009002_1 ras homolog gene family, member e – *Mus musculus*	DRFF12074	9.96E−03	−2.5
|ak002102|ak002102_1 unnamed protein product – *Homo sapiens*	DRFF06834	3.19E−03	−2.4
|ab066373|ab066373_1 glyceraldehyde 3-phosphate dehydrogenase – *Oncorhynchus mykiss*	DRFF07043	2.11E−03	−2.4

(*Continued*)

Table 5. Continued.

Zebrafish gene or sequence	Accession reference	*p*	fc
|u46837|u46837_1 srb7; srb7 – *Homo sapiens*	DRFF11213	2.76E−03	−2.4
|x65372|x65372_1 polypirimidine tract binding protein; ptb-4 – *Homo sapiens*	DRFF04394	6.31E−03	−2.4
|af396879|af396879_1 bullous pemphigoid antigen 1-b; bpag1 – *Mus musculus*	DRFF06457	3.06E−03	−2.3
|m14644|m14644_1 alpha-tubulin 2 – *Drosophila melanogaster*	DRFF12999	6.01E−04	−2.3
|bc014875|bc014875_1 unknown protein for mgc:6920 – *Mus musculus*	DRFF10440	6.81E−03	−2.3
|l11369|l11369_1 protocadherin 42 – *Homo sapiens*	DRFF00490	8.40E−03	−2.3
|x03018|x03018_5 histone h4 aa 1-103 – *Xenopus laevis*	DRFF00341	4.71E−03	−2.3
|m86918|m86918_1 keratin type i – *Carassius auratus*	DRFF00018	8.41E−03	−2.3
|x91637|x91637_1 brg1 protein; brg1 – *Gallus gallus*	DRFF09211	3.84E−03	−2.2
|af208159|af208159_1 electroneutral potassium-chloride cotransporter kcc2; kcc2 – *Homo sapiens*	DRFF08109	2.24E−03	−2.2
|af036717|af036717_1 fgfr signalling adaptor snt-1; snt-1 – *Homo sapiens*	DRFF03189	2.73E−03	−2.2
|ab013381|ab013381_1 homologue of high affinity ige receptor gamma subunit – *Cyprinus carpio*	DRFF05287	1.43E−03	−2.2
|ak021361|ak021361_1 data source:mgd, source key:mgi:1353557, evidence:iss neurofibromatosis 2 interacting protein putative – *Mus musculus*	DRFF04385	7.58E−03	−2.2
|ab028980|ab028980_1 kiaa1057 protein; kiaa1057 – *Homo sapiens*	DRFF12292	4.62E−03	−2.1
|af015253|af015253_1 ionotropic glutamate recetor subunit 3 alpha precursor; fglur3a – *Oreochromis mossambicus*	DRFF11380	2.62E−03	−2.1
|x74904|x74904_1 alpha-2-macroglobulin receptor – *Gallus gallus*	DRFF01962	1.27E−03	−2.1
|ak001734|ak001734_1 unnamed protein product – *Homo sapiens*	DRFF09369	1.09E−03	−2.1
|x64712|x64712_1 collagen-alpha-3 type ix – *Gallus gallus*	DRFF11424	8.65E−03	−2.1
|d43948|d43948_1 kiaa0097 protein; kiaa0097 – *Homo sapiens*	DRFF11348	2.13E−03	−2.1
|af151886|af151886_1 cgi-128 protein – *Homo sapiens*	DRFF11878	2.12E−04	−2.0
|af016042|af016042_1 beta-2 microglobulin precursor; b2m – *Ictalurus punctatus*	DRFF02503	8.55E−03	−2.0
|af302502|af302502_1 pellino 2; peli2 – *Homo sapiens*	DRFF10982	3.50E−04	−2.0
|bc004809|bc004809_1 pdz and lim domain 1 elfin – *Mus musculus*	DRFF11269	7.19E−04	−2.0
|ak001035|ak001035_1 unnamed protein product – *Homo sapiens*	DRFF00292	2.10E−03	−2.0
|af104261|af104261_1 pax transcription activation domain interacting protein ptip – *Mus musculus*	DRFF03893	9.02E−03	−2.0
|u90236|u90236_1 myosin vi; myo6 – *Homo sapiens*	DRFF01442	2.19E−03	−2.0
|af302127|af302127_1 pkc-regulated kinase pkk – *Mus musculus*	DRFF10226	1.52E−03	−2.0
|af126246|af126246_1 tip-associated protein tap – *Homo sapiens*	DRFF12902	1.29E−03	−2.0
|af057146|af057146_1 putative deubiquitinating enzyme ubpy; ubpy – *Mus musculus*	DRFF04851	8.25E−04	−2.0
|aj308233|aj308233_6 amylase-2 protein; amylase-2 – *Tetraodon nigroviridis*	DRFF09910	2.03E−03	−2.0
|bc003745|bc003745_1 similar to dead/h asp-glu-ala-asp/his box polypeptide 15 – *Mus musculus*	DRFF08012	6.37E−04	−2.0
|u13674|u13674_1 xcap-e – *Xenopus laevis*	DRFF04332	4.54E−03	−2.0

(*Continued*)

Table 5. Continued.

Zebrafish gene or sequence	Accession reference	*p*	fc
|af401579|af401579_1 ribosomal protein l24 – *Ictalurus punctatus*	DRFF09068	9.40E−03	−1.9
|x53456|x53456_1 plasma membrane ca2+ pump pmca1b – *Sus scrofa*	DRFF07079	6.70E−03	−1.9
|ab011132|ab011132_1 kiaa0560 protein; kiaa0560 – *Homo sapiens*	DRFF00189	9.36E−03	−1.9
|l78010|l78010_1 es2 – *Homo sapiens*	DRFF09720	6.53E−03	−1.9
|bc000198|bc000198_1 similar to cg11985 gene product – *Homo sapiens*	DRFF04729	6.87E−03	−1.9
|bc020051|bc020051_1 unknown protein for mgc:28179 – *Mus musculus*	DRFF11345	9.63E−03	−1.9
|u91844|u91844_1 glucose-6-phosphatase – *Canis familiaris*	DRFF06093	4.63E−03	−1.8
|aj276373|aj276373_1 mitogen inducible gene mig-6 product; mig-6 – *Homo sapiens*	DRFF12532	8.08E−03	−1.8
|ab034702|ab034702_1 xotx5; xotx5 – *Xenopus laevis*	DRFF09994	2.89E−03	−1.8
|x03104|x03104_1 h3 aa 1-135 – *Xenopus laevis*	DRFF04533	5.31E−03	−1.8
|u79284|u79284_1 sec14l – *Homo sapiens*	DRFF02299	1.76E−03	−1.8
|af255554|af255554_1 beta tubulin – *Notothenia coriiceps*	DRFF07181	4.49E−03	−1.8
|af125175|af125175_1 angiopoietin-related protein-2 – *Homo sapiens*	DRFF00655	2.52E−03	−1.8
|af310676|af310676_1 e3 ubiquitin ligase smurf2 – *Homo sapiens*	DRFF11286	1.88E−03	−1.8
|af275309|af275309_1 transcription factor foxp1; foxp1 – *Homo sapiens*	DRFF06284	1.55E−04	−1.8
|m83680|m83680_1 rab14; rab14 – *Rattus norvegicus*	DRFF01797	5.40E−03	−1.7
|aj278065|aj278065_1 putative alanine:glyoxylate aminotransferase; agt – *Xenopus laevis*	DRFF09499	9.99E−03	−1.7
|af353674|af353674_1 btb domain protein; bdpl – *Homo sapiens*	DRFF09639	2.81E−03	−1.7
|ak023742|ak023742_1 unnamed protein product – *Homo sapiens*	DRFF10475	4.10E−04	−1.6
|y08923|y08923_1 beta-globin – *Salmo salar*	DRFF07359	9.34E−03	−1.6
similar to pir-nref|nf00099164 transcription factor nf-kappa-b chain p65 – *Homo sapiens*	DRFF01700	9.28E−03	−1.6
|ak025429|ak025429_1 unnamed protein product – *Homo sapiens*	DRFF10726	5.64E−03	−1.5
|m83233|m83233_1 transcription factor; htf4 – *Homo sapiens*	DRFF01407	5.47E−03	−1.5
|al136761|al136761_1 hypothetical protein; dkfzp434i0515 – *Homo sapiens*	DRFF06298	4.69E−04	−1.5
|ak004962|ak004962_1 data source:sptr, source key:o00308, evidence:iss homolog to wwp2 putative – *Mus musculus*	DRFF06912	8.49E−03	−1.5
|ax287596|ax287596_1 unnamed protein product – *Homo sapiens*	DRFF08279	3.71E−03	−1.5
|al133174|al133174_3 dj470l14.2.1 staufen RNA binding protein isoform 1; stau – *Homo sapiens*	DRFF01151	2.60E−03	−1.5
|ab035443|ab035443_1 glycogen-debranching enzyme; agl – *Homo sapiens*	DRFF07750	3.75E−03	−1.5
similar to pir-nref|nf00474818 actin (fragment) – *Mortierella polycephala*	DRFF03447	4.34E−03	−1.5
4755345 1-474 forward	DRFF00320	2.74E−03	1.4
16375245 65-575 forward	DRFF00059	4.44E−03	9.9
17035762 16-529 forward	DRFF11021	3.13E−02	9.4
8712984 1-461 forward	DRFF12562	1.32E−03	3.5
19106447 167-667 forward	DRFF12956	1.35E−03	1.6
14389691 19-595 forward	DRFF10822	2.62E−03	−2.0
13105175 1-594 forward	DRFF06335	1.84E−03	−1.8
17241543 46-655 forward	DRFF09708	3.86E−03	−1.7

developmental and reproductive disruptions. Toxicity assessments by bioassays usually aim at known target structures and are suitable tools for risk assessment of estrogenic compounds in environmental compartments. Pollutants generally exist as mixtures of several substances in the environment. Effects of these mixtures on organisms may differ from those caused by single substances. Multi-marker approaches are crucial to analyze these effects.

One possibility to verify influences of estrogenic compounds on the expression of individual genes is the use of qPCR. This approach is a particularly useful, fast and affordable method for the examination of marker genes like vitellogenin. Our studies showed a strong expression of the gene coding for vitellogenin in liver tissue of male zebrafish exposed to 200 ng L^{-1} E2. At lower concentrations no induction of vitellogenin expression could be detected. However, it is possible that the extended exposure time (11 days) resulted in a lower sensitivity toward E2 and related compounds with respect to gene expression. Moreover, it is likely that adult zebrafish are less sensitive than juvenile ones.

The limitation of single marker approaches such as qPCR relates to the fact that they cover only a small part of possible bio-effects. In contrast, microarray experiments offer the possibility to examine a wide gene expression profile, even in individual tissues or organisms. This technique is able to examine changes of gene expression due to pollutants or other stress situations. Comparing a number of individuals allows reliable assessment of changes in gene expression profiles caused by estrogenic compounds. At the same time multiple effects can be detected.

Up to now environmental analyses by microarray experiments are rare. The impact of EDCs on *C. elegans* has been examined by Custodia *et al.* [16]. These studies showed that exposure of *C. elegans* to estrogens at a concentration of 10^{-7} M (27.2 µg L^{-1}) in liquid culture lead to an up-regulation of vitellogenin genes (*vtg*). The *vtg-2* gene showed an up-regulation of 2.7-fold and *vtg-6* of 3.3-fold.

Our studies focused on E2 responses as an initial step in a broader survey on EDCs. They confirmed the up-regulation of genes already known to be induced by estrogens. The marker gene coding for vitellogenin (*vtg1*) showed a 129-fold up-regulation at a concentration of 500 ng L^{-1} 17β-estradiol. Larkin *et al.* reported an up-regulation of *vtg* in the largemouth bass (*Micropterus salmoides*) when exposed to EDCs [17]. In addition we found a whole set of genes that are regulated by E2. A 6.7-fold up-regulation of the expression of the estrogen receptor represents a strong response to the exposure to E2. Nothepsin is a liver specific aspartic protease, known to process vitellogenin before this egg yolk precurser protein is released into the blood [18]. This protease is similar to mammalian cathepsin E and D [19] and cathepsin D is known to be induced by E2 in MCF-7 cells [20]. Our studies showed that the expression of nothepsin was up-regulated 6.8-fold by E2.

In addition to genes involved in Vtg biosynthesis, an impressive number of genes was identified, strongly linked to crucial events in embryonic development. Among them are homeo box genes. *Hoxa2a* was up-regulated 31-fold; it codes for transcription factors required for organ development [21]. The gene *dvr1* codes for a growth factor, which is assumed to play an important role in early embryogenesis [22] and is up-regulated 5.7-fold. A further gene with a related function is *mespb*, which is usually expressed in the presomitic mesoderm and controls the growth of somites [23].

An entire set of developmental genes was found to be down-regulated by E2. *plei1* is down-regulated 2.7-fold, it codes for a growth factor and could be involved

in carcinogenesis [24]. *Pou50* is down-regulated 1.8-fold and is known to play an important role in development of embryonic zebrafish forebrain [25]. *Efnb2*, down-regulated 1.9-fold, codes for Ephrin b2 and was shown to contribute to segmentation of the somites [26].

Considering the fact that EDCs interfering with the estrogen pathway particularly affect embryonic development as shown by vom Saal *et al.* [27] and Colborn *et al.* [28], the identification of underlying genes opens a better understanding of the molecular mechanisms. This is especially relevant for developmental aberrations, such as cancerogenesis.

Our results underline the role of zebrafish as an useful model organism for screening exposure to estrogenic substances including EDCs by microarray analysis. In addition, microarrays can contribute to the identification of new target genes for EDCs and a better understanding of biological pathways and mechanisms involved in EDC responses. The next step will be exposure studies with EDCs and the integration of genes regulated by EDCs into signal transduction chains. Presently it is not clear, yet, how many pathways and branching points are involved. Tools of bioinformatics are expected to clarify this issue.

Work is in progress to compare E2 effects with responses of EDCs known to interfere with the estrogen pathway. It remains to be seen whether characteristic response patterns can be identified for individual groups of EDCs.

Acknowledgement

We are indebted to the EU for supporting the EDEN project (Endocrine disrupters: Exploring Novel Endpoints, Exposure, Low-Dose and Mixture-Effects in Humans, Aquatic Wildlife and Laboratory Animals). Furthermore we would like to thank the Lehrstuhl fuer Tierhygiene at the TU Munich for inviting us to use the LightCycler.

References

[1] J.P. Sumpter, S. Jobling. *Environ. Health Perspect.*, **103**, 173 (1995).
[2] W. Wahli. *Trends Genet.*, **4**, 227 (1988).
[3] U. Langheinrich. *Bioassays*, **25**, 904 (2003).
[4] A.E. Davidson, D. Balciunas, D. Mohn, J. Schaffer, S. Hermanson, S. Sivasubbu, M.P. Cliff, P.B. Hackett, S.C. Ekker. *Dev. Biol.*, **263**, 191 (2003).
[5] [http://www.ncbi.nlm.nih.gov/genome/guide/zebrafish/index.html].
[6] M. Seifert, S. Haindl, B. Hock. *Adv. Exp. Med. Biol.*, **444**, 113 (1998).
[7] M.W. Pfaffl, G.W. Horgan, L. Dempfle. *Nucleic Acids Res.*, **30**, 30 (2002).
[8] R. Ihaka, R.C. Gentleman. *J. Comput. Graph. Statist.*, **5**, 299 (1996) [http://www.R-project.org].
[9] R.C. Gentleman, V.J. Carey, D.M. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. Leisch, C. Li, M. Maechler, A.J. Rossini, G. Sawitzki, C. Smith, G. Smyth, L. Tierney, J.Y. Yang, J. Zhang. *Genome Biol.*, **5**, R80 (2004) [http://www.bioconductor.org].
[10] R.J.A. Little, D.B. Rubin. *Statistical Analysis with Missing Data*, Wiley, New York (1987).
[11] W.S. Cleveland. *J. Amer. Statist. Assoc.*, **74**, 829 (1979).
[12] K. Fent. *Oekotoxikologie*, Georg Thieme Verlag, Stuttgart, New York (1998).
[13] J.E. Cavaco, C. Vilrokx, V.L. Trudeau, R.W. Schulz, H.J. Goos. *Am. J. Physiol.*, **275**, 1793 (1998).
[14] C. Stoker, F. Rey, H. Rodriguez, J.G. Ramos, P. Sirosky, A. Larriera, E.H. Luque, M. Munoz-de-Toro. *Gen. Comp. Endocrinol.*, **133**, 287 (2003).
[15] R.L. Hill, D.M. Janz. *Aquat. Toxicol.*, **63**, 417 (2003).
[16] N. Custodia, S.J. Won, A. Novillo, M. Wieland, C. Li, I.P. Callard. *Ann. NY Acad. Sci.*, **948**, 32 (2001).

[17] P. Larkin, T. Sabo-Attwood, J. Kelso, N.D. Denslow. *Ecotoxicology*, **12**, 463 (2003).
[18] M. Riggio, R. Scudiero, S. Filona, E. Parisi. *Gene*, **260**, 67 (2000).
[19] C. Capasso, M. Riggio, R. Scudiero, V. Carginale, G. di Prisco, J. Kay, P. Kille, E. Parisi. *Biochim. Biophy. Acta*, **1387**, 457 (1998).
[20] F. Wang, W. Porter, W. Xing, T.K. Archer, S. Safe. *Biochemistry*, **36**, 7793 (1997).
[21] A.H. Charpentier, A.K. Bednarek, R.L. Daniel, K.A. Hawkins, K.J. Laflin, S. Gaddis, M.C. MacLeoad, C.M. Aldaz. *Cancer Res.*, **60**, 5977 (2000).
[22] K.A. Helde, D.J. Grunwald. *Dev. Biol.*, **159**, 418 (1993).
[23] A. Sawada, A. Fritz, Y.J. Jiang, A. Yamamoto, K. Yamasu, A. Kuroiwa, Y. Saga, H. Takeda. *Development*, **127**, 1691 (2000).
[24] K. Kadomatsu, T. Muramatsu. *Cancer Lett.*, **204**, 127 (2002).
[25] G. Hauptmann, T. Gerster. *Mech. Dev.*, **91**, 105 (2000).
[26] L. Durbin, C. Brennan, K. Shiomi, J. Cooke, A. Barrios, S. Shanmugalingam, B. Guthrie, R. Lindberg, N. Holder. *Genes Dev.*, **12**, 3096 (1998).
[27] F.S. vom Saal, S.C. Nagel, P. Palanza, M. Boechler, S. Parmigiani, W.V. Welshons. *Toxicol. Lett.*, **77**, 343 (1995).
[28] T. Colborn, D. Dumanowski, J.P. Myers. *Environment*, **3**, 42 (1996).