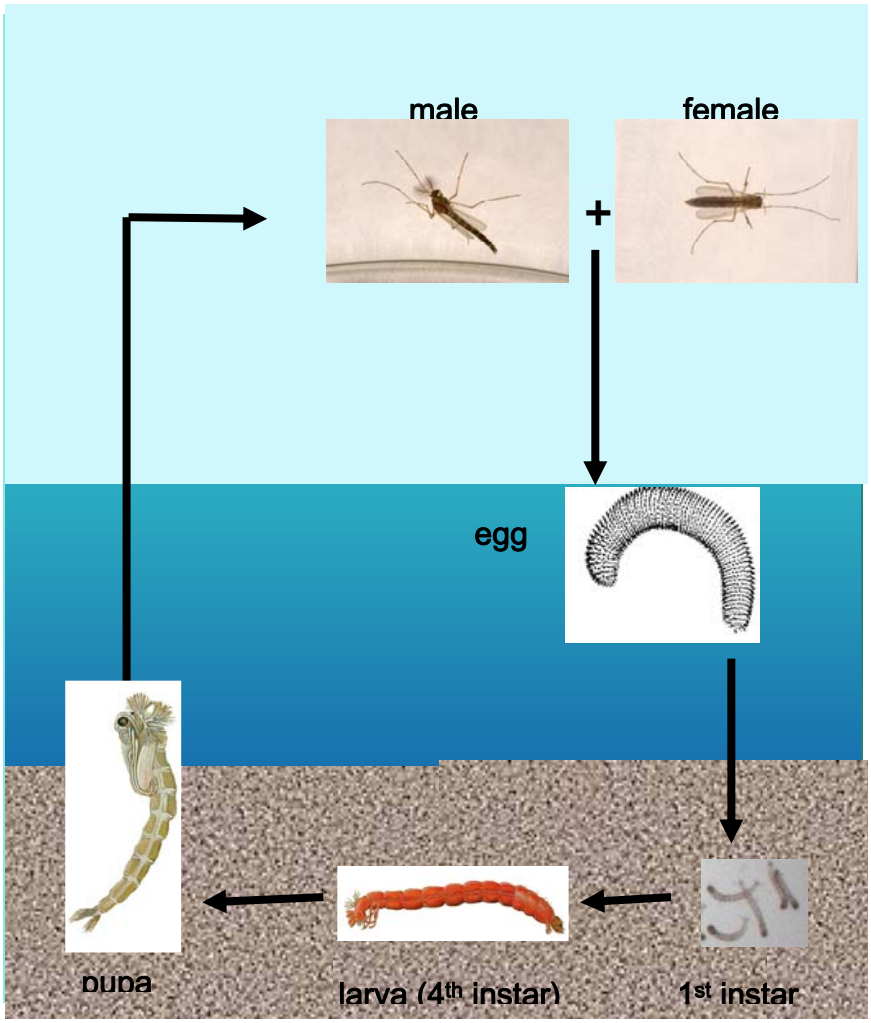
The chironomid full life-cycle test
Validation report

First draft

Lennart Weltje[1] and Eric Bruns[2]

[1] BASF SE, Crop Protection - Ecotoxicology, Limburgerhof, Germany
[2] Bayer CropScience AG, BCS AG-D-ETX, Monheim, Germany



November 2009

Preface

This validation report contains the results of various experiments that were conducted in support of the chironomid life cycle test, which is currently under development at the OECD. The supporting member state is Germany.

Some of the experiments following the draft guideline have been published in peer reviewed journals (Taenzler et al., 2007; Tassou and Schulz, 2009) and/or were presented at scientific meetings (Pupp and Weltje, 2006; Weltje et al., 2004a, 2004b, 2004c, 2006a, 2006b, 2009). Progress in the development of the guideline and interim experimental results were presented by BIAC at the OECD ad hoc invertebrate expert group meetings held in 2005 (Paris, France), 2007 (Columbia, SC, USA) and in 2009 (Paris, France).

At the 2009 meeting in Paris it was decided by the expert group that sufficient experimental data were available, so that the final ring test would only have to consist of a control group, testing the compliance with the proposed validity criteria.
Part of these ringtest results are now available and reported here. It is expected that by the end of 2009 a significantly higher number of test results is available and as a result this report will be updated accordingly.

Acknowledgements

In the past years many people contributed to the development of this guideline, by providing comments, reviewing papers, discussing and last but not least conducting studies.

Particularly we would like to mention the people in the laboratories of Syngenta, BASF, Bayer CropScience and the University of Landau, including the students from the Universities of Landau and Frankfurt who either did a practicum or their MSc thesis work on various aspects of chironomid life-cycle testing. The initial proposal for this guideline was formulated at the Industrieverband Agrar (IVA) and principal discussions took place in the OECD Invertebrate Expert Group.

In no particular order we would like to mention: Verena Taenzler, Marcus Ebeling, Michael Dorgerloh, Gabriele Silke, Dominik Reinhard, Verena Pfeifle, Kate Benyon, James Wheeler, Sabine Zok, Peter Dohmen, Daniela Belz, Annika Pupp, Philipp Egeler, Rebecca Pierstorf, Tania Jarosz, Daniela Schroth, Ribana Seliger, Lucas Jagodzinski, Christian Vogt, Matthias Oetken, Jörg Oehlmann, Candida Shinn, Torsten Hahn, Koffi Tassou and Ralf Schulz for their contributions.

Also we would like to thank Petra Greiner and Jörn Wogram of the UBA, Germany for acting as the supporting member state and finally Anne Gourmelon of OECD for her continuous support and coordination.

Lennart Weltje and Eric Bruns

Table of Contents

1. Introduction

A higher tier full life cycle test using chironomids (Insecta, Diptera, Chironomidae) has been developed. The increasing interest in potential adverse effects due to endocrine modes of action made the development of an appropriate invertebrate assay a primary goal of national and international testing programs, such as the one under development at the OECD. Thereby it should be considered that the various invertebrate assays (including the presented chironomid test) are not mechanistic in nature, i.e. they do not reveal if chemicals act on the endocrine system. Instead the assays are designed to study effects of chemicals on reproduction and development by covering the complete life cycle of the organisms and thus would pick up any adverse effect.

The presented chironomid full life cycle test completes the existing acute and chronic test cascade for these widely used test organisms. The acute and the chronic chironomid tests are part of the standard test procedures for registration of agrochemicals under the European plant protection products directive 91/414/EEC and also for biocides. Therefore, the chironomid life cycle test yields results, which directly fit into the general risk assessment. The life cycle test is an extension of the existing OECD guidelines 218 and 219 for the testing of chemicals ”Sediment water chironomid toxicity test using spiked sediment“ (OECD, 2004a) and ”Sediment water chironomid toxicity test using spiked water“ (OECD, 2004b), respectively. The development of the presented method also took the development and validation of the long term toxicity test with chironomids (edited by Streloke and Köpp, 1995) and the existing US-EPA guidelines (OPPTS 850.1790, OPPTS 850.1755, EPA 600R-991064) into consideration.

Chironomids have been chosen for the development of a freshwater invertebrate life cycle test due to numerous reasons. As freshwater insects, chironomids belong to one of the most abundant, species-rich and ecologically relevant taxonomic groups.
As the endocrine system of insects is the best described among the invertebrates, potential endocrine related effects can be investigated in this sexually reproducing organism. Specific third generation insecticides, juvenile hormone (ant)agonists and ecdysone (ant)agonists, have been developed to target the insect endocrine system and can be used as reference chemicals for potential endocrine effects. As holometabolic insects, chironomids undergo a full metamorphosis during a chronic test and after emergence the separate sexes are easily identified. The life cycle of *Chironomus riparius* at 20°C is completed in about 25 days. This time period is long enough to represent a chronic exposure in the environment and at the same time feasible from a laboratory perspective.

Chemicals can enter the aquatic environment by different routes, such as run-off, drainage and spray drift. The bioavailability of a chemical in the aquatic compartment is related to its physico-chemical properties. The exposure of sediment dwelling organisms can occur via the overlying water, the pore water and the sediment. Different exposure scenarios can be investigated by the use of chronic chironomid tests. The inserted chironomid larvae are exposed to the overlying water, the pore water and the sediment. The exposure situation to be investigated can therefore be

adapted according to the exposure scenario of interest. Both, spiked water and spiked sediment tests can be performed using the chironomid life cycle protocol.

The results observed in the chironomid life cycle tests can be compared to historical data obtained in chironomid studies conducted according to the OECD guidelines 218 and 219. Chironomids are standard test organisms which are easy to handle and to culture. The knowledge which is needed to perform life cycle tests on chironomids is already existing in a vast number of academic, industry and contracting laboratories, that perform ecotoxicological studies.

In comparison with the other aquatic arthropod species for which there are currently life-cycle tests under discussion, initiated or being validated at the OECD, the chironomid test has some distinct features:

- Chironomids are insects, while the other species (daphnids, mysids, copepods) belong to the crustaceans. Crustaceans are already represented in the aquatic testing suite by *Daphnia magna* with a reproduction test (OECD TG 211), while aquatic insects are not, although the insects are representing a higher number of species (in freshwaters).
- Chironomids, like daphnids, are freshwater species, while copepods and mysids are saltwater species (some requiring filtered seawater as test and/or culture medium). For most chemicals, freshwaters are more important than marine waters in terms of exposure concentrations.
- Chironomids reproduce sexually like mysids and copepods, but unlike daphnids, which reproduce parthenogenetically (unless a stress factor induces male production and thereafter sexual reproduction).
- Exposure of chironomids via both water and sediment is relevant and possible, like in the copepod test, but this is not a relevant route in the daphnid and in the mysid test.
- Chironomids can easily be fed with TetraPhyll/TetraMin, but daphnids and copepods feed on algae, which need to be cultured additionally. Mysids feed on neonates of the crustacean *Artemia salina*, which also require to be cultured in parallel.
- Whilst chironomids and daphnids are established test species in many laboratories all over the world, mysids are only established in a few laboratories in the US, which are based on the coast. Copepods present a relatively new test system, which is not established yet for regulatory testing.
- Finally, chironomids and daphnids have a reasonable size to work with and do not require specialized optical equipment, unlike copepods, whilst mysid testing requires a flow-through testing system.

The above given information as well as the comparative data presented in the following chapters are aimed to demonstrate the relevance, usefulness and feasibility of the chironomid life cycle test as an additional higher-tier tool for the environmental risk assessment.

2. Materials and methods

The experimental work presented in the next chapter was conducted according to a draft guideline, which is based on the existing test guidelines 218 Sediment-Water Chironomid Toxicity Test Using Spiked Sediment (OECD, 2004a) and 219 Sediment-Water Chironomid Toxicity Test Using Spiked Water (OECD, 2004b). The validation data and background of the 218 and 219 guidelines is described in a BBA booklet edited by Streloke and Köpp (1995).

The life-cycle test is an extension of the OECD TG 218/219 and unlike these does not end with the emergence of the adult midges, but instead collects the adults (P generation) carefully and puts them into a breeding cage, so that they can swarm, mate and oviposit in crystallising dishes containing the same spiked water-sediment sytem they emerged from. Apart from quantitatively following reproduction, fertile egg ropes are selected at the peak of oviposition and allowed to hatch. After hatching, 1st instar larvae (F1 generation) are collected and added to spiked water-sediment vessels again and followed until emergence.

So, in essence the proposed life cycle test covers two generations (P and F1) and exists of a series of two OECD TG 218 or 219 tests. The additional information provided by conducting a life-cycle test (in comparison with the current chronic study) is therefore:

- Addition of a second (filial) generation endpoints (to account for possible higher sensitivity, due to accumulation, carry over effects etc.
- Reproduction assessments, both fecundity and fertility.

Since the first part of the test was already validated in 1995, the ringtest should focus on the additional endpoints as outlined above.

Full experimental details for study conduct can be obtained from the draft guideline (current revised version, including spiked sediment scenario of November 15, 2009).

3. Experimental work

3.1. overview

After the writing of the draft guideline, various experiments have been conducted, all using the species *Chironomus riparius* MEIGEN. The experiments all followed a NOEC design, since there is no agreement on the effect level to be assessed with the EC_X approach.
The first small scale ringtest was a test in spiked-water design using the general toxicant and reference compound 3,5-dichlorophenol (3,5-DCP). Thereafter a spiked water test was conducted with the juvenile hormone analogue pyriproxifen (PPF) in anticipation of specific effects on chironomids and also because this compound was selected for validation of other arthropod reproduction tests (daphnids, copepods, mysids) under development at the OECD (see Gourmelon and Ahtiainen, 2007). The 3,5-DCP and PPF experiments were conducted as small scale ringtests by Syngenta (UK), Bayer CropScience (Germany) and BASF (Germany) and published by Taenzler et al. (2007). Later, PPF was tested following the same design at the University of Landau (Germany) and published by Tassou and Schulz (2009).

Since it was commented by the UBA that the guideline should also cover other designs, such as a spiked sediment scenario, a small scale ringtest with the insecticide lindane (γ-HCH) was conducted by Bayer CropScience, BASF and the University of Landau (on behalf of Syngenta). Lindane was selected since it had been used in the original validation studies for development of the OECD TG 219 (see Streloke and Köpp, 1995). The used concentrations were based on a range finding test according to the OECD TG 218 conducted at BASF.

As a side project, the effect of the sex ratio of adult chironomids in the swarm on egg rope fertility was tested at the BASF laboratory using a slightly different design than described in the draft guideline. Results were presented at a scientific conference by Pupp and Weltje (2006). The results will only be briefly discussed.

3.2. Spiked water test with 3,5-dichlorophenol

The first test carried out according to the draft protocol was a spiked water test using 3,5-dichlorophenol (3,5-DCP), which is a general toxicant that is recommended for reference testing in OECD guidelines such as OECD TG 201 – algae and OECD TG 221 - *Lemna*. The small scale ringtest ($n = 3$) with 3,5-DCP was conducted in the laboratories of Syngenta, Bayer CropScience and BASF. The results were published in Taenzler et al. (2007).
In short, the selected concentrations (0.0313, 0.125, 0.5 and 2.0 mg/L) were too low to induce a significant inhibition of emergence, development or reproduction. Hence, all three laboratories concluded that the NOEC for all endpoints was ≥ 2.0 mg/L. Although no real effect threshold was achieved with the used concentrations, the results were consistent among the participating laboratories. In addition, all experiments were compliant with the validity criteria (possibly one endpoint invalid - see footnote of Table 1). This was judged by the performance of the solvent control containing dimethylformamide (DMF), since not all laboratories employed a water control. A comparison of the results of the solvent controls are presented below (adapted from Taenzler et al., 2007).

Table 1. Mean endpoint values of the solvent control in the 3,5-DCP experiment

Endpoint	**Lab 1**	**Lab 2**	**Lab 3**
P emergence ratio (both sexes summed) *n* = 8	0.91	0.94	0.81
P development rate (both sexes summed) *n* = 8	0.0667	0.0652	0.0631
P male fraction	0.50	0.58	0.51
Fecundity (egg ropes/female) *n* = 2	0.82	0.79	0.82
Fraction fertile egg ropes *n* = 2	0.97	0.73	0.89
F1 emergence ratio (both sexes summed) *n* = 8	0.88	0.88	0.83
F1 development rate (both sexes summed) *n* = 8	0.0694	0.0592	0.0652
F1 male fraction	0.50	0.52	0.39*

* not valid, when male fraction is interpreted as 0.40 (guideline says 0.4)

3.3. Spiked water test with pyriproxifen

The second test that was carried out according to the draft protocol was again a spiked water test using the insecticide pyriproxifen (PPF). PPF is a juvenile hormone analogue, which is a specific arthropod toxicant that interferes with the endocrine regulation of the larval phase and the transition into metamorphosis (the pupal phase). PPF has been selected as reference compound for validating reproduction tests with aquatic arthropods under OECD flag (see Gourmelon and Ahtiainen, 2007). This compound can be regarded as a positive control. The small scale ringtest (n = 3) with PPF was conducted in the laboratories of Syngenta, Bayer CropScience and BASF. The first run of the test at Syngenta resulted in a P emergence slightly below 70%, thus invalidating the test. The second run yielded a valid P emergence, but now the F1 emergence was slightly below 70%, thus invalidating the test. However, considering the consistent trend in the data and the valid assessments of reproduction the second test was used for comparison. The selected concentrations were 0.16, 0.8, 4, 20, 50/100 µg/L (spacing factor 5), which induced the full range of effects. The results were published in Taenzler et al. (2007).
Afterwards an identical test was conducted at the University of Landau, using slightly different concentrations (1, 3, 10, 30, 100 µg/L, spacing factor 3) so that in total four experiments are available for comparison (n = 4). The data from the University of Landau were published in Tassou and Schulz (2009). The results are presented in Table 2.

Table 2. Endpoint values (NOECs in µg/L) of the PPF experiment

Endpoint	Tassou and Schulz (2009)	Lab 1	Lab 2	Lab 3
P emergence ratio (both sexes summed)	10	4	4	4
P development rate (both sexes summed)	10	4	4	4
P male fraction	≥ 30	≥ 50	4[1]	≥ 20
Fecundity (egg ropes/female)	10	20	4	≥ 20
Fraction fertile egg ropes	10	20	4	≥ 20
F1 emergence ratio (both sexes summed)	3	0.8	≥ 4	0.8
F1 development rate (both sexes summed)	3	0.8	≥ 4	≥ 20[2]
F1 male fraction	≥ 10	4	≥ 4	≥ 20

[1] NOEC seems rather low in comparison to the other tests
[2] NOEC seems rather high in comparison to the other tests

The test concentrations induced the full range of effects, with for the highest treatment a massive inhibition of P generation emergence, so that reproduction assessments for this treatment were not possible anymore.
In comparing the endpoints, a consistent pattern is seen. With the exception of two single cases (highlighted in the footnotes of Table 2), the calculated NOEC values are not more than one treatment level apart. For the evaluation of the appropriateness of the endpoints it was extremely helpful to have an additional data set that was generated outside of the framework of the small scale ringtest and which employed different concentrations. From the work it can be concluded that the second (F1) generation is more sensitive than the first (P) generation, considering emergence and development. Further, there was a trend (only a few significances) in the sex ratio of both generations towards a lower fraction of males. Finally, reproduction was a less sensitive endpoint than emergence or development.

At the same time, chironomid life cycle experiments were conducted at the University of Aachen (Germany) with the insecticide fenoxycarb (Shinn et al., 2007), which also belongs to the class of juvenile hormone analogues. A comparison with the PPF data could be of interest to study if certain endpoints react sensitively to compounds with this specific mode of action.
In summary, Shinn et al. (2007) concluded for fenoxycarb that addition of a second generation to the 28-day chronic test increases its sensitivity, as F1 generation NOECs were one level below the P generation NOECs for emergence and development. The same trend that was also observed for PPF. A lower male fraction due to fenoxycarb exposure is consistent with the findings on PPF as described by Taenzler et al. (2007). Finally, emergence was more sensitive than reproduction, which is also confirmed for PPF. Taken all the information together it appears that both juvenile hormone analogues display similar patterns in the effects they induce in the two exposed chironomid generations.

3.4. Spiked-sediment test with lindane

Since the guideline should also enable the conduct of other designs, such as a spiked sediment scenario, a small scale ringtest (n = 3) with the insecticide lindane (γ-HCH) was conducted by Bayer CropScience, BASF and the University of Landau (on behalf of Syngenta). Lindane, a GABA-gated chloride channel antagonist (i.e. neurotoxic to insects) was selected, because it had been used in the original validation studies for development of the chronic chironomid test guideline (see Streloke and Köpp, 1995) and is since then used as a reference compound for chironomid studies. The used concentrations (0.03125 to 0.5 mg/kg, spacing factor of 2) were based on a range finding test according to the OECD TG 218 conducted at BASF.
The results are presented in Table 3.

Table 3. Endpoint values (NOECs in mg/kg dry sediment) of the lindane experiment

Endpoint	**Lab 3**	**Lab 1**	**Lab 2**
P emergence ratio (both sexes summed)	0.15	0.075	0.125
P development rate (both sexes summed)	0.075	0.0375	0.03125
P male fraction	0.15	0.075	0.125
Fecundity (egg ropes/female)	≥ 0.3	≥ 0.15	≥ 0.25
Fraction fertile egg ropes	0.0375[1]	≥ 0.15	≥ 0.25
F1 emergence ratio (both sexes summed)	0.15	≥ 0.15	0.0625[1]
F1 development rate (both sexes summed)	0.075	0.075	≥ 0.25[2]
F1 male fraction	≥ 0.15	≥ 0.15	≥ 0.25

[1] NOEC seems rather low in comparison to the other tests
[2] NOEC seems rather high in comparison to the other tests

The test concentrations induced the full range of effects, with for the highest treatment a massive inhibition of P generation emergence, so that reproduction assessments for this treatment were not possible anymore.

In comparing the endpoints, again a consistent pattern is seen. With the exception of three single cases (highlighted in the footnotes of Table 3), the calculated NOEC values are not more than one treatment level apart.
From the experiments it can be concluded that the second (F1) generation is not more sensitive than the first (P) generation, considering emergence ratio and development rate. Further, there was a trend (only a few significances) in the sex ratio of the P generation towards a lower fraction of males. Finally, reproduction was a less sensitive endpoint than development. This pattern is different from that obtained with PPF or fenoxycarb.

3.5. Swarm sex ratio

The work on the effect of sex ratio shifts in the P generation on egg rope fertility was conducted to better understand and interpret the results of the PPF experiments, where effects had been observed on the sex ratio (fewer males) and on egg rope fertility in the higher concentrations. Hence, a lower fertility may be caused by direct toxicant effects or by indirect effects (shift in sex ratio).

Normally the sex ratio in *C. riparius* is evenly distributed. i.e. 1:1 (50%:50%) with rather low deviations, provided the number of animals is sufficient.

For these experiments breeding cages were filled with swarms in which the sex ratios were manipulated to 3:1, 2:1, 1:1, 1:2 and 1:3. In addition, swarms were limited to maximally 100 animals (both sexes summed) and the smallest swarm consisted of 36 animals (both sexes summed). With the conducted experiments it was also possible to look at the effect of swarm size on fertility, starting from the hypothesis that in a larger swarm there is a better chance of meeting a partner and thus achieve a higher fertility. There were no substances used for these experiments and the sediment consisted of pure quartz sand only.

The results showed clearly and significantly that when the fraction of females in the swarm increases, that the fertility of the egg ropes decreases (see Figure 1 below). Also, the larger swarm size correlates significantly with a higher fertility of the egg ropes (see Figure 2).
In consideration of the results, swarm size should be at least 50 animals (both sexes summed) with a female percentage smaller than 60% to gain a proportion of more then 80% fertile egg ropes.

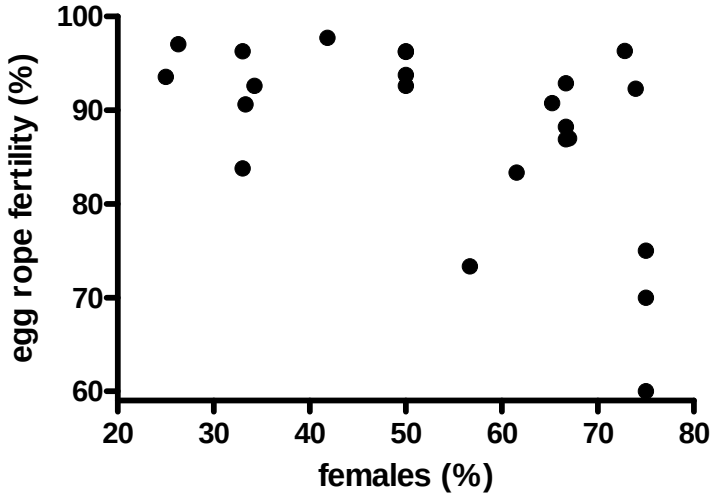


Figure 1. Egg rope fertility against percentage of females in the swarm (Pearson correlation, *r* = -0.533, *P* = 0.0088)

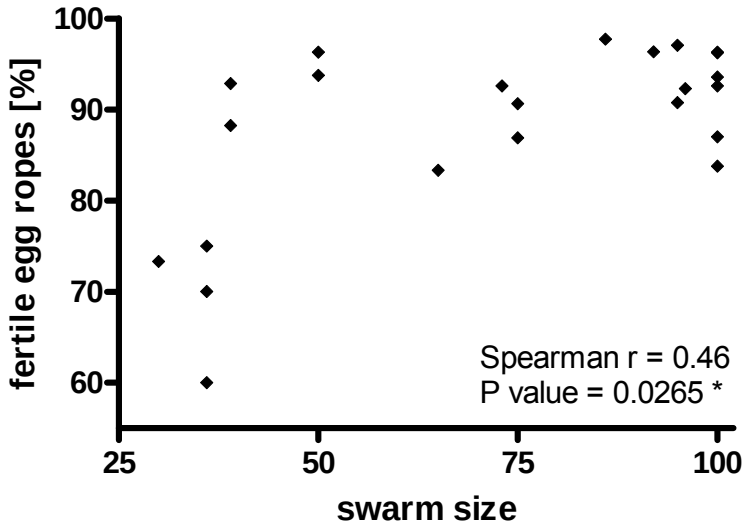


Figure 2. Egg rope fertility against swarm size (Pearson correlation, *r* = 0.46, *P* = 0.0265)

4. Ring test on controls

4.1. General

Around 10 laboratories are involved as participants in the ringtest. Currently there are three data sets available for evaluation; more results are under way and will be considered in the next draft of this report.

The following considers only the three available data sets. The physico-chemical data are not presented in detail here, but it can be stated that the dissolved O_2, pH and temperature were well under control and within the limits and/or validity criteria set in the draft guideline. There was only one issue with the water temperature in Lab 3 at the start of the test (P generation). The value was too low (minimum 18.3°C, maximum 18.9°C) but it apparently did not affect the development of the animals during the test (average for P generation 19.33°C, standard deviation 0.62°C). In fact, the development rate for the P generation in Lab 3 was the highest. Therefore, this measurement at the start of the test is not considered to invalidate the test.

4.2. Emergence and development

The emergence ratio and development rate for the P and F1 generation are presented in Table 4 below. They show that the validity criteria for emergence were satisfied.

Table 4. Detailed data on development rate and emergence ratio of both generations as obtained in the ring test (for all observations *n* = 8)

P gen	Lab 1		Lab 2		Lab 3	
	Development rate		Development rate		Development rate	
cage	mean	st dev	mean	st dev	mean	st dev
A	0,0600	0,0017	0,0612	0,0009	0,0631	0,0017
B	0,0596	0,0007	0,0601	0,0011	0,0633	0,0013
C	0,0619	0,0011	0,0601	0,0024	0,0645	0,0015
D	0,0631	0,0021	0,0614	0,0007	0,0650	0,0022
	Emergence ratio		Emergence ratio		Emergence ratio	
cage	mean	st dev	mean	st dev	mean	st dev
A	0,8875	0,0750	0,8250	0,0866	0,7375	0,2016
B	0,9500	0,0707	0,8625	0,1377	0,8750	0,0645
C	0,9250	0,0500	0,9375	0,0750	0,8125	0,1493
D	0,9250	0,0957	0,9250	0,0645	0,8500	0,0707

F1 gen	Development rate		Development rate		Development rate	
cage	mean	st dev	mean	st dev	mean	st dev
A	0,0605	0,0045	0,0640	0,0026	0,0604	0,0013
B	0,0612	0,0030	0,0622	0,0067	0,0588	0,0038
C	0,0603	0,0016	0,0648	0,0023	0,0619	0,0033
D	0,0627	0,0021	0,0663	0,0049	0,0594	0,0033
	Emergence ratio		Emergence ratio		Emergence ratio	
cage	mean	st dev	mean	st dev	mean	st dev
A	0,9250	0,0500	1,0000	0,0000	0,8125	0,0479
B	0,8625	0,0250	0,8250	0,3500	0,8125	0,0750
C	0,8750	0,1041	1,0000	0,0000	0,7500	0,1633
D	0,8875	0,0750	0,9250	0,0957	0,9875	0,0250

There is no fixed number for judging the validity the development rate itself, but instead the guideline says that emergence in the control(s) should occur between day 12 and 23 after inserting the larvae into the test vessels. With the exception of very few individuals (see below) this was the case.

Lab 1. P gen: 3 on day 23; F1 gen: 1 on day 23, 8 on day 25.
Lab 2. none.
Lab 3. P gen: 2 on day 23, 1 on day 24; F1 gen: 5 on day 23

4.3. Sex ratio

The sex ratio in the P generation and in the F1 generation are presented in Table 5 together with the reproduction data (i.e. egg rope production and fertility). As expected, no large deviations from the natural sex ratio were observed. The only deviation from the validity criteria set by the draft guideline (minimum fraction of males or females 0.4) was observed for Lab 1 in cage B, where 38% males ocurred. This would constitute a breaching of the validity criterion for sex ratio, when male fraction is interpreted as 0.40 (guideline says 0.4 – a rounding issue). However, this did not lead to a breaching of the reproduction validity criteria (see section 4.4)

4.4. Reproduction

According to the draft guideline, the test is valid when at least 0.6 egg ropes per female are produced (in each breeding cage) and in addition when 60% of the produced egg ropes are fertile. So, taking these two criteria together would mean that (0.6 x 0.6 =) 0.36 fertile egg ropes per female should be produced.
The number of egg ropes per female and the number of fertile egg ropes per female for each of the laboratories are presented in Table 5 below.

Table 5. Detailed data on reproduction and sex ratio of the P generation as obtained in the ringtest

cage	Egg ropes/female	Fertile egg ropes/female	Fraction fertile	Male fraction
Lab 1				
A	1.26	0.89	0.71	0.50
B	0.83	0.53	0.64	0.38*
C	1.00	0.82	0.82	0.47
D	0.87	0.77	0.89	0.48
Lab 2				
A	1.09	1.06	0.97	0.47
B	1.30	1.24	0.95	0.46
C	1.12	1.12	1.00	0.55
D	1.11	1.11	1.00	0.49
Lab 3				
A	0.77	0.73	0.95	0.56
B	0.34[1]	0.18	0.53[1]	0.45
C	0.55[1]	0.52	0.95	0.49
D	0.85	0.73	0.86	0.51

* not valid, when male fraction is interpreted as 0.40 (guideline says 0.4 – rounding issue)
[1] not valid

Obviously, Lab 3 had problems in achieving a sufficient number of egg ropes in cages B and C. In the group that produced the lowest number of egg ropes, the lowest fertility percentage ocurred. A possible explanation for this low performance in Lab 3 is the deviation from the draft guideline in the size of the breeding cage: 15 x 15 x 22 cm, instead of the recommended 30 x 30 x 30 cm. Possibly the cage was too small for swarming. Further, the data in Table 5 show that the lowest fertile egg mass fractions coincide with the lowest male fractions.

5. Conclusions

5.1. General

The motivation for the development of a chironomid full life-cycle toxicity test has been detailed in the *Introduction* section and will not be repeated here. More important are the conclusions that may be drawn from the supporting experimental work that has been conducted over the last years. Firstly, it was demonstrated that both spiked water and spiked sediment designs can be conducted succesfully.

5.2. Sex ratio and swarm size

Further, the sex ratio in the swarm may have an influence on the egg rope fertility (when fewer males than females are present) and thus to set a validity criterion on the male fraction is justified. Accordingly, when in a test effects on fertility are observed an inspection of the sex ratio may assist in determining if these effects are caused by the lower fraction of males (indirect effect) or by direct compound toxicity. This aspect is highlighted, since males tend to be more sensitive than females (as observed in the lindane, PPF and fenoxycarb experiments) and thus fertility may be impacted by this. The higher sensitivity of males may be related to their accelerated development rate as compared to females (i.e. the characteristic protandry, causing males to emerge earlier than females). The natural phenomenon of protandry also causes a gradual shift in the sex ratio in the breeding cage from 100% males at the beginning of emergence to 100% females at the end (when all males have died and the last emerged females are still alive). This natural shift may be the cause for the observation of infertile egg ropes, which mostly ocurr at the end of the reproduction phase.
When laboratories had problems with the validity criterion that was proposed for the sex ratio (i.e. male fraction) it was always the case that fewer males were present, never too many males (i.e. more than 60% males never occurred). Therefore, it may be considered to adjust this validity criterion.

Additionally it was shown that the swarm size has an influence on the egg rope fertility (in a larger swarm more fertile egg ropes are produced). This aspect is covered by setting a validity criterion on the emergence rate of the P generation, ensuring a swarm of at least (4 x 20 x 0.7 =) 56 individuals in one cage.

5.3. Effect patterns and sensitivity

The pattern of effects induced by juvenile hormone mimics (PPF and fenoxycarb) was different from that induced by the insecticide lindane. As further data become available in the future it would be interesting to study if specific effect pattern correlate with specific (insecticidal) modes of action.

In terms of sensitivity, the addition of a 2nd generation made the test more sensitive than the OECD test 218 or 219 as demonstrated with the test on juvenile hormone mimics. Interestingly, reproduction was mostly not the most sensitive endpoint, but

equally sensitive or less sensitive than emergence and/or development (see PPF and lindane experiments). Due to statistical issues with regard to the reduced number of replicates, an EC_X approach may be more useful for the reproduction assessments, but there needs to be an agreement on the level of X (the size of the effect) first to facilitate a comparison with obtained NOEC values.
For emergence and development, it should also be considered that this test is potentially more sensitive than the classic chronic OECD chironomid test, because it employs eight replicates per treatment instead of the usual six. Thus, its discriminating power to pick up smaller effects is increased.
Further, since the life-cycle test is designed to assess in total eight endpoints there is the increased possibility of statistical significant effects, which may be merely due to chance (i.e. type II error). Therefore, the final outcome of the test should be based on an integrated interpretation of the collected endpoints, instead of focussing on a single value.

References

Gourmelon, A. and J. Ahtiainen (2007) Developing test guidelines on invertebrate development and reproduction for the assessment of chemicals, including potential endocrine active substances – the OECD perspective. Ecotoxicology, 16(1): 161-167.

OECD (2004a) Test guideline 218 Sediment-Water Chironomid Toxicity Test Using Spiked Sediment.

OECD (2004b) Test guideline 219 Sediment-Water Chironomid Toxicity Test Using Spiked Water.

Pupp, A and L. Weltje (2006) Effect of adult sex ratio and swarm size on reproduction of the non-biting midge *Chironomus riparius*. SETAC-GLB, Landau, Germany.

Shinn, C., Weltje, L., Maes, H. and T. Ratte (2007) Impact of the juvenile hormone analogue fenoxycarb on the life-cycle of the non-biting midge *Chironomus riparius*. Micropol & Ecohazard 2007, Frankfurt am Main, Germany.

Streloke M, Köpp H, editors. 1995. Long-term toxicity test with *Chironomus riparius*: Development and validation of a new test system. Mitteilungen aus der Biologischen Bundesanstalt für Land- und Forstwirtschaft, No. 315, BBA. Berlin, Germany.

Taenzler, V., Bruns, E., Dorgerloh, M., Pfeifle, V. and Weltje, L. (2007) Chironomids: suitable test organisms for risk assessment investigations on the potential endocrine disrupting properties of pesticides. Ecotoxicology, 16(1): 221-230.

Tassou K, Schulz R. 2009. Effects of the insect growth regulator pyriproxyfen in a two generation test with *Chironomus riparius*. Ecotoxicol Environ Safety 72: 1058–1062.

Weltje, L., Hahn, T., Vogt, C. and R. Schulz (2004a) Extending toxicity testing guidelines for chironomids to assess endocrine disruption in invertebrates. In: Ecological Relevance of Chemically-Induced Endocrine Disruption in Wildlife. CREDO-Cluster Workshop, Exeter, England, p. 21-22.

Weltje, L., Hahn, T., Vogt, C. and R. Schulz (2004b) Development of a guideline for testing endocrine disruption in chironomids. In: Ecological Relevance of Chemically-Induced Endocrine Disruption in Wildlife. CREDO-Cluster Workshop, Exeter, England, p. 30-31.

Weltje, L., Hahn, T., Vogt, C. and R. Schulz (2004c) Extending toxicity testing guidelines for chironomids to assess endocrine disruption in invertebrates. In: Crossing Borders - Molecular Mechanisms and Ecosystem Processes. SETAC-GLB, Aachen, Germany, p. 45.

Weltje, L., Dorgerloh, M., Pfeifle, V., Belz, D., Tänzler, V. and D. Reinhard (2006a) Development of a life-cycle test with the non-biting midge *Chironomus riparius*. In: Abstracts of the 16th Annual Meeting of SETAC-Europe. The Hague, The Netherlands.

Weltje, L., Dorgerloh, M., Pfeifle, V., Belz, D., Tänzler, V. and D. Reinhard (2006b) Development of a life-cycle test with the non-biting midge *Chironomus riparius*. In: Abstract book SETAC-GLB. Landau, Germany.

Weltje, L., Schroth, D. and R. Pierstorf (2009) Chronic toxicity of lindane in a full life-cycle test with non-biting midges (*Chironomus riparius* Meigen). In: Abstracts of the 19th Annual Meeting of SETAC-Europe. Göteborg, Sweden.