

Final Report

Validation of an Endocrine Disrupting Chemical (EDC) Microarray Gene Chip for the Detection of EDC Mixtures in Ambient Water

Submitted to:

John Borkovich and Jan Stepek

**SWRCB Contract# 04-111-250-0
Groundwater Ambient Monitoring Program
State Water Resources Control Board
California Environmental Protection Agency
Sacramento, California 95814**

Date:

March 31st 2009

By

Swee J. Teh, Petra Kunz, and Ping-Shi Wu

**Aquatic Toxicology Program
Department of Anatomy, Physiology, and Cell Biology
School of Veterinary Medicine,
University of California at Davis
Davis, California 95616
Phone (530)-754-8183
Fax (530)-752- 7690
Email: sjteh@ucdavis.edu**

EXECUTIVE SUMMARY

Scientists at the Aquatic Toxicology Program at UC Davis have developed a medaka fish model DNA microarray to evaluate the expression of genes associated with the presence of endocrine disrupting chemicals (EDCs) in surface and groundwater. There is a tremendous potential in the application of DNA microarray technology as a screening tool to determine the presence of EDCs in the aquatic environment. Developing and using this approach will greatly benefit the Groundwater Ambient Monitoring Program (GAMA) in their effort to monitor the occurrence of EDCs that have profound implications when present in ground and surface waters that affect both human and wildlife.

The Qurt medaka (*Oryzias latipes*) is an ideal model for studying EDCs due to its short generation time (6-8 wks) and daily year-round spawning under controlled photoperiod and temperature conditions. Gonad histopathology in medaka is an essential core endpoint for the rapid detection of endocrine disrupting activity. The Qurt strain medaka is a gonochoristic species with an XX/XY sex determination system. The genotypic sex can be identified at early embryonic stages due to sex-linked occurrence of leucophores. As leucophore differentiation occurs at the 2-day-old embryo stage in males, but not in females, the genotypic sex can be identified before sexual differentiation, which facilitates the examination of gender-specific effects of EDCs during critical stages of development.

We selected three receptor agonist/antagonist as prototype EDCs: estrogen/anti-estrogen [i.e., 17- β -estradiol/faslodexTM (ICI 182,780)]; androgen/anti-androgen (i.e., 11-ketotestosterone/flutamide); and thyroidogen/ anti-thyroidogen (i.e., thyroid hormone (T3)/amiodarone). Biologically effective concentrations (BEC) for each of the six prototypic EDCs were identified by conducting a series of range-finding experiments using 1-week-old male and female Qurt medaka larvae.

In coordination efforts with LLNL (Bradley Esser and Alan Grayson) and SWRCB (Jan Stepek and John Borkovich), we collected 4, 4-L of UV-treated post-secondary non-chlorinated wastewater (2°-wastewater) from a wastewater treatment plant and 4, 4-L bottles of groundwater from a domestic well near a septic system in Livermore (L-groundwater), California to test for the presence of endocrine disrupting chemicals. Standard medaka acute toxicity testing revealed no mortality in L-groundwater at the end of 96 hr and >95% mortality in 2°-wastewater within 24 hrs of exposure. In addition, 92.5% mortality was observed at 96 hr when 2°-wastewater was diluted with 50% medaka culture water. Therefore, microarray analyses were performed on 50%-2°-wastewater at 6 hr and on L-groundwater at 6 and 96hr. DNA from a pooled sample size of 200, 1-week-old medaka larvae per gender was extracted and subjected to microarray analysis for gene expression evaluation. A Functional Classification-Kernel Density Approach model was used to predict and classify L-groundwater and 50%-2°-wastewater as either no EDC activity or belong to one of the six prototypic EDCs.

Our results indicate that L-groundwater has anti-thyroidogenic effects in both female and male medaka larvae at the end of 96 hr of exposure while 50%-2°-wastewater has anti-androgenic effects on female and thyroidogenic effects on male at the end of 6 hr. These results strongly suggest the presence of complex mixtures of EDCs in L-groundwater and 50%-2°-wastewater. Finally, initial findings indicate EDCs in surface and groundwater may be detected using the

medaka biologically-based screening assay with 75% certainty. While preliminary results are promising, further study is warranted to validate the consistency of our methodology and findings prior to dissemination and implementation of medaka DNA microarray as a diagnostic tool.

INTRODUCTION AND PROJECT OVERVIEW

This project uses medaka (*Oryzias latipes*) as a fish model system to identify microarray-based gene expression profiles that are characteristic of endocrine-disrupting chemicals (EDCs). The overall goal is to detect the presence of EDCs in ambient water – whether the EDCs are present as a component of a mixture or as an individual compound – based on comparisons to gene expression patterns established with prototypical EDCs.

Task 1-EDC Mixture Validation

1-1. Development of Medaka Microarray Gene Chip

A custom 12K oligonucleotide array for medaka was developed in collaboration with CombiMatrix, Corporation (WA, USA). The medaka microarray elements include: EDC-affected genes (n=131), medaka cDNAs (n=1075), medaka genome project sequences (n=890), and unigene database probes (n=8154). In general, there are 9,379 unique features and more than 3,000 control probes. Specific probes were designed by submitting GenBank accession numbers of gene sequences identified in the medaka databases to the CombiMatrix, Corp. web-interfaced probe design software. The medaka CustomArray™ probes have intermediate lengths (35 nucleotides) that combine the benefits of enhanced specificity and strong signals. There are four replicates of each probe distributed across each array which were synthesized on a semiconductor chip using phosphoramidite chemistry under electrochemical control.

Biological samples were prepared from Qurt medaka larvae poly A+ RNA after 6-h exposure to each of the six prototypic EDCs at biological effective concentrations (BECs). Briefly, total RNA was isolated from whole tissue of 7 day (± 1) old larvae (n= 200 larvae/gender and three replicates/treatment) using TRIZOL reagent (Invitrogen). The integrity of the isolated RNA was verified on a 1% agarose gel. Subsequently, poly A+ RNA was isolated using Qiagen Oligotex midi-kit (Valencia, CA). First strand cDNA incorporating amino allyl-dUTP was synthesized from poly A+ RNA, and CyDye-labelled cDNA was purified by using the Amersham Biosciences (USA) CyScribe cDNA Post Labelling Kit. UV/Visible spectrophotometry measurements were used to calculate the concentration of fluorescent Cy3- and Cy5-Dye incorporated in the cDNA sample. Spike-in control transcripts were prepared similarly using PCR fragments, containing a T7 RNA Polymerase promoter site, as template for transcription.

The microarrays were assembled with hybridization caps and rehydrated with RNase-free water (Ambion) at 65°C for 10 minutes. After rehydrating, blocking solution (6X SSPE, 20mM EDTA, 0.05% Tween-20, 5X Denhardt's Solution, 0.05% SDS, 100ng/ μ l sonicated salmon sperm DNA) was added and the arrays were incubated at hybridization temperature (50°C) for 30 minutes. The labelled cDNA sample was added to hybridization solution (6X SSPE, 0.05% Tween-20, 20mM

EDTA, 0.05% SDS, 100 ng/μl sonicated salmon sperm DNA) and denatured for 3 minutes at 95°C. Samples were placed briefly on ice followed by centrifugation at 14,000 rpm (maximum) for 3 minutes. Blocking solution was removed from the hybridization chamber and 100μl of hybridization solution was applied to the arrays. Hybridization was carried out in a Fisher Scientific Isotemp hybridization incubator for 18 hours at 45°C under gentle rotation. Following hybridization, arrays were washed at hybridization temperature (50°C) for 5 minutes with 6X SSPE, 0.05% Tween-20 pre-heated to 50°C. Washing continued with 3X SSPE, 0.05% Tween-20 for 5 minutes at room temperature (RT), 0.5X SSPE, 0.05% Tween-20 for 5 minutes at RT and 2X PBS, 0.1% Tween-20. Final washing steps were performed at RT for 5 minutes with two rounds of 2X PBS with no detergent. The microarrays were imaged with Cy5 filter sets on an Applied Precision (Issaquah, WA) arrayW0Rx™ Biochip Reader. Imaging was performed while the array was wet with Imaging Solution (2XPBS) under a LifterSlip™ glass cover slip (Erie Scientific, Portsmouth, NH). Images were analyzed using ScanAlyze® v. 2.5 (Stanford University, USA).

ScanAlyze® v.2.5 was used to estimate the gene expression levels by obtaining the median intensity of replicate spots. An initial set of differentially expressed genes in each treatment was obtained by conducting a one sample t-test with pooled variance. Differentially expressed genes were identified by the Benjamini and Hochberg (1995) procedure for False Discovery Rate (FDR) control at 0.01 and 0.05 values. Differential expression of a set of genes involved in sex differentiation and growth in medaka and other vertebrates was summarized by heat map visualization using GenePattern® 2.0 (MIT and Harvard, USA).

Loess normalization and functional embedding were used for high dimensional data analyses and pattern discovery. First, the array data was normalized with the loess normalization scheme (degree=2; span=1/6). The argument “degree” indicates that the fit is a local quadratic fit while “span” indicates the neighborhoods window size in terms of proportion of the sample. For each array, a loess fit of logarithm of intensity ratio (denoted by M) $\log(\text{Cy5}/\text{Cy3})$ on averaged logged intensity (denoted by A) $\frac{1}{2} [\log(\text{Cy3}) + \log(\text{Cy5})]$ was removed from $\log(\text{Cy5}/\text{Cy3})$. Before the functional embedding step, a gene-screening procedure was conducted for each array. For each gene, a two sample t-test was performed to obtain a significance value (p-value) for the difference between chemical treatments. A working set of diagnostic genes was defined by selecting the top 50 genes with the most significant differences (smallest p-values) between treatments. This working set of genes is believed to provide the most powerful means to discriminate the chemical treatments at the transcription level. The latent ordering of the working set of genes was retrieved and used to re-express the subject profiles (columns) in functional form. Latent ordering of genes allowed for pattern identification and analysis with the technique of functional embedding. This technique was applied to discover the pattern of gene expression for Qurt medaka under exposure to the selected EDCs. To further illustrate the EDC-specific gene patterns, the response of each group of genes was represented by smoothed mean curves with standard error bars. The mean curves were smoothed with scatter smoother which is more robust than the cross section mean approach.

1-2 Functional Approach to Predict EDCs

We have combined several statistical methods, called Functional Classification- Kernel Density Approach (FCKDA) to predict and classify EDCs (Appendix I).

1-2.1 Gene Screening by pairwise t-tests

All the entities are first filtered/screened by the significance of t-tests to initiate the basis of genetic discrimination foundation.

1-2.2 Functional Embedding:

Gene expression profiles consisting of the pre-screened entities are further embedded into functional space according to the correlation structure.

1-2.3 Functional Principal Component Analysis.

After Functional Embedding, each gene expression profile is regarded as a realization from the underlying random function. In other words, $X_{ip} \xrightarrow{\text{i.i.d.}} X_p$ with $E(X_p) = \mu(t)$ and the variance-covariance surface $\text{Cov}(X_p(s), X_p(t)) = G(s, t)$

Define the Covariance Operator: $(A_G f)(t) = \int f(s)G(s, t)ds$, for $f \in L^2$, any functional in L^2 Space. We adopt L^2 orthogonal decomposition of the covariance surface, i.e. $G(s, t) = \sum_{k=1}^{\infty} \lambda_k \phi_k(s) \phi_k(t)$ where

$$(A_G \phi_k(t)) = \lambda_k \phi_k(t) \text{ and } \int \phi_j(t) \phi_k(t) dt = \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases}.$$

The L^2 -orthogonal representation of functional expression profile is:

$$X_{ip}(t) = \mu(t) + \sum_{k=1}^{\infty} \xi_{ik} \phi_k(t), \text{ where } \xi_{ik} = \int (X_{ip}(t) - \mu(t)) \phi_k(t) dt.$$

Considering the measurement error, the contaminated functional expression profile can be formed by the following:

$$Y_{ij} = X_{ip}(t_{ij}) + \varepsilon_{ij} = \mu(t_{ij}) + \sum_{k=1}^{\infty} \xi_{ik} \phi_k(t_{ij}) + \varepsilon_{ij}, t_{ij} \in T.$$

This contaminated version will be used in the real implementation. The resulting functional principal component scores that indicating the projection of functional expression profiles onto the feature space will then be used for downstream classification task.

1-2.4 Kernel Density Classification

The posterior probability for assigning a case in feature space to group k is defined via the usage of spherical Kernel function that uses L^2 norm as the metric, and can be written as the following..

$$P(C(\vec{\xi}) = k | \vec{\xi}) = E(I\{C(\vec{\xi}) = k\} | I\{\vec{\xi} \in S(\vec{\xi})\}) = \frac{\sum_{C_k} K(\|\vec{\xi}_i - \vec{\xi}\|)}{\sum_{S(\vec{\xi})} K(\|\vec{\xi}_i - \vec{\xi}\|)}$$

The labeling of the target case will then be determined by maximizing the posterior probability, i.e. the label is determined as

$$C^*(\vec{\xi}) = \text{Argmax}_{1 \leq i \leq 6} P(C^*(\vec{\xi}) = i | \vec{\xi})$$

One will mark the target case as “doubt”, if $\text{Max}_{1 \leq i \leq 6} P(C(\vec{\xi}) = i | \vec{\xi}) < \frac{1}{6}$

1-2.5 Leave-One-Out Cross Validation:

To furnish the whole task of FCKDA depends upon the choice of truncation parameter that determines the dimension and the complexity of the feature space. We use Leave-One-Out Cross Validation to fulfill this task. Define the corresponding Variance-Covariance surface, contaminated L^2 -orthogonal representation, functional principal component scores, posterior probability, classifier as the following

$$G^{(-i)}(s, t) = \sum_{k=1}^{\infty} \lambda_k^{(-i)} \phi_k^{(-i)}(s) \phi_k^{(-i)}(t)$$

$$Y_{ij} = \mu^{(-i)}(t_{ij}) + \sum_{k=1}^{\infty} \xi_{ik}^{(-i)} \phi_k^{(-i)}(t_{ij}) + \varepsilon_{ij}, t_{ij} \in T$$

$$\xi_{ik}^{(-i)} = \int (Y_{ip}(t) - \mu^{(-i)}(t)) \phi_k^{(-i)}(t) dt$$

$$P^{(-i)}(C(\vec{\xi}_i) = k | \vec{\xi}_i) = E(I\{C(\vec{\xi}_i) = k\} | I\{\vec{\Xi}_j^{(-i)} \in S(\vec{\xi}_i)\}) = \frac{\sum_{C_k} K(\|\vec{\Xi}_j^{(-i)} - \vec{\xi}_i\|)}{\sum_{S(\vec{\xi}_i)} K(\|\vec{\Xi}_j^{(-i)} - \vec{\xi}_i\|)}$$

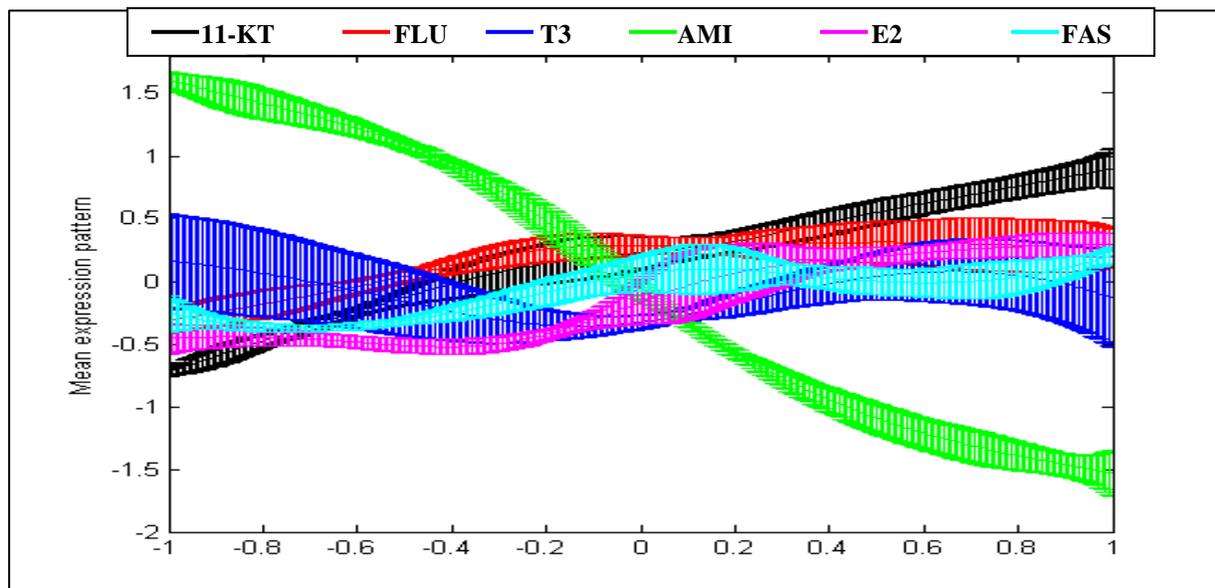
$$C^{*(-i)}(\vec{\xi}_i) = \text{Argmax}_{1 \leq k \leq 6} P^{(-i)}(C(\vec{\xi}_i) = k | \vec{\xi}_i)$$

The prediction error corresponding to truncation parameter M is basically the proportion of misclassified cases.

$$LOOCV(M) = \frac{\sum_i I\{C^{*(-i)}(\vec{\xi}_i) \neq C_0(\vec{\xi}_i)\}}{n}$$

We here just choose the value that minimizes the prediction error, i.e. $M^* = \text{Argmin}_M LOOCV(M)$.

Figure 1 shows the grouped mean expression profiles of six prototypic EDCs with error bars indicating the within group variation.



Six prototypic EDCs used in this study: 11-Ketotestosterone (11-KT) is the primary endogenous fish androgen. Since 11-KT is not converted to estradiol by P450 aromatase, the differential gene expression observed in exposed medaka fish should be that expected from a pure androgenic agent. Flutamide (FLU) is a non-steroidal compound widely used in endocrinology research. FLU is a therapeutic agent for prostate cancer treatment that can function as an antiandrogen in mammals. 17- β Estradiol (E2) is one of the most studied natural estrogens. FaslodexTM (FAS), ICI 182,780AstraZeneca, UK is a novel, steroidal estrogen antagonist that was specifically designed to be devoid of estrogen agonist activity. 3, 3', 5-Triiodo-L-thyronine (T3) is one of two iodinated hormones secreted by the thyroid gland. T3 is essential for growth, differentiation, and reproductive system development. T3 primarily exerts its biological effects through a receptor-mediated mechanism. Furthermore, T3 modulates its own function by regulating the thyroid receptor (TR) levels. Amiodarone (AMI) is a diiodinated benzofuran derivative that functions as a thyroid receptor antagonist.

In summary, the medaka microarray, coupled with robust statistical techniques, can differentiate between six biological categories of EDCs on the basis of gene expression patterns. For instance, the expression profile for vinclozolin, an antiandrogenic fungicide, is remarkably similar to the prototypical antiandrogen flutamide, indicating that our system can prospectively identify EDCs with characteristic modes of action (Appendix II). However, when tested with the microarray chip using a simple mixture of two prototypic EDCs [i.e., (17- β -estradiol (E2) + faslodex (FAS); 11-ketotestosterone (11-KT) + flutamide (FLU); and thyroid hormone (T3) + amiodarone (AMI)], the outcomes were different from when medaka larvae were exposed to individual prototypic EDC (Appendix III). Except for E2+FAS mixture which had similar estrogenic expression profile to fish exposed to E2 alone, the expression profiles in medaka larvae exposed to 11-KT+FLU and T3+AMI mixtures were remarkably different from those resulting from exposures to any of the four prototypic EDCs alone, suggesting the possible occurrence of antagonistic, competitive, inhibitory, and/or synergistic effects of EDC mixtures (Appendix III).

Task 2- EDCs in Ambient Groundwater

2-1. Exposure of Qurt Medaka Larvae to Ground- and 2° Wastewater

In order to assess water quality and acute toxicity, we collected 4 liters of UV-treated post-secondary non-chlorinated wastewater with the assistance of Mr. Bradley Esser and Mr. Allen Grayson, both from LLNL, on January 15th 2008. Due to high ammonia concentration (0.2 mg/L) in the water, 100% of the 7 day post-hatched (dph) medaka larvae (n=200) and 3 month-old adults (n=10) died within 24h of the exposure.

On a second attempt on January 30th 2008, with support from LLNL (Mr. Bradley Esser and Mr. Allen Grayson, we collected the following water samples for 6h EDC screening and 96h acute toxicity testing using 7 dph medaka larvae:

1. 4, 4-L bottles of UV-treated post-secondary non-chlorinated wastewater (2°-wastewater)
2. 4, 4-L bottles of Livermore groundwater (L-groundwater)

Prior to the actual EDC screening experiment, 100 medaka larvae (7 dph) were initially exposed to the 2°-wastewater and L-groundwater for 96 hr to determine acute toxicity of the water samples. No toxicity towards the larvae was observed for the L-groundwater, whereas >95 % mortality was observed for the larvae exposed to the 2°-wastewater after 24 hr.

Based on these findings, we decided to use a 50% dilution of the 2°-wastewater, and undiluted L-groundwater for the EDC and toxicity testing. Survival and microarray analysis after 6 and 96 hr in male and female medaka larvae (200 each, 7 dph) were carried out as follows:

6-hour exposure:

CONTROL: - 200 male Qurt Medaka larvae, 7 dph
- 200 female Qurt Medaka larvae, 7 dph

L-groundwater: - 200 male Qurt Medaka larvae, 7 dph
- 200 female Qurt Medaka larvae, 7 dph

50%-2°-wastewater: - 200 male Qurt Medaka larvae, 7 dph
- 200 female Qurt Medaka larvae, 7 dph

96-hour exposure:

CONTROL: - 200 male Qurt Medaka larvae, 7 dph
- 200 female Qurt Medaka larvae, 7 dph

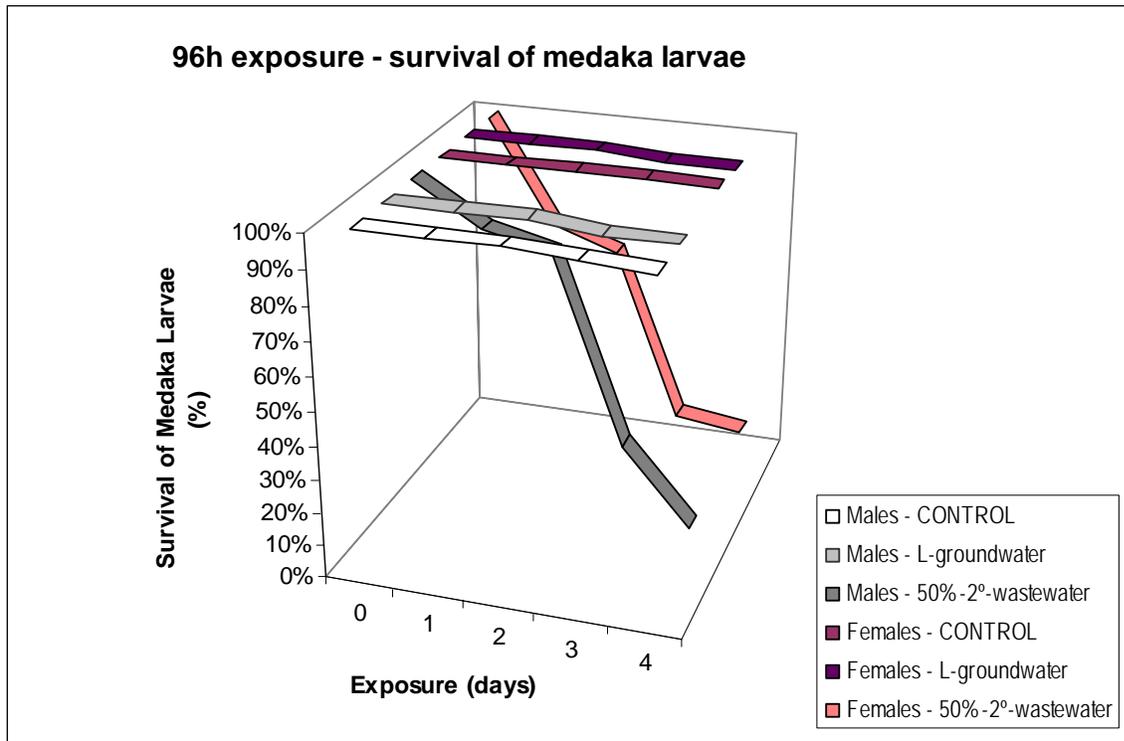
L-groundwater: - 200 male Qurt Medaka larvae, 7 dph
- 200 female Qurt Medaka larvae, 7 dph

50%-2°-wastewater: - 200 male Qurt Medaka larvae, 7 dph
- 200 female Qurt Medaka larvae, 7 dph

Larvae were not fed during the exposure trials to avoid absorption of toxic or endocrine disrupting chemicals, which may be present in these water samples, onto food particles. During the 96h exposure, 80% of the exposure waters were renewed and mortalities were recorded daily.

2-2. Results on water sample toxicity

During the 96 hr exposure, mortalities were recorded daily.



After 96 hr of exposure, larvae exposed to CONTROL and L-groundwater showed only little mortality rates (female =0.5-1.5% and male = 2.0-2.5 %), whereas mortality rate for 50%-2°-wastewater were 92.5 (female) and 94% (male), respectively, possibly because of the 100-times higher ammonia levels (0.22 mg/L NH₃ at pH 6.9) compared to the L-groundwater (0.0021 mg/L NH₃ at pH 7.7) and Control (not detectable at pH 7.7).

7 day old Qurt medaka Larvae after 96h exposure	Males				Females			
	survival		mortality		survival		mortality	
	no.	%	no.	%	no.	%	no.	%
CONTROL	195	97.5	5	2.5	199	99.5	1	0.5
L-groundwater	196	98.0	4	2.0	197	98.5	3	1.5
50%2°-wastewater	12	6.0	188	94.0	15	7.5	185	92.5

2-3. Analysis for EDC gene expression profiles

At the end of the exposures (6 hr and 96 hr), 200 larvae/treatment were collected from each replicate for gene expression analysis using the medaka gene chip. Gene expression profiles were characterized and genes affected by L-groundwater at 6 and 96 hr and 50% 2°-wastewater at 6 hr were identified. Because of the high mortality observed after 24 hours, gene expression profiles for 50%-2°-wastewater were assessed after only 6h of exposure as the surviving larvae after 96h did not provide enough tissue for analysis.

2-3-1 Identification and characterization of gene expression in medaka

Gene expression profiles in L-groundwater and 50%-2°-wastewater-exposed medaka were developed by first using Trizol to isolate total RNA from each replicate. cDNA was then synthesized from the total mRNA, purified, labeled, and hybridized to our medaka gene chips (Please see Task 1-1 for details of experimental design). A computational approach was used to identify the unique gene response patterns of the L-groundwater and 50%-2°-wastewater samples and to categorize the type of EDC which are present in these samples based on the prototypic EDCs gene patterns developed in task 1-2. Briefly, unique gene patterns for L-groundwater and 50%-2°-wastewater were identified by using ScanAlyze v.2.5 (Stanford University, USA) to estimate the gene expression levels by obtaining the median intensity of replicate spots from the image. An initial set of differentially expressed genes in each treatment was obtained by conducting a one sample t-test with pooled variance. Differentially expressed genes were then identified for False Discovery Rate (FDR) control at 0.01 and 0.05 values. Differential expression of a set of genes involved in sex differentiation and growth in medaka and other vertebrates were summarized by heat map visualization using GenePattern© 2.0 (MIT and Harvard, USA). Loess normalization and functional embedding was used for high dimensional data analyses and pattern discovery. First, the array data was normalized with the loess normalization scheme. For each gene, a two sample t-test was performed to obtain a significant value (p-value) for the difference between treatments and controls.

For further analysis on single gene expression, only significantly affected genes that were present after 6 and 96 hrs of exposure were taken into account. This was done in order to filter out genes that reflect nonspecific properties of the compound administration and non-specific stress responses. Biological information on significantly affected genes was obtained by functional profiling of genes using sequence similarity analyses with Basic Local Alignment Search Tool (BLAST). Gene annotation was based on the lowest expectation values (e-value) obtained from sequence similarity searches. The threshold value for establishing gene identity and function was $E\text{-value} \leq e^{-10}$. Non-medaka sequences were considered orthologous to a medaka gene if the e-value was within the established threshold.

2-4. Results of L-Groundwater and 2°-wastewater exposure

2-4-1. L-Groundwater exposure for 6 and 96 hrs:

2-4-1-1 Patterning and Classification of EDCs through Gene Expression Profile

The classification task of endocrine disrupting chemicals is fulfilled through the application of series of functional data analytic approaches [see 1-2. Functional Classification- Kernel Density

Approach (FCKDA) to predict and classify EDCs]. Briefly, to facilitate the construction of the classifier for EDCs, the data harvested in Figure 1 was used. Genes on the medaka chip are first screened to yield a refined set of genes, which were chosen in order to have a most significant distinction between EDC treatments. In our study, a set of 81 genes were screened, each yielding at least 9 out of 15 significant pair-wise comparisons. The gene expression profiles based on this refined set of genes are then processed by functional embedding to be embedded into functional space for overcoming the difficulty that arises whenever high dimension and low sample size data is analyzed. The mean gene expression profile could then be retrieved through scatter smoothing separately for each of the EDC treatment, and used as the prototype for the EDC gene expression patterns. As for the purpose of classification, a functional multinomial logit model is adopted to predict polychotomous labeling based on truncated functional principal component scores. The truncation is determined to yield the least misclassification error through a leave-one-out-cross validation in the training data set.

After the functional multinomial logit classifier (FMLC) was harvested from the training data set, Medaka gene expression data of L-Groundwater samples were then processed and embedded into the same functional space delineated from the training data where the classifier was built. The posterior probabilities calculated from FMLC then indicate the likelihoods for the Livermore expression data to be assigned to one of the 6 prototypical EDCs. The prototypic EDC with the highest posterior probability will then be assigned as the type of EDC that is most likely contaminating the water sample. Following the above procedure, the posterior probabilities rounded to two decimal points of the Livermore groundwater samples are shown in the following table.

Medaka Samples-hrs post exposure	11-KT	Flu	T3	Ami	E2	Fas
Male 6h groundwater	0.00	0.01	0.00	0.00	0.99	0.00
Male 96h groundwater	0.00	0.00	0.00	1.00	0.00	0.00
Female 6h groundwater	0.05	0.00	0.00	0.00	0.00	0.95
Female 96h groundwater	0.00	0.00	0.00	1.00	0.00	0.00

11-KT= androgenic; Flu= anti-androgenic; T3= thyroidogenic or goitrogenic; AMI= anti-thyroidogenic or anti-goitrogenic; E2=estrogenic; and Fas= anti-estrogenic

2-4-1-2 Differentially Expressed Genes

The exposure of male and female medaka larvae to L-Groundwater led to the identification of 104 and 19 differentially expressed genes (≥ 2 -times up- or down-regulated), respectively that were present after 6 and 96 hrs of exposure (Table 1). Males had a higher number of L-Groundwater responsive genes and they expressed far more down-regulated genes. Overall, much more genes were down- than up-regulated, with predominant inhibition responses for both sexes. We found ratios of up-regulation to down-regulation after 6hr of 0.12 for females and 0.02 for males, and 0.19 for females and 0.09 for males after 96hr, respectively.

Of the 104 (males) and 19 (females) differentially expressed genes, we were able to obtain biological information by functional profiling of 11 (males) and 7 (females) genes, respectively.

For all other genes, the function is either currently unknown or e-values obtained from the sequence similarity searches were too high to result in a reliable biological function. For male medaka larvae the exposure to L-Groundwater altered the transcription and led to a down-regulation of genes involved in metabolism, transport, multidrug resistance, vitellogenesis, xenobiotic biotransformation, and the endocrine system (Table 2). Most of these genes were significantly down-regulated after 6hr and remained down-regulated after 96hr of exposure, pointing towards an organismic response elicited by the exposure to L-Groundwater. Female genes were differently expressed with an up-regulated genes involved in the regulation of cytoskeleton stability, and down-regulation of genes involved in osmotic and cardiovascular homeostasis, iron storage, RNA/DNA metabolism and transcription and the endocrine system (Table 2) after 6 and 96 hrs of exposure.

Only one differentially expressed gene was present in males and females after 6 and 96 hrs, the ZPC domain containing protein. This is a known zona pellicula protein, which was previously found to be a suitable biomarker for anti-thyroid activity when significantly down-regulated (Leon-Cardona et al. 2008). In addition, females exposed to L-groundwater showed a significant down-regulation of the ferritin H3 gene, which plays a role in iron storage and was suggested by Leon-Cardona et al. (2008) to be a suitable biomarker for anti-androgenic activity when down-regulated. In males, three genes with potential as anti-estrogenic biomarkers were significantly down-regulated, the vitellogenin-1-, L-SF precursor- and choriogenin L gene, in addition to the down regulation of glutathione S-transferase, a putative biomarker for anti-androgenic activity (Leon-Cardona et al. 2008).

Table 1 Number of differentially expressed genes identified in medaka larvae 6hr after exposure to UV-treated post-secondary non-chlorinated wastewater (50%-2°-wastewater, Livermore Water Reclamation Plant) and 6 and 96 hr after exposure to Livermore groundwater (L-groundwater)

Treatment		Up-regulated	Down-regulated	Total
50%-2°-wastewater:	Males	22	36	58
	Females	12	134	146
L-groundwater 6h:	Males	2	102	104
	Females	2	17	19
L-groundwater 96h:	Males	9	95	104
	Females	3	16	19

Table 2 Functional profiles of a subset of genes differentially expressed after exposure of male and female Qurt medaka to L-Groundwater

Gene Name	Accession#	E-value	Biological Function	Regulation	
				males	
				6h	96h
Putative: Coatamer protein complex, subunit alpha	BJ7506743	3E-118	Golgi-endoplasmic reticulum Transport	↓	↑
ATP-binding cassette, sub-family E (OABP), member 1 (abce1)	BJ710431	8E-162	Mainly regulation of protein synthesis or expression (belongs to ABC transporter, multidrug resistance)	↓	↑
Predicted: <i>Danio rerio</i> amyloid beta (A4) precursor protein-binding, family A, member 2 (X11-like, mRNA)	BJ719387	8E-158	Integral membrane protein, primary function unknown, possibly regulator of synapse formation and neural plasticity	↓	↑
ATP-ase protein	AU168249	2E-114	Transport, metabolism, energy	↓	↓
Vitellogenin 1 (ol-vit1)	AU179789 AU180172	0	Vitellogenesis.	↓	↓
Glutathione S-transferase, mRNA	AV670256	7E-99	Phase II biotransformation of xenobiotics	↓	↓
L-SF precursor (1-sf), mRNA	AU 180097	0	hypothalamic-pituitary-gonadal (HPG) axis	↓	↓
ZPC domain containing protein 5	AF128813 AU241961	0	Zona pellicula protein	↓	↓
Choriogenin L, mRNA	AU179813	0	Zona radiate protein	↓	↓

Gene Name	Accession#	E-value	Biological Function	Regulation	
				females	
				6h	96h
Microtubule associated protein RP/EB family (EB1)	BJ740475	4E-74	Microtubule dynamics, cytoskeleton stability regulation, chromosome stability	↑	↑
Predicted: similar kelch-like protein 20	BJ748092	8E-26	Function unknown, associated with actin tails	↑	↑
ZPC domain containing protein 5	AU241961	0	Zona pellicula protein	↓	↓
Membrane guanylyl cyclase OIGC1	AU169372	0	Natriuretic peptide receptor subfamily, osmotic and cardiovascular homeostasis	↓	↓
Ferritin H3, mRNA	AU242757	5E-107	Iron storage	↓	↓
Putative: THO complex 3 (<i>D. rerio</i>)	BJ746432	0	RNA DNA metabolism, transcription	↓	↓
rpS16 mRNA for ribosomal protein S16	BJ543498	0	Ribosomal subunit, catalysis of mRNA-directed protein synthesis (overexpression in tumors)	↓	↓

↑ and ↓ indicate upregulation and downregulation of genes, respectively. BLAST similarity determined by the lowest expectation (E)-values on NR Database using accession numbers as identifiers.

2-4-2. 50%-2^o-wastewater exposure for 6 hours

2-4-2-1 Patterning and Classification of EDCs through Gene Expression Profiles

The same process of patterning and classification described in section 2-4-1-2 for L-Groundwater sample was used for 50%-2^o-wastewater sample. The medaka gene expression data for 50%-2^o-wastewater samples were first processed and embedded into the same functional space delineated by the training data where the classifier was built. The posterior probabilities calculated from FMLC then indicate the likelihoods for the 50%-2^o-wastewater expression data to be assigned to one of the 6 prototypical EDCs. The EDC with the highest posterior probability will then be assigned as the type of EDC that tends to contaminate the water sample. Our findings for the 50%-2^o-wastewater are listed below:

Sample	11-KT	Flu	T3	Ami	E2	Fas
Male 6h 50%-2 ^o -wastewater	0.00	0.00	1.00	0.00	0.00	0.00
Female 6h 50%-2 ^o -wastewater	0.00	1.00	0.00	0.00	0.00	0.00

11-KT= androgenic; Flu= anti-androgenic; T3= thyroidogenic or goitrogenic; Ami= anti-thyroidogenic or anti-goitrogenic; E2=estrogenic; and Fas= anti-estrogenic

2-4-2-2 Differentially Expressed Genes

For male and female Qurt medaka larvae exposed to a 50%-2^o-wastewater, gene expression patterns could only be analyzed after 6 hours of exposure because of the acute toxicity of the water sample, resulting in 92.5 to 94% mortality after 96 hours of exposure.

6hr exposure of medaka larvae showed 58 differentially expressed genes in males and to 146 differentially expressed genes in females (Table 1). Contrary to the L-groundwater exposure, females had a higher number of responsive genes after exposure to 50%-2^o-wastewater and they expressed far more down-regulated genes. Again, as already observed for L-groundwater, much more genes were down- than up-regulated, with predominant inhibition responses for both sexes. After 6hr the ratios of up-regulation to down-regulation was 0.09 for females and 0.61 for males, respectively.

We were able to obtain biological information by functional profiling of 16 of the 58 genes differentially expressed by males and for 32 of 146 genes differentially expressed by females. In males and females, we found genes involved in metabolism, immune, inflammatory and stress responses, regulation of cytoskeleton stability, osmotic and cardiovascular homeostasis, neuropeptide regulation, RNA/DNA metabolism and transcription, and the hypothalamic-pituitary-gonadal axis (Table 3). In addition, females showed differentially expressed genes involved in morphogenesis, sex differentiation and determination, oogenesis and vitellogenesis, and transport (Table 3).

Of those differentially expressed genes with known functions, 9 were present in males and females after 6hr of exposure to 50%-2^o-wastewater. Among these genes, up-regulation of glycoprotein (rh50) and cytochrome P4501A (CYP1A) genes may indicate anti-androgenic activity and increased immune response and metabolism. The down-regulation of L-SF precursor (1-sf) and choriogenin Hminor (chm) suggests the presence of anti-estrogenic components in the

water sample. In addition we observed an up-regulation of the microtubule associated protein (EB1), which is involved in the regulation of cytoskeleton and chromosome stability. The function of the gene for predicted kelch-like protein 20 (klp20) is currently unknown but is most likely associated with actin tails, which play a role in cytoskeleton stability. The down-regulated B-type natriuretic peptide precursor (cnp) is involved in the cardiovascular system, whereas the down-regulated heme-oxygenase 1 gene (ho1) is a stress response gene, indicating the presence of stressors such as oxidative stress, hypoxia, heavy metals or cytokines when up-regulated. The putative THO complex 3 gene may indicate a decrease in RNA/DNA metabolism and transcription activity.

Table 3 Functional profiles of a subset of genes differentially expressed in medaka larvae after 6hr of exposure to 50%-2°-wastewater

Gene Name	Accession#	E-value	Biological Function	Regulation Males
Glycoprotein (rh50)	AB036511	0.00	Immune response	↑
cytochrome P450 1A (CYP1A) mRNA	AY297923	0.00	Phase I metabolism	↑
CCAAT- enhancer binding protein delta-like protein mRNA	BJ710640	1E-105	Inflammatory mediator regulation	↑
Pituitary adenylate cyclase activation polypeptide receptor (PACAP1A gene)	BJ736738	2E-53	Regulatory neuropeptide	↑
Microtubule associated protein RP/EB family (EB1)	BJ740475	4E-74	Microtubule dynamics, cytoskeleton stability regulation, chromosome stability	↑
Predicted: similar kelch-like protein 20	BJ748092	8E-26	Function unknown, associated with actin tails	↑
KFH-R protein (kfh-r mRNA)	AB001604	0	Putative visual pigment	↓
C-type natriuretic peptide-3 (cnp-3, mRNA)(paracrine, autocrine)	AB091698	0	Cardiovascular and osmoregulatory hormones	↓
B-type natriuretic peptide precursor (cnp, mRNA)(circulating)	AB099700	0	Cardiovascular precursor	↓
Heme-oxygenase 1, mRNA	AB163431	6E-154	Stress response (oxidative stress, hypoxia, heavy metals, cytokines)	↓
L-SF precursor (1-sf), mRNA	AU179140 AU 180097	0	Hypothalamic-pituitary-gonadal (HPG) axis	↓

Choriogenin Hminor gene	AU179899 AU 180137	0	Hypothalamic-pituitary-gonadal (HPG) axis	↓
Cytochrome P450 monooxygenase CYP2K1	BJ512702	3E-120	Phase I metabolism (estrogenicity)	↓
Putative: THO complex 3 (D.rerio)	BJ746432	0.0	RNA DNA metabolism, transcription	↓
Gene Name	Accession#	E-value	Biological Function	Regulation females
Glycoprotein (rh50)	AB036511	0.00	Immune response	↑
cytochrome P450 1A (CYP1A) mRNA	AY297923	0.00	Phase I metabolism	↑
Microtubule associated protein RP/EB family (EB1)	BJ740475	4E-74	Microtubule dynamics, cytoskeleton stability regulation, chromosome stability	↑
Predicted: similar kelch-like protein 20	BJ748092	8E-26	Function unknown, associated with actin tails	↑
hoxA2a	AB027025	9E-54	Morphogenesis, developmental patterns	↓
MF-Hox9-1	AB027039	2E-55	Morphogenesis, developmental patterns	↓
DMY protein (dmy), mRNA	AB071534	0	Sex differentiation	↓
B-type natriuretic peptide precursor (cnp, mRNA)(circulating)	AB099700	0	Cardiovascular precursor	↓
Heme-oxygenase 1, mRNA	AB163431	6E-154	oxidative stress, hypoxia, heavy metals, cytokines	↓

ZPC domain containing protein 5	AF128813 AU241961	0	Zona pellicula protein	↑
Pax6 gene	AJ000938	0	Eye development, sensory organs	↓
Hox gene cluster	AU169510	7E-66	Morphogenesis	↓
Alpha tubulin	AU176604	2E-111	Microtubule dynamics, cytoskeleton stability	↓
L-SF precursor (1-sf), mRNA	AU179828	0	hypothalamic-pituitary-gonadal (HPG) axis	↓
Choriogenin Hminor gene	AU179899 AU 180137	0	hypothalamic-pituitary-gonadal (HPG) axis	↓
Progesterin receptor membrane component	AV668306	0	Down-expression during late vitellogenesis	↓
Cytochrome P450 26A1 (CYP26a1)	AV668461	0	Phase I metabolism	↓
Sp3 transcription factor	BJ009046	3E-105	bifunctional transcription factor, either stimulates or represses the transcription of numerous genes	↓
14-alpha demethylase (CYP51)	BJ016320	3E-149	sterol biosynthetic pathways, cholesterol biosynthesis, gonads	↓
Sox17 (SRY (sex determining region Y)-box 17), mRNA	BJ493205	6E-152	Transcription factor, gut endoderm development	↓
Interphotoreceptor retinoid-binding protein (IRBP2) gene, Exon 1	BJ714260	7E-27	Visual cycle, retinoid transport	↓

Putative: Interleukin enhancer binding factor (ILF-2), mRNA	BJ717119	0.0	Immune system	↓
6-Phosphogluconate dehydrogenase (pgd gene)	BJ720150	2E-53	Red blood cell metabolism	↓
Type I keratin isoform	BJ728005	1E-99	intracytoplasmatic cytoskeleton	↓
MHC Class I region	BJ728138	9E-77		↓
Chromatin assembly factor 1	BJ728607	3E-116	immune system, autoimmunity, reproductive success	↓
Beta-3-glucuronyltransferase (b3gat1 gene)	BJ729999	1E-134	Immune response	↓
HSP putative Zinc finger transcription factor (DMRT1a)	BJ730394	9E-12	Male determining gene	↓
membrane guanylyl cyclase OIGC1	BJ738711	8E-60	Downregulation possible because of high salt or lead	↓
Putative: THO complex 3 (D.rerio)	BJ746432	0.0	RNA DNA metabolism, transcription	↓

↑ and ↓ indicate upregulation and downregulation of genes, respectively. BLAST similarity determined by the lowest expectation (E)-values on NR Database using accession numbers as identifiers.

CONCLUSIONS, DISCUSSIONS AND SUGGESTIONS

A custom medaka microarray chip was developed and we tested its ability to discriminate between six prototypical EDCs evaluated at biological effective concentrations in a robust experimental design. At least three elements of the microarray experiment design of this study eliminated various intrinsic sources of variation. First, all of the exposed fish were at the same developmental stage and it is likely that they shared similar physiological states. Second, inbred Qurt strain medaka colonies had been maintained for over 20 years prior to this study which reduces the effect of genetic polymorphisms on the observed alterations in gene expression. In addition, the genders of all of the fish used were determined prior to exposure to EDCs allowing a reliable testing for gender-specific alterations in gene expression response. By pooling the RNA samples of fish within each gender, we characterized the mean gene expression response of each gender to EDC exposure.

We tested and validated the expression profile for vinclozolin which is remarkably similar to the prototypical anti-androgen flutamide, indicating that our system can prospectively identify EDCs with characteristic modes of action. We also tested and validated the expression profile for simple EDC mixtures, i.e., exposure to mixtures of either 17- β -estradiol + faslodexTM or 11-ketotestosterone + flutamide; or thyroid hormone (T3) + amiodarone at per-determined biologically effective concentrations, indicating that effects of simple mixtures are different from those exerted by exposure to single EDCs. For example: exposure to 11-KT + FLU mixture has an anti-estrogenic outcome, exposure to T3 + AMI mixture has anti-androgenic outcome, and exposure to E2 + FAS mixture has estrogenic outcome. These results further suggest complex mechanisms that may be associated with EDCs when present in aquatic environments. In addition, to overcome the cost of microarray analysis, we have designed and built a new 1620 medaka microarray gene chip at UC Davis. This new 1620 medaka microarray gene chip has been tested and has similarly proven sensitivity in discriminating the six prototypic EDCs compared to the 12,000 custom chip initially built by CombiMatrix Corp. This new gene chip will significantly reduce our cost and time for future EDC microarray analysis.

Our results indicate: 1) L-groundwater has estrogenic effects on male medaka larvae at 6hr but anti-thyroidogenic effects at the end of 96hr exposure and anti-estrogenic effects on female medaka larvae at 6hr but anti-thyroidogenic effects at the end of 96hr exposure, and 2) 50%-2^o-wastewater has thyroidogenic effects on male and anti-androgenic effects on female medaka exposure to the 50%-2^o-wastewater for 6 hr. The transition from estrogenic to anti-thyroidogenic effects in female and anti-estrogenic to anti-thyroidogenic effects in male may indicate gender-specific effects of medaka to the L-groundwater at 6hr. However, prolonged exposure (96hr) further illustrated that some genes may persist while other genes subside indicating a complex cascade of up- and down-regulation of genes in female and male medaka during the duration of exposure. In addition to the endocrine disrupting effects, the occurrence of significantly higher number of down-regulated genes compared to up-regulated genes further suggest the long-term effects of L-groundwater and 50%-2^o-wastewater on the growth of medaka larvae. Our preliminary results also indicate that the L-groundwater and 50%-2^o-wastewater consist of complex contaminant mixtures with endocrine disruptor effects on 1-wk-old medaka larvae.

Endocrine disrupting chemicals (EDCs) have been a major issue in environmental science and policy for more than ten years. Although there has been considerable debate over true concentrations, sources, identity, and effects of potential EDCs, overwhelming evidence indicates that EDCs can critically compromise reproductive and developmental processes in both humans and wildlife. Unfortunately, even with these profound impacts, there are no specific restrictions currently in place to regulate EDC discharges due, in part, to the lack of effective and standardized tests to accurately detect the presence of such chemicals in the environment. As EDCs primarily target gene transcriptional activity, microarray-based alterations in gene expression can be used as markers of exposure. EDCs can interfere with development of the brain and reproductive systems in fish, humans and wildlife by interfering with neural and endocrine signals and cellular function of the hypothalamus-pituitary-gonadal and liver (HPG/L) axis. Identifying the molecular and functional responses of the HPG/L to EDCs is critical to understanding the mechanisms involved in the adverse health effects caused by exposures to EDCs.

Our medaka microarray chip will provide a powerful screening tool that can simultaneously yield information not only on the parent chemical's endocrine disruptor activity but also on the presence of any active metabolites. This basic screening system also has profound benefits to future characterization of organ-specific responses to EDCs as well as characterizing differences in EDC gene-expression profiles associated with exposure of different life stages of fish. Furthermore, assessment of complex contaminant mixtures of EDCs can be accomplished using this novel method that could greatly facilitate the identification of diverse aquatic environmental EDCs that may have synergistic as well as individual effects.

Our future goal is to develop the medaka microarray chip system into a comprehensive tool for screening and identifying previously uncharacterized EDCs. Additional funding from GAMA, SWRCB will provide an opportunity to investigate the biological effective concentrations (BEC) of several pharmaceutical chemicals previously detected by LLNL in wastewater or groundwater. Using 7 dph medaka larvae, we propose to perform 1) time-course and, 2) dose-response experiments with the prototypical EDCs at LC10 and LC25 of the BEC using individual or combined mixtures of the pharmaceutical chemicals. The medaka genome annotation which we have already developed will identify the function of affected genes and their mode of action. DNA microarrays provide a snapshot of transcriptional activity in tissue samples and demonstrate which genes are actively expressed within the cells at one point in time. Using this technology will allow relatively rapid and straightforward assays to identify EDCs by observing the changes in gene expression patterns in response to exposure. Further development of this technique with the proposed studies outlined above will build the essential groundwork needed to support cross species comparisons on shared mechanisms and vulnerabilities to EDC compounds.

ACKNOWLEDGEMENTS

We would like to thank John Borkovich, Jan Stepek, Lisa Babcock, and James Giannopoulos of SWRCB for their advice, guidance, and patient, Bradley Esser and Alan Grayson of LLNL for their assistance in the collection of wastewater and groundwater, Ching Teh for medaka culturing and separating thousands of females and male medaka eggs for exposure study, and Janice Pfeiff at ArrayCore facility at UCD for microarray analysis. Funding was provided by Groundwater Ambient Monitoring and Assessment Program, State Water Resources Control Board, Sacramento, California 95814 (Contract# 04-111-250-0).

ACRONYMS AND ABBREVIATION

11-KT = 11-Ketotestosterone

AMI = Amiodarone

BEC = Biologically effective concentration

BLAST = Basic Local Alignment Search Tool

Dph = Day post-hatched

EDC = Endocrine disrupting chemical

E2 = 17- β -Estradiol

FAS = FaslodexTM (ICI 182,780)

FCKDA = Functional Classification- Kernel Density Approach

FLU = Flutamide

GAMA = Groundwater Ambient Monitoring and Assessment

HPG/L = hypothalamus-pituitary-gonadal and liver axis.

LLNL = Lawrence Livermore National Laboratory

SWRCB = State Water Board Resources Control Board

T3 = Thyroid hormone

APPENDICES

- I. A Functional Approach to Predict Endocrine Disruption from Gene Expression Profiles Using a Medaka Fish Model
- II. Using a Sensitive Japanese Medaka (*Oryzias latipes*) Fish Model for Endocrine Disruptors Screening
- III. A Computational Approach to Predict Endocrine Disrupting Activity in Water Using a Medaka Fish Model
- IV. Global Gene Expression Profiling of Androgen Disruption in Qurt Strain Medaka

A Functional Approach to Predict Endocrine Disruption from Gene Expression Profiles using a Medaka Fish Model

Ping-Shi Wu¹, Abimael León², Wei-Min Huang¹, and Swee J. Teh²

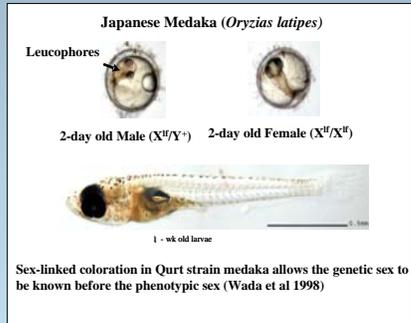
¹Department of Mathematics, Lehigh University

²Department of Anatomy, Physiology, and Cell Biology, School of Veterinary Medicine

INTRODUCTION

Global gene expression profiles in a Qurt Strain Medaka fish model were successfully used to identify patterns predictive of Androgen Disrupting Chemicals (ADC) with agonist (i.e. 11-ketotestosterone; 11-KT) and antagonist (i.e. Flutamide; FLU) activity on a custom oligonucleotide array (León et al. 2007). The gene-profiling Medaka model serves as a rapid, sensitive, and biologically-integrated screening assay for the identification and classification of environmental ADC activities.

In this study, we included six Endocrine Disrupting Chemicals (EDC), 11-KT and FLU from the previous study along with 3, 3', 5-Triiodo-L-thyronine (T3), Amiodarone, 17-β Estradiol (E2) and ICI 182,780 (Faslodec™, AstraZeneca, UK). These six prototypic EDCs have been extensively studied and are known to have unique agonist or antagonist activity in three hormone receptor-mediated systems.



OBJECTIVES

- Build up the gene-expression pattern for identification and classification of EDC's, using Medaka Fish model.
- Construct statistical model for predicting environmental EDC activities.

METHODS

- Oligonucleotide array design: Please see León et al. (2007) for details.

- Functional Embedding (FEM): Muller and Wu (2007) proposed a novel approach for embedding high dimensional data into functional space. Gene-expression profiles are re-expressed in functional form via FEM. The effect of ordination after FEM can be easily seen in Figure.1. Here gene profiling consists of 281 genes yielding most significant difference among EDC groups.

- Functional Discrimination – Kernel Density Estimation Approach–(FDKDE).

$$X_{ij} \stackrel{iid}{\sim} N_{\mu_j} \text{ with } E(X_{ij}) = \mu_j \text{ and } \text{cov}(X_{ij}(s), X_{ij}(t)) = G(s, t)$$

$$(A_{ij}f)(t) = \int f(s)G(s, t)ds, \text{ defined for } f \in L^2$$

$$(A_{ij}\phi_k)(t) = \lambda_k \phi_k(t),$$

$$\int \phi_j(t)\phi_k(t)dt = 1 \text{ for } j = k \text{ and } = 0 \text{ for } j \neq k$$

$$G(s, t) = \sum_{k=1}^{\infty} \lambda_k \phi_k(s)\phi_k(t)$$

$$X_{ij}(t) = \mu_j(t) + \sum_{k=1}^{\infty} \xi_{ik} \phi_k(t),$$

$$\xi_{ik} = \int (X_{ij}(t) - \mu_j(t))\phi_k(t)dt$$

$$Y_{ij} = X_{ij}(t_{ij}) + \epsilon_{ij} = \mu_j(t_{ij}) + \sum_{k=1}^{\infty} \xi_{ik}\phi_k(t_{ij}) + \epsilon_{ij}, \quad -1 \leq t_{ij} \leq 1,$$

- Truncated FPC score $(\xi_{i1}, \xi_{i2}, \dots, \xi_{iM})$ then provide the new configuration of the gene expression profiles in feature space.

$$P(C|\hat{C}) = k | \hat{C} | E[1(C|\hat{C}) = k | \hat{C} = \hat{C}]] = \frac{\sum_{C:|C|=k} K(|\hat{C} - C|)}{\sum_{C:|C| \leq M} K(|\hat{C} - C|)}$$

$$C^*(\hat{C}) = \text{Argmax}_{C:|C| \leq M} P(C|\hat{C}) = i | \hat{C} |$$

- Mark subject as "Doubt" if $\text{Max}_{C:|C| \leq M} P(C|\hat{C}) = i | \hat{C} | < \frac{1}{5}$

- Leave-One-Out-Cross-Validation for the evaluation of FDKDE.

$$G(x, t) = \sum_{k=1}^M \lambda_k^{-1} \phi_k^{-(-)}(x) \phi_k^{-(-)}(t)$$

$$Y_{ij} = \mu_j^{-(-)}(t_{ij}) + \sum_{k=1}^M \xi_{ik}^{-(-)} \phi_k^{-(-)}(t_{ij}) + \epsilon_{ij}, \quad -1 \leq t_{ij} \leq 1,$$

$$\xi_{ik}^{-(-)} = \int (Y_{ij}(t) - \mu_j^{-(-)}(t)) \phi_k^{-(-)}(t) dt$$

$$P^{(-)}(C|\hat{C}) = k | \hat{C} | E[1(C|\hat{C}) = k | \hat{C} = \hat{C}]) = \frac{\sum_{C:|C|=k} K(|\hat{C}^{-(-)} - C|)}{\sum_{C:|C| \leq M} K(|\hat{C}^{-(-)} - C|)}$$

$$C^*(\hat{C}) = \text{Argmax}_{C:|C| \leq M} P^{(-)}(C|\hat{C}) = k | \hat{C} |$$

$$LOOCV = \frac{\sum_{i=1}^n 1(C^*(\hat{C}_i^{-(-)}) \neq C_i(\hat{C}_i))}{n}$$

, where the superscript (-) indicates that the i-th observation is left out in the constructions of the functional covariance surface, the mean function, the eigenfunctions, and the functional principal component score. Moreover, $C_i(\hat{C}_i)$ is the true group membership, and the dimension of truncation M is determined by using Functional Variance Explanation criterion (FVE, hereafter).

RESULTS

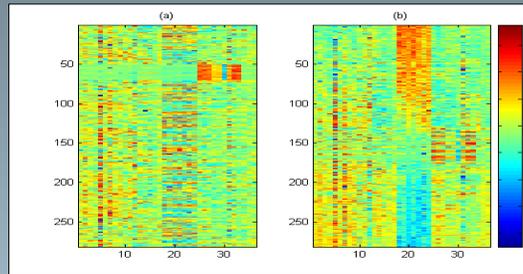


Figure.1 (a) Before FEM; (b) After FEM (281 genes are used for coordinate configuration)

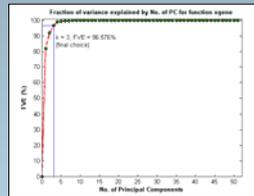


Figure.2 FVE scree plot

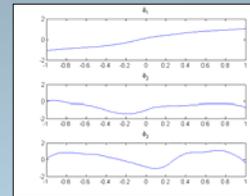


Figure.3 Three estimated eigenfunctions

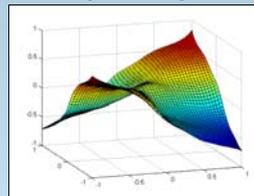


Figure.4 Covariance Surface

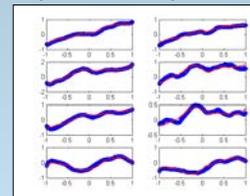


Figure.5 Eight estimated gene profiles

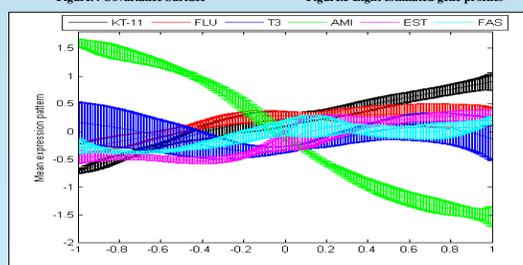


Figure.6 Mean profiles with error bars indicating the within group variation.

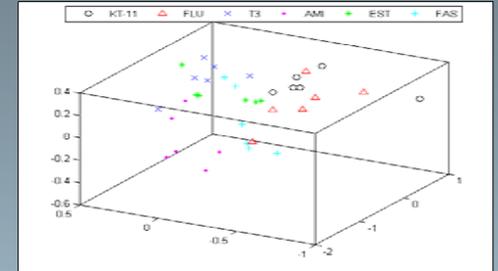


Figure.7 Scatter plot for FPCS

The LOOCV misclassification rate based on FDKDE with 3 Functional principle component scores (FPCS) is 14/35=0.4 with one subject classified as Doubt. The Choice of number of FPCS is by Functional Variance Explained ratio (FVE) illustrated in figure.2.

DISCUSSION

The use of Medaka Fish model for identification and classification of EDC activities based on gene expression profiles sheds a new light on aquatic-toxicological study. Our findings show that gene expression profiling with medaka has significant potential for environmental monitoring. It is a rapid, sensitive, and biologically-integrated screening assay. In the future, a follow-up study can be developed in the following perspectives:

- Combining information from gene pathways can help with annealing the assay.
- A statistical model with variable trimming feature can help with simplifying the screening assay without loss of sensitivity.
- To utilize the medaka microarray and computational approach to test for the activity of endogenous and synthetic hormones in ambient water.
- Evaluate how EDC gene patterns change over time and dose.

REFERENCES

- León, A., Wu, P.S., Hall, L., Johnson, M., Teh, S. (2007). "Global Gene Expression Profiling of Androgen Disruption in Qurt Strain Medaka", *Environmental Science & Technology*, accepted for publication.
- Muller, H.G. and Wu, P.S. (2007) "Functional Embedding for High-dimensional Data" submitted.
- Hall, P., Poskitt, D.S., Presnell, B. (2001) "A Functional Data-Analytic Approach to Signal Discrimination", *JASA* 43, no. 1.
- Ripley, B.D. (1996) *Pattern Recognition and Neural Networks*
- Wada, H., A. Shimada, S. Fukamachi, K. Naruse, A. Shima. Sex-linked inheritance of the *lf* locus in the medaka fish (*Oryzias latipes*). *Zoological Science* 15, 123-126, (1998).

ACKNOWLEDGEMENTS

This study was funded in part by the US Environmental Protection Agency (US EPA) STAR Research Grant RD R83129901-0 and California State Water Resources Control Board (Contract# 04-111-250-0). Additional funding was provided by the Ecotoxicology Lead Campus Program (ELCP) Graduate Fellowship 2003-2005 at the University of California, Davis.

Using a Sensitive Japanese Medaka (*Oryzias latipes*) Fish Model for Endocrine Disruptors Screening

Swee J. Teh, Abimael León-Cardona, Ping-Shi Wu, Linda C. Hall, and Michael Johnson

U.S EPA, ORD, Computational Toxicology Research Program

Project ID: Swee Teh (IIC)

Science Question

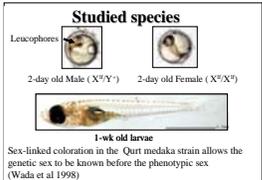
The endocrine system consists of a highly integrated set of glands and widely distributed neuroendocrine cells in exocrine glands whose primary function is the control of homeostasis, including growth, reproduction and fertility, which are essential for propagation of the species. Therefore, the disruption of the endocrine system by any anthropogenic chemical may eventually have profound effects on both individuals and the population of a species. Endocrine disrupting chemicals (EDCs) are under intense scientific scrutiny because of the increasing number of environmental contaminants linked to disruption of one or more components of the endocrine system. Perhaps the greatest challenge to scientists studying EDCs are that (1) many of these chemicals can exert their effects at very low doses, and (2) harmful effects may not manifest for months or years after exposure. Thus, it is possible that many of the pesticides and industrial products in widespread use today may have harmful long-term effects that interfere with some aspect of the endocrine system. Although there is no single approach that can simultaneously identify EDCs and characterize their toxicity, the development of a rapid, sensitive, biologically-integrated screening assay that can identify and classify EDCs by category of endocrine activity is of utmost importance in beginning to understand the scope of the problem posed by these chemicals.

Research Goals

To develop and validate a high-throughput EDC screening assay utilizing the microarray gene chip, four objectives are proposed:

1. Develop a microarray chip designed to identify endocrine disrupting chemicals.
2. Identify gene expression profiles associated with all categories of endocrine disrupting activity.
3. Conduct a statistical analysis of gene expression profiles to develop response criteria that identify patterns predictive of endocrine disruption, and
4. Validate the microarray chip using a set of chemicals selected to represent both positive and negative controls, as well as chemicals with previously-undefined endocrine activity.

- ### Six Prototypic EDCs
1. 17- β Estradiol (E2) is one of the most studied natural estrogens
 2. ICI 162,780 (FaslodexTM, AstraZeneca, UK) is a novel, steroidal estrogen antagonist that was specifically designed to be devoid of estrogen agonist activity (Howell et al., 2000).
 3. 11-Ketotestosterone is the primary endogenous fish androgen and a non-aromatizable androgen (Borg, 1994). Since it is not converted to estradiol by the P450 aromatase, the biological effects observed in exposed Qurt medaka fish should be those expected from a pure androgenic agent
 4. Flutamide is a non-steroidal compound widely used in endocrinology research. It has been thoroughly established that this therapeutic agent can function as an antiandrogen in mammals (Boyley et al., 2002)
 5. 3, 3', 5-Triiodo-L-thyronine (T3) is one of two iodinated hormones secreted by the thyroid gland. T3 is essential for growth, differentiation, and reproductive system development (Dace et al., 2000; Claski et al., 2003). T3 exerts its biological effects through a receptor-mediated mechanism. Furthermore, T3 modulates its own function by regulating the thyroid receptor (TR) level (Dace et al., 2000)
 6. Amiloradone is a diiodinated benzofuran derivative that functions as a potent antiarrhythmic drug (Shahara and Drozda, 1999; Latham et al., 1987)



Methods/Approach

Experimental Design I

Determining biological effective concentrations (BECs) of the six prototypic EDCs

Type of experiment: 96-h static renewal method

Sample size: 50 larvae per gender per four replicates

Time of Sampling: Three months after exposure

Endpoints:

1. Health indices (body length, body weight, and condition factor)
2. Histopathology
3. Occurrence of intersex

Experimental Design II

Developing Custom Medaka L2K Oligonucleotide Array

Features of array:

1. EDC-affected genes (n=131)
2. Medaka cDNAs (n=1075)
3. Medaka genome project (n=890)
4. Unigene database (n=8154)

9379 unique and ~3000 control probes

Experimental Design III

Identifying unique gene response patterns of the six prototypic EDCs

Type of experiment: Static 6-h individually exposure to

1. Ethanol control
2. 10 μ g/L 17- β Estradiol
3. 10 μ g/L ICI 162,780 (FaslodexTM)
4. 34 μ g/L Amiloradone
5. 100 μ g/L 11-ketotestosterone
6. 500 μ g/L 3, 3', 5-Triiodo-L-thyronine
7. 1000 μ g/L Flutamide

Sample size = 200 larvae per gender per three replicates

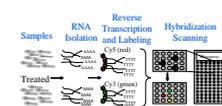
Total RNA isolation

Microarray analysis

Analysis of microarray images

Loess normalization and functional embedding for high dimension data analyses

Latent ordering of genes for patterning analysis



Results

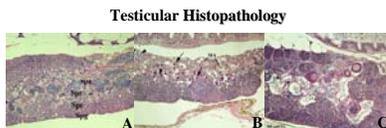


Figure 1. (A) Normal testis of control male showing several reproductive stages in spermatogenesis. (Spg = spermatogonia; Ssp = spermatocytes; Spt = spermatids; Sps = spermatozoa). (B) Ovaritis in a male exposed to 320 μ g/L of 3, 3', 5-Triiodo-L-thyronine. (C) Ovaritis in a male exposed to 500 μ g/L of 3, 3', 5-Triiodo-L-thyronine. Notes: Both severe germ cell necrosis (arrowheads) and ova cell necrosis (arrows) were observed in this intersex fish (MA = macrophage aggregations), and (C) Ovaritis in a male exposed to 500 μ g/L of 3, 3', 5-Triiodo-L-thyronine



Figure 2. (A) Normal thyroid gland in control male showing single-layered of low cuboidal follicular cells. (B) Follicular cell atrophy (arrow) in a male exposed to 50 μ g/L of 3, 3', 5-Triiodo-L-thyronine, and (C) Thyroid follicular cells hyperplasia and hypertrophy were observed in a male exposed to 500 μ g/L of 3, 3', 5-Triiodo-L-thyronine. Arrows point to necrotic follicular cells.

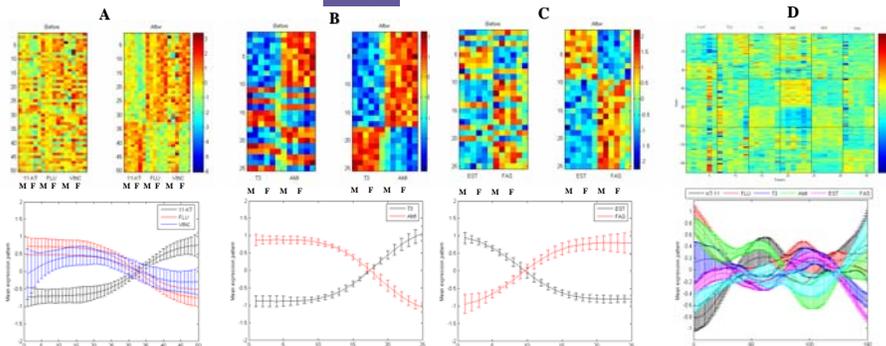


Figure 3 A-D. Heat map images of the six prototypic EDCs before and after functional embedding technique (Muller and Wu, in preparation). To perform the clustering/classification analysis, the latent ordering of genes with the smallest ρ -values based on t -test was used to express the subject profiles (columns) in functional form. The Heat map images were then transformed to line curve using the scattered-smoothed mean curves with standard error bar. 11-ketotestosterone = 11-KT, Flutamide = FLU, Vinclozolin (VIN), 3, 3', 5-Triiodo-L-thyronine = T3, Amiloradone = AMI, 17- β Estradiol = EST, and ICI 162,780 (FaslodexTM) = FAS.

Gene Name	Biological Function
<i>Arhgap10</i>	GTPase-activating protein
<i>Arhgap11</i>	GTPase-activating protein
<i>Arhgap12</i>	GTPase-activating protein
<i>Arhgap13</i>	GTPase-activating protein
<i>Arhgap14</i>	GTPase-activating protein
<i>Arhgap15</i>	GTPase-activating protein
<i>Arhgap16</i>	GTPase-activating protein
<i>Arhgap17</i>	GTPase-activating protein
<i>Arhgap18</i>	GTPase-activating protein
<i>Arhgap19</i>	GTPase-activating protein
<i>Arhgap20</i>	GTPase-activating protein
<i>Arhgap21</i>	GTPase-activating protein
<i>Arhgap22</i>	GTPase-activating protein
<i>Arhgap23</i>	GTPase-activating protein
<i>Arhgap24</i>	GTPase-activating protein
<i>Arhgap25</i>	GTPase-activating protein
<i>Arhgap26</i>	GTPase-activating protein
<i>Arhgap27</i>	GTPase-activating protein
<i>Arhgap28</i>	GTPase-activating protein
<i>Arhgap29</i>	GTPase-activating protein
<i>Arhgap30</i>	GTPase-activating protein
<i>Arhgap31</i>	GTPase-activating protein
<i>Arhgap32</i>	GTPase-activating protein
<i>Arhgap33</i>	GTPase-activating protein
<i>Arhgap34</i>	GTPase-activating protein
<i>Arhgap35</i>	GTPase-activating protein
<i>Arhgap36</i>	GTPase-activating protein
<i>Arhgap37</i>	GTPase-activating protein
<i>Arhgap38</i>	GTPase-activating protein
<i>Arhgap39</i>	GTPase-activating protein
<i>Arhgap40</i>	GTPase-activating protein
<i>Arhgap41</i>	GTPase-activating protein
<i>Arhgap42</i>	GTPase-activating protein
<i>Arhgap43</i>	GTPase-activating protein
<i>Arhgap44</i>	GTPase-activating protein
<i>Arhgap45</i>	GTPase-activating protein
<i>Arhgap46</i>	GTPase-activating protein
<i>Arhgap47</i>	GTPase-activating protein
<i>Arhgap48</i>	GTPase-activating protein
<i>Arhgap49</i>	GTPase-activating protein
<i>Arhgap50</i>	GTPase-activating protein
<i>Arhgap51</i>	GTPase-activating protein
<i>Arhgap52</i>	GTPase-activating protein
<i>Arhgap53</i>	GTPase-activating protein
<i>Arhgap54</i>	GTPase-activating protein
<i>Arhgap55</i>	GTPase-activating protein
<i>Arhgap56</i>	GTPase-activating protein
<i>Arhgap57</i>	GTPase-activating protein
<i>Arhgap58</i>	GTPase-activating protein
<i>Arhgap59</i>	GTPase-activating protein
<i>Arhgap60</i>	GTPase-activating protein
<i>Arhgap61</i>	GTPase-activating protein
<i>Arhgap62</i>	GTPase-activating protein
<i>Arhgap63</i>	GTPase-activating protein
<i>Arhgap64</i>	GTPase-activating protein
<i>Arhgap65</i>	GTPase-activating protein
<i>Arhgap66</i>	GTPase-activating protein
<i>Arhgap67</i>	GTPase-activating protein
<i>Arhgap68</i>	GTPase-activating protein
<i>Arhgap69</i>	GTPase-activating protein
<i>Arhgap70</i>	GTPase-activating protein
<i>Arhgap71</i>	GTPase-activating protein
<i>Arhgap72</i>	GTPase-activating protein
<i>Arhgap73</i>	GTPase-activating protein
<i>Arhgap74</i>	GTPase-activating protein
<i>Arhgap75</i>	GTPase-activating protein
<i>Arhgap76</i>	GTPase-activating protein
<i>Arhgap77</i>	GTPase-activating protein
<i>Arhgap78</i>	GTPase-activating protein
<i>Arhgap79</i>	GTPase-activating protein
<i>Arhgap80</i>	GTPase-activating protein
<i>Arhgap81</i>	GTPase-activating protein
<i>Arhgap82</i>	GTPase-activating protein
<i>Arhgap83</i>	GTPase-activating protein
<i>Arhgap84</i>	GTPase-activating protein
<i>Arhgap85</i>	GTPase-activating protein
<i>Arhgap86</i>	GTPase-activating protein
<i>Arhgap87</i>	GTPase-activating protein
<i>Arhgap88</i>	GTPase-activating protein
<i>Arhgap89</i>	GTPase-activating protein
<i>Arhgap90</i>	GTPase-activating protein
<i>Arhgap91</i>	GTPase-activating protein
<i>Arhgap92</i>	GTPase-activating protein
<i>Arhgap93</i>	GTPase-activating protein
<i>Arhgap94</i>	GTPase-activating protein
<i>Arhgap95</i>	GTPase-activating protein
<i>Arhgap96</i>	GTPase-activating protein
<i>Arhgap97</i>	GTPase-activating protein
<i>Arhgap98</i>	GTPase-activating protein
<i>Arhgap99</i>	GTPase-activating protein
<i>Arhgap100</i>	GTPase-activating protein

Table 1. Selected Functional Profiles of Differentially Expressed Genes in Males and Females Exposed to Flutamide (FLA), 11-ketotestosterone (11-KT), and Vinclozolin (VIN). Red and green indicate upregulation and downregulation of genes, respectively. Blasts similarity determined by the best hit (lowest E-values) on NR Database using accession numbers and clone IDs as identifiers. M-male and F-female.

Discussions/Conclusions

1. We have determined the BECs of the six prototypic EDCs.
2. We have developed a custom medaka gene chip that is capable of identifying and differentiating gene expression patterns of six categories of EDCs (Figure 3).
3. Figure 3A shows vinclozolin has a similar pattern as flutamide therefore confirm the anti-androgenic activity of vinclozolin and the predictive ability of the medaka gene chip.
4. The predictive ability of the microarray gene chip with additional EDC compounds and with simple EDC mixtures is in progress.

Impact and Outcomes

The medaka microarray chip will provide a powerful screening tool that can simultaneously yield information on not only the parent chemical's ED activity but also any active metabolites. The basic screening system is also directly amenable to the future characterization of organ-specific responses to EDCs, as well as to characterizing differences in EDC gene-expression profiles associated with exposure of different life stages. Furthermore, assessment of complex contaminant mixtures of EDCs will be possible with this novel method and could greatly facilitate future aquatic environmental EDC identification and evaluation.

Future Directions

As the medaka genome is annotated, we can identify the function of affected genes and gene pathways to further our understanding of the underlying causes of EDC-induced toxicity (see Table 1). Our future goals

References

Borg, B. (1994) *Comparative Biochemistry and Physiology*. 109C (3): 219-245.

Bayley, M., M. Jungs, and E. Baatrup. (2002) *Aquatic Toxicology*. 56: 227-239.

Dace, A., L. et al. (2000) *Proceedings of the National Academy of Sciences, USA*. 97(16): 8968-8970.

Howell, A., et al. (2000) *Cancer* 89 (4): 817-825.

Latham, K. R. et al. (1987) *Journal of the American College of Cardiology* 9(4): 872-876.

Shahara, S., and Drozda, V. (1999) *Journal of Cardiovascular Pharmacology* 34(2): 261-267.

Wada, H. A., et al. (1998) *Zoological Science*. 15, 123-126.

A Computational Approach to Predict Endocrine Disrupting Activity in Water Using a Medaka Fish Model

TOXICOGENOMICS

Swee J. Teh, Abimael León, Ping-Shi Wu, and Linda C. Hall

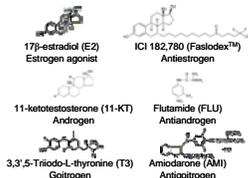
University of California-Davis/VM:APC Aquatic Toxicology Program • Lehigh University/Department of Mathematics • Lawrence Livermore National Laboratory/Environmental Restoration Division, US

Background

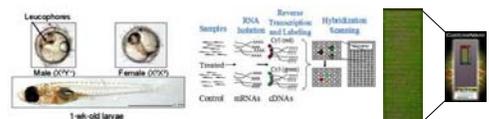
Endocrine disrupting chemicals (EDCs) occur within a suite of surface water, groundwater, and wastewater contaminants that could be compromising human and environmental health. Aquatic EDCs from natural and anthropogenic origins including agricultural runoff and industrial effluents significantly reduce the quality of water resources. In many cases, potential EDCs exist in the environment as mixtures that could act synergistically (Falconer et al. 2006).

A novel computational approach was utilized to characterize the patterns of genomic response of Qurt strain Japanese medaka (*Oryzias latipes*) larvae after short term exposures to biologically effective concentrations of three receptor agonist/antagonist EDCs and their mixtures.

Three pairs of prototypic agonist/antagonist EDCs



Methodology



Custom oligonucleotide array = ~9.4K unique probes
Sample size = 200 larvae per gender per three replicates
Static 6h exposure of males and females
RNA isolation and purification
Microarray hybridization and scanning

Data/Analysis

Identification of unique gene response patterns of EDCs

Six Prototypical EDCs: E2, ICI, 11-KT, FLU, T3 and AMI

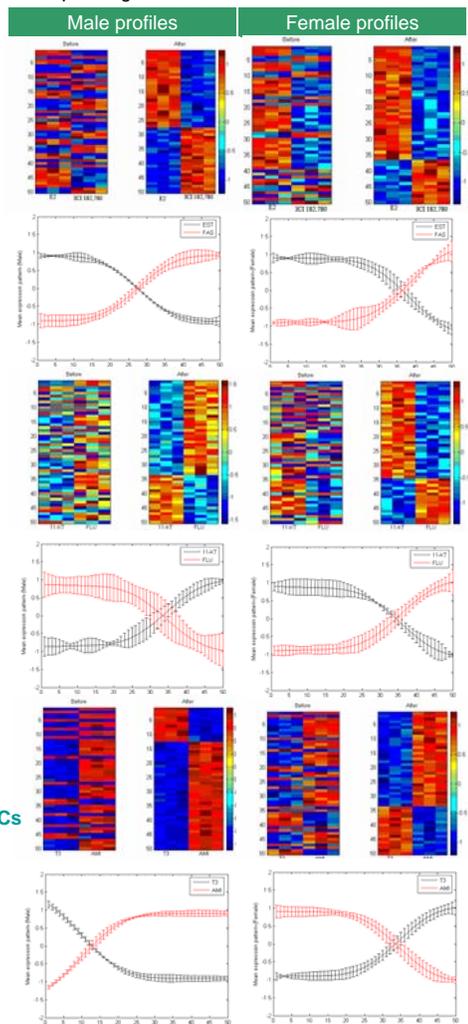
Three EDC mixtures: E2+ICI, 11-KT+FLU, and T3+AMI

Clustering/classification analysis:

Microarray feature extraction with ScanAlyze™ software
Loess normalization and functional embedding for high dimension data analyses (Muller and Wu 2007)
Latent ordering of genes for pattern analysis and identification

Annotation of unique EDC probes in Qurt medaka

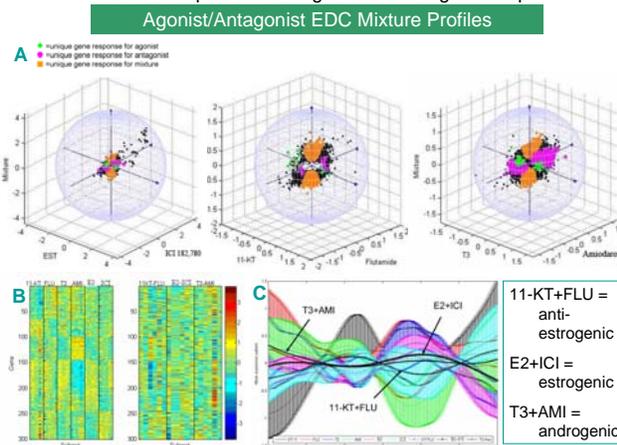
Figure 1. Gender-specific expression profiles of EDCs: Heat map images before and after application of a functional embedding technique and corresponding scattered smoothed mean curves.



Results

Figure 2. Synergistic effects of EDC mixtures:

(A) Vector Analysis results for genomic responses to prototypic EDCs single and paired (B) Heat map and (C) smoothed curves comparison of single and mixed genomic profiles.



Potential Biomarkers for prototypic EDCs in medaka

Gene name and rank	E-value	11KT	FLU
Regulation in ♂			
cytP450 (18)	0	↑	↓
Regulation in ♀			
putative:ferritin(7)	0	↑	↓
putative:ER-binding cyclophilin(22)	0	↑	↓
putative:GSH-T gene(23)	0	↑	↓
cyclin B1(36)	0	↓	↑
Regulation in ♂			
vtg I (14)	0	↑	↓
choriogenin H (16)	0	↑	↓
vtg I precursor (30)	0	↑	↓
putative: PR gene (33)	0	↓	↑
HSC70 gene (37)	0	↓	↑
transcription factor DMRT1a (39)	5E-17	↓	↑
GnRH-R1 (40)	5E-49	↓	↑
GnRH-R3 (41)	5E-16	↓	↑
Regulation in ♀			
AR-α (37)	0	↓	↑
Regulation in ♂			
tyrosine PTPs (17)	0	↓	↑
vtg II (28)	0	↓	↑
Regulation in ♀			
choriogenin H minor (4)	0	↓	↑
tyrosinase precursor (10)	2E-10	↓	↑
dmrt1Y (32)	0	↓	↑
ZPC gene (40)	0	↓	↑
mhc class II antigen gene (50)	0	↓	↑

Discussion and Conclusion

- Diagnostic probes identified in this study are a powerful means to discriminate between different types of endocrine disrupting activity at the transcription level.
- Annotation of diagnostic probes revealed alterations of potential biomarkers associated to metabolism, stress responses, molecular repair, neural networks, and reproductive physiology in male and female medaka.
- Critics of EDC screening and testing programs have indicated that the current approach of regulatory agencies does not allow a thorough examination of synergistic interactions among chemicals (Bienen 2006). Our computational approach could be an effective tool for the assessment of EDC mixture activity and their synergistic effects.

Future Directions

To utilize the microarray assay in rapid screening and prioritizing chemicals for EDC activity.

To utilize the medaka microarray and computational approach to test for the activity of endogenous and synthetic hormones, pharmaceuticals and personal care products in ambient water.

Evaluate how EDC gene patterns change over time and dose.

References

- Bienen, L. (2006) *Frontiers Ecol. And the Environ.* 4(4):174.
Falconer et al. (2006) *Environ. Toxicol.* 21(2):181-191.
Muller and Wu (2007) In preparation.

Partners



Global Gene Expression Profiling of Androgen Disruption in Qurt Strain Medaka

ABIMAELEÓN,[†] PING-SHI WU,[‡]
LINDA C. HALL,[§]
MICHAEL L. JOHNSON,^{†,||} AND
SWEET J. TEH^{*†}

Aquatic Toxicology Program, School of Veterinary Medicine, and Center for Watershed Sciences, University of California, Davis, California 95616, Lehigh University, Bethlehem, Pennsylvania, and Lawrence Livermore National Laboratory, Livermore, California

Received July 19, 2007. Revised manuscript received October 30, 2007. Accepted November 5, 2007.

Androgen disrupting chemicals (ADCs) are endocrine disrupting chemicals (EDCs) that mimic or antagonize the effect of physiological androgens. Microarray-based detection of altered gene expression can be used as a biomarker of EDC exposure. Therefore, the purpose of this study was to identify and compare gene expression profiles of the androgen 11-ketotestosterone (11-KT), the antiandrogen flutamide (FLU), and the antiandrogenic fungicide vinclozolin (VIN), on Qurt medaka (*Oryzias latipes*). Biologically effective concentrations for 11-KT (100 $\mu\text{g/L}$), VIN (100 $\mu\text{g/L}$), and FLU (1000 $\mu\text{g/L}$) determined in range-finding studies were used for exposures. The oligonucleotide microarray included 9379 probes for EDC-affected genes, medaka cDNAs, sequences from the medaka genome project, and the UniGene database. We found that treatment with FLU, VIN, and 11-KT caused significant (false discovery rate = 0.01) differential expression of at least 87, 82, and 578 genes, respectively. Two sets of responsive genes are associated to vertebrate sex differentiation and growth, and 50 genes were useful in discriminating between ADC classes. The discriminating capacity was confirmed by a remarkable similarity of the antiandrogenic expression profiles of VIN and FLU, which were distinct from the androgenic profile of 11-KT. Gene expression profiles characterized in this study allow for reliable screening of ADC activity.

Introduction

Patterns of gene expression are sensitive indicators of toxicant exposure that can form a genetic signature of the toxin. This genetic signature can be used to identify toxic substances (1), including different classes of endocrine disrupting chemicals (EDCs) (2). In addition, patterns of gene expression could yield insight into modes of action and mechanisms of toxicity (3). There are several classes of EDCs including substances that can disrupt thyroid hormones, androgens,

and estrogens. These groups of compounds have been implicated as causes for a variety of effects, including increased rates of breast cancer, feminization, masculinization, and decreased fertility and hatching success in animal populations. Androgen disrupting chemicals (ADCs), namely, androgens and antiandrogens, are EDCs that mimic or antagonize the effect of physiological androgens by inducing or inhibiting androgen-receptor-mediated gene transcription, respectively. Environmental ADCs include plasticizers, chlorinated hydrocarbon insecticides, herbicides, and fungicides (e.g., vinclozolin, VIN) that can potentially impair early development of organisms (4, 5). The development of ADC-specific gene signatures in expression profiling could help identify ADC activity by screening previously uncharacterized or putative toxicants for their similarity to prototypic ADCs.

A gene expression profile could differ on the basis of the gender exposed (6). It is expected that some genes may be upregulated in males and downregulated in females, or vice versa. Consequently, correct interpretation of a gene expression profile depends on knowing the gender of the test organism. The sex-linked coloration in the Qurt medaka strain allows the genetic sex to be known before the phenotypic sex (7). Consequently, fish of a known gender can be exposed early in development and the gene expression profile can be interpreted correctly.

The aims of this study were as follows: to design a custom medaka microarray to detect EDCs; to identify gene expression profiles associated with androgen and antiandrogen activity; to conduct a statistical analysis of gene expression profiles to develop response criteria that identify patterns predictive of androgen disruption; and to validate the medaka custom microarray using a putative antiandrogen to evaluate its potential to discriminate androgenic and antiandrogenic chemicals. The specific objective of the present study is to characterize alterations in gene expression induced by 11-ketotestosterone (11-KT), a prototypic androgen; flutamide (FLU), a prototypic antiandrogen; and vinclozolin (VIN), a putative antiandrogen, on a sensitive medaka fish model. 11-KT is the primary endogenous fish androgen (8). Since 11-KT is not converted to estradiol by P450 aromatase, the differential gene expression observed in exposed medaka fish should be that expected from a pure androgenic agent. FLU is a therapeutic agent for prostate cancer treatment (9) that can function as an antiandrogen in mammals (10). We hypothesized that androgen- and antiandrogen-specific toxicity can be detected by analyzing changes in global gene expression that would indicate a molecular response following exposure. Overall, our results shed light on the genetic changes that may underlie the action of ADCs during early development.

Experimental Section

Determination of Biologically Effective Concentrations for ADCs. The prototypic androgen (11-KT; 98% purity), antiandrogen (FLU; 99% purity), and putative antiandrogen (VIN; 99.6% purity) were obtained from Sigma-Aldrich Co. (St. Louis, MO). Biologically effective concentrations (BECs) were determined by range-finding experiments in which male or female Qurt strain medaka were exposed separately to 11-KT, FLU (11), or VIN. A 96 h static renewal method was used for the aquatic exposures to each chemical by following standardized guidelines from the United States Environmental Protection Agency (12). A sample size of 50 1-week-old larvae (stage 40 first fry stage (13)) per gender per replicate was used in 11-KT and FLU exposures. After treatment, larvae were grown out for a 3 month period to approximate sexual

* Corresponding author phone: (530) 754-8183; fax: (530) 752-7690; e-mail: sjteh@ucdavis.edu.

[†] Aquatic Toxicology Program, School of Veterinary Medicine, University of California, Davis.

[‡] Lehigh University.

[§] Lawrence Livermore National Laboratory.

^{||} Center for Watershed Sciences, University of California, Davis.

maturity. The end points evaluated in these range-finding studies were growth performance indices (body length, body weight, and condition factor), gonad and thyroid histopathology, and occurrence of intersex (11).

Development of a Custom Medaka 12K Oligonucleotide Array.

A custom 12K oligonucleotide array for medaka was developed in collaboration with CombiMatrix Corp (14). The medaka microarray elements include: EDC-affected genes identified from searches of the scientific literature ($n = 131$), medaka cDNAs ($n = 1075$), medaka genome project sequences ($n = 890$), and unigene database probes ($n = 8154$). The microarray contains 9379 DNA probes including redundant sequences extracted from more than one source. Specific probes were designed by CombiMatrix Corp. using proprietary probe design software. The medaka CustomArray probes have intermediate lengths (34–39 nt). There are one to two replicates of each probe distributed across each array.

Microarray Experiment Design. Biological samples were prepared from Qurt medaka larvae poly A+ RNA after exposure to 11-KT, FLU, or VIN at BECs of 100, 100, and 1000 $\mu\text{g/L}$, respectively. Briefly, the total RNA was isolated from whole tissue of 7 day (± 1) post hatching (7 dph) larvae ($n = 200$ larvae per gender per each of three replicates) using TRIZOL reagent (Invitrogen). Integrity of the isolated RNA was verified on a 1% agarose gel. Poly A+ RNA was isolated using Qiagen (Valencia, CA) Oligotex midi-kits. First strand cDNA incorporating amino allyl-dUTP was synthesized from poly A+ RNA, and CyDye-labeled cDNA was purified by using the Amersham Biosciences (USA) CyScribe cDNA post labeling kit. UV/visible spectrophotometry measurements were used to calculate the concentration of fluorescent Cy3- and Cy5-dye incorporated in the cDNA sample. Spike-in control transcripts were prepared similarly using polymerase chain reaction (PCR) fragments, containing a T7 RNA polymerase promoter site, as a template for transcription.

The microarrays were assembled with hybridization caps and rehydrated with RNase-free water (Ambion) at 65 °C for 10 min. After rehydrating, blocking solution (6X SSPE, 20 mM ethylenediaminetetraacetic acid (EDTA), 0.05% Tween-20, 5X Denhardt's solution, 0.05% sodium dodecyl sulfate (SDS), and 100 ng/ μL sonicated salmon sperm DNA) was added, and the arrays were incubated at hybridization temperature (50 °C) for 30 min. A 1.5 μg labeled cDNA sample from each treatment was added to the hybridization solution (6X SSPE, 20 mM EDTA, 0.05% Tween-20, 0.05% SDS, and 100 ng/ μL sonicated salmon sperm DNA) and denatured for 3 min at 95 °C. Samples were placed briefly on ice followed by centrifugation at 14000g (maximum) for 3 min. Blocking solution was removed from the hybridization chamber, and 100 μL of hybridization solution was applied to the arrays. Hybridization was carried out in a Fisher Scientific Isotemp hybridization incubator for 18 h at 45 °C under gentle rotation. Following hybridization, arrays were washed at hybridization temperature (50 °C) for 5 min with 6X SSPE and 0.05% Tween-20 preheated to 50 °C. Washing continued with 3X SSPE and 0.05% Tween-20 for 5 min at room temperature (RT), 0.5X SSPE and 0.05% Tween-20 for 5 min at RT, and 2X phosphate-buffered saline (PBS) and 0.1% Tween-20. Final washing steps were performed at RT for 5 min with two rounds of 2X PBS with no detergent. The microarrays were imaged with Cy5 filter sets on an Applied Precision (Issaquah, WA) arrayW0Rx Biochip Reader. Imaging was performed while the array was wet with Imaging Solution (2X PBS) under a LifterSlip glass coverslip (Erie Scientific, Portsmouth, NH).

Identifying Unique Gene Patterns of Each Class of ADC.

ScanAlyze v.2.5 (Stanford University) was used to estimate the gene expression levels by obtaining the median intensity of replicate spots from the image. An initial set of differentially expressed genes in each treatment was obtained by conducting a one-sample t test with pooled variance. Differ-

TABLE 1. Numerical Distribution of Differentially Expressed Genes Identified in Qurt Medaka Larvae after Exposure to BECs of 11-KT and FLU^a

treatment	upregulated genes	downregulated genes	total
100 $\mu\text{g/l}$ 11-KT Males	518 ^b (216) ^c	650 (362)	1168 (578)
100 $\mu\text{g/l}$ 11-KT Females	27 (17)	116 (65)	143 (82)
1000 $\mu\text{g/l}$ FLU Males	75 (17)	260 (70)	335 (87)
1000 $\mu\text{g/l}$ FLU Females	853 (492)	962 (518)	1815 (1010)

^a BECs = biologically effective concentrations; 11-KT = 11-ketotestosterone; FLU = flutamide. ^b Differentially expressed genes identified with an FDR = 0.05 ^c Numbers in parenthesis represent unique genes identified with an FDR = 0.01

entially expressed genes were identified by the Benjamini and Hochberg (15) procedure for false discovery rate (FDR) control at 0.01 and 0.05 values. Differential expression of a set of genes involved in sex differentiation and growth in medaka and other vertebrates was summarized by heat map visualization using GenePattern 2.0 (MIT and Harvard) (16).

Loess normalization (17) and functional embedding (18) were used for high-dimensional data analyses and pattern discovery. First, the array data were normalized with the loess normalization scheme. For each gene, a two-sample t test was performed to obtain a significance value (p -value) for the difference between chemical treatments. A working set of diagnostic genes, defined by selecting the top 50 genes with the most significant differences (smallest p -values) between treatments, is believed to provide the most powerful means of discriminating between the chemical treatments at the transcription level. A latent ordering of the working set of genes allowed for pattern identification and analysis with the technique of functional embedding (18). This technique was applied to discover the patterning of gene expression for Qurt medaka after exposure to the selected ADCs. The response of each group of genes was represented by smoothed mean curves with standard error bars. The mean curves were smoothed with scatter smoother.

Biological information on significantly affected genes was obtained by functional profiling of genes using sequence similarity analyses with BLAST (19, 20). Gene annotation was based on the lowest expectation values (e -value) obtained from sequence similarity searches. The threshold value for establishing gene identity and function was E -value $< e^{-10}$. Non-medaka sequences were considered orthologous to a medaka gene if the e -value was within the established threshold.

Results

Chemical-Specific and Gender-Specific Effects. Up to 1815 (FDR = 0.05) differentially expressed genes were identified after androgen and antiandrogen treatment of Qurt medaka (Table 1). Overall, the ratio of upregulation to downregulation was close to unity (1:1.34) but inhibition responses were predominant. Males had a higher number of androgen-responsive genes (1168), while females had a higher number of antiandrogen-responsive genes (1815). Moreover, there were far more downregulated genes in females treated with antiandrogen than in males treated with androgens.

There was considerable variation in expression levels of affected genes in spite of rigid control of experimental conditions. Nevertheless, the microarray analysis revealed a series of chemical- and gender-specific alterations of gene expression profiles induced by 11-KT and FLU. Different sets of unique genes were identified within each gender after treatment with 11-KT or FLU (Supporting Information Figure S1A,B), and each gender showed unique responses to the

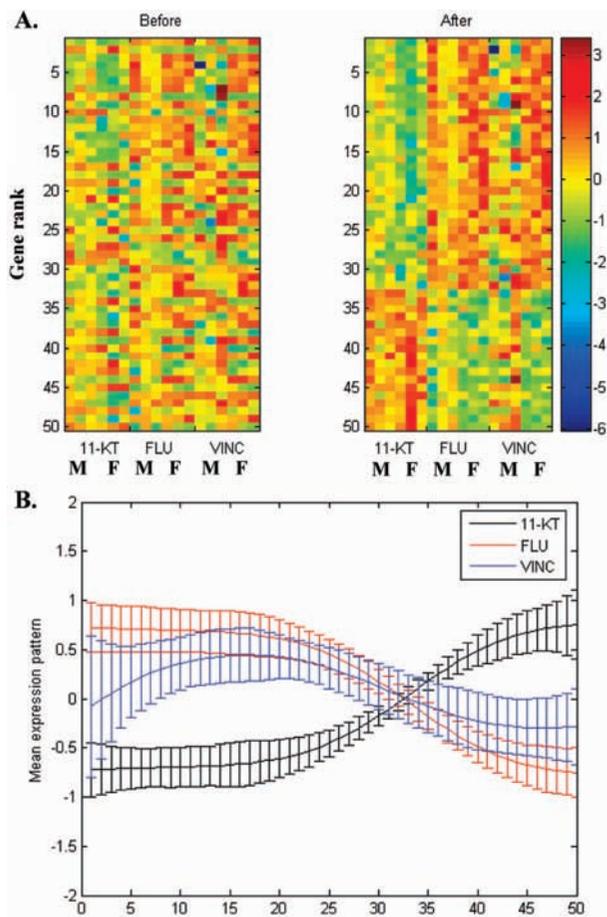


FIGURE 2. Expression profiles of 50 diagnostic genes: (A) heat map images of genes (rows) significantly altered by 11-KT, FLU, and VIN treatments before and after application of a functional embedding clustering algorithm (Müller and Wu 2006). Three replicate treatments for males (M) and females (F) are represented in tandem. The heat map images were then transformed to B line curves using scattered smoothed mean curves with standard error bars.

the androgen receptor (AR), estrogen receptor alpha (ER α), and sox9 protein in males. FLU treatment inhibited sGnRH, cGnRH-II, dmrt1Y (Dmy) gene, and growth hormone (Gh) in males. On the other hand, FLU stimulated the expression of mdGnRH in females while inhibiting the expression of GnRH-R3, P450c17, CYP19, estrogen receptor beta (ER β), sox9 protein, and Wnt4 gene in females.

Diagnostic Genes. The medaka microarray can differentiate between the two biological categories of ADCs (Figure 2). Functional embedding was quite successful in terms of classifying ADCs (Figure 2A). A set of 50 genes with the most significant differential expression allowed discrimination between medaka treated with an androgen or an antiandrogen. Evaluated collectively, the expression level of diagnostic genes can be monitored to predict or distinguish the class of unknown or putative ADCs. For example, our results show that the gene patterns revealed by functional embedding of 50 genes after 11-KT, FLU, and VIN treatment make it feasible to classify the latter as an antiandrogen.

The set of 50 genes are variably expressed within the classes formed by 11-KT and FLU treatments. Twenty-two out of the 50 genes constitute sequences with no known function. The genes form two distinct clusters after statistical analysis. One cluster is formed by the genes in positions 1–32, and a second cluster is formed by genes in positions 33–50 after latent ordering (Figure 2B). The first cluster includes transcripts for the following genes (in alphabetical

order) whose expression were downregulated by 11-KT and upregulated by FLU: Cyfip, calpain, WOHU α -2-HS-glycoprotein precursor (human), ubiquitin C1, ferritin chain H, elf-6, calbindin 2, C-type natriuretic peptide-1(cnp-1), homeobox gene, hox gene cluster, and ubiquitin polyprotein (heat shock related). A second cluster includes transcripts for genes which were upregulated by 11-KT and downregulated by FLU: elastase 2-like protein A, transforming acidic coiled coil 1a, MHC Class I, tropomyosin 1 alpha (brain), DMRT1a, keratin type I. A subset of gender-specific expression of functionally meaningful genes is presented in Table S1 in the Supporting Information.

Discussion

Chemical-Specific and Gender-Specific Effects. A custom medaka microarray chip was developed and tested for its ability to discriminate between ADC classes. At least three elements of the microarray experiment design of this study eliminate various intrinsic sources of variation. First, all of the exposed fish were at the same developmental stage, and the uniformity of culture conditions indicates it is likely that they shared similar physiological states. Second, inbred Qurt strain medaka colonies have been maintained for more than 10 years in our laboratory prior to this study which reduces the effect of genetic polymorphisms on the observed alterations in the gene expression. In addition, all of the fish were sexed prior to exposure to ADCs to allow testing for gender-specific alterations in gene expression response of larvae. Finally, by pooling the RNA samples by gender from a large number of fish, we characterized the mean gene expression response of each gender to ADC exposure without concern for individual variation in a few individuals.

Androgen treatment caused gender-specific alterations in gene expression of Qurt medaka larvae. Males exposed to 11-KT showed upregulation of genes related to cell metabolism, central nervous system (CNS) signaling and steroidogenesis. Exposure to this androgen caused downregulation of specific stress protein genes in male (HSP70 mRNA) and female medaka (HSP90 alpha mRNA). In contrast, hsp 90 beta was upregulated in males exposed to both androgens and antiandrogens. Females exposed to 11-KT showed inhibition of genes involved in phase I metabolism (CYP2K3), zona pellucida and vitellogenin 1. In a related study, Blum et al. (21) found that 11-KT stimulated the expression of vitellogenin 2 in liver of adult male largemouth bass (*Micropterus salmoides*) after exposure to natural androgens. Androgen treatment has gender-specific effects on vitellogenin expression, a transcript that is normally synthesized by oviparous females.

Antiandrogen treatment decreased the synthesis of at least three gene transcripts required for normal male development. First, FLU inhibited the production of male sex determining protein gene (dmrt1y). DMRT1Y (DMY) is expressed during male embryonic and larval development of medaka (22) and is responsible for testicular differentiation. However, our study shows that FLU inhibits expression of the gene that directs testicular differentiation which had not been previously reported for Qurt strain medaka. Second, FLU downregulates the production of GnRH which could alter major pathways of steroidogenesis in medaka. Third, FLU altered the expression of the cap-binding subunit of the elongation factor (eIF-4E) gene previously described as female-specific and ovary-specific during early oogenesis in Qurt medaka (23). However, our results suggest that the eIF-4E transcript is present in 7dph male medaka. Moreover, the inhibition of eIF-4E by FLU could disrupt translational regulation with subsequent effects on cell growth and development in males.

Sex Determination, Sex Differentiation, and Growth Genes. As noted above, dmy was downregulated by FLU treatment in males. The DMY gene is a DNA sequence present

in the sex-determining region of the Y chromosome of medaka (24, 25), which was first isolated from inbred HNI medaka strain (26). The *dmy* gene is believed to result from recent evolution of a duplicated copy of *dmrt1* and has not been identified in other teleost species (22, 27–31). Similar to the mammalian *SRY/Sry* gene, *dmy* is the master gene at the top of the male sex determination cascade in medaka. Scholz et al. (32) suggested that the key regulatory role of the *DMY* gene in sex determination is apparently hormone insensitive. On the other hand, the present study shows that *DMY* expression is labile to xenobiotic disruption. Therefore, *dmy* downregulation in medaka is a possible mechanism leading to FLU-induced intersex (11) and/or sex reversed individuals due to disruption of the sex determination pathway.

DM domain genes, including *dmrt1Y* (*dmy*) and the autosomal *dmrt1* gene, are a family of genes involved in sex determination and differentiation. We found that a Qurt strain medaka ortholog (Accession No. BJ730394) that is highly similar to HNI strain medaka *dmrt1a* (Accession No. AY157712) was upregulated by 11-KT and downregulated by FLU (Figure 1). Androgenic treatment may upregulate *DMRT1* (24). *Dmrt1* is an autosomal gene required for testicular differentiation that may promote testicular development and spermatogenesis in medaka (24, 25, 33). Besides the DM-domain genes, we found differential expression of other genes with putative roles in sex differentiation and growth, as described below (see also Figure 1).

Three GnRH forms were differentially regulated in Qurt medaka. Potential alterations to signals mediated by GnRHs and GnRH-Rs reveal top-down effects of ADC exposure in Qurt medaka reproductive physiology. GnRH is a decapeptide that acts on the pituitary to stimulate the synthesis and secretion of gonadotropins. At least three molecular forms of GnRH have been identified in medaka: medaka-type GnRH (mdGnRH), chicken-II-type GnRH (cGnRH-II), and salmon-type GnRH (sGnRH) (34, 35). FLU inhibited both sGnRH and cGnRH-II gene expression in males, while it stimulated mdGnRH expression in females. Each form has a distinct role in medaka. While mdGnRH is a hypophysiotropic factor, cGnRH-II and sGnRH are presumably neuromodulators or neurotransmitters commonly referred to as brain GnRHs and are both physically and functionally linked (36). They have been implicated in the modulation of excitability and development of neuronal cells, the proliferation of tumor cells, and the control of reproductive behavior. sGnRH is found only in the teleost lineage (36). In the present study, 11-KT treatment had no effect on sGnRH expression. On the other hand, treatment with FLU inhibited sGnRH mRNA expression in male Qurt medaka. Treatment with 11-KT caused no effect on cGnRH-II production in African catfish (37). In contrast, we observed an inhibition of cGnRH-II in medaka by 11-KT treatment.

The expression of the gene for GnRH receptor (GnRH-R3) in Qurt medaka was inhibited by 11-KT and FLU in males and females, respectively. All GnRH-Rs are expressed in the pituitary. Therefore, GnRH-R inhibition was likely due to a negative feedback mechanism that would lead to a reduction of the responsiveness of the medaka pituitary to GnRH stimulation. GnRH-R3 has been shown to exhibit higher and approximately equal selectivity for two of the three native forms of medaka GnRH; cGnRH-II, and sGnRH (38). Therefore, the neuromodulatory functions of the brain GnRHs could be inhibited as a result of ADC treatment of Qurt medaka with potential reproductive effects via disruption of gonad growth and development and endogenous steroid secretion.

Steroidogenesis pathways may have been altered by 11-KT and FLU exposure of Qurt medaka. Steroidogenesis is modulated in an autoregulatory manner by androgens (39).

In fish, 11-KT generally triggers an upregulatory response that augments the steroidogenic capacity of juveniles; this effect can be blunted by high dosages (see ref 40).

P450c17 gene expression was inhibited in female Qurt medaka exposed to FLU. P450c17 is a steroidogenic enzyme which normally occurs at high levels in the ovary. P450c17 catalyzes key steps in the metabolism of pregnenolone to the 11-oxygenated androgens, including 11-KT, in teleost fish (39).

The expression of the gene for another steroidogenesis enzyme, P45011 β , was inhibited by 11-KT in male medaka. P45011 β is the first enzyme in the synthesis of 11-oxygenated androgens from precursors (41). For instance, P45011 β is a key enzyme in the production of 11-KT, and its mRNA is highly abundant in the medaka testis (42). It was previously found that xenoestrogen exposure inhibits 11-KT production through downregulation of P45011 β mRNA in male medaka (42). Similarly, we found that exposure to exogenous 11-KT downregulates the P45011 β gene in males. Therefore, we hypothesize that exogenous 11-KT may trigger a negative feedback mechanism that inhibits its endogenous production in male fish.

Endogenous estrogen production may have been altered by ADC exposure of Qurt medaka. The enzyme aromatase (CYP19) catalyzes the conversion of testosterone to estrogen. The CYP19 gene was inhibited in males treated with 11-KT and in females treated with FLU. In both cases, CYP19 inhibition could lead to a lower rate of conversion of endogenous androgens to estrogens. There are two isoforms of aromatase in teleost species, ovarian (CYP19A) and brain (CYP19B). CYP19A activity is regulated at the transcriptional level in medaka vitellogenic follicles (43, 44). High aromatase activity is associated with ovarian differentiation (45). Hence, CYP19 gene inhibition in Qurt females treated with FLU is a likely mechanism for antiandrogen-induced intersexuality and sex reversal in females due to an alteration of sex steroid homeostasis by ADCs.

Similar to the CYP19 gene, the *sox9* gene was downregulated by 11-KT in males and by FLU in females. Our results confirm previous studies indicating that *Sox* mRNA is normally expressed in both male and female medaka (46) and that *Sox9* expression is similar to that of aromatase (45). The *sox9* gene is essential for cartilage and testis development in vertebrates. Gonad expression of *sox9* gene has diversified in fish species (47, 48). In medaka, *sox9* was expressed in oocytes in the developing ovary but not dominantly expressed in the developing testis (48). According to Nakamoto et al. (49), testicular type *Sox9* might not be involved in sex determination or early sex differentiation in medaka but has a crucial role in the maintenance of male gonadal differentiation. Nakamoto et al. (49) suggest that it is likely that estrogen downregulates *sox9* gene expression. In the present study, downregulation of *sox9* after treatment of females with FLU has an effect similar to that expected from estrogen treatment. On the other hand, *Sox9* participates in the development of testicular structures in the medaka but we observed *sox9* inhibition after androgen treatment of males. This apparently contradictory observation suggests a negative feedback of exogenous androgens on the signaling pathways that lead to testicular development.

Sex steroid receptors were also a target of ADC disruption in Qurt medaka. Sex steroids are absent in medaka during sex differentiation but the sex steroid receptors appear early in development (50). In the present study, 11-KT and FLU affected expression of two different isoforms of the ER. 11-KT treatment inhibited ER α in males, while FLU treatment inhibited both ER α and ER β in females. The ER expression pattern in medaka exposed to estrogenic EDCs had been attributed to direct receptor induction by the chemicals (51) while both ADCs in our study elicited an inhibitory response

on the ER. On the other hand, AR expression was inhibited by 11-KT exposure in medaka males.

Besides affecting the mRNA of genes with roles in early reproductive physiology, ADC exposure interfered with other genes with putative roles in growth and development in medaka. Wnt4 and growth hormone (GH) genes were downregulated by antiandrogen treatment in females and males, respectively. Wnt genes encode glycoproteins (WNTs) that mediate cell–cell signaling, control cell differentiation and proliferation, and could be useful as molecular markers for development and organogenesis in medaka (52). The wnt4 gene is expressed during embryogenesis and has been implicated in brain formation. Therefore, it is possible that FLU treatment impaired brain development in female medaka. On the other hand, the GH encoding gene plays a major role in the growth and development of vertebrates and could potentially enhance the growth rate of fishes in aquaculture (53). Downregulation of the GH gene by FLU treatment could indicate a molecular mechanism for the previously observed growth inhibition in medaka (11).

Diagnostic Genes. The expression profile for VIN is remarkably similar to the prototypical antiandrogen FLU, indicating that our system appears to be capable of prospectively identifying ADCs with characteristic modes of action. A group of 32 genes was downregulated by 11-KT and upregulated by FLU. A subset of genes within this cluster has diverse putative biological roles related to growth and larval development (elf6); body plan development (MF-Hox8–1 and hox genes) (54, 55); neuronal connectivity (Cyfip) (56, 57); homeostasis of bone (α_2 -HS-glycoprotein precursor), body fluids (cnp-1), and cell homeostasis (calpain) (58–60); heat shock response (ubiquitin polyprotein); calcium transport (calbindin 2); protein stability; and iron regulation (ferritin chain H) (61). A second group of 28 genes was upregulated by androgen treatment and downregulated by antiandrogen treatment. The putative roles of genes in this subset include sex differentiation (dmrt1a) (27), immune function (MHC class I) (62), muscular function (tropomyosin 1 alpha) (63), formation of intermediate filaments (keratin type I), cell adhesion (tgf BIGH3 precursor); degradation of extracellular matrix components (elastase 2) (64), and chromatin remodeling (TACC1a) (65).

In summary, the present study illustrates that gene expression profiling with a customized medaka oligonucleotide array provides a preliminary understanding of androgen and antiandrogen disruption of sex differentiation and development at the transcript level. Chemical-specific altered gene expression after 11-KT and FLU treatment in Qurt medaka indicates that microarray-based screening to detect androgenic and antiandrogenic activity appears to be a viable technique. The custom-designed array constitutes a powerful screening tool that can simultaneously yield information on the ADC activity of parent chemicals and their active metabolites. This screening tool could be used to characterize organ-, stage-, and dose-specific differences in ADC gene expression profiles or to obtain gene expression profiles of complex contaminant mixtures to facilitate environmental ADC identification and evaluation. Functional profiling of affected genes and gene pathways will lay the groundwork for cross-species comparisons on shared mechanisms and vulnerabilities to these chemicals.

Acknowledgments

This study was funded by the Science to Achieve Results (STAR) grant (No. R831299) of the United States Environmental Protection Agency (U.S. EPA), Office of Research and Development (ORD), Computational Toxicology Program, and the California State Water Resources Control Board (Contract No. 04-111-250-0). The authors declare that they have no competing financial interests. We thank Ching Teh for her assistance with

the culture and care of the experimental fish and the separation of 2-day-old male and female medaka embryo. We also thank Dr. Akihiro Shima of Japan for his generous gift of Qurt medaka and Drs. Kiyoshi Naruse and Kazuei Mita of Japan for providing expressed sequence tags.

Supporting Information Available

Figures showing different sets of unique genes and a table listing a subset of gender-specific expression of functionally meaningful genes. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Literature Cited

- (1) Bartosiewicz, M. J.; Jenkins, D.; Penn, S.; Emery, J.; Buckpitt, A. Unique Gene Expression Patterns in Liver and Kidney Associated with Exposure to Chemical Toxicants. *J. Pharmacol. Exp. Ther.* **2001**, *297*, 895–905.
- (2) Iguchi, T.; Watanabe, H.; Katsu, Y. Application of Ecotoxicogenomics for Studying Endocrine Disruption in Vertebrates and Invertebrates. *Environ. Health Perspect.* **2006**, *114*, 101–105.
- (3) Lettieri, T. Recent Applications of DNA Microarray Technology to Toxicology and Ecotoxicology. *Environ. Health Perspect.* **2006**, *114*, 4–9.
- (4) Daxenberger, A. Pollutants with Androgen-Disrupting Potency. *Eur. J. Lipid Sci. Technol.* **2002**, *104*, 124–130.
- (5) Gray, L. E., Jr.; Wilson, V. S.; Stoker, T.; Lambright, C.; Furr, J.; Noriega, N.; Howdeshell, K.; Ankley, G. T.; Guillette, L. Adverse Effects of Environmental Antiandrogens and Androgens on Reproductive Development in Mammals. *Int. J. Androl.* **2006**, *29*, 96–104.
- (6) Kishi, K.; Kitagawa, E.; Onikura, N.; Nakamura, A.; Iwahashi, H. Expression Analysis of Sex-Specific and 17 β -Estradiol-Responsive Genes in the Japanese Medaka, *Oryzias Latipes*, Using Oligonucleotide Microarrays. *Genomics* **2006**, *88*, 241–251.
- (7) Wada, H.; Shimada, A.; Fukamachi, S.; Naruse, K.; Shima, A. Sex-Linked Inheritance of the If Locus in the Medaka Fish (*Oryzias latipes*). *Zool. Sci.* **1998**, *15*, 123–126.
- (8) Borg, B. Androgens in Teleost Fish. *Comp. Biochem. Physiol.* **1994**, *109C*, 219–245.
- (9) Kolvenbag, G. J. C. M.; Iversen, P.; Newling, D. W. W. Antiandrogen Monotherapy: A New Form of Treatment for Patients with Prostate Cancer. *Urology* **2001**, *58* (2A), 16–22.
- (10) Bayley, M.; Junge, M.; Baatrup, E. Exposure of Juvenile Guppies to Three Antiandrogens Causes Demasculinization and a Reduced Sperm Count in Males. *Aquat. Toxicol.* **2002**, *56*, 227–239.
- (11) León, A.; Teh, S. J.; Hall, L. C.; Teh, F. C. Androgen Disruption of Early Development in Qurt sStrain Medaka (*Oryzias latipes*). *Aquat. Toxicol.* **2007**, *82*, 195–203.
- (12) U.S. EPA. Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms 5th ed.; Office of Water, U.S. Environmental Protection Agency: Washington, D. C., 2002; EPA-821-R-02-012.
- (13) Iwamatsu, T. Stages of Normal Development in the Medaka *Oryzias latipes*. *Zool. Sci.* **1994**, *11*, 825–839.
- (14) Orbus, R.; Cooper, J.; Strathmann, M. Assessing the Sensitivity, Dynamic Range and Reproducibility of the CombiMatrix CustomArray Platform, 2005. <http://www.combimatrix.com/docs/CustomArraySensitivity.pdf> (Accessed Jul 16, 2007).
- (15) Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Statist. Soc. B* **1995**, *57* (1), 289–300.
- (16) Reich, M.; Liefeld, T.; Gould, J.; Lerner, J.; Tamayo, P.; Mesirov, J. P. GenePattern 2.0. *Nat. Genet.* **2006**, *38* (5), 500–501.
- (17) Cleveland, W. S.; Devlin, S. J. Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *J. Am. Stat. Assoc.* **1988**, *83* (403), 596–610.
- (18) Wu, P. S. Time-Dynamic Density Estimation and Functional Discrimination for High-Dimensional Data. Ph.D. Dissertation, University of CA–Davis, 2005.
- (19) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic Local Alignment Search Tool (BLAST). *J. Mol. Biol.* **1990**, *215*, 403–410.
- (20) Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs. *Nucleic Acids Res.* **1997**, *25* (17), 3389–3402.
- (21) Blum, J. L.; Knoebel, I.; Larkin, P.; Kroll, K. J.; Denslow, N. D. Use of Suppressive Subtractive Hybridization and cDNA Arrays To

- Discover Patterns of Altered Gene Expression in the Liver of Dihydrotestosterone and 11-Ketotestosterone Exposed Adult Male Largemouth Bass (*Micropterus salmoides*). *Mar. Environ. Res.* **2004**, *58*, 565–569.
- (22) Nanda, I.; Kondo, M.; Hornung, U.; Asakawa, S.; Winkler, C.; Shimizu, A.; Shan, Z.; Haaf, T.; Shimizu, N.; Shima, A.; Schmid, M.; Scharl, M. A Duplicated Copy of DMRT1 in the Sex-Determining Region of the Y Chromosome of the Medaka, *Oryzias latipes*. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (18), 11778–11783.
- (23) Kanamori, A. Systematic Identification of Genes Expressed during Early Oogenesis in Medaka. *Mol. Reprod. Dev.* **2000**, *55*, 31–36.
- (24) Matsuda, M. Sex Determination in the Teleost Medaka, *Oryzias latipes*. *Annu. Rev. Genet.* **2005**, *39*, 293–307.
- (25) Ohmuro-Matsuyama, Y.; Matsuda, M.; Kobayashi, T.; Ikeuchi, T.; Nagahama, Y. Expression of DMY and DMRT1 in Various Tissues of the Medaka (*Oryzias latipes*). *Zool. Sci.* **2003**, *20*, 1395–1398.
- (26) Matsuda, M.; Kusama, T.; Oshiro, T.; Kurihara, Y.; Hamaguchi, S.; Sakaizumi, M. Isolation of a Sex Chromosome-Specific DNA Sequence in the Medaka, *Oryzias latipes*. *Genes Genet. Syst.* **1997**, *72*, 263–268.
- (27) Kondo, M.; Nanda, I.; Hornung, U.; Asakawa, S.; Shimizu, N.; Mitani, H.; Schmid, M.; Shima, A.; Scharl, M. Absence of the Candidate Sex-Determining Gene *dmrt1b(Y)* of Medaka from Other Fish Species. *Curr. Biol.* **2003**, *13*, 416–420.
- (28) Lutfalla, G.; Crollins, H. R.; Brunet, F. G.; Laudet, V.; Robinson-Rechavi, M. Inventing a Sex-Specific Gene: A Conserved Role of DMRT1 in Teleost Fishes plus a Recent Duplication in the Medaka *Oryzias latipes* Resulted in DMY. *J. Mol. Evol.* **2003**, *57*, S148–S153.
- (29) Matsuda, M.; Sato, T.; Toyazaki, Y.; Nagahama, Y.; Hamaguchi, S.; Sakaizumi, M. *Oryzias curvinotus* Has DMY, a Gene That Is Required for Male Development in the Medaka, *O. latipes*. *Zool. Sci.* **2003**, *20*, 159–161.
- (30) Volf, J.; Kondo, M.; Scharl, M. Medaka *dmY/dmrt1Y* Is Not the Universal Primary Sex-Determining Gene in Fish. *Trends Genet.* **2003**, *19* (4), 196–199.
- (31) Zhang, J. Evolution of DMY, a Newly Emergent Male Sex-Determination Gene of Medaka Fish. *Genetics* **2004**, *166*, 1887–1895.
- (32) Scholz, S.; Rösler, S.; Schäffer, M.; Hornung, U.; Scharl, M.; Gutzeit, H. O. Hormonal Induction and Stability of Monosex Populations in the Medaka (*Oryzias latipes*): Expression of Sex-Specific Marker Genes. *Biol. Reprod.* **2003**, *69*, 673–678.
- (33) Kobayashi, T.; Matsuda, M.; Kajiuira-Kobayashi, H.; Suzuki, A.; Saito, N.; Nakamoto, M.; Shibata, N.; Nagahama, Y. Two DM Domain Genes, DMY and DMRT1, Involved in Testicular Differentiation and Development in the Medaka, *Oryzias latipes*. *Dev. Dyn.* **2004**, *231*, 518–526.
- (34) Okubo, K.; Amano, M.; Yoshiura, Y.; Suetake, H.; Aida, K. A Novel Form of Gonadotropin-Releasing Hormone in the Medaka, *Oryzias latipes*. *Biochem. Biophys. Res. Commun.* **2000**, *276*, 298–303.
- (35) Okubo, K.; Nagata, S.; Ko, R.; Kataoka, H.; Yoshiura, Y.; Mitani, H.; Kondo, M.; Naruse, K.; Shima, A.; Aida, K. Identification and Characterization of Two Distinct GnRH Receptor Subtypes in a Teleost, the Medaka *Oryzias latipes*. *Endocrinology* **2001**, *142* (11), 4729–4739.
- (36) Okubo, K.; Mitani, H.; Naruse, K.; Kondo, M.; Shima, A.; Tanaka, M.; Asakawa, S.; Shimizu, N.; Yoshiura, Y.; Aida, K. Structural Characterization of GnRH Loci in the Medaka Genome. *Gene* **2002**, *293*, 181–189.
- (37) Dubois, E. A.; Florijn, M. A.; Zandbergen, M. A.; Peute, J.; Goos, H. J. Th. Testosterone Accelerates the Development of the Catfish GnRH System in the Brain of Immature African Catfish (*Clarias gariepinus*). *Gen. Comp. Endocrinol.* **1998**, *112*, 383–393.
- (38) Okubo, K.; Ishii, S.; Ishida, J.; Mitani, H.; Naruse, K.; Kondo, M.; Shima, A.; Tanaka, M.; Asakawa, S.; Shimizu, N.; Aida, K. A Novel Third Gonadotropin-Releasing Hormone Receptor in the Medaka *Oryzias latipes*: Evolutionary and Functional Implications. *Gene* **2003**, *314*, 121–131.
- (39) Schulz, R. W.; Vischer, H. F.; Cavaco, J. E. B.; Santos, E. M.; Tyler, C. R.; Goos, H. J. T.; Bogerd, J. Gonadotropins, Their Receptors, and the Regulation of Testicular Functions in Fish. *Comp. Biochem. Physiol.* **2001**, *129B*, 407–417.
- (40) Cavaco, J. E. B.; van Blijswijk, B.; Leatherland, J. F. Th.; Goos, H. J.; Schulz, R. W. Androgen-Induced Changes of Circulating and Testicular Androgens in African Catfish, *Clarias gariepinus*. *Fish Physiol. Biochem.* **1999**, *17*, 155–162.
- (41) Kime, D. E. “Classical” and “Non-classical” Reproductive Steroids in Fish. *Rev. Fish Biol. Fish.* **1993**, *3*, 160–180.
- (42) Yokota, H.; Abe, T.; Nakai, M.; Murakami, H.; Eto, C.; Yakabe, Y. Effects of 4-tert-Pentylphenol on the Gene Expression of P450 11 β -Hydroxylase in the Gonad of Medaka (*Oryzias latipes*). *Aquat. Toxicol.* **2005**, *71*, 121–132.
- (43) Suzuki, A.; Tanaka, M.; Shibata, N. Expression of Aromatase mRNA and Effects of Aromatase Inhibitor during Ovarian Development in the Medaka, *Oryzias latipes*. *J. Exp. Zool.* **2004**, *301A*, 266–273.
- (44) Fukada, S.; Tanaka, M.; Matsuyama, M.; Kobayashi, D.; Nagahama, Y. Isolation, Characterization, and Expression of cDNAs Encoding the Medaka (*Oryzias latipes*) Ovarian Follicle Cytochrome P-450 Aromatase. *Mol. Reprod. Dev.* **1996**, *45*, 285–290.
- (45) Watanabe, M.; Tanaka, M.; Kobayashi, D.; Yoshiura, Y.; Oba, Y.; Nagahama, Y. Medaka (*Oryzias latipes*) FTZ-F1 Potentially Regulates the Transcription of P-450 Aromatase in Ovarian Follicles: cDNA Cloning and Functional Characterization. *Mol. Cell Endocrinol.* **1999**, *149*, 221–228.
- (46) Kuhl, A. J.; Manning, S.; Brouwer, M. Brain Aromatase in Japanese Medaka (*Oryzias latipes*): Molecular Characterization and Role in Xenostrogen-Induced Sex Reversal. *J. Steroid Biochem. Mol. Phys.* **2005**, *96*, 67–77.
- (47) Fukada, S.; Tanaka, M.; Iwaya, M.; Nakajima, M.; Nagahama, Y. The Sox Gene Family and Its Expression during Embryogenesis in the Teleost Fish, Medaka (*Oryzias latipes*). *Dev. Growth Differ.* **1995**, *37*, 379–385.
- (48) Klüver, N.; Kondo, M.; Herper, A.; Mitani, H.; Scharl, M. Divergent Expression Patterns of Sox9 Duplicates in Teleosts Indicate Lineage Specific Subfunctionalization. *Dev. Genes Evol.* **2005**, *215*, 297–305.
- (49) Yokoi, H.; Kobayashi, T.; Tanaka, M.; Nagahama, Y.; Wakamatsu, Y.; Takeda, H.; Araki, K.; Morohashi, K.; Ozato, K. *sox9* in a Teleost Fish, Medaka (*Oryzias latipes*): Evidence for Diversified Function of Sox9 in Gonad Differentiation. *Mol. Reprod. Dev.* **2002**, *63*, 5–16.
- (50) Nakamoto, M.; Suzuki, A.; Matsuda, M.; Nagahama, Y.; Shibata, N. Testicular Type Sox9 Is Not Involved in Sex Determination but Might Be in the Development of Testicular Structures in the Medaka, *Oryzias latipes*. *Biochem. Biophys. Res. Commun.* **2005**, *333*, 729–736.
- (51) Kawahara, T.; Okada, H.; Yamashita, I. Cloning and Expression of Genomic and Complementary DNAs Encoding an Estrogen Receptor in the Medaka Fish, *Oryzias latipes*. *Zool. Sci.* **2000**, *17*, 643–649.
- (52) Min, J.; Lee, S. K.; Gu, M. B. Effects of Endocrine Disrupting Chemicals on Distinct Expression Patterns of Estrogen Receptor, Cytochrome P-450 Aromatase and p53 Genes in *Oryzias latipes*. *J. Biochem. Mol. Toxicol.* **2003**, *17* (5), 272–277.
- (53) Yokoi, H.; Nishimatsu, A.; Ozato, K.; Yoda, K. Cloning and Embryonic Expression of Six *wnt* Genes in the Medaka (*Oryzias latipes*) with Special Reference to Expression of *wnt5a* in the Pectoral Fin Buds. *Dev. Growth Differ.* **2003**, *45*, 51–61.
- (54) Venkatesh, B.; Brenner, S. Genomic Structure and Sequence of the Pufferfish (*Fugu rubripes*) Growth Hormone-Encoding Gene: A Comparative Analysis of Teleost Growth Hormone Genes. *Gene* **1997**, *187*, 211–215.
- (55) Kurosawa, G.; Yamada, K.; Ishiguro, H.; Hori, H. *Hox* Gene Complexity in Medaka Fish May Be Similar to That in Pufferfish Rather Than Zebrafish. *Biochem. Biophys. Res. Commun.* **1999**, *260*, 66–70.
- (56) Kurosawa, G.; Takamatsu, N.; Takahashi, M.; Sumitomo, M.; Sanaka, E.; Yamada, K.; Nishii, K.; Matsuda, M.; Asakawa, S.; Ishiguro, H.; Miura, K.; Kurosawa, Y.; Shimizu, N.; Kohara, Y.; Hori, H. Organization and Structure of *hox* Gene Loci in Medaka Genome and Comparison with Those of Pufferfish and Zebrafish Genomes. *Gene* **2006**, *370*, 75–82.
- (57) Schenck, A.; Bardoni, B.; Moro, A.; Bagni, C.; Mandel, J. L. A Highly Conserved Protein Family Interacting with the Fragile X Mental Retardation Protein (FMRP) and Displaying Selective Interactions with FMRP-Related Proteins FXR1P and FXR2P. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98* (15), 8844–8849.
- (58) Schenck, A.; Bardoni, B.; Langmann, C.; Harden, N.; Mandel, J. L.; Giangrande, A. CYFIP/Sra-1 Controls Neuronal Connectivity in *Drosophila* and Links the Rac1 GTPase Pathway to the Fragile X Protein. *Neuron* **2003**, *38*, 887–898.
- (59) Azam, M.; Andrabi, S. S.; Sahr, K. E.; Kamath, L.; Kuliopulos, A.; Chishti, A. H. Disruption of the Mouse μ -Calpain Gene Reveals an Essential Role in Platelet Function. *Mol. Cell. Biol.* **2001**, *21* (6), 2213–2220.
- (60) Inoue, K.; Naruse, K.; Yamagami, S.; Mitani, H.; Suzuki, N.; Takei, Y. Four Functionally Distinct C-Type Natriuretic Peptides Found in Fish Reveal Evolutionary History of the Natriuretic Peptide System. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (17), 10079–10084.

- (61) Yang, F.; Chen, Z.; Bergeron, J. M.; Cupples, R. L.; Friedrichs, W. E. Human α 2-HS-Glycoprotein/Bovine Fetuin Homologue in Mice: Identification and Development Regulation of the Gene. *Biochim. Biophys. Acta* **1992**, *1130*, 149–156.
- (62) Ingrassia, R.; Gerardi, G.; Biasotto, G.; Arosio, P. Mutations of Ferritin H Chain c-Terminus Produced by Nucleotide Insertions Have Altered Stability and Functional Properties. *J. Biochem.* **2006**, *139*, 881–885.
- (63) Matsuo, M. Y.; Asakawa, S.; Shimizu, N.; Kimura, H.; Nonaka, M. Nucleotide Sequence of the MHC Class I Genomic Region of a Teleost, the Medaka (*Oryzias latipes*). *Immunogenetics* **2002**, *53*, 930–940.
- (64) Huang, M.; Ochiai, Y.; Watabe, S. Structural and Thermodynamic Characterization of Tropomyosin from Fast Skeletal Muscle of Bluefin Tuna. *Fish. Sci.* **2004**, *70*, 667–674.
- (65) Sarropoulou, E.; Power, D. M.; Magoulas, A.; Geisler, R.; Kotoulas, G. Comparative Analysis and Characterization of Expressed Sequence Tags in Gilthead Sea Bream (*Sparus aurata*) Liver and Embryos. *Aquaculture* **2005**, *243*, 69–81.
- (66) Still, I. H.; Vettaikorumakankau, A. K.; DiMatteo, A.; Liang, P. Structure-Function Evolution of the Transforming Acidic Coiled Coil Genes Revealed by Analysis of Phylogenetically Diverse Organisms. *BMC Evol. Biol.* **2004**, *4*, 1–16.

ES071785C