

# UC Irvine

## UC Irvine Previously Published Works

### Title

Influence of ovarian stage on transcript profiles in fathead minnow (*Pimephales promelas*) ovary tissue

### Permalink

<https://escholarship.org/uc/item/7121b2s5>

### Journal

Aquatic Toxicology, 98(4)

### ISSN

0166-445X

### Authors

Villeneuve, Daniel L  
Garcia-Reyero, Natàlia  
Martinović, Dalma  
[et al.](#)

### Publication Date

2010-07-01

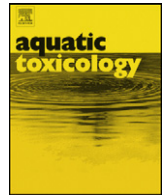
### DOI

10.1016/j.aquatox.2010.03.006

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed



## Influence of ovarian stage on transcript profiles in fathead minnow (*Pimephales promelas*) ovary tissue

Daniel L. Villeneuve<sup>a,\*</sup>, Natàlia Garcia-Reyero<sup>b</sup>, Dalma Martinović<sup>a,c</sup>, Jenna E. Cavallin<sup>a</sup>, Nathaniel D. Mueller<sup>a</sup>, Leah C. Wehmas<sup>a</sup>, Michael D. Kahl<sup>a</sup>, Anne L. Linnum<sup>a</sup>, Edward J. Perkins<sup>d</sup>, Gerald T. Ankley<sup>a</sup>

<sup>a</sup> US Environmental Protection Agency, Office of Research and Development, National Health and Environmental Effects Research Laboratory, Mid-Continent Ecology Division, 6201 Congdon Blvd, Duluth, MN 55804, USA

<sup>b</sup> Jackson State University, Department of Chemistry, Jackson, MS 39217, USA

<sup>c</sup> University of St. Thomas, Department of Biology, St. Paul, MN 55105, USA

<sup>d</sup> US Army Engineering Research and Development Center, Environmental Laboratory, 3909 Halls Ferry Rd, Vicksburg, MS 39108, USA

### ARTICLE INFO

#### Article history:

Received 16 December 2009

Received in revised form 3 March 2010

Accepted 8 March 2010

#### Keywords:

Oogenesis

Toxicogenomics

Reproduction

Histology

### ABSTRACT

Interpretation of toxicogenomic experiments conducted with ovary tissue from asynchronous-spawning small fish species is complicated by background variation in the relative abundance and proportion of follicles at different stages within the ovary tissue sample. This study employed both real-time quantitative polymerase chain reaction and a 15,000 gene oligonucleotide microarray to examine variation in the fathead minnow (*Pimephales promelas*) ovarian transcriptional profile as a function of quantitative and qualitative differences in ovarian histology. The objectives were to provide data that could potentially aid interpretation of future toxicogenomics experiments, identify putative stage-related transcriptional markers, and generate insights into basic biological regulation of asynchronous oocyte development. Multiple lines of evidence from the present study indicate that variation in the transcriptional profile is primarily dependent on the relative abundance of previtellogenic versus vitellogenic follicles in the ovary. Due to the relatively small proportions of mature ovulated follicles or atretic follicles in the overall follicle population, few potential transcriptional markers of maturation, ovulation, or atresia could be identified. However, among the 460 differentially expressed genes identified in the present study, several targets, including HtrA serine peptidase 3 (*htra3*), tissue inhibitor of metalloproteinase 3 (*timp3*), aquaporin 8 (*aqp8*), transgelin 2 like (*tagln2*), Nedd4 family interacting protein 2 (*ndfip2*), chemokine ligand 12a (*cxcl12a*), midkine-related growth factor (*mdka*), and jagged 1b (*jag 1b*) exhibited responses and functional properties that support them as candidate molecular markers of significant shift in gross ovarian stage. Genes associated with a diversity of functions including cellular development, morphogenesis, coated vesicle transport, sexual reproduction, and neuron development, among others, were statistically enriched within the list of 460 genes differentially expressed among different ovarian classes. Overall, results of this study provide insights into background variation in ovary transcript profiles that should aid and enhance the interpretation of toxicogenomic data generated in experiments conducted with small, asynchronous-spawning fish species.

Published by Elsevier B.V.

### 1. Introduction

Small fish species are commonly used model organisms in aquatic ecotoxicology research and regulatory testing. The fathead minnow (*Pimephales promelas*), zebrafish (*Danio rerio*), and Japanese medaka (*Oryzias latipes*), in particular, are being developed as models for endocrine disruptor screening and testing

programs worldwide (Ankley and Johnson, 2004), [<http://www.oecd.org/dataoecd/55/59/38662884.pdf>]. Compared to larger fish species and many vertebrate models, favorable traits of these organisms for testing include their small size (allowing for exposure of large numbers of individuals using comparatively small solution volumes and masses of test chemical), ease and cost-effectiveness of their husbandry, rapid generation times (making them well suited for developmental and life-cycle testing), and ability to reproduce continuously, year-around, under laboratory conditions (Ankley and Johnson, 2004). This latter attribute is associated with their asynchronous ovarian development (Tyler

\* Corresponding author.

E-mail address: [villeneuve.dan@epa.gov](mailto:villeneuve.dan@epa.gov) (D.L. Villeneuve).

and Sumpter, 1996). Specifically, rather than having oocytes develop homogeneously as a single batch, as occurs in synchronous ovulator/spawners, oocytes in asynchronous spawners develop in a heterogeneous manner, such that ovarian follicles representing multiple different stages of development are distributed throughout the ovarian tissue (Wallace and Selman, 1981; Tyler and Sumpter, 1996).

While the asynchronous reproductive strategy makes these species convenient model organisms for routine testing, it introduces certain challenges when it comes to applying genomic approaches to study the effects of endocrine-active chemicals and other stressors on the reproductive axis. Oocytes, and their surrounding follicular cells, representing different stages of follicle development, both produce and respond to different types of signaling molecules during maturation (Jalabert, 2005; Lubzens et al., 2009), and thus would be expected to have different transcription profiles depending on their relative stage of maturation. Consequently, when studying the effects of chemical contaminants or other stressors on ovarian transcript profiles, chemical effects must be interpreted over a background signal which is dependent on the relative number and proportions of different staged follicles present within the ovary. Transcriptional changes associated with chemical mechanisms of action that preferentially target signaling pathways dominant in certain follicular stages, but not others, may be diluted by the heterogeneous stage and distribution of follicles resulting in a reduction of the overall sensitivity to detect direct effects. Furthermore, direct effects of chemical treatment on gene expression might not be readily discriminated from the effects of the chemical on the asynchronous reproductive cycle and potential shifts in the overall representation of different stages (e.g., pre-vitellogenic, vitellogenic, mature ovulated oocytes, atretic follicles) present.

Previous research on endocrine-active chemicals has shown that chemical exposure can markedly alter the relative proportions of different staged follicles. For example, in female fathead minnows exposed to the aromatase inhibitor fadrozole for 21 d there was a significant reduction in histologically determined ovarian stage, and reduced vitellogenin uptake into developing oocytes (Ankley et al., 2002). Jensen et al. (2004) reported increased numbers of early stage follicles and atretic follicles in female fathead minnows exposed to the anti-androgen flutamide for 21 d. Weisbrod et al. (2007) reported increased numbers of pre-vitellogenic oocytes and decreased proportions of early and late vitellogenic oocytes in female fathead minnows exposed to >1.2 mg benzophenone-2/L or to 50 ng 17 $\alpha$ -ethynylestradiol for 15 d. Similar types of effects have been observed in zebrafish and medaka (Kiparissis et al., 2003; van der Ven et al., 2003). Such shifts in the relative proportions of differently staged follicles may reflect effects of the chemicals on the fish reproductive axis, and at the same time obscure aspects of the transcript profile that could yield mechanistic insights.

The present study was conducted to enhance the interpretation of toxicogenomic data generated in experiments with small, asynchronous-spawning fish species (e.g., Ankley et al., 2009). The intent was to identify groups of genes whose expression was altered by differences in gross ovarian stage, specifically, the relative proportions of different-staged oocytes within the ovary, in the absence of any chemical stressor. The overall approach was to collect ovary tissue from a large sample population (76 fish). Histological examination was then used to classify and select a set of ovary samples for microarray and quantitative real-time PCR (QPCR) analyses. While more sophisticated stereological methods, follicle sorting, laser capture microdissection, and/or in situ hybridization approaches would be better equipped to determine associations between specific ovarian structural domains and gene expression, for the purposes of this study we chose to focus on

samples that would be representative of those typically collected as part of routine ecotoxicological testing with adult fathead minnows and their associated heterogeneity. Similarly, histological examinations were based on a guidance document for the evaluation of endocrine disrupting chemical-induced gonad pathology in fathead minnows [<http://www.epa.gov/endo/pubs/att-h-histopathologyguidelines.fhm.pdf>], rather than more specialized methods.

Results of this study could potentially aid interpretation of future toxicogenomic experiments in several respects. First, alteration of genes identified as differentially expressed in the present study may serve as potential molecular markers of chemically induced shifts in gross morphological stage, as qualitatively defined in current histopathology guidance documents. Second, genes known to vary significantly with stage, regardless of chemical treatment, could be excluded from analyses aimed at inferring specific chemical mechanism(s) of action based on the ovarian transcript profile. Finally the data set may provide some general insights into the basic biological regulation of asynchronous oocyte development and/or help identify specific molecular targets on which to focus more targeted investigation.

## 2. Materials and methods

### 2.1. Fish and sample collection

All fish sampled in the present study were sexually differentiated female fathead minnows obtained from an on-site culture facility at the US EPA Mid-Continent Ecology Division, Duluth, MN, USA. The fish sampled ranged in age from 4 to 6 months. Most fish sampled were held at 25 °C under a 16:8 h light:dark photoperiod. However, to further enhance the likelihood of obtaining diverse ovarian stages, particularly ovaries containing significant numbers of atretic follicles, temperature and photoperiod manipulations were also used as part of sample generation. Specifically, one group of 16 females was held in tanks (4 fish per tank) in which the temperature was reduced 2 °C per day over 5 d, from 25 to 15 °C by altering the temperature of the external water bath in which the tanks were immersed, then held at 15 °C for an additional 5 d prior to sampling. Tank temperatures were monitored daily. Another group of 16 females (four tanks of four fish per tank) was held under a photoperiod that was shifted by 2 h each day over a period of 4 d from a 16:8 h light:dark photoperiod to a 8:16 h light:dark photoperiod, and held under that inverted photoperiod for an additional 6 d prior to sampling. At the time of sample collection, fish were humanely killed in a buffered (200 mg Na<sub>2</sub>HCO<sub>3</sub>/L) solution of tricaine methanesulfonate (MS-222; Finquel; Argent, Redmond WA, USA). Gentle pressure was applied to the abdomen of each fish. If eggs were released, the fish was noted to contain ovulated eggs. Blood was collected from the caudal peduncle using heparinized microcapillary tubes, and plasma was separated by centrifugation and subsequently stored at –80 °C. Ovaries were removed, weighed, and laid side by side, oriented anterior to posterior as in the intact fish. The anterior and posterior tip of each pair of ovaries was removed and discarded and the remaining tissue was then divided into three subsamples of approximately equal size. The anterior- and posterior-most subsamples were transferred to labeled microcentrifuge tubes and snap-frozen in liquid nitrogen, then stored at –80 °C until extracted for metabolomic (not reported here) and transcriptomic analyses, respectively. The center subsample was placed in a tissue cassette and immersed in Davidson's fixative [<http://www.epa.gov/endo/pubs/att-h-histopathologyguidelines.fhm.pdf>] for histological analysis. Ovary tissue was collected from a total of 76 fish.

**Table 1**

Criteria used to assign samples to histological classifications (histoclasses). To be included in class, samples had to meet both the quantitative and qualitative criteria listed. Samples that were evaluated but did not meet any of the defined criteria were not included in the transcriptomic and real-time PCR analyses.

| Histoclass | Quantitative criteria  | Qualitative criteria   |
|------------|--|--|
| Prevgtg    | ≥90% of oocytes/follicles in the section(s) characterized as previtellogenic   | Few or no atretic follicles observed   |
| Mature     | Minimal atretic follicles (<2.5%)  | From fish that were noted to release eggs upon applying gentle pressure to their abdomen |
| Atretic    | >7% atretic follicles in the section(s) characterized                          | Stage 3–4  |
| Vtg        | > average % vitellogenic follicles<br>No atretic follicles observed in section | From fish that did not release eggs upon applying gentle pressure to the abdomen         |

## 2.2. Histological analysis and sample classification

Samples stored in Davidson's fixative overnight were transferred to, and stored in, neutral buffered formalin until embedded. Embedding, sectioning, and hematoxylin and eosin (H&E) staining was conducted as described elsewhere [[http://www.epa.gov/endo/pubs/att-h\\_histopathologyguidelines\\_fhm.pdf](http://www.epa.gov/endo/pubs/att-h_histopathologyguidelines_fhm.pdf)]. Slides were arranged in random order by personnel not involved in the histological examination, and scored blindly without reference to sampling groups. Numbers of perinucleolar, cortical alveolar, early vitellogenic, late vitellogenic, atretic, and post-ovulatory oocytes/follicles [[http://www.epa.gov/endo/pubs/att-h\\_histopathologyguidelines\\_fhm.pdf](http://www.epa.gov/endo/pubs/att-h_histopathologyguidelines_fhm.pdf)] were counted. Initially, three sections of each ovary sample were examined; however, after scoring approximately 20% of the sample set in that manner, it was determined that data for a single section was representative of the three sections mounted and stained. Therefore, counts were performed on a single section for the remainder (approximately 80%) of the samples. After counts were completed, the total number of oocytes/follicles was summed and the percent of the total in each class calculated. Quantitative histology data were used to calculate mean percentages of previtellogenic follicles (including perinucleolar and cortical alveolar stage), vitellogenic follicles (including early and late vitellogenic), atretic follicles, and post-ovulatory follicles, across all samples examined. The mean values were used to help set quantitative criteria for classifying the samples. In addition to the quantitative data, a qualitative stage was assigned to each sample based on criteria described elsewhere [[http://www.epa.gov/endo/pubs/att-h\\_histopathologyguidelines\\_fhm.pdf](http://www.epa.gov/endo/pubs/att-h_histopathologyguidelines_fhm.pdf)]. Stage and several other qualitative criteria were considered when classifying the samples. Only samples that met both the quantitative and qualitative criteria defined (Table 1) were used for QPCR and/or microarray analysis (Supplemental Table S.1).

## 2.3. RNA extraction

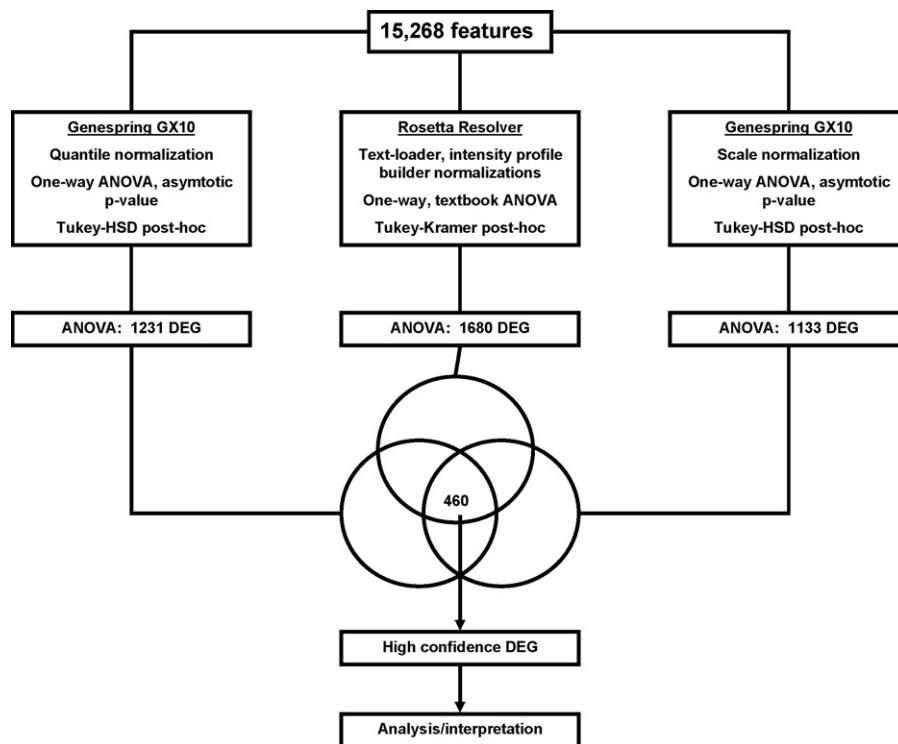
Based on histological classification, RNA was extracted from eight samples (six in the case of the previtellogenic [prevgtg] class) fitting the criteria for each of the four classes considered (Table 1; Supplemental Table S.1). Total RNA was isolated from the ovary samples using RNeasy kits (Qiagen, Valencia, CA, USA). RNA quality was assessed with an Agilent 2100 Bioanalyzer (Agilent, Wilmington, DE, USA) and quantity was determined using a Nanodrop® ND-1000 spectrophotometer (Nanodrop Technologies, Wilmington, DE, USA). Total RNA was stored at –80 °C until analyzed using QPCR or oligonucleotide microarrays.

## 2.4. QPCR analyses

Relative abundance of transcripts coding for ten proteins known to play important roles in the regulation of gonad development and reproduction was evaluated in samples representing four different histoclasses (Table 1). All QPCR assays were performed in duplicate using Taqman EZ RT-PCR kits (Applied Biosystems, Foster City, CA, USA). Each 12 µl reaction contained 20 ng total RNA, 250 nM of a gene-specific, dual-labeled Taqman probe (5'-FAM™ [6-carboxyl-fluorescein], 3'-Black Hole Quencher-1®; Integrated DNA Technologies, Coralville, IA, USA), 250 nM forward primer, and 250 nM reverse primer. Samples were reverse transcribed (50 °C for 2 min, 60 °C for 30 min, 95 °C for 5 min) followed immediately by 40 cycles of PCR amplification (melt 94 °C for 20 s, anneal and extend 58 °C for 60 s) using a 7500 Real-Time PCR System (Applied Biosystems). A standard curve of known quantities of a gene-specific mRNA standard (10-fold dilution series; generally 20–2 × 10<sup>7</sup> copies) was used to calibrate the QPCR data and estimate the number of gene-specific transcripts per ng total RNA. Details regarding the specific primer and probe sequences used and the methods for preparing gene-specific mRNA standards have been reported previously (Villeneuve et al., 2006, 2007, 2009a; Martinovic et al., 2008). Due to potential differences in amplification and probe binding efficiencies between total RNA samples and purified standards, the number of copies per ng total RNA is regarded as an approximate estimate of the number of transcripts. Therefore, data are reported as relative number of copies, rather than absolute quantities. Mean intraassay CVs for duplicate measurements of the same sample were less than 10% for all genes examined, and all samples were run on a single 96 well plate, such that interassay variability was not relevant to the study. Most QPCR data conformed to parametric assumptions, as tested using a Kolmogorov–Smirnov test for normality and Bartlett's test for homogeneity of variance, and were analyzed by one-way analysis of variance (ANOVA) followed by a Tukey–Kramer post hoc test, *p* < 0.05. However, data for luteinizing hormone receptor (LHR) and cytochrome P45011a (CYP11A) transcripts, which did not conform to parametric assumptions, were analyzed by a non-parametric Kruskal–Wallis test followed by Dunn's post hoc test.

## 2.5. Microarray hybridizations

Relative abundance of approximately 15,000 RNA transcripts in 26 ovary samples representing the four different histoclasses defined for the present study, was evaluated using fathead minnow oligonucleotide microarrays. Fathead minnow 15,000 gene microarrays (GEO Platform Accession GPL9248) were purchased from Agilent Technologies (Palo Alto, CA, USA). The Agilent one-color microarray hybridization protocol (One-Color Microarray-Based Gene Expression Analysis, version 5.7, Agilent) was used for microarray hybridizations following the manufacturer's protocol and recommendations. One microgram of total RNA was used for all hybridizations. Complementary DNA synthesis, cRNA labeling, amplification, and hybridization were performed following the manufacturer's kits and protocols (Quick Amp Labeling kit; Agilent). An Axon GenePix® 4000B Microarray Scanner (Molecular Devices Inc., Concord, Ontario, Canada) was used to scan microarray images at 5 µm resolution. Data were resolved from microarray images using Agilent Feature Extraction software (Agilent Technologies, Palo Alto, CA). Consistent with the Minimum Information About a Microarray Experiment (MIAME) standards (Brazma et al., 2001), text versions of the Agilent raw data from this study have been deposited to the Gene Expression Omnibus (GEO: <http://www.ncbi.nlm.nih.gov/geo/>; Accession series record number GSE18254).



**Fig. 1.** Flow chart summarizing a data analysis approach that employed the consensus among three microarray data analysis pipelines to identify the set of high confidence differentially expressed genes (DEG) considered in the present study.

## 2.6. Microarray data analysis

Data normalization in microarray experiments can have significant effects on the detection of differentially expressed genes (Hoffmann et al., 2002; Steinhoff and Vingron, 2006). Different data analysis software allow for implementation of many different normalizations, several of which may be appropriate for a given data set, including the single-color data set evaluated in the present study. Consequently, in the absence of a means for objectively selecting the optimum normalization for our data set, we normalized the data using three different algorithms, implemented in two different software programs (Fig. 1). The normalization techniques included scale normalization and quantile normalization, implemented in Genespring GX10 (Agilent) and default text-loader, intensity profile builder, normalization implemented in Rosetta Resolver 7.2 (Rosetta Biosoftware, Seattle, WA, USA). Normalized data were then analyzed by one-way ANOVA (textbook option in Resolver; asymptotic  $p$ -value option in Genespring) followed by a Tukey-HSD (Genespring) or Tukey-Kramer (Resolver) post hoc test (Fig. 1). The statistical threshold for both the ANOVA and post hoc test was set to  $p < 0.05$ . No corrections for multiple testing were applied. Only those genes identified as differentially expressed using all three data analysis pipelines were considered high confidence differentially expressed genes (DEG) and subjected to additional analysis and interpretation (Fig. 1). Determination of relative up- or down-regulation and fold change in expression between histoclasses was based on the gene-specific processed intensity data for each class (mean  $\pm$  error). Processed intensity data came directly from the Agilent feature extracted text files, without further scaling or normalization. Means and errors across replicates within each histoclass were calculated in Rosetta Resolver.

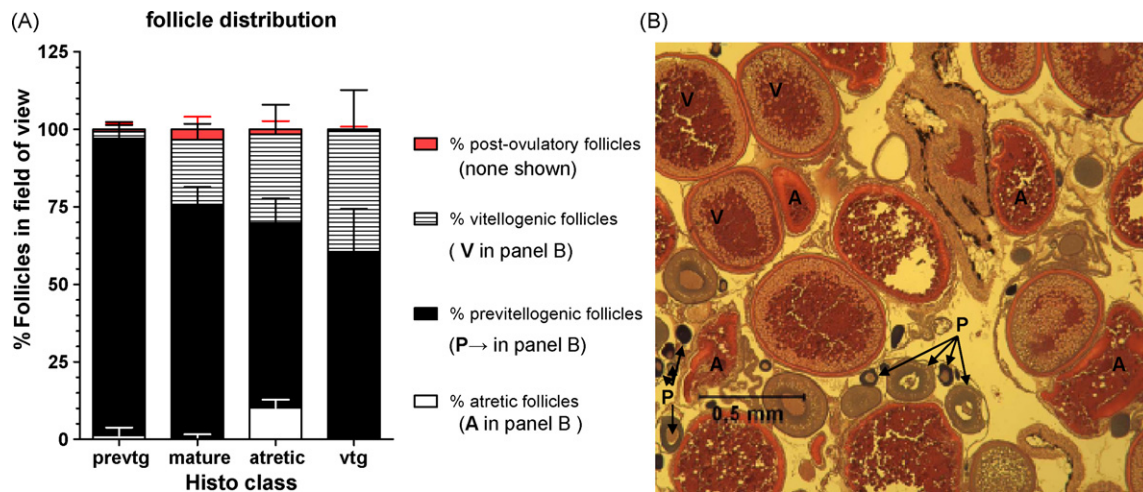
The overall similarity and differences in the expression profile of the high confidence differentially expressed genes among histoclasses was examined using two-dimensional agglomerative

and divisive clustering, as well as principal components analysis (PCA), implemented in Rosetta Resolver. Both agglomerative and divisive clustering algorithms were implemented using cosine correlation as a similarity measure, weighted by error. The heuristic criterion for agglomerative clustering was “average link”. Analysis parameters for PCA were 95% of variation captured in three principal components. Both clustering and PCA were conducted on the Rosetta Resolver normalized intensity data converted to Z scores (see Rosetta Resolver Plug-in Pipeline and Statistical Reference manual for details), using only those features (sequences) identified as high confidence DEG.

Over-representation of functional annotations like gene ontology (GO) categories, or biological pathways (e.g., Kyoto Encyclopedia of Genes and Genomes [KEGG] pathways; Kanehisa et al., 2008) within various lists of differentially expressed genes was examined using functional annotation clustering within DAVID (<http://david.abcc.ncifcrf.gov/home.jsp>; Huang et al., 2009). Accession numbers associated with all features on the array were submitted and those that mapped to DAVID were used as the background list for the analyses. In general, approximately 50–60% of the submitted accession numbers were successfully mapped to homologous features in the DAVID database. Analyses were conducted using the default selections and medium classification stringency. Enrichment analyses and clustering utilized DAVID's default EASE score, a modified Fisher's Exact  $p$ -value, of 0.1 as a statistical cutoff and a count threshold of 2 (the minimum number of genes belonging to any given annotation term considered). However, for the purposes of this analysis, only those clusters containing terms with  $p$ -values less than 0.05 were considered, and only terms with  $p < 0.05$  are shown.

High confidence DEG identified in this study were compared with a list of genes identified as differentially expressed in the ovaries of female zebrafish (*Danio rerio*) exposed to the aromatase inhibitor fadrozole (25  $\mu\text{g/L}$ ) for 48 or 96 h (Villeneuve et al., 2009a,b). Homologous features between the fathead minnow





**Fig. 2.** (A) Average proportions of atretic, previtellogenic, vitellogenic, and post-ovulatory follicles represented in ovary samples assigned to the prevtg, mature, atretic, and vtg histological classifications (Histo class; see Table 1), and (B) a representative micrograph illustrating examples of atretic, previtellogenic, and vitellogenic follicles. Error bars indicate standard deviation.  $n = 8$  except for prevtg  $n = 6$ .

microarray used in the present study and the Agilent zebrafish microarray (product 015064) used in the previous work were identified using a reciprocal best hit approach (Moreno-Hagelsieb and Latimer, 2008). In total, 9171 homologous features (not all unique) were identified. The list of homologous features was used to convert the differentially expressed gene lists from Villeneuve et al. (2009b) into lists of corresponding fathead minnow microarray features. This facilitated the cross-species comparison of results from the two experiments.

### 3. Results

#### 3.1. Histological scoring and staging

Ovary tissue from 74 of the 76 fish sampled was examined histologically. The total number of oocytes/follicles present in the sections examined ranged from 20 to 212, with 97 and 94 being the mean and median number of follicles per section, respectively. Averaged ( $\pm$ SD) across all samples, previtellogenic oocytes accounted for  $74.7 \pm 13.4\%$  of the total number of oocytes present, while vitellogenic, atretic, and post-ovulatory oocytes/follicles accounted for  $21.7 \pm 12.4$ ,  $2.5 \pm 3.4$ , and  $1.1 \pm 1.9\%$ , respectively. Neither temperature nor photoperiod manipulations had a significant effect on the distribution of different classes of oocytes/follicles in the ovary tissue (data not shown). Therefore, these pre-sampling manipulations were disregarded and all sample selection and classification was based on quantitative histological criteria and other qualitative factors (Table 1). Following sample classification and selection, the average distribution of oocytes/follicles in the prevtg histoclass was  $95.7 \pm 4.3\%$  previtellogenic and  $2.2 \pm 3.1\%$  vitellogenic (Supplemental Table S.1; Fig. 2) with relatively few atretic or post-ovulatory follicles. The mature histoclass contained, on average ( $\pm$ SD),  $73.1 \pm 2.6\%$  previtellogenic oocytes/follicles and  $22.6 \pm 4.4$  vitellogenic, with below average percentages of atretic follicles and above average percentages of post-ovulatory follicles (Supplemental Table S.1, Fig. 2). In the atretic histoclass, an average of  $10.9 \pm 2.3\%$  of follicles were atretic, while vitellogenic and previtellogenic follicles accounted for  $30.0 \pm 8.6$  and  $57.9 \pm 5.2\%$ , respectively (Supplemental Table S.1, Fig. 2). While the vitellogenic (vtg) histoclass had the greatest proportion of vitellogenic follicles,  $40.5 \pm 13.7\%$  on average, this class still contained  $58.8 \pm 14.2\%$  previtellogenic oocytes, by number. On an area basis, vitellogenic follicles occupied the majority of the area of the section, with exception of samples representing the prevtg

class. Gross ovarian stage, which is based on relative area occupied by follicles of different stages [<http://www.epa.gov/endo/pubs/att-h.histopathologyguidelines.fhm.pdf>] ranged from 0 to 2 for samples in the prevtg class, 3 to 4 for those in the atretic class, 2 to 4.5 for the vtg class, and 2 to 3.5 for those in the mature class (Supplemental Table S.1).

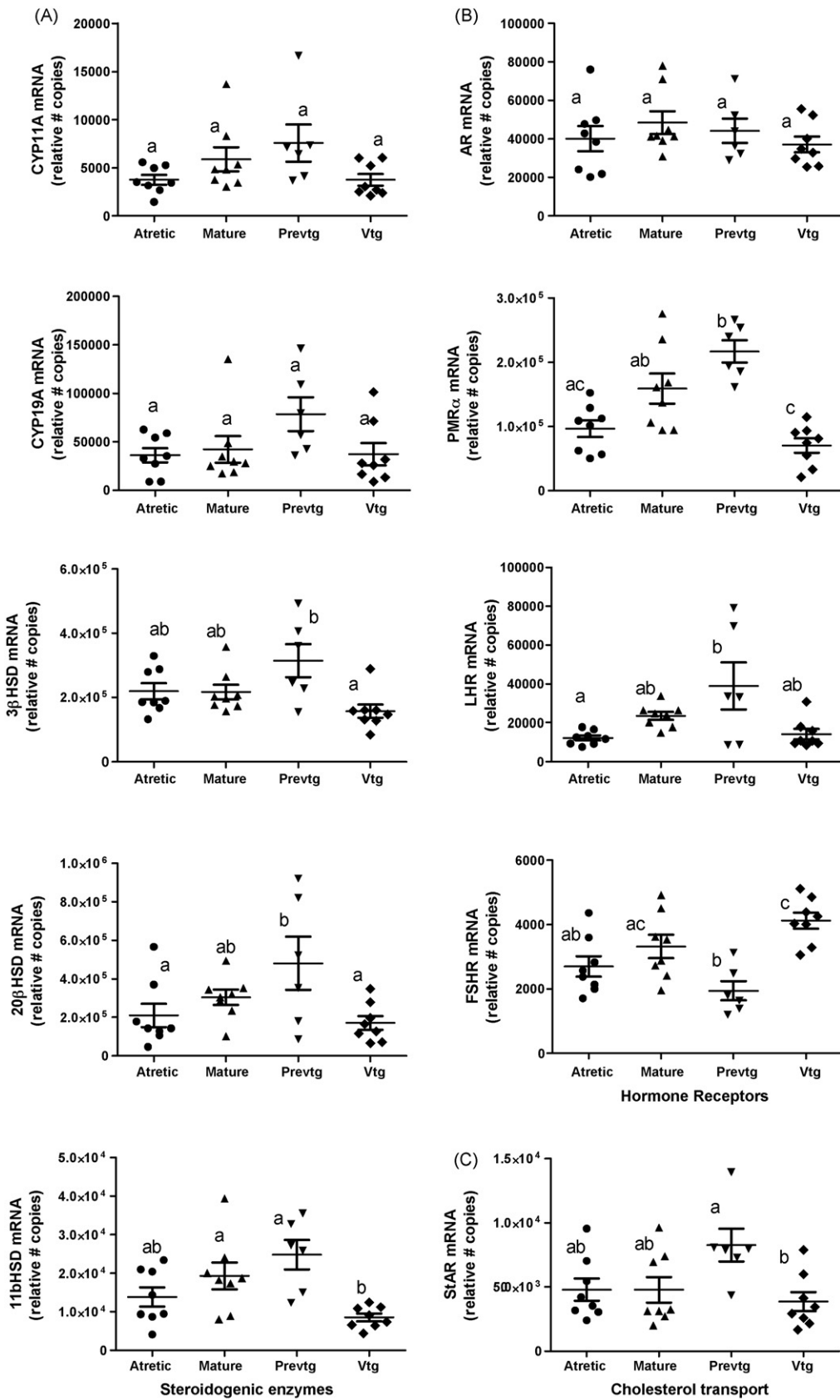
#### 3.2. QPCR

The relative abundance of transcripts coding for five steroidogenic enzymes, including cytochrome P450 cholesterol side chain cleavage (CYP11A), aromatase (CYP19A1a), and hydroxysteroid dehydrogenases (HSD)  $3\beta$ -,  $11\beta$ -, and  $20\beta$ , and the cholesterol transporter StAR, among the four histoclasses was examined (Fig. 3A and C). Mean expression of all five genes coding for steroidogenic enzymes was greatest in the prevtg class samples (Fig. 3A). In the case of  $3\beta$ HSD,  $20\beta$ HSD, and  $11\beta$ HSD, statistically significant differences between the prevtg and vtg classes were detected (Fig. 3A). The same was true for StAR where a significant difference between the prevtg and vtg classes was detected, but neither differed significantly from either the mature or atretic class (Fig. 3C).

Expression of mRNA transcripts coding for four important hormone receptors was also examined (Fig. 3B). The abundance of transcripts coding for fathead minnow androgen receptor (AR) did not vary significantly as a function of histoclass. However, expression of progesterin membrane receptor  $\alpha$  (PMR $\alpha$ ) was significantly elevated in the prevtg histoclass compared to the vtg and atretic classes. Expression of LHR was significantly different only between the prevtg and atretic classes. As for  $20\beta$  HSD, variability in LHR expression was particularly high among individuals within the prevtg class. Expression of follicle stimulating hormone receptor (FSHR) was unusual compared to the other genes analyzed by QPCR. The greatest FSHR expression was in the vtg class, which was significantly elevated compared to the atretic and prevtg classes.

#### 3.3. Microarray

Fathead minnow oligonucleotide microarrays containing 15,268 features were used to conduct a more global, unsupervised, analysis of differences in transcript profiles among ovary samples representing four different histoclasses: prevtg, mature, atretic, and vtg (Table 1). The three data normalization and analysis approaches applied to the data identified between 1133 and 1680



**Fig. 3.** Real-time quantitative PCR results for genes coding for (A) five steroidogenic enzymes including cytochrome P450 (CYP) 11A, CYP19A (aromatase; ovarian isoform), and 3 $\beta$ -, 20 $\beta$ -, and 11 $\beta$ -hydroxysteroid dehydrogenase (HSD); (B) four hormone receptors including androgen receptor (AR), progestin membrane receptor  $\alpha$  (PMR $\alpha$ ), luteinizing hormone receptor (LHR), and follicle-stimulating hormone receptor (FSHR); and (C) steroidogenic acute regulatory protein (Star). Points indicate relative abundance determined for each individual fish. Horizontal line indicates the mean for each group, error bars indicate the standard error. Different letters indicate statistically significant differences between groups. Sample size = 8, except for prevgt class (n = 6).

**Table 2**

Number of genes confidently identified as differentially expressed between each binary pair of histological classifications (histoclasses), based on the union of Tukey post hoc test results from three analysis pipelines (see Fig. 1). Numbers in parentheses indicate the number of genes up-regulated and down-regulated, respectively, in the row class (italicized) compared to the column class.

| Histoclass            | Mature    | Atretic     | Vtg           |
|-----------------------|-----------|-------------|---------------|
| <i>Prevgt</i> (>, <)  | 73(3, 70) | 75 (47, 28) | 226 (36, 190) |
| <i>Mature</i> (>, <)  |           | 20 (15, 5)  | 41 (26, 15)   |
| <i>Atretic</i> (>, <) |           |             | 15 (1, 14)    |

features as being significantly differentially expressed among the four classes (Fig. 1). However, when only the intersection of all three analysis pipelines was considered, only 460 genes (27–40% of those identified by any single approach) were identified as high confidence DEG (Fig. 1). The number of genes confidently identified as differentially expressed between any two of the histoclasses, based on the intersection of Tukey post hoc results from the three analysis pipelines, ranged from a minimum of 15 between the atretic and vtg group, to a maximum of 226 between the prevtg and vtg group (Table 2). The prevtg group had the largest number of DEG compared to each other histoclass (Table 2). Overall, the results suggest there were differences in gene expression between classes, and that the prevtg was the most distinct from the other groups.

The similarities and differences in the expression profiles of the 460 high confidence DEGs between histoclasses were also evaluated using a combination of cluster analyses and PCA. Agglomerative hierarchical clustering, which uses pair-wise similarity measures to identify the most similar objects (Rosetta Resolver, Analysis Guide), showed the prevtg group and mature group to be the most similar, with the vtg and atretic group clustering together (Supplementary Figure S.1A). Divisive hierarchical clustering, which takes all the objects into one cluster and then splits them by optimal partitioning (Rosetta Resolver, Analysis Guide) such that groups most different from one another branch first, showed the vtg and prevtg classes to be the most different from one another (Supplementary Figure S.1B). PCA showed clear separation of the vtg and prevtg class along principal component (PC) 1 (Fig. 4). The mature class was intermediate between the vtg and prevtg class along PC1, and was clearly separated from the prevtg

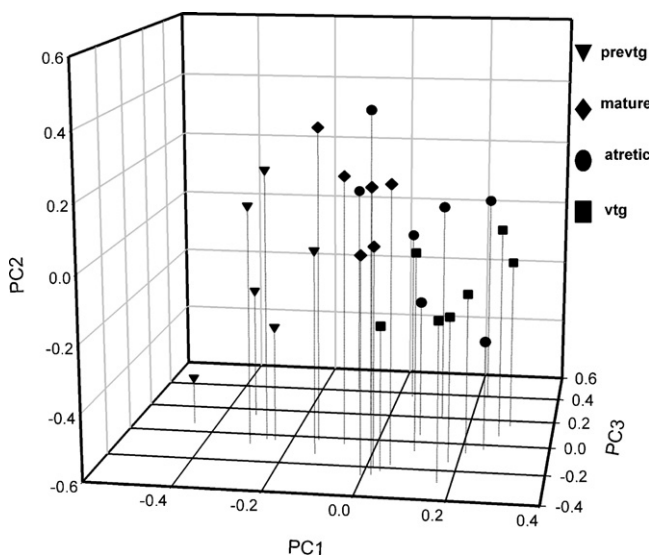
and vtg classes along PC1 and PC2 (Fig. 4). Conversely, the atretic class was largely intermingled among the three other classes along PC1 and PC2 and only separated well along PC3 (Fig. 4). Taken together, the results provide a consistent impression of the degree of similarity/difference in the expression profile of the 460 DEG among classes, which would order as prevtg-mature-atretic-vtg.

Over-representation of differentially expressed genes within GO categories, or biological pathways (e.g., KEGG pathways; Kanehisa et al., 2008) was examined using functional annotation clustering within DAVID (Huang et al., 2009). As a first level of analysis, GO categories and pathways over-represented in the list of 460 high confidence DEG were evaluated (Supplemental Table S.2). Of the 460 features loaded (as Accession numbers) into DAVID, 270 (58.6%) were successfully mapped to the DAVID database. Significantly enriched functional annotation clusters included those associated with metal/cation binding, cellular development/morphogenesis/neuron morphogenesis, coated vesicles, ubiquitin cycle, metal/cation transport, sexual reproduction, wd repeat domain, coiled coil/cytoskeleton, and several others (Supplemental Table S.2). Functional annotation clustering in DAVID was also conducted using the 190 genes identified as more highly expressed in the vtg class compared to the prevtg class (Table 2; Supplemental Table S.3). Most of the enriched clusters were similar to those identified for the 460 gene set (e.g., metal/cation binding; development/morphogenesis/neuron development; phosphoric ester hydrolase activity), but fewer significant clusters were detected (as is typical when fewer input genes are provided). For the list of 36 (25/36; 69.4% mapped to DAVID) genes more highly expressed in the prevtg class, compared to the vtg class, there were no significantly enriched ( $p < 0.05$ ) functional annotation terms. Given the small numbers of genes up or down-regulated for the other binary class comparisons, additional functional annotation clustering analyses were not performed.

The intersection of the three sets of Tukey test results for each histoclass (i.e., binary class comparisons) was examined to identify genes with potential utility as class/stage specific markers. There were no genes common to all three comparisons between the vtg or atretic class and the other three classes. However, for the prevtg class, three genes, RNA binding motif 39b (*rbm39b*), BX293562 (unknown), and *c1galt1c1* (C1 galactosyl transferase 1), were commonly identified as differentially expressed compared to all other classes. For the mature class, two genes, similar to HtrA serine peptidase 3 (*htra3*) and LOC100001650 (unknown) were identified as differentially expressed relative to all other classes.

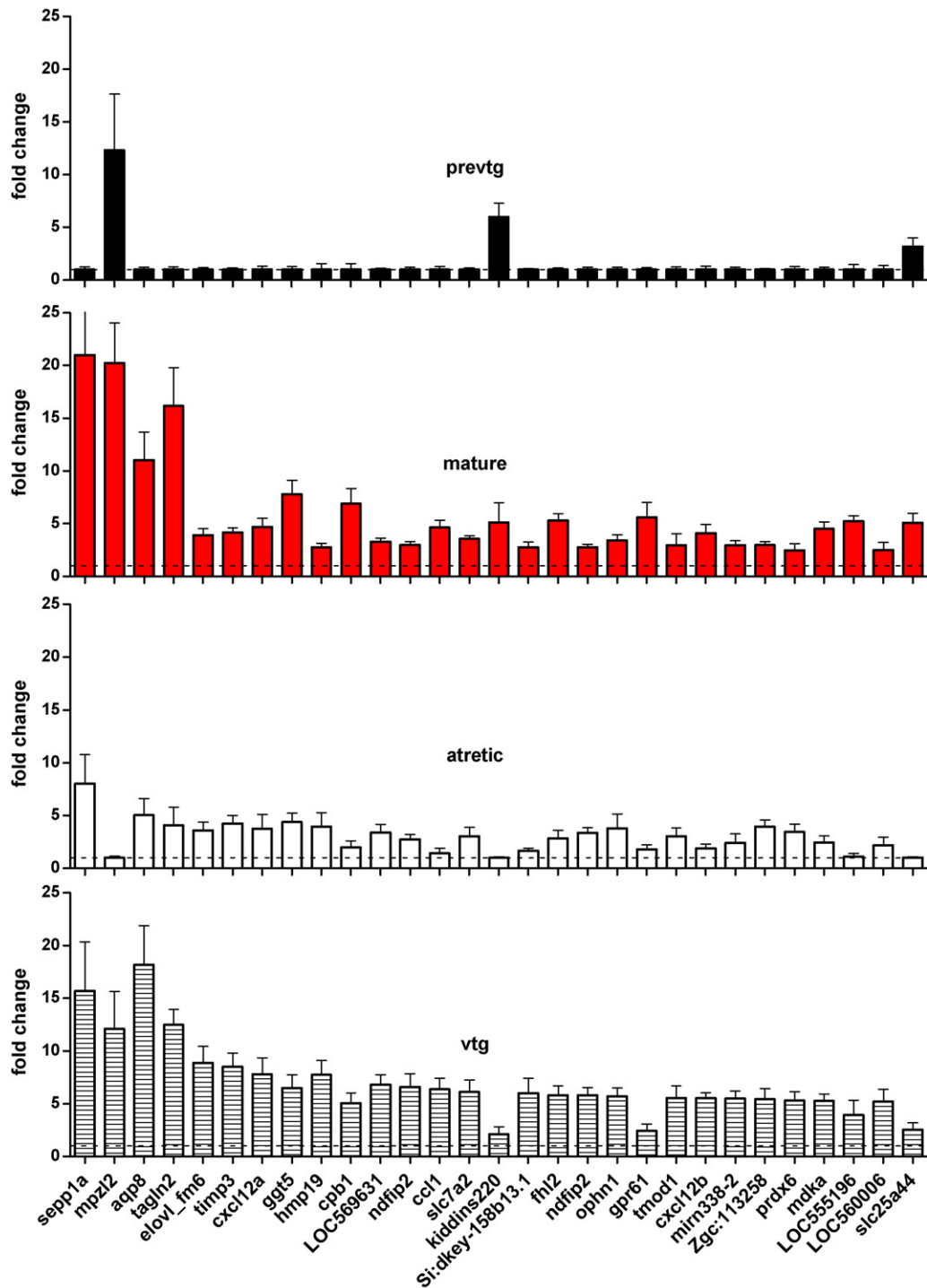
In terms of fold-change, the greatest difference among any of the two classes was observed for selenoprotein P (*sepp1a*), which had a 21-fold greater intensity in the mature class than the prevtg class (Fig. 5). Only three other genes had greater than 10-fold differences in intensity between histoclasses. Aquaporin 8 (*aqp8*; zgc:158752) was expressed 18.2-fold more intensely in the vtg class than in the pre-vtg class and 11- and 5-fold more intensely in the mature and atretic classes, respectively, than in the prevtg class, although the latter differences were not statistically significant. Transgelin 2 like (*tagln2*; zgc:153186) was expressed 16.2- and 12.5-fold more intensely in the mature and vtg groups, respectively, than in the prevtg group. Finally, myelin protein zero like 2 (*mpzl2*) was expressed 20-, 12- and 12-fold more intensely in mature, vtg, and the prevtg groups, respectively, than in the atretic group. Twenty-five genes had 5–10-fold differences in intensity between at least two classes (Fig. 5), with the bulk of those (19/25) being genes detected as significantly more highly expressed in the vtg class compared to the prevtg class.

Genes with the greatest fold change difference between two or more histoclasses were not necessarily those detected as significantly different with the greatest confidence (i.e., lowest  $p$ -value). Only three of the genes with fold change of five or greater between



**Fig. 4.** Principal components (PC) analysis of 460 genes identified as significantly differentially expressed as a function of histological classification (histoclass; see Table 1). Each point represents a separate biological replicate within each histoclass. Histoclasses are differentiated by shape.



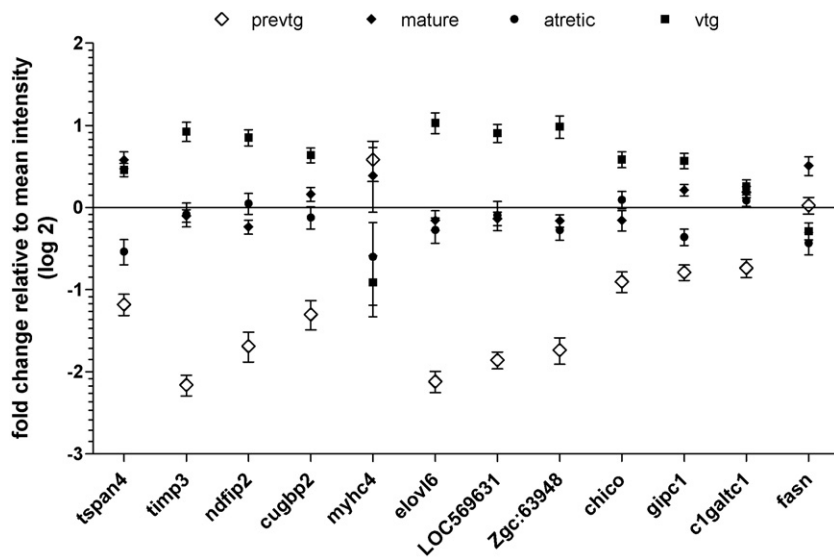


**Fig. 5.** Genes with a 5-fold or greater difference between two or more histoclasses (see Table 1). Fold-change is expressed relative to the minimum average intensity for any one histoclass. Error bars indicate the standard error among biological replicates within the class. Abbreviations are established gene symbols.

two or more classes, tissue inhibitor of metalloproteinase 3 (*timp3*), Nedd4 family interacting protein 2 (*ndfip2*) and elongation of long chain fatty acids family member 6 (*elovl6*), were among the dozen genes with the lowest overall ANOVA *p*-values (averaged across all three analysis pipelines; Fig. 6). The other 9 genes identified as differentially expressed with greatest statistical confidence were tetraspanin 4 (*tspan4*), CUG triplet repeat RNA binding protein 2 (*cugbp2*), myosin heavy chain 4 (*myhc4*), *chico*, GIPC PDZ domain containing family member 1 (*gipc1*), *c1galt1c1*, fatty acid synthase (*fasn*), and two sequences without known orthologs, LOC56963

and *zgc:63948*. Genes in this group had average ANOVA *p*-values ranging from 0.000006 to 0.0003. Among this group of 12 genes, expression was lowest in the prevtg class for 10 of the 12 (Fig. 6). The exceptions were *myhc4* which had greatest expression in the prevtg and vtg classes and lesser expression in the other two, and *fasn* which was most highly expressed in the vtg class, intermediate in the prevtg class, and least expressed in the atretic class (Fig. 6).

Finally, in a previous study, Villeneuve et al. (2009b) hypothesized that the profile of genes differentially expressed in the ovaries of zebrafish exposed to the aromatase inhibitor fadrozole were, at



**Fig. 6.** Twelve genes with the lowest average (among 3 analysis pipelines; see Fig. 1) ANOVA *p*-values among those identified as differentially expressed as a function of histological classification (histoclass; see Table 1). Relative expression intensity presented as fold change relative to the mean intensity across all groups, expressed on a log 2 scale. Error bars indicate standard error for each histoclass. Abbreviations are established gene symbols.

least in part, reflective of a deterred follicle development as a result of impaired vitellogenesis. Thus, comparison between the lists of differentially expressed genes identified in the present study with those from the previous zebrafish study was also used to identify potential transcriptional markers of interest. Eleven features differentially expressed as a function of both ovary stage (present study) and fadrozole exposure (Villeneuve et al., 2009b) were identified (Table 3). Of those, nine were significantly differentially expressed between the prevtg and vtg class.

## 4. Discussion

### 4.1. Primary source of transcriptional variability

Results of this study support the conclusion that gross differences in the relative proportions of previtellogenic versus vitellogenic follicles in an ovary sample from a sexually differentiated, asynchronous-spawning, fish will likely be the major determinant of background variability in the ovarian transcript profile. This conclusion was based on multiple lines of evidence. First, considering the 460 genes confidently identified as differen-

tially expressed between histoclasses, cluster analysis indicated the greatest difference between prevtg and vtg, with the greatest similarity between prevtg and mature and between vtg and atretic (Supplementary Figure S.1). Principal components analysis supported that ordering (Fig. 4), as did Tukey test results which found the greatest number of DEG when comparing prevtg and vtg and the fewest when comparing vtg and atretic (Table 2). Even when only a small subset of genes was considered (Figs. 3, 5 and 6), prevtg was the class most often different from the others, with mature being different from prevtg less frequently than either atretic or vtg. On average, ovaries assigned to the prevtg class were from younger fish (i.e., mostly 4-month old) than those from other classes. Thus, we cannot rule out the possibility that age was a co-factor that influenced the distinction of the prevtg class from the other histoclasses. Nonetheless, the consistent ordering based on similarity/difference of transcript-level response directly corresponded with ordering based on relative numbers of either previtellogenic follicles or vitellogenic follicles observed through histological analysis (Fig. 2). Thus, while there were exceptions for individual genes, overall the ratio of previtellogenic to vitellogenic follicles appeared to significantly influence overall differences in transcript profiles.

**Table 3**  
Genes identified as differentially expressed in the ovaries of zebrafish (ZF) exposed to the aromatase inhibitor fadrozole (Villeneuve et al., 2009b) that were also differentially expressed among the histological classifications identified in the present fathead minnow (FHM) study. With the exception of entries in italics, all were identified as differentially expressed between the previtellogenic and vitellogenic histological classes.

| FHM sequence code       | Homologous ZF probe id(s) | Gene symbol  | Accession #      | Sequence description                                   |
|-------------------------|---------------------------|--------------|------------------|--|
| UF_Ppr_AF_119042        | A.15.P115553              | fam110b      | NM_001025174     | Family with sequence similarity 110, member B          |
| UF_Ppr_AF_117752        | A.15.P110961              | tmc6         | NM_001002705     | Transmembrane channel-like 6                           |
| UF_Ppr_AF_116150        | A.15.P120003              | zgc:158236   | NM_001080590     | WAS protein family, member 3                           |
| UF_Ppr_AF_115695        | A.15.P113017              | hmp19        | BC063320         | HMP19 protein  |
| <i>UF_Ppr_AF_115424</i> | <i>A.15.P117619</i>       | <i>hhex</i>  | <i>NM_130934</i> | <i>Hematopoietically expressed homeobox</i>            |
| UF_Ppr_AF_112884        | A.15.P116901              | ndfip2       | NM_213347        | Nedd4 family interacting protein 2                     |
| UF_Ppr_AF_111217        | A.15.P109272              | ppp1cb       | BC095825         | Protein phosphatase 1, catalytic subunit, beta isoform |
| UF_Ppr_AF_108285        | A.15.P108110              | mdka         | NM_131070        | Midkine-related growth factor                          |
|                         | A.15.P112924              |              |                  |  |
| UF_Ppr_AF_107108        | A.15.P111152              | cxcl2a       | AJ627274         | Chemokine ligand 12a                                   |
|                         | A.15.P113844              |              |                  |  |
| UF_Ppr_AF_102557        | A.15.P101691              | jag1b        | NM_131863        | Jagged 1   |
|                         | A.15.P116455              |              |                  |  |
| <i>UF_Ppr_AF_101394</i> | <i>A.15.P104600</i>       | <i>krt18</i> | <i>AF537368</i>  | <i>Keratin 18</i>                                      |
|                         | <i>A.15.P108350</i>       |              |                  |  |
|                         | <i>A.15.P120999</i>       |              |                  |  |

This conclusion has several implications relative to the use of ovary tissue from asynchronous-spawning fish species in ecotoxicogenomic studies. Histological examination in the present study demonstrates that previtellogenic and vitellogenic follicles constitute the majority of follicles in the ovary. Even for fish that have ovulated or with significantly increased incidence of follicle atresia, the bulk of follicles would still be classified as previtellogenic or vitellogenic. This suggests that chemical impacts on vitellogenesis and associated ovarian signaling processes may have a more profound (or at least more discernible) effect on ovarian transcript profiles than impacts on ovulation or modest increases in oocyte atresia. Our results also suggest that the degree of difference in transcript profiles is likely to be dependent on the degree to which relative proportions of vitellogenic versus previtellogenic follicles are shifted by a chemical treatment or other stressor. For example, in the present study, although both mature and atretic classes had significant numbers of vitellogenic follicles, the mature class, with proportionally fewer vitellogenic follicles was detected as being more different from the vtg class than the atretic class. Thus, there may be merit in identifying specific targets that most sensitively differentiate the prevtg and vtg classes, in terms of statistical significance or fold-change, to develop as potential markers of chemically induced alterations in gross ovarian stage. Finally, even when using fish of known, uniform, age cultured under the same conditions, it is not unusual for a few fish with immature ovaries to be inadvertently included in experiments with reproductively mature fish, particularly when the total number of fish included in the experiment is large. Results from this study suggest that samples from immature fish should be excluded as outliers among sample populations in which most of the ovary samples are collected from fish with reproductively mature ovaries containing significant numbers of vitellogenic oocytes. At least for molecular endpoints, the variability that such samples would introduce into a relatively small sample population would be significant and would likely limit overall power to detect chemical effects. Thus, unless there is evidence to suggest that the immature status is caused by the stressor of interest, it would seem both prudent and scientifically sound to exclude immature animals as outliers.

#### 4.2. Functional analysis

Functional analysis was employed to gain a sense of what biological processes may be involved in the process of asynchronous ovarian development. A number of the enriched annotation clusters associated with the high confidence DEG could reasonably be related to changes that occur over the course of oocyte/follicle maturation, particularly the transition from a previtellogenic to vitellogenic state. For example, cluster 2 for both the overall list of DEG, and those specifically identified as more highly expressed in the vtg class than in the prevtg class relate to morphogenesis, developmental processes, etc. (Tables S.2, S.3). This group of terms was, perhaps, not surprising given that oogenesis is a highly dynamic process involving structural and functional changes that are wide and complex in scope (Lubzens et al., 2009). The abundance of terms associated with neuron/neurite development may reflect alterations in innervation of the ovarian tissue as follicles develop, or may suggest other roles for genes typically associated with nerve development. For example, a variety of neurotrophins and their receptors have been shown to be expressed in ovary and appear to play important roles in growth, differentiation, and cell–cell communication between ovarian cell types (Ojeda et al., 2000). The prominence of cell projection organization and morphogenesis among the functional terms in cluster 2 (Tables S.2, S.3) is an interesting observation as well, in light of the bidirectional microvilli and gap junctional contacts known to develop between the oocytes and adjacent follicular cells (Cerda et al., 1999). The enrichment of

functional annotations associated with membrane bound coated vesicles (Table S.1, cluster 3) may relate to the fact that vitellogenin uptake into the developing oocyte is mediated through membrane associated receptors localized to clathrin-coated pits (Lubzens et al., 2009). However, similar vesicular uptake processes are thought to function at synapses (Granseth et al., 2007) and in a variety of other cellular roles. The ubiquitin cycle (see cluster 4, Table S.2) is a fairly general protein degradation process known to play a role in degrading proteins critical for cell cycle regulation during the process of oocyte maturation in fish (Tokumoto, 1999). Three other DEG identified as orthologous to *hect*, *citron*, and *nanos* (Table S.4), constituted significant enrichment of genes involved in reproductive processes (cluster 6; Table S.2), which is certainly consistent with ovarian function.

Unfortunately, most of the other functional annotation clusters (e.g., clusters 1, 8, 11; Table S.2) resulting from analysis in DAVID were very general and, as such, not very informative. A number of genes such as those with HECT domains, *citron*, myc binding protein 2, *jagged 1b*, *nanos*, etc., were associated with multiple enriched clusters (Tables S.2–4, figures S.1–S.6). This suggests a potential role for these genes as key signaling molecules in the process of asynchronous ovarian development. However, one must be cautious in ascribing extra weight or importance to these genes. The association of these genes with multiple clusters may be as much a reflection of how well mammalian orthologs of these genes have been studied and annotated as their importance in regulating fish oogenesis. Targets which did not map to orthologs in the DAVID database may play roles that are just as important. For example, since the DAVID functional annotation tool was built based on mammalian literature there are no annotations or pathways within DAVID that specifically relate to the process of vitellogenesis, despite the critical role it plays in the reproductive physiology of oviparous vertebrates. Additionally, while the functional annotation clustering gives some impression of the putative functions of a number of the high confidence DEG identified in the present study, it has not necessarily been established that the orthologous targets perform the same functions in fish.

#### 4.3. Notable differentially expressed genes

Beyond generalized functional analysis (described above) it is not feasible to consider all high confidence DEG in the context of a single paper. However, with recognition of the probability of false positives in microarray experiments, more detailed consideration of individual gene responses can potentially provide additional insight into reproductive biology of the fish. Relative to ontology-based analyses, interpretation of the potential functional significance of individual DEG can draw from a broader range of literature and is not limited by the annotation, or lack thereof, of orthologous genes. Therefore, in terms of specific gene responses, we focused on those DEG with the greatest fold change between classes and those detected with the greatest statistical confidence (i.e., lowest ANOVA *p*-value) to look for clues as to some of the important processes/functions that may be involved in regulating ovarian development in an asynchronous spawner. We also identified several genes that were potentially characteristic markers of a specific histoclass.

Consideration of some of the genes with the largest fold-change between histoclasses provides insights into the diversity of pathways and processes involved in the coordination of oogenesis. For example, *sepp1a*, the gene with the largest fold change observed in the present study, codes for a protein thought to function primarily in selenium distribution which has been reported to carry up to approximately one third of plasma selenium in humans (Akesson et al., 1994; Renko et al., 2008). Deficiency of *sepp1a* has been linked to impaired spermatogenesis and infertility in male mammals (e.g.,

Renko et al., 2008). While a similar linkage has not been established for ovary, selenium has been reported to modulate granulosa cell proliferation and estradiol production (Basini and Tamanini, 2000). The large fold change in expression between the prevtg and mature classes suggests that *sepp1a* may play an important role in oogenesis and/or be involved in distributing Se, as a key micronutrient, to the oocytes. An aquaporin ortholog, similar to *aqp8* characterized in humans, was another DEG with a large fold-change. Aquaporins are a family of integral membrane proteins that facilitate transport of water across the cell membrane (King et al., 2004). In the marine fish species, gilthead sea bream (*Sparus aurata*), a specific aquaporin, *aqp10*, has been shown to play an important role in oocyte hydration (Fabra et al., 2006; Litman et al., 2009). However, *aqp8* is an ammonia-permeable aquaporin in humans (Litman et al., 2009), suggesting that the aquaporin differentially expressed in the fathead minnow ovary may play a somewhat different role. A third DEG with a large fold-change, *tagln2*, is known as a marker of differentiated smooth muscle (Assinder et al., 2009). Its 12–16-fold greater expression in the vtg and mature groups than in the prevtg group is notable in light of reports that theca cells undergo differentiation into smooth muscle-like cells during oocyte maturation (Jalabert and Szollosi, 1975). Additionally, other smooth muscle markers have been associated with cord-like structures extending along ovarian vasculature in several ectotherms, including zebrafish (Van Nassauw et al., 1991). It is hypothesized that these cord-like structures or differentiated theca cells may play a role in follicular contraction during ovulation (Jalabert and Szollosi, 1975; Van Nassauw et al., 1991). The large fold-change in *mpzl2* transcripts, which codes for an adhesion molecule recognized as one of the most abundant myelin proteins of the peripheral nervous system (Filbin et al., 1990), concurs with the enrichment of functional annotation terms related to neuron morphogenesis and development, even though it was not among the features that mapped to an orthologous gene in DAVID. Thus, the four genes with the largest fold change among classes in the present study have four very different functions which all could be plausibly linked to the dynamic process of oogenesis. Given their large magnitude of change between classes, these targets may be useful as markers of gross shifts in ovarian follicle composition.

While genes with large fold-change may be particularly attractive markers in terms of the potential magnitude of signal they produce, DEG identified with high statistical confidence may also be favorable in terms of their signal to noise. Among the dozen genes with the lowest ANOVA *p*-value across histoclasses, several have particularly favorable properties as potential indicators of ovarian status. Perhaps first among those is *timp3*, which had an 8.5-fold difference in expression between the vtg and prevtg histoclasses and an overall *p*-value of 0.000019. Functionally, matrix metalloproteinases are known to play a key role in the extracellular matrix remodeling that accompanies follicular growth, atresia, and/or ovulation (Smith et al., 2002). Tissue inhibitors of metalloproteinases, like TIMP3, modulate the activity matrix metalloproteinases (Smith et al., 2002). In rats, TIMP3 has been specifically implicated in regulation of fatty acid synthesis, steroidogenesis, and protein turnover in granulosa cells through siRNA modulation of its expression (Li and Curry, 2009). Thus, *timp3* has both favorable signal to noise characteristics and a potentially prominent role in oogenesis. Several other genes/features including *ndfip2*, *elovl6*, LOC569631, zgc:63948 had similarly favorable signal to noise characteristics but their functions could not be as tightly linked to oogenesis. The latter two features had no identifiable orthologs. Among the former two, *ndfip2* codes for a protein thought to activate the activity of HECT domain E3 ubiquitin ligases, and thus be involved in regulation of protein degradation (Mund and Pelham, 2009). While this feature did not map to DAVID, it lends further support for the observation that ubiquitin cycle and HECT-related genes were among those

differentially expressed as a function of ovarian follicular development. The protein coded by *elovl6* is presumably involved in fatty acid metabolism, but a more specific role in ovary development has not been reported. Proteins similar to that coded for by *tspan4* are thought to play an important role in the cellular adhesion process that mediates sperm-oocyte fusion and thus fertility in mammals (Hemler, 2003; Li et al., 2004). However, such a function as not been specifically ascribed to TSPAN4, to date. The other DEG with low *p*-values are thought to be involved in a variety of functions ranging from glucose transport (*cugbp2*), to fatty acid synthesis (*fasn*), to skeletal muscle contraction (*myhc4*), but have no known regulatory or specific functional role in the regulation of oogenesis. Thus, while they may have favorable signal to noise properties as potential markers of gross ovarian stage, their biological relevance in terms of a mechanistic linkage to female gamete development is less clear.

As a third approach to identifying specific genes of potential interest, it was postulated that genes differentially expressed in a given class (e.g., prevtg) versus the three other histoclasses would be reasonable candidates as class-specific marker transcripts. Based on this criterion, five putative class-specific marker genes were identified. However, only two were orthologous to genes with known (or at least hypothesized) functions. The first, *c1galt1c1* (also known as *cosmc*) codes for a chaperone protein that appears to be key to proper folding and function of a specific galactosyltransferase involved in the synthesis of O-glycan glycoproteins (Ju et al., 2008). Recent evidence suggests that this chaperone may play an important role in regulating protein glycosylation (Ju et al., 2008). Expression of *c1galt1c1* was 1.8–2.0-fold lower in the prevtg class than in the mature, atretic, or vtg class. However, the functional significance of this difference in transcript abundance is unclear. Htra serine peptidase 3, which was expressed 3–4-fold more abundantly in ovary tissue from the mature class than in other three classes, has also been reported to increase with the maturation of follicles in both rat and rhesus monkey ovaries (Bowden et al., 2008, 2009). Evidence suggests that *htra3* can act as an inhibitor of a number of proteins in the transforming growth factor  $\beta$  (TGF $\beta$ ) superfamily, many of which (e.g., various bone morphogenic proteins, growth differentiation factor 9, fibroblast growth factors, activin, inhibin, follistatin, etc.) are known to play a role in ovarian development (Tocharus et al., 2004; Knight and Glistner, 2006). Given its functional significance, and apparent conservation in both sequence and ovarian expression profile, *htra3* may warrant further investigation as a potentially relevant and useful marker of follicle maturation.

Among the 11 DEG that overlapped between the present study and a previous study with a chemical stressor (the aromatase inhibitor fadrozole) hypothesized to impact follicular development (Villeneuve et al., 2009b), three genes (*ndfip2*, *mdka*, and *cxc12a*) were among those with at least 5-fold differences in intensity between at least two histoclasses (Fig. 5). As noted earlier, *ndfip2* likely plays a role in regulation of ubiquitin cycle and protein degradation (Mund and Pelham, 2009). Regulatory roles for *cxc12a* (*cxl12a*; also known as stromal cell derived factor 1 [*sdf1*]) in mammalian reproduction have been suggested (Kryczek et al., 2005; Holt et al., 2006), however these ligands likely play a diversity of signaling roles within various tissues. Midkine has been purported to play a role in angiogenesis during follicle develop and its expression has been associated with granulosa cell proliferation and estradiol production (Hirota et al., 2007). Thus all three of the high fold-change genes that were also affected by fadrozole have putative regulatory roles in ovarian follicle development. Jagged 1b, one of the genes that featured prominently in the DAVID functional annotation clustering, was also among the DEG impacted by fadrozole (Villeneuve et al., 2009b). Jagged is involved in regulating cell–cell and cellular matrix interactions, cell differentiation, has been implicated in regulation of ovarian vascular development, and



is regulated by 17 $\beta$ -estradiol (Soares et al., 2004; Vorontchikhina et al., 2005). These genes may also be candidates for further investigation and possible development as molecular markers of gross ovarian stage and/or impacts of chemicals on follicle development.

Overall, *htra3*, *timp3*, *aqp8*, *tagln2*, *ndfip2*, *cxcl12a*, *mdka*, and *jag1b* suggested themselves as top candidates for further investigation and possible development as molecular markers of gross ovarian stage in the fathead minnow. As with any microarray study, given the large number of dependent variables examined, there is a reasonable likelihood that at least some of the individual DEGs identified are false positives, even among those with the largest fold change, lowest *p*-values, or consistency across multiple post hoc comparisons. Therefore, follow-up studies are needed to evaluate the reproducibility of the results reported here and to establish their potential utility and reliability as markers. A more detailed analysis of other individual DEG in the context of the scientific literature and/or comparison to other studies should prove a fertile ground for hypothesis generation to direct future experimentation regarding the role of various genes and proteins in oogenesis and the extent to which they are impacted or altered by chemical stressors.

## 5. Conclusions

As anticipated, gene expression varied significantly with the gross proportions of different-staged oocytes and follicles within the ovary of an asynchronous-spawning fish. Previtellogenic and vitellogenic oocytes/follicles made up the bulk of the follicular population within the ovary samples, and their relative proportion was concluded to be the major discernible driver of variation in the transcript profiles among samples. A number of candidate genes that might serve as potential markers of shifts in gross morphological stage were identified. The sensitivity and reproducibility of these responses, as a function of shifts in follicle proportions can be examined in future studies using targeted approaches like QPCR. The present study was less successful in identifying putative markers of either mature (ovulated) follicles, or follicular atresia. This was likely due to the fact that these were minority populations within the overall follicle distribution of the ovary samples. The sampling and structural characterization methods employed in this study were intentionally crude, and designed to represent sample collection and histological examination methods typical of ecotoxicological studies conducted with fathead minnow. More precise sampling and analytical methods such as laser capture microdissection, in situ hybridization, and/or advanced stereological characterization of ovarian structural domains, would likely be needed to accurately identify transcriptional changes associated with follicular maturation, ovulation, and/or atresia and understand their overall mechanistic role in oogenesis. However, the list of DEGs generated in this study provides a basis for the development of hypotheses regarding some of the genes/proteins that may be involved in regulating the process of vitellogenesis and oocyte growth in asynchronous-spawning teleosts that can be tested in future work. Finally, we demonstrated how the list of DEG generated in this study can be compared to lists of genes differentially expressed following chemical exposure. Overall, results of the present study can be employed to enhance interpretation of toxicogenomic data generated in experiments with small, asynchronous-spawning fish.

## Acknowledgements

The authors thank Lynn Escalon for technical assistance with microarray hybridizations. The microarrays used in the present study were designed and annotated by Dr. Nancy Denslow and Dr.

Li Liu of the University of Florida. Dr. Liu and colleagues also conducted reciprocal best hit analyses to identify homologous features between the fathead minnow microarray (GEO accession GPL9248) and the zebrafish microarray used by Villeneuve et al. (2009b). We thank Dr. Rodney Johnson and Dr. Chris Martyniuk for comments on an earlier version of this manuscript. This work was supported in part by the US EPA National Center for Computational Toxicology and the Army Environmental Quality Program of the US Army Corps of Engineers. The manuscript has been reviewed in accordance with US EPA and US ACE guidelines and approved for publication. Approval does not indicate that the contents reflect the views of either Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.aquatox.2010.03.006.

## References

- Akesson, B., Bellew, T., Burk, R.F., 1994. Purification of selenoprotein P from human plasma. *Biochim. Biophys. Acta* 1204, 243–249.
- Ankley, G., Kahl, M., Jensen, K., Hornung, M., Korte, J., Makynen, E., Leino, R., 2002. Evaluation of the aromatase inhibitor fadrozole in a short-term reproduction assay with the fathead minnow (*Pimephales promelas*). *Toxicol. Sci.* 67, 121–130.
- Ankley, G.T., Bencic, D., Breen, M., Collette, T.W., Connolly, R., Denslow, N.D., Edwards, S., Ekman, D.R., Jensen, K.M., Lazorchak, J., Martinovic, D., Miller, D.H., Perkins, E.J., Orlando, E.F., Garcia-Reyero, N., Villeneuve, D., Wang, R.L., Watanabe, K., 2009. Endocrine disrupting chemicals in fish: developing exposure indicators and predictive models of effects based on mechanisms of action. *Aquat. Toxicol.* 92, 168–178.
- Ankley, G.T., Johnson, R.D., 2004. Small fish models for identifying and assessing the effects of endocrine-disrupting chemicals. *ILAR J.* 45, 469–483.
- Assinder, S.J., Stanton, J.A., Prasad, P.D., 2009. Transgelin: an actin-binding protein and tumour suppressor. *Int. J. Biochem. Cell Biol.* 41, 482–486.
- Basini, G., Tamanini, C., 2000. Selenium stimulates estradiol production in bovine granulosa cells: possible involvement of nitric oxide. *Domest. Anim. Endocrinol.* 18, 1–17.
- Bowden, M., Drummond, A.E., Salamonsen, L.A., Findlay, J.K., Nie, G., 2009. Evolutionary conservation of mammalian HTRA3 and its developmental regulation in the rat ovary. *J. Exp. Zool. B: Mol. Dev. Evol.* 312, 701–713.
- Bowden, M.A., Li, Y., Liu, Y.X., Findlay, J.K., Salamonsen, L.A., Nie, G., 2008. HTRA3 expression in non-pregnant rhesus monkey ovary and endometrium, and at the maternal-fetal interface during early pregnancy. *Reprod. Biol. Endocrinol.* 6, 22.
- Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C.A., Causton, H.C., Gaasterland, T., Glenisson, P., Holstege, F.C., Kim, I.F., Markowitz, V., Matese, J.C., Parkinson, H., Robinson, A., Sarkans, U., Schulze-Kremer, S., Stewart, J., Taylor, R., Vilo, J., Vingron, M., 2001. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat. Genet.* 29, 365–371.
- Cerda, J., Reidenbach, S., Pratzel, S., Franke, W.W., 1999. Cadherin-catenin complexes during zebrafish oogenesis: heterotypic junctions between oocytes and follicle cells. *Biol. Reprod.* 61, 692–704.
- Fabra, M., Raldua, D., Bozzo, M.G., Deen, P.M., Lubzens, E., Cerda, J., 2006. Yolk proteolysis and aquaporin-1 $\alpha$  play essential roles to regulate fish oocyte hydration during meiosis resumption. *Dev. Biol.* 295, 250–262.
- Filbin, M.T., Walsh, F.S., Trapp, B.D., Pizzey, J.A., Tennekoon, G.I., 1990. Role of myelin P0 protein as a homophilic adhesion molecule. *Nature* 344, 871–872.
- Granseth, B., Odermatt, B., Royle, S.J., Lagnado, L., 2007. Clathrin-mediated endocytosis: the physiological mechanism of vesicle retrieval at hippocampal synapses. *J. Physiol.* 585, 681–686.
- Hemler, M.E., 2003. Tetraspanin proteins mediate cellular penetration, invasion, and fusion events and define a novel type of membrane microdomain. *Annu. Rev. Cell Dev. Biol.* 19, 397–422.
- Hirota, Y., Osuga, Y., Nose, E., Koga, K., Yoshino, O., Hirata, T., Yano, T., Tsutsumi, O., Sakuma, S., Muramatsu, T., Taketani, Y., 2007. The presence of midline and its possible implication in human ovarian follicles. *Am. J. Reprod. Immunol.* 58, 367–373.
- Hoffmann, R., Seidl, T., Dugas, M., 2002. Profound effect of normalization on detection of differentially expressed genes in oligonucleotide microarray data analysis. *Genome Biol.* 3, 11, research0033. 1–0033.
- Holt, J.E., Jackson, A., Roman, S.D., Aitken, R.J., Koopman, P., McLaughlin, E.A., 2006. CXCR4/SDF1 interaction inhibits the primordial to primary follicle transition in the neonatal mouse ovary. *Dev. Biol.* 293, 449–460.
- Huang, D., Sherman, B., Lempicki, R., 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57.

- Jalabert, B., 2005. Particularities of reproduction and oogenesis in teleost fish compared to mammals. *Reprod. Nutr. Dev.* 45, 261–279.
- Jalabert, B., Szollosi, D., 1975. In vitro ovulation of trout oocytes: effect of prostaglandins on smooth muscle-like cells of the theca. *Prostaglandins* 9, 765–778.
- Jensen, K.M., Kahl, M.D., Makynen, E.A., Korte, J.J., Leino, R.L., Butterworth, B.C., Ankley, G.T., 2004. Characterization of responses to the antiandrogen flutamide in a short-term reproduction assay with the fathead minnow. *Aquat. Toxicol.* 70, 99–110.
- Ju, T., Aryal, R.P., Stowell, C.J., Cummings, R.D., 2008. Regulation of protein O-glycosylation by the endoplasmic reticulum-localized molecular chaperone Cosmc. *J. Cell. Biol.* 182, 531–542.
- Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T., Yamanishi, Y., 2008. KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* 36, D480–484.
- King, L.S., Kozono, D., Agre, P., 2004. From structure to disease: the evolving tale of aquaporin biology. *Nat. Rev. Mol. Cell. Biol.* 5, 687–698.
- Kiparissis, Y., Balch, G.C., Metcalfe, T.L., Metcalfe, C.D., 2003. Effects of the isoflavones genistein and equol on the gonadal development of Japanese medaka *Oryzias latipes*. *Environ. Health Perspect.* 111, 1158–1163.
- Knight, P.G., Glistler, C., 2006. TGF-beta superfamily members and ovarian follicle development. *Reproduction* 132, 191–206.
- Kryczek, I., Frydman, N., Gaudin, F., Krzysiek, R., Fanchin, R., Emilie, D., Chouaib, S., Zou, W., Machelon, V., 2005. The chemokine SDF-1/CXCL12 contributes to T lymphocyte recruitment in human pre-ovulatory follicles and coordinates with lymphocytes to increase granulosa cell survival and embryo quality. *Am. J. Reprod. Immunol.* 54, 270–283.
- Li, F., Curry Jr., T.E., 2009. Regulation and function of tissue inhibitor of metalloproteinase (TIMP) 1 and TIMP3 in periovulatory rat granulosa cells. *Endocrinology* 150, 3903–3912.
- Li, Y.H., Hou, Y., Ma, W., Yuan, J.X., Zhang, D., Sun, Q.Y., Wang, W.H., 2004. Localization of CD9 in pig oocytes and its effects on sperm-egg interaction. *Reproduction* 127, 151–157.
- Litman, T., Sogaard, R., Zeuthen, T., 2009. Ammonia and urea permeability of mammalian aquaporins. In: Beitz, E. (Ed.), *Aquaporins, Handbook of Experimental Pharmacology*, vol. 190. Springer-Verlag, Berlin, pp. 327–358.
- Lubzens, E., Young, G., Bobe, J., Cerda, J., 2009. Oogenesis in teleosts: how fish eggs are formed. *Gen. Comp. Endocrinol.* 165, 367–389.
- Martinovic, D., Blake, L.S., Durhan, E.J., Greene, K.J., Kahl, M.D., Jensen, K.M., Makynen, E.A., Villeneuve, D.L., Ankley, G.T., 2008. Reproductive toxicity of vinclozolin in the fathead minnow: confirming an anti-androgenic mode of action. *Environ. Toxicol. Chem.* 27, 478–488.
- Moreno-Hagelsieb, G., Latimer, K., 2008. Choosing BLAST options for better detection of orthologs as reciprocal best hits. *Bioinformatics* 24, 319–324.
- Mund, T., Pelham, H.R., 2009. Control of the activity of WW-HECT domain E3 ubiquitin ligases by NDFIP proteins. *EMBO Rep.* 10, 501–507.
- Ojeda, S.R., Romero, C., Tapia, V., Dissen, G.A., 2000. Neurotrophic and cell-cell dependent control of early follicular development. *Mol. Cell. Endocrinol.* 163, 67–71.
- Renko, K., Werner, M., Renner-Muller, I., Cooper, T.G., Yeung, C.H., Hollenbach, B., Scharpf, M., Kohrle, J., Schomburg, L., Schweizer, U., 2008. Hepatic selenoprotein P (SePP) expression restores selenium transport and prevents infertility and motor-incoordination in Sepp-knockout mice. *Biochem. J.* 409, 741–749.
- Smith, M.F., Ricke, W.A., Bakke, L.J., Dow, M.P., Smith, G.W., 2002. Ovarian tissue remodeling: role of matrix metalloproteinases and their inhibitors. *Mol. Cell. Endocrinol.* 191, 45–56.
- Soares, R., Balogh, G., Guo, S., Gartner, F., Russo, J., Schmitt, F., 2004. Evidence for the notch signaling pathway on the role of estrogen in angiogenesis. *Mol. Endocrinol.* 18, 2333–2343.
- Steinhoff, C., Vingron, M., 2006. Normalization and quantification of differential expression in gene expression microarrays. *Brief. Bioinform.* 7, 166–177.
- Tocharus, J., Tsuchiya, A., Kajikawa, M., Ueta, Y., Oka, C., Kawaichi, M., 2004. Developmentally regulated expression of mouse HtrA3 and its role as an inhibitor of TGF-beta signaling. *Dev. Growth Differ.* 46, 257–274.
- Tokumoto, T., 1999. Nature and role of proteasomes in maturation of fish oocytes. *Int. Rev. Cytol.* 186, 261–294.
- Tyler, C., Sumpter, J., 1996. Oocyte growth and development in teleosts. *Rev. Fish Biol. Fish.* 6, 287–318.
- van der Ven, L.T.M., Wester, P.W., Vos, J.G., 2003. Histopathology as a tool for the evaluation of endocrine disruption in zebrafish (*Danio rerio*). *Environ. Toxicol. Chem.* 22, 908–913.
- Van Nassauw, L., Harrisson, F., Callebaut, M., 1991. Localization of smooth-muscle markers in the ovaries of some ectothermic vertebrates. *Anat. Rec.* 229, 439–446.
- Villeneuve, D.L., Blake, L.S., Brodin, J.D., Greene, K.J., Knoebl, I., Miracle, A.L., Martinovic, D., Ankley, G.T., 2007. Transcription of key genes regulating gonadal steroidogenesis in control and ketoconazole- or vinclozolin-exposed fathead minnows. *Toxicol. Sci.* 98, 395–407.
- Villeneuve, D.L., Knoebl, I., Kahl, M.D., Jensen, K.M., Hammermeister, D.E., Greene, K.J., Blake, L.S., Ankley, G.T., 2006. Relationship between brain and ovary aromatase activity and isoform-specific aromatase mRNA expression in the fathead minnow (*Pimephales promelas*). *Aquat. Toxicol.* 76, 353–368.
- Villeneuve, D.L., Mueller, N.D., Martinovic, D., Makynen, E.A., Kahl, M.D., Jensen, K.M., Durhan, E.J., Cavallin, J.E., Bencic, D., Ankley, G.T., 2009a. Direct effects, compensation, and recovery in female fathead minnows exposed to a model aromatase inhibitor. *Environ. Health Perspect.* 117, 624–631.
- Villeneuve, D.L., Wang, R.L., Bencic, D.C., Biales, A.D., Martinovic, D., Lazorchak, J.M., Toth, G., Ankley, G.T., 2009b. Altered gene expression in the brain and ovaries of zebrafish exposed to the aromatase inhibitor fadrozole: microarray analysis and hypothesis generation. *Environ. Toxicol. Chem.* 28, 1767–1782.
- Vorontchikhina, M.A., Zimmermann, R.C., Shawber, C.J., Tang, H., Kitajewski, J., 2005. Unique patterns of Notch1, Notch4 and Jagged1 expression in ovarian vessels during folliculogenesis and corpus luteum formation. *Gene Expr. Patterns* 5, 701–709.
- Wallace, R.A., Selman, K., 1981. Cellular and dynamic aspects of oocyte growth in teleosts. *Am. Zool.* 21, 325–343.
- Weisbrod, C.J., Kunz, P.Y., Zenker, A.K., Fent, K., 2007. Effects of the UV filter benzophenone-2 on reproduction in fish. *Toxicol. Appl. Pharmacol.* 225, 255–266.