

DRAFT REPORT OF PHASE 1 OF THE VALIDATION OF THE FISH SEXUAL DEVELOPMENT TEST

TABLE OF CONTENTS

INTRODUCTION.....	3
Scientific rationale for the endpoints.....	4
OBJECTIVES OF PHASE 1	5
ORGANISATION OF PHASE 1.....	7
Introduction.....	7
Overview of the test method	7
Test fish.....	7
Test chemicals and concentrations.....	8
Participating laboratories.....	9
Time Schedule	9
Preparation of test solutions.....	9
Analytical methods for determination of test concentrations	9
Test conditions	9
Test acceptability criteria	10
Endpoints studied.....	10
Other observations	10
Data collection.....	10
Vitellogenin measurement and supply of kits.....	10
Statistical analysis:	11
Vtg.....	11
Sex ratio	11
RESULTS	11
Analytical chemistry.....	11
Analytical chemistry of 4-tert pentylphenol and prochloraz	11
Vitellogenin analysis	15
Vitellogenin analysis of 4-tert pentylphenol studies.....	15
Vitellogenin analysis of prochloraz studies.....	19
Sex ratio analysis	25
Sex ratio analysis of 4-tert pentylphenol studies.....	25
Sex ratio analysis of prochloraz studies.....	28
Statistical analysis of sex ratio:	31
Part 1: Analysis of two experiments with fathead minnow conducted at Bayer CropScience.....	31
Part 2: Analysis of two experiments with fathead minnow conducted at Bayer CropScience. Pooled control groups.....	31
Part 3: Analysis of three experiments with zebrafish conducted at DHI and Heidelberg	31
DISCUSSION	33
Literature review of Fathead minnow vitellogenin.....	33
Phase 1 Vitellogenin data	35
4-tert pentylphenol	36

Prochloraz.....	37
Phase 1 Sex ratio data	38
4-tert pentylphenol	39
Prochloraz.....	39
Sex ratio statistics	39
Green and Springer parts 1 and 2 conclusions.....	39
Green and Springer Part 3 conclusions.....	39
John Green’s final conclusions	40
Revised sex ratio NOECs using the improved statistical procedure	40
OVERALL CONCLUSIONS AND RECOMMENDATIONS ON THE FISH SEXUAL DEVELOPMENT TEST (FSDT).....	42
Results from Phase 1:	42
Possible options for Phase 2	43
Recommendations.....	43
REFERENCES	44
APPENDIX	46
Appendix 1.....	46
Preparation of stock solutions of 17 β -estradiol, 4-tert-pentylphenol and prochloraz without the use of carrier solvent.....	46
Appendix 2.....	51
Method for quantification of Prochloraz by LC-MS	51
Appendix 3.....	53
Method for quantification of 4-tert-pentylphenol by LC-MS	53
Appendix 4.....	54
<u>ANALYTICAL METHOD FOR DETERMINATION OF PROCHLORAZ CONCENTRATIONS</u> (provided by GT Ankley, US-EPA Duluth, 2004)	54
Appendix 5.....	56
4-Tert pentylphenol (4-tPP) and estradiol (E2): Analytical methods provided by Wildlife International, LTD.	56
Appendix 6.....	58
SOP: Homogenisation of head & tail from juvenile zebrafish, fathead minnow, three spined stickleback and Japanese medaka.....	58
Appendix 7.....	59
SOP: Quantification of head & tail homogenate vitellogenin in zebrafish (Danio rerio) (modified from Holbech et al., 2001)	59
Appendix 8.....	60
SOP: Preparation of tissue sections for sex determination and staging of gonads.....	60
Euthanasia, Necropsy, and Tissue Fixation.....	60
Tissue Processing.....	60
Embedding.....	61
Microtomy	61
Staining, Cover slipping, and Slide Labelling	62
Appendix 9.....	63
Description of Experimental Data	63
Utility of Regression Modeling for Fish Sex Ratio Data.....	63
Statistical Protocol - NOEC.....	63
Appendix 10.....	71

INTRODUCTION

The need to develop and validate a fish assay capable of measuring the hazards of endocrine disrupting chemicals (or EDCs) originates from concerns that environmental levels of certain chemicals may be causing adverse effects in both humans and wildlife due to the interaction of these chemicals with the endocrine system. Several cases have been reported where exposures to exogenous chemicals have indeed resulted in effects on wildlife, in particular fish ([Jensen et al. 2006;Orlando et al. 2004;Milnes et al. 2006]). In 1997, OECD member countries advised that existing test methods were insufficient to identify such substances and characterize their effects. As part of the OECD Test Guidelines Program a Special Activity on the Testing and Assessment of Endocrine Disrupters was therefore initiated to revise existing, and develop new, OECD Test Guidelines for the screening and testing of potential EDCs. A Task Force on Endocrine Disrupters Testing and Assessment (EDTA) was subsequently established to provide a focal point within OECD to consider and recommend priorities for the development of testing methods.

The Fish Sexual Development Test (FSDT) fits into the OECD Conceptual Framework for the Testing and Assessment of Endocrine Disrupters, discussed and agreed at the sixth Meeting of the EDTA Task Force. This framework identifies approaches, assays and long-term tests of increasing biological complexity, meant to gather information on potential EDCs. Each of the tools added to the framework will require validation to ensure its relevance and reliability, the two main validation principles. OECD Guidance Document 34 on Validation and Acceptance of New and Updated Test Methods for Hazard Assessment provides definitions, principles and concrete examples of validation, applied in different areas of hazard assessment.

The FSDT (earlier named the Fish Partial Life-Cycle Test) is a modified version of OECD guideline 210 adopted in 1992, the Fish Early-Life Stage Toxicity Test with added end-points for the detection of endocrine disrupters (secondary sex characteristics, vitellogenin (Vtg) concentration and sex ratio) (OECD, 2005). The main idea behind this assay is that exposure of fish to certain EDCs during the sensitive window for sexual development will alter the plasma Vtg concentration and/or phenotypic sex ratio and/or secondary sexual characteristics. The FSDT was first developed for zebrafish (*Danio rerio*), which possess a sensitive window of exposure from 20-60 days post hatch (DPH). It was designed as a relatively short test and the window of exposure was chosen to avoid zebrafish exposure during an oversensitive stage, between 10-20 DPH, when high larval mortality can occur [Andersen et al. 2003]. After discussion of the test at OECD in 2003, it was decided that other OECD candidate species such as Japanese medaka (*Oryzias latipes*) and fathead minnow (*Pimephales promelas*), and in 2006 stickleback (*Gasterosteus aculeatus*), should also be able to be used in the test. Since the sensitive period has not been fully explored in these species, newly fertilized eggs (instead of larvae) were therefore proposed to be used at the test start for all species including zebrafish.

The FSDT was included in the work program of the OECD test guideline program in 2003, with high approval from member countries. During autumn 2005 and early spring 2006 the test proposal was put through an in-depth statistical evaluation. The outcome was presented at the 9th EDTA meeting held in Stockholm April 26-27, 2006. The group agreed to take the FSDT proposal through an experimental validation. The first round (Phase 1) of this validation exercise took place during summer/autumn 2006 and spring/summer 2007. The test substances were 4-tert-pentylphenol and prochloraz. The participants in the validation were: Bayer CropScience, Germany; Heidelberg University, Germany; Swedish University of Agricultural Sciences (SLU), Sweden; University of Southern Denmark (SDU), Denmark; and DHI, Denmark.

The FSDT, evaluated in the Phase 1 validation trial, is intended for the hazard evaluation of individual chemicals acting as estrogens, anti-estrogens, androgens, anti-androgens and aromatase inhibitors. The aim of the validation has been to develop a robust, relevant and reliable test method for the assessment of chemicals mentioned above. It is also the purpose of the validation to understand and define the area of application of the assay and any limitations on its use.

Scientific rationale for the endpoints

Vitellogenin

Vitellogenin (Vtg) is a phospholipoglycoprotein precursor to egg yolk protein that normally occurs in sexually active females of all oviparous species; the production of Vtg is controlled by interaction of estrogenic hormones, predominantly 17β -estradiol, with the estrogen receptor (4). Males retain the capacity to produce Vtg in response to stimulation with estrogen receptor agonists; as such, induction of Vtg in males has been successfully exploited as a biomarker specific for estrogenic compounds in a variety of fish species, including fathead minnow, Japanese medaka and zebrafish. A number of Vtg measurement methods have been developed and standardized for each of these species ([Kunz et al. 2006; Lange et al. 2001; Panter et al. 2002; Panter et al. 2006]). The criteria for selecting the methods used in this validation programme are described further in the report.

Sex ratio

Phenotypic sex ratio (proportions of males, females etc) is determined by gonadal examination. Sex is defined as female, male, intersex or undifferentiated.

Secondary sex characteristics

Nuptial tubercles in fathead minnow and papillary processes in Japanese medaka are among the sexual characteristics observed under normal conditions in adult males. Chemicals with certain endocrine-mediated action will cause the abnormal occurrence of these secondary sex characteristics in the opposite sex. For example, androgen receptor agonists will induce formation of nuptial tubercles in female fathead minnow [Ankley et al. 2001; Jensen et al. 2006] and formation of papillary processes in female Japanese medaka [Asahina et al. 1989]. No defined secondary sex characteristics are available for the zebrafish so far.

Gonad histopathology

There is an interest in developing and applying toxicological pathology for lower vertebrates, such as fish, particularly in the case of EDCs. Examination of gonadal histopathology has been beneficial in understanding and assessing the effects of EDCs in fish [Patyna et al. 1999; Hutchinson and Pickford 2002; Nash et al. 2004] because histopathological changes integrate a large number of interactive physiological processes. However, it is recognized that there is inherent subjectivity in histopathological evaluation, caused by difficulty in consistently defining effects across species. Efforts are being made to reduce the bias associated with individual pathologists, to optimize histological techniques, to describe exposure-related diagnoses and to standardize the reporting of observed changes. These aspects are critical for the regulatory acceptance and use of histopathology. The diagnostic value of the endpoint, the amount of effort necessary and the time required for a reliable evaluation will, together, determine the role of gonadal histopathology in a screening assay for endocrine active substances. This endpoint represents a major challenge in the validation effort and gonad histopathology is therefore not mandatory but an option in the FSDT.

OBJECTIVES OF PHASE 1

The objective of Phase 1 was primarily to experimentally investigate the sensitivity and reproducibility of the assay for assessing the effects of weakly active EDCs in several test fish species. Here is a summary of the key goals of Phase 1:

- Obtain additional information on the relevance of assay endpoints and in particular their ability to identify weakly active substances with diverse modes of action
- Collect a set of data for two model EDCs;
- Obtain additional information on possible differences in species sensitivity to a weak estrogen and an aromatase inhibitor;
- Check that the protocol contains enough details to enable laboratories to conduct the assay in a reproducible manner;
- Check the reproducibility of test results in laboratories located in diverse geographical areas, and with diverse levels of experience in conducting this type of assay.
- Compare the FSDT validation results to the Fish Screening Assay (FSA) validation results.

OECD Guidance Document 34 (64) provides the following important definitions:

Relevance: Description of relationship of the test to the effect of interest and whether it is meaningful and useful for a particular purpose. It is the extent to which the test correctly measures or predicts the biological effect of interest. Relevance incorporates consideration of the accuracy (concordance) of a test method.

Reliability: Measures of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and inter-laboratory reproducibility and intra-laboratory repeatability.

Repeatability: The agreement among test results obtained within a single laboratory when the procedure is performed repeatedly on the same substance under identical conditions. (see Reliability)

Reproducibility: The degree of agreement among results obtained from testing the same substance using the same test protocol in different laboratories (see reliability).

Robust(ness): Sensitivity of test results in response to departures from the specified test conditions when conducted in different laboratories or over a range of conditions under which the test method might normally be used. If a test is not robust, it will be difficult to use in a reproducible manner within and between laboratories.

Transferability: The ability of a test procedure to be accurately and reliably performed in independent, competent laboratories.

Validation: The process by which the reliability and relevance of a particular approach, method, process or assessment is established for a defined purpose.

These concepts represent components of the validation. The mechanistic relevance of endpoints used in the FSDT assay, i.e. vitellogenin, sex ratio, and secondary sex characteristics, and to a certain extent gonad histopathology, has been described on several occasions. However the established relationship between e.g. vitellogenin induction and adverse health effects in fish is limited. In particular the predictive value of low to moderate vitellogenin induction or abnormal secondary sex characteristics for the reproductive health of fish is not yet well defined. The relationship between skewed sex ratio and adverse health effects is, on the other hand, more clear. Sex ratio is a population-relevant endpoint because it has direct implications for reproductive success.

ORGANISATION OF PHASE 1

Introduction

For Phase 1, DHI and SDU served as technical lead laboratories. Before initiation of Phase 1 a protocol including standard operating procedures for vitellogenin measurements and sex ratio determinations was provided. Each participating laboratory (see Table 0-2 for contact details), was asked to submit a study plan to the lead laboratories for its participation in Phase 1, and to conduct experimental work in compliance with the agreed protocol.

Overview of the test method

The experimental work was conducted according to the protocol prepared for Phase 1 of the validation of the Fish Sexual Development Test for Endocrine Active Substances. A summary of the protocol is provided below.

The protocol was designed to detect the effects of EDCs in fish exposed during their sex differentiation period.

The test is initiated with newly fertilized eggs, transferred to the exposure beakers. However, note that the sensitive window of exposure in zebrafish is 20-60 dph, and is it possible that different results would have been obtained if the zebrafish experiment had been started at day 20.

Two different exposure scenarios were tested in Phase 1: One scenario used five test concentrations as well as appropriate controls, with two replicate aquaria per concentration and 45 individuals per aquarium. The second scenario used three test concentrations as well as appropriate controls, with at least four replicate aquaria per concentration and 45 individuals per aquarium. For Phase 1 one laboratory used three exposure concentrations of both chemicals and three laboratories used three exposure concentrations with one chemical and five concentrations with the second chemical.

Based on the preliminary outcome of the validation studies with fathead minnow performed by Bayer CropScience, the length of the exposure phase was discussed at the VMG-eco Meeting in Madrid 2007. Here it was decided that Denmark should undertake a literature review on the Vtg response in young fathead minnows. At the meeting it was agreed that Vtg is an important endpoint in the FSDT assay because it demonstrates the endocrine-related mode of action of the responses. VMG-eco agreed that further validation studies on the fathead minnow should take into consideration the need for possible extended exposure duration. Therefore, the length of the exposure phase in the studies with fathead minnow is different between the different participating laboratories (60 and 120 days, respectively). At the meeting it was also agreed that gonadal histology should be kept outside the scope of the validation of the FSDT for the time being.

Exposure to the test chemical is aqueous, with or without carrier solvent. Monitoring continues for up to 60 (120) days post hatch (dph) and includes hatching rate, development, survival, growth (total length and body weight), sexual differentiation, secondary sex differentiation, gonadal development, and Vtg levels.

Test fish

One fathead minnow (*Pimephales promelas*) strain (Bayer Crop Science) and three zebrafish (*Danio rerio*) strains (Heidelberg, Uppsala and DHI) were used in the Phase 1 validation work. These are commonly used species for regulatory work in OECD member countries. The fish strains were not defined by the protocol. The exposure phase was started with newly fertilized eggs obtained from a breeding

culture. The breeding culture did not receive any treatment for diseases in the 2-week acclimation period preceding the test, or during the exposure period.

Test chemicals and concentrations

The Fish Drafting Group followed the recommendation from the EDTA Task Force to use weak test substances. Two chemicals, representing different modes of action, were proposed, with the addition of 17 β -estradiol as a positive control:

Table 0-1: Test substances used in Phase 1.

Test substance	CAS-number	Lot no	purity., Supplier	Mode of action
prochloraz	67747-09-5	2226X	99.5 % Sigma Aldrich, US	Aromatase inhibitor
4-tert-pentylphenol	80-46-6	13625 LB	99.9 % Sigma Aldrich, US	Weak estrogen
17 β -estradiol	50-28-2	103K1 117	100 % Sigma Aldrich, US	Strong estrogen

All substances were managed through SDU, Denmark. The choice of weakly-acting test substances and test concentrations was based on the selections for the FSA validation. Prochloraz is an aromatase inhibitor (interfering with the conversion of testosterone into 17 β -estradiol in females), and 4-tert-pentylphenol is a weak estrogen. The chemical lot numbers etc are presented in Table 0-1.

Concentrations of both test substances in Phase 1 were as follows:

- Three concentration scenario (NOEC approach): 32, 100 and 320 μ g/l (+ water control);
- Five concentration scenario (ECx-approach): 38, 75, 150, 300 and 600 μ g/l (+ water control).
- 17 β -estradiol: 100 ng/l.

Participating laboratories

Five laboratories participated in the Phase 1 validation. The names of the laboratories and their role in the validation exercise are given in Table 0-2.

Table 0-2: Participating laboratories, test species and statistical approach

Laboratory (lab-no.)	Species	Statistical approach
University of Southern Denmark (SDU), Odense, Denmark (Lab 1)	Vtg analysis for all participating laboratories. Water sample analysis for Lab. 2, 3 and 4. Histology for Lab 2.	-
DHI, Institute of Water and Environment, Hoersholm, Denmark (Lab 2)	Zebrafish Fathead minnow (120 days)	NOEC and ECx NOEC
Swedish University of Agricultural Sciences, Uppsala, Sweden (Lab 3)	Zebrafish	NOEC and ECx
University of Heidelberg, Heidelberg, Germany (Lab 4)	Zebrafish	NOEC and ECx
Bayer CropScience AG, Mannheim, Germany (Lab 5)	Fathead minnow (60 days)	NOEC

Time Schedule

The experimental work started in August 2006 until mid-2007.

Preparation of test solutions

Test solutions of the selected concentrations were prepared by dilution of a stock solution. The stock solutions were prepared by simply mixing the test substance in the dilution water by mechanical means (e.g. stirring or ultrasonication) or by solid-liquid saturation. Detailed standard operating procedures (SOP) for each substance are presented in Appendix 1. LAB 2 encountered difficulties in achieving full solubility of 4-tert-pentylphenol in the stock solution without solvent; this possibly explains the low measured concentrations of 4-tert-pentylphenol in the LAB 2 dataset (see Table 0-3).

Analytical methods for determination of test concentrations

Water samples were collected on a weekly basis in each tank and reported in a spreadsheet. The water samples from Lab 2, 3 and Lab 4 was analysed at SDU, Odense by use of LC-MS. Samples from Lab 5 were analysed at Bayer CropScience. Analysis methods are described in the appendixes.

Test conditions

A flow-through test system was used. Such a system continually dispenses and dilutes a stock solution of the test substance (e.g. metering pump, proportional diluter, saturator system) in order to deliver a series of concentrations to the test chambers. The flow rates of stock solutions and dilution water were checked at intervals and were not intended to vary by more than 10% throughout the 60 (120) days of the test.

Test acceptability criteria

For the test results to be acceptable, the following conditions applied:

- The dissolved oxygen concentration was between 60 and 100 per cent of the air saturation value (ASV) throughout the exposure period.
- The water temperature did not differ by more than $\pm 2.0^{\circ}\text{C}$ between test vessels at any one time during the exposure period.
- The following criteria were fulfilled in the controls:

		Fathead minnow	Zebrafish
Hatchability:		>66%	>80%
Survival		>70%	>70%
Growth	Weight	≥ 275 mg	>50-75mg
	Length	≥ 25 mm	>10-15 mm
Sex ratio (% males)		Not yet determined	35-65%

Endpoints studied

Two endpoints were indicators of endocrine activity of the test substances: Sex ratio (proportions of female, male, intersex and undifferentiated) and Vtg concentration in head/tail homogenate. Besides, hatching success, survival and growth were registered during exposure or measured during termination of the test. Histology (e.g. gonadal staging) was optional. For fathead minnow secondary sexual characteristics were also recorded.

Other observations

Behaviour and external abnormalities

Fish were examined daily during the test period and any external abnormalities (such as haemorrhage, discolouration, signs of general toxicity including hyperventilation, uncoordinated swimming, loss of equilibrium, and atypical quiescence or feeding) noted. Mortality was recorded and the dead fish were removed as soon as possible.

Data collection

Participating laboratories recorded the raw experimental data from Phase 1 on standardized Excel spreadsheets developed specifically for the validation study. A workbook (collection of Excel spreadsheets for each endpoint) was prepared for each species. In addition to raw data, means and standard deviations were calculated. All completed data sheets with laboratory results are available from the Lead laboratory and from the OECD Secretariat.

Vitellogenin measurement and supply of kits

Standard operating procedures (SOPs) for the sampling procedure (homogenisation of head and tail from juvenile fish) have been added as appendices to the Phase 1 protocol. The measurement of zebrafish Vtg was performed according to the method published by Holbech et al, 2001. The fathead minnow Vtg concentration was measured in head/tail homogenate by use of a homologous commercial fathead minnow vitellogenin ELISA kit (Biosense, Norway). All Vtg measurements were performed at the same laboratory: Lab 1: SDU, Denmark.

Statistical analysis:

Vtg

The statistical analysis of Vtg concentrations was performed as described in the FSDT protocol: if possible, data were transformed to obtain normality and homogeneity of variances. Results were evaluated using a one-way analysis of variance (ANOVA) followed by a Bonferroni adjusted Fisher's Least Significant Difference test or Kruskal–Wallis One Way Analysis of Variance on Ranks followed by Multiple Comparisons versus Control Group (Dunn's Method).

Sex ratio

The statistical analysis of sex ratio has been discussed for several years because there are some decisions to be made on the FSDT test design that affect the analysis of sex ratio i.e. number of test concentrations and replicates. Therefore the FSDT Phase 1 was organized to test two different designs: A NOEC design with three test concentrations and four replicates and an ECx design with five test concentrations in two replicates. Both approaches have been tested in the analysis of the first results:

For the data treatment shown in Figures 4-11 to 4-19, the proportions of each sex category were the unit analysed and the Chi-square test followed by Bonferroni correction was used to compare control mean with treatment mean and determine significance.

Proportions of each sex were determined according to the mode of action of the test substances: for 4-tert pentylphenol experiments, proportions were defined as females or “not females” due to the estrogenic mode of action of 4-tert pentylphenol and the same was done for prochloraz due to the aromatase inhibiting mode of action of prochloraz.

John W. Green, DuPont, USA and Timothy A. Springer, Wildlife International, USA have produced three documents on how to treat sex ratio data statistically, based on the first results of the Phase 1 validation. Their overall approach is described in section 4. A summary of the conclusions of the three documents is shown in section 6. A final evaluation of almost the whole dataset has also been produced by John Green, is summarized in section 6, and is reproduced in full in Appendix 9.

RESULTS

Analytical chemistry

Water samples were collected in accordance with the FSDT protocol. Due to logistical problems not all water samples from Lab 4 (Heidelberg) and Lab 3 (Uppsala) were analyzed.

Analytical chemistry of 4-tert pentylphenol and prochloraz

Fathead minnow

Lab 5: Fathead minnow

Table 0-1

Lab 5: 4-tert pentylphenol			Lab 5: Prochloraz		
Nominal Concentration	Mean Measured	% nominal	Nominal	Mean Measured	% nominal

32 µg/l	36.1	113	32 µg/l	29.2	91
100 µg/l	93.0	93	100 µg/l	96.2	96
320 µg/l	295.6	92	320 µg/l	284.0	89

Lab 2: Fathead minnow

Table 0-2

Lab 2: Prochloraz		
Nominal	Mean Measured	% nominal
32 µg/l	30.9	97
100 µg/l	106.0	106
320 µg/l	300.5	94

Zebrafish

Lab 2: Zebrafish

Table 0-3

Lab 2: 4-tert pentylphenol			Lab 2: Prochloraz		
Nominal Conc.	Mean Measured	% nominal	Nominal Conc.	Mean Measured	% nominal
38 µg/l	12.9 µg/l	34	32 µg/l	14.5 µg/l	45
75 µg/l	17.3 µg/l	23	100 µg/l	48.0 µg/l	48
150 µg/l	34.0 µg/l	23	320 µg/l	319.6 µg/l	100
300 µg/l	62.2 µg/l	21			
600 µg/l	96.8 µg/l	16			

Lab 3: Zebrafish

Table 0-4

* Mean measured concentrations are based on only four samples from day 30 and 60 post hatch.

**mean measured concentrations are based on only two samples from day 30 and 60 post hatch.

Lab 3: 4-tert pentylphenol			Lab 3: Prochloraz		
Nominal Conc.	Mean measured	% nominal	Nominal Conc.	Mean measured	% nominal
38 µg/l	4.9 µg/l**	12.9	38 µg/l	22.1* µg/l	58
75 µg/l	4.7 µg/l**	6.3	75 µg/l	43.8* µg/l	58
150 µg/l	117 µg/l**	78	150 µg/l	98.6* µg/l	66
300 µg/l	13.0 µg/l**	4.3	300 µg/l	197.4* µg/l	66
600 µg/l	23.2 µg/l**	3.9	600 µg/l	434.1* µg/l	72
Stock solutions		49.5			

Lab 4: Zebrafish

Table 0-5

Lab 4: 4-tert pentylphenol			Lab 4: Prochloraz		
Nominal Conc.	Mean Measured	% nominal	Nominal Conc.	Mean Measured	% nominal
32 µg/l	2.0 µg/l*	6.2	38 µg/l	59.7 µg/l	159
100 µg/l	0.9 µg/l*	0.9	75 µg/l	134.9 µg/l	180
320 µg/l	1.8 µg/l*	0.6	150 µg/l	182.7 µg/l	122
Stock solutions		71	300 µg/l	233.4 µg/l	78
			600 µg/l	1166.2 µg/l	194

The water samples were measured at SDU, Denmark. Due to preservation problems with the 4-t-pp samples (see discussion) these results are not reliable and the nominal concentrations will be used.

Vitellogenin analysis

Vitellogenin analysis of 4-tert pentylphenol studies

Fathead minnow (Lab 5)

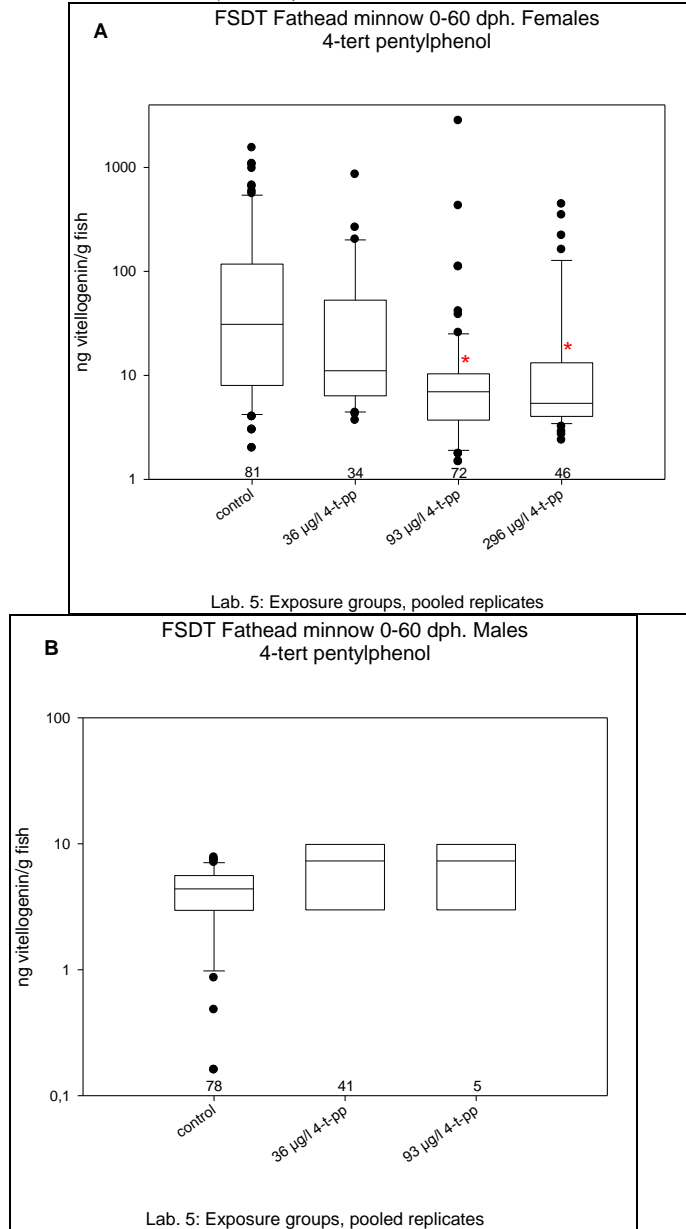


Figure 0-1: Vitellogenin concentrations in head/tail homogenate from fathead minnow after exposure to 4-tert pentylphenol (measured concs) from 0-60 DPH at Lab 5. Female and male vitellogenin concentrations are presented in Figure A and B respectively. The boundary of the box closest to zero indicates the 25th percentile, a line within the box marks the median, and the boundary of the box farthest from zero indicates the 75th percentile. Whiskers above and below the box indicate the 90th and 10th percentiles. * Indicate significantly different from the control group ($p < 0.05$). The number of fish from each group is marked at the bottom of the figure.

Zebrafish

Lab 2:

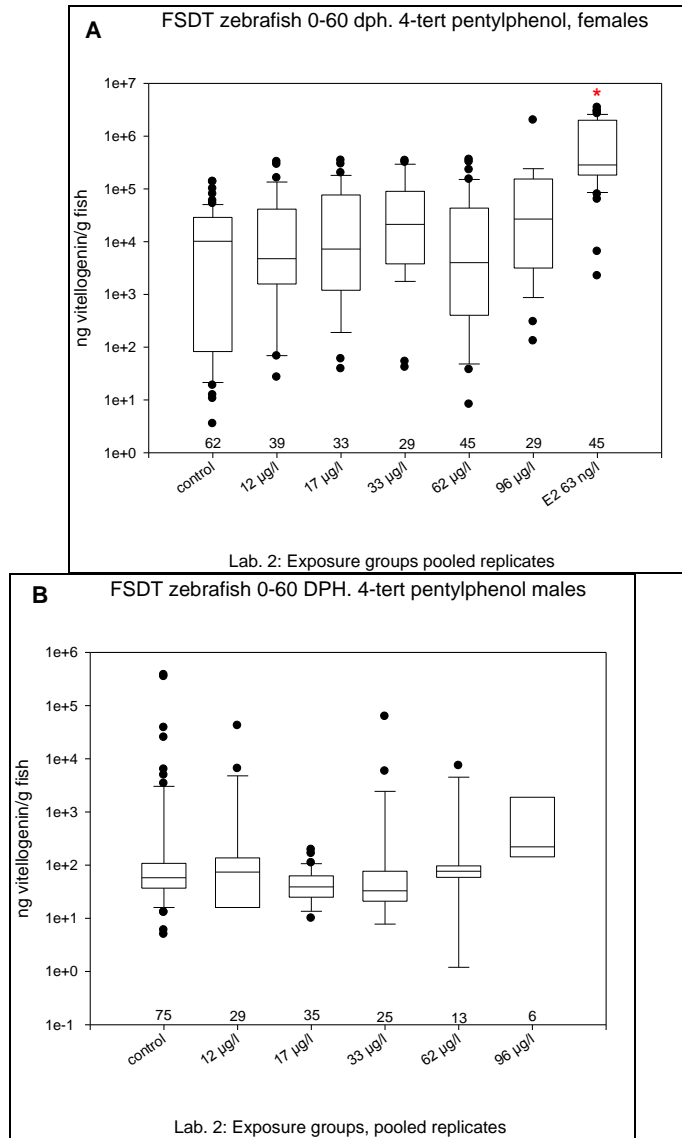


Figure 0-2: Vitellogenin concentrations in head/tail homogenate from zebrafish after exposure to 4-tert pentylphenol (measured concs) from 0-60 DPH at Lab 2. Female and male vitellogenin concentrations are presented in Figure A and B respectively. The boundary of the box closest to zero indicates the 25th percentile, a line within the box marks the median, and the boundary of the box farthest from zero indicates the 75th percentile. Whiskers above and below the box indicate the 90th and 10th percentiles. * Indicate significantly different from the control group ($p < 0.05$). The number of fish from each group is marked at the bottom of the figure.

Lab 3:

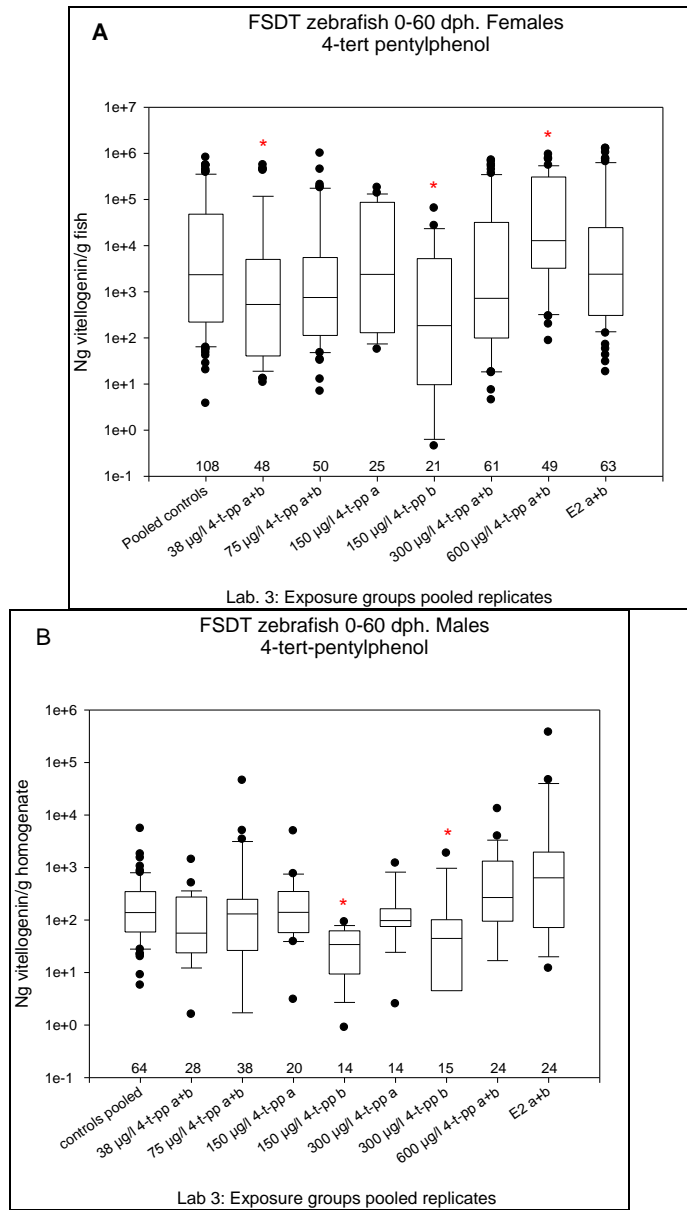


Figure 0-3: Vitellogenin concentrations in head/tail homogenate from zebrafish after exposure to 4-tert-pentylphenol (nominal concs) from 0-60 DPH at Lab 3. Female and male vitellogenin concentrations are presented in Figure A and B respectively. The boundary of the box closest to zero indicates the 25th percentile, a line within the box marks the median, and the boundary of the box farthest from zero indicates the 75th percentile. Whiskers above and below the box indicate the 90th and 10th percentiles. Only nominal concentrations are shown as very few concentration measurements were made. * Indicate significantly different from the control group ($p < 0.05$). The number of fish from each group is marked at the bottom of the figure.

Lab 4:

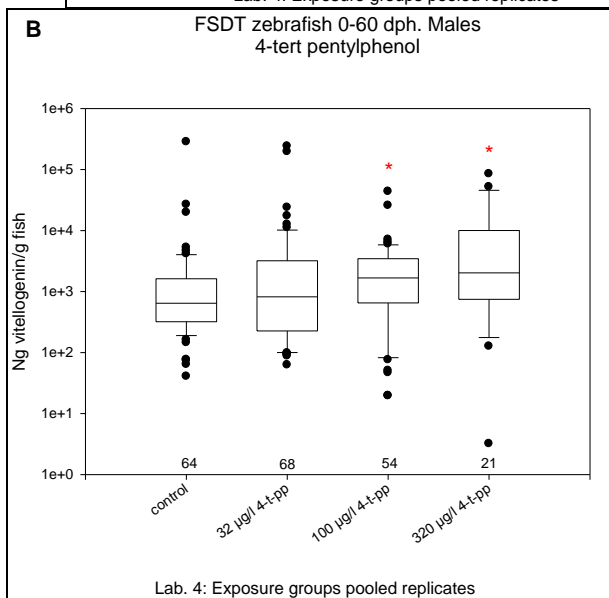
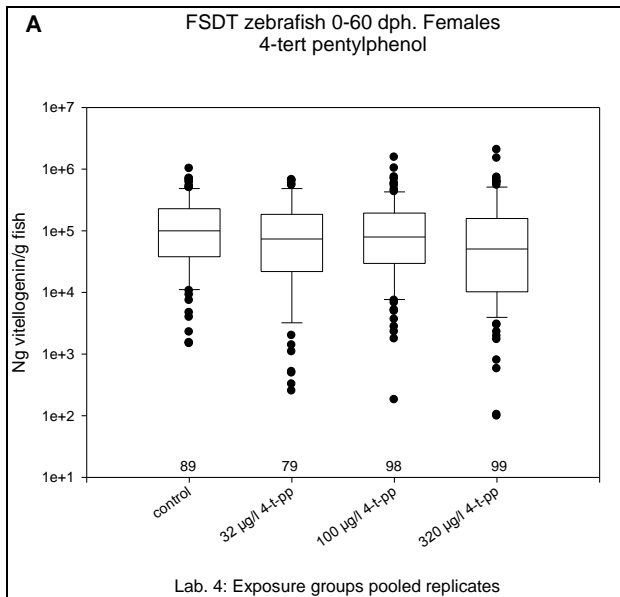


Figure 0-4: Vitellogenin concentrations in head/tail homogenate from zebrafish after exposure to 4-tert pentylphenol (nominal concs) from 0-60 DPH at Lab 4. Female and male vitellogenin concentrations are presented in Figure A and B respectively. The boundary of the box closest to zero indicates the 25th percentile, a line within the box marks the median, and the boundary of the box farthest from zero indicates the 75th percentile. Whiskers above and below the box indicate the 90th and 10th percentiles. * Indicate significantly different from the control group ($p < 0.05$). The number of fish from each group is marked at the bottom of the figure.

Vitellogenin analysis of prochloraz studies

Fathead minnow

Lab 5:

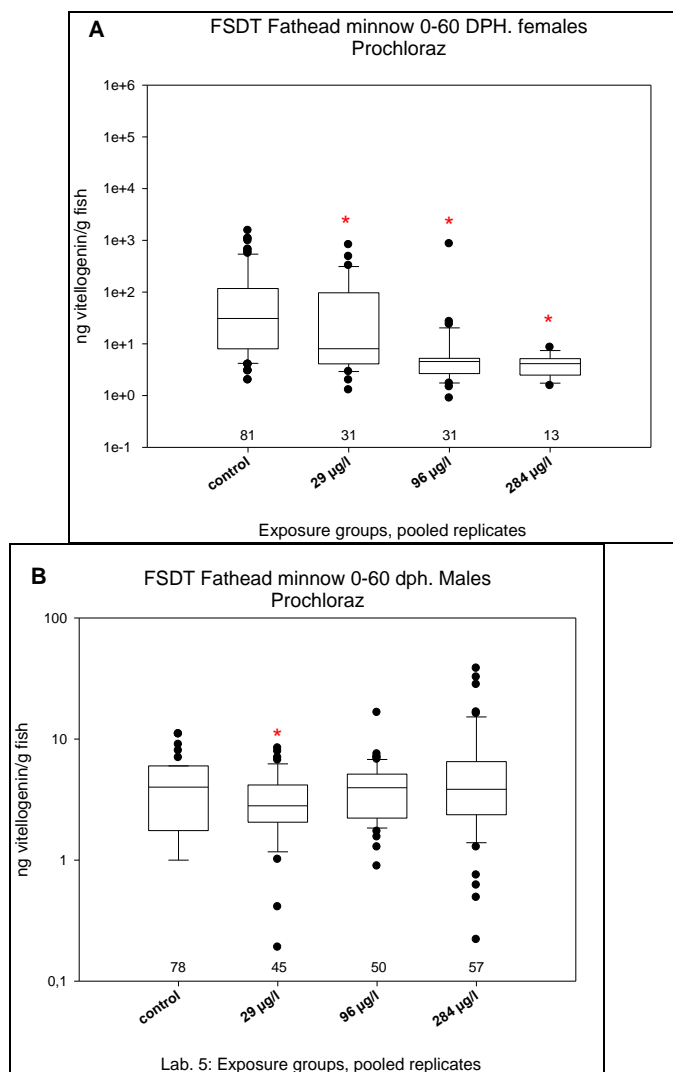


Figure 0-5: Vitellogenin concentrations in head/tail homogenate from fathead minnow after exposure to prochloraz (measured concs) from 0-60 DPH at Lab 5. Female and male vitellogenin concentrations are presented in Figure A and B respectively. The boundary of the box closest to zero indicates the 25th percentile, a line within the box marks the median, and the boundary of the box farthest from zero indicates the 75th percentile. Whiskers above and below the box indicate the 90th and 10th percentiles. * Indicate significantly different from the control group ($p < 0.05$). The number of fish from each group is marked at the bottom of the figure.

Lab 2:

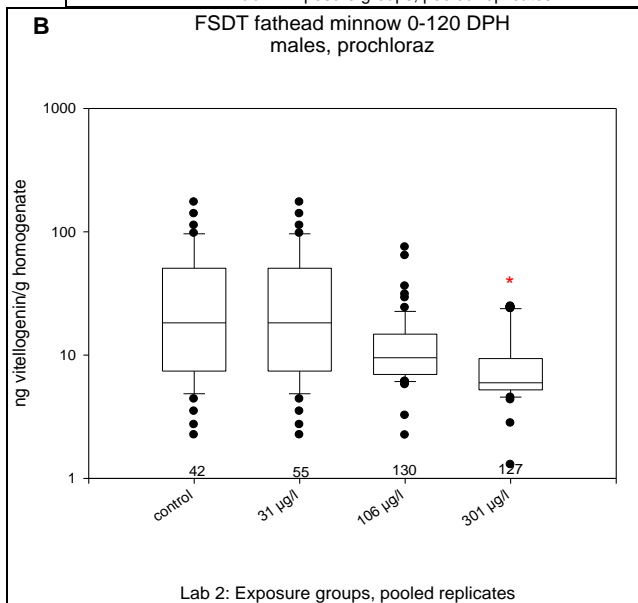
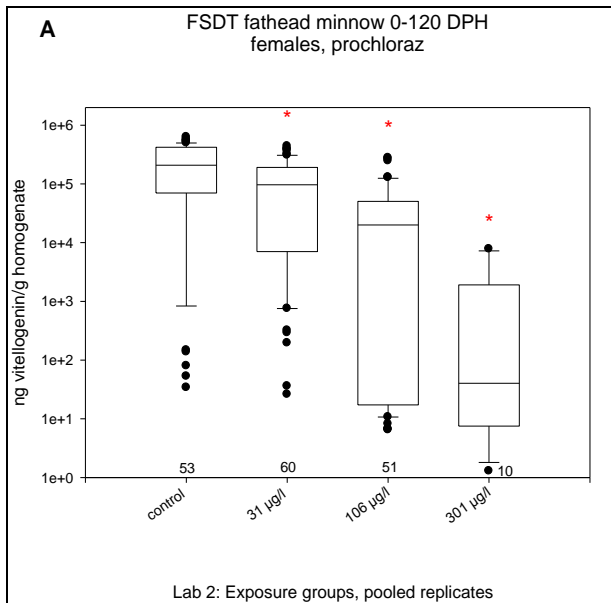


Figure 0-6: Vitellogenin concentrations in head/tail homogenate from fathead minnow after exposure to prochloraz (measured concs) from 0-120 DPH at Lab 2. Female and male vitellogenin concentrations are presented in Figure A and B respectively. The boundary of the box closest to zero indicates the 25th percentile, a line within the box marks the median, and the boundary of the box farthest from zero indicates the 75th percentile. Whiskers above and below the box indicate the 90th and 10th percentiles. * Indicate significantly different from the control group ($p < 0.05$). The number of fish from each group is marked at the bottom of the figure.

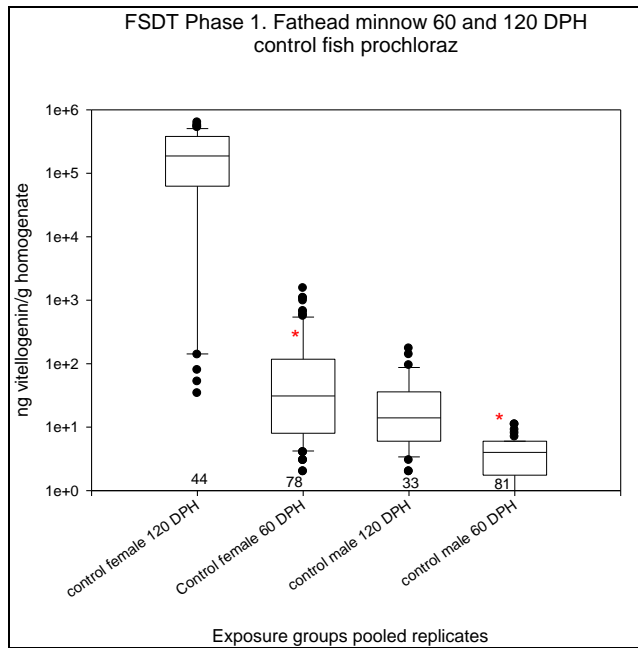


Figure 0-7: Vitellogenin concentrations in head/tail homogenate from fathead minnow age 60 DPH (Lab 5) or 120 DPH (Lab 2). All groups are based on pooled replicates. The boundary of the box closest to zero indicates the 25th percentile, a line within the box marks the median, and the boundary of the box farthest from zero indicates the 75th percentile. Whiskers above and below the box indicate the 90th and 10th percentiles. * Indicate significantly different from the control group ($p < 0.05$). The number of fish from each group is marked at the bottom of the figure

Zebrafish

Lab 2

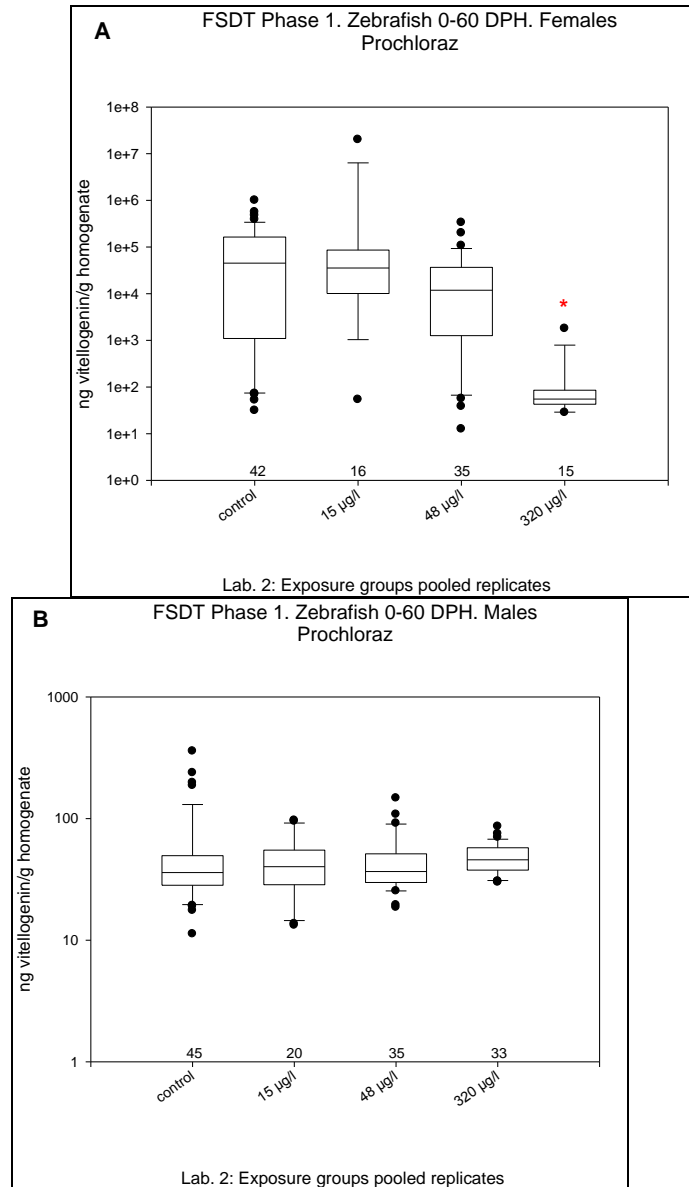


Figure 0-8: Vitellogenin concentrations in head/tail homogenate from zebrafish after exposure to prochloraz (measured concs) from 0-60 DPH at Lab 2. Female and male vitellogenin concentrations are presented in Figure A and B respectively. The boundary of the box closest to zero indicates the 25th percentile, a line within the box marks the median, and the boundary of the box farthest from zero indicates the 75th percentile. Whiskers above and below the box indicate the 90th and 10th percentiles. * Indicate significantly different from the control group ($p < 0.05$). The number of fish from each group is marked at the bottom of the figure.

Lab. 3:

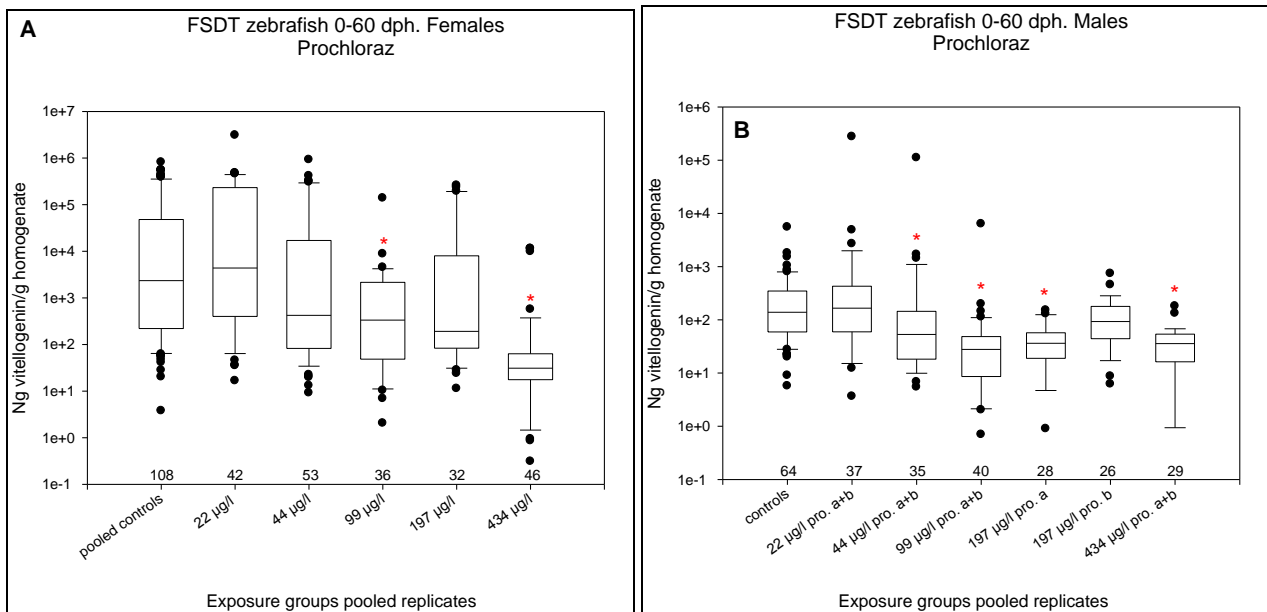


Figure 0-9: Vitellogenin concentrations in head/tail homogenate from zebrafish after exposure to prochloraz (measured concs) from 0-60 DPH at Lab 3. Female and male vitellogenin concentrations are presented in Figure A and B respectively. The boundary of the box closest to zero indicates the 25th percentile, a line within the box marks the median, and the boundary of the box farthest from zero indicates the 75th percentile. Whiskers above and below the box indicate the 90th and 10th percentiles. * Indicate significantly different from the control group ($p < 0.05$). The number of fish from each group is marked at the bottom of the figure.

Lab 4:

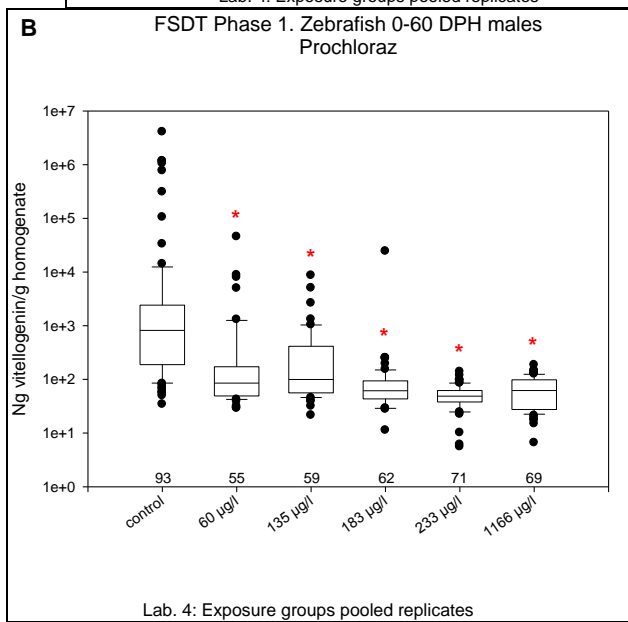
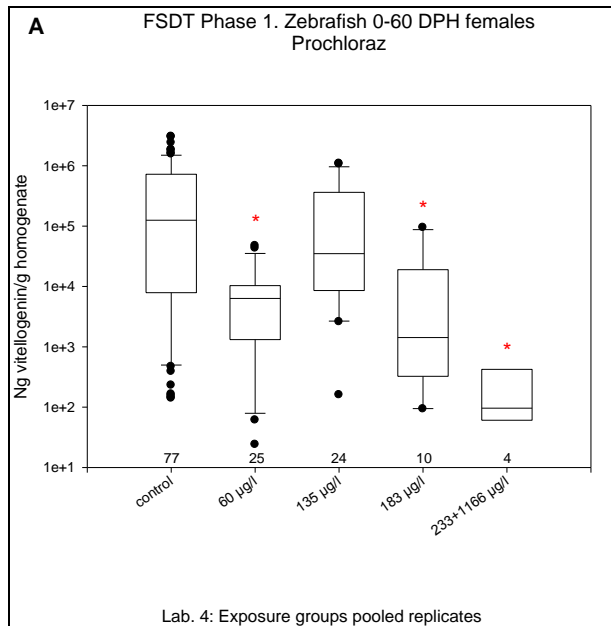


Figure 0-10: Vitellogenin concentrations in head/tail homogenate from zebrafish after exposure to prochloraz (measured concs) from 0-60 DPH at Lab 3. Female and male vitellogenin concentrations are presented in Figure A and B respectively. The boundary of the box closest to zero indicates the 25th percentile, a line within the box marks the median, and the boundary of the box farthest from zero indicates the 75th percentile. Whiskers above and below the box indicate the 90th and 10th percentiles. * Indicate significantly different from the control group ($p < 0.05$). The number of fish from each group is marked at the bottom of the figure. The data for the 4 female fish exposed to 233 and 1166 $\mu\text{g/l}$ prochloraz have been combined.

Sex ratio analysis

Sex ratio analysis of 4-tert pentylphenol studies

Fathead minnow

Lab 5

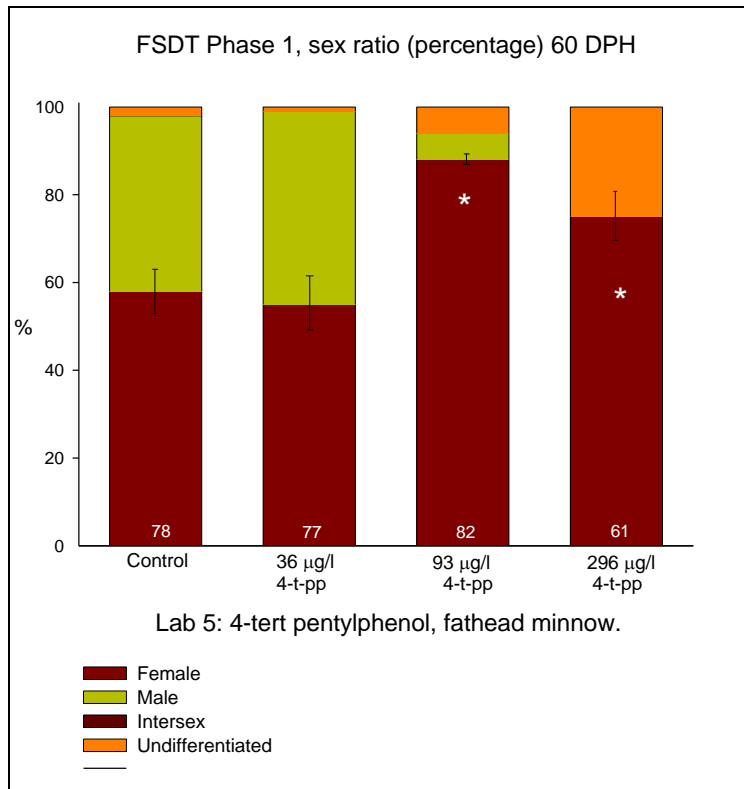


Figure 0-11 Replicate mean percentage of females, males, intersex and undifferentiated fathead minnow at 60 DPH (measured concs of tPP). Error bars denote standard deviation of percentage females from replicates. Asterisk symbolise significant different from control ($P=0.05$) calculated with Bonferroni adjusted χ^2 analysis.

Zebrafish

Lab 2:

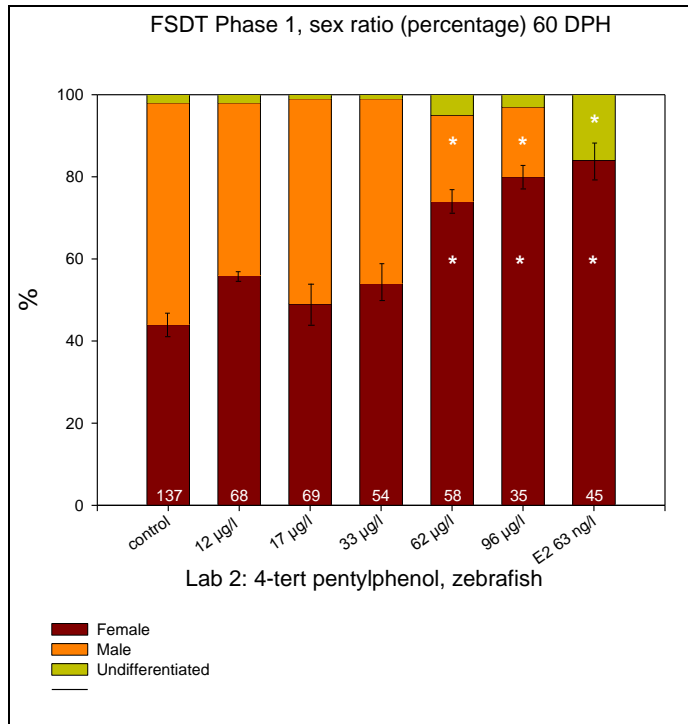


Figure 0-12 Replicate mean percentage of females, males, intersex and undifferentiated zebrafish at 60 DPH (measured concs of tPP). Error bars denote standard deviation of percentage females from replicates. Asterisk symbolise significant different from control (P=0.05) calculated with Bonferroni adjusted Chi² analysis.

Lab 3:

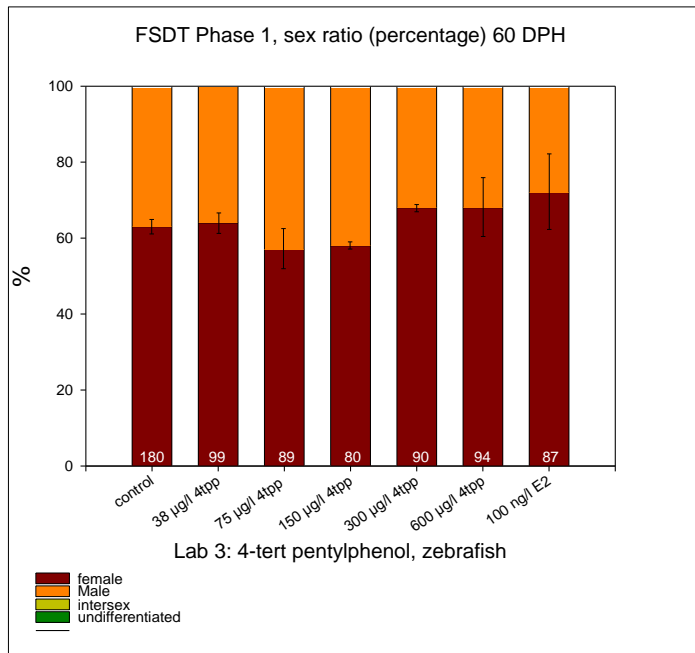


Figure 0-13 Replicate mean percentage of females, males, intersex and undifferentiated zebrafish at 60 DPH (nominal concs of tPP). Only nominal concentrations are shown as very few concentration measurements were made. Error bars denote standard deviation of percentage females from replicates. Asterisk symbolise significant different from control (P=0.05) calculated with Bonferroni adjusted Chi² analysis.

Lab 4:

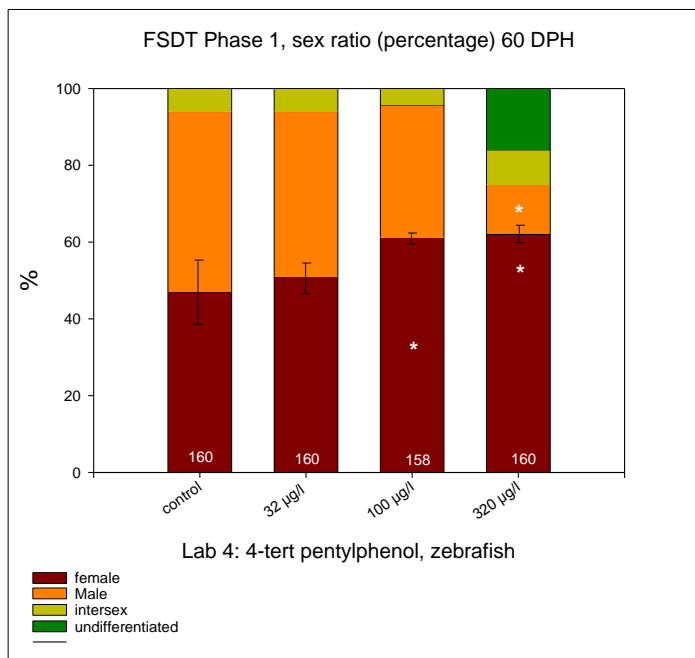


Figure 0-14 Replicate mean percentage of females, males, intersex and undifferentiated zebrafish at 60 DPH (nominal concs of tPP). Error bars denote standard deviation of percentage females from replicates. Asterisk symbolise significant different from control (P=0.05) calculated with Bonferroni adjusted Chi² analysis.

Sex ratio analysis of prochloraz studies

Fathead minnow

Lab 5

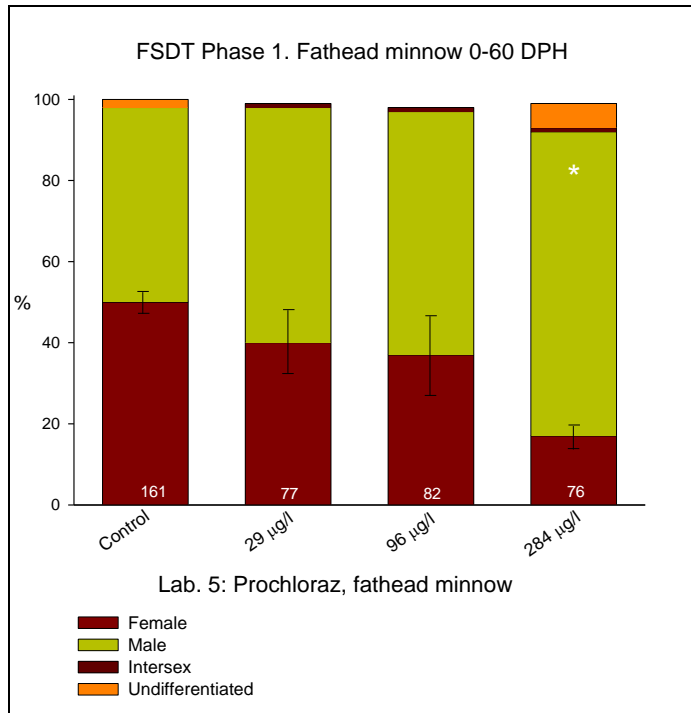


Figure 0-15 Replicate mean percentage of females, males, intersex and undifferentiated fathead minnow at 60 DPH (measured concs of prochloraz). Error bars denote standard deviation of percentage females from replicates. Asterisk symbolise significant different from control (P=0.05) calculated with Bonferroni adjusted Chi² analysis.

Lab 2:

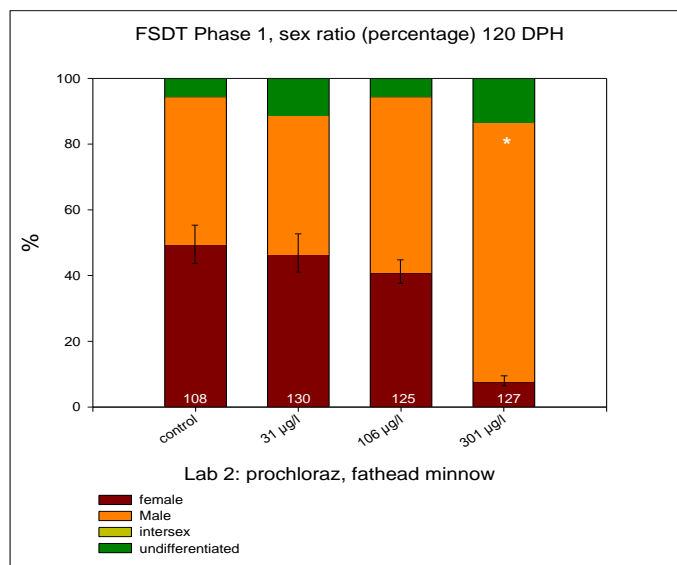


Figure 0-16 Replicate mean percentage of females, males, intersex and undifferentiated fathead minnow at 120 DPH (measured concs of prochloraz). Error bars denote standard deviation of percentage females from replicates. Asterisk symbolise significant different from control (P=0.05) calculated with Bonferroni adjusted Chi² analysis.

Zebrafish

Lab 2:

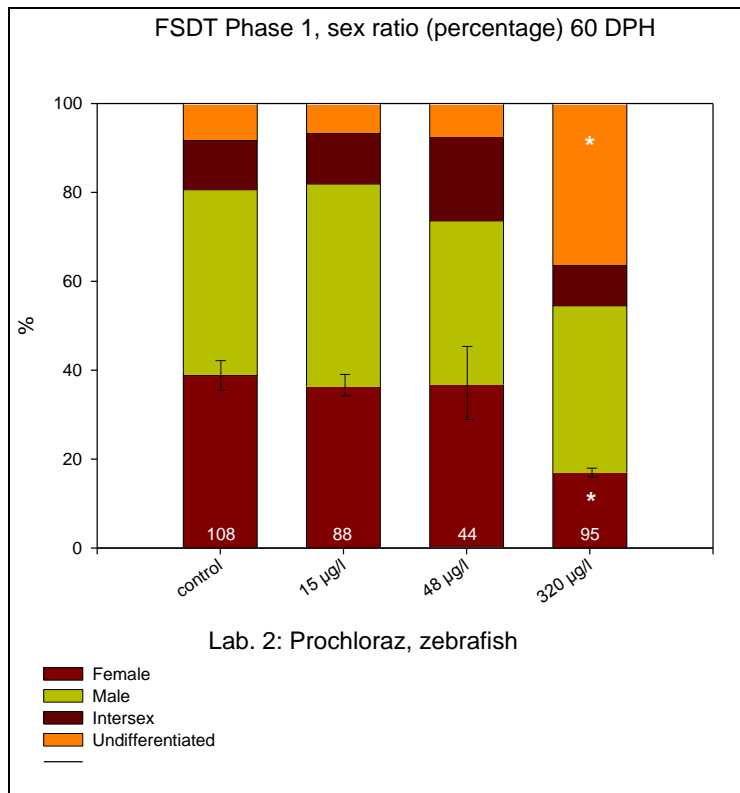


Figure 0-17 Replicate mean percentage of females, males, intersex and undifferentiated zebrafish at 60 DPH (measured concs of prochloraz). Error bars denote standard deviation of percentage females from replicates. Asterisk symbolise significant different from control ($P=0.05$) calculated with Bonferroni adjusted χ^2 analysis.

Lab 3:

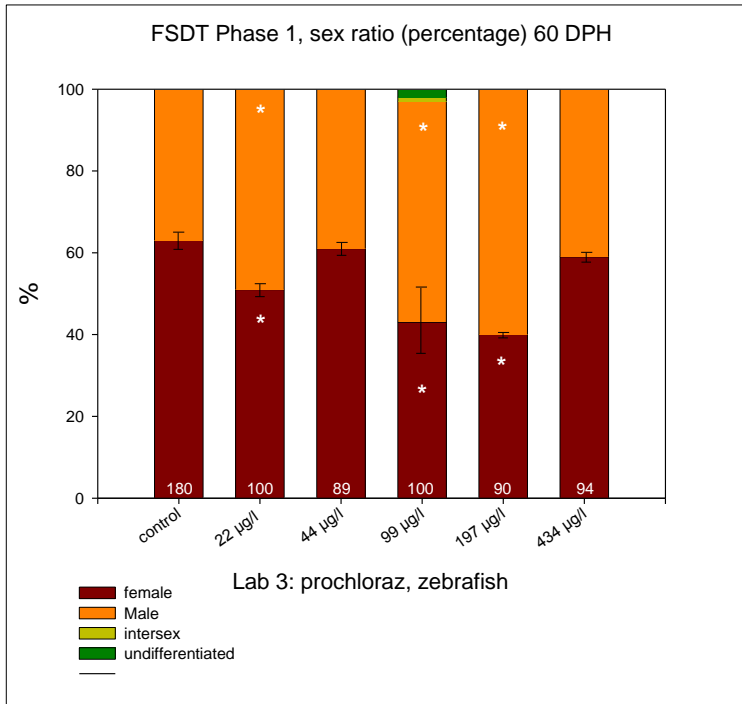


Figure 0-18 Replicate mean percentage of females, males, intersex and undifferentiated zebrafish at 60 DPH (measured concs of prochloraz). Error bars denote standard deviation of percentage females from replicates. Asterisk symbolise significant different from control (P=0.05) calculated with Bonferroni adjusted Chi² analysis.

Lab 4:

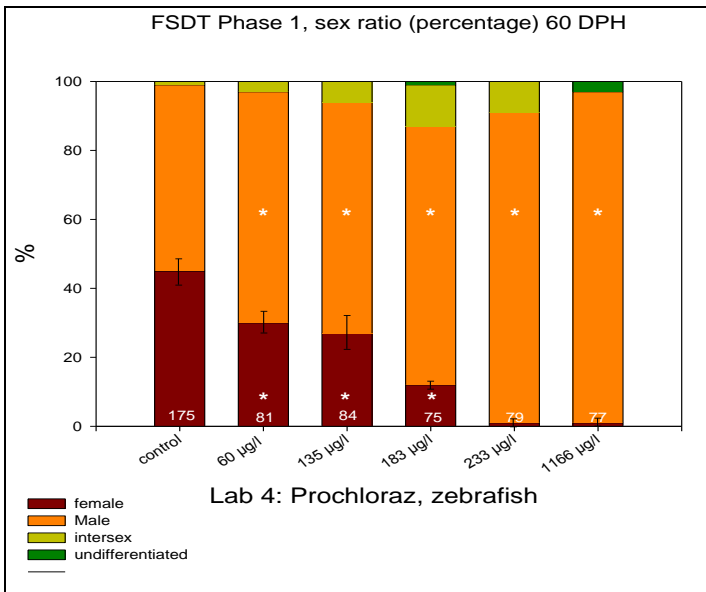


Figure 0-19 Replicate mean percentage of females, males, intersex and undifferentiated zebrafish at 60 DPH (measured concs of prochloraz). Error bars denote standard deviation of percentage females from replicates. Asterisk symbolise significant different from control (P=0.05) calculated with Bonferroni adjusted Chi² analysis.

Statistical analysis of sex ratio:

A complex re-evaluation of how to treat sex ratio data statistically has been produced by John W. Green, Ph.D., Ph.D., DuPont and Timothy A. Springer, Ph.D., Wildlife International. This data evaluation was made in three steps covering different data and data setups. The three initial reports are attached as separate pdf-files. A fourth and final analysis of data from all experiments (except for lab 2, 120 dph fathead minnow prochloraz exposure) was done by John W Green and is included in full as appendix 9 (section 0). The outcome of all Green and Springer's investigations is summarised in section 6.

Part 1: Analysis of two experiments with fathead minnow conducted at Bayer CropScience.

The experiments followed the NOEC/LOEC approach with three test concentrations and four replicates per concentration (plus 8 control replicates) and N=20. One experiment used prochloraz and one used 4-tert-pentylphenol as test chemical. The main purpose of the analysis was to compare NOEC with ECx approaches, and to identify the most robust statistical method.

Part 2: Analysis of two experiments with fathead minnow conducted at Bayer CropScience. Pooled control groups.

This re-analysis was conducted primarily to incorporate an improved understanding of how the 4-tert-pentylphenol and prochloraz experiments were conducted. In Part 1, each statistical analysis of these chemicals used the controls identified for that chemical. Subsequently, it was realised that the experiments were conducted in such a way that the controls identified for the two chemicals could equally well be applied to both experiments. Thus, all analyses were redone using the combined controls, keeping replicate vessel information. In addition, some simulation results were added and the discussions was expanded and re-organized.

The impact of the combined controls was seen in several ways: (a) The increase in the number of replicate vessels in the control increases the degrees of freedom of all statistical tests, tending to make the tests more sensitive. (b) The increased number of replicate vessels in the control is expected to give a better estimate of the control mean response. This can increase or decrease the likelihood of a treatment effect being found statistically significant, according to whether the combined control mean is farther from, or closer to, the treatment mean. The new control mean also affects the regression model's fit to the data and thus affects the estimated ECx. The revised ECx value could increase or decrease and a different regression model could be selected. (c) The increase in replication impacts the estimated pooled within-vessel variance, which in turn can increase or decrease the sensitivity of the tests according to whether the estimated variance is smaller, or larger, than before. All of the above effects were observed for some of the responses analyzed.

The conclusions of the statistical data analyses in Part 1 and 2 are presented in section 0

Part 3: Analysis of three experiments with zebrafish conducted at DHI and Heidelberg

The objective of Part 3 was essentially the same as Parts 1 and 2, but using zebrafish data. Three experimental designs were used to explore the sex ratio responses in zebrafish.

At DHI, both NOEC/LOEC and ECx designs were used. In the NOEC/LOEC design, there was a control with four replicates with variable numbers (22-34) of fish per replicate and three test concentrations (32, 100, and 320 ug/L) of prochloraz, with 2 replicates (at 32 ug/L) or 4 replicates (at 100 and 320 ug/L) each with variable numbers (16-28) of fish. In the ECx design, there was a control with four replicates and five test concentrations (38, 75, 150, 300, and 600 ug/L) of 4-t-PP with two replicates each. The number of fish per replicate varied from 18 to 35, with most replicates having 35 fish at the end of the experiment.

There was also a positive control with two replicates. The positive control was ignored in the statistical analyses to determine the LOEC or estimate an ECx.

In the Heidelberg laboratory, they used a control with four replicates, and five test concentrations (37.5, 75, 150, 300, and 600 ug/L) of prochloraz with two replicates each. The number of fish per replicate varied from 38 to 48.

There was only one control group reported in each experiment, so no comparison of controls was needed.

The conclusions of Part 3 are presented in section 6.3.

DISCUSSION

Literature review of Fathead minnow vitellogenin

Based on the outcome of validation measurements with fathead minnow performed at 60 dph by Lab 5 (Figure 0-1 and Figure 0-5), which in several cases showed a lack of expected responsiveness to 4-t-PP and prochloraz, the length of the exposure phase was discussed at the VMG-eco Meeting in Madrid 2007, and it was suggested that 60 dph fathead minnows may not be sufficiently mature to make a full Vtg response. VMG-eco decided that Denmark should undertake a literature review of the vitellogenin response in young fathead minnows. The outcome of the literature review is given below in Tables 6-1 and 6-2.

Table 6-1. Outcome of literature review concerning Vtg induction in fathead minnows at various times post-hatch.

Fathead minnow vitellogenin		Most data are visual estimated concentrations based on figures from the publications!				
Age at sampling	Exposure regime	mean control Vtg concentration	Mean exposure Vtg concentration	Method	sex	reference
30 dph	17β-estradiol 0-30 dph	100-300 ng/ml WBH	10000-30000 ng/ml WBH	Carp ELISA (Tyler et al 1999)	undifferentiated	Tyler et al 1999
172 dph	EE2 (10 ng/l) 0-20 dph	~100 ng/g WBH		Carp RIA (Tyler et al 1996)	male	Länge et al 2001
172 dph	EE2 (10 ng/l) 0-20 dph	~10000 ng/g WBH		Carp RIA (Tyler et al 1996)	female	Länge et al 2001
30 dph	EE2 (10 ng/l) 0-20 dph	917 ng/ml WBH	7000 ng/ml WBH	Carp ELISA (Tyler et al 1999)	undifferentiated	Van Aerle et al 2002
< 3 month		0-5000 ng/ml WBH		Carp ELISA (Tyler et al 1999)	undifferentiated (mixed sex)	Zerulla et al 2002
45-100 days	DES 4-21 days	400 ng/ml WBH	51000 ng/ml WBH	Carp ELISA (Tyler et al 1999)	undifferentiated (mixed sex)	Panter et al 2006
107 dph	4 tert pentylphenol (56-560 µg/l) 0-107 dph	~500 ng/ml plasma	500~1000 ng/ml plasma	Carp ELISA (Tyler et al 1999) (Standards and IgG from Biosense)	male	Panter et al 2006
107 dph	4 tert pentylphenol 0-107 dph	~2300 ng/ml plasma	~2000 µg/ml plasma	Carp ELISA (Tyler et al 1999) (Standards and IgG from Biosense)	female	Panter et al 2006
107 dph	EE2 (10 ng/l) 0-107 dph	~2300 ng/ml plasma	~17000 µg/ml plasma	Carp ELISA (Tyler et al 1999) (Standards and IgG from Biosense)	female	Panter et al 2006
63 dph	EE2 (10ng/l) 0-63 dph	<5 ng/ml plasma	~300 µg/ml plasma	Biosense Carp ELISA	male	Bogers et al 2006
63 dph	EE2 (10ng/l) 0-63 dph	~100 ng/ml plasma	~700 µg/ml plasma	Biosense Carp ELISA	female	Bogers et al 2006
114 dph	EE2 (10ng/l) 0-114 dph	~50 ng/ml plasma	no males	Biosense Carp ELISA	male	Bogers et al 2006
114 dph	EE2 (10ng/l) 0-114 dph	200000 ng/ml plasma	~3000 µg/ml plasma	Biosense Carp ELISA	female	Bogers et al 2006
60 dph	4 tert pentylphenol (32-320 µg/l) 0-60 dph	~4 ng/ml head/tail homogenate	~6 ng/ml head/tail homogenate	Biosense FHM ELISA	male	Bayer CropScience 2006 FSDT Phase I
60 dph	4 tert pentylphenol (32-320 µg/l) 0-60 dph	~148 ng/ml head/tail homogenate	~36 ng/ml head/tail homogenate	Biosense FHM ELISA	female	Bayer CropScience 2006 FSDT Phase I
60 dph	prochloraz (32-320 µg/l) 0-60 dph	~4 ng/ml head/tail homogenate	~3 ng/ml head/tail homogenate	Biosense FHM ELISA	male	Bayer CropScience 2006 FSDT Phase I
60 dph	prochloraz (32-320 µg/l) 0-60 dph	~148 ng/ml head/tail homogenate	~4 ng/ml head/tail homogenate	Biosense FHM ELISA	female	Bayer CropScience 2006 FSDT Phase I
Adult		~50 ng/ml plasma		Carp ELISA (Tyler et al 1999)	male	Pawlowski et al 2007
Adult		~50000 ng/ml plasma		Carp ELISA (Tyler et al 1999)	female	Pawlowski et al 2007
Adult		~300 ng/ml plasma		Carp ELISA (Tyler et al 1999)	male	Brian et al 2007
Adult		~100000 ng/ml plasma		Carp ELISA (Tyler et al 1999)	female	Brian et al 2007
Adult		1500 ng/ml plasma		Biosense FHM ELISA	male	Kunz et al 2006
Adult		2000000 ng/ml plasma		Biosense FHM ELISA	female	Kunz et al 2006
Adult		~4000 ng/ml plasma		EnBioTec FHM ELISA	male	Seki et al 2006
Adult		~10000000 ng/ml plasma		EnBioTec FHM ELISA	female	Seki et al 2006
Adult		~15000000 ng/ml plasma		FHM ELISA (Parks et al 1999)	female	Jensen et al 2006

Note 1: Eidem et al 2006 observed no significant difference between Vtg concentrations in plasma versus whole body homogenate (WBH), whereas Länge et al 2001 observed a higher concentration in plasma than in WBH.

Note 2: Most of the presented Vtg values are rough estimates based on figures from the respective publications

Table 6-2. Summary of the literature review of control fathead minnow Vtg data.

summary	
Sex and age	Vtg
males 60 dph	4-5 ng/ml homogenate
males 100-170 dph	50-500 ng/ml plasma
adult males	50-4000 ng/ml plasma
females 60 dph	100-150 ng/ml
females 100-170 dph	2300-200000 ng/ml
adult females	50-15000 µg/ml plasma
mixed sex 30 dph	200-900 ng/ml homogenate
mixed sex 45-100 dph	400-2500 ng/ml homogenate

Table 6-2 summarises variations in Vtg levels from the different fathead minnow studies listed in Table 6-1. Because the Vtg concentration is dependent on more factors than age and sex, e.g. the growth of the fish, these results can only be used as rough indicators of the “Vtg to age relationship”.

The outcome of the review of fathead minnow Vtg can be summarised as follows:

- Females at 60 dph contain in general very low levels of Vtg and are not suitable for experiments where a reduction in the Vtg concentration is expected (e.g. aromatase inhibition). Their VTG levels are responsive to strong estrogens, but they do not appear very responsive to weak estrogens.
- Males at 60 dph contain very low levels (<10 ng/ml) of Vtg and a rise in the vitellogenin concentration would be easily registered. Their Vtg levels appear to be responsive to strong estrogens, but not to weak estrogens or weak aromatase inhibitors.
- Females at 110 dph contain Vtg in the µg/ml range and are suitable for vitellogenin reduction studies.
- Males at 110 dph contain Vtg in the range below 1 µg/ml and are suitable for Vtg induction studies.
- Adult females contain Vtg from the µg/ml to mg/ml range.
- Adult males in general contain Vtg below 1µg/ml as do the 110 dph males.

As much information and data might not be published yet, we would encourage laboratories working with fathead minnow to comment and extend the data-set presented in this review.

Based on the current information it is recommended that the next FSDT fathead minnow study should be run for 120 days post hatch at which stage the fish have reached maturity (according to Länge et al., 2001) and are fully responsive to strong and weak estrogens and aromatase inhibitors. Such a study was finalized by Lab 2 in October 2007 and the results are presented in Figure 0-6 and Figure 0-7. The results support the recommended 120 day exposure period for fathead minnow because female median control Vtg levels increased by more than 4 orders of magnitude between 60 and 120 dph, and median Vtg levels in 120 dph females decreased by 4 orders of magnitude when the fish had been exposed to the top concentration of prochloraz. When fully analysed, these data will supply us with the needed information regarding Vtg concentration versus age in juvenile fathead minnow.

Phase 1 Vitellogenin data

It should be noted that all data (both Vtg and sex ratio) were analysed in a blinded fashion.

4-tert pentylphenol

Induction of Vtg by estrogens in juvenile fish is well known [Andersen et al. 2003; Bogers et al. 2006; Holbech et al. 2006]. Unfortunately, only two laboratories (Labs 2 and 3) submitted positive control data for estradiol (E2). In Lab 2, female zebrafish Vtg concentrations increased in response to E2 as expected, but no data for males were submitted. In Lab 3, neither males nor females gave a Vtg response to E2. This to some extent throws doubt on the responsiveness of the experimental fish used by several laboratories.

[Difficulties in maintaining test concentrations]

The results obtained so far on Vtg induction in response to estrogen and Vtg depression due to aromatase inhibition are summarised in Table 6-3.

Table 0-3: Vitellogenin response (NOEC and LOEC values) obtained in zebrafish and fathead minnow studies, respectively. M = measured; N = nominal; ♀ = based on females; ♂ = based on males. * = non-monotonic concentration-response

Lab No.	4-tert pentylphenol (µg/L)				Prochloraz (µg/L)			
	Zebrafish		Fathead minnow		Zebrafish		Fathead minnow	
	NOEC	LOEC	NOEC	LOEC	NOEC	LOEC	NOEC	LOEC
1	-	-	-	-	-	-	-	-
2	-	>96 (M) (♀+♂)	-	-	4 8 (M) (♀)	3 20 (M) (♀)	<31 (M) (♀)	31 (M) (♀)
3	300* (N) (♀)	600* (N) (♀)	-	-	22 (M) (♂) 44* (M) (♀)	44 (M) (♂) 99* (M) (♀)	-	-
4	32 (N) (♂)	100 (N) (♂)	-	-	<60 (M) (♂) <60* (M) (♀)	60 (M) (♂) 60* (M) (♀)	-	-
5	-	-	36 (M) (♀)	93 (M) (♀)	-	-	<29 (M) (♀)	29 (M) (♀)

Females

Due to problems with maintaining concentrations of 4-tert pentylphenol in the present experiments combined with the natural high concentrations of Vtg in females, only one experiment (Lab 3) could define a LOEC for induced Vtg in females (zebrafish) exposed to 4-tert pentylphenol: 600 µg/l (nominal concentration), although even in this case there was a non-monotonic concentration-response relationship. Reduced Vtg was observed (LOEC 93 µg/l) in female fathead minnow by Lab 5. This effect cannot be explained as an estrogenic effect of 4-tert pentylphenol and one might be tempted to explain the results as being caused by some other endocrine mode of action, because substances with multiple modes of endocrine action are not unknown. However, if the observation of declining Vtg is combined with the sex ratio results (Figure 0-11) it becomes clear that the reduced Vtg concentration is caused by the shift in sex ratio toward females. The new females (probably “genetic” males) have less-developed gonads and also less Vtg than “normal” females, thereby reducing the mean Vtg concentration.

Males

Induction of Vtg in males was observed at 100 and 320 µg/l, with a NOEC of 32 µg/l, (nominal concentrations) by Lab 4, but otherwise the problems with maintaining the concentrations of 4-tert pentylphenol resulted in failed induction. The highest measured concentrations of 93 µg/l (Lab 5) and 96 µg/l (Lab 2) are lower than 4-tert pentylphenol concentrations known to induce Vtg in male fathead minnow [Panter et al. 2006] and in juvenile zebrafish (38 dph (unpublished results, SDU, Denmark)).

Prochloraz

Inhibition of Vtg production in females is known to be a biomarker of exposure to aromatase inhibitors [Kinnberg et al. 2007]. The conversion of endogenous androgen to estrogen is blocked via inhibition of CYP19 aromatase, thus stopping estrogen receptor activation, itself controlling Vtg production.

Females

The present results from both fathead minnow and zebrafish experiments verified vitellogenin as a biomarker for aromatase inhibition in females. A significant reduction in female Vtg concentration was seen in all Phase 1 prochloraz experiments. The LOECs were 29 µg/l (Lab 5) and 31 µg/l (Lab 2) for fathead minnow and between 60 and 320 µg/l for zebrafish, which for both species are below LOECs resulting from adult exposure to prochloraz [Ankley et al. 2005; OECD 2006]. However, note that the concentration-response relationships for Labs 3 and 4 were non-monotonic.

As discussed in paragraph 0, the Vtg concentration in female fathead minnow at 60 DPH were lower than expected and not much differentiated from male concentrations. This observation gave rise to the 0-120 DPH experiment performed by Lab 2 and in Figure 0-7 it can be seen that female control Vtg levels rose by a factor of 1000 or more from 60 to 120 DPH whereas male Vtg also rose by about a factor of 5. These results support the recommended 120 day exposure period for fathead minnow because the higher female control vitellogenin levels increase the sensitivity to detect both aromatase inhibitors and anti-estrogenic chemicals.

Males

A decline in the Vtg concentration were observed at 29 µg/l prochloraz in fathead minnow but with no concentration-response (Lab 5) and at 44 µg/l prochloraz (Lab 3 LOEC) and 60 µg/l prochloraz (Lab 4 LOEC) in zebrafish. A consistent concentration response relationship was not observed overall which might be due to the concentration-dependent change in sex ratio: any specimens that appeared to be phenotypic males but in reality were genetic females because of the prochloraz exposure might have different Vtg metabolism and therefore these individuals could have caused a rise in the Vtg concentration variance compared to the low natural male Vtg concentration in control groups. Another

explanation could be the known multiple modes of endocrine action of prochloraz other than aromatase inhibition (i.e. anti-androgenic ([Vinggaard et al. 2002])).

Phase 1 Sex ratio data

The NOEC/LOEC values shown in this section are derived using the original statistical methods described in section 3.15.2 (i.e. pairwise comparison with controls of the proportion of females, using the chi-squared test with Bonferroni’s correction). The improved statistical re-evaluation of these data by John Green which is shown in full in Appendix 9 produced differences in some NOECs which will be described in section 6.4.

Sex ratio or proportions of sex were determined according to the mode of action of the test substances: for 4-tert pentylphenol experiments, proportions were defined as females or “not females” due to the estrogenic mode of action of 4-tert pentylphenol and the same was done for prochloraz due to the aromatase inhibiting mode of action of prochloraz. A more detailed discussion of the endpoint can be found in section 0. Unfortunately, only two laboratories (Labs 2 and 3) submitted positive control data for E2. In one case (Lab 2), E2 caused the expected skewing of the zebrafish sex ratio towards females, but E2 caused no response in Lab 3 zebrafish. As with the Vtg measurements (see above), this to some extent throws doubt on the responsiveness of the fish used by several laboratories.

The results obtained so far on changes in sex ratio are summarised in Table 6-4.

Table 0-4: Response in Sex ratio (NOEC and LOEC values) obtained in zebrafish and fathead minnow studies, respectively. M = measured concentrations; N = nominal concentrations; ** = non-monotonic concentration-response. Note that these data were analysed using the original statistical procedure described in section 3.15.2 and not the improved methods recommended by Green and Springer.

Lab No.	4-tert pentylphenol				Prochloraz			
	Zebrafish		Fathead minnow		Zebrafish		Fathead minnow	
	N OEC	L OEC	N OEC	L OEC	N OEC	L OEC	N OEC	L OEC
1	-	-	-	-	-	-	-	-
2	3 3 (M)	6 2 (M)	-	-	4 8 (M)	3 20 (M)	1 06 (M)	3 01 (M)
3	-	> 600 (N)	-	-	> 434** (M)	> 434** (M)	-	-
4	3 2 (N)	1 00 (N)	-	-	< 60 (M)	6 0 (M)	-	-
5	-	-	3 6 (M)	9 3 (M)	-	-	9 6 (M)	2 84 (M)

*The experiment has not yet been finally analysed (December 2007)

4-tert pentylphenol

A sex ratio skewed toward females was seen at 93 µg/l 4-tert pentylphenol for fathead minnow (Lab 5 LOEC) and at 62 µg/l (Lab 2 measured LOEC) and 100 µg/l (Lab 4 nominal LOEC) for zebrafish. No effect on sex ratio was observed by Lab 3 (LOEC > 600 µg/l nominal). In general sex ratio tends to be more sensitive and appears to react more consistently to a weak estrogen such as 4-tert pentylphenol than the Vtg concentration.

Prochloraz

A concentration-dependent skewing of the sex ratio toward fewer females was observed in all experiments except Lab 3, where the highest concentration of 434 µg/l prochloraz did not skew the zebrafish sex ratio (although some concentrations viz. 22, 99 and 197 µg/l prochloraz did appear to cause a skew). For fathead minnows, the LOECs for skewed sex ratio were 284 and 301 µg/l prochloraz, and for zebrafish the LOECs derived from monotonic concentration-responses were 320 and 60 µg/l prochloraz. It is important to notice that only four water samples were analysed from each exposure concentration in this experiment compared to more than the double from the other experiments.

Sex ratio statistics

This section will briefly present the initial conclusions reached by John Green and Tim Springer in their report parts 1-3. It will then present John Green's final conclusions based on almost all of the Phase 1 dataset, the report on which is reproduced in full in Appendices 9 and 10.

Green and Springer parts 1 and 2 conclusions

- These conclusions are based on analysis of two fathead minnow datasets, one for 4-tPP and one for prochloraz.
- The preferred endpoint is either the proportion of males or females, and not the sex ratio as such. The proportion of intersex or undifferentiated fish would also be appropriate endpoints. However, a definitive decision to use proportions requires analysis of more datasets.
- The ECx approach for the proportion of males or females is not recommended as it generally gives very unreliable results with these data. The reason for this is that the confidence bounds on the regression model are so wide as to render the ECx estimates meaningless. This in turn is a result of the high background incidence rate (approximately 50%) of phenotypically male (or female) fish. Were the genetic sex of the fish known, then analysis of the proportion of fish whose phenotypic sex differs from their genetic sex would not suffer from this high background incidence.
- If ECx must be used, then it is most appropriate with this type of information to use a probit model or one of the standard models (e.g. Bruce-Versteeg) with replicate proportions or replicate sex ratios. However, none of these models are free of problems.
- The preferred NOEC approach is best conducted with the Jonckheere or Williams tests based on replicate means, providing the number of organisms does not vary much and the concentration-response is approximately monotonic.
- If the concentration-response is not approximately monotonic, Fisher's exact test with Bonferroni-Holm adjusted p-values would be the simplest alternative method for determining the NOEC, although other tests (e.g. the more powerful and scientifically more appropriate Rao-Scott version of the Cochran-Armitage test) could be considered.

Green and Springer Part 3 conclusions

- These conclusions are based on analysis of 3 zebrafish datasets, one for 4-tPP and two for prochloraz.

- Only one control group was reported in each experiment, so no comparison of controls was undertaken.
- Replicate tanks were the unit of analysis, and the endpoints used were the proportions of the 4 sex categories.
- It was not possible to fit a regression (i.e. EC_x) model to some of the data, and for other data the model was unstable, giving very wide confidence bounds.
- The tests used for determining the NOEC/LOEC were Dunnett's test on transformed proportions, the Jonckheere-Terpstra test applied in step-down fashion, or Dunn's test if no normalising transform was found. Serious deviations from monotonicity precluded the use of Jonckheere's test.
- As with the part 1 and 2 analyses, the Jonckheere test generally gave acceptable results with transformed proportions, thus confirming the earlier tentative conclusions. Power tests of this approach were not undertaken, but earlier power simulations suggest that it has a power of greater than 70% to detect a change of 25% in the proportion of males or females.

John Green's final conclusions

John Green's complete analysis of all Phase 1 datasets (except the fathead minnow/ prochloraz data from Lab 2) is presented in full in Appendices 9. The conclusions of this analysis do not differ substantially from those arrived at from the partial analyses described above. In summary, the best approach is to use a NOEC/LOEC design, to combine the control data if the solvent and water controls do not differ according to a Mann-Whitney test (retaining separate identification of replicate controls), to use the arcsine square root-transformed proportions of each sex category (rather than sex ratio) as endpoints, and to employ the Jonckheere-Terpstra test (if the data are at least approximately monotonic) to identify the NOEC/LOEC. If the data are not monotonic (such as is often the case in these data for the proportion of intersex or undifferentiated fish), then the best approach for identifying the NOEC/LOEC is to conduct a pairwise Dunnett's test (or if the requirements for Dunnett's are not met, to use alternatives – the Tamhane-Dunnett test or the exact version of the Mann-Whitney test with Bonferroni-Holm adjustment).

These procedures have been clearly set out by John Green in the flow chart which is reproduced in Appendix 10.

Revised sex ratio NOECs using the improved statistical procedure

The revised sex ratio NOECs (taken from John Green's final report, and identified largely by Jonckheere's test) based on the proportion of females, are shown below in Tables 6-5 and 6-6. The original NOECs derived using the Chi-squared test are also shown for comparison.

Table 6-5. Revised results for prochloraz based on the proportion of females, and using Jonckheere's test (2-sided).

Laboratory Number and Species	Revised NOEC (µg/l)	Original NOEC (µg/l)	Comments
2 (zebrafish)	48.1	48	Original NOEC: nominal 100 µg/l; measured 48.0 µg/l
3 (zebrafish)	>197.4	>434	Original NOEC: nominal 600 µg/l; measured 434.1 µg/l Significant (chi-squared) responses at 22, 99 and 197 µg/l; no responses at 44 and 434 µg/l
4 (zebrafish)	59.7	<60	Original LOEC: nominal 38 µg/l; measured 59.7 µg/l
5 (fatheads)	96.4	96	Original NOEC: nominal 100 µg/l;

			measured 96.2 µg/l
--	--	--	--------------------

Table 6-6. Revised results for 4-tert-pentylphenol based on the proportion of females, and using Jonckheere's test (2-sided).

Laboratory Number	Revised NOEC (µg/l)	Original NOEC (µg/l)	Comments
2 (zebrafish)	34.0	33*	Original NOEC: nominal 150 µg/l; measured 34.0 µg/l
3 (zebrafish)	>306.5	>600	Original NOEC: nominal 600 µg/l; measured 23.2 µg/l.
4 (zebrafish)	>227.2	32	Original NOEC: nominal 32 µg/l; measured 2.0 µg/l Note that the revised NOEC based on the proportion of males is 71.0 µg/l
5 (fatheads)	36.1	36	Original NOEC: nominal 32 µg/l; measured 36.1 µg/l

*Taken from Fig. 4-12. This is probably a typographical error and should be 34.0 µg/l

The revised and original NOECs based on the proportions of females appear broadly comparable, with the notable exception of the Lab 4 experiment in which zebrafish were exposed to 4-t-PP. In this case, Jonckheere's test failed to establish a NOEC (>227 µg/l nominal), while the Chi-squared test identified a NOEC of 32 µg/l nominal. Interestingly, the revised NOEC based on the proportion of males was 71 µg/l, which is much closer to the original NOEC (based on the proportion of females) of 32 µg/l.

In general terms, therefore, the two statistical approaches appear to be producing broadly similar results when all data are considered, but the use of Jonckheere's test is considered to be more statistically sound in that (1) it is the only test discussed that makes use of the important monotone dose-response, (2) it has been shown through extensive computer simulations to be have greater power to detect real effects, (3) it is not sensitive to aberrations in the data at intermediate test concentrations that are not reflected in higher test concentrations, and (4) it has minimal requirements for its use other than approximate dose-response monotonicity (e.g., normality and variance homogeneity are not required, nor is any transformation of the data).

John Green has also demonstrated that a smaller experimental design will provide satisfactory statistical power to detect 10-25% *sexual inversion*, rather than change in the proportion of phenotypic sex. That is, where genetic sex is determined (which is now possible for stickleback, medaka and fathead minnow), the analysis of the proportion of fish whose phenotypic sex differs from their genetic sex is more sensitive than analysis based on proportion of phenotypically male or female fish. In particular, 4 replicate tanks of 20 fish each are sufficient. This assumes 10 fish of each sex, as changes from male to female or female to male may both be of interest. It has been shown that changes in either sex of 10-25% will be found statistically significant with high probability using the step-down Jonckheere-Terpstra test. Lower power is achieved by Dunn's test, the highest power theoretically acceptable test to use when a monotone dose-response cannot be assumed.

OVERALL CONCLUSIONS AND RECOMMENDATIONS ON THE FISH SEXUAL DEVELOPMENT TEST (FSDT)

FSDT exposures start with newly fertilized eggs and continue until gonadal sex differentiation has finished. Besides hatching success, growth and survival, two endocrine biomarkers are measured: i) Vitellogenin (Vtg) levels and ii) Gonadal sex (determined as male, female, intersex or undifferentiated). Gonadal sex is presented as sex ratio or proportions (e.g. males or females). Gonadal histopathology can be included (as an option), as can secondary sexual characteristics in fathead minnow and medaka.

Results from Phase 1:

- The FSDT Phase 1 validation has demonstrated that this protocol can detect estrogenic effects of weak estrogens (4-tert pentylphenol) as well as aromatase inhibitors (prochloraz), although variability of endpoints is generally high.
- However, there is some uncertainty about the sensitivity to oestrogens of the fish used in Phase 1, as few positive control data (E2) were reported, and some positive controls failed to produce either a Vtg or sex ratio response. Furthermore, other endpoints (i.e. hatching success, survival, growth, secondary sex characteristics and gonadal histopathology) have not yet been reported.
- The sex ratio- and vitellogenin concentration changes are in combination promising biomarkers for estrogenic and aromatase inhibiting exposure in zebrafish and fathead minnow. However, the optimum numbers of concentrations and replicates required to achieve acceptable discriminatory power have not yet been finalised
- Overall, sex ratio seems to be the most robust and population relevant endpoint although it appears to be of similar sensitivity to Vtg. Sex ratio is less variable than Vtg, but it may nevertheless be desirable to increase the numbers of replicates in order to minimise the effects of variability caused by the way in which sexual condition is categorised.
- Fathead minnow should be exposed until 120 days post hatch to guarantee a female control vitellogenin concentration significantly above male control vitellogenin concentrations.
- Fathead minnow and zebrafish are comparably sensitive to 4-tert pentylphenol and prochloraz under FSDT exposure scenarios.
- ~~Sex ratio change toward females was the most sensitive endpoint for detecting the estrogenic effect of 4-tert pentylphenol.~~
- Vitellogenin concentration decline in female fathead minnow appears to be the most sensitive endpoint for detecting the aromatase inhibiting effect of prochloraz, although the sample on which this statement is based is very small. Furthermore, it should be noted that variability of Vtg measurements between replicate fish is very high, which reduces the discriminatory power of this endpoint.
- The NOEC/LOEC approach is preferred to the ECx approach because sex ratio (or proportions of sex categories) is impossible to analyze with regression analysis in most cases unless genetic sex can be determined. This suggests that the number of concentrations tested may need to be at least 5 plus a control unless the threshold effect level has already been well-characterised.
- The preferred method of identifying the NOEC is to use Jonckheere's test on the proportions of males, females, intersexes or indeterminate sexes, and not to analyse the sex ratio as such. However, Jonckheere's test should only be used for data that are at least approximately monotonic. If data are non-monotonic but normally distributed, one should use the Dunnett or Tamhane-Dunnett test. If data are neither monotonic nor normally distributed, one should use non-parametric pairwise comparisons such as Dunn's test with a Bonferroni-Holm adjustment.

- Some of the water sample analyses failed due to incorrect storage of samples. Chemical analysis of water samples is important and preservation (e.g. freezing) of water samples must be performed immediately after removal from exposure units.
- Actual concentrations of 4-tert pentylphenol within the range 80-120% of the nominal concentrations were difficult to measure in several experiments. This might partly be due to inadequate storage conditions.
- Inter laboratory coefficient of variation: Due to the Phase 1 test design, where two species and two experimental designs were conducted by only four laboratories (Lab 1 did not perform exposure experiments) a specific inter laboratory comparison has not been possible.

Possible options for Phase 2

Recommendations

- The most important objective for Phase 2 is to firmly establish the inter-laboratory reproducibility of the main endpoints of the FSDT, as no reliable data on this point have yet been generated.
- Another major objective of future validation studies should be to establish the optimum number of concentrations and replicates required to achieve acceptable statistical power. A NOEC/LOEC approach with at least 5 test concentrations should be followed by all laboratories in Phase 2, positive controls should always be used, and test concentrations should always be verified by adequate numbers of measurements employing reliable sampling, sample storage and analytical techniques.
- Based on the Phase 1 validation, the exposure period for fathead minnow should be extended to 120 days.
- The ability of the FSDT to detect androgenic substances has previously been shown in pre-validation studies performed by Denmark. It is therefore recommended that the FSDT is further validated in a Phase 2 with a negative substance (one laboratory), a weak androgen (probably di-hydro testosterone (DHT) and 4-tert octylphenol (to follow FSA validation).
- Ways of reducing the variability in Vtg measurements should be explored. For example, it may be possible to minimise variability by only measuring Vtg in a single organ such as liver, or by confining measurements to those fish whose phenotypic sex has not been altered (identified by using genetic sex markers available for medaka and fathead minnow). The tradeoff between reduced variability and smaller sample sizes on the power to detect effects would have to be evaluated carefully.
- Only two species have been used in the Phase 1 validation and the capability of the FSDT in relation to other species as e.g. Japanese medaka and stickleback should be evaluated during Phase 2.
- The additional value (if any) of measuring secondary sexual characteristics (and possibly also gonadosomatic index) should be investigated. The value of measuring gonadal histopathology should also be investigated if resources permit. Initial investigations on limited datasets were done by John Green which indicated low statistical power to detect all but large effects on secondary sex characteristics using the current experimental design.

REFERENCES

1. Andersen,L., H.Holbech, A.Gessbo, L.Norrgrén, and G.I.Petersen, 2003. Effects of exposure to 17 alpha-ethinylestradiol during early development on sexual differentiation and induction of vitellogenin in zebrafish (*Danio rerio*). *Comparative Biochemistry and Physiology C-Toxicology & Pharmacology* **134**: 365-374.
2. Ankley,G.T., K.M.Jensen, E.J.Durhan, E.A.Makynen, B.C.Butterworth, M.D.Kahl, D.L.Villeneuve, A.Linnum, L.E.Gray, M.Cardon, and V.S.Wilson, 2005. Effects of two fungicides with multiple modes of action on reproductive endocrine function in the fathead minnow (*Pimephales promelas*). *Toxicological Sciences* **86**: 300-308.
3. Ankley,G.T., K.M.Jensen, M.D.Kahl, J.J.Korte, and E.A.Makynen, 2001. Description and evaluation of a short-term reproduction test with the fathead minnow (*Pimephales promelas*). *Environmental Toxicology and Chemistry* **20**: 1276-1290.
4. Asahina,K., A.Urabe, T.Sakai, H.Hirose, and T.Hibiya, 1989. Effects of Various Androgens on the Formation of Papillary Processes on the Anal Fin Rays in the Female Medaka *Oryzias-Latipes*. *Nippon Suisan Gakkaishi* **55**: 1871.
5. Bogers,R., E.Mutsaerds, J.Druke, D.F.De Roode, A.J.Murk, B.Van der Burg, and J.Legler, 2006. Estrogenic endpoints in fish early life-stage tests: Luciferase and vitellogenin induction in estrogen-responsive transgenic zebrafish. *Environmental Toxicology and Chemistry* **25**: 241-247.
6. Holbech,H., K.Kinnberg, G.I.Petersen, P.Jackson, K.Hylland, L.Norrgrén, and P.Bjerregaard, 2006. Detection of endocrine disrupters: Evaluation of a Fish Sexual Development Test (FSDT). *Comparative Biochemistry and Physiology C-Toxicology & Pharmacology* **144**: 57-66.
7. Hutchinson,T.H. and D.B.Pickford, 2002. Ecological risk assessment and testing for endocrine disruption in the aquatic environment. *Toxicology* **181**: 383-387.
8. Jensen,K.M., E.A.Makynen, M.D.Kahl, and G.T.Ankley, 2006. Effects of the feedlot contaminant 17 alpha-trenholone on reproductive endocrinology of the fathead minnow. *Environmental Science & Technology* **40**: 3112-3117.
9. Kinnberg,K., H.Holbech, G.I.Petersen, and P.Bjerregaard, 2007. Effects of the fungicide prochloraz on the sexual development of zebrafish (*Danio rerio*). *Comparative Biochemistry and Physiology C-Toxicology & Pharmacology* **145**: 165-170.
10. Kunz,P.Y., T.Gries, and K.Fent, 2006. The ultraviolet filter 3-benzylidene camphor adversely affects reproduction in fathead minnow (*Pimephales promelas*). *Toxicological Sciences* **93**: 311-321.
11. Lange,R., T.H.Hutchinson, C.P.Croudace, and F.Siegmund, 2001. Effects of the synthetic estrogen 17 alpha-ethinylestradiol on the life-cycle of the fathead minnow (*Pimephales promelas*). *Environmental Toxicology and Chemistry* **20**: 1216-1227.

12. Milnes, M.R., D.S. Bermudez, T.A. Bryan, T.M. Edwards, M.P. Gunderson, I.L.V. Larkin, B.C. Moore, and L.J. Guillette, 2006. Contaminant-induced feminization and demasculinization of nonmammalian vertebrate males in aquatic environments. *Environmental Research* **100**: 3-17.
13. Nash, J.P., D.E. Kime, L.T.M. van der Ven, P.W. Wester, F. Brion, G. Maack, P. Stahlschmidt-Allner, and C.R. Tyler, 2004. Long-term exposure to environmental concentrations of the pharmaceutical ethynylestradiol causes reproductive failure in fish. *Environmental Health Perspectives* **112**: 1725-1733.
14. OECD. OECD, 2006. OECD Final Draft Report of Phase 1B of the Validation of the 21-day Fish Assay for Detection of Endocrine Active Substances. 131 pp. 2006.
Ref Type: Generic
15. Orlando, E.F., A.S. Kolok, G.A. Binzick, J.L. Gates, M.K. Horton, C.S. Lambright, L.E. Gray, A.M. Soto, and L.J. Guillette, 2004. Endocrine-disrupting effects of cattle feedlot effluent on an aquatic sentinel species, the fathead minnow. *Environmental Health Perspectives* **112**: 353-358.
16. Panter, G.H., T.H. Hutchinson, K.S. Hurd, J. Bamforth, R.D. Stanley, S. Duffell, A. Hargreaves, S. Gimeno, and C.R. Tyler, 2006. Development of chronic tests for endocrine active chemicals - Part 1. An extended fish early-life stage test for oestrogenic active chemicals in the fathead minnow (*Pimephales promelas*). *Aquatic Toxicology* **77**: 279-290.
17. Panter, G.H., T.H. Hutchinson, R. Lange, C.M. Lye, J.P. Sumpter, M. Zerulla, and C.R. Tyler, 2002. Utility of a juvenile fathead minnow screening assay for detecting (anti-)estrogenic substances. *Environmental Toxicology and Chemistry* **21**: 319-326.
18. Patyna, P.J., R.A. Davi, T.F. Parkerton, R.P. Brown, and K.R. Cooper, 1999. A proposed multigeneration protocol for Japanese medaka (*Oryzias latipes*) to evaluate effects of endocrine disrupters. *Science of the Total Environment* **233**: 211-220.
19. Vinggaard, A.M., C. Nellemann, M. Dalgaard, E.B. Jorgensen, and H.R. Andersen, 2002. Antiandrogenic effects in vitro and in vivo of the fungicide prochloraz. *Toxicological Sciences* **69**: 344-353.

APPENDIX

Appendix 1

Preparation of stock solutions of 17 β -estradiol, 4-tert-pentylphenol and prochloraz without the use of carrier solvent

17 β -Estradiol:

To prepare a 1 mg l⁻¹ stock solution:

1 mg of chemical is dissolved in one litre of deionised water, which has been autoclaved.

Magnetically stir for 5 mins.

Ultrasonicate for 5 mins.

The stock solution should then be clear and colourless.

This stock may be further diluted in deionised water, which has been autoclaved, to give a more preferable stock concentration.

The stock solution is prepared daily and at Brixham Environmental Laboratory (AstraZeneca) has been shown to be stable for 24 hours.

4-tert-Pentylphenol:

To prepare a 100 mg l⁻¹ stock solution of 4-tert-pentylphenol:

100 mg of chemical is dissolved in one litre of dechlorinated water.

Magnetically stir for 5 minutes.

Heat to approximately 50°C, using a water bath, provide overhead stirring and leave overnight.

Allow the solution to cool to test temperature before use.

The solution should be clear and colourless after cooling.

The stock solution is prepared every 1-2 days and at Brixham Environmental Laboratory has been shown to be stable over that time period.

Prochloraz:

(Michael Kahl and Gerald Ankley, US Environmental Protection Agency Duluth, MN, USA)

Preparation of chemical stocks for use in aquatic bioassays can be achieved using several techniques (Kahl *et al.* 1999; *Chemosphere* 39: 539). The method of choice for preparation of stock solution(s) must provide for an adequate volume of water to achieve desired test concentrations for a flow-through test. When possible, it is recommended that stock solutions be prepared in water without the use of solvents or

solubility enhancers. Different approaches have been used to successfully achieve appropriate test concentrations for chemicals of various physico-chemical characteristics in the absence of solvents.

Prochloraz requires the use of a solid-liquid slow-stir saturator for chemical delivery, if solvent carriers are not to be used. It has physical characteristics which make this method appropriate for chemical stock generation.

A solid-liquid saturator is designed to use the surface area of the inside of a glass container (e.g., carboy) as an area which, when coated with a test chemical and filled with water, generates a stock of solution at or near aqueous water solubility. Compounds which are solids and moderately water-soluble (e.g., $\log K_{ow} < 4.5$) are best suited for this method. Generally, these types of chemicals are not soluble enough to mix in stock vessels (e.g., as can be done for fadrozole) for immediate use, but not so insoluble that it is necessary to coat substrates with a large surface area, such as a glass-wool column saturation unit. Below is a protocol describing in detail the preparation of solid-liquid saturators appropriate for flutamide and prochloraz.

Solid-Liquid Saturator Protocol

Overview

The test chemical is dissolved in a solvent (usually acetone) before application on the inside of a glass vessel, with the amount of solvent varying with vessel size and amount of chemical used. Inlet and outlet ports are fitted with a neoprene stopper which, when secured, provides an air tight and constant working volume within the vessel (Fig. 1). The capacity of any saturator to maintain a constant concentration of chemical in solution will diminish with time of use. A general rule is to use at least twice, sometimes up to ten-fold, the amount of chemical the test may require to maintain a consistent concentration over the entire exposure period. The successful dosing of chemicals may require multiple saturators for each chemical to supply stock for use over the duration of the tests. The following detailed procedures are what is used at the Duluth EPA lab for stock solution generation via a solid-liquid saturator. Minor variations in materials and procedures should produce comparable results in terms of generating stocks for testing.

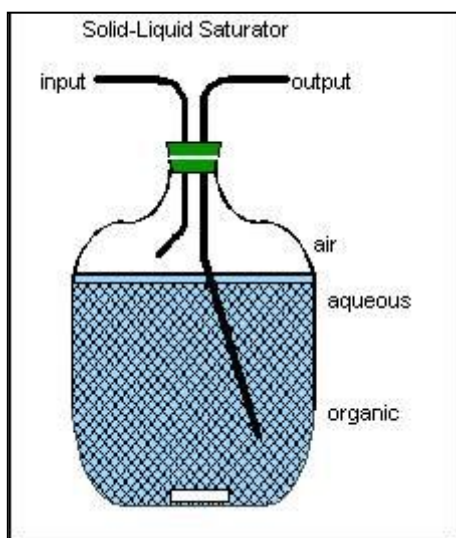


Figure 1. Schematic of glass carboy assembled as a solid-liquid saturator

Supplies:

19 L glass carboy - Pyrexplus (size can vary)

Neoprene stopper

Stainless steel tubing, 60 cm x 3.2 mm OD x 1.6 mm ID

TFE tubing, 60 cm x 3.2 mm OD x 1.6 mm ID

Magnetic stir plate

5.1 cm x 1 cm stir bar

Rolling mill - adaptable to carboy use

Delivery pump - low flow variable speed

Chemical Coating Procedure:

(using the carboy in a vented hood)

- Place 40 ml of acetone in 150 ml beaker
- Add preweighed chemical to acetone, dissolve
- Transfer dissolved chemical solution to carboy; rinse excess chemical from beaker to carboy with a small amount of acetone
- Place carboy on rolling mill; low speed

- Rinse chemical out of the neck area to the main body of the carboy with a small amount of acetone
- Meter in a small but constant flow of compressed air or nitrogen to facilitate evaporation
- When acetone has vaporized, discontinue rolling and slightly increase air/nitrogen flow for an additional 60 minutes

Preparation of Stock Solution:

Coating the glass with the chemical requires an even distribution, which is achieved by rolling the container, while supplying a low flow of gas to aid solvent evaporation. Evaporation which is too rapid will cause uneven coating and subsequently shorter life span of the saturator vessel. Upon total solvent evaporation, identified by a uniform crystalline coating or filming on the wall surface of the carboy, the vessel is filled to the desired level (just above the coated chemical; Figure 1) with test water, fitted with the stopper assembly and secured. It is important that the air space in the saturator be completely sealed from air exit or entry during use, and that it is possible to add new (clean) water at the same rate that the test solution is removed. Water input occurs just above the water level in the vessel, with the stainless steel output tube orifice located within 6 cm of the bottom. Prior to use in the test system, solid liquid saturators must be flushed to waste for a period of time. This procedure allows the saturator to come to a relative state of equilibrium, removes most of the chemical particulates not sufficiently coated to the vessel walls and allows for detection of leaks in the vessel.

Additional Operational Notes

The saturators are a flow-through system and require replacement of test water. Replacement water can be drawn from a reservoir which is filled daily or plumbed to a supply line which continually supplies make-up water. Input water, when added slightly above the water layer allows for monitoring of system function. Water in, must equal stock out.

Pumping rate of solution from the saturator (and introduction of clean water) can be increased gradually over time to supply greater quantities of toxicant should analytical determinations indicate that efficiency of the saturator is diminishing.

Stir rate should be adequate to continually move solution in the carboy over the chemical interface. It is important to avoid too rapid a stir rate which could damage the chemical layer causing it to peel from the vessel sides. The presence of a slight vortex at the air/liquid interface is generally sufficient for adequate stock mixing.

Securing the stopper/tubing assembly can be accomplished by the use of a metal strap type hose clamp (20-30 cm) and flexible, but rigid wire.

New replacement carboys for a test should be made and flushed in advance of use with pre-sampling and analytical verification of stock solutions advised.

Stock Solutions Used in Duluth EPA Lab Exposures

Stock Preparations: Prochloraz

The number of coated carboys used will vary dependent on the target test concentrations and demands of the test system (flow rates, replicate tanks, and tank volume additions per day).

Prochloraz

2.5 g chemical is dissolved in 40 ml of acetone

19 L carboy coated and flushed with water at 3.5 ml/min for 1 day

Magnetic stir plate with 5.1 x 1 cm stir bar at moderate stir rate

1.5 to 2.75 ml/min flow rates will generated a 22 to 40 mg/L stock for 11 days

(Pump rate may need to be adjusted up slightly every 2 days - this allows supply volumes of stock to meet the requirements of the test system)

Two coated vessels should supply a usable stock solution for 21 day exposure

Appendix 2

Method for quantification of Prochloraz by LC-MS

(Propiconazole as internal standard)

Provided by Henrik Holbech, SDU, Odense July 2006

Solid Phase Extraction:

Column: Strata-X 100 mg/6 ml Polymeric RP sorbent (8b-S100-ECH, Supware)

conditioning: 3 ml MeOH

Equilibration: 3 ml H₂O

Wash: 3 ml 15% MeOH (The column is emptied and dried for 1 min)

Elution: 3 ml MeOH (The column is closed when 2 ml has been applied and then soaked for 1 min)

Drying: 10 min. in a TurboVap

Re-solution: 300 µl MeOH

LCMS:

Internal standard: Propiconazole (Tilt TM; PS-1075 Supelco. Supelco park, bellafonte USA). Add 50 µl of a 13 µg/ml solution

Column: C₁₈ 250x4.6mm (e.g. 5µm. Prodigy from Phenomenex)

Column temperature: 25 °C

Eluents: A: 85% MeOH

Flow: 0.6 ml/min

Stoptime: 15 min

Gradient: Isocratic

Injektion: 50 µl

Mode: APCI, positive

Ions: Prochloraz 376; Propiconazole 342

Fragmentor: 50 V

Drying gas flow: 2.0 l/min

Nebulizer pressure: 40 psig

Drying gas temperature: 350 °C

Vaporizer Temperature: 400°C

Capillary Voltage: 5000 V

Corona Current: 4 µA

Standards: Prochloraz 6.25; 12.5; 25; 50; 62.5; 100; 500,00 ng

Control: 200 ng on LCMS

Appendix 3

Method for quantification of 4-tert-pentylphenol by LC-MS (4-tert butylphenol as internal standard)

Provided by Henrik Holbech, SDU, Odense July 2006

Solid Phase Extraction:

column: Sep Pak Vac 6cc C18.1 g (Waters wat 036905)

Conditioning: 5 ml MeOH

Equilibration: 5 ml H₂O

Wash: 5 ml H₂O (The column is emptied and dried for 1 min)

Elution: 2 ml MeOH (The column is closed when 2 ml has been applied and then soaked for 1 min)

Drying: none

Re-solution: none

LCMS:

Internal standard: 4-tert butylphenol (B 9.990-1; Aldrich). Add 50 µl of a 6.6 µg/ml solution

Column: C₁₈ 150x2.1 mm (e.g. 5µm Luna from Phenomenex)

Column temperature: 25 °C

Eluents: A: 20 % MeOH; B: 100 % MeOH

Flow: 0.4 ml/min

Stoptime: 25 min

Gradient:

Time (min)	% B
5	25
7	60
20	100
25	100
25.1	25

Injection: 20 µl

Mode: APCI, negative

Ions: Pentylphenol: 163; Butylphenol 149

Fragmentor: 90 V

Drying gas flow: 2.8 l/min

Nebulizer pressure: 20 psig

Drying gas temperature: 300 °C

Vaporizer Temperature: 325°C

Capillary Voltage: 2500 V

Corona Current: 20 µA

Standards: 4-tert-pentylphenol (15.384-2, Aldrich)

5.0; 25.0; 125.0; 625.0; 2500.0 ng

Control: 400 ng/ml, 100 ng on LCMS; 2000 ng/l 500 ng on LCMS

Appendix 4

ANALYTICAL METHOD FOR DETERMINATION OF PROCHLORAZ CONCENTRATIONS *(provided by GT Ankley, US-EPA Duluth, 2004)*

Routine quality assurance analyses (i.e., procedural blanks, spiked matrix and duplicate samples) were conducted with each sample set and constitute approximately 10% of all samples. Prochloraz standards were prepared in 10% methanol/water and spiked matrix samples were prepared by diluting an aliquot of a prochloraz solution (prepared in methanol) with Lake Superior water (resulting concentration of about 300 ppb). During the course of the study all exposure tanks were sampled in duplicate at least once. Prochloraz was obtained from Sigma, St Louis, MO, and all solvents were Chromatography and HPLC grade or better.

Water samples were collected from the control and treatment tanks two to three times a week during the course of the study. These samples were placed into crimp top amber vials and immediately analyzed

for prochloraz by reversed phase HPLC. The stock/saturator solution was diluted 20 fold (to get within standard range) with DI water before analysis. The Agilent (Wilmington, DE, USA) model 1100 HPLC consisted of a capillary pump, chilled auto sampler (4°C), heated column compartment (25°C), and a diode-array detector. An aliquot of sample (20 µl) was injected onto a Zorbax (Agilent, Wilmington, DE) SB-C18 column (2.1 x 75 mm) and eluted isocratically with 75% methanol/water at a flow rate of 0.4 ml/min. Prochloraz concentrations were determined using the response at wavelength 220 nm and an external standard method of quantitation. No prochloraz was detected in the control tanks (n = 24) or procedural blanks (n = 7). The mean (\pm SD) recovery of prochloraz in the spiked matrix samples was 97 ± 2 % (n = 6) and the mean (\pm SD) percentage agreement among duplicate samples was 98 ± 1 % (n = 11). The analytical quantification limit was 15 ng/ml. Concentrations of prochloraz in the water samples were recovery corrected based on an average of all the spiked matrix data in the study.

Appendix 5

4-Tert pentylphenol (4-tPP) and estradiol (E2):

Analytical methods provided by Wildlife International, LTD.

The analytical method for 4-tert-pentylphenol consisted of diluting the samples in freshwater, as necessary, and analyzing by high performance liquid chromatography (HPLC) with UV detection at 220 nm. The analytical method for Estradiol consisted of extracting the samples with dichloromethane, rotary evaporation and reconstitution in 50% methanol : 50% water, followed by analysis by high performance liquid chromatography with mass spectrometric detection (LC/MS/MS).

Concentrations of 4-tert-pentylphenol in the samples were determined by high performance liquid chromatography using an Agilent Series 1100 High Performance Liquid Chromatograph (HPLC) equipped with an Agilent Series 1100 Variable Wavelength Detector. Chromatographic separations were achieved using an Agilent SB-phenyl column (250 mm x 4.6 mm, 5 µm particle size). A method flowchart is provided in Appendix 3.1, and typical instrumental parameters are summarized in Appendix 3.3.

Five calibration standards of 4-tert-pentylphenol, ranging in concentration from 50.0 to 500 µg/L, were prepared using a stock solution of 4-tert-pentylphenol test substance in methanol (Appendix 3.5). The calibration standards were analyzed with each sample set. Linear regression equations were generated using the peak area responses versus the respective concentrations of the calibration standards. The concentration of 4-tert-pentylphenol in the samples was determined by substituting the peak area responses of the samples into the applicable linear regression equation. An example of the calculations for a representative sample is included in Appendix 3.6.

The method limit of quantitation (LOQ) for 4-tert-pentylphenol was defined as 50.0 µg/L, calculated as the product of the concentration of the lowest calibration standard (50.0 µg/L) and the dilution factor of the matrix blank samples (1.00). Four matrix blank samples were analyzed to determine possible interferences. No interferences were observed at or above the LOQ during the sample analyses (Appendix 3.7).

For 4-tert-pentylphenol, samples of freshwater were fortified at 100 and 1000 µg/L using a stock solution of 4-tert-pentylphenol test substance in methanol (Appendix 3.5), and were analyzed concurrently with the test samples. The measured concentrations for the matrix fortification samples ranged from 94.9 to 101% of fortified concentrations (Appendix 3.7).

Concentrations of Estradiol in the samples were determined using a PE SCIEX API 3000 LC/MS/MS coupled with an Agilent Series 1100 HPLC system. The Mass Spectrometer was equipped with a PE SCIEX Heated Nebulizer ion source. Chromatographic separations were achieved using a Keystone Betasil C-18 (50 x 2 mm, 5 µm particle size) analytical column fitted with a Keystone Javelin

C-18 guard column. A method flowchart is provided in Appendix 3.2, and instrumental parameters are summarized in Appendix 3.4.

Five calibration standards of Estradiol, ranging in concentration from 0.500 to 5.00 µg/L, were prepared using a stock solution of Estradiol reference substance in methanol (Appendix 3.5). The calibration standards were analyzed with each sample set. Linear regression equations were generated using the peak area responses versus the respective concentrations of the calibration standards. The concentration of Estradiol in the samples was determined by substituting the peak area responses of the samples into the applicable linear regression equation.

The method limit of quantitation (LOQ) for Estradiol was defined as 0.020 µg/L, calculated as the product of the concentration of the lowest calibration standard (0.500 µg/L) and the dilution factor of the matrix blank samples (0.04). Four matrix blank samples were analyzed to determine possible interferences. No interferences were observed at or above the LOQ during the sample analyses (Appendix 3.8).

For Estradiol, freshwater samples were fortified at 0.100 µg/L using a stock solution of Estradiol reference substance in methanol (Appendix 3.5), and were analyzed concurrently with the test samples. The measured concentrations for the matrix fortification samples ranged from 95.4 to 101% of fortified concentrations (Appendix 3.8).

A representative calibration curve for 4-tert-pentylphenol is presented in Appendix 3.9. Representative chromatograms of low and high-level calibration standards for 4-tert-pentylphenol are presented in Appendices 3.10 and 3.11, respectively. Representative chromatograms of a freshwater matrix blank sample and a matrix fortification sample for 4-tert-pentylphenol are presented in Appendices 3.12 and 3.13, respectively. A representative chromatogram of a test sample for 4-tert-pentylphenol is presented in Appendix 3.14.

A representative calibration curve for Estradiol is presented in Appendix 3.15. Representative chromatograms of low and high-level calibration standards for Estradiol are presented in Appendices 3.16 and 3.17, respectively. Representative chromatograms of a freshwater matrix blank sample and a matrix fortification sample for Estradiol are presented in Appendices 3.18 and 3.19, respectively. A representative chromatogram of a test sample for Estradiol is presented in Appendix 3.20.

Appendix 6

SOP: Homogenisation of head & tail from juvenile zebrafish, fathead minnow, three spined stickleback and Japanese medaka

- The fish are anaesthetised and euthanised in accordance with the test description.
- The head and tail are cut off the fish in accordance with the test description. **Important:** *All dissection instruments, and the cutting board must be rinsed and cleaned properly (e.g. with 96% ethanol) between handling of each single fish to prevent “vitellogenin pollution” from females or induced males to un-induced males.*
- The weight of the pooled head and tail from each fish is measured to the nearest mg.
- After being weighed, the parts are placed in appropriate tubes (e.g. 1.5 ml eppendorf) and frozen at -80°C until homogenisation or directly homogenised on ice with two plastic pestles. (Other methods can be used if they are performed on ice and the result is a homogenous mass). **Important:** *The tubes must be numbered properly so that the head and tail from the fish can be related to their respective body-section used for gonad histology.*
- When a homogenous mass is achieved, 10 x the tissue weight of ice-cold **homogenisation buffer*** is added. Keep working with the pestles until the mixture is homogeneous. **Important note:** *New pestles are used for each fish.*
- The samples are placed on ice until centrifugation at 4°C at 50000 x g for 30 min.
- Use a pipette to dispense portions of 50 μl supernatant into **at least two** tubes by dipping the tip of the pipette below the fat layer on the surface and carefully sucking up the supernatant without fat- or pellet fractions.
- The tubes are stored at -80°C until use.
- ***Homogenisation buffer:**
- (50 mM Tris-HCl pH 7,4; 1% Protease inhibitor cocktail (Sigma)): 12 ml Tris-HCl pH 7,4 + 120 μl Protease inhibitor cocktail.

TRIS: TRIS-ULTRA PURE (ICN) e.g. from Bie & Berntsen, Denmark.

Protease inhibitor cocktail: From Sigma (for mammalian tissue) Product number **P 8340**.

NOTE: The homogenisation buffer has to be used the same day as manufactured. Place on ice during use.

Appendix 7

SOP: Quantification of head & tail homogenate vitellogenin in zebrafish (*Danio rerio*) (modified from Holbech et al., 2001)

- Microtiterplates (certified Maxisorp F96, Nunc, Roskilde Denmark) previously coated with 5 µg/ml anti zebrafish lipovitellin-IgG are thawed and washed 3 times with washing buffer*.
- Battelle zebrafish Standard AP4.6.04 (1.18 mg/ml (AAA)) is serially diluted to 0.2, 0.5, 1, 2, 5, 10 and 20 ng/ml in dilution buffer** and samples are diluted at least 200 times (to prevent matrix effects) in dilution buffer and applied to the plates. An assay control is applied in duplicate. 150 µl are applied to each well. Standards are applied in duplicate and samples in triplicate. Incubate over night at 4°C on a shaker.
- The plates are washed 5 times with washing buffer*
- AMDEX (HRP coupled to a dextran chain (AMDEX A/S, Denmark)) conjugated antibodies are diluted in washing buffer; actual dilution differs by batch and age. 150 µl are applied to each well and the plates are incubated for 1 hour at room temperature on a shaker.
- The plates are washed 5 times with washing buffer* and the bottom of the plates is carefully cleaned with ethanol.
- 150 µl TMB plus*** are applied to each well. Protect the plate against light with tinfoil, and watch the colour development on a shaker.
- When the standard curve is fully developed the enzyme activity is stopped by adding 150 µl 0.2 M H₂SO₄ to each well.
- The absorbance is measured at 450 nm (e.g. on a Molecular Devices Thermomax plate reader). Data are analysed on the associated software (e.g. Softmax).

*** Washing buffer:**

PBS-stock****	500.0 ml
BSA	5.0 g
Tween 20	5.0 ml

Adjust pH to 7.3 and fill to 5 l with millipore H₂O. Store at 4° C.

****Dilution buffer**

PBS-Stock***	100.0 ml
BSA	3.0 g
Tween 20	1.0 ml

Adjust pH to 7.3 and fill to 1 l with millipore H₂O. Store at 4° C.

*** TMB plus is a "ready-to-use" substrate produced by KemEnTec (Denmark). It is sensitive to light. Store at 4° C.

****** PBS stock**

NaCl	160.0 g
KH ₂ PO ₄	4.0 g
Na ₂ HPO ₄ ·2H ₂ O	26.6 g
KCl	4.0 g

Adjust pH to 6.8 and fill with millipore H₂O to 2 l. Store at room temperature.

Appendix 8

SOP: Preparation of tissue sections for sex determination and staging of gonads

The purpose of this section is to describe the procedures that occur prior to the evaluation of histological sections.

With a few exceptions these procedures are similar for fathead minnow (FHM), Japanese medaka (JMD), zebrafish (ZBF), and three spined stickleback (TSS).

Euthanasia, Necropsy, and Tissue Fixation

Objectives:

- Provide for the humane sacrifice of fish.
- Obtain necessary body weights and measurements.
- Evaluate secondary sex characteristics.
- Dissect tissues for vitellogenin analysis.
- Fixation of the gonads.

Procedures:

- Fish should be sacrificed immediately prior to necropsy. Therefore, unless multiple processors are available, multiple fish should not be sacrificed simultaneously.
- Using a small dip net, a fish is removed from the experimental chamber and transported to the necropsy area in the transport container.
- The fish is placed in the euthanasia solution. The fish is removed from the solution when there is cessation of respiration and the fish is unresponsive to external stimuli.
- The fish is wet weighed.
- For preparation of tissues for vitellogenin analysis the fish can be placed on a corkboard on the stage of a dissecting microscope.
- For FHM and ZBF the head is cut right behind the pectoral fin and tail is cut right behind the dorsal fin.
- For JMD the abdomen is opened via a carefully made incision that extends along the ventral midline from the pectoral girdle to a point just cranial to the anus. Using the small forceps and small scissors, the liver is carefully removed.
- Specimen for vitellogenin analysis are placed in eppendorf tubes and immediately frozen in liquid nitrogen.
- The carcass including the gonads is placed into a pre-labelled plastic tissue cassette, which is transferred into Davidson's or Bouin's fixative. The volume of fixative should be at least 10 times the approximate volume of the tissues. The fixative container is gently agitated for five seconds to dislodge air bubbles from the cassette.
- All tissues remain in Davidson's fixative overnight, followed by transfer to individual containers of 10 % neutral buffered formalin the next day. Containers with cassettes are gently agitated for 5 seconds to ensure adequate penetration of formalin into cassettes.
- Tissues remain in Bouin's fixative for 24 h, followed by transfer to 70 % ethanol.

Tissue Processing

Objectives:

- Dehydrate tissue for adequate penetration of paraffin.

- Impregnate the tissue with paraffin to maintain tissue integrity and create a firm surface for microtomy.

Procedures:

- Labelled tissue cassettes are removed from formalin/ethanol storage and the cassettes are placed in the processing basket(s). The processing basket is loaded in the tissue processor.
- The processing schedule is selected.
- After the tissue processor has completed the processing cycle, the basket(s) may be transferred to the embedded station.

Embedding

Objective:

Properly orient the specimen in solidified paraffin for microtomy.

Procedures:

- The basket(s) of cassettes is/are removed from the processor and immersed in the paraffin-filled front chamber of the embedding station thermal console or the cassettes are moved to a separate paraffin heater.
- The first cassette to be embedded is removed from the front chamber of the thermal console or the paraffin heater. The cassette lid is removed and discarded, and the cassette label is checked against the animal records to resolve potential discrepancies prior to embedding.
- An appropriately sized embedding mould is selected.
- The mould is held under the spout of the dispensing console and filled with molten paraffin.
- The specimen is removed from the cassette and placed in the molten paraffin in the mould. This is repeated with 6-8 specimens for each paraffin mould. The position of individual fish is marked by putting fish no 1 in 180 degrees to fish 2-6/8.
- Additional paraffin is added to cover the specimen.
- The mould with the cassette base is placed on the cooling plate of the cryo console.
- After the paraffin has solidified, the block (i.e., the hardened paraffin containing the tissues and the cassette base) is removed from the mould.

Microtomy

Objective:

Cut and mount histological sections for staining.

Procedures:

- The initial phase of microtomy termed “facing” is conducted as follows:
- The paraffin block is placed in the chuck of the microtome.
- The chuck is advanced by rotating the microtome wheel and thick sections are cut from the paraffin surface of the block until the knife reaches the embedded tissues.
- The section thickness on the microtome is set between 4 – 10 microns. The chuck is advanced and multiple sections are cut from the block to remove any artefacts created on the cut surface of the tissue during rough trimming.
- The block can be removed from the chuck and placed facedown on ice to soak the tissue.

- The next phase of microtomy is final sectioning and mounting of tissue sections on slides. These procedures are conducted as follows:
- If the block has been placed on ice, the block is removed from the ice and replaced in the chuck of the microtome.
- With the section thickness on the microtome set to 4 – 5 microns, the chuck is advanced by rotating the microtome wheel. Sections are cut from the block until a “ribbon” containing at least one acceptable section including the gonads has been produced. (As necessary during sectioning, the block may be removed from the chuck, placed on ice to soak the tissue, and replaced in the chuck.)
- The sections are floated flat on the surface of the water in the water bath. An attempt is made to obtain at least one section that contains no wrinkles and has no air bubbles trapped beneath it.
- A microscope slide is immersed beneath the best section, which is lifted out of the water using the slide. This process is referred to as “mounting” the section on the slide.
- Three sections are prepared for a set of fish. The second and third sections are taken at 50 micron intervals following the first section. If the fish are not embedded with their gonads in the same sectioning level, more sections are to be made to ensure that at least three sections including the gonads are obtained from each fish.
- With a slide-marking pen, the block number from which the slide was produced is recorded on the slide.
- The slide is placed in a staining rack.
- The block is removed from the chuck and placed facedown for storage.

Staining, Cover slipping, and Slide Labelling

Objectives:

- Stain the sections for histopathological examination
- Permanently seal mounted and stained tissues.
- Permanently identify stained sections in a manner that allows complete traceability.

Procedures:

- Staining
- Slides are air-dried overnight before staining.
- The sections are stained by Hematoxylin-Eosin.
- Cover slipping
- Cover slips can be applied manually or automatically.
- A slide is dipped in xylene or TissueClear, and the excess xylene/TissueClear is gently knocked off the slide.
- Approximately 0.1 ml of mounting medium is applied near the end of the slide opposite to the frosted end or on the cover slip.
- The cover slip is tilted at a shallow angle as it is applied to the slide.
- Labelling
- Each slide label should contain the following information.
- Laboratory name
- Species
- Specimen No. / Slide No.
- Chemical / Treatment group
- Date

Appendix 9

Statistical Analysis of Fish Sexual Development Sex Ratio Data John W. Green, Ph.D., Ph.D. Senior Consultant: Biostatistics, DuPont Applied Statistics Group

Description of Experimental Data

Four laboratories participated in experiments with fish exposed to various concentrations of two chemicals, prochloraz and 4tPP. Three of the Labs used zebrafish and one lab, Lab 5, used fathead minnows. For prochloraz, labs 2, 3 and 4 used nominal test concentrations of 38, 75, 150, 300, and 600 µg/L, whereas Lab 5 used nominal test concentrations of 32, 100, and 320 µg/L. For 4tPP, Labs 2 and 3 used test concentrations of 38, 75, 150, 300, and 600 µg/L, while Labs 4 and 5 used test concentrations of 32, 75, 100, and 320 µg/L. Each test concentration in Labs 2-4 was expressed in two replicate tanks, while Lab 5 had 4 reps of each test concentration.

For both chemicals and all labs, there was also a zero nominal concentration control of 4 or 8 reps. Lab 3 had two reps of an acetone control and 2 reps of water control. The other labs did not report in an unambiguous fashion the use of a solvent control, though Lab 5 reported 4 reps of control 1 and 4 reps of control 2 which may be an indication of use of a solvent.

In addition to the nominal concentration, each lab also reported the mean measured concentration on each replicate. Within each lab, the number of fish in each replicate was roughly uniform, but different labs used different numbers of fish, ranging from about 20 to about 45.

In each replicate vessel, the counts of males, females, intersex, and undifferentiated fish at study end were reported. These counts or the corresponding proportions of the total number of fish per replicate, were the responses analyzed. Since it was not clear whether to expect a chemical effect to manifest itself as an increase or a decrease in males or females, 2-sided hypothesis tests were used to determine the NOEC. For regression model fitting to estimate an EC_x value, the data dictated an increasing or decreasing dose-response.

Utility of Regression Modeling for Fish Sex Ratio Data

In previous data on sex ratio where only nominal test concentrations were available, it was demonstrated that the regression approach provided little useful information because of the great variability associated with a background (i.e., control) incidence of approximately 50% for males and females. Further discussion of this is given in Annex 1. Since the present dataset contains the mean measured test concentrations on a per-rep basis, the question of regression versus hypothesis testing was revisited for the prochloraz experiments. As is demonstrated in Annex 2 and summarized in Table 2 below, regression on replicate mean measured test concentrations did not alter the earlier conclusion on the value, or lack thereof, of that approach for fish male and female proportions. A statistical protocol for regression is given in Annex 1.

Statistical Protocol - NOEC

All of the tests described in this section, with the exception of the Rao-Scott test, are discussed in detail in OECD (2006) and, as in that reference, all tests were done at the overall 0.05 significance level. All analyses were done on replicate mean proportions or counts. Where a

solvent control was used, it was compared to the water control using the Mann-Whitney test. If no significant difference between the water and solvent controls was found, then the two controls were combined, retaining separate rep identification, for further analysis. If a statistically significant difference between the water and solvent controls was found, then the water control was discarded and all further analysis used only the solvent control. Some discussion of this procedure is given in Annex 1. In the event, no difference was found between the two controls.

Since it is well known that proportion data tend to have heterogeneous variances, the proportions were transformed using the arcsin square-root transform. This is a standard transformation to use for proportion data and usually stabilizes the variance while maintaining a normal distribution. Nevertheless, formal tests for normality (Shapiro-Wilk) and variance homogeneity (Levene) were conducted for each response.

A parametric analysis (Williams or Dunnett) of male and female proportions is generally valid, since the number of fish and number of males or females per replicate vessel vary and are mostly different from zero. These data accordingly do not have large numbers of zeros or tied values, making normality feasible. Intersex and undifferentiated present a different problem, as in some experiments there are few fish or none of either type. For these endpoints, the normality requirement for a standard parametric analysis such as the Dunnett or Williams test is highly suspect. Non-parametric methods (Dunn, Mann-Whitney, Jonckheere-Terpstra) are still valid. As demonstrated elsewhere, the Williams and Jonckheere-Terpstra step-down trend tests have superior power properties to pairwise tests (Dunnett, Dunn, Mann-Whitney) provided the data follow a monotone dose-response. Furthermore, the power properties of the Williams and Jonckheere-Terpstra tests are very similar under widely varying conditions, so that restriction to the latter results in no loss of power to detect effects. Furthermore, the Jonckheere-Terpstra test is readily available in commercial software. Therefore, to keep the statistical protocol simple and avoid undue programming difficulty, the only trend test used on all four proportions (male, female, intersex, undifferentiated) was the Jonckheere-Terpstra test.

Since trend tests assume dose-response monotonicity, a test for that was done for each response from each experiment. If a significant departure from dose-response monotonicity was found, then a pairwise test was done. The pairwise test done was Dunnett's test if the data satisfied the requirements for that test. If the data were normally distributed but had heterogeneous variances, then a robust form of the Dunnett test called the Tamhane-Dunnett test was used. Otherwise, an exact version of the Mann-Whitney test with Bonferroni-Holm adjusted p-values was done.

There is an alternative test that could be applied to these responses, namely the Cochran-Armitage test with the Rao-Scott adjustments for extra-binomial variability (Rao and Scott (1992)). This test has been demonstrated elsewhere to have very desirable properties. However, no commercial software is known that implements this test. For this very practical reason and the fact that the Jonckheere test also has attractive properties, this test was not used on the data for this report. Finally, for comparison, Dunn's test was reported. However, when a pairwise test is needed, an exact version of the Mann-Whitney test is recommended. This test is available in the commercially available packages StatXact (any version) and SAS (version 8 or later).

Summary of Results for Prochloraz

Table 1a: Prochloraz NOEC Results Using Trend Test

Chemical	Lab	Response	NOEC	Test	Direction	Comment
Prochloraz	2	Females	48.1	JT	2-sided	
Prochloraz		Intersex	>319.6	JT	2-sided	
Prochloraz		Male	>319.6	JT	2-sided	
Prochloraz		Undif	>319.6	JT	2-sided	
Prochloraz	3	Females	>197.4	JT	2-sided	
Prochloraz		Intersex	>197.4	JT	2-sided	
Prochloraz		Male	>197.4	JT	2-sided	
Prochloraz		Undif	>197.4	JT	2-sided	
Prochloraz	4	Females	59.7	JT	2-sided	
Prochloraz		Intersex	>434.1	Dunnett	2-sided	significant non-monotonicity
Prochloraz		Male	135.0	JT	2-sided	
Prochloraz		Undif	>434.1	JT	2-sided	
Prochloraz	5	Females	96.4	JT	2-sided	
Prochloraz		Intersex	>286.4	JT	2-sided	
Prochloraz		Male	96.4	JT	2-sided	
Prochloraz		Undif	>286.4	JT	2-sided	

Table 1b: Prochloraz NOEC Results Using Dunnett or Dunn Pairwise Test

Chemical	Lab	Response	NOEC	Test	Direction	Comment
Prochloraz	2	Females	48.1	Dunnett	2-sided	
Prochloraz		Intersex	14.5	Dunnett	2-sided	Higher conc not significant
Prochloraz		Male	>319.6	Dunnett	2-sided	
Prochloraz		Undif	48.1	Dunnett	2-sided	
Prochloraz	3	Females	98.6	Dunnett	2-sided	
Prochloraz		Intersex	>197.4	Dunn	2-sided	
Prochloraz		Male	98.6	Dunnett	2-sided	
Prochloraz		Undif	>197.4	Dunn	2-sided	
Prochloraz	4	Females	0.0	Dunnett	2-sided	
Prochloraz		Intersex	>434.1	Dunn	2-sided	
Prochloraz		Male	135.0	Dunnett	2-sided	
Prochloraz		Undif	>434.1	Dunn	2-sided	
Prochloraz	5	Females	96.4	Dunnett	2-sided	
Prochloraz		Intersex	96.4	Dunn	2-sided	
Prochloraz		Male	96.4	Dunnett	2-sided	
Prochloraz		Undif	>286.4	Dunn	2-sided	

The step-down Jonckheere-Terpstra and Dunnett tests usually agree for these experiments. Lab 3 results are exceptions, where Dunnett's test finds the effect in the highest test concentration significant for males and females but the Jonckheere-Terpstra test does not. There is definite non-monotonicity in the dose responses for females, but the formal test for non-monotonicity is significant only at the 0.06 level. There is also evidence of non-monotonicity in the responses for males, but the formal test is not significant.

Lab 2 also exhibits some differences between the two types of test. Intersex is significant by Dunnett's test at the second highest test concentration but not at the highest test concentration, making the one significant finding suspect. Undifferentiated is significant at the highest test concentration by a 2-sided Dunnett's test and by a 1-sided Jonckheere test, but not by a 2-sided Jonckheere test. The effect appears to be real.

In Lab 4, Dunnett's test finds a significant effect at every test concentration for females, whereas the Jonckheere does not find a significant effect at the lowest tested concentration. The latter test is significant at the 0.07 level. There is a 16 percentage point difference in the percent of females in the lowest test concentration compared to the control.

In Lab 5, the data for intersex failed the normality requirement, so pairwise methods were used. The above table reports the result of Dunn's test. The Mann-Whitney test, reported in Annex 3, gave the same result. However, the Jonckheere-Terpstra test was applicable and should be the result reported.

Differences in results from different tests are expected and the differences observed are not cause for concern. In general, a definite statistical protocol should be decided in advance of the experiments and then should be followed. The two sets of results are included here only for informational purposes. The recommended protocol is to use the Jonckheere test unless there is evidence of a non-monotone dose-response. Such evidence exists in the data from Lab 2.

The arcsine square-root transform is irrelevant when non-parametric tests are done, such as the Jonckheere-Terpstar, Dunn, and Mann-Whitney. For the tests reported, it is only important for Dunnett's test.

Details of the NOEC tests are given in Annexes 3 and 4.

Table 2: Summary of Prochloraz Regression Results

Lab	Response	AEC10	LCB	UCB	AEC25	LCB	UCB	Model
2	Male	654.6	0.0	3147.3	2184.9	0.0	10658.0	OECD2
		672.0	~	~	5932055.0	~	~	Probit
	Females	109.9	1.0	218.8	369.7	0.0	753.9	OECD2
		109.3	~	~	265.5	~	~	Probit
	Intersex	~	~	~	~	~	~	OECD2
		~	~	~	~	~	~	Probit
	Undiff	91.8	57.2	126.4	240.4	149.0	331.8	OECD2
		136.5	~	~	105.7	~	~	Probit
3	Male	~	~	~	~	~	~	OECD2
		~	~	~	~	~	~	Probit
	Females	943.6	0.0	4924.1	2832.1	0.0	14844.1	OECD2
		~	~	~	~	~	~	Probit
	Intersex	~	~	~	~	~	~	OECD2
		~	~	~	~	~	~	Probit
	Undiff	~	~	~	~	~	~	OECD2
		~	~	~	~	~	~	Probit
4	Male	61.1	31.7	90.5	195.1	95.7	294.5	OECD2
		55.9	16.9	98.0	166.8	94.0	247.4	Probit
	Females	39.3	20.4	58.2	126.5	62.2	190.8	OECD2
		44.5	14.7	71.7	100.9	58.1	136.2	Probit
	Intersex	~	~	~	~	~	~	OECD2
		~	~	~	~	~	~	Probit
	Undiff	~	~	~	~	~	~	OECD2
		~	~	~	~	~	~	Probit
5	Male	81.5	0.0	165.8	253.4	0.0	523.3	OECD2
		40.9	~	~	288.0	~	~	Probit
	Females	67.6	3.4	131.8	210.6	4.2	417.0	OECD2
		63.7	~	~	193.8	~	~	Probit
	Intersex	~	~	~	~	~	~	OECD2
		~	~	~	~	~	~	Probit
	Undiff	~	~	~	~	~	~	OECD2
		~	~	~	~	~	~	Probit

With only a few exceptions, these regression estimates either could not be obtained or had confidence intervals so wide as to be useless. Where the lower confidence bound is zero, the AECx estimate is not significantly different from zero, underscoring the lack of information provided by these estimates. Estimates from the two models (probit and OECD 2) are reasonably consistent.

All labs except Lab 5 used zebrafish, whereas Lab 5 used fathead minnows. Labs 3 and 4 had nominal test concentrations of 38, 75, 150, 300, and 600 µg/L Prochloraz, Lab 2 had all of these except the high concentration, while Lab 5 had nominal test concentrations of 32, 100, and 320 µg/L Prochloraz. There is an order of magnitude range among AEC10 values for males in Labs 2-4 and a 20-fold spread among AEC10 values for females. Only Lab 2 exhibited sufficiently many intersex or undifferentiated fish to fit a model.

Table 3a: 4tPP NOEC Results Using Trend Test

Chemical	Lab	Response	NOEC	Test	Direction	Comment
4TPP	2	Female	34	JT	2-Sided	
4TPP		Intersex	>96.8			no obs
4TPP		Male	34	JT	2-Sided	
4TPP		Undif	>96.8	JT	2-Sided	
4TPP	3	Female	>306.5	JT	2-Sided	
4TPP		Intersex	>306.5			no obs
4TPP		Male	>306.5	JT	2-Sided	
4TPP		Undif	>306.5			no obs
4TPP	4	Female	>227.2	JT	2-Sided	
4TPP		Intersex	>227.3	JT	2-Sided	
4TPP		Male	71	JT	2-Sided	
4TPP		Undif	71	JT	2-Sided	
4TPP	5	Female	36.1	JT	2-Sided	
4TPP		Intersex	>296.2			no data
4TPP		Male	36.1	JT	2-Sided	
4TPP		Undif	94	JT	2-Sided	

Table 3b: 4tPP NOEC Results Using Pairwise Test

Chemical	Lab	Response	NOEC	Test	Direction	Comment
4TPP	2	Female	62.2	Dunnett	2-sided	Heterogeneous variances
4TPP		Intersex	>96.8			no observations
4TPP		Male	>96.8	Dunnett	2-sided	Heterogeneous variances
4TPP		Undif	>96.8	Dunnett	2-sided	
4TPP	3	Female	>306.48	Dunnett	2-sided	Heterogeneous variances
4TPP		Intersex	>306.48			no observations
4TPP		Male	>306.48	Dunnett	2-sided	Heterogeneous variances
4TPP		Undif	>306.48			no observations
4TPP	4	Female	>227.2	Dunnett	2-sided	Heterogeneous variances
4TPP		Intersex	>227.2	Dunnett	2-sided	Het variances, non-monotone
4TPP		Male	71	Dunnett	2-sided	
4TPP		Undif	71	Dunn	2-sided	
4TPP	5	Female	36.1	Dunnett	2-sided	
4TPP		Intersex	>296.2			no observations
4TPP		Male	36.1	Dunnett	2-sided	
4TPP		Undif	94	Dunn	2-sided	

There were only two differences in results when pairwise tests are used. In Lab 2, the Jonckheere-Terpstra test found significantly elevated females and significantly reduced males in the two highest test concentrations but Dunn's test found no significant effect in males and found only the highest test concentration significant in females, while the Mann-Whitney test found no significant effect in either response at any test concentration. In Lab 5, the Jonckheere-Terpstra and Dunn tests found significantly elevated undifferentiated fish only in the highest test concentration, while the Mann-Whitney test found significantly elevated undifferentiated fish in the highest two test concentrations.

Power properties for these tests are reported elsewhere and indicate greater power to detect effects from the Jonckheere-Terpstra test whenever the data are consistent with dose-response monotonicity. When the data do not meet that requirement, the pairwise test of choice is Dunnett's provided the normality and variance homogeneity requirements are satisfied. With intersex and undifferentiated fish, these requirements are often not satisfied due to the prevalence of zero proportions in many reps. In such instances, the alternatives are the Dunn and the Mann-Whitney test, neither of which has good power for designs with only three reps per test concentration. Of the two, the exact permutation version of the Mann-Whitney test is recommended. There is no exact permutation version of Dunn's test available in commercial software and development of a program within a lab to carry out this test is a daunting undertaking.

References for Regression and NOEC

- Bruce R.D. and Versteeg D.J.** (1992); A statistical procedure for modeling continuous toxicity data. *Environ. Toxicol. Chem.* 11: 1485-1494.
- Collett D.** (1991); *Modelling Binary Data*, Chapman and Hall, London.
- Dunn O. J.** (1964); Multiple Comparisons Using Rank Sums, *Technometrics* 6, 241-252.
- Hochberg Y. and Tamhane A.C.** (1987); Multiple comparison procedures, Wiley, New York.
- Marcus R., Peritz E. and Gabriel, K.R.** (1976); On closed testing procedures with special reference to ordered analysis of variance, *Biometrika* 63, 655-660.
- OECD (2006)**; *Current Approaches in the Statistical Analysis of Ecotoxicity Data: A Guidance to Applications*, OECD Series on Testing and Assessment, Number 54, ENV/JM/MONO(2006)18, Environment Directorate, Organisation For Economic Co-Operation And Development, Paris 2006
- Rao J.N.K. and Scott A.J.** (1992); A simple method for the analysis of clustered binary data, *Biometrics* 48, 577-585.
- Schabenberger, O. ; B.E. Tharp, J.J. Kells, D. Penner** (1999); Statistical Tests for Hormesis and Effects Dosages in Herbicide Dose Response, *Agron. J.* 91, 713-721.
- Selwyn M. R.** (1988); Preclinical Safety Assessment, in K. E. Peace, ed., *Biopharmaceutical Statistics for Drug Development*, Marcel Dekker, New York.
- Slob W.** (2002); Dose-response modelling of continuous endpoints. *Toxicol. Sci.*, 66, 298-312
- Smejkal1, G. B., W. K. Thomas, D. Bauer, M. Kinter, A. Kwan, F. Witzmann, and H. Ringham** (2007); Proteomic analysis of individual *Daphnia* microcrustaceans, (http://www.pressurebiosciences.com/downloads/posters/2007/DGC_Smejkal.ppt)
- Westfall, P.H., Tobias R.D., Rom D., Wolfinger R.D. and Hochberg Y.** (1999); Multiple comparisons and multiple tests, SAS Institute, Cary, North Carolina.
- Williams D.A.** (1971); A test for differences between treatment means when several dose levels are compared with a zero dose control. *Biometrics*, 27, 103-117.

Appendix 10

