Draft validation report for the new Test Guideline for the Short-term Juvenile Hormone Activity Screening Assay using Daphnia magna (JHASA)

For first WNT-review by 22 December 2023
VALIDATION OF SHORT-TERM JUVENILE HORMONE ACTIVITY SCREENING ASSAY USING *Daphnia magna* (JHASA)

DRAFT REPORT

October 19, 2023
Foreword

This document presents the design and results validation exercise for the short-term Juvenile Hormone Activity Screening Assay (JHASA) using Daphnia magna. This method was developed to detect the potency of endocrine disrupting chemicals (EDCs) with juvenile hormone activity in cladoceran using proportion of male offspring (sex ratio) of Daphnia magna as endpoint. The project was proposed by Japan to the Working Group of the National Coordinators of the Test Guidelines Programme (WNT) in 2017.

First part of this report described the results of an international ring-test conducted in 2018-2019. Following the comments on the results of the ring-test provided in the Meeting of the Validation Management Group for Ecotoxicity Testing, another validation exercise was conducted in 2020-2021 to investigate the effect of non-chemical stress on the induction of male offspring production as described in the second part of this report. Based on the results of these validation exercises, the conclusions of the recommendations for the draft test protocol were summarized.
Table of Contents

40 Foreword ................................................................................................................. 3
41 Table of Contents ..................................................................................................... 4
42 Acknowledgments .................................................................................................... 6
43 1. Introduction .......................................................................................................... 8
44 2. Principle of the test .............................................................................................. 9
45 3. Report of the international ring test of JHASA .................................................. 11
46 3.1 Overview of the ring test .................................................................................. 11
47 3.1.1 Specific Goals ............................................................................................... 11
48 3.1.2 Participants of the international ring test ...................................................... 11
49 3.1.3 Overview of the Test Conditions .................................................................. 11
50 3.1.4 Test organism and culturing ......................................................................... 12
51 3.1.5 Test substances ............................................................................................. 12
52 3.1.6 Test concentrations ...................................................................................... 13
53 3.1.7 Test media ..................................................................................................... 13
54 3.1.8 Water quality parameters ............................................................................ 13
55 3.1.9 Observations and measurement variables ................................................... 13
56 3.1.10 Chemical analysis ....................................................................................... 14
57 3.1.11 Data assessment and validity of the test ...................................................... 14
58 3.1.12 Statistical analysis ....................................................................................... 14
59 3.2 Results ................................................................................................................ 16
60 3.2.1 Chemical analysis ......................................................................................... 16
61 3.2.2 Parent mortality ............................................................................................ 17
62 3.2.3 Reproduction ................................................................................................. 17
63 3.2.4 Male ratio ..................................................................................................... 18
64 3.2.5 Overall data evaluation ................................................................................ 22
65 3.3 Conclusions of the ring test .............................................................................. 23
66 4. Report of the validation study for non-chemical stress on male offspring production
   in JHASA .................................................................................................................. 24
67 4.1 Overview of the validation study of non-chemical stress .................................... 24
68 4.1.1 Specific Goals ............................................................................................... 24
69 4.1.2 Test Conditions ............................................................................................. 24
70 4.2 Results ................................................................................................................ 26
Acknowledgments

We would like to thank the following laboratories for kindly participating in this validation study coordinated by National Institute for Environmental Studies (NIES), Japan:

- L’Institut national de l’environnement industriel et des risques (INERIS), France
- Norwegian Institute for Water Research (NIVA), Norway
- Chemicals Evaluation and Research Institute (CERI), Japan
- IDEA Consultants, Inc., Japan
- KUMIAI CHEMICAL INDUSTRY CO., LTD., Japan
- Nisso Chemical Analysis Service Co. (NCAS), Japan

We are also deeply grateful to Prof. Taisen Iguchi (Yokohama City University) and Prof. Norihisa Tatarazako (Ehime University) for their advice and generous support and all who made constructive discussion at the VMG-eco expert meeting at the OECD.

This work was supported partly by a grant from the Ministry of the Environment, Japan.
Tables

Table 1 Test conditions for JHASA ................................................................. 9
Table 2 Participating laboratories in the ring test ......................................... 11
Table 3 Overview of key testing conditions in seven laboratories participated the inter-laboratory test ................................................................. 12
Table 4 Information of test substances .......................................................... 13
Table 5 Measured concentrations of 3,5-DCP ............................................. 16
Table 6 Measured concentrations of diofenolan ........................................ 16
Table 7 LOECs for parent mortality, reproduction, male ratio in diofenolan and 3,5-DCP test ............................................................... 22
Table 8 Overview of culture and JHASA conditions in non-chemical stress tests ........... 25
Table 9 Summary of male ratio in culturing period (Clone A) ......................... 28
Table 10 Summary of male ratio in 17 days culturing period (NIES, Clone A) ... 30

Figures

Figure 1 24-hour-old male (left) and female (right) of D. magna. Males can be distinguished from females by the length and morphology of the first antennae as shown in the circles (OECD, 2011) .................. 10
Figure 2 Statistical flow chart for male ratio analysis ...................................... 15
Figure 3 Number of number offspring (female and male), male ratio, and parent mortality in diofenolan test .............................. 19
Figure 4 Number of number offspring (female) and parent mortality in 3,5-DCP test .............................................................. 20
Figure 5 Boxplot of number of offspring in control and solvent control ......... 21
Figure 6 Mean time to second brood reproduction in each treatment ............. 21
Figure 7 Temperatures change test (culturing at 20°C and JHASA at 20, 28, 30°C) ................................................................. 26
Figure 8 Temperatures change test (culturing at 30 or 10°C and JHASA at 20°C) 26
Figure 9 Hardness change test (culturing at 250 mg CaCO3/L and JHASA at at 250, 500, 1000 mg CaCO3/L) .......................... 27
Figure 10 JHASA test from (a) the high-density culture and (b) the high-density culture with low feeding ... 27
Figure 11 No. of male and female offspring and male ratio in each observation day in culturing period ...... 29
Figure 12 No. of male and female offspring and male ratio in each observation day under different photoperiod conditions (NIES strain) ......................................................................................................................... 31
Figure 13 No. of male and female offspring and male ratio in each observation day under different photoperiod conditions (Clone A strain) ......................................................................................................................... 32
Figure 14 JHASA test under different photoperiod conditions (NIES, Clone A) ................. 33
Figure 15 Reproduction cycle and the critical period and sex determination of the second brood ....................... 36
Figure 16 The demonstration of statistical analysis ........................................... 37
1. Introduction

The cladoceran crustacean Daphnia magna is the recommended test organism for OECD Test Guideline (TG) 202 and TG 211 (OECD, 1998; 2004) due to relatively short life cycles and ease handling in the laboratory. The reproduction of D. magna is cyclic parthenogenetic, with alternating asexual and sexual reproduction cycles; females reproduce asexually only genetically identical females under favorable conditions, whereas sexual reproduction producing males under changing environmental conditions, such as shortening photoperiod, decreasing food concentration, and increasing population density (Hobaek and Larson, 1990; Kleiven et al., 1992). Females also produce haploid sexual eggs, which form resting egg after fertilization by males. Resting eggs in the ephippium can survive harsh environment and hatch once the environmental conditions improve (Hebert, 1978).

Juvenile hormones (JHs) are considered to play key roles in the developmental and reproductive processes of D. magna. Several studies have demonstrated that JHs are involved in the regulation of sex determination. Exposure to JHs and their analogs such as insect growth regulators (IGRs) pyriproxyfen and fenoxycarb induces the production of male neonates in the following brood of D. magna (Olmstead and LeBlanc, 2002, 2003; Tatarazako et al., 2003; Oda et al., 2005a). This male offspring induction can lead to adverse outcomes of regulatory concern on the population level (i.e., decline in population), as males cannot produce offspring (Tanaka et al., 2018). The same effects induced by JH analogs have also been observed in other cladoceran taxonomic groups: the genera Moina, Ceriodaphnia, Bosmina, and Oxyurellá (Oda et al., 2005b; Kim et al., 2006; Sinex and Sanoamuang, 2011). Therefore, male offspring production in cladoceran can be used as an endpoint for detecting endocrine disrupting chemicals with JH activities.

The sex of D. magna offspring can be distinguished by the length of the first antenna (longer in males) under a light microscope. Addition of this endpoint to the OECD TG 211 Daphnia magna reproduction test was proposed by Japan in 2004 and the validation studies including an international ring test were conducted from 2005 to 2007. Twelve laboratories from 6 countries participated in the ring test using JH agonist pyriproxyfen as the positive control and 3,5-dichlorophenol as the negative control. Using a genetically identical strain (provided by National Institute for Environmental Studies, Japan), concentration-dependent increases in the proportion of male offspring by pyriproxyfen were observed in all participating laboratories (OECD, 2008). After the discussion in the OECD Invertebrate Expert Group meeting in 2007, Annex 7 was added to TG 211 in 2008 to describe the procedure for determining neonate sex, if required. This is in line with the previous versions of this TG where sex ratio was an optional endpoint.

TG 211 Annex 7 can provide sufficient data for hazard and risk assessment; however, this test requires additional resources and cost to identify sex of all offspring produced for 21 days. For efficient screening of chemicals with JH activity, a short-term JH activity screening assay (JHASA) has been developed based on the knowledge that sex determination of offspring only occurs during the critical period of oocyst development in the ovary, i.e., approximately 7-10h before ovulation to the brood chamber (Kato et al., 2010). The JHASA protocol can significantly shorten the test duration and reduce the workload, as the test starts with exposure of matured females and only requires observation of the offspring sex from the second brood. The first brood in the brood chamber at the start of the test is not included in the observation, as the offspring sex has been already determined prior to the exposure. By employing this method, we successfully demonstrated that exposure to JH and JH agonists induced male offspring in the second broods in D. magna offspring (Abe et al., 2015a).
2. **Principle of the test**

The assay is initiated with adult females with developing embryos that are clearly observable in their brood chambers. Since the sex of ovulated eggs in the brood chamber at the start of this test has already been determined prior to the exposure, the subsequent brood (second brood during the exposure) was observed for offspring sex. An overview of the assay conditions are provided in Table 1. Identification of offspring sex is conducted by observing the length of the first antennae under a stereomicroscope (Figure 1). In principle, the test duration is set for 7 days to allow sufficient time for production of two consecutive broods. However, the test can be terminated as soon as the second brood is produced in each exposure unit (e.g., a beaker), whereas extended if not all test organisms have had the second brood within 7 days.

<table>
<thead>
<tr>
<th>Table 1. Test conditions for the ring test of JHASA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Test species (recommended) <strong>Daphnia magna.</strong> Other <strong>Daphnia</strong> sp. (e.g., <strong>Daphnia pulex</strong>) can also be used.</td>
</tr>
<tr>
<td>2. Test type</td>
</tr>
<tr>
<td>3. Test duration</td>
</tr>
<tr>
<td>4. Water temperature</td>
</tr>
<tr>
<td>5. Illumination quality</td>
</tr>
<tr>
<td>6. Photoperiod</td>
</tr>
<tr>
<td>7. Test chamber size</td>
</tr>
<tr>
<td>8. Age of test organisms at initiation</td>
</tr>
<tr>
<td>9. Number of organisms per replicate</td>
</tr>
<tr>
<td>10. Number of replicates per treatment</td>
</tr>
<tr>
<td>11. Number of treatments</td>
</tr>
<tr>
<td>11. Number of organisms per test</td>
</tr>
<tr>
<td>12. Feeding regime</td>
</tr>
<tr>
<td>13. Aeration</td>
</tr>
</tbody>
</table>
14. **pH**  
**Within the range 6 - 9**

15. **Dilution water**  
Reconstituted hard fresh water (e.g., Elendt M4 and M7 media)

17. **Biological endpoints**  
Male ratio of the second brood offspring  
The number of offspring (female, male) (optional)

18. **Test acceptability criteria**  
Mortality of parent daphnia <20% in the control(s)

---

**Figure 1.** 24-hour-old male (left) and female (right) of *D. magna*. Males can be distinguished from females by the length and morphology of the first antennae as shown in the circles (OECD, 2012).
3. Report of the international ring test of JHASA

3.1 Overview of the ring test

3.1.1 Specific Goals

1. Phase 1 validation study was conducted in five laboratories in Japan with positive (diofenolan) and negative (3,5-dichlorophenol) control substances, and phase 2 with the two other laboratories from France and Norway.

2. Applicability to the different *D. magna* strains was investigated in addition to compare the sensitivity difference in reproduction and the proportion of male offspring. As a preliminary validation, four other strains were also investigated with diofenolan (Annex 1).

3. Performance was compared between the laboratories. Reliability, reproducibility across laboratories, and sensitivity of the assay were determined.

3.1.2 Participants of the international ring test

Validation study was carried out by seven laboratories from three countries under the initiative of NIES (Table 2).

<table>
<thead>
<tr>
<th>Laboratories</th>
<th>Countries</th>
</tr>
</thead>
<tbody>
<tr>
<td>L’Institut national de l’environnement industriel et des risques (INERIS)</td>
<td>France</td>
</tr>
<tr>
<td>Norwegian Institute for Water Research (NIVA)</td>
<td>Norway</td>
</tr>
<tr>
<td>Chemicals Evaluation and Research Institute (CERI)</td>
<td>Japan</td>
</tr>
<tr>
<td>IDEA Consultants, Inc. (IDEA)</td>
<td>Japan</td>
</tr>
<tr>
<td>Kumiai Chemical Industry Co., Ltd.</td>
<td>Japan</td>
</tr>
<tr>
<td>Nisso Chemical Analysis Service Co. (NCAS)</td>
<td>Japan</td>
</tr>
<tr>
<td>National Institute for Environmental Studies (NIES)</td>
<td>Japan</td>
</tr>
</tbody>
</table>

3.1.3 Overview of the Test Conditions

Table 3 shows the overview of key testing conditions in seven participant laboratories: strain, age of test daphnids at the initiation of the test, culture density of daphnids in the mass culture, test medium, test temperature, feeding and renewal of the medium during the test, renewal of medium. Five laboratories (Lab A, B, C, D, E) followed almost the same test conditions except for minor differences in the test age and the culture density before the test. Lab F and Lab G used different strain, test medium (only Lab F) and food type.
Table 3. Overview of key testing conditions in seven laboratories participated in the ring test

<table>
<thead>
<tr>
<th>Draft protocol</th>
<th>Strain</th>
<th>Age</th>
<th>Culture density (/L)</th>
<th>Test medium</th>
<th>Temp (°C)</th>
<th>Feeding (/daphnid/d)</th>
<th>Renewal of medium</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lab A</td>
<td>NIES</td>
<td>11 d</td>
<td>20</td>
<td>M4</td>
<td>19.1-20.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lab B</td>
<td>NIES</td>
<td>13 d</td>
<td>35</td>
<td>M4</td>
<td>19.6-20.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lab C</td>
<td>USEPA a</td>
<td>14 d</td>
<td>31</td>
<td>M4</td>
<td>19.8-20.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lab D</td>
<td>NIES</td>
<td>10 d</td>
<td>35</td>
<td>M4</td>
<td>20.7-20.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lab E</td>
<td>NIES</td>
<td>16-18 d</td>
<td>35</td>
<td>M4</td>
<td>19.0-21.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lab F</td>
<td>Clone A</td>
<td>12 d</td>
<td>40</td>
<td>M4</td>
<td>19.1-20.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lab G</td>
<td>DHI b</td>
<td>14-15 d</td>
<td>18</td>
<td>M7</td>
<td>19.2-20.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a: USEPA is original strain of NIES strain. b: from Denmark. c: 2 volumes was added on Friday (at the test initiation), no feeding for weekend.

3.1.4 Test organism and culturing

All laboratories used *D. magna* Straus. In the draft protocol, NIES strain is recommended but Lab C, F, and G used other strains (Table 3). Lab A, B, D, and E used NIES strain which had been maintained for more than 30 years at NIES. Lab C used USEPA strain, which is the original strain of NIES strain. Lab F used Clone A strain, which is named as “Clone A” in the ring test of the OECD TG211 *Daphnia magna* reproduction test (OECD, 1997). Lab G used DHI strain which was originally obtained from a laboratory in Denmark (DHI Water & Environment) and has been sub-cultured at Lab G for more than 15 years.

Culturing density where the test adult daphnids were obtained (mass culture) was also described in Table 3. Lab A maintained daphnids in the lowest density (20 per 1 L) and Lab G in the highest density (40 per 1 L).

3.1.5 Test substances

Diofenolan (CAS 63837-33-2), a juvenile hormone agonist that was developed as an insect growth regulator (IGR), was chosen as a positive control based on the previous studies (Abe et al., 2015b, Annex 1).

3,5-dichlorophenol (CAS 591-35-5) (3,5-DCP) was chosen as the negative control because it has been widely used as a reference test compound in other international ring tests and has been proven to NOT induce any male offspring in the previous TG 211 ring tests (OECD, 2008).

NIES aliquoted each test substances from the same lot and distributed without indicating substance name and concentration (Table 4).
Table 4. Information of test substances

<table>
<thead>
<tr>
<th>Name</th>
<th>Diophenolan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>3,5-Dichlorophenol (3,5-DCP)</td>
</tr>
<tr>
<td>Molecular formula</td>
<td>C₁₈H₂₀NO₄</td>
</tr>
<tr>
<td>Molecular weight</td>
<td>300.35</td>
</tr>
<tr>
<td>Molecular weight</td>
<td>163.0</td>
</tr>
<tr>
<td>CAS no.</td>
<td>63837-33-2</td>
</tr>
<tr>
<td>CAS no.</td>
<td>591-35-5</td>
</tr>
<tr>
<td>Purity (%)</td>
<td>98.0%</td>
</tr>
<tr>
<td>Purity (%)</td>
<td>99.2%</td>
</tr>
<tr>
<td>Supplier</td>
<td>FUJIFILM Wako Pure Chemical Industries Corporation</td>
</tr>
</tbody>
</table>

3.1.6 Test concentrations

Three test concentrations of the two test substances were determined by NIES: 87.5, 175, 350 μg/L for diophenolan and 250, 500, 1000 μg/L for 3,5-DCP based on the *Daphnia* acute immobilization tests (OECD, 2004) which was pre-tested at NIES. The highest concentration was approximately half of the acute EC50, and two additional concentrations were set with a spacing factor of 2. Stock solution of 10,000 × each test concentration was prepared in dimethyl formamide (DMF, special grade chemical, FUJIFILM Wako Pure Chemical Industries Corporation, Osaka, Japan) and added to the test medium at a concentration of 0.01% (v/v). A solvent control (DMF only, .0.01%, v/v) and control were also subjected to the test. Lab F changed test concentrations of diophenolan: 62.5, 125, and 250 μg/L based on their preliminary test. In Lab G, additional two lower concentrations of diophenolan (21.9 and 43.8 μg/L) were tested.

3.1.7 Test media

Elendt M4 (OECD, 2011) was used as the test medium in the laboratories except for Lab G. Lab G used Elendt M7 (OECD, 2011).

3.1.8 Water quality parameters

Water quality parameters (temperature, pH, dissolved oxygen/DO) were measured at the start and end of each renewal period as recommended in TG 211.

3.1.9 Observations and measurement variables

The observations and measurement variables included:

- Mortality of the parent animals,
- Total number of the alive male and female neonates (offspring) produced in the second brood during an exposure period.
- Total number of the dead neonates and aborted eggs from the second brood and the parent ecdysis per replicate over the exposure period.
The test duration is basically 7 days. However, Lab B and Lab C expanded the observation until all daphnids produced the second brood or died.

The sex of neonates was identified under a stereomicroscope using the length and morphology of the first antenna as defining characteristics (OECD TG211 Annex 7, OECD, 2012). The offspring male ratio was defined as the proportion of males to the total number of neonates produced in second brood during the exposure at each concentration.

### 3.1.10 Chemical analysis

Chemical analysis of test substance concentrations was performed once at least to confirm the concentrations during the test except for Lab G. The samples from the same solution - when freshly prepared (new) and at renewal (old) - were subjected to solid phase extraction (SPE) in each participant and the eluates were transferred to NIES for GCMS (diofenolan) and LC/MS (3,5-DCP) analysis.

### 3.1.11 Data assessment and validity of the test

The main endpoint is the male ratio in the second brood offspring. The number of living male and female offspring from the second brood per parent animal were recorded for each test vessel and male ratio (the number of males/the number of total offspring) was calculated per replicate. Treatment of data with parental mortality was in accordance with TG211. First, Cochran-Armitage trend test was conducted to investigate whether the mortality follows a concentration-response pattern. If the test was significant (monotonic), the replicate with parental mortality was included to the analysis (assigned zero if parent animal died before second brood reproduction) whereas it was not (non-monotonic), the replicate with parental mortality was excluded in the reproduction analysis. Parental death after the second brood production was not considered.

In the case of male ratio, the replicates without second brood production were excluded from the analysis (The analysis will then be based on a reduced number of replicates).

In the ring test, the mortality of the parent animals in the controls should not exceed 20% at the end of the test as a solid validity criterion. In addition, if male offspring was observed in the controls, participants were requested to repeat the test.

### 3.1.12 Statistical analysis

The purpose of this JHASA is to detect the presence or absence of potential juvenile hormone activities of the test substance, not to obtain LOEC/NOEC. As a definitive test, TG211 Annex 7 (OECD, 2011) should be used when LOEC/NOEC are required to determine. However, if induction of male offspring is also observed in control(s) or the proportion of male offspring in the treatments is not substantial, statistical analysis may be required to evaluate a statistical significance of the male ratio in the exposure group compared to the control.

A statistical analysis to determine the NOEC/LOEC of reproduction (the number of total offspring) was performed in accordance with OECD TG 211 (OECD, 2011). Parent mortality was analyzed by step-down Cochran-Armitage test.

For male offspring, Figure 2 shows an example of statistical flow chart. First, a water control and a solvent control were statistically compared for each response (e.g., t test) and combined if no significant difference is found between the controls. Otherwise, the solvent control is used for NOEC determination.

If the data (normally, replicate means) are assumed to be a monotone trend (in an increase or a decrease), it is recommended to analyse the data by a trend-based step-down methods (e.g., step-down Williams test,
Jonckheere-Terpstra test). To assess monotonicity, data is evaluated by using linear and quadratic contrasts, although a visual check from a scatter plot can be used. Unless the quadratic contrast is significant and the linear contrast was not significant, the trend test is done. If the data is not monotonic, Dunnett’s test is used to determine treatment effects if the data were normally distributed with homogeneous variances. Use Tamhane-Dunnett if the data shows normality but heterogeneous variance. If assumptions are not met, Dunn’s test or Mann-Whitney test with a Bonferroni-Holm adjustment should be used. All tests are done independently of any overall F- or Kruskal-Wallis’s test. Further details are provided in OECD (2006).

Figure 2. Statistical flow chart for male ratio analysis
### 3.2 Results

#### 3.2.1 Chemical analysis

The time-weighted mean of measured concentration of 3,5-DCP and diofenolan are presented in Table 5 and Table 6, respectively. For 3,5-DCP, the time-weighted mean concentrations were 76–110% of the nominal concentrations and the concentration decrease over 2 days renewal period was only 81–96%. For diofenolan, the time-weighted mean concentrations were 65–135% of the nominal concentrations and the concentration change (Old/New) over 2-3 days renewal period was 59–120%. The slight increase during renewal was observed at 87.5 µg/L in Lab E and it may be due to the differences in recovery rate of SPE. The mean concentration change over renewal period was 86%, so that the concentrations of diofenolan were considered maintained well during the test. Note that the differences of mean concentrations among the laboratories (concentration ratio in Lab C or D/Lab A) were 2.1, 2.0, and 1.8 in 87.5, 175, and 350 µg/L of diofenolan, respectively.

Lab G did not perform chemical analysis, however, based on the measured concentrations in other laboratories, the concentrations of test substance were considered maintained well during the test.

The following test results were reported based on nominal concentrations.

#### Table 5. Measured concentrations of 3,5-DCP

<table>
<thead>
<tr>
<th>Laboratory</th>
<th>Control</th>
<th>Solvent</th>
<th>Control</th>
<th>250 µg/L</th>
<th>500 µg/L</th>
<th>1000 µg/L</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>New</td>
<td>Old</td>
<td>Mean</td>
</tr>
<tr>
<td>Lab A</td>
<td>N.D.</td>
<td>N.D.</td>
<td></td>
<td>270</td>
<td>240</td>
<td>250</td>
</tr>
<tr>
<td>Lab B</td>
<td>N.D.</td>
<td>N.D.</td>
<td></td>
<td>250</td>
<td>210</td>
<td>230</td>
</tr>
<tr>
<td>Lab C</td>
<td>N.D.</td>
<td>N.D.</td>
<td></td>
<td>240</td>
<td>200</td>
<td>220</td>
</tr>
<tr>
<td>Lab D</td>
<td>N.D.</td>
<td>N.D.</td>
<td></td>
<td>210</td>
<td>170</td>
<td>190</td>
</tr>
<tr>
<td>Lab E</td>
<td>N.D.</td>
<td>N.D.</td>
<td></td>
<td>260</td>
<td>210</td>
<td>230</td>
</tr>
<tr>
<td>Lab F</td>
<td>N.D.</td>
<td>N.D.</td>
<td></td>
<td>198</td>
<td>161</td>
<td>159</td>
</tr>
</tbody>
</table>

N.D.: Not detected. The data was not available for Lab G. a: concentration at freshly prepared, b: concentration at renewal, c: time weighted mean concentration (In Lab F, the mean concentrations of New and Old were shown (n=3)).

#### Table 6. Measured concentrations of diofenolan

<table>
<thead>
<tr>
<th>Laboratory</th>
<th>Control</th>
<th>Solvent</th>
<th>Control</th>
<th>87.5 µg/L</th>
<th>175 µg/L</th>
<th>350 µg/L</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>New</td>
<td>Old</td>
<td>Mean</td>
</tr>
<tr>
<td>Lab A</td>
<td>N.D.</td>
<td>N.D.</td>
<td></td>
<td>60.9</td>
<td>53.1</td>
<td>56.9</td>
</tr>
<tr>
<td>Lab B</td>
<td>N.D.</td>
<td>N.D.</td>
<td></td>
<td>95.5</td>
<td>74.2</td>
<td>84.4</td>
</tr>
<tr>
<td>Lab C</td>
<td>N.D.</td>
<td>N.D.</td>
<td></td>
<td>94.7</td>
<td>77.7</td>
<td>85.9</td>
</tr>
<tr>
<td>Lab D</td>
<td>N.D.</td>
<td>N.D.</td>
<td></td>
<td>133</td>
<td>103</td>
<td>117</td>
</tr>
<tr>
<td>Lab E</td>
<td>N.D.</td>
<td>N.D.</td>
<td></td>
<td>78.4</td>
<td>94.0</td>
<td>86.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Laboratory</th>
<th>Control</th>
<th>Solvent</th>
<th>Control</th>
<th>62.5 µg/L</th>
<th>125 µg/L</th>
<th>250 µg/L</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>New</td>
<td>Old</td>
<td>Mean</td>
</tr>
<tr>
<td>Lab F</td>
<td>N.D.</td>
<td>N.D.</td>
<td></td>
<td>59.0</td>
<td>44.3</td>
<td>51.7</td>
</tr>
</tbody>
</table>

The concentrations were time-weighted mean. N.D.: Not detected. The nominal concentrations in Lab F were lower than those in Lab A-E based on their preliminary testing. The measured data was not available for Lab G. a: concentration at freshly prepared, b: concentration at renewal, c: time weighted mean concentration (In Lab F, the mean concentrations of New and Old were shown (n=3)).
3.2.2 Parent mortality

Performance in the controls

Except for Lab E, the two substances were tested simultaneously using common water and solvent controls. In both controls and solvent controls, no parent female died until the release of the second brood, as reported by all seven laboratories. The validity criterion was therefore met in the ring test. In Lab F, 1 daphnid died the day after the release of the second brood (on Day 7), but it could be excluded in data analysis because the calculations were made based on the data until the release of the second brood.

Effects of the test substances

No parent mortality was observed in Lab A, B, C, and G in all test concentration of diofenolan (Figure 3). In Lab D, 10 and 40% mortality were observed at 87.5 and 350 µg/L, respectively. In Lab E, 20% mortality was observed at 350 µg/L. In Lab F, 10% mortality was observed at 125 µg/L. Since the parent mortality in Lab F was not dose-dependent, the replicate with parent mortality was excluded from reproduction and male ratio analysis.

The lethal toxicity of 3,5-DCP varied among laboratories (Figure 4). In Lab A, high parent mortality (>70%) was observed even in the lowest concentration (250 µg/L), whereas only 20% mortality at 500 and 1000 µg/L was observed in Lab G. In Lab A, B, D, E, F and G, concentration-dependent mortality was observed, therefore, the replicates containing dead adults were included in the following analysis. In contrast, Lab C observed concentration-independent mortality, therefore the replicates with parental mortality were excluded from the data analysis.

3.2.3 Reproduction

Performance in the controls

The mean number of living offspring produced in the second brood in the controls and solvent controls was greater than 40 in Lab A and C (Figures 3, 4). In contrast, an average of 20 was reported by Lab B, E, F and G, while the smallest was 12.4 ± 2.5 (Mean ± SD) in Lab D. Even with the same strain (Lab A, B, D, and E), the reproduction in controls was different by a factor of 3.7. The mean number of offspring among seven laboratories was 27.3 in control and 28.1 in solvent control, and the median was 25.3 and 26.0, respectively (Figure 5).

Effects of the test substances

The effects of diofenolan and 3,5-DCP on the number of offspring are presented in Figure 3 and Figure 4, respectively. The total number of offspring (reproduction) decreased in a concentration-dependent manner after exposure to diofenolan in all laboratories (Figure 3). Except for Lab D, no living offspring from the second brood was observed at the highest concentration (350 µg/L in Lab A, B, C, E, G, 250 µg/L in Lab F). Most of the laboratory observed aborted egg or dead offspring at the highest concentration. Even at 87.5 µg/L, reproduction inhibition was 30% (Lab E) – 61% (Lab B). Lab G tested additional two lower concentrations (43.8 and 21.9 µg/L) and no significant difference was observed in these concentrations.

In the diofenolan exposure, a delay in reproduction was also observed. Figure 6 showed the mean duration (days) to the second brood in all treatments. Since Lab A, E, F, and G did not observe the release of the second brood in some replicates during a 7-day exposure, 8 days was assigned for mean time calculation. In Lab B and C, second brood production in the diofenolan treatment was observed on Day 8 and Day 9 while the second brood production in controls occurred on Day 6 or Day 7. The delay of mean time of the second brood production compared with solvent control was 0.5–2.0 days (Lab G–Lab F), 0.7–2.6 days (Lab G–Lab F), and 1.0–3.4 days (Lab D–Lab F) in 87.5, 175, and 350 µg/L, respectively.
In 3,5-DCP exposure, Lab A, B and C demonstrated significant decreases in reproduction in all the three concentrations and Lab D and F demonstrated significant decreases at 500 and 1000 µg/L, whereas Lab E and G did not show significant effects on reproduction. Lab A hardly observed any reproduction due to high parental mortality. Therefore, reproduction NOEC of 3,5-DCP was <250 µg/L for Lab A, B and C, 250 µg/L for Lab D and F, and 1000 µg/L for Lab E and G.

3.2.4 Male ratio

Controls

Male offspring was not found in either controls or solvent controls in all laboratories (Figures 3, 4) (Participants were requested to repeat the test when the male was not induced in the controls.).

Effects of the test substances

The effects of diofenolan and 3,5-dichlorophenol on the male ratio are presented in Figures 3 and 4, respectively.

The production of male neonates was induced in response to diofenolan exposure in all concentrations in all the laboratories. The three concentrations (87.5, 175, 300 µg/L in Lab A-E, G, 62.5, 125, 250 µg/L in Lab F) demonstrated 100% male ratio in all the laboratories and Lab G also demonstrated 81% and 98% male ratio at additional lower concentrations, 43.8 and 21.9 µg/L, respectively.

In contrast, no male offspring was observed in all the concentration of 3,5-DCP in all laboratories.
Figure 3. Number of offspring (female and male), male ratio, and parent mortality in diofenolan test

Solvent: solvent control. Laboratories were named randomly with alphabet letters, which correspond to those in other Tables and Figures. Mean ± SD. Asterisks indicate statistically significant difference compared with the controls (*: p < 0.05, **: p < 0.01; step-down Jonckheere-Terpstra test for the total number of offspring and male ratio except for the total number of offspring in Lab G (Dunn test), step-down Cochran-Armitage test for parent mortality).
Laboratories were named randomly with alphabet letters, which correspond to those in other Tables and Figures. Mean ± SD. Asterisks indicate statistically significant difference compared with the controls (*: p < 0.05, **: p < 0.01; step-down Jonckheere-Terpstra test for the total number of offspring except for Lab D (Dunn’s test) and step-down Cochran-Armitage test for parent mortality shown in red color).
Figure 5. Boxplot of number of offspring in control and solvent control

Figure 6. Mean time to the second brood reproduction in each treatment

">" indicates there are replicates where the second brood was not reproduced within 7 days (It was assumed to be 8 days for these replicates to calculate the mean value). DFN: diofenolan. DCP: 3,5-DCP. In Lab F, concentrations of DFN were 62.5, 125, 250 ppb.
3.2.5 Overall data evaluation

The LOECs for all endpoints are summarized in Table 7. Although the effect on parent mortality and reproduction varied among laboratories, response to male offspring production was reproducible in both negative and positive controls; No male offspring was observed in the control, solvent control, or any concentrations of 3,5-DCP (negative control) and diofenolan (positive control) approximately 100% male offspring was observed in all the laboratories.

Most of studies on JH/JHA observed reduced reproduction at concentrations lower than that could induce male offspring (Tatarazako, 2003, Oda et al., 2005a, 2005b, 2006, 2007, Olmstead and LeBlanc, 2002, 2003). Therefore, sufficient range finders for the test concentrations is needed to ensure that enough second brood offspring are produced and the measurement of this endpoint is not influenced by other types of effects. In the ring test, Lab A to Lab E followed the same concentrations which were preliminarily determined by NIES based on the acute immobilization tests, however, strain-related sensitivities need to be taken into account and the test concentrations should be carefully selected based on the acute immobilization test with their corresponding strains in different laboratories, as suggested in the draft test guideline.

In addition, mean time to the release of the second brood was different among treatments and laboratories. The difference in controls was probably dependent on the status of reproduction cycle (approximately 72 h in general) at the initiation of the test. If the parent animals containing embryos at the late stage of embryonic development in the brood chamber were used, the first brood reproduction normally occurs 1-2 days after the initiation of the exposure, and second broods can normally be observed after 4-5 days. If the parent animals containing embryos at the early developmental stage were used, the first brood is normally released 2-3 days after initiation of the test and the second broods may be observed after 5-6 days. The treatment effects were also affected by the strain sensitivity differences to different chemicals. Diofenolan exposure delayed reproduction for maximum 3.4 days at the highest concentration in Lab F. Therefore, the test duration in the draft protocol (i.e. 7 days) should be modified to allow possible extension of the test duration, if necessary. However, a maximum test duration should be determined as reference considering the concept for the test as a screening assay. Assuming a 3-4 days delay in reproduction is acceptable, the test could be terminated at the end of 10 days (mean time to second brood production 6 days + 3-4 days =9-10 days).

Table 7. LOECs for parent mortality, reproduction, male ratio in diofenolan and 3.5-DCP test

<table>
<thead>
<tr>
<th>Laboratory</th>
<th>Diofenolan (positive control)</th>
<th>Reproduction</th>
<th>Male ratio</th>
<th>3.5-DCP (negative control)</th>
<th>Reproduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lab A</td>
<td>&gt;350 (234)</td>
<td>87.5 (56.9)</td>
<td>87.5 (56.9)</td>
<td>250 (250)</td>
<td>250 (250)</td>
</tr>
<tr>
<td>Lab B</td>
<td>&gt;350 (384)</td>
<td>87.5 (84.4)</td>
<td>87.5 (84.4)</td>
<td>1000 (1090)</td>
<td>250 (230)</td>
</tr>
<tr>
<td>Lab C</td>
<td>&gt;350 (414)</td>
<td>87.5 (85.9)</td>
<td>87.5 (85.9)</td>
<td>&gt;1000 (980)</td>
<td>250 (220)</td>
</tr>
<tr>
<td>Lab D</td>
<td>350 (370)</td>
<td>87.5 (117)</td>
<td>87.5 (117)</td>
<td>1000 (900)</td>
<td>500 (400)</td>
</tr>
<tr>
<td>Lab E</td>
<td>350 (374)</td>
<td>87.5 (86.0)</td>
<td>87.5 (86.0)</td>
<td>500 (490)</td>
<td>&gt;1000 (1070)</td>
</tr>
<tr>
<td>Lab F</td>
<td>&gt;250 (210)</td>
<td>62.5 (51.3)</td>
<td>62.5 (51.3)</td>
<td>1000 (756)</td>
<td>1000 (757)</td>
</tr>
<tr>
<td>Lab G</td>
<td>&gt;350</td>
<td>87.5</td>
<td>21.9</td>
<td>500 (756)</td>
<td>&gt;1000</td>
</tr>
</tbody>
</table>

Male ratio was not available for 3,5-DCP because no male offspring was observed. The values in parentheses were measured mean concentrations.
3.3 Conclusions of the ring test

Juvenile hormones and their analogs are involved in sex determination of the daphnids and thus exposure to these chemical substances induce *D. magna* to produce male neonates. Based on the existing OECD Test Guideline No. 211 (Daphnia reproduction test) Annex 7, a short-term screening assay to detect potential juvenile hormone activity of the test substance was developed, and an inter-laboratory study was conducted to investigate the reproducibility of the test protocol. Two substances were tested: a positive control (juvenile hormone receptor agonist, diofenolan) and a negative control (3,5- dichlorophenol). Results showed that the effects on the proportion of male offspring (sex ratio) produced in the second brood during exposure period were the same in all seven laboratories using four different strains of *D. magna*. No male offspring were observed in the control, solvent control or the 3,5-dichlorophenol (negative control) treatments in all the laboratories, while diofenolan (positive control) treatments induced male offspring at all nominated concentration in all laboratories. Although the total number of living offspring in the second brood and survival of parent daphnids were not the endpoints of interest, the effects on reproduction and parent survival exposed to each test substance were different across the participating laboratories, suggesting that the test concentrations should be carefully selected according to the strain sensitivity to the test chemical (i.e. acute immobilization test) in each laboratory.
4. Report of the validation study for non-chemical stress on male offspring production in JHASA

4.1 Overview of the validation study of non-chemical stress

4.1.1 Specific Goals

1) The purpose of this validation study is to confirm the effect of non-chemical stress (i.e. temperature, hardness, density, food amount, photoperiod) on the induction of male offspring production.

2) Frequency and proportion of male offspring production in the stock culture were also investigated to understand the environmental conditions (i.e. photoperiod and density + food amount) that induces male offspring.

4.1.2 Test Conditions

Table 8 shows the overview of the culture condition and JHASA conditions.

For temperature, in the first experiment, we changed only JHASA condition to higher temperature (28°C, 30°C). In the second experiment, we changed culture conditions to higher (28°C) and lower (10°C) temperature from first instar offspring (<24 h old) to 17 d old and returned to normal test temperature (20°C) at JHASA.

For high hardness, we usually use Element M4 media with harness 250 mgCaCO₃/L for culturing. To evaluate the effect of increased hardness, we prepared the media with 2-fold and 4-fold concentrations (500, 1000 mgCaCO₃/L) and conduct JHASA using these medium.

For high density, we prepared three culturing condition with different density: 35 daphnids/L (normal culturing condition in NIES), 70 daphnids/L (twice the normal) and 105 daphnids/L (3 times the normal). In the first experiment, we also increased food amount so that ration per daphnid remains the same (feed 1, 2, 3 mL of 5×10⁸ cells/mL Chlorella sp and YCT to 35, 70, 105 daphnids, respectively). In the second experiment, we added the same food amount (1 mL of 5×10⁸ cells/mL Chlorella sp and 1 mL YCT/L/day) to each density condition so that the food amount per daphnid decreased (×1, ×1/2, ×1/3 of normal amount). Following JHASA was conducted as usual.

Additional experiments to observe the effect of high density with lower food amount on induction of male offspring in the mass culture were conducted since JHASA cannot be conducted at high density condition (1 daphnid/vessel). These experiments were carried out by Clone A stain which sometimes produces male offspring under high density condition. First, twice the normal density of daphnids (70 daphnids/L) were housed in 1 L of culture medium from <24 h old offspring and 100 offspring was randomly collected from the culture and identified sex in weekday for 28 days. During the culturing, ×1/2 of normal amount of food/daphnid (i.e. 1 mL of 5×10⁸ cells/mL Chlorella sp and 1 mL YCT/L/day) was added every weekday. Secondly, a small-scale culture with 10 daphnids per 50 mL culture medium was investigated for 21 days. The density of this small-scale culture (200 daphnids/L) was equivalent to ×6 of the normal condition. Food amount per daphnid was 5, 10, and 20 µL of Chlorella suspension, which were equivalent to ×0.18, ×0.35, ×0.70 of the normal food amount per daphnid in the mass culture. The small-scale test was conducted in four replicates and sex identification was conducted for all offspring produced in 23 days.

For photoperiod, it was reported that short day-length lower than 12 h increased male ratio in D. pulex WTN6 strain (Toyota et al., 2015, 2017) and D. magna LRV13.2 strain (Toyota et al., 2021). Therefore, four different light conditions (0, 10, 16, 24 h) were tested using NIES strain and Clone A strain. Since switch of male/female offspring production occurred 3-7 days after photoperiod change in D. pulex WTN6 strain (Toyota et al 2017), the daphnids were acclimated from <24 h offspring to each photoperiod and start JHASA from 16-18 d old. During the acclimation, the sex of 100 offspring randomly collected from the mass culture was checked in weekdays.
<table>
<thead>
<tr>
<th>Stress type</th>
<th>Stain</th>
<th>Culture conditions</th>
<th>JHASA conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>NIES</td>
<td>20°C (Normal)</td>
<td>• 20°C (normal)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30°C (High)</td>
<td>• 28°C (higher than normal)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10°C (Low)</td>
<td>• 30°C (higher than normal)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• 20°C (normal)</td>
</tr>
<tr>
<td>High hardness</td>
<td>NIES</td>
<td>250 mgCaCO₃/L (M4 medium, normal)</td>
<td>• 250 mgCaCO₃/L (M4 medium, normal)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• 500 mgCaCO₃/L (×2 hardness of M4 medium)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• 1000 mgCaCO₃/L (×4 hardness of M4 medium)</td>
</tr>
<tr>
<td>High density</td>
<td>NIES</td>
<td>• 35 daphnids/L (Normal density at NIES)</td>
<td>• 1 daphnid/vessel (normal)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 70 daphnids/L (×2 of normal density)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 105 daphnids/L (×3 of normal density)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Same food amount per daphnids</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NIES</td>
<td>• 35 daphnids/L (Normal density at NIES)</td>
<td>• 1 daphnid/vessel (normal)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 70 daphnids/L (×2 of normal density)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 105 daphnids/L (×3 of normal density)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 1 daphnid/vessel (normal)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Clone A</td>
<td>70 daphnids/L (density: ×2, food per daphnids: ×1/2 of normal condition), from offspring to adult (28 days)</td>
<td>Not performed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10 daphnids/50 mL (density: ×6, food per daphnids: ×0.18, ×0.35, ×0.70 of normal condition), from offspring to adult (23 days)</td>
<td></td>
</tr>
<tr>
<td>Photoperiod</td>
<td>NIES, Clone A</td>
<td>0 h light: 24 h dark</td>
<td>same photoperiod condition as in culture (other test conditions were as usual)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10 h light: 14 h dark</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>16 h light: 8 h dark (normal)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>24 h light: 0 h dark</td>
<td></td>
</tr>
</tbody>
</table>
4.2 Results

4.2.1 Temperature

After rearing at 20°C until 17 d old, the JHASA tests were conducted at 20, 28, and 30°C (Figure 7) using NIES strain. No males were induced in either the first test or second test although increase of the number of offspring was observed at 28°C. On the other hand, daphnids cultured at 30°C and followed by the JHASA test at 20°C could not have reproduction due to the heat stress during the culturing (Figure 8). When the JHASA test 20°C after rearing at 10°C, male offspring was not observed (Figure 8). These results suggest that temperature stress does not induce male offspring in the JHASA test.

![Figure 7. Temperatures change test (culturing at 20°C and JHASA at 20, 28, 30°C) (NIES strain)](image)

No. of female and male offspring (Mean ± SD, n=10) and parent animal mortality. No males were observed.

![Figure 8. Temperatures change test (culturing at 30 or 10°C and JHASA at 20°C) (NIES strain)](image)

No. of female and male offspring (Mean ± SD, n=10) and parent animal mortality. No males were observed.

4.2.2 Hardness

The JHASA tests using ×2 and ×4 hardness of M4 did not observe male offspring in either the first test or second test. Therefore, the hardness increase did not affect male offspring induction (Figure 9).
Figure 9. Hardness change test (culturing at 250 mg CaCO$_3$/L and JHASA at 250, 500, 1000 mg CaCO$_3$/L) (NIES strain)

No. of female and male offspring (Mean ± SD, n=10) and parent animal mortality. No males were observed.

4.2.3 High density and low feeding

Figure 10 shows the JHASA test results using NIES strain cultured at high density. The first experiment (Figure 10(a)) kept the same amount of food per daphnid and there was no significant different in reproduction compared with the normal density condition (35 daphnids/L). In the second experiment, the food amount per daphnid was decreased (×1, ×1/2, ×1/3 of normal amount) so that the reproduction was decreased compared with the normal density and feeding condition. However, no male offspring was observed either high density condition or high density + lower food amount condition. Thus, stress of high density and lower nutritional status of culturing conditions immediately prior to the test is not likely to induce male offspring, however, it is possible that the results may differ from this study if daphnids are reared under these stresses for a longer period.

Figure 10. JHASA test from (a) the high-density culture and (b) the high-density culture with low feeding (NIES strain)

No. of female and male offspring (Mean ± SD, n=10) and parent animal mortality. No males were observed.
The above experiment only observed sex ratio of offspring during the JHASA under the normal density and feeding condition. Clone A strain may produce male offspring during high-density and low-feeding culturing condition. Therefore, the male ratio of offspring of Clone A strain cultured at high-density and low-feeding condition was observed. The mean male ratio was increased 2-3 factors (3.8% and 4.0% in ×2 density mass culture and small-scale) compared with that in the normal condition (1.3%) (Table 9). Highest male ratio among each observation day was 17% and 86% in ×2 density mass culture and small-scale (×6 density), respectively (Figure 11).

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Density (daphnids/L)</th>
<th>Culturing period</th>
<th>No of observations</th>
<th>Male ratio (%) mean ± SD</th>
<th>range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal mass culture</td>
<td>35</td>
<td>17 days</td>
<td>7</td>
<td>1.3 ± 2.3</td>
<td>0–6</td>
</tr>
<tr>
<td>×2 density mass culture</td>
<td>70</td>
<td>28 days</td>
<td>15</td>
<td>3.8 ± 5.6</td>
<td>0–17</td>
</tr>
<tr>
<td>Small-scale (×6 density)</td>
<td>200</td>
<td>23 days</td>
<td>11</td>
<td>4.0 ± 6.8</td>
<td>0–86</td>
</tr>
</tbody>
</table>

Mean and standard deviation (SD) of male ratio were calculated from each observation (n=7~11). In small-scale test, replicate mean was first calculated in each observation and then calculated the mean during the observation.
Figure 11. No. of male and female offspring and male ratio in each observation day in culturing period (Clone A strain)

(a) Normal mass culture, (b) x2 density mass culture, (c) Small-scale (x6 density) (mean± SD, n=4), 100 offspring was randomly collected from the mass culture and all offspring was observed in the small-scale test.
4.2.4 Photoperiod

The mean ± SD of male ratio observed during culturing period under different photoperiod conditions was summarized in Table 10 and data of each observation day was shown in Figure 12 and Figure 13. The mean male ratio of NIES strain was slightly increased to 5.0% in the 1st test of 0 h light condition (24 h darkness) but it was 0.8% in the 2nd test, which was comparable to normal 16 h light condition (0%). Clone A demonstrated 13.5 and 12.0% under 0 h light condition, which were higher than those in NIES strain. Highest male ratio in 0 h light condition among each observation day was 22 and 25% in the 1st test of NIES and Clone A and 4% and 29% in the 2nd test of NIES and Clone A (Figures 12, 13). The subsequent JHASA using NIES under 0 h light condition observed 10 ± 32% male ratio (Mean ± SD) in the 1st test, however, it was not reproducible in the 2nd test (Figure 14). Although slightly higher male ratio was observed in the mass culture of Clone A, the JHASA using this mass culture did not induce any male offspring. Under 10 and 24 h light condition, the mean male ratio in the mass culture was approximately 1-2% in both NIES and Clone A, which was comparable to normal 16 h light condition (NIES: 0%, Clone A: 1.3%), and the JHASA under 10 and 24 h light condition did not observe male offspring induction. Therefore, day-length changes did not induce male offspring production in JHASA although more than 10% male was observed in the mass culture of Clone A under the darkness condition.

Table 10. Summary of mean male ratio (minimum and maximum values in parentheses) in 17 days culturing period (NIES and Clone A strains)

<table>
<thead>
<tr>
<th>Photoperiod condition</th>
<th>NIES</th>
<th>Clone A</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 h light (1st test)</td>
<td>5.0% (0–22%)</td>
<td>13.5% (0–25%)</td>
</tr>
<tr>
<td>0 h light (2nd test)</td>
<td>0.8% (0–4%)</td>
<td>12.0% (0–29%)</td>
</tr>
<tr>
<td>10 h light</td>
<td>1.1% (0–6%)</td>
<td>0.9% (0–2%)</td>
</tr>
<tr>
<td>16 h light</td>
<td>0%</td>
<td>1.3% (0–6%)</td>
</tr>
<tr>
<td>24 h light (1st test)</td>
<td>1.7% (0–7%)</td>
<td>1.7% (0–11%)</td>
</tr>
<tr>
<td>24 h light (2nd test)</td>
<td>1.0% (0–5%)</td>
<td>0.8% (0–3%)</td>
</tr>
</tbody>
</table>
Figure 12. No. of male and female offspring and male ratio in each observation day under different photoperiod conditions (NIES strain)

(a) 0 h light (1st test), (b) 0 h light (2nd test), (c) 10 h light, (d) 16 h light, (e) 24 h light (1st test), (f) 24 h light (2nd test). 100 neonates were randomly collected from the mass culture and all offspring was observed in the small-scale test.
Figure 13. No. of male and female offspring and male ratio in each observation day under different photoperiod conditions (Clone A strain)

(a) 0 h light (1st test), (b) 0 h light (2nd test), (c) 10 h light, (d) 16 h light, (e) 24 h light (1st test), (f) 24 h light (2nd test). 100 neonates were randomly collected from the mass culture and all offspring was observed in the small-scale test.
Figure 14. JHASA test under different photoperiod conditions (NIES and Clone A strains)

(a) 0 h light (1st test), (b) 0 h light (2nd test), (c) 10 h light, (d) 16 h light, (e) 24 h light.

No. of female and male offspring (Mean ± SD, n=10) and parent animal mortality. No males were observed.
4.3 Conclusions of non-chemical stress

- Higher or lower temperature, high hardness, high density, and high density with low feeding did not induce males in JHASA.

- Constant darkness condition produced approx. 20% male in Clone A strain 2 weeks after acclimation in mass culture, however, male offspring was hardly observed in JHASA where the daphnids individually held in vessels.

- There was strain difference in male induction by day-length condition and density. Clone A strain is relatively sensitive to these stress in male offspring induction compared to NIES strain.

- It is desirable to find appropriate culturing conditions where male offspring is hardly induced in each strain.
5. Discussion

5.1 Sources of variability in ring test

In the ring test, there were no significant difference between laboratories in the presence or absence of male offspring although the four different strains were used. However, the total number of offspring in controls and the sensitivity to 3,5-DCP were different among laboratories even though Lab A, B, D, and E used the same NIES strain. The results of three different strains (USEPA, Clone A, DHI) used in Lab C, F, G were not significantly different from NIES strain, suggesting that the other factors affected the variability. In any biological test, the condition of the organisms and their acclimatization is a potentially large source of variability. The age of test organisms ranged from 11 days old to 18 days old among laboratories, but analysis of control data did not show correlation between the total number of offspring and the age. The culturing condition of the daphnids, such as feeding regime, might affect the control reproduction.

5.2 Recommended strain

Draft protocol described that “the recommended strains to be used is NIES”, however, as described in section 6.1, the results of the other strains (USEPA, Clone A, DHI) were not significantly different from NIES strain in the ring test, indicating that these strains also can be used in JHASA. Preliminary validation using four other strains also demonstrated that male offspring production was induced by the exposure to diofenolan (Annex 1). As suggested in the validation study for non-chemical stress and the preliminary validation in Annex 1, Clone A strain is relatively sensitive to environmental stress and more likely to induce male offspring compared to NIES strain, so that careful culturing is recommended to avoid producing male. Therefore, the protocol should be revised as follows

Paragraph 15: In the property of the test, strains that are less likely produce male offspring in control condition are suitable. For example, the strain from NIES (National Institute for Environmental Studies) and USEPA (United States Environmental Protection Agency), Clone A strain, and DHI strain (DHI Water & Environment, Denmark) are acceptable, which did not produce male offspring in the controls in the ring test. However, Clone A sometimes induces males in controls probably due to high density in the mass culture just before the test, so care should be taken to maintain healthy stock without males by adjusting optimal density and feeding condition.

5.3 Test duration

Draft protocol determined the test duration as 7 days because the second brood reproduction are usually observed within 7 days. In the ring test, five laboratories observed the second brood in the controls within 6 days whereas six laboratories did not observe the second brood in the diofenolan exposure groups within 7 days (Figure 5). The objective of the JHASA is to determine the sex of offspring from the second brood of which sex is determined after the exposure, test duration can be extended until the daphnids produce the second brood. In the diofenolan exposure of the ring test, two laboratories observed the second brood production on Day 8 and Day 9. Given the properties as a screening assay, the reproduction inhibition may also be induced in the higher concentrations, it is not appropriate to extend the test duration for too long. Therefore, we suggest that the typical test duration is 7 days, and it is recommended to end the test within 10 days even the second brood cannot be observed in some replicates in the higher concentrations.

5.4 Selection of daphnids at the initiation of test

Draft protocol suggests that 10–17-day old females with eggs visible in the brood chamber are used. Neonates from this brood (first brood in the test) are not used for the sex ratio scoring as sex determination may have occurred prior to the test. The susceptible period of sex determination is late oocytes developmental stage in the ovary, which is typically 7-10 h before releasing of the previous brood in the brood chamber (Kato et al., 2010, Abe et al., 2015). When neonates from the first brood have hatched in the brood chamber, it is possible that the
sex of the neonates from the second brood in the ovary has already determined. Therefore, we should avoid using daphnids with hatched neonates in the brood chamber (>48 h after ovulation) because the susceptible period of sex determination of the second brood may be missed.

In addition, if a readily degradable chemical substance is tested, we should take care to maintain the exposure concentrations during the sex determination period. If we start the test with females that have just ovulated egg to the brood chamber and the renewal period of the solution is every two days as in the ring test, the sex determination of the second brood occurs around just before the water renewal (concentration shown in black line in Figure 15) suggesting that the concentration may be lowest during the sex determination period. To avoid this situation, one is to renew the solution frequently until the reproduction of the first brood (concentration shown in red line in Figure 14), and the other is to use females with offspring 40-50 h post ovulation. It is difficult to obtain adult females of which ovulation time are known and synchronized, frequent renewal of the solution is a practical measure. These recommendations were added to the paragraphs “Preparation of Test animals” and the paragraph “Test medium renewal” of the draft test guideline.

![Figure 15. Reproduction cycle and the critical period and sex determination of the second brood.](image)

5.5 Statistical power analysis

Since the JHASA is not intended to determine NOEC for risk assessment, it is not necessary to perform statistical analysis for male ratio. However, if induction of male offspring is also observed in control(s) or the proportion of male offspring in the treatments is not substantial, statistical significance of the proportion of male offspring in the treatments help interpretation of the result.

Discussion of statistical issues for appropriate analysis of proportion response (i.e. male ratio) have been reported in the validation report of an enhancement of OECD TG 211 Daphnia reproduction test (OECD, 2008) and the schematic diagram of the statistical decision procedure was suggested. In addition, following the diagram for sex ratio used in fish sexual development test (OECD, 2011b), an example of statistical flow chart is shown in Figure 2.

Demonstration of analysis using the suggested analytical methods was performed to investigate the detection power on male ratio. Figure 16 shows examples of demonstration analysis following the statistical flow chart suggested in Figure 1. Usually, as suggested in the ring test, the proportion of males in the control(s) and one or more low concentrations is usually zero or close to zero in every replicate. This means that there will be little or no variation among replicate means, so that the assumption of normality required by Dunnett and Williams tests cannot be satisfied regardless of data transformation. Therefore, appropriate alternative tests considered are step-down Jonckheere-Terpstra test, and the Dunn or Mann-Whitney U test with a Bonferroni-Holm adjustment.

If no male offspring is induced in the controls and the two low test concentrations (C1, C2), Jonckheere-Terpstra test can detect significant difference when 3 and more replicates produce male offspring (Figure 16(1)ab). Since the proportion of male in each replicate does not affect the results in non-parametric test, very small male ratio, for example, only 0.1% of male ratio in the three replicates in C3 (0.03 ± 0.05%, replicate mean ± SD) shows statistically significant difference by Jonckheere-Terpstra test. If pooled 20 replicates from a dilution water control and a solvent control are used, statistically significance is found in C3 with 2 replicates producing male offspring (Figure 16(1)c).
(1) No male in controls

(2) 1 control daphnid produced only males (10 ± 32%, mean ± SD)

(3) 1 control daphnid produced males (a, b: 1 ± 3%, c: 5 ± 16%, mean ± SD)

(4) 2 control daphnids produced only males (20 ± 42%, mean ± SD)

Figure 16. The demonstration of statistical analysis
(5) 2 control daphnids produced 10% males (2 ± 4%, mean ± SD)

Figure 16. The demonstration of statistical analysis (continued)

Therefore, in the case of no males in controls, expert judgement is required to determine what % or more of mean male ratio is considered positive JH activity.

Next, we consider the case of male offspring induction in one replicate in the control. As a worst-case scenario, we suppose that one replicate produces only male offspring (100% male ratio), which sometimes occurs (mean ± SD: 10 ± 32%, n=10). Significant differences are found when 4 and more replicate produce males by Jonckheere-Terpstra test and 5 or more replicates produce male by Dunn test (Figure 16(2)). When the male ratio of 4 replicates in C3 is lower than 100% in the control replicate, the data does not meet monotonicity. Thus, Dunn test should be used rather than Jonckheere-Terpstra test, however Dunn test cannot detect significant (Figure 16(2)c). Significant difference is shown if >4 replicates produce males regardless of the male ratio in C3 (e.g. 10%×5 replicates, 5 ± 5%). In the case of the male ratio in the control replicate is <100%, significant difference is detected at C3 with 4 and more replicates produce male offspring (Figure 16(3)). Figure 16(3)bc also suggests that significant difference can be detected if the male ratio of 4 replicates in C3 is greater than or equal to that of the control replicate.

If two replicates produce only male offspring in the control (mean ± SD: 20 ± 42%, n=10), significant differences are found when 6 and more replicate produce males by Jonckheere-Terpstra test (Figure 15(4)). When the male ratio in 6 replicates in C3 is lower than 100% in the control, the data does not meet monotonicity and Dunn test could not detect significant (Figure 15(4)c). Significant difference is shown if >6 replicates produce males regardless of the male ratio in C3 (e.g. 10%×7 replicates, 7 ± 5%). In the case of the male ratio in the control replicates is <100%, significant difference is detected at C3 with 5 and more replicates produce male offspring as long as the male ratio in C3 is higher than that in control (Figure 16(5)).

Detection power of Jonckheere-Terpstra test and Dunn test are dependent not only on the male ratio but also the number of replicates with male offspring out of 10 replicates (incidence). To detect significant difference of male ratio in C3 that have 4/10 replicates with male induction, more than 1 replicate should not produce male offspring and the male ratio in the control should not exceed that in C3 replicates with male offspring. For example, to detect 20% male ratio (replicates mean) in C3 with 4 replicates producing males (i.e. 50% × 4 replicates, 20 ± 26%), the male ratio in the control replicate should not exceed 50% (i.e. mean ± SD: ≤ 5 ± 16%).

Based on the simulation analysis conducted in the validation study of an enhancement of OECD TG 211 Daphnia reproduction test (OECD, 2008), both Jonckheere-Terpstra test and Dunn test demonstrated 80% power to detect an increase of <10% males in the highest concentration (5 concentrations, 10 replicates). The rarity of male offspring in controls causes the statistical tests to be very sensitive to extremely small changes in percent males, making interpretation of the results problematic. The report also examined the quality of regression model and concluded that it is more appropriate to analyze male ratio, which has a 0% background incidence rate, using a suitable regression model to estimate ECx. It is a biologically based decision which value of x in ECx is most appropriate for this screening experiment.

Alternative methods can be used, such as a generalized linear model (GLM) or generalized linear mixed model (GLMM) with binomial errors for proportion of males, if justified statistically (Cameron and Trivedi, 2013).
However, because male ratio in the controls is basically zero in all replicates (zero variance), it would be an informal analysis. Step-down Cochran-Armitage test or Fisher’s exact test with Bonferroni-Holm adjustment by pooling replicates per treatment is not recommended because sex determination occurs per brood so that the proportion of male should be calculated per replicate.

5.6 Validity criteria

The draft protocol only suggests the criterion for the control mortality (paragraph 10): For a test to be valid, the following performance criteria should be met in the control(s): The mortality of the parent animals (female Daphnia) does not exceed 20% at the end of the test.

Based on the simulation of statistical analysis, we suggest following additional requirement as optional to ensure the statistical detection power on male ratio:

- The mean number of living offspring produced per alive parent animal in control(s) is ≥12.
- No more than 1 parent animal produce male offspring in the control(s).
- Mean male ratio preferably does not exceed 5% in the control(s).

The male ratio is calculated by dividing the number of male offspring by the number of the total offspring and enough number of offspring and the number of total offspring is not directly taken into account to the statistical analysis. However, considering the ratio per one male offspring, it is desirable to have a sufficient number of offspring. The mean number of living offspring produced per parent animal surviving at the observation in the ring test was 27.3±12.7 (min: 12.9, median: 25.3, max: 45.6) in control and 28.1±13.5 (min: 12.4, median: 26.0, max: 48.6) in solvent control. Moreover, in TG 211, approx. 4-5 broods are observed for 21 days and the validity criteria require ≥60 living offspring per parent, indicating that the mean number of offspring per brood is 12 at least. Therefore, we recommend ≥12 per alive daphnids as the criterion for the baseline reproduction performance. One male per 12 offspring is equal to 8% male ratio. In the ring test, the results from all seven laboratories met this value.

The discussion of biologically significant incidence of male induction and mean male ratio are not fully exhausted, but if we want to detect ≥40% incidence of male induction (4 out of 10 replicates), more than 1 parent animal should not produce male offspring. Moreover, to detect ≥20% mean male ratio in the exposure group, mean male ratio should not exceed 5% (i.e. ≤50% male ratio in one replicate in control, Figure 16(3)c) in the case of 40% incidence, but it is not required in the case of >40% incidence. This additional test validity criterion of no more than 5% males in the controls is also suggested in OECD TG 211 Annex 7 because the previous ring test for OECD TG 211 Annex 7 indicated that 5% of male offspring in control replicate is common and can represent the background noise (OECD, 2008). Note that the mean male ratio in the normal mass culture condition was 0% and 1.3% for NIES and Clone A strains, respectively (Table 9, 10).

Considering the concept for this screening assay, statistical significance is not always necessary to determine positive JH activity. It is up to regulatory authority to decide whether to introduce these additional criteria.
6. Conclusions and recommendations for TG

- Test strain is not specified; however, it is necessary to confirm that the strain rarely produces male offspring in normal culturing and testing conditions.
- Test duration is not fixed to 7 days but can terminate earlier or extend until the second brood is produced. For reference, the maximum test duration could be 10 days.
- Clarify the condition of the test adult females. To ensure the exposure during the sex determination period, test should be initiated more than 7-10 h before ovulation of the second brood egg to the brood chamber indicating that we should avoid using females that are ready to release the first brood offspring out of the brood chamber. To test a readily degradable chemical, care should be taken to keep the concentration during the sex determination period (e.g. frequent renewal, use females with offspring 40-50 h post ovulation).
- The additional validity criteria for male production are needed to ensure the statistical detection power on the male ratio. More than 1 parent animal does not produce male offspring and mean male ratio preferably does not exceed 5% in the control(s) in order to detect 20% male ratio derived from 40% of replicates with male production.
- Environmental conditions such as temperature, hardness, density, feeding, photoperiod do not generally induce male offspring production if they are within the test condition. Because the sensitivity to the environmental stimuli varies among strains, it is desirable to find appropriate conditions where male offspring is hardly induced in each strain to exclude the effect of non-chemical stress.
7. Literature


ANNEX 1 Preliminary validation study

Even in the same *D. magna* species, Oda et al. (2006, 2007) demonstrated different strains of *D. magna* had different sensitivity in the proportion of male offspring induced by fenoxycarb, JH receptor agonist. Some strains (e.g. *D. magna* LRV13.5–1 (Toyota et al., 2021)) constantly produce male offspring even in the laboratory culturing condition, whereas the NIES strain that is maintained at NIES more than 30 years rarely produces males in in the laboratory culturing condition (see Chapter 5). To understand the strains suitable for JHASA and compare the sensitivity difference of the strains, a pre-validation study was conducted using five strains of *D. magna* (Table S1) which were provided from different laboratories and have been maintained at NIES for more than ten years. NIES strain was originally provided from US Environmental Protection Agency in 1990s. Clone A provided from Chemicals Evaluation and Research Institute (CERI), Japan and Clone B provided from the United Kingdom are the same strain as “Clone A” identified in the ring test of the OECD TG211 *Daphnia magna* reproduction test (OECD, 1997). Clone C was provided from the laboratory in Finland when the validation study was performed for OECD TG 211 Annex 7 (OECD, 2008). Clone D was obtained from dormant eggs in Daphtoxkit F (MicroBioTests, Gent, Belgium) and have been culturing by asexual reproduction at NIES more than ten years.

As a test substance, JH receptor agonist, diofenolan was used. Firstly, acute immobilization test (OECD, 2004) was conducted with diofenolan. EC50 (50% effective concentration) values for diofenolan varied by a factor of 2.4 among strains (287–713 µg/L). Based on EC50 values for each strain, test concentrations in JHASA were determined as shown in Table S1.

Diofenolan exposure induced the production of male neonates in all the strains in all concentrations where offspring was observed. In Clone A and Clone B, originally the genetic identical strains European “Clone A”, male offspring was also produced in control and solvent control at 6.5±21% (pooled controls) and 14±33% (solvent control, respectively (pooled controls). This may be due to the high density in the culturing condition of the daphnids used in the test. In particular, the mean total number of offspring per daphnids in control and solvent control were only 7.6 and 11.2, respectively in Clone B, it is possible that the culture conditions prior to the test were not good. For Clone C, no offspring was observed at 175 µg/L and higher concentration in all concentrations where offspring was observed. For Clone D, no offspring was observed at 175 µg/L and higher concentration in the first experiment although acute EC50 was 713 µg/L. Acute immobilization test was used as range finding to avoid parent mortality, but it is note that the effect on reproduction decrease could not be estimated from the acute toxicity data.

### Table S1. The result of preliminary validation study using diofenolan

<table>
<thead>
<tr>
<th>Strain</th>
<th>Source</th>
<th>48 h EC50 (µg/L)</th>
<th>Test concentrations (µg/L)</th>
<th>Male ratio in controls (incidence)</th>
<th>LOEC for male induction (µg/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIES</td>
<td>USEPA</td>
<td>685</td>
<td>87.5, 175, 350, 700</td>
<td>0%</td>
<td>87.5</td>
</tr>
<tr>
<td>Clone A</td>
<td>CERI (Clone A from Europe)</td>
<td>427</td>
<td>12.5, 25, 50, 100</td>
<td>6.5±21% (2/10)</td>
<td>12.5</td>
</tr>
<tr>
<td>Clone B</td>
<td>United Kingdom (Clone A)</td>
<td>287</td>
<td>25, 50, 100, 200</td>
<td>14±33% (2/10)</td>
<td>25</td>
</tr>
<tr>
<td>Clone C</td>
<td>Finland</td>
<td>713</td>
<td>1st: 87.5, 175, 350, 700</td>
<td>0%</td>
<td>1st: 87.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2nd: 12.5, 25, 50, 100</td>
<td></td>
<td>2nd: 12.5</td>
</tr>
<tr>
<td>Clone D</td>
<td>Daphtoxkit (MicroBioTests)</td>
<td>308</td>
<td>12.5, 25, 50, 100</td>
<td>0%</td>
<td>12.5</td>
</tr>
</tbody>
</table>

a: No. of replicates with male offspring out of 10 replicates.